Microsoft Purview

Microsoft Purview is a comprehensive portfolio of products spanning data governance, data security, and data compliance solutions.



OVERVIEW
Learn about Microsoft Purview



GET STARTED
Set up data governance solutions



GET STARTED
Set up data security solutions



GET STARTED
Set up data compliance solutions



QUICKSTART
Try the free version of data governance solutions



QUICKSTART
Try the data security and data compliance solutions trial for free



WHAT'S NEW
Catch up on new functionality and documentation updates



REFERENCE
Build data security and governance into AI apps with the Microsoft Purview SDK

Explore Microsoft Purview solutions

Guidance to help you explore and get started with Microsoft Purview solutions in the Microsoft Purview portal.







Microsoft Purview data governance solutions

- Unified Catalog
- Data Map
- Data governance planning
- Data governance roles and permissions

Microsoft Purview data security solutions

- Data Loss Prevention
- Data Security Investigations
- Data Security Posture Management
- Information Barriers
- Information Protection
- Insider Risk Management

Microsoft Purview data compliance solutions

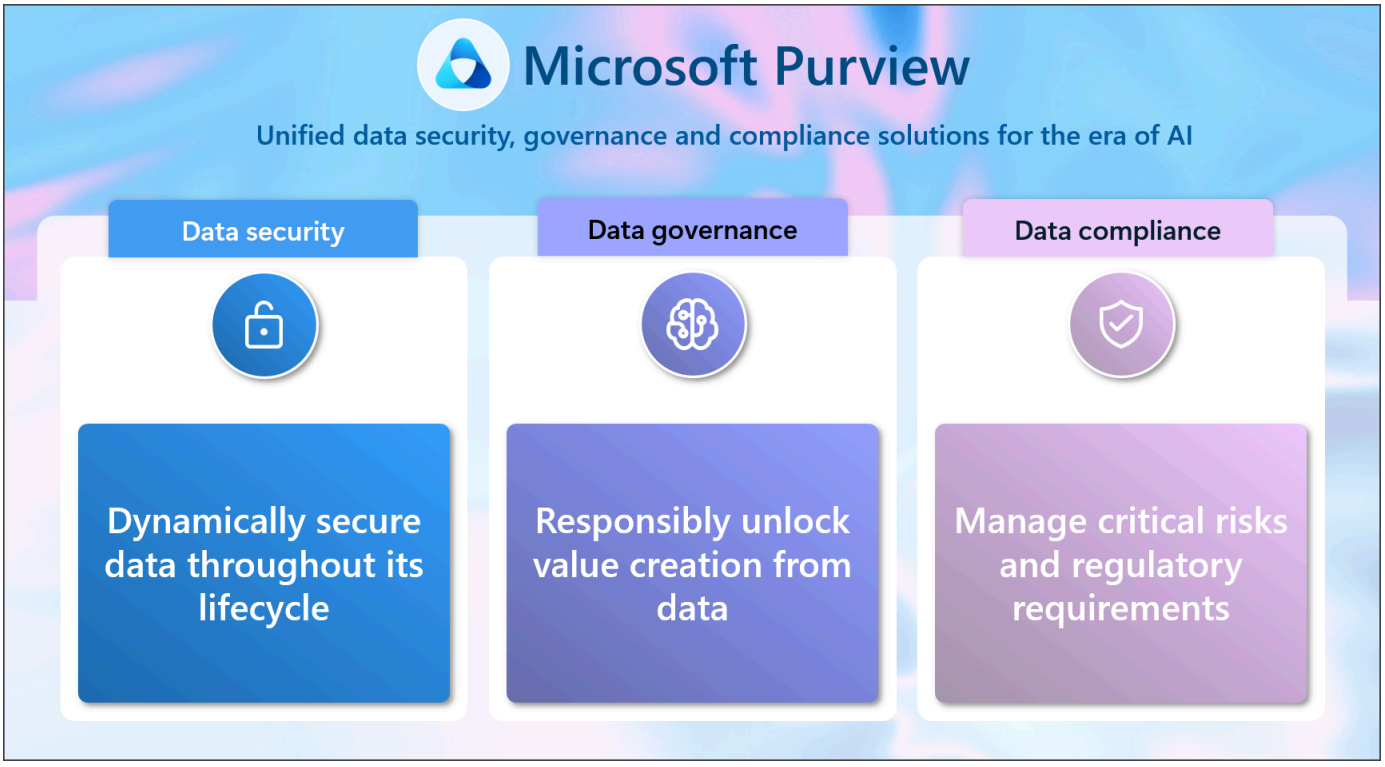
- Audit
- Communication Compliance
- Compliance Manager
- Data Lifecycle Management
- eDiscovery
- Records Management

Learn about Microsoft Purview

Microsoft Purview is a comprehensive set of solutions that helps your organization govern, protect, and manage data in the era of AI, wherever your data lives. Microsoft Purview solutions provide integrated coverage and help address the fragmentation of data across organizations, the lack of visibility that hampers data protection and governance, the rapid acceleration of AI transformation, and the blurring of traditional IT management roles.

Microsoft Purview combines solutions and services into a unified platform that helps your organization:

- Gain visibility into data across your organization
- Safeguard and manage sensitive data across its lifecycle, wherever it lives
- Govern data seamlessly in new, comprehensive ways
- Manage critical data risks and regulatory requirements
- Protect against accidental oversharing and sensitive data leakage in generative AI



Unified data security, governance, and compliance solutions for the era of AI

Microsoft Purview provides organizations with a powerful platform for seamlessly securing and governing data. This unified approach helps organizations reduce complexity and improve

outcomes related to mitigating risks in the era of AI.

Adopt AI with confidence

AI transformation is accelerating at an incredible pace, which compounds organizations' challenges around securing their data. It's never been more important for organizations to mitigate and manage the risks associated with AI usage, and implement corresponding protection and governance controls.

Microsoft Purview data security and compliance protections for generative AI apps helps you protect data however you use AI, whether it's:

- Copilot experiences and agents,
- Enterprise AI apps that you build, or
- Other AI apps you use.

Explore the full range of Microsoft Purview capabilities for data security and compliance for AI.

Data security

Microsoft Purview provides a robust and coordinated set of data security solutions to help you dynamically secure data throughout its lifecycle. The solutions include:

- Microsoft Purview Data Loss Prevention
- Microsoft Purview Data Security Investigations
- Microsoft Purview Information Barriers
- Microsoft Purview Information Protection
- Microsoft Purview Insider Risk Management
- Microsoft Purview Privileged Access Management

Now in preview, Microsoft Data Security Posture Management (DSPM) helps you discover, protect, and investigate sensitive data risks across your digital estate. Unifying other Microsoft Purview solutions, DSPM provides visibility and control for both traditional applications and AI apps and agents, so you can monitor, assess, and remediate data risks regardless of where sensitive data resides.

Data governance

Microsoft Purview includes unified data governance solutions for managing data services across your data estate and developing a federated approach to governing your data and unlocking business innovation. The solutions include:

- Microsoft Purview Data Map
- Microsoft Purview Unified Catalog

Data compliance

Microsoft Purview includes data compliance compliance solutions to help your organization minimize compliance risks and meet regulatory requirements. The solutions include:

- Microsoft Purview Audit
- Microsoft Purview Communication Compliance
- Microsoft Purview Compliance Manager
- Microsoft Purview Data Lifecycle Management
- Microsoft Purview eDiscovery
- Microsoft Purview Records Management

Shared capabilities

Many capabilities that help you protect and govern data across Microsoft Purview are shared among multiple solutions. Some of these capabilities include:

- Classifiers
- Connectors
- Data explorer and activity explorer
- Sensitivity labels

The Microsoft Purview portal

The Microsoft Purview portal is designed to make securing and governing your data easy and efficient. The updated experience provides a consistent look and feel, a streamlined information architecture, and simplified navigation into Microsoft Purview solutions.

- **What's in the portal?** The Microsoft Purview portal provides access to data governance, data security, and data compliance solutions.
- **Start exploring**: Get details about accessing and navigating the portal by visiting Learn about the new Microsoft Purview portal.

Ready to get started?

- Start exploring the capabilities in the Microsoft Purview portal.

- Plan your AI data security strategy with Microsoft Purview data security and compliance for AI.
- Get started with Microsoft Purview data security solutions.
- Get started with Microsoft Purview data compliance solutions.
- Get started with Microsoft Purview data governance solutions.

Last updated on 05/15/2026

What's new in Microsoft Purview

Whether it's adding new solutions, updating existing features based on your feedback, or rolling out fresh and updated documentation, Microsoft Purview helps you stay on top of the ever-changing data governance, data security, and risk and compliance areas. Take a look at the following information to see what's new in Microsoft Purview.

What's planned for Microsoft Purview

Microsoft Purview continues to add new solutions and features to help with data governance, data security, and risk and compliance in your organization. Check out the following roadmap sites to learn more about what's planned for Microsoft Purview:

- Roadmap for data governance solutions.
- Roadmap for data security and risk and compliance solutions.

September 2026

Data Loss Prevention

- **General availability**: Integrate Microsoft Entra Global Secure Access with Purview to protect text, files, and AI interactions at the network layer, enforce restrictive actions based on DLP policies, and detect risky user activity through Insider Risk Management. It helps prevent sensitive data from being shared with untrusted cloud applications through browsers, apps, APIs, and add-ins, including generative AI platforms, social media, and collaborative platforms. See Learn about Microsoft Purview Network Data Security.

Information Barriers

- **New**: Create an Information Barriers policy compliance report to identify SharePoint sites, OneDrive accounts, and user-owned SharePoint Embedded containers that no longer comply after Information Barriers policy changes.

August 2026

Data Governance

- **Updated**: Data Quality capability requires the Microsoft Purview account and its data sources to be in the same Azure region.

Data Loss Prevention

- **In preview**: Use data loss prevention policies for non-Microsoft connected apps. Create DLP policies that protect sensitive data at rest in non-Microsoft connected applications, such as Box and Google Workspace. These policies use the existing Microsoft Defender for Cloud Apps connectors and support the same classification engine available for Microsoft 365 locations.
- **Updated**: The Endpoint DLP capability comparison now clarifies that policies for access by apps not included in a restricted apps list or restricted app group are supported on Windows but not on macOS.
- **Updated**: If an Exchange DLP policy uses sensitivity labels as a condition and policy tips while mandatory labeling is enabled, selecting a label from the mandatory-labeling prompt doesn't reevaluate the policy tip. The configured DLP action is still enforced.
- **Updated**: In cross-tenant Microsoft Teams chats and channels, DLP policy enforcement follows the tenant that initiated the conversation. A user's home-tenant DLP policies don't automatically apply when another tenant hosts the chat or thread.
- **In preview**: Prepare Endpoint DLP for permission changes in macOS 27. Deploy client version 101.26072 or later, configure protection modes, monitor permission status, and notify users if they disable the required **Device Control and Data Access** permission.
- **Updated**: Guidance for DLP policies in Microsoft Edge for Business now explains policy sync status, where to investigate managed and unmanaged app activity, and coverage limitations caused by duplicate app catalog entries, encoded traffic, dynamic endpoints, and shared consumer and enterprise URLs.

Data lifecycle management

- **Updated**: When you remove the last hold from an inactive mailbox, Microsoft 365 automatically applies a 30-day delay hold. Learn how to identify and remove delay holds before recalculating mailbox hold status.
- **New**: Remove a retention hold from an unlicensed OneDrive site by using Security & Compliance PowerShell. The guidance covers tenant-level and site-level holds, prerequisites, and common errors that can block site deletion.
- **In preview**: Permanently delete SharePoint and OneDrive files with Priority Cleanup. After simulation, review, and required approvals, selected files bypass the recycle bins and are no longer discoverable in SharePoint search, Microsoft 365 Copilot, or eDiscovery.

Data Security Posture Management

- **Updated**: Microsoft Purview support for Anthropic Claude Enterprise now includes data classification, Insider Risk Management, Communication Compliance, eDiscovery, Data Lifecycle Management, and Compliance Manager.

eDiscovery

- **Updated**: App-only authentication for eDiscovery cmdlets in Security & Compliance PowerShell is unsupported. Transition automations to Microsoft Graph APIs where available; the documented certificate-based authentication steps are best-effort guidance for existing automations and don't change the support status.

Sensitivity labels

- **New**: Before enforcing an auto-labeling policy, run it in simulation mode to identify which items it would label without making any changes. Review the match results and source distribution to determine whether the policy is ready to enforce.
- **New**: The **Insights** tab in the policy details panel provides an at-a-glance view of an auto-labeling policy's performance. The information shown varies depending on whether the policy is running in simulation or enforcement mode, helping you understand how it identifies or labels content.
- **In preview**: Sensitivity label policies for non-Microsoft connected apps. Create auto-labeling policies that protect sensitive data at rest in non-Microsoft connected apps, including Box and Google Workspace. These policies use the existing Microsoft Defender for Cloud Apps connectors and support the same classification engine available for Microsoft 365 locations.
- **Updated**: Added guidance describing how to provide external users access to open files protected by user-defined permissions sensitivity labels. For more information, see Limitations.
- **Updated**: Auto-labeling policies can target more than 100 SharePoint sites by using an existing adaptive scope with the SharePointAdaptiveScopes parameter with New-AutoSensitivityLabelPolicy.
- **New**: Review the minimum OneNote versions and supported sensitivity label capabilities for Windows, Mac, iOS, Android, and the web.
- **Updated**: If an active auto-labeling policy stops matching Exchange email because a rule failed to load, contact Microsoft Support to identify the affected rule.

Shared capabilities

- **New**: Run on-demand classification scans on Windows endpoints to discover sensitive information without waiting for files to be opened or modified. The new guidance covers billing and best practices, monitoring and troubleshooting, and frequently asked questions.
- **New**: Microsoft 365 audit log activities now include Microsoft Defender for Cloud activities, People Skills activities, Microsoft 365 Archive file-level policy activities, and Microsoft Purview permission activities.

July 2026

Data Governance

- **Updated**: Microsoft Purview protection policies no longer support Azure SQL Database.
- **New**: Configure a manual business continuity and disaster recovery environment for Microsoft Purview Data Map. The guidance covers creating a secondary account, mirroring scans, validating the environment, and requesting account promotion during a regional outage.

Data Loss Prevention

- **In preview**: Protect sensitive data in text and prompts by integrating with Microsoft Entra Global Secure Access (GSA). This integration enables organizations to intercept and inspect text and AI interactions at the network layer, enforce restrictive actions based on DLP policies, and detect risky user activity through Insider Risk Management. It helps prevent sensitive data from being shared with untrusted cloud applications through browsers, apps, APIs, and add-ins, including generative AI platforms, social media, and collaborative platforms. See Learn about Microsoft Purview Network Data Security.
- **In preview**: Exchange Online DLP policies can detect classification failures caused by timeouts, throttling, and other scanning errors. Administrators can enable classification-failure detection and use the **DocumentScanFailures** condition to apply different protection actions for specific failure types.
- **Updated**: Administrators can use Microsoft Fabric governance experiences in the OneLake catalog to understand DLP activity and adoption. These insights help identify evaluated workspaces, locate sensitive information, track policy adoption, and prioritize high-risk assets.

Data lifecycle management

- **Updated**: Priority cleanup policies now require three separate approvals before content can be permanently deleted: one Priority Cleanup administrator, one retention manager, and one eDiscovery administrator.

Information Protection

- **Updated**: Added a checklist to help administrators understand auditing, regional availability, label scope, encryption settings, sensitive information types, file eligibility, and required roles before creating an auto-labeling policy.
- **Updated**: Troubleshoot why a specific SharePoint or OneDrive file was or wasn't labeled. The diagnostic guidance covers unsupported formats, checked-out files, indexing delays, existing higher-priority labels, encryption prerequisites, transient failures, and unexpected labeling.
- **Updated**: Auto-labeling policies can remove or downgrade sensitivity labels. The documentation explains behaviors that are by design and often interpreted as bugs. These behaviors include activation results that differ from simulation, encryption restrictions, cross-tenant label identities, conflicts with label-application policies, and the effect of label removal on encryption.
- **Updated**: Auto-labeling simulation isn't a completely silent dry run. Simulation can generate activity alerts, duplicated policies start in simulation, management actions can remain unavailable during provisioning, and the review experience can mix auto-labeling and DLP results in test mode.
- **Updated**: Licensing and configuration locations are documented for client-side Office auto-labeling, service-side auto-labeling policies, and container-level labels. The guidance also describes common symptoms caused by missing licenses, roles, regional availability, or transient backend failures.
- **Updated**: Auto-labeling policy guidance now identifies supported and unsupported group types and explains membership-propagation delays. Mail-enabled security groups, distribution groups, and Microsoft 365 groups are supported, while mail-disabled, dynamic distribution, and deeply nested groups have limitations.

Insider Risk Management

- **In preview**: Unified alert experience combines the Triage Agent and Standard alert dashboards into a single alerts list page. View and manage both classic and agent-triaged alerts from one location, with the ability to preview agent summaries, alert and user details directly on the alerts list page.

- **In preview**: Expanded user profile details in the unified alert experience add additional user profile signals from Microsoft Entra, including office location, employee type, department, and last working date.
- **In preview**: Expanded note capabilities across alerts and cases. Analysts and investigators can now add and view notes on both alerts and cases. System-generated notes are automatically applied when there's a change in alert or case status, assigned user, closure, or case escalation.

Shared capabilities

- **Update**: Microsoft Purview role-group assignments can be configured with an expiration date so that access is revoked automatically. Temporary assignments can last from one day to two years and are supported by most built-in and custom role groups, except eDiscovery Administrator and eDiscovery Manager.

June 2026

Copilot Cowork

- **General availability (GA)**: Data security and compliance protections for Microsoft 365 Copilot Cowork

Data Loss Prevention

- **New**: Access Endpoint DLP device attribute data using Advanced Hunting. Query Endpoint DLP device configuration and policy sync attributes at scale through the DeviceInfo table's DlpInfo column in Advanced hunting in the Microsoft Defender portal, instead of relying on point-in-time exports from the Microsoft Purview portal.

- **New**: Create a DLP policy that uses device scoping. Scope an Endpoint DLP policy to specific device groups — for example, enforce policy only when Finance users access data from Windows devices, and not when the same users work from macOS — using dynamic device groups defined in Microsoft Entra ID.

- **In preview**: New **Email is received from** > **External users** condition for the **Microsoft 365 Copilot and Copilot Chat** policy location lets DLP policies prevent Copilot from using external email as grounding data, helping reduce prompt injection risk from untrusted senders. See Block external email from being processed (preview).

- **In preview**: Enhanced matched conditions for Exchange DLP events surfaces detailed non-sensitive information type (SIT) condition matches in DLP alerts and Activity Explorer for Exchange Online. Each matched condition includes the condition name, matched value, and source.

Data Security Investigations

- **In preview**: Endpoint DLP evidence collection is now available as a data source in Data Security Investigations. Investigators can query data captured by endpoint Data Loss Prevention (DLP) policies on onboarded devices and add the associated content to an investigation scope for AI-powered analysis. This integration enables aggregate analysis of endpoint exfiltration events instead of per-alert triage. For more information, see Search, review, and refine results in Data Security Investigations.
- **General availability (GA)**: Email and portal notifications for Data Security Investigations. Investigators receive notifications through the Microsoft Purview Notification Center and email when setup completes and investigations are ready to use.
- **Updated**: Data preparation in Data Security Investigations now runs automatically in the background as items are added to scope. You no longer need to manually initiate vectorization before using AI features.

Device Onboarding

- **New**: Monitor device health with the device health reports dashboard. Use the device health reports dashboard to monitor device onboarding status, policy update readiness, and feature readiness for Endpoint DLP.

eDiscovery

- **New**: A new *Convert supported file formats to HTML* option is available when adding search results to a review set and when exporting items from a review set in eDiscovery. When enabled, cloud-native file formats such as `.loop` and `.page` files are converted to HTML, making the content indexed and keyword searchable in the review set and easier to process in post-export workflows.

Information Protection client

- **In preview**: View and label files with the Information Protection client using macOS.

Insider Risk Management

- **General availability (GA)**: Select which generative AI apps to monitor in Insider Risk Management policy indicators. For Microsoft Copilot experiences and Enterprise AI apps, you can now select or deselect generative AI apps for monitoring, reducing alert noise, and avoiding unnecessary pay-as-you-go billing charges.

Sensitive information types

- **New**: Added definitions for the following sensitive information types:
  - China physical addresses
  - Colombia national ID
  - Colombia tax identification number
  - Greenland physical addresses
  - Russia physical addresses
  - Russia taxpayer identification number
  - Singapore physical addresses
  - South Africa physical addresses
  - Ukraine physical addresses

Sensitivity labels

- **Preview**: Rolling out, the sensitivity label setting to prevent connected experiences that analyze content now prevents all connected experiences in Word, Excel, and PowerPoint for Windows, rather than a subset of these experiences.
- **Preview**: Rolling out, the sensitivity label setting to prevent connected experiences that analyze content extends to Word, Excel, and PowerPoint across MacOS, iOS, and Android.

May 2026

Agent 365

- **General availability (GA)**: Data security and compliance protections for Microsoft Agent 365.

Data Governance

- **General availability (GA)**: Standalone data asset data quality scan is now generally available.
- **General availability (GA)**: Incremental data quality scan is now generally available.

- **General availability (GA)**: Configurable data quality thresholds for data quality rules and data assets are now generally available.

Data Loss Prevention

- **Updated**: Added the admin permissions (Directory Reader, Microsoft Edge administration, and Microsoft Intune administration) required to activate DLP policies for unmanaged cloud apps in Microsoft Edge for Business. For more information, see Learn about Data Loss Prevention for Cloud Apps in Edge for Business.
- **Updated**: Clarified Edge browser profile scope for cloud app DLP policies. Policies for unmanaged cloud apps on managed devices apply across all Edge profiles (work, personal, and InPrivate); policies for managed apps apply only in the Edge work profile. See Help prevent sharing via Microsoft Edge for Business to unmanaged AI apps from managed devices and Help prevent users from sharing sensitive info with cloud apps in Edge for Business.
- **Changed**: Removed DeepL and Zapier from the list of unmanaged AI apps supported by browser policies in Edge for Business.
- **In preview**: New **Block access for specific external domains or users** sub-option for the **Restrict access or encrypt the content in Microsoft 365 locations** action lets DLP policies for SharePoint and OneDrive block access to sensitive files for specific external domains or user SMTPs. See Actions and Help prevent sharing sensitive items via SharePoint and OneDrive with external users.

Data Security Investigations

- **New**: OCR support in Data Security Investigations. Image files are automatically processed with optical character recognition (OCR), and the extracted text is merged and vectorized for AI analysis.
- **New**: Custom examinations in Data Security Investigations. Define your own examination focus with custom prompts to analyze investigation content beyond the built-in examination areas.
- **Updated**: New guidance for working with large audit search results in Data Security Investigations. When audit searches exceed the approximately 3,000-item limit, use the Audit solution to analyze the full result volume, then split searches into smaller time-based slices for ingestion into an investigation.

Data Security Posture Management

- **General availability (GA)**: The new version of Data Security Posture Management is now generally available. Partner solutions for non-Microsoft data sources remain in preview, as does the Data Security Posture Agent. This current version provides guided workflows for proactive risk management and streamlines data security operations so you can more confidently adopt AI across your digital estate.
- **New**: Support for administrative units, to bring parity with the classic versions of DSPM and DSPM for AI.
- **New**: To optimize resources, processing is paused for Microsoft 365 data when tenants are inactive for more than 60 days, and automatically resume when you return to the solution. For more information, see Data updates paused for inactive tenants.
- **New**: The "Responsible AI FAQ for Data Security Posture Management" is replaced with the more detailed Application card for Data Security Posture Management to better help you understand this solution's AI capabilities, intended uses, limitations, evaluations, safety components, and best practices.
- **New**: Support for Anthropic Claude (Enterprise) when you add and configure the Anthropic Claude data connector, now in preview. Claude then displays as another AI application alongside Copilot, Copilot Studio, ChatGPT Enterprise, and other AI apps. Use activity explorer to see individual Claude interactions, such as who used Claude, when they used it, and what kinds of content were involved, just as you do for other AI apps. For more information about Purview support for Claude, see Use Microsoft Purview to manage data security & compliance for Anthropic Claude (Enterprise).

Information protection scanner

- **In preview**: Administrators can now enable, disable, and configure cluster-level scanner features from PowerShell.
- **In preview**: Custom Reporting populates additional columns and tables in the scanner cluster database so administrators can build their own reports directly against scan results in Power BI or any SQL-based reporting tool, without stitching together per-scan CSV reports.

Reports

- **Preview**: Custom posture reports let admins build tailored views of information protection and DLP activity. Assemble metric and chart cards in sections to answer organization-specific questions that complement the built-in posture reports.

Sensitivity labels

- **In preview**: Rolling out, manual labeling support for MP4 files in SharePoint and OneDrive. For more information, see Video support (MP4 files.
- **In preview**: Rolling out, a new label policy setting for meetings, **Apply meeting label to artifacts**, automatically applies the meeting's sensitivity label to recordings and their transcripts (.mp4 files), and to meeting notes (.loop files).
- **In preview**: You can now see the sync status of your sensitivity label publishing policies on the **Label policies** page, giving you visibility into when label policy updates are fully synced across Microsoft 365.
- **Updated**: The documentation section How to disable sensitivity labels for SharePoint and OneDrive (opt-out) now includes labeling behavior if you disable sensitivity labels for SharePoint and Onedrive after they've been enabled.
- **New**: For auto-labeling policies that are turned on and target SharePoint and OneDrive, per-policy review pages let you monitor daily labeling activity, spot-check labeled and failed files, and investigate labeling failures. For more information, see Policy-level labeling activity for SharePoint and OneDrive.

April 2026

Collection Policies

- **Preview**: Collection policies support sensitivity labels as a condition for scoping detection to items with specific sensitivity labels applied. This condition is supported with browser and network cloud apps detection.

Data Lifecycle Management

- **New**: Newly created Teams call data records (often abbreviated to CDRs, and sometimes also called *call detail records* or just *call records*) are no longer included with Teams chat retention policies. Instead, they are included in the new support for Teams call logs retention policies that you create by using PowerShell. These newly supported retention policies let you manage the deletion of calling-related data when this is required for compliance and regulatory requirements. Call data records previously included in Teams chat retention policies continue to be managed by those same policies.

Data Governance

- **In preview**: Use a one-time glossary migration and asset enablement process to curate data assets and columns with glossary terms. This process allows you to centralize the

management of glossary terms by migrating glossary terms created in the classic governance experience into Unified Catalog. When you complete the process, you can curate data assets and columns.

- **In preview**: New bulk import, editing, and moving capabilities can help you quickly scale operations in Unified Catalog:
  - Create data products in bulk
  - Create critical data elements in bulk
  - Create glossary terms in bulk and bulk edit glossary terms
  - Move multiple glossary terms between governance domains

- **General availability (GA)**: Now rolling out, the advanced resource sets capability is available to all customers. Pricing for advanced resource sets is consistent with existing rates for classic Microsoft Purview data governance.

- **In preview**: Data quality provides on-premises support for Oracle and SQL server. On-premises databases are scanned in on-premises infrastructure to ensure that data doesn't move beyond organizations' premises. You need to set up a Kubernetes cluster to host the runtime that scans these databases.

- **In preview**: Data quality thresholds help stakeholders and data consumers know when data quality scores fall below standards. You can configure alerts for a rule level threshold and a data asset level threshold.

Data Loss Prevention

- **Updated**: Restructured the just-in-time (JIT) protection documentation.
- **Updated**: The Get started with just-in-time protection article now focuses on deployment and configuration steps.
- **New**: A new conceptual article, Learn about just-in-time protection, now covers JIT concepts, terms, supported activities, device compatibility, and includes a detailed JIT workflow diagram.
- **Preview**: DLP policies for unmanaged cloud apps support a new URL contains text condition that detects when the URL of the cloud app contains specified text strings. You can use it as a condition to scope DLP rules to specific URLs, or as an exception to exclude specific URLs from policy enforcement.
- **Preview**: Email notifications for browser and network DLP rules notify end users via email when their activity is blocked. Notifications use a rolling 10-minute batching window to prevent excessive emails.

- **Preview**: Data loss prevention policy tip reference for Outlook for Android, iOS, and macOS. A new reference article covering DLP policy tips, supported conditions, oversharing dialogs, and override capabilities for Outlook on Android, iOS, and macOS.

Data Security Investigations

- **In preview**: Proactive AI insights from Data Security Posture Management (DSPM) automatically create and refresh a single investigation for your tenant every 24 hours. The DSPM exfiltration objective card displays risk counts across five fixed categories, giving security teams continuous visibility into recently exfiltrated sensitive data without manual investigation creation.
- **New**: A new Data Security Investigation Contributor role automatically provides Data Security Investigations access to members of several Microsoft Purview role groups. Members of the *Compliance Administrator* and *Organization Management* role groups have administrative and contributor access, while members of the *Data Security Management* and *Insider Risk Management* role groups have contributor access without needing explicit role assignment.

Data Security Posture Management (preview)

- **New**: Microsoft Sentinel with partner solutions now also supports Varonis to provide holistic data insights for Salesforce.

Developers

- **Documentation update:** Added *Scenarios and API overview* section. This new section helps developers identify which APIs to use for specific scenarios. For more information, see Scenarios and API overview.

eDiscovery

- **In preview**: Organizations using Customer Key can now enable customer-managed key (CMK) encryption for direct export packages in eDiscovery. When enabled, exported investigation data is encrypted at rest using tenant-specific encryption scopes backed by customer-managed keys.
- **Updated**: The maximum number of review sets per case has increased from 20 to 100 for eDiscovery with premium feature support.
- **New**: The Advanced review set explorer (preview) includes a new left navigation pane to browse the review set schema, insert KQL operators, and run sample queries, and a new

Getting started tab with basic and advanced query templates to help you build queries faster.

Insider Risk Management

- **In preview**: Preview content while triaging alerts to quickly identify false positives, confirm the presence of sensitive data, and decide whether the alert warrants escalation.

Sensitivity labels

- **General availability (GA)**: Auto-labeling policies introduce a new flow where you must decide whether to automatically apply a sensitivity label, or remove a label when the configured conditions apply for files in SharePoint and OneDrive. When you chose to automatically apply a sensitivity label, you can now optionally choose to always overriding an existing label that has a lower priority label, even if it was manually applied. This option was previously available for emails only and now extends to files in SharePoint and OneDrive.
- **New**: Rolling out, users can now apply sensitivity labels configured for user-defined permissions while using Office for the web. A prerequisite for this functionality is that co-authoring is enabled for the tenant. If this prerequisite isn't in place, users still see the message that they must use a desktop app to apply the label.
- **New**: The **Label policies** page for label publishing policies has a new **Export policies** option, where the **Export to CSV** selection acts similarly to **Export** on the **Sensitivity labels** page. The **Export to Zip** selection includes more detailed information about the policies and all sensitivity labels in your tenant. For more information, see Export policy configuration in Microsoft Purview.

Shared capabilities

- **General availability (GA)**: Export policy configuration as a ZIP file containing a point-in-time snapshot of all policy configurations in XML format for DLP and sensitivity label publishing policies. Use the export for support requests, configuration reference, and local analysis by using PowerShell or Microsoft 365 Copilot.

March 2026

Data Governance

- **General availability (GA)**: Authoring custom data quality rules using SQL expression language is now generally available. Users can create custom rules using both Azure Data Factory expression and SQL expression languages.
- **In preview**: Configurable Data Quality thresholds allows users to define minimum acceptable quality scores at the data quality rule and data asset levels to align quality evaluation with business criticality.

Data Loss Prevention

- **Preview**: DLP supports adaptive scopes for scoping SharePoint policies. SharePoint location scoping

Data Security Investigations

- **New**: Categorization now includes a **Standard** and **Advanced** option. **Standard** categorization can significantly reduce the time it takes to complete processing and the amount of Data Security Investigation Compute Units (compute unit) needed for categorization.
- **In preview**: New support for the Data Security Posture Agent in Microsoft Purview. The Data Security Posture agent (preview) in Data Security Investigations helps your organization proactively surface credentials buried in data across your organization at scale.
- **Updated**: New guidance for how categorization processes data in Data Security Investigations. Categorization uses relevance scoring to prioritize the most relevant content for each selected category. Updated documentation includes considerations for results, content volume effects, and recommendations for using examination tools for comprehensive analysis.
- **New**: Data Security Investigations now supports soft purge for Exchange mailbox items. Soft purge moves items to the recoverable items folder, preserving the ability to restore items based on retention settings. Updated documentation includes guidance for choosing between soft purge and hard purge methods.
- **New**: Audit search in Data Security Investigations is now generally available. Use audit search to identify and collect content based on user activities recorded in the Microsoft Purview unified audit log, such as accessing, copying, or downloading files, and pull the associated content into your investigation.
- **Updated**: Data Security Investigations searches now respect compliance boundaries configured with search permissions filters. Investigators whose accounts are scoped by a compliance boundary only see search results for content locations within their permitted boundary.

- **New**: Personal data examinations in Data Security Investigations identify and extract personally identifiable information from selected data items in an investigation scope. Quickly assess which personal data types were exposed after a data security incident, including names, email addresses, financial account numbers, and Social Security numbers, with severity classification and AI-generated reasoning to support regulatory compliance reporting.

Data Security Posture Management (preview)

- **New**: Extend coverage of data insights to third-party SaaS and IaaS platforms by using Microsoft Sentinel with partner solutions to provide holistic data insights across Google Cloud Platform, Snowflake, and Databricks.
- **New**: You can now use federated credentials as a more secure method of authentication to run Fabric data risk assessments. This change is also available for data risk assessments in Data Security Posture Management for AI (classic). For more information, see Prerequisites for Fabric data risk assessments.

Deployment models

- **Updated**: Expanded Microsoft Purview deployment models with comprehensive step-by-step inline guides for five deployment scenarios: Prevent data leak to shadow AI, Secure and govern Microsoft 365 Copilot agents, Deploy and use Data Security Posture Management, Lightweight guide to mitigate data leakage, and Reduce false positives with SITs and advanced classifiers. Previously available only as downloadable PPTX and PDF files, each model now includes detailed articles following a step-based workflow.

eDiscovery

- **In preview**: Use the new Advanced review set explorer to query review set data with Kusto Query Language (KQL). Build advanced queries with complex filtering, pattern-based text extraction, and data visualization to analyze and find key information in your review sets.
- **New**: Configure sampling options when adding search results to a review set in eDiscovery. Choose confidence-based or percentage-based sampling to add a statistically representative subset of search results instead of all items. Completing the **Generate statistics** process is required to enable sampling.

Insider Risk Management

- **In preview**: Disable content download to create cases without content to reduce triage time. To get started, see Enable or disable content download.
- **General availability (GA)**: Microsoft Fabric indicators now include Lakehouse indicators.
- **General availability (GA)**: A new quick policy template for detecting data theft from non-Microsoft 365 apps by users leaving your organization is now available.
- **General availability (GA)**: Pay-as-you-go usage reports provide transparency and enable more accurate budget planning and policy tuning.
- **In preview**: The agent summary tab in the Triage Agent in Insider Risk Management has been enhanced to intelligently distill user activity risk into meaningful risk pattern narratives, contextual filtering options, granular activity signals, and provide specific files within alerts.

Sensitivity labels

- **General availability (GA)**: Manual labeling for OneNote, supported at the section level. Because retention labels also support labeling at the section level, you might find it helpful to review the architecture diagram for retention policies and retention labels. To enable the support for sensitivity labels, SharePoint and OneDrive must already be enabled for sensitivity labels, and then run the following PowerShell command: *Set-SPOTenant -EnableSensitivityLabelforOneNote $true*
- **In preview**: Auto-labeling policies introduce a new flow where you must decide whether to automatically apply a sensitivity label, or remove a label when the configured conditions apply for files in SharePoint and OneDrive. When you chose to automatically apply a sensitivity label, you can now optionally choose to always overriding an existing label that has a lower priority label, even if it was manually applied. This option was previously available for emails only and now extends to files in SharePoint and OneDrive.
- **In preview**: Viva Engage communities now support sensitivity labels applied to their underlying Microsoft 365 groups and connected SharePoint sites. The container label settings supported are privacy and guest access controls, to help you manage consistent protection with other groups and sites that support sensitivity labels. You can manually apply the label in Engage communities, and configure it in label policies as a default label for newly created communities.
- **Documentation update**: Now that container-level support for sensitivity labels include new collaborative workspaces such as Viva Engage communities and Loop workspaces, the previously titled article "Use sensitivity labels to protect content in Microsoft Teams, Microsoft 365 groups, and SharePoint sites" is now renamed Use sensitivity labels to protect collaborative workspaces (groups and sites).

February 2026

Data Governance

- **In preview**: Microsoft Purview Data Quality supports incremental data quality scans by using time-based filtering. By using incremental scans, you can choose between full scans, incremental scans, or both when you run data quality rules on data assets.
- **In preview**: Data quality scans for standalone data assets enable organizations to measure and improve data quality immediately without associating a data asset to a data product, significantly speeding up governance adoption. By running scans on assets, organizations can decide whether to associate a data asset to a data product if the data asset is in poor quality.
- **General availability (GA)**: Azure SQL Managed Instance support (SQL MI) is now generally available. You can now measure, understand, and improve the quality of your data in your SQL MI. Both public network and private endpoint configurations are supported. To get started, create a connection similar to Azure SQL database with supported port number, and assess the quality of your data using the Microsoft Purview Data Quality scanner.

Developers

- Microsoft Purview enables software development companies to integrate governance, protection, and compliance capabilities into their applications using SDKs and APIs. A new list of partner integrations is now available, with links to partner documentation. For more information, see Software Developer Partner Integrations.

Insider Risk Management

- **In preview**: Microsoft Fabric indicators now include Lakehouse indicators.
- **In preview**: A new quick policy template for detecting data theft from non-Microsoft 365 apps by users leaving your organization is now available.

Sensitivity labels

- **New**: The client-side improvements for sensitivity labels that extend SharePoint permissions to downloaded documents that started to roll out to Windows version 2601+ in January for the Current Channel is complete, and now also available with the Monthly Enterprise Channel.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

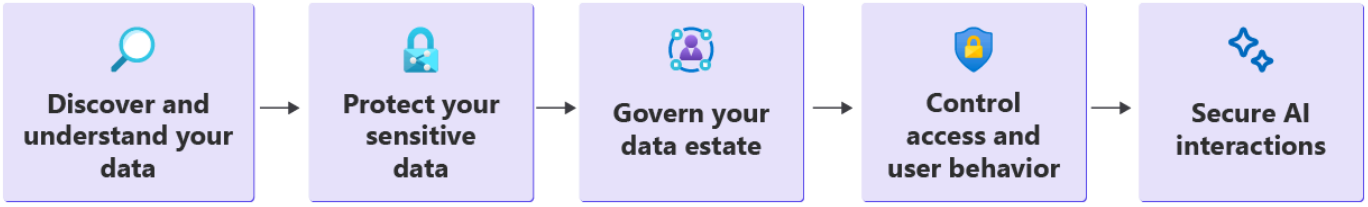
---

Last updated on 09/21/2026

Where to start with Microsoft Purview

Microsoft Purview brings together data security, data governance, and data compliance solutions in a unified platform, the Microsoft Purview portal. These solutions share capabilities, exchange signals, and reinforce each other to give your organization comprehensive visibility and control over its data.

But with its breadth of solutions and capabilities, it can be hard to know where to begin with Purview. There's no single path that works for every organization because priorities differ based on your industry, regulatory requirements, and how far along you are in your AI adoption. However, most organizations follow a similar progression:



1. **Discover and understand your data**: Find out what data you have, where it lives, and how much of it is sensitive.
2. **Protect your sensitive data**: Apply labels, encryption, and policies to prevent data loss.
3. **Govern your data estate**: Catalog your assets, ensure data quality, and establish accountability.
4. **Control access and user behavior**: Make sure users can only reach the data they need and that they interact with it safely.
5. **Secure AI interactions**: Extend your protections and compliance controls to Copilot, agents, and other AI apps.

This article walks through each stage and explains which Purview solutions apply, how they connect, and where to start.

> **Tip**
>
> Before you begin configuring solutions, make sure the right people in your organization have the roles and permissions they need. Most exploration tasks, like viewing Data Security Posture Management dashboards, require only reader roles.

Discover and understand your data

Before you can protect or govern anything, you need to know what data you have and identify which data is sensitive. Discovery is the foundation for every other stage, and Purview provides several ways to build that picture.

Assess your data security posture

Data Security Posture Management (DSPM) brings together signals from across Purview to give you a unified view of sensitive data risks. DSPM helps you discover where sensitive data lives, identify oversharing risks, and get personalized recommendations for remediation. If your organization uses AI apps, DSPM provides extra visibility into how AI interacts with your data.

DSPM is a good starting point if you want a single dashboard that shows what needs attention first.

Classify sensitive data

Sensitive information types and trainable classifiers identify sensitive items across your Microsoft 365 environment by using pattern matching, machine learning, or examples of data you provide. These classifiers are *shared* across Purview, so a sensitive information type you configure here is also available to Data Loss Prevention policies, Insider Risk Management, Communication Compliance, and data classification reports.

Data classification provides dashboards and explorers that show how data is classified and used across your organization. Use activity explorer to see how users interact with sensitive data, including labeling activity and AI interactions.

To learn more, see Get started with data classification.

Scan and catalog data assets

Data Map scans your data sources and multicloud environments to capture metadata about your data assets. This process creates a comprehensive inventory that spans your entire data estate and feeds the Unified Catalog, where data owners and stewards curate data products and data consumers discover trusted datasets. Data Map and Purview share data to support various shared capabilities, as detailed in How data flows in Microsoft Purview.

To learn more, see Plan for data governance.

Protect your sensitive data

After you understand what sensitive data you have and where it lives, the next step is to apply protections. Purview data security solutions work together to protect data throughout its lifecycle.

Apply sensitivity labels

Sensitivity labels are the foundational protection mechanism in Purview. Labels provide visibility into the sensitivity of data and can apply encryption, access restrictions, and visual markings. When you apply a label to a document or email, the protection travels with the content wherever it goes - across services, devices, and into AI interactions.

Many solutions use sensitivity labels. Data Loss Prevention enforces them, AI apps honor them, Microsoft Purview Audit logs capture them, Microsoft Purview eDiscovery makes them searchable, and DSPM makes them visible. When you define and apply your labels, you build the foundation for protections that extend across the entire platform.

To learn more, see Get started with sensitivity labels.

Prevent data loss

Data Loss Prevention (DLP) policies monitor for sensitive information across Microsoft 365 services and endpoints. In a DLP policy, you define the sensitive information to monitor, the locations to monitor, the conditions that must be matched, and the actions to take when a match is found. DLP uses the same classifiers and sensitivity labels that other solutions rely on, so your protection policies are consistent across the platform.

To learn more, see Learn about data loss prevention.

Manage data retention and deletion

Data Lifecycle Management provides retention and deletion policies that help you keep content you need and delete content you don't. Retaining data often addresses compliance and regulatory requirements. Deleting data that no longer has business value reduces your exposure to cyberattacks and data breaches. These policies apply across Exchange, SharePoint, OneDrive, Teams, and other Microsoft 365 services.

To learn more, see Get started with data lifecycle management.

Govern your data estate

With sensitive data protected, your organization can focus on making sure all data is well-organized, trustworthy, and accessible to the right people. Purview data governance solutions help your organization establish a federated governance model where a central data office sets the rules and domain experts manage the data they know best.

Curate data in Unified Catalog

Unified Catalog is a searchable catalog powered by the metadata captured in Data Map. Data owners and stewards use it to curate data products, manage data quality, and grant access to data consumers. Data consumers discover trusted datasets through a self-service experience with natural language search.

Governance domains organize data products and business concepts to make data easier to find and use. Built-in data quality capabilities help data owners measure and improve the health of their data, which is especially important in the era of AI because AI systems that ingest trusted, high-quality data produce more accurate and reliable results.

How governance supports security and compliance

Governance and security complement each other. Data governance helps you understand *what* data you have and *where* it lives. Data security helps you *protect* that data. For example, when Data Map identifies sensitive data sources, your security team can prioritize applying sensitivity labels and DLP policies to those sources.

Governance also supports compliance. Data lineage capabilities in Unified Catalog help you see the relationships between data products and their sources, which is useful for responding to regulatory inquiries and demonstrating compliance.

To learn more, see Plan for data governance and Get started with data governance.

Control access and user behavior

Protecting data and governing data assets addresses the data itself. But you also need to make sure that *people* interact with data in safe, secure, and compliant ways. Purview provides several solutions that monitor user behavior, enforce access boundaries, and help you respond to incidents.

Detect and act on insider risks

Insider Risk Management identifies potentially risky user activities by correlating signals from Microsoft 365 and third-party sources. It can detect behaviors such as unusual data downloads, sensitive data sharing, and risky AI usage patterns. Insights from insider risk signals are integrated into Microsoft Defender XDR for a comprehensive view of security threats.

To learn more, see Learn about insider risk management.

Restrict communication and collaboration

Information Barriers restrict two-way communication and collaboration between groups and users in Teams, SharePoint, and OneDrive. Often used in highly regulated industries, Microsoft Purview Information Barriers help avoid conflicts of interest and safeguard internal information between users and organizational areas.

To learn more, see Learn about information barriers.

Limit privileged access

Privileged Access Management helps protect your organization by limiting standing access to sensitive data or critical configuration settings. Instead of administrators having constant access, just-in-time access rules are implemented for tasks that need elevated permissions.

To learn more, see Learn about privileged access management.

Monitor organizational activity

Audit captures, records, and retains user and admin activities across Microsoft 365 services. Audit logs are essential for security investigations, compliance obligations, and forensic analysis. Audit records include AI interactions, such as user prompts and responses with Copilot and other AI apps.

To learn more, see Get started with auditing solutions.

Detect inappropriate communications

Communication Compliance monitors email and Teams communications for regulatory violations, harassment, threats, and inappropriate sharing of sensitive information. You can also set policies to monitor interactions with generative AI apps.

For more information, see Learn about communication compliance.

Investigate data security incidents

Data Security Investigations uses generative AI to help your security team analyze and respond to data security incidents, identify risks from sensitive data exposure, and collaborate more effectively on remediation.

Meet regulatory requirements

Compliance Manager helps your organization assess its compliance posture against more than 360 industry and regional regulatory standards. It provides risk-based scoring, improvement actions, and prebuilt assessment templates, including templates for AI regulations.

eDiscovery helps your organization identify, collect, preserve, and export electronic information for legal cases and regulatory investigations. Records Management helps declare content as records or regulatory records, manage retention schedules, and conduct disposition reviews with proof of deletion.

For more information, see Get started with data compliance solutions.

Secure AI interactions

The rapid adoption of generative AI adds a new dimension to every stage described in this article. AI apps can surface content faster and at greater scale, which amplifies the risk of oversharing and sensitive data leakage. Purview provides data security and compliance capabilities for AI that span Copilot experiences and agents, enterprise AI apps, and other AI apps.

The protections you built in earlier stages carry forward into AI scenarios:

- **Sensitivity labels** applied to content are honored by AI apps. If a user doesn't have permission to access data, AI apps don't return that data. Users must have the EXTRACT usage right, as well as VIEW, for AI apps to return labeled, encrypted data.
- **DLP policies** can block sensitive information from being shared with third-party AI apps. Endpoint DLP policies can warn or block users from pasting sensitive data into AI sites accessed through a browser.
- **Insider Risk Management** includes a risky AI usage policy template that can detect prompt injection attacks and attempts to access protected materials.
- **DSPM** provides visibility into how AI is used across your organization, surfaces AI-specific risks, and offers one-click policies to address them.

- **Audit** captures user prompts and AI responses, and these records flow into activity explorer for monitoring and analysis.
- **Communication Compliance** can monitor AI-generated communications for regulatory violations.
- **Retention policies** from Data Lifecycle Management apply to AI interaction data, ensuring it's kept for the required period and then deleted.
- **Compliance Manager** provides assessment templates specifically for AI regulations.
- **Data governance** ensures the data feeding your AI apps is high quality, well-curated, and trustworthy, so AI systems produce accurate and reliable results.

To learn more, see Microsoft Purview data security and compliance for AI.

Shared capabilities across solutions

Purview solutions work well together because they share a common foundation of capabilities. When you configure these capabilities in one solution, they're available everywhere:

Expand table

Shared capability	How it's used
Classifiers (sensitive information types and trainable classifiers)	Used by DLP, Information Protection, Insider Risk Management, Communication Compliance, and data classification reports. Define once, use everywhere.
Sensitivity labels	Applied by Information Protection, enforced by DLP, honored by AI apps, captured by Audit, searchable in eDiscovery, and visible in DSPM.
Connectors	Bring third-party data into Microsoft 365 for compliance, governance, and security scenarios.
Activity explorer	Surfaces user and admin activities related to sensitive data, labeling, and AI interactions across solutions.

Iterate and improve over time

Deploying Purview isn't a one-time project. The deployment guides in Microsoft Purview deployment models consistently emphasize an operational phase where you expand scope, refine policies, and act on what you learn. You don't need to get everything right upfront. Start with foundational protections and improve over time.

As your deployment matures, revisit these key areas:

- **Refine your classifiers.** Start with built-in sensitive information types and expand to Exact Data Match, trainable classifiers, and document fingerprinting as you learn which detections produce false positives and which miss important content.
- **Enable Adaptive Protection.** Adaptive Protection connects Insider Risk Management signals to DLP policies so that protections adjust dynamically based on user risk level. This approach reduces friction for low-risk users while applying stricter controls when risk increases.
- **Expand to new locations and workloads.** Most organizations start with a subset of locations (for example, SharePoint and Exchange). Over time, expand DLP policies to endpoints, Teams chat, and third-party apps. Extend sensitivity labels retroactively to existing content using auto-labeling policies.
- **Use data security objectives.** DSPM provides guided data security objectives that group relevant Purview solutions into end-to-end workflows. Each objective focuses on a specific security goal, such as preventing oversharing or data exfiltration, with prioritized actions, progress tracking, and one-click policy configuration.

The key principle is to start simple, monitor results, and expand when ready. Each improvement you make builds on the foundation you already have in place.

Next steps

- Learn about Microsoft Purview and explore the full range of solutions.
- Get started with data security solutions to protect sensitive data.
- Plan for data governance and get started with Unified Catalog and Data Map.
- Get started with data compliance solutions to meet regulatory requirements.
- Explore Microsoft Purview data security and compliance for AI to secure your AI adoption.
- Explore Microsoft Purview deployment models for scenario-based step-by-step guides.

ⓘ **Note:** This article contains content created with AI. Learn more

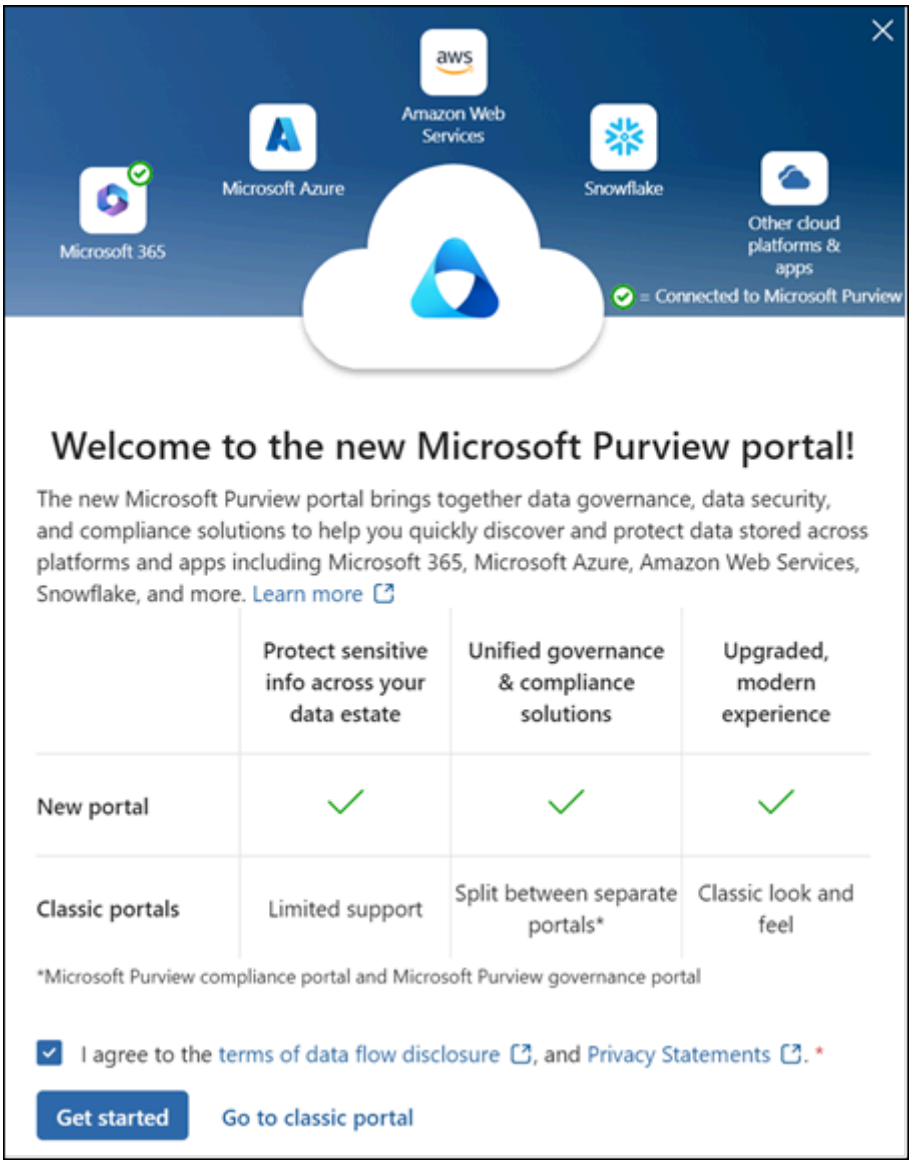
Last updated on 06/17/2026

Learn about the Microsoft Purview portal

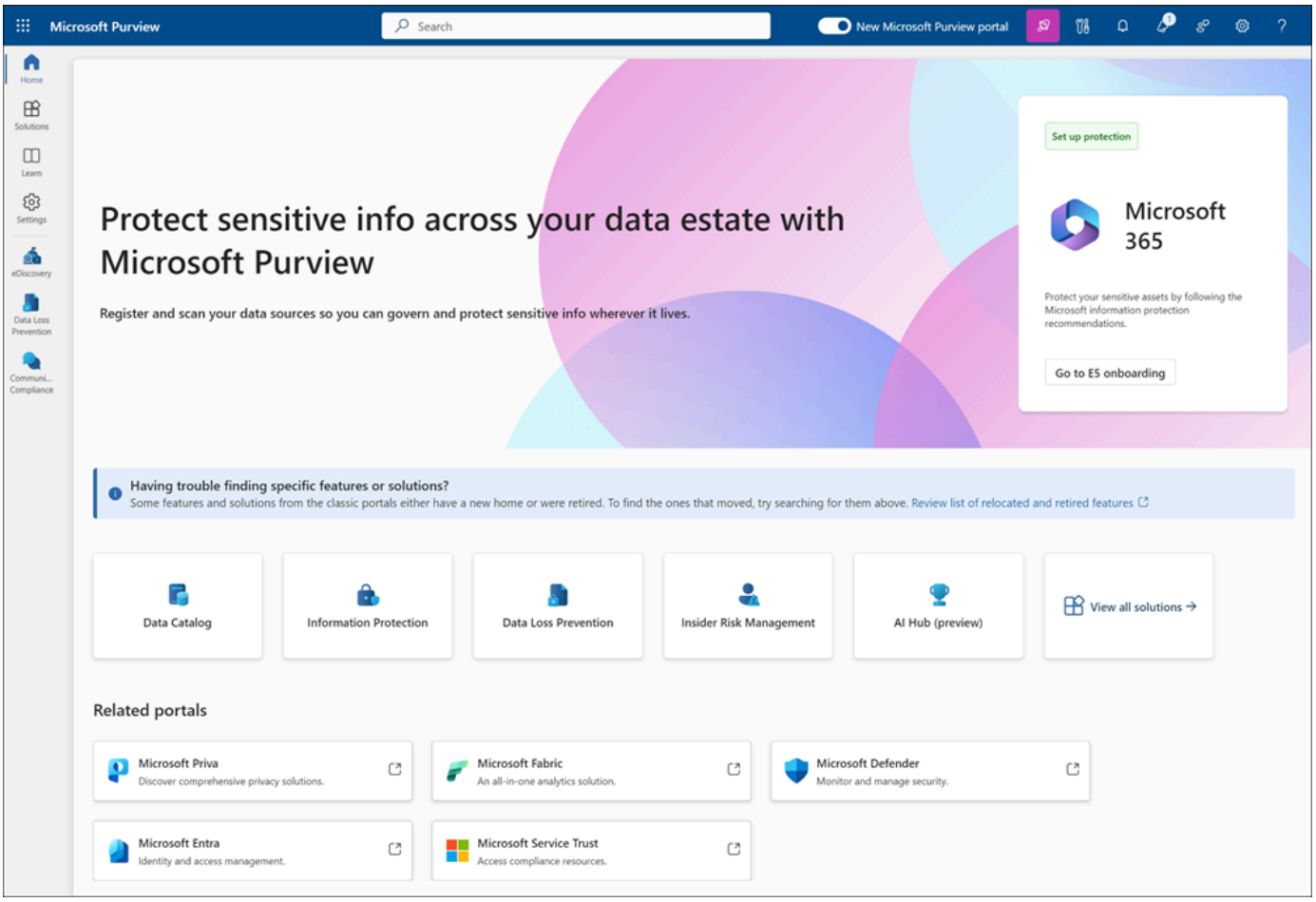
Microsoft Purview is a comprehensive set of solutions that helps you govern, protect, and manage data in your organization. The Microsoft Purview portal has a streamlined design and unified experience that helps you discover and access data security, data governance, and risk and compliance solutions for all your data, wherever it lives across your data estate. The unified experience streamlines navigation for all Purview solutions and provides a single-entry point for settings, search, and roles and permissions management.

Get started with the portal

When you first navigate to and open the Microsoft Purview portal, a welcome dialog appears that provides a brief overview of the new portal experience. After you agree to the terms and privacy conditions, select **Get started** to follow teaching bubbles that highlight key areas of the new portal experience.



Next, you see the new home page. The home page is your starting point for accessing all Microsoft Purview solutions, portal-wide settings, summary information about your data, and more.



Relocated portal features

Some features and capabilities that you're familiar with when using the retired Microsoft Purview compliance portal and the classic Microsoft Purview governance portal are relocated or retired in the Microsoft Purview portal.

Expand table

Feature or capability	Summary of changes
Classification	Classification is renamed to *Classifiers* and moved to the left-navigation area for each solution.
Content Search	Content Search features are now included as core components in eDiscovery.
Data asset search and browse	Data asset search and browse has moved to **Unified Catalog** > **Discovery** > **Data assets**.
Data Estate Insights	Data Estate Insights is renamed to *Reports* and moved to **Unified Catalog** > **Health management** > **Reports**.
Explorers	Explorers moved to the left-navigation area for each solution.
Glossary	The glossary moved to **Unified Catalog** > **Discovery** > **Enterprise glossary**.

Feature or capability	Summary of changes
Policies	Policies moved to the left-navigation area for each solution.
Privacy Risk Management	Privacy Risk Management is available in the Microsoft Priva portal.
Reports	Reports moved to the left-navigation area for each solution.
Roles and scopes	Roles and scopes are located in **Settings**.
Settings	Settings are located in the global left-navigation or at the top command bar of the portal. Select **Settings** to configure global and solution-level setting options.

Permissions and subscriptions

Depending on your permissions and your Microsoft Purview subscription, you see different solutions, home page cards, and features in the portal. For example, if you have permissions and a supported subscription to access a specific data governance or risk and compliance solution (such as **Data Loss Prevention**, **Insider Risk Management**, and others), you see these solution cards on the home page. If you don't have permissions or a supported subscription, you don't see these solution cards on the home page.

> **Note**
>
> You can grant Microsoft Purview permissions through security group membership. If you use Microsoft Entra Privileged Identity Management (PIM) for just-in-time membership for security groups in Microsoft Purview role groups, it can take up to two hours after activation for eligible administrators to have effective permissions applied in Microsoft Purview. For more information about PIM, see What is Microsoft Entra Privileged Identity Management.

For more information about permissions and role groups in Microsoft Purview, see:

- Permissions in the Microsoft Purview portal
- Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview portals

For more information about Microsoft Purview subscription requirements, see:

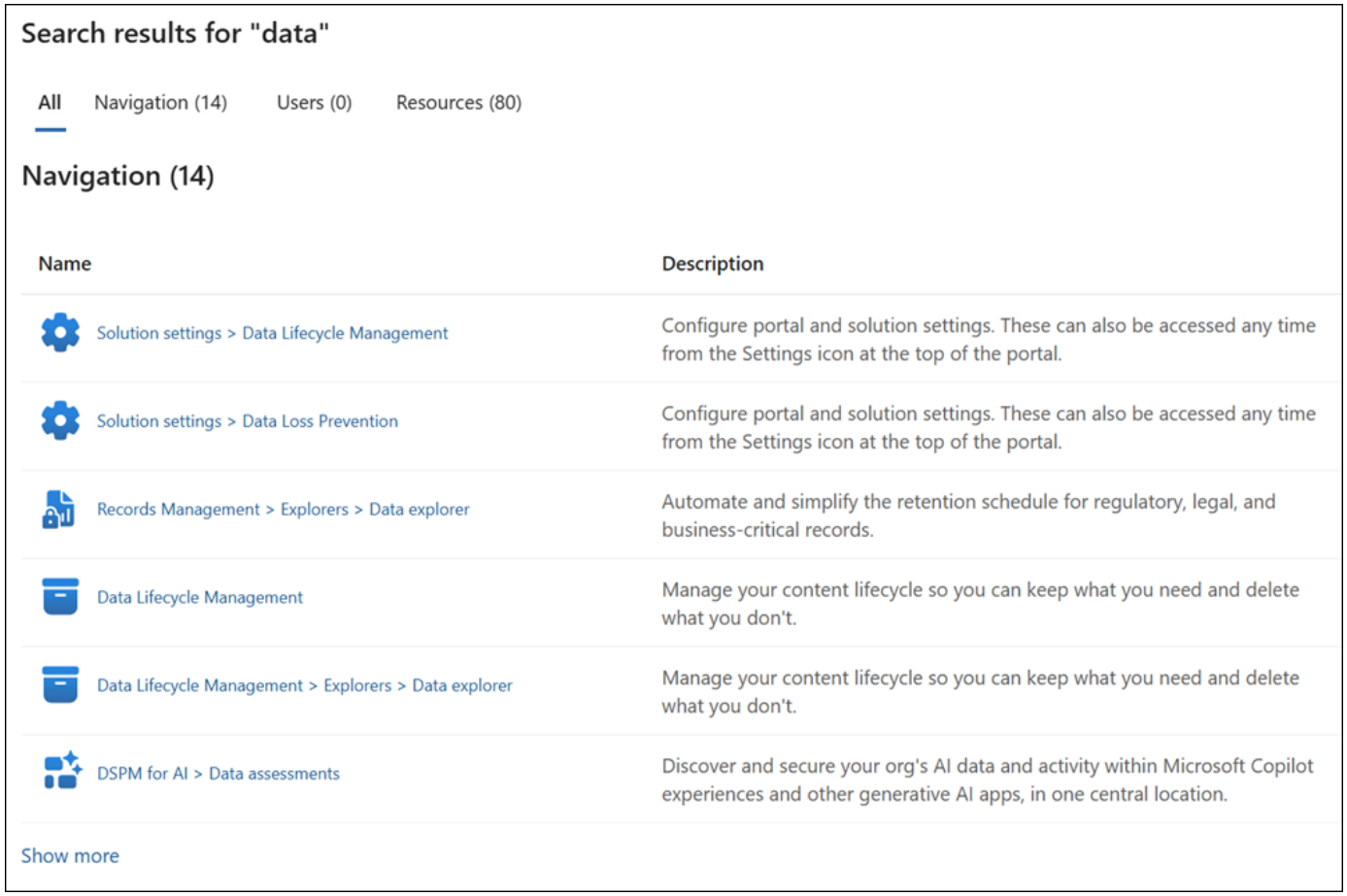
- Purview compliance solution requirements
- Purview governance solution plan requirements

Global search

Use global search at the top of the portal to search for navigation, users, and resources across all your solutions and data estate. For terms you enter in the search bar, you see applicable results grouped by the following sections in the search results drop down:

- **Navigation**: Results for items associated with specific Microsoft Purview solutions and solution features.
- **Users**: Results for users included in your organization. When you select a user, you see user details, including email, role groups, and administrative units assigned. To view role groups for a user, you must have the *Role management* role assigned that isn't restricted to an administrative unit.
- **Resources**: Results for helpful resources associated with the searched term.

To view the consolidated search results page, select the links to view all results at the bottom of any section listed in the search results drop down or select the **Search** arrow on the right side of the search bar. To filter the displayed results by section on the search results page, select the **All**, **Navigation**, **Users**, or **Resources** link at the top of the page. To view all results in each section, select **Show more**.



The five most recently entered search terms are retained as quick links in the search bar for quick access to current results for these terms.

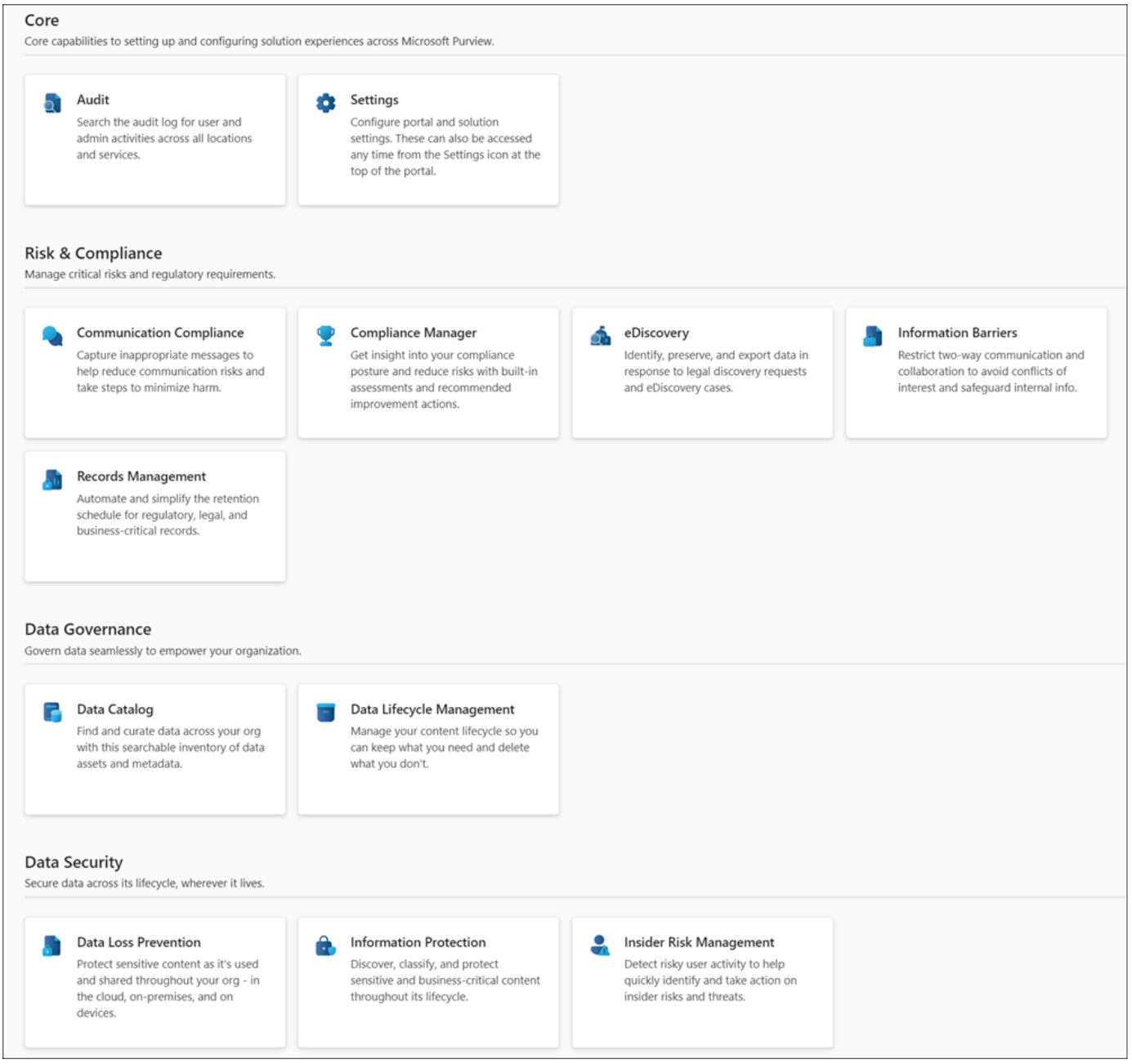
Sections and cards

Solution cards and the Solutions page

Solution cards on the portal home page allow you to quickly access and open the Microsoft Purview solutions that you have access to. You can't customize the displayed view for these cards.



To view and manage all solutions you have access to that aren't listed in the solution card list, select **View all solutions**. From the **Solutions** page, you can select solutions in the **Core** section to set up and configure app experiences across Microsoft Purview. You can also use the **Risk & Compliance**, **Data Governance**, and **Data Security** sections to directly view and manage specific Microsoft Purview solutions. Use the predefined filters to view solutions by area, or use the search box to find a specific solution by keyword. To explore learning materials (documentation, blogs, videos, and tutorials), select **Knowledge Center** in the **Resources** section.



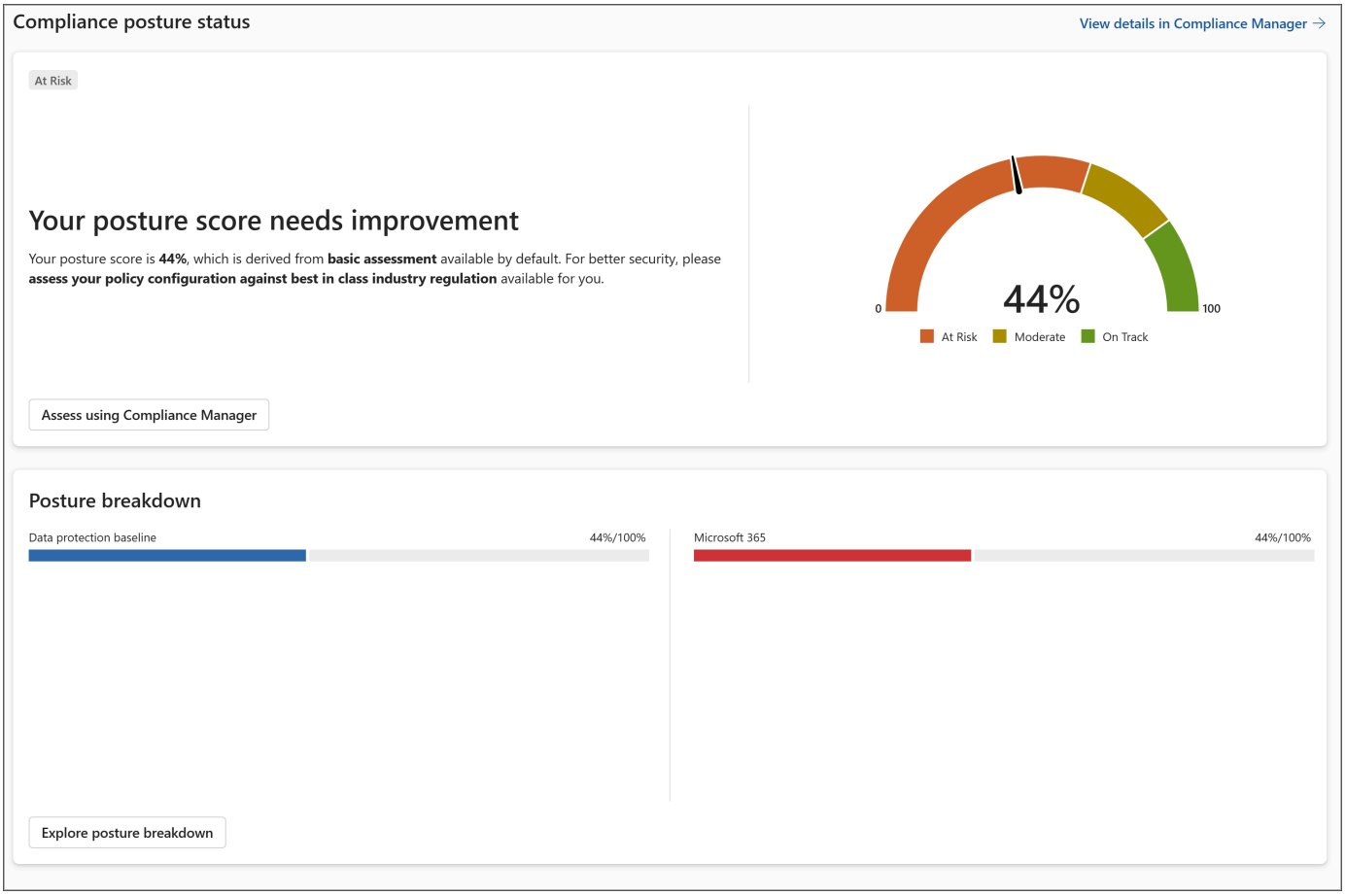
Related portals

Access related portals and solutions to help you manage all aspects of data analytics, data privacy, user identity, and more in your organization. You can access the following management and resource portals by selecting **Solutions** > **Related portals** in the left navigation:

- Microsoft Defender
- Microsoft Entra
- Microsoft Fabric
- Microsoft Priva
- Microsoft Service Trust

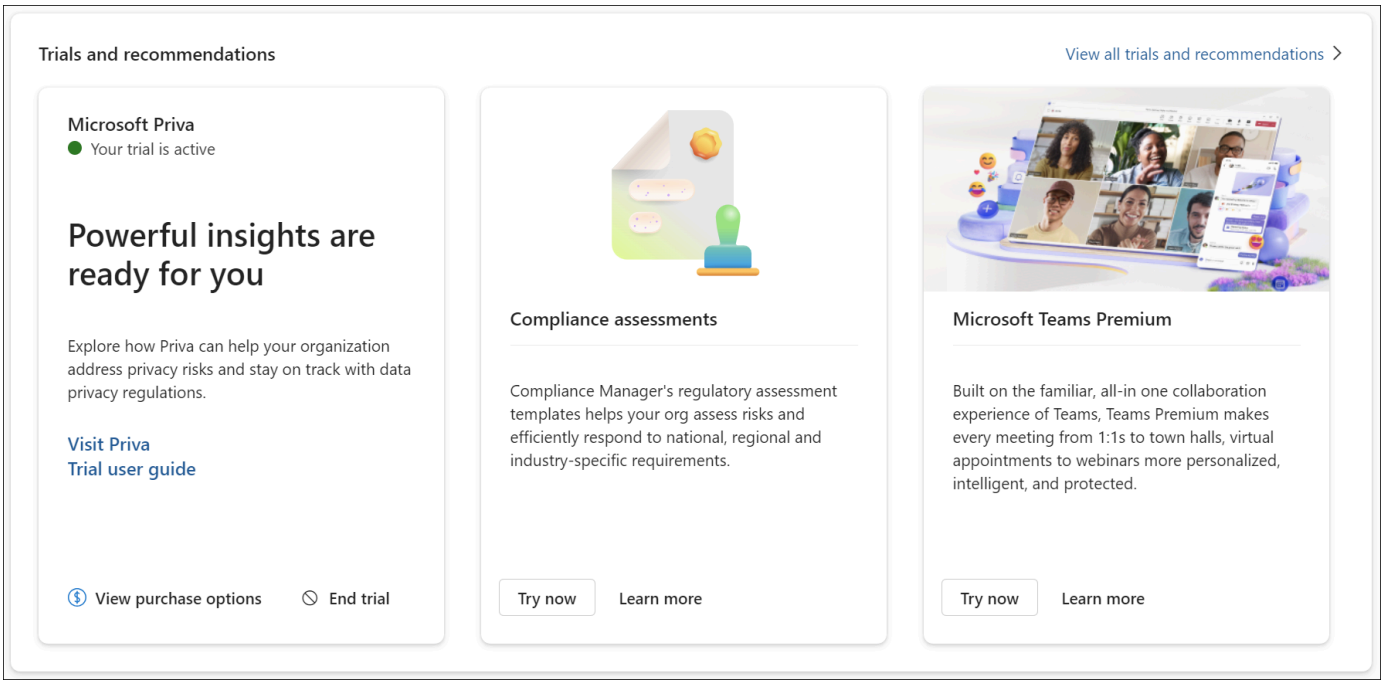
Compliance posture status

The **Compliance posture status** section includes cards that display summary information from the Compliance Manager solution about your organization's compliance posture. The compliance posture percentages reflect progress toward completing the requirements of regulatory assessments. The **Posture breakdown** section highlights completion rates for your organization's top assessments and the services covered by the assessments. Select **Assess using Compliance Manager** or **Explore posture breakdown** to visit Compliance Manager, where you can view, add, and manage assessments.



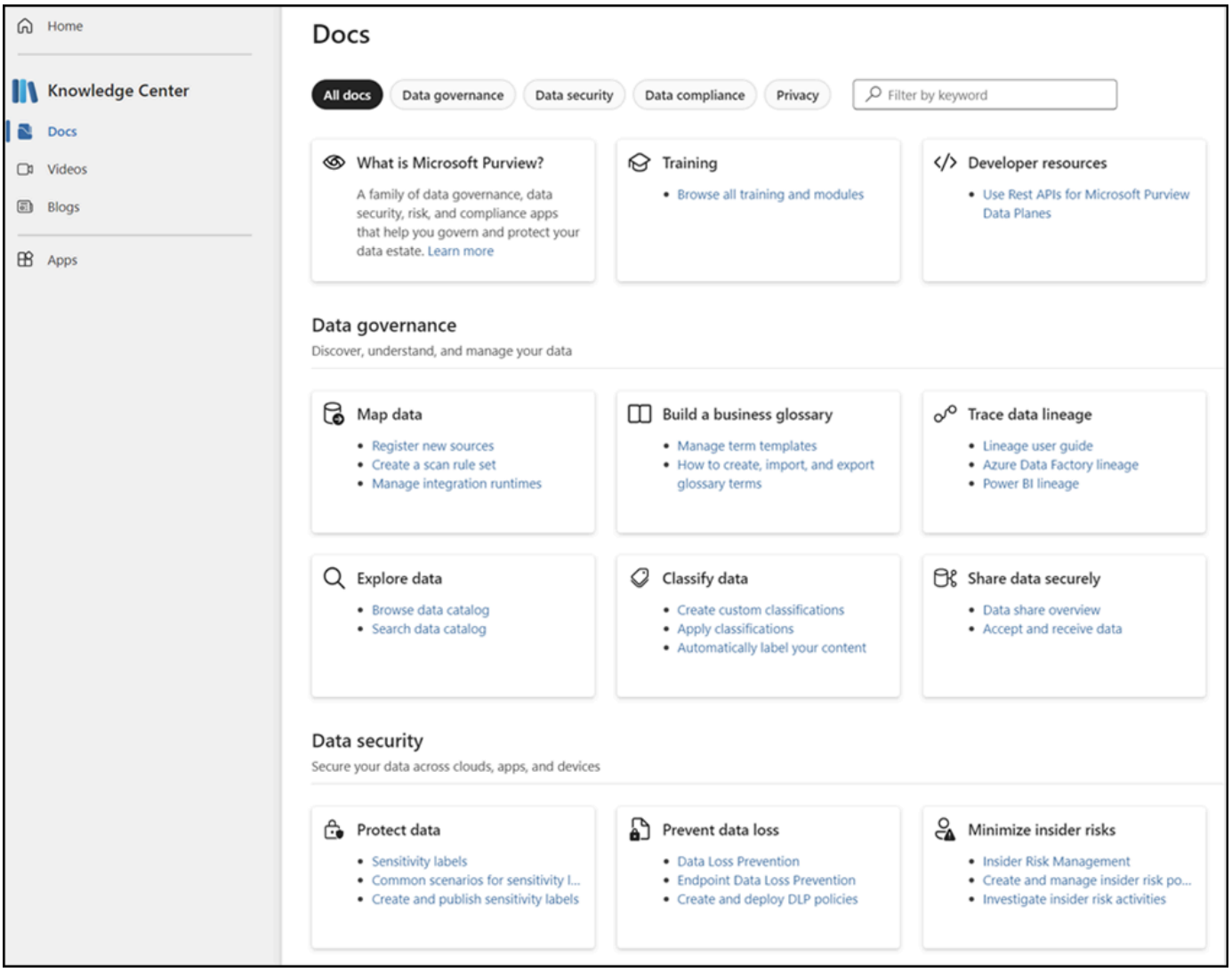
Trials and recommendations

The **Trials and recommendations** card displays information and links to help you get started with trial solutions for specific Microsoft Purview solutions. Select **View all trials and recommendations** to discover more security and compliance capabilities in Microsoft Purview you can try for free.



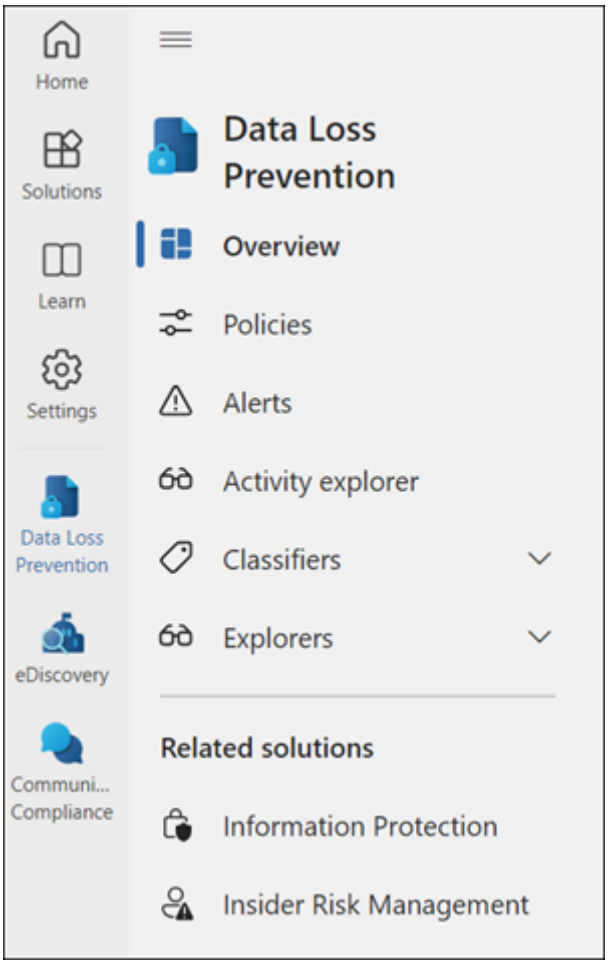
Knowledge Center

The **Knowledge Center** card highlights articles, product demos, tutorials, and other learning materials to help you succeed with Microsoft Purview. Select **Go to the Knowledge Center** to open the **Knowledge Center** page and view all available learning materials. Use the predefined filters to view knowledge center items by solution area, or use the search box to find a specific item by keyword.



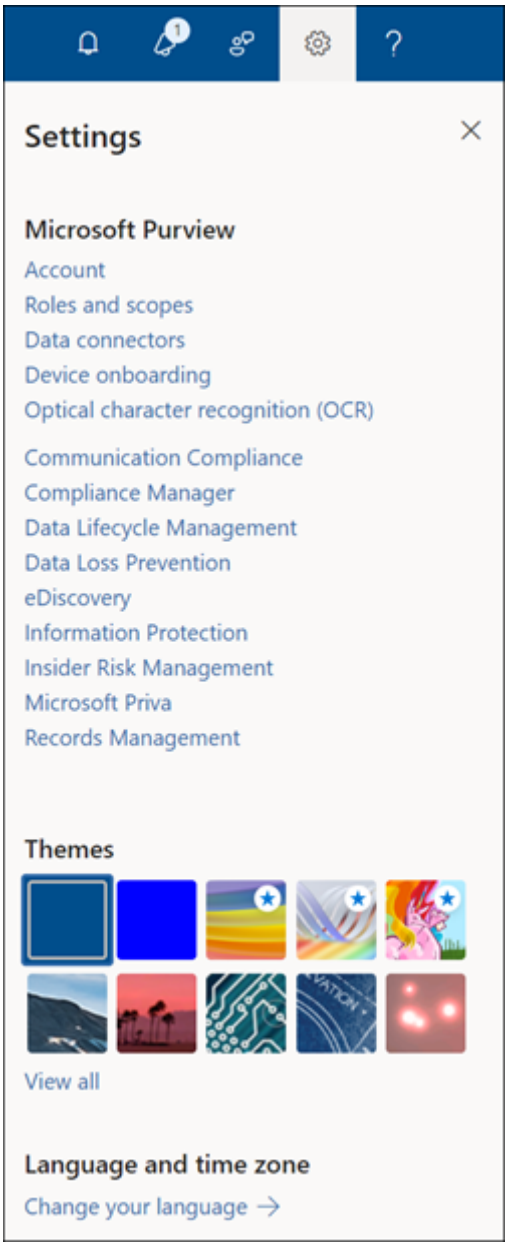
New left navigation for solutions

When you select a Microsoft Purview solution in the portal, you see a solution-specific home page and a new left-navigation experience that lets you access all solution features, settings, and more. The new left-navigation displays links to the **Home**, **Solutions**, **Agents**, **Learn**, **Usage center**, and **Settings** pages for quick and easy access to these areas for managing solutions, learning more about Microsoft Purview, configuring solution settings, and more. Additionally, links to the five most recent Microsoft Purview solutions also display for quick access to these solutions.



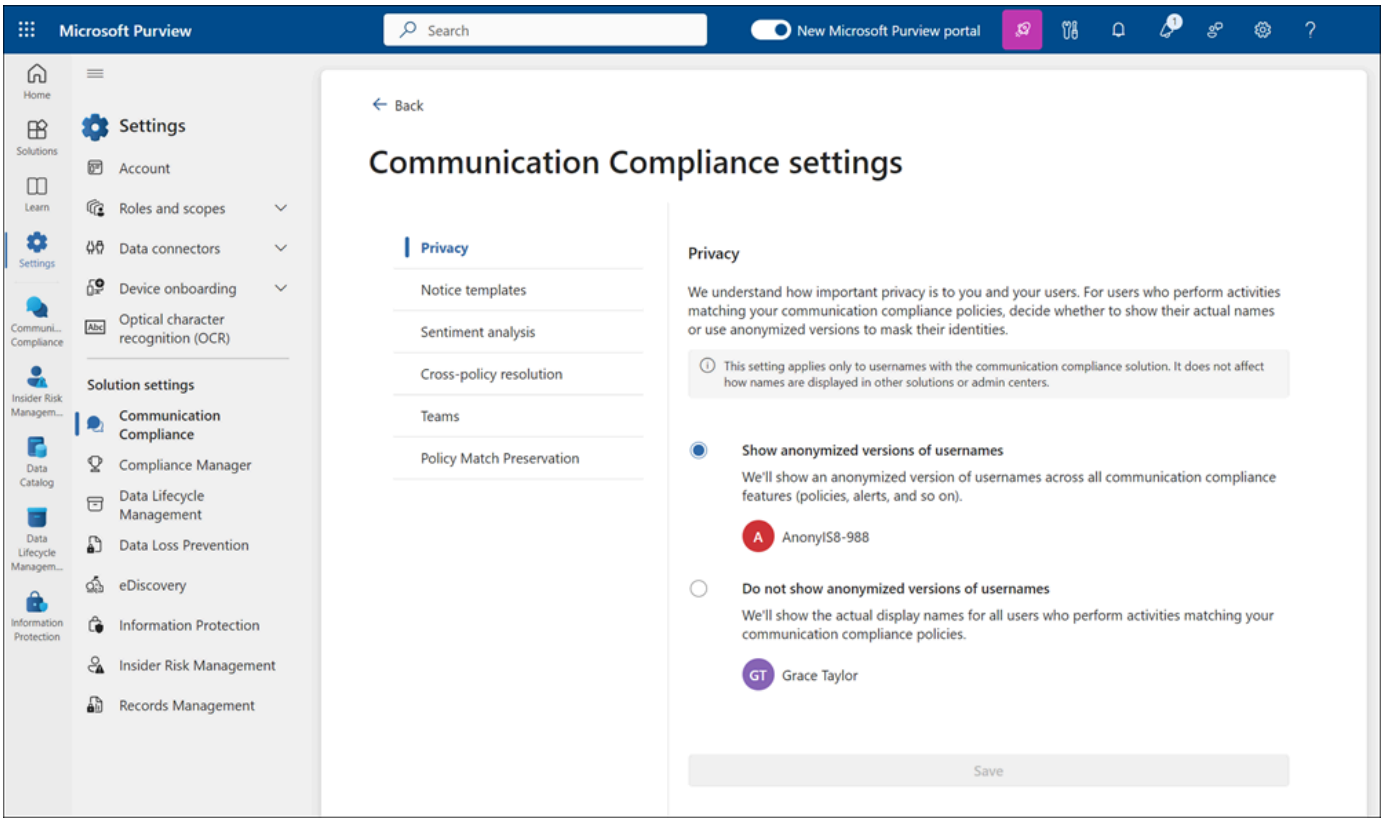
Settings in the portal

Settings in the Microsoft Purview portal are now centralized in the portal. By selecting the **Settings** icon in the left navigation or at the top-right of the page, you can quickly manage solution and global portal-wide settings, no matter where you are in the portal.



To view and manage portal-wide settings, select options in the **Themes**, **Language and time zone**, **Password**, and **Contact preferences** sections.

To view and manage solution settings, select a solution. This action opens the settings page for the solution where you can update settings for the selected solution. From any solutions settings page, you can also select other Microsoft Purview solutions in the **Solution settings** area in the left navigation to view and update settings in these solutions. To view and manage account settings, select **Account**.



Release notes and updates

View the latest release notes for Microsoft Purview solutions in a centralized location by selecting the **Release notes & updates** icon in the top right command bar. You can view Message center announcements for solutions that give you a high-level overview of a planned update or change and how it might affect your users or organization. Select the associated **Learn more** link to view the announcement and more details on the Message Center.

> **Note**
>
> You must have the `microsoft.office365.messageCenter/messages/read` permission to view release notes in the Microsoft Purview portal. The Message Center Reader, Global Reader, and Global Administrator roles include this permission. If you don't see an expected release note, check with your solution administrator or see What's new in Microsoft Purview.

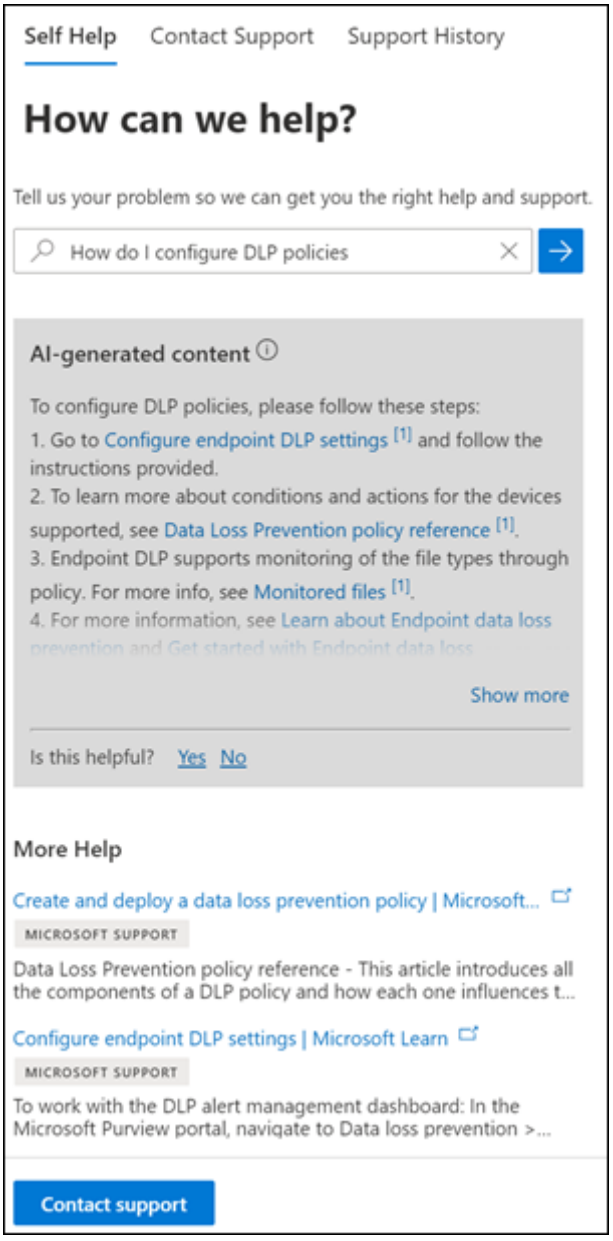
Help and support

Select the **Help** icon in the upper-right command bar to access specific help and support options for Microsoft Purview solutions:

- **Guided tours**: Start an integrated dialog-driven overview of key features and capabilities in the Microsoft Purview portal.

- **Knowledge Center**: Quickly access the Knowledge Center that contains highlighted articles, product demos, tutorials, and other learning materials to help you succeed with Microsoft Purview.
- **Need help?**: Select specific solution areas for self-help or share what you need help with so the right help and support can be provided.

For solution-specific help and support, select a specific solution in the drop-down field in the **Need help** section and select **Get help**. For example, if you need help with data loss prevention (DLP) policies, select *Data Loss Prevention*, and then select **Get help**. On the **Help** tab, enter the details for your help request.



Depending on your request, you see AI-generated guidance customized for your request and links to articles scoped to your request. If neither of these options helps resolve your questions, select **Contact support** to select email or telephone options to get support from a Microsoft agent.

> **Note**
>
> AI and article suggestions in help and support might not be available for some Microsoft Purview solutions. Check back frequently for updates about upcoming availability for these solutions in help and support.

Submit feedback about the new portal

Select the **Feedback** icon in the upper-right command bar to provide your feedback to Microsoft about the new Microsoft Purview portal. Provide as much detail as you'd like, but don't include any private or sensitive information. After entering your feedback, select **Submit** to send your suggestions to Microsoft to help improve the portal experience.

Get started with data governance solutions

Getting started with data governance solutions in the Microsoft Purview portal depends on your organization's current relationship with Microsoft Purview data governance solutions.

- If your organization didn't create a Microsoft Purview account in the Azure portal, read about the new experience for new customers.
- If your organization has Microsoft Purview accounts in the Azure portal, read about the new experience for existing data governance customers.

New experience for new customers

If your organization doesn't have any Microsoft Purview accounts in any subscriptions under your Microsoft Entra tenant, you can get started with our governance solutions right away. Use the new Microsoft Purview portal to start your journey with the free version of Microsoft Purview.

Free version of Microsoft Purview data governance solutions

If you're new to Microsoft Purview data governance solutions, you can start in the free version of Microsoft Purview. Because the free version has limited capabilities, use it only for initial evaluation, development, and test scenarios. To use the full capabilities of Microsoft Purview data governance, upgrade to the enterprise version.

Only a core subset of Microsoft Purview's governance solutions with limited capabilities are currently available in the free version:

- Catalog: Browse and search for your data assets.
- Data Map history: A log of updates made to assets.

Currently, the free version of Microsoft Purview governance solutions supports these governance data sources using live view:

- Azure Blob Storage
- Azure Data Lake Storage Gen 2
- Azure SQL Database
- Azure subscriptions
- Microsoft Fabric

You can annotate and curate assets that are available via live view, or use our APIs to create new entities.

Check out the guide to getting started with the free version of Microsoft Purview data governance solutions.

Enterprise version

When you upgrade to the enterprise version of Microsoft Purview data governance, you have access to all Microsoft Purview data governance features, including:

- All data sources supported by Microsoft Purview data governance supported governance sources: Manage data sources across your data estate.
- Collections: Manage fine grained access control for users across your environment.
- Scanning: Gather your sources' technical metadata and lineage information.
- Automatic classification: Classify your data to identify important and sensitive information.
- Data owner policies: Govern data at the source from Microsoft Purview.
- All reports: Understand your data estate at a glance.
- Workflows: Automate governance in your environment.

Learn more about data governance in Microsoft Purview.

New experience for existing data governance customers

The new experience enhances the current Microsoft Purview data governance experience. It doesn't affect the information already stored in your accounts or your ability to use the APIs. When you upgrade to the new experience, you automatically get the enterprise version of Microsoft Purview. The enterprise version has all the features you already use, plus these new features:

- This new Microsoft Purview portal: the latest platform to manage your organization's data governance.
- Live view: view your Azure data assets in the Microsoft Purview Unified Catalog automatically and in real time with no need to register or set up a scan for the source.
- Tenant-level administration: roles that give you the ability to delegate administration from the tenant level, while maintaining fine-grained access control options to limit permissions by role and collections. This feature gives your team flexibility when managing access and responsibilities.
- Data Map history: a log of updates made to data assets, so you can audit changes in your Data Map.

For more information about the new experience, see Guide to getting started with data governance in the new Microsoft Purview portal.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 08/05/2026

Microsoft Purview setup guides

Microsoft Purview setup guides offer a guided experience for planning and deploying data security and compliance solutions in your organization.

These guides organize the basic solution configuration steps and help point you directly to primary guidance for these steps in Microsoft Purview documentation. By using the authenticated experience in the admin center guides, you can also use automation for some solution configuration steps to help you quickly get started with your solution deployment.

Setup guides in the admin center require authentication to Microsoft Purview as an administrator or other role with access to the Microsoft 365 admin center. However, anyone can access the guides in the Microsoft 365 Setup portal.

Setup guides

Expand table

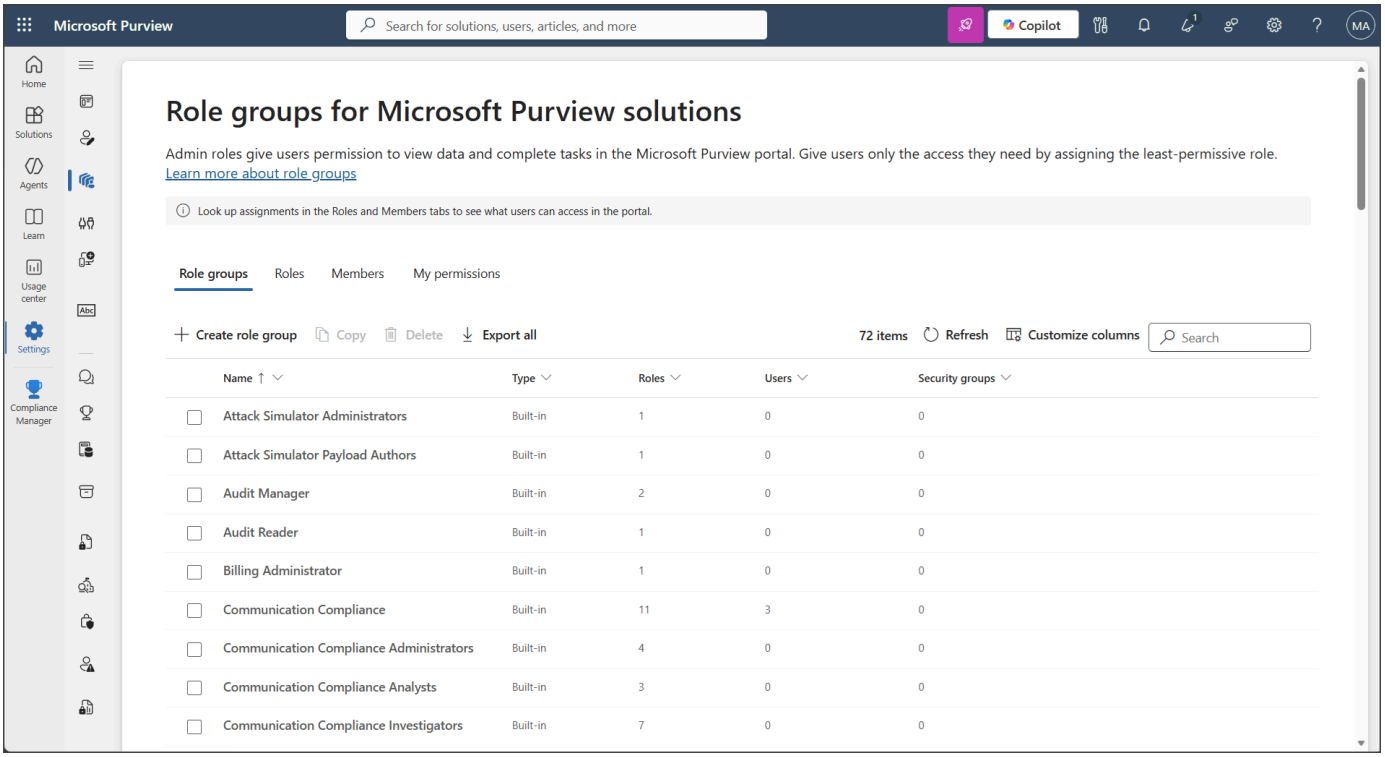
Guide: Setup Portal	Guide: Admin Center	Description
Audit setup guide	Audit setup guide	**Microsoft Purview Auditing Solutions setup guide** provides guidance to enable and configure auditing in Microsoft 365. The guide covers verifying licensing, assigning permissions, enabling Audit (Premium) events, creating audit retention policies, and using the unified audit log to search and review user and admin activities for investigations.
Manage data exfiltration with Purview	Manage data exfiltration with Purview	The **Manage data exfiltration with Microsoft Purview setup guide** provides steps to configure Insider Risk Management policies, link Data Loss Prevention alerts, and enable auditing and labeling features to prevent unauthorized data transfers. The guide highlights built-in tools that streamline permission setup for Microsoft 365 groups, reducing manual steps and improving configuration efficiency.
Secure by default with Purview	Secure by default with Purview	The **Secure by default with Microsoft Purview setup guide** provides guidance to enable default security configurations and apply encryption and sensitivity labels automatically across Microsoft 365. The guide covers publishing default labels, configuring auto-labeling for Office apps, setting contextual defaults for SharePoint and Microsoft Teams, and enabling built-in labeling for Office for the web to achieve broad coverage with minimal user interaction.

Last updated on 04/01/2026

Permissions in the Microsoft Purview portal

The Microsoft Purview portal supports direct management of permissions for users who perform tasks within Microsoft Purview. By using the **Roles and scopes** area in **Settings** for the portal, you can manage permissions for users across your Purview data security, data governance, and risk and compliance solutions. You can limit users to perform only specific tasks that you explicitly grant them access to.

> **Important**
>
> To view **Role groups** in the **Roles and scopes** area in the Microsoft Purview portal, users need to be a global administrator or need to be assigned the *Role Management* role (assigned only to the *Organization Management* role group). The *Role Management* role allows users to view, create, and modify role groups.



> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. When planning your access control strategy, follow the best practice to manage access with the least privilege for your users.

Microsoft Purview permissions

The Microsoft Purview portal uses the role-based access control (RBAC) permissions model. Most Microsoft 365 services use RBAC, so if you're familiar with the permission structure in these services, granting permissions in the Microsoft Purview portal is similar. The permissions you manage in the Microsoft Purview portal don't cover the management of all the permissions needed in each individual service. You still need to manage certain service-specific permissions in the admin center for the specific service. For example, if you need to assign permissions for archiving, auditing, and MRM retention policies, you need to manage these permissions in the Exchange admin center.

Solution-specific permissions

For specific Microsoft Purview solution permission guidance, see the following articles:

Data security

- Data Loss Prevention
- Data Security Investigations
- Data Security Posture Management
- Device Onboarding
- Sensitive Information Types - Custom
- Sensitive Information Types - Exact Data Match
- Sensitivity Labels

Data governance

- Collection Policies
- Data Classification - Activity Explorer
- Data Classification - Content Explorer
- Data Classification - Data Explorer
- Data governance - Classic Data Catalog
- Data governance - Unified Catalog
- Data Lifecycle Management
- Records Management

Risk and compliance

- Adaptive Protection - Insider Risk Management
- Administrative Units

- Audit
- Communication Compliance
- Compliance Manager
- eDiscovery
- Information Barriers
- Insider Risk Management
- Privileged Access Management

AI and Copilot

- Data Security Posture Management for AI
- Security Copilot for Purview
- Triage Agent in Data Loss Prevention
- Triage Agent in Insider Risk Management

To view all of the default role groups that are available in the Microsoft Purview portal and the roles that are assigned to the role groups by default, see Roles and role groups in the Microsoft Defender XDR and Microsoft Purview portals.

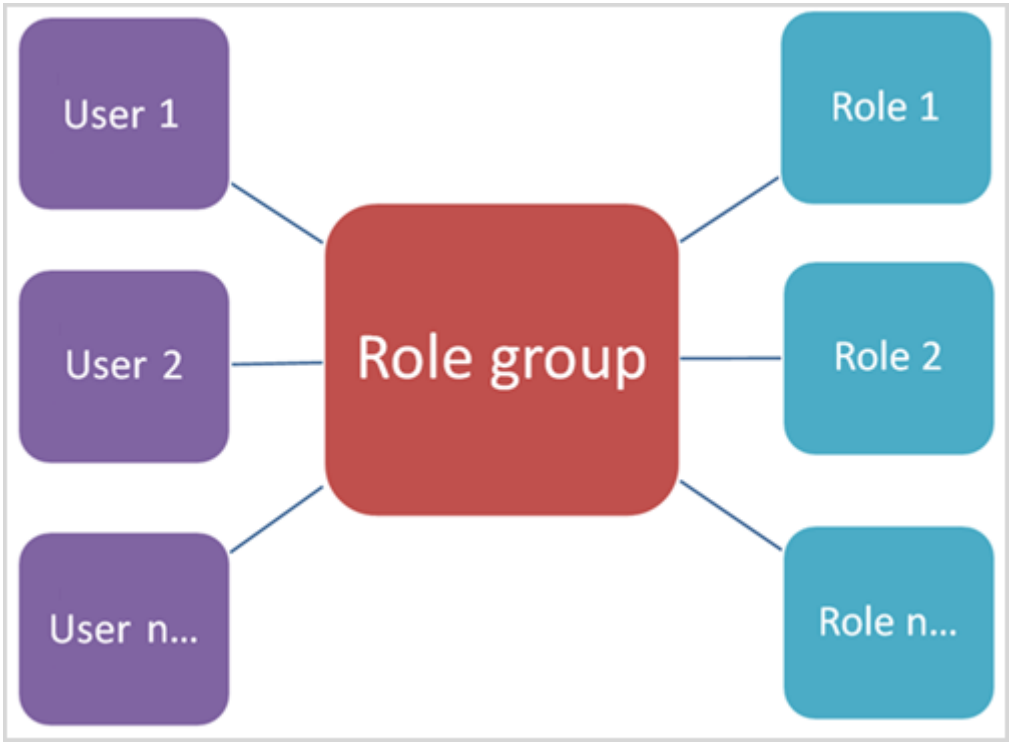
Managing permissions in the Microsoft Purview portal only gives users access to the compliance and governance features that are available within the Microsoft Purview portal. To grant permissions to other features that aren't in the Microsoft Purview portal, such as Exchange mail flow rules (also known as transport rules), use the Exchange admin center.

The current governance roles and role groups cover only broad access to the Microsoft Purview Data Map and Unified Catalog. For more access, Microsoft Purview governance uses a combination of role groups, data access, and solution-specific permissions.

Relationship of members, roles, and role groups

A role grants permissions to perform a set of tasks. For example, the *Case Management* role lets users work with eDiscovery cases.

A role group is a set of roles that enable users to do their jobs across compliance and governance solutions in the Microsoft Purview portal. For example, by adding users to the *Insider Risk Management* role group, designated administrators, analysts, investigators, and auditors get the necessary insider risk management permissions in a single group. The Microsoft Purview portal includes default role groups for tasks and functions for each compliance and governance solution that you need to assign people to. Generally, add individual users as members to the default role groups as needed.



Microsoft Entra roles in the Microsoft Purview portal

The roles that appear in the Microsoft Entra ID section of the **Roles and scopes** area are Microsoft Entra roles, and global administrators can see this section. These roles align with job functions in your organization's IT group, making it easy to give a person all the permissions necessary to get their job done. You can view the users currently assigned to each role by selecting an admin role and viewing the role panel details. To manage members of a Microsoft Entra role, select **Manage members** in Microsoft Entra ID. This choice redirects you to the Azure management portal.

For more information about Microsoft Entra roles, see Microsoft Entra built-in roles.

Expand table

Entra role	Description	Mapped Purview role groups
AI Administrator	Manage all aspects of Microsoft 365 Copilot and AI-related enterprise services in Microsoft 365. For more information, see AI Administrator.	Compliance Manager Administrators Data Security AI Admins Data Security AI Viewers

Entra role	Description	Mapped Purview role groups
Attack Payload Author	Create attack payloads but not actually launch or schedule them. For more information, see Attack Payload Author.	Attack Simulator Payload Authors
Attack Simulation Administrator	Create and manage all aspects of attack simulation creation, launch/scheduling of a simulation, and the review of simulation results. For more information, see Attack Simulation Administrator.	Attack Simulator Administrators
Compliance Administrator	Help your organization stay compliant with any regulatory requirements, manage eDiscovery cases, and maintain data governance policies across Microsoft 365 locations, identities, and apps. For more information, see Compliance Administrator.	Compliance Administrator
Compliance Data Administrator	Keep track of your organization's data across Microsoft 365, make sure it's protected, and get insights into any issues to help mitigate risks. For more information, see Compliance Data Administrator.	Compliance Data Administrator
Global Administrator	Access to all administrative features in all Microsoft 365 services. Only global administrators can assign other administrator roles. For more information, see Global Administrator / Company Administrator.	Data Catalog Curators Data Estate Insights Readers Organization Management Purview Administrators
Global Reader	The read-only version of the **Global administrator** role. View all settings and administrative information across Microsoft 365. For more information, see Global Reader.	Global Reader
Security Administrator	Control your organization's overall security by managing security policies, reviewing security analytics and reports across Microsoft 365 products, and staying up-to-speed on the threat landscape. For more information, see Security Administrator.	Security Administrator
Security Operator	View, investigate, and respond to active threats to your Microsoft 365 users, devices, and content. For more information, see Security Operator.	Security Operator
Security Reader	View and investigate active threats to your Microsoft 365 users, devices, and content, but (unlike the Security operator) they don't have permissions to respond by taking action. For more information, see Security Reader.	Security Reader

For more information about other Entra roles and user assignment synchronization from Purview for authorizing long-running operations, see the Purview Role Assignment Migrator.

Privileged Identity Management for Groups in the Microsoft Purview portal

Microsoft Entra Privileged Identity Management (PIM) for Groups provides user eligibility for just-in-time membership to groups assigned to Microsoft Entra roles. By using PIM for Groups, you can assign Microsoft Entra security groups to Microsoft Purview role groups in the Microsoft Purview portal. This assignment allows users to activate just-in-time access and then perform tasks in the Microsoft Purview portal. Unlike Microsoft Entra role assignments with Privileged Identity Management (PIM) in Microsoft Entra ID, you can't manage just-in-time activations by using direct user assignments to role groups in the Microsoft Purview portal.

Role precedence and scope behavior

If you assign both an Entra role and a scoped Microsoft Purview role group assignment to a user or group (for example, a role group scoped to an Administrative Unit), the Entra role takes precedence at runtime. As a result, the user's effective permissions are unscoped, even if a scoped Microsoft Purview role group assignment also exists.

- **Example 1**: If you assign the *Compliance Administrator* role in Entra to a user and also assign a scoped *Compliance Administrator* role in Microsoft Purview, the Entra role overrides the scoped assignment. The user's effective access is *unscoped*, and they can access all entities available to that role without Administrative Unit scoping.
- **Example 2**: If you assign the *Global Reader* role in Entra to a user and assign a scoped *DLP Compliance Management* role in Microsoft Purview, any features or APIs that overlap between these roles grant the user *unscoped access* to the corresponding data.
- **Example 3**: When both scoped Microsoft Purview role assignments and Entra roles are present, Entra roles always take precedence, and scoping applied through Administrative Units doesn't restrict the user's effective permissions for overlapping capabilities.

Temporary permissions

You can assign role groups and grant temporary access that automatically revokes permissions when they're no longer needed. Automatically expiring permissions let administrators assign an expiration date to Microsoft Purview role group assignments. When the configured expiration date is reached, the assignment is automatically removed and access is revoked without requiring

manual action from an administrator. Actions that are already completed or in progress aren't affected, but any new operation is denied after the assignment expires. This approach helps organizations implement least-privilege access by reducing long-term access that's no longer needed.

Automatic expiring permissions are optional and are configured on a per-user assignment basis. Administrators can create temporary assignments with an expiration date, update an existing expiration date, extend an assignment, or remove the expiration to make the assignment permanent. A minimum expiration period of one day and maximum expiration period of two years can be selected from the current date in the local time zone of the administrator who sets the expiration.

Role group assignments are evaluated independently. If a user receives the same role group through both an individual assignment and a security group assignment, each assignment retains its own expiration date. Access remains active as long as one of the assignments is still valid. Assigning a role group to a security group also means changes to that group's membership update any added users with the expiration date associated with the assigned role group. The **My Permissions** page displays the latest expiration date across active assignments. The expiration date is recorded as part of the audit logs for Microsoft Purview permissions when the date is added or updated by the admin. You don't receive a notification before permissions expire, nor are any additional audit logs are generated when a user's temporary permissions expires.

> ⓘ **Important**
>
> All built-in and custom role groups except the following support automatic expiring permissions:
>
> - eDiscovery Administrator
> - eDiscovery Manager

Manage role groups

The following sections describe how to add, remove, create, update, and delete role groups in the Microsoft Purview portal.

Add users or groups to a Microsoft Purview built-in role group

Complete the following steps to add users or groups to a Microsoft Purview role group:

1. Sign in to the Microsoft Purview portal by using credentials for an admin account assigned the *Role management* role. Go to **Settings** > **Roles and scopes** to view and manage compliance and governance roles in your organization.
2. Select **Role groups**.
3. On **Role groups for Microsoft Purview solutions**, select a Microsoft Purview role group you want to add users to. Select **Edit** on the control bar.
4. On **Edit members of the role group**, select **Choose users** or **Choose groups**.

   > **Important**
   >
   > Security groups are supported only in Microsoft 365 commercial cloud organizations.

5. Select the check box for all users or groups you want to add to the role group.
6. Select **Select**.
7. If the selected users or groups need organization-wide access as part of this role group assignment, go to step 11.
8. If the selected users or groups need temporary access, select the users or groups and select **Edit expiration**.
9. If the selected users or groups need to be assigned to administrative units, select the users or groups and select **Assign admin units**.
10. On **Assign admin units**, select the check box for all the administrative units you want to assign to the users or groups. Select **Select**.
11. Select **Next** and **Save** to add the users or groups to the role group. Select **Done** to complete the steps.

Remove users or groups from a Microsoft Purview built-in role group

To remove users or groups from a Microsoft Purview role group, complete the following steps:

1. Sign in to the Microsoft Purview portal by using credentials for an admin account assigned the *Role management* role. Go to **Settings** > **Roles and scopes** to view and manage compliance and governance roles in your organization.

2. Select **Role groups**.
3. On **Role groups for Microsoft Purview solutions**, select a Microsoft Purview role group that you want to remove users or groups from. Select **Edit** on the control bar.
4. On **Edit members of the role group**, select the check box for all users or groups you want to remove from the role group.
5. Select **Remove members**, and then select **Next**.
6. Select **Save** to remove the users or groups from the role group. Select **Done** to complete the steps.

Create a custom Microsoft Purview role group

Complete the following steps to create a custom Microsoft Purview role group:

1. Sign in to the Microsoft Purview portal by using credentials for an admin account assigned the *Role management* role. Go to **Settings** > **Roles and scopes** to view and manage compliance and governance roles in your organization.

2. Select **Role groups**.

3. On **Role groups for Microsoft Purview solutions**, select **Create role group**.

4. On **Name the role group**, enter a name for the custom role group in the **Name** field. You can't change the name of the role group after you create it. If needed, enter a description for the custom role group in the **Description** field. Select **Next** to continue.

5. On **Add roles to the role group**, select **Choose roles**.

6. Select the checkboxes for the roles to add to the custom role group. Select **Select**.

7. Select **Next** to continue.

8. On **Add members to the role group**, select **Choose users** (or **Choose groups** if applicable).

   > **Important**
   >
   > Security groups are supported only in Microsoft 365 commercial cloud organizations.

9. Select the checkboxes for the users (or groups) to add to the custom role group. Select **Select**.

10. Select **Next** to continue.

11. If the selected users or groups need organization-wide access as part of this role group assignment, go to step 15.

12. If the selected users or groups need temporary access, select the users or groups and select **Edit expiration**.

13. If the selected users or groups need to be assigned to administrative units, select the users or groups and select **Assign admin units**.

14. On **Assign admin units**, select the check box for all the administrative units you want to assign to the users or groups. Select **Select**.

15. Select **Next**.

16. On **Review the role group and finish**, review the details for the custom role group. If you need to edit the information, select **Edit** in the appropriate section. When all the settings are correct, select **Create** to create the custom role group or select **Cancel** to discard the changes and not create the custom role group.

Update a custom Microsoft Purview role group

Complete the following steps to update a custom Microsoft Purview role group:

1. Sign in to the Microsoft Purview portal by using credentials for an admin account assigned the *Role management* role. Go to **Settings** > **Roles and scopes** to view and manage compliance and governance roles in your organization.
2. Select **Role groups**.
3. On **Role groups for Microsoft Purview solutions**, select a Microsoft Purview role group you want to update, and then select **Edit** on the control bar.
4. On **Name the role group**, update the description for the custom role group in the **Description** field. You can't change the name of the custom role group. Select **Next**.
5. On **Edit roles of the role group**, select **Choose roles** to add roles to update the roles assigned to the role group. You can also select any of the currently assigned roles and select **Remove roles** to remove the roles from the role group. After you update the roles, select **Next**.
6. On **Edit members of the role group**, select **Choose users** or **Choose groups** to add users or groups assigned to the role group.
   a. To change the expiration for the selected users or groups, select any of the currently assigned user or groups and select **Edit expiration**.

b. To update the administrative units for users or groups, select any of the currently assigned user or groups and select **Assign admin units**. You can also select any of the currently assigned users and groups and select **Remove members** to remove the users or groups from the role group. After you update the members, select **Next**.

7. On **Review the role group and finish**, review the details for the custom role group. If you need to edit the information, select **Edit** in the appropriate section. When all the settings are correct, select **Save** to update the custom role group or select **Cancel** to discard the changes and not update the custom role group.

Delete a custom Microsoft Purview role group

Complete the following steps to delete a custom Microsoft Purview role group:

1. Sign in to the Microsoft Purview portal by using credentials for an admin account assigned the *Role management* role. Go to **Settings** > **Roles and scopes** to view and manage compliance and governance roles in your organization.
2. Select **Role groups**.
3. On **Role groups for Microsoft Purview solutions**, select a Microsoft Purview role group you want to delete, then select **Delete** on the control bar.
4. On **Delete role group**, select **Delete** to delete the role group or select **Cancel** to cancel the deletion process.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 07/30/2026

Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview

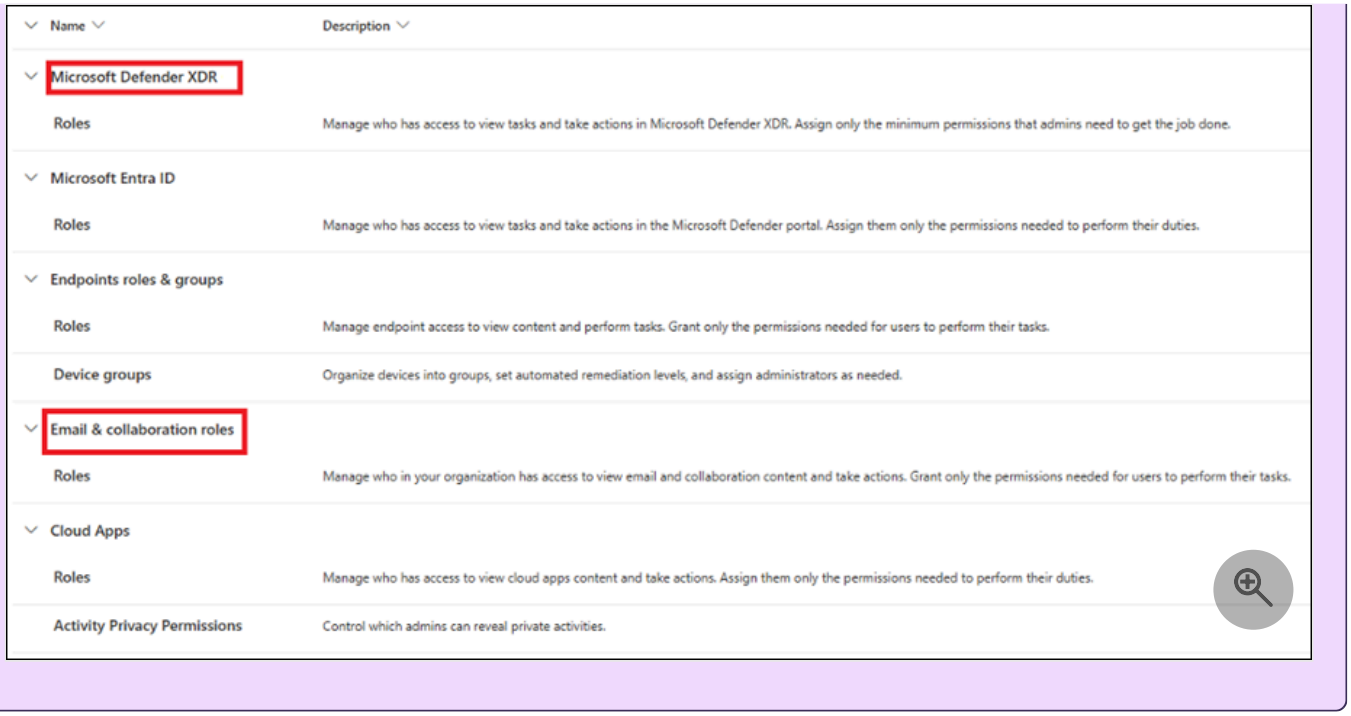
> **Tip**
>
> *Did you know you can try the features in Microsoft Defender for Office 365 Plan 2 for free?* Use the 90-day Defender for Office 365 trial at the Microsoft Defender portal trials hub. Learn about who can sign up and trial terms on Try Microsoft Defender for Office 365.

This article lists the default roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview.

Permissions in these products are based on role-based access control (RBAC). Users need to be a global admin or a member of one or more role groups to access the Microsoft Defender portal or Microsoft Purview portals. RBAC is the same permissions model that's used by Exchange Online, but role groups in Exchange Online don't share membership or permissions with role groups in Defender for Office 365 or Purview.

The Security Dashboard for AI at `ai.security.microsoft.com` is also available to eligible tenants and provides a unified, real-time view of AI-related risk signals from Microsoft Defender, Microsoft Entra, and Microsoft Purview.

> **Note**
>
> Microsoft Defender unified role-based access control (Unified RBAC) is a separate permissions model that's also available. It's also the recommended permission model for Microsoft Defender for Office 365 Plan 2, and it's the default for new Defender for Office 365 Plan 2 organizations starting July 2026. The permissions in this RBAC model are different from the Defender for Office 365 permissions as described in this article. For more information, see Microsoft Defender XDR role-based access control (RBAC).
>
> **If you activate Defender XDR RBAC for Email & collaboration, the permissions page at https://security.microsoft.com/emailandcollabpermissions is no longer available in the Defender portal, so you need to ensure that you configure or import your roles *before* you activate Defender unified RBAC.**



For more information about permissions, see the following articles:

- Email & collaboration permissions in the Microsoft Defender portal
- Microsoft Defender unified role-based access control (RBAC)
- Permissions in the Microsoft Purview portal
- Permissions in Microsoft Purview solutions
- Permissions in the Microsoft Purview governance portal

Role groups in Microsoft Defender for Office 365 and Microsoft Purview

The table in this section lists the default role groups that are available in the Microsoft Defender portal and the Microsoft Purview portals, and the roles that are assigned to the role groups by default. To grant permissions to a user to perform tasks in Defender for Office 365 or Microsoft Purview, add them to the appropriate role group.

Managing permissions in Defender for Office 365 or Microsoft Purview gives users access to security and compliance and governance features that are available within their respective portals. To grant permissions to other features, such as Exchange mail flow rules (also known as transport rules), you need to grant permissions in Exchange Online. For more information, see Permissions in Exchange Online.

> ⓘ **Note**

To view the **Permissions** tab as described in this article, you need to be an admin. Specifically, you need to be assigned the **Role Management** role, and that role is assigned only to the **Organization Management** and **Purview Administrators** role groups by default. The **Role Management** role also allows you to view, create, and modify role groups.

Expand table

Role group	Description	Default roles assigned
AI Administrators	In addition to the capabilities listed for this role in Microsoft Entra built-in roles, use this group to assign read-only permissions to users for Data Security Posture Management for AI (classic) and AI-related data in the preview version of Data Security Posture Management.	AI Administrator
Attack Simulator Administrators	Don't use this role group. Use the Attack Simulation Administrator role in Microsoft Entra ID.	Attack Simulator Admin
Attack Simulator Payload Authors	Don't use this role group. Use the Attack Payload Author role in Microsoft Entra ID.	Attack Simulator Payload Author
Audit Manager	Manage Audit log settings and Search, View, and Export Audit logs.	Audit Logs View-Only Audit Logs
Audit Reader	Search, View, and Export Audit logs.	View-Only Audit Logs
Billing Administrator	Configure Billing features.	Billing Admin
Communication Compliance	Provides permission to all the communication compliance roles: administrator, analyst, investigator, and viewer.	Case Management Communication Compliance Admin Communication Compliance Analysis Communication Compliance Case Management Communication Compliance Investigation Communication Compliance Viewer

Role group	Description	Default roles assigned
		Data Classification Feedback Provider Data Connector Admin Purview Copilot Workspace Contributor Scope Manager View-Only Case
Communication Compliance Administrators	Administrators of communication compliance that can create/edit policies and define global settings.	Communication Compliance Admin Communication Compliance Case Management Data Connector Admin Scope Manager
Communication Compliance Analysts	Analysts of communication compliance that can investigate policy matches, view message meta data, and take remediation actions.	Communication Compliance Analysis Communication Compliance Case Management Purview Copilot Workspace Contributor
Communication Compliance Investigators	Analysts of communication compliance that can investigate policy matches, view message content, and take remediation actions.	Case Management Communication Compliance Analysis Communication Compliance Case Management Communication Compliance Investigation Data Classification

Role group	Description	Default roles assigned
		Feedback Provider Purview Copilot Workspace Contributor View-Only Case
Communication Compliance Viewers	Viewer of communication compliance that can access the available reports and widgets.	Communication Compliance Case Management Communication Compliance Viewer
Compliance Administrator[1]	Members can manage settings for device management, data loss prevention, reports, and preservation.	Admin Unit Extension Manager Case Management Communication Compliance Admin Communication Compliance Case Management Compliance Administrator Compliance Manager Administration Compliance Search Credential Reader Credential Writer Data Classification Feedback Provider Data Classification Feedback Reviewer Data Connector Admin Data Investigation Management

Role group	Description	Default roles assigned
		Data Map Reader Data Security Investigation Admin Device Management Disposition Management DLP Compliance Management Hold IB Compliance Management Information Protection Admin Information Protection Analyst Information Protection Reader Insider Risk Management Admin Insights Reader Manage Alerts Organization Configuration Purview Agent Analysis Purview Agent Deployment Purview Copilot Workspace Contributor RecordManagement

Role group	Description	Default roles assigned
		Retention Management Scan Reader Scan Writer Scope Manager Source Reader Source Writer View-Only Audit Logs View-Only Case View-Only Device Management View-Only DLP Compliance Management View-Only IB Compliance Management View-Only Manage Alerts View-Only Recipients View-Only Record Management View-Only Retention Management
Compliance Data Administrator	Members can manage settings for device management, data protection, data loss prevention, reports, and preservation.	Compliance Administrator Compliance Manager Administration Compliance Search Device Management

Role group	Description	Default roles assigned
		Disposition Management DLP Compliance Management IB Compliance Management Information Protection Admin Information Protection Analyst Information Protection Reader Manage Alerts Organization Configuration Purview Copilot Workspace Contributor RecordManagement Retention Management Scope Manager Sensitivity Label Administrator View-Only Audit Logs View-Only Device Management View-Only DLP Compliance Management View-Only IB Compliance Management

Role group	Description	Default roles assigned
		View-Only Manage Alerts View-Only Recipients View-Only Record Management View-Only Retention Management
Compliance Manager Administrators	Manage template creation and modification.	Compliance Manager Administration Compliance Manager Assessment Compliance Manager Contribution Compliance Manager Reader Data Connector Admin
Compliance Manager Assessors	Create assessments, implement improvement actions, and update test status for improvement actions.	Compliance Manager Assessment Compliance Manager Contribution Compliance Manager Reader Data Connector Admin
Compliance Manager Contributors	Create assessments and perform work to implement improvement actions.	Compliance Manager Contribution Compliance Manager Reader Data Connector Admin
Compliance Manager Readers	View all Compliance Manager content except for administrator functions.	Compliance Manager Reader

Role group	Description	Default roles assigned
Content Explorer Content Viewer	View the contents of files in content explorer, and the prompts and response in Data Security Posture Management for AI (classic) and the preview version of Data Security Posture Management.	Data Classification Content Viewer
Content Explorer List Viewer	View all items in content explorer in list format only.	Data Classification List Viewer
Data Catalog Curators	Perform create, read, modify, and delete actions on catalog data objects and establish relationships between objects.	Data Map Reader Data Map Writer
Data Estate Insights Admins	Provides admin access to all insights reports across platforms and providers.	Data Map Reader Insights Reader Insights Writer
Data Estate Insights Readers	Provides read-only access to all insights reports across platforms and providers.	Data Map Reader Insights Reader
Data Governance	Grants access to data governance roles within Microsoft Purview.	Data Governance Administrator
Data Investigator	Perform searches on mailboxes, SharePoint sites, and OneDrive locations.	Communication Compliance Search Custodian Data Investigation Management Export Preview Review RMS Decrypt Search And Purge
Data Security AI Admins	Use this group to assign editing capabilities for Data Loss Prevention policies and viewing AI content in Data Security Posture Management. Review the role description for access details.	Data Security AI Admin

Role group	Description	Default roles assigned
Data Security AI Content Viewers	Use this group to grant permissions for users to view extended prompt and response details in AI application interactions within Data Security Posture Management for AI (classic) and the preview version of Data Security Posture Management.	Data Security AI Content Viewer
Data Security AI Viewers	Use this group to assign read-only permissions to users for Data Security Posture Management for AI (classic) and the preview version of Data Security Posture Management.	Data Security AI Viewer
Data Security DLP Triage Agent	Intended exclusively for non-interactive agent users to triage and remediate data loss prevention alerts.	Data Classification Content Download Data Classification Content Viewer Information Protection Analyst Purview Content Analyst Purview Copilot Workspace Contributor
Data Security DSPM Posture Agent	Intended exclusively for non-interactive agent users to search and classify the data based on the selected scope provided by Posture management administrators.	Purview Content Analyst Purview Copilot Workspace Contributor
Data Security Investigation Admins	Administrators for Data Security Investigation that can create and manage all investigations, processes, and settings.	Data Security Investigation Admin Data Security Investigation Analyst Data Security Investigation Investigator Data Security Investigation Reviewer
Data Security Investigation Investigators	Investigators for Data Security Investigation that can create and manage assigned investigations, processes, and settings.	Data Security Investigation Analyst Data Security

Role group	Description	Default roles assigned
		Investigation Investigator Data Security Investigation Reviewer
Data Security Investigation Reviewers	Reviewers for Data Security Investigation that can create and manage all assigned investigations.	Data Security Investigation Reviewer
Data Security IRM Triage Agent	Intended exclusively for non-interactive agent users to triage and remediate insider risk management alerts.	Insider Risk Management Analysis Insider Risk Management Investigation Purview Copilot Workspace Contributor
Data Security Management	View all Data Security Posture Management insights, use Security Copilot (including within the Security Dashboard for AI to prioritize and investigate AI risks and view tailored remediation suggestions), and manage Microsoft Purview data security solutions (Data Loss Prevention, Information Protection, and Insider Risk Management).	Case Management Custodian Data Classification Content Download Data Classification Content Viewer Data Classification List Viewer Data Connector Admin Data Map Reader Data Security Investigation Contributor Data Security Viewer Information Protection Admin Information Protection Analyst

Role group	Description	Default roles assigned
		Information Protection Investigator Information Protection Reader Insider Risk Management Admin Insider Risk Management Analysis Insider Risk Management Approval Insider Risk Management Audit Insider Risk Management Graph Reader Insider Risk Management Investigation Insider Risk Management Reports Administrator Insider Risk Management Sessions Insights Reader Purview Agent Analysis Purview Agent Deployment Purview Copilot Workspace Contributor Purview Evaluation Administrator Review

Role group	Description	Default roles assigned
		Scan Reader Source Reader View-Only Case
Data Security Viewers	Read-only access to Data Security Posture Management (DSPM) dashboard insights and the Security Dashboard for AI, including the AI risk scorecard and AI inventory (Microsoft Copilot, Copilot Studio agents, Microsoft Foundry applications and agents, and third-party models and agents). Members can view unmanaged and shadow AI agents and use Security Copilot to view detailed information.	Data Security Viewer
Data Source Administrators	Manage data sources and data scans.	Credential Reader Credential Writer Scan Reader Scan Writer Source Reader Source Writer
eDiscovery Manager	Members can perform searches and place holds on mailboxes, SharePoint sites, and OneDrive locations. Members can also create and manage eDiscovery cases, add and remove members to a case, create and edit Content Searches associated with a case, and access case data in eDiscovery (Premium). An eDiscovery Administrator is a member of the eDiscovery Manager role group who has been assigned additional permissions. In addition to the tasks that an eDiscovery Manager can perform, an eDiscovery Administrator can: • View all eDiscovery cases in the organization. • Manage any eDiscovery case after they add themselves as a member of the case. The primary difference between an eDiscovery Manager and an eDiscovery Administrator is that an eDiscovery	Case Management Communication Compliance Search Custodian Export Hold Manage Review Set Tags Preview Review RMS Decrypt

Role group	Description	Default roles assigned
	Administrator can access all cases that are listed on the **eDiscovery cases** page in the Microsoft Purview portal. An eDiscovery manager can only access the cases they created or cases they're a member of. For more information about making a user an eDiscovery Administrator, see Assign eDiscovery permissions in the Microsoft Purview portal.	
Exact Data Match Upload Admins	Upload data for Exact Data Match.	Exact Data Match Upload Admin
Global Reader	Members have read-only access to reports, alerts, and can see all the configuration and settings. The primary difference between Global Reader and Security Reader is that a Global Reader can access **configuration and settings**.	Compliance Manager Reader Security Reader Sensitivity Label Reader Service Assurance View View-Only Audit Logs View-Only Device Management View-Only DLP Compliance Management View-Only IB Compliance Management View-Only Manage Alerts View-Only Recipients View-Only Record Management View-Only Retention Management
Information Protection	Full control over all information protection features, including sensitivity labels and their policies, DLP, all classifier types, activity and content explorers, and all related reports.	Data Classification Content Download Data Classification Content Viewer

Role group	Description	Default roles assigned
		Data Classification List Viewer Data Map Reader Information Protection Admin Information Protection Analyst Information Protection Investigator Information Protection Reader Insights Reader Purview Agent Analysis Purview Agent Deployment Purview Copilot Workspace Contributor Purview Evaluation Administrator Scan Reader Source Reader
Information Protection ABAC Attribute Assignment Administrators[2]	Assign and remove M365 Info Protect Attributes for supported objects such as users and service principals.	Information Protection ABAC Attribute Assignment Administrator Information Protection ABAC Attribute Definition Reader
Information Protection ABAC Attribute	Read M365 Info Protect Attributes for supported objects.	Information Protection ABAC Attribute Assignment Reader

Role group	Description	Default roles assigned
Assignment Readers[2]		
Information Protection ABAC Attribute Definition Administrators[2]	Define a valid set of M365 Info Protect Attributes that can be assigned to supported objects. Activate and deactivate these attributes.	Information Protection ABAC Attribute Definition Administrator
Information Protection ABAC Attribute Definition Readers[2]	Read the definitions of M365 Info Protect Attributes.	Information Protection ABAC Attribute Definition Reader
Information Protection ABAC Policy Administrators[2]	Define M365 Info Protect ABAC Policies and Rules to control M365 ABAC Mandatory Access Control decisions. Create, view, edit, delete, enable, and disable those policies and rules.	Information Protection ABAC Policy Administrator Information Protection ABAC Attribute Definition Reader
Information Protection ABAC Policy Readers[2]	View M365 Info Protect ABAC Policies and Rules.	Information Protection ABAC Policy Reader
Information Protection Admins	Create, edit, and delete DLP policies, sensitivity labels and their policies, and all classifier types. Manage endpoint DLP settings and simulation mode for auto-labeling policies.	Data Map Reader Information Protection Admin Insights Reader Purview Copilot Workspace Contributor Purview Evaluation Administrator Scan Reader Source Reader
Information Protection Analysts	Access and manage DLP alerts and activity explorer. View-only access to DLP policies, sensitivity labels and their policies, and all classifier types.	Data Classification List Viewer Data Map Reader Information Protection Analyst

Role group	Description	Default roles assigned
		Insights Reader Purview Agent Analysis Purview Agent Deployment Purview Copilot Workspace Contributor Purview Evaluation Administrator
Information Protection Investigators	Access and manage DLP alerts, activity explorer, and content explorer. View-only access to DLP policies, sensitivity labels and their policies, and all classifier types.	Data Classification Content Download Data Classification Content Viewer Data Classification List Viewer Data Map Reader Information Protection Analyst Information Protection Investigator Insights Reader Purview Agent Analysis Purview Agent Deployment Purview Copilot Workspace Contributor Purview Evaluation Administrator Scan Reader Source Reader

Role group	Description	Default roles assigned
Information Protection Readers	View-only access to reports for DLP policies and sensitivity labels and their policies.	Information Protection Reader
Insider Risk Management	Use this role group to manage insider risk management for your organization in a single group. By adding all user accounts for designated administrators, analysts, and investigators, you can configure insider risk management permissions in a single group. This role group contains all the insider risk management permission roles. This role group is the easiest way to quickly get started with insider risk management and is a good fit for organizations that don't need separate permissions defined for separate groups of users.	Case Management Custodian Data Connector Admin Data Security Investigation Contributor Insider Risk Management Admin Insider Risk Management Analysis Insider Risk Management Approval Insider Risk Management Audit Insider Risk Management Graph Reader Insider Risk Management Investigation Insider Risk Management Reports Administrator Insider Risk Management Sessions Purview Agent Analysis Purview Agent Deployment Purview Copilot Workspace Contributor

Role group	Description	Default roles assigned
		Review View-Only Case
Insider Risk Management Admins	Use this role group to initially configure insider risk management and later to segregate insider risk administrators into a defined group. Users in this role group can create, read, update, and delete insider risk management policies, global settings, and role group assignments.	Case Management Data Connector Admin Insider Risk Management Admin Purview Copilot Workspace Contributor View-Only Case
Insider Risk Management Analysts	Use this group to assign permissions to users that act as insider risk case analysts. Users in this role group can access all insider risk management alerts, cases, and notices templates. They can't access the insider risk Content Explorer.	Case Management Insider Risk Management Analysis Insider Risk Management Graph Reader Purview Agent Analysis Purview Agent Deployment Purview Copilot Workspace Contributor View-Only Case
Insider Risk Management Approvers	For internal approval use only.	Insider Risk Management Approval Purview Copilot Workspace Contributor
Insider Risk Management Auditors	Use this group to assign permissions to users that audit insider risk management activities. Users in this role group can access the insider risk audit log.	Insider Risk Management Audit Purview Copilot Workspace Contributor

Role group	Description	Default roles assigned
Insider Risk Management Investigators	Use this group to assign permissions to users that act as insider risk data investigators. Users in this role group can access all insider risk management alerts, cases, notices templates, and the Content Explorer for all cases.	Case Management Custodian Data Security Investigation Contributor Insider Risk Management Graph Reader Insider Risk Management Investigation Purview Agent Analysis Purview Agent Deployment Purview Copilot Workspace Contributor Review View-Only Case
Insider Risk Management Session Approvers	For internal approval use only.	Insider Risk Management Sessions
IRM Contributors	This role group is visible, but is used by background services only.	Insider Risk Management Permanent contribution Insider Risk Management Temporary contribution
Knowledge Administrators	Configure knowledge, learning, assign trainings and other intelligent features.	Knowledge Admin
MailFlow Administrator	Members can monitor and view mail flow insights and reports in the Defender portal. Global admins can add ordinary users to this group, but, if the user isn't a member of the Exchange Admin group, the user doesn't have access to Exchange admin-related tasks.	Exchange Administrator View-Only Recipients

Role group	Description	Default roles assigned
Organization Management[1]	Members can control permissions for accessing features in these portals, and also manage settings for device management, data loss prevention, reports, and preservation. Users who aren't global administrators must be Exchange administrators to see and take action on devices that are managed by Basic Mobility and Security for Microsoft 365 (formerly known as Mobile Device Management or MDM). Global admins are automatically added as members of this role group, but you don't see them in the output of the Get-RoleGroupMember cmdlet in Security & Compliance PowerShell. **Important**: Microsoft strongly advocates for the principle of least privilege. Assigning accounts only the minimum permissions necessary to perform their tasks helps reduce security risks and strengthens your organization's overall protection. Global Administrator is a highly privileged role that you should limit to emergency scenarios or when you can't use a different role.	Admin Unit Extension Manager Audit Logs Case Management Communication Compliance Admin Communication Compliance Case Management Compliance Administrator Compliance Manager Administration Compliance Search Data Connector Admin Data Security Investigation Admin Data Security Investigation Contributor Device Management DLP Compliance Management Hold IB Compliance Management Insider Risk Management Admin License Usage Reader Manage Alerts

Role group	Description	Default roles assigned
		Organization Configuration Priority Cleanup Admin Priority Cleanup Viewer Purview Consumption Admin Purview Copilot Workspace Contributor Quarantine RecordManagement Retention Management Role Management Scope Manager Search And Purge Security Administrator Security Reader Sensitivity Label Administrator Sensitivity Label Reader Service Assurance View Tag Contributor Tag Manager Tag Reader View-Only Audit Logs View-Only Case View-Only Device

Role group	Description	Default roles assigned
		Management View-Only DLP Compliance Management View-Only IB Compliance Management View-Only Manage Alerts View-Only Recipients View-Only Record Management View-Only Retention Management
Privacy Management	Manage access control for Privacy Management solution in the Microsoft Purview portal.	Case Management Compliance Manager Contribution Compliance Manager Reader Compliance Search Custodian Data Classification Content Viewer Data Classification List Viewer Data Map Reader Export Insights Reader Manage Review Set Tags

Role group	Description	Default roles assigned
		Preview Privacy Management Admin Privacy Management Analysis Privacy Management Investigation Privacy Management Permanent contribution Privacy Management Temporary contribution Privacy Management Viewer Review Source Reader Subject Rights Request Admin View-Only Case
Privacy Management Administrators	Administrators of privacy management solution that can create/edit policies and define global settings.	Case Management Compliance Manager Contribution Compliance Manager Reader Compliance Search Custodian Data Map Reader Export Insights Reader Manage Review Set

Role group	Description	Default roles assigned
		Tags Preview Privacy Management Admin Review Source Reader View-Only Case
Privacy Management Analysts	Analysts of privacy management solution that can investigate policy matches, view messages meta data, and take remediation actions.	Case Management Compliance Manager Reader Compliance Search Custodian Data Classification List Viewer Data Map Reader Export Insights Reader Manage Review Set Tags Preview Privacy Management Analysis Review View-Only Case
Privacy Management Contributors	Manage contributor access for privacy management cases.	Compliance Manager Reader Compliance Search

Role group	Description	Default roles assigned
		Custodian Data Map Reader Export Manage Review Set Tags Preview Privacy Management Permanent contribution Privacy Management Temporary contribution Review
Privacy Management Investigators	Investigators of privacy management solution that can investigate policy matches, view message content, and take remediation actions.	Case Management Compliance Manager Reader Compliance Search Custodian Data Classification Content Viewer Data Classification List Viewer Data Map Reader Export Insights Reader Manage Review Set Tags Preview Privacy Management Investigation

Role group	Description	Default roles assigned
		Review View-Only Case
Privacy Management Viewers	Viewer of privacy management solution that can access the available dashboards and widgets.	Compliance Manager Reader Data Classification List Viewer Data Map Reader Export Insights Reader Preview Privacy Management Viewer Review
Purview Administrators	Create, edit, and delete domains and perform role assignments.	Admin Unit Extension Manager Purview Domain Manager Role Management
Purview Agent Management	To enable, setup and manage the AI agents within Purview	Purview Agent Deployment Purview Content Analyst
Purview Consumption Management	Assign Purview Consumption Billing reports access capabilities.	Purview Consumption Admin
Quarantine Administrator	Members can access all Quarantine actions. For more information, see Manage quarantined messages and files as an admin	Quarantine
Records Management	Members can configure all aspects of records management, including retention labels and disposition reviews.	Disposition Management RecordManagement

Role group	Description	Default roles assigned
		Retention Management Scope Manager
Reviewer	Members can access review sets in eDiscovery (Premium) cases. Members of this role group can see and open the list of cases on the **eDiscovery** > **Advanced** page in the Microsoft Purview portal that they're members of. After the user accesses an eDiscovery (Premium) case, they can select **Review sets** to access case data. This role doesn't allow the user to preview the results of a collection search that's associated with the case or do other search or case management tasks. Members of this role group can only access the data in a review set.	Review
Security Administrator	Members have access to many security features of Identity Protection Center, Privileged Identity Management, Monitor Microsoft 365 Service Health, and the Defender and compliance portals. By default, this role group may not appear to have any members. However, the Security Administrator role from Microsoft Entra ID is assigned to this role group. Therefore, this role group inherits the capabilities and membership of the Security Administrator role from Microsoft Entra ID. To manage permissions centrally, add and remove group members in the Microsoft Entra admin center. For more information, see Microsoft Entra built-in roles. If you edit this role group in these portals (membership or roles), those changes apply only to the security and compliance areas and not to any other services. This role group includes all of the read-only permissions of the Security reader role, plus many additional administrative permissions for the same services: Azure Information Protection, Identity Protection Center, Privileged Identity Management, Monitor Microsoft 365 Service Health, and the Defender and compliance portals.	Audit Logs Compliance Manager Administration Data Security Investigation Contributor Device Management DLP Compliance Management IB Compliance Management Manage Alerts Purview Copilot Workspace Contributor Quarantine Security Administrator Sensitivity Label Administrator Tag Contributor Tag Manager

Role group	Description	Default roles assigned
		Tag Reader View-Only Audit Logs View-Only Device Management View-Only DLP Compliance Management View-Only IB Compliance Management View-Only Manage Alerts
Security Operator	Members can manage security alerts, and also view reports and settings of security features.	Compliance Search Data Security Investigation Contributor Manage Alerts Purview Copilot Workspace Contributor Security Reader Tag Contributor Tag Reader Tenant AllowBlockList Manager View-Only Audit Logs View-Only Device Management View-Only DLP Compliance Management

Role group	Description	Default roles assigned
		View-Only IB Compliance Management View-Only Manage Alerts
Security Reader	Members have read-only access to many security features of Identity Protection Center, Privileged Identity Management, Monitor Microsoft 365 Service Health, and the Defender and compliance portals. By default, this role group may not appear to have any members. However, the Security Reader role from Microsoft Entra ID is assigned to this role group. Therefore, this role group inherits the capabilities and membership of the Security Reader role from Microsoft Entra ID. To manage permissions centrally, add and remove group members in the Microsoft Entra admin center. For more information, see Microsoft Entra built-in roles. If you edit this role group in the portals (membership or roles), those changes apply only to security and compliance areas and not to any other services.	Compliance Manager Reader Security Reader Sensitivity Label Reader Tag Reader View-Only Device Management View-Only DLP Compliance Management View-Only IB Compliance Management View-Only Manage Alerts
Service Assurance User	Members can access the Service assurance section in the Microsoft Purview portal. Service assurance provides reports and documents that describe Microsoft's security practices for customer data that's stored in Microsoft 365. It also provides independent non-Microsoft audit reports on Microsoft 365. For more information, see Service assurance in the Microsoft Purview portal.	Service Assurance View
Subject Rights Request Administrators	Create subject rights requests.	Case Management Compliance Manager Contribution Compliance Manager Reader Compliance Search

Role group	Description	Default roles assigned
		Custodian Export Manage Review Set Tags Preview Review Subject Rights Request Admin View-Only Case
Subject Rights Request Approvers	Approvers who are able to approve subject rights requests.	Compliance Manager Reader Compliance Search Custodian Export Manage Review Set Tags Preview Review Subject Rights Request Approver
Supervisory Review	Members can create and manage the policies that define which communications are subject to review in an organization. For more information, see Configure communication compliance policies for your organization.	Supervisory Review Administrator

Note

[1] This role group doesn't assign members the permissions necessary to search the audit log or to use any reports that might include Exchange data, such as the DLP or Defender for Office 365 reports. To search the audit log or to view all reports, a user has to be assigned permissions in Exchange Online. This action is required because the underlying cmdlet that's

used to search the audit log is an Exchange Online cmdlet. Global admins can search the audit log and view all reports because they're automatically added as members of the Organization Management role group in Exchange Online. For more information, see Search the audit log in the Microsoft Purview portal.

> **Note**
>
> [2] This role or role group is in preview. While it may be visible in commercial Microsoft 365 tenants, it is not supported and has no operational effect outside the M365 ABAC private preview. At this time, support is planned only for Microsoft U.S. government cloud environments (GCC High and DoD, see Understand Microsoft U.S. government cloud environments for Microsoft 365 and Microsoft Copilot).

Roles in Microsoft Defender for Office 365 and Microsoft Purview

The table in this section lists the available roles and the role groups that they're assigned to by default.

Roles that aren't assigned to the Organization Management role group by default are marked with *

Expand table

Role	Description	Default role group assignments
Admin Unit Extension Manager	Assigns permissions to use AUE.	Compliance Administrator Organization Management Purview Administrators
***AI Administrator**	In addition to the capabilities listed for this role in Microsoft Entra built-in roles, use this role for read-only access to all pages in Data Security Posture Management for AI (classic) and for read-only access to AI-related data in the preview version of Data Security Posture	AI Administrators

Role	Description	Default role group assignments
	Management. This role does not have access to read prompts and responses of AI interactions.	
*Attack Simulator Admin	Don't use this role. Use the Attack Simulation Administrator role in Microsoft Entra ID.	Attack Simulator Administrators
*Attack Simulator Payload Author	Don't use this role in the portals. Use the corresponding role in Microsoft Entra ID.	Attack Simulator Payload Authors
Audit Logs	Turn on and configure auditing for the organization, view the organization's audit reports, and then export these reports to a file.	Audit Manager Organization Management Security Administrator
*Billing Admin	Allows billing admin for selected feature.	Billing Administrator
Case Management	Create, edit, delete, and control access to eDiscovery cases.	Communication Compliance Communication Compliance Investigators Compliance Administrator Data Security Management eDiscovery Manager Insider Risk Management Insider Risk Management Admins Insider Risk Management Analysts Insider Risk

Role	Description	Default role group assignments
		Management Investigators Organization Management Privacy Management Privacy Management Administrators Privacy Management Analysts Privacy Management Investigators Subject Rights Request Administrators
*Communication	Manage all communications with the custodians identified in an eDiscovery (Premium) case. Create hold notifications, hold reminders, and escalations to management. Track custodian acknowledgment of hold notifications and manage access to the custodian portal that's used by each custodian in a case to track communications for the cases where they were identified as a custodian.	Data Investigator eDiscovery Manager
Communication Compliance Admin	Used to manage policies in the Communication Compliance feature.	Communication Compliance Communication Compliance Administrators Compliance Administrator Organization Management

Role	Description	Default role group assignments
*Communication Compliance Analysis	Used to perform investigation, remediation of the message violations in the Communication Compliance feature. Can only view message meta data.	Communication Compliance Communication Compliance Analysts Communication Compliance Investigators
Communication Compliance Case Management	Used to access Communication Compliance cases.	Communication Compliance Communication Compliance Administrators Communication Compliance Analysts Communication Compliance Investigators Communication Compliance Viewers Compliance Administrator Organization Management
*Communication Compliance Investigation	Used to perform investigation, remediation, and review message violations in the Communication Compliance feature. Can view message meta data and message.	Communication Compliance Communication Compliance Investigators
*Communication Compliance Viewer	Used to access reports and widgets in the Communication Compliance feature.	Communication Compliance Communication

Role	Description	Default role group assignments
		Compliance Viewers
Compliance Administrator	View and edit settings and reports for compliance features.	Compliance Administrator Compliance Data Administrator Organization Management
Compliance Manager Administration	Manage template creation and modification.	Compliance Administrator Compliance Data Administrator Compliance Manager Administrators Organization Management Security Administrator
*__Compliance Manager Assessment__	Create assessments, implement improvement actions, and update test status for improvement actions.	Compliance Manager Administrators Compliance Manager Assessors
*__Compliance Manager Contribution__	Create assessments and perform work to implement improvement actions.	Compliance Manager Administrators Compliance Manager Assessors Compliance Manager Contributors Privacy

Role	Description	Default role group assignments
		Management Privacy Management Administrators Subject Rights Request Administrators
*Compliance Manager Reader	View all Compliance Manager content except for administrator functions.	Compliance Manager Administrators Compliance Manager Assessors Compliance Manager Contributors Compliance Manager Readers Global Reader Privacy Management Privacy Management Administrators Privacy Management Analysts Privacy Management Contributors Privacy Management Investigators Privacy Management

Role	Description	Default role group assignments
		Viewers Security Reader Subject Rights Request Administrators Subject Rights Request Approvers
Compliance Search	Perform searches across mailboxes and get an estimate of the results.	Compliance Administrator Compliance Data Administrator Data Investigator eDiscovery Manager Organization Management Privacy Management Privacy Management Administrators Privacy Management Analysts Privacy Management Contributors Privacy Management Investigators Security Operator Subject Rights

Role	Description	Default role group assignments
		Request Administrators Subject Rights Request Approvers
*Credential Reader	Read the different credentials created in the tenant.	Compliance Administrator Data Source Administrators
*Credential Writer	Create and edit credentials.	Compliance Administrator Data Source Administrators
*Custodian	Identify and manage custodians for eDiscovery (Premium) cases and use the information from Microsoft Entra ID and other sources to find data sources associated with custodians. Associate other data sources such as mailboxes, SharePoint sites, and Teams with custodians in a case. Place a legal hold on the data sources associated with custodians to preserve content in the context of a case.	Data Investigator Data Security Management eDiscovery Manager Insider Risk Management Insider Risk Management Investigators Privacy Management Privacy Management Administrators Privacy Management Analysts Privacy Management Contributors

Role	Description	Default role group assignments
		Privacy Management Investigators Subject Rights Request Administrators Subject Rights Request Approvers
*Data Classification Content Download	When evidence collection is turned on from Endpoint DLP settings, this role lets admins download endpoint-related evidence files from activity explorer and DLP alerts.	Data Security DLP Triage Agent Data Security Management Information Protection Information Protection Investigators
*Data Classification Content Viewer	View in-place rendering of files in Content explorer.	Content Explorer Content Viewer Data Security DLP Triage Agent Data Security Management Information Protection Information Protection Investigators Privacy Management Privacy Management Investigators

Role	Description	Default role group assignments
*Data Classification Feedback Provider	Allows providing feedback to classifiers in content explorer.	Communication Compliance Communication Compliance Investigators Compliance Administrator
*Data Classification Feedback Reviewer	Allows reviewing feedback from classifiers in feedback explorer.	Compliance Administrator
*Data Classification List Viewer	View the list of files in content explorer.	Content Explorer List Viewer Data Security Management Information Protection Information Protection Analysts Information Protection Investigators Privacy Management Privacy Management Analysts Privacy Management Investigators Privacy Management Viewers
Data Connector Admin	Create and manage connectors to import and archive non-Microsoft data in Microsoft 365.	Communication Compliance

Role	Description	Default role group assignments
		Communication Compliance Administrators Compliance Administrator Compliance Manager Administrators Compliance Manager Assessors Compliance Manager Contributors Data Security Management Insider Risk Management Insider Risk Management Admins Organization Management
*Data Governance Administrator	Delegates the first level of access for business domain creators and other application-level permissions.	Data Governance
*Data Investigation Management	Create, edit, delete, and control access to data investigation.	Compliance Administrator Data Investigator
*Data Map Reader	Read actions on data map objects.	Compliance Administrator Data Catalog Curators Data Estate Insights Admins

Role	Description	Default role group assignments
		Data Estate Insights Readers Data Security Management Information Protection Information Protection Admins Information Protection Analysts Information Protection Investigators Privacy Management Privacy Management Administrators Privacy Management Analysts Privacy Management Contributors Privacy Management Investigators Privacy Management Viewers
*Data Map Writer	Create, read, modify, and delete actions on data map objects and establish relationships between objects.	Data Catalog Curators

Role	Description	Default role group assignments
*Data Security AI Admin	Role for editing Data Loss Prevention policies and viewing AI content in Data Security Posture Management. This role does not have access to read prompts and responses of AI interactions.	Data Security AI Admins
*Data Security AI Content Viewer	Role for read-only access to prompts and responses of AI interactions in Data Security Posture Management for AI (classic) and the preview version of Data Security Posture Management.	Data Security AI Content Viewers
*Data Security AI Viewer	Role for read-only access to all pages in Data Security Posture Management for AI (classic) and AI-related data in the preview version of Data Security Posture Management. This role does not have access to read prompts and responses of AI interactions.	Data Security AI Viewers
*Data Security Investigation Admin	Used to create and manage investigations, processes, and settings in Data Security Investigation.	Compliance Administrator Data Security Investigation Admins Organization Management
*Data Security Investigation Analyst	All of the capabilities within the investigations they own or are members of. Includes mitigation actions to reduce or contain data security risks such as purging content.	Data Security Investigation Admins Data Security Investigation Investigators
Data Security Investigation Contributor	Create investigations and list or get the investigations that they own or are members of.	Data Security Management Insider Risk Management Insider Risk Management Investigators Organization Management

Role	Description	Default role group assignments
		Security Administrator Security Operator
*Data Security Investigation Investigator	Used to create and manage assigned investigations, processes, and settings in Data Security Investigation.	Data Security Investigation Admins Data Security Investigation Investigators
*Data Security Investigation Reviewer	Used to review assigned investigations in Data Security Investigation.	Data Security Investigation Admins Data Security Investigation Investigators Data Security Investigation Reviewers
*Data Security Viewer	Read access to Data Security Posture Management insights and the Security Dashboard for AI (AI risk scorecard, AI inventory for Microsoft Copilot, Copilot Studio, Microsoft Foundry, and third-party models and agents). Allows users to use Security Copilot to view details and run prompts. Viewing prompts and responses requires the Data Security AI Content Viewer role.	Data Security Management Data Security Viewers
Device Management	View and edit settings and reports for device management features.	Compliance Administrator Compliance Data Administrator Organization Management Security Administrator
*Disposition Management	Control permissions for accessing Manual Disposition in the Defender and compliance portals.	Compliance Administrator

Role	Description	Default role group assignments
		Compliance Data Administrator Records Management
DLP Compliance Management	View and edit settings and reports for data loss prevention (DLP) policies.	Compliance Administrator Compliance Data Administrator Organization Management Security Administrator
*Exact Data Match Upload Admin	Lets users upload data for Exact Data Match.	Exact Data Match Upload Admins
*Exchange Administrator	Allows Exchange administrator for selected features.	MailFlow Administrator
*Export	Export mailbox and site content that's returned from searches.	Data Investigator eDiscovery Manager Privacy Management Privacy Management Administrators Privacy Management Analysts Privacy Management Contributors Privacy Management Investigators

Role	Description	Default role group assignments
		Privacy Management Viewers Subject Rights Request Administrators Subject Rights Request Approvers
Hold	Place content in mailboxes, sites, and public folders on hold. When on hold, a copy of the content is stored in a secure location. Content owners are still able to modify or delete the original content.	Compliance Administrator eDiscovery Manager Organization Management
IB Compliance Management	View, create, remove, modify, and test Information Barrier policies.	Compliance Administrator Compliance Data Administrator Organization Management Security Administrator
*__Information Protection ABAC Attribute Assignment Administrator__[2]	Assign and remove M365 Info Protect Attributes for supported objects such as users and service principals.	Information Protection ABAC Attribute Assignment Administrators
*__Information Protection ABAC Attribute Assignment Reader__[2]	Read M365 Info Protect Attributes for supported objects.	Information Protection ABAC Attribute Assignment Readers

Role	Description	Default role group assignments
*Information Protection ABAC Attribute Definition Administrator[2]	Define a valid set of M365 Info Protect Attributes that can be assigned to supported objects. Activate and deactivate these attributes.	Information Protection ABAC Attribute Definition Administrators
*Information Protection ABAC Attribute Definition Reader[2]	Read the definitions of M365 Info Protect Attributes.	Information Protection ABAC Attribute Assignment Administrators Information Protection ABAC Attribute Definition Readers Information Protection ABAC Policy Administrators
*Information Protection ABAC Policy Administrator[2]	Define M365 Info Protect ABAC Policies and Rules to control M365 ABAC Mandatory Access Control decisions. Create, view, edit, delete, enable, and disable those policies and rules.	Information Protection ABAC Policy Administrators
*Information Protection ABAC Policy Reader[2]	View M365 Info Protect ABAC Policies and Rules.	Information Protection ABAC Policy Readers
*Information Protection Admin	Create, edit, and delete DLP policies, sensitivity labels and their policies, and all classifier types. Manage endpoint DLP settings and simulation mode for auto-labeling policies.	Compliance Administrator Compliance Data Administrator Data Security Management Information Protection Information Protection Admins
*Information Protection Analyst	Access and manage DLP alerts and activity explorer. View-only access to DLP policies, sensitivity labels and their	Compliance Administrator

Role	Description	Default role group assignments
	policies, and all classifier types.	Compliance Data Administrator Data Security DLP Triage Agent Data Security Management Information Protection Information Protection Analysts Information Protection Investigators
*Information Protection Investigator	Access and manage DLP alerts, activity explorer, and content explorer. View-only access to DLP policies, sensitivity labels and their policies, and all classifier types.	Data Security Management Information Protection Information Protection Investigators
*Information Protection Reader	View-only access to reports for DLP policies and sensitivity labels and their policies.	Compliance Administrator Compliance Data Administrator Data Security Management Information Protection Information Protection Readers
Insider Risk Management Admin	Create, edit, delete, and control access to Insider Risk Management feature.	Compliance Administrator

Role	Description	Default role group assignments
		Data Security Management Insider Risk Management Insider Risk Management Admins Organization Management
*Insider Risk Management Analysis	Access all insider risk management alerts, cases, and notices templates.	Data Security IRM Triage Agent Data Security Management Insider Risk Management Insider Risk Management Analysts
*Insider Risk Management Approval	Perform investigation, remediation, and review message violations in Privacy Management solution. Can view message metadata and full messages.	Data Security Management Insider Risk Management Insider Risk Management Approvers
*Insider Risk Management Audit	Allow viewing Insider Risk audit trails.	Data Security Management Insider Risk Management Insider Risk Management Auditors

Role	Description	Default role group assignments
*Insider Risk Management Graph Reader	Allow viewing to data risk graphs for alerts.	Data Security Management Insider Risk Management Insider Risk Management Analysts Insider Risk Management Investigators
*Insider Risk Management Investigation	Access all insider risk management alerts, cases, notices templates, and the Content Explorer for all cases.	Data Security IRM Triage Agent Data Security Management Insider Risk Management Insider Risk Management Investigators
*Insider Risk Management Permanent contribution	This role group is visible, but is used by background services only.	IRM Contributors
*Insider Risk Management Reports Administrator	Allows viewing Insider Risk reports related to cases and events for Insider Risk Management.	Data Security Management Insider Risk Management
*Insider Risk Management Sessions	Allows for controlled approval and oversight of user session-based activities within Microsoft Purview Insider Risk Management, without granting access to investigations, alerts, cases, or sensitive content.	Data Security Management Insider Risk Management Insider Risk Management Session Approvers

Role	Description	Default role group assignments
*Insider Risk Management Temporary contribution	This role group is visible, but is used by background services only.	IRM Contributors
*Insights Reader	Provides read-only access to all Insights reports in the Data Estate Insights app. Insights readers need to have at least data reader role access to a collection to view reports about that specific collection.	Compliance Administrator Data Estate Insights Admins Data Estate Insights Readers Data Security Management Information Protection Information Protection Admins Information Protection Analysts Information Protection Investigators Privacy Management Privacy Management Administrators Privacy Management Analysts Privacy Management Investigators Privacy

Role	Description	Default role group assignments
		Management Viewers
*Insights Writer	Admin access to all insights reports in the Data Estate Insights app.	Data Estate Insights Admins
*Knowledge Admin	Configure knowledge, learning, assign trainings and other intelligent features.	Knowledge Administrators
License Usage Reader	Access to view the License Utilization Reporting Dashboard for Compliance Solutions.	Organization Management
Manage Alerts	View and edit settings and reports for alerts.	Compliance Administrator Compliance Data Administrator Organization Management Security Administrator Security Operator
*Manage Review Set Tags	This role lets users create, edit, and delete review set tags for cases they can access.	eDiscovery Manager Privacy Management Privacy Management Administrators Privacy Management Analysts Privacy Management Contributors Privacy Management Investigators

Role	Description	Default role group assignments
		Subject Rights Request Administrators Subject Rights Request Approvers
Organization Configuration	Run, view, and export audit reports and manage compliance policies for DLP, devices, and preservation.	Compliance Administrator Compliance Data Administrator Organization Management
*__Preview__	View a list of items that are returned from content searches, and open each item from the list to view its contents.	Data Investigator eDiscovery Manager Privacy Management Privacy Management Administrators Privacy Management Analysts Privacy Management Contributors Privacy Management Investigators Privacy Management Viewers Subject Rights Request Administrators

Role	Description	Default role group assignments
		Subject Rights Request Approvers
Priority Cleanup Admin	Access Priority Cleanup tab within Data Lifecycle Management to create, update, and delete policies and modify Priority Cleanup settings for the tenant	Organization Management
Priority Cleanup Viewer	Access Priority Cleanup tab within Data Lifecycle Management to view policies	Organization Management
*Privacy Management Admin	Manage policies in Privacy Management and has access to all functionality of the solution.	Privacy Management Privacy Management Administrators
*Privacy Management Analysis	Perform investigation and remediation of the message violations in Privacy Management. Can only view messages metadata.	Privacy Management Privacy Management Analysts
*Privacy Management Investigation	Perform investigation, remediation, and review message violations in Privacy Management. Can view message metadata and the full message.	Privacy Management Privacy Management Investigators
*Privacy Management Permanent contribution	Access Privacy Management cases as a permanent contributor.	Privacy Management Privacy Management Contributors
*Privacy Management Temporary contribution	Access Privacy Management cases as a temporary contributor.	Privacy Management Privacy Management Contributors
*Privacy Management Viewer	Access dashboards and widgets in Privacy Management.	Privacy Management

Role	Description	Default role group assignments
		Privacy Management Viewers
*Purview Agent Analysis	Configure, Customize and View Purview agent data and analysis	Compliance Administrator Data Security Management Information Protection Information Protection Analysts Information Protection Investigators Insider Risk Management Insider Risk Management Analysts Insider Risk Management Investigators
*Purview Agent Deployment	Deploy Purview Agents	Compliance Administrator Data Security Management Information Protection Information Protection Analysts Information Protection Investigators

Role	Description	Default role group assignments
		Insider Risk Management Insider Risk Management Analysts Insider Risk Management Investigators Purview Agent Management
Purview Consumption Admin	Access the Purview consumption billing transparency report	Organization Management Purview Consumption Management
***Purview Content Analyst**	Enable Purview Agents	Data Security DLP Triage Agent Data Security DSPM Posture Agent Purview Agent Management
Purview Copilot Workspace Contributor	Access Purview Copilot Workspace as a Contributor.	Communication Compliance Communication Compliance Analysts Communication Compliance Investigators Compliance Administrator Compliance Data

Role	Description	Default role group assignments
		Administrator Data Security DLP Triage Agent Data Security DSPM Posture Agent Data Security IRM Triage Agent Data Security Management Information Protection Information Protection Admins Information Protection Analysts Information Protection Investigators Insider Risk Management Insider Risk Management Admins Insider Risk Management Analysts Insider Risk Management Approvers Insider Risk Management Auditors

Role	Description	Default role group assignments
		Insider Risk Management Investigators Organization Management Security Administrator Security Operator
*Purview Domain Manager	Create, edit, and delete domains and perform role assignments.	Purview Administrators
*Purview Evaluation Administrator	Create and manage the Microsoft 365 Purview Evaluation lab.	Data Security Management Information Protection Information Protection Admins Information Protection Analysts Information Protection Investigators
Quarantine	Allows viewing and releasing quarantined email.	Organization Management Quarantine Administrator Security Administrator
RecordManagement	View and edit the configuration of the records management feature.	Compliance Administrator Compliance Data Administrator Organization

Role	Description	Default role group assignments
		Management Records Management
Retention Management	Manage retention policies, retention labels, and retention label policies. Includes permissions to add and remove adaptive scopes from these policies, and to create, delete, and modify adaptive scopes.	Compliance Administrator Compliance Data Administrator Organization Management Records Management
*Review	This role lets users access review sets in eDiscovery (Premium) cases. Users who are assigned this role can see and open the list of cases on the **eDiscovery** > **Advanced** page in the Microsoft Purview portal that they're members of. After the user accesses an eDiscovery (Premium) case, they can select **Review sets** to access case data. This role doesn't allow the user to preview the results of a collection search that's associated with the case or do other search or case management tasks. Users with this role can only access the data in a review set.	Data Investigator Data Security Management eDiscovery Manager Insider Risk Management Insider Risk Management Investigators Privacy Management Privacy Management Administrators Privacy Management Analysts Privacy Management Contributors

Role	Description	Default role group assignments
		Privacy Management Investigators Privacy Management Viewers Reviewer Subject Rights Request Administrators Subject Rights Request Approvers
*RMS Decrypt	Decrypt RMS-protected content when exporting search results.	Data Investigator eDiscovery Manager
Role Management	Manage role group membership and create or delete custom role groups.	Organization Management Purview Administrators
*Scan Reader	Read the different scans created in the tenant.	Compliance Administrator Data Security Management Data Source Administrators Information Protection Information Protection Admins Information Protection Investigators

Role	Description	Default role group assignments
*Scan Writer	Create, update, and delete scans in the tenant.	Compliance Administrator Data Source Administrators
Scope Manager	Enables administrators to create, edit, delete, and control access to scoping features such as Adaptive Scopes in the organization.	Communication Compliance Communication Compliance Administrators Compliance Administrator Compliance Data Administrator Organization Management Records Management
Search And Purge	Lets people bulk-remove data that matches the criteria of a content search.	Data Investigator Organization Management
Security Administrator	View and edit the configuration and reports for Security features.	Organization Management Security Administrator
Security Reader	View the configuration and reports for Security features.	Global Reader Organization Management Security Operator Security Reader
Sensitivity Label Administrator	View, create, modify, and remove sensitivity labels.	Compliance Data Administrator

Role	Description	Default role group assignments
		Organization Management Security Administrator
Sensitivity Label Reader	View the configuration and usage of sensitivity labels.	Global Reader Organization Management Security Reader
Service Assurance View	Download the available documents from the Service Assurance section. Content includes independent auditing, compliance documentation, and trust-related guidance for using Microsoft 365 features to manage regulatory compliance and security risks.	Global Reader Organization Management Service Assurance User
*Source Reader	Read the different sources created in the tenant.	Compliance Administrator Data Security Management Data Source Administrators Information Protection Information Protection Admins Information Protection Investigators Privacy Management Privacy Management Administrators

Role	Description	Default role group assignments
*Source Writer	Create, update, and delete sources in the tenant.	Compliance Administrator Data Source Administrators
*Subject Rights Request Admin	Manage supervisory review policies, including which communications to review and who should perform the review.	Privacy Management Subject Rights Request Administrators
*Subject Rights Request Approver	Create, edit, delete, and control access to custodian.	Subject Rights Request Approvers
*Supervisory Review Administrator	Manage supervisory review policies, including which communications to review and who should do the review.	Supervisory Review
Tag Contributor	Enables viewing and updating of existing tags.	Organization Management Security Administrator Security Operator
Tag Manager	View, update, create, and delete user tags.	Organization Management Security Administrator
Tag Reader	Read-only access to existing user tags.	Organization Management Security Administrator Security Operator Security Reader
*Tenant AllowBlockList Manager	Manage Tenant Allow/Block List settings.	Security Operator
View-Only Audit Logs	View and export audit reports. Because these reports might contain sensitive information, you should only	Audit Manager

Role	Description	Default role group assignments
	assign this role to people with an explicit need to view this information.	Audit Reader Compliance Administrator Compliance Data Administrator Global Reader Organization Management Security Administrator Security Operator
View-Only Case		Communication Compliance Communication Compliance Investigators Compliance Administrator Data Security Management Insider Risk Management Insider Risk Management Admins Insider Risk Management Analysts Insider Risk Management Investigators Organization

Role	Description	Default role group assignments
		Management Privacy Management Privacy Management Administrators Privacy Management Analysts Privacy Management Investigators Subject Rights Request Administrators
View-Only Device Management	View the configuration and reports for the Device Management feature.	Compliance Administrator Compliance Data Administrator Global Reader Organization Management Security Administrator Security Operator Security Reader
View-Only DLP Compliance Management	View the settings and reports for data loss prevention (DLP) policies.	Compliance Administrator Compliance Data Administrator Global Reader

Role	Description	Default role group assignments
		Organization Management Security Administrator Security Operator Security Reader
View-Only IB Compliance Management	View the configuration and reports for the Information Barriers feature.	Compliance Administrator Compliance Data Administrator Global Reader Organization Management Security Administrator Security Operator Security Reader
View-Only Manage Alerts	View the configuration and reports for the Manage Alerts feature.	Compliance Administrator Compliance Data Administrator Global Reader Organization Management Security Administrator Security Operator Security Reader

Role	Description	Default role group assignments
View-Only Recipients	View information about users and groups.	Compliance Administrator Compliance Data Administrator Global Reader MailFlow Administrator Organization Management
View-Only Record Management	View the configuration of the records management feature.	Compliance Administrator Compliance Data Administrator Global Reader Organization Management
View-Only Retention Management	View the configuration of retention policies, retention labels, and retention label policies.	Compliance Administrator Compliance Data Administrator Global Reader Organization Management

> **Note**
>
> [2] This role or role group is in preview. While it may be visible in commercial Microsoft 365 tenants, it is not supported and has no operational effect outside the M365 ABAC private preview. At this time, support is planned only for Microsoft U.S. government cloud environments (GCC High and DoD, see **Understand Microsoft U.S. government cloud environments for Microsoft 365 and Microsoft Copilot**).

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 09/24/2026

Purview Role Assignment Migrator

The **Microsoft Purview Role Assignment Migrator** is a Microsoft first-party enterprise application in Microsoft Entra ID that runs the background synchronization *from* Microsoft Purview role assignments *to* corresponding Entra roles. Microsoft 365 services use these Entra roles to verify that the administrator has the correct level of permissions before allowing long-running Microsoft Purview operations, such as content search and export, to proceed.

Role mapping

The following table shows which Purview roles map to each of the following Entra roles. These Entra roles only grant the ability to run Purview operations, and no other operations outside Purview are granted.

Expand table

Purview roles	Mapped Entra role	Description
Insider Risk Management Analysis Insider Risk Management Investigation Compliance Search Export Privacy Management Admin Privacy Management Analysis Privacy Management Investigation Privacy Management Permanent Contribution Privacy Management	Purview Workload Content Reader	Allows Microsoft Purview operations to read content from Microsoft 365 services (for example, read a file in SharePoint).

Purview roles	Mapped Entra role	Description
Temporary Contribution Privacy Management Viewer Data Security Investigation Reviewer		
Hold Privacy Management Investigation Data Security Investigation Investigator	Purview Workload Content Writer	Allows Microsoft Purview operations to read and write content to Microsoft 365 services (for example, store mail items in Exchange).
Search and Purge Data Security Investigation Admin Data Security Investigation Analyst	Purview Workload Content Administrator	Allows Microsoft Purview operations to read, write, and delete content in Microsoft 365 services (for example, purge messages in Microsoft Teams).

If an administrator holds multiple Purview roles that map to different Entra roles, they receive the highest-privilege Entra role. The precedence order is: **Administrator** > **Writer** > **Reader**.

> ⓘ **Note**
>
> The Purview Role Assignment Migrator manages the Purview Workload Content Reader, Purview Workload Content Writer, and Purview Workload Content Administrator Entra roles, and only the listed Purview roles synchronize to Entra. Don't assign the Entra roles in the Entra portal. These roles don't appear in the Microsoft Purview portal like the ones that are shown on the settings page.

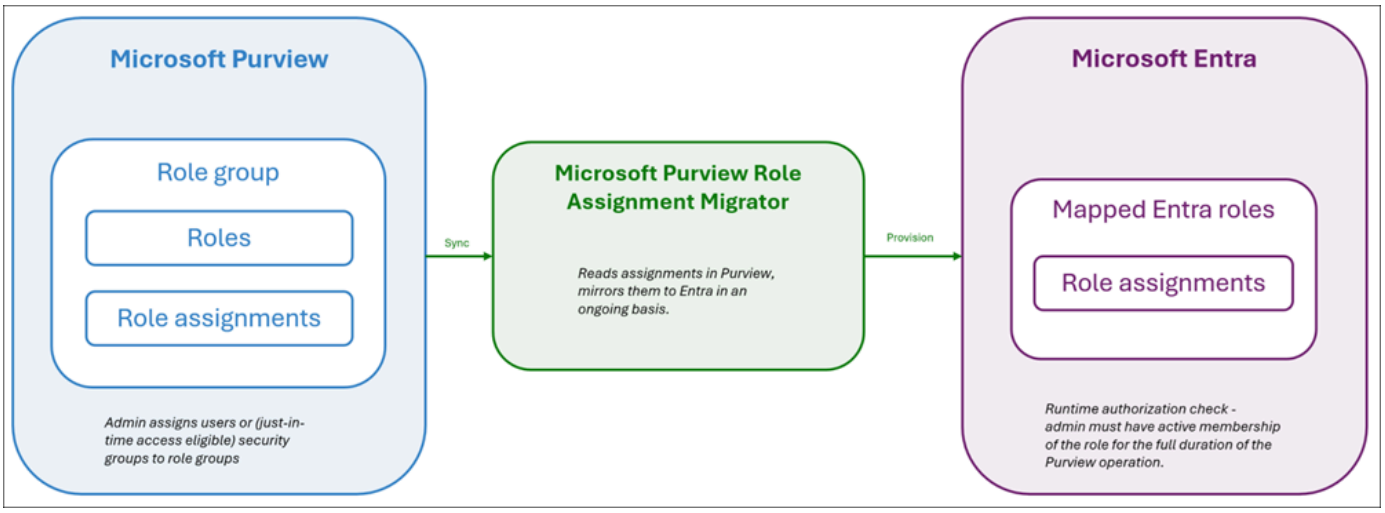
Synchronization audit logs

The Purview Role Assignment Migrator works in two modes:

- **Initial bulk sync:** When the Purview Role Assignment Migrator first activates for your tenant, it syncs all existing Purview role assignments to Microsoft Entra in a single pass. This process

generates a burst of activity in your Microsoft Entra audit logs.

- **Continuous sync:** All subsequent changes to Purview role memberships trigger synchronization to Microsoft Entra.



> **Important**
>
> Keep the Purview Role Assignment Migrator enterprise application enabled in Microsoft Entra ID. The application ID is `7fe3d988-4f3b-4f33-83bd-1fb921a35ed2`. Disabling this app stops synchronization. New Purview role assignments don't propagate to Microsoft Entra, and long-running Purview operations fail authorization checks at runtime.

The Purview Role Assignment Migrator synchronization activity appears in Microsoft Entra audit logs with the display name **PurviewRoleAssignmentMigrator**. The **New Value** field for each log entry shows the Microsoft Entra role that was assigned.

You see two distinct patterns of activity:

Expand table

Activity pattern	When it occurs	Volume
Bulk sync	Once, when the Purview Role Assignment Migrator first activates for your tenant	High — all existing Purview role assignments sync at once
Continuous sync	Ongoing, after each change to a Purview role membership	Low — proportional to the rate of Purview role changes in your tenant

If you see a sudden spike of `PurviewRoleAssignmentMigrator` entries in your Microsoft Entra audit logs, this spike in audit log entries comes from the initial bulk sync and isn't a sign of

unauthorized activity.

> **Important**
>
> Don't assign users to these Microsoft Entra roles directly in Microsoft Entra ID. The Purview Role Assignment Migrator manages these assignments exclusively from Purview and it overwrites any manual assignments in Microsoft Entra at the next synchronization.

Just-in-time access

By default, the Purview Role Assignment Migrator synchronizes the role assignments to the Microsoft Entra roles as **active (permanent)** assignments. Microsoft Entra Privileged Identity Management (PIM) for Groups allows eligible users to activate just-in-time membership to groups. When you add these Microsoft Entra **security groups** with eligible assignments to Purview roles, the Purview Role Assignment Migrator synchronizes that same security group membership into the corresponding Microsoft Entra roles. This process allows organizations that require role assignments to be just-in-time in Microsoft Entra to apply the same model for Microsoft Purview and enforce role activation.

> **Important**
>
> Size your just-in-time membership activation window on the security group to cover the full duration of your operations. If an activation expires while a long-running operation is still running, the operation might fail at its next runtime authorization check. For example, to identify how long your eDiscovery operations typically take, use **Process Manager in eDiscovery** and set activation durations accordingly.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 06/04/2026

Learn about Microsoft Purview billing models

Microsoft Purview offers a set of integrated solutions that help customers secure and govern their data. It's composed of data security and data governance solutions for:

- Microsoft 365 and non-Microsoft 365 workloads and data systems.
- Microsoft 365 and non-Microsoft 365 endpoints.
- Compliance solutions for data used with Generative AI applications.

Microsoft Purview supports two complementary billing models to support this diverse environment:

- Per-user license for Microsoft 365 and Windows/macOS endpoint sources
- Pay-as-you-go model for non-Microsoft 365 data sources and certain other capabilities

This article takes you through:

- An overview of the two billing models.
- Links to information on the solution specific capabilities.

> **Note**
>
> The pay-as-you-go billing model builds on the per-user licensing model. The two are complementary, not mutually exclusive.

Per-user licensing model

The per-user licensing model is the familiar Microsoft 365 E3/E5/A5/F5/G5 model. It remains unchanged. This licensing model enables you to apply Microsoft Purview controls and protections to Microsoft 365 and Windows/macOS endpoint-based assets. It's described in the Microsoft Purview service description.

Pay-as-you-go billing model

The pay-as-you-go billing model is a consumption billing model. It extends Microsoft Purview data security, data governance, and data risk and compliance protection capabilities beyond Microsoft 365 and Windows/macOS environments to non-Microsoft 365 locations. The billing mechanisms are Azure based, so you must associate your Microsoft 365 tenant with an active

Azure subscription. It also protects data moving through networks, DataOps, application architectures, and AI apps and agents, as well as certain other user-agnostic capabilities.

In the pay-as-you-go model, you're charged based on your usage of various pay-as-you-go features, based on the unit of consumption for the feature. An Azure bill is generated at the end of the month for the total number of pay-as-you-go charges across all Microsoft solutions.

> **Note**
>
> For guidance on how to associate a Microsoft 365 tenant to an Azure subscription, see Enable Microsoft Purview pay-as-you-go features for new customers.

Capabilities that use the pay-as-you-go billing model

> **Note**
>
> Some Microsoft Purview solutions use the per-user licensing model. Some require the per-user licensing model to be enabled before you can use the pay-as-you-go billing model. Some are pay-as-you-go only and don't require a per-user license.

Data Security and Data Governance

The pay-as-you-go billing model extends Microsoft Purview data security and governance capabilities beyond Microsoft 365 and Windows/macOS to environments like:

- Amazon Web Services (AWS)
- Azure ADLS
- Azure SQL
- Box
- Dropbox
- Google Drive
- Microsoft Fabric

For details on data governance assets, see Microsoft Purview data governance billing.

Data security capabilities

Expand table

Solution	Applies to	Unit of Measure	Details
Data Security Investigations	Data storage meter based on the storage associated each investigation	Number of gigabytes of stored data for all investigations/month Number of Security Compute units consumed	Learn more about billing for Data Security Investigations
Information Protection	Sensitivity labels that you apply to non-Microsoft 365 data sources	Number of assets in scope of protection policy/day	Refer to details below on how assets are defined and calculated
Insider Risk Management	Detect risky behavior for non-Microsoft 365 locations when using Cloud and generative AI policy indicators	Data Security processing unit as measured on a daily basis	Learn more about Insider Risk Management policy indicators

Data governance capabilities

Expand table

Solution	Applies	Unit of Measure	Details
Unified Catalog data curation	Applies when you actively curate and manage the technical assets in Microsoft Purview Unified Catalog	Number of unique assets governed/day	Learn more about data governance billing
Unified Catalog data health management	Applies when you manage data quality and take health management actions	Number of data governance processing units (DGPUs) consumed	Learn more about data governance billing

Data governance processing units: For more information on data governance processing units, see Data governance processing units explained.

Data risk and compliance capabilities

For generative AI apps and agents, Microsoft Purview offers the following capabilities on a pay-as-you-go billing model. Microsoft 365 Copilot experiences aren't charged.

Expand table

Solution	Applies to	Unit of Measure	Details
Audit solutions	Audit logs for user interactions with non-Microsoft generative AI applications	Number of audit records processed	Learn more about generative AI applications supported by Audit solutions.
Communication compliance	Detect inappropriate or risky interactions for non-Microsoft 365 AI interactions when using AI policy indicators	Number of text records scanned	Billing meters for communication compliance are broken down into two categories, standard, and premium. Learn more about the channels and generative AI applications supported by Communication Compliance
Data Lifecycle Management	Retention policies for AI interactions	Number of non-Microsoft 365 Copilot or AI App interactions (prompts and responses) under a retention policy	Each non-Microsoft 365 generative AI prompt and response count as a separate interaction and is retained and deleted according to the retention policy settings you configure in Microsoft Purview. For more information on retention policies in DLM for Microsoft 365 Copilot and AI apps, see Learn about retention for Copilot & AI apps
eDiscovery	Storage of non-Microsoft 365 AI application data and usage of Microsoft Graph APIs for standard licensed tenants	Number gigabytes stored/day for storage of non-Microsoft 365 AI data. Gigabytes/export for usage of Export API by Standard licensed tenants	Learn more about eDiscovery billing

Other Microsoft Purview solutions that use pay-as-you-go pricing

Expand table

Solution	Applies to	Unit of Measure	Details
On-demand classification (preview)	Applies when you run a scan to identify and classify sensitive content in data stored in SharePoint and OneDrive	Asset, based on the number classified per scan	Learn more about On-demand classification (preview)

Solution	Applies to	Unit of Measure	Details
Microsoft Security Copilot	Applies for all Security Copilot functions	Security Compute Units	Get started with Security Copilot
Network Data Security (preview)	Requests from an endpoint device to a website, cloud app, or generative AI app	Number of requests sent from the endpoint device to the website, cloud app, or generative AI app	Network data security only counts requests that are outbound from the device.
Data Security for Gen AI Applications	Applies to classification and protection of sensitive content for non-M365 AI interactions (prompts/responses)	Number of requests or messages for non-Microsoft 365 AI interactions (prompts or responses)	Learn more about data security for AI interactions
Data Loss Prevention for Cloud Apps in Edge for Business browser	Applies to requests from the Edge for Business browser to a website, cloud app, or generative AI app	Number of requests sent from the Edge for Business browser to the website, cloud app, or generative AI app	Data loss prevention for cloud apps only counts requests for unmanaged cloud apps.

Cost estimator tools

Understand pricing and estimate your expected monthly costs for pay-as-you-go capabilities:

- Microsoft pay-as-you-go pricing
- Estimate your costs for Microsoft Purview pay-as-you-go feature use

Units of measure for pay-as-you-go capabilities

Assets

An asset is any Microsoft 365 item that a Microsoft Purview policy protects. Some examples are a table for structured data or any file for unstructured data. You're charged for each day that a policy covers an asset. For Microsoft Data Lifecycle Management, assets include interactions (prompts and responses) between a user and a Copilot or AI App.

Here are some examples of assets for Microsoft Purview Data Security and Governance and the protection policies that you can apply to them.

Expand table

Cloud provider	Data source	Asset	Can be protected by
Azure	SQL DB	Table	Protection policy and auto-labeling policy from Microsoft Purview Information Protection
Azure	ADLS	File or resource set	Protection policy and auto-labeling policy from Microsoft Purview Information Protection
Azure	Blob	File or resource set	Protection policy and auto-labeling policy from Microsoft Purview Information Protection
Azure	Fabric	Supported item types	Protection policies from Microsoft Purview Information Protection
Azure	Fabric	Supported item types	Microsoft Purview Data Loss Prevention

For details on data governance assets, see Microsoft Purview data governance billing.

Counting assets

Assets are counted based on the number of items that are in the scope of a policy. The asset doesn't have to match a policy's conditions to be counted, it just has to be in a location that's in the scope of a policy. An asset is only counted once, regardless of how many solutions or protection policies cover it.

Here are some examples:

- **Protection policy**: A tenant has SQL server with 100 tables, and 50 of those tables are labeled **Confidential** and 20 are labeled **General**. The protection policy only applies to assets labeled with **Confidential**, so the asset count is 50.
- **Data Loss Prevention policy**: A tenant has 10 Fabric workspaces, each with 50 assets. The policy only applies to five of these Fabric workspaces. The asset count for this policy is 250.

Processing units

A processing unit measures the amount of compute resources used to process signals from workloads that are included in pay-as-you-go.

- **Data governance processing units**: A data governance processing unit (DGPU) is a fully managed compute unit that runs compute capabilities such as data quality and data health management. Each DGPU is 60 minutes of compute time that runs across varying

sets of nodes based on the workload need. For more information on data governance processing units, see Data governance processing units explained.

- **Data security processing units**: Microsoft Purview data security processing units represent the compute required to process user activities from non-Microsoft 365 data sources to generate insights. The number of units required can vary depending on the Purview solution and the complexity of the data being processed. Insider risk management is a pay-as-you-go feature that uses a processing unit based meter. It includes pay-as-you-go indicators. For more information, see Configure policy indicators in insider risk management.
- **Compute units**: Data Security Investigation compute units (compute units) represent the compute required to process and generate AI-based insights in Data Security Investigations. The number of units required can vary depending on the process and is a pay-as-you-go feature that uses a processing unit based meter.For more information, see Billing in Data Security Investigations.

Counting processing units

The number of processing units you need varies depending on the Microsoft Purview solution and the complexity of the data being processed.

For example, Insider Risk Management processes user activities corresponding to the non-Microsoft 365 indicators selected in data theft and data leak policies to generate insights, alerts, and cases. For Insider Risk Management, a processing unit represents the compute required to process 10,000 such user activities. Billing is based on the number of processing units utilized. Insider Risk Management indicators that detect activity in cloud storage apps (Box, Dropbox, Google Drive), cloud services (AWS, Azure), and Microsoft Fabric (Power BI) are billed on data security processing units.

Data storage meter (GBs)

Microsoft Purview data storage meter is defined by a gigabyte per month storage amount billed at a specific rate for applicable solutions. The total amount of data within a solution subject to data storage meter is automatically calculated from the current amount of data in solution-related containers. For example, pay-as-you-go in Data Security Investigations and eDiscovery use the data storage meter.

eDiscovery also uses a storage meter defined by the gigabytes per export through Graph API for E3 eDiscovery licensed customers. Each organization receives an included amount of free storage per month, with additional usage billed at a set price per GB.

For more information about API storage and pricing, see Microsoft Purview pricing.

Text records

Microsoft Purview text records meter is defined by text records, where one text record equals 1,000 characters. Each message converts into multiple text records based on character length. For example, pay-as-you-go in Communication Compliance uses a text records-based meter.

1 text record = 1,000 characters

If a message has more than 1,000 characters, it counts as one text record for each unit of 1,000 characters. For instance, if a message contains 7,500 characters, it counts as eight text records. If a message contains 500 characters, it counts as one text record.

Counting text records for Communication Compliance Premium

Counting text records for premium is based on the type of detections that the messages are being evaluated for. If the evaluation includes the following detections, they're charged at the premium rate:

1. **Code of conduct**: Hate, self-harm, violence, sexual
2. **Risky Gen AI:** Prompt shields, protected materials

Calculating text records for Communication Compliance Standard

If messages are evaluated positive for any other detections, such as other trainable classifiers or sensitive info types, you pay standard charges. Messages that evaluate messages for both premium and standard detections are charged only for premium detections.

Security Compute Units (SCU)

A security compute unit (SCU) is a bundled unit of measure. It combines all the resources, like network, processor, and storage, that Microsoft Security Copilot needs to provide service into a single, billable quantity.

For more information on SCUs, see Get started with security compute units.

Counting SCUs

The number of SCUs you have available is your capacity. Your available capacity consists of two types:

- **Provisioned SCUs:** These are units that you pre-allocate. You use the cost estimators to project the number of SCUs you anticipate using. Use of provisioned SCUs is measured in hourly blocks. For example, if you provision an SCU at 9:05 a.m., then deprovision it at 9:35 am, and then provision another SCU at 9:45 am, you are charged for two units within the 9:00 a.m. to 10:00 a.m. hour. To maximize usage, make SCU provisioning changes at the beginning of the hour. For more information, see Manage usage.
- **Overage SCUs:** These units cover unexpected spikes in usage. To manage unexpected demand spikes, you can allocate an overage amount to ensure that additional SCUs are available when you run out of provisioned units. Overage units are billed on-demand. You can set limits for overage units as unlimited or set a maximum upper amount. This approach enables predictable billing while providing the flexibility to handle both regular and unexpected usage.

Requests

For pay-as-you-go billing, a request is each network call a device or browser makes to a website or API. This definition doesn't include the responses to the requests. Azure counts requests in the monthly pay-as-you-go bill you receive. Microsoft Purview network data security pay-as-you-go uses requests as its unit of measure. Here are some examples:

Expand table

Activity	Data type	Example
Text sent to or shared with cloud or AI app	human readable strings transmitted inline	- submitting a form with textual information - Sending raw text or a prompt to a generative AI - the body of an email - sending JDSON data to an API
File uploaded to or shared with cloud or AI app	Byte streams, including text based file, binary files, txt files, source code, documents, images, videos, .exe's, .pdf's, archive files	- Uploading a profile picture to social media - sending a document or PDF file as an email attachment - sharing a document with generative AI - transferring a document or .ZIP files to a cloud storage solution

See also

- Enable Microsoft Purview pay-as-you-go features via subscription
- Learn more about pricing for Microsoft Purview pay-as-you-go features
- Calculate your estimated costs for Microsoft Purview pay-as-you-go features

---

Last updated on 01/06/2026

Consent to use use pay-as-you-go capabilities

Article • 04/08/2025

Applies to:

- Existing Azure Purview customers who are using the classic data governance pricing model.
- Microsoft Purview customers using Data Security pay-as-you-go features in free public preview.

> **Important**
>
> You can only use this procedure if your organization is in one of the above categories.

This feature is in preview.

Consent to use pay-as-you-go capabilities

If your organization is in one of the above categories, you must give your consent to switch from your current billing model to the pay-as-you-go model.

If you aren't familiar with the pay-as-you-go model see, Learn about Microsoft Purview billing models.

Prerequisites for pay-as-you-go

To consent to changing to the pay-as-you-go model, you must have these prerequisites.

Roles and permissions to consent to the pay-as-you-go pricing structure

Only an account in the Global Administrator role can provide consent to switch to the pay-as-you-go model.

Consent experience

Microsoft Purview has a single unified consent experience that applies to all pay-as-you-go apps and capabilities. Your consent applies to all current and future pay-as-you-go features, as

they transition from free to paid public preview.

> **Note**
>
> Although the consent experience is common for all apps and capabilities, certain apps may not be available in all regions. For instance, data governance is currently available in specific regions. For more information on region availability, see Supported regions for Microsoft Purview Data Governance.

1. Open the Microsoft Purview portal > **Settings** > **Account overview** in the new Microsoft Purview portal.
2. Select **Review Consent** and switch to the pay-as-you-go model.
3. Microsoft Purview Unified Catalog app customers can learn more here.
4. A confirmation is displayed that you've successfully switched to the new business model.

See also

- Learn about Microsoft Purview billing models
- Enable Microsoft Purview pay-as-you-go features via subscription

Enable Microsoft Purview pay-as-you-go features for new customers

Applies to:

- Organizations that aren't current Azure Purview customers using the classic data governance pricing model customers.

In this procedure, you enable the Microsoft Purview pay-as-you-go features by associating your Microsoft 365 tenant with an Azure subscription for billing purposes.

If you aren't familiar with the pay-as-you-go model see, Learn about Microsoft Purview billing models.

This featue is in preview.

Prerequisites for pay-as-you-go

To enable pay-as-you go, you must have:

- An Azure subscription in the same tenant as Microsoft Purview.
- An Azure resource group in that subscription.
- Either of the following service plan
  - MIP_S_Exchange
  - EXCHANGE_S_FOUNDATION

If you already have these resources, you can use them with Microsoft Purview.

If you don't already have these resources, see:

- Create your initial Azure subscriptions
- Manage Azure resource groups by using the Azure portal.

Roles and permissions to consent to the new pricing structure

Only an account in the Global Administrator role can enable the pay-as-you-go model:

Subscription based

The enablement flow requires you to associate your Microsoft Purview tenant to an Azure subscription. All consumption and billing information is available in the Azure portal under **Cost Management + Billing**.

You can use an existing Azure subscription and resource group or create a new one. The Azure subscription must be active to bill against before you can enable the pay-as-you-go billing model.

1. Open the Microsoft Purview portal and select the rocket icon on the top right in the header pane, near the **Settings** icon.
2. Select **Get Started** in the ribbon menu to start the process. This can also be done from **Settings** > **Account details**.
3. Select your active Azure subscription from the list. If you don't have an active Azure subscription, you must create a new one before you can proceed. See Create your initial Azure subscriptions.
4. Select the resource group. If you don't have a **Resouce group** in your Azure tenant, you must create a new one before you can proceed. See - Manage Azure resource groups by using the Azure portal. The billing resource is created under the resource group. This is where the pay-as-you-go version of Microsoft Purview is hosted in the Azure portal.

The *Country/Region* value is fixed for the tenant.

Upon completion, it can take a few hours for changes to complete. When completed, you can use pay-as-you-go features to protect your non-Microsoft 365 assets on the **Solutions** page.

See also

- Learn about Microsoft Purview billing models
- Consent to use use pay-as-you-go capabilities

Last updated on 02/03/2026

Manage pay-as-you-go and per-user licensing usage

> **Note**
>
> The Microsoft Purview **Usage center** feature is in preview.

Microsoft Purview solutions are licensed through the **Premium licensing model**, which is the per-seat model, the **pay-as-you-go billing model** model, or a combination of both. If you need to understand the billing models in Microsoft Purview, see Learn about Microsoft Purview billing models.

The **Usage center** in the Microsoft Purview portal helps you understand and manage your usage for both billing models. You can view your pay-as-you-go billing usage and your per-user licensing usage for Microsoft Purview capabilities.

Permissions

To access the **Usage center** > **Pay-as-you-go** and **Usage center** > **Premium usage** pages, the account you use must have one of the following roles:

- Global Reader
- Security Reader
- Global Admin

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

Pay-as-you-go usage

The **Usage center** provides detailed information on **Total Units Used** about your pay-as-you-go billing usage by feature and by workload. For definitions on the **Unit of Measure** see Units of measure for pay-as-you-go capabilities. The pay-as-you-go usage page also lets you toggle the features listed between paused and active states.

Expand table

Feature name	Pausable	Detailed report available
Communication compliance Standard	Yes	Yes
Communication compliance Premium	Yes	Yes
Enterprise Data Management Basic	No	Yes
Enterprise Data Management Standard	No	Yes
Enterprise Data Management Advanced	No	Yes
Audit	Yes	Yes
Insider Risk Management	No	Yes
On-Demand Classification	Yes	Yes
Information Protection	Yes	Yes
Data lifecycle management	Yes	Yes
Data Catalogue Governed Assets	No	Yes
Data Security Investigations	No	No
In-transit protection of data to AI apps	No	Yes
In-transit protection of data through Browser	No	Yes
In-transit protection of data at the network layer	No	Yes
eDiscovery	Yes	No

Premium usage model

The **Premium usage** model is useful for identifying underlicensing situations and under utilization of purchased licenses.

Policy Scoping

Use the **Policy Scoping** card tells you how many unlicensed users are in scope of one or more Microsoft Purview premium features. This helps you identify potential under-licensing situations so you can optimize your licensing usage.

The **Policy Scoping** card shows three key metrics:

- Policy scoped users: The total number of users in scope of one or more Microsoft Purview premium features.
- Policy scope users with assigned licenses: The number of users in scope of one or more, which are licensed one or more of Microsoft Purview premium features.
- Policy scoped users without assigned licenses: The number of users in scope of one or more Microsoft Purview premium features, which aren't licensed for any of those features.

The unlicensed user metric is calculated like this:

Policy scoped users - **Policy scoped users with assigned licenses** = **Policy scoped users without assigned licenses**

License Assignment

There may be users who are assigned licenses for Microsoft Purview premium features but aren't in scope of any of the features. So essentially, you're paying for usage of a feature, but not using the feature. The **License Assignment** card helps you identify these users so you can optimize your licensing usage.

The **License Assignment** card shows two key metrics:

- **License assigned users**: The number of users that are assigned Microsoft Purview Premium licenses.
- **License assigned users not under policy scope**: The number of users that are assigned Microsoft Purview Premium licenses but aren't in scope of any Microsoft Purview premium features.

The **License assigned users not under policy scope** metric is calculated like this:

License assigned users - **Policy scoped users with assigned licenses** = **License assigned users not under policy scope**

See also

- Enable Microsoft Purview pay-as-you-go features via subscription
- Learn more about pricing for Microsoft Purview pay-as-you-go features
- Calculate your estimated costs for Microsoft Purview pay-as-you-go features

---

Last updated on 11/19/2025

Free version of Microsoft Purview data governance solutions

Article • 04/04/2025

> **Important**
>
> This feature is currently in preview. The Supplemental Terms of Use for Microsoft Azure Previews include additional legal terms that apply to Azure features that are in beta, in preview, or otherwise not yet released into general availability.

Microsoft Purview offers your administrators the ability to govern your data estate and your users the ability to discover data across your organization. The free version of Microsoft Purview governance solutions is an automatically available instance, ready for your users without needing to set it up. It allows your organization to try Microsoft Purview's basic capabilities and begin your governance journey.

The free version of Microsoft Purview governance solutions uses a single, organization-wide instance of Microsoft Purview that users can access to explore their Azure and Microsoft Fabric resources. Metadata for these resources is automatically inventoried and can be searched using the catalog.

Access to this version of Microsoft Purview depends on your organization's current relationship with Microsoft Purview; whether you're a new Microsoft Purview user, if you have existing accounts, or if you're already using other services like Microsoft Fabric. For more information, see, how to access the free version of Microsoft Purview governance solutions.

Free version limitations

- In the free version of Microsoft Purview governance solutions, there's a limitation of 1,000 annotated assets. An annotated asset applies both to live view assets that have been curated, or custom assets added using the Microsoft Purview Data Map REST APIs. For example, if your account has 50 live view data assets, and you add descriptions and classifications to 15 of those assets, you have 15 annotated assets, and can have 985 more before you reach your limit. A description and a classification on the same asset count only as one annotated asset.
- In the free version, only data curators are able to access any resources added to the data map using the REST APIs.

- If your free version tenant shows three months of no activity, we have the right to clean up the account and might remove the unused resources.

What's available in the free version

Users who have access to data stored in supported Azure resources have access to its metadata in Microsoft Purview. Data readers can search for and browse this metadata and data owners can augment the metadata to provide more context.

Only a core subset of Microsoft Purview's governance solutions with limited capabilities are currently available in the free version:

- Catalog: Browse and search for your data assets.
- Data Map history: A log of updates made to assets.

Catalog

Users can browse and search for their data assets in Azure. You can explore your Azure assets to discover reliable data sources in their environment using live view without any setup needed. The catalog automatically records data schema so users can explore data sources before accessing them. Data owners can add metadata and information to their assets, such as description, classifications, and contacts, so users have a more complete picture of their data estate.

Supported data sources

Currently, the free version of Microsoft Purview governance solutions supports these governance data sources using live view:

- Azure Blob Storage
- Azure Data Lake Storage Gen 2
- Azure SQL Database
- Azure subscriptions
- Microsoft Fabric

You can annotate and curate assets that are available via live view, or use our APIs to create new entities.

Access the free version of Microsoft Purview governance solutions

The free version of Microsoft Purview governance solutions uses a single, organization-wide instance of Microsoft Purview. Access to the latest Microsoft Purview experience depends on your current relationship with Microsoft Purview; whether you're a new Microsoft Purview customer or an existing Microsoft Purview customer.

The free version of Microsoft Purview governance solutions uses the same location mappings as Microsoft 365, and maps to a location based on your Microsoft Entra ID home region. For more information, see the FAQ.

New Microsoft Purview customers

If you're an Azure Enterprise, Dynamics 365, Microsoft 365 Enterprise, Microsoft Fabric, Power Apps, or Power BI customer, to access the free version of Microsoft Purview governance solutions for your organization go to https://purview.microsoft.com. For more information about this access, see the FAQ.

If you're not an Azure Enterprise, Dynamics 365, Microsoft 365 Enterprise, Microsoft Fabric, Power Apps, or Power BI customer, we're currently making the free version of Microsoft Purview governance solutions available to organizations in batches. Go to https://purview.microsoft.com to check if it's available to your organization. If not:

- You can create a Microsoft Purview account to access the **enterprise** version of Microsoft Purview.
- You can wait until the free version of Microsoft Purview governance solutions is available to your organization, and check this document for updates.

Existing Microsoft Purview customers

If you're already using Microsoft Purview, your account is likely already upgraded to the enterprise version of Microsoft Purview.

You can check your organization's account type in the Microsoft Purview portal under the **Settings** gear icon, and **View all settings**.

Permissions

Access control in the free version of Microsoft Purview governance solutions is managed at the organization/tenant level. All users are able to view data assets for sources where they have at least **Read** permissions already. For other functionality, like metadata curation, you can assign this role:

Expand table

Role	Description
Data Catalog Curator	Catalog curators can perform the create, read, modify, and delete actions on catalog data objects and establish relationships between objects.

Get details about governance roles and permissions.

Upgrade to enterprise

The enterprise version of Microsoft Purview has access to the rest of Microsoft Purview applications and data sources. Learn how to upgrade to the enterprise version of Microsoft Purview data governance.

When you upgrade, all your data from the free version of Microsoft Purview governance solutions will be migrated into the enterprise version. The enterprise version will be hosted in the same location as the free version of Microsoft Purview governance solutions.

What's in the the free and enterprise versions

Here's a brief comparison of the free and enterprise versions of Microsoft Purview data governance:

Expand table

Free	Enterprise
Auto discovery of Azure data sources	Collections, and glossary terms to organize your data map
Role group access control to platform and apps	Fine-grained, collection-level access control to platform and apps
Definition and manual application of classifications and terms	Automated application of classifications and terms
Manual descriptions, tags, and contacts	Manual and Bulk descriptions, tags, and contacts.
Manual creation of assets using the data map APIs	Full use of Microsoft Purview's REST APIs
	Automated scans for the hybrid data estate

Free	Enterprise
	Workflows
	Business rules
	Support for business assets and managed attributes

Applications supported between versions of Microsoft Purview:

Expand table

Application	Free	Enterprise	Notes
Catalog	x	x	Free supports browse and search, while enterprise supports advanced search and self-service data access.
Data Estate Insights		x	Enterprise supports all available reports
Data Policy		x	
Data Map	x	x	Data map history is available in both the free and enterprise versions of Microsoft Purview.

For more information about the applications available in the enterprise version of Microsoft Purview, see the overview article.

Support for the free version of Microsoft Purview governance solutions

With the free version of Microsoft Purview governance solutions, you can't open support tickets for Microsoft Purview. You need to upgrade to the enterprise version of Microsoft Purview for support.

Next steps

- Get started with the free version of Microsoft Purview governance solutions
- New Microsoft Purview portal
- Upgrade to enterprise
- Roles and permissions

Feedback

Was this page helpful? Yes No

Get started with the free version of Microsoft Purview data governance solutions

Article • 04/04/2025

> **Important**
>
> This feature is currently in preview. The Supplemental Terms of Use for Microsoft Azure Previews include additional legal terms that apply to Azure features that are in beta, in preview, or otherwise not yet released into general availability.

The free version of Microsoft Purview governance solutions provides an automatically available Microsoft Purview instance, ready for your users without setup. It allows your organization to try Microsoft Purview's basic governance capabilities and begin your governance journey. This article offers some best practices to help you quickly get started with Microsoft Purview data governance solutions, the Data Map, and the catalog.

Access to this version of Microsoft Purview depends on your organization's current relationship with Microsoft Purview; whether you're a new Microsoft Purview user, if you have existing accounts, or if you're using other services like Microsoft Fabric. For more information, see how to access the free version of Microsoft Purview governance solutions..

How to start your governance journey

The free version of Microsoft Purview governance solutions lets you get started on your governance journey even as you begin planning. Once your organization has access to the free version of Microsoft Purview governance solutions, use this guide to build your governance experience.

1. Explore the Microsoft Purview portal.
2. See what metadata for supported Azure sources is already available in the catalog.
3. Search or browse the catalog to understand your initial assets.
4. Engage your data owners, and add data curators to manage metadata in your environment.

5. Work with your data owners and curators to augment the metadata for existing data sources. This gives your users more context and information about their data sources.

6. Add your other critical Azure resources to the data map using our REST APIs.

   > **Important**
   >
   > Only **data curators** will be able to access any resources added to the data map using the **REST APIs**. In order for other users to be able to access, **upgrade to the enterprise version of Microsoft Purview**.

7. Engage your users to discover your users to access your resources in the catalog.

When you're ready to expand

Once you've added and curated your critical Azure resources in the Microsoft Purview portal, if you're looking for more governance features, you're ready to upgrade to the enterprise version of Microsoft Purview.

When you upgrade, all the information you have and have added to the free version of Microsoft Purview governance solutions will be available in the enterprise version of Microsoft Purview within minutes, so you can get started using all of the enterprise features right away. Here's some steps you can take to prepare your environment for upgrade.

1. Review our deployment best practices to start the right conversations in your organization.
2. Consider your network architecture for Microsoft Purview.
3. Plan your collections architecture for your data map.
4. Upgrade to the enterprise version of Microsoft Purview.
5. Review our scanning and classification.

> **Tip**
>
> If you used the **REST APIs** to add sources and assets in the free version of Microsoft Purview governance solutions that you can now scan in the enterprise version, Microsoft Purview will **normalize** the newly scanned information and add it to your existing resources.

Next steps

- New Microsoft Purview portal
- Upgrade to enterprise
- Roles and permissions

Feedback

Was this page helpful? Yes No

Upgrade from the free to enterprise version of Microsoft Purview data governance

Article • 05/01/2025

> **Note**
>
> If you already have a Microsoft Purview account and enabled the new Microsoft Purview portal experience, you are using the enterprise version of Microsoft Purview. You can check your account type in the Microsoft Purview portal by selecting the **Settings** gear icon, then selecting **View all settings**.

When you upgrade from the free version of Microsoft Purview governance solutions to the enterprise version, the data in the free version will automatically be available in the enterprise version within minutes. Your experience in the portal remains the same, and you gain access to all Microsoft Purview data governance solutions' capabilities, including:

- All Microsoft Purview supported sources
- Collections
- Scanning
- Automatic classification
- Microsoft Purview policies
- All reports in Data Estate Insights
- Workflows

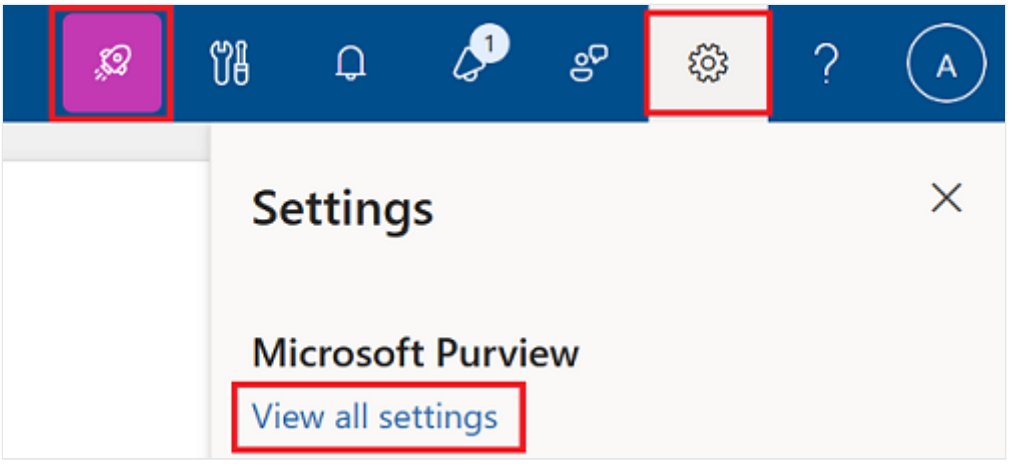
Learn more about getting started with Microsoft Purview data governance solutions.

> **Important**
>
> If you create a classic Microsoft Purview instance for the first time in your organization, the instance will automatically be upgraded to the enterprise version of Microsoft Purview.

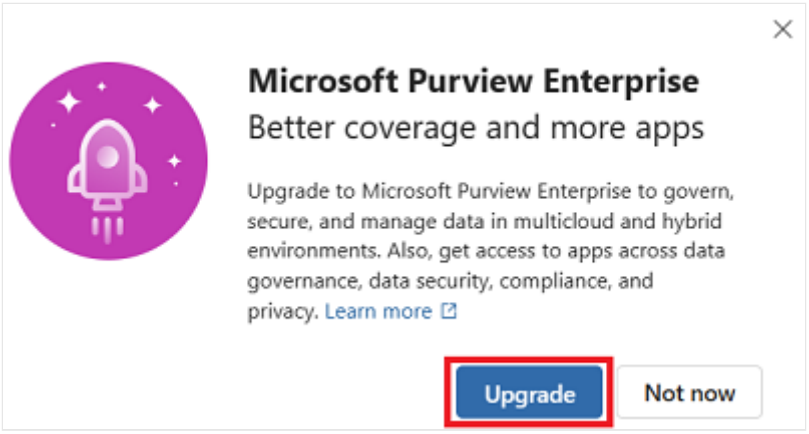
Upgrade

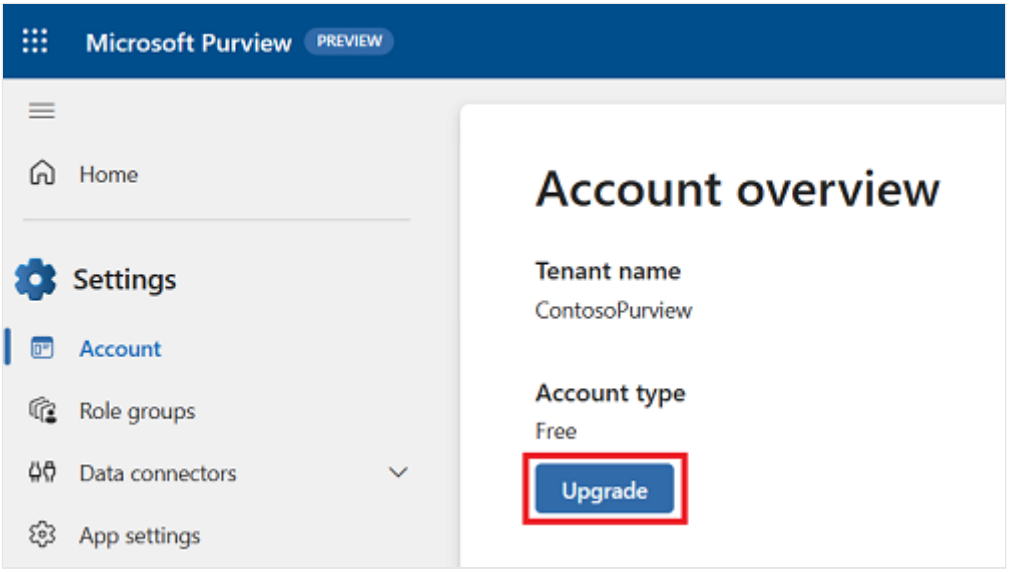
You can upgrade to the enterprise version of Microsoft Purview using these steps:

1. Open the Microsoft Purview portal
2. Select the upgrade button in the ribbon menu, or select **Settings**, then **View all settings**.



3. Select **Upgrade**





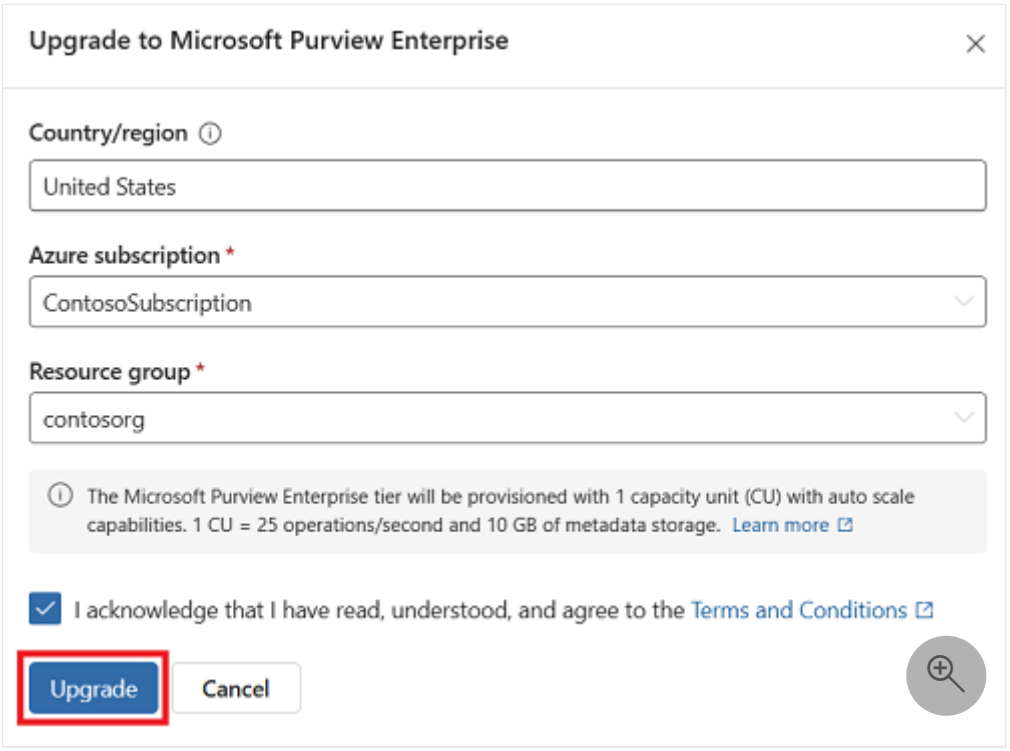
4. Your region is the same region as the free version of your Microsoft Purview account. For more information, see the FAQ.

5. Select the Azure subscription where you would like to house the Azure linked Microsoft Purview resource.

6. Select the resource group where you would like to house the Azure linked resource.

7. Select the check box to acknowledge terms and conditions.

8. Select **Upgrade**



9. You've upgraded! Select **Use default settings** to get started, or **Continue with setup** to manage settings.

Protect your account

In order to protect your Microsoft Purview account from accidental deletion, we recommend creation of the Azure Policy 'DenyAction'

JSON

```
{
  "if": {
    "allOf": [
      {
        "field": "type",
        "equals": "Microsoft.Purview/accounts"
      },
      {
        "field": "tags.environment",
        "equals": "Production"
      }
    ]
  },
  "then": {
    "effect": "denyAction",
    "details": {
      "actionNames": [
        "delete"
      ],
```

```
        "cascadeBehaviors": {
          "resourceGroup": "deny"
        }
      }
    }
  }
```

Settings

Governance settings in the new Microsoft Purview portal include:

- Account type - either free or enterprise. The free version only includes a subset of applications and sources. The enterprise version of Microsoft Purview includes all available governance applications and sources.
- Kafka configurations - configure Microsoft Purview to be able to send and receive Atlas Kafka topics events with Azure Event Hubs

Possible exceptions

When you're upgrading to the enterprise version of Microsoft Purview, there are a few possible exceptions you can encounter. These guides address them.

Private endpoints

If your Microsoft Purview account had disabled all public network access and is using private endpoints, you can still upgrade to the new experience.

You can continue to use your *ingestion* private endpoints as is, but the new experience also uses a *platform* private endpoint.

After upgrading to the new experience, you can follow this guide to set up platform the private endpoint for governance solutions in the Microsoft Purview portal.

For more information, see what to do if your account has disabled all public network access.

Resource providers

Before upgrading to the enterprise version of Microsoft Purview, you need to have these resource providers registered in Azure:

- Microsoft.Purview
- Microsoft.Storage
- Microsoft.EventHub

If all of them aren't already registered, you might receive an error.

Default name isn't allowed

It's possible that default names have been disallowed in your organization, and you need to enter a new resource name.

To resolve, enter a new resource name and then complete the upgrade process.

For more information about default naming policies, you can see the resource naming guidance documentation.

Region

If your account is in a different region than your tenant, an admin needs to confirm setup.

> **Note**
>
> If you want to use a different region, you will either need to cancel and wait for upcoming region related features, or cancel and create a new Microsoft Purview account in one of the available regions.

Nonadmin request

If a user who isn't a Purview Admin selects the upgrade button, a request to upgrade is sent to an administrator. The administrator can update from directly within the request email, or they can upgrade using the above steps.

If you don't have access to the Microsoft 365 role groups to assign Purview Admins, contact support.

Switch back to classic portal

Even after upgrading to the enterprise version of Microsoft Purview, you can switch back to the classic portal at any time using the toggle.



Billing

> **Note**
>
> The enterprise version of Microsoft Purview uses the same pricing dimensions as the classic Microsoft Purview governance portal experience. You can check your account type in the Microsoft Purview portal under **Settings**, then **View all settings**.

Billing for the enterprise version of Microsoft Purview is pay-as-you-go.

For more information, see these articles:

- Pricing guidelines
- Microsoft Purview elastic data map
- Pricing calculator

Linked resource

The enterprise version of Microsoft Purview is hosted by a linked resource in the Azure portal. If you delete this linked resource in the Azure portal, you lose all your Microsoft Purview information.

After deletion, you can reinstate the free version for your environment if it's been enabled for your organization by going to the Microsoft Purview portal. But you're starting your environment from scratch.

Next steps

- Microsoft Purview portal
- Roles and permissions
- Scans and ingestion
- Microsoft Purview policies

Learn about the Microsoft Purview Suite trial

The Microsoft Purview Suite trial is a free and easy way to try compliance capabilities of Microsoft Purview solutions. After a quick setup taking only a couple of minutes, the features of the Microsoft E5 license package are available for you to use for up to 90 days.

Where to sign up

Sign up in the Microsoft Purview portal. When you start a trial, the trial is effective immediately for 90 days. It can take up to two hours for all Microsoft Purview solutions to appear in your left navigation. Sign out and sign back in to see the updates.

- Sign in to the Microsoft Purview portal.
- On the home page, go to the **Trials and recommendations** section and select **View all trials and recommendations**.
- In the **Purview and Priva trials** section, find the **Microsoft Purview setup** card and select **Get started**.

At the hub you can see the status of your trials, sign up for other trials, and view purchase options.

Eligibility and licensing

Customers who don't already have a Microsoft E5 license package are eligible for the trial if they meet one of these two requirements:

- Have Microsoft 365 E3, **or**
- Have Office 365 E3 **and** Enterprise Mobility and Security E3

The trial isn't available for Microsoft 365 Government customers.

As part of the trial setup, 25 Microsoft Purview Suite (formerly known as Microsoft 365 E5 Compliance) licenses are automatically applied to your organization. The licenses are active for 90 days. You might want to refer to the Microsoft Purview service description for licensing details for each solution.

Roles for starting the trial

A user with one of the roles listed below can start or end a trial:

- Billing Administrator
- Compliance Administrator
- Compliance Data Administrator
- Global Administrator

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

Get more details about roles for starting trials.

Getting started, finding support, and ending the trial

Getting started

You can start using Microsoft Purview solutions as soon as you see them in the Microsoft Purview portal. Nothing is created automatically and users won't be affected. When you navigate to each solution, you may be guided to make extra setup configurations to start using features.

Finding support

Organizations can use FastTrack for trial deployment support if the tenant has at least 150 licenses. Get more details about FastTrack.

Extending the trial

You can extend the trial within the last 15 days of the trial period. You're limited to a maximum of two trial periods. If you don't extend by the time your trial period ends, you need to wait at least 30 days before signing up for a second trial.

Ending the trial

Admins can disable the trial anytime by selecting **End trial** on the **Purview setup** card on the **Microsoft Purview trials and recommendations** page in the Microsoft Purview portal.

If you decide not to upgrade to the E5 level solutions, unless stated otherwise for the solution your trial data will be maintained for a period of time, usually 180 days, before being permanently deleted. You may continue to access the data gathered during the trial until that time.

Learn more about Microsoft Purview solutions

The Microsoft Purview Suite trial includes the solutions listed below.

> **Note**
>
> Certain Microsoft Purview features may appear in other Microsoft products for which a license is needed. Refer to each Microsoft Purview solution's licensing details at **Microsoft 365 licensing guidance for security & compliance**.

- **Audit**

  Microsoft Purview Audit (Premium) helps organizations to conduct forensic and compliance investigations by increasing audit log retention required to conduct an investigation, providing access to intelligent insights that help determine scope of compromise, and faster access to Office 365 Management Activity API. Learn more about Audit (Premium).

- **Communication Compliance**

  Microsoft Purview Communication Compliance helps you overcome modern compliance challenges associated with internal and external communications by helping you automatically capture inappropriate messages, investigate possible policy violations, and take steps to remediate. Learn more about Communication Compliance.

- **Compliance Manager**

  Microsoft Purview Compliance Manager can help you throughout your compliance journey, from taking inventory of your data protection risks to managing the complexities of implementing controls, staying current with regulations and certifications, and reporting to auditors. Learn more about Compliance Manager.

- **Data Lifecycle Management**

  Automate your retention coverage by using adaptive policy scopes. This feature allows you to dynamically target policies for retention to specific users, groups, or sites. These policies automatically update when changes occur in your organization. In addition,

policies that use adaptive scopes aren't subject to location limits. Learn more about adaptive policy scopes.

Learn more about Microsoft Purview Data Lifecycle Management.

- **Data Loss Prevention**

  Microsoft Purview Data Loss Prevention (DLP) helps you control sharing and use of sensitive info on devices, apps, and services. With trials, you can extend coverage of DLP policies to Teams and devices, protect even more sensitive content using trainable classifiers and exact data match (EDM), and get started quickly with our recommended DLP policies. Learn more about Data Loss Prevention.

  When the trial ends:
  - Your existing policies stay turned on unless you turn them off.
  - You won't be able to edit or add Teams or Devices locations in existing or new policies.

- **Data Security Posture Management**

  Microsoft Purview Data Security Posture Management (DSPM) helps organizations discover, protect, and investigate sensitive data risks across their digital estate. This solution provides unified visibility and control for both traditional applications and AI apps and agents, supporting data governance across Microsoft 365 and integrated third-party SaaS platforms. Monitor, assess, and remediate data risks, regardless of where sensitive data resides.

  Learn more about Microsoft Purview Data Security Posture Management.

- **eDiscovery**

  Take advantage of an end-to-end workflow for preserving, collecting, analyzing, and exporting content that's responsive to your organization's internal and external investigations. Legal teams can also manage the entire legal hold notification process by communicating with custodians involved in a case. Learn more about eDiscovery.

- **Information Protection**

  Implement Microsoft Purview Information Protection with sensitivity labels to help you discover, classify, and protect your sensitive content wherever it lives or travels. With the trials, you can recommend users apply sensitivity labels to their emails and docs, discover and automatically label sensitive content, take advantage of machine learning based classifiers, and get started quickly with some recommended labels and policies.

  When the trial ends:

    - All files and emails labeled during your trial stay labeled. You can manually remove the labels.
    - Your existing policies stay turned on unless you turn them off.
    - Any autolabeling policies can't be edited after the trial ends, but can be deleted.

    For more information about each of these preconfigured features and how they affect users, see Learn about the free trial for Microsoft Purview Information Protection.

Get more information about the full range of features for Microsoft Purview Information Protection.

- **Insider Risk Management**

  Use artificial intelligence to help you quickly identify, triage, and remediate internal risks. Using logs from Microsoft 365 and Azure services, you can define policies that monitor for risk signals, then take remediation actions such as promoting user education or initiating an investigation. Learn more about Insider Risk Management.

- **Records Management**

  Use integrated Microsoft Purview Records Management features to automate the retention schedule for organizational regulatory, legal, and business-critical records. Get full content lifecycle support, from creation to collaboration, record declaration, retention, and disposition:
    - Classify content as a record to prevent users from editing, as required by regulations, laws, or organizational policy
    - Apply retention labels to content automatically when it matches criteria you specify, using autoapply label policies
    - Use adaptive scope policies to dynamically target your retention label policies to locations, with no limit on how many locations are included
    - Get full content lifecycle support, including the ability to perform disposition review on contents before they're permanently deleted at the end

Get more information on the full range of features for Microsoft Purview Records Management.

Terms and conditions

See the terms and conditions for Microsoft 365 trials.

Last updated on 03/30/2026

Learn about the Microsoft Purview Compliance Manager premium assessments trial

07/24/2025

Compliance Manager helps you automatically assess and manage compliance across your multicloud environment. The assessments you build are based on templates that correspond to governmental regulations and industry standards around the world.

Your licensing agreement to use Compliance Manager may include one or more built-in regulatory templates for building assessments. An extensive library of premium regulations is also available for building assessments specific to your organization's needs. To try out these regulatory templates before purchasing licenses, sign up for the free Compliance Manager premium assessments trial.

> **Tip**
>
> If you're not an E5 customer, use the 90-day Microsoft Purview solutions trial to explore how additional Purview capabilities can help your organization manage data security and compliance needs. Start now at the **Microsoft Purview trials hub**. Learn details about **signing up and trial terms**.

Trial setup instructions

- Sign in to the Microsoft Purview portal.
- On the home page, go to the [**Trials and recommendations**](purview-portal.md#trials-and-recommendations section and select **View all trials and recommendations**.
- In the **Purview and Priva trials** section, find the **Compliance assessments** card and select **Try now**.

The premium assessment trial is available to organizations using Compliance Manager under a commercial license. For GCC and DOD information and trial options, see Learn about regulations in Compliance Manager.

Full access to Compliance Manager features depends on your licensing agreement for Compliance Manager. To learn more about the free compliance trial that includes Compliance Manager, see Learn about the Microsoft Purview risk and compliance solutions trial.

> **Note**
>
> Check your Microsoft 365 admin center settings to facilitate enrollment in the trial. In the Microsoft 365 admin center, go to the **Settings** > **Org settings** > **Services** page, and select **User owned apps and services**. Then check the box to let users start trials on behalf of your organization. When setup finishes in about two minutes, you can go back and uncheck the box to disable this setting. You can also start the trial without adjusting your org settings by going to the **Billing** > **Purchase services** page and searching for the Compliance Manager premium assessment trial.

Regulations available to use during the trial

When you start your one-time trial for premium assessments, you can use 25 premium regulatory templates of your choice for free for 90 days. You can create as many assessments from the 25 regulations as you wish. View the full list of Compliance Manager regulations.

After starting your trial, go to your **Assessments** tab in Compliance Manager and view the **Purchased regulation licenses used** counter to see how many regulations are available and are in use. This counter shows your 25 licensed regulatory templates. As you create assessments for these regulations, your activated number updates. Multiple assessments based on the same regulation count as one regulatory template use. For more information, see Learn about regulations in Compliance Manager.

Roles for starting the trial

A user with one of the roles listed below can start or end a trial:

- Billing Administrator
- Compliance Administrator
- Compliance Data Administrator
- Global Administrator

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

Get more details about roles for starting trials.

During your trial

Throughout your trial, you can review the **Trial summary** section of the **Overview** page in Compliance Manager to see how many assessments you created, the number of improvement actions taken, how your compliance score has changed, and how much time remains in your trial. From here, you can also review the list of available regulations and learn more about purchasing them for ongoing use.

After the trial

After your trial ends, the regulatory templates you used won't receive automatic updates and might fall out of date with any regulatory changes. We recommend that you delete assessments built from regulatory templates that you don't intend to purchase.

Ending the trial

Admins can disable the trial anytime by selecting **End trial** on the **Compliance assessments trial** card on the **Microsoft Purview trials and recommendations** page in the Microsoft Purview portal.

Learn more and explore other trials

To learn more about the features of Compliance Manager and how it can help your organization meet compliance and regulatory requirements, see Microsoft Purview Compliance Manager.

You can also learn about other Microsoft Purview solutions and how to start trials at Learn about the Microsoft Purview risk and compliance solutions trial.

Terms and conditions

See the terms and conditions for Microsoft 365 trials.

How data flows in Microsoft Purview

This article describes how Microsoft Purview shares data between Purview solutions. This information supplements the commitments made to you in the relevant sections of the Microsoft Product Terms and the Microsoft Products and Services Data Protection Addendum.

The Data Map provides a unified map of your data assets and can identify the assets' relationships for effective governance. Microsoft Defender for Cloud Apps gives you full visibility of your SaaS app landscape to help protect your apps.

When you create an enterprise account, either by upgrading a free tier account or by creating a new enterprise account, you can register data sources within your data estate. By registering data sources, users in the enterprise account can scan the registered data sources. After you complete registration and scan activities for a data source, data flows between Data Map and some Microsoft Purview solutions.

> **Important**
>
> If you don't want to authorize data flows throughout Microsoft Purview, don't register or scan any of the data sources in your data estate into Data Map.

The following table describes how Microsoft Purview solutions share data with Data Map and receive data from Data Map and Microsoft Defender for Cloud Apps.

Expand table

Microsoft Purview Solution	What gets shared to Data Map	What is received from Data Map	What is received from Microsoft Defender for Cloud Apps
Microsoft Purview Audit	N/A	- Service configuration data - Audited activities data (including audit records and log query permissions) - Admin operations data	N/A
Microsoft Purview Unified Catalog	N/A	- Asset details, including asset metadata, classification data,	N/A

Microsoft Purview Solution	What gets shared to Data Map	What is received from Data Map	What is received from Microsoft Defender for Cloud Apps
		and label data	
Microsoft Purview Data Loss Prevention (DLP)	- DLP policies and configuration	- Classification data - Policy hit audit events and policy distribution acknowledgment status	N/A
Microsoft Purview Information Protection	- Label definitions, policies, and configurations -Label insights and aggregations for Microsoft 365 - Microsoft Information Protection policies and configurations - Classification data for Microsoft 365 - Asset metadata for Microsoft 365 - Telemetry data for protection policies and labels	- Label insights and aggregations - Classification data - Tenant roles, role groups, and license information - Scan setup and execution details - Asset metadata (owner and additional information including name, path, consent status) - Policy hit audit events and policy distribution acknowledgment status	N/A
Microsoft Purview Insider Risk Management	N/A	N/A	- Cloud app events (activity log) and connection status
Microsoft Purview Policy	N/A	- Label names, policies and Ids - Policy definition	N/A

Last updated on 04/01/2026

Microsoft Purview extensibility and Development Partners

Microsoft Purview solutions help organizations intelligently assess their compliance risks, govern and protect sensitive data, and effectively respond to regulatory requirements. Microsoft Purview offers many extensibility scenarios and enables organizations to adapt, extend, integrate, accelerate, and support their compliance solutions.

Three key building blocks support extensibility in Microsoft Purview:

- **Data connectors**. Use these connectors to import and archive non-Microsoft data so you can apply Microsoft 365 protection and governance capabilities to third-party data.
- **APIs**. APIs enable programmatic access to Microsoft Purview capabilities.
- **Software Developer Partner Integrations:** Microsoft Purview enables software development companies to embed advanced data security capabilities directly into their applications.

Data connectors

Microsoft provides third-party data connectors that you can configure in the Microsoft Purview portal. For a list of data connectors provided by Microsoft, see the Third-party data connectors table. The table of third-party data connectors also summarizes the compliance solutions that you can apply to third-party data after you import and archive data in Microsoft 365, and links to the step-by-step instructions for each connector.

Prerequisites for data connectors

Many of the data connectors available in the Microsoft Purview portal allow you to import and archive third-party data but require you to prepare and perform configuration tasks in the third-party data source. The documentation details these prerequisites for each third-party data connector.

For data connectors in the Microsoft Purview portal or the Microsoft Purview portal provided by one of Microsoft's partners, your organization needs a business relationship with the partner before you can deploy a connector.

For guidance and requirements for third-party data connectors, see the "Data connectors" section in Microsoft 365 guidance for security & compliance.

APIs

Microsoft Purview and Microsoft Priva APIs are available in the Microsoft Information Protection SDK, Microsoft Graph API, and the Office 365 Management Activity API. Some compliance APIs are part of a new set of security and compliance APIs that enable developers for Microsoft 365 customers, independent software publishers, system integrators, and managed security service providers to build high-value security and compliance solutions.

To learn more about how to access Graph APIs, see Overview of Microsoft Graph.

Microsoft Information Protection (MIP) SDK

The MIP SDK exposes the labeling and encryption services from the Microsoft Purview portal to third-party applications and services. Developers can use the SDK to build native support for applying labels and protection to files. Developers can determine which actions to take when specific labels are detected, and reason over MIP-encrypted information.

High-level MIP SDK use cases include:

- A line-of-business application that applies classification labels to files on export.
- A CAD/CAM design application that provides native support for sensitivity labels.
- A cloud access security broker or data loss prevention solution that can encrypt data with rights management.

For more information about the MIP SDK, prerequisites, additional scenarios, and samples, see MIP SDK Overview.

Microsoft Graph API for Teams DLP

Data loss prevention (DLP) capabilities are widely used in Microsoft Teams, particularly as organizations shift to remote work. Recently, we announced the general availability of the Microsoft Graph Change Notification API for messages in Teams. This API enables developers to build apps that can listen to Microsoft Teams messages in near-real time and then implement DLP scenarios for both customers and partners. Additionally, Microsoft Graph Patch API lets you apply DLP actions to Teams messages.

These two APIs form the Microsoft Graph API for Teams DLP. You can get started by trying out the sample app. For more information about Microsoft Teams messaging webhooks, see the documentation.

For the licensing requirements for Teams DLP, see Microsoft 365 licensing guidance for security & compliance.

Microsoft Graph API for eDiscovery

With eDiscovery, organizations can discover data where it lives and manage more end-to-end eDiscovery workflows with intelligent machine learning and analytics capabilities to reduce data to the relevant set – all while the data stays within the Microsoft 365 security and compliance boundary.

You can use Graph APIs for eDiscovery to create and manage cases, review sets, and review set queries in a scalable and repeatable manner. This approach enables customers and partners to create apps and workflows to automate common and repetitive processes such as creating cases and managing custodians and legal holds.

For the licensing requirements for eDiscovery and the API, see the "eDiscovery" section in the Microsoft 365 licensing guidance for security & compliance.

Microsoft Graph API for Teams Export

Enterprise Information Archiving (EIA) for Microsoft Teams is a key scenario for our customers as it allows them to solve for regulatory requirements. In addition to our built-in capabilities for archiving content in Microsoft Teams, customers and partners can now use Teams Export APIs to solve for custom application and integration scenarios. The Teams Export APIs support bulk-export (up to 200 requests per second/per app/per tenant) of Teams messages and message attachments. Deleted messages are also accessible by the API for up to 30 days after they're deleted. For more information about these Teams Export APIs and how to use them in your applications, see Export content with the Microsoft Teams Export APIs.

For the licensing requirements for the use of the Teams Export APIs, see Microsoft 365 licensing guidance for security & compliance.

Microsoft Graph Connector APIs (preview)

With Microsoft Graph connectors, organizations can index third-party data so it appears in Microsoft Search results. This feature expands the types of content sources that are searchable in

your Microsoft 365 productivity apps and the broader Microsoft ecosystem. The third-party data can be hosted on-premises or in public or private clouds. Starting with eDiscovery (Premium), we're enabling developer preview of built-in compliance value of Microsoft 365 connected apps. This capability enables compliance for apps integrating into the Microsoft 365 ecosystem to empower users with seamless compliance experiences. To learn more about how to incorporate Microsoft Graph Connector APIs in your apps, see Create, update, and delete connections in the Microsoft Graph.

Microsoft Graph API for records management

Organizations of all types need a records management solution to manage critical records across their data. Microsoft Purview Records Management helps an organization manage its legal obligations, provides the ability to demonstrate compliance with regulations, and increases efficiency with regular disposition of items that are no longer required.

Organizations use the records management solution to protect, label, retain, or delete their data. The Microsoft Graph APIs for records management help organizations manage retention labels and their associated actions more efficiently, automate repetitive tasks, and equip customers with flexibility in options.

The first release of Graph APIs for records management supports the management of retention labels and event-based retention. Example scenarios include:

- **Managing retention labels**

  Record management admins and developers need to maintain their record management systems with labels that they periodically create, update, and delete.

  Developers and compliance admins use the Graph APIs for records management to perform CRUD operations on the label entity to maintain their systems.

- **Triggering an event for an existing label**

  When an employee leaves an organization, the information is updated in the HR management system. From the date of leaving, confidential documents need to be retained for seven years. These documents already have the retention label "Employee_departure" applied to them.

  Developers and compliance admins use the Graph APIs for records management to read the label “Employee_departure” and look up the associated event type "Event-employee_departure".

They then use the Graph APIs for records management to create an event for the associated event type. The retention period for the confidential documents starts after this event is created.

For more information about the Graph APIs for records management, see Use the Microsoft Graph Records Management API.

For licensing requirements to use these APIs, see the records management information from the Microsoft 365 guidance for security & compliance, Microsoft Purview Data Lifecycle Management & Microsoft Purview Records Management section.

Microsoft Graph API for retention labels in SharePoint and OneDrive for Business

Retention labels are part of the Microsoft Purview Data Lifecycle Management solution and apply governance at the driveitem level. Retention labels support more capabilities than retention policies and can handle exceptions within a location. For more information on retention labels, see Create retention labels for exceptions to your retention policies.

With the Microsoft Graph APIs for retention labels, organizations can programmatically apply and manage these labels on items in SharePoint and OneDrive for Business to automate their processes.

These Graph APIs support the following tasks:

- **Apply a retention label to an item**
- **Remove a retention label from an item**
- **Get metadata information on the retention label applied to an item**
- **Lock or unlock the record label applied to an item**

For licensing requirements to use these APIs, see the records management information from the Microsoft 365 guidance for security & compliance, Microsoft Purview Data Lifecycle Management & Microsoft Purview Records Management section.

Software Developer Partner Integrations

Microsoft Purview enables software development companies to embed advanced data security capabilities directly into their applications. Using available SDKs and APIs, developers can build solutions that apply classification, labeling, protection, and retention policies natively within their workflows. These integrations enable automation, consistent policy enforcement, and unified

governance across diverse environments—helping companies meet regulatory requirements and streamline compliance processes at scale.

Network Data Security

Microsoft Purview Network Data Security enables real-time protection by intercepting and inspecting network traffic using the Purview DLP engine. Through deep integration with SASE and next-generation firewall platforms, it extends Purview classification and policy enforcement capabilities to the network layer based on customer-defined data loss prevention and collection policies. This enables:

- Granular control over sensitive data flows across both sanctioned and unsanctioned applications.
- Blocking exfiltration and unauthorized sharing with generative AI tools, cloud services, storage platforms, and social media.
- Deep integration with Purview insights and solutions, including Data Security Posture Management (DSPM) and Insider Risk Management (IRM).
- Stronger compliance for AI interactions through content capture that supports eDiscovery, retention, deletion, and related regulatory workflows.

The following list of partners have successfully built a Network Data Security integration:

Expand table

Partner	Partner documentation	Description
iboss	iboss deployment guide	iboss Zero Trust Integrates with Purview DLP for secure access and network data classification.
Netskope	Netskope deployment guide	Netskope One Integrates with Purview DLP to inspect and classify network traffic for data protection.

For more information on building integrations with Microsoft Purview, see Use Microsoft Purview APIs to support data security and governance in your apps. To get listed in the above table of integration partners, register your interest using the following form: Microsoft Purview Network Data Security partner integration checklist.

Microsoft Purview for Generative AI Apps integration

As generative AI tools become embedded in business processes, Microsoft Purview enables security, compliance, and data teams to discover and monitor AI usage, safeguard sensitive information in prompts and outputs, and enforce organizational policies—helping organizations innovate responsibly while reducing risk and meeting regulatory requirements. Developers can use the Microsoft Purview SDK to integrate these protections directly into their AI solutions, extending Purview's coverage across the enterprise AI ecosystem. Visit the following documentation for guidance on building this integration: Use the Microsoft Purview SDK with Agent Framework.

The following partner has successfully built a Microsoft Purview for Generative AI App integration.

Expand table

Partner	Partner documentation
Miro	Miro deployment guide

To get listed in the above table of integration partners, register your interest using the following form: Microsoft Purview for Generative AI apps partner integration checklist.

Microsoft Information Protection (MIP)

Many vendors integrate with Purview Information Protection to unify classification and protection across platforms. For a comprehensive list of MIP development partners, visit Microsoft Intelligent Security Association.

For information on building with Microsoft Information Protection see the Microsoft Information Protection SDK documentation.

Last updated on 05/11/2026

Implementing Zero Trust with Microsoft Purview

Microsoft Purview solutions can help you implement a Zero Trust security strategy that is based on the following security principles:

Expand table

Verify explicitly	Use least privilege access	Assume breach
Always authenticate and authorize based on all available data points.	Limit user access with Just-In-Time and Just-Enough-Access (JIT/JEA), risk-based adaptive policies, and data protection.	Minimize blast radius and segment access. Verify end-to-end encryption and use analytics to get visibility, drive threat detection, and improve defenses.

Microsoft Purview is a primary component of the **Use least privilege access** principle by providing data protection solutions. Use Purview capabilities to help you safeguard your data across platforms, apps, and clouds.

Zero Trust to protect data

Microsoft Purview provides the following capabilities and options for a data defense in depth strategy and a Zero Trust implementation for data protection:

- **Data classification**: So you can know your data

  Discover and detect sensitive data across your entire organization, so that you can better protect it. For more information, see Classifiers overview and Reports.

- **Information protection**: So you can protect your data

  Apply sensitivity labels to integrate with Microsoft 365 Copilot and other apps and services to provide access control guardrails, and encryption with rights management for your most sensitive data. Content markings, such as footers and watermarks, can increase awareness and security policy compliance. While users create or update content, the highly visible labels and labeling recommendations support user education about sensitive data. For more information, see Learn about sensitivity labels.

  When you use sensitivity labels with protection policies, you can automatically enforce access restrictions across your data estate the moment sensitive information is detected.

- **Data loss prevention (DLP)**: So you can prevent data loss

Users sometimes take risks with your organization's sensitive data, which might result in a data security or compliance incident. Data loss prevention helps you monitor for and protect against risky oversharing of sensitive data. As with sensitivity labeling, policy tips support user education about sensitive data. For more information, see Learn about data loss prevention.

- **Insider Risk Management**: So you can identify and take action against risky security-related user activities and data activity patterns

  Correlate various signals to identify and take action on potential malicious or inadvertent insider risks, such as IP theft, data leakage, and security violations. For more information, see Learn about Insider Risk Management.

- **Data lifecycle management**: So you can delete what you don't need, and safeguard important data

  Deploy policies to manage the lifecycle of sensitive data to reduce data exposure. Limit the number of copies or propagation of sensitive data by automatically and permanently deleting it when it's no longer needed. Or conversely, protect important data from malicious or accidental deletes by automatically retaining a copy in a secured location after a user deletes the data. For more information, see Learn about data lifecycle management.

Supporting tools and technologies:

> **Note**
>
> Now in preview, use Microsoft Data Security Posture Management (DSPM) to help you discover, protect, and investigate sensitive data risks across your digital estate. Unifying other Microsoft Purview solutions, DSPM provides visibility and control for both traditional applications and AI apps and agents, so you can monitor, assess, and remediate data risks regardless of where sensitive data resides.

- As you implement these capabilities, use appropriate role-based permissions and administrative units to provide Just-Enough-Access and segment access. Augment these protective measures with privileged access management for Just-In-Time access.

- Consider your encryption requirements for specific scenarios, for example:
  - Microsoft 365 Copilot
  - Use your sensitivity labels to apply Double Key Encryption to selected documents and emails when only your organization and no cloud services should be able to decrypt them.

    - Use Advanced Message Encryption if you need to keep sensitive content within your organization boundary, log external mail access, or revoke access to encrypted emails.
    - Use Customer Key if you need to control the root encryption keys for Microsoft 365 data at-rest.
    - If you use Conditional Access or cross-tenant access settings, these services need specific configurations to support encrypted content.

- For high-value documents and emails, records management supports additional restrictions and a disposition review process.

- Consider using information barriers if you need to segment access between specific users by restricting two-way communication and collaboration between groups and users in Microsoft Teams, SharePoint, and OneDrive.

- Additional governance options:
    - Access policies in the Microsoft Purview Unified Catalog allow users to request access to data, but also give you the tools to ensure they only have access to the data they need, and only for as long as they need it.
    - Use the data sharing app to minimize data duplication and instead, provide read-only access that you can time-limit or remove the access.

- Consider using Compliance Manager to help drive the adoption of and monitor the implementation of security features and configurations. Easy-to-build assessments with automatic monitoring help you stay on track with requirements across your multicloud environment.

- Use auditing solutions to help you monitor Microsoft 365 data and respond to security events.

- Use Customer Lockbox to ensure Microsoft service engineers must obtain approval before accessing any Microsoft 365 data you own during a support investigation.

Next steps

Solution guidance to help you implement a Zero Trust strategy for data protection by using Microsoft Purview:

- Deploy an information protection solution with Microsoft Purview
- Deploy a data governance solution with Microsoft Purview

Because data protection helps protect personal data stored and managed by your organization, see also Manage data privacy and data protection with Microsoft Priva and Microsoft Purview.

Zero Trust solution guidance:

- For data protection concepts and deployment objectives, see Secure data with zero trust
- For other capabilities that contribute to a strong Zero Trust strategy and architecture for your Microsoft 365 data, see Zero Trust deployment plan with Microsoft 365
- To learn how to apply Zero Trust to Microsoft 365 Copilot, see Apply principles of Zero Trust to Microsoft 365 Copilot.

Learn more about Zero Trust and how to build an enterprise-scale strategy and architecture with the Zero Trust Guidance Center.

Last updated on 12/15/2025

Help dynamically mitigate risks with Adaptive Protection

> **Important**
>
> Microsoft Purview Insider Risk Management correlates various signals to identify potential malicious or inadvertent insider risks, such as IP theft, data leakage, and security violations. Insider Risk Management enables customers to create policies to manage security and compliance. Built with privacy by design, users are pseudonymized by default, and role-based access controls and audit logs are in place to help ensure user-level privacy.

Adaptive protection in Microsoft Purview uses machine learning to identify the most critical risks and proactively and dynamically apply protection controls from:

- Microsoft Purview Data Loss Prevention (DLP)
- Microsoft Purview Data Lifecycle Management (preview)
- Microsoft Entra Conditional Access (preview)

Integration with data loss prevention, data lifecycle management, and Conditional Access helps organizations automate their response to insider risks and reduces the time required to identify and remediate potential threats. By using the capabilities of all four solutions, organizations can create a more comprehensive security framework that addresses both internal and external threats.

> **Important**
>
> To learn about licensing requirements for using Adaptive Protection with data loss prevention, data lifecycle management, and Conditional Access, see Microsoft 365, Office 365, Enterprise Mobility + Security, and Windows 11 Subscriptions.

Adaptive protection helps mitigate potential risks by using:

- **Context-aware detection**. Helps identify the most critical risks with ML-driven analysis of both content and user activities.
- **Dynamic controls**. Helps enforce effective controls on high-risk users while others maintain productivity.

- **Automated mitigation**. Helps minimize the impact of potential data security incidents and reduce admin overhead.

Adaptive protection dynamically assigns appropriate data loss prevention, data lifecycle management, and Conditional Access policies to users based on the insider risk levels defined and analyzed by the machine learning models in Insider Risk Management. Policies become adaptive based on user context, ensuring that the most effective policy, such as blocking data sharing through data loss prevention or blocking application access through Conditional Access, is applied only to high-risk users while low-risk users maintain productivity. Data loss prevention and Conditional Access policy controls constantly adjust, so when a user's insider risk level changes, an appropriate policy dynamically applies to match the insider new risk level. In the case of data lifecycle management, Insider Risk Management detects *Elevated risk level* users and preserves their deleted data for 120 days by using an automatically created data lifecycle management retention label policy.

> **Important**
>
> Insider Risk Management is currently available in tenants hosted in geographical regions and countries supported by Azure service dependencies. To verify that the Insider Risk Management solution is supported for your organization, see Azure dependency availability by country/region. Insider Risk Management is available for commercial clouds but isn't available for US Government cloud programs at this time.

Watch the following video for a summary of how Adaptive Protection can help identify and mitigate the most critical risks in your organization:

https://learn-video.azurefd.net/vod/player?id=9487d97b-4306-4a17-abb8-1008d4c63325&locale=en-us&embedUrl=%2Fpurview%2Finsider-risk-management-adaptive-protection

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Insider risk levels and preventive controls

With Adaptive Protection, administrators can configure risk factors or activities for customizable insider risk levels based on your organization's needs. The insider risk levels for Adaptive Protection update continuously and automatically based on the users' risk factors and insights. When users' data security risks increase or decrease, their insider risk levels adjust accordingly. Based on the insider risk levels, data loss prevention policies and Conditional Access policies automatically apply the right level of preventative controls as configured by admins, such as *block*, *block with override*, or *warning*.

> **Tip**
>
> *Insider risk levels* in Adaptive Protection differ from *alert severity levels* that assign Insider Risk Management users based on activity detected in Insider Risk Management policies.
>
> - Insider risk levels (Elevated, Moderate, or Minor), as described in this article, measure risk determined by admin-defined conditions such as the number of exfiltration activities users perform in a day or whether their activity generated a high-severity insider risk alert.
> - Alert severity levels (Low, Medium, or High) assign to users based on activity detected in Insider Risk Management policies. These levels calculate based on alert risk scores assigned to all active alerts associated with the user. These levels help insider risk analysts and investigators prioritize and respond to user activity accordingly.

For data lifecycle management, the data lifecycle management policy monitors users that assign an *Elevated risk level* by Insider Risk Management Adaptive Protection. When a risky user deletes any content from SharePoint, OneDrive, or Exchange Online, the contents are automatically preserved for 120 days. Admins can contact Microsoft support to restore any preserved content.

Depending on the Insider Risk Management policy assigned in Adaptive Protection, different criteria (users, groups, indicators, thresholds, etc.) are used to determine applicable insider risk levels. Insider risk levels are based on user insights, not solely on the number of instances of specific user activities. Insights are a calculation of the aggregate number of activities and the severity level of these activities.

For example, insider risk levels for User A wouldn't be determined by User A performing a potentially risky activity more than three times. The insider risk levels for User A would be determined by an insight of the aggregate number of activities and risk scores would be assigned to the activity based on the thresholds configured in the selected policy.

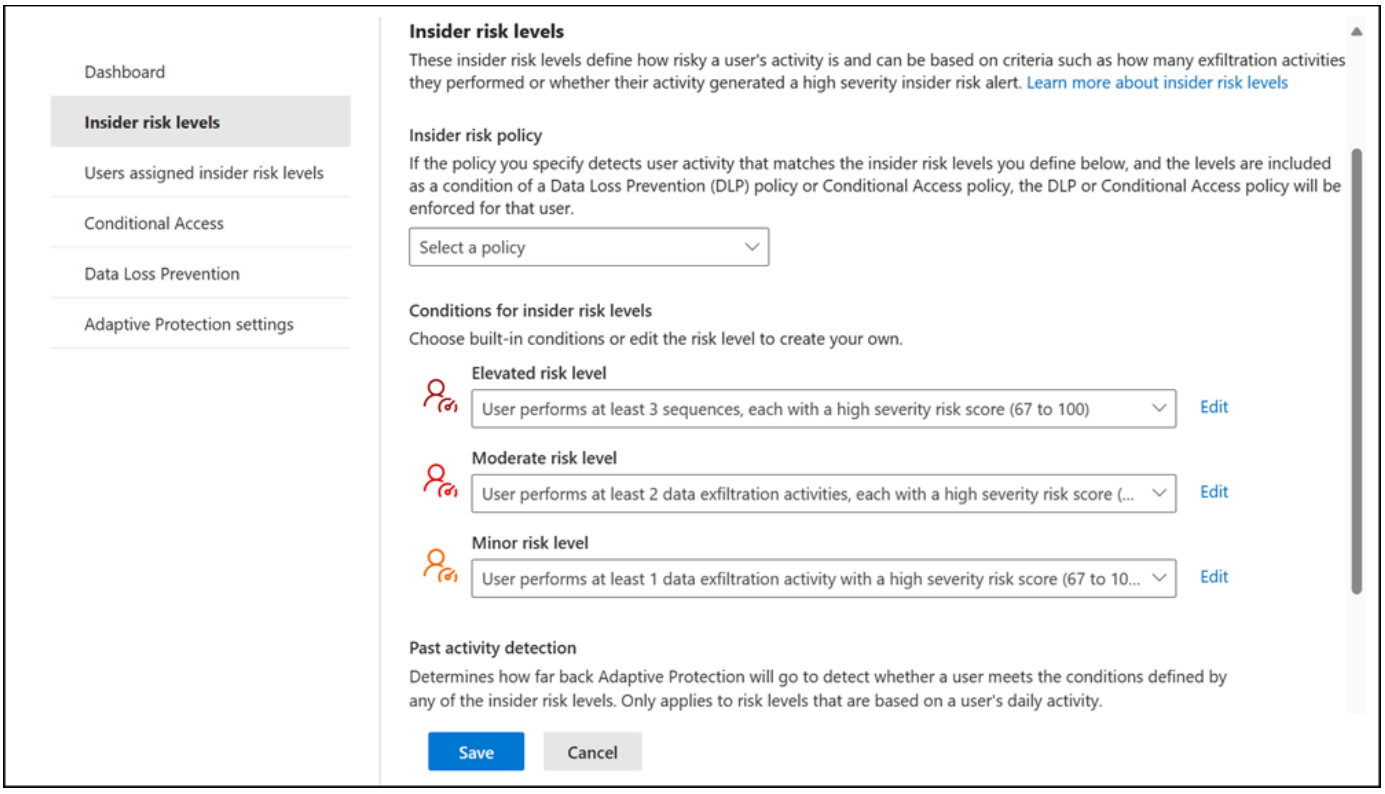
Insider risk levels

Insider risk levels in Adaptive Protection define how risky a user's activity is and can base on criteria such as how many exfiltration activities they performed or whether their activity generated a high severity insider risk alert. These insider risk levels have built-in insider risk level definitions, but you can customize these definitions as needed:

- **Elevated risk level**: This risk level includes built-in definitions for users with high severity alerts, users with at least three sequence insights that each have a high severity alert for specific risk activities, or one or more confirmed high severity alerts.
- **Moderate risk level**: This risk level includes built-in definitions for users with medium severity alerts or users with at least two data exfiltration activities with high severity scores.
- **Minor risk level**: This risk level includes built-in definitions for users with low severity alerts or users with at least one data exfiltration activity with a high severity score.

For an insider risk level to assign to a user, the number of insights and the severity assigned to the activity need to match the definition for the insider risk level. The number of activities for an insight might be a single activity or multiple activities accruing to the single insight. The number of insights evaluates for the insider risk level definition, not the number of activities contained in an insight.

For example, suppose the conditions in the Insider Risk Management policy assigned to Adaptive Protection scope for identifying downloads from SharePoint sites in your organization. If the policy detects that a user downloaded 10 files from a SharePoint site in a single day that are determined to be high severity, this counts as a single insight that consists of 10 activity events. To qualify for assigning an *Elevated risk level* to the user, two additional insights (with high severity) require for the user. The additional insights might or might not contain one or more activities.



Customizing insider risk levels

Custom insider risk levels enable you to create insider risk levels that fit your organization's needs. You can customize the criteria for the insider risk level and define conditions for when to assign the insider risk level to users.

Consider the following examples for using Adaptive Protection together with data loss prevention, data lifecycle management, and Conditional Access policies.

- **Data loss prevention policies**:
  - Allow users with the *Minor risk level* or *Medium risk level* to receive policy tips and education on best practices for handling sensitive data. With this approach, you can influence positive behavior changes over time and reduce organizational data risks.
  - Block users with the *Elevated risk level* from saving or sharing sensitive data to minimize the impact of potential data incidents.
- **Data lifecycle management policy**:
  - Preserve any SharePoint, OneDrive, or Exchange Online content deleted by a risky user for 120 days. A risky user in this case is a user that Adaptive Protection assigns the *Elevated risk level*.
- **Conditional Access policies**:
  - Require *Minor risk level* users to acknowledge Terms of Use before using an application.
  - Block *Medium risk level* users from accessing certain applications.

- Completely block *Elevated risk level* users from using *any* applications. Learn more about commonly applied Conditional Access policies

Insider risk level criteria and conditions

You can customize insider risk level criteria and conditions based on the following areas:

- **Alerts generated or confirmed for a user**: Choose conditions based on the severity level for alerts that the selected Insider Risk Management policy generates or confirms for a user. Conditions for alerts aren't additive. The insider risk level is assigned to a user if one of the conditions is met.
- **Specific user activity**: Choose conditions for activity to detect, its severity, and the number of daily occurrences during the past activity detection window (optional). Conditions for user activity are additive. The insider risk level is assigned to a user only if all the conditions are met.

Past activity detection

The **Past activity detection** setting determines how many days back Adaptive Protection examines to detect whether a user meets the conditions defined by any of the insider risk levels. The default setting is 7 days, but you can choose between 5 and 30 days of previous activity to apply insider risk level conditions. This setting only applies to insider risk levels that are based on a user's daily activity and excludes insider risk levels based on alerts.

The following example illustrates how past activity detection settings and insider risk levels interact to determine if a user's past activity is in-scope:

- **Elevated risk level** setting: User performs at least three sequences, each with a high severity risk score (67 to 100)
- **Past activity detection** setting: 3 days

Expand table

User activity	Activity in-scope for insider risk level
User has 1 high severity sequence each day on Day T-3, T-2, T-1	Yes
User has 1 high severity sequence on Day T-4 and 2 high severity sequences on Day T-3	No
User has 3 high severity sequences on Day T-3	Yes

Insider risk level timeframe

The **Insider risk level timeframe** setting determines how long an insider risk level remains assigned to a user before it's automatically reset. The default setting is 7 days, but you can choose between 5 and 30 days before resetting the insider risk level for a user.

Insider risk levels also reset for a user when:

- The associated alert for the user is dismissed.
- The associated case for the user is resolved.
- The insider risk level end date is manually expired.

> **Note**
>
> If a user is currently assigned an insider risk level and that user meets the criteria again for that insider risk level, the insider risk level timeframe extends for the defined number of days for the user.

Risk level expiration options

When you enable the automatic expiration option (enabled by default), Adaptive Protection risk levels for a user automatically expire when the associated alert for a user is dismissed or the associated case for the user is closed. If you want to keep the Adaptive Protection risk level for a user even if the alert is dismissed or the case is closed, disable the option.

Permissions for Adaptive Protection

> **Important**
>
> To learn about licensing requirements for using Adaptive Protection with data loss prevention, data lifecycle management, and Conditional Access, see Microsoft 365, Office 365, Enterprise Mobility + Security, and Windows 11 Subscriptions.

Depending on how you use Insider Risk Management built-in role groups and role groups for data loss prevention or Conditional Access, you might need to update Insider Risk Management permissions for administrators, analysts, and investigators in your organization.

The following table describes the permissions required for specific Adaptive Protection tasks.

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

Expand table

Task	Required role group
Configure Adaptive Protection and update settings	*Insider Risk Management* or *Insider Risk Management Admins*
Create and manage Conditional Access policies with the Adaptive Protection condition	One of the following: *Global administrator*, *Conditional Access Administrator*, *Security Administrator*
Create and manage data loss prevention policies with the Adaptive Protection condition	One of the following: *Compliance Administrator*, *Compliance Data Administrator*, *DLP Compliance Management*, *Global Administrator*
View details on users' or agents' assigned insider risk levels	*Insider Risk Management*, *Insider Risk Management Analysts*, or *Insider Risk Management Investigators*

> **Important**
>
> The four categories of role groups correspond to the following tabs on the Adaptive Protection page: **Insider risk levels**, **Users assigned insider risk levels**, **Data Loss Prevention**, **Conditional Access**. If you're not assigned to the appropriate role group, the tab doesn't appear on the **Adaptive protection** page.

Learn more about role groups in Microsoft Defender for Office 365 and Microsoft Purview compliance

Configure Adaptive Protection

Depending on the needs of your organization or where you're currently configured with Insider Risk Management, data loss prevention, data lifecycle management, and Conditional Access, you have two options to get started with Adaptive Protection:

- Quick setup

- Custom setup

Quick setup

The quick setup option is the fastest way to get started with Adaptive Protection. With this option, you don't need any preexisting Insider Risk Management, data loss prevention, data lifecycle management, or Conditional Access policies. You also don't need to preconfigure any settings or features. If your organization doesn't have a current subscription or license that supports Insider Risk Management, data loss prevention, or data lifecycle management, sign up for a Microsoft Purview risk and compliance solutions trial before starting the quick setup process. You can also sign up for a Microsoft Entra trial for Conditional Access.

To get started, select **Turn on Adaptive Protection** from the Adaptive Protection cards on the Microsoft Purview portal home page or the data loss prevention Overview page. You can also start the quick setup process by going to **Insider Risk Management** > **Adaptive protection** > **Dashboard** > **Quick setup**.

> **Note**
>
> If you're already a scoped admin for Microsoft Purview, you can't turn on quick setup.

Here's what the quick setup process configures for Adaptive Protection:

Expand table

Area	Configuration
A new insider risk policy	- Policy template: Data leaks - Policy name: Adaptive Protection policy for Insider Risk Management - Policy scope for users and groups: All users and groups - Priority content: None - Triggering events: Selected exfiltration events (you can view in Insider Risk Management settings) - Policy indicators: A subset of Office indicators (you can view in Insider Risk Management settings) - Risk score boosters: Activity is over user's usual activity for that day
Adaptive protection insider risk levels	- *Elevated risk level*: Users must have at least three high severity exfiltration sequences - *Moderate risk level*: Users must have at least two high severity activities (excluding some types of downloads) - *Minor risk level*: Users must have at least one high severity activity (excluding some types of downloads)

Area	Configuration
Insider risk settings (if already configured)	- Policy indicators: Office indicators not already configured (you can view in Insider Risk Management settings) - All other settings previously configured aren't updated or changed. - Analytics: On (thresholds for triggering events in policies are the default settings determined by Analytics recommendations)
Insider risk settings (if not already configured)	- Privacy: Show anonymized versions of user names. **Note:** User names aren't anonymized in Conditional Access or data loss prevention. - Policy timeframes: Defaults - Policy indicators: A subset of Office indicators (you can view in Insider Risk Management settings) - Risk score boosters: All - Intelligent Detections: Alert Volume = Default volume - Analytics: On - Admin notifications: Send notification email when first alert is generated to all
New Conditional Access policy (created in Report-only mode so users won't be blocked)	**1-Block access for users with Insider Risk (Preview)** - Included users: All users - Excluded guest or external users: B2bDirectConnect User; OtherExternalUser; ServiceProvider - Cloud apps: Office 365 apps - Insider risk levels: *Elevated* - Block access: Selected
New data lifecycle management policy	An organization-wide autoapply label policy that monitors users assigned the *Elevated risk level*. Any SharePoint, OneDrive, or Exchange Online content deleted by a risky user is preserved for 120 days. If you haven't already turned on Adaptive Protection, the policy is automatically created and applied when you turn on Adaptive Protection. If you have already turned on Adaptive Protection, you need to explicitly opt in to the data lifecycle management policy that's created automatically.
Two new data loss prevention policies	Adaptive Protection policy for Endpoint DLP - *Elevated risk level* rule: Blocked - *Moderate/Minor risk level* rule: Audit - Policy starts in test mode (audit only) Adaptive Protection policy for Teams and Exchange DLP - *Elevated risk level* rule: Blocked - *Moderate/Minor risk level* rules: Audit - Policy starts in test mode (audit only)

After you start the quick setup process, it can take up to 72 hours before the following items are completed:

- Analytics
- Associated Insider Risk Management, data loss prevention, data lifecycle management, and Conditional Access policies
- Adaptive Protection insider risk levels, data loss prevention, data lifecycle management, and Conditional Access actions applied to applicable user activities.

> **Note**
>
> Once you enable Adaptive Protection, don't disable it before the setup process completes. Disabling Adaptive Protection before the setup process completes might lead to policy errors.

Administrators receive a notification email when the quick setup process completes.

Custom setup

The custom setup option lets you customize the Insider Risk Management policy, the insider risk levels, and the data loss prevention and Conditional Access policies configured for Adaptive Protection.

> **Note**
>
> For data lifecycle management, if you don't turn on Adaptive Protection, the policy isn't created and applied automatically. If you turn on Adaptive Protection for your organization, you need to **explicitly opt in to apply the data lifecycle management policy**.

The custom setup option also lets you configure these items before enabling the Adaptive Protection connections between Insider Risk Management and data loss prevention. In most cases, use this option if your organization already has Insider Risk Management and data loss prevention policies in place.

Complete the following steps to configure Adaptive Protection by using the custom setup.

Step 1: Create Insider Risk Management policy

Assign insider risk levels to users or agents when a policy assigned in Adaptive Protection either detects user or agent activity or generates alerts that match the insider risk level conditions you define in the next step. If you don't want to use an existing Insider Risk Management policy,

create a new Insider Risk Management policy. Your Insider Risk Management policy for Adaptive Protection should include:

- **Users whose activity you want to detect**. This group can include all users and groups in your organization or just a subset for specific risk mitigation scenarios or testing purposes.
- **Activities you consider risky and custom thresholds that influence an activity's risk score**. Risky activities might include emailing people outside your organization or copying files to USB devices.

Select **Create insider risk policy** to start the new policy workflow. The Data leaks policy template is automatically selected in the workflow, but you can select any policy template.

> **Important**
>
> Depending on the policy template you select, you might need to configure additional settings for the policy to properly detect potentially risky activities and to create applicable alerts.

Step 2: Configure insider risk level settings

Select the **Insider risk levels** tab. Start by selecting the Insider Risk Management policy you want to use for Adaptive Protection. This policy can be the new policy you created in Step 1 or an existing policy or policies that you already configured.

Next, accept the applicable built-in insider risk level conditions or create your own. Depending on the type of policy you select, the insider risk level conditions reflect the applicable conditions associated with indicators and activities you configure in the policy.

For example, if you choose a policy based on the Data leaks policy template, the built-in insider risk level condition choices apply to indicators and activities available in that policy. If you select a policy based on the Security policy violations policy template, the built-in insider risk level conditions automatically scope to indicators and activities available in that policy.

Customize an insider risk level for your policy

1. Sign in to the Microsoft Purview portal by using credentials for an admin account in your Microsoft 365 organization.
2. Go to the **Insider Risk Management** solution.
3. Select **Adaptive protection** in the left navigation, then select **Insider risk levels**.

4. On the **Insider risk levels** page, select **Edit** for the insider risk level you want to customize (*Elevated*, *Moderate*, or *Minor*).

5. On the **Custom insider risk level** pane, select an option in the **Insider risk level based on** section:

   - **Alert generated or confirmed for a user**
   - **Specific user activity**

6. If you select the **Alert generated or confirmed for a user** option, choose the severity levels for alerts that are generated or confirmed for a user that should use this insider risk level. You can keep the *Severity for generated alerts* and the *Severity for confirmed alerts* conditions or remove one of these conditions if you only want to use one. If you need to add one of these conditions back, select **Add condition**, then select the condition. For each condition, choose the severity level that should be applied for the condition (*High*, *Medium*, or *Low*). If the **any** of the conditions are met, the insider risk level is assigned to a user.

7. If you select the **Specific user activity** option, choose the activity to detect, its severity, and number of daily occurrences during the past activity detection window. You must configure the *Activities*, *Activity severity*, and *Activity occurrences during detection window* conditions for this insider risk level.

   For the *Activities* condition, the options you can choose from automatically update for the types of activities you define with the indicators configured in the associated policy. If needed, select the **Assign this insider risk level to any user who has a future alert confirmed, even if conditions above aren't met** checkbox. If the **all** of the conditions are met, the insider risk level is assigned to a user.

   For the *Activity severity* condition, specify the severity level for the activities included in the daily activity insight. The options are *High*, *Medium*, and *Low*, and are based on risk score ranges.

   For the *Activity occurrences during detection window* condition, specify the number of times selected activities must be detected within the specified *Past activity detection* period. This number isn't related to the number of events that might occur for an activity. For example, if the policy detects that a user downloaded 20 files from SharePoint in one day, that count is one daily activity insight consisting of 20 events.

8. Select **Confirm** to apply the custom insider risk level conditions.

How insider risk level is assigned if a user is in scope for multiple policies

If a user is in scope for multiple policies and receives alerts of different severity levels, the user is assigned the highest severity level received. For example, consider a policy that assigns an *Elevated risk level* if users receive a high severity alert. If a user receives a low-severity alert from policy 1, a medium-severity alert from policy 2, and a high-severity alert from policy 3, the user is assigned an *Elevated risk level* - the level for the highest alert severity received.

Insider risk level conditions must be present in selected policies to be detected. For example, if you select the *Copy to USB* activity to assign a *Moderate risk level*, but the activity is selected in just one of three selected policies, then only activity from that one policy assigns a *Moderate risk level* for that activity.

Step 3: Create or edit a data loss prevention policy

Next, create a new data loss prevention policy or edit an existing policy to restrict actions for users who match your insider risk level conditions in Adaptive Protection. Use the following guidelines for your data loss prevention policy configuration:

- You must include the *User's insider risk level for Adaptive Protection is* condition in your data loss prevention policy. This data loss prevention policy can include other conditions as needed.
- Although you can include other locations in the data loss prevention policy, Adaptive Protection currently supports only Exchange, Microsoft Teams, and devices.

Select **Create policy** to start the data loss prevention policy workflow and create a new data loss prevention policy. If you have an existing data loss prevention policy that you'd like to configure for Adaptive Protection, go to **Data loss prevention** > **Policies** in the Microsoft Purview portal and select the data loss prevention policy you'd like to update for Adaptive Protection. For guidance on how to configure a new data loss prevention policy or update an existing data loss prevention policy for Adaptive Protection, see Learn about Adaptive Protection in data loss prevention: Manual configuration.

> **Tip**
>
> We recommend testing the data loss prevention policy (with policy tips) so you can review data loss prevention alerts to verify that the policy is working as expected before enabling Adaptive Protection.

Step 4: Create or edit a Conditional Access policy

Next, create a new Conditional Access policy or edit an existing policy to restrict actions for users who match your insider risk level conditions in Adaptive Protection. Use the following guidelines for your Conditional Access policy configuration:

- On the Conditional Access page where you control access based on signals from conditions, set the *Insider risk* condition to *Yes*, and then select an insider risk level (*Elevated*, *Moderate*, or *Minor*). This is the insider risk level that users must have in order for the policy to be enforced.

Select **Create policy** to start the Conditional Access policy workflow and create a new Conditional Access policy. If you have an existing Conditional Access policy that you'd like to configure for Adaptive Protection, go to **Protection** > **Conditional Access** in the Microsoft Entra admin center and select the Conditional Access policy you'd like to update for Adaptive Protection. For guidance on how to configure a new Conditional Access policy or update an existing Conditional Access policy for Adaptive Protection, see Common Conditional Access policy: Insider risk based policy.

Step 5: Turn on Adaptive Protection

After you complete the Insider Risk Management policy, insider risk level, DLP policy, and Conditional Access policy configuration steps, you're ready to enable Adaptive Protection. When you turn on Adaptive Protection:

- The Insider Risk Management policy starts looking for user activity matching your insider risk level conditions. If detected, the insider risk levels are assigned to users.
- Users who are assigned insider risk levels appear on the **Users assigned insider risk levels** tab in Adaptive Protection.
- The data loss prevention policy applies protection actions for any user assigned to insider risk levels included in the data loss prevention policy. The data loss prevention policy is added to the **Data Loss Prevention** tab in **Adaptive protection**. You can view details about the data loss prevention policy and edit policy conditions from the dashboard.
- The data lifecycle management policy, which is automatically created, applies protection actions for any user assigned to the *Elevated risk level*. The following message appears with a green background on the **Adaptive Protection** tab to let users know that proactive data preservation is being applied: "Your organization is also being dynamically protected from users who might potentially delete critical data." The message also provides a link to the

data lifecycle management setting where you can turn off the proactive data preservation feature, if desired.

- The Conditional Access policy applies protection actions for any user assigned to insider risk levels included in the Conditional Access policy. The Conditional Access policy is added to the **Conditional Access policies** tab in **Adaptive protection**. You can view details about the Conditional Access policy and edit policy conditions from the dashboard.

To enable Adaptive Protection, select the **Adaptive Protection settings** tab, and then turn **Adaptive Protection** to *On*. It might take up to 36 hours before you can expect to see Adaptive Protection insider risk levels and data loss prevention, data lifecycle management, and Conditional Access actions applied to applicable user activities.

Watch the following video on the Microsoft Mechanics channel to see how Adaptive Protection can automatically adjust the strength of data protection based on calculated data security insider risk levels of users.

Manage Adaptive Protection

When you enable Adaptive Protection and configure your Insider Risk Management, data loss prevention, and Conditional Access policies, you can access information about policy metrics, users currently included, and insider risk levels currently in scope.

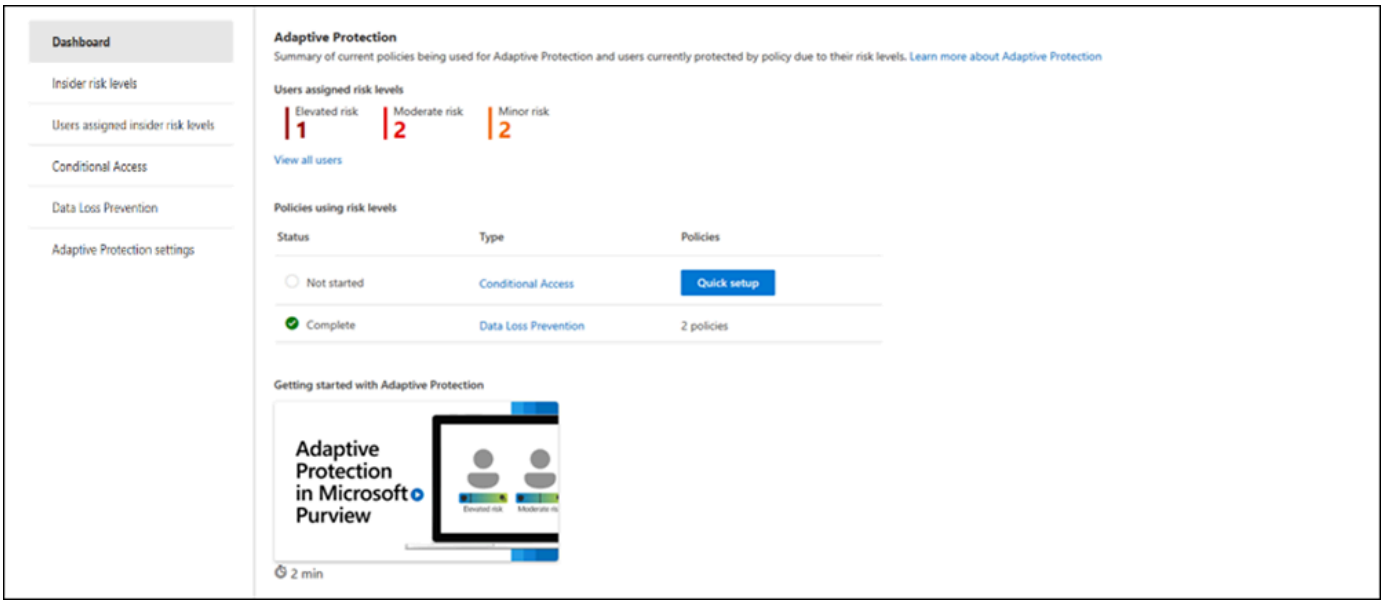
> **Note**
>
> Data lifecycle management metrics don't currently appear on the dashboard. However, if proactive data preservation is turned on, you see the following message on the **Adaptive Protection** tab: "Your organization is also being dynamically protected from users who might potentially delete critical data."

Use the Adaptive Protection dashboard

After you complete either the Quick setup or Custom setup process, the **Dashboard** tab in **Adaptive protection** displays widgets for summary information about user insider risk levels, Conditional Access policies, and data loss prevention policies.

- **Users assigned insider risk levels**: Shows the number of users for each insider risk level (*Elevated*, *Moderate*, and *Minor*).
- **Policies using insider risk levels**: Shows the status of policies (*Not started* or *Complete*), the type of policy (*Conditional Access* or *Data Loss Prevention*), and the number of configured

policies for each type of policy. If a policy type isn't configured, select **Quick setup** to configure the policy.



View users assigned insider risk levels

The **Users assigned insider risk levels** tab shows users assigned an insider risk level in Adaptive Protection. You can review the following information for each user:

- **Users**: Lists the user name. For data loss prevention policies, if the Show anonymized versions of usernames option is selected in Insider Risk Management settings, you see anonymized user names. For Conditional Access policies, user names aren't anonymized even if the **Show anonymized versions of usernames** setting is selected.

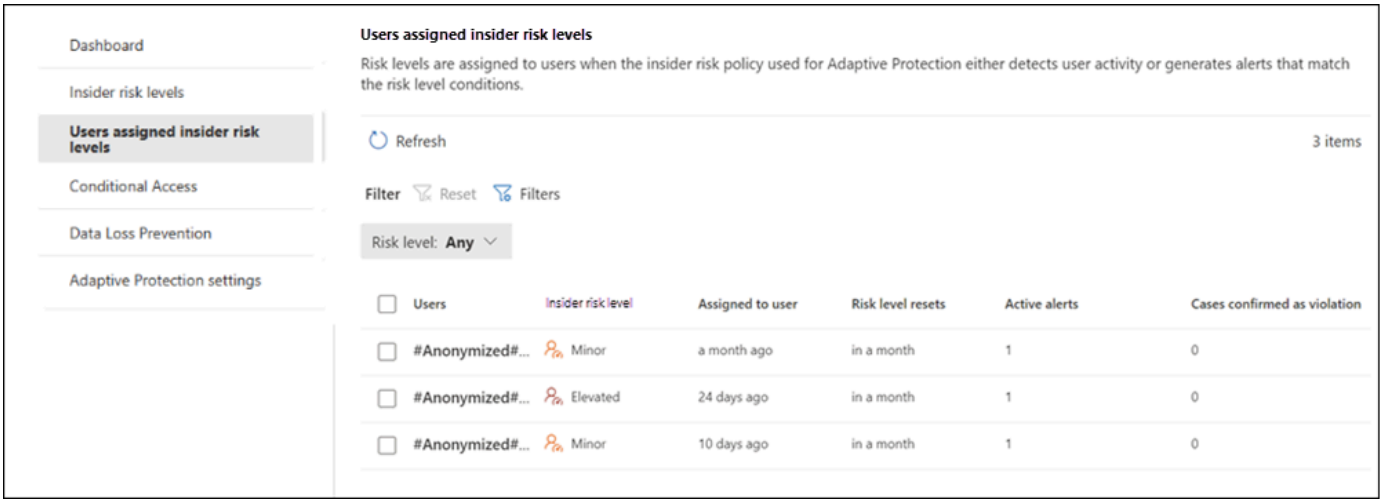
  > **Important**
  >
  > To maintain referential integrity, anonymization of usernames (if turned on) isn't preserved for users from Adaptive Protection who have alerts or activity that appear outside Insider Risk Management. Actual usernames appear in related data loss prevention alerts and the activity explorer.

- **Insider risk level**: The current insider risk level assigned to the user.
- **Assigned to user**: The number of days or months that pass since assigning an insider risk level to the user.
- **Insider risk level resets**: The number of days until the insider risk level automatically resets for the user.

To manually reset the insider risk level for a user, select the user, then select **Expire**. This user no longer has an assigned insider risk level. Existing alerts or cases for this user aren't removed. If this user is included in the selected Insider Risk Management policy, an insider risk level is assigned again if a triggering event is detected.

- **Active alerts**: The number of current Insider Risk Management alerts for the user.
- **Cases confirmed as violation**: The number of confirmed cases for the user.
- **Case**: The name of the case.

You can filter users by *Insider risk level*.

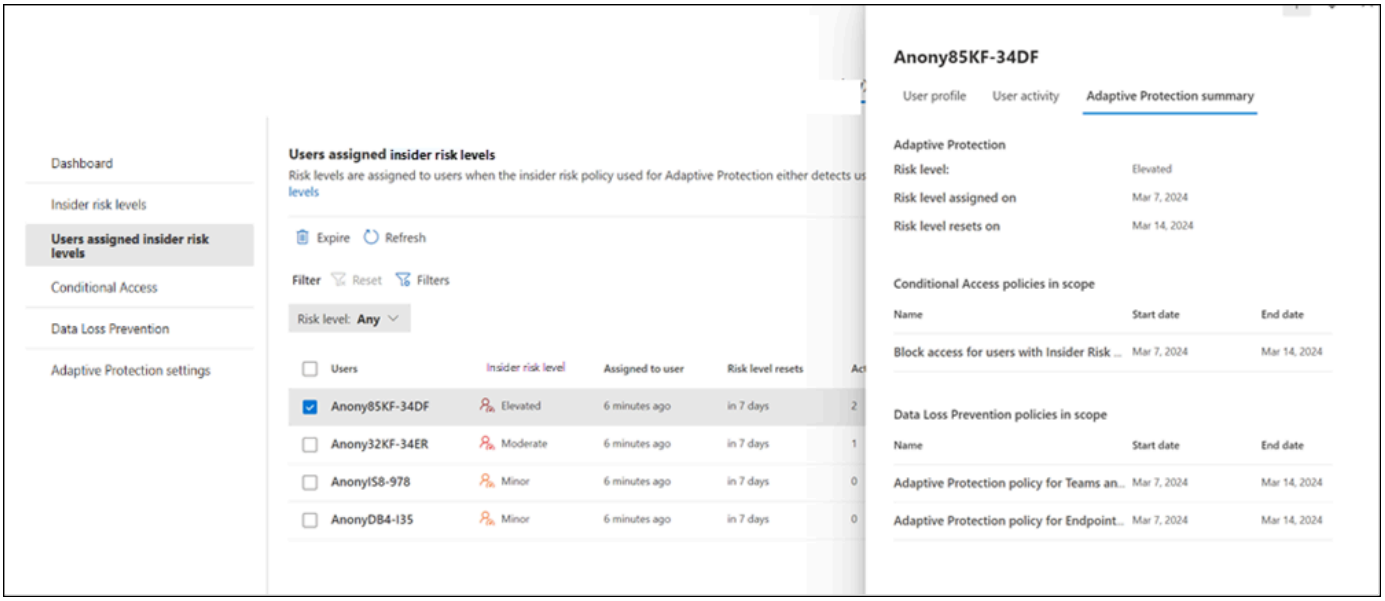


To view detailed insider risk and Adaptive Protection information for a specific user, select the user to open the user details pane. The details pane contains three tabs: **User profile**, **User activity**, and **Adaptive protection summary**. For information about the **User profile** and **User activity** tabs, see View user details.

The **Adaptive protection summary** tab aggregates information in four sections:

- **Adaptive protection**: Displays information about the current *Risk level*, *Risk level assigned on*, and *Risk level reset on* for the user.
- **Data Loss Prevention policies in scope (dynamic)**: Displays all the data loss prevention policies currently in scope for the user and the start and end date for the policy. This information is based on the insider risk level for the user and data loss prevention policy configuration for insider risk levels. For example, if a user has activities that are defined as *Elevated risk level* for Insider Risk Management policies, and two data loss prevention policies are configured with the *Elevated risk level* condition, these two data loss prevention policies are displayed here for the user.

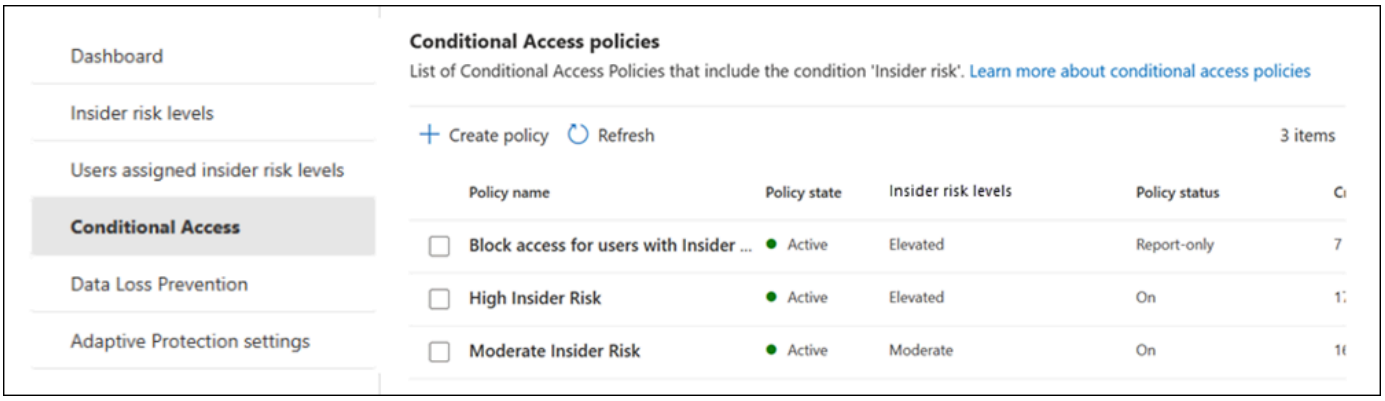
- **Conditional Access policies in scope (dynamic)**: Displays all the Conditional Access policies currently in scope for the user and the start and end date for the policy. This information is based on the insider risk level for the user and Conditional Access policy configuration for insider risk levels. For example, if a user has activities that are defined as *Elevated risk level* for Insider Risk Management policies, and a Conditional Access policy is configured with the *Elevated risk level* condition, the Conditional Access policy is displayed here for the user.
- **Insider risk policy for Adaptive protection**: Displays any Insider Risk Management policy where the user is currently in-scope.



Manage Conditional Access policies for Adaptive Protection

The **Conditional Access policies** page shows all Conditional Access policies that use the **Insider risk** condition. You can review the following information for each policy:

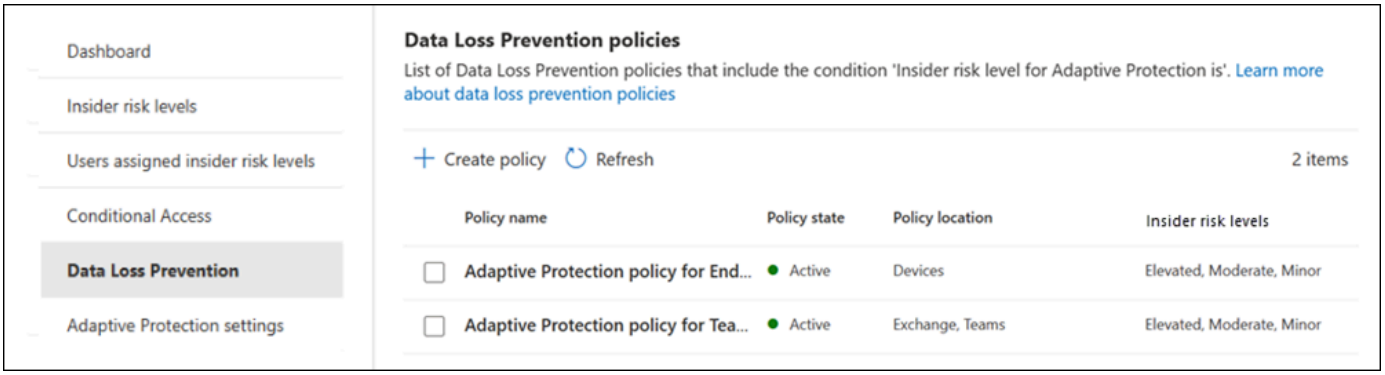
- **Policy name**: The name of the Conditional Access policy.
- **Policy state**: The current state of the policy. Values are *Active* or *Inactive*.
- **Insider risk levels**: The insider risk levels included in the Conditional Access policy by using the **Insider Risk** condition. The options are *Elevated*, *Moderate*, or *Minor*.
- **Policy status**: The current status of the Conditional Access policy. The options are *On* or *Test with notifications*.
- **Created on**: The date you created the Conditional Access policy.
- **Last modified**: The date you last edited the Conditional policy.



Manage Data Loss Prevention policies for Adaptive Protection

The **Data Loss Prevention policies** page shows all data loss prevention policies that use the **User's insider risk level for Adaptive Protection is** condition. You can review the following information for each policy:

- **Policy name**: The name of the data loss prevention policy.
- **Policy state**: The current state of the policy. Values are *Active* or *Inactive*.
- **Policy location**: The DLP policy locations included in the data loss prevention policy. Currently, Adaptive Protection supports Exchange, Teams, and Devices.
- **Insider risk levels**: The insider risk levels included in the data loss prevention policy by using the **Insider risk level for Adaptive Protection is** condition. The options are *Elevated*, *Moderate*, or *Minor*.
- **Policy status**: The current status of the data loss prevention policy. The options are *On* or *Test with notifications*.
- **Created**: The date you created the data loss prevention policy.
- **Last modified**: The date you last edited the data loss prevention policy.



Tune your insider risk level settings

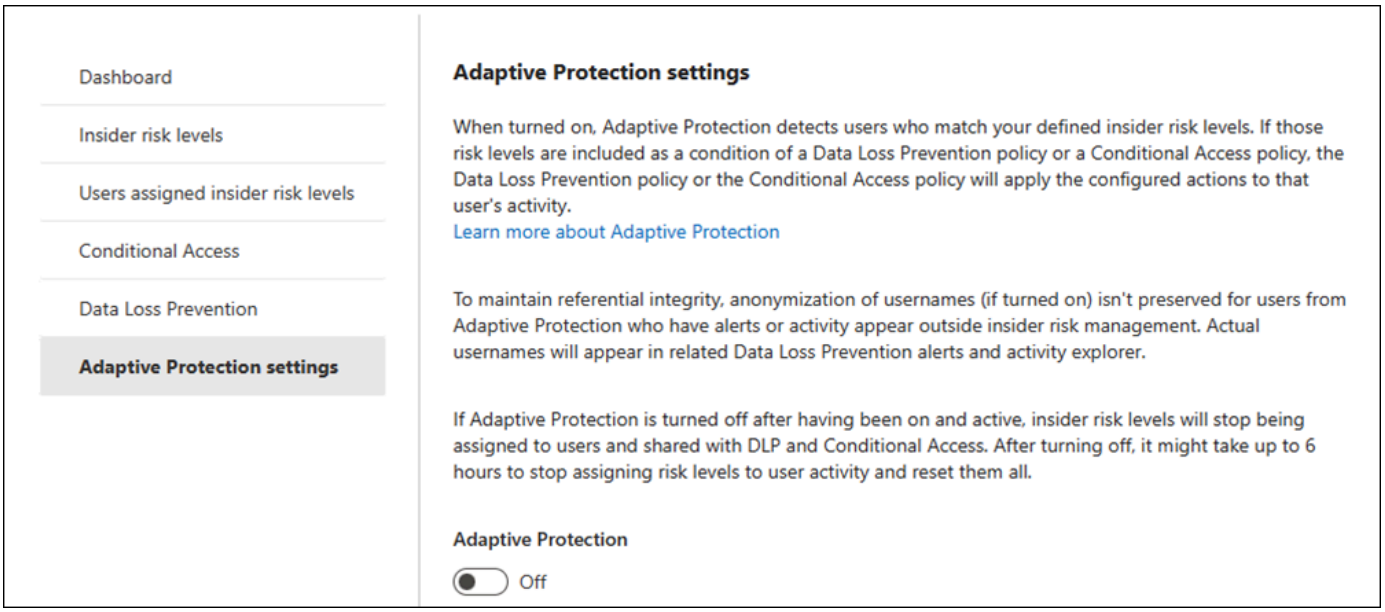
After reviewing users with insider risk levels, you might find too many or too few users assigned an insider risk level. Use these two methods to tune your policy configurations and adjust the number of users assigned insider risk levels:

- Modify insider risk level settings. Adjust your thresholds to assign an insider risk level to a user:
  - Increase or decrease the severity of the activity required to assign an insider risk level. For example, reduce the activity or alert severity if you're seeing too few users with insider risk levels.
  - If the insider risk level is based on a specific user activity, increase or decrease the activity occurrences during the detection window. For example, reduce the activity occurrences if you're seeing too few users with insider risk levels.
  - Change what the insider risk level is based on. For example, assign an insider risk level only if the alert is confirmed to decrease the number of users if you're seeing too many users with insider risk levels.
- Modify policy thresholds. Since policy detections assign insider risk levels, modify your policy to change the requirements to assign an insider risk level. Modify a policy by increasing or decreasing the policy thresholds that lead to high, medium, or low severity activities and alerts.

Disable Adaptive Protection

In certain scenarios, you might need to temporarily disable Adaptive Protection. To disable Adaptive Protection, select the **Adaptive Protection settings** tab and turn **Adaptive Protection** *Off*.

If you turn off Adaptive Protection after it's been on and active, the system stops assigning insider risk levels to users and stops sharing them with data loss prevention, data lifecycle management, and Conditional Access. All existing insider risk levels for users are reset. After turning off Adaptive Protection, it might take up to six hours to stop assigning insider risk levels to user activity and reset them all. The Insider Risk Management, data loss prevention, data lifecycle management, and Conditional Access policies aren't automatically deleted.



> **Note**
>
> You can opt out of data lifecycle management protection without disabling Insider Risk Management Adaptive Protection by turning off the **Adaptive protection in Data Lifecycle Management** setting in data lifecycle management. If you turn off this setting, the data lifecycle management policy is deleted. The setting isn't enabled again unless you turn it back on. Learn more about the Adaptive protection in Data Lifecycle Management setting

Note: The author created this article with assistance from AI. Learn more

Last updated on 06/26/2026

Learn about Adaptive Protection in Data Loss Prevention

Article • 03/31/2025

Adaptive Protection in Microsoft Purview integrates Microsoft Purview Insider Risk Management with Microsoft Purview Data Loss Prevention (DLP). When insider risk identifies a user who is engaging in risky behavior, they are dynamically assigned to a inside risk level. Then Adaptive Protection can automatically create a DLP policy to help protect the organization against the risky behavior that's associated with that inside risk level. As users insider risk levels change in insider risk management, the DLP policies applied to users can adjust.

You can manually create DLP policies that help protect against risky behaviors that insider risk identifies too.

Refer to Help dynamically mitigate risks with Adaptive Protection to learn about Adaptive Protection and how to configure it.

How Adaptive Protection shows up in DLP policies

If you're unfamiliar with DLP policies, you should review these articles before working with Adaptive Protection:

- Learn about data loss prevention
- Plan for data loss prevention (DLP)
- Data Loss Prevention policy reference
- Design a data loss prevention policy

Once Adaptive Protection is configured in insider risk, a condition called **User's risk level for Adaptive Protection is** will be available to use in rules that are configured for policies scoped to Exchange Online, Devices, and Teams locations.

The condition **Insider risk level for Adaptive Protection is** has three values:

- **Elevated risk level**
- **Moderate risk level**
- **Minor risk level**

These insider risk level profiles are defined in insider risk. You can select one, two or all three in a policy rule. Learn more about insider risk levels.

You can manually configure DLP policies that are part of Adaptive Protection and also use the quick setup configuration in insider risk to create DLP policies automatically from a template.

Manual configuration

You manually configure an Adaptive Protection DLP policy just like you would configure any other policy. Just select the **Insider risk level for Adaptive Protection is** condition and the insider risk level profiles that you want, configure all the other policy options and deploy the policy according to your normal procedures.

Quick setup configuration

If quick setup is used to configure Adaptive Protection in insider risk, DLP policies are created automatically, so you should be on the lookout for them. Quick setup will create one policy for Teams and Exchange Online with two rules, one for the elevated risk profile and one for the moderate and minor insider risk levels. It will also create one policy for Devices with two rules, one for the elevated risk profile and one for the moderate and minor insider risk levels.

> **Tip**
>
> Insider risk presents a view of just the DLP policies that use the **Insider risk level for Adaptive Protection is** condition. Open **Microsoft Purview portal** > **Insider risk management** > **Adaptive protection** to see the list. You'll need DLP to be in one of these roles to access the insider risk node:

- Compliance administrator
- Compliance Data administrator
- Organization management (Users who are not global administrators must be Exchange administrators to see and take action on devices that are managed by Basic Mobility and Security for Microsoft 365)
- Global administrator
- DLP compliance management
- View-only DLP compliance management

> **Important**

Microsoft recommends that you use roles with the fewest permissions. This helps improve security for your organization. Global Administrator is a highly privileged role that should only be used in scenarios where a lesser privileged role can't be used.

Policy values for Teams and Exchange online DLP policy

This is the configuration for the Teams and Exchange DLP policy created during Quick Setup. The policy name is **Adaptive Protection policy for Teams and Exchange DLP**.

Rule: Adaptive Protection block rule for Teams and Exchange DLP

Expand table

DLP policy element	Configured value
Conditions	**Insider risk level for Adaptive Protection is** **- Elevated Risk Level** AND **- Content is Shared from Microsoft 365 With people outside my organization**
Actions	**Restrict access or encrypt the content in Microsoft 365 locations** **- Block only people outside your organization**
User Notification	**On** **- Notify user with a policy tip** **– Notify the user who sent, shared, or last modified the content**
User Override	**Off**
Incident reports	**On** **- Severity Level – Low** **- Send alert every time an activity matches the rule**
Additional Options	**Off**
Status	**Run the policy in simulation mode** **- Policy Tips** not selected

Rule: Adaptive Protection audit rule for Teams and Exchange DLP

Expand table

DLP policy element	Configured value
Conditions	**Insider risk level for Adaptive Protection is** **- Moderate Risk Level, Minor Risk Level** AND **- Content is Shared from Microsoft 365 With people outside my organization**
Actions	None
User Notification	**On** **- Notify user with a policy tip** **- Notify the user who sent, shared, or last modified the content**
User Override	**Off**
Incident reports	**On** **- Severity Level – Low** **- Send alert every time an activity matches the rule**
Additional Options	**Off**
Status	**Run the policy in simulation mode** - **Policy tips** not selected

Policy values for Devices DLP policy

This is the configuration for the Devices DLP policy created during Quick Setup. The policy name is **Adaptive Protection policy for Endpoint DLP**.

> **ⓘ Important**
>
> For Adaptive Protection to work on Devices, you must either enable Advanced classification scanning and protection or if you are manually creating the Adaptive Protection policy, select the **File Type is** condition.

> **ⓘ Important**
>
> If a user is targeted by a default Adaptive Protection Device DLP policy and is targeted by an independent Device DLP policy, only the actions of the *most restrictive* policy will be applied.

Rule: Adaptive Protection block rule for Endpoint DLP

Expand table

DLP policy element	Configured value
Conditions	**Insider risk level for Adaptive Protection is** **- Elevated Risk Level** AND **- File Type is** **- Word processing** **- Spreadsheet** **- Presentation** **- Archive** **- Mail**
Actions	**Audit or Restrict activities on Devices** **- Upload to a restricted cloud service domain or access from unallowed browsers - Block** **File activities for all apps** **- Apply restrictions to specific activity** **- Copy to clipboard – Block** **- Copy to removable USB device – Block** **- Copy to network share – Block** **- Print – Block** **Restricted App activities - Access by restricted apps - Block**
User Notification	**Off**
User Override	**Off**
Incident reports	**On** **- Severity Level – Low** **- Send alert every time an activity matches the rule**
Additional Options	**Off**
Status	**Run the policy in simulation mode** **- Policy Tips** option not selected

Rule: Adaptive Protection rule for Endpoint DLP

Expand table

DLP policy element	Configured value
Conditions	**Insider risk level for Adaptive Protection is** **- Moderate Risk Level, Minor Risk Level**

DLP policy element	Configured value
	AND - **File Type is** - **Word processing** - **Spreadsheet** - **Presentation** - **Archive** - **Mail**
Actions	**Audit or Restrict activities on Devices** **- Upload to a restricted cloud service domain or access from unallowed browsers – Audit** **File activities for all apps** **- Apply restrictions to specific activity** **- Copy to clipboard – Audit** **- Copy to removable USB device – Audit** **- Copy to network share – Audit** **-Print – Audit** **Restricted App activities** **- Access by restricted apps - Audit**
User Notification	**Off**
User Override	**Off**
Incident reports	**On** **- Severity Level – Low** **- Send alert every time an activity matches the rule**
Additional Options	**Off**
Status	**Run the policy in simulation mode** - **Policy tips** option not selected

See Also

- Adaptive Protection in Microsoft Purview

Feedback

Was this page helpful? Yes No

Adaptive Protection configuration guide

> **Important**
>
> Microsoft Purview Insider Risk Management correlates various signals to identify potential malicious or inadvertent insider risks, such as IP theft, data leakage, and security violations. Insider Risk Management enables customers to create policies to manage security and compliance. Built with privacy by design, users are pseudonymized by default, and role-based access controls and audit logs are in place to help ensure user-level privacy.

Adaptive Protection is highly configurable, from defining risk levels to setting up mitigation actions in data loss prevention (DLP) or conditional access policies. How you configure your risk levels and your policies that use the Adaptive Protection condition depends on your organization's security posture and your willingness to disrupt the end user experience for users and agents in your organization. You can define separate risk levels for users and agents by selecting the appropriate tab.

This configuration is our recommendation for organizations wanting to minimize false positives and potential disruptions to users and agents.

Expand table

Adaptive Protection settings	Elevated	Moderate	Minor
Risk level configuration	Confirmed alert of any severity for a user	High severity alert generated for a user	Low or medium severity alert generated for a user

Data loss prevention

Teams and Exchange configuration

With this configuration, your organization creates two DLP policies for Teams and Exchange locations. One policy is for elevated risk users and agents, and the other is for moderate and minor risk users and agents.

Elevated risk users and agents can't share content on Teams and Exchange with people outside of your organization. Moderate and minor risk users and agents have their attempts to share content externally on Teams and Exchange audited. Users and agents of all risk levels receive

an email notification that their attempt to share content externally conflicts with your organization's policy.

You can modify these configurations to fit your organization's security posture and policies. Policy tips aren't currently supported for Adaptive Protection in DLP.

Expand table

DLP policy element	Elevated	Moderate
Actions	Restrict access or encrypt the content in Microsoft 365 locations - Block only people outside your organization	None
Additional options	Off	Off
Conditions	Insider risk level for Adaptive Protection is - Elevated Risk AND - Content is Shared from Microsoft 365 with people outside my organization	Insider risk level for Adaptive Protection is - Moderate Risk AND - Content is Shared from Microsoft 365 with people outside my organization
Incident reports	On - Severity Level-Low - Send an alert every time an activity matches the rule	On - Severity Level-Low - Send an alert every time an activity matches the rule
Status	Run the policy in simulation mode - Policy tips not selected	Run the policy in simulation mode - Policy tips not selected
User notification	On - Notify the user who sent, shared, or last modified the content	On - Notify the user who sent, shared, or last modified the content
User override	Off	Off

Endpoint configuration

With this configuration, your organization creates two policies for devices. One policy is for elevated risk users and the other is for moderate and minor risk users.

Elevated risk users are blocked from uploading to restricted cloud service domains or accessing content from unallowed browsers. They're also blocked from copying content to the clipboard, removable USB devices, and network share. They're blocked from printing content that matches the policy conditions. They can't access content with restricted apps. For moderate and minor risk users, all of these activities are audited.

Expand table

DLP policy element	Elevated	Moderate
Actions	Audit or restrict activities on devices - Upload to a restricted cloud service domain or access from unallowed browsers (Block) File activities for all apps - Apply restrictions to specific activity - Copy to clipboard (Block) - Copy to removable USB device (Block) - Copy to network share (Block) - Print (Block) Restricted App activities - Access by restricted apps (Block)	Audit or restrict activities on devices - Upload to a restricted cloud service domain or access from unallowed browsers (Audit) File activities for all apps - Apply restrictions to specific activity - Copy to clipboard (Audit) - Copy to removable USB device (Audit) - Copy to network share (Audit) - Print (Audit) Restricted App activities - Access by restricted apps (Audit)
Additional options	Off	Off
Conditions	Insider risk level for Adaptive Protection is - Elevated Risk AND - File Type is *Word processing*, *Spreadsheet*, *Presentation*, *Archive*, and *Mail*	Insider risk level for Adaptive Protection is - Moderate Risk, Minor risk AND - Content is Shared from Microsoft 365 with people outside my organization
Incident reports	On - Severity Level-Low - Send an alert every time an activity matches the rule	On - Severity Level-Low - Send an alert every time an activity matches the rule

DLP policy element	Elevated	Moderate
Status	Run the policy in simulation mode - Policy tips not selected	Run the policy in simulation mode - Policy tips not selected
User notification	Off	Off
User override	Off	Off

Insider Risk Management

Conditional access policy configuration

To minimize disruption to user productivity while protecting against risky users, we recommend creating the following Conditional Access policies with the Insider risk condition when configuring Adaptive Protection.

Elevated

For elevated risk users and agents, minimize potential impact by taking a multipronged approach to limit how they can access sensitive data. Depending on your organization's risk posture, choose any combination of these policies.

- Block from Office 365 Apps: When used with the Insider risk condition, this configuration blocks elevated risk users from signing in to Office 365 apps.
- Block from access to specific labeled sites: When used with the Conditional Access Insider risk condition, this configuration blocks elevated users and agents from accessing those SharePoint sites and prevents Copilot from generating content grounded on these sites. By using the authentication context in Conditional Access, this configuration helps limit how risky users and agents can exploit access to sensitive information by using generative AI.
- Block downloads using Conditional Access application control: Prevent data exfiltration by using the Insider risk condition with this configuration to block elevated risk users and agents from downloading from specific cloud applications. This configuration might be useful for organizations that don't want to block access from cloud apps entirely but still want to apply additional controls to elevated users and agents.

Moderate

- Require Terms of Use at sign-in to Microsoft Admin Portals: Use the Insider risk condition in Conditional Access with a Terms of Use requirement to remind moderate risk users and agents of their security and privacy commitments while preserving productivity.

Minor

Get Conditional Access insights for minor risk users and agents while protecting their productivity. Using a policy with the Insider risk condition in Report-Only mode provides increased visibility on minor risk users and agents while preserving their productivity.

- Conditional Access insights and reporting workbook

Last updated on 11/21/2025

Adaptive scopes

> *Microsoft Purview service description*

When you create a communication compliance policy, a policy for retention, or an auto-labeling policy, you can add an adaptive scope for your policy. A single policy can have one or many adaptive scopes.

- An adaptive scope uses a query that you specify, so you can define the membership of users or groups included in that query. These dynamic queries run daily against the attributes or properties that you specify for the selected scope. You can use one or more adaptive scopes with a single policy.
- For example, you can assign different policy settings to users according to their department by using existing Microsoft Entra attributes without the administrative overhead of creating and maintaining groups for this purpose.

Advantages of using adaptive scopes

The advantages of using adaptive scopes include:

- No limits on the number of items per policy. Although adaptive policies are still subject to the maximum number of policies per tenant limitations, the more flexible configuration likely results in far fewer policies.
- More powerful targeting for your policies. For example, you can assign different settings to users according to their geographical location without the administrative overhead of creating and maintaining groups.
- Query-based scopes provide resilience against business changes that might not be reliably reflected in group membership or external processes that rely on cross-department communication.
- A single policy can include locations for both Microsoft Teams and Viva Engage, whereas when you don't use an adaptive scope, each location requires its own policy.
- Support for Microsoft Entra administrative units.

For additional advantages of using adaptive scopes specific to policies for retention, see Adaptive or static policy scopes for retention.

For configuration information, see Configuring adaptive scopes.

How adaptive scopes work with Microsoft Entra administrative units

Whereas adaptive scopes are created and configured in Microsoft Purview to support the dynamic targeting of policies for compliance, administrative units are created and configured in Microsoft Entra ID. They provide the ability to assign admins to one or more administrative units, with the result that these now restricted admins can manage only the users in their assigned administrative units. This configuration supports the security best practice of least privilege. Typically, administrative units are designed around specific geography, department, or business divisions.

This boundary of management flows into Microsoft Purview for supported solutions to ensure that restricted admins can manage only the users they have been assigned to manage.

As an example to show how administrative units integrate with adaptive scopes, consider a restricted compliance admin who wants to create an adaptive user scope just for users in France:

1. A compliance admin is assigned two administrative units, **All users in Europe** and **All users in North America**. When they create an adaptive scope, they can select and assign only these administrative units. They can't create an adaptive scope to manage users from other administrative units.
2. They create a new adaptive scope for users and select the administrative unit **All users in Europe**. Then, because they want the adaptive scope to be only for users in France, they use the Microsoft Entra ID **Country or region** attribute to specify France (CountryOrRegion = France). If they misconfigure this attribute and specify a valid value in Microsoft Entra ID, such as India, but the users with that value aren't included in the **All users in Europe** administrative unit, the scope won't contain any users.
3. When just this adaptive scope is selected for a policy that's targeted to all users, the policy is applied just to users in France.
4. As a reusable configuration element, the same adaptive scope can be used for other compliance policies.

If the compliance admin adds both of the administrative units to this adaptive scope, the end result is still the same because users in the North America administrative scope don't have France specified as their country or region attribute. However, the compliance admin knows that they need to target just users in France, so it's more efficient to run the query against just the Europe administrative unit. If requirements change, you can always add or remove administrative units from an existing adaptive scope.

Maximums for adaptive policy scopes

There's no limit to the number of adaptive policy scopes that you can add to a policy, but there are some maximum limits for the query that defines each adaptive scope:

- String length for attribute or property values: 200
- Number of attributes or properties without a group, or within a group: 10
- Number of groups: 10
- Number of characters in an advanced query: 10,000
- Grouping attributes or properties within a group isn't supported. This limitation means that the maximum number of properties or attributes supported within a single adaptive scope is 100.
- For auto-labeling policies, a SharePoint adaptive scope supports up to 50,000 sites per policy.

Configure adaptive scopes

When you choose to use adaptive scopes, you're prompted to select the type of adaptive scope you want. There are three different types of adaptive scopes, and each one supports different attributes or properties:

Expand table

Adaptive scope type	Attributes or properties supported
Users - applies to: - Exchange mailboxes - OneDrive accounts - Teams chats and Copilot interactions - Teams private channel messages (applicable pre-migration only) - Viva Engage user messages - Team chats - Teams call logs - Microsoft Copilot experiences - Enterprise AI apps - Other AI apps	First Name Last name Display name Job title Department Office Street address City State or province Postal code Country or region Email addresses Alias Exchange custom attributes: CustomAttribute1 - CustomAttribute15 Exchange Extension custom attributes: ExtensionCustomAttribute1 - ExtensionCustomAttribute5

Adaptive scope type	Attributes or properties supported
SharePoint sites * - applies to: - SharePoint sites and OneDrive accounts for retention policies - SharePoint sites for auto-labeling policies	Site URL Site name Site template Custom properties (SharePoint only): RefinableString00 - RefinableString99
Microsoft 365 Groups - applies to: - Microsoft 365 Group mailboxes & sites - Teams channel messages (standard, shared, and private post-migration) - Viva Engage community messages	Name Display name Description Email addresses Alias Exchange custom attributes: CustomAttribute1 - CustomAttribute15 Exchange Extension custom attributes: ExtensionCustomAttribute1 - ExtensionCustomAttribute5

* Records Center sites (legacy feature) in SharePoint aren't supported.

> **Note**
>
> For communication compliance policies:
>
> - SharePoint sites and OneDrive accounts aren't supported.
> - Excluded users and Microsoft 365 groups are supported.

The property names for sites are based on SharePoint site managed properties. For information about the custom attributes, see Using Custom SharePoint Site Properties to Apply Microsoft 365 Retention with Adaptive Policy Scopes.

The attribute names for users and groups are based on filterable recipient properties that map to Microsoft Entra attributes. For example:

- **Alias** maps to the LDAP name **mailNickname** that displays as **Email** in the Microsoft Entra admin center.
- **Email addresses** maps to the LDAP name **proxyAddresses** that displays as **Proxy address** in the Microsoft Entra admin center.

You can easily specify the attributes and properties listed in the table when you configure an adaptive scope by using the simple query builder. The advanced query builder supports additional attributes and properties, as described in the following section.

How to configure an adaptive scope

Before you configure your adaptive scope, use the previous section to identify what type of scope to create and what attributes and values to use. You might need to work with other administrators to confirm this information.

You need to assign the correct role groups to admins to create an adaptive scope. Any role group with the *Scope Manager* role can create an adaptive scope. The *Scope Manger* role is included in the following built-in role groups:

- Compliance Administrator
- Compliance Data Administrator
- Organization Management
- Records Management
- Communication Compliance
- Communication Compliance Admins

For SharePoint sites, there might be additional SharePoint configuration needed if you plan to use custom site properties.

To create and configure adaptive scopes, use the following instructions.

1. Sign in to the Microsoft Purview portal > **Settings** > **Roles and scopes** > **Adaptive scopes**.

2. On the **Adaptive scopes** page, select **+ Create scope**.

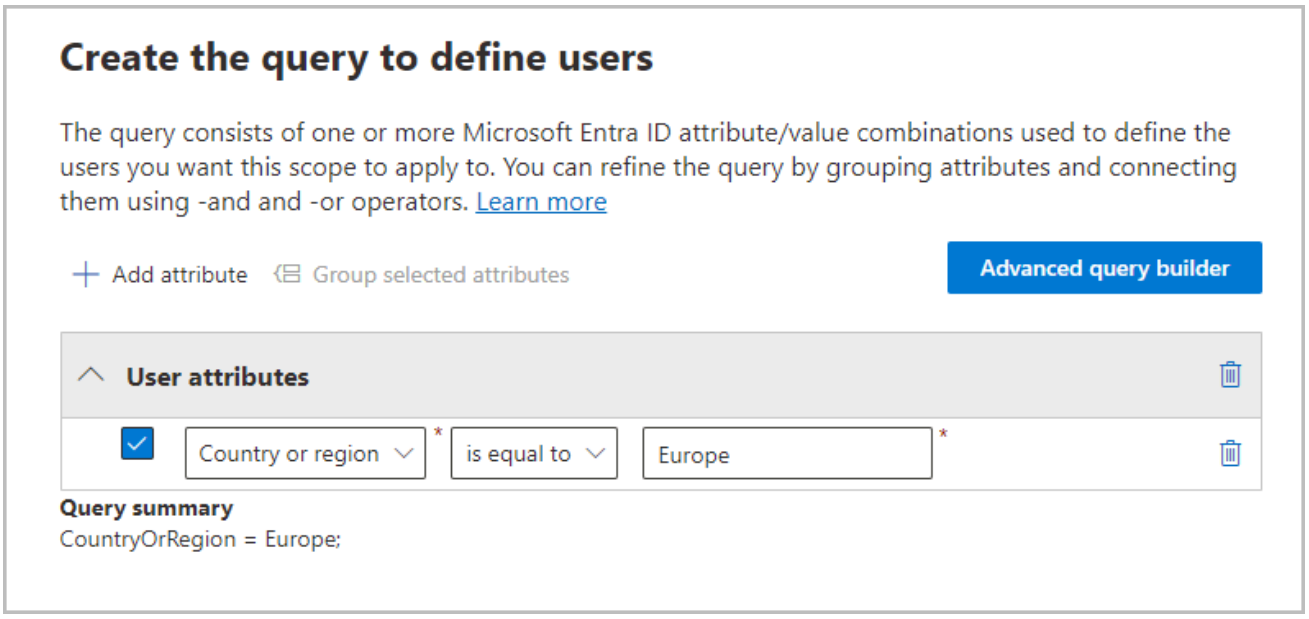
3. Follow the prompts in the configuration where you're first asked to assign an administrative unit. If your account is assigned administrative units, you must select one administrative unit that restricts the scope membership.

   > **Note**
   >
   > Because administrative units don't yet support SharePoint sites, you can't create an adaptive scope for SharePoint sites if you select administrative units.

   If you don't want to restrict the adaptive scope by using administrative units, or your organization hasn't configured administrative units, keep the default of **Full directory**.

4. Select the type of scope, and then select the attributes or properties you want to use to build the dynamic membership. Type in the attribute or property values.

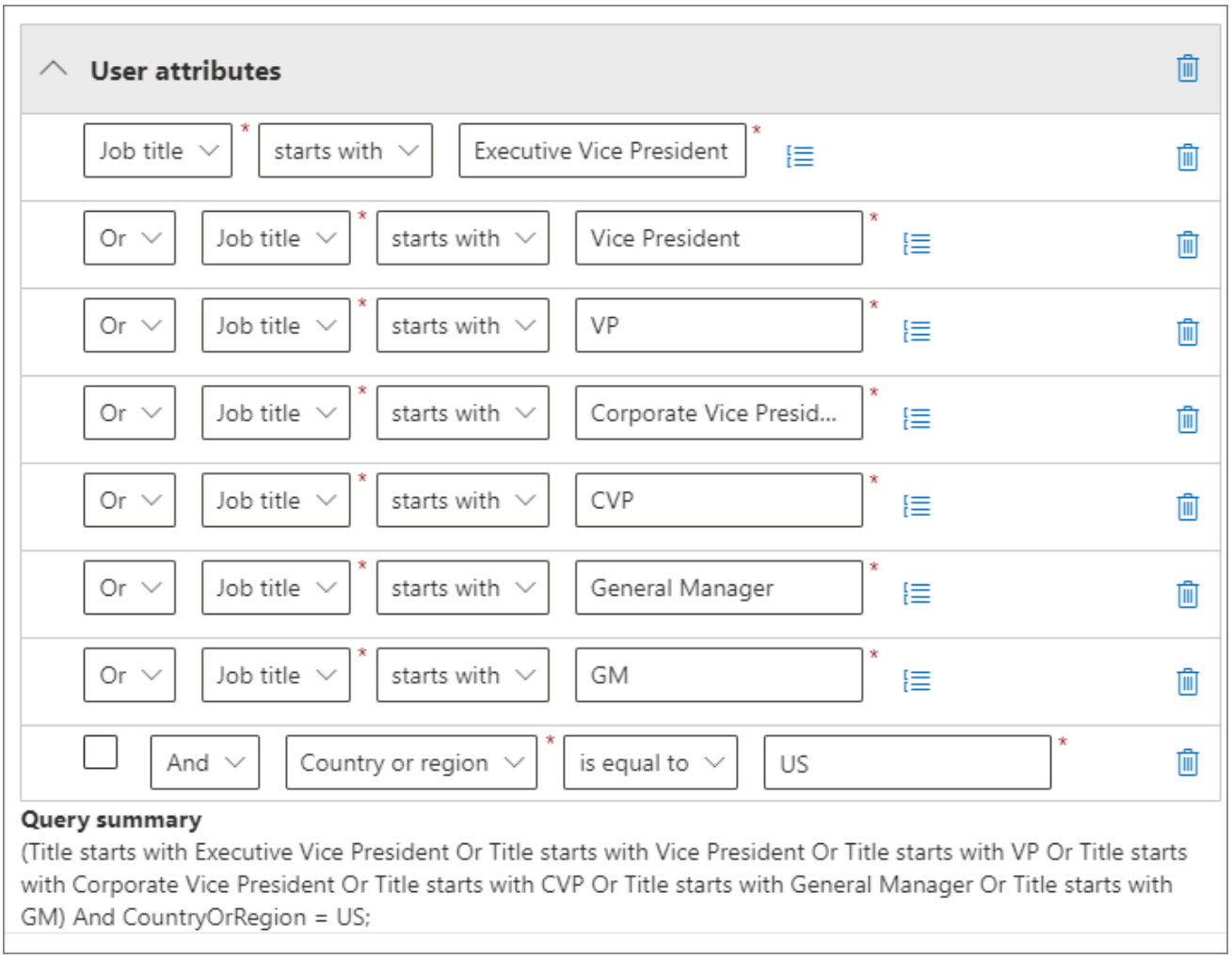
For example, to configure an adaptive scope that identifies users in Europe, first select **Users** as the scope type. Then select the **Country or region** attribute, and type in **Europe**:



Once daily, this query runs against Entra ID and identifies all users who have the value **Europe** specified for in their account for the **Country or region** attribute.

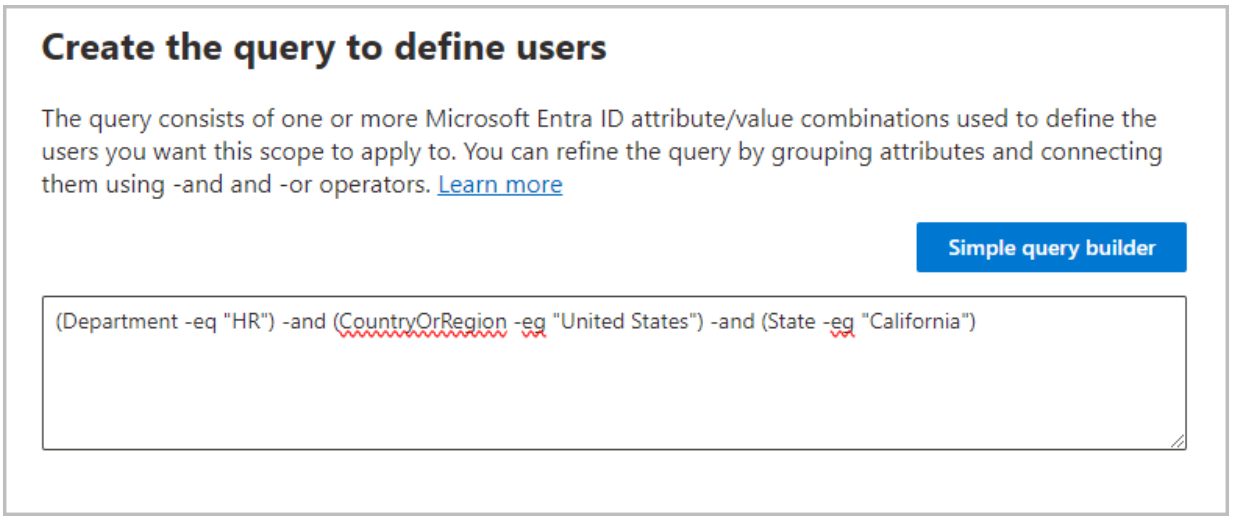
> **Important**
>
> Because the query doesn't run immediately, there's no validation that you typed in the value correctly.

Select **Add attribute** (for users and groups) or **Add property** (for sites) to use any combination of attributes or properties that are supported for their scope type. Use logical operators to build queries. The operators supported are **is equal to**, **is not equal to**, **starts with**, and **not starts with**. You can group selected attributes or properties. For example:



Alternatively, you can select **Advanced query builder** to specify your own queries:

- For **User** and **Microsoft 365 Group** scopes: Use OPATH filtering syntax. For example, to create a user scope that defines its membership by department, country/region, and state:

  One of the advantages of using the advanced query builder for these scopes is a wider choice of query operators:

  - **and**

- **or**
- **not**
- **eq** (equals)
- **ne** (not equals)
- **lt** (less than)
- **gt** (greater than)
- **like** (string comparison)
- **notlike** (string comparison)

- For **SharePoint sites** scopes: Use Keyword Query Language (KeyQL). You might already be familiar with using KeyQL to search SharePoint by using indexed site properties. To help you specify these KeyQL queries, see Keyword Query Language (KeyQL) syntax reference.

  For example, because SharePoint site scopes automatically include all SharePoint site types, which include Microsoft 365 group-connected and OneDrive sites, you can use the indexed site property **SiteTemplate** to include or exclude specific site types. The templates you can specify:

  - `SITEPAGEPUBLISHING` for modern communication sites
  - `GROUP` for Microsoft 365 group-connected sites
  - `TEAMCHANNEL` for Microsoft Teams private channel sites
  - `STS` for a classic SharePoint team site
  - `SPSPERS` for OneDrive sites

  So to create an adaptive scope that includes only modern communication sites and excludes Microsoft 365 group-connected and OneDrive sites, specify the following KeyQL query:

  ```
  Console
  SiteTemplate=SITEPAGEPUBLISHING
  ```

  You can validate these advanced queries independently from the scope configuration.

  > **Note**
  >
  > You must use the advanced query builder if you want to exclude inactive mailboxes. Or conversely, target just inactive mailboxes. For this configuration, use the OPATH property *IsInactiveMailbox*:

- To exclude inactive mailboxes, make sure the query includes: `(IsInactiveMailbox -eq "False")`
- To target just inactive mailboxes, specify: `(IsInactiveMailbox -eq "True")`

5. Create as many adaptive scopes as you need. You can select one or more adaptive scopes when you create your policy.

It can take up to five days for the queries to fully populate and changes aren't immediate. Factor in this delay by waiting a few days before you add a newly created scope to a policy.

> **Tip**
>
> For the data lifecycle management and records management solutions, use the **policy lookup option** to help you identify policies that are currently assigned to specific users, sites, and Microsoft 365 groups.

Validate advanced queries

You can manually validate advanced queries by using PowerShell and SharePoint search:

- Use PowerShell for the scope types **Users** and **Microsoft 365 Groups**
- Use SharePoint search for the scope type **SharePoint sites**

To run a query by using PowerShell

1. Connect to Exchange Online PowerShell by using an account with appropriate Exchange Online Administrator permissions.

2. Use either Get-Recipient, Get-Mailbox, or Get-User with the *-Filter* parameter and your OPATH query for the adaptive scope enclosed in curly brackets (`{`,`}`). If your attribute values are strings, enclose these values in double or single quotes.

   You can determine whether to use Get-Mailbox, Get-Recipient, or Get-User for validation by identifying which cmdlet is supported by the OPATH property that you choose for your query.

   > **Important**

Get-Mailbox doesn't support the *MailUser* recipient type, so you must use Get-Recipient or Get-User to validate queries that include on-premises mailboxes in a hybrid environment.

To validate a **User** scope, use the appropriate command:

- `Get-Mailbox` with *-RecipientTypeDetails UserMailbox,SharedMailbox,RoomMailbox,EquipmentMailbox*
- `Get-Recipient` with *-RecipientTypeDetails UserMailbox,MailUser,SharedMailbox,RoomMailbox,EquipmentMailbox*

To validate a **Microsoft 365 Group** scope, use:

- `Get-Mailbox` with *-GroupMailbox* or `Get-Recipient` with *-RecipientTypeDetails GroupMailbox*

For example, to validate a **User** scope related to the **Department** attribute set to the value **Marketing**, you could use:

PowerShell

```
Get-Recipient -RecipientTypeDetails UserMailbox,MailUser -Filter {Department -eq "Marketing"} -ResultSize Unlimited
```

To validate a **User** scope related to the **EmailAddresses** attribute generally requires the value to include the **smtp:** prefix. For example to **exclude** a user where EmailAddresses includes **admin@contoso.com**:

PowerShell

```
Get-Mailbox -RecipientTypeDetails UserMailbox -Filter {EmailAddresses -notlike "smtp:admin@contoso.com"} -ResultSize Unlimited
```

To validate a **Microsoft 365 Group** scope related to the group's **CustomAttribute15** attribute set to **Marketing** as the value, you could use:

PowerShell

```
Get-Mailbox -RecipientTypeDetails GroupMailbox -Filter {CustomAttribute15 -eq "Marketing"} -ResultSize Unlimited
```

> **Tip**
>
> When you use these commands to validate a user scope, if the number of recipients returned is higher than expected, it might be because it includes users who don't have a valid license for adaptive scopes. These users don't have the policy settings applied to them.
>
> For example, in a hybrid environment, you might have unlicensed synchronized user accounts without an Exchange mailbox on-premises or in Exchange Online. You can identify these users by running the following command: `Get-User -RecipientTypeDetails User`

3. Verify that the output matches the expected users or groups for your adaptive scope. If it doesn't, check your query and the values with the relevant administrator for Microsoft Entra ID or Exchange.

> **Note**
>
> The output from these commands might not match the member list for the same filter within the scope details on the Adaptive Scopes page. The member list in scope details shows arbitration mailboxes in scope whereas the command output doesn't. Moreover, adaptive scope query changes can take up to five days to become effective and reflected in the scope details view of the Adaptive Scopes page.

To run a query by using SharePoint search

1. Using an account that has the SharePoint admin role, go to `https://<your_tenant>.sharepoint.com/search`.
2. Use the search bar to specify your KeyQL query.
3. Verify that the search results match the expected site URLs for your adaptive scope. If they don't, check your query and the URLs with the relevant administrator for SharePoint.

View adaptive scope membership details

After creating a query for an adaptive scope in Microsoft Purview, you can view the member details of the scope.

Complete the following steps to access the member details page of an adaptive scope in Microsoft Purview:

1. Sign in to the Microsoft Purview portal, if you're not already signed in.
2. Select **Settings** > **Roles and scopes** > **Adaptive scopes**.
3. Select an existing adaptive scope from the list.
4. From the **Details** flyout pane, select **Scope details**.
5. A list of scope members is shown. Review the **State** column in the list that shows **Added** for members within the adaptive scope or **Removed** if the member was previously in the adaptive scope but removed now they no longer match the scope query.
6. Optionally, use the **Export** option to download a list of the displayed scope members to a CSV file.

   If the scope membership is large (1,000 or more), or you want to customize the export, use the PowerShell cmdlet Get-AdaptiveScopeMembers.

> **Note**
>
> It might take up to five days for the member details list to update added or removed members. The adaptive scope can apply to more than 1 million members. However, the member details can only show up to 1 million members across both added and removed members.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 09/11/2026

Administrative units in Microsoft Purview

By using administrative units in Microsoft Entra ID, you can subdivide your organization into smaller units and restrict administrative permissions to specific parts of your Microsoft Entra organization. You can create, delete, and edit administrative units in Microsoft Entra, and manage the users or groups that are members of each unit.

Microsoft Purview integrates with administrative units by enabling you to assign role group members to specific units. Microsoft Purview solutions that support administrative units restrict visibility and management permissions to members of the assigned unit.

For example, you can use administrative units to delegate permissions to administrators for each geographic region in a large multinational organization or to group administrator access by department. You can then create region-specific or department-specific policies, and view user activity as a result of those policies and administrative unit assignments. You can also use administrative units as an initial scope for a policy, where the users eligible for the policy depend on administrative unit membership.

If you're using adaptive scopes for compliance policies, see How adaptive scopes work with Microsoft Entra administrative units.

Supported solutions

The following Microsoft Purview solutions support administrative units:

Expand table

Solution	Configuration support
Data lifecycle management	Role groups, retention policies, and retention label policies
Data Loss Prevention (DLP)	Role groups and DLP policies
Communication Compliance	Role groups and policies
Insider Risk Management	Role groups and policies
Records management	Role groups, retention policies, retention label policies, and adaptive scopes
Sensitivity labeling	Role groups, sensitivity label policies, and auto-labeling policies

The configuration for administrative units automatically flows down to the following features:

- **Alerts**:
  - DLP alerts are visible only to users in assigned administrative units.
- **Activity explorer**:
  - Activity events are visible only to users in assigned administrative units.
- **Adaptive scopes**:
  - Restricted administrators can select, create, edit, and view adaptive scopes only for users in those administrators' assigned administrative units.
  - When a restricted administrator configures a policy that uses adaptive scopes, the administrator can select only adaptive scopes that are assigned to their administrative units.
- **Audit log**:
  - Search access.
- **Communication Compliance**:
  - Policy lookup and configuration: Restricted administrators can create or manage policies only for users assigned to their administrative units.
  - Alerts and policy matches: Restricted administrators can investigate user activities only for users within their assigned administrative units.
- **Data lifecycle management and records management**:
  - Policy lookup: Restricted administrators see policies only from users within their assigned administrative units.
  - Disposition review and verification: Restricted administrators can add reviewers only from within their assigned administrative units, and see disposition reviews and items disposed only from users within their assigned administrative units.
- **Insider Risk Management**:
  - Policy lookup and configuration: Restricted administrators can create or manage policies only for users assigned to their administrative units.
  - User activities: Restricted administrators can start scoring activities or investigate user activities only for users within their assigned administrative units.
  - Alerts and cases: Restricted administrators can view and investigate alerts and cases only for users within their assigned administrative units.

> **Note**
>
> Although DLP and Insider Risk Management both surface **alerts** in Microsoft Defender XDR, Microsoft Defender XDR only supports administrative units for DLP. Microsoft Defender XDR supports up to 100 administrative units.

Permissions for administrative units

To assign a role group member to an administrative unit, administrators must be assigned the *Role management* role. For more information about managing role groups and roles in Microsoft Purview, see Permissions in the Microsoft Purview portal. To view the complete list of default role groups and the roles assigned to each, see Role groups in Microsoft Purview.

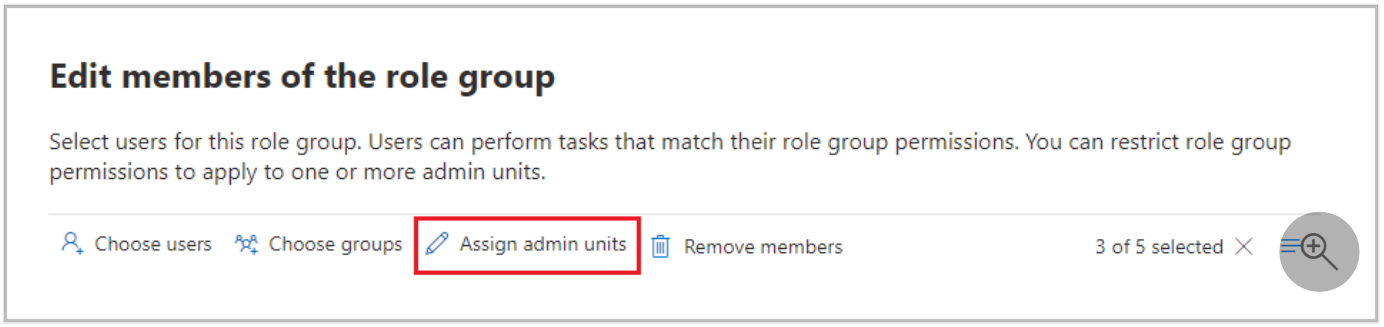
You can assign role group members to administrative units within the following built-in role groups:

Expand table

Role group	Solution area
Communication Compliance	Communication Compliance
Communication Compliance Administrators	Communication Compliance
Communication Compliance Analysts	Communication Compliance
Communication Compliance Investigators	Communication Compliance
Compliance Administrator	Cross-solution
Compliance Data Administrators	Cross-solution
Global Reader	Cross-solution
Information Protection	Information Protection
Information Protection Admins	Information Protection
Information Protection Analyst	Information Protection
Information Protection Investigators	Information Protection
Information Protection Readers	Information Protection
Insider Risk Management	Insider Risk Management
Insider Risk Management Admins	Insider Risk Management
Insider Risk Management Analysts	Insider Risk Management
Insider Risk Management Approvers	Insider Risk Management
Insider Risk Management Investigators	Insider Risk Management
Insider Risk Management Session Approvers	Insider Risk Management

Role group	Solution area
Organization Management	Cross-solution
Records Management	Records Management
Security Administrator	Cross-solution
Security Operator	Cross-solution
Security Reader	Cross-solution

When you assign role groups, you can select individual members or groups, and then select the **Assign admin units** option to select administrative units that you define in Microsoft Entra ID:



> **Important**
>
> **Assign admin units** is always available when you create custom role groups. You can assign administrative units for any custom role group.

These administrators, referred to as *restricted administrators*, can select one or more of their assigned administrative units to automatically define the initial scope of policies that they create or edit. If administrators don't have administrative units assigned (*unrestricted administrators*), they can assign policies to the entire directory without being required to select individual administrative units.

> **Important**
>
> After you assign administrative units to members of the role groups, these restricted administrators can't see and edit existing policies. However, there's no operational change to these policies and they remain visible and can be edited by unrestricted administrators.
>
> Restricted administrators also can't see historical data by using features that support administrative units, such as activity explorer and alerts. They remain visible to unrestricted

administrators. Going forward, restricted administrators can see this related data for their assigned administrative units only.

Prerequisites for administrative units

Before configuring administrative units for Microsoft Purview solutions, make sure your organization and users meet the following subscription and licensing requirements:

- Microsoft Entra ID P1 or P2 licensing
- Microsoft Purview licensing (one of the following subscriptions):
  - Microsoft 365 E5/A5/G5
  - Microsoft 365 E5/A5/G5/F5 Compliance or F5 Security & Compliance
  - Microsoft 365 E5/A5/G5/F5 Information Protection & Governance
  - Microsoft 365 E5/A5/F5 Insider Risk Management

  For more information about licensing options for Microsoft Purview features, see the Microsoft Purview service description.

Configure administrative units

Complete the following steps to configure and use administrative units with Microsoft Purview solutions:

1. Create administrative units to restrict the scope of role permissions in Microsoft Entra ID.
2. Add users and distribution groups to administrative units.

   > **Important**
   >
   > Members of Dynamic Distribution Groups don't automatically become members of an administrative unit.

3. If you create geographic region or department-based administrative units, configure administrative units by using dynamic membership rules.

   > **Note**

> You can't add groups to an administrative unit that uses dynamic membership rules. If needed, create two administrative units, one for users and one for groups.

4. Use any of the role groups from the Microsoft Purview solutions that support administrative units to assign administrative units to members.

When restricted administrators create or edit policies that support administrative units, they can select administrative units so that only the users in those administrative units are eligible for the policy:

- *Unrestricted administrators* don't have to select administrative units as part of the policy configuration. They can keep the default of the entire directory or select one or more administrative units.
- *Restricted administrators* must select one or more administrative units as part of the policy configuration.

Later in the policy configuration, administrators who select administrative units must include or exclude (if supported) individual users and groups from the administrative units that they previously selected for the policy.

For information about administrative units that is specific to each supported solution, see the following sections:

- For audit: Scoping access to audit logs using administrative units
- For communication compliance: Consider administrative units if you want to scope user permissions to a region or department
- For data lifecycle management: Support for administrative units
- For DLP: Administrative Unit restricted policies
- For insider risk management: Consider administrative units if you want to scope user permissions to a region or department
- For records management: Support for administrative units
- For sensitivity labeling: Support for administrative units

SharePoint site support for administrative units

Microsoft Purview now supports adding SharePoint sites to administrative units. By using this feature, you can tailor visibility and management control of Microsoft Purview administrators within the Microsoft Purview portal. This support is only available for Microsoft Information

Protection auto-labeling policies and Microsoft Data Loss Prevention policies that support application to SharePoint sites.

> **Note**
>
> You manage administrative unit creation, deletion, and Microsoft Entra resource membership from Microsoft Entra ID. Any administrative unit you create appears in the listing for Microsoft Purview.

It can take up to five days for a SharePoint site query to fully populate. Changes aren't immediate. Wait until the query fully populates before you associate a policy with an administrative unit.

Configure SharePoint sites for administrative units

Complete the following steps to configure administrative units. These steps guide you through administrative unit creation, how to associate resources to that administrative unit, and how to assign Microsoft Purview role group members to administrative units. After you create administrative units, use the following steps to associate SharePoint sites to the administrative unit.

Microsoft Purview customers entitled to this feature can now access **Administrative units** under the **Roles and scopes** in Microsoft Purview portal settings. Within **Administrative units**, select an administrative unit and edit it to associate SharePoint sites to the administrative unit.

To complete the following steps to configure SharePoint sites within administrative units for use in Microsoft Purview, you must have the admin unit extension manager role assigned. An administrator with role management rights in Microsoft Purview assigns this role, either by using a built-in role group with this role or by using a custom role group.

1. Go to the Microsoft Purview portal.
2. Select **Settings** and select **Roles and scopes**.
3. Select **Administrative units**.
4. Select an existing administrative unit from the list.
5. Select **Edit**.
6. Create a query to associate SharePoint sites with the administrative unit.
7. Use any of the role groups from the Microsoft Purview solutions that support administrative units to assign administrative units to members.

Queries to associate SharePoint sites with administrative units support site properties for Site URL, site name, and `RefinableString00`-`RefinableString99`. These queries apply both to SharePoint sites and OneDrive accounts, with the exception of shared channel SharePoint sites. The property names for sites are based on SharePoint site managed properties. For more information about associating custom properties to managed properties (`RefinableString00`-`RefinableString99`), see Using Custom SharePoint Site Properties to Apply Microsoft 365 Retention with Adaptive Policy Scopes.

Test your query

To test your query by using SharePoint search, complete the following steps:

1. With an account assigned the SharePoint admin role, go to `https://<your_tenant>.sharepoint.com/search`.
2. Use the search bar to enter the query shown in the **query summary** for your SharePoint sites in the administrative unit.
3. Verify that the search results match the expected site URLs for your administrative unit. If they don't, check your query and the URLs with the relevant administrator for SharePoint.

The same restricted and unrestricted administrator behavior that applies to user-scoped policies also applies to policies that include SharePoint sites. Additionally, administrators selecting administrative units can't include or exclude individual SharePoint sites. SharePoint sites support application to all sites associated with the administrative unit.

> **Important**
>
> To remove associated sites from an administrative unit, edit a SharePoint query for an administrative unit to result in no site memberships. To clear site membership, make the query result in no sites.

View SharePoint site membership

After you create a SharePoint query for an administrative unit in Microsoft Purview, you can view the site members of the administrative unit. You can only view user and group membership of an administrative unit from the Microsoft Entra ID portal.

Complete the following steps to access the member details page and see SharePoint site members of an administrative unit in Microsoft Purview:

1. Go to the Microsoft Purview portal.
2. Select **Settings** and select **Roles and scopes**.
3. Select **Administrative units**.
4. Select an existing administrative unit from the list.
5. Select **Member Details**.
6. View the list of SharePoint site members.
7. Check the **State** column in the list, which shows *Added* for sites added to the administrative unit or *Removed* if the site was previously in the administrative unit but is removed after it's no longer a match for the SharePoint query.
8. Use the **Export** function to download a list of the displayed sites to a CSV file.

> **Note**
>
> It might take up to five days for the member details list to update added or removed sites.

For information about administrative units and the SharePoint site in Microsoft Purview solutions, see:

- **DLP**: Administrative Unit restricted policies
- **Sensitivity labels**: Support for administrative units

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 06/04/2026

Get started with data explorer

Data explorer allows you to natively view the items that were summarized on the overview page.

Document preview and download is disabled if the document is encrypted to mitigate security risks.

Prerequisites

Licensing

For information on licensing, see

- Microsoft 365 Enterprise Plans
- Microsoft 365 Service Descriptions

Permissions

> **Important**
>
> Access to data explorer is highly restricted because it lets you read the contents of scanned files. These permissions supersede permissions that are locally assigned to the items, which allow viewing of the content.

Microsoft Entra ID roles to access the data explorer tab

In order to get access to the data explorer tab, an account must be assigned membership in any one of these roles or role groups. Membership in these roles doesn't allow you to view the list of items in content explorer or to view the contents of the items in data explorer.

- Compliance administrator
- Security administrator
- Compliance data administrator
- Global administrator

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your

organization. Learn more about Microsoft Purview roles and permissions.

Required Microsoft Purview permissions to access items in data explorer

Access to data explorer is configured in the Microsoft Purview portal. There are two roles that grant access to data explorer. **Data Explorer List viewer** and **Data Explorer Content viewer**.

These are independent role groups and the permissions they grant aren't cumulative. For example, if you want to grant an account the ability to view the items and their locations only, grant Data Explorer List viewer rights. If you want that same account to also be able to view the contents of the items in the list, grant Data Explorer Content viewer rights as well.

You can also assign either or both of the roles to a custom role group to tailor access to content explorer.

- **Data Explorer List viewer**: Membership in this role group allows you to see each item and its location in list view. The `data classification list viewer` role has been pre-assigned to this role group.
- **Data Explorer Content viewer**: Membership in this role group allows you to view the contents of each item in the list. The `data classification content viewer` role has been pre-assigned to this role group. Additionally, this role is also required to view name of items in list view, which may contain sensitive data.

> **Note**
>
> In preview, data explorer supports administrative units. Members of role groups that have the *data classification list viewer* or *data classification content viewer* role will receive these respective role permissions can be restricted by administrative unit assignments within data explorer. For more on administrative unit support in Purview, see Administrative units support in Microsoft Purview.

Role management role holders in Microsoft Purview can assign the necessary data explorer list viewer, and data explorer content viewer role group membership.

Fine tune Microsoft Purview Roles and Role Groups

There are roles and role groups that you use to fine tune your access controls.

Here's a list of applicable roles. To learn more about them, see Permissions in the Microsoft Purview portal.

- Information Protection Admin
- Information Protection Analyst
- Information Protection Investigator
- Information Protection Reader

Here's a list of applicable role groups. To learn more, see Permissions in the Microsoft Purview portal.

- Information Protection
- Information Protection Admins
- Information Protection Analysts
- Information Protection Investigators
- Information Protection Readers

Data explorer

Data explorer shows a current snapshot of the items that have a sensitivity label, a retention label or have been classified as a sensitive information type in your organization.

Sensitive information types

A DLP policy can help protect sensitive information, which is defined as a **sensitive information type**. Microsoft 365 includes definitions for many common sensitive information types from across many different regions that are ready for you to use. For example, a credit card number, bank account numbers, and national ID numbers.

Sensitivity labels

A sensitivity label is simply a tag that indicates the value of the item to your organization. It can be applied manually, or automatically. Once applied, the label gets embedded in the document and will follow the document everywhere it goes. A sensitivity label enables various protective behaviors, such as mandatory watermarking or encryption.

Sensitivity labels must be enabled for files that are in SharePoint and OneDrive in order for the corresponding data to surface in the data classification page. For more information, see Enable sensitivity labels for files in SharePoint and OneDrive.

Encrypted sensitivity labels are not surfaced on content explorer for SharePoint and OneDrive.

Retention labels

A retention label allows you to define how long a labeled item is kept and the steps to be taken prior to deleting it. They're applied manually or automatically via policies. They can play a role in helping your organization stay in compliance with legal and regulatory requirements.

Trainable Classifiers

A trainable classifiers allows you to identify sensitive data by using examples of the data you're interested in rather than identifying elements in the item (pattern matching). You can use built-in classifiers or train a classifier with your own content.

How to use data content explorer

1. Sign in to the Microsoft Purview portal > **Solutions** > **Information Protectionn** > **Explorers** > **Data explorer**.
2. If you know the name of the label, classifiers or the sensitive information type, you can type that into the filter box.
3. Alternately, you can browse for the item by expanding the type and selecting the type from the list.
4. Select a data source under **Data source** and drill down the folder structure to the item.
5. Double-click to open the item natively in data explorer.

Export

The **Export** control will create a .csv file that contains a listing of whatever the focus of the pane is.

> **Note**
>
> It can take up to *seven days* for counts to be updated in data content explorer, *fourteen days* for files that are in SharePoint.

Filter

When you drill down into a location, such as an Exchange or Teams folder, or a SharePoint or OneDrive site, the **Filter** tool appears.

The scope of the search tool is what is displaying in the **All locations** pane and what you can search on varies depending on the selected location.

When **Exchange** or **Teams** is the selected location, you can search on the full email address of the mailbox, for example `user@domainname.com`.

When either **SharePoint** or **OneDrive** are selected location, the search tool will appear when you drill down to site names, folders and files.

You can search on:

Expand table

value	example
full site name	`https://contoso.onmicrosoft.com/sites/sitename`
file name	`RES_Resume_1234.txt`
text at the beginning of file name	`RES`
text after an underscore character (_) in file name	`Resume` or `1234`
file extension	`txt`

Provide match/not a match accuracy feedback in data explorer

You can view the number of matches a SIT or trainable classifier has in **Data explorer**. You can also provide feedback on whether an item is actually a match or not using the **Match**, **Not a Match** feedback mechanism and use that feedback to tune your classifiers. See, Increase classifier accuracy for more information.

See also

- Learn about sensitivity labels
- Learn about retention policies and retention labels
- Sensitive information type entity definitions.md
- Learn about Microsoft Purview Data Loss Prevention

Last updated on 01/23/2026

Get started with Content Explorer (classic)

Content Explorer (classic) lets you natively view the items summarized on the overview page. When a document is encrypted, you can't access the document preview, and download is disabled to mitigate any security risks.

Prerequisites

Licensing

For information on licensing, see

- Microsoft 365 Enterprise Plans
- Microsoft 365 Service Descriptions

Permissions

One set of permissions grants access to the content explorer tab, and a more restrictive set grants access to the items in content explorer.

Permissions to access the content explorer tab

To access the content explorer tab, assign an account membership in any one of these Entra ID roles or Purview roles. Membership in these roles doesn't grant permission to view the list of items in content explorer or to view the contents of the items in content explorer.

Entra ID roles

- Compliance administrator
- Security administrator
- Compliance data administrator

Microsoft Purview Roles and Role Groups

You can use roles and role groups to fine tune your access controls beyond the above guidance. Here's a list of applicable roles. To learn more about them, see Permissions in the Microsoft Purview portal.

- Information Protection Admin
- Information Protection Analyst

- Information Protection Investigator
- Information Protection Reader

Here's a list of applicable role groups. To learn more, see Permissions in the Microsoft Purview portal.

- Information Protection
- Information Protection Admins
- Information Protection Analysts
- Information Protection Investigators
- Information Protection Readers

Required permissions to access items in content explorer

Access to content explorer is highly restricted because it lets you read the contents of scanned files. Therefore, you need additional permissions to access content explorer data.

These permissions supersede permissions that are locally assigned to the items, which allow viewing of the content. The roles that grant access to content explorer are:

- **Content Explorer List viewer**: Membership in this role group allows you to see each item and its location in list view. The `data classification list viewer` role is preassigned to this role group.

- **Content Explorer Content viewer**: Membership in this role group allows you to view the contents of each item in the list. The `data classification content viewer` role is preassigned to this role group. Additionally, this role is also required to view name of items in list view, which might contain sensitive data.

The account you use to access content explorer **must be in one or both of the role groups**. These are independent role groups and aren't cumulative. For example, if you want to grant an account the ability to view the items and their locations only, grant Content Explorer List viewer rights. If you want that same account to also be able to view the contents of the items in the list, grant Content Explorer Content viewer rights as well.

You can also assign either or both of the roles to a custom role group to tailor access to content explorer.

Role management role holders in Microsoft Purview can assign the necessary Content Explorer List Viewer, and Content Explorer Content Viewer role group membership.

> **Note**
>
> Content Explorer support administrative units. Members of role groups that have the *data classification list viewer* or *data classification content viewer* role receive these respective role permissions can be restricted by administrative unit assignments within content explorer. For more on administrative unit support in Purview, see Administrative units support in Microsoft Purview.

Content explorer

Content explorer shows a current snapshot of the items that have a sensitivity label, a retention label, or are classified as a sensitive information type in your organization.

Sensitive information types

A DLP policy helps protect sensitive information, which is defined as a **sensitive information type**. Microsoft 365 includes definitions for many common sensitive information types from across many different regions that are ready for you to use. For example, credit card numbers, bank account numbers, and national ID numbers.

Sensitivity labels

A sensitivity label is a tag that indicates the value of the item to your organization. You can apply it manually or automatically. Once applied, the label gets embedded in the document and follows the document everywhere it goes. A sensitivity label enables various protective behaviors, such as mandatory watermarking or encryption.

You must enable sensitivity labels for files in SharePoint and OneDrive for the corresponding data to surface in the data classification page. For more information, see Enable sensitivity labels for files in SharePoint and OneDrive.

Encrypted sensitivity labels don't surface on content explorer for SharePoint and OneDrive.

Retention labels

A retention label lets you define how long to keep a labeled item and the steps to take before deleting it. You can apply them manually or automatically through policies. They help your organization stay in compliance with legal and regulatory requirements.

How to use content explorer

1. Sign in to the Microsoft Purview portal > **Solutions** > **Data Lifecycle Management** > **Explorers** > **Content explorer**.
2. If you know the name of the label or the sensitive information type, type it into the filter box.
3. Alternately, you can browse for the item by expanding the label type and selecting the label from the list.
4. Select a location under **All locations** and drill down the folder structure to the item.
5. Double-click to open the item natively in content explorer.

Export

The **export** control creates a .csv file that contains a listing of whatever the focus of the pane is.

> **Note**
>
> It can take up to *seven days* for counts to update in content explorer and *14 days* for files that are in SharePoint.

Filter

When you drill down into a location, such as an Exchange or Teams folder, or a SharePoint or OneDrive site, the **Filter** tool appears.

The scope of the search tool is what displays in the **All locations** pane. What you can search on varies depending on the selected location.

When you select **Exchange** or **Teams**, you can search on the full email address of the mailbox, for example `user@domainname.com`.

When you select **SharePoint** or **OneDrive**, you can search on site names, folders, and files.

You can search on:

Expand table

Value	Example
Full site name	`https://contoso.onmicrosoft.com/sites/sitename`

Value	Example
File name	`RES_Resume_1234.txt`
Text at the beginning of file name	`RES`
Text after an underscore character (_) in file name	`Resume` or `1234`
File extension	`txt`

Save views in Content Explorer

You can create customized views in content explorer, by configuring what SITs, Sensitivity Labels, Trainable Classifiers, and Retention Labels which are important for you. A view can be made default, to quickly load the view next time content explorer is opened.

Provide match or not a match accuracy feedback in content explorer

In **Content explorer**, you can see how many matches a sensitive information type (SIT) or trainable classifier has. You can also give feedback on whether an item is actually a match by using the **Match** and **Not a Match** feedback options. Use this feedback to improve your classifiers. For more information, see Increase classifier accuracy.

> **Note**
>
> Updating a custom sensitive information type (SIT) definition doesn't by itself reclassify unchanged historical files. For SharePoint and OneDrive files that need updated classification, use on-demand classification, and review its scope, costs, and active-policy effects first. A file's last-modified date alone doesn't establish when it was classified. For the distinction between classification results and auto-labeling policy matches, see Auto-labeling classification and simulation concepts.

Email attachment preview

For Exchange items, content explorer supports the preview of email attachments without the need to download the email. The following email attachment file types are supported: `.doc, .docx, .xlsx, .xls, .xlsm, .xlsb, .pptx, .pdf, .txt, .csv, .html, .xml, .json, .rtf, .cpp, .cs, .py, .jpg, .png, .jpeg, and .eml`

Content explorer document preview supports Excel files up to 25 MB.

See also

- Learn about sensitivity labels
- Learn about retention policies and retention labels
- Sensitive information type entity definitions.md
- Learn about Microsoft Purview Data Loss Prevention

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 10/14/2026

Get started with activity explorer

Activity explorer lets you monitor what's being done with your labeled content. Activity explorer provides a historical view of activities on your labeled content. The activity information comes from the Microsoft 365 unified audit logs. It's transformed and then made available in the activity explorer UI. Activity explorer reports on up to 30 days worth of data.

Activity explorer gives you multiple ways to sort and view the data.

When activities appear

Activity explorer isn't a real-time view. Because its data comes from the Microsoft 365 unified audit log, activities appear on the same schedule as audit records — not the moment they occur.

- For core services (Exchange, SharePoint, OneDrive, and Teams), allow 60 to 90 minutes after an activity for it to appear in activity explorer. Other services can take longer, and Microsoft doesn't guarantee a specific time for an activity to become available. For more information, see Search the audit log.
- After you enable or change a policy, allow time for the policy to reach the workload *and* for the resulting activities to flow in. An empty view right after you turn on a policy usually means the data pipeline hasn't caught up yet, not that the policy failed.
- Devices report endpoint activities only while they're online. An offline device backfills its activities after it reconnects.

> **Note**
>
> If a recent activity isn't showing yet, confirm that it falls within the 30 days of data that activity explorer reports on, and allow the audit-log availability time described earlier before you treat a missing activity as a defect. To confirm an activity independently, search for it in Audit.

Filters

Filters are the building blocks of activity explorer. Each filter focuses on a different dimension of the collected data. You can use about 50 different individual filters, including:

- Date range

- Activity type
- Location
- Sensitivity label
- User
- Client IP
- Device name
- Is protected

To see all the filters, open the filter pane in activity explorer and look at the dropdown list.

> **Note**
>
> Filter options are generated based on the first 500 records to ensure optimal performance. This limitation might cause some values to not appear in the filter dropdown. For endpoint events, only the most restrictive DLP rule appears. Filters you apply in activity explorer also operate based on this most restrictive rule.

Filter sets

Activity explorer comes with predefined sets of filters to help save time when you want to focus on a specific activity. Use filter sets to quickly provide you with a view of higher level activities than individual filters do. Some of the predefined filter sets are:

- Endpoint DLP activities
- Sensitivity labels applied, changed, or removed
- Egress activities
- DLP policies that detected activities
- Network DLP activities
- Protected Browser

You can also create and save your own filter sets by combining individual filters.

Microsoft Security Copilot in activity explorer (preview)

In *preview*, Microsoft Security Copilot in Microsoft Purview is embedded in activity explorer. It can help efficiently drill down into Activity data and help you identify activities, files with sensitive info, users, and other details that are relevant to an investigation.

> **Important**
>
> Be sure to check the responses from Security Copilot for accuracy and completeness before taking any action based on the information provided. You can provide feedback to help improve the accuracy of the responses.

Data hunting

Security Copilot skills use all the data available to Microsoft Purview, filters, and filter sets available in activity explorer and use machine learning to provide you with insights into the activity (sometimes referred to as *data hunting*) on your data that is most important to you.

- **Show me the top 5 activities from the past week**
- **Filter and investigate activities**
- **Find files used in specific activities**

Selecting a prompt automatically opens the Security Copilot side card and shows you the results of the query. You can then further refine the query.

Natural language to filter set generation

Use the prompt box to enter complex natural language queries to generate filter sets. For example, you can enter:

```
Filter and investigate files copied to cloud with sensitive info type credit card number for past 30 days.
```

Security Copilot generates a filter set for your query. Review the filter to make sure it fits your needs, then apply it to the data.

Prerequisites

SKU/subscriptions licensing

For information on licensing, see

- Microsoft 365 Enterprise Plans
- Microsoft 365 Service Descriptions

Permissions

An account must be explicitly assigned membership in any one of these role groups, or must be explicitly granted the role.

Roles and role groups

Use roles and role groups to fine-tune your access controls. For more information, see Permissions in the Microsoft Purview portal.

Microsoft Purview roles

- Information Protection Admin
- Information Protection Analyst
- Information Protection Investigator
- Information Protection Reader

Microsoft Purview role groups

- Information Protection
- Information Protection Investigators
- Information Protection Analysts
- Information Protection Admins & Information Protection Readers (both role groups needs to be assigned)

Microsoft 365 roles

- Compliance Admins
- Security Admins
- Compliance Data Admins

Microsoft 365 role groups

- Compliance Administrator
- Security Administrator
- Security Reader

Activity types

Activity explorer gathers information from the audit logs of multiple sources of activities.

Some examples of the **Sensitivity label activities** and **Retention labeling activities** from applications native to Microsoft Office, the Microsoft Purview Information Protection client and

scanner, SharePoint, Exchange (sensitivity labels only), and OneDrive include:

- Label applied
- Label changed (upgraded, downgraded, or removed)
- Auto-labeling simulation
- File read

For the current list of activities listed in Activity explorer, go into Activity explorer and open the activity filter. The list of activities is available in the dropdown list.

Labeling activity specific to the Microsoft Purview Information Protection client and scanner that comes into Activity explorer includes:

- Protection applied
- Protection changed
- Protection removed
- Files discovered

For more detailed information on what labeling activity makes it into Activity explorer, see Labeling events available in Activity explorer.

Additionally, Activity Explorer gathers DLP policy match events from Microsoft 365 workloads such as Exchange, SharePoint, OneDrive, Teams chat and channels, and on-premises SharePoint folders, libraries, and file shares. When you enable Endpoint data loss prevention (DLP), Activity Explorer also includes device-level activities from onboarded Windows 10, Windows 11, and the three most recent major macOS versions.

Enhanced matched conditions for Exchange DLP events (preview)

For Exchange Online DLP events, Activity Explorer surfaces enhanced matched condition details for non-sensitive information type (SIT) conditions in addition to SIT matches. Every non-SIT condition that contributed to a DLP policy match is displayed with three levels of detail:

- **Condition name**: The specific policy condition that was matched, for example, *Sender domain is* or *Attachment's file extension is*.
- **Matched value**: The actual value that triggered the condition match, for example, `contoso.com` or `.docx`.
- **Source**: The part of the message where the match was found, for example, the message header, envelope, or attachment.

Enhanced matched condition details appear on the event detail flyout for Exchange DLP events in Activity Explorer and the DLP Alerts dashboard. The following condition categories are supported:

- **Sender conditions**: Sender is, Sender domain is, Sender address contains words, Sender address matches patterns, Sender is a member of, Sender IP address is, Sender AD attribute
- **Recipient conditions**: Recipient is, Recipient domain is, Recipient address contains words, Recipient address matches patterns, Recipient is a member of, Recipient AD attribute
- **Attachment conditions**: Attachment's file extension is, Document name contains words, Document name matches patterns, Document property is, Document size equals or is greater than, Document is password protected, Document could not be scanned, Document didn't complete scanning
- **Content conditions**: Content character set contains words, Content is shared from M365, Content is received from
- **Header conditions**: Header contains words, Header matches patterns
- **Message property conditions**: Message size over, Message type is, Message importance is, Subject contains words, Subject matches patterns, Subject or body contains words, Subject or body matches patterns, Has sender overridden the policy tip

For the full reference of Exchange DLP conditions, see Data loss prevention Exchange conditions and actions reference.

Some example events gathered from devices include the following actions taken on files:

- Deletion
- Creation
- Copy to clipboard
- Modify
- Read
- Print
- Rename
- Copy to network share
- Access by an unallowed app

Understanding the actions that are taken on content with sensitivity labels helps you determine whether the controls that you have in place, such as Microsoft Purview Data Loss Prevention policies, are effective. If not, or if you discover something unexpected (such as a large number of items labeled `highly confidential` that are downgraded to `general`), you can manage your policies and take new actions to restrict the undesired behavior.

> **Note**
>
> - Activity explorer doesn't monitor retention activities for Exchange.
> - Legacy Azure Information Protection (AIP) labels are not supported; as a result, they may appear as GUIDs in Activity Explorer.

> **Note**
>
> If a user reports the Teams DLP verdict as a false positive, the activity shows as **DLP info** in the list on Activity explorer. The entry doesn't have any rule and policy match details but shows synthetic values. There's also no incident report generated for false positive reporting.

Activity type events and alerts

This table shows the events that Activity Explorer triggers for three sample policy configurations. The events depend on whether a policy match is detected.

Expand table

Policy configuration	Activity Explorer event triggered for this action type	Activity Explorer event triggered when a DLP rule is matched	Activity Explorer alert triggered
Policy contains a single rule allowing the activity without auditing it.	Yes	No	No
Policy contains two rules: Matches for Rule #1 are allowed; policy matches for Rule #2 are audited.	Yes (Rule #2 only)	Yes (Rule #2 only)	Yes (Rule #2 only)
Policy contains two rules: Matches for both rules are allowed and not audited.	Yes	No	No

See also

- Learn about sensitivity labels
- Learn about retention policies and retention labels
- Learn about sensitive information types

Note: The author created this article with assistance from AI. Learn more

Last updated on 07/23/2026

Labeling activities available in Activity explorer

09/10/2025

> **Tip**
>
> If you're not an E5 customer, use the 90-day Microsoft Purview solutions trial to explore how additional Microsoft Purview capabilities can help your organization manage data security and compliance needs. Start now at the Microsoft Purview trials hub. Learn details about signing up and trial terms.

Sensitivity label applied

This event is generated each time you label an unlabeled document or send an email with a sensitivity label.

- Activity explorer captures this event when you save a labeled document in native Microsoft Office applications and web applications.
- It captures this event when you apply a label for the Microsoft Purview Information Protection client.
- You can also monitor upgrade and downgrade label actions through the *Label event type* field and filter.

Expand table

Source	Reported in Activity explorer	Note
Word, Excel, PowerPoint	Yes	
Outlook	Yes	If you apply a sensitivity label to any email in an email thread, the same label is automatically applied to subsequent replies on the thread. These labeling events appear in activity explorer as automatic labeling events, even when automatic labeling isn't configured.
SharePoint, OneDrive	Yes	
Exchange	Yes	

Source	Reported in Activity explorer	Note
Microsoft Purview Information Protection client and scanner	Yes	The *new label* action maps to *label applied* in activity explorer
Microsoft Information Protection (MIP) SDK	Yes	The *new label* action maps to *label applied* in activity explorer
Rights Management Service (RMS)	Not applicable	
Power BI desktop and web	No	Accessible in the Microsoft 365 audit logs
Microsoft Defender for Cloud Apps	No	

Sensitivity label changed

This event is generated each time a sensitivity label is updated on the document or email.

- For the Microsoft Purview Information Protection client and scanner, and MIP SDK sources, the *upgrade label* and *downgrade label* action maps to Activity explorer *label changed*.
- The event is captured when the action is saved in native Microsoft Office applications and web applications.
- The event is captured when the action occurs for the Microsoft Purview Information Protection client and scanner enforcements.
- You can also monitor upgrade and downgrade labels actions via the *Label event type* field and filter. The *justification* text is also captured, except for SharePoint and OneDrive.
- Sensitivity labeling in Microsoft Office apps: Outlook collects the last action that was generated before *file save/email send* actions. For example, if the user changes label on an email multiple times before sending, the last label found on the email when it is sent is captured in the audit log and then reported in activity explorer.

Expand table

Source	Reported in Activity explorer	Note
Word, Excel, PowerPoint	Yes	

Source	Reported in Activity explorer	Note
Outlook	Yes	
SharePoint, OneDrive	Yes	
Exchange	Yes	
Microsoft Purview Information Protection client	Yes	
Microsoft Purview Information Protection scanner	Yes	
MIP SDK	Yes	
RMS service	Not applicable	
Power BI desktop and Web	No	Accessible in the Microsoft 365 audit logs
Microsoft Defender for Cloud Apps	No	

Sensitivity label removed

This event is generated each time you remove a sensitivity label from a file or document.

- Microsoft Office applications and web applications capture this event when you save the document.
- Microsoft Purview Information Protection client captures this event when you remove the label.
- Sensitivity labeling in Microsoft Office: Outlook collects the last labeling event that was generated before *file save/email send* actions.

Expand table

Source	Reported in Activity explorer	Note
Word, Excel, PowerPoint	Yes	
Outlook	Yes	
SharePoint, OneDrive	Yes	
Exchange	Yes	

Source	Reported in Activity explorer	Note
Microsoft Purview Information Protection client	Yes	The *remove label* action is mapped to the *label removed* action in activity explorer
Microsoft Purview Information Protection scanner	Yes	The *remove label* action is mapped to the *label removed* action in activity explorer
MIP SDK	Yes	The *remove label* action is mapped to the *label removed* action in activity explorer
RMS service	Not applicable	
Power BI desktop and Web	No	Accessible in the Microsoft 365 audit logs
Microsoft Defender for Cloud Apps	No	

Sensitivity label file read

This event is generated each time a sensitivity labeled or protected document is opened.

Expand table

Source	Reported in Activity explorer	Note
Word, Excel, PowerPoint	Yes	
Outlook	No	
SharePoint, OneDrive	No	
Exchange	No	
Microsoft Purview Information Protection client	Yes	The *access* action maps to the *file read* action in activity explorer
Microsoft Purview Information Protection scanner	Yes	The *access* action maps to the *file read* action in activity explorer
MIP SDK	Yes	The *access* action maps to the *file read* action in activity explorer
RMS service	Yes	The *access* action maps to the *file read* action in activity explorer
Power BI desktop and Web	No	Accessible in the Microsoft 365 audit logs

Source	Reported in Activity explorer	Note
Microsoft Defender for Cloud Apps	No	

Files discovered

The Microsoft Purview Information Protection scanner generates this event each time it discovers files while scanning sensitive data in various locations.

Expand table

Source	Reported in Activity explorer	Note
Word, Excel, PowerPoint	Not applicable	
Outlook	Not applicable	
SharePoint, OneDrive	Not applicable	
Exchange	Not applicable	
Microsoft Purview Information Protection client	Not applicable	
Microsoft Purview Information Protection scanner	Yes	The *discover* action maps to the *files discovered* action in activity explorer
MIP SDK	Yes	The *discover* action maps to the *file discovered* action in Activity explorer
RMS service	Not applicable	
Power BI desktop and Web	Not applicable	
Microsoft Defender for Cloud Apps	Not applicable	

Sensitivity label file renamed

This event is generated each time you rename a document with a sensitivity label.

Expand table

Source	Reported in Activity explorer	Note
Word, Excel, PowerPoint	Yes	
Outlook	Not applicable	
SharePoint, OneDrive	No	
Exchange	Not applicable	
Microsoft Purview Information Protection client	No	
Microsoft Purview Information Protection scanner	No	
MIP SDK	No	
RMS service	No	
Power BI desktop and Web	No	
Microsoft Defender for Cloud Apps	No	

File removed

The Microsoft Purview Information Protection scanner generates this event each time it detects that a previously scanned file is removed.

Expand table

Source	Reported in Activity explorer	Note
Word, Excel, PowerPoint	Not applicable	
Outlook	Not applicable	
SharePoint, OneDrive	Not applicable	
Exchange	Not applicable	
Microsoft Purview Information Protection client	Not applicable	
Microsoft Purview Information Protection scanner	Yes	
MIP SDK	Not applicable	
RMS service	Not applicable	
Power BI desktop and Web	Not applicable	
Microsoft Defender for Cloud Apps	Not applicable	

Protection applied

This event is generated the first time you manually add protection to an item that doesn't have a label.

Expand table

Source	Reported in Activity explorer	Note
Word, Excel, PowerPoint	No	
Outlook	No	
SharePoint, OneDrive	Not applicable	
Exchange	No	
Microsoft Purview Information Protection client	Yes	
Microsoft Purview Information Protection scanner	Not applicable	
MIP SDK	Yes	
RMS service	Not applicable	
Power BI desktop and Web	Not applicable	
Microsoft Defender for Cloud Apps	Not applicable	

Protection changed

This event is generated each time you manually change the protection on an unlabeled document.

Expand table

Source	Reported in Activity explorer
Word, Excel, PowerPoint	No
Outlook	No
SharePoint, OneDrive	Not applicable
Exchange	No
Microsoft Purview Information Protection client	Yes
Microsoft Purview Information Protection scanner	Not applicable

Source	Reported in Activity explorer
MIP SDK	Yes
RMS service	Not applicable
Power BI desktop and Web	Not applicable
Microsoft Defender for Cloud Apps	Not applicable

Protection removed

This event is generated each time you manually change the protection on an unlabeled document.

Expand table

Source	Reported in Activity explorer
Word, Excel, PowerPoint	No
Outlook	No
SharePoint, OneDrive	Not applicable
Exchange	No
Microsoft Purview Information Protection client	Yes
Microsoft Purview Information Protection scanner	Not applicable
MIP SDK	Yes
RMS service	Not applicable
Power BI desktop and Web	Not applicable
Microsoft Defender for Cloud Apps	Not applicable

Data loss prevention (DLP) policy matched

This event is generated each time a DLP policy matches a document or an email.

Expand table

Source	Reported in Activity explorer
Exchange	Yes
SharePoint	Yes
OneDrive	Yes
Teams	Yes
Windows 10, Windows 11, and macOS devices	Yes
MAC	No
On-premises	No
Microsoft Defender for Cloud Apps	No

The Endpoint DLP events for devices running Windows 10, Windows 11, and any of the three most recently released major versions of macOS are:

- File deleted
- File created
- File copied to clipboard
- File modified
- File read
- File printed
- File renamed
- File copied to network share
- File accessed by unallowed app

Retention label applied

This event is generated each time you label an unlabeled document or send an email with a retention label.

- The event is captured when you save a document or send an email.

Expand table

Source	Reported in Activity explorer
Exchange	No
SharePoint	Yes

Source	Reported in Activity explorer
OneDrive	Yes

Retention label changed

This event is generated each time you update a label on a document or email.

- The event is captured when you save a document or send an email.

Expand table

Source	Reported in Activity explorer
Exchange	No
SharePoint	Yes
OneDrive	Yes

Retention label removed

This event is generated each time you remove a label from a file or document.

- The event is captured when you save a document or send an email.

Expand table

Source	Reported in Activity explorer
Exchange	No
SharePoint	Yes
OneDrive	Yes

Considerations

- When the recommended label tool tip appears to an end user, the event isn't captured. However, if the user chooses to apply the recommended label, the label appears under the *How applied* field as *Recommended*.
- Sensitive information types aren't currently available for auto-labeling activities from Word, Excel, PowerPoint, and Outlook, SharePoint, and OneDrive.

- When you label items with a sensitivity label and then delete that label, any events involving those items still appear in activity explorer. Instead of showing the name of the deleted sensitivity label, the GUID of the deleted sensitivity label appears.

Learn about connectors for third-party data

Microsoft 365 lets administrators use data connectors to import and archive non-Microsoft, third-party data from social media platforms, instant messaging platforms, and document collaboration platforms, to mailboxes in your Microsoft 365 organization. One primary benefit of using data connectors to import and archive third-party data in Microsoft 365 is that you can apply various Microsoft Purview solutions to the data after you import it. This approach helps you ensure that your organization's non-Microsoft data complies with the regulations and standards that affect your organization.

This article describes the available third-party data connectors, the Microsoft Purview solutions that support imported third-party data, and how to provision a service principal for partner connectors.

Third-party data connectors

The Microsoft Purview portal provides native third-party data connectors from Microsoft to import data from various data sources, such as LinkedIn, Instant Bloomberg, and Twitter. It also provides data connectors that support the Insider risk management solution.

In addition to these native data connectors, Microsoft works with the following partners to provide many more third-party data connectors in the Microsoft Purview portal. Your organization works with these partners to set up their archiving service before creating a corresponding data connector in the Microsoft Purview portal.

- TeleMessage
- 17a-4 LLC
- CellTrust

> **Important**
>
> All partner third-party data connector configurations require a provisioned service principal for the partner application. This requirement isn't applicable for native Microsoft third-party data connector configurations.
>
> To provision a service principal using PowerShell, see Provision a service principal for partner third-party data connectors.

The third-party data connectors described in this article (except for HR data and physical badging data that the Microsoft Purview Insider Risk Management solution uses) import into user mailboxes. The Microsoft Purview solutions that support third-party data apply to the user mailbox where the data is stored.

> **Important**
>
> All Veritas data connectors in Microsoft Purview were retired on June 30, 2024. Non-Veritas data connectors in your organization aren't affected by this change. Contact your Veritas account representative if you have questions about Veritas archiving services.

Microsoft Purview data connectors

The following table lists the native third-party data connectors available in the Microsoft Purview portal. The table also summarizes the compliance solutions that you can apply after you import and archive third-party data in Microsoft 365. For a more detailed description of each compliance solution and how it supports third-party data, see the Overview of compliance solutions that support third-party data section.

Select the link in the **Third-party data** column to go to the step-by-step instructions for creating a connector for that data type.

Expand table

Third-party data	Litigation Hold	eDiscovery	Retention settings	Records management	Communication Compliance	Insider Risk Management
Anthropic Claude						
Bloomberg Message	✔	✔	✔	✔	✔	
ChatGPT Enterprise AI interactions		✔	✔	✔	✔	✔
Facebook	✔	✔	✔	✔	✔	
Generic EHR healthcare						✔

Third-party data	Litigation Hold	eDiscovery	Retention settings	Records management	Communication Compliance	Insider Risk Management
Human resources (HR)						✔
ICE Chat	✔	✔	✔	✔	✔	
Instant Bloomberg	✔	✔	✔	✔	✔	
LinkedIn	✔	✔	✔	✔		
Physical badging						✔
Third-party insider risk detections (preview)						✔
Twitter	✔	✔	✔	✔		

TeleMessage data connectors

> **Important**
>
> All partner third-party data connector configurations require a provisioned service principal for the partner application. This requirement isn't applicable for native Microsoft third-party data connector configurations.
>
> To provision a service principal using PowerShell, see Provision a service principal for partner third-party data connectors.

The following table lists the third-party data connectors available through the partnership with TeleMessage. The table also summarizes the compliance solutions you can apply to third-party data after you import and archive it in Microsoft 365. For more information about each compliance solution and how it supports third-party data, see the Overview of compliance solutions that support third-party data section.

Before you can archive third-party data in Microsoft 365, you need to work with TeleMessage to set up their archiving service for your organization. For more information, select the link in the

Third-party data column to go to the step-by-step instructions for creating a connector for that data type.

You can also use TeleMessage data connectors in GCC environments in the Microsoft 365 US Government cloud. For more information, see the Data connectors in the US Government cloud section in this article.

Expand table

Third-party data	Litigation Hold	eDiscovery	Retention settings	Records management	Communication Compliance	Insider Risk Management
Android	✔	✔	✔	✔	✔	
AT&T Network	✔	✔	✔	✔	✔	
Bell Network	✔	✔	✔	✔	✔	
Enterprise Number	✔	✔	✔	✔	✔	
Hong Kong CSL Network	✔	✔	✔	✔	✔	
O2 Network	✔	✔	✔	✔	✔	
Rogers Network	✔	✔	✔	✔	✔	
Signal	✔	✔	✔	✔	✔	
T-Mobile	✔	✔	✔	✔	✔	
Telegram	✔	✔	✔	✔	✔	
TELUS Network	✔	✔	✔	✔	✔	
Verizon Network	✔	✔	✔	✔	✔	
WeChat	✔	✔	✔	✔	✔	

Third-party data	Litigation Hold	eDiscovery	Retention settings	Records management	Communication Compliance	Insider Risk Management
WhatsApp	✔	✔	✔	✔	✔	

17a-4 data connectors

> **Important**
>
> All partner third-party data connector configurations require a provisioned service principal for the partner application. This requirement isn't applicable for native Microsoft third-party data connector configurations.
>
> To provision a service principal using PowerShell, see Provision a service principal for partner third-party data connectors.

The following table lists the third-party data connectors available through a partnership with 17a-4 LLC. The table also summarizes the compliance solutions you can apply to third-party data after you import and archive it in Microsoft 365. For a detailed description of each compliance solution and how it supports third-party data, see the Overview of compliance solutions that support third-party data section.

Before you can archive third-party data in Microsoft 365, you need to work with 17a-4 LLC to set up their archiving service (called *DataParser*) for your organization. For more information, select the link in the **Third-party data** column to go to the step-by-step instructions for creating a connector for that data type.

You can also use 17a-4 data connectors in GCC environments in the Microsoft 365 US Government cloud. For more information, see the Data connectors in the US Government cloud section in this article.

Expand table

Third-party data	Litigation Hold	eDiscovery	Retention settings	Records management	Communication Compliance	Insider Risk Management
BlackBerry	✔	✔	✔	✔	✔	
Bloomberg	✔	✔	✔	✔	✔	

Third-party data	Litigation Hold	eDiscovery	Retention settings	Records management	Communication Compliance	Insider Risk Management
Cisco Jabber	✔	✔	✔	✔	✔	
Cisco Webex	✔	✔	✔	✔	✔	
FactSet	✔	✔	✔	✔	✔	
Fuze	✔	✔	✔	✔	✔	
FX Connect	✔	✔	✔	✔	✔	
ICE Chat	✔	✔	✔	✔	✔	
InvestEdge	✔	✔	✔	✔	✔	
LivePerson Conversational Cloud	✔	✔	✔	✔	✔	
Quip	✔	✔	✔	✔	✔	
Refinitiv Eikon Messenger	✔	✔	✔	✔	✔	
ServiceNow	✔	✔	✔	✔	✔	
Skype for Business Server	✔	✔	✔	✔	✔	
Slack	✔	✔	✔	✔	✔	
SQL	✔	✔	✔	✔	✔	
Symphony	✔	✔	✔	✔	✔	
Zoom	✔	✔	✔	✔	✔	

CellTrust data connectors

> ⓘ **Important**

All partner third-party data connector configurations require a provisioned service principal for the partner application. This requirement isn't applicable for native Microsoft third-party data connector configurations.

To provision a service principal using PowerShell, see Provision a service principal for partner third-party data connectors.

The following table lists the third-party data connector available through a partnership with CellTrust. The table also summarizes the compliance solutions you can apply to third-party data after you import and archive it in Microsoft 365. For a more detailed description of each compliance solution and how it supports third-party data, see the Overview of compliance solutions that support third-party data section.

Before you can archive third-party data in Microsoft 365, you need to work with CellTrust to set up their archiving service, *CellTrust SL2*, for your organization. For more information, select the link in the **Third-party data** column to go to the step-by-step instructions for creating a CellTrust SL2 connector.

Expand table

Third-party data	Litigation Hold	eDiscovery	Retention settings	Records management	Communication Compliance	Insider Risk Management
CellTrust SL2	✔	✔	✔	✔	✔	

The CellTrust SL2 data connector is also available in GCC environments in the Microsoft 365 US Government cloud. For more information, see the Data connectors in the US Government cloud section in this article.

Provision a service principal for third-party data connectors

For partner third-party data connectors, you must configure a service principal. Before you begin, confirm the Azure CLI is installed on your computer.

Complete the following steps to configure a service principal for your organization:

1. Open PowerShell and verify the Azure CLI installation:

```
PowerShell
```

```
az version
```

2. Sign in with the administrator account for your organization:

```PowerShell
az login
```

3. Run the following cmdlet to create the service principal for your organization. Replace `<PartnerAppId>` with the application ID provided by the partner (TeleMessage, 17a-4 LLC, or CellTrust):

```PowerShell
az ad sp create --id <PartnerAppId>
```

Overview of Microsoft Purview solutions that support third-party data

The following sections describe how Microsoft Purview solutions help you manage imported third-party data, including Communication Compliance, eDiscovery, Insider Risk Management, Litigation Hold, Records management, and Retention settings.

Communication Compliance

Use Microsoft Purview Communication Compliance to examine third-party data to make sure it complies with your organization's data standards. You can detect, capture, and take remediation actions for inappropriate messages in your organization. For example, you can monitor the third-party data that you import for offensive language, sensitive information, and regulatory compliance.

eDiscovery

Microsoft Purview eDiscovery is the primary legal investigation tool in Microsoft 365. Use eDiscovery to add custodians to a case, place a custodian's data on hold, and load a custodian's third-party data into a review set for further analysis such as themes and duplicate detection. After you load third-party data into a review set, you can query and filter it to a narrow result set. Use eDiscovery to manage third-party data that might be relevant to your organization's legal or internal investigations.

You can also use eDiscovery to search for and create query-based holds to preserve third-party data. To search for third-party data imported to user mailboxes, use the `kind:externaldata` property in your search query. To search for a specific type of third-party data, use the `itemclass:ipm.externaldata.<third-party data type>` mailbox property. For more information, see Find content in mailboxes and Use Keyword Query Language to create search queries.

Insider Risk Management

By using signals from third-party data, such as selective HR data, Microsoft Purview Insider Risk Management helps you minimize internal risks by detecting, investigating, and acting on risky activities in your organization. For example, data imported by the HR data connector is used as risk indicators to help detect departing user data theft. The healthcare data connector supports risk indicators for patient data misuse, and the third-party insider risk detections connector lets you extend detections to non-Microsoft workloads.

Litigation Hold

Place a Litigation Hold on a user mailbox to retain third-party data. When you create a hold, specify a hold duration (also called a *time-based hold*) so that deleted and modified third-party data is retained for a specified period and then permanently deleted from the mailbox. Or retain content indefinitely (called an *infinite hold*) or until the Litigation Hold is removed.

Records management

The Microsoft Purview Records Management solution lets you declare third-party data as a record. Users can manually apply a retention label that marks third-party data in their mailbox as a record. Or you can auto-apply retention labels by identifying sensitive information, keywords, or content types in third-party data.

Retention settings

Apply a retention policy to user mailboxes to retain and then delete third-party data (and other mailbox content) after the retention period expires. Use retention policies to delete third-party data of a certain age or use retention labels to trigger a disposition review when the retention period for third-party data expires.

Data connectors in the US Government cloud

Some data connectors are available in the US Government cloud. The following tables show the specific government environments that support third-party data connectors. For more information about US Government clouds, see Microsoft 365 US Government.

Microsoft data connectors in the US Government cloud

The following table lists Microsoft data connectors available in US Government cloud environments.

Expand table

Data connector	GCC	GCC High	DoD
Human resources (HR)	Yes	Yes	Yes

TeleMessage data connectors in the US Government cloud

The following table lists TeleMessage data connectors available in US Government cloud environments.

Expand table

Data connector	GCC	GCC High	DoD
Android Archiver	Yes	No	No
AT&T SMS/MMS Network Archiver	Yes	No	No
Bell SMS/MMS Network Archiver	Yes	No	No
Enterprise Number Archiver	Yes	No	No
O2 SMS and Voice Network Archiver	Yes	No	No
Rogers Network Archiver	Yes	No	No
Signal Archiver	Yes	No	No
Telegram Archiver	Yes	No	No
TELUS SMS Network Archiver	Yes	No	No
Verizon SMS/MMS Network Archiver	Yes	No	No
WeChat Archiver	Yes	No	No
WhatsApp Archiver	Yes	No	No

17a-4 data connectors in the US Government cloud

The following table lists 17a-4 data connectors available in US Government cloud environments.

Expand table

Data connector	GCC	GCC High	DoD
BlackBerry DataParser	Yes	No	No
Bloomberg DataParser	Yes	No	No
Cisco Jabber DataParser	Yes	No	No
Cisco Webex DataParser	Yes	No	No
FactSet DataParser	Yes	No	No
Fuze DataParser	Yes	No	No
FX Connect DataParser	Yes	No	No
ICE DataParser	Yes	No	No
InvestEdge DataParser	Yes	No	No
LivePerson Conversational Cloud DataParser	Yes	No	No
Quip DataParser	Yes	No	No
Refinitiv Eikon Messenger DataParser	Yes	No	No
ServiceNow DataParser	Yes	No	No
Skype for Business Server DataParser	Yes	No	No
Slack DataParser	Yes	No	No
SQL DataParser	Yes	No	No
Symphony DataParser	Yes	No	No
Zoom DataParser	Yes	No	No

CellTrust data connectors in the US Government cloud

The following table lists CellTrust data connector availability in US Government cloud environments.

Expand table

Data connector	GCC	GCC High	DoD
CellTrust SL2	Yes	No	No

ⓘ **Note:** The author created this article with assistance from AI. Learn more

Last updated on 06/04/2026

Set up a connector to manage Anthropic Claude data

Use a connector to manage the activity feed and chat conversations from Anthropic Claude. After you set up and configure the connector, it connects by using the Anthropic Compliance API. Microsoft Purview ingests the data, which begins showing up in Data Security Posture Management reports after about one day.

Before you begin

Before you set up the Anthropic Claude data connector, make sure the following prerequisites are in place.

Subscription

To learn more about Microsoft 365 tenant-level services licensing requirements, see Microsoft 365 Security.

Anthropic Enterprise plan

Your organization must have an **Enterprise** plan from Anthropic. The primary owner must enable the Compliance API in **Organization Settings** > **Data and Privacy**. On this page, you can also retrieve your Compliance Access Key by going to **Compliance access keys** > **Create key**. You need this key in the next section where you set up the connector.

For more information about the Anthropic Compliance API, see Access the Compliance API.

Permissions

Assign the Data Connector Admin role to the user who creates the Anthropic Claude data connector in the Microsoft Purview portal. This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions, see Permissions in the Microsoft Purview portal and Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview.

Set up the Anthropic Claude data connector

To set up the Anthropic Claude data connector, perform the following steps.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **All connectors**.
4. From the list, select **Anthropic Claude**.
5. On the **Anthropic Claude** product description page, select **Add connector**.
6. On the **Terms of service** page, review and select **Accept**.
7. On the **Connector name** page, enter a descriptive name for the connector and select **Next**.
8. On the **Authentication** page, enter the following information:
    - **API Key**: Enter the Compliance API key generated from your Anthropic Organization Settings. Select **Validate connection**.
    - **Organization IDs**: Select your Anthropic organization identifier.
    - **API Base URL**: Confirm the Anthropic API base URL (the default is `https://api.anthropic.com`).
9. Select **Next**.
10. On the **Data scope** page, select the data types to import:
    - **Activity Feed** — Authentication events, chat interactions, file uploads, admin actions, and compliance audit events across Claude.ai, Console, and API. *(Scope: read:compliance_activities)*
    - **Chat Conversations** — Full message history between users and Claude, including user prompts, AI responses, and metadata for DSPM analysis. *(Scope: read:compliance_user_data)*
    - Select a **Sync frequency** to determine how often the connector pulls new data from the Anthropic Compliance API.
11. Select **Next**.
12. On the **Review your connector** page, verify your connector configuration and select **Finish**.

After you create the connector, you can see the following data:

- **Ongoing sync**: View new activities and chats that the connector imports based on your selected frequency.
- **DSPM**: View Claude interaction data in the DSPM Dashboard and Activity Explorer.
- **Monitoring**: View connector health and sync status in **Data connectors** > **My connectors**.

ⓘ **Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 05/28/2026

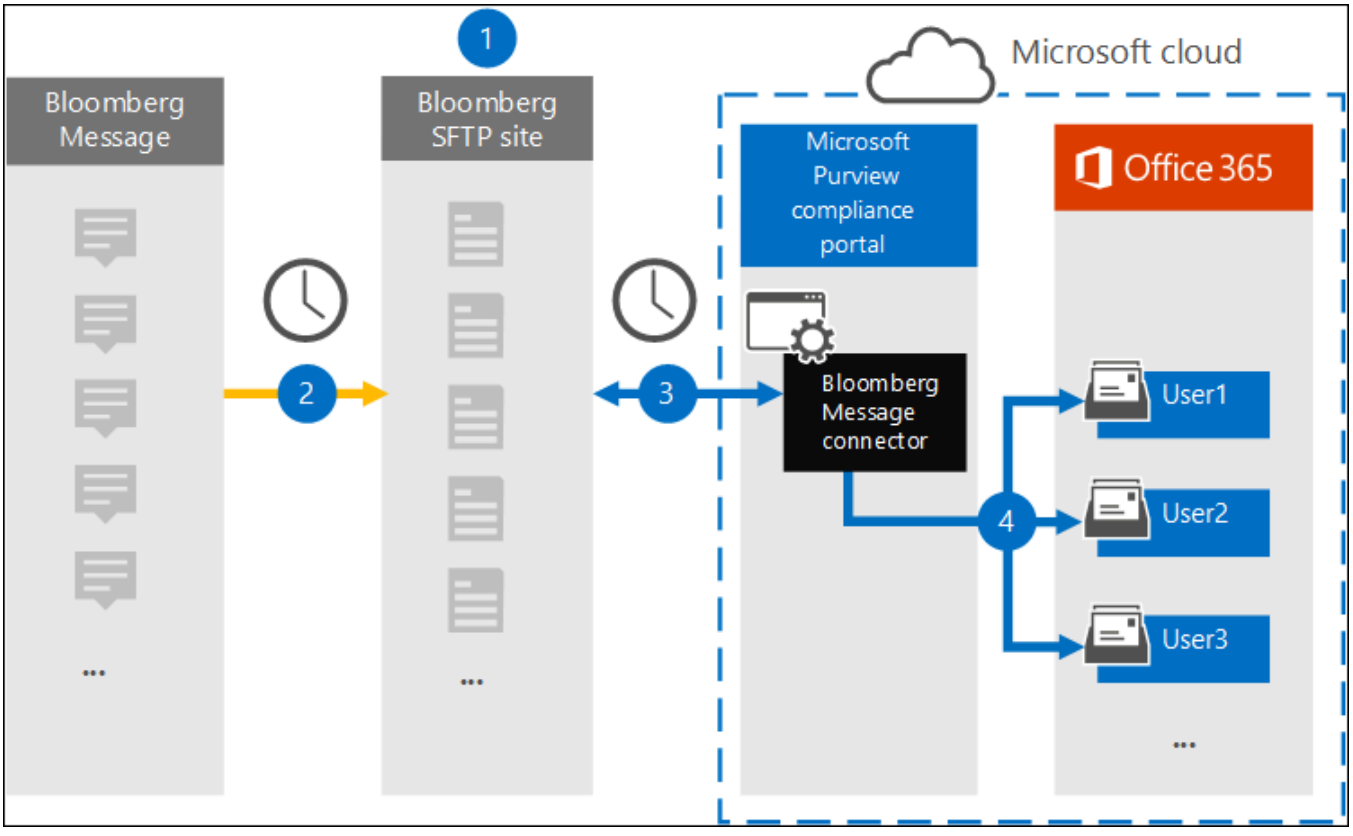
Set up a connector to archive Bloomberg Message data

Use a data connector to import and archive financial services email data from the Bloomberg Message collaboration tool. After you set up and configure a connector, it connects to your organization's Bloomberg secure FTP (SFTP) site once every day, and imports email items to mailboxes in Microsoft 365.

After storing Bloomberg Message data in user mailboxes, you can apply Microsoft Purview features such as Litigation Hold, eDiscovery, In-Place Archiving, Auditing, Communication Compliance, and Microsoft 365 retention policies to Bloomberg Message data. For example, you can search Bloomberg Message emails by using eDiscovery or associate the mailbox that contains the Bloomberg Message data with a custodian in an eDiscovery (Premium) case. Using a Bloomberg Message connector to import and archive data in Microsoft 365 can help your organization stay compliant with government and regulatory policies.

Overview of archiving Bloomberg Message data

The following overview explains the process of using a connector to archive Bloomberg Message data in Microsoft 365.



1. Your organization works with Bloomberg to set up a Bloomberg SFTP site. You also work with Bloomberg to configure Bloomberg Message to copy email messages to the

Bloomberg SFTP site.

2. Every 24 hours, Bloomberg Message copies email messages to the Bloomberg SFTP site.

3. The Bloomberg Message connector that you create in the Microsoft Purview portal connects to the Bloomberg SFTP site every day and transfers the email messages from the previous 24 hours to a secure Azure Storage area in the Microsoft Cloud.

4. The connector imports the email message items to the mailbox of a specific user. A new folder named BloombergMessage is created in the specific user's mailbox and the items are imported to it.

   The connector uses the value of the CorporateEmailAddress property. Every email message contains this property, which is populated with the email address of every participant of the email message. In addition to automatic user mapping by using the value of the *CorporateEmailAddress* property, you can also define a custom mapping by uploading a CSV mapping file. This mapping file contains a Bloomberg UUID and the corresponding Microsoft 365 mailbox address for each user in your organization. If you enable automatic user mapping and provide a custom mapping, for every email item the connector first looks at the custom-mapping file. If it doesn't find a valid Microsoft 365 user that corresponds to a user's Bloomberg UUID, the connector uses the *CorporateEmailAddress* property of the email item. If the connector doesn't find a valid Microsoft 365 user in either the custom-mapping file or the *CorporateEmailAddress* property of the email item, the item won't be imported.

Before you begin

Some of the implementation steps for archiving Bloomberg Message data are external to Microsoft 365. You must complete these steps before you create the connector in the Microsoft Purview portal.

- To set up a Bloomberg Message connector, you need keys and key passphrases for Pretty Good Privacy (PGP) and Secure Shell (SSH). Use these keys to configure the Bloomberg SFTP site. The connector uses these keys to connect to the Bloomberg SFTP site and import data to Microsoft 365. Use the PGP key to configure the encryption of data that's transferred from the Bloomberg SFTP site to Microsoft 365. Use the SSH key to configure secure shell to enable a secure remote authentication when the connector connects to the Bloomberg SFTP site.

  When setting up a connector, you can use public keys and key passphrases provided by Microsoft or use your own private keys and passphrases. We recommend that you use the public keys provided by Microsoft. However, if your organization already configured a

Bloomberg SFTP site by using private keys, you can create a connector by using these same private keys.

- Subscribe to Bloomberg Anywhere. This subscription is required so that you can sign in to Bloomberg Anywhere to access the Bloomberg SFTP site that you set up and configure.

- Set up a Bloomberg SFTP (Secure file transfer protocol) site. After working with Bloomberg to set up the SFTP site, data from Bloomberg Message is uploaded to the SFTP site every day. The connector you create in Step 2 connects to this SFTP site and transfers the email data to Microsoft 365 mailboxes. SFTP also encrypts the Bloomberg Message data that is sent to mailboxes during the transfer process.

  For information about Bloomberg SFTP (also called *BB-SFTP*):

  - See the "SFTP Connectivity Standards" document at Bloomberg Support.
  - Contact Bloomberg customer support.

- After you work with Bloomberg to set up an SFTP site, Bloomberg provides some information to you after you respond to the Bloomberg implementation email message. Save a copy of the following information. You use it to set up a connector in Step 3.

  - Firm code, which is an ID for your organization and is used to sign in to the Bloomberg SFTP site.
  - Password for your Bloomberg SFTP site
  - URL for Bloomberg SFTP site (for example, sftp.bloomberg.com). In addition, Bloomberg might also provide a corresponding IP address for the Bloomberg SFTP site, which you can also use to set up the connector.
  - Port number for Bloomberg SFTP site

- The Bloomberg Message connector can import a total of 200,000 items in a single day. If there are more than 200,000 items on the SFTP site, none of those items are imported to Microsoft 365.

- The user who creates a Bloomberg Message connector in Step 3 (and who downloads the public keys and IP address in Step 1) must be assigned the Data Connector Admin role. This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For

instructions, see:- Permissions in the Microsoft Purview portal - Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview compliance

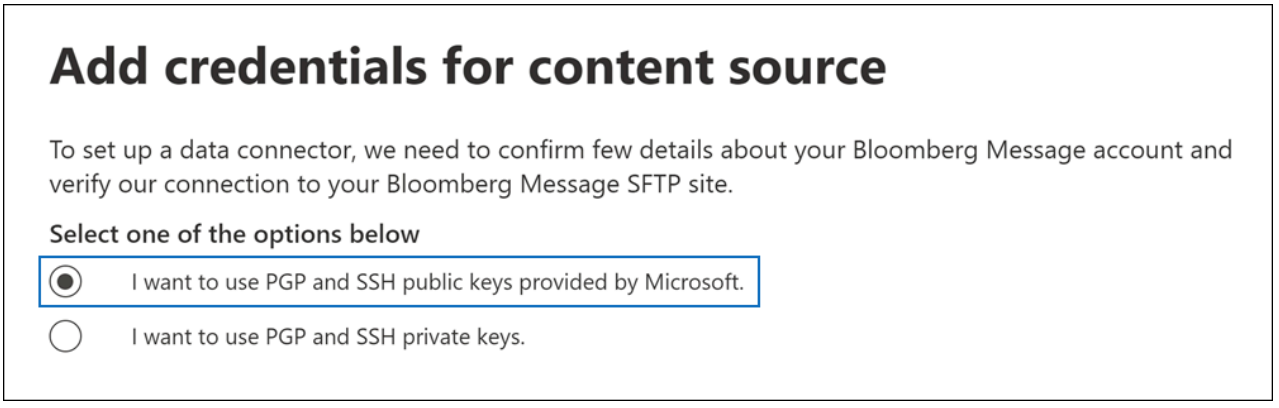
Set up a connector by using public keys

The following steps show you how to set up a Bloomberg Message connector by using the public keys for Pretty Good Privacy (PGP) and Secure Shell (SSH).

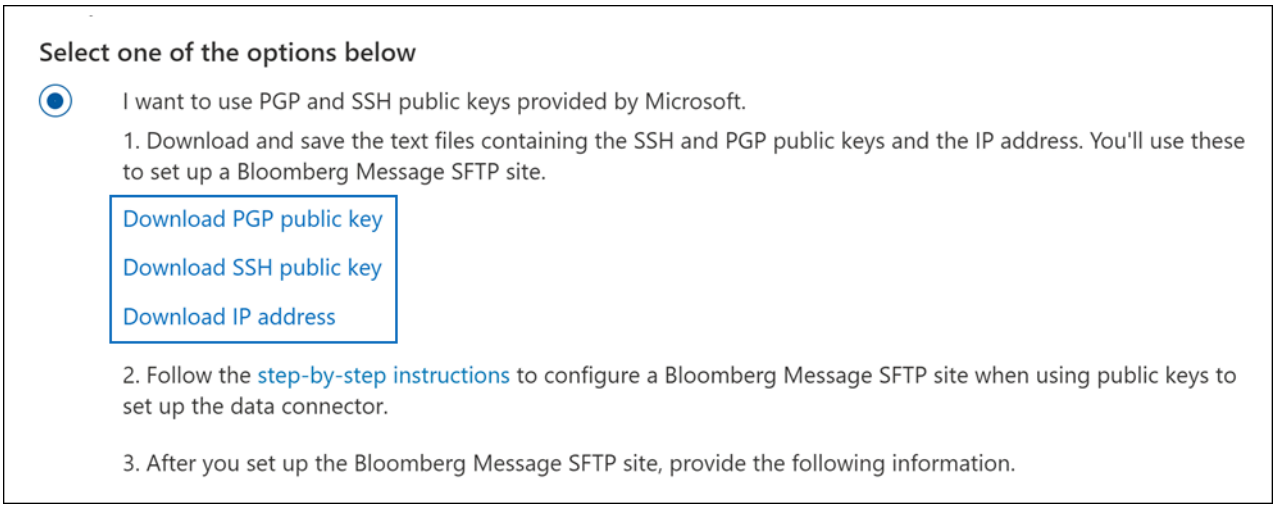
Step 1: Obtain PGP and SSH public keys

First, get a copy of the PGP and SSH public keys. You use these keys in Step 2 to set up the Bloomberg SFTP site so it allows the connector (that you create in Step 3) to connect to the SFTP site and transfer the Bloomberg Message email data to Microsoft 365 mailboxes. You also get an IP address in this step, which you use when configuring the Bloomberg SFTP site.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. From the list, choose **Bloomberg Message**, select **View**.
5. On the **Bloomberg Message** product description page, select **Add connector**.
6. On the **Terms of service** page, select **Accept**.
7. On the **Add credentials for content source** page, select **I want to use PGP and SSH public keys provided by Microsoft**.



8. Under step 1, select the **Download SSH key**, **Download PGP key**, and **Download IP address** links to save a copy of each file to your local computer.



These files contain the following items that you use to configure the Bloomberg SFTP site in Step 2:

- PGP public key: Use this key to configure the encryption of data that's transferred from the Bloomberg SFTP site to Microsoft 365.
- SSH public key: Use this key to configure secure shell to enable a secure remote authentication when the connector connects to the Bloomberg SFTP site.
- IP address: Configure the Bloomberg SFTP site to accept connection requests from this IP address. The same IP address is used by the Bloomberg Message connector to connect to the SFTP site and transfer Bloomberg Message data to Microsoft 365.

9. Select **Cancel** to close the workflow. You come back to this workflow in Step 3 to create the connector.

Step 2: Configure the Bloomberg SFTP site

> **Note**
>
> If your organization previously set up a Bloomberg SFTP site to archive Instant Bloomberg data by using public PGP and SSH keys, you don't need to set up another one. You can specify the same SFTP site when you create the connector in Step 3.

Use the PGP and SSH public keys and the IP address that you got in Step 1 to configure PGP encryption and SSH authentication for the Bloomberg SFTP site. This configuration lets the Bloomberg Message connector that you create in Step 3 connect to the Bloomberg SFTP site and transfer Bloomberg Message data to Microsoft 365. You need to work with Bloomberg customer support to set up your Bloomberg SFTP site. Contact Bloomberg customer support for assistance.

> **Important**
>
> Bloomberg recommends that you attach the three files that you download in Step 1 to an email message and send it to their customer support team when working with them to set up your Bloomberg SFTP site.

Step 3: Create a Bloomberg Message connector

In this step, you create a Bloomberg Message connector in the Microsoft Purview portal. The connector uses the information you provide to connect to the Bloomberg SFTP site and transfer email messages to the corresponding user mailboxes in Microsoft 365.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. Choose **Bloomberg Message** from the list.
5. On the **Terms of service** page, select **Accept**.
6. On the **Add credentials for content source** page, select **I want to use PGP and SSH public keys provided by Microsoft**.
7. Under Step 3, enter the required information in the following boxes and then select **Validate connection**.
   - **Name:** The name for the connector. It must be unique in your organization.
   - **Firm code:** The ID for your organization that is used as the username for the Bloomberg SFTP site.
   - **Password:** The password for your organization's Bloomberg SFTP site.
   - **SFTP URL:** The URL for the Bloomberg SFTP site (for example, `sftp.bloomberg.com`). You can also use an IP address for this value.
   - **SFTP port:** The port number for the Bloomberg SFTP site. The connector uses this port to connect to the SFTP site.
8. After the connection is successfully validated, select **Next**.
9. On the **Define user** page, specify the users to import data for.
   - **All users in your organization**. Select this option to import data for all users.
   - **Only users on Litigation hold**. Select this option to import data only for users whose mailboxes are placed on Litigation hold. This option imports data to user mailboxes

that have the LitigationHoldEnabled property set to True. For more information, see Create a Litigation hold.

10. On the **Map Bloomberg Message users to Microsoft 365 users** page, enable automatic user mapping and provide custom user mapping as required.

    > **Note**
    >
    > The connector imports message items to the mailbox of a specific user. A new folder named **BloombergMessage** is created in the specific user's mailbox and the items are imported to it. The connector uses the value of the *CorporateEmailAddress* property. Every chat message contains this property, and the property is populated with the email address of every participant of the chat message. In addition to automatic user mapping using the value of the *CorporateEmailAddress* property, you can also define custom mapping by uploading a CSV mapping file. The mapping file should contain the Bloomberg UUID and corresponding Microsoft 365 mailbox address for each user. If you enable automatic user mapping and provide a custom mapping, for every message item the connector will first look at custom mapping file. If it doesn't find a valid Microsoft 365 user that corresponds to a user's Bloomberg UUID, the connector will use the *CorporateEmailAddress* property of the chat item. If the connector doesn't find a valid Microsoft 365 user in either the custom mapping file or the *CorporateEmailAddress* property of the message item, the item won't be imported.

11. Select **Next**, review your settings, and then select **Finish** to create the connector.

Set up a connector by using private keys

The steps in this section show you how to set up a Bloomberg Message connector by using PGP and SSH private keys. Use this connector setup option if your organization already configured a Bloomberg SFTP site by using private keys.

Step 1: Obtain an IP address to configure the Bloomberg SFTP site

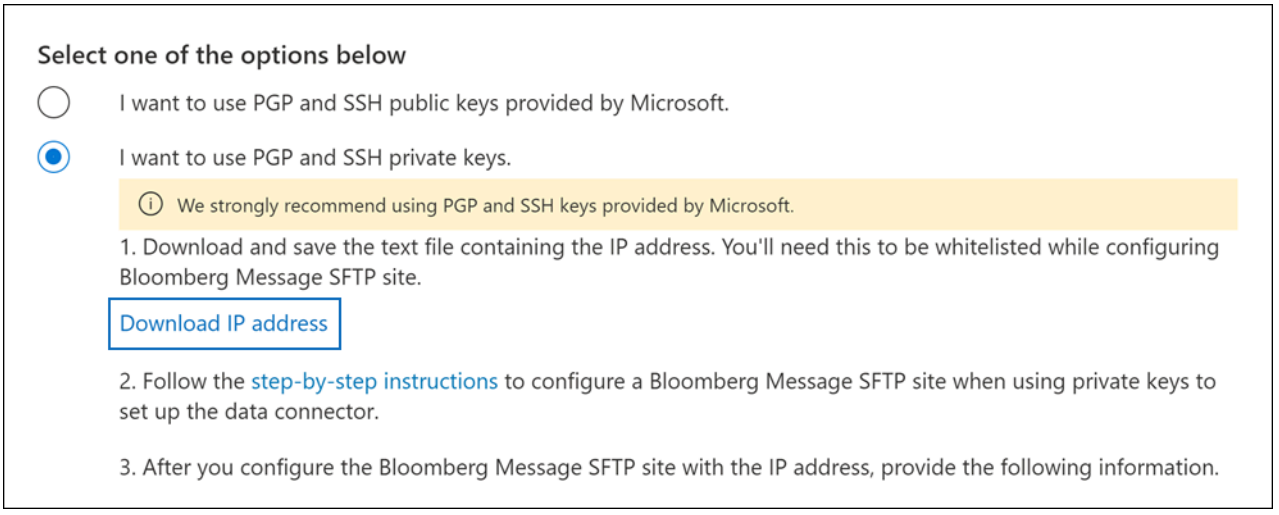
> **Note**

If your organization previously configured a Bloomberg SFTP site to archive Instant Bloomberg data by using PGP and SSH private keys, you don't need to configure another one. You can specify the same SFTP site when you create the connector in Step 2.

If your organization used PGP and SSH private keys to set up a Bloomberg SFTP site, you need to obtain an IP address and provide it to Bloomberg customer support. The Bloomberg SFTP site must be configured to accept connection requests from this IP address. The same IP address is used by the Bloomberg Message connector to connect to the SFTP site and transfer Bloomberg Message data to Microsoft 365.

To obtain the IP address:

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. Choose **Bloomberg Message** from the list.
5. On the **Terms of service** page, select **Accept**.
6. On the **Add credentials for content source** page, select **I want to use PGP and SSH private keys**.
7. Under step 1, select **Download IP address** to save a copy of the IP address file to your local computer.



8. Select **Cancel** to close the workflow. You come back to this workflow in Step 4 to create the connector.

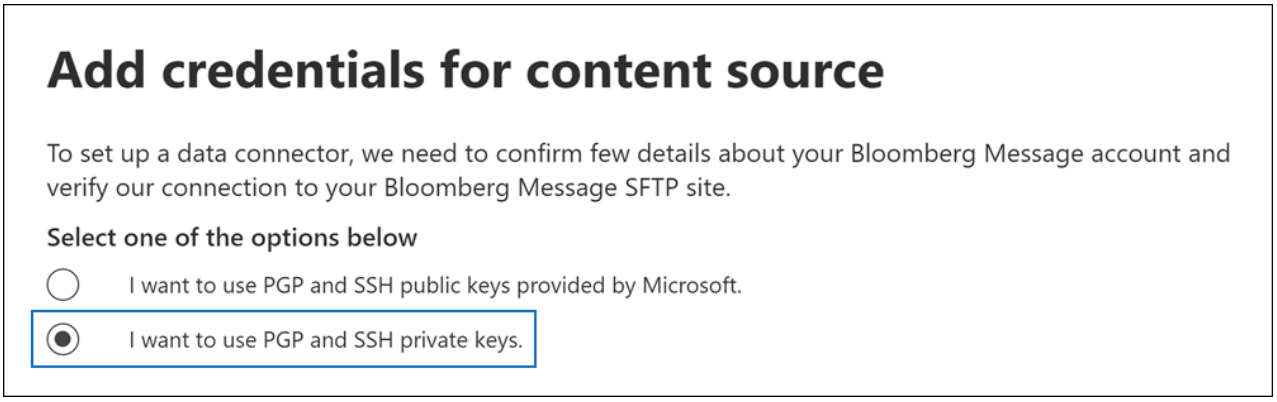
You need to work with Bloomberg customer support to configure your Bloomberg SFTP site to accept connection requests from this IP address. Contact Bloomberg customer support for

assistance.

Step 2: Create a Bloomberg Message connector

After configuring your Bloomberg SFTP site, create a Bloomberg Message connector in the Microsoft Purview portal. The connector uses the information you provide to connect to the Bloomberg SFTP site and transfer email messages to the corresponding user mailboxes in Microsoft 365. To complete this step, have copies of the same private keys and key passphrases that you used to set up your Bloomberg SFTP site.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. Choose **Bloomberg Message** from the list.
5. On the **Terms of service** page, select **Accept**.
6. On the **Add credentials for content source** page, select **I want to use PGP and SSH private keys**.



7. Under Step 3, enter the required information in the following boxes and then select **Validate connection**.

   - **Name:** The name for the connector. It must be unique in your organization.
   - **Firm code:** The ID for your organization that is used as the username for the Bloomberg SFTP site.
   - **Password:** The password for your organization's Bloomberg SFTP site.
   - **SFTP URL:** The URL for the Bloomberg SFTP site (for example, `sftp.bloomberg.com`). You can also use an IP address for this value.
   - **SFTP port:** The port number for the Bloomberg SFTP site. The connector uses this port to connect to the SFTP site.

- **PGP private key:** The PGP private key for the Bloomberg SFTP site. Include the entire private key value, including the beginning and ending lines of the key block.
- **PGP key passphrase:** The passphrase for the PGP private key.
- **SSH private key:** The SSH private key for the Bloomberg SFTP site. Include the entire private key value, including the beginning and ending lines of the key block.
- **SSH key passphrase:** The passphrase for the SSH private key.

8. After the connection is successfully validated, select **Next**.

9. On the **Define user** page, specify the users to import data for.

   - **All users in your organization**. Select this option to import data for all users.
   - **Only users on Litigation hold**. Select this option to import data only for users whose mailboxes are placed on Litigation hold. This option imports data to user mailboxes that have the LitigationHoldEnabled property set to True. For more information, see Create a Litigation hold.

10. On the **Map Bloomberg Message users to Microsoft 365 users** page, enable automatic user mapping and provide custom user mapping as required.

    > ⓘ Note
    >
    > The connector imports message items to the mailbox of a specific user. A new folder named **BloombergMessage** is created in the specific user's mailbox and the items are imported to it. The connector uses the value of the *CorporateEmailAddress* property. Every chat message contains this property, and the property is populated with the email address of every participant of the chat message. In addition to automatic user mapping using the value of the *CorporateEmailAddress* property, you can also define custom mapping by uploading a CSV mapping file. The mapping file should contain the Bloomberg UUID and corresponding Microsoft 365 mailbox address for each user. If you enable automatic user mapping and provide a custom mapping, for every message item the connector will first look at custom mapping file. If it doesn't find a valid Microsoft 365 user that corresponds to a user's Bloomberg UUID, the connector will use the *CorporateEmailAddress* property of the chat item. If the connector doesn't find a valid Microsoft 365 user in either the custom mapping file or the *CorporateEmailAddress* property of the message item, the item won't be imported.

11. Select **Next**, review your settings, and then select **Finish** to create the connector.

Limits

- Microsoft 365 doesn't support threading for Bloomberg Message email. The system imports individual messages sent to a person, but it doesn't show them in a threaded conversation.

Last updated on 03/30/2026

Set up a connector to archive Facebook data (preview)

Use a connector to import and archive data from Facebook Business pages to Microsoft 365. After you set up and configure the connector, it connects to the Facebook Business page on a scheduled basis, converts the content of Facebook items to an email message format, and then imports those items to a mailbox in Microsoft 365.

After the Facebook data is imported, you can apply Microsoft Purview features such as Litigation Hold, eDiscovery, In-Place Archiving, Auditing, Communication Compliance, and Microsoft 365 retention policies to the Facebook data. For example, when you place a mailbox on Litigation Hold or assign it to a retention policy, the Facebook data is preserved. You can search third-party data by using eDiscovery or associate the mailbox where the Facebook data is stored with a custodian in a Microsoft Purview eDiscovery (Premium) case. Using a connector to import and archive Facebook data in Microsoft 365 can help your organization stay compliant with government and regulatory policies.

Before you begin for setting up a connector for Facebook Business pages

Complete the following prerequisites before setting up and configuring a connector in the Microsoft Purview portal to import and archive data from your organization's Facebook Business pages.

- You need a Facebook account for your organization's business pages. You need to sign in to this account when setting up the connector. Currently, you can only archive data from Facebook Business pages; you can't archive data from individual Facebook profiles.
- Your organization must have a valid Azure subscription. If you don't have an existing Azure subscription, you can sign up for one of these options:
  - Sign up for a free one year Azure subscription
  - Sign up for a Pay-As-You-Go Azure subscription
- The connector for Facebook Business pages can import a total of 200,000 items in a single day. If there are more than 200,000 Facebook Business items in a day, none of those items

are imported to Microsoft 365.

- The user who sets up the custom connector in the Microsoft Purview portal (in Step 5) must be assigned the Data Connector Admin role. This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions, see Permissions in the Microsoft Purview portal and Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview.

Step 1: Create an app in Microsoft Entra ID

First, register a new app in Microsoft Entra ID. This app corresponds to the Azure web app resource used by the Facebook connector, which you deploy in Step 2 and configure in Step 4.

For more information, see Create an app in Microsoft Entra ID.

When you complete this step, save the following information to a text file. You'll use these values in later steps in the deployment process.

- Microsoft Entra application ID
- Microsoft Entra application secret
- Tenant Id

Step 2: Deploy the connector web service from GitHub to your Azure account

Next, deploy the source code for the Facebook Business pages connector app. This app uses the Facebook API to connect to your Facebook account and extract data so you can import it to Microsoft 365. The Facebook connector that you deploy for your organization uploads the items from your Facebook Business pages to the Azure Storage location that you create in this step. After you create a Facebook business pages connector in the Microsoft Purview portal (in Step 5), the Import service copies the Facebook business pages data from the Azure Storage location to a mailbox in your Microsoft 365 organization. As explained in the Before you begin section, you must have a valid Azure subscription to create an Azure Storage account.

For more information, see Deploy the connector web service from GitHub to your Azure account.

When you complete this step, provide the following information:

- APISecretKey: Create this secret during the completion of this step. It's used in Step 5.
- TenantId: The tenant ID of your Microsoft 365 organization that you copied after creating the Facebook connector app in Microsoft Entra ID in Step 1.

After you complete this step, copy the Azure app service URL (for example, `https://fbconnector.azurewebsites.net`). You need to use this URL to complete Step 3, Step 4, and Step 5.

Step 3: Register the web app on Facebook

Create and configure a new app on Facebook. The Facebook business pages connector that you create in Step 5 uses the Facebook web app to interact with the Facebook API to obtain data from your organization's Facebook Business pages.

For more information, see Register the Facebook app.

When you complete this step, save the following information to a text file. You use these values to configure the Facebook connector app in Step 4.

- Facebook application ID
- Facebook application secret
- Facebook webhooks verify token

Step 4: Configure the Facebook connector app

Add configuration settings to the Facebook connector app that you deployed as an Azure web app resource in Step 2. Go to the home page of the Facebook connector Azure web app and configure the connector app.

For more information, see Configure the Facebook connector app.

When you complete this step, provide the following information (that you copied to a text file after completing the previous steps):

- Facebook application ID (obtained in Step 3)
- Facebook application secret (obtained in Step 3)
- Facebook webhooks verify token (obtained in Step 3)
- Microsoft Entra application ID (the Microsoft Entra application ID obtained in Step 1)
- Microsoft Entra application secret (the Microsoft Entra application secret obtained in Step 1)

Step 5: Set up a Facebook Business pages connector in the Microsoft Purview portal

In this step, you set up the connector in the Microsoft Purview portal that imports data from your Facebook Business pages to a specified mailbox in Microsoft 365. When you finish this step, the Microsoft 365 Import service starts importing data from your Facebook Business pages to Microsoft 365.

For more information, see Step 5: Set up a Facebook connector in the Microsoft Purview portal.

As you complete this step, you provide the following information, which you copy to a text file after completing the steps:

- Microsoft Entra application ID (obtained in Step 1)
- Azure app service URL (obtained in Step 2; for example, `https://fbconnector.azurewebsites.net`)
- APISecretKey (created in Step 2)

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/19/2026

Set up a connector to import healthcare EHR audit data (preview)

You can set up a data connector to import auditing data for user activity in your organization's Electronic Healthcare Records (EHR) system. Auditing data from your healthcare EHR system includes data for events related to accessing a patient's health records. Microsoft Purview Insider Risk Management solution uses healthcare EHR auditing data to help protect your organization from unauthorized access to patient information.

Setting up a Healthcare connector consists of the following tasks:

- Creating an app in Microsoft Entra ID to access an API endpoint that accepts a tab-separated text file containing healthcare EHR auditing data.
- Creating a text file with all the required fields as defined in the connector schema.
- Creating a Healthcare connector instance in the Microsoft Purview portal.
- Running a script to push healthcare EHR auditing data to the API endpoint.
- Optionally, scheduling the script to run automatically to import the auditing data.

Before you begin

- Assign the Data Connector Admin role to the user who creates the Healthcare connector in Step 3. This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions, see:- Permissions in the Microsoft Purview portal - Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview compliance

- Determine how to retrieve or export the data from your organization's healthcare EHR system (on a daily basis) and create a text file that's described in Step 2. The script that you run in Step 4 pushes the data in the text file to the API endpoint.

- Understand that the sample script you run in Step 4 pushes the healthcare EHR auditing data from a text file to the connector API so that the insider risk management solution can use it. This sample script isn't supported under any Microsoft standard support program or service. The sample script is provided AS IS without warranty of any kind. Microsoft further disclaims all implied warranties including, without limitation, any implied warranties of merchantability or of fitness for a particular purpose. You assume all risk arising from the use or performance of the sample script and documentation. In no event shall Microsoft, its authors, or anyone else involved in the creation, production, or delivery

of the scripts be liable for any damages whatsoever (including, without limitation, damages for loss of business profits, business interruption, loss of business information, or other pecuniary loss) arising out of the use of or inability to use the sample scripts or documentation, even if Microsoft has been advised of the possibility of such damages.

Step 1: Create an app in Microsoft Entra ID

First, create and register a new app in Microsoft Entra ID. The app corresponds to the Healthcare connector that you create in Step 3. When you create the app, Microsoft Entra ID can authenticate the push request for a text file containing healthcare EHR auditing data. During the creation of this Microsoft Entra app, save the following information. You use these values in later steps.

- Microsoft Entra application ID (also called the *app Id* or *client Id*)
- Microsoft Entra application secret (also called the *client secret*)
- Tenant Id (also called the *directory Id*)

For step-by-step instructions for creating an app in Microsoft Entra ID, see Register an application with the Microsoft identity platform.

Step 2: Prepare a text file with healthcare EHR auditing data

Next, create a text file that contains information about employees' access to patient health records in your organization's healthcare EHR system. As previously explained, determine how to generate this text file from your healthcare EHR system. The Healthcare connector workflow requires a text file with tab-separated values to map that data in the text file with required connector schema. The supported file formats are comma (.csv), pipe (.psv), or tab (.tsv) separated text files.

> **Note**
>
> The maximum size of the text file that contains the auditing data is 3 GB. The maximum number of rows is 5 million. Also, be sure to only include the relevant auditing data from your healthcare EHR system.

The following table lists the fields that are required to enable insider risk management scenarios. A subset of these fields is mandatory. These fields are highlighted with an asterisk (*). If the text file is missing any of the mandatory fields, the file isn't validated and the data in the file isn't imported.

Expand table

Field	Category
Creation Time* Event Name* Workstation Id Event Section Event Category	These fields are used to identify access activity events in your healthcare EHR system.
Patient Reg Id Patient First Name* Patient Middle Name Patient Last Name* Patient Address Line 1* Patient Address Line 2 Patient City* Patient Zip Code* Patient State Patient Country Patient Department	These fields are used to identify patient profile information.
Restricted Access Reason* Restricted Access Comment	These fields are used to identify access to restricted records.
Email Address (UPN) or SamAccountName* Employee User Name Employee Id Employee Last Name [1] Employee First Name [1]	These fields are used to identify employee profile information for address and name matching required to determine access to Family/Neighbor/Employee records.

> **Note**
>
> [1]This field might not be available by default in your healthcare EHR system. You need to configure the export to ensure the text file contains this field.

Step 3: Create the Healthcare connector

In this step, you create a Healthcare connector in the Microsoft Purview portal. After you run the script in Step 4, the connector processes the text file you created in Step 2 and pushes it to the API endpoint you set up in Step 1. Be sure to copy the JobId generated when you create the connector. You need the JobId when you run the script.

1. Sign in to the Microsoft Purview portal.

2. Select **Settings** > **Data connectors**.

3. Select **My connectors**, then select **Add connector**.

4. From the list, choose **Healthcare (preview)**.

5. Accept the terms of service.

6. On the **Authentication credentials** page, enter the following information and select **Next**:
   a. Enter or paste the Microsoft Entra application ID for the Azure app that you created in Step 1.
   b. Enter a name for the healthcare connector.

7. On the **File mapping method** page, select one of the following options and select **Next**:

   - **Upload a sample file**. Select **Upload sample file** to upload the file that you prepared in Step 2. With this option, you can quickly select column names in your text file from a drop-down list to map the columns to the required schema for the healthcare connector.

   Or

   - **Manually provide the mapping details**. Type the name of the columns in your text file to map the columns to the required schema for the healthcare connector.

8. On the **File mapping details** page, complete one of the following steps, depending on whether you uploaded a sample file in the previous step:

   - Use the dropdown lists to map the columns from the sample file to each required field for the healthcare connector.

   Or

   - For each field, type the column name from the file that you prepared in Step 2 that corresponds to the field for the healthcare connector.

9. On the **Review** page, review your settings and select **Finish** to create the connector.

   A status page confirms the connector was created. This page contains two important items that you need to complete the next step to run the sample script to upload your healthcare EHR auditing data.

   - **Job ID.** You need this job ID to run the script in the next step. You can copy it from this page or from the connector flyout page.
   - **Link to sample script.** Select the **here** link to go to the GitHub site to access the sample script. The link opens a new window. Keep this window open so that you can

copy the script in Step 4. Alternatively, you can bookmark the destination or copy the URL so you can access it again when you run the script. This link is also available on the connector flyout page.

10. Select **Done**.

    The new connector appears in the list on the **Connectors** tab.

11. Select the Healthcare connector that you just created to display the flyout page, which contains properties and other information about the connector.

If you haven't already done so, copy the values for the **Azure App ID** and **Connector job ID**. You need these values to run the script in the next step. You can also download the script from the flyout page or using the link in the next step.

Select **Edit** to change the Azure App ID or the column header names that you defined on the **File mapping** page.

Step 4: Run the sample script to upload your healthcare EHR auditing data

The last step in setting up a Healthcare connector is to run a sample script that uploads the healthcare EHR auditing data in the text file (that you created in Step 1) to the Microsoft cloud. Specifically, the script uploads the data to the Healthcare connector. After you run the script, the Healthcare connector that you created in Step 3 imports the healthcare EHR auditing data to your Microsoft 365 organization where other compliance tools, such as the Insider Risk Management solution, can access it. After you run the script, consider scheduling a task to run it automatically on a daily basis so the most current employee termination data is uploaded to the Microsoft cloud. For more information, see (Optional) Step 6: Schedule the script to run automatically.

> **Note**
>
> As previously stated, the maximum size of the text file that contains the auditing data is 3 GB. The maximum number of rows is 5 million. The script that you run in this step takes about 30 to 40 minutes to import the auditing data from large text files. Additionally, the script divides large text files into smaller blocks of 100K rows, and then imports those blocks sequentially.

1. Go to the window that you left open from the previous step to access the GitHub site with the sample script. Alternatively, open the bookmarked site or use the URL that you

copied. You can also access the script here.

2. Select the **Raw** button to display the script in text view.
3. Copy all the lines in the sample script and save them to a text file.
4. Modify the sample script for your organization, if necessary.
5. Save the text file as a Windows PowerShell script file by using a filename suffix of `.ps1`; for example, `HealthcareConnector.ps1`.
6. Open a Command Prompt on your local computer, and go to the directory where you saved the script.
7. Run the following command to upload the healthcare audit data in the text file to Microsoft cloud; for example:

```
PowerShell

.\HealthcareConnector.ps1 -tenantId <tenantId> -appId <appId>  -appSecret
<appSecret>  -jobId <jobId>  -filePath '<filePath>'
```

The following table describes the parameters to use with this script and their required values. Use the information you obtained in the previous steps for these values.

Expand table

Parameter	Description
tenantId	This is the Id for your Microsoft 365 organization that you obtained in Step 1. You can also obtain the tenant Id for your organization on the **Overview** blade in the Microsoft Entra admin center. This value identifies your organization.
appId	This is the Microsoft Entra application Id for the app that you created in Microsoft Entra ID in Step 1. Microsoft Entra ID uses this value for authentication when the script attempts to access your Microsoft 365 organization.
appSecret	This is the Microsoft Entra application secret for the app that you created in Microsoft Entra ID in Step 1. This value is also used for authentication.
jobId	This is the job ID for the Healthcare connector that you created in Step 3. This value associates the healthcare EHR auditing data that you upload to the Microsoft cloud with the Healthcare connector.
filePath	This is the file path for the text file (stored on the same system as the script) that you created in Step 2. Try to avoid spaces in the file path; otherwise use single quotation marks.

Here's an example of the syntax for the Healthcare connector script using actual values for each parameter:

```PowerShell
.\HealthcareConnector.ps1 -tenantId aaaabbbb-0000-cccc-1111-dddd2222eeee -appId 00001111-aaaa-2222-bbbb-3333cccc4444 -appSecret Aa1Bb~2Cc3.-Dd4Ee5Ff6Gg7Hh8Ii9_Jj0Kk1Ll2 -jobId 00001111-aaaa-2222-bbbb-3333cccc4444 -filePath 'C:\Users\contosoadmin\Desktop\Data\healthcare_audit_records.csv'
```

If the upload is successful, the script displays the **Upload Successful** message.

> **Note**
>
> If you have problems running the previous command because of execution policies, see About Execution Policies and Set-ExecutionPolicy for guidance about setting execution policies.

Step 5: Monitor the Healthcare connector

After you create the Healthcare connector and push your EHR auditing data, you can view the connector and upload status in the Microsoft Purview portal. If you schedule the script to run automatically on a regular basis, you can also view the current status after the last time the script runs.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select the Healthcare connector that you created to display the flyout page. This page contains the properties and information about the connector.
4. Under **Last import**, select the **Download log** link to open (or save) the status log for the connector. This log contains information about each time the script runs and uploads the data from the text file to the Microsoft cloud.

   The `RecordsSaved` field indicates the number of rows in the text file that the script uploads. For example, if the text file contains four rows, the value of the `RecordsSaved` fields is 4, if the script successfully uploads all the rows in the text file.

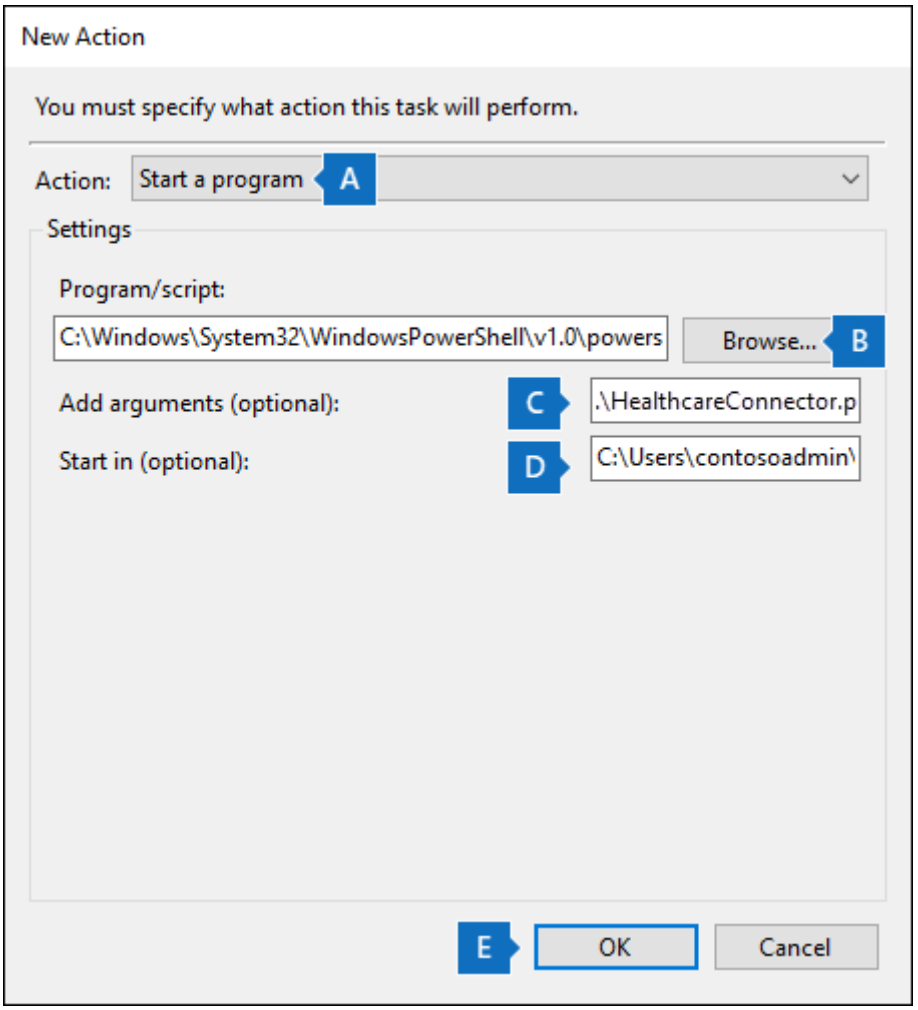
If you don't run the script in Step 4, a link to download the script appears under **Last import**. You can download the script and then follow the steps to run the script.

(Optional) Step 6: Schedule the script to run automatically

To make sure tools like the Insider Risk Management solution always have the latest auditing data from your healthcare EHR system, schedule the script to run automatically every day. You also need to update the EHR auditing data in the same text file on a similar schedule so it contains the latest information about patient records access activities by your employees. The goal is to upload the most current auditing data so the Healthcare connector can make it available to the insider risk management solution.

Use the Task Scheduler app in Windows to automatically run the script every day.

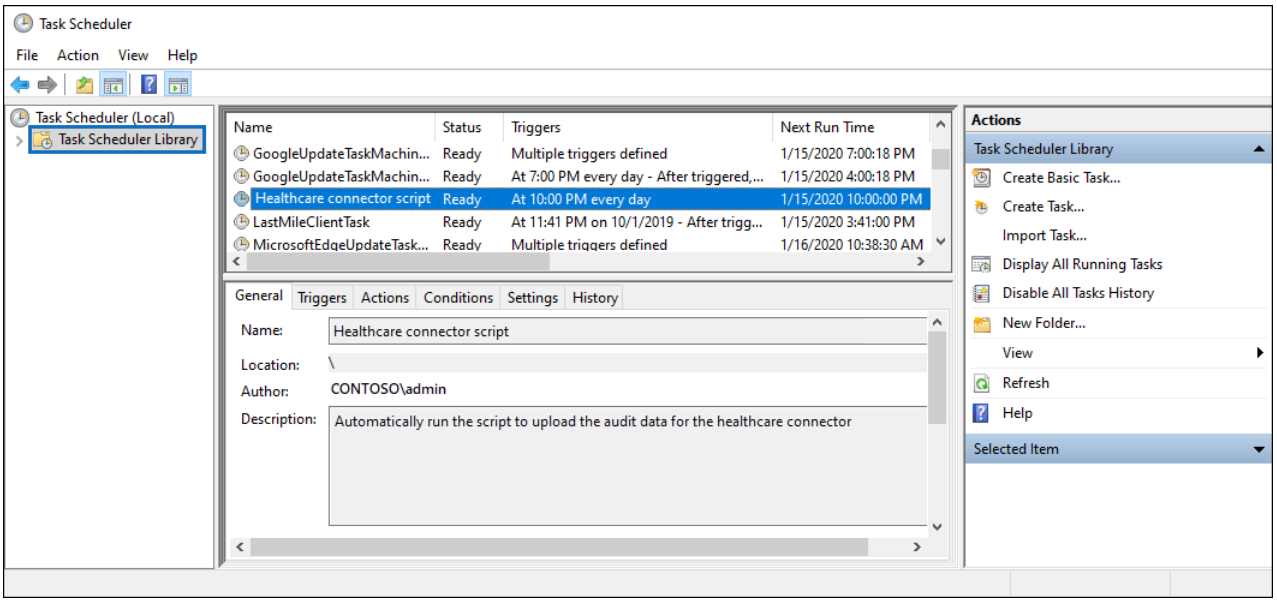
1. On your local computer, select the Windows **Start** button and type **Task Scheduler**.
2. Select the **Task Scheduler** app to open it.
3. In the **Actions** section, select **Create Task**.
4. On the **General** tab, type a descriptive name for the scheduled task, such as **Healthcare connector script**. You can also add an optional description.
5. Under **Security options**, do the following things:
   a. Decide whether to run the script only when you're logged on to the computer or run it when you're logged on or not.
   b. Make sure the **Run with the highest privileges** checkbox is selected.
6. Select the **Triggers** tab, select **New**, and then do the following things:
   a. Under **Settings**, select the **Daily** option, then choose a date and time to run the script for the first time. The script runs every day at the same specified time.
   b. Under **Advanced settings**, make sure the **Enabled** checkbox is selected.
   c. Select **Ok**.
7. Select the **Actions** tab, select **New**, and then do the following things:



a. In the **Action** dropdown list, make sure that **Start a program** is selected.
b. In the **Program/script** box, select **Browse**, go to the following location, and select it so the path is displayed in the box: `C:\Windows\System32\WindowsPowerShell\v1.0\powershell.exe`.
c. In the **Add arguments (optional)** box, paste the same script command that you ran in Step 4. For example, `.\HealthcareConnector.ps1 -tenantId "aaaabbbb-0000-cccc-1111-dddd2222eeee" -appId "00001111-aaaa-2222-bbbb-3333cccc4444" -appSecret "Aa1Bb~2Cc3.-Dd4Ee5Ff6Gg7Hh8Ii9_Jj0Kk1Ll2" -jobId "00001111-aaaa-2222-bbbb-3333cccc4444" -filePath "C:\Healthcare\audit\records.txt"`
d. In the **Start in (optional)** box, paste the folder location of the script that you ran in Step 4. For example, C:\Healthcare\audit.
e. Select **Ok** to save the settings for the new action.

8. In the **Create Task** window, select **Ok** to save the scheduled task. You might be prompted to enter your user account credentials.

   The new task appears in the Task Scheduler Library.



The last time the script ran and the next time it's scheduled to run is displayed. You can double-select the task to edit it.

You can also verify the last time the script ran on the flyout page of the corresponding Healthcare connector in the Microsoft Purview portal.

Last updated on 03/30/2026

Set up a connector to import HR data

You can set up a data connector to import human resources (HR) data related to events such as a user's resignation or a change in a user's job level. The insider risk management solution uses the HR data to generate risk indicators that help you identify possible malicious activity or data theft by users inside your organization.

Setting up a connector for HR data that insider risk management policies use to generate risk indicators involves creating a CSV file that contains the HR data, creating an app in Microsoft Entra for authentication, creating an HR data connector in the Microsoft Purview portal, and running a script (on a scheduled basis) that ingests the HR data in CSV files to the Microsoft cloud so it's available to the insider risk management solution.

Before you begin

- Determine which HR scenarios and data to import to Microsoft 365. This determination helps you decide how many CSV files and HR connectors you need to create, and how to generate and structure the CSV files. The insider risk management policies you want to implement determine the HR data you import. For more information, see Step 1.
- Determine how to retrieve or export the data from your organization's HR system (regularly) and add it to the CSV files that you create in Step 1. The script that you run in Step 4 uploads the HR data in the CSV files to the Microsoft cloud.
- Assign the Data Connector Admin role to the user who creates the HR connector in Step 3. This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions, see:- Permissions in the Microsoft Purview portal - Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview compliance
- Understand that the sample script you run in Step 4 uploads your HR data to the Microsoft cloud so that the insider risk management solution can use it. This sample script isn't supported under any Microsoft standard support program or service. The sample script is provided AS IS without warranty of any kind. Microsoft further disclaims all implied warranties including, without limitation, any implied warranties of merchantability or of fitness for a particular purpose. You assume all risk arising from the use or performance of the sample script and documentation. In no event shall Microsoft, its authors, or anyone else involved in the creation, production, or delivery of the scripts be liable for any damages whatsoever (including, without limitation, damages for loss of

business profits, business interruption, loss of business information, or other pecuniary loss) arising out of the use of or inability to use the sample scripts or documentation, even if Microsoft has been advised of the possibility of such damages.

- This data connector is available in GCC environments in the Microsoft 365 US Government cloud. Third-party applications and services might involve storing, transmitting, and processing your organization's customer data on third-party systems that are outside of the Microsoft 365 infrastructure and therefore aren't covered by the Microsoft Purview and data protection commitments. Microsoft makes no representation that use of this product to connect to third-party applications implies that those third-party applications are FedRAMP compliant. For step-by-step instructions for setting up an HR connector in a GCC environment, see Set up a connector to import HR data in US Government.
- Add the *webhook.ingestion.office.com* domain to your firewall allowlist for your organization.

Step 1: Prepare a CSV file with your HR data

First, create a CSV file that contains the HR data the connector imports to Microsoft 365. The insider risk solution uses this data to generate potential risk indicators. You can import data for the following HR scenarios to Microsoft 365:

- Employee resignation. Information about employees who leave your organization.
- Job level changes. Information about job level changes for employees, such as promotions and demotions.
- Performance reviews. Information about employee performance.
- Performance improvement plans. Information about performance improvement plans for employees.
- Employee profile (preview). General information about an employee.

The type of HR data to import depends on the insider risk management policy and corresponding policy template you want to implement. The following table shows which HR data type each policy template requires:

Expand table

Policy template	HR data type
Data theft by departing users	Employee resignations

Policy template	HR data type
Data leaks	Not applicable
Data leaks by priority users	Not applicable
Data leaks by risky users	Job level changes, Performance reviews, Performance improvement plans
Security policy violations	Not applicable
Security policy violations by departing users	Employee resignations
Security policy violations by priority users	Not applicable
Security policy violations by risky users	Job level changes, Performance reviews, Performance improvement plans
Offensive language in email	Not applicable
Healthcare policy	Employee profile

For more information about policy templates for insider risk management, see Insider risk management policies.

For each HR scenario, provide the corresponding HR data in one or more CSV files. The number of CSV files to use for your insider risk management implementation is discussed later in this section.

After you create the CSV file with the required HR data, store it on the local computer where you run the script in Step 4. Implement an update strategy to make sure the CSV file always contains the most current information so that whatever you run the script, the most current HR data is uploaded to the Microsoft cloud and accessible to the insider risk management solution.

> **ⓘ Important**
>
> The column names described in the following sections aren't required parameters, but only examples. You can use any column name in your CSV files. However, you *must* map the column names you use in a CSV file to the data type when you create the HR connector in Step 3. Also note that the sample CSV files in the following sections are shown in NotePad view. It's much easier to view and edit CSV files in Microsoft Excel. The following sections describe the required CSV data for each HR scenario.

CSV file for employee resignation data

Here's an example of a CSV file for employee resignation data.

text

```
UserPrincipalName,ResignationDate,LastWorkingDate
sarad@contoso.com,2019-04-23T15:18:02.4675041+05:30,2019-04-
29T15:18:02.4675041+05:30
pilarp@contoso.com,2019-04-24T09:15:49Z,2019-04-29T15:18:02.7117540
```

The following table describes each column in the CSV file for employee resignation data.

Expand table

Column	Description
UserPrincipalName	The Microsoft Entra UserPrincipalName (UPN) used to identify the terminated user.
ResignationDate	Specifies the date the user's employment is officially terminated or the user resigns from your organization. For example, this date might be when the user gives their notice about leaving your organization. This date might be different from the date of the person's last day of work. Use the following date format: `yyyy-mm-ddThh:mm:ss.nnnnnn| -hh:mm`, which is the ISO 8601 date and time format.
LastWorkingDate	Specifies the last day of work for the terminated user. This date can't be more than six months prior or one year in advance from the time of upload. Use the following date format: `yyyy-mm-ddThh:mm:ss.nnnnnn| -hh:mm`, which is the ISO 8601 date and time format.

CSV file for job level changes data

Here's an example of a CSV file for job level changes data.

text

```
UserPrincipalName,EffectiveDate,OldLevel,NewLevel
sarad@contoso.com,2019-04-23T15:18:02.4675041+05:30,Level 61 - Sr. Manager,Level
60- Manager
pillar@contoso.com,2019-04-23T15:18:02.4675041+05:30,Level 62 - Director,Level 60-
Sr. Manager
```

The following table describes each column in the CSV file for job level changes data.

Expand table

Column	Description
UserPrincipalName	The Microsoft Entra UserPrincipalName (UPN) used to specify the user's email address.
EffectiveDate	Specifies the date that the user's job level is officially changed. Use the following date format: `yyyy-mm-ddThh:mm:ss.nnnnnn+ | -hh:mm`, which is the ISO 8601 date and time format.
Remarks	Specifies the remarks that evaluator provides about the change of job level. You can enter a limit of 200 characters. This parameter is optional. You don't have to include it in the CSV file.
OldLevel	Specifies the user's job level before it was changed. This is a free-text parameter and can contain hierarchical taxonomy for your organization. This parameter is optional. You don't have to include it in the CSV file.
NewLevel	Specifies the user's job level after it was changed. This is a free-text parameter and can contain hierarchical taxonomy for your organization. This parameter is optional. You don't have to include it in the CSV file.

CSV file for performance review data

Here's an example of a CSV file for performance data.

text

```
UserPrincipalName,EffectiveDate,Remarks,Rating
sarad@contoso.com,2019-04-23T15:18:02.4675041+05:30,Met expectations but bad
attitude,2-Below expectation
pillar@contoso.com,2019-04-23T15:18:02.4675041+05:30, Multiple conflicts with the
team
```

The following table describes each column in the CSV file for performance review data.

Expand table

Column	Description
UserPrincipalName	The Microsoft Entra UserPrincipalName (UPN) used to specify the user's email address.
EffectiveDate	Specifies the date that the user is officially informed about the result of their performance review. This date can be the date when the performance review cycle ends. Use the following date format: `yyyy-mm-ddThh:mm:ss.nnnnnn+ | -hh:mm`, which is the ISO 8601 date and time format.

Column	Description
Remarks	Specifies any remarks that evaluator provides to the user for the performance review. This is a text parameter with a limit of 200 characters. This parameter is optional. You don't have to include it in the CSV file.
Rating	Specifies the rating provided for the performance review. This is a text parameter and can contain any free-form text that your organization uses to recognize the evaluation. For example, "3 Met expectations" or "2 Below average". This is a text parameter with a limit of 25 characters. This parameter is optional. You don't have to include it in the CSV file.

CSV file for performance improvement plan data

Here's an example of a CSV file for the data for the performance improvement plan data.

text

```
UserPrincipalName,EffectiveDate,ImprovementRemarks,PerformanceRating
sarad@contoso.com,2019-04-23T15:18:02.4675041+05:30,Met expectation but bad
attitude,2-Below expectation
pillar@contoso.com,2019-04-23T15:18:02.4675041+05:30, Multiple conflicts with the
team
```

The following table describes each column in the CSV file for performance review data.

Expand table

Column	Description
UserPrincipalName	The Microsoft Entra UserPrincipalName (UPN) used to specify the user's email address.
EffectiveDate	Specifies the date when the user is officially informed about their performance improvement plan. You must use the following date format: `yyyy-mm-ddThh:mm:ss.nnnnnn+ | -hh:mm`, which is the ISO 8601 date and time format.
Remarks	Specifies any remarks that evaluator provides about the performance improvement plan. This is a text parameter with a limit of 200 characters. This is an optional parameter. You don't have to include it in the CSV file.
Rating	Specifies any rating or other information related to the performance review. This is a text parameter and can contain any free form text that your organization uses to recognize the evaluation. For example, "3 Met expectations" or "2 Below average". This is a text parameter with limit of 25 characters. This is an optional parameter. You don't have to include it in the CSV file.

CSV file for employee profile data (preview)

> **Note**
>
> The capability to create an HR connector for employee profile data is in public preview. To create an HR connector that supports employee profile data, go to the **Data connectors** page in the Microsoft Purview portal, select the **Connectors** tab, then select **Add a connector** > **HR**. Follow the steps to create a connector in Step 3: Create the HR connector.

Here's an example of a CSV file for employee profile data.

text

```
UserPrincipalName,UserName,EmployeeFirstName,EmployeeLastName,EmployeeAddLine1,Empl
oyeeAddLine2,EmployeeCity,EmployeeState,EmployeeZipCode,EmployeeDept,EmployeeType,E
mployeeRole
jackq@contoso.com,jackq,jack,qualtz,50 Oakland
Ave,#206,City,Florida,32104,Orthopaedic,Regular,Nurse
```

The following table describes each column in the CSV file for employee profile data.

Expand table

Column	Description
UserPrincipalName*	The Microsoft Entra UserPrincipalName (UPN) used to specify the user's email address.
EmployeeFirstName*	First name of the employee.
EmployeeLastName*	Last name of the employee.
EmployeeAddressLine1*	Street address of the employee.
EmployeeAddressLine2	Secondary address information, such as apartment number, for employee.
EmployeeCity	City of residence for employee.
EmployeeState	State of residence for employee.
EmployeeZipCode*	Zip code of residence for employee.
EmployeeCountry	Country of residence for employee.
EmployeeDepartment	Employee's department in the organization.

Column	Description
EmployeeType	Employment type for employee, such as Regular, Exempt, or Contractor.
EmployeeRole	Employees's role, designation, or job title in the organization.

> **Note**
>
> * This column is mandatory. If a mandatory column is missing, the CSV file isn't validated and other data in the file isn't imported.

We recommend that you create an HR connector that only imports employee profile data. For this connector, frequently refresh the employee profile data, preferably every 15 to 20 days. Employee profile records are deleted if you don't update them in the past 30 days.

Determining how many CSV files to use for HR data

In Step 3, you can choose to create separate connectors for each HR data type or a single connector for all data types. You can use separate CSV files that contain data for one HR scenario (like the examples of the CSV files described in the previous sections). Alternatively, you can use a single CSV file that contains data for two or more HR scenarios. Here are some guidelines to help you determine how many CSV files to use for HR data.

- If the insider risk management policy that you want to implement requires multiple HR data types, consider using a single CSV file that contains all the required data types.
- The method for generating or collecting the HR data might determine the number of CSV files. For example, if the different types of HR data used to configure an HR connector are located in a single HR system in your organization, then you might be able to export the data to a single CSV file. But if data is distributed across different HR systems, then it might be easier to export data to different CSV files. For example, employee resignation data might be located in a different HR system than job level or performance review data. In this case, it might be easier to have separate CSV files rather than having to manually combine the data into a single CSV file. So, how you retrieve or export data from your HR systems might determine how many CSV files you need.
- As a general rule, the data types in a CSV file determine the number of HR connectors that you need to create. For example, if a CSV file contains all the data types required to support your insider risk management implementation, then you only need one HR connector. But if you have two separate CSV files that each contain a single data type, then you have to create two HR connectors. An exception to this rule is that if you add an

HRScenario column to a CSV file (see the next section), you can configure a single HR connector that can process different CSV files.

- For each CSV file, you can ingest up to 500 records at once. To ingest a larger number of records, upload multiple CSV files, each with fewer than 500 records.

Configuring a single CSV file for multiple HR data types

You can add multiple HR data types to a single CSV file. This configuration is useful if the insider risk management solution you're implementing requires multiple HR data types or if the data types are located in a single HR system in your organization. Having fewer CSV files always allows you to have fewer HR connectors to create and manage.

Here are requirements for configuring a CSV file with multiple data types:

- Add the required columns (and optional columns if you use them) for each data type and the corresponding column name in the header row. If a data type doesn't correspond to a column, leave the value blank.

- Add an **HRScenario** column to the CSV file. The HR connector uses this column to identify which rows in the CSV file contain which type HR data. The values in this column identify the type of HR data in each row. For example, values that correspond to the HR scenarios could be `Resignation`, `Job level change`, `Performance review`, `Performance improvement plan`, and `Employee profile`.

- If you have multiple CSV files that contain an **HRScenario** column, make sure that each file uses the same column name and the same values that identify the specific HR scenarios.

The following example shows a CSV file that contains the **HRScenario** column. The values in the HRScenario column identify the type of data in the corresponding row. The following sample covers four HR scenarios `Resignation`, `Job level change`, `Performance review`, and `Performance improvement plan`.

text

```
HRScenario,EmailAddress,ResignationDate,LastWorkingDate,EffectiveDate,Remarks,Ratin
g,OldLevel,NewLevel
Resignation,sarad@contoso.com,2019-04-23T15:18:02.4675041+05:30,2019-04-
29T15:18:02.4675041+05:30,,,,
Resignation,pilarp@contoso.com,2019-04-24T09:15:49Z,2019-04-29T15:18:02.7117540,,,,
Job level change,sarad@contoso.com,2019-04-23T15:18:02.4675041+05:30,,,,,Level 61
Sr. Manager, Level 60 Manager
Job level change,pillarp@contoso.com,2019-04-23T15:18:02.4675041+05:30,,,,,Level 62
Director,Level 60 Sr Manager
Performance review,sarad@contoso.com,,,2019-04-23T15:18:02.4675041+05:30,Met
```

```
expectation but bad attitude,2 Below expectations,,
Performance review,pillarp@contoso.com,,,2019-04-23T15:18:02.4675041+05:30,
Multiple conflicts with the team,,
Performance improvement plan,sarad@contoso.com,,,2019-04-
23T15:18:02.4675041+05:30,Met expectations but bad attitude,2 Below expectations,,
Performance improvement plan,pillarp@contoso.com,,,2019-04-
23T15:18:02.4675041+05:30,Multiple conflicts with the team,,
```

> **Note**
>
> You can use any name for the column that identifies HR data type because you map the name of the column in your CSV file as the column that identifies the HR data type when you set up the connector in Step 3. You also map the values used for the data type column when you set up the connector.

Adding the HRScenario column to a CSV file that contains a single data type

Based on your organization's HR systems and how you export HR data to a CSV file, you might need to create multiple CSV files that each contain a single HR data type. In this case, you can still create a single HR connector to import data from different CSV files. To do this, add an HRScenario column to the CSV file and specify the HR data type. Then, you can run the script for each CSV file but use the same job ID for the connector. For more information, see Step 4.

Step 2: Create an app in Microsoft Entra ID

Next, create and register a new app in Microsoft Entra ID. The app corresponds to the HR connector that you create in Step 3. When you create this app, Microsoft Entra ID can authenticate the HR connector when it runs and attempts to access your organization. This app also authenticates the script that you run in Step 4 to upload your HR data to the Microsoft cloud. During the creation of this Microsoft Entra app, save the following information. You use these values in Step 3 and Step 4.

- Microsoft Entra application ID (also called the *app ID* or *client ID*)
- Microsoft Entra application secret (also called the *client secret*)
- Tenant ID (also called the *directory Id*)

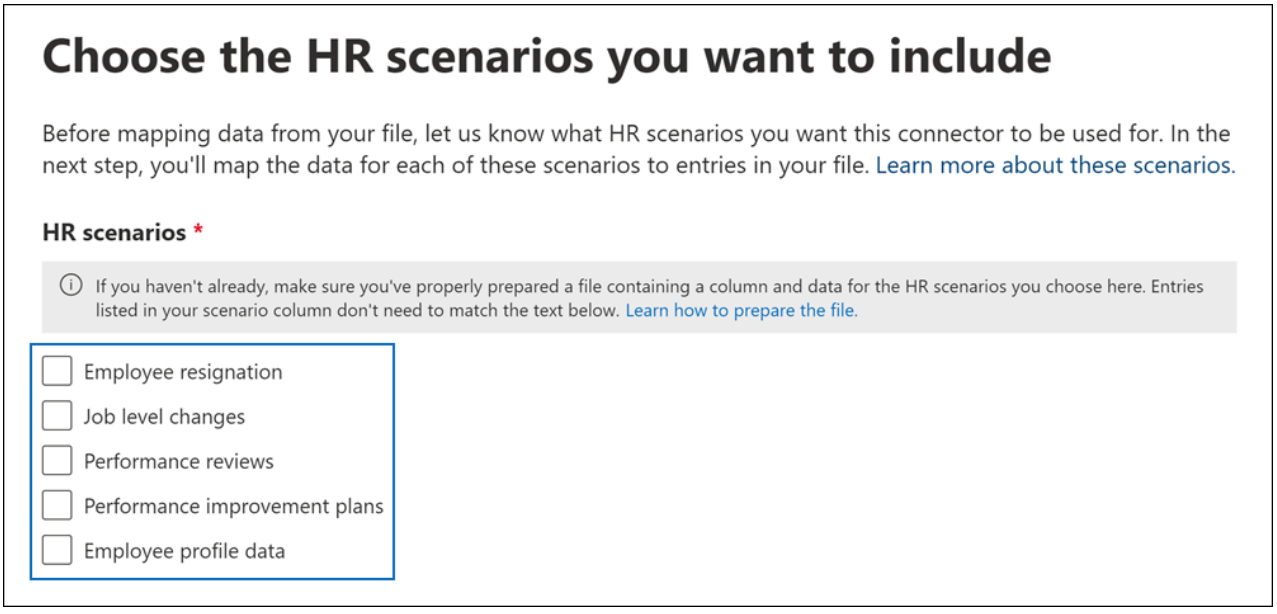
For step-by-step instructions for creating an app in Microsoft Entra ID, see Register an application with the Microsoft identity platform.

Step 3: Create the HR connector

The next step is to create an HR connector in the Microsoft Purview portal. After you run the script in Step 4, the HR connector that you create will ingest the HR data from the CSV file to your Microsoft 365 organization. Before you create a connector, be sure that you have a list of the HR scenarios and the corresponding CSV column names for each one. You have to map the data required for each scenario to the actual column names in your CSV file when configuring the connector. Alternatively, you can upload a sample CSV file when configuring the connector and the workflow help you map the name of the columns to the required data types.

After you complete this step, be sure to copy the job ID that's generated when you create the connector. You use the job ID when you run the script.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. From the list, choose **HR (preview)**.
5. On the **Setup the connection** page, do the following and then select **Next**:
   a. Type or paste the Microsoft Entra application ID for the Azure app that you created in Step 2.
   b. Type a name for the HR connector.
6. On the HR scenarios page, select one or more HR scenarios that you want to import data for and then select **Next**.



7. On the file mapping method page, select a file type if necessary, and then select one of the following options and then select **Next**.

- **Upload a sample file**. If you select this option, select **Upload sample file** to upload the CSV file that you prepared in Step 1. This option allows you to quickly select column names in your CSV file from a drop-down list to map them to the data types for the HR scenarios that you previously selected.

OR

- **Manually provide the mapping details**. If you select this option, you have to type the name of the columns in your CSV file to map them to the data types for the HR scenarios that you previously selected.

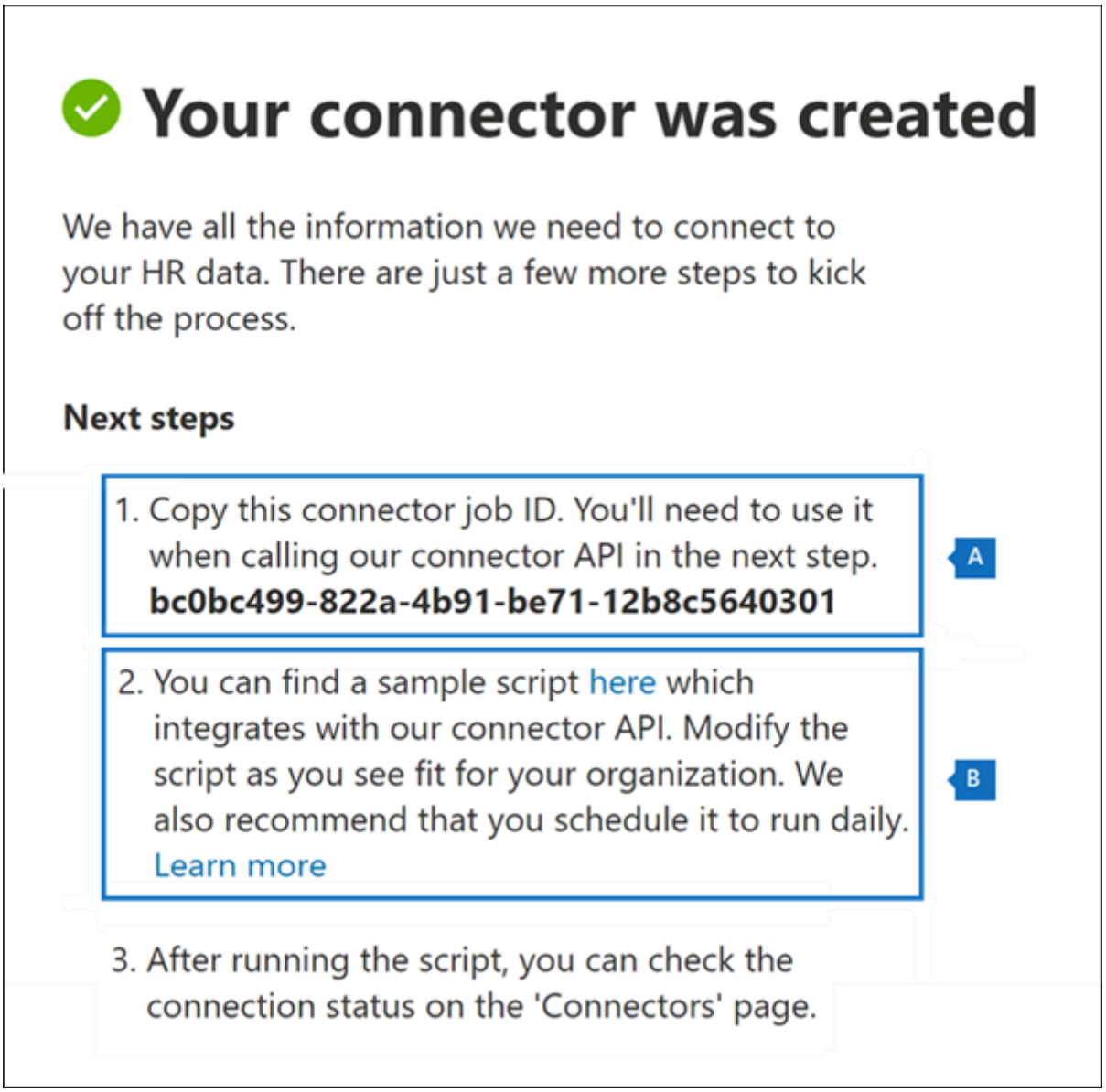
8. On the File mapping details page, do one of the following, depending on whether you uploaded a sample CSV file and whether you're configuring the connector for a single HR scenario or for multiple scenarios. If you uploaded a sample file, you don't have to type the column names. You pick them from a dropdown list.

   - If you selected a single HR scenario in the previous step, then type the column header names (also called *parameters*) from the CSV file that you created in Step 1 in each of the appropriate boxes. The column names that you type aren't case-sensitive, but be sure to include spaces if the column names in your CSV file include spaces. As previously explained, the names you type in these boxes must match the parameter names in your CSV file. For example, the following screenshot shows the parameter names from the sample CSV file for the employee resignation HR scenario shown in Step 1.

   - If you selected multiple data types in the previous step, then you need to enter identifier column name that will identify the HR data type in your CSV file. After entering the identifier column name, type the value that identifies this HR data type, and type the column header names for selected data types from the CSV file(s) that you created in Step 1 in each of the appropriate boxes for each selected data type. As previously explained, the names that you type in these boxes must match the column names in your CSV file.

9. On the **Review** page, review your settings and then select **Finish** to create the connector.

   A status page is displayed that confirms the connector was created. This page contains two important things that you need to complete the next step to run the sample script to upload your HR data.



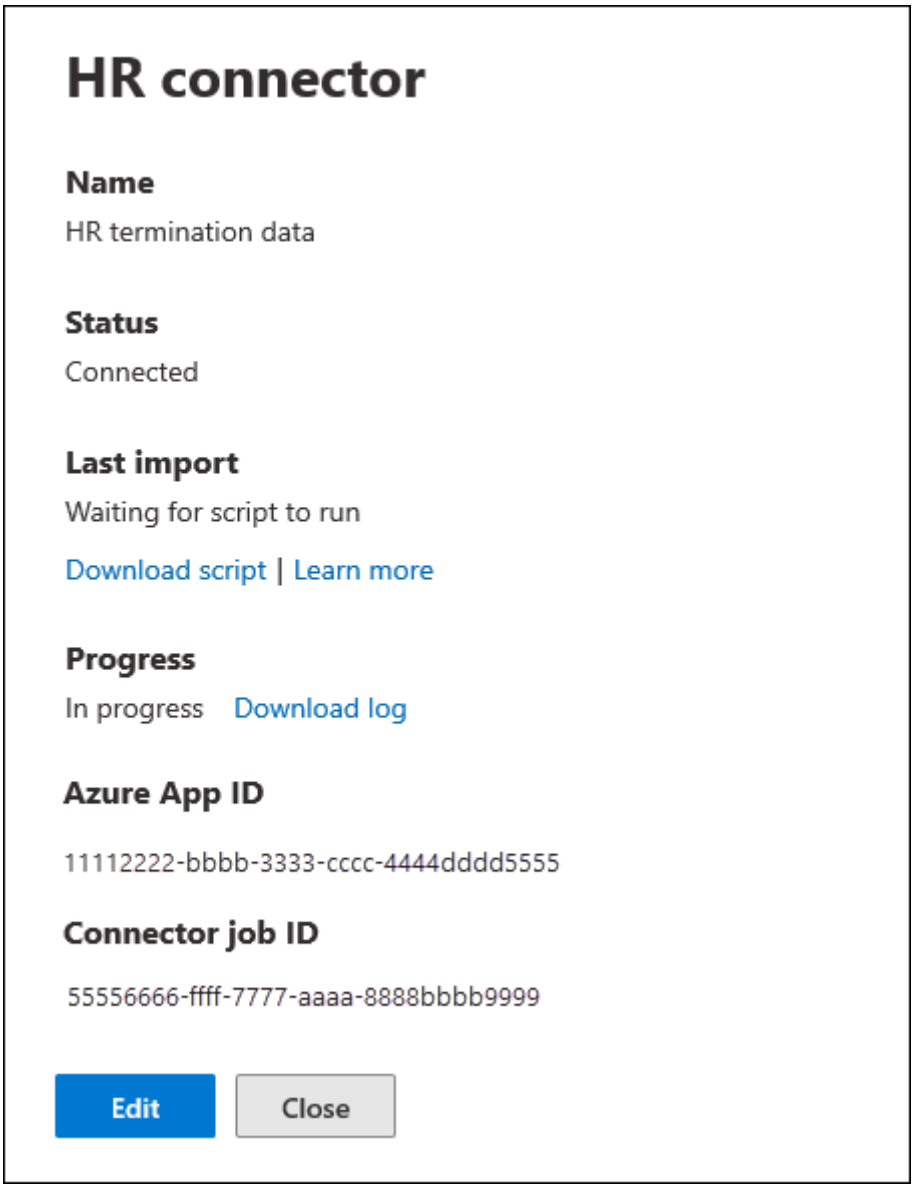
a. **Job ID.** You'll need this job ID to run the script in the next step. You can copy it from this page or from the connector flyout page.

b. **Link to sample script.** Select the **here** link to go to the GitHub site to access the sample script (the link opens a new window). Keep this window open so that you can copy the script in Step 4. Alternatively, you can bookmark the destination or copy the URL so you can access it again when you run the script. This link is also available on the connector flyout page.

10. Select **Done**.

    The new connector is displayed in the list on the **Connectors** tab.

11. Select the HR connector that you just created to display the flyout page, which contains properties and other information about the connector.



If you haven't already done so, you can copy the values for the **Azure App ID** and **Connector job ID**. You'll need these to run the script in the next step. You can also download the script from the flyout page (or download it using the link in the next step.)

You can also select **Edit** to change the Azure App ID or the column header names that you defined on the **File mapping** page.

Step 4: Run the sample script to upload your HR data

> **Important**
>
> You must add the *webhook.ingestion.office.com* domain to your firewall allowlist for your organization. If you block this domain, the script doesn't run.

The last step in setting up an HR connector is to run a sample script that uploads the HR data in the CSV file (that you created in Step 1) to the Microsoft cloud. Specifically, the script uploads the data to the HR connector. After you run the script, the HR connector that you created in Step 3 imports the HR data to your Microsoft 365 organization where other compliance tools, such as the Insider risk management solution, can access it. After you run the script, consider scheduling a task to run it automatically on a daily basis so the most current employee termination data is uploaded to the Microsoft cloud. For more information, see Schedule the script to run automatically.

1. Go to the window that you left open from the previous step to access the GitHub site with the sample script. Alternatively, open the bookmarked site or use the URL that you copied. You can also access the script here.
2. Select the **Raw** button to display the script in text view.
3. Copy all the lines in the sample script and save them to a text file.
4. Modify the sample script for your organization, if necessary.
5. Save the text file as a Windows PowerShell script file by using a filename suffix of `.ps1`; for example, `HRConnector.ps1`. Alternatively, you can use the GitHub filename for the script, which is `upload_termination_records.ps1`.
6. Open a command prompt on your local computer, and go to the directory where you saved the script.
7. Run the following command to upload the HR data in the CSV file to the Microsoft cloud; for example:

   PowerShell

   ```
   .\HRConnector.ps1 -tenantId <tenantId> -appId <appId>  -appSecret <appSecret>
   -jobId <jobId>  -filePath '<filePath>'
   ```

   The following table describes the parameters to use with this script and their required values. The information you obtained in the previous steps is used in the values for these parameters.

   Expand table

Parameter	Description
`tenantId`	This is the ID for your Microsoft 365 organization that you obtained in Step 2. You can also obtain the tenant ID for your organization on the **Overview** blade in the Microsoft Entra admin center. This value identifies your organization.

Parameter	Description
`appId`	This is the Microsoft Entra application ID for the app that you created in Microsoft Entra ID in Step 2. This value is used by Microsoft Entra ID for authentication when the script attempts to access your Microsoft 365 organization.
`appSecret`	This is the Microsoft Entra application secret for the app that you created in Microsoft Entra ID in Step 2. This value is also used for authentication.
`jobId`	This is the job ID for the HR connector that you created in Step 3. This value associates the HR data that is uploaded to the Microsoft cloud with the HR connector.
`filePath`	This is the file path for the file (stored on the same system as the script) that you created in Step 1. Try to avoid spaces in the file path; otherwise use single quotation marks.

Here's an example of the syntax for the HR connector script using actual values for each parameter:

PowerShell

```
.\HRConnector.ps1 -tenantId aaaabbbb-0000-cccc-1111-dddd2222eeee -appId
00001111-aaaa-2222-bbbb-3333cccc4444 -appSecret Aa1Bb~2Cc3.-
Dd4Ee5Ff6Gg7Hh8Ii9_Jj0Kk1Ll2 -jobId 00001111-aaaa-2222-bbbb-3333cccc4444 -
filePath 'C:\Users\contosoadmin\Desktop\Data\employee_termination_data.csv'
```

If the upload is successful, the script displays the **Upload Successful** message.

> **Note**
>
> If you have problems running the previous command because of execution policies, see About Execution Policies and Set-ExecutionPolicy for guidance about setting execution policies.

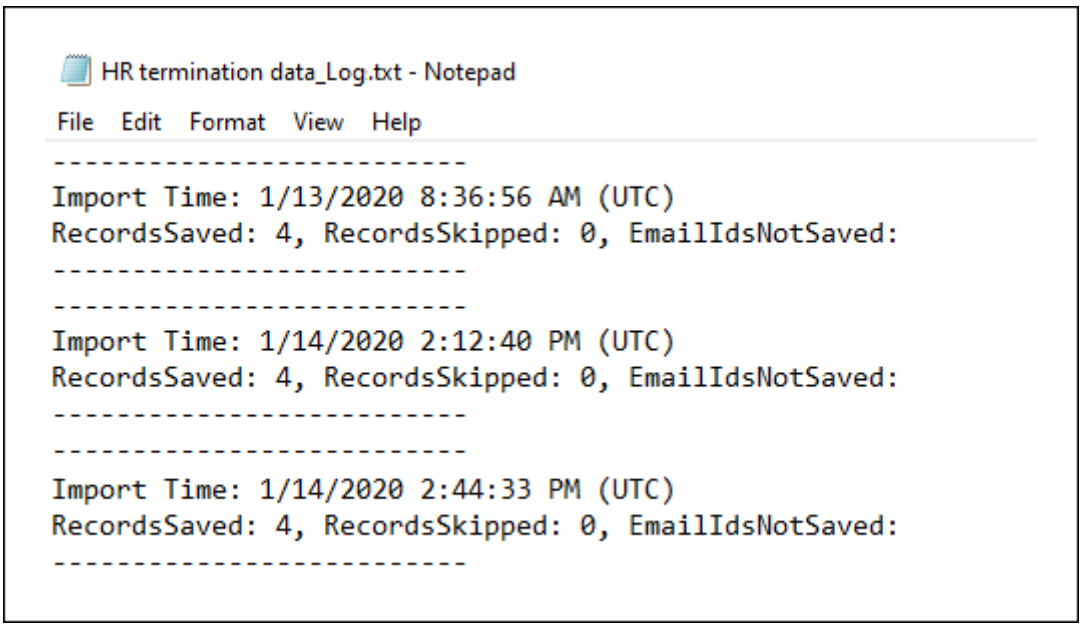
Step 5: Monitor the HR connector

After you create the HR connector and run the script to upload your HR data, you can view the connector and upload status in the Microsoft Purview portal. If you schedule the script to run automatically on a regular basis, you can also view the current status after the last time the script ran.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.

3. Select **My connectors**, then select the HR connector that you created to display the flyout page. This page contains the properties and information about the connector.

4. Under **Progress**, select the **Download log** link to open (or save) the status log for the connector. This log contains information about each time the script runs and uploads the data from the CSV file to the Microsoft cloud.



   The `RecordsSaved` field indicates the number of rows in the CSV file that you uploaded. For example, if the CSV file contains four rows, the value of the `RecordsSaved` fields is 4, if the script successfully uploads all the rows in the CSV file.

If you didn't run the script in Step 4, a link to download the script is displayed under **Last import**. You can download the script and then follow the steps to run the script.

(Optional) Step 6: Schedule the script to run automatically

To make sure tools like the insider risk management solution always have access to the latest HR data from your organization, schedule the script to run automatically on a recurring basis, such as once a day. This schedule requires updating the HR data in the CSV file on a similar (if not the same) schedule so that it contains the latest information about employees who leave your organization. The goal is to upload the most current HR data so that the HR connector can make it available to the insider risk management solution.

Use the Task Scheduler app in Windows to automatically run the script every day.

1. On your local computer, select the Windows **Start** button and then type **Task Scheduler**.

2. Select the **Task Scheduler** app to open it.

3. In the **Actions** section, select **Create Task**.

4. On the **General** tab, type a descriptive name for the scheduled task, such as **HR Connector Script**. You can also add an optional description.

5. Under **Security options**, do the following steps:

   a. Decide whether to run the script only when you're logged on to the computer or run it when you're logged on or not.

   b. Make sure that the **Run with the highest privileges** checkbox is selected.

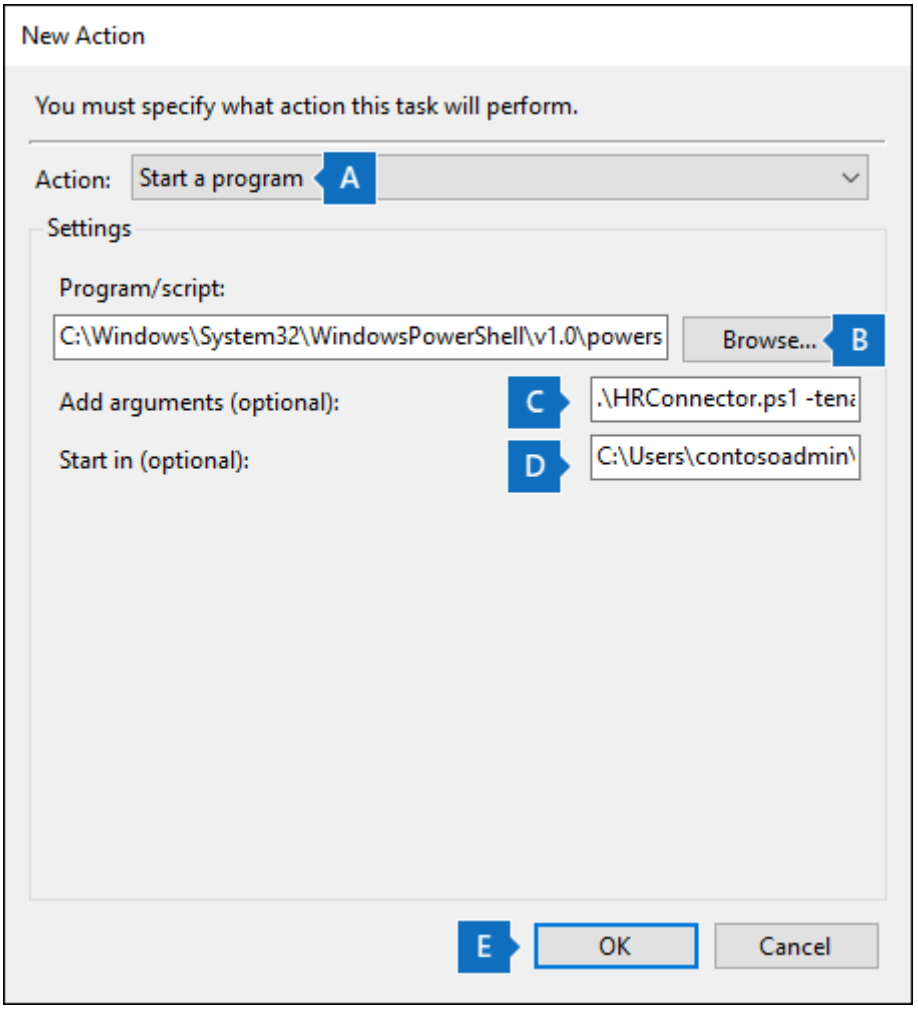
6. Select the **Triggers** tab, select **New**, and then do the following steps:

   a. Under **Settings**, select the **Daily** option, then choose a date and time to run the script for the first time. The script runs every day at the same specified time.

   b. Under **Advanced settings**, make sure the **Enabled** checkbox is selected.

   c. Select **Ok**.

7. Select the **Actions** tab, select **New**, and then do the following steps:



   a. In the **Action** dropdown list, make sure that **Start a program** is selected.

b. In the **Program/script** box, select **Browse**, go to the following location, and select it so the path displays in the box: `C:\Windows\System32\WindowsPowerShell\v1.0\powershell.exe`.

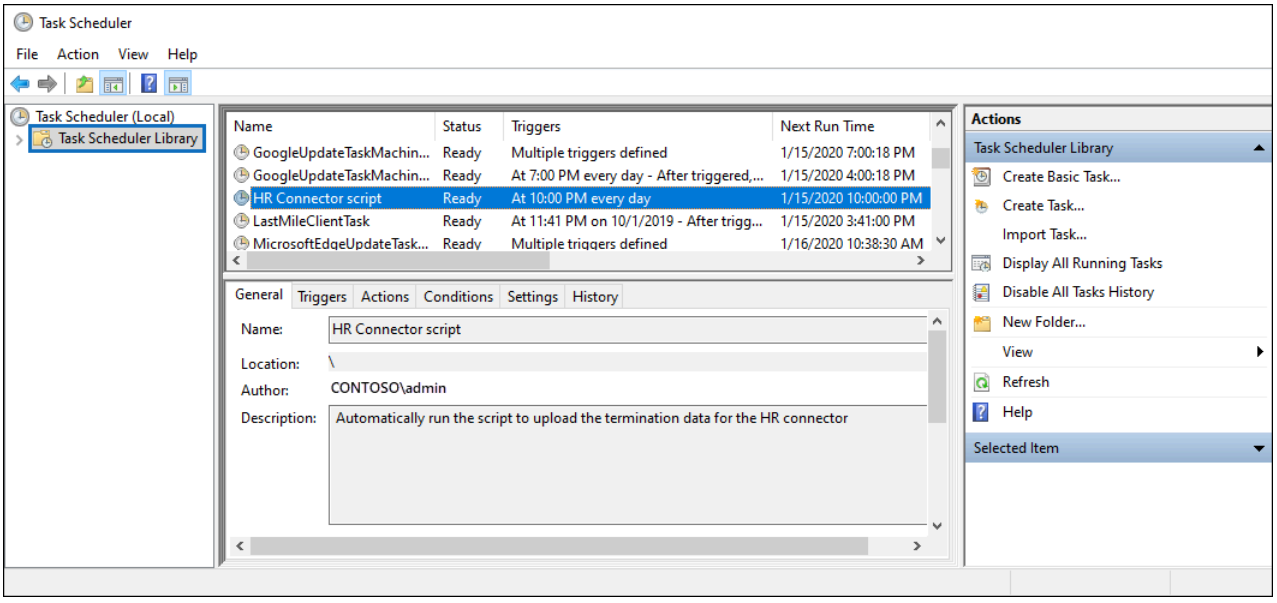
c. In the **Add arguments (optional)** box, paste the same script command that you ran in Step 4. For example, `.\HRConnector.ps1 -tenantId "aaaabbbb-0000-cccc-1111-dddd2222eeee" -appId "00001111-aaaa-2222-bbbb-3333cccc4444" -appSecret "Aa1Bb~2Cc3.-Dd4Ee5Ff6Gg7Hh8Ii9_Jj0Kk1Ll2" -jobId "00001111-aaaa-2222-bbbb-3333cccc4444" -filePath "C:\Users\contosoadmin\Desktop\Data\employee_termination_data.csv"`

d. In the **Start in (optional)** box, paste the folder location of the script that you ran in Step 4. For example, `C:\Users\contosoadmin\Desktop\Scripts`.

e. Select **Ok** to save the settings for the new action.

8. In the **Create Task** window, select **Ok** to save the scheduled task. You might be prompted to enter your user account credentials.

   The new task appears in the Task Scheduler Library.



   The last time the script ran and the next time it's scheduled to run displays. You can double-select the task to edit it.

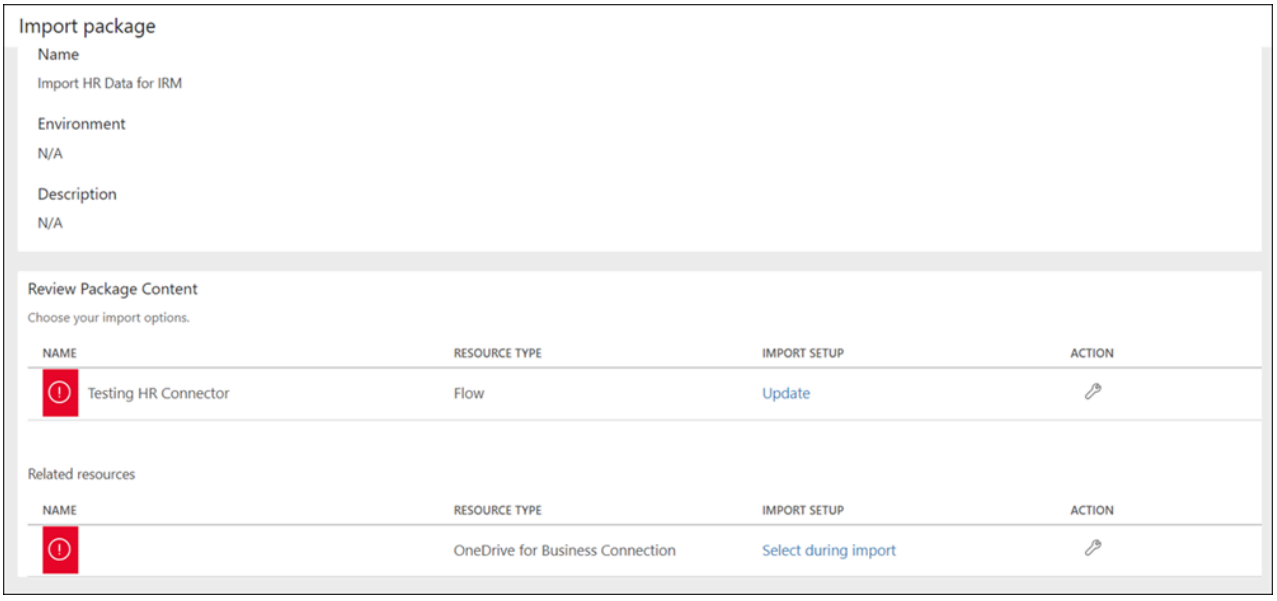
   You can also verify the last time the script ran on the flyout page of the corresponding HR connector in the Microsoft Purview portal.

(Optional) Step 7: Upload data by using Power Automate templates

You can upload HR data by using Power Automate templates and define triggers. For example, you can configure a Power Automate template to trigger when new HR connector files are available in SharePoint or OneDrive locations. You can also streamline this process by storing confidential information like Microsoft Entra application secret (created in **Step 2**) in Azure Key Vault and use it with Power Automate for authentication.

Complete the following steps to automatically upload HR data when new files become available on OneDrive for Business:

1. Download the *ImportHRDataforIRM.zip* package from the GitHub site.
2. In Power Automate, go to **My flows**.
3. Select **Import** and upload the *ImportHRDataforIRM.zip* package.
4. After the package is uploaded, update the content (name and OneDrive for Business connection), and select **Import**.



5. Select **Open flow** and update the parameters. The following table describes the parameters to use in this Power Automate Flow and their required values. Use the information you obtained in the previous steps for these values.

Expand table

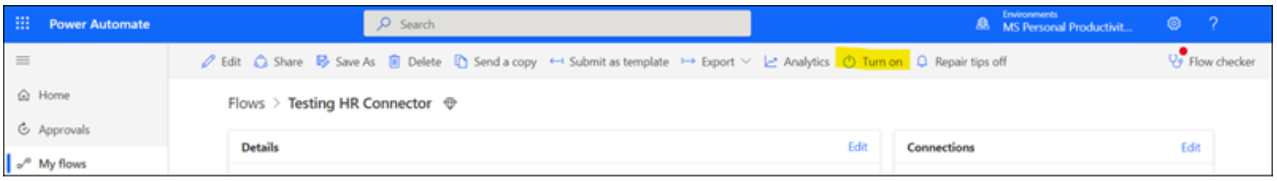
Parameter	Description
App ID	This is the Microsoft Entra application ID for the app that you created in Microsoft Entra ID in **Step 2**. Microsoft Entra ID uses this application ID for authentication when the script attempts to access your Microsoft 365 organization.
App Secret	This is the Microsoft Entra application secret for the app that you created in Microsoft Entra ID in **Step 2**. This application secret is used for authentication.

Parameter	Description
File location	This is the OneDrive for Business location where Power Automate monitors for 'new file created' activities to trigger this flow.
Job ID	Identifier for the HR connector created in **Step 3**. This identifier associates the HR data uploaded to the Microsoft cloud with the HR connector.
Tenant ID	Identifier for your Microsoft 365 organization obtained in **Step 2**. You can also obtain the tenant ID for your organization on the **Overview** blade in the Microsoft Entra admin center. This identifier is used to identify your organization.
URI	Verify that the value for this parameter is *https://webhook.ingestion.office.com/api/signals*

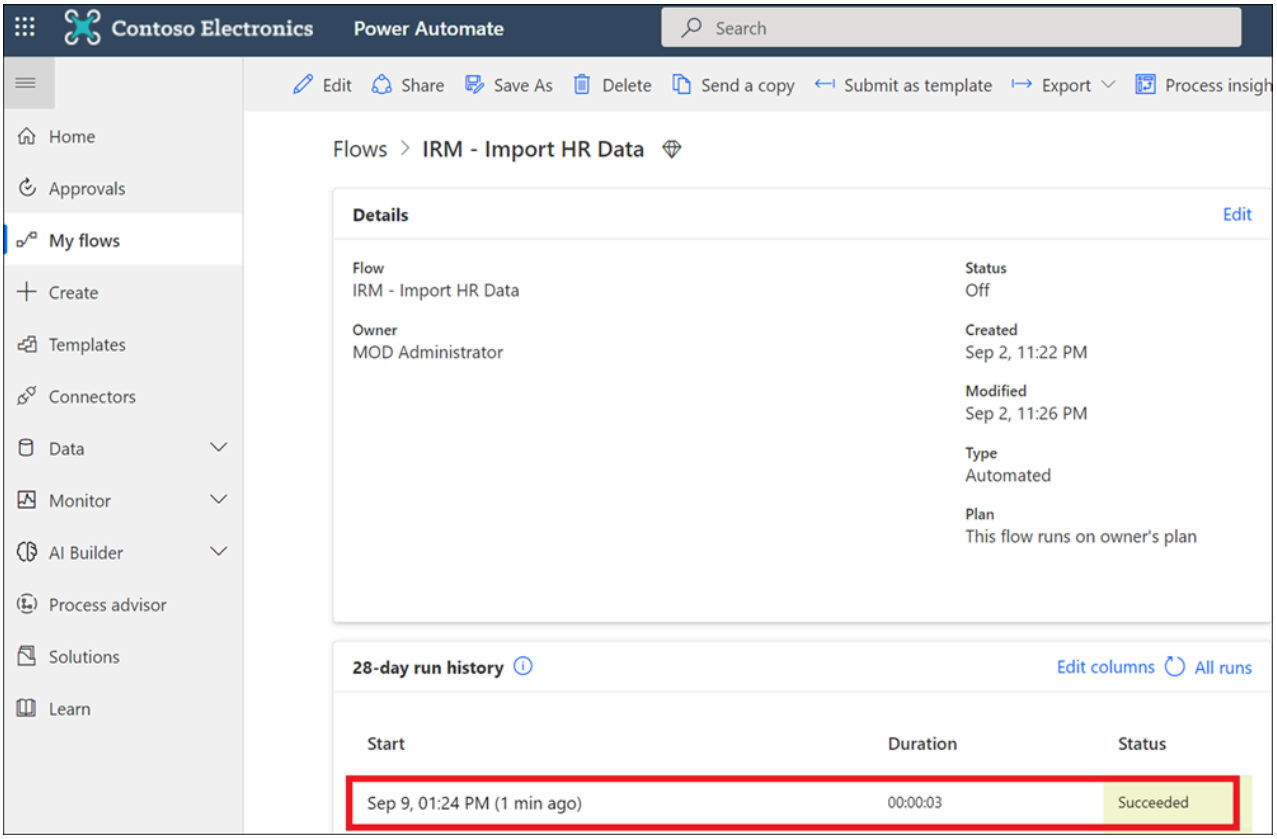


1. Select **Save**.

2. Go to **Flow overview** and select **Turn on**.



3. Test the flow manually by uploading a new file to your OneDrive for Business folder and verify that it ran successfully. This process might take a few minutes after the upload before the flow is triggered.



4. You can now monitor the HR connector as described in **Step 5**.

If needed, you can update the flow to create triggers based on file availability and modification events on SharePoint and other data sources supported by Power Automate Flows.

Existing HR connectors

On December 13, 2021, we released the employee profile data scenario for HR connectors. If you created an HR connector before this date, we migrate the existing instances or your organization's HR connectors so your HR data continues to be imported to the Microsoft cloud. You don't have to do anything to maintain this functionality. You can keep using these connectors without disruption.

If you want to implement the employee profile data scenario, create a new HR connector and configure it as required. After you create a new HR connector, run the script by using the job ID

of the new connector and CSV files with employee profile data previously described in this article.

Last updated on 03/30/2026

Set up a connector to import HR data in US Government

You can set up a data connector to import human resources (HR) data to your US Government organization. HR-related data includes the date an employee submitted their resignation and date of the employee's last day. Microsoft information protection solutions, such as the insider risk management solution, use this HR data to help protect your organization from malicious activity or data theft inside your organization. Setting up an HR connector consists of creating an app in Microsoft Entra that's used for authentication by connector, creating a CSV mapping files that contains your HR data, creating a data connector in the Microsoft Purview portal, and then running a script (on a scheduled basis) that ingests the HR data in the CSV file to the Microsoft cloud. Then the data connector is used by the insider risk management tool to access the HR data that you imported to your Microsoft 365 US Government organization.

Before you begin

- The user who creates the HR connector in Step 3 must be assigned the Data Connector Admin role. This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions, see Permissions in the Microsoft Purview portal and Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview.

  > **Note**
  >
  > The Data Connector Admin role currently isn't supported in US Government GCC High and DoD environments. Therefore, the user who creates the HR connector in GCC High and DoD environments must be assigned the Mailbox Import Export role in Exchange Online. By default, no role group in Exchange Online has this role assigned. You can add the Mailbox Import Export role to the Organization Management role group in Exchange Online. Or you can create a new role group, assign the Mailbox Import Export role, and then add the appropriate users as members. For more information, see

the Create role groups or Modify role groups sections in the article "Manage role groups in Exchange Online".

- Determine how to retrieve or export the data from your organization's HR system (on a regular basis) and add it to the CSV file that's described in Step 2. The script that you run in Step 4 uploads the HR data in the CSV file to the Microsoft cloud.

- The sample script that you run in Step 4 uploads HR data to the Microsoft cloud so that other Microsoft tools, such as the insider risk management solution, can use it. This sample script isn't supported under any Microsoft standard support program or service. The sample script is provided AS IS without warranty of any kind. Microsoft further disclaims all implied warranties including, without limitation, any implied warranties of merchantability or of fitness for a particular purpose. You assume all risk arising from the use or performance of the sample script and documentation. In no event shall Microsoft, its authors, or anyone else involved in the creation, production, or delivery of the scripts be liable for any damages whatsoever (including, without limitation, damages for loss of business profits, business interruption, loss of business information, or other pecuniary loss) arising out of the use of or inability to use the sample scripts or documentation, even if Microsoft has been advised of the possibility of such damages.

Step 1: Create an app in Microsoft Entra ID

First, create and register a new app in Microsoft Entra ID. The app corresponds to the HR connector that you create in Step 3. When you create the app, Microsoft Entra ID can authenticate the HR connector. The app also authenticates the script that you run in Step 4 to upload your HR data to the Microsoft cloud. During the creation of this Microsoft Entra app, save the following information. You use these values in later steps.

- Microsoft Entra application ID (also called the *app Id* or *client Id*)
- Microsoft Entra application secret (also called the *client secret*)
- Tenant Id (also called the *directory Id*)

For step-by-step instructions for creating an app in Microsoft Entra ID, see Register an application with the Microsoft identity platform.

Step 2: Prepare a CSV file with your HR data

In this step, you create a CSV file that contains information about employees who left your organization. As explained in the Before You Begin section, you need to decide how to generate

this CSV file from your organization's HR system. The following example shows a completed CSV file (opened in Note Pad) that contains the three required parameters (columns). It's much easier to edit the CSV file in Microsoft Excel.

```text
EmailAddress,TerminationDate,LastWorkingDate
sarad@contoso.com,2019-04-23T15:18:02.4675041+05:30,2019-04-29T15:18:02.4675041+05:30
pilarp@contoso.com,2019-04-24T09:15:49Z,2019-04-29T15:18:02.7117540
```

The first row, or header row, of the CSV file lists the required column names. You choose the name used in each column header (the ones in the following example are suggestions). However, you must specify the same column names when you create the HR connector in Step 3. Don't include spaces in the column names.

The following table describes each column in the CSV file:

Expand table

Column name	Description
EmailAddress	Specifies the email address of the terminated employee.
TerminationDate	Specifies the date the person's employment was officially terminated in your organization. For example, this date might be when the employee gives their notice about leaving your organization. This date might be different from the date of the person's last day of work. Use the following date format: `yyyy-mm-ddThh:mm:ss.nnnnnn+|-hh:mm`, which is the ISO 8601 date and time format.
LastWorkingDate	Specifies the last day of work for the terminated employee. Use the following date format: `yyyy-mm-ddThh:mm:ss.nnnnnn+|-hh:mm`, which is the ISO 8601 date and time format.

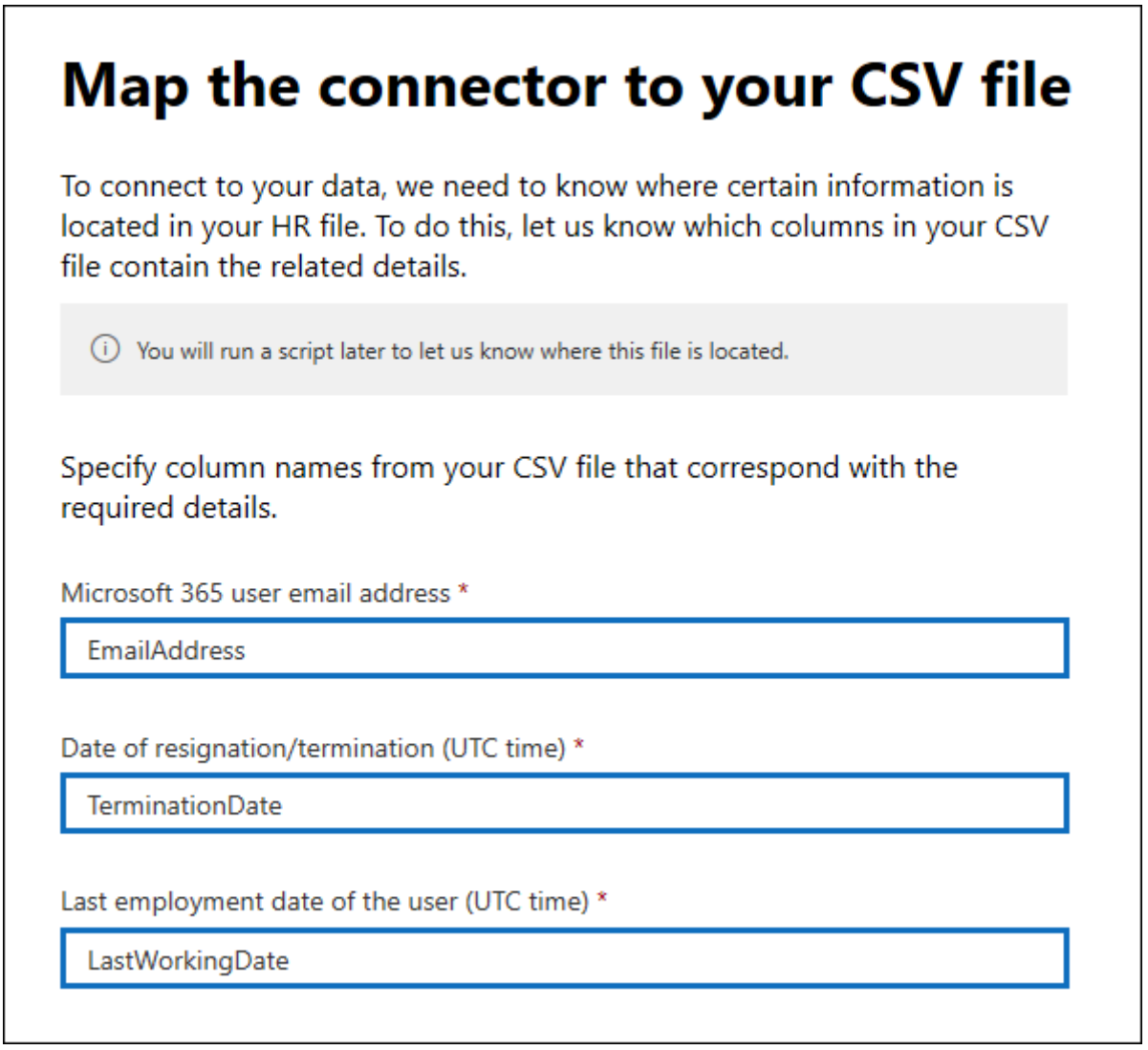
After you create the CSV file with the required HR data, store it on the same system as the script that you run in Step 4. Be sure to implement an update strategy so the CSV file always contains the most current information. This strategy ensures that whatever script you run, the most current employee termination data is uploaded to the Microsoft cloud.

Step 3: Create the HR connector

In this step, you create an HR connector in the Microsoft Purview portal. When you run the script in Step 4, the HR connector ingests the HR data from the CSV file to your Microsoft 365

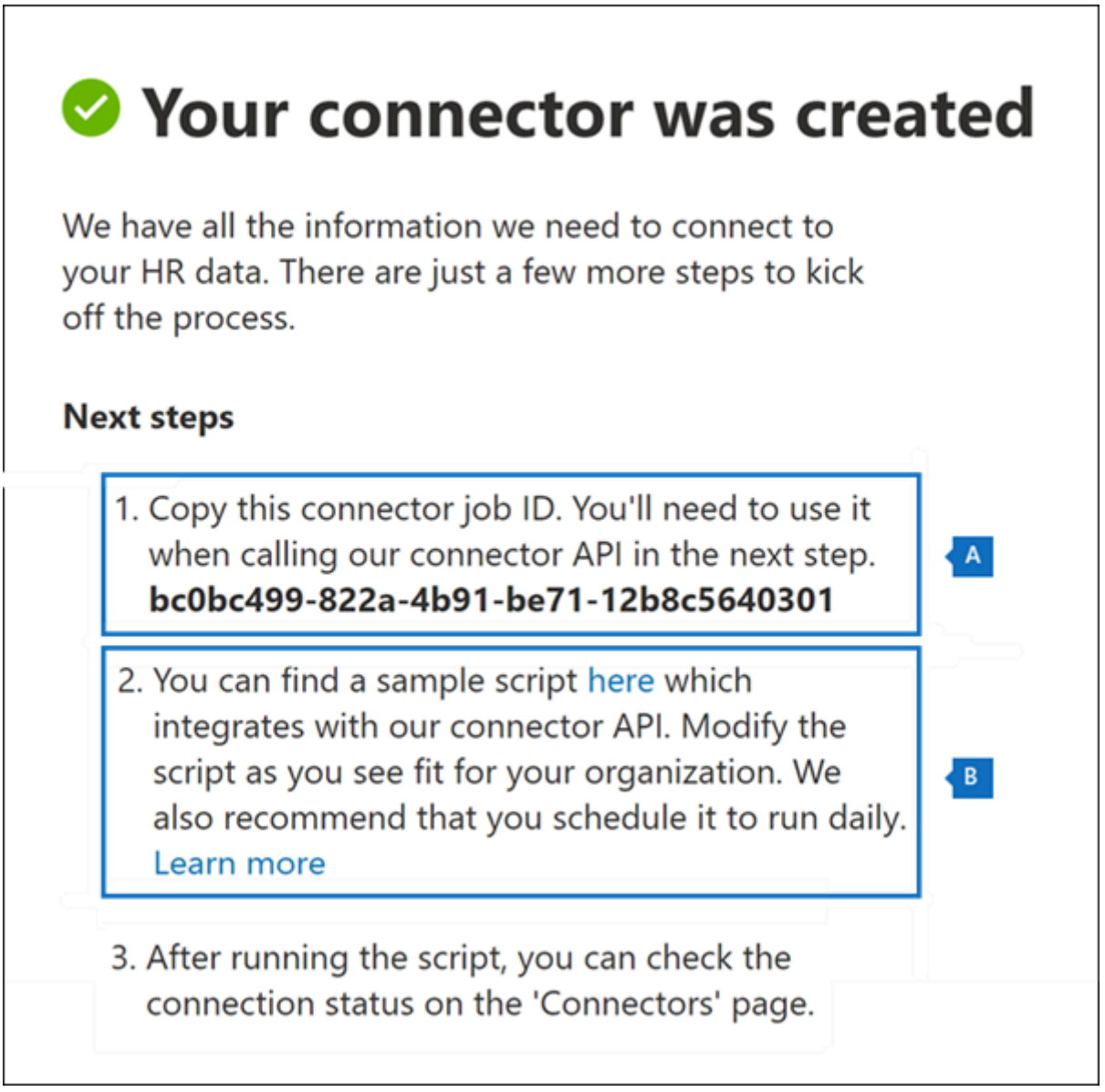
organization. Be sure to copy the job ID that's generated when you create the connector. You'll use the job ID when you run the script.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. From the list, select **HR**.
5. On the **Authentication credentials** page, enter the following information and then select **Next**:
   a. Enter or paste the Microsoft Entra application ID for the Azure app that you created in Step 1.
   b. Enter a name for the HR connector.
6. On the **File mapping** page, enter the names of the three column headers (also called *parameters*) from the CSV file that you created in Step 2 in each of the appropriate boxes. The names aren't case-sensitive. As previously explained, the names that you enter in these boxes must match the parameter names in your CSV file. For example, the following screenshot shows the parameter names from the example in sample CSV file shown in Step 2.



7. On the **Review** page, review your settings and then select **Finish** to create the connector.

   A status page is displayed that confirms the connector was created. This page contains two important things that you need to complete the next step to run the sample script to upload your HR data.



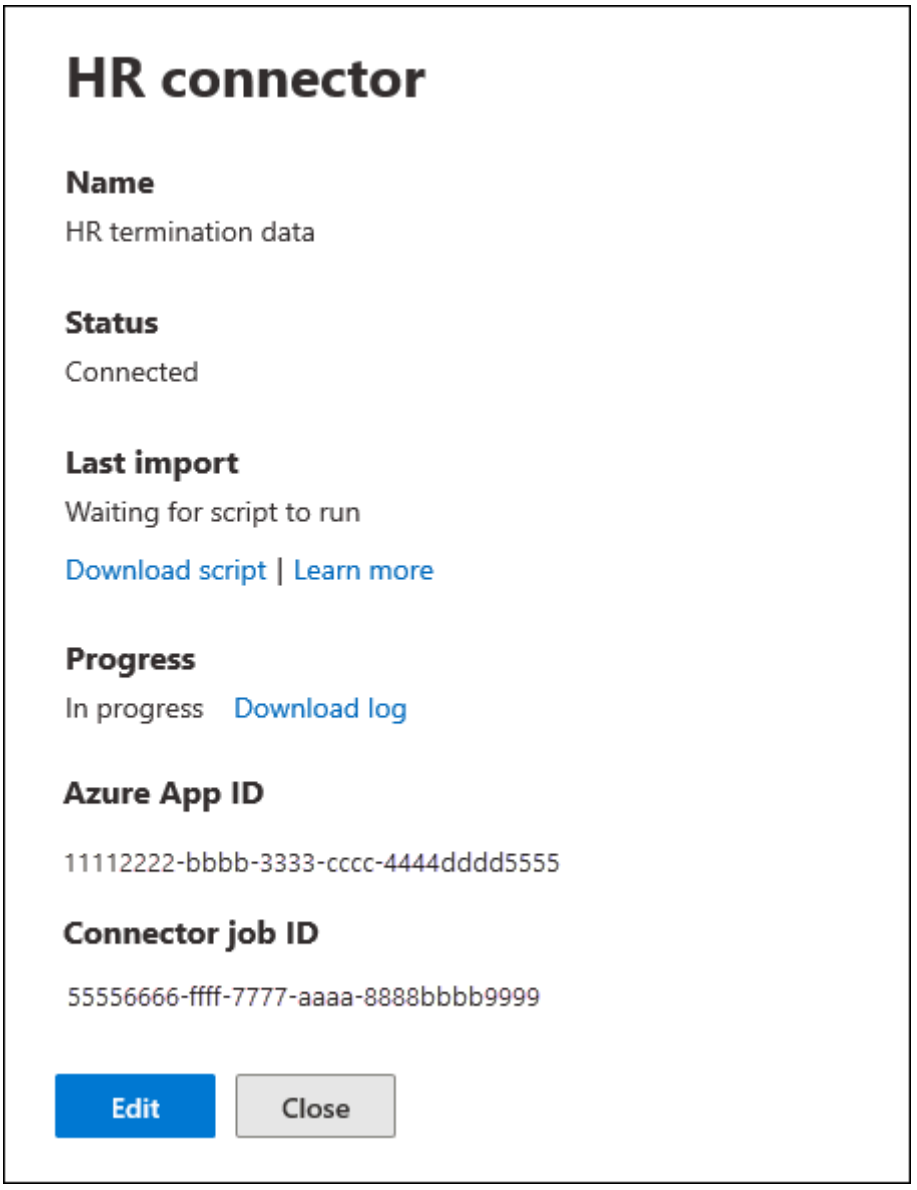
a. **Job ID.** You need this job ID to run the script. You can copy it from this page or from the connector flyout page.

b. **Link to sample script.** Select the **here** link to go to the GitHub site to access the sample script (the link opens a new window). Keep this window open so that you can copy the script in Step 4. Alternatively, you can bookmark the destination or copy the URL so you can access it again in Step 4. This link is also available on the connector flyout page.

8. Select **Done**.

   The new connector is displayed in the list on the **Connectors** tab.

9. Select the HR connector that you just created to display the flyout page, which contains properties and other information about the connector.



If you didn't already, copy the values for the **Azure App ID** and **Connector job ID**. You need these values to run the script in the next step. You can also download the script from the flyout page (or download it using the link in the next step).

Select **Edit** to change the Azure App ID or the column header names that you defined on the **File mapping** page.

Step 4: Run the sample script to upload your HR data

To set up an HR connector, run a sample script that uploads the HR data in the CSV file (created in Step 2) to the Microsoft cloud. The script specifically uploads the data to the HR connector. After you run the script, the HR connector that you created in Step 3 imports the HR data to your Microsoft 365 organization where other compliance tools, such as the Insider risk management solution, can access it. Consider scheduling a task to run the script automatically on a daily basis so the most current employee termination data is uploaded to the Microsoft cloud. For more information, see Schedule the script to run automatically.

1. Go to the window you left open from the previous step to access the GitHub site with the sample script. Alternatively, open the bookmarked site or use the URL that you copied.
2. Select the **Raw** button to display the script in text view.
3. Copy all the lines in the sample script and save them to a text file.
4. Modify the sample script for your organization, if necessary.
5. Save the text file as a Windows PowerShell script file by using a filename suffix of `.ps1`; for example, `HRConnector.ps1`.
6. Open a Command Prompt on your local computer, and go to the directory where you saved the script.
7. Run the following command to upload the HR data in the CSV file to the Microsoft cloud; for example:

PowerShell

```
.\HRConnector.ps1 -tenantId <tenantId> -appId <appId>  -appSecret <appSecret>  -jobId <jobId>  -csvFilePath '<csvFilePath>'
```

The following table describes the parameters to use with this script and their required values. Use the information you obtained in the previous steps for these values.

Expand table

Parameter	Description
`tenantId`	The Id for your Microsoft 365 organization that you obtained in Step 1. You can also get the tenant Id for your organization on the **Overview** blade in the Microsoft Entra admin center. This value identifies your organization.
`appId`	The Microsoft Entra application Id for the app that you created in Microsoft Entra ID in Step 1. Microsoft Entra ID uses this value for authentication when the script attempts to access your Microsoft 365 organization.
`appSecret`	The Microsoft Entra application secret for the app that you created in Microsoft Entra ID in Step 1. This value is also used for authentication.
`jobId`	The job ID for the HR connector that you created in Step 3. This value associates the HR data that you upload to the Microsoft cloud with the HR connector.
`csvFilePath`	The file path for the CSV file (stored on the same system as the script) that you created in Step 2. Try to avoid spaces in the file path; otherwise use single quotation marks.

Parameter	Description

Here's an example of the syntax for the HR connector script using actual values for each parameter:

PowerShell

```
 .\HRConnector.ps1 -tenantId aaaabbbb-0000-cccc-1111-dddd2222eeee -appId
00001111-aaaa-2222-bbbb-3333cccc4444 -appSecret Aa1Bb~2Cc3.-
Dd4Ee5Ff6Gg7Hh8Ii9_Jj0Kk1Ll2 -jobId 00001111-aaaa-2222-bbbb-3333cccc4444 -
csvFilePath 'C:\Users\contosoadmin\Desktop\Data\employee_termination_data.csv'
```

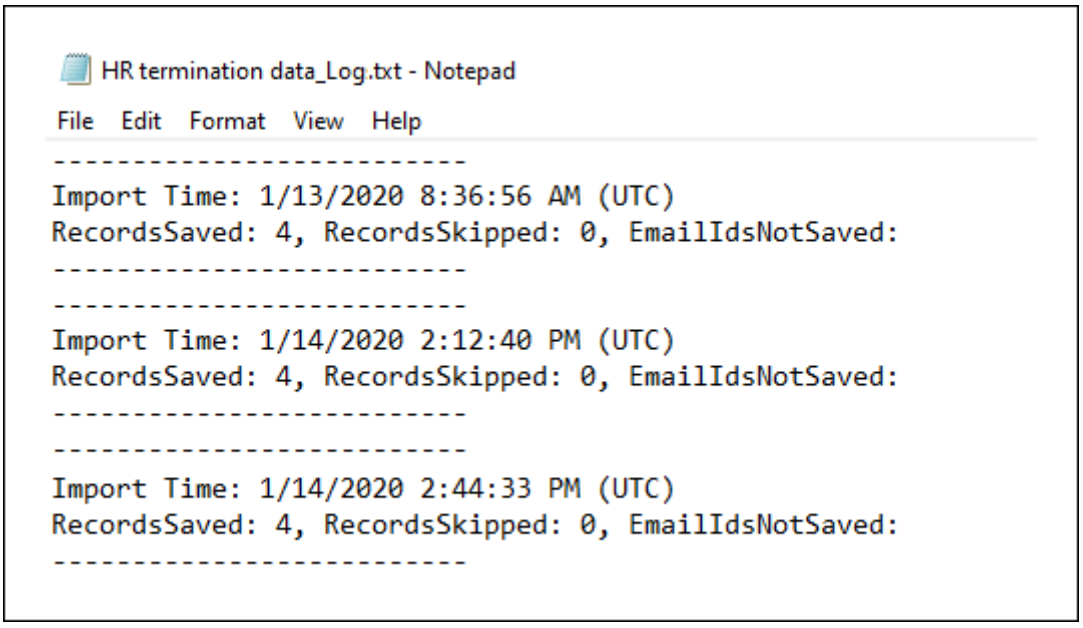
If the upload is successful, the script displays the **Upload Successful** message.

> **Note**
>
> If you have problems running the previous command because of execution policies, see About Execution Policies and Set-ExecutionPolicy for guidance about setting execution policies.

Step 5: Monitor the HR connector

After you create the HR connector and run the script to upload your HR data, you can view the connector and upload status in the Microsoft Purview portal. If you schedule the script to run automatically on a regular basis, you can also view the current status after the last time the script ran.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select the HR connector that you created to display the flyout page. This page contains the properties and information about the connector.
4. Under **Progress**, select the **Download log** link to open (or save) the status log for the connector. This log contains information about each time the script runs and uploads the data from the CSV file to the Microsoft cloud.



The `RecordsSaved` field indicates the number of rows in the CSV file that uploaded. For example, if the CSV file contains four rows, the value of the `RecordsSaved` fields is 4, if the script successfully uploaded all the rows in the CSV file.

If you didn't run the script in Step 4, a link to download the script is displayed under **Last import**. You can download the script and then follow the steps in Step 4 to run it.

(Optional) Step 6: Schedule the script to run automatically

To make sure tools like the insider risk management solution always have the latest HR data from your organization, schedule the script to run automatically on a recurring basis, such as once a day. This schedule requires that you update the HR data in the CSV file on a similar (if not the same) schedule so that it contains the latest information about employees who leave your organization. The goal is to upload the most current HR data so that the HR connector can make it available to the insider risk management solution.

Use the Task Scheduler app in Windows to automatically run the script every day.

1. On your local computer, select the Windows **Start** button and then type **Task Scheduler**.
2. Select the **Task Scheduler** app to open it.
3. In the **Actions** section, select **Create Task**.
4. On the **General** tab, type a descriptive name for the scheduled task, such as **HR Connector Script**. You can also add an optional description.

5. Under **Security options**, complete the following steps:

   a. Decide whether to run the script only when you're logged on to the computer or run it when you're logged on or not.

   b. Make sure that the **Run with the highest privileges** checkbox is selected.

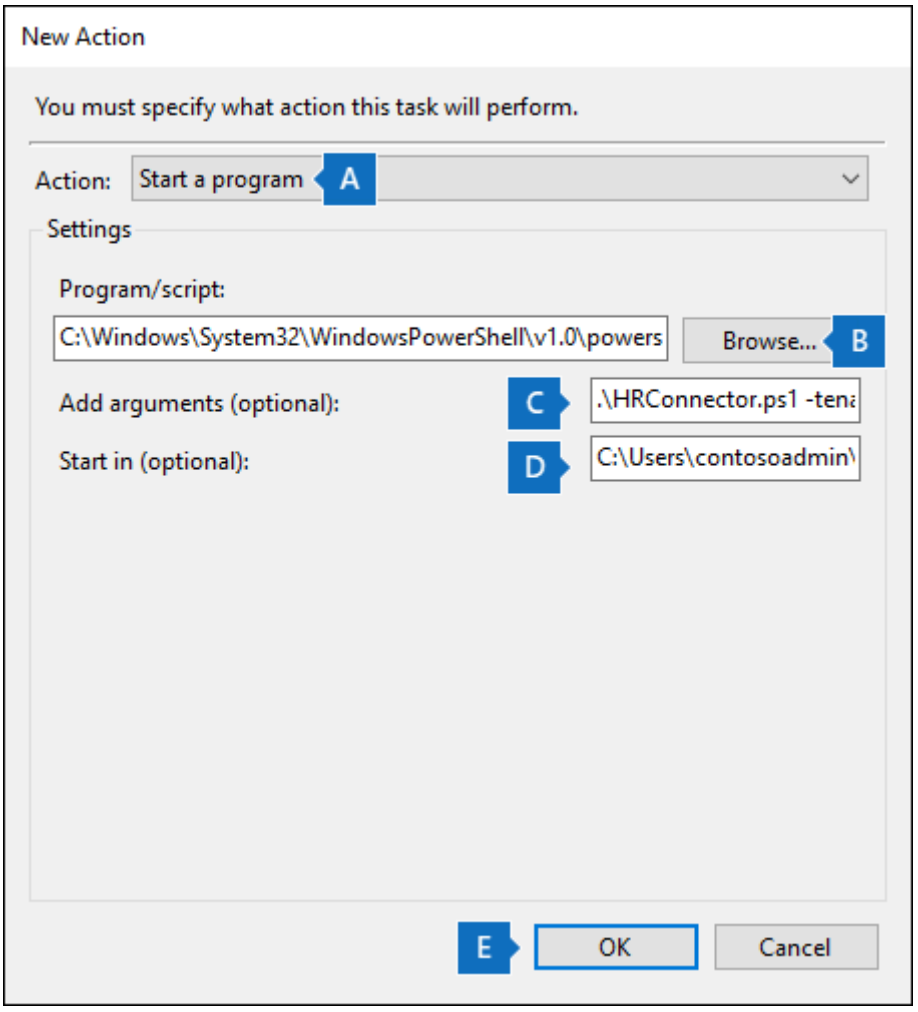
6. Select the **Triggers** tab, select **New**, and then complete the following steps:

   a. Under **Settings**, select the **Daily** option, and then select a date and time to run the script for the first time. The script runs every day at the same specified time.

   b. Under **Advanced settings**, make sure the **Enabled** checkbox is selected.

   c. Select **Ok**.

7. Select the **Actions** tab, select **New**, and then complete the following steps:



   a. In the **Action** dropdown list, make sure that **Start a program** is selected.

   b. In the **Program/script** box, select **Browse**, and go to the following location and select it so the path is displayed in the box:

`C:\Windows\System32\WindowsPowerShell\v1.0\powershell.exe`.

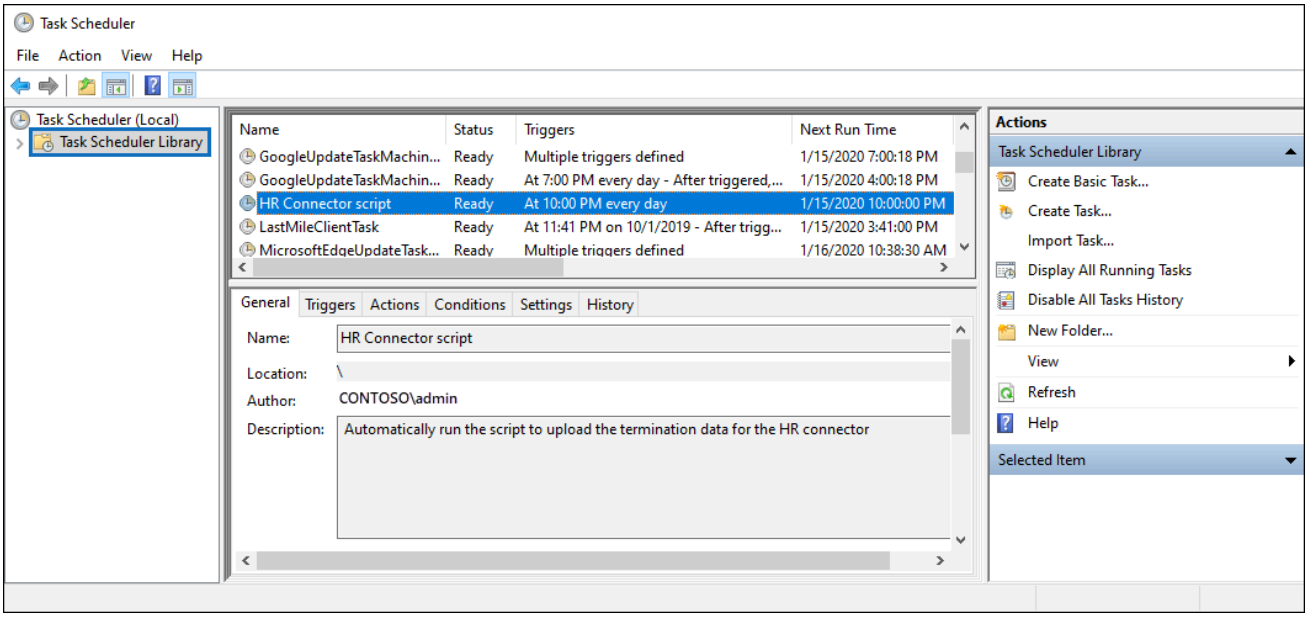
c. In the **Add arguments (optional)** box, paste the same script command that you ran in Step 4. For example, `.\HRConnector.ps1 -tenantId "aaaabbbb-0000-cccc-1111-dddd2222eeee" -appId "00001111-aaaa-2222-bbbb-3333cccc4444" -appSecret "Aa1Bb~2Cc3.-Dd4Ee5Ff6Gg7Hh8Ii9_Jj0Kk1Ll2" -jobId "00001111-aaaa-2222-bbbb-3333cccc4444" -csvFilePath "C:\Users\contosoadmin\Desktop\Data\employee_termination_data.csv"`

d. In the **Start in (optional)** box, paste the folder location of the script that you ran in Step 4. For example, `C:\Users\contosoadmin\Desktop\Scripts`.

e. Select **Ok** to save the settings for the new action.

8. In the **Create Task** window, select **Ok** to save the scheduled task. You might be prompted to enter your user account credentials.

   The new task is displayed in the Task Scheduler Library.



   The last time the script ran and the next time it's scheduled to run is displayed. You can double-select the task to edit it.

   You can also verify the last time the script ran on the flyout page of the corresponding HR connector in the Microsoft Purview portal.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 03/30/2026

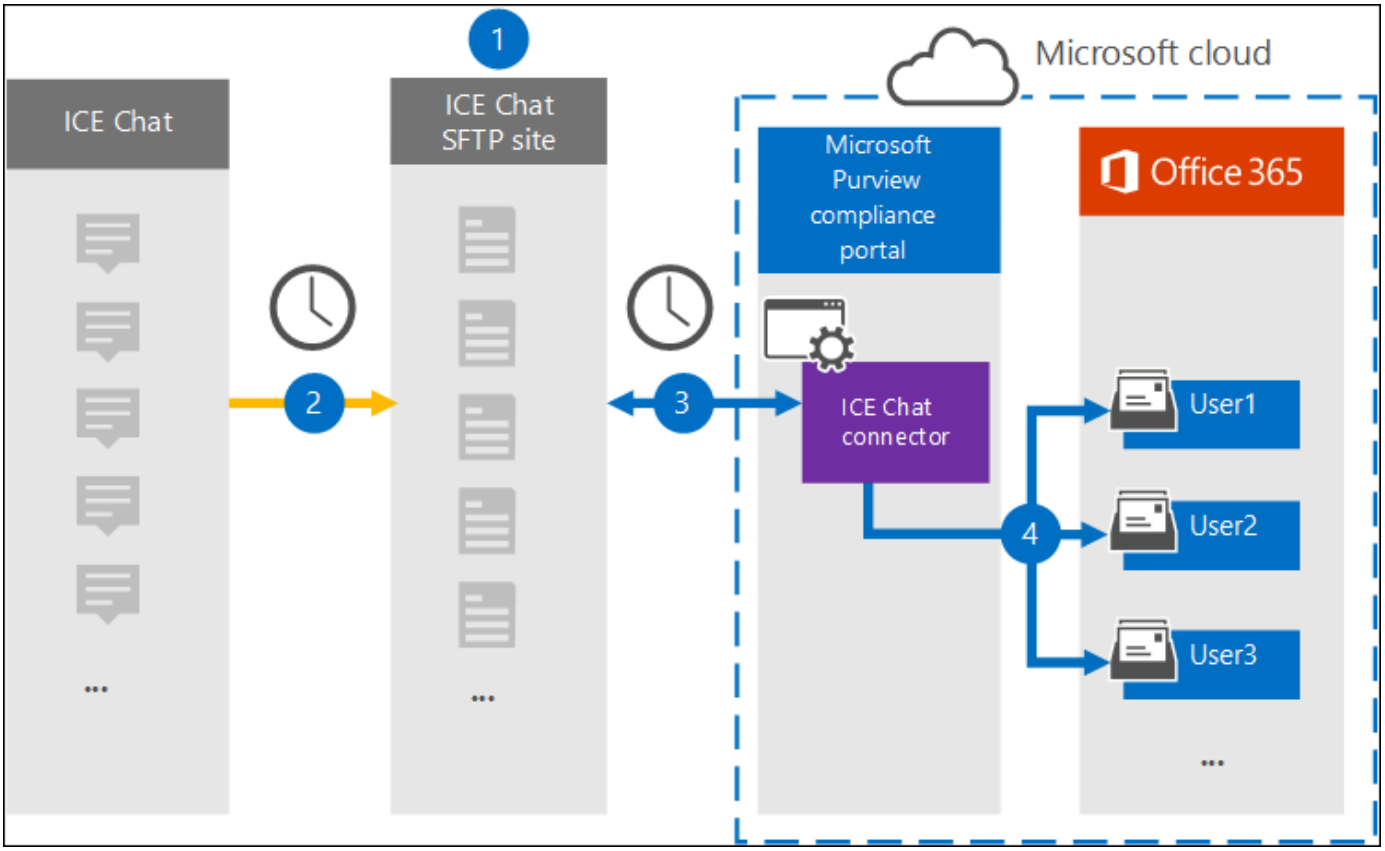
Set up a connector to archive ICE Chat data

Use a native connector to import and archive financial services chat data from the ICE Chat collaboration tool. After you set up and configure a connector, it connects to your organization's ICE Chat secure FTP (SFTP) site once every day, converts the content of chat messages to an email message format, and then imports those items to mailboxes in Microsoft 365.

After storing ICE Chat data in user mailboxes, you can apply Microsoft Purview features such as Litigation Hold, eDiscovery, In-Place Archiving, Auditing, Communication Compliance, and Microsoft 365 retention policies to ICE Chat data. For example, you can search ICE Chat messages by using eDiscovery or associate the mailbox that contains the ICE Chat data in an eDiscovery case. Using an ICE Chat connector to import and archive data in Microsoft 365 can help your organization stay compliant with government and regulatory policies.

Overview of archiving ICE Chat data

The following overview explains the process of using a connector to archive ICE chat data in Microsoft 365.



1. Your organization works with ICE Chat to set up an ICE Chat SFTP site. You also work with ICE Chat to configure ICE Chat to copy chat messages to your ICE Chat SFTP site.

2. Once every 24 hours, ICE Chat copies chat messages to your ICE Chat SFTP site.

3. The ICE Chat connector that you create in the Microsoft Purview portal connects to the ICE Chat SFTP site every day and transfers the chat messages from the previous 24 hours to a secure Azure Storage location in the Microsoft cloud. The connector also converts the content of a chat massage to an email message format.

4. The connector imports chat message items to the mailboxes of specific users. A new folder named **ICE Chat** is created in the user mailboxes and the chat message items are imported to that folder. The connector uses the value of the *SenderEmail* and *RecipientEmail* properties. Every chat message contains these properties, which are populated with email address of the sender and every recipient/participant of the chat message.

   In addition to automatic user mapping that uses the values of the *SenderEmail* and *RecipientEmail* property (which means that the connector imports a chat message to the sender's mailbox and the mailboxes of every recipient), you can also define custom user mapping by uploading a CSV mapping file. This mapping file contains the ICE Chat *ImId* and the corresponding Microsoft 365 mailbox address for every user in your organization. If you enable automatic user mapping and provide a custom-mapping file, for every chat item the connector first looks at the custom-mapping file. If it doesn't find a valid Microsoft 365 user account that corresponds to a user's ICE Chat ImId, the connector uses the *SenderEmail* and *RecipientEmail* properties of the chat item to import the item to the mailboxes of the chat participants. If the connector doesn't find a valid Microsoft 365 user in either the custom-mapping file or the *SenderEmail* and *RecipientEmail* properties, the item won't be imported.

Before you begin

Some implementation steps for archiving ICE Chat data are external to Microsoft 365. You must complete these steps before you create the connector in the Microsoft Purview portal.

- ICE Chat charges customers a fee for external compliance. Your organization should contact the ICE Chat sales group to discuss and sign the ICE Chat data services agreement, which you can get at ICE Data Services Agreement (PDF). This agreement is between ICE Chat and your organization and doesn't involve Microsoft. After you set up an ICE Chat SFTP site (see Configure the ICE Chat SFTP site), ICE Chat provides the FTP credentials directly to your organization. Then you provide those credentials to Microsoft when you create the connector (see Create an ICE Chat connector).

- You must set up an ICE Chat SFTP site before creating the connector (see Create an ICE Chat connector). After working with ICE Chat to set up the SFTP site, data from ICE Chat is uploaded to the SFTP site every day. The connector you create in the Create an ICE Chat connector step connects to this SFTP site and transfers the chat data to Microsoft 365 mailboxes. SFTP also encrypts the ICE Chat data that's sent to mailboxes during the transfer process.

- To set up an ICE Chat connector, you need keys and key passphrases for Pretty Good Privacy (PGP) and Secure Shell (SSH). These keys configure the ICE Chat SFTP site and help the connector connect to the ICE Chat SFTP site to import data to Microsoft 365. The PGP key configures the encryption of data that's transferred from the ICE Chat SFTP site to Microsoft 365. The SSH key configures secure shell to enable a secure remote authentication when the connector connects to the ICE Chat SFTP site.

  When setting up a connector, you can use public keys and key passphrases provided by Microsoft or your own private keys and passphrases. Use the public keys provided by Microsoft. However, if your organization already configured an ICE Chat SFTP site by using private keys, you can create a connector by using these same private keys.

- The ICE Chat connector can import up to 200,000 items in a single day. If there are more than 200,000 items on the SFTP site, none of those items are imported to Microsoft 365.

- The admin who creates the ICE Chat connector in Step 3 (and who downloads the public keys and IP address in Step 1) must be assigned the Data Connector Admin role. This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions, see Permissions in the Microsoft Purview portal and Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview.

Set up a connector by using public keys

The steps in this section show you how to set up an ICE Chat connector by using the public keys for Pretty Good Privacy (PGP) and Secure Shell (SSH).

Step 1: Obtain PGP and SSH public keys

First, get a copy of the public keys for Pretty Good Privacy (PGP) and Secure Shell (SSH). You use these keys in Step 2 to configure the ICE Chat SFTP site to allow the connector (that you create in Step 3) to connect to the SFTP site and transfer the ICE Chat data to Microsoft 365 mailboxes. You also get an IP address during the public-key download process, which you use when configuring the ICE Chat SFTP site.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. Select **ICE Chat** from the list.
5. On the **Terms of service** page, select **Accept**.
6. On the **Add credentials for content source** page, select **I want to use PGP and SSH public keys provided by Microsoft**.
7. Under step 1, select the **Download SSH key**, **Download PGP key**, and **Download IP address** links to save a copy of each file to your local computer.

   These files contain the following items that you use to configure the ICE Chat SFTP site in Step 2:

   - PGP public key: Use this key to configure the encryption of data that's transferred from the ICE Chat SFTP site to Microsoft 365.
   - SSH public key: Use this key to configure Secure SSH to enable a secure remote authentication when the connector connects to the ICE Chat SFTP site.
   - IP address: Configure the ICE Chat SFTP site to accept a connection request only from this IP address, which is used by the ICE Chat connector that you create in Step 3.
8. Select **Cancel** to close the workflow. You return to this workflow in the Create an ICE Chat connector section to create the connector.

Step 2: Configure the ICE Chat SFTP site

Next, use the PGP and SSH public keys and the IP address that you obtained in the Obtain PGP and SSH public keys step to configure PGP encryption and SSH authentication for the ICE Chat SFTP site. This configuration lets the ICE Chat connector that you create in the Create an ICE Chat connector section connect to the ICE Chat SFTP site and transfer ICE Chat data to Microsoft 365. You need to work with ICE Chat customer support to set up your ICE Chat SFTP site.

Step 3: Create an ICE Chat connector

In the last step, you create an ICE Chat connector in the Microsoft Purview portal. The connector uses the information you provide to connect to the ICE Chat SFTP site and transfer chat messages to the corresponding user mailboxes in Microsoft 365.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. Select **ICE Chat** from the list.
5. On the **Terms of service** page, select **Accept**.
6. On the **Add credentials for content source** page, select **I want to use PGP and SSH public keys**.
7. Under Step 3, enter the required information in the following boxes and then select **Validate connection**.
    - **Firm code:** The ID for your organization, which you use as the username for the ICE Chat SFTP site.
    - **Password:** The password for your ICE Chat SFTP site.
    - **SFTP URL:** The URL for the ICE Chat SFTP site (for example, `sftp.theice.com`). You can also use an IP address for this value.
    - **SFTP port:** The port number for the ICE Chat SFTP site. The connector uses this port to connect to the SFTP site.
8. After the connection is successfully validated, select **Next**.
9. On the **Define user** page, specify the users to import data for.
    - **All users in your organization**. Select this option to import data for all users.
    - **Only users on Litigation hold**. Select this option to import data only for users whose mailboxes are placed on Litigation hold. This option imports data to user mailboxes that have the LitigationHoldEnabled property set to True. For more information, see Create a Litigation hold.
10. On the **Map external users to Microsoft 365 users** page, enable automatic user mapping and provide custom user mapping as required. You can download a copy of the user-

mapping CSV file on this page. You can add the user mappings to the file and then upload it.

> **Note**
>
> The custom mapping file contains the ICE Chat imid and corresponding Microsoft 365 mailbox address for each user. If you enable automatic user mapping and provide a custom mapping, for every chat item, the connector first looks at the custom mapping file. If it doesn't find a valid Microsoft 365 user that corresponds to a user's ICE Chat imid, the connector imports the item to the mailboxes for the users specified in the *SenderEmail* and *RecipientEmail* properties of the chat item. If the connector doesn't find a valid Microsoft 365 user by either automatic or custom user mapping, the item isn't imported.

11. Select **Next**, review your settings, and then select **Finish** to create the connector.

Set up a connector by using private keys

The steps in this section show you how to set up an ICE Chat connector by using PGP and SSH private keys. Use this connector setup option if your organization already configured an ICE Chat SFTP site by using private keys.

Step 1: Get an IP address to configure the ICE Chat SFTP site

If your organization uses PGP and SSH private keys to set up an ICE Chat SFTP site, get an IP address and give it to ICE Chat customer support. The ICE Chat SFTP site must accept connection requests from this IP address. The IP address you download in this step is also used by the ICE Chat connector to connect to the SFTP site and transfer ICE Chat data to Microsoft 365.

To get the IP address:

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. Select **ICE Chat** from the list.
5. On the **Terms of service** page, select **Accept**.
6. On the **Add credentials for content source** page, select **I want to use PGP and SSH private keys**.

7. Under step 1, select **Download IP address** to save a copy of the IP address file to your local computer.
8. Select **Cancel** to close the workflow. You return to this workflow in the Create an ICE Chat connector section to create the connector.

Work with ICE Chat customer support to configure your ICE Chat SFTP site to accept connection requests from this IP address.

Step 2: Create an ICE Chat connector

After you configure your ICE Chat SFTP site, create an ICE Chat connector in the Microsoft Purview portal. The connector uses the information you provide to connect to the ICE Chat SFTP site and transfer email messages to the corresponding user mailboxes in Microsoft 365. To complete this step, make sure you have copies of the same private keys and key passphrases that you used to set up your ICE Chat SFTP site.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. Select **ICE Chat** from the list.
5. On the **Terms of service** page, select **Accept**.
6. On the **Add credentials for content source** page, select **I want to use PGP and SSH private keys**.
7. Under Step 3, enter the required information in the following boxes and then select **Validate connection**.
    - **Name:** The name for the connector. It must be unique in your organization.
    - **Firm code:** The ID for your organization that is used as the username for the ICE Chat SFTP site.
    - **Password:** The password for your organization's ICE Chat SFTP site.
    - **SFTP URL:** The URL for the ICE Chat SFTP site (for example, `sftp.theice.com`). You can also use an IP address for this value.
    - **SFTP port:** The port number for the ICE Chat SFTP site. The connector uses this port to connect to the SFTP site.

- **PGP private key:** The PGP private key for the ICE Chat SFTP site. Include the entire private key value, including the beginning and ending lines of the key block.
- **PGP key passphrase:** The passphrase for the PGP private key.
- **SSH private key:** The SSH private key for the ICE Chat SFTP site. Include the entire private key value, including the beginning and ending lines of the key block.
- **SSH key passphrase:** The passphrase for the SSH private key.

8. After the connection is successfully validated, select **Next**.

9. On the **Define user** page, specify the users to import data for.

   - **All users in your organization**. Select this option to import data for all users.
   - **Only users on Litigation hold**. Select this option to import data only for users whose mailboxes are placed on Litigation hold. This option imports data to user mailboxes that have the LitigationHoldEnabled property set to True. For more information, see Create a Litigation hold.

10. On the **Map ICE Chat users to Microsoft 365 users** page, enable automatic user mapping and provide custom user mapping as required.

    > ⓘ Note
    >
    > The custom mapping file contains the ICE Chat imid and corresponding Microsoft 365 mailbox address for each user. If you enable automatic user mapping and provide a custom mapping, for every chat item, the connector first looks at the custom mapping file. If it doesn't find a valid Microsoft 365 user that corresponds to a user's ICE Chat imid, the connector imports the item to the mailboxes for the users specified in the *SenderEmail* and *RecipientEmail* properties of the chat item. If the connector doesn't find a valid Microsoft 365 user by either automatic or custom user mapping, the item isn't imported.

11. Select **Next**, review your settings, and then select **Finish** to create the connector.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/24/2026

Set up a connector to import third-party insider risk detections (preview)

You can set up a connector to extend the Microsoft Purview Insider Risk Management solution to include third-party (non-Microsoft) detections. For example, you might want to extend your detections to include Salesforce and Dropbox activities and use them alongside the built-in detections provided by insider risk management, which focuses on Microsoft services like SharePoint Online and Exchange Online.

To bring your own detections to the insider risk management solution, import preprocessed, aggregated detections from security information and event management (SIEM) solutions such as Microsoft Sentinel or Splunk. Import a sample file into the Insider Risk Indicators connector workflow. The connector workflow analyzes the sample file and configures the required schema for insider risk management.

> ⓘ **Note**
>
> Currently, you can't import "raw" detection signals into insider risk management. You can only import preprocessed aggregations as a file.

Overall process

Bringing your own detections to insider risk management is a three-step process:

1. In Microsoft Purview, create the Insider Risk Indicators (preview) connector as described in this article.
2. In the insider risk management solution, create custom indicators.
3. In the insider risk management solution, use the custom indicators in policies as triggers or indicators and define thresholds.

When user activity crosses the threshold value that you specify for the policy, the user is brought into scope of the insider risk management policy and is scored for risk. An alert is generated and analysts can investigate the alert using custom indicator details.

> ⓘ **Note**
>
> You can only use custom indicators with the *Data theft* and *Data leaks* templates.

Before you begin

- **Determine the scenarios and data you want to import to Microsoft 365**. This determination helps you decide how many CSV files and Insider Risk Indicator connectors you need to create and how to structure the CSV files. The types of triggers and indicators you want to create determine the imported data. See Determining how many CSV files to prepare for indicator data.
- **Determine how to retrieve or export the data from your internal system and add it to the CSV files that you prepare in Step 2**. The script that you run in Step 4 uploads the data in the CSV files to the insider risk management solution.
- **Assign the *Data Connector Admin* role** to the user who creates the connector in Step 3. This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions, see:- Permissions in the Microsoft Purview portal - Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview compliance

> ⓘ **Important**
>
> The sample script that you run in Step 4 uploads your data to the Microsoft cloud so that the insider risk management solution can use it. This sample script isn't supported under any Microsoft standard support program or service. The sample script is provided AS IS without warranty of any kind. Microsoft further disclaims all implied warranties including, without limitation, any implied warranties of merchantability or of fitness for a particular purpose. You assume all risk arising from the use or performance of the sample script and documentation. In no event shall Microsoft, its authors, or anyone else involved in the creation, production, or delivery of the scripts be liable for any damages whatsoever (including, without limitation, damages for loss of business profits, business interruption, loss of business information, or other pecuniary loss) arising out of the use of or inability to use the sample scripts or documentation, even if Microsoft has been advised of the possibility of such damages.

Determining how many CSV files to prepare for indicator data

In Step 3, you can prepare separate CSV files that contain data for each indicator or prepare a single CSV file that contains data for two or more indicators.

Here are some guidelines to help you determine how many CSV files to prepare:

- If the insider risk management policy you want to implement requires multiple indicators, consider using a single CSV file that contains the data for all the indicators. As a general

rule, the number of connectors that you need to create is determined by the services in a CSV file. For example, if a CSV file contains all the services required to support your insider risk management implementation, you only need one connector. Having fewer CSV files allows you to have fewer connectors to create and manage. If you have two separate CSV files that each contain a single service, you need to create two connectors.

- The method for generating or collecting the data might determine the number of CSV files. For example, if the different types of data used to configure a connector are located in a single system in your organization, you might be able to export the data to a single CSV file. But if data is distributed across different systems, it might be easier to export data to different CSV files. How you retrieve or export data from your systems might determine the number of CSV files you need.

Step 1: Create an app in Microsoft Entra ID

First, create and register a new app in Microsoft Entra ID for the connector that you create in Step 3. When you create this app, Microsoft Entra ID can authenticate the connector when it runs and attempts to access your organization. Use this app to authenticate the script that you run in Step 4 to upload your data to the Microsoft cloud. When you create the Microsoft Entra app, save the following information:

- Microsoft Entra application ID (app ID or client ID)
- Microsoft Entra application secret (client secret)
- Tenant ID (directory ID)

Use these values in Steps 3 and 4. For step-by-step instructions on creating an app in Microsoft Entra ID, see Register an application with the Microsoft identity platform.

Step 2: Prepare CSV file(s) with your insider risk indicators data

Next, prepare a CSV file that contains the indicator data that the connector imports to Microsoft 365. The insider risk management solution uses this data. You can import data for the following scenarios:

- Create a trigger that, when activated, brings a user into scope for a policy. Example 1 below shows how to prepare a CSV file for a 'home-grown' trigger that predicts the probability of an employee leaving an organization.
- Create a policy indicator that monitors user activities. Example 2 below shows how to prepare a single CSV file for multiple indicators (one for Dropbox and one for Salesforce).

For each scenario, provide the corresponding indicator data in one or more CSV files. See Determining how many CSV files to use for indicator data.

After you create the CSV file with the required indicator data, store it on the local computer that you run the script on in Step 4. Implement an update strategy to make sure that the CSV file always contains the most current information so that whenever you run the script, the most current indicator data is uploaded to the Microsoft cloud and accessible to the insider risk management solution.

> **Important**
>
> The column names described in the following sections are examples, not required parameters. You can use any column names in your CSV files. However, the column names that you use in a CSV file must be mapped to a data type when you create the connector in Step 3. Also note that the sample CSV files in the following sections are shown in NotePad. It's much easier to view and edit CSV files in Microsoft Excel.

Example 1: Prepare a CSV file for a simple trigger that brings a user into scope for a policy

This example shows how to structure a CSV file to create a 'home-grown' trigger that could be used to predict the probability of an employee leaving an organization. This example uses the following sample data:

text

```
UserPrincipalName,PredictionTime,PredictionScore,ModelInfo
sarad@contoso.com,2023-04-20T05:52:56.962686Z,6,Model accuracy: 67%, Model name:
LeaverPrediction_M1
sarad@contoso.com,2023-04-24T05:52:56.962686Z,9,Model accuracy: 67%, Model name:
LeaverPrediction_M1
sarad@contoso.com,2023-04-24T05:52:56.962686Z,3,Model accuracy: 67%, Model name:
LeaverPrediction_M1
```

The following table describes each column in the CSV file.

Expand table

Column	Description
UserPrincipalName	The Microsoft Entra UserPrincipalName (UPN) used to identify the user.
Prediction Time	Mandatory field that displays the date/time that the activity occurred. Use the following date format: `yyyy-mm-ddThh:mm:ss.nnnnnn+|-hh:mm`, which is the ISO 8601

Column	Description
	date and time format
Prediction Score	Risky activity score. This field is used for the trigger threshold setting. Only *Number* fields can be used for threshold settings.
Model Info	Extra field used to track information about the prediction model.

> **Note**
>
> Only the UserPrincipalName and date/time fields are mandatory. All other fields are optional but can be helpful for the analyst or investigator in decision making when they triage alerts (these fields appear in the Activity explorer and in alerts and cases).

When you create the connector in Step 3, you'll use the data in the `PredictionScore` field as a threshold value for the trigger. If a user crosses the threshold value that you set later in the policy, the user is brought into the scope of the policy.

Example 2: Prepare a single CSV file to create multiple policy indicators

This example shows how to create multiple policy indicators (one for Dropbox and one for Salesforce) from a single CSV file. This example uses the following sample data:

text

```
User_Principal_Name,Display_Name,Alert_Severity,Alert_Count,Aggregation_Date,Source
_Workload,AdditionalInfo_Salesforce,AdditionalInfo_Dropbox
sarad@contoso.com,Salesforce - Sensitive report downloaded and emailed
externally,High,10,2023-04-24T05:52:56.962686Z,Salesforce,text,text
sarad@contoso.com,Salesforce - Anomalous download of sales lead
reports,Medium,6,2023-04-24T05:52:56.962686Z,Salesforce,text,text
bradh@contoso.com,Salesforce - Printing sales reports,Low,50,2023-04-
24T05:52:56.962686Z,Salesforce,text,text
bradh@contoso.com,Salesforce - Excessive modifications to sensitive
reports,Medium,3,2023-04-24T05:52:56.962686Z,Salesforce,text,text
sarad@contoso.com,Dropbox - Sensitive files saved to personal Dropbox,High,14,2023-
04-24T05:52:56.962686Z,Dropbox,text,text
bradh@contoso.com,Dropbox - Anomalous file copy activity,Medium,5,2023-04-
24T05:52:56.962686Z,Dropbox,text,text
```

The following table describes each column in the CSV file.

Expand table

Column	Description
UserPrincipalName	The Microsoft Entra UserPrincipalName (UPN) used to identify the user.
Display Name	Name of the risky activity.
Alert Severity	Severity categories: *Low*, *Medium*, and *High*.
Alert Count	Number of incidences of each activity. Data in this field is used for the indicator threshold setting.
Aggregation Date	Mandatory field that displays the date/time that the activity occurred. Use the following date format: `yyyy-mm-ddThh:mm:ss.nnnnnn+|-hh:mm`, which is the ISO 8601 date and time format
Source Workload	This is the key field for the multiple indicators scenario. Select this field for the **Source column** field when you create the connector, and use the values in this field (**Dropbox** and **Salesforce**) in the **Related values in source column** field in the connector.
Additional Info Salesforce	Any additional info that you want to note about the Salesforce indicator
Additional Info Dropbox	Any additional info that you want to note about the Dropbox indicator

See the following example to see how you use this CSV file when creating the data connector.

Step 3: Create the Insider Risk Indicators connector

Create a connector in the Microsoft Purview portal. After you run the script in Step 4, the connector imports the data from the CSV file and uploads it to your Microsoft 365 organization.

> **Note**
>
> Before you create a connector, make sure that you have a list of the scenarios and the corresponding CSV column names for each scenario.

Example 1: Create a connector file for a simple trigger

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.

4. From the list, choose **Insider Risk Indicators (preview)**.

5. Review the terms of service, then select **Accept** if you want to continue creating the connector.

6. On the **Authentication** page, complete the following steps:
   a. Enter a name for the connector.
   b. Paste the Microsoft Entra application ID for the Azure app that you created in Step 1.
   c. Select **Next**.

7. On the **Sample file** page:
   a. Select **Upload sample file**, then select the CSV file that you want to upload.
   b. In the **Source column** list, select **None (Single source)**.
   c. In the **Verify sample data and data type** section, review each field to make sure that the right data types are assigned to each field. If a field is used later as a threshold value, make sure that it has a *Number* data type. For example, in this scenario, the `PredictionScore` field is used as a threshold value and the data type is set appropriately to *Number*.

8. Select **Next**.

9. On the **Data mapping** page:
   a. Enter the values for **Event time (UTC time)** and **Microsoft 365 user email address** based on the appropriate values from the CSV file. These fields are mandatory for the connector.
   b. In the **Default** field, use the list to select each field you want to include from the CSV file. For example, select a Number field to use later as a threshold value for the indicator or select other fields to use as supporting information.

10. Select **Next**.

11. On the **Finish** page, review all the information, and if everything looks correct, select **Finish**.

12. Copy the job ID for the connector. You need it for the next step.

13. Go to Step 4 to run the script that uploads the data to Microsoft 365.

Example 2: Create a connector that includes multiple policy indicators

This example shows how to set up a single connector to create multiple policy indicators for Salesforce and Dropbox. You could create two separate connectors, but creating a single connector that works for both can reduce overall file maintenance.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. From the list, choose **Insider Risk Indicators (preview)**.
5. Review the terms of service, then select **Accept** if you want to continue creating the connector.
6. On the **Authentication** page, enter the following information:
   a. Enter a name for the connector.
   b. Paste the Microsoft Entra application ID for the Azure app that you created in Step 2.
   c. Select **Next**.
7. On the **Sample file** page:
   a. Select **Upload sample file**, then select the CSV file that you want to upload.
   b. In the **Source column** list, select the column to use as the source. In the example CSV file, the source column is `SourceWorkload`, since it stores the values for the two separate workloads (Salesforce and Dropbox).
   c. In the **Related values in source column** field, enter the related values. For this example, enter `Salesforce,Dropbox`. Don't include spaces between values.

   > **Important**
   >
   > Make sure that the values you enter in the **Related values in source column** field match the values in the **Source column** list. The connector fails if the column values don't match.

   a. In the **Verify sample data and data type** section, review each field to make sure that the right data types are assigned for each field. If a field is used later as a threshold value, make sure that it has a Number data type. For example, in this example scenario, the `AlertCount` field is used as a threshold value and the data type is set appropriately to *Number*.
   b. Select **Next**.
8. On the **Data mapping** page:
   a. Enter the values for **Event time (UTC time)** and **Microsoft 365 user email address** based on the appropriate values from the CSV file. These fields are mandatory and common to both indicators you're creating for this example.
   b. Select the columns from your sample file that you want to map to the two workloads, Salesforce and Dropbox.

> **Tip**
>
> You can use the same process described previously to create multiple policy indicators based on severity levels. For example, you can use a single connector to create separate Low, Medium, and High indicators. In the **Source column** list, select the field that holds the values for the separate workloads (Low, Medium, High). Enter those workload values in the **Related values in source column** field. Map the appropriate fields in the **Data mapping** page.

9. Select **Next**.
10. On the **Finish** page, review all the information, and if everything looks correct, select **Finish**.
11. Copy the job ID for the connector. You need it for the next step.
12. Go to the next step (Step 4) to run the script that uploads the data to Microsoft 365.

Step 4: Run the sample script to upload your data

The last step in setting up a connector is running a sample script that uploads the data in the CSV file. When you run the script, the connector you created in Step 3 imports the data to your Microsoft 365 organization where the insider risk management solution can access it. After you run the script, consider scheduling a task to run it automatically on a daily basis so the most current data is uploaded to the Microsoft cloud. For more information, see Schedule the script to run automatically.

Before running the script

- Make sure to add the *webhook.ingestion.office.com* domain to your firewall allowlist for your organization. If this domain is blocked, the script doesn't run.
- Make sure to wait 24 hours before uploading the data after you update the custom indicators and the associated policies. This waiting period ensures all components sync. If you immediately upload the data while the updates are syncing, some data might not be scored for risk.
- Make sure that all combinations of UPN and timestamp to be imported are unique. If any record in the uploaded CSV file contains the same timestamp and UPN as other records in the file, the record is dropped.

Run the sample script

1. Go to the window that you left open from the previous step to access the GitHub site with the sample script. Alternatively, open the bookmarked site or use the URL that you copied. You can also access the script at https://github.com/microsoft/m365-compliance-connector-sample-scripts/blob/main/sample_script.ps1 .
2. Select the **Raw** button to display the script in text view.
3. Copy all the lines in the sample script, and save them to a text file.
4. Modify the sample script for your organization, if necessary.
5. Save the text file as a Windows PowerShell script file by using a filename suffix of `.ps1`; for example, `HRConnector.ps1`. Alternatively, you can use the GitHub filename for the script, which is `upload_termination_records.ps1`.
6. Open a command prompt on your local computer, and then go to the directory where you saved the script.
7. Run the following command to upload the data in the CSV file to the Microsoft cloud; for example:

   PowerShell

   ```
   .\HRConnector.ps1 -tenantId <tenantId> -appId <appId>  -appSecret <appSecret>
   -jobId <jobId>  -filePath '<filePath>'
   ```

   The following table describes the parameters to use with this script and their required values. The information you obtained in the previous steps is used in the values for these parameters.

   Expand table

Parameter	Description
`tenantId`	This is the ID for your Microsoft 365 organization that you obtained in Step 1. You can also obtain the tenant ID for your organization on the **Overview** blade in the Microsoft Entra admin center. This value identifies your organization.
`appId`	This is the Microsoft Entra application ID for the app that you created in Microsoft Entra ID in Step 1. This value is used by Microsoft Entra ID for authentication when the script attempts to access your Microsoft 365 organization.
`appSecret`	This is the Microsoft Entra application secret for the app that you created in Microsoft Entra ID in Step 1. This value is also used for authentication.
`jobId`	This is the job ID for the connector that you created in Step 3. This value associates

Parameter	Description
	the data that is uploaded to the Microsoft cloud with the connector.
`filePath`	This is the file path for the file (stored on the same system as the script) that you created in Step 1. Try to avoid spaces in the file path; otherwise use single quotation marks.

Here's an example of the syntax for the connector script using actual values for each parameter:

PowerShell

```
 .\HRConnector.ps1 -tenantId aaaabbbb-0000-cccc-1111-dddd2222eeee -appId
00001111-aaaa-2222-bbbb-3333cccc4444 -appSecret Aa1Bb~2Cc3.-
Dd4Ee5Ff6Gg7Hh8Ii9_Jj0Kk1Ll2 -jobId 00001111-aaaa-2222-bbbb-3333cccc4444 -
filePath 'C:\Users\contosoadmin\Desktop\Data\insider_risk_indicator_data.csv'
```

If the upload is successful, the script displays the **Upload Successful** message.

> ⓘ Note
>
> If you have problems running the previous command because of execution policies, see About Execution Policies and Set-ExecutionPolicy for guidance about setting execution policies.

Step 5: Monitor the connector

After you create the connector and run the script to upload your data, view the connector and upload status in the Microsoft Purview portal. If you schedule the script to run automatically on a regular basis, you can view the current status after the last time the script runs.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select the HR connector that you created to display the flyout page. This page contains the properties and information about the connector.
4. Under **Progress**, select the **Download log** link to open (or save) the status log for the connector. This log contains information about each time the script runs and uploads the data from the CSV file to the Microsoft cloud.

The `RecordsSaved` field indicates the number of rows in the CSV file that uploaded. For example, if the CSV file contains four rows, the value of the `RecordsSaved` field is 4 if the script successfully uploads all the rows in the CSV file.

If you don't run the script in Step 4, a link to download the script is displayed under **Last import**. You can download the script and then follow the steps to run the script.

(Optional) Step 6: Schedule the script to run automatically

To make sure the insider risk management solution always has the latest data from your organization, schedule the script to run automatically on a recurring basis, such as once a day. This schedule requires updating the data in the CSV file on a similar (if not the same) schedule so that it contains the latest information. The goal is to upload the most current data so that the connector can make it available to the insider risk management solution.

Use the Task Scheduler app in Windows to automatically run the script every day.

1. On your local computer, select the Windows **Start** button and type **Task Scheduler**.
2. Select the **Task Scheduler** app.
3. In the **Actions** section, select **Create Task**.
4. On the **General** tab, enter a descriptive name for the scheduled task. For example, **HR Connector Script**. You can also add an optional description.
5. Under **Security options**, complete the following steps:
   a. Decide whether to run the script only when you're logged on to the computer or when you're logged on or not.
   b. Make sure that the **Run with the highest privileges** check box is selected.
6. Select the **Triggers** tab, select **New**, and complete the following steps:
   a. Under **Settings**, select the **Daily** option, and choose a date and time to run the script for the first time. The script runs every day at the same specified time.
   b. Under **Advanced settings**, make sure that the **Enabled** check box is selected.
   c. Select **OK**.
7. Select the **Actions** tab, select **New**, and complete the following steps:
   a. In the **Action** dropdown list, make sure that **Start a program** is selected.
   b. In the **Program/script** box, select **Browse**, then go to the following location and select it so the path is displayed in the box:

`C:\Windows\System32\WindowsPowerShell\v1.0\powershell.exe`.

c. In the **Add arguments (optional)** box, paste the same script command that you ran in Step 4. For example, `.\HRConnector.ps1 -tenantId "aaaabbbb-0000-cccc-1111-dddd2222eeee" -appId "00001111-aaaa-2222-bbbb-3333cccc4444" -appSecret "Aa1Bb~2Cc3.-Dd4Ee5Ff6Gg7Hh8Ii9_Jj0Kk1Ll2" -jobId "00001111-aaaa-2222-bbbb-3333cccc4444" -filePath "C:\Users\contosoadmin\Desktop\Data\insider_risk_indicator_data.csv"`

d. In the **Start in (optional)** box, paste the folder location of the script that you ran in Step 4. For example, `C:\Users\contosoadmin\Desktop\Scripts`.

e. Select **OK** to save the settings for the new action.

8. In the **Create Task** window, select **OK** to save the scheduled task. You might be prompted to enter your user account credentials.

   The new task appears in the Task Scheduler Library. You see the last time the script ran and the next time it's scheduled to run. Double-click the task to edit it.

   You can also verify the last time the script ran on the flyout page of the corresponding connector in the Microsoft Purview portal.

(Optional) Step 7: Upload data using Power Automate templates

You can upload the CSV data by using Power Automate templates and define triggers. For example, you can configure a Power Automate template to trigger when new connector files are available in SharePoint or OneDrive locations. You can also streamline this process by storing confidential information like the Microsoft Entra application secret (created in Step 1) in Azure Key Vault and use it with Power Automate for authentication.

Complete the following steps to automatically upload data when new files become available on OneDrive for Business:

1. Download the *ImportHRDataforIRM.zip* package from the GitHub site.
2. In Power Automate, go to **My flows**.
3. Select **Import** and upload the *ImportHRDataforIRM.zip* package.
4. After the package is uploaded, update the content (name and OneDrive for Business connection) and select **Import**.
5. Select **Open flow** and update the parameters. The following table describes the parameters to use in this Power Automate flow and their required values. The information

you obtained in the previous steps is used in the values for these parameters.

Expand table

Parameter	Description
`appId`	This is the Microsoft Entra application ID for the app that you created in Microsoft Entra ID in Step 1. This value is used by Microsoft Entra ID for authentication when the script attempts to access your Microsoft 365 organization.
`appSecret`	This is the Microsoft Entra application secret for the app that you created in Microsoft Entra ID in Step 1. This secret is used for authentication.
`fileLocation`	This is the OneDrive for Business location where Power Automate monitors for 'new file created' activities to trigger this flow.
`jobId`	Identifier for the connector created in Step 3. This identifier associates the data uploaded to the Microsoft cloud with the connector.
`tenantId`	Identifier for your Microsoft 365 organization obtained in Step 1. You can also obtain the tenant ID for your organization on the **Overview** blade in the Microsoft Entra admin center. Microsoft uses this tenant ID to identify your organization.
`URI`	Verify that the value for this parameter is *https://webhook.ingestion.office.com/api/signals*

1. Select **Save**.
2. Go to **Flow overview** and select **Turn on**.
3. Test the flow manually by uploading a new file to your OneDrive for Business folder and verify that it runs successfully. This process might take a few minutes after the upload before the flow is triggered.
4. You can now monitor the connector as described in Step 5.

If needed, you can update the flow to create triggers based on file availability and modification events on SharePoint and other data sources supported by Power Automate flows.

Next steps

- Create custom indicators in insider risk management
- Use the custom indicators in policies as triggers or indicators and define thresholds

Last updated on 03/30/2026

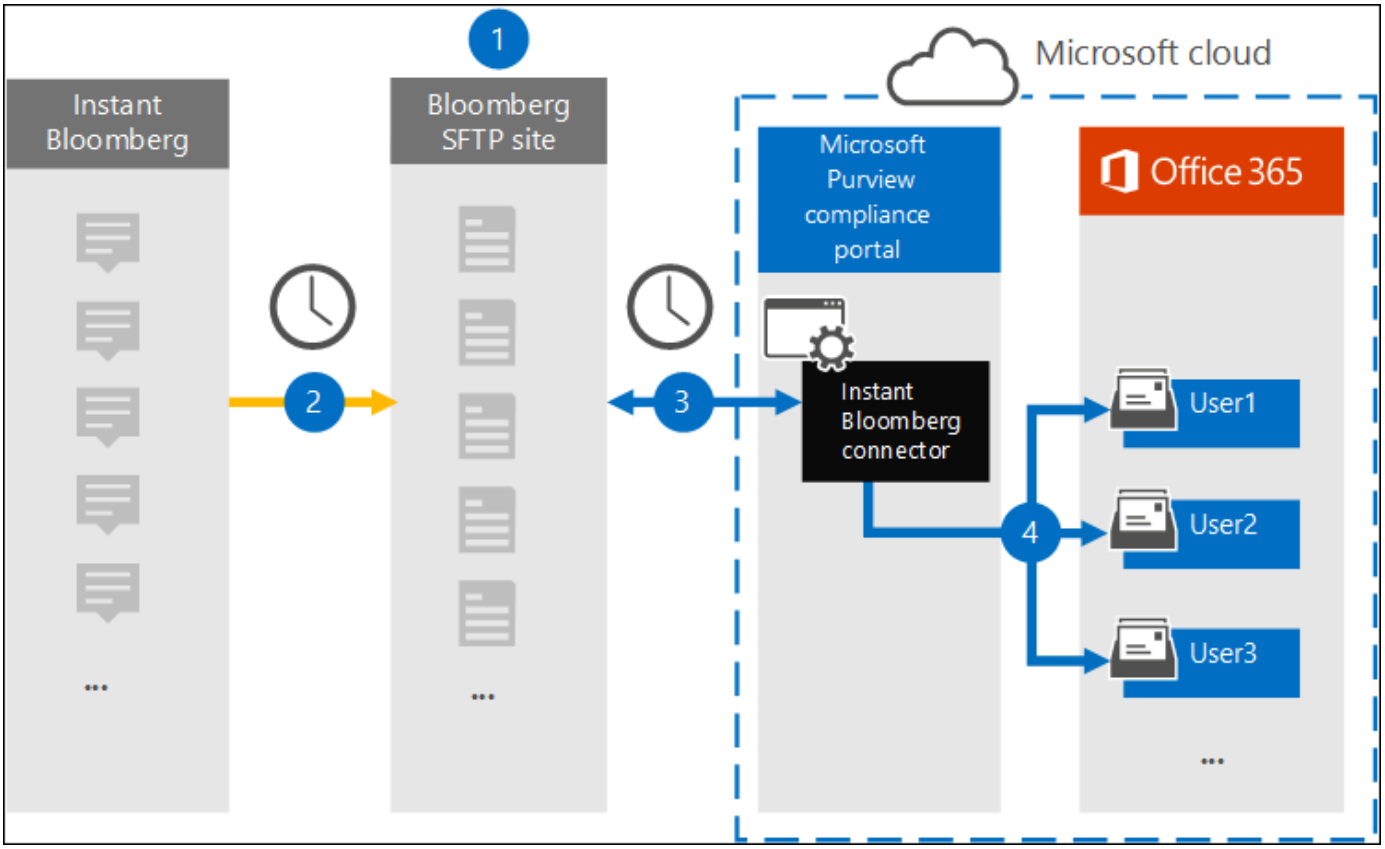
Set up a connector to archive Instant Bloomberg data

Use a connector to import and archive financial services chat data from the Instant Bloomberg collaboration tool. After you set up and configure a connector, it connects to your organization's Bloomberg secure FTP site (SFTP) once every day, converts the content of chat messages to an email message format, and then imports those items to mailboxes in Microsoft 365.

After storing Instant Bloomberg data in user mailboxes, you can apply Microsoft Purview features such as Litigation Hold, eDiscovery, In-Place Archiving, Auditing, Communication Compliance, and Microsoft 365 retention policies to Instant Bloomberg data. For example, you can search Instant Bloomberg chat messages by using eDiscovery or associate the mailbox that contains the Instant Bloomberg data with a custodian in a Microsoft Purview eDiscovery (Premium) case. By using an Instant Bloomberg connector to import and archive data in Microsoft 365, you can help your organization stay compliant with government and regulatory policies.

Overview of archiving Instant Bloomberg data

The following overview explains the process of using a connector to archive Instant Bloomberg chat data in Microsoft 365.



1. Your organization works with Bloomberg to set up a Bloomberg SFTP site. You also work with Bloomberg to configure Instant Bloomberg to copy chat messages to your Bloomberg SFTP site.

2. Every 24 hours, Instant Bloomberg copies chat messages to the Bloomberg SFTP site.

3. The Instant Bloomberg connector that you create in the Microsoft Purview portal connects to the Bloomberg SFTP site every day and transfers the chat messages from the previous 24 hours to a secure Azure Storage area in the Microsoft Cloud. The connector also converts the content of a chat massage to an email message format.

4. The connector imports the chat message items to the mailbox of a specific user. A new folder named InstantBloomberg is created in the specific user's mailbox and the items are imported to it. The connector uses the value of the *CorporateEmailAddress* property. Every chat message contains this property, which is populated with the email address of every participant of the chat message. In addition to automatic user mapping by using the value of the *CorporateEmailAddress* property, you can also define a custom mapping by uploading a CSV mapping file. This mapping file should contain a Bloomberg UUID and the corresponding Microsoft 365 mailbox address for each user. If you enable automatic user mapping and provide a custom mapping, for every chat item the connector first looks at custom-mapping file. If it doesn't find a valid Microsoft 365 user that corresponds to a user's Bloomberg UUID, the connector uses the *CorporateEmailAddress* property of the chat item. If the connector doesn't find a valid Microsoft 365 user in either the custom-mapping file or the *CorporateEmailAddress* property of the chat item, the item isn't imported.

Before you begin

Some of the implementation steps required to archive Instant Bloomberg data are external to Microsoft 365. You must complete these steps before you can create the connector in the Microsoft Purview portal.

- To set up an Instant Bloomberg connector, you need keys and key passphrases for Pretty Good Privacy (PGP) and Secure Shell (SSH). Use these keys to configure the Bloomberg SFTP site. The connector uses these keys to connect to the Bloomberg SFTP site and import data to Microsoft 365. Use the PGP key to configure the encryption of data that's transferred from the Bloomberg SFTP site to Microsoft 365. Use the SSH key to configure secure shell to enable a secure remote authentication when the connector connects to the Bloomberg SFTP site.

When setting up a connector, you can use public keys and key passphrases provided by Microsoft or use your own private keys and passphrases. Use the public keys provided by Microsoft. However, if your organization already configured a Bloomberg SFTP site by using private keys, you can create a connector by using these same private keys.

- Subscribe to Bloomberg Anywhere. This subscription is required so that you can sign in to Bloomberg Anywhere to access the Bloomberg SFTP site that you set up and configure.

- Set up a Bloomberg SFTP (Secure file transfer protocol) site. After working with Bloomberg to set up the SFTP site, data from Instant Bloomberg is uploaded to the SFTP site every day. The connector you create in Step 2 connects to this SFTP site and transfers the chat data to Microsoft 365 mailboxes. SFTP also encrypts the Instant Bloomberg chat data that you send to mailboxes during the transfer process.

  For external information about Bloomberg SFTP (also called *BB-SFTP*), use the following resources:
  - See the "SFTP Connectivity Standards" document at Bloomberg Support.
  - Contact Bloomberg customer support.

  After you work with Bloomberg to set up an SFTP site, Bloomberg provides some information to you after you respond to the Bloomberg implementation email message. Save a copy of the following information. You use it to set up a connector in Step 3.
  - Firm code, which is an ID for your organization and is used to sign in to the Bloomberg SFTP site.
  - Password for your Bloomberg SFTP site
  - URL for Bloomberg SFTP site (for example, sftp.bloomberg.com)
  - Port number for Bloomberg SFTP site

- The Instant Bloomberg connector can import a total of 200,000 items in a single day. If there are more than 200,000 items on the SFTP site, none of those items are imported to Microsoft 365.

- The user who downloads the public keys and IP address in the "Get PGP and SSH and public keys" procedure and creates the connector in the "Create an Instant Bloomberg connector" procedure must be assigned the Data Connector Admin role. This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions, see Permissions in the Microsoft Purview

portal and Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview.

Set up a connector by using public keys

The steps in this section show you how to set up an Instant Bloomberg connector by using the public keys for Pretty Good Privacy (PGP) and Secure Shell (SSH).

Step 1: Get PGP and SSH and public keys

First, get a copy of the public keys for Pretty Good Privacy (PGP) and Secure Shell (SSH). You use these keys to configure the Bloomberg SFTP site and to allow the Instant Bloomberg connector that you create later to connect to the SFTP site and transfer the Instant Bloomberg chat data to Microsoft 365 mailboxes. You also download an IP address during this key-download procedure, which you use when configuring the Bloomberg SFTP site.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. From the list, select **Instant Bloomberg**.
5. On the **Terms of service** page, select **Accept**.
6. On the **Add credentials for content source** page, select **I want to use PGP and SSH public keys provided by Microsoft**.
7. Under step 1, select the **Download SSH key**, **Download PGP key**, and **Download IP address** links to save a copy of each file to your local computer.

   These files contain the following items that you use to configure the Bloomberg SFTP site in Step 2:

   - PGP public key: Use this key to configure the encryption of data that's transferred from the Bloomberg SFTP site to Microsoft 365.
   - SSH public key: Use this key to configure secure shell to enable a secure remote authentication when the connector connects to the Bloomberg SFTP site.
   - IP address: Configure the Bloomberg SFTP site to accept connection requests from this IP address. The same IP address is used by the Instant Bloomberg connector to connect to the SFTP site and transfer Instant Bloomberg data to Microsoft 365.

8. Select **Cancel** to close the Add credentials workflow. You return to this connector setup workflow later to create the Instant Bloomberg connector.

Step 2: Configure the Bloomberg SFTP site

Use the PGP and SSH public keys and the IP address that you downloaded earlier to set up PGP encryption and SSH authentication for the Bloomberg SFTP site. The PGP encryption and SSH authentication configuration lets the Instant Bloomberg connector connect to the Bloomberg SFTP site and transfer Instant Bloomberg data to Microsoft 365 after you create the connector later in this procedure. You need to work with Bloomberg customer support to set up your Bloomberg SFTP site. Contact Bloomberg customer support for assistance.

> **Important**
>
> Bloomberg recommends that you attach the downloaded SSH key, PGP key, and IP address files to an email message and send them to their customer support team when working with them to set up your Bloomberg SFTP site.

Step 3: Create an Instant Bloomberg connector

In this step, you create an Instant Bloomberg connector in the Microsoft Purview portal. The connector uses the information you provide to connect to the Bloomberg SFTP site and transfer chat messages to the corresponding user mailboxes in Microsoft 365.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. From the list, select **Instant Bloomberg**.
5. On the **Terms of service** page, select **Accept**.
6. On the **Add credentials for Bloomberg SFTP site** page, in the SFTP credentials section, enter the required information in the following boxes and then select **Next**.
    - **Firm code:** The ID for your organization that Bloomberg uses as the username for the SFTP site.
    - **Password:** Password for Bloomberg SFTP site.

- **SFTP URL:** The URL for Bloomberg SFTP site (for example, `sftp.bloomberg.com`). You can also use an IP address for this value.
- **SFTP port:** The port number for Bloomberg SFTP site. The connector uses this port to connect to the SFTP site.

7. On the **Define user** page, select one of the following options to specify the users whose data you want to import.

   - **All users in your organization**. Select this option to import data for all users.
   - **Only users on Litigation hold**. Select this option to import data only for users whose mailboxes are placed on Litigation hold. This option imports data to user mailboxes that have the LitigationHoldEnabled property set to True. For more information, see Create a Litigation hold.

8. On the **Select data types to import** page, select the required data types to be imported apart from **Messages**

9. On the **Map Instant Bloomberg users to Microsoft 365 users** page, enable automatic user mapping and provide custom user mapping as required

   > ⓘ **Note**
   >
   > The connector imports the chat message items to the mailbox of a specific user. A new folder named **InstantBloomberg** is created in the specific user's mailbox and the items are imported to it. The connector uses the value of the *CorporateEmailAddress* property. Every chat message contains this property, and the property is populated with the email address of every participant of the chat message. In addition to automatic user mapping using the value of the *CorporateEmailAddress* property, you can also define custom mapping by uploading a CSV mapping file. The mapping file should contain the Bloomberg UUID and corresponding Microsoft 365 mailbox address for each user. If you enable automatic user mapping and provide a custom mapping, for every chat item the connector first looks at custom mapping file. If it doesn't find a valid Microsoft 365 user that corresponds to a user's Bloomberg UUID, the connector uses the *CorporateEmailAddress* property of the chat item. If the connector doesn't find a valid Microsoft 365 user in either the custom mapping file or the *CorporateEmailAddress* property of the chat item, the item won't be imported.

10. Select **Next**, review your settings, and then select **Finish** to create the connector.

Set up a connector by using private keys

The following procedure shows how to set up an Instant Bloomberg connector by using PGP and SSH private keys. Use this connector setup option if your organization already configured a Bloomberg SFTP site by using private keys.

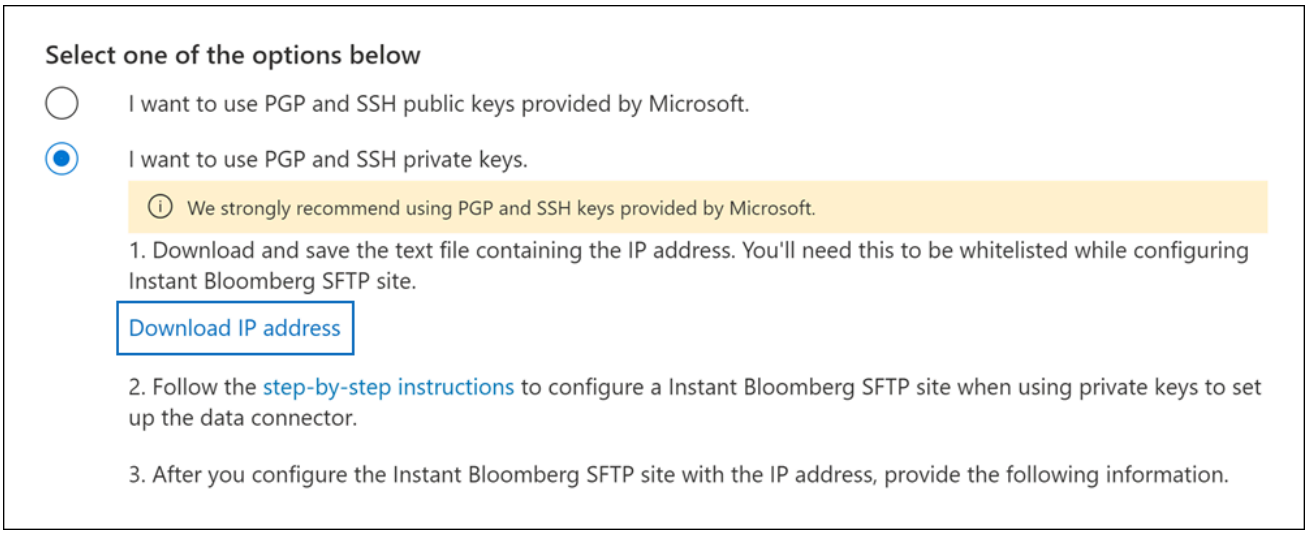
Step 1: Get an IP address to configure the Bloomberg SFTP site

> **Note**
>
> If your organization previously configured a Bloomberg SFTP site to archive Bloomberg Message data by using PGP and SSH private keys, you don't need to configure another one. You can specify the previously configured Bloomberg SFTP site when you create the Instant Bloomberg connector later in this procedure.

If your organization uses PGP and SSH private keys to set up a Bloomberg SFTP site, you need to get an IP address and give it to Bloomberg customer support. The Bloomberg SFTP site must accept connection requests from this IP address. The same IP address is used by the Instant Bloomberg connector to connect to the SFTP site and transfer Instant Bloomberg data to Microsoft 365.

To get the IP address:

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. From the list, select **Instant Bloomberg**.
5. On the **Terms of service** page, select **Accept**.
6. On the **Add credentials for content source** page, select **I want to use PGP and SSH private keys**.
7. Under step 1, select **Download IP address** to save a copy of the IP address file to your local computer.



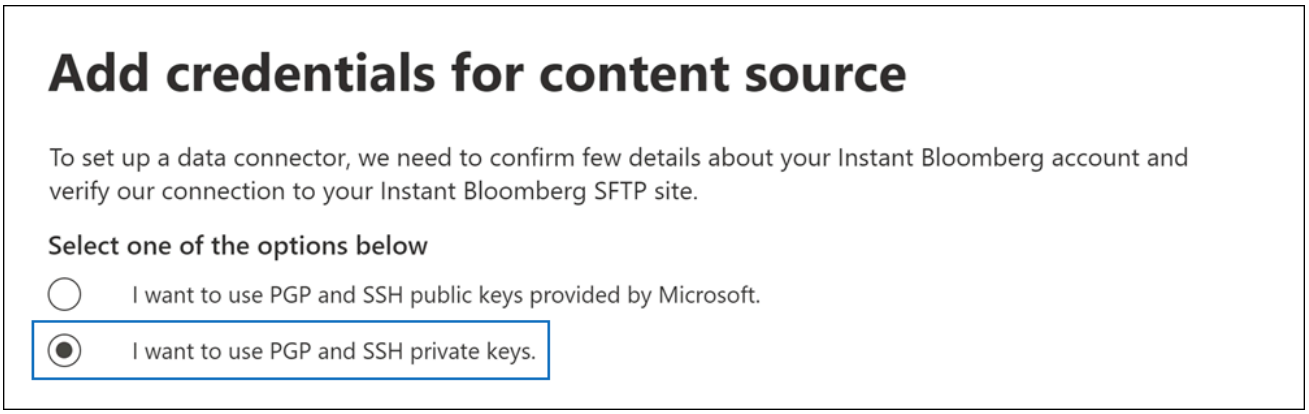
8. Select **Cancel** to close the workflow. You return to this workflow in Step 2 to create the connector.

Work with Bloomberg customer support to configure your Bloomberg SFTP site to accept connection requests from this IP address. Contact Bloomberg customer support for assistance.

Step 2: Create an Instant Bloomberg connector

After you configure your Bloomberg SFTP site, create an Instant Bloomberg connector in the Microsoft Purview portal. The connector uses the information you provide to connect to the Bloomberg SFTP site and transfer email messages to the corresponding user mailboxes in Microsoft 365. Before creating the connector, be sure to have copies of the PGP and SSH private keys and key passphrases that you used to set up your Bloomberg SFTP site.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. From the list, select **Instant Bloomberg**.
5. On the **Terms of service** page, select **Accept**.
6. On the **Add credentials for content source** page, select **I want to use PGP and SSH private keys**.



7. In the connection details section, enter the required information in the following boxes and then select **Validate connection**.

    - **Name:** The name for the connector. It must be unique in your organization.
    - **Firm code:** The ID for your organization that Bloomberg uses as the username for the SFTP site.
    - **Password:** The password for your organization's Bloomberg SFTP site.
    - **SFTP URL:** The URL for the Bloomberg SFTP site (for example, `sftp.bloomberg.com`). You can also use an IP address for this value.
    - **SFTP port:** The port number for the Bloomberg SFTP site. The connector uses this port to connect to the SFTP site.
    - **PGP private key:** The PGP private key for the Bloomberg SFTP site. Be sure to include the entire private key value, including the beginning and ending lines of the key block.
    - **PGP key passphrase:** The passphrase for the PGP private key.
    - **SSH private key:** The SSH private key for the Bloomberg SFTP site. Be sure to include the entire private key value, including the beginning and ending lines of the key block.
    - **SSH key passphrase:** The passphrase for the SSH private key.

8. After the connection is successfully validated, select **Next**.

9. On the **Define user** page, select one of the following options to specify the users whose data you want to import.

    - **All users in your organization**. Select this option to import data for all users.
    - **Only users on Litigation hold**. Select this option to import data only for users whose mailboxes are placed on Litigation hold. This option imports data to user mailboxes that have the LitigationHoldEnabled property set to True. For more information, see Create a Litigation hold.

10. On the **Map Instant Bloomberg users to Microsoft 365 users** page, enable automatic user mapping and provide custom user mapping as required.

> **Note**
>
> The connector imports the chat message items to the mailbox of a specific user. A new folder named **InstantBloomberg** is created in the specific user's mailbox and the items are imported to it. The connector uses the value of the *CorporateEmailAddress* property. Every chat message contains this property, and the property is populated with the email address of every participant of the chat message. In addition to automatic user mapping using the value of the *CorporateEmailAddress* property, you can also define custom mapping by uploading a CSV mapping file. The mapping file should contain the Bloomberg UUID and corresponding Microsoft 365 mailbox address for each user. If you enable automatic user mapping and provide a custom mapping, for every chat item the connector first looks at custom mapping file. If it doesn't find a valid Microsoft 365 user that corresponds to a user's Bloomberg UUID, the connector uses the *CorporateEmailAddress* property of the chat item. If the connector doesn't find a valid Microsoft 365 user in either the custom mapping file or the *CorporateEmailAddress* property of the chat item, the item won't be imported.

11. Select **Next**, review your settings, and then select **Finish** to create the connector.

Note: The author created this article with assistance from AI. Learn more

Last updated on 06/24/2026

Set up a connector to archive LinkedIn data

Use a connector to import and archive data from LinkedIn Company pages. After you set up and configure a connector, it connects to the account for the specific LinkedIn Company page once every 24 hours. The connector converts the messages posted to the Company page to an email message, then imports those items to a mailbox in Microsoft 365.

After storing the LinkedIn Company page data in a mailbox, you can apply Microsoft Purview features such as Litigation Hold, eDiscovery, In-Place Archiving, Auditing, and Microsoft 365 retention policies to LinkedIn data. For example, you can search for these items by using eDiscovery or associate the storage mailbox with a custodian in a Microsoft Purview eDiscovery (Premium) case. Creating a connector to import and archive LinkedIn data in Microsoft 365 can help your organization stay compliant with government and regulatory policies.

Before you begin

Before you create the connector, make sure you meet the following requirements:

- Assign the Data Connector Admin role to the user who creates a LinkedIn Company Page connector. This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions, see Permissions in the Microsoft Purview portal and Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview.
- Have the sign-in credentials (email address or phone number and password) of a LinkedIn user account that is an admin for the LinkedIn Company Page that you want to archive. You use these credentials to sign into LinkedIn when setting up the connector.
- Understand that the LinkedIn connector can import a total of 200,000 items in a single day. If there are more than 200,000 LinkedIn items in a day, none of those items are imported to Microsoft 365.

Create a LinkedIn connector

Perform the following steps to create a LinkedIn connector in Microsoft Purview.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. Select **LinkedIn company pages** from the list.
5. On the **Terms of service** page, select **Accept**.
6. On the **Sign in with LinkedIn** page, select **Sign in with LinkedIn**.
7. On the LinkedIn sign in page, enter the email address (or phone number) and password for the LinkedIn account associated with the company page that you want to archive, then select **Sign in**. A workflow page displays with a list of all LinkedIn Company Pages associated with the account that you signed in to. You can only configure a connector for one company page. If your organization has multiple LinkedIn Company Pages, you need to create a connector for each one.
8. Select the company page that you want to archive items from, then select **Next**.
9. On the **Choose storage location** page, select the email address of a Microsoft 365 mailbox that the LinkedIn items will be imported to, then select **Next**. Items are imported to the inbox folder in this mailbox. The mailbox used must have an Exchange Online Plan 1 or Plan 2 license.
10. Select **Next** to review the connector settings and then select **Finish** to complete the connector setup.

> **Note**
>
> LinkedIn items are imported to the LinkedIn subfolder in the inbox of the storage mailbox in Microsoft 365. They appear as email messages.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 06/24/2026

Set up a connector to import physical badging data

Set up a data connector to import physical badging data, such as an employee's raw physical access events or any physical access alarms generated by your organization's badging system. Examples of physical access points are an entry to a building or an entry to server room or data center. Microsoft Purview Insider Risk Management solution uses physical badging data to help protect your organization from malicious activity or data theft inside your organization.

Setting up a physical badging connector consists of the following tasks:

- Create an app in Microsoft Entra ID to access an API endpoint that accepts a JSON payload that contains physical badging data.
- Create the JSON payload with a schema defined by physical badging data connector.
- Create a physical badging data connector in the Microsoft Purview portal.
- Run a script to push the physical badging data to the API endpoint.
- Optionally, schedule the script to run automatically to import currently physical badging data.

Before you begin

- Assign the Data Connector Admin role to the user who creates the physical badging connector in Step 3. This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions, see:- Permissions in the Microsoft Purview portal - Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview compliance

  > ⓘ **Note**
  >
  > The Data Connector Admin role isn't currently supported in US Government GCC High and DoD environments. Therefore, assign the Mailbox Import Export role in Exchange Online to the user who creates the HR connector in GCC High and DoD environments. By default, no role group in Exchange Online has this role assigned. You can add the Mailbox Import Export role to the Organization Management role group in Exchange Online. Or, you can create a new role group, assign the Mailbox Import Export role, and add the appropriate users as members. For more

information, see the **Create role groups** or **Modify role groups** sections in the article "Manage role groups in Exchange Online".

- Determine how to retrieve or export data from your organization's physical badging system (on a daily basis) and create a JSON file as described in Step 2. The script that you run in Step 4 pushes the data in the JSON file to the API endpoint.
- Understand that the sample script you run in Step 4 pushes the physical badging data from a JSON file to the connector API so the insider risk management solution can use it. This sample script isn't supported under any Microsoft standard support program or service. The sample script is provided AS IS without warranty of any kind. Microsoft further disclaims all implied warranties including, without limitation, any implied warranties of merchantability or of fitness for a particular purpose. You assume all risk arising from the use or performance of the sample script and documentation. In no event shall Microsoft, its authors, or anyone else involved in the creation, production, or delivery of the scripts be liable for any damages whatsoever (including, without limitation, damages for loss of business profits, business interruption, loss of business information, or other pecuniary loss) arising out of the use of or inability to use the sample scripts or documentation, even if Microsoft has been advised of the possibility of such damages.
- This data connector is available in GCC environments in the Microsoft 365 US Government cloud. Third-party applications and services might involve storing, transmitting, and processing your organization's customer data on third-party systems that are outside of the Microsoft 365 infrastructure and therefore aren't covered by the Microsoft Purview and data protection commitments. Microsoft makes no representation that use of this product to connect to third-party applications implies that those third-party applications are FedRAMP compliant.

Step 1: Create an app in Microsoft Entra ID

First, create and register a new app in Microsoft Entra ID. The app corresponds to the physical badging connector that you create in Step 3. When you create the app, Microsoft Entra ID can authenticate the push request for a JSON payload containing physical badging data. During the creation of this Microsoft Entra app, save the following information. You use these values in later steps.

- Microsoft Entra application ID (also called the *app Id* or *client Id*)
- Microsoft Entra application secret (also called the *client secret*)
- Tenant Id (also called the *directory Id*)

For step-by-step instructions for creating an app in Microsoft Entra ID, see Register an application with the Microsoft identity platform.

Step 2: Prepare a JSON file with physical badging data

In this step, you create a JSON file that contains information about employees' physical access data. As explained in the before you begin section, you need to determine how to generate this JSON file from your organization's physical badging system.

> **Note**
>
> Don't add non-English characters to the JSON file. The connector supports only English characters. If the JSON contains non-English characters, data ingestion might fail.

The JSON file must conform to the schema definition required by the connector. The following table describes the required schema properties for the JSON file:

Expand table

Property	Description	Data type
UserId	An employee can have multiple digital identities across the systems. The input needs to have the Microsoft Entra ID already resolved by the source system.	UPN or email address
AssetId	The reference ID of the physical asset or physical access point.	Alphanumeric string
AssetName	The friendly name of the physical asset or physical access point.	Alphanumeric string
EventTime	The time stamp of access.	Date and time, in UTC format
AccessStatus	Value of `Success` or `Failed`	String

The following example shows a JSON file that conforms to the required schema:

JSON

```
[
    {
        "UserId":"sarad@contoso.com",
```

```
        "AssetId":"Mid-Sec-7",
        "AssetName":"Main Building 1st Floor Mid Section",
        "EventTime":"2019-07-04T01:57:49",
        "AccessStatus":"Failed"
    },
    {
        "UserId":"pilarp@contoso.com",
        "AssetId":"Mid-Sec-7",
        "AssetName":"Main Building 1st Floor Mid Section",
        "EventTime":"2019-07-04T02:57:49",
        "AccessStatus":"Success"
    }
]
```

You can also download the schema definition for the JSON file from the workflow when you create the physical badging connector in Step 3.

text

```
{
   "title" : "Physical Badging Signals",
   "description" : "Access signals from physical badging systems",
   "DataType" : {
      "description" : "Identify what is the data type for input signal",
      "type" : "string",
   },
   "type" : "object",
   "properties": {
      "UserId" : {
         "description" : "Unique identifier AAD Id resolved by the source system",
         "type" : "string",
      },
      "AssetId": {
         "description" : "Unique ID of the physical asset/access point",
         "type" : "string",
      },
      "AssetName": {
         "description" : "friendly name of the physical asset/access point",
         "type" : "string",
      },
      "EventTime" : {
         "description" : "timestamp of access",
         "type" : "string",
      },
      "AccessStatus" : {
         "description" : "what was the status of access attempt - Success/Failed",
         "type" : "string",
      },
   }
   "required" : ["UserId", "AssetId", "EventTime" "AccessStatus"]
}
```

Step 3: Create the physical badging connector

In this step, you create a physical badging connector in the Microsoft Purview portal. When you run the script in Step 4, it processes the JSON file you created in this step and pushes it to the API endpoint you configured in Step 1. Be sure to copy the JobId generated when you create the connector. You'll use the JobId when you run the script.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. From the list, choose **Physical badging**.
5. On the **Authentication credentials** page, enter the following information, then select **Next**:
   a. Enter or paste the Microsoft Entra application ID for the Azure app that you created in Step 1.
   b. Download the sample schema for your reference to create the JSON file.
   c. Enter a unique name for the physical badging connector.
6. On the **Review** page, review your settings and select **Finish** to create the connector.
7. View the status page that confirms the connector was created. This page also contains the job ID. You can copy the job ID from this page or from the flyout page for the connector. You need this job ID when running the script.

   The status page also contains a link to the script. Refer to this script to understand how to post the JSON file to the API endpoint.
8. Select **Done**.

   The new connector appears in the list on the **Connectors** tab.
9. Select the physical badging connector that you just created to display the flyout page, which contains properties and other information about the connector.

Step 4: Run the script to upload your physical badging data

To set up a physical badging connector, run a script that pushes the physical badging data in the JSON file (created in Step 2) to the API endpoint (created in Step 1). We provide a sample script for your reference. You can choose to use it or create your own script to post the JSON file to the API endpoint.

After you run the script, the JSON file containing the physical badging data is pushed to your Microsoft 365 organization where the insider risk management solution can access it. We recommend you post physical badging data daily. You can automate the process to generate the JSON file every day from your physical badging system and then schedule the script to push the data.

> **Note**
>
> The API can process a JSON file with up to 50,000 records.

1. Go to this GitHub site to access the sample script.
2. Select the **Raw** button to display the script in text view.
3. Copy all the lines in the sample script and save them to a text file.
4. Modify the sample script for your organization, if necessary.
5. Save the text file as a Windows PowerShell script file by using a filename suffix of .ps1; for example, PhysicalBadging.ps1.
6. Open a Command Prompt on your local computer, and go to the directory where you saved the script.
7. Run the following command to push the physical badging data in the JSON file to the Microsoft cloud; for example:

   PowerShell

   ```
   .\PhysicalBadging.ps1 -tenantId "<Tenant Id>" -appId "<Azure AD App Id>" -appSecret "<Azure AD App Secret>" -jobId "Job Id" -jsonFilePath "<records file path>"
   ```

   The following table describes the parameters to use with this script and their required values. Use the information you obtained in the previous steps for these values.

Expand table

Parameter	Description
tenantId	This is the Id for your Microsoft 365 organization that you obtained in Step 1. You can also get the tenantId for your organization on the **Overview** blade in the Microsoft Entra admin center. This value identifies your organization.
appId	This is the Microsoft Entra application Id for the app that you created in Microsoft Entra ID in Step 1. Microsoft Entra ID uses this value for authentication when the script attempts to access your Microsoft 365 organization.
appSecret	This is the Microsoft Entra application secret for the app that you created in Microsoft Entra ID in Step 1. Microsoft Entra ID also uses this value for authentication.
jobId	This is the Job Id for the physical badging connector that you created in Step 3. The script uses this value to associate the physical badging data that it pushes to the Microsoft cloud with the physical badging connector.
JsonFilePath	This is the file path on the local computer (the one you're using to run the script) for the JSON file that you created in Step 2. This file must follow the sample schema described in Step 3.

Here's an example of the syntax for the physical badging connector script using actual values for each parameter:

PowerShell

```
.\PhysicalBadging.ps1 -tenantId aaaabbbb-0000-cccc-1111-dddd2222eeee -appId
00001111-aaaa-2222-bbbb-3333cccc4444 -appSecret Aa1Bb~2Cc3.-
Dd4Ee5Ff6Gg7Hh8Ii9_Jj0Kk1Ll2 -jobId 00001111-aaaa-2222-bbbb-3333cccc4444 -
jsonFilePath 'C:\Users\contosoadmin\Desktop\Data\physical_badging_data.json'
```

If the upload is successful, the script displays the **Upload Successful** message.

If you have multiple JSON files, run the script for each file.

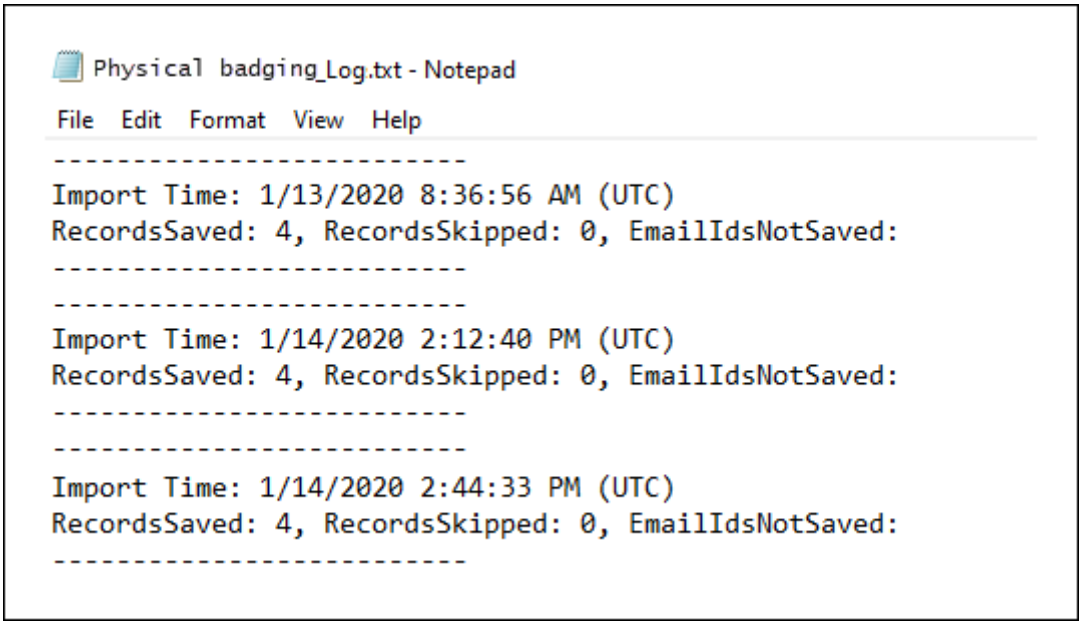
Step 5: Monitor the physical badging connector

After you create the physical badging connector and push your physical badging data, you can view the connector and upload status in the Microsoft Purview portal. If you schedule the script to run automatically on a regular basis, you can also view the current status after the last time the script runs.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.

3. Select **My connectors**, then select the physical badging connector that you created to display the flyout page. This page contains the properties and information about the connector.

4. Under **Last import**, select the **Download log** link to open (or save) the status log for the connector. This log contains information about each time the script runs and uploads the data from the JSON file to the Microsoft cloud.



   The **RecordsSaved** field shows the number of records in the JSON file that the script uploads. For example, if the JSON file contains four records, the value of the **RecordsSaved** fields is 4 when the script successfully uploads all the records in the JSON file. The **RecordsSkipped** field shows the number of records in the JSON file that the script skips. Before uploading records in the JSON file, the script validates the Email IDs. Any record with an invalid Email ID is skipped and the corresponding Email ID displays in the field **EmailIdsNotSaved**.

If you didn't run the script in Step 4, a link to download the script appears under **Last import**. You can download the script and then follow the steps in Step 4 to run it.

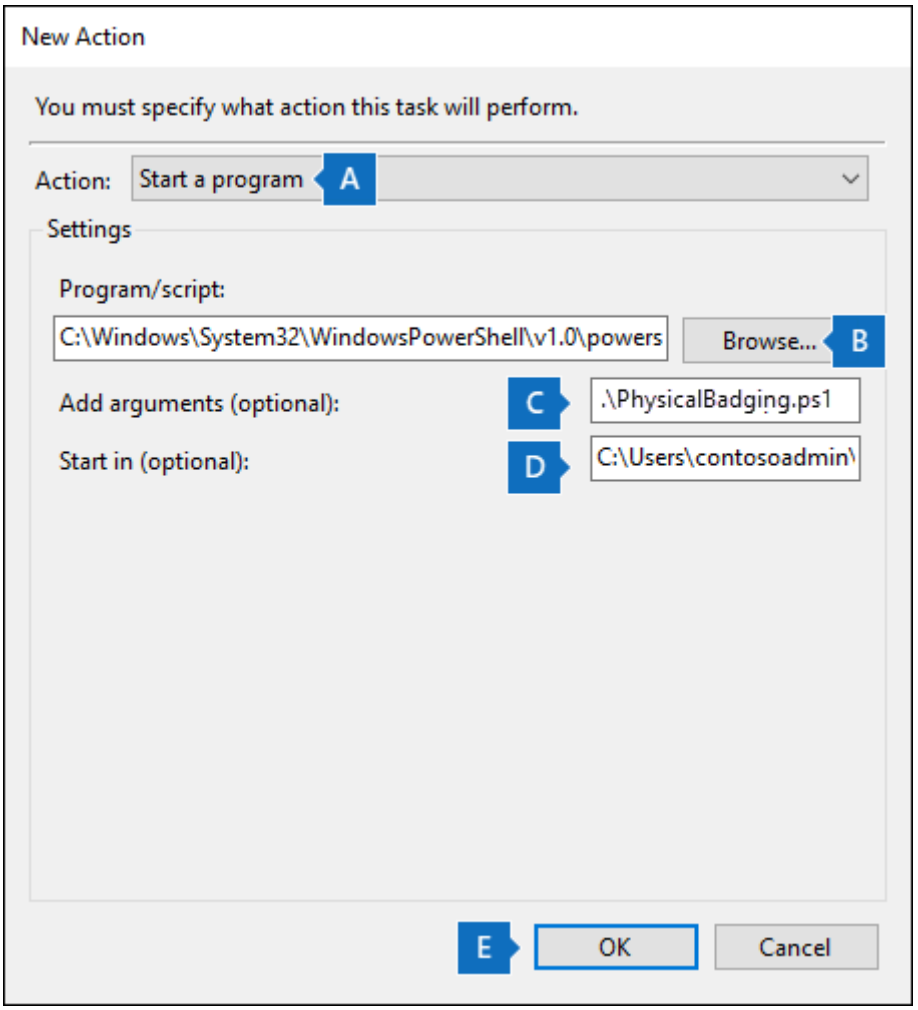
(Optional) Step 6: Schedule the script to run automatically

To make sure tools like the insider risk management solution always have the latest physical badging data from your organization, schedule the script to run automatically on a recurring basis, such as once a day. This schedule requires updating the physical badging data to JSON file on a similar (if not the same) schedule so that it contains the latest information about employees who leave your organization. The goal is to upload the most current physical

badging data so that the physical badging connector can make it available to the insider risk management solution.

Use the Task Scheduler app in Windows to automatically run the script every day.

1. On your local computer, select the Windows **Start** button and then type **Task Scheduler**.
2. Select the **Task Scheduler** app to open it.
3. In the **Actions** section, select **Create Task**.
4. On the **General** tab, type a descriptive name for the scheduled task, such as **physical badging connector Script**. You can also add an optional description.
5. Under **Security options**, do the following things:
    a. Decide whether to run the script only when you're logged on to the computer or run it when you're logged on or not.
    b. Make sure that the **Run with the highest privileges** checkbox is selected.
6. Select the **Triggers** tab, select **New**, and then do the following things:
    a. Under **Settings**, select the **Daily** option, and then choose a date and time to run the script for the first time. The script runs every day at the same specified time.
    b. Under **Advanced settings**, make sure the **Enabled** checkbox is selected.
    c. Select **Ok**.
7. Select the **Actions** tab, select **New**, and then do the following things:



a. In the **Action** dropdown list, make sure that **Start a program** is selected.

b. In the **Program/script** box, select **Browse**, and go to the following location and select it so the path is displayed in the box: `C:\Windows\System32\WindowsPowerShell\v1.0\powershell.exe`.

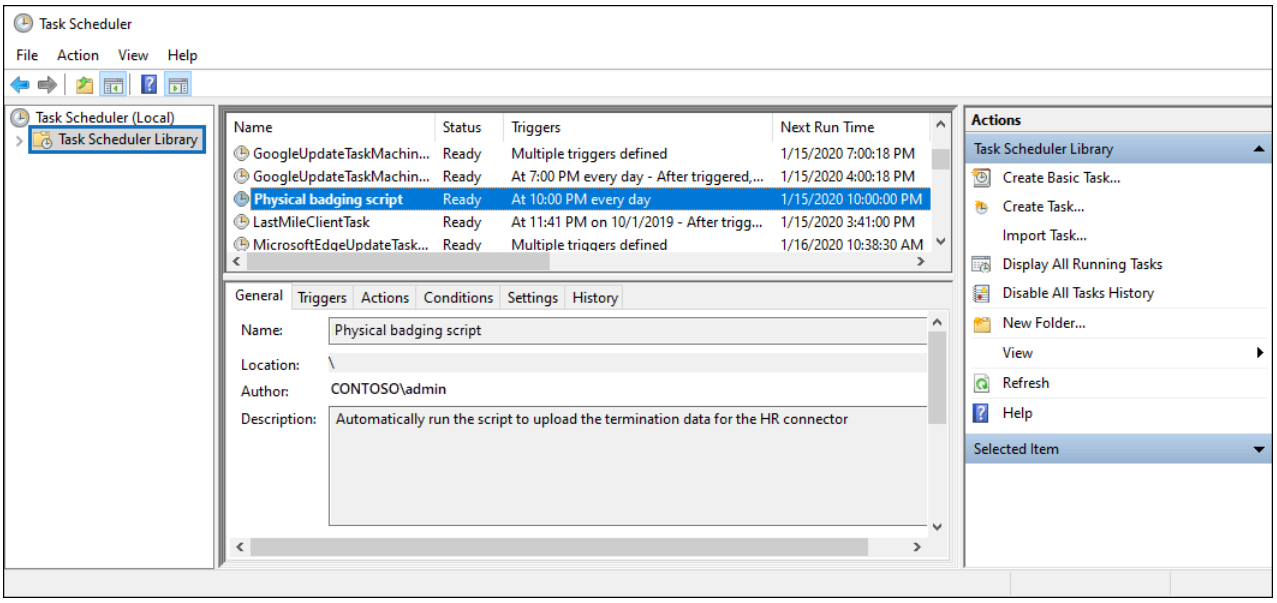
c. In the **Add arguments (optional)** box, paste the same script command that you ran in Step 4. For example, `.\PhysicalBadging.ps1 -tenantId "aaaabbbb-0000-cccc-1111-dddd2222eeee" -appId "00001111-aaaa-2222-bbbb-3333cccc4444" -appSecret "Aa1Bb~2Cc3.-Dd4Ee5Ff6Gg7Hh8Ii9_Jj0Kk1Ll2" -jobId "00001111-aaaa-2222-bbbb-3333cccc4444" -jsonFilePath "C:\Users\contosoadmin\Desktop\Data\physical_badging_data.json"`

d. In the **Start in (optional)** box, paste the folder location of the script that you ran in Step 4. For example, `C:\Users\contosoadmin\Desktop\Scripts`.

e. Select **Ok** to save the settings for the new action.

8. In the **Create Task** window, select **Ok** to save the scheduled task. You might be prompted to enter your user account credentials.

   The new task is displayed in the Task Scheduler Library.



The last time the script ran and the next time it's scheduled to run is displayed. You can double-select the task to edit it.

You can also verify the last time the script ran on the flyout page of the corresponding physical badging connector in the Microsoft Purview portal.

Last updated on 03/30/2026

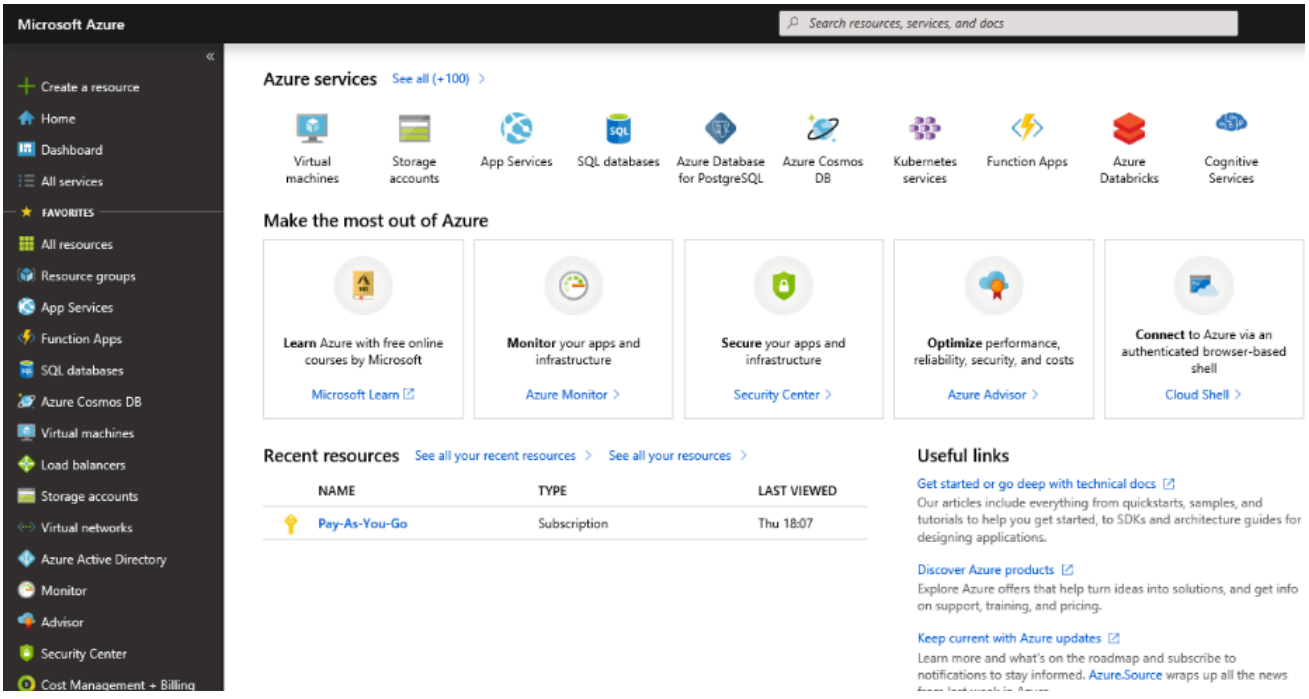
Deploy a connector to archive Twitter data

This article describes the process to deploy a connector that uses the Microsoft 365 Import service to import data from your organization's Twitter account to Microsoft 365. For a high-level overview of this process and a list of prerequisites required to deploy a Twitter connector, see Set up a connector to archive Twitter data.

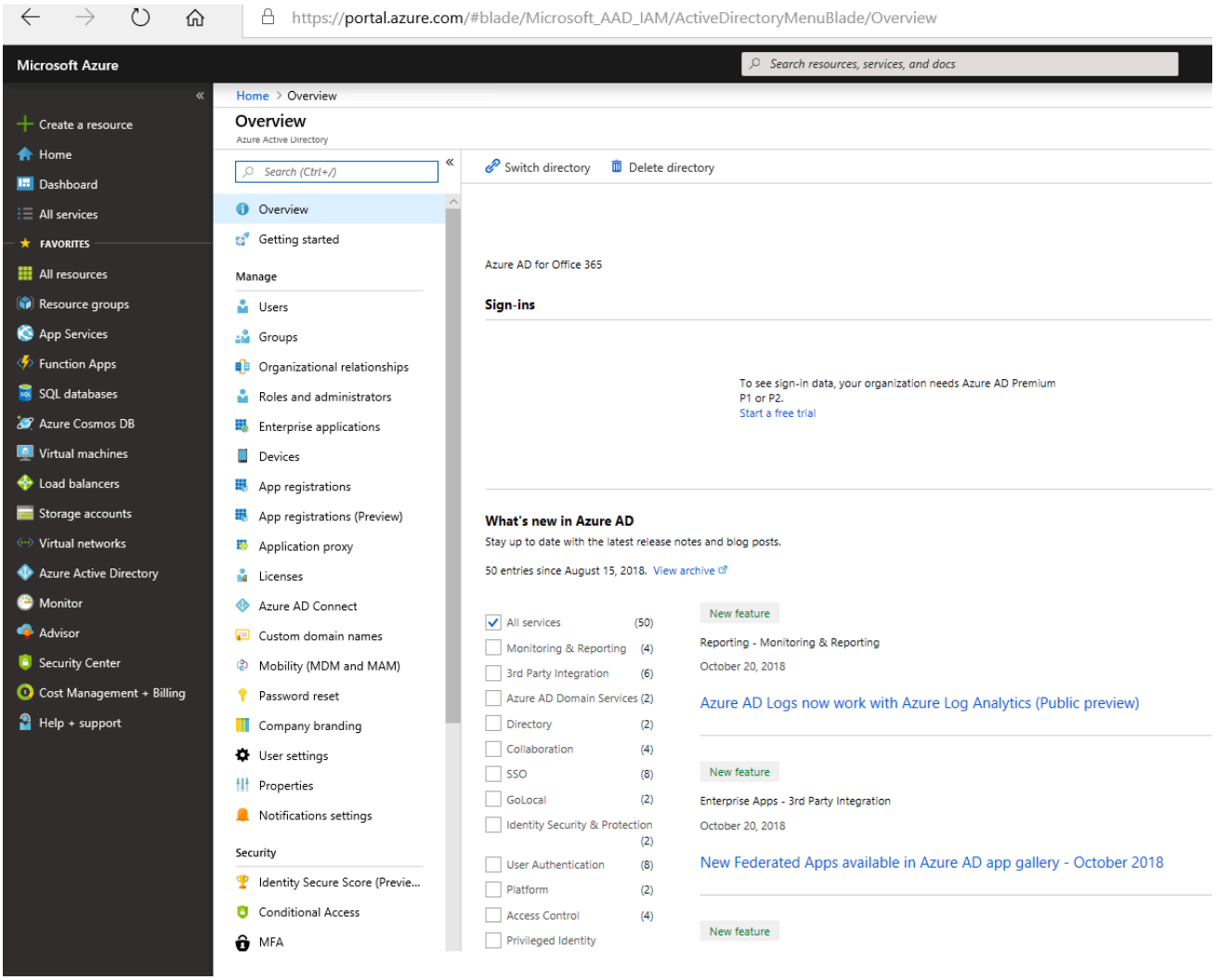
Step 1: Create an app in Microsoft Entra ID

1. Go to https://portal.azure.com and sign in with the credentials of a global admin account.

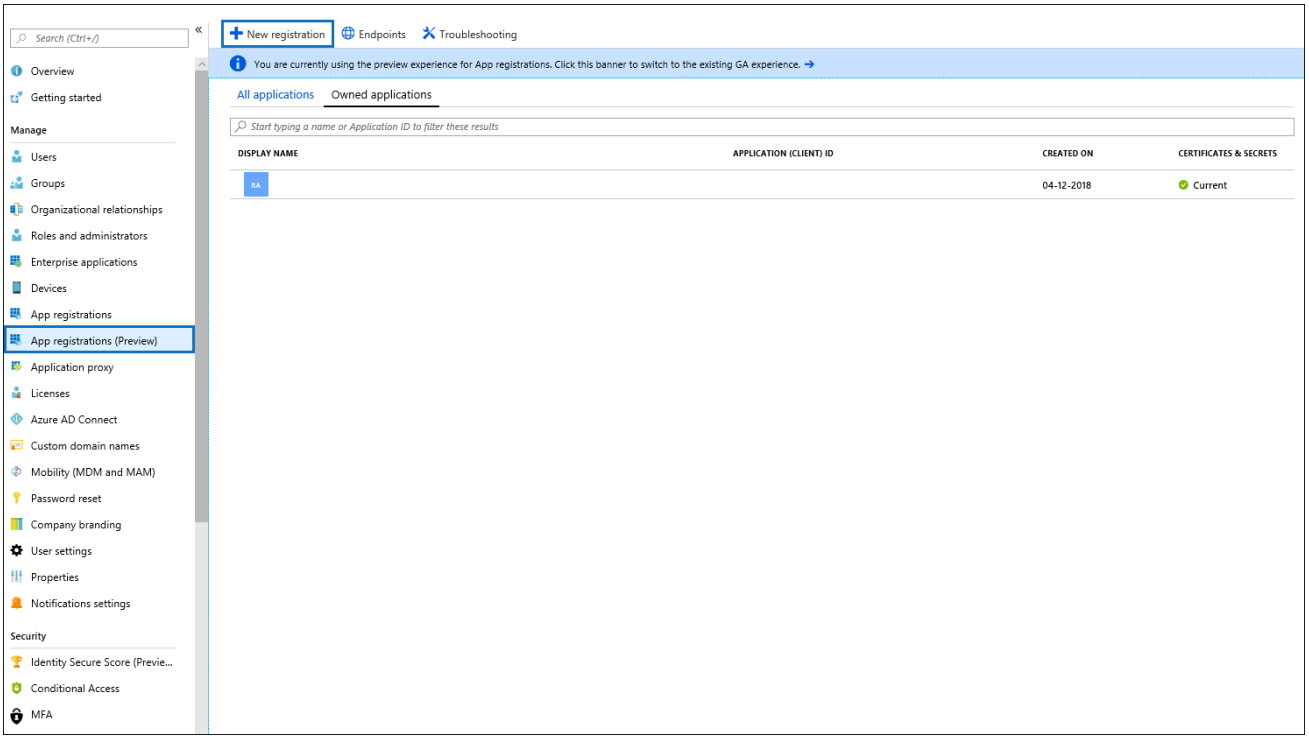
   > **Important**
   >
   > Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.



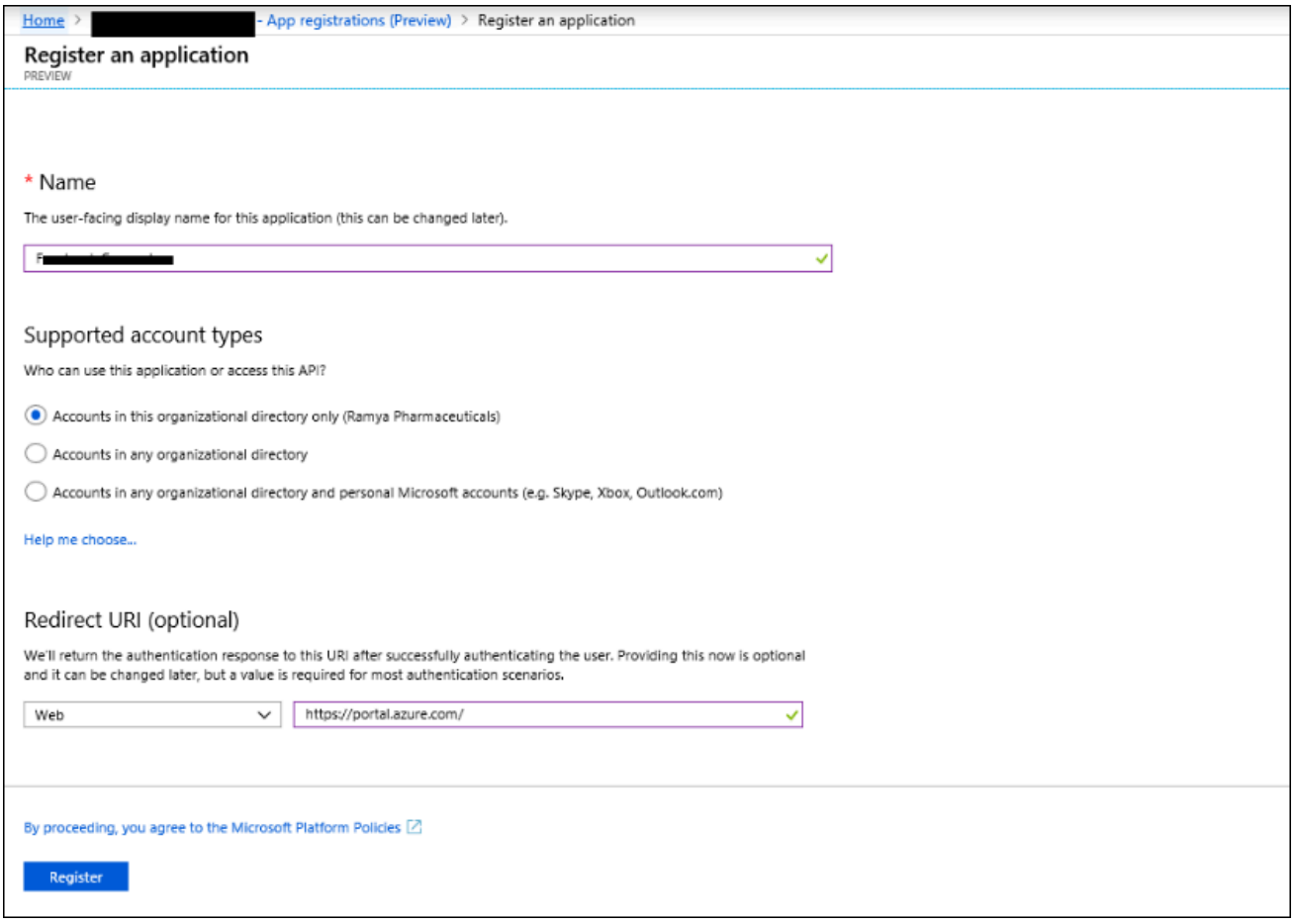
2. In the left navigation, select **Microsoft Entra ID**.



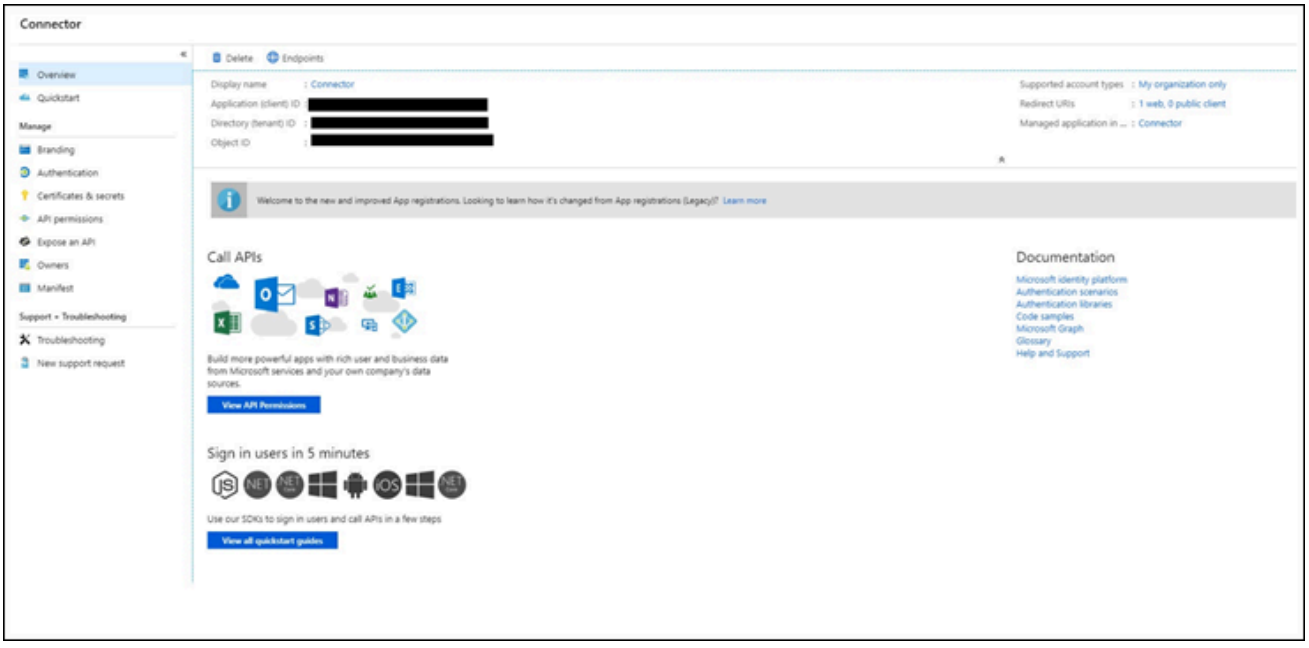
3. In the left navigation, select **App registrations (Preview)**, then select **New registration**.



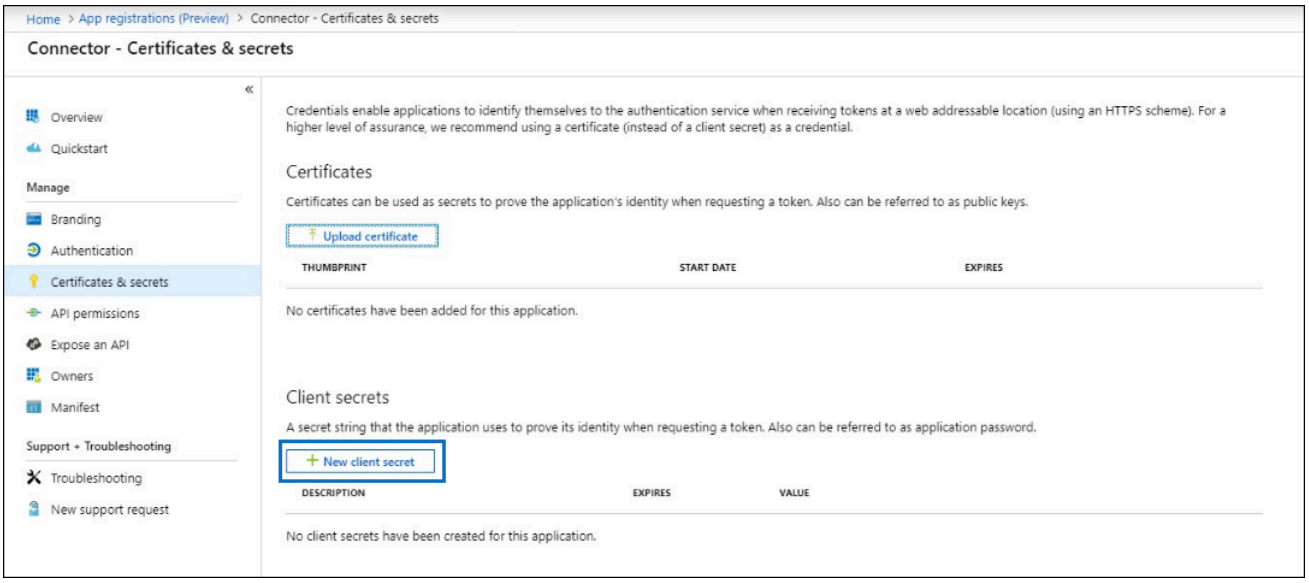
4. Register the application. Under **Redirect URI (optional)**, select **Web** in the application type dropdown list, then type `https://portal.azure.com` in the box for the URI.



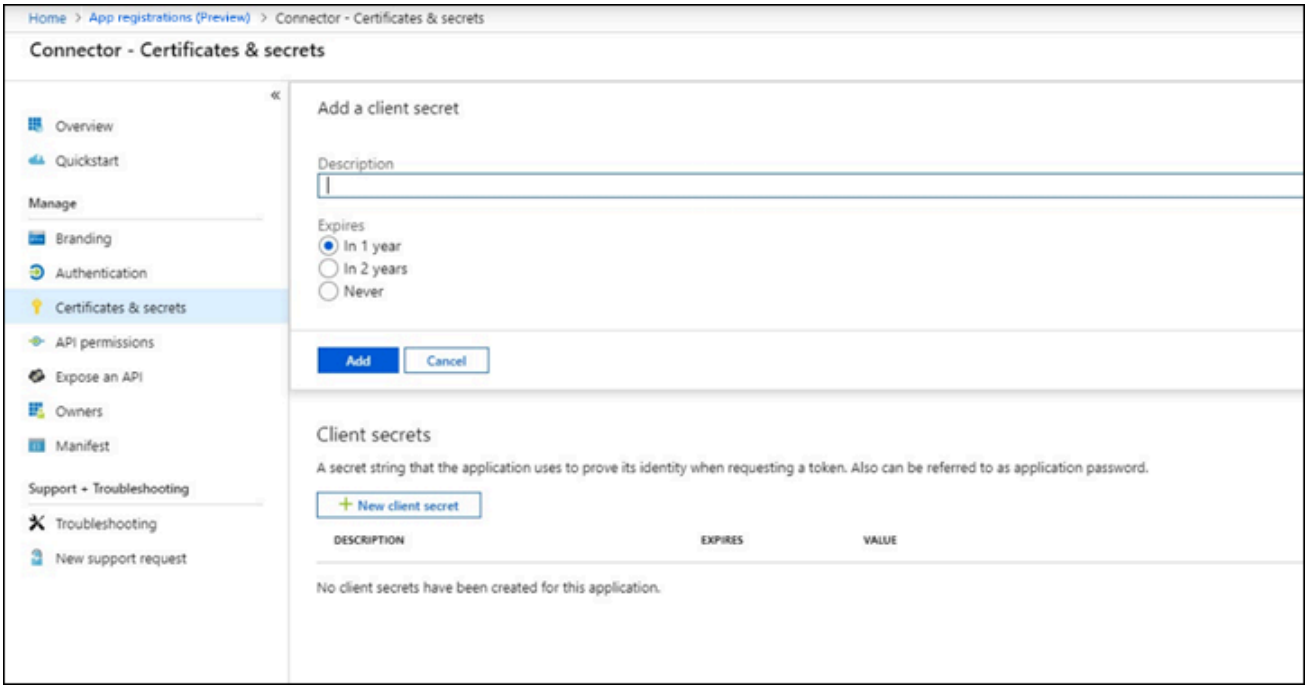
5. Copy the **Application (client) ID** and **Directory (tenant) ID** and save them to a text file or other safe location. You use these IDs in later steps.



6. Go to **Certificates & secrets for the new app** and under **Client secrets**, select **New client secret**.



7. Create a new secret. In the description box, type the secret and then select an expiration period.

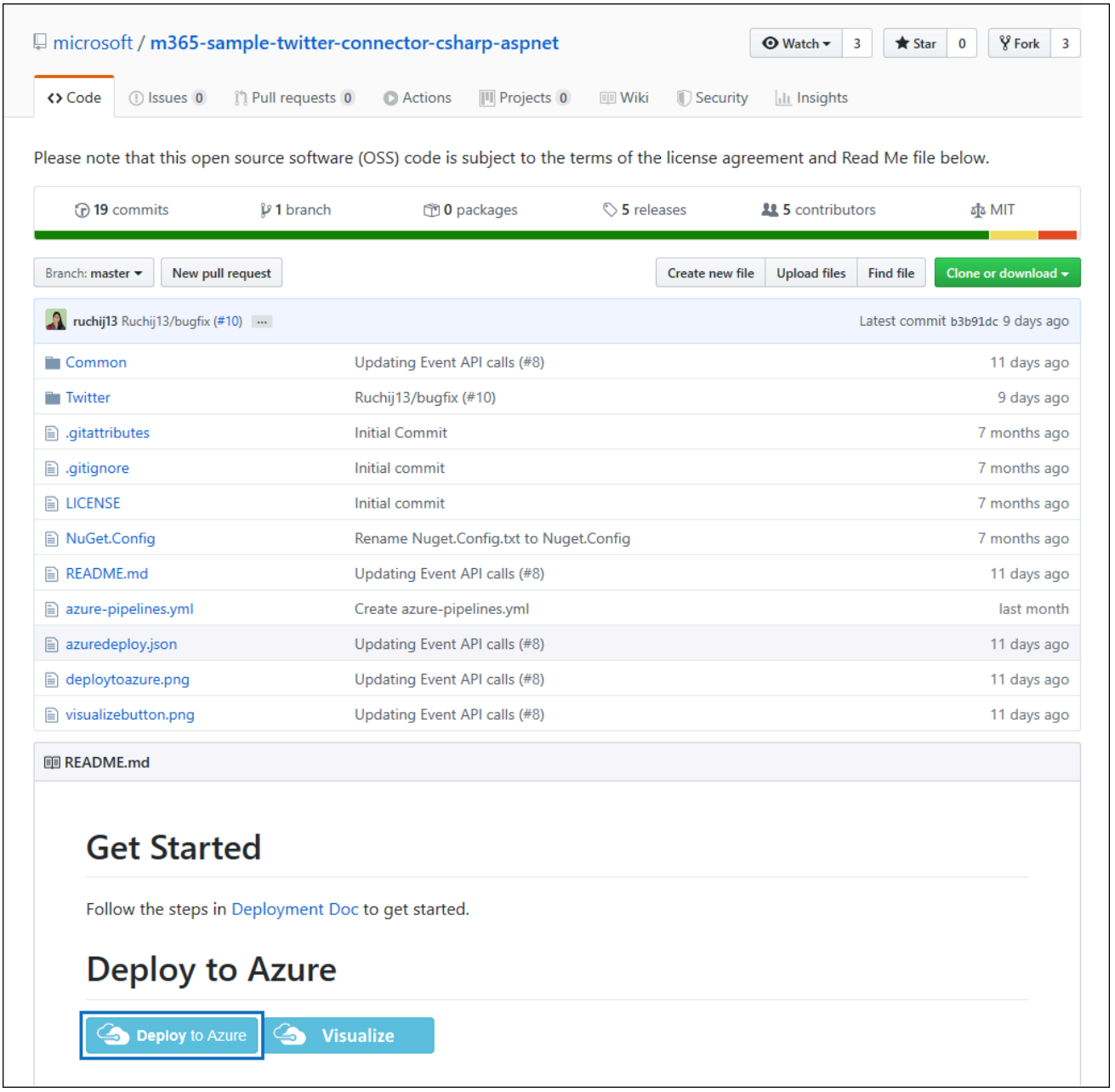


8. Copy the value of the secret and save it to a text file or other storage location. This is the Microsoft Entra application secret that you use in later steps.

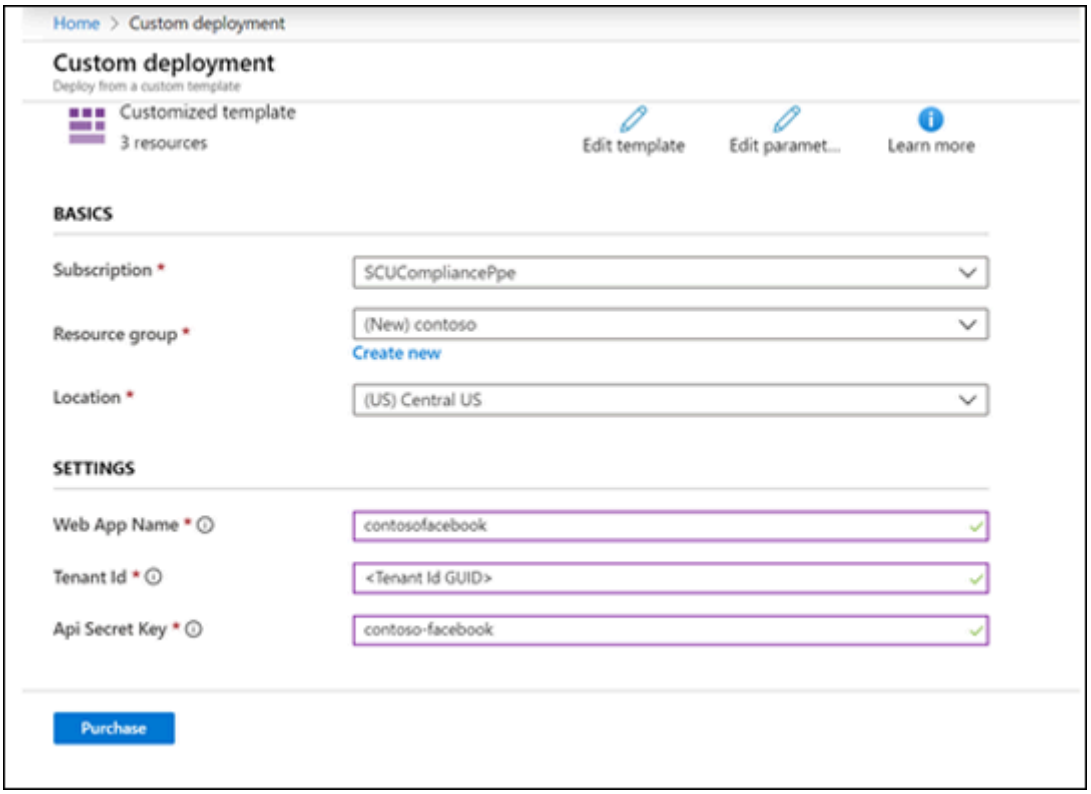


Step 2: Deploy the connector web service from GitHub to your Azure account

1. Go to this GitHub site and select **Deploy to Azure**.

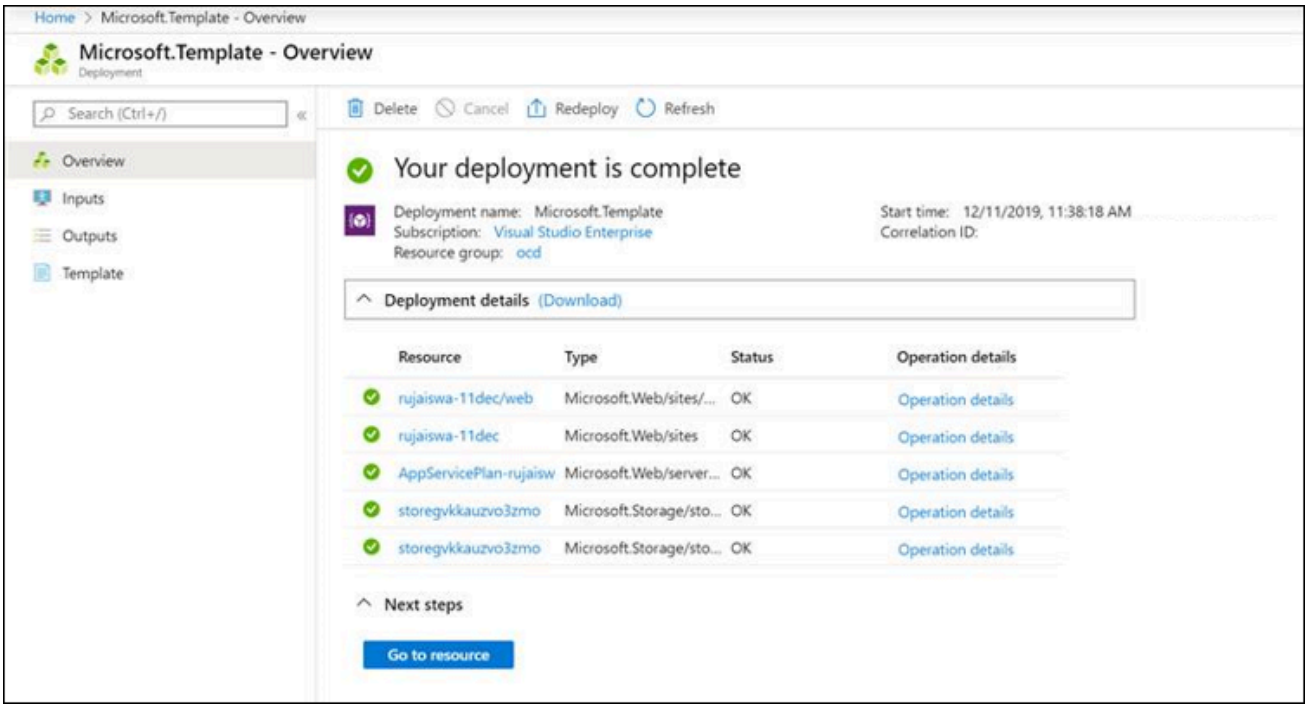


2. After you select **Deploy to Azure**, you're redirected to an Azure portal with a custom template page. Fill in the **Basics** and **Settings** details, then select **Purchase**.



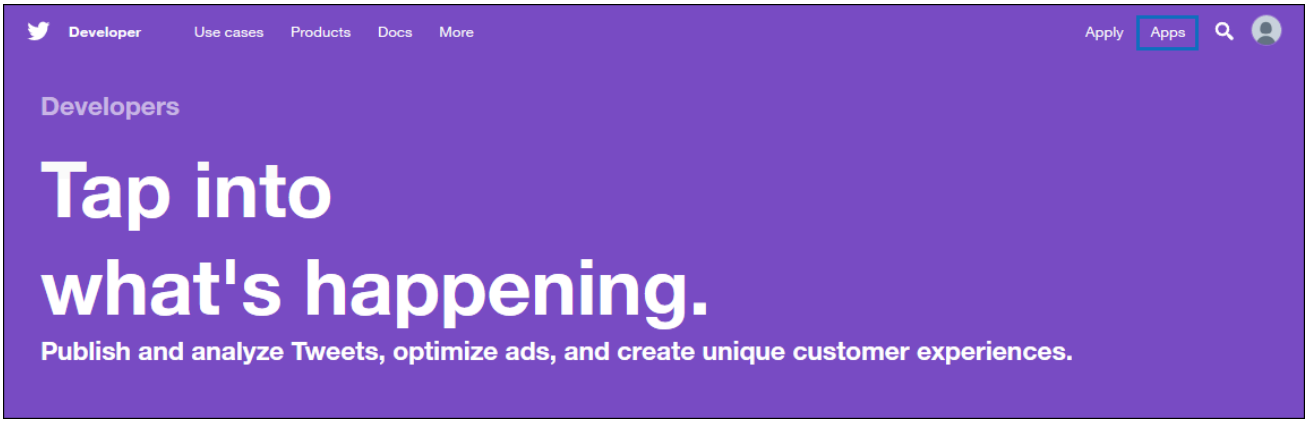
- **Subscription:** Select your Azure subscription that you want to deploy the Twitter connector web service to.
- **Resource group:** Select or create a new resource group. A resource group is a container that holds related resources for an Azure solution.
- **Location:** Select a location.
- **Web App Name:** Provide a unique name for the connector web app. The name must be between 3 and 18 characters in length. This name is used to create the Azure app service URL; for example, if you provide the Web app name of **twitterconnector** then the Azure app service URL is **twitterconnector.azurewebsites.net**.
- **tenantId:** The tenant ID of your Microsoft 365 organization that you copied after creating the Twitter connector app in Microsoft Entra ID in Step 1.
- **APISecretKey:** You can type any value as the secret. This value is used to access the connector web app in Step 5.

3. When the deployment succeeds, the page looks similar to the following screenshot:

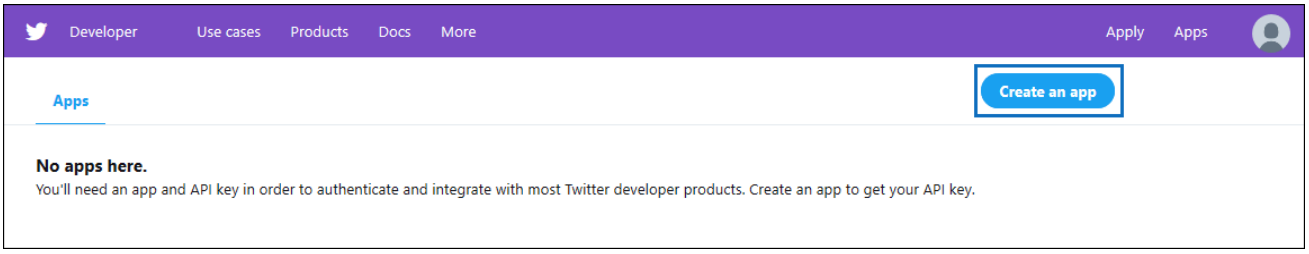


Step 3: Create the Twitter app

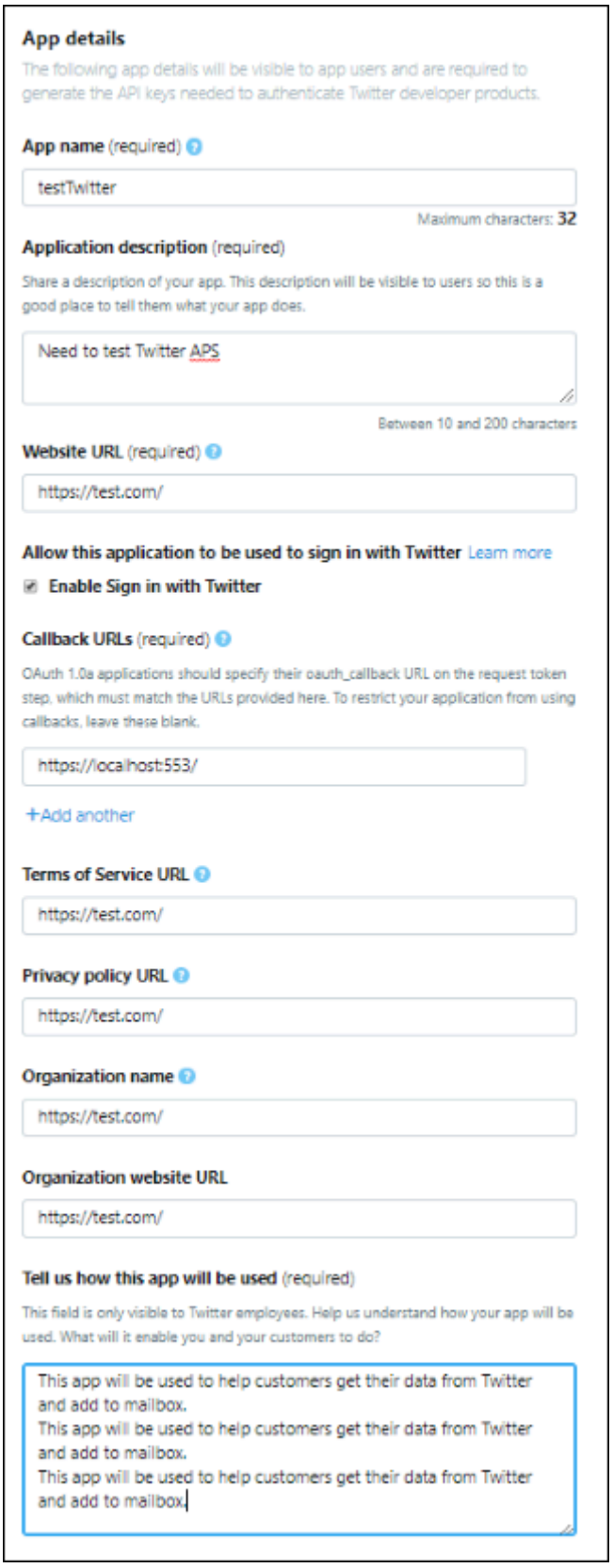
1. Go to https://developer.twitter.com , sign in with the credentials for the developer account for your organization, and select **Apps**.



2. Select **Create an app**.



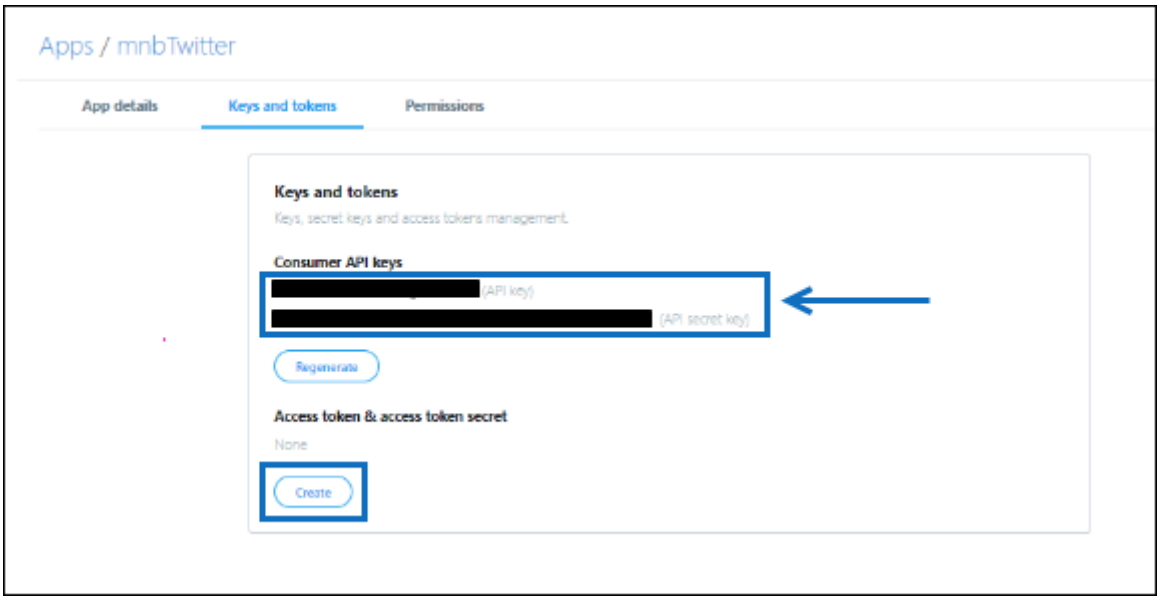
3. Under **App details**, enter information about the application.



4. On the Twitter developer dashboard, select the app that you just created, then select **Details**.

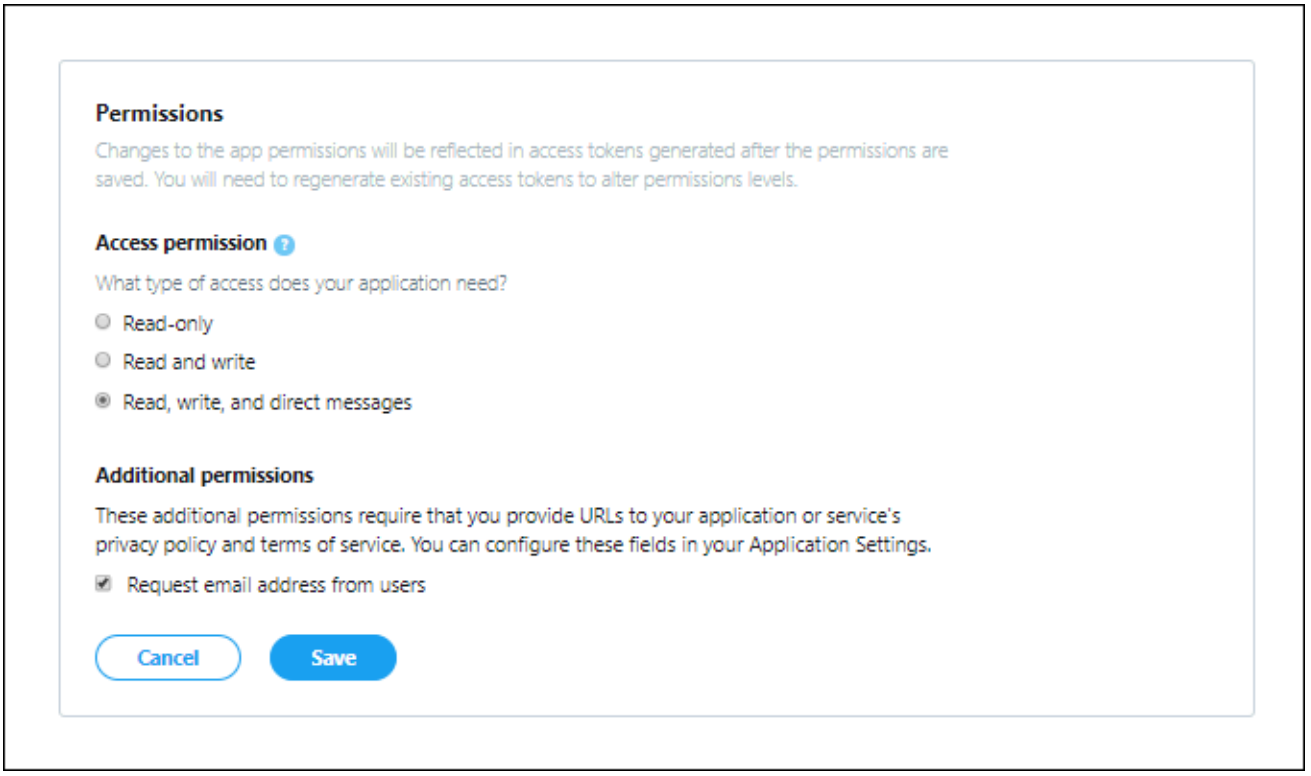


5. On the **Keys and tokens** tab, under **Consumer API keys**, copy both the API Key and the API secret key and save them to a text file or other storage location. Then select **Create** to generate an access token and an access token secret and copy these to a text file or other storage location.

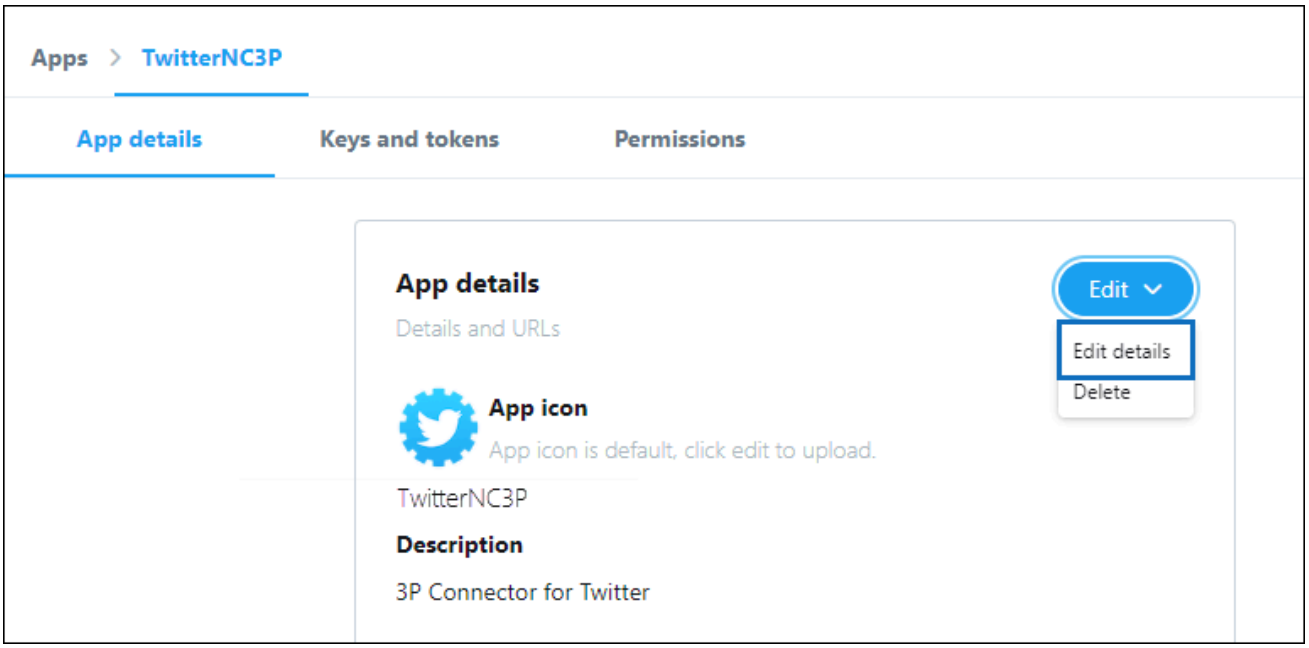


Select **Create** to generate an access token and an access token secret, and copy these to a text file or other storage location.

6. Select the **Permissions** tab and configure the permissions as shown in the following screenshot:

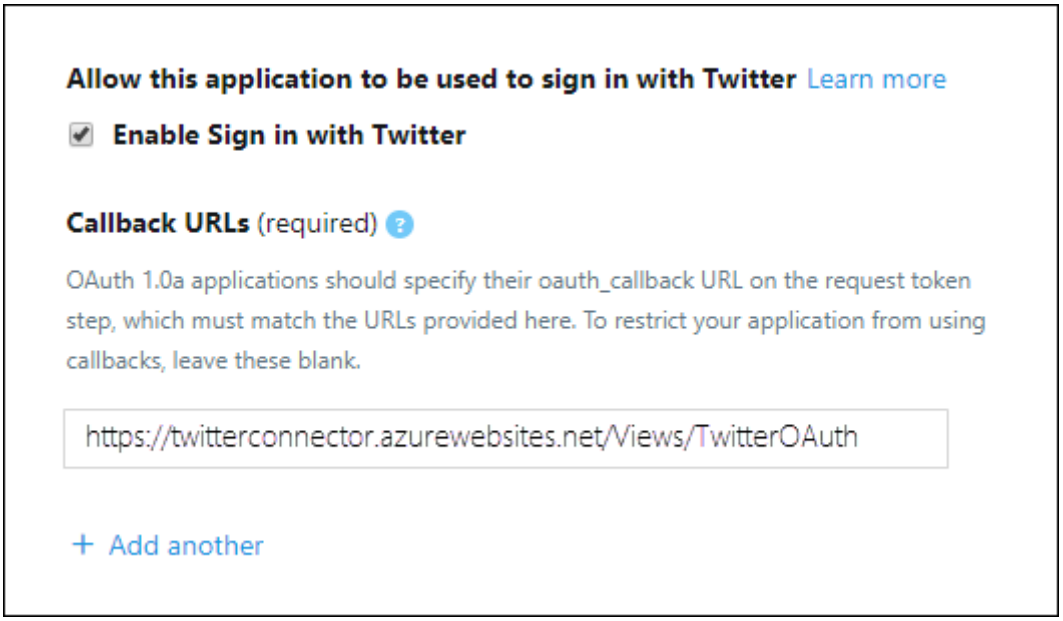


7. After you save the permission settings, select the **App details** tab, then select **Edit** > **Edit details**.



8. Complete the following tasks:

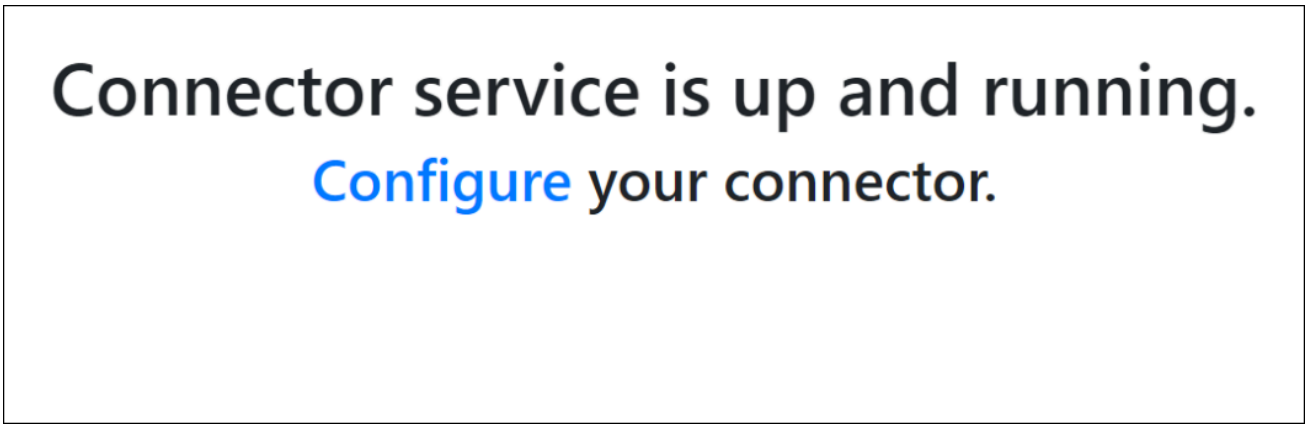
   - Select the checkbox to allow the connector app to sign in to Twitter.
   - Add the OAuth redirect Uri using the following format: **<connectorserviceuri>/Views/TwitterOAuth**, where the value of *connectorserviceuri* is the Azure app service URL for your organization.



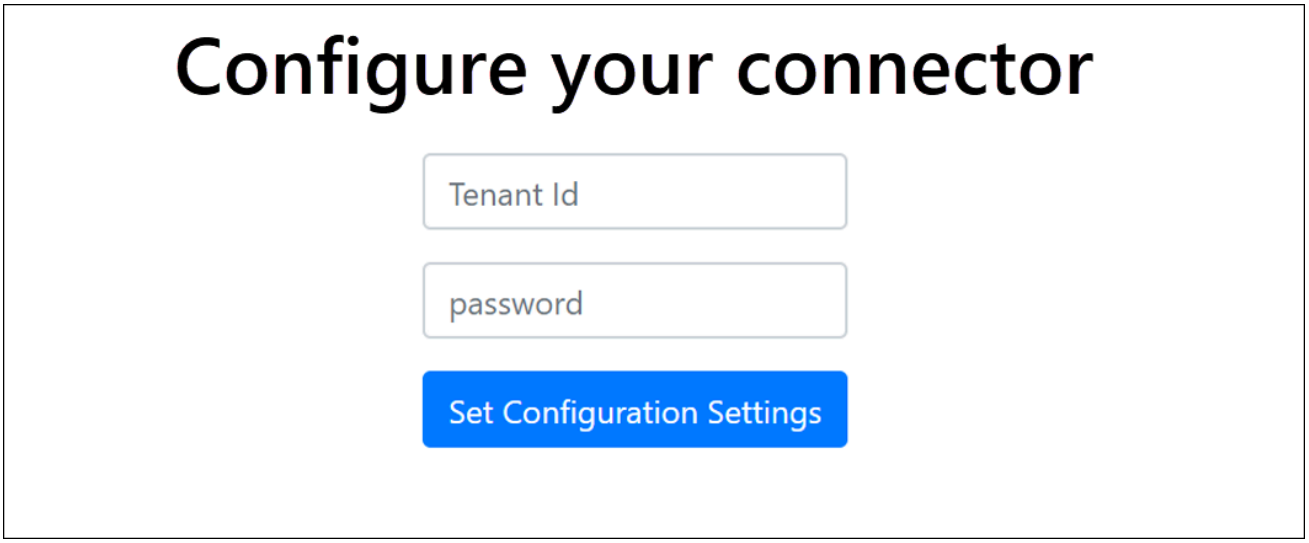
The Twitter developer app is now ready to use.

Step 4: Configure the connector web app

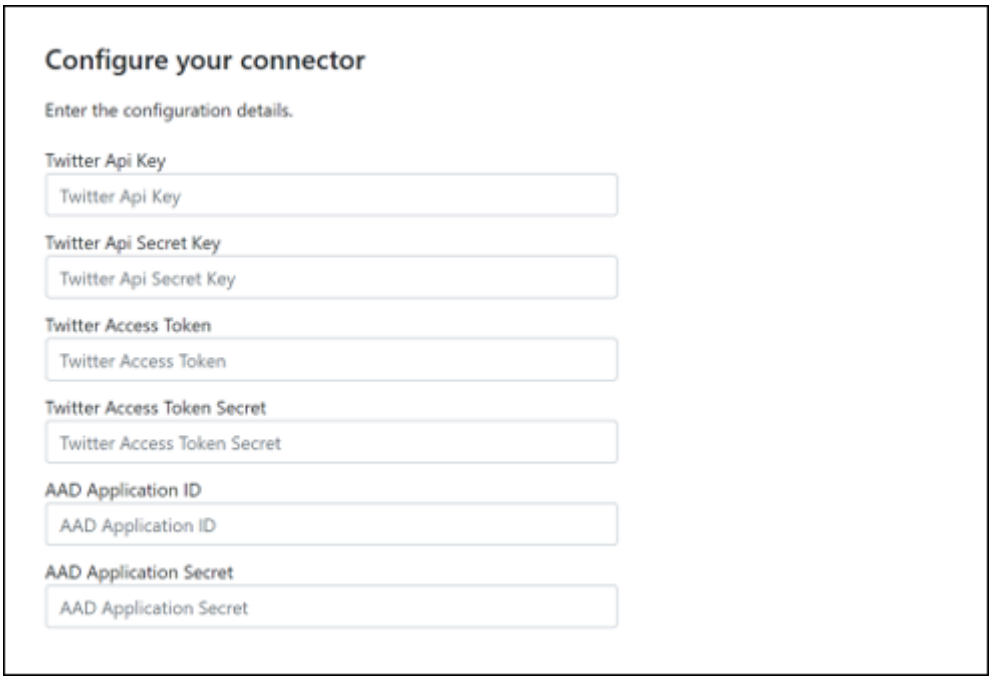
1. Go to `https://<AzureAppResourceName>.azurewebsites.net` (where **AzureAppResourceName** is the name of your Azure app resource that you named in Step 2). The home page of the app looks like the following screenshot:



2. Select **Configure** to display a sign in page.



3. In the Tenant Id box, type or paste your Microsoft Entra tenant ID (the Directory (tenant) ID from the Microsoft Entra app registration). In the password box, type or paste the APISecretKey value from the Azure deployment template settings, then select **Set Configuration Settings** to display the configuration details page.



4. Enter the following configuration settings:

   - **Twitter Api Key:** The API key for the Twitter application that you created in Step 3.
   - **Twitter Api Secret Key:** The API secret key for the Twitter application that you created in Step 3.
   - **Twitter Access Token:** The access token that you created in Step 3.
   - **Twitter Access Token Secret:** The access token secret that you created in Step 3.
   - **Microsoft Entra Application ID:** The application ID for the Microsoft Entra app that you created in Step 1
   - **Microsoft Entra Application Secret:** The value of the Microsoft Entra application client secret that you created during the app registration in Microsoft Entra ID.

5. Select **Save** to save the connector settings.

Step 5: Set up a Twitter connector

1. Sign in to the Microsoft Purview portal.

2. Select **Settings** > **Data connectors**.

3. Select **My connectors**, then select **Add connector**.

4. From the list, select **Twitter**.

5. On the **Terms of service** page, select **Accept**.

6. On the **Add credentials for your connector app** page, enter the following information and then select **Validate connection**.

    - In the **Name** box, type a name for the connector, such as **Twitter help handle**.
    - In the **Connector URL** box, type or paste the Azure app service URL.
    - In the **Password** box, type or paste the APISecretKey value from the Azure deployment template settings.
    - In the **Azure App ID** box, type or paste the value of the Azure Application App Id (also called the *client ID*) that you obtained in Step 1.

7. After the connection is successfully validated, select **Next**.

8. On the **Authorize Microsoft 365 to import data** page, type or paste the APISecretKey again and then select **Login web app**.

9. Select **Login with Twitter**.

10. On the Twitter sign in page, sign in by using the credentials for your organization's Twitter account. After you sign in, the Twitter page displays the following message, "Twitter Connector Job Successfully set up."

11. Select **Continue** to complete setting up the Twitter connector.

12. On the **Set filters** page, you can apply a filter to initially import items that are a certain age. Select an age, and then select **Next**.

13. On the **Choose storage location** page, type the email address of Microsoft 365 mailbox that the Twitter items are imported to, and then select **Next**.

14. Select **Next** to review the connector settings and then select **Finish** to complete the connector setup.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/24/2026

Set up a Microsoft connector to archive Twitter data (preview)

Use a connector to import and archive data from Twitter to Microsoft 365. After you set up and configure the connector, it connects to your organization's Twitter account (on a scheduled basis), converts the content of an item to an email message format, and then imports those items to a mailbox in Microsoft 365. Before you begin, make sure you have a Twitter account for your organization, a valid Azure subscription, and the Data Connector Admin role in the Microsoft Purview portal.

After Twitter data is imported, you can apply Microsoft Purview features such as Litigation Hold, eDiscovery, In-Place Archiving, Auditing, Communication Compliance, and Microsoft 365 retention policies to the data stored in the mailbox. For example, you can search Twitter data using eDiscovery or associate the mailbox where the data is stored with a custodian in an eDiscovery case. Using a connector to import and archive Twitter data in Microsoft 365 can help your organization stay compliant with government and regulatory policies.

Before you begin

Complete the following prerequisites before you set up and configure a connector in the Microsoft Purview portal to import and archive data from your organization's Twitter account.

- You need a Twitter account for your organization. You need to sign in to this account when setting up the connector.
- Your organization must have a valid Azure subscription. If you don't have an existing Azure subscription, you can sign up for one of these options:
  - Sign up for a free one year Azure subscription
  - Sign up for a Pay-As-You-Go Azure subscription
- The Twitter connector can import a total of 200,000 items in a single day. If there are more than 200,000 Twitter items in a day, none of those items are imported to Microsoft 365.
- The user who sets up the Twitter connector in the Microsoft Purview portal (in Step 5) must be assigned the Data Connector Admin role. This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role

by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions, see Permissions in the Microsoft Purview portal and Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview.

Step 1: Create an app in Microsoft Entra ID

First, register a new app in Microsoft Entra ID. This app corresponds to the web app resource that you implement in Step 2 for the Twitter connector.

For more information, see Create an app in Microsoft Entra ID.

When you complete this step, save the following information to a text file. You'll use these values in later steps in the deployment process.

- Microsoft Entra application ID
- Microsoft Entra application secret
- Tenant Id

Step 2: Deploy connector web service from GitHub repository to your Azure account

Next, deploy the source code for the Twitter connector app that uses the Twitter API to connect to your Twitter account and extract data so you can import it to Microsoft 365. The Twitter connector that you deploy for your organization uploads the items from your organization's Twitter account to the Azure Storage location that you create in this step. After you create a Twitter connector in the Microsoft Purview portal (in Step 5), the Microsoft 365 Import service copies the Twitter data from the Azure Storage location to a mailbox in Microsoft 365. You must have a valid Azure subscription to create an Azure Storage account.

To deploy the source code for the Twitter connector app:

1. Go to the Microsoft 365 sample Twitter connector repository on GitHub.
2. Select **Deploy to Azure**.

For more information, see Deploy the connector web service from GitHub to your Azure account.

When you complete this step, provide the following information:

- APISecretKey: Create this secret during the completion of this step. You use it in Step 5.

- tenantId: The tenant ID of your Microsoft 365 organization that you copied after creating the Twitter app in Microsoft Entra ID in Step 1.

After you complete this step, be sure to copy the app Service URL (for example, `https://twitterconnector.azurewebsites.net`). You need to use this URL to complete Step 3, Step 4, and Step 5.

Step 3: Create developer app on Twitter

Create and configure a developer app on Twitter. The custom connector you create in Step 7 uses the Twitter app to interact with the Twitter API and obtain data from your organization's Twitter account.

For step-by-step instructions, see Create the Twitter app.

When you complete this step, save the following information to a text file. You'll use these values to configure the Twitter connector app in Step 4.

- Twitter API Key
- Twitter API Secret Key
- Twitter Access Token
- Twitter Access Token Secret

Step 4: Configure the Twitter connector app

Add configuration settings to the Twitter connector app that you deployed in Step 2. Go to the home page of your connector app and configure it.

For step-by-step instructions, see Configure the connector web app.

When you complete this step, provide the following information. You copied this information to a text file after completing the previous steps:

- Twitter API Key (obtained in Step 3)
- Twitter API Secret Key (obtained in Step 3)
- Twitter Access Token (obtained in Step 3)
- Twitter Access Token Secret (obtained in Step 3)
- Microsoft Entra application ID (the Microsoft Entra application ID obtained in Step 1)
- Microsoft Entra application secret (the Microsoft Entra application secret obtained in Step 1)

Step 5: Set up a Twitter connector in the Microsoft Purview portal

In this step, you set up the Twitter connector in the Microsoft Purview portal. The connector imports data from your organization's Twitter account to a specified mailbox in Microsoft 365. When you finish this step, the Microsoft 365 Import service starts importing data from your organization's Twitter account to Microsoft 365.

For step-by-step instructions, see Set up a Twitter connector.

As you complete this step, provide the following information and save it in a text file:

- Azure app service URL (obtained in Step 2; for example, `https://twitterconnector.azurewebsites.net`)
- APISecretKey (created in Step 2)

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/19/2026

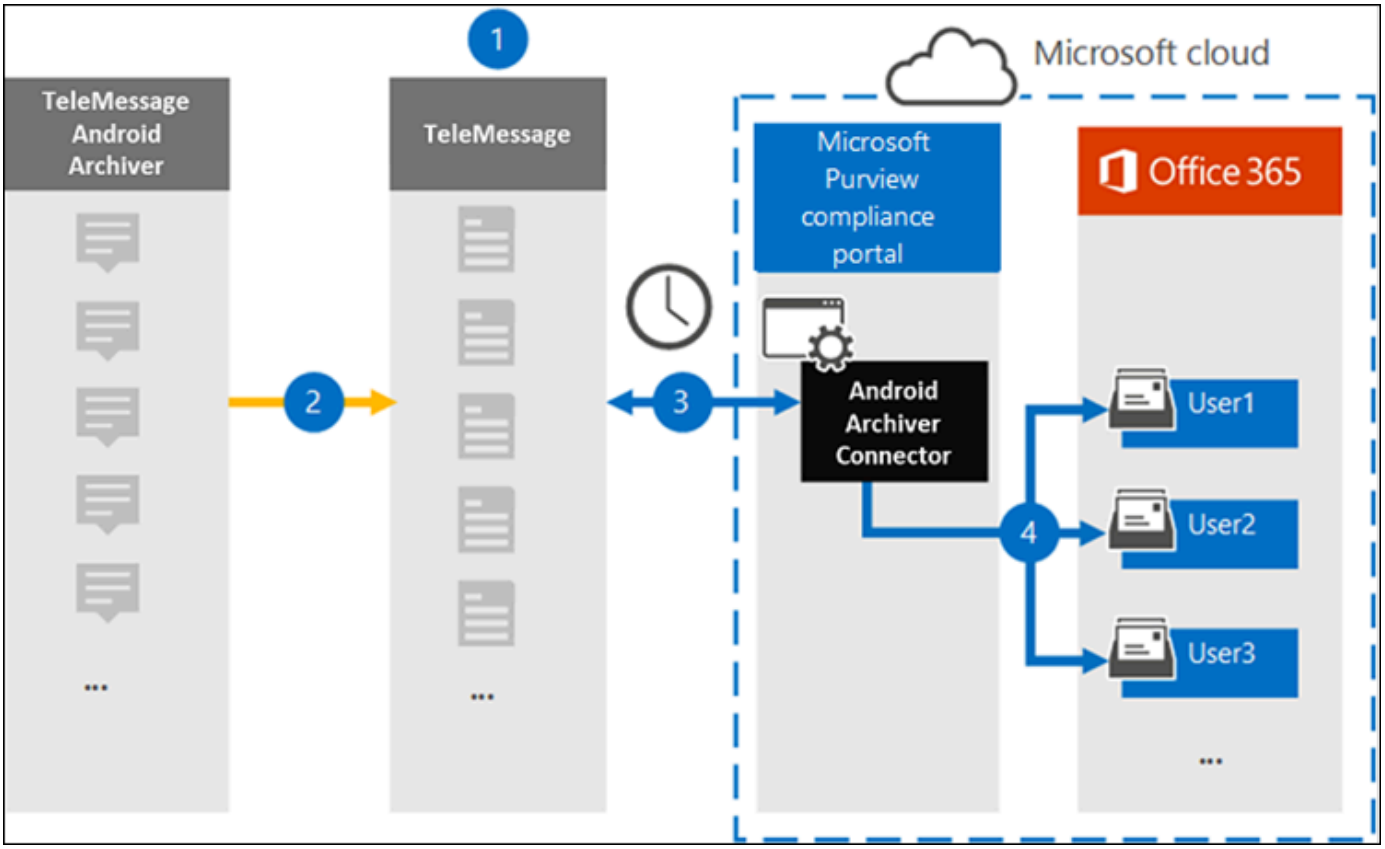
Set up a connector to archive Android mobile data

Use a TeleMessage connector to import and archive SMS, MMS, voice calls, and call logs from Android mobile phones. After you set up and configure a connector, it connects to your organization's TeleMessage account once every day and imports the mobile communication of employees using the TeleMessage Android Archiver to mailboxes in Microsoft 365.

After storing data from Android mobile phones in user mailboxes, you can apply Microsoft Purview features such as Litigation Hold, eDiscovery, and Microsoft 365 retention policies to Android Archiver data. For example, you can search Android Archiver mobile communication by using eDiscovery or associate the mailbox that contains the Android Archiver connector data with a custodian in an eDiscovery (Premium) case. Using an Android Archiver connector to import and archive data in Microsoft 365 can help your organization stay compliant with government and regulatory policies.

Overview of archiving Android mobile data

The following overview explains the process of using a connector to archive Android mobile data in Microsoft 365.



1. Your organization works with TeleMessage to set up an Android Archiver connector. For more information, see Android Archiver.

2. In real time, SMS, MMS, voice calls, and call logs from your organization's Android mobile phones are copied to the TeleMessage site.

3. The Android Archiver connector that you create in the Microsoft Purview portal connects to the TeleMessage site every day and transfers the Android data from the previous 24 hours to a secure Azure Storage location in the Microsoft cloud. The connector also converts the Android data to an email message format.

4. The connector imports the mobile communication items to the mailbox of a specific user. A new folder named Android Archiver is created in the specific user's mailbox, and the items are imported to the folder. The connector does mapping by using the value of the *User's Email address* property. Every email message contains this property, which is populated with the email address of every participant of the email message. In addition to automatic user mapping by using the value of the *User's Email address* property, you can also define a custom mapping by uploading a CSV mapping file. This mapping file should contain the mobile number and corresponding Microsoft 365 mailbox address for each user. If you enable automatic user mapping and provide a custom mapping, for every email item the connector first looks at custom mapping file. If it doesn't find a valid Microsoft 365 user that corresponds to a user's mobile number, the connector uses the user's email address property of the email item. If the connector doesn't find a valid Microsoft 365 user in either the custom mapping file or the *User's email address* property of the email item, the item won't be imported.

Before you begin

> **Important**
>
> All partner third-party data connector configurations require a provisioned service principal for the partner application. This requirement isn't applicable for native Microsoft third-party data connector configurations.
>
> To provision a service principal using PowerShell, see Provision a service principal for partner third-party data connectors.

Some of the implementation steps required to archive Android communication data are external to Microsoft 365. You must complete these steps before you can create the connector in the

Microsoft Purview portal.

- Order the Android Archiver service from TeleMessage and get a valid administration account for your organization. You need to sign in to this account when you create the connector.
- Register all users that require the Android Archiver service in the TeleMessage account. When registering users, use the same email address that's used for their Microsoft 365 account.
- Install and activate the TeleMessage Android Archiver app on the mobile phones of your employees.
- Assign the Data Connector Admin role to the user who creates a Android Archiver connector. This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions, see Permissions in the Microsoft Purview portal and Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview.
- This data connector is available in GCC environments in the Microsoft 365 US Government cloud. Third-party applications and services might involve storing, transmitting, and processing your organization's customer data on third-party systems that are outside of the Microsoft 365 infrastructure and therefore aren't covered by the Microsoft Purview and data protection commitments. Microsoft makes no representation that use of this product to connect to third-party applications implies that those third-party applications are FedRAMP compliant.

Create an Android Archiver connector

After completing the prerequisite tasks, create an Android Archiver connector in the Microsoft Purview portal. The connector uses the information you provide to connect to the TeleMessage site and transfer Android communication to the corresponding user mailboxes in Microsoft 365.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.

4. From the list, select **Android Archiver**.
5. On the **Terms of service** page, select **Accept**.
6. On the **Login to TeleMessage** page, enter the required information in the following boxes and then select **Next**.
    - **Username:** Your TeleMessage username.
    - **Password:** Your TeleMessage password.
7. After you create the connector, close the pop-up window and select **Next**.
8. On the **User mapping** page, enable automatic user mapping and select **Next**. If you need custom mapping, upload a CSV file, and select **Next**.
9. Review your settings, then select **Finish** to create the connector.

Android Archiver connector limits

The connector supports importing attachments or items up to 10 MB.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/19/2026

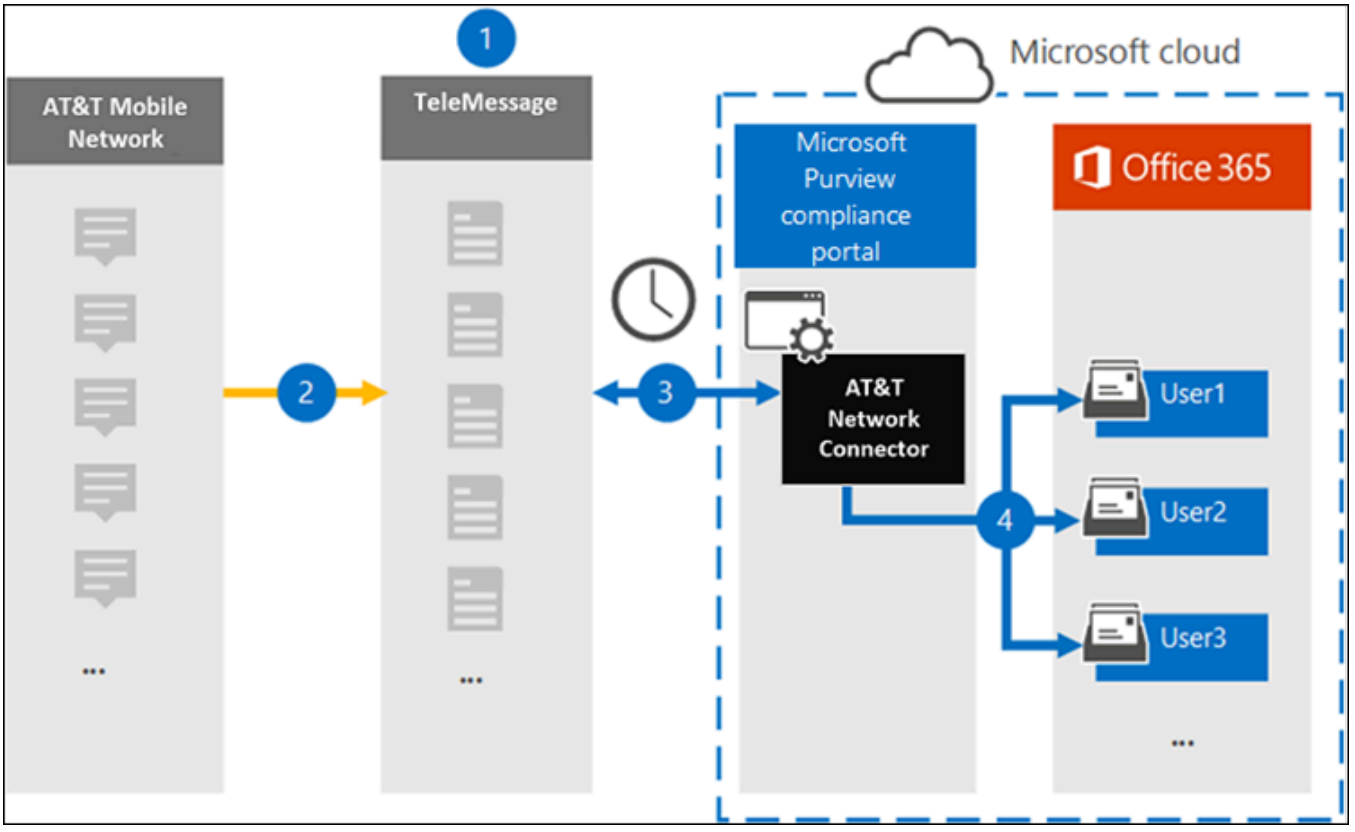
Set up a connector to archive AT&T SMS/MMS data

Use a TeleMessage connector to import and archive SMS and MMS data from the AT&T Mobile Network. After you set up and configure a connector, it connects to your organization's AT&T Network once every day and imports SMS and MMS data to mailboxes in Microsoft Purview.

After storing SMS and MMS messages in user mailboxes, you can apply Microsoft Purview features such as Litigation Hold, eDiscovery, and Microsoft 365 retention policies to AT&T Network data. For example, you can search AT&T Network data by using eDiscovery or associate the mailbox that contains the AT&T Network connector data with a custodian in an eDiscovery (Premium) case. Using an AT&T Network connector to import and archive data in Microsoft 365 can help your organization stay compliant with government and regulatory policies.

Overview of archiving AT&T Network data

The following overview explains the process of using a connector to archive AT&T Network data in Microsoft 365.



1. Your organization works with TeleMessage to set up an AT&T Network connector. For information, see AT&T Network Archiver.

2. In real time, SMS and MMS messages from your organization's AT&T Network are copied to the TeleMessage site.

3. The AT&T Network connector that you create in the Microsoft Purview portal connects to the TeleMessage site every day and transfers the SMS and MMS messages from the previous 24 hours to a secure Azure Storage location in the Microsoft cloud. The connector also converts the content of SMS and MMS messages to an email message format.

4. The connector imports the mobile communication items to the mailbox of specific users. A new folder named **AT&T SMS/MMS Network Archiver** is created in the user's mailbox and the items are imported to it. The connector does this mapping by using the value of the *User's Email address* property. Every SMS and MMS message contains this property, which is populated with the email address of every participant of the message.

   In addition to automatic user mapping by using the value of the *User's Email address* property, you can also define a custom mapping by uploading a CSV mapping file. This mapping file contains the mobile phone number and corresponding Microsoft 365 email address for users in your organization. If you enable both automatic user mapping and custom mapping, for every email item the connector first looks at the custom mapping file. If it doesn't find a valid Microsoft 365 user that corresponds to a mobile phone number, the connector uses the values in the email address property of the item it's trying to import. If the connector doesn't find a valid Microsoft 365 user in either the custom mapping file or in the email address property of the email item, the item won't be imported.

Before you begin

> **Important**
>
> All partner third-party data connector configurations require a provisioned service principal for the partner application. This requirement isn't applicable for native Microsoft third-party data connector configurations.
>
> To provision a service principal using PowerShell, see Provision a service principal for partner third-party data connectors.

Some of the implementation steps required to archive AT&T Network data are external to Microsoft 365. You must complete these steps before you can create the connector in the Microsoft Purview portal.

- Order the mobile archiver service from TeleMessage and get a valid administration account for your organization. You need to sign in to this account when you create the connector in the Microsoft Purview portal.
- Obtain your AT&T account and billing contact details so you can fill out the TeleMessage onboarding forms and order the message archiving service from AT&T.
- Register all users that require AT&T SMS/MMS Network archiving in the TeleMessage account. When registering users, use the same email address that's used for their Microsoft 365 account.
- Ensure your employees have corporate-owned and corporate-liable mobile phones on the AT&T mobile network. Archiving messages in Microsoft 365 isn't available for employee-owned or "Bring Your Own Devices (BYOD) devices.
- Assign the Data Connector Admin role to the user who creates a AT&T Network connector. This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions, see:- Permissions in the Microsoft Purview portal - Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview compliance
- This data connector is available in GCC environments in the Microsoft 365 US Government cloud. Third-party applications and services might involve storing, transmitting, and processing your organization's customer data on third-party systems that are outside of the Microsoft 365 infrastructure and therefore aren't covered by the Microsoft Purview and data protection commitments. Microsoft makes no representation that use of this product to connect to third-party applications implies that those third-party applications are FedRAMP compliant.

Create an AT&T Network connector

After you complete the prerequisites described in the previous section, you can create an AT&T Network connector in the Microsoft Purview portal. The connector uses the information you provide to connect to the TeleMessage site and transfer SMS and MMS messages to the corresponding user mailboxes in Microsoft 365.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.

3. Select **My connectors**, then select **Add connector**.
4. From the list, choose **AT&T Network**.
5. On the **Terms of service** page, select **Accept**.
6. On the **Login to TeleMessage** page, under Step 3, enter the required information in the following boxes and then select **Next**.
    - **Username:** Your TeleMessage username.
    - **Password:** Your TeleMessage password.
7. After the connector is created, you can close the pop-up window and go to the next page.
8. On the **User mapping** page, enable automatic user mapping. To enable custom mapping, upload a CSV file that contains the user mapping information, and then select **Next**.
9. Review your settings, and then select **Finish** to create the connector.

Limits

Importing attachments or items supports files up to 10 MB.

Last updated on 03/13/2026

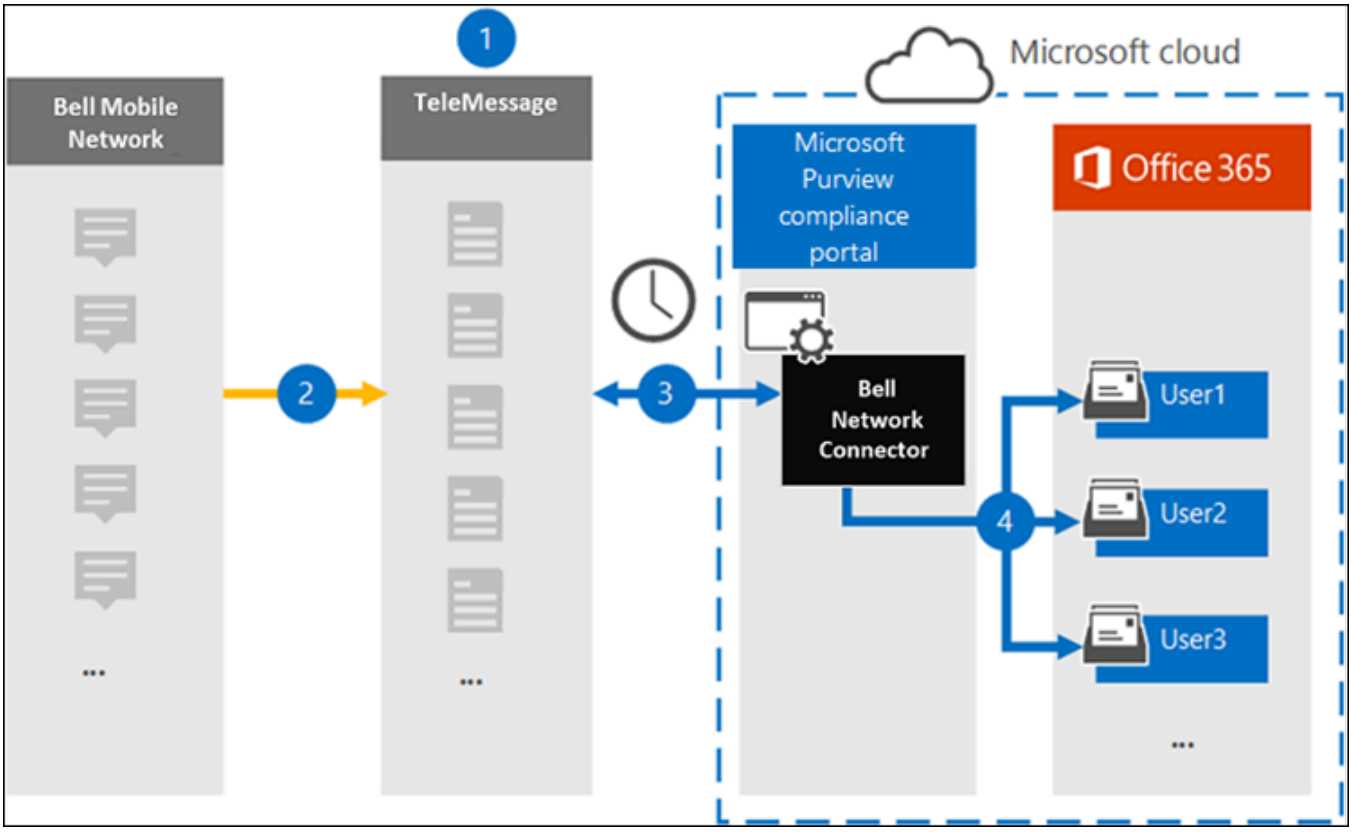
Set up a connector to archive Bell Network data

Use a TeleMessage connector to import and archive Short Messaging Service (SMS) and Multimedia Messaging Service (MMS) messages from the Bell Network. After you set up and configure a connector, it connects to your organization's Bell Network once every day, and imports SMS and MMS messages to mailboxes in Microsoft 365.

After storing the SMS and MMS messages in user mailboxes, you can apply Microsoft Purview features such as Litigation Hold, eDiscovery, and Microsoft 365 retention policies to Bell Network data. For example, you can search Bell Network SMS/MMS using eDiscovery or associate the mailbox that contains the Bell Network connector data with a custodian in an eDiscovery (Premium) case. Using a Bell Network connector to import and archive data in Microsoft 365 can help your organization stay compliant with government and regulatory policies.

Overview of archiving Bell Network data

The following overview explains the process of using a connector to archive Bell Network data in Microsoft 365.



1. Your organization works with TeleMessage and Bell to set up a Bell Network connector. For more information, see Bell Network Archiver.

2. In real time, SMS and MMS messages from your organization's Bell Network are copied to the TeleMessage site.

3. The Bell Network connector that you create in the Microsoft Purview portal connects to the TeleMessage site every day and transfers the SMS and MMS messages from the previous 24 hours to a secure Azure Storage location in the Microsoft cloud. The connector also converts the content of SMS and MMS messages to an email message format.

4. The connector imports the mobile communication items to the mailbox of specific users. A new folder named **Bell SMS/MMS Network Archiver** is created in a specific user's mailbox and the items are imported to it. The connector does this mapping by using the value of the *User's Email address* property. Every SMS and MMS message contains this property, which is populated with the email address of every participant of the message.

   In addition to automatic user mapping by using the value of the *User's Email address* property, you can also define a custom mapping by uploading a CSV mapping file. This mapping file contains the mobile phone number and corresponding Microsoft 365 email address for users in your organization. If you enable both automatic user mapping and custom mapping, for every Bell Network item the connector first looks at custom mapping file. If it doesn't find a valid Microsoft 365 user that corresponds to a user's mobile phone number, the connector uses the values in the email address property of the item it's trying to import. If the connector doesn't find a valid Microsoft 365 user in either the custom mapping file or in the email address property of the Bell Network item, the item won't be imported.

Before you begin

> **Important**
>
> All partner third-party data connector configurations require a provisioned service principal for the partner application. This requirement isn't applicable for native Microsoft third-party data connector configurations.
>
> To provision a service principal using PowerShell, see Provision a service principal for partner third-party data connectors.

Some of the implementation steps required to archive Bell Network data are external to Microsoft 365. You must complete these steps before you can create a connector in the Microsoft Purview portal.

- Order the Bell Network Archiver service from TeleMessage and get a valid administration account for your organization. You need to sign in to this account when you create the connector in the Microsoft Purview portal.
- Obtain your Bell Network account and billing contact details so you can fill out the TeleMessage onboarding forms and order the message archiving service from Bell.
- Register all users that require Bell SMS/MMS Network archiving in the TeleMessage account. When registering users, use the same email address that's used for their Microsoft 365 account.
- Ensure your employees have corporate-owned and corporate-liable mobile phones on the Bell mobile network. Archiving messages in Microsoft 365 isn't available for employee-owned or "Bring Your Own Devices (BYOD) devices.
- Assign the Data Connector Admin role to the user who creates a Bell Network connector. This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions, see:- Permissions in the Microsoft Purview portal - Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview compliance
- This data connector is available in GCC environments in the Microsoft 365 US Government cloud. Third-party applications and services might involve storing, transmitting, and processing your organization's customer data on third-party systems that are outside of the Microsoft 365 infrastructure and therefore aren't covered by the Microsoft Purview and data protection commitments. Microsoft makes no representation that use of this product to connect to third-party applications implies that those third-party applications are FedRAMP compliant.

Create a Bell Network connector

The last step is to create a Bell Network connector in the Microsoft Purview portal. The connector uses the information you provide to connect to the TeleMessage site and transfer SMS and MMS messages to the corresponding user mailboxes in Microsoft 365.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.

4. From the list, choose **Bell SMS/MMS Network Archiver**.
5. On the **Terms of service** page, select **Accept**.
6. On the **Login to TeleMessage** page, under Step 3, enter the required information in the following boxes and then select **Next**.
    - **Username:** Your TeleMessage username.
    - **Password:** Your TeleMessage password.
7. After the connector is created, you can close the pop-up window and go to the next page.
8. On the **User mapping** page, enable automatic user mapping. To enable custom mapping, upload a CSV file that contains the user mapping information, and then select **Next**.
9. Review your settings, and then select **Finish** to create the connector.

Limits

Importing attachments or items supports files up to 10 MB.

Last updated on 03/13/2026

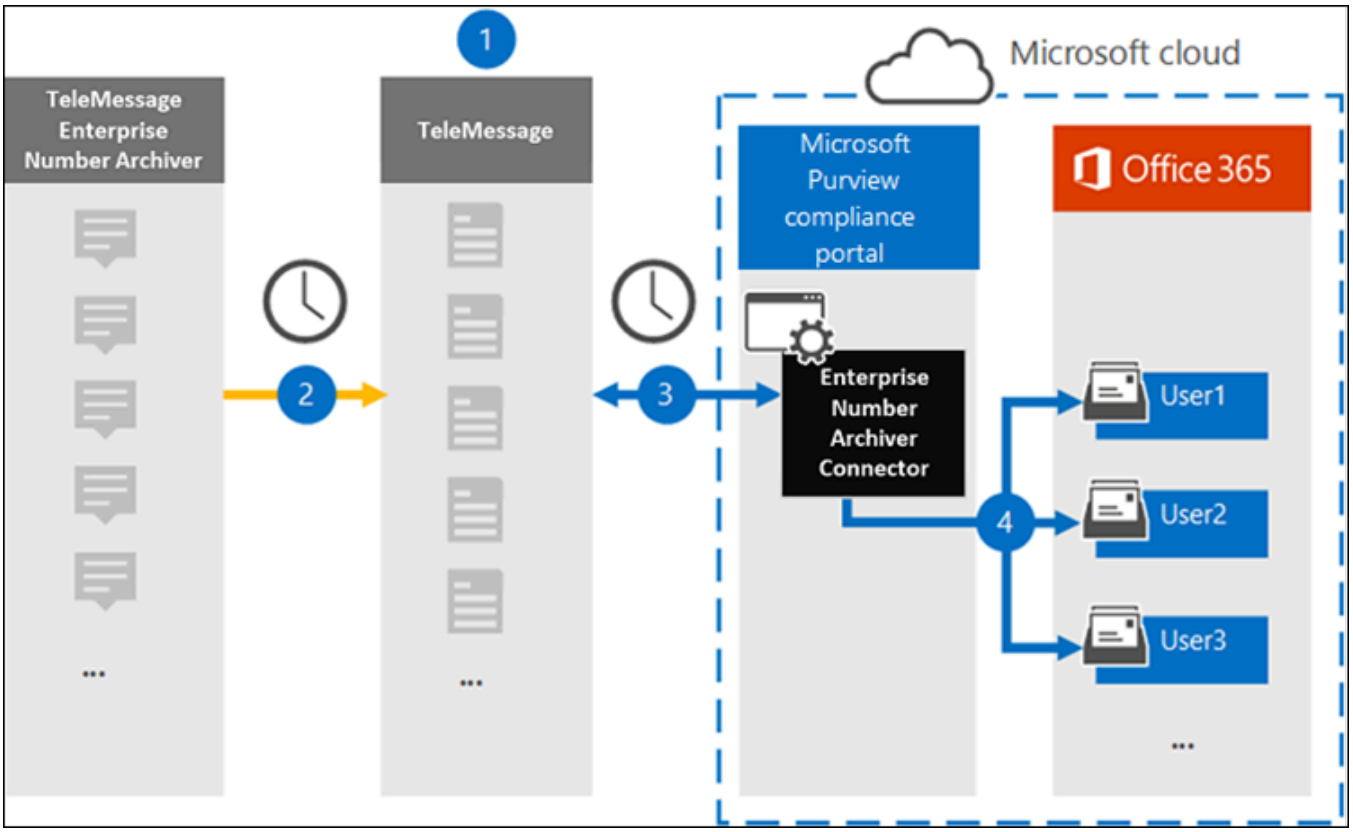
Set up a connector to archive Enterprise Number data

Use a TeleMessage connector to import and archive Short Messaging Service (SMS) and Multimedia Messaging Service (MMS) messages, chat messages, voice call recordings, and voice call logs from the Enterprise Number Archiver. After you set up and configure a connector, it connects to your organization's TeleMessage account once every day and imports the mobile communication data of employees using the TeleMessage Enterprise Number Archiver to mailboxes in Microsoft 365.

After storing the TeleMessage Enterprise Number Archiver connector data in user mailboxes, you can apply Microsoft Purview features such as Litigation Hold, eDiscovery, In-Place Archiving, Auditing, Communication Compliance, and Microsoft 365 retention policies to Enterprise Number Archiver data. For example, you can search the TeleMessage Enterprise Number Archiver SMS, MMS, and Voice Call using eDiscovery or associate the mailbox that contains the Enterprise Number Archiver connector data with a custodian in an eDiscovery (Premium) case. Using an Enterprise Number Archiver connector to import and archive data in Microsoft 365 can help your organization stay compliant with government and regulatory policies.

Overview of archiving Enterprise Number data

The following overview explains the process of using a connector to archive Enterprise Network data in Microsoft 365.



1. Your organization works with TeleMessage to set up an Enterprise Number Archiver connector. For more details refer to here.
2. The Enterprise Number Archiver connector that you create in the Microsoft Purview portal connects to the TeleMessage site every day and transfers the email messages from the previous 24 hours to a secure Azure Storage area in the Microsoft Cloud.
3. The connector imports the mobile communication items to the mailbox of a specific user. A new folder named Enterprise Number Archiver is created in the specific user's mailbox and the items are imported to it. The connector does mapping by using the value of the *User's Email address* property. Every email message contains this property, which is populated with the email address of every participant of the email message. In addition to automatic user mapping using the value of the *User's Email address* property, you can also define a custom mapping by uploading a CSV mapping file. This mapping file should contain User's mobile Number and the corresponding Microsoft 365 mailbox address for each user. If you enable automatic user mapping and provide a custom mapping, for every email item the connector will first look at custom mapping file. If it doesn't find a valid Microsoft 365 user that corresponds to a user's mobile number, the connector will use the User 's email address property of the email item. If the connector doesn't find a valid Microsoft 365 user in either the custom mapping file or the *user's email address* property of the email item, the item won't be imported.

Before you begin

Important

All partner third-party data connector configurations require a provisioned **service principal** for the partner application. This requirement isn't applicable for native Microsoft third-party data connector configurations.

To provision a service principal using PowerShell, see **Provision a service principal for partner third-party data connectors**.

Some of the implementation steps required to archive Enterprise Number Archiver data are external to Microsoft 365. You must complete these steps before you can create the connector in the Microsoft Purview portal.

- Order the Enterprise Number Archiver service from TeleMessage and get a valid administration account for your organization. You need to sign in to this account when you create the connector in the Microsoft Purview portal.
- Register all users that require Enterprise Number SMS/MMS Network archiving in the TeleMessage account. When registering users, use the same email address that's used for their Microsoft 365 account.
- Install and activate the TeleMessage Enterprise Number Archiver app on the mobile phones of your employees.
- Assign the Data Connector Admin role to the user who creates a Enterprise Number Archiver connector. This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions, see:- Permissions in the Microsoft Purview portal - Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview compliance
- This data connector is available in GCC environments in the Microsoft 365 US Government cloud. Third-party applications and services might involve storing, transmitting, and processing your organization's customer data on third-party systems that are outside of the Microsoft 365 infrastructure and therefore aren't covered by the Microsoft Purview and data protection commitments. Microsoft makes no representation that use of this product to connect to third-party applications implies that those third-party applications are FedRAMP compliant.

Create an Enterprise Number Archiver connector

After you complete the prerequisites described in the previous section, create an Enterprise Number Archiver connector in the Microsoft Purview portal. The connector uses the information you provide to connect to the TeleMessage site and transfer SMS, MMS, and voice call messages to the corresponding user mailboxes in Microsoft 365.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. From the list, choose **Enterprise Number Archiver**.
5. On the **Terms of service** page, select **Accept**.
6. On the **Login to TeleMessage** page, under Step 3, enter the required information in the following boxes and then select **Next**.
    - **Username:** Your TeleMessage username.
    - **Password:** Your TeleMessage password.
7. After the connector is created, you can close the pop-up window and go to the next page.
8. On the **User mapping** page, enable automatic user mapping. To enable custom mapping, upload a CSV file that contains the user mapping information, and then select **Next**.
9. Review your settings, and then select **Finish** to create the connector.

Limits

Importing attachments or items supports files up to 10 MB.

Last updated on 03/13/2026

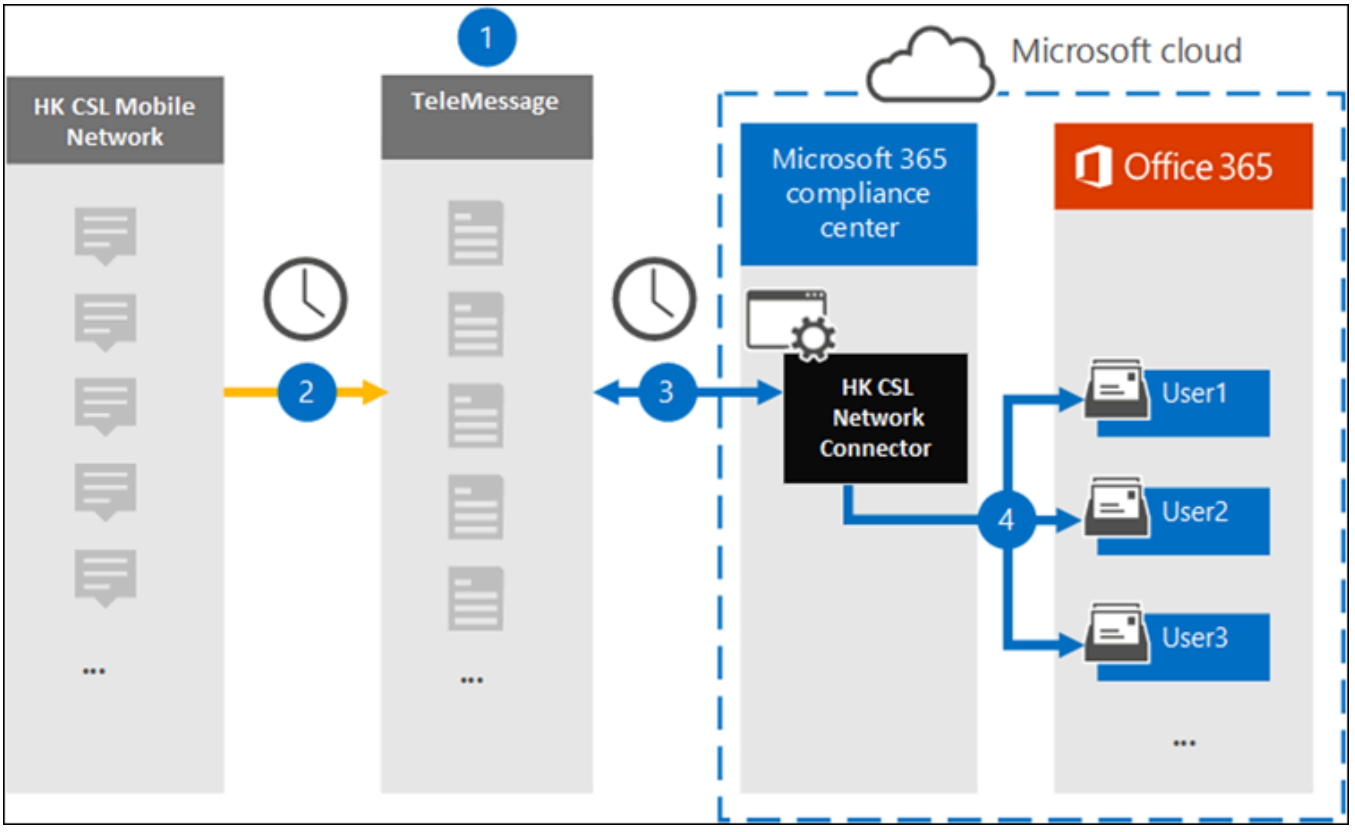
Set up a connector to archive Hong Kong CSL SMS/MMS data

Use a TeleMessage connector in the Microsoft Purview portal to import and archive SMS and MMS data from the Hong Kong (HK) CSL Mobile Network. After you set up and configure a connector, it connects to your organization's HK CSL Network once every day, and imports SMS and MMS data to mailboxes in Microsoft Purview.

After storing SMS and MMS messages in user mailboxes, you can apply Microsoft Purview features such as Litigation Hold and Microsoft 365 retention policies to HK CSL Network data. For example, you can search HK CSL Network data or associate the mailbox that contains the HK CSL Network connector data with a custodian in an eDiscovery case. Using a HK CSL Network connector to import and archive data in Microsoft 365 can help your organization stay compliant with government and regulatory policies.

Overview of archiving HK CSL Network data

The following overview explains the process of using a connector to archive HK CSL Network data in Microsoft 365.



1. Your organization works with TeleMessage to set up a HK CSL Network connector. For information, see HK CSL Network Archiver.

2. In real time, SMS and MMS messages from your organization's HK CSL Network are copied to the TeleMessage site.
3. The HK CSL Network connector that you create in the Microsoft Purview portal connects to the TeleMessage site every day and transfers the SMS and MMS messages from the previous 24 hours to a secure Azure Storage location in the Microsoft cloud. The connector also converts the content of SMS and MMS messages to an email message format.
4. The connector imports the mobile communication items to the mailbox of specific users. A new folder named **HK CSL SMS/MMS Network Archiver** is created in the user's mailbox and the items are imported to it. The connector does this mapping by using the value of the *User's Email address* property. Every SMS and MMS message contains this property, which is populated with the email address of every participant of the message.

In addition to automatic user mapping by using the value of the *User's Email address* property, you can also define a custom mapping by uploading a CSV mapping file. This mapping file contains the mobile phone number and corresponding Microsoft 365 email address for users in your organization. If you enable both automatic user mapping and custom mapping, for every email item the connector first looks at the custom mapping file. If it doesn't find a valid Microsoft 365 user that corresponds to a mobile phone number, the connector uses the values in the email address property of the item it's trying to import. If the connector doesn't find a valid Microsoft 365 user in either the custom mapping file or in the email address property of the email item, the item won't be imported.

Before you begin

> **Important**
>
> All partner third-party data connector configurations require a provisioned service principal for the partner application. This requirement isn't applicable for native Microsoft third-party data connector configurations.
>
> To provision a service principal using PowerShell, see Provision a service principal for partner third-party data connectors.

Some of the implementation steps required to archive HK CSL Network data are external to Microsoft 365. You must complete these steps before you can create the connector in the Microsoft Purview portal.

- Order the mobile archiver service from TeleMessage and get a valid administration account for your organization. You need to sign in to this account when you create the connector in the Pruview portal.

- Obtain your HK CSL account and billing contact details so you can fill out the TeleMessage onboarding forms and order the message archiving service from HK CSL.
- Register all users that require HK CSL SMS/MMS Network archiving in the TeleMessage account. When registering users, be sure to use the same email address that's used for their Microsoft 365 account.
- Ensure users have corporate-owned and corporate-liable mobile phones on the HK CSL mobile network. Archiving messages in Microsoft 365 isn't available for employee-owned or "Bring Your Own Devices (BYOD) devices.
- Assign the Data Connector Admin role to the user who creates a HK CSL Network connector. This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions, see:- Permissions in the Microsoft Purview portal - Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview compliance

This data connector is available in GCC environments in the Microsoft 365 US Government cloud. Third-party applications and services might involve storing, transmitting, and processing your organization's customer data on third-party systems that are outside of the Microsoft 365 infrastructure and therefore aren't covered by the Microsoft Purview and data protection commitments. Microsoft makes no representation that use of this product to connect to third-party applications implies that those third-party applications are FedRAMP compliant.

Create an HK CSL Network connector

After you complete the prerequisites described in the previous section, you can create an HK CSL Network connector in the Microsoft Purview portal. The connector uses the information you provide to connect to the TeleMessage site and transfer SMS and MMS messages to the corresponding user mailboxes in Microsoft 365.

1. Go to the Microsoft Purview portal and select **Data connectors** > **HK CSL Network**.

2. On the **HK CSL Network product** description page, select **Add connector**.

3. On the **Terms of service** page, select **Accept**.

4. On the **Login to TeleMessage** page, under Step 3, enter the required information in the following boxes and select **Next**.

    - **Username:** Your TeleMessage username.
    - **Password:** Your TeleMessage password.

5. After you create the connector, you can close the pop-up window and go to the next page.
6. On the **User mapping** page, enable automatic user mapping. To enable custom mapping, upload a CSV file that contains the user mapping information, and select **Next**.
7. Review your settings, and select **Finish** to create the connector.
8. Go to the **Connectors** tab on the **Data connectors** page in the Microsoft Purview portal to see the progress of the import process for the new connector.

Limits

Importing attachments or items supports files up to 10 MB.

Last updated on 03/13/2026

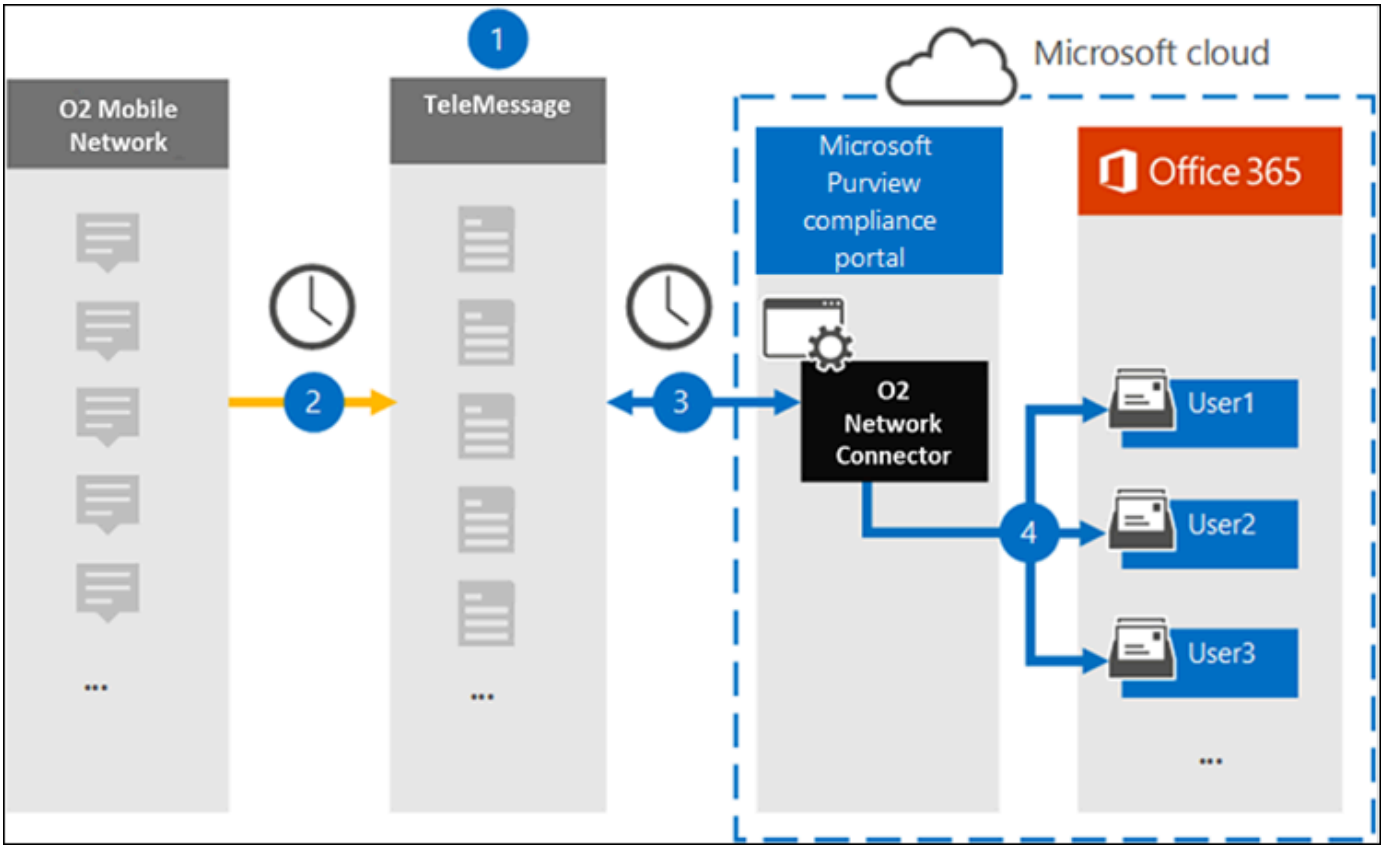
Set up a connector to archive O2 Network data

Use a TeleMessage connector to import and archive Short Messaging Service (SMS) messages and voice calls from the O2 mobile network. After you set up and configure a connector, it connects to your organization's O2 Network once every day, and imports SMS and voice calls to mailboxes in Microsoft 365.

After you store SMS messages and voice calls in user mailboxes, you can apply Microsoft Purview features such as Litigation Hold, eDiscovery, and Microsoft 365 retention policies to O2 Network data. For example, you can search O2 Network SMS messages and voice calls by using eDiscovery or associate the mailbox that contains O2 Network data with a custodian in an eDiscovery (Premium) case. Using an O2 Network connector to import and archive data in Microsoft 365 can help your organization stay compliant with government and regulatory policies.

Overview of archiving O2 Network data

The following overview explains the process of using a connector to archive O2 Network data in Microsoft 365.



1. Your organization works with TeleMessage and O2 to set up an O2 Network connector. For more information, see O2 Network Archiver.

2. Once every 24 hours, the TeleMessage site copies SMS messages and voice calls from your organization's O2 Network.

3. The O2 Network connector that you create in the Microsoft Purview portal connects to the TeleMessage site every day and transfers the SMS messages and voice calls from the previous 24 hours to a secure Azure Storage location in the Microsoft cloud. The connector also converts the content of SMS messages and voice calls to an email message format.

4. The connector imports the mobile communication items to the mailbox of specific users. A new folder named **O2 SMS and Voice Network Archiver** is created in a specific user's mailbox and the items are imported to it. The connector does this mapping by using the value of the *User's Email address* property. Every SMS message and voice call contains this property, which is populated with the email address of every participant of the message.

   In addition to automatic user mapping by using the value of the *User's Email address* property, you can also define a custom mapping by uploading a CSV mapping file. This mapping file contains the mobile phone number and corresponding Microsoft 365 email address for users in your organization. If you enable both automatic user mapping and custom mapping, for every O2 item the connector first looks at custom mapping file. If it doesn't find a valid Microsoft 365 user that corresponds to a user's mobile phone number, the connector uses the values in the email address property of the item it's trying to import. If the connector doesn't find a valid Microsoft 365 user in either the custom mapping file or in the email address property of the O2 item, the item won't be imported.

Before you begin

> **Important**
>
> All partner third-party data connector configurations require a provisioned service principal for the partner application. This requirement isn't applicable for native Microsoft third-party data connector configurations.
>
> To provision a service principal using PowerShell, see Provision a service principal for partner third-party data connectors.

Some of the implementation steps required to archive O2 Network data are external to Microsoft 365. You must complete these steps before you can create a connector in the Microsoft Purview portal.

- Order the O2 Network Archiver service from TeleMessage and get a valid administration account for your organization. You need to sign in to this account when you create the connector in the Microsoft Purview portal.
- Get your O2 Network account and billing contact details so you can fill out the TeleMessage onboarding forms and order the message archiving service from O2.
- Register all users that require O2 SMS and Voice Network archiving in the TeleMessage account. When registering users, be sure to use the same email address that's used for their Microsoft 365 account.
- Ensure your employees have corporate-owned and corporate-liable mobile phones on the O2 mobile network. Archiving messages in Microsoft 365 isn't available for employee-owned or "Bring Your Own Devices (BYOD) devices.
- Assign the Data Connector Admin role to the user who creates an O2 Network connector. This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions, see Permissions in the Microsoft Purview portal and Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview.
- This data connector is available in GCC environments in the Microsoft 365 US Government cloud. Third-party applications and services might involve storing, transmitting, and processing your organization's customer data on third-party systems that are outside of the Microsoft 365 infrastructure and therefore aren't covered by the Microsoft Purview and data protection commitments. Microsoft makes no representation that use of this product to connect to third-party applications implies that those third-party applications are FedRAMP compliant.

Create an O2 Network connector

After you complete the prerequisites described in Before you begin, you can create an O2 Network connector in the Microsoft Purview portal. The connector uses the information you

provide to connect to the TeleMessage site and transfer SMS messages and voice calls to the corresponding user mailboxes in Microsoft 365.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. Select **O2 Network** from the list.
5. On the **Terms of service** page, select **Accept**.
6. On the **Login to TeleMessage** page, under Step 3, enter the required information in the following boxes and then select **Next**.
    - **Username:** Your TeleMessage username.
    - **Password:** Your TeleMessage password.
7. After the connector is created, you can close the pop-up window and go to the next page.
8. On the **User mapping** page, enable automatic user mapping and select **Next**. If you need custom mapping, upload a CSV file, and select **Next**.
9. Review your settings, and then select **Finish** to create the connector.

O2 Network connector limits

Importing attachments or items supports files up to 10 MB.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/19/2026

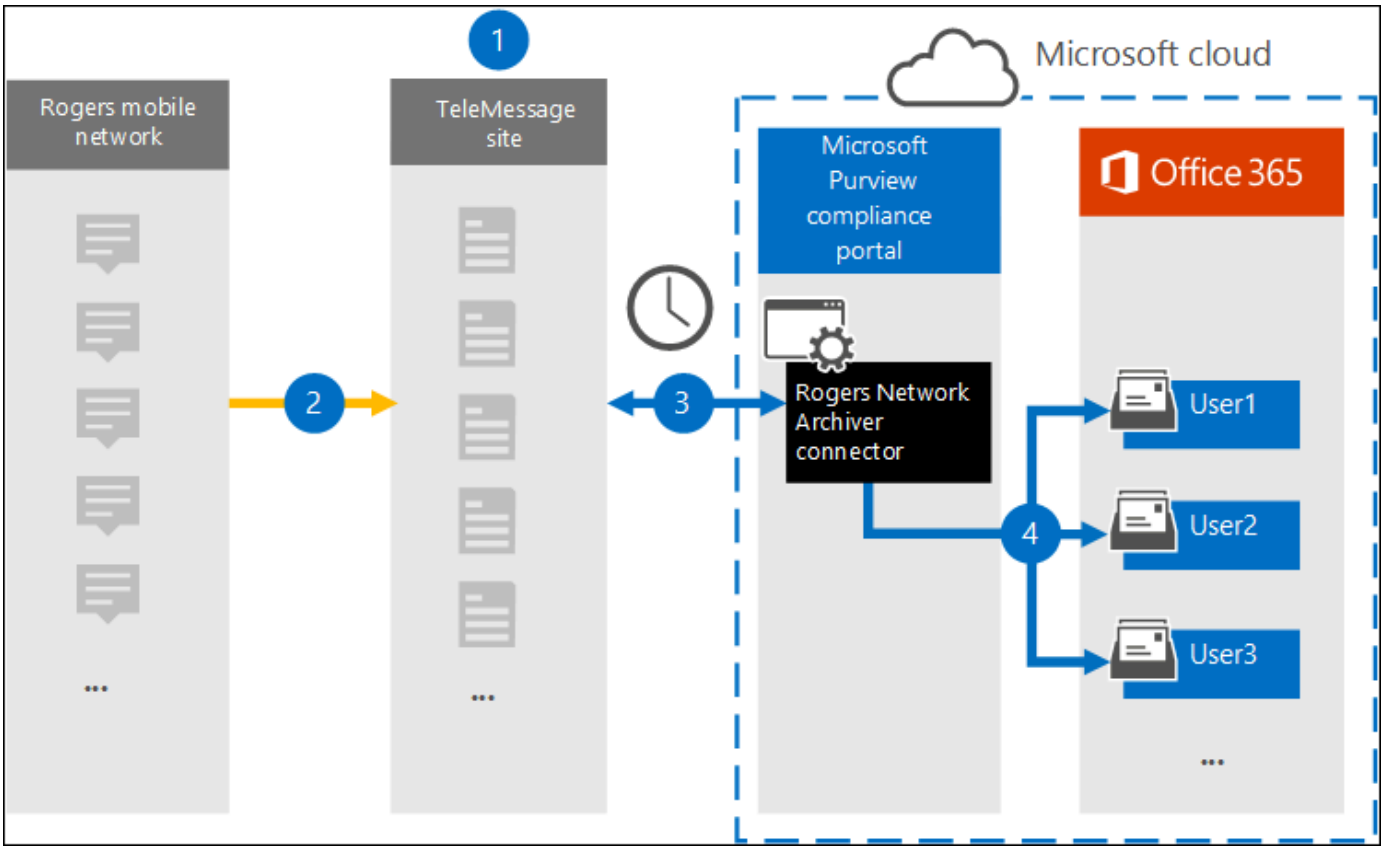
Set up a connector to archive Rogers Network data

Use the TeleMessage connector to import and archive SMS and MMS data from the Rogers mobile network. When you set up and configure a Rogers Network Archiver connector, it connects to your organization's Rogers mobile network and imports SMS and MMS data to mailboxes in Microsoft 365. Before you set up the connector, you need a TeleMessage account, Rogers account details, and the Data Connector Admin role. For the full list of prerequisites, see Before you begin.

After storing data from the Rogers mobile network in user mailboxes, you can apply Microsoft Purview features such as Litigation Hold, eDiscovery, and Microsoft 365 retention policies to the data. For example, you can search for SMS and MMS messages from the Rogers mobile network by using eDiscovery or a search associated with a Microsoft Purview eDiscovery case. Using a Rogers Network Archiver connector to import and archive data in Microsoft 365 can help your organization stay compliant with corporate governance regulations and regulatory policies.

Overview of archiving Rogers mobile network data

The following overview explains the process of using a connector to archive Rogers SMS and MMS data in Microsoft 365.



1. Your organization works with TeleMessage to set up a Rogers Network Archiver connector. For more information, see Activating the TeleMessage Rogers Network Archiver for Microsoft 365.

2. In real time, your organization's Rogers mobile network copies data to the TeleMessage site.

3. The Rogers Network Archiver connector that you create in the Microsoft Purview portal connects to the TeleMessage site every day and transfers the email messages from the previous 24 hours to a secure Azure Storage area in the Microsoft Cloud.

4. The connector imports the mobile communication items to the mailbox of a specific user. A new folder named Rogers SMS/MMS Network Archiver is created in the specific user's mailbox and the items are imported to it. The connector does the mapping by using the value of the *User's Email address* property. Every email message contains this property, which is populated with the email address of every participant of the email message.

   In addition to automatic user mapping by using the value of the *User's Email address* property, you can also define a custom mapping by uploading a CSV mapping file. This mapping file should contain User's mobile Number and the corresponding Microsoft 365 mailbox address for each user. If you enable automatic user mapping and provide a custom mapping, for every email item the connector first looks at custom mapping file. If it doesn't find a valid Microsoft 365 user that corresponds to a user's mobile number, the connector uses the user's email address property of the email item. If the connector doesn't find a

valid Microsoft 365 user in either the custom mapping file or the *user's email address* property of the email item, the item isn't imported.

Before you begin

> **Important**
>
> All partner third-party data connector configurations require a provisioned service principal for the partner application. This requirement isn't applicable for native Microsoft third-party data connector configurations.
>
> To provision a service principal using PowerShell, see Provision a service principal for partner third-party data connectors.

- Order the Rogers Network Archiver service from TeleMessage and get a valid administration account for your organization. You need to sign in to this account when you create the connector in the Microsoft Purview portal.
- Register all users that require Rogers Network archiving in the TeleMessage account. When registering users, use the same email address that's used for their Microsoft 365 account.
- Ensure your employees have corporate-owned and corporate-liable mobile phones on the O2 mobile network. Archiving messages in Microsoft 365 isn't available for employee-owned or "Bring Your Own Devices (BYOD) devices.
- Get the Rogers account and billing contact details for your organization so that you can complete the onboarding forms and order the message archiving service from Rogers.
- Assign the Data Connector Admin role to the user who creates the Rogers Network Archiver connector in Create a Rogers Network Archiver connector. This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions, see Permissions in the Microsoft Purview portal and Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview.
- This data connector is available in GCC environments in the Microsoft 365 US Government cloud. Third-party applications and services might involve storing, transmitting, and

processing your organization's customer data on third-party systems that are outside of the Microsoft 365 infrastructure and therefore aren't covered by the Microsoft Purview and data protection commitments. Microsoft makes no representation that use of this product to connect to third-party applications implies that those third-party applications are FedRAMP compliant.

Create a Rogers Network Archiver connector

After you complete the prerequisites described in Before you begin, you can create the Rogers Network Archiver connector in the Microsoft Purview portal. The connector uses the information you provide to connect to the TeleMessage site and transfer Rogers SMS/MMS data to the corresponding user mailboxes in Microsoft 365.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. From the list, select **Rogers Network Archiver**.
5. On the **Terms of service** page, select **Accept**.
6. On the **Login to TeleMessage** page, in the credentials section of the login form, enter the required information in the following boxes and then select **Next**.
   - **Username:** Your TeleMessage username.
   - **Password:** Your TeleMessage password.
7. After the connector is created, you can close the pop-up window and go to the next page.
8. On the **User mapping** page, enable automatic user mapping. To enable custom mapping, upload a CSV file that contains the user mapping information, and then select **Next**.
9. Review your settings, and then select **Finish** to create the connector.

Rogers Network Archiver limits

Importing attachments or items supports files up to 10 MB.

Note: The author created this article with assistance from AI. Learn more

Last updated on 06/24/2026

Set up a connector to archive Signal communications data

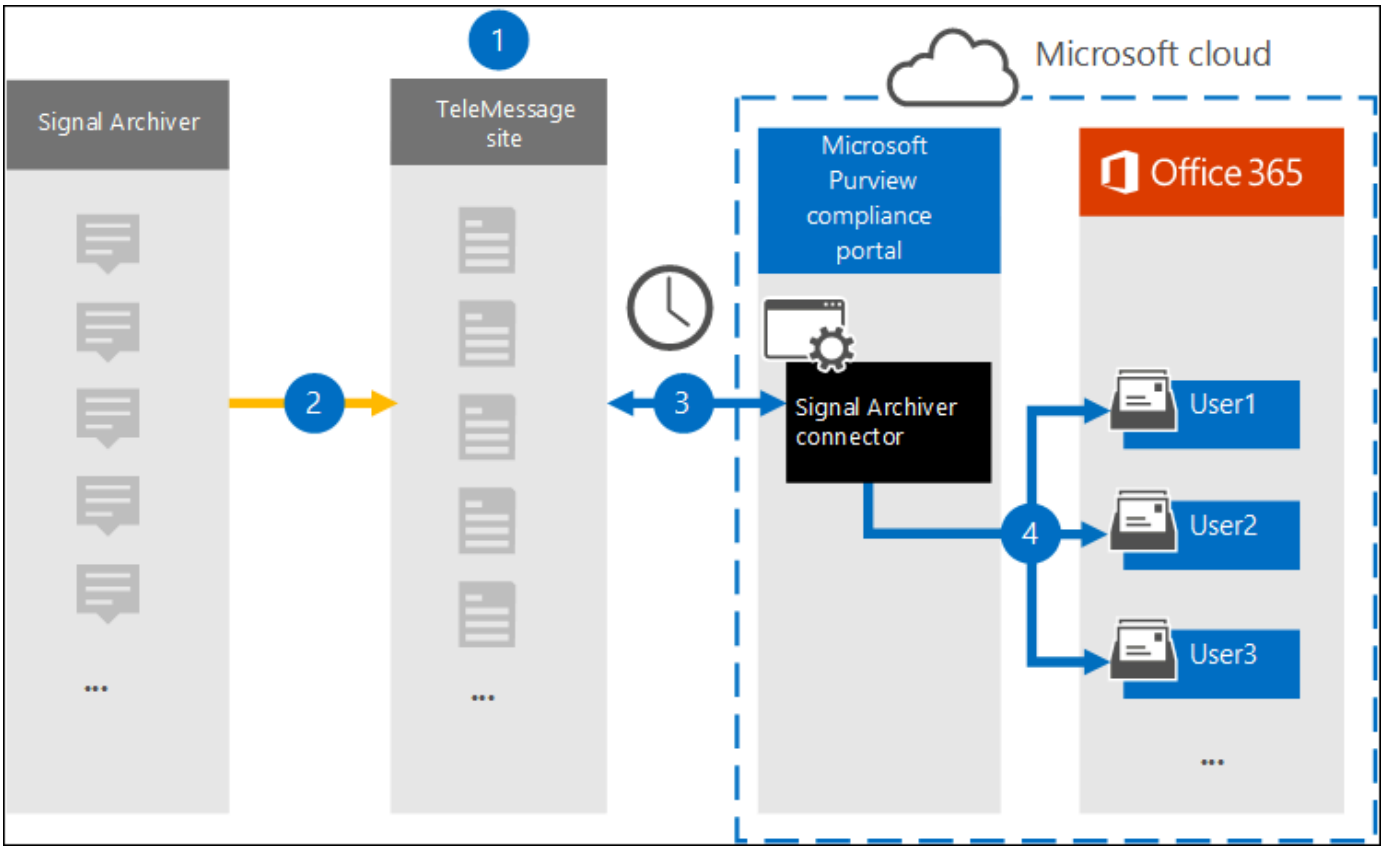
Signal Archiver connector for TeleMessage

Use the TeleMessage connector to import and archive Signal chats, attachments, files, and deleted messages and calls. After you set up and configure a connector, it connects to your organization's TeleMessage account and imports the mobile communication of employees using the TeleMessage Signal Archiver to mailboxes in Microsoft 365.

After Signal Archiver connector data is stored in user mailboxes, you can apply Microsoft Purview features such as Litigation Hold, eDiscovery, and Microsoft 365 retention policies to Signal communications data. For example, you can search Signal communication by using eDiscovery or associate the mailbox that contains the Signal Archiver connector data with a custodian in an eDiscovery (Premium) case. Using a Signal Archiver connector to import and archive data in Microsoft 365 can help your organization stay compliant with corporate governance regulations and regulatory policies.

Overview of archiving Signal communications data

The following overview explains the process of using a connector to archive Signal communication data in Microsoft 365.



1. Your organization works with TeleMessage to set up a Signal Archiver connector. For more information, see Activating the TeleMessage Signal Archiver for Microsoft 365.

2. In real time, your organization's Signal data is copied to the TeleMessage site.

3. The Signal Archiver connector that you create in the Microsoft Purview portal connects to the TeleMessage site every day and transfers the email messages from the previous 24 hours to a secure Azure Storage area in the Microsoft Cloud.

4. The connector imports the mobile communication items to the mailbox of a specific user. A new folder named Signal Archiver is created in the specific user's mailbox and the items are imported to it. The connector uses the value of the *User's Email address* property for mapping. Every email message contains this property, which is populated with the email address of every participant of the email message.

   In addition to automatic user mapping by using the value of the *User's Email address* property, you can also define a custom mapping by uploading a CSV mapping file. This mapping file should contain User's mobile Number and the corresponding Microsoft 365 mailbox address for each user. If you enable automatic user mapping and provide a custom mapping, for every email item the connector first looks at custom mapping file. If it doesn't find a valid Microsoft 365 user that corresponds to a user's mobile number, the connector uses the User's email address property of the email item. If the connector doesn't find a

valid Microsoft 365 user in either the custom mapping file or the *user's email address* property of the email item, the item won't be imported.

Before you begin

> **Important**
>
> All partner third-party data connector configurations require a provisioned service principal for the partner application. This requirement isn't applicable for native Microsoft third-party data connector configurations.
>
> To provision a service principal using PowerShell, see Provision a service principal for partner third-party data connectors.

- Order the Signal Archiver service from TeleMessage and get a valid administration account for your organization. You need to sign in to this account when you create the connector in the Microsoft Purview portal.
- Register all users that require Signal archiving in the TeleMessage account. When registering users, use the same email address that's used for their Microsoft 365 account.
- Install the Signal Archiver app on the mobile phones of your employees and activate it. The Signal Archiver app allows them to communicate and chat with other Signal users.
- Assign the Data Connector Admin role to the user who creates a Signal Archiver connector in Step 3. This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions, see Permissions in the Microsoft Purview portal and Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview.
- This data connector is available in GCC environments in the Microsoft 365 US Government cloud. Third-party applications and services might involve storing, transmitting, and processing your organization's customer data on third-party systems that are outside of the Microsoft 365 infrastructure and therefore aren't covered by the Microsoft Purview and data protection commitments. Microsoft makes no representation that use of this product to

connect to third-party applications implies that those third-party applications are FedRAMP compliant.

Create a Signal Archiver connector

After you complete the prerequisites described in Before you begin, you can create the Signal Archiver connector in the Microsoft Purview portal. The connector uses the information you provide to connect to the TeleMessage site and transfers Signal communications data to the corresponding user mailboxes in Microsoft 365.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. From the list, select **Signal Archiver**.
5. On the **Terms of service** page, select **Accept**.
6. On the **Login to TeleMessage** page, under Step 3, enter the required information in the following boxes and then select **Next**.
    - **Username:** Your TeleMessage username.
    - **Password:** Your TeleMessage password.
7. After the connector is created, close the pop-up window and go to the **User mapping** page.
8. On the **User mapping** page, enable automatic user mapping. To enable custom mapping, upload a CSV file that contains the user mapping information, and then select **Next**.
9. Review your settings, and then select **Finish** to create the connector.

Signal Archiver connector limits

Importing attachments or items supports files up to 10 MB.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 06/19/2026

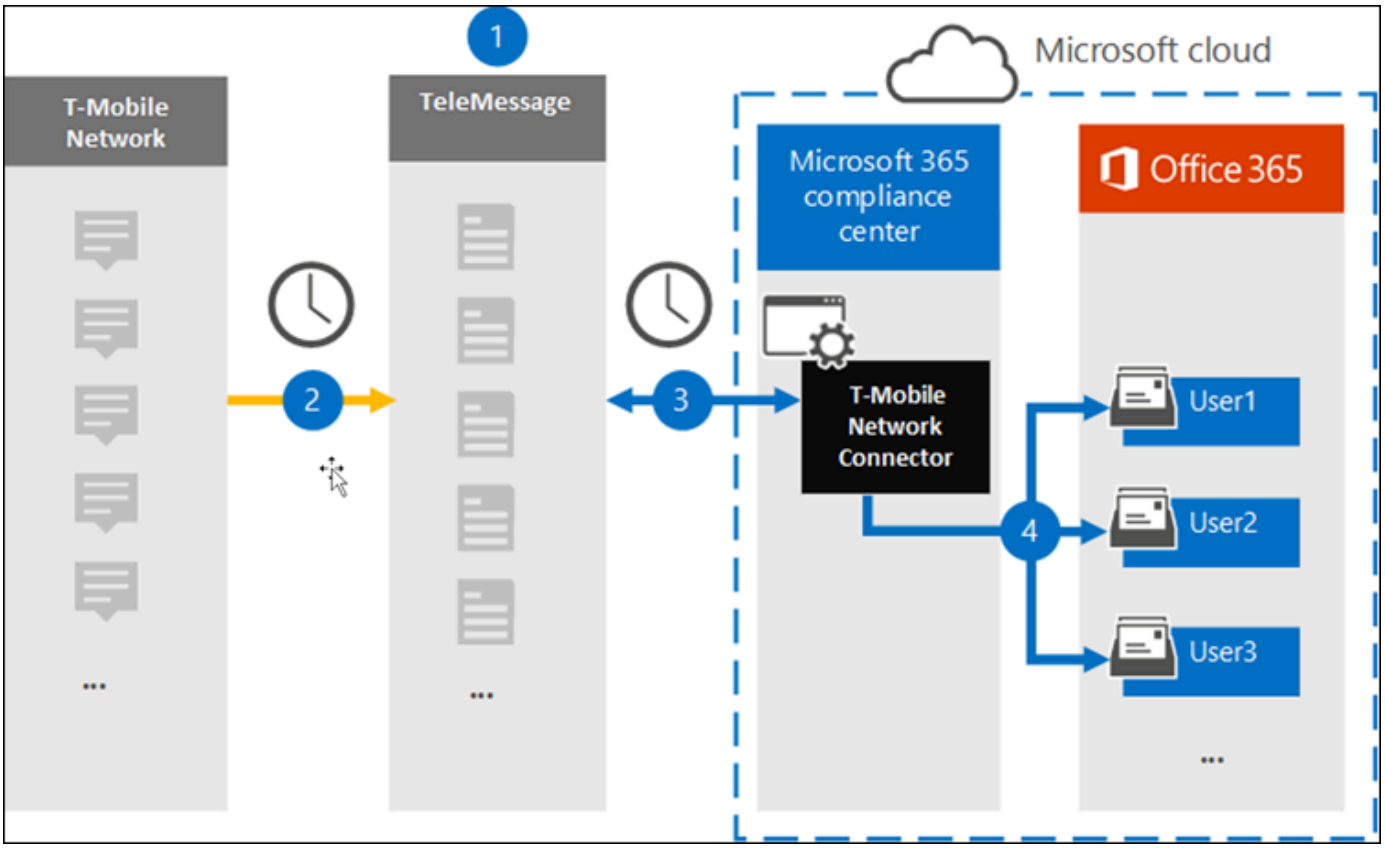
Set up a connector to archive T-Mobile SMS/MMS data

Use a TeleMessage connector in the Microsoft Purview portal to import and archive SMS and MMS data from T-Mobile Mobile Network. After you set up and configure a connector, it connects to your organization's T-Mobile Network once every day, and imports SMS and MMS data to mailboxes in Microsoft Purview. Before you begin, you'll need a TeleMessage account, T-Mobile account details, corporate-owned devices on the T-Mobile network, and the appropriate Microsoft Purview permissions. For the full list of prerequisites, see Before you begin.

After storing SMS and MMS messages in user mailboxes, you can apply Microsoft Purview features such as Litigation Hold, eDiscovery, and Microsoft 365 retention policies to T-Mobile Network data. For example, you can search T-Mobile Network data by using eDiscovery or associate the mailbox that contains the T-Mobile Network connector data with a custodian in an eDiscovery (Premium) case. Using a T-Mobile Network connector to import and archive data in Microsoft 365 can help your organization stay compliant with government and regulatory policies.

Overview of archiving T-Mobile Network data

The following overview explains the process of using a connector to archive T-Mobile Network data in Microsoft 365.



1. Your organization works with TeleMessage to set up a T-Mobile Network connector. For information, see T-Mobile Network Archiver.
2. In real time, SMS and MMS messages from your organization's T-Mobile Network are copied to the TeleMessage site.
3. The T-Mobile Network connector that you create in the Microsoft Purview portal connects to the TeleMessage site every day and transfers the SMS and MMS messages from the previous 24 hours to a secure Azure Storage location in the Microsoft cloud. The connector also converts the content of SMS and MMS messages to an email message format.
4. The connector imports the mobile communication items to the mailbox of specific users. A new folder named **T-Mobile SMS/MMS Network Archiver** is created in the user's mailbox and the items are imported to it. The connector does this mapping by using the value of the *User's Email address* property. Every SMS and MMS message contains this property, which is populated with the email address of every participant of the message.

In addition to automatic user mapping by using the value of the *User's Email address* property, you can also define a custom mapping by uploading a CSV mapping file. This mapping file contains the mobile phone number and corresponding Microsoft 365 email address for users in your organization. If you enable both automatic user mapping and custom mapping, for every email item the connector first looks at the custom mapping file. If it doesn't find a valid Microsoft 365 user that corresponds to a mobile phone number, the connector uses the values in the email address property of the item it's trying to import. If the connector doesn't find a valid Microsoft

365 user in either the custom mapping file or in the email address property of the email item, the item won't be imported.

Before you begin

> **Important**
>
> All partner third-party data connector configurations require a provisioned service principal for the partner application. This requirement isn't applicable for native Microsoft third-party data connector configurations.
>
> To provision a service principal using PowerShell, see Provision a service principal for partner third-party data connectors.

Some of the implementation steps required to archive T-Mobile Network data are external to Microsoft 365. You must complete these steps before you can create the connector in the Microsoft Purview portal.

- Order the mobile archiver service from TeleMessage and get a valid administration account for your organization. You need to sign in to this account when you create the connector in the Microsoft Purview portal.
- Obtain your T-Mobile account and billing contact details so you can fill out the TeleMessage onboarding forms and order the message archiving service from T-Mobile.
- Register all users that require T-Mobile SMS/MMS Network archiving in the TeleMessage account. When registering users, be sure to use the same email address that's used for their Microsoft 365 account.
- Ensure users have corporate-owned and corporate-liable mobile phones on the T-Mobile mobile network. Archiving messages in Microsoft 365 isn't available for user-owned or "Bring Your Own Devices (BYOD)" devices.
- Assign the Data Connector Admin role to the user who creates a T-Mobile Network connector. This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions, see Permissions in the Microsoft Purview portal and Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview.

This data connector is available in GCC environments in the Microsoft 365 US Government cloud. Third-party applications and services might involve storing, transmitting, and processing your organization's customer data on third-party systems that are outside of the Microsoft 365 infrastructure and therefore aren't covered by the Microsoft Purview and data protection commitments. Microsoft makes no representation that use of this product to connect to third-party applications implies that those third-party applications are FedRAMP compliant.

Create a T-Mobile Network connector

After you complete the prerequisites described in Before you begin, you can create a T-Mobile Network connector in the Microsoft Purview portal. The connector uses the information you provide to connect to the TeleMessage site and transfer SMS and MMS messages to the corresponding user mailboxes in Microsoft 365.

1. Go to Microsoft Purview portal and then select **Data connectors** > **T-Mobile Network**.
2. On the **T-Mobile Network product** description page, select **Add connector**.
3. On the **Terms of service** page, select **Accept**.
4. On the **Login to TeleMessage** page, under Step 3, enter the required information in the following boxes and then select **Next**.
    - **Username:** Your TeleMessage username.
    - **Password:** Your TeleMessage password.
5. After the connector is created, you can close the pop-up window and go to the next page.
6. On the **User mapping** page, enable automatic user mapping. To enable custom mapping, upload a CSV file that contains the user mapping information, and then select **Next**.
7. Review your settings, and then select **Finish** to create the connector.
8. Go to the **Connectors** tab on the **Data connectors** page in the Microsoft Purview portal to see the progress of the import process for the new connector.

T-Mobile connector limits

Importing attachments or items supports files up to 10 MB.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

Last updated on 06/24/2026

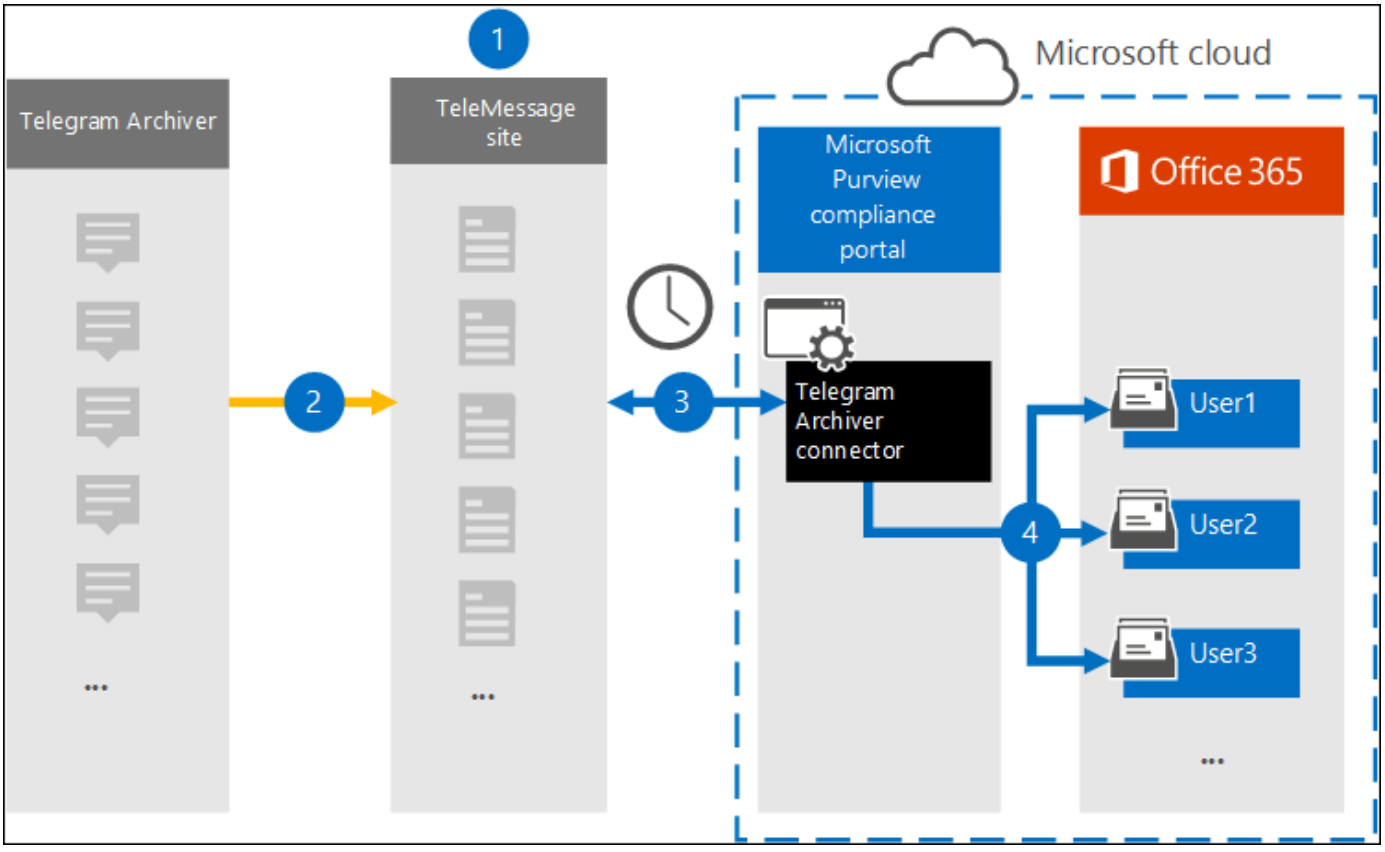
Set up a connector to archive Telegram communications data

Use the TeleMessage connector to import and archive Telegram chats, attachments, files, and deleted messages and calls. After you set up and configure a connector, it connects to your organization's TeleMessage account and imports the mobile communication of employees using the Telegram Archiver to mailboxes in Microsoft 365.

After storing Telegram Archiver connector data in user mailboxes, you can apply Microsoft Purview features such as Litigation Hold, eDiscovery, and Microsoft 365 retention policies to Telegram communication data. For example, you can search Telegram communication by using eDiscovery or associate the mailbox that contains the Telegram Archiver connector data with a custodian in an eDiscovery (Premium) case. By using a Telegram Archiver connector to import and archive data in Microsoft 365, your organization can stay compliant with corporate governance regulations and regulatory policies.

Overview of archiving Telegram communications data

The following overview explains the process of using a connector to archive Telegram communications data in Microsoft 365.



1. Your organization works with TeleMessage to set up a Telegram Archiver connector. For more information, see Activating the TeleMessage Telegram Archiver for Microsoft 365.

2. In real time, your organization's Telegram data is copied to the TeleMessage site.

3. The Telegram Archiver connector that you create in the Microsoft Purview portal connects to the TeleMessage site every day and transfers the email messages from the previous 24 hours to a secure Azure Storage area in the Microsoft Cloud.

4. The connector imports the mobile communication items to the mailbox of a specific user. A new folder named Telegram Archiver is created in the specific user's mailbox and the items are imported to it. The connector uses the value of the *User's Email address* property for this mapping. Every email message contains this property, which is populated with the email address of every participant of the email message.

> In addition to automatic user mapping by using the value of the *User's Email address* property, you can also define a custom mapping by uploading a CSV mapping file. This mapping file should contain User's mobile Number and the corresponding Microsoft 365 mailbox address for each user. If you enable automatic user mapping and provide a custom mapping, for every email item the connector will first look at custom mapping file. If it doesn't find a valid Microsoft 365 user that corresponds to a user's mobile number, the connector will use the User's email address property of the email item. If the connector doesn't find a valid Microsoft 365 user in either the custom mapping file or the *user's email address* property of the email item, the item won't be imported.

Before you begin

> **Important**
>
> All partner third-party data connector configurations require a provisioned **service principal** for the partner application. This requirement isn't applicable for native Microsoft third-party data connector configurations.
>
> To provision a service principal using PowerShell, see **Provision a service principal for partner third-party data connectors**.

- Order the Telegram archiving service from TeleMessage and get a valid administration account for your organization. You need to sign in to this account when you create the connector in the Microsoft Purview portal.

- Register all users that require Telegram archiving in the TeleMessage account. When registering users, use the same email address that's used for their Microsoft 365 account.
- Install the Telegram Archiver app on the mobile phones of your employees and activate it. The Telegram Archiver app allows them to communicate and chat with other Telegram users.
- Assign the Data Connector Admin role to the user who creates a Telegram Archiver connector in Create a Telegram Archiver connector. This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions, see Permissions in the Microsoft Purview portal and Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview.
- This data connector is available in GCC environments in the Microsoft 365 US Government cloud. Third-party applications and services might involve storing, transmitting, and processing your organization's customer data on third-party systems that are outside of the Microsoft 365 infrastructure and therefore aren't covered by the Microsoft Purview and data protection commitments. Microsoft makes no representation that use of this product to connect to third-party applications implies that those third-party applications are FedRAMP compliant.

Create a Telegram Archiver connector

After you complete the prerequisites described in Before you begin, create the Telegram Archiver connector in the Microsoft Purview portal. The connector uses the information you provide to connect to the TeleMessage site and transfers Telegram communications data to the corresponding user mailboxes in Microsoft 365.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. From the list, select **Telegram Archiver**.
5. On the **Terms of service** page, select **Accept**.

6. On the **Login to TeleMessage** page, under Step 3, enter the required information in the following boxes and then select **Next**.

    - **Username:** Your TeleMessage username.
    - **Password:** Your TeleMessage password.

7. After the connector is created, you can close the pop-up window and go to the next page.

8. On the **User mapping** page, enable automatic user mapping. To enable custom mapping, upload a CSV file that contains the user mapping information, and then select **Next**.

9. Review your settings, and then select **Finish** to create the connector.

Limits for archiving Telegram communications data

Importing attachments or items supports files up to 10 MB.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/19/2026

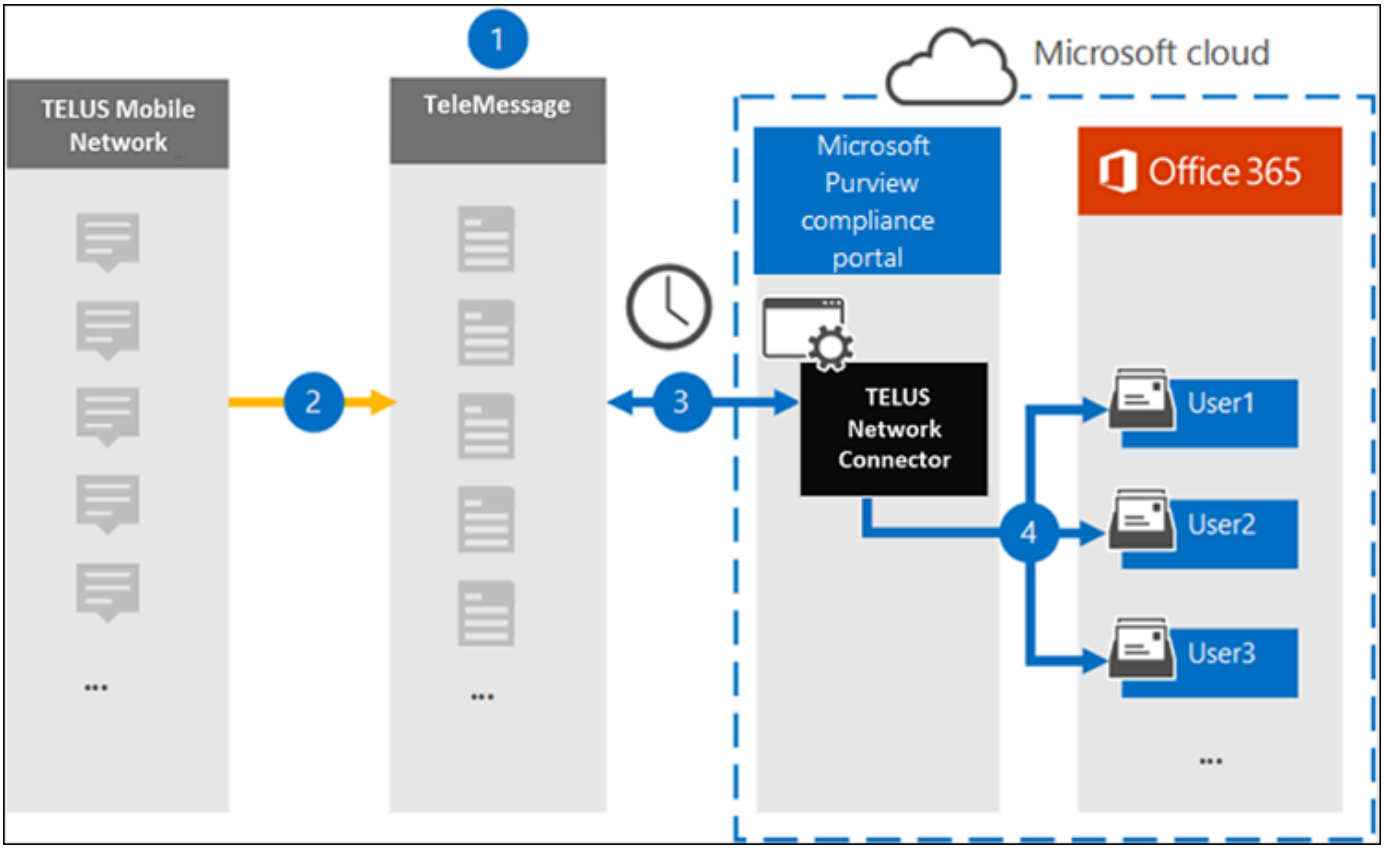
Set up a connector to archive TELUS Network data

Use the TeleMessage connector to import and archive Short Messaging Service (SMS) data from your organization's TELUS Network. After you set up and configure a connector, it connects to your organization's TELUS Network once every day and imports SMS data to mailboxes in Microsoft 365.

After storing SMS messages in user mailboxes, you can apply Microsoft Purview features such as Litigation Hold, eDiscovery, and Microsoft 365 retention policies to TELUS data. For example, you can search TELUS SMS messages by using eDiscovery or associate the mailbox that contains the TELUS data with a custodian in an eDiscovery (Premium) case. Using a TELUS Network connector to import and archive data in Microsoft 365 can help your organization stay compliant with government and regulatory policies.

Overview of archiving TELUS Network data

The following overview explains the process of using a connector to archive TELUS Network data in Microsoft 365.



1. Your organization works with TeleMessage and TELUS to set up a TELUS Network connector. For more information, see TELUS Network Archiver.

2. In real time, SMS messages from your organization's TELUS Network are copied to the TeleMessage site.

3. The TELUS Network connector that you create in the Microsoft Purview portal connects to the TeleMessage site every day and transfers the SMS messages from the previous 24 hours to a secure Azure Storage location in the Microsoft cloud. The connector also converts the content of SMS messages to an email message format.

4. The connector imports the mobile communication items to the mailbox of a specific user. A new folder named **TELUS SMS Network Archiver** is created in the specific user's mailbox and the items are imported to it. The connector does mapping by using the value of the *User's Email address* property. Every SMS message contains this property, which is populated with the email address of every participant of the SMS message.

   In addition to automatic user mapping by using the value of the *User's Email address* property, you can also implement custom mapping by uploading a CSV mapping file. This mapping file contains the mobile phone number and corresponding Microsoft 365 email address for users in your organization. If you enable both automatic user mapping and custom mapping, for every TELUS item the connector first looks at custom mapping file. If it doesn't find a valid Microsoft 365 user that corresponds to a user's mobile phone number, the connector uses the values in the email address property of the item it's trying to import. If the connector doesn't find a valid Microsoft 365 user in either the custom mapping file or in the email address property of the TELUS item, the item won't be imported.

Before you begin

> **Important**
>
> All partner third-party data connector configurations require a provisioned service principal for the partner application. This requirement isn't applicable for native Microsoft third-party data connector configurations.
>
> To provision a service principal using PowerShell, see Provision a service principal for partner third-party data connectors.

Some of the implementation steps required to archive TELUS Network data are external to Microsoft 365. You must complete these steps before you can create a connector in the Microsoft Purview portal.

- Order the TELUS Network Archiver service from TeleMessage and get a valid administration account for your organization. You need to sign in to this account when you create the connector in the Microsoft Purview portal.
- Obtain your TELUS Network account and billing contact details so you can fill out the TeleMessage onboarding forms and order the message archiving service from TELUS.
- Register all users that require TELUS SMS Network archiving in the TeleMessage account. When registering users, be sure to use the same email address that's used for their Microsoft 365 account.
- Ensure your employees have corporate-owned and corporate-liable mobile phones on the TELUS mobile network. Archiving messages in Microsoft 365 isn't available for employee-owned or Bring Your Own Devices (BYOD) devices.
- Assign the Data Connector Admin role to the user who creates a TELUS Network connector. This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions, see Permissions in the Microsoft Purview portal and Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview.
- This data connector is available in GCC environments in the Microsoft 365 US Government cloud. Third-party applications and services might involve storing, transmitting, and processing your organization's customer data on third-party systems that are outside of the Microsoft 365 infrastructure and therefore aren't covered by the Microsoft Purview and data protection commitments. Microsoft makes no representation that use of this product to connect to third-party applications implies that those third-party applications are FedRAMP compliant.

Create a TELUS Network connector

After you complete the prerequisites described in Before you begin, create a TELUS Network connector in the Microsoft Purview portal. The connector uses the information you provide to

connect to the TeleMessage site and transfer SMS messages to the corresponding user mailboxes in Microsoft 365.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. From the list, select **TELUS Network**.
5. On the **Terms of service** page, select **Accept**.
6. On the **Login to TeleMessage** page, under Step 3, enter the required information in the following boxes and then select **Next**.
    - **Username:** Your TeleMessage username.
    - **Password:** Your TeleMessage password.
7. After you create the connector, close the pop-up window.
8. On the **User mapping** page, enable automatic user mapping and select **Next**. To use custom mapping, upload a CSV file, and select **Next**.
9. Review your settings, and then select **Finish** to create the connector.

TELUS Network connector limits

Importing attachments or items supports files up to 10 MB.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/19/2026

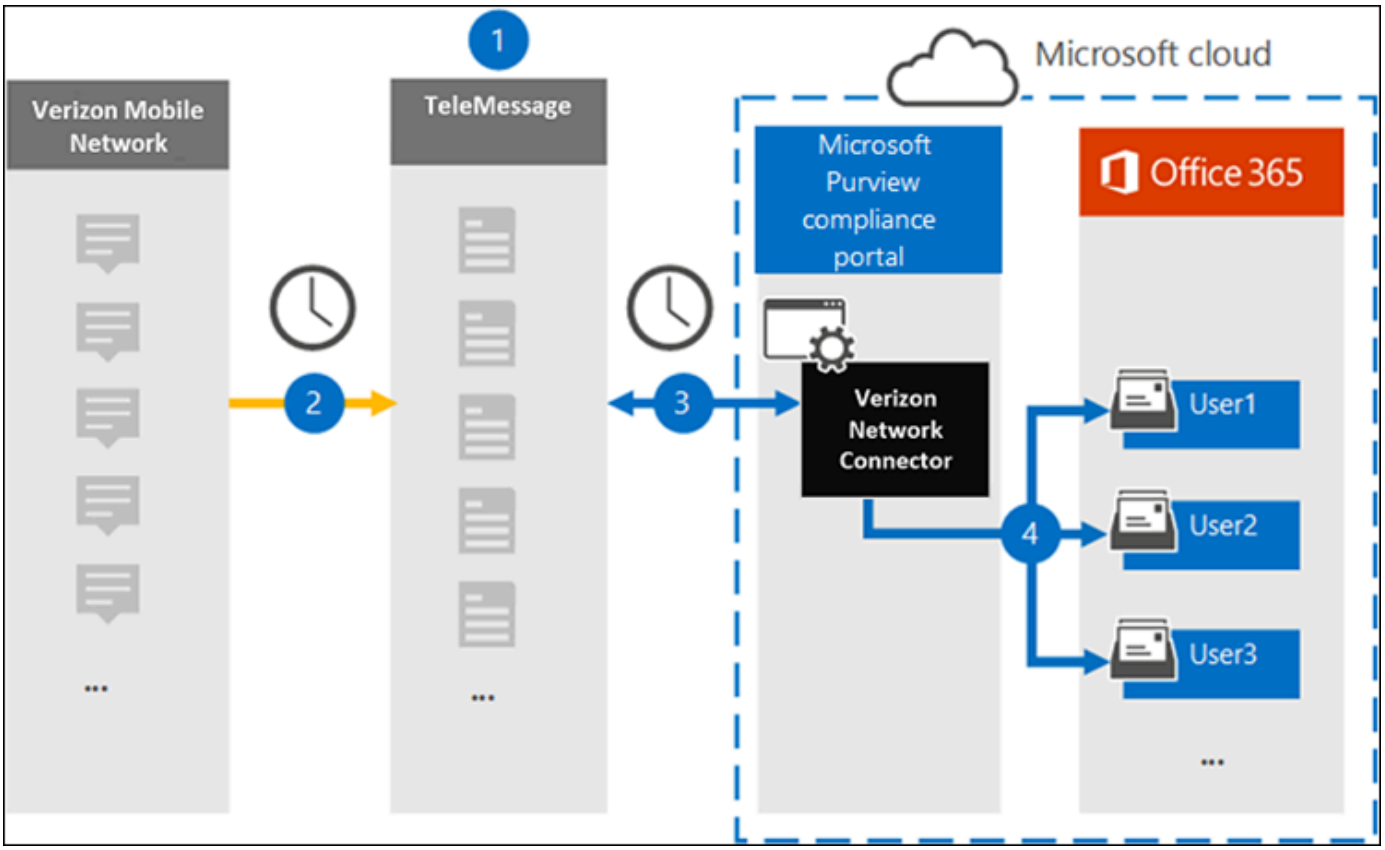
Set up a connector to archive Verizon Network data

Use the TeleMessage connector to import and archive Short Messaging Service (SMS) and Multimedia Messaging Service (MMS) data from Verizon Network. Before you begin, review the prerequisites for setting up the Verizon Network connector. After you set up and configure a connector, it connects to your organization's Verizon Network once every day and imports SMS and MMS data to mailboxes in Microsoft 365.

After storing Verizon Network connector data in user mailboxes, you can apply Microsoft Purview features such as Litigation Hold, eDiscovery, and Microsoft 365 retention policies to Verizon data. For example, you can search Verizon SMS and MMS messages by using eDiscovery or associate the mailbox that contains Verizon Network data with a custodian in a Microsoft Purview eDiscovery (Premium) case. Using a Verizon Network connector to import and archive data in Microsoft 365 can help your organization stay compliant with government and regulatory policies.

Overview of archiving Verizon Network data

The following overview explains the process of using a connector to archive Verizon Network data in Microsoft 365.



1. Your organization works with TeleMessage and Verizon to set up a Verizon Network connector. For more information, see Verizon Network Archiver.

2. Once every 24 hours, your organization's Verizon Network copies SMS and MMS messages to the TeleMessage site.

3. The Verizon Network connector that you create in the Microsoft Purview portal connects to the TeleMessage site every day and transfers the SMS and MMS messages from the previous 24 hours to a secure Azure Storage location in the Microsoft cloud. The connector also converts the content of SMS and MMS messages to an email message format.

4. The connector imports the mobile communication items to the mailbox of a specific user. A new folder named **Verizon SMS/MMS Network Archiver** is created in the specific user's mailbox and the items are imported to it. The connector does this mapping by using the value of the *User's Email address* property. Every SMS and MMS message contains this property, which is populated with the email address of every participant of the message.

   In addition to automatic user mapping by using the value of the *User's Email address* property, you can also implement custom mapping by uploading a CSV mapping file. This mapping file contains the mobile phone number and corresponding Microsoft 365 email address for users in your organization. If you enable both automatic user mapping and custom mapping, for every Verizon item the connector first looks at custom mapping file. If it doesn't find a valid Microsoft 365 user that corresponds to a user's mobile phone number,

the connector uses the values in the email address property of the item it's trying to import. If the connector doesn't find a valid Microsoft 365 user in either the custom mapping file or in the email address property of the Verizon item, the item won't be imported.

Before you begin

> **Important**
>
> All partner third-party data connector configurations require a provisioned service principal for the partner application. This requirement isn't applicable for native Microsoft third-party data connector configurations.
>
> To provision a service principal using PowerShell, see Provision a service principal for partner third-party data connectors.

Some of the implementation steps required to archive Verizon Network data are external to Microsoft 365. You must complete these steps before you can create a connector in the Microsoft Purview portal.

- Order the Verizon Network Archiver service from TeleMessage and get a valid administration account for your organization. You need to sign in to this account when you create the connector in the Microsoft Purview portal.
- Obtain your Verizon Network account and billing contact details so you can fill out the TeleMessage onboarding forms and order the message archiving service from Verizon.
- Register all users that require Verizon SMS and MMS archiving in the TeleMessage account. When registering users, use the same email address that's used for their Microsoft 365 account.
- Ensure your employees have corporate-owned and corporate-liable mobile phones on the Verizon mobile network. Archiving messages in Microsoft 365 isn't available for employee-owned or Bring Your Own Devices (BYOD) devices.
- Assign the Data Connector Admin role to the user who creates a Verizon Network connector. This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions,

see Permissions in the Microsoft Purview portal and Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview.

- This data connector is available in GCC environments in the Microsoft 365 US Government cloud. Third-party applications and services might involve storing, transmitting, and processing your organization's customer data on third-party systems that are outside of the Microsoft 365 infrastructure and therefore aren't covered by the Microsoft Purview and data protection commitments. Microsoft makes no representation that use of this product to connect to third-party applications implies that those third-party applications are FedRAMP compliant.

Create a Verizon Network connector

After you complete the prerequisites described in Before you begin, you can create a Verizon Network connector in the Microsoft Purview portal. The connector uses the information you provide to connect to the TeleMessage site and transfer SMS and MMS messages to the corresponding user mailboxes in Microsoft 365.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. From the list, select **Verizon Network**.
5. On the **Terms of service** page, select **Accept**.
6. On the **Login to TeleMessage** page, under Step 3, enter the required information in the following boxes and then select **Next**.
    - **Username:** Your TeleMessage username.
    - **Password:** Your TeleMessage password.
7. After the connector is created, you can close the pop-up window and go to the next page.
8. On the **User mapping** page, enable automatic user mapping and select **Next**. If you need custom mapping, upload a CSV file, and select **Next**.
9. Review your settings, and then select **Finish** to create the connector.

Limits for archiving Verizon Network data

Importing attachments or items supports files up to 10 MB.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/19/2026

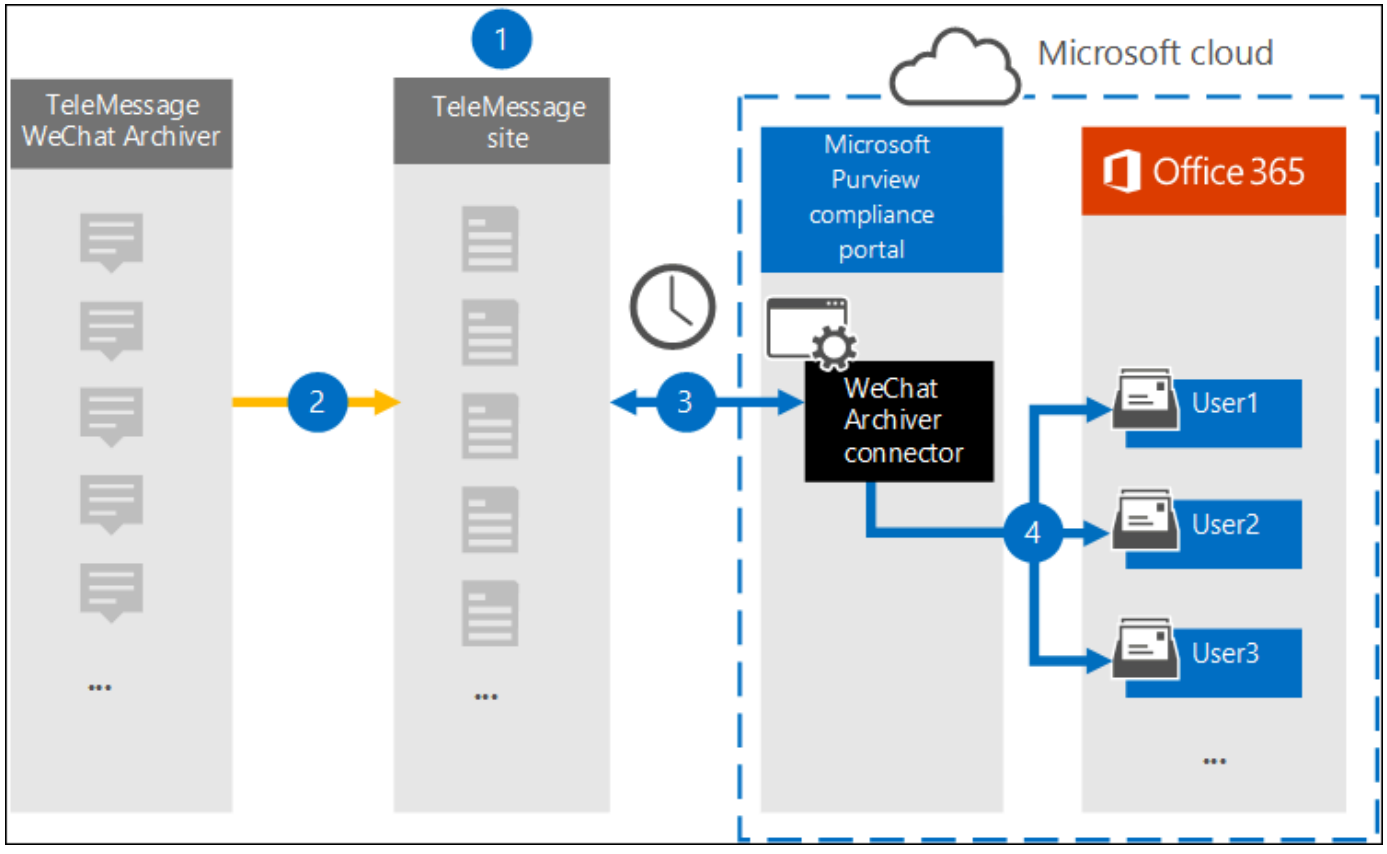
Set up a connector to archive WeChat data

Use the TeleMessage connector to import and archive WeChat and WeCom calls, chats, attachments, files, and recalled messages. When you set up and configure a connector, it connects to your organization's TeleMessage account and imports the mobile communication of employees who use the TeleMessage WeChat Archiver to mailboxes in Microsoft 365.

After storing WeChat Archiver connector data in user mailboxes, you can apply Microsoft Purview features such as Litigation Hold, eDiscovery, In-Place Archiving, Auditing, Communication Compliance, and Microsoft 365 retention policies to WeChat communication data. For example, you can search WeChat communication by using eDiscovery or associate the mailbox that contains the WeChat Archiver connector data with a custodian in an eDiscovery (Premium) case. Using a WeChat Archiver connector to import and archive data in Microsoft 365 can help your organization stay compliant with corporate governance regulations and regulatory policies.

Overview of archiving WeChat communication data

The following overview explains the process of using a connector to archive WeChat communications data in Microsoft 365.



1. Your organization works with TeleMessage to set up a WeChat Archiver connector.

2. In real time, your organization's WeChat data is copied to the TeleMessage site.

3. The WeChat Archiver connector that you create in the Microsoft Purview portal connects to the TeleMessage site every day and transfers the email messages from the previous 24 hours to a secure Azure Storage area in the Microsoft Cloud.

4. The connector imports the mobile communication items to the mailbox of a specific user. A new folder named WeChat Archiver is created in the specific user's mailbox and the items are imported to it. The connector does mapping by using the value of the *User's Email address* property. Every email message contains this property, which is populated with the email address of every participant of the email message. In addition to automatic user mapping using the value of the *User's Email address* property, you can also define a custom mapping by uploading a CSV mapping file. This mapping file should contain User's mobile Number and the corresponding Microsoft 365 mailbox address for each user. If you enable automatic user mapping and provide a custom mapping, for every email item the connector will first look at custom mapping file. If it doesn't find a valid Microsoft 365 user that corresponds to a user's mobile number, the connector will use the User ‘s email address property of the email item. If the connector doesn't find a valid Microsoft 365 user in either the custom mapping file or the *user's email address* property of the email item, the item won't be imported.

Before you begin

> **Important**
>
> All partner third-party data connector configurations require a provisioned service principal for the partner application. This requirement isn't applicable for native Microsoft third-party data connector configurations.
>
> To provision a service principal using PowerShell, see Provision a service principal for partner third-party data connectors.

- Work with TeleMessage to set up a WeChat archive connector. For activation steps, see Activating the TeleMessage WeChat Archiver for Microsoft 365.

- Set up a TeleMessage connector for Microsoft 365 and get a valid company administration account. To purchase the service, see Order Microsoft 365 Mobile Archiving.

- Register all users that require WeChat archiving in the TeleMessage account with the same email address that is used for the user's Microsoft 365 account.
- Install the Tencent WeCom app on the mobile phones of users in your organization and activate it. The WeCom app lets users communicate and chat with other WeChat and WeCom users.
- Assign the Data Connector Admin role to the user who creates a WeChat Archiver connector in the Microsoft Purview portal. This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions, see Permissions in the Microsoft Purview portal and Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview.
- This data connector is available in GCC environments in the Microsoft 365 US Government cloud. Third-party applications and services might involve storing, transmitting, and processing your organization's customer data on third-party systems that are outside of the Microsoft 365 infrastructure and therefore aren't covered by the Microsoft Purview and data protection commitments. Microsoft makes no representation that use of this product to connect to third-party applications implies that those third-party applications are FedRAMP compliant.

Create a WeChat Archiver connector

Complete the following steps to create a WeChat Archiver connector in the Microsoft Purview portal. The connector uses the information you provide to connect to the TeleMessage site and transfer WeChat communications data to the corresponding user mailboxes in Microsoft 365.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. From the list, select **WeChat Archiver**.
5. On the **Terms of service** page, select **Accept**.

6. On the **Login to TeleMessage** page, under Step 3, enter the required information in the following boxes and then select **Next**.

    - **Username**: Your TeleMessage user name.
    - **Password**: Your TeleMessage password.

7. After you create the connector, you can close the TeleMessage sign-in pop-up window.

8. On the **User mapping** page, enable automatic user mapping. You can also upload a custom user mapping CSV file.

9. Select **Next**, review your settings, and then select **Finish** to create the connector.

WeChat connector limits

Importing attachments or items supports files up to 10 MB.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/19/2026

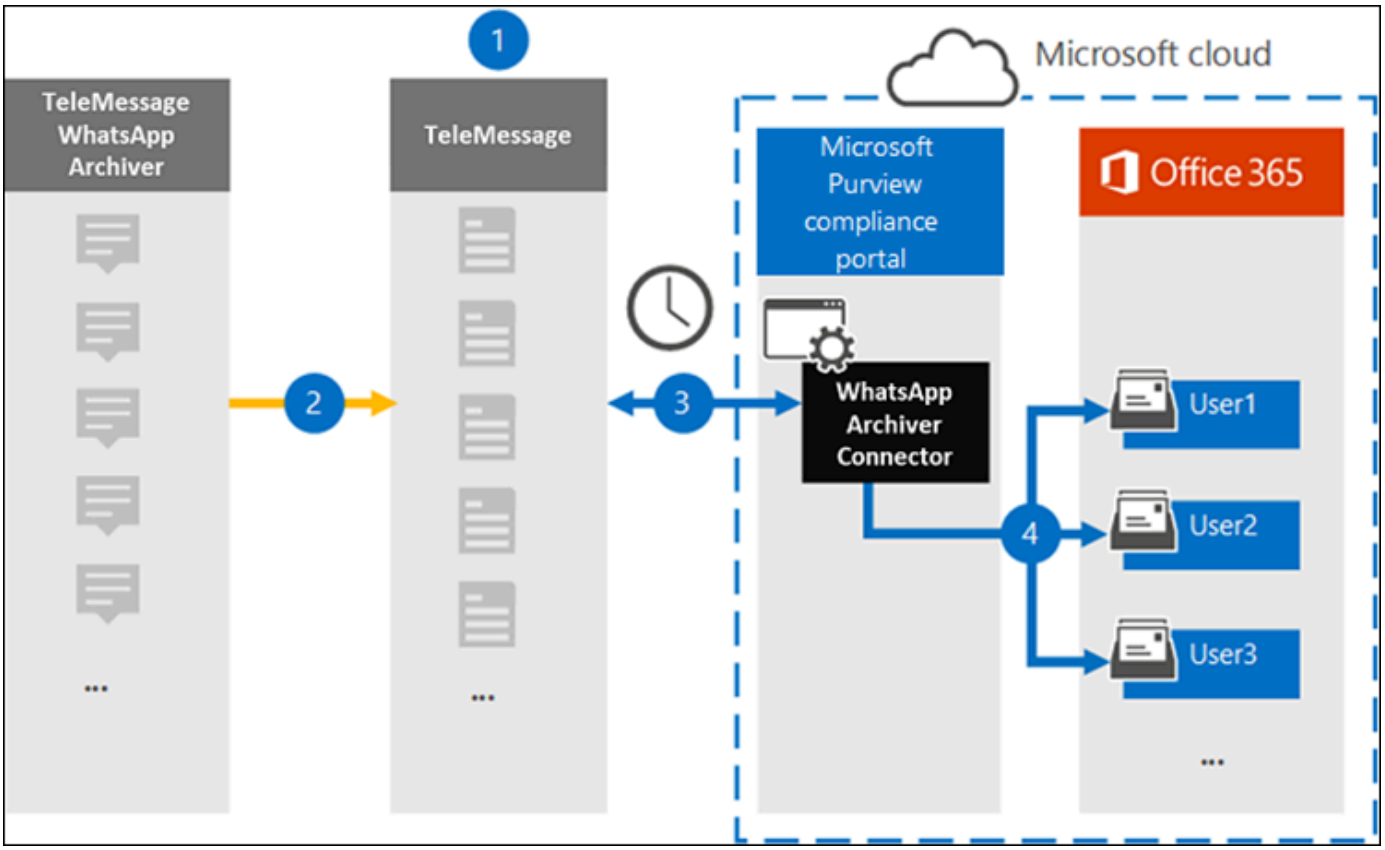
Set up a connector to archive WhatsApp data

Use the TeleMessage connector to import and archive WhatsApp calls, chats, attachments, files, and deleted messages. After you set up and configure a connector, it connects to your organization's TeleMessage account once every day, and imports the mobile communication of employees using the TeleMessage WhatsApp Phone Archiver or TeleMessage WhatsApp Cloud Archiver to mailboxes in Microsoft 365.

After WhatsApp data is stored in user mailboxes, you can apply Microsoft Purview features such as Litigation Hold, eDiscovery, and Microsoft 365 retention policies to WhatsApp data. For example, you can search WhatsApp messages using eDiscovery or associate the mailbox that contains WhatsApp messages with a custodian in an eDiscovery (Premium) case. Using a WhatsApp connector to import and archive data in Microsoft 365 can help your organization stay compliant with government and regulatory policies.

Overview of archiving WhatsApp data

The following overview explains the process of using a connector to archive WhatsApp data in Microsoft 365. A connector is a configuration in Microsoft Purview that connects to TeleMessage and imports WhatsApp data into Microsoft 365 mailboxes.



1. Your organization works with TeleMessage to set up a WhatsApp Archiver connector. For more information, see WhatsApp Archiver.

2. In real time, your organization's WhatsApp data is copied to the TeleMessage site.

3. The WhatsApp connector that you create in the Microsoft Purview portal connects to the TeleMessage site every day and transfers WhatsApp data from the previous 24 hours to a secure Azure Storage location in the Microsoft cloud. The connector also converts the content WhatsApp data to an email message format.

4. The connector imports WhatsApp data to the mailbox of a specific user. A new folder named **WhatsApp Archiver** is created in the specific user's mailbox and the items are imported to it. The connector does this mapping by using the value of the *User's Email address* property. Every WhatsApp message contains this property, which is populated with the email address of every participant of the message.

   In addition to automatic user mapping using the value of the *User's Email address* property, you can also implement custom mapping by uploading a CSV mapping file. This mapping file contains the mobile phone number and corresponding Microsoft 365 email address for users in your organization. If you enable both automatic user mapping and custom mapping, for every WhatsApp item the connector first looks at custom mapping file. If it doesn't find a valid Microsoft 365 user that corresponds to a user's mobile phone number, the connector will use the values in the email address property of the item it's trying to

import. If the connector doesn't find a valid Microsoft 365 user in either the custom mapping file or in the email address property of the WhatsApp item, the item won't be imported.

Before you begin

> **Important**
>
> All partner third-party data connector configurations require a provisioned service principal for the partner application. This requirement isn't applicable for native Microsoft third-party data connector configurations.
>
> To provision a service principal using PowerShell, see Provision a service principal for partner third-party data connectors.

You need to complete some implementation steps before you can create the connector in the Microsoft Purview portal. These steps are external to Microsoft 365.

- Order the WhatsApp Archiver service from TeleMessage and get a valid administration account for your organization. You need to sign in to this account when you create the connector in the Microsoft Purview portal.
- Register all users that require WhatsApp archiving in the TeleMessage account. When registering users, use the same email address that's used for their Microsoft 365 account.
- Install the TeleMessage WhatsApp Phone Archiver app on the mobile phones of your employees and activate it. Alternatively, you can install the regular WhatsApp or WhatsApp Business apps on the mobile phones of your employees and activate the WhatsApp Cloud Archiver service by scanning a QR code on the TeleMessage website. For more information, see WhatsApp Cloud Archiver.
- Assign the Data Connector Admin role to the user who creates a WhatsApp Archiver connector. This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions, see Permissions in the Microsoft Purview portal and Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview.

- This data connector is available in GCC environments in the Microsoft 365 US Government cloud. Third-party applications and services might involve storing, transmitting, and processing your organization's customer data on third-party systems that are outside of the Microsoft 365 infrastructure and therefore aren't covered by the Microsoft Purview and data protection commitments. Microsoft makes no representation that use of this product to connect to third-party applications implies that those third-party applications are FedRAMP compliant.

Create a WhatsApp Archiver connector

After you complete the prerequisites described in Before you begin, create the WhatsApp connector in the Microsoft Purview portal. The connector uses the information you provide to connect to the TeleMessage site and transfer the WhatsApp data to the corresponding user mailboxes in Microsoft 365.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. From the list, select **WhatsApp Archiver**.
5. On the **Terms of service** page, select **Accept**.
6. On the **Login to TeleMessage** page, under Step 3, enter the required information in the following boxes and then select **Next**.
    - **Username:** Your TeleMessage username.
    - **Password:** Your TeleMessage password.
7. After the connector is created, you can close the pop-up window and go to the next page.
8. On the **User mapping** page, enable automatic user mapping and select **Next**. If you need custom mapping, upload a CSV file, and select **Next**.
9. Review your settings, and then select **Finish** to create the connector.

Limits for archiving WhatsApp data

Importing attachments or items supports files up to 10 MB.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/19/2026

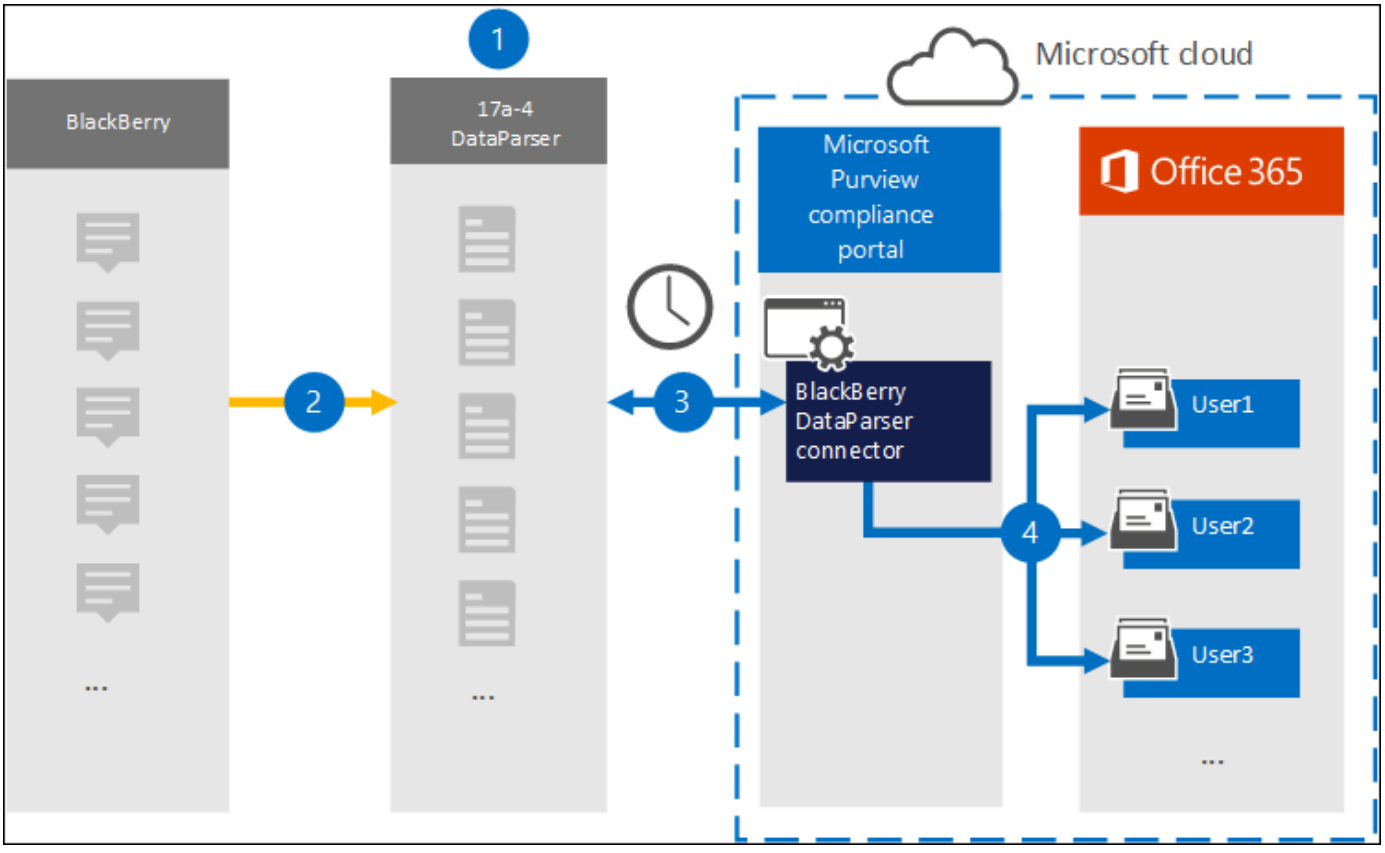
Set up a connector to archive BlackBerry data

Use the BlackBerry DataParser from 17a-4 LLC to import and archive BlackBerry enterprise data to user mailboxes in your Microsoft 365 organization. The DataParser includes a BlackBerry connector that you configure to capture items from a third-party data source and import those items to Microsoft 365. The BlackBerry DataParser connector converts BlackBerry data to an email message format and then imports those items to user mailboxes in Microsoft 365.

After you store BlackBerry data in user mailboxes, you can apply Microsoft Purview features such as Litigation Hold, eDiscovery, retention policies and retention labels, and Communication Compliance. By using a BlackBerry connector to import and archive data in Microsoft 365, your organization can stay compliant with government and regulatory policies.

Overview of archiving BlackBerry data

The following overview explains the process of using a data connector to archive BlackBerry data in Microsoft 365.



1. Your organization works with 17a-4 to set up and configure the BlackBerry DataParser.

2. The DataParser regularly collects BlackBerry items and converts the content of a message to an email message format.

3. You create the BlackBerry DataParser connector in the Microsoft Purview portal. The connector connects to DataParser and transfers the messages to a secure Azure Storage location in the Microsoft cloud.

4. The connector creates a subfolder in the Inbox folder named **BlackBerry DataParser** in the user mailboxes and imports the BlackBerry items to that folder. The connector uses the value of the *Email* property to determine which mailbox to import items to. Every BlackBerry item contains the *Email* property, which is populated with the email address of every participant.

Before you begin

> **Important**
>
> All partner third-party data connector configurations require a provisioned service principal for the partner application. This requirement isn't applicable for native Microsoft third-party data connector configurations.
>
> To provision a service principal using PowerShell, see Provision a service principal for partner third-party data connectors.

- Create a DataParser account for Microsoft connectors. To create an account, contact 17a-4 LLC. You need to sign in to this account when you create the connector.

- Assign the Data Connector Admin role to the user who creates the BlackBerry DataParser connector and completes the connector setup. This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions, see Permissions in the Microsoft Purview portal and Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview.

- This data connector is available in GCC environments in the Microsoft 365 US Government cloud. Third-party applications and services might involve storing, transmitting, and processing your organization's customer data on third-party systems that are outside of the

Microsoft 365 infrastructure and therefore aren't covered by the Microsoft Purview and data protection commitments. Microsoft makes no representation that use of this product to connect to third-party applications implies that those third-party applications are FedRAMP compliant.

Step 1: Set up a BlackBerry DataParser connector

Create a 17a-4 connector for BlackBerry data in the Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. Select **BlackBerry DataParser** from the list.
5. Select **Accept** on the **Terms of service** page.
6. Enter a unique name that identifies the connector in the **Name** field and select **Next**.
7. Sign in to your 17a-4 account and complete the steps in the BlackBerry DataParser connection workflow.

Step 2: Configure the BlackBerry DataParser connector

Work with 17a-4 Support to configure the BlackBerry DataParser connector.

Step 3: Map users

The BlackBerry DataParser connector automatically maps users to their Microsoft 365 email addresses before importing data to Microsoft 365.

Step 4: Monitor the BlackBerry DataParser connector

After you create a BlackBerry DataParser connector, you can view the connector status in the Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select the BlackBerry DataParser connector that you created to display the flyout page, which contains the properties and information about the connector.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

Last updated on 06/20/2026

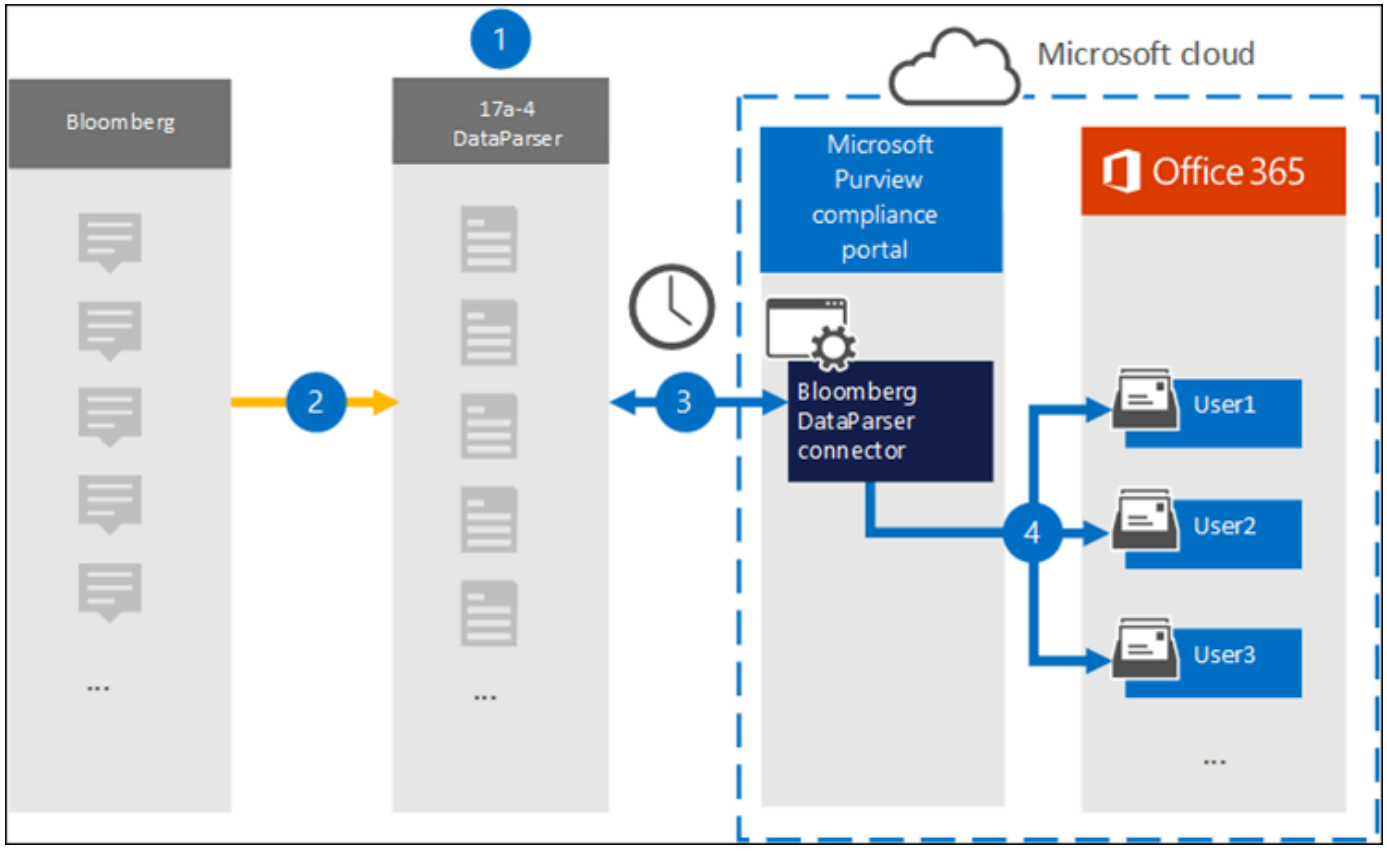
Set up a connector to archive Bloomberg data

Use the Bloomberg DataParser from 17a-4 LLC to import and archive data from Bloomberg to user mailboxes in your Microsoft 365 organization. The DataParser includes a Bloomberg connector that's configured to capture items from a third-party data source and import those items to Microsoft 365. The Bloomberg DataParser connector converts Bloomberg data to an email message format and then imports those items to user mailboxes in Microsoft 365.

After you store Bloomberg data in user mailboxes, you can apply Microsoft Purview features such as Litigation Hold, eDiscovery, retention policies and retention labels, and Communication Compliance. By using a Bloomberg connector to import and archive data in Microsoft 365, you can help your organization stay compliant with government and regulatory policies.

Overview of archiving Bloomberg data

The following overview explains the process of using a data connector to archive Bloomberg data in Microsoft 365.



1. Your organization works with 17a-4 to set up and configure the Bloomberg DataParser.

2. The DataParser regularly collects Bloomberg items and converts the content of a message to an email message format.

3. You create the Bloomberg DataParser connector in the Microsoft Purview portal. The connector connects to DataParser and transfers the converted Bloomberg messages to a secure Azure Storage location in the Microsoft cloud.

4. The connector creates a subfolder in the Inbox folder named **Bloomberg DataParser** in the user mailboxes and imports the Bloomberg items to that folder. The connector uses the value of the *Email* property to determine which mailbox to import items to. Every Bloomberg item contains the *Email* property, which is populated with the email address of every participant.

Before you begin

> **Important**
>
> All partner third-party data connector configurations require a provisioned service principal for the partner application. This requirement isn't applicable for native Microsoft third-party data connector configurations.
>
> To provision a service principal using PowerShell, see Provision a service principal for partner third-party data connectors.

- Create a DataParser account for Microsoft connectors. To create an account, contact 17a-4 LLC. You need to sign in to this account when you create the connector.

- Assign the Data Connector Admin role to the user who creates the Bloomberg DataParser connector. This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions, see Permissions in the Microsoft Purview portal and Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview.

- This data connector is available in GCC environments in the Microsoft 365 US Government cloud. Third-party applications and services might involve storing, transmitting, and processing your organization's customer data on third-party systems that are outside of the

Microsoft 365 infrastructure and therefore aren't covered by the Microsoft Purview and data protection commitments. Microsoft makes no representation that use of this product to connect to third-party applications implies that those third-party applications are FedRAMP compliant.

Step 1: Set up a Bloomberg DataParser connector

Create a 17a-4 connector for Bloomberg data in the Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. Select **Bloomberg DataParser** from the list.
5. Select **Accept** on the **Terms of service** page.
6. Enter a unique name that identifies the connector and select **Next**.
7. Sign in to your 17a-4 account and complete the steps in the Bloomberg DataParser connection workflow.

Step 2: Configure the Bloomberg DataParser connector

Work with 17a-4 Support to configure the Bloomberg DataParser connector.

Step 3: Map Bloomberg users to Microsoft 365 users

The Bloomberg DataParser connector automatically maps users to their Microsoft 365 email addresses before importing data to Microsoft 365.

Step 4: Monitor the Bloomberg DataParser connector

After you create a Bloomberg DataParser connector, you can view the connector status in the Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select the Bloomberg DataParser connector that you created to display the flyout page, which contains the properties and information about the connector.
4. Under **Connector status with source**, select the **Download log** link to open (or save) the status log for the connector. This log contains information about the data that's imported to the Microsoft cloud. For more information, see View admin logs for data connectors.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/24/2026

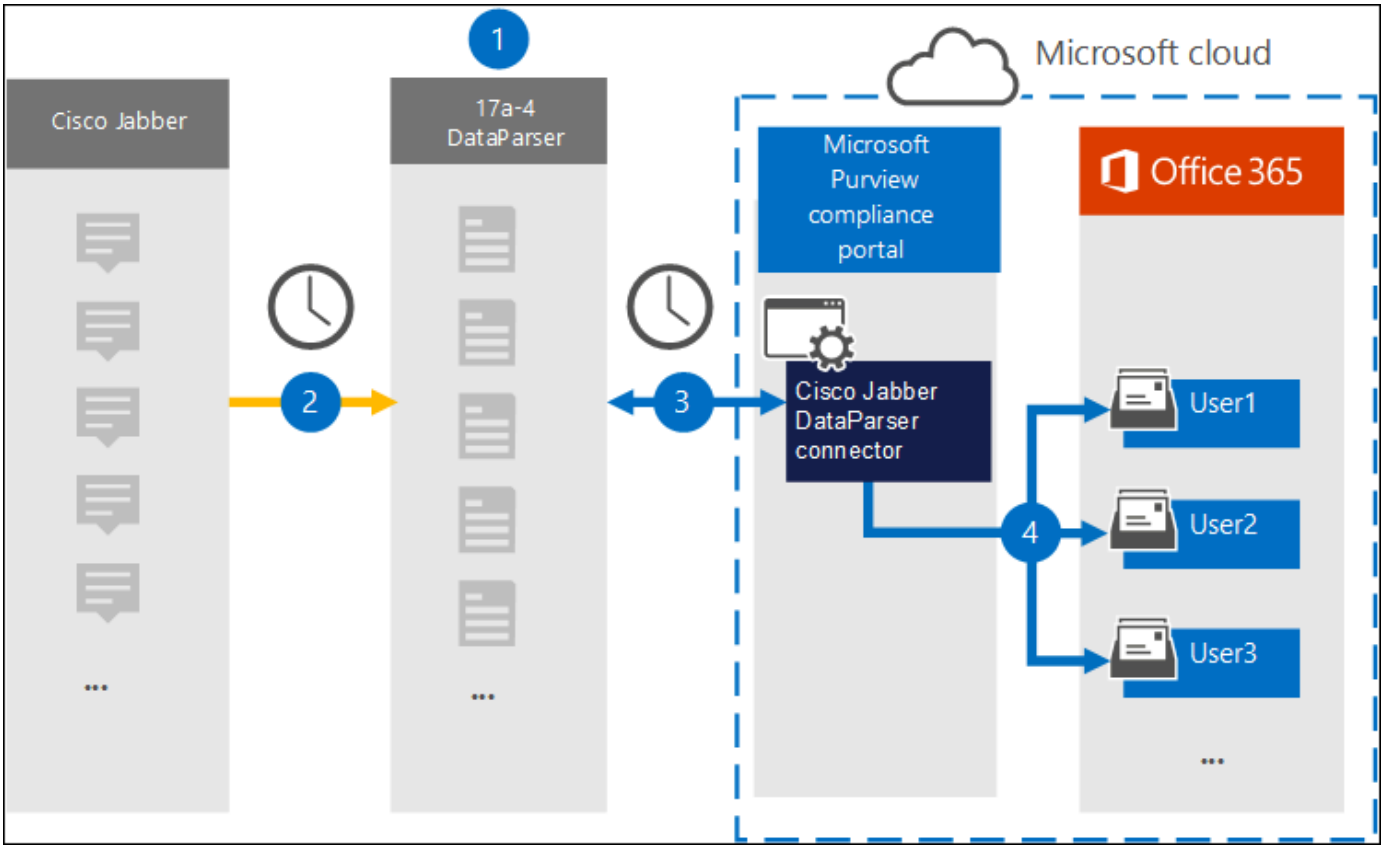
Set up a connector to archive Cisco Jabber data

Use the Cisco Jabber DataParser from 17a-4 LLC to import and archive data from Cisco Jabber to user mailboxes in your Microsoft 365 organization. The DataParser includes a Cisco Jabber connector that you configure to capture items from a third-party data source and import those items to Microsoft 365. The Cisco Jabber DataParser connector converts Cisco Jabber data to an email message format and then imports those items to user mailboxes in Microsoft 365.

After you store Cisco Jabber data in user mailboxes, you can apply Microsoft Purview features such as Litigation Hold, eDiscovery, retention policies and retention labels, and Communication Compliance. By using a Cisco Jabber connector to import and archive data in Microsoft 365, your organization can stay compliant with government and regulatory policies.

Overview of archiving Cisco Jabber data

The following overview explains the process of using a data connector to archive Cisco Jabber data in Microsoft 365.



1. Your organization works with 17a-4 to set up and configure the Cisco Jabber DataParser.

2. The DataParser regularly collects Cisco Jabber items and converts the content of a message to an email message format.

3. You create the Cisco Jabber DataParser connector in the Microsoft Purview portal. The connector connects to DataParser and transfers the messages to a secure Azure Storage location in the Microsoft cloud.

4. The connector creates a subfolder in the Inbox folder named **Cisco Jabber DataParser** in the user mailboxes and imports the Cisco Jabber items to that folder. The connector uses the value of the *Email* property to determine which mailbox to import items to. Every Cisco Jabber item contains this property, which is populated with the email address of every participant.

Before you begin

> **Important**
>
> All partner third-party data connector configurations require a provisioned service principal for the partner application. This requirement isn't applicable for native Microsoft third-party data connector configurations.
>
> To provision a service principal using PowerShell, see Provision a service principal for partner third-party data connectors.

- Create a DataParser account for Microsoft connectors. To create an account, contact 17a-4 LLC. You need to sign in to this account when you set up the Cisco Jabber DataParser connector in the Microsoft Purview portal.

- Assign the Data Connector Admin role to the user who sets up the Cisco Jabber DataParser connector in the Microsoft Purview portal and completes the user mapping configuration. This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions, see Permissions in the Microsoft Purview portal and Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview.

- This data connector is available in GCC environments in the Microsoft 365 US Government cloud. Third-party applications and services might involve storing, transmitting, and processing your organization's customer data on third-party systems that are outside of the Microsoft 365 infrastructure and therefore aren't covered by the Microsoft Purview and data protection commitments. Microsoft makes no representation that use of this product to connect to third-party applications implies that those third-party applications are FedRAMP compliant.

Step 1: Set up a Cisco Jabber DataParser connector

Create a 17a-4 connector for Cisco Jabber data in the Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. From the list, select **Cisco Jabber DataParser**.
5. On the **Terms of service** page, select **Accept**.
6. Enter a unique name that identifies the connector and then select **Next**.
7. Sign in to your 17a-4 account and complete the steps in the Cisco Jabber DataParser connection workflow.

Step 2: Configure the Cisco Jabber DataParser connector

Work with 17a-4 Support to configure the Cisco Jabber DataParser connector.

Step 3: Map Cisco Jabber users to Microsoft 365 users

The Cisco Jabber DataParser connector automatically maps users to their Microsoft 365 email addresses before importing data to Microsoft 365.

Step 4: Monitor the Cisco Jabber DataParser connector

After you create a Cisco Jabber DataParser connector, you can view the connector status in the Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.

3. Select **My connectors**, then select the Cisco Jabber DataParser connector that you created to display the flyout page, which contains the properties and information about the connector.
4. Under **Connector status with source**, select the **Download log** link to open (or save) the status log for the connector. This log contains information about the data that's been imported to the Microsoft cloud. For more information, see View admin logs for data connectors.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/24/2026

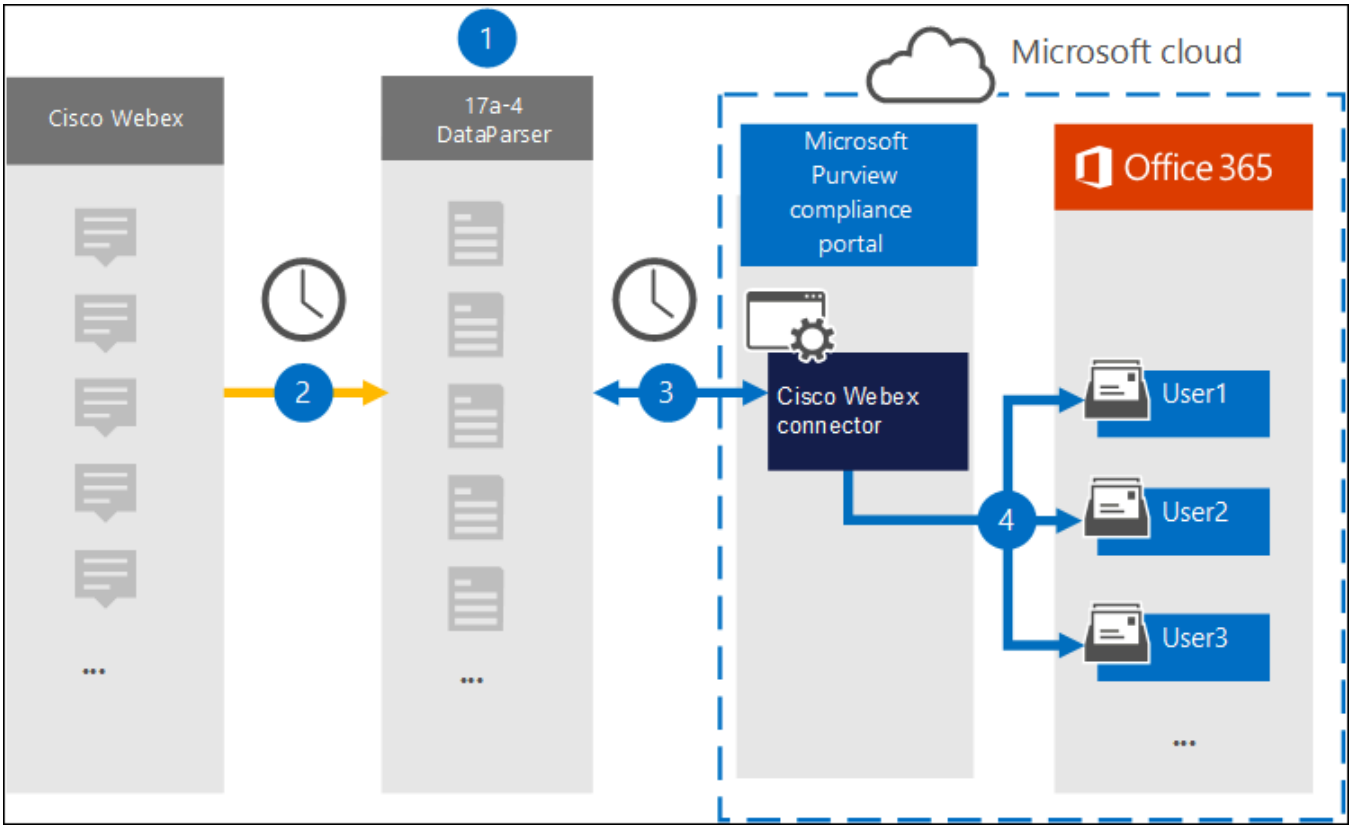
Set up a connector to archive Cisco Webex data

Use the Cisco Webex DataParser from 17a-4 LLC to import and archive data from the Cisco Webex platform to user mailboxes in your Microsoft 365 organization. The DataParser includes a Cisco Webex connector that you configure to capture items from a third-party data source and import those items to Microsoft 365. The Cisco Webex DataParser connector converts Cisco Webex data to an email message format and then imports those items to user mailboxes in Microsoft 365.

After you store Cisco Webex data in user mailboxes, you can apply Microsoft Purview features such as Litigation Hold, eDiscovery, retention policies and retention labels, and Communication Compliance. By using a Cisco Webex connector to import and archive data in Microsoft 365, you can help your organization stay compliant with government and regulatory policies.

Overview of archiving Cisco Webex data

The following overview explains the process of using a data connector to archive Cisco Webex data in Microsoft 365.



1. Your organization works with 17a-4 to set up and configure the Cisco Webex DataParser.
2. The DataParser regularly collects Cisco Webex items and converts the content of a message to an email message format.

3. You create the Cisco Webex DataParser connector in the Microsoft Purview portal. The connector connects to DataParser and transfers the messages to a secure Azure Storage location in the Microsoft cloud.

4. The connector creates a subfolder in the Inbox folder named **Cisco Webex DataParser** in the user mailboxes and imports the Cisco Webex items to that folder. The connector uses the value of the *Email* property to determine which mailbox to import items to. Every Cisco Webex item contains this property, which is populated with the email address of every participant.

Before you begin

> **Important**
>
> All partner third-party data connector configurations require a provisioned service principal for the partner application. This requirement isn't applicable for native Microsoft third-party data connector configurations.
>
> To provision a service principal using PowerShell, see Provision a service principal for partner third-party data connectors.

- Create a DataParser account for Microsoft connectors. To create an account, contact 17a-4 LLC. You need to sign in to this account when you create the connector in Step 1.

- Assign the Data Connector Admin role to the user who creates the Cisco Webex DataParser connector in Step 1 (and completes it in Step 3). This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions, see:- Permissions in the Microsoft Purview portal - Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview compliance

- This data connector is available in GCC environments in the Microsoft 365 US Government cloud. Third-party applications and services might involve storing, transmitting, and processing your organization's customer data on third-party systems that are outside of the Microsoft 365 infrastructure and therefore aren't covered by the Microsoft Purview and data protection commitments. Microsoft makes no representation that use of this product to connect to third-party applications implies that those third-party applications are FedRAMP compliant.

Step 1: Set up a Cisco Webex DataParser connector

Create a 17a-4 connector for Cisco Webex data in the Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. From the list, choose **Cisco Webex DataParser**.
5. On the **Terms of service** page, select **Accept**.
6. Enter a unique name that identifies the connector and then select **Next**.
7. Sign in to your 17a-4 account and complete the steps in the Cisco Webex DataParser connection workflow.

Step 2: Configure the Cisco Webex DataParser connector

Work with 17a-4 Support to configure the Cisco Webex DataParser connector.

Step 3: Map users

The Cisco Webex DataParser connector automatically maps users to their Microsoft 365 email addresses before importing data to Microsoft 365.

Step 4: Monitor the Cisco Webex DataParser connector

After you create a Cisco Webex DataParser connector, you can view the connector status in the Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select the Cisco Webex DataParser connector that you created to display the flyout page, which contains the properties and information about the connector.
4. Under **Connector status with source**, select the **Download log** link to open (or save) the status log for the connector. This log contains information about the data that's imported to the Microsoft cloud. For more information, see View admin logs for data connectors.

Limits

Importing attachments or items supports files up to 10 MB.

Last updated on 03/13/2026

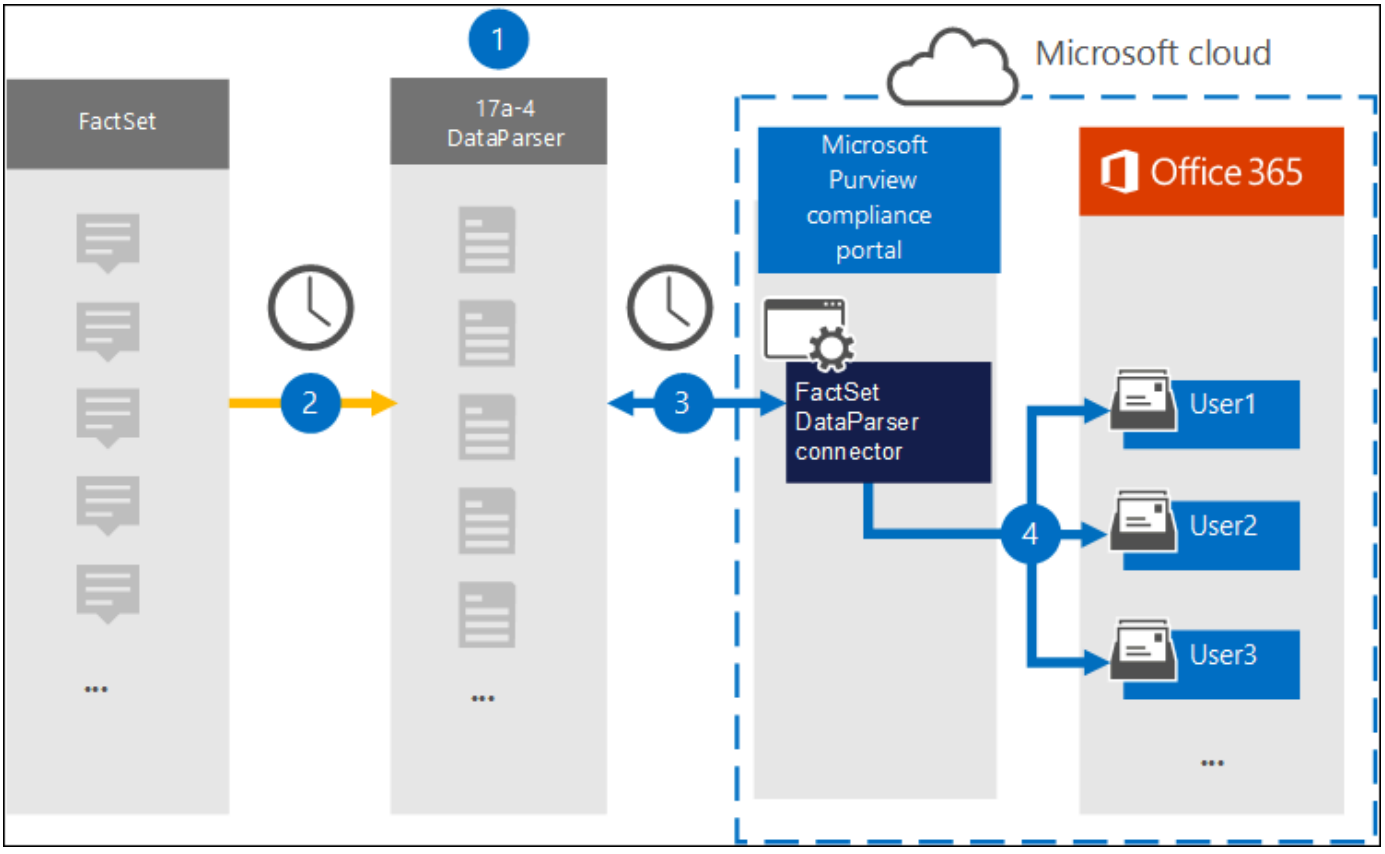
Set up a connector to archive FactSet data

Use the FactSet DataParser from 17a-4 LLC to import and archive data from the FactSet platform to user mailboxes in your Microsoft 365 organization. The DataParser includes a FactSet connector that you configure to capture items from a third-party data source and import those items to Microsoft 365. The FactSet DataParser connector converts FactSet data to an email message format and then imports those items to user mailboxes in Microsoft 365.

After storing FactSet data in user mailboxes, you can apply Microsoft Purview features such as Litigation Hold, eDiscovery, retention policies and retention labels, and Communication Compliance. By using a FactSet connector to import and archive data in Microsoft 365, your organization can stay compliant with government and regulatory policies.

Overview of archiving FactSet data

The following overview explains the process of using a data connector to archive FactSet data in Microsoft 365.



1. Your organization works with 17a-4 to set up and configure the FactSet DataParser.

2. The DataParser regularly collects FactSet items and converts the content of a message to an email message format.

3. You create the FactSet DataParser connector in the Microsoft Purview portal. The connector connects to DataParser and transfers the messages to a secure Azure Storage location in the Microsoft cloud.

4. The connector creates a subfolder in the Inbox folder named **FactSet DataParser** in the user mailboxes and imports the FactSet items to that folder. The connector uses the value of the *Email* property to determine which mailbox to import items to. Every FactSet item contains this property, which is populated with the email address of every participant.

Before you begin

> **Important**
>
> All partner third-party data connector configurations require a provisioned service principal for the partner application. This requirement isn't applicable for native Microsoft third-party data connector configurations.
>
> To provision a service principal using PowerShell, see Provision a service principal for partner third-party data connectors.

- Create a DataParser account for Microsoft connectors. To do this, contact 17a-4 LLC. You need to sign in to this account when you create the connector in the Microsoft Purview portal.

- Assign the Data Connector Admin role to the user who creates the FactSet DataParser connector in the Microsoft Purview portal and completes the user-mapping configuration. This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions, see Permissions in the Microsoft Purview portal and Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview.

- This data connector is available in GCC environments in the Microsoft 365 US Government cloud. Third-party applications and services might involve storing, transmitting, and

processing your organization's customer data on third-party systems that are outside of the Microsoft 365 infrastructure and therefore aren't covered by the Microsoft Purview and data protection commitments. Microsoft makes no representation that use of this product to connect to third-party applications implies that those third-party applications are FedRAMP compliant.

Step 1: Set up a FactSet DataParser connector

Create a 17a-4 connector for FactSet data in the Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. From the list, select **FactSet DataParser**.
5. On the **Terms of service** page, select **Accept**.
6. Enter a unique name that identifies the connector and then select **Next**.
7. Sign in to your 17a-4 account and complete the steps in the FactSet DataParser connection workflow.

Step 2: Configure the FactSet DataParser connector

Work with 17a-4 Support to configure the FactSet DataParser connector.

Step 3: Map FactSet users to Microsoft 365 users

The FactSet DataParser connector automatically maps users to their Microsoft 365 email addresses before importing data to Microsoft 365.

Step 4: Monitor the FactSet DataParser connector

After you create a FactSet DataParser connector, you can view the connector status in the Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select the FactSet DataParser connector that you created to display the connector properties and information.
4. Under **Connector status with source**, select the **Download log** link to open or save the status log for the connector. This log contains information about the data imported to the

Microsoft cloud. For more information, see View admin logs for data connectors.

Data connector limits for FactSet archiving

Importing attachments or items supports files up to 10 MB.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 06/24/2026

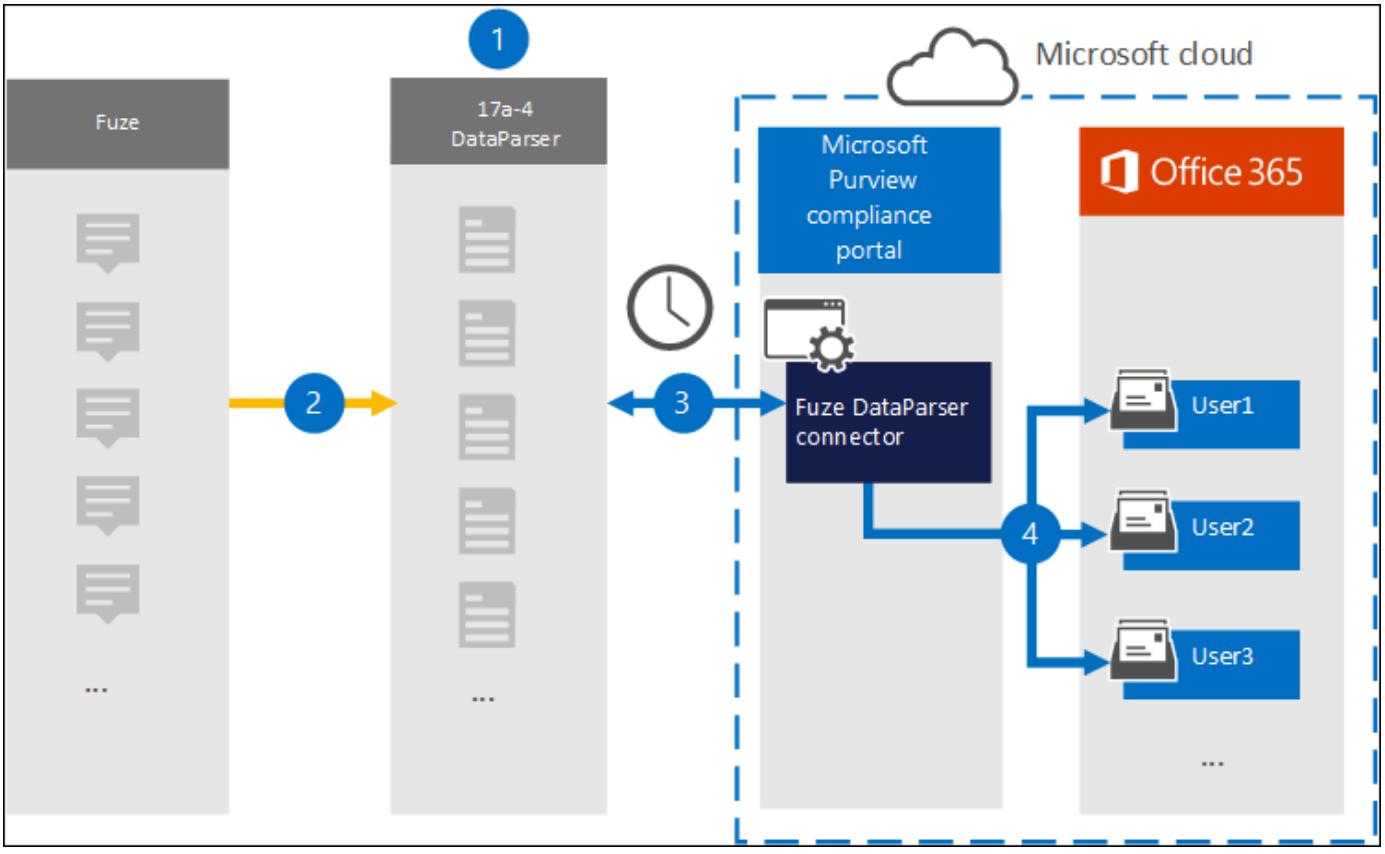
Set up a connector to archive Fuze data

Use the Fuze DataParser from 17a-4 LLC to import and archive data from Fuze to user mailboxes in your Microsoft 365 organization. The DataParser includes a Fuze connector that you configure to capture items from a third-party data source and import those items to Microsoft 365. The Fuze DataParser connector converts Fuze data to an email message format and then imports those items to user mailboxes in Microsoft 365.

After storing Fuze data in user mailboxes, you can apply Microsoft Purview features such as Litigation Hold, eDiscovery, retention policies and retention labels, and Communication Compliance. By using a Fuze connector to import and archive data in Microsoft 365, your organization can stay compliant with government and regulatory policies.

Overview of archiving Fuze data

The following overview explains the process of using a data connector to archive Fuze data in Microsoft 365.



1. Your organization works with 17a-4 to set up and configure the Fuze DataParser.

2. The DataParser regularly collects Fuze items and converts the content of a message to an email message format.

3. You create the Fuze DataParser connector in the Microsoft Purview portal. The connector connects to DataParser and transfers the messages to a secure Azure Storage location in the Microsoft cloud.

4. The connector creates a subfolder in the Inbox folder named **Fuze DataParser** in the user mailboxes and imports the Fuze items to that folder. The connector uses the value of the *Email* property to determine which mailbox to import items to. Every Fuze item contains this property, which is populated with the email address of every participant.

Before you begin

> **Important**
>
> All partner third-party data connector configurations require a provisioned service principal for the partner application. This requirement isn't applicable for native Microsoft third-party data connector configurations.
>
> To provision a service principal using PowerShell, see Provision a service principal for partner third-party data connectors.

- Create a DataParser account for Microsoft connectors. To create an account, contact 17a-4 LLC. You need to sign in to this account when you set up the Fuze DataParser connector in the Microsoft Purview portal.

- Assign the Data Connector Admin role to the user who creates the Fuze DataParser connector in the Microsoft Purview portal and completes the user-mapping configuration. This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions, see Permissions in the Microsoft Purview portal and Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview.

- This data connector is available in GCC environments in the Microsoft 365 US Government cloud. Third-party applications and services might involve storing, transmitting, and

processing your organization's customer data on third-party systems that are outside of the Microsoft 365 infrastructure and therefore aren't covered by the Microsoft Purview and data protection commitments. Microsoft makes no representation that use of this product to connect to third-party applications implies that those third-party applications are FedRAMP compliant.

Step 1: Set up a Fuze DataParser connector

Create a 17a-4 connector for Fuze data in the Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. Select **Fuze DataParser** from the list.
5. Select **Accept** on the **Terms of service** page.
6. Enter a unique name that identifies the connector, then select **Next**.
7. Sign in to your 17a-4 account and complete the steps in the Fuze DataParser connection workflow.

Step 2: Configure the Fuze DataParser connector

Work with 17a-4 Support to configure the Fuze DataParser connector.

Step 3: Map Fuze users to Microsoft 365 users

The Fuze DataParser connector automatically maps users to their Microsoft 365 email addresses before importing data to Microsoft 365.

Step 4: Monitor the Fuze DataParser connector

After you create a Fuze DataParser connector, you can view the connector status in the Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select the Fuze DataParser connector that you created to display the flyout page, which contains the properties and information about the connector.
4. Under **Connector status with source**, select the **Download log** link to open (or save) the status log for the connector. This log contains information about the data that's imported to

the Microsoft cloud. For more information, see View admin logs for data connectors.

Fuze data connector limits

Importing attachments or items supports files up to 10 MB.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/24/2026

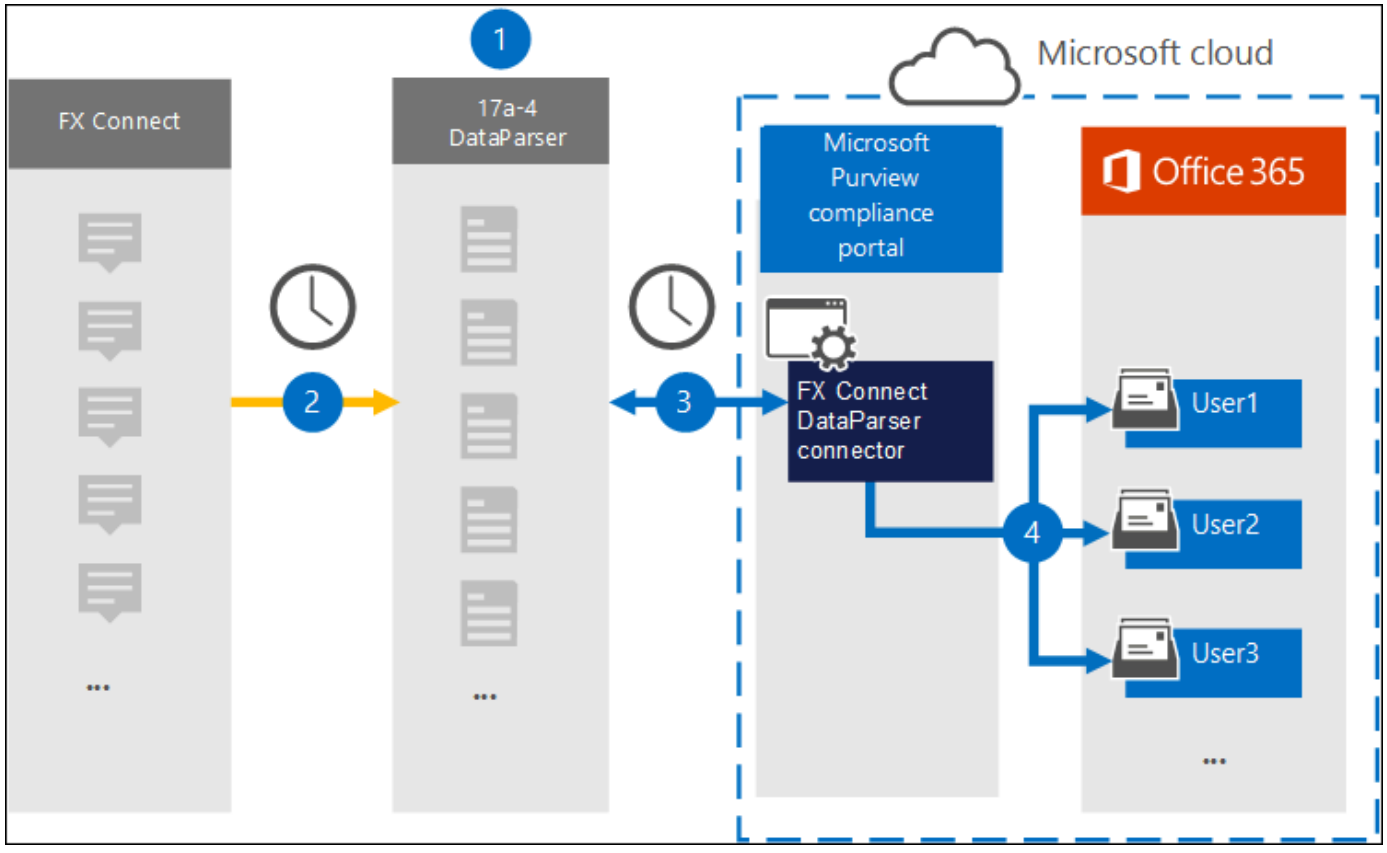
Set up a connector to archive FX Connect data

Use the FX Connect DataParser from 17a-4 LLC to import and archive data from FX Connect to user mailboxes in your Microsoft 365 organization. The DataParser includes a FX Connect connector that's configured to capture items from a third-party data source and import those items to Microsoft 365. The FX Connect DataParser connector converts FX Connect data to an email message format and then imports those items to user mailboxes in Microsoft 365.

After you store FX Connect data in user mailboxes, you can apply Microsoft Purview features such as Litigation Hold, eDiscovery, retention policies and retention labels, and Communication Compliance. By using a FX Connect connector to import and archive data in Microsoft 365, your organization can stay compliant with government and regulatory policies.

Overview of archiving FX Connect data

The following overview explains the process of using a data connector to archive FX Connect data in Microsoft 365.



1. Your organization works with 17a-4 to set up and configure the FX Connect DataParser.

2. The DataParser regularly collects FX Connect items. The DataParser also converts the content of a message to an email message format.

3. You create the FX Connect DataParser connector in the Microsoft Purview portal. The connector connects to DataParser and transfers the messages to a secure Azure Storage location in the Microsoft cloud.

4. The connector creates a subfolder in the Inbox folder named **FX Connect DataParser** in the user mailboxes, and imports the FX Connect items to that folder. The connector uses the value of the *Email* property to determine which mailbox to import items to. Every FX Connect item contains this property, which is populated with the email address of every participant.

Before you begin

> **Important**
>
> All partner third-party data connector configurations require a provisioned **service principal** for the partner application. This requirement isn't applicable for native Microsoft third-party data connector configurations.
>
> To provision a service principal using PowerShell, see **Provision a service principal for partner third-party data connectors**.

- Create a DataParser account for Microsoft connectors. To do this, contact 17a-4 LLC. You need to sign in to this account when you create the connector in the Set up a FX Connect DataParser connector procedure.

- Assign the Data Connector Admin role to the user who creates the FX Connect DataParser connector in Step 1 (and completes it in Step 3). This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions, see Permissions in the Microsoft Purview portal and Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview.

- This data connector is available in GCC environments in the Microsoft 365 US Government cloud. Third-party applications and services might involve storing, transmitting, and

processing your organization's customer data on third-party systems that are outside of the Microsoft 365 infrastructure and therefore aren't covered by the Microsoft Purview and data protection commitments. Microsoft makes no representation that use of this product to connect to third-party applications implies that those third-party applications are FedRAMP compliant.

Step 1: Set up a FX Connect DataParser connector

Create a 17a-4 connector for FX Connect data in the Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. From the list, select **FX Connect DataParser**.
5. On the **Terms of service** page, select **Accept**.
6. Enter a unique name that identifies the connector and then select **Next**.
7. Sign in to your 17a-4 account and complete the steps in the FX Connect DataParser connection workflow.

Step 2: Configure the FX Connect DataParser connector

Work with 17a-4 Support to configure the FX Connect DataParser connector.

Step 3: Map FX Connect users to Microsoft 365 users

The FX Connect DataParser connector automatically maps users to their Microsoft 365 email addresses before importing data to Microsoft 365.

Step 4: Monitor the FX Connect DataParser connector

After you create a FX Connect DataParser connector, you can view the connector status in the Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select the FX Connect DataParser connector that you created to display the flyout page, which contains the properties and information about the connector.
4. Under **Connector status with source**, select the **Download log** link to open (or save) the status log for the connector. This log contains information about the data that's been

imported to the Microsoft cloud. For more information, see View admin logs for data connectors.

Connector limits for archiving FX Connect data

Importing attachments or items supports files up to 10 MB.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/20/2026

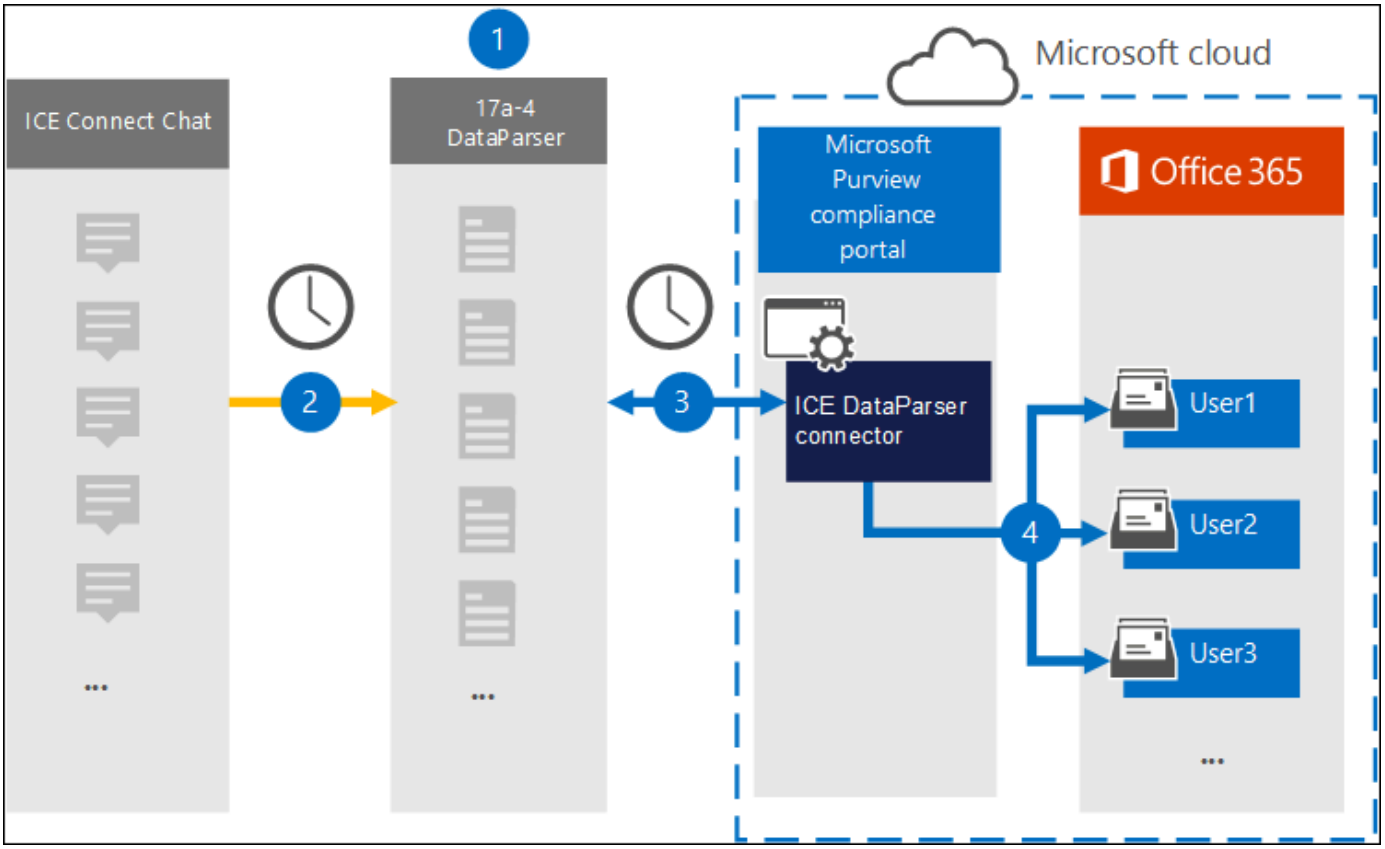
Set up a connector to archive ICE Connect Chat data

Use the ICE DataParser from 17a-4 LLC to import and archive data from ICE Connect Chat to user mailboxes in your Microsoft 365 organization. The DataParser includes an ICE Chat connector that's configured to capture items from a third-party data source and import those items to Microsoft 365. The ICE DataParser connector converts ICE Connect Chat data to an email message format and then imports those items to user mailboxes in Microsoft 365.

After you store ICE Connect Chat data in user mailboxes, you can apply Microsoft Purview features such as Litigation Hold, eDiscovery, retention policies and retention labels, and Communication Compliance. By using an ICE DataParser connector to import and archive data in Microsoft 365, you can help your organization stay compliant with government and regulatory policies.

Overview of archiving ICE Chat data

The following overview explains the process of using a data connector to archive ICE Connect Chat data in Microsoft 365.



1. Your organization works with 17a-4 to set up and configure the ICE DataParser.
2. The DataParser regularly collects ICE Connect Chat items. The DataParser also converts the content of a message to an email message format.
3. You create the ICE DataParser connector in the Microsoft Purview portal. The connector connects to DataParser and transfers the messages to a secure Azure Storage location in the Microsoft cloud.
4. The connector creates a subfolder in the Inbox folder named **ICE DataParser** in the user mailboxes, and imports the ICE Connect Chat items to that folder. The connector uses the value of the *Email* property to determine which mailbox to import items to. Every ICE Connect Chat item contains this property, which is populated with the email address of every participant.

Before you begin

> **Important**
>
> All partner third-party data connector configurations require a provisioned service principal for the partner application. This requirement isn't applicable for native Microsoft third-party data connector configurations.
>
> To provision a service principal using PowerShell, see **Provision a service principal for partner third-party data connectors**.

- Create a DataParser account for Microsoft connectors. To create an account, contact 17a-4 LLC. You need to sign in to this account when you set up the ICE DataParser connector in the Microsoft Purview portal.
- Assign the Data Connector Admin role to the user who creates the ICE DataParser connector in the Microsoft Purview portal and completes the user-mapping configuration. This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions, see Permissions in the Microsoft Purview portal and Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview.

- This data connector is available in GCC environments in the Microsoft 365 US Government cloud. Third-party applications and services might involve storing, transmitting, and processing your organization's customer data on third-party systems that are outside of the Microsoft 365 infrastructure and therefore aren't covered by the Microsoft Purview and data protection commitments. Microsoft makes no representation that use of this product to connect to third-party applications implies that those third-party applications are FedRAMP compliant.

Step 1: Set up an ICE DataParser connector

Create a 17a-4 connector for ICE Connect Chat data in the Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. Select **ICE DataParser** from the list.
5. Select **Accept** on the **Terms of service** page.
6. Enter a unique name that identifies the connector, then select **Next**.
7. Sign in to your 17a-4 account and complete the steps in the ICE DataParser connection workflow.

Step 2: Configure the ICE DataParser connector

Work with 17a-4 Support to configure the ICE DataParser connector.

Step 3: Map ICE Connect Chat users to Microsoft 365 users

The ICE DataParser connector automatically maps users to their Microsoft 365 email addresses before importing data to Microsoft 365.

Step 4: Monitor the ICE DataParser connector

After you create an ICE DataParser connector, you can view the connector status in the Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.

3. Select **My connectors**, then select the ICE DataParser connector that you created to display the flyout page, which contains the properties and information about the connector.
4. Under **Connector status with source**, select the **Download log** link to open (or save) the status log for the connector. This log contains information about the data that's been imported to the Microsoft cloud. For more information, see View admin logs for data connectors.

ICE Connect Chat connector limits

Importing attachments or items supports files up to 10 MB.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/24/2026

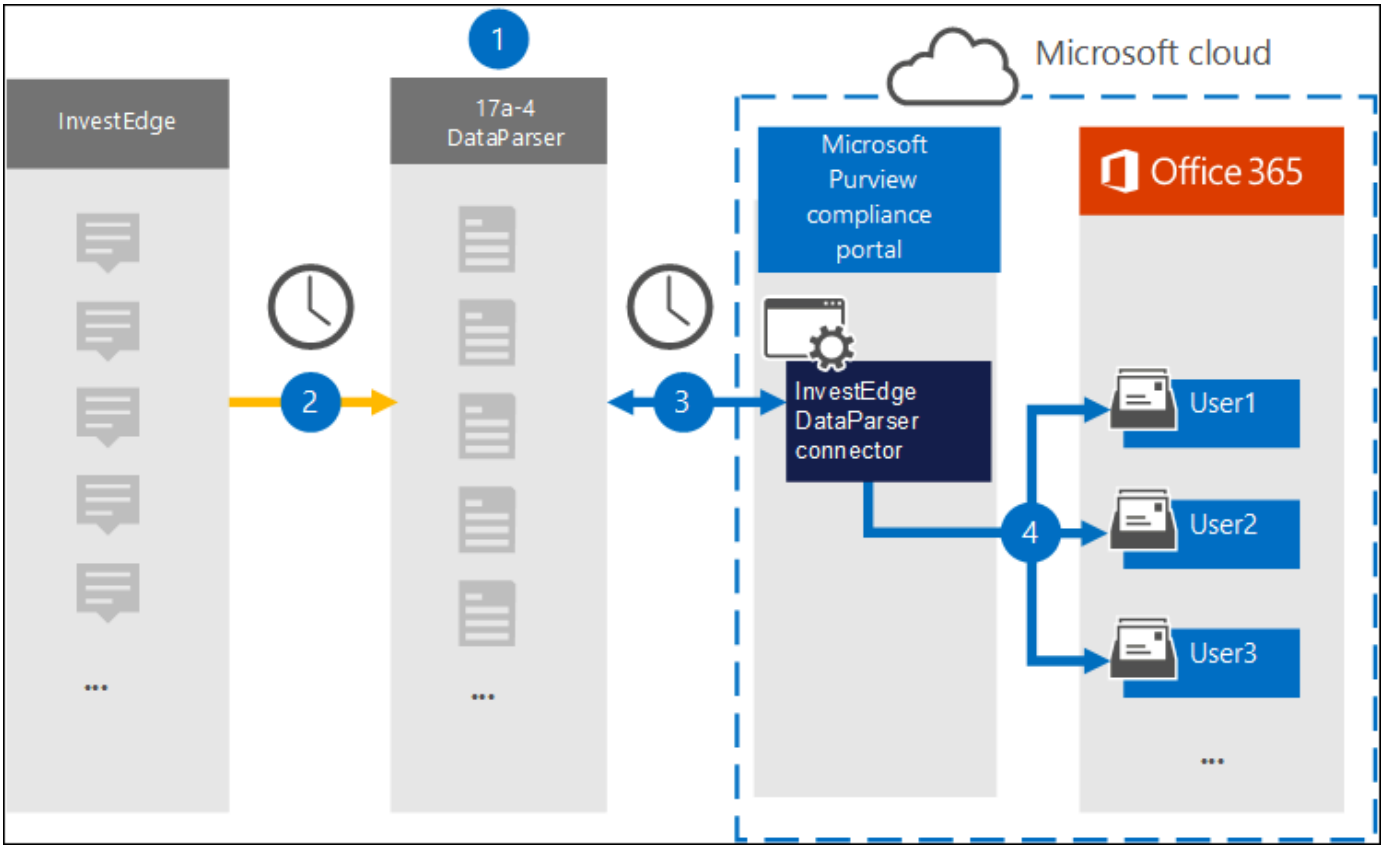
Set up a connector to archive InvestEdge data

Use the InvestEdge DataParser from 17a-4 LLC to import and archive data from InvestEdge to user mailboxes in your Microsoft 365 organization. The DataParser includes an InvestEdge connector that's configured to capture items from a third-party data source and import those items to Microsoft 365. The InvestEdge DataParser connector converts InvestEdge data to an email message format and then imports those items to user mailboxes in Microsoft 365.

After storing InvestEdge data in user mailboxes, you can apply Microsoft Purview features such as Litigation Hold, eDiscovery, retention policies and retention labels, and Communication Compliance. By using an InvestEdge connector to import and archive data in Microsoft 365, your organization can stay compliant with government and regulatory policies.

Overview of archiving InvestEdge data

The following overview explains the process of using a data connector to archive InvestEdge data in Microsoft 365.



1. Your organization works with 17a-4 to set up and configure the InvestEdge DataParser.

2. The DataParser regularly collects InvestEdge items and converts the content of a message to an email message format.

3. You create the InvestEdge DataParser connector in the Microsoft Purview portal. The connector connects to DataParser and transfers the messages to a secure Azure Storage location in the Microsoft cloud.

4. The connector creates a subfolder in the Inbox folder named **InvestEdge DataParser** in the user mailboxes and imports the InvestEdge items to that folder. The connector uses the value of the *Email* property to determine which mailbox to import items to. Every InvestEdge item contains this property, which is populated with the email address of every participant.

Before you begin

> **Important**
>
> All partner third-party data connector configurations require a provisioned service principal for the partner application. This requirement isn't applicable for native Microsoft third-party data connector configurations.
>
> To provision a service principal using PowerShell, see Provision a service principal for partner third-party data connectors.

- Create a DataParser account for Microsoft connectors. To create an account, contact 17a-4 LLC. You need to sign in to this account when you set up the InvestEdge DataParser connector in the Microsoft Purview portal.

- Assign the Data Connector Admin role to the user who sets up the InvestEdge DataParser connector in the Microsoft Purview portal and completes the user-mapping configuration. This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions, see Permissions in the Microsoft Purview portal and Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview.

- This data connector is available in GCC environments in the Microsoft 365 US Government cloud. Third-party applications and services might involve storing, transmitting, and processing your organization's customer data on third-party systems that are outside of the Microsoft 365 infrastructure and therefore aren't covered by the Microsoft Purview and data protection commitments. Microsoft makes no representation that use of this product to connect to third-party applications implies that those third-party applications are FedRAMP compliant.

Step 1: Set up an InvestEdge DataParser connector

Create a 17a-4 connector for InvestEdge data in the Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. From the list, select **InvestEdge DataParser**.
5. On the **Terms of service** page, select **Accept**.
6. Enter a unique name that identifies the connector and then select **Next**.
7. Sign in to your 17a-4 account and complete the steps in the InvestEdge DataParser connection workflow.

Step 2: Configure the InvestEdge DataParser connector

Work with 17a-4 Support to configure the InvestEdge DataParser connector.

Step 3: Map InvestEdge users to Microsoft 365 users

The InvestEdge DataParser connector automatically maps users to their Microsoft 365 email addresses before importing data to Microsoft 365.

Step 4: Monitor the InvestEdge DataParser connector

After you create an InvestEdge DataParser connector, you can view the connector status in the Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select the InvestEdge DataParser connector that you created to display the flyout page, which contains the properties and information about the connector.

4. Under **Connector status with source**, select the **Download log** link to open (or save) the status log for the connector. This log contains information about the data that's imported to the Microsoft cloud. For more information, see View admin logs for data connectors.

InvestEdge DataParser connector limits

Importing attachments or items supports files up to 10 MB.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/24/2026

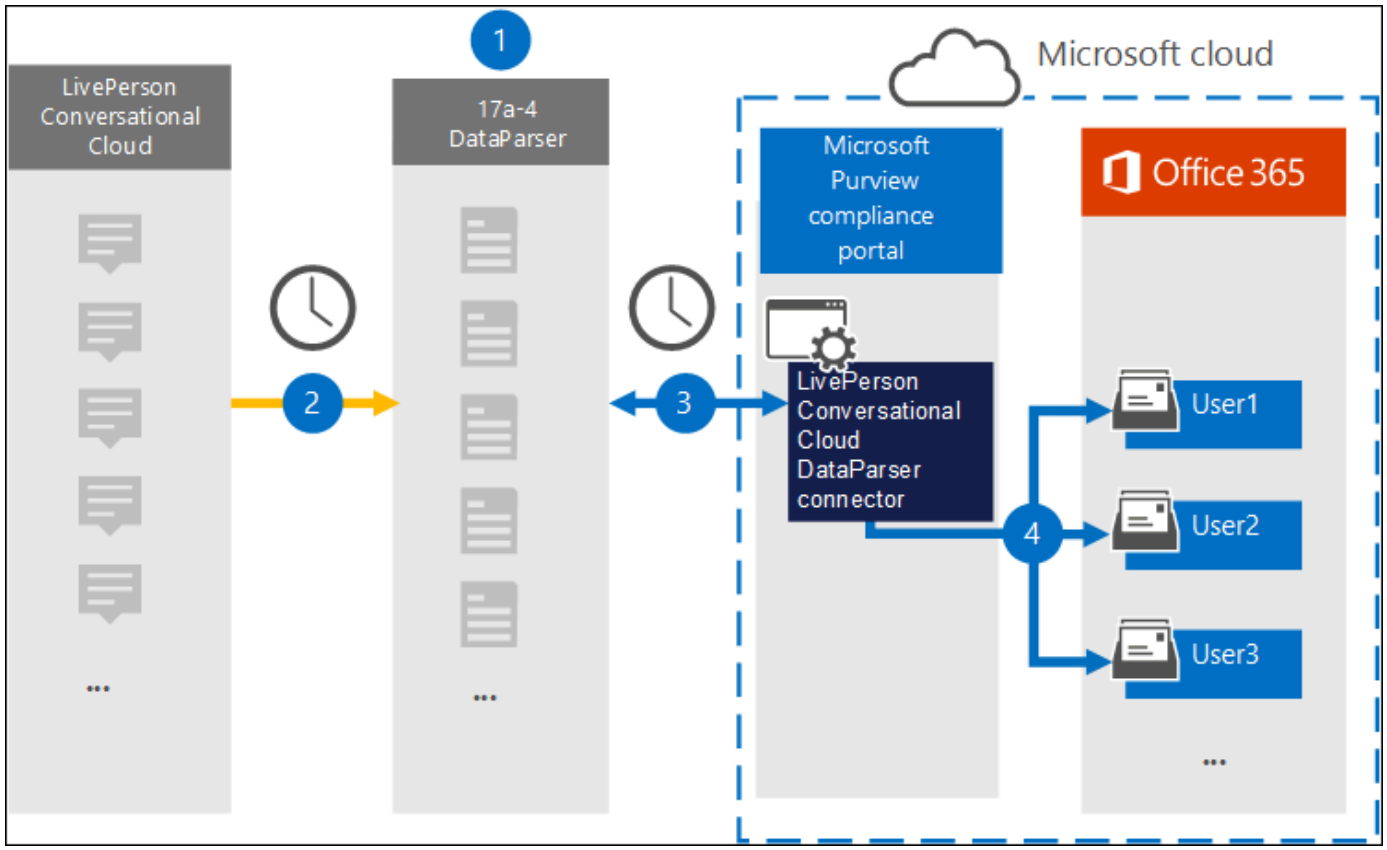
Set up a connector to archive LivePerson Conversational Cloud data

Use the LivePerson Conversational Cloud DataParser from 17a-4 LLC to import and archive data from LivePerson Conversational Cloud to user mailboxes in your Microsoft 365 organization. The DataParser includes a LivePerson Conversational Cloud connector that's configured to capture items from a third-party data source and import those items to Microsoft 365. The LivePerson Conversational Cloud DataParser connector converts data to an email message format and then imports those items to user mailboxes in Microsoft 365.

After you store data in user mailboxes, you can apply Microsoft Purview features such as Litigation Hold, eDiscovery, retention policies and retention labels, and Communication Compliance. By using a LivePerson Conversational Cloud connector to import and archive data in Microsoft 365, you help your organization stay compliant with government and regulatory policies.

Overview of archiving LivePerson Conversational Cloud data

The following overview explains the process of using a data connector to archive LivePerson Conversational Cloud data in Microsoft 365.



1. Your organization works with 17a-4 to set up and configure the LivePerson Conversational Cloud DataParser.

2. The DataParser regularly collects LivePerson Conversational Cloud items and converts the content of a message to an email message format.

3. You create the LivePerson Conversational Cloud DataParser connector in the Microsoft Purview portal. The connector connects to DataParser and transfers the messages to a secure Azure Storage location in the Microsoft cloud.

4. The connector creates a subfolder in the Inbox folder named **LivePerson Conversational Cloud DataParser** in the user mailboxes and imports the items to that folder. The connector uses the value of the *Email* property to determine which mailbox to import items to. Every item contains this property, which is populated with the email address of every participant.

Before you begin

> **Important**
>
> All partner third-party data connector configurations require a provisioned service principal for the partner application. This requirement isn't applicable for native Microsoft third-party data connector configurations.

To provision a service principal using PowerShell, see **Provision a service principal for partner third-party data connectors**.

- Create a DataParser account for Microsoft connectors. To create an account, contact 17a-4 LLC. You need to sign in to this account when you set up the LivePerson Conversational Cloud DataParser connector in the Microsoft Purview portal.
- Assign the Data Connector Admin role to the user who sets up and completes the LivePerson Conversational Cloud DataParser connector configuration in the Microsoft Purview portal. This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions, see Permissions in the Microsoft Purview portal and Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview.
- This data connector is available in GCC environments in the Microsoft 365 US Government cloud. Third-party applications and services might involve storing, transmitting, and processing your organization's customer data on third-party systems that are outside of the Microsoft 365 infrastructure and therefore aren't covered by the Microsoft Purview and data protection commitments. Microsoft makes no representation that use of this product to connect to third-party applications implies that those third-party applications are FedRAMP compliant.

Step 1: Set up a LivePerson Conversational Cloud DataParser connector

Create a 17a-4 connector for LivePerson Conversational Cloud data in the Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. From the list, select **LivePerson Conversational Cloud DataParser**.
5. On the **Terms of service** page, select **Accept**.
6. Enter a unique name that identifies the connector and then select **Next**.

7. Sign in to your 17a-4 account and complete the steps in the LivePerson Conversational Cloud DataParser connection workflow.

Step 2: Configure the LivePerson Conversational Cloud DataParser connector

Work with 17a-4 Support to configure the LivePerson Conversational Cloud DataParser connector.

Step 3: Map LivePerson Conversational Cloud users to Microsoft 365 users

The LivePerson Conversational Cloud DataParser connector automatically maps users to their Microsoft 365 email addresses before importing data to Microsoft 365.

Step 4: Monitor the LivePerson Conversational Cloud DataParser connector

After you create a LivePerson Conversational Cloud DataParser connector, you can view the connector status in the Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select the LivePerson Conversational Cloud DataParser connector that you created to display the flyout page, which contains the properties and information about the connector.
4. Under **Connector status with source**, select the **Download log** link to open (or save) the status log for the connector. This log contains information about the data that's been imported to the Microsoft cloud. For more information, see View admin logs for data connectors.

Connector limits for archiving LivePerson Conversational Cloud data

Importing attachments or items supports files up to 10 MB.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

Last updated on 06/20/2026

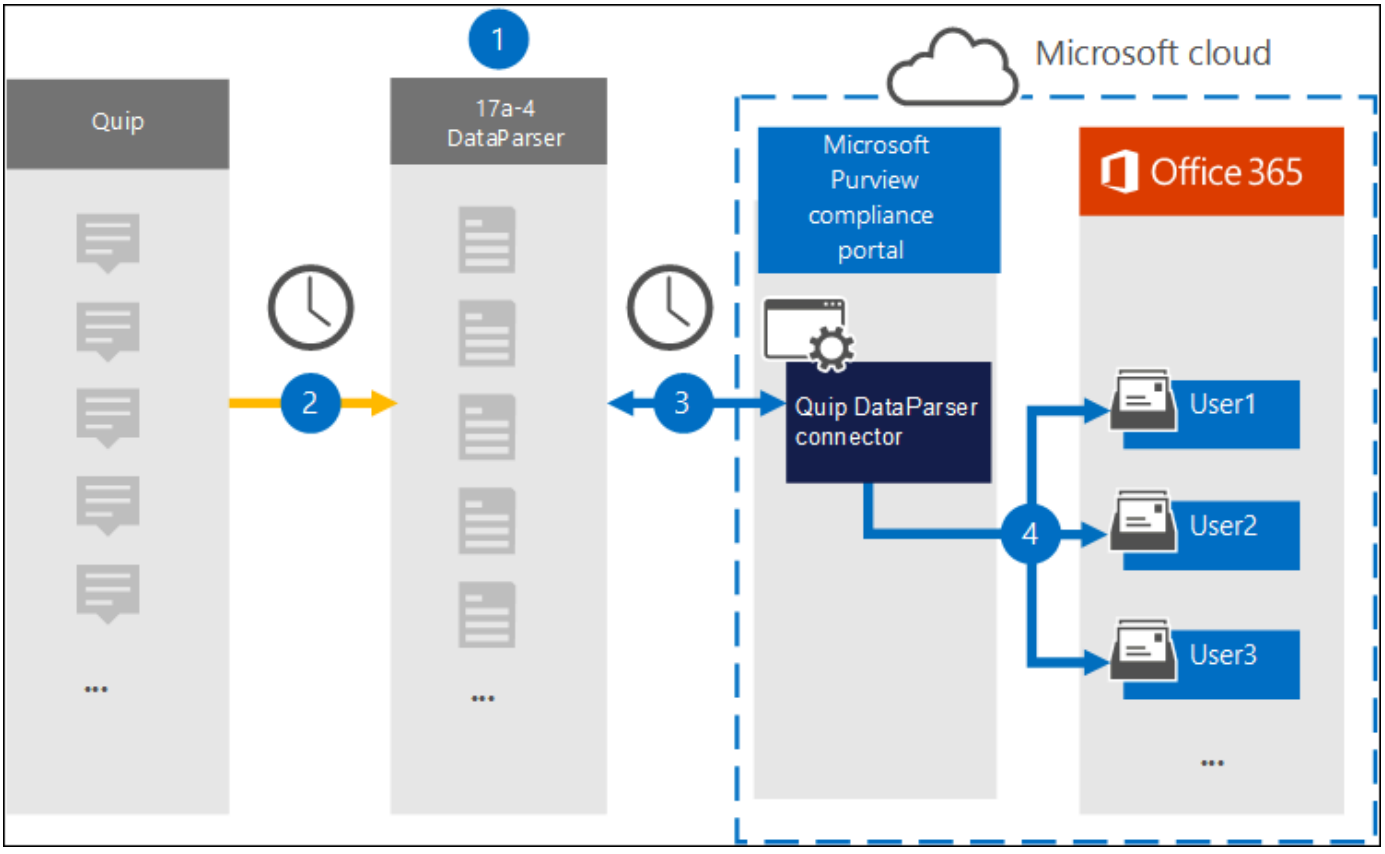
Set up a connector to archive Quip data

Use the Quip DataParser from 17a-4 LLC to import and archive data from Quip to user mailboxes in your Microsoft 365 organization. The DataParser includes a Quip connector that's configured to capture items from a third-party data source and import those items to Microsoft 365. The Quip DataParser connector converts Quip data to an email message format and then imports those items to user mailboxes in Microsoft 365.

After storing Quip data in user mailboxes, you can apply Microsoft Purview features such as Litigation Hold, eDiscovery, retention policies and retention labels, and Communication Compliance. By using a Quip connector to import and archive data in Microsoft 365, your organization can stay compliant with government and regulatory policies.

Overview of archiving Quip data

The following overview explains the process of using a data connector to archive Quip data in Microsoft 365.



1. Your organization works with 17a-4 to set up and configure the Quip DataParser.

2. The DataParser regularly collects Quip items and converts the content of a message to an email message format.

3. You create the Quip DataParser connector in the Microsoft Purview portal. The connector connects to DataParser and transfers the messages to a secure Azure Storage location in the Microsoft cloud.

4. The connector creates a subfolder in the Inbox folder named **Quip DataParser** in the user mailboxes and imports the Quip items to that folder. The connector uses the value of the *Email* property to determine which mailbox to import items to. Every Quip item contains the *Email* property, which is populated with the email address of every participant.

Before you begin

> **Important**
>
> All partner third-party data connector configurations require a provisioned service principal for the partner application. This requirement isn't applicable for native Microsoft third-party data connector configurations.
>
> To provision a service principal using PowerShell, see Provision a service principal for partner third-party data connectors.

- Create a DataParser account for Microsoft connectors. To create an account, contact 17a-4 LLC. You need to sign in to the DataParser account when you create the connector in Step 1.

- Assign the Data Connector Admin role to the user who creates the Quip DataParser connector in Step 1 (and completes it in Step 3). This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions, see Permissions in the Microsoft Purview portal and Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview.

- This data connector is available in GCC environments in the Microsoft 365 US Government cloud. Third-party applications and services might involve storing, transmitting, and processing your organization's customer data on third-party systems that are outside of the

Microsoft 365 infrastructure and therefore aren't covered by the Microsoft Purview and data protection commitments. Microsoft makes no representation that use of this product to connect to third-party applications implies that those third-party applications are FedRAMP compliant.

Step 1: Set up a Quip DataParser connector

Create the Quip DataParser connector

Create a 17a-4 connector for Quip data in the Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. From the list, select **Quip DataParser**.
5. On the **Terms of service** page, select **Accept**.
6. Enter a unique name that identifies the connector and then select **Next**.
7. Sign in to your 17a-4 account and complete the steps in the Quip DataParser connection workflow.

Step 2: Configure the Quip DataParser connector

Work with 17a-4 Support to configure the Quip DataParser connector.

Step 3: Map Quip users to Microsoft 365 users

The Quip DataParser connector automatically maps users to their Microsoft 365 email addresses before importing data to Microsoft 365.

Step 4: Monitor the Quip DataParser connector

After you create a Quip DataParser connector, you can view the connector status in the Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select the Quip DataParser connector that you created to display the flyout page, which contains the properties and information about the connector.

4. Under **Connector status with source**, select the **Download log** link to open (or save) the status log for the connector. This log contains information about the data that's been imported to the Microsoft cloud. For more information, see View admin logs for data connectors.

Connector limits for archiving Quip data

Importing attachments or items supports files up to 10 MB.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/20/2026

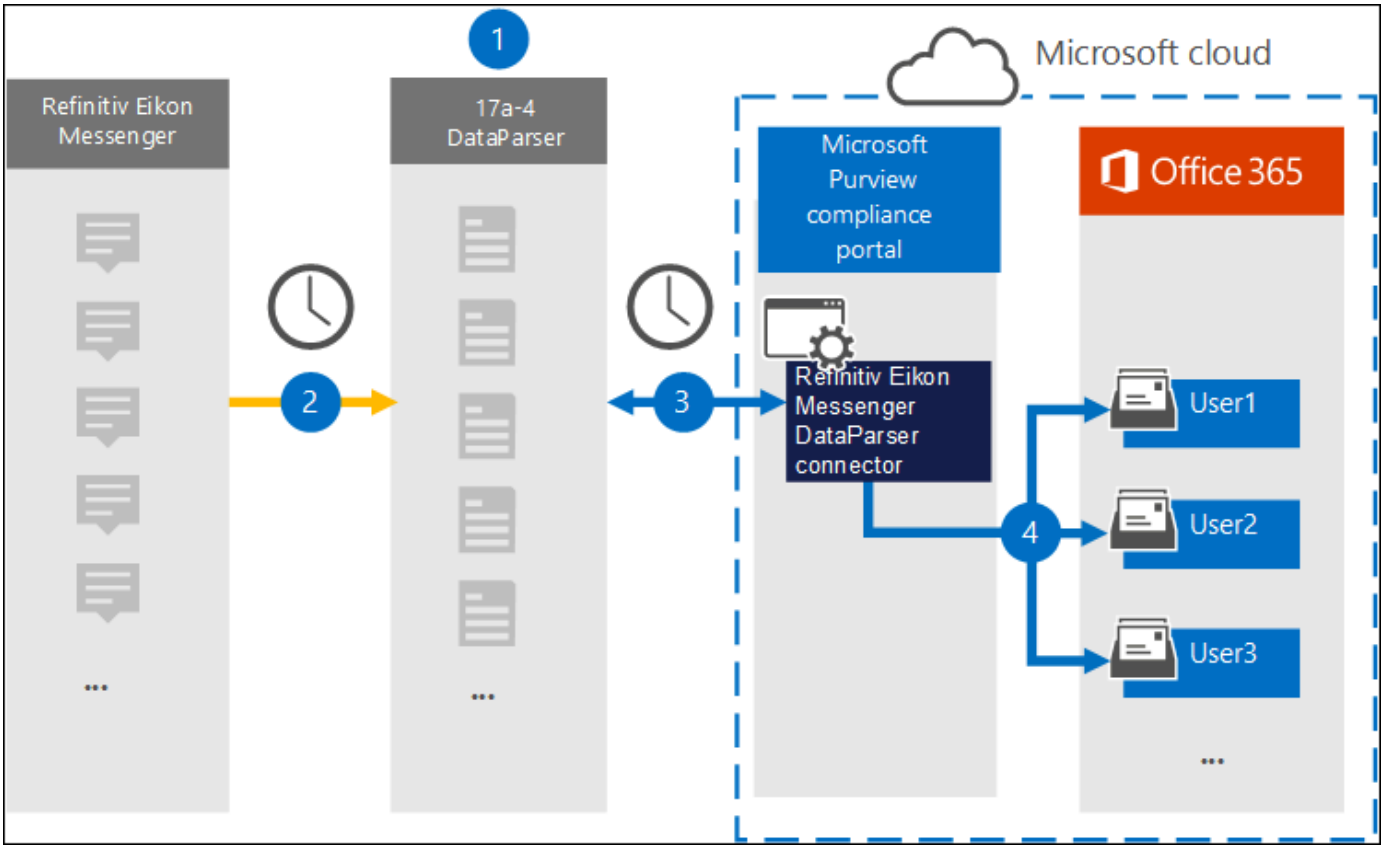
Set up a connector to archive Refinitiv Eikon Messenger data

Use the Refinitiv Eikon Messenger DataParser from 17a-4 LLC to import and archive data from Refinitiv Eikon Messenger to user mailboxes in your Microsoft 365 organization. The DataParser includes a Refinitiv Eikon Messenger connector that's configured to capture items from a third-party data source and import those items to Microsoft 365. The Refinitiv Eikon Messenger DataParser connector converts Refinitiv Eikon Messenger data to an email message format and then imports those items to user mailboxes in Microsoft 365.

After you store Refinitiv Eikon Messenger data in user mailboxes, you can apply Microsoft Purview features such as Litigation Hold, eDiscovery, retention policies and retention labels, and Communication Compliance. By using a Refinitiv Eikon Messenger connector to import and archive data in Microsoft 365, you can help your organization stay compliant with government and regulatory policies.

Overview of archiving Refinitiv Eikon Messenger data

The following overview explains the process of using a data connector to archive Refinitiv Eikon Messenger data in Microsoft 365.



1. Your organization works with 17a-4 to set up and configure the Refinitiv Eikon Messenger DataParser.
2. The DataParser regularly collects Refinitiv Eikon Messenger items. The DataParser also converts the content of a message to an email message format.
3. The Refinitiv Eikon Messenger DataParser connector that you create in the Microsoft Purview portal connects to DataParser and transfers the messages to a secure Azure Storage location in the Microsoft cloud.
4. The connector creates a subfolder in the Inbox folder named **Refinitiv Eikon Messenger DataParser** in the user mailboxes and imports the Refinitiv Eikon Messenger items to that folder. The connector uses the value of the *Email* property to determine which mailbox to import items to. Every Refinitiv Eikon Messenger item contains this property, which is populated with the email address of every participant.

Before you begin

> **Important**
>
> All partner third-party data connector configurations require a provisioned service principal for the partner application. This requirement isn't applicable for native Microsoft third-party data connector configurations.
>
> To provision a service principal using PowerShell, see Provision a service principal for partner third-party data connectors.

- Create a DataParser account for Microsoft connectors. To create an account, contact 17a-4 LLC. You need to sign in to this account when you create the connector in the Microsoft Purview portal.
- Assign the Data Connector Admin role to the user who creates and completes the Refinitiv Eikon Messenger DataParser connector setup in the Microsoft Purview portal. This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions, see Permissions in the

Microsoft Purview portal and Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview.

- This data connector is available in GCC environments in the Microsoft 365 US Government cloud. Third-party applications and services might involve storing, transmitting, and processing your organization's customer data on third-party systems that are outside of the Microsoft 365 infrastructure and therefore aren't covered by the Microsoft Purview and data protection commitments. Microsoft makes no representation that use of this product to connect to third-party applications implies that those third-party applications are FedRAMP compliant.

Step 1: Set up a Refinitiv Eikon Messenger DataParser connector

Create a 17a-4 connector for Refinitiv Eikon Messenger data in the Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. From the list, select **Refinitiv Eikon Messenger DataParser**.
5. On the **Terms of service** page, select **Accept**.
6. Enter a unique name that identifies the connector and then select **Next**.
7. Sign in to your 17a-4 account and complete the steps in the Refinitiv Eikon Messenger DataParser connection workflow.

Step 2: Configure the Refinitiv Eikon Messenger DataParser connector

Work with 17a-4 Support to configure the Refinitiv Eikon Messenger DataParser connector.

Step 3: Map Refinitiv Eikon Messenger users to Microsoft 365 users

The Refinitiv Eikon Messenger DataParser connector automatically maps users to their Microsoft 365 email addresses before importing data to Microsoft 365.

Step 4: Monitor the Refinitiv Eikon Messenger DataParser connector

After you create a Refinitiv Eikon Messenger DataParser connector, you can view the connector status in the Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select the Refinitiv Eikon Messenger DataParser connector that you created to display the flyout page, which contains the properties and information about the connector.
4. Under **Connector status with source**, select the **Download log** link to open (or save) the status log for the connector. This log contains information about the data that's imported to the Microsoft cloud. For more information, see View admin logs for data connectors.

Connector limits for Refinitiv Eikon Messenger data import

Importing attachments or items supports files up to 10 MB.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 06/19/2026

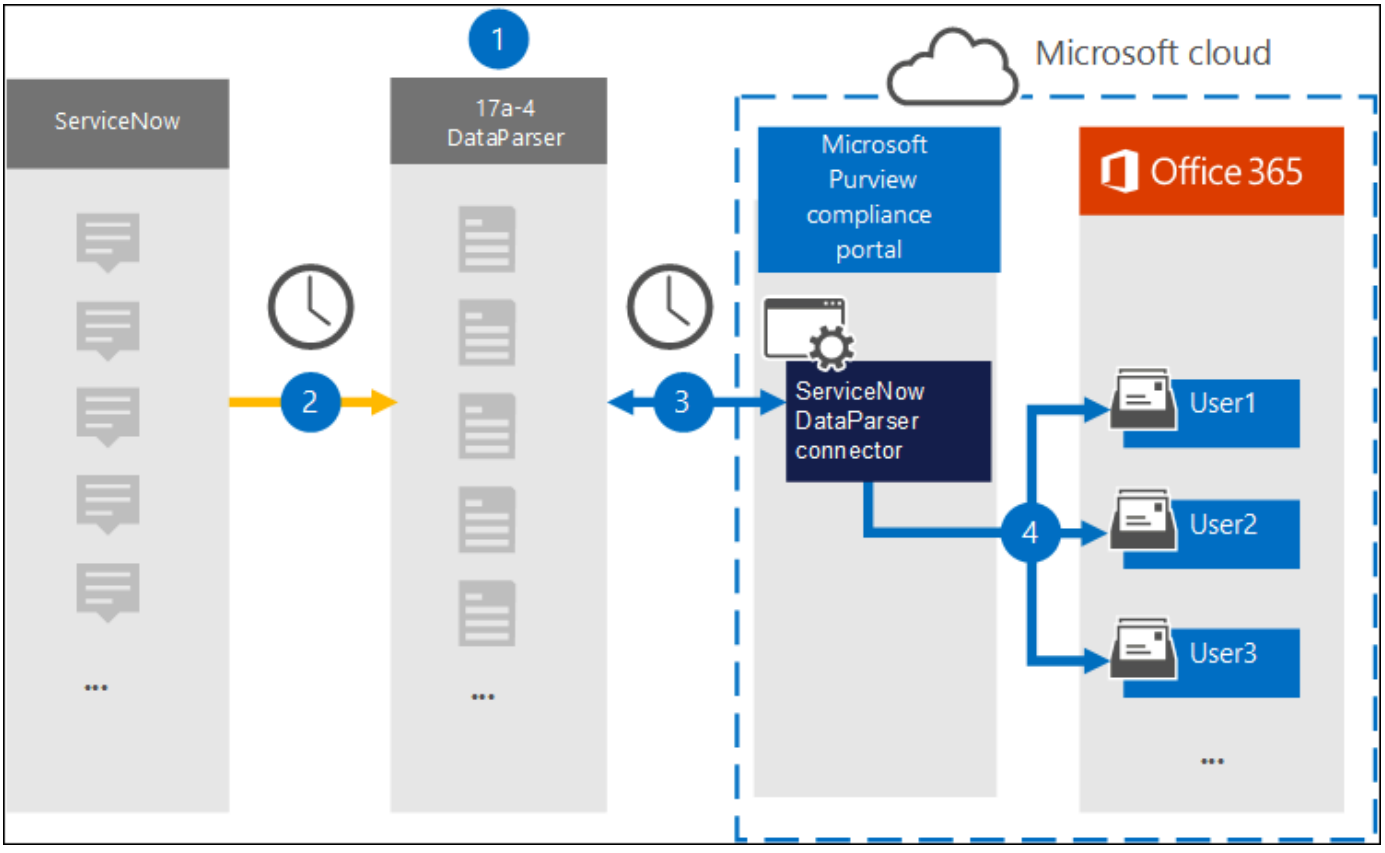
Set up a connector to archive ServiceNow data

Use the ServiceNow DataParser from 17a-4 LLC to import and archive data from ServiceNow to user mailboxes in your Microsoft 365 organization. The DataParser includes a ServiceNow connector that you configure to capture items from a third-party data source and import those items to Microsoft 365. The ServiceNow DataParser connector converts ServiceNow data to an email message format and then imports those items to user mailboxes in Microsoft 365.

After storing ServiceNow data in user mailboxes, you can apply Microsoft Purview features such as Litigation Hold, eDiscovery, retention policies and retention labels, and Communication Compliance. By using a ServiceNow connector to import and archive data in Microsoft 365, your organization can stay compliant with government and regulatory policies.

Overview of archiving ServiceNow data

The following overview explains the process of using a data connector to archive ServiceNow data in Microsoft 365.



1. Your organization works with 17a-4 to set up and configure the ServiceNow DataParser.

2. The DataParser regularly collects ServiceNow items and converts the content of a message to an email message format.

3. You create the ServiceNow DataParser connector in the Microsoft Purview portal. The connector connects to DataParser and transfers the messages to a secure Azure Storage location in the Microsoft cloud.

4. The connector creates a subfolder in the Inbox folder named **ServiceNow DataParser** in the user mailboxes and imports the ServiceNow items to that folder. The connector uses the value of the *Email* property to determine which mailbox to import items to. Every ServiceNow item contains this property, which is populated with the email address of every participant.

Before you begin

> **Important**
>
> All partner third-party data connector configurations require a provisioned service principal for the partner application. This requirement isn't applicable for native Microsoft third-party data connector configurations.
>
> To provision a service principal using PowerShell, see Provision a service principal for partner third-party data connectors.

- Create a DataParser account for Microsoft connectors. To create an account, contact 17a-4 LLC. You need to sign in to this account when you set up the ServiceNow DataParser connector in the Microsoft Purview portal.

- Assign the Data Connector Admin role to the user who creates the ServiceNow DataParser connector in the Microsoft Purview portal and completes the user-mapping configuration. This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions, see Permissions in the Microsoft Purview portal and Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview.

- This data connector is available in GCC environments in the Microsoft 365 US Government cloud. Third-party applications and services might involve storing, transmitting, and processing your organization's customer data on third-party systems that are outside of the Microsoft 365 infrastructure and therefore aren't covered by the Microsoft Purview and data protection commitments. Microsoft makes no representation that use of this product to connect to third-party applications implies that those third-party applications are FedRAMP compliant.

Step 1: Set up a ServiceNow DataParser connector

Create a 17a-4 connector for ServiceNow data in the Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. From the list, select **ServiceNow DataParser**.
5. On the **Terms of service** page, select **Accept**.
6. Enter a unique name that identifies the connector and then select **Next**.
7. Sign in to your 17a-4 account and complete the steps in the ServiceNow DataParser connection workflow.

Step 2: Configure the ServiceNow DataParser connector

Work with 17a-4 Support to configure the ServiceNow DataParser connector.

Step 3: Map ServiceNow users to Microsoft 365 users

The ServiceNow DataParser connector automatically maps users to their Microsoft 365 email addresses before importing data to Microsoft 365.

Step 4: Monitor the ServiceNow DataParser connector

After you create a ServiceNow DataParser connector, you can view the connector status in the Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select the ServiceNow DataParser connector that you created to display the flyout page, which contains the properties and information about the connector.

4. Under **Connector status with source**, select the **Download log** link to open (or save) the status log for the connector. This log contains information about the data that's imported to the Microsoft cloud. For more information, see View admin logs for data connectors.

ServiceNow DataParser connector limits

Importing attachments or items supports files up to 10 MB.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/24/2026

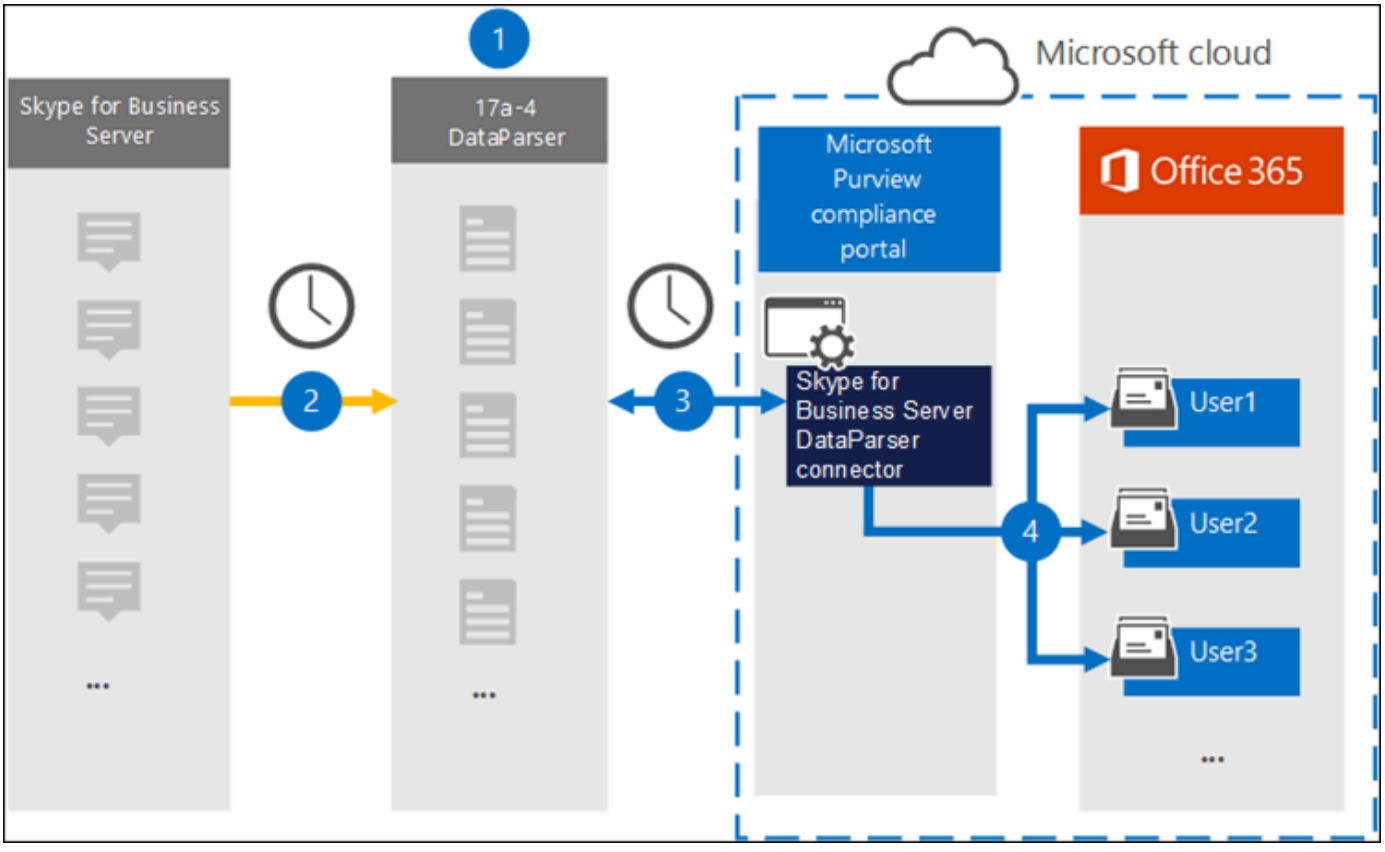
Set up a connector to archive Skype for Business Server data

Use the Skype Server DataParser from 17a-4 LLC to import and archive data from a Skype for Business Server to user mailboxes in your Microsoft 365 organization. The DataParser includes a Skype for Business connector that you configure to capture items from a third-party data source and import those items to Microsoft 365. The Skype for Business Server DataParser connector converts Skype for Business Server data to an email message format and then imports those items to user mailboxes in Microsoft 365.

After storing Skype for Business Server data in user mailboxes, you can apply Microsoft Purview features such as Litigation Hold, eDiscovery, retention policies and retention labels, and Communication Compliance. By using a Skype for Business Server connector to import and archive data in Microsoft 365, your organization can stay compliant with government and regulatory policies.

Overview of archiving Skype for Business Server data

The following overview explains the process of using a data connector to archive Skype for Business Server data in Microsoft 365.



1. Your organization works with 17a-4 to set up and configure the Skype for Business Server DataParser.

2. The DataParser regularly collects Skype for Business Server items and converts the content of a message to an email message format.

3. You create the Skype for Business Server DataParser connector in the Microsoft Purview portal. The connector connects to DataParser and transfers the messages to a secure Azure Storage location in the Microsoft cloud.

4. The connector creates a subfolder in the Inbox folder named **Skype for Business Server DataParser** in the user mailboxes and imports the Skype for Business Server items to that folder. The connector uses the value of the *Email* property to determine which mailbox to import items to. Every Skype for Business Server item contains this property, which is populated with the email address of every participant.

Before you begin

> **Important**
>
> All partner third-party data connector configurations require a provisioned service principal for the partner application. This requirement isn't applicable for native Microsoft third-party data connector configurations.
>
> To provision a service principal using PowerShell, see Provision a service principal for partner third-party data connectors.

- Create a DataParser account for Microsoft connectors. To create an account, contact 17a-4 LLC. You need to sign in to this account when you create the connector in the Microsoft Purview portal.

- Assign the Data Connector Admin role to the user who creates and completes the Skype for Business Server DataParser connector setup in the Microsoft Purview portal. This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions, see Permissions in the

Microsoft Purview portal and Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview.

- This data connector is available in GCC environments in the Microsoft 365 US Government cloud. Third-party applications and services might involve storing, transmitting, and processing your organization's customer data on third-party systems that are outside of the Microsoft 365 infrastructure and therefore aren't covered by the Microsoft Purview and data protection commitments. Microsoft makes no representation that use of this product to connect to third-party applications implies that those third-party applications are FedRAMP compliant.

Step 1: Set up a Skype for Business Server DataParser connector

Create a 17a-4 connector for Skype for Business Server data in the Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. From the list, select **Skype for Business Server DataParser**.
5. On the **Terms of service** page, select **Accept**.
6. Enter a unique name that identifies the connector and then select **Next**.
7. Sign in to your 17a-4 account and complete the steps in the Skype for Business Server DataParser connection workflow.

Step 2: Configure the Skype for Business Server DataParser connector

Work with 17a-4 Support to configure the Skype for Business Server DataParser connector.

Step 3: Map Skype for Business Server users to Microsoft 365 users

The Skype for Business Server DataParser connector automatically maps users to their Microsoft 365 email addresses before importing data to Microsoft 365.

Step 4: Monitor the Skype for Business Server DataParser connector

After you create a Skype for Business Server DataParser connector, you can view the connector status in the Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select the Skype for Business Server DataParser connector that you created to display the connector properties and information.
4. Under **Connector status with source**, select the **Download log** link to open or save the status log for the connector. This log contains information about the data that's imported to the Microsoft cloud. For more information, see View admin logs for data connectors.

Connector limits for archiving Skype for Business Server data

Importing attachments or items supports files up to 10 MB.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 06/19/2026

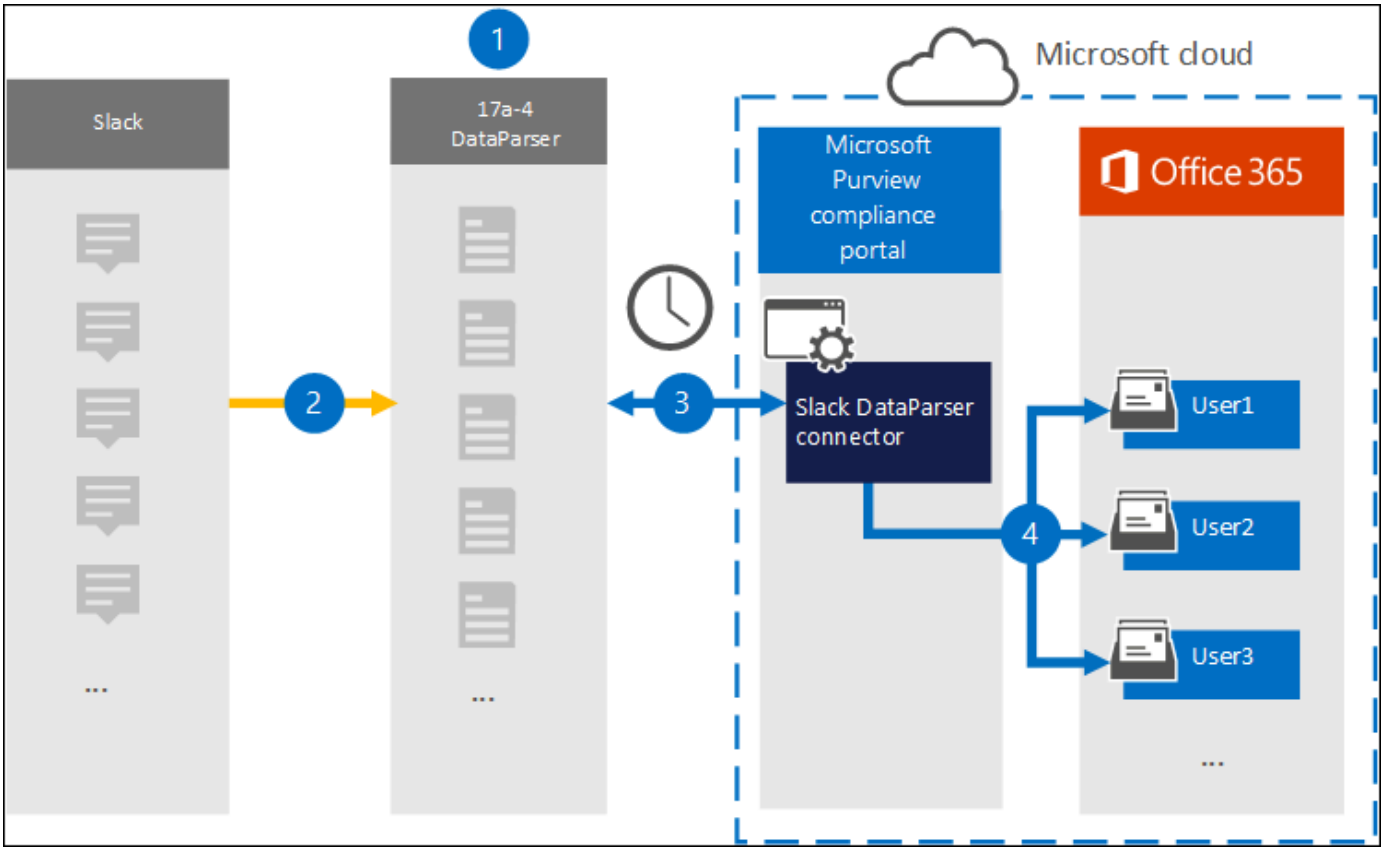
Set up a connector to archive Slack data

Use DataParser from 17a-4 LLC to import and archive data from the Slack platform to user mailboxes in your Microsoft 365 organization. DataParser includes a Slack connector that's configured to capture items from a third-party data source and import those items to Microsoft 365. The Slack DataParser connector converts Slack data to an email message format and then imports those items to user mailboxes in Microsoft 365.

After you store Slack data in user mailboxes, you can apply Microsoft Purview features such as Litigation Hold, eDiscovery, retention policies and retention labels, and Communication Compliance. By using a Slack connector to import and archive data in Microsoft 365, you help your organization stay compliant with government and regulatory policies.

Overview of archiving Slack data

The following overview explains the process of using a data connector to archive Slack data in Microsoft 365.



1. Your organization works with 17a-4 to set up and configure the Slack DataParser.

2. The DataParser regularly collects Slack items and converts the content of a message to an email message format.

3. You create the Slack DataParser connector in the Microsoft Purview portal. The connector connects to DataParser and transfers the messages to a secure Azure Storage location in the Microsoft cloud.

4. The connector creates a subfolder in the Inbox folder named **Slack DataParser** in the user mailboxes and imports the Slack items to that folder. The connector uses the value of the *Email* property to determine which mailbox to import items to. Every Slack item contains this property, which is populated with the email address of every participant.

Before you begin

> **Important**
>
> All partner third-party data connector configurations require a provisioned service principal for the partner application. This requirement isn't applicable for native Microsoft third-party data connector configurations.
>
> To provision a service principal using PowerShell, see Provision a service principal for partner third-party data connectors.

- Create a DataParser account for Microsoft connectors. To do this, contact 17a-4 LLC. You need to sign in to this account when you create the connector in the Microsoft Purview portal.

- Assign the Data Connector Admin role to the user who creates the Slack DataParser connector in the Microsoft Purview portal and completes the user-mapping configuration. This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions, see Permissions in the Microsoft Purview portal and Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview.

- This data connector is available in GCC environments in the Microsoft 365 US Government cloud. Third-party applications and services might involve storing, transmitting, and

processing your organization's customer data on third-party systems that are outside of the Microsoft 365 infrastructure and therefore aren't covered by the Microsoft Purview and data protection commitments. Microsoft makes no representation that use of this product to connect to third-party applications implies that those third-party applications are FedRAMP compliant.

Step 1: Set up a Slack DataParser connector

Create a 17a-4 connector for Slack data in the Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. Select **Slack DataParser** from the list.
5. Select **Accept** on the **Terms of service** page.
6. Enter a unique name that identifies the connector, then select **Next**.
7. Sign in to your 17a-4 account and complete the steps in the Slack DataParser connection workflow.

Step 2: Configure the Slack DataParser connector

Work with 17a-4 Support to configure the Slack DataParser connector.

Step 3: Map Slack users to Microsoft 365 users

The Slack DataParser connector automatically maps users to their Microsoft 365 email addresses before importing data to Microsoft 365.

Step 4: Monitor the Slack DataParser connector

After you create a Slack DataParser connector, you can view the connector status in the Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select the Slack DataParser connector that you created to display the flyout page, which contains the properties and information about the connector.
4. Under **Connector status with source**, select the **Download log** link to open (or save) the status log for the connector. This log contains information about the data that's been

imported to the Microsoft cloud. For more information, see View admin logs for data connectors.

Connector limits for archiving Slack data

Importing attachments or items supports files up to 10 MB.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/19/2026

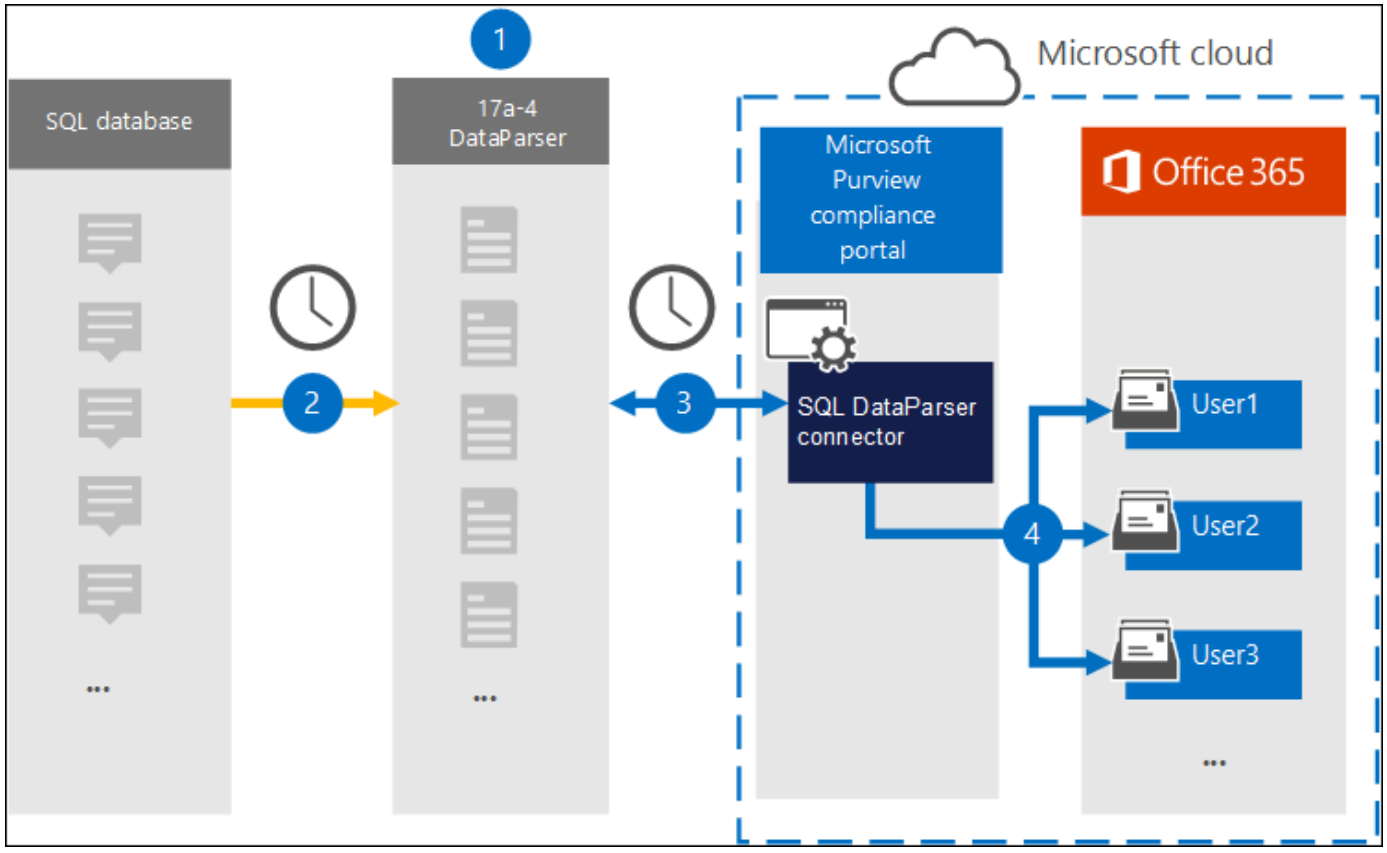
Set up a connector to archive SQL data

Use the SQL DataParser from 17a-4 LLC to import and archive data from a SQL database. The connector stores that data in user mailboxes in your Microsoft 365 organization. The DataParser captures items from the SQL source and converts them to an email message format. It then imports those items to Microsoft 365.

After SQL data is stored in user mailboxes, you can apply Microsoft Purview features. These features include Litigation Hold, eDiscovery, retention policies, retention labels, and Communication Compliance. A SQL connector helps your organization stay compliant with government and regulatory policies.

Overview of archiving SQL data

The following overview explains the process of using a data connector to archive SQL data in Microsoft 365.



1. Your organization works with 17a-4 to set up and configure the SQL DataParser.
2. The DataParser regularly collects SQL items and converts the content of a message to an email message format.

3. You create the SQL DataParser connector in the Microsoft Purview portal. The connector connects to DataParser and transfers the messages to a secure Azure Storage location in the Microsoft cloud.

4. The connector creates a subfolder in the Inbox folder named **SQL DataParser** in the user mailboxes and imports the SQL items to that folder. The connector uses the value of the *Email* property to determine which mailbox to import items to. Every SQL item contains this property, which is populated with the email address of every participant.

Before you begin

Complete the following prerequisites before you create the SQL DataParser connector:

> **Important**
>
> All partner third-party data connector configurations require a provisioned service principal for the partner application. This requirement isn't applicable for native Microsoft third-party data connector configurations.
>
> To provision a service principal using PowerShell, see Provision a service principal for partner third-party data connectors.

- Create a DataParser account for Microsoft connectors. To create an account, contact 17a-4 LLC. You need to sign in to this account when you create the connector in the Microsoft Purview portal.

- Assign the Data Connector Admin role to the user who creates the SQL DataParser connector in the Microsoft Purview portal and completes the user mapping configuration. This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions, see Permissions in the Microsoft Purview portal and Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview.

- This data connector is available in GCC environments in the Microsoft 365 US Government cloud. Third-party applications and services might involve storing, transmitting, and processing your organization's customer data on third-party systems that are outside of the

Microsoft 365 infrastructure and therefore aren't covered by the Microsoft Purview and data protection commitments. Microsoft makes no representation that use of this product to connect to third-party applications implies that those third-party applications are FedRAMP compliant.

Step 1: Set up a SQL DataParser connector

Create a 17a-4 connector for SQL data in the Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. From the list, select **SQL DataParser**.
5. On the **Terms of service** page, select **Accept**.
6. Enter a unique name that identifies the connector and then select **Next**.
7. Sign in to your 17a-4 account and complete the steps in the SQL DataParser connection workflow.

Step 2: Configure the SQL DataParser connector

Work with 17a-4 Support to configure the SQL DataParser connector.

Step 3: Map SQL source users to Microsoft 365 users

The SQL DataParser connector automatically maps users to their Microsoft 365 email addresses before importing data to Microsoft 365.

Step 4: Monitor the SQL DataParser connector

After you create a SQL DataParser connector, you can view the connector status in the Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select the SQL DataParser connector that you created to display the flyout page, which contains the properties and information about the connector.
4. Under **Connector status with source**, select the **Download log** link to open or save the status log for the connector. This log contains information about the data that's been

imported to the Microsoft cloud. For more information, see View admin logs for data connectors.

SQL data connector limits

Importing attachments or items supports files up to 10 MB.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/24/2026

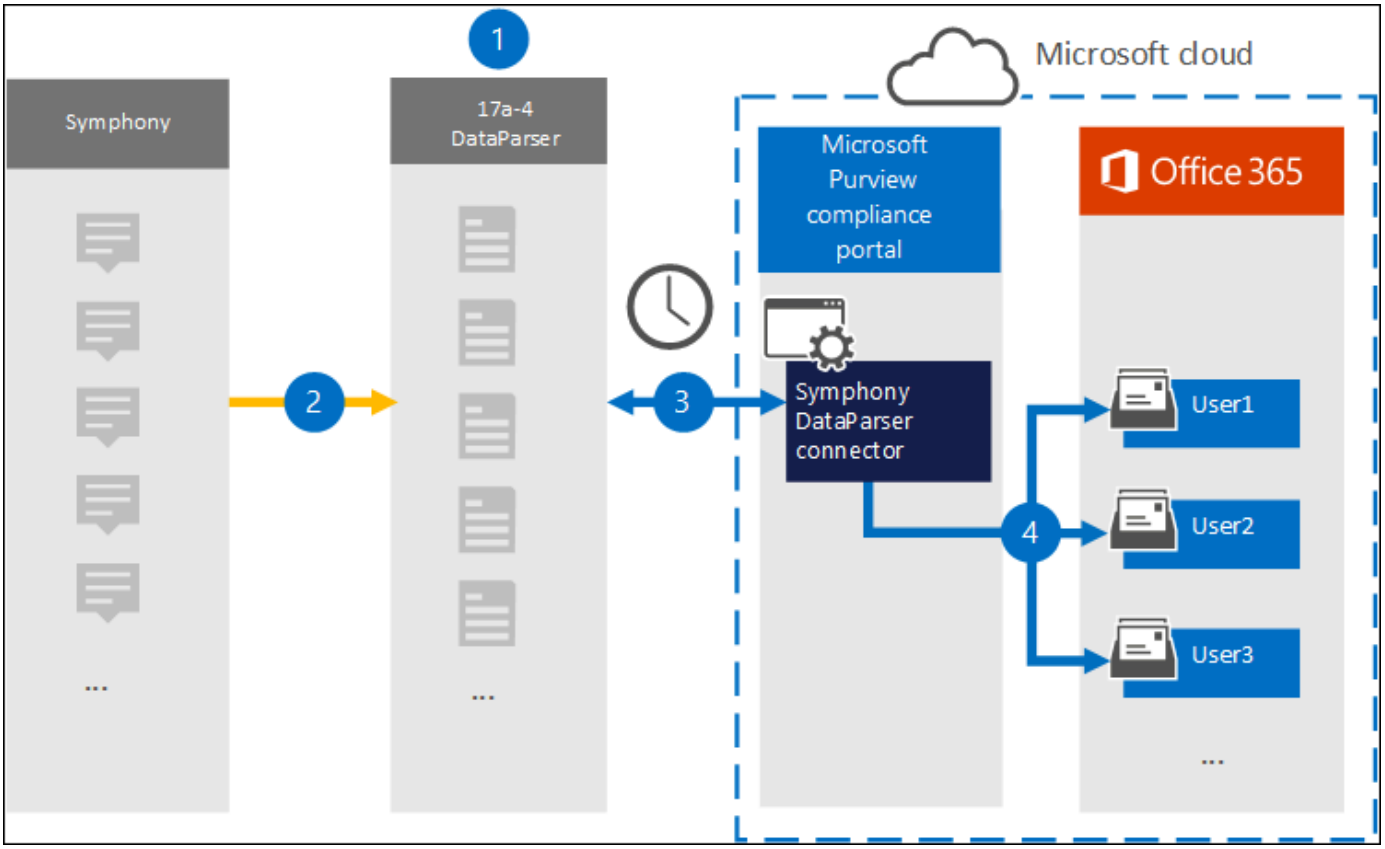
Set up a connector to archive Symphony data

Use the Symphony DataParser from 17a-4 LLC to import and archive Symphony communications data to user mailboxes in your Microsoft 365 organization. The DataParser includes a Symphony connector that you configure to capture items from a third-party data source and import those items to Microsoft 365. The Symphony DataParser connector converts Symphony data to an email message format and then imports those items to user mailboxes in Microsoft 365.

After storing Symphony data in user mailboxes, you can apply Microsoft Purview features such as Litigation Hold, eDiscovery, retention policies and retention labels, and Communication Compliance. By using a Symphony connector to import and archive data in Microsoft 365, you help your organization stay compliant with government and regulatory policies.

Overview of archiving Symphony data

The following overview explains the process of using a data connector to archive Symphony data in Microsoft 365.



1. Your organization works with 17a-4 to set up and configure the Symphony DataParser.
2. The DataParser regularly collects Symphony items and converts the content of a message to an email message format.
3. You create the Symphony DataParser connector in the Microsoft Purview portal. The connector connects to DataParser and transfers the messages to a secure Azure Storage location in the Microsoft cloud.
4. The connector creates a subfolder in the Inbox folder named **Symphony DataParser** in the user mailboxes and imports the Symphony items to that folder. The connector uses the value of the *Email* property to determine which mailbox to import items to. Every Symphony item contains this property, which is populated with the email address of every participant.

Before you begin

> **Important**
>
> All partner third-party data connector configurations require a provisioned service principal for the partner application. This requirement isn't applicable for native Microsoft third-party data connector configurations.
>
> To provision a service principal using PowerShell, see Provision a service principal for partner third-party data connectors.

- Create a DataParser account for Microsoft connectors. To create an account, contact 17a-4 LLC. You need to sign in to this account when you set up the Symphony DataParser connector.
- Assign the Data Connector Admin role to the user who sets up the Symphony DataParser connector and completes the user-mapping configuration. This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions, see Permissions in the Microsoft Purview portal and Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview.

- This data connector is available in GCC environments in the Microsoft 365 US Government cloud. Third-party applications and services might involve storing, transmitting, and processing your organization's customer data on third-party systems that are outside of the Microsoft 365 infrastructure and therefore aren't covered by the Microsoft Purview and data protection commitments. Microsoft makes no representation that use of this product to connect to third-party applications implies that those third-party applications are FedRAMP compliant.

Step 1: Set up a Symphony DataParser connector

Create a 17a-4 connector for Symphony data in the Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. From the list, select **Symphony DataParser**.
5. On the **Terms of service** page, select **Accept**.
6. Enter a unique name that identifies the connector and then select **Next**.
7. Sign in to your 17a-4 account and complete the steps in the Symphony DataParser connection workflow.

Step 2: Configure the Symphony DataParser connector

Work with 17a-4 Support to configure the Symphony DataParser connector.

Step 3: Map Symphony users to Microsoft 365 users

The Symphony DataParser connector automatically maps users to their Microsoft 365 email addresses before importing data to Microsoft 365.

Step 4: Monitor the Symphony DataParser connector

After you create a Symphony DataParser connector, you can view the connector status in the Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select the Symphony DataParser connector that you created to display the flyout page, which contains the properties and information about the connector.

4. Under **Connector status with source**, select the **Download log** link to open (or save) the status log for the connector. This log contains information about the data that's been imported to the Microsoft cloud. For more information, see View admin logs for data connectors.

Connector limits for archiving Symphony data

Importing attachments or items supports files up to 10 MB.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/24/2026

Set up a connector to archive Zoom data

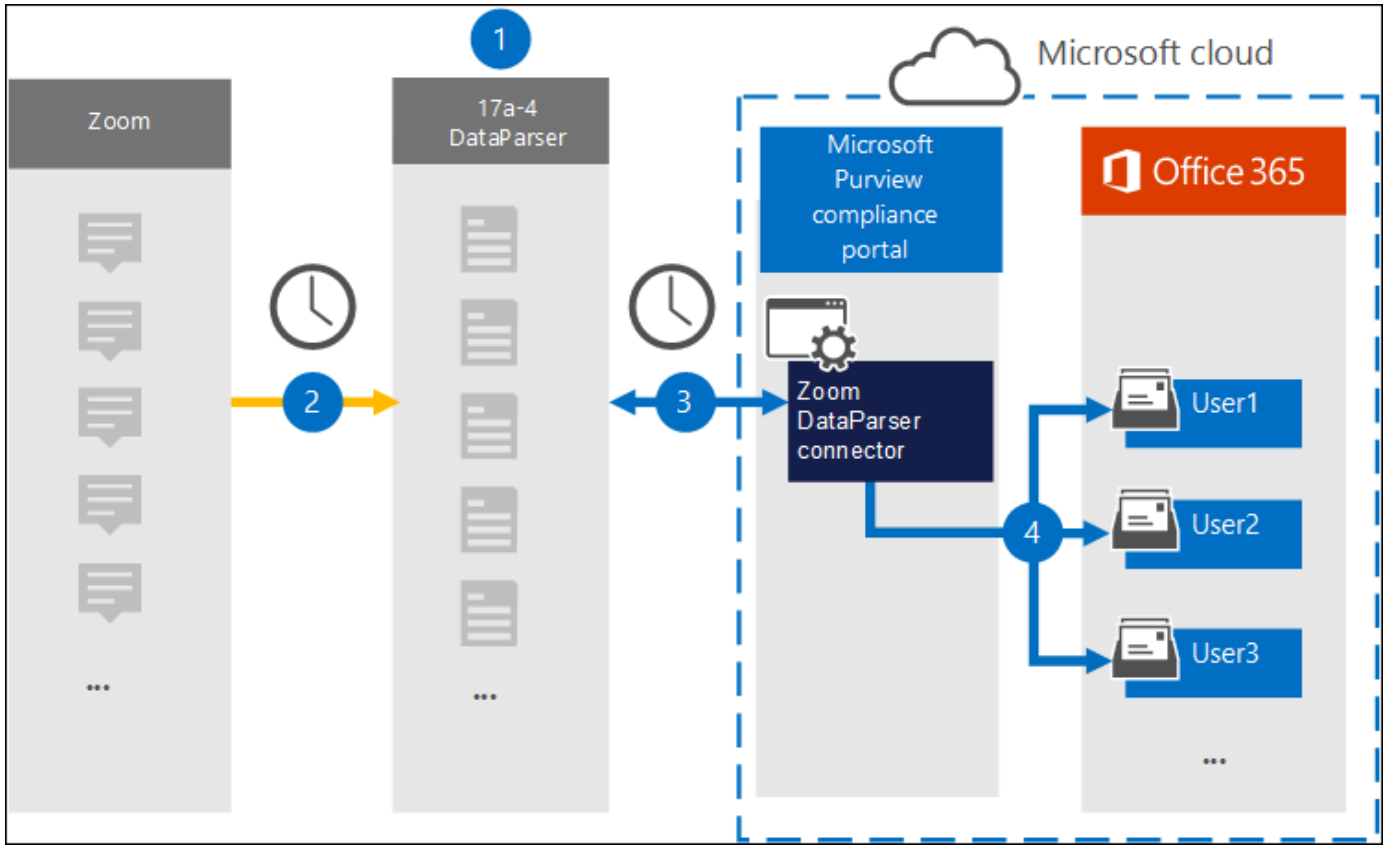
Archive Zoom data with the 17a-4 Zoom DataParser connector

Use the Zoom DataParser from 17a-4 LLC to import and archive data from the Zoom platform to user mailboxes in your Microsoft 365 organization. The DataParser includes a Zoom connector that captures items from a third-party data source. It then converts Zoom data to an email message format and imports those items to user mailboxes in Microsoft 365.

After Zoom data is stored in user mailboxes, you can apply Microsoft Purview features. These features include Litigation Hold, eDiscovery, retention policies, retention labels, and Communication Compliance. Using a Zoom connector to import and archive data in Microsoft 365 helps your organization stay compliant with government and regulatory policies.

Overview of archiving Zoom data

The following overview explains the process of using a data connector to archive Zoom data in Microsoft 365.



1. Your organization works with 17a-4 to set up and configure the Zoom DataParser.
2. The DataParser regularly collects Zoom items and converts the content of a message to an email message format.
3. You create the Zoom DataParser connector in the Microsoft Purview portal. The connector connects to DataParser and transfers the messages to a secure Azure Storage location in the Microsoft cloud.
4. The connector creates a subfolder in the Inbox folder named **Zoom DataParser** in the user mailboxes and imports the Zoom items to that folder. The connector uses the value of the *Email* property to determine which mailbox to import items to. Every Zoom item contains this property, which is populated with the email address of every participant.

Before you begin

> **Important**
>
> All partner third-party data connector configurations require a provisioned service principal for the partner application. This requirement isn't applicable for native Microsoft third-party data connector configurations.
>
> To provision a service principal using PowerShell, see Provision a service principal for partner third-party data connectors.

- Create a DataParser account for Microsoft connectors. To do this, contact 17a-4 LLC. You need to sign in to this account when you create the connector in Step 1: Set up a Zoom DataParser connector.
- Assign the Data Connector Admin role to the user who creates the Zoom DataParser connector in Step 1 (and completes it in Step 3). This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions, see Permissions in the Microsoft Purview portal and Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview.
- This data connector is available in GCC environments in the Microsoft 365 US Government cloud. Third-party applications and services might involve storing, transmitting, and

processing your organization's customer data on third-party systems that are outside of the Microsoft 365 infrastructure and therefore aren't covered by the Microsoft Purview and data protection commitments. Microsoft makes no representation that use of this product to connect to third-party applications implies that those third-party applications are FedRAMP compliant.

Step 1: Set up a Zoom DataParser connector

Create a 17a-4 connector for Zoom data in the Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. Select **Zoom DataParser** from the list.
5. Select **Accept** on the **Terms of service** page.
6. Enter a unique name that identifies the connector, then select **Next**.
7. Sign in to your 17a-4 account and complete the steps in the Zoom DataParser connection workflow.

Step 2: Configure the Zoom DataParser connector

Work with 17a-4 Support to configure the Zoom DataParser connector.

Step 3: Map Zoom users to Microsoft 365 users

The Zoom DataParser connector automatically maps users to their Microsoft 365 email addresses before importing data to Microsoft 365.

Step 4: Monitor the Zoom DataParser connector

After you create a Zoom DataParser connector, you can view the connector status in the Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select the Zoom DataParser connector that you created to display the flyout page, which contains the properties and information about the connector.
4. Under **Connector status with source**, select the **Download log** link to open (or save) the status log for the connector. This log contains information about the data that's imported to

the Microsoft cloud. For more information, see View admin logs for data connectors.

Zoom data connector limits

Zoom items are the messages and related content imported by the connector. Importing attachments or items supports files up to 10 MB.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 06/24/2026

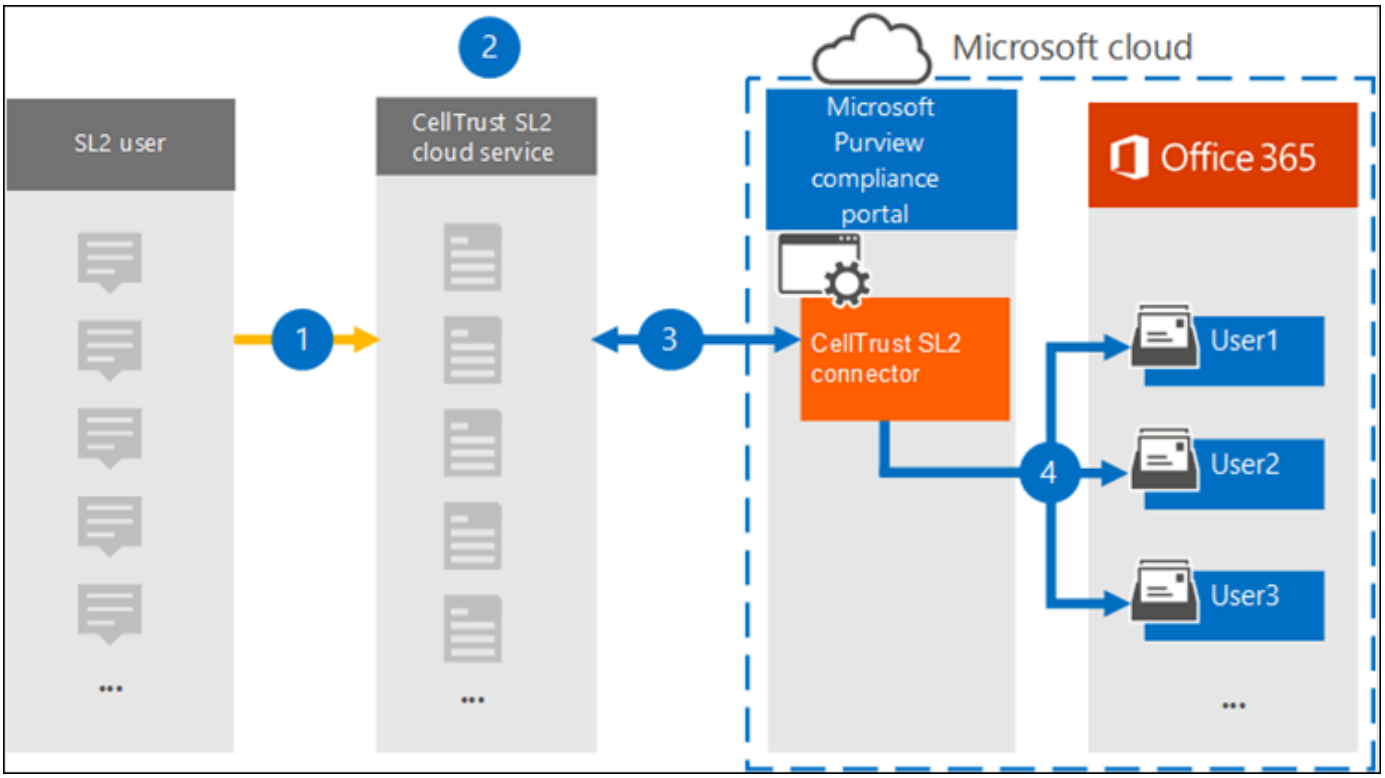
Archive data from CellTrust SL2 to Microsoft 365

CellTrust SL2 captures mobile communications data and integrates with the leading archiving technologies to meet the electronic discovery requirements for regulations such as FINRA, HIPAA, FOIA, and TCPA. The SL2 Data Connector imports mobile communication items to Microsoft 365. This article describes the process for integrating SL2 with Microsoft 365 by using the CellTrust SL2 Data Connector for archiving. To complete this process, you need a subscription to the CellTrust SL2 service and familiarity with the SL2 architecture. For information about CellTrust SL2, see https://www.celltrust.com.

After you import data to user mailboxes in Microsoft 365, you can apply Microsoft Purview features such as Litigation Hold, eDiscovery, Microsoft 365 retention policies, and Communication Compliance. By using the CellTrust SL2 Data Connector to import and archive data in Microsoft 365, your organization can stay compliant with government and regulatory policies.

Overview of archiving with the CellTrust SL2 Data Connector

CellTrust's SL2 platform captures communication data from multiple sources. SL2 data sources are either Person-to-Person (P2P) or Application-to-Person (A2P). The process described in this article pertains only to P2P data sources. For all P2P data sources, at least one party in the collaboration is an SL2 user who subscribes to the SL2 service. The following overview explains the process of using the CellTrust SL2 Data Connector in Microsoft 365.



1. SL2 users send and receive data to and from SL2 services in Microsoft Azure.

2. Your organization has an SL2 domain in CellTrust's SL2 Cloud Service environment. Your domain can have one or more organizational units (OUs). The SL2 Cloud Service transfers your data to a highly secure area in the Microsoft Azure platform, so that your data never leaves the Microsoft Azure environment. Depending on your SL2 plan (Enterprise, SMB, or Government), Microsoft Azure Global or Microsoft Azure Government hosts your domain.

3. After you create the CellTrust SL2 Data Connector in the Microsoft Purview portal, your domain and OUs (regardless of your SL2 plan) begin sending data to Microsoft 365. The data feed is structured to support reporting based on data sources, OUs, or the domain by itself. As a result, your organization needs only one connector to feed all your data sources to Microsoft 365.

4. The connector creates a folder under each mapped user with an appropriate Microsoft 365 license titled **CellTrust SL2**. This mapping connects a CellTrust SL2 user to an Microsoft 365 mailbox by using an email address. If a user ID in CellTrust SL2 has no match in Microsoft 365, the user's data isn't archived.

Before you begin

> **Important**

All partner third-party data connector configurations require a provisioned **service principal** for the partner application. This requirement isn't applicable for native Microsoft third-party data connector configurations.

To provision a service principal using PowerShell, see **Provision a service principal for partner third-party data connectors**.

- Verify that you have a domain in the CellTrust SL2 cloud service environment. For more information about obtaining a production or trial SL2 domain, Contact CellTrust.
- Get the credentials to access the administrator account for your SL2 domain.
- Assign the Data Connector Admin role to the user who creates the CellTrust SL2 data connector in the Microsoft Purview portal and completes the connector setup, including user mapping. This role is required to add connectors on the **Data connectors** page in the Microsoft Purview portal. Multiple role groups include this role by default. For a list of these role groups, see Roles in Microsoft Defender for Office 365 and Microsoft Purview. Alternatively, an admin in your organization can create a custom role group, assign the Data Connector Admin role, and then add the appropriate users as members. For instructions, see Permissions in the Microsoft Purview portal and Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview.
- This data connector is available in GCC environments in the Microsoft 365 US Government cloud. Third-party applications and services might involve storing, transmitting, and processing your organization's customer data on third-party systems that are outside of the Microsoft 365 infrastructure and therefore aren't covered by the Microsoft Purview and data protection commitments. Microsoft makes no representation that use of this product to connect to third-party applications implies that those third-party applications are FedRAMP compliant.
- A domain administrator can set up a connector for the domain or any OUs in that domain. If you use the OU Administrator account, you can only set up a connector for that specific OU.
- To successfully complete the setup steps, you must be assigned a Microsoft 365 E5 license and have the proper Microsoft Office admin rights.

Step 1: Create a CellTrust SL2 connector

Create a data connector for CellTrust SL2 data in the Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select **My connectors**, then select **Add connector**.
4. Select **CellTrust SL2 (preview)** from the list.
5. Select **Accept** on the **Terms of service** page.
6. Enter a unique name that identifies the connector and select **Next**. The name you enter identifies the connector on the **Data connectors** page after you create it.
7. Select **Sign into CellTrust** on the **Sign in to your CellTrust account** page. You're redirected to the **CellTrust Portal for Microsoft 365** in a new browser window.

Step 2: Select the domains or OUs to archive

To select domains and OUs for archiving, sign in to an administrator account for your CellTrust SL2 domain and choose which domains and OUs to archive in Microsoft 365.

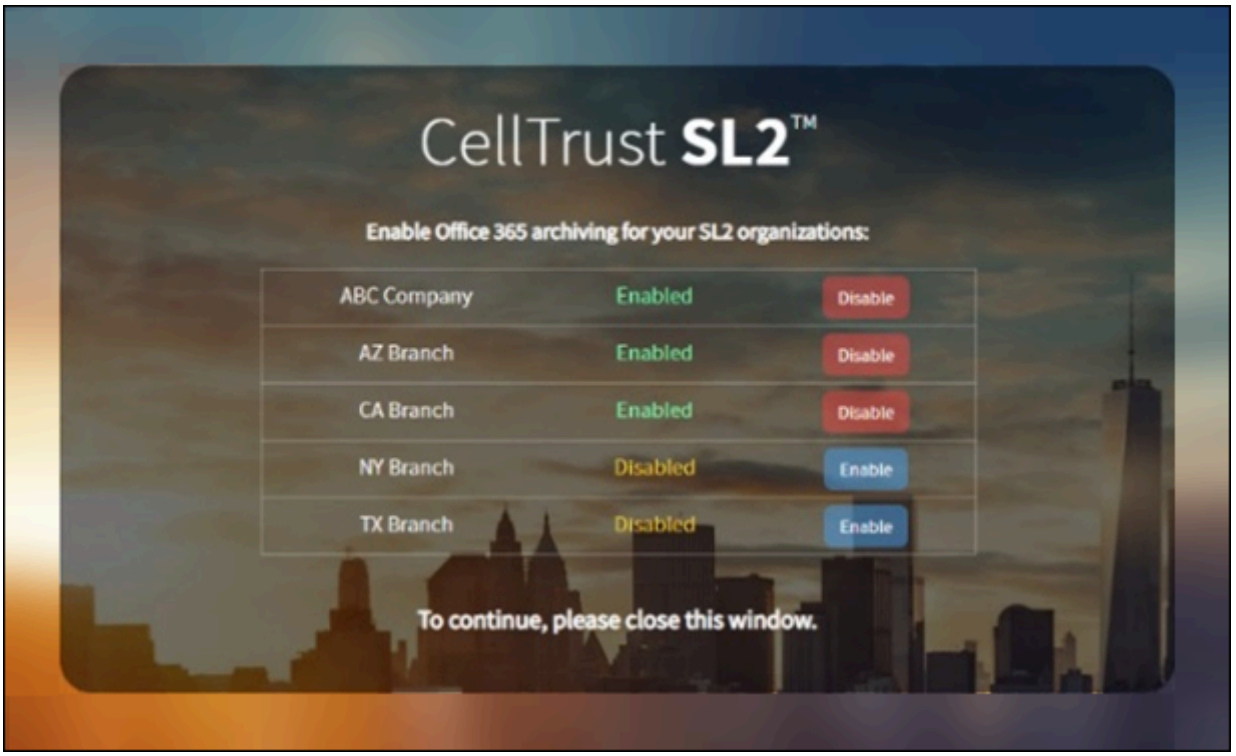
1. On the CellTrust **Microsoft 365 Connector** page, select your environment in the SL2 cloud service to display a sign-in page.

   Typically, you see one option representing your environment. However, if you have domains in more than one environment, you see options for each environment. After you make a selection, you're redirected to the SL2 authentication page.

2. Sign in with your Domain or OU Administrator account credentials.

   If you sign in as an SL2 domain administrator, you see the name of your domain and the OUs in that domain. If you don't have OUs, you only see the name of your domain. If you sign in as OU Administrator, you only see the name of your OU.

3. Enable the business units you want to archive. Selecting the domain doesn't automatically select the OUs. You must enable each OU separately to archive it.



4. When you finish your selections, close the browser window and return to the workflow page in Microsoft Purview portal. After a few seconds, the workflow automatically advances to the **User mapping** page.

Step 3: Map users and complete the connector setup

To complete the connector setup, map users and finalize the configuration in the Microsoft Purview portal.

1. On the **User mapping** page, select **Enable automatic user mapping** if the email address for users is the same in both SL2 and Microsoft 365. Otherwise, manually provide user email addresses by uploading a CSV file that maps users' SL2 address to their Microsoft 365 address.

2. Select **Next**, review your settings, and then select **Finish** to create the connector.

   The new connector is added to the list on the **Data connectors** page.

> **Tip**
>
> To test the new connector, send a text message using your SL2 mobile app or from your SL2 portal. Go to your Microsoft 365 mailbox and open the **CellTrust SL2** folder in your Inbox. It might take a few minutes for the text messages to show up in your mailbox.

Get help from CellTrust

For details about contacting CellTrust for help with setting up a CellTrust SL2 data connector, see the CellTrust Customer Support page.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 06/24/2026

Use Content Search to search third-party data imported by a custom partner connector

> **⊗ Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

You can use the Content search eDiscovery tool in the Microsoft Purview portal to search for items imported to mailboxes in Microsoft 365 from a third-party data source. You can create a query to search all imported third-party data items or you can create a query to search specific third-party data items. Also, you can also create a query-based retention policy or a query-based eDiscovery hold to preserve third-party data.

> **ⓘ Important**
>
> The guidance in this article only applies to third-party data imported by a custom partner connector. This article doesn't apply to third-party data that is imported by using the third-party data connectors in the Microsoft Purview portal.

Creating a query to search all third-party data

To search (or place on hold) any type of third-party data that you've imported to Office 365, you can use the `kind:externaldata` message property-value pair in the keyword box for a Content Search or when creating a query-based hold. For example, to search for items imported from any third-party data source and contain the word "contoso" in the Subject property of the imported item, you would use the following query:

PowerShell

```
kind:externaldata AND subject:contoso
```

When creating queries to search and hold third-party data, you can also use conditions to narrow the search results. For more information about creating Content Search queries, see Keyword queries and search conditions for Content Search.

Creating a query to search specific types of third-party data

Instead of searching all types of third-party data, you can create queries that only search for a specify type of third-party data by using the following message *property: value* pair in the keyword box for a Content Search:

```
PowerShell
itemclass:ipm.externaldata.<third-party data type>*
```

For example, to search Facebook data that contains the word "contoso" in the Subject property, you would use the following query:

```
PowerShell
itemclass:ipm.externaldata.Facebook* AND subject:contoso
```

The following table lists the third-party data types that you can search, and the value to use for the `itemclass:` message property to specifically search for that type of third-party data. The query syntax isn't case-sensitive.

Expand table

Third-party data type	Value for `itemclass:` property
AIM	`ipm.externaldata.AIM*`
American Idol	`ipm.externaldata.AmericanIdol*`
AOL with Pivot Client	`ipm.externaldata.Pivot.IM`
Apple Juice	`ipm.externaldata.AppleJuice*`
Ares	`ipm.externaldata.Ares*`
Axs Encrypted	`ipm.externaldata.AxsEncrypted*`
Axs Exchange	`ipm.externaldata.AxsExchange*`
Axs Local Archive	`ipm.externaldata.AxsLocalArchive*`

Third-party data type	Value for `itemclass:` property
Axs Placeholder	`ipm.externaldata.AxsPlaceHolder*`
Axs Signed	`ipm.externaldata.AxsSigned*`
Bazaarvoice	`ipm.externaldata.Bazaarvoice*`
Bearshare	`ipm.externaldata.Bearshare*`
BitTorrent	`ipm.externaldata.BitTorrent*`
Blackberry	`ipm.externaldata.Blackberry*`
BlackBerry Call Logs	`ipm.externaldata.BlackBerryCall*`
BlackBerry Messenger	`ipm.externaldata.BlackBerryMessenger*`
BlackBerry PIN	`ipm.externaldata.BlackBerryPIN*`
BlackBerry SMS	`ipm.externaldata.BlackBerrySMS*`
Bloomberg	`ipm.externaldata.Bloomberg*`
Bloomberg Message	`ipm.externaldata.conversation.Bloomberg Message*`
Bloomberg Messaging	`ipm.externaldata.BloombergMessaging*`
Box	`ipm.externaldata.Box*`
Cisco IM & Presence Server	`ipm.externaldata.Jabber.IM`
Cisco Jabber	`ipm.externaldata.Jabber*`
CipherCloud for Salesforce Chatter	`ipm.externaldata.Chatter.Post` `ipm.externaldata.Chatter.Comment`
Direct Connect	`ipm.externaldata.DirectConnect*`
Facebook	`ipm.externaldata.Facebook*`
FastTrack	`ipm.externaldata.FastTrack*`
FXConnect	`ipm.externaldata.FXConnect.chat`
Flickr	`ipm.externaldata.Flickr*`
Gnutella	`ipm.externaldata.Gnutella*`
Google+	`ipm.externaldata.GooglePlus*`
Google Talk	`ipm.externaldata.GoogleTalk*`

Third-party data type	Value for `itemclass:` property
GoToMyPC	`ipm.externaldata.GoToMyPC*`
HipChat	`ipm.externaldata.HipChat*`
Hopster	`ipm.externaldata.Hopster*`
HubConnex	`ipm.externaldata.HubConnex*`
IBM Connections	`ipm.externaldata.Connections*`
IBM SameTime	`ipm.externaldata.Sametime*`
ICE Chat	`ipm.externaldata.conversation.Ice Chat*`
Indii Messenger	`ipm.externaldata.Indii*`
Instagram	`ipm.externaldata.Instagram*`
Instant Bloomberg	`ipm.externaldata.InstantBloomberg*`
InvestEdge	`ipm.externaldata.InvestEdge*`
IRC	`ipm.externaldata.IRC*`
Jive	`ipm.externaldata.Jive*`
JiveApiRetention	`ipm.externaldata.JiveApiRetention*`
JXTA	`ipm.externaldata.JXTA*`
LinkedIn	`ipm.externaldata.LinkedIn*`
MFTP	`ipm.externaldata.MFTP*`
Microsoft UC	`ipm.externaldata.MicrosoftUC*`
Mind Align	`ipm.externaldata.MindAlign*`
Mobile Guard	`ipm.externaldata.MobileGuard*`
MSN	`ipm.externaldata.MSN*`
MySpace	`ipm.externaldata.MySpace*`
NEONetwork	`ipm.externaldata.NEONetwork*`
OpenNap	`ipm.externaldata.OpenNap*`
Pinterest	`ipm.externaldata.Pinterest*`
Pivot	`ipm.externaldata.Pivot*`

Third-party data type	Value for `itemclass:` property
QQ	`ipm.externaldata.QQ*`
Microsoft SharePoint	`ipm.externaldata.SharePoint*`
Salesforce Chatter	`ipm.externaldata.Chatter*`
Skype for Business	`ipm.externaldata.Skype*`
Slack Enterprise Grid	`ipm.externaldata.Slack.IM`
SoftEther	`ipm.externaldata.SoftEther*`
Squawker	`ipm.externaldata.Squawker*`
Symphony	`ipm.externaldata.Symphony*`
Thomson Reuters	`ipm.externaldata.Reuters*`
Thomson Reuters Eikon Messenger	`ipm.externaldata.ReutersEikon*`
Tor	`ipm.externaldata.Tor*`
TTT	`ipm.externaldata.TTT*`
Twitter	`ipm.externaldata.Twitter*`
UBS Chat	`ipm.externaldata.UBS*`
Vimeo	`ipm.externaldata.Vimeo*`
WinMX	`ipm.externaldata.WinMX*`
Winny	`ipm.externaldata.Winny*`
Yahoo!	`ipm.externaldata.Yahoo!*`
Viva Engage	`ipm.externaldata.yammer*`
YellowJacket	`ipm.externaldata.YellowJacket*`
YouTube	`ipm.externaldata.YouTube*`

Last updated on 11/18/2025

View admin logs for data connectors

09/26/2025

After you create a data connector to import non-Microsoft data to Microsoft Purview, you can monitor the daily import status of the connector by downloading the admin logs for the data connector.

Before you view admin logs

> **Important**
>
> We highly recommended that you review admin logs daily to ensure any data ingestion failures caused by erroneous data or service issues are detected early and remediated. Without daily monitoring, you could end up with irreversible loss of data.

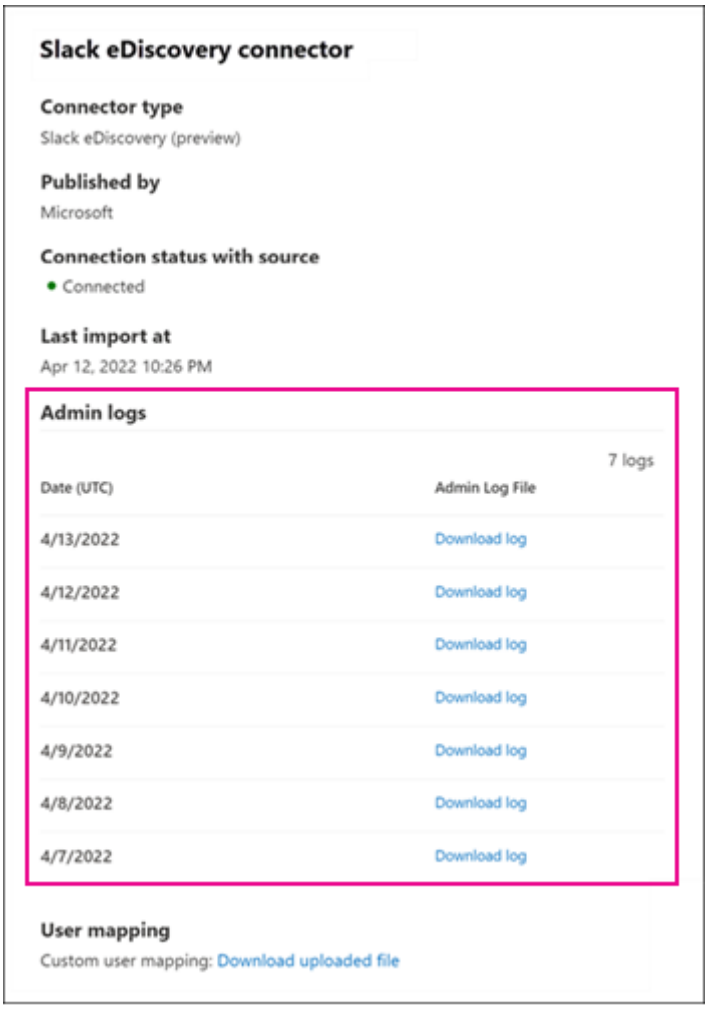
Before you can view admin logs, make sure that you've confirmed the following:

- Auditing must be enabled for your organization to generate and view admin log for your organization. Auditing is enabled by default in Microsoft Purview. However, we recommend that you verify auditing status of your organization. For instructions, see Verify the auditing status for your organization. If you need to enable auditing for your organization, see Turn on auditing.
- After auditing is turned on, it could take up to 48 hours to generate admin logs for data connectors. We recommend that you enable auditing before you create data connectors.
- After a data connector is created, it can take up to 72 hours to generate admin logs that contain import status summary.
- Admin logs are available for the previous seven days.

Download admin logs for data connectors

To download admin logs for date connectors, complete the following steps:

1. Sign into the Microsoft Purview portal.
2. Select **Settings** > **Data connectors**.
3. Select the **My connectors** tab and then select a data connector to display the fly out page, which contains information and properties about the data connector.
4. Under **Admin logs**, select the **Download log** link to open an admin log.



5. View the following import status information in the admin log:

- **Import completion time**: timestamp (in UTC) when connector completes import from data source and all the below events are computed/updated against the import.
- **Items available from source**: count of items downloaded by Connector from data source.
- **Items available for import**: count of items available for import by Connector after fanout. Fanout represents the act of writing a message to all associated participants (sender, receiver etc.).
- **Items imported successfully**: count of items imported successfully by Connector into user mailboxes after fanout.
- **Items partially imported**: count of items imported successfully by Connector into user mailboxes after fanout but had attachments dropped.
- **Items skipped**: count of items skipped from being imported into user mailboxes after fanout due to it being duplicate items.
- **Items failed**: count of items that failed to be imported into user mailboxes after fanout due to it errors (like user mapping, item size exceeded etc.). Event is logged once per user for user mapping failures.

> **Note**
>
> Items available for import should be a sum of items imported successfully, partially imported, skipped, and failed.

- Summary of Failed item details:
  - **Item ID**: unique identifier of the item
  - **Source User ID**: user ID at source application
  - **Microsoft 365 User ID**: User Principal Name in Microsoft 365
  - **Failure Reason**: indicates the reason why the connector couldn't import the item

> **Note**
>
> If no items are imported on a particular day, the admin log contains the following message: `No items ingested on mm/dd/yyyy.`

Reingesting data after failures

Reingesting of data after failures is supported if completed within 90 days. To reingest data, select the **Help** icon in the Microsoft Purview portal to open a support case. Be prepared to provide all relevant logs to help troubleshoot any issues more quickly.

Collection Policies solution overview

Collection policies are an event collection and filtering tool in Microsoft Purview that enables monitoring and classification of events from apps and locations that lay both inside of and beyond your organizations trust boundaries. They let you filter which events from both untrusted and trusted sources are ingested into Purview. Once ingested, that data can be classified and used by various Microsoft Purview signal consuming solutions, such as Microsoft Purview Activity explorer, Microsoft Purview Insider Risk Management, Microsoft Purview eDiscovery, Microsoft Purview Data Lifecycle Management.

Collection policies can help you achieve these data security outcomes:

Only ingest the events that you want

The primary purpose of collection policies is to enable you to define the events that your organization wants to bring into Microsoft Purview, so you can focus on the data that you care about. Here are some examples of how this capability can help you.

Regulatory compliance by geopolitical location

Your organization may span geopolitical and regulatory boundaries with different regions having different data security and privacy requirements. One may require all events be ingested into Purview, while another area may be subject to restrictions on what events can be collected. Collection policies let you define the events that are collected for each region. This allows you to meet the data security and privacy requirements of each region while still having a single consolidated view into your organization's sensitive information.

Filter out unnecessary events

Your organization may have many events that are being collected, but you only want to see the events that are relevant to your needs. For example, say your organization has onboarded all devices and enabled **Always audit file activity for devices** in Endpoint DLP settings. This means that all file activity is being collected and placed into activity explorer. Sifting through all events may slow event investigation as you sort out the unwanted data to get to the data you need. With collection policies, you can filter out the events that you don't want to see before they get to activity explorer. This can help reduce noise and make it easier to find the events that are relevant to your organization.

How collection policies work

The key to understanding collection policies is to know what an event is.

An *event* is made up of two things:

- A *condition*, like **Content contains**, or **File extension is**. This requires the use of classifiers, like sensitive information types. For example, a condition could be **Content contains** a sensitive information type, like **U.S. Social Security Number** or **File extension is** > **.docx**.
- An *activity* that a user can take on a sensitive item, like **File copied to removable media**

Collection policies filter events based on:

- A match to one or more conditions and activities defined in the event.
- A match to the *data source* or location where the event is happening. For example, a collection policy can be scoped to only collect events from devices, or only collect events from Copilot in Fabric.

A policy is considered matched when one or more of the conditions are met (unless you AND the conditions) and the activity is matched on the data source.

Multiple collection policies for single data source

It's possible your organization will need to create multiple collection policies for any given data source. Collection policies handle the multiple policy scenario differently than DLP or other Purview policies, which assign a priority to policy evaluation based on the order in which the policies are created.

On the collection polices page, you see all the individual policies that have been created. But on the back end, during policy evaluation, Microsoft Purview combines all the collection policies for a data source into a single collection policy for that data source. So all the conditions for a data source are used when identifying matches.

Accessing collection policies

Collection policies are a component of the Microsoft Purview portal **Classifiers** feature. You can access them from anywhere in the Purview portal where the **Classifiers** feature is available. For example,

- In the Microsoft Purview portal, select **Solutions** > **Data Loss Prevention** > **Classifiers** > **Collection policies**.

- Alternately, you can access collection policies in the Microsoft Purview portal from **Solutions** > **Information Protection** > **Classifiers** > **Collection policies**.

Collection policy configuration overview

You have flexibility in how you create and configure your collection policies. All collection policies require the same information from you. You must know this information *before* you start creating a collection policy. For additional details about each step, including supported configuration, see Collection policies policy reference.

1. Define **data to detect** by configuring conditions required for a policy match. There are four supported conditions:

    - **Content contains classifiers**, which makes use of sensitive information types and trainable classifiers. Can be scoped to all classifiers, all classifiers except selected ones, or only specific classifiers.

      > ⓘ Note
      >
      > The devices location doesn't support using trainable classifiers.

    - **Document size equals or is greater than**, as measured in bytes, kilobytes (KB), megabytes (MB), gigabytes (GB), or terabytes (TB).
    - **Document is equal to or smaller than**, as measured in bytes, kilobytes (KB), megabytes (MB), gigabytes (GB), or terabytes (TB).
    - **File extension is**, which makes use of file extensions.

2. Choose the **activities** required for a policy match - The supported activities are specific for the data source. For devices, there are 19 supported activities, such as **File copied to removable media**, **File printed**, and **File uploaded to cloud**. A full list of the supported activities is available in the Collection policies policy reference.

3. Choose **where to apply** the policy by selecting a data source.

4. Depending on the policy settings you choose, you may have additional settings to configure, such as whether or not to enable content capture for AI interactions and the cloud apps detection method.

Content capture for AI interactions

To help comply with regulatory requirements, you can decide whether to capture and store all detected prompts and responses from any generative AI data sources added to the policy. This makes it easy to discover and protect the captured content later with additional Microsoft Purview policies and solutions. This capability doesn't include content in files shared with generative AI, and only applies to:

- Copilot experiences
- Enterprise AI
- Unmanaged cloud apps categorized as generative AI
- All unmanaged AI apps adaptive app scope

Without the setting enabled, content detected in prompts and responses are limited to sensitive information only.

> ⓘ **Note**
>
> To capture AI content, you must have the **Content contains classifiers** condition set to **All**.

Collection Policies and Microsoft Purview Insider Risk Management

Depending on your configuration, Microsoft Purview Insider Risk Management policy behavior might be impacted by collection policies. When an insider risk policy is configured, it monitors service and third-party indicators to help you quickly identify, triage, and act on risk activity. Using logs from Microsoft 365 and Microsoft Graph, Insider Risk Management allows you to define specific policies to identify risk indicators. These policies allow you to identify risky activities and to act to mitigate these risks.

> ⚠ **Warning**
>
> When a device collection policy is configured and deployed, if there's a conflict between what the Insider Risk Management policy is configured to monitor and what the collection policy is configured to filter for, the collection policy configuration takes precedence. This means that if an Insider Risk Management policy is configured to monitor a specific activity, but the collection policy is configured to filter out that activity, the activity isn't collected and isn't available for review in Insider Risk Management. This only applies to device indicators in Insider Risk Management policies, and only when a device collection policy is deployed. If a collection policy is configured to collect something that an Insider Risk Management policy isn't configured for, then the indicator is collected, but Insider Risk Management ignores it.

Collection policies and Endpoint Data Loss Prevention

When **Always audit file activity for devices** is enabled all activities for Office, PDF, and CSV files are collected by default. If you want to modify which activities are collected for devices in this case, you can do so by configuring a collection policy targeted to devices. If **Always audit file activity for devices** isn't enabled, activities aren't collected from devices, even if a collection policy is created scoped to devices.

Collection policies can't impact the behavior of DLP policies, only what is collected and recorded for audit file events on devices.

Collection policies and Microsoft Purview Data Security Posture Management and Purview Data Security Posture Management for AI

When you use one-click policies in Microsoft Data Security Posture Management or Microsoft Data Security Posture Management for AI (DSPM for AI), you can take action on recommendations that automatically create collection policies on your behalf.

You can view and monitor the outcomes of these policies in tailored experiences within DSPM and DSPM for AI, as well as in activity explorer. These policies are named automatically by the quick policy experience. Post-creation, these collection policies can be viewed and edited in the same manner as collection policies created using the workflow.

Last updated on 11/12/2025

Create and deploy collection policies

There are many configuration options in a Microsoft Purview collection policy. Each option changes the policy's behavior. This article presents some common intent scenarios for policies that map to configuration options. Then it walks you through configuring those options. Once you familiarize yourself with the scenarios in this article, you'll have the foundational skills that you need to use the collection policy creation UX to create your own policies.

Before you begin

If you're new to Microsoft Purview collection policies, here's a list of the core articles you should be familiar before.

- Collection policies solution overview.
- Collection policies policy reference

SKU/subscriptions licensing

For information on licensing, see

- Microsoft 365 Enterprise Plans
- Microsoft 365 Service Descriptions

Learn about Microsoft Purview billing models

Billing for collection policies

This feature uses pay-as-you-go billing or per-user licensing for Microsoft Purview capabilities or both. To help you understand and manage your usage, Microsoft Purview provides a **Usage center** in the Microsoft Purview portal. For more information, see Manage pay-as-you-go and per-user licensing usage. To learn more about Purview billing, see Purview billing models..

Granular Roles and Role Groups

There are roles and role groups that you can use to fine tune your access controls.

Here's a list of applicable roles.

- DLP Compliance Management

To learn more, see Permissions in the Microsoft Purview portal.

Here's a list of applicable role groups.

- Organization Management
- Compliance Administrator
- Insider Risk Management Admin
- Information Protection Admins
- Data Security Viewers (read-only support)

Policy creation scenarios

There are so many configuration options in the policy creation flow that it's not possible to cover every, or even most configurations. So this article covers the most common collection policy scenarios. Going through the following scenarios gives you hands-on experience across a broad range of configurations.

- Scenario 1 Detect sensitive data shared with unmanaged cloud apps via network

Scenario 1 Detect sensitive data shared with unmanaged cloud apps via network

This scenario detects sensitive data shared with unmanaged cloud apps at the network level. Visibility into the network traffic is provided by an integrated Secure Access Service Edge (SASE) or non-Microsoft secure browser provider, so it requires the integration to be configured and enabled. The steps to configure and enable the integration are covered in this scenario.

> **Note**
>
> This is a hypothetical scenario with hypothetical values. It's only for illustrative purposes. You should substitute your own sensitive information types, sensitivity labels, distribution groups and users.

Scenario 1 prerequisites and assumptions

- You either have an existing relationship with one or more supported SASE or non-Microsoft secure browser providers or are interested in procuring one as part of your deployment. See the Security Store in Microsoft Purview Data Loss Prevention to understand which providers are supported.

- If you choose a non-Microsoft SASE or secure browser integration, you have configured Microsoft Purview pay-as-you-go billing for your Microsoft 365 tenant. Pay-as-you-go is not required for integration with Entra Global Secure Access. For more information, see Learn about Microsoft Purview Network Data Security.

Configuring network data security is a multi-phase process.

1. Setup Purview integration with one or more SASE providers.
2. Configure a collection policy to detect sensitive data shared with unmanaged cloud apps at the network level.
3. Viewing the network data security data

Scenario 1 policy intent statement and mapping

We need to detect data that is sensitive to our organization in text or files shared by any user in the "US Sales" group through browsers, applications, add-ins, APIs, and more, with unmanaged cloud apps, such as ChatGPT, Dropbox, Slack, or Gmail.

Expand table

Statement	Configuration question answered and configuration mapping
"We need to detect data that is sensitive to our organization"	- Data to detect: **All classifiers** excluding **All Full Names** and **All Physical Addresses**
"...in text or files..."	- Activities to detect: **Text sent to or shared with cloud or AI app** and **File uploaded to or shared with cloud or AI app**
"...shared by any user in the "US Sales" group through browsers, applications, add-ins, APIs, and more, with unmanaged cloud apps, such as ChatGPT, Dropbox, Slack, or Gmail..."	- Where to apply: **Unmanaged cloud apps: ChatGPT, Dropbox, Slack, Gmail** - Scope: **Include "US Sales" group**

Steps to create policy for scenario 1

SASE provider integration

Perform the following steps to integrate a supported SASE or secure browser provider with Microsoft Purview:

1. In the Microsoft Purview portal, open the **Data Loss Prevention** solution and choose **Security Store** in the left navigation bar.

2. Select **Get solution** for the SASE or non-Microsoft secure browser solution provider you want to integrate with.
3. Complete the steps required by the integration. The steps vary depending on the SASE provider you choose. For Entra Global Secure Access integration, see Create content policies for network content filtering.

Configure collection policy for network data security

Perform the following steps to create the collection policy for network data security:

1. Sign in to the Microsoft Purview portal.
2. Open the **Data loss prevention** solution and navigate to **Classifiers** > **Collection policies**.
3. Select + **Create policy**.
4. Give the policy a **Name** and an optional **Description**. You can use the policy intent statement here.
5. Choose **Next**.
6. Select + **Add condition** and choose **Content contains**.
7. Select **Classifiers** and then **Edit**.
8. In the **Scope for classifiers** flyout, check **Exclude classifiers** and then select + **Exclude classifiers**.
9. In the **Choose Classifiers** flyout, select **All Full Names** and **All Physical Addresses** then select **Done**.
10. Select **Done** then select **Next**.

> **Tip**
>
> You can also add **Sensitivity labels** as a condition under **Content contains** to scope detection to items with specific sensitivity labels applied. For example, you can focus detection on items labeled as "Internal" or "Confidential" while excluding publicly labeled content. Sensitivity labels are only supported with browser and network cloud apps detection.

11. Select + **Add activities**.

12. Select **Text sent to or shared with cloud or AI app** and **File uploaded to or shared with cloud or AI app**.
13. Select **Add** then **Next**.
14. Select **+ Add data source**.
15. Select the **Unmanaged cloud apps** tab.
16. Search for and select **ChatGPT**, **Dropbox**, **Slack**, and **Gmail** and select **Add**.
17. Select **Edit scope** on the **Unmanaged cloud apps** row.
18. Select the **Include only specific** option, and on the **Included** tab select **+ Add inclusions**.
19. Search for the **US Sales** group, select it, and select **Add**.
20. Select **Save and close** then select **Next**.
21. Leave the default option of **Don't capture content** and select **Next**.
22. Ensure **Network and non-Microsoft secure browsers** is enabled and select **Next**.
23. Ensure **Turn on** is enabled and select **Next**.
24. Review the policy settings and select **Create policy**.
25. Review the **Next steps** and select **Done**.

Viewing the network data security data

You can view network sensitive data matches in two Purview tools:

- Activity explorer in Data loss prevention
- Data Security Posture Management (DSPM).

Activity explorer

Use the following steps to view network data security matches in Activity explorer:

1. Open the **Data loss prevention** solution and navigate to **Explorers** > **Activity explorer**.
2. Select **+ Add filter** and choose **Enforcement plane**.
3. Click the **Enforcement plane** filter and select **Network**.
4. Select **Apply**.

DSPM

Use the following steps to review collection policy results in DSPM for AI:

1. Open the DSPM solution.
2. Navigate to **Reports** to see high level reports and graphs breaking down the interactions detected by the collection policy.
3. Navigate to **Activity explorer** to see individual AI interactions and sensitive information detection events.

> **Note**
>
> Unless you've enabled content capture for AI interactions in the collection policy, only sensitive content matches will be available to view.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 09/16/2026

Collection policy reference

Microsoft Purview collection policies have many components to configure. To create an effective policy, you need to understand the purpose of each component and how its configuration alters the behavior of the policy. This article provides a detailed anatomy of a collection policy.

Before you begin

If you're new to collection policies, here's a list of the core articles you need as you implement them in your organization:

1. Collection Policies solution overview
2. Collection policy reference - this article that you're reading now introduces all the components of a DLP policy and how each one influences the behavior of a policy
3. Create and Deploy collection policies.

Conditions

Specify conditions to define what **data to detect**. Conditions are optional, but some might be required for other settings. If you don't add conditions, what gets detected depends on the data sources you select later:

- **Devices**: All data is detected, even if it doesn't match your organization's classifiers.
- **All other data sources**: Only data that matches your organization's classifiers is detected.

Collection policies support five conditions:

Expand table

Condition	More information
Content contains classifiers	Sensitive information types and trainable classifiers to detect. Can be scoped to all classifiers, all classifiers except selected ones, or specific classifiers. The devices data source *doesn't support trainable classifiers*. If you select one, it's ignored.
Content contains sensitivity labels	Sensitivity labels to scope detection to items with specific sensitivity labels applied. Only supported with browser and network cloud apps detection.
Document size equals or is greater than	Detect files with a size that is greater than a specified number of bytes, kilobytes (KB), megabytes (MB), gigabytes (GB), or terabytes (TB).

Condition	More information
Document is equal to or smaller than	Detect files with a size that is smaller than a specified number of bytes, kilobytes (KB), megabytes (MB), gigabytes (GB), or terabytes (TB).
File extension is	Detect files with specified file extensions.

Activities

Choose which **activities to detect**. Supported activities are specific to the data sources you want to include.

> **Tip**
>
> You can mix activities that support different data sources in a single policy, but you must add all applicable data sources to the policy to support the selected activities.

Expand table

Activity	Description	Data source
Text sent to or shared with cloud or AI app	When raw text is uploaded to a cloud app, including generative AI prompts, form submissions, and messages	- Cloud apps - Generative AI
File uploaded to or shared with cloud or AI app	When a binary file is uploaded to a cloud app or generative AI services	- Cloud apps - Generative AI
Text received from cloud or AI app	When raw text is downloaded from a cloud app, including generative AI responses	- Cloud apps - Generative AI
File downloaded from cloud or AI app	When a binary file is downloaded from a cloud app or generative AI service	- Cloud apps - Generative AI
Archive created	When an archive file is created on an onboarded endpoint device	Devices
File accessed by unallowed app	When a file is accessed by a restricted app or app group on an onboarded endpoint device	Devices

Activity	Description	Data source
File archived	When a file is added to an archive on an onboarded endpoint device	Devices
File copied to network share	When a file is copied to a network share on an onboarded endpoint device	Devices
File copied to remote desktop session	When a file is copied to a remote computer through a remote desktop session on an onboarded endpoint device	Devices
File copied to removable media	When a file is copied to a removable media, such as a USB flash drive, on an onboarded endpoint device	Devices
File created	When a file is created on an onboarded endpoint device	Devices
File created on network share	When a file is created on a network share from an onboarded endpoint device	Devices
File created on removable media	When a file is created on removable media, such as a USB flash drive, from an onboarded endpoint device	Devices
File deleted	When a file is deleted from an onboarded endpoint device	Devices
File modified	When a file is modified from an onboarded endpoint device	Devices
File printed	When a file is printed from an onboarded endpoint device	Devices
File read	When a file is read from an onboarded endpoint device	Devices
File renamed	When a file is renamed from an onboarded endpoint device	Devices
File transferred by Bluetooth	When a file is transferred by Bluetooth from an onboarded endpoint device	Devices
File uploaded to cloud	When a file is uploaded to the cloud from an onboarded endpoint device	Devices
Removable media mount	When removable media, such as a USB flash drive, is mounted on an onboarded endpoint device	Devices
Removable media unmount	When removable media, such as a USB flash drive, is unmounted on an onboarded endpoint device	Devices

Data sources

Data sources define **where to apply** the policy, and directly correlate to the activities added to the policy.

The following data sources are supported:

Expand table

Data source	More information	Supported activities
Devices	Devices onboarded to Microsoft 365 and managed by your org.	Windows devices onboarded into Microsoft 365.
Copilot experiences	Includes Copilot in Microsoft Fabric and Microsoft Security Copilot only, with support for more experiences coming soon.	- Text sent to or shared with cloud or AI app - Text received from cloud or AI app
Enterprise AI	Non-Copilot AI apps that are onboarded or connected to your org by using methods like Microsoft Entra registration, Microsoft Foundry, or Microsoft Purview Data Map connectors. Policies can be applied to all Enterprise AI apps or scoped to specific apps.	- Text sent to or shared with cloud or AI app - Text received from cloud or AI app
Unmanaged cloud apps	Cloud apps sourced in the Defender for Cloud Apps catalog that aren't set up for single sign-on (SSO), allowing users to access personal data through a browser, app, add-in, or API. Policies only detect data while its being shared or transferred (data in motion) via browser and network detection.	**Browser & Network**: - Text sent to or shared with cloud or AI app - File uploaded to or shared with cloud or AI app **Network only**: - Text received from cloud or AI app -File downloaded from cloud or AI app
Adaptive app scopes	Groups of apps, whose membership is determined based on app metadata, such as category. Currently only "All unmanaged AI apps" - unmanaged cloud apps categorized as generative AI - is supported via browser and network detection.	**Browser & Network**: - Text sent to or shared with cloud or AI app - File uploaded to or shared with cloud or AI app **Network only**:

Data source	More information	Supported activities
		- Text received from cloud or AI app -File downloaded from cloud or AI app

> **Note**
>
> Some unmanaged AI apps aren't supported in Edge for Business. Adaptive app scopes apply only to the supported unmanaged apps in Edge for Business. To learn more, see Learn more about which apps are supported.

Scoping data sources to users and groups

For each data source, you can choose to scope by:

- **All** users and groups (default)
- **Specific** users and groups
- **All except** specific users and groups

> **Note**
>
> Excluded users and groups take precedence over any included users or groups.

Other collection policy settings

Depending on the conditions, activities, and data sources you specify, you might need to configure other collection policy settings. When these settings are disabled or grayed out, the policy configuration isn't compatible with the setting.

Content capture for AI interactions

To help comply with regulatory requirements, you can decide whether to capture and store all detected prompts and responses from any generative AI data sources you add to the policy. This feature is called *content capture*. By using this feature, you can easily discover and protect the captured content later by using other Microsoft Purview policies and solutions. This capability

doesn't include content in files shared with generative AI, and only applies to the following data sources:

- Copilot experiences
- Enterprise AI
- Unmanaged cloud apps categorized as generative AI
- All unmanaged AI apps adaptive app scope

Without this setting enabled, content detected in prompts and responses is limited to sensitive information only.

> **Note**
>
> To capture AI content, you must set the **Content contains classifiers** condition to `All`.

Cloud apps detection

If you add unmanaged cloud apps or adaptive app scopes data sources to the policy, including generative AI, you must choose how to detect this data. You can choose from the following options:

- **Browser** - Detect sensitive data shared with unmanaged cloud apps through the Microsoft Edge browser when on Intune-managed work devices. See supported unmanaged apps that can be targeted in policies, see supported activities for what user activities can be captured, and see supported browsers to confirm your version of the Microsoft Edge browser supports browser detection.
- **Network** - Detect sensitive data shared with unmanaged cloud apps through browsers, apps, APIs, and more, by using an integrated Secure Service Edge (SSE) provider and Microsoft Purview network data security.

> **Important**
>
> Microsoft Purview browser and network data security policies don't apply to B2B guest users.

Considerations for unmanaged cloud apps policies

Review this information when you configure policies that include unmanaged cloud apps, or when you troubleshoot unexpected policy behavior:

- When multiple catalog entries exist for the same app with differences in the entry name (for example, QwenAI and Qwen Chat), include all entries for the app to avoid unintended coverage gaps.
- Some unmanaged AI apps like Runway and Meta AI might intermittently send content in encoded form to dynamically generated endpoints, which can impact policy enforcement.
- Policies that target unmanaged apps can capture interactions from both the consumer and enterprise versions of an app when the app's URL is shared across instances (for example, ChatGPT consumer and ChatGPT enterprise).
- When you target an unmanaged cloud app in an Edge for Business browser policy, detection is based on the destination app's traffic, which isn't always exactly the same as the originating app.
- Inline evaluation size limits: Microsoft Purview evaluates up to 4 MB of content for uploadText and downloadText, and files up to 3 MB for uploadFile and downloadFile.

> **Note**
>
> The unmanaged cloud app features only apply to the consumer version of Microsoft 365 Copilot. For more information on the enterprise version and available Purview features see, Learn more about Microsoft 365 Copilot Enterprise protections.

Next steps

After creating a collection policy, you might need to take additional steps depending on the configured settings.

- If you enable **Browser** detection, the Microsoft Edge management service automatically creates the required configuration policies to activate collection policies in Edge for Business. To learn more, see Activate your Microsoft Purview policy in Microsoft Edge.
- If you enable **Network** detection, you must add and configure one or more Secure Access Service Edge (SASE) or non-Microsoft secure enterprise browser integrations to begin detecting network traffic. To learn more, see SASE provider integrations.

> **Important**

If the **automatic behaviors** fail to sync for browser detection, Microsoft Purview shows an error message and policies aren't applied in Edge for Business. An admin with the **required permissions** must resync to resolve the error. For more information, see **activate your Microsoft Purview policy in Microsoft Edge**.

Pay-as-you-go features

Some collection policy data sources and features are pay-as-you-go and require an Azure subscription to be linked before creating a policy. Learn more about Microsoft Purview features that use pay-as-you-go-billing here..

- Copilot experiences
- Enterprise AI
- Unmanaged cloud app activity detected through non-Microsoft SASE and secure enterprise browser integrations via Microsoft Purview network data security
  - Microsoft Entra Global Secure Access integration with Purview network data security is licensed per seat and does not require pay-as-you-go. For more information, see Learn about Microsoft Purview Network Data Security.
- Unmanaged cloud app activity detected through Microsoft Purview policies using inline data protections for the browser.

Privacy notice for Enterprise AI and Network Data Security

Enterprise AI data sources and network data security integrations might require integration with a third-party app or provider. If you choose to enable any third-party integration, it gains access to and might store some policy configuration, including user identifiers. In this case, the third party's terms, conditions, and privacy policy govern the usage and storage of this data.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 09/15/2026

Monitor device health with the device health reports dashboard

The device health reporting dashboard in Microsoft Purview gives compliance and security administrators a centralized view of device readiness for Endpoint Data Loss Prevention (Endpoint DLP). Use the dashboard to monitor device onboarding status, policy update readiness, and feature readiness for Endpoint DLP.

The dashboard helps you find devices that might not be ready to receive or enforce Endpoint DLP policies because of configuration, connectivity, or version issues. It consolidates device health insights, reducing manual troubleshooting across multiple pages.

Before you begin

To access the device health reporting dashboard, you must have one of the following roles:

- Compliance Administrator
- Security Administrator
- Global Administrator

> **Tip**
>
> Use the least-privileged role that meets your needs. For most organizations, the Compliance Administrator or Security Administrator role is sufficient.

Devices must also:

- Be onboarded to Endpoint DLP
- Be actively reporting device telemetry

For information about onboarding devices, see Get started with Endpoint data loss prevention.

Open the device health reporting dashboard

To open the device health reporting dashboard, follow these steps:

1. Go to the Microsoft Purview portal.

2. Select **Settings** > **Device onboarding** > **Device report**.

What the dashboard shows

The dashboard displays data from devices that reported within the past 30 days. It updateevery hour. Each section is interactive — select **View devices** to open a side pane with device categories, then select a category to open a filtered list of matching devices.

Expand table

Visualization	Description	Use case
Device onboarding	Provides an overview of onboarded devices, including: total onboarded devices, onboarded devices running Endpoint DLP, onboarded devices with configuration issues.	Use this section to quickly understand overall onboarding coverage and identify devices that may require configuration remediation.
Device readiness to receive policy updates	Shows devices that may not be ready to receive future Endpoint DLP policy updates because of: offline status during the past day, outdated Microsoft Defender Antivirus versions, configuration issues. These indicators don't confirm whether a device failed to receive a policy update, but instead highlight devices that may be at risk of not receiving future updates.	Use this section to proactively identify devices that may require remediation to remain ready for future policy updates.
Last time devices were seen online	Provides a breakdown of devices based on when each device was last seen online. Devices are grouped based on recency of activity, including devices seen in the last 24 hours, past 3 days, past 7 days, and past 30 days. This visualization helps you understand how recently devices have been active across your organization and provides visibility into overall device connectivity and reporting recency.	Use this section to understand how recently devices have been active or reporting activity across your organization.
Device readiness for feature	Displays readiness information for supported Endpoint DLP features. Currently, this section supports readiness reporting for: Just-in-time protection, Paste to supported browsers. This section shows how many devices are ready for each feature, and how many require updates before a feature can be used. Additional Endpoint DLP features will be added to this section in future releases.	Use this section to validate readiness before enabling or deploying Endpoint DLP features across devices.

Monitor device readiness

Use the dashboard to monitor device posture at scale, find misconfigured devices, detect devices that might not receive Endpoint DLP policy updates, and validate readiness before deploying new features.

To investigate device health issues, follow this typical workflow:

1. Review dashboard summaries to identify problem areas.
2. Select a dashboard segment to open a filtered list of affected devices.
3. Investigate individual devices.
4. Identify root cause issues such as:
    - Outdated versions
    - Connectivity issues
    - Invalid configurations

Common scenarios

Monitor device posture across your organization

Organizations need to monitor device health and readiness at scale to ensure devices remain ready to receive the latest policies and supported Endpoint DLP features. Use the dashboard to continuously monitor onboarding status, policy update readiness, feature readiness, and recent device activity. The dashboard helps you find devices that need remediation.

Identify onboarded devices with configuration issues

Onboarded devices may still have configuration issues that affect Endpoint DLP enforcement or feature readiness. Use the **Device onboarding** visualization to review Endpoint DLP coverage and find devices with configuration issues that need remediation.

Identify devices that may not be ready to receive policy updates

Some devices may not be ready to receive Endpoint DLP policy updates because of connectivity, Defender version, or configuration issues. These indicators don't confirm whether a device failed to receive an update, but highlight devices at risk of not receiving future updates. Use the **Device readiness to receive policy updates** visualization to find these devices and plan remediation.

Monitor device connectivity

Device connectivity affects readiness for policy updates and Endpoint DLP features. Use the **Last time devices were seen** visualization to view when devices were last seen online — in the past 24 hours, 3 days, 7 days, or 30 days. This breakdown helps you evaluate device activity patterns across your organization.

Validate readiness before enabling Endpoint DLP features

Some Endpoint DLP features require devices to meet specific configuration or version requirements. Use the **Device readiness for feature** visualization to validate whether devices are ready for supported features, including Just-in-time protection and Paste to supported browsers.

Best practices

- **Review the dashboard regularly.** Frequent checks help you catch device health trends before they become widespread issues.
- **Investigate devices that haven't reported recently.** Devices offline for extended periods may miss critical policy updates.
- **Validate readiness before deploying new policies or features.** Confirm that devices meet configuration and version requirements to avoid failed deployments.
- **Prioritize remediation for devices with outdated Defender versions.** Outdated Microsoft Defender Antivirus versions can prevent devices from receiving Endpoint DLP policies.

ⓘ **Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/05/2026

Onboard Windows devices into Microsoft 365 overview

Applies to:

- Endpoint data loss prevention (DLP)
- Insider risk management

Endpoint data loss prevention (Endpoint DLP) and insider risk management require that Windows 10 and Windows 11 devices be onboarded into the service so that they can send monitoring data to the services.

Endpoint DLP allows you to monitor Windows 10 or Windows 11 devices and detect when sensitive items are used and shared. This gives you the visibility and control you need to ensure that they're used and protected properly, and to help prevent risky behavior that might compromise them. For more information about all of Microsoft's DLP offerings, see Learn about data loss prevention. To learn more about Endpoint DLP, see Learn about Endpoint data loss prevention.

Endpoint DLP also allows you to onboard devices running Windows Server 2019 and later versions.

> **Note**
>
> Installing the supported Windows Server KBs disables the *Classification* feature on the server. This means that Endpoint DLP won't classify files on the server. However, Endpoint DLP will still protect those files on the server that were classified before those KBs were installed on server. To ensure this protection, install Microsoft Defender version 4.18.23100 (October 2023) or later.

By default, Endpoint DLP isn't enabled for Windows servers when they're initially onboarded. Before you can see Endpoint DLP events for your servers in Activity Explorer, you must first turn on Endpoint DLP support for onboarded servers.

Once properly configured, the same data loss protection policies can be automatically applied to both Windows PCs and Windows servers.

Insider risk management uses the full breadth of service and 3rd-party indicators to help you quickly identify, triage, and act on risky user activity. By using logs from Microsoft 365 and Microsoft Graph, insider risk management allows you to define specific policies to identify risk indicators and to take action to mitigate these risks. For more information, see Learn about insider risk management.

Device onboarding is shared across Microsoft 365 and Microsoft Defender for Endpoint (MDE). If you've already onboarded devices to MDE, they appear in the managed devices list and no further steps are necessary to onboard those specific devices. Onboarding devices in Microsoft Purview portal also onboards them into MDE.

Before you begin

SKU/subscriptions licensing

For information on licensing, see

- Microsoft 365 Enterprise Plans
- Microsoft 365 Service Descriptions

> **Important**
>
> For information on managing user licensing for devices see Available plans.

Permissions

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. This helps improve security for your organization. Global Administrator is a highly privileged role that should only be used in scenarios where a lesser privileged role can't be used.

To enable device management, the account you use must be a member of any one of these roles:

- Security admin
- Compliance admin
- Global admin

If you want to use a custom account to view the device management settings, it must be in one of these roles:

- Compliance admin
- Compliance data admin
- Global reader
- Security admin
- Global admin

If you want to use a custom account to access the onboarding/offboarding page, it must be in one of these roles:

- Compliance admin
- Security admin
- Global admin

If you want to use a custom account to turn on/off device monitoring, it must be in one of these roles:

- Compliance admin
- Security admin
- Global admin

Prepare your Windows devices

Make sure that the Windows devices that you need to onboard meet these requirements.

1. Must be running one of the following builds of Windows or Windows Server:

    a. Windows (X64):
        i. Windows 11 24H2 (See update details)
        ii. Windows 11 23H2 (See update details)
        iii. Windows 11 22H2 update (See update details)
        iv. Windows 11 21H2 (See update details)
        v. Windows 10 22H2 update (See update details)

    b. Windows (ARM64):
        i. Windows 11 24H2 (See update details)
        ii. Windows 11 23H2 (See update details)
        iii. Windows 11 22H2 update (See update details)
        iv. Windows 11 21H2 (See update details)

c. Windows Server 2019 OS: 1809 onwards or Windows Server 2022 OS: 21H2 onwards.

d. Anti-malware Client Version is 4.18.2110 or newer. Check your current version by opening Windows Security app, select the Settings icon, and then select About. The version number is listed under anti-malware Client Version. Update to the latest anti-malware Client Version by installing Windows Update KB4052623. For more information, see: Microsoft Defender Antivirus in Windows. After a new anti-malware Client version or a new anti-malware Engine version is released, versions older than N-2 are no longer supported.

> ⓘ **Important**
>
> None of Windows Security components need to be active, but the **Real-time protection and Behavior monitor** must be enabled.
>
> After a new anti-malware Client version or a new anti-malware Engine version is released, versions older than N-2 are no longer supported.

2. All devices must be one of these:

   - Microsoft Entra joined
   - Microsoft Entra hybrid joined
   - Microsoft Entra registered

3. A supported version of Microsoft 365 Apps is installed and up to date. For the most robust protection and user experience, ensure Microsoft 365 Apps version 16.0.14701.0 or later is installed.

> ⓘ **Note**
>
> - If you're running Office 365 - KB 4577063 is required.
> - If you are on Monthly Enterprise Channel of Microsoft 365 Apps versions 2004-2008, you need to update to version 2009 or later. See **Update history for Microsoft 365 Apps (listed by date)** for current versions. To learn more about known issue, see the Office Suite section of **Release notes for Current Channel releases in 2020**.

4. If you have endpoints that use a device proxy to connect to the internet, follow the procedures in Configure device proxy and internet connection settings for Microsoft Purview Endpoint DLP.

   > **Important**
   >
   > Make sure you allow *MpDlpService.exe* through your firewall, third-party antivirus software, or application control.

5. Here are the prerequisites for each feature.

Expand table

Feature	Prerequisite
Add hyperlink support in warn and block toast messages (Roadmap ID 480733)	Anti-malware Client Version 4.18.25010 or newer is required.
IP or IP address ranges support within Sensitive service domain groups (Roadmap ID 479759)	Anti-malware Client Version 4.18.25010 or newer is required.
Allow and Off modes in the policy (Roadmap ID 481356)	anti-malware Client Version 4.18.25010 or newer is required.
Device and Device group-based policy scoping support for Endpoint DLP (Roadmap ID 480732)	Windows 10: See update details Windows 11 22H2: See update details Windows 11 23H2: See update details Windows 11 24H2: See update details
Full file evidence support for restricted apps (Roadmap ID 479757)	Anti-malware Client Version 4.18.25020 or newer is required.
Pause and Resume support for printer (Roadmap ID 486369)	Anti-malware Client Version 4.18.25030 or newer is required.
Increase file type coverage to 100+ (Roadmap ID 416481)	Windows 10: See update details Windows 11 22H2: See update details Windows 11 23H2: See update details Windows 11 24H2: See update details

Feature	Prerequisite
Endpoint DLP support for Teams	Anti-malware Client Version 4.18.25050 or newer is required.
Advanced label-based protection for all files on devices (Roadmap ID 487859)	Information protection client version: 3.1.309 or newer is required. Anti-malware Client Version 4.18.25050 or newer is required. Windows 10: See update details Windows 11 22H2: See update details Windows 11 23H2: See update details Windows 11 24H2: See update details
Network share/Removable media/Printer group support on Edge (Roadmap ID 486370)	Edge: 137 or newer is required. Anti-malware Client Version 4.18.25050 or newer is required.
OneDrive shadow file protection support (Roadmap ID 489842)	Anti-malware Client Version 4.18.26020 or newer is required.
Restricted apps and app groups support on Edge (Roadmap ID 485769)	Anti-malware Client Version 4.18.25080 or newer is required.
Logged-in user details (user email) on the device onboarding page (Roadmap ID 548669)	Anti-malware Client Version 4.18.26030 or newer is required. Windows 10: See update details Windows 11 22H2: See update details Windows 11 23H2: See update details Windows 11 24H2: See update details
Device health details in the export function on the device onboarding page and in Advanced hunting on Security portal (Roadmap ID 542792)	Anti-malware Client Version 4.18.26030 or newer is required. Windows 10: See update details Windows 11 22H2: See update details Windows 11 23H2: See update details Windows 11 24H2: See update details
Network share/Removable media/Printer group support on Edge (Roadmap ID 486370)	Edge: 137 or newer is required. Anti-malware Client Version

Feature	Prerequisite
	4.18.25050 or newer is required.
Hyperlink within toast messages for the Edge browser (Roadmap ID 558688)	Anti-malware Client Version 4.18.26040 or newer is required.
Print resume support for Just-in-time protection (JIT)	Anti-malware Client Version 4.18.26040 or newer is required.

Onboarding Windows 10 or Windows 11 devices

You must enable device monitoring and onboard your endpoints before you can monitor and protect sensitive items on a device. Both of these actions are done in the Microsoft Purview portal.

When you want to onboard devices that haven't been onboarded yet, you download the appropriate script and deploy it to those devices. Follow the device onboarding procedures below.

If you already have devices onboarded into Microsoft Defender for Endpoint, they'll already appear in the managed devices list.

In this deployment scenario, you onboard Windows 10 or Windows 11 devices that haven't been onboarded yet.

1. Open the Microsoft Purview portal. Choose **Settings** > **Device onboarding** > **Devices**.

   > **Note**
   >
   > If you have previously deployed Microsoft Defender for Endpoint, all the devices that were onboarded during that process will be listed in the **Devices** list. There's no need to onboard them again. While it usually takes about 60 seconds for device onboarding to be enabled, allow up to 30 minutes before engaging with Microsoft support.

2. Choose **Turn on device onboarding**.

3. Choose **Onboarding** to begin the onboarding process.

4. Choose the way you want to deploy to these other devices from the **Deployment method** list and then **download package**.

5. Choose the appropriate procedure to follow from the table below:

Expand table

Article	Description
Intune	Use Mobile Device Management tools or Microsoft Intune to deploy the configuration package on device.
Configuration Manager	You can use either use Microsoft Endpoint Configuration Manager (current branch) version 1606 or Microsoft Endpoint Configuration Manager (current branch) version 1602 or earlier to deploy the configuration package on devices.
Group Policy	Use Group Policy to deploy the configuration package on devices.
Local script	Learn how to use the local script to deploy the configuration package on endpoints.
Virtual desktop infrastructure (VDI) devices	Learn how to use the configuration package to configure VDI devices.

Check device status

After onboarding your devices, you can monitor **Configuration status** and **Policy Sync status** to ensure devices are properly set up to receive policy updates. To learn what Configuration status and Policy sync status mean, see: Understand device configuration and policy sync status values

To view these details, go to **Device onboarding** > **Devices**. From there, you can review each device's current configuration state and confirm whether policies have successfully synced. For guidance on resolving **Configuration status** or **Policy Sync status** issues, see: Troubleshooting endpoint data loss prevention configuration and policy sync

> **Note**
>
> Onboarded devices will continue to appear for 180 days after the device is offline.

See also

- Learn about insider risk management
- Learn about Endpoint data loss prevention
- Using Endpoint data loss prevention

- Learn about data loss prevention
- Create and Deploy data loss prevention policies
- Get started with Activity explorer
- Microsoft Defender for Endpoint
- Onboarding tools and methods for Windows 10 machines
- Microsoft 365 subscription
- Microsoft Entra joined devices
- Download the new Microsoft Edge based on Chromium

Last updated on 06/26/2026

Onboard Windows 10 and Windows 11 devices using Mobile Device Management tools

09/26/2025

Applies to:

- Endpoint data loss prevention (DLP)
- Insider risk management

You can use mobile device management (MDM) solutions to configure devices. Microsoft 365 information protection supports MDM solutions by providing OMA-URIs to create policies to manage devices.

Before you begin

If you're using Microsoft Intune, the device must be enrolled in MDM.

For more information on enabling MDM with Microsoft Intune, see Device enrollment (Microsoft Intune).

Onboard devices using Microsoft Intune

Follow the instructions from Intune.

> **Note**
>
> The **Health Status for onboarded devices** policy uses read-only properties and can't be remediated.

Offboard and monitor devices using Mobile Device Management tools

For security reasons, the package used to offboard devices will expire 30 days after the date it was downloaded. Expired offboarding packages sent to a device will be rejected. When you download an offboarding package, you're notified of the package's expiry date. The expiry date is also included in the package name.

> **Note**
>
> Onboarding and offboarding policies must not be deployed on the same device at the same time. If they are, unpredictable collisions result.

1. Get the offboarding package from the Microsoft Purview portal.
2. In the navigation pane, select **Settings** > **Device onboarding** > **Offboarding**.
3. In the **Deployment method** field, select **Mobile Device Management / Microsoft Intune**.
4. Select **Download package**, and save the .zip file.
5. Extract the contents of the .zip file to a shared, read-only location that can be accessed by the network administrators who deploy the package. You should have a file named *DeviceCompliance_valid_until_YYYY-MM-DD.offboarding*.
6. Use the Microsoft Intune custom configuration policy to deploy the following supported OMA-URI settings.

   ```
   Console

   OMA-URI: ./Device/Vendor/MSFT/WindowsAdvancedThreatProtection/Offboarding
   Date type: String
   Value: [Copy and paste the value from the content of the
   DeviceCompliance_valid_until_YYYY-MM-DD.offboarding file]
   ```

   > **Note**
   >
   > If Microsoft Defender for Endpoint is already configured, you can **Turn on device onboarding**. If you do this, step 6 isn't required.

> **Note**
>
> The **Health Status for offboarded devices** policy uses read-only properties and can't be remediated.

> **Important**

Offboarding causes the device to stop sending sensor data to the portal. However, data from the device, including reference to any alerts received, will be retained for up to six months.

Related articles

- Onboard Windows 10 devices using Group Policy
- Onboard Windows 10 devices using Microsoft Endpoint Configuration Manager
- Onboard Windows 10 devices using a local script
- Onboard non-persistent virtual desktop infrastructure (VDI) devices
- Troubleshoot Microsoft Defender Advanced Threat Protection onboarding issues

Onboard Windows 10 and Windows 11 devices using Microsoft Configuration Manager

Article • 02/18/2025

Applies to:

- Endpoint data loss prevention (DLP)
- Insider risk management

Supported client operating systems

You can onboard the following operating systems using Configuration Manager:

- Windows 11
- Windows 10, version 1709 or newer

> **Important**
>
> When onboarding systems to Microsoft Purview for endpoint, they will also report to Defender for Endpoint, even if a different antivirus or anti-malware solution is in use. In such cases, these systems will report to Defender for Endpoint in **passive mode**, ensuring there is no interference with the existing antivirus or anti-malware solution.

Onboarding to Microsoft Purview for Endpoint with Configuration Manager

In Configuration Manager, configure the Endpoint Protection settings as follows to ensure the client recognizes the configuration file.

- Set **Manage Endpoint protection client on the computers** to **Yes**.
- The **Install Endpoint Protection Client on Client computers** can be set to **No**.

> **Note**

This is to allow clients to receive the configuration file from Configuration Manager. If you do not turn this on the Configuration file may not deploy on to the clients.

- Set the Microsoft Defender for Endpoint client on Windows server 2012 R2 and Windows Server 2016 to **MDE client (recommended)**.

Deploy this Client Settings configuration to the collections that are used to onboard systems.

Onboard devices using Configuration Manager

1. Get the configuration package .zip file (*DeviceComplianceOnboardingPackage.zip*) from Microsoft Purview portal.
2. In the navigation pane, select **Settings** > **Device Onboarding** > **Onboarding**.
3. In the **Select Operating system to start onboarding process**: select **Windows 10** and then in the **Deployment method** field, select **Microsoft Endpoint Configuration Manager**.
4. Select **Download package** and save the .zip file.
5. Extract the contents of the .zip file to a shared, read-only location accessible by the Configuration Manager console for deployment. Ensure the file named ***DeviceCompliance.onboarding*** is present.

Onboard the devices

1. In the Configuration Manager console, navigate to **Assets and Compliance** > **Endpoint Protection** > **Microsoft Defender for Endpoint Policies**.

   > **Note**
   >
   > Microsoft Purview and Microsoft Defender for Endpoint use the same method to connect to Microsoft cloud security environments.

2. Select **Create Microsoft Defender for Endpoint Policy** to open the policy wizard.
3. Select **I accept these license terms and automatic updates of both Agents**. Then type the **Name** and **Description** for the Microsoft Defender for Endpoint policy and select **Onboarding**.

4. Select **Browse** and browse to the configuration file you extracted from the downloaded .zip file.

   > ⓘ **Note**
   >
   > You do not need the Workspace key and Workspace ID for Windows 10 and 11.

   Select **Next**.

5. Specify the file samples that are collected and shared from managed devices for analysis.

   - **None**
   - **All file types**

6. Review the summary and complete the wizard.

7. Right-click on the policy you created, select **Deploy**, and then choose a collection to which you want to deploy the Microsoft Defender for Endpoint policy.

Monitor

1. In the Configuration Manager console, navigate **Monitoring** > **Deployments** and then select the deployment you created for the **Defender for Endpoint policy** deployment.

2. Click on **View Status** to review the information. Under **Compliant** you can see the status of the onboarded systems. Systems that have not yet been onboarded may appear under the **Unknown** tab.

> ⓘ **Note**
>
> This process may take some time, and the system might need to reboot for the changes to take effect.

Offboard devices using Configuration Manager

For security reasons, the package used to Offboard devices will expire 30 days after the date it was downloaded. Expired offboarding packages sent to a device will be rejected.

When downloading an offboarding package, you'll be notified of the packages expiry date and it will also be included in the package name.

> **Note**
>
> Onboarding and offboarding policies must not be deployed on the same device at the same time, otherwise this causes unpredictable collisions.

Offboard devices using Microsoft Configuration Manager current branch

If you use Microsoft Configuration Manager current branch, see Create an offboarding configuration file.

Create an offboarding configuration file

1. Sign in to the Microsoft Purview console.
2. Select **Settings**, then select **Device onboarding** and then select **Offboarding** under the onboarding heading.
3. Select **Windows 10** for the operating system and **Microsoft Endpoint Configuration Manager** for the deployment method.
    - Using the **Windows 10** option ensures that all devices in the collection are off boarded, and the MMA is uninstalled when needed.
4. Download the compressed archive (.zip) file and extract the contents. Offboarding files are valid for 30 days.
5. In the Configuration Manager console, navigate to **Assets and Compliance** > **Endpoint Protection** > **Microsoft Defender For Endpoint Policies** and select **Create Microsoft Defender for Endpoint Policy**. The policy wizard opens.
6. Type the **Name** and **Description** for the Microsoft Defender for Endpoint policy and select **Offboarding**.
7. **Browse** to the configuration file you extracted from the downloaded .zip file.
8. Review the summary and complete the wizard.
9. Select Deploy to target the Microsoft Defender for Endpoint policy to clients you wish to offboard from Purview.

> **Important**
>
> The Microsoft Purview for Endpoint configuration files contains sensitive information which should be kept secure.

> **Tip**
>
> If you're not an E5 customer, use the 90-day Microsoft Purview solutions trial to explore how additional Purview capabilities can help your organization manage data security and compliance needs. Start now at the **Microsoft Purview trials hub**. Learn details about **signing up and trial terms**.

Monitor device configuration

With Microsoft Configuration Manager current branch, use the built-in Microsoft Defender for Endpoint dashboard in the Configuration Manager console. For more information, see Microsoft Defender Advanced Threat Protection - Monitor.

Check that the devices are compliant with the Endpoint data loss prevention service

You can set a compliance rule for configuration item in Configuration Manager to monitor your deployment.

> **Note**
>
> This procedure and registry entry applies to Endpoint DLP and Defender for Endpoint.

This rule should be a *non-remediating* compliance rule configuration item that monitors the value of a registry key on targeted devices.

Monitor the following registry key entry:

Console

```
Path: "HKLM\SOFTWARE\Microsoft\Windows Advanced Threat Protection\Status"
Name: "OnboardingState"
Value: "1"
```

For more information, see Plan for and configure compliance settings.

Related articles

- Onboard Windows 10 and Windows 11 devices using Group Policy
- Onboard Windows 10 and Windows 11 devices using Mobile Device Management tools
- Onboard Windows 10 and Windows 11 devices using a local script
- Onboard non-persistent virtual desktop infrastructure (VDI) devices
- Run a detection test on a newly onboarded Microsoft Defender for Endpoint device
- Troubleshoot Microsoft Defender for Endpoint onboarding issues

---

Feedback

Was this page helpful? Yes No

Onboard Windows 10 devices and Windows 11 using Group Policy

Applies to:

- Endpoint data loss prevention (DLP)
- Insider risk management
- Group Policy

> **Note**
>
> To use Group Policy (GP) updates to deploy the package, you must be on Windows Server 2008 R2 or later.
>
> For Windows Server 2019, you may need to replace NT AUTHORITY\Well-Known-System-Account with NT AUTHORITY\SYSTEM of the XML file that the Group Policy preference creates.

This article explains how to use Group Policy to deploy the Microsoft Purview device configuration package on Windows 10 and Windows 11 devices. After onboarding, devices can be monitored by endpoint data loss prevention and insider risk management features. The procedures cover both onboarding and offboarding scenarios. This guidance is intended for IT administrators who manage device policies through Group Policy Objects (GPOs).

Onboard devices using Group Policy

Use the following steps to download and deploy the Group Policy onboarding package:

1. Open the Purview portal.
2. In the navigation pane, select **Settings** > **Device Onboarding**.
3. In the **Deployment method** field, select **Group policy**.
4. Click **Download package** and save the .zip file.
5. Extract the contents of the .zip file to a shared, read-only location that can be accessed by the device. You should have a folder called *OptionalParamsPolicy* and the file

DeviceComplianceLocalOnboardingScript.cmd.

6. Open the Group Policy Management Console (GPMC), right-click the Group Policy Object (GPO) you want to configure and click **Edit**.

7. In the **Group Policy Management Editor**, go to **Computer configuration**, then **Preferences**, and then **Control panel settings**.

8. Right-click **Scheduled tasks**, point to **New**, and then click **Immediate Task (At least Windows 7)**.

9. In the **Task** window that opens, go to the **General** tab. Under **Security options** click **Change User or Group** and type SYSTEM and then click **Check Names** then **OK**. NT AUTHORITY\SYSTEM appears as the user account the task will run as.

10. Select **Run whether user is logged on or not** and check the **Run with highest privileges** check box.

11. Go to the **Actions** tab and click **New...** Ensure that **Start a program** is selected in the **Action** field. Enter the file name and location of the shared *WindowsDefenderATPOnboardingScript.cmd* file.

12. Click **OK** and close any open GPMC windows.

Offboard devices using Group Policy

For security reasons, the package used to offboard devices will expire 30 days after the date it was downloaded. Expired offboarding packages sent to a device will be rejected. When downloading an offboarding package you will be notified of the packages expiry date and it will also be included in the package name.

> **ⓘ Important**
>
> Offboarding causes the device to stop sending sensor data to the portal. However, data from the device, including references to any alerts it has had, will be retained for up to six months.

> **ⓘ Note**

Onboarding and offboarding policies must not be deployed on the same device at the same time. Deploying both policies simultaneously causes unpredictable collisions.

1. Get the offboarding package from Microsoft Purview portal.
2. In the navigation pane, select **Settings** > **//Device onboarding** > **Offboarding**.
3. In the **Deployment method** field, select **Group policy**.
4. Click **Download package** and save the .zip file.
5. Extract the contents of the .zip file to a shared, read-only location that can be accessed by the device. You should have a file named *DeviceComplianceOffboardingScript_valid_until_YYYY-MM-DD.cmd*.
6. Open the Group Policy Management Console (GPMC), right-click the Group Policy Object (GPO) you want to configure and click **Edit**.
7. In the **Group Policy Management Editor**, go to **Computer configuration,** then **Preferences**, and then **Control panel settings**.
8. Right-click **Scheduled tasks**, point to **New**, and then click **Immediate task**.
9. In the **Task** window that opens, go to the **General** tab. Choose the local SYSTEM user account (BUILTIN\SYSTEM) under **Security options**.
10. Select **Run whether user is logged on or not** and check the **Run with highest privileges** check-box.
11. Go to the **Actions** tab and click **New...**. Ensure that **Start a program** is selected in the **Action** field. Enter the file name and location of the shared *DeviceComplianceOffboardingScript_valid_until_YYYY-MM-DD.cmd* file.
12. Click **OK** and close any open GPMC windows.

Monitor device configuration

With Group Policy there isn't an option to monitor deployment of policies on the devices. Monitoring can be done directly in the Microsoft Purview portal, or by using other deployment tools such as Microsoft Endpoint Configuration Manager or Mobile Device Management tools.

Monitor devices using the portal

To verify that onboarded devices are reporting correctly, use the following steps:

1. Go to the Microsoft Purview portal.
2. Click **Devices** list.
3. Verify that devices are appearing.

> **Note**
>
> It can take several days for devices to start showing on the **Devices list**. This includes the time it takes for the policies to be distributed to the device, the time it takes before the user logs on, and the time it takes for the endpoint to start reporting.

Related content

For related onboarding methods and follow-up tasks, see the following articles:

- Onboard Windows 10 and Windows 11 devices using Microsoft Endpoint Configuration Manager
- Onboard Windows 10 and Windows 11 devices using Mobile Device Management tools
- Onboard Windows 10 and Windows 11 devices using a local script
- Onboard non-persistent virtual desktop infrastructure (VDI) devices
- Run a detection test on a newly onboarded Microsoft Defender for Endpoint devices
- Troubleshoot Microsoft Defender Advanced Threat Protection onboarding issues

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Onboard Windows 10 and Windows 11 devices using a local script

09/26/2025

Applies to:

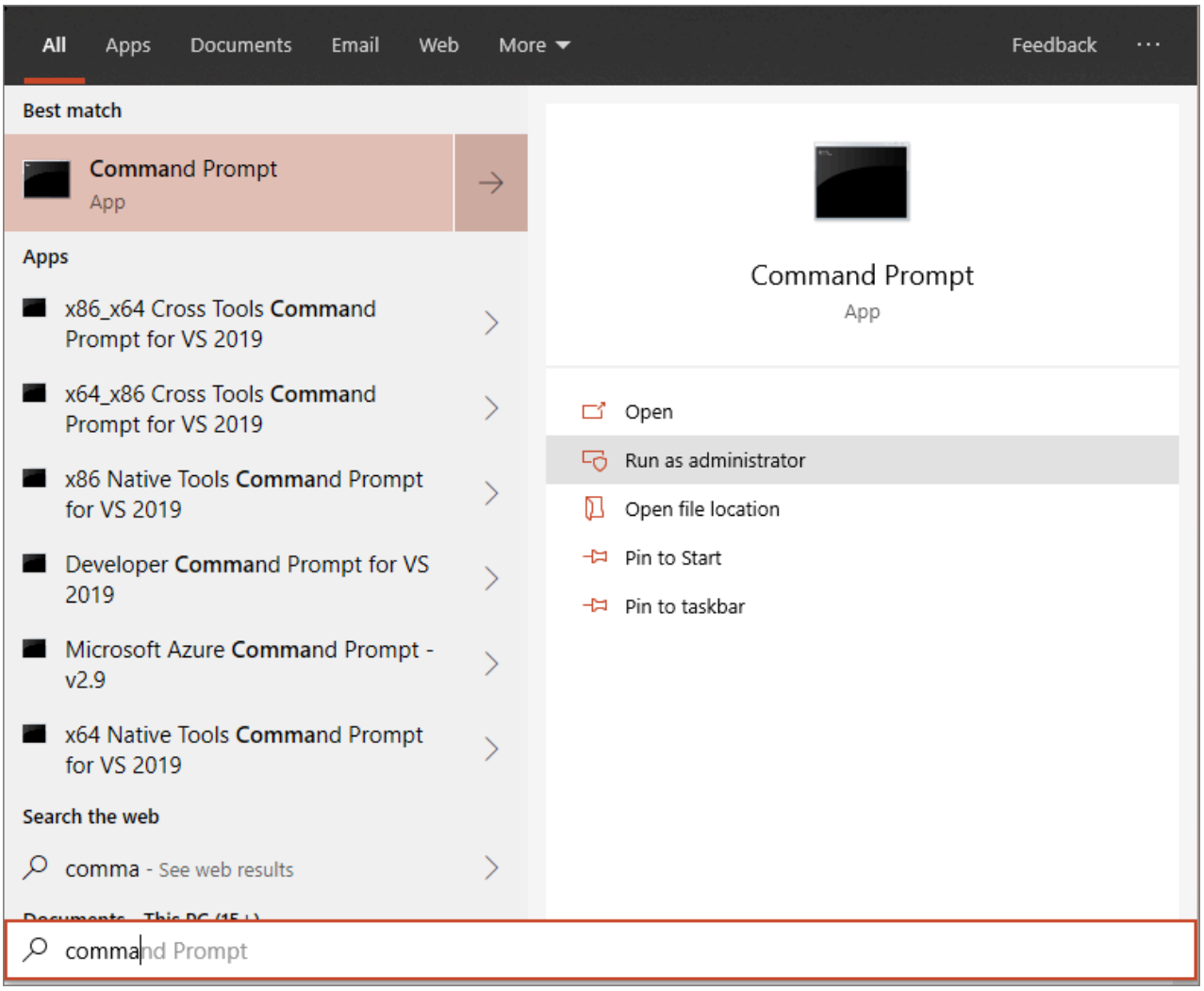
- Endpoint data loss prevention (DLP)
- Insider risk management

You can also manually onboard individual devices to Microsoft 365. You might want to do this first when testing the service before you commit to onboarding all devices in your network.

> **Important**
>
> This script has been optimized for use on up to 10 devices.
>
> To deploy at scale, use other deployment options. For example, you can deploy an onboarding script to more than 10 devices in production with the script available in Onboard Windows 10 devices using Group Policy.

Onboard devices

1. Get the configuration package .zip file (*DeviceComplianceOnboardingPackage.zip*) package from Microsoft Purview portal
2. In the navigation pane, select **Settings** > **Device onboarding**.
3. In the **Deployment method** field, select **Local Script**.
4. Click **Download package** and save the .zip file.
5. Extract the contents of the configuration package to a location on the device you want to onboard (for example, the Desktop). You should have a file named *DeviceComplianceLocalOnboardingScript.cmd*.
6. Open an elevated command-line prompt on the device and run the script:
7. Go to **Start** and type **cmd**.
8. Right-click **Command prompt** and select **Run as administrator**.



9. Type the location of the script file. If you copied the file to the desktop, type: *%userprofile%\Desktop\DeviceComplianceLocalOnboardingScript.cmd*

10. Press the **Enter** key or click **OK**.

For information on how you can manually validate that the device is compliant and correctly reports sensor data see, Troubleshoot Microsoft Defender Advanced Threat Protection onboarding issues.

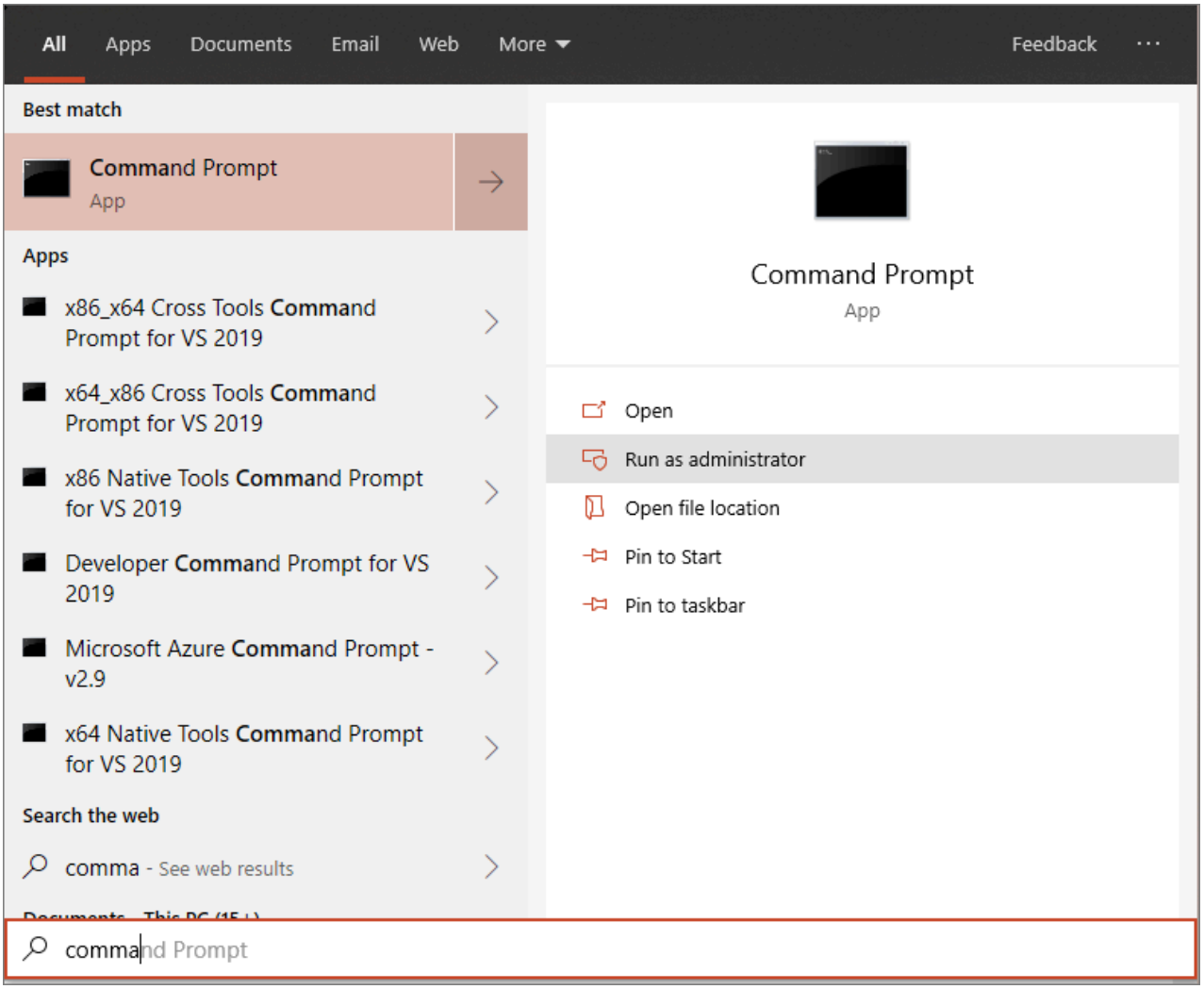
Offboard devices using a local script

For security reasons, the package used to Offboard devices will expire 30 days after the date it was downloaded. Expired offboarding packages sent to a device will be rejected. When downloading an offboarding package you will be notified of the packages expiry date and it will also be included in the package name.

> ⓘ **Note**

Onboarding and offboarding policies must not be deployed on the same device at the same time, otherwise this will cause unpredictable collisions.

1. Get the offboarding package from Microsoft Purview portal.
2. In the navigation pane, select **Settings** > **Device offboarding**.
3. In the **Deployment method** field, select **Local Script**.
4. Click **Download package** and save the .zip file.
5. Extract the contents of the .zip file to a shared, read-only location that can be accessed by the devices. You should have a file named *DeviceComplianceOffboardingScript_valid_until_YYYY-MM-DD.cmd*.
6. Open an elevated command-line prompt on the device and run the script:
7. Go to **Start** and type **cmd**.
8. Right-click **Command prompt** and select **Run as administrator**.



9. Type the location of the script file. If you copied the file to the desktop, type: *%userprofile%\Desktop\WindowsDefenderATPOffboardingScript_valid_until_YYYY-MM-DD.cmd*

10. Press the **Enter** key or click **OK**.

> **Important**
>
> Offboarding causes the device to stop sending sensor data to the portal.

Monitor device configuration

You can follow the different verification steps in the [Troubleshoot onboarding issues] ((/windows/security/threat-protection/microsoft-defender-atp/troubleshoot-onboarding) to verify that the script completed successfully and the agent is running.

Monitoring can also be done directly on the portal, or by using the different deployment tools.

Monitor devices using the portal

1. Go to Microsoft Purview portal.

2. Choose **Settings** > **Device onboarding** > **Devices**.

3. Go to Microsoft Purview portal, and select **Settings** > **Device onboarding** > **Devices**.

4. Verify that devices are appearing.

Related topics

- Onboard Windows 10 and Windows 11 devices using Group Policy
- Onboard Windows 10 and Windows 11 devices using Microsoft Endpoint Configuration Manager
- Onboard Windows 10 and Windows 11 devices using Mobile Device Management tools
- Onboard non-persistent virtual desktop infrastructure (VDI) devices
- Run a detection test on a newly onboarded Microsoft Defender for Endpoint device
- Troubleshoot Microsoft Defender Advanced Threat Protection onboarding issues

Onboard non-persistent virtual desktop infrastructure devices

09/26/2025

Applies to:

- Endpoint data loss prevention (DLP)
- Insider risk management
- Virtual desktop infrastructure (VDI) devices

Onboard VDI devices

Microsoft 365 supports non-persistent virtual desktop infrastructure (VDI) session onboarding.

> **Note**
>
> To onboard non-persistent VDI sessions, VDI devices must be on Windows 10 1809 or higher.

There might be associated challenges when onboarding VDIs. The following are typical challenges for this scenario:

- Instant early onboarding of short-lived sessions, which must be onboarded to Microsoft 365 prior to the actual provisioning.
- The device name is typically reused for new sessions.

VDI devices can appear in the Microsoft Purview portal as either:

- Single entry for each device. In this case, the *same* device name must be configured when the session is created, for example using an unattended answer file.
- Multiple entries for each device - one for each session.

The following steps guide you through onboarding VDI devices and will highlight steps for single and multiple entries.

> **Warning**
>
> Endpoint data loss prevention support for Windows Virtual Desktop supports both single session and multi-session scenarios. For environments where there are low resource

configurations, the VDI boot procedure might slow the device onboarding process.

1. Get the VDI configuration package .zip file (*DeviceCompliancePackage.zip*) from Microsoft Purview portal.
2. In the navigation pane, select **Settings** > **Device onboarding** > **Onboarding**.
3. In the **Deployment method** field, select **VDI onboarding scripts for non-persistent endpoints**.
4. Select **Download package** and save the .zip file.
5. Copy the files from the DeviceCompliancePackage folder extracted from the .zip file into the `golden` image under the path `C:\WINDOWS\System32\GroupPolicy\Machine\Scripts\Startup`.
6. If you aren't implementing a single entry for each device, copy DeviceComplianceOnboardingScript.cmd.
7. If you're implementing a single entry for each device, copy both Onboard-NonPersistentMachine.ps1 and DeviceComplianceOnboardingScript.cmd.

   > ⓘ Note
   >
   > If you don't see the `C:\WINDOWS\System32\GroupPolicy\Machine\Scripts\Startup` folder, it might be hidden. You need to choose the **Show hidden files and folders** option from File Explorer.

8. Open a Local Group Policy Editor window and navigate to **Computer Configuration** > **Windows Settings** > **Scripts** > **Startup**.

   > ⓘ Note
   >
   > Domain Group Policy may also be used for onboarding non-persistent VDI devices.

9. Depending on the method you'd like to implement, follow the appropriate steps:

   For single entry for each device

   Select the **PowerShell Scripts** tab, then select **Add** (Windows Explorer opens directly in the path where you copied the onboarding script earlier). Navigate to onboarding PowerShell script `Onboard-NonPersistentMachine.ps1`.

For multiple entries for each device:

Select the **Scripts** tab, then select **Add** (Windows Explorer opens directly in the path where you copied the onboarding script earlier). Navigate to the onboarding bash script `DeviceComplianceOnboardingScript.cmd`.

10. Test your solution:
    a. Create a pool with one device.
    b. Sign in device.
    c. Sign out from device.
    d. Sign in device with another user.
    e. **For single entry for each device**: Check only one entry in Microsoft Defender Security Center. **For multiple entries for each device**: Check multiple entries in Microsoft Defender Security Center.
11. Select **Devices list** on the Navigation pane.
12. Use the search function by entering the device name and select **Device** as search type.

Updating non-persistent virtual desktop infrastructure (VDI) images

As a best practice, we recommend using offline servicing tools to patch golden images.

For example, you can use the below commands to install an update while the image remains offline:

Windows Command Prompt

```
DISM /Mount-image /ImageFile:"D:\Win10-1909.vhdx" /index:1
/MountDir:"C:\Temp\OfflineServicing"
DISM /Image:"C:\Temp\OfflineServicing" /Add-Package
/Packagepath:"C:\temp\patch\windows10.0-kb4541338-x64.msu"
DISM /Unmount-Image /MountDir:"C:\Temp\OfflineServicing" /commit
```

For more information on DISM commands and offline servicing, please refer to the articles below:

- Modify a Windows image using DISM
- DISM Image Management Command-Line Options
- Reduce the Size of the Component Store in an Offline Windows Image

If offline servicing isn't a viable option for your non-persistent VDI environment, the following steps should be taken to ensure consistency and sensor health:

1. After booting the golden image for online servicing or patching, run an offboarding script to turn off the Microsoft 365 device monitoring sensor. For more information, see Offboard devices using a local script.

2. Ensure the sensor is stopped by running the command below in a CMD window:

   Windows Command Prompt

   ```
   sc query sense
   ```

3. Service the image as needed.

4. Run the below commands using PsExec.exe (which can be downloaded from https://download.sysinternals.com/files/PSTools.zip) to clean up the cyber folder contents that the sensor may have accumulated since boot:

   Windows Command Prompt

   ```
   PsExec.exe -s cmd.exe
   cd "C:\ProgramData\Microsoft\Windows Defender Advanced Threat
   Protection\Cyber"
   del *.* /f /s /q
   REG DELETE "HKLM\SOFTWARE\Microsoft\Windows Advanced Threat Protection" /v
   senseGuid /f
   exit
   ```

5. Reseal the golden image as you normally would.

Related articles

- Onboard Windows 10 and Windows 11 devices using Group Policy
- Onboard Windows 10 and Windows 11 devices using Microsoft Endpoint Configuration Manager
- Onboard Windows 10 and Windows 11 devices using Mobile Device Management tools
- Onboard Windows 10 and Windows 11 devices using a local script
- Troubleshoot Microsoft Defender Advanced Threat Protection onboarding issues

Onboard macOS devices into Microsoft 365 overview

MacOS devices can be onboarded into Microsoft Purview solutions using either Intune, JAMF Pro, or any other MDM solutions. The onboarding procedures differ depending on which management solution you use. If your macOS devices have already been onboarded into Microsoft Defender for Endpoint (MDE), there are fewer steps. See Next steps for links to the appropriate procedures for you.

Applies to:

- Endpoint data loss prevention (DLP)
- Insider risk management

Before you begin

Before you get started with Endpoint DLP on macOS devices (the three latest versions), familiarize yourself with these articles:

- Learn about Endpoint data loss prevention
- Get started with Endpoint data loss prevention

If you aren't familiar with DLP at all, you should familiarize yourself with these articles as well:

- Learn about data loss prevention
- Plan for data loss prevention (DLP)
- Data loss prevention policy reference

If you aren't familiar with Insider Risk, read these articles:

- Insider risk management
- Plan for insider risk management

Your macOS devices must already be managed through Intune, JAMF Pro or any other MDM solutions.

- To onboard into Intune, see Deployment guide: Manage macOS devices in Microsoft Intune and Enroll your Mac with Intune Company Portal.

- To onboard into JAMF Pro see, JAMF Pro administrators guide and JAMF Pro Installation and Configuration Guide for Mac
- To onboard into other MDM solution see,

To take advantage of the classification improvements on file read access on macOS devices, roadmap ID 476099, the device must have Microsoft Defender Antimalware client version 101.25012.008 or newer installed.

Supported processors

macOS devices with x64 and M1, M2 and M3 (ARM64) processors are supported.

Supported browsers

Endpoint DLP supports these browsers on macOS (three latest released versions):

- Microsoft Edge (latest version)
- Safari (latest version, macOS only)
- Chrome (latest version)
- Firefox (latest version)

Licensing guidance

See, Microsoft 365 licensing guidance for information protection.

Prerequisites for each feature

Here are the prerequisites for each feature.

Expand table

Feature	Prerequisite
Device and Device group-based policy scoping support for Endpoint DLP (Roadmap ID 526792)	101.25072 or newer build is required

Conditions supported on macOS

Once a macOS device is onboarded into Microsoft Purview solutions, you can use the following condition with data loss prevention (DLP) policies:

Content Contains – Applies to documents that contain sensitive information types and sensitivity labels.

Content is not labeled - Applies to documents that do not contain any sensitivity labels as defined by Microsoft Purview Data Loss Prevention (DLP) policies.

> ⓘ **Note**
>
> '**Content is not labeled**' condition only works for doc/docx, xls/xlsx, ppt/pptx, and pdf file for Endpoint DLP.

File type is - This condition allows you to specify which types of files you want to monitor, restrict, or apply certain rules to, based on their format.

File extension is - This condition allows you to create policies that target files based on their extensions such as pdf, docx etc.

Document name contains words - Detects documents where the file name contains specific words or phrases such as 'credit'.

Document could not be scanned - Applies to documents when their content wasn't scanned. Examples include password protected files, files for which text extraction failed, files that exceed file size limit (64 MB for non-archive, 256 MB for archive)

Document or Attachment is password protected - Detects documents that are password protected.

Document couldn't complete scanning - Applies to documents when their content was scanned, however the entire document was not scanned. Examples include files for which extracted text exceeded the limits/thresholds.

Document Size equals or is greater than - Detects documents with file sizes that are greater than or equal to the specified value.

Activities that can be audited and restricted on macOS

Once a macOS device is onboarded into Microsoft Purview solutions, you can monitor and restrict the following actions using data loss prevention (DLP) policies.

Copy to a USB removable media – When enforced, this action blocks, warns, or audits the copying or moving of protected files from an endpoint device to USB removable media.

Copy to network shares – When enforced, this action blocks, warns, or audits the copying or moving of protected files from an endpoint device to any network share.

Print – When enforced, this action blocks, warns, or audits when protected files are printed from an endpoint device.

Copy to clipboard – When enforced, this action blocks, warns, or audits data in protected file that is being copied to a clipboard on an endpoint device.

Upload to cloud – This action blocks, warns, or audits when protected files are uploaded or prevented from being uploaded to cloud services based on the allow/unallowed domains list in global settings. When this action is set to warn or block, other browsers (defined on the unallowed browsers list under Global settings) are blocked from accessing the file.

Accessed by unallowed apps – When enforced, this action prevents applications that are on the unallowed apps list (as defined in Global settings) from accessing protected files on an endpoint device.

Onboarding devices into device management

You must enable device monitoring and onboard your endpoints before you can monitor and protect sensitive items on a device. Both of these actions are done in the Microsoft Purview portal.

When you want to onboard devices that haven't been onboarded yet, download the appropriate script and deploy it to those devices.

1. Open the Microsoft Purview portal **Settings** page, expand **Device onboarding**, select **Devices**, and choose **Enable device monitoring**.

> **Note**
>
> While it usually takes about 60 seconds for device onboarding to be enabled, please allow up to 30 minutes before engaging with Microsoft support.

1. Once device Monitoring is enabled, choose **Turn on macOS device monitoring**.

Next steps

Getting devices onboarded into Microsoft Purview solutions is required in order to receive DLP sensor telemetry and to enforce data loss prevention policies. As mentioned, macOS devices can be onboarded into Microsoft Purview solutions using Intune, JAMF Pro or any other MDM solutions. Review the following articles for the procedures appropriate to your situation.

Expand table

Topic	Description
Intune	For macOS devices that are managed through Intune
Intune for Microsoft Defender for Endpoint customers	For macOS devices that are managed through Intune and that have Microsoft Defender for Endpoint (MDE) deployed to them
JAMF Pro	For macOS devices that are managed through JAMF Pro
JAMF Pro for Microsoft Defender for Endpoint	For macOS devices that are managed through JAMF Pro and that have Microsoft Defender for Endpoint (MDE) deployed to them
Any MDM	For macOS devices that are managed through Any MDM
Any MDM for Microsoft Defender for Endpoint	For macOS devices that are managed through any MDM that have Microsoft Defender for Endpoint (MDE) deployed to them

Device configuration and policy sync status

You can check the **Configuration status** and the **Policy sync status** of all your onboarded devices in the **Devices** list. For macOS devices, the Microsoft Defender for Endpoint (MDE) minimum version is 101.95.07. For more information on the configuration and policy status, select an onboarded device and then open the details pane.

Configuration status shows you whether the device is configured correctly, meets DLP configuration requirements, and the last time the configuration was validated. For macOS devices, configuration includes:

- If you use Intune, check the UPN configuration by making sure your devices are onboarded into Intune.
- If you use Intune, make sure that your devices are enrolled in the Company Portal
- If you use JAMF Pro, make sure your devices are onboarded before checking the UPN configuration.
- If you use other MDM solutions, make sure your devices are onboarded in your solution before checking the UPN configuration.

For more information see: https://learn.microsoft.com/en-us/entra/identity-platform/apple-sso-plugin

Policy sync status shows you whether the most current versions of the endpoint DLP policies have been synchronized to the device and the last time a policy sync occurred.

Expand table

Field value	Configuration status	Policy sync status
Updated	Device health parameters are enabled and correctly set.	Device has been updated with the current versions of policies.
Not updated	You need to enable the configuration settings for this device. Follow the procedures for your environment: - Onboard and offboard macOS devices into Microsoft Purview solutions using Intune - Onboard and offboard macOS devices into Purview solutions using Intune for Microsoft Defender for Endpoint customers - Onboard and offboard macOS devices into Microsoft Purview solutions using JAMF Pro - Onboard and offboard macOS devices into Purview solutions using JAMF Pro for Microsoft Defender for Endpoint customers	This device hasn't synced the latest policy updates. If the policy update was made within the last 2 hours, wait for the policy to reach your device.
Not available	Device properties aren't available in the device list. This could be because the device doesn't meet the minimum OS version or configuration, or because the device was just onboarded.	Device properties aren't available in the device list. This could be because the device doesn't meet the minimum OS version or configuration, or because the device was just onboarded.

Related articles

- Use Endpoint data loss prevention
- Support Matrix for DLP policy tips across Microsoft apps

Last updated on 06/26/2026

Onboard and offboard macOS devices into Microsoft Purview solutions using Intune

09/26/2025

You can use Microsoft Intune to onboard macOS devices into Microsoft Purview solutions.

> **Important**
>
> Use this procedure if you ***do not*** have Microsoft Defender for Endpoint (MDE) deployed to your macOS devices

Applies to:

- Endpoint data loss prevention (DLP)
- Insider risk management

Before you begin

- Make sure your macOS devices are onboarded into Intune and are enrolled in the Company Portal app.
- Make sure you have access to the Microsoft Intune admin center.
- Create the user groups that you're going to assign the configuration updates to.
- OPTIONAL: Install the v95+ Microsoft Edge browser on your macOS devices to have native Endpoint DLP support on Microsoft Edge.

> **Note**
>
> The three most recent major releases of macOS are supported.

Onboard macOS devices into Microsoft Purview solutions using Microsoft Intune

Onboarding a macOS device into Purvoiew solutions is a multi-phase process.

1. Get the device onboarding package
2. Deploy the mobileconfig and onboarding packages
3. Publish the application

Prerequisites

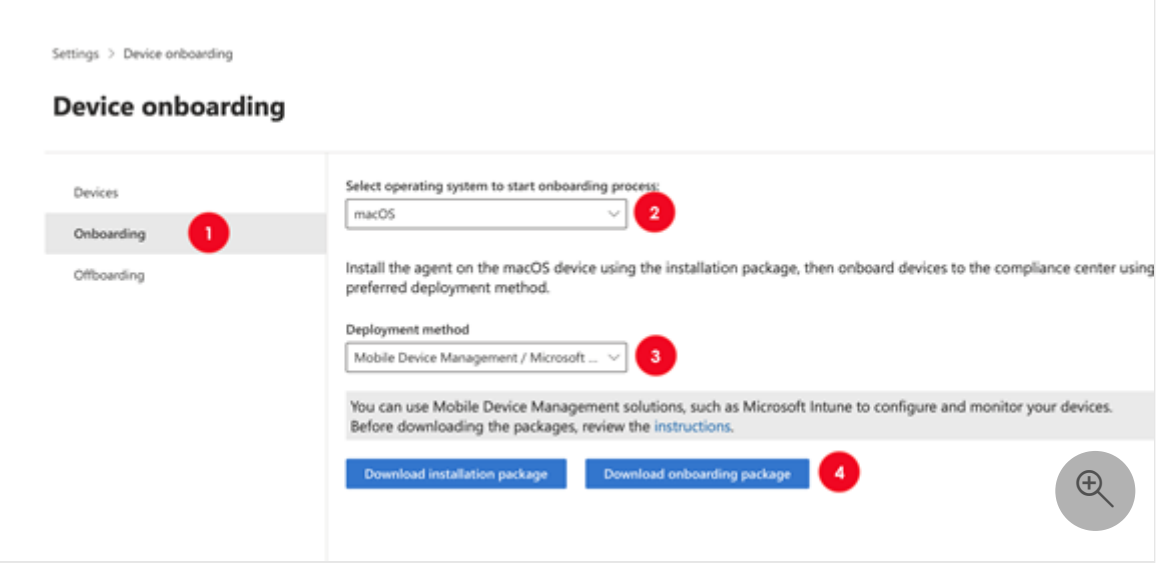
Download the following files:

Expand table

File	Description
mdatp.mobileconfig	System mobile config file
com.microsoft.wdav.mobileconfig.	MDE preferences

> **Tip**
>
> We recommend downloading the bundled **mdatp.mobileconfig** file, rather than the **individual** *.mobileconfig* files. The bundled file includes the following required files:
>
> - accessibility.mobileconfig
> - fulldisk.mobileconfig
> - netfilter.mobileconfig
> - sysext.mobileconfig
>
> If any of these files are updated, you need to either download the updated bundle, or download each updated file individually.

Get the device onboarding package



1. In the **Microsoft Purview portal** open **Settings** > **Device Onboarding** and then choose **Onboarding**.

2. For the **Select operating system to start onboarding process** option, choose **macOS**.
3. For **Deployment method**, choose **Mobile Device Management/Microsoft Intune**.
4. Choose **Download onboarding package**.
5. Extract the .ZIP file and open the *Intune* folder. This contains the onboarding code in the *DeviceComplianceOnboarding.xml* file.

Deploy the mobileconfig and onboarding packages

1. Open the **Microsoft Intune admin center** and navigate to **Devices** > **macOS** > **Configuration**.
2. Choose: **+ Create** and then choose **New policy**.
3. Select the following values:
   a. **Platform** = macOS
   b. **Profile type** = Templates
   c. **Template name** = Custom
4. Choose **Create**.
5. On the **Basics** page, Enter a name for the profile, such as *Microsoft Purview System MobileConfig*, and then Choose **Next**.
6. On the **Configuration settings** page, Choose the `mdatp.mobileconfig` file that you downloaded in the pre-requisites section as the configuration profile file.
7. Choose **Next**.
8. On the **Assignments** tab, add the group you want to deploy these configurations to and then choose **Next**.
9. Review your settings and then choose **Create** to deploy the configuration.
10. Repeat steps 2-9 to create profiles for the:
    a. **DeviceComplianceOnboarding.xml** file. Name it *Microsoft Purview Device Onboarding Package*
    b. **com.microsoft.wdav.mobileconfig** file. Name it *Microsoft Endpoint Device Preferences*
11. In the macOS **Configuration policies** page, the policies you created now display. Select **Refresh** if your changes don't show up right away.
12. Choose the profile that you just created. Next, view the **Device and user check-in status** report of the configuration policy.

Publish the application

Microsoft Endpoint data loss protection is installed as a component of Microsoft Defender for Endpoint on macOS. This procedure applies to onboarding devices into Microsoft Purview solutions

1. In the Microsoft Intune admin center, open **Apps**.
2. Select **macOS**.
3. On the **macOS** > **macOS apps** page, select **Create**.
4. On the **Select app type** > **Create app** slider, select the App type drop-down and then **macOS** under **Microsoft Defender for Endpoint**.
5. Choose the **Select** button.
6. On the **App Information** tab, keep the default values and then choose **Next**.
7. On the **Assignments** tab, add assignments and then choose **Next**.
8. You can visit **Apps** > **macOS** to see the new application listed.

Offboard macOS devices using Intune

> **Note**
>
> Offboarding causes the device to stop sending sensor data to the portal. However, data from the device, including reference to any alerts it has had, will be retained for up to six months.

1. In the **Microsoft Intune admin center**, open **Devices** > **macOS** > **Configuration**. The policies you created are listed.
2. Select **Microsoft Endpoint Device Preferences** policy.
3. Under **Properties** > **Assignments,** choose **Edit**.
4. Remove the group from the assignment. This will uninstall the com.microsoft.wdav.mobileconfig configuration profile file and offboard the macOS device from Purview solutions.

Onboard and offboard macOS devices into Purview solutions using Intune for Microsoft Defender for Endpoint customers

Summarize this article for me

You can use Microsoft Intune to onboard macOS devices into Microsoft Purview solutions.

> **Important**
>
> Use this procedure ***if you have already deployed*** Microsoft Defender for Endpoint (MDE) to your macOS devices.

Applies to:

- Customers who have MDE deployed to their macOS devices.
- Endpoint data loss prevention (DLP)
- Insider risk management

Before you begin

- Make sure your macOS devices are onboarded to Intune and enrolled in the Company Portal app.
- Make sure you have access to the Microsoft Intune admin center.
- OPTIONAL: Install the v95+ Microsoft Edge browser on your macOS devices.

> **Note**
>
> The three most recent major releases of macOS are supported.

Onboard macOS devices into Microsoft Purview solutions using Microsoft Intune

Given that Microsoft Defender for Endpoints (MDE) has already been deployed to your macOS devices, you can still onboard the devices into Purview solutions by competing the following steps:

1. Validate existing FDA system configuration profile

2. Update MDE Preferences

Validate existing system configuration profile

1. A full disk access (FDA) configuration profile should have been created and deployed previously for MDE. (For details, see Intune-based deployment for Microsoft Defender for Endpoint on Mac).

2. Endpoint data loss prevention (DLP) *requires additional FDA permission* for the new application (`com.microsoft.dlp.daemon`). Validate that the `com.microsoft.dlp.daemon` is in the fulldisk.mobileconfig file

Update MDE Preferences

1. Find the existing **MDE Preferences** configuration profile. See Intune-based deployment for Microsoft Defender for Endpoint on Mac for details.

2. Add the following key to the `com.microsoft.wdav` configuration profile file, then save the file.

```xml
<key>features</key>
<dict>
  <key>dataLossPrevention</key>
  <string>enabled</string>
</dict>
```

Offboard macOS devices using Microsoft Intune

> **Important**
>
> Offboarding causes the device to stop sending sensor data to the portal. However, data received from the device, including references to any alerts it has had, will be retained for up to six months.

1. In the **Microsoft Intune admin center**, open **Devices** > **Configuration**. The policies you created display.

2. Choose the **MDE Preferences** policy.

3. Under **Properties** > **Configuration Settings**, choose **Edit**

4. Remove these settings:

```xml
<key>features</key>
<dict>
    <key>dataLossPrevention</key>
    <string>enabled</string>
</dict>
```

5. Choose **Review+**Save**. This will offboard the macOS devices from Purview solutions while not changing the macOS device enrollment in MDE.

Last updated on 02/17/2026

Onboard and offboard macOS devices into Microsoft Purview solutions using JAMF Pro

09/26/2025

You can use JAMF Pro to onboard macOS devices into Microsoft Purview solutions such as Endpoint data loss prevention (DLP).

> **Important**
>
> Use this procedure if you ***do not*** have Microsoft Defender for Endpoint (MDE) deployed to your macOS devices.

Applies to:

- Endpoint data loss prevention (DLP)
- Insider risk management

Before you begin

- Make sure your macOS devices are managed through JAMF pro and are associated with an identity (Microsoft Entra joined UPN) through JAMF Connect.
- OPTIONAL: Install the v95+ Microsoft Edge browser on your macOS devices for native Endpoint DLP support on Microsoft Edge.

> **Note**
>
> The three most recent major releases of macOS are supported.

Onboard devices into Microsoft Purview solutions using JAMF Pro

Prerequisites

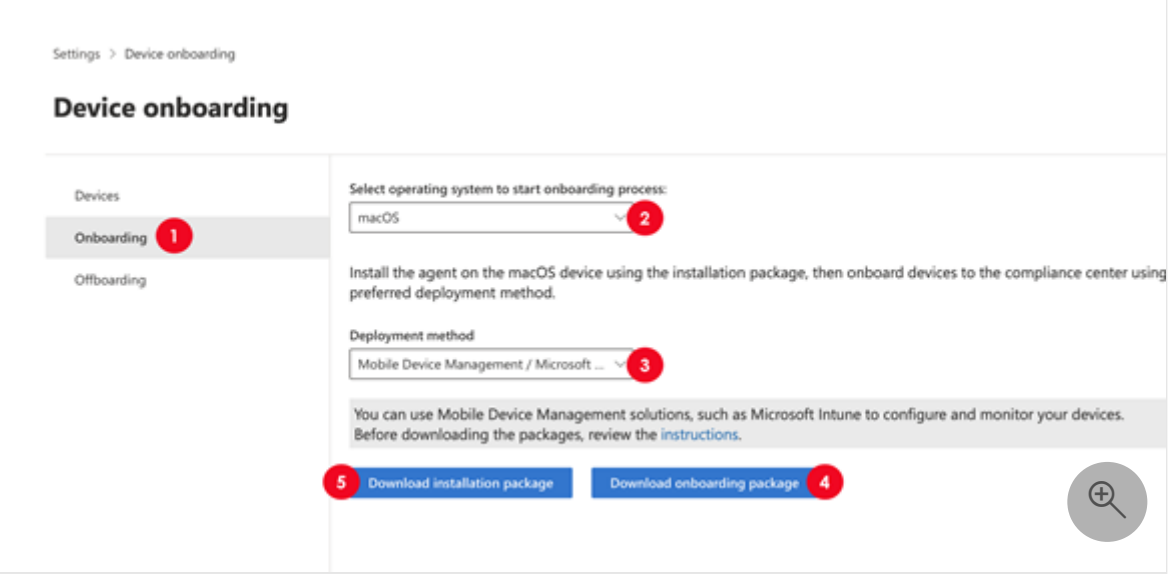
Download the following files.

Expand table

File	Description
mdatp.mobileconfig	This is the bundled file.
schema.json	This is the MDE preference file.

> **Tip**
>
> We recommend downloading the bundled **mdatp.mobileconfig** file, rather than the **individual** *.mobileconfig* files. The bundled file includes the following required files:
>
> - accessibility.mobileconfig
> - fulldisk.mobileconfig
> - netfilter.mobileconfig
> - sysext.mobileconfig
>
> If any of these files are updated, you need to either download the updated bundle, or download each updated file individually.

Get the device onboarding and installation packages



1. In the Purview portal, open **Settings** > **Device Onboarding** and then choose **Onboarding**.
2. For the **Select operating system to start onboarding process** value, choose **macOS**.
3. For **Deployment method**, choose **Mobile Device Management/Microsoft Intune**.

4. Choose **Download onboarding package** and then extract the contents of the device onboarding package. the *DeviceComplianceOnboarding.plist* file is downloaded to the JAMF folder.

5. Choose **Download installation package**.

Deploy onboarding packages

Create a new configuration profile in JAMF Pro. Use the following values:

- **Name:** *MDATP onboarding for macOS*
- **Description:** **MDATP EDR onboarding for macOS*
- **Category:** *none*
- **Distribution method:** *`*install automatically*
- **Level:** *computer level*

For help with performing steps within the Jamf Pro console, refer to the JAMF Pro documentation, or contact Jamf Pro support.

Configure application preferences

> **Important**
>
> You must use *com.microsoft.wdav* as the **Preference Domain** value. Microsoft Defender for Endpoint uses this name and *com.microsoft.wdav.ext* to load the managed settings.

For information about how to create a new configuration profile in JAMF Pro, refer to the JAMF Pro documentation. Or contact Jamf Pro support.

Use these values:

- **Name:** *MDATP MDAV configuration settings*
- **Description:** *Leave this blank*
- **Category:** *none*
- **Distribution method:** *install automatically*
- **Level:** *computer level*

Deploy system configuration profiles

For information about how to deploy system configuration profiles in JAMF Pro, refer to the JAMF Pro documentation, or contact Jamf Pro support.

Offboard macOS devices using JAMF Pro

> **Important**
>
> Offboarding causes the device to stop sending sensor data to the portal. However, data from the device, including references to any alerts it has had, will be retained for up to six months.

1. If you aren't using MDE, uninstall the application. See the **Package Deployment** section in the JAMF Pro documentation.

2. Restart the macOS device. (Some applications may lose printing functionality until they're restarted.)

Onboard and offboard macOS devices into Purview solutions using Jamf Pro for Microsoft Defender for Endpoint customers

09/26/2025

You can use Jamf Pro to onboard macOS devices into Microsoft Purview solutions.

> **Important**
>
> Use this procedure ***if you have*** deployed Microsoft Defender for Endpoint (MDE) to your macOS devices

Applies to:

- Customers who have MDE deployed to their macOS devices.
- Endpoint data loss prevention (DLP)
- Insider risk management

Before you begin

- Make sure your macOS devices are managed through Jamf pro and are associated with an identity (Microsoft Entra joined UPN) through Jamf Connect.
- OPTIONAL: Install the v95+ Edge browser on your macOS devices to have native Endpoint DLP support on Edge.

> **Note**
>
> The three most recent major releases of macOS are supported.

Onboard devices into Microsoft Purview solutions using Jamf Pro

Onboarding a macOS device into Purview solutions is a multi-phase process.

1. Update the existing MDE Preference domain profile using the Jamf Pro console
2. Enable full-disk access
3. Enable accessibility access to Microsoft Purview data loss prevention

4. Check the macOS device

Prerequisites

Download the following files:

Expand table

File	Description
accessibility.mobileconfig	Accessibility
fulldisk.mobileconfig	Full disk access (FDA)
schema.json	MDE preference

If any of these individual files are updated, you must download the updated bundled file and redeploy as described.

> **Tip**
>
> We recommend downloading the bundled **mdatp.mobileconfig** file, rather than the **individual** *.mobileconfig* files. The bundled file includes the following required files:
>
> - accessibility.mobileconfig
> - fulldisk.mobileconfig
> - netfilter.mobileconfig
> - sysext.mobileconfig
>
> If any of these files are updated, you need to either download the updated bundle, or download each updated file individually.

Update the existing MDE Preference domain profile using the Jamf Pro console

For information about how to replace the *schema.json* file in the MDE deployment with the updated version, and configure the settings in the Jamf Pro console, refer to the Jamf Pro documentation, or contact Jamf Pro support.

Enable full-disk access

To update the existing full disk access profile with the `fulldisk.mobileconfig` file, upload `fulldisk.mobileconfig` to Jamf. For more information, refer to Set up the Microsoft Defender for Endpoint on macOS policies in Jamf Pro.

Enable accessibility access to Microsoft Purview data loss prevention

To grant accessibility access to DLP, upload the `accessibility.mobileconfig` file you downloaded previously to Jamf, as described in Deploy system configuration profiles.

Check the macOS device

1. Restart the macOS device.
2. Verify that the following profiles are listed under Profiles in System Settings:

    - Accessibility
    - MAU
    - MDATP Onboarding
    - MDE Preferences
    - Management profile
    - Network filter
    - Notifications
    - System extension profile

Offboard macOS devices using Jamf Pro

> **Important**
>
> Offboarding causes the device to stop sending sensor data to the portal. However, data from the device, including references to any alerts it has had, will be retained for up to six months.

For information about how to offboard a macOS device using Jamf Pro, refer to the Jamf Pro documentation, or contact Jamf Pro support.

Any mobile device management tool for customers without Microsoft Defender for Endpoint

09/26/2025

You can use any mobile device management (MDM) solution to onboard macOS devices into Microsoft Purview solutions such as Endpoint data loss prevention (DLP).

> **Important**
>
> Use this procedure if you ***do not*** have Microsoft Defender for Endpoint (MDE) deployed to your macOS devices.

Applies to:

- Endpoint data loss prevention (DLP)

Before you begin

- OPTIONAL: Install the v95+ Microsoft Edge browser on your macOS devices for native Endpoint DLP support on Microsoft Edge.

> **Note**
>
> The three most recent major releases of macOS are supported.

Onboard devices into Microsoft Purview solutions using any MDM

Prerequisites

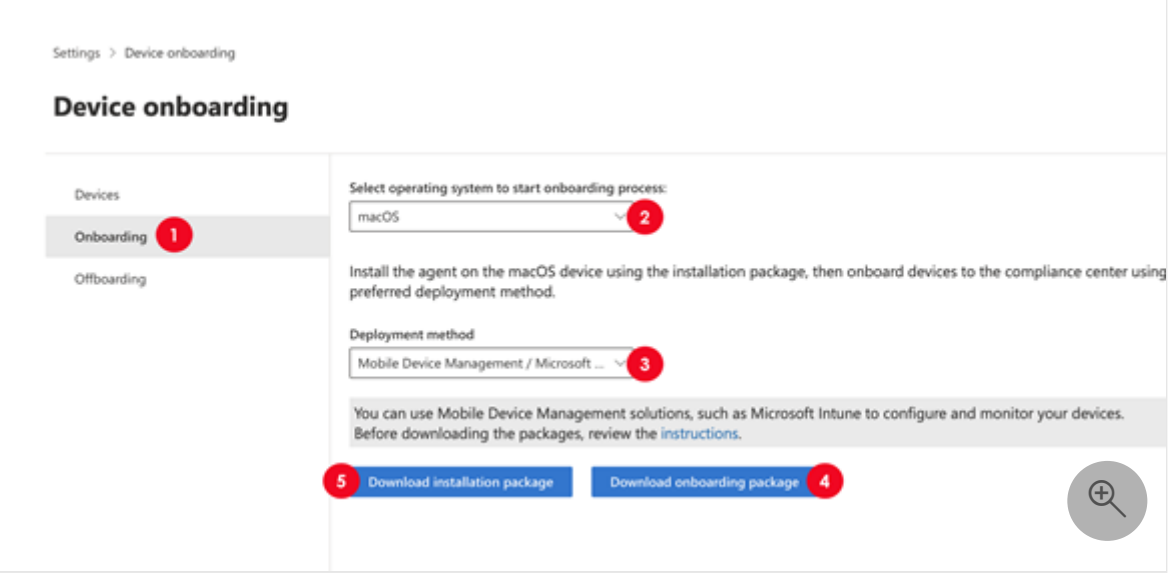
- You must have Microsoft Defender Antimalware client version 101.25012.0005 or newer installed on the device.
- Download the following files.

Expand table

File	Description
mdatp.mobileconfig	This is the bundled file.
schema.json	This is the MDE preference file.

> **Tip**
>
> We recommend downloading the bundled **mdatp.mobileconfig** file, rather than the **individual** *.mobileconfig* files. The bundled file includes the following required files:
>
> - accessibility.mobileconfig
> - fulldisk.mobileconfig
> - netfilter.mobileconfig
> - sysext.mobileconfig
>
> If any of these files are updated, you need to either download the updated bundle, or download each updated file individually.

Get the device onboarding and installation packages



1. In the Microsoft Purview portal, open **Settings** > **Device Onboarding** and then choose **Onboarding**.
2. For the **Select operating system to start onboarding process** value, choose **macOS**.
3. For **Deployment method**, choose **Mobile Device Management/Microsoft Intune**.
4. Choose **Download onboarding package** and then extract the contents of the device onboarding package. the *DeviceComplianceOnboarding.plist* file is downloaded to the

JAMF folder.

5. Choose **Download installation package**.

Deploy onboarding packages

1. Create a new configuration profile in your MDM. Use the following values (if applicable):

- **Name:** *MDATP onboarding for macOS*
- **Description:** **MDATP EDR onboarding for macOS*
- **Category:** *none*
- **Distribution method:** *`*install automatically*
- **Level:** *computer level*

1. Upload *DeviceComplianceOnboarding.plist* as a custom settings
2. Enter `com.microsoft.wdav.atp` as the **Preference Domain**,
3. Choose **Save**. Refer to your MDM documentation on creating configuration profile.

Configure application preferences

> **Important**
>
> You must use *com.microsoft.wdav* as the **Preference Domain** value. Microsoft Defender for Endpoint uses this name and *com.microsoft.wdav.ext* to load the managed settings.

For information about how to create a new configuration profile, refer to your MDM documentation.

1. Create a new configuration profile. Use these values (if applicable):

- **Name:** *MDATP MDAV configuration settings*
- **Description:** *Leave this blank*
- **Category:** *none*
- **Distribution method:** *install automatically*
- **Level:** *computer level*

2. Add the `schema.json` as a custom schema in the configuration profile.

3. Under **Preference Domain Properties**, locate **Data Loss Prevention** and select `enabled`.

4. Save the profile after adjusting other MDE settings (if necessary).

Deploy system configuration profiles

For information about how to deploy system configuration profiles, refer to your MDM documentation.

1. Create a new configuration profile by uploading the `mdatp.mobileconfig` file into your MDM.

Register Identity

Depending on your configuration, you register the Microsoft Entra identity of the using one of these two methods.

Method 1: Register without the Microsoft Intune Company Portal

1. Install any Microsoft application such as Microsoft Office, or Teams, or Microsoft Edge. When the user logs in using valid Microsoft Entra account, that account is registered as the UPN for DLP policy evaluation. Personal Microsoft accounts aren't considered as valid. The device may need to restart to complete the registration.

Method 2: Register with the Microsoft Intune Company Portal

This method is for users whose macOS devices don't use Microsoft apps.

You can Get the Microsoft Intune Company Portal if you don't already have it.

1. Deploy Microsoft Enterprise SSO plug-in for Apple devices to their managed macOS devices using your MDM solution. The Microsoft Intune Company Portal for macOS provides the SSO plugin for smooth user experience. So, the Microsoft Intune Company Portal must be installed on the macOS device, but the user doesn't need to be signed into the Intune Company portal. When the user signs-in to any non-Microsoft Authentication Library (MSAL) application, like Safari browser, the SSO plugin is enabled to register the UPN. The device may need to restart to complete the registration.

For information on troubleshooting the Microsoft Enterprise SSO Extension, go to Troubleshooting the Microsoft Enterprise SSO Extension plugin on Apple devices.

Offboard macOS devices using any MDM

> **Important**

Offboarding causes the device to stop sending sensor data to the portal. However, data from the device, including references to any alerts it has had, will be retained for up to six months.

1. If you aren't using MDE, uninstall the application. Refer to your MDM documentation for any MDM specific instructions for uninstallation.
2. Restart the macOS device. (Some applications may lose printing functionality until they're restarted.)

Any mobile device management tool for Microsoft Defender for Endpoint customers

09/26/2025

You can use any MDM to onboard macOS devices into Microsoft Purview solutions.

> **Important**
>
> Use this procedure ***if you have*** deployed Microsoft Defender for Endpoint (MDE) to your macOS devices

Applies to:

- Customers who have MDE deployed to their macOS devices.
- Endpoint data loss prevention (DLP)

Before you begin

- OPTIONAL: Install the v95+ Microsoft Edge browser on your macOS devices to have native Endpoint DLP support on Microsoft Edge.

> **Note**
>
> The three most recent major releases of macOS are supported.

Onboard devices into Microsoft Purview solutions using any MDM solution

Onboarding a macOS device into Purview solutions is a multi-phase process.

1. Update any existing MDE Preference domain profile by setting `com.microsoft.wdav` as Preference domain to enable Data Loss Prevention. Refer to update preference domain profile.
2. Enable full-disk access
3. Enable accessibility access to Microsoft Purview Data Loss Prevention
4. Register identity
5. Check the macOS device

Prerequisites

- Device must have Microsoft Defender Antimalware client version 101.25012.0005 or newer installed.
- Download the following files:

Expand table

File	Description
accessibility.mobileconfig	Accessibility
fulldisk.mobileconfig	Full disk access (FDA)
schema.json	MDE preference

If any of these individual files are updated, you must download the updated bundled file and redeploy as described.

> **Tip**
>
> We recommend downloading the bundled **mdatp.mobileconfig** file, rather than the **individual** *.mobileconfig* files. The bundled file includes the following required files:
>
> - accessibility.mobileconfig
> - fulldisk.mobileconfig
> - netfilter.mobileconfig
> - sysext.mobileconfig
>
> If any of these files are updated, you need to either download the updated bundle, or download each updated file individually.

Update the existing MDE Preference domain profile

1. Replace the *schema.json* file in the MDE deployment with the updated version that you just downloaded. For information about how to replace the *schema.json* file in the MDE deployment with the updated version, and configure the settings in your MDM solution, refer to your MDM solution documentation.

Enable full-disk access

To update the existing full disk access profile with the `fulldisk.mobileconfig` file, upload `fulldisk.mobileconfig` to your MDM solution to enable full disk access for Data Loss Prevention.

Enable accessibility access to Microsoft Purview Data Loss Prevention

To grant accessibility access to DLP, upload the `accessibility.mobileconfig` file you downloaded previously to your MDM solution. This should be similar to what is described in Deploy system configuration profiles.

Register Identity

Depending on your configuration, you'll register the Microsoft Entra identity of the using one of these two methods.

Method 1: Register without the Microsoft Intune Company Portal

1. Install any Microsoft application such as Microsoft Office, or Teams, or Microsoft Edge. When the user logs in using valid Microsoft Entra account, that account is registered as the UPN for DLP policy evaluation. Personal Microsoft accounts aren't considered as valid.

Method 2: Register with the Microsoft Intune Company Portal

This method is for users whose macOS devices don't use Microsoft apps.

You can Get the Microsoft Intune Company Portal if you don't already have it.

1. Deploy Microsoft Enterprise SSO plug-in for Apple devices to their managed macOS devices using your MDM solution. The Microsoft Intune Company Portal for macOS provides the SSO plugin for smooth user experience. So, the Microsoft Intune Company Portal must be installed on the macOS device, but the user doesn't need to be signed into the Intune Company portal. When the user signs-in to any non-Microsoft Authentication Library (MSAL) application, like Safari browser, the SSO plugin is enabled to register the UPN.

For information on troubleshooting the Microsoft Enterprise SSO Extension, go to Troubleshooting the Microsoft Enterprise SSO Extension plugin on Apple devices.

Check the macOS device

1. Restart the macOS device.
2. Verify that the following profiles are listed under Profiles in System Settings:

    - Accessibility
    - MAU
    - MDATP Onboarding
    - MDE Preferences
    - Management profile
    - Network filter
    - Notifications
    - System extension profile

Offboard macOS devices

> **ⓘ Important**
>
> Offboarding causes the device to stop sending sensor data to the portal. However, data from the device, including references to any alerts it has had, will be retained for up to six months.

For information about how to offboard a macOS using any MDM solution, refer to the documetentation of your MDM solution.

Security Copilot in Microsoft Purview Overview

Summarize this article for me

Microsoft Security Copilot is a cloud-based AI platform that can assist security and compliance professionals in protecting their organization's data. Security and compliance professionals can use Security Copilot to identify, summarize, triage, and remediate issues within the following Microsoft Purview solutions:

- Data Security Posture Agent (DSPM)
- Microsoft Purview Data Loss Prevention (DLP)
- Microsoft Purview Insider Risk Management
- Microsoft Purview Communication Compliance
- Microsoft Purview eDiscovery

For more information about what Security Copilot can do and the different scenarios it supports, read What is Microsoft Security Copilot?.

Know before you begin

If you're new to Security Copilot, you should familiarize yourself with it by reading these articles:

- What is Microsoft Security Copilot?
- Microsoft Security Copilot experiences
- Get started with Microsoft Security Copilot
- Understand authentication in Microsoft Security Copilot
- Prompting in Microsoft Security Copilot
- Configure Owner settings
- Learn about Security Copilot in Microsoft 365 E5

Security Copilot integration in Microsoft Purview

When you sign up for Security Copilot in the same tenant as Microsoft Purview, you can use both the Security Copilot embedded and standalone experiences. Security Copilot capabilities, like summarizing DLP or insider risk management alerts, are embedded into Microsoft Purview features.

Copilot in Microsoft Purview embedded experiences is a set of capabilities that are embedded in Microsoft Purview features. For more information, see standalone and embedded

experiences.

Copilot in Microsoft Purview standalone experience is a chat-like experience that you can use to ask questions and get answers about your data. For more information, see standalone and embedded experiences.

Key features in the embedded experience

You can open Security Copilot in Microsoft Purview by selecting the **Copilot** icon in the top navigation bar. It's available across all Microsoft Purview solutions.



The embedded experience in Purview can help you:

- *Get AI-generated summaries from Microsoft Purview product documentation*.
- *Use suggested or customized prompts in DSPM*. Learn more about the data security posture in your organization. For more information, see Use Microsoft Security Copilot with Data Security Posture Management.
- *Summarize alerts in DLP*. For more information on this and how to access Security Copilot in DLP, see Investigate a DLP alert.
- *Get insights into policies*. Security copilot can help you understand what your policies are doing in your organization, and where they're active. For more information, see Get Insights with Microsoft Security Copilot.
- *Summarize alerts in insider risk management*. For more information on this and how to access Security Copilot in Insider Risk Management, see Investigate insider risk management activities.
- *Summarize policy matches based on trainable classifiers for communication compliance*. Communication compliance also provides an interactive prompt experience to dig deeper into the summary. For more information on this and the steps to access Copilot in Communication Compliance, see Investigate and remediate communication compliance alerts.
- *Get contextual summary for eDiscovery cases*. For more information, see Group and view documents in a review set in eDiscovery (Premium).
- *Gain insights in activity explorer data and generate filters from natural language prompts (preview)*. For more information, see Microsoft Security Copilot in activity explorer (preview).

Key features in the standalone experience

The Security Copilot in Microsoft Purview standalone experience has many built in capabilities. You can use these capabilities to get insights from your Purview data and make connections between datapoints. This information can help you understand your information security and compliance posture and triage alerts.

- *DLP and Insider Risk Management data and user risk promptbook*. Security Copilot in Microsoft Purview analyzes data from DLP and insider risk management and, by running multiple, sequential prompts contained in a promptbook, presents you with integrated results. For more information, see Microsoft Copilot in Microsoft Purview prompts and promptbooks.

System capabilities of Security Copilot

In the standalone experience, there are built-in capabilities (prompts) that are available once the Microsoft Purview plugin is enabled.

Security Copilot in Microsoft Purview brings three types of capabilities:

- Summarize Microsoft Purview alerts.
- Triage Microsoft Purview alerts.
- Drill down into your Microsoft Purview data.

Enable the Microsoft Purview source in Microsoft Security Copilot

> **Important**
>
> Security Copilot in Purview must be enabled for both the standalone and embedded experiences to work.

Security Copilot in Purview is enabled by default. To enable or disable the Microsoft Purview source in Microsoft Security Copilot, follow these steps:

1. Ensure that you have permissions.
2. Open https://securitycopilot.microsoft.com/.
3. Open **Sources** in the prompt bar.



4. On the **Manage plugins** page, set the **Purview** toggle to **On** to enable or **Off** to disable.

Review the Microsoft Purview system capabilities

1. Select the capabilities control in the prompt bar.



2. Select **See all system capabilities** to see all the system capabilities that are available for Microsoft Purview. Here are a few:

    - Get Data Risk Summary
    - Get User Risk Summary
    - Summarize Purview Alert
    - Triage Purview Alerts
    - Zoom into Purview Data and User Risk

Sample prompts

For guidance on writing effective prompts, see Prompting in Microsoft Security Copilot. Here are some examples:

- *Show me the top five DLP alerts from the past 24 hours.*
- *Summarize the DLP alert with ID <12345>.*
- *What's the risk profile of the user that's associated with the DLP alert <12345>.*
- *Show me the Insider Risk Management alerts for <user> .*
- *What items did user <user> exfiltrate in the past 30 days.*

Provide feedback

Your feedback is vital to guide the current and planned development of the product. The best way to provide this feedback is directly in the product. Select **How's this response?** at the bottom of each completed prompt and choose any of the following options:

- **Looks right** - Select if the results are accurate, based on your assessment.
- **Needs improvement** - Select if any detail in the results is incorrect or incomplete, based on your assessment.
- **Inappropriate** - Select if the results contain questionable, ambiguous, or potentially harmful information.

For each feedback option, you can provide more information in the next dialog box that appears. Whenever possible, and especially when the result is **Needs improvement**, write a few words explaining what can be done to improve the outcome. If you entered prompts specific to Microsoft Purview and the results aren't related, then include that information.

Privacy and data security in Microsoft Security Copilot

To understand how Microsoft Security Copilot in Purview handles your prompts and the data that's retrieved from the service (prompt output), see the Privacy and data security in Microsoft Security Copilot.

Related articles

- Privacy and data security in Microsoft Security Copilot
- Microsoft Security Copilot Responsible AI FAQ

Last updated on 02/24/2026

Microsoft Security Copilot in Microsoft Purview prompts and promptbooks

Summarize this article for me

Microsoft Security Copilot is a stand-alone chat-based experience that helps you get insights from your security data and make connections between datapoints. You interact with the stand-alone experience by entering prompts, and Security Copilot returns results. This article focuses on the stand-alone experience and how to use it to get insights from your Microsoft Purview data.

Know before you begin

If you're new to Microsoft Security Copilot, read these articles before you begin creating your own prompts and promptbooks:

- What is Microsoft Security Copilot?
- Microsoft Security Copilot experiences
- Get started with Microsoft Security Copilot
- Microsoft Security Copilot in Microsoft Purview
- Understand authentication in Microsoft Security Copilot
- Prompting in Microsoft Security Copilot

> **Important**
>
> You must enable the Security Copilot in Purview plugin to use the Microsoft Purview prompts and promptbooks. Use the steps in **Enable the Microsoft Purview source in Microsoft Security Copilot**

Prompting in Security Copilot in Microsoft Purview

You can use prompts in three different ways:

Open ended

Open ended prompts are ones that you write yourself without any suggestions from Security Copilot. You can use these prompts to ask questions about your data that are specific to your needs at a certain point in time or when you need a custom prompt that isn't covered by the

prompt suggestions or a promptbook. For more information on creating solid prompts, see Create effective prompts.

Prompt suggestions

Prompt suggestions are prewritten prompts that you can use as a starting point and customize to fit your needs. They're listed under all system capabilities. Once you enable the Purview add-in, these prompt suggestions are available to you.

Get Data Risk Summary fetches information from Microsoft Purview Information Protection and Microsoft Purview Data Loss Prevention (DLP), and summarizes the risk associated with data impacted by a security incident or a DLP alert.

Get User Risk Summary summarizes the risk associated with any user by using their risk profile from Microsoft Purview Insider Risk Management.

Summarize Purview Alert provides more details about a Purview alert, including a DLP alert and an insider risk management alert.

Triage Purview Alerts retrieves the top or most recent DLP alerts, organizing DLP alerts based on user preferences.

Zoom Into Purview Data Risk fetches the information from Microsoft Purview Information Protection and DLP and helps identify the risks and attributes related to the data.

Zoom Into Purview User Risk fetches information about the user's activities, including operations and actions done by the user, user activity over a time span, data leakage or exfiltration activities by the user, sequential activities by the user, or any signs of anomalous or unusual behavior by the user.

Promptbooks

A promptbook is a collection of prompts that are in sequence. The stand-alone experience can use the output from one prompt as the input for the next prompt, and it can run multiple prompts in a sequence. As you use Security Copilot with the Microsoft Purview add-in enabled, you may find that you keep using the same prompts over and over. You can save time by collecting all these prompts into a promptbook. Then, you can run the promptbook to get all the insights you need in one go. You can limit the use of a promptbook to yourself, or you can share it with others in your organization.

This article provides two sample sets of prompts you can use for investigating Microsoft Purview Data Loss Prevention (DLP) and Microsoft Purview Insider Risk Management data. You

can start with these prompts, then customize them to fit your specific needs. Once you have them tuned the way you want, you can create promptbooks from them.

Run a prompt

1. Go to Microsoft Security Copilot and sign in with your credentials.
2. By default, the **Purview** plugin should be enabled. To confirm, select **Sources**. In **Plugins**, confirm Microsoft Purview is on. Close **Sources**.
3. Navigate to the prompt bar.
4. Enter your prompt.

For example, enter:

- *Show me the five most recent high severity DLP alerts* or
- *Show me the risk associated with user user@contoso.com*.

For more information on creating solid prompts, see Create effective prompts

Promptbooks for Microsoft Purview features

Here are two sequences of prompts that you can use to investigate Microsoft Purview Data Loss Prevention (DLP) and Microsoft Purview Insider Risk Management data. Refer to Build your own promptbooks for the procedures on how to create, share, and edit a promptbook.

Sample prompt sequences

Microsoft Purview Data Loss Prevention

1. Show me the top five high severity DLP alerts.
2. Summarize the first alert in the list.
3. What actions were performed on the file in the alert?
4. Who is the user related to the alert?
5. What are the data risks related to the alert?
6. What is the risk summary of the user?

Microsoft Purview Insider Risk Management

1. Show me the risk associated with user *user@contoso.com*
2. Show me the exfiltration activities for this user.
3. Show me the sequential activities for this user.

4. Show me obfuscation activities for this user.
5. Show me all the activities that this user performed over the past 30 days.

Related articles

- Privacy and data security in Microsoft Security Copilot
- Responsible AI overview

Last updated on 02/24/2026

Security Copilot agents in Microsoft Purview overview

Microsoft Security Copilot agents are AI-powered processes that help you with specific role-based tasks. Microsoft Purview offers:

- Microsoft Purview Data Loss Prevention (DLP) triage agent
- Microsoft Purview Insider Risk Management triage agent
- Data Security Posture Management posture agent in preview
- Data Security Investigations posture agent in preview

Data Security Triage Agent

The triage agent provides a managed alert queue where it identifies and prioritizes the highest risk activities. The agent analyzes the content and potential intent involved in the activity based on the organization's chosen parameters and level of risk tolerance. The agent offers a comprehensive explanation for the logic behind the categorization.

You can find these agents in the **Microsoft Purview embedded experiences**. For more information, see embedded experiences.

Triaging and assigning a priority to alerts can be complex and time consuming. When you set an agent to triage and prioritize alerts, it reduces the amount of time required to complete the task. The agent helps you focus on the most important alerts by sifting them out from the noise of lower risk alerts. This prioritization improves your response time and helps increase the efficiency and effectiveness of your team.

For information on deploying, configuring, and using the triage agents, see:

- Get started with the Microsoft Purview Triage Agent in Data Loss Prevention
- Get started with the Microsoft Purview Triage Agent in Insider Risk Management

Data Security Posture Agent

Searching for relevant data and identifying its protection is a time-consuming, tedious, and highly manual task. The agent's primary role is to help you discover sensitive data across your data estate by using a simple natural language-based search. Instead of relying on keyword- or filter-based search tools, it uses large language models (LLMs) to understand user intent. It searches scoped content across your Microsoft 365 ecosystem—including documents, emails, messages, and Copilot interactions—and returns results quickly. The experience includes a concise summary and an associated risk analysis generated by the LLM.

In Data Security Investigations, the Posture agent extends these capabilities to help you proactively discover exposed credentials at scale across your tenant. The agent automates credential scanning across Microsoft 365 data locations, produces AI-generated risk assessments, and tracks tasks on a Kanban-style task board for streamlined review and remediation. For more information, see Get started with the Data Security Investigations Posture Agent (preview).

For information on deploying, configuring, and using the posture agents, see:

- Get started with the Posture Agent in Data Security Posture Management
- Get started with the Posture agent (preview) in Data Security Investigations

Before you begin

If you're new to Security Copilot or Security Copilot agents, familiarize yourself with the information in these articles:

- Microsoft Security Copilot agents overview
- What is Microsoft Security Copilot?
- Microsoft Security Copilot experiences
- Get started with Microsoft Security Copilot
- Understand authentication in Microsoft Security Copilot
- Prompting in Microsoft Security Copilot
- Configure Owner settings
- Learn about Security Copilot in Microsoft 365 E5

Security Copilot agent concepts

Microsoft Purview agents run on Security Compute Units (SCU). Your organization must have SCUs provisioned for the agents to run. For more information, see:

- Get started with the Microsoft Purview Triage Agent in Data Loss Prevention
- Get started with the Microsoft Purview Triage Agent in Insider Risk Management
- Get started with the Microsoft Purview Posture Agent in Data Security Posture Management
- Get started with the Data Security Investigations Posture Agent (preview)

The following sections apply only to the Triage agent. They don't apply to the Posture Agent.

Triggers

Triggers are groupings of parameters whose values must be met for the agent to triage any given alert. Triggers include:

- Time frame: You can define the time scope that alerts are generated in for triaging. See, Select Alert timeframe.
- Custom Instructions: You can define certain instructions for the agent to learn and prioritize from.

Run automatically or manually

When you deploy an agent, and when you edit triggers, you can select whether the agent runs **automatically based on a set schedule** or **Agent runs manually on one alert at a time** . If you select **Run automatically based on a set schedule**, the agent triages the alerts that are included in the **Select Alert timeframe** setting.

Select alert timeframe

When you deploy an agent and edit an agent's triggers, you can pick the timeframe the agent uses to decide which alerts to triage. The options are:

- Only triage new alerts
- Last 24 hours
- Last 48 hours
- Last 72 hours
- Last 7 days
- Last 14 days
- Last 21 days
- Last 30 days

If you select **Only triage new alerts**, the agent only triages alerts that are generated after the agent is deployed. The agent doesn't triage any alerts that were generated before the agent was deployed, which means that all the *Last # hours or days* options are ignored.

If you select any of the *Last # hours or days* options, the agent triages alerts that were generated in the selected timeframe. The selection lets you triage all alerts generated before the agent was deployed. The agent also triages all newly generated alerts.

> **ⓘ Important**
>
> The timeframe scope for alerts to be triaged is anchored to the moment of successful agent enablement. Essentially, the clock starts ticking when you enable the agent. So, *Last*

number of hours or days refers to the time period before agent deployment. ***This timeframe isn't rolling***.

Agent identity

You can set up agents, including the Posture agent, by using two options:

- **Create and use agent identity (recommended):** Create a separate Microsoft Entra agent identity for the agent. Associate all agent activities with this agent identity. To learn more, see Microsoft Entra agent identity and Microsoft Entra agent identities for AI agents.
- **Assign and use my identity:** Set up the agent by using the identity of the user or admin who sets up the agent. Associate all activities with the user's identity.

> **Important**
>
> Agents aren't administrative unit aware. But if the agent runs in the context of an administrative-unit restricted admin, and there are policies that are administrative-unit scoped to that admin, the agent only sees alerts from the policies that scoped to the admin unit.

Custom instructions

During the agent configuration process, you can give the agent *custom instructions*. (Custom instructions only apply to the Triage agent in DLP). Microsoft Purview agents use these natural language instructions to enhance alert triage by:

- Translating your input into structured classification logic.
- Running this logic against the document content associated with each alert.
- Raising the alert's priority if the content matches your custom instruction.

For custom instructions, the agent only analyzes document content, not metadata or behavioral attributes. This means the agent supports content categories like:

- Tax, financial, or legal information.
- Named entities, like credit card numbers, Social Security numbers, and names.
- Logical conditions like *more than five SSNs, contains financial documents*.

If you're unsure whether the condition you want to use is supported, check whether the condition is marked as a content condition. If it is, the agent can use it in custom instructions.

For example, if you enter: *I want to focus on alerts with content that is tax or finance related and contains more than five credit card numbers or SSNs*. The agent interprets this condition as:

JSON

```
{
  "logic_expression": "MATCH(content, 'Tax') or MATCH(content, 'Finance') and
(COUNT(content, 'U.S. social security number') >= 5) or (COUNT(content, 'U.S.
credit card number') >= 5)"
}
```

Triaged alerts

The agent triages alerts generated in the timeframe and using policies you select. The agent doesn't triage all alerts.

The agent groups triaged alerts into four categories:

All: This category includes all the alerts that the agent triaged. The count indicated in the category might not accurately reflect the true number of alerts until you go into that view and scroll down to load all the alerts. If the conditions that caused the alert to be raised change, or if the alert isn't triaged yet, you can select the alert and then select **Run agent** to manually run the agent on the alert.

Needs attention: This category contains alerts that the agent evaluated and determined pose the greatest risk to your organization. When you select one of these alerts, the details flyout opens to show a summary of the alert and other details.

Less Urgent: This category contains alerts that the agent evaluated and determined to pose a lower risk to your organization. When you select one of these alerts, the details flyout opens to show a summary of the alert and other details.

Not categorized: This category contains alerts that the agent wasn't able to successfully triage. This condition can happen for many reasons, including:

- Server error
- In process of reviewing
- Other error
- Unsupported error for alerts that contain activities that the agent doesn't support.

Agents triage files up to 2 MB in size.

How agents prioritize

The Triage Agent in DLP prioritizes alerts based on these risk factors:

- Content Risk: This category is the primary risk factor used during agent triage. It covers sensitive content based on Microsoft provided sensitive information types (SIT), trainable classifiers, and default sensitivity labels. For more information, see default sensitivity labels.
- Exfiltration Risk: Exfiltration of sensitive data shared externally.
- Policy Risk: Policy mode and rules with actions affect the prioritization of alerts.
- Content Risk: Label removed or downgraded.

The Triage Agent in Insider Risk Management prioritizes alerts based on:

- Activity risk: The agent identifies activities with the highest risk of exfiltration and reports historical alert insights.
- User risk: Attributes of the user that can affect the prioritization of alerts such as priority user group configuration or number of currently active cases.

Alert triage details

> **Important**
>
> The Triage agent in DLP only supports alerts from policies that are in active mode. It doesn't triage alerts from DLP policies that are running in simulation mode.

Agents can review alerts that were generated up to 30 days before enablement if the tenant has sufficient SCUs. Alerts that were generated more than 30 days before agent enablement are out of scope.

The Triage agent in DLP triages alerts from Exchange, Devices, SharePoint, OneDrive, and Teams. To triage alerts from Devices, you must enable evidence collection for file activities on devices.

In DLP, the agent doesn't triage alerts by non-SITs or non-trainable classifier policy conditions only, such as `Email subject match`.

You should do manual analysis on alerts that the agent doesn't evaluate.

Partially triaged alerts

Here are some examples of situations where you might partially triage alerts.

- DLP rule contains some conditions that aren't supported, such as `The user accessed a sensitive site from Edge`.
- The DLP rule includes certain conditions, but the system can't retrieve the corresponding properties of the email or files, such as `Document couldn't be scanned`.
- User not found.
- During preview, the insider risk management agent analyzes only SharePoint file content. It doesn't analyze email and device activities with accompanying files. If an alert contains only email or device activities, the agent doesn't analyze it.
- An alert generated from a policy that looks only at endpoint or email activities for Insider Risk Management.

Content analysis

There are some situations where content analysis might be limited.

The content risk prioritization of an alert is based on Sensitive Information Types (SITs), trainable classifiers, and sensitivity labels in content. When an agent evaluates content risk, it only looks for SITs and trainable classifiers that are defined in the policy.

When a DLP alert is associated with fewer than 10 files, the agent scans all the files and uses them to generate the content summary. When an alert includes 10 or more files, the agent uses the potentially top 10 files to generate the file risk summary. In DLP, the triage agent selects the top 10 risky files based on the number of policy classifier hits, the file size, and the most recent file access time. Once this selection is done, the triage agent adds a note stating that the content summary doesn't include all the files in the alert.

Related articles

- Security Copilot in Microsoft Purview Overview
- Get started with the Microsoft Purview Triage Agent in Data Loss Prevention
- Get started with the Microsoft Purview Triage Agent in Insider Risk Management
- Get started with the Microsoft Purview Posture Agent in Data Security Posture Management
- Get started with the Data Security Investigations Posture Agent (preview)

Last updated on 03/23/2026

Get started with the Microsoft Purview Triage Agent in Insider Risk Management

Use these steps to deploy the Microsoft Purview triage agent in Microsoft Purview Insider Risk Management. Before you begin, make sure you meet the licensing, permissions, and infrastructure prerequisites described in this article.

Before you begin

If you're new to Microsoft Purview triage agent in Insider Risk Management, you should read this article.

- Security Copilot Agents in Microsoft Purview Overview (preview)

SKU/subscriptions licensing

The Triage Agent in Insider Risk Management requires both the standard per-seat licensing model and the pay-as-you-go billing model. For more information, see Purview billing models. Your organization must be licensed for:

- Microsoft Purview Insider Risk Management to use the Insider Risk Management triage agent.

Microsoft Purview triage agents consume security compute units (SCUs) as they perform their tasks. You must have SCUs provisioned for the triage agents to work. The number of SCUs consumed depends on the number and type of alerts that are processed. For more information about SCUs, see Security compute units (SCUs). You can track your SCU consumption in the Security Copilot usage monitoring tool. For more information about onboarding into Microsoft Security Copilot, see Get started with Microsoft Security Copilot.

For information on Security Copilot licensing in E5 see, Learn about Security Copilot in Microsoft 365 E5

For information on licensing, see

- Microsoft 365 Enterprise Plans
- Microsoft 365 Service Descriptions

Permissions and roles

There are different permissions and roles needed to perform different functions. For more information, see Permissions in the Microsoft Purview portal, and Roles and role groups in the Microsoft Purview portals.

Assign these roles for setting up the agent using a recommended agent identity, actions like configuring or customizing, or deactivating and removing agents:

For admins (any one from each group):

Group A

Expand table

Role	Role Groups containing the role
Role Management	Purview Administrators Organization Management
Insider Risk Management Admin	Compliance Administrator Insider Risk Management Insider Risk Management Admins Organization Management

Group B

Expand table

Role	Role Groups containing the role
Security Copilot Contributor	(Managed in Security Copilot)
Purview Copilot Workspace Contributor	Included in multiple role groups. Some examples Organization Management Compliance Administrator Security Administrator Compliance Data Administrator Insider Risk Management Insider Risk Management Admins Insider Risk Management Analysts Insider Risk Management Investigators

For analysts (any one from each group):

Group A

Expand table

Role	Role Groups containing the role
Insider Risk Management Analysis	Insider Risk Management Insider Risk Management Analysts
Insider Risk Management Investigation	Insider Risk Management Insider Risk Management Investigators

Group B

Expand table

Role	Role Groups containing the role
Purview Agent Deployment	Included in multiple role groups. Some examples Compliance Administrator Data Security Management Insider Risk Management Insider Risk Management Analysts Insider Risk Management Investigators Purview Agent Management

Group C

Expand table

Role	Role Groups containing the role
Security Copilot Contributor	(Managed in Security Copilot)
Purview Copilot Workspace Contributor	Included in multiple role groups. Some examples Organization Management Compliance Administrator Security Administrator Compliance Data Administrator Insider Risk Management Insider Risk Management Admins Insider Risk Management Analysts Insider Risk Management Investigators

Important

Agents run in the security context of the last user that saved the agent configuration. This authentication is good for 90 days. After 90 days, the agent will stop running until the

configuration is manually saved again.

Permissions for viewing triaged Insider Risk Management alerts

The account that you use to view the Insider Risk Management triage agent alert activity must be assigned the roles (any one from each group):

Group A

Expand table

Role	Role Groups containing the role
Insider Risk Management Analysis	Insider Risk Management Insider Risk Management Analysts
Insider Risk Management Investigation	Insider Risk Management Insider Risk Management Investigators

Group B

Expand table

Role	Role Groups containing the role
Purview Agent Analysis	Included in multiple role groups. Some examples Compliance Administrator Data Security Management Insider Risk Management Insider Risk Management Analysts Insider Risk Management Investigators

Group C

Expand table

Role	Role Groups containing the role
Security Copilot Contributor	(Managed in Security Copilot)
Purview Copilot Workspace Contributor	Included in multiple role groups. Some examples Organization Management Compliance Administrator Security Administrator Compliance Data Administrator Insider Risk Management

Role	Role Groups containing the role
	Insider Risk Management Admins Insider Risk Management Analysts Insider Risk Management Investigators

Deployment and configuration roadmap

Deploying the Microsoft Purview Insider Risk Management triage agent involves these phases:

1. Infrastructure prerequisites
2. Enabling agents
3. Set up agents
4. Deactivate agents
5. Remove agents

Infrastructure prerequisites

Microsoft Purview triage agents run on Microsoft Security Copilot.

1. Your tenant must be onboarded to Microsoft Security Copilot. For more information on how to onboard, see Get started with Microsoft Security Copilot.
2. You must enable Microsoft 365 data sharing in Security Copilot. For more information, see Accessing data from Microsoft 365 services .
3. You must enable the Microsoft Purview plug-in in Microsoft Security Copilot. For more information, see Enable the Microsoft Purview source in Microsoft Security Copilot.

Enabling agents

This procedure is for organizations that haven't enabled any of the Microsoft Purview agents or have previously removed the agents and want to enable them again. Once you enable the agents, they're available for use in Microsoft Purview. There can be only one instance of each agent in a tenant. This procedure works for both the Purview DLP triage agent and the Insider Risk Management triage agent.

1. Sign in to the Microsoft Purview portal with an account that has the required Purview agent permissions.
2. In the left hand navigation pane, select **Agents**.
3. Select **Explore agents**.
4. Select the agent that you want to enable and select **Add**. This opens a page that shows the requirements to enable the agent.

5. Select **Setup**, this opens the **Deploy agent** global configuration page. You can:
    a. Choose to **Run automatically based on a set schedule**. If you don't choose this option, you must run the agent manually one at a time. The scheduled is set by Microsoft and isn't configurable by organizations. You can change this setting later when you edit the agent.
    b. **Select the alert timeframe**, which is how far back the agent looks for alerts to triage. Analysts can shorten the timeframe when they edit the agent but not lengthen it. For more information, see Select Alert timeframe.
6. Select **Deploy**. You see the *Alert Triage Agent in <solution> is deployed* message when the agent is successfully deployed.

Set up agents

Once an agent is enabled, you need to set specific triggers for the agent. The triggers are used to determine which alerts the agent triages. You can do this either in the **Agents** page or, for first run experience on the **Alerts** page for the solution. For this procedure, we'll use the first run experience **Alerts** page method. This procedure assumes that you still have the Microsoft Purview portal open to the **Explore agents** page from the Enabling agents procedure.

> **Important**
>
> The most recent agent configuration is always used.

1. Select **Go to <solution>**. This opens the **Alerts** page for the solution.
2. Because you are in the first run experience, you see dialog box that with a **Customize** button which, when selected, opens the **Customize Alert Triage Agent** flyout.
3. Here, choose either to accept the default global setting for **Select alert timeframe** or change it to be shorter than what was configured during agent deployment.
4. Enter custom instructions for the agent. The agent interprets your natural language input and uses it to better identifyw hich types of alerts matter most to you. This helps you identify and respond faster to the most relevant alerts.
5. Choose **Select policies** to select the policies whose alerts will be triaged by the agent. In preview, all policies are selected by default, you can change that here.
6. Select **Review**.
7. Select **Start agent**. Allow up to 2 hours for the agent to complete triaging the in scope alerts and enabled manual runs from this initial setup.

Deactivate agents

Use the following steps to deactivate an agent without removing it.

1. Sign in to the Microsoft Purview portal with an account that has the required Purview agent permissions.
2. In the left hand navigation pane, select **Agents**.
3. Select **Explore agents**.
4. Select **View agent** for the agent you want to deactivate. This opens the agent overview page.
5. On the far right upper right hand corner of the agent overview page, select the ellipses (three dots) that are located *next to* the **Edit agent** button.
6. Select **Deactivate agent**.

Deactivating the agent stops it from triaging alerts. It doesn't remove the agent and it doesn't reset the **Select alert timeframe** reference point in time.

Remove agents

Use the following steps to permanently remove an agent from Microsoft Purview.

1. Sign in to the Microsoft Purview portal with an account that has the required Purview agent permissions.
2. In the left hand navigation pane, select **Agents**.
3. Select **Explore agents**.
4. Select **View agent** for the agent you want to pause. This opens the agent overview page.
5. On the far right upper right hand corner of the agent overview page, select the ellipses (three dots) that are located *next to* the **Edit agent** button.
6. Select **Remove agent**.

> ⚠ **Warning**
>
> Removing the agent deletes it from Microsoft Purview and resets the **Select Alert timeframe** reference point in time. If you want to use the agent again, you must re-enable and set it up again.

Edit agents

Use the following steps to edit an existing agent's configuration and triggers.

1. Sign in to the Microsoft Purview portal with an account that has the required Purview agent permissions.
2. In the left hand navigation pane, select **Agents**.
3. Select **Explore agents**.
4. Select **View agent** for agent you want to edit. This opens the agent overview page.
5. Select **Edit agent**. This opens the **Edit agent** page.
6. Select **Triggers**.
7. Here you can change when the agent runs, either **Agent will run manually on one alert at a time** or **Agent will run automatically based on a set schedule**.
8. Select **Edit** to change the **Select alert timeframe** value and the policies that the agent will triage alerts from.
9. If you select **Agent will run manually on one alert at a time**, you can select a single alert in the **Alerts** page for the solution. Set the toggle to **Alert Triage Agent** and select **Run Agent**.
10. You can also change the **Intent** if you want.

Monitoring SCU usage

Use the following steps to view security compute unit (SCU) usage for an agent.

1. Sign in to the Microsoft Purview portal with an account that has the required Purview agent permissions.
2. In the left hand navigation pane, select **Agents**.
3. Select **Explore agents**.
4. Select **View agent** for agent you want to edit. This opens the agent overview page.
5. Select **Edit agent**. This opens the **Edit agent** page.
6. Select **Usage monitoring**.
7. You can track your SCU consumption in the Security Copilot usage monitoring tool.

Alerts page overview

Use the following steps to open the Alerts page and view alerts triaged by the agent.

1. Sign in to the Microsoft Purview portal with an account that has the required Purview agent permissions.
2. Open the solution you want to view the triaged alerts for.
3. Open the **Alerts** page for the solution.

4. In the top right hand area of the page, there's a new toggle that lets you choose between the **Standard** view of the alerts page and a **Triage Agent** view of the alerts page. Set the toggle to the **Triage Agent** view. This view shows the alerts that have been triaged by the agent. The alerts are grouped by the agent into four categories:

    - **All**
    - **Needs attention**
    - **Less Urgent**
    - **Not categorized**

Next steps

For more information, see these articles:

- Learn about investigating data loss prevention alerts
- Get started with the data loss prevention Alerts dashboard
- Triage Agent in Insider Risk Management dashboard
- Indicators and the Alert Triage Agent for Insider Risk Management

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Get started with the Microsoft Purview Triage Agent in Data Loss Prevention

Use these steps to deploy the Microsoft Purview triage agent in Data Loss Prevention (DLP) in both the Microsoft Purview and Microsoft Defender XDR portals.

The Microsoft Purview Triage Agent in DLP triages alerts from policies that are scoped to Exchange, Teams, OneDrive, SharePoint, and devices (Endpoint) locations, *only*. If you want to triage alerts from devices, Evidence collection for file activities on devices must be enabled.

Before you begin

If you're new to Microsoft Security Copilot Agent in Microsoft, read this article.

- Security Copilot agents in Microsoft Purview overview

Subscriptions and licensing

This agent requires both the standard per seat licensing model and the pay-as-you-go billing model. Your organization must be licensed for:

- Microsoft Data Loss Prevention (DLP) to use the DLP triage agent.

The agent consumes security compute units (SCUs) as it performs its tasks. You must have SCUs provisioned for the triage agent to work. The number of SCUs consumed depends on the number and type of alerts that are processed. For more information about SCUs, see Security compute units (SCUs). You can track your SCU consumption in the usage monitoring tool. For more information about onboarding into Microsoft Security Copilot, see Get started with Microsoft Security Copilot.

For information on Security Copilot licensing in E5 see, Learn about Security Copilot in Microsoft 365 E5.

For information on licensing, see

- Microsoft 365 Enterprise Plans
- Microsoft 365 Service Descriptions

Permissions and Roles

There are different permissions and roles needed to perform different functions with the agent. For more information, Permissions and roles needed are specific to the agent's various functions. For more information, see Permissions in the Microsoft Purview portal, and Roles and role groups in the Microsoft Purview portals.

> **Important**
>
> An agent now has its own agent identity (read more: https://learn.microsoft.com/en-us/entra/agent-id/identity-platform/what-is-agent-id-platform) or can be set up by assigning the user's identity who is setting the agent. These settings can be changed anytime by visiting the deployment configurations tab for the agent in Explore agents page. Regarding an agent that that was set up using a user's identity, migration to the agent identity in agent settings is recommended.

Permissions for enabling and configuring the Triage Agent in DLP from Microsoft Purview and Defender XDR portals

Assign these roles for setting up the agent using an recommended agent identity, actions like configuring or customizing, or deactivating and removing agents:

For admins (any one from each group):

Group A

Expand table

Role	Role Groups containing the role
Role Management	Microsoft Purview Administrators Organization Management
DLP Compliance Management	Compliance Administrator Compliance Data Administrator Organization Management Security Administrator
Information Protection Admin	Compliance Administrator Compliance Data Administrator Information Protection Information Protection Admins
Data security AI admin	Data Security AI admins

Group B

Expand table

Role	Role Groups containing the role
Security Copilot Contributor	(Managed in Security Copilot)
Purview Copilot Workspace Contributor	Included in multiple role groups. Some examples are Organization Management Compliance administrator Security Administrator Compliance Data Administrator Information Protection Information Protection Admins Information Protection Analysts Information Protection Investigators

For analysts (any one from each group):

Group A

Expand table

Role	Role Groups containing the role
Information Protection Analyst	Compliance Administrator Compliance Data Administrator Information Protection Information Protection Analysts Information Protection Investigators
Information Protection Investigator	Information Protection Information Protection Investigators

Group B

Expand table

Role	Role Groups containing the role
Purview Agent Deployment	Included in multiple role groups. Some examples include Compliance Administrator Data Security Management Information Protection
Information Protection Analysts	Information Protection Investigators

Role	Role Groups containing the role
	Purview Agent Management

Group C

Expand table

Role	Role Groups containing the role
Security Copilot Contributor	(Managed in Security Copilot)
Purview Copilot Workspace Contributor	Included in multiple role groups. Some examples include: Organization Management Compliance administrator Security Administrator Compliance Data Administrator Information Protection Information Protection Admins Information Protection Analysts Information Protection Investigators

See Microsoft Entra Agent ID documentation | Microsoft Learn for more on agent identity.

Wait for 30-min post setting the agent with agent identity for the roles based access to get reflected before initiating any manual run on an alert.

> **Important**
>
> After setup is complete, the agent starts running within 30–60 minutes and begins triaging DLP alerts. By default, the agent processes alerts from all DLP policies generated within the past 30 days. You can modify this behavior at any time by updating the agent's settings.

Roles assigned to the agent

These roles are assigned to the Triage agent in DLP when agent identity is used:

- Information Protection Analyst
- Purview content Analyst
- Data classification content viewer (needed to get the contextual summary)
- Data classification content downloader (needed for endpoint scenario)

- Purview Copilot Workspace Contributor
- Security Copilot Contributor

> **Important**
>
> The agent starts running within 30–60 minutes after setup is complete, and begins triaging DLP alerts. By default, the agent processes alerts from all DLP policies generated in the past 30 days. You can modify this behavior at any time by updating the agent's settings.

Permissions for viewing alert results in Purview and Defender XDR and managing Microsoft Purview Triage Agent for DLP

The account you use to view results along with update triggers, give feedback, or custom instruction for the Triage Agent in DLP must be assigned the roles (any one from each group):

Group A

Expand table

Role	Role Groups containing the role
Information Protection Analyst	Compliance Administrator Compliance Data Administrator Information Protection Information Protection Analysts Information Protection Investigators
Information Protection Investigator	Information Protection Information Protection Investigators
Manage Alerts	Compliance Administrator Compliance Data Administrator Organization Management Security Administrator Security Operator
View-Only Manage alerts	Compliance Administrator Compliance Data Administrator Global Reader Organization Management Security Administrator Security Operator Security Reader

Group B

Expand table

Role	Role Groups containing the role
Purview Agent Analysis	Included in multiple role groups. Some examples Compliance administrator Data Security Management Information Protection Information Protection Analysts Information Protection Investigators

Group C

Expand table

Role	Role Groups containing the role
Security Copilot Contributor	(Managed in Security Copilot)
Purview Copilot Workspace Contributor	Included in multiple role groups. Some examples Organization Management Compliance administrator Security Administrator Compliance Data Administrator Information Protection Information Protection Admins Information Protection Analysts Information Protection Investigators

Deployment and configuration roadmap

Implementing Microsoft Purview agents involves several phases:

1. Infrastructure prerequisites for agents
2. Enable agents
3. Edit agents
4. Deactivate agents
5. Remove agents

Infrastructure prerequisites

Microsoft Purview triage agent in DLP runs on Microsoft Security Copilot.

- Your tenant must be onboarded to Microsoft Security Copilot. For more information on how to onboard, see Get started with Microsoft Security Copilot.

- You must enable Microsoft 365 data sharing in Security Copilot. For more information, see Accessing data from Microsoft 365 services .
- You must enable the Microsoft Purview plug-in in Microsoft Security Copilot. For more information, see Enable the Microsoft Purview source in Microsoft Security Copilot.
- To use the Remediation Capability of the Triage Agent in Data Loss Prevention, you must configure the Microsoft Teams Admin Center to allow the Data Security Triage Agent app to send messages to users. For detailed configuration steps, see Configure Microsoft Teams for the Data Security Triage Agent remediation messages.

Enable the agent

This procedure is for organizations that didn't enable any of the Microsoft Purview agents or removed agents and want to enable them again. Once you enable the Microsoft Purview Triage Agent in Data Loss Prevention, they're available for use in Microsoft Purview and Microsoft Defender XDR.

> **Note**
>
> There can be only one instance of each agent in a tenant, regardless of the portal in which it was deployed.

This process works for the Triage Agent in Microsoft Purview.

1. Sign in to the Microsoft Purview portal with an account that has the required permissions.
2. In the left hand navigation pane, select **Agents**.
3. Select **Explore agents**.
4. On the card for the Triage agent in Data Loss Prevention, select **View Details** to open a page that shows the details about the agent.
5. Select **Setup** to open the **Deploy agent** global configuration page. You can:
    a. Choose to **Run automatically based on a set schedule**. If you don't choose this option, you must run the agent manually one at a time. Microsoft sets the schedule and it isn't configurable by organizations. You can change this setting later when you edit the agent.

b. Select the **alert timeframe**, which is how far back the agent looks for alerts to triage. Analysts can shorten the timeframe when they edit the agent but not lengthen it. For more information, see Select Alert timeframe.

c. Enable **Remediation reminders in Microsoft Teams**. For Sharepoint/OneDrive alerts triaged as Needs Attention, the agent sends a Microsoft Teams chat to the last modified user of the file, and ask them to remove the sensitive information found in the file related to the alert. (Preview)

d. Select **Reminder duration**, which is the number of days the agent sends the reminders on Microsoft Teams for a particular file requiring action from last modified user. (Preview)

6. The agent is assigned **Agent Identity**, which allows it to access data and take actions

7. Select **Start**. You see the *Triage Agent in <solution> is active* message when the agent is successfully deployed.

This process works for the Triage Agent in Defender XDR.

> **ⓘ Important**
>
> The Triage Agent can be deployed from both Microsoft Purview and Defender XDR portals, but can only be managed, edited, and disabled from the Microsoft Purview portal. After deployment from either portal, Agent summaries and outputs can be viewed in both the Microsoft Purview and Defender XDR portals.

1. Sign in to the Microsoft Defender XDR portal with an account that has the required permissions.

2. In the left hand navigation pane, select "Investigation & response" then select "Incidents & alerts" and then select "Alerts".

3. Click into any DLP alert the alert queue.

4. If you don't have the Triage Agent deployed, you will find a banner inside a DLP alert suggesting you deploy it.

   a. Click on "Get started" to start the agent with default settings.

   b. Click on "Learn more" to understand more about the configuration options available to you.

c. Finish deploying the Agent

Deploy from the Defender XDR Security Store

You can also deploy the Data Security Triage Agent from the Security Copilot Security Store in the Defender XDR portal.

> **Important**
>
> The Triage Agent can be deployed from both Microsoft Purview and Defender XDR portals, but can only be managed, edited, and disabled from the Microsoft Purview portal. After deployment from either portal, agent summaries and outputs can be viewed in both portals.

1. Sign in to the Microsoft Defender XDR portal with an account that has the required permissions.
2. Select **Security Copilot** > **Security Store**.
3. Search for or browse to the **Triage Agent in Data Loss Prevention** card.
4. Select the agent card, then select **Get agent** to begin deployment.
5. After purchase, select **Security Copilot** > **Agents**, find the agent in the **Ready for setup** section, and select **Set up**.
6. Choose whether to run automatically or manually, set the alert timeframe, and optionally enable remediation reminders in Microsoft Teams. For more information on these settings, see Triggers and settings.
7. Select **Start**. The agent begins triaging DLP alerts within 30–60 minutes.

'One-Click' deployment and enabling the agent

A quick setup flow with default triggers was introduced. Any admin or analyst for a tenant (with no agent deployed) sees a banner or card on alerts page for Triage agent setup. Select **Start agent** to deploy and enable the agent. The roles needed for this setup align with the roles mentioned previously for normal setup flow.

The default settings for 'One-Click' setup are:

- Identity: Create and use agent identity
- Trigger: Run automatically based on a set schedule
- Alert timeframe: Triage alerts from the last 30 days

- Remediation reminder: Turned OFF (Preview)
- Policy scope: All policies in scope for triage.

These settings can be customized during 'One-Click' setup by selecting the Customize option.

Edit agent

Once an agent is enabled, it's ready to triage alerts. Edit and set specific triggers for the agent to act upon. These triggers are used to determine which alerts the agent will triage.

Select which triggers are used on either the **Agents** page or, for first run experience, on the **Alerts** page for the solution. Select the **Agent details** button on the top right on alerts page or:

1. In the Microsoft Purview navigation pane (left hand side of the screen), select **Agents**.
2. Select **Explore agents**.
3. Select **Go to agent** for agent you want to edit.
4. Select **Edit agent**.
5. Select **Agent configuration** for editing and configuring settings like Trigger, alert timeframe, remediation reminders, reminder duration, and policy scope.
6. Go to **Agent Custom Instruction** for editing and configuring the customized instruction on how the agent should work in your tenant.
7. Select **Save**

> **Important**
>
> The most recent agent configuration is always used. Agent settings can only be edited from the Microsoft Purview portal even if it was deployed from the Defender XDR portal.

Allow up to 2 hours for the agent to complete triaging in-scope alerts and enabled manual runs after this initial setup.

Deactivate agents

1. Sign in to the Microsoft Purview portal with an account that has the required permissions.
2. In the left hand navigation pane, select **Agents**.
3. Select **Explore agents**.
4. Select **Go to agent** for the agent you want to deactivate.
5. On the far right upper right hand corner of the agent overview page, select the ellipses (three dots) that are located *next to* the **Edit agent** button.

6. Select **Deactivate agent**. Deactivating the agent stops the agent from triaging alerts. It doesn't remove the agent and it doesn't reset the **Select alert timeframe** reference point in time.

> **Important**
>
> Agent can only be deactivated from the Microsoft Purview portal even if it was deployed from the Defender XDR portal.

Remove agents

1. Sign in to the Microsoft Purview portal with an account that has the required permissions.
2. In the left hand navigation pane, select **Agents**.
3. Select **Explore agents**.
4. Select **Go to agent** for the agent you want to remove.
5. On the far right upper right hand corner of the agent overview page, select the ellipses (three dots) that are located *next to* the **Edit agent** button.
6. Select **Remove agent**. Removing the agent deletes it from Microsoft Purview. If you want to use it again you must go through the Enable the Agents, and Edit agents procedures again.

> **Important**
>
> Agent can only be removed from the Microsoft Purview portal even if it was deployed from the Defender XDR portal.

Triggers and settings

Triggers are groupings of parameters whose values must be met in order for the agent to triage any given alert. Triggers include:

Run automatically or manually

Once an agent is deployed, and triggers are edited, select whether the agent runs **automatically based on a set schedule** or the **Agent runs manually on one alert at a time** . If you select **Run automatically based on a set schedule**, the agent triages the alerts that are included in the **Select Alert timeframe** setting.

Select alert timeframe

Choose an alert timeframe. The options are:

- Only triage new alerts
- Last 24 hours
- Last 48 hours
- Last 72 hours
- Last 7 days
- Last 14 days
- Last 21 days
- Last 30 days

If **Only triage new alerts** is selected, the agent only triages alerts that are generated after the agent is deployed. The agent won't triage any alerts that were generated before the agent was deployed, which means that all the *Last # hours or days* options are ignored.

If any of the *Last # hours or days* options are selected, the agent triages alerts that were generated in the selected timeframe. This selection triages all alerts generated before the agent was deployed. All newly generated alerts are also triaged.

Remediation reminder (Preview)

The Remediation Capability extends the effectiveness of the Alert Triage Agent in Data Loss Prevention by reaching out to the user who violated DLP Policies--by including sensitive data in SharePoint and OneDrive files, for example--via Microsoft Teams. This is applicable only for the alerts categorized as "Needs Attention".

Enabling this setting also enables the bot in Microsoft Teams for the tenant. Reminders are sent once daily to the user who is last modifier of the file in context (combining up to 10 files if multiple files require remediation).

The remediation message includes the file name and a description of what sensitive information requires attention. Once the last person who modified, remediates the file, the file is removed from the remediation reminder for the next day. If no files are found, no reminder is sent.

> **ⓘ Important**
>
> If your organization's settings for Microsoft Teams do not allow the deployment of new Microsoft Teams apps, the Remediation reminders will not be sent to users in your

organization. For configuration steps, see [Configure Microsoft Teams for the Data Security Triage Agent remediation messages](#).

Reminder duration (Preview)

The reminder duration reports the number of days the Agent should follow up with the user on files that aren't remediated.

Policy Scope

A DLP policy must meet certain criteria for it to appear in the **Specify Policy Scope** list and be eligible to be included triage.

Specify the policy scope of the triage agent to control which DLP policy alerts and which policies are eligible for remediation reminders. Remediation reminders must be enabled, and eligible policies must be a subset of the policies defined for agent triage.

Full eligibility

A policy is fully *eligible* for inclusion if the DLP Alert Triage Agent supports all the conditions in the policy. All alerts from a fully eligible policy are triaged.

Limited eligibility

If a policy has some unsupported conditions--a state called *limited eligibility*--the Microsoft Purview DLP triage agent might not be able to triage all the alerts from that policy. Based on the criteria listed some alerts from the policy won't be triaged in limited eligibility. The reason for policy showing as *limited* can be seen by selecting **Limited** state in the **Specify Policy Scope** picker.

Here are some reasons why a policy can be in a limited eligibility state:

- Get started with collecting files that match data loss prevention policies from devices isn't enabled or the storage associated with it isn't Microsoft storage. Also the File collection setting in policy rule is turned OFF for policies targeting endpoint or devices workload.
- The policy doesn't include one of these workloads: SharePoint, OneDrive, Exchange, Teams, and devices (Endpoint).

The same eligibility status is shown on the DLP policy list page. The eligibility column appears when the DLP Triage Agent is enabled and set up. Once the policies are included in agent scope,

only future alerts are checked for scoped policies. Existing triaged alerts, or alerts that have already been generated, aren't impacted.

Custom instructions (Preview)

During the agent config process, you can give the agent *custom instructions*. Custom instructions only apply to the Triage agent in DLP. Microsoft Purview agents use these natural language instructions to enhance alert triage by:

- Translating your input into structured classification logic.
- Running this logic against the alert content associated with each alert.
- Raising or lowering the alert’s priority if the content matches your custom instruction.

For custom instructions, the agent now supports both content, and metadata-based instruction, such as:

- Focus on alerts related to tax, financial, or legal information from last seven days
- Don't focus on alerts containing credit card numbers, social security numbers, and names of users like Mary@contoso.com
- Deprioritize alerts where the related assets have .pdf or .jpg files

Users can also use provided templates when unsure how to provide instruction. These templates are:

- *Generic instruction*: Helps define the clear classifiers targeted, or prioritized conditions versus those classifiers that aren’t prioritized. Separate complex conditions to avoid agent generating complex inferences.
- *Focus on*: Helps define clear conditions to prioritize. Other conditions are marked as Less Urgent.
- *Reduce overload*: Helps define conditions to deprioritize.
- *Definitions*: Defines definitions for classifiers for the agent's understanding.

Edit custom instructions

- The user can add multiple templates to the instruction as required. In other words, they can select a generic instruction and add all needed details or remove whatever isn't required. However, they can also add these templates on demand and enter definitions, then add a couple of *Reduce overload conditions*.
- The user can add more than one condition by clicking on multiple templates, or add separate lines after single template select.

- They can also add conditions in any order and repeat conditions if necessary. For example, the user can add template for definition, then add a *Focus on*, a *Reduce overload*, a *Focus on*, and any needed *Definitions*. Compounding these conditions is like adding things to a cart by selecting and removing the items as required and in any sequence.
- The existing text in the input box isn't removed or replaced when a template is selected. The template-based text is always appended to existing text.
- Users can always edit the input box even with the template as required. They can remove the complete template-provided input or make changes.

On review, the agent combines all instances and presents agentic understanding for confirmation. The output is shown in three segments:

- *Classifiers identified from instructions*: Shows list of classifiers the agent identifies from the instruction.
- *Agent-inferred instructions*: Agent interpretation converts user-provided natural language instruction. Users can deselect any instruction they don't want.
- *Unsupported instructions*: Highlights the instruction if any part of the instruction isn't supported. Users can edit the custom instruction at any time and review it again to remove any unsupported conditions.

Once the instruction is saved, the agent uses it to triage all future alerts. If an alert matches the instruction, a separate summary is generated. The agent also highlights the custom-instruction–matched classifiers for each related document associated with the alert.

Monitor SCU usage

1. Sign in to the Microsoft Purview portal with an account that has the required permissions.
2. In the left hand navigation pane, select **Agents**.
3. Select **Explore agents**.
4. Select **View agent** for agent you want to edit.
5. Select **Edit agent**.
6. Select **Usage monitoring**.
7. You can track your SCU consumption in the usage monitoring tool.

Microsoft Purview Alerts page overview

1. Sign in to the Microsoft Purview portal with an account that has the required permissions.
2. Open the solution you want to view the triaged alerts for.
3. Open the **Alerts** page for the solution.

4. In the top right hand area of the page, there's a new toggle that lets you choose between the **Standard** view of the alerts page and a **Triage Agent** view of the alerts page. Set the toggle to the **Triage Agent** view. This view shows agent triaged alerts. The agent groups alerts into four categories:

- **All**
- **Needs attention**
- **Less urgent**
- **Not categorized**

Defender XDR Alert queue

1. Sign in to the Microsoft Defender XDR portal with an account that has the required permissions.
2. In the left hand navigation pane, select "Investigation & response" then select "Incidents & alerts" and then select "Alerts".
3. Agent-provided summaries and risk factors will be available on DLP alerts triaged by the Agent as they are in Microsoft Purview.

Triaged alerts

The agent triages alerts generated in the timeframe and using policies you selected. Not all alerts are triaged. Triaged alerts are available to review in both the Microsoft Purview and Defender XDR portals.

Triaged alerts in Microsoft Purview are grouped into four categories:

1. **All**: This category includes all the alerts that the agent triaged. The count indicated in the category might not accurately reflect the true number of alerts until you go into that view and scroll down to load all the alerts. If the conditions that caused the alert to be raised in the first place changed, or if the alert isn't triaged yet, you can select the alert and then select **Run agent** to manually run the agent on the alert.

2. **Needs attention**: This category contains alerts that the agent evaluated and determined pose the greatest risk to your organization. When you select one of these alerts, the details flyout opens to show a summary of the alert and other details.

3. **Less Urgent**: These less urgent results are the alerts the agent evaluated and determined to pose a lower risk to your organization. When you select one of these alerts, the details flyout opens to show a summary of the alert and other details.

4. **Not categorized**: Uncategorized results are the alerts that the agent wasn't able to successfully triage, which can happen for many reasons, including:

- Server error
- In process of reviewing
- other error
- Unsupported error for alerts that contain activities that the agent doesn't support.

Agents triage files up to 2 MB in size.

How agents prioritize

The Triage Agent in DLP prioritizes alerts based these risk factors:

- *Content Risk*: This category is the primary risk factor used during agent triage, it covers sensitive content based on Microsoft provided sensitive information types (SIT), trainable classifiers, and default sensitivity labels. For more information, see default sensitivity labels.
- *Exfiltration Risk*: Exfiltration of sensitive data shared externally.
- *Policy Risk*: Policy mode and rules with actions affect the prioritization of alerts.
- *Content Risk*: Label removed or downgraded.

Additionally, any custom instruction or feedback is applied for the final verdict categorization.

Alert Triage details

> **Important**
>
> The Triage agent in DLP only supports alerts from policies that are in active mode. It doesn't triage alerts from DLP policies that are running in simulation mode.

Agents can review alerts that were generated up to 30 days before enablement if the tenant has sufficient SCUs. Alerts that were generated more than 30 days before agent enablement are out of scope.

The Triage agent in DLP triages alerts from Exchange, Devices, SharePoint, OneDrive, Teams. To triage alerts from Devices, you must enable evidence collection for file activities on devices.

Do manual analysis on alerts not evaluated by an agent.

Content Analysis

There are some situations where content analysis can be limited.

The content risk prioritization of an alert is based on Sensitive Information Types (SITs), trainable classifiers, and sensitivity labels in content. When an agent evaluates content risk, it only looks for SITs, and trainable classifiers that are defined in the policy.

When a DLP alert is associated with fewer than 10 files, the agent scans all the files and uses them to generate the content summary. When an alert includes 10 or more files, the agent uses the potentially top 10 files to generate the file risk summary. In DLP, the triage agent selects the top 10 risky files based on the number of policy classifier hits, the file size, and the most recent file access time. Once this selection is done, the triage agent adds a note stating that the content summary doesn't include all the files in the alert.

Remediation capability within alert view (Preview)

There are two ways to view the agent's remediation progress.

You can view remediation status at the alert level on the Triage Agent alert queue in the DLP solution. This view lets a user see fully remediated alerts, meaning users remediated all files implicated in the alert. *In progress* means that the Agent reached out to the user, but the user didn't remediate all the files yet. *Error* indicates that the Teams message couldn't be sent.

Use the second view to see files in a DLP alert that are remediated. You can see the status inside each DLP alert in the *Related assets* section that lists all the files implicated in the alert and the respective remediation status. *Remediated* indicates that the user remediated all detected SITs in the file. If the user remediated some or none of the SITs, the status indicates *In progress*.

> **Note**
>
> The Remediation Capability doesn't cost anything extra. The only costs associated with using this capability are the SCUs needed to triage the DLP alert first.

Feedback for alerts triaged

> **Important**
>
> You can only provide feedback on alerts in that are in the **Needs attention** or **Less urgent** categories. If you disagree with the agent's initial categorization of an alert, you can use the feedback mechanism to revise the categorization. If the initial value is **Needs attention**, you

can change it to **Less urgent**. If the initial categorization is **Less urgent**, you can reprioritize it as **Needs attention**.

You can also add properties, such as a **Sensitive info type** or a **File path** that the agent should use in future evaluations to improve performance. These properties are supported in preview.

- **User email address**: the user who performed the activity that triggered the alert
- **External recipient email address**: to track alerts triggered by Exchange email or Teams interactions with external recipients
- **Sensitive information type**: shows all the SITs within the subscription, the SITs involved in the policy rule triggering the alert are preselected
- **Trainable classifier**: shows all trainable classifiers, the trainable classifiers that are involved in the policy rule triggering the alert are preselected
- **Sensitivity label**: The labels present in the subscription
- **File path**: in case the alert is related to a file and the file path is available. The file path isn't enabled for Endpoint DLP (device) alerts but Full file evidence wasn't enabled at the time of alert triaging
- **Target domain**: for endpoint devices DLP alerts where target domain is present

1. Select any **Needs attention**, or **Less urgent** alert. The selection opens a flyout with the agent provided summary and other settings.
2. Select **Agent feedback**.
3. The **Revise to** field shows the value of the recategorization.
4. Select **+ Add property** and add one or more properties. The added properties are used to improve triaging performance.
5. If you want to **Apply the feedback to all policies**, select that option. Otherwise, the feedback is applied only to the policy that triggered the alert.
6. Select **Review** to see a summary of the changes.
7. Select **Submit** to save the feedback.
8. The current alert doesn't change immediately. Admin has to manually run a triage pass on the alert to change the categorization based on the feedback provided.

Managing feedback

You can view and manage all the feedback given for a triaged alert, including exporting, editing, and deleting feedback.

1. In the solution for the alert, open the **Alerts** page.
2. Select the **Needs attention** or **Less urgent** category.

3. Select the alert that you want to manage the feedback for.
4. Select **Agent feedback**.
5. Select **View all feedback**.
6. Select the alert and feedback entry that you want to manage.
7. **Edit**, **Delete**, or **Export** the feedback as needed.

Feedback conflict resolution

Feedback conflict can occur when multiple admins provide conflicting feedback for the same user and policy combination on different alerts. Feedback conflict generates an error.

For example:

- Admin 1 provides feedback to change *all* alert's categorization to **Less urgent** if the alert is for User A and Policy P.
- Admin 2 provides feedback to change *an* alert's categorization to **Needs attention** if the alert is for User A and Policy P.

Other considerations

- For the alert count, only the alerts where single instance of match is enabled are counted. In case an alert is generated for a threshold scenario--for example, generate alert if there are 10 matches in last 24 hours--those alerts aren't counted.

Next steps

Refer to solution specific articles for information on reviewing the triaged alerts.

- Learn about investigating data loss prevention alerts
- Get started with the data loss prevention Alerts dashboard
- Alert Triage Agent dashboard (preview)
- Indicators and the Alert Triage Agent for Insider Risk Management

---

Last updated on 06/25/2026

Configure Microsoft Teams for the Data Security Triage Agent remediation messages

The Data Security Triage Agent in Data Loss Prevention (DLP) can send remediation messages through Microsoft Teams to users who last modified files containing sensitive information. Before the agent can deliver these messages, you must configure two settings in the Microsoft Teams Admin Center.

This article covers the required Microsoft Teams configuration, how to control which users receive remediation messages, and how to troubleshoot common delivery failures.

> **Important**
>
> Remediation reminders for the Data Security Triage Agent are currently in **Preview**. Features and steps described in this article might change before general availability.

For information about deploying and configuring the triage agent itself, see Get started with the Microsoft Purview Triage Agent in Data Loss Prevention.

Prerequisites

Before configuring Teams for remediation messages, ensure these requirements are met:

- The Data Security Triage Agent in DLP must be deployed and configured.
- **Remediation reminders** must be enabled in the agent's settings. For more information, see Remediation reminder (Preview).
- You must have Teams admin permissions to configure the Microsoft Teams Admin Center settings described in this article.

> **Note**
>
> The Purview admin or analyst who configures the triage agent might not have permissions to access the Microsoft Teams Admin Center. If that's the case, coordinate with a user in your organization who has Teams admin permissions to configure these settings.

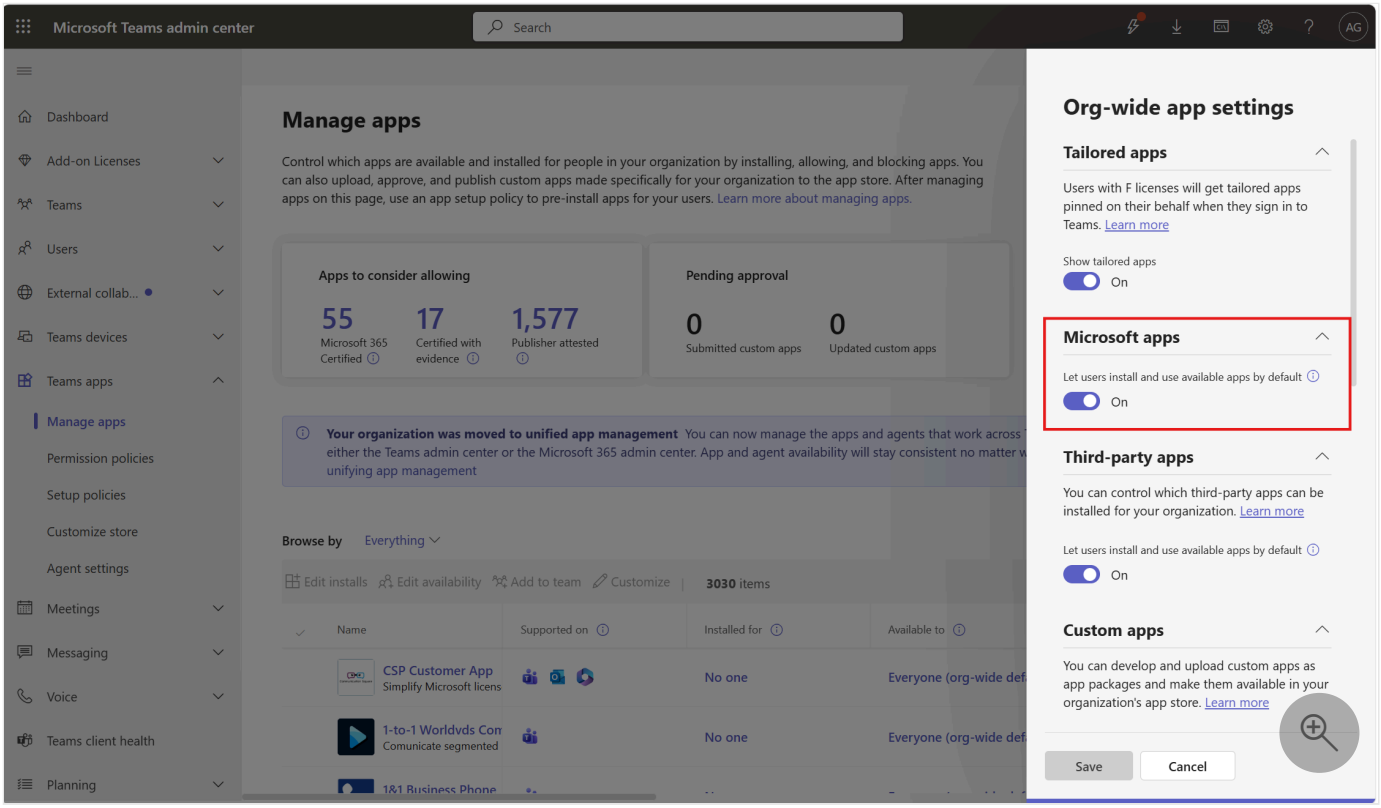
Required Teams Admin Center settings

Two settings must be correctly configured for the Data Security Triage Agent to send remediation messages through Microsoft Teams.

Enable org-wide Microsoft apps

The org-wide Microsoft apps setting controls whether Microsoft-published apps can be used across your organization. This toggle must be set to **On** for the Data Security Triage Agent to function.

1. Sign in to the Teams Admin Center.
2. Go to **Teams apps** > **Manage apps**.
3. Select **Actions** > **Org-wide app settings**.
4. In the **Org-wide app settings** panel, set the **Microsoft apps** toggle to **On**.
5. Select **Save**.

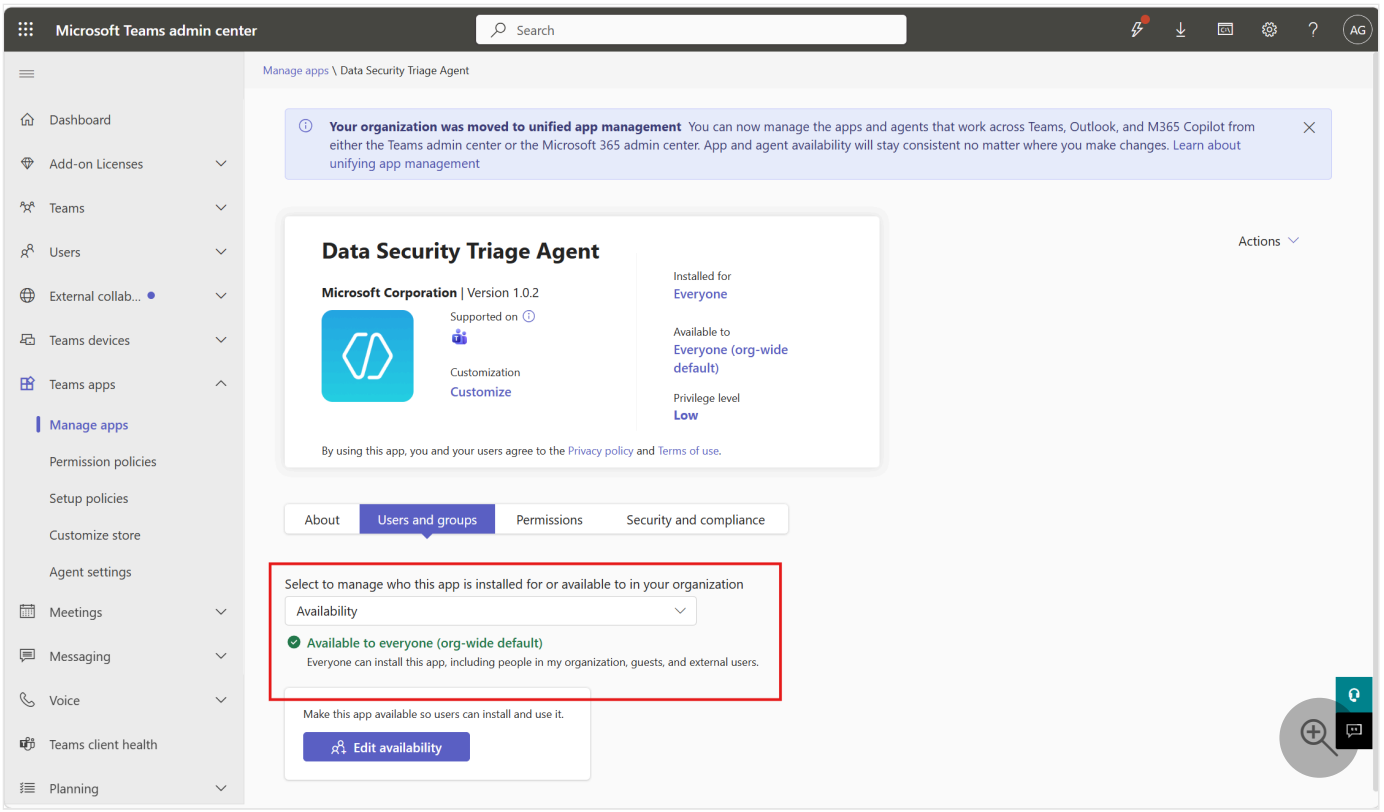


Make the Data Security Triage Agent app available

The Data Security Triage Agent app must not be blocked for users who should receive remediation messages.

1. Sign in to the Teams Admin Center.
2. Go to **Teams apps** > **Manage apps**.
3. Search for **Data Security Triage Agent** and select it.
4. Go to the **Users and groups** tab.

5. Verify that **Availability** is set to **Available to everyone (org-wide default)**.



> **Important**
>
> The Data Security Triage Agent app doesn't need to be preinstalled for users. Microsoft Purview automatically installs the app for a specific user at the time a remediation message is sent. The app only needs to be *available* (not blocked) in your organization.

Control who receives remediation messages

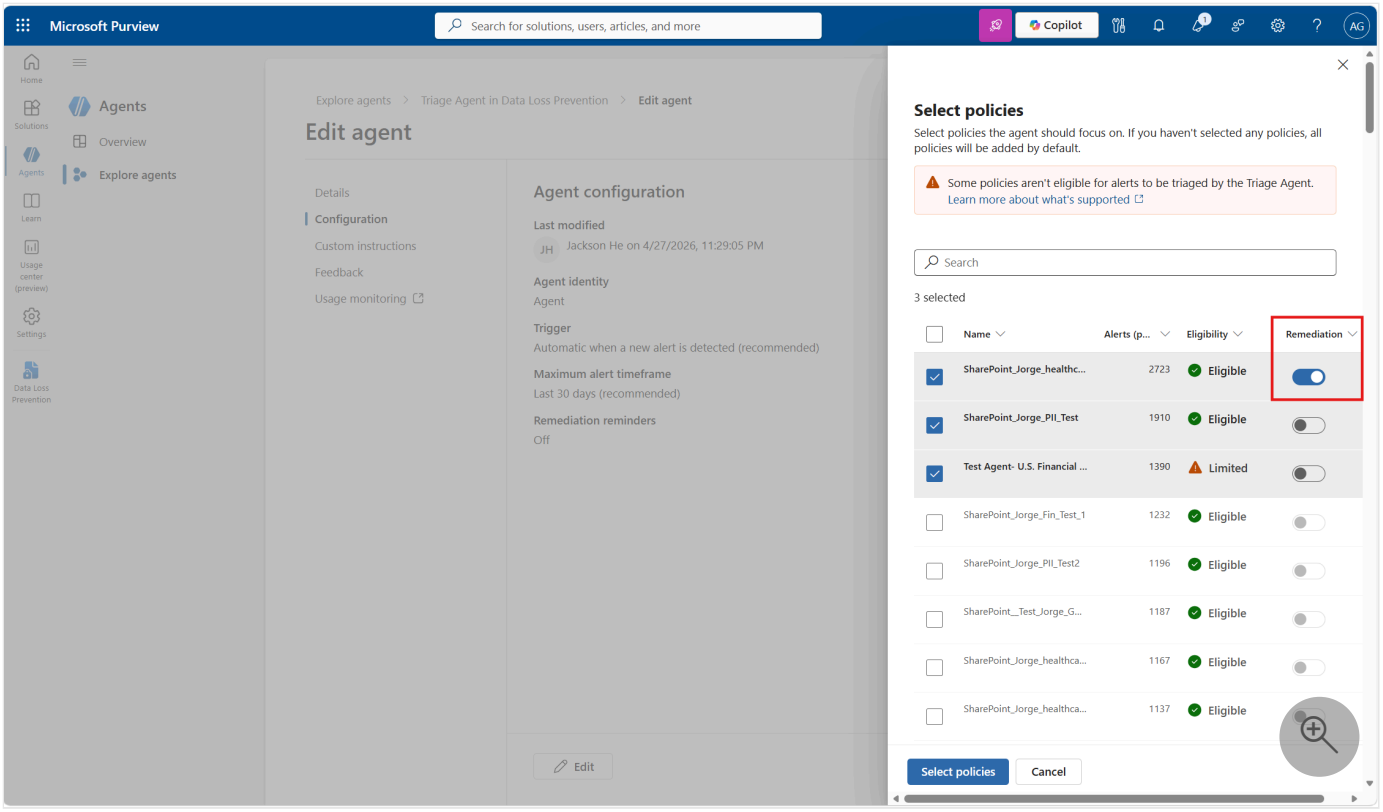
When remediation reminders are enabled, the Data Security Triage Agent sends Microsoft Teams messages to users who last modified files containing sensitive information. You can control which users receive these messages through two mechanisms.

Limit remediation to specific policies

To limit the scope of remediation to specific files or locations, edit the agent's policy scope to enable remediation on only a subset of triage-eligible policies. Policies enabled for remediation are always a subset of policies enabled for triage.

1. In the Microsoft Purview portal, go to **Agents** > **Explore agents**.
2. Select **Go to agent** for the Triage Agent in Data Loss Prevention.
3. Select **Edit agent** > **Agent configuration**.

4. Under **Set agent scope**, select **Edit**.
5. In the **Select policies** panel, enable the **Remediation** toggle only for the policies where you want remediation messages sent to users.
6. Select **Select policies** to save.

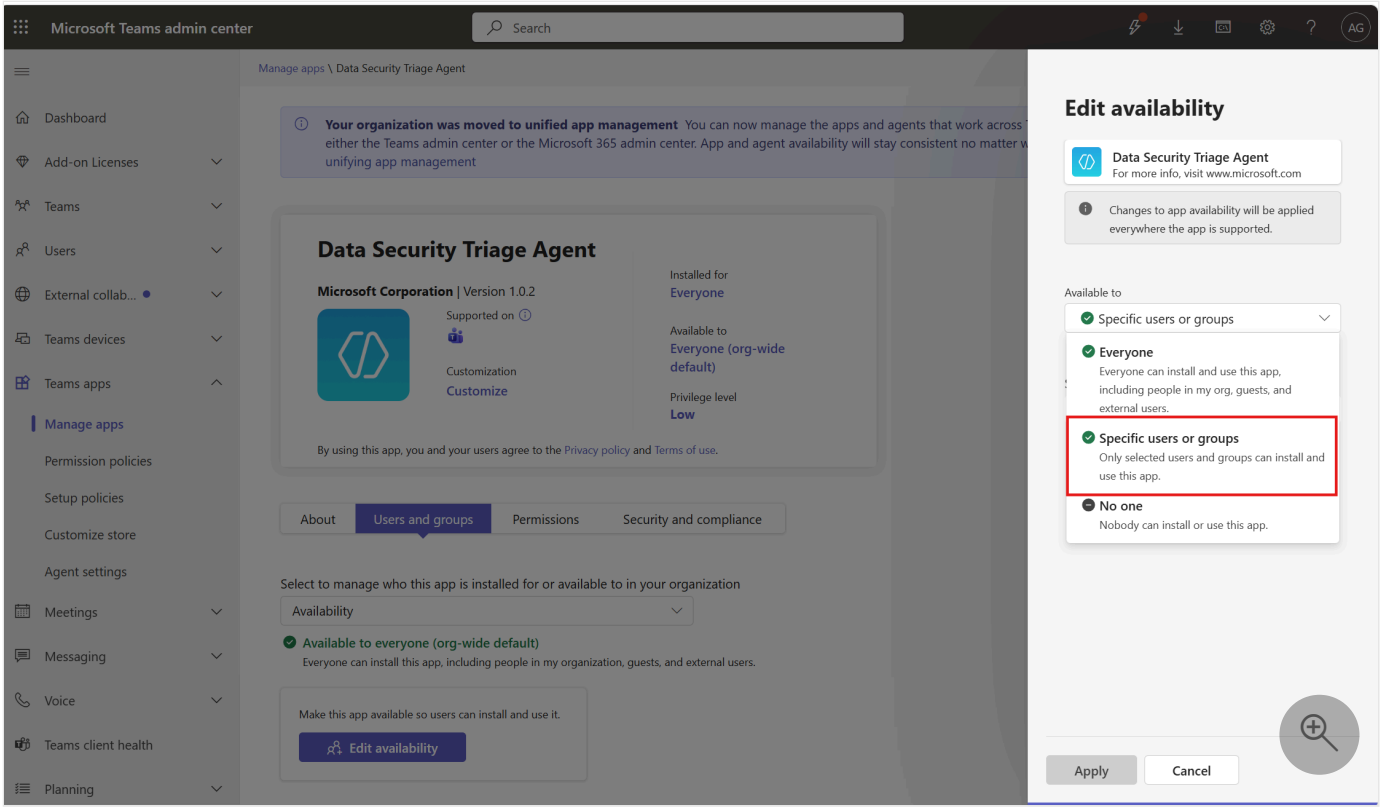


This approach limits remediation by controlling which DLP policies trigger remediation messages. Users who violate policies that aren't enabled for remediation don't receive Teams messages from the agent.

Block the app for specific users or groups

To prevent the agent from sending messages to specific users--such as executives, sensitive roles, or specific organizational groups--change the app's availability in the Microsoft Teams Admin Center.

1. Sign in to the Teams Admin Center.
2. Go to **Teams apps** > **Manage apps**.
3. Search for **Data Security Triage Agent** and select it.
4. Go to the **Users and groups** tab.
5. Select **Edit availability**.
6. Change the availability from **Everyone** to **Specific users or groups**, then add only the users or groups who should receive remediation messages.



> **Important**
>
> Blocking the app for a user is a hard block. Even if an alert is triaged and a remediation message is generated, the message *won't* be sent to users for whom the app is blocked. This result is a failsafe mechanism--if you need to *immediately* stop all remediation messages, set availability to **No one**.

> **Tip**
>
> Use policy scoping in Purview for routine control over what files and locations trigger remediation. Use Teams Admin Center blocking as another safeguard when you need to guarantee that specific users never receive messages from this agent, regardless of which policies are in scope.

Troubleshoot remediation message delivery

If remediation messages aren't being sent, check:

Expand table

Symptom	Likely cause	Resolution
No users receive messages	Org-wide Microsoft apps toggle is off	Enable org-wide Microsoft apps
No users receive messages	Data Security Triage Agent app is blocked	Make the app available
Some users don't receive messages	App is blocked for specific users or groups	Check the app's **Users and groups** settings in Teams Admin Center
Messages sent but not appearing	User hasn't opened Teams recently	Messages appear once the user signs in to Teams

> **Important**
>
> A common misconfiguration occurs when the org-wide Microsoft apps toggle is *enabled* but the Data Security Triage Agent app is *individually blocked*. Both settings must be correctly configured for remediation to work.

Related content

- Get started with the Microsoft Purview Triage Agent in Data Loss Prevention
- Security Copilot agents in Microsoft Purview overview
- Manage access to Teams apps using app permission policies

Last updated on 05/11/2026

Get started with the Data Security Posture Agent

This feature is in preview.

The Microsoft Purview Posture Agent in Data Security Posture Management uses natural language prompts to help you find sensitive information across your organization. Follow these steps to deploy the Microsoft Purview Posture Agent in the following Microsoft Purview solutions:

- Data Security Investigations
- Data Security Posture Management

Before you begin

If you're new to Microsoft Security Copilot Agent in Microsoft Purview, read this article.

- Security Copilot Agents in Microsoft Purview Overview

SKU/subscriptions and licensing

This agent requires both the standard per seat licensing model and the pay-as-you-go billing model. Your organization must be licensed for:

- Microsoft 365 E5, with security compute units (SCUs) provisioned to use the Microsoft Purview posture agent in Data Security Posture Management.

The agent consumes security compute units (SCUs) as it performs its tasks. You must have SCUs provisioned for the agent to work. The agent consumes SCUs every time it runs based on the complexity of analysis it performs. For more information about SCUs, see Security compute units (SCUs). You can track your SCU consumption in the usage monitoring tool. For more information about onboarding into Microsoft Security Copilot, see Get started with Microsoft Security Copilot.

For information on Security Copilot licensing in E5 see, Learn about Security Copilot in Microsoft 365 E5.

For information on licensing, see

- Microsoft 365 Enterprise Plans
- Microsoft 365 Service Descriptions

Permissions and Roles

You can enable and deploy the Microsoft Purview Posture Agent in Data Security Posture Management or Data Security Investigations with either an organizational user account or an *agent identity*.

We recommend that you deploy the agent using an *agent identity*. The account used to deploy the agent with an agent identity must have the *Role Management* role. To get started with agent identities, see Governing Agent Identities (preview).

If you enable and deploy the agent with your organizational user account, see Permissions for deploying the Posture Agent.

There are different permissions and roles needed to perform different functions with the agent. For more information, see Permissions in the Microsoft Purview portal, and Roles and role groups in the Microsoft Purview portals.

Permissions for deploying the Posture Agent

Use an account that has **the following roles as described** in the table.

Expand table

One from Group A	OR Both from Group B	AND One from Group C
- Compliance Admin OR - Security Reader OR Data Security Viewer OR - Data Security AI Viewer OR - Data Security AI Admin	- Purview Content Analyst - Purview Agent Deployment	- Purview Copilot Workspace Contributor

Permissions for running and viewing results from the Posture Agent

Use an account that has **the following roles as described** in the table.

Expand table

Role Name	Can run Posture Agent and view summary	Can view content snippet that matched intent	Can view file matched intent	Can label file
Compliance Admin	Yes	No	No	No
Security Reader	Yes	No	No	No

Role Name	Can run Posture Agent and view summary	Can view content snippet that matched intent	Can view file matched intent	Can label file
Data Security Viewer	Yes	No	No	No
Data Security AI Viewer	Yes	No	No	No
Data Security AI Admin	Yes	No	No	No
Data Classification Content Viewer	Yes	Yes	Yes	Yes
Data Classification List Viewer	Yes	Yes	No	No
Data Security Investigation Reviewer	Yes	Yes	No	No
Data Security Investigation Investigator	Yes	Yes	Yes	Yes
Data Security Investigation Admin	Yes	Yes	Yes	Yes
Purview Content Analyst	Yes	Yes	Yes	Yes

Deployment and configuration roadmap

Implementing the Microsoft Purview agents involves several phases:

1. Infrastructure prerequisites
2. Enable the agent
3. Manage the agent configuration
4. Monitor SCU usage

Infrastructure prerequisites

- Your tenant must be onboarded to Microsoft Security Copilot. For more information on how to onboard, see Get started with Microsoft Security Copilot.
- You must enable Microsoft 365 data sharing in Security Copilot. For more information, see Accessing data from Microsoft 365 services.
- You must enable the Microsoft Purview plug-in in Microsoft Security Copilot. For more information, see Enable the Microsoft Purview source in Microsoft Security Copilot.

> **Important**
>
> You can only run two prompts concurrently.

Enabling the agent

This procedure is for organizations that without the Microsoft Purview Posture Agent for Data Security Posture Management enabled or organizations who removed it and want to enable it again. After you enable the agent, it's available in Microsoft Purview. There can be only one instance of each agent in a tenant.

1. Sign in to the Microsoft Purview portal with an account that has the required permissions.
2. In the left hand navigation pane, select **Agents**.
3. Select **Explore agents**.
4. Select the agent to enable, and then select **Add**. A page opens that shows the requirements to enable the agent.
5. Select **Setup**, this action opens the **Deploy agent** global configuration page. You can:
   a. Choose **Create and use agent identity (recommended)** or **Assign and use my identity**.
6. Select **Start**. You see the **Agent is now active** message when the agent is successfully deployed.

Managing the agent configuration

After the agent is deployed, make minor changes to the agent configuration as needed.

1. Sign in to the Microsoft Purview portal with an account that has the required permissions.
2. In the left hand navigation pane, select **Agents**.
3. Select **Explore agents**.
4. Select **Go to agent** for the agent you want to manage. This action opens the agent overview page.
5. Select the ellipses (three dots) on the upper right hand corner of the agent overview page, next to the **Edit agent** button. From here you can **Deactivate agent** which makes it inactive, but doesn't uninstall it. If you want to uninstall it, select **Remove agent**.
6. Select **Edit agent** to update **Agent identity** under **Deployment configuration**.

Deactivate agent

1. Sign in to the Microsoft Purview portal with an account that has the required permissions.

2. In the left hand navigation pane, select **Agents**.
3. Select **Explore agents**.
4. Select **View agent** for the agent you want to pause. This action opens the agent overview page.
5. In the upper-right corner of the agent overview page, select the ellipsis (three dots) next to the **Edit agent** button.
6. Select **Deactivate agent**. Deactivating the agent stops it from functioning. It doesn't remove the agent.

Remove agent

1. Sign in to the Microsoft Purview portal with an account that has the required permissions.
2. In the left hand navigation pane, select **Agents**.
3. Select **Explore agents**.
4. Select **View agent** for the agent you want to remove. This action opens the agent overview page.
5. On the far right upper right hand corner of the agent overview page, select the ellipses (three dots) that are located *next to* the **Edit agent** button.
6. Select **Remove agent**. Removing the agent deletes it from Microsoft Purview. To use it again, follow Enabling the agent.

Monitoring SCU usage

1. Sign in to the Microsoft Purview portal with an account that has the required permissions.
2. In the left hand navigation pane, select **Agents**.
3. Select **Explore agents**.
4. Select **View agent** for the agent you want to edit. This action opens the agent overview page.
5. Select the **Performance** tab.
6. Track your SCU consumption in the usage monitoring tool.

In some cases, the Agent downloads and scans max 5 GB or 10K items from the selected data sources. The Agent only returns the top 1,000 matched items.

Labeling files discovered by the Agent

You can apply sensitivity labels directly to discovered files. This application means the Data Security Posture Agent not only helps you identify sensitive content, but also helps you start protecting it faster without switching to a separate workflow.

The sensitivity labels available in this feature are the same labels that you define in the Information Protection solution. You can select up to 10 files at a time and apply the same label to all selected files. For supported file types, see File types supported - Microsoft Information Protection (MIP) SDK | Microsoft Learn.

Prerequisites:

1. To view the latest status for labels applied to files, make sure auditing is turned on: https://learn.microsoft.com/en-us/purview/audit-log-enable-disable
2. Users with the Data Classification Content Viewer or the Purview Content Analyst role can apply sensitivity labels.

Steps to apply the label

1. Select a discovery that includes items on the **Files** tab.
2. Review the items in the list, then select up to 10 files by using the checkboxes.
3. After you select the files, select **Label** to open the list of sensitivity labels available in your organization. you can apply sensitivity labels directly to discovered files.
4. Choose a label.
5. Refresh the page after 30 to 60 minutes to check the labeling status.

Run Agent on a schedule

Schedule prompts to run automatically on a recurring basis—daily, weekly, or monthly—so you can continuously monitor data risks without manual effort. Instead of rerunning the same investigations, the agent re-executes the prompt on the same prompt and data sources and surfaces updated results based on changes in your data estate. This scheduling helps you stay on top of new or evolving sensitive data over time while building a consistent monitoring workflow.

You can set a schedule for a prompt at the time of or after running a prompt, if the prompt is eligible to be scheduled.

Run prompts the Agent intelligently suggests

The Agent generates rich insights based on what the Agent observed in your environment. These insights may help you ask the right questions to the Agent and run more targeted prompts.

Instead of starting from scratch, you can build on what the agent already found, making investigations feel more relevant.

Each recommendation includes a **suggested prompt** and **pre-selected data source**, so you can launch a discovery directly from the card with a single click. The intent and sources are prefilled, but you can also edit them further.

Tips and Tricks to write better prompts and get sharper results

Help the Agent understand your intent better. The Agent particularly understands these five search parameters well:

Expand table

Search parameter	How it helps
Identifiers	Basic filters such as people, titles, file names, file characteristics, time ranges, and sharing status. Example: Emails from john@contoso.com; files modified last seven days; documents with attachments; externally shared items.
Sensitive information types (SITs)	Filters for regulated or business-sensitive content using named sensitive information types or their identifiers. Example: Documents containing credit card numbers; emails with passport numbers; files with at least five SSNs and confidence above 85%.
Sensitivity labels	Filters based on the sensitivity labels already applied to content. Example: Documents labeled Confidential; items with the Public label.
Keywords	Searches for specific words or phrases that appear in the body text of an item. Example: Emails containing the words "merger" and "acquisition".
Intent of the file/data asset	Meaning-based search for broader themes or concepts rather than exact keywords. Example: Documents discussing strategies for improving employee retention; communications about frustration with project delays.

Example of a prompt that may deliver sharp results

A single request can combine multiple filter types in one query for sharper results.

Example: "PDF documents larger than 1 MB, labeled Confidential, that explain data privacy policies, from the last 90 days."

This prompt contains search parameters Identifiers, Sensitivity labels, Keywords, and Intent of the file/data asset.

Examples of prompts that may not deliver sharp results

Sometimes a request is too vague, contradictory, or asks for something the search service doesn't support. In those cases, you see an invalid result with a reason explaining why the request couldn't be interpreted.

Expand table

Prompt	Why it may not deliver sharp results
"Find stuff", "Sensitive data", "Emails"	Too vague
"What's the weather today?"	Unrelated to search
"Files created in 2025 and modified in the last week"	Conflicting time ranges
"Emails from the legal team", "messages from the finance team"	Abstract groups instead of concrete people
A request that doesn't yield any keywords, conditions, or topic to search on	Asks for something with no actionable detail at all

See also

- Using the data security posture agent.

Last updated on 06/23/2026

Get started with the Data Security Posture agent (preview) in Data Security Investigations

> **Important**
>
> Data Security Investigations uses generative artificial intelligence (AI), large language models, and orchestration in the analysis of data in your organization. Results generated by AI might not always be accurate or complete. While we strive to provide reliable and helpful information, AI systems can produce incorrect or false results. It is important to verify the information and use it with caution. Microsoft makes no warranties, express, implied, or statutory, as to the information provided by AI systems.

The Data Security Posture agent (preview) in Data Security Investigations helps your organization proactively surface credentials buried in data across your organization at scale. Rather than manually investigating individual items, the Posture agent automates credential scanning across Microsoft 365 data locations, including SharePoint sites, OneDrive accounts, Exchange mailboxes, and Teams content. After the agent completes scanning, it produces AI-generated risk assessments with confidence scores and supporting reasoning to help your security team quickly prioritize and remediate credential exposure risks.

> **Important**
>
> The Posture agent in Data Security Investigations is currently in preview. Features and functionality might change before general availability.

How the Posture agent works

The Posture agent in Data Security Investigations uses the same agent as Data Security Posture Management (DSPM) (preview). When you enable the Posture agent, you make it available in both DSPM and Data Security Investigations. If you already enabled the Posture agent in DSPM, you don't need to onboard again to use it in Data Security Investigations.

By using the Posture agent in Data Security Investigations, you can:

- **Scan for credentials at scale**: Run tenant-wide or scoped credential scanning tasks to discover exposed credentials across your organization's Microsoft 365 data.

- **Track tasks on a task board**: Monitor the progress of scanning tasks through a Kanban-style task board with statuses like *In progress* and *Ready for review*.
- **Review AI-generated risk assessments**: Examine AI-produced risk scores, confidence levels, and reasoning for flagged items to help prioritize remediation efforts.
- **Explore data with KQL queries**: Use the Data Explorer to run Kusto Query Language (KQL) queries against credential findings for advanced analysis and investigation.
- **Download reports**: Generate and download reports of credential scan results for offline review and compliance reporting.

Prerequisites

Before you can use the Posture agent in Data Security Investigations, confirm the following prerequisites are in place:

- **Permissions**: You must be assigned to one of the role groups that includes Posture agent access. For more information, see Permissions for deploying, running and viewing results from the Posture Agent.
- **Agent enabled**: You must enable the Posture agent in the Microsoft Purview portal. For more information, see Enable the Posture agent.
- **Security Compute Units (SCUs)**: Your organization must have Security Copilot SCUs provisioned for the Posture agent to run.

> **Tip**
>
> You don't need to configure or enable Data Security Investigations to use the Posture agent. The Posture agent is independent of the Data Security Investigations solution and doesn't require Data Security Investigations billing, permissions, or setup to be completed. As long as the Posture agent is enabled and your organization has SCUs provisioned, the Posture agent experience is available in the Data Security Investigations navigation. Other Data Security Investigations features, such as creating investigations and AI analysis, still require separate Data Security Investigations configuration and billing.

Enable the Posture agent

If your organization hasn't already enabled the Posture agent, use the following steps to enable it:

1. Go to the Microsoft Purview portal and sign in by using the credentials for a user account assigned the appropriate permissions.
2. Select **Agents** in the left navigation.

3. Select **Explore agents**, and then select **View details** for the Posture agent in Data Security Investigations.
4. Select **Set up** to enable the agent for your organization.

After you enable the Posture agent, it appears in Agents, DSPM, and Data Security Investigations. In Data Security Investigations, you can access the agent by selecting **Posture agent** in the left navigation.

> **Note**
>
> If you already enabled the Posture agent in DSPM, you don't need to complete these steps. The agent is automatically available in Data Security Investigations.

Create a credential scanning task

Create a credential scanning task so the Posture agent can scan your organization's data for exposed credentials. You can scope the scan to your entire tenant or narrow it to specific users, sites, or groups.

Complete the following steps to create a credential scanning task:

1. Go to Data Security Investigations in the Microsoft Purview portal and sign in by using the credentials for a user account assigned Data Security Investigations permissions.
2. Select **Posture agent** in the left navigation.
3. Select **Assign to agent** to create a credential scanning task.
4. For the **Credential Scanning Task**, enter a name for the task in the **Task name** field.
5. Configure the data sources for the scan. You can select a tenant-wide scan or narrow the scope to your organization, specific sites, users, mailboxes, or groups.
6. Select **Save**.
7. In the **Additional context for AI**, use natural language guidance to help AI scope the agents to applicable areas.
8. Select **Create** to start the task.

After you create the task, it appears in the **In progress** column on the task board.

> **Tip**
>
> If you realize you scoped a task incorrectly after it starts, you can stop the task and create a new one with the correct scope.

Monitor tasks on the task board

The Posture agent uses a Kanban-style task board to help you track the progress of credential scanning tasks. Each task appears in one of the following columns:

- **In progress**: The agent actively scans the specified data locations and analyzes content.
- **Ready for review**: The agent finished scanning and the results are ready for a human reviewer to examine.
- **Closed**: A human reviewer has marked a tasked as completed.

The task board displays key information for each task, including the scan scope, the number of locations scanned, and the number of items analyzed. You can also assign a security operations center (SOC) administrator to review the results for a specific task.

Review scan results

After a credential scanning task finishes and moves to the **Ready for review** column, you can review the results to identify and prioritize credential exposure risks.

Summary view

The summary view provides a high-level overview of the credential scan results. In the summary view, you can:

- **Sort by risk level**: View results sorted by *High*, *Medium*, or *Low* risk to prioritize the most critical findings.
- **Group by credential category**: Organize results by credential type to understand the distribution of exposed credentials.
- **Review AI reasoning**: For each flagged item, the AI provides its reasoning for the assigned risk level, helping you understand why a specific item is flagged as high, medium, or low risk.
- **Download reports**: Generate a report of the scan results for offline review, compliance reporting, or sharing with stakeholders.

Item details

Select a specific item in the scan results to view detailed information, including:

- The detected credential type.
- The data location where the credential was found.
- The AI-generated risk and confidence scores.

- The reasoning behind the risk classification.

Explore data with Data Explorer

Data Explorer provides advanced analysis capabilities by allowing you to run Kusto Query Language (KQL) queries against the credential findings from the Posture agent scans. This experience is similar to advanced hunting in Microsoft Defender XDR and provides a schema-based metadata framework for querying credential data.

To use Data Explorer, complete the following steps:

1. Go to Data Security Investigations in the Microsoft Purview portal.
2. Select **Posture agent** in the left navigation.
3. Select a completed task, and then select the **Data Explorer** tab.
4. Write and run KQL queries against the credential findings schema.

> **Tip**
>
> If you're new to KQL, see **Kusto Query Language overview** to learn the basics of writing queries.

ⓘ **Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 03/23/2026

Learn about Microsoft Sentinel in Microsoft Purview

Microsoft Sentinel is a scalable, cloud-native security information and event management (SIEM) solution that delivers cost-efficient security across multicloud and multiplatform environments with built-in AI, automation, threat intelligence, and a modern data lake and graph architecture. The Microsoft Sentinel data lake and graph are tenant-wide repositories for collecting, storing, and managing large volumes of security-related data from various sources. These solutions integrate with Microsoft Purview solutions and support data risk graphs.

> ⓘ **Important**
>
> **Data risk graph is being deprecated.** The data risk graph in Microsoft Purview Insider Risk Management will be retired on **November 24, 2026**. After that date, the data risk graph experience will no longer be available.

Data risk graph capabilities are available in the following Microsoft Purview solutions:

- Microsoft Purview Data Security Investigations (preview)
- Microsoft Purview Insider Risk Management

Partner integrations are available in the following Microsoft Purview solution:

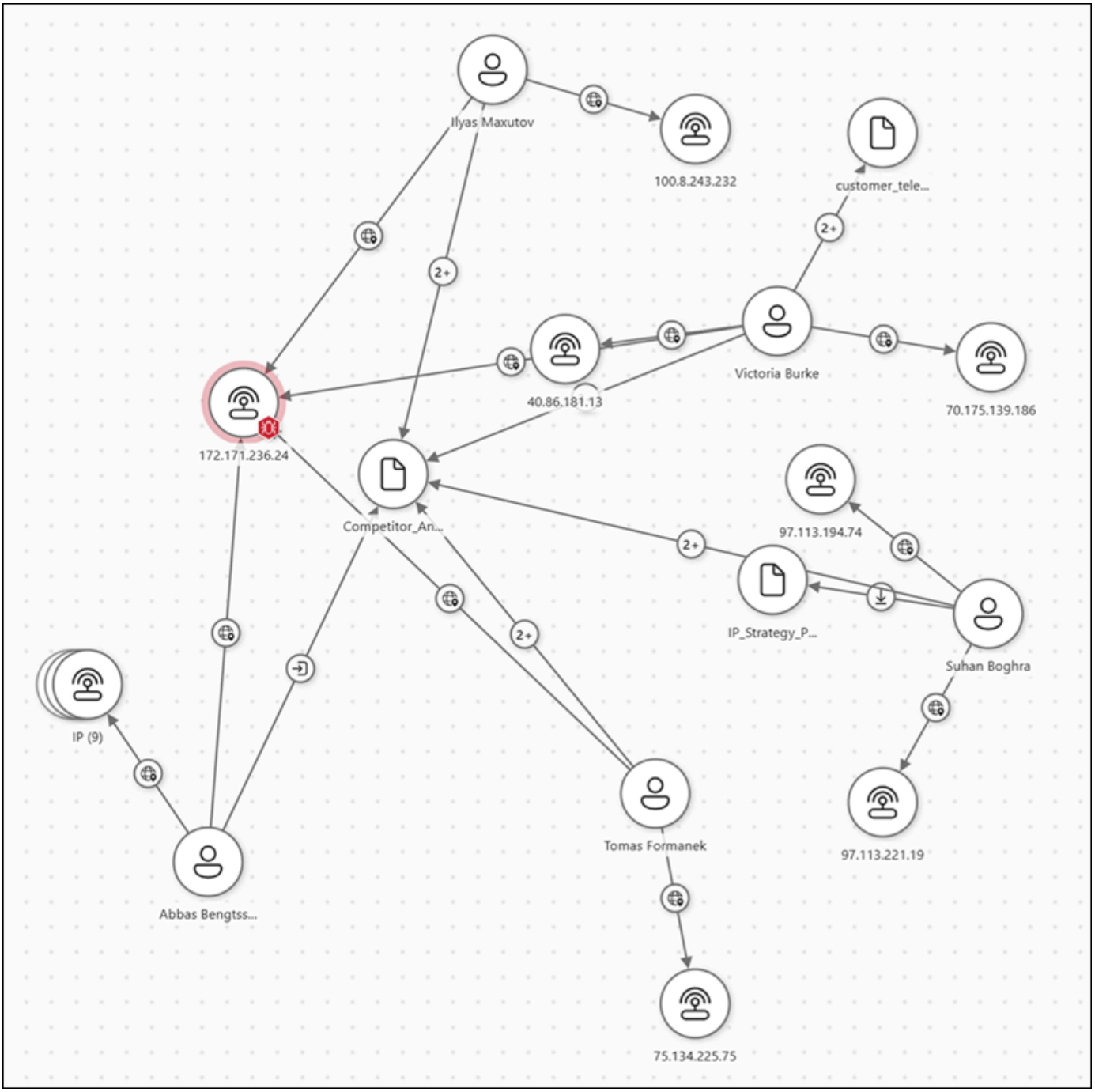
- Microsoft Purview Data Security Posture Management

Data risk graph

Powered by Microsoft Sentinel graph, data risk graphs in Microsoft Purview solutions allow you to view connections between impacted assets, users, and their activities in an interactive graph experience. Microsoft Sentinel graph is a unified graph analytics capability within Microsoft Sentinel that powers graph-based experiences across Microsoft Purview solutions. For more information about Microsoft Sentinel graph, see What is Microsoft Sentinel graph?

Data risk graphs natively represent the real-world web of users, devices, cloud resources, data flows, activities, and attacker actions associated with activities identified in Microsoft Purview solutions. By connecting activities to assets, analysts can:

- **Understand context**: View a user's role and file locations to assess whether behavior is typical or suspicious.
- **Assess impact**: Visualize which assets might be impacted by tracing activity paths within sensitive sites.
- **Collaborate effectively**: Use the data risk graph to communicate findings with stakeholders by showing connections between users, actions, and assets.



Data risk graph controls

Data risk graphs in Microsoft Purview solutions have the following controls to help you discover and work with information about item relationships and connected nodes:

- **Movable nodes**: Change the presentation of the graph by moving nodes to customize the view of nodes, relationships, and connections.
- **Layers**: The **Vulnerability** layer is presented as the default **Insights layer** in a data risk graph. To change layers, select the **Layers** control in the lower-left corner of a graph.
- **Maximize the graph view**: To quickly maximize a graph, select the **Zoom to fit** control in the lower-left corner of a graph.
- **Scale the graph to any size**: To enlarge or reduce the size of the graph, select the + or - controls in the lower-left corner of a graph. Select the **Settings** control in the lower-left corner of a graph to enable or disable the **Zoom slider** control.

Data risk graph features

Data risk graphs in Microsoft Purview solutions include the following features to help you discover and work with information about item relationships and connected nodes:

- **Nodes**: Nodes include users, sites, files, and IP addresses associated with each other.
- **Relationships**: Connections between nodes describing how they interact with each other. Badges on each relationship describe actions and meanings. For example, badges might include *Signed in from*, *Downloaded*, *Uploaded*, and more.

Prerequisites and onboarding changes

To use the data risk graph in supported Microsoft Purview solutions, you must onboard to the Microsoft Sentinel data lake and graph and meet specific requirements. Additionally, when you onboard to data lake and graph, the onboarding process makes changes in several areas and services in your organization. The onboarding process includes changes to your existing data lake, primary and other connected workspaces in Microsoft Defender, and more.

Data risk graphs use a pay-as-you-go billing model to ingest and store data, but data risk graph usage doesn't incur any extra charges. If data needed to populate the data risk graph is already enabled or stored, you don't incur any extra charges.

For step-by-step guidance for all prerequisites, billing information, and changes made in your onboarding process, see the following articles:

- Prerequisites for Microsoft Sentinel data lake and Microsoft Sentinel graph
- Changes made when onboarding to Microsoft Sentinel data lake and Microsoft Sentinel graph

Configure data risk graph

After you meet the prerequisites for Microsoft Sentinel data lake and Microsoft Sentinel graph and understand the changes made during the onboarding process, you're ready to configure data risk graph in Microsoft Purview solutions. The data lake for your organization stores all activity and asset data from Microsoft Purview data risk graphs.

To configure data risk graph in Microsoft Purview solutions, see the following articles:

- Configure data risk graph in Data Security Investigations (preview)
- Configure data risk graph in Insider Risk Management

Partner integrations

Set up Microsoft Sentinel data lake to integrate partner solutions and centralize security posture monitoring across all data sources. To see the available partner solutions and then select one from Data Security Posture Management, you must first connect to the Microsoft Sentinel content hub. The Microsoft Sentinel content hub is the centralized catalog where you discover and deploy partner solutions. Initial partners include Cyera (for Snowflake), BigID (for Databricks), OneTrust (for Google Cloud Platform), and Varonis (for Salesforce).

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 08/27/2026

Data risk graph in Data Security Investigations

> **Important**
>
> We are retiring Data risk graph (Explore insights) in Microsoft Purview Data Security Investigations (DSI). **Beginning September 30, 2026**, organizations will no longer be able to configure Data risk graph. The feature will be fully retired and no longer available after **November 30, 2026**.

Data risk graphs in Data Security Investigations provide a visual investigation experience that combines asset and activity data into a single view. Powered by Microsoft Sentinel integration, it summarizes the previous 30 days of activity for any file scoped in your investigation. It helps you identify risky user accounts that interacted with content of interest and analyze the sequence of events leading up to recent incidents. By enriching analysis with activity insights, data risk graphs can help you resolve your data security incidents faster and with greater confidence.

For more information about Microsoft Sentinel integration, see Learn about Microsoft Sentinel in Microsoft Purview.

> **Note**
>
> Admin units aren't supported in the data risk graph. If you're scoped to an admin unit, data doesn't appear in data risk graphs.

Supported activities in data risk graph

The data risk graph currently supports the following exfiltration activities:

- Anonymous links created in SharePoint or OneDrive
- Anonymous links used via SharePoint or OneDrive
- Company links created in SharePoint or OneDrive
- File downloads from SharePoint and OneDrive
- Files renamed in SharePoint and OneDrive

Before you begin

Before you can use data risk graphs in Data Security Investigations, complete the following steps:

- Learn about pay-as-you-go billing in Microsoft Sentinel for your organization.
- Configure the prerequisites for Microsoft Sentinel data lake and Microsoft Sentinel graph.
- Review the changes made when onboarding to Microsoft Sentinel data lake and Microsoft Sentinel graph.
- Create an investigation in Data Security Investigations.

Configure data risk graph

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview **roles and permissions**.

After you complete the prerequisites, complete the following steps to configure data risk graph in Data Security Investigations:

1. Go to Data Security Investigations in the Microsoft Purview portal with an account assigned the *Global Administrator* or *Security Administrator* role.

2. Select **Overview** in the left navigation.

3. In the **Setup tasks** section, select **Set up data lake and data risk graph**.

   If you don't have the correct permissions to set up the data lake, a notification appears to contact a global or billing administrator in your organization.

4. On the **Setup** tab on **Set up data lake & risk graph (preview)**, configure the following settings:
   a. **Subscription**: Enter the Azure subscription you want to use. Your selected subscription must be valid and accessible. You must be a subscription owner.
   b. **Resource group**: Enter the resource group you want to use. Your selected resource group must be valid and accessible.

   > **Important**

After the data lake is provisioned for a specific Azure subscription and resource group, you can't migrate it to a different subscription or resource group.

5. Select **Setup**.

> **Important**
>
> When the data risk graph is first created, it includes the most recent seven days of data. As the data risk graph refreshes over time, the look-back window expands up to a maximum of 30 days. It might take some time for the data risk graph to reach the full 30-day window, so expect gradual growth after initial setup.

The set up process begins and onboarding might take up to 60 minutes to complete. You can close the setup panel while the process is running. After the onboarding process completes, you see *Complete* on **Set up data lake & risk graph (preview)**. The data risk graph shows events that occur after your Microsoft Sentinel data lake was created or onboarded.

Initial processing of data and data risk graph availability for an investigation might take 24-48 hours.

> **Important**
>
> You always have one data lake. If you already onboarded to the data lake with another Microsoft service, your existing data lake is used. If you never onboarded to the data lake, onboarding in Data Security Investigations enables Microsoft Sentinel data lake and graph in the Defender portal.

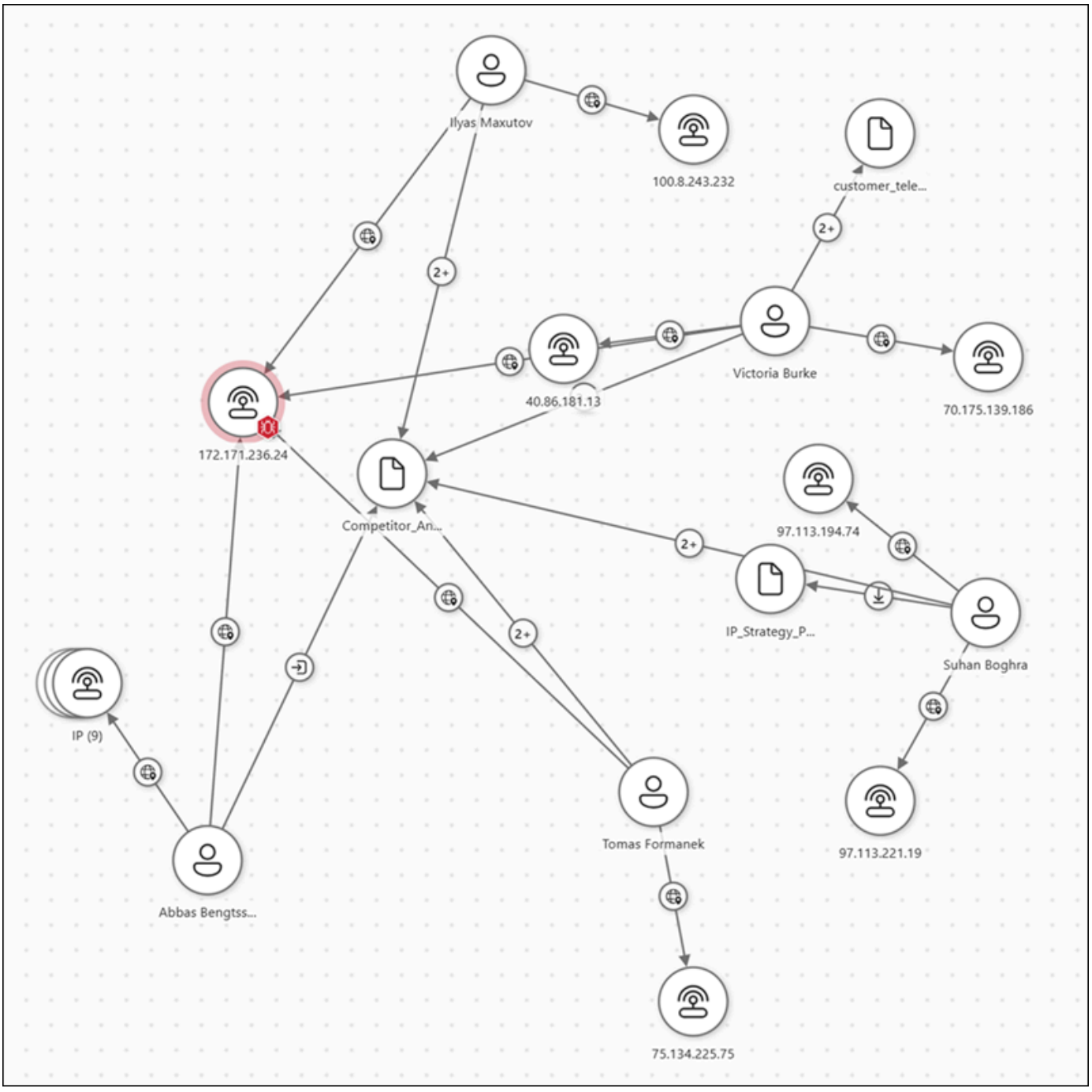
Use data risk graph

Data risk graph in Data Security Investigations uniquely visualizes correlations between impacted data, users, and their activities. It provides critical context to guide mitigation and next steps. For example, upon uncovering a highly sensitive document, data risk graphs give you visibility into which users downloaded it or if they accessed it from a risky IP address. This visibility lets you uncover new nodes to a data security incident, like additional users or new content that requires investigation.

To view the data risk graph for an investigation, you must be assigned to at least one built-in Data Security Investigations role group:

- *Data Security Investigations Admins*
- *Data Security Investigations Investigators*
- *Data Security Investigations Reviewers*

For more information about role groups, see Assign permissions in Data Security Investigations.



To use the data risk graph in Data Security Investigations, complete the following steps:

1. Go to Data Security Investigations in the Microsoft Purview portal with an account assigned to a Data Security Investigations role group.

2. Select **Investigations** in the left navigation.

3. Select an investigation to review, then select one of the following options:

a. On **Summary**, select **Manage scope** on the **Investigation scope** card.
b. Select the **Analysis** tab to list items included in the **Investigation scope**.

4. Select items (up to 100) included in the scope, and then select **Explore insights**.

> ⓘ Note
>
> **Explore insights** isn't available unless you configure Data Security Investigations for Microsoft Sentinel data lake and graph.

5. Filter the time as applicable and select individual data risk graph nodes for more information about each connection during triage.

6. Expand relationships in the graph by selecting the + icon on users or assets.

The data risk graph contains several nodes. Selecting a node shows details about the node:

- **User details**: Shows information about users associated with the investigation. This information includes organization information for each user, including User Principal Name (UPN), labels, location, job title, and more.
- **Relationships**: Shows information about starting and ending connections between nodes. This information includes starting and last connection dates, relationship types, and more.
- **IP addresses**: Shows information about each IP address node.
- **Data details**: Shows information about files and sites. This information includes the object location, type, labels, and more.

You can view data risk graph information for users based on the last 30 days of activity. This duration is fixed and can't be extended.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 09/02/2026

Data risk graph in Insider Risk Management

> **Important**
>
> Microsoft Purview Insider Risk Management correlates various signals to identify potential malicious or inadvertent insider risks, such as IP theft, data leakage, and security violations. Insider Risk Management enables customers to create policies to manage security and compliance. Built with privacy by design, users are pseudonymized by default, and role-based access controls and audit logs are in place to help ensure user-level privacy.

Data risk graphs in Insider Risk Management provide a visual investigation experience that combines asset and activity data into a single view. Powered by Microsoft Sentinel integration, it summarizes alert-related activity by a potential insider over the past 30 days, helping analysts triage alerts by revealing hidden contextual connections.

> **Important**
>
> **Data risk graph is being deprecated.** The data risk graph in Microsoft Purview Insider Risk Management will be retired on **November 24, 2026**. After that date, the data risk graph experience will no longer be available.

For more information about Microsoft Sentinel integration, see Learn about Microsoft Sentinel in Microsoft Purview.

> **Important**
>
> You cannot use this feature with the anonymized usernames privacy setting enabled. If you have anonymized user names, the graph will not load.

> **Note**
>
> Admin units aren't supported in the data risk graph. If you're scoped to an admin unit, data doesn't appear in data risk graphs.

Supported activities in data risk graph

The data risk graph currently supports the following exfiltration activities:

- Anonymous links created in SharePoint or OneDrive
- Anonymous links used via SharePoint or OneDrive
- Company links created in SharePoint or OneDrive
- File downloads from SharePoint and OneDrive
- Files renamed in SharePoint and OneDrive

For more information about exfiltration detection indicators, see Configure policy indicators in Insider Risk Management.

Before you begin

Before you can use data risk graphs in Insider Risk Management, complete the following steps:

- Learn about pay-as-you-go billing in Microsoft Sentinel for your organization.
- Configure the prerequisites for Microsoft Sentinel data lake and Microsoft Sentinel graph
- Review the changes made when onboarding to Microsoft Sentinel data lake and Microsoft Sentinel graph
- Configure at least one Insider Risk Management policy for supported data exfiltration activities.

Configure data risk graph

> **Important**
>
> You always have one data lake. If you already onboarded to the data lake with another Microsoft service, your existing data lake is used. If you never onboarded to the data lake, onboarding in Insider Risk Management enables Microsoft Sentinel data lake and graph in the Defender portal.

After you complete the prerequisites and understand the onboarding changes that Microsoft Sentinel data lake and graph make to your organization, complete the following steps to configure data risk graph in Insider Risk Management:

> **Important**

Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

1. Go to the Microsoft Purview portal with an account assigned the *Global Administrator* or *Security Administrator* role.
2. Select the **Insider Risk Management** solution card, then select **Recommended actions** in the left nav.
3. In the **Comprehensive protection** section, select **Set up data lake and data risk graph**.

   If you don't have the correct permissions to set up the data lake, a notification appears to contact a global or billing administrator in your organization.
4. On the **Setup** tab on **Set up data lake & risk graph**, configure the following settings:
   a. **Subscription**: Enter the Azure subscription you want to use. Your selected subscription must be valid and accessible and you must be the subscription owner.
   b. **Resource group**: Enter the resource group you want to use. Your selected resource group must be valid and accessible.

   > **Important**
   >
   > After the data lake is provisioned for a specific Azure subscription and resource group, you can't migrate it to a different subscription or resource group.
5. Select **Setup**.

The data lake and data risk graph setup process begins, and onboarding might take up to 60 minutes to complete. You can close the setup panel while the process is running. After the onboarding process completes, you see *Complete* on **Set up data lake & risk graph**. The data risk graph shows events that occur after your Microsoft Sentinel data lake was created or onboarded.

Initial processing of data and data risk graph availability for an investigation can take 24–48 hours.

> **Important**
>
> When the data risk graph is first created, it includes the most recent seven days of data. As the data risk graph refreshes over time, the look-back window expands up to a maximum of

30 days. It might take some time for the data risk graph to reach the full 30-day window, so expect gradual growth after initial setup.

Use data risk graph

Data risk graph in Insider Risk Management uniquely visualizes correlations between impacted data, users, and their activities. It provides critical context to guide mitigation and next steps. For example, when you uncover an alert related to a highly sensitive document, data risk graphs give you visibility into which users downloaded it or if they accessed it from an IP address flagged as risky by your organization's security signals. This visibility lets you uncover new nodes to a data security incident, like additional users or new content related to an alert.

To view the data risk graph for an alert, you must be assigned the *Insider Risk Management Graph Reader* role. This role is included by default in several Insider Risk Management built-in role groups. For more information, see Assign permissions in Insider Risk Management.

> **Important**
>
> All user names are visible in data risk graph, even if pseudonymization is enabled in Insider Risk Management privacy settings.

To use the data risk graph in Insider Risk Management, complete the following steps:

1. Go to the Microsoft Purview portal with an account assigned the *Insider Risk Management Graph Reader* role.
2. Select the **Insider Risk Management** solution card, then select **Alerts** in the left nav.
3. Select an alert to review, then select the **Data risk graph** tab.
4. Filter the time as applicable and select individual data risk graph nodes for more information about each connection during triage.
5. Expand relationships in the graph by selecting the + icon on users or assets.

The data risk graph contains several nodes. Selecting a node shows details about the node:

- **User details**: Shows information about users associated with the alert. This information includes organization information for each user, including User Principal Name (UPN), labels, location, job title, and more.
- **Data details**: Shows information about files and sites. This information includes the object location, type, labels, and more.

You can view data risk graph information for users based on the last 30 days of activity. This fixed duration isn't extendable.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 08/27/2026

Export policy configuration in Microsoft Purview

The **Export to zip** option on the policies page downloads a point-in-time snapshot of all your policy configurations as a ZIP file in XML format. The export captures the full configuration details for every policy in the current solution at the moment of export, including rules, conditions, actions, scoping, and priority settings where relevant. For example, Label policies include the labels in the export, but not rules or conditions as they don't have any. You can't select individual policies for export—the ZIP file always contains all policies.

> **Important**
>
> The exported ZIP file is a read-only, point-in-time snapshot of your policy configuration. It's intended for reference only. You can't import it back into your tenant to restore or create policies, and it doesn't automatically update to reflect subsequent configuration changes.

> **Note**
>
> The **Export to CSV** option exports only the summary information shown in the policies table, such as policy name, status, and last modified date.

Common scenarios

Common scenarios for using the ZIP export include:

- **Support requests** — Attach the ZIP file to a support ticket so that support engineers can review your complete policy configuration without a live troubleshooting session.
- **Configuration backup** — Save periodic snapshots of your policy configuration for version control and change tracking purposes.
- **Configuration analysis** — Extract the XML file and analyze your policies locally by using PowerShell or by uploading the file to Microsoft 365 Copilot.

Export your policy configuration

1. Sign in to the Microsoft Purview portal.

2. Navigate to the policies page for the solution whose policies you want to export. If exporting is supported for that solution, the **Export to zip** option will be available.
3. Select **Export to zip**.
4. Save the ZIP file when prompted. The file downloads to your default downloads folder.
5. Extract the ZIP file to access the XML configuration file.

View the exported configuration by using PowerShell

You can use the standard PowerShell `Import-Clixml` cmdlet to deserialize the exported XML file into PowerShell objects for local analysis. This cmdlet reads the file locally and doesn't connect to any service.

1. Open a PowerShell window.

2. Import the XML file into a variable by using the `Import-Clixml` cmdlet. Replace the path with the location of your extracted XML file.

   PowerShell

   ```
   $p = Import-Clixml -Path "<path to extracted XML file>"
   ```

3. List the available properties in the exported configuration.

   PowerShell

   ```
   $p | Get-Member -MemberType Properties
   ```

4. View the data for a specific property by using `Format-Table` for a readable summary.

   PowerShell

   ```
   $p.<PropertyName> | Format-Table
   ```

   For example, if the export contains DLP policies:

   PowerShell

   ```
   $p.DlpCompliancePolicies | Format-Table
   $p.DlpComplianceRules | Format-Table
   ```

> **Tip**

The property names in the exported XML vary by solution. Use `$p | Get-Member -MemberType Properties` to discover which properties are available in your export.

See also

- Export-PurviewConfig

ⓘ **Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 05/27/2026

Microsoft Purview data security and compliance protections for generative AI apps

> *Microsoft Purview service description*

Use Microsoft Purview to mitigate and manage the risks associated with AI usage, and implement corresponding protection and governance controls.

The following sections on this page provide an overview of Data Security Posture Management for AI and the Microsoft Purview capabilities that provide additional data security and compliance controls to accelerate your organization's adoption of Copilots and agents, and other generative AI apps.

In some Microsoft Purview solutions, you might see the supported AI apps grouped by the following category names:

- **Copilot experiences and agents** or **Microsoft Copilot experiences** for supported Copilots and agents that include:
  - Microsoft 365 Copilot
  - Security Copilot
  - Copilot in Fabric
  - Copilot Studio

- **Enterprise AI apps** for non-Copilot AI apps and agents connected to your organization through Entra registration, data connectors, Microsoft Foundry, and other methods, and include:
  - Microsoft Foundry
  - Entra-registered AI apps
  - Anthropic Claude (Enterprise)
  - ChatGPT Enterprise

- **Other AI apps** that are detected through browser activity and categorized as "Generative AI" in the Defender for Cloud Apps catalog. This category can include AI apps and agents from the other two categories but also uniquely includes AI apps and agents from third-party LLMs, such as:

- ChatGPT
- Google Gemini
- Microsoft Copilot (consumer version)
- DeepSeek

For a breakdown of Microsoft Purview security and compliance supported capabilities for AI interactions by app, see the additional pages identified in the following table. Where these AI apps support agents, they inherit the same security and compliance capabilities as their parent AI app. However, for a quick summary, see Use Microsoft Purview to manage data security & compliance for AI agents.

Expand table

Copilot experiences and agents	Enterprise AI apps	Other AI apps
Microsoft 365 Copilot & Microsoft 365 Copilot Chat	Entra-registered AI apps	Other AI apps
Microsoft Security Copilot	Microsoft Foundry	
Copilot in Fabric	ChatGPT Enterprise	
Microsoft Copilot Studio	Anthropic Claude (Enterprise)	
Microsoft Facilitator		
Channel Agent in Teams		

For a list of supported Microsoft Purview security and compliance supported capabilities for Microsoft Agent 365, see Use Microsoft Purview to manage data security & compliance for Microsoft Agent 365.

If you're new to Microsoft Purview, you might also find an overview of the product helpful: Learn about Microsoft Purview.

DSPM and DSPM for AI (classic)

Use Data Security Posture Management or Data Security Posture Management for AI (classic) as your front door to discover, secure, and apply compliance controls for AI usage across your enterprise. Both DSPM versions use existing controls from Microsoft Purview information protection and compliance management with easy-to-use graphical tools and reports to quickly gain insights into AI use within your organization. With personalized recommendations, and one-click policies help you protect your data and comply with regulatory requirements.

Microsoft Purview strengthens information protection for AI apps

Because of the power and speed AI can proactively surface content, generative AI amplifies the problem and risk of oversharing or leaking data. Learn how information protection capabilities from Microsoft Purview can help to strengthen your existing data security solutions.

Sensitivity labels and AI interactions

AI apps that Microsoft Purview support use existing controls to ensure that data stored in your tenant is never returned to the user or used by a large language model (LLM) if the user doesn't have access to that data. When the data has sensitivity labels from your organization applied to the content, there's an extra layer of protection:

- When a file is open in Word, Excel, PowerPoint, or similarly an email or calendar event is open in Outlook, the sensitivity of the data is displayed to users in the app with the label name and content markings (such as header or footer text) that have been configured for the label. Loop components and pages also support the same sensitivity labels.
- When the sensitivity label applies encryption, users must have the EXTRACT usage right, as well as VIEW, for the AI apps to return the data.
- This protection extends to data stored outside your Microsoft 365 tenant when it's open in an Office app (data in use). For example, local storage, network shares, and cloud storage.

> **Tip**
>
> If you haven't already, we recommend you enable sensitivity labels for SharePoint and OneDrive and also familiarize yourself with the file types and label configurations that these services can process. When sensitivity labels aren't enabled for these services, the encrypted files that Copilot and agents can access are limited to data in use from Office apps on Windows.
>
> For instructions, see **Enable sensitivity labels for Office files in SharePoint and OneDrive**.

If you're not already using sensitivity labels, see Get started with sensitivity labels.

Encryption without sensitivity labels and AI interactions

Even if a sensitivity label isn't applied to content, services and products might use the encryption capabilities from the Azure Rights Management service. As a result, AI apps can still check for the VIEW and EXTRACT usage rights before returning data and links to a user, but there's no automatic inheritance of protection for new items.

> **Tip**
>
> You'll get the best user experience when you always use sensitivity labels to protect your data, and encryption is applied by a label.

Examples of products and services that can use the encryption capabilities from the Azure Rights Management service without sensitivity labels:

- Microsoft Purview Message Encryption
- Microsoft Information Rights Management (IRM)
- Microsoft Rights Management connector
- Microsoft Rights Management SDK

For other encryption methods that don't use the Azure Rights Management service:

- S/MIME protected emails won't be returned by Copilot, and Copilot isn't available in Outlook when an S/MIME protected email is open.
- Password-protected documents can't be accessed by AI apps unless they're already opened by the user in the same app (data in use). Passwords aren't inherited by a destination item.

As with other Microsoft 365 services, such as eDiscovery and search, items encrypted with Microsoft Purview Customer Key or your own root key (BYOK) are supported and eligible to be returned by Copilot.

Data loss prevention and AI interactions

Microsoft Purview Data Loss Prevention (DLP) helps you identify sensitive items across Microsoft 365 services and endpoints, monitor them, and helps protect against leakage of those items. It uses deep content inspection and contextual analysis to identify sensitive items and it enforces policies to protect sensitive data such as financial records, health information, or intellectual property.

Windows computers that are onboarded to Microsoft Purview can be configured for Endpoint data loss prevention (DLP) policies that warn or block users from sharing sensitive information

with third-party generative AI sites that are accessed via a browser. For example, a user is prevented from pasting credit card numbers into ChatGPT, or they see a warning that they can override. For more information about the supported DLP actions and which platforms support them, see the first two rows in the table from Endpoint activities you can monitor and take action on.

Insider Risk Management and AI interactions

Microsoft Purview Insider Risk Management helps you detect, investigate, and mitigate internal risks such as IP theft, data leakage, and security violations. It leverages machine learning models and various signals from Microsoft 365 and third-party indicators to identify potential malicious or inadvertent insider activities. The solution includes privacy controls like pseudonymization and role-based access, ensuring user-level privacy while enabling risk analysts to take appropriate actions.

Use the Risky AI usage policy template to detect risky usage that includes prompt injection attacks and accessing protected materials. Insights from these signals are integrated into Microsoft Defender XDR to provide a comprehensive view of AI-related risks.

Data classification and AI interactions

Microsoft Purview data classification provides a comprehensive framework for identifying and tagging sensitive data across various Microsoft services, including Office 365, Dynamics 365, and Azure. Classifying data is often the first step to ensure compliance with data protection regulations and safeguard against unauthorized access, alteration, or destruction. You can use built-in system classifications or create your own.

Sensitive information types and trainable classifiers can be used to find sensitive data in user prompts and responses when they use AI apps. The resulting information then surfaces in Microsoft Purview Reports overview and activity explorer in DSPM for AI and the **AI activities** tab in activity explorer from the preview version of DSPM.

Microsoft Purview supports compliance management for AI apps

Interactions using supported AI apps can be monitored for each user in your tenant. As such, together with data classification, you can use Microsoft Purview's auditing, communication compliance, eDiscovery with content search, and automatic retention and deletion capabilities from Data Lifecycle Management to manage this AI usage.

Auditing and AI interactions

Microsoft Purview Audit solutions provide comprehensive tools for searching and managing audit records of activities performed across various Microsoft services by users and admins, and help organizations to effectively respond to security events, forensic investigations, internal investigations, and compliance obligations.

Like other activities, prompts and responses are captured in the unified audit log. Events include how and when users interact with the AI app, and can include in which Microsoft 365 service the activity took place, and references to the files stored in Microsoft 365 that were accessed during the interaction. If these files have a sensitivity label applied, that's also captured.

These events flow into activity explorer in the **AI activities** tab in the current version of DSPM and in DSPM for AI, where the data from prompts and responses can be displayed. You can also use the **Audit** solution from the Microsoft Purview portal to search and find these auditing events.

For more information, see Audit logs for Copilot and AI activities.

Communication compliance and AI interactions

Microsoft Purview Communication Compliance provides tools to help you detect and manage regulatory compliance and business conduct violations across various communication channels, which include user prompts and responses for AI apps. It's designed with privacy by default, pseudonymizing usernames and incorporating role-based access controls. The solution helps identify and remediate inappropriate communications, such as sharing sensitive information, harassment, threats, and adult content.

To learn more about using communication compliance policies for AI apps, see Configure a communication compliance policy to detect for generative AI interactions.

eDiscovery with content search and AI interactions

Microsoft Purview eDiscovery lets you identify and deliver electronic information that can be used as evidence in legal cases. The eDiscovery tools in Microsoft Purview support searching for content in Exchange Online, OneDrive for Business, SharePoint Online, Microsoft Teams, Microsoft 365 Groups, and Viva Engage teams. You can then prevent the information from deletion and export the information.

Because user prompts and responses for AI apps are stored in a user's mailbox, you can create a case and use search when a user's mailbox is selected as the source for a search query. For

example, select and retrieve this data from the source mailbox by selecting from the query builder **Add condition** > **Type** > **Contains any of** > **Edit** > *Copilot activity*. This query condition includes all Copilot and other AI application activity.

After the search is refined, you can export the results or add to a review set. You can review and export information directly from the review set.

For more information about identifying and deleting user AI interaction data, see Search for and delete Copilot data in eDiscovery.

Data Lifecycle Management and AI interactions

Microsoft Purview Data Lifecycle Management provides tools and capabilities to manage the lifecycle of organizational data by retaining necessary content and deleting unnecessary content. These tools ensure compliance with business, legal, and regulatory requirements.

Use retention policies to automatically retain or delete user prompts and responses for AI apps. For detailed information about this retention works, see Learn about retention for Copilot & AI apps.

As with all retention policies and holds, if more than one policy for the same location applies to a user, the principles of retention resolve any conflicts. For example, the data is retained for the longest duration of all the applied retention policies or eDiscovery holds.

Compliance Manager and AI interactions

Microsoft Purview Compliance Manager is a solution that helps you automatically assess and manage compliance across your multicloud environment. Compliance Manager can help you throughout your compliance journey, from taking inventory of your data protection risks to managing the complexities of implementing controls, staying current with regulations and certifications, and reporting to auditors.

To help you keep compliant with AI regulations, Compliance Manager provides regulatory templates to help you assess, implement, and strengthen your compliance requirements for all generative AI apps. For example, monitoring AI interactions and preventing data loss in AI applications. For more information, see Assessments for AI regulations.

Other documentation to help you secure and manage generative AI apps

Blog post announcement: Accelerate AI adoption with next-gen security and governance capabilities

Microsoft 365 Copilot:

- Microsoft 365 Copilot documentation
- Apply principles of Zero Trust to Microsoft 365 Copilot

Related resources:

- Secure Generative AI with Microsoft Entra
- Govern AI apps and data for regulatory compliance
- Blog posts (April 2025):
  - How to deploy Microsoft Purview DSPM for AI to secure your AI apps
  - How to use DSPM for AI Data Risk Assessment to Address Internal Oversharing

---

Last updated on 05/27/2026

Use Microsoft Purview to manage data security & compliance for Microsoft 365 Copilot & Microsoft 365 Copilot Chat

> *Microsoft Purview service description*

Use the following sections to identify the Microsoft Purview capabilities that are supported for AI interactions with Microsoft 365 Copilot & Microsoft 365 Copilot Chat, and some get started recommendations for you to manage these AI interactions for security and compliance.

Capabilities supported

Use the following table to see at a glance the Microsoft Purview capabilities that are supported with Microsoft 365 Copilot & Microsoft 365 Copilot Chat.

Expand table

Capability or solution in Microsoft Purview	Supported for AI interactions
DSPM and DSPM for AI (classic)	✓
Auditing	✓
Data classification	✓
Sensitivity labels	✓
Encryption without sensitivity labels	✓
Data loss prevention	✓
Insider Risk Management	✓
Communication compliance	✓
eDiscovery	✓
Data Lifecycle Management	✓
Compliance Manager	✓

DSPM and DSPM for AI (classic)

Use Data Security Posture Management or Data Security Posture Management for AI (classic) as your front door to discover, secure, and apply compliance controls for AI usage across your enterprise. Both DSPM versions use existing controls from Microsoft Purview information protection and compliance management with easy-to-use graphical tools and reports to quickly gain insights into AI use within your organization. With personalized recommendations, and one-click policies help you protect your data and comply with regulatory requirements.

AI app-specific information:

- Data risk assessments help you identify and fix issues that could results in oversharing of data. From the recommendations:
  - **Protect your data from potential oversharing risks** for the default weekly data risk assessment.
  - **Protect sensitive data references in Copilot and agent responses** for a custom data risk assessment.
- Recommendation: **Get guided assistance to AI regulations**, which uses control-mapping regulatory templates from Compliance Manager.
- Additional guidance from **Overview** > **Microsoft 365 Copilot** view.
- One-click policies available:
  - **Sensitivity labels and policies** from the recommendation or remediation action **Protect your data with sensitivity labels**.
  - **DSPM for AI - Detect risky AI usage** from the recommendation or remediation action **Detect risky interactions in AI apps**.
  - **DSPM for AI - Unethical behavior in AI apps** from the recommendation or remediation action **Detect unethical behavior in AI**.
  - **DSPM for AI - Protect sensitive data from Copilot processing** from the recommendation or remediation action **Protect items with sensitivity labels from Microsoft 365 Copilot and agent processing**
  - **DSPM for AI - Detect sensitive info shared with AI via network** from the recommendation or remediation action **Extend insights into sensitive data in AI app interactions**.

Auditing and AI interactions

Microsoft Purview Audit solutions provide comprehensive tools for searching and managing audit records of activities performed across various Microsoft services by users and admins, and

help organizations to effectively respond to security events, forensic investigations, internal investigations, and compliance obligations.

Like other activities, prompts and responses are captured in the unified audit log. Events include how and when users interact with the AI app, and can include in which Microsoft 365 service the activity took place, and references to the files stored in Microsoft 365 that were accessed during the interaction. If these files have a sensitivity label applied, that's also captured.

These events flow into activity explorer in DSPM for AI and the **AI activities** tab in activity explorer from the preview version of DSPM, where the data from prompts and responses can be displayed. You can also use the **Audit** solution from the Microsoft Purview portal to search and find these auditing events.

For more information, see Audit logs for Copilot and AI activities.

Data classification and AI interactions

Microsoft Purview data classification provides a comprehensive framework for identifying and tagging sensitive data across various Microsoft services, including Office 365, Dynamics 365, and Azure. Classifying data is often the first step to ensure compliance with data protection regulations and safeguard against unauthorized access, alteration, or destruction. You can use built-in system classifications or create your own.

Sensitive information types and trainable classifiers can be used to find sensitive data in user prompts and responses when they use AI apps. The resulting information then surfaces in Microsoft Purview Reports overview and activity explorer in DSPM for AI and the **AI activities** tab in activity explorer from the preview version of DSPM.

Sensitivity labels and AI interactions

AI apps that Microsoft Purview support use existing controls to ensure that data stored in your tenant is never returned to the user or used by a large language model (LLM) if the user doesn't have access to that data. When the data has sensitivity labels from your organization applied to the content, there's an extra layer of protection:

- When a file is open in Word, Excel, PowerPoint, or similarly an email or calendar event is open in Outlook, the sensitivity of the data is displayed to users in the app with the label name and content markings (such as header or footer text) that have been configured for the label. Loop components and pages also support the same sensitivity labels.

- When the sensitivity label applies encryption, users must have the EXTRACT usage right, as well as VIEW, for the AI apps to return the data.
- This protection extends to data stored outside your Microsoft 365 tenant when it's open in an Office app (data in use). For example, local storage, network shares, and cloud storage.

> **Tip**
>
> If you haven't already, we recommend you enable sensitivity labels for SharePoint and OneDrive and also familiarize yourself with the file types and label configurations that these services can process. When sensitivity labels aren't enabled for these services, the encrypted files that Copilot and agents can access are limited to data in use from Office apps on Windows.
>
> For instructions, see **Enable sensitivity labels for Office files in SharePoint and OneDrive**.

If you're not already using sensitivity labels, see Get started with sensitivity labels.

AI app-specific information:

- Microsoft 365 Copilot Chat displays the sensitivity label for items listed in the response and citations. Using the sensitivity labels' priority number that's defined in the Microsoft Purview portal, the latest response in Microsoft 365 Copilot Chat and threaded summaries in Copilot in Teams chat and channels display the highest priority sensitivity label from the data used for that Copilot chat.

  Although compliance admins define a sensitivity label's priority, a higher priority number usually denotes higher sensitivity of the content, with more restrictive permissions. As a result, Copilot responses display the most restrictive sensitivity label to educate the user about the sensitivity of the data.

- Copilot in Word, Copilot in PowerPoint, and Copilot in Outlook support sensitivity label inheritance for newly created content. See the following section for more information.

Sensitivity label inheritance

If you use Copilot in Word, Copilot in PowerPoint, and Copilot in Outlook to create new content based on an item that has a sensitivity label applied, the sensitivity label from the source file is automatically inherited, with the label's protection settings.

For example, a user selects **Draft with Copilot** in Word and then **Reference a file**. Or a user selects **Create presentation from file** in PowerPoint, or **Edit in Pages** from Microsoft 365 Copilot Chat. The source content has the sensitivity label **Confidential\Anyone (unrestricted)** applied and that label is configured to apply a footer that displays "Confidential". The new content is automatically labeled **Confidential\Anyone (unrestricted)** with the same footer.

To see an example of this in action, watch the following demo from the Ignite 2023 session, "Getting your enterprise ready for Microsoft 365 Copilot". The demo shows how the default sensitivity label of **General** is replaced with a **Confidential** label when a user drafts with Copilot and references a labeled file. The information bar under the ribbon informs the user that content created by Copilot resulted in the new label being automatically applied:

https://learn-video.azurefd.net/vod/player?id=e24ee590-7684-49aa-ab7e-b0e5652171b6&locale=en-us&embedUrl=%2Fpurview%2Fai-m365-copilot

If multiple files are used to create new content, the sensitivity label with the highest priority is used for label inheritance.

As with all automatic labeling scenarios, the user can always override and replace an inherited label (or remove, if you're not using mandatory labeling).

Encryption without sensitivity labels and AI interactions

Even if a sensitivity label isn't applied to content, services and products might use the encryption capabilities from the Azure Rights Management service. As a result, AI apps can still check for the VIEW and EXTRACT usage rights before returning data and links to a user, but there's no automatic inheritance of protection for new items.

> **Tip**
>
> You'll get the best user experience when you always use sensitivity labels to protect your data, and encryption is applied by a label.

Examples of products and services that can use the encryption capabilities from the Azure Rights Management service without sensitivity labels:

- Microsoft Purview Message Encryption
- Microsoft Information Rights Management (IRM)
- Microsoft Rights Management connector

- Microsoft Rights Management SDK

For other encryption methods that don't use the Azure Rights Management service:

- S/MIME protected emails won't be returned by Copilot, and Copilot isn't available in Outlook when an S/MIME protected email is open.
- Password-protected documents can't be accessed by AI apps unless they're already opened by the user in the same app (data in use). Passwords aren't inherited by a destination item.

As with other Microsoft 365 services, such as eDiscovery and search, items encrypted with Microsoft Purview Customer Key or your own root key (BYOK) are supported and eligible to be returned by Copilot.

Data loss prevention and AI interactions

Microsoft Purview Data Loss Prevention (DLP) helps you identify sensitive items across Microsoft 365 services and endpoints, monitor them, and helps protect against leakage of those items. It uses deep content inspection and contextual analysis to identify sensitive items and it enforces policies to protect sensitive data such as financial records, health information, or intellectual property.

Windows computers that are onboarded to Microsoft Purview can be configured for Endpoint data loss prevention (DLP) policies that warn or block users from sharing sensitive information with third-party generative AI sites that are accessed via a browser. For example, a user is prevented from pasting credit card numbers into ChatGPT, or they see a warning that they can override. For more information about the supported DLP actions and which platforms support them, see the first two rows in the table from Endpoint activities you can monitor and take action on.

AI app-specific information:

- Use the **Microsoft 365 Copilot and Copilot Chat** policy location to restrict the processing of prompts that contain sensitive information types, or processing files and emails that have specific sensitivity labels applied. For more information, see Learn about using Microsoft Purview Data Loss Prevention to protect interactions with Microsoft 365 Copilot and Copilot Chat.
- Microsoft 365 Copilot Chat supports the following Endpoint DLP capabilities only:
  - Block paste of sensitive content
  - Block files based on a specified sensitivity label

Insider Risk Management and AI interactions

Microsoft Purview Insider Risk Management helps you detect, investigate, and mitigate internal risks such as IP theft, data leakage, and security violations. It leverages machine learning models and various signals from Microsoft 365 and third-party indicators to identify potential malicious or inadvertent insider activities. The solution includes privacy controls like pseudonymization and role-based access, ensuring user-level privacy while enabling risk analysts to take appropriate actions.

Use the Risky AI usage policy template to detect risky usage that includes prompt injection attacks and accessing protected materials. Insights from these signals are integrated into Microsoft Defender XDR to provide a comprehensive view of AI-related risks.

Communication compliance and AI interactions

Microsoft Purview Communication Compliance provides tools to help you detect and manage regulatory compliance and business conduct violations across various communication channels, which include user prompts and responses for AI apps. It's designed with privacy by default, pseudonymizing usernames and incorporating role-based access controls. The solution helps identify and remediate inappropriate communications, such as sharing sensitive information, harassment, threats, and adult content.

To learn more about using communication compliance policies for AI apps, see Configure a communication compliance policy to detect for generative AI interactions.

eDiscovery and AI interactions

Microsoft Purview eDiscovery lets you identify and deliver electronic information that can be used as evidence in legal cases. The eDiscovery tools in Microsoft Purview support searching for content in Exchange Online, OneDrive for Business, SharePoint Online, Microsoft Teams, Microsoft 365 Groups, and Viva Engage teams. You can then prevent the information from deletion and export the information.

Because user prompts and responses for AI apps are stored in a user's mailbox, you can create a case and use search when a user's mailbox is selected as the source for a search query. For example, select and retrieve this data from the source mailbox by selecting from the query builder **Add condition** > **Type** > **Contains any of** > **Edit** > *Copilot activity*. This query condition includes all Copilot and other AI application activity.

After the search is refined, you can export the results or add to a review set. You can review and export information directly from the review set.

For more information about identifying and deleting user AI interaction data, see Search for and delete Copilot data in eDiscovery.

Data Lifecycle Management and AI interactions

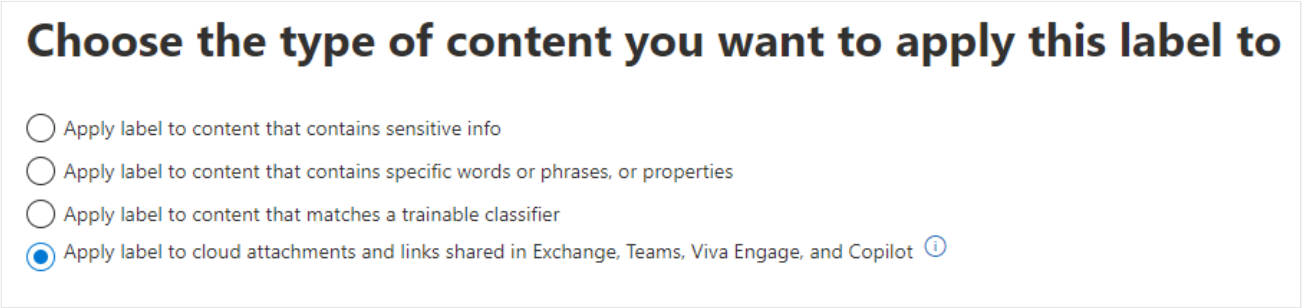
Microsoft Purview Data Lifecycle Management provides tools and capabilities to manage the lifecycle of organizational data by retaining necessary content and deleting unnecessary content. These tools ensure compliance with business, legal, and regulatory requirements.

Use retention policies to automatically retain or delete user prompts and responses for AI apps. For detailed information about this retention works, see Learn about retention for Copilot & AI apps.

As with all retention policies and holds, if more than one policy for the same location applies to a user, the principles of retention resolve any conflicts. For example, the data is retained for the longest duration of all the applied retention policies or eDiscovery holds.

AI app-specific information:

- For retention policies, select the option for **Microsoft Copilot Experiences**.
- Retention labels can automatically retain files referenced in Microsoft 365 Copilot when you select the option for cloud attachments with an auto-apply retention label policy: **Apply label to cloud attachments and links shared in Exchange, Teams, Viva Engage, and Copilot**. As with all retained cloud attachments, the file version at the time it's referenced is retained.

  For detailed information about how this retention works, see How retention works with cloud attachments.

Compliance Manager and AI interactions

Microsoft Purview Compliance Manager is a solution that helps you automatically assess and manage compliance across your multicloud environment. Compliance Manager can help you throughout your compliance journey, from taking inventory of your data protection risks to managing the complexities of implementing controls, staying current with regulations and certifications, and reporting to auditors.

To help you keep compliant with AI regulations, Compliance Manager provides regulatory templates to help you assess, implement, and strengthen your compliance requirements for all generative AI apps. For example, monitoring AI interactions and preventing data loss in AI applications. For more information, see Assessments for AI regulations.

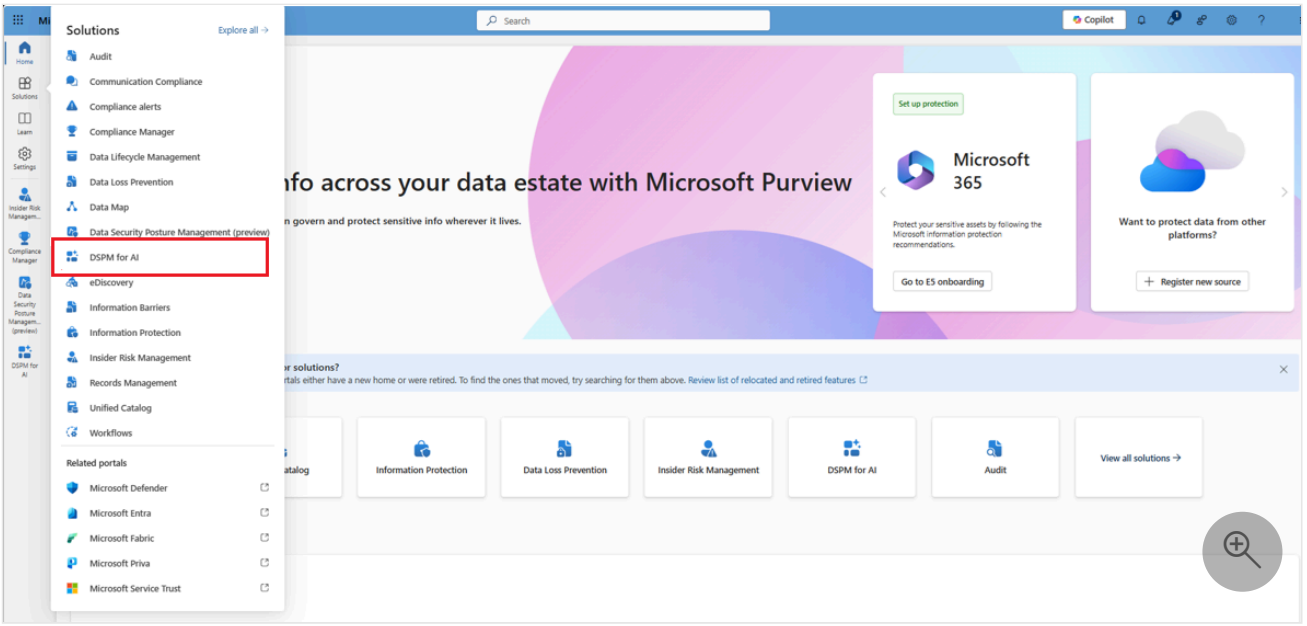
Getting started recommended steps

Use the following steps to get started with managing data security & compliance for AI interactions from Microsoft 365 Copilot & Microsoft 365 Copilot Chat.

> **Note**
>
> These steps focus on managing a specific AI app or agent. For broader coverage that uses security objectives with guided workflows, use the current version of **Data Security Posture Management**.

Because Data Security Posture Management for AI is your front door for securing and managing AI interactions, most of the following instructions use that solution:

- Sign in to the Microsoft Purview portal > **Solutions** > **DSPM for AI (classic)** with an account that has appropriate permissions. For example, an account that's a member of the **Microsoft Entra Compliance Administrator** group role.



Confirm that auditing is turned on

From **DSPM for AI (classic)** > **Overview** > **All AI apps** view > **Get Started** section, look to see if auditing is on for your tenant. If not, select **Activate Microsoft Purview Audit**.

Use the Microsoft 365 Copilot view to discover, protect, and apply compliance controls

Change the view from **All AI apps** to **Microsoft 365 Copilot** and work your way through the sections:

- **Assess and prevent oversharing of sensitive data**
- **Secure your data in Microsoft 365 Copilot**
- **Discover Microsoft 365 Copilot activity**

These sections include recommendations that are specific to Microsoft 365 Copilot. Wait at least a day for data to display for the reports on this page.

> **Tip**
>
> **Secure your data in Microsoft 365 Copilot** includes the following recommendations with manual configuration instructions for their respective Microsoft Purview solutions: **Create sensitivity labels for your organization** and **Protect items with sensitivity labels from Copilot processing**

If you prefer, you can use DSPM for AI one-click policies for these recommendations, as described in the next section.

Use one-click policies to increase coverage

Select **Recommendations** from the navigation and use one-click policies to automatically create policies that help you discover, protect, and apply compliance controls. Specific to Microsoft 365 Copilot:

- **Protect your data with sensitivity labels**
- **Detect risky interactions in AI apps**
- **Detect unethical behavior in AI**
- **Protect items with sensitivity labels from Microsoft 365 Copilot and agent processing**

View data from your policies

1. Wait at least a day for data, and then navigate to the **Reports** page to view the results of your policies. Select **Copilot experiences & agents** and view information such as:

    - Total interactions over time (Microsoft Copilot and agents)
    - Sensitive interactions per AI app
    - Top unethical AI interactions
    - Top sensitivity labels references in Microsoft 365 Copilot and agents
    - Insider Risk severity
    - Insider risk severity per AI app
    - Potential risky AI usage

2. Select **View details** for each of the report graphs to view detailed activities in the activity explorer.

    From the filters, select the **AI app category** of **Copilot experiences & agents**, and then use the other filters if you need to further refine the displayed data. Then drill down to each activity to view details that include displaying the prompts and response when you're a member of the Microsoft Purview Content Explorer Content Viewer role group. For more information about this requirement, see Permissions for Data Security Posture Management for AI.

Apply additional compliance controls to Microsoft 365 Copilot interactions

1. If you need to retain the exact version of files referenced in Microsoft 365 Copilot interactions:

   In the Microsoft Purview portal, navigate to **Data Lifecycle Management** > **Retention labels** > and either locate or create a retention label with the required retention period. Then, navigate to **Label policies** to auto-apply that label, and select the option **Apply label to cloud attachments and links shared in Exchange, Teams, Viva Engage, and Copilot**. For more information, see Automatically apply a retention label to retain or delete content.

2. If you need to preserve, collect, analyze, review, or export Microsoft 365 Copilot interactions:

   In the Microsoft Purview portal, navigate to **eDiscovery** > **Cases** > **Create case**. In the case, create a search and use the *ItemClass* property and the `IPM.SkypeTeams.Message.Copilot.*` value to search for these interactions in your organization.

Routinely review the reports and data risk assessments in DSPM for AI to determine if you need to make changes, and use activity explorer and events for deeper analysis of how users are interacting with Microsoft 365 Copilot & Microsoft 365 Copilot Chat.

Other documentation to help you secure and manage generative AI apps

For more detailed information, see Considerations to manage Microsoft 365 Copilot for security and compliance.

Microsoft 365 Copilot documentation:

- Microsoft 365 Copilot documentation
- Microsoft 365 Copilot admin guide for E5 + SAM licenses
- Apply principles of Zero Trust to Microsoft 365 Copilot

Last updated on 05/01/2026

Use Microsoft Purview to manage data security & compliance for Microsoft 365 Copilot Cowork

> *Microsoft Purview service description*

Use the following sections to identify the Microsoft Purview capabilities that are supported for Microsoft 365 Copilot Cowork.

> **Note**
>
> Although Copilot Cowork is part of Microsoft 365 Copilot, Cowork introduces long-running, multi-step work across apps. Because this model can result in different behaviors and feature support, its capabilities are documented separately to provide accurate guidance for administrators to better understand security and compliance scenarios.

Capabilities supported

Use the following table to see at a glance the Microsoft Purview capabilities that are supported with Copilot Cowork.

Expand table

Capability or solution in Microsoft Purview	Supported for AI interactions
DSPM and DSPM for AI (classic)	✓
Auditing	✓
Data classification	✗
Sensitivity labels	✓
Encryption without sensitivity labels	✓
Data loss prevention	✓
Insider Risk Management	✓

Capability or solution in Microsoft Purview	Supported for AI interactions
Communication compliance	✓
eDiscovery	✓
Data Lifecycle Management	✓
Compliance Manager	✗

DSPM and DSPM for AI (classic)

Use Data Security Posture Management or Data Security Posture Management for AI (classic) as your front door to discover, secure, and apply compliance controls for AI usage across your enterprise. Both DSPM versions use existing controls from Microsoft Purview information protection and compliance management with easy-to-use graphical tools and reports to quickly gain insights into AI use within your organization. With personalized recommendations, and one-click policies help you protect your data and comply with regulatory requirements.

AI app-specific information:

- Data risk assessments help you identify and fix issues that could results in oversharing of data. From the recommendations:
  - **Protect your data from potential oversharing risks** for the default weekly data risk assessment.
  - **Protect sensitive data references in Copilot and agent responses** for a custom data risk assessment.

- One-click policies available:
  - **Sensitivity labels and policies** from the recommendation or remediation action **Protect your data with sensitivity labels**.
  - **DSPM for AI - Detect risky AI usage** from the recommendation or remediation action **Detect risky interactions in AI apps**.
  - **DSPM for AI - Unethical behavior in AI apps** from the recommendation or remediation action **Detect unethical behavior in AI**.

- Although interactions from Cowork display in activity explorer (**AI activities** tab), they don't display in the **Apps and agents** dashboard or the **AI observability** page.

Auditing and AI interactions

Microsoft Purview Audit solutions provide comprehensive tools for searching and managing audit records of activities performed across various Microsoft services by users and admins, and help organizations to effectively respond to security events, forensic investigations, internal investigations, and compliance obligations.

Like other activities, prompts and responses are captured in the unified audit log. Events include how and when users interact with the AI app, and can include in which Microsoft 365 service the activity took place, and references to the files stored in Microsoft 365 that were accessed during the interaction. If these files have a sensitivity label applied, that's also captured.

These events flow into activity explorer in the **AI activities** tab in the current version of DSPM and in DSPM for AI, where the data from prompts and responses can be displayed. You can also use the **Audit** solution from the Microsoft Purview portal to search and find these auditing events.

For more information, see Audit logs for Copilot and AI activities.

AI app-specific information:

Typical auditing events:

- A user starts a new Cowork conversation
- A user adds, removes, or shares a skill or plugin
- A user runs a scheduled prompt
- A user starts a browser task
- A user uploads a file to Cowork
- Cowork creates an artifact in OneDrive on behalf of a user

Sensitivity labels and AI interactions

AI apps that Microsoft Purview support use existing controls to ensure that data stored in your tenant is never returned to the user or used by a large language model (LLM) if the user doesn't have access to that data. When the data has sensitivity labels from your organization applied to the content, there's an extra layer of protection:

- When a file is open in Word, Excel, PowerPoint, or similarly an email or calendar event is open in Outlook, the sensitivity of the data is displayed to users in the app with the label name and content markings (such as header or footer text) that have been configured for the label. Loop components and pages also support the same sensitivity labels.

- When the sensitivity label applies encryption, users must have the EXTRACT usage right, as well as VIEW, for the AI apps to return the data.

- This protection extends to data stored outside your Microsoft 365 tenant when it's open in an Office app (data in use). For example, local storage, network shares, and cloud storage.

> **Tip**
>
> If you haven't already, we recommend you enable sensitivity labels for SharePoint and OneDrive and also familiarize yourself with the file types and label configurations that these services can process. When sensitivity labels aren't enabled for these services, the encrypted files that Copilot and agents can access are limited to data in use from Office apps on Windows.
>
> For instructions, see **Enable sensitivity labels for Office files in SharePoint and OneDrive**.

If you're not already using sensitivity labels, see Get started with sensitivity labels.

AI app-specific information:

- Microsoft 365 Copilot Cowork displays the sensitivity label for items listed in the response and citations. Using the sensitivity labels' priority number that's defined in the Microsoft Purview portal, the latest response displays the highest priority sensitivity label from the data used for that Copilot chat.

  Although compliance admins define a sensitivity label's priority, a higher priority number usually denotes higher sensitivity of the content, with more restrictive permissions. As a result, Copilot responses display the most restrictive sensitivity label to educate the user about the sensitivity of the data.

- Sensitivity label inheritance is supported for newly created content. See the following section for more information.

Sensitivity label inheritance

If you use Cowork to create new content in Word, PowerPoint, and Outlook based on an item that has a sensitivity label applied, the sensitivity label from the source file is automatically inherited, with the label's protection setting.

For example, Cowork creates a new Word document based on labeled data. The source content has the sensitivity label **Confidential\Anyone (unrestricted)** applied and that label is configured to apply a footer that displays "Confidential". The new content is automatically labeled **Confidential\Anyone (unrestricted)** with the same footer.

If multiple sources are used to create new content, the sensitivity label with the highest priority is used for label inheritance.

As with all automatic labeling scenarios, a user can always override and replace an inherited label (or remove, if you're not using mandatory labeling).

Encryption without sensitivity labels and AI interactions

Even if a sensitivity label isn't applied to content, services and products might use the encryption capabilities from the Azure Rights Management service. As a result, AI apps can still check for the VIEW and EXTRACT usage rights before returning data and links to a user, but there's no automatic inheritance of protection for new items.

> **Tip**
>
> You'll get the best user experience when you always use sensitivity labels to protect your data, and encryption is applied by a label.

Examples of products and services that can use the encryption capabilities from the Azure Rights Management service without sensitivity labels:

- Microsoft Purview Message Encryption
- Microsoft Information Rights Management (IRM)
- Microsoft Rights Management connector
- Microsoft Rights Management SDK

For other encryption methods that don't use the Azure Rights Management service:

- S/MIME protected emails won't be returned by Copilot, and Copilot isn't available in Outlook when an S/MIME protected email is open.
- Password-protected documents can't be accessed by AI apps unless they're already opened by the user in the same app (data in use). Passwords aren't inherited by a destination item.

As with other Microsoft 365 services, such as eDiscovery and search, items encrypted with Microsoft Purview Customer Key or your own root key (BYOK) are supported and eligible to be returned by Copilot.

Insider Risk Management and AI interactions

Microsoft Purview Insider Risk Management helps you detect, investigate, and mitigate internal risks such as IP theft, data leakage, and security violations. It leverages machine learning models and various signals from Microsoft 365 and third-party indicators to identify potential malicious or inadvertent insider activities. The solution includes privacy controls like pseudonymization and role-based access, ensuring user-level privacy while enabling risk analysts to take appropriate actions.

Use the Risky AI usage policy template to detect risky usage that includes prompt injection attacks and accessing protected materials. Insights from these signals are integrated into Microsoft Defender XDR to provide a comprehensive view of AI-related risks.

Communication compliance and AI interactions

Microsoft Purview Communication Compliance provides tools to help you detect and manage regulatory compliance and business conduct violations across various communication channels, which include user prompts and responses for AI apps. It's designed with privacy by default, pseudonymizing usernames and incorporating role-based access controls. The solution helps identify and remediate inappropriate communications, such as sharing sensitive information, harassment, threats, and adult content.

To learn more about using communication compliance policies for AI apps, see Configure a communication compliance policy to detect for generative AI interactions.

eDiscovery and AI interactions

Microsoft Purview eDiscovery lets you identify and deliver electronic information that can be used as evidence in legal cases. The eDiscovery tools in Microsoft Purview support searching for content in Exchange Online, OneDrive for Business, SharePoint Online, Microsoft Teams, Microsoft 365 Groups, and Viva Engage teams. You can then prevent the information from deletion and export the information.

Because user prompts and responses for AI apps are stored in a user's mailbox, you can create a case and use search when a user's mailbox is selected as the source for a search query. For example, select and retrieve this data from the source mailbox by selecting from the query builder **Add condition** > **Type** > **Contains any of** > **Edit** > *Copilot activity*. This query condition includes all Copilot and other AI application activity.

After the search is refined, you can export the results or add to a review set. You can review and export information directly from the review set.

For more information about identifying and deleting user AI interaction data, see Search for and delete Copilot data in eDiscovery.

AI app-specific information:

In scope for eDiscovery:

- Conversation transcripts — The full text of a user's Cowork conversations, including prompts and responses.
- Files Cowork created — Documents, decks, PDFs, and images saved to the user's OneDrive.
- Scheduled prompts — The prompt text and the conversations created by each scheduled run.
- Uploaded files — Files a user uploaded to Cowork that are retained in the user's OneDrive or in the conversation.

Data Lifecycle Management and AI interactions

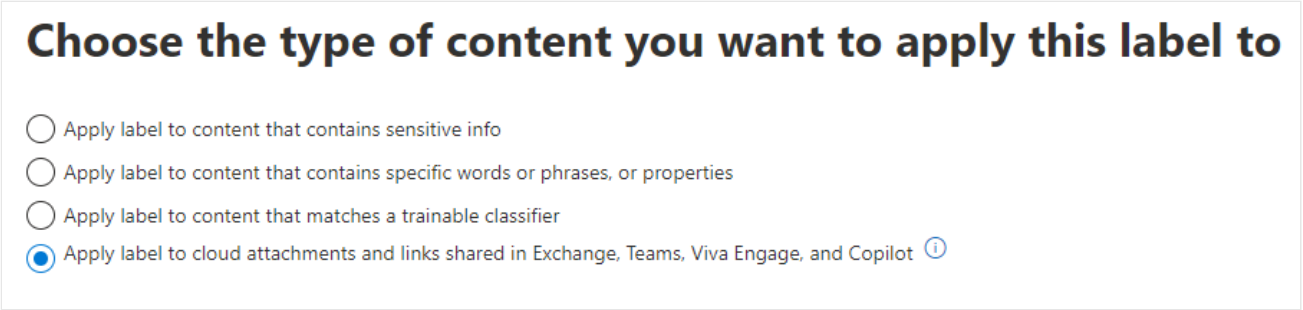
Microsoft Purview Data Lifecycle Management provides tools and capabilities to manage the lifecycle of organizational data by retaining necessary content and deleting unnecessary content. These tools ensure compliance with business, legal, and regulatory requirements.

Use retention policies to automatically retain or delete user prompts and responses for AI apps. For detailed information about this retention works, see Learn about retention for Copilot & AI apps.

As with all retention policies and holds, if more than one policy for the same location applies to a user, the principles of retention resolve any conflicts. For example, the data is retained for the longest duration of all the applied retention policies or eDiscovery holds.

AI app-specific information:

- For retention policies, select the option for **Microsoft Copilot Experiences**.
- Retention labels can automatically retain files referenced in Microsoft 365 Copilot Cowork when you select the option for cloud attachments with an auto-apply retention label policy: **Apply label to cloud attachments and links shared in Exchange, Teams, Viva Engage, and Copilot**. As with all retained cloud attachments, the file version at the time it's referenced is retained.



For detailed information about how this retention works, see How retention works with cloud attachments.

Other documentation to help you secure and manage generative AI apps

- Microsoft Purview data security and compliance protections for generative AI apps
- Microsoft 365 Copilot Cowork Overview

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 09/30/2026

Use Microsoft Purview to manage data security & compliance for Microsoft Security Copilot

> *Microsoft Purview service description*

Use the following sections to identify the Microsoft Purview capabilities that are supported for AI interactions with Microsoft Security Copilot, and some get started recommendations for you to manage these AI interactions for security and compliance.

Before Microsoft Purview can manage these AI interactions, you must first turn on the Security Copilot option **Allow Microsoft Purview to access, process, copy, and store Customer Data from your Security Copilot Service**. For more information, see Enable Security Copilot for Purview Audit and DSPM Solutions.

Managing these AI interactions with Microsoft Purview requires you to enable pay-as-you-go billing in your organization.

Capabilities supported

Use the following table to see at a glance the Microsoft Purview capabilities that are supported with Microsoft Security Copilot.

Expand table

Capability or solution in Microsoft Purview	Supported for AI interactions
DSPM and DSPM for AI (classic)	✓
Auditing	✓
Data classification	✓
Sensitivity labels	✗
Encryption without sensitivity labels	✗
Data loss prevention	✗
Insider Risk Management	✓

Capability or solution in Microsoft Purview	Supported for AI interactions
Communication compliance	✓
eDiscovery	✓
Data Lifecycle Management	✓
Compliance Manager	✓

DSPM and DSPM for AI (classic)

Use Data Security Posture Management or Data Security Posture Management for AI (classic) as your front door to discover, secure, and apply compliance controls for AI usage across your enterprise. Both DSPM versions use existing controls from Microsoft Purview information protection and compliance management with easy-to-use graphical tools and reports to quickly gain insights into AI use within your organization. With personalized recommendations, and one-click policies help you protect your data and comply with regulatory requirements.

AI app-specific information:

- Recommendation **Get guided assistance to AI regulations**, which uses control-mapping regulatory templates from Compliance Manager.
- One-click policies available:
  - **DSPM for AI - Capture interactions for Copilot experiences** from the recommendation or remediation action **Secure interactions in Microsoft Copilot experiences**.
  - **DSPM for AI - Detect risky AI usage** from the recommendation or remediation action **Detect risky interactions in AI apps**.
  - **DSPM for AI - Unethical behavior in AI apps** from the recommendation or remediation action **Detect unethical behavior in AI apps**
  - **DSPM for AI - Detect sensitive info shared with AI via network** from the recommendation or remediation action **Extend insights into sensitive data in AI app interactions**.

Auditing and AI interactions

Microsoft Purview Audit solutions provide comprehensive tools for searching and managing audit records of activities performed across various Microsoft services by users and admins, and help organizations to effectively respond to security events, forensic investigations, internal investigations, and compliance obligations.

Like other activities, prompts and responses are captured in the unified audit log. Events include how and when users interact with the AI app, and can include in which Microsoft 365 service the activity took place, and references to the files stored in Microsoft 365 that were accessed during the interaction. If these files have a sensitivity label applied, that's also captured.

These events flow into activity explorer in DSPM for AI and the **AI activities** tab in activity explorer from the preview version of DSPM, where the data from prompts and responses can be displayed. You can also use the **Audit** solution from the Microsoft Purview portal to search and find these auditing events.

For more information, see Audit logs for Copilot and AI activities.

Data classification and AI interactions

Microsoft Purview data classification provides a comprehensive framework for identifying and tagging sensitive data across various Microsoft services, including Office 365, Dynamics 365, and Azure. Classifying data is often the first step to ensure compliance with data protection regulations and safeguard against unauthorized access, alteration, or destruction. You can use built-in system classifications or create your own.

Sensitive information types and trainable classifiers can be used to find sensitive data in user prompts and responses when they use AI apps. The resulting information then surfaces in Microsoft Purview Reports overview and activity explorer in DSPM for AI and the **AI activities** tab in activity explorer from the preview version of DSPM.

Insider Risk Management and AI interactions

Microsoft Purview Insider Risk Management helps you detect, investigate, and mitigate internal risks such as IP theft, data leakage, and security violations. It leverages machine learning models and various signals from Microsoft 365 and third-party indicators to identify potential malicious or inadvertent insider activities. The solution includes privacy controls like pseudonymization and role-based access, ensuring user-level privacy while enabling risk analysts to take appropriate actions.

Use the Risky AI usage policy template to detect risky usage that includes prompt injection attacks and accessing protected materials. Insights from these signals are integrated into Microsoft Defender XDR to provide a comprehensive view of AI-related risks.

AI app-specific information:

- For prompts and responses, requires a collection policy such as **DSPM for AI - Capture interactions for Copilot experiences** from the DSPM for AI recommendation **Secure interactions in Microsoft Copilot experiences**.

Communication compliance and AI interactions

Microsoft Purview Communication Compliance provides tools to help you detect and manage regulatory compliance and business conduct violations across various communication channels, which include user prompts and responses for AI apps. It's designed with privacy by default, pseudonymizing usernames and incorporating role-based access controls. The solution helps identify and remediate inappropriate communications, such as sharing sensitive information, harassment, threats, and adult content.

To learn more about using communication compliance policies for AI apps, see Configure a communication compliance policy to detect for generative AI interactions.

AI app-specific information:

- For prompts and responses, requires a collection policy such as **DSPM for AI - Capture interactions for Copilot experiences** from the DSPM for AI recommendation **Secure interactions in Microsoft Copilot experiences**.

eDiscovery and AI interactions

Microsoft Purview eDiscovery lets you identify and deliver electronic information that can be used as evidence in legal cases. The eDiscovery tools in Microsoft Purview support searching for content in Exchange Online, OneDrive for Business, SharePoint Online, Microsoft Teams, Microsoft 365 Groups, and Viva Engage teams. You can then prevent the information from deletion and export the information.

Because user prompts and responses for AI apps are stored in a user's mailbox, you can create a case and use search when a user's mailbox is selected as the source for a search query. For example, select and retrieve this data from the source mailbox by selecting from the query builder **Add condition** > **Type** > **Contains any of** > **Edit** > *Copilot activity*. This query condition includes all Copilot and other AI application activity.

After the search is refined, you can export the results or add to a review set. You can review and export information directly from the review set.

For more information about identifying and deleting user AI interaction data, see Search for and delete Copilot data in eDiscovery.

AI app-specific information:

- For prompts and responses, requires a collection policy such as **DSPM for AI - Capture interactions for Copilot experiences** from the DSPM for AI recommendation **Secure interactions in Microsoft Copilot experiences**.

Data Lifecycle Management and AI interactions

Microsoft Purview Data Lifecycle Management provides tools and capabilities to manage the lifecycle of organizational data by retaining necessary content and deleting unnecessary content. These tools ensure compliance with business, legal, and regulatory requirements.

Use retention policies to automatically retain or delete user prompts and responses for AI apps. For detailed information about this retention works, see Learn about retention for Copilot & AI apps.

As with all retention policies and holds, if more than one policy for the same location applies to a user, the principles of retention resolve any conflicts. For example, the data is retained for the longest duration of all the applied retention policies or eDiscovery holds.

AI app-specific information:

- For retention policies, select the option for **Microsoft Copilot Experiences**.
- For prompts and responses, requires a collection policy such as **DSPM for AI - Capture interactions for Copilot experiences** from the DSPM for AI recommendation **Secure interactions in Microsoft Copilot experiences**.

Compliance Manager and AI interactions

Microsoft Purview Compliance Manager is a solution that helps you automatically assess and manage compliance across your multicloud environment. Compliance Manager can help you throughout your compliance journey, from taking inventory of your data protection risks to managing the complexities of implementing controls, staying current with regulations and certifications, and reporting to auditors.

To help you keep compliant with AI regulations, Compliance Manager provides regulatory templates to help you assess, implement, and strengthen your compliance requirements for all

generative AI apps. For example, monitoring AI interactions and preventing data loss in AI applications. For more information, see Assessments for AI regulations.

Getting started recommended steps

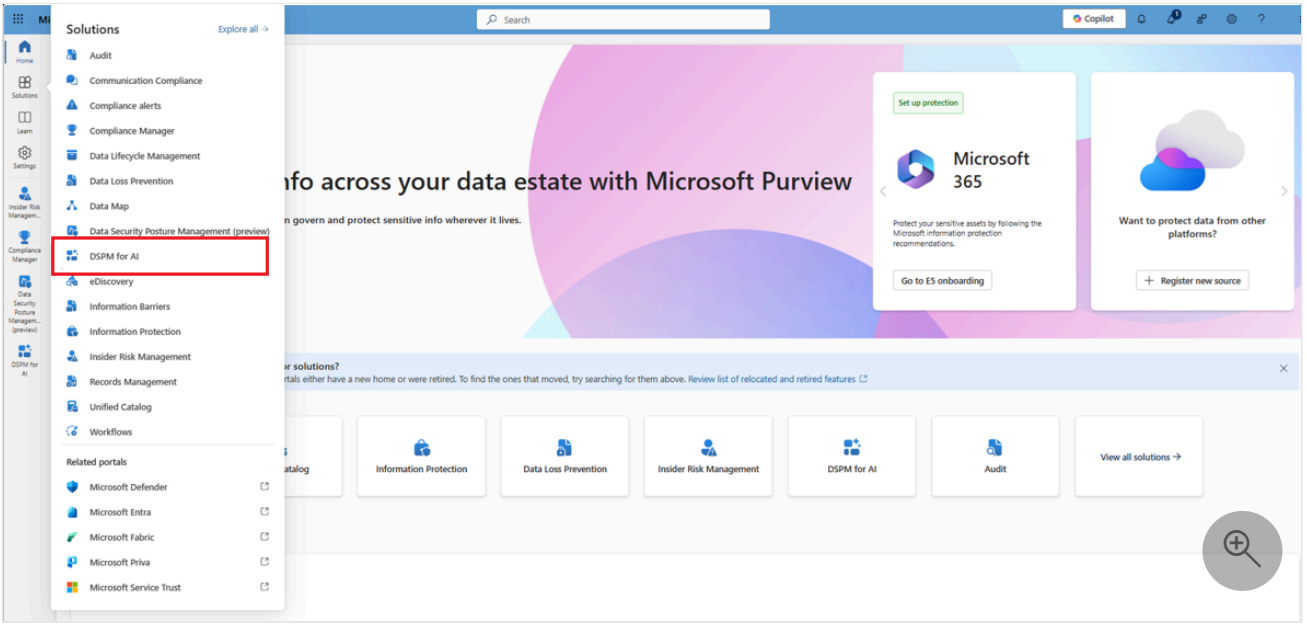
Use the following steps to get started with managing data security & compliance for AI interactions from Microsoft Security Copilot.

> **Note**
>
> These steps focus on managing a specific AI app or agent. For broader coverage that uses security objectives with guided workflows, use the current version of Data Security Posture Management.

Managing data security & compliance for Copilot in Fabric requires the enterprise version of Microsoft Purview data governance, to support the required APIs.

Because Data Security Posture Management for AI is your front door for securing and managing AI interactions, most of the following instructions use that solution:

- Sign in to the Microsoft Purview portal > **Solutions** > **DSPM for AI (classic)** with an account that has appropriate permissions. For example, an account that's a member of the **Microsoft Entra Compliance Administrator** group role.



Discover potential security risks in Microsoft Security Copilot interactions

1. From **DSPM for AI (classic)** > **Overview**, in the **Get Started** section, look to see if auditing is on for your tenant. If not, select **Activate Microsoft Purview Audit**.

2. In the **Recommendations** section, select **View all recommendations** to take you to the **Recommendations** page. If they have a status of **Not Started**, take action on the following recommendations that create one-click policies:

   - **Secure interactions in Microsoft Copilot experiences**: This recommendation creates a policy to policy captures prompts and responses for data security posture and regulatory compliance from Copilot in Fabric, and Security Copilot.

   - **Detect risky interactions in AI apps**: This recommendation creates a policy to help calculate user risk by detecting risky prompts and responses in Microsoft 365 Copilot, agents, and other generative AI apps.

3. Wait at least a day for data, and then navigate to the **Reports** page to view the results of your policy. Select **Copilot experiences & agents** and view information such as:

   - Total interactions over time (Microsoft Copilot and agents)
   - Sensitive interactions per AI app
   - Top unethical AI interactions
   - Insider Risk severity
   - Insider risk severity per AI app
   - Potential risky AI usage

4. Select **View details** for each of the report graphs to view detailed activities in the activity explorer.

   From the filters, select the **AI app category** of **Copilot experiences & agents**, and then use the other filters if you need to further refine the displayed data. Then drill down to each activity to view details that include displaying the prompts and response when you're a member of the Microsoft Purview Content Explorer Content Viewer role group. For more information about this requirement, see Permissions for Data Security Posture Management for AI.

Apply compliance controls to Microsoft Security Copilot

1. From **DSPM for AI (classic)** > **Recommendations** page, under **Not Started**, locate, and select and take actions on the following recommendation:

- **Detect unethical behavior in AI**: This recommendation creates a policy to detect sensitive information in prompts and responses in Microsoft Security Copilot and agents, and other generative AI apps.

2. If you need to ensure that Microsoft Security Copilot interactions are retained for compliance reasons:

   In the Microsoft Purview portal, navigate to **Data Lifecycle Management** > **Policies** > **Retention Policies** and create a retention policy to retain Microsoft Security Copilot interactions by selecting the location **Microsoft Copilot Experiences** and specify the required retention period. For more information, see Create and configure retention policies.

3. If you need to preserve, collect, analyze, review, or export Microsoft Security Copilot interactions:

   In the Microsoft Purview portal, navigate to **eDiscovery** > **Cases** > **Create case**. In the case, create a search and use the *ItemClass* property and the `IPM.SkypeTeams.Message.Copilot.Security.SecurityCopilot` value to search for these interactions in your organization.

Routinely review the reports and data risk assessments in DSPM for AI to determine if you need to make changes, and use activity explorer and events for deeper analysis of how users are interacting with Microsoft Security Copilot.

Related resources

- Blog post (July 2025): Microsoft Purview Powering Data Security and Compliance for Security Copilot

---

Last updated on 05/01/2026

Use Microsoft Purview to manage data security & compliance for Copilot in Fabric

> *Microsoft Purview service description*

Use the following sections to identify the Microsoft Purview capabilities that are supported for AI interactions with Copilot in Fabric, and some get started recommendations for you to manage these AI interactions for security and compliance.

The first Copilot in Fabric that's supported by Microsoft Purview is Copilot for Power BI. Current support extends to all prompts and all text-based responses.

> **Note**
>
> Microsoft Purview support for other Copilot experiences in Fabric and Fabric Data Agent is currently in preview.

As a prerequisite for Microsoft Purview to manage these AI interactions, the **Allow Microsoft Purview to secure AI interactions** option in the Microsoft Fabric admin portal must be **Enabled**, which is the default setting. For more information about this option, see Tenant settings index.

Managing these AI interactions with Microsoft Purview requires you to enable pay-as-you-go billing in your organization.

Capabilities supported

Use the following table to see at a glance the Microsoft Purview capabilities that are supported with Copilot in Fabric.

Expand table

Capability or solution in Microsoft Purview	Supported for AI interactions
DSPM and DSPM for AI (classic)	✓
Auditing	✓
Data classification	✓

Capability or solution in Microsoft Purview	Supported for AI interactions
Sensitivity labels	✗
Encryption without sensitivity labels	✗
Data loss prevention	✗
Insider Risk Management	✓
Communication compliance	✓
eDiscovery	✓
Data Lifecycle Management	✓
Compliance Manager	✓

DSPM and DSPM for AI (classic)

Use Data Security Posture Management or Data Security Posture Management for AI (classic) as your front door to discover, secure, and apply compliance controls for AI usage across your enterprise. Both DSPM versions use existing controls from Microsoft Purview information protection and compliance management with easy-to-use graphical tools and reports to quickly gain insights into AI use within your organization. With personalized recommendations, and one-click policies help you protect your data and comply with regulatory requirements.

AI app-specific information:

- Recommendation **Get guided assistance to AI regulations**, which uses control-mapping regulatory templates from Compliance Manager.
- One-click policies available:
  - **DSPM for AI - Capture interactions for Copilot experiences** from the recommendation or remediation action **Secure interactions in Microsoft Copilot experiences**.
  - **DSPM for AI - Detect risky AI usage** from the recommendation or remediation action **Detect risky interactions in AI apps**.
  - **DSPM for AI - Unethical behavior in AI apps** from the recommendation or remediation action **Detect unethical behavior in AI apps**.
  - **DSPM for AI - Detect sensitive info shared with AI via network** from the recommendation or remediation action **Extend insights into sensitive data in AI app interactions**.

Auditing and AI interactions

Microsoft Purview Audit solutions provide comprehensive tools for searching and managing audit records of activities performed across various Microsoft services by users and admins, and help organizations to effectively respond to security events, forensic investigations, internal investigations, and compliance obligations.

Like other activities, prompts and responses are captured in the unified audit log. Events include how and when users interact with the AI app, and can include in which Microsoft 365 service the activity took place, and references to the files stored in Microsoft 365 that were accessed during the interaction. If these files have a sensitivity label applied, that's also captured.

These events flow into activity explorer in DSPM for AI and the **AI activities** tab in activity explorer from the preview version of DSPM, where the data from prompts and responses can be displayed. You can also use the **Audit** solution from the Microsoft Purview portal to search and find these auditing events.

For more information, see Audit logs for Copilot and AI activities.

Data classification and AI interactions

Microsoft Purview data classification provides a comprehensive framework for identifying and tagging sensitive data across various Microsoft services, including Office 365, Dynamics 365, and Azure. Classifying data is often the first step to ensure compliance with data protection regulations and safeguard against unauthorized access, alteration, or destruction. You can use built-in system classifications or create your own.

Sensitive information types and trainable classifiers can be used to find sensitive data in user prompts and responses when they use AI apps. The resulting information then surfaces in Microsoft Purview Reports overview and activity explorer in DSPM for AI and the **AI activities** tab in activity explorer from the preview version of DSPM.

Insider Risk Management and AI interactions

Microsoft Purview Insider Risk Management helps you detect, investigate, and mitigate internal risks such as IP theft, data leakage, and security violations. It leverages machine learning models and various signals from Microsoft 365 and third-party indicators to identify potential malicious or inadvertent insider activities. The solution includes privacy controls like pseudonymization and role-based access, ensuring user-level privacy while enabling risk analysts to take appropriate actions.

Use the Risky AI usage policy template to detect risky usage that includes prompt injection attacks and accessing protected materials. Insights from these signals are integrated into Microsoft Defender XDR to provide a comprehensive view of AI-related risks.

AI app-specific information:

- For prompts and responses, requires a collection policy such as **DSPM for AI - Capture interactions for Copilot experiences** from the DSPM for AI recommendation **Secure interactions in Microsoft Copilot experiences**.

Communication compliance and AI interactions

Microsoft Purview Communication Compliance provides tools to help you detect and manage regulatory compliance and business conduct violations across various communication channels, which include user prompts and responses for AI apps. It's designed with privacy by default, pseudonymizing usernames and incorporating role-based access controls. The solution helps identify and remediate inappropriate communications, such as sharing sensitive information, harassment, threats, and adult content.

To learn more about using communication compliance policies for AI apps, see Configure a communication compliance policy to detect for generative AI interactions.

AI app-specific information:

- For prompts and responses, requires a collection policy such as **DSPM for AI - Capture interactions for Copilot experiences** from the DSPM for AI recommendation **Secure interactions in Microsoft Copilot experiences**.

eDiscovery and AI interactions

Microsoft Purview eDiscovery lets you identify and deliver electronic information that can be used as evidence in legal cases. The eDiscovery tools in Microsoft Purview support searching for content in Exchange Online, OneDrive for Business, SharePoint Online, Microsoft Teams, Microsoft 365 Groups, and Viva Engage teams. You can then prevent the information from deletion and export the information.

Because user prompts and responses for AI apps are stored in a user's mailbox, you can create a case and use search when a user's mailbox is selected as the source for a search query. For example, select and retrieve this data from the source mailbox by selecting from the query

builder **Add condition** > **Type** > **Contains any of** > **Edit** > *Copilot activity*. This query condition includes all Copilot and other AI application activity.

After the search is refined, you can export the results or add to a review set. You can review and export information directly from the review set.

For more information about identifying and deleting user AI interaction data, see Search for and delete Copilot data in eDiscovery.

AI app-specific information:

- For prompts and responses, requires a collection policy such as **DSPM for AI - Capture interactions for Copilot experiences** from the DSPM for AI recommendation **Secure interactions in Microsoft Copilot experiences**.

Data Lifecycle Management and AI interactions

Microsoft Purview Data Lifecycle Management provides tools and capabilities to manage the lifecycle of organizational data by retaining necessary content and deleting unnecessary content. These tools ensure compliance with business, legal, and regulatory requirements.

Use retention policies to automatically retain or delete user prompts and responses for AI apps. For detailed information about this retention works, see Learn about retention for Copilot & AI apps.

As with all retention policies and holds, if more than one policy for the same location applies to a user, the principles of retention resolve any conflicts. For example, the data is retained for the longest duration of all the applied retention policies or eDiscovery holds.

AI app-specific information:

- For retention policies, select the option for **Microsoft Copilot Experiences**.
- For prompts and responses, requires a collection policy such as **DSPM for AI - Capture interactions for Copilot experiences** from the DSPM for AI recommendation **Secure interactions in Microsoft Copilot experiences**.

Compliance Manager and AI interactions

Microsoft Purview Compliance Manager is a solution that helps you automatically assess and manage compliance across your multicloud environment. Compliance Manager can help you throughout your compliance journey, from taking inventory of your data protection risks to

managing the complexities of implementing controls, staying current with regulations and certifications, and reporting to auditors.

To help you keep compliant with AI regulations, Compliance Manager provides regulatory templates to help you assess, implement, and strengthen your compliance requirements for all generative AI apps. For example, monitoring AI interactions and preventing data loss in AI applications. For more information, see Assessments for AI regulations.

Getting started recommended steps

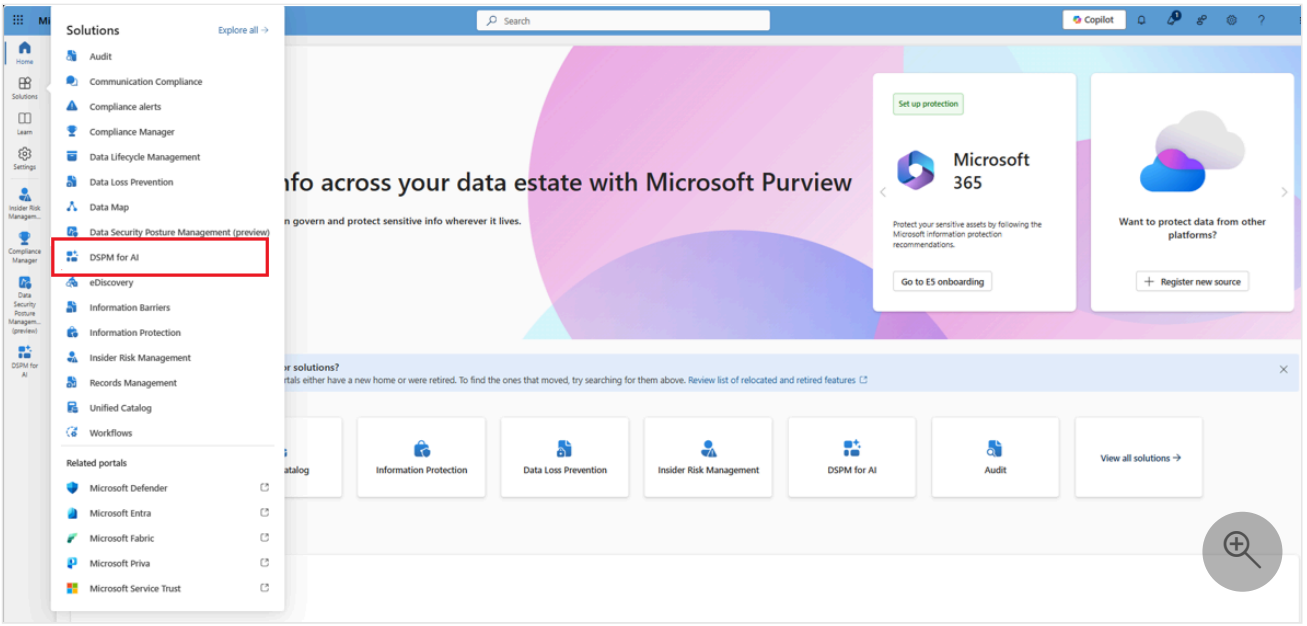
Use the following steps to get started with managing data security & compliance for AI interactions from Copilot in Fabric.

> **Note**
>
> These steps focus on managing a specific AI app or agent. For broader coverage that uses security objectives with guided workflows, use the current version of Data Security Posture Management.

Managing data security & compliance for Copilot in Fabric requires the enterprise version of Microsoft Purview data governance, to support the required APIs.

Because Data Security Posture Management for AI is your front door for securing and managing AI interactions, most of the following instructions use that solution:

- Sign in to the Microsoft Purview portal > **Solutions** > **DSPM for AI (classic)** with an account that has appropriate permissions. For example, an account that's a member of the **Microsoft Entra Compliance Administrator** group role.



Discover potential security risks in Copilot in Fabric interactions

1. From **DSPM for AI (classic)** > **Overview**, in the **Get Started** section, look to see if auditing is on for your tenant. If not, select **Activate Microsoft Purview Audit**.

2. In the **Recommendations** section, select **View all recommendations** to take you to the **Recommendations** page. If they have a status of **Not Started**, take action on the following recommendations that create one-click policies:

    - **Secure interactions in Microsoft Copilot experiences**: This recommendation creates a policy to policy captures prompts and responses for data security posture and regulatory compliance from Copilot in Fabric, and Security Copilot.

    - **Detect risky interactions in AI apps**: This recommendation creates a policy to help calculate user risk by detecting risky prompts and responses in Microsoft 365 Copilot, agents, and other generative AI apps.

3. Wait at least a day for data, and then navigate to the **Reports** page to view the results of your policy. Select **Copilot experiences & agents** and view information such as:

    - Total interactions over time (Microsoft Copilot and agents)
    - Sensitive interactions per AI app
    - Top unethical AI interactions
    - Insider Risk severity
    - Insider risk severity per AI app
    - Potential risky AI usage

4. Select **View details** for each of the report graphs to view detailed activities in the activity explorer.

   From the filters, select the **AI app category** of **Copilot experiences & agents**, and then use the other filters if you need to further refine the displayed data. Then drill down to each activity to view details that include displaying the prompts and response when you're a member of the Microsoft Purview Content Explorer Content Viewer role group. For more information about this requirement, see Permissions for Data Security Posture Management for AI.

Apply compliance controls to Copilot in Fabric

1. From **DSPM for AI (classic)** > **Recommendations** page, under **Not Started**, locate, and select and take actions on the following recommendation:

   - **Detect unethical behavior in AI**: This recommendation creates a policy to detect sensitive information in prompts and responses in Copilot in Fabric and agents, and other generative AI apps.

2. If you need to ensure that Copilot in Fabric interactions are retained for compliance reasons:

   In the Microsoft Purview portal, navigate to **Data Lifecycle Management** > **Policies** > **Retention Policies** and create a retention policy to retain Microsoft Security Copilot interactions by selecting the location **Microsoft Copilot Experiences** and specify the required retention period. For more information, see Create and configure retention policies.

3. If you need to preserve, collect, analyze, review, or export Microsoft Fabric interactions:

   In the Microsoft Purview portal, navigate to **eDiscovery** > **Cases** > **Create case**. In the case, create a search and use the *ItemClass* property and the `IPM.SkypeTeams.Message.Copilot.Fabric.*` value to search for these interactions in your organization.

Routinely review the reports in DSPM for AI to determine if you need to make changes, and use activity explorer and events for deeper analysis of how users are interacting with Copilot in Fabric.

Related resources

- Blog post (July 2025): Using Copilot in Fabric with Confidence: Data Security, Compliance & Governance with DSPM for AI

Last updated on 05/01/2026

Use Microsoft Purview to manage data security & compliance for Microsoft Copilot Studio

> *Microsoft Purview service description*

Use the following sections to identify the Microsoft Purview capabilities that are supported for AI interactions with agents built in Microsoft Copilot Studio, and some get started recommendations for you to manage these AI interactions for security and compliance.

Managing these AI interactions for Copilot Studio agents published to non-Microsoft channels requires you to enable pay-as-you-go billing in your organization.

Capabilities supported

Use the following table to see at a glance the Microsoft Purview capabilities that are supported with agents built in Copilot Studio.

Expand table

Capability or solution in Microsoft Purview	Supported for AI interactions
DSPM and DSPM for AI (classic)	✓
Auditing	✓
Data classification	✓
Sensitivity labels	✓
Encryption without sensitivity labels	✗
Data loss prevention	✓
Insider Risk Management	✓
Communication compliance	✓
eDiscovery	✓
Data Lifecycle Management	✓

Capability or solution in Microsoft Purview	Supported for AI interactions
Compliance Manager	✓

DSPM and DSPM for AI (classic)

Use Data Security Posture Management or Data Security Posture Management for AI (classic) as your front door to discover, secure, and apply compliance controls for AI usage across your enterprise. Both DSPM versions use existing controls from Microsoft Purview information protection and compliance management with easy-to-use graphical tools and reports to quickly gain insights into AI use within your organization. With personalized recommendations, and one-click policies help you protect your data and comply with regulatory requirements.

AI app-specific information:

- Recommendation **Get guided assistance to AI regulations**, which uses control-mapping regulatory templates from Compliance Manager.
- One-click policies available:
  - **DSPM for AI - Detect risky AI usage** from the recommendation or remediation action **Detect risky interactions in AI apps**.
  - **DSPM for AI - Unethical behavior in AI apps** from the recommendation or remediation action **Detect unethical behavior in AI apps**
  - **DSPM for AI - Detect sensitive info shared with AI via network** from the recommendation or remediation action **Extend insights into sensitive data in AI app interactions**.

Auditing and AI interactions

Microsoft Purview Audit solutions provide comprehensive tools for searching and managing audit records of activities performed across various Microsoft services by users and admins, and help organizations to effectively respond to security events, forensic investigations, internal investigations, and compliance obligations.

Like other activities, prompts and responses are captured in the unified audit log. Events include how and when users interact with the AI app, and can include in which Microsoft 365 service the activity took place, and references to the files stored in Microsoft 365 that were accessed during the interaction. If these files have a sensitivity label applied, that's also captured.

These events flow into activity explorer in DSPM for AI and the **AI activities** tab in activity explorer from the preview version of DSPM, where the data from prompts and responses can be displayed. You can also use the **Audit** solution from the Microsoft Purview portal to search and find these auditing events.

For more information, see Audit logs for Copilot and AI activities.

Data classification and AI interactions

Microsoft Purview data classification provides a comprehensive framework for identifying and tagging sensitive data across various Microsoft services, including Office 365, Dynamics 365, and Azure. Classifying data is often the first step to ensure compliance with data protection regulations and safeguard against unauthorized access, alteration, or destruction. You can use built-in system classifications or create your own.

Sensitive information types and trainable classifiers can be used to find sensitive data in user prompts and responses when they use AI apps. The resulting information then surfaces in Microsoft Purview Reports overview and activity explorer in DSPM for AI and the **AI activities** tab in activity explorer from the preview version of DSPM.

Sensitivity labels and AI interactions

AI apps that Microsoft Purview support use existing controls to ensure that data stored in your tenant is never returned to the user or used by a large language model (LLM) if the user doesn't have access to that data. When the data has sensitivity labels from your organization applied to the content, there's an extra layer of protection:

- When a file is open in Word, Excel, PowerPoint, or similarly an email or calendar event is open in Outlook, the sensitivity of the data is displayed to users in the app with the label name and content markings (such as header or footer text) that have been configured for the label. Loop components and pages also support the same sensitivity labels.
- When the sensitivity label applies encryption, users must have the EXTRACT usage right, as well as VIEW, for the AI apps to return the data.
- This protection extends to data stored outside your Microsoft 365 tenant when it's open in an Office app (data in use). For example, local storage, network shares, and cloud storage.

> **Tip**

If you haven't already, we recommend you enable sensitivity labels for SharePoint and OneDrive and also familiarize yourself with the file types and label configurations that these services can process. When sensitivity labels aren't enabled for these services, the encrypted files that Copilot and agents can access are limited to data in use from Office apps on Windows.

For instructions, see **Enable sensitivity labels for Office files in SharePoint and OneDrive**.

If you're not already using sensitivity labels, see Get started with sensitivity labels.

AI app-specific information:

- Supported when the knowledge source is SharePoint (includes OneDrive), and Dataverse with auto-labeling policies for the Data Map.
- Encryption is supported only with sensitivity labels, and only for the SharePoint knowledge source.
- Like sensitivity labels in Microsoft 365 Chat, agents built in Copilot Studio display the sensitivity label for items listed in the response and citations. Using the sensitivity labels' priority number that's defined in the Microsoft Purview portal, the latest response in Copilot displays the highest priority sensitivity label from the data used for that Copilot chat.
- Also like sensitivity label inheritance for Microsoft 365 Copilot, newly created content from Word and PowerPoint inherits the highest priority sensitivity label.

Data loss prevention and AI interactions

Microsoft Purview Data Loss Prevention (DLP) helps you identify sensitive items across Microsoft 365 services and endpoints, monitor them, and helps protect against leakage of those items. It uses deep content inspection and contextual analysis to identify sensitive items and it enforces policies to protect sensitive data such as financial records, health information, or intellectual property.

Windows computers that are onboarded to Microsoft Purview can be configured for Endpoint data loss prevention (DLP) policies that warn or block users from sharing sensitive information with third-party generative AI sites that are accessed via a browser. For example, a user is prevented from pasting credit card numbers into ChatGPT, or they see a warning that they can override. For more information about the supported DLP actions and which platforms support them, see the first two rows in the table from Endpoint activities you can monitor and take action on.

AI app-specific information:

- The following Endpoint DLP capability is supported only:
  - When the knowledge source is SharePoint, a DLP policy scoped to the Microsoft 365 Copilot location can restrict agents built in Copilot Studio for Microsoft Teams, SharePoint, and Microsoft 365 Copilot from processing sensitive content based on a specified sensitivity label.

Insider Risk Management and AI interactions

Microsoft Purview Insider Risk Management helps you detect, investigate, and mitigate internal risks such as IP theft, data leakage, and security violations. It leverages machine learning models and various signals from Microsoft 365 and third-party indicators to identify potential malicious or inadvertent insider activities. The solution includes privacy controls like pseudonymization and role-based access, ensuring user-level privacy while enabling risk analysts to take appropriate actions.

Use the Risky AI usage policy template to detect risky usage that includes prompt injection attacks and accessing protected materials. Insights from these signals are integrated into Microsoft Defender XDR to provide a comprehensive view of AI-related risks.

Communication compliance and AI interactions

Microsoft Purview Communication Compliance provides tools to help you detect and manage regulatory compliance and business conduct violations across various communication channels, which include user prompts and responses for AI apps. It's designed with privacy by default, pseudonymizing usernames and incorporating role-based access controls. The solution helps identify and remediate inappropriate communications, such as sharing sensitive information, harassment, threats, and adult content.

To learn more about using communication compliance policies for AI apps, see Configure a communication compliance policy to detect for generative AI interactions.

eDiscovery and AI interactions

Microsoft Purview eDiscovery lets you identify and deliver electronic information that can be used as evidence in legal cases. The eDiscovery tools in Microsoft Purview support searching for content in Exchange Online, OneDrive for Business, SharePoint Online, Microsoft Teams, Microsoft 365 Groups, and Viva Engage teams. You can then prevent the information from deletion and export the information.

Because user prompts and responses for AI apps are stored in a user's mailbox, you can create a case and use search when a user's mailbox is selected as the source for a search query. For example, select and retrieve this data from the source mailbox by selecting from the query builder **Add condition** > **Type** > **Contains any of** > **Edit** > *Copilot activity*. This query condition includes all Copilot and other AI application activity.

After the search is refined, you can export the results or add to a review set. You can review and export information directly from the review set.

For more information about identifying and deleting user AI interaction data, see Search for and delete Copilot data in eDiscovery.

Data Lifecycle Management and AI interactions

Microsoft Purview Data Lifecycle Management provides tools and capabilities to manage the lifecycle of organizational data by retaining necessary content and deleting unnecessary content. These tools ensure compliance with business, legal, and regulatory requirements.

Use retention policies to automatically retain or delete user prompts and responses for AI apps. For detailed information about this retention works, see Learn about retention for Copilot & AI apps.

As with all retention policies and holds, if more than one policy for the same location applies to a user, the principles of retention resolve any conflicts. For example, the data is retained for the longest duration of all the applied retention policies or eDiscovery holds.

AI app-specific information:

- For retention policies, select the option for **Microsoft Copilot Experiences**.

Compliance Manager and AI interactions

Microsoft Purview Compliance Manager is a solution that helps you automatically assess and manage compliance across your multicloud environment. Compliance Manager can help you throughout your compliance journey, from taking inventory of your data protection risks to managing the complexities of implementing controls, staying current with regulations and certifications, and reporting to auditors.

To help you keep compliant with AI regulations, Compliance Manager provides regulatory templates to help you assess, implement, and strengthen your compliance requirements for all

generative AI apps. For example, monitoring AI interactions and preventing data loss in AI applications. For more information, see Assessments for AI regulations.

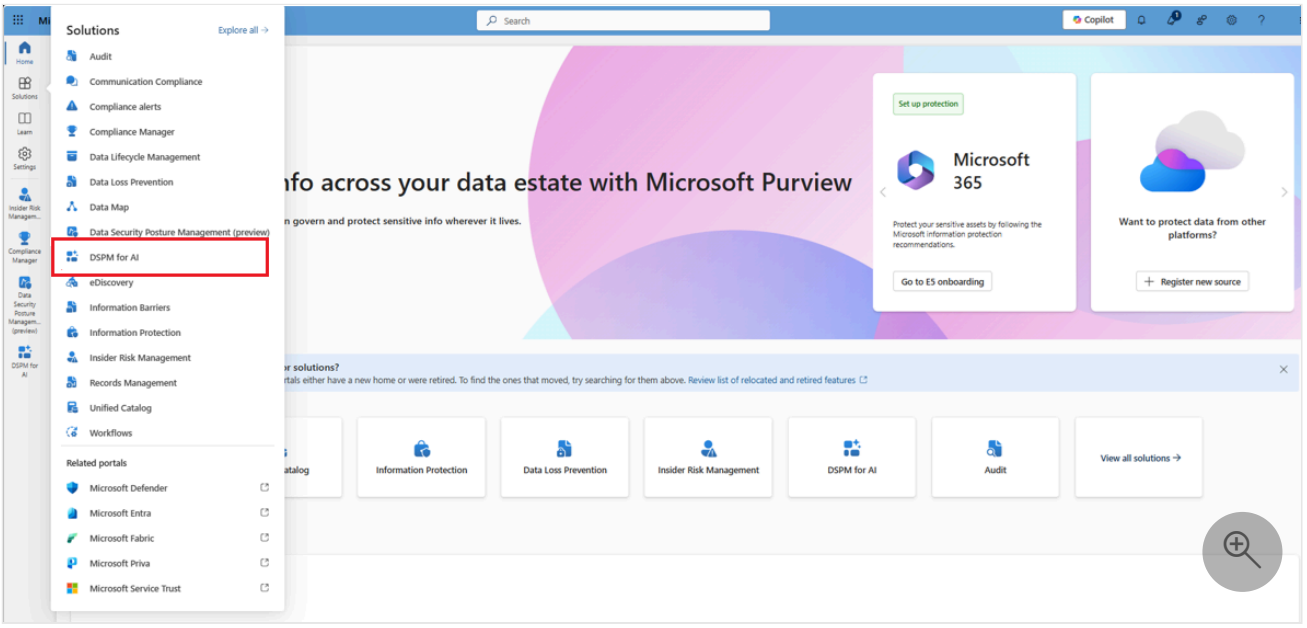
Getting started recommended steps

Use the following steps to get started with managing data security & compliance for interactions from agents built in Microsoft Copilot Studio.

> **Note**
>
> These steps focus on managing a specific AI app or agent. For broader coverage that uses security objectives with guided workflows, use the current version of Data Security Posture Management.

Because Data Security Posture Management for AI is your front door for securing and managing AI interactions, most of the following instructions use that solution:

- Sign in to the Microsoft Purview portal > **Solutions** > **DSPM for AI (classic)** with an account that has appropriate permissions. For example, an account that's a member of the **Microsoft Entra Compliance Administrator** group role.



Discover potential security risks in Microsoft Copilot Studio interactions

1. From **DSPM for AI (classic)** > **Overview**, in the **Get Started** section, look to see if auditing is on for your tenant. If not, select **Activate Microsoft Purview Audit**.

2. In the **Recommendations** section, select **View all recommendations** to take you to the **Recommendations** page. If the following recommendation has a status of **Not Started**, take action on it to create a one-click policy:

    - **Detect risky interactions in AI apps**: This recommendation creates a policy to help calculate user risk by detecting risky prompts and responses in Microsoft 365 Copilot, agents, and other generative AI apps.

3. Wait at least a day for data, and then navigate to the **Reports** page to view the results of your policy. Select **Copilot experiences & agents** and view information such as:

    - Total interactions over time (Microsoft Copilot and agents)
    - Sensitive interactions per AI app
    - Top unethical AI interactions
    - Insider Risk severity
    - Insider risk severity per AI app
    - Potential risky AI usage

4. Select **View details** for each of the report graphs to view detailed activities in the activity explorer.

    From the filters, select the **AI app category** of **Copilot experiences & agents**, and then use the other filters if you need to further refine the displayed data. Then drill down to each activity to view details that include displaying the prompts and response when you're a member of the Microsoft Purview Content Explorer Content Viewer role group. For more information about this requirement, see Permissions for Data Security Posture Management for AI.

Apply compliance controls to Microsoft Copilot Studio

1. From **DSPM for AI (classic)** > **Recommendations** page, under **Not Started**, locate, and select and take actions on the following recommendation:

    - **Detect unethical behavior in AI**: This recommendation creates a policy to detect sensitive information in prompts and responses in agents built in Copilot Studio, and other generative AI apps.

2. If you need to ensure that interactions from agents built in Copilot Studio are retained for compliance reasons:

In the Microsoft Purview portal, navigate to **Data Lifecycle Management** > **Policies** > **Retention Policies** and create a retention policy to retain Microsoft Security Copilot interactions by selecting the location **Microsoft Copilot Experiences** and specify the required retention period. For more information, see Create and configure retention policies.

3. If you need to preserve, collect, analyze, review, or export Microsoft Copilot Studio interactions:

   In the Microsoft Purview portal, navigate to **eDiscovery** > **Cases** > **Create case**. In the case, create a search and use the *ItemClass* property and the `IPM.SkypeTeams.Message.Copilot.Studio.*` value to search for these interactions in your organization.

Routinely review the reports and data risk assessments in DSPM for AI to determine if you need to make changes, and use activity explorer and events for deeper analysis of how users are interacting with agents built in Copilot Studio.

Last updated on 05/01/2026

Use Microsoft Purview to manage data security & compliance for Microsoft Facilitator

> *Microsoft Purview service description*

Use the following sections to identify the Microsoft Purview capabilities that are supported for AI interactions with Microsoft Facilitator, and some get started recommendations for you to manage these AI interactions for security and compliance.

For additional information, see Set up Facilitator in Microsoft Teams.

Capabilities supported

Use the following table to see at a glance the Microsoft Purview capabilities that are supported with Microsoft Facilitator.

Expand table

Capability or solution in Microsoft Purview	Supported for AI interactions
DSPM and DSPM for AI (classic)	✓
Auditing	✓
Data classification	✓
Sensitivity labels	✗
Encryption without sensitivity labels	✗
Data loss prevention	✗
Insider Risk Management	✗
Communication compliance	✓
eDiscovery	✓
Data Lifecycle Management	✓
Compliance Manager	✓

DSPM and DSPM for AI (classic)

Use Data Security Posture Management or Data Security Posture Management for AI (classic) as your front door to discover, secure, and apply compliance controls for AI usage across your enterprise. Both DSPM versions use existing controls from Microsoft Purview information protection and compliance management with easy-to-use graphical tools and reports to quickly gain insights into AI use within your organization. With personalized recommendations, and one-click policies help you protect your data and comply with regulatory requirements.

AI app-specific information:

- One-click policies available:
  - **DSPM for AI - Detect risky AI usage** from the recommendation or remediation action **Detect risky interactions in AI apps**.
  - **DSPM for AI - Unethical behavior in AI apps** from the recommendation or remediation action **Detect unethical behavior in AI**.
  - **DSPM for AI - Detect sensitive info shared with AI via network** from the recommendation or remediation action **Extend insights into sensitive data in AI app interactions**.

Auditing and AI interactions

Microsoft Purview Audit solutions provide comprehensive tools for searching and managing audit records of activities performed across various Microsoft services by users and admins, and help organizations to effectively respond to security events, forensic investigations, internal investigations, and compliance obligations.

Like other activities, prompts and responses are captured in the unified audit log. Events include how and when users interact with the AI app, and can include in which Microsoft 365 service the activity took place, and references to the files stored in Microsoft 365 that were accessed during the interaction. If these files have a sensitivity label applied, that's also captured.

These events flow into activity explorer in DSPM for AI and the **AI activities** tab in activity explorer from the preview version of DSPM, where the data from prompts and responses can be displayed. You can also use the **Audit** solution from the Microsoft Purview portal to search and find these auditing events.

For more information, see Audit logs for Copilot and AI activities.

AI app-specific information:

- Facilitator AI-generated notes, moderation, and questions and answers for meetings are supported.

Data classification and AI interactions

Microsoft Purview data classification provides a comprehensive framework for identifying and tagging sensitive data across various Microsoft services, including Office 365, Dynamics 365, and Azure. Classifying data is often the first step to ensure compliance with data protection regulations and safeguard against unauthorized access, alteration, or destruction. You can use built-in system classifications or create your own.

Sensitive information types and trainable classifiers can be used to find sensitive data in user prompts and responses when they use AI apps. The resulting information then surfaces in Microsoft Purview Reports overview and activity explorer in DSPM for AI and the **AI activities** tab in activity explorer from the preview version of DSPM.

Insider Risk Management and AI interactions

Microsoft Purview Insider Risk Management helps you detect, investigate, and mitigate internal risks such as IP theft, data leakage, and security violations. It leverages machine learning models and various signals from Microsoft 365 and third-party indicators to identify potential malicious or inadvertent insider activities. The solution includes privacy controls like pseudonymization and role-based access, ensuring user-level privacy while enabling risk analysts to take appropriate actions.

Use the Risky AI usage policy template to detect risky usage that includes prompt injection attacks and accessing protected materials. Insights from these signals are integrated into Microsoft Defender XDR to provide a comprehensive view of AI-related risks.

Communication compliance and AI interactions

Microsoft Purview Communication Compliance provides tools to help you detect and manage regulatory compliance and business conduct violations across various communication channels, which include user prompts and responses for AI apps. It's designed with privacy by default, pseudonymizing usernames and incorporating role-based access controls. The solution helps identify and remediate inappropriate communications, such as sharing sensitive information, harassment, threats, and adult content.

To learn more about using communication compliance policies for AI apps, see Configure a communication compliance policy to detect for generative AI interactions.

eDiscovery and AI interactions

Microsoft Purview eDiscovery lets you identify and deliver electronic information that can be used as evidence in legal cases. The eDiscovery tools in Microsoft Purview support searching for content in Exchange Online, OneDrive for Business, SharePoint Online, Microsoft Teams, Microsoft 365 Groups, and Viva Engage teams. You can then prevent the information from deletion and export the information.

Because user prompts and responses for AI apps are stored in a user's mailbox, you can create a case and use search when a user's mailbox is selected as the source for a search query. For example, select and retrieve this data from the source mailbox by selecting from the query builder **Add condition** > **Type** > **Contains any of** > **Edit** > *Copilot activity*. This query condition includes all Copilot and other AI application activity.

After the search is refined, you can export the results or add to a review set. You can review and export information directly from the review set.

For more information about identifying and deleting user AI interaction data, see Search for and delete Copilot data in eDiscovery.

AI app-specific information:

- Facilitator AI-generated notes, moderation, and questions and answers for meetings are supported with a KeyQL query.

Data Lifecycle Management and AI interactions

Microsoft Purview Data Lifecycle Management provides tools and capabilities to manage the lifecycle of organizational data by retaining necessary content and deleting unnecessary content. These tools ensure compliance with business, legal, and regulatory requirements.

Use retention policies to automatically retain or delete user prompts and responses for AI apps. For detailed information about this retention works, see Learn about retention for Copilot & AI apps.

As with all retention policies and holds, if more than one policy for the same location applies to a user, the principles of retention resolve any conflicts. For example, the data is retained for the longest duration of all the applied retention policies or eDiscovery holds.

AI app-specific information:

- Configuration for retention and deletion:
  - For Facilitator AI interactions in meetings, use a retention policy with the **Teams chats** option.
  - For Facilitator AI-generated notes in meetings, use a retention policy or retention labels with the **OneDrive accounts** option.

Compliance Manager and AI interactions

Microsoft Purview Compliance Manager is a solution that helps you automatically assess and manage compliance across your multicloud environment. Compliance Manager can help you throughout your compliance journey, from taking inventory of your data protection risks to managing the complexities of implementing controls, staying current with regulations and certifications, and reporting to auditors.

To help you keep compliant with AI regulations, Compliance Manager provides regulatory templates to help you assess, implement, and strengthen your compliance requirements for all generative AI apps. For example, monitoring AI interactions and preventing data loss in AI applications. For more information, see Assessments for AI regulations.

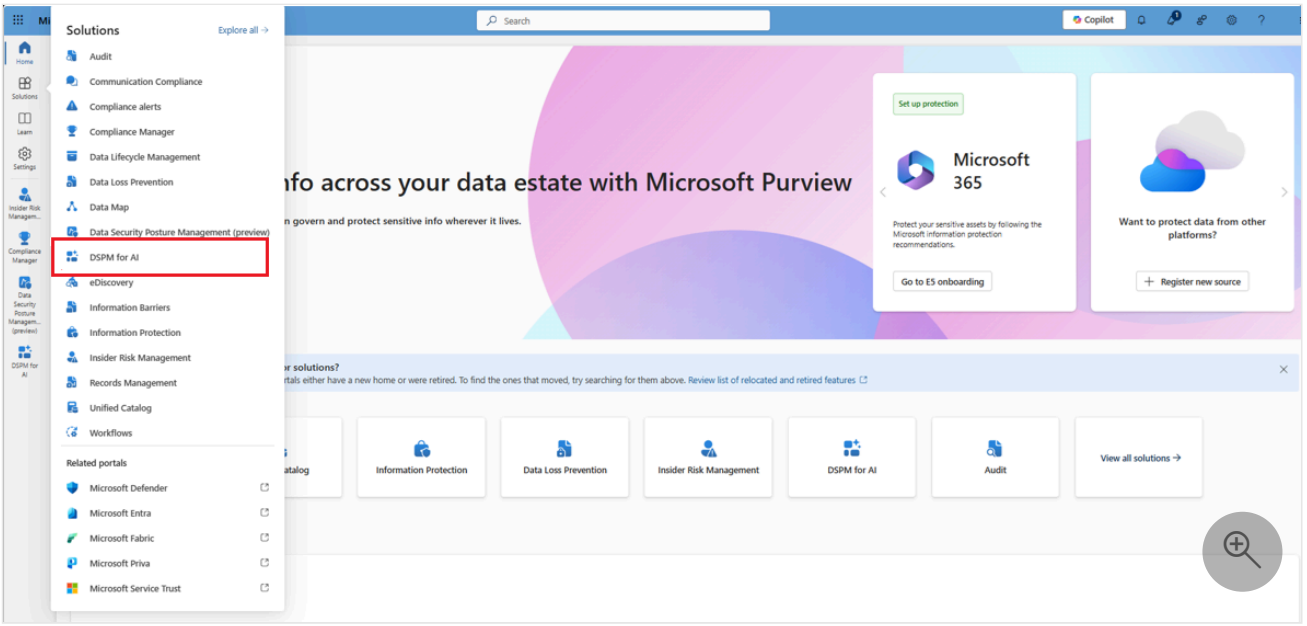
Getting started recommended steps

Use the following steps to get started with managing data security & compliance for Microsoft Facilitator AI interactions.

> **Note**
>
> These steps focus on managing a specific AI app or agent. For broader coverage that uses security objectives with guided workflows, use the current version of Data Security Posture Management.

Because Data Security Posture Management for AI is your front door for securing and managing AI interactions, the following instructions use that solution:

- Sign in to the Microsoft Purview portal > **Solutions** > **DSPM for AI (classic)** with an account that has appropriate permissions. For example, an account that's a member of the **Microsoft Entra Compliance Administrator** group role.



Discover interactions from Microsoft Facilitator

1. From **DSPM for AI (classic)** > **Overview**, in the **Get Started** section, look to see if auditing is on for your tenant. If not, select **Activate Microsoft Purview Audit**.

2. In the **Recommendations** section, select **View all recommendations** to take you to the **Recommendations** page. If the following recommendation has a status of **Not Started**, select it to create a one-click policy:

   - **Detect risky interactions in AI apps**: This recommendation creates a policy to help calculate user risk by detecting risky prompts and responses in Microsoft 365 Copilot, agents, and other generative AI apps.

3. Wait at least a day for data, and then navigate to the **Reports** page. Select the AI app category of **Copilot experiences & agents** and view information such as:

   - Total interactions over time (Microsoft Copilot and agents)
   - Sensitive interactions per AI app
   - Insider Risk severity
   - Insider risk severity per AI app

4. Select **View details** for each of the report graphs to view detailed activities in the activity explorer.

   Select the available filters to see the results from **Copilot experiences & agents** based on the different **Activity type**, **AI app category**, **App**, and **Scope** for administrative units, and more. Then drill down to each activity to view details that include displaying the prompts and response when you're a member of the Microsoft Purview Content Explorer Content

Viewer role group. For more information about this requirement, see Permissions for Data Security Posture Management for AI.

Apply compliance controls to interactions with Microsoft Facilitator

1. If you need to ensure that interactions with Microsoft Facilitator are retained for compliance reasons:

   In the Microsoft Purview portal, navigate to **Data Lifecycle Management** > **Policies** > **Retention Policies** and create a retention policy with the following policy locations: - **Teams chats** for Facilitator interactions in meetings - **OneDrive accounts** for Facilitator AI-generated notes in meetings

2. If you need to preserve, collect, analyze, review, or export Microsoft Facilitator interactions:

   In the Microsoft Purview portal, navigate to **eDiscovery** > **Cases** > **Create case**. In the case, create a search and use the *ItemClass* property and the `IPM.SkypeTeams.Message` value to search for these interactions in your organization.

Last updated on 05/01/2026

Use Microsoft Purview to manage data security & compliance for Channel Agent in Teams

> *Microsoft Purview service description*

Use the following sections to identify the Microsoft Purview capabilities that are supported for AI interactions with agents in Teams channels, and some get started recommendations for you to manage these AI interactions for security and compliance.

For additional information, see Get started with Channel Agent in Microsoft Teams.

Capabilities supported

Use the following table to see at a glance the Microsoft Purview capabilities that are supported with Channel Agent in Teams.

Expand table

Capability or solution in Microsoft Purview	Supported for AI interactions
DSPM and DSPM for AI (classic)	✓
Auditing	✓
Data classification	✓
Sensitivity labels	✓
Encryption without sensitivity labels	✗
Data loss prevention	✓
Insider Risk Management	✗
Communication compliance	✓
eDiscovery	✓
Data Lifecycle Management	✓
Compliance Manager	✓

DSPM and DSPM for AI (classic)

Use Data Security Posture Management or Data Security Posture Management for AI (classic) as your front door to discover, secure, and apply compliance controls for AI usage across your enterprise. Both DSPM versions use existing controls from Microsoft Purview information protection and compliance management with easy-to-use graphical tools and reports to quickly gain insights into AI use within your organization. With personalized recommendations, and one-click policies help you protect your data and comply with regulatory requirements.

AI app-specific information:

- One-click policies available:
  - **DSPM for AI - Detect risky AI usage** from the recommendation or remediation action **Detect risky interactions in AI apps**.
  - **DSPM for AI - Unethical behavior in AI apps** from the recommendation or remediation action **Detect unethical behavior in AI**.
  - **DSPM for AI - Detect sensitive info shared with AI via network** from the recommendation or remediation action **Extend insights into sensitive data in AI app interactions**.

Auditing and AI interactions

Microsoft Purview Audit solutions provide comprehensive tools for searching and managing audit records of activities performed across various Microsoft services by users and admins, and help organizations to effectively respond to security events, forensic investigations, internal investigations, and compliance obligations.

Like other activities, prompts and responses are captured in the unified audit log. Events include how and when users interact with the AI app, and can include in which Microsoft 365 service the activity took place, and references to the files stored in Microsoft 365 that were accessed during the interaction. If these files have a sensitivity label applied, that's also captured.

These events flow into activity explorer in DSPM for AI and the **AI activities** tab in activity explorer from the preview version of DSPM, where the data from prompts and responses can be displayed. You can also use the **Audit** solution from the Microsoft Purview portal to search and find these auditing events.

For more information, see Audit logs for Copilot and AI activities.

Data classification and AI interactions

Microsoft Purview data classification provides a comprehensive framework for identifying and tagging sensitive data across various Microsoft services, including Office 365, Dynamics 365, and Azure. Classifying data is often the first step to ensure compliance with data protection regulations and safeguard against unauthorized access, alteration, or destruction. You can use built-in system classifications or create your own.

Sensitive information types and trainable classifiers can be used to find sensitive data in user prompts and responses when they use AI apps. The resulting information then surfaces in Microsoft Purview Reports overview and activity explorer in DSPM for AI and the **AI activities** tab in activity explorer from the preview version of DSPM.

Sensitivity labels and AI interactions

AI apps that Microsoft Purview support use existing controls to ensure that data stored in your tenant is never returned to the user or used by a large language model (LLM) if the user doesn't have access to that data. When the data has sensitivity labels from your organization applied to the content, there's an extra layer of protection:

- When a file is open in Word, Excel, PowerPoint, or similarly an email or calendar event is open in Outlook, the sensitivity of the data is displayed to users in the app with the label name and content markings (such as header or footer text) that have been configured for the label. Loop components and pages also support the same sensitivity labels.
- When the sensitivity label applies encryption, users must have the EXTRACT usage right, as well as VIEW, for the AI apps to return the data.
- This protection extends to data stored outside your Microsoft 365 tenant when it's open in an Office app (data in use). For example, local storage, network shares, and cloud storage.

> **Tip**
>
> If you haven't already, we recommend you enable sensitivity labels for SharePoint and OneDrive and also familiarize yourself with the file types and label configurations that these services can process. When sensitivity labels aren't enabled for these services, the encrypted files that Copilot and agents can access are limited to data in use from Office apps on Windows.
>
> For instructions, see **Enable sensitivity labels for Office files in SharePoint and OneDrive**.

If you're not already using sensitivity labels, see Get started with sensitivity labels.

AI app-specific information:

- Channel Agent in Teams responses display the sensitivity label for items listed in the response and citations. Using the sensitivity labels' priority number that's defined in the Microsoft Purview portal, the latest response from Channel Agent displays the highest priority sensitivity label from the data used for that channel chat.

  Although compliance admins define a sensitivity label's priority, a higher priority number usually denotes higher sensitivity of the content, with more restrictive permissions. As a result, the Channel Agent responses display the most restrictive sensitivity label to educate the user about the sensitivity of the data.

- When Channel Agent creates newly created content that supports sensitivity labels, that action supports sensitivity label inheritance. See the following section for more information.

- Currently, Channel Agent could summarize for an authorized user the content of encrypted items that one or more users in the channel don't have permissions to open. The requesting user is warned of this risk and administrators always have the ability to turn off Channel Agent in Teams. For additional information about this and other deployment considerations, see Information protection considerations for Channel Agent in Teams.

Sensitivity label inheritance

If you use Channel Agent in Teams to create new content based on an item that has a sensitivity label applied, the sensitivity label from the source item is automatically inherited, with the label's protection settings.

For example, you ask an agent to create a Loop component based on a file that's uploaded to the channel. The file has the sensitivity label **Confidential\Project Bluebird** applied and that label is configured to apply encryption that restricts access to people involved with the project. The new Loop component is automatically labeled **Confidential\Project Bluebird** and can't be opened by people who aren't involved with the project.

AI-generated items that currently support sensitivity label inheritance for Channel Agent in Teams: Loop components and pages, and files in OneDrive and SharePoint.

If multiple labeled sources are used to create new content, the sensitivity label with the highest priority is used for label inheritance.

As with all automatic labeling scenarios, a user can always override and replace an inherited label (or remove, if you're not using mandatory labeling).

Data loss prevention and AI interactions

Microsoft Purview Data Loss Prevention (DLP) helps you identify sensitive items across Microsoft 365 services and endpoints, monitor them, and helps protect against leakage of those items. It uses deep content inspection and contextual analysis to identify sensitive items and it enforces policies to protect sensitive data such as financial records, health information, or intellectual property.

Windows computers that are onboarded to Microsoft Purview can be configured for Endpoint data loss prevention (DLP) policies that warn or block users from sharing sensitive information with third-party generative AI sites that are accessed via a browser. For example, a user is prevented from pasting credit card numbers into ChatGPT, or they see a warning that they can override. For more information about the supported DLP actions and which platforms support them, see the first two rows in the table from Endpoint activities you can monitor and take action on.

AI app-specific information:

- The following DLP capability only is supported for Channel Agent in Teams:
  - When an agent interacts with a user in a Teams conversation, a DLP policy can scan the messages for sensitive data.

Insider Risk Management and AI interactions

Microsoft Purview Insider Risk Management helps you detect, investigate, and mitigate internal risks such as IP theft, data leakage, and security violations. It leverages machine learning models and various signals from Microsoft 365 and third-party indicators to identify potential malicious or inadvertent insider activities. The solution includes privacy controls like pseudonymization and role-based access, ensuring user-level privacy while enabling risk analysts to take appropriate actions.

Use the Risky AI usage policy template to detect risky usage that includes prompt injection attacks and accessing protected materials. Insights from these signals are integrated into Microsoft Defender XDR to provide a comprehensive view of AI-related risks.

Communication compliance and AI interactions

Microsoft Purview Communication Compliance provides tools to help you detect and manage regulatory compliance and business conduct violations across various communication channels, which include user prompts and responses for AI apps. It's designed with privacy by default,

pseudonymizing usernames and incorporating role-based access controls. The solution helps identify and remediate inappropriate communications, such as sharing sensitive information, harassment, threats, and adult content.

To learn more about using communication compliance policies for AI apps, see Configure a communication compliance policy to detect for generative AI interactions.

eDiscovery and AI interactions

Microsoft Purview eDiscovery lets you identify and deliver electronic information that can be used as evidence in legal cases. The eDiscovery tools in Microsoft Purview support searching for content in Exchange Online, OneDrive for Business, SharePoint Online, Microsoft Teams, Microsoft 365 Groups, and Viva Engage teams. You can then prevent the information from deletion and export the information.

Because user prompts and responses for AI apps are stored in a user's mailbox, you can create a case and use search when a user's mailbox is selected as the source for a search query. For example, select and retrieve this data from the source mailbox by selecting from the query builder **Add condition** > **Type** > **Contains any of** > **Edit** > *Copilot activity*. This query condition includes all Copilot and other AI application activity.

After the search is refined, you can export the results or add to a review set. You can review and export information directly from the review set.

For more information about identifying and deleting user AI interaction data, see Search for and delete Copilot data in eDiscovery.

Data Lifecycle Management and AI interactions

Microsoft Purview Data Lifecycle Management provides tools and capabilities to manage the lifecycle of organizational data by retaining necessary content and deleting unnecessary content. These tools ensure compliance with business, legal, and regulatory requirements.

Use retention policies to automatically retain or delete user prompts and responses for AI apps. For detailed information about this retention works, see Learn about retention for Copilot & AI apps.

As with all retention policies and holds, if more than one policy for the same location applies to a user, the principles of retention resolve any conflicts. For example, the data is retained for the longest duration of all the applied retention policies or eDiscovery holds.

AI app-specific information:

- Configuration for retention and deletion:
  - For Channel Agent in Teams interactions, use a retention policy with the **Teams chats** option.

Compliance Manager and AI interactions

Microsoft Purview Compliance Manager is a solution that helps you automatically assess and manage compliance across your multicloud environment. Compliance Manager can help you throughout your compliance journey, from taking inventory of your data protection risks to managing the complexities of implementing controls, staying current with regulations and certifications, and reporting to auditors.

To help you keep compliant with AI regulations, Compliance Manager provides regulatory templates to help you assess, implement, and strengthen your compliance requirements for all generative AI apps. For example, monitoring AI interactions and preventing data loss in AI applications. For more information, see Assessments for AI regulations.

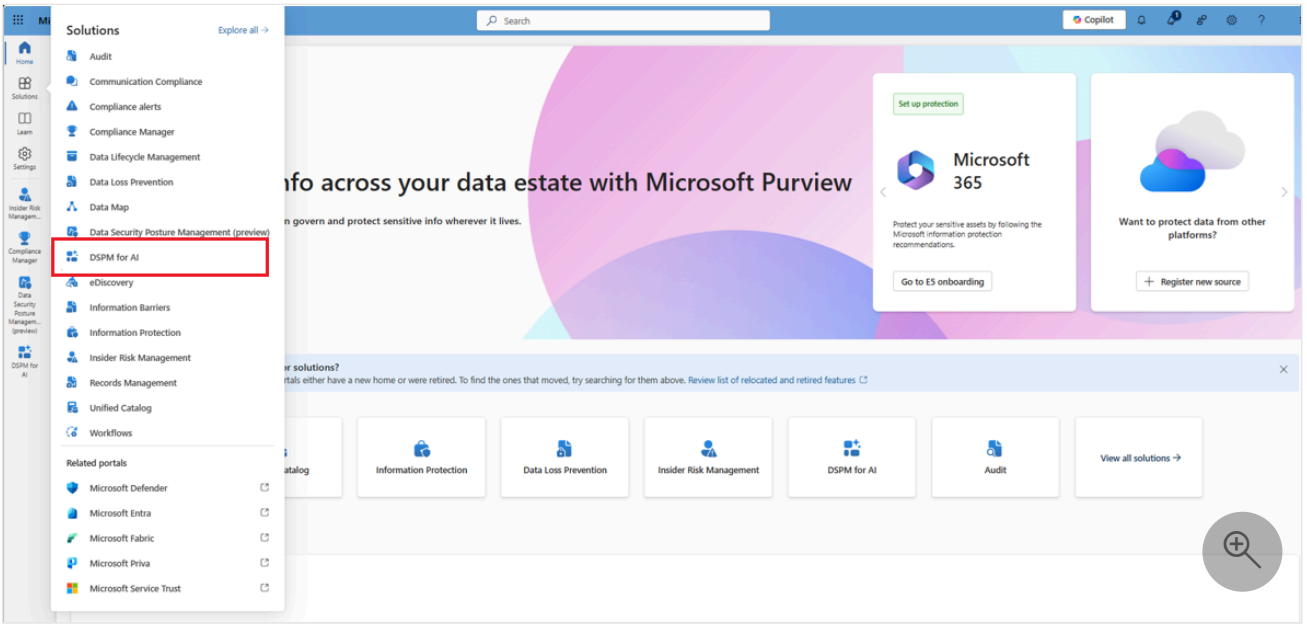
Getting started recommended steps

Use the following steps to get started with managing data security & compliance for AI interactions with Channel Agent in Teams.

> **Note**
>
> These steps focus on managing a specific AI app or agent. For broader coverage that uses security objectives with guided workflows, use the current version of Data Security Posture Management.

Because Data Security Posture Management for AI is your front door for securing and managing AI interactions, the following instructions use that solution:

- Sign in to the Microsoft Purview portal > **Solutions** > **DSPM for AI (classic)** with an account that has appropriate permissions. For example, an account that's a member of the **Microsoft Entra Compliance Administrator** group role.



Discover interactions from Channel Agent in Teams

1. From **DSPM for AI (classic)** > **Overview**, in the **Get Started** section, look to see if auditing is on for your tenant. If not, select **Activate Microsoft Purview Audit**.

2. In the **Recommendations** section, select **View all recommendations** to take you to the **Recommendations** page. If the following recommendation has a status of **Not Started**, select it to create a one-click policy:

   - **Detect risky interactions in AI apps**: This recommendation creates a policy to help calculate user risk by detecting risky prompts and responses in Microsoft 365 Copilot, agents, and other generative AI apps.

3. Wait at least a day for data, and then navigate to the **Reports** page. Select the AI app category of **Copilot experiences & agents** and view information such as:

   - Total interactions over time (Microsoft Copilot and agents)
   - Sensitive interactions per AI app
   - Insider Risk severity
   - Insider risk severity per AI app

4. Select **View details** for each of the report graphs to view detailed activities in the activity explorer.

   Select the available filters to see the results from **Copilot experiences & agents** based on the different **Activity type**, **AI app category**, **App**, and **Scope** for administrative units, and more. Then drill down to each activity to view details that include displaying the prompts and response when you're a member of the Microsoft Purview Content Explorer Content

Viewer role group. For more information about this requirement, see Permissions for Data Security Posture Management for AI.

Apply compliance controls to interactions with Channel Agent in Teams

1. If you need to ensure that interactions with Channel Agent in Teams are retained for compliance reasons:

   In the Microsoft Purview portal, navigate to **Data Lifecycle Management** > **Policies** > **Retention Policies** and create a retention policy with the following policy location: - **Teams chats** for interactions with a channel agent

2. If you need to preserve, collect, analyze, review, or export interactions for Channel Agent in Teams:

   In the Microsoft Purview portal, navigate to **eDiscovery** > **Cases** > **Create case**. In the case, create a search and use the *ItemClass* property and the `IPM.SkypeTeams.Message` value to search for these interactions in your organization.

Other documentation to help you secure and manage generative AI apps

Securing and managing AI interactions with Channel Agent in Teams has much in common with securing and managing AI interactions with Microsoft 365 Copilot. For more detailed information, see Considerations for Microsoft Purview to manage Microsoft 365 Copilot and Channel Agent in Teams for security and compliance.

Last updated on 05/01/2026

Use Microsoft Purview to manage data security & compliance for Microsoft Agent 365

> *Microsoft Purview service description*

Capabilities supported

Use the following table to see at a glance the Microsoft Purview capabilities that are supported for Microsoft Agent 365.

Expand table

Capability or solution in Microsoft Purview	Supported for AI interactions
DSPM and DSPM for AI (classic)	✓
Auditing	✓
Data classification	✓
Sensitivity labels	✓
Encryption without sensitivity labels	✗
Data loss prevention	✓
Insider Risk Management	✓
Communication compliance	✓
eDiscovery	✓
Data lifecycle management	✓
Compliance manager	✓

When you create an agent instance for Agent 365, it's automatically enabled for audit, detection of sensitive data with data classification, and included with the Assessments for AI regulations from Compliance Manager. For other capabilities, include the agent instance in policies as you would a user.

DSPM and DSPM for AI (classic)

Use Data Security Posture Management or Data Security Posture Management for AI (classic) as your front door to discover, secure, and apply compliance controls for AI usage across your enterprise. Both DSPM versions use existing controls from Microsoft Purview information protection and compliance management with easy-to-use graphical tools and reports to quickly gain insights into AI use within your organization. With personalized recommendations, and one-click policies help you protect your data and comply with regulatory requirements.

AI app-specific information:

- Supported for Data Security Posture Management only, use the **AI observability** page to gain visibility into active agent instances to better understand potential risks and how to remediate them. For more information, see the Getting started recommended steps.

Auditing and AI interactions

Microsoft Purview Audit solutions provide comprehensive tools for searching and managing audit records of activities performed across various Microsoft services by users and admins, and help organizations to effectively respond to security events, forensic investigations, internal investigations, and compliance obligations.

Like other activities, prompts and responses are captured in the unified audit log. Events include how and when users interact with the AI app, and can include in which Microsoft 365 service the activity took place, and references to the files stored in Microsoft 365 that were accessed during the interaction. If these files have a sensitivity label applied, that's also captured.

These events flow into activity explorer in DSPM for AI and the **AI activities** tab in activity explorer from the preview version of DSPM, where the data from prompts and responses can be displayed. You can also use the **Audit** solution from the Microsoft Purview portal to search and find these auditing events.

For more information, see Audit logs for Copilot and AI activities.

AI app-specific information:

- Audit an agent instance as you would a user.
- Supported interactions: All agent-to-human, human-to-agent, agent-to-tools, and agent-to-agent interactions.
- In addition to displaying in activity explorer in DSPM, search the audit log for Agent 365 activities.

Data classification and AI interactions

Microsoft Purview data classification provides a comprehensive framework for identifying and tagging sensitive data across various Microsoft services, including Office 365, Dynamics 365, and Azure. Classifying data is often the first step to ensure compliance with data protection regulations and safeguard against unauthorized access, alteration, or destruction. You can use built-in system classifications or create your own.

Sensitive information types and trainable classifiers can be used to find sensitive data in user prompts and responses when they use AI apps. The resulting information then surfaces in Microsoft Purview Reports overview and activity explorer in DSPM for AI and the **AI activities** tab in activity explorer from the preview version of DSPM.

Sensitivity labels and AI interactions

AI apps that Microsoft Purview support use existing controls to ensure that data stored in your tenant is never returned to the user or used by a large language model (LLM) if the user doesn't have access to that data. When the data has sensitivity labels from your organization applied to the content, there's an extra layer of protection:

- When a file is open in Word, Excel, PowerPoint, or similarly an email or calendar event is open in Outlook, the sensitivity of the data is displayed to users in the app with the label name and content markings (such as header or footer text) that have been configured for the label. Loop components and pages also support the same sensitivity labels.
- When the sensitivity label applies encryption, users must have the EXTRACT usage right, as well as VIEW, for the AI apps to return the data.
- This protection extends to data stored outside your Microsoft 365 tenant when it's open in an Office app (data in use). For example, local storage, network shares, and cloud storage.

> **Tip**
>
> If you haven't already, we recommend you enable sensitivity labels for SharePoint and OneDrive and also familiarize yourself with the file types and label configurations that these services can process. When sensitivity labels aren't enabled for these services, the encrypted files that Copilot and agents can access are limited to data in use from Office apps on Windows.

For instructions, see **Enable sensitivity labels for Office files in SharePoint and OneDrive**.

If you're not already using sensitivity labels, see Get started with sensitivity labels.

AI app-specific information:

- For agent instances to access files, the files must be explicitly shared with them.
- If an applied sensitivity label is configured for encryption, that encryption must explicitly grant the agent instance VIEW and EXTRACT usage rights. For example, a sensitivity label that's configured for **Add all users and groups in your organization** or **Add any authenticated users** wouldn't be sufficient.
- Newly created content from Agent 365 doesn't inherit sensitivity labels from the source items. As a result, that content isn't automatically labeled or encrypted.

Data loss prevention and AI interactions

Microsoft Purview Data Loss Prevention (DLP) helps you identify sensitive items across Microsoft 365 services and endpoints, monitor them, and helps protect against leakage of those items. It uses deep content inspection and contextual analysis to identify sensitive items and it enforces policies to protect sensitive data such as financial records, health information, or intellectual property.

Windows computers that are onboarded to Microsoft Purview can be configured for Endpoint data loss prevention (DLP) policies that warn or block users from sharing sensitive information with third-party generative AI sites that are accessed via a browser. For example, a user is prevented from pasting credit card numbers into ChatGPT, or they see a warning that they can override. For more information about the supported DLP actions and which platforms support them, see the first two rows in the table from Endpoint activities you can monitor and take action on.

AI app-specific information:

- Supported by explicitly specifying agent instances in the DLP policy as you would a user, or by specifying a security group that includes agent instances.
- Supported interactions: Block or audit agent-to-human and human-to-agent for Microsoft Teams, OneDrive or SharePoint, and emails.
- Because an agent instance is unaware of the block action, the agent owner must actively monitor a DLP policy that uses this configuration and understand the impact to subsequent workflows.

Insider Risk Management and AI interactions

Microsoft Purview Insider Risk Management helps you detect, investigate, and mitigate internal risks such as IP theft, data leakage, and security violations. It leverages machine learning models and various signals from Microsoft 365 and third-party indicators to identify potential malicious or inadvertent insider activities. The solution includes privacy controls like pseudonymization and role-based access, ensuring user-level privacy while enabling risk analysts to take appropriate actions.

Use the Risky AI usage policy template to detect risky usage that includes prompt injection attacks and accessing protected materials. Insights from these signals are integrated into Microsoft Defender XDR to provide a comprehensive view of AI-related risks.

AI app-specific information:

- Supported by explicitly specifying agent instances in the Insider Risk Management policy as you would a user.
- Supports built-in trigger events, such as data exfiltration.

Communication compliance and AI interactions

Microsoft Purview Communication Compliance provides tools to help you detect and manage regulatory compliance and business conduct violations across various communication channels, which include user prompts and responses for AI apps. It's designed with privacy by default, pseudonymizing usernames and incorporating role-based access controls. The solution helps identify and remediate inappropriate communications, such as sharing sensitive information, harassment, threats, and adult content.

To learn more about using communication compliance policies for AI apps, see Configure a communication compliance policy to detect for generative AI interactions.

AI app-specific information:

- Supports detection of unethical communication for Microsoft Teams and emails, for agent-to-human and human-to-agent interactions.

eDiscovery and AI interactions

Microsoft Purview eDiscovery lets you identify and deliver electronic information that can be used as evidence in legal cases. The eDiscovery tools in Microsoft Purview support searching for content in Exchange Online, OneDrive for Business, SharePoint Online, Microsoft Teams,

Microsoft 365 Groups, and Viva Engage teams. You can then prevent the information from deletion and export the information.

Because user prompts and responses for AI apps are stored in a user's mailbox, you can create a case and use search when a user's mailbox is selected as the source for a search query. For example, select and retrieve this data from the source mailbox by selecting from the query builder **Add condition** > **Type** > **Contains any of** > **Edit** > *Copilot activity*. This query condition includes all Copilot and other AI application activity.

After the search is refined, you can export the results or add to a review set. You can review and export information directly from the review set.

For more information about identifying and deleting user AI interaction data, see Search for and delete Copilot data in eDiscovery.

AI app-specific information:

- Identify an agent instance as you would a user.
- Supports agent-to-human and human-to-agent interactions.

Data Lifecycle Management and AI interactions

Microsoft Purview Data Lifecycle Management provides tools and capabilities to manage the lifecycle of organizational data by retaining necessary content and deleting unnecessary content. These tools ensure compliance with business, legal, and regulatory requirements.

Use retention policies to automatically retain or delete user prompts and responses for AI apps. For detailed information about this retention works, see Learn about retention for Copilot & AI apps.

As with all retention policies and holds, if more than one policy for the same location applies to a user, the principles of retention resolve any conflicts. For example, the data is retained for the longest duration of all the applied retention policies or eDiscovery holds.

AI app-specific information:

- Supports locations that include Microsoft Teams, OneDrive and SharePoint, and emails.
- Supports agent-to-human and human-to-agent interactions.

Compliance Manager and AI interactions

Microsoft Purview Compliance Manager is a solution that helps you automatically assess and manage compliance across your multicloud environment. Compliance Manager can help you throughout your compliance journey, from taking inventory of your data protection risks to managing the complexities of implementing controls, staying current with regulations and certifications, and reporting to auditors.

To help you keep compliant with AI regulations, Compliance Manager provides regulatory templates to help you assess, implement, and strengthen your compliance requirements for all generative AI apps. For example, monitoring AI interactions and preventing data loss in AI applications. For more information, see Assessments for AI regulations.

Getting started recommended steps

To get started with managing data security & compliance for Agent 365. use the Microsoft Purview portal. You need an account that has appropriate permissions for security and compliance management, such as an account that's a member of the Entra **Compliance Administrator** role, or the **Microsoft Purview Compliance Administrator** role group.

Then navigate to Data Security Posture Management: **DSPM** > **AI observability**.

> ⓘ **Note**
>
> Don't confuse this current version with the older version, that's named **Data Security Posture Management (classic)**. The classic version doesn't support Agent 365.

From the **AI observability** page:

- Review the overview of all agents in your organization with activities in the last 30 days, prioritized with the highest risk level determined by Insider Risk Management.
- Analyze top risky activities: oversharing, exfiltration, and unethical behavior.
- Select a specific agent to see details:
  - Agent details that include Entra-enabled status, created date, owner, agent user ID, and which agent it's an instance of.
  - Agent activities that include the risk level determined by Insider Risk Management and any risky activities from agent interactions.
  - Recommendations for remediation that use Microsoft Purview solutions, based on the identified risks.

Last updated on 05/01/2026

Use Microsoft Purview to manage data security & compliance for Entra-registered AI apps

> *Microsoft Purview service description*

Use the following sections to identify the Microsoft Purview capabilities that are supported for AI interactions for custom AI apps that are registered with Entra for user authentication, and integrated with the Microsoft Purview SDK. Then refer to the get started recommendations for you to manage these AI interactions for security and compliance.

Managing these AI interactions with Microsoft Purview requires you to enable pay-as-you-go billing in your organization.

Capabilities supported

Use the following table to see at a glance the Microsoft Purview capabilities that are supported for Entra-registered AI apps.

Expand table

Capability or solution in Microsoft Purview	Supported for AI interactions
DSPM and DSPM for AI (classic)	✓
Auditing	✓
Data classification	✓
Sensitivity labels	✓
Encryption without sensitivity labels	✓
Data loss prevention	✓
Insider Risk Management	✓
Communication compliance	✓
eDiscovery	✓
Data Lifecycle Management	✓

Capability or solution in Microsoft Purview	Supported for AI interactions
Compliance Manager	✓

DSPM and DSPM for AI (classic)

Use Data Security Posture Management or Data Security Posture Management for AI (classic) as your front door to discover, secure, and apply compliance controls for AI usage across your enterprise. Both DSPM versions use existing controls from Microsoft Purview information protection and compliance management with easy-to-use graphical tools and reports to quickly gain insights into AI use within your organization. With personalized recommendations, and one-click policies help you protect your data and comply with regulatory requirements.

AI app-specific information:

- Recommendation **Get guided assistance to AI regulations**, which uses control-mapping regulatory templates from Compliance Manager.
- One-click policies available:
  - **DSPM for AI - Capture interactions for enterprise AI apps** from the recommendation or remediation action **Secure interactions from enterprise apps**.
  - **DSPM for AI - Detect sensitive info shared with AI via network** from the recommendation or remediation action **Extend insights into sensitive data in AI app interactions**.

Auditing and AI interactions

Microsoft Purview Audit solutions provide comprehensive tools for searching and managing audit records of activities performed across various Microsoft services by users and admins, and help organizations to effectively respond to security events, forensic investigations, internal investigations, and compliance obligations.

Like other activities, prompts and responses are captured in the unified audit log. Events include how and when users interact with the AI app, and can include in which Microsoft 365 service the activity took place, and references to the files stored in Microsoft 365 that were accessed during the interaction. If these files have a sensitivity label applied, that's also captured.

These events flow into activity explorer in DSPM for AI and the **AI activities** tab in activity explorer from the preview version of DSPM, where the data from prompts and responses can be

displayed. You can also use the **Audit** solution from the Microsoft Purview portal to search and find these auditing events.

For more information, see Audit logs for Copilot and AI activities.

Data classification and AI interactions

Microsoft Purview data classification provides a comprehensive framework for identifying and tagging sensitive data across various Microsoft services, including Office 365, Dynamics 365, and Azure. Classifying data is often the first step to ensure compliance with data protection regulations and safeguard against unauthorized access, alteration, or destruction. You can use built-in system classifications or create your own.

Sensitive information types and trainable classifiers can be used to find sensitive data in user prompts and responses when they use AI apps. The resulting information then surfaces in Microsoft Purview Reports overview and activity explorer in DSPM for AI and the **AI activities** tab in activity explorer from the preview version of DSPM.

Sensitivity labels and AI interactions

AI apps that Microsoft Purview support use existing controls to ensure that data stored in your tenant is never returned to the user or used by a large language model (LLM) if the user doesn't have access to that data. When the data has sensitivity labels from your organization applied to the content, there's an extra layer of protection:

- When a file is open in Word, Excel, PowerPoint, or similarly an email or calendar event is open in Outlook, the sensitivity of the data is displayed to users in the app with the label name and content markings (such as header or footer text) that have been configured for the label. Loop components and pages also support the same sensitivity labels.
- When the sensitivity label applies encryption, users must have the EXTRACT usage right, as well as VIEW, for the AI apps to return the data.
- This protection extends to data stored outside your Microsoft 365 tenant when it's open in an Office app (data in use). For example, local storage, network shares, and cloud storage.

> **Tip**
>
> If you haven't already, we recommend you enable sensitivity labels for SharePoint and OneDrive and also familiarize yourself with the file types and label configurations that these

services can process. When sensitivity labels aren't enabled for these services, the encrypted files that Copilot and agents can access are limited to data in use from Office apps on Windows.

For instructions, see **Enable sensitivity labels for Office files in SharePoint and OneDrive**.

If you're not already using sensitivity labels, see Get started with sensitivity labels.

Encryption without sensitivity labels and AI interactions

Even if a sensitivity label isn't applied to content, services and products might use the encryption capabilities from the Azure Rights Management service. As a result, AI apps can still check for the VIEW and EXTRACT usage rights before returning data and links to a user, but there's no automatic inheritance of protection for new items.

> **Tip**
>
> You'll get the best user experience when you always use sensitivity labels to protect your data, and encryption is applied by a label.

Examples of products and services that can use the encryption capabilities from the Azure Rights Management service without sensitivity labels:

- Microsoft Purview Message Encryption
- Microsoft Information Rights Management (IRM)
- Microsoft Rights Management connector
- Microsoft Rights Management SDK

For other encryption methods that don't use the Azure Rights Management service:

- S/MIME protected emails won't be returned by Copilot, and Copilot isn't available in Outlook when an S/MIME protected email is open.
- Password-protected documents can't be accessed by AI apps unless they're already opened by the user in the same app (data in use). Passwords aren't inherited by a destination item.

As with other Microsoft 365 services, such as eDiscovery and search, items encrypted with Microsoft Purview Customer Key or your own root key (BYOK) are supported and eligible to be returned by Copilot.

Data loss prevention and AI interactions

Microsoft Purview Data Loss Prevention (DLP) helps you identify sensitive items across Microsoft 365 services and endpoints, monitor them, and helps protect against leakage of those items. It uses deep content inspection and contextual analysis to identify sensitive items and it enforces policies to protect sensitive data such as financial records, health information, or intellectual property.

Windows computers that are onboarded to Microsoft Purview can be configured for Endpoint data loss prevention (DLP) policies that warn or block users from sharing sensitive information with third-party generative AI sites that are accessed via a browser. For example, a user is prevented from pasting credit card numbers into ChatGPT, or they see a warning that they can override. For more information about the supported DLP actions and which platforms support them, see the first two rows in the table from Endpoint activities you can monitor and take action on.

AI app-specific information:

- Support today is only available for a DLP policy that blocks prompts based on sensitive information types. This requires the configuration of a PowerShell cmdlet that's scoped to a specific Entra-registered AI app. The AI app can honor this configuration by integration with the Microsoft Purview APIs. The AI app developer can use this GitHub sample to help achieve this integration.

Insider Risk Management and AI interactions

Microsoft Purview Insider Risk Management helps you detect, investigate, and mitigate internal risks such as IP theft, data leakage, and security violations. It leverages machine learning models and various signals from Microsoft 365 and third-party indicators to identify potential malicious or inadvertent insider activities. The solution includes privacy controls like pseudonymization and role-based access, ensuring user-level privacy while enabling risk analysts to take appropriate actions.

Use the Risky AI usage policy template to detect risky usage that includes prompt injection attacks and accessing protected materials. Insights from these signals are integrated into Microsoft Defender XDR to provide a comprehensive view of AI-related risks.

AI app-specific information:

- For prompts and responses, requires a collection policy such as **DSPM for AI - Capture interactions for enterprise AI apps** from the DSPM for AI recommendation **Secure interactions from enterprise apps**.

Communication compliance and AI interactions

Microsoft Purview Communication Compliance provides tools to help you detect and manage regulatory compliance and business conduct violations across various communication channels, which include user prompts and responses for AI apps. It's designed with privacy by default, pseudonymizing usernames and incorporating role-based access controls. The solution helps identify and remediate inappropriate communications, such as sharing sensitive information, harassment, threats, and adult content.

To learn more about using communication compliance policies for AI apps, see Configure a communication compliance policy to detect for generative AI interactions.

AI app-specific information:

- For prompts and responses, requires a collection policy such as **DSPM for AI - Capture interactions for enterprise AI apps** from the DSPM for AI recommendation **Secure interactions from enterprise apps**.

eDiscovery and AI interactions

Microsoft Purview eDiscovery lets you identify and deliver electronic information that can be used as evidence in legal cases. The eDiscovery tools in Microsoft Purview support searching for content in Exchange Online, OneDrive for Business, SharePoint Online, Microsoft Teams, Microsoft 365 Groups, and Viva Engage teams. You can then prevent the information from deletion and export the information.

Because user prompts and responses for AI apps are stored in a user's mailbox, you can create a case and use search when a user's mailbox is selected as the source for a search query. For example, select and retrieve this data from the source mailbox by selecting from the query builder **Add condition** > **Type** > **Contains any of** > **Edit** > *Copilot activity*. This query condition includes all Copilot and other AI application activity.

After the search is refined, you can export the results or add to a review set. You can review and export information directly from the review set.

For more information about identifying and deleting user AI interaction data, see Search for and delete Copilot data in eDiscovery.

AI app-specific information:

- For prompts and responses, requires a collection policy such as **DSPM for AI - Capture interactions for enterprise AI apps** from the DSPM for AI recommendation **Secure interactions from enterprise apps**.

Data Lifecycle Management and AI interactions

Microsoft Purview Data Lifecycle Management provides tools and capabilities to manage the lifecycle of organizational data by retaining necessary content and deleting unnecessary content. These tools ensure compliance with business, legal, and regulatory requirements.

Use retention policies to automatically retain or delete user prompts and responses for AI apps. For detailed information about this retention works, see Learn about retention for Copilot & AI apps.

As with all retention policies and holds, if more than one policy for the same location applies to a user, the principles of retention resolve any conflicts. For example, the data is retained for the longest duration of all the applied retention policies or eDiscovery holds.

AI app-specific information:

- For retention policies, select the option for **Enterprise AI apps**.
- For prompts and responses, requires a collection policy such as **DSPM for AI - Capture interactions for enterprise AI apps** from the DSPM for AI recommendation **Secure interactions from enterprise apps**.

Compliance Manager and AI interactions

Microsoft Purview Compliance Manager is a solution that helps you automatically assess and manage compliance across your multicloud environment. Compliance Manager can help you throughout your compliance journey, from taking inventory of your data protection risks to managing the complexities of implementing controls, staying current with regulations and certifications, and reporting to auditors.

To help you keep compliant with AI regulations, Compliance Manager provides regulatory templates to help you assess, implement, and strengthen your compliance requirements for all

generative AI apps. For example, monitoring AI interactions and preventing data loss in AI applications. For more information, see Assessments for AI regulations.

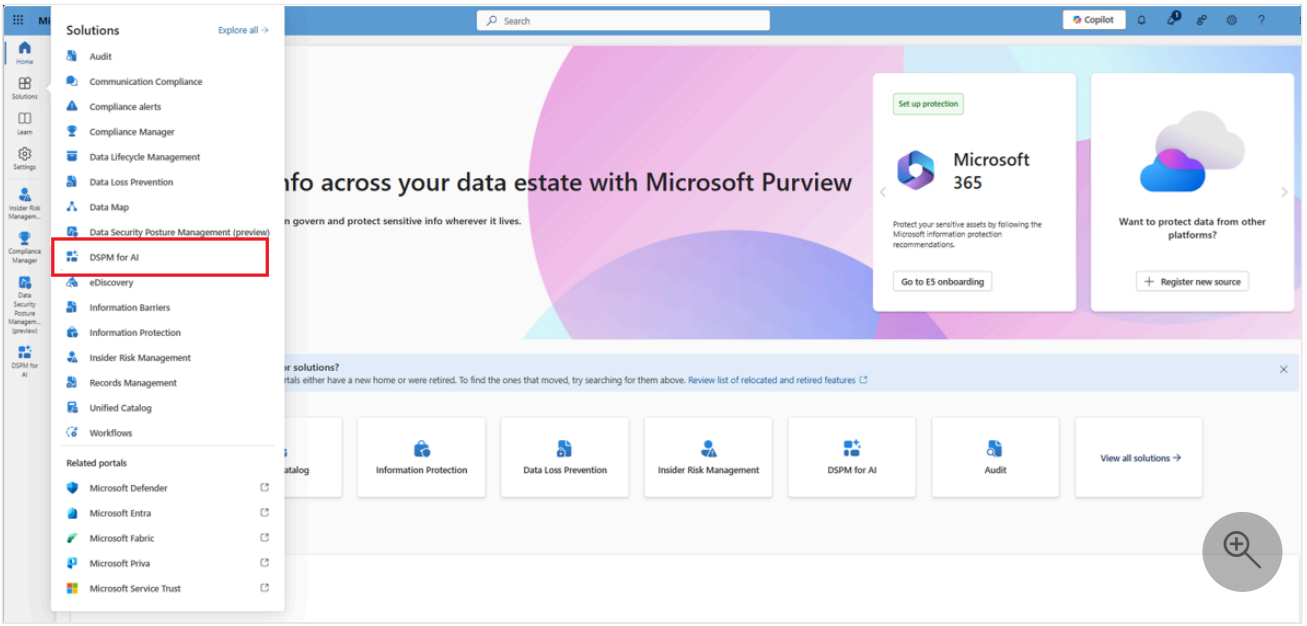
Getting started recommended steps

Use the following steps to get started with managing data security & compliance for AI interactions that use Entra-registered AI apps.

> **Note**
>
> These steps focus on managing a specific AI app or agent. For broader coverage that uses security objectives with guided workflows, use the current version of Data Security Posture Management.

Because Data Security Posture Management for AI is your front door for securing and managing AI interactions, most of the following instructions use that solution:

- Sign in to the Microsoft Purview portal > **Solutions** > **DSPM for AI (classic)** with an account that has appropriate permissions. For example, an account that's a member of the **Microsoft Entra Compliance Administrator** group role.



1. From **DSPM for AI (classic)** > **Recommendations** page, locate and select **Secure interactions from enterprise apps**, which creates a one-click policy to capture prompts and responses from enterprise AI apps.

2. Wait at least a day for data, and then navigate to the **Reports** page to view the results of your policy. Select the AI app category of **Enterprise AI apps** and view information such as:

- Total interactions over time (enterprise AI apps)
- Sensitive interactions per AI app
- Insider Risk severity
- Insider risk severity per AI app

3. Select **View details** for each of the report graphs to view detailed activities in the activity explorer.

   From the filters, select the **AI app category** of **Enterprise AI apps** with the **App** filter of **??**, and then use the other filters if you need to further refine the displayed data. Drill down to each activity to view details that include displaying the prompts and response when you're a member of the Microsoft Purview Content Explorer Content Viewer role group. For more information about this requirement, see Permissions for Data Security Posture Management for AI.

4. If you need to ensure that interactions from Entra-registered AI apps are retained for compliance reasons:

   In the Microsoft Purview portal, navigate to **Data Lifecycle Management** > **Policies** > **Retention Policies** and create a retention policy to retain interactions with Entra-registered AI apps by selecting the location **Enterprise AI apps** and specify the required retention period. For more information, see Create and configure retention policies.

5. If you need to preserve, collect, analyze, review, or export Entra-registered AI app interactions:

   In the Microsoft Purview portal, navigate to **eDiscovery** > **Cases** > **Create case**. In the case, create a search and use the *ItemClass* property and the `IPM.SkypeTeams.Message.ConnectedAIApp.Entra.<AppID>` value to search for these interactions in your organization.

Routinely review the reports in DSPM for AI to determine if you need to make changes, and use activity explorer and events for deeper analysis of how users are interacting with Entra-registered AI apps.

Last updated on 05/01/2026

Use Microsoft Purview to manage data security & compliance for Microsoft Foundry

> *Microsoft Purview service description*

Use the following sections to identify the Microsoft Purview capabilities that are supported for AI interactions that use Microsoft Foundry, and some get started recommendations for you to manage these AI interactions for security and compliance.

Before Microsoft Purview can manage these AI interactions, ensure you perform either of the following configuration steps:

1. Enable Microsoft Purview Data Security in the Microsoft Foundry Control Plane using the Microsoft Foundry portal. For more information, see Manage compliance and security in Microsoft Foundry.
2. Enable Microsoft Purview Data Security for Microsoft Foundry in Microsoft Defender for Cloud in your Microsoft Azure portal. For more information, see Enable Data Security for Azure AI with Microsoft Purview.

Managing policies for Microsoft Foundry AI interactions with Microsoft Purview requires you to enable pay-as-you-go billing in your organization. Note that Microsoft Purview Audit includes Microsoft Foundry data as part of the Microsoft Purview license.

Capabilities supported

Use the following table to see at a glance the Microsoft Purview capabilities that are supported for Microsoft Foundry.

Expand table

Capability or solution in Microsoft Purview	Supported for AI interactions
DSPM and DSPM for AI (classic)	✓
Auditing	✓
Data classification	✓
Sensitivity labels	✓

Capability or solution in Microsoft Purview	Supported for AI interactions
Encryption without sensitivity labels	✗
Data loss prevention	✓
Insider Risk Management	✓
Communication compliance	✓
eDiscovery	✓
Data Lifecycle Management	✓
Compliance Manager	✓

DSPM and DSPM for AI (classic)

Use Data Security Posture Management or Data Security Posture Management for AI (classic) as your front door to discover, secure, and apply compliance controls for AI usage across your enterprise. Both DSPM versions use existing controls from Microsoft Purview information protection and compliance management with easy-to-use graphical tools and reports to quickly gain insights into AI use within your organization. With personalized recommendations, and one-click policies help you protect your data and comply with regulatory requirements.

AI app-specific information:

- Recommendation **Secure data in Azure AI apps and agents**, which has set up instructions to capture prompts and responses for AI apps that use one or more Azure AI subscriptions.
- Recommendation **Get guided assistance to AI regulations**, which uses control-mapping regulatory templates from Compliance Manager.
- One-click policies available:
  - **DSPM for AI - Capture interactions for enterprise AI apps** from the recommendation or remediation action **Secure interactions from enterprise apps**
  - **DSPM for AI - Detect sensitive info shared with AI via network** from the recommendation or remediation action **Extend insights into sensitive data in AI app interactions**.

Auditing and AI interactions

Microsoft Purview Audit solutions provide comprehensive tools for searching and managing audit records of activities performed across various Microsoft services by users and admins, and

help organizations to effectively respond to security events, forensic investigations, internal investigations, and compliance obligations.

Like other activities, prompts and responses are captured in the unified audit log. Events include how and when users interact with the AI app, and can include in which Microsoft 365 service the activity took place, and references to the files stored in Microsoft 365 that were accessed during the interaction. If these files have a sensitivity label applied, that's also captured.

These events flow into activity explorer in DSPM for AI and the **AI activities** tab in activity explorer from the preview version of DSPM, where the data from prompts and responses can be displayed. You can also use the **Audit** solution from the Microsoft Purview portal to search and find these auditing events.

For more information, see Audit logs for Copilot and AI activities.

Data classification and AI interactions

Microsoft Purview data classification provides a comprehensive framework for identifying and tagging sensitive data across various Microsoft services, including Office 365, Dynamics 365, and Azure. Classifying data is often the first step to ensure compliance with data protection regulations and safeguard against unauthorized access, alteration, or destruction. You can use built-in system classifications or create your own.

Sensitive information types and trainable classifiers can be used to find sensitive data in user prompts and responses when they use AI apps. The resulting information then surfaces in Microsoft Purview Reports overview and activity explorer in DSPM for AI and the **AI activities** tab in activity explorer from the preview version of DSPM.

> **Note**
>
> Microsoft Purview Data Security Policies for Foundry Services interactions apply to API calls that use Microsoft Entra ID authentication with a user-context token, or for API calls that explicitly include user context. To learn more, see Gain end-user context for Azure AI API calls.
>
> For all other authentication scenarios, user interactions are displayed in Microsoft Purview Audit and AI Interactions with classifications in DSPM for AI Activity Explorer only.
>
> Ensure your AI application developers are aware of the above requirements.

Sensitivity labels and AI interactions

AI apps that Microsoft Purview support use existing controls to ensure that data stored in your tenant is never returned to the user or used by a large language model (LLM) if the user doesn't have access to that data. When the data has sensitivity labels from your organization applied to the content, there's an extra layer of protection:

- When a file is open in Word, Excel, PowerPoint, or similarly an email or calendar event is open in Outlook, the sensitivity of the data is displayed to users in the app with the label name and content markings (such as header or footer text) that have been configured for the label. Loop components and pages also support the same sensitivity labels.
- When the sensitivity label applies encryption, users must have the EXTRACT usage right, as well as VIEW, for the AI apps to return the data.
- This protection extends to data stored outside your Microsoft 365 tenant when it's open in an Office app (data in use). For example, local storage, network shares, and cloud storage.

> **Tip**
>
> If you haven't already, we recommend you enable sensitivity labels for SharePoint and OneDrive and also familiarize yourself with the file types and label configurations that these services can process. When sensitivity labels aren't enabled for these services, the encrypted files that Copilot and agents can access are limited to data in use from Office apps on Windows.
>
> For instructions, see **Enable sensitivity labels for Office files in SharePoint and OneDrive**.

If you're not already using sensitivity labels, see Get started with sensitivity labels.

AI app-specific information:

- RAG-based Foundry apps and agents built with AI search as the knowledge retrieval service can honor sensitivity labels the same way as Microsoft 365 Copilot does. For encrypted items, a user must have the EXTRACT usage right, as well as VIEW, for these items to be returned in search results. For more information, see Query-Time Microsoft Purview Sensitivity Label Enforcement in Azure AI Search.

Data loss prevention and AI interactions

Microsoft Purview Data Loss Prevention (DLP) helps you identify sensitive items across Microsoft 365 services and endpoints, monitor them, and helps protect against leakage of those items. It uses deep content inspection and contextual analysis to identify sensitive items and it enforces policies to protect sensitive data such as financial records, health information, or intellectual property.

Windows computers that are onboarded to Microsoft Purview can be configured for Endpoint data loss prevention (DLP) policies that warn or block users from sharing sensitive information with third-party generative AI sites that are accessed via a browser. For example, a user is prevented from pasting credit card numbers into ChatGPT, or they see a warning that they can override. For more information about the supported DLP actions and which platforms support them, see the first two rows in the table from Endpoint activities you can monitor and take action on.

AI app-specific information:

- Support today is only available for a DLP policy that blocks prompts based on sensitive information types. This requires the configuration of a PowerShell cmdlet that's scoped to a specific Entra-registered AI app. The AI app can honor this configuration by integration with the Microsoft Purview APIs. The AI app developer can use this GitHub sample to help achieve this integration.

Insider Risk Management and AI interactions

Microsoft Purview Insider Risk Management helps you detect, investigate, and mitigate internal risks such as IP theft, data leakage, and security violations. It leverages machine learning models and various signals from Microsoft 365 and third-party indicators to identify potential malicious or inadvertent insider activities. The solution includes privacy controls like pseudonymization and role-based access, ensuring user-level privacy while enabling risk analysts to take appropriate actions.

Use the Risky AI usage policy template to detect risky usage that includes prompt injection attacks and accessing protected materials. Insights from these signals are integrated into Microsoft Defender XDR to provide a comprehensive view of AI-related risks.

AI app-specific information:

- For prompts and responses, requires a collection policy such as **DSPM for AI - Detect sensitive info shared with AI via network** from the DSPM for AI recommendation **Extend insights into sensitive data in AI app interactions**.

Communication compliance and AI interactions

Microsoft Purview Communication Compliance provides tools to help you detect and manage regulatory compliance and business conduct violations across various communication channels, which include user prompts and responses for AI apps. It's designed with privacy by default, pseudonymizing usernames and incorporating role-based access controls. The solution helps identify and remediate inappropriate communications, such as sharing sensitive information, harassment, threats, and adult content.

To learn more about using communication compliance policies for AI apps, see Configure a communication compliance policy to detect for generative AI interactions.

AI app-specific information:

- For prompts and responses, requires a collection policy such as **DSPM for AI - Detect sensitive info shared with AI via network** from the DSPM for AI recommendation **Extend insights into sensitive data in AI app interactions**.

eDiscovery and AI interactions

Microsoft Purview eDiscovery lets you identify and deliver electronic information that can be used as evidence in legal cases. The eDiscovery tools in Microsoft Purview support searching for content in Exchange Online, OneDrive for Business, SharePoint Online, Microsoft Teams, Microsoft 365 Groups, and Viva Engage teams. You can then prevent the information from deletion and export the information.

Because user prompts and responses for AI apps are stored in a user's mailbox, you can create a case and use search when a user's mailbox is selected as the source for a search query. For example, select and retrieve this data from the source mailbox by selecting from the query builder **Add condition** > **Type** > **Contains any of** > **Edit** > *Copilot activity*. This query condition includes all Copilot and other AI application activity.

After the search is refined, you can export the results or add to a review set. You can review and export information directly from the review set.

For more information about identifying and deleting user AI interaction data, see Search for and delete Copilot data in eDiscovery.

AI app-specific information:

- For prompts and responses, requires a collection policy such as **DSPM for AI - Detect sensitive info shared with AI via network** from the DSPM for AI recommendation **Extend insights into sensitive data in AI app interactions**.

Data Lifecycle Management and AI interactions

Microsoft Purview Data Lifecycle Management provides tools and capabilities to manage the lifecycle of organizational data by retaining necessary content and deleting unnecessary content. These tools ensure compliance with business, legal, and regulatory requirements.

Use retention policies to automatically retain or delete user prompts and responses for AI apps. For detailed information about this retention works, see Learn about retention for Copilot & AI apps.

As with all retention policies and holds, if more than one policy for the same location applies to a user, the principles of retention resolve any conflicts. For example, the data is retained for the longest duration of all the applied retention policies or eDiscovery holds.

AI app-specific information:

- For retention policies, select the option for **Enterprise AI apps**.
- For prompts and responses, requires a collection policy such as **DSPM for AI - Detect sensitive info shared with AI via network** from the DSPM for AI recommendation **Extend insights into sensitive data in AI app interactions**.

Compliance Manager and AI interactions

Microsoft Purview Compliance Manager is a solution that helps you automatically assess and manage compliance across your multicloud environment. Compliance Manager can help you throughout your compliance journey, from taking inventory of your data protection risks to managing the complexities of implementing controls, staying current with regulations and certifications, and reporting to auditors.

To help you keep compliant with AI regulations, Compliance Manager provides regulatory templates to help you assess, implement, and strengthen your compliance requirements for all generative AI apps. For example, monitoring AI interactions and preventing data loss in AI applications. For more information, see Assessments for AI regulations.

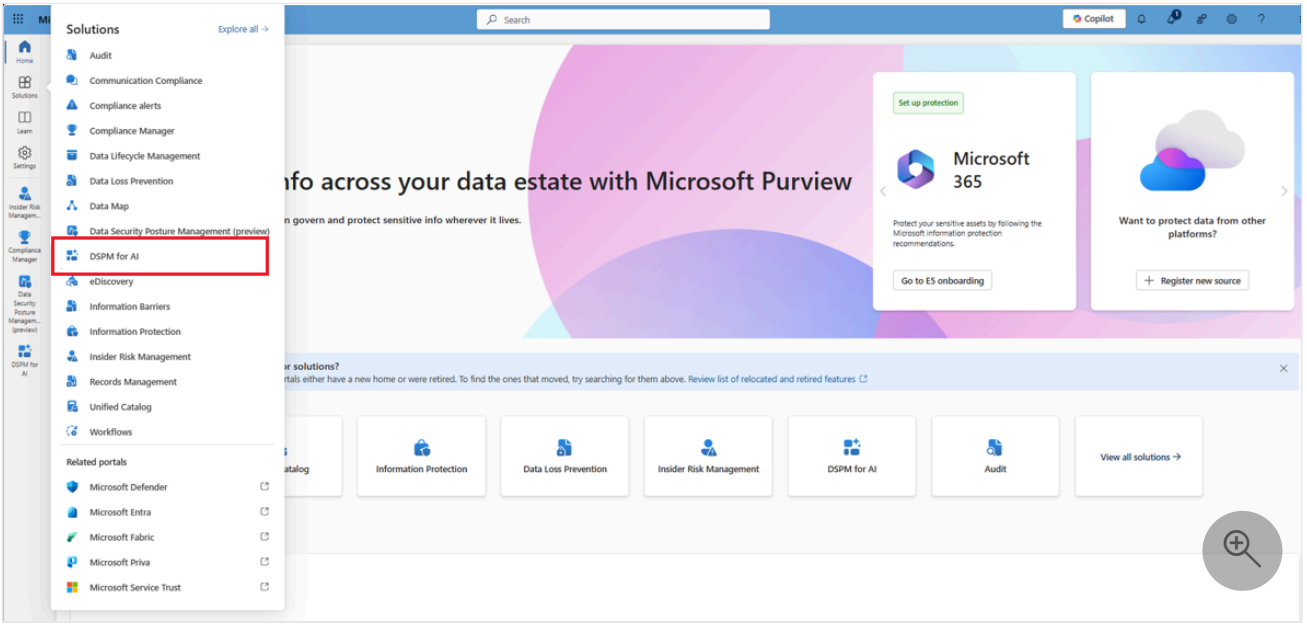
Getting started recommended steps

Use the following steps to get started with managing data security & compliance for AI interactions that use Microsoft Foundry.

> **Note**
>
> These steps focus on managing a specific AI app or agent. For broader coverage that uses security objectives with guided workflows, use the current version of Data Security Posture Management.

Because Data Security Posture Management for AI is your front door for securing and managing AI interactions, most of the following instructions use that solution:

- Sign in to the Microsoft Purview portal > **Solutions** > **DSPM for AI (classic)** with an account that has appropriate permissions. For example, an account that's a member of the **Microsoft Entra Compliance Administrator** group role.



1. From **DSPM for AI (classic)** > **Recommendations** page, locate and select **Secure data in Azure AI apps and agents**, which has set up instructions to capture prompts and responses for AI apps that use one or more Azure AI subscriptions.

2. Also from the recommendations list, select **Secure interactions from enterprise apps**, which creates a one-click policy to capture prompts and responses from enterprise AI apps.

3. Wait at least a day for data, and then navigate to the **Reports** page to view the results of your policy. Select **Enterprise AI apps** and view information such as:

    - Total interactions over time (enterprise AI apps)

- Sensitive interactions per AI app
- Insider Risk severity
- Insider risk severity per AI app

4. Select **View details** for each of the report graphs to view detailed activities in the activity explorer.

   From the filters, select the **AI app category** of **Enterprise AI apps** with the **App** filter of **Azure AI**, and then use the other filters if you need to further refine the displayed data. Drill down to each activity to view details that include displaying the prompts and response when you're a member of the Microsoft Purview Content Explorer Content Viewer role group. For more information about this requirement, see Permissions for Data Security Posture Management for AI.

5. If you need to ensure that interactions from Microsoft Foundry are retained for compliance reasons:

   In the Microsoft Purview portal, navigate to **Data Lifecycle Management** > **Policies** > **Retention Policies** and create a retention policy to retain interactions with Microsoft Foundry by selecting the location **Enterprise AI apps** and specify the required retention period. For more information, see Create and configure retention policies.

6. If you need to preserve, collect, analyze, review, or export Azure AI service interactions:

   In the Microsoft Purview portal, navigate to **eDiscovery** > **Cases** > **Create case**. In the case, create a search and use the *ItemClass* property and the `IPM.SkypeTeams.Message.ConnectedAIApp.AzureAI.<AzureResourceName>` value to search for these interactions in your organization.

Routinely review the reports in DSPM for AI to determine if you need to make changes, and use activity explorer and events for deeper analysis of how users are interacting with Microsoft Foundry.

Last updated on 05/01/2026

Use Microsoft Purview to manage data security & compliance for ChatGPT Enterprise

> *Microsoft Purview service description*

Use the following sections to identify the Microsoft Purview capabilities that are supported for AI interactions with ChatGPT Enterprise, and some get started recommendations for you to manage these AI interactions for security and compliance.

For setup information and prerequisites, see Connect to and manage ChatGPT Enterprise AI interactions in Microsoft Purview. You must run the connector scan before Microsoft Purview can manage the AI interactions for ChatGPT Enterprise.

Managing these AI interactions with Microsoft Purview requires you to enable pay-as-you-go billing in your organization.

Capabilities supported

Use the following table to see at a glance the Microsoft Purview capabilities that are supported for ChatGPT Enterprise.

Expand table

Capability or solution in Microsoft Purview	Supported for AI interactions
DSPM and DSPM for AI (classic)	✓
Auditing	✓
Data classification	✓
Sensitivity labels	✗
Encryption without sensitivity labels	✗
Data loss prevention	✗
Insider Risk Management	✓
Communication compliance	✓

Capability or solution in Microsoft Purview	Supported for AI interactions
eDiscovery	✓
Data Lifecycle Management	✓
Compliance Manager	✓

> **Note**
>
> These features only apply to the enterprise version of ChatGPT. For more information on the available Purview features for the consumer version, see Learn about Microsoft Purview Network Data Security and Learn about Data Loss Prevention for Cloud Apps in Edge for Business.

DSPM and DSPM for AI (classic)

Use Data Security Posture Management or Data Security Posture Management for AI (classic) as your front door to discover, secure, and apply compliance controls for AI usage across your enterprise. Both DSPM versions use existing controls from Microsoft Purview information protection and compliance management with easy-to-use graphical tools and reports to quickly gain insights into AI use within your organization. With personalized recommendations, and one-click policies help you protect your data and comply with regulatory requirements.

AI app-specific information:

- Recommendation **Get guided assistance to AI regulations**, which uses control-mapping regulatory templates from Compliance Manager.
- One-click policies available:
  - **DSPM for AI - Capture interactions for enterprise AI apps** from the recommendation or remediation action **Secure interactions from enterprise apps**.
  - **DSPM for AI - Detect sensitive info shared with AI via network** from the recommendation or remediation action **Extend insights into sensitive data in AI app interactions**.

Auditing and AI interactions

Microsoft Purview Audit solutions provide comprehensive tools for searching and managing audit records of activities performed across various Microsoft services by users and admins, and

help organizations to effectively respond to security events, forensic investigations, internal investigations, and compliance obligations.

Like other activities, prompts and responses are captured in the unified audit log. Events include how and when users interact with the AI app, and can include in which Microsoft 365 service the activity took place, and references to the files stored in Microsoft 365 that were accessed during the interaction. If these files have a sensitivity label applied, that's also captured.

These events flow into activity explorer in the **AI activities** tab in the current version of DSPM and in DSPM for AI, where the data from prompts and responses can be displayed. You can also use the **Audit** solution from the Microsoft Purview portal to search and find these auditing events.

For more information, see Audit logs for Copilot and AI activities.

Data classification and AI interactions

Microsoft Purview data classification provides a comprehensive framework for identifying and tagging sensitive data across various Microsoft services, including Office 365, Dynamics 365, and Azure. Classifying data is often the first step to ensure compliance with data protection regulations and safeguard against unauthorized access, alteration, or destruction. You can use built-in system classifications or create your own.

Sensitive information types and trainable classifiers can be used to find sensitive data in user prompts and responses when they use AI apps. The resulting information then surfaces in Microsoft Purview Reports overview and activity explorer in DSPM for AI and the **AI activities** tab in activity explorer from the preview version of DSPM.

Insider Risk Management and AI interactions

Microsoft Purview Insider Risk Management helps you detect, investigate, and mitigate internal risks such as IP theft, data leakage, and security violations. It leverages machine learning models and various signals from Microsoft 365 and third-party indicators to identify potential malicious or inadvertent insider activities. The solution includes privacy controls like pseudonymization and role-based access, ensuring user-level privacy while enabling risk analysts to take appropriate actions.

Use the Risky AI usage policy template to detect risky usage that includes prompt injection attacks and accessing protected materials. Insights from these signals are integrated into Microsoft Defender XDR to provide a comprehensive view of AI-related risks.

AI app-specific information:

- For prompts and responses, requires a collection policy such as **DSPM for AI - Capture interactions for enterprise AI apps** from the DSPM for AI recommendation **Secure interactions from enterprise apps**.

Communication compliance and AI interactions

Microsoft Purview Communication Compliance provides tools to help you detect and manage regulatory compliance and business conduct violations across various communication channels, which include user prompts and responses for AI apps. It's designed with privacy by default, pseudonymizing usernames and incorporating role-based access controls. The solution helps identify and remediate inappropriate communications, such as sharing sensitive information, harassment, threats, and adult content.

To learn more about using communication compliance policies for AI apps, see Configure a communication compliance policy to detect for generative AI interactions.

AI app-specific information:

- For prompts and responses, requires a collection policy such as **DSPM for AI - Capture interactions for enterprise AI apps** from the DSPM for AI recommendation **Secure interactions from enterprise apps**.

eDiscovery and AI interactions

Microsoft Purview eDiscovery lets you identify and deliver electronic information that can be used as evidence in legal cases. The eDiscovery tools in Microsoft Purview support searching for content in Exchange Online, OneDrive for Business, SharePoint Online, Microsoft Teams, Microsoft 365 Groups, and Viva Engage teams. You can then prevent the information from deletion and export the information.

Because user prompts and responses for AI apps are stored in a user's mailbox, you can create a case and use search when a user's mailbox is selected as the source for a search query. For example, select and retrieve this data from the source mailbox by selecting from the query builder **Add condition** > **Type** > **Contains any of** > **Edit** > *Copilot activity*. This query condition includes all Copilot and other AI application activity.

After the search is refined, you can export the results or add to a review set. You can review and export information directly from the review set.

For more information about identifying and deleting user AI interaction data, see Search for and delete Copilot data in eDiscovery.

AI app-specific information:

- For prompts and responses, requires a collection policy such as **DSPM for AI - Capture interactions for enterprise AI apps** from the DSPM for AI recommendation **Secure interactions from enterprise apps**.

Data Lifecycle Management and AI interactions

Microsoft Purview Data Lifecycle Management provides tools and capabilities to manage the lifecycle of organizational data by retaining necessary content and deleting unnecessary content. These tools ensure compliance with business, legal, and regulatory requirements.

Use retention policies to automatically retain or delete user prompts and responses for AI apps. For detailed information about this retention works, see Learn about retention for Copilot & AI apps.

As with all retention policies and holds, if more than one policy for the same location applies to a user, the principles of retention resolve any conflicts. For example, the data is retained for the longest duration of all the applied retention policies or eDiscovery holds.

AI app-specific information:

- For retention policies, select the option for **Enterprise AI apps**.
- For prompts and responses, requires a collection policy such as **DSPM for AI - Capture interactions for enterprise AI apps** from the DSPM for AI recommendation **Secure interactions from enterprise apps**.

Compliance Manager and AI interactions

Microsoft Purview Compliance Manager is a solution that helps you automatically assess and manage compliance across your multicloud environment. Compliance Manager can help you throughout your compliance journey, from taking inventory of your data protection risks to managing the complexities of implementing controls, staying current with regulations and certifications, and reporting to auditors.

To help you keep compliant with AI regulations, Compliance Manager provides regulatory templates to help you assess, implement, and strengthen your compliance requirements for all

generative AI apps. For example, monitoring AI interactions and preventing data loss in AI applications. For more information, see Assessments for AI regulations.

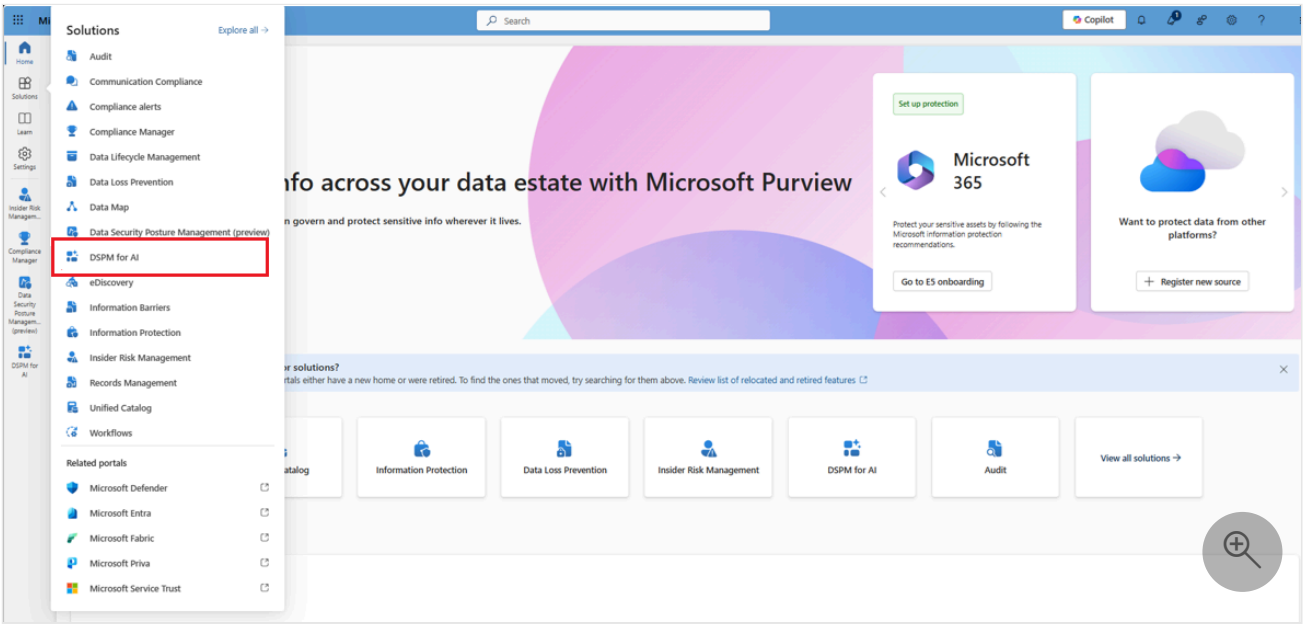
Getting started recommended steps

Use the following steps to get started with managing data security & compliance for AI interactions that use ChatGPT Enterprise.

> **Note**
>
> These steps focus on managing a specific AI app or agent. For broader coverage that uses security objectives with guided workflows, use the current version of Data Security Posture Management.

Because Data Security Posture Management for AI is your front door for securing and managing AI interactions, most of the following instructions use that solution:

- Sign in to the Microsoft Purview portal > **Solutions** > **DSPM for AI (classic)** with an account that has appropriate permissions. For example, an account that's a member of the **Microsoft Entra Compliance Administrator** group role.



1. From **DSPM for AI (classic)** > **Recommendations** page, locate and select **Discover and govern interactions with ChatGPT Enterprise AI**, which has set up instructions to capture prompts and responses for AI apps that use one or more Azure AI subscriptions.

2. Also from the recommendations list, select **Secure interactions from enterprise apps**, which creates a one-click policy to capture prompts and responses from enterprise AI apps.

3. Wait at least a day for data, and then navigate to the **Reports** page to view the results of your policy. Select the AI app category of **Enterprise AI apps** and view information such as:

    - Total interactions over time (enterprise AI apps)
    - Sensitive interactions per AI app
    - Insider Risk severity
    - Insider risk severity per AI app

4. Select **View details** for each of the report graphs to view detailed activities in the activity explorer.

    From the filters, select the **AI app category** of **Enterprise AI apps** with the **App** filter of **ChatGPT Enterprise**, and then use the other filters if you need to further refine the displayed data. Drill down to each activity to view details that include displaying the prompts and response when you're a member of the Microsoft Purview Content Explorer Content Viewer role group. For more information about this requirement, see Permissions for Data Security Posture Management for AI.

5. If you need to ensure that interactions from ChatGPT Enterprise are retained for compliance reasons:

    In the Microsoft Purview portal, navigate to **Data Lifecycle Management** > **Policies** > **Retention Policies** and create a retention policy to retain interactions with ChatGPT Enterprise by selecting the location **Enterprise AI apps** and specify the required retention period. For more information, see Create and configure retention policies.

6. If you need to preserve, collect, analyze, review, or export ChatGPT Enterprise interactions:

    In the Microsoft Purview portal, navigate to **eDiscovery** > **Cases** > **Create case**. In the case, create a search and use the *ItemClass* property and the `IPM.SkypeTeams.Message.ConnectedAIApp.Connector.<ChatGPTEnterprise>` value for local machine activities or the `IPM.SkypeTeams.Message.CloudAIApp.SaaS.<AppID>` value for browser-based activities to search for these interactions in your organization.

Routinely review the reports in DSPM for AI to determine if you need to make changes, and use activity explorer and events for deeper analysis of how users are interacting with ChatGPT Enterprise.

Last updated on 07/29/2026

Use Microsoft Purview to manage data security & compliance for Anthropic Claude (Enterprise)

> *Microsoft Purview service description*

Use the following sections to identify the Microsoft Purview data security and compliance protections for generative AI apps with Anthropic Claude, and recommendations to get started to help you manage these AI interactions for security and compliance.

For setup information and prerequisites, see Set up a connector to manage Anthropic Claude data. You must set up and configure the connector before Microsoft Purview can manage the AI interactions for Anthropic Claude.

> **Important**
>
> Managing these AI interactions with Microsoft Purview requires you to enable pay-as-you-go billing in your organization, which is subject to the Microsoft Purview billing model. For more information, see Consent to use pay-as-you-go capabilities and Learn about Microsoft Purview billing models.

> **Note**
>
> The Anthropic Claude connector is currently in preview.

Capabilities supported

The following table lists the Microsoft Purview capabilities supported for Anthropic Claude.

Expand table

Capability or solution in Microsoft Purview	Supported for AI interactions
DSPM and DSPM for AI (classic)	✓
Auditing	✓

Capability or solution in Microsoft Purview	Supported for AI interactions
Data classification	✓
Sensitivity labels	✗
Encryption without sensitivity labels	✗
Data loss prevention	✗
Insider Risk Management	✓
Communication compliance	✓
eDiscovery	✓
Data Lifecycle Management	✓
Compliance Manager	✓

DSPM and DSPM for AI (classic)

Use Data Security Posture Management or Data Security Posture Management for AI (classic) as your front door to discover, secure, and apply compliance controls for AI usage across your enterprise. Both DSPM versions use existing controls from Microsoft Purview information protection and compliance management with easy-to-use graphical tools and reports to quickly gain insights into AI use within your organization. With personalized recommendations, and one-click policies help you protect your data and comply with regulatory requirements.

AI app-specific information:

- Currently, there are no recommendations or one-click policies applicable to Anthropic Claude, but you can view and monitor the AI interactions by using the **Reports** page, **activity explorer** (**AI activities** tab), and the **Apps and agents** dashboard.
- Because agents for Anthropic Claude aren't currently supported, they aren't displayed in **AI observability**.

Auditing and AI interactions

Microsoft Purview Audit solutions provide comprehensive tools for searching and managing audit records of activities performed across various Microsoft services by users and admins, and help organizations to effectively respond to security events, forensic investigations, internal investigations, and compliance obligations.

Like other activities, prompts and responses are captured in the unified audit log. Events include how and when users interact with the AI app, and can include in which Microsoft 365 service the activity took place, and references to the files stored in Microsoft 365 that were accessed during the interaction. If these files have a sensitivity label applied, that's also captured.

These events flow into activity explorer in the **AI activities** tab in the current version of DSPM and in DSPM for AI, where the data from prompts and responses can be displayed. You can also use the **Audit** solution from the Microsoft Purview portal to search and find these auditing events.

For more information, see Audit logs for Copilot and AI activities.

Recommended steps to get started

Use the following steps to get started with managing data security & compliance for AI interactions that use Anthropic Claude.

> **Note**
>
> These steps focus on managing a specific AI app or agent. For broader coverage that uses security objectives with guided workflows, use the current version of Data Security Posture Management.

Because Data Security Posture Management for AI is your front door for securing and managing AI interactions, most of the following instructions use that solution:

1. Sign in to the Microsoft Purview portal > **Solutions** > **DSPM for AI (classic)** with an account that has appropriate permissions. For example, an account that's a member of the **Microsoft Entra Compliance Administrator** group role. For more information, see Permissions for Data Security Posture Management for AI - (classic).

2. If you haven't already set up and configured the connector to manage AI interactions for Anthropic Claude, see Set up a connector to manage Anthropic Claude data.

3. After setting up the Anthropic Claude connector, wait at least a day for data to populate. Go to the **Reports** page to view information such as:

    - Total interactions over time.
    - Interaction trends by AI app.

4. Select **View details** for each of the report graphs to view detailed activities in activity explorer.

   From the filters, select the **AI app category** of **Enterprise AI apps** with the **App** filter of **Anthropic Claude**, and then use the other filters if you need to further refine the displayed data.

Regularly review the reports in DSPM for AI to determine whether you need to make changes, and use activity explorer for deeper analysis of how users are interacting with Anthropic Claude.

Last updated on 08/04/2026

Use Microsoft Purview to manage data security & compliance for other AI apps

> *Microsoft Purview service description*

Use the following sections to identify the Microsoft Purview capabilities that are supported for AI interactions that use a browser to third-party AI sites, and some get started recommendations for you to manage these AI interactions for security and compliance.

Managing these AI interactions with Microsoft Purview requires you to enable pay-as-you-go billing in your organization.

Capabilities supported

Use the following table to see at a glance the Microsoft Purview capabilities that are supported for third-party AI sites, such as ChatGPT, Google Gemini, and DeepSeek.

Unless noted for specific solutions, the list of supported AI sites is the same as those supported for DSPM for AI: List of AI sites supported by Microsoft Purview Data Security Posture Management for AI.

Expand table

Capability or solution in Microsoft Purview	Supported for AI interactions
DSPM and DSPM for AI (classic)	✓
Auditing	✓
Data classification	✓
Sensitivity labels	✗
Encryption without sensitivity labels	✗
Data loss prevention	✓
Insider Risk Management	✓
Communication compliance	✓
eDiscovery	✓

Capability or solution in Microsoft Purview	Supported for AI interactions
Data Lifecycle Management	✓
Compliance Manager	✓

Most of the supported capabilities and solutions require the Microsoft Purview browser extension and devices onboarded to Microsoft Purview. The exceptions are, with their own prerequisites:

- Network data security to detect sensitive information types at the network layer, using a Secure Access Service Edge or Security Service Edge integration.
- Browser data security to detect sensitive information types at the Edge browser layer, using a direct integration from Purview into the Edge browser.

DSPM and DSPM for AI (classic)

Use Data Security Posture Management or Data Security Posture Management for AI (classic) as your front door to discover, secure, and apply compliance controls for AI usage across your enterprise. Both DSPM versions use existing controls from Microsoft Purview information protection and compliance management with easy-to-use graphical tools and reports to quickly gain insights into AI use within your organization. With personalized recommendations, and one-click policies help you protect your data and comply with regulatory requirements.

AI app-specific information:

- The **Getting started** section from the **Overview** page contains the required steps and information to install the Microsoft Purview browser extension and devices onboarding process.

- Recommendation **Get guided assistance to AI regulations**, which uses control-mapping regulatory templates from Compliance Manager.

- One-click policies available:
  - **DSPM for AI: Detect sensitive info added to AI sites** from the recommendation or remediation action **Extend your insights for data discovery**
  - **DSPM for AI - Detect when users visit AI sites** from the recommendation or remediation action **Extend your insights for data discovery**
  - **DSPM for AI - Detect sensitive info shared in AI prompts in Edge** from the recommendation or remediation action **Extend your insights for data discovery**
  - **DSPM for AI - Detect sensitive info shared with AI via network** from the recommendation or remediation action **Extend insights into sensitive data in AI app**

interactions.

Auditing and AI interactions

Microsoft Purview Audit solutions provide comprehensive tools for searching and managing audit records of activities performed across various Microsoft services by users and admins, and help organizations to effectively respond to security events, forensic investigations, internal investigations, and compliance obligations.

Like other activities, prompts and responses are captured in the unified audit log. Events include how and when users interact with the AI app, and can include in which Microsoft 365 service the activity took place, and references to the files stored in Microsoft 365 that were accessed during the interaction. If these files have a sensitivity label applied, that's also captured.

These events flow into activity explorer in DSPM for AI and the **AI activities** tab in activity explorer from the preview version of DSPM, where the data from prompts and responses can be displayed. You can also use the **Audit** solution from the Microsoft Purview portal to search and find these auditing events.

For more information, see Audit logs for Copilot and AI activities.

AI app-specific information:

- Use network data security to audit prompts and responses at the network layer, using a Secure Access Service Edge or Security Service Edge integration.

Data classification and AI interactions

Microsoft Purview data classification provides a comprehensive framework for identifying and tagging sensitive data across various Microsoft services, including Office 365, Dynamics 365, and Azure. Classifying data is often the first step to ensure compliance with data protection regulations and safeguard against unauthorized access, alteration, or destruction. You can use built-in system classifications or create your own.

Sensitive information types and trainable classifiers can be used to find sensitive data in user prompts and responses when they use AI apps. The resulting information then surfaces in Microsoft Purview Reports overview and activity explorer in DSPM for AI and the **AI activities** tab in activity explorer from the preview version of DSPM.

AI app-specific information:

- Use network data security to detect sensitive info types at the network layer, using a Secure Access Service Edge or Security Service Edge integration.
- Use browser data security to detect sensitive info types at the Edge browser layer, using a direct integration from Purview into the Edge browser.

Data loss prevention and AI interactions

Microsoft Purview Data Loss Prevention (DLP) helps you identify sensitive items across Microsoft 365 services and endpoints, monitor them, and helps protect against leakage of those items. It uses deep content inspection and contextual analysis to identify sensitive items and it enforces policies to protect sensitive data such as financial records, health information, or intellectual property.

Windows computers that are onboarded to Microsoft Purview can be configured for Endpoint data loss prevention (DLP) policies that warn or block users from sharing sensitive information with third-party generative AI sites that are accessed via a browser. For example, a user is prevented from pasting credit card numbers into ChatGPT, or they see a warning that they can override. For more information about the supported DLP actions and which platforms support them, see the first two rows in the table from Endpoint activities you can monitor and take action on.

AI app-specific information:

- Endpoint DLP supports an option to block the upload of files based on a specified sensitivity label.

Insider Risk Management and AI interactions

Microsoft Purview Insider Risk Management helps you detect, investigate, and mitigate internal risks such as IP theft, data leakage, and security violations. It leverages machine learning models and various signals from Microsoft 365 and third-party indicators to identify potential malicious or inadvertent insider activities. The solution includes privacy controls like pseudonymization and role-based access, ensuring user-level privacy while enabling risk analysts to take appropriate actions.

Use the Risky AI usage policy template to detect risky usage that includes prompt injection attacks and accessing protected materials. Insights from these signals are integrated into Microsoft Defender XDR to provide a comprehensive view of AI-related risks.

AI app-specific information:

- For prompts and responses, requires a collection policy that includes the setting to capture content. For more information, see Collection Policies solution overview.

Communication compliance and AI interactions

Microsoft Purview Communication Compliance provides tools to help you detect and manage regulatory compliance and business conduct violations across various communication channels, which include user prompts and responses for AI apps. It's designed with privacy by default, pseudonymizing usernames and incorporating role-based access controls. The solution helps identify and remediate inappropriate communications, such as sharing sensitive information, harassment, threats, and adult content.

To learn more about using communication compliance policies for AI apps, see Configure a communication compliance policy to detect for generative AI interactions.

AI app-specific information:

- Support restricted to the Edge browser, for ChatGPT, Microsoft Chat (consumer version), Google Gemini, and DeepSeek.
- For prompts and responses, requires a collection policy that includes the setting to capture content. For more information, see Collection Policies solution overview.
- Use network data security to analyze prompts and responses at the network layer, using a Secure Access Service Edge or Security Service Edge integration.

eDiscovery and AI interactions

Microsoft Purview eDiscovery lets you identify and deliver electronic information that can be used as evidence in legal cases. The eDiscovery tools in Microsoft Purview support searching for content in Exchange Online, OneDrive for Business, SharePoint Online, Microsoft Teams, Microsoft 365 Groups, and Viva Engage teams. You can then prevent the information from deletion and export the information.

Because user prompts and responses for AI apps are stored in a user's mailbox, you can create a case and use search when a user's mailbox is selected as the source for a search query. For example, select and retrieve this data from the source mailbox by selecting from the query builder **Add condition** > **Type** > **Contains any of** > **Edit** > *Copilot activity*. This query condition includes all Copilot and other AI application activity.

After the search is refined, you can export the results or add to a review set. You can review and export information directly from the review set.

For more information about identifying and deleting user AI interaction data, see Search for and delete Copilot data in eDiscovery.

AI app-specific information:

- Support restricted to the Edge browser, for ChatGPT, Microsoft Chat (consumer version), Google Gemini, and DeepSeek.
- For prompts and responses, requires a collection policy that includes the setting to capture content. For more information, see Collection Policies solution overview.
- Use network data security to analyze prompts and responses at the network layer.

Data Lifecycle Management and AI interactions

Microsoft Purview Data Lifecycle Management provides tools and capabilities to manage the lifecycle of organizational data by retaining necessary content and deleting unnecessary content. These tools ensure compliance with business, legal, and regulatory requirements.

Use retention policies to automatically retain or delete user prompts and responses for AI apps. For detailed information about this retention works, see Learn about retention for Copilot & AI apps.

As with all retention policies and holds, if more than one policy for the same location applies to a user, the principles of retention resolve any conflicts. For example, the data is retained for the longest duration of all the applied retention policies or eDiscovery holds.

AI app-specific information:

- Support restricted to the Edge browser, for ChatGPT, Microsoft Chat (consumer version), Google Gemini, and DeepSeek.
- For retention policies, select the option for **Other AI apps**.
- For prompts and responses, requires a collection policy that includes the setting to capture content. For more information, see Collection Policies solution overview.
- Use network data security to retain prompts and responses at the network layer.
- Use browser data security to retain prompts and responses at the Edge browser layer.

Compliance Manager and AI interactions

Microsoft Purview Compliance Manager is a solution that helps you automatically assess and manage compliance across your multicloud environment. Compliance Manager can help you throughout your compliance journey, from taking inventory of your data protection risks to managing the complexities of implementing controls, staying current with regulations and certifications, and reporting to auditors.

To help you keep compliant with AI regulations, Compliance Manager provides regulatory templates to help you assess, implement, and strengthen your compliance requirements for all generative AI apps. For example, monitoring AI interactions and preventing data loss in AI applications. For more information, see Assessments for AI regulations.

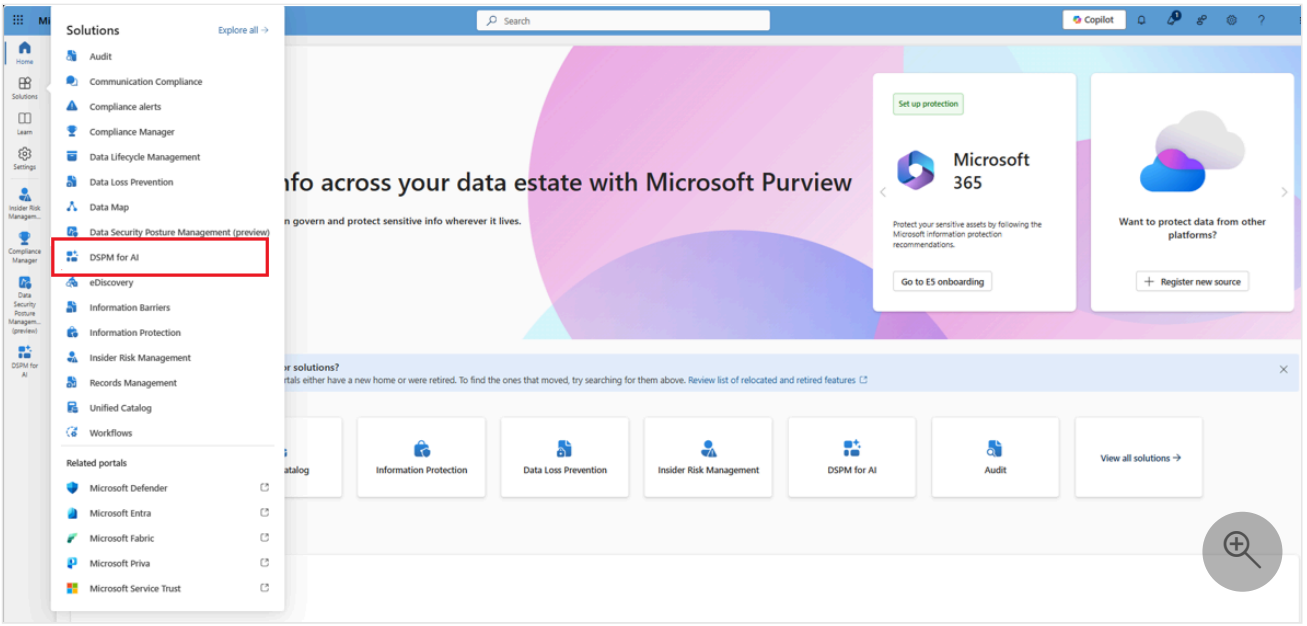
Getting started recommended steps

Use the following steps to get started with managing data security & compliance for AI interactions that use a browser to third-party AI sites.

> **Note**
>
> These steps focus on managing a specific AI app or agent. For broader coverage that uses security objectives with guided workflows, use the current version of Data Security Posture Management.

Because Data Security Posture Management for AI is your front door for securing and managing AI interactions, the following instructions use that solution:

- Sign in to the Microsoft Purview portal > **Solutions** > **DSPM for AI (classic)** with an account that has appropriate permissions. For example, an account that's a member of the **Microsoft Entra Compliance Administrator** group role.



Discover potential security risks in interactions with other AI apps

1. From **DSPM for AI (classic)** > **Overview**, in the **Get Started** section, locate and take action on the following required steps:

   - **Install Microsoft Purview browser extension**
   - **Onboard devices to Microsoft Purview**
   - **Extend your insights for data discovery**

   If you need more information about the browser extension, or other onboarding requirements, see Prerequisites for Data Security Posture Management for AI.

   For more information about the one-click policies that are automatically created with the option to extend your insights for data discovery, see Default policies for data discovery using Data Security Posture Management for AI and the policies that have the source of **Extend your insights for data discovery**.

2. Wait at least a day for data, and then navigate to the **Reports** page to view the results of your policy. Select **Other AI apps** and view information such as:

   - Total interactions over time (other AI apps)
   - Total visits (other AI apps)
   - Sensitive interactions per AI app
   - Insider Risk severity
   - Insider risk severity per AI apps

3. Select **View details** for each of the report graphs to view detailed activities in the activity explorer.

   Select the available filters to see the results from **Copilot experiences & agents** based on the different **Activity type**, **AI app category**, **App**, and **Scope** for administrative units, and more. Then drill down to each activity to view details that include displaying the prompts and response when you're a member of the Microsoft Purview Content Explorer Content Viewer role group. For more information about this requirement, see Permissions for Data Security Posture Management for AI.

Protect sensitive data in interactions with other AI apps

1. From **DSPM for AI (classic)** > **Recommendations** page, under **Not Started**, select **Fortify your data security** to create the following one-click policies:

   - **Block elevated risk users from pasting or uploading sensitive info on AI sites**: This recommendation creates a policy that uses Adaptive Protection to give a block-with-override to elevated risky users attempting to paste or upload sensitive information to other AI apps in Edge, Chrome, and Firefox. This policy covers all users and groups in your org in test mode.

   - **Block elevated risk users from submitting prompts to AI apps in Microsoft Edge**: This recommendation guides you through the process of creating a Purview Data Loss Prevention (DLP) policy to block elevated, moderate, and minor risk users attempting to put information in AI apps while using Microsoft Edge.

   - **Block sensitive info sharing to AI apps in Edge**: This recommendation guides you through the process of creating a Purview Data Loss Prevention (DLP) policy to detect inline for a selection of common sensitive information types and blocks prompts from being sent in AI apps while using Microsoft Edge.

Apply compliance controls to interactions with other AI apps

If you need to ensure that interactions with supported other AI apps are retained for compliance reasons:

- In the Microsoft Purview portal, navigate to **Data Lifecycle Management** > **Policies** > **Retention Policies** and create a retention policy to retain interactions with other AI apps by selecting the location **Other AI apps** and specify the required retention period. For more information, see Create and configure retention policies.

> **Note**
>
> Retention is supported only for the Edge browser, and for ChatGPT, Microsoft Chat (consumer version), Google Gemini, and DeepSeek.

- In the Microsoft Purview portal, navigate to **eDiscovery** > **Cases** > **Create case**. In the case, create a search and use the *ItemClass* property and the `IPM.SkypeTeams.Message.ConnectedAIApp.Connector.<AppName>` value for local machine activities or the `IPM.SkypeTeams.Message.CloudAIApp.SaaS.<AppID>` value for browser-based activities to search for these interactions in your organization.

Routinely review the reports in DSPM for AI to determine if you need to make changes, and use activity explorer and events for deeper analysis of how users are interacting with other AI apps when they use a browser to third-party AI sites.

Last updated on 05/01/2026

Use Microsoft Purview to manage data security & compliance for AI agents

> *Microsoft Purview service description*

Use the following table to see at a glance the Microsoft Purview capabilities that are supported for AI interactions for AI agents from Microsoft AI apps and supported enterprise AI apps.

Managing these AI interactions with Microsoft Purview might require you to enable pay-as-you-go billing in your organization.

Capabilities supported

AI agents typically have the same security and compliance protections for AI interactions as their parent AI app. Select the name of each app to learn the support details at the solution or capabilities level.

All the listed agents are supported by Data Security Posture Management and Data Security Posture Management for AI (classic), and have their own dedicated **Apps and agents** page. Only Data Security Posture Management has an additional **AI observability** page that includes more information and support for Agent 365. Use the dashboard from those pages to view how the agents and their apps are being used across your organization.

Expand table

AI app and agents	Information protection for AI interactions	Compliance management for AI interactions
Microsoft 365 Copilot agents	✓	✓
Microsoft Security Copilot agents	✓ [1]	✓
Copilot in Fabric agents	✓ [1]	✓
Microsoft Copilot Studio agents	✓ [2]	✓
Microsoft Facilitator agents	✓ [1]	✓

AI app and agents	Information protection for AI interactions	Compliance management for AI interactions
Channel Agent in Teams agents	✓ [3]	✓
Entra-registered agents	✓	✓
Microsoft Foundry agents	✓ [2]	✓
ChatGPT Enterprise agents	✓ [4]	✓
Anthropic Claude (Enterprise) agents	✗	✗

Footnotes

The following capabilities are supported:

[1] Data classification

[2] Data classification, sensitivity labels, data loss prevention, Insider Risk Management

[3] Data classification, sensitivity labels, data loss prevention

[4] Data classification, Insider Risk Management

Last updated on 05/28/2026

Considerations to manage Microsoft 365 Copilot and Channel Agent in Teams for security and compliance

> *Microsoft Purview service description*

When you understand how you can use Microsoft Purview to manage data security and compliance protections for Microsoft 365 Copilot and Channel Agent in Teams, use the following detailed information for any prerequisites, considerations, and exemptions that might apply to your organization. Be sure to read these in conjunction with Microsoft 365 Copilot requirements and Enterprise data protection in Microsoft 365 Copilot and Microsoft Copilot 365 Chat.

For licensing information to use these capabilities for Copilot, see the licensing and service description links at the top of the page. For licensing information for Copilot, see the service description for Microsoft 365 Copilot.

Information protection considerations for Microsoft 365 Copilot

Microsoft 365 Copilot, Microsoft 365 Copilot Chat, and agents have the capability to access data stored within your Microsoft 365 tenant, including mailboxes in Exchange Online and documents in SharePoint or OneDrive.

In addition to accessing Microsoft 365 content, Copilot and agents can also use content from the specific file you're working on in the context of an Office app session, regardless of where that file is stored. For example, local storage, network shares, cloud storage, or a USB stick. When files are open by a user within an app, access is often referred to as *data in use*.

Before you deploy licenses for Copilot, make sure you're familiar with the following details that help you strengthen your data protection solutions:

- If content grants a user VIEW usage rights but not EXTRACT:
  - When a user has this content open in an app, they won't be able to use Copilot.
  - Copilot won't summarize this content but can reference it with a link so the user can then open and view the content outside Copilot.

- You can further prevent Copilot and agents from summarizing labeled files and emails (whether encrypted or not) that are identified by a Microsoft Purview Data Loss Protection policy. Copilot and agents won't summarize this content but can reference it with a link so the user can then open and view the content outside Copilot.
- Just like your Office apps, Copilot and agents can access sensitivity labels from your organization, but not other organizations. For more information about labeling support across organizations, see Support for external users and labeled content.
- An advanced PowerShell setting for sensitivity labels can prevent Office apps from sending content to connected experiences, which includes Microsoft 365 Copilot and agents.
- Copilot and agents can't access unopened documents in SharePoint and OneDrive when they're labeled and encrypted with user-defined permissions, unless the label is used as a default sensitivity label that extends SharePoint permissions to downloaded documents or they're directly referenced by typing "/" (and in both cases, the user has EXTRACT permissions). Copilot and agents can always access these documents for a user when they're open in the app (data in use).
- Sensitivity labels that are applied to containers (groups and sites) aren't inherited by items in those containers. As a result, the items won't display their container label in Copilot and can't support sensitivity label inheritance. For example, Teams channel chat messages that are summarized from a team that's labeled as Confidential won't display that label for sensitivity context in Microsoft 365 Copilot Chat. Similarly, content from SharePoint site pages and lists won't display the sensitivity label of their container label. Containers that support sensitivity labels include Teams, SharePoint sites, Micrsosoft 365 Groups, Loop workspaces, and Viva Engage.
- If you're using SharePoint information rights management (IRM) library settings that restrict users from copying text, be aware that usage rights are applied when files are downloaded, and not when they're created or uploaded to SharePoint. If you don't want Copilot and agents to summarize these files when they're at rest, use sensitivity labels that apply encryption without the EXTRACT usage right.
- Unlike other automatic labeling scenarios, an inherited label when you create new content will replace a lower priority label that was manually applied.
- When an inherited sensitivity label can't be applied, the text won't be added to the destination item. For example:
  - The destination item is read-only

    - The destination item is already encrypted and the user doesn't have permissions to change the label (requires EXPORT or FULL CONTROL usage rights)
    - The inherited sensitivity label isn't published to the user

- If a user asks Copilot or agents to create new content from labeled and encrypted items, label inheritance isn't supported when the encryption is configured for user-defined permissions or if the encryption was applied independently from the label. The user won't be able to send this data to the destination item.

- Because Double Key Encryption (DKE) is intended for your most sensitive data that's subject to the strictest protection requirements, Copilot and agents can't access this data. As a result, items protected by DKE won't be returned by Copilot, and if a DKE item is open (data in use), you won't be able to use Copilot in the app.

- Sensitivity labels that protect Teams meetings and chat aren't currently recognized by Copilot and agents. For example, data returned from a meeting chat or channel chat won't display an associated sensitivity label, copying chat data can't be prevented for a destination item, and the sensitivity label can't be inherited. This limitation doesn't apply to meeting invites, responses, and calendar events that are protected by sensitivity labels.

- For Microsoft 365 Copilot Chat (formerly known as Business Chat, Graph-grounded chat, and Microsoft 365 Chat):
    - When meeting invites have a sensitivity label applied, the label is applied to the body of the meeting invite but not to the metadata, such as date and time, or recipients. As a result, questions based just on the metadata return data without the label. For example, "What meetings do I have on Monday?" Questions that include the meeting body, such as the agenda, return the data as labeled.
    - If content is encrypted independently from its applied sensitivity label, and that encryption doesn't grant the user EXTRACT usage rights (but includes the VIEW usage right), the content can be returned by Copilot and therefore sent to a source item. An example of when this configuration can occur if a user has applied Office restrictions from Information Rights Management when a document is labeled "General" and that label doesn't apply encryption.
    - If you're using the extension capabilities that include plugins and the Microsoft Graph Connector, sensitivity labels and encryption that are applied to this data from external sources aren't recognized by Microsoft 365 Copilot Chat. Most of the time this limitation won't apply because the data is unlikely to support sensitivity labels and encryption, although one exception is Power BI data. You can always disconnect the external data

sources by using the Microsoft 365 admin center to turn off those plugins for users, and disconnect connections that use a Graph API connector.

App-specific exceptions:

- **Microsoft 365 Copilot in Outlook**: You must have a minimum version of Outlook to use Microsoft 365 Copilot for encrypted items in Outlook:
  - **Outlook (Classic) for Windows**: Starting with version 2408 in Current Channel and Monthly Enterprise Channel
  - **Outlook for Mac**: Version 16.86.609+
  - **Outlook for iOS**: Version 4.2420.0+
  - **Outlook for Android**: Version 4.2420.0+
  - **Outlook on the web**: Yes
  - **New Outlook for Windows**: Yes
- **Microsoft 365 Copilot in Edge**, **Microsoft 365 Copilot in Windows**: Unless data loss prevention (DLP) is used in Edge, Copilot can reference encrypted content from the active browser tab in Edge when that content doesn't grant the user EXTRACT usage rights. For example, the encrypted content is from Office for the web or Outlook for the web.

Will an existing label be overridden for sensitivity label inheritance?

Summary of outcomes when Microsoft 365 Copilot automatically applies protection with sensitivity label inheritance:

Expand table

Existing label	Override with sensitivity label inheritance
Manually applied, lower priority	Yes
Manually applied, higher priority	No
Automatically applied, lower priority	Yes
Automatically applied, higher priority	No
Default label from policy, lower priority	Yes
Default label from policy, higher priority	No
Default sensitivity label for a document library, lower priority	Yes

Existing label	Override with sensitivity label inheritance
Default sensitivity label for a document library, higher priority	No

Copilot honors existing protection with the EXTRACT usage right

Although you might not be very familiar with the individual usage rights for encrypted content, they've been around a long time. From Windows Server Rights Management, to Active Directory Rights Management, to the cloud version that became Azure Information Protection with the Azure Rights Management service.

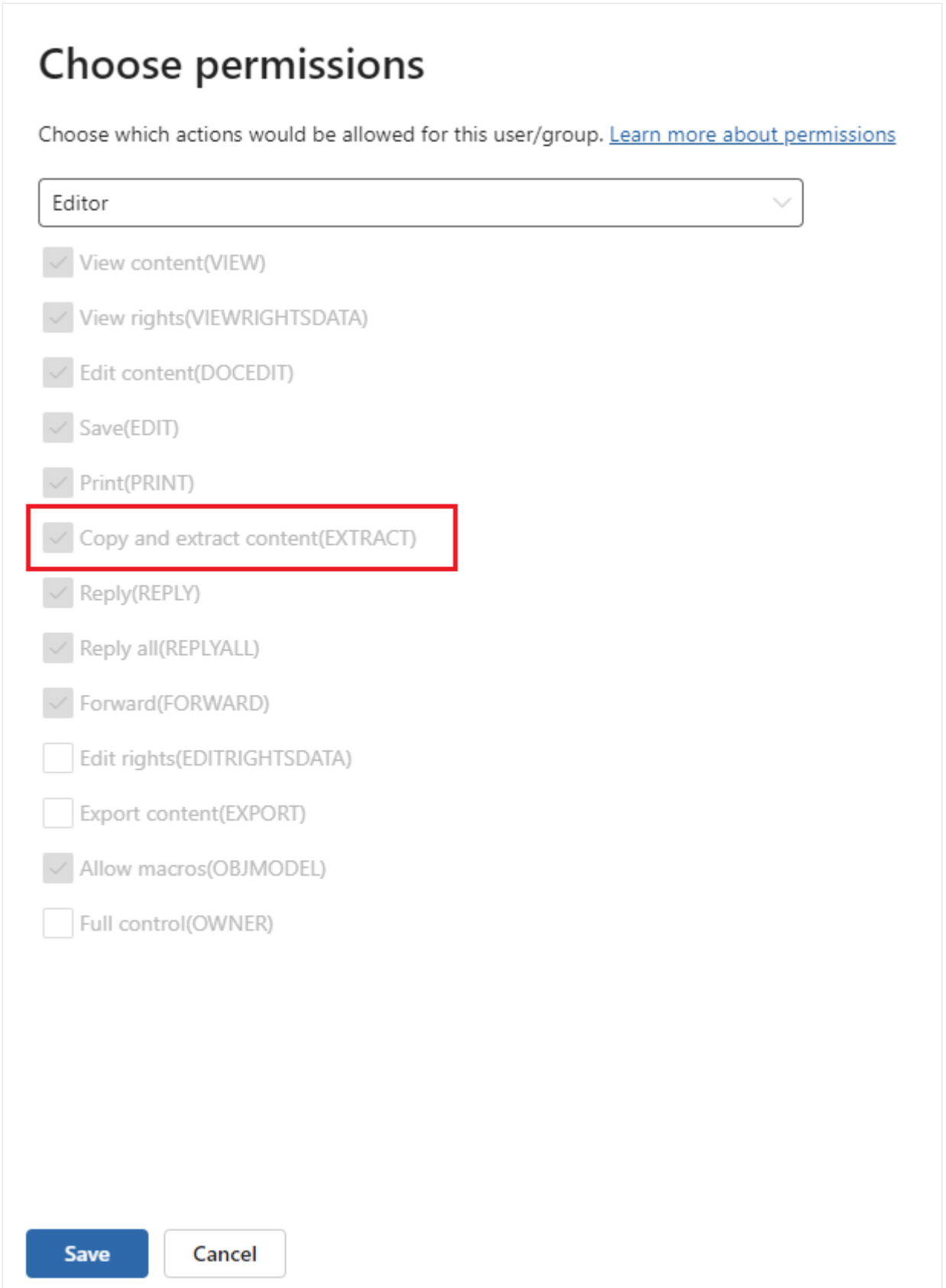
If you've ever received a "Do Not Forward" email, it's using usage rights to prevent you from forwarding the email after you've been authenticated. As with other bundled usage rights that map to common business scenarios, a Do Not Forward email grants the recipient usage rights that control what they can do with the content, and it doesn't include the FORWARD usage right. In addition to not forwarding, you can't print this Do Not Forward email, or copy text from it.

The usage right that grants permission to copy text is **EXTRACT**, with the more user-friendly, common name of **Copy**. It's this usage right that determines whether Copilot or agents can display text to the user from encrypted content.

> ⓘ **Note**
>
> Because the Full control (OWNER) usage right includes all the usage rights, EXTRACT is automatically included with Full control.

When you use the Microsoft Purview portal to configure a sensitivity label to apply encryption, the first choice is whether to assign the permissions now, or let users assign the permissions. If you assign now, you configure the permissions by either selecting a predefined permission level with a preset group of usage rights, such as Co-Author or Reviewer. Or, you can select custom permissions where you can individually select available usage rights.

In the Microsoft Purview portal, the EXTRACT usage right is displayed as **Copy and extract content(EXTRACT)**. For example, the default permission level selected is **Editor**, where you see **Copy and extract content(EXTRACT)** is included. As a result, content protected with this encryption configuration can be returned by Copilot and agents



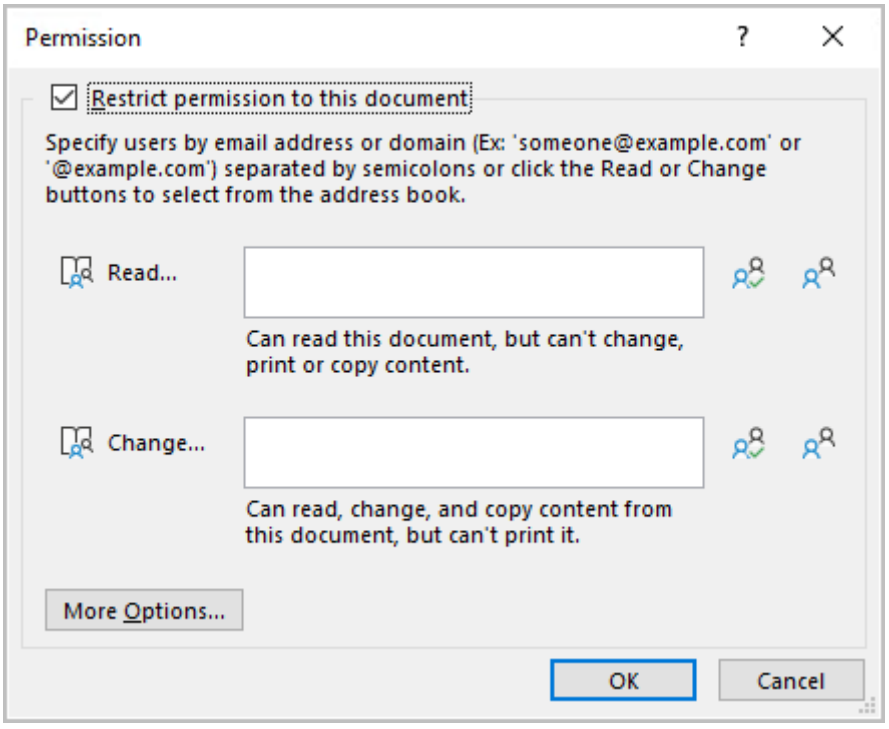
If you select **Custom** from the dropdown box, and then **Full control(OWNER)** from the list, this configuration will also grant the EXTRACT usage right.

> ⓘ **Note**

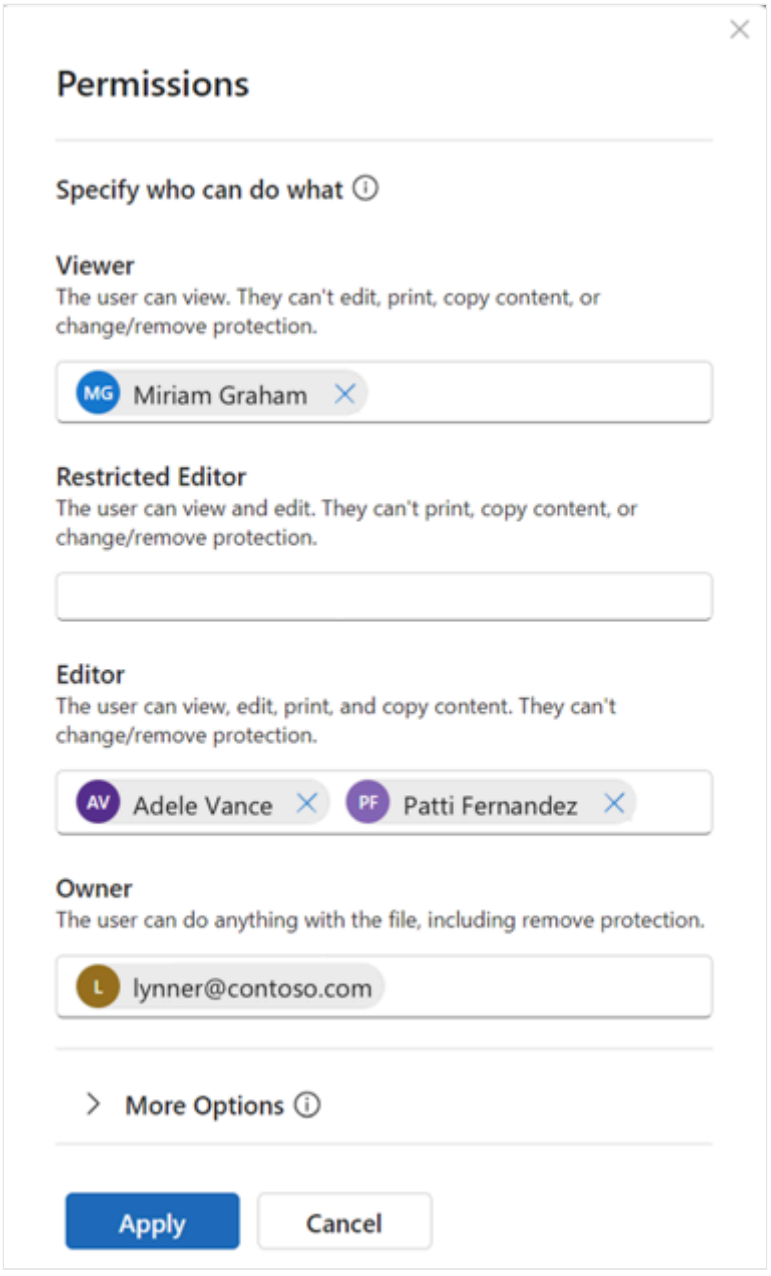
The person applying the encryption always has the EXTRACT usage right, because they are the Rights Management owner. This special role automatically includes all usage rights and some other actions, which means that content a user has encrypted themselves is always eligible to be returned to them by Microsoft 365 Copilot, Microsoft 365 Copilot Chat, and agents. The configured usage restrictions apply to other people who are authorized to access the content.

Alternatively, if you select the encryption configuration to let users assign permissions, for Outlook, this configuration includes predefined permissions for Do Not Forward and Encrypt-Only. The Encrypt-Only option, unlike Do Not Forward, does include the EXTRACT usage right.

When you select custom permissions for Word, Excel, and PowerPoint, users select their own permissions in the Office app when they apply the sensitivity label. There are currently two versions of the dialog box. In the older version, they're informed that from the two selections, **Read** doesn't include the permission to copy content, but **Change** does. These references to copy refer to the EXTRACT usage right. If the user selects **More Options**, they can add the EXTRACT usage right to Read by selecting **Allow users with read access to copy content**.



In the latest version of the dialog box, **Read** and **Change** is replaced with permission levels where the descriptions that state copy isn't allowed won't include the EXTRACT usage right. The permission levels that do include EXTRACT are **Editor** and **Owner**. For example:



> **Tip**
>
> If you need to check whether a document you're authorized to view includes the EXTRACT usage right, open it in the Windows Office app and customize the status bar to show **Permissions**. Select the icon next to the sensitivity label name to display **My Permission**. View the value for **Copy**, which maps to the EXTRACT usage right, and confirm whether it displays **Yes** or **No**.
>
> For emails, if the permissions aren't displayed at the top of the message in Outlook for Windows, select the information banner with the label name, and then select **View Permission**.

Copilot and agents honor the EXTRACT usage right for a user, however it was applied to the content. Most times, when the content is labeled, the usage rights granted to the user match

those from the sensitivity label configuration. However, there are some situations that can result in the content's usage rights being different from the applied label configuration:

- The sensitivity label is applied after usage rights are already applied
- Usage rights are applied after a sensitivity label has been applied

For more information about configuring a sensitivity label for encryption, see Restrict access to content by using sensitivity labels to apply encryption.

For technical details about the usage rights, see Configure usage rights for the Azure Rights Management service.

Information protection considerations for Channel Agent in Teams

Channel Agent in Teams uses sensitivity labeled information to answer user questions that might be derived from its channel files in SharePoint, channel messages and meetings, and standard meetings. For these sensitivity labels, they have many of the same information protection considerations as those listed in the previous section for Microsoft 365 Copilot. For example:

- The highest priority label is displayed in the chat and is used for label inheritance when new AI-generated content is created.
- The content from items encrypted with user-defined permissions and Double Key Encryption isn't accessible.
- Sensitivity labels applied to the channel from the labels' groups and sites configuration aren't used by Channel Agent in Teams.

However, unlike sensitivity labels in Microsoft 365 Copilot:

- Although Channel Agent in Teams honors the EXTRACT permission like Copilot does for the requesting user, the permissions aren't checked for all users in the channel. As a result, Channel Agent could summarize the content of items that one or more users in the channel don't have permissions to open. The requesting user is warned of this risk and administrators have the ability to turn off Channel Agent in Teams.
- You can't prevent Channel Agent from summarizing labeled files that are identified by a Microsoft Purview Data Loss Protection policy.
- You can't prevent users from copying chat responses from Channel Agent or prevent these responses from being used by another channel agent in Teams channels.
- Information barriers aren't supported for Channel Agent and as a result, these it might return information that crosses barriers.

- Channel Agent in Teams isn't supported with Microsoft Purview Customer Key.

Compliance management considerations for Microsoft 365 Copilot

Compliance management for interactions with Microsoft 365 Copilot and agents extend across the following apps: Word, Excel, PowerPoint, Outlook, Teams, Loop and Copilot Pages, Whiteboard, OneNote, and Microsoft 365 Copilot Chat (formerly Business Chat, Graph-grounded chat, and Microsoft 365 Chat).

> ⓘ **Note**
>
> Compliance management for Microsoft 365 Copilot Chat includes prompts and responses to and from the public web when users are signed in and select the **Work** option rather than **Web**.

Compliance tools identify the source Microsoft 365 Copilot interactions by the name of the app. For example, **Copilot in Word**, **Copilot in Teams**, and **Microsoft 365 Chat** (for Microsoft 365 Copilot Chat).

Before your organization uses Microsoft 365 Copilot, make sure you're familiar with the following details to support your compliance management solutions:

- The audit data is designed for data security and compliance purposes, and provides comprehensive visibility into Microsoft 365 Copilot interactions for these use cases. For example, to discover data oversharing risks or to collect interactions for regulatory compliance or legal purposes. It's not intended to be used as the basis for Copilot usage reporting.

  > ⓘ **Note**
  >
  > Any aggregated metrics that you build on top of this audit data such as "prompt count" or "active user count" might not be consistent with the corresponding data points in the official Copilot usage reports provided by Microsoft. Microsoft can't provide guidance on how to use audit log data as the basis for usage reporting, nor can Microsoft guarantee that aggregated usage metrics built on top of audit log data will match similar usage metrics reported in other tools.

To access accurate information about Copilot usage, use one of the following reports: the **Microsoft 365 Copilot usage report** in the Microsoft 365 Admin Center or the **Copilot Dashboard** in Viva Insights.

- Auditing captures the Microsoft 365 Copilot activity of search, but not the actual user prompt or response. For this information, use eDiscovery. Or from Microsoft Purview Data Security Posture Management for AI, use **AI interaction** from the **Activity explorer** page.
- Admin-related changes for auditing Copilot aren't yet supported.
- Device identify information isn't currently included in the audit details.
- Retention policies for Copilot interactions don't inform users when messages are deleted as a result of a retention policy.

App-specific exceptions:

- **Microsoft 365 Copilot in Teams**:
  - If transcripts are turned off, the capabilities for auditing, eDiscovery, and retention aren't supported
  - If transcripts are referenced, this action isn't captured for auditing
  - For Microsoft 365 Copilot Chat, Copilot currently can't retain as cloud attachments, referenced files that it returns to users. Files referenced by users are supported to be retained as cloud attachments.

Last updated on 06/30/2026

Learn about Data Security Posture Management

> *Microsoft Purview service description*

> **Note**
>
> This documentation is for the **current** version of Data Security Posture Management that expands coverage to more data sources, introduces guided workflows for proactive risk management, and streamlines data security operations so you can more confidently adopt AI across your digital estate.
>
> Most new features will be added to this version only but you can still access the previous versions and their documentation:
>
> - DSPM for AI (classic)
> - DSPM (classic)

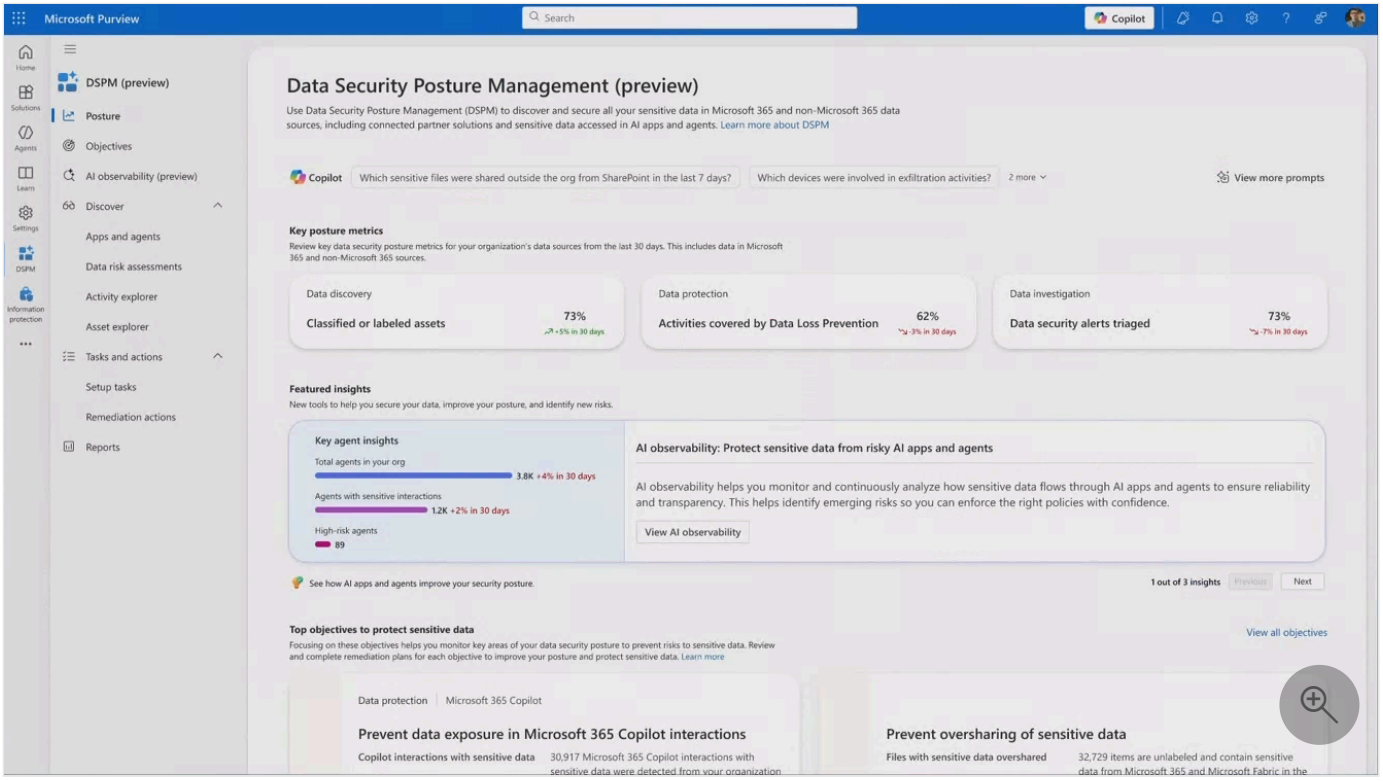
Microsoft Purview Data Security Posture Management (DSPM) helps organizations discover, protect, and investigate sensitive data risks across their digital estate. This solution provides unified visibility and control for both traditional applications and AI apps and agents, supporting data governance across Microsoft 365, Azure, Fabric, and integrated third-party SaaS platforms. Monitor, assess, and remediate data risks, regardless of where sensitive data resides.

Instead of focusing on infrastructure or endpoints, Data Security Posture Management centers on the data itself—identifying where it resides, who can access it, how it's used, and whether it's adequately protected. This is especially important as data becomes more distributed and exposed in today's AI-driven workplaces where data is constantly moving and changing, making it harder to keep track of and control.

Data Security Posture Management continuously scans your environment to identify sensitive data, assess risk, and recommend actions to reduce exposure. It consolidates insights from the Microsoft Purview solutions data loss prevention (DLP), Insider Risk Management, information protection with sensitivity labels, and Data Security Investigations. These insights provide a single view for monitoring data risks, policy coverage, and posture trends. This version of Data Security Posture Management extends coverage to third-party SaaS and IaaS platforms, such as Google

Cloud Platform, Snowflake, and Databricks, and integrates with partner solutions such as Varonis, Cyera, BigID, and OneTrust for comprehensive risk insights.

Dedicated dashboards and metrics monitor risks associated with AI apps and agents. AI observability enables you to track agent-specific activities, such as oversharing, exfiltration, and unusual access patterns, across both Microsoft and third-party environments. Enhanced reporting provides advanced filtering and customizable views, supporting granular analysis of sensitive data usage, DLP activity, and posture trends. The dashboards and reports can be tailored to align with your organization's compliance requirements and operational priorities. Audit logs and activity explorer features help track interactions with AI apps and agents, supporting compliance investigations and incident response. Integrated investigation and forensics tools support rapid incident response and root cause analysis for data security events.



These insights help uncover shadow data, classify sensitive information, analyze access patterns, and highlight risks. The unified approach replaces the need for multiple solutions and manual audits, answering the four practical questions:

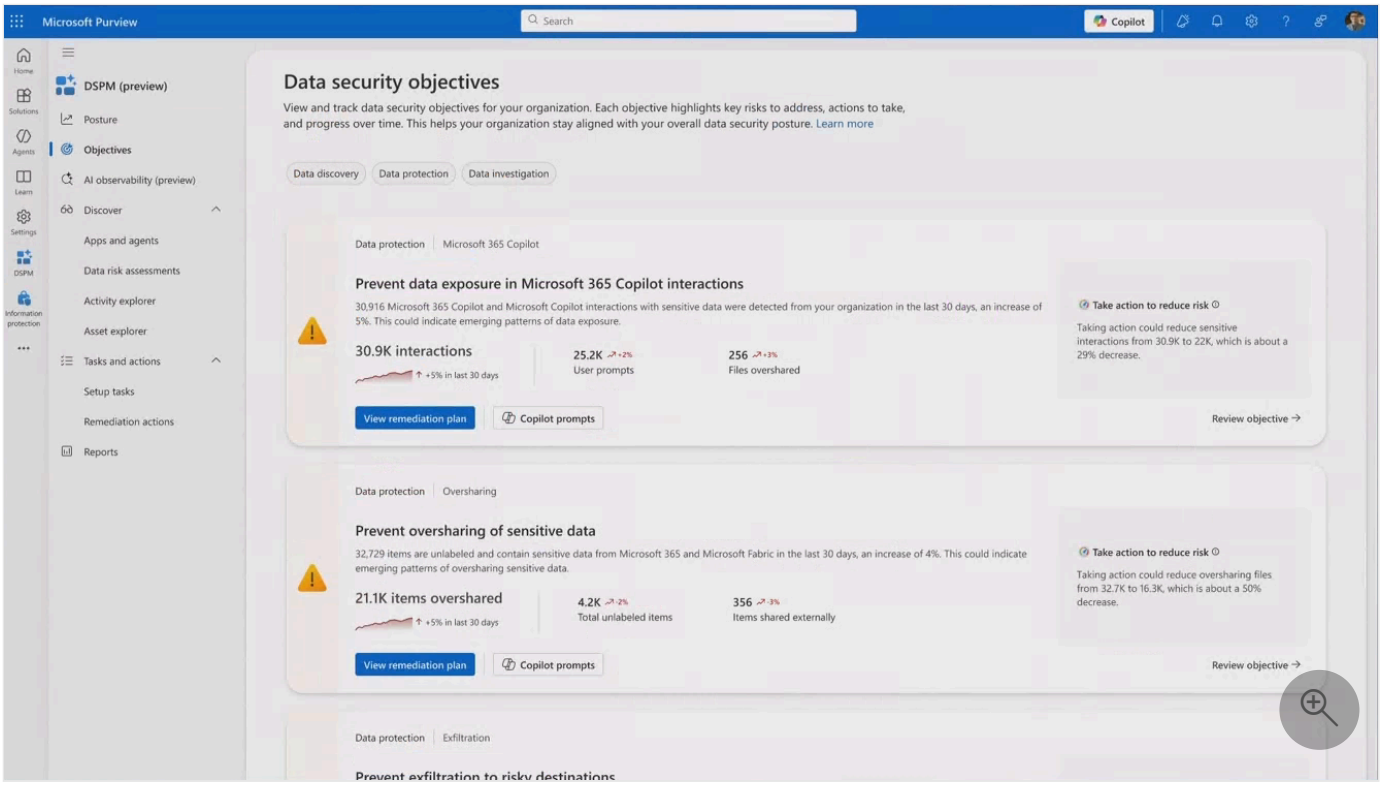
- What data do we have?
- Where is it stored?
- Who can access it?
- How is it protected?

When these questions are addressed, Data Security Posture Management can automate remediation steps, such as removing public sharing links or applying data loss prevention

policies, to help prevent incidents before they happen. Investigation, alerting, analytics, and reporting are all integrated, supporting a layered security approach. Key metrics are tracked over time, supporting continuous improvement of your organization's data security posture.

Data security objectives

New in this version of Data Security Posture Management, after the initial setup tasks, you see **Data security objectives** displayed as prominent, selectable cards or modules. They're also available in context from the initial **Posture** page. Each security objective represents a specific security goal, such as **Prevent data exposure in Microsoft 365 Copilot and Microsoft Copilot interactions**, **Prevent oversharing of sensitive data**, **Prevent exfiltration to risky locations**, and **Discover sensitive data in your organization**.



Selecting each one guides you through an end-to-end workflow by grouping together the most relevant Microsoft Purview solutions, such as information protection, data loss prevention (DLP), Insider Risk Management, and eDiscovery. As a result, you can focus on achieving specific data security outcomes rather than navigating separate solutions. Each **Outcome** card displays key metrics, such as the percentage of data covered by policies, number of risky sharing incidents, or improvements over time. This information lets you quickly see your current security posture and track progress as you remediate risks.

Within each outcome, you see suggested prioritized actions, such as applying sensitivity labels, configuring DLP policies, or investigating alerts, all tailored to your organzation's data. You can

take action directly from the workflow, such as remediating oversharing, configuring one-click policies, or launching investigations into suspicious activity. Reporting and analytics are also organized by outcome, making it easier to identify and report improvements, compliance, and risk reduction.

The **Prevent exfiltration to risky destinations** objective also supports proactive AI insights powered by Data Security Investigations. When enabled, Data Security Investigations automatically creates and refreshes an investigation that continuously analyzes recently exfiltrated sensitive data across five risk categories.

Operational insights are surfaced throughout Data Security Posture Management, including:

- Impact prediction visuals and progress tracking for remediation steps
- Role-based access controls to provide granular access to features and AI content for delegated administration and compliance

How AI helps achieve data security outcomes

This version of Data Security Posture Management not only secures and governs AI apps and agents, but also uses Microsoft Security Copilot and AI agents to help secure and govern your data. This AI integration helps you more quickly identify, protect, and investigate data that resides across cloud, SaaS, on-premises, and AI environments. Data security becomes more adaptive and less reliant on manual intervention. AI analyzes access patterns, sharing behaviors, and policy gaps to surface actionable risks. It can also detect unusual activity, such as excessive sharing or suspicious downloads, and prioritize incidents that seem to need immediate attention.

You're always in control of your data and under your guidance, AI agents can take direct action on detected risks, such as removing public sharing links, applying DLP policies, or revoking permissions. To streamline investigations, AI-driven triage agents review alerts from the DLP and insider risk management solutions, filtering out noise and highlighting the most critical threats. You review, approve, or customize automated actions taken by AI agents, and these actions are always audited. Use the **View agent activity** options throughout the data security objectives for easy access to the agents' activity.

For example, the proactive AI insights feature uses Data Security Investigations to automatically analyze exfiltrated data and surface risk counts by category on the exfiltration objective card, without requiring manual investigation creation.

These AI capabilities from Data Security Posture Management help ensure that sensitive data is governed, labeled, and monitored, with streamlined management. For more information:

- How AI is used within Data Security Posture Management, see Application card: Microsoft Purview Data Security Posture Management
- For the best experience using Security Copilot prompts, see Tips for custom Security Copilot prompts in Data Security Posture Management

How to use Data Security Posture Management

To get started with Data Security Posture Management, use the Microsoft Purview portal. You need an account that has appropriate permissions for security and compliance management, such as an account that's a member of the Entra **Compliance Administrator** role, or the **Microsoft Purview Compliance Administrator** role group. Some Data Security Posture Management activities need additional permissions, such as the Data Security Viewer role to use Security Copilot. For more information, see Permissions for Data Security Posture Management.

If you're new to Data Security Posture Management, the following steps provide a recommended walkthrough.

1. Sign in to the Microsoft Purview portal > **Solutions** > **DSPM**.

   Don't confuse this with the previous versions, that are now named **Data Security Posture Management (classic)** and **DSPM for AI (classic)**.

2. With the first-time use of DSPM, you're prompted to accept some initial setup tasks if they're not already turned on for your tenant. Allow a day or so before you start to see data from your tenant that you can take action on.

   In the meantime, watch the Microsoft Mechanics video, New Data Security Posture Management for an overview introduction with same sample data to see how you might use the following key pages from DSPM to discover, protect, and investigate sensitive data:

   - **Posture**: The dashboard that lets you immediate interact with Security Copilot with suggested prompts, or work your way down the key posture metrics, top objectives to address based on risk, a snapshot of data use across your data estate, and a 30-day trending graph for your organization's data security posture.

   - **Objectives**: Data security objectives to help you address data security risks that have been identified for your organization, each with a remediation plan that includes one-click policies and recommended actions.

   - **AI observability**: An inventory of all AI apps and agents (including recently released Microsoft Agent 365) with activity in the last 30 days, how many are high risk and the

total with sensitive interactions, with a breakdown of individual agents and policies to govern them.

- **Asset explorer**: Use the **Agent** tab for the Data Security Posture Agent, and the **Standard** tab to see a unified view across your organization of unlabeled or classified data by workload, filtered by Microsoft and non-Microsoft locations. Microsoft locations currently include Microsoft 365 only. Non-Microsoft locations are made possible by integration with partner solutions.

- **Reports**: To help you track sensitive data usage and labeling, policy usage, and risky behavior of users and AI agents.

- **Setup tasks**: To identify and complete configuration steps independently from the security objectives. They include, in preview, setting up Microsoft Sentinel data lake to integrate with partner solutions for non-Microsoft data sources, such as Google Cloud Platform and Snowflake.

3. Explore the following DSPM pages that aren't directly covered in the video:

   - **Discover** > **Apps and agents**: To view a dashboard of AI apps and their agents that have been used across your organization, although it doesn't include Agent 365 (use the **AI observability** page, instead). For the top 20 most recently used agents, view details about sensitive data that they accessed and how they're protected by policies from Microsoft Purview.

   - **Discover** > **Activity explorer**: To understand activity related to content that contains sensitive information or has sensitivity labels applied. Use the **AI activities** tab to see data specific to AI interactions, such as when a user browsed to a generative AI site, the prompts and responses and if they contained sensitive information, and when a data loss prevention rule was matched during one of these interactions. For more information about these events, see Activity explorer events in Data Security Posture Management.

   - **Discover** > **Data risk assessments**: To help prevent oversharing of sensitive data in your organization with the default or custom assessments that identify and fix potential data oversharing risks. For more information, see Prevent oversharing with data risk assessments from Data Security Posture Management.

   - **Tasks and actions** > **Remediation actions**: To identify and create automatic policies independently from the initial setup tasks and security objectives. Most of these

policies originate from the previous version, DSPM for AI. For more information about them, see:

- Default policies for data discovery using Data Security Posture Management for AI
- Default policies from data security to help you protect sensitive data used in generative AI

If you're familiar with the previous versions that are now named DSPM for AI (classic) and DSPM (classic), you might find it helpful to refer to Find familiar tasks that you did in DSPM for AI or in DSPM.

Next steps

Read the blog post announcements for this new release of Data Security Posture Management:

- Empowering organizations with integrated data security: What's new in Microsoft Purview
- Beyond Visibility: The new Microsoft Purview Data Security Posture Management (DSPM) experience

Learn how Microsoft Purview secures and governs AI apps and agents: Microsoft Purview data security and compliance protections for generative AI apps

Learn about Security Copilot and agents in Microsoft Purview:

- Security Copilot in Microsoft Purview Overview
- Security Copilot Agent in Microsoft Purview Overview

Last updated on 05/26/2026

Find familiar tasks that you did in DSPM for AI or in DSPM

> *Microsoft Purview service description*

> **Note**
>
> This documentation is for the **current** version of Data Security Posture Management that expands coverage to more data sources, introduces guided workflows for proactive risk management, and streamlines data security operations so you can more confidently adopt AI across your digital estate.
>
> Most new features will be added to this version only but you can still access the previous versions and their documentation:
>
> - DSPM for AI (classic)
> - DSPM (classic)

Microsoft Purview Data Security Posture Management (DSPM) for AI, and Microsoft Purview Data Security Posture Management (DSPM), have converged into a single, central solution. This unified version helps you more easily discover, protect, and investigate your data by using security and governance solutions from Microsoft Purview.

Many of tasks that you used to do in previous versions are now embedded in the context of data security objectives, although they can still be accessed independently. Data security objectives are new in Data Security Posture Management, and you find them front and center with **DSPM** > **Objectives**.

Data security objectives focus on the business outcomes you want to achieve. Choose an objective and we recommend a combination of Microsoft Purview solutions and specific actions to help you meet that outcome, with as few clicks as possible to achieve it. For more information, see Objectives and outcomes.

To help you navigate the new Data Security Posture Management for the tasks you used to do in previous versions, use the following tables.

Tasks that you did in DSPM for AI

Expand table

Task in DSPM for AI	Navigation for the new Data Security Posture Management
Setup tasks that include auditing and extending your insights for data discovery	1. When you first open **DSPM** > **Getting Started** 2. **DSPM** > **Actions** > **Setup tasks**
Recommendations for AI apps	**DSPM** > **Actions** > **Remediation actions**
Microsoft 365 Copilot overview page	**DSPM** > **Reports** > **Microsoft 365 Copilot**
Reports to see results of one-click policies	**DSPM** > **Reports**
Policies for AI apps	**DSPM** > **Reports** > **Policies with AI workloads**
Apps and agents	**DSPM** > **Discover** > **Apps and agents** See also: **AI observability** that includes more information and support for Agent 365
Activity explorer	**DSPM** > **Discover** > **Activity explorer** > **AI activities** tab
Data risk assessments	**DSPM** > **Discover** > **Data risk assessments**

Tasks that you did in DSPM

Expand table

Task in DSPM	Navigation for the new Data Security Posture Management
Opt in to analytics processing and scanning	When you first open **DSPM** > **Getting Started**
Recommendations	**DSPM** > **Actions** > **Remediation actions**
Reports and trends	**DSPM** > **Reports** See also: **DSPM** > **Posture**
Security Copilot promptbooks	**DSPM** > **Posture** Security Copilot promptbooks are surfaced contextually within objectives and their workflows

Get started with the new Data Security Posture Management

Sign in to the Microsoft Purview portal > **Solutions** > **DSPM**.

Use an account with one of the following roles or role groups:

- **Microsoft Entra Compliance Administrator** role
- **Microsoft Entra Global Administrator** role
- **Microsoft Purview Compliance Administrator** role group

> **Note**
>
> Some activities require additional roles. For more information, see Permissions for Data Security Posture Management.

For more information about this latest version of Data Security Posture Management, with a guided walkthrough, see Learn about Data Security Posture Management.

Last updated on 05/01/2026

Prevent oversharing with data risk assessments from Data Security Posture Management

> *Microsoft Purview service description*

> **Note**
>
> This documentation is for the **current** version of Data Security Posture Management that expands coverage to more data sources, introduces guided workflows for proactive risk management, and streamlines data security operations so you can more confidently adopt AI across your digital estate.
>
> Most new features will be added to this version only but you can still access the previous versions and their documentation:
>
> - DSPM for AI (classic)
> - DSPM (classic)

Data risk assessments from Data Security Posture Management help you identify and fix potential data oversharing risks in your organization. Because of the power and speed AI can proactively surface content that might be obsolete, over-permissioned, or lack governance controls, generative AI amplifies the problem of oversharing data. Use data risk assessments to both identify and remediate issues.

You can access data risk assessments from the Microsoft Purview portal > **DSPM** > **Discover** > **Data risk assessments**, and insights from these assessments are surfaced from the security objective **Prevent oversharing of sensitive data**.

Default data risk assessments

For oversharing insights in SharePoint and OneDrive, use the **Microsoft 365** tab, and the **Fabric** tab for items in Fabric workspaces.

Microsoft 365

A default data risk assessment automatically runs weekly for the top 100 SharePoint sites based on usage in your organization, and you might have already run a custom assessment as one of the recommendations. However, come back regularly to this option to check the latest weekly results of the default assessment and run custom assessments when you want to check for different users or specific sites. After a custom assessment has run, wait at least 48 hours to see the results that don't update again. You'll need a new assessment to see any changes in the results.

The **Default assessment** displays at the top of the page with a quick summary, such as the total number of items found, the number of sensitive data detected, and the number of links sharing data with anyone. The first time the default assessment is created there's a 4-day delay before results are displayed.

After you select **View details** for more in-depth information, from the list, select each site to access the flyout pane that has tabs for **Overview**, **Identify**, **Protect**, and **Monitor**. Use the information on each tab to learn more, and take recommended actions. For example:

Use the **Identify** tab to identify how much data has been scanned or not scanned for sensitive information types, with an option to initiate an on-demand classification scan as needed.

Use the **Protect** tab to select options to remediate oversharing, which include:

- **Restrict access by label**: Use Microsoft Purview Data Loss Prevention to create a DLP policy that prevents Microsoft 365 Copilot and agents from summarizing data when it has sensitivity labels that you select. For more information about how this works and supported scenarios, see Learn about using Microsoft Purview Data Loss Prevention to protect interactions with Microsoft 365 Copilot and Copilot Chat.
- **Restrict all items**: Use SharePoint Restricted Content Discovery to list the SharePoint sites to be exempt from Microsoft 365 Copilot. For more information, see Restrict discovery of SharePoint sites and content.
- **Create an auto-labeling policy**: When sensitive information is found for unlabeled files, use Microsoft Purview Information Protection to create an auto-labeling policy to automatically apply a sensitivity label for sensitive data. For more information about how to create this policy, see How to configure auto-labeling policies for SharePoint, OneDrive, and Exchange.
- **Create retention policies**: When content hasn't been accessed for at least 3 years, use Microsoft Purview Data Lifecycle Management to automatically delete it. For more

information about how to create the retention policy, see Create and configure retention policies.

Use the **Monitor** tab to view the number of items in the site shared with anyone, shared with everyone in the organization, shared with specific people, and shared externally. Select **Start a SharePoint site access review** for information how to use the SharePoint data access governance reports.

Custom data risk assessments

To create your own custom data risk assessment, select **Create custom assessment** to identify potential oversharing issues for all or selected users, the data sources to scan, and run the assessment.

> **Note**
>
> Microsoft 365 custom assessments optionally support item-level scanning with remediation actions when you complete a one-time authentication process. To specify the required settings, you must have a registered Entra application. For more information about the Entra application, see Prerequisites for Microsoft 365 item-level scanning for data risk assessments.
>
> Select **Item-level** on the **Scan level** page, and then **Authenticate** to specify the Entra application information.

This data risk assessment is created in the **Custom assessments** category. Wait for the status of your assessment to display **Completed**, and select it to view details. To rerun a custom data risk assessment, and to see results after the 30-day expiration, use the duplicate option to create a new assessment with the same selections.

Item-level scanning and remediation for items that are potentially overshared is applicable to Microsoft 365 only, and currently restricted to SharePoint sites. This scan identifies items as potentially overshared if they have a sharing link for external or anonymous users, and also shows any applied sensitivity label and the owner of each item.

When the scan is complete, open the custom assessment and view item-level insights from the **Potentially overshared items** tab or the **View items** button from the **Potentially overshared**

items banner. The following remediation actions can be taken on the identified potentially overshared items:

- **Resolve**, for example if you decide the item isn't at risk of oversharing.
- **Apply sensitivity label** for items that are identified as currently unlabeled or should have a different sensitivity label.
- **Notify** the site owner with an email notification (not customizable).
- **Remove sharing link** to remove the existing sharing link so it can no longer be used to access the item. This is an action to be used sparingly because it could prevent people from accessing the item legitimately. In this scenario, the site owner or item owner must then set a less permissive sharing link type for authorized users. If you haven't already, consider configuring sensitivity labels for a default sharing link.

The email notification contains information and instructions about the potentially overshared items in the site. It also contains options for them to go to the SharePoint site to manage access to the items from there, or to the Microsoft Purview portal for a read-only view of the item-level scan results that only show items and sites that they own.

> **Tip**
>
> Both the default and custom data risk assessments provide an **Export** option that let you save and customize the data into a choice of file formats (Excel, CSV, JSON, TSV).

Limits for Microsoft 365:

- A maximum of 200,000 items per location, which applies to both a custom data risk assessment and a default data risk assessment. The count of files reported might not be accurate when there are more than 100,000 files per location.
- OneDrive currently isn't supported for item-level scanning.
- A current maximum of 10 SharePoint sites for item-level scanning.

Last updated on 05/01/2026

Setup tasks for Data Security Posture Management

> *Microsoft Purview service description*

> **Note**
>
> This documentation is for the **current** version of Data Security Posture Management that expands coverage to more data sources, introduces guided workflows for proactive risk management, and streamlines data security operations so you can more confidently adopt AI across your digital estate.
>
> Most new features will be added to this version only but you can still access the previous versions and their documentation:
>
> - DSPM for AI (classic)
> - DSPM (classic)

When you first access Data Security Posture Management (DSPM) for your tenant, you see an initial setup task for **Audting and analytics** that's required to see meaningful data that you can take action on. Optionally, create collection policies to capture AI interactions. Other setup tasks can be revealed in the context of data security objectives.

All setup tasks can found from **DSPM** > **Actions** > **Setup tasks**. They include step-by-step instructions, status, estimated time to complete, and other helpful information such as any impact to users and links to additional information.

Some of these setup tasks include creating preconfigured policies with just one click so you can more quickly gain insights and protect your data. These one-click policies originated from the previous version, and so have the prefix of **DSPM for AI**. For more information about these policies, see the DSPM for AI documentation, One-click policies from Data Security Posture Management for AI.

> **Note**

Although not listed as setup tasks, you might also want to check for **prerequisites**, such as permissions, and other requirements for specific tasks.

Activate Microsoft Purview Audit

This setup task is likely already complete because auditing is on by default for new tenants, and checked when you first run Data Security Posture Management. If you need more information, see Learn about auditing solutions in Microsoft Purview and Turn auditing on or off.

Extend insights into sensitive data in AI interactions

Create a one-click policy that helps you gain visibility into sensitive data shared wtih AI apps through network activity.

Extend insights for data discovery

Create one-click policies for collecting information about users visiting third-party generative AI sites and sending sensitive information to them.

Install Microsoft Purview browser extension

The Microsoft Purview browser extension must be deployed to Windows users to discover visits to third-party generative AI sites by using an Insider Risk Management policy. The browser extension is also required for endpoint DLP policies on Windows when you use Chrome.

Integrate with partner solutions

This options is in preview.

Set up Microsoft Sentinel data lake to integrate partner solutions and centralize security posture monitoring across all data sources. To see the available partner solutions and then select one from this task, you must first connect to the Microsoft Sentinel content hub. Initial partners include Varonis (for Salesforce), Cyera (for Snowflake), BigID (for Databricks), and OneTrust (for Google Cloud Platform).

Onboard devices to Microsoft Purview

Devices must be onboarded to Microsoft Purview for:

- Gaining visibility into sensitive information that's shared with third-party generative AI sites. For example, a user pastes credit card numbers into ChatGPT.
- Applying endpoint DLP policies to warn or block users from sharing sensitive information with third-party generative AI sites. For example, a user identified as elevated risk in Adaptive Protection is blocked with the option to override when they paste credit card numbers into ChatGPT.

Protect your data with sensitivity labels

Create default sensitivity labels and sensitivity label policies, if they don't already exist in your tenant.

Secure interactions from enterprise AI apps

Create a one-click policy that captures prompts and responses for regulatory compliance from enterprise AI apps, such as Chat GPT Enterprise and AI apps connected through Microsoft Entra or Azure AI Foundry, so they can be managed in Microsoft Purview solutions like eDiscovery, Data Lifecycle Management, and more.

Secure interactions in Microsoft Copilot experiences

Create a one-click policy that captures prompts and responses for data security posture and regulatory compliance from Copilot in Fabric and Security Copilot.

Set up pay-as-you-go billing for data storage

For some options you must set up pay-as-you-go billing for your organization. When this billing model is applicable for specific configurations, you'll see notifications and instructions in the UI.

You might find it helpful to use the **Usage center** to help you understand and manage your usage. For more information, see Manage pay-as-you-go and per-user licensing usage.

Last updated on 05/01/2026

Permissions for Data Security Posture Management

> *Microsoft Purview service description*

> **Note**
>
> This documentation is for the **current** version of Data Security Posture Management that expands coverage to more data sources, introduces guided workflows for proactive risk management, and streamlines data security operations so you can more confidently adopt AI across your digital estate.
>
> Most new features will be added to this version only but you can still access the previous versions and their documentation:
>
> - DSPM for AI (classic)
> - DSPM (classic)

Members of your security and compliance teams who are responsible for managing AI apps in Microsoft Purview Data Security Posture Management need appropriate permissions when they sign in to the Microsoft Purview portal.

Roles and role groups that can view, create, and edit in Data Security Posture Management:

- **Microsoft Entra Compliance Administrator** role
- **Microsoft Entra Global Administrator** role
- **Microsoft Purview Compliance Administrator** role group

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

Roles and role groups that can view-only in Data Security Posture Management:

- **Microsoft Purview Security Reader** role group
- **Purview Data Security AI Viewer** role
- **Data Security Viewers** role group and **Data Security Viewer** role
- AI Administrator role from Entra
- **Purview Data Security AI Content Viewer** role for AI interactions only

To help you assign the right permissions to users, use the following guidance, depending on the portal you're using:

- Permissions in the Microsoft Purview portal
- Assign Microsoft Entra roles to users
- Permissions in Exchange Online

Use the following table to understand the detailed permissions for different activities in Data Security Posture Management.

Permissions by activities

✓: Supported. The role or role group have permissions to do the specified activities.

✕: Not supported. The role or role group don't have permissions to do the specified activities.

Expand table

Activities	Microsoft Entra Compliance Administrator role	Microsoft Entra Global Administrator role	Microsoft Purview Compliance Administrator role group	Roles or role groups that are view-only[1]	Roles or role groups that are view-only for AI data[2]	When not supported additional groups required
Complete one-click get started steps	✓	✓	✓	✕	✕	Not applic
View all get started steps	✓	✓	✓	✓	✕	Not applic
View key posture metrics, data snapshot, and posture trends chart on the Posture page	✓	✓	✓	✓	✕	Not applic
View Security Copilot insights and run prompts	✕	✕	✕	✕	✕	Data Secu Viewers
Complete action on getting started steps	✓	✓	✓ Excludes **Activate Audit**	✕	✕	Microsoft Exchange Compliand Managem Microsoft Exchange Records Managem Microsoft

Activities	Microsoft Entra Compliance Administrator role	Microsoft Entra Global Administrator role	Microsoft Purview Compliance Administrator role group	Roles or role groups that are view-only[1]	Roles or role groups that are view-only for AI data[2]	When not supported additional groups required
						Exchange Organizat Managem
View completion status of getting started steps	✓	✓	✓ Excludes status of **Activate Audit**	✓ Excludes: Status of **Activate Audit** Status of **Extend Your Insights**	✓ Excludes: Status of **Activate Audit** Status of **Extend Your Insights**	For **Activa Audit**: Microsoft Exchange Only Organizat Managem Microsoft Exchange Hygiene Managem Microsoft Exchange Compliand Managem Microsoft Exchange Records Managem Microsoft Exchange Organizat Managem For **Extend Your Insig** Microsoft Purview In Risk Managem Administra Microsoft Purview In Risk Managem Analyst Microsoft Purview In

Activities	Microsoft Entra Compliance Administrator role	Microsoft Entra Global Administrator role	Microsoft Purview Compliance Administrator role group	Roles or role groups that are view-only[1]	Roles or role groups that are view-only for AI data[2]	When not supported additional groups required
						Risk Managem Investigat
Connect to partner solutions in Microsoft Sentinel content hub	✗	✗	✗	✗	✗	Microsoft Sentinel Contribut
Create data security investigation	✗	✗	✗	✗	✗	Compliand Administra Data Secu Managem Insider Ris Managem Admin Organizat Managem Data Secu Investigat Admin Data Secu Investigat Investigat
View all recommendations	✓	✓	✓	✓	✓	Not applic
Complete actions on recommendation cards	✓	✓	✓	✗	✗	Not applic
View completion status of recommendation cards	✓	✓	✓	✓ Excludes **Unethical Behavior** card	✓ Excludes **Unethical Behavior** card	Communi Compliand Administra
View all activity and posture reports	✓	✓	✓	✓	✓	Not applic
View all policies in the policy list, **Reports** page	✓	✓	✓	✓ Excludes:	✓ Excludes:	For inside managem polices:

Activities	Microsoft Entra Compliance Administrator role	Microsoft Entra Global Administrator role	Microsoft Purview Compliance Administrator role group	Roles or role groups that are view-only[1]	Roles or role groups that are view-only for AI data[2]	When not supported additional groups required
				Insider risk management policies Communication compliance policies	Insider risk management policies Communication compliance policies	Microsoft Purview In Risk Managem Administra Microsoft Purview In Risk Managem Analyst Microsoft Purview In Risk Managem Investigat For communic compliand policies: Communi Compliand Administra
View all events in activity explorer, AI activities tab	✓ Excludes browse to URL (**AI Visit**) from insider risk management	✓ Excludes browse to URL (**AI Visit**) from insider risk management	✓ Excludes browse to URL (**AI Visit**) from insider risk management	✓ Excludes browse to URL (**AI Visit**) from insider risk management	✓ Excludes browse to URL (**AI Visit**) from insider risk management	Microsoft Purview In Risk Managem Analyst Microsoft Purview In Risk Managem Investigat
View all events in activity explorer, **All activity types** tab	✓ Excludes insider risk management events	✓ Excludes insider risk management events	✓ Excludes insider risk management events	✓ Excludes insider risk management events	✕	Microsoft Purview In Risk Managem Analyst Microsoft Purview In Risk

Activities	Microsoft Entra Compliance Administrator role	Microsoft Entra Global Administrator role	Microsoft Purview Compliance Administrator role group	Roles or role groups that are view-only[1]	Roles or role groups that are view-only for AI data[2]	When not supported additional groups required
						Managem Investigat
View user risk level of an individual user in all events from activity explorer	✗	✗	✗	✗	✗	Microsoft Purview In Risk Managem Analyst Microsoft Purview In Risk Managem Investigat
View user risk level of an individual user in all events from activity explorer, both tabs	✗	✗	✗	✗	✗	Microsoft Purview In Risk Managem Analyst Microsoft Purview In Risk Managem Investigat
View the prompts and responses within **AI Interaction** events from activity explorer	✗	✗	✗	✗	✗	Content Explorer Content V Microsoft Purview D Security A Content V
Create data risk assessments	✓	✓	✓	✗	✗	Not applic
View data risk assessments	✓	✓	✓	✓	✓	Not applic
View file details for data risk assessments	✗	✗	✗	✗	✗	Content Explorer Content V Content Explorer L Viewer

Activities	Microsoft Entra Compliance Administrator role	Microsoft Entra Global Administrator role	Microsoft Purview Compliance Administrator role group	Roles or role groups that are view-only[1]	Roles or role groups that are view-only for AI data[2]	When not supported additional groups required
View **Apps and agents** page	✓	✓	✓	✓	✓	Not applic
View objectives	✓	✓	✓	✓	✓	Not applic
View remediation plan for objectives	✓	✓	✓	✓	✕	Not applic
View remediation plan completion status	✓	✓	✓	✓	✕	Not applic
View risk patterns	✓	✓	✓	✓	✓	Not applic
View asset explorer	✓	✓	✓	✓	✕	Not applic
View file metadata in asset explorer	✕	✕	✕	✕	✕	Content Explorer Content V

[1] Includes Microsoft Purview Security Reader role group, Microsoft Purview Data Security AI Viewer role, the AI Administrator role from Entra, and the Data Security Viewers role group and Data Security Viewer role.

[2] The activities that display AI data security information, for view-only. For example, AI data from objectives such as exfiltration to AI apps and prevent oversharing, with their AI-related policies and metrics. Includes Microsoft Purview Data Security AI Viewer role and the AI Administrator role from Entra.

Custom role groups

Instead of granting access to Data Security Posture Management by using the built-in role groups, you can grant access by including the Microsoft Purview Compliance Administrator role in a custom role group. For read-only permissions, include the the Microsoft Purview Security Reader role, the Purview Data Security AI Viewer role, or the AI Administrator role from Entra.

If a custom role group includes the Microsoft Purview Compliance Administrator role, the user has the same access to Data Security Posture Management as the Microsoft Purview Compliance Administrator role group, except for the following:

- Create, view, update, and delete policies for insider risk management and communication compliance

If a custom role group includes the Microsoft Purview Security Reader role, the Purview Data Security AI Viewer role, or the AI Administrator role, the user has the same access to Data Security Posture Management the Microsoft Purview Security Reader role group, except for the following:

- View information protection policies

Last updated on 06/15/2026

Considerations for Microsoft Purview Data Security Posture Management

> *Microsoft Purview service description*

> **Note**
>
> This documentation is for the **current** version of Data Security Posture Management that expands coverage to more data sources, introduces guided workflows for proactive risk management, and streamlines data security operations so you can more confidently adopt AI across your digital estate.
>
> Most new features will be added to this version only but you can still access the previous versions and their documentation:
>
> - DSPM for AI (classic)
> - DSPM (classic)

For the most part, Microsoft Purview Data Security Posture Management is easy to use and self-explanatory, guiding you through setup and configuration tasks to help secure and govern your data. Use this section to complement that information and provide additional details that you might need.

Prerequisites for Data Security Posture Management

Use this section to identify prerequisites that aren't included in setup tasks.

- You have a supporting license for this Microsoft Purview solution, and your region is supported. For more information, see the Microsoft Purview service description.
- You have the right permissions.
- For Microsoft 365 Copilot and agents, users are assigned a license for Microsoft 365 Copilot.
  - For Copilot in Fabric and Security Copilot:
    - The enterprise version of Microsoft Purview data governance, to support the required APIs.

    - A collection policy, such as the one created from the setup task **Secure interactions in Microsoft Copilot experiences**.

- For monitoring interactions and applying DLP policies to other AI apps in Edge:
    - An Edge configuration policy is required to activate the Microsoft Purview integration in Edge. For configuration information, see Activate your DLP policy in Microsoft Edge.

- For Fabric data risk assessments, shared with DSPM for AI (classic): Configuration in both Entra and the Fabric admin portal.

- For Microsoft 365 custom data risk assessments and item-level scanning, shared with DSPM for AI (classic): Configuration in the Entra admin portal.

- For Entra-registered AI apps, they are integrated with the Microsoft Purview SDK to support all the Microsoft Purview capabilities.

- For Security Copilot promptbooks and agents that support data discovery and remediation for sensitive data, Enable Security Copilot for Purview Audit and DSPM Solutions.

Support for administrative units in Data Security Posture Management

When your tenant is configured for administrative units, restricted administrators who view policies, reports, activity explorer or asset explorer see data for users in their assigned administrative units only.

Because the one-click policies apply to all users, restricted administrators can't create these policies. You must be an unrestricted administrator to create one-click policies.

Activity explorer events in Data Security Posture Management

Use the following information to help you understand the events you might see in the activity explorer from Data Security Posture Management and the **AI activities** tab. References to a generative AI site can include Microsoft 365 Copilot, Microsoft 365 Copilot Chat, agents, other Microsoft copilots, and third-party AI sites.

Expand table

Event	Description
AI interaction	User interacted with a generative AI site. Details include the prompts and responses, except for unmanaged AI apps in Edge where text prompts only are included. For Microsoft 365 Copilot and Microsoft 365 Copilot Chat, this event requires auditing to be turned on. For Copilot in Fabric and Security Copilot, and for non-Copilot AI apps, prompts and responses require a collection policy with content capture selected to capture these interactions.
AI website visit	User browsed to a generative AI site.
DLP rule match	A data loss prevention rule was matched when a user interacted with a generative AI site. Includes DLP for Microsoft 365 Copilot.
Sensitive info types	Sensitive information types were found while a user interacted with a generative AI site. For Microsoft 365 Copilot and Microsoft 365 Copilot Chat, this event requires auditing to be turned on but doesn't require any active policies.

Known issues:

- The **AI interaction** event doesn't always display text for the prompt and response. Sometimes, the prompt and response spans consecutive entries. Other scenarios can include:
  - Microsoft Facilitator AI-generated notes, no prompt or response is displayed
  - When a user doesn't have a mailbox hosted in Exchange Online, no prompt or response is displayed
- The **Sensitive info types detected** event doesn't display the user risk level.
- For Microsoft Facilitator AI-generated notes, **AI interaction** events can't be linked to **Sensitive info types detected** events.
- For collection policies, no prompt or response is displayed if the option to capture content isn't selected in the policy. For example, the one-click policy **DSPM for AI - Detect sensitive info shared with AI via network** doesn't select this option when the policy is automatically created, but you can manually edit the policy and select this option after the policy is created.

Data updates paused for inactive tenants

To optimize resources,Microsoft pauses the background processing that refreshes Microsoft 365 data in Data Security Posture Management when a tenant hasn't accessed the solution for more than 60 days. While paused, you won't see any information for Microsoft 365 data in places such as in asset explorer, the data snapshot, and the data scanned posture metric.

Processing automatically resumes when you or another admin for your tenant returns to Data Security Posture Management. Data begins displaying again, although you can expect an initial delay. You might see an in-product message to indicate that updates were paused due to inactivity. This message disappears after processing resumes and no action is required from you to restart the updates.

Using the Data Security Posture Agent

The Data Security Posture Agent is currently in preview.

This agent acts as a quick, intent-based search mechanism for sensitive data. It's designed for pre-investigation checks rather than formal cases that still require other Microsoft Purview solutions, such as eDiscovery, audit, and Data Security Investigations. It uses natural-language search and doesn't rely on keywords, sensitive information types, or classifiers to search across files in SharePoint and OneDrive, messages in Teams, emails in Exchange, and interactions with Copilot.

The Data Security Posture Agent is exclusive to the current version of Data Security Posture Management and isn't available in Data Security Posture Management (classic). Before you can use it, you must first activate and configure it. For more information, see Get started with the Microsoft Purview Posture Agent in Data Security Posture Management.

Run searches from **DSPM** > **Discover** > **Asset explorer** > **Agent** tab. As with all Security Copilot agents, you can use this agent with preconfigured prompts and custom prompts. Before you run a custom prompt, you must first select **Add data sources**. In the flyout pane, enter at least 3 characters to search for users, groups, or sites. If you specify a user or group, the agent automatically selects the related mailbox, OneDrive account, or Sharepoint site. Keep in mind the following:

- For more efficient and faster processing, use a restricted search rather than tenant-wide searches.
- If you don't specify a time period for your prompt, search results are returned for the last 7 days.
- Search supports content up to 1 GB of data, but not metadata. For example, the agent can search all files containing financial information but not all files containing financial information and shared externally by a specific user in the last 7 days.

When a prompt is processed, you see a count of items matched against each item with any applied sensitivity labels and its risk level. The risk level is based on how closely related the match

of the content is from the prompt provided to the agent.

The **View Insights** option provides a report that includes a summary, risk assessment, findings, labeled or unlabeled items, and the top unlabeled items. You can download this report as a Word document to share with others for further analysis.

For auditing events as a result of running this agent, see Data Security Posture Agent activities from the auditing documentation.

Tips for custom Security Copilot prompts in Data Security Posture Management

For an enhanced experience with Security Copilot in Data Security Posture Management, use the following tips for higher accuracy for your Copilot responses:

- Always include the user's UPN for questions involving a specific user.
- Always specify the complete name for the sensitive info type or label for questions that involve a specific type of sensitive info type or label.
- Clearly list the sorting criteria for questions for top users, activities, and alerts.
- Always specify the date period for questions for data in a specific date period. If you don't specify a date period, only data from the last 10 days from current date is included. The maximum lookback is 30 days from the current date.
- Put all items (classifiers or labels) in single quotes in your prompt.
- Use "/" as a separator for any path (for example, a file path) in a user prompt.
- Scope the prompt to a single intent for higher accuracy of responses. Break complex prompts into single intent questions and enter the prompts one by one.
- Make questions self-contained. Avoid referring to previous questions or responses.
- Avoid using generic terms.
- Support prompts for data security across Information Protection, DLP, Insider Risk Management, or from public documentation.

For more information about creating Security Copilot prompts, see Create effective prompts.

Last updated on 05/01/2026

Application card: Microsoft Purview Data Security Posture Management

What is an Application or Platform Card?

Microsoft's Application and Platform cards are intended to help you understand how our AI technology works, the choices application owners can make that influence application performance and behavior, and the importance of considering the whole application, including the technology, the people, and the environment. Application cards are created for AI applications and platform cards are created for AI platform services. These resources can support the development or deployment of your own applications and can be shared with users or stakeholders impacted by them.

As part of its commitment to responsible AI, Microsoft adheres to six core principles: fairness, reliability and safety, privacy and security, inclusiveness, transparency, and accountability. These principles are embedded in the Responsible AI Standard, which guides teams in designing, building, and testing AI applications. Application and Platform Cards play a key role in operationalizing these principles by offering transparency around capabilities, intended uses, and limitations. For further insight, readers are encouraged to explore Microsoft's Responsible AI Transparency Report and Code of Conduct, which outline how enterprise customers and individuals can engage with AI responsibly.

Overview

Microsoft Purview Data Security Posture Management is a data security application within the Microsoft Purview portal that helps organizations discover, protect, and investigate sensitive data risks across their digital estate. Rather than focusing on infrastructure or endpoints, Data Security Posture Management centers on the data itself, identifying where it resides, who can access it, how it's used, and whether it's adequately protected. It uses AI to analyze access patterns, sharing behaviors, and policy gaps in order to surface actionable risks and recommend remediation steps.

Data Security Posture Management addresses a growing challenge in today's AI-driven workplaces: the increasing complexity and volume of data makes it difficult for security teams to maintain visibility and control. The application solves this problem by consolidating insights from multiple Microsoft Purview solutions, including data loss prevention (DLP), Insider Risk

Management, information protection with sensitivity labels, and Data Security Investigations, into a single view for monitoring data risks, policy coverage, and posture trends. This unified approach replaces the need for multiple tools and manual audits, helping administrators close data security gaps more efficiently.

Data Security Posture Management is designed to perform reliably in enterprise environments where organizations need to monitor and protect sensitive data across multiple locations. The application includes an embedded Microsoft Security Copilot experience that allows these users to ask natural language questions about their data security posture, as well as AI agents from Security Copilot that can take guided remediation actions on detected risks. For an overview introduction, see the Microsoft Mechanics video: New Data Security Posture Management.

Key terms

The following list provides a glossary of key terms related to Data Security Posture Management:

AI observability: A feature within Data Security Posture Management that provides an inventory of AI apps and agents with recent activity, showing how many are high risk and the total with sensitive interactions, along with a breakdown of individual agents and the policies governing them.

Collection policy: A policy that captures prompts and responses from AI interactions for Copilot in Fabric and Security Copilot, and for non-Copilot AI apps, so they can be managed in Microsoft Purview solutions.

Copilot interactions: Prompts and responses from Copilots and agents that Microsoft Purview supports for data security and compliance protections. Includes Microsoft 365 Copilot, Security Copilot, Copilot in Fabric, Copilot Studio.

Data loss prevention (DLP): A Microsoft Purview solution that helps prevent users from inappropriately sharing sensitive data by using content analysis techniques including keyword matching, expression evaluation, and machine learning algorithms.

Data risk assessment: An automated or custom scan that identifies and helps fix potential data oversharing risks, particularly for SharePoint sites and Fabric workspaces. Default assessments run weekly for the top 100 sites based on usage.

Data security objectives: Guided workflows within Data Security Posture Management that each represent a specific security goal, such as preventing data exposure in Copilot interactions or

preventing oversharing. Each objective groups together relevant Microsoft Purview solutions and prioritized actions.

Data Security Posture Agent: A Microsoft Security Copilot agent exclusive to the current version of Data Security Posture Management that uses natural language search across files in SharePoint, OneDrive, Teams, Exchange, and Copilot interactions. It is designed for pre-investigation checks rather than formal cases.

Insider Risk Management: A Microsoft Purview solution that uses built-in service and third-party indicators to help identify, triage, and act on potentially risky activity by users in an organization.

One-click policy: A preconfigured policy within Data Security Posture Management that can be activated with a single click to quickly gain insights and protect data, without requiring manual policy configuration.

Promptbook: A built-in sequence of prompts for Microsoft Security Copilot that helps users quickly investigate specific data security scenarios, such as risky user behavior or sensitive data protection.

Security Copilot: Microsoft Security Copilot and its agents is a generative AI-powered security solution that helps data security professionals investigate and respond to security incidents. Within Data Security Posture Management, it provides an open-ended prompt experience for asking natural language questions about data security.

Sensitivity label: A label from Microsoft Purview Information Protection that can be applied to content to define and enforce protection policies for sensitive data across clouds, apps, and devices.

Key features or capabilities

The key features and capabilities outlined here describe what Data Security Posture Management is designed to do and how it performs across supported tasks.

- **Data security posture dashboard**: The central landing page that provides immediate access to key posture metrics, top objectives to address based on risk, a snapshot of data usage across the data estate, and a 30-day trending graph of the organization's data security posture. Users can interact directly with Security Copilot through suggested prompts on this page.
- **Data security objectives with guided workflows**: Selectable cards that each represent a specific security goal, such as preventing data exposure in Copilot interactions, preventing

oversharing of sensitive data, preventing exfiltration to risky locations, and discovering sensitive data. Each objective provides an end-to-end workflow with prioritized actions, one-click policies, and progress tracking, so users can focus on achieving outcomes rather than navigating separate solutions.

- **AI observability**: An inventory of AI apps and agents with activity in the last 30 days, showing how many are high risk and the total with sensitive interactions. Provides a breakdown of individual agents, their activities, and the policies governing them, enabling administrators to monitor risks such as oversharing, exfiltration, and unusual access patterns.

- **Data risk assessments**: Automated and custom scans that identify potential data oversharing risks. Default assessments run weekly for the top 100 SharePoint sites based on usage. Custom assessments can target specific users, sites, or Fabric workspaces. Results include remediation options such as restricting access by label, creating auto-labeling policies, or creating retention policies.

- **Data security recommendations and remediation actions**: Insights and recommendations generated from processed data that help administrators create or refine DLP and Insider Risk Management policies. Includes one-click policies for quick deployment directly from the Data Security Posture Management workflow.

- **Reports and analytics**: Tracks the organization's data security posture over time with reports summarizing sensitivity label usage, DLP policy coverage, risky user behavior, and AI app activity. Enhanced reporting provides advanced filtering, customizable views, and export capabilities to support compliance requirements.

- **Activity explorer**: Provides detailed visibility into content-related activity, including AI interactions (prompts, responses, and sensitive information detected), DLP rule matches, AI website visits, and sensitive information type detections. Supports filtering by workload categories such as Copilot experiences, enterprise AI apps, and other AI apps.

- **Data Security Posture Agent**: An AI-powered agent from Microsoft Security Copilot and that uses natural language search to find sensitive data across SharePoint, OneDrive, Teams, Exchange, and Copilot interactions. It provides item counts, sensitivity label classifications, and risk-level assessments, along with exportable insight reports. This agent is designed for quick, pre-investigation checks.

- **Embedded Security Copilot experience**: An open-ended prompt experience where users can ask natural language questions about their data security posture. Includes built-in

promptbooks for risky user investigation and sensitive data protection, as well as a prompt gallery with categorized prompts for alerts, data at risk, risky users, suspicious activity, and sensitive data.

- **AI agent-driven remediation**: Under user guidance, AI agents can take direct action on detected risks, such as removing public sharing links, applying DLP policies, or revoking permissions. AI-driven triage agents review alerts from DLP and Insider Risk Management, filtering noise and highlighting critical threats. All agent actions are audited and subject to user review and approval.

Intended uses

Data Security Posture Management can be used in multiple scenarios across a variety of industries. Some examples of use cases include:

- **Monitoring data security posture across the enterprise**: A compliance officer at a financial services organization uses Data Security Posture Management to get a unified view of their sensitive data across different locations. The posture dashboard surfaces key metrics and trends, enabling the officer to quickly identify where unprotected sensitive data exists and track improvements over time. This eliminates the need to manually review multiple separate tools and dashboards.

- **Preventing oversharing before deploying Microsoft 365 Copilot**: An IT administrator preparing for a Microsoft 365 Copilot deployment runs data risk assessments to identify SharePoint sites with potentially overshared content. Data Security Posture Management surfaces sites where sensitive files have broad sharing links or lack sensitivity labels, and provides one-click remediation options such as restricting access or creating auto-labeling policies. This helps the organization reduce data exposure before AI tools begin processing that content.

- **Investigating risky user behavior with AI-assisted analysis**: A data security analyst receives an alert about a user performing unusual file-sharing activities. Using the embedded Security Copilot experience, the analyst runs the risky user investigation promptbook, which automatically analyzes the user's sensitive data activities, identifies potential exfiltration patterns, checks for anomalies, and suggests protective actions, all through a sequence of natural language prompts.

- **Governing AI app usage across the organization**: A security team at a healthcare organization uses AI observability to monitor which AI apps and agents employees are

interacting with, which of those interactions involve sensitive patient data, and whether appropriate DLP policies are in place. The team uses data security objectives to create targeted policies that prevent sensitive health information from being shared with unauthorized AI applications.

- **Responding to compliance requirements**: A compliance team at a government agency uses Data Security Posture Management reports and activity explorer to demonstrate that sensitive data is appropriately labeled, protected by DLP policies, and monitored for risky activity. The export capabilities and customizable views allow the team to generate audit-ready documentation aligned with their regulatory obligations.

- **Discovering unprotected sensitive data in new environments**: An organization just getting started with Microsoft Purview uses Data Security Posture Management to automatically scan data and user activities, gaining baseline insights and recommendations focused on unprotected data. This helps the organization quickly establish DLP, information protection, and Insider Risk Management policies without conducting deep manual analysis.

Data Security Posture Management is not intended for use as a general-purpose AI assistant, for scenarios unrelated to data security within Microsoft Purview, or for formal legal investigations (which require dedicated solutions such as eDiscovery and Data Security Investigations).

Models and training data

Data Security Posture Management leverages existing AI agents and AI functionality. As such, it's dependent on any external configuration that these provide. For example, model choice for Microsoft Security Copilot for the embedded natural language prompt experience. Data Security Posture Management also uses content analysis within Data Loss Prevention for detecting sensitive information through keyword matching, expression evaluation, and machine learning algorithms. To learn more, refer to the linked application cards and documentation.

Performance

Data Security Posture Management is designed to perform reliably in enterprise environments where organizations need to monitor and protect sensitive data across multiple locations. The application processes and correlates data from existing Microsoft Purview solutions, including DLP, Insider Risk Management, and information protection, to generate insights, recommendations, and posture metrics. Performance is optimized for organizations using supported Microsoft cloud services, and results improve as more Microsoft Purview solutions are configured and more data is available for analysis.

The application accepts text-based inputs across its features. Users interact with Data Security Posture Management through the Microsoft Purview portal interface, where they can navigate dashboards, configure policies, and run data risk assessments. The embedded Security Copilot experience accepts natural language text prompts and returns text-based responses, including data summaries, risk analysis, and recommended actions. The Data Security Posture Agent accepts natural language search queries and returns structured results, including item counts, sensitivity label classifications, and risk-level assessments, with an option to export reports as Word documents.

Data Security Posture Management supports the languages available in the Microsoft Purview portal interface. The Security Copilot experience is designed and evaluated primarily for English-language prompts and responses. Sensitive information type detection supports multiple languages as documented in the individual sensitive information type definitions. Users should be aware that performance of the natural language prompt and search features may vary when using languages other than English.

After initial setup, Data Security Posture Management requires time to process organizational data. Scanning times vary based on the size of the organization and the amount of data and activities to process, and it may take up to three days for initial processing to complete. New policies require at least 24 hours to collect data before results appear. Data risk assessments require at least 48 hours after completion before updated results are available. The Data Security Posture Agent can search up to 1 GB of content per query, and restricted searches targeting specific users or sites perform significantly faster than tenant-wide searches.

Limitations

Understanding Data Security Posture Management's limitations is crucial to determine it is used within safe and effective boundaries. While we encourage customers to leverage Data Security Posture Management in their innovative solutions or applications, it's important to note that Data Security Posture Management was not designed for every possible scenario. We encourage users to refer to either the Microsoft Enterprise AI Services Code of Conduct (for organizations) or the Code of Conduct section in the Microsoft Services Agreement (for individuals) as well as the following considerations when choosing a use case:

- **Data security scope**: The Security Copilot experience in Data Security Posture Management is designed to answer questions about data security associated with Insider Risk Management, Information Protection, and Data Loss Prevention in Microsoft Purview. All other Purview solutions are currently out of scope for Data Security Posture Management

insights. Users should direct questions outside of these areas to the appropriate solution-specific tools.

- **Data coverage**: The asset explorer Standard tab is filtered by Microsoft and non-Microsoft locations. Microsoft locations currently include Microsoft 365 only. Non-Microsoft locations are made possible by integration with partner solutions. Organizations with significant data in environments that aren't included should be aware that visibility through the asset explorer may be limited and should supplement with other tools as needed.

- **Data risk assessment volume limits**: A maximum of 200,000 items per location applies to both custom and default data risk assessments. The reported file count may not be accurate when there are more than 100,000 files per location. Organizations with very large SharePoint sites should consider running multiple targeted assessments. Additionally, OneDrive is not currently supported for item-level scanning, and custom assessments support a maximum of 10 SharePoint sites for item-level scanning.

- **Processing time requirements**: After a custom data risk assessment completes, results require at least 48 hours to become available and do not update after that point. A new assessment is needed to see changes. New policies require at least 24 hours before data appears. Users should plan ahead and avoid expecting real-time results from newly created assessments or policies.

- **Activity explorer data completeness**: The AI interaction event may not always display prompt and response text. In some cases, prompts and responses span consecutive entries. When a user does not have a mailbox hosted in Exchange Online, no prompt or response is displayed. Microsoft Facilitator AI-generated notes may not display prompts or responses. Users should be aware of these gaps when relying on activity explorer for compliance or investigative purposes.

- **Data Security Posture Agent constraints**: The agent supports content searches up to 1 GB of data but does not support metadata-based searches. If no time period is specified, results default to the last 7 days. The agent is designed for pre-investigation checks and is not a replacement for formal investigation tools such as eDiscovery, audit, or Data Security Investigations.

- **Language support**: The Security Copilot experience and the Data Security Posture Agent are designed and evaluated primarily for English-language prompts. Using other languages may result in reduced accuracy or incomplete responses. Users should exercise caution when operating outside the intended language scope.

- **Generative AI response accuracy**: As with all generative AI systems, Security Copilot responses in Data Security Posture Management may occasionally contain inaccurate or incomplete information. Users should verify AI-generated insights against actual data before taking consequential actions based on those responses.

Evaluations

Performance and safety evaluations assess whether AI applications are operating reliably and securely by examining factors like groundedness, relevance, and coherence while identifying the risks of generating harmful content. The following evaluations were conducted with safety components already in place, which are also described in Safety Components and Mitigations.

Data Security Posture Management was evaluated using custom evaluation methods developed by the product team. The evaluation focused on the accuracy of the embedded Security Copilot experience when responding to natural language questions about data security.

The evaluation data set consisted of relevant prompts identified by the product team and customers, including an expanded test set of prompts based on table schemas and descriptions. Microsoft evaluated performance using an accuracy rate metric: a response is considered accurate if the generated query provides the exact information asked in the prompt. The team tested across a range of data security scenarios covering DLP, information protection, and Insider Risk Management insights to verify that Security Copilot responses are factually supported by the underlying data and contextually appropriate to the user's question.

An ideal result is one where the Security Copilot response returns exactly the information requested, grounded in the organization's actual data, with no fabricated or misleading content. A suboptimal result would be a response that returns inaccurate data, misinterprets the user's question, or provides information outside the scope of the supported data security solutions. The evaluation process is iterative, with the product team refining the prompt sets and accuracy thresholds based on customer feedback and real-world usage patterns.

Safety components and mitigations

- **Role-based access control**: Data Security Posture Management enforces granular permissions through Microsoft Entra and Microsoft Purview role groups. Different activities require specific roles, such as Compliance Administrator for editing, Data Security Viewer for using Security Copilot, and Content Explorer Content Viewer for viewing AI interaction prompts and responses. This ensures that users can only access data and take actions appropriate to their role, reducing the risk of unauthorized access to sensitive information.

- **Comprehensive audit logging**: All automated actions taken by AI agents in Data Security Posture Management are recorded in audit logs. This includes actions such as removing sharing links, applying policies, and revoking permissions. Audit logs and activity explorer features provide a full trail of interactions with AI apps and agents, supporting compliance investigations, incident response, and accountability.

- **Human review and approval for agent actions**: Users always maintain control over AI agent behavior within Data Security Posture Management. AI-driven triage agents and the Data Security Posture Agent surface recommendations and findings, but users review, approve, or customize all automated actions before they take effect. The "View agent activity" options throughout the interface provide easy access to agent activity for oversight.

- **Opt-in analytics processing**: Data Security Posture Management requires explicit opt-in before processing organizational data. Analytics in both Insider Risk Management and DLP must be enabled before Data Security Posture Management can generate insights. This ensures that organizations make a deliberate decision to share data for analysis and are aware of the processing that takes place.

- **Scoped AI capabilities**: The Security Copilot experience in Data Security Posture Management is deliberately scoped to data security topics associated with Insider Risk Management, Information Protection, and Data Loss Prevention. Questions outside this scope are not answered, reducing the risk of the AI generating responses about topics where it lacks reliable data. The Data Security Posture Agent is similarly scoped to content searches within Microsoft 365 and limited to 1 GB of content per query.

- **Risk-level classification for search results**: The Data Security Posture Agent assigns a risk level to each search result based on how closely the content matches the user's prompt. Results are categorized to help users prioritize and focus on the most relevant findings. When the agent cannot determine the risk level, items are marked as "not categorized," providing transparency about the agent's confidence level.

- **Alert triage and noise reduction**: AI-driven triage agents review DLP and Insider Risk Management alerts and categorize them as "needs attention," "less urgent," or "not categorized." This filtering reduces alert fatigue and helps security analysts focus on genuine threats rather than false positives.

- **Privacy-preserving design**: Insider Risk Management, which feeds data into Data Security Posture Management, includes privacy controls such as pseudonymization and role-based access to ensure user-level privacy while enabling risk analysis. The Security Copilot

component follows the data privacy and security practices documented in the Microsoft Security Copilot privacy and data security documentation.

- **In-product feedback mechanism**: Users can provide feedback on each Security Copilot response by marking it as "looks right," "needs improvement," or "inappropriate." This feedback loop helps Microsoft continuously improve the quality and safety of AI-generated responses.

Best practices for deploying and adopting Data Security Posture Management

Responsible AI is a shared commitment between Microsoft and its customers. While Microsoft builds AI applications with safety, fairness, and transparency at the core, customers play a critical role in deploying and using these technologies responsibly within their own contexts. To support this partnership, we offer the following best practices for deployers and end users to help customers implement responsible AI effectively.

Deployers and end-users should:

- **Exercise caution and evaluate outcomes when using Data Security Posture Management for consequential decisions or in sensitive domains**: Consequential decisions are those that may have a legal or significant impact on a person's access to education, employment, financial platforms, government benefits, healthcare, housing, insurance, legal platforms, or that could result in physical, psychological, or financial harm. Sensitive domains—such as financial platforms, healthcare, and housing—require particular care due to the potential for disproportionate impact on different groups of people. When using AI for decisions in these areas, make sure that impacted stakeholders can understand how decisions are made, appeal decisions, and update any relevant input data.

- **Evaluate legal and regulatory considerations**: Customers need to evaluate potential specific legal and regulatory obligations when using any AI platforms and solutions, which may not be appropriate for use in every industry or scenario. Additionally, AI platforms or solutions are not designed for and may not be used in ways prohibited in applicable terms of service and relevant codes of conduct.

End-users should:

- **Exercise human oversight when appropriate**: Human oversight is an important safeguard when interacting with AI applications. While we continuously improve our AI applications, AI might still make mistakes. The outputs generated may be inaccurate, incomplete, biased,

misaligned, or irrelevant to your intended goals. This could happen due to various reasons, such as ambiguity in the inputs or limitations of the underlying models. As such, users should review the responses generated by Data Security Posture Management and verify that they match their expectations and requirements.

- **Be aware of the risk of overreliance**: Overreliance on AI happens when users accept incorrect or incomplete AI outputs, mainly because mistakes in AI outputs may be hard to detect. For the end-user, overreliance could result in decreased productivity, loss of trust, application abandonment, financial loss, psychological harm, physical harm, among others. In the context of Data Security Posture Management, overreliance could mean acting on an inaccurate Security Copilot insight without verifying it against actual data, potentially leading to misguided policy changes or missed security risks.

- **Scope Security Copilot questions to supported areas**: For the best results, ask questions about data security topics associated with Insider Risk Management, Information Protection, and Data Loss Prevention. Questions outside these areas may return incomplete or inaccurate responses because they fall outside the data Data Security Posture Management is designed to analyze.

- **Use specific and targeted prompts**: When using the Security Copilot experience or the Data Security Posture Agent, provide clear and specific prompts that include relevant details such as user names, time periods, and data locations. For example, instead of asking "show me security risks," ask "show all sensitive data activities performed by user@contoso.com in the last 30 days." More specific prompts produce more accurate and actionable results.

- **Provide feedback to improve quality**: Use the in-product feedback option to mark Security Copilot responses as "looks right," "needs improvement," or "inappropriate." This feedback helps Microsoft identify and address quality issues in AI-generated responses.

Deployers should:

- **Configure appropriate permissions before granting access**: Assign the minimum required roles to each user based on their responsibilities. Use the detailed permissions table in the Data Security Posture Management permissions documentation to determine which role groups are needed for each activity. Restricting access reduces the risk of unauthorized users viewing sensitive data or taking unintended actions.

- **Complete all setup tasks before relying on insights**: Ensure that auditing, analytics, and collection policies are properly configured before evaluating Data Security Posture Management insights. Incomplete setup can result in missing data and inaccurate posture

metrics. Allow sufficient processing time (up to three days for initial scans, at least 24 hours for new policies) before making decisions based on Data Security Posture Management data.

- **Use restricted searches for efficient processing**: When using the Data Security Posture Agent, configure searches to target specific users, groups, or sites rather than running tenant-wide scans. Restricted searches are significantly faster and more efficient, reducing processing time and producing more focused results.

- **Review and refine one-click policies after deployment**: One-click policies provide a useful starting point, but they should be reviewed and customized based on your organization's specific requirements. After initial deployment, monitor policy results through Data Security Posture Management reports and activity explorer, and adjust policy settings in the corresponding solution (DLP, Insider Risk Management, or Information Protection) to match your organization's risk tolerance and compliance needs.

- **Establish a regular review cadence**: Use Data Security Posture Management posture trends and reports to track your organization's data security posture over time. Establish a regular cadence for reviewing data risk assessments, policy coverage, and AI observability metrics. Regular reviews help identify emerging risks, policy gaps, and changes in user behavior before they become security incidents.

- **Plan for collection policy configuration**: If your organization needs to capture prompts and responses from Copilot in Fabric and Security Copilot, and for non-Copilot AI apps, configure collection policies with the content capture option selected. For apps that use these collection policies, AI interaction events in activity explorer will not display prompt and response text, limiting the usefulness of the data for compliance monitoring.

Learn more about Data Security Posture Management

For additional guidance or to learn more about the responsible use of Data Security Posture Management, we recommend reviewing the following documentation:

- Learn about Data Security Posture Management
- Data Security Posture Management (classic) - Overview of the previous version
- Setup tasks for Data Security Posture Management
- Considerations for Data Security Posture Management
- Permissions for Data Security Posture Management
- Prevent oversharing with data risk assessments

- Microsoft Purview data security and compliance protections for generative AI apps
- Data, privacy, and security for Microsoft Security Copilot
- Microsoft Purview service description

Learn more about responsible AI

- Microsoft AI principles
- Microsoft responsible AI resources
- Microsoft Azure Learning courses on responsible AI

ⓘ **Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 05/01/2026

List of AI sites supported by Microsoft Purview Data Security Posture Management (DSPM) and DSPM for AI (classic)

Summarize this article for me

> *Microsoft Purview service description*

This article lists the generative AI sites that are supported by Microsoft Purview Data Security Posture Management and Microsoft Purview Data Security Posture Management for AI (classic).

You can expect this list to increase over time.

Supported AI sites

```
*.10web.cloud
*.10web.io
*.10web.me
*.10zebra.com
*.20ma.cn
*.4b3.com
*.7aai.com
*.a0.dev
*.a1d.ai
*.a2e.ai
*.abridge.com
*.adcreative.ai
*.addy.ai
*.addy.so
*.adventureai.gg
*.agentgpt.reworkd.ai
*.agpt.co
*.ai.poweredtemplate.com
*.ai.yiios.com
*.ai2006.io
*.ai21.com
*.ai2sql.io
*.ai-blogkun.com
*.aibuddy.chat
*.aichat.gif99.cn
*.aichatbot.co
*.aichatting.net
*.ai-claude.net
```

```
*.aiclip.net
*.aicomicgenerator.net
*.aicontentfy.com
*.aicontentstudio.ai
*.aicopywriter.ai
*.aicotravel.com
*.aicoverletter.com
*.aicut.pro
*.aider.chat
*.aidesign.center
*.aidocmaker.com
*.aidungeon.com
*.aidungeon.io
*.ai-ebooks.co.uk
*.aiexcelbot.com
*.aifaceswap.io
*.aifree.site
*.aigca.org
*.aigcdeep.com
*.aigen.biz.id
*.ai-ghostwriter.com
*.aiheadshotgenerator.com
*.aihomedesign.com
*.aihumanize.com
*.aihumanize.io
*.aiimagegenerator.io
*.aiisajoke.com
*.ailabtools.com
*.ailessonplan.com
*.aimagicface.com
*.aimdoc.ai
*.aimreply.com
*.aimusic.so
*.aimusicfactory.ai
*.aipassportphoto.com
*.aipdf.ai
*.aipoemgenerator.org
*.ai-portraits.org
*.ai-q.in
*.aireception.io
*.airpost.ai
*.aiseo.ai
*.aisozai.com
*.aissistify.com
*.aistudios.com
*.aiswers.com
*.aiter.io
*.aitestkitchen.withgoogle.com
*.aitextpromptgenerator.com
*.aithor.com
*.aitubo.ai
*.aiuncensored.info
*.aiva.ai
*.ai-writer.com
*.aiwritingpal.com
```

```
*.aixtsy.com
*.ajelix.net
*.ajelix.org
*.akool.com
*.allenai.org
*.anagram.ai
*.analogenie.com
*.anchorbrowser.io
*.animaker.com
*.anthropic.com
*.anythingllm.com
*.anyword.com
*.aphra.me
*.app.aleph-alpha.com
*.app.alice.tech
*.app.instantdeco.ai
*.app.iokoo.co
*.app.koah.ai
*.app.longshot.ai
*.app.neuro-flash.com
*.app.questionwell.ai
*.app.submagic.co
*.app.suissechat.ch
*.app.surferseo.com
*.app.synthesia.io
*.appaca.ai
*.applaime.com
*.appointwise.io
*.aptlystar.ai
*.aragon.ai
*.archie.8base.com
*.architectgpt.io
*.archivinci.com
*.arria.com
*.artbreeder.com
*.artflow.ai
*.article.audio
*.articlefiesta.com
*.articleforge.com
*.articlereword.com
*.artiphoria.ai
*.askaichat.app
*.askan.ai
*.askbrian.ai
*.askgpt.app
*.askjack.io
*.askmany.cn
*.askmanyai.cn
*.asksocrates.app
*.askyourpdf.com
*.assignmentgpt.ai
*.atomai.fr
*.audiobox.metademolab.com
*.audiostack.ai
*.auraai.app
```

```
*.aurwy.com
*.autogenai.com
*.automateed.com
*.autoppt.com
*.autoreelapp.com
*.autoshorts.ai
*.avatarwithai.com
*.aws.amazon.com/bedrock/titan
*.aws.amazon.com/q
*.azure.microsoft.com/en-us/products/ai-services/openai-service
*.bagoodex.io
*.bannerbear.com
*.bard.google.com
*.basedlabs.ai
*.baz.co
*.bearly.ai
*.beatoven.ai
*.beautiful.ai
*.beehive.ai
*.beewriter.com
*.befunky.com
*.beingai.com
*.bertha.ai
*.beta.picjam.ai
*.betterpic.io
*.bettersynonyms.com
*.bewerbung.ai
*.bhuman.ai
*.bilmans.com
*.bing.com/chat
*.bizplanner.ai
*.blackink.ai
*.blinkvideo.ai
*.blogaudio.co
*.bluewillow.ai
*.bolt.new
*.bomml.ai
*.bookbud.ai
*.bookread.ai
*.bookwizardai.com
*.boomy.com
*.boostpixels.com
*.botowski.com
*.bots.miku.gg
*.botstar.com
*.brandmark.io
*.bravostudio.app
*.browsegpt.ai
*.buhuw.com
*.buildglare.com
*.bulkgpt.ai
*.business.adobe.com/products/sensei/adobe-sensei-genai.html
*.buster.ai
*.buzznews.today
*.bwana19.com
```

```
*.bypassgpt.ai
*.bytecap.io
*.c.ai
*.c5i.ai
*.caboodleai.net
*.camb.ai
*.cantina.com
*.capcut.com
*.captions.ai
*.caribbeans.ai
*.censusgpt.com
*.chadgpt.ru
*.chai-research.com
*.chanjing.cc
*.channel.io
*.character.ai
*.characternsfw.ai
*.charisma.ai
*.charley.ai
*.charshift.com
*.chat.bing.com
*.chat.com
*.chat.lmsys.org
*.chat.mymap.ai
*.chat.nbox.ai
*.chat.novaapp.ai
*.chat2db.ai
*.chatai.com
*.chataigpt.net
*.chat-ai-israel.com
*.chatba.com
*.chatbase.co
*.chatbit.co
*.chatbotgen.com
*.chatbotgf.com
*.chatboxapp.ai
*.chatcxc.com
*.chatdoc.com
*.chatespanolaigratis.com
*.chatfai.com
*.chatgate.ai
*.chatgbt.live
*.chatglm.cn
*.chatgot.io
*.chatgpt.at
*.chatgpt.ch
*.chatgpt.com
*.chatgpt.com.ua
*.chatgpt.es
*.chatgpt.fr
*.chatgpt.it
*.chatg-pt.org
*.chat-gpt.photos
*.chatgpt.pt
*.chatgpt4.org
```

```
*.chatgpt4google.com
*.chatgpt4online.org
*.chatgpt4rus.ru
*.chatgpt4youtube.com
*.chatgptbot.ru
*.chatgptchatapp.com
*.chatgptdemo.ai
*.chatgptdemo.net
*.chatgptduo.com
*.chatgptdutch.org
*.chatgptextension.ai
*.chatgpt-francais.com
*.chatgptfrancais.org
*.chatgptfree.ai
*.chatgpt-free.ai
*.chatgptgratuit.org
*.chatgpthuman.com
*.chatgptimagegenerator.org
*.chatgpt-israel.org
*.chat-gpt-israel.org
*.chatgptlogindeutsch.de
*.chatgptonline.io
*.chatgptonline.tech
*.chatgptpersian.com
*.chatgpt-quorum.com
*.chatgptspanish.org
*.chatgptsv.se
*.chatgptsvenska.io
*.chatgptwriter.ai
*.chatgptx.com
*.chatgpt-x.de
*.chatgptx.fr
*.chatgptx.ru
*.chatgtp.com
*.chathub.gg
*.chatmost.ru
*.chaton.ai
*.chatop.ai
*.chatopen.app
*.chatpdf.com
*.chatpdf.de
*.chatplayground.ai
*.chatroboten.se
*.chatsimple.ai
*.chattab.app
*.chattab.framer.website
*.chatthing.ai
*.chatuml.com
*.chatx.de
*.chatyy.xingmengsw.cn
*.chub.ai
*.cicube.io
*.civitai.com
*.claid.ai
*.claralabs.com
```

```
*.claude.ai
*.clausebuilder.ai
*.cleanup.pictures
*.cleoai.tech
*.clickable.so
*.clipchamp.com
*.clipdrop.co
*.cliprun.com
*.cmdj.app
*.coachvox.ai
*.code2.ai
*.codegate.ai
*.codesnippets.ai
*.cognition.ai
*.cohere.com
*.cohesive.so
*.coinvent.ai
*.coloringninja.com
*.colormind.io
*.combopix.com
*.comfy.nordy.ai
*.comicsmaker.ai
*.compose.ai
*.consensus.app
*.consistentcharacter.ai
*.contentbot.ai
*.contentforge.ai
*.contentpie.ai
*.contentvillain.com
*.convergence.ai
*.conversionblitz.com
*.copilot.microsoft.com
*.copy.ai
*.copymatic.ai
*.copymonkey.ai
*.copysmith.ai
*.copyter.com
*.copytruck.com
*.copywriting.ai
*.corina.ai
*.cotypist.app
*.coursebox.ai
*.coursepro.ai
*.coverler.com
*.covric.com
*.craftly.ai
*.craiyon.com
*.crammer.app
*.crankoutai.com
*.cre8tiveai.com
*.creaitor.ai
*.createaicourse.com
*.created.app
*.creatify.ai
*.creator.nightcafe.studio
```

```
*.cronbot.ai
*.csm.ai
*.ctrlsheet.ai
*.cubox.cc
*.cursor.com
*.cutout.pro
*.d5next.ai
*.daftart.ai
*.dall-efree.com
*.dante-ai.com
*.davinci.ai
*.dazi.co
*.decidable.com
*.decktopus.com
*.decopy.ai
*.deemos.com
*.deepai.org
*.deepbrain.io
*.deepbrainai.io
*.deepcode.ai
*.deepdreamgenerator.com
*.deepfakesweb.com
*.deep-image.ai
*.deepreview.eu
*.deep-seek.chat
*.deepseek.com
*.deepseek.net
*.deepseek-ai.ru
*.deepseekr1.jfsmaa.com
*.deepseekr1.org
*.deepseekv3.com
*.deepswap.ai
*.deevid.ai
*.describely.ai
*.descrii.tech
*.descript.com
*.designer.microsoft.com
*.designs.ai
*.devin.ai
*.devinapps.com
*.dezgo.com
*.differentdimensionme.net
*.diffusionart.co
*.docgpt.ai
*.docky.ai
*.docuhelp.ai
*.domoai.app
*.donakosy.com
*.dopeloop.ai
*.dopple.ai
*.dorkgpt.com
*.doubao.com
*.draftly.so
*.dream.ai
*.dreamhairai.com
```

```
*.dreamily.ai
*.dreampress.ai
*.dreamstudio.ai
*.dreamtavern.ai
*.dubbing-ai.com
*.dubbingai.io
*.dubverse.ai
*.dxm.com
*.dzine.ai
*.east.im
*.easymusic.ai
*.easy-peasy.ai
*.echo.win
*.echt-lu.de
*.editapp.ai
*.editee.com
*.editgpt.app
*.edit-videos-online.com
*.eduwiz.ai
*.edwardbot.com
*.effectssdk.ai
*.egger.ai
*.eightify.app
*.eilla.ai
*.elai.io
*.elephas.app
*.eleuther.ai
*.elevenlabs.io
*.elicit.com
*.engagely.ai
*.essai.pro
*.essayailab.com
*.essay-builder.ai
*.essaygrader.ai
*.essaypal.ai
*.ev5.ai
*.excelmaster.ai
*.explainpaper.com
*.faceswapper.ai
*.facetuneapp.com
*.fadr.com
*.faible.ai
*.fakeyou.com
*.falconllm.tii.ae
*.fanfuel.app
*.fantasy.ai
*.fashionphoto.ai
*.fermat.app
*.figflow.io
*.finechat.ai
*.finepixel.lusion.ai
*.fineshare.com
*.finito.ai
*.fireflies.ai
*.firefly.adobe.com
```

```
*.firetexts.co
*.fitnessai.com
*.flair.ai
*.flatai.org
*.flexclip.com
*.fliki.ai
*.flowgpt.com
*.flowrite.com
*.flux1.ai
*.fluxpro.ai
*.fomo.ai
*.fontjoy.com
*.forethought.ai
*.formulabot.com
*.formwise.ai
*.fotor.com
*.foxyai.ai
*.frase.io
*.freedomgpt.com
*.freeimagegenerator.com
*.freepass.ai
*.fuelix.ai
*.fullstackdata.com
*.fusionartai.com
*.gajix.com
*.galichat.com
*.gamma.app
*.gamma.work
*.ganstudio.com
*.gemini.google.com
*.genai-go.com
*.genaizurich.ch
*.gencraft.com
*.genei.io
*.generated.photos
*.generatelyrics.io
*.generativeai.pub
*.generatorxyz.com
*.genesismedical.ai
*.gerwin.app
*.get-ai.world
*.getaprototype.com
*.getchunky.io
*.getconch.ai
*.getcoralai.com
*.getcube.one
*.getgptapi.com
*.getimg.ai
*.getliner.com
*.getmerlin.in
*.getpoppy.ai
*.getsmartgpt.com
*.gettrove.co
*.getvoila.ai
*.getwoord.com
```

```
*.ghibliai.io
*.giga.chat
*.gingerkeyboard.com
*.gista.co
*.github.com/features/copilot
*.giti.ai
*.giz.ai
*.gizzmo.ai
*.glasp.co
*.glass-imaging.com
*.glif.app
*.gliglish.com
*.godinabox.co
*.golf.ai
*.goneural.ai
*.goodwill.ribbit.works
*.gozen.io
*.gpt.h2o.ai
*.gpt3demo.com
*.gpt4all.io
*.gpt4free.io
*.gpt6.ai
*.gptassistant.app
*.gptexcel.uk
*.gptfy.co
*.gptgame.app
*.gptgo.ai
*.gptinf.com
*.gptkit.ai
*.gptnederlands.nl
*.gptonline.ai
*.gpt-persona.com
*.gpt-ppt.neftup.app
*.gpttranslator.co
*.gptunnel.ru
*.gptzero.me
*.grammarly.ai
*.grammarly.com
*.grammarly.net
*.grammatikai.com
*.grock.com
*.grok.com
*.grok.x.ai
*.grok-2.ai
*.guidde.com
*.hailuoai.com
*.hailuoai.video
*.hal9.com
*.hap.ai
*.harpa.ai
*.harvey.ai
*.headlime.com
*.headshotaigenerator.com
*.headshotpro.com
*.hebbia.com
```

```
*.hedra.com
*.hei.io
*.heimdallapp.org
*.hellohistory.ai
*.heuristi.ca
*.heyboss.xyz
*.heyfriday.ai
*.heygen.com
*.heygpt.chat
*.higgsfield.ai
*.high-eleven.com
*.hinative.com
*.hippocraticai.com
*.hiretruffle.com
*.hitalk.ai
*.hoppycopy.co
*.hotpot.ai
*.hourone.ai
*.huatuogpt.cn
*.huggingface.co/spaces/tiiuae/falcon-180b-demo
*.humanizeai.io
*.humanizeai.pro
*.humanpal.io
*.humata.ai
*.hypotenuse.ai
*.ichatwithgpt.com
*.icon.me
*.icongenerate.com
*.id.freepik.com
*.ideabuddy.com
*.ideapicker.io
*.ideasai.com
*.ideiaschatgpt.com.br
*.ideogram.ai
*.illuminate.google.com
*.illustroke.com
*.image-to-caption.com
*.image-upscaling.net
*.imaginario.ai
*.imagine.art
*.inbox-ai.app
*.infinitecraft.ai
*.infinity.ai
*.inkforall.com
*.inputai.com
*.insmind.com
*.instantanswers.xyz
*.instatext.io
*.interiorai.com
*.interview.lol
*.invideo.io
*.iris.ai
*.ivrit.ai
*.ixcoach.com
*.jammable.com
```

```
*.janusai.pro
*.jasper.ai
*.jellypod.ai
*.jiangziai.com
*.jocondeai.com
*.jogg.ai
*.kafkai.com
*.kaiber.ai
*.kapwing.com
*.kb-whisper.mesu.re
*.kerlig.com
*.kibo.vercel.app
*.kickresume.com
*.kittensai.com
*.kittl.com
*.klingai.com
*.koala.sh
*.kodus.io
*.kokorottsai.com
*.komo.ai
*.korbit.ai
*.krater.ai
*.krea.ai
*.kudoway.eu
*.lalal.ai
*.lambdatest.com
*.lamini.ai
*.landbot.io
*.langchain.com
*.laragpt.com
*.learn.xyz
*.learnitive.com
*.learnt.ai
*.letsenhance.io
*.letstrip.co.il
*.letz.ai
*.lexalytics.com
*.lexica.art
*.ligo.bio
*.linke.ai
*.linkfox.com
*.llama.com
*.llama.meta.com
*.llama2.ai
*.llamachat.app
*.logoai.com
*.lolgames.net
*.loudly.com
*.loudme.ai
*.lovo.ai
*.lp.ai-copywriter.jp
*.lp.comic-copilot.ai
*.lucie.chat
*.luis.ai
*.lumen5.com
```

```
*.lummi.ai
*.machinetranslation.com
*.maester.app
*.maestra.ai
*.mage.space
*.magicregexgenerator.xyz
*.magicslides.app
*.magicstudio.com
*.magisto.com
*.magnific.ai
*.mailshake.com/ai-email-writer
*.makeform.ai
*.mammouth.ai
*.mapify.so
*.markcopy.ai
*.marketingblocks.ai
*.marketingcopy.ai
*.masterpiecestudio.com
*.mebook.ai
*.meetmaya.world
*.meetz.ai
*.memfree.me
*.memorylane.co
*.mentara.io
*.merlin.foyer.work
*.meta.ai
*.micvoice.ai
*.midjourney.co
*.mieux.ai
*.mightygpt.com
*.mimicpc.com
*.mind.am
*.mindmapai.app
*.mintoai.com
*.miramuseai.net
*.modulify.ai
*.moonvalley.com
*.mosaicml.com
*.motiff.cn
*.mubert.com
*.mureka.ai
*.murf.ai
*.muse.fun
*.musely.ai
*.musicful.ai
*.myaiteam.com
*.myaskai.com
*.mygptwizard.com
*.mykin.ai
*.myninja.ai
*.mytales.io
*.myths-of-gallantium.pressdeck.io
*.namesnack.com
*.nando.ai
*.napkin.ai
```

```
*.narakeet.com
*.narrationbox.com
*.nat.dev
*.naturalreaders.com
*.netcrumb.com
*.netus.ai
*.neural.love
*.neuraltext.com
*.newswriter.ai
*.nextbrain.ai
*.noluai.com
*.notebooklm.cloud.google.com
*.notion.so
*.novelai.net
*.noxus.cc
*.nsfwai.pro
*.numind.ai
*.ocoya.com
*.octocom.ai
*.office-ai.cn
*.ohai.ai
*.ohmygpt.com
*.oji.ai
*.ollama.ai
*.olympia.chat
*.omagic.ai
*.omneky.com
*.openai.cha-tai.cn
*.openai.com
*.openart.ai
*.openevidence.com
*.opus.pro
*.opusclip.com
*.ora.ai
*.ordermaid.ai
*.ossa.ai
*.outerbase.com
*.outwrite.com
*.pagelines.com
*.palmate.ai
*.paperpal.com
*.paragraphai.com
*.parallelgpt.ai
*.paraphrasingtool.ai
*.partyrock.aws
*.pdf.wondershare.com
*.pencilsketchgenerator.com
*.peoplealsoask.ai
*.peppercontent.io
*.perfectessaywriter.ai
*.perfecttense.com
*.perplexity.ai
*.personal.ai
*.petal.org
*.pfpmaker.com
```

```
*.phaii.net
*.phedra.ai
*.phind.com
*.photoaistudio.com
*.photoleapapp.com
*.photomath.com
*.photomyne.com
*.photopacks.ai
*.phrasee.co
*.pica-ai.com
*.piccopilot.com
*.pickaxeproject.com
*.piclumen.com
*.picmaker.com
*.pictory.ai
*.pika.art
*.pineapplebuilder.com
*.ping.tm
*.piufoto.com
*.pixai.art
*.pixelcut.ai
*.pixlr.com
*.pixu.ai
*.play.ai
*.play.ht
*.playgroundai.com
*.plusai.com
*.podcast.adobe.com
*.podcastle.ai
*.podnotes.app
*.podpally.com
*.poe.com
*.pollo.ai
*.polybuzz.ai
*.polylm.com
*.popai.pro
*.portraitpal.ai
*.postwise.ai
*.predis.ai
*.premai.io
*.prepai.io
*.preppally.com
*.presentationgpt.com
*.presentations.ai
*.prezo.ai
*.print-ai-masterpiece.com
*.privatellm.app
*.proai.co
*.productai.photo
*.profilepicture.ai
*.projectdecember.net
*.promeai.pro
*.prometai.app
*.promptclub.ai
*.promptfolder.com
```

```
*.prompthero.com
*.prompthunt.com
*.promptitude.io
*.promptkit.tools
*.promptpictures.tv
*.prophetes.ai
*.punchlines.ai
*.pyxl.pro
*.qianlu.cc
*.qopywriter.ai
*.questmate.com
*.quickchat.ai/emerson
*.quicktok.ai
*.quillbot.com
*.quizme.live
*.qwen.ai
*.qwenlm.ai
*.raita-kun.arvo.net
*.randomgenerate.io
*.rask.ai
*.rawshorts.com
*.read.ai
*.read2me.online
*.rebecc.ai
*.refraction.dev
*.regem.in/ai-writer
*.regex.ai
*.regie.ai
*.regisai.com
*.reimaginehome.ai
*.relevanceai.com
*.remini.ai
*.rencoach.com
*.rendair.ai
*.rendernet.ai
*.rendora.ai
*.replicate.com
*.replika.com
*.replit.com
*.resemble.ai
*.retellio.com
*.retresco.de
*.revid.ai
*.revoicer.com
*.riku.ai
*.rizzai.com
*.roamaround.app
*.roomdesigner.ai
*.room-genius.com
*.rootly.com
*.rovioai.com
*.rtst.ai
*.ru-chatgpt.ru
*.runwayml.com
*.ryrob.com/paragraph-rewriter
```

```
*.rytr.me
*.saffronintell.com
*.saga.so
*.sakubun.ai
*.salesforge.ai
*.samwell.ai
*.sapling.ai
*.sassbook.com
*.saufter.io
*.scholarcy.com
*.scholarlyai.com
*.scifocus.ai
*.scisummary.com
*.scite.ai
*.scribblediffusion.com
*.scribbyo.com
*.sdxlturbo.ai
*.seaart.ai
*.seaart.me
*.search.hebbia.ai
*.searchgpt.net
*.segment-anything.com
*.segmind.com
*.sensibleaimanifesto.com
*.seowriting.ai
*.serif.ai
*.shakespearetoolbar.com
*.shakker.ai
*.sheetai.app
*.sheetgpt.ai
*.shopia.ai
*.shortlistiq.com
*.shortlyai.com
*.shyuhuankj.com
*.sieve.video
*.sievedata.com
*.simbla.com
*.simpleshow.com
*.simpliciter.ai
*.simplified.com
*.sindresorhus.com/aiko
*.sindresorhus.com/amazing-ai
*.sindresorhus.com/quickgpt
*.sinonimos.com.br
*.sintra.ai
*.site.xaris.ai
*.sitegpt.ai
*.skybox.blockadelabs.com
*.skyengine.ai
*.skylum.com
*.skyvern.com
*.slazzer.com
*.slider-ai.ru
*.slidesgpt.com
*.smart-copy.io
```

```
*.smartlifeskills.ai
*.smartling.com
*.smartsolve.ai
*.smartwriter.ai
*.snapheadshots.com
*.snowpixel.app
*.sofer.ai
*.sologo.ai
*.soloist.ai
*.sonix.ai
*.sonofa.ai
*.soofy.io
*.soulgen.net
*.soundful.com
*.soundraw.io
*.sourcetable.com
*.soynerdito.com
*.speechgen.io
*.speechify.com
*.speechmatics.com
*.spicychat.ai
*.spikes.studio
*.splice.com
*.stability.ai
*.stableaudio.com
*.stablediffusion.com
*.stablediffusionweb.com
*.staccato.ai
*.starbyface.com
*.starryai.com
*.steve.ai
*.stockimg.ai
*.stork.ai
*.storyblocker.com
*.storybrand.ai
*.storyd.ai
*.storynest.ai
*.storyscapeai.app
*.storytailor.ai
*.streamlit.io/generative-ai
*.strella.io
*.studio.tune.app
*.studytexter.com
*.studytexter.de
*.styleadvice.ai
*.subscribr.ai
*.sudowrite.com
*.summarizer.org
*.suno.ai
*.suno.com
*.suno-ai.org
*.supafriends.com
*.supercraft.ai
*.supermeme.ai
*.svg.io
```

```
*.swiftspec.ai
*.switchboard.ai
*.sydologie.ai
*.syllaby.io
*.symbolab.com
*.synesthesia.io
*.synthesys.io
*.t3.chat
*.tabnine.com
*.talkai.info
*.talkie-ai.com
*.talkpal.ai
*.talkstack.ai
*.talktext.io
*.talktowalle.com
*.tana.inc
*.taplio.com
*.teachablemachine.withgoogle.com
*.tealhq.com
*.team-gpt.com
*.teammately.ai
*.telborg.com
*.template.net
*.tensorart.green
*.tethered.dev
*.texta.ai
*.textadviser.com
*.textalky.com
*.textcortex.com
*.texteo.ai
*.textie.ai
*.text-summary.com
*.textsynth.com
*.textwise.qutesoft.com
*.texxto.ai
*.theogrowth.com
*.there.do
*.thesummary.ai
*.thirdai.com/pocketllm
*.thispersonnotexist.org
*.thisresumedoesnotexist.com
*.threadcreator.com
*.thunderbit.com
*.thundercontent.com
*.tinyalbert.ai
*.tinywow.com
*.tldrthis.com
*.tome.app
*.tool.tool123.fun
*.toolhouse.ai
*.tools.picsart.com
*.toolsaday.com
*.topazlabs.com
*.to-teach.ai
*.transfercompass.com
```

```
*.trint.com
*.tryblend.ai
*.trypencil.com
*.trypromptly.com
*.ttsmaker.com
*.tugan.ai
*.tungsten.run
*.turbolearn.ai
*.turbular.com
*.tutorai.me
*.tweetyai.com
*.twoslash.ai
*.txtplay.ai
*.typecast.ai
*.typeface.ai
*.typeright.com
*.typli.ai
*.uberduck.ai
*.udio.com
*.uimagic.io
*.uizard.io
*.uminal.com
*.unbounce.com
*.uncode.vercel.app
*.undatas.io
*.undress.vip
*.unifygpt.ai
*.uniglobalcareers.com
*.unsloth.ai
*.upscale.media
*.upscayl.org
*.useadrenaline.com
*.usechat.ai
*.usegalileo.ai
*.usemano.com
*.usemotion.com
*.v0.dev
*.vanceai.com
*.vectorize.io
*.vectorizer.ai
*.veed.io
*.venice.ai
*.vertexaisearch.cloud.google
*.vgu.community
*.vidby.com
*.videoblog.ai
*.videocandy.com
*.videomuse.app
*.vidext.app
*.vidu.studio
*.vidyard.com
*.vidyo.ai
*.viggle.ai
*.vinteo.ai
*.virtualghostwriter.com
```

```
*.visualelectric.com
*.vivago.ai
*.vizcom.ai
*.vizly.fyi
*.vmfi.net
*.vobox.io
*.vocalid.ai
*.vogent.ai
*.voicemaker.in
*.voicemod.net
*.voicenotes.com
*.voiceovermaker.io
*.voyageai.com
*.vozfly.com
*.walles.ai
*.wan.video
*.wand.ai
*.warmer.ai
*.warpvideo.ai
*.watermarkremover.io
*.watermelon.ai
*.wavoai.com
*.webllm.ai
*.webllm.mlc.ai
*.websim.ai
*.webspellchecker.com
*.webtastic.ai
*.weel.co.jp
*.wellsaidlabs.com
*.wepik.com
*.we-spots.com
*.whatsgoodapps.com
*.whispertranscribe.com
*.wireply.ai
*.wishesai.com
*.wisprflow.ai
*.witsyai.com
*.wnr.ai
*.wolframalpha.com
*.wordai.com
*.wordhero.co
*.wordplay.ai
*.word-spinner.com
*.wordtune.com
*.wordybot.ai
*.workflos.ai
*.woxo.tech
*.wpaibot.com
*.wpkubio.com
*.writecream.com
*.writefull.com
*.writegpt.ai
*.writeholo.com
*.writeme.ai
*.writer.com
```

```
*.writersbrew.app
*.writerx.co
*.writesonic.com
*.writesparkle.ai
*.writier.io
*.wrizzle.ai
*.wui.ai
*.www.getlid.co
*.xccelerateai.com
*.xerac.ir
*.xiaoin.com.cn
*.xiaoshuoai.com
*.xinva.ai
*.xlscout.ai
*.xn--e1aalohlij4g.net
*.yaara.ai
*.yarnit.app
*.yeschat.ai
*.yeting.fun
*.yiyan.baidu.com
*.yodayo.com
*.youchuan.cn
*.yourjoi.com
*.zevbot.com
*.zipwp.dev
*.zoc.ai
*.zomani.ai
*.zomgsounds.com
*.zonerai.com
```

Last updated on 02/27/2026

Learn about Data Security Posture Management for AI - (classic)

> *Microsoft 365 licensing guidance for security & compliance*

> **Note**
>
> This article is for the **classic** version of Data Security Posture Management that's being replaced with a new version that introduces guided workflows for proactive risk management, and streamlines data security operations so you can more confidently adopt AI across your digital estate.
>
> These improvements won't be added to this classic version so we invite you to try the new Data Security Posture Management, currently in preview.

Microsoft Purview Data Security Posture Management (DSPM) for AI from the Microsoft Purview portal provides a central management location to help you quickly secure data for AI apps and proactively monitor AI use. These apps include Copilots, agents, and other AI apps that use third-party large language modules (LLMs).

Data Security Posture Management for AI offers a set of capabilities so you can safely adopt AI without having to choose between productivity and protection:

- Insights and analytics into AI activity in your organization
- Ready-to-use policies to protect data and prevent data loss in AI prompts
- Data risk assessments to identify, remediate, and monitor potential oversharing of data.
- Compliance controls to apply optimal data handling and storing policies

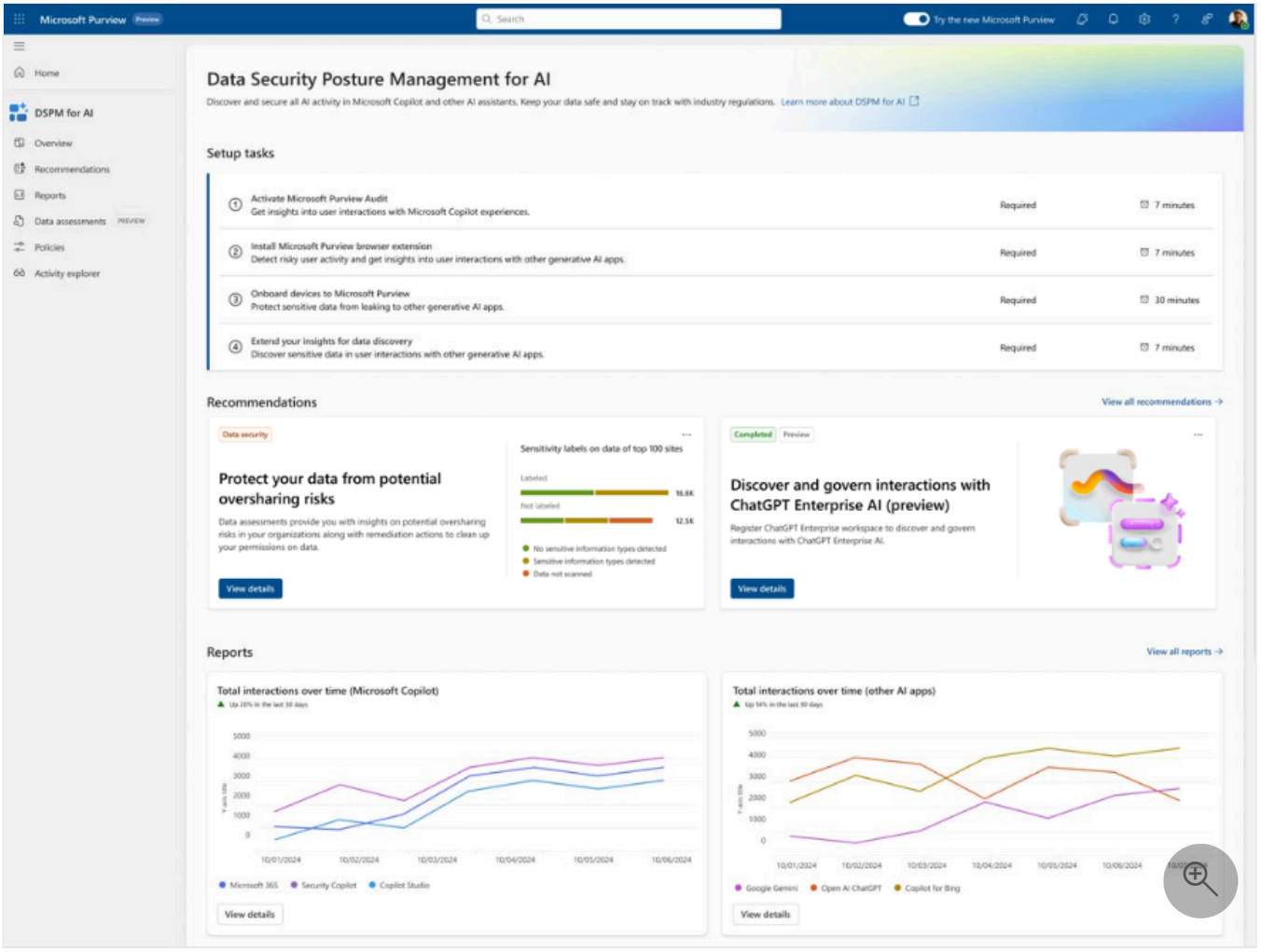
https://learn-video.azurefd.net/vod/player?id=e3504512-f69e-455e-847c-798d92839d8a&locale=en-us&embedUrl=%2Fpurview%2Fdspm-for-ai

For a list of supported third-party AI sites, such as those used for Gemini and ChatGPT, see Supported AI sites by Microsoft Purview for data security and compliance protections.

How to use Data Security Posture Management for AI

To help you more quickly gain insights into AI usage and protect your data, Data Security Posture Management for AI provides some recommended preconfigured policies that you can activate with a single click. If you're using administrative units, you must be an unrestricted administrator to create these policies that apply to all users in your organization.

Allow at least 24 hours for these new policies to collect data to display the results in Data Security Posture Management for AI, or reflect any changes that you make to the default settings.



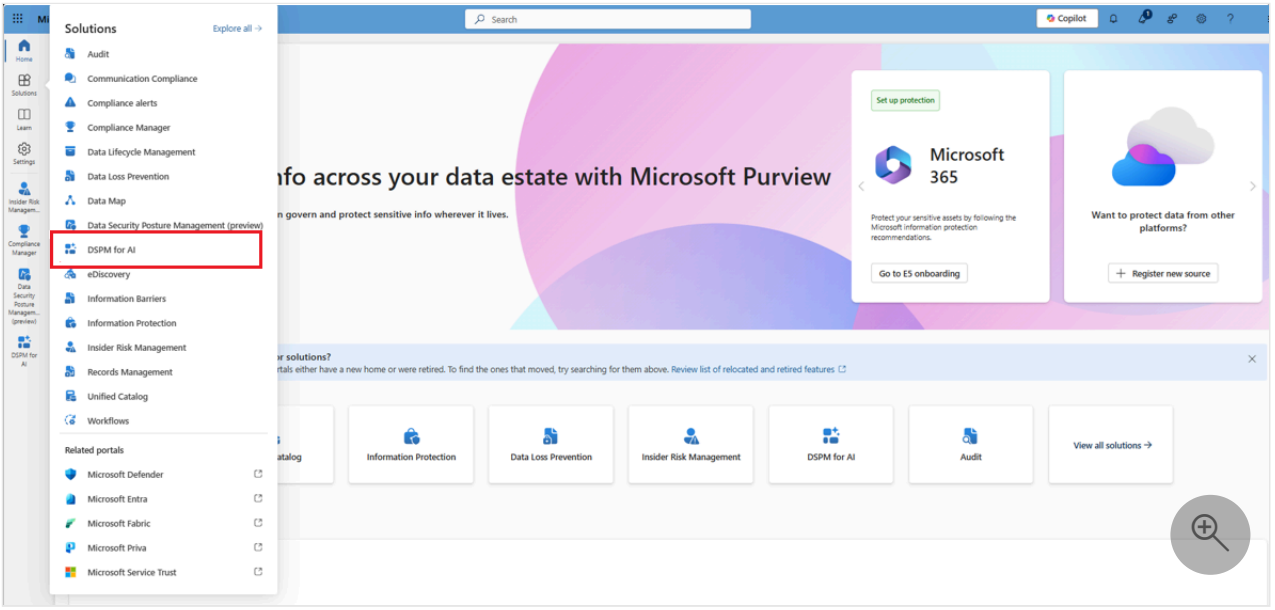
No activation needed, Data Security Posture Management for AI automatically runs a weekly data risk assessment for the top 100 SharePoint sites based on usage in your organization. You can supplement this with your own custom data risk assessments, currently in preview. These assessments are designed specifically to help you identify, remediate, and monitor potential oversharing of data, so you can be more confident about your organization using Microsoft 365 Copilot and agents, and Microsoft 365 Copilot Chat.

https://learn-video.azurefd.net/vod/player?id=d68ad746-d47d-4e62-b5a5-de8dac5c940e&locale=en-us&embedUrl=%2Fpurview%2Fdspm-for-ai

To get started with Data Security Posture Management for AI, use the Microsoft Purview portal. You need an account that has appropriate permissions for compliance management, such as an account that's a member of the **Microsoft Entra Compliance Administrator** group role.

> **Tip**
>
> The following steps provide a recommended walkthrough to use all the available capabilities from DSPM for AI. If you're interested in using DSPM for AI just for a specific AI app, use the links in the **Next steps** section.

1. Sign in to the Microsoft Purview portal > **Solutions** > **DSPM for AI (classic)**.



2. From **Overview**, keep the **All AI apps** view selected. Review the **Get started** section to learn more about Data Security Posture Management for AI, and the immediate actions you can take. Select each one to display the flyout pane to learn more, take actions, and verify your current status.

Expand table

Action	More information
Turn on Microsoft Purview Audit	Auditing is on by default for new tenants, so you might already meet this prerequisite. If you do, and users are already assigned licenses for Microsoft 365 Copilot, you start to see insights about Copilot and agents activities from the **Reports** section further down the page.
Install Microsoft Purview browser extension	A prerequisite for third-party AI sites.

Action	More information
Onboard devices to Microsoft Purview	Also a prerequisite for third-party AI sites.
Extend your insights for data discovery	One-click policies for collecting information about users visiting third-party generative AI sites and sending sensitive information to them. The option is the same as the **Extend your insights** button in the AI data analytics section further down the page.

For more information about the prerequisites, see Prerequisites for Data Security Posture Management for AI.

For more information about the preconfigured policies that you can activate, see One-click policies from Data Security Posture Management for AI.

3. Then, review the **Recommendations** section and decide whether to implement any options that are relevant to your tenant. View each recommendation to understand how they're relevant to your data and learn more.

   These options include running a data risk assessment across SharePoint sites, creating sensitivity labels and policies to protect your data, and creating some default policies to immediately help you detect and protect sensitive data sent to generative AI sites. Examples of recommendations:

   - **Protect your data from potential oversharing risks** by viewing the results of your default data risk assessment to identify and fix issues to help you more confidently deploy Microsoft 365 Copilot.
   - **Protect your data with sensitivity labels** by creating a set of default sensitivity labels if you don't yet have sensitivity labels, and policies to publish them and automatically label documents and emails.
   - **Protect items with sensitivity labels from Microsoft 365 Copilot and agent processing** by prompting you to select one or more sensitivity labels that prevent Microsoft 365 Copilot and agents summarizing the labeled data. For more information about this capability, see Learn about using Microsoft Purview Data Loss Prevention to protect interactions with Microsoft 365 Copilot and Copilot Chat.
   - **Detect risky interactions in AI apps** to calculate user risk by detecting risky prompts and responses in Microsoft 365 Copilot, agents, and other generative AI apps. For more information, see Risky AI usage.
   - **Discover and govern interactions with ChatGPT Enterprise AI** by registering ChatGPT Enterprise workspaces, you can identify potential data exposure risks by detecting sensitive information that's shared with ChatGPT Enterprise.

- **Get guided assistance to AI regulations**, which uses control-mapping regulatory templates from Compliance Manager.
- **Secure interactions for Microsoft Copilot experiences**, which captures prompts and responses for Copilot in Fabric, and Security Copilot. Without a policy similar to the one that's created with this recommendation, auditing events are captured for Copilot in Fabric and Security Copilot, but not the prompts and responses.
- **Detect sensitive info shared with AI via network**, which uses network data security to detect sensitive info types shared with AI apps in browsers, applications, APIs, add-ins, and more, using a Secure Access Service Edge or Security Service Edge integration.
- **Secure interactions from enterprise apps**, which captures prompts and responses for regulatory compliance from Entra-registered AI apps, ChatGPT Enterprise Connector, and applications built on Microsoft Foundry.
- **Secure data in Azure AI apps and agents**, which has set up instructions to capture prompts and responses for AI apps that use one or more Azure AI subscriptions.

You can use the **View all recommendations** link, or **Recommendations** from the navigation pane to see all the available recommendations for your tenant, and their status. When a recommendation is complete or dismissed, you no longer see it on the **Overview** page.

4. Use the **Reports** section or the **Reports** page from the navigation pane to view the results of the default policies created. You need to wait at least a day for the reports to be populated. Select the categories of **Copilot experiences and agents**, **Enterprise AI apps**, and **Other AI apps** to help you identify the specific generative AI app.

5. Use the **Policies** page to monitor the status of the default one-click policies created and AI-related policies from other Microsoft Purview solutions. To edit the policies, use the corresponding management solution in the portal. For example, for **DSPM for AI - Unethical behavior in Copilot**, you can review and remediate the matches from the **Communication Compliance** solution.

   > ⓘ Note
   >
   > If you have the older retention policies for the location **Teams chats and Copilot interactions**, they aren't included on this page. Learn how to create separate retention policies for Microsoft Copilot Experiences that will be included on this **Policy** page.

6. Select **Apps and agents** to view a dashboard of AI apps and their agents used across your organization so you can identify and manage any potential data security risks. For each

agent, view details about sensitive data that they accessed and how they are protected by policies from Microsoft Purview.

7. Select **Activity explorer** to see details of the data collected from your policies.

   This more detailed information includes activity type and user, date and time, AI app category and app, app accessed in, any sensitive information types, files referenced, and sensitive files referenced.

   Examples of activities include **AI interaction**, **Sensitive info types**, and **AI website visit**. Prompts and responses are included in the **AI interaction** events when you have the right permisisons. You can use the **Web queries** filter in activity explorer to more easily identify which AI interactions have a web query, with the AI interaction flyout pane diplaying the search query text in the **Web references AI app used** section. For more information about the events, see Activity explorer events.

   Similar to the report categories, workloads include **Copilot experiences and agents**, **Enterprise AI apps**, and **Other AI Apps**. Examples of an app ID and app host for **Copilot experiences and agents** include **Microsoft 365 Copilots** and **Copilot Studio**. **Enterprise AI apps** includes **ChatGPT Enterprise**. **Other AI Apps** include the apps from the supported third-party AI sites, such as those used for Gemini and ChatGPT.

8. Select **Data risk assessments** to identify and fix potential data oversharing risks in your organization. Because of the power and speed AI can proactively surface content that might be obsolete, over-permissioned, or lacking governance controls, generative AI amplifies the problem of oversharing data. Use data risk assessements to both identify and remediate issues.

   For oversharing insights in SharePoint and OneDrive, use the **Microsoft 365** tab, and the **Fabric** tab for items in Fabric workspaces.

   Microsoft 365

   A default data risk assessment automatically runs weekly for the top 100 SharePoint sites based on usage in your organization, and you might have already run a custom assessment as one of the recommendations. However, come back regularly to this option to check the latest weekly results of the default assessment and run custom assessments when you want to check for different users or specific sites. After a custom assessment has run, wait at least 48 hours to see the results that don't update again. You'll need a new assessment to see any changes in the results.

   The **Default assessment** displays at the top of the page with a quick summary, such as the total number of items found, the number of sensitive data detected, and the

number of links sharing data with anyone. The first time the default assessment is created there's a 4-day delay before results are displayed.

After you select **View details** for more in-depth information, from the list, select each site to access the flyout pane that has tabs for **Overview**, **Identify**, **Protect**, and **Monitor**. Use the information on each tab to learn more, and take recommended actions. For example:

Use the **Identify** tab to identify how much data has been scanned or not scanned for sensitive information types, with an option to initiate an on-demand classification scan as needed.

Use the **Protect** tab to select options to remediate oversharing, which include:

- **Restrict access by label**: Use Microsoft Purview Data Loss Prevention to create a DLP policy that prevents Microsoft 365 Copilot and agents from summarizing data when it has sensitivity labels that you select. For more information about how this works and supported scenarios, see Learn about using Microsoft Purview Data Loss Prevention to protect interactions with Microsoft 365 Copilot and Copilot Chat.
- **Restrict all items**: Use SharePoint Restricted Content Discovery to list the SharePoint sites to be exempt from Microsoft 365 Copilot. For more information, see Restrict discovery of SharePoint sites and content.
- **Create an auto-labeling policy**: When sensitive information is found for unlabeled files, use Microsoft Purview Information Protection to create an auto-labeling policy to automatically apply a sensitivity label for sensitive data. For more information about how to create this policy, see How to configure auto-labeling policies for SharePoint, OneDrive, and Exchange.
- **Create retention policies**: When content hasn't been accessed for at least 3 years, use Microsoft Purview Data Lifecycle Management to automatically delete it. For more information about how to create the retention policy, see Create and configure retention policies.

Use the **Monitor** tab to view the number of items in the site shared with anyone, shared with everyone in the organization, shared with specific people, and shared externally. Select **Start a SharePoint site access review** for information how to use the SharePoint data access governance reports.

9. To create your own custom data risk assessment, select **Create custom assessment** to identify potential oversharing issues for all or selected users, the data sources to scan, and run the assessment.

> **Note**
>
> Microsoft 365 custom assessments optionally support item-level scanning with remediation actions when you complete a one-time authentication process. To specify the required settings, you must have a registered Entra application. For more information about the Entra application, see **Prerequisites for Microsoft 365 item-level scanning for data risk assessments**.
>
> Select **Item-level** on the **Scan level** page, and then **Authenticate** to specify the Entra application information.

This data risk assessment is created in the **Custom assessments** category. Wait for the status of your assessment to display **Completed**, and select it to view details. To rerun a custom data risk assessment, and to see results after the 30-day expiration, use the duplicate option to create a new assessment with the same selections.

Item-level scanning and remediation for items that are potentially overshared is applicable to Microsoft 365 only, and currently restricted to SharePoint sites. This scan identifies items as potentially overshared if they have a sharing link for external or anonymous users, and also shows any applied sensitivity label and the owner of each item.

When the scan is complete, open the custom assessment and view item-level insights from the **Potentially overshared items** tab or the **View items** button from the **Potentially overshared items** banner. The following remediation actions can be taken on the identified potentially overshared items:

- **Resolve**, for example if you decide the item isn't at risk of oversharing.
- **Apply sensitivity label** for items that are identified as currently unlabeled or should have a different sensitivity label.
- **Notify** the site owner with an email notification (not customizable).
- **Remove sharing link** to remove the existing sharing link so it can no longer be used to access the item. This is an action to be used sparingly because it could prevent people from accessing the item legitimately. In this scenario, the site owner or item owner must then set a less permissive sharing link type for authorized users. If you haven't already, consider configuring sensitivity labels for a default sharing link.

The email notification contains information and instructions about the potentially overshared items in the site. It also contains options for them to go to the SharePoint site to manage access to the items from there, or to the Microsoft Purview portal for a read-only view of the item-level scan results that only show items and sites that they own.

> **Tip**
>
> Both the default and custom data risk assessments provide an **Export** option that let you save and customize the data into a choice of file formats (Excel, CSV, JSON, TSV).

Limits for Microsoft 365:

- A maximum of 200,000 items per location, which applies to both a custom data risk assessment and a default data risk assessment. The count of files reported might not be accurate when there are more than 100,000 files per location.
- OneDrive currently isn't supported for item-level scanning.
- A current maximum of 10 SharePoint sites for item-level scanning.

Next steps

Use the policies, tools, and insights from DSPM for AI in conjunction with additional protections and compliance capabilities from Microsoft Purview. If you're interested in a specific AI app, use the following links:

- Microsoft 365 Copilot & Microsoft 365 Copilot Chat
- Microsoft Security Copilot
- Copilot in Fabric
- Microsoft Copilot Studio
- Microsoft Facilitator
- Channel Agent in Microsoft Teams
- Entra-registered AI apps
- Microsoft Foundry
- ChatGPT Enterprise
- Other AI apps
- Agents

Last updated on 12/15/2025

Considerations for DSPM for AI to manage data security and compliance protections for AI interactions - (classic)

> *Microsoft Purview service description*

> **Note**
>
> This article is for the **classic** version of Data Security Posture Management for AI that's now replaced with a new version that incorporates support for AI apps and agents, new features and broader reach, with simplified management.
>
> These improvements won't be added to this classic version so we invite you to try the new Data Security Posture Management.

For the most part, Data Security Posture Management for AI is easy to use and self-explanatory, guiding you through prerequisites and preconfigured reports and policies. Use this section to complement that information and provide additional details that you might need.

Prerequisites for Data Security Posture Management for AI

To use Data Security Posture Management for AI from the Microsoft Purview portal, you must have the following prerequisites:

- You have the right permissions.
- Required for monitoring interactions with Copilot and agents:
  - Microsoft Purview auditing is enabled for your organization. Although this is the default, you might want to check the instructions for Turn auditing on or off.
  - For Microsoft 365 Copilot and agents, users are assigned a license for Microsoft 365 Copilot.
  - For Copilot in Fabric and Security Copilot:

    - The enterprise version of Microsoft Purview data governance, to support the required APIs.
    - A collection policy, such as the one created from the recommendation or remediation action **Secure interactions for Microsoft Copilot experiences**.

- Required for monitoring interactions and applying DLP policies to other AI apps in Edge:
    - An Edge configuration policy is required to activate the Microsoft Purview integration in Edge. For configuration information, see Activate your DLP policy in Microsoft Edge.

- Required for monitoring interactions with third-party generative AI sites:

    - Devices are onboarded to Microsoft Purview, required for:
        - Gaining visibility into sensitive information that's shared with third-party generative AI sites. For example, a user pastes credit card numbers into ChatGPT.
        - Applying endpoint DLP policies to warn or block users from sharing sensitive information with third-party generative AI sites. For example, a user identified as elevated risk in Adaptive Protection is blocked with the option to override when they paste credit card numbers into ChatGPT.

    - The Microsoft Purview browser extension is deployed to Windows users and required to discover visits to third-party generative AI sites by using an Insider Risk Management policy. The browser extension is also required for Endpoint DLP policies on Windows when you use Chrome.

- For Entra-registered AI apps, they are integrated with the Microsoft Purview SDK to support all the Microsoft Purview capabilities.

- For AI apps other than Microsoft 365 Copilot and Microsoft Facilitator, you've set up pay-as-you-go billing for your organization. When this billing model is applicable for specific configurations, you'll see notifications and instructions in the UI.

You'll find more information about the prerequisites for auditing, device onboarding, and the browser extension in Data Security Posture Management for AI: Navigate to **Overview** > **Get started** section.

For a list of currently supported third-party AI apps, see Supported AI sites by Microsoft Purview for data security and compliance protections.

> ⓘ **Note**

If you're using **administrative units**, restricted administrators won't be able to create the one-click policies that apply to all users. You must be an unrestricted administrator to create these policies. Restricted administrators see the results only for users in their assigned administrative unit for the **Policies** page, and in the reports and activity explorer in Microsoft Purview Data Security Posture Management for AI.

Prerequisites for Fabric data risk assessments

Use the following instructions to meet the prerequisites to run data risk assessments for Fabric workspaces, and before you select **Set config** in DSPM for AI.

For steps 1 and 2, use one of the following Entra roles for the required permissions:

- Cloud Application Administrator
- Application Administrator
- Privileged Role Administrator

For step 3, you must be a Fabric administrator.

1. In the Microsoft Entra admin portal, create and configure a registered app to be the service principal authentication for Fabric. You will authenticate the app in one of two ways: With **Federated Credentials** (recommended) or **Client Secret**. The federated credentials option is more secure because it lets Kubernetes workloads authenticate to Azure services without hard-coding secrets, while ensuring each pod only has the permissions it needs. For more information about this authentication method, see Overview of federated identity credentials in Microsoft Entra ID.

   After you've created this registered app and still in the Entra admin portal, copy and store the **Application (client) ID**, which is always required to setup data risk assessments for Fabric. Then, to authenticate with federated credentials, work with your compliance administrator to identify and enter the **OIDC Issuer URL**, **Namespace**, and **Service Account Name** under **Edit connection**. Enter these values into the corresponding fields in **Your App name** > **Certificates & Secrets** > **Federated Credentials** > **Kubernetes accessing Azure resources**. Alternatively, to authenticate with a client secret, store the **Client Secret** value to then enter in the Microsoft Purview portal:

   - To locate the **Application (client) ID**: In the Microsoft Entra admin portal, **Identity** > **Applications** > **App registrations**, select the app you created for Fabric. The **Application (client) ID** is displayed on the **Overview** tab.

- Create a client secret for this app. Copy and securely save the value immediately. It will only be shown once.

2. Now navigate to **Groups** to add the registered app as a group member to a new or existing security group. Its type displays as **Enterprise application**.

3. In the Fabric admin portal, navigate to **Tenant settings** > **Admin API settings** and enable the options **Service principals can access read-only admin APIs** and **Service principals can access admin APIs used for updates**. For both of these, in the **Apply to:** sections, select **Specific security groups** and specify the security group that you configured in the previous step. Select **Apply** to save and apply these tenant settings.

For more information about steps 1-2, see Enable service principal authentication for admin APIs.

For more information about the tenant settings and support for information protection in Fabric, see Tenant settings index and Information protection in Microsoft Fabric.

Prerequisites for Microsoft 365 item-level scanning for data risk assessments

Use the following instructions to create an Entra application with the required permissions for Microsoft 365 item-level scanning in a custom data risk assessment.

You must have one of the following Entra roles to create and configure the required registered app:

- Cloud Application Administrator
- Application Administrator
- Privileged Role Administrator

In the Entra admin portal, create and configure a registered app with the following settings:

- **Supported account types**: **Accounts in this organizational directory only**

- **API permissions**: **Microsoft APIs** > **Microsoft Graph** > **Application permissions**
  - **Application: Application.Read.All**
  - **Directory: Directory.Read.All**
  - **Files: Files.ReadWrite.All**
  - **SensitivityLabels: SensitivityLabels.Read.All**
  - **Sites: Sites.ReadWrite.All**
  - **User: User.Read.All**

- Grant admin consent for your tenant

Now create a client secret for this app. Copy and securely save the value immediately. It will only be shown once.

You'll also need the registered application's ID, which is displayed as **Application (client) ID** in the **Overview** tab.

One-click policies from Data Security Posture Management for AI

After the default policies are created, you can view and edit them at any time from their respective solution areas in the portal. For example, you want to scope the policies to specific users during testing, or for business requirements. Or, you want to add or remove classifiers that are used to detect sensitive information. Use the **Policies** page to quickly navigate to the right place in the portal.

Some policies, such as **DSPM for AI - Capture interactions for Copilot experiences** and **DSPM for AI - Detect sensitive info shared with AI via network** are collection policies that you can edit, and if necessary, delete like any other collection policy. For more information, see Accessing collection policies.

If you delete any of the policies, their status on the **Policies** page displays **PendingDeletion** and continues to show as created in their respective recommendation cards until the deletion process is complete.

For sensitivity labels and their policies, view and edit these independently from Data Security Posture Management for AI, by navigating to **Information Protection** in the portal. For more information, use the configuration links in Default labels and policies to protect your data.

For more information about the supported DLP actions and which platforms support them, see the first two rows in the table from Endpoint activities you can monitor and take action on.

For the default policies that use Adaptive Protection, this capability is turned on if it's not already on, using default risk levels for all users and groups to dynamically enforce protection actions. For more information, see Quick setup

> ⓘ **Note**

Any default policies created while Data Security Posture Management for AI was in preview and named Microsoft Purview AI Hub won't be changed. For example, policy names will retain their **Microsoft AI Hub -** prefix.

Default policies for data discovery using Data Security Posture Management for AI

- DLP policy: **DSPM for AI: Detect sensitive info added to AI sites**

  Source: **Extend your insights for data discovery**

  This policy discovers sensitive content pasted or uploaded in Microsoft Edge, Chrome, and Firefox to AI sites. This policy covers all users and groups in your org in audit mode only.

- Insider risk management policy: **DSPM for AI - Detect when users visit AI sites**

  Source: **Extend your insights for data discovery**

  This policy detects when users use a browser to visit AI sites.

- Insider risk management policy: **DSPM for AI - Detect risky AI usage**

  Source: Recommendation **Detect risky interactions in AI apps**

  This policy helps calculate user risk by detecting risky prompts and responses in Microsoft 365 Copilot, agents, and other generative AI apps.

- Communication Compliance: **DSPM for AI - Unethical behavior in AI apps**

  Source: Recommendation **Detect unethical behavior in AI apps**

  This policy detects sensitive information in prompts and responses in Microsoft 365 Copilot, agents, and other generative AI apps. This policy covers all users and groups in your organization.

- Collection policy: **DSPM for AI - Capture interactions for Copilot experiences**

  Source: Recommendation **Secure interactions in Microsoft Copilot experiences**

  This policy captures prompts and responses for data security posture and regulatory compliance from Copilot in Fabric, and Security Copilot. Manage them in Microsoft Purview solutions like eDiscovery, Data Lifecycle Management, and more.

- Collection policy: **DSPM for AI - Detect sensitive info shared with AI via network**

  Source: Recommendation **Extend insights into sensitive data in AI app interactions**

  This policy detects sensitive information shared with AI apps in browsers, applications, APIs, add-ins, and more, using a Secure Access Service Edge (SASE) or Security Service Edge (SSE) integration.

  > **Important**
  >
  > This policy requires that you manually add one or more Secure Access Service Edge (SASE) or Security Service Edge (SSE) integrations in Data Loss Prevention settings. The detection of AI interactions is dependent on the network partner implementation.
  >
  > If you want to capture prompts and responses in addition to detecting sensitive information, you must edit the collection policy and select the option to capture content.

- Collection policy: **DSPM for AI - Capture interactions for enterprise AI apps**

  Source: Recommendation **Secure interactions from enterprise apps**

  This policy captures prompts and responses for regulatory compliance from enterprise AI apps, such as Chat GPT Enterprise and AI apps connected through Microsoft Entra or Microsoft Foundry, so they can be managed in Microsoft Purview solutions like eDiscovery, Data Lifecycle Management, and more.

- Collection policy: **DSPM for AI - Detect sensitive info shared in AI prompts in Edge**

  Source: **Extend your insights for data discovery**

  This policy detects prompts sent to generative AI apps in Microsoft Edge and discovers sensitive information shared in prompt contents. This policy covers all users and groups in your organization in audit mode only and doesn't capture content.

Default policies from data security to help you protect sensitive data used in generative AI

- DLP policy **DSPM for AI - Block sensitive info from AI sites**

  Source: Recommendation **Fortify your data security**

This policy uses Adaptive Protection to give a block-with-override to elevated risky users attempting to paste or upload sensitive information to other AI apps in Edge, Chrome, and Firefox. This policy covers all users and groups in your org in test mode.

- DLP policy **DSPM for AI - Block elevated risk users from submitting prompts to AI apps in Microsoft Edge**

  Source: Recommendation **Fortify your data security**

  Using Adaptive Protection, this policy blocks elevated, moderate, and minor risk users attempting to put information in AI apps while using Microsoft Edge. Users included in this policy will be blocked from using unprotected browsers because this policy includes unmanaged apps.

- DLP policy **DSPM for AI - Block sensitive info from AI apps in Edge**

  Source: Recommendation **Fortify your data security**

  This policy detects inline for a selection of common sensitive information types and blocks prompts from being sent to AI apps while using Microsoft Edge. Users included in this policy will be blocked from using unprotected browsers because this policy includes unmanaged apps.

- DLP policy **DSPM for AI - Protect sensitive data from Copilot processing**

  Source: Recommendation **Protect items with sensitivity labels from Microsoft 365 Copilot and agent processing**

  This policy blocks Microsoft 365 Copilot and agents from processing items with the sensitivity labels selected in this policy.

- Information Protection - **Sensitivity labels and policies**

  Source: Recommendation **Protect your data with sensitiivity labels**

  This recommendation creates default sensitivity labels and sensitivity label policies. If you've already configured sensitivity labels and their policies, this configuration is skipped.

Activity explorer events

Use the following information to help you understand the events you might see in the activity explorer from Data Security Posture Management for AI. References to a generative AI site can

include Microsoft 365 Copilot, Microsoft 365 Copilot Chat, agents, other Microsoft copilots, and third-party AI sites.

Expand table

Event	Description
AI interaction	User interacted with a generative AI site. Details include the prompts and responses, except for unmanaged AI apps in Edge where text prompts only are included. For Microsoft 365 Copilot and Microsoft 365 Copilot Chat, this event requires auditing to be turned on. For Copilot in Fabric and Security Copilot, and for non-Copilot AI apps, prompts and responses require a collection policy with content capture selected to capture these interactions.
AI website visit	User browsed to a generative AI site.
DLP rule match	A data loss prevention rule was matched when a user interacted with a generative AI site. Includes DLP for Microsoft 365 Copilot.
Sensitive info types	Sensitive information types were found while a user interacted with a generative AI site. For Microsoft 365 Copilot and Microsoft 365 Copilot Chat, this event requires auditing to be turned on but doesn't require any active policies.

Known issues:

- The **AI interaction** event doesn't always display text for the prompt and response. Sometimes, the prompt and response spans consecutive entries. Other scenarios can include:
  - Microsoft Facilitator AI-generated notes, no prompt or response is displayed
  - When a user doesn't have a mailbox hosted in Exchange Online, no prompt or response is displayed
- The **Sensitive info types detected** event doesn't display the user risk level.
- For Microsoft Facilitator AI-generated notes, **AI interaction** events can't be linked to **Sensitive info types detected** events.
- For collection policies, no prompt or response is displayed if the option to capture content isn't selected in the policy. For example, the one-click policy **DSPM for AI - Detect sensitive info shared with AI via network** doesn't select this option when the policy is automatically created, but you can manually edit the policy and select this option after the policy is created.

Last updated on 05/01/2026

Permissions for Data Security Posture Management for AI - (classic)

> *Microsoft Purview service description*

> **Note**
>
> This article is for the **classic** version of Data Security Posture Management that's being replaced with a new version that expands coverage to more data sources, introduces guided workflows for proactive risk management, and streamlines data security operations so you can more confidently adopt AI across your digital estate.
>
> These improvements won't be added to this classic version so we invite you to try the new Data Security Posture Management, currently in preview.

Members of your security and compliance teams who are responsible for managing AI apps in the Microsoft Purview Data Security Posture Management for AI need appropriate permissions when they sign in to the Microsoft Purview portal.

Roles and role groups that can view, create, and edit in Data Security Posture Management for AI:

- **Microsoft Entra Compliance Administrator** role
- **Microsoft Entra Global Administrator** role
- **Microsoft Purview Compliance Administrator** role group

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

Roles and role groups that can view-only in Data Security Posture Management for AI:

- **Microsoft Purview Security Reader** role group
- **Purview Data Security AI Viewer** role

- **AI Administrator** role from Entra
- **Purview Data Security AI Content Viewer** role for AI interactions only

To help you assign the right permissions to users, use the following guidance, depending on the portal you're using:

- Permissions in the Microsoft Purview portal
- Assign Microsoft Entra roles to users
- Permissions in Exchange Online

Use the following table to understand the detailed permissions for different activities in Data Security Posture Management for AI.

Permissions by activities

✓: Supported. The role or role group have permissions to do the specified activities.

✕: Not supported. The role or role group don't have permissions to do the specified activities.

Expand table

Activities	Microsoft Entra Compliance Administrator role	Microsoft Entra Global Administrator role	Microsoft Purview Compliance Administrator role group	Roles or role groups that are view-only*	When not supported, additional role groups required
View all get started steps	✓	✓	✓	✓	Not applicable
Complete action on getting started steps	✓	✓	✓ Excludes **Activate Audit**	✕	Microsoft Exchange Compliance Management Microsoft Exchange Records Management Microsoft Exchange Organization Management

Activities	Microsoft Entra Compliance Administrator role	Microsoft Entra Global Administrator role	Microsoft Purview Compliance Administrator role group	Roles or role groups that are view-only*	When not supported, additional role groups required
View completion status of getting started steps	✓	✓	✓ Excludes status of **Activate Audit**	✓ Excludes: Status of **Activate Audit** Status of **Extend Your Insights**	For **Activate Audit**: Microsoft Exchange View-Only Organization Management Microsoft Exchange Hygiene Management Microsoft Exchange Compliance Management Microsoft Exchange Records Management Microsoft Exchange Organization Management For **Extend Your Insights**: Microsoft Purview Insider Risk Management Administrator Microsoft Purview Insider Risk Management Analyst

Activities	Microsoft Entra Compliance Administrator role	Microsoft Entra Global Administrator role	Microsoft Purview Compliance Administrator role group	Roles or role groups that are view-only*	When not supported, additional role groups required
					Microsoft Purview Insider Risk Management Investigator
View all recommendations	✓	✓	✓	✓	Not applicable
Complete actions on recommendation cards	✓	✓	✓	✕	Not applicable
View completion status of recommendation cards	✓	✓	✓	✓ Excludes **Unethical Behavior** card	Communication compliance Administrator
View all graphs from the **Reports** page	✓	✓	✓	✓	Not applicable
View all policies in the policy list	✓	✓	✓	✓ Excludes: Insider risk management policies Communication compliance policies	For insider risk management polices: Microsoft Purview Insider Risk Management Administrator Microsoft Purview Insider Risk Management Analyst Microsoft Purview Insider Risk Management Investigator

Activities	Microsoft Entra Compliance Administrator role	Microsoft Entra Global Administrator role	Microsoft Purview Compliance Administrator role group	Roles or role groups that are view-only*	When not supported, additional role groups required
					For communication compliance policies: Communication Compliance Administrator
View all events in activity explorer	✓ Excludes browse to URL (**AI Visit**) from insider risk management	✓ Excludes browse to URL (**AI Visit**) from insider risk management	✓ Excludes browse to URL (**AI Visit**) from insider risk management	✓ Excludes browse to URL (**AI Visit**) from insider risk management	Microsoft Purview Insider Risk Management Analyst Microsoft Purview Insider Risk Management Investigator
View user risk level of an individual user in all events from activity explorer	✕	✕	✕	✕	Microsoft Purview Insider Risk Management Analyst Microsoft Purview Insider Risk Management Investigator
View link to view user details in insider risk management in all events from activity explorer	✕	✕	✕	✕	Microsoft Purview Insider Risk Management Analyst Microsoft Purview Insider Risk

Activities	Microsoft Entra Compliance Administrator role	Microsoft Entra Global Administrator role	Microsoft Purview Compliance Administrator role group	Roles or role groups that are view-only*	When not supported, additional role groups required
					Management Investigator
View the prompts and responses within **AI Interaction** events from activity explorer	✗	✗	✗	✗	Content Explorer Content Viewer Microsoft Purview Data Security AI Content Viewer
Create data risk assessments	✓	✓	✓	✗	Not applicable
View data risk assessments	✓	✓	✓	✓	Not applicable
View file details for data risk assessments	✗	✗	✗	✗	Content Explorer Content Viewer Content Explorer List Viewer
View **Apps and agents** page	✓	✓	✓	✓	Not applicable

* Includes Microsoft Purview Security Reader role group, Microsoft Purview Data Security AI Viewer role, and the AI Administrator role from Entra

Custom role groups

Instead of granting access to Data Security Posture Management for AI by using the built-in role groups, you can grant access by including the Microsoft Purview Compliance Administrator role in a custom role group. For read-only permissions, include the the Microsoft Purview Security Reader role, the Purview Data Security AI Viewer role, or the AI Administrator role from Entra.

If a custom role group includes the Microsoft Purview Compliance Administrator role, the user has the same access to Data Security Posture Management for AI as the Microsoft Purview Compliance Administrator role group, except for the following:

- Create, view, update, and delete policies for insider risk management and communication compliance

If a custom role group includes the Microsoft Purview Security Reader role, the Purview Data Security AI Viewer role, or the AI Administrator role, the user has the same access to Data Security Posture Management for AI as the Microsoft Purview Security Reader role group, except for the following:

- View information protection policies

Last updated on 04/01/2026

List of AI sites supported by Microsoft Purview Data Security Posture Management (DSPM) and DSPM for AI (classic)

Summarize this article for me

> *Microsoft Purview service description*

This article lists the generative AI sites that are supported by Microsoft Purview Data Security Posture Management and Microsoft Purview Data Security Posture Management for AI (classic).

You can expect this list to increase over time.

Supported AI sites

```
*.10web.cloud
*.10web.io
*.10web.me
*.10zebra.com
*.20ma.cn
*.4b3.com
*.7aai.com
*.a0.dev
*.a1d.ai
*.a2e.ai
*.abridge.com
*.adcreative.ai
*.addy.ai
*.addy.so
*.adventureai.gg
*.agentgpt.reworkd.ai
*.agpt.co
*.ai.poweredtemplate.com
*.ai.yiios.com
*.ai2006.io
*.ai21.com
*.ai2sql.io
*.ai-blogkun.com
*.aibuddy.chat
*.aichat.gif99.cn
*.aichatbot.co
*.aichatting.net
*.ai-claude.net
```

```
*.aiclip.net
*.aicomicgenerator.net
*.aicontentfy.com
*.aicontentstudio.ai
*.aicopywriter.ai
*.aicotravel.com
*.aicoverletter.com
*.aicut.pro
*.aider.chat
*.aidesign.center
*.aidocmaker.com
*.aidungeon.com
*.aidungeon.io
*.ai-ebooks.co.uk
*.aiexcelbot.com
*.aifaceswap.io
*.aifree.site
*.aigca.org
*.aigcdeep.com
*.aigen.biz.id
*.ai-ghostwriter.com
*.aiheadshotgenerator.com
*.aihomedesign.com
*.aihumanize.com
*.aihumanize.io
*.aiimagegenerator.io
*.aiisajoke.com
*.ailabtools.com
*.ailessonplan.com
*.aimagicface.com
*.aimdoc.ai
*.aimreply.com
*.aimusic.so
*.aimusicfactory.ai
*.aipassportphoto.com
*.aipdf.ai
*.aipoemgenerator.org
*.ai-portraits.org
*.ai-q.in
*.aireception.io
*.airpost.ai
*.aiseo.ai
*.aisozai.com
*.aissistify.com
*.aistudios.com
*.aiswers.com
*.aiter.io
*.aitestkitchen.withgoogle.com
*.aitextpromptgenerator.com
*.aithor.com
*.aitubo.ai
*.aiuncensored.info
*.aiva.ai
*.ai-writer.com
*.aiwritingpal.com
```

```
*.aixtsy.com
*.ajelix.net
*.ajelix.org
*.akool.com
*.allenai.org
*.anagram.ai
*.analogenie.com
*.anchorbrowser.io
*.animaker.com
*.anthropic.com
*.anythingllm.com
*.anyword.com
*.aphra.me
*.app.aleph-alpha.com
*.app.alice.tech
*.app.instantdeco.ai
*.app.iokoo.co
*.app.koah.ai
*.app.longshot.ai
*.app.neuro-flash.com
*.app.questionwell.ai
*.app.submagic.co
*.app.suissechat.ch
*.app.surferseo.com
*.app.synthesia.io
*.appaca.ai
*.applaime.com
*.appointwise.io
*.aptlystar.ai
*.aragon.ai
*.archie.8base.com
*.architectgpt.io
*.archivinci.com
*.arria.com
*.artbreeder.com
*.artflow.ai
*.article.audio
*.articlefiesta.com
*.articleforge.com
*.articlereword.com
*.artiphoria.ai
*.askaichat.app
*.askan.ai
*.askbrian.ai
*.askgpt.app
*.askjack.io
*.askmany.cn
*.askmanyai.cn
*.asksocrates.app
*.askyourpdf.com
*.assignmentgpt.ai
*.atomai.fr
*.audiobox.metademolab.com
*.audiostack.ai
*.auraai.app
```

```
*.aurwy.com
*.autogenai.com
*.automateed.com
*.autoppt.com
*.autoreelapp.com
*.autoshorts.ai
*.avatarwithai.com
*.aws.amazon.com/bedrock/titan
*.aws.amazon.com/q
*.azure.microsoft.com/en-us/products/ai-services/openai-service
*.bagoodex.io
*.bannerbear.com
*.bard.google.com
*.basedlabs.ai
*.baz.co
*.bearly.ai
*.beatoven.ai
*.beautiful.ai
*.beehive.ai
*.beewriter.com
*.befunky.com
*.beingai.com
*.bertha.ai
*.beta.picjam.ai
*.betterpic.io
*.bettersynonyms.com
*.bewerbung.ai
*.bhuman.ai
*.bilmans.com
*.bing.com/chat
*.bizplanner.ai
*.blackink.ai
*.blinkvideo.ai
*.blogaudio.co
*.bluewillow.ai
*.bolt.new
*.bomml.ai
*.bookbud.ai
*.bookread.ai
*.bookwizardai.com
*.boomy.com
*.boostpixels.com
*.botowski.com
*.bots.miku.gg
*.botstar.com
*.brandmark.io
*.bravostudio.app
*.browsegpt.ai
*.buhuw.com
*.buildglare.com
*.bulkgpt.ai
*.business.adobe.com/products/sensei/adobe-sensei-genai.html
*.buster.ai
*.buzznews.today
*.bwana19.com
```

```
*.bypassgpt.ai
*.bytecap.io
*.c.ai
*.c5i.ai
*.caboodleai.net
*.camb.ai
*.cantina.com
*.capcut.com
*.captions.ai
*.caribbeans.ai
*.censusgpt.com
*.chadgpt.ru
*.chai-research.com
*.chanjing.cc
*.channel.io
*.character.ai
*.characternsfw.ai
*.charisma.ai
*.charley.ai
*.charshift.com
*.chat.bing.com
*.chat.com
*.chat.lmsys.org
*.chat.mymap.ai
*.chat.nbox.ai
*.chat.novaapp.ai
*.chat2db.ai
*.chatai.com
*.chataigpt.net
*.chat-ai-israel.com
*.chatba.com
*.chatbase.co
*.chatbit.co
*.chatbotgen.com
*.chatbotgf.com
*.chatboxapp.ai
*.chatcxc.com
*.chatdoc.com
*.chatespanolaigratis.com
*.chatfai.com
*.chatgate.ai
*.chatgbt.live
*.chatglm.cn
*.chatgot.io
*.chatgpt.at
*.chatgpt.ch
*.chatgpt.com
*.chatgpt.com.ua
*.chatgpt.es
*.chatgpt.fr
*.chatgpt.it
*.chatg-pt.org
*.chat-gpt.photos
*.chatgpt.pt
*.chatgpt4.org
```

```
*.chatgpt4google.com
*.chatgpt4online.org
*.chatgpt4rus.ru
*.chatgpt4youtube.com
*.chatgptbot.ru
*.chatgptchatapp.com
*.chatgptdemo.ai
*.chatgptdemo.net
*.chatgptduo.com
*.chatgptdutch.org
*.chatgptextension.ai
*.chatgpt-francais.com
*.chatgptfrancais.org
*.chatgptfree.ai
*.chatgpt-free.ai
*.chatgptgratuit.org
*.chatgpthuman.com
*.chatgptimagegenerator.org
*.chatgpt-israel.org
*.chat-gpt-israel.org
*.chatgptlogindeutsch.de
*.chatgptonline.io
*.chatgptonline.tech
*.chatgptpersian.com
*.chatgpt-quorum.com
*.chatgptspanish.org
*.chatgptsv.se
*.chatgptsvenska.io
*.chatgptwriter.ai
*.chatgptx.com
*.chatgpt-x.de
*.chatgptx.fr
*.chatgptx.ru
*.chatgtp.com
*.chathub.gg
*.chatmost.ru
*.chaton.ai
*.chatop.ai
*.chatopen.app
*.chatpdf.com
*.chatpdf.de
*.chatplayground.ai
*.chatroboten.se
*.chatsimple.ai
*.chattab.app
*.chattab.framer.website
*.chatthing.ai
*.chatuml.com
*.chatx.de
*.chatyy.xingmengsw.cn
*.chub.ai
*.cicube.io
*.civitai.com
*.claid.ai
*.claralabs.com
```

```
*.claude.ai
*.clausebuilder.ai
*.cleanup.pictures
*.cleoai.tech
*.clickable.so
*.clipchamp.com
*.clipdrop.co
*.cliprun.com
*.cmdj.app
*.coachvox.ai
*.code2.ai
*.codegate.ai
*.codesnippets.ai
*.cognition.ai
*.cohere.com
*.cohesive.so
*.coinvent.ai
*.coloringninja.com
*.colormind.io
*.combopix.com
*.comfy.nordy.ai
*.comicsmaker.ai
*.compose.ai
*.consensus.app
*.consistentcharacter.ai
*.contentbot.ai
*.contentforge.ai
*.contentpie.ai
*.contentvillain.com
*.convergence.ai
*.conversionblitz.com
*.copilot.microsoft.com
*.copy.ai
*.copymatic.ai
*.copymonkey.ai
*.copysmith.ai
*.copyter.com
*.copytruck.com
*.copywriting.ai
*.corina.ai
*.cotypist.app
*.coursebox.ai
*.coursepro.ai
*.coverler.com
*.covric.com
*.craftly.ai
*.craiyon.com
*.crammer.app
*.crankoutai.com
*.cre8tiveai.com
*.creaitor.ai
*.createaicourse.com
*.created.app
*.creatify.ai
*.creator.nightcafe.studio
```

```
*.cronbot.ai
*.csm.ai
*.ctrlsheet.ai
*.cubox.cc
*.cursor.com
*.cutout.pro
*.d5next.ai
*.daftart.ai
*.dall-efree.com
*.dante-ai.com
*.davinci.ai
*.dazi.co
*.decidable.com
*.decktopus.com
*.decopy.ai
*.deemos.com
*.deepai.org
*.deepbrain.io
*.deepbrainai.io
*.deepcode.ai
*.deepdreamgenerator.com
*.deepfakesweb.com
*.deep-image.ai
*.deepreview.eu
*.deep-seek.chat
*.deepseek.com
*.deepseek.net
*.deepseek-ai.ru
*.deepseekr1.jfsmaa.com
*.deepseekr1.org
*.deepseekv3.com
*.deepswap.ai
*.deevid.ai
*.describely.ai
*.descrii.tech
*.descript.com
*.designer.microsoft.com
*.designs.ai
*.devin.ai
*.devinapps.com
*.dezgo.com
*.differentdimensionme.net
*.diffusionart.co
*.docgpt.ai
*.docky.ai
*.docuhelp.ai
*.domoai.app
*.donakosy.com
*.dopeloop.ai
*.dopple.ai
*.dorkgpt.com
*.doubao.com
*.draftly.so
*.dream.ai
*.dreamhairai.com
```

```
*.dreamily.ai
*.dreampress.ai
*.dreamstudio.ai
*.dreamtavern.ai
*.dubbing-ai.com
*.dubbingai.io
*.dubverse.ai
*.dxm.com
*.dzine.ai
*.east.im
*.easymusic.ai
*.easy-peasy.ai
*.echo.win
*.echt-lu.de
*.editapp.ai
*.editee.com
*.editgpt.app
*.edit-videos-online.com
*.eduwiz.ai
*.edwardbot.com
*.effectssdk.ai
*.egger.ai
*.eightify.app
*.eilla.ai
*.elai.io
*.elephas.app
*.eleuther.ai
*.elevenlabs.io
*.elicit.com
*.engagely.ai
*.essai.pro
*.essayailab.com
*.essay-builder.ai
*.essaygrader.ai
*.essaypal.ai
*.ev5.ai
*.excelmaster.ai
*.explainpaper.com
*.faceswapper.ai
*.facetuneapp.com
*.fadr.com
*.faible.ai
*.fakeyou.com
*.falconllm.tii.ae
*.fanfuel.app
*.fantasy.ai
*.fashionphoto.ai
*.fermat.app
*.figflow.io
*.finechat.ai
*.finepixel.lusion.ai
*.fineshare.com
*.finito.ai
*.fireflies.ai
*.firefly.adobe.com
```

```
*.firetexts.co
*.fitnessai.com
*.flair.ai
*.flatai.org
*.flexclip.com
*.fliki.ai
*.flowgpt.com
*.flowrite.com
*.flux1.ai
*.fluxpro.ai
*.fomo.ai
*.fontjoy.com
*.forethought.ai
*.formulabot.com
*.formwise.ai
*.fotor.com
*.foxyai.ai
*.frase.io
*.freedomgpt.com
*.freeimagegenerator.com
*.freepass.ai
*.fuelix.ai
*.fullstackdata.com
*.fusionartai.com
*.gajix.com
*.galichat.com
*.gamma.app
*.gamma.work
*.ganstudio.com
*.gemini.google.com
*.genai-go.com
*.genaizurich.ch
*.gencraft.com
*.genei.io
*.generated.photos
*.generatelyrics.io
*.generativeai.pub
*.generatorxyz.com
*.genesismedical.ai
*.gerwin.app
*.get-ai.world
*.getaprototype.com
*.getchunky.io
*.getconch.ai
*.getcoralai.com
*.getcube.one
*.getgptapi.com
*.getimg.ai
*.getliner.com
*.getmerlin.in
*.getpoppy.ai
*.getsmartgpt.com
*.gettrove.co
*.getvoila.ai
*.getwoord.com
```

```
*.ghibliai.io
*.giga.chat
*.gingerkeyboard.com
*.gista.co
*.github.com/features/copilot
*.giti.ai
*.giz.ai
*.gizzmo.ai
*.glasp.co
*.glass-imaging.com
*.glif.app
*.gliglish.com
*.godinabox.co
*.golf.ai
*.goneural.ai
*.goodwill.ribbit.works
*.gozen.io
*.gpt.h2o.ai
*.gpt3demo.com
*.gpt4all.io
*.gpt4free.io
*.gpt6.ai
*.gptassistant.app
*.gptexcel.uk
*.gptfy.co
*.gptgame.app
*.gptgo.ai
*.gptinf.com
*.gptkit.ai
*.gptnederlands.nl
*.gptonline.ai
*.gpt-persona.com
*.gpt-ppt.neftup.app
*.gpttranslator.co
*.gptunnel.ru
*.gptzero.me
*.grammarly.ai
*.grammarly.com
*.grammarly.net
*.grammatikai.com
*.grock.com
*.grok.com
*.grok.x.ai
*.grok-2.ai
*.guidde.com
*.hailuoai.com
*.hailuoai.video
*.hal9.com
*.hap.ai
*.harpa.ai
*.harvey.ai
*.headlime.com
*.headshotaigenerator.com
*.headshotpro.com
*.hebbia.com
```

```
*.hedra.com
*.hei.io
*.heimdallapp.org
*.hellohistory.ai
*.heuristi.ca
*.heyboss.xyz
*.heyfriday.ai
*.heygen.com
*.heygpt.chat
*.higgsfield.ai
*.high-eleven.com
*.hinative.com
*.hippocraticai.com
*.hiretruffle.com
*.hitalk.ai
*.hoppycopy.co
*.hotpot.ai
*.hourone.ai
*.huatuogpt.cn
*.huggingface.co/spaces/tiiuae/falcon-180b-demo
*.humanizeai.io
*.humanizeai.pro
*.humanpal.io
*.humata.ai
*.hypotenuse.ai
*.ichatwithgpt.com
*.icon.me
*.icongenerate.com
*.id.freepik.com
*.ideabuddy.com
*.ideapicker.io
*.ideasai.com
*.ideiaschatgpt.com.br
*.ideogram.ai
*.illuminate.google.com
*.illustroke.com
*.image-to-caption.com
*.image-upscaling.net
*.imaginario.ai
*.imagine.art
*.inbox-ai.app
*.infinitecraft.ai
*.infinity.ai
*.inkforall.com
*.inputai.com
*.insmind.com
*.instantanswers.xyz
*.instatext.io
*.interiorai.com
*.interview.lol
*.invideo.io
*.iris.ai
*.ivrit.ai
*.ixcoach.com
*.jammable.com
```

```
*.janusai.pro
*.jasper.ai
*.jellypod.ai
*.jiangziai.com
*.jocondeai.com
*.jogg.ai
*.kafkai.com
*.kaiber.ai
*.kapwing.com
*.kb-whisper.mesu.re
*.kerlig.com
*.kibo.vercel.app
*.kickresume.com
*.kittensai.com
*.kittl.com
*.klingai.com
*.koala.sh
*.kodus.io
*.kokorottsai.com
*.komo.ai
*.korbit.ai
*.krater.ai
*.krea.ai
*.kudoway.eu
*.lalal.ai
*.lambdatest.com
*.lamini.ai
*.landbot.io
*.langchain.com
*.laragpt.com
*.learn.xyz
*.learnitive.com
*.learnt.ai
*.letsenhance.io
*.letstrip.co.il
*.letz.ai
*.lexalytics.com
*.lexica.art
*.ligo.bio
*.linke.ai
*.linkfox.com
*.llama.com
*.llama.meta.com
*.llama2.ai
*.llamachat.app
*.logoai.com
*.lolgames.net
*.loudly.com
*.loudme.ai
*.lovo.ai
*.lp.ai-copywriter.jp
*.lp.comic-copilot.ai
*.lucie.chat
*.luis.ai
*.lumen5.com
```

```
*.lummi.ai
*.machinetranslation.com
*.maester.app
*.maestra.ai
*.mage.space
*.magicregexgenerator.xyz
*.magicslides.app
*.magicstudio.com
*.magisto.com
*.magnific.ai
*.mailshake.com/ai-email-writer
*.makeform.ai
*.mammouth.ai
*.mapify.so
*.markcopy.ai
*.marketingblocks.ai
*.marketingcopy.ai
*.masterpiecestudio.com
*.mebook.ai
*.meetmaya.world
*.meetz.ai
*.memfree.me
*.memorylane.co
*.mentara.io
*.merlin.foyer.work
*.meta.ai
*.micvoice.ai
*.midjourney.co
*.mieux.ai
*.mightygpt.com
*.mimicpc.com
*.mind.am
*.mindmapai.app
*.mintoai.com
*.miramuseai.net
*.modulify.ai
*.moonvalley.com
*.mosaicml.com
*.motiff.cn
*.mubert.com
*.mureka.ai
*.murf.ai
*.muse.fun
*.musely.ai
*.musicful.ai
*.myaiteam.com
*.myaskai.com
*.mygptwizard.com
*.mykin.ai
*.myninja.ai
*.mytales.io
*.myths-of-gallantium.pressdeck.io
*.namesnack.com
*.nando.ai
*.napkin.ai
```

```
*.narakeet.com
*.narrationbox.com
*.nat.dev
*.naturalreaders.com
*.netcrumb.com
*.netus.ai
*.neural.love
*.neuraltext.com
*.newswriter.ai
*.nextbrain.ai
*.noluai.com
*.notebooklm.cloud.google.com
*.notion.so
*.novelai.net
*.noxus.cc
*.nsfwai.pro
*.numind.ai
*.ocoya.com
*.octocom.ai
*.office-ai.cn
*.ohai.ai
*.ohmygpt.com
*.oji.ai
*.ollama.ai
*.olympia.chat
*.omagic.ai
*.omneky.com
*.openai.cha-tai.cn
*.openai.com
*.openart.ai
*.openevidence.com
*.opus.pro
*.opusclip.com
*.ora.ai
*.ordermaid.ai
*.ossa.ai
*.outerbase.com
*.outwrite.com
*.pagelines.com
*.palmate.ai
*.paperpal.com
*.paragraphai.com
*.parallelgpt.ai
*.paraphrasingtool.ai
*.partyrock.aws
*.pdf.wondershare.com
*.pencilsketchgenerator.com
*.peoplealsoask.ai
*.peppercontent.io
*.perfectessaywriter.ai
*.perfecttense.com
*.perplexity.ai
*.personal.ai
*.petal.org
*.pfpmaker.com
```

```
*.phaii.net
*.phedra.ai
*.phind.com
*.photoaistudio.com
*.photoleapapp.com
*.photomath.com
*.photomyne.com
*.photopacks.ai
*.phrasee.co
*.pica-ai.com
*.piccopilot.com
*.pickaxeproject.com
*.piclumen.com
*.picmaker.com
*.pictory.ai
*.pika.art
*.pineapplebuilder.com
*.ping.tm
*.piufoto.com
*.pixai.art
*.pixelcut.ai
*.pixlr.com
*.pixu.ai
*.play.ai
*.play.ht
*.playgroundai.com
*.plusai.com
*.podcast.adobe.com
*.podcastle.ai
*.podnotes.app
*.podpally.com
*.poe.com
*.pollo.ai
*.polybuzz.ai
*.polylm.com
*.popai.pro
*.portraitpal.ai
*.postwise.ai
*.predis.ai
*.premai.io
*.prepai.io
*.preppally.com
*.presentationgpt.com
*.presentations.ai
*.prezo.ai
*.print-ai-masterpiece.com
*.privatellm.app
*.proai.co
*.productai.photo
*.profilepicture.ai
*.projectdecember.net
*.promeai.pro
*.prometai.app
*.promptclub.ai
*.promptfolder.com
```

```
*.prompthero.com
*.prompthunt.com
*.promptitude.io
*.promptkit.tools
*.promptpictures.tv
*.prophetes.ai
*.punchlines.ai
*.pyxl.pro
*.qianlu.cc
*.qopywriter.ai
*.questmate.com
*.quickchat.ai/emerson
*.quicktok.ai
*.quillbot.com
*.quizme.live
*.qwen.ai
*.qwenlm.ai
*.raita-kun.arvo.net
*.randomgenerate.io
*.rask.ai
*.rawshorts.com
*.read.ai
*.read2me.online
*.rebecc.ai
*.refraction.dev
*.regem.in/ai-writer
*.regex.ai
*.regie.ai
*.regisai.com
*.reimaginehome.ai
*.relevanceai.com
*.remini.ai
*.rencoach.com
*.rendair.ai
*.rendernet.ai
*.rendora.ai
*.replicate.com
*.replika.com
*.replit.com
*.resemble.ai
*.retellio.com
*.retresco.de
*.revid.ai
*.revoicer.com
*.riku.ai
*.rizzai.com
*.roamaround.app
*.roomdesigner.ai
*.room-genius.com
*.rootly.com
*.rovioai.com
*.rtst.ai
*.ru-chatgpt.ru
*.runwayml.com
*.ryrob.com/paragraph-rewriter
```

```
*.rytr.me
*.saffronintell.com
*.saga.so
*.sakubun.ai
*.salesforge.ai
*.samwell.ai
*.sapling.ai
*.sassbook.com
*.saufter.io
*.scholarcy.com
*.scholarlyai.com
*.scifocus.ai
*.scisummary.com
*.scite.ai
*.scribblediffusion.com
*.scribbyo.com
*.sdxlturbo.ai
*.seaart.ai
*.seaart.me
*.search.hebbia.ai
*.searchgpt.net
*.segment-anything.com
*.segmind.com
*.sensibleaimanifesto.com
*.seowriting.ai
*.serif.ai
*.shakespearetoolbar.com
*.shakker.ai
*.sheetai.app
*.sheetgpt.ai
*.shopia.ai
*.shortlistiq.com
*.shortlyai.com
*.shyuhuankj.com
*.sieve.video
*.sievedata.com
*.simbla.com
*.simpleshow.com
*.simpliciter.ai
*.simplified.com
*.sindresorhus.com/aiko
*.sindresorhus.com/amazing-ai
*.sindresorhus.com/quickgpt
*.sinonimos.com.br
*.sintra.ai
*.site.xaris.ai
*.sitegpt.ai
*.skybox.blockadelabs.com
*.skyengine.ai
*.skylum.com
*.skyvern.com
*.slazzer.com
*.slider-ai.ru
*.slidesgpt.com
*.smart-copy.io
```

```
*.smartlifeskills.ai
*.smartling.com
*.smartsolve.ai
*.smartwriter.ai
*.snapheadshots.com
*.snowpixel.app
*.sofer.ai
*.sologo.ai
*.soloist.ai
*.sonix.ai
*.sonofa.ai
*.soofy.io
*.soulgen.net
*.soundful.com
*.soundraw.io
*.sourcetable.com
*.soynerdito.com
*.speechgen.io
*.speechify.com
*.speechmatics.com
*.spicychat.ai
*.spikes.studio
*.splice.com
*.stability.ai
*.stableaudio.com
*.stablediffusion.com
*.stablediffusionweb.com
*.staccato.ai
*.starbyface.com
*.starryai.com
*.steve.ai
*.stockimg.ai
*.stork.ai
*.storyblocker.com
*.storybrand.ai
*.storyd.ai
*.storynest.ai
*.storyscapeai.app
*.storytailor.ai
*.streamlit.io/generative-ai
*.strella.io
*.studio.tune.app
*.studytexter.com
*.studytexter.de
*.styleadvice.ai
*.subscribr.ai
*.sudowrite.com
*.summarizer.org
*.suno.ai
*.suno.com
*.suno-ai.org
*.supafriends.com
*.supercraft.ai
*.supermeme.ai
*.svg.io
```

```
*.swiftspec.ai
*.switchboard.ai
*.sydologie.ai
*.syllaby.io
*.symbolab.com
*.synesthesia.io
*.synthesys.io
*.t3.chat
*.tabnine.com
*.talkai.info
*.talkie-ai.com
*.talkpal.ai
*.talkstack.ai
*.talktext.io
*.talktowalle.com
*.tana.inc
*.taplio.com
*.teachablemachine.withgoogle.com
*.tealhq.com
*.team-gpt.com
*.teammately.ai
*.telborg.com
*.template.net
*.tensorart.green
*.tethered.dev
*.texta.ai
*.textadviser.com
*.textalky.com
*.textcortex.com
*.texteo.ai
*.textie.ai
*.text-summary.com
*.textsynth.com
*.textwise.qutesoft.com
*.texxto.ai
*.theogrowth.com
*.there.do
*.thesummary.ai
*.thirdai.com/pocketllm
*.thispersonnotexist.org
*.thisresumedoesnotexist.com
*.threadcreator.com
*.thunderbit.com
*.thundercontent.com
*.tinyalbert.ai
*.tinywow.com
*.tldrthis.com
*.tome.app
*.tool.tool123.fun
*.toolhouse.ai
*.tools.picsart.com
*.toolsaday.com
*.topazlabs.com
*.to-teach.ai
*.transfercompass.com
```

```
*.trint.com
*.tryblend.ai
*.trypencil.com
*.trypromptly.com
*.ttsmaker.com
*.tugan.ai
*.tungsten.run
*.turbolearn.ai
*.turbular.com
*.tutorai.me
*.tweetyai.com
*.twoslash.ai
*.txtplay.ai
*.typecast.ai
*.typeface.ai
*.typeright.com
*.typli.ai
*.uberduck.ai
*.udio.com
*.uimagic.io
*.uizard.io
*.uminal.com
*.unbounce.com
*.uncode.vercel.app
*.undatas.io
*.undress.vip
*.unifygpt.ai
*.uniglobalcareers.com
*.unsloth.ai
*.upscale.media
*.upscayl.org
*.useadrenaline.com
*.usechat.ai
*.usegalileo.ai
*.usemano.com
*.usemotion.com
*.v0.dev
*.vanceai.com
*.vectorize.io
*.vectorizer.ai
*.veed.io
*.venice.ai
*.vertexaisearch.cloud.google
*.vgu.community
*.vidby.com
*.videoblog.ai
*.videocandy.com
*.videomuse.app
*.vidext.app
*.vidu.studio
*.vidyard.com
*.vidyo.ai
*.viggle.ai
*.vinteo.ai
*.virtualghostwriter.com
```

```
*.visualelectric.com
*.vivago.ai
*.vizcom.ai
*.vizly.fyi
*.vmfi.net
*.vobox.io
*.vocalid.ai
*.vogent.ai
*.voicemaker.in
*.voicemod.net
*.voicenotes.com
*.voiceovermaker.io
*.voyageai.com
*.vozfly.com
*.walles.ai
*.wan.video
*.wand.ai
*.warmer.ai
*.warpvideo.ai
*.watermarkremover.io
*.watermelon.ai
*.wavoai.com
*.webllm.ai
*.webllm.mlc.ai
*.websim.ai
*.webspellchecker.com
*.webtastic.ai
*.weel.co.jp
*.wellsaidlabs.com
*.wepik.com
*.we-spots.com
*.whatsgoodapps.com
*.whispertranscribe.com
*.wireply.ai
*.wishesai.com
*.wisprflow.ai
*.witsyai.com
*.wnr.ai
*.wolframalpha.com
*.wordai.com
*.wordhero.co
*.wordplay.ai
*.word-spinner.com
*.wordtune.com
*.wordybot.ai
*.workflos.ai
*.woxo.tech
*.wpaibot.com
*.wpkubio.com
*.writecream.com
*.writefull.com
*.writegpt.ai
*.writeholo.com
*.writeme.ai
*.writer.com
```

```
*.writersbrew.app
*.writerx.co
*.writesonic.com
*.writesparkle.ai
*.writier.io
*.wrizzle.ai
*.wui.ai
*.www.getlid.co
*.xccelerateai.com
*.xerac.ir
*.xiaoin.com.cn
*.xiaoshuoai.com
*.xinva.ai
*.xlscout.ai
*.xn--e1aalohlij4g.net
*.yaara.ai
*.yarnit.app
*.yeschat.ai
*.yeting.fun
*.yiyan.baidu.com
*.yodayo.com
*.youchuan.cn
*.yourjoi.com
*.zevbot.com
*.zipwp.dev
*.zoc.ai
*.zomani.ai
*.zomgsounds.com
*.zonerai.com
```

Last updated on 02/27/2026

Learn about Data Security Posture Management - (classic)

> **Note**
>
> This article is for the **classic** version of Data Security Posture Management that's being replaced with a new version that introduces guided workflows for proactive risk management, and streamlines data security operations so you can more confidently adopt AI across your digital estate.
>
> These improvements won't be added to this classic version so we invite you to try the new Data Security Posture Management, currently in preview.

Strengthening data security and minimizing exposure to data risks for sensitive information is often challenging for organizations in the modern workplace. The increasing complexity of data, the variety of data sources and platforms, limited visibility into sensitive data, and fragmented security solutions can pose significant challenges for administrators and data security professionals. Adding multicloud platforms and generative AI applications to these areas makes it even more difficult to assess data security coverage and to correlate insights from user data points. To help discover and mitigate data risks, organizations must address the following questions:

- What is my sensitive data?
- Where is it located?
- What data is currently unprotected?
- How is unprotected sensitive data being handled and accessed?
- How can I help lower the risk and help secure unprotected sensitive data?

Microsoft Purview Data Security Posture Management (DSPM) enables you to quickly and easily monitor cross-cloud data and user risk through dynamic reports and trend analysis. By processing and correlating data across other Microsoft Purview data security and risk and compliance solutions, DSPM helps you identify vulnerabilities with unprotected data and quickly take action to help you improve your data security posture and minimize risk. DSPM provides:

- **Data security recommendations**: Gain insights into your data security posture and get recommendations for creating Insider Risk Management and data loss prevention (DLP) policies to help protect sensitive data and to close data security gaps. For example, some recommendations might include creating policies to prevent users from printing sensitive files or to prevent users from copying sensitive files to other network locations.

- **Data security analytic trends and reports**: Track your organization's data security posture over time with reports summarizing sensitivity label usage, DLP policy coverage, changes in risky user behavior, and more.
- **Microsoft Security Copilot**: Use Security Copilot promptbooks and custom prompts to help you investigate alerts, identify risk patterns, and pinpoint the top data security risks in your organization.

Integration with Microsoft Purview solutions

DSPM processes and correlates data state, signals, and user activities based on the current configuration of other data security and risk and compliance solutions in Microsoft Purview. For optimal coverage and deeper insights, consider using the features and capabilities of the following solutions:

- **Data loss prevention (DLP)**: Microsoft Purview Data Loss Prevention policies help you prevent users from inappropriately sharing sensitive data. DLP detects sensitive information by using content analysis that includes keyword matching, evaluations of expressions, internal validation, machine learning algorithms, and more.

  Depending on the specific recommendations for the data, you can choose to quickly create an applicable DLP policy directly in the DSPM workflow.

- **Information protection**: Microsoft Purview Information Protection provides a framework, process, and capabilities you can use to protect sensitive data across clouds, apps, and devices. By using sensitivity labels, trainable classifiers, and sensitive information types, you can define and apply protection policies to sensitive data.

- **Insider risk management**: Microsoft Purview Insider Risk Management uses the full breadth of built-in service and third-party indicators to help you quickly identify, triage, and act on potentially risky activity by users in your organization. By using logs from Microsoft 365 and Microsoft Graph, Insider Risk Management allows you to define specific policies to identify risk indicators. After identifying the risks, you can take action to mitigate these risks, and if necessary open investigation cases and take appropriate legal action.

  Depending on the specific recommendations for the data, you can choose to quickly create an applicable Insider Risk Management policy directly in the DSPM workflow.

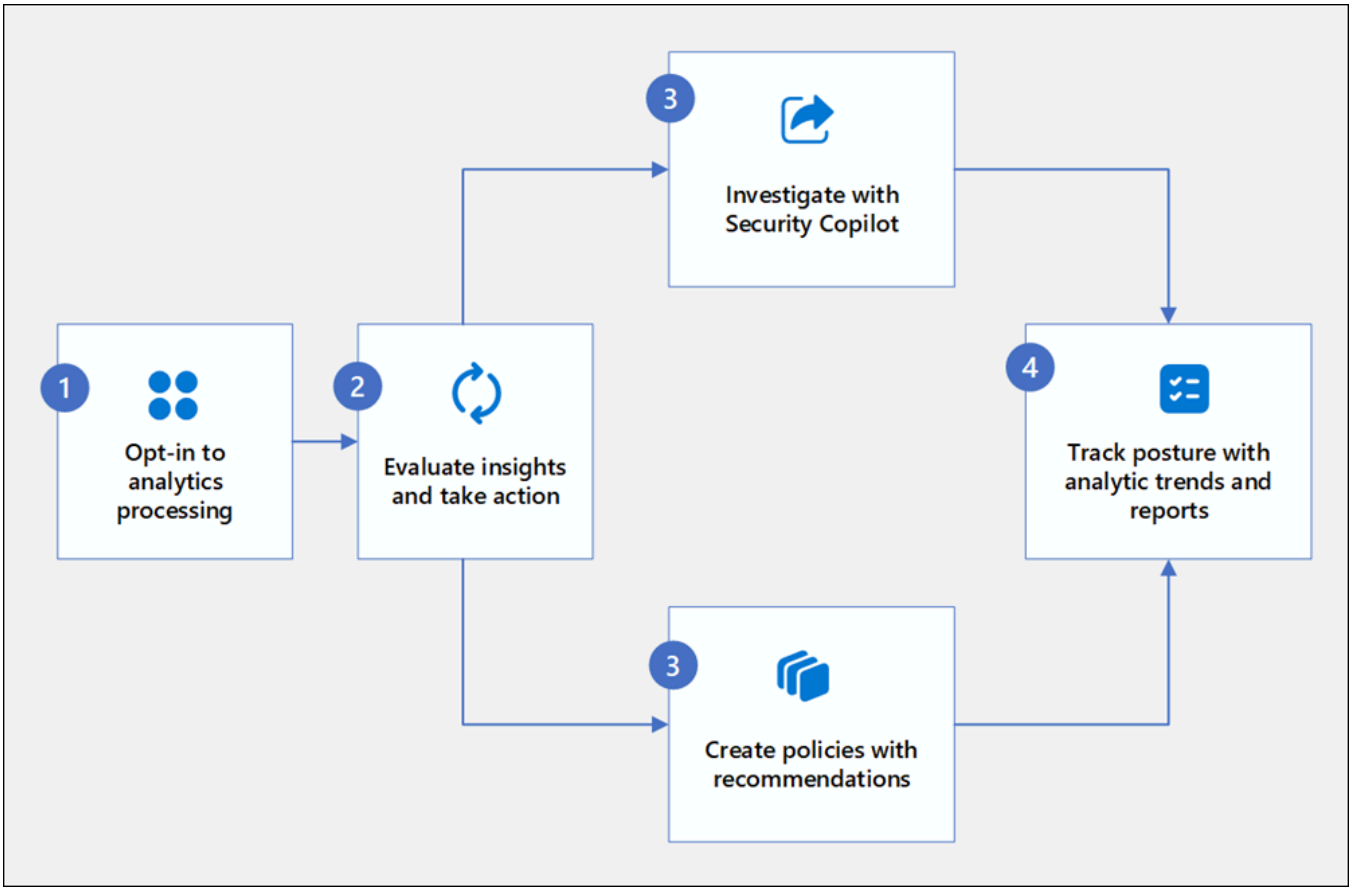
Working with Adaptive Protection

Adaptive protection in Microsoft Purview uses machine learning to identify the most critical risks and proactively and dynamically apply protection controls from data loss prevention, data

lifecycle management, and Microsoft Entra conditional access. If Adaptive Protection isn't currently enabled for unprotected sensitive data, you can create a new Adaptive Protection policy directly in the DSPM workflow.

Workflow

The DSPM workflow helps you investigate and take action to address potential security concerns with unprotected data across your organization.



1. **Opt in to analytics processing**: To get started with DSPM, enable and opt in to:
   a. Insider risk management analytics
   b. DLP analytics
   c. Analytics processing in DSPM to scan for unprotected data in your organization.
2. **Evaluate insights and take action**: After the automated analytics processing completes, evaluate the insights created by DSPM to help mitigate risks for unprotected data.
3. **Actions**:
   a. **Investigate with Security Copilot**: Use built-in promptbooks and custom prompts with Security Copilot to help identify specific areas of risk.
   b. **Create policies with recommendations**: Use recommendations to quickly create Insider Risk Management and DLP policies to help mitigate data security risks for unprotected data assets.

4. **Track posture with analytic trends and reports**: Use analytic trends and reports to view your posture over time and for data locations across your organization.

Using DSPM (classic) in new environments

For organizations just getting started with Microsoft Purview, DSPM helps provide easy configuration and new policy creation across several data security and risk and compliance solutions.

By automatically scanning data and activities across your organization, you quickly get baseline insights and recommendations focused on unprotected data that can help you quickly get started in DLP, information protection, and Insider Risk Management.

Using data security in existing environments

If your organization is already using DLP, information protection, and Insider Risk Management in Microsoft Purview, DSPM helps identify any gaps in your current policy coverage.

Without having to conduct deeper dives into your existing policies, DSPM can quickly determine and highlight any unprotected data at risk, saving you the time and effort associated with reviewing and testing numerous individual data security and risk and compliance policies.

Ready to get started?

- Get started with Data Security Posture Management - (classic)
- Take action with Data Security Posture Management recommendations - (classic)
- Use Microsoft Security Copilot with Data Security Posture Management - (classic)
- Use Data Security Posture Management analytics trends and reports - (classic)

Last updated on 12/15/2025

Get started with Data Security Posture Management - (classic)

> **Note**
>
> This article is for the **classic** version of Data Security Posture Management that's being replaced with a new version that introduces guided workflows for proactive risk management, and streamlines data security operations so you can more confidently adopt AI across your digital estate.
>
> These improvements won't be added to this classic version so we invite you to try the new Data Security Posture Management, currently in preview.

Use Data Security Posture Management (DSPM) to quickly identify unprotected sensitive data assets and potentially risky user activities in your organization. Complete the steps in this article to get started with the DSPM workflow.

Subscriptions and licensing

Before getting started with DSPM, confirm your Microsoft 365 subscription and any add-ons. Depending on the licensing and subscriptions for your organization, you can access other data security and risk and compliance solutions in the Microsoft Purview portal.

To access and use DSPM, you need a Microsoft 365 E5 or Microsoft Purview Suite (formerly known as Microsoft 365 E5 Compliance) license. Administrators need to verify that their organization has a supported subscription and assign the appropriate licenses to users. For more information about subscriptions and licensing, see the subscription requirements for DSPM.

Step 1: Assign permissions

> **Important**
>
> If your permissions are restricted in specific administrative units in your organization, you can't access DSPM.

To continue with these configuration steps, you must be assigned to one of the following roles or role groups:

- *Data Security Management* role group
- *Data Security Viewer* role. If the DSPM dashboard is opted in, this role grants permission to view all the insights. This role is required to use Security Copilot in DSPM.
- Insider Risk Management Admins role
- Microsoft Entra ID *Global Administrator* role
- Microsoft Entra *Compliance Administrator* role

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview **roles and permissions**.

Step 2: Opt in to analytics processing and scanning

To get started with DSPM, you need to enable analytics in Insider Risk Management and data loss prevention (DLP). When you opt into DSPM, the system automatically enables analytics in Insider Risk Management and analytics in DLP if they're not already active. If you enable analytics in Insider Risk Management and analytics in DLP, the system automatically enables DSPM.

For more information about analytics in these solutions, see:

- Enable analytics in Insider Risk Management
- Enable analytics in DLP

When you complete the opt-in process, the system starts an automated scan of data and user activities for your organization. Scanning times vary based on the size of your organization and the amount of data and activities that need processing. It might take up to three days for this processing to complete.

Step 3: Configure Security Copilot

To access Copilot in DSPM, you must configure your organization for Security Copilot. For step by step guidance, see Get started with Microsoft Copilot for Security.

Step 4: Evaluate insights and take action

Review and take action with recommendations

DSPM generates recommendations directly from the processed data, user activities, and current state of unprotected assets in your organization. Specific recommendations let you take action and quickly create data loss prevention (DLP) and Insider Risk Management policies to help you mitigate data security risks.

For more information about recommendations, see Take action with Data Security Posture Management recommendations - (classic).

Use Microsoft Security Copilot investigate data security details

Use Security Copilot promptbooks and custom prompts to quick dive into the details and get answers about unprotected sensitive data assets and potentially risky user activities in your organization. Data security insights are generated from scanned data across data loss prevention (DLP), information protection, and Insider Risk Management solutions.

For more information about using Security Copilot, see Use Microsoft Security Copilot with Data Security Posture Management - (classic).

Step 5: Review analytics and trends to track your data security posture

DSPM analytics trends and reports help you quickly see the unprotected and protected sensitive assets and potentially risky user activities in your organization.

For more information about using trends and reports, see Use Data Security Posture Management analytics trends and reports - (classic).

Last updated on 12/15/2025

Take action with Data Security Posture Management recommendations - (classic)

> **Note**
>
> This article is for the **classic** version of Data Security Posture Management that's being replaced with a new version that introduces guided workflows for proactive risk management, and streamlines data security operations so you can more confidently adopt AI across your digital estate.
>
> These improvements won't be added to this classic version so we invite you to try the new Data Security Posture Management, currently in preview.

Data Security Posture Management (DSPM) generates recommendations directly from the processed data, current state of unprotected sensitive assets in your organization, and user activities that put the unprotected sensitive assets at risk. Specific recommendations enable you to take action and quickly create data loss prevention (DLP) and Insider Risk Management policies to help mitigate data security risks. DSPM recommendations can also help you identify coverage gaps in existing Insider Risk Management and DLP policies.

Using recommendations

To display and review recommendations, go to **Data Security Posture Management (classic)** > **Overview** or **Data Security Posture Management (classic)** > **Recommendations**.

- From the **Overview** page, you can review the top two data security recommendations or select **View all recommendations** for a complete listing of all recommendations.
- From the **Recommendations** page, you can directly review a complete listing of all recommendations.

Recommendations come from the last 30 days of user activity and the state of unprotected sensitive assets. As processing continues, the list of recommendations automatically updates and removes recommendations older than 30 days.

Each recommendation provides a short description of the risky activity or state of the unprotected sensitive asset and includes a recommended policy to configure to help mitigate ongoing and future data security risks. The recommendation displays metrics for the number of activities, the associated type of protection, and the number of users involved.

Select **View recommendation** to take action and create a new policy to help mitigate the data security risks associated with the recommendation.

Creating policies from recommendations

After selecting **View recommendation** for a specific recommendation, you can create one or more policies in data loss prevention (DLP) and/or Insider Risk Management to help mitigate the data security risks associated with the recommendation.

For example, if you have a recommendation to prevent users from copying sensitive files to network shares, the recommendation might be to create both an Insider Risk Management policy to detect when certain users perform risky activities containing sensitive data and a DLP policy to prevent those same users from sharing the data with others.

To create policies from a recommendation, complete the following steps:

1. Go to the Microsoft Purview portal and sign in with the credentials for a user account assigned DSPM permissions.
2. Select the **Data Security Posture Management (classic)** solution card, then select **Recommendations** in the left nav.
3. Select a recommendation, then select **View recommendation**.
4. On the recommendation flyout pane, choose the option to create either one type of policy or multiple policies. By default, if more than one policy applies to the recommendation, the option to create multiple policies is selected by default.
5. Each policy type has a dedicated section for configuring the policy options:
6. For the **Insider Risk Management policy** section options, configure the following policy options:
   a. **Policy name**: Enter a unique policy name or accept the suggested policy name.
   b. **User scope**: Select **Include all users and groups (recommended for best coverage)** or **I want to choose specific users and groups**. If you choose to include specific users or groups, enter the user and group names in the picker field.
   c. **Settings we filled in for you**: The settings for the policy are automatically scoped to the type of activity and unprotected data asset insights associated to the recommendation. These settings include specific policy templates, triggering events, and indicators. If you need to change the suggested settings, select **Customize** > **Insider risk policy** to create the policy in the Insider Risk Management policy workflow.
7. For the **Data Loss Prevention policy** section options, configure the following policy options:
   a. **Policy name**: Enter a unique policy name.
   b. **Mode**: Select **Simulation** or **On**.
   c. **Settings we filled in for you**: The settings for the policy are automatically scoped to the type of activity and unprotected data asset insights associated to the recommendation. These settings include specific sensitive information types, enforcement actions, triggering events, and data locations. If you need to change the

suggested settings, select **Customize** > **DLP policy** to create the policy in the DLP policy workflow.

8. If the suggested policy options meet your needs, select **Create policy**. This step creates one or more policies in the applicable solutions.

It takes a few minutes to create each policy. The new policies appear on the **Policies** tab in the applicable solutions. Once the policy is active, it could take at least 24 hours for the triggering event to occur and score user activity, then the first alert is generated. If administrator notifications are turned on, you get an email when this alert occurs.

You can triage the alert and confirm it to a case for further investigation or dismiss it as normal behavior. After you review a few alerts, fine tune your policy to control how many alerts are generated, what activities are detected, and more. Additional recommendations might be generated for this type of activity that can help you with policy updates.

Updating policies in other solutions

To update policies you created from recommendations, use the policy management tools in each solution. The tools in the solutions let you fully customize policies beyond the quick policy settings included with recommendations.

- Manage Insider Risk Management policies
- Manage DLP policies

Verifying the configuration of existing policies

If your organization already has Insider Risk Management or DLP policies configured, recommendations in DSPM help you identify and correct any gaps in your current policy configurations. Instead of using the policy creation options directly in the recommendation, you might choose to use the insights associated with the recommendation to update or refine an existing Insider Risk Management or DLP policy.

Last updated on 12/15/2025

Use Data Security Posture Management analytics trends and reports - (classic)

> **Note**
>
> This article is for the **classic** version of Data Security Posture Management that's being replaced with a new version that introduces guided workflows for proactive risk management, and streamlines data security operations so you can more confidently adopt AI across your digital estate.
>
> These improvements won't be added to this classic version so we invite you to try the new Data Security Posture Management, currently in preview.

Data security posture management (DSPM) analytics trends and reports help you quickly view the unprotected and protected sensitive assets and potentially risky user activities in your organization.

Working with reports

For each report, use the following features to help filter, review, evaluate, and export DSPM insights:

- **Charts**: Hover on specific areas of report information for metrics and more details about insights.
- **Customize columns**: Select **Customize columns** to add or remove columns and associated information from a report.
- **Export**: Select **Export** to create a .csv file that contains the values included in a report.

Analytics reports

- **Unprotected sensitive assets across data sources**: This report displays the location of sensitive assets containing one or more classifiers (sensitive information types, exact data match classifiers, trainable classifiers) and:
  - Aren't protected by a data loss prevention (DLP) policy that restricts exfiltration activities.
  - Don't have a sensitivity label applied that controls access.

- **Users performing top risk-related activities on unprotected sensitive assets**: This report displays the number of users flagged by Insider Risk Management (departing users,

potentially disgruntled users, high risk users, potential high impact users, potentially risky users) who recently performed activities containing sensitive assets that:

- Contain any type of classifier (sensitive information types, exact data match classifiers, trainable classifiers) that aren't protected by a DLP policy that blocks access or don't have a sensitivity label applied that control access.

Trends

You can view trends on the **Overview** page. They help you see the recent history of sensitive data assets and user activities in your organization. Trends help you monitor your data security posture over time and track the status of protected assets. When you hover over specific points in the trends, you can view details about that point in time or area.

- **Sensitive assets labeled (automatically + manually)**: Percentage of assets per week that have sensitivity labels applied, either manually by users or automatically by auto-labeling policies.
- **Sensitive assets protected by at least one DLP policy**: Percentage of sensitive assets in your organization that contain any type of classifier (sensitive information types, exact data match classifiers, trainable classifiers) and are protected by at least one DLP policy.
- **Potentially risky users**: Number of users per week who are assigned insider risk severity levels (*Low*, *Medium*, *High*) based on activity detected by Insider Risk Management policies.

Last updated on 12/15/2025

Use Microsoft Security Copilot with Data Security Posture Management - (classic)

> **Note**
>
> This article is for the **classic** version of Data Security Posture Management that's being replaced with a new version that introduces guided workflows for proactive risk management, and streamlines data security operations so you can more confidently adopt AI across your digital estate.
>
> These improvements won't be added to this classic version so we invite you to try the new Data Security Posture Management, currently in preview.

Use Microsoft Security Copilot and Data Security Posture Management (DSPM) to quickly explore the details and get answers about unprotected sensitive data assets and potentially risky user activities in your organization. Data security insights come from scanned data across Data Loss Prevention (DLP), Information Protection, and Insider Risk Management solutions in Microsoft Purview.

Get started with Copilot

After you configure DSPM, onboard your organization to Security Copilot, and complete the automated scanning, use Security Copilot featured promptbooks and the new prompt gallery to help you get started. Promptbooks are a built-in sequence of prompts that help you quickly learn more about your data security posture. Choose from the following Security Copilot promptbooks or built-in prompts in the Copilot Prompt Gallery:

Risky user investigation promptbook

This promptbook is a six-prompt sequence that helps you investigate users handling sensitive data and shows their data activities, anomalies, and related alerts. By providing the User Principal Name (UPN) for a user and a duration (in days), this promptbook automatically runs the following prompts in order:

1. Show all sensitive data activities performed by <upn> in the last <duration> days
2. Was <upn> involved in any potential sensitive data exfiltration (for example, email forwarding, external file sharing, USB transfers, cloud uploads) in the last <duration> days?
3. Summarize <upn>'s sensitive data interactions over the last <duration> days, highlighting the most accessed classifiers, labels, SharePoint sites, common upload

domains, and primary email recipient domains and users.
4. Did <upn> exhibit unusual behavior or take uncommon actions like excessive access or downloads in the last <duration> days?
5. Are there any alerts associated with <upn> in the last <duration> days, and what is the user's current risk level?
6. What actions can be taken to prevent <upn> from leaking sensitive data? Include policies, data loss prevention controls, and access restriction strategies.

Sensitive data protection promptbook

This promptbook is a sequence of six prompts that helps you identify and protect sensitive data across your organization. It suggests recommended policy changes and data loss prevention rules. By using the full name of the trainable classifier, sensitivity label, or sensitive information type (SIT) and a duration in days, this promptbook automatically runs the following prompts in order:

1. Where is data labeled as <label_or_classifier_or_SIT> stored?
2. Provide an overview of activities involving <label_or_classifier_or_SIT> data in the last <duration> days.
3. Identify instances where <label_or_classifier_or_SIT> data was transferred outside of the organization in the last <duration> days.
4. Who are the top five users with the most <label_or_classifier_or_SIT> data exfiltration in the last <duration> days?
5. Are there any alerts for users who interacted with <label_or_classifier_or_SIT> data in the last <duration> days?
6. How can I prevent unauthorized transfers of <label_or_classifier_or_SIT> data?

To get started with Security Copilot promptbooks, complete the following steps:

1. Go to the Microsoft Purview portal and sign in with the credentials for a user account assigned DSPM permissions.

2. Select the **Data Security Posture Management (classic)** solution card, then select **Overview** in the left nav.

3. Select one of the suggested prompts for Security Copilot:

    - Risky user investigation promptbook
    - Sensitive data protection promptbook

4. Enter the requested inputs for the user, classifier, sensitivity label, or SIT and duration in days.

5. Select **Submit**.

The promptbook responses automatically scope insight data and provide quick answers in a separate flyout pane. You can select additional built-in prompts to automatically update and generate new responses in the flyout pane. Select **New chat** to clear previous responses to suggested prompts. Create additional custom prompts directly in Copilot to generate responses from AI-driven analytics based on the scanning results from your organization.

Copilot prompt gallery

Data Security Posture Management also offers built-in prompts to help you investigate sensitive areas, potentially risky users, potentially suspicious activity, and more in your organization. To view these built-in Copilot prompts, select **View more in prompt gallery** in the **Get started with Security Copilot featured promptbooks** section on the **Overview** page. You can customize the built-in prompt or immediately run the prompt directly in Copilot.

Select built-in Copilot prompts from the following categories:

- Alerts and policies
- Data at risk
- Potentially risky users
- Potentially suspicious activity
- Sensitive data

Tips for custom Copilot prompts in DSPM

For an enhanced experience with Security Copilot in Data Security Posture Management, use the following tips for higher accuracy for your Copilot responses:

- Always include the user's UPN for questions involving a specific user.
- Always specify the complete name for the sensitive info type or label for questions that involve a specific type of sensitive info type or label.
- Clearly list the sorting criteria for questions for top users, activities, and alerts.
- Always specify the date period for questions for data in a specific date period. If you don't specify a date period, only data from the last 10 days from current date is included. The maximum lookback is 30 days from the current date.
- Put all items (classifiers or labels) in single quotes in your prompt.
- Use "/" as a separator for any path (for example, a file path) in a user prompt.
- Scope the prompt to a single intent for higher accuracy of responses. Break complex prompts into single intent questions and enter the prompts one by one.
- Make questions self-contained. Avoid referring to previous questions or responses.
- Avoid using generic terms.

- Support prompts for data security across Information Protection, DLP, Insider Risk Management, or from public documentation.

For more information about creating Security Copilot prompts, see Create effective prompts.

Using Copilot in other solutions

Security Copilot is also available directly in other Purview solutions to help you quickly find answers for specific scenarios or to generate insights scoped to specific solution areas that aren't related to unprotected assets.

- Using Copilot to summarize an alert in Insider Risk Management
- Use Copilot to summarize user activities in Insider Risk Management

---

Last updated on 12/15/2025

Application card: Microsoft Purview Data Security Posture Management

What is an Application or Platform Card?

Microsoft's Application and Platform cards are intended to help you understand how our AI technology works, the choices application owners can make that influence application performance and behavior, and the importance of considering the whole application, including the technology, the people, and the environment. Application cards are created for AI applications and platform cards are created for AI platform services. These resources can support the development or deployment of your own applications and can be shared with users or stakeholders impacted by them.

As part of its commitment to responsible AI, Microsoft adheres to six core principles: fairness, reliability and safety, privacy and security, inclusiveness, transparency, and accountability. These principles are embedded in the Responsible AI Standard, which guides teams in designing, building, and testing AI applications. Application and Platform Cards play a key role in operationalizing these principles by offering transparency around capabilities, intended uses, and limitations. For further insight, readers are encouraged to explore Microsoft's Responsible AI Transparency Report and Code of Conduct, which outline how enterprise customers and individuals can engage with AI responsibly.

Overview

Microsoft Purview Data Security Posture Management is a data security application within the Microsoft Purview portal that helps organizations discover, protect, and investigate sensitive data risks across their digital estate. Rather than focusing on infrastructure or endpoints, Data Security Posture Management centers on the data itself, identifying where it resides, who can access it, how it's used, and whether it's adequately protected. It uses AI to analyze access patterns, sharing behaviors, and policy gaps in order to surface actionable risks and recommend remediation steps.

Data Security Posture Management addresses a growing challenge in today's AI-driven workplaces: the increasing complexity and volume of data makes it difficult for security teams to maintain visibility and control. The application solves this problem by consolidating insights from multiple Microsoft Purview solutions, including data loss prevention (DLP), Insider Risk

Management, information protection with sensitivity labels, and Data Security Investigations, into a single view for monitoring data risks, policy coverage, and posture trends. This unified approach replaces the need for multiple tools and manual audits, helping administrators close data security gaps more efficiently.

Data Security Posture Management is designed to perform reliably in enterprise environments where organizations need to monitor and protect sensitive data across multiple locations. The application includes an embedded Microsoft Security Copilot experience that allows these users to ask natural language questions about their data security posture, as well as AI agents from Security Copilot that can take guided remediation actions on detected risks. For an overview introduction, see the Microsoft Mechanics video: New Data Security Posture Management.

Key terms

The following list provides a glossary of key terms related to Data Security Posture Management:

AI observability: A feature within Data Security Posture Management that provides an inventory of AI apps and agents with recent activity, showing how many are high risk and the total with sensitive interactions, along with a breakdown of individual agents and the policies governing them.

Collection policy: A policy that captures prompts and responses from AI interactions for Copilot in Fabric and Security Copilot, and for non-Copilot AI apps, so they can be managed in Microsoft Purview solutions.

Copilot interactions: Prompts and responses from Copilots and agents that Microsoft Purview supports for data security and compliance protections. Includes Microsoft 365 Copilot, Security Copilot, Copilot in Fabric, Copilot Studio.

Data loss prevention (DLP): A Microsoft Purview solution that helps prevent users from inappropriately sharing sensitive data by using content analysis techniques including keyword matching, expression evaluation, and machine learning algorithms.

Data risk assessment: An automated or custom scan that identifies and helps fix potential data oversharing risks, particularly for SharePoint sites and Fabric workspaces. Default assessments run weekly for the top 100 sites based on usage.

Data security objectives: Guided workflows within Data Security Posture Management that each represent a specific security goal, such as preventing data exposure in Copilot interactions or

preventing oversharing. Each objective groups together relevant Microsoft Purview solutions and prioritized actions.

Data Security Posture Agent: A Microsoft Security Copilot agent exclusive to the current version of Data Security Posture Management that uses natural language search across files in SharePoint, OneDrive, Teams, Exchange, and Copilot interactions. It is designed for pre-investigation checks rather than formal cases.

Insider Risk Management: A Microsoft Purview solution that uses built-in service and third-party indicators to help identify, triage, and act on potentially risky activity by users in an organization.

One-click policy: A preconfigured policy within Data Security Posture Management that can be activated with a single click to quickly gain insights and protect data, without requiring manual policy configuration.

Promptbook: A built-in sequence of prompts for Microsoft Security Copilot that helps users quickly investigate specific data security scenarios, such as risky user behavior or sensitive data protection.

Security Copilot: Microsoft Security Copilot and its agents is a generative AI-powered security solution that helps data security professionals investigate and respond to security incidents. Within Data Security Posture Management, it provides an open-ended prompt experience for asking natural language questions about data security.

Sensitivity label: A label from Microsoft Purview Information Protection that can be applied to content to define and enforce protection policies for sensitive data across clouds, apps, and devices.

Key features or capabilities

The key features and capabilities outlined here describe what Data Security Posture Management is designed to do and how it performs across supported tasks.

- **Data security posture dashboard**: The central landing page that provides immediate access to key posture metrics, top objectives to address based on risk, a snapshot of data usage across the data estate, and a 30-day trending graph of the organization's data security posture. Users can interact directly with Security Copilot through suggested prompts on this page.
- **Data security objectives with guided workflows**: Selectable cards that each represent a specific security goal, such as preventing data exposure in Copilot interactions, preventing

oversharing of sensitive data, preventing exfiltration to risky locations, and discovering sensitive data. Each objective provides an end-to-end workflow with prioritized actions, one-click policies, and progress tracking, so users can focus on achieving outcomes rather than navigating separate solutions.

- **AI observability**: An inventory of AI apps and agents with activity in the last 30 days, showing how many are high risk and the total with sensitive interactions. Provides a breakdown of individual agents, their activities, and the policies governing them, enabling administrators to monitor risks such as oversharing, exfiltration, and unusual access patterns.

- **Data risk assessments**: Automated and custom scans that identify potential data oversharing risks. Default assessments run weekly for the top 100 SharePoint sites based on usage. Custom assessments can target specific users, sites, or Fabric workspaces. Results include remediation options such as restricting access by label, creating auto-labeling policies, or creating retention policies.

- **Data security recommendations and remediation actions**: Insights and recommendations generated from processed data that help administrators create or refine DLP and Insider Risk Management policies. Includes one-click policies for quick deployment directly from the Data Security Posture Management workflow.

- **Reports and analytics**: Tracks the organization's data security posture over time with reports summarizing sensitivity label usage, DLP policy coverage, risky user behavior, and AI app activity. Enhanced reporting provides advanced filtering, customizable views, and export capabilities to support compliance requirements.

- **Activity explorer**: Provides detailed visibility into content-related activity, including AI interactions (prompts, responses, and sensitive information detected), DLP rule matches, AI website visits, and sensitive information type detections. Supports filtering by workload categories such as Copilot experiences, enterprise AI apps, and other AI apps.

- **Data Security Posture Agent**: An AI-powered agent from Microsoft Security Copilot and that uses natural language search to find sensitive data across SharePoint, OneDrive, Teams, Exchange, and Copilot interactions. It provides item counts, sensitivity label classifications, and risk-level assessments, along with exportable insight reports. This agent is designed for quick, pre-investigation checks.

- **Embedded Security Copilot experience**: An open-ended prompt experience where users can ask natural language questions about their data security posture. Includes built-in

promptbooks for risky user investigation and sensitive data protection, as well as a prompt gallery with categorized prompts for alerts, data at risk, risky users, suspicious activity, and sensitive data.

- **AI agent-driven remediation**: Under user guidance, AI agents can take direct action on detected risks, such as removing public sharing links, applying DLP policies, or revoking permissions. AI-driven triage agents review alerts from DLP and Insider Risk Management, filtering noise and highlighting critical threats. All agent actions are audited and subject to user review and approval.

Intended uses

Data Security Posture Management can be used in multiple scenarios across a variety of industries. Some examples of use cases include:

- **Monitoring data security posture across the enterprise**: A compliance officer at a financial services organization uses Data Security Posture Management to get a unified view of their sensitive data across different locations. The posture dashboard surfaces key metrics and trends, enabling the officer to quickly identify where unprotected sensitive data exists and track improvements over time. This eliminates the need to manually review multiple separate tools and dashboards.

- **Preventing oversharing before deploying Microsoft 365 Copilot**: An IT administrator preparing for a Microsoft 365 Copilot deployment runs data risk assessments to identify SharePoint sites with potentially overshared content. Data Security Posture Management surfaces sites where sensitive files have broad sharing links or lack sensitivity labels, and provides one-click remediation options such as restricting access or creating auto-labeling policies. This helps the organization reduce data exposure before AI tools begin processing that content.

- **Investigating risky user behavior with AI-assisted analysis**: A data security analyst receives an alert about a user performing unusual file-sharing activities. Using the embedded Security Copilot experience, the analyst runs the risky user investigation promptbook, which automatically analyzes the user's sensitive data activities, identifies potential exfiltration patterns, checks for anomalies, and suggests protective actions, all through a sequence of natural language prompts.

- **Governing AI app usage across the organization**: A security team at a healthcare organization uses AI observability to monitor which AI apps and agents employees are

interacting with, which of those interactions involve sensitive patient data, and whether appropriate DLP policies are in place. The team uses data security objectives to create targeted policies that prevent sensitive health information from being shared with unauthorized AI applications.

- **Responding to compliance requirements**: A compliance team at a government agency uses Data Security Posture Management reports and activity explorer to demonstrate that sensitive data is appropriately labeled, protected by DLP policies, and monitored for risky activity. The export capabilities and customizable views allow the team to generate audit-ready documentation aligned with their regulatory obligations.

- **Discovering unprotected sensitive data in new environments**: An organization just getting started with Microsoft Purview uses Data Security Posture Management to automatically scan data and user activities, gaining baseline insights and recommendations focused on unprotected data. This helps the organization quickly establish DLP, information protection, and Insider Risk Management policies without conducting deep manual analysis.

Data Security Posture Management is not intended for use as a general-purpose AI assistant, for scenarios unrelated to data security within Microsoft Purview, or for formal legal investigations (which require dedicated solutions such as eDiscovery and Data Security Investigations).

Models and training data

Data Security Posture Management leverages existing AI agents and AI functionality. As such, it's dependent on any external configuration that these provide. For example, model choice for Microsoft Security Copilot for the embedded natural language prompt experience. Data Security Posture Management also uses content analysis within Data Loss Prevention for detecting sensitive information through keyword matching, expression evaluation, and machine learning algorithms. To learn more, refer to the linked application cards and documentation.

Performance

Data Security Posture Management is designed to perform reliably in enterprise environments where organizations need to monitor and protect sensitive data across multiple locations. The application processes and correlates data from existing Microsoft Purview solutions, including DLP, Insider Risk Management, and information protection, to generate insights, recommendations, and posture metrics. Performance is optimized for organizations using supported Microsoft cloud services, and results improve as more Microsoft Purview solutions are configured and more data is available for analysis.

The application accepts text-based inputs across its features. Users interact with Data Security Posture Management through the Microsoft Purview portal interface, where they can navigate dashboards, configure policies, and run data risk assessments. The embedded Security Copilot experience accepts natural language text prompts and returns text-based responses, including data summaries, risk analysis, and recommended actions. The Data Security Posture Agent accepts natural language search queries and returns structured results, including item counts, sensitivity label classifications, and risk-level assessments, with an option to export reports as Word documents.

Data Security Posture Management supports the languages available in the Microsoft Purview portal interface. The Security Copilot experience is designed and evaluated primarily for English-language prompts and responses. Sensitive information type detection supports multiple languages as documented in the individual sensitive information type definitions. Users should be aware that performance of the natural language prompt and search features may vary when using languages other than English.

After initial setup, Data Security Posture Management requires time to process organizational data. Scanning times vary based on the size of the organization and the amount of data and activities to process, and it may take up to three days for initial processing to complete. New policies require at least 24 hours to collect data before results appear. Data risk assessments require at least 48 hours after completion before updated results are available. The Data Security Posture Agent can search up to 1 GB of content per query, and restricted searches targeting specific users or sites perform significantly faster than tenant-wide searches.

Limitations

Understanding Data Security Posture Management's limitations is crucial to determine it is used within safe and effective boundaries. While we encourage customers to leverage Data Security Posture Management in their innovative solutions or applications, it's important to note that Data Security Posture Management was not designed for every possible scenario. We encourage users to refer to either the Microsoft Enterprise AI Services Code of Conduct (for organizations) or the Code of Conduct section in the Microsoft Services Agreement (for individuals) as well as the following considerations when choosing a use case:

- **Data security scope**: The Security Copilot experience in Data Security Posture Management is designed to answer questions about data security associated with Insider Risk Management, Information Protection, and Data Loss Prevention in Microsoft Purview. All other Purview solutions are currently out of scope for Data Security Posture Management

insights. Users should direct questions outside of these areas to the appropriate solution-specific tools.

- **Data coverage**: The asset explorer Standard tab is filtered by Microsoft and non-Microsoft locations. Microsoft locations currently include Microsoft 365 only. Non-Microsoft locations are made possible by integration with partner solutions. Organizations with significant data in environments that aren't included should be aware that visibility through the asset explorer may be limited and should supplement with other tools as needed.

- **Data risk assessment volume limits**: A maximum of 200,000 items per location applies to both custom and default data risk assessments. The reported file count may not be accurate when there are more than 100,000 files per location. Organizations with very large SharePoint sites should consider running multiple targeted assessments. Additionally, OneDrive is not currently supported for item-level scanning, and custom assessments support a maximum of 10 SharePoint sites for item-level scanning.

- **Processing time requirements**: After a custom data risk assessment completes, results require at least 48 hours to become available and do not update after that point. A new assessment is needed to see changes. New policies require at least 24 hours before data appears. Users should plan ahead and avoid expecting real-time results from newly created assessments or policies.

- **Activity explorer data completeness**: The AI interaction event may not always display prompt and response text. In some cases, prompts and responses span consecutive entries. When a user does not have a mailbox hosted in Exchange Online, no prompt or response is displayed. Microsoft Facilitator AI-generated notes may not display prompts or responses. Users should be aware of these gaps when relying on activity explorer for compliance or investigative purposes.

- **Data Security Posture Agent constraints**: The agent supports content searches up to 1 GB of data but does not support metadata-based searches. If no time period is specified, results default to the last 7 days. The agent is designed for pre-investigation checks and is not a replacement for formal investigation tools such as eDiscovery, audit, or Data Security Investigations.

- **Language support**: The Security Copilot experience and the Data Security Posture Agent are designed and evaluated primarily for English-language prompts. Using other languages may result in reduced accuracy or incomplete responses. Users should exercise caution when operating outside the intended language scope.

- **Generative AI response accuracy**: As with all generative AI systems, Security Copilot responses in Data Security Posture Management may occasionally contain inaccurate or incomplete information. Users should verify AI-generated insights against actual data before taking consequential actions based on those responses.

Evaluations

Performance and safety evaluations assess whether AI applications are operating reliably and securely by examining factors like groundedness, relevance, and coherence while identifying the risks of generating harmful content. The following evaluations were conducted with safety components already in place, which are also described in Safety Components and Mitigations.

Data Security Posture Management was evaluated using custom evaluation methods developed by the product team. The evaluation focused on the accuracy of the embedded Security Copilot experience when responding to natural language questions about data security.

The evaluation data set consisted of relevant prompts identified by the product team and customers, including an expanded test set of prompts based on table schemas and descriptions. Microsoft evaluated performance using an accuracy rate metric: a response is considered accurate if the generated query provides the exact information asked in the prompt. The team tested across a range of data security scenarios covering DLP, information protection, and Insider Risk Management insights to verify that Security Copilot responses are factually supported by the underlying data and contextually appropriate to the user's question.

An ideal result is one where the Security Copilot response returns exactly the information requested, grounded in the organization's actual data, with no fabricated or misleading content. A suboptimal result would be a response that returns inaccurate data, misinterprets the user's question, or provides information outside the scope of the supported data security solutions. The evaluation process is iterative, with the product team refining the prompt sets and accuracy thresholds based on customer feedback and real-world usage patterns.

Safety components and mitigations

- **Role-based access control**: Data Security Posture Management enforces granular permissions through Microsoft Entra and Microsoft Purview role groups. Different activities require specific roles, such as Compliance Administrator for editing, Data Security Viewer for using Security Copilot, and Content Explorer Content Viewer for viewing AI interaction prompts and responses. This ensures that users can only access data and take actions appropriate to their role, reducing the risk of unauthorized access to sensitive information.

- **Comprehensive audit logging**: All automated actions taken by AI agents in Data Security Posture Management are recorded in audit logs. This includes actions such as removing sharing links, applying policies, and revoking permissions. Audit logs and activity explorer features provide a full trail of interactions with AI apps and agents, supporting compliance investigations, incident response, and accountability.

- **Human review and approval for agent actions**: Users always maintain control over AI agent behavior within Data Security Posture Management. AI-driven triage agents and the Data Security Posture Agent surface recommendations and findings, but users review, approve, or customize all automated actions before they take effect. The "View agent activity" options throughout the interface provide easy access to agent activity for oversight.

- **Opt-in analytics processing**: Data Security Posture Management requires explicit opt-in before processing organizational data. Analytics in both Insider Risk Management and DLP must be enabled before Data Security Posture Management can generate insights. This ensures that organizations make a deliberate decision to share data for analysis and are aware of the processing that takes place.

- **Scoped AI capabilities**: The Security Copilot experience in Data Security Posture Management is deliberately scoped to data security topics associated with Insider Risk Management, Information Protection, and Data Loss Prevention. Questions outside this scope are not answered, reducing the risk of the AI generating responses about topics where it lacks reliable data. The Data Security Posture Agent is similarly scoped to content searches within Microsoft 365 and limited to 1 GB of content per query.

- **Risk-level classification for search results**: The Data Security Posture Agent assigns a risk level to each search result based on how closely the content matches the user's prompt. Results are categorized to help users prioritize and focus on the most relevant findings. When the agent cannot determine the risk level, items are marked as "not categorized," providing transparency about the agent's confidence level.

- **Alert triage and noise reduction**: AI-driven triage agents review DLP and Insider Risk Management alerts and categorize them as "needs attention," "less urgent," or "not categorized." This filtering reduces alert fatigue and helps security analysts focus on genuine threats rather than false positives.

- **Privacy-preserving design**: Insider Risk Management, which feeds data into Data Security Posture Management, includes privacy controls such as pseudonymization and role-based access to ensure user-level privacy while enabling risk analysis. The Security Copilot

component follows the data privacy and security practices documented in the Microsoft Security Copilot privacy and data security documentation.

- **In-product feedback mechanism**: Users can provide feedback on each Security Copilot response by marking it as "looks right," "needs improvement," or "inappropriate." This feedback loop helps Microsoft continuously improve the quality and safety of AI-generated responses.

Best practices for deploying and adopting Data Security Posture Management

Responsible AI is a shared commitment between Microsoft and its customers. While Microsoft builds AI applications with safety, fairness, and transparency at the core, customers play a critical role in deploying and using these technologies responsibly within their own contexts. To support this partnership, we offer the following best practices for deployers and end users to help customers implement responsible AI effectively.

Deployers and end-users should:

- **Exercise caution and evaluate outcomes when using Data Security Posture Management for consequential decisions or in sensitive domains**: Consequential decisions are those that may have a legal or significant impact on a person's access to education, employment, financial platforms, government benefits, healthcare, housing, insurance, legal platforms, or that could result in physical, psychological, or financial harm. Sensitive domains—such as financial platforms, healthcare, and housing—require particular care due to the potential for disproportionate impact on different groups of people. When using AI for decisions in these areas, make sure that impacted stakeholders can understand how decisions are made, appeal decisions, and update any relevant input data.

- **Evaluate legal and regulatory considerations**: Customers need to evaluate potential specific legal and regulatory obligations when using any AI platforms and solutions, which may not be appropriate for use in every industry or scenario. Additionally, AI platforms or solutions are not designed for and may not be used in ways prohibited in applicable terms of service and relevant codes of conduct.

End-users should:

- **Exercise human oversight when appropriate**: Human oversight is an important safeguard when interacting with AI applications. While we continuously improve our AI applications, AI might still make mistakes. The outputs generated may be inaccurate, incomplete, biased,

misaligned, or irrelevant to your intended goals. This could happen due to various reasons, such as ambiguity in the inputs or limitations of the underlying models. As such, users should review the responses generated by Data Security Posture Management and verify that they match their expectations and requirements.

- **Be aware of the risk of overreliance**: Overreliance on AI happens when users accept incorrect or incomplete AI outputs, mainly because mistakes in AI outputs may be hard to detect. For the end-user, overreliance could result in decreased productivity, loss of trust, application abandonment, financial loss, psychological harm, physical harm, among others. In the context of Data Security Posture Management, overreliance could mean acting on an inaccurate Security Copilot insight without verifying it against actual data, potentially leading to misguided policy changes or missed security risks.

- **Scope Security Copilot questions to supported areas**: For the best results, ask questions about data security topics associated with Insider Risk Management, Information Protection, and Data Loss Prevention. Questions outside these areas may return incomplete or inaccurate responses because they fall outside the data Data Security Posture Management is designed to analyze.

- **Use specific and targeted prompts**: When using the Security Copilot experience or the Data Security Posture Agent, provide clear and specific prompts that include relevant details such as user names, time periods, and data locations. For example, instead of asking "show me security risks," ask "show all sensitive data activities performed by user@contoso.com in the last 30 days." More specific prompts produce more accurate and actionable results.

- **Provide feedback to improve quality**: Use the in-product feedback option to mark Security Copilot responses as "looks right," "needs improvement," or "inappropriate." This feedback helps Microsoft identify and address quality issues in AI-generated responses.

Deployers should:

- **Configure appropriate permissions before granting access**: Assign the minimum required roles to each user based on their responsibilities. Use the detailed permissions table in the Data Security Posture Management permissions documentation to determine which role groups are needed for each activity. Restricting access reduces the risk of unauthorized users viewing sensitive data or taking unintended actions.

- **Complete all setup tasks before relying on insights**: Ensure that auditing, analytics, and collection policies are properly configured before evaluating Data Security Posture Management insights. Incomplete setup can result in missing data and inaccurate posture

metrics. Allow sufficient processing time (up to three days for initial scans, at least 24 hours for new policies) before making decisions based on Data Security Posture Management data.

- **Use restricted searches for efficient processing**: When using the Data Security Posture Agent, configure searches to target specific users, groups, or sites rather than running tenant-wide scans. Restricted searches are significantly faster and more efficient, reducing processing time and producing more focused results.

- **Review and refine one-click policies after deployment**: One-click policies provide a useful starting point, but they should be reviewed and customized based on your organization's specific requirements. After initial deployment, monitor policy results through Data Security Posture Management reports and activity explorer, and adjust policy settings in the corresponding solution (DLP, Insider Risk Management, or Information Protection) to match your organization's risk tolerance and compliance needs.

- **Establish a regular review cadence**: Use Data Security Posture Management posture trends and reports to track your organization's data security posture over time. Establish a regular cadence for reviewing data risk assessments, policy coverage, and AI observability metrics. Regular reviews help identify emerging risks, policy gaps, and changes in user behavior before they become security incidents.

- **Plan for collection policy configuration**: If your organization needs to capture prompts and responses from Copilot in Fabric and Security Copilot, and for non-Copilot AI apps, configure collection policies with the content capture option selected. For apps that use these collection policies, AI interaction events in activity explorer will not display prompt and response text, limiting the usefulness of the data for compliance monitoring.

Learn more about Data Security Posture Management

For additional guidance or to learn more about the responsible use of Data Security Posture Management, we recommend reviewing the following documentation:

- Learn about Data Security Posture Management
- Data Security Posture Management (classic) - Overview of the previous version
- Setup tasks for Data Security Posture Management
- Considerations for Data Security Posture Management
- Permissions for Data Security Posture Management
- Prevent oversharing with data risk assessments

- Microsoft Purview data security and compliance protections for generative AI apps
- Data, privacy, and security for Microsoft Security Copilot
- Microsoft Purview service description

Learn more about responsible AI

- Microsoft AI principles
- Microsoft responsible AI resources
- Microsoft Azure Learning courses on responsible AI

ⓘ **Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 05/01/2026

Data governance with Microsoft Purview

Article • 02/20/2025

Data governance is how you ensure that the data you use in your business operations, reports, and analysis is discoverable, accurate, trusted, and protected. With ever-evolving regulations combined with the proliferation of AI, and regulations for AI, data governance is as critical as data security.

Managing the quality of your data in the era of AI is essential. When your AI systems ingest trusted, high quality data, you have accurate insights for making sound business decisions. But managing data quality is time-consuming. Data professionals often spend most their time simply finding, cleaning, and organizing data, instead of using the data that's clean and trusted.

Microsoft Purview data governance solutions, Unified Catalog and Data Map, deliver a modern data governance experience to provide comprehensive visibility, data confidence, and responsible innovation by streamlining metadata from disparate catalogs and sources. With Unified Catalog at the center of your day-to-day operations, your organization can experience comprehensive visibility, data confidence, and responsible innovation to help you achieve greater business value in the era of AI.

Data governance solutions

Microsoft Purview data governance involves two primary solutions:

- Microsoft Purview Data Map
- Microsoft Purview Unified Catalog

Data Map is what you use to scan all your assets and multicloud sources to capture metadata. Unified Catalog is a searchable catalog of your scanned data where you curate, grant access to, and improve the health of your data. The catalog is a software as a service (SaaS) experience based on a single-tenant model, providing end-to-end capabilities for curating data, managing data quality and health, and granting users in your organization access to data. With a user-friendly and modern experience, Unified Catalog helps organizations achieve greater business value in the era of AI through high quality data.

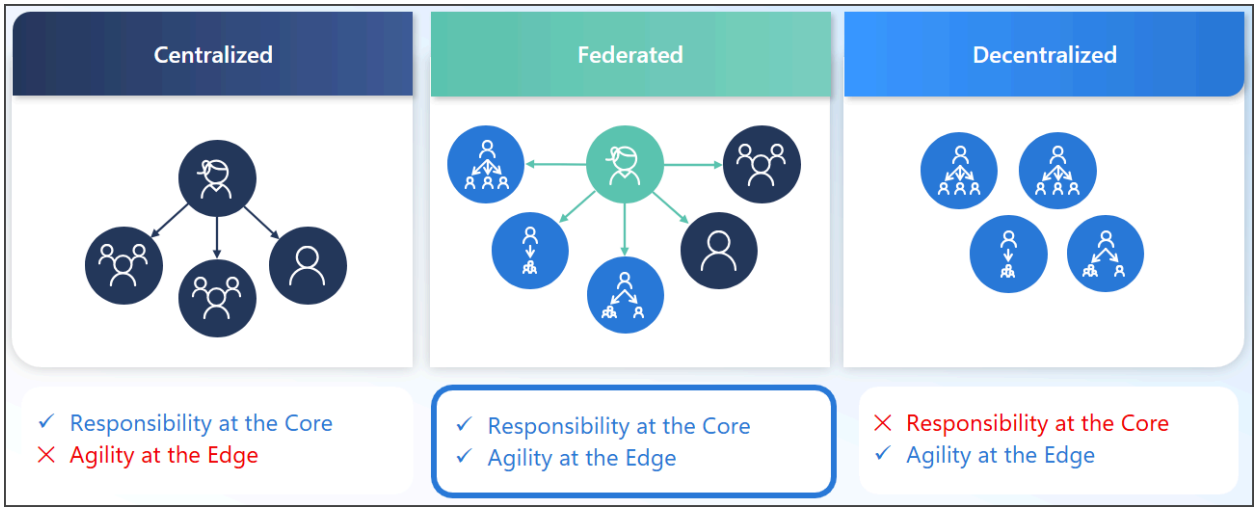
To get an overview, check out this video explaining Microsoft Purview data governance capabilities, how to set up for success, and the day-to-day management

experience.

> **Important**
>
> As you work with Data Map and Unified Catalog, keep these points in mind:
>
> - All data in Data Map and Unified Catalog is **metadata**, not the underlying data itself.
> - None of the permissions or roles in Data Map or Unified Catalog provide access to underlying data itself.

Federated approach to data governance

A federated approach to data governance is a middle ground between a centralized and a decentralized approach. An organization's central data office sets the rules, while individuals in various roles and departments who use the data and understand its function are entrusted to govern that data appropriately. The federated model ensures that there's centralized oversight and helps unlock business innovation by facilitating how users request access to the data they need.



User-friendly tools for data professionals

An effective system of data governance empowers your organization's data professionals to easily manage, discover, and access data. Microsoft Purview data governance solutions are built for each role performed in your organization. Business domain experts can accelerate data value while your data office and stewards provide the standards for consistent operations and can maintain commonalities.

- **Central data office**: Establishes and ensures governance policies, active metadata, compliance, and insights into overall governance health.
- **Data consumers**: Quickly find and use relevant, trusted datasets through streamlined access request workflow.
- **Data owners**: Register data assets for use, manage classifications and access, and ensure high quality standards.
- **Data stewards**: Ensure data quality, seamless data discovery, glossary, consistency, and lineage. Work with central data office.

Key capabilities

Comprehensive visibility

Simplify visibility across disparate catalogs, data sources, data management, and data security with these capabilities:

- Unified Catalog, providing AI-enabled recommendations for data quality and data curation.
- Data products within the Unified Catalog, which are groups of assets with a use case that can be shard with data consumers.
- Actionable oversight that helps you analyze and track data health while ensuring users have access to the data they need.

Data confidence

Maintain confidence in your organization's data with built-in data quality and data lineage capabilities aligned to industry standards and policies:

- Governance domains of data products and business concepts for easier access and use. Use governance controls across your data estate to complement the domains set up in other technical catalogs. Establish accountability for the area of your data estate.
- Data lineage helps you easily see the relationship between the data products and any assets associated with them to identify the root cause of quality issues.
- Data quality experiences built-in to empower data owners to oversee the quality of their data ecosystem and facilitate actions for improvement.

Responsible innovation

Responsibly unlock business value with easier data discovery and actionable insights with these capabilities:

- Role-based access controls to ensure the right people have access to the right data through reader and owner level rights.
- Simplified access and discovery by searching on business context and natural language, with help from Security Copilot.
- Quick access to analytics with Purview Analytics in OneLake, so you can export data from Unified Catalog and import it into OneLake for business users running analytics in Microsoft Fabric.

> **Tip**
>
> Visit **Learn about Unified Catalog** for detailed information about features and how to get started using the catalog.

Workflow at a glance

There are different approaches your organization can take to start building a data governance practice. Listed below are the overarching steps involved in setting up data governance with Microsoft Purview:

1. Assign a user in your organization the Data Governance Administrator role, which grants permissions to assign roles to other users in your organization for key governance tasks and data curation.
2. Using Data Map, scan your data assets and multicloud sources in Data Map to capture metadata.
3. In Unified Catalog, build governance domains and curate data products.
4. Connect your data with business concepts, such as objectives and key results (OKRs) and glossary terms.
5. Improve data quality and remove data issues to improve data health.

> **Tip**
>
> We've mapped out suggested paths to help you get started depending on your situation at **Plan for data governance**.

Billing

Before you get started, review the pricing model for Microsoft Purview data governance solutions.

Getting started

If you're ready to begin, use these resources to get started:

- Learn how to plan for data governance.
- Visit Get started to view detailed instructions with links to deeper technical content.
- Review the roles and permissions needed for users in your organization who play a role in data governance.

Feedback

Was this page helpful? Yes No

Plan for data governance with Microsoft Purview

Article • 02/20/2025

Before you get started with data governance, there are important planning activities and considerations that should be reviewed by your information technology and data management teams. Thoroughly understanding and planning for deployment will help ensure that your implementation and use of features in Microsoft Purview Unified Catalog and Microsoft Purview Data Map go smoothly and align with the best practices.

> **Note**
>
> Check to see which regions are supported for Microsoft Purview data governance solutions.

Relationship between Unified Catalog and Data Map

The Data Map is the technical layer of Microsoft Purview that stores the data about the assets. Data sources are scanned to create the inventory of what assets are available in the different data sources across your organization's data estate.

The Unified Catalog is the business-friendly layer where data stewards curate data, and data consumers find the right data to use. The Unified Catalog is powered by the asset inventory in the Data Map to create a marketplace of data products, ensuring consumers can find the right data and use it responsibly.

> **Important**
>
> As you work with Data Map and Unified Catalog, keep these points in mind:
>
> - All data in Data Map and Unified Catalog is **metadata**, not the underlying data itself.
> - None of the permissions or roles in Data Map or Unified Catalog provide access to underlying data itself.

Who to involve in data governance

When you start your data governance planning with data owners, data stewards, and other business users with expertise, you'll ensure that the Unified Catalog has all of the information needed to build new value with data. IT teams are critical organizational partners in their data use, and help to ensure the right data sources are prioritized for onboarding to the Data Map.

Your organization's data office will oversee the creation of the catalog with company-wide standards, and help build bridges between teams to provide the best governance framework. Having both the data publishing teams and the data consumers included in the process of building out the catalog will provide the most value from the data, and build the collaboration to include all of the experts across your organization.

> **Tip**
>
> For the best success, **it's critical to involve all parties early**. This is because the practice of data governance isn't done by a single person, and no one knows everything about the data, the business, and what it means to use that data responsibly.

Where to start

There's more than one way to start building your data governance practice with Microsoft Purview. How you start is up to you. The sequence in which you perform your initial setup tasks might depend on your business strategy and who's doing the work. Regardless of the order, the full value of Microsoft Purview for data governance isn't realized until all key tasks are completed. The tabs below provide guidance based on your situation or preferred approach.

New to Microsoft Purview or data governance

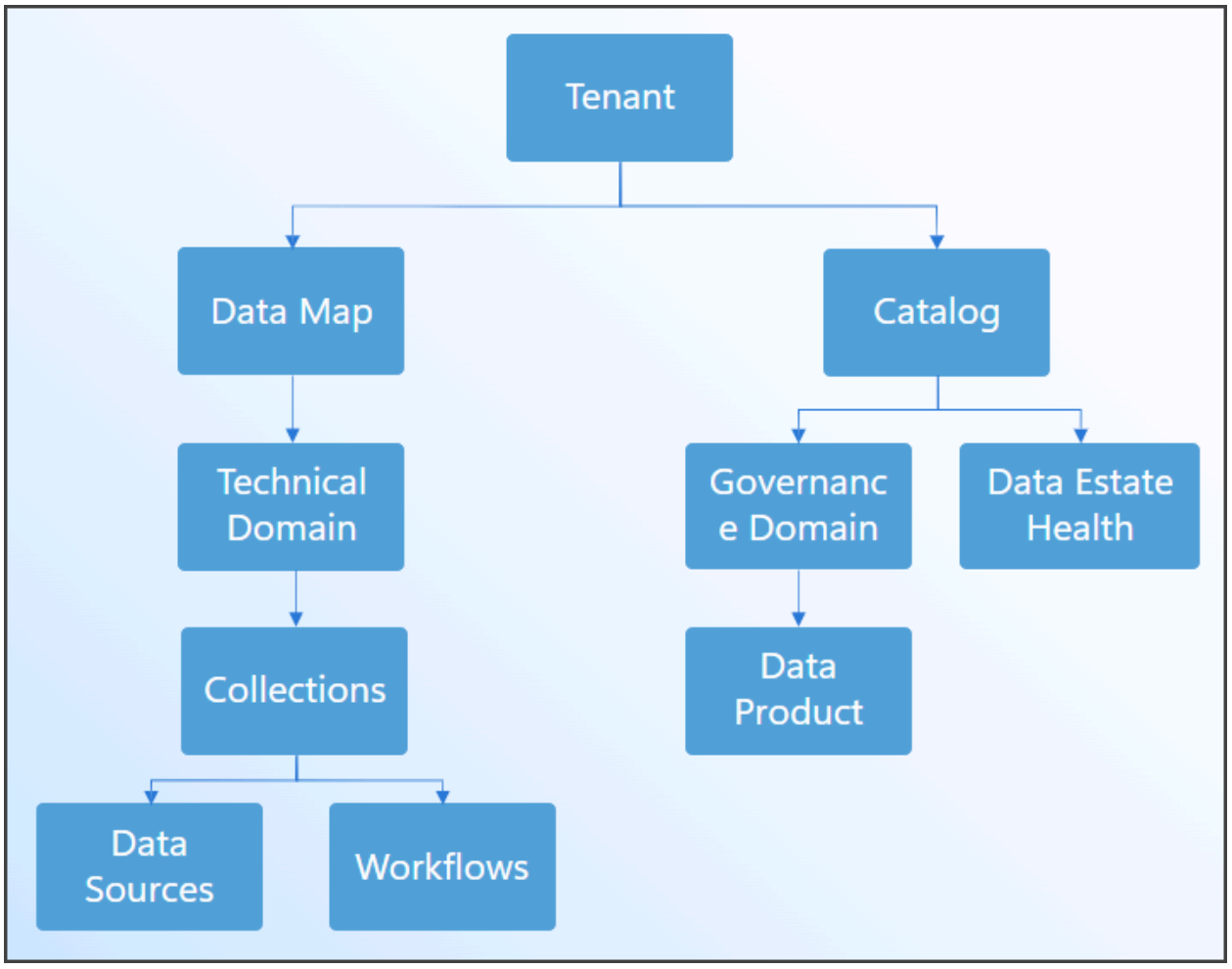
If you're new to data governance or to Microsoft Purview, we suggest starting with the stakeholders who are most passionate about data governance, or have the need to create a data catalog. Typically, business intelligence teams or data lake teams can be a good place to start because they have experts on the data and have invested in building data that others can use.

When you're ready to get started, we recommend doing the tasks outlined in Getting started with data governance in the following order:

1. Create a governance domain (Unified Catalog).
2. Add a domain owner who will be accountable for that domain (Unified Catalog).
3. Add the experts for that part of business to build business concepts (Unified Catalog).
4. Create a collection (Data Map).
5. Add a data source administrator from the data platform to onboard first (Microsoft Purview settings).
6. Register and scan the data source (Data Map).
7. Build out the data products for the most important and frequently used data in that source in the governance domain (Unified Catalog).

Permissions hierarchy

Before you start assigning roles and permissions to users in your organization, it's helpful to understand the relationship between the primary data governance solutions, Unified Catalog and Data Map, and the permissions granted by the roles for each solution.



- **Tenant, or organization**: Users who own the Microsoft Purview instance and need access to maintain its permissions and information.
- **Data Map**: Grants permissions to create and administrate technical domains and collections.
- **Unified Catalog**: Grants users permission to create governance domains or access data estate health.
- **Governance domain**: Grants permissions for objects within a governance domain, such as glossary terms or data products.

When you're ready to start assigning roles and permissions, Data governance roles and permissions gives detailed instructions. To guide you through the process, we recommend beginng with Getting started with data governance.

Next steps

- Ready to grant access and assign roles? Visit Roles and permissions.
- If you're ready to start setting up things, visit Get started with data governance.
- Want to follow an example? Review a walkthrough of a sample data governance setup.

Feedback

Was this page helpful? Yes No

Get started with Microsoft Purview data governance

This article takes you through technical steps to get started building your data governance solution in Microsoft Purview and integrating data governance with your day-to-day business operations.

You might find it helpful to review Plan for Unified Catalog as an orientation to the material below. For a step-by-step illustration of how to set up your environment, visit the sample setup walkthrough.

Prerequisites

1. You need a Microsoft Purview Enterprise instance, either by:
   a. Upgrading an existing account to the new experience
   b. Upgrading from the free version of Microsoft Purview to enterprise
2. Get your data sources registered and scanned in Microsoft Purview. View the list of supported data sources.

Data Governance Administrator role

The Data Governance Administrator role delegates the first level of access for Microsoft Purview Unified Catalog users. Granting this role to a user in your organization is your first step as you get started. Find details on how to assign roles for data governance.

Business concepts

Business concepts are elements in Unified Catalog that help you organize and understand your data in ways that make sense to the data consumers and curators in your organization. When we refer to business concepts in Unified Catalog, we're talking about these five elements:

1. Governance domains
2. Data products
3. Critical data elements (preview)
4. Glossary terms
5. Objectives and key results (OKRs)

The Plan for Unified Catalog article has detailed guidance to help you understand these concepts and how to start working with them. As a starting point, it's helpful to have a good understanding of **governance domains** and **data products**:

- **Governance domain**: A boundary that enables the common governance, ownership, and discovery of data products and other business concepts. The goal is to empower a governance domain owner to manage their data products, and establish rules for their access, usage, and distribution. Here are some examples of how you could set up your governance domains:
  - Corporate or business areas, such as human resources, sales, and finance.
  - Overarching subject areas, such as product or parties.
  - Regulatory domains, such as SOX, PCI, or anti-money laundering.
  - Boundaries based on organizational functions, such as customer experience, cloud supply chain, or business intelligence.

- **Data product**: A group of data assets, such as tables, files, or Microsoft Power BI reports, that are packaged together with a proper use case to be shared to data consumers in the organization. A governance domain can house many data products, but a data product is managed by a single governance domain and can be discovered across many governance domains.

Once you set up governance domains and data products, you can further curate your data by:

- Connecting to the other business concepts.
- Creating and applying attributes to your business concepts.

Basic setup steps

1. Assign users to the Data Governance Administrator role - Gives the admin permissions to grant users application-level-permissions like governance domain creation.
2. Build governance domains:
   a. Assign users the governance domain creator role
   b. Create at least one governance domain

      > **Tip**
      >
      > Use **our guide** to strategize your governance domain structure.

   c. Assign at least one governance domain owner on each governance domain. This user is the point of support and decision authority for data being consumed in this governance

domain.

3. Register your data sources in Microsoft Purview Data Map, and then scan the data sources to capture metadata.
4. Build data products:
   a. Assign data product owners to create data products in your governance domains. These should be business and data experts who can pair data with day-to-day scenarios in their governance domain.
   b. Create at least one data product in your governance domains.

   > **Tip**
   >
   > Use **our guide** for tips on how to build good data products.

5. Assign users catalog reader permissions in your governance domains so they can view and explore your data products in the Unified Catalog.

Connect your data with business concepts

Glossary terms provide vocabulary for business users. These terms let users discover and work with data in the vocabulary that's familiar to them, instead of using abstract technical terminology inherited from physical data sources.

Objectives and key results (OKRs) are the goals or desired outcomes of a governance domain; for example, 10% rise in sales or 3% reduction in support cases. Objectives should relate to everything an organization does and should define how they're achieving their outcomes.

Health management actions give you and your users steps to improve data health and governance across your data estate. These actions correspond to the checks made to calculate a data product's data governance health control score. Addressing these actions raises your health score and promotes an overall more usable and discoverable Unified Catalog. Understanding the value of your data products improves the trust others take in that data and helps with prioritizing which data to focus on improving first.

Connect action steps

1. Assign users data steward permission in your governance domains to create and manage glossary terms. These users should be data and business experts. They increase the value of data products by curating the information and attaching glossary terms to make the data understandable.

2. Create at least one OKR in your governance domain.
    a. Link your OKR to a data product to connect your data to your business goals.
3. Assign users data health reader permissions in Unified Catalog. These users monitor Unified Catalog use and current governance scores, and take actions to build trust in your data products.
4. Publish some glossary terms in your governance domain.
    a. Link a glossary term to a data product to improve discoverability.
5. Assign users data steward permission in your governance domains to create and manage OKRs. These users should be business strategy experts to ensure business leaders appreciate the value of their data, and the importance of data governance. They drive prioritization and strategize how teams build, maintain, and govern their data to create insights.
6. Review your health controls to get a baseline for your current health management.
7. Review health actions to start considering next steps for your data governance journey.
8. Create custom metadata and apply attributes to your business concepts.

Improve data quality and remove data issues

Data quality measures the quality of data in an organization. It uses data quality rules that you configure and define in Unified Catalog.

Data quality rules describe the state of the data with dimensions like accuracy, completeness, conformity, consistency, timeliness, and uniqueness. Each rule produces a score that describes how close the data is to its desired state.

Data profiling examines the data available in your data sources. It collects statistics and information, and assesses the quality level of the data according to a defined set of goals. If data is of poor quality or managed in structures that can't be integrated to meet the needs of the organization, it can affect business processes and decision-making.

Quality action steps

1. Assign users(s) data quality steward permissions to use all data quality features.
2. Set up a data source connection to prepare your source for data quality assessment.
3. Configure and run data profiling for an asset in your data source.
    a. When profiling is complete, browse the results for each column in the data asset to understand your data's current structure and state.
4. Set up data quality rules based on the profiling results, and apply them to your data asset.

5. Configure and run a data quality scan on a data product to assess the quality of all supported assets in the data product.
6. Review your scan results to evaluate your data product's current data quality.

Weekly planning

The following is a reference example to assist with planning the new Microsoft Purview data governance solution areas, scenarios, tasks, and personas with key stakeholders.

Week 1-2

Expand table

Area	Scenario	Task	Description/Outcomes	Persona
Data management	Catalog setup	Set up first governance domain	Identify governance domain scope, usage, and owners. Assign accountability to governance domain owner, define and create your first governance domain, description, and assign data owners. Capture feedback on establishing the governance domain.	Governance domain owner
	Catalog curation	Create data products in the governance domain	Identify scope of data to manage, publish, and owners. Create data products, descriptions, use cases, assign ownership, and create and assign glossary terms to help increase usability for data consumers. Map data assets to data products, create access policies for data consumers to attest to when requesting access – capture feedback (ease of use for business unit to manage, understand, and own curation)	Data product owner/data steward
	Publication	Publish the governance domain and data products	Publish the governance domain and associated data products to make available for discovery, understanding, and access through the Unified Catalog experience. Assign data consumers permissions to access and view the first governance domain by adding them to the Global Catalog reader role and capture feedback with publication.	Governance domain owner and data product owner
	Operations	Data governance and	Assess operational tasks, stakeholders, processes, and procedures to enable data governance and management. Evaluate	Data governance office

Area	Scenario	Task	Description/Outcomes	Persona
		management operations	against current state data governance policies, practice, and culture to identify potential areas for improvement or change.	

Week 2-3

Expand table

Area	Scenario	Task	Description/Outcomes	Persona
Data discovery, understanding, and access	Discover and access	Unified Catalog product search	Exercise the Unified Catalog product search experience to help data consumers and users discover and understand data products that are curated and developed for a specific business purpose. Assess data product metadata to determine proper usage, data quality, and applicability to data consumer business outcomes, and then request access. Assess ease of use to data products the user recently reviewed, and data products subscribed to – capture feedback on full data consumer experience.	Data consumer
Data management	Access management	Access request management	Review access requests for data products in the first governance domain and approve or reject. Engage with IT owners for approvals (as appropriate) to data assets.	Data product owner
	Catalog curation	Review data product discoverability	Review discoverability and usability of published data products along with data consumer feedback to inform semantic knowledge improvement opportunities (for example, glossary terms, attention to attention items in the action center, etc.).	Data product owner/data steward

Week 3-4

Expand table

Area	Scenario	Task	Description/Outcomes	Persona
Data management	Data quality	Improve data quality and reduce data issues	Assess top-level data quality by the first governance domain, and evaluate and set up data quality for associated data assets by data product (via connections). Use data quality profiling data to inform quality rules and dimensions to establish key data assets in the data product. Run data quality scans (ad-hoc or scheduled), monitor data quality activity and scans, and set up alerts to be informed of changes with data asset health (via target thresholds). Capture feedback on overall data quality experience.	Data quality steward
	Operations	Data governance and management operations	Assess operational tasks, stakeholders, processes, and procedures to enable data quality in the context of data governance and management. Evaluate against current state data governance policies, practice, and culture to identify potential areas for improvement or change.	Data governance office

Week 4-5

Expand table

Area	Scenario	Task	Description/Outcomes	Persona
Health management	Reports	Manage data governance	Review the controls with business data domain owners and set up regular review of reporting on those controls. The goal of the meeting is to review issues and prioritize solutions or data products that are needed to meet business needs.	Data governance office
	Health actions	Improve data governance	Take actions based on the controls to improve data governance and ensure standards are met.	Data stewards/data product owners

Unified Catalog overview page

The **Overview page** in Unified Catalog helps users in an organization get started with their data governance journey, understand, and navigate the different steps outlined in this document, using step by step instructions and video demos.

Resources

- Sample setup walkthrough for data governance using Unified Catalog
- Data governance roles and permissions

Last updated on 04/24/2026

Sample setup for data governance

10/13/2025

Microsoft Purview data governance, featuring Microsoft Purview Unified Catalog and Microsoft Purview Data Map, delivers comprehensive visibility, data confidence, and responsible innovation to help organizations achieve greater business value in the era of AI. Using an example of managing health data, follow the steps in this article to help you understand how to set up Unified Catalog and use its functionality to build a sound data governance practice for your organization.

Step 1: Set up your governance domains in Unified Catalog

Governance domains are the key to establishing accountability for your data and help to federate governance of that data across the company. When you create governance domains, start with the proper owner to ensure you can effectively identify and collaborate with experts for all of the data in the data estate. Governance domains can be many different types to align to the type of data boundary for the team that governs that data. For example: functional domains (finance, HR, sales), or data domains (product, customer, health).

Prerequisites

- You need the Data governance administrator role for the Unified Catalog.

Grant permissions and build the first governance domain

1. Open the Microsoft Purview portal.
2. Sign in to the Microsoft Purview portal using credentials for an admin account that has the *Role management* role (for example, a Purview administrator). Go to **Settings** > **Roles and scopes** to view and manage.
3. Select **Role groups**.
4. On the **Role groups for Microsoft Purview solutions** page, select the **Data Governance** role group.
5. On the **Edit member of the role group** page, select **Choose users** or **Choose groups**.
6. Select the check box for all users or groups you want to add to the role group.

7. Select **Select**.

8. In Unified Catalog, select **Catalog management**, then select **Governance domains**.

9. On the **Governance domains** page, you can set up the rest of your catalog to enable others to federate the ownership of data, empower teams to build out their knowledge, and establish business value of your data.
    a. Start by selecting **New governance domain**.
        i. You can update the name of your governance domain. For this tutorial, name it '(Tutorial) Personal Health' and give it a description of 'Personal health data refers to any information related to an individual's physical or mental health that is collected and used within the healthcare sector. This data can include a wide range of types, such as medical records, treatment histories, diagnostic images, and laboratory test results. It's often protected under various laws and regulations to ensure privacy and confidentiality.'
        ii. Select the type as a 'data domain'.
        iii. Leave the parent blank (if this is the first governance domain in the catalog, it doesn't have a parent).
        iv. Select **Create**.
        v. Now create two more domains on your own. These domains are key points of federation for collaboration and governance in your organization. Think about who might be the owners of your domains when you implement Unified Catalog.
            i. You can follow these examples:

                - A **Corporate** functional domain represents the highly controlled assets and terms that an entire company uses. - **Sales** is a functional domain that most organizations have as a child domain of Corporate.

10. Select the governance domain you created.

11. Select the **Roles** tab of the governance domain.

12. By default, when you create the governance domain, you're added to all roles in the governance domain. As a governance domain owner, you add the data stewards (business experts in your domain) and the data product owners (who know which data assets are best for others to consume).

13. Switch back to the **Details** tab.

14. Select **Manage policies** to apply a domain level policy. This policy applies to all data products in the domain. By enabling the automatic application of a policy, data experts don't have to be policy experts as well.

15. In the **Manage access policies** tab, select the checkbox next to **Permit data copies.** By selecting this policy option, it automatically applies an attestation requiring all users who request access to your data products to attest that they understand the data copy policy for your data.

16. Select **Save changes** to confirm the policy is set by the governance domain.

17. Select **Publish** on the governance domain, which publishes all other concepts within the domain.

Create glossary terms

Adding glossary terms to your governance domain helps others understand how the business uses and interprets the data. Glossary terms also ensure insights use common terms, and generally your knowledge across your governance domain.

1. On the page for your governance domain, find the **Glossary terms** card and select **View all**.

2. On the **Glossary terms** page, select **New term**.

3. Enter details: 1. Name: "Outbreak."
   a. Description: A disease that affects or has the potential to affect a large portion of the population.
   b. You can leave the rest blank for now but there are fields to collect: the term owner responsible for defining the term for your company, acronyms to share common also known as names of the term, lastly you can provide links to the resources that would have even more information about the term.

4. Select **Create**.

5. Select **Manage Policies**. Similar to the domain level policies, you can create term level policies that apply wherever the term is in use.

6. Check the box next to **Manager approval required**. This policy enforces a secondary approval from the user's manager in Microsoft Entra ID when access is requested to the data products.

7. Select **Publish** for the Outbreak term created. Published terms are filterable in Unified Catalog and ensure others that use the term to describe their data product can see that description in Unified Catalog while browsing the data product.

8. Now create two more terms. This time, select the "Outbreak" term as the parent term for the terms you created. Try building relationships between these child terms in the related

tab on either term to help build out the network of how these terms work together to explain the entirety of a topic.
   a. Pandemic: A global outbreak of a disease that affects a large number of people across multiple countries/regions or continents.
   b. Epidemic: A countrywide or regional outbreak of a disease that is highly contagious and affects a large portion of the population.

9. Try creating a couple of other terms in any other domains you created earlier. If you're not sure what to add, select **Get suggested terms** to have generative AI propose a few based on the description and name of the domain you already provided.

Add an OKR

Add an OKR (objective and key result) for your Personal Health domain to help others understand the business value of your data. This step builds a direct connection between your data and the business value it provides.

1. Select the **OKR** box from the governance domain page.

2. Select **New OKR**.

3. Enter the details of the objective first:
   a. Objective: Reduce pandemic risk by enabling effective patient vaccine uptake.
   b. Owner: Enter your name
   c. Target date: '2024-12-31'

4. Select **Create**.

5. Add key results to your objective to make the goals measurable and to monitor progress towards the goal. Select **Add key result**.

6. Enter the Key result details:
   a. Key result: Ensure 80% older age groups(>65 years) that are most likely to be affected by the pandemic receive full vaccination by end of the calendar year 2024.
   b. Progress status: On track
   c. Progress Amount: 70
   d. Goal amount: 80
   e. Maximum amount: 100

7. Select **Create**.

8. Select **Publish**.

Create critical data elements

Create a critical data element (CDE) in Personal Health to ensure the most important columns of data have a consistent definition and understanding. The CDE always meets business expectations for how that data is formed and stored.

1. From the governance domains page with the Personal Health domain selected, select the **Critical data elements** box.
2. Select **New critical data element**.
3. Enter the basic CDE metadata:
   a. Name: Age groups
   b. Description: Common grouping of person ages used to ensure needed analytical reports follow a reference that others can depend on and removing individual ages to improve anonymity of the data. The age group is divided into eight groups: <2 years, 2-4 years, 5-11 years, 12-17 years, 18-24 years, 25-49 years, 50-64 years, 65+ years.
   c. Owner: enter your name
   d. Expected Data Type: Text
   e. Select **Create**.

The real power of the CDE is that it maps directly to the physical data columns where this data is stored. This connection ensures common understanding and enables the evaluation of Data Quality rules and policies at scale.

1. From the CDE you just created, select **Add column**.
2. Search for the **Covid 19 Vaccine and Case Trends** data asset from the gold container of the data lake
3. Select the box, not the name, of the **Covid 19 Vaccine and Case Trends** asset.

   > **Tip**
   >
   > If you select the blue name of the asset, it opens a new window showing you the asset details.

4. Select the radio button next to the AgeGroupVacc column.
5. Select **Add**.
6. Select the **Data quality** tab at the top of the CDE you just created to apply data quality rules to the CDE. It's similar to how you added policies for glossary terms and governance domains.

7. Select **New rule**
8. Select **Data type match**
9. Enter Rule name: Confirm Age group formatting
10. Select **Create**.
11. Select **Publish** on the CDE

This CDE now automatically applies a data quality rule to every data product that uses the Covid 19 Vaccine and Case Trends asset, which you see in the next section.

1. Try creating a couple of other CDEs in your other domains. Here are some ideas:

    - Sales: Revenue and Seller Name
    - Corporate: Product ID

Step 2: Set up and register your data in Data Map

If you don't have data sources available for scanning, follow these steps to fully deploy an Azure Data Lake Storage (ADLS Gen2) example.

> **Tip**
>
> If you already have a data source in the same tenant as your Microsoft Purview account, move ahead the next part of this section to scan your assets.

In a real data estate, you find many different systems in use for different data applications. There are reporting environments like Fabric and Snowflake where teams use copies of data to build analytical solutions and power their reports and dashboards. There are operational data systems that power the applications teams or customers use to complete business processes that collect or add data based on decisions made during the process.

To create a more realistic data estate, show many sources of data in the catalog, which can cover the breadth of different data uses any company might have. The types of data required to power a use case can be vastly different with business users that need reports and dashboards, analysts need conformed dimensions and facts to build reports, data scientists or data engineers need raw source data that came directly from the system that collects the data. All of these and more enable different users to see the importance of finding, understanding, and accessing data in the same place.

For some other tutorials to add data to your estate, follow these guides:

- Fabric Lakehouse Tutorial – provides the base of a reporting environment
- Azure SQL Database (Sample) – provides a well structured example of an operational data store

Prerequisites

- Subscription in Azure: Create Your Azure Free Account Today
- Microsoft Entra ID for your tenant: Microsoft Entra ID Governance
- A Microsoft Purview account and administrator privileges (this is the default if you created the Microsoft Purview account).
- All resources; Microsoft Purview, your data source, and Microsoft Entra ID have to be in the same cloud tenant.

Set up your data estate

A. Create and populate a storage account

1. Follow this guide to create a storage account: Create a storage account for Azure Data Lake Storage Gen2
2. Create containers for your new data lake:
   a. Go to the **Overview** page of your Storage Account.
   b. Select the **Containers** tab under the Data storage section.
      i. Select **Container**.
      ii. Name the container "bronze" and select **Create**.
      iii. Repeat these steps to create a "gold" container.
3. Download some example CSV data from data.gov: Covid-19 Vaccination And Case Trends by Age Group, United States
4. Upload the CSV to the container named 'bronze' in the storage account you created.
5. Select the container named 'bronze' and select **Upload**.
6. Browse the location where you saved the CSV and select the **Covid-19_Vaccination_Case_Trends** file.
7. Select **Upload**.

B. Create an Azure Data Factory

This step demonstrates how data moves between layers of a medallion data lake and ensures the data is in a standardized format that consumers expect to use. This step is a prerequisite for running data quality.

1. Follow this guide to create an Azure Data Factory: Create an Azure Data Factory

2. Copy the data from the CSV in the 'bronze' container to the 'gold' container as a Delta format table using this Azure Data Factory guide: Transform data using a mapping data flow

3. Open the Azure Data Factory (ADF) experience from the Azure portal by selecting **Launch studio** on the **Overview** tab of the ADF resource created.

4. Select the **Author** tab in ADF studio.

5. Select the + command, then select **Data flow**.

6. Name the dataflow 'CSVtoDeltaC19VaxTrends'.

7. Select **Add Source** in the empty box.

8. Set Source settings to:
   a. Output stream name: 'C19csv'
   b. Description: leave blank
   c. Source type: Inline
   d. Inline dataset type: Delimited Text
   e. Linked Service: Select the data lake where you stored the csv

9. Set Source options to:
   a. File mode: File
   b. File path: /bronze/ Covid-19_Vaccination_Case _Trends
   c. Allow no files found: leave unchecked
   d. Change data capture: leave unchecked
   e. Compression type: None
   f. Encoding: Default(UTF-8)
   g. Column delimiter: Comma (,)
   h. Row delimiter: Default(\r, \n, or\r\n)
   i. Quote character: Double quote (")
   j. Escape character: Backslash ()
   k. First row as header: CHECKED
   l. Leave the rest as defaults

10. Select **Next** by the source created and select **Sink**.

11. Create the sink where the format and location of the data to be stored to move the data from a csv in 'bronze' to a delta table in 'gold'.
    a. Set the Sink values (leave all settings as default unless specified).
    b. Sink type: Inline.
    c. Inline dataset type: Delta.

d. Linked service: the same data lake as used in the source, because you are storing in a different container.

12. Set the Setting values (leave all settings as default unless specified)
    a. Folder path: gold/Covid19 Vaccine and Case Trends.
13. Enter the value because this name is how you want to store the data and it doesn't exist to select.
14. Select **Validate**. This action checks your data flow and provides instructions to fix any errors.
15. Select **Publish all**.
16. Select the + command, then select **Pipeline**.
17. Name your pipeline 'CSV to Delta C19 Vax Trends'.
18. Select the dataflow created in the previous steps CSV to Delta (C19VaxTrends) and drag and drop it on the open pipeline tab.
19. Select **Validate**.
20. Select **Publish**.
21. Select **Debug** (use activity runtime) to run the pipeline.

    > **Tip**
    >
    > If you get errors for spaces or inappropriate characters for delta format, open the downloaded CSV and make corrections. Then re-upload and overwrite the CSV in the bronze zone. Then rerun your pipeline.

22. Navigate to your gold container in the data lake and you should now see the new Delta table created during the pipeline.

Scan your assets

If you haven't scanned data assets into your Data Map, follow these steps to populate your data map.

Scanning sources in your data estate automatically collects the metadata of the data assets (tables, files, folders, reports, and more) in those sources. When you register a data source and create the scan, you establish the technical ownership over the sources and assets that appear

in the catalog. You also control who can access which metadata in Microsoft Purview. When you register and store sources and assets at the domain level, you store them at the highest level of access hierarchy. Typically, it's best to create some collections where you scan the asset metadata and establish the correct access hierarchy for that data.

- Provide reader access for Microsoft Purview Managed Identity (MSI) to your data lake or other data store.

  > **Tip**
  >
  > The MSI is the account name of the Microsoft Purview instance.

If you choose to use Fabric or SQL, use these guides to provide access:

- Microsoft Fabric
- Azure SQL Database

Register your data lake and scan your assets

1. In Data Map, under the domains tab, select the **Role assignments** for the domain (it's the name of Microsoft Purview account):
   a. Add yourself as the data source admin and the data curator to the domain.
      i. Select the person icon next to the role **Data source admin**. 1. Search your name as it is in Microsoft Entra ID (you might need to enter your full name spelled exactly as it is in Microsoft Entra ID).
      ii. Select **OK**.
      iii. Repeat these steps for data curator.

2. Register the data lake:
   a. Select the **Data sources** tab.
   b. Select **Register**.
   c. Select the Azure Data Lake Storage Gen2 storage type.

3. Provide the details to connect:
   a. Subscription (optional)
   b. Data Source Name (this is the name of the ADLS Gen2 source)
   c. Collection where asset metadata should be stored (optional)
   d. Select **Register**

4. Once registration of the data source is complete, you can configure the scan. Registration signifies that Microsoft Purview is connected to the data source and has placed it in the

correct collection for ownership. Scanning reads the metadata from the source and populates the assets in the data map.

5. Select the source you registered in data sources tab

6. Select new scan and provide details:
   a. Use the default integration runtime for this scan
   b. Credential should be Microsoft Purview MSI (system)
   c. Scan level is Auto Detect
   d. Select a collection or use the domain (collection must be the same collection or child collection of where the data source was registered)
   e. Select Continue

   > **Tip**
   >
   > At this point the connection is tested to validate a scan can be done. If you don't grant the Microsoft Purview MSI reader access on the data source, it fails. If you're not the data source owner or don't have user access contributor, the scan fails since it expects you have authorization to create the connection.

7. Now only select the container "gold" where you placed the delta table in the building data section of the tutorial. This selection prevents scanning any other data assets that are in your data store.
   a. Should have only one blue check next to gold, you can leave checks next to everything as it will scan the full source and still create the assets we'll use and more.
   b. Select **Continue**

8. In the select a scan rule set screen you should use the default scan rule set.

9. Select **Continue**

10. In set a scan trigger you'll set the frequency of the scanning so as you continue to add data assets to the gold container of the lake it will continue to populate the data map. Select **Once**.

11. Select **Continue**.

12. Select **Save and Run**. This action creates a scan that reads the metadata from the gold container of your data lake and populates the table we'll use in Unified Catalog in the next sections. If you only select save, it doesn't run the scan, and you don't see the assets. Once the scan is running, you see the scan you created with a **Last run** status of **Queued**. When the scan reads complete your assets are ready for the next section. This process

could take a few minutes or hours depending on how many assets you have in your source.

Step 3: Publish your data products

Creating data products is essential to ensure that your organization can discover the right data. Data products help prevent overgoverning data that has low or no value in your data estate because it has no use or limited value. When your data experts publish data products, you activate your most valuable data and build the right level of governance based on that value. Curating assets that technical teams don't know the business purpose of, or trying to govern everything in your complex and growing data estate, causes extra time and lost productivity chasing down the details of data that might never be used or could just be removed from the estate. Instead, focus on the pieces of data that have value and that people need to discover and build even more value. As teams use more data and gain a better understanding of what is needed, you can create more useful data products to meet those demands. Governance can adapt to ensure it always stays the right size based on the value and sensitivity of the data.

Prerequisites

- Be a data product owner for the governance domain you're using.
- Have data assets in the Data Map. If you don't, see section 2 of this tutorial to add some.
- Publish a governance domain to publish a data product. If you don't have one, see section 1 of this tutorial to create one.

Create and publish a data product

1. Open the Microsoft Purview portal.
2. Select **Unified Catalog**.
3. Select **Catalog management** and then **Governance domains**.
4. From the **Governance domains** page, select the Personal Health domain.
5. Select **Go to data products** under **Business concepts**.
6. Here's where the data experts called data product owners will identify the data assets that are intended to be consumed by others in your organization, and provide the necessary information to make them usable.
7. Select **New data product**.

8. Provide details about the data product: 1. Name: "Covid-19 Vaccination and Case Trending by Age" 1. Description: "This data comes from the CDC as a part of the U.S. Department of Health & Human Services. The data contains trends in vaccinations and cases by age group, at the US national level. Data is stratified by at least one dose and fully vaccinated. Data also represents all vaccine partners including jurisdictional partner clinics, retail pharmacies, long-term care facilities, dialysis centers, Federal Emergency Management Agency and Health Resources and Services Administration partner sites, and federal entity facilities."
    a. Type: Dataset
    b. Select **Next**.
    c. Use cases: `This data is provided for public use and is intended to help understand the trends of vaccination up take and new cases by different age groups. The ages are banded into two groups ranging from <2 years to 65+ years. Similarly the trends are provided in daily numbers that provide seven day average of new cases by age group.`
    d. Mark as **Endorsed** as checked.
    e. Select **Save**.

9. Now you have the base metadata of the data product built out. Next, add some properties and map the asset from Data Map.

10. Select **Add data assets**.

11. You see the assets you scanned into Data Map, including all folders and layers of the data source.

12. Search for the **Covid19 Vaccine and Case Trends** asset you added to the gold container of your data lake and select this resource set.

13. Select **Add**. You can select as many assets as needed for a data product but here only one is needed.

    > **Tip**
    >
    > Select **Get suggestions** to have generative AI help pick from the assets in your data map and select the Covid19 Vaccine and Case Trends from a reduced list of results.

14. You can now see the asset added to your data product.

15. Select **Add term** next to the glossary terms title.

16. Select the Outbreak term created earlier and select **Add**.

17. You should see the critical data element for age group from the asset mapped to the data product now.

18. Select **Add OKR** next to the OKR title.

19. Select the Reduce pandemic risk by enabling effective patient vaccine uptake. It's the objective you created in the first section.

Manage data product access request policies

At the top of the page, the last step before publishing the data product is to select **Manage policies**. Here, you configure the access policies and request access workflow by making selections and providing the names for approval. You can also use the **Inherited policies** tab to see the governance domain policy applied for data copies attestation you applied earlier. It's the same for the **Manager approval** required coming from the Outbreak glossary term.

1. Select the **Manage policies** tab.

2. Under **Access time limit**, provide details for how long the request for access is good before needing to be renewed. Set this value to grant access for up to one year.

3. In the box, enter 1.

4. Select years in the drop-down.

5. Under approval requirements, provide your name in the approvers box. (It requires the name registered in Microsoft Entra ID.)

   > ⓘ **Note**
   >
   > You don't need to check manager approval because that policy is inherited from the outbreak glossary term.

6. Select **Preview request form** to see what the catalog consumers view when requesting access. You see the data copy attestation and manager approval required because they were set by the governance domain and glossary term.

7. Select **Save changes**.

Once you map the data assets and configure the access policies, you're ready to publish your data product to the catalog.

1. Select **Publish** on the data product.

2. Try creating a Profit Report in other domains you created earlier
   a. Profit Report, Type: Dashboards/reports.
   b. Product Master, Type: Maser data and reference data.

> **Note**
>
> You can add many assets to these and see how a data product with many assets looks. Add the data products to the terms from any domain to see how the glossary is used to describe the data using a consistent set of terms.

Step 4: Run data quality

Now that you have a data product available in the catalog, running data quality rules tells everyone that the data is in good shape and ready to use. As you learn more about the data, add new data quality rules to make sure it's fit for all use cases. Ensuring data products are of the highest quality helps build trust in your data and shows others that you're monitoring and improving it. As the value of data increases, you need to more closely monitor and control the quality of that data. Poorly managed data quality issues can cause significant negative effects.

Prerequisites

- Data quality rules can only be run on delta format tables in ADLS Gen2 and Microsoft Fabric.
- The Managed Identity from Microsoft Purview must be enabled to read the data source as it is the only supported credential for data quality today.
- You must have the data quality steward role in the governance domain you're running data quality in.
- You must be the owner or have user access administrator access to the data source you're connecting data quality scanning to ensure proper security authorization to scan the data.
- You must have the data profile steward role to run profiles on your data.

Create and run data quality rules

1. Open the Microsoft Purview portal.
2. Select **Unified Catalog**.
3. Select the **Data quality** tab under **Data management**.
4. Select the Personal Health Domain created in section 1.

5. Select **Manage**, then select **Connections**. When you build this connection, you can run data quality scans on your data source in that governance domain. This step prevents teams from gaining access to knowledge of the data without proper authorization.

6. Select **New** on the connections screen to create a new connection: 1. Enter the display name "Personal Health ADLSg2 DQ".
   a. Select source type of Azure Data Lake Storage Gen2.
      i. Enter details of the data source created in section 2.

      > ⓘ **Note**
      >
      > Credential must be Microsoft Purview MSI (system) for a data quality connection.

      ii. Select **Test connection**.
      iii. Once the connection is tested, select **Submit**.

Once the connection is established, you're ready to run profiles and start building data quality rules. This step ensures that the experts who know the business rules and appropriate rules are running on the most important data products.

1. Go back to the **Data quality** page.
2. Select the Personal Health governance domain.
3. Select the **Covid-19 Vaccination and Case Trending by Age** data product built in section 3.
4. Select the asset that you added to the data product. (It must be in delta format from section 2 or data quality won't run).
5. Apply data quality rules to the columns of the data to measure if it meets your expectation of quality:
   a. Select **Rules** tab on the asset selected.
   b. Select **New rule**.
   c. Select Empty/blank fields rule.
   d. Enter details:

      - Select AgeGroupVacc column from the column drop-down
      - Rule Name: Confirm Vaccination Age Group Exists

   e. Select **Create**.
   f. Select **New rule**.
   g. Select **Data type match**.
   h. Enter details.
   i. Select DateAdministered column.
   j. Select **Create**.

6. Select **Run Data quality scan**.

Profile data

Create a profile for your data to see the high-level statistics of each column and discover any anomalies that could require a new rule.

1. In Unified Catalog, select **Health management**, then select **Data quality**.
2. Select **Profile data**.
3. Check the top box next to **Column name** to profile all columns. The system recommends which columns to profile, and you can select columns that you know are worth profiling to help prevent profiles on highly sensitive data or data you know is sparsely populated.
4. Select **Run profile**.

When the scan completes, you can review the data quality score and profile for your new data product. All users of the catalog can see the data quality score, so everyone knows the status of the data.

Create a schedule for your data quality scans to ensure you're continuously monitoring for data quality issues. Set alerts to make sure you're addressing data quality issues before consumers are affected.

1. Under **Health management**, select **Data quality**.
2. Select the Personal Health domain where you configured the data quality rules.
3. From the **Manage** dropdown list, select **Scheduled scans**.
4. On the **Scheduled scans** page, select **New**.
5. Add **Overview** details
    a. Name: Personal Health DQ Monthly Evaluation
    b. Description: Monthly scan of DQ rules for continuous improvement.
6. Select **Continue**
7. Select the scope of the scan
8. Check the box next to **Covid-19 Vaccination and Case Trending by Age** data product
9. Select **Continue**
10. Schedule the scan to ensure it runs on the last day of every month
    a. Select **Recurring**
    b. Recurrence: Every one Month
    c. Month days: Last
    d. Schedule scan time (UTC): 12:00:00
    e. Start recurrence at (UTC): leave as default
11. Select **Continue**
12. Review details of the scan to see if there are any changes you want to make before saving.

13. Select **Save**. Because you triggered a manual scan earlier, you don't need to trigger another scan now. If you need a new scan, select **Save and run**.

Configure alerts

After scheduling scans for data quality, you can set up alerts to notify stewards about issues or when data quality problems or scan failures need attention. Configure a data quality alert for failed scans and when the score decreases by more than 5%.

1. Return to the Personal Health domain on the **Data quality** page.
2. From the **Manage** dropdown list, select **Alerts**.
3. Select **New**.
4. Enter alert details
   a. Display Name: Personal Health DQ Monthly Scan
   b. Description: To ensure minimum DQ thresholds are meeting consumer expectations.
   c. Target: Score decreases by more than
   d. Threshold: 5
   e. Turn off notifications: leave unchecked
   f. Turn on notification for failed quality scans: leave checked
   g. Recipient: enter your name
5. Select **Continue**.

> **Tip**
>
> When implementing in Unified Catalog, send the alerts to the stewards who can notify consumers of the issue and work with the technical owner of the data to make corrections.

At the end of this section, you have a functioning Unified Catalog with operational data quality to manage the data you're offering to organizational data consumers. Everything is set up to get the most valuable data to the consumer and build trust in the data that they use. As the value of the data grows and new data strategies emerge, the next section shows how you can manage the entire catalog or go deeper into specific data management with master data.

Step 5: Master data management

Master data management (MDM) is the practice of conforming the most important data entities that must be accurate, unique, and consistently applied in all areas of the business because errors and issues in this data can impact the whole business. Through one of our MDM partners, you can integrate your choice of MDM solutions with Microsoft Purview to

enable data unification, standardization, and cleansing that enables golden record creation and the publication of master data as data products.

Follow the tutorials here for your solution of choice: Master data management in Microsoft Purview

Step 6: Manage data health

In the **Health management** area of Unified Catalog, the central data office and other data managers can evaluate the status of the data against their company standards and effectively manage progress towards their strategy. To make sure that everyone in the company knows what they can do to increase the value of their data, it's essential that the standards are understood and scalable to the whole organization without needing to make everyone a data governance expert. Starting from an industry standard set of controls that are available out of the box, each data office can customize the controls to meet their expectations and ensure it aligns with their data goals. Critical to the effectiveness of these controls isn't only the measurement of these standards but also ensuring those responsible for the data can take action on their own and be held accountable for making the improvements that affect the value of data. In Data Estate Health, you can set and manage all of these critical capabilities.

Prerequisites

- Data products, glossary terms, and other business concepts published in Unified Catalog. You can follow the previous sections to create these concepts:
  - Section 1 - Set up a governance domain
  - Section 3 - Publish your data products
- At least 24 hours since the curation of data products.
- You must have the Data Health Owner role in Unified Catalog.

Evaluate your data governance with data estate health

1. Open the Microsoft Purview portal.
2. Select **Unified Catalog**.
3. Under **Data Estate Health** in the left navigation, select **Health controls**.
4. Select the carrot > next to the **Value Creation** control group.
5. While hovering over a control title, select the pencil icon to edit the control. By editing the control, you change the threshold of the control to set expectations for what the score should be and set the color scoring to demonstrate the progress stages.

6. The details enable you to provide a description of the control and what it means to your organization and set an owner for a specific control.

7. Select the **Rules** tab of the control to change the threshold. This setting has a high target and if it isn't healthy it's critical to follow up on.
   a. Inherit from group: toggle to switch off (should turn grey).
   b. Target score: 90
   c. Select **New rule**.
   d. Set the box next to the score to GreaterThanOrEqual
   e. Set the percentage to 90
   f. Status = Health (green)
   g. Else Box Status = Critical (Purple)
   h. Select **Save**.

8. Under data estate health, select **Metadata quality**.

Here you can change or add rules that create the scores of the control. Here you want to change the severity of the actions for **Value Creation** to ensure all users know the importance of this action.

1. Select **Configure severity**
2. Select the **Value Creation** control group
3. Select the **Business OKRs alignment** control title
4. Change the Severity from Medium to High and select **Save**
5. Select the **Health actions** tab
6. Filter **Assigned to:** to your name
7. Select an action where you can see what the owner of the action needs to do to ensure governance expectations are met or they can assign a new owner to get the best expert to provide their input. There's also a status that lets others know what work is ongoing and where other actions could need prioritization.

Step 7: Data democratization

Data democratization enables users to find and access the data they need in a compliant manner. It ensures people can find the data they need to build business value. Unified Catalog provides a clean and easy experience to discover data. It empowers stewards to update and manage the data made available in the catalog at scale. In this section, you learn how users can find and request access to data and ensure that the appropriate approvers can track and provide inputs on those access requests.

Prerequisites

- Completed steps 1-4 at minimum:
  - Section 1 - Set up a governance domain
  - Section 2 - Set up your data estate
  - Section 3 - Publish your data products
- A catalog reader role in one governance domain

Discover data products

1. In Unified Catalog, select **Discovery**, then select **Data products**.
2. On the **Data products** page, use the search bar to search for vaccination rates by age.
3. Here you see the data products you published in section 2. This view shows how users only see the data intended for them and prevents users from having to navigate a highly technical data estate.
4. Select the **Covid-19 Vaccination and Case Trending by Age** data product
   a. Here, consumers can see the metadata you provided and any of the other properties that you configured during setup. The data quality score is here as well so consumers know the quality before they even get access to the data.
   b. Select the asset and the consumer can see all of the columns that are available in the data asset.
   c. Select the Outbreak glossary term and the consumer can see the description and other information about the term to gain a deeper understanding of the data.
5. Once the consumer is confident that they want to use that data, they need to get approved access to the data.
   a. Select **Request access**
   b. Fill in the form detail to submit a request.
   c. User: leave your name
   d. Manager Approval: automatically required and directed to the Microsoft Entra ID manager.
   e. Purpose: select a purpose
   f. Business justification: OKR monitoring
   g. Check the box next to the attestation to say you understand the expectations to use this data.
   h. Select *Send*.

The access request is now sent to the listed manager in Microsoft Entra ID. From here, the manager can access the requests by opening the email and selecting a link or coming into Microsoft Purview. Approving and managing access can be done directly in Microsoft Purview.

1. In Unified Catalog, select **Catalog management**, then select **Requests**.
2. Select the **Personal Health** domain.
3. Select the request you submitted.

4. Now the approvers are able to approve or decline by selecting **Respond** on the request.

Data governance roles and permissions in Microsoft Purview

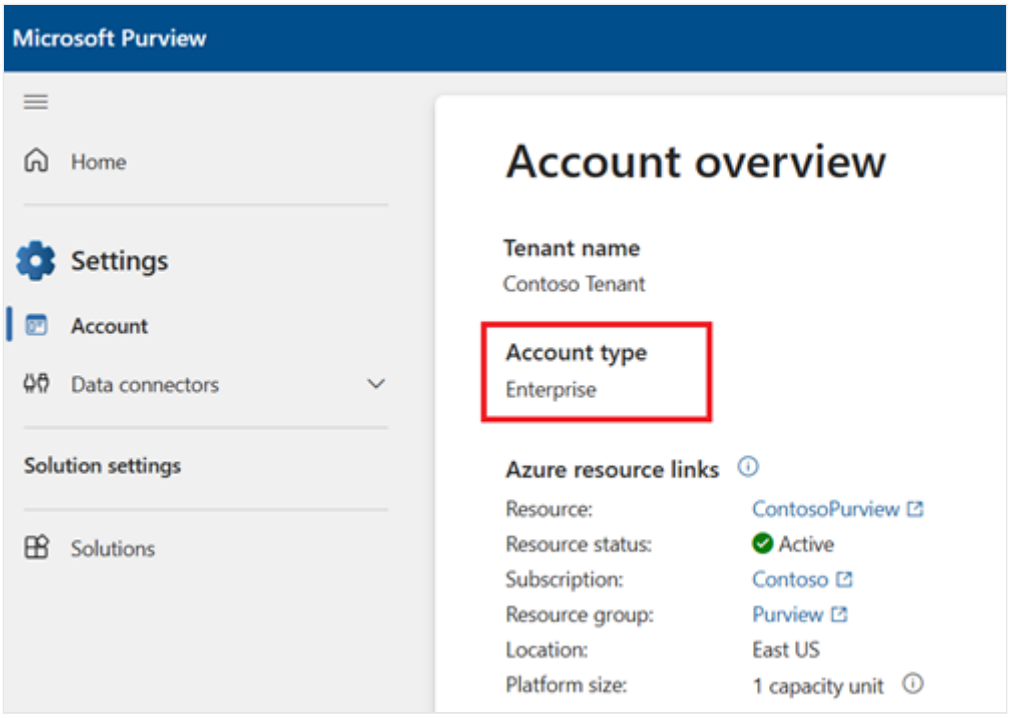
> **Important**
>
> This article covers Microsoft Purview data governance permissions in the new Microsoft Purview portal using Microsoft Purview Unified Catalog.
>
> - If you're using the classic Data Catalog, see governance permissions for the classic Data Catalog.
> - For data governance permissions in the classic Microsoft Purview portal, see permissions in the classic Microsoft Purview governance portal.

Microsoft Purview data governance has two solutions in the Microsoft Purview Portal: Data Map and Unified Catalog. These solutions use tenant/organizational-level permissions, existing data access permissions, and domain/collection permissions to provide users access to governance tools and data assets.

The kind of permissions available to use depends on your Microsoft Purview account type.

Check your account type

Check whether your organization has the free or enterprise type of account. To check your account type, go to the Microsoft Purview portal, select the **Settings** card. Under **Account**, view your **Account type**.



> **Important**
>
> For users newly created in Microsoft Entra ID, it might take some time for permissions to propagate even after correct permissions are applied.

Permissions in the enterprise version

All users can view data assets for available sources where they have at least **Read** permissions. Users who are owners can manage the metadata for assets where they have at least **Owner/Write** permissions. Learn more about Microsoft Azure roles.

Permissions types

- Tenant/organization permissions: Assigned at the organizational level, these permissions provide general and administrative permissions.
- Unified Catalog permissions: These permissions allow users to browse Unified Catalog and build out their data governance solutions.
- Data Map domain and collection level permissions: Permissions in Data Map that grant access to data assets in Microsoft Purview.

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your

organization. Learn more about Microsoft Purview **roles and permissions**.

Tenant level role groups

Assigned at the organizational level, tenant level role groups provide general and administrative permissions for both Microsoft Purview Data Map and Unified Catalog. If you're managing your Microsoft Purview account or your organization's data governance strategy, you probably need to belong to one or more of these role groups:

- Purview Administrators
- Data Source Administrators
- Data Governance

For more information, see descriptions of these roles.

For a full list of all available roles and role groups, not just for data governance, see roles, and role groups in the Microsoft Defender XDR and Microsoft Purview portals.

How to assign and manage role groups

A user must hold the *Role management* role in order to add users or groups to a Microsoft Purview role group. For instructions on assigning and managing roles in Microsoft Purview, see permissions in Microsoft Purview.

Unified Catalog permissions

Three levels of permissions grant users access to information in Unified Catalog:

1. Data governance tenant level: A tenant/organizational-level role group that has the Data Governance Admin role. That role delegates the first level of access for Governance Domain Creators. (This role isn't surfaced in Unified Catalog, but affects your ability to assign permissions in Unified Catalog.)

2. Catalog level permissions: Permissions that grant ownership to governance domains and access to health management.

3. Governance domain level permissions: Permissions to access and manage resources inside specific governance domains.

Permissions to search the full Unified Catalog

You don't need specific permissions in Unified Catalog to search the Unified Catalog. However, searching Unified Catalog only returns relevant data assets that you have permissions to view in Data Map. Users can find a data asset in Unified Catalog when:

- They search data products in a governance domain where they have catalog reader permissions.
- They have data reader permissions on a domain or collection in Data Map where the asset is stored.
- They have at least Read permissions on an available Azure or Microsoft Fabric resource.

> **Important**
>
> Users who already have **Read** permissions in Azure might have access to assets in Unified Catalog that you didn't intend for them to have. If you don't want them to have such access, you need to **remove their permissions**.

You manage permissions to these assets at the resource level and at the Data Map level.

If your catalog is well-curated, day-to-day business users shouldn't need to search the full catalog. They should be able to find the data they need in data products. For more information about setting up the Unified Catalog, see get started with data governance, and plan for Unified Catalog.

Catalog level permissions

Catalog level permissions provide high-level access within Unified Catalog and include these roles:

- Governance Domain Creator
- Global Catalog Reader
- Data Health Owner
- Data Health Reader

For more information, see descriptions of these roles.

How to assign catalog level roles

1. Sign in to the Microsoft Purview portal using credentials for an admin account that is assigned the *Data Governance* role.
2. Go to **Settings** and select **Unified Catalog**.

3. Select **Roles and permissions**.
4. Select **Governance Domain Creators** or another role, then select the add user icon.
5. Search for the user you want to add, then select the user.
6. Select **Save**.

Governance domain level permissions

Governance domain permissions provide access within a specific governance domain. Grant these permissions to data experts and business users so they can read and manage objects within the governance domain. Only a user with the **Governance domain owner** role can grant governance domain permissions.

See the list of these roles and descriptions.

> **Tip**
>
> A user in your organization who runs data governance or Unified Catalog should hold the **Governance domain owner** role. This role is essential for building entities such as governance domains, data products, and glossary terms. Assign at least two people as governance domain owners.

How to assign governance domain roles

A user needs to be a Governance Domain Owner in the governance domain to assign governance domain roles. Data Governance Administrators or Governance Domain Creators assign this role.

Assign governance domain roles under the **Roles** tab in a governance domain. For instructions on assigning roles, see how to manage governance domains.

Unified Catalog roles

> **Important**
>
> Be sure you understand the difference between the two Catalog Reader roles:
>
> **Global Catalog Reader** allows users to access published business concepts from published governance domains.
>
> **Local Catalog Reader** greatly reduces access to governance domains. When you assign users to this role within a governance domain:

- Only those users have reader access to published concepts within that governance domain.
- A user with this role can also hold other roles, such as Data Product Owner or Data Steward, that let them view published business concepts from other governance domains.

The **Local Catalog Reader** role is helpful when, for example, you need to limit access to a governance domain to meet regulatory or legal requirements. However, overuse of this role hinders a federated approach to data governance.

> **Note**
>
> Data observability uses the same roles as the rest of the Unified Catalog. Catalog Readers can't access the broader Data Observability explorer view in data estate health and they can only view published concepts. The Data Steward, Data Product Owners, Governance Domain Owners, and Governance Domain Creators have access to data observability views in both the concepts that they can view and the data health management view that enables broader exploration and view across the full data estate.

Expand table

Permission level	Role	Description
Tenant	Data Governance role group	Grants access to data governance roles within Microsoft Purview.
	Data Source Administrators role group	Manage data sources and data scans in the Microsoft Purview Data Map.
	Purview Administrators role group	Create, edit, and delete domains and perform role assignments.
Catalog	Data Governance Administrator	Delegates the first level of access for Governance Domain Creators and other catalog level permissions.
	Data Health Owner	Can create, update, and read artifacts in the **Health management** area of Unified Catalog.
	Data Health Reader	Can read artifacts in the **Health management** area of Unified Catalog.

Permission level	Role	Description
	Global Asset Curator	Can add published glossary terms to data assets and columns.
	Global Catalog Reader	Can read published artifacts across governance domains that don't have a Local Catalog Reader specified.
	Governance Domain Creator	Can create domains and delegate governance domain owner (or remains governance domain owner by default).
Governance domain	Data Product Owner*	Can create, update, and read data products only within their governance domain. Can read and build relationships with concepts in governance domains.
	Data Profile Reader	Has access to browse data profile insights and can drill down the profiling results to browse the statistics in column level. This is a sub-role that requires the user to also hold the **Governance Domain Reader** *and* either the **Global Catalog Reader** or **Local Catalog Reader** roles.
	Data Profile Steward	Can run data profiling jobs and access profiling insight details. This role can also browse through all data quality insights and can monitor profiling jobs. This role can't create rules and can't run data quality scanning. This is a sub-role that requires the user to also hold the **Governance Domain Reader** *and* **Data Product Owner** roles.
	Data Quality Metadata Reader	Can browse data quality insights (except profiling results column level insight), data quality rule definition, and rule level scores. This role can't access error records and can't run profiling and data quality scanning jobs. This is a sub-role that requires the user to also hold the **Governance Domain Reader** *and* either the **Global Catalog Reader** or **Local Catalog Reader** roles.
	Data Quality Reader	Can browse all data quality insights and data quality rules definitions. This role can't run data quality scanning and data profiling jobs, and can't access data profiling column level insight. This is a sub-role that requires the user to also hold the **Governance Domain Reader** *and* either the **Global Catalog Reader** or **Local Catalog Reader** roles.
	Data Quality Steward	Can use data quality features like data quality rule management, data quality scanning, browsing data quality insights, data quality scheduling, job monitoring, and configuring thresholds and alerts. This is a sub-role that requires the user to also hold the **Governance Domain Reader** *and* **Data Product Owner** roles.
	Data Steward*	Can create, update, and read artifacts and policies within their governance domain. Can also read artifacts from other governance domains.

Permission level	Role	Description
	Governance Domain Owner	Can delegate all other governance domain permissions, configure domain level data quality scan alerts, set up domain level schedule for data quality scanning job, and set domain level access policies.
	Governance Domain Reader	Can read governance domain metadata for published domains they're added to.
	Local Catalog Reader	Can read published concepts only in the governance domain they're granted access to. Because this role greatly limits the scope of who can access and manage data, it's recommended to limit use of this role so that you can better utilize a federated approach to data governance.

*To be able to add data assets to a data product, Data Product Owners and Data Stewards also need Data Map permissions to read those data assets in Data Map.

Data Map permissions

Data Map uses a set of predefined roles to control who can access what within the account. Domains and collections are tools used by Data Map to group assets, sources, and other artifacts into a hierarchy for discoverability, and to manage access control within Data Map.

Find a description of each role and learn how to add roles and restrict access through collections.

For more detailed information about the roles available in collections, see who should be assigned what roles or the collections example.

> **Tip**
>
> If you're a data steward or data product owner for Unified Catalog, it's a good idea to have Data Map permissions as well.

Data asset lifecycle example

To understand how permissions work between Data Map and Unified Catalog, review the table below on the full lifecycle of an Azure SQL table in the environment:

Expand table

Step	Role	Permission level
1. The Azure SQL Database is registered in Data Map	Data Source Administrator	Data Map
2. The Azure SQL Database is scanned in Data Map	Data Curator or Data Source Administrator	Data Map
3. The Azure SQL table is curated and certified	Data Curator	Data Map
4. A governance domain is created in the Microsoft Purview Account	Governance Domain Creator	Catalog
5. A data product is created in the governance domain	Governance Domain Owner and/or Data Product Owner	Governance domain
6. The Azure SQL table is added as an asset to the data product	Data Product Owner and/or Steward	Governance domain
7. An access policy is added to the data product	Data Product Owner and/or Steward	Governance domain
8. A user searches Unified Catalog, looking for data assets that match their needs	Asset permissions or data reader permission	Asset permissions or Data Map permissions
9. A user searches data products, looking for a product that matches their needs	Global Catalog Reader	Catalog
10. A user requests access to the resources in the data product	Global Catalog Reader	Catalog
11. A user views Data Health Insights to track the health of their Data Catalog	Data Health Reader	Catalog
12. A user wants to develop a new report to track data health progress in their catalog	Data Health Owner	Catalog

Last updated on 04/24/2026

Learn about data governance in the Microsoft Purview portal

09/18/2025

The modern governance experience is available in the Microsoft Purview portal for anyone who didn't create a Microsoft Purview account in the Azure portal before, and for existing customers. This article explains how to use this new experience.

The new governance experience

The Microsoft Purview portal, https://purview.microsoft.com, is a single, organization-wide instance of Microsoft Purview that's the next step towards unifying your organization's governance, policy, compliance, risk, and security. It's a prerequisite to a single platform used to manage and govern any data, both structured and unstructured, across your data estate. Including Azure, Microsoft 365, Azure, on-premises, or multicloud and SaaS applications in future.

- If you're a new Microsoft Purview customer, read about the new governance experience for new customers.
- If you're an existing Microsoft Purview customer, read about the new governance experience for existing customers.

Learn more about all the governance, data security, and risk and compliance features available in the Microsoft Purview portal.

How to use the new experience

If you're a new Microsoft Purview governance customer (if your organization doesn't have any Microsoft Purview accounts in any Azure subscriptions under your Microsoft Entra tenant), you can start using the Microsoft Purview portal right now for your governance journey.

If your organization already has Microsoft Purview accounts in Azure, after your organization is enabled, getting started with the new experience depends on your organization's current structure:

- You only have one Microsoft Purview account:
  - Your account region matches your tenant region
  - Your account region doesn't match your tenant region
  - Your account has disabled all public access
- You have multiple Microsoft Purview accounts

For more information, visit the new experience FAQ.

One Microsoft Purview account

Your account has disabled all public network access

If your Microsoft Purview account had disabled all public network access and is using private endpoints, you need to make some changes before you upgrade it. You can continue to use *ingestion* private endpoints, but not *account* or *portal* private endpoints. The new experience uses a new private endpoint called a *platform* private endpoint.

To check your current public network settings:

1. Go to the Azure portal.
2. Search for your Microsoft Purview account.
3. Select **Networking** under the **Settings** menu.
4. If your firewall is set to **Disabled from all networks**, you need to follow the next steps to upgrade. If your firewall is set to **Enabled from all networks** or **Disabled for ingestion only**, you should be able to upgrade without these steps.

To upgrade an account using private endpoints, you need to:

1. Confirm that your firewall allows these global and tenant-specific endpoints (replacing the `accountname` and `tenantid` with your values):
   a. `api.purview-service.microsoft.com`
   b. `accountname.purview.azure.com`
   c. `tenantid-api.purview-service.microsoft.com`
2. After upgrading to the new experience, you can follow this guide to set up private endpoints for governance solutions in the Microsoft Purview portal.

> **Important**
>
> 1. Currently, Azure Data Factory, Azure Machine Learning, and Azure Synapse lineage connections are not supported with the platform private endpoint, and might not work after switching. If your production replies on these lineage connections, wait until these connections are supported. (Azure Synapse is supported as a data source using a managed Virtual Network Integration Runtime and platform private endpoint.)
> 2. Scans that run after upgrading but before you set up the platform private endpoints might fail and need to be re-run after private endpoint set up.

Your region matches your tenant region

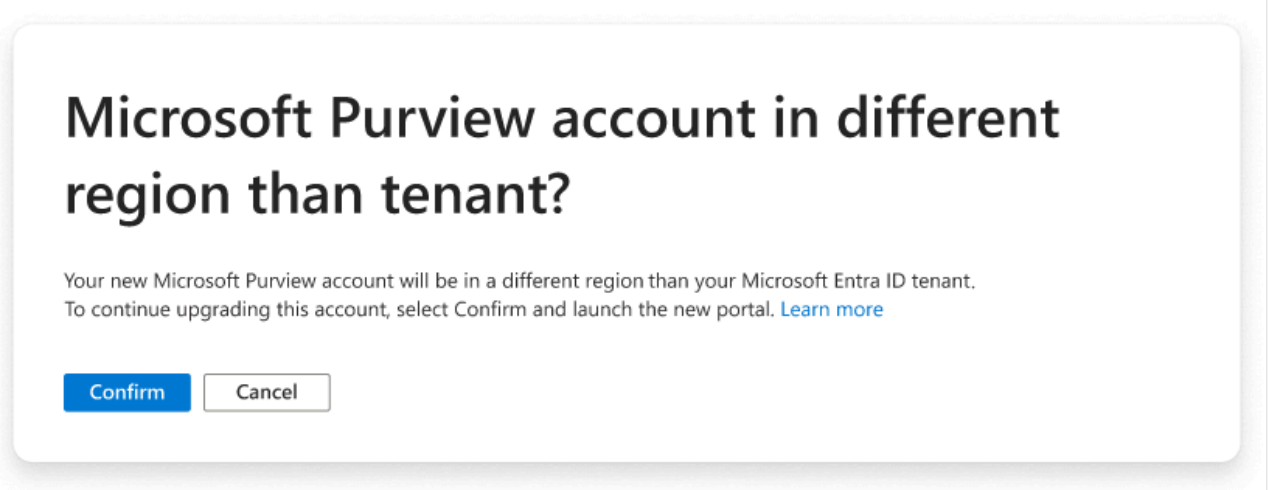
If your Microsoft Purview account's region matches your Microsoft Entra tenant region, and your networking settings are either **enabled from all networks** or **disabled for ingestion only**, you can use the new experience; no upgrade is required.

Your region doesn't map to your tenant region

1. A Microsoft Purview Admin selects the **Upgrade** button in the Azure portal, Microsoft Purview portal, or upgrade email.

   > **Note**
   >
   > If you don't have access to the Microsoft 365 role groups to assign users to the Microsoft Purview Admin role, contact support.

2. If your account is in a different region than your tenant, an admin **confirms** setup, as seen in this image:

3. After confirmation, the new portal launches.

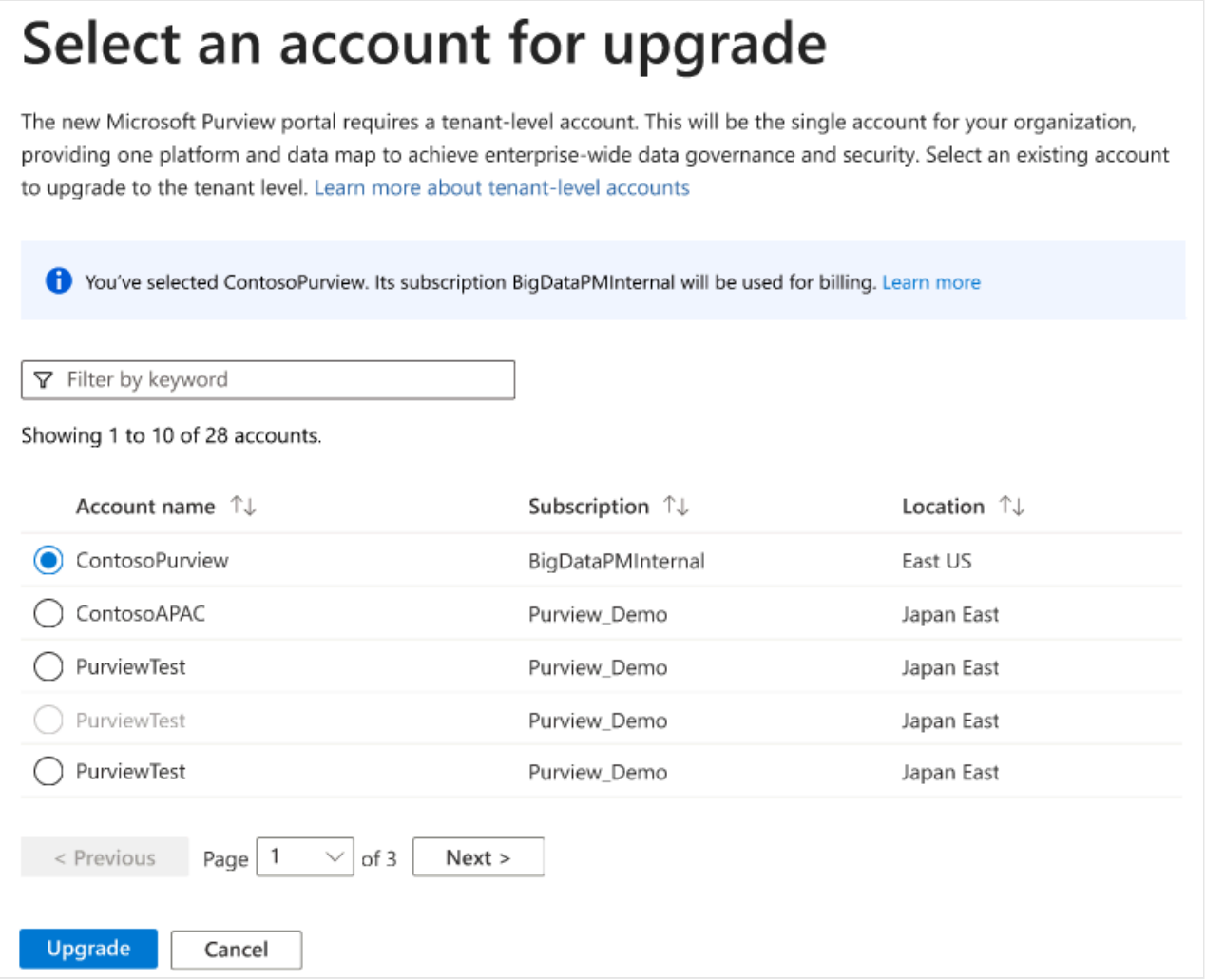
Multiple Microsoft Purview accounts

The new experience requires a single, tenant-level/organization-wide account that's the primary account for your organization. If you have multiple Microsoft Purview accounts in your tenant, you can upgrade one to the new experience and merge up to four other accounts into your primary account as domains.

1. If your Microsoft Purview account has disabled public network access and is using account or portal private endpoints, follow the steps at Disabled all public network access to upgrade.

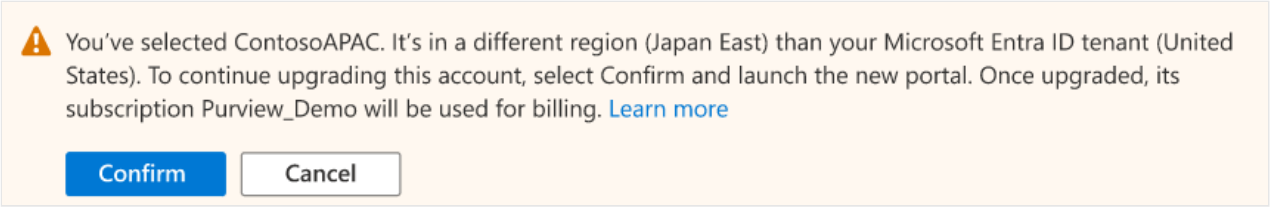
2. A Microsoft Purview Admin selects the **Upgrade** button in the Azure portal, Microsoft Purview portal, or upgrade email.

3. Select an existing account to upgrade it to be your organization-wide account. See How to choosing a primary account.



> **Note**
>
> Upgrading one account doesn't affect your other Microsoft Purview accounts or their data. After upgrading, you can merge your secondary accounts into your primary account as domains.
>
> If you don't have access to the Microsoft 365 role groups to assign users to the Microsoft Purview Admin role, contact support.

4. If the account you select is in a different region than your tenant, an admin needs to confirm setup.



5. After confirmation, the new portal launches.
6. You can take the tour and begin exploring the new Microsoft Purview experience!
7. Merge your secondary accounts into your primary account as domains.

Disabled all public network access

If your Microsoft Purview account had disabled all public network access and is using private endpoints, you need to make some changes before you upgrade it. You can continue to use *ingestion* private endpoints, but not *account* or *portal* private endpoints. The new experience uses a new private endpoint called a *platform* private endpoint.

To check your current public network settings:

1. Go to the Azure portal.
2. Search for your Microsoft Purview account.
3. Select **Networking** under the **Settings** menu.
4. If your firewall is set to **Disabled from all networks**, you need to follow the next steps to upgrade. If your firewall is set to **Enabled from all networks** or **Disabled for ingestion only**, you should be able to upgrade without these steps.

To upgrade an account using private endpoints, you need to:

1. Confirm that your firewall allows these global and tenant-specific endpoints (replacing the `accountname` and `tenantid` with your values):
   a. `api.purview-service.microsoft.com`
   b. `accountname.purview.azure.com`
   c. `tenantid-api.purview-service.microsoft.com`
2. After upgrading to the new experience, you can follow this guide to set up private endpoints for governance solutions in the Microsoft Purview portal.

> **Important**
>
> 1. Currently, Azure Data Factory, Azure Machine Learning, and Azure Synapse lineage connections are not supported with the platform private endpoint, and might not work after switching. If your production replies on these lineage connections, wait

until these connections are supported. (Azure Synapse is supported as a data source using a managed Virtual Network Integration Runtime and platform private endpoint.)
2. Scans that run after upgrading but before you set up the platform private endpoints might fail and need to be re-run after private endpoint set up.

Why upgrade to the new experience

If you're already using the classic Microsoft Purview governance portal, upgrading your Microsoft Purview account to the new experience offers several benefits, including:

- **Prerequisites for unified experience**: The upgrade is a prerequisite and one step to bring your environment closer to use data governance, security, risk, and compliance capabilities under the same platform.
- **Improved user experience**: The new experience provides a more streamlined and intuitive user interface, making it easier for users to access and manage their data assets in the new Microsoft Purview portal.
- **Enhanced features**: Upgrading unlocks new features and capabilities, such as live view, preset scans, and more.
- **Consolidated management**: The enhanced experience offers a centralized management interface for all your Microsoft Purview resources, simplifying administration tasks.
- **Access to new capabilities**: All new Microsoft Purview features are only available in the upgraded tenant level account.

After upgrading an existing account to the new experience, here's what to expect:

- **Account consolidation**: You access all other Microsoft Purview accounts in your tenant through the classic portal. Be prepared to merge those accounts with the primary account when the feature becomes available. Microsoft notifies customers who need to perform other steps to merge these accounts.
- **Roles and permissions**: Fine grained access control via roles and permissions at collection scopes continues to function as-is after your accounts are upgraded. In addition, there are new tenant-level roles that you can manage in the new portal. For more information, see our tenant-level permissions documentation.
- **Adapting to new features:** Educate users within your organization on the new features and capabilities introduced in the upgraded account, enabling them to maximize the

benefits of these enhancements. This education includes understanding and using the following features:
    - Live view
    - Tenant-level permissions

Accessing support

With the free version of Microsoft Purview governance solutions, you can't open support tickets for Microsoft Purview. You need to upgrade to enterprise access for support.

How to choose a primary account

If you have a single Microsoft Purview account, your account is likely automatically upgraded. When you have multiple Microsoft Purview accounts under the same Microsoft Entra tenant, consider the following factors while selecting an account for the upgrade. Select an account as primary tenant-level account by following guidelines below:

- **Data assets**: Choose the account with the most valuable or frequently used data assets, as this account becomes the primary data map after the upgrade.
- **Account usage**: Evaluate how each account is currently being used and its role within your organization. Accounts that are considered "production" are most viable to select for this purpose.
- **Permissions and access controls**: Consider the existing permissions and access controls for each account, as they're carried over to the upgraded environment (as an isolated Domain).
- **System Assigned Managed Identity (SAMI) and User Assigned Managed Identity (UAMI)**: The existing system and user assigned managed identities for the primary upgraded account can be used across all domains in the Microsoft Purview instance.

Other considerations

- You can access the upgraded account through current endpoints and new endpoints, so there's no immediate effect if you use Microsoft Purview APIs. Any updates you perform in the upgraded account also reflect in the classic portal for the same Microsoft Purview account.
- New Microsoft Purview features are only available for the upgraded account.

- For other accounts in your tenant, you can continue to access through the classic governance portal.

Regions

> **Note**
>
> The regions listed here are for the single, organization-wide instance of Microsoft Purview that uses the Microsoft Purview portal.
>
> Regions available for Microsoft Purview Unified Catalog are different. Visit Supported regions for Unified Catalog.

This table lists the regions that are currently available to use Microsoft Purview in the new portal:

Expand table

Purview Account Location	Mapped Microsoft Entra Country/Region	Tenant Location Code
Australia East	Australia	AU
	Fiji	FJ
	New Zealand	NZ
Brazil South	Argentina	AR
	Bolivia (Plurinational State of)	BO
	Bonaire Sint Eustatius and Saba	BQ
	Brazil	BR
	Chile	CL
	Colombia	CO
	Curaçao	CW
	Ecuador	EC
	Falkland Islands (the)	FK
	French Guiana	GF

Purview Account Location	Mapped Microsoft Entra Country/Region	Tenant Location Code
	Guyana	GY
	Peru	PE
	Paraguay	PY
	Suri	SR
	Sint Maarten (Dutch part)	SX
	Uruguay	UY
	Venezuela (Bolivarian Republic of)	VE
Canada Central	Canada	CA
Central India	India	IN
	Sri Lanka	LK
	Nepal	NP
Central US	United States of America (the)	US
East US	Antigua and Barbuda	AG
	Anguilla	AI
	Aruba	AW
	Barbados	BB
	Bermuda	BM
	Bahamas (the)	BS
	Cuba	CU
	Dominica	DM
	Dominican Republic (the)	DO
	Grenada	GD
	Guadeloupe	GP
	Haiti	HT
	Jamaica	JM
	Saint Kitts and Nevis	KN

Purview Account Location	Mapped Microsoft Entra Country/Region	Tenant Location Code
	Cayman Islands (the)	KY
	Saint Lucia	LC
	Martinique	MQ
	Panama	PA
	Puerto Rico	PR
	Trinidad and Tobago	TT
	Saint Vincent and the Grenadines	VC
	Virgin Islands (British)	VG
	Virgin Islands (U.S.)	VI
	United States of America (the)	US
East US 2	United States of America (the)	US
France Central	France	FR
Germany West Central	Germany	DE
Japan East	Japan	JP
Korea Central	Korea (the Republic of)	KR
North Europe	Finland	FI
	Faroe Islands (the)	FO
	Ireland	IE
	Iceland	IS
	Moldova (the Republic of)	MD
	Sweden	SE
Qatar Central	Qatar	QA
South Africa North	Angola	AO
	Burkina Faso	BF
	Burundi	BI
	Benin	BJ

Purview Account Location	Mapped Microsoft Entra Country/Region	Tenant Location Code
	Botswana	BW
	Congo (the Democratic Republic of the)	CD
	Central African Republic (the)	CF
	Congo (the)	CG
	Côte d'Ivoire	CI
	Cameroon	CM
	Cabo Verde	CV
	Djibouti	DJ
	Algeria	DZ
	Egypt	EG
	Eritrea	ER
	Ethiopia	ET
	Gabon	GA
	Ghana	GH
	Gambia (the)	GM
	Guinea	GN
	Equatorial Guinea	GQ
	Guinea-Bissau	GW
	Kenya	KE
	Comoros (the)	KM
	Liberia	LR
	Lesotho	LS
	Libya	LY
	Morocco	MA
	Madagascar	MG
	Mali	ML

Purview Account Location	Mapped Microsoft Entra Country/Region	Tenant Location Code
	Mauritania	MR
	Mauritius	MU
	Malawi	MW
	Mozambique	MZ
	Namibia	NA
	Niger (the)	NE
	Nigeria	NG
	Réunion	RE
	Rwanda	RW
	Seychelles	SC
	Sudan (the)	SD
	Saint Helena Ascension and Tristan da Cunha	SH
	Sierra Leone	SL
	Senegal	SN
	Somalia	SO
	South Sudan	SS
	São Tomé and Príncipe	ST
	Eswatini	SZ
	Chad	TD
	Togo	TG
	Tunisia	TN
	Tanzania United Republic of	TZ
	Uganda	UG
	Mayotte	YT
	South Africa	ZA
	Zambia	ZM

Purview Account Location	Mapped Microsoft Entra Country/Region	Tenant Location Code
	Zimbabwe	ZW
Southeast Asia	Armenia	AM
	American Samoa	AS
	Brunei Darussalam	BN
	Cocos (Keeling) Islands (the)	CC
	Cook Islands (the)	CK
	China	CN
	Christmas Island	CX
	Micronesia (Federated States of)	FM
	Guam	GU
	Hong Kong SAR	HK
	Indonesia	ID
	Cambodia	KH
	Kiribati	KI
	Lao People's Democratic Republic (the)	LA
	Marshall Islands (the)	MH
	Myanmar	MM
	Macao SAR	MO
	Northern Mariana Islands (the)	MP
	Malaysia	MY
	New Caledonia	NC
	Norfolk Island	NF
	Nauru	NR
	Niue	NU
	French Polynesia	PF
	Papua New Guinea	PG

Purview Account Location	Mapped Microsoft Entra Country/Region	Tenant Location Code
	Philippines (the)	PH
	Pitcairn	PN
	Palau	PW
	Solomon Islands	SB
	Singapore	SG
	Thailand	TH
	Tokelau	TK
	Tonga	TO
	Tuvalu	TV
	Taiwan (Province of China)	TW
	United States Minor Outlying Islands (the)	UM
	Viet Nam	VN
	Vanuatu	VU
	Wallis and Futuna	WF
	Samoa	WS
South Central US	United States of America (the)	US
Switzerland North	Switzerland	CH
UAE North	United Arab Emirates (the)	AE
UK South	United Kingdom of Great Britain and Northern Ireland (the)	GB
West Central US	United States of America (the)	US
West Europe	Andorra	AD
	Albania	AL
	Austria	AT
	Åland Islands	AX
	Azerbaijan	AZ
	Bosnia and Herzegovina	BA

Purview Account Location	Mapped Microsoft Entra Country/Region	Tenant Location Code
	Belgium	BE
	Bulgaria	BG
	Bahrain	BH
	Belarus	BY
	Cyprus	CY
	Czechia	CZ
	Denmark	DK
	Estonia	EE
	Spain	ES
	Georgia	GE
	Guernsey	GG
	Gibraltar	GI
	Greece	GR
	Croatia	HR
	Hungary	HU
	Israel	IL
	Isle of Man	IM
	Italy	IT
	Jersey	JE
	Jordan	JO
	Kuwait	KW
	Kazakhstan	KZ
	Lebanon	LB
	Liechtenstein	LI
	Lithuania	LT
	Luxembourg	LU

Purview Account Location	Mapped Microsoft Entra Country/Region	Tenant Location Code
	Latvia	LV
	Monaco	MC
	Montenegro	ME
	Republic of North Macedonia	MK
	Malta	MT
	Netherlands (the)	NL
	Norway	NO
	Oman	OM
	Pakistan	PK
	Poland	PL
	Portugal	PT
	Romania	RO
	Serbia	RS
	Russian Federation (the)	RU
	Saudi Arabia	SA
	Slovenia	SI
	Svalbard and Jan Mayen	SJ
	Slovakia	SK
	San Marino	SM
	Türkiye	TR
	Ukraine	UA
	Holy See (the)	VA
	Yemen	YE
West US	United States of America (the)	US
West US 2	United States of America (the)	US
West US 3	Costa Rica	CR

Purview Account Location	Mapped Microsoft Entra Country/Region	Tenant Location Code
	Guatemala	GT
	Honduras	HN
	Mexico	MX
	Nicaragua	NI
	El Salvador	SV
	New Mexico	NM
	United States of America (the)	US

For information about all available Microsoft Purview regions, see Azure products by regions.

> **Important**
>
> Microsoft 365 GCC regions aren't currently supported.

Resources

- New Microsoft Purview portal experience for data governance frequently asked questions
- Share and receive data
- Data lineage in classic Data Catalog
- Data owner policies
- Classification in Data Map

New Microsoft Purview portal experience for data governance FAQ

Where is the new portal?

The new Microsoft Purview portal is located at https://purview.microsoft.com. The classic data governance portal is located at https://web.purview.azure.com.

How can I use the new portal experience?

The new portal experience is available to all organizations. If you're using Microsoft Purview for the first time and you use Power BI or Microsoft Fabric, you'll automatically have access. If not, create a new account and you're automatically upgraded to the enterprise version of Microsoft Purview within a few weeks. For existing Microsoft Purview users, learn more about using the new experience if you have one Microsoft Purview account or if you have multiple Microsoft Purview accounts.

Is there a cost to use the enterprise version of Microsoft Purview?

Yes. Learn more about pricing at Microsoft Purview pricing.

Can I create new Microsoft Purview accounts in my subscription?

It depends:

- If your organization created an account after August 1, 2023, you can't create more accounts.
- If your organization hasn't created a Microsoft Purview account before, if you create an account it will automatically be reconciled to the new experience, and you can't create more accounts after.
- If your organization has a single Microsoft Purview account that has already been reconciled to the new experience, you can't create more accounts.
- If you're using the free version of Microsoft Purview governance solutions and haven't created a Microsoft Purview account, if you upgrade to the enterprise version, you can't create more

accounts.

- If you're using the free version of Microsoft Purview governance solutions and create an account, the newly created account will be automatically reconciled to the new experience and you can't create more accounts after that.
- If your organization already has multiple Microsoft Purview accounts, you can create Microsoft Purview accounts up to the preexisting quota, even if you have already reconciled an account to the new experience.

We already had a Microsoft Purview account, and are using the new portal experience. Is our account the free or enterprise version of Microsoft Purview?

If you're already using a Microsoft Purview account that's upgraded to use the new portal experience, your organization is using the enterprise version of Microsoft Purview. You can check your organization's account type in the Microsoft Purview portal by going to **Settings** > **View all settings**.

When I start using the new Microsoft Purview experience, does it create an account for my organization or just for me?

The new Microsoft Purview experience uses a single, organization-wide instance of Microsoft Purview.

What does this mean for my existing Microsoft Purview governance portal accounts?

One of your Microsoft Purview governance portal accounts are elevated to be the primary tenant/organization level account with all its current metadata and information. If you have other Microsoft Purview accounts, in the future their information will also be made available in the primary account.

Can multiple accounts in the same tenant have access to the new experience?

Not currently. The new experience uses a single, primary Microsoft Purview account. If you have multiple accounts, you select a primary account when you upgrade to the new experience. In the future, the information in your other accounts will also be made available in this environment.

What is the impact of upgrading the production account to the enhanced experience?

There's no impact on upgrading the account to the enhanced experience. Users can navigate the capabilities in the new Microsoft Purview portal and classic governance portal in Azure. If users browse Microsoft Purview portal from the Azure portal, the upgraded account provides two links to browse the portals.

How long does the upgrade take?

If your account is eligible, upgrading to the new experience only takes a few minutes.

What if my organization blocked the new portal endpoints on the firewall?

You can still access Microsoft Purview through the classic portal by launching the old endpoints through https://web.purview.azure.com.

What should I do if my organization can't use preview features?

Upgrading to enhanced experience doesn't change the status of generally available (GA) features.

My primary Microsoft Purview instance has private endpoints enabled. Can I upgrade it to enhanced experience?

If your Microsoft Purview account had disabled all public network access and is using private endpoints, you can still upgrade to the new experience. You can continue to use your *ingestion* private endpoints as is, but the new experience also uses a *platform* private endpoint. After upgrading to the new experience, you can follow this guide to set up platform the private endpoint for governance solutions in the Microsoft Purview portal. For more information, see what to do if your account has disabled all public network access.

How is access to the Microsoft Purview portal granted if you're a Dynamics 365, Microsoft 365 Enterprise, Microsoft Fabric, Power Apps, or Power BI user?

If your organization uses Dynamics 365, Microsoft 365 Enterprise, Microsoft Fabric, Power Apps, or Power BI, your tenant has the Microsoft Purview Discovery Plan added. Administrators can find this plan under **Your products** in the admin center. This plan has no cost and grants access to the new experience.

What if I'm not an Azure Enterprise, Dynamics 365, Microsoft 365 Enterprise, Microsoft Fabric, Power Apps, or Power BI customer?

If you're an existing Microsoft Purview customer (your Microsoft Entra tenant already has a Microsoft Purview account), a user holding a Purview Administrator role can upgrade your instance to the new experience. When that happens, you're prompted to confirm the upgrade. When you confirm, your tenant is added to the Microsoft Purview Discovery Plan. Administrators can find this plan under **Your products** in the admin center. This plan has no cost and grants access to the new experience. If a global or billing admin attempts this upgrade, they receive an error. If you don't have access to the Microsoft 365 role groups to assign Purview Admins, contact support.

My organization wants to continue using the classic portal. What link can we use to access it?

Access the classic portal at https://web.purview.azure.com.

If I continue to scan or curate assets in the existing classic portal, will the new or updated assets show up in the new portal?

Yes. The upgraded account is accessible via both portals, and changes in metadata and configurations are visible in the classic and new portals.

Can I opt out of migration to the new experience?

Existing Microsoft Purview accounts stay the same and remain operational in parallel with the new experience for a time. Eventually users are encouraged to transition to the new experience.

Should I start the upgrade now?

We recommend all Microsoft Purview customers to start the upgrade as soon as possible. New features and capabilities are only available on the new portal.

What do I need to prepare before the upgrade and what are some areas of considerations to include in my plan?

1. Prepare your change request. Before the upgrade, Microsoft Purview users can navigate to the classic Microsoft Purview governance portal only through legacy endpoints:

Portal: `web.purview.azure.com`
Endpoint: `https://[accountname].purview.azure.com/`

After the upgrade, the upgraded account is accessible via legacy experience and the new experience. To use the new experience, make sure your corporate firewalls allow outbound traffic to new endpoints.

Access to classic portal and endpoints:

Portal: `web.purview.azure.com`
Endpoint - `https://[accountname].purview.azure.com/`

Access to new portal and endpoints:
Portal: `purview.microsoft.com`
Endpoint: `api.purview-service.microsoft.com`

2. Add new endpoints to your corporate firewalls. Even though the classic portal and endpoints are still available, as part of retiring these endpoints, Microsoft suggests planning to include new URLs and endpoints to your firewall. These updates make sure your users' access and applications aren't impacted when the legacy endpoints are retired.

3. Adjust permissions. The upgraded account can be accessed by users through Microsoft Purview roles. New roles are also added to your Microsoft Purview account at the tenant level. You can manage these roles in new portal. Learn more about data governance roles and permissions.

What if my organization has multiple Microsoft Purview accounts?

You can upgrade your primary account. All other Microsoft Purview accounts in your tenant can still be accessed through the classic portal. Be prepared to merge those accounts with the primary account when the feature becomes available. Microsoft will notify customers who require to perform more steps to merge these accounts.

We have multiple Microsoft Purview accounts and want to upgrade to use the new portal. Which account should we select as the "primary"?

We suggest you select the production account, the account that is used by most of your end users, or the account that's the most critical to your daily operations. See our guide to choosing your primary account.

If I can't create a new account, what are my other options?

Use collections and assign role-based access on those collections to provide granular access control.

What can I access in the free version of Microsoft Purview governance solutions?

The free version allows users to explore their Azure and Microsoft Fabric resources. See what's available in the free version.

What region is Microsoft Purview in if we use the free version?

The free version of Microsoft Purview governance solutions uses the same location mappings as Microsoft 365, and maps to a location based on your Microsoft Entra ID home region.

How does Microsoft Purview choose the region for the enterprise version of Microsoft Purview?

If your tenant has no Microsoft Purview account: The free version of Microsoft Purview governance solutions is mapped to the same location as your Microsoft 365 instance. If that's unavailable, then it falls back to using the location geographically closest to your Microsoft Entra ID region. When you upgrade from free to enterprise, the enterprise version uses the same location.

If your tenant has a Microsoft Purview account: If your Microsoft Purview account is already upgraded to the new experience, it's using the region it was created in or the region that was selected during upgrade. If you haven't upgraded a Microsoft Purview account to the new experience, if you upgrade an existing account or create a new account to upgrade then the enterprise version will use that account's region.

The Microsoft Purview region is incidental; it doesn't affect your configurations or how you use your data, even if your sources are in another region. Microsoft Purview does all data processing

local to your data sources, and doesn't move your data to its own region or store your data. Your data remains securely in its own location, and ingress and egress charges that other Azure products sometimes incur aren't applicable. Regardless of where your data sources are housed, you don't need to change the region your Microsoft Purview account is using to interact with your sources.

If you're trying to upgrade from the free version to enterprise and can't upgrade due to security policy or data residency constraints around Microsoft Purview's mapped region, contact support. Also contact support if you're trying to create a first Microsoft Purview account in the Azure portal and can't create the account in the selected region due to security policy or data residency constraints.

If you're moving because of a security policy, you can only move to another region in your tenant's current Azure Geography. If you need to move to a region outside your tenant's current geography, you need to create a tenant in that geography.

How can I get more governance capabilities?

All governance features and capabilities are available by upgrading to the enterprise version of Microsoft Purview.

Last updated on 10/10/2025

Classic Microsoft Purview governance portal

Article • 01/28/2025

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

This article gives an overview of some of the features in the classic Microsoft Purview governance portal.

> **Tip**
>
> Microsoft Purview has a new experience! If you're a new Microsoft Purview customer, or if you want more information, see our **article on the new portal** or the article about our new **data governance experience**.

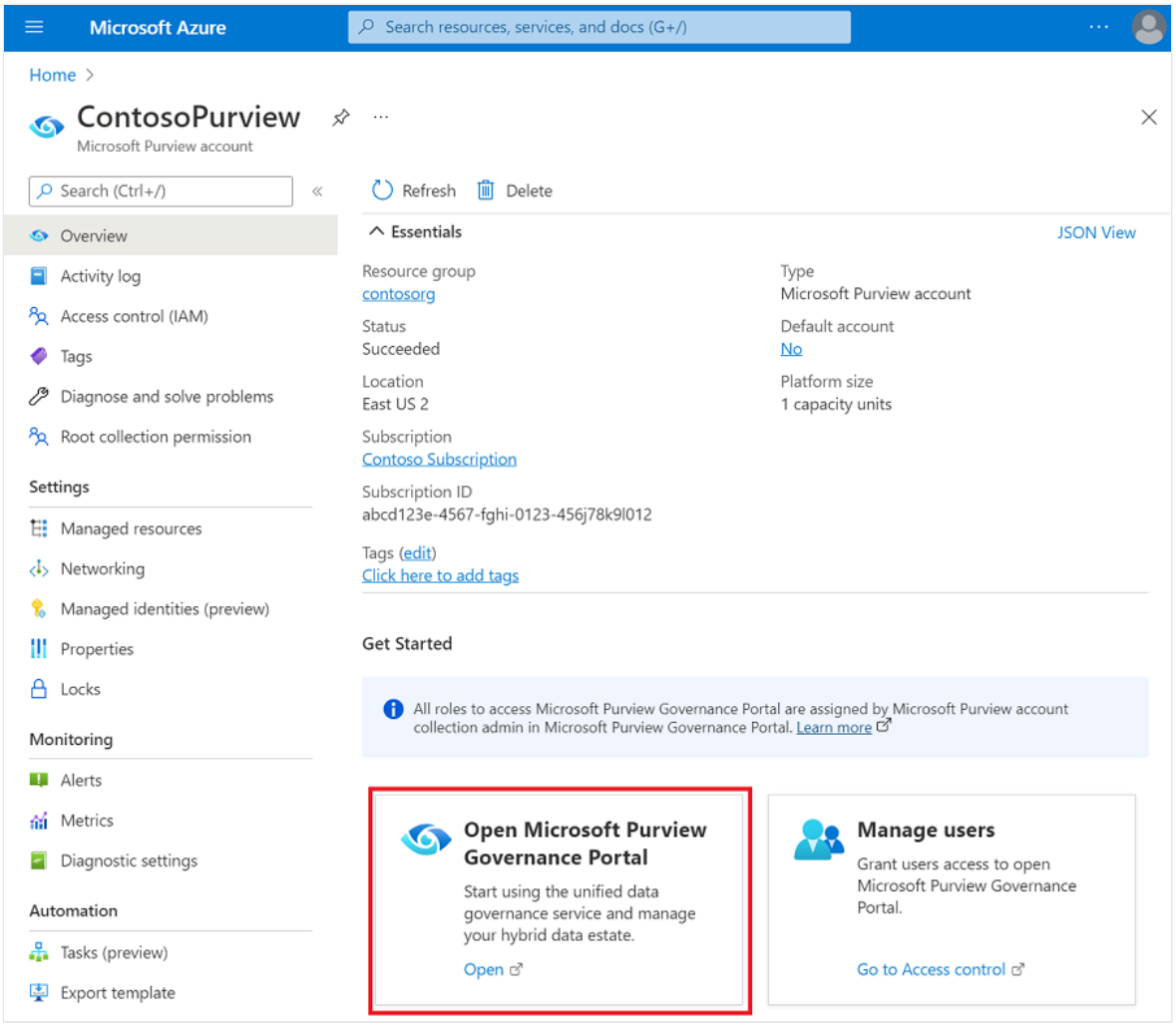
Prerequisites

- An Active Microsoft Purview account is already created in Azure portal
- The user has permissions to access the classic Microsoft Purview governance portal.

Launch Microsoft Purview account

- You can launch the Microsoft Purview account directly by going to `https://web.purview.azure.com`, selecting **Microsoft Entra ID** and the account name. Or by going to `https://web.purview.azure.com/resource/yourpurviewaccountname`

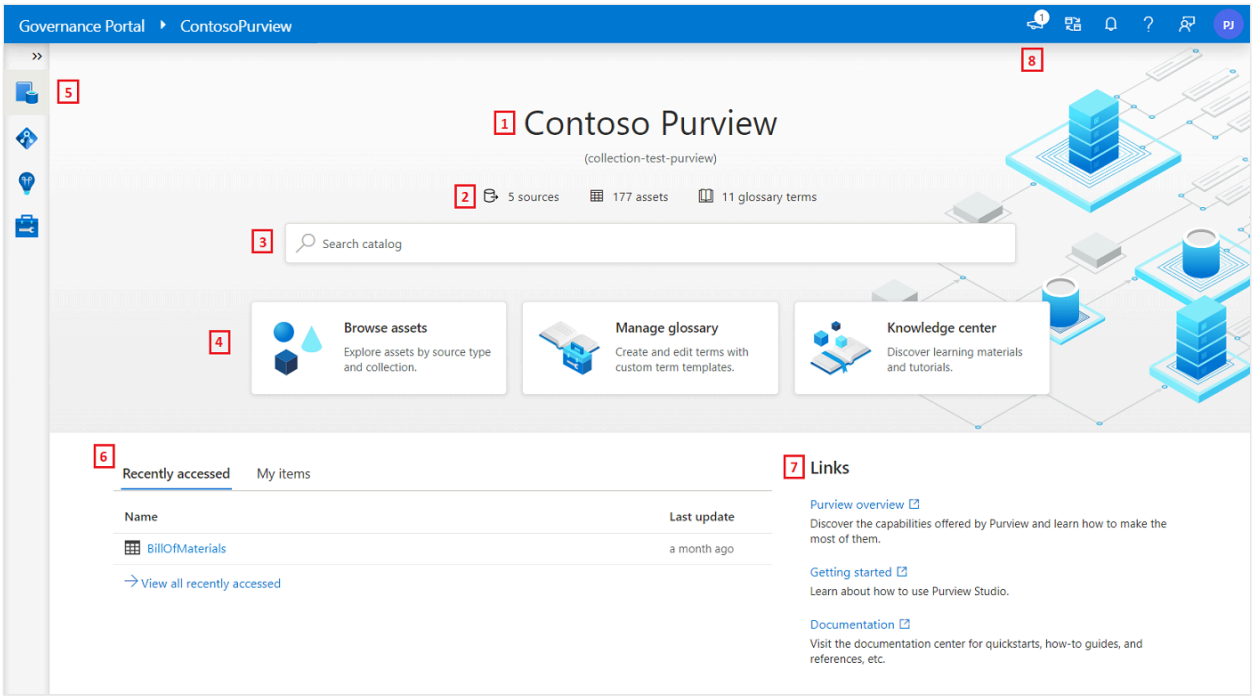
- To launch your Microsoft Purview account from the Azure portal, go to Microsoft Purview accounts in Azure portal, select the account you want to launch and launch the account.



> **Tip**
>
> If you can't access the portal, confirm you have the necessary permissions.

Home page

Home is the starting page for the Microsoft Purview governance portal.



The following list summarizes the main features of **Home page**. Each number in the list corresponds to a highlighted number in the preceding screenshot.

1. Friendly name of the catalog. You can set catalog name in **Management** > **Account information**.

2. Catalog analytics shows the number of:

    - Data sources
    - Assets
    - Glossary terms

3. The search box allows you to search for data assets across the data catalog.

4. The quick access buttons give access to frequently used functions of the application. The buttons that are presented, depend on the role assigned to your user account at the root collection.

    - For *collection admin*, the available button is **Knowledge center**.
    - For *data curator*, the buttons are **Browse assets**, **Manage glossary**, and **Knowledge center**.
    - For *data reader*, the buttons are **Browse assets**, **View glossary**, and **Knowledge center**.
    - For *data source admin* + *data curator*, the buttons are **Browse assets**, **Manage glossary**, and **Knowledge center**.
    - For *data source admin* + *data reader*, the buttons are **Browse assets**, **View glossary**, and **Knowledge center**.

> **Note**
>
> For more information about Microsoft Purview roles, see Access control in Microsoft Purview.

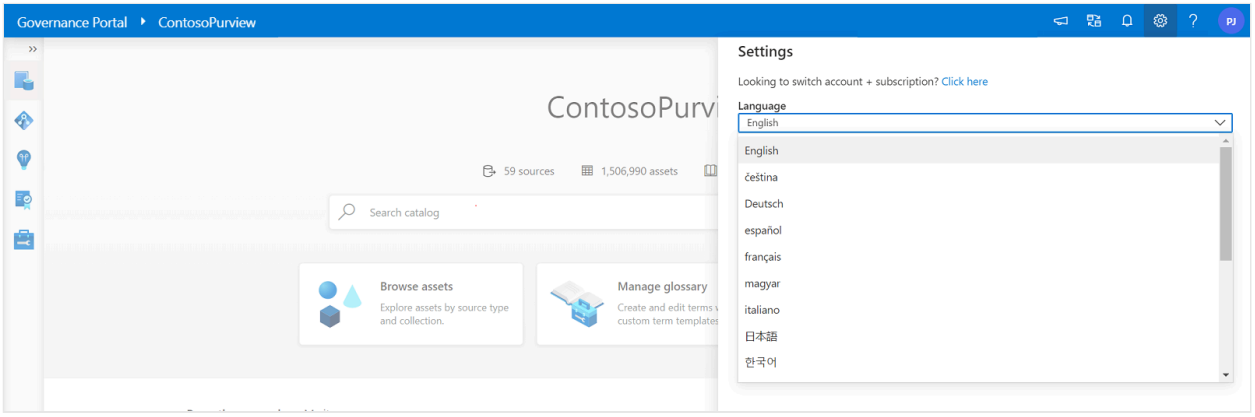
5. The left navigation bar helps you locate the main pages of the application.
6. The **Recently accessed** tab shows a list of recently accessed data assets. For information about accessing assets, see Search the Data Catalog and Browse by asset type. **My items** tab is a list of data assets owned by the logged-on user.
7. **Links** contains links to region status, documentation, pricing, overview, and Microsoft Purview status
8. The top navigation bar contains information about release notes/updates, change purview account, notifications, help, and feedback sections.

Knowledge center

The knowledge center is where you can find all the videos and tutorials related to Microsoft Purview.

Localization

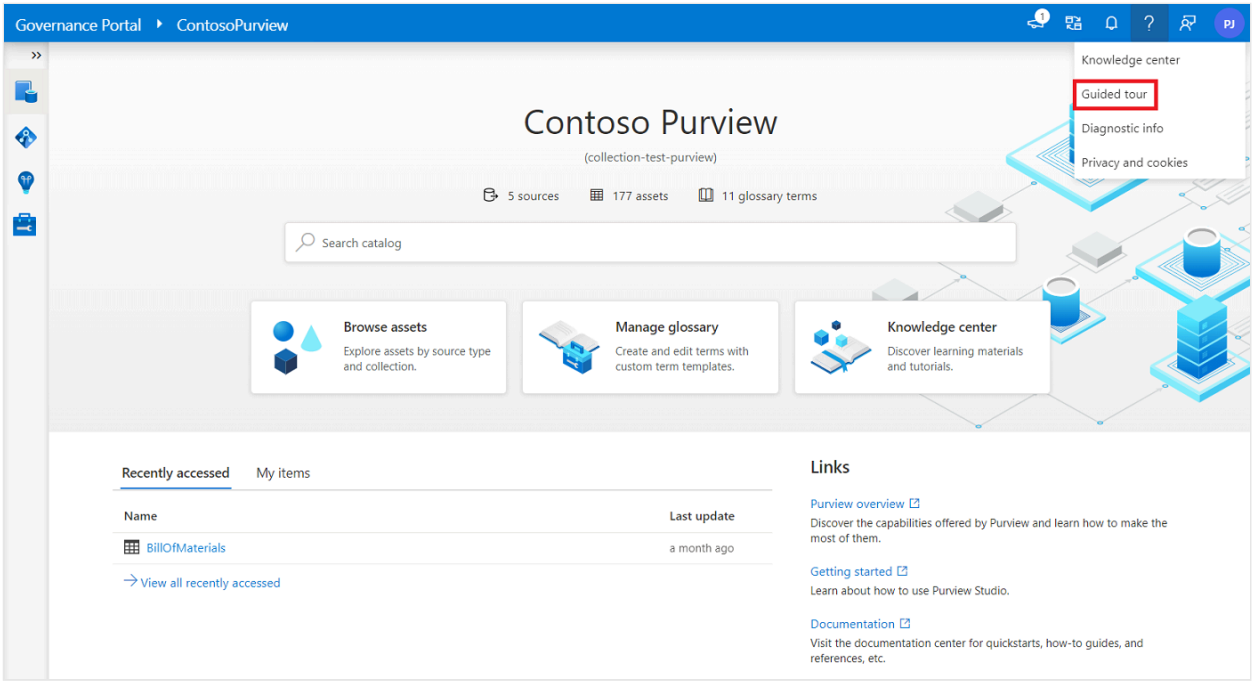
Microsoft Purview is localized in 18 languages. To change the language used, go to the **Settings** from the top bar and select the desired language from the dropdown.



> **Note**
>
> Only generally available features are localized. Features still in preview are in English regardless of which language is selected.

Guided tours

Each UX in the Microsoft Purview governance portal will have guided tours to give overview of the page. To start the guided tour, select **help** on the top bar and select **guided tours**.



Next steps

Scan data

Feedback

Was this page helpful? Yes No

Learn about data governance billing

Microsoft Purview data governance solutions include Microsoft Purview Unified Catalog and Microsoft Purview Data Map. For Unified Catalog, you pay for governed assets and data health management capabilities. The billing for these components operates on a pay-as-you-go model. There's no cost to scan assets into Microsoft Purview Data Map when certain billing conditions are met. This article explains how to prepare for billing and how data governance capabilities are charged.

For complete pricing information, with pricing calculators, see the Microsoft Purview pricing site.

To learn more about the pay-as-you-go billing model for Microsoft Purview, including the consent experience, see Microsoft Purview billing models.

Prerequisites

To use Microsoft Purview pay-as-you go, you need:

- An Azure subscription in the same tenant as Microsoft Purview.
- An Azure resource group in that subscription.

If you already have these resources for other purposes, you can also use them with Microsoft Purview.

Billing setup

To use the data governance experience in the Microsoft Purview portal, you need to be set up for Microsoft Purview pay-as-you-go billing. Pay-as-you-go took effect for Microsoft Purview on January 6, 2025. For details about required roles and the setup experience, visit one of the resources below based on your situation:

- If you're a new data governance customer using Unified Catalog in the Microsoft Purview portal, see Enable Microsoft Purview pay-as-you-go features for new customers.
- If you're an existing Azure Purview customer using the classic data governance pricing model, see Consent to use pay-as-you-go capabilities.

If you don't consent to the pay-as-you-go pricing model, you can still use the classic Data Catalog in the classic data Microsoft Purview portal.

Your Data Map assets remain intact regardless of whether you consent to the new billing experience for Unified Catalog.

Pricing components

Two pricing components support data governance in Microsoft Purview. Unified Catalog has two meters that run based on:

1. Number of unique governed assets per day; see Unified Catalog billing.
2. Data governance processing units per run; see Data health management.

Unified Catalog billing

A single meter starts for an organization when users begin to govern unique data assets. A unique data asset can be a table or a view linked to a governance concept, such as a data product or a critical data element.

Assets that are collected in Data Map but aren't linked to a governance concept don't count as governed assets.

> **Tip**
>
> - To view specific pricing information and a calculator for your region, see the **Data Governance** tab on the Microsoft Purview pricing site.
> - For an admin view into your organization's consumption across governance domains, see Usage monitoring (preview).

What are governed assets and how are you billed

In Unified Catalog, when you actively manage and curate data assets, you turn them into **governed assets**. A governed asset is a data asset that you attach or associate to governance concepts, such as data products, critical data elements, glossary terms, and data quality. If you don't associate a technical asset to a governance concept, it isn't a governed asset.

For example, if you have a Microsoft SQL Server with 200 tables, and only 20 tables are in the gold tier of the medallion structure, you created 20 data products with each data product linked to a table. Every day, the Unified Catalog billing system counts the tables that are associated to the data product. Since only 20 tables link to data products, only 20 tables are governed assets, instead of the full 200 tables.

Over time, you might want to attach a glossary term and a policy to the data product, which then also apply to the governed assets. In each case, a governed asset is still a single governed asset regardless of how many concepts you attach to the asset.

- **Example 1:** A SQL table is referenced within a data product. You use the same SQL table in five other data products. The SQL table is counted one time only, once a day.
- **Example 2:** You create 50 governance domains and data products, but don't attach any tables, files, reports, or dashboards. In this case, you aren't charged for any governed assets.
- **Example 3:** You accidentally attach a server to the data product. Microsoft Purview counts just the server as a single asset, not all the children tables within the server.

View pricing information on the **Data Governance** tab on the Microsoft Purview pricing site.

Data health management usage

The Microsoft Purview Data Governance Enterprise Data Management meter starts when you run data quality and health management actions on your governed data estate. You pay for data management usage based on the data governance processing unit (DGPU) pay-as-you-go meters. A DGPU is the amount of service performance consumed for 60 minutes and comes in three different performance options: Basic, Standard, and Advanced. The Basic SKU option is the default performance option until you select a higher option.

View pricing information at each performance level on the **Data Governance** tab on the Microsoft Purview pricing site.

Data governance processing units explained

A DGPU is a fully managed compute unit that runs compute-heavy capabilities such as data quality and data health management. Each DGPU provides 60 minutes of compute time, running across varying sets of nodes based on the workload need.

The number of DGPUs you consume depends on:

1. Data quality or health (metadata quality) rule type: Out of the box or custom.
2. The volume of data.
3. Source type: The same volume of data might generate different DGPUs across two different source types.

Some ranges of DGPU generation are as follows:

Expand table

Data Source	Row count	SKU type	Rule type	DGPU per rule per run
Azure SQL DB	1 Million	Basic	Empty/blank check	0.02
Azure SQL DB	1 Million	Basic	StringFormat=Regex/Like check	0.02
Azure SQL DB	1 Million	Basic	Table lookup (1 mill row reference table)	0.03
Azure SQL DB	1 Million	Basic	Unique check	0.02
Azure SQL DB	1 Million	Basic	Duplicate check = 3 column-combo	0.02

Data health controls pricing

Data health controls, self-serve analytics, and reports within **Health management** use the Basic SKU, as these capabilities don't require higher processing capacity.

The total DGPU consumption depends on the data volume, the calculation of control scores, and report generation. The control job is scheduled to run daily when you initially provision it and triggers only if at least one business domain exists. If you remove all business domains, the job doesn't run.

You can adjust the schedule as needed; the job always runs based on your configured schedule. If you don't want to refresh a specific control, you can edit the control and toggle it to deactivate.

This flexibility lets you fully manage schedules and selectively inactivate controls to optimize costs and manage DGPU consumption.

> **Tip**
>
> - Default control refreshes daily. Scheduled controls run based on the schedule you configure.
> - To stop using data health controls, manually disable them from the schedule page. You can activate or deactivate individual controls separately, so you can focus on the most relevant controls.

Data quality pricing

Data quality usage is billed based on DGPU pay-as-you-go meters. For example, if you run 100 data quality jobs in a single day, and each run produces 0.02 DGPU, then the total DGPU for that day is two DGPUs. Data quality for on-premises data sources uses the Basic SKU for DGPU consumption. To calculate the cost based on performance option and region, visit the Microsoft Purview pricing site.

You can also use a data quality and data health control pricing calculator on the Microsoft Azure Pricing calculator page to calculate pricing for governed assets, DGPU consumption price for data quality, and data health controls. Use the calculator to budget the investment for the data governance program of your organization. On the calculator page, search for Microsoft Purview to calculate price and budget the data governance investment.

Here's example of consumed processing units for basic to complex rules for different data volumes, based on the standard performance option.

Expand table

Rule complexity	10,000 records	-	100,000 records	-	1,000,000 records	-	10,000,000 records	-	100,000,000 records	-	1,000,000,00 records
	Duration	PU	Duration	PU	Duration	PU	Duration	PU	Duration	PU	Duration
Simple	Elapsed time: 1m	0.02	Elapsed time: 1m	0.02	Elapsed time: 1m	0.02	Elapsed time: 1m	0.02	Elapsed time: 1m	0.02	Elapsed time 1m 31s

Rule complexity	10,000 records	-	100,000 records	-	1,000,000 records	-	10,000,000 records	-	100,000,000 records	-	1,000,000,00 records
	1s		1s		1s		16s		16s		
Medium	Elapsed time: 1m 1s	0.02	Elapsed time: 1m 1s	0.02	Elapsed time: 1m 1s	0.02	Elapsed time: 1m 16s	0.02	Elapsed time: 1m 31s	0.03	Elapsed time 2m 1s
High	Elapsed time: 1m 1s	0.02	Elapsed time: 1m 1s	0.02	Elapsed time: 1m 31s	0.03	Elapsed time: 1m 32s	0.03	Elapsed time: 2m 1s	0.03	Elapsed time 2m 51s

Tip

- Increase data volume: For testing, you can use the same data asset by adding more data, or you can create multiple data assets with different data volumes; see the example above.
- Example of rule complexity
  - Simple rule: Blank or empty value check.
  - Medium complex rule: Uniqueness rule, data type, or format check rule.
  - High complex rule: Duplicate check with multiple selected columns; credit card, national ID, table lookup, and SSN validation are in the category of complex rule.

Data Map and scanning charges

Unified Catalog currently runs on the existing Data Map. Data Map and scan charges don't apply to Unified Catalog customers once they either consent to the Microsoft Purview billing model or upgrade from free to enterprise tier. The new Unified Catalog is rolling out in multiple regions; check the supported regions for Unified Catalog.

Customers using the classic Data Catalog can check pricing on the **Data Map (Classic)** tab of the Microsoft Purview pricing site.

Resources

- Data governance billing FAQ
- Learn about Microsoft Purview billing models
- Microsoft Purview pricing site
- Unified Catalog supported regions

Last updated on 04/24/2026

Data governance billing frequently asked questions

When did the new pricing model go into effect?

The new pricing model took effect on January 6, 2025.

Is the asset counting done at a monthly level or daily level?

Assets are counted on a daily basis. Each day, unique governed assets are counted. If you ungovern any assets mid-month, the right asset count is picked for the day and multiplied by unique governed assets per day.

I have a single data product with three data assets attached. Am I charged for each data asset for 30 days?

Yes. When attaching three data assets to a data product, you're charged for each data asset each day.

I have the same data asset attached to both a data product and a critical data element. How am I charged?

You aren't charged for a data asset attached to both a data product and a critical data element. You're charged for a data asset once.

How am I charged if I scanned 10 data assets and only three are attached to data products?

You're charged for the three data assets attached to the data product. You aren't charged for the remaining seven data assets in Data Map.

If I use the classic glossary in Data Map and attach classic glossary terms to a data asset, am I charged?

No. Assets with classic glossary terms attached to them aren't charged.

If I attach a SQL server with 200 tables in it to a data product, how many governed assets am I charged for?

You're charged for a single asset as governed asset for a SQL server with many tables. Children won't be counted since the intent is to govern the server as a whole. We recommend putting an individual table or a selected group of tables in a data product, so the experience is more curated and meaningful. If the data product is data consumer facing, then limited data consumers would want access to a data product with 200 tables.

If I attach a Power BI semantic model into a data product that has 200 tables, and then I intentionally hand-pick 10 tables to be part of the data product, how many assets do I get charged for?

If the semantic model is still attached to the data product, we charge it as a single asset. If you have picked five tables to be attached to the data product, removing the semantic model we charge for the five tables only.

Can I run data quality rules on a database or its individual tables and columns?

Data quality runs are set on columns of tables. Running data quality on a data asset's columns or tables turns that data asset into a governed asset.

What type of billing transparency will I be able to see?

Usage monitoring (preview) in Unified Catalog settings provides Data Governance Administrators with a view of governed assets and data governance processing unit (DGPU) consumption across governance domains. Based on the processing units, you can decide if you're over-using, under-using, or if there are opportunities to optimize data quality rules. You can also charge back costs to the different governance domains using the consumption details.

Are there tools and guidance to help estimate costs with new pricing model?

Yes. You can find complete pricing information and a pricing calculator on the Microsoft Purview pricing site.

Is there a volume discount?

No, but based on usage at scale, we can work on volume-based pricing for future releases.

What cost reporting will be available, and where can I view this?

Costs can be viewed from January 6, 2025, in the Cost management section of the subscription tied to your Microsoft Purview tenant (in the Microsoft Purview portal, go to **Settings** > **Account** > **Account overview**). Data Governance Admininisrators can also view the usage monitoring (preview) report.

What meters are zeroed out from classic to Microsoft Purview Unified Catalog?

The following meters are zeroed out from classic to Microsoft Purview Unified Catalog in the following areas:

- Data Map population - Scanning including SHIR
- Data Map enrichment - Insights Generation, Data Map capacity units
- Data Map consumption - including insights consumption

Does the advanced resource set meter zero out for me?

The advanced resource set isn't offered as part of Microsoft Purview, so the meter is ineffective.

Last updated on 02/13/2026

Microsoft Purview data governance glossary

This glossary provides a brief description of important terms and concepts for the Microsoft Purview data governance experience.

A

Expand table

Term	Description
Advanced resource sets	A set of features activated at the Microsoft Purview instance level that, when enabled, enrich resource set assets by computing extra aggregations on the metadata to provide information such as partition counts, total size, and schema counts. Resource set pattern rules are also included.
Annotation	Information that is associated with data assets in the Microsoft Purview Data Map, for example, glossary terms and classifications. After they're applied, annotations can be used within Search to aid in the discovery of the data assets.
Approved	The state given to any request that has been accepted as satisfactory by the designated individual or group who has authority to change the state of the request.
Asset	Any single object that is stored within a Microsoft Purview Unified Catalog. Note: A single object in the catalog could potentially represent many objects in storage, for example, a resource set is an asset but it's made up of many partition files in storage.

C

Expand table

Term	Description
Capacity unit	A measure of data map usage. All Microsoft Purview Data Maps include one capacity unit by default, which provides up to 10 GB of metadata storage and has a throughput of 25 data map operations per second.
Classification report	A report that shows key classification details about the scanned data.
Classification rule	A set of conditions that determine how scanned data should be classified when content matches the specified pattern.

Term	Description
Classified asset	An asset where Microsoft Purview extracts schema and applies classifications during an automated scan. The scan rule set determines which assets get classified. If the asset is considered a candidate for classification and no classifications are applied during scan time, an asset is still considered a classified asset.
Classifier	The technology used to identify, categorize, classify, and label data. Microsoft Purview offers several classifiers, including sensitive info types, trainable classifiers, and exact data match (EDM) classifiers.
Collection	An organization-defined grouping of assets, terms, annotations, and sources. Collections allow for easier fine-grained access control and discoverability of assets within a Unified Catalog.
Collection admin	A role that can assign roles in the Microsoft Purview governance portal. Collection admins can add users to roles on collections where they're admins. They can also edit collections, their details, and add subcollections.
Column pattern	A regular expression included in a classification rule that represents the column names that you want to match.
Contact	An individual who is associated with an entity in Unified Catalog.
Control plane operation	An operation that manages resources in your subscription, such as role-based access control and Azure policy that are sent to the Azure Resource Manager end point. Control plane operations can also apply to resources outside of Azure across on-premises, multicloud, and SaaS sources.
Credential	A verification of identity or tool used in an access control system. Credentials can be used to authenticate an individual or group to grant access to a data asset.

D

Expand table

Term	Description
Data curator	A role that provides access to Unified Catalog to manage assets, configure custom classifications, set up glossary terms, and view insights. Data curators can create, read, modify, move, and delete assets. They can also apply annotations to assets.
Data dictionary	A list of canonical names of database columns and their corresponding data types. Often used to describe the format and structure of a database, and the relationship between its elements.
Data Estate Insights	An area of the Microsoft Purview governance portal that provides up-to-date reports and actionable insights about the data estate.

Term	Description
Data Map	A metadata repository that is the foundation of the Microsoft Purview governance portal. The data map is a graph that describes assets across a data estate and is populated through scans and other data ingestion processes. This graph helps organizations understand and govern their data by providing rich descriptions of assets, representing data lineage, classifying assets, storing relationships between assets, and housing information at both the technical and semantic layers. The data map is an open platform that can be interacted with and accessed through Apache Atlas APIs or the Microsoft Purview governance portal.
Data Map operation	A create, read, update, or delete action performed on an entity in the data map. For example, creating an asset in the data map is considered a data map operation.
Data owner	An individual or group responsible for managing a data asset.
Data pattern	A regular expression that represents the data that is stored in a data field. For example, a data pattern for employee ID could be Employee{GUID}.
Data plane operation	An operation within a specific Microsoft Purview instance, such as editing an asset or creating a glossary term. Each instance has predefined roles, such as "data reader" and "data curator" that control which data plane operations a user can perform.
Data reader	A role that provides read-only access to data assets, classifications, classification rules, collections, glossary terms, and insights.
Data Share contributor	A role that can share data within an organization and with other organizations using data share capabilities in Microsoft Purview. Data share contributors can view, create, update, and delete sent and received shares.
Data Sharing	Microsoft Purview Data Sharing is a set of features in Microsoft Purview that enables you to securely share data across organizations.
Data source admin	A role that can manage data sources and scans. A user in the Data source admin role doesn't have access to Microsoft Purview governance portal. Combining this role with the Data reader or Data curator roles at any collection scope provides Microsoft Purview governance portal access.
Data steward	An individual or group responsible for maintaining nomenclature, data quality standards, security controls, compliance requirements, and rules for the associated object.
Discovered asset	An asset that the Microsoft Purview Data Map identifies in a data source during the scanning process. The number of discovered assets includes all files or tables before resource set grouping.
Distinct match threshold	The total number of distinct data values that need to be found in a column before the scanner runs the data pattern on it. For example, a distinct match threshold of eight for employee ID requires that there are at least eight unique data values among the sampled values in the column that match the data pattern set for employee ID.

E

Expand table

Term	Description
Expert	An individual within an organization who understands the full context of a data asset or glossary term.

F

Expand table

Term	Description
Full scan	A scan that processes all assets within a selected scope of a data source.
Fully Qualified Name (FQN)	A path that defines the location of an asset within its data source.

G

Expand table

Term	Description
Glossary	A searchable list of specialized terms that an organization uses to describe key business words and their definitions. Using a business glossary can provide consistent data usage across the organization.
Glossary term	An entry in the glossary that defines a concept specific to an organization. Glossary terms can contain information on synonyms, acronyms, and related terms.

I

Expand table

Term	Description
Incremental scan	A scan that detects and processes assets that have been created, modified, or deleted since the previous successful scan. To run an incremental scan, at least one full scan must be completed on the source.

Term	Description
Ingested asset	An asset that has been scanned, classified (when applicable), and added to the Microsoft Purview Data Map. Ingested assets are discoverable and consumable within Unified Catalog through automated scanning or external connections, such as Azure Data Factory and Azure Synapse.
Insight reader	A role that provides read-only access to Data Estate Insights reports. Insight readers must have at least data reader role access to a collection to view reports about that specific collection.
Integration runtime	The compute infrastructure used to scan in a data source.
Item	Any single object that's stored within the Microsoft Purview Data Map.

L

Expand table

Term	Description
Lineage	How data transforms and flows as it moves from its origin to its destination. Understanding this flow across the data estate helps organizations see the history of their data, and aid in troubleshooting or impact analysis.

M

Expand table

Term	Description
Management	An area within the Microsoft Purview Governance Portal where you can manage connections, users, roles, and credentials. Also referred to as "Management center."
Microsoft Fabric	Microsoft's unified analytics solution that delivers an integrated and simplified experience for all analytics workloads and users on an enterprise grade data foundation with pervasive data governance.
Microsoft Purview instance	A single Microsoft Purview (formerly Azure Purview) account.
Minimum match threshold	The minimum percentage of matches among the distinct data values in a column that must be found by the scanner for a classification to be applied. For example, a minimum match threshold of 60% for employee ID requires that 60% of all distinct

Term	Description
	values among the sampled data in a column match the data pattern set for employee ID. If the scanner samples 128 values in a column and finds 60 distinct values in that column, then at least 36 of the distinct values (60%) must match the employee ID data pattern for the classification to be applied.

O

Expand table

Term	Description
Object type	A categorization of assets based upon common data structures. For example, an Azure SQL Server table and Oracle database table both have an object type of table.
On-premises data	Data that is in a data center controlled by a customer, for example, not in the cloud or software as a service (SaaS).
Owner	An individual or group in charge of managing a data asset.

P

Expand table

Term	Description
Pattern rule	A configuration that overrides how the Microsoft Purview Data Map groups assets as resource sets and displays them within the catalog.
Physical asset	An asset that represents a physical data object. Physical assets are different from business assets because they represent real data. For example, a database is a physical asset.
Policy	A statement or collection of statements that controls how access to data and data sources should be authorized.

R

Expand table

Term	Description
Registered source	A source that has been added to a Microsoft Purview instance and is now managed as a part of Unified Catalog.

Term	Description
Related terms	Glossary terms that are linked to other terms within the organization.
Resource set	A single asset that represents many partitioned files or objects in storage. For example, the Microsoft Purview Data Map stores partitioned Apache Spark output as a single resource set instead of unique assets for each individual file.
Role	Permissions assigned to a user within a Microsoft Purview instance. Roles, such as Microsoft Purview Data Curator or Microsoft Purview Data Reader, determine what can be done within the product.
Root collection	A system-generated collection that has the same friendly name as the Microsoft Purview account. All assets belong to the root collection by default.

S

Expand table

Term	Description
Scan	A Microsoft Purview Data Map process that discovers and examines metadata in a source or set of sources to populate the data map. A scan automatically connects to a source, extracts metadata, captures lineage, and applies classifications. Scans can be run manually or on a schedule.
Scan rule set	A set of rules that define which data types and classifications a scan ingests into a catalog.
Scan trigger	A schedule that determines the recurrence of when a scan runs.
Schema classification	A classification applied to one of the columns in an asset schema.
Search	A feature that allows users to find items in Unified Catalog by entering in a set of keywords.
Search relevance	The scoring of data assets that determine the order search results are returned. Multiple factors determine an asset's relevance score.
Self-hosted integration runtime	An integration runtime installed on an on-premises machine or virtual machine inside a private network that is used to connect to data on-premises or in a private network.
Sensitive info type (SIT)	A pattern-based classifier that helps detect sensitive information, such as credit card numbers, in items.

Term	Description
Sensitivity label	A label managed in the Microsoft Purview portal that defines how confidential an item is. Sensitivity labels can be configured to apply associated protection settings, helping users to remain compliant with organizational information protection policies.
Sensitivity label report	A summary that describes which sensitivity labels are applied across the data estate.
Service	A product that provides standalone functionality and is available to customers by subscription or license.
Share	A group of assets that are shared as a single entity.
Source	A system where data is stored. Sources can be hosted in various places such as a cloud or on-premises. You register and scan sources so that you can manage them in the Microsoft Purview governance portal.
Source type	A categorization of the registered sources used in Microsoft Purview; for example, Microsoft 365, Azure SQL Database, Azure Blob Storage, Amazon S3, Google Cloud, or SAP ECC.
Steward	An individual who defines the standards for a glossary term. They're responsible for maintaining quality standards, nomenclature, and rules for the assigned entity.

T

Expand table

Term	Description
Term template	A definition of attributes included in a glossary term. Users can either use the system-defined term template or create their own to include custom attributes.
Trainable classifier	A type of classifier that helps identify and categorize content that isn't easily identified by manual or automated pattern-matching methods. Unlike a sensitive info type, this method of categorization identifies an item based on what the item is (for example, a resume), not solely by elements included in the item (pattern matching). Trainable classifiers can be included in several Microsoft Purview solutions to detect, protect, and govern sensitive data.

U

Expand table

Term	Description
Unified Catalog	A searchable inventory of assets and their associated metadata that allows users to find and curate data across a data estate. Unified Catalog also includes a business glossary where subject matter experts can provide terms and definitions to add a business context to an item.

W

Expand table

Term	Description
Workflow	An automated process that coordinates the creation and modification of catalog entities, including validation and approval. Workflows define repeatable business processes to achieve high quality data, policy compliance, and user collaboration across an organization.

Last updated on 11/07/2025

Learn about Microsoft Purview Unified Catalog

10/14/2025

> **Tip**
>
> The new Unified Catalog experience is rolling out across environments. If you don't see it yet, check the schedule for your region. Make sure you've upgraded to the enterprise version of the new Microsoft Purview experience in order to access the new experience when it's available in your region.

The goal of Microsoft Purview Unified Catalog is to not only provide a platform for data governance but also to drive business value creation in your organization.

Historically, data governance was a defense mechanism, a way to make sure your data is secure and compliant. But good data governance makes your data more visible to your users and provides many opportunities to reunite your business with the data that fuels it.

Most importantly, all of these new features are available in a single, integrated SaaS framework. Users shouldn't need to switch between applications to practice good data governance. We're working toward a solution where everything is in one place with experiences for data consumers, data stewards, and data owners.

Data governance for the era of AI

This era of artificial intelligence means that you have more data than ever, more ways to use it, and even more motivation to make sure you properly use and secure it. It's a task many organizations already struggle with, without the added complexity of using AI tools.

Governing data for an entire organization requires both rigor and flexibility. Clean, secure data requires consistency, but your teams' needs are unique for access and management. Because of this, use a federated governance approach: provide a centralized place to develop data safety, quality, and standards, but provide tools to create self-service access control, discoverability, and maintenance. Federated data governance spreads ownership across your business, reducing bottlenecks and encouraging participation in the life cycle of managing, governing, consuming, and applying data.

Data governance isn't just a gatekeeping function. When practiced well, it also accelerates data value creation. Your data users, stakeholders, and subject matter experts have crucial insights to the data-to day operations of your organization. Good data governance can apply that

expertise to reasonably scale the practice of data governance across your entire organization, including your nontechnical functions and business users. Said another way, good data governance utilizes your whole team and matches your data with your data-to-day business functions.

The point is both to reveal your data's business value and simplify its management even as your organization and data estate grow. Here's how it can help each sector of your business:

- For organization-wide data consumers:
  - Data discovery - help you easily find the data you need
  - Secure access - facilitates safe access to your data
  - Data understanding - providing what you need to know about the data before and while you use it

- For data owners and stewards:
  - Data curation and management - helps you deliver high quality data that's easy to understand and safely access for organization-wide applications
  - Responsible data use - helps you ensure that your data is used by intended users for intended purposes
  - Impact analysis - understand action anomaly states impacting your data

- For data officers and CxO stakeholders:
  - Data value creation - maximize value creation from your data while reducing operations spend
  - Data estate standardization - create common controls across your data estate with federated accountability so your data is healthy and safe.

Governance with the Unified Catalog

The new Unified Catalog experience lets you explore and understand your data categorized by governance domains. You can search through an AI-powered copilot and subscribe to data products that come equipped with all the data you need and the tools to safely access it. Over the last couple of years, we invested in a strong platform that has an inventory of all your data assets, their metadata, and their lineage so you can understand the topography of your data estate. Now we're providing better tools to manage it as it grows, and more points to surface that data to your business, to make use of it in the day-to-day.

Here are the tools that Unified Catalog provides to meet these data governance basics:

Expand table

Data governance principle	Catalog solution	Description
Data access - quickly provide right access and enforce right use to balance safety and innovation.	Unified Catalog access policies	Provides tools to manage self-service access requests that pairs data with compliance standards and right-use requirements.
	Critical data elements	Establish access policies on critical types of information to be applied across your data estate.
	Glossary terms	Attach access policies with your business vocabulary to promote right-use.
Data curation - organizing, annotating, and publishing your data so that it's safely accessible, reusable, and protected.	Governance domains	Organize data by business concepts, to make data more accessible and distribute ownership.
	Data products	Group related data assets so users can easily find the full data picture.
	Health actions	Take action to bring your data up to good governance standards.
Data discovery - users can find the data they need for day-to-day business and innovation.	Search & browse Unified Catalog	Search by governance domain, by data product, by keyword, or use the AI powered copilot to find what you need.
	Self-service access requests	Get access to all the data you need with a single request from inside Unified Catalog.
Data health - data quality standards are maintained across your estate, and there's an active data lifecycle keeping your data fresh and secure.	Health management	Ready-made reports provide the status of your data estate, and your data governance progress, at a glance.
	Critical data elements	Track important information to standardize and govern use.
	Data quality	Set quality rules from the top down to simplify distribution and provide snapshot progress tracking.
	Health controls	See how your data estate measures up to governance standards.
	OKRs	Map data health and governance goals to your business objectives.

Data governance principle	Catalog solution	Description
Data understanding - data has quality descriptors that help users understand what the data is and how it should be used.	Data products	Provide business context for a set of data assets.
	Glossary terms	Attach your day-to-day business vocabulary to your data assets.
	OKRs	Link data usage directly to your business objectives.

Get started

Ready to get started making your information clean, valuable, and accessible to your business users? We've created a guide to take you from zero to a fully integrated Unified Catalog in Get started with data governance.

Unified Catalog features

Now that you understand how Unified Catalog supports good data governance practice, and how the features promote data governance principals, explore the features in more detail:

- Governance domains: An organizational object that provides context for your data assets and makes it easier to scale data governance practices.
  - Unified Catalog access policies: Provide secure, self-service access to data products.
  - Critical data elements: Logically describe your key business data to govern them.
  - Glossary terms: Active values that provide context but also apply policies that determine how your data should be managed, governed, and made discoverable for use.
- Data products: A kit of data assets (tables, files, Microsoft Power BI reports, and more) that provides assets with a use case for ease of discovery and understanding.
- OKRs: Drive business value from your data and directly promote data governance objectives.
- Data estate health updates: Expanded features to provide new insights and encourage governance ownership.
  - Health controls: Track your governance progress with a health score.
  - Health actions: Follow these actions in your data estate to improve your governance score.
- Data quality: Set rules and get scores to get insight across your data estate.

Governance domains

Governance domains let you explore data through business concepts, like Marketing or Finance. This approach makes data more accessible to everyone.

Governance domains offer a new way to organize your data estate. You want and need to access your entire data estate, but a single, uncategorized list is overwhelming. A governance domain is a boundary that aligns your data estate to your organization. Think of it as a mini catalog inside Unified Catalog.

Inside your governance domain are the terms that define users' day-to-day work. There are objectives and key results (OKRs) to align your organization's goals with your data. There are data products that make data discovery more straightforward for your users.

The goal of the governance domain is to organize your Unified Catalog so that it not only drives healthy data governance but also ties your data directly to its business value and your objectives.

For more information, see the overview of governance domains.

Access policies

Unified Catalog access policies let you manage access to your data products and set up a system to provide access to users who request it. Promote innovation and flexibility in your data estate by creating self-service access opportunities, while upholding security and right-use standards. All in the Unified Catalog.

For more information, see how to create and manage Unified Catalog access policies.

Critical data elements

Critical data elements are a logical grouping of important pieces of information across your data estate. For example, a "Customer ID" critical data element can map "CustID" from one table and "CID" from another table into the same logical container. These groupings can make data easier to understand and promote standardization. You can attach data quality rules and access policies to these elements to further secure sensitive information across your data estate.

For more information, see how to create and manage critical data elements.

Glossary terms

If you use Microsoft Purview, you're familiar with glossary terms and how they provide critical business context to your data assets. We've taken them from static objects to active objects that help define how you should manage, govern, and make your data assets discoverable.

Policies within these terms allow data stewards to scale governance across your entire data estate. Terms applied to data products trickle down to the data assets and automatically secure those resources with their attached policies.

For more information, see the overview of glossary terms.

Data products

Data products improve catalog organization by grouping data assets together, such as tables, files, and reports, so users can easily discover them. You no longer need to request access to 15 different tables to build a data model. When one user researches and creates a viable data product, all other users benefit from that work. They can find and request access to the data in that product and have everything they need in one place.

For more information, see the overview of data products.

OKRs

Objectives and key results (OKRs) link data products directly to your objectives. They bridge the gap between your business and Unified Catalog. Use data to discover and track objectives in your business. Unified Catalog makes it easy to see those connections and track your goals.

For more information, see the overview of OKRs.

New data estate health features

Health management has a few new features for you to implement to enhance your data governance strategy and management.

Health controls

Track your journey to complete data governance by monitoring health controls. Health controls measure your current governance practices against standards and give your data estate a score.

Learn more about data health controls.

Health actions

Health actions are concrete steps you can take to improve data governance across your data estate. The actions are provided in a single list that can focus your data governance journey

and democratize ownership. Completing these actions improves data quality and discoverability across your data estate.

Learn more about viewing data health management actions.

Data quality

Data quality enables your organization to set rules through your governance domains, data products, and the data assets themselves. These rules trickle down across your environment so you can better evaluate your data across your data estate. You get data quality scores at the asset, data product, and governance domain levels, so you have full insight into your data estate.

Get an overview of data quality.

Related content

- Get started with the Unified Catalog
- How to create and manage governance domains
- How to create and manage data products
- How to create and manage glossary terms

Supported regions for Unified Catalog

> **Note**
>
> To access Microsoft Purview data governance solutions, make sure you upgrade to the enterprise version of the new Microsoft Purview experience when available in your region.

The following table outlines product availability or the deployment plan for the general availability of Unified Catalog across different global regions. The table also includes the roll out of the name change from Data Catalog to Unified Catalog in the Microsoft Purview portal.

Expand table

Geo-location	Region	Availability	Billing start date	Catalog name
United Arab Emirates	UAE North	Available	January 6, 2025	Unified Catalog
France	France Central	Available	January 6, 2025	Unified Catalog
India	Central India	Available	January 6, 2025	Unified Catalog
Japan	Japan East	Available	January 6, 2025	Unified Catalog
Brazil	Brazil South	Available	January 6, 2025	Unified Catalog
United States	West US2	Available	January 6, 2025	Unified Catalog
United States	South Central US	Available	January 6, 2025	Unified Catalog
Asia Pacific	Southeast Asia	Available	January 6, 2025	Unified Catalog
United States	West US	Available	January 6, 2025	Unified Catalog
Canada	Canada Central	Available	January 6, 2025	Unified Catalog
United States	East US2	Available	January 6, 2025	Unified Catalog
United Kingdom	UK South	Available	January 6, 2025	Unified Catalog
Australia	Australia East	Available	January 6, 2025	Unified Catalog
Europe	North Europe	Available	January 6, 2025	Unified Catalog
United States	East US	Available	January 6, 2025	Unified Catalog
Europe	West Europe	Available	January 6, 2025	Unified Catalog
South Africa	South Africa North	Available	January 6, 2025	Unified Catalog

Geo-location	Region	Availability	Billing start date	Catalog name
Germany	Germany West Central	Available	January 6, 2025	Unified Catalog
Korea	Korea Central	Available	January 6, 2025	Unified Catalog
Switzerland	Switzerland North	Available	January 6, 2025	Unified Catalog
Canada	Canada East	Available	January 6, 2025	Unified Catalog
Sweden	Sweden Central	Available	January 6, 2025	Unified Catalog
United States	West Central US	Available	November 20, 2025	Unified Catalog
United States	West US 3	TBD	TBD	Unified Catalog
Qatar	Qatar Central	TBD	TBD	Unified Catalog

As we continue to evolve the plan and deploy to new regions, we regularly update this page.

Next steps

- Get started with the Unified Catalog
- How to create and manage governance domains
- How to create and manage data products
- How to create and manage glossary terms

Last updated on 11/20/2025

Plan for Unified Catalog with best practices

07/16/2025

In this article, we review the data governance strategies you can implement to maintain healthy, valuable, discoverable data. For a list of technical steps to set up Microsoft Purview Unified Catalog, see the get started guide.

Know your data with business concepts

Business concepts in Unified Catalog are the tools used to unite your data with your day-to-day business practices. This not only makes it easier for your data consumers to understand the data they're using, but it allows you to democratize the data governance of those resources. Use your existing experts and data champions to build Unified Catalog into a rich resource.

When we refer to business concepts in Unified Catalog, we're talking about these five elements:

1. Governance domains
2. Data products
3. Critical data elements (preview)
4. Glossary terms
5. Objectives and key results (OKRs)

Jump to a section below for best practices and scenario-based guidance on:

- ✓ Governance domain structure models and development planning
- ✓ Planning for and creating data products
- ✓ Glossary term development planning and defining
- ✓ Creating and using custom metadata

Create governance domains

Governance domains are used to both distribute ownership and maintenance tasks, and make it easier for users to find the data that they need. Distributing information by governance domains allows your users to reach the right level of information they need, without needing to traverse the entirety of your data estate.

When creating governance domains, or reviewing your governance domain structure, here are some things to consider:

1. Governance domain structure model
2. Development planning

Governance domain structure models

- **Central domain** (good): Using a single domain can be efficient for small organizations, but might not scale well and is prone to bottlenecks during growth.
- **Department-based domains** (good): Departments don't make decisions consistently, and if departments regularly shift, you might need to shift your Unified Catalog structure.
- **Functional/line-of-governance domains** (better): Grants flexibility to teams, and aligns with the existing business model. This can be difficult to manage at scale, and might need many sub domains to empower data decision makers. It can also create data use silos, which is the antithesis of the Unified Catalog's governance approach.
- **Domain mix** (best): Having a combination of domains across subject areas/data domains, functional domains, regulatory domains, and project domains align your data to its experts. In Unified Catalog, your data experts are your most powerful resource; they know what policies need to be applied and what others need to know to make the most use of the data. This structure will also be the most durable to organizational updates, since it's based on how the data is used in the day-to-day, instead of on business structures.

Governance domain development planning

> **Tip**
>
> It's not recommended to align your governance domains with your platform domains. IT is typically aligned with a technology structure or service/application, and isn't aligned with how data is used by business teams. Platform domains in Data Map likely align to these technology teams, instead of your business teams. The goal of governance domains is to align business users with the information that's most useful to them. Focus on data use, instead of data structure, to develop your governance domains.

1. When you begin creating your governance domains, start with a few domains aligned to teams that already have strong data stewardship:
   a. Assign data stewards and data product owners to your governance domains, and have them begin development on a glossary and data products that align with their current practices.
   b. If needed, scan data into the data map in parallel to supplement your data products.
   c. Leave your governance domain in a *draft* state until a few data products have been developed and are ready for users.
2. Publish your governance domain, and assign Global Catalog Reader permissions to your first users to let them start exploring.
3. With the feedback from your first batch of users, iterate on existing data products, or expand to your next data products or governance domains.

4. Starting with a few governance domains that have mostly complete coverage with data products ensures data consumers that the Unified Catalog has what they need and that they can continue to come back.

> **Tip**
>
> Learn more about governance domains.

Create data products

Much of the data that is stored today has little to no known value and can take time and manual effort to evaluate and understand before it can be removed or improved. Focusing on the data with known value and use will enable more teams to build consistent value and show the benefits to having well understood and highly utilized data. This drives further adoption of data governance practices and makes the effort to clean up data estates easier as the value of each data asset becomes clearer.

Focus on data resources that already exist in your organization. Adding these as data products in Unified Catalog makes it easier for your users to discover them. It will also make access more scalable, and improve trustworthiness with lineage, data quality, and accountability. Some examples of existing data resources are:

1. Gold zone data lakes, highly curated SQL stores, curated data warehouses/data lakehouses that teams use to support their day-to-day practices.
2. Reports that are used to make decisions.
3. Data tables that are used in reporting environments.
4. Master and reference data.

Data product development planning

1. Planning data products should be part of your intake process when data sources are added to the Microsoft Purview Data Map. Data Product Owners should know which data stores are being registered and scanned, and which have data assets ready to be added to Unified Catalog.
2. Build your first data products from core data assets that have been scanned into the data map.
3. Publish your first data products when your users are ready to consume data with that domain.

> **Tip**

Learn more about **data products**.

Define glossary terms

When you're building terms, start with what you already know and continue to build value from your data to show where effort is the most impactful. Here are some tips you can follow when creating and managing glossary terms to create the most value.

1. Giving data to the most passionate users demonstrates the ability to continue growing value and provides prioritization for more governance.
2. Many business teams already have a glossary to help new employees to orient themselves to the business. Use these as some of your first term candidates to describe a governance domain and its data.
3. If you aren't sure if a term represents another concept (like an entity or business process) adding a term is a good place to start so the most basic metadata is collected. If needed, th term can be expired, and a new concept can be used to collect more metadata and drive the intended end-to-end experience.
4. Once glossary terms are added, linking these terms to data products improves the discoverability of data products and enhance the consumers knowledge of the data.
5. Periodically checking the data products that are mapped to a term to enable data stewards to better understand their use across the data estate.
6. Term definitions can always be improved and edited. Waiting to publish a term until it's fully aligned will delay teams use of the term and prevent new value creation or escalation of potential improvements.

Term development planning

1. Data stewards should learn the framework of the governance domain, and then begin to add known terms and start to develop new ones.
2. Term definitions should be developed and contain valuable information for consumers to understand their context and use.
3. The first set of terms and data products should be published together for consumers to start their data use cases and discovery of data in the catalog.
4. Building semantic knowledge never stops, so make a plan on how you can enable your team to continue to contribute terms throughout your governance lifecycle.

> **Tip**
>
> Learn more about **glossary terms**.

Create and apply custom metadata

Custom metadata plays a crucial role in empowering organizations to govern their data effectively. By enabling users to define attributes tailored to their specific needs, custom metadata fosters better organization, accessibility, and utilization of data assets. In an era where data is the backbone of decision-making, custom metadata allows organizations to streamline processes, ensure compliance, and unlock the full potential of their data resources.

Unified Catalog > **Catalog management** > **Custom metadata** is where you create and manage business concept attributes and data asset attributes. Currently you can manage your custom attributes for business concept types and data asset types.

> **Tip**
>
> Learn more about **custom metadata**.

Unlock business value by making data accessible

Now that your basic Unified Catalog structures are in place, it's time to start unlocking the value of your data by making it accessible to your users, and tying it directly to your business goals. Creating value from data comes from using that data, but using data means every person in the company needs to find the right data at the right time and in the right format to provide the needed insights or functionality. Data consumers are the key to making new business value from data.

Jump to a section below for best practices and scenario-based guidance on:

- ✓ Helping your users search governance domains and data products
- ✓ Building out objectives and key results (OKRs)
- ✓ Establishing compliant data access
- ✓ Building logical data models with critica data elements (preview)

Let users search and browse your governance domains and data products

You've taken the time to build out governance domains and data products, so give your data consumers access to use them and see how they do. Business users might be looking for strategic reports that are already available with the insights they need to make business decisions in a timely and well informed manner.

Here's how you can think about granting access to your users strategically:

- **Do not start by granting access to Unified Catalog to everyone at your company.** Enable the teams that need the data you have in your catalog first. If your data products aren't available in the format data scientists need or the data isn't in predefined reports for business users, they'll lose trust in your catalog. Enabling the right roles to use the catalog first to build the pathway to success.
- **Do start with the teams that need the data you have in your catalog first.** Who did you build your data products for? Which teams helped develop your glossary terms? Those are good initial candidates.
- **Do start with analysts and data experts that can tell you where gaps exist in the catalog.** They can help point to experts and business owners that can contribute to Unified Catalog. Over time the completeness of Unified Catalog will be great enough everyone in the company will be able to find most of their data needs.

> **Tip**
>
> Learn more about managing access to data in Unified Catalog.

Create OKRs

Demonstrate the business value of your data by building objectives and key results (OKRs) and tying them to the data products that help to drive or measure that value. Ensuring business leaders appreciate the value of their data and the importance of governance will drive prioritization and new synergies in how teams build, maintain and govern their data to create insights.

Building out an objective provides immediate recognition of the importance of the data to the users and the business it drives. This greatly enhances the understanding of the role certain data plays in business processes or in the ability to achieve their goals.

1. Consider OKRs for process improvements, quality issues, major strategic goals, and anything else that you would measure with data to demonstrate business value and change.
2. Make sure to create key result for each objective to show how the objective is being measured and evaluated, and create accountability to meet that goal.
3. There can complex objectives that have many key results required to be able to accomplish. Key results might progress independently of the other key results that the measurement can show the areas that need prioritization or help to get back on track.

> **Tip**

Learn more about **OKRs**.

Compliant data access

Providing access to data can introduce risk to your company. Following known standards and policies is essential for ensuring access is granted appropriately and that there's responsible use of data. Users in Unified Catalog can complete a form for data access at the time of discovery or data use. Keeping this form and process as a part of the catalog makes access, secure, quick, and consistent for a highly variable and technical data estate. Here are some ways you can successfully set up access in your catalog:

- Ensure the appropriate approvers are in place on data products and that they understand the processing needs of the data products.
- Some data products might have hundreds or thousands of access requests, so having a team in distributed time zones could be required to ensure timely access approval and provisioning.
- Prepare groups or back up approvers in case there are vacations or unplanned time off.
- Governance domain owners should check on the access requests summary periodically to validate expectations and see if changes to controls monitoring the access request process are driving the desired response times.

Build logical data models with critical data elements (preview)

Improving deep technical understanding and expectations of data entities and elements will include new controls to assert if the data is meeting those expectations. Creating data dictionaries and logical models of data provides the structure and deep business expectations of the data to ensure that the data is fit for its purpose. By incorporating this knowledge into the Unified Catalog teams will immediately gain an understanding of how data is structured and why, and how what is available in the physical data estate might differ.

- Focus on the data elements that are most important for your domain. Critical data elements show the deep expertise and importance that data has on your business.
- **Don't** focus on the completeness of the elements across an entire domain. Not every column needs this level of control and many data elements might be self-explanatory for users.
- Evaluating critical data elements across different teams ensures that business teams have a common understanding of their data and how what one team creates is impacting many other areas of the business.
- Aligning access policies with critical data elements ensures proper access controls are in place for critical data across your whole data estate.

- Building data quality rules for critical data elements ensures that data meets expectations no matter where or how it's being used.

> **Tip**
>
> Learn more about critical data elements (preview).

Enhance data maturity

Improve your data estate and governance to fill gaps and remove bottlenecks to value creation:

- Monitor your health actions to improve governance incrementally across your entire Unified Catalog.
- Optimize for new uses of the data and eliminate data issues by improving data quality.
- Create best-in-class data products for single sources of truth with master data management.
- Evaluate your data health and prioritize for the greatest value impact.

Investing deeply in the core data that runs your company ensures this data is usable across the entire business consistently eliminating data issues and providing a stable base for insights creation. Having evidence of data issues helps to ensure actionability of data governance, and drives improvements that will immediately unlock new value without investing in data areas that have low value aren't fully understood yet. Continuously improving data maturity helps teams share learnings with each other and show the proof of the improvement as changes take place.

Jump to a section below for best practices and scenario-based guidance on:

- ✓ Improving data products with governance-focused actions
- ✓ Using data quality to improve the trustworthiness of your data
- ✓ Creating source of truth data products with master data management
- ✓ Using data health controls to measure governance maturity
- ✓ Building domain-specific standards

> **Tip**
>
> Start exploring how to manage data health and data quality.

Improve data products with governance-focused actions

Building trust in data requires continuous improvement and support. While consumers take time to find and apply data, bringing attention to issues or support needs there are easy actions that can be taken ahead of time based on best practices. Health actions in health management provide a complete list of these useful actions for Unified Catalog, to help you focus on what you can do next to improve your governance. Here are some best practices for using health actions to get the most value:

- Check the actions of your data products while they're still in a draft state. This ensures that when you do publish it has the basics covered and provide comfort to consumers that this data was published with care.
- Not all actions have to be taken at the same time. Some actions could take time to resolve as you learn more about the data or work with stewards to create more clarity. Keep checking on actions to see where new improvements are ready to be made.
- If actions begin to seem overwhelming, unnecessary, or distracting, consider making changes to your health controls. Optimizing the number of actions any person takes ensures that the right level of governance is being applied to data.

Improve the trustworthiness of your data with data quality

Too often data quality is a one-off project to fix a particular problem in the data. These improvements but don't last. Good data quality requires continuous evaluation and improvement to ensure problems won't return or new problems aren't created.

- Once a baseline of data quality expectations has been defined, building a plan to remediate issues in a timely manner is essential to keeping the business functioning with data fit for use.
- Scheduling your data quality scans to run regularly will help to ensure consumers that data is continuously being improved and is highly supported.
- Setting alerts on the critical rules and score changes enables data providers to correct issues before a consumer finds or experiences an issue. Alerting can also be used to share transparently with consumers issues before they find it in an experience or by making a decision based on poor quality data.

Create source of truth data products with master data management

Some data is so critical to nearly every process and the entire business that it deserves exceptional levels of management and governance. These data entities are usually cross-cutting entities like customer lists or employee profiles and they can require deep business expertise and experience in many business processes. Some data is highly usable but low scale, and still benefits from the deeper level of control and management; like reference data

attributes of country or region, currency, or industry segments. Each of these data types would benefit from master data management solutions to build a source of truth that is fit for use across your entire business.

- Practicing master data management with data quality is crucial to ensure that this vital data is clean and consistent.
- This level of data management is high effort, so choose valuable data elements or high-risk data elements to ensure your effort produces high value.
- Creating a critical data element and a data product for master data. These partner objects help to elevate your master data in the Unified Catalog and increase its use and understanding.
- Build new health controls for master data to continuously evaluate its use at scale and prevent new unmastered data from gaining use and causing confusion in a quickly evolving data estate.

> **Tip**
>
> Learn more about **master data management**.

Measure governance maturity with data health controls

Evaluating the maturity of data governance at scale across the entire business is required to ensure governance is effective and creating business value. By applying the built-in measurement of controls, health management enables the central data office or an individual governance domain to see where there's more that can be done. Collecting this evidence at scale quickly elevates the most critical data issues impacting the business and where one issue maybe impacting many areas of the business. This evidence helps to resolve prioritization issues with making data management changes and quickly demonstrates the value of having the right level of governance in place.

1. Establish a rhythm of business to review health management practices:
    a. Have a monthly review with governance domain leaders and the central data office to discuss priorities and needs for new governance or technical solutions.
    b. Empower teams to dig deep into their health management reports to make sure they can make the best decisions to create the value they need in their business.
    c. Bring health management to all levels of the business from the SLT to the individual steward to ensure that governance is right leveled and consistently actionable.
    d. Where data has larger issues requiring cross-business collaboration or deeper governance, consider creating a new governance domain and defining ownership for driving the governance of that data.

2. Don't expect all governance domains to have the same level of maturity or be focused on the same aspects of governance:
    a. Enabling governance at the right level empowers business owners to make the most valuable decisions about what to do with their data.
    b. Not all parts of the business have the same needs of their data and forcing deeper levels of governance might not help to create business value when the focus is elsewhere.
    c. Some data is less valuable or emergent in the data estate and the value isn't yet fully known. Enabling teams to move fast and adapt to their needs they can mature their governance with the value of the data.
3. Consistently evaluate health management to look for large changes that can indicate large issues or new learnings that need attention.
4. Share your health management scores. Sharing can bring teams together to learn what works for them or how they're finding new controls to build new value within a domain. Seeing what 'good' health looks like can motivate other teams to improve and ensure they're also delivering valuable data to their consumers.

Build domain-specific standards

Ensuring data governance is right sized for the level of value and control required is best handled by the business owners of the data. These business teams already have dependencies on the data and are in the best position to define their expectations and needs to make sure data is valuable.

- Empower governance domains to create new controls for their data regardless of where the data is used.
- Don't expect all governance domains to need the same level of controls or to adopt all controls. Data that is confined for use to a single part of the business by design might not benefit from a high level of control. Creating more control over data that doesn't have the appropriate value might prevent teams from collecting or keeping data that isn't fully utilized.
- Use the right level of control to help prioritize where low value data can be removed from the governance domain to eliminate risk and increase the value of the data estate.

Ready to get started?

To build a data governance practice in your organization using Unified Catalog and Data Map, see Get started with data governance or check out the sample setup walkthrough for more of a scenario-based example of how to perform key functions in Unified Catalog.

Manage Unified Catalog with the data governance admin role

In the data governance admin experience, you can manage the data governance settings that apply to Microsoft Purview Unified Catalog in one place. Users who hold the Data Governance Administrator role can set Unified Catalog roles and permissions, configure self-service analytics (preview), and manage other settings described in this article.

Prerequisites

To use the data governance admin experience, you need to be in the tenant-level data governance admin role.

Where to find Unified Catalog solution settings

To use the Data Governance Administrator experience in the Microsoft Purview portal:

1. In the Microsoft Purview portal, select **Settings**.
2. Under **Solution settings**, select **Unified Catalog**.
3. Select a setting name from the navigation. Each setting is explained in more detail in the following sections.

Roles and permissions

The roles and permissions item moved from the Unified Catalog solution to the solution settings, but the steps for applying those permissions are still the same. For more information about setting up roles and permissions for your users, see the data governance permissions documentation.

Solution integrations for self-serve analytics and insights

The **Solution integrations (preview)** page is where you set up self-service analytics and insights. These insights are data, tools, and platforms that enable business users to access, analyze, and generate insights from data independently by bringing their own tools and compute. For details, see Self-serve analytics and insights.

Usage settings (preview)

Data Governance Administrators can understand and adjust the usage settings available for Unified Catalog in this experience. The Data Health Management SKU can be changed for data quality jobs, and data quality monitoring can be used to monitor and understand the effect on processing. In the rare event of capacity not being available for a higher SKU, a basic SKU is defaulted to. Learn more about data governance billing.

Usage monitoring (preview)

> **Note**
>
> Usage monitoring is rolling out across regions. If you don't see it yet in your tenant, it will appear in the coming weeks.

The **Usage monitoring** experience is the Data Governance Administrator's view of governed assets and data governance processing unit (DGPU) consumption across different governance domains. This view lets the admin take the overall bill and use the consumption breakdown by governance domains to create a chargeback model in their organization across governed assets and DGPU meters.

The data reported in the charts reflects daily totals. The counts are by each governance domain and aren't aggregated to the parent.

Consumption data is split across governance domains as follows:

- **Governed asset**: Asset used in more than one governance domain split evenly across each domain that uses the asset. This count can show decimals in asset count.
- **Data quality**: Each data quality job maps to a governance domain, not split.
- **Data health management**: Split evenly across all governance domains.

You can use any other advanced split logic through self-serve analytics, when all DGPU data is available in self-serve analytics.

Other points to know:

- Cumulative governed assets are the cumulative governed assets for the time period you choose that is used to calculate your bill, not a unique count. For more information, see how governed assets are billed.
- The **Usage monitoring** settings don't show the actual bill; you can find your bill in the Microsoft Azure portal.
- If you rename a governance domain, the previous name shows on the chart for the date on which it changed. The new name shows on the date after the name change.

- When you delete a governance domain, only the GUID shows, not the domain name and owner name.

> **Tip**
>
> Find detailed information about **data governance billing** and a **billing FAQ**.

Data observability (preview)

Data observability (preview) provides a single view of your data estate to help you understand lineage and data health, and identify improvements in metadata, ownership, data quality, or data architecture. In this area of Unified Catalog settings, admins handle initial setup steps to generate data visualizations. For more information, see Set up data observability.

Virtual network

Data Governance Administrators can provision a virtual network compute location in supported Microsoft Azure regions. Get details at Manage virtual network provisioning

Related content

- How to set up data governance roles and permissions for the Unified Catalog
- How to set up Unified Catalog self service analytics and reporting
- Data governance billing

Last updated on 03/29/2026

Governance domains in Unified Catalog

10/03/2025

A governance domain in Microsoft Purview Unified Catalog provides context for your data assets and makes it easier to scale data governance practices.

Data ownership is a central aspect of data governance. Controlling access and data use are core tenets of any successful governance program. Often, IT teams store and maintain data assets even though business teams own and use the data. That creates a gap between how data should be discovered and maintained and the teams that actually use it. Unified Catalog helps bridge that gap, but business owners still need clear context about their data assets and repeatable ways to scale governance across a growing data estate.

To address these needs, Unified Catalog introduces governance domains.

Overview and elements

At its simplest, a governance domain is a boundary that enables the common governance, ownership, and discovery of data products and business concepts like glossary terms, objectives and key results (OKRs), or critical data. The goal is to empower a governance domain owner to manage their data products and concepts, and establish rules for their access, use, and distribution. With this goal in mind, you could establish many kinds of governance domains:

- Fundamental business areas, such as human resources, sales, finance, and supply chain.
- Overarching subject areas, such as product and parties.
- Boundaries based on organizational functions, such as customer experience, cloud supply chain, and business intelligence.

In Microsoft Purview, governance domains are flexible so your organization can build the boundaries that make the most sense for your data. But once you've created them, how do they provide context or assist with data governance?

The parts of a governance domain: business concepts

A governance domain has a name and description, which helps users quickly understand what parts of your organization the governance domain represents. The domain also has owners who govern and maintain the domain and its assets.

Governance domains also contain the following business concepts:

- Data products: Packaged sets of data assets (tables, files, Power BI reports) grouped for discovery and reuse.
- Glossary terms: Business terms that provide context and can carry access and data-handling policies that determine how data should be managed and discovered.
- Objectives and key results (OKRs): Metrics that describe the value of your data with measurable values and goals.
- Critical data elements (preview): Logical groupings of important columns (for example, mapping "CustID" and "CID" to a single "Customer ID" concept) that require elevated governance.

Types of governance domains

These are the supported out-of-the-box governance domain types:

- Functional unit: Organizations or business units, such as sales, marketing, or finance.
- Line of business: Products or services being sold, such as Xbox, Office, or Azure, and different markets or subsidiaries.
- Data domain: Key organization-wide entities, such as customers or employees.
- Regulatory: Compliance-related domains such as GDPR, SOX, or HIPAA.
- Project: Collaborative programs across the organization.

How governance domains provide business context

Governance domains provide context for your data products, which are combinations of files, tables, reports, and other assets, by defining an overarching category for them. For example, someone who needs supply delivery schedules can limit their search to the Supply Chain domain instead of scanning the entire catalog.

Each governance domain can define OKRs to describe the business value of its data products and set measurable goals. Domain owners can prioritize data products based on those goals so users understand how data supports organizational strategy.

This structure provides familiar landmarks for business users and lets domain owners tailor and scale governance as the data estate grows.

How governance domains scale data governance

In a flat catalog, governing each asset individually is costly and unsustainable. Governance domains provide boundaries where policies can be applied broadly to data products.

The key to propagating governance policies is the use of glossary terms. Glossary terms provide business-context labels that improve discovery. Glossary terms are defined within a governance domain and can be applied to any data product in that governance domain

Data owners and data stewards don't have to traverse an entire data estate to maintain governance. They can apply terms across a governance domain they understand and know that when these terms are attached to a data product, the right policies will automatically fall into place. As an example: In the Human Resources governance domain, the "Feedback Results" term denotes returned company-wide feedback queries that are considered sensitive. A designated review team evaluates access requests. A data access policy on the Feedback Results term is applied automatically to any data product labeled with that term when users request access.

Governance domain owners and data stewards define policies once and rely on terms to propagate those practices across the data estate, ensuring consistent enforcement even as new assets are added.

Governance domains use existing business context to make data more discoverable and consistently governed as it grows.

Related content

- How to create and manage governance domains.

Create and manage governance domains

Ownership of data is an important aspect of data governance. However, today IT infrastructure maintains and supports data without always having the full context for how data can and should be used in an organization. The use of governance domains brings users and data together with a data-use driven structure. This article covers how to create and manage governance domains in Microsoft Purview Unified Catalog.

Prerequisites

- To create a governance domain, you must have the governance domain creator role.
- To assign data stewards or data product owners, you must have governance domain owner permissions.
- To curate data, such as adding glossary terms or objectives and key results (OKRs) to a governance domain, you must have Data Steward permissions.
- To add data products to a governance domain, you must have Data Product Owner permissions.

Access governance domains

To see governance domains, you need the Global Catalog Reader role or the Local Catalog Reader role, which is a more restrictive role. For details about each role, see Catalog level permissions and Governance domain level permissions.

1. In the Microsoft Purview portal, open **Unified Catalog**.
2. Under **Catalog management**, select **Governance domains**. You see a list of all the governance domains in your organization. You can filter the governance domains by keyword.
3. Select a governance domain to see its details, roles, aggregate data quality score, health actions, associated data products, glossary terms, OKRs, and critical data.

Details

The **Details** tab shows key information about the domain and displays all associated business concepts.

Custom attributes

Under **Custom attributes**, you can see any custom attribute groups and attributes within groups, and the values set for each attribute. Only attributes with values are shown by default, but you can toggle **Show attributes without a value** on to view all attributes.

Roles

The **Roles** tab displays the users in your organization (grouped by role) who are associated to the domain. For details, see edit roles for governance domains.

Data estate mappings

Use the **Data estate mappings** tab to map your domain to a specific Data Map collection. For more information, see setting up data estate mappings.

Data observability (preview)

The **Observability** (preview) tab provides a detailed view of data health and lineage. This view requires a brief setup process. For more information, see the observability view for governance domains.

Monitoring

Select **Monitoring** on the **Details** page of a governance domain to view the status of bulk import jobs. The **Monitoring** page displays the status of bulk imports of glossary terms, data products, and critical data elements. The page will also list any upload errors and how to fix them.

Edit roles

To edit the users, groups, or service principals assigned to a role:

1. Go to the governance domain where you want to edit the roles.
2. On the **Roles** tab, select the + button next to the role name where you want to add the user.
3. Search for users, groups, or service principals to add to the role and select them.
4. To remove users from the list, select the **X** next to their name.
5. When you're done, select **Save**.

For more information, see roles assigned at the governance domain level.

View your business concepts

Your governance domain has many business concepts associated with it: data products, glossary terms, OKRs, and critical data.

You can find and view each of these concepts in your governance domain in the **Business concepts** area on the governance domain's **Details** tab. Select the card for each concept to view its details.

To add any of these concepts to your governance domain, see:

- Add data products.
- Add glossary terms.
- Add OKRs.
- Add critical data.

Filter your view by custom attribute

You can use business concept attributes as a filter when exploring business concepts in a governance domain.

1. On a governance domain details page, select **View all** on a business concept card.
2. Select **Add filter**.
3. At **Filter**, select an attribute name from the dropdown list.
4. At **Operator**, select a condition, such as **Equals** or **Starts with**, which varies based on the kinds of values allowed by the attribute.
5. At **Value**, enter your value.
6. Select **Apply**.

Create governance domains

> **Note**
>
> To create a governance domain, you need the governance domain creator role.

1. In Unified Catalog, select **Catalog management**, then select **Governance domains**.
2. Select **New Governance domain**.
3. On the **Basic details** page, enter a unique name for your governance domain and a description.

4. Select the governance domain **Type**: Functional unit, line of business, data domain, regulatory, or project. **Type** is an extra descriptor to indicate your governance domain's purpose. There aren't any behavioral differences between types.
5. Select an existing domain as a parent domain (optional). Parent/child relationships appear in a tree-like hierarchy.

   > **Note**
   >
   > Governance domain hierarchies in Unified Catalog can support as many as 200 domains, including a maximum of five levels of depth.

6. Select **Next**.
7. On the **Custom attributes** page, you can set values for all custom attributes defined by your admin in the **Custom attributes** section of the catalog. If the attribute is marked as required, you need to fill out all values on this page before you can complete the process.
8. Select **Create**.

You see your governance domain in the list of domains. Other users can't see this governance domain until you set its status to **Published**. You must also publish your domain before you can publish business concepts, like glossary terms or data products, within it.

> **Tip**
>
> Find guidance on what governance domains to create in your organization.

Manage governance domains

> **Note**
>
> To edit a governance domain, you need the governance domain owner role.

1. On the governance domain's details page, select **Edit**.
2. You can edit the domain's name, description, owner, and status.
3. When you're done, select **Save**.

To add data products to a governance domain, see Create and manage data products.

> **Note**

If a governance domain becomes uneditable because the owner leaves the organization, Data Governance admins can add new governance domain owners in the **Roles** tab.

Configure data estate mappings

You can map your governance domains to collections in your Data Map to direct your data stewards and data product owners to build business concepts from data assets belonging to specific collections. This mapping is meant to be recommended guidance, as these Data Governance roles require Data Map permissions to access the data assets.

To map your governance domain to a specific Data Map collection:

1. Go to the **Data estate mappings** tab in a governance domain.
2. Use the **Select a mapping** dropdown menu to select the collection you want to associate your governance domain with.

   > ⓘ **Note**
   >
   > You can only see the collections that you have access to in Data Map.

3. Select **Save**.

Delete governance domains

To delete a governance domain, unpublish it and delete all business concepts within it, including any subdomains. Then select the **Delete** button to delete the governance domain.

Manage domain policies

To manage domain policies, you need data owner or data steward permissions.

1. On your governance domain's page, select **Configure access**.
2. From the policy configuration window, you can create and manage your access policies. For more information, see how to manage access policies.

Access classic experiences in the governance domain

You can access classic glossaries, business assets, or managed attributes by going to **Catalog management**, then selecting **Classic types**.

Select a tab to access your classic experiences: **Glossaries**, **Business assets**, **Asset types**, or **Managed attributes**.

Next steps

- Create and manage data products

Last updated on 04/21/2026

Data products in Unified Catalog

10/06/2025

A data product is a group of data assets, such as tables, files, and Power BI reports, with a defined use case that you can share with other users.

In Microsoft Purview, data governance isn't only a way to make sure your data is secure and compliant, but it's also a tool to accelerate your data's business value. Cataloging data in your estate makes it possible to better manage data for right use, but it also provides a complete picture of your data landscape. Now that there's a list of every available data asset, users no longer have to rely on networking or team knowledge to find what they need; they can search the catalog themselves. But giving every user a raw list of all available data is overwhelming and not inherently useful. Even with good descriptions, tagging, and glossary terms, it can be hard to know what you're looking for. And for a complete data visualization you probably need several data assets, and not just one. As Microsoft Purview Unified Catalog grows, context needs to grow alongside it to make it easier for your users to find and request access to the data they need.

To provide scalable data context and access management, Microsoft Purview is introducing the data product.

What's a data product?

A data product is a business concept with a name, description, owners, and most importantly a list of associated data assets. The data product provides context for these assets, grouping them under a use case for data consumers. A governance domain can house many data products but a data product is managed by a single governance domain and can be discovered across many domains.

A successful data product makes it easy for data consumers to recognize valuable data using their day-to-day language, and at the same time streamlines ownership responsibilities for those data assets. Let's explore what that looks like.

Scalable data context

As an example, a data scientist creates a set of data assets for their data model to consume and they want others to be able to use the same dataset.

Without data products, the data scientist can use Unified Catalog to add a glossary term to all the relevant data assets. A user might not know which glossary term to search, so it might be best to add a description to each data asset to make it more relevant in searches for similar

information. But both additions don't guarantee that other users see all the associated data assets. They might group in other assets that aren't as relevant, or miss a critical data piece, and spend time repeating research the original data scientist already performed.

With a data product, a data scientist can create a data product that lists all the assets used to create their data model. The description provides a full use case, with examples or suggestions on how to use the data. The data scientist is now a data product owner and they've improved their data consumer's search experience by helping them get everything they need in this one data product.

Scalable data governance

Data products streamline governance for data assets. Using the same example of a data scientist who creates a set of data assets:

Without data products, if a user wants access to the data assets for the data set, they must request access to each data asset individually. A data owner might know that these assets are used for machine learning models, but if they make any changes to policies around their security and use cases, they must update each asset individually.

With data products, a user finds the data product and requests access to the data product. After approval, they get access to all the associated data assets. If a data owner puts more approval or data use policies in place around datasets for machine learning, they only need to apply the new policies to the data product. The policies automatically trickle down to the assets.

Data products are also associated with business health controls and OKRs. These controls allow data owners to assess data health and prioritize assets that need attention. They also help assess which data assets provide business value. This support not only helps progress towards complete data governance in your estate but also encourages developing business value from your data. Assets are no longer abstract but tied to real use cases and business objectives that your team can focus on.

Data access policies

Data security and access is the core tenant of successful data governance. But to implement data governance and successfully drive data use (and therefore value), the data access process needs to be secure, convenient, and customizable to all scenarios across your data estate. Some data should be widely useable and accessible, and some needs to be under rigorous approval and monitoring to ensure right use.

Each data product has an access policy that determines how users request access, the terms of use for the data, and who should approve access to the data. Each of these access policies is customizable for appropriate use and will evolve to cover more use cases in the future. All users need to do is select **Request Access** inside a data product. They automatically go through the process to agree to terms of use and get approval from correct parties.

For more information about access for data products, see the article on managing Unified Catalog access policies in Microsoft Purview.

Next steps

- Create and manage data products
- Set up access to data products
- Request access to a data product

Create and manage data products

A data product in Microsoft Purview Unified Catalog is a set of information with a defined use case that you share with other users. For example, a data product can be a sales report, an ML model, or a data model and its associated tables. A data product is a logical grouping of related physical assets that you create for a specific purpose.

It can be hard to find the value of data presented by itself, but data already associated with a purpose is easier to find and use. Data products in Unified Catalog provide practical context for your users and for intelligence systems, to help users identify what data is useful to them.

To create data products for your organization and ensure those products have enough definition to be useful to your user base, follow these instructions.

Prerequisites

- To view data products, users need catalog reader permissions.
- To manage or edit data products, users need data product owner permissions.

View data products

1. In the Microsoft Purview portal, open **Unified Catalog**.
2. Under **Catalog management**, select **Data product**. You see a list of all the data products you have access to based on your permissions. You can scroll, sort, filter, and search through these data products to find the ones you're looking for.
3. Select a **Data product name** from the list to view more details on the data project's details page.

Filter your view by custom attribute

You can use business concept attributes as a filter when exploring data products.

1. Under **Catalog management**, select **Data products**.
2. Select **Add filter**.
3. At **Filter**, select an attribute name from the dropdown list.
4. At **Operator**, select a condition, such as **Equals** or **Starts with**, which varies based on the kinds of values allowed by the attribute.
5. At **Value**, enter your value.

6. Select **Apply**.

Details

The **Details** tab on the data product detail page provides a description and use cases. Here you find the governance domain, update frequency, status, owner, subscribers, terms of use, aggregate data quality score, health actions, and documentation.

1. You see a list of all the data products you have access to based on your permissions.
2. You can scroll, sort, filter, and search through these data products to find the ones you're looking for.
3. For more details about a specific data product, select that data product.
4. On the data product's detail page, you can view the governance domain, update frequency, status, owner, subscribers, terms of use, aggregate data quality score, health actions, and documentation.

Custom attributes

Under **Custom attributes**, you can see any custom attribute groups and attributes within groups, and the values set for each attribute. Only attributes with values are shown by default, but you can toggle **Show attributes without a value** on to view all attributes.

Data observability

The **Data observability** (preview) tab provides a more detailed view into data health and lineage. This view requires a brief setup process. Get details about the observability view for governance domains.

Create data products

> **Note**
>
> To create, edit, and manage data products, users need the Data Product Owner role. You can only create data products in governance domains where you have data product owner access.

There are two ways you can create data products:

1. Individually, by creating one term at a time.

2. In bulk, by importing a formatted CSV file (preview).

Create a single data product

1. In Unified Catalog, select **Catalog management**, select **Data products**, and then select **New data product**.

2. On **Basic details**, enter a **Name** and **Description**. If you use a name that already exists, you'll see a warning during the creation process. While duplicating names isn't recommended, you won't be blocked from using a duplicate name. The description should be a business narrative about the data, where it came from, and why you captured it. The goal is for a new user to understand the basics of when, what, why, and how the data came into existence. Provide clarity about the meaning of the specific data to the business. The description is limited to 10,000 characters.

3. Select a **Type** from the dropdown menu. Get details about data product types.

4. Make an optional selection from the **Audience** dropdown list.

5. Enter the **Owner** and any other owners for the data product.

6. Select **Next**.

7. On **Business details**, select the governance domain your data product should be associated with.

8. Enter a **Use case**. The use case describes what the data is used for today and how a user can effectively apply it to their own scenario. Add information about what filters or dimensions are available, which asset contains the view that's easiest to parse, or any other details that can accelerate the usage of the data. If the data is only appropriate for specific purposes, include these details so users don't request access for data that they won't be approved for or won't help them to achieve their goals.

9. If you decide to mark the data product as **Endorsed**, select the **Mark as Endorsed** box. Learn more about endorsing a data product.

10. Select **Next**.

11. On the **Custom attributes** page, you can set values for all custom attributes defined by your admin in the **Custom attributes** section of the catalog. If the attribute is marked as required, you need to fill out all values on this page before you can complete the process.

12. Select **Create**.

You arrive at the new data product's details page. This is where you can make other edits to the data product, which is currently in a **draft** state.

> **Important**
>
> Other users can't see the data product yet. To make it visible to other users, you need to add data assets, create an access policy, and then publish the data product.

Bulk import data products (preview)

You can create many data products at once and import them into a governance domain by uploading a formatted CSV file. During the bulk import process, you download a sample CSV file containing a header row. You fill out the file according to the instructions in this article, and then reupload it. The data products appear in the governance domain in a draft state.

You can't use the bulk import process to edit or update data products.

> **Note**
>
> - You must download the sample CSV (.csv) file and reupload it according to the following instructions. The file must have at least one row of data, not including the header row.
> - The CSV file can only include 1,000 rows. More rows cause an upload error. If you need to import more data products, fill out and upload a separate CSV file.
> - Special characters are allowed in all fields.
> - Each data product name must be unique. Duplicate names aren't supported for import.

To import data products in bulk to a governance domain, follow these steps:

1. In Unified Catalog, select **Catalog management**, then select **Governance domains**.
2. Select the name of the governance domain where you want to add data products.
3. On the **Details** tab of the governance domain's details page, find the **Data products** card and select **View all**.
4. On the **Data products** page, select **Import**.

5. On the **Import** flyout pane, select **Download a sample CSV file**.

6. Open the downloaded file. Fill out the worksheet as described in this article to avoid upload errors.

    - **Column A - name** (required): Enter the data product name. Every name must be unique; you can't have multiple rows with the same name. Maximum length is 256 characters.

    - **Column B - owners** (required): Enter a valid email address that belongs to the same tenant that you're importing from. You can add multiple email addresses separated by a semicolon. For example: `owner1@contoso.com;owner2@econtoso.com`. Groups without email aren't supported.

    - **Column C - type** (required): Enter a data product type; see list of types.

    - **Column D - audience**: Enter one of these allowed values to describe the intended audience for the data product:
        - BI engineer
        - Business analyst
        - Business user
        - Data analyst
        - Data engineer
        - Data scientist
        - Executive
        - Software engineer

    - **Column E - endorsed**: Enter **True** or **False**. Endorsed is a signal of confidence to data consumers that the data product meets the organization's quality and governance expectations. It shows as a label on the data product for data consumers.

    - **Column F - description** (required): Enter a description for the data product. Maximum length is 500,000 characters.

    - **Column G - business use** (required): This entry shows in the **Use case** area of the data product. Enter a description of what the data is used for today and how a user can effectively apply it to their own scenario.

    - **Column H and onward**: These columns are for custom attributes, which differ by governance domain based on how they're defined and scoped. Note these formatting requirements for specific entry types:

    - **Multiselect options**: Enter the objects inside square brackets and in quotes. For example, `["a"]`, or `["a", "b"]`.
    - **Single select options**: Enter the raw value. For example, `a1`, `b2`.
    - **Booleans**: Enter `true` or `false`. You can use upper or lowercase.
    - **String or rich text**: Use regular text; don't add quotes unless you want quotes reflected in the attribute.
    - **Numeric - short, int, long, double, float, bit, byte**: Don't add quotes. Just enter the number.

7. After formatting your CSV file, save it locally. Return to the **Import** flyout pane, select **Browse**, and then find and select your CSV file to upload it.

8. On the **Import** flyout pane, select **OK** to start the import process.

If there's a problem with the CSV file, an error message appears on the **Import** pane explaining the error so you can fix the CSV file and reupload it.

If there are no problems with the file, the import process begins. Depending on the size of the CSV file, it might take some time to complete the import process. If the file is small, the import happens immediately and you see a success message. Larger files take longer for the import process to complete.

Check the status by returning to the governance domain details page. On the **Details** tab, select **Monitoring** to view the status of your bulk import jobs, including any errors, on the **Monitoring** page.

Data product types

When you create a data product, you can optionally identify it as one of the following types. These types help identify the kinds of users that might want to use the data product, and you can use them to filter data products in a search.

- **Analytics model**: Data that's transformed from raw data to help build analytics from, and might be distinct fact and dimension tables.

- **Business system/Application**: A common set of consumable tables from a specific source, or available in a certain application that you consume together.

- **Dashboards/Reports**: Data visualizations that you use to make decisions and view common metrics. They might or might not include the data used to power the visualization.

- **Dataset**: An individual asset or group of assets used as a common data product type to make data discoverable for specific use cases.
- **Master and reference data**: Highly reusable data that you should consistently reference and highly control due to its common use and greater importance to the health of the data estate.
- **ML training data**: Data approved for use as training a new model. It's typically not production data and is desensitized to prevent potential risks.
- **ML testing data**: Data built to specifically test ML models for edge cases, fairness, and potential harms.
- **Model types**: Typically used for AI models and analytical or semantic models used with AI.
- **Operational**: A data product that you use for data quality, or to operate a portion of the data estate but not planned for high consumption or governance. It could be an investigatory data product of transitory data assets or data used to run a process, but isn't modeled or consolidated for analytical purposes.
- **Semantic model**: A group of related data tables and their relationships and meanings for interpretation and reuse.
- **Transactional data**: Data from an originating source that might have limited use cases, but is required for engineering scenarios or system-to-system integrations. It's not typically used for most business data consumption, but it could be important for governance and control scenarios.

Managing data products

Data product owners can view and modify the properties of a data product by following these steps:

1. In Unified Catalog, select **Catalog management**, and then select **Data products**.
2. Select the data product to open its details page.
3. Select **Edit**.

From the data product page, you can:

- Edit details about the data product.
- Update frequency.
- Change the data product's status.

- Manage data product access policies.
- Add or remove associated data assets.
- Remove data assets that are deleted from Data Map (preview).
- Manage associated terms or OKRs.
- Manage terms of use.
- Manage associated documentation.
- Endorse a data product.
- Edit values for custom attributes.

Edit data product

To edit a data product, you need data product owner permissions. You can only edit a data product when it's in **Draft** status.

1. On your data product page, select the **Edit** button.
2. On the **Basic details** page, update the name, description, type, and owners. If you change the name to one that already exists, you'll see a warning about the duplication but you can proceed to use the duplicate name.
3. Select **Save**.
4. On the **Business details** page, update use cases and endorsement status.
5. Select **Save**.
6. On the **Custom attributes** page, you can set values for all custom attributes defined by your admin in the **Custom attributes** section of the catalog. If the attribute is marked as required, you need to fill out all values on this page before you can complete the process.
7. To save your changes, select **Save**.

Update contacts

When you add owners to data products during the creation or editing process, the owners automatically appear in the **Contacts** section. You can customize the label by selecting the pencil icon in each contact card.

Update frequency

Use the update frequency to show how regularly you manage a data product. This indicator isn't currently automated.

1. To edit the update frequency, select the **Update frequency** attribute on the right side of the page.

2. Select the update frequency you expect for your data product.
3. Select **Done**.

Publish, draft, and expire

To publish a drafted data product, select **Publish** on the data product's details page. Before you can publish, you need to add data assets to your data product and set up a data access policy so users can request access to your data product.

> ⓘ **Note**
>
> Ensure your governance domain is published before you publish your data products.

After completing the requirements, you can publish your data product, which makes the data product available for users in Unified Catalog.

You can also set a previously published data product to a draft by **unpublishing** it:

1. Select the **Unpublish** button in the data product.
2. Select **Set to draft** which allows only stewards, governance domain owners, and data product owners to view and manage it.

You can also set your data product to **Expired**:

1. Select the **Unpublish** button in the data product.
2. Select **Set to expired** which makes it visible only to stewards, governance domain owners, and data product owners.

Manage data product access policies

To manage access policies, you need data product owner or data steward permissions.

1. If your data product is published, first unpublish it to manage policies.

2. On your data product page, select **Manage policies**.

3. From the policy configuration window, you can create and manage your data product's access policy. Learn how to set up data product access policies.

Add and remove data assets

A data product groups together data assets. Grouping assets in this way makes it easier to discover them. To add or remove data assets from a data product, see the following sections.

Add data assets

> **Note**
>
> When you add data assets, the search only shows assets that the governance domain is scoped for, or that the user has access to in Microsoft Purview Data Map.

To add data assets to a data product:

1. In Unified Catalog, select **Catalog management**, and then select **Data products**.
2. Select the data product you want to add assets to.
3. Select **Add data assets** under the description and use cases.
4. Search for your data assets by using keywords or filters. To add filters, select **Add filter**.
5. Select any assets you want to add to the data product.
6. Select **Selected assets** to edit your selected asset list.
7. When you're done selecting assets, select **Add**.

You see your newly added assets in your data product.

Remove data assets

> **Tip**
>
> Collaborate with your data quality team to perform necessary cleanup work, as described in the following section, before removing data assets and deleting data products.

Before you can remove a data asset from a data product, you must remove all data quality rules and previous data quality scans from the asset. For more information, see deleting data quality history.

Although data quality stewards might not be involved in creating data products, they're the subject matter experts in data quality rules and scans. Only users holding the Data Quality

Steward role can remove data quality rules and scans, so data product owners should work with data quality stewards to get rules and scans removed before the asset in question can be removed from the data product.

Follow these steps to remove a data asset:

1. In Unified Catalog, select **Catalog management**, and then select **Data products**.
2. Select the data product you want to remove assets from.
3. If the assets you want to remove are on the front page, select the ellipsis (...) on the data asset's card, and then select **Remove**.
4. If you don't see the assets on the main product page, select **View all data assets**. There you can search and filter the assets. To remove them, select their ellipsis button and select **Remove**.

   > **Note**
   >
   > The remove action is unavailable if the data asset has data quality rules running on it or failed data quality run history. You need the Data Quality Steward role to resolve these issues before the asset can be removed.

Remove assets that are deleted from Data Map and metadata sync from Data Map (preview)

If you delete a data asset in a data product in Data Map, a banner appears in the data product to inform you of the asset deletion. However, the asset still appears in the data product until you remove it.

Open the data product and select **View all data assets** to view the list of assets. Assets with a caution symbol on its row are deleted from Data Map. To remove the deleted asset from your data product, select ellipsis (...) in the row, and then select **Remove**.

When you open the asset detail page from assets in a data product, the metadata is retrieved from the Data map. Metadata changes in Data Map are now reflected in Unified Catalog after they're updated in Data Map or initiated by selecting the refresh button.

For metadata represented within parent assets, such as columns within tables and views, the changes may not appear immediately. However, in all cases, metadata changes are reflected in

Unified Catalog within 24 hours.

Manage linked resources

Link glossary terms and OKRs directly to your data products to improve understanding, apply policies, and associate your data products with your business goals.

When you map assets in your data product to critical data elements, the system automatically adds critical data elements.

Add linked resources

1. On the data products page, select the data product you want to add terms or OKRs to.
2. Under **Glossary terms** or **OKRs**, select the + button next to the listing of those items.
3. Search for your terms or OKRs by using keywords, governance domain, or filters by selecting **Filter**.
4. Select any terms or OKRs you want to add to the data product.
5. Select the **Add** button.
6. You can now see the term or OKR in the list for your data products.

Remove linked resources

1. On the data products page, select the data product you want to remove terms or OKRs from.
2. Under **Glossary terms** or **OKRs**, find the item you want to remove, and select its ellipsis button.
3. Select **Remove**.

> **Note**
>
> The system automatically adds critical data elements based on the assets in your data product. For more information, see managing critical data elements.

Terms of use

1. On the data products page, select the data product you want to manage terms of use for.
2. Select the **Terms of use** attribute, on the right side of the page.
3. To add more terms of use:
   a. Select + **Add link**.

b. Choose a particular data asset to link the use to (optional).
c. Enter a friendly name for the terms.
d. Enter the link to the terms of use.
e. Select **Create**.
f. Select **Done**.

4. To remove any terms of use:
   a. Hover over the term you want to remove.
   b. Select the trashcan **remove** button.
5. When you're finished, select **Done**.

Documentation

1. On the data products page, select the data product you want to manage documentation for.
2. Select the **Documentation** attribute on the right side of the page.
3. To add documentation:
   a. Select **Add link**.
   b. Choose a particular data asset to link the documentation to (optional).
   c. Enter a friendly name for the documentation.
   d. Enter the link to the documentation.
   e. Select **Create**.
   f. Select **Done**.
4. To remove documentation:
   a. Hover over the documentation you want to remove.
   b. Select the trashcan **Remove** button.
5. When you're finished, select **Done**.

Endorse a data product

As a Unified Catalog grows in size, it's important for data consumers to understand what data they can trust. Data consumers need to know if a data product meets their organization's quality standards and can be regarded as reliable.

Data product owners can now set the **Endorsed** flag for their data products to indicate that they certified their data products, and build confidence in the quality of their data product.

To endorse a data product, you need data product owner permissions.

1. On your data product page, select **Edit**.

2. Select **Details**.
3. Select the **Mark as 'Endorsed'** checkbox.
4. Select **Save**.

Delete data products

Plan carefully before deleting a data product, as the process requires several steps.

1. First, unpublish the data product.
   a. Also, remove links to all related business concepts and remove any data assets within the data product.
   b. If any assets have data quality rules running on them or failed data quality run history, delete the data quality history.

2. Then, inform subscribers of the data product about the upcoming deletion and delete their access request to the data product.

3. When you complete all tasks, select **Delete** on the data product's page to delete the data product.

Next steps

- Set up access to data products
- Request access to a data product
- Learn about glossary terms

Last updated on 09/09/2026

Critical data elements (preview)

Not all data elements have the same importance or sensitivity. Dedicating resources to manage the quality of all data indiscriminately can be impractical and costly. Critical data elements are a logical grouping of important columns across tables in your data sources that you can use to focus your efforts.

These columns should be critical pieces of information that are necessary for decision making and so need to be governed with the highest care. For example, a "Customer ID" critical data element can map "CustID" from one table and "CID" from another table into the same logical container. Users can match this value across data assets to make connections, and when data producers create a new asset, they can use this element as a blueprint to provide quality information in the correct format.

Critical data elements (preview) in Microsoft Purview Unified Catalog are created within governance domains and can have policies set to manage these important pieces of information. When you create a critical data element, your organization can allocate resources strategically, focusing governance effort on areas that have the most significant effect on the business.

Prerequisites

- You must be using Microsoft Purview Unified Catalog.
- To create critical data elements and add columns to them, you must have data steward and data product owner permissions. For more information, see governance roles and permissions.
- To access data assets in the data map, you need at least data reader permissions on the collection where the asset is housed. For more information about data map permissions, see Data Map permissions.

Access critical data

Critical data is defined in governance domains. To see critical data in Unified Catalog:

1. In Unified Catalog, select **Catalog management**, then select **Governance domains**. You see a list of all the governance domains you have access to based on your permissions.
2. Select the governance domain where you want to see your critical data.
3. On the **Details** tab, find the **Critical data elements** card and select **View all**.

4. The **Critical data elements (preview)** page lists all of your critical data elements. Select a name to view its details.

You can also view a list of all critical data elements:

1. In Unified Catalog, under **Discovery**, select **Enterprise glossary**.
2. Select the **Critical data elements** tab.
3. Select the name of a critical data element to open its details page.

Filter your view by custom attribute

Use business concept attributes as a filter when exploring critical data elements.

1. On the **Critical data elements (preview)** page, select **Add filter**.
2. At **Filter**, select an attribute name from the dropdown list.
3. At **Operator**, select a condition, such as **Equals** or **Starts with**, which varies based on the kinds of values allowed by the attribute.
4. At **Value**, enter your value.
5. Select **Apply**.

Critical data element details

Select the name of a critical data element to view its details page. You can create and manage access policies for the critical data element by selecting **Manage policies**.

Overview

On the **Overview** tab, you see a description and:

- **Columns**: All data columns mapped from your assets to this critical data element. You can add columns.
- **Associated data products**: An automatically generated list of data products that have assets associated with your critical data element.

> **Note**
>
> The list of associated data products only refreshes when you add or remove columns.

Custom attributes

Under **Custom attributes**, you can see any custom attribute groups and attributes within groups, and the values set for each attribute. Only attributes with values are shown by default, but you can toggle **Show attributes without a value** on to view all attributes.

Create critical data

> ⓘ Note
>
> To create and upload critical data elements, you need the Data Steward role.

There are two ways you can create critical data elements:

1. Individually
2. In bulk, by importing a formatted CSV file (preview)

Create a single critical data element

1. In Unified Catalog, select **Catalog management**, then select **Governance domains**.
2. You see a list of all the governance domains you have access to based on your permissions. Select the governance domain where you want to add critical data.
3. On the **Details** tab, find the **Critical data elements** card and select **View all**.
4. Select **New critical data element**.
5. On the **Basic details** page, enter a name for your critical data. If you use a name that already exists, you'll see a warning during the creation process. While duplicating names isn't recommended, you won't be blocked from using a duplicate name.
6. Add a description, owner, and select an expected data type. Then select **Next**.
7. On the **Custom attributes** page, you can set values for all custom attributes defined by your admin in the **Custom attributes** section of the catalog. If the attribute is marked as required, you need to fill out all values on this page before you can complete the process.
8. When done, select **Create**.

You see your critical data in the list of critical data elements.

> ⓘ Note
>
> - Ensure your governance domain is published before you publish your critical data element.

- Other users can't see this critical data as a filter in the data product search until its status is set to **Published**.

Bulk import critical data elements (preview)

You can import multiple critical data elements at once into a governance domain by uploading a formatted CSV file. During the bulk import process, you download a sample CSV file containing a header row. You fill out the file according to the instructions below, then reupload it. The critical data elements then appear in the governance domain in a draft state.

The bulk import process can't be used to edit or update critical data elements.

> ⓘ **Note**
>
> - You must download the sample CSV (.csv) file and reupload it according to the following instructions. The file must have at least one row of data, not including the header row.
> - The CSV file can only include 1,000 rows; more rows cause an upload error. If you need to import more critical data elements, fill out and upload a separate CSV file.
> - Special characters are allowed in all fields.
> - Each critical data element name must be unique; duplicate names aren't supported for import.

Follow these steps to import critical data elements in bulk to a governance domain:

1. In Unified Catalog, select **Catalog management**, then select **Governance domains**.
2. Select the name of the governance domain you want to add critical data elements to.
3. On the **Details** tab of the governance domain's details page, find the **Critical data elements (preview)** card and select **View all**.
4. On the **Critical data elements (preview)** page, select **Import**.
5. On the **Import** flyout pane, select **Download a sample CSV file**.
6. Open the downloaded file. Fill out the worksheet as described below to avoid upload errors.
    - **Column A - name** (required): Enter the critical data element name. Every name must be unique; there can't be multiple rows with the same name. Maximum length is 256

characters.

- **Column B - description** (required): Enter a description for the critical data element. Maximum length is 10,000 characters.
- **Column C - owners** (required): Enter a valid email address that belongs to the same tenant that you're importing from. You can add multiple email addresses separated with a semicolon; for example: `owner1@contoso.com;owner2@econtoso.com`. Groups without email aren't supported.
- **Column D - data type** (required): Enter one of the possible values: number, text, date/time, Boolean.

7. **Column E and onward**: These are custom attributes, which differ by governance domain based on how they're defined and scoped. Note these formatting requirements for specific entry types: - **Multiselect options**: Enter the objects inside square brackets and in quotes; for example, `["a"]`, or `["a", "b"]`. - **Single select options**: Enter the raw value; for example, `a1`, `b2`. - **Booleans**: Enter `true` or `false`; can be upper or lowercase. - **String or rich text**: Use regular text; don't add quotes unless you want quotes reflected in the attribute. - **Numeric - short, int, long, double, float, bit, byte**: Don't add quotes, just enter the number.

8. After formatting your CSV file, save it locally. Return to the **Import** flyout pane, select **Browse**, then find and select your CSV file to upload it.

9. On the **Import** flyout pane, select **OK** to start the import process.

If there's a problem with the CSV file, an error message appears on the **Import** pane explaining the error so you can fix the CSV file and reupload it.

If there are no problems with the file, the import process begins. Depending on the size of the CSV file, it might take some time to complete the import process. If the file is small, the import happens immediately and you see a success message. Larger files take longer for the import process to complete.

Check the status by returning to the governance domain details page. On the **Details** tab, select **Monitoring** to view the status of your bulk import jobs, including any errors, on the **Monitoring** page.

Add columns

After you create your critical data, you can add data columns to map critical columns from your assets to this logical concept.

1. Select your critical data element from the list of critical data elements in your governance domain.
2. Select the + **Add column** button.
3. Search for your data sources by using keywords and filters, and select a data source.
4. In that data source, select any relevant column.
5. Select **Add**.
6. You should now see all your columns mapped to your critical data element.
7. Repeat for any number of data sources and columns.

   Below the columns you add, you also see the list of data products that are associated with the assets you add to your critical data element.

Remove column

To remove a column from a critical data element, select the ellipsis **...** button in the row where the column is listed, and select **Remove**.

If you can't find the column you want to remove, select the **View all columns** button to see the full list of columns.

Update and edit critical data

> ⓘ **Note**
>
> To edit a critical data element, you need the data steward role.

1. On the critical data's details page, select **Edit**.
2. On the **Basic details** page, you can edit the name, description, owner, and expected data type. If you change the name to one that already exists, you'll see a warning about the duplication but you can proceed to use the duplicate name.
3. On the **Custom attributes** page, you can set values for all custom attributes defined by your admin in the **Custom attributes** section of the catalog. If the attribute is marked as required, you need to fill out all values on this page before you can complete the process.
4. Select **Save** to save your changes.

To update the status of your critical data element, select the **Status** button to change between: **Draft**, **Published**, and **Expired**.

Manage related terms

You can link your critical data elements to related glossary terms.

1. Select the governance domain where your critical data element resides.
2. Select the **Critical data elements** button.
3. Search or browse for the critical data element and select it.
4. To link your critical data element to a glossary term, select the **+ Add term** button.
5. Search for the term, or terms, you want to link and select them.
6. To add your selected terms, select **Add**.
7. To remove a related term, select the term, then select the **...** ellipsis button, and then the **Remove** button.

Delete critical data

To delete a critical data element, you need to unpublish it and delete all columns within it, and any links to glossary terms. You can then select the **Delete** button and proceed with deleting the critical data element.

Remove data that's deleted from Data Map (preview)

If the parent data asset of a column in a critical data element is deleted in Data Map, a banner appears in the critical data element to inform you of the asset deletion. However, the column still appears in the critical data element until you remove it.

Open the critical data element to view the list of columns. Columns with a caution symbol on its row have been deleted from Data Map due to their parent data asset deletion. To remove the deleted column from your critical data element, select ellipsis (...) in the row, then select **Remove**.

Manage access policies

To manage access policies for critical data, you need data steward permissions.

1. On your governance domain's page, select your critical data elements.

2. Select a critical data element.

3. Select **Manage policies**.
4. From the policy configuration window, you're able to create and manage access policies for your critical data elements.
5. For more information about setting up access policies, see the documentation about managing access policies.

   These policies apply to all associated data products in your critical data element.

Last updated on 04/20/2026

Glossary terms in Unified Catalog

Glossary terms in Microsoft Purview Unified Catalog are active values that provide context. They also apply policies that determine how you manage, govern, and make your data discoverable for use.

If you already use Microsoft Purview, you're familiar with the business glossary in general. It provides a vocabulary for business users. It consists of business terms that can be related to each other and allows them to be categorized so that they can be understood in different contexts. This structure helps abstract the technical jargon associated with the data repositories and allows the business user to discover and work with data in the vocabulary that is more familiar to them. Microsoft Purview takes these terms a step further by allowing them to play an active role in data governance. Now data stewards can use terms to apply policies and scale data governance as their data estate grows.

> **Note**
>
> You can now bring classic glossary terms into Unified Catalog and link all of your terms to data assets and columns. Follow the instructions to migrate glossary terms from the classic data governance experience and enable the curation of data assets (preview).

What is a glossary term?

Terms are individual concepts that define the business, processes, and systems used in an organization. You can apply them across a data estate, relating to data assets and data products to provide business context to your users. A technically named SQL table becomes less mysterious once it's linked to your "Account" business term, since users can immediately see what aspect of the business it's related to.

Create glossary terms under governance domains to impart context specific to each part of your organization. Both sales and marketing might use the same term to mean different things, and your governance domains help your team differentiate between those meanings.

After you create glossary terms, you can link them to data products, data assets, columns, and critical data elements to provide context and to provide specific data governance based on business context. It's easy to understand how a term and its definition provide business context

to a data product or asset, but how do terms provide data governance? Terms are an active part of Microsoft Purview that provide data governance by now containing policies.

Policies in a business term apply specific business health goals, data governance requirements, and terms of use to any data product or asset that you apply the term to.

How does this scale data governance?

Glossary terms were previously static terms that you applied to define a concept. But now they can be activated as objects that attach to policies that determine how you manage, govern, and make your data discoverable for use. Terms for data access and data health controls can trickle down to data products. Today, data stewards apply policies on single data objects. That's not scalable. Now a data steward can create a single term and create its policies. They might not know where to apply that term, but any time they apply the term to a data product or asset, all the associated policies are automatically applied. Active terms allow your data estate to naturally scale, and maintain security and discoverability without overburdening your data stewards.

Next steps

- Create and manage glossary terms

Last updated on 04/24/2026

Create and manage glossary terms

Glossary terms in Microsoft Purview Unified Catalog provide a vocabulary for business users. These terms allow users to discover and work with data in the vocabulary that is more familiar to them, rather than using abstract technical jargon inherited from data sources.

Use business terms to define a shared vocabulary for your organization. By creating terms, identifying their synonyms, acronyms, related terms, and more, you can create a flexible controlled taxonomy organized in a hierarchical way. Glossaries of terms help bridge the communication gap between various departments in your company by providing consistent definitions for concepts, metrics, and other important elements across the organization.

This article describes how to manage glossary terms in Microsoft Purview. It explains how to create a glossary term in a governance domain, and how to link glossary terms to data products, data assets, columns, and critical data elements.

Prerequisites

- Any user can read published glossary terms.
- To create, edit, and manage glossary terms, you need the Data Steward role in the governance domain where the terms reside.
- To link glossary terms to data assets and columns, in addition to needing the Data Steward role, you also need the Data Reader role in the assets' collections.

Access glossary terms

1. In the Microsoft Purview portal, open **Unified Catalog**.
2. Under **Catalog management**, select **Governance domains**.
3. Select the governance domain you want to see glossary terms for.
4. On the **Glossary terms** card, select **View all**.

You see a list of all the glossary terms for that governance domain. You can search or sort through these glossary terms, filter them, or change the view between a list, a compact list, or a tree.

> **Note**

There might be multiple pages of terms. Use the arrow buttons and the page selector to toggle between these pages.

You can also view a list of all glossary terms:

1. In Unified Catalog, under **Discovery**, select **Enterprise glossary**.
2. Select the **Glossary terms** tab.
3. Select the name of a glossary term to open its details page.

Filter your view by custom attribute

You can use business concept attributes as a filter when exploring glossary terms.

1. In Unified Catalog, go to **Discovery** > **Enterprise glossary**.
2. On the **Glossary terms** tab, select **Add filter**.
3. At **Filter**, select an attribute name from the dropdown list.
4. At **Operator**, select a condition, such as **Equals** or **Starts with**, which varies based on the kinds of values allowed by the attribute.
5. At **Value**, enter your value.
6. Select **Apply**.

Glossary term details

Select a term from the list to view its details page. To create and manage access policies for the critical data element, select **Manage policies**.

Overview

The **Overview** tab contains a description of the term.

Custom attributes

Under **Custom attributes**, you can see any custom attribute groups and attributes within groups, and the values set for each attribute. Only attributes with values are shown by default, but you can toggle **Show attributes without a value** on to view all attributes.

Related

The **Related** tab displays any related elements, such as acronyms, synonyms, and resources. It also shows you which data products, data assets, and columns the term is linked to.

Observability

If you set up data observability (preview), the **Data observability** tab is where you see its visualizations.

Create glossary terms

> **Note**
>
> To create and upload glossary terms, you need the Data Steward role.

There are two ways you can create glossary terms:

1. Individually, by creating one term at a time.
2. In bulk, by importing a formatted CSV file (preview).

Create a single glossary term

1. In Unified Catalog, select **Catalog management**, and then select **Governance domains**.
2. Select the name of the governance domain where you want to add a term.
3. On the **Details** tab for the domain, find the **Glossary terms** card and select **View all**.
4. Select **New term** to create a single new term.
5. On **Basic details**, enter a name and a definition for your term. If you use a name that already exists, you'll see a warning during the creation process. While duplicating names isn't recommended, you won't be blocked from using a duplicate name.. The description is limited to 10,000 characters.
6. Select an owner or several owners for the term.
7. For **Parent term**, select an existing term within the same domain as a parent term. Then select **Next**.

   > **Note**
   >
   > You can visualize parent and child relationships in a tree-like hierarchy within the terms list page by selecting **Tree** in the views drop-down.
8. On **Acronyms**, add any related acronyms or leave it blank, and then select **Next**.
9. On **Resources**, add any related resources, such as links to documentation or other sources that provide context for your term, or leave it blank. Then select **Next**.

10. On the **Custom attributes** page, you can set values for all custom attributes defined by your admin in the **Custom attributes** section of the catalog. If the attribute is marked as required, you need to fill out all values on this page before you can complete the process.
11. When done, select **Create**.

Your term is created in a draft state, where only stewards and domain owners can see it. To make it visible to all users, you need to publish the term.

You arrive at the details page for your new glossary term, where you can edit or manage your term, publish your term, or link the term to data products, data assets, columns, and critical data elements.

Bulk import glossary terms (preview)

You can create many glossary terms at once and import them into a governance domain by uploading a formatted CSV file. During the bulk import process, you download a sample CSV file containing a header row. You fill out the file according to the instructions below, then reupload it. The terms then appear in the governance domain in a draft state.

The bulk import process can't be used to edit or update glossary terms.

> **Note**
>
> - You must download the sample CSV (.csv) file and reupload it according to the following instructions. The file must have at least one row of data, not including the header row.
> - The CSV file can only include 1,000 rows; more rows cause an upload error. If you need to import more terms, fill out and upload a separate CSV file.
> - Special characters are allowed in all fields.
> - Each term name must be unique; duplicate names aren't supported for import.

Follow these steps to import glossary terms in bulk to a governance domain:

1. In Unified Catalog, select **Catalog management**, then select **Governance domains**.
2. Select the name of the governance domain you want to add glossary terms to.
3. On the **Details** tab of the governance domain's details page, find the **Glossary terms** card and select **View all**.

4. On the **Glossary terms** page, select **Import**.

5. On the **Import terms** flyout pane, select **Download a sample CSV file**.

6. Open the downloaded file. Fill out the worksheet as described below to avoid upload errors.

   - **Column A - name** (required): Enter the glossary term name. Every term name must be unique; there can't be multiple rows with the same name. Maximum length is 256 characters.

   - **Column B - description** (required): Enter a description for the term. Maximum length is 500,000 characters.

   - **Column C - owners** (required): Enter a valid email address that belongs to the same tenant that you're importing from. You can add multiple email addresses separated with a semicolon; for example: `owner1@contoso.com;owner2@econtoso.com`. Groups without email aren't supported.

   - **Column D - experts**: Enter the email address of one or more users who are considered experts of the term.

   - **Column E - acronyms**: Enter any common abbreviations or shortened versions of the term.

   - **Column F - resources**: Enter related resources, such as links to documentation or other sources that provide context for the term. Resource links must be a valid URL in this format: `https://www.contoso.com`; both `http://` and `https://` are acceptable. List multiple resources by following this example: `resourcename1|value1;resourcename2|value2`.

   - **Column G and onward**: These are custom attributes, which differ by governance domain based on how they're defined and scoped. Note these formatting requirements for specific entry types:
     - **Multiselect options**: Enter the objects inside square brackets and in quotes; for example, `["a"]`, or `["a", "b"]`.
     - **Single select options**: Enter the raw value; for example, `a1`, `b2`.
     - **Booleans**: Enter `true` or `false`; can be upper or lowercase.
     - **String or rich text**: Use regular text; don't add quotes unless you want quotes reflected in the attribute.
     - **Numeric - short, int, long, double, float, bit, byte**: Don't add quotes, just enter the number.

7. After formatting your CSV file, save it locally. Return to the **Import terms** flyout pane, select **Browse**, then find and select your CSV file to upload it.

8. On the **Import terms** flyout pane, select **OK** to start the import process.

If there's a problem with the CSV file, an error message appears on the **Import terms** pane explaining the error so you can fix the CSV file and reupload it.

If there are no problems with the file, the import process begins. Depending on the size of the CSV file, it might take some time to complete the import process. If the file is small, the import happens immediately and you see a success message. Larger files take longer for the import process to complete.

Check the status by returning to the governance domain details page. On the **Details** tab, select **Monitoring** to view the status of your bulk import jobs, including any errors, on the **Monitoring** page.

Publish terms

When your term is ready to use in your governance domain, publish it by following these steps:

1. Select the domain where your term resides.
2. On the **Glossary terms** card, select **View all**.
3. Search or browse for the glossary term and select it.
4. On the glossary term's details page, select **Publish**.

> **Note**
>
> Ensure your governance domain is published before you publish your glossary terms.

Manage term policies

To manage term policies, you need data steward permissions.

1. On your glossary term page, select **Manage policies**.
2. From the policy configuration window, create and manage your term policies. For more information, see the documentation about managing access policies.

Link terms to data products, assets, and critical data elements (preview)

You can link glossary terms to data products, data assets, columns, and critical data elements to add business context to your data. Glossary terms must be in **Draft** state in order to add links; if the term is published, select **Unpublish** on the term's page to put it in **Draft** state.

> **Note**
>
> - When working with data assets and columns, *in addition to* needing the **Data Steward** role, you need the **Data Reader** role in the collection to which the asset or column belongs. All users with the **Catalog Reader** role can see published glossary terms associated to data assets and columns.
> - You need to enable asset curation before you can attach terms to assets and columns.

1. Select the governance domain where the term you want to link resides.
2. On the **Glossary terms** card, select **View all**.
3. Search or browse for the glossary term and select it.
4. Select **Related**.
5. Depending on what you want to link it to, select any of the following options: **Add data product**, **Add data assets**, **Add column**, **Add critical data element**.
6. In the flyout pane, search for and select the items you want to link to, then select **Add**.

> **Tip**
>
> Your search results might span multiple pages. Check the page selector to view all pages.

You can view all your linked data products, data assets, and columns from within your glossary term. The **Related** tab lists the first 10 linked items of each type. Select **View all** to see and search the entire list.

To remove a link, select the element on the **Related** tab of the glossary term, then select **...**, then **Remove**.

Manage links to terms from data products and data assets

There are two other pathways through which you can manage links to glossary terms:

- Within a data product: Data Stewards can add or remove links to published and unpublished glossary terms from a data product's details page.

- Within a data asset: Data Stewards can add or remove links to published and unpublished glossary terms residing in their governance domains. Users with the Global Asset Curator role can add published glossary terms to data assets.

Manage related terms

Add related glossary terms from across all your governance domains to connect similar terms.

1. Select the governance domain where your term resides.
2. On the **Glossary terms** card, select **View all**.
3. Search or browse for the glossary term and select it.
4. On the glossary term's details page, select the **Related** tab.
5. To add a term, select **Add term** and choose whether to add the term as a synonym or a related term.

   > ⓘ **Note**
   >
   > You can link terms from different governance domains.

6. To remove a related term, select the **X** next to the term.

Update contacts

When you add owners to glossary terms during the creation or editing process, you automatically add them to the **Contacts** section on the glossary term's details page. You can add experts to glossary terms as a contact type. To update contacts, edit the glossary term and at **Basic details**, enter a user name at **Owner** to add them as a contact.

Edit glossary terms

1. Select the governance domain where you want to edit one of the terms.
2. On the **Glossary terms** card, select **View all**.
3. Search or browse for the glossary term and select it.
4. If the term is published, select **Unpublish**, and then select **Edit** to edit the name, definition, owners, or custom attributes. If you change the name to one that already exists, you'll see a

warning about the duplication but you can proceed to use the duplicate name. When you're done, select **Save**.

5. To edit an owner's description, select the pencil icon on the card with the owner's name, edit the text, and then select the check mark to save.
6. To update status, select **...**, and then select the new desired status.
7. When done editing, publish the term.

Bulk edit glossary terms (preview)

Use the bulk edit process to edit up to 50 glossary terms at one time. You can edit **Attributes**, such as owners and descriptions, and **Term relationships** by adding and removing synonyms and related terms. All glossary terms must be in **Draft** state for bulk editing.

1. Select the governance domain where you want to bulk edit terms.
2. On the **Glossary terms** card, select **View all**.
3. On the **Glossary terms** page, select **Bulk edit.**
4. Check the boxes next to the name of the **Draft** terms you want to edit, and then select **Edit items**.
5. On the **Editing terms** flyout pane, select **Select a new attribute**.
6. From the **Attribute** dropdown menu, select the attribute you want to edit:
   a. Description (required)
   b. Acronyms
   c. Owner (required)
   d. Expert
   e. Resources
   f. Parent term
7. For **Operation**, select one of these options:
   a. **Replace with**: Replace the value of the existing attribute.
   b. **Add**: Add a value to the attribute. This option isn't available for required attributes or for parent terms.
   c. **Remove**: Delete the attribute. This option isn't available for required attributes.
8. For **Replace with** and **Add** operations, enter a value at **New value**.
9. To edit more attributes, select **Select a new attribute**.
10. Expand **Term relationships** if you want to edit synonyms or related terms.
    a. Under **Synonyms** or **Related terms**, select **Select terms for operation**.
    b. Select **Add** or **Remove**.
    c. On the **Search and select** pane, check the box next to the synonyms and related terms you want to add or remove, and then select **Select**.

11. When you finish making edits, select **Save**.

Your changes are applied to all the terms you selected to edit.

> **Tip**
>
> You can also select **Publish** and **Unpublish** to publish draft glossary terms or unpublish published glossary terms in bulk.

Move terms between governance domains (preview)

You can move multiple glossary terms at once from one governance domain to another domain. To move terms, you need to be a Data Steward in both the domain you're moving terms from, and in the domain you're moving the terms to. You can only move terms in a **Draft** state. If you select a parent term to move, the term's associated child terms move with it.

Use these steps to move glossary terms in bulk:

1. Select the governance domain that contains the terms you want to move out.
2. On the **Glossary terms** card, select **View all**.
3. On the **Glossary terms** page, select **Move to**.
4. Select the checkbox next to the **Draft** terms that you want to move, then select **Move items**.
5. In the **Move to** window, select the governance domain from the dropdown that you want to move the terms into.
6. Select **Move**.

The glossary terms appear in their new governance domain and are removed from their previous domain.

Expire glossary terms

When you expire a glossary term, only Data Stewards and Governance Domain Owners can see it.

1. Select the governance domain where you want to remove or retire a term.
2. On the **Glossary terms** card, select **View all**.
3. Search or browse for the glossary term and select it.
4. Select **...**, and then select **Set to Expired**.
5. Select **Save**.

Delete glossary terms

To delete a glossary term, first unpublish it and remove any links to related business concepts such as glossary terms, data products, data assets, columns, and critical data elements. Then select **Delete** to delete the term.

Accessing the classic business glossary

You can still access the classic business glossary by following these steps:

1. In the Microsoft Purview portal, open **Unified Catalog**.
2. Select the **Catalog management** drop-down.
3. Select **Classic types**.
4. Select the **Glossaries** tab.

Last updated on 05/20/2026

Migrate classic glossary terms and enable asset curation (preview)

Microsoft Purview Unified Catalog provides a centralized place for creating and managing glossary terms for both business concepts and data assets. If you created glossary terms for data assets and columns in the classic data governance glossary experience, you can start migrating them into the unified glossary experience in Unified Catalog.

This article explains the one-time migration process and the benefits of managing all your glossary terms for business concepts and data assets in one centralized place.

> **Note**
>
> The process outlined in this article enables you to curate data assets with glossary terms. Therefore, even if you don't plan to migrate classic terms, you still need to complete the workflow described in the steps below to enable asset curation.

Overview

You can bring classic glossary terms into the Unified Catalog for centralized creation and management of all glossary terms. This migration process is also what enables curation of data assets. After your terms are migrated to the **.StagingDomain** domain (a system domain automatically created for migration), you can move them into the governance domain that manages them going forward.

With the new unified glossary experience, any glossary term can apply to a business concept and data asset and column. The **Related** tab on each glossary term page shows its relationship to other glossary terms, data products, critical data elements, data assets, and columns.

When you migrate your glossary terms, you can:

- Use Unified Catalog as your single entry point for glossary terms within governance domains.
- Connect your terms to logical business concepts and physical concepts, or data assets (tables and columns).
- Invoke Copilots to enhance your glossary term creation and management experience.

Who needs to complete this process

This process is a required step in order to curate, or attach glossary terms to, data assets.

If your organization has glossary terms in the classic data governance experience, the migration process brings those terms into the updated experience, Unified Catalog.

If you're new to Unified Catalog in the Microsoft Purview portal, or never created glossary terms in the classic data governance experience, you still need to follow the steps of the enablement process by confirming use of current terms in Unified Catalog.

Plan for the migration and enablement process

> **Important**
>
> **Review this section closely and plan early.** The migration and enablement process is a one-time action that you can't reverse. When the process finishes, terms in the classic governance experience become read-only, and term templates convert to custom attributes that you can't edit or delete. Use this information to help you develop a thorough migration plan that involves the right people and ensures the most efficient outcomes for your organization.

What gets migrated and what doesn't

Glossaries and terms

The migration process moves glossaries and glossary terms that you created in the classic data governance experience. You can find these classic terms in the new Purview portal in **Unified Catalog**, under **Catalog management**, on the **Classic types** page.

When you migrate terms, this metadata is migrated into Unified Catalog:

- Glossaries and terms within, including all out-of-the-box metadata such as acronyms and parent information.
- All existing relationships to assets and columns in Microsoft Purview Data Map.

Term templates

The migration process moves all term templates when you migrate terms. During this process, term templates convert to business concept attributes. The attributes have the **.StagingDomain**

domain scope, which you can extend to more domains.

You can't edit or delete business concept attributes, so be sure to review term templates and delete unwanted ones before you migrate them.

Workflows

Any workflows you have on classic terms are **not** migrated.

Cleanup work to do before migration

- **Review and delete unwanted terms.** Delete any classic terms that you don't want to migrate *before* beginning the migration process, so that unwanted terms don't clutter the **.StagingDomain** domain. This step greatly facilitates and speeds up your process of relocating terms once they're brought over into Unified Catalog.
- **Review and delete unwanted term templates.** Term templates convert to business concept attributes after migration. These attributes can't be edited or deleted, so be sure not to migrate anything you don't want to appear or manage in Unified Catalog.

What happens during migration

The migration process creates a system-generated **.StagingDomain** governance domain and adds the migrated terms to it. Depending on the number of terms, it can take up to 48 hours for all migrated terms to appear in the staging domain. All glossary terms are added to the staging domain as drafts.

Only users who have both the Governance Domain Creator and Data Curator roles can see the **.StagingDomain** domain. You can't edit or publish this domain.

What happens after migration

When migration finishes, all terms are added to the **.StagingDomain** domain. The classic terms no longer appear on the **Classic types** page in Unified Catalog.

The **.StagingDomain** domain stays in your list of governance domains, even after you reassign the terms to the appropriate governance domains. You can't delete it.

The classic glossary experience becomes read-only, and you can't create any new terms there. Classic glossary terms don't appear when you search for them in data assets in Data Map.

Classic term APIs become read-only after migration finishes.

Plan who moves and manages the migrated terms

Identify a user in your organization to conduct the migration. This user needs both the **Governance Domain Creator** and **Data Curator** roles. This user also needs access to the **.StagingDomain** domain.

Identify the Data Stewards in your organization who move the terms into the correct governance domains. Add each of those Data Stewards as owners of the **.StagingDomain** domain.

Required user roles

Expand table

Role	Necessary for	Description
Governance Domain Creator and **Data Curator** (both are required)	Performing the migration and accessing the **.StagingDomain** domain.	Only users with this role can start the migration process. The user who starts the migration process becomes the owner of the **.StagingDomain** domain. Users with this role can access the **.StagingDomain** domain to view and move migrated terms.
Data Steward	Adding glossary terms to a governance domain.	Data Stewards must be assigned to the governance domain that they're moving migrated glossary terms into. Data Stewards must also be added to the **.StagingDomain** domain so they can move terms from that domain to the appropriate governance domain.

Learn how to assign data governance roles and permissions.

Steps for migrating glossary terms and enabling asset curation

This process enables your organization to add glossary terms to data assets and columns. During this process, you choose whether to migrate your existing glossary terms in the classic experience into Unified Catalog, or to use the glossary terms you already have in the Unified Catalog without migrating classic terms.

Step 1: Assign the necessary roles

See the required roles for the migration process.

Step 2: Choose how to enable asset curation

1. In the Microsoft Purview portal, go to **Unified Catalog** > **Catalog management** > **Classic types**.
2. Select **Enable asset curation**.
3. On the **Enable asset curation** flyout pane, select one of the following options, then select **Enable**:

    - **Migrate classic terms to Unified Catalog**: Select this option to migrate classic terms into Unified Catalog. After migration, classic terms are deactivated and placed in a read-only state for reference. When full migration is complete, you can curate data assets and columns with Unified Catalog terms.
    - **Switch directly to terms in Unified Catalog**: Select this option if you don't have any classic terms, or you intend to no longer use your classic terms and don't want to migrate them. If you have classic terms and select this option, your classic terms are deactivated and placed in a read-only state.

If you select **Migrate classic terms to Unified Catalog**, you see a summary of what will be migrated on the **Migrate glossary terms** flyout pane. Migration is a one-time process that can't be reversed, so you have an option of migrating a small sample of terms to see how the process goes, or to do the full migration all at once. At **Mode**, select how you want to proceed:

- **Sample migration**: This option migrates up to five glossaries so you can see how the migration works. Select this option, and then select **Start sample migration** to start the sample migration.

- **Full migration**: This option moves all glossaries and terms at once. Select this option, and then select **Start full migration** to start the migration process. This process can't be stopped or reversed after you start it.

    > ⓘ **Note**
    >
    > After you start the full migration process, access to terms in the classic governance experience is restricted to read-only. Any classic terms added by using APIs going forward are automatically added to the **.StagingDomain** domain.

Step 3: View terms in the staging domain and add more users to roles

After you select **Start full migration**, the glossary terms gradually appear in the **.StagingDomain** domain on the governance domains page. Depending on the number of terms migrated, the

migration process can take up to 48 hours.

To locate the **.StagingDomain** domain:

1. In Unified Catalog, go to **Catalog management** > **Governance domains**.
2. At the top of the list, find the domain named **.StagingDomain**.

When you begin migration, assign the necessary **Data Stewards** to the **.StagingDomain** domain so they can move terms to governance domains. On the staging domain's page, go to the **Roles** tab. Select the icon next to the **Data Steward** header to add users to this role for the domain. You might also want to add more users as owners of the **.StagingDomain** domain to help manage the domain.

Step 4: Move terms from the staging domain to owning domains

> **Important**
>
> A single term can only have 100 relationships in Unified Catalog. Review the relationships and clean up before migrating.

When migration is complete, reassign the glossary terms from the **.StagingDomain** domain to the appropriate governance domains for ongoing ownership. To assign migrated terms to a governance domain:

1. In **Unified Catalog**, under **Catalog management**, select **Governance domains**.
2. Select **.StagingDomain**, at the top of the list of governance domains.
3. On the **Details** tab, find the **Glossary terms** card and select **View all**.
4. Select **Move to**, and then select the box next to the terms you want to move. Selected terms are moved into one governance domain. To move terms into separate domains, select the terms going into one domain at a time.
5. After selecting the terms to move, select **Move item**.
6. In the **Move to** window, select the governance domain from the dropdown list that you want to move the terms into.
7. Select **Move**.

The terms are moved into their new governance domain. You can now add glossary terms to data assets and columns. Data Stewards see a notice at the top of the glossary terms list page that new terms were added to their governance domain.

> **Note**
>
> - Term hierarchies always move as a unit. If you select a parent term, all of its children move with the parent. If you don't want the hierarchy to be moved, remove the parent terms first.
> - The migration process doesn't automatically move any terms with duplicate names in the destination governance domain.

Next steps

Get started linking glossary terms to your data assets and columns:

- Start from a glossary term and link it to an asset or column
- Start from a data asset and attach glossary terms

Last updated on 06/23/2026

Objectives and key results (OKRs) in Unified Catalog

10/09/2025

Objectives and key results (OKRs) (preview) in Microsoft Purview Unified Catalog are trackable business objectives tied to governance domains and data products to emphasize the value of business data.

Data governance is an important practice that organizations often overlook unless they have strict compliance requirements. While organizations seem to understand the value of data, they don't always recognize the role data governance plays in promoting healthy, useful data. As data estates become overgrown with ungoverned information, valuable data gets neglected.

OKRs link data products directly to real business objectives to bridge the gap between the business and the data estate. Data governance isn't just an IT task or engineering best practice, it's a critical part of value generation. To make use of your data, you need to maintain your data estate, and you can share the work among the experts who know the data best.

What is an OKR

An OKR is the goal or desired outcome of a governance domain; for example, a 10% raise in sales or a 3% reduction in support cases. Objectives should relate to everything an organization does and should define how they achieve their outcomes.

Parts of an OKR

Unified Catalog OKRs have several parts:

- Owners: The users responsible for maintaining the objective.
- Definition: An overarching description of your objective, which can be as broad as “Increase sales,” or as specific as "Use marketing campaigns to increase sales in the first quarter."
- Target date: When you expect to achieve your goal.
- Key results: The measurable, time-bound goals associated with your objective. These should be how you track progress toward your objective.

Owner

These users are responsible for maintaining a business objective. They could be the leaders of the business and should have knowledge of how the business functions and where it's headed. Multiple owners are encouraged to provide a complete perspective on the business and to support each other in owning the objective.

Definition

Good objectives are specific and achievable in a time frame. The definition is a high-level text description of the mission or current goal of the team or process, and why it matters. The description itself should be succinct, but it isn't enough on its own. It requires a way to measure the goal's achievement and progress, which are the key results.

Target date

The target date is when you expect to meet the goal. This date helps drive progress toward the goal and encourages achievable goals, since you must complete them within a specific time frame.

Key results

Key results show how you measure progress toward your objective. Your objective might have several key results that roll up to your overall goals.

A good metric for goals and objectives uses the SMART acronym:

- **S**pecific
- **M**easurable
- **A**chievable
- **R**ealistic
- **T**ime-bound

To tie your objectives to real results, key results have several features:

- Result: A statement that encompasses the specific, measurable result you expect.
- Progress status: Is your progress on track, behind, or at risk? This status helps your business identify and prioritize issues.
- Progress amount: Where do you stand in relation to your goal? You need to use specific numbers or percentages to represent this amount.
- Goal amount: What's your goal? You need to use specific numbers or percentages to represent this amount.
- Maximum amount: What's the maximum achievable?

You can edit these values so business users and owners can update their goals and results over time. These values are also converted to graphics, so business owners can see at a glance how their organization is progressing toward their goals and where there might be blockages.

Related data products

You can connect an OKR to all data products that are relevant to that objective. It can be the data used to measure that OKR, or the data used to investigate new business opportunities. Data consumers can use this connection to see all data related directly to their business objectives, and business owners can track the health of the data directly associated with their goals.

How OKRs generate business value

Data governance directly supports these business objectives. After all, data helps you both track and achieve these objectives. You can also use data to uncover new areas for improvement and to guide which objectives need prioritization. OKRs help to directly tie data to its business value.

For example: Your marketing team develops an email campaign results data product that they use to track the effectiveness of their email marketing campaigns. Your marketing governance domain has a "Customer Response" OKR, with the goal of at least 20% customer response to email, telephone, and mail campaigns sent out. A business owner connects the "Customer Response" OKR to the email campaign results data product and creates an "Email Campaign Response" key result to measure the current percentage of customer response. When the marketing team tracks their quarterly progress, their objectives directly tie to the data they use to track that information. They're directly aware of the health and value of the data associated with their OKRs and naturally encouraged to invest in the health of the data.

This approach not only encourages your users to focus on their objectives but also promotes health, security, right use, and data discovery directly related to your business objectives. Your business and your data estate grow together naturally to support your goals.

Next steps

- How to create and manage OKRs
- Learn about data products

Create and manage OKRs in Unified Catalog

10/09/2025

Objectives and key results (OKRs) in Microsoft Purview Unified Catalog are trackable business objectives tied to governance domains and data products to emphasize the value of business data.

This article describes how to manage OKRs in Microsoft Purview. It provides steps to create OKRs in a governance domain, create key results, and connect them to your data products.

Prerequisites

- Any user can read published OKRs.
- To create and edit OKRs, you need the steward role.

Access OKRs

1. In Unified Catalog, select **Catalog management**, then select **Governance domains**.
2. Select the governance domain you'd like to see OKRs for.
3. On the **Details** tab, find the **OKRs** card and select **View all**.

The full list of OKRs for that governance domain is displayed. Select an OKR to see its detailed view, and get more information about its key results.

You can also view a list of all OKRs:

1. In Unified Catalog, under **Discovery**, select **Enterprise glossary**.
2. Select the **OKRs** tab.
3. Select the name of an OKR to open its details page.

Filter your view by custom attribute

Use business concept attributes as a filter when exploring OKRs.

1. In Unified Catalog, go to **Discovery** > **Enterprise glossary**.
2. On the **OKRs** tab, select **Add filter**.
3. At **Filter**, select an attribute name from the dropdown list.
4. At **Operator**, select a condition, such as **Equals** or **Starts with**, which varies based on the kinds of values allowed by the attribute.

5. At **Value**, enter your value.
6. Select **Apply**.

Overview

The **Overview** tab shows key results, associated data products, target date, and current overall progress.

Custom attributes

Under **Custom attributes**, you can see any custom attribute groups and attributes within groups, and the values set for each attribute. Only attributes with values are shown by default, but you can toggle **Show attributes without a value** on to view all attributes.

Create OKRs

You can create OKRs directly in the Microsoft Purview portal.

1. In Unified Catalog, select **Catalog management**, then select **Governance domains**.
2. Select the governance domain you want to add an OKR to.
3. On the **Details** tab, find the **OKRs** card and select **View all**.
4. On the **OKRs** page, select **New OKR**.
5. On the **Basic details** page, at **Objective**, enter a statement describing the action you'll take and why it matters. For example: Revamp customer service to improve satisfaction and retention rates. If you use a name that already exists, you'll see a warning during the creation process. While duplicating names isn't recommended, you won't be blocked from using a duplicate name.
6. Select an owner or owners for your OKR.
7. Select a target date to meet the OKR.
8. Select **Next**.
9. On the **Custom attributes** page, you can set values for all custom attributes defined by your admin in the **Custom attributes** section of the catalog. If the attribute is marked as required, you need to fill out all values on this page before you can complete the process.
10. Select **Create**.

You see the details page for your new OKR, where you can edit or manage your OKR, view key results, and view linked data products.

> ⓘ **Note**

Only stewards and governance domain owners can see your OKR until you **publish it**.

Key results

Key results are how you track the overall progress of an OKR. They should be specific, measurable outcomes so you and your team can accurately gauge progress or completion.

The top key results are listed on the details page. To view any others, select **View all key results**.

Select a specific key result to see its details.

Create key result

To create a new key result for an OKR:

1. In your OKR's details page, select **Add key result**.
2. Add your key result statement - it should be a specific, measurable outcome that uses metrics. For example: Increase customer retention rate by 25%.
3. Select current progress status - on track, behind, or at risk.
4. Set progress amount as a percentage or a value.
5. Set goal amount as a percentage or a value.
6. Set the maximum allowed amount as a percentage or value.
7. Select **Create**.

Your newly created key result is now listed under key results in your OKR.

Update key result

1. In your OKR's details page, select the key result you want to update, or select **View all key results** to see a full list of key results.
2. Select the key result you want to edit, and select the **Edit** button.
3. Edit the result, progress status, progress amount, goal amount, or maximum amount.
4. Select **Save** to save your updates.

Data products

Add, remove, and view data product links all from within your business OKR.

View linked data products

You can view all your linked data products from within your OKR. The details page lists the first 10 linked data products. Select **View all data products** to see and search the entire list.

Add data product link

1. In your OKR's details page, select the **+ Link data product** button.
2. Search for the data products you want to link.
3. Select the data products you want to link.
4. Remove any selected products from the **Selected** list by selecting the **X** button next to their name.
5. To add your selected data products, select **Save**.

Remove linked data products

To remove a linked data product:

1. Open the OKR where the data product is linked.
2. Find the data product on the details page or select **View all data products** to search for it.
3. Select the ellipsis button (**...**) next to the data product and select **Remove link**.

Edit and manage OKRs

1. On the OKR's details page, select **Edit**.
2. On the **Basic details** page, update the objective, owner, and target date. If you change the name to one that already exists, you'll see a warning about the duplication but you can proceed to use the duplicate name.
3. Select **Save**.
4. On the **Custom attributes** page, you can set values for all custom attributes defined by your admin in the **Custom attributes** section of the catalog. If the attribute is marked as required, you need to fill out all values on this page before you can complete the process.
5. Select **Save**.

Draft, publish, or close

You can publish a drafted OKR, or unpublish a previously published OKR and set it to draft or closed. Closing an OKR limits its visibility to stewards and governance domain owners.

To publish or unpublish an OKR:

Closing an OKR limits its visibility to stewards and governance domain owners. Deleted OKRs are removed from the governance domain.

1. Select the governance domain you'd like to publish or unpublish an OKR from.
2. Select the **OKRs** card.
3. Search or browse for the OKR and select it.
4. Select **Publish** if in draft state and **Unpublish** if already published.
5. You can either **Set to draft** for further edits or **Set to closed** if you don't wish to use it any longer.

> **Note**
>
> Ensure your governance domain is published before you publish your OKRs.

Delete OKRs

To delete an OKR, first unpublish it and delete any key results and links to related data products. Then select **Delete** to delete the OKR.

Custom metadata (preview) in Unified Catalog

10/09/2025

Custom metadata plays a crucial role in data governance by providing you with the flexibility to add user-defined context to describe and organize your data. In Microsoft Purview Unified Catalog, you can define and apply organization specific business attributes to your data assets and business concepts that foster better organization, accessibility, and utilization of your data. The benefits of using custom metadata include:

- **Enhanced data discovery**: Custom metadata simplifies the search and retrieval of data by allowing users to apply specific tags, which are **attributes** in Unified Catalog, that make datasets easier to locate.
- **Help with compliance**: Organizations can use custom attributes to track regulatory requirements or flag sensitive data to help ensure adherence to legal and industry standards.
- **Streamlined communication**: Custom metadata facilitates better collaboration among teams by providing clear descriptions and classifications of data assets.
- **Optimized data utilization**: By categorizing data assets effectively, organizations can prioritize the use of high-value data for analytics and decision-making.

Custom metadata usage scenarios

Here are some examples of how you can use custom metadata to help achieve better data governance and compliance:

Example 1: Healthcare industry

A hospital's data catalog includes patient records, medical research, and operational data. By implementing custom metadata, administrators add attributes such as "Patient Age Group," "Disease Category," and "Research Priority Level" to the datasets. This approach enables the hospital to segment data effectively for clinical studies, prioritize urgent cases, and comply with healthcare regulations like HIPAA.

Example 2: Financial services

A financial institution uses custom metadata to govern its risk assessment data. Attributes like "Risk Rating," "Investment Sector," and "Regulatory Compliance Status" help ensure that

analysts can quickly identify high-risk investments, provide insights into sector-specific trends, and maintain legal compliance with financial regulations.

Example 3: Academic institutions

Universities use custom metadata to manage their research databases. Attributes like "Research Field," "Funding Source," and "Publication Status" enable academic departments to track progress, allocate resources effectively, and collaborate across institutions.

Example 4: AI and compliance

Organizations can include AI and compliance related attributes to define various data product types. AI-related attributes might cover aspects such as algorithm transparency, explainability, and model performance metrics, while compliance attributes might address regulatory requirements, data privacy, and ethical considerations. By incorporating these factors, organizations can ensure their data products are robust, compliant with laws and standards, and aligned with responsible AI practices.

Example 5: Tracking translations

Language can be a pivotal attribute to describe and discover glossary terms in different languages. By incorporating multilingual capabilities into metadata attributes, organizations can ensure seamless communication and understanding across diverse linguistic groups. This approach not only facilitates global collaboration but also enhances accessibility, enabling users to interact with data in their preferred language while maintaining consistency and clarity in terminology usage.

Types of custom metadata: attributes

Currently, you can use two types of custom metadata:

1. Business concept attributes: Attributes you create and apply to business concepts, such as governance domains, data products, critical data elements, glossary terms, and objectives and key results.
2. Data asset attributes: Attributes you create and apply to data assets.

You can find, create, and manage your custom metadata at **Unified Catalog** > **Catalog management** > **Custom metadata (preview)**.

Each type of attribute appears on its own tab on the **Custom metadata (preview)** page. The attribute list on each tab is organized by attribute group, and each attribute is nested within its group.

Get started

- Create business concept attributes
- Create data asset attributes

Create and manage business concept attributes (preview)

07/16/2025

Business concept attributes are a type of user-defined custom metadata in Microsoft Purview Unified Catalog that you can create and apply to your organization's business concepts for better organization, governance, and discovery.

The **Custom metadata (preview)** area in **Unified Catalog** > **Catalog management** is where you create and manage business concept attributes.

Workflow and tasks by role

Creating and managing business concept attributes involves three basic steps, each involving certain user roles and functions in your organization:

1. Admins create and maintain the attributes and attribute group definitions.
2. Data stewards and data product owners update the values for business concept attributes defined by their admins.
3. Data consumers or catalog readers can use business concept attributes to filter their view of published business concepts.

Use this table to guide you to relevant tasks depending on your role in the organization. See Data governance roles and permissions for role descriptions and how to assign roles.

Expand table

Role	Task
Admins - Governance Domain Creators and Data Curators	Create attribute groups Create attributes Set the scope for an attribute group Edit groups to add attributes Require values for attributes Expire attributes and attribute groups
Data Stewards Data Product Owners	Edit attributes Add values to attributes Use attributes as filters for viewing business concepts Delete attribute values
Catalog readers (local and global)	Use attributes as filters for viewing business

Role	Task
	concepts

Know before you start

Before you start creating attributes, we recommend reviewing these sections:

- Understand attributes and groups
- Field types for attribute values
- Scope settings for attribute groups

> **Note**
>
> As you start planning your attribute strategy, keep in mind that after you create a business concept attribute, **edits** are limited. Also, attributes and attribute groups can't be deleted; they can only be **expired**.

Understand attributes and groups

- **Business concept attributes**: Business concepts, such as governance domains and glossary terms, help you organize and understand your data in ways that make sense for your organization. An attribute is user-defined metadata that you can apply to your business concepts to better describe your organization's unique needs.
- **Attribute groups**: A category by which you group attributes, allowing for easier organization and consumption.

Field types for attribute values

Business concept attributes can include several types of values, such as text-based descriptions, numerical identifiers, or date-based identifiers. During the attribute creation process, the field type options are:

- Text
- Single choice: When you select this type, a **Choices** field appears, where you enter a name for the choice and can add more choices.
- Multiple choice: When you select this type, a **Choices** field appears, where you enter a name for the choice and can add more choices.
- Date
- Boolean

- Byte
- Double
- Float
- Integer
- Long
- Rich text
- Short

See Add values to attributes for more details.

Scope setting for attribute groups

The **Scope** setting for attribute groups allows you to apply attributes to a broad array of governance domains and business concepts in Unified Catalog at once.

When you create an attribute group, the **Scope** settings allow you to select all or some governance domains, business concepts, and data product types. Setting a scope means that every attribute in that group will apply to the governance domains, business concepts, and data products that you selected for that group.

Scope defaults to all governance domains and business concepts during the attribute group creation process. If scope needs to be narrowed, select each dropdown within **Scope** and set it to the desired scope.

When creating or editing an attribute, you can accept the **Inherit Group Scope** setting, which applies the group scope to the attribute, or you can override the inherited setting and select individual governance domains and business concepts as the scope for that particular attribute within the group.

Create and edit attribute groups

All business concept attributes must be connected to an attribute group. You might want to develop a grouping strategy ahead of time, then create one or more attribute groups before creating the attributes you'll assign to the groups. Alternatively, you can create an attribute group at the point in the attribute creation process where you assign a group to the attribute.

> **Note**
>
> When you apply an attribute group to a business concept, all attributes in that group are added to the concept.

Create an attribute group

1. Sign in to the Microsoft Purview portal > **Solutions** > **Unified Catalog** with an account that has a role in a tenant level role group.
2. Under **Catalog management**, select **Custom metadata**.
3. On the **Business concept attributes** tab, select **New**, then select **New attribute group**.
4. On the **New attribute group** flyout pane, enter an **Attribute group name** and an optional description.
5. You can create a new business concept attribute at this stage by selecting **New attribute**. This is the **only** stage at which you can mark an attribute as required.
6. Under **Scope**, select the **Governance Domains**, **Business Concepts**, and **Data product types** that will apply to the group and all attributes assigned to the group. All values are checked by default, so be sure to uncheck any values you don't want. Get more details on scope settings.
7. When done, select **Create**.

The group is now listed on the **Business concept attributes** tab of the **Custom metadata (preview)** page.

View and edit an attribute group

Find the name of an attribute group on the **Business concept attributes** tab, hover over its row and select the **Edit** pencil icon. A flyout pane appears, which contains all the details of the attribute group, including a list of all attributes in the group.

Add a new attribute to the group by selecting **New attribute**.

To change the scope of an attribute in the group or edit any details of the attribute, select the attribute name, make edits on the **Edit attribute** flyout pane, then select **Apply**.

Mark an attribute as required

When you create a business concept attribute as part of a business concept attribute group, you can make the attribute required by selecting **Mark as required**. Marking as required means that a value must be provided for the attribute. When a business concept is created or edited, the value for the required attribute must be filled out before you can complete the creation or editing process.

> **Important**
>
> - You can only select **Mark as required** when you create an attribute during the process of creating an attribute group. You can't mark the attribute required when creating just an attribute on its own.
> - If an existing attribute isn't marked as required, you can't mark it as required during the editing process.
> - Once you select **Mark as required** and create the attribute, you can't unselect it during the editing process.

Create attributes

1. Sign in to the Microsoft Purview portal > **Solutions** > **Unified Catalog** with an account that has the Data Steward role.
2. Under **Catalog management**, select **Custom metadata**.
3. On the **Business concept attributes** tab, select **New**, then select **New attribute**.
4. On the **New attribute** flyout pane, enter an **Attribute name** and an optional description.

   > **Note**
   >
   > You can create more than one attribute with the same name and link them to different attribute groups. Providing a detailed description for identically named attributes can help avoid confusion or the misapplication of attributes.

5. Select **Mark as required** if you want to require a value be provided for the attribute. This option is only available when creating an attribute during the process of creating an attribute group (learn more).
6. At **Attribute group**, select a group name from the dropdown list.

   - You can create a new attribute group at this step by selecting **New**, then follow the steps to create an attribute group.
   - If you're creating the attribute at the same time as you're creating the group, the **Attribute group** selector doesn't show; the attribute is automatically associated to the group you're creating.

7. At **Field type**, select an attribute field type.

8. The **Inherit Group Scope** toggle is on by default, which means that the attribute will have the same scope as the group to which you assign it. To override scope inheritance, toggle off the switch. Then select the individual governance domains, business concepts, and data product types that you want the attribute to apply to. Get details on scope settings.
9. When done, select **Create**.

Your attribute now appears on the list of attributes on the **Business concept attributes** tab, nested under its group.

Edit attributes

An admin can make edits to attribute groups and attributes. You can update the description, the scope, or set an attribute or group as expired. When you edit an attribute group, you can add attributes to it.

Follow these steps to make edits:

1. Find the attribute or group from the list on the **Custom metadata** page.
2. Hover over the name and select the **Edit** pencil icon.
3. Make your edits, then select **Update**.

Add values to attributes

Data stewards and data product owners can set and update values for attributes that are defined by their admins. The values don't have to be filled in at the time of business concept creation. However, if an attribute is marked as required, then whenever a user creates or edits an attribute, the values must be filled in before the attribute editing process can be completed.

Delete attribute values

Attributes that are expired display a trash can icon next to them when viewed on a business concept's **Custom attributes** page. To delete an attribute value:

- Open the business concept details page and select **Edit**.
- On the **Custom attributes** page, find the attribute you want to delete (expired attributes appear as grayed out).
- Select the trash can icon next to the attribute. The attribute is removed from the view.
- Select **Save**.

Expire an attribute or attribute group

A business concept attribute can't be deleted, and a business concept attribute group can't be deleted. They can only be expired. When an attribute or attribute group is expired, you can't apply them to other business concepts, and they can't be reactivated.

- **When you expire an attribute**: Any business concepts that had the attribute applied will continue to show that attribute.
- **When you expire an attribute group**: All the attributes within the group are also expired. Any business concepts that had an attribute within this group applied to it will continue to show that attribute.

To expire an attribute or attribute group:

1. On the **Business concept attributes** tab of the **Custom metadata (preview)** page, find the attribute or attribute group you want to expire and select the **Edit** pencil icon.
2. On the **Edit attribute** or **Edit attribute group** pane, select the **Mark as expired** checkbox.
3. Select **Update**.

View attributes and use as filters

Attributes are viewable, and can be used as filters, in various ways:

- List view of all attributes on the Custom metadata (preview) page.
- Displayed on a business concept's details page
- As a filter when exploring business concepts.
- As a filter for exploring data products.

On the Custom metadata (preview) page

Business concept attributes are listed in **Unified Catalog** > **Catalog management** > **Custom metadata** on the **Business concept attributes** tab. The list is organized by group name. Use the arrow next to the group name to expand or collapse the list of attributes underneath a group.

On a business concept details page

You can view details about attributes on the details pages for all business concepts and data products:

- Governance domains
- Data products
- Critical data elements
- Glossary terms

- Objectives and key results (OKRs)

Filters for viewing business concepts

You can use business concept attributes as a filter when browsing and searching for business concepts.

1. Go to **Unified Catalog** > **Discovery** > **Enterprise glossary**.
2. Select the tab of the business concept you want to search in. If you're searching in a governance domain, select **View all** on a business concept card.
3. Select **Add filter**.
4. At **Filter**, select an attribute name from the dropdown list.
5. At **Operator**, select a condition, such as **Equals** or **Starts with**, which will vary based on the kinds of values allowed by the attribute.
6. At **Value**, enter your value.
7. Select **Apply**.

Filter for exploring data products

When you explore data products by governance domains, you can filter your view by data asset attributes.

1. In Unified Catalog, go to **Discovery** > **Data products**.
2. Select **Explore all**. You arrive at the **Explore data products by governance domain** page.
3. Select **Add filter**.
4. At **Filter**, select an attribute name from the dropdown list.
5. At **Operator**, select a condition, such as **Equals** or **Starts with**, which will vary based on the kinds of values allowed by the attribute.
6. At **Value**, enter your value.
7. Select **Apply**.

Create and manage data asset attributes

10/09/2025

Data asset attributes are a type of custom metadata in Microsoft Purview Unified Catalog that you can create and apply to your organization's data assets. These user-defined attributes provide a business or organization level context to an asset, helping data consumers understand the role an asset plays in the organization.

Prerequisites

To create a **data asset attribute**, you must hold the Data Curator role in the Data Map.

Manage attributes on the Custom metadata (preview) page

> **Note**
>
> Data asset attributes were previously called "managed attributes" and were located under **Catalog management** > **Classic types**.

The **Custom metadata (preview)** area under **Unified Catalog** > **Catalog management** is where you create and manage two types of attributes:

1. Data asset attributes
2. Business concept attributes

The **Data asset attributes** tab is where you find all of your data asset attributes nested within each attribute group.

The **Business concept attributes** tab lists all of your attributes nested within each attribute group. Learn how to work with business concept attributes.

Know before you start

Before you start creating attributes, review these sections:

- Understand attributes and groups
- Attribute field types
- Learn when you can mark an attribute as required

Understand attributes and groups

- **Data asset attributes:** A set of user-defined attributes that provide a business or organization level context to a data asset. Each attribute has a name and a value. For example, "Department" is an attribute name and "Finance" is its value.
- **Attribute group:** A grouping of data asset attributes that makes it easier to organize and consume them.

Attribute field types

Data asset attributes can include several types of values, such as text-based descriptions, numerical identifiers, or date-based identifiers.

During the attribute creation process, you can choose from the following field type options:

- Text
- Single choice: When you select this type, a **Choices** field appears, where you enter a name for the choice and can add more choices.
- Multiple choice: When you select this type, a **Choices** field appears, where you enter a name for the choice and can add more choices.
- Date
- Boolean
- Byte
- Double
- Float
- Integer
- Long
- Short

Create and edit attribute groups

You must connect all data asset attributes to an attribute group. You might want to develop a grouping strategy ahead of time, then create one or more attribute groups before creating the attributes you assign to the groups. Alternatively, you can create an attribute group at the point in the attribute creation process where you assign a group to the attribute.

As you develop a grouping strategy, consider which attributes in a group you want to mark as required. **You can only set attributes to be required when you create an attribute group first, and then add assets *while you're in the process of creating the group*.**

Create an attribute group

1. Sign in to the Microsoft Purview portal > **Solutions** > **Unified Catalog**.
2. Under **Catalog management**, select **Custom metadata**.
3. On the **Data asset attributes** tab, select **New**, then select **New attribute group**.
4. On the **New attribute group** flyout pane, enter an **Attribute group name** and an optional description. You can't change a group name after you create the group.
5. You can create a new data asset attribute at this stage by selecting **New attribute**. This is the **only** stage at which you can mark an attribute as required.
6. When done, select **Create**.

The group is now listed on the **Data asset attributes** tab of the **Custom metadata (preview)** page.

View and edit an attribute group

Find the name of an attribute group on the **Data asset attributes** tab, hover over its row, and select the **Edit** pencil icon. A flyout pane appears, which contains all the details of the attribute group, including a list of all attributes in the group.

Add a new attribute to the group by selecting **New attribute**.

You can't edit an attribute group name after you create the group.

Mark an attribute as required

When you create a data asset attribute as part of a data asset attribute group, you can make it required by selecting **Mark as required**. Marking an attribute as required means that a value must be provided for this attribute. When you edit a data asset, you must fill out the required attribute before you can close the editor.

> **Important**
>
> - You can only select **Mark as required** when you create an attribute during the process of creating an attribute group. You can't mark the attribute as required when creating just an attribute on its own.
> - If an existing attribute isn't marked as required, you can't mark it as required during the editing process.

- Once you select **Mark as required** and create the attribute, you can't unselect it during the editing process.

1. Sign in to the Microsoft Purview portal > **Solutions** > **Unified Catalog**.
2. Under **Catalog management**, select **Custom metadata**.
3. On the **Data asset attributes** tab, select **New**, then select **New attribute group**.
4. Select **New** and select **New attribute**.
5. Fill out your attribute details, and select **Mark as required**.
6. Select **Apply** and finish adding other attributes to complete your attribute group.

Create data asset attributes

1. In the Microsoft Purview portal, under **Solutions**, select **Unified Catalog**.
2. Under **Catalog management**, select **Custom metadata**, then select the **Data asset attributes** tab.
3. Select **New**, then select **New attribute**.
4. On the **New attribute** flyout pane, enter an **Attribute name** and an optional description.

   > **Note**
   >
   > You can create more than one attribute with the same name and link them to different attribute groups. Providing a detailed description for identically named attributes can help avoid confusion or the misapplication of attributes.

5. Select **Mark as required** if you want to require a value for the attribute. This option is only available when creating an attribute during the process of creating an attribute group (learn more).
6. At **Attribute group**, select a group name from the dropdown list.

   - You can create a new attribute group at this step by selecting **New**, then follow the steps to create an attribute group.
   - If you're creating the attribute at the same time as you're creating the group, the **Attribute group** selector doesn't show; the attribute is automatically associated to the group you're creating.

7. At **Field type**, select an attribute field type.
8. At **Applicable asset types**, select one or more data asset types that the attribute can be applied to. For example, if you select "Azure SQL Table" for an attribute, it applies to Azure SQL Table assets, but not Azure Synapse Analytics Dedicated Table assets. When done, select **Apply**.
9. Select **Create**.

These attributes are automatically applied to the specified resources without a value. You need to apply a value for each asset.

Edit a data asset attribute

After you create a data asset attribute, you can't change its name, the attribute group it belongs to, or the field type. To edit other details, such as the description or to mark the attribute as expired, follow these steps:

- On the **Data asset attributes** tab of the **Custom metadata (preview)** page, hover over the attribute name and select the **Edit** pencil icon.
- Make your edits, then select **Update**.

Add a value to a data asset attribute

After you create a data asset attribute, you need to add a value for each of your assets. To add values to your assets, follow these steps:

1. Search for your data asset.
2. On the **Overview** tab of the asset's details page, see the **Data asset attributes** section with all attributes that have values. (You can see attributes without values by using the **Show attributes without a value** toggle.)
3. Select **Edit**.
4. Under **Data asset attributes**, add values for each of your attributes.
5. If any attributes are **Required**, you can't save until you add a value for that attribute.

Expire a data asset attribute or attribute group

You can't delete a data asset attribute or a data asset attribute group. You can only expire them. When you expire an attribute or attribute group, you can't apply them to other data assets, and you can't reactivate them.

- **When you expire an attribute**: Any data assets that had the attribute applied continue to show that attribute.
- **When you expire an attribute group**: All the attributes within the group are also expired. Any data assets that had an attribute within this group applied to it continue to show that attribute.

To expire an attribute or attribute group:

1. On the **Data assets attributes** tab of the **Custom metadata (preview)** page, find the attribute or attribute group you want to expire and select the **Edit** pencil icon.
2. On the **Edit attribute** or **Edit attribute group** pane, select the **Mark as expired** checkbox.
3. Select **Update**.

Create data asset attributes using APIs

You can programmatically create and apply data asset attributes by using the business metadata APIs in Apache Atlas 2.2. For more information, see the Use Atlas 2.2 APIs tutorial.

Best practices for using data asset attributes

Use data asset attributes to extend the fields available for an asset in the Microsoft Purview Unified Catalog. Data asset attributes are key-value pairs that add structured metadata to your catalog. When Microsoft Purview scans data, it adds technical information about the data, such as data type and classification. If you want to add more fields, you need to define data asset attributes.

For example, a data asset attribute lets a user tag tables with the department that publishes them. Use data asset attributes when you want to make sure assets are always tagged in exactly the same way with this information. The user also wants to filter by the publisher field when they search for data.

The data asset attribute in this example helps users quickly find all data published by a team, but it doesn't help someone understand the definition of a publisher or what it means if the team is the publisher of the data. For any information that needs a business explanation, use glossary terms.

Search for data products in Unified Catalog

This article outlines how to search and browse Microsoft Purview Unified Catalog to find data products.

Prerequisites

To search for data products, you need a Catalog Reader role.

How to search

You can search for data products by using keywords or natural language (preview). To find data products that best fit your business needs, use **Search with natural language (preview)**. Searching with natural language supports data products created in English, French, and Spanish languages, and search terms in the corresponding languages.

The data product search is based on these attributes:

- Data product name, description, and business use.
- Name and description of associated governance domain, glossary terms, CDEs, and OKRs.
- Name and description of contained data assets.

To search for data products in Unified Catalog, select **Discovery**, then select **Data products**.

In the search bar, search by keywords, partial keywords, or natural language search:

- **Partial keyword match**: Enter a partial keyword, and the search returns results that match the partial term. For example, entering *cust* returns results that match *customer*.
- **Exact match with double quotes**: Use double quotes to search for an exact phrase. For example, "*customer profile*" searches for the phrase as a single unit, rather than matching *customer* and *profile* separately.
- **Natural language (preview)**: Select the magnifying glass icon at the end of the search bar, then enter a description of the data you need to find, then press **Enter**. Examples:
  - *I need three years of North American revenue data from finance department to analyze sales trends.*
  - *Find certified master data for customer accounts, which includes customer ID, name, and address.*

- *I'm working on improving customer satisfaction. Show me Power BI reports of customer rating and feedback.*

A list of search results appears. Select a data product to view its details page. Hover over a search result to see why it was returned. The highlight shows which attributes matched your search query, such as name, description, or term.

You can filter search results by governance domain, data product type, owner, glossary terms, or critical data elements.

> **Note**
>
> Search isn't case sensitive, but spaces within search terms matter. Underscores and other non-space characters are treated as part of the keyword.

Explore data products by governance domain

You can browse data products in Unified Catalog by governance domain.

1. In the Microsoft Purview portal, go to **Discovery** > **Data products**.
2. Select a governed domain under **Explore data products by governance domain**, or select **Explore all** to navigate to a governance domain.
3. When you select a governance domain, a list of its data products appears. Select a data product name from the list to view that product's details page.

Explore data products by filtering on custom attributes

When you explore data products by governance domains, you can filter your view by business concept attributes.

1. In Unified Catalog, go to **Discovery** > **Data products**.
2. Select **Explore all**. You arrive at the **Explore data products by governance domain** page.
3. Select **Add filter**.
4. At **Filter**, select an attribute name from the dropdown list.
5. At **Operator**, select a condition, such as **Equals** or **Starts with**, which varies based on the kinds of values allowed by the attribute.
6. At **Value**, enter your value.
7. Select **Apply**.

View data product details

A data product details page contains a description, use case, terms of use, and contacts. The page lists the glossary terms, critical data elements (CDE), and data assets that are associated to the data product. Select an asset, term, or CDE to display a flyout with a **View details** button for further details.

View data asset details

Select a data asset from the data product details page to view a flyout pane with more information. Select **View details** on the flyout pane to open the details page for the data asset.

The data asset details page includes a description and owner, lists the data asset columns, and displays lineage information. The page also lists other data products that contain the data asset.

Next steps

- Request access to data product
- Manage Unified Catalog access policies

Related articles

- Microsoft Purview Unified Catalog Responsible AI FAQ

ⓘ **Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 05/21/2026

Search for data assets

After data is scanned and ingested into the Microsoft Purview Data Map, data consumers need to easily find the data they need for their analytics or governance workloads. Data discovery can be time consuming because you might not know where to find the data that you want. Even after finding the data, you might have doubts about whether you can trust the data and take a dependency on it.

This article outlines how to search Microsoft Purview Unified Catalog to quickly find the data you're looking for. You can search across all data assets, or focus on governed assets that have been curated with governance metadata such as glossary terms, data products, critical data elements, or data quality scores.

Permissions to search the Unified Catalog

Searching Unified Catalog only returns relevant data assets that you have permissions to view.

You can find a data asset in Unified Catalog when:

- You have at least read permissions on an available Azure or Microsoft Fabric resource.
- You have at least data reader permissions on a domain or collection in the Microsoft Purview Data Map where the asset is stored.

Manage permissions to these assets at the resource level and at the Microsoft Purview Data Map level, respectively. For more information on providing this access, see the links above.

To search for curated, ready-to-use data assets, search for data products your organization has developed.

> **Tip**
>
> If your Unified Catalog is well-curated, day-to-day business users shouldn't need to search the full catalog. They should be able to find data they need in data products, or by using the governed assets search to find curated data. For more information about setting up Unified Catalog, see Get started with data governance, and plan for Unified Catalog.

Permissions to search governed assets

Governed assets are visible to users with the Global Catalog Reader role. Local Catalog Readers can view governed assets in the governance domain they have access to. Published governance concepts, such as glossary terms, data products, and critical data elements, are also visible to catalog readers. Draft concepts are visible only to Data Product Owners and Data Stewards.

Existing Data Map reader and curator roles aren't changed by the governed assets search experience.

Searching the catalog

To search for data assets in Unified Catalog:

1. In the Microsoft Purview portal, open **Unified Catalog**.
2. Under **Discovery**, select **Data assets**.
3. In the search bar, enter keywords to narrow your search, such as name, data type, classifications, and glossary terms. Search suggestions appear as you enter keywords. Your search history is also presented so you can resume a previous data exploration.
4. To complete your search, select **View search results** or select **Enter**. For full search syntax information, see keyword search syntax.

The search results page shows a list of data assets and glossary terms that match the search terms. The keyword is highlighted in the return results, so you can see where the term was found in the asset. The **Governed assets** tab displays only governed assets search results, and the **All assets** tab displays the search results in all assets in your data estate, whether they're governed or not.

> **Note**
>
> Search only returns items in collections where you're a data reader or curator. For more information, see roles and permissions in collections.

The Microsoft Purview relevance engine sorts through all the matches and ranks them based on what it believes is useful to a user. For example, a data consumer is likely more interested in a table curated by a data steward that matches on multiple keywords than an unannotated folder. Many factors determine an asset's relevance score. The Microsoft Purview search team constantly tunes the relevance engine to ensure the top search results have value to you.

Keywords and keyword search syntax

Search queries consist of keywords and operators such as `AND`, `OR`, `NOT`, exact phrase quotes, field-scoped search, and parentheses for grouping. You can combine these operators to build precise queries that return exactly the assets you need.

For the full list of supported operators, field names, and detailed syntax guidance, see Keyword search syntax.

Filter results

If the top results don't include the assets you're looking for, you can filter your results by using the filter pane to the left of the search results.

On the search results page, select **Filter** to open and close the filter pane. With the filter pane open, select a filter category, then select the checkbox next to the values you want to filter on. For some filters, you can select the ellipsis to choose **Or** or **And**.

Available filters

You can add filters by selecting **Add filter** on the filter pane and selecting one of the following options:

- **Activity**: Refine your search to attributes created or updated within a certain timeframe.
- **Asset type**: Refine your search to specified asset types. For example: dashboards, files, glossary terms, or metamodel assets.
- **Assigned term**: Refine your search to assets with the selected terms applied.
- **Classification**: Refine your search to assets with certain classifications.
- **Collection**: Refine your search by assets in a specific collection.
- **Contact**: Refine your search to assets that have selected users listed as a contact.
- **Data source type**: Refine your search to assets from specified source types. For example: Azure Blob Storage or Power BI.
- **Endorsement**: Refine your search to assets with specified endorsements, like **Certified** or **Promoted**.
- **Label**: Refine your search to assets with specific security labels.
- **Data asset attributes**: Refine your search to assets with specified data asset attributes. Attributes are listed under their attribute group, and use operators to help search for specific values. For example: Equals or Doesn't equal.
- **Rating**: Refine your search to only data assets with a specified rating.
- **Tags**: Refine your search to assets with selected tags.

Search for governed assets (preview)

The governed assets search provides a focused view of data assets that are actively governed in Unified Catalog. Instead of searching across all assets, you can search only for assets that have governance metadata, such as glossary terms, data products, critical data elements, or data quality scores. Learn more about governed assets.

The governed assets search is helpful because organizations have large data estates but only a small percentage of assets are curated and ready for consumption. Data consumers can use this search to find meaningful, governed data without navigating the entire data estate.

To search for governed assets:

1. In the Microsoft Purview portal, open **Unified Catalog**.
2. Under **Discovery**, select **Data assets**.
3. Enter keywords in the search bar and select **Enter** or **View search results**.
4. On the search results page, select the **Governed assets** tab.

The search results show only assets that are governed in Unified Catalog. Each asset has a **Governed** label next to its name. Governed assets search supports semantic search, so you can use natural language queries to find assets.

Search capabilities

Governed assets search includes the following capabilities:

- **Partial keyword match**: Enter a partial keyword, and the search returns results that match the partial term. For example, entering *cust* returns results that match *customer*.
- **Exact match with double quotes**: Use double quotes to search for an exact phrase. For example, "*customer profile*" searches for the phrase as a single unit, rather than matching *customer* and *profile* separately.
- **Search highlights**: Hover over a search result to see why it was returned. The highlight shows which attributes matched your search query, such as name, description, or term.

> **Note**
>
> Search isn't case sensitive, but spaces within search terms matter. Underscores and other non-space characters are treated as part of the keyword.

Filter governed asset results

You can filter governed asset search results by using the following filters (see filter definitions):

- Asset type
- Data source type
- Classification
- Contact
- Glossary term
- Data product
- Critical data element

> **Note**
>
> The **Terms** filter in governed assets search uses Unified Catalog terms. After you migrate glossary terms and enable data asset curation, the **Assigned term** filter is no longer available in the all data assets search, and classic terms no longer appear on asset details pages.

View governed asset details

Select a governed asset from the search results to view its details. On the asset details page, select the **Governance** tab to see how the asset is governed. The **Governance** tab shows:

- Data products the asset is part of.
- Glossary terms linked to the asset.
- Critical data elements associated with columns in the asset.
- The latest data quality score, if a data quality run was performed on the asset.

The **Governance** tab follows existing permission models. If you have a Catalog Reader role, you see only published governance concepts. Draft concepts, such as draft glossary terms, are visible only to Data Product Owners and Data Stewards.

If an asset is deleted from Data Map, a warning icon appears next to the asset in search results. The asset continues to appear until the Data Product Owner removes it.

> **Note**

If an asset is ungoverned (all links to data products, glossary terms, and critical data elements are removed), the asset currently still appears in governed assets search results.

Bulk edit search results

If you want to make changes to multiple assets returned by search, you can modify glossary terms, classifications, and contacts in bulk. For more information, see How to bulk edit assets.

View assets

From the search results page, select an asset to view details, such as description, attributes, schema, lineage, and classifications.

Browse the Unified Catalog

Unified Catalog offers a browse experience that enables users to explore what data is available to them either by collection or by exploring the hierarchy of each data source in the catalog.

In Unified Catalog, go to **Discovery** > **Data assets**. Under **Explore your data**, select a card for how you want to explore your data: in Azure, in Microsoft Fabric, by source type, or by collection.

You can only see the data sources you have permission to view.

Searching Microsoft Purview in connected services

After you register your Microsoft Purview instance to an Azure Data Factory or an Azure Synapse Analytics workspace, you can search Unified Catalog directly from those services. For more information, see Discover data in ADF using Microsoft Purview and Discover data in Synapse using Microsoft Purview.

Link glossary terms to assets and columns (preview)

You can add business context to your data by linking glossary terms to data assets and columns.

Users need two roles in order to attach terms to assets and columns:

1. **Data Curator**: All users need to hold this role in the collection where the asset belongs. Only Data Curators can edit data assets. Learn more about this role and collections permissions in Data Map.

2. In addition, users need either one of these roles:
    a. **Data Steward**: Users can be a Data Steward in any governance domian. These users can see published and unpublished glossary terms, and can attach both kinds of terms to assets and columns.
    b. **Global Asset Curator**: This role allows users to attach published glossary terms to assets and columns. Users with this role can only see published terms and attach published terms to assets and columns.

See Data governance roles and permissions for role descriptions and how to assign roles.

Link terms to a data asset:

1. Search for an asset or column in Unified Catalog.
2. Select the data asset to open its details page.
3. Select **Edit**.
4. On the **Overview** page, at **Glossary terms**, select **Add term**.
5. Search for a term by entering keywords, or select the checkbox next to a term listed on the page. You can select multiple terms.
6. Select **Add**.
7. Select **Save**.

The terms are attached to the asset.

Link terms to a column within an asset:

1. Search for an asset or column in Unified Catalog.
2. Select the data asset to open its details page.
3. Select **Edit**.
4. On the **Schema** page, find the row for the column you want to add a term to, and select **Add term**.
5. Search for a term by entering keywords, or select the checkbox next to a term listed on the page. You can select multiple terms.
6. Select **Add**.
7. Select **Save**.

The terms are attached to the column.

ⓘ **Note:** The author created this article with assistance from AI. Learn more

Last updated on 05/21/2026

Keyword search syntax

When you search for data assets, you can use keywords, operators, and field-scoped queries to find exactly the data you need. This article describes the supported query syntax so you can build precise, effective searches.

For a general overview of searching and browsing for data assets, see Search for data assets in Unified Catalog.

> **Note**
>
> The syntax described in this article applies to the **Unified Catalog** > **Discovery** > **Data assets** search experience in the Microsoft Purview portal. Search results are filtered based on your permissions, so you only see assets you have access to. For more information about permissions, see Search for data assets in Unified Catalog.

Quick reference

The following table summarizes the query syntax options available in Unified Catalog search.

Expand table

Syntax	Example	Description
Keyword	`customer`	Search all fields for the keyword.
Multiple keywords	`customer sales`	Search for assets matching either word. Assets matching both words are ranked higher.
AND	`customer AND sales`	Both terms must be present.
OR	`customer OR sales`	Either term can be present.
NOT	`customer NOT draft`	Must contain first term, must not contain second.
Exact phrase	`"sales report"`	Match the exact phrase in order.
Grouping	`(A OR B) AND C`	Control evaluation order.

Syntax	Example	Description
Field search	`name:customer`	Search only within a specific field.
Match all	`*` or empty	Return all assets.

Simple keyword search

A single keyword searches across all searchable fields in the catalog, including asset name, description, qualified name, schema columns, glossary terms, contacts, classifications, custom attributes, collections, domains, and sensitivity labels.

The following example returns any asset where `customer` appears in any searchable field:

```
customer
```

Results are ranked by relevance. An exact match on the asset name scores higher than a partial match in the description.

Multiple keywords

When you enter multiple words separated by spaces without an explicit operator, the search returns assets that contain *any* of the words. Assets that contain more of the keywords are ranked higher in the results.

The following example finds assets that contain `customer`, `sales`, or both:

```
customer sales
```

> **Important**
>
> A space-separated keyword search is **not** the same as an AND search. Consider the following differences:
>
> Expand table

Query	Behavior	Example result
`customer sales`	Matches `customer` OR `sales`. Assets with both score higher.	An asset named "Sales Dashboard" matches even without "customer."
`customer AND sales`	**Both** `customer` AND `sales` must be present.	"Sales Dashboard" doesn't match unless it also contains "customer."
`"customer sales"`	The exact phrase `customer sales` must appear in that order.	Only matches if `customer` is immediately followed by `sales`.

Boolean operators

Boolean operators let you combine search terms with explicit logic. Operators must be written in uppercase (`AND`, `OR`, `NOT`).

AND operator

Use the AND operator when all terms must be present in the matching assets. The terms don't need to be adjacent or in the same field.

```
customer AND sales
azure AND sql AND database
```

OR operator

Use the OR operator when any of the terms can be present.

```
customer OR client
sales OR marketing OR operations
```

NOT operator

Use the NOT operator to exclude assets that contain a specific term.

```
customer NOT draft
```

```
sales NOT internal NOT test
```

The following example returns assets that contain `customer` but don't contain `draft`:

```
customer NOT draft
```

Combine operators

You can chain operators in a single query:

```
customer AND sales NOT draft
```

This query returns assets that contain `customer` and `sales`, but don't contain `draft`.

Exact phrase search

Wrap your query in double quotes to search for an exact phrase. The words must appear in the specified order and adjacent to each other.

```
"sales report"
"customer data analysis"
```

The following table shows how phrase search differs from a keyword search:

Expand table

Query	Matches	Doesn't match
`"sales report"`	"Q1 **sales report** summary"	"report on sales" (wrong order)
`"sales report"`	"Annual **sales report**"	"sales quarterly report" (words not adjacent)

Use `sales report` (no quotes) when you want any asset that mentions `sales` or `report`. Use `"sales report"` (with quotes) when you want only assets that contain the exact phrase.

Grouping with parentheses

Parentheses control the order of evaluation, just like in math expressions.

```
(customer OR client) AND sales
```

Without parentheses, operator precedence can produce unexpected results. The following examples show how parentheses change the query interpretation:

Expand table

Query	Interpretation
`(customer OR client) AND sales`	Assets that contain `customer` or `client`, and also contain `sales`.
`customer OR (client AND sales)`	Assets that contain `customer`, or assets that contain both `client` and `sales`.

You can nest parentheses for more complex queries:

```
(A AND B) OR (C AND D) NOT E
((alpha OR beta) AND gamma) NOT delta
```

Field-scoped search

By default, a keyword is searched across all fields. Field-scoped search lets you restrict the search to a specific field.

Use the syntax `fieldName:value` or `fieldName:"phrase value"`:

```
name:customer
classification:Confidential
entityType:azure_sql_table
contact:alice@company.com
name:"sales report"
```

Supported fields

The following fields can be used with the `field:value` syntax:

Expand table

Field name	Description	Example
`name`	Asset name	`name:customer`
`qualifiedName`	Fully qualified name	`qualifiedName:sales_db.customers`
`description`	Asset description	`description:quarterly`
`userDescription`	User-provided description	`userDescription:important`
`displayText`	Display text	`displayText:customer`
`entityType`	Entity type	`entityType:azure_sql_table`
`assetType`	Asset category	`assetType:Tables`
`classification`	Sensitivity labels or tags	`classification:Confidential`
`endorsement`	Endorsement status	`endorsement:Certified`
`glossaryType`	Glossary type	`glossaryType:AtlasGlossaryTerm`
`termStatus`	Term status	`termStatus:Approved`
`termTemplate`	Term template	`termTemplate:BusinessTerm`
`glossary`	Glossary name	`glossary:Finance`
`fileExtension`	File extension	`fileExtension:csv`
`term`	Assigned glossary terms	`term:Revenue`
`contact`	Owner or expert contact	`contact:john@company.com`
`abbreviation`	Term abbreviation	`abbreviation:ROI`
`objectType`	Object type	`objectType:Tables`
`tag`	Tags	`tag:important`

Field names are case-insensitive. For example, `Name:customer` and `name:customer` produce the same results.

If you use a field name that isn't in the Supported fields table, the search engine doesn't return an error. Instead, it treats the entire expression as an unscoped keyword. For example, `alice AND report:myreport` searches for `alice` AND the literal text `report:myreport` across all fields, rather than restricting `myreport` to a field called `report`.

> **Important**
>
> Only one value token is captured after the colon. To search for a multiword value, use quotes:
>
> Expand table
>
> | Query | Behavior |
> | --- | --- |
> | `name:sales report` | Searches `name` for `sales`, then separately searches all fields for `report`. |
> | `name:"sales report"` | Searches `name` for the exact phrase `sales report`. |

Nested fields

Some supported fields, such as `contact` and `term`, represent composite data with multiple properties. For example, a contact has a display name, an email address, and an ID.

You can't use dot notation to target a specific property. The system only recognizes the top-level field names listed in the Supported fields table.

Expand table

Query	Supported	Behavior
`contact:Alice`	Yes	Searches across all contact properties (display name, email, ID).
`contact.displayName:Alice`	No	Not recognized as a field-scoped search.
`term:Revenue`	Yes	Searches across all term properties (name, glossary).
`term.name:Revenue`	No	Not recognized as a field-scoped search.

Use the top-level field name. For example, `contact:Alice` searches across all contact properties and returns assets where any property matches.

Combine keywords with field-scoped search

When you mix a plain keyword with a field-scoped search, a space doesn't imply AND. Space-separated clauses are treated as OR.

Expand table

Query	Behavior
`alice contact:bob`	`alice` OR contact contains `bob`.
`alice AND contact:bob`	`alice` AND contact contains `bob`.

> **Tip**
>
> Always use explicit `AND` or `OR` operators when combining keyword search with field-scoped search.

Match all results

An empty query or `*` returns all assets, subject to your permissions and any applied filters.

```
*
```

This option is useful as a starting point when you want to browse assets by using only filters and facets.

Case sensitivity

Search is generally case-insensitive. Searching for `Customer`, `customer`, or `CUSTOMER` returns the same results.

However, the casing and special characters in your search query can affect the number of results returned. Certain fields are processed by the search engine, which splits terms at specific boundaries before matching.

Token-splitting delimiters

The following boundaries cause a term to be split into separate tokens during both indexing and searching:

Expand table

Delimiter	Example input	Tokens produced
Case change (lowercase to uppercase)	`CustomerOrder`	`customer`, `order`, `customerorder`

Delimiter	Example input	Tokens produced
Underscore (`_`)	`customer_order`	`customer`, `order`, `customer_order`
Hyphen (`-`)	`customer-order`	`customer`, `order`, `customer-order`
Dot (`.`)	`customer.order`	`customer`, `order`, `customer.order`
Comma (`,`)	`item1,item2`	`item1`, `item2`
Colon (`:`)	`db:schema`	`db`, `schema`
Whitespace	`customer order`	`customer`, `order`
Letter-to-number transition	`version2release`	`version`, `2`, `release`
Number-to-letter transition	`2ndEdition`	`2`, `nd`, `edition`

> **Note**
>
> Splitting is cumulative. A name like `Customer_Order_Details` is split at every underscore and at every case-change boundary within each segment, producing tokens such as `customer`, `order`, `details`, `customer_order`, and `order_details`.

How casing affects results

When your search query contains mixed case (such as PascalCase or camelCase), the search engine splits it at case-change boundaries, producing multiple tokens. Each token is independently matched, which broadens the result set.

When your query is all lowercase, there are no case-change boundaries to split on, so it stays as a single token and must match as a contiguous string.

The following example illustrates the difference:

Expand table

Search query	Behavior	Effect on results
`CustomerOrder`	Split into tokens: `customer`, `order`, `customerorder`	Matches assets containing `customer` or `order` or `customerorder`. Broader results.
`customerorder`	Stays as one token: `customerorder`	Matches only assets containing the contiguous string `customerorder`. Narrower results.

Phrase search and casing

Quoted phrase search also interacts with casing. With `"CustomerOrder"` in quotes, the query tokens are still `customer`, `order`, `customerorder`, but because it's a phrase search, the tokens must appear adjacent and in order.

Expand table

Phrase query	Matches	Doesn't match
`"CustomerOrder"`	`CustomerOrder`, `FAC_Customer_OrderLine` (where `customer` and `order` are adjacent tokens)	`legacy customer database` (no adjacent `order` token)
`"customerorder"`	`CustomerOrder` (indexed as contiguous `customerorder`)	`FAC_Customer_OrderLine` (no contiguous `customerorder` token)

Recommendations for casing

- For the broadest results, use the natural casing of the term (for example, `CustomerOrder` or `SalesReport`). The case-change splitting finds partial matches.
- For the most precise results, use all lowercase (for example, `customerorder`) or use an exact phrase search with quotes.
- When in doubt, search for both forms or use explicit Boolean operators: `CustomerOrder OR customerorder`.

Result ranking

When a search returns multiple results, they're ordered by a relevance score. The following factors influence ranking, from highest to lowest impact. The exact ranking logic is subject to change as the relevance engine is continually tuned.

Where the match occurs

Not all fields carry equal weight. A match in the asset name is considered more relevant than a match in the description or other fields:

Expand table

Priority	Match type	Example
Highest	Exact match on asset name	Searching `customers` when the asset is named exactly `customers`.

Priority	Match type	Example
High	Case-insensitive exact match on name	Searching `Customers` when the asset is named `customers`.
Medium	Partial match on name or qualified name	Searching `customer` when the asset is named `CustomerOrderTable`.
Standard	Match in description, display text, classification, entity type, or other fields	Searching `customer` when the word appears in the asset description.
Lower	Match in schema columns, glossary terms, contacts, or custom attributes	Searching `customer` when a column named `customer_id` exists on the asset.

Within each level, assets matching more of the search tokens score higher than assets matching fewer.

Asset type

Leaf-level assets, such as individual tables, files, and glossary terms, receive a ranking boost over container assets like databases or storage accounts. When you search for `sales`, a table named `sales_data` typically ranks above the database that contains it.

Data completeness

Assets that are more fully curated receive a ranking boost. Each of the following signals independently contributes to a higher score:

- Has a description
- Has one or more classifications (sensitivity labels)
- Has one or more glossary terms assigned
- Has owners or experts defined
- Has schema or column information
- Has custom attributes populated
- Has sensitivity labels applied

Endorsement status

Endorsed assets rank higher than nonendorsed ones. **Promoted** assets receive a moderate boost, and **Certified** assets receive a boost that stacks with Promoted when both apply.

Freshness

Recently updated assets receive a slight ranking boost over stale ones. The boost decays gradually.

Popularity signals

Usage-based signals contribute to ranking, including modification frequency, data quality score, and user rating.

Personalization

When you perform a search, assets where you're listed as an owner or expert receive a small boost, making your own assets easier to find.

Smart keyword detection

The search system recognizes certain keywords and adjusts ranking:

- `my`: Boosts assets where you're an owner, expert, or recent editor.
- **Object type keywords**: Searching for a recognized asset category (such as `Tables`, `Files`, or `Folders`) boosts assets of that type.
- **Classification keywords**: Searching for a term that matches a known classification name (such as `PII` or `Confidential`) boosts assets with that classification.

> **Note**
>
> When an explicit sort order is applied (for example, sort by name or update time), relevance scoring is only used as a tiebreaker when two assets have the same sort value.

Query limits

Expand table

Limit	Value
Maximum keyword length	1,024 characters. Longer queries are truncated.

Known limitations

Grouping isn't supported within a field search. For example, `name:(alice AND bob)` is invalid syntax. Use `name:alice AND name:bob` instead.

Resources

- Search for data assets in Unified Catalog
- Discovery Query API reference

ⓘ **Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 04/24/2026

Enterprise glossary (preview)

10/10/2025

This article helps you understand the Enterprise glossary (preview) area of Microsoft Purview Unified Catalog, where you can browse and search for published business concepts, such as governance domains, glossary terms, critical data elements, and objectives and key results (OKRs).

As a consumer of the catalog, you can discover data products through the search experience and through the enterprise glossary experience that links different business concepts to data products. This experience has three tabs for glossary terms, critical data elements, and OKRs, and gives you a consolidated view of all the published concepts within. Once you find the concept you're looking for, you can view the related data products and request access to the ones you need.

Prerequisites

To browse the enterprise glossary, you need a catalog reader role.

Search for business concepts in the catalog

You can browse and search for published business concepts within Unified Catalog.

1. In the Microsoft Purview portal, open **Unified Catalog**.
2. Under **Discovery**, select **Enterprise glossary**.
3. Select the **Governance domains**, **Glossary terms**, **Critical data elements**, or **OKRs** tab to learn more about published concepts under each concept.

> **Note**
>
> - You can visualize any parent/child relationships in a tree-like hierarchy within the **Glossary terms** tab by selecting **Tree** in the views drop-down.
> - The list view displays all published glossary terms but the tree visualization doesn't display terms with unpublished parents.

Search for any concept by name in the search bar. You can sort the results in ascending or descending alphabetical order, and filter results by governance domain, owner, or any other available filter.

Select a published concept to view its related concepts, such as data products.

Filter your view by custom attributes

Use business concept attributes as a filter when browsing and searching for business concepts.

1. In the Enterprise glossary, select the tab for the concept you want to search. If you're searching in a governance domain, select **View all** on a business concept card.
2. Select **Add filter**.
3. At **Filter**, select an attribute name from the dropdown list.
4. At **Operator**, select a condition, such as **Equals** or **Starts with**, which varies based on the kinds of values allowed by the attribute.
5. At **Value**, enter your value.
6. Select **Apply**.

Related content

Request access to linked data products

Manage data product access policies

> **Note**
>
> This article explains how to set access policies on business concepts including governance domains, data products, critical data elements, and glossary terms, and how to manage access requests. If you need to request access for yourself, see Request access to data products.

Access to a data product in Microsoft Purview Unified Catalog grants users permissions to the data product and permissions to the data assets within the data product. This article explains how you can manage access to your data products and set up a system to provide access to users who request it.

This article explains:

- How to set up a data product access policy.
- How to manage and approve access to data products.

Permissions

- To view and manage policies at the governance domain level, you need governance domain owner permissions.
- To view and manage policies at the data product level, you need data product owner permissions.
- To view and manage policies in critical data elements or glossary terms, you need data steward permissions.
- Unified Catalog readers can view and request access to data products.
- Managers and privacy approvers of access requests also need to have a minimum of a catalog reader to be able to use the Unified Catalog and approve requests as part of the tiered approval.

Considerations

In this experience, request approvers must provide access to the individual data assets manually as part of the approval process, or specify an access provider to do so.

The product doesn't enforce policies such as attestations (including the no copy attestation). The data consumer attests that they'll follow these policies while requesting access.

Skip the access policies and workflow setup

If you use an external solution for access management or don't want to have access requests on a data product, you can disable this feature during data product curation. When you disable access management on the data product, no access policies from business concepts, workflows, or the request access button apply to that data product.

To disable the access process on a data product:

1. Select the data product in a draft state.
2. Select the **Manage policies** tab.
3. Scroll down to the workflows section of the tab to the checkbox **Disable Access Management** and make sure it's checked.

Set up data product access policies

To build access policies, you usually need data product owner or data steward permissions.

> **Tip**
>
> Your data product needs to be in an unpublished state to manage access policies.

1. In the Microsoft Purview portal, open Unified Catalog.
2. Under **Catalog management**, select **Data products**, then select a data product.
3. On the data product's details page, select **Manage policies**. Here you can view and edit the default values assigned to the default set of data product access policies. The selected values affect what the data consumers see on their access request form and actions they need to take.
4. Under **Permitted access**, add your usage purposes, which are the authorized purposes for accessing and using the data product. Three purposes are provided by default. You can add more purposes and their descriptions, which the consumer can select in the request access form.
5. At **Approval requirements**, determine if manager approval or a privacy and compliance review is required.

   > **Note**

If you select these options, the access request form shows that a manager approver or a privacy and compliance review is required. The consumer's manager in Microsoft Entra ID is notified for the first tier of approval. The privacy reviewer named by the consumer in the request form is notified for their approval. The request is first reviewed by manager, followed by the privacy reviewer, and then by the access request approvers for approval and granting of access. The optional access provider tier is last, if selected. The request status is final only when all configured levels of approval are complete.

6. Under the **Approval requirements** dropdown, in **Access request approvers**, select the users that need to approve the access request. The first approver to take action approves the request and proceeds to grant access to the data assets manually, records the status of access provisioned at each asset, and specifies data access instructions or comments to the data consumer. By default, data product owners are populated and you can add more approvers. You can also add Microsoft Entra ID groups or security groups as approvers. Approvers can see detailed status in the request view.

7. An optional tier called **Access provider** can be set explicitly to grant access to the data assets manually, record the status of access provisioned at each asset, and specify data access instructions or comments to the data consumer. This practice ensures the data consumer is aware of the next steps they can take to access and use the data.

   > ⓘ **Note**
   >
   > You can select a role group to grant approver status to all the users in a group at once. However, users who belong to nested groups within the role group aren't supported at this time.

8. Under **Attestations**, determine if copies of the data are permitted. The attestations are reflected on the access request form for the consumer to attest to.

9. If terms of use are present on a data product, those terms are also reflected on the access request form for the consumer to attest to.

10. Add any more attestations by selecting **Add attestation** and adding a display name and the file location. These attestations are reflected on the request access form for the consumer to attest to.

11. If there are policies inherited from the governance domain, critical data element, or glossary term, you can see them in the **Inherited policies** tab. See the following section in this document for details: Policies on governance domains and glossary terms (inherited policies).

12. Select **Preview request form** to see what users see when they request access.

13. Select **Save changes** to save the access policy for the governance domain.

Policies on governance domains, glossary terms, and critical data elements (inherited policies)

You can set policies on governance domains, glossary terms, and critical data elements. Data products in the governance domain, or data products that have glossary terms or critical data elements applied, inherit and aggregate these policies.

You can see the inherited policies in the data product manage policies view, in a separate tab called **Inherited policies**. You can see the aggregated view by selecting the **Preview** button. Your data consumers see this aggregated view when they request access.

For example, if you set a manager approval policy on a glossary term applied to the data product, the data product also requires manager approval. Any attestation set on a business concept (governance domain, data product, glossary term, or critical data element) is aggregated on the data product, and the data consumer attests to all required attestations when requesting access.

To build access policies on governance domains, glossary terms, or critical data elements, you need governance domain owner or data steward permissions.

1. In the Microsoft Purview portal, open Unified Catalog.
2. Select the **Catalog management** dropdown and select **Governance domains**.
3. Select a governance domain.
4. On the governance domain page, select **Manage policies**.
5. Alternately on a glossary term or critical data element page, select **Manage policies**.
6. From the Policy configuration window, you're able to create and manage relevant access policies.

Configure inherited policies

In the **Manage policies** window, you can view and manage the default set of access policies on each business concept. The selected values affect what the data consumers see on their access request form and actions they need to take.

1. Determine if manager approval is required. The data consumer's manager, configured in Microsoft Entra ID, is notified as first tier of approval. If the request is approved, the request proceeds to the next tier.

2. Determine if copies of the data are permitted. This selection is reflected on the access request form for the consumer to attest to.

3. Add any more attestations you want by selecting **Add attestation** and adding a display name and the file location. These attestations are reflected on the request access form for the consumer to attest to.

View and manage access requests

> **Important**
>
> The **Requests** page in **Catalog management** shows only the access requests from the existing published data products before the new workflow feature became available on 11/12/2025. All access requests and other approvals moving forward are in the **Requests and approvals** tab under the **Waiting for my response** screen. The process and approvals are the same.

To view access requests in Unified Catalog, follow these steps. When you select an access request, the **Access Request** flyout pane opens. It contains details about the request and its status. You see only requests for which you're an approver.

1. In Unified Catalog, select **Catalog management**, then select **Requests**.

2. In the status column of the governance domain table, sort by any domains that have open access requests.

3. Select the governance domain that you want to manage access requests for.

4. On the access requests tab, you see a list of the most recent access requests.

5. Select the access request you want to respond to.

6. View the details submitted by the consumer.

7. Select **Approve** or **Deny**.

8. If you're the last approver in the sequence of approvers (a data product access request approver or an explicit access provider), record the status of access provisioned at each asset and specify data access instructions or comments to the data consumer. This practice ensures the data consumer is aware of the next steps they can take to access and use the data.

9. The data consumer is notified once you respond to the request.

10. The requestor is notified via email and can also view the status on the Microsoft Purview Unified Catalog **Discovery** drop-down, **Data products** page, in the **My data access tab**.

Request statuses

- **Pending**: The request is sent to an approver for review.
- **Declined**: The request is denied by the approver.
- **Approved**: The request is approved, but provisioning work isn't yet completed to grant access to the requestor.
- **Completed**: The data asset is provisioned to grant access for the requestor. This request is now fully complete.

Respond to access requests through email notification

1. In the email you receive as a request to approve an access request, select the **Approval** request link.

2. You're brought to the request detail view in **Requests** page in **Catalog management** .

3. View the details submitted by the consumer.

4. Select **Approve** or **Deny**.

5. If you're the last approver in the sequence of approvers, a data product access request approver, or an explicit access provider, record the status of access provisioned at each asset and specify data access instructions or comments to the data consumer. This step ensures the data consumer is aware of the next steps they can take to access and use the data.

6. If you're the last approver in the sequence of approvers, and there are multiple data assets in the data product, you can record the status and save intermittently and mark the request as complete only when you're done with all the data assets.

7. The data consumer is notified once the request is responded to and in its final state.

8. The requestor is notified via email and can also view the status on the Microsoft Purview Unified Catalog **Data product search** page, in the **My data access** tab.

Sequential or tiered approvals and request statuses

As noted in the section to manage policies on data product, depending on the approver selections you make, a sequential or tiered approval is in effect. The following sequence is maintained for data product access request approvals.

1. If you select manager approval, the consumer's manager in Microsoft Entra is notified for the first tier of approval.
2. Only after the manager approves the request, the privacy reviewer if selected is notified for their approval or can take action.
3. Once approved by the privacy reviewer, the access request approver is notified for approval and provisioning of access to the data assets in the data product. This tier approves and completes the request if there's no access provider specified, else they only approve.
4. If there's a last tier of access provider specified, then that tier provisions access to the data assets and records the status and instructions for data consumer. They complete the request. Complete is the final status.
5. The workflow proceeds until the last approval and completion occurs or the first denial occurs.
6. The data consumer is only notified after all approvers take their respective actions.
7. At any point the data consumer can see the status of the request in the Microsoft Purview Unified Catalog **Discovery** drop-down, **Data products** page, in the **My data access** tab.
8. A request status can go from pending to declined or pending to approved and then to completed.

Delete access requests

To revoke a user's access to a data product and the data assets within the data product, the request approver needs to complete both of the steps listed below.

> **Note**
>
> The only request approver who can remove access to a data product and delete an access request is the owner of that same data product. This data owner must also hold the Data Product Owner role.

1. Return to the individual data asset (such as in Azure Data Lake, Microsoft Fabric, or SQL) and manually remove the provisioning for the user.

2. Delete the request in Unified Catalog by following these steps:
    a. In Unified Catalog, select **Catalog management**, then select **Requests**.
    b. Select a governance domain, then select the name of the request you want to delete.
    c. On the **Access Request** flyout pane, select **Delete**.

If you remove the provisioning on the asset but don't delete the access request in Unified Catalog, the data product shows that the user's access is removed, but the user still has access to the underlying data assets.

Last updated on 11/19/2025

Request access to data products

10/10/2025

> **Note**
>
> This article explains how to request access to a data product. To manage access to data products, see Manage data products access policies.

Follow the steps in this article to request access to a data product and its assets in Microsoft Purview Unified Catalog.

Considerations

In this experience, the approvers of the request must provide access to the individual data assets manually as part of the approval process or specify an access provider to do so and record the status of access provisioned.

When you receive confirmation that your request is approved or see the status change on the Unified Catalog **Data product search** page in the **My data access** tab, you can view the details on how to access and use the data. Learn how to access and use data products.

How to request access to a data product

If you discover a data product in the catalog that you want to access, you can request access directly through Microsoft Purview. The request triggers a workflow that asks the owners of the data resource to grant you access to the data product.

This article outlines how to make an access request.

1. In Unified Catalog, select **Discovery**, then select **Data products**.
2. To find a data product, search or browse for it by entering a keyword or natural language in the search bar.
3. Select the data product to open the data product details page.
4. Select **Request access**.
5. Fill in the form. The form is different depending on your organization and the data product, but provide all information.
6. When you're finished, select **Send**.

Approval requests are sent to all involved parties. When your request is approved or rejected, you get an email notification.

How to view data product access requests

You can see all your requests in Unified Catalog by selecting **Discovery**, then **Data products**, and selecting the **My data access** tab.

Select a data product name to open the **Access Request** flyout pane. This detailed request view helps you understand the approvals or actions taken, and view the asset level access provisioning status and any instructions, if provided.

Request statuses

- **Pending**: The approver receives the request for review.
- **Declined**: The approver denies the request.
- **Approved**: The approver grants the request but hasn't finished provisioning access for the requestor.
- **Completed**: The data asset is provisioned, and the request is fully complete.

How to cancel an access request

You can cancel an access request if it isn't in the **Completed** status. To cancel an access request:

1. In Unified Catalog, select **Discovery**, then **Data products**, then the **My data access** tab.
2. On the **Access Request** flyout pane, select **Delete**.

If your request is approved and completed, only the data product owner can delete it. The data owner must manually deprovision access to individual assets before deleting a request.

Create workflows to automate processes in Unified Catalog

Workflows in Microsoft Purview Unified Catalog provide a centralized, automated way to manage approval scenarios, such as granting access to data products and publishing data products and glossary terms.

Workflows are automated, repeatable sets of actions or business processes that users can establish to streamline operations in their organization. The available workflows can automate the approval processes for requesting access to, and publishing, data products. Governance teams can use the workflows to automate their desired approval processes.

Workflow types

You can set up customizable workflows in Unified Catalog for these processes:

- Granting access to data products
- Publishing data products and glossary terms

You set up these workflows directly within Unified Catalog. The setup process follows a similar process as the existing data product access request, but gives more options to customize. The new workflows go through a new requests and approvals tab, but with the same out-of-the-box options available for configuration.

> **Note**
>
> Access workflows can be applied to a governance domain or a single data product. Publication workflows can only apply to governance domains.

Prerequisites

To create and manage workflows, users need to hold the **Governance Domain Creator** role for publication workflows, and the **Data Product Owner** role for data product access workflows. Learn more about Unified Catalog roles.

Data product access workflow

A data product access workflow allows you to manage data product access requests and generate approvals. A data product owner can only create a workflow for the data products in their governance domain.

If you previously created workflows in the classic catalog experience, you should update your workflow to the new experience by creating a new workflow for data access at the data product level. If you don't create a new workflow, your existing workflow continues to work.

For any new data products you create, you can only use the new workflow experience described in this article.

Steps to create workflow

You can create workflows from either a data product's details page or from the **Workflows** page.

1. In Unified Catalog, start from either of these locations:

    - **From a data product**:
        a. Under **Catalog management**, select **Data products**, then select the data product you want to create a workflow for.
        b. On the data product's details page, select **Manage policies**.
        c. On the **Manage access policies** flyout pane, under **Approval requirements**, either add your approvers like you're used to **or** select **New Workflow** to edit the approver process. If you choose to use the approvers list in the flyout page, the workflow is automatically generated for you when you select **Save**.

    - **From the Workflows page**:
        a. Under **Process automation**, select **Workflows**.

2. Starting from either location in step 1, you arrive at the **Workflows** page. Select **New**.

3. In the **New workflow** flyout pane, enter a **Name** and an optional description.

4. At **Workflow category**, select **Data product access** from the list.

5. Select **Create**. A base template for the workflow opens, which is where you build your workflow.

6. At **Start and wait for an approval**:

    - Select an **Approval type**:

    - **Pending on all**: If there are multiple approvers for the access request, all approvers must approve the request.
    - **Pending on any**: Only one approver needs to approve the request.
  - The **Title** is automatically filled in as **Approval for data access request**.
  - At **Assigned to**, enter the user name.

7. At **Condition**, select the condition you want to apply, such as **and** or **or**, and set the approval outcome.

8. The **Condition** leads to **If yes** and **If no** outcomes. **If yes** results in the approval of the access request. **If no** results in the rejection of the access request. Within the **Yes** and **No** outcomes, you can add more actions to create extra layers based on conditions.

9. When you're done setting conditions, select **Set scope** at the top right of the page.

10. On the **Set scope** flyout pane, expand the list under the governance domains listed to view its data products. Select one data product or an entire domain of data products. Then select **Save**.

Limitations

1. The current version doesn't support variables for contacts, metadata during the workflow, or email templates.
2. This set of workflows doesn't include external connectors.
3. You can only configure data product access workflows for a single data product or the whole governance domain of data products for the same list of approvers. You can't use these workflows for multiple data products or many domains.
4. If requests exist for a user on a data product in either workflow experience, the user can't create a new request. When you switch to the new workflow, remove all previously approved requests and have users recreate their requests if needed.
5. If the privacy approver is hard coded in a workflow, it won't prepopulate in the request form when filled in by the data consumer.
6. Security groups can be added for approvals, but someone from that group needs to manually approve the access request. This requirement applies for both the privacy approver and data product approver.
7. A security group added to be any approver must be email enabled to receive emails notification during workflow runs.
8. A group can be added to the access request if there are **zero** attestations on that data product. If attestations are required for a data product, only users and service principal

requests are supported.
9. Data access workflows experience and APIs only support user principals. Service principals won't work in the workflow roles.

Data product and glossary term publish workflow

A **Catalog curation** publish workflow lets you automate the publishing of data products and glossary terms for a specific governance domain. It enables governance domain data product owners and data stewards to collaborate more broadly during the drafting process but ensures a final review before publication to the catalog.

If you previously created workflows in the classic catalog experience, you can update your workflow to the new experience by creating a new workflow. If you don't create a new workflow, your existing workflow continues to work.

For any new data products and glossary terms you create, you can only use the new workflow experience described in this article.

Steps to create workflow

1. In Unified Catalog, under **Process automation**, select **Workflows**. You arrive at the **Workflows** page.

2. Select **New**.

3. In the **New workflow** flyout pane, enter a **Name** and an optional description.

4. At **Workflow category**, select **Catalog curation** from the list.

5. At **Workflow type**, select **Data product publish** or **Term publish**.

6. Select **Create**. A base template for the workflow opens, which is where you build your workflow.

7. At **Start and wait for an approval**:

    - Select an **Approval type**:
        - **Pending on all**: If there are multiple approvers for publishing, all approvers must approve the publish.
        - **Pending on any**: Only one approver needs to approve the publish.
    - The **Title** is automatically filled in as **Approval for data product publish**.
    - At **Assigned to**, enter the user name XYZ.

8. At **Condition**, select the condition you want to apply, such as **and** or **or**, and set the approval outcome.

9. The **Condition** leads to **If yes** and **If no** outcomes. **If yes** results in publishing the data product. **If no** results in the data product remaining in a draft state. Within the **Yes** and **No** outcomes, you can add more actions to create extra layers based on conditions.

10. When you're done setting conditions, select **Set scope** at the top right of the page.

11. On the **Set scope** flyout pane, expand the list under the governance domains listed. Select one governance domain for the publication workflow. Then select **Save**.

Adding actions

At the **If yes** and **If no** outcomes, add other actions to the workflow to build multiple levels and tiers based on the conditions. The available actions are:

- Start and wait for approval
- Condition
- Approve data subscription
- Complete data subscription
- Reject data subscription
- Send email notifications

Last updated on 07/31/2026

Manage data health in Microsoft Purview Unified Catalog

08/20/2025

Health management refers to the ongoing processes and practices that ensure an organization's data remains accurate, complete, consistent, secure, and accessible throughout its lifecycle. It's a critical aspect of data governance, focusing on maintaining the overall "health" of data, much like how healthcare practices aim to maintain the health of individuals.

Importance of health management:

- **Informed decision-making:** High-quality, reliable data is essential for making informed business decisions.
- **Operational efficiency:** Healthy data reduces the time and resources spent on correcting errors and reconciling inconsistencies.
- **Risk management:** Proper data health management helps mitigate risks associated with data breaches, regulatory noncompliance, and inaccurate reporting.
- **Competitive advantage:** Organizations with healthy data are better positioned to leverage analytics, AI, and machine learning, leading to better insights and innovations.
- **Customer trust:** Ensuring data privacy and accuracy builds trust with customers and partners.

In summary, health management is about maintaining the integrity, reliability, and security of data throughout its lifecycle, ensuring that it remains a valuable asset to the organization.

Unified Catalog health management provides four key components and tools to help you understand your data and develop your data governance strategies:

- Controls
- Data quality
- Actions
- Reports

You can find these tools by opening **Unified Catalog** in the Microsoft Purview portal and selecting **Health management**.

This article provides a brief overview of each tool, its scope, and how you can use it to understand your data and develop your data governance strategies. Users need the **Data health owner** and **Data health reader** roles to access and configure rules for controls and to browse built-in reports.

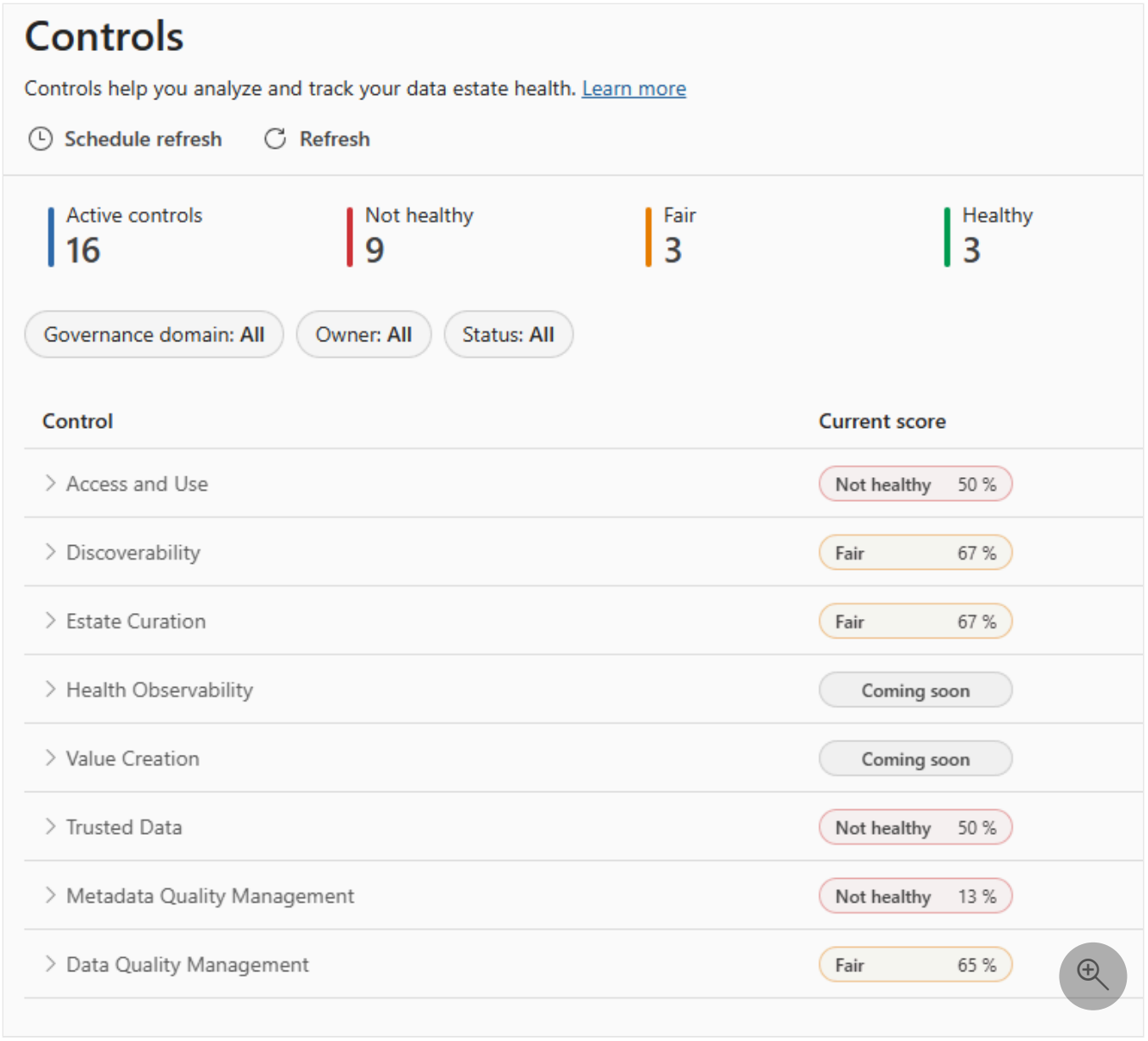
Expand table

Role	Description	Application
Data health owner	Create, update, and read artifacts in health management.	Health management
Data health reader	Can read artifacts in health management.	Health management

Controls

Controls help data stewards and chief data officers evaluate the health of their data estate through industry-recognized standards and controls. Set control rules at a global level but leave execution to individual governance domains. This approach lets you set broad standards while meeting the needs of individual groups in your organization.

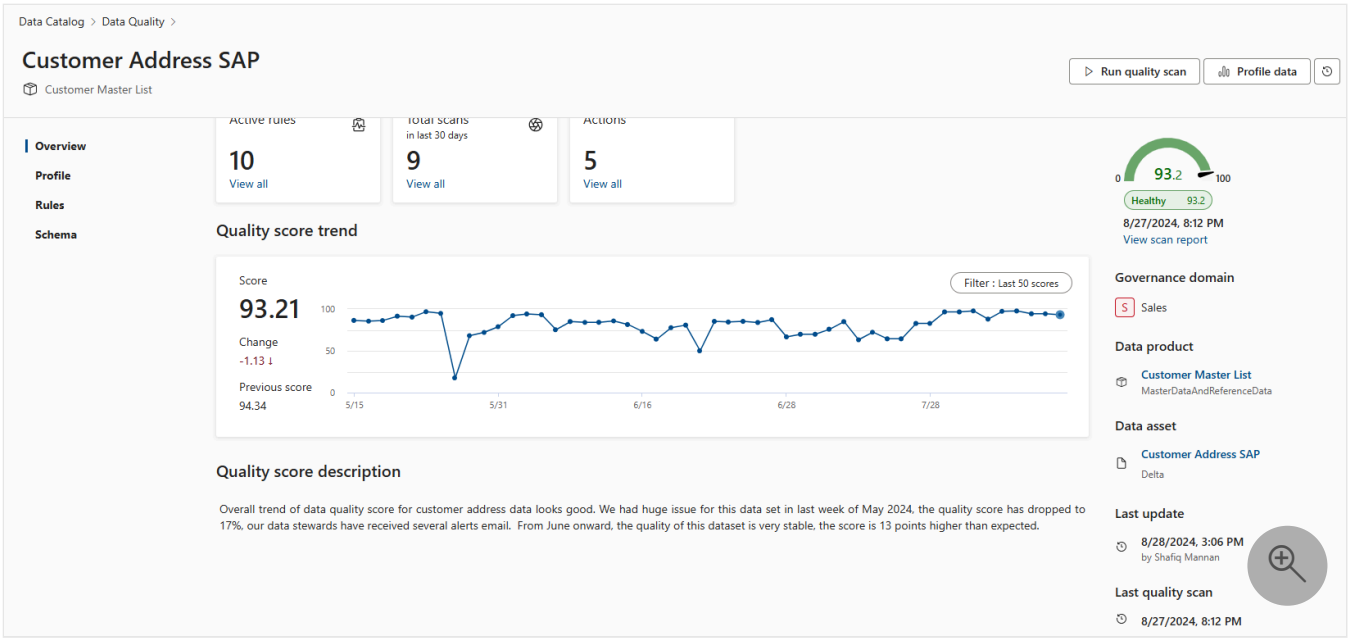
If you're a data health owner, you can edit controls and change the preset health ranges and thresholds to customize how you measure the success of your data estate. You can also edit the schedule for the overall data refresh rate for all active controls. For more information, see the Health management controls article.



Data quality

Microsoft Purview Data Quality is a comprehensive solution that empowers governance domain and data owners to assess and oversee the quality of their data ecosystem, facilitating targeted actions for improvement. In today's AI-driven landscape, the reliability of data directly impacts the accuracy of AI-driven insights and recommendations. Without trustworthy data, there's a risk of eroding trust in AI systems and hindering their adoption.

Microsoft Purview Data Quality also incorporates AI-powered data profiling capabilities, recommending columns for profiling while allowing human intervention to refine these recommendations. This iterative process not only enhances the accuracy of data profiling but also contributes to the continuous improvement of the underlying AI models.

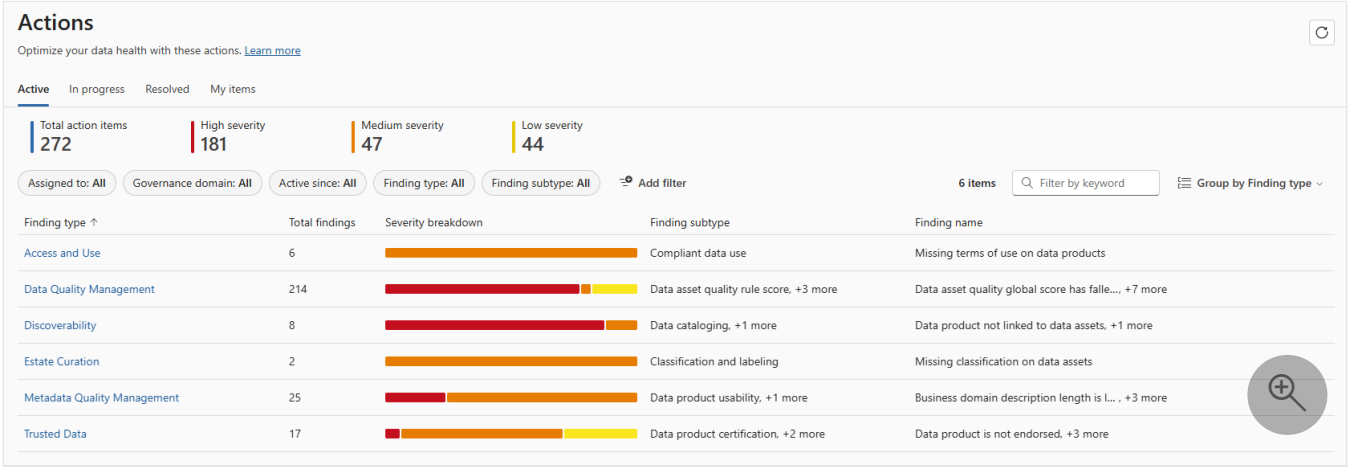


Actions (preview)

Health actions give data governance practitioners truly actionable insights. All the anomalies that health controls note become actions, alerts with owners, and recommendations that you can track and address from within Unified Catalog.

As a data steward, you can review actions and filter the list by finding type, finding subtype, and finding name. You can assign a task to anyone in your organization, and that user sees the action in their list. These actions update on the same schedule as the health controls, since actions come from the controls themselves.

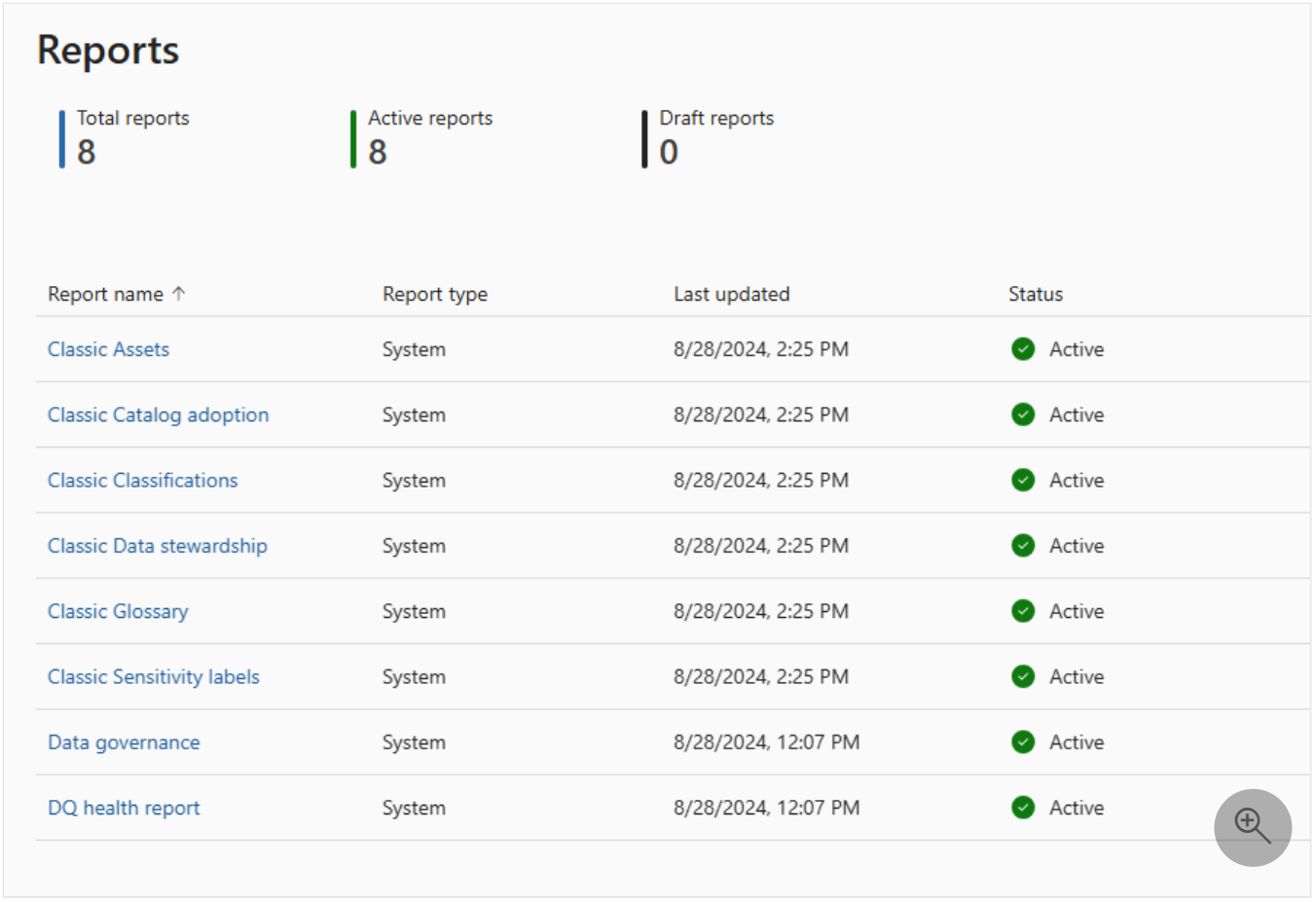
Get details about viewing all actions on the Actions (preview) page.



Reports

Data analysts and stewards in the federated data governance office can view built-in reports for deeper analysis of the data estate. These reports give summaries of individual business

concepts like governance domains or data products. They also address data governance goals, catalog adoption, and asset information like classification. For more information, see the health management reports article.



View all data health management actions on the Actions (preview) page

10/15/2025

The **Actions (preview)** page in Microsoft Purview Unified Catalog displays all health management actions related to catalog metadata, controls, and data quality. Health management actions give you and your users steps to take to improve data health and governance across your data estate. These actions correspond to the checks made to calculate a data product's data governance health control score. Addressing these actions raises your health score and promotes a more useful and discoverable Unified Catalog.

Prerequisites

- You need data health reader permissions to view data health information.
- You need data health owner permissions to update and assign health actions.

View health management actions

To view health management actions, go to **Unified Catalog** > **Health Management** > **Actions**. The **Actions (preview)** page shows a list of actions, how long they're active, the target for each action, and the owner.

You can search by keyword or filter these actions by many values, including: assigned to, governance domains, active since, action types, and more.

You can also group actions by selecting **Group by** to view actions by:

- Severity
  - High
  - Medium
  - Low
- Finding types and subtypes
- Finding subtype
- Finding name

> **Tip**
>
> - If you don't see the actions you expect, check the applied filters and group.

- You need the **Data Quality Metadata Reader** role to see the **Target entity** related to the action. The target entity provides a deep link to browse for remediation action.
- You won't see the **Target entity** if it's deleted from the governance domain or data product, or if you don't have the **Data Quality Metadata Reader** role.

Available actions

The actions available are determined by control rules or data quality rules. Learn more about data quality actions.

Here's a list of actions that are available from metadata quality:

Expand table

Finding type	Finding subtype	Finding Name	Severity
Access and use	Compliant data use	Missing terms of use on data products	Medium
	Self-serve access enablement	Missing access policy on data products	Medium
Discoverability	Data cataloging	Data product not linked to data assets	High
		Missing description on data products	High
		Missing published glossary terms on data products	High
		Missing use case on data products	High
	Data products connection	Data product not linked to data assets	Medium
Estate curation	Classification and labeling	Missing classification on data assets	Medium
Health observability	Health management monitoring, alerting, and insights		
Metadata quality management	Data product usability	Governance domain description length is fewer than 100 characters	Medium
		Data product description length is fewer than 100 characters	Medium

Finding type	Finding subtype	Finding Name	Severity
		Published glossary term's description length is fewer than 25 characters	Medium
	Linked assets	Data product description length is fewer than 100 characters	Medium
		Missing published glossary terms on data product	High
Trusted data	Data product certification	Data product isn't endorsed	Low
	Data product ownership	Missing owners on data assets	High
		Missing owners on data products	High
	Data quality enablement	Missing data quality scores on data assets	Medium
Value creation	Business OKRs alignment	Missing OKRs on data products	Medium

> **Tip**
>
> If the actions you see don't match up with these values and severities, check metadata quality. Your organization might have implemented custom checks and severities.

Action details

To see an action's details, select it. You see the action owners, the associated data product and governance domain, as well as its reason and any potential recommendations for resolving it.

Resolve an action

To resolve an action:

1. Assign owners who take steps to resolve the recommendations.
2. The owners investigate the issue based on the provided information and recommendation.
3. The owners update the status to reflect their current progress.
4. When they make the suggested updates, they update the action's status to **Resolved** and save the action.

As the entire health management refreshes in the next scheduled run, it reconciles all resolved actions from the catalog. If you artificially mark an action as resolved but the issue in the catalog remains, the action status changes and is added back to the health actions queue.

Next steps

- Metadata quality
- Health actions best practices

Data health controls in Unified Catalog

10/15/2025

Data health controls help your team analyze and track your health management. Track your journey to complete data governance by monitoring your governance health, and use the provided health controls to track your progress. Data health controls are specific measures, processes, and tools that you implement to monitor, maintain, and improve the quality, security, and overall health of your organization's data. These controls are part of a broader data governance framework and are essential for ensuring that data remains accurate, consistent, and usable across the organization.

Data health controls are essential components of a comprehensive data governance strategy. They help organizations maintain the integrity, security, and usability of their data assets. Some benefits include:

- **Improved data quality:** Ensures that data remains accurate, consistent, and reliable for decision-making.
- **Enhanced security:** Protects sensitive data from breaches, unauthorized access, and corruption.
- **Regulatory compliance:** Helps organizations adhere to legal and industry standards for data management.
- **Operational efficiency:** Reduces the time and resources spent on correcting data issues and ensures that data is readily available and usable.
- **Risk mitigation:** Prevents costly errors and data-related risks that can arise from poor data management.

Prerequisites

- You need data health reader permissions to view data estate health information.
- You need data health owner permissions to update or edit any health controls.

Limitations

- Currently, you can't create custom controls.

Access data health controls

To view controls, go to **Unified Catalog** > **Health management** > **Controls**. The **Controls** page shows your governance scores at the domain level, so you can understand your organization's progress at a glance.

Health controls are grouped into several categories that you can open to view their parts. Category scores are an aggregate of the controls underneath them. They show their current score and their target, so you can see areas for growth.

Select any category or individual health control for a detailed view.

Available health controls

The currently available health controls are:

Expand table

Control grouping	Control	Description
Access and use	Self-serve access enablement	Percentage of data products that have self-serve policies and workflows enabled.
	Compliant data use	Percentage of data subscriptions with a provided data use purpose.
Discoverability	Data cataloging	Percentage of data products that a data product owner publishes in the governance domain.
	Data products connection	Percentage of data products that have data assets mapped.
Estate Curation	Classification and labeling	Percentage of data products and governance domains that are classified and labeled.
	Critical data identification	Percentage of business domains with at least one critical data element.
Value creation	Business OKRs alignment	Percentage of data products aligned with one or more business objectives and key results (OKRs).
Trusted data	Data product ownership	Percentage of data products that have owners assigned.
	Data quality enablement	Percentage of data products with a data quality score.
	Data product certification	Percentage of data products certified by an authorized data product owner in the governance domain.
Metadata quality management	Data product useability	Percentage of data products that have: a description greater than 100 characters, a parent domain with a description of over 100 characters, and a glossary term with at least a 25 character description.

Control grouping	Control	Description
	Linked assets	Percentage of data products with at least one linked asset and published glossary term.

Edit health controls

If you're a data health owner, you can edit your health controls to:

- Edit control details, such as owners, description, and governance domains.
- Activate or deactivate a control to manage scores.
- Configure thresholds to reflect your business goals.

Edit control details

1. On the **Controls** page, expand a control group name to view individual controls.
2. Hover over the name of the control you want to edit and select the **Edit** pencil button. You arrive at the **Details** tab of the **Edit control** page.
3. At **Owner**, add or remove owner names.
4. At **Description**, edit the control's description.
5. At **Domains**, select all governance domains you want to include in the score, and deselect any domains you don't want to include.
6. Select **Save** at the top right of the page to save your changes.

Activate or deactivate a control

You can deactivate any control so that new scores no longer roll up to a control group.

1. On the **Controls** page, expand a control group name to view individual controls.
2. Hover over the name of the control you want to edit and select the **Edit** pencil button. You arrive at the **Details** tab of the **Edit control** page.
3. Select the **Active** toggle to switch the control to the active or inactive state.
4. Select **Save** at the top right of the page to save your changes.

Configure control thresholds

Adjust your health control's target score status, or threshold, to reflect your business goals.

1. On the **Controls** page, expand a control group name to view individual controls.
2. Hover over the name of the control you want to edit and select the **Edit** pencil button.

3. On the **Edit control** page, select **Thresholds**.

4. Make the following adjustments:
   a. Set the overall target score to reflect your goals.
   b. Adjust existing threshold conditions to show any of these statuses:

      - Undefined
      - Healthy
      - Fair
      - Not healthy
      - Critical

   c. Add and remove conditions on existing rules to calculate status based on score.
   d. Add or remove thresholds to calculate your status based on target score.

   > **Tip**
   >
   > To return to the saved calculation before your edits, select the **Reset** command at the top right of the page.

5. Select **Save** at the top right of the page to save your changes.

Configure control rules

Add rules to controls to measure their performance.

1. On the **Controls** page, expand a control group name to view individual controls.
2. Hover over the name of the control you want to edit and select the **Edit** pencil button.
3. On the **Edit control** page, select **Rules**.
4. Select **Add rule**. The **Add rule** flyout pane opens.
5. Search for a rule from the **Data product rules** list and select the rule to add it.
6. Select **OK**.
7. Add the needed parameter and accept it by selecting the check symbol.
8. Select **Save changes**.

Configure rules severity

1. On the **Controls** page, expand a control group name to view individual controls.
2. Hover over the name of the control you want to edit and select the **Edit** pencil button.
3. On the **Edit control** page, select **Rules**.
4. Select **Configure severity**.

5. Select a control name, and update the severity for the **Finding name**. Finding names are the rule names, and you see these finding names as action name if the defined rule fails for the applied severity.

Schedule control refresh

The schedule for the control refresh affects your health management actions and reports, as the controls provide the basis for many of these values. Consider how often you want to evaluate your overall data estate health, and schedule the controls refresh accordingly.

1. On the **Controls** page, select **Schedule refresh**.
2. At **Start**, enter a date, time, and time zone.
3. Set the frequency by selecting a number and choosing between **Days**, **Weeks**, or **Months**.
4. At **Until**, set an end date for the schedule, or leave blank if you want it to continue on the schedule indefinitely.
5. Save the configured schedule.

Data health controls pricing

Data health controls usage is billed based on the Data Governance Processing Unit (DGPU) pay-as-you-go meters. The Basic SKU is used for compute of data health controls. Find details on how pricing is computed for data health controls.

Metadata quality in Unified Catalog

Metadata quality is a low-code or no-code experience for data stewards and members of the Chief Data Officer's office, to write any logic to test the metadata health and quality. Data stewards can write complex logic without worrying about coding, and immediately see results in Microsoft Purview Unified Catalog. Metadata quality comes with a set of predefined logic for each control, that you can read about in the health controls article. As a data steward, you can add more rules to the existing logic, and at the next refresh of the health controls, the new metadata quality logic applies to the metadata based on the scope, such as data product, governance domain, or other entity.

Prerequisites

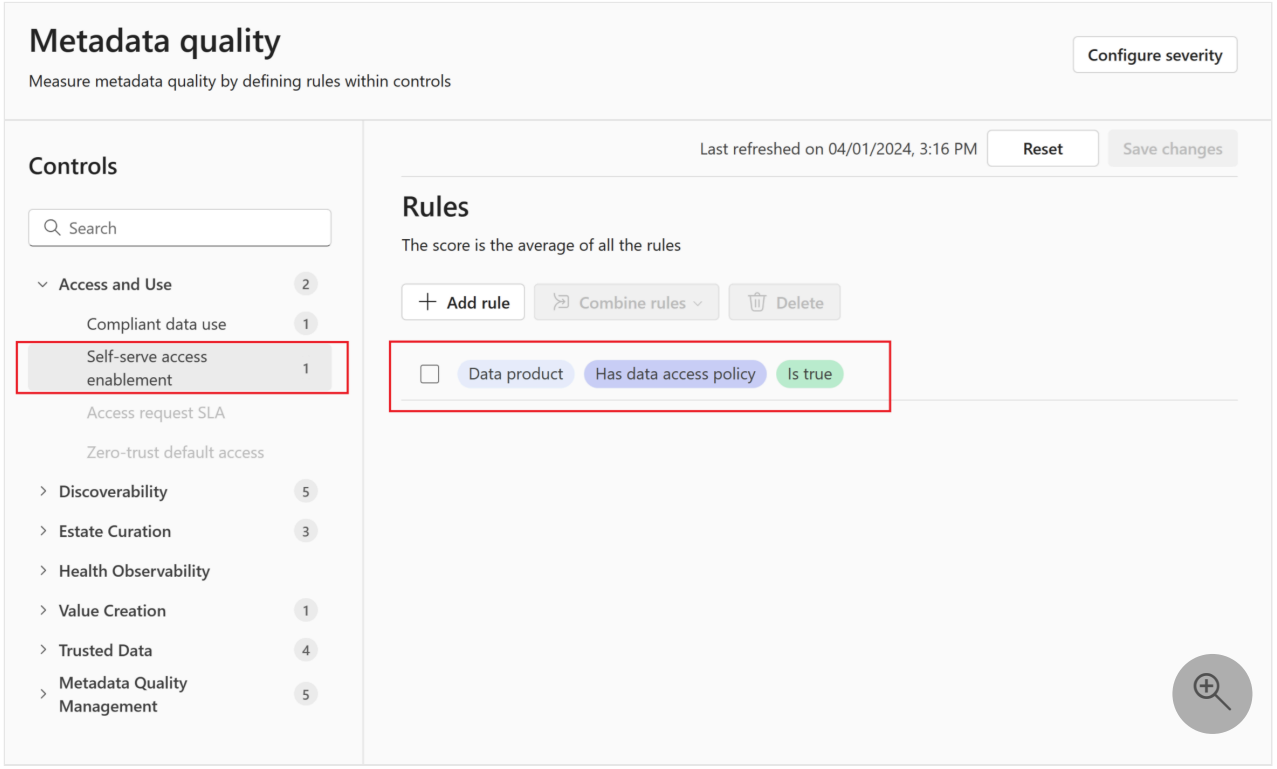
Users need the data health owner role to update and manage metadata quality.

Access metadata quality

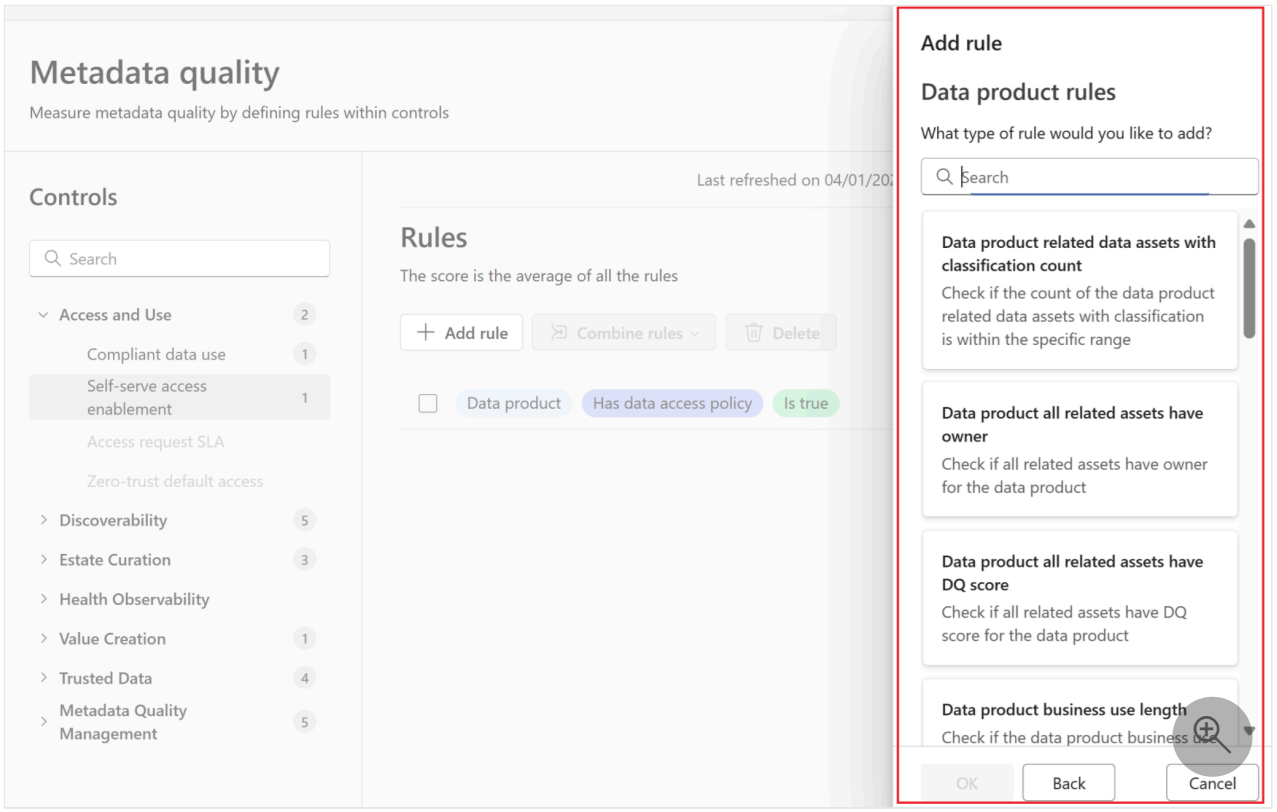
1. In **Unified Catalog**, select **Health management**, then select **Controls**.
2. On the **Controls** page, expand **Metadata Quality Management**.

Add a metadata quality control rule

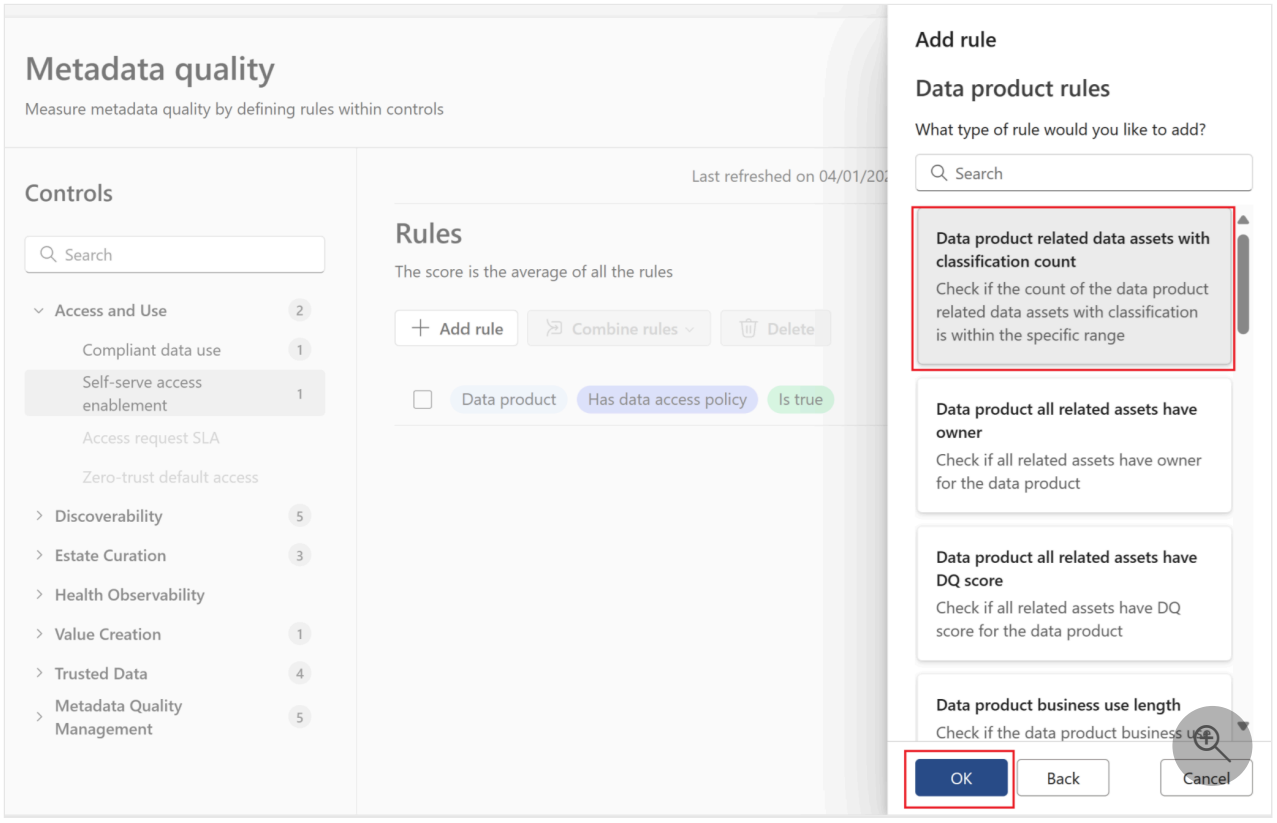
1. On the metadata quality page, select any control to see its list of controls.



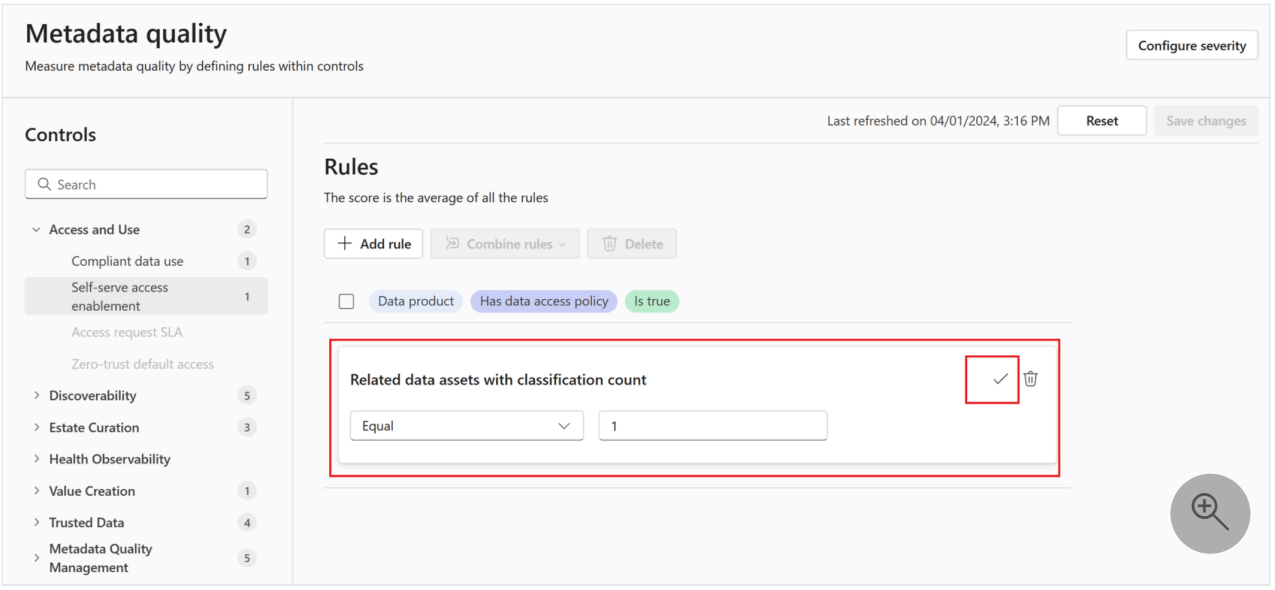
2. Add a new rule by selecting **Add rule**. The **Add rule** flyout pane opens with a list of all available rules for your current scope. (Example scope: Data product)



3. Select a rule, and select **OK**.

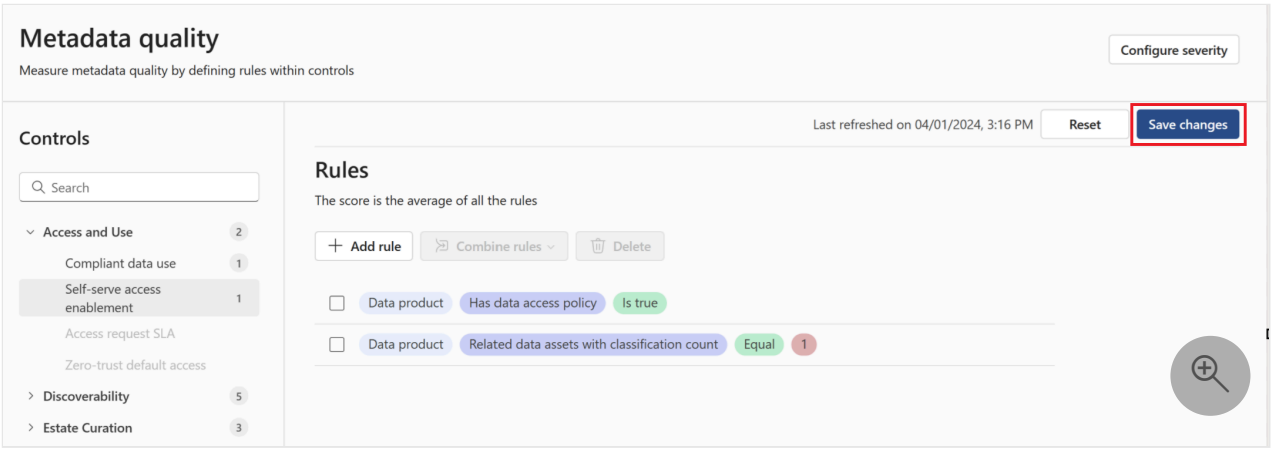


4. Select any more parameters required to complete the logic (For example, select the number of data assets with a classification count)



5. Select the checkmark button.

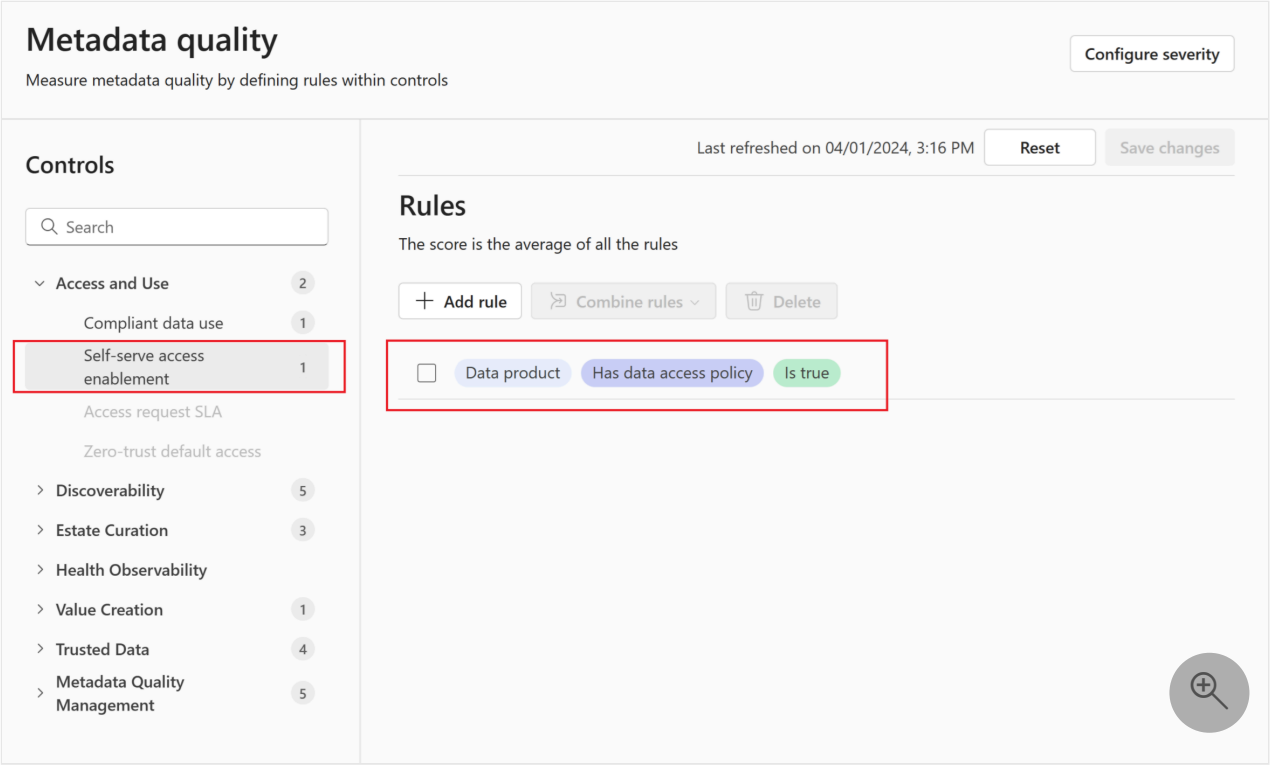
6. Select **Save changes** to save your updates.



> **Note**
>
> As the metadata quality rules run, all pass checks contribute to health control scores, and all failed checks contribute to health control actions.

Edit a metadata quality control rule

1. On the metadata quality page, select any control to see its list of controls.



2. Hover over any existing rule and select the pencil icon to edit, or the trash can icon to delete.

3. If you edit, update any rule logic and select the check mark button.

4. Select **Save changes** to save your changes.

Combine rules

1. Select the checkbox next to more than one rule, then select the **Combine rules** button.
2. Select whether you want to combine the rules with the OR or AND operator.
3. Select **Save changes** to save your updates.

Reset all rules

At any time, you can reset all rules back to their default using the **Reset** button. If you do, you lose any customizations you made.

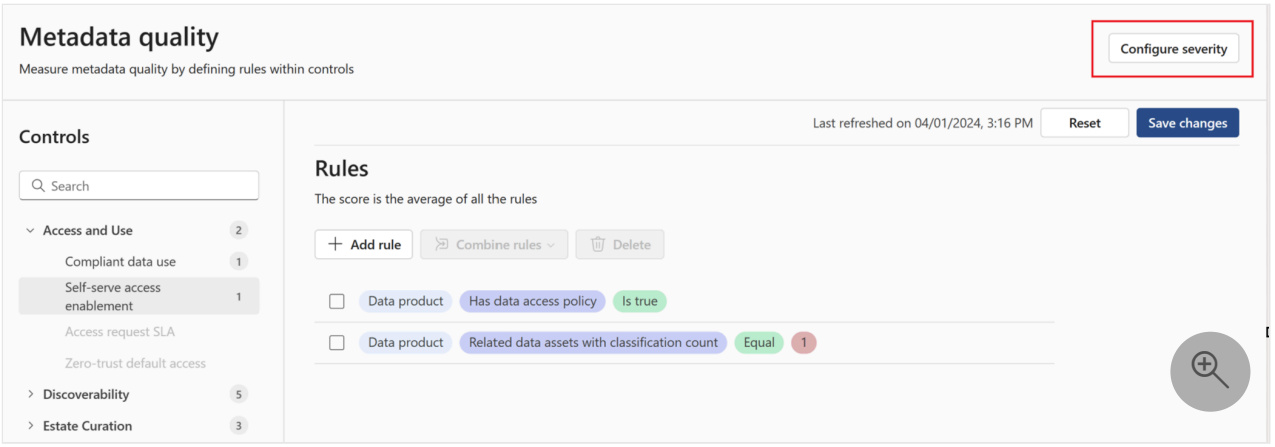
Set severity

Set the severity of your controls to change how they appear in the health actions when rule failures occur.

1. In **Unified Catalog**, select **Health management**.

2. Select **Metadata quality**.

3. Select **Configure severity**.

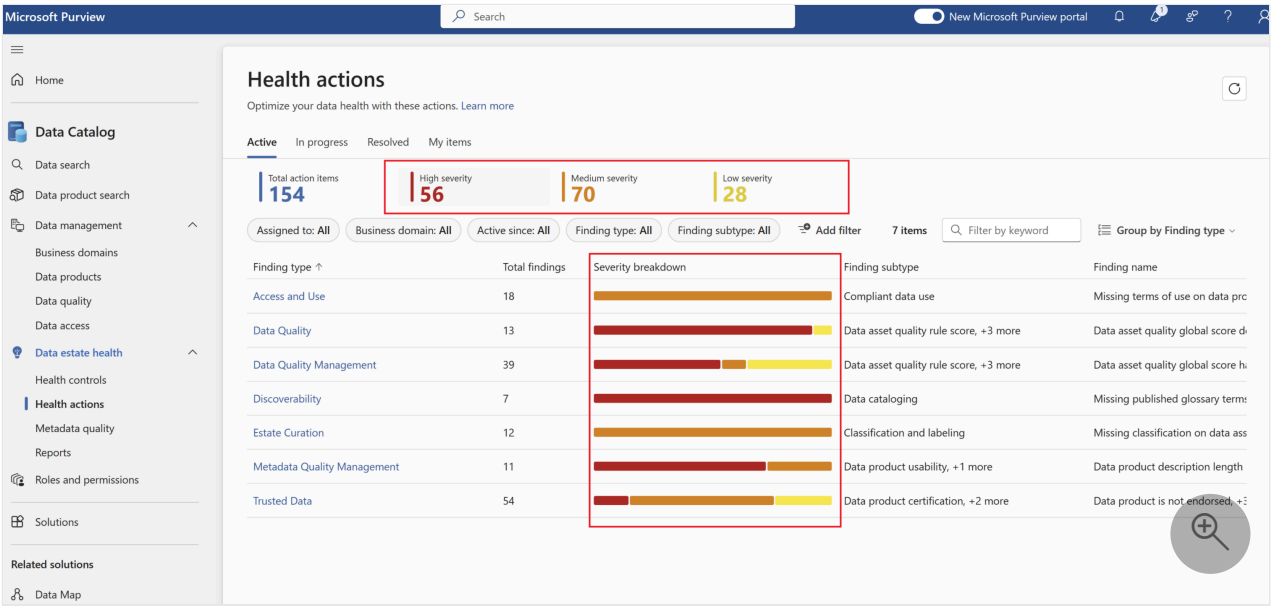


4. Select a control category, then select your control.

5. For each control, set the severity to one of these options: Low, Medium, High.

6. Select **Save** to save your changes.

7. When metadata quality rules run, your health actions update according to your new settings.



Default actions

Here's a list of the default actions that are available out of the box:

Expand table

Finding type	Finding subtype	Finding Name	Severity
Access and use	Compliant data use	Missing terms of use on data	Medium

Finding type	Finding subtype	Finding Name	Severity
		products	
	Self-serve access enablement	Missing access policy on data products	Medium
Discoverability	Data cataloging	Data product not linked to data assets	High
		Missing description on data products	High
		Missing published glossary terms on data products	High
		Missing use case on data products	High
	Data products connection	Data product not linked to data assets	Medium
Estate curation	Classification and labeling	Missing classification on data assets	Medium
Health observability	Health management monitoring, alerting, and insights		
Metadata quality management	Data product usability	Governance domain description length is fewer than 100 characters	Medium
		Data product description length is fewer than 100 characters	Medium
		Published glossary term's description length is fewer than 25 characters	Medium
	Linked assets	Data product description length is fewer than 100 characters	Medium
		Missing published glossary terms on data product	High
Trusted data	Data product certification	Data product isn't endorsed	Low
	Data product ownership	Missing owners on data assets	High
		Missing owners on data products	High
	Data quality enablement	Missing data quality scores on data assets	Medium
Value creation	Business OKRs alignment	Missing OKRs on data products	Medium

Limitations

Currently, you can edit and manage existing health control metadata quality, but you can't create new controls.

Next steps

- Health actions

Last updated on 11/10/2025

Learn about data observability in Unified Catalog (preview)

Data observability in Microsoft Purview Unified Catalog provides a bird's eye view into the health of the data estate as data flows between sources. Data stewards, experts, and those accountable for data health can view their data estate in a single screen, providing the ability to view the lineage of data assets all the way up to the structure of the governance domains. This single view helps users understand where data starts and ends, what the issues are, and where there's an effect on reporting obligations.

How data observability works

The visualizations in data observability bring together all of the metadata from Microsoft Purview Unfiied Catalog for things like technical data lineage and data quality. Data observability provides contextual views and on-demand custom views in a single diagram.

Having the right view of the metadata in a single place allows you to dig into questions about how data quality is changing across your data estate, and how your data's being used and by whom. Viewing the entire data estate at one time enables the central data team or federated data teams to quickly identify improvements in metadata, ownership, data quality, or data architecture.

> **Note**
>
> Data observability doesn't create any of the lineage or metadata used in the visualization. If you haven't added lineage, or haven't yet created governance domains and business concepts, the screen shows a limited amount of the data estate to view.

Where observability shows up in Unified Catalog

You see an **Observability** tab when you look at a details page for a governance domain, a data product, and a glossary term. Also, the **Data observability (preview)** page under **Health management** in the left navigation provides a custom view of your entire data estate. See Observability views for details.

> **Note**

You need certain roles to see and work with data observability. See Unified Catalog roles for details.

Next steps

- Get set up for observability
- Understand observability views

Last updated on 11/11/2025

Set up data observability in Unified Catalog (preview)

To start generating visualizations in Microsoft Purview Unified Catalog, complete the following setup steps. This setup enables two snapshot jobs that collect metadata needed to render the diagrams.

Prerequisites

1. Assign the Data Governance Administrator role in the Data Governance role group.
2. Create assets in Microsoft Purview Data Map, and create governance domains and concepts in the Unified Catalog.

Setup steps

1. Sign in to the Microsoft Purview portal using credentials for a *Data Governance Administrator* role.
2. Go to **Settings**, then select **Unified Catalog**.
3. Select **Data Observability**.
4. Select **Enable** next to the **Catalog Snapshot Job** heading.
5. Select **Enable** next to the **Data Quality Snapshot Job** heading.

The snapshot jobs take a few minutes to run before the visualizations are enabled. The catalog snapshot job refreshes the view of metadata and lineage every 1440 minutes. The data quality snapshot job refreshes the view of data quality scores every 10 minutes.

> **Important**
>
> There's a known limitation in browser performance that prevents rendering a large snapshot of more than 500 MB.

> **Tip**
>
> You don't have to enable the data quality snapshot job if you're not working on data quality. You can turn on this job later.

Last updated on 06/25/2026

Observability views (preview)

Data observability in Microsoft Purview Unified Catalog provides specific views focused on a business concept, such as a governance domain or data product, and the user who views that concept. Each view is scoped to a specific concept and provides alternate views in a dropdown menu to help give different perspectives of the data. You can also create your own view by using saved views in the data observability experience under **Health management**.

Saved views

When you create a saved view, you create a diagram that you can reference later or share with others. After building a diagram by providing minimal details and deciding on sharability, you can reuse the same view at any time.

To create a saved view:

1. In Unified Catalog, go to **Health management** > **Data observability**.
2. Create a diagram by searching, filtering, or selecting the items you want to focus on.
3. Select **Save this view** at the top right of the screen.
4. In the **Save view** window, provide details about the view:
   a. Enter a **View name**.
   b. Enter a **Description**.
   c. At **Owner**, you can add more users. Only the view owners can manage the view.
   d. Select **Visible to everyone** if you want other users to see the view.
5. Select **Save**.

To delete saved views:

1. In **Data observability**, select **Manage saved views** from the **Save this view** dropdown.
2. On the flyout pane, select the trash can icon next to the view you want to remove.
3. Select **OK**.

Known limitations:

- Only 25 shared views can exist at one time.
- Views aren't a point in time picture; they're based on what is in the most current snapshot for the view.
- Views are either personal or shared to all users of **Data observability**. Views don't impact the permissions the user has or what they can see in the catalog.
- You can't edit views. If you need to edit a view, start from the previous saved view, make the needed changes, create a new saved view, and delete the previous view.

Governance domains data observability

At the highest level of business concepts, the governance domain provides the structure and accountability for a portion of the data estate.

1. In Unified Catalog, under **Catalog management**, select **Governance domains**.
2. Select a governance domain that you have access to, then select the **Observability** tab.
3. Select a view from the **Current view** dropdown menu:
   a. **Data Products/Assets** displays all of the data products in the domain with all of the assets mapped to each product.
   b. **Data Products** displays only the data products of the domain.
   c. **Assets** displays only the assets that are mapped to data products of that domain.
   d. **Data Product Lineage** displays the inferred lineage between data products of that domain.
   e. **Data Asset Linage** displays the lineage between the assets used in the data products of the domain.
   f. **Data Assets with Data Quality issues** displays the assets that have at least one data quality dimension that isn't healthy.
   g. **Data Products with Data Quality issues** displays the data products that have at least one data quality dimension that isn't healthy.

Data product data observability

For data product owners, the **Observability** tab on a product's details page provides a quick view of everything within a single data product.

1. In Unified Catalog, select a data product by going to **Catalog management** > **Data products**.

2. On the data product's details page, select the **Observability** tab.

3. Select a view from the **Current View** dropdown menu:
   a. **Assets** displays the assets contained within the data product.
   b. **Data Assets with Data Quality issues** displays the assets with data quality issues that are in the data product.

Glossary term data observability

Glossary terms can link to data and give better context to where certain topics have data available in the governance domains across the data estate.

1. On the glossary term's page, select **Observability**.

2. The default view displays the glossary term and the connections to the data products from any domain using links to that glossary term.

Data observability in Health management

To view your entire data estate in Unified Catalog, go to **Health management**, then select **Data observability (preview)** for a custom view of data observability. This view provides a view of the entire data estate by governance domain by default, called the **Data Estate** view.

Two other views let you focus on either data assets or data products with data quality issues. The difference is that you can search for specific assets, columns, or business concepts, then add them to the diagram or add them to the diagram with lineage. Adding anything with lineage to the diagram includes all of the lineage connections from end to end for that specific asset, column, or business concept.

Other functions available for custom view:

- **Undo**: Go back a change in the diagram to the previous view.
- **Redo**: Go forward a change in the diagram to the view before a user clicked undo.
- **Scope to selection**: Focus the diagram to the single item selected on the screen.
- **Show Lineage**: Include all related items that are connected to the focus item at the same level of grain. If you're focused on a column, the diagram includes the column level lineage. If you're focused on an asset, the diagram includes asset level lineage.
- **Show non-governed assets**: Include all assets included in the lineage of an item even if they're not included in the governed assets of any domain.
- **Save Screenshot**: Download a picture of the current view for use offline.
- **Hierarchy Filter**: Select which levels of hierarchy in the concepts are shown. The hierarchy filter enables you to include deeper level of details in any currently scoped view. If the diagram is focused on data products, the hierarchy can be set to include data assets. If it has data assets, the asset schema can be added.

Last updated on 01/09/2026

Overview of data quality in Microsoft Purview Unified Catalog

Data quality in Microsoft Purview Unified Catalog empowers governance domain and data owners to assess and oversee the quality of their data ecosystem, facilitating targeted actions for improvement. In today's AI-driven landscape, the reliability of data directly impacts the accuracy of AI-driven insights and recommendations. Without trustworthy data, there's a risk of eroding trust in AI systems and hindering their adoption.

Poor data quality or incompatible data structures can hamper business processes and decision-making capabilities. Data quality in Unified Catalog addresses these challenges by offering users the ability to evaluate data quality using no-code or low-code rules, including out-of-the-box (OOB) rules and AI-generated rules. These rules are applied at the column level and aggregated to provide scores at the levels of data assets, data products, and governance domains, ensuring end-to-end visibility of data quality within each domain.

Data quality in Microsoft Purview also incorporates AI-powered data profiling capabilities, recommending columns for profiling while allowing human intervention to refine these recommendations. This iterative process not only enhances the accuracy of data profiling but also contributes to the continuous improvement of the underlying AI models.

By applying data quality, organizations can effectively measure, monitor, and enhance the quality of their data assets, bolstering the reliability of AI-driven insights and fostering confidence in AI-based decision-making processes.

Data quality life cycle

1. Assign users(s) data quality steward permissions in Unified Catalog to use all data quality features.
2. Register and scan a data source in Microsoft Purview Data Map.
3. Add your data asset to a data product
4. Set up a data source connection to prepare your source for data quality assessment.
5. Configure and run data profiling for an asset in your data source.
    a. When profiling is complete, browse the results for each column in the data asset to understand your data's current structure and state.
6. Set up data quality rules based on the profiling results, and apply them to your data asset.

7. Configure and run a data quality scan on a data product to assess the quality of all supported assets in the data product.
8. Review your scan results to evaluate your data product's current data quality.
9. Repeat steps 5-8 periodically over your data asset's life cycle to ensure it's maintaining quality.
10. Continually monitor your data quality
    a. Review data quality actions to identify and resolve problems.
    b. Set data quality notifications to alert you to quality issues.

Supported data quality regions

Data quality is currently supported in the following regions.

> ⓘ **Important**
>
> Data Quality capability is supported only when the accounts and the data sources are in the Purview supported regions. Data sources and purview account need to be in the same azure region. Concurrent DQ job execution limits are:
>
> - **Manual scans**: Max 10 concurrent jobs.
> - **Scheduled scans**: Max 25 concurrent jobs.
> - **Manual profile**: Max 10 concurrent profiling jobs.
> - **DQ Rule suggestion**: Max 10 concurrent jobs.

Supported multicloud data sources

View the list of supported data sources.

> ⓘ **Important**
>
> Data quality for Parquet files is designed to support:
>
> 1. A directory with Parquet Part File. For example: **./Sales/{Parquet Part Files}**. The fully qualified name must follow `https://(storage account).dfs.core.windows.net/(container)/path/path2/{SparkPartitions}`. Make sure the directory and subdirectory structure doesn't include {n} patterns. Instead, use a direct FQN leading to {SparkPartitions}.

2. A directory with partitioned Parquet files, partitioned by columns within the dataset like sales data partitioned by year and month. For example: **./Sales/{Year=2018}/{Month=Dec}/{Parquet Part Files}.**

Both of these essential scenarios, which present a consistent Parquet dataset schema, are supported. **Limitation:** Data quality isn't designed to support arbitrary hierarchies of directories with Parquet files. **We recommend presenting data in the (1) or (2) constructed structure.**

Currently, Microsoft Purview can only run data quality scans by using Managed Identity as an authentication option. Data quality services run on **Apache Spark 3.5** and **Delta Lake 3.2.1**.

Data quality features

- **Data source connection configuration**
  - Configure connection to allow Microsoft Purview data quality SaaS application to have read access to data for quality scanning and profiling.
  - Microsoft Purview uses Managed Identity as an authentication option.
- **Data profiling**
  - AI-enabled data profiling experience.
  - Industry standard statistical snapshot (distribution, min, max, standard deviation, uniqueness, completeness, duplicate, and more).
  - Drill down column level profiling measures.
- **Data quality rules**
  - Out of box rules to measure six industry standards data quality dimensions (completeness, consistency, conformity, accuracy, freshness, and uniqueness).
  - Custom rules creation features include number of out of the box functions and expression values.
  - Auto generated rules with AI integrated experience.
- **Data quality scanning**
  - Select and assign rules to columns for data quality scan.
  - Apply data freshness rule in the entity or table level to measure the data freshness SLA.
  - Scheduling data quality scanning job for time period (hourly, daily, weekly, monthly, and more).
- **Data quality job monitoring**
  - Enable monitoring data quality job status (active, completed, failed, and more).
  - Enable browsing the data quality scanning history.

- **Data quality scoring**
  - Data quality score in rule level (what is the quality score for a rule that applied to a column).
  - Data quality score for data assets, data products, and governance domains (one governance domain can have many data products, one data product can have many data assets, one data asset can have many data columns).
- **Data quality alerts**
  - Configure alerts to notify data owners and data stewards if data quality threshold missed the expectation.
  - Configure email alias or distribution group to send the notification about data quality issues.
- **Data quality actions**
  - Actions center for data quality with actions to address data quality anomaly states, including diagnostic queries for data quality steward to zero in on the specific data to fix for each anomaly state.
- **Data quality managed virtual network**
  - A virtual network managed by data quality that connects with private endpoints to your Microsoft Azure data sources.

Data residency and encryption

Microsoft Managed Storage account stores data quality metadata and profiling summary. It stores them in the same region as the data source, so data residency remains intact. All data is encrypted. The Purview Resource Provider regional user data store is used for metadata. It handles all the encryption and is common across all Purview services. If you want more control over your data encryption with a customer-managed encryption key (CMK), use a separate process. Learn more about Microsoft Purview Customer Key.

Data quality compute pricing

Data quality usage is billed based on the Data Governance Processing Unit (DGPU) pay-as-you-go meters. Find details on how pricing is computed for data quality.

Limitations

- Virtual network isn't supported for Google Big Query yet.
- You can apply a maximum of 200 data quality rules per data asset for a data quality scan. Get details about this limit and workarounds.

Related content

- Data quality for Fabric data estate
- Data quality for Fabric mirrored data sources
- Data quality for Fabric shortcut data sources
- Data quality for Azure Synapse serverless and data warehouses
- Data quality for Azure Databricks Unity Catalog
- Data quality for Snowflake data sources
- Data quality for Google BigQuery
- Data quality native support for Iceberg data
- Data quality for on-premises data sources (preview)

Next steps

1. Assign users data quality steward permissions in Unified Catalog so they can use all data quality features.
2. Set up a data source connection to prepare your source for a data quality assessment.
3. Configure and run data profiling for an asset in your data source.

Last updated on 09/08/2026

Data sources and file formats supported for data quality in Unified Catalog

Supported sources

Expand table

Data source	Data profiling	Data quality scan	Virtual network support	Note
Azure Data Lake Storage Gen2	Yes	Yes	Yes	
Azure Databricks Unity Catalog	Yes	Yes	Yes	
Azure Synapse serverless	Yes	Yes	Yes	
Azure Synapse Data Warehouse	Yes	Yes	Yes	
Azure SQL Database	Yes	Yes	Yes	
Azure SQL Managed Instance	Yes	Yes	Yes	
Azure Dedicated SQL Pool (formarly SQL DW)	Yes	Yes	No	
Google BigQuery	Yes	Yes	No	
Snowflake	Yes	Yes	Yes	
Fabric	Yes	Yes	Yes	Lakehouse, Shortcut to other filesystem, and mirroring with other database
Amazon S3	Yes	Yes	No	Supported via Fabric shortcut
Dataverse	Yes	Yes	No	Supported via Fabric shortcut
Google Cloud Storage	Yes	Yes	No	Supported via Fabric shortcut
Oracle	No	Yes	NA	On-premises infrastructure

Data source	Data profiling	Data quality scan	Virtual network support	Note
SQL Server	No	Yes	NA	On-premises infrastructure

Supported file formats

Expand table

File format	Data profiling	Data quality scan	vNet support
Delta	Yes	Yes	NA
Parquet	Yes	Yes	NA
Iceberg Avro	Yes	Yes	NA
Iceberg Orc	Yes	Yes	NA

Supported resource set pattern

Expand table

Pattern Name	Details
SparkPartitions	All patterns ending with {SparkPartitions} are supported provided that they do not contain any other mixed non-column patterns in their folder path.
Column partitions	All column partition patterns for parquet, delta and iceberg datasets are supported.

- **Example 1 (supported)**: Standard resource-set folder path: `https://myblob.blob.core.windows.net/sample-data/name-of-folder-output/{SparkPartitions}`
- **Example 2 (supported)**: Column partitioned resource-set folder path: `https://myblob.blob.core.windows.net/my-partitioned-data/Year={Year}/Month={Month}/Day={Day}/{SparkPartitions}`
- **Example 3 (not supported)**: Mixed resource-set path: `https://myblob.blob.core.windows.net/sample-data/data{N}.parquet`

For more information on Microsoft Data Map resource sets, see Understanding resource set.

Related content

- Data quality for Fabric data estate
- Data quality for Fabric Mirrored data sources
- Data quality for Fabric shortcut data sources
- Data quality for Azure Synapse serverless and data warehouses
- Data quality for Azure Databricks Unity Catalog
- Data quality for Snowflake data sources
- Data quality for Google BigQuery
- Data quality native support for Iceberg data

Last updated on 04/09/2026

Kubernetes supported self-hosted data integration runtime for on-premises data sources (preview)

Microsoft Purview Data Quality for on-premises data sources enables organizations to assess, monitor, and improve the quality of data stored within internal systems such as databases and legacy platforms. It supports rule-based validation, error detection, and remediation workflows, while ensuring compliance with organizational policies. By integrating with existing infrastructure, it delivers consistent data quality insights and governance across both on-premises and cloud environments.

By using a self-hosted data integration runtime, you can scale data quality processes by securely connecting on-premises data sources to Purview. This article covers the Kubernetes-based, Linux self-hosted data integration runtime, which enhances the underlying infrastructure and provides several key benefits:

- Scalability: Ability to scale to hundreds of machines.
- Performance: Improved performance for scanning workloads.
- Security (containerized): Enables containerized deployment on a Kubernetes cluster, eliminating the need to host the data integration runtime directly on a Windows machine.

Supported data sources

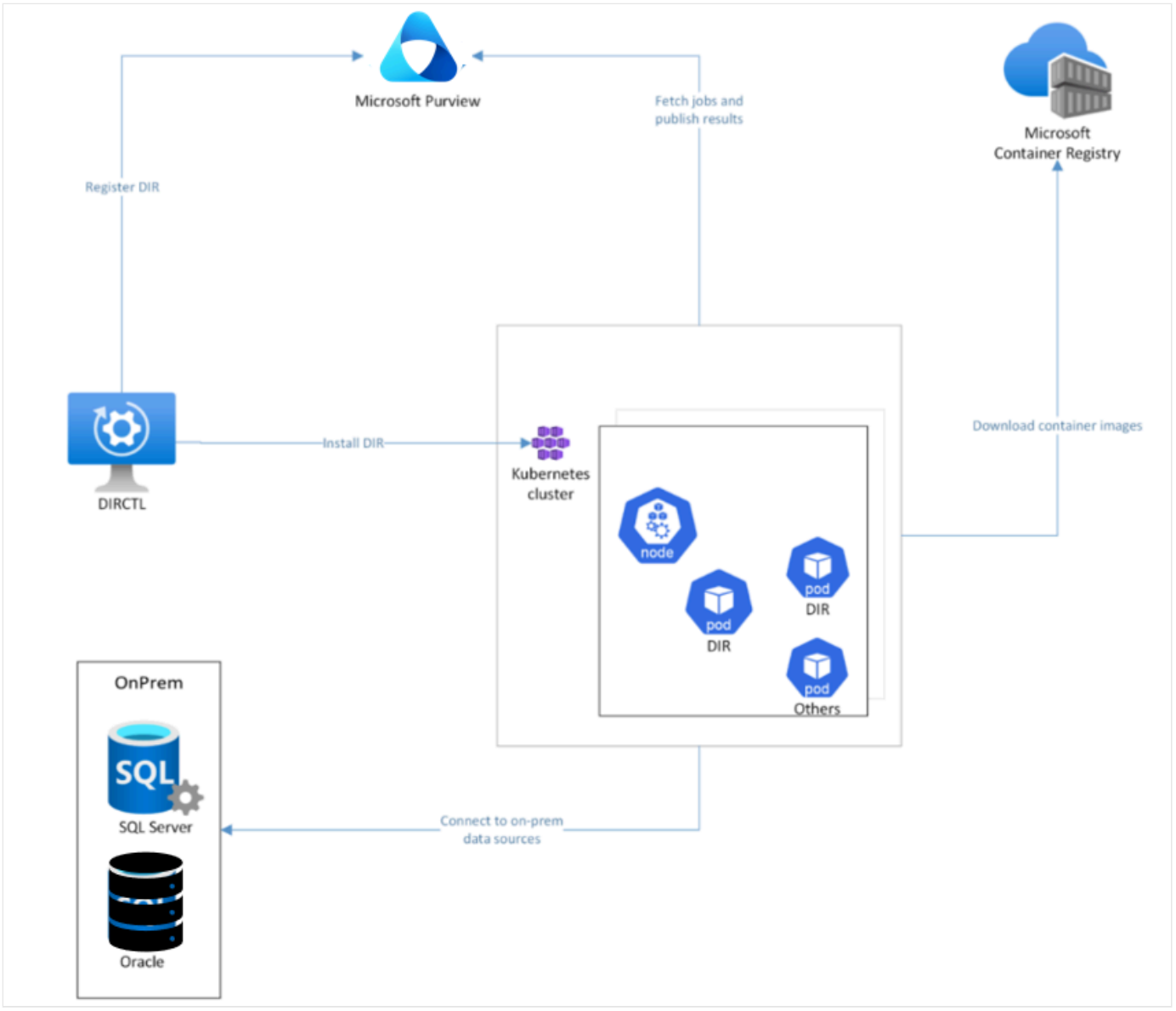
- Oracle
- SQL Server

Architecture

At a high-level architectural view, when you install a Kubernetes-based data integration runtime, several pods are autocreated on the nodes of your Kubernetes cluster. A command line tool named DIRCTL triggers this installation. DIRCTL connects to the Microsoft Purview Service to register the data integration runtime and connect to the Kubernetes cluster to install the self-hosted data integration runtime.

During the installation, the process downloads data integration runtime images from MCR (Microsoft Container Registries) to the data integration runtime pods. After installation is done,

the pods in your cluster connect to the Purview service to pull scan jobs. As a scan job is pulled, it can connect your on-premises data source for Data Quality scanning.



Prerequisites

Data Integration Runtime Command Line tool (DIRCTL)

You need the Data Integration Runtime Command Line tool (DIRCTL) to set up data integration runtime. For download and installation instructions, see Set up DIRCTL tool for self-hosted integration runtime (preview).

Roles

To set up a self-hosted integrated runtime in Purview, you need the Data Governance Administrator role.

Kubernetes cluster

You need an existing Linux-based Kubernetes cluster, or you need to prepare one. Identify the nodes by using a node selector, which follows the definition of Kubernetes node selector. Minimum configuration:

- Container type: Linux
- Kubernetes version: 1.24.9 or later
- Node OS: Linux based OS running on x86 architecture
- Node spec: minimal eight cores CPU, 32-GB memory, and at least 80 GB of available hard disk space
- Node count: 1 or more (fixed, not enabled cluster auto scaler)
- Pod number per Node: 20 or more (max Pod number – count of other Pods not belonging to Self-Hosted IR)

> **Note**
>
> The folder `/var/irstorage/` of each node is reserved for self-hosted integrated runtime. It's readable and writable to data integration runtime. You can get logs from this folder or upload external drivers to this folder. Data integration runtime creates the folder if it doesn't exist, and it doesn't delete the folder after data integration runtime is deleted. The container images used by data integration runtime are managed by Kubernetes Garbage Collection, which isn't cleaned up by data integration runtime. Configure the proper threshold for your Kubernetes cluster.
>
> Outbound connectivity is required to pull in container images as well as for further operation, which includes activities such as pulling data quality jobs and pushing generated stats.

Kubernetes context

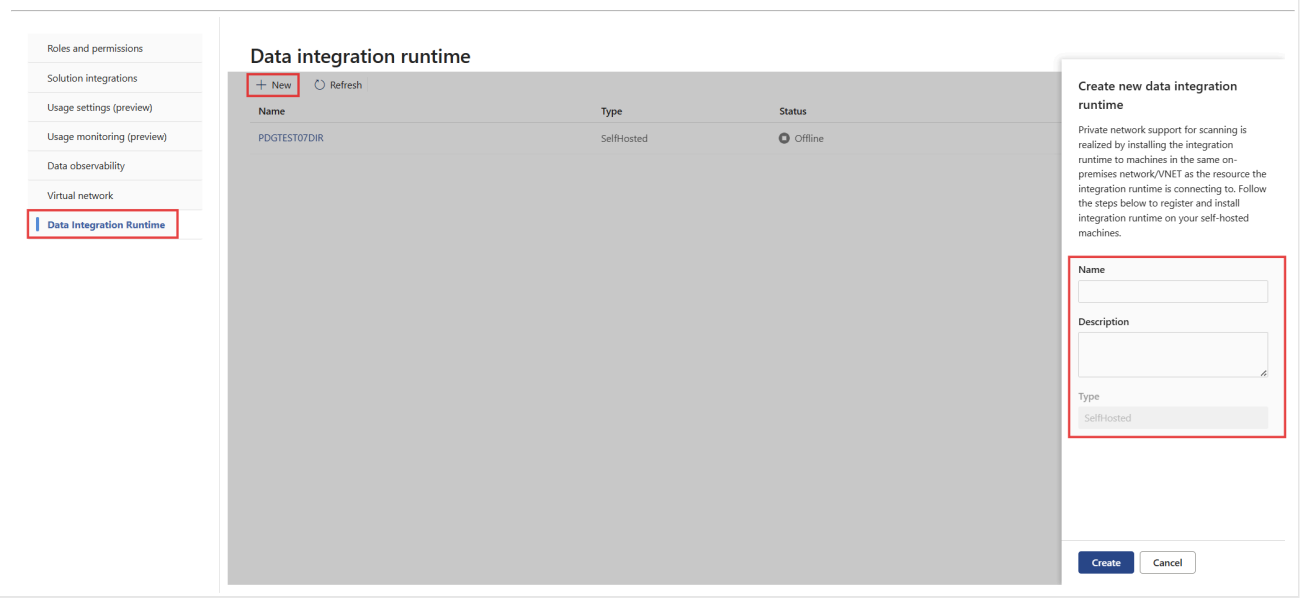
Kubernetes context, which contains Kubernetes cluster information and user's permissions and credential for this cluster, is needed to talk to your Kubernetes cluster. To ease the configuration for the user's permissions for DIR management, you can start with Kubernetes Admin role. This context is generated with the setup of your Kubernetes cluster and saved in a config file. Where and how you can get this file depends on your setting up the Kubernetes cluster.

- If you use kubeadm init to set up the Kubernetes cluster, you can find the config file under `/etc/Kubernetes/admin.conf`.

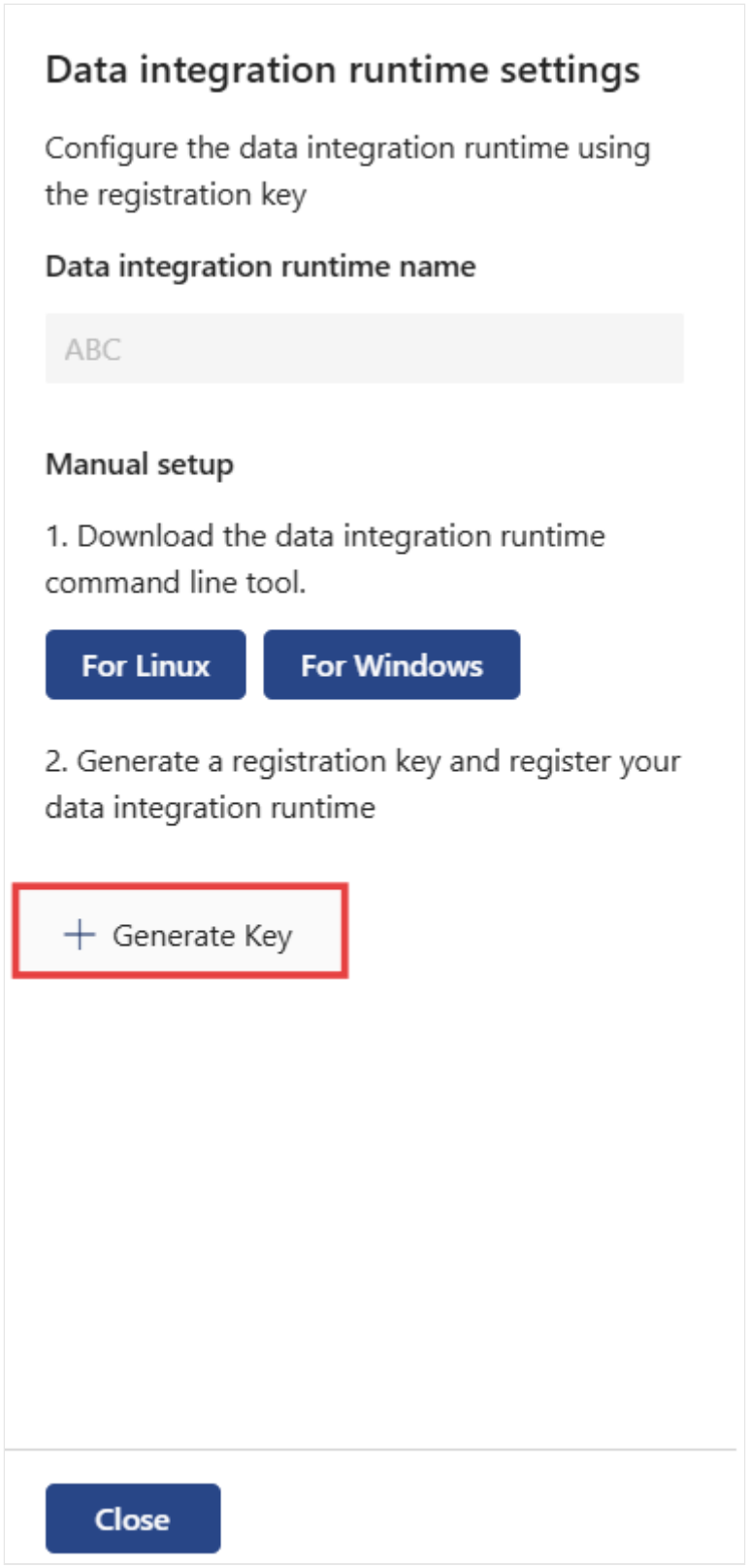
- If you use AKS, you can follow the guidance of AKS to use Az PowerShell module command to get credentials of this cluster to your local machine. You can merge the context to the config file under `$HOME/.kube/config` directly.
- If you're using other tools setting up a Kubernetes cluster, refer to the Kubernetes documentation.
- After you get the config file for the Kubernetes context, merge it into the config file `$HOME/.kube/config` on the machine where you want to run the IRCTL command. Or, you can set the config file of the Kubernetes context in an environment variable named `KUBECONFIG`. For more information about the Kubernetes context, see how to configure access to multiple clusters.

Set up a self-hosted data integration runtime

1. Go to **Settings** > **Microsoft Purview Unified Catalog** > **Data Integration Runtime**, and then select **New** to create a data integration runtime.
2. Enter a **Name** and **Description** for the self-hosted integrated runtime, and then select **Create**.



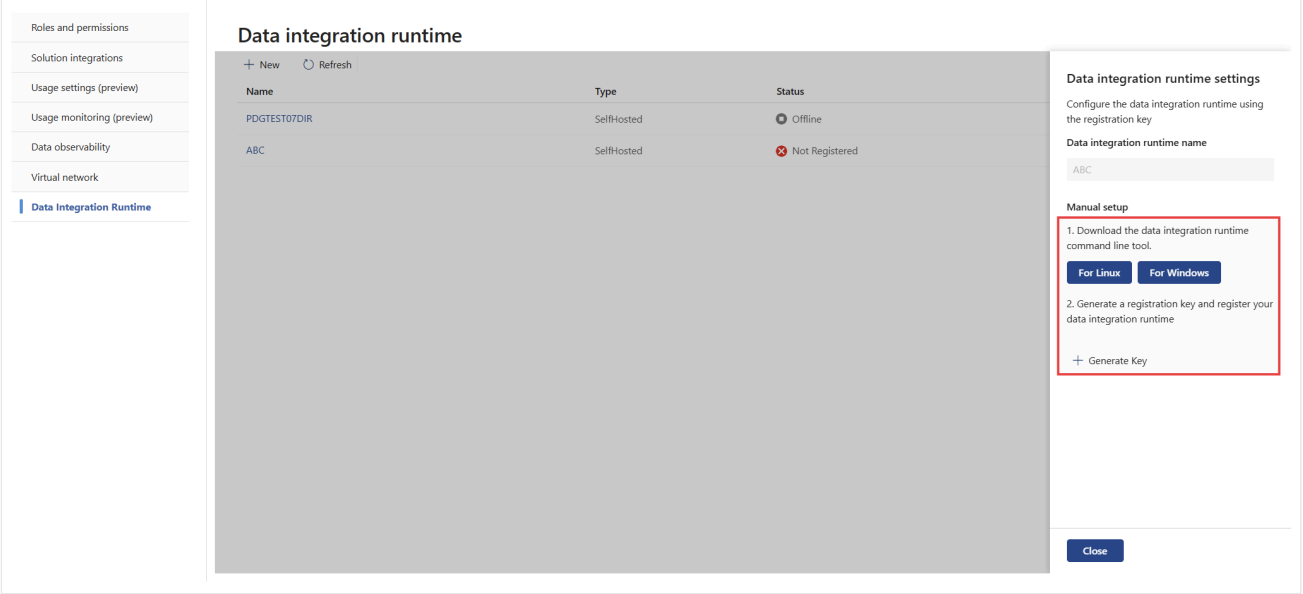
3. Select **Generate key** to generate a registration key and register your data integration runtime.



4. Copy the key value and select **Done**.

> **Tip**
>
> If needed, you can regenerate a key or revoke a generated key.

5. Select **For Linux** to download the Data Integration Runtime Command Line (DIRCTL) tool. Get details on how to install and manage DIRCTL.



6. On the machine where you want to run the DIRCTL command line, install DIRCTL from the download. DIRCTL connects to your Kubernetes cluster by context of the Kube config. If you don't specify a context, DIRCTL uses the current context. You can set the context in one of two ways:
   a. Run `kubectl` command line and execute this command to confirm the current context:
      - `kubectl config get-contexts`: List all contexts configured on the machine
      - `kubectl config current-context`: Get the current context name
      - `kubectl config use-context <name of context>`
   b. Run DIRCTL and execute `-context` to specify the context in the Kube config.
7. Run the DIRCTL Create command: `./DIRCTL create - -registration-key <registration-key copied from the portal>`. The DIRCTL Create command registers a new data integration runtime with Data Quality and initiates the creation of an application in Kubernetes as a pod specific to the registered data integration runtime. It handles the provisioning of resources and configuration essential for data integration runtime functionality while maintaining compatibility with existing system requirements.

After registration finishes, you can check the status of the data integration runtime on the **Data Integration Runtime** page in **Settings**. The status is displayed as **Online**. You can also check the status of your data integration runtime by running this command: `./DIRCTL describe`.

> **Tip**

These are the public endpoints that data integration runtime connects to that need to be allow listed:

- <purview_account_name>.purview.azure.com
- Mcr.microsoft.com
- *.data.mcr.microsoft.com

Create the namespace

1. Run the following command to create the namespace:

   Bash

   ```
   kubectl create namespace dirctl
   ```

2. Create and apply the `rolebindingv2.yaml` file with the following content:

   YAML

   ```yaml
   apiVersion: rbac.authorization.k8s.io/v1
   kind: RoleBinding
   metadata:
     name: readwrite-mongodb-secret
     namespace: dirctl
   subjects:
   - kind: ServiceAccount
     name: default
     namespace: dirctl
   roleRef:
     kind: Role
     name: secret-manager
     apiGroup: rbac.authorization.k8s.io
   ```

   Apply the file:

   Bash

   ```
   kubectl apply -f rolebindingv2.yaml
   ```

3. Create and apply the `secret-manager-role.yaml` file with the following content:

   YAML

   ```yaml
   apiVersion: rbac.authorization.k8s.io/v1
   kind: Role
   ```

```
metadata:
  namespace: dirctl
  name: secret-manager
rules:
- apiGroups: [""]
  resources: ["secrets"]
  verbs: ["get", "list", "create", "update", "patch", "delete"]
- apiGroups: [""]
  resources: ["configmaps"]
  verbs: ["get", "list", "watch"]
- apiGroups: ["apps"]
  resources: ["deployments"]
  verbs: ["get", "list", "watch","update"]
```

Apply the file:

Bash

```
kubectl apply -f secret-manager-role.yaml
```

Set up on-premises data source connection with data integration runtime

Connect to your Oracle database

Create connections by associating them with a data integration runtime instance.

- In Unified Catalog, go to **Health management** > **Data quality**.
- Select the governance domain where you created your data product with Oracle data asset.
- Select **Manage**, and then select **Connection** to set up connection for your Oracle database.

Add the following information to set up the connection:

- Enter a **Display name** for the connection.
- Enter a **Description**.
- At **Source type**, select **Oracle**.
- Select the **Data integration runtime** that you created as part of the prerequisite.
- Enter the **Host** name.
- Enter the **Port** number.
- Enter the **Service name**.
- Enter the **Schema name**.
- Select an **Authentication method**.
- Enter the **Username**.

- At **Credential**, enter the **Azure subscription**, **Azure Key Vault connection**, **Secret name**, and **Secret version**.
- Select **Submit** to complete the connection setup.

> **Tip**
>
> If you don't have all the required information, select **Save as draft** to continue later when you have the rest of the information to complete the connection setup.

This image illustrates how to create a connection:

Overview

Display name *

Oracle OnPrem

Show ID

Description

Oracle on premises data source data quality measure.

Source type *

Oracle

Data Integration Runtime *

VaibhavDIRSql (Online)

Host * ⓘ

10.0.0.17

Port * ⓘ

1521

Service Name * ⓘ

ORCL

Schema Name ⓘ

c##testuser

Authentication method *

Basic authentication

Username *

ma

Credential *

Azure subscription *

[redacted]

Key Vault connection *



Connect to your SQL Server database

Create connections by associating them with a data integration runtime instance, just like you do for Oracle. In SQL Server, a single database can contain tables belonging to multiple schemas, so you can use a single connection to scan all schemas in a single database. A connection accepts only database information but not schema. Create connections for SQL Server just like you do for other data source types.

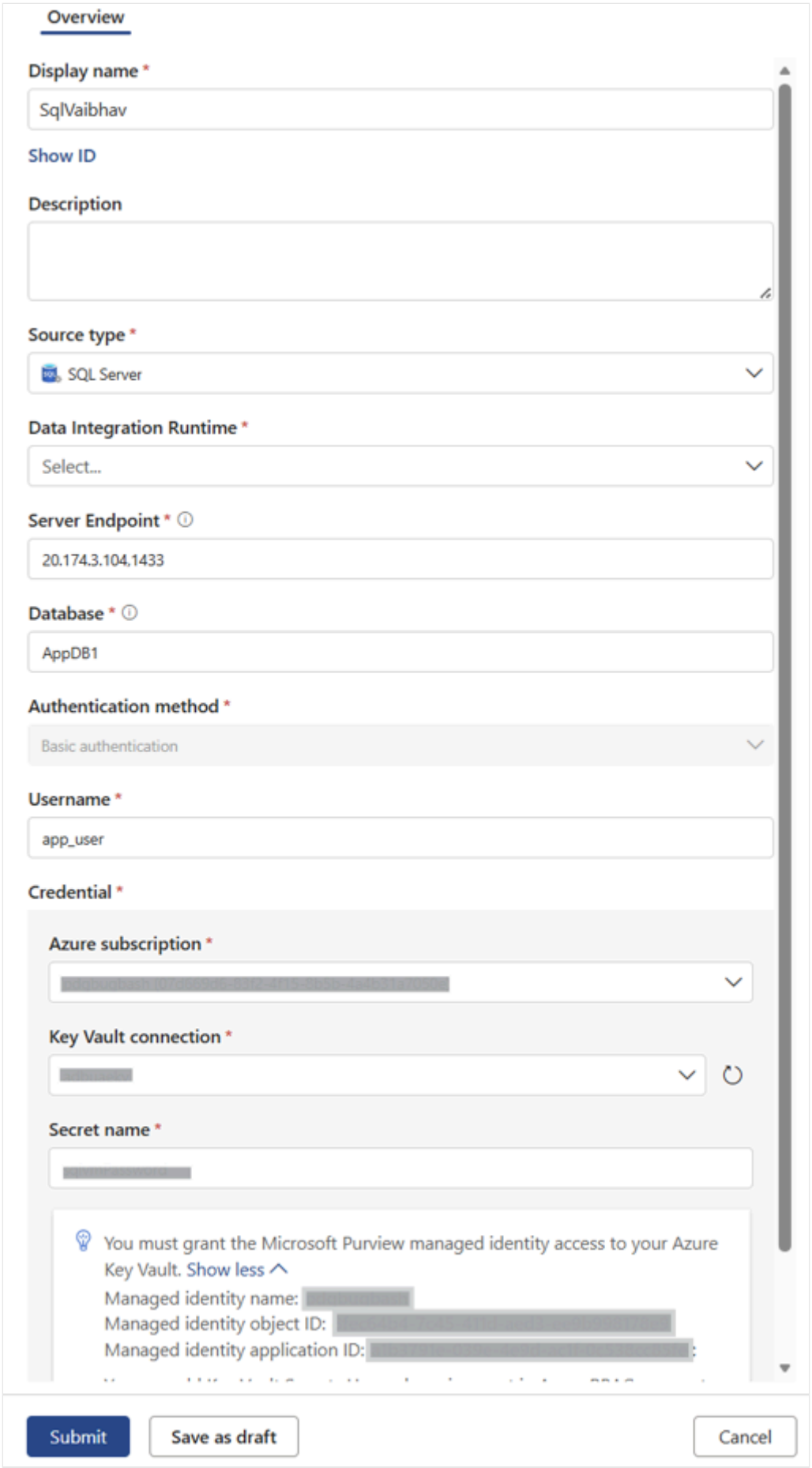
- In Unified Catalog, go to **Health management** > **Data quality**.
- Select the governance domain where you created your data product with Oracle data asset.
- Select **Manage**, and then select **Connection** to set up connection for your Oracle database.

Add the following information to set up connection successfully:

- Enter a **Display name** for the connection.
- Enter a **Description**.
- At **Source type**, select **SQL Server**.
- Select the **Data integration runtime** that you created as part of the prerequisite.
- Enter the **Server Endpoint**.
- Enter the **Database** name.
- Select an **Authentication method**.
- Enter the **Username**.
- At **Credential**, enter the **Azure subscription**, **Azure Key Vault connection**, and **Secret name**.
- Select **Submit** to complete the connection setup.

> **Tip**
>
> If you don't have all the required information, select **Save as draft** to continue later when you have the rest of the information to complete the connection setup.

This image illustrates how to create a connection:



Data quality scan

After you complete the connection setup, follow data quality profiling and scanning documents to measure and monitor data quality of Oracle and SQL Server on-premises data sources.

- Data quality overview
- Supported data sources
- Set up data source connection
- Data product quality scan
- Data asset quality scan
- Incremental data quality scan

High availability and scalability

Assign multiple nodes in the Kubernetes cluster for high availability by using the node selector during the installation of the Kubernetes supported self-hosted integration runtime. The benefits of having multiple nodes include:

- Higher availability of the self-hosted integration runtime so that it isn't a single point of failure for scans.
- More concurrent scans. Each node can handle many scan runs at the same time. You can manually scale out nodes of the Kubernetes cluster if you need more concurrent scans.
- When scanning some sources like Azure Blob, Azure Data Lake Storage Gen2, and Azure Files, each scan run can use multiple nodes to boost the scan performance. For other sources, scans run on only one of the nodes.

You can update the capabilities of the Kubernetes supported self-hosted integration runtime by manually scaling out or scaling in nodes of the Kubernetes cluster.

> **Note**
>
> You must upload all necessary drivers for scanning on each new node.

Networking requirement

Expand table

Domain name	Outbound port	Description
Public cloud: `<tenantID>-`api.purview-service.microsoft.com Azure Government: `<tenantID>-`api.purview-service.microsoft.us China: `<tenantID>-`api.purview-service.microsoft.cn	443	Required to connect to the Microsoft Purview service. If you use Microsoft Purview Private Endpoints, account private endpoint covers this endpoint.
Public cloud: `<purview_account>`.purview.azure.com Azure Government: `<purview_account>`.purview.azure.us China: `<purview_account>`.purview.azure.cn	443	Required to connect to the Microsoft Purview service. If you use Microsoft Purview Private Endpoints, account private endpoint covers this endpoint.
mcr.microsoft.com	443	Required to download images.
*.data.mcr.microsoft.com	443	Required to download images.

ⓘ **Note:** The author created this article with assistance from AI. Learn more

Last updated on 05/10/2026

Set up DIRCTL tool for self-hosted integration runtime (preview)

Data integration runtime command line (DIRCTL) is a .NET-based command line interface (CLI) tool that automates data integration runtime management within Kubernetes clusters. It simplifies complex Kubernetes operations with straightforward commands and ensures secure communication with Microsoft Purview for registration and token lifecycle management.

To manage your Microsoft Purview Kubernetes self-hosted data integration runtime locally, you need the DIRCTL command line tool. You can download this tool during the data integration runtime creation process. See the setup instructions below for Windows and Linux operating systems.

Windows setup for DIRCTL CLI

1. Download and install the Kubectl CLI tool.
2. Place the generated kube config file in the user-specific folder according to this path: `C:/Users/<user-folder>/.kube/`. Copy the file and name it **config**.

   > **Note**
   >
   > This config file is required for DIRCTL authentication on Kubernetes clusters.

3. Download the DIRCTL CLI package .zip for Windows.
4. Extract the contents to a preferred directory on your local machine.
5. Open a terminal on Windows.
6. Navigate to the extracted DIRCTL directory.
7. Use the cd command `<path-to-dirctl>` to move into the folder.
8. Run commands using `./`DIRCTL```.

Linux setup for DIRCTL CLI

1. Download and install the kubectl CLI for your Linux distribution.
2. Place the kubeconfig file, for authentication to Kubernetes, in the default kube folder: `~/.kube/`.

3. Copy your generated kubeconfig file to the user-specific directory. Make sure the file is named **config**: `mkdir -p ~/.kube cp /path/to/your/kubeconfig ~/.kube/config`

   > ⓘ **Note**
   >
   > This config file is required for DIRCTL authentication on Kubernetes clusters.

4. Download the DIRCTL CLI package for Windows or Linux.
5. Open a terminal and navigate to the DIRCTL directory: `cd /your/target/directory/DIRCTL`
6. Run commands by using: `./DIRCTL` .

DIRCTL commands

1. **DIRCTL Help**: The Help command provides a comprehensive overview of all the available commands and functionalities within the DIRCTL tool. It ensures you can use the tool effectively by providing guidance and context for your data integration tasks.

2. **DIRCTL Re-register**: `./DIRCTL re-register –registration-key <re-registration-key-value>` Use the Re-register command to re-register an existing data integration runtime resource, if it's inactive for a long time, by using a new registration key.

3. **DIRCTL Delete**: `./DIRCTL delete` When you want to deregister a data integration runtime from the system and release all associated application Kubernetes pods, the Delete command provides a streamlined solution. This command completely removes the DIR from the SQL Server Data Quality Services (DQS), ensuring that no remnants of the runtime or its application pods remain. The command also handles the termination of all running Kubernetes pods associated with the DIR application, freeing up resources within the system.

4. **DIRCTL Upgrade**: `./DIRCTL upgrade –v <image version>` The Upgrade command enables you to seamlessly upgrade the version of a data integration runtime to the latest available version specified by the user. This process ensures that the data integration runtime operates with the most up-to-date features, enhancements, and security patches, maintaining its compatibility and efficiency within the system.

5. **DIRCTL Describe:** `./DIRCTL describe` The Describe command provides detailed insights into the status and performance of the running DIR and DIRCTL applications. It checks the health of application pods, ensuring that any problems or anomalies are promptly identified. By offering visibility into the runtime environment, this functionality supports troubleshooting and proactive system maintenance.

6. **DIRCTL Log**: `./DIRCTL log download -d <relative-destination-folder-path>` Use the Log command to download logs of all active data integration runtime application pods into a specified destination folder path. This functionality is especially vital for debugging, performance monitoring, or audit activities as it ensures you have access to detailed runtime records. By retaining comprehensive logs, you can effectively identify patterns, anomalies, and potential problems, making it a key tool in system maintenance and optimization.
7. **DIRCTL Config**: `./DIRCTL config auto-update-enable <true/false>` Update the configuration for auto-update-enable to false or true.

Last updated on 04/09/2026

Troubleshoot data integration runtime and DIRCTL for Unified Catalog (preview)

This article helps you resolve common setup and configuration issues with the data integration runtime and the DIRCTL command-line tool in Microsoft Purview Unified Catalog. For setup instructions, see Set up a Kubernetes self-hosted data integration runtime. Learn how to download, install, and manage DIRCTL.

Verify prerequisites

Before deployment, verify the following prerequisites to avoid setup errors:

- Linux machine with Kubernetes installed
- Hardware requirements. Minimum specs:
  - CPU – eight cores
  - RAM – 32 GB
  - HDD – 80 GB
- Kubernetes tools
  - To access the `K8s` cluster from a Windows machine, install Kubectl.
  - To check the installation, run: `kubectl version --client`
- Permissions: You need the Data Governance Administrator role to run the DIRCTL CLI tool.

Networking: Firewall and ports

Expand table

Domain Name	Outbound Ports	Description
Public cloud: https://<purview_account>.purview.azure.com https://mcr.microsoft.com	443	Required to connect to Microsoft Purview service. If you use Microsoft Purview private endpoints, the account private endpoint covers this endpoint.

Common issues and solutions

Issue 1: Kubernetes setup isn't available

If you don't have a Kubernetes cluster set up, follow these steps to prepare the environment. You can set up Kubernetes in two ways:

1. Install Kubernetes on any Linux-based virtual machine, such as Ubuntu; or
2. Enable Hyper-V on an existing Windows Server, create a Linux-based virtual machine, and install Kubernetes.

Issue 2: Data integration runtime can't reach the database

Cause

Data integration runtime and the database might be running on different machines or networks. As a result, you might see this error message as the failure reason:

`TNS: destination host unreachable.`

Resolution

Ensure that the firewall in the database network allows inbound traffic on the configured database listener port from the IP address of the machine hosting the data integration runtime.

Issue 3: Config file isn't found in host machine

After installing Kubernetes inside the Linux-based machine, it's essential that the Kubernetes configuration file (config) is present in the `.kube` directory on the host machine.

1. Verify that the config file exists in the `.kube` directory.
2. If it's missing, copy the Kubernetes admin config file `~/.kube/config` to the required user location.

This file enables kubectl to connect to the Kubernetes cluster.

Issue 4: DIRCTL connectivity error to Microsoft Purview service endpoint

When you try to register the Kubernetes supported self-hosted data integration runtime, the DIRCTL create command throws errors.

Cause

DIRCTL can't connect to Microsoft Purview service endpoint. This issue is usually caused by network settings or a firewall.

Resolution

- Review the network topology of the DIRCTL host machine.
- Allow port 443 outbound to https://<Accountname>.purview.azure.com

Issue 5: Kubernetes connectivity error with Microsoft Purview service endpoint

When you try to register the Kubernetes supported self-hosted data integration runtime, the DIRCTL create command could run long until it eventually times out. Or, after a successful installation, the self-hosted data integration runtime status shows as unhealthy or offline in the Microsoft Purview portal.

Resolution

Verify network reachability to the Microsoft Purview service endpoint as specified in the documentation. Ensure that outbound traffic on port 443 to the Microsoft Purview service endpoint https://<AccountName>.purview.azure.com is allowed through the firewall.

Issue 6: Kubernetes self-hosted data integration runtime creation timed out error

If the DIRCTL create command doesn't complete all operations successfully within 10 minutes, it throws a timeout error. Check the status of pods under the namespaces mentioned by the kubectl command.

```
Kubectl get pods –n dirctl
```

Resolution

You should reinitiate the process of DIRCTL delete, wait for pods to destroy completely, and then do create command.

Issue 7: Registration key isn't authorized

When you try to register the Kubernetes supported self-hosted data integration runtime, the DIRCTL create command returns this error:

```
[Error] DIR failed for registration.
```

Cause

The registration key might have expired past its 24 hour validity time, or was manually revoked.

Resolution

Regenerate the key from the data integration runtime's page in the Microsoft Purview portal and register again.

Issue 8: Self-hosted data integration runtime isn't functional or has been offline for more than 14 days

The refresh token expires if data integration runtime isn't functional for more than 14 days due to any reason. In that case, reregister the data integration runtime using this command:

```
./DIRCTL create --registration-key <registration-key>
```

Issue 9: Download data integration runtime logs locally for troubleshooting

Each day's data integration runtime log file only persists for 14 days. After 14 days, the log files are deleted. Download logs locally using DIRCTL log download command if necessary:

```
e.g. ./DIRCTL log download –d <relative path>
```

Last updated on 04/09/2026

Set up data source connection for data quality in Unified Catalog

Data source connections set up the authentication needed to profile your data for a statistical snapshot, or scan your data for data quality anomalies and scoring.

Setting up data source connections is the **fourth** step in the data quality life cycle for a data asset. Previous steps are:

1. Assign users data quality steward permissions in Unified Catalog to use all data quality features.
2. Register and scan a data source in your Microsoft Purview Data Map.
3. Add data assets to a data product.

Prerequisites

1. To create connections to data assets, your users must be in the data quality steward role.
2. You need at least read access to the data source for which you're setting up the connection.

Supported multicloud data sources

Browse the supported data source document to view the list of supported data sources, including file formats for data profiling and data quality scanning, with and without virtual network support.

Currently, data quality scans can only run by using managed identity as an authentication option for Microsoft native data sources, such as Microsoft Fabric, Azure Data Lake Storage Gen2, Azure SQL, Synapse, and Azure SQL Managed Instance. **For Azure Databricks, Snowflake, and Google BigQuery, a managed identity can't be used as an authentication option**.

Data quality services run on **Apache Spark 3.4** and **Delta Lake 2.4**.

> **Important**
>
> To access these sources, either you need to set your Azure Storage sources to have an open firewall, to Allow Trusted Azure Services, or to use private endpoints follow the guideline documented in the data quality managed virtual network configuration guide.

Microsoft Purview only needs read-level permissions to discover metadata, run profiling, and execute Data Quality scans.

Set up data source connection

Follow these steps to create a new connection for the data products and data assets in a governance domain.

1. In Unified Catalog, select **Health management**, and then select **Data quality**.
2. Select a governance domain from the list.
3. From the **Manage** dropdown list, select **Connections**.
4. On **Connections**, select **New**.
5. On **Create connection**, enter a **Display name** and an optional **Description**.
6. Select a **Source type**.
7. Select one of the data sources: Azure subscription, Data Map, or enter a data source manually. Depending on which data source you choose, enter the required access details. The connection is then tested.
8. If the test connection is successful, select **Submit** to complete the connection setup.

Tip

- You can also create a connection to your resources by using private endpoints and a Microsoft Purview Data Quality managed virtual network. Learn more about setting up managed virtual networks for data quality.
- Connection setup steps vary for native connectors. Check the connection setup steps from native connectors articles to set up connection for Azure Databricks, Snowflake, Google BigQuery, and Azure Synapse connectors.
- To set up Azure Dedicated SQL Pool (formerly SQL DW) connection, select source type as **Azure SQL Database** and add `sqldatawarehouse.database.windows.net` as **endpoint** name.
- The virtual network region is auto populated from the selected source region. Find details on managing virtual network provisioning.
- For Azure SQL Managed Instance, you need to provide the port number for connection. Public endpoint port number is 3342 and private endpoint port number is 1433.

Grant Microsoft Purview permissions on the source

After you create the connection, grant Microsoft Purview managed identity permissions on your data sources to scan them:

- To scan Azure Data Lake Storage Gen2, assign the storage blob data reader role to Microsoft Purview Managed Identity. Follow the steps to assign managed identity permissions.
- To scan an Azure SQL database, assign the db_datareader role to the Microsoft Purview Managed Identity. Follow the steps to assign managed identity permissions.

Related content

- Data quality for Fabric data estate
- Data quality for Fabric Mirrored data sources
- Data quality for Fabric shortcut data sources
- Data quality for Azure Synapse serverless and data warehouses
- Data quality for Azure Databricks Unity Catalog
- Data quality for Snowflake data sources
- Data quality for Google BigQuery
- Data quality native support for Iceberg data

Next steps

1. Configure and run data profiling for an asset in your data source.
2. Set up data quality rules based on the profiling results, and apply them to your data asset.
3. Configure and run a data quality scan on a data product to assess the quality of all supported assets in the data product.
4. Review your scan results to evaluate your data product's current data quality.

Last updated on 05/08/2026

Set up data quality for Azure Databricks Unity Catalog

This article explains how to configure Microsoft Purview to scan Azure Databricks Unity Catalog metadata and set up a data quality connection for profiling and scanning your Databricks data. It covers prerequisites, supported assets, Data Map scan configuration, and connection setup. Use this guide if you're a data quality steward or admin who manages Azure Databricks data governance in Microsoft Purview.

Prerequisites

Before you begin, enable your Azure Databricks workspace for Unity Catalog. Attach the workspace to a Unity Catalog metastore. Recent workspaces are enabled automatically. For older workspaces, an account admin might need to enable Unity Catalog manually.

After Unity Catalog is enabled, complete these steps:

- Create catalogs and schemas to contain database objects like tables and volumes.
- Create managed storage locations to store the managed tables and volumes in these catalogs and schemas.
- Grant user access to catalogs, schemas, and database objects.

Workspaces that are automatically enabled for Unity Catalog provision a workspace catalog with broad privileges granted to all workspace users. The workspace catalog is a convenient starting point for trying out Unity Catalog.

Before you configure a scan or connection, store your Azure Databricks Access Token in Azure Key Vault and grant the product (service) MSI read (secret) access to the Key Vault. Microsoft Purview only needs read-level permissions to discover metadata, run profiling, and execute data quality scans.

For detailed setup instructions, see Set up and manage Unity Catalog.

Supported Unity Catalog assets

When you scan Azure Databricks Unity Catalog, Microsoft Purview supports the following assets:

- Metastore

- Catalogs
- Schemas
- Tables including the columns
- Views including the columns

When you set up a scan, you can choose to scan the entire Unity Catalog or scope the scan to a subset of catalogs.

Configure a Data Map scan for Databricks Unity Catalog

To catalog your Azure Databricks Unity Catalog data in Microsoft Purview, configure a Data Map scan.

1. Register an Azure Databricks workspace in Microsoft Purview.
2. Create a scan for the registered workspace:
   a. Enter the name of the scan.
   b. Select **Unity Catalog** as the extraction method.
   c. Connect through an integration runtime (Azure Integration runtime, Managed Virtual Network IR, or a Kubernetes-supported self-hosted integration runtime).
   d. Select **Access Token** authentication while creating a credential. For more information, see Credentials for source authentication in Data Map.
   e. Specify the Databricks SQL Warehouse HTTP path that Microsoft Purview uses to perform the scan.
3. On the **Scope your scan** page, select the catalogs you want to scan.
4. Select a scan rule set for classification. You can choose the system default, existing custom rule sets, or create a new rule set inline. For more information, see Data classification in Data Map.
5. For **Scan trigger**, choose whether to set up a schedule or run the scan once.
6. Review your scan and select **Save and run**.
7. View your scans and scan runs to verify that cataloging completed successfully.

After scanning, the scanned Unity Catalog asset is available in Microsoft Purview Unified Catalog search. For more information, see Connect to and manage Azure Databricks Unity Catalog in Microsoft Purview.

> ⓘ **Note**

Data quality scans for Microsoft native data sources (Microsoft Fabric, Azure Data Lake Storage Gen2, Azure SQL, Azure Synapse Analytics, and Azure SQL Managed Instance) use managed identity authentication. For Azure Databricks, Snowflake, and Google BigQuery, managed identity isn't available as an authentication option — use an Access Token stored in Key Vault instead.

Set up a connection for a data quality scan

After the Azure Databricks Unity Catalog scan completes, the scanned asset is ready for cataloging and governance. To run data quality scans, create a data quality connection for your Azure Databricks source and associate it with a data product in a governance domain.

> **Important**
>
> - Data quality stewards need **read only** access to Azure Databricks Unity Catalog to set up a data quality connection.
> - If public access is disabled, select the **Allow trusted Microsoft services** checkbox for Key Vault. This requirement applies only to Key Vault, not to your Azure Databricks workspace.
> - Virtual network support is generally available to all supported Azure regions. It's temporarily included in the Data Governance SKUs to maintain flexibility during this phase. Virtual network pricing isn't yet available to include in billing.

1. In the Microsoft Purview portal, open **Unified Catalog**.
2. Under **Health management**, select **Data quality**.
3. Select a governance domain from the list, then select **Connections** from the **Manage** dropdown list.
4. Configure the connection on the **Connections** page:
    - Add connection name and description.
    - Select source type **Azure Databricks**.
    - Select Azure subscription.
    - Select workspace URL.
    - Add Databricks metastore ID.
    - Select Unity catalog as extraction method.

- Select HTTP path.
- Select unity catalog name.
- Select schema name.
- Select table name.
- Select authentication method: **Access Token**.
  - Add Azure subscription.
  - Key Vault connection.
  - Secret name.
  - Secret version.
- Select the **Enable managed V-Net** checkbox if your Databricks is running in a virtual network.
- Region is selected automatically.
- Create a new virtual network if a virtual network storage hasn't yet been created.

5. If your Databricks storage is in a virtual network, you can't test the data quality connection. Otherwise, select **Test connection** to validate your configuration.

> ⓘ **Note**
>
> - The fully qualified domain name (FQDN) for a data asset follows a pattern like `databricks://(metastore-id)/catalogs/(catalog-name)/schemas/(schema-name)/tables/(table-name)`. You can find the FQDN details for your Azure Databricks data asset on the Data Map asset page.
> - If your connection parameters don't match the FQDN, your connection might still work but you see a connection error on the data quality overview page. Ensure that you fill in all corresponding fields correctly.

Create connection

Overview

Databricks UC

Show ID

Description

Setup databricks UC connection in with or without vNet

Microsoft Purview is located in the region **East US**. Any sourced data through this for DQ will be processed in the same region.

Source type *

Azure Databricks

Azure subscription *

PDG_Test_07 (27ef0436-f648-4bad-be15-3e872e16318b)

Workspace name *

dbricko7

Workspace URL

adb-349180787161256.16.azuredatabricks.net

MetastoreId *

Add metastore id

Extraction method

Unity catalog

HTTP path *

Add http path

Catalog name *

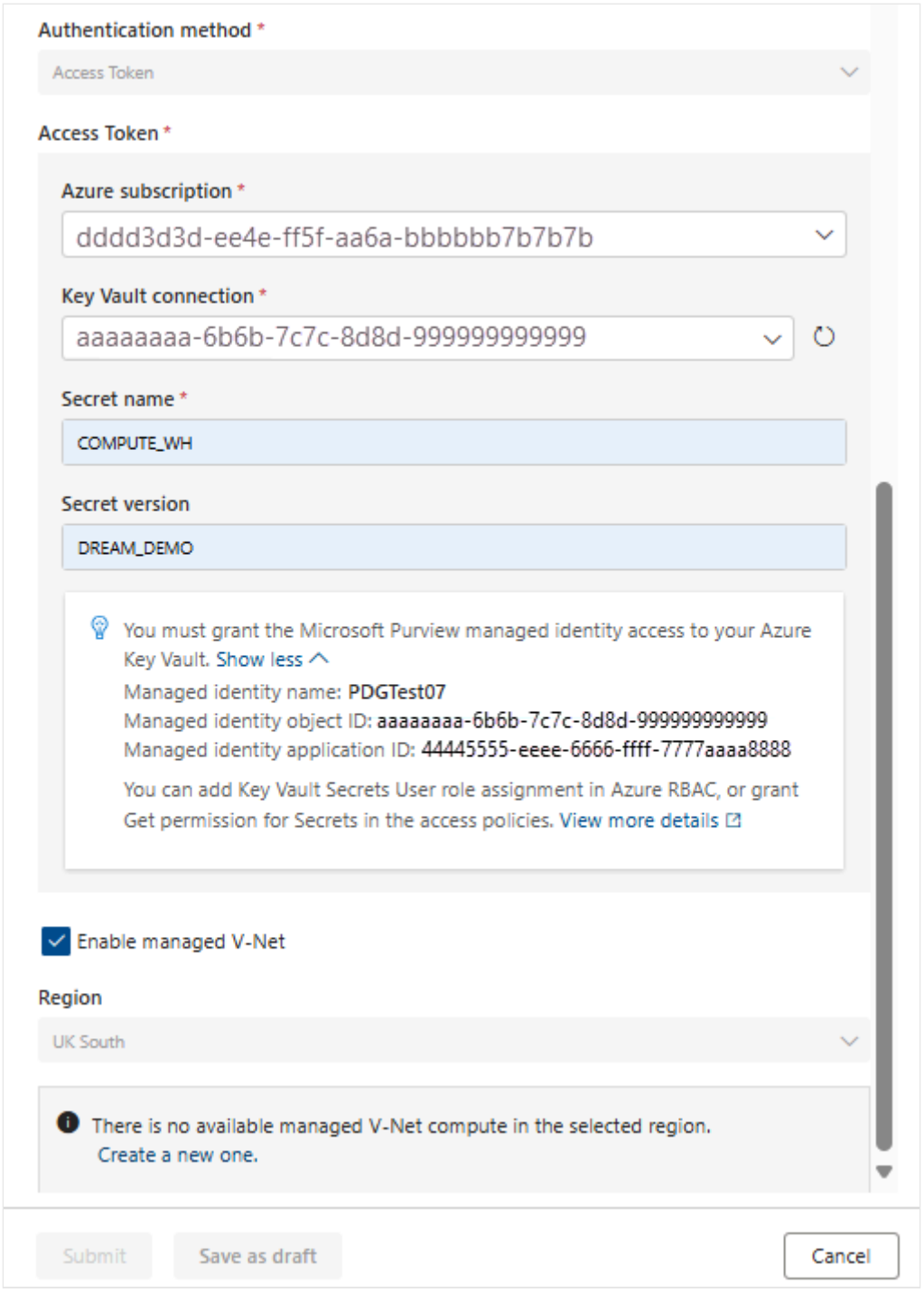
Unity catalog name

Schema name (for testing connection only) *

Add schema name

Table name (for testing connection only) *

Add table name



Run profiling and data quality scans

After you set up the connection, you can profile your data, create and apply rules, and run a data quality scan for your data in Azure Databricks Unity Catalog databases. For step-by-step guidance, see Configure and run data profiling and Configure and run a data quality scan.

> **Note**
>
> The default XS SQL Warehouse with one node isn't suitable for production use with medium or large datasets. Review Azure Databricks warehouse behavior and adopt appropriate

vertical scaling (XS, S, M, L, XL) and horizontal scaling (8, 16, 32, 64 nodes). Start with M (1-8) SQL Warehouse.

Related content

Learn more about Azure Databricks and data quality setup:

- Set up Azure Databricks Unity Catalog
- Data quality troubleshooting
- Set up a data quality connection
- Configure virtual network
- Azure Databricks warehouse behavior
- Create an Azure Databricks SQL warehouse

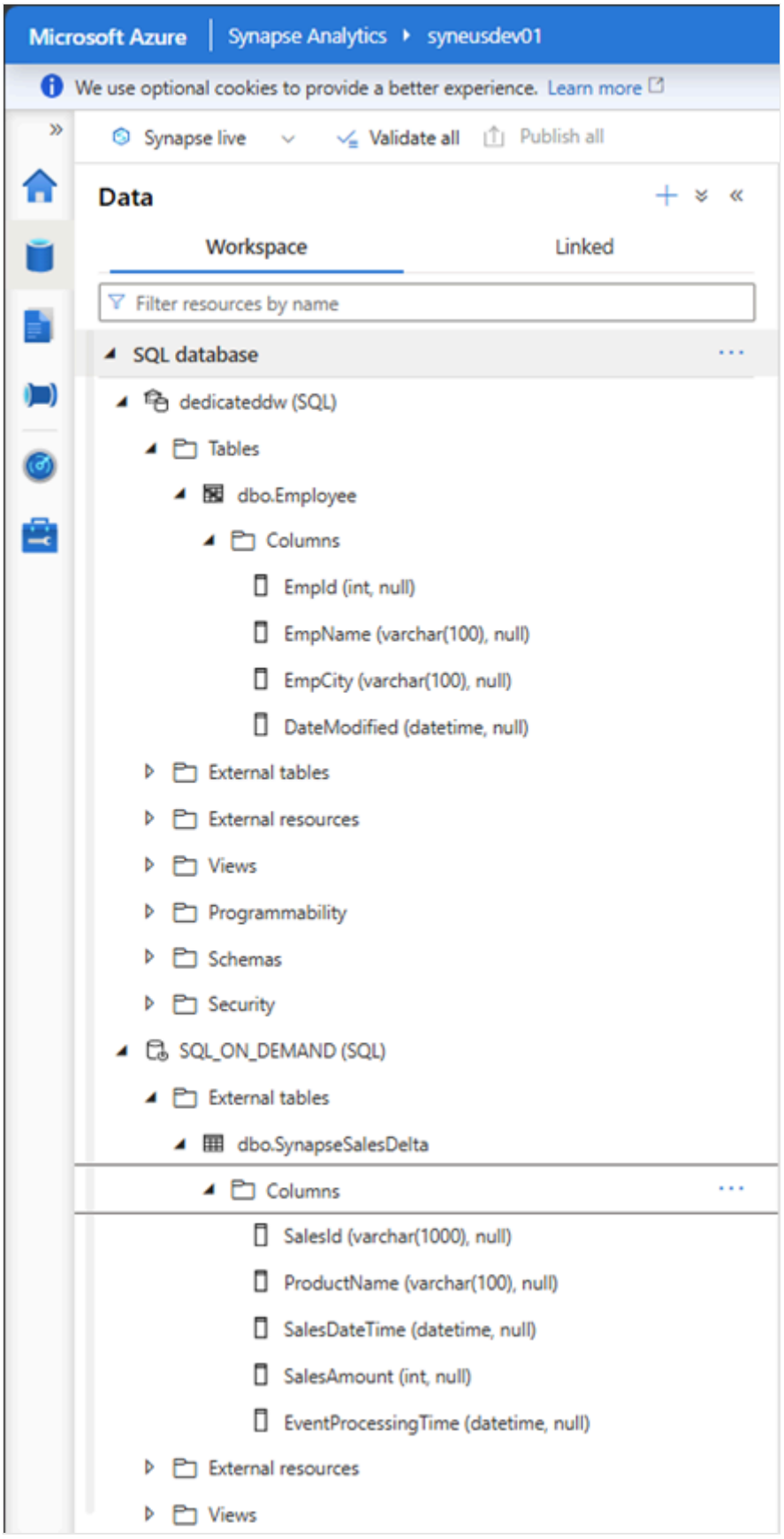
Last updated on 05/26/2026

Set up data quality for Microsoft Azure Synapse serverless and data warehouse

This article explains how to configure data quality scans for Azure Synapse Analytics assets in Microsoft Purview Unified Catalog. You'll learn how to set up Data Map scans, create connections, and run data profiling and quality scans for both Synapse dedicated data warehouse and Synapse serverless SQL pools.

Microsoft Azure Synapse Analytics is an enterprise analytics service that accelerates time to insight across data warehouses and big data systems. For more details, see the Azure Synapse Analytics documentation.

The following example shows a Synapse workspace with a Dedicated Synapse Data Warehouse (DWH) table called **EMPLOYEE** and a Serverless Database (SQL_ON_DEMAND) with a **SynapseSalesDelta** table.



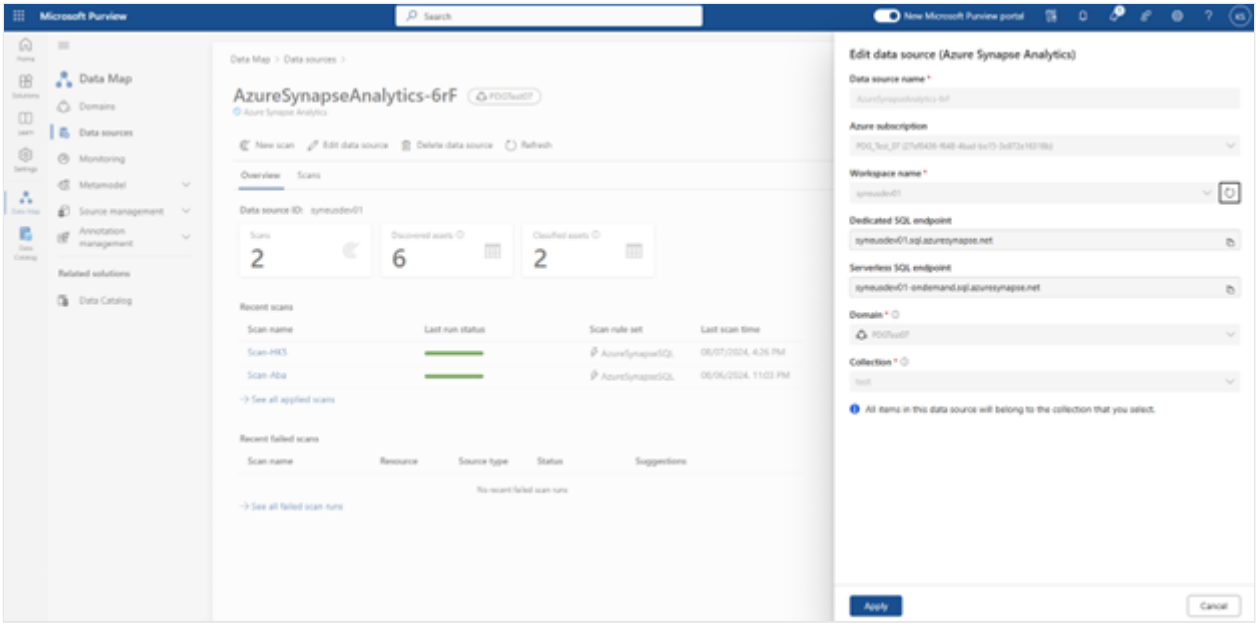
After you scan Azure Synapse assets, the scanned assets appear in Microsoft Purview. The following example shows an Employee table on a Synapse Analytics Dedicated instance.

Set up data quality for Azure Synapse Analytics dedicated SQL pool

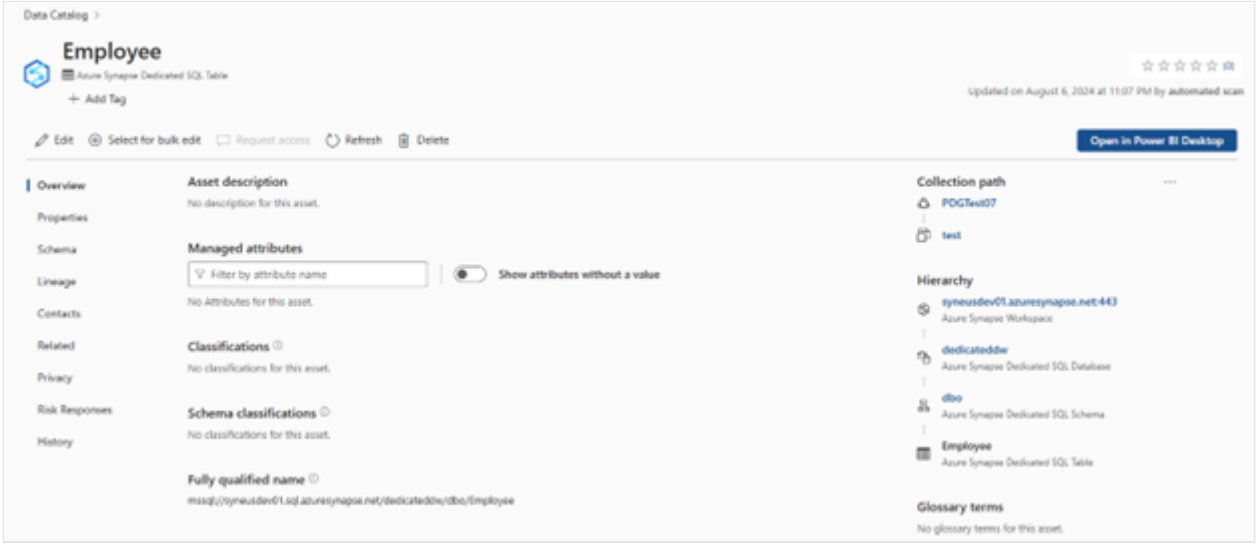
To configure data quality for a Synapse dedicated SQL pool, set up a Data Map scan, create a connection, and then run profiling and data quality scans.

Set up Data Map scan for Synapse dedicated SQL

To scan Azure Synapse Analytics dedicated SQL pool, see Register and scan a Synapse workspace. To grant the necessary managed identity permissions on the dedicated SQL instance, see Use a managed identity for dedicated SQL databases.



After the Data Map scan completes, the Synapse dedicated SQL assets appear in Microsoft Purview Unified Catalog. The following example shows an Employee table on a Synapse Analytics dedicated instance:



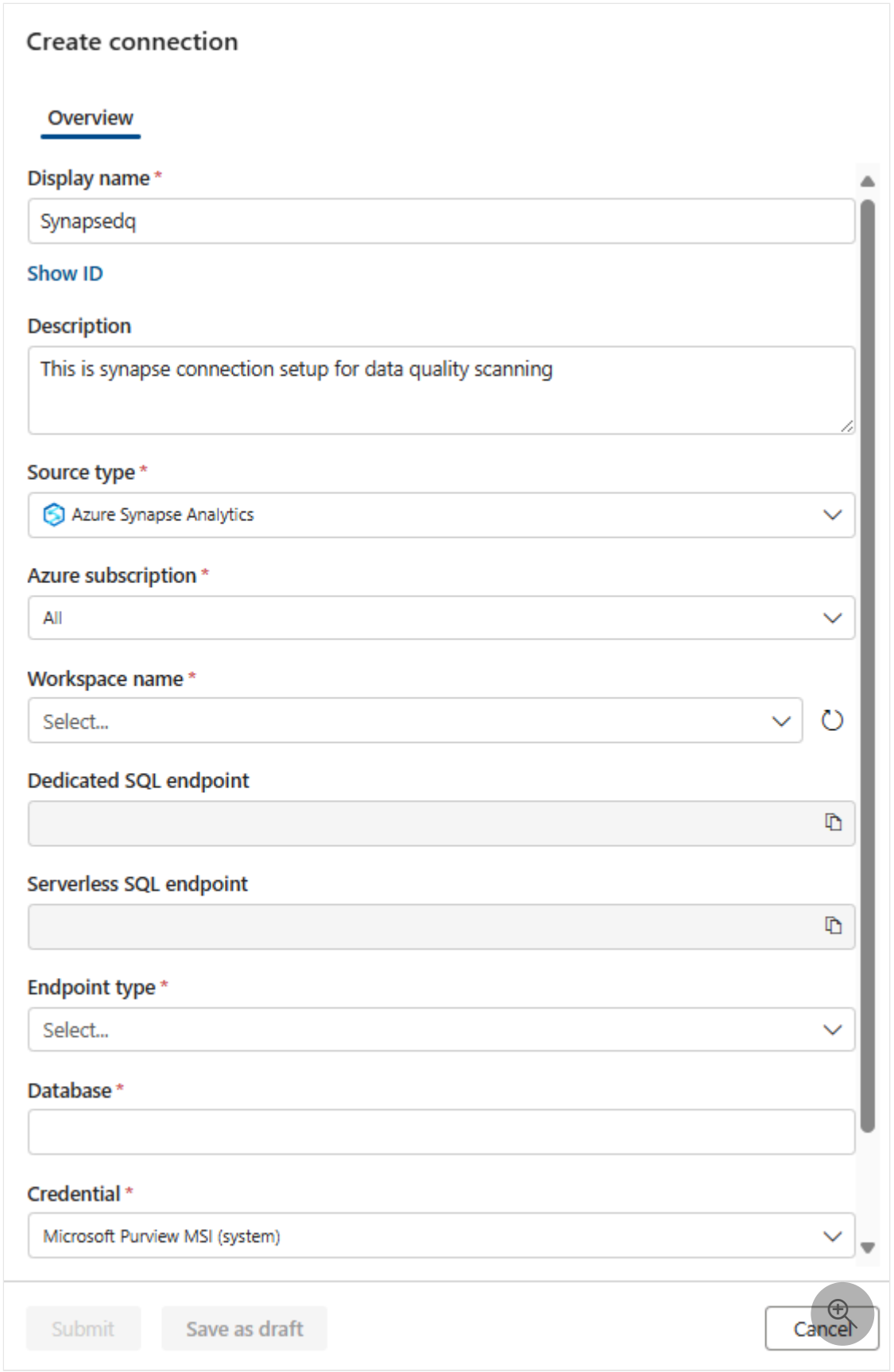
Set up connection to your Synapse dedicated data warehouse

After the Data Map scan completes and the dedicated SQL asset appears in Unified Catalog, the asset is ready for cataloging and governance.

> **Important**
>
> - Data Quality stewards need read-only access to the Synapse dedicated data warehouse to set up the data quality connection.
> - If your Synapse data source is behind a private endpoint, you must first enable managed virtual networks. See Set up managed virtual networks. In managed virtual network setups, you can't test the connection.

Associate the scanned asset to the data product in a governance domain. At the Data Quality tab, add a new Azure SQL Database connection and enter the database name manually.

1. In Unified Catalog, go to **Health management** > **Data quality** and select a governance domain.
2. On the governance domain's details page, select **Manage**, then select **Connections**.
3. On the **Connections** page, select **New** and configure connection with the following information, as seen in the example below:
    - Add connection name and description.
    - Select source type **Azure Synapse Analytics**.
    - Select **Azure subscription**.
    - Select **Workspace name**.
    - Select **Dedicated SQL endpoint**.
    - Select **serverless SQL endpoint**.
    - Select **Endpoint type**.
    - Select **Database**.
    - Add MSI as Credential.



4. Test the connection. After configuring the data source connection and successfully testing it, you can proceed to configure and run data profiling and data quality scans.

Profiling and data quality scanning for data in Synapse dedicated data warehouse

After you set up the connection, you can profile your data, create and apply rules, and run a data quality scan for your data in Synapse dedicated data warehouse.

> **Important**
>
> Before you run profiling or data quality scans, verify that your dedicated SQL pool configuration and concurrency limits can support these jobs. Data quality assessment jobs open connections on the dedicated SQL pool and might fail if the instance is underprovisioned or exceeds concurrency limits. Concurrency limits might lead to job termination. Higher DW limits (such as DWU 1000) provide the capacity to run the queries. For more information, see **Memory and concurrency limits**.

Use the following articles for step-by-step guidance:

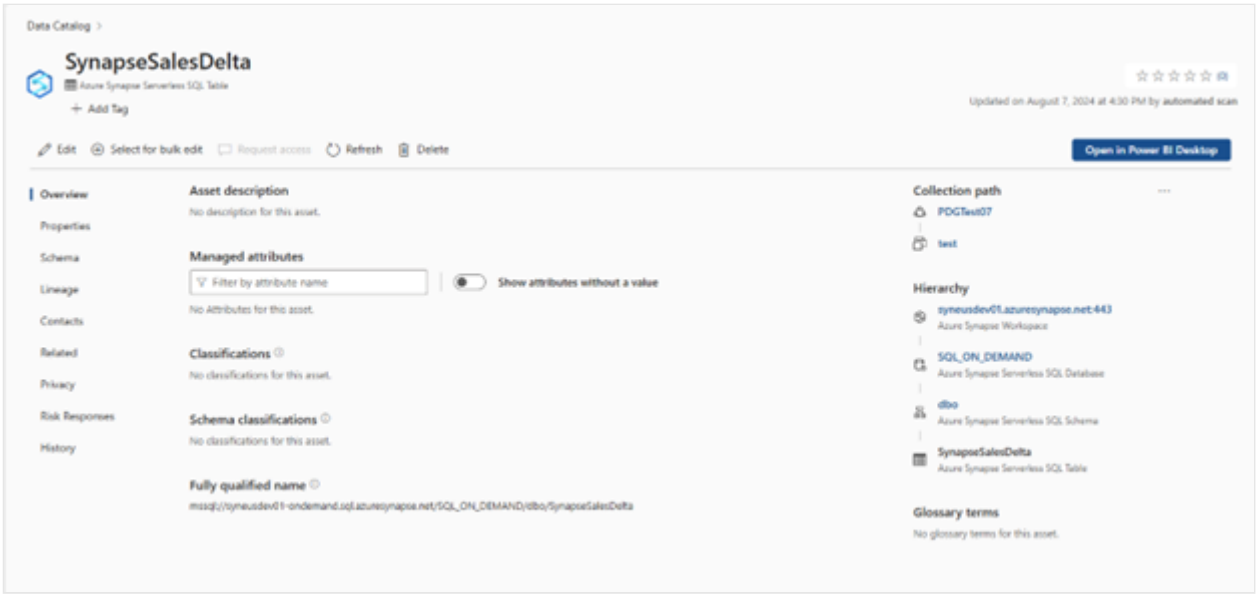
- How to configure and run data profiling of your data
- How to configure and run data quality scan

Set up data quality for Azure Synapse Analytics serverless SQL pool

To configure data quality for a Synapse serverless SQL pool, set up a Data Map scan, create a connection, and then run profiling and data quality scans.

Set up Data Map scan for Synapse serverless SQL

To scan Azure Synapse Analytics serverless SQL pool, see Register and scan a Synapse workspace. To grant the necessary managed identity permissions, see Use a managed identity for dedicated SQL databases. After the Data Map scan completes, the serverless assets appear in Unified Catalog.



Set up connection to your Synapse serverless SQL pool

After the Data Map scan completes and the serverless SQL asset appears in Unified Catalog, the asset is ready for cataloging and governance.

> **ⓘ Important**
>
> - Data Quality stewards need read-only access to the Synapse dedicated data warehouse to set up the data quality connection.
> - If your Synapse data source is behind a private endpoint, you must first enable managed virtual networks. See Set up managed virtual networks.
> - In Synapse serverless setup, the external table points to Delta formatted data stored in ADLS Gen2.
> - The Synapse connector only detects and supports **sql.azuresynapse.net**. If the Fully Qualified Name (FQN) generated by your Data Map scan contains **database.windows.net**, the Synapse connection for data quality scan will fail.

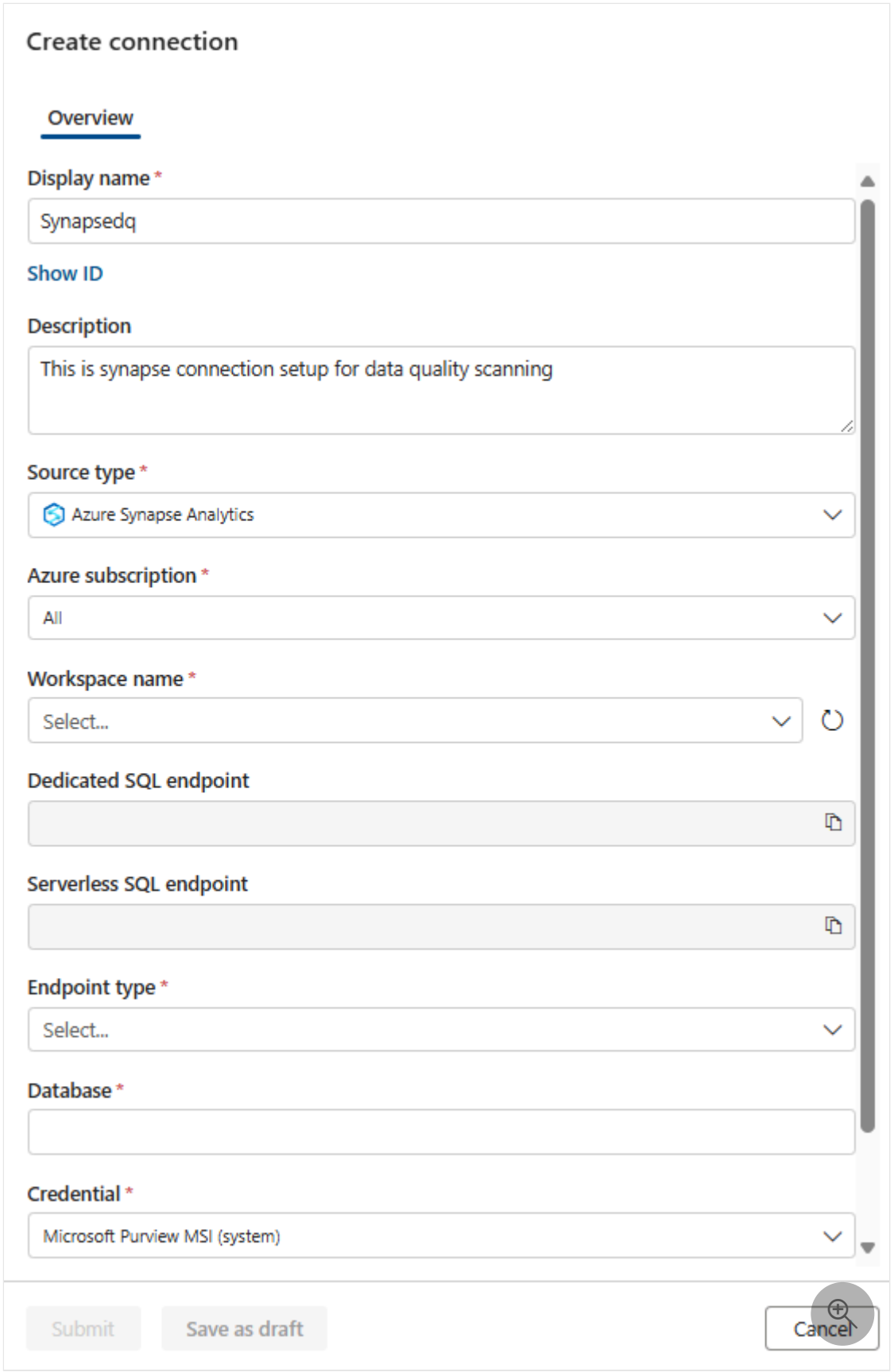
Associate the scanned asset to the data product in a governance domain. In **Data Quality**, add a new SQL Database connection and enter the database name manually.

1. In Unified Catalog, go to **Health management** > **Data quality** and select a governance domain.

2. On the governance domain's details page, select **Manage**, then select **Connections**.

3. On the **Connections** page, select **New** and configure connection with the following information, as seen in the example below

- Add connection name and description.
- Select source type **Azure Synapse Analytics**.
- Select **Azure subscription**.
- Select **Workspace name**.
- Select **Dedicated SQL endpoint**.
- Select **serverless SQL endpoint**.
- Select **Endpoint type**.
- Select **Database**.
- Add MSI as Credential.



4. Test the connection. After configuring the data source connection and successfully testing it, you can proceed to configure and run data profiling and data quality scans.

Profiling and data quality scanning for data in Synapse serverless

After you set up the connection, you can profile your data, create and apply rules, and run a data quality scan for your data in Synapse serverless SQL pool.

> **Important**
>
> Data quality assessments and profiling run on Spark in the background, where each Spark node opens a connection. For Azure Synapse serverless SQL, no concurrency limits apply. Performance depends on the Delta Parquet optimizations on your ADLS Gen2 instance. The engine works with external lakehouse sources such as Delta format tables.

Use the following articles for step-by-step guidance:

- How to configure and run data profiling of your data
- How to configure and run data quality scan

Performance and connection reference for Synapse serverless

Use the following references for serverless performance and connection details:

- Memory and concurrency limits for Azure Synapse Analytics
- Set up data source connection for data quality

Last updated on 05/27/2026

Set up data quality for Fabric Lakehouse data

This article explains how to register, scan, and run data quality assessments on Fabric Lakehouse data using Microsoft Purview. Before you begin, review the prerequisites.

Overview of Fabric OneLake

Fabric OneLake is a single, unified, logical data lake for your whole organization. A data lake processes large volumes of data from various sources. Like Microsoft OneDrive, OneLake comes automatically with every Microsoft Fabric tenant and is designed to be the single place for all your analytics data. OneLake brings customers:

- One data lake for the entire organization
- One copy of data for use with multiple analytical engines

OneLake aims to give you the most value possible out of a single copy of data without data movement or duplication. You no longer need to copy data just to use it with another engine or to break down silos so you can analyze the data with data from other sources. You can use Microsoft Purview to catalog fabric data estate and measure data quality to govern and drive improvement action.

You can use shortcut for referencing to data that is stored in other file locations. These file locations can be within the same workspace or across different workspaces, within OneLake or external to OneLake in Azure Data Lake Storage (ADLS), Amazon Web Services (AWS) S3, or Dataverse with more target locations coming soon. Data source location doesn't matter that much, OneLake shortcuts make files and folders appear as if they are stored locally. When teams work independently in separate workspaces, shortcuts enable you to combine data across different business groups and domains into a virtual data product to fit a user's specific needs.

You can use mirroring to bring data from various sources together into Fabric. Mirroring in Fabric is a low-cost and low-latency solution to bring data from various systems together into a single analytics platform. You can continuously replicate your existing data estate directly into Fabric's OneLake, including data from Azure SQL Database, Azure Cosmos DB, and Snowflake. With the most up-to-date data in a queriable format in OneLake, you can now use all the different services in Fabric. For example, running analytics with Spark, executing notebooks, data engineering,

visualizing through Power BI Reports, and more. The Delta tables created from mirrored data can then be used throughout Fabric, allowing users to accelerate their journey into Fabric.

Register Fabric OneLake as a data source

To configure Data Map scan, you need to first register data source that you want to scan. To scan a Fabric workspace, there are no changes to the existing experience to register a Fabric tenant as a data source. To register a new data source, follow these steps:

- In the Microsoft Purview portal, go to **Data Map**.
- Select **Register**.
- On **Register** sources, select **Fabric**.

For setup instructions, see same-tenant Fabric scan setup and cross-tenant Fabric scan setup.

Set up a Data Map scan for Fabric Lakehouse

To scan Lakehouse subartifacts, you don't need to change your existing experience in Data Map to set up a scan. A viewer role is sufficient to scan metadata and item-level lineage.

Currently, only service principal is supported as an authentication method. Managed identity (MSI) support is still in the backlog.

For setup instructions, see same-tenant Fabric scan setup and cross-tenant Fabric scan setup.

Set up connection for Fabric Lakehouse scan

After you register Fabric Lakehouse as a source, you can select Fabric from the list of your registered data sources in Data Map and select **New scan**. Add a **Data source ID**, then follow these steps:

1. Create a security group and a service principal.

2. Add both this service principal and the Purview Managed Identity to this security group, then specify that security group in the Fabric tenant Admin API settings.

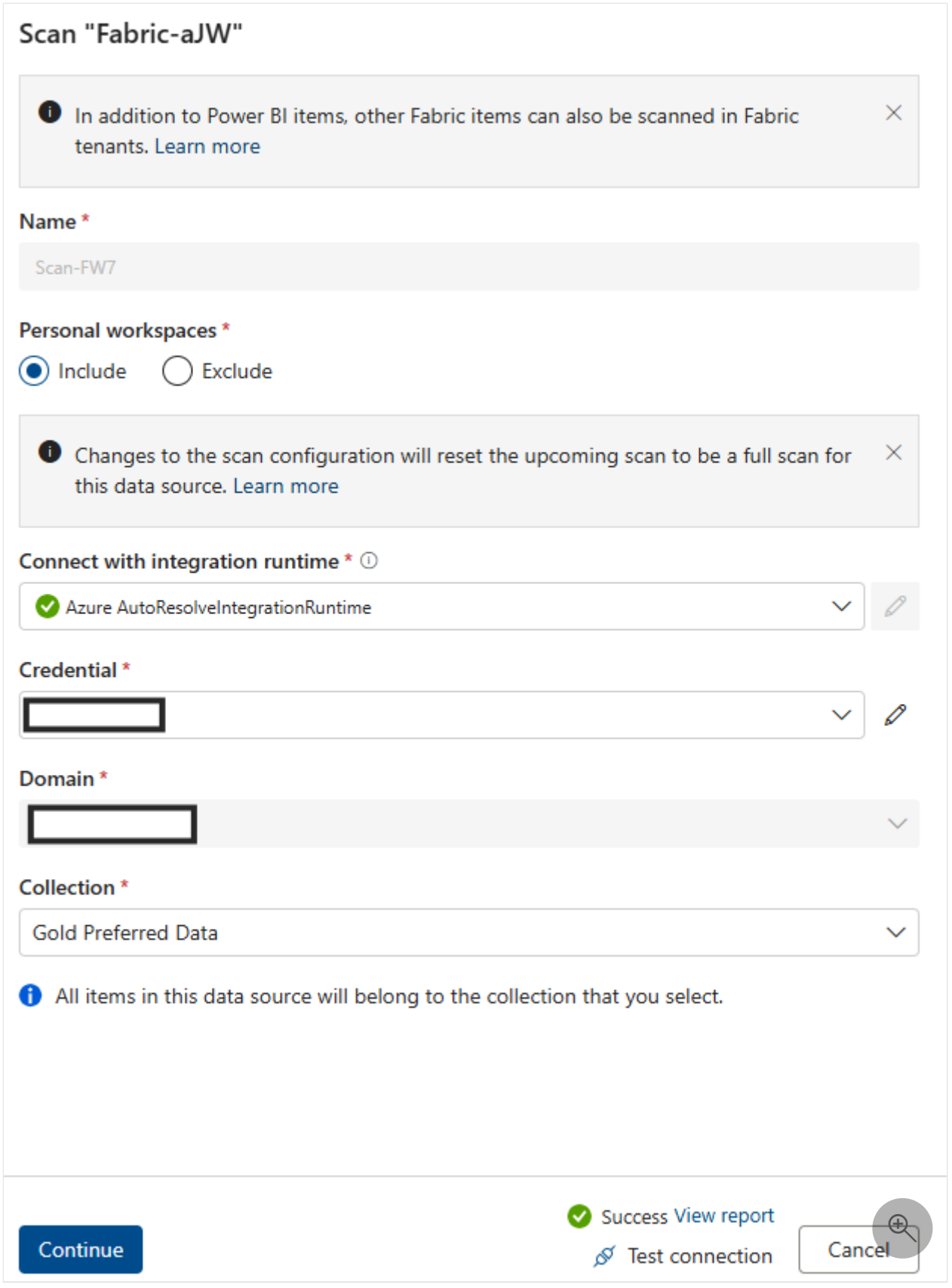
3. Associate the security group with Fabric tenant.
   a. Sign in to the Fabric admin portal.
   b. Select the Tenant settings page. **You need to be a Fabric Admin to see the tenant settings page.**
   c. Select Admin API settings > Allow service principals to use read-only admin APIs.

d. Select Specific security groups.
e. Select Admin API settings > Enhance admin APIs responses with detailed metadata and Enhance admin APIs responses with DAX and mashup expressions > Enable the toggle to allow Data Map to automatically discover the detailed metadata of Fabric datasets as part of its scans. **After you update the Admin API settings on your Fabric tenant, wait about 15 minutes before registering a scan and test connection.**

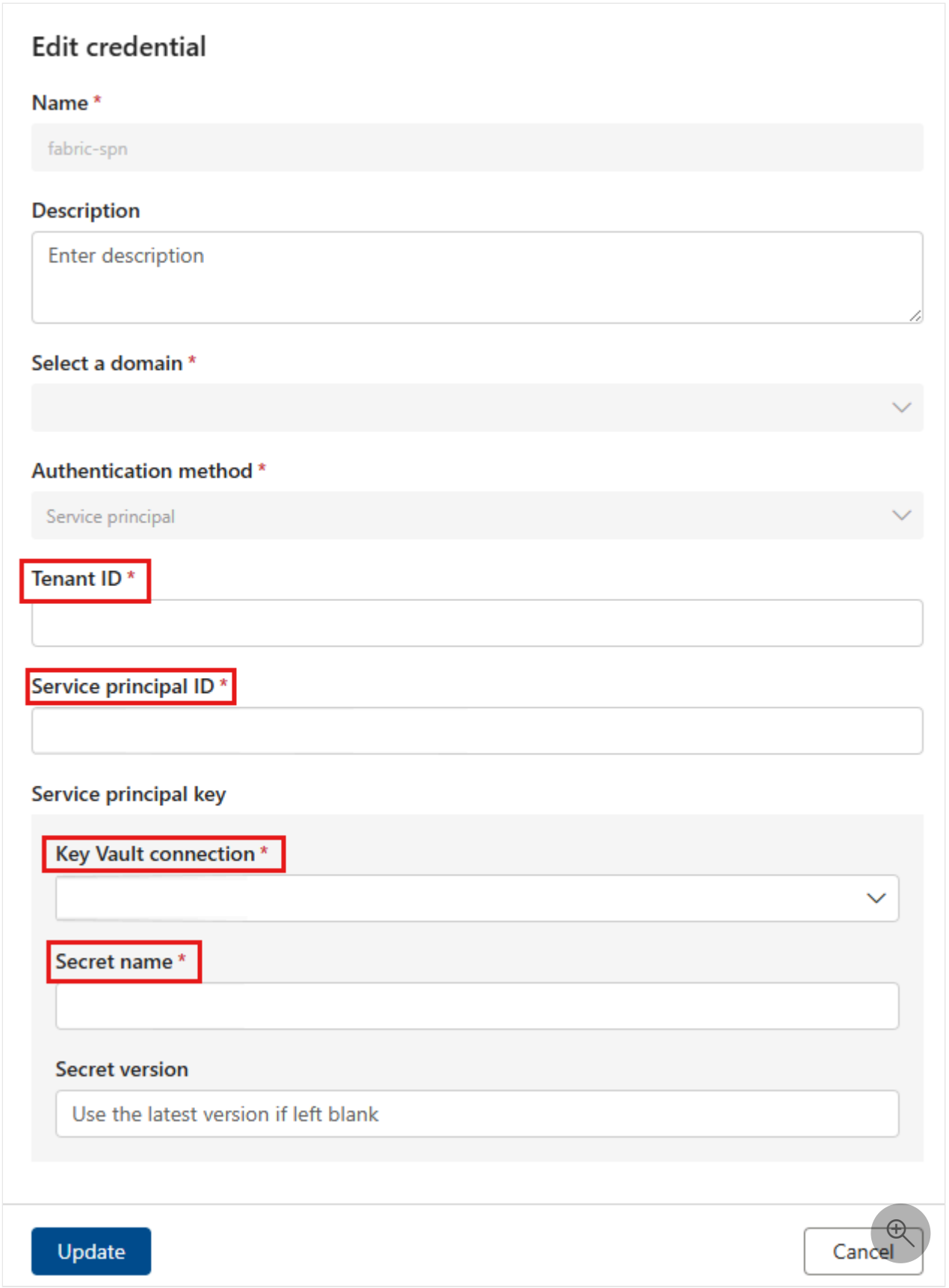
4. Provide Admin API settings read-only API permission to the security group that contains the service principal and Purview Managed Identity.

5. Add **SPN** into **Credential** field.

6. Add Azure resource name.



7. Add **Tenant ID**.
8. Add **Service Principle ID**.
9. Add **Key Vault connection**.
10. Add **Secret name**.



After you finish the Data Map scan, locate a Lakehouse instance in Unified Catalog.

1. In the Microsoft Purview portal, open **Unified Catalog**.
2. Select **Discovery**, then **Data assets**.
3. On the **Data assets** page, select **Microsoft Fabric**.
4. Select **Fabric workspaces**, then select a workspace from the list.

5. On the workspace's page, find the Lakehouse instance under **Item name**.

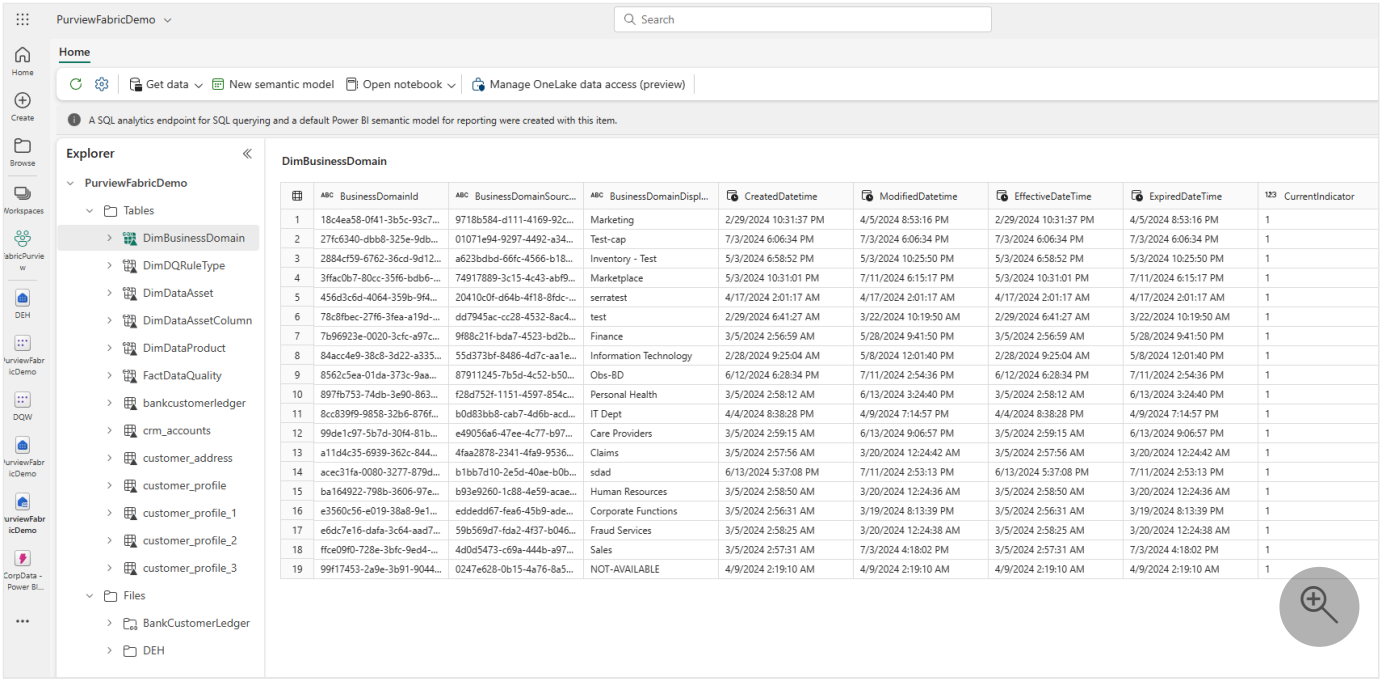
To browse Lakehouse tables:

1. On a workspace page, select the item name **Tables**.
2. Select a Lakehouse table asset listed under **Item name**.
3. View the asset's details page to find metadata such as schema, lineage, and properties.

Review Fabric Lakehouse data quality scan prerequisites

Before you run a data quality scan on Fabric Lakehouse data, complete the following prerequisites:

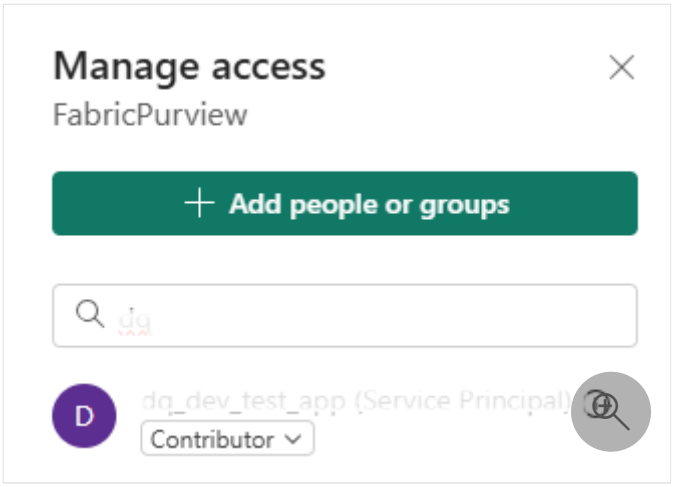
1. Shortcut, mirror, or load your data to Fabric Lakehouse in delta format.



	BusinessDomainId	BusinessDomainSourc...	BusinessDomainDispl...	CreatedDatetime	ModifiedDatetime	EffectiveDateTime	ExpiredDateTime	CurrentIndicator
1	18c4ea58-0f41-3b5c-93c7...	9718b584-d111-4169-92c...	Marketing	2/29/2024 10:31:37 PM	4/5/2024 8:53:16 PM	2/29/2024 10:31:37 PM	4/5/2024 8:53:16 PM	1
2	27fc6340-dbb8-325e-9db...	01071e94-9297-4492-a34...	Test-cap	7/3/2024 6:06:34 PM	7/3/2024 6:06:34 PM	7/3/2024 6:06:34 PM	7/3/2024 6:06:34 PM	1
3	2884cf59-6762-36cd-9d12...	a623bdbd-66fc-4566-b18...	Inventory - Test	5/3/2024 6:58:52 PM	5/3/2024 10:25:50 PM	5/3/2024 6:58:52 PM	5/3/2024 10:25:50 PM	1
4	3ffac0b7-80cc-35f6-bdb6-...	74917889-3c15-4c43-abf9...	Marketplace	5/3/2024 10:31:01 PM	7/11/2024 6:15:17 PM	5/3/2024 10:31:01 PM	7/11/2024 6:15:17 PM	1
5	456d3c6d-4064-359b-9f4...	20410c0f-d64b-4f18-8fdc-...	serratest	4/17/2024 2:01:17 AM	4/17/2024 2:01:17 AM	4/17/2024 2:01:17 AM	4/17/2024 2:01:17 AM	1
6	78c8fbec-27f6-3fea-a19d-...	dd7945ac-cc28-4532-8ac4...	test	2/29/2024 6:41:27 AM	3/22/2024 10:19:50 AM	2/29/2024 6:41:27 AM	3/22/2024 10:19:50 AM	1
7	7b96923e-0020-3cfc-a97c...	9f88c21f-bda7-4523-bd2b...	Finance	3/5/2024 2:56:59 AM	5/28/2024 9:41:50 PM	3/5/2024 2:56:59 AM	5/28/2024 9:41:50 PM	1
8	84acc4e9-38c8-3d22-a335...	55d373bf-8486-4d7c-aa1e...	Information Technology	2/28/2024 9:25:04 AM	5/8/2024 12:01:40 PM	2/28/2024 9:25:04 AM	5/8/2024 12:01:40 PM	1
9	8562c5ea-01da-373c-9aa...	87911245-7b5d-4c52-b50...	Obs-BD	6/12/2024 6:28:34 PM	7/11/2024 2:54:36 PM	6/12/2024 6:28:34 PM	7/11/2024 2:54:36 PM	1
10	897fb753-74db-3e90-863...	f28d752f-1151-4597-854c...	Personal Health	3/5/2024 2:58:12 AM	6/13/2024 3:24:40 PM	3/5/2024 2:58:12 AM	6/13/2024 3:24:40 PM	1
11	8cc839f9-9858-32b6-876f...	b0d83bb8-cab7-4d6b-acd...	IT Dept	4/4/2024 8:38:28 PM	4/9/2024 7:14:57 PM	4/4/2024 8:38:28 PM	4/9/2024 7:14:57 PM	1
12	99de1c97-5b7d-30f4-81b...	e49056a6-47ee-4c77-b97...	Care Providers	3/5/2024 2:59:15 AM	6/13/2024 9:06:57 PM	3/5/2024 2:59:15 AM	6/13/2024 9:06:57 PM	1
13	a11d4c35-6939-362c-844...	4faa2878-2341-4fa9-9536...	Claims	3/5/2024 2:57:56 AM	3/20/2024 12:24:42 AM	3/5/2024 2:57:56 AM	3/20/2024 12:24:42 AM	1
14	acec31fa-0080-3277-879d...	b1bb7d10-2e5d-40ae-b0b...	sdad	6/13/2024 5:37:08 PM	7/11/2024 2:53:13 PM	6/13/2024 5:37:08 PM	7/11/2024 2:53:13 PM	1
15	ba164922-798b-3606-97e...	b93e9260-1c88-4e59-acae...	Human Resources	3/5/2024 2:58:50 AM	3/20/2024 12:24:36 AM	3/5/2024 2:58:50 AM	3/20/2024 12:24:36 AM	1
16	e3560c56-e019-38a8-9e1...	eddedd67-fea6-45b9-ade...	Corporate Functions	3/5/2024 2:56:31 AM	3/19/2024 8:13:39 PM	3/5/2024 2:56:31 AM	3/19/2024 8:13:39 PM	1
17	e6dc7e16-dafa-3c64-aad7...	59b569d7-fda2-4f37-b046...	Fraud Services	3/5/2024 2:58:25 AM	3/20/2024 12:24:38 AM	3/5/2024 2:58:25 AM	3/20/2024 12:24:38 AM	1
18	ffce09f0-728e-3bfc-9ed4-...	4d0d5473-c69a-444b-a97...	Sales	3/5/2024 2:57:31 AM	7/3/2024 4:18:02 PM	3/5/2024 2:57:31 AM	7/3/2024 4:18:02 PM	1
19	99f17453-2a9e-3b91-9044...	0247e628-0b15-4a76-8a5...	NOT-AVAILABLE	4/9/2024 2:19:10 AM	4/9/2024 2:19:10 AM	4/9/2024 2:19:10 AM	4/9/2024 2:19:10 AM	1

> **Important**
>
> If you add new tables, files, or new sets of data to Fabric Lakehouse through mirroring or shortcut, run a Data Map scope scan to catalog those new sets of data before adding those data assets to a data product for data quality assessment.

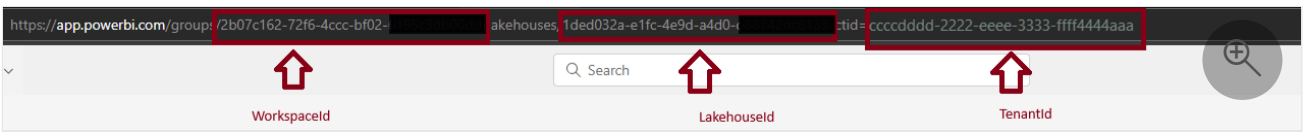
2. Grant Contributor right to your workspace for Purview MSI



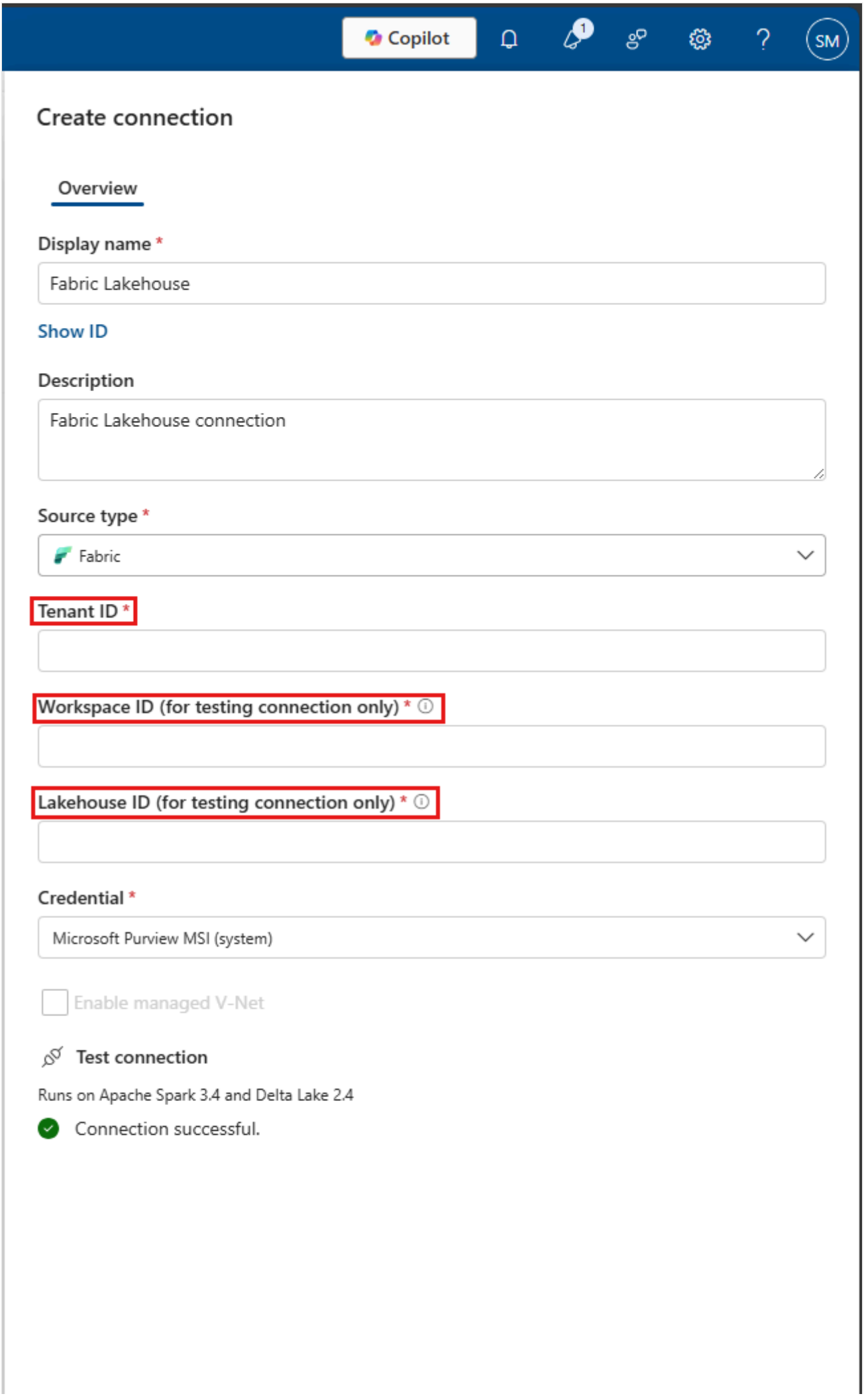
3. Add scanned data asset from Lakehouse to the data products of the governance domain. On a data product's page in Unified Catalog, locate the data assets and select **Add data assets**. You can only perform data profiling and data quality scanning for the data assets associated with the data products under the governance domain.

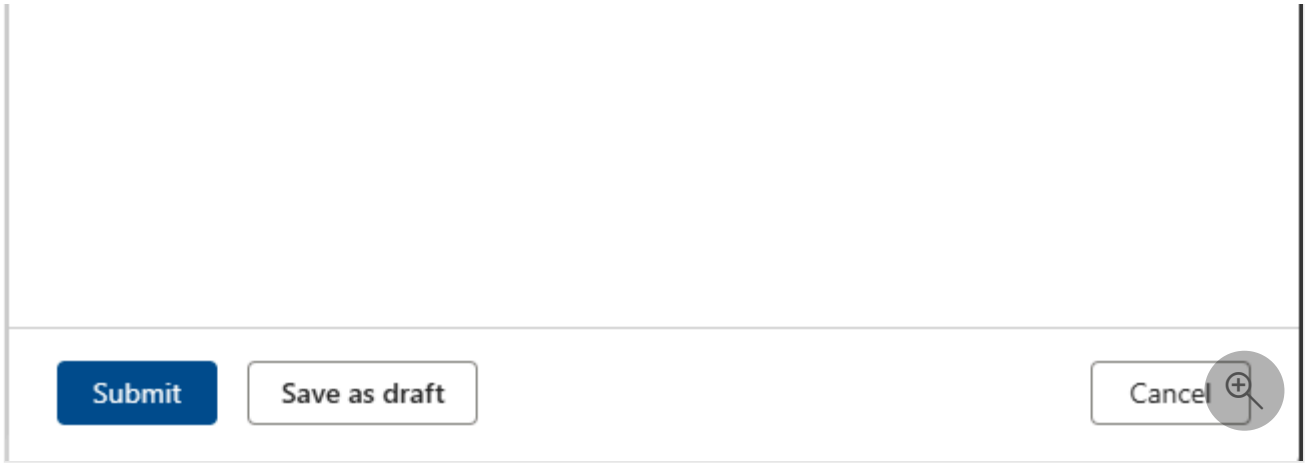
For data profiling and data quality scanning, you need to create a separate data source connection. Data Map uses a service principal scan connection, while Unified Catalog data quality uses a Fabric connection with Microsoft Purview MSI. To set up the data quality connection:

1. In Unified Catalog, select **Health management**, then select **Data quality**.
2. Select a governance domain, and from the **Manage** dropdown list, select **Connections**.
3. Select **New** to open connection configuration page.
4. Add connection display name and a description.
5. Add Source type **Fabric**.
6. Add **Tenant ID**.
7. Add **Workspace ID**.
8. Add **Lakehouse ID**.
9. Add **Credential** - **Microsoft Purview MSI**.



10. Test connection to make sure that the configured connection is successful.


Copilot
SM
Create connection
Overview
Display name *
Fabric Lakehouse
Show ID
Description
Fabric Lakehouse connection
Source type *
Fabric
Tenant ID *
Workspace ID (for testing connection only) *
Lakehouse ID (for testing connection only) *
Credential *
Microsoft Purview MSI (system)
Enable managed V-Net
Test connection
Runs on Apache Spark 3.4 and Delta Lake 2.4
Connection successful.



If your Fabric tenant is configured to run on Virtual Network or in Private Link, check the Virtual Network checkbox in the connection screen. Then select the Azure region where your Fabric and Microsoft Purview are running; for example, UAE North, West Europe, or Canada Central.

In addition to Azure region, you need to add the **Private Link Resource ID**. The following example shows the format:

```
/subscriptions/07d669d6-83f2-4f15-8b5b-4a4b31a7050e/resourceGroups/pdgbugbashfabricvnet/providers/Microsoft.PowerBI/privateLinkServicesForPowerBI/fabricvnetpl.
```

To set up compute for virtual networks, refer to Set up data quality for managed virtual networks. And get details on how to set up and use private link for your Fabric tenant. You need to set up the private link for the Fabric tenant before configuring data quality virtual network connection and compute allocation.

> **Important**
>
> - For a data quality scan, Microsoft Purview MSI must have contributor access to Fabric workspace to connect Fabric workspace. To grant contributor access, open your Fabric workspace, select three dots (...), select **Workspace access**, then **Add people or group**, then add Purview MSI as **Contributor**.
> - Fabric tables must be in Delta format or Iceberg format.

Profiling and data quality scanning for data in Fabric Lakehouse

After completed connection setup successfully, you can profile, create and apply rules, and run a data quality (DQ) scan of your data in Fabric Lakehouse. Follow these steps:

1. Associate a Lakehouse table to a data product for curation, discovery, and subscription. Learn how to create and manage data products.

2. Profile Fabric Lakehouse table. Learn how to configure and run data profiling for a data asset.

3. Configure and run data quality scan to measure data quality of a Fabric Lakehouse table. Learn how to configure and run a data quality scan.

> **Important**
>
> - Make sure that your data are in Delta format or Iceberg format.
> - Make sure the Data Map scan ran successfully. If it didn't, then rerun the scan.

Limitations

Data Quality for Parquet file is designed to support:

- A directory with Parquet Part File. For example: *./Sales/{Parquet Part Files}*. The Fully Qualified Name must follow `https://(storage account).dfs.core.windows.net/(container)/path/path2/{SparkPartitions}`. Make sure there are no {n} patterns in directory/sub-directory structure; it must rather be a direct FQN leading to {SparkPartitions}.
- A directory with Partitioned Parquet Files, partitioned by columns within the dataset like sales data partitioned by year and month. For example: *./Sales/{Year=2018}/{Month=Dec}/{Parquet Part Files}*.

Both supported Parquet directory patterns present a consistent Parquet dataset schema. Limitation: It isn't designed to or won't support N arbitrary Hierarchies of Directories with Parquet Files. We advise the customer to present data in (1) or (2) constructed structure. So recommending customer to follow the supported parquet standard or migrate their data to **ACID** compliant delta format.

Tips for initial setup

Use the following tips when setting up data quality scanning for Fabric Lakehouse data for the first time.

For Data Map:

- Ensure the service principal (SPN) has workspace permissions.
- Ensure scan connection uses the service principal.
- Running a full scan is suggested if you're setting up a Lakehouse scan for the first time.
- Check that the ingested assets have been updated or refreshed.

For Unified Catalog:

- The data quality connection needs to use MSI credentials.
- Ideally, create a new data product for first-time testing of Lakehouse data quality scans.
- Add the ingested data assets, and check that the data assets are updated.
- Try running a profile first. If successful, try running a data quality rule. If not successful, try refreshing the asset schema (**Schema** > **Schema management** > **Import schema**).
- Some users also had to create a new Lakehouse and sample data to verify everything works from scratch. In some cases, working with assets that were previously ingested in the Data Map, the experience isn't consistent.

Resources

Use the following resources for related tasks and concepts.

Data quality tasks:

- Configure and run data profiling of your data
- Configure and run data quality scan
- Set up data quality connection
- FAQ and self-help

Fabric data ingestion:

- Learn about Mirroring in Fabric
- OneLake shortcuts

Related data quality scenarios:

- Data Quality for Fabric mirrored data sources
- Data Quality for Fabric shortcut databases

Last updated on 05/27/2026

Set up data quality for Fabric mirrored databases

This article explains how to configure data quality scans in Microsoft Purview for data assets that are mirrored from external sources into Microsoft Fabric.

Mirroring in Fabric is a low-cost, low-latency solution that replicates data from sources like Azure SQL Database, Azure Cosmos DB, and Snowflake into Fabric's OneLake. For more information, see the Fabric mirroring documentation.

Configure data quality for a Fabric mirrored database

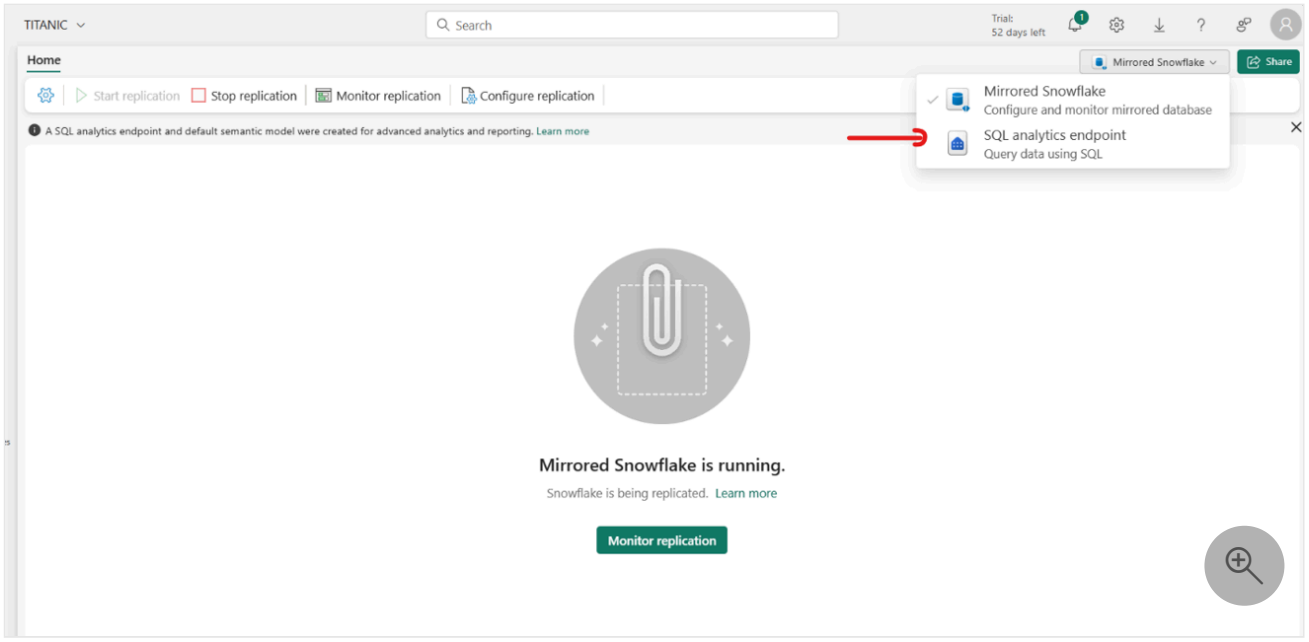
To set up data quality for a Fabric mirrored database, complete the following steps:

> **Note**
>
> Before you begin, make sure you have:
>
> - A Microsoft Fabric workspace with mirroring enabled.
> - A Lakehouse in your Fabric workspace (or permissions to create one).
> - Access to Microsoft Purview Data Map with service principal authentication configured.

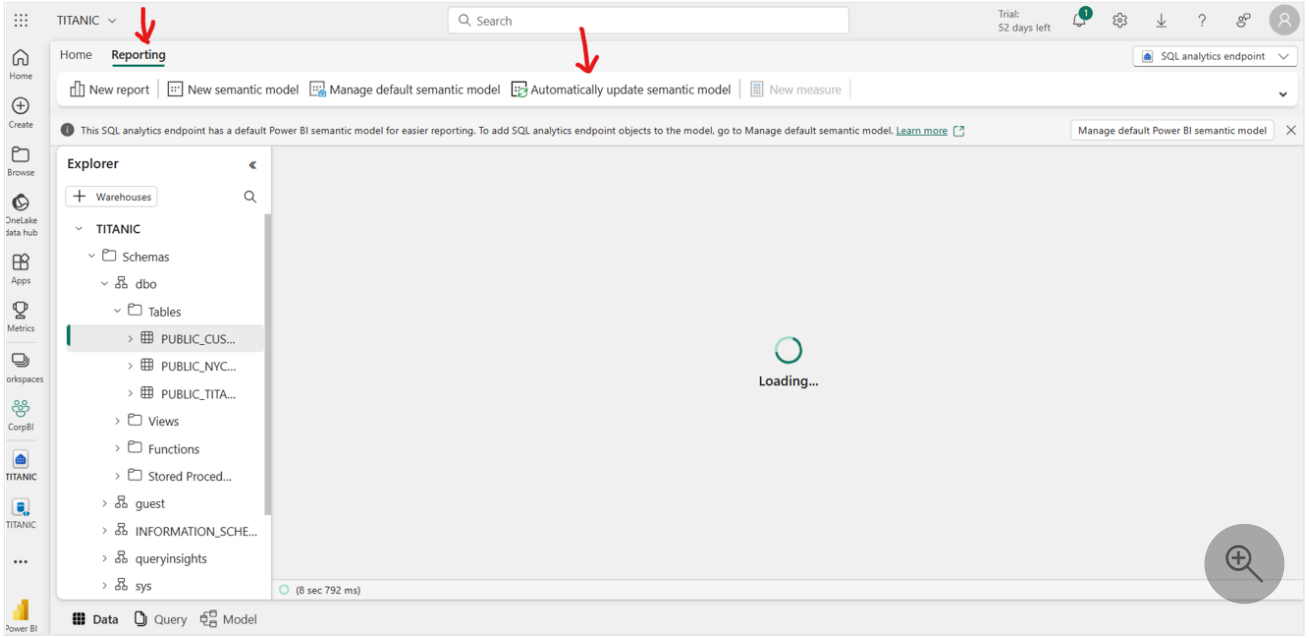
1. Enable mirroring in your Fabric tenant. A Power BI administrator can enable mirroring for the entire organization or for specific security groups in the Power BI admin portal. You can replicate an entire database or individual tables.

2. After enabling mirroring and initiating replication, confirm that replication completes successfully.

3. Open the SQL analytics endpoint.



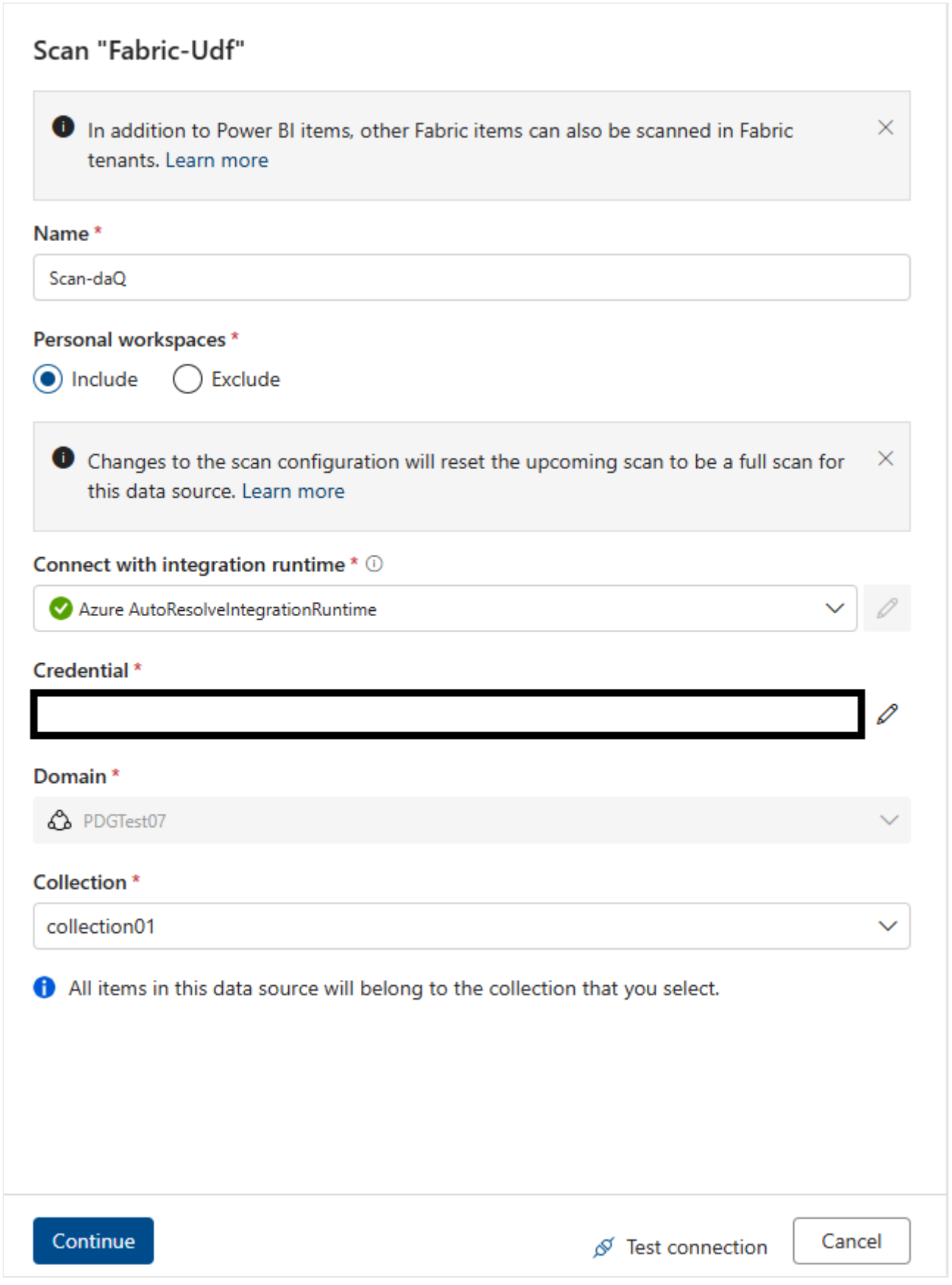
4. On the **Reporting** tab, select **Automatically update semantic model**.



5. Create a Lakehouse in your Fabric workspace if you don't have one created.

6. Create a Fabric shortcut from your mirrored database to the lakehouse.

7. Go to Microsoft Purview Data Map and run a Data Map scan on the lakehouse you created in the previous step. Use service principal authentication.



8. When the scan is completed, associate the discovered Lakehouse table assets with a data product. Make sure to select the **Lakehouse tables** to associate to your data product.

9. After associating the Lakehouse tables with your target data product in Unified Catalog, you can profile and measure the data quality of all mirrored tables as Lakehouse tables in Microsoft Purview.

10. In the **Data quality** area of **Health management** in Unified Catalog, run a data quality scan or profile your data as usual.

> **Important**
>
> - Use service principal for Data Map scans, and use a managed identity for data quality scans.
> - Select the mirrored database instead of individual tables.
> - Update the semantic model every time when you add new table to the mirrored database.
> - If your mirrored database tables aren't available in the Fabric Lakehouse, contact Fabric support.
> - Data quality scanning is supported only for Lakehouse Delta, Iceberg, and Parquet file formats.

Related content

- How to configure and run data profiling of your data
- How to configure and run data quality scan

Last updated on 06/01/2026

Set up data quality for Fabric shortcut databases

Shortcuts are objects in Microsoft OneLake that point to other storage locations. The location can be internal or external to OneLake. The shortcut target path is the location the shortcut points to. The shortcut path is where the shortcut appears. Shortcuts appear as folders in OneLake, and any workload or service with access to OneLake can use them.

This article explains how to create Fabric shortcuts for external data sources and configure data quality scans for the shortcut-based assets in Microsoft Purview Unified Catalog.

With OneLake shortcuts, you can unify your data across domains, clouds, and accounts in a single virtual data lake. All Microsoft Fabric analytical engines connect directly to data sources like Azure, Amazon Web Services (AWS), and OneLake through a unified namespace. OneLake manages permissions and credentials, so you don't need to configure each Fabric workload separately.

For more information about Fabric shortcuts, see the OneLake shortcuts documentation.

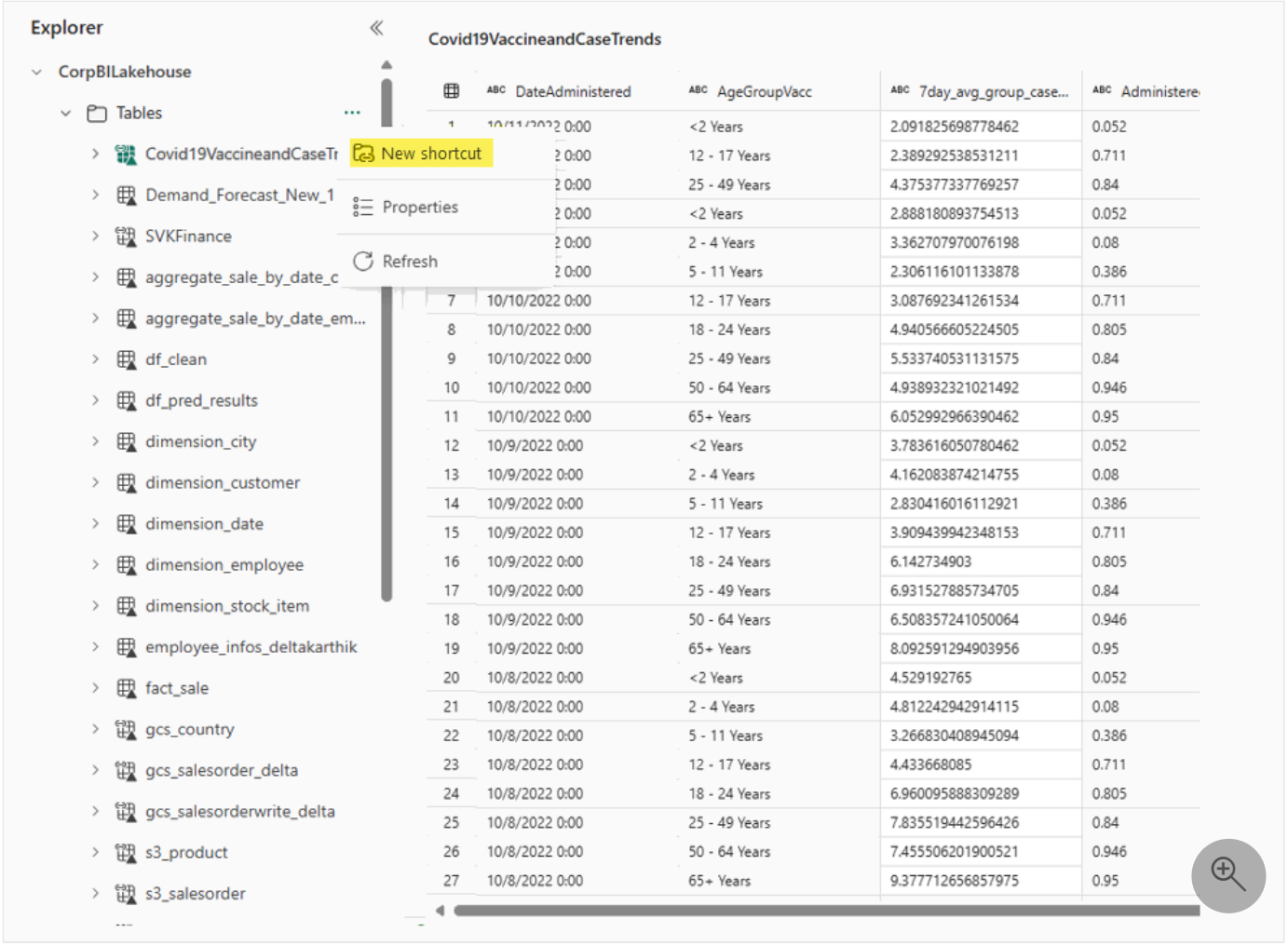
Prerequisites

Before you can run data quality scans on shortcut databases, create a shortcut for your external data source. After you create the shortcut, run a Data Map scan to register the asset in Unified Catalog, then run a data quality scan.

> **ⓘ Important**
>
> - Use a service principal or managed identity for Data Map scans, and managed identity for data quality scans.
> - Any data sourced through a shortcut is processed in the same region as the Fabric workspace.
> - The Fabric team needs to differentiate shortcut items from native items in the Microsoft OneLake SDK for Lakehouse subartifacts. For now, all shortcut items (tables and files) are considered as native items during Data Map scanning.

Sign in to your Microsoft Fabric workspace. Select the ellipsis button under **Tables**, and select **New Shortcut**. On the **New shortcut** page, you can create shortcuts for several external sources. This article covers the following:

- Azure Data Lake Gen2 shortcuts
- Amazon S3 shortcuts
- Google Cloud Storage shortcuts

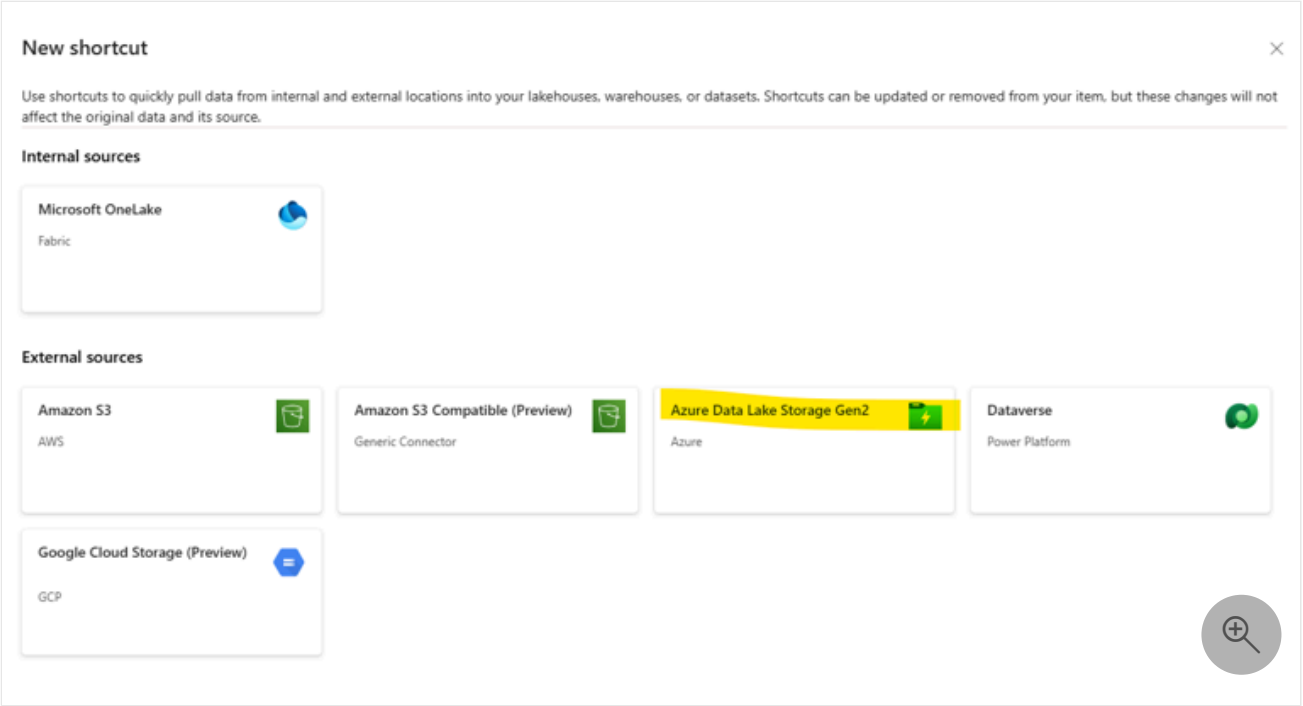


	DateAdministered	AgeGroupVacc	7day_avg_group_case...	Administere
1	10/11/2022 0:00	<2 Years	2.091825698778462	0.052
2	10/10/2022 0:00	12 - 17 Years	2.389292538531211	0.711
3	10/10/2022 0:00	25 - 49 Years	4.375377337769257	0.84
4	10/10/2022 0:00	<2 Years	2.888180893754513	0.052
5	10/10/2022 0:00	2 - 4 Years	3.362707970076198	0.08
6	10/10/2022 0:00	5 - 11 Years	2.306116101133878	0.386
7	10/10/2022 0:00	12 - 17 Years	3.087692341261534	0.711
8	10/10/2022 0:00	18 - 24 Years	4.940566605224505	0.805
9	10/10/2022 0:00	25 - 49 Years	5.533740531131575	0.84
10	10/10/2022 0:00	50 - 64 Years	4.938932321021492	0.946
11	10/10/2022 0:00	65+ Years	6.052992966390462	0.95
12	10/9/2022 0:00	<2 Years	3.783616050780462	0.052
13	10/9/2022 0:00	2 - 4 Years	4.162083874214755	0.08
14	10/9/2022 0:00	5 - 11 Years	2.830416016112921	0.386
15	10/9/2022 0:00	12 - 17 Years	3.909439942348153	0.711
16	10/9/2022 0:00	18 - 24 Years	6.142734903	0.805
17	10/9/2022 0:00	25 - 49 Years	6.931527885734705	0.84
18	10/9/2022 0:00	50 - 64 Years	6.508357241050064	0.946
19	10/9/2022 0:00	65+ Years	8.092591294903956	0.95
20	10/8/2022 0:00	<2 Years	4.529192765	0.052
21	10/8/2022 0:00	2 - 4 Years	4.812242942914115	0.08
22	10/8/2022 0:00	5 - 11 Years	3.266830408945094	0.386
23	10/8/2022 0:00	12 - 17 Years	4.433668085	0.711
24	10/8/2022 0:00	18 - 24 Years	6.960095888309289	0.805
25	10/8/2022 0:00	25 - 49 Years	7.835519442596426	0.84
26	10/8/2022 0:00	50 - 64 Years	7.455506201900521	0.946
27	10/8/2022 0:00	65+ Years	9.377712656857975	0.95

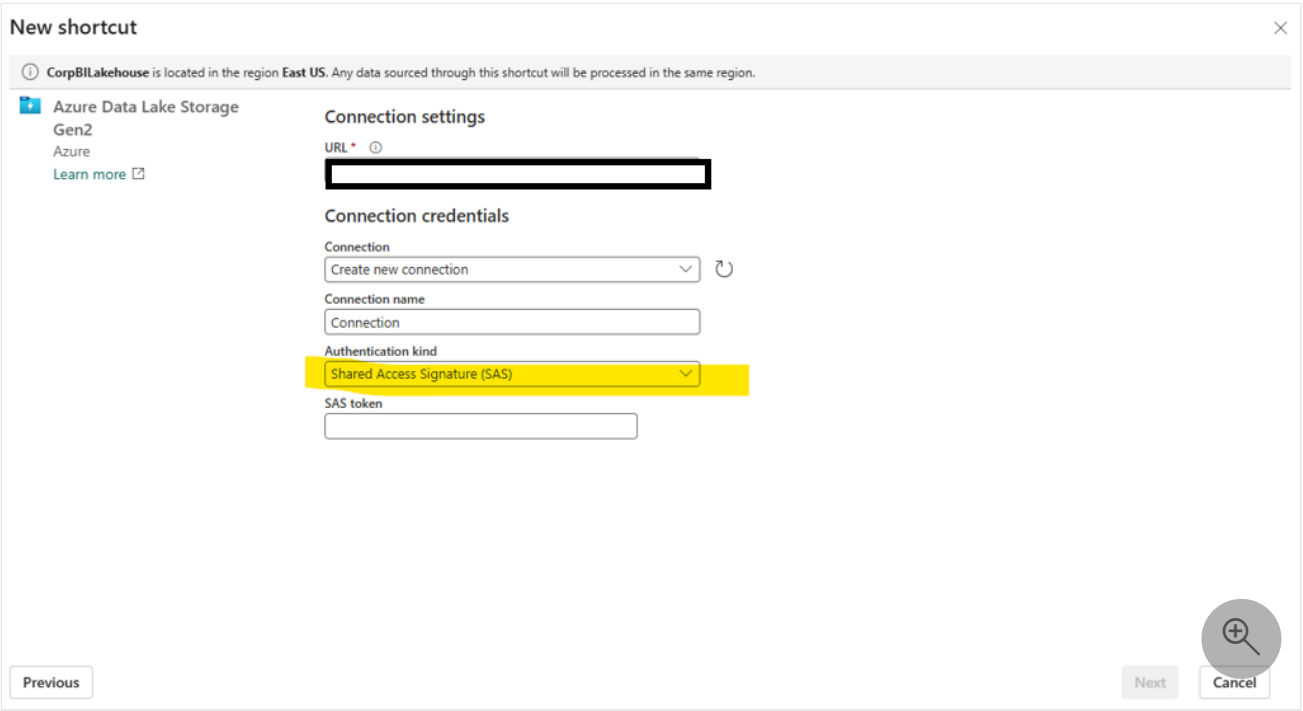
Set up an Azure Data Lake Gen2 shortcut

To create an Azure Data Lake Gen2 shortcut and scan it for data quality, complete the following steps:

1. Select **Azure Data Lake Storage Gen2** on the Fabric workspace **New shortcut** page.

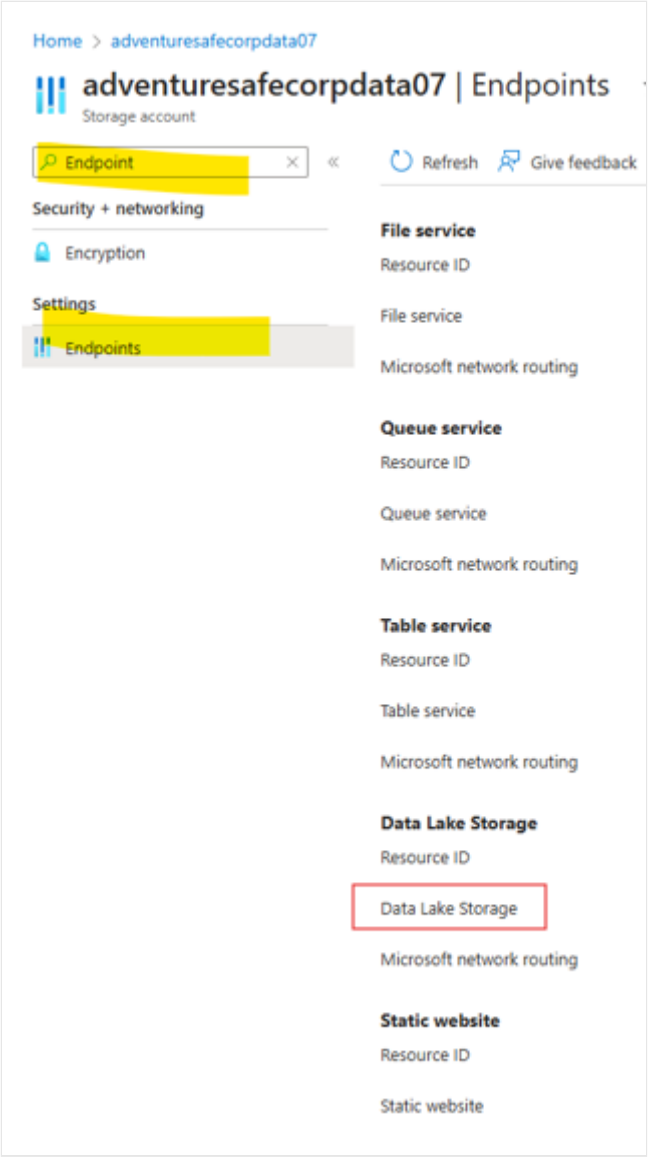


2. Select ADLS Gen2 SAS authentication.

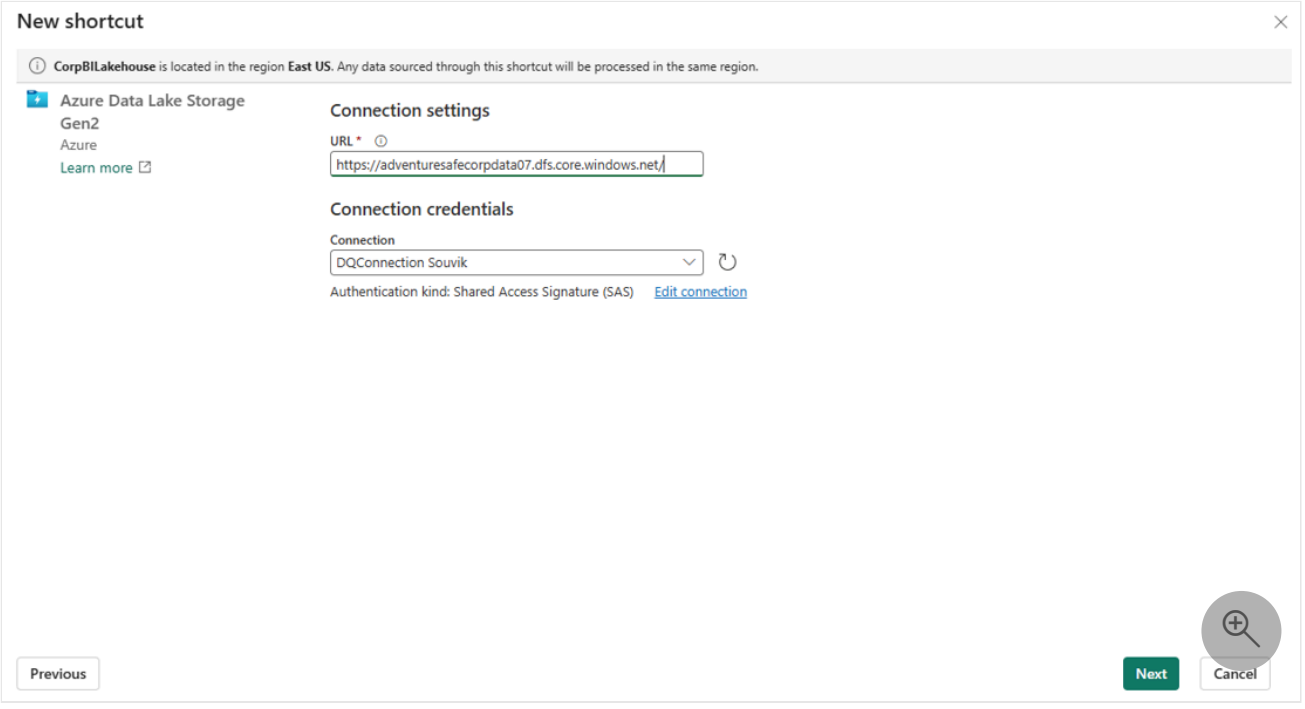


3. In the Azure portal, generate a shared access signature (SAS) and connection string for your ADLS Gen2 resource.

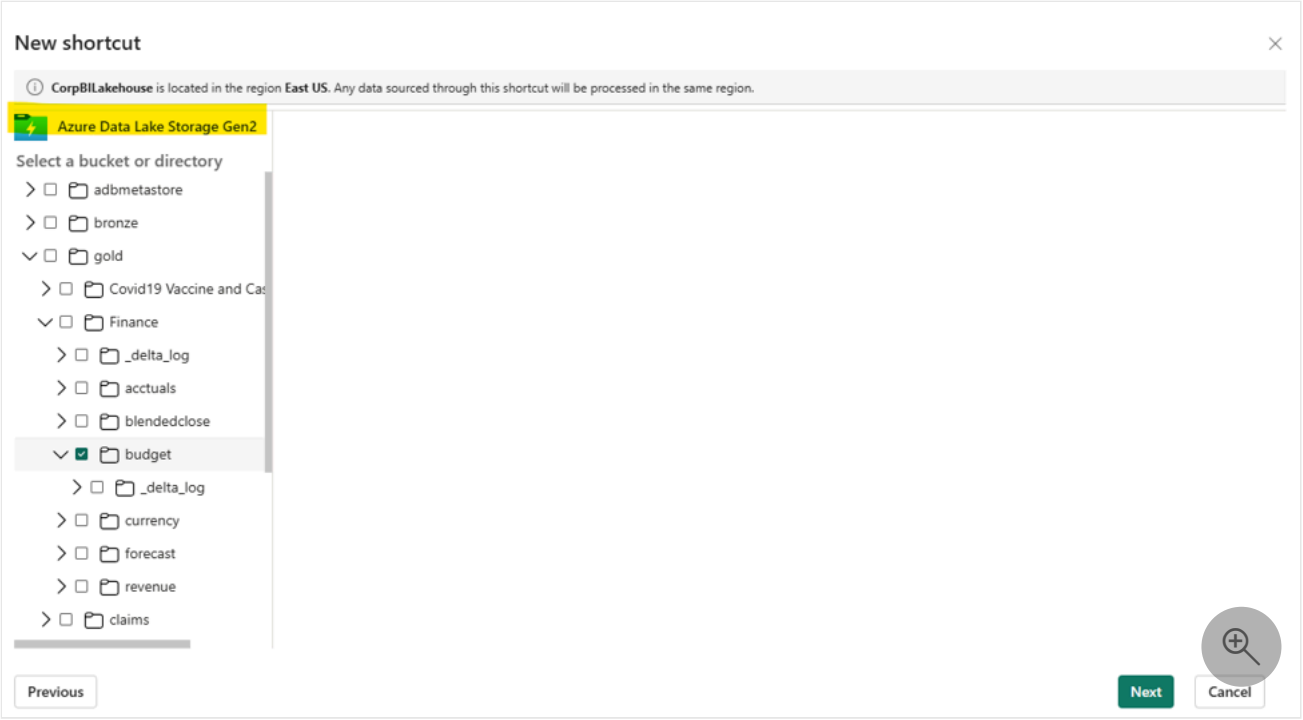
4. **Copy** the endpoint of the data lake.



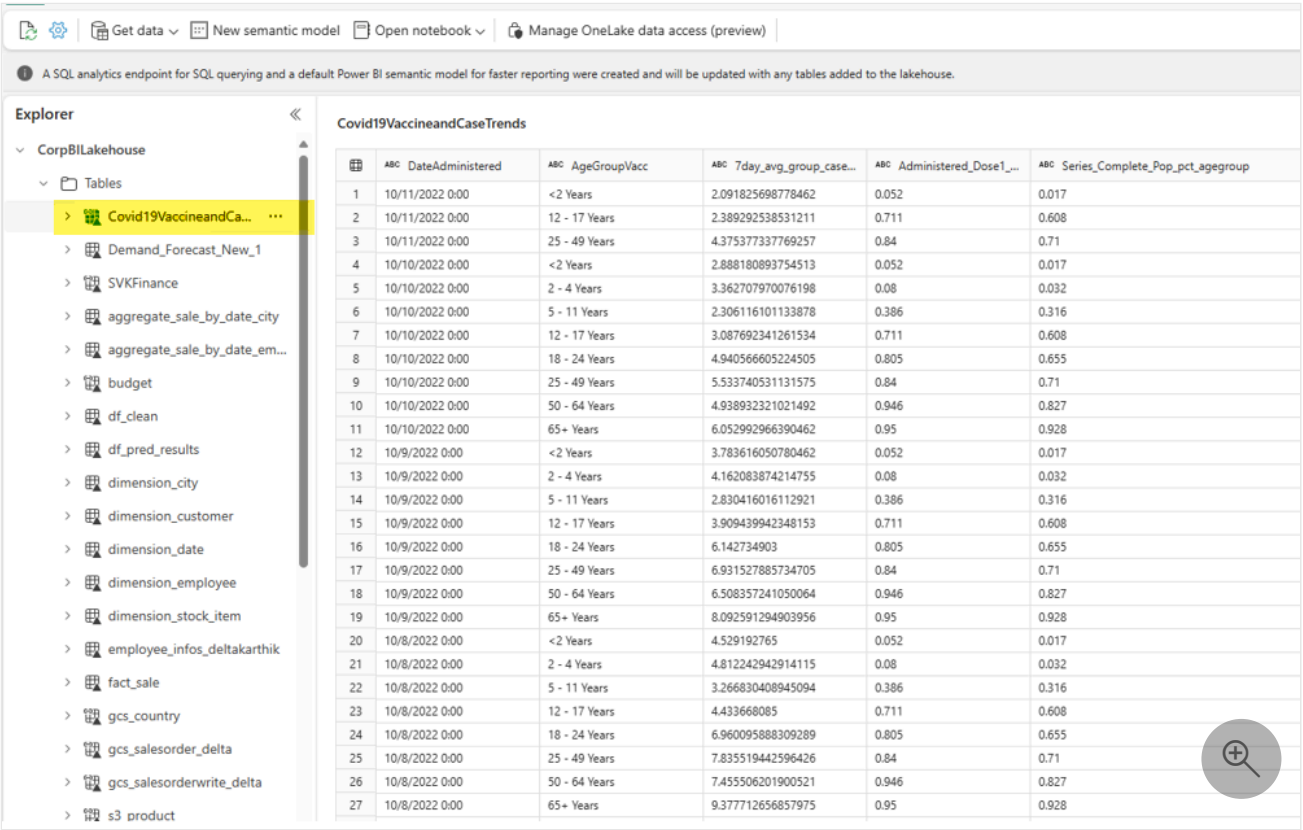
5. Add storage details for the shortcut storage.



6. Navigate to and choose the correct delta folder.

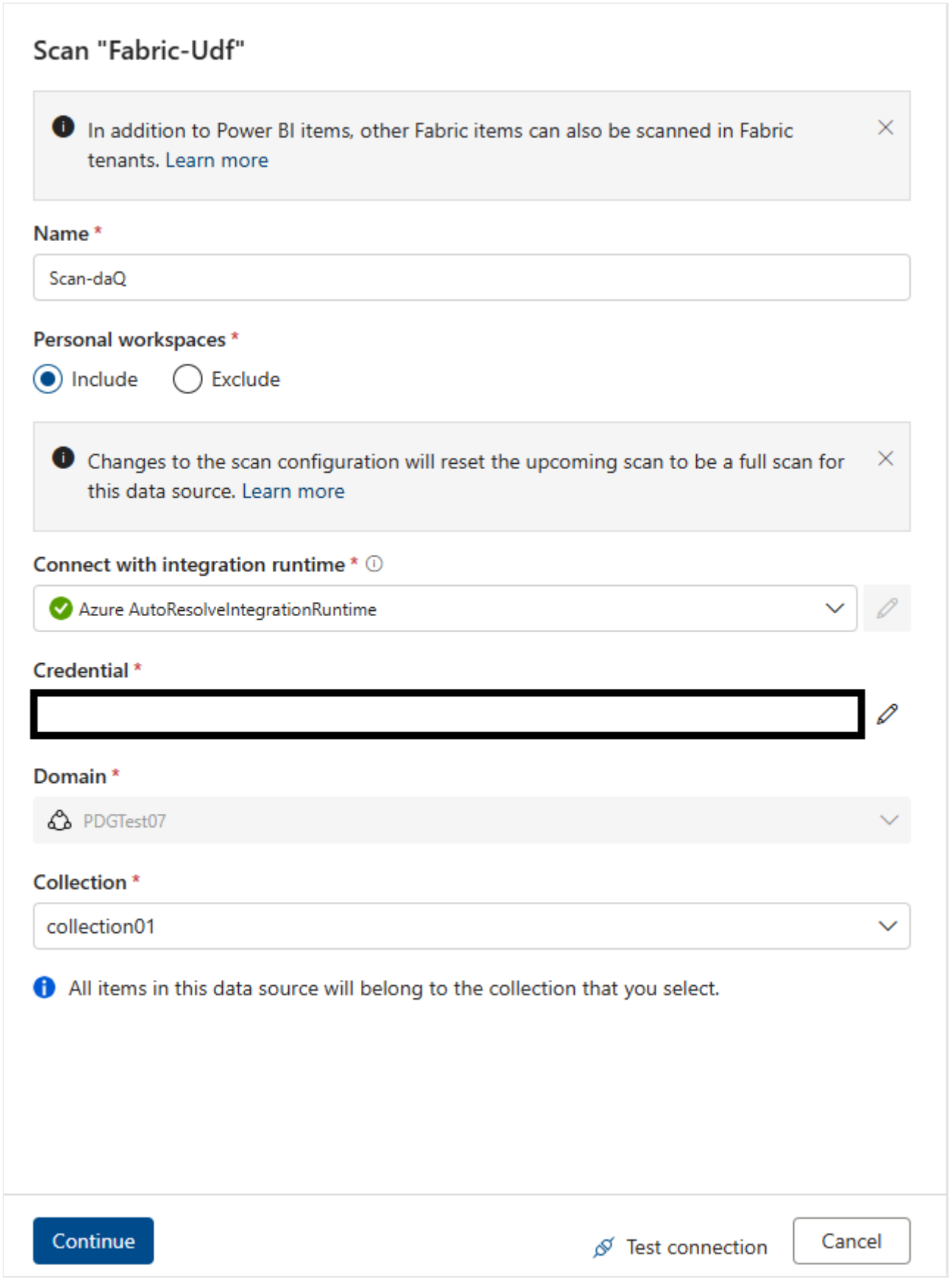


7. Preview the shortcut delta table in your Fabric workspace.



	DateAdministered	AgeGroupVacc	7day_avg_group_case...	Administered_Dose1_...	Series_Complete_Pop_pct_agegroup
1	10/11/2022 0:00	<2 Years	2.091825698778462	0.052	0.017
2	10/11/2022 0:00	12 - 17 Years	2.389292538531211	0.711	0.608
3	10/11/2022 0:00	25 - 49 Years	4.375377337769257	0.84	0.71
4	10/10/2022 0:00	<2 Years	2.888180893754513	0.052	0.017
5	10/10/2022 0:00	2 - 4 Years	3.362707970076198	0.08	0.032
6	10/10/2022 0:00	5 - 11 Years	2.306116101133878	0.386	0.316
7	10/10/2022 0:00	12 - 17 Years	3.087692341261534	0.711	0.608
8	10/10/2022 0:00	18 - 24 Years	4.940566605224505	0.805	0.655
9	10/10/2022 0:00	25 - 49 Years	5.5337405311315 75	0.84	0.71
10	10/10/2022 0:00	50 - 64 Years	4.938932321021492	0.946	0.827
11	10/10/2022 0:00	65+ Years	6.052992966390462	0.95	0.928
12	10/9/2022 0:00	<2 Years	3.783616050780462	0.052	0.017
13	10/9/2022 0:00	2 - 4 Years	4.162083874214755	0.08	0.032
14	10/9/2022 0:00	5 - 11 Years	2.830416016112921	0.386	0.316
15	10/9/2022 0:00	12 - 17 Years	3.909439942348153	0.711	0.608
16	10/9/2022 0:00	18 - 24 Years	6.142734903	0.805	0.655
17	10/9/2022 0:00	25 - 49 Years	6.931527885734705	0.84	0.71
18	10/9/2022 0:00	50 - 64 Years	6.508357241050064	0.946	0.827
19	10/9/2022 0:00	65+ Years	8.092591294903956	0.95	0.928
20	10/8/2022 0:00	<2 Years	4.529192765	0.052	0.017
21	10/8/2022 0:00	2 - 4 Years	4.812242942914115	0.08	0.032
22	10/8/2022 0:00	5 - 11 Years	3.266830408945094	0.386	0.316
23	10/8/2022 0:00	12 - 17 Years	4.433668085	0.711	0.608
24	10/8/2022 0:00	18 - 24 Years	6.960095888309289	0.805	0.655
25	10/8/2022 0:00	25 - 49 Years	7.835519442596426	0.84	0.71
26	10/8/2022 0:00	50 - 64 Years	7.455506201900521	0.946	0.827
27	10/8/2022 0:00	65+ Years	9.377712656857975	0.95	0.928

8. Start a Data Map scan of your Azure Data Lake Gen2 resource. Use service principal or managed identity authentication.



9. When the Data Map scan finishes, your asset appears in Unified Catalog as a Fabric Lakehouse table.

10. Associate the asset with a data product for curation and data quality assessment.

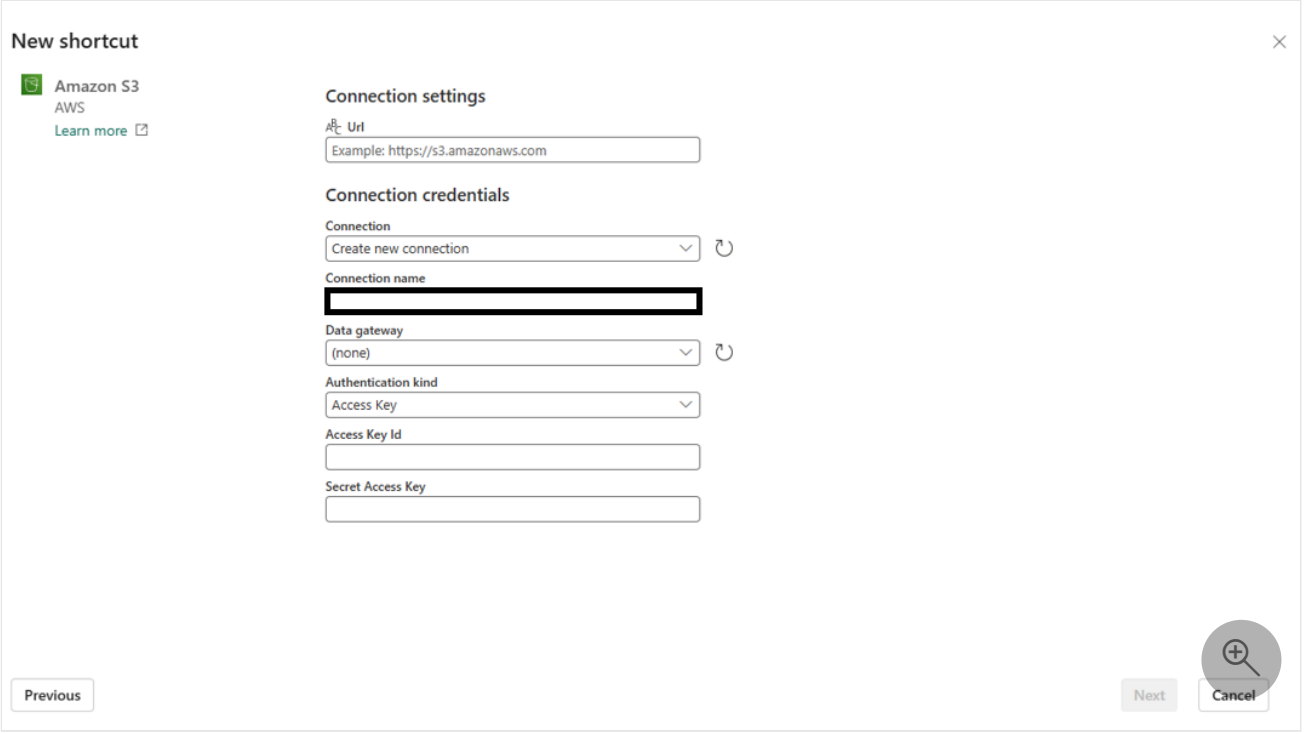
11. In Unified Catalog, configure and run a data quality scan or run data profiling as usual.

Set up an Amazon S3 shortcut

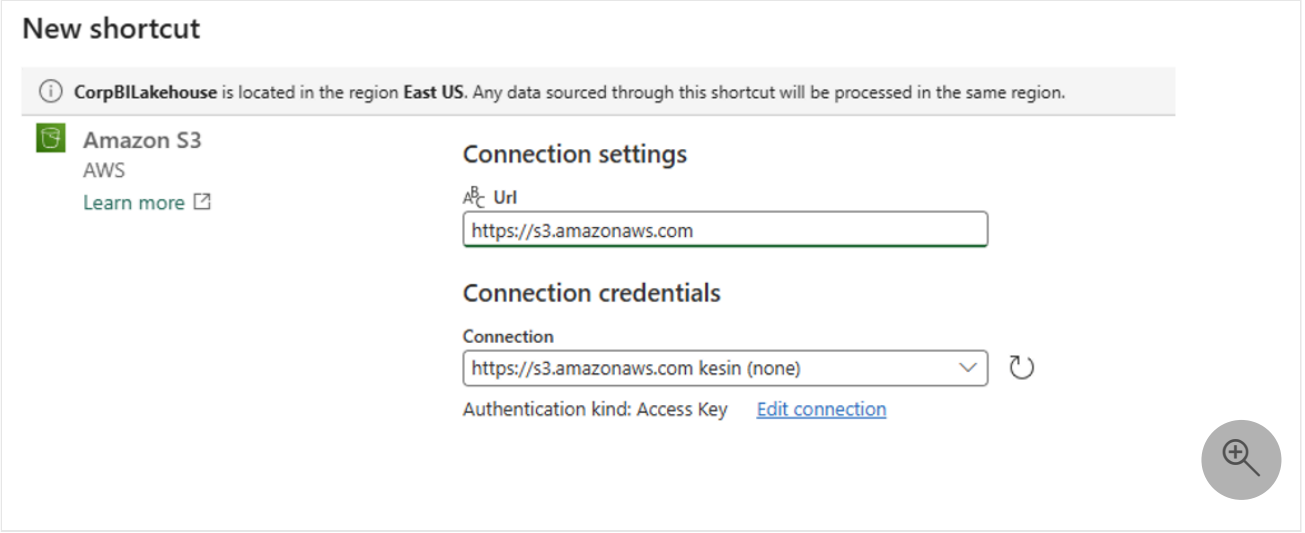
To create an Amazon S3 shortcut and scan it for data quality, you need the S3 bucket URL, access key ID, and secret access key. Complete the following steps:

1. Select **New shortcut** in the Microsoft Fabric workspace.

2. Select **Amazon S3** and add the URL, access key ID, and secret access key.



3. Add the connection URL and storage details.



4. Preview the shortcut in your Fabric workspace.

5. Start a Data Map scan of your Amazon S3 resource. Use service principal or managed identity authentication.

6. When the Data Map scan finishes, your data asset appears in Unified Catalog.

7. Associate the asset with a data product for curation and data quality assessment.

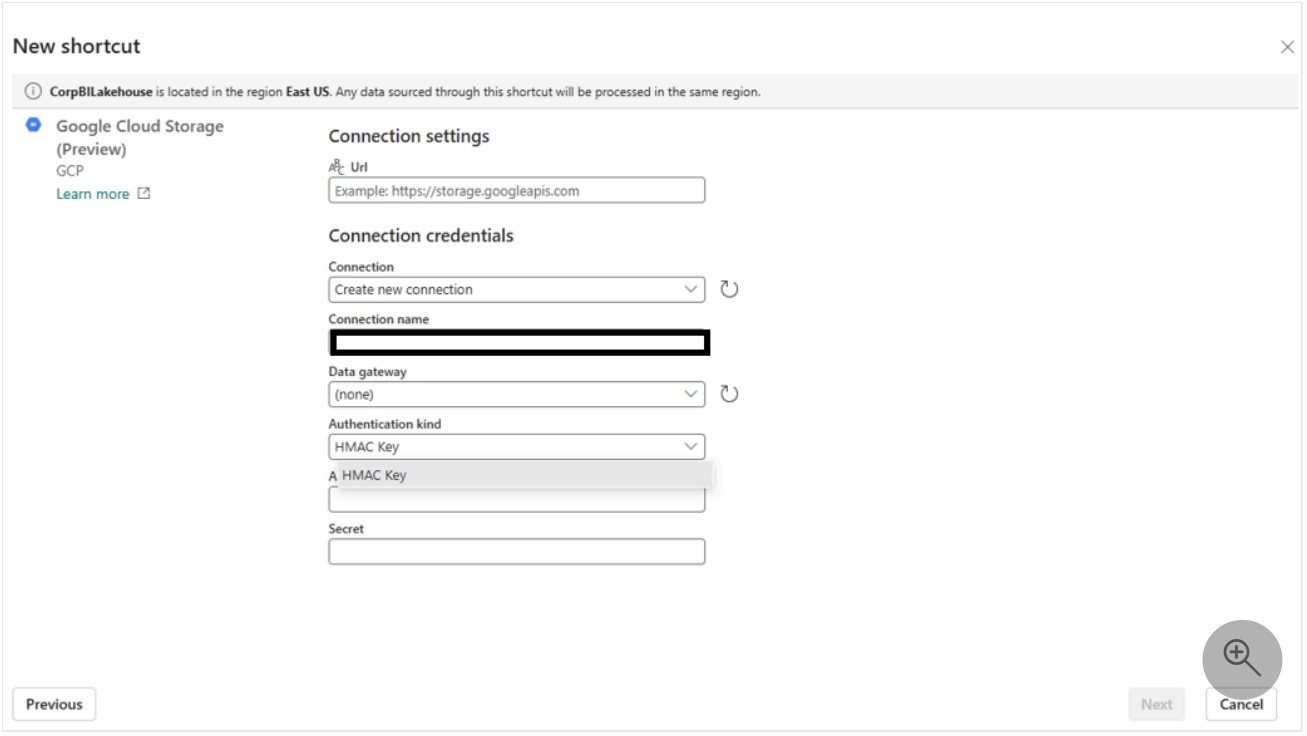
8. In Unified Catalog, configure and run a data quality scan or run data profiling as usual.

Set up a Google Cloud Storage (GCS) shortcut

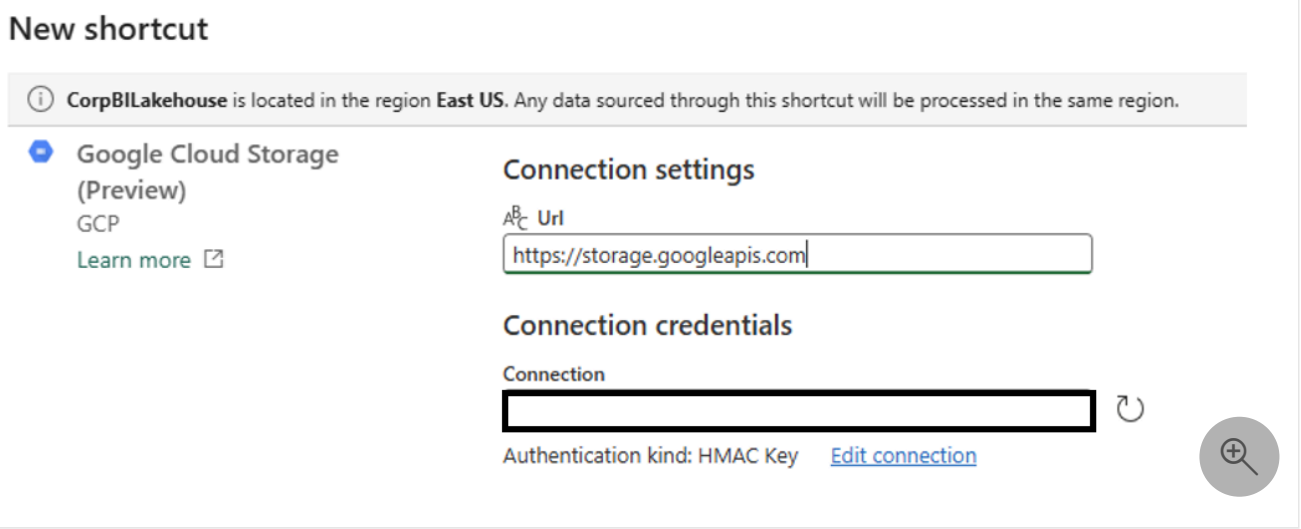
To create a Google Cloud Storage shortcut and scan it for data quality, you need the storage URL and HMAC access key credentials. Complete the following steps:

1. Select **New shortcut** in the Fabric workspace.

2. Select **Google Cloud Storage** and add the URL, access key ID, and access key shortcut.



3. Add the connection URL and storage details.



4. Preview the shortcut in your Fabric workspace.

5. Start a Data Map scan of your Google Cloud Storage resource. Use service principal or managed identity authentication.

6. When the Data Map scan finishes, your data asset appears in Unified Catalog.

7. Associate the asset with a data product for curation and data quality assessment.

8. In Unified Catalog, configure and run a data quality scan or run data profiling as usual.

Related content

- Configure and run data profiling in Unified Catalog
- Configure and run a data quality scan in Unified Catalog

Last updated on 06/02/2026

Set up data quality for Google BigQuery (preview)

This article provides step-by-step guidance to register, scan, and configure data quality for Google BigQuery sources in Microsoft Purview Unified Catalog.

Supported capabilities

When scanning a Google BigQuery source, Microsoft Purview supports:

- Extracting technical metadata including:
  - Projects and datasets.
  - Tables including the columns.
  - Views including the columns.
- Fetching static lineage on assets relationships among tables and views.

When setting up a scan, you can choose to scan an entire Google BigQuery project. You can also scope the scan to a subset of datasets matching the given names or name patterns.

Known limitations

Be aware of the following limitations when scanning Google BigQuery sources:

- Microsoft Purview only supports scanning Google BigQuery datasets in a US multiregional location. If a dataset is in another location (for example, "us-east1" or "EU"), the scan completes but no assets appear in Microsoft Purview.

- When you delete an object from the data source, the subsequent scan doesn't automatically remove the corresponding asset in Microsoft Purview.

Configure Data Map scan to catalog Google BigQuery data in Microsoft Purview

Register a Google BigQuery project

To register a Google BigQuery project in Microsoft Purview, complete the following steps:

1. In the Microsoft Purview portal, select **Data Map**, then select **Register**.

2. At **Register sources**, select **Google BigQuery**, then select **Continue**.
3. Enter a **Name** for the data source that appears in the catalog.
4. Enter the **ProjectID**. The ProjectID should be a fully qualified project ID. For example, `mydomain.com: myProject`.
5. Select a collection from the list.
6. Select **Register**.

Set up a Data Map scan for Google BigQuery project

To configure a Data Map scan for the registered Google BigQuery project, follow these steps:

1. Make sure a self-hosted integration runtime is set up. If the runtime isn't set up, use the steps in Google BigQuery connection prerequisites.

2. Navigate to **Sources**.

3. Select the registered BigQuery project.

4. Select **New scan**.

5. Enter these details:

    - **Name**: The name of the scan.
    - **Connect via integration runtime**: Select the configured self-hosted integration runtime.
    - **Credential**: While configuring BigQuery credential, make sure to:
        - Select **Basic Authentication** as the authentication method.
        - Provide the email ID of the service account in the **User name** field. For example, `xyz\@developer.gserviceaccount.com`.
        - Store the service account private key as a Key Vault secret (see the following steps).
    - **JDBC driver location**: Specify the path to the JDBC (Java Database Connectivity) driver on the machine where the self-hosted integration runtime is running. For example: `D:\Drivers\GoogleBigQuery`.
    - **Datasets to import**: Specify a list of BigQuery datasets to import. For example, `dataset1`; `dataset2`. When the list is empty, all available datasets are imported.
    - **Maximum memory (in GB)**: The required memory value depends on the size of the Google BigQuery project being scanned.

    Create a service account private key:

a. In the Google Cloud console navigation menu, select **IAM & Admin** > **Service Accounts**, then select a project.
b. Select the email address of the service account that you want to create a key for.
c. Select the **Keys** tab.
d. Select **Add key** > **Create new key**.
e. Choose JSON format.
f. Copy the entire JSON key file and store it as the value of a Key Vault secret.

6. Select **Test connection**.

7. Select **Continue**.

8. Choose your scan trigger. You can set up a schedule or run the scan once.

9. Review your scan and select **Save and Run**.

Once scanned, the data assets in Google BigQuery project are available in the Unified Catalog search. For more information, see how to connect and manage Google BigQuery in Microsoft Purview.

> ⓘ **Important**
>
> Deleting your scan doesn't delete catalog assets created from previous scans.

Set up connection to Google BigQuery project for data quality scan

The scanned assets are now ready for cataloging and governance. Associate the scanned assets to the data products in a governance domain to set up a data quality scan.

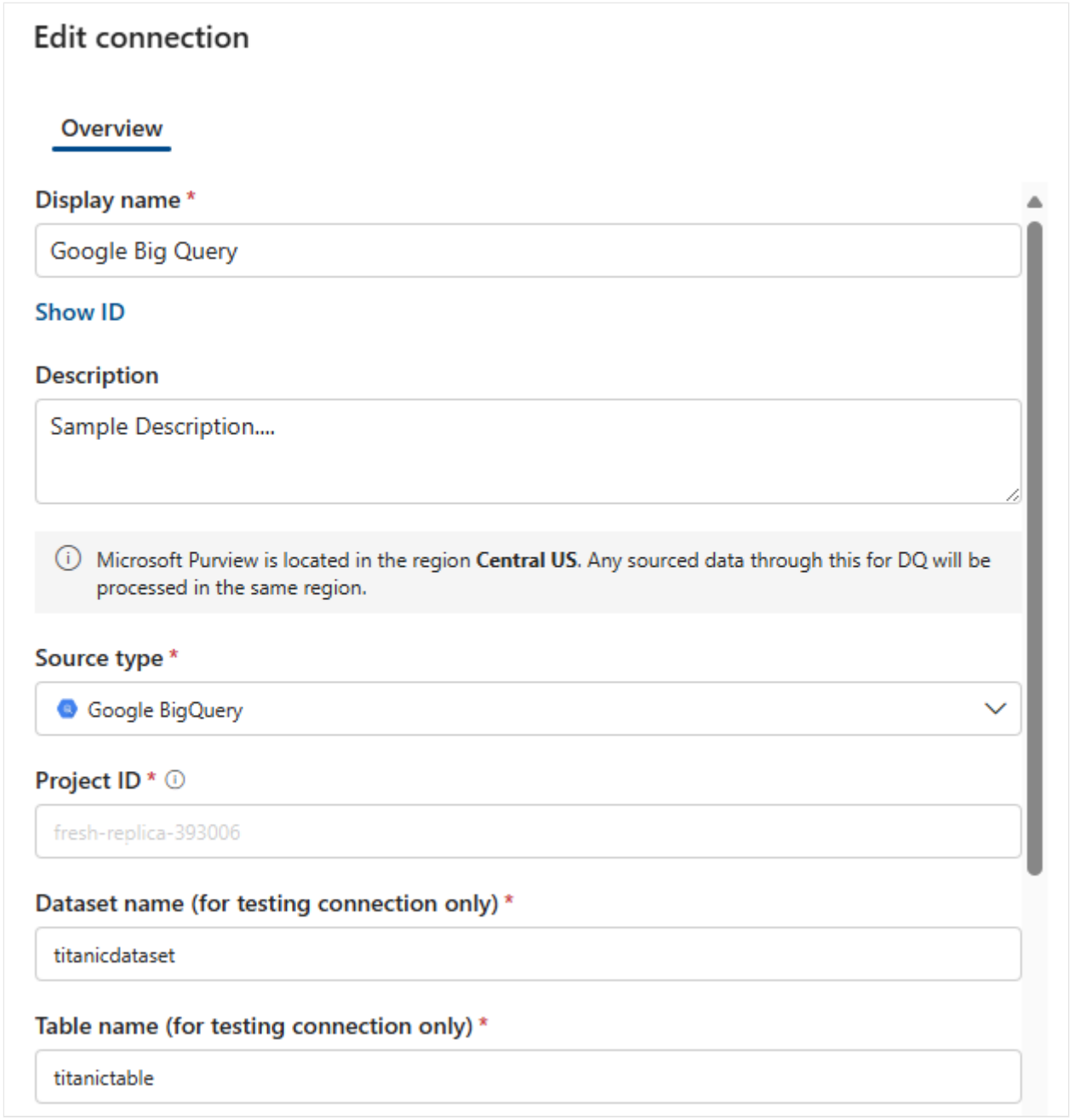
> ⓘ **Important**
>
> Data quality stewards need **read only** access to Google BigQuery to set up a data quality connection. Virtual network and private endpoint aren't supported for Google BigQuery data source yet for data quality scanning service.

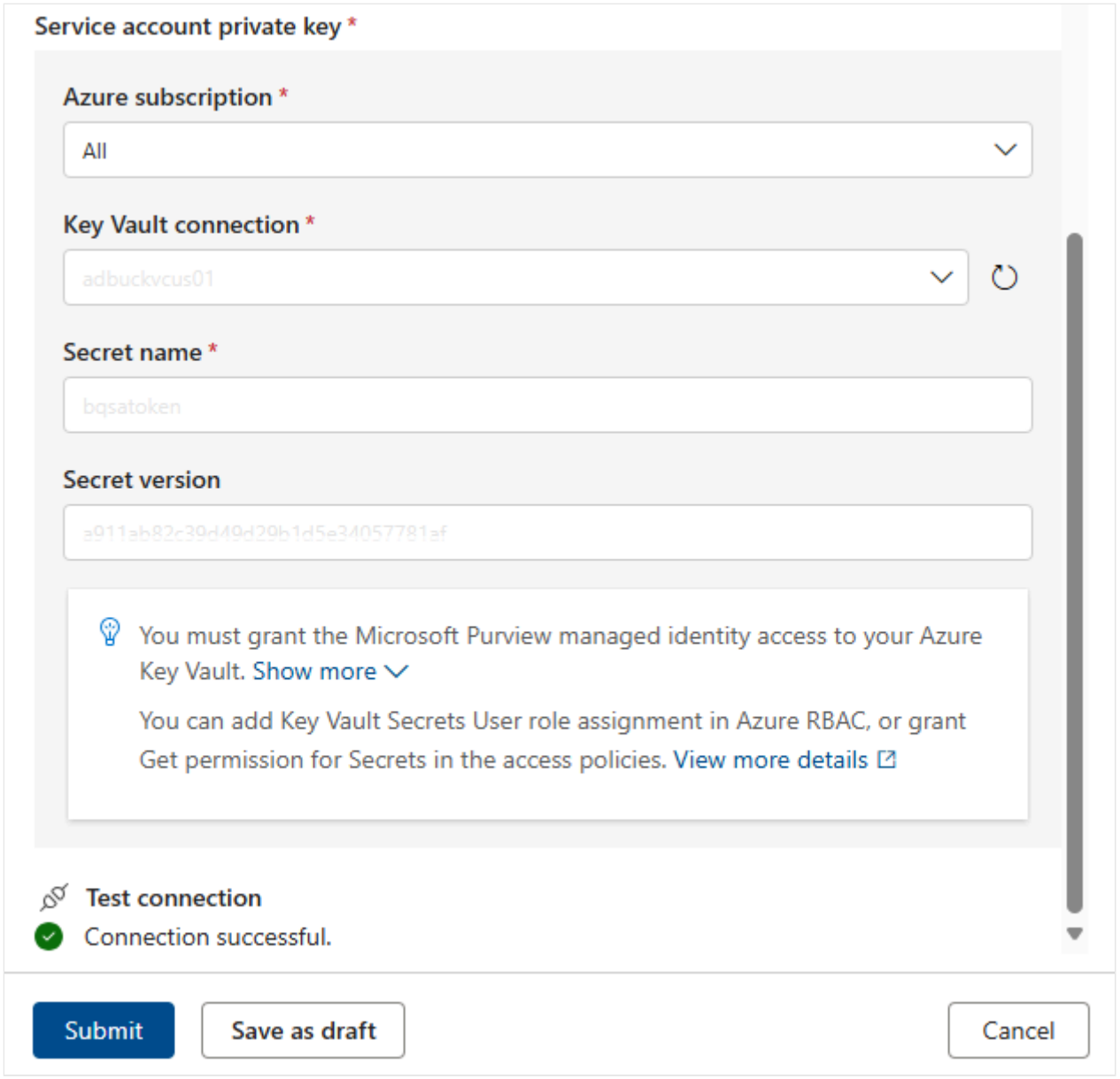
1. In Unified Catalog, go to **Health management** > **Data quality**. Select a governance domain to open its details page, then select **Manage** to create a connection.

2. Set up the connection:

- Add connection name and description.
- Select source type **Google BigQuery**.
- Add **Project ID**, **Dataset name**, and **Table name**.
- Enter details at **Service account private key**:
  - Add the Microsoft Azure subscription.
  - Add the Microsoft Azure Key Vault connection.
  - Enter the **Secret name**.
  - Enter the **Secret version**.

3. Test the connection to make sure the data source connection is successfully configured.





Profiling and data quality scanning for data in Google BigQuery

After you set up the connection, you can profile your data, create and apply rules, and run a data quality scan for your data in Google BigQuery. Follow the step-by-step guidelines in the following articles:

- How to configure and run data profiling of your data
- How to configure and run data quality scan

Related content

- Connect to and manage Google BigQuery projects in Microsoft Purview
- Data quality troubleshooting
- Set up data source connection for data quality in Unified Catalog

Last updated on 06/04/2026

Set up managed virtual networks for data quality scans in virtual network storage

This article outlines the steps to create protected data source connections for data profiling and data quality scans in Microsoft Purview.

Virtual networks and private endpoints enhance the security and isolation of Azure resources. Virtual network protected endpoints let you access Azure services while keeping traffic on the Azure backbone network. Private endpoints go further by assigning a private IP address in your virtual network, so all traffic stays within the network and doesn't use the public internet. Virtual network protected endpoints and private endpoints are essential when security and network isolation are critical, because they minimize exposure to threats from the public internet.

Review user permission requirements

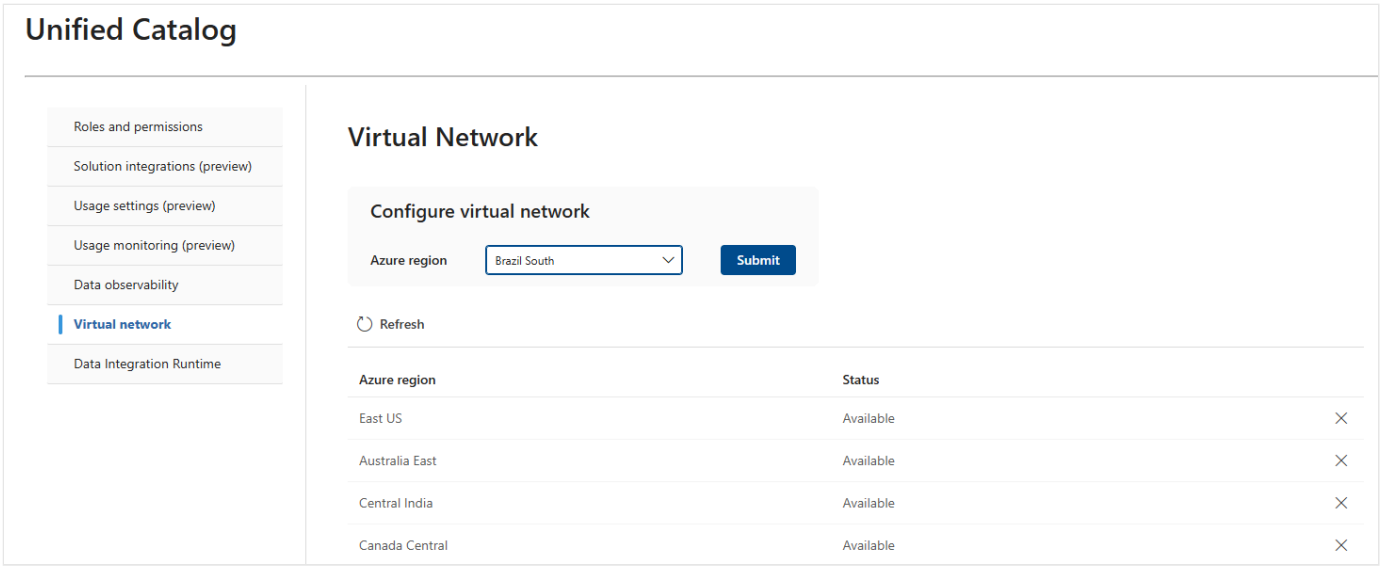
You need the following permissions to set up managed virtual networks for data quality:

- Compute provision – Governance Domain Owner
- Managed private endpoint creation - Governance Domain Owner
- Private endpoint approval – Owner of the Azure storage source

> **⊗ Caution**
>
> Compute and managed private endpoint connections are shared across all governance domains in the same Purview account. These connections are shared per region and data source.

Manage virtual network provisioning

Data Governance Administrators can provision a virtual network compute location in supported Azure regions. To provision a compute location, go to **Settings** > **Unified Catalog** > **Virtual network** in the Unified Catalog Admin page. Select a region from the **Region** dropdown. Regions that are already set up are greyed out.



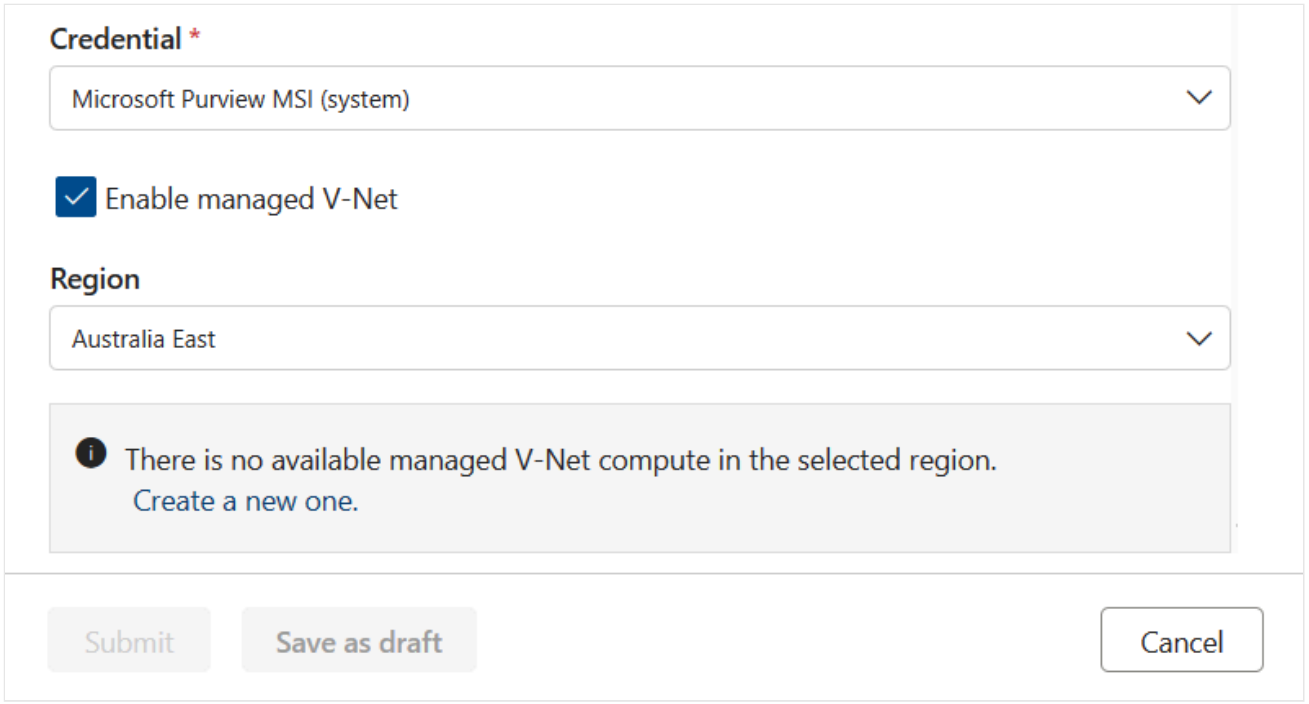
Deleting a configured region removes all associated data quality connections linked to that region's virtual network across business domains. To delete a region, select **X** on the region's row. You can't delete a region while it's being provisioned. Provisioning has three stages: Provisioning, Completed, and Available.

Configure a data quality managed virtual network

Configure a data quality managed virtual network by creating a connection to a protected data source.

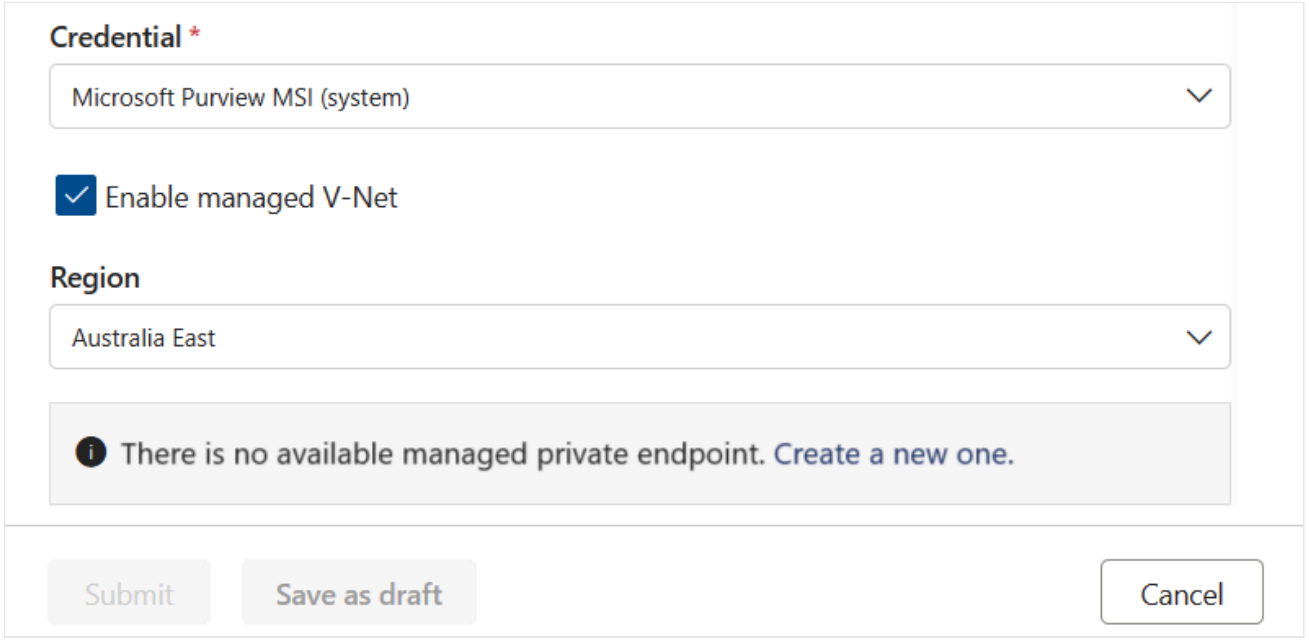
1. In the Microsoft Purview Unified Catalog, select **Health Management**, and then select **Data quality**.
2. Select a governance domain from the list.
3. From the **Manage** dropdown list, select **Connections** to open the connections page.
4. Select the **New** tab to create a new connection for the data products and data assets in your governance domain.
5. In the connection page, add a display name and description, and select the data source type.
6. Add other data source details like **Subscription** and **Storage Account** name or **Server Name** and **database name**, depending on the source.
7. Select the **Enable managed V-Net** checkbox.
8. Select the region where the data source is housed.

9. Purview Data Quality checks if a compute infrastructure already exists for the account in that region. If not, you're prompted to create a new **virtual network dedicated compute**.

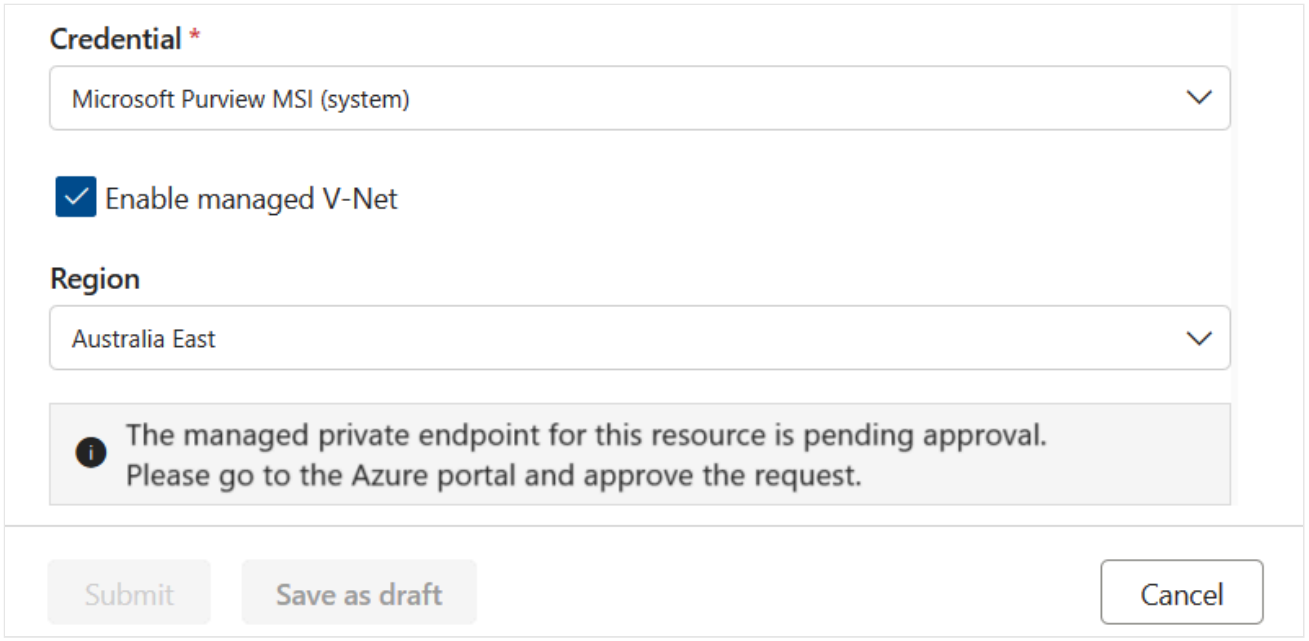


> **Tip**
>
> Provisioning compute can take up to 10 minutes. After requesting compute provisioning, you can save the connection creation request in draft mode and edit it later.

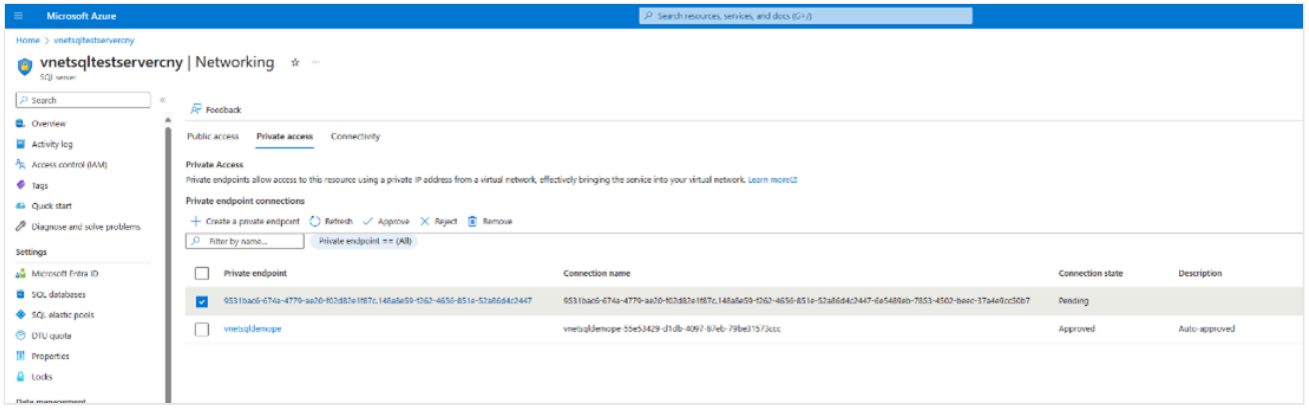
10. After the compute is provisioned, data quality checks if a **private endpoint** connection to the asset already exists. If not, you're prompted to create a private endpoint connection.



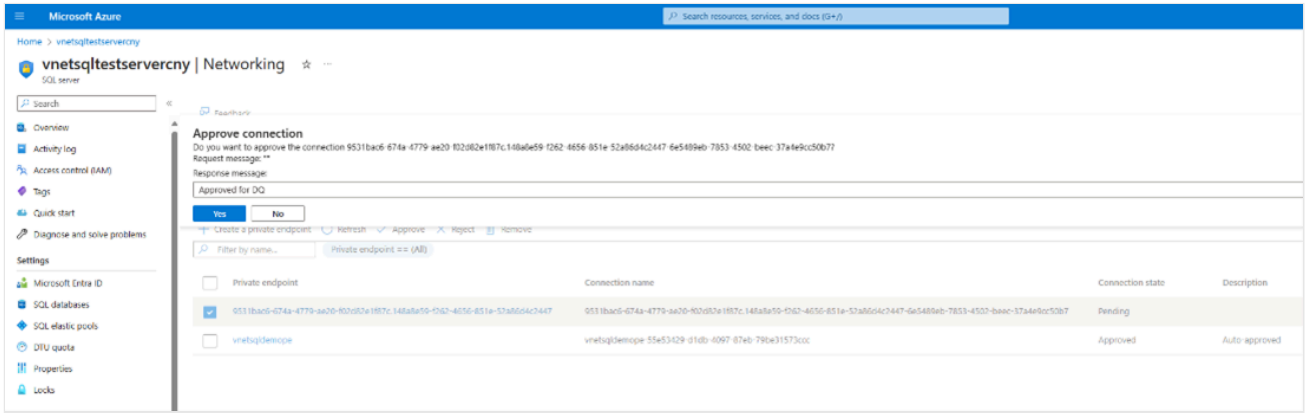
11. After the private endpoint is created, or if a private endpoint already exists but wasn't approved, you're prompted to approve the private endpoint connection request.



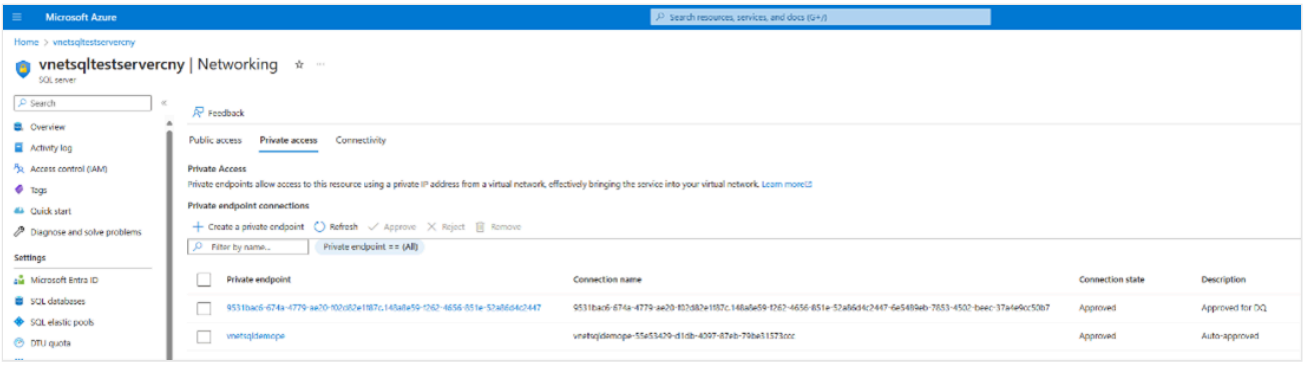
12. To approve this request, go to the Networking tab in Storage Account or SQL Server. Select the **Private access** tab, select a pending connection, and select **Approve**.



13. Select **Yes** to approve the connection.

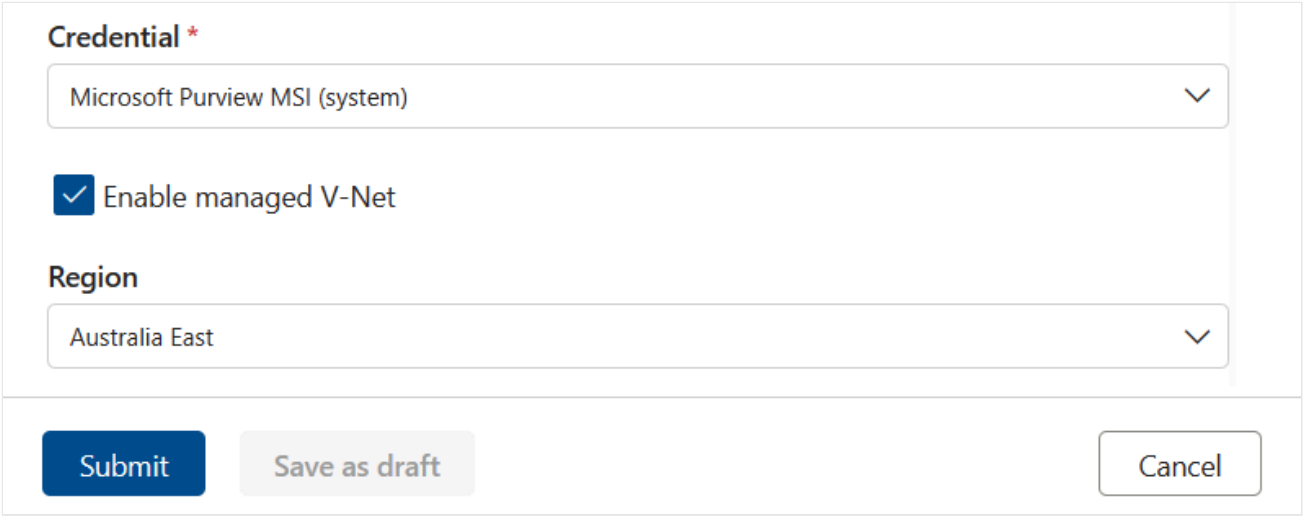


14. The request now shows as **Approved**.



> **Tip**
>
> After you request the private endpoint connection, you can save the connection as a draft and resume after the request is approved.

15. After the private endpoint connection is created and approved, submit the connection.



> **Caution**
>
> Test connection isn't currently supported for virtual network protected assets.

16. After the connection is completed, you can run data quality jobs as usual against the virtual network protected assets.

> **Note**
>
> Virtual Network currently supports only these data sources:
>
> - **Azure Data Lake storage**
> - **Azure SQL**

- **Synapse serverless and Synapse Data Warehouse**
- **Azure Databricks**
- **Snowflake**
- **Microsoft Fabric OneLake (in tenant level)**
- **Azure SQL Managed Instance (SQL MI)**

Set up data source-specific connections

To set up virtual network connections for specific data sources, see the following resources:

- Data quality for Azure Databricks UC
- Data quality for Azure Synapse serverless and data warehouse
- Data quality for Snowflake data
- Data quality for Fabric Lakehouse

For Azure SQL and Data Lake storage, follow the steps in Configure a data quality managed virtual network.

Import a schema for data quality

If the data type in a schema is undefined, incorrectly defined, or changed in the source, your data quality job might fail. If the job fails, re-import the schema by using the schema import capability. You can import schemas for data sources on both public networks and those behind private endpoints. For supported data sources, see Configure a data quality managed virtual network. To import a schema from your data sources, follow these steps:

- In Unified Catalog, under **Health Management**, select **Data quality**.
- Select a business domain, then select a data product, then select a data asset from that data product. You arrive at the data quality overview page.
- Select **Schema**, then select the **Schema management** toggle.
- Select **Import schema** to import the schema.

Last updated on 06/05/2026

Data quality native support for Iceberg format (preview)

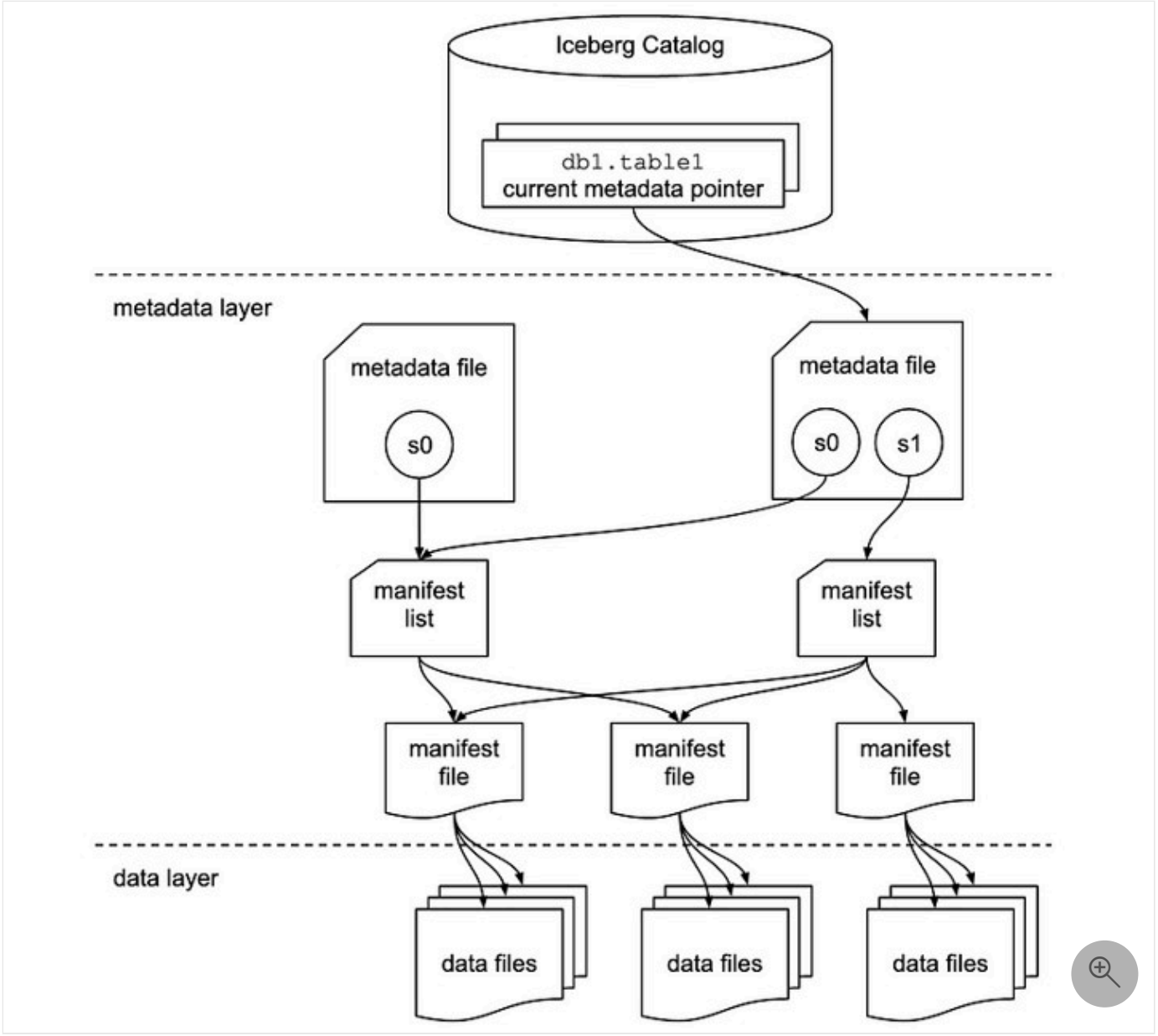
Microsoft Purview supports data quality assessment on Apache Iceberg data assets. You can curate, govern, and scan Iceberg data on these storage sources:

- Azure Data Lake Storage Gen2
- Microsoft Fabric Lakehouse
- Amazon Web Services (AWS) S3
- Google Cloud Platform (GCP) Cloud Storage (GCS)

This article explains the Iceberg file structure, shows how to set up data quality scans, and lists tips for each catalog and storage type.

Iceberg file structure

An Iceberg table is more than just a collection of data files. It includes various metadata files that track the state of the table and facilitate operations like reads, writes, and schema evolution. An Iceberg table includes a catalog, a metadata layer, and a data layer. The data files in an Iceberg table are typically stored in columnar formats like Apache Parquet, Apache Avro, or Apache Optimized Row Columnar (ORC). These files contain the actual data that users interact with during queries.



Catalog layer overview

The Iceberg catalog sits at the top of the hierarchy. It stores the current metadata pointer for each table. The catalog enables tracking the most recent state of a table by referencing the current metadata file.

Metadata layer overview

The metadata layer is central to Iceberg's functionality. It includes these key elements:

- **Metadata file:** Contains information about the table's schema, partitioning, and snapshots. In the diagram, s0 refers to a snapshot, a record of the table's state at a given point in time. If multiple snapshots exist (such as s0 and s1), the metadata file tracks both.
- **Manifest list:** Points to one or more manifest files. A manifest list acts as a container of references to these manifests. It helps Iceberg manage which data files should be read or

written during different operations. Each snapshot might have its own manifest list.

Data layer overview

In the data layer, the manifest files act as an intermediary between the metadata and the actual data files. Each manifest file points to a collection of data files, providing a map of the physical files stored in the data lake.

- **Manifest files**: Store the metadata for a group of data files, including row counts, partition information, and file paths. They allow Iceberg to prune and access specific files quickly.
- **Data files**: Contain the actual data, stored in formats like Parquet, ORC, or Avro. Iceberg organizes data files based on partitions, which minimizes unnecessary data scans during query execution.

How Iceberg components work together

When you query or update data, Iceberg looks up the table's metadata file via the catalog. The metadata file references the current snapshot (or multiple snapshots). Each snapshot points to a manifest list, which references manifest files. The manifest files list the data files. This layered lookup lets Iceberg manage large datasets while keeping reads and writes consistent. Readers and writers always see a coherent table state. The snapshot-based design also enables time travel (querying earlier table states) and schema evolution.

The same layered approach boosts performance for batch and streaming operations. Only the needed data files are read, and updates use snapshots without changing the full dataset.

Iceberg data in OneLake

> **Important**
>
> Iceberg data in AWS S3 and GCS also needs to be auto synced as Delta to curate, govern, and to measure and monitor data quality.

You can seamlessly consume Iceberg-formatted data across Microsoft Fabric with no data movement or duplication. Use OneLake shortcuts to directly point to a data layer.

Iceberg data is stored in OneLake and written using Snowflake or another Iceberg writer. OneLake virtualizes the table as a Delta Lake table, which ensures broad compatibility across Fabric engines. For example, you can create a volume in Snowflake and point it directly to the

Fabric Lakehouse. Once the Iceberg table is created in OneLake, autosync reflects data updates in real time. Get further details in Configure an Iceberg external volume for Azure in the Snowflake documentation.

Data quality for Iceberg data

For all users natively hydrating data in Iceberg on Parquet, ORC, or Avro on Data Lake Storage Gen2 or Fabric Lakehouse, configure a scan pointing to the location of the directory hosting the data and metadata Iceberg directories. To configure Iceberg data quality support, complete these steps:

1. Configure and run a scan in Microsoft Purview Data Map.
2. Configure the directory hosting data and metadata as a data asset and associate it to a data product. Associating the directory as a data asset and linking it to a data product forms the Iceberg dataset. Learn to associate data assets to a data product.
3. In Unified Catalog, under **Health management** select **Data quality** view to find your Iceberg files (data asset) and to set up data source connection.
   a. To set up a Data Lake Storage Gen2 connection, follow the steps in Set up data source connection for data quality.
   b. To set up a Fabric OneLake connection, follow the steps in Set up data quality for Fabric Lakehouse data.
4. On the **Schema** page of the selected Iceberg file (data asset), select **Import schema** to import the schema from the data source.
5. On the Iceberg file's **Overview** page, at the **Data asset** dropdown menu, select **Iceberg**.
6. Apply data quality rules, and run data quality scans for column- and table-level data quality scoring.

Profiling and data quality scanning

> **Important**
>
> Before you run data profiling or data quality scans, you need to retrieve and set the schema from the Data Quality Schema page. Consumers don't see the schema in the data asset view because Data Map doesn't yet support the Iceberg open table format. Data Quality stewards can import the schema from the Data Quality schema page.

After you finish the connection setup and data asset file format selection, you can profile your data, create and apply rules, and run a data quality scan of your data in Iceberg open format files.

For step-by-step guidance, see the following articles:

- Configure and run data profiling for a data asset
- Configure and run a data quality scan

> **Important**
>
> Support for the Iceberg open format in catalog discovery, curation, data profiling, and data quality scanning features is in **preview**.

Current limitations for Iceberg data quality

The current preview release of Microsoft Purview supports data created in Iceberg format with **Apache Hadoop** catalog (a file-based catalog implementation) only. Snowflake catalog scenarios are supported only through Delta virtualization using OneLake shortcuts.

Lakehouse Path and Data Lake Storage Gen2 Path

- Iceberg Metadata stores the complete path for the data and metadata. Ensure you use the complete path for Data Lake Storage Gen2 and Fabric Lakehouse. Additionally for Fabric Lakehouse path during the write, ensure operating (WRITES, UPSERTS) with the ID paths. The following example shows the required ABFSS path format. Replace `<filesystem-ID>` and `<lakehouse-ID>` with your actual GUIDs:

  ```text
  abfss://<filesystem-ID>@onelake.dfs.fabric.microsoft.com/<lakehouse-ID>/Files/CustomerData
  ```

- Filesystem as ID and Lakehouse as ID. Absolute and not relative paths are necessary for Microsoft Purview to perform data quality on Iceberg. To validate, ensure you check the snapshots path to point as complete Fully Qualified Name (FQN) paths.

Schema detection

- Data Map can't detect the Iceberg schema. When curating the Iceberg directories on Fabric Lakehouse or Data Lake Storage Gen2, you can't review the schema. However, the data quality fetch schema can pull up the schema for the curated asset.

Recommendations for Iceberg data quality configuration

Choose the approach that matches your catalog and storage:

Expand table

Source/Catalog	Storage	Approach	Supported formats
Snowflake Catalog	Data Lake Storage Gen2, AWS S3, or GCP GCS (VOLUME)	Use a Fabric OneLake Table shortcut. Run data quality as a Delta table.	Parquet only
Hadoop catalog	Data Lake Storage Gen2	Scan the directory directly. Use the Iceberg engine for data quality.	Parquet, ORC, Avro
Snowflake	Fabric Lakehouse (VOLUME pointed to Lakehouse)	Use OneLake Table to create a Delta-compatible version.	Parquet only
Hadoop catalog	Fabric Lakehouse	Scan the Lakehouse directory directly. Use the Iceberg engine for data quality.	Parquet, ORC, Avro

Related resources

These articles cover data quality setup and scanning:

- Set up data source connection for data quality
- Configure and run data profiling for a data asset
- Configure and run a data quality scan
- Set up data quality for Fabric Lakehouse shortcut data sources
- Data quality troubleshooting

Last updated on 06/12/2026

Set up data quality for Snowflake data

This article walks you through how to register a Snowflake source, run a Data Map scan, and configure a data quality connection in Microsoft Purview Unified Catalog. After you complete these steps, you can profile your Snowflake data and run data quality scans.

Supported Snowflake metadata

When you scan Snowflake data sources, Microsoft Purview extracts technical metadata. The scan captures the following objects:

- Servers, databases, schemas, and tables
- Views, foreign keys, and unique constraints
- Stored procedures and their parameters
- Functions and their parameters
- Pipes, stages, and streams
- Tasks and sequences

Register a Snowflake source in Data Map

To register a new Snowflake source in Microsoft Purview Unified Catalog, follow these steps:

1. Sign in to the Microsoft Purview portal.
2. Select the **Data Map** solution card. If the **Data Map** solution card isn't displayed, select **View all solutions** and then select **Data Map** from the **Core** section.
3. Select **Register**.
4. On **Register sources**, select **Snowflake**.

Set up a Data Map scan

To set up a scan for your registered Snowflake source, follow these steps:

1. Select the registered Snowflake source.
2. Select **New scan**.
3. Provide the following details:

    - Name: The name of the scan.

- Connect via integration runtime: Select the autoresolved integration runtime, Managed Virtual Network IR, or SHIR for your scenario.
- Host for connection: Select the server URL or another host you configured in the data source.
- Credential: Select the credential for your data source. Use Basic Authentication, provide the Snowflake user name, and store the password in the secret key.
- Warehouse: Specify the warehouse name in uppercase. The credential's default role must have USAGE rights on this warehouse.
- Databases: Specify one or more database names in uppercase, separated by semicolons (for example, `DB1;DB2`). The credential's default role must have rights on these database objects.
- Schema: List the schemas to import, separated by semicolons.

4. Select **Test connection** to validate the settings (available when using Azure Integration Runtime).
5. Select **Continue**.
6. Select a scan rule set for classification. You can choose between the system default, existing custom rule sets, or create a new rule set inline.
7. Review your scan and select **Save and Run**.

After the scan finishes, the data asset in Snowflake is available on the Unified Catalog search. For more information about how to connect and manage Snowflake in Microsoft Purview, see Connect to and manage Snowflake in Microsoft Purview.

> ⓘ **Important**
>
> When you delete the object from the data source, the subsequent scan doesn't automatically remove the corresponding asset in Microsoft Purview.

Set up a Snowflake data quality connection

After the scan completes, the scanned asset is ready for cataloging and governance. To run data quality scans on the asset, set up a Snowflake connection in Microsoft Purview.

> ⓘ **Important**
>
> - Data quality stewards need **read only** access to Snowflake to set up the data quality connection.

- The Snowflake connector doesn't accept **https://**. Remove **https://** when you add the server name to configure the data source connection.
- If you disable public access, you need to select **Allow trusted Microsoft services** for Key Vault. This requirement applies only to Key Vault, not to your Snowflake workspace.

1. In the Microsoft Purview portal, open **Unified Catalog**.

2. Under **Health management**, select **Data quality**.

3. Select a governance domain from the list, then select **Connections** from the **Manage** dropdown list.

4. Configure the connection on the **Connections** page:

    - Add a connection name and description.
    - Select source type **Snowflake**.
    - Add the server name, warehouse name, database name, schema name, and table name.
    - Select an authentication method: **Basic**, **Private key (preview)**, or **OAuth (preview)**.
    - Add the user name.
    - Add credentials: Azure subscription, Key Vault connection, secret name, and secret version.
    - Select **Enable managed VNet** if your Snowflake instance runs on Azure Virtual Network, then add the Azure region, Private Link Resources ID, and Fully Qualified Domain Name.

5. Test the connection to make sure it works. If you're using Virtual Network, the test connection feature isn't supported.

Overview

Display name *

Snowflake-Connection

Show ID

Description

Snowflake data source connection

Source type *

Snowflake

Server * ⓘ

bta81782.east-us-2.azure.snowflakecomputing.com

Warehouse * ⓘ

DQWAREHOUSE

Database * ⓘ

EMPLOYEEDB

Schema name (for testing connection only) *

EMP

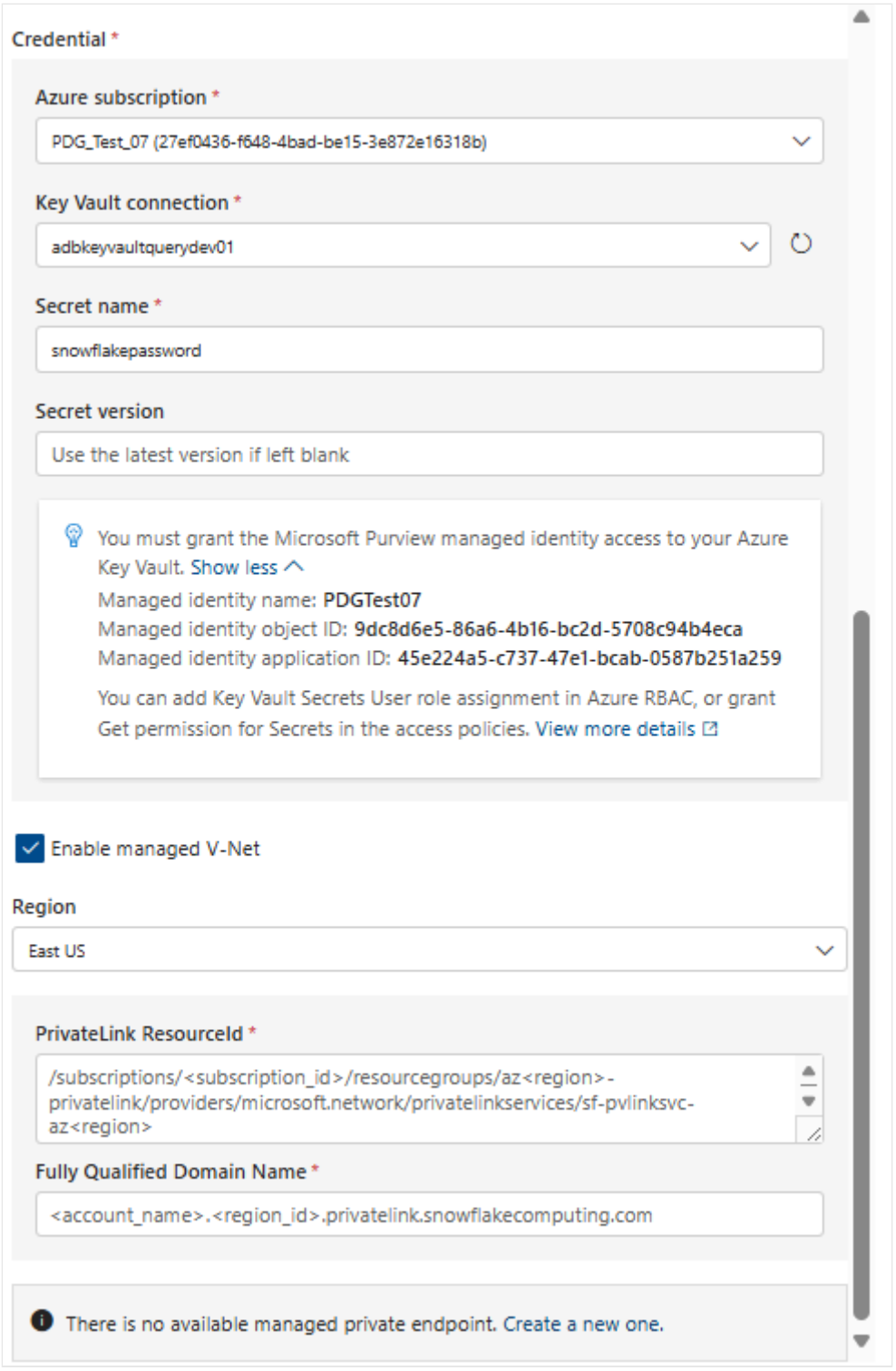
Table name (for testing connection only) *

TITANIC

Authentication method *

Basic authentication

User name *



The resource ID for the target Snowflake private link has the following format: `/subscriptions/(subscription_id)/resourcegroups/az(region)-privatelink/providers/microsoft.network/privatelinkservices/sf-pvlinksvc-az(region)`.

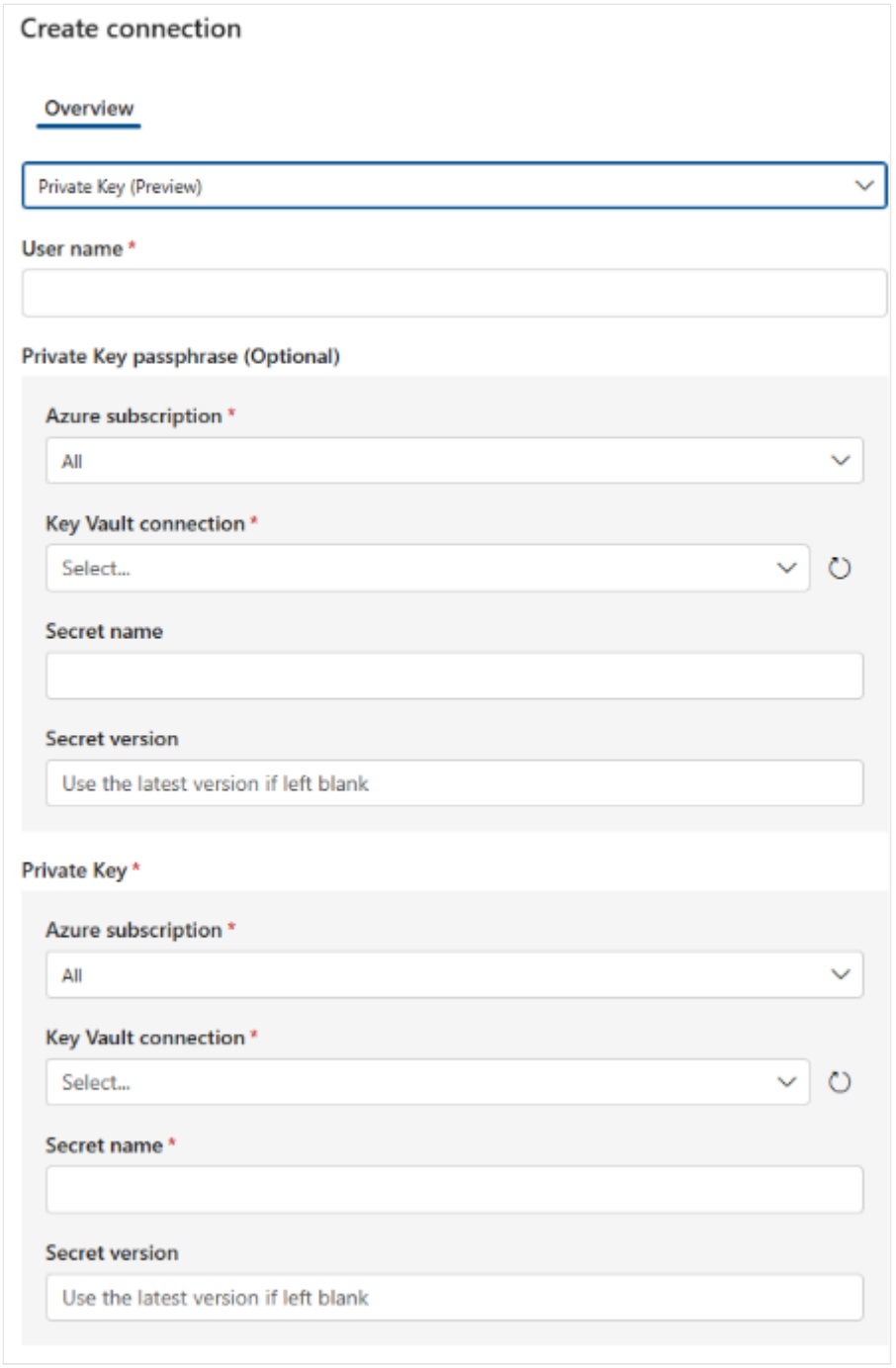
- To obtain region-ID and fully qualified name, run `SYSTEM_WHITE_LIST` and `SYSTEM_WHITE_LIST_PRIVATELINK` to get the `SNOWFLAKE_DEPLOYMENT`, `SNOWFLAKE_DEPLOYMENT_REGIONLESS`, and `OCSP_CACHE` values for public and allow list hosts.
- To obtain Subscription ID, run `SYSTEM$GET_SNOWFLAKE_PLATFORM_INFO()` as `ACCOUNTADMIN` to obtain the **snowflake-vnet-subnet-ids** values. The Subscription ID for the Private Link of the Snowflake Azure tenant is obtained from the **snowflake-vnet-subnet-ids** output.

Private key authentication (preview)

You can create private keys with or without a passphrase by using OpenSSL, a command-line tool for generating cryptographic keys. If you generate a private key with a passphrase, configure both the key and passphrase with the connection as shown in the following screenshot. If you don't provide a passphrase during key generation, omit it in the connection.

The private key generated through the **openssl** command contains `headers/footers` that indicate the beginning and ending of the `key/encrypted` key. Save the entire contents of the file, including headers and new lines, as-is in the Key Vault as the secret value. You can use a command-line tool such as Azure CLI to do this. Without the private key file headers, footers, and original newline formatting, authentication with Snowflake fails.

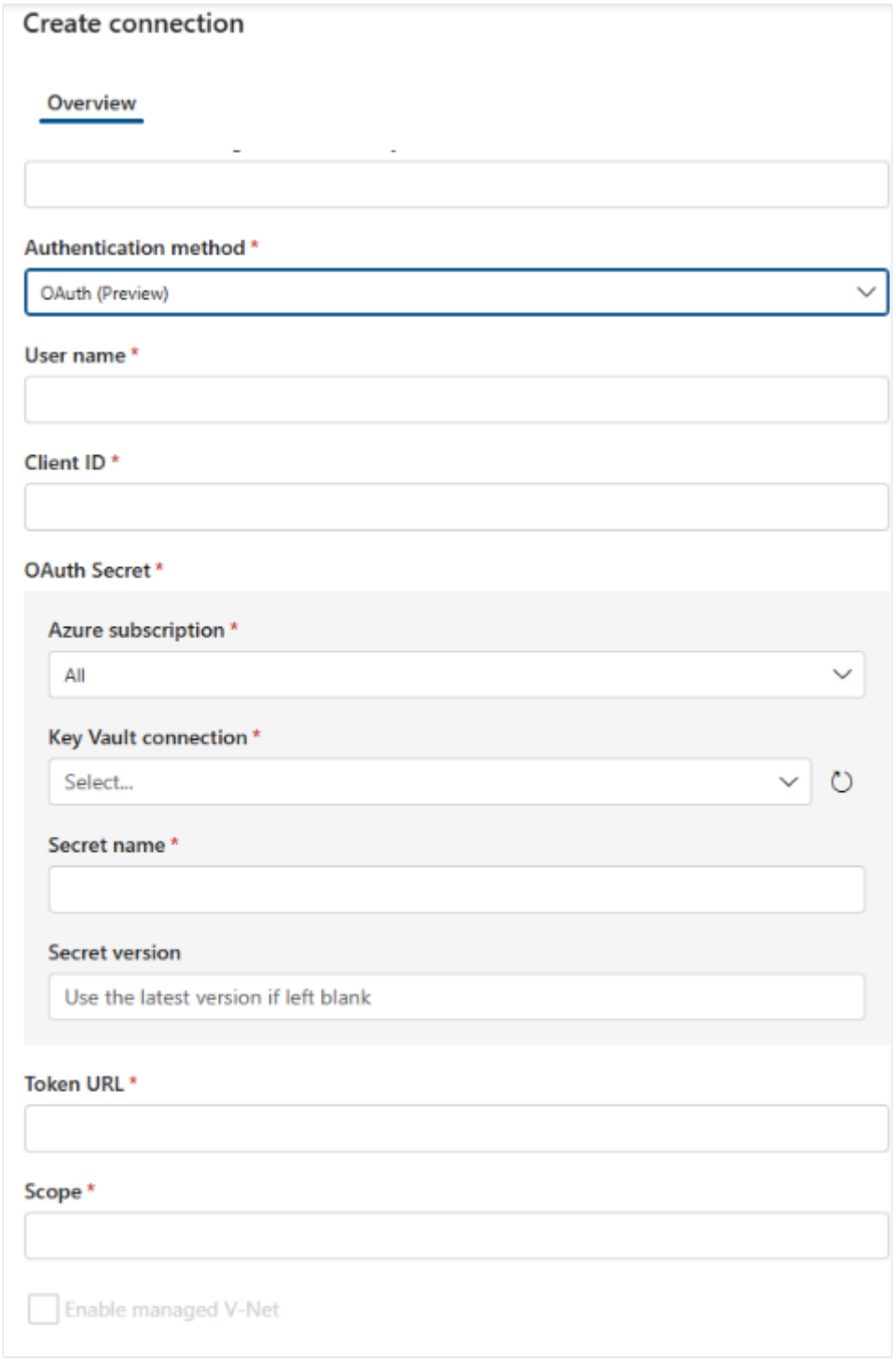
For security, rotate your keys regularly. When you rotate a key, update it in Key Vault and in the Microsoft Purview connection. New jobs use the updated key. Keep the older key active until any in-progress jobs complete.



OAuth authentication (preview)

You can configure an OAuth-based connection by using Microsoft Entra ID as the OAuth provider. OAuth access tokens are typically short-lived. But data quality assessment jobs can run for a long time, depending on data volume. Set the OAuth access token lifetime to at least six hours, which is the average maximum job duration. Jobs fail if the access token expires before the job completes.

If you can't increase the access token lifetime, or if your policy doesn't allow it, use key-pair-based authentication because key-pair auth doesn't have token lifetime constraints.



Sample values for the OAuth connection parameters:

- User name: DB41C9CD-96E7-47B0-ADB6-0A4D080229A0 (Value of `login_name` property of Snowflake)
- Client ID: 632ce887-41bb-4975-b6a9-ae798bf41956 (Client ID from Azure portal)
- Token URL: `https://login.microsoftonline.com/ddfd8861-7cbd-43df-ad13-4f9e958cf53b/oauth2/v2.0/token (Token URL from Azure app)`
- Scope: `api://d682fa4a-d784-4d28-93e2-7156ff129ae0/.default (Scope as registered with the Azure app)`

Understand connection selection

You don't explicitly choose connections when you start assessments or configure schedules. The system selects a connection automatically based on the following rules:

- If only user-password connections exist, the system selects the first one created.
- If key-pair or OAuth connections exist, the system ignores all user-password connections and selects the first key-pair or OAuth connection created.

The following illustration shows how the system selects a connection based on type and creation order.

Suppose there are only user-password-based connections:

- SnowflakeConnectionBob (User-password)
- SnowflakeConnectionAlice (User-password)
- SnowflakeConnectionEve (User-password)

The system selects the **SnowflakeConnectionBob** connection.

Suppose all types of connections exist and you created them in the following order:

- SnowflakeConnectionBob (User-password)
- SnowflakeConnectionAlice (Key-pair)
- SnowflakeConnectionCharlie (OAuth)
- SnowflakeConnectionEve (User-password)

In this case, the system ignores **SnowflakeConnectionBob** and **SnowflakeConnectionEve**. The system considers only the following connections for selection:

- SnowflakeConnectionAlice (Key-pair)
- SnowflakeConnectionCharlie (OAuth)

Typically, the system selects **SnowflakeConnectionAlice** for assessment and scheduled runs.

> **Important**
>
> After you create the request for a private endpoint connection, raise a support ticket with Snowflake support to approve the private endpoint connection. Provide the details of the managed private endpoint resource ID in this support ticket for the Snowflake team to approve.

> **Note**

Open Authorization (OAuth) and key pair authentication for Snowflake database connections will come. With these options, you can create multiple connections that each use a different authentication method. When multiple connections use different authentication mechanisms, the system picks the last created connection. For example, if you first create a key pair-based connection followed by an OAuth connection, the system picks the OAuth connection. If you want the key pair connection to be picked, you need to delete the OAuth connection. Contact your Microsoft account representative to preview the new Snowflake authentication model for Microsoft Purview data quality scans.

Next steps

After you set up the connection, you can profile your data, create and apply rules, and run a data quality scan. For step-by-step guidance, see the data profiling and data quality scan articles:

- Configure and run data profiling
- Configure and run a data quality scan

Related content

- Connect to and manage Snowflake in Microsoft Purview
- Set up data source connections for data quality
- Configure Azure Virtual Network
- Configure private endpoint to Snowflake Private Link
- Data quality troubleshooting

ⓘ **Note:** The author created this article with assistance from AI. Learn more

Last updated on 06/02/2026

Configure and run data profiling for a data asset

Data profiling is the process of examining the data available in different data sources and collecting statistics and information about this data. Data profiling helps you assess the quality level of the data according to a defined set of goals. If data is of poor quality, or managed in structures that can't be integrated to meet the needs of the organization, business processes and decision making suffer. Data profiling allows you to understand the trustworthiness and quality of your data, which is a prerequisite for making data-driven decisions that boost revenue and foster growth.

Prerequisites

- To run and schedule data quality assessment scans, your users must be in the data quality steward role.
- Currently, you can set the Microsoft Purview account to allow public access or managed virtual network access so data quality scans can run.

Data quality life cycle

Data profiling is the **fifth** step in the data quality life cycle for a data asset. The previous steps are:

1. Assign users(s) data quality steward permissions in Unified Catalog to use all data quality features.
2. Register and scan a data source in your Purview Data Map.
3. Add your data asset to a data product.
4. Set up a data source connection to prepare your source for data quality assessment.

Supported multicloud data sources

Browse the supported data source document to view the list of supported data sources, including file formats for data profiling and data quality scanning, with and without virtual network support.

> **Important**

Data quality for Parquet file is designed to support:

1. A directory with Parquet Part File. For example: **./Sales/{Parquet Part Files}**. The Fully Qualified Name must follow `https://(storage account).dfs.core.windows.net/(container)/path/path2/{SparkPartitions}`. Make sure not to have {n} patterns in directory or subdirectory structure; it must be a direct FQN leading to {SparkPartitions}.
2. A directory with Partitioned Parquet Files, partitioned by columns within the dataset like sales data partitioned by year and month. For example: **./Sales/{Year=2018}/{Month=Dec}/{Parquet Part Files}.**

Both of these essential scenarios, which present a consistent Parquet dataset schema, are supported. **Limitation:** It isn't designed to or won't support N arbitrary hierarchies of directories with Parquet files. **We recommend presenting data in (1) or (2) constructed structure.**

Supported authentication methods

Currently, Microsoft Purview can only run data quality scans by using Managed Identity as authentication option. Data quality services run on **Apache Spark 3.5** and **Delta Lake 3.2.1**. For more information about supported regions, see the data quality overview.

ⓘ Important

- If you update the schema on the data source, you need to import schema from data quality overview page by using **schema import feature** before running data profiling.
- Virtual network isn't supported for Google BigQuery.
- In the current version, you can profile 50 columns per batch. If your data asset has more than 50 columns, you can profile extra columns in more batches.
- Profiling is using random sample from the selected data asset, the sample size is 1 million records.
- If a column contains distinct value, don't profile that column. A column with distinct values can't create a normal distribution.

Steps to configure data profiling job

1. Configure a data source connection to the assets you're scanning for data quality if you haven't already done so.
2. From Microsoft Purview Unified Catalog, select **Health Management**, then select **Data quality**.
3. Select a **governance domain** from the list.
4. Select a **data product** to profile a **data asset** linked to that product.
5. Select a data asset to navigate into data quality **Overview** page for profiling.
6. Select **Profile** button to run profiling job for the selected data asset.
7. The AI recommendation engine suggests potentially important columns to run data profiling against. You can deselect recommended columns and/or select more columns to be profiled.
8. Once you've selected the relevant columns, select **Run Profile**.
9. While the job is running, you can track its progress from the data quality monitoring page in the governance domain.
10. When the job is complete, select the **Profile** tab from left menu of the asset's data quality page to list browse the profiling result and statistical snapshot. There could be several profile result pages depending on how many columns your data assets have.
11. Browse the profiling results and statistical measures for each column.

Understanding profiling results

- **Minimum**: Minimum length of the characters in the column.
- **Maximum**: Maximum length of the characters in the column.
- **Distribution**: Normal distribution of your data.
- **Unique values**: Values that don't repeat; they're unique.
- **Duplicate values**: Values that repeat more than once.
- **Empty/blank values**: The column is blank or empty.
- **Null**: Null values in the column.
- **Average**: Mid value or arithmetic average.
- **Standard deviation**: A statistical measure that quantifies the amount of variation or dispersion in a set of values.

Related content

- Data quality for Fabric data estate
- Data quality for Fabric Mirrored data sources
- Data quality for Fabric shortcut data sources
- Data quality for Azure Synapse serverless and data warehouses
- Data quality for Azure Databricks Unity Catalog
- Data quality for Snowflake data sources
- Data quality for Google BigQuery

Next steps

1. Set up data quality rules based on the profiling results, and apply them to your data asset.
2. Configure and run a data quality scan on a data product to assess the quality of all supported assets in the data product.
3. Review your scan results to evaluate your data product's current data quality.

Last updated on 09/09/2026

Create data quality rules

Data quality measures the integrity of data in an organization. You assess data quality by using data quality scores. Microsoft Purview Unified Catalog generates scores based on the assessment of the data against rules that you define.

Data quality rules are essential guidelines that organizations establish to ensure the accuracy, consistency, and completeness of their data. These rules help maintain data integrity and reliability.

Here are some key aspects of data quality rules:

- **Accuracy**: Data should accurately represent real-world entities. Context matters! For example, if you're storing customer addresses, ensure they match the actual locations.
- **Completeness**: This rule identifies empty, null, or missing data. It validates that all values are present, though not necessarily correct.
- **Conformity**: This rule ensures that the data follows data formatting standards such as representation of dates, addresses, and allowed values.
- **Consistency**: This rule checks that different values of the same record conform to a given rule and that there are no contradictions. Data consistency ensures that the same information is represented uniformly across different records. For instance, if you have a product catalog, consistent product names and descriptions are crucial.
- **Timeliness**: This rule aims to ensure that the data is accessible in as short a time as possible. It ensures that the data is up to date.
- **Uniqueness**: This rule checks that values aren't duplicated. For example, if there's supposed to be only one record per customer, then there aren't multiple records for the same customer. Each customer, product, or transaction should have a unique identifier.

Data quality life cycle

Creating data quality rules is the **sixth** step in the data quality lifecycle. The previous steps are:

1. Assign users data quality steward permissions in Unified Catalog to use all data quality features.
2. Register and scan a data source in Microsoft Purview Data Map.

3. Add your data asset to a data product.
4. Set up a data source connection to prepare your source for data quality assessment.
5. Configure and run data profiling for an asset in your data source.

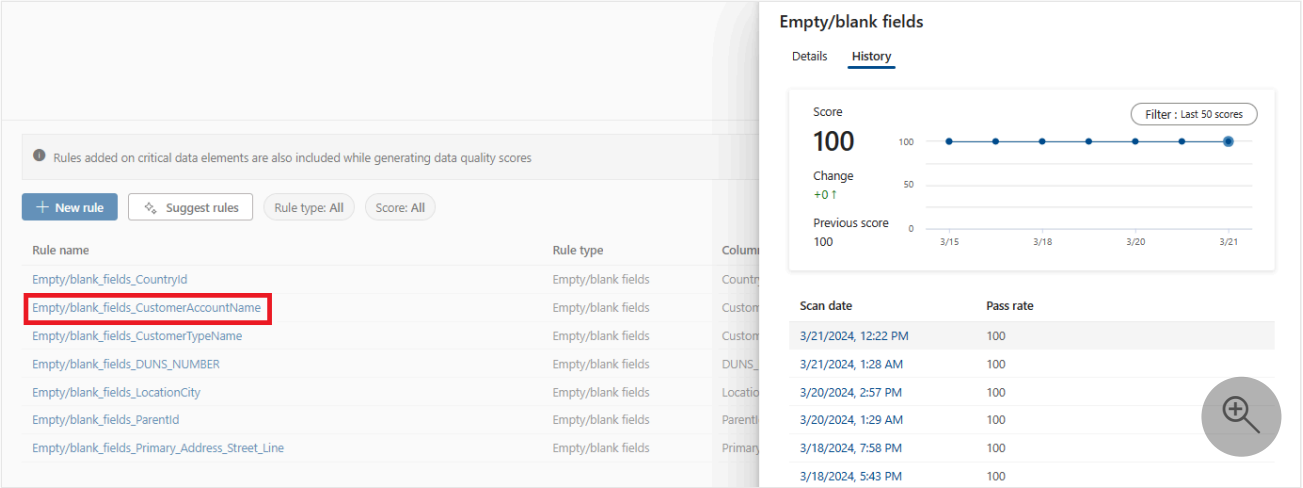
Required roles

- To create and manage data quality rules, users need the data quality steward role.
- To view existing quality rules, users need the data quality reader role.

> **Note**
>
> When you create a rule through the API, first create the associated asset metadata. Rule creation depends on this metadata. If the asset metadata doesn't exist, the rule creation request fails. For a successful workflow:
>
> 1. Create the asset metadata.
> 2. Verify that the asset is successfully registered.
> 3. Create the data quality rule and associate it with the asset.

View existing data quality rules

1. In Unified Catalog, select **Health Management**, then select **Data quality**.
2. Select a **governance domain**, then select a data product.
3. Select a data asset from the **Data assets** list.
4. Select the **Rules** tab to see the existing rules applied to the asset.
5. Select a rule to browse the performance history of the applied rule to the selected data asset.



Available data quality rules

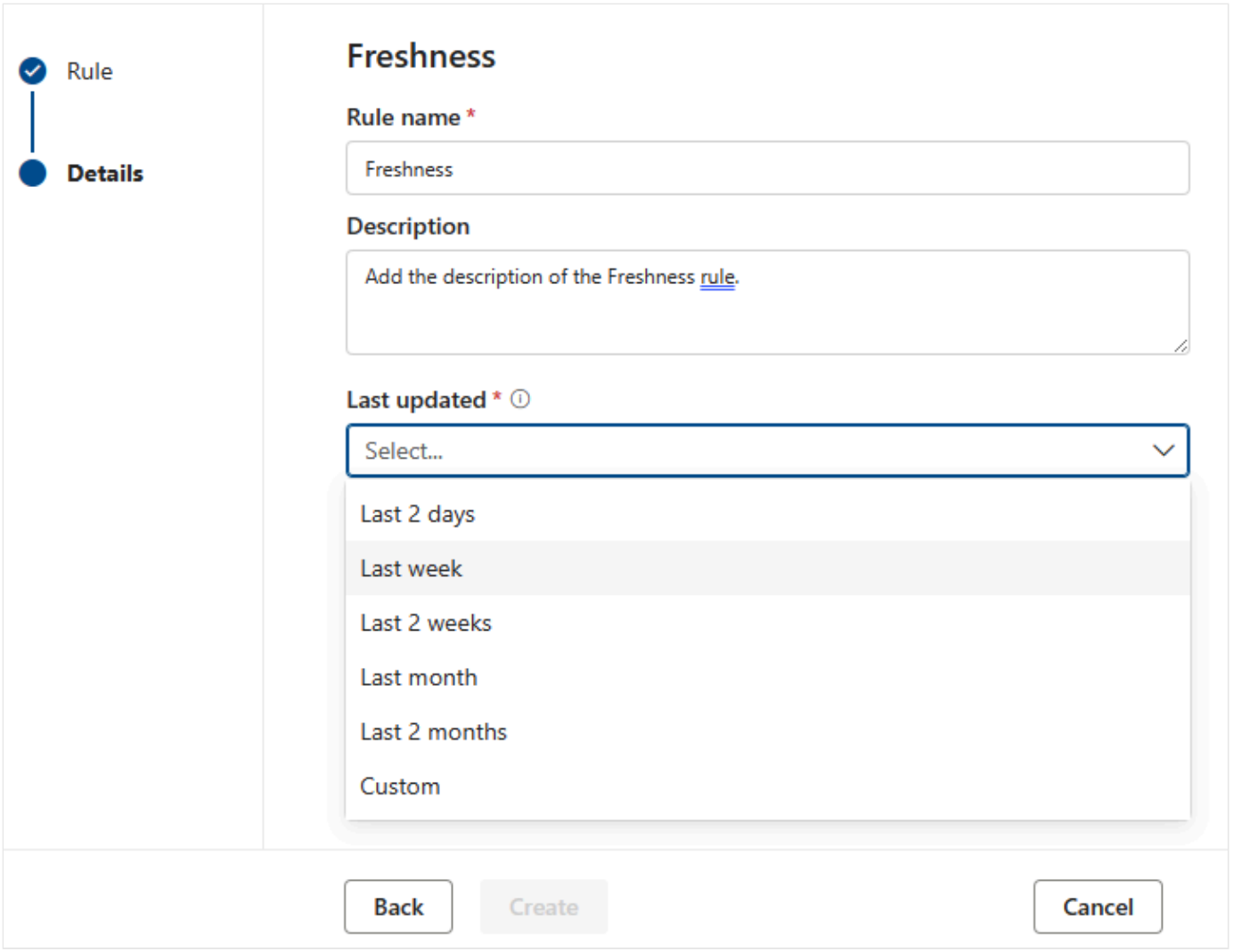
Microsoft Purview Data Quality enables configuration of the following rules. These rules are available out of the box and offer a low-code to no-code way to measure the quality of your data.

Expand table

Rule	Definition
Freshness	Confirms that all values are up to date.
Unique values	Confirms that the values in a column are unique.
String format match	Confirms that the values in a column match a specific format or other criteria.
Data type match	Confirms that the values in a column match their data type requirements.
Duplicate rows	Checks for duplicate rows with the same values across two or more columns.
Empty/blank fields	Looks for blank and empty fields in a column where there should be values.
Table lookup	Confirms that a value in one table can be found in the specific column of another table.
Custom	Create a custom rule with the visual expression builder.

Freshness

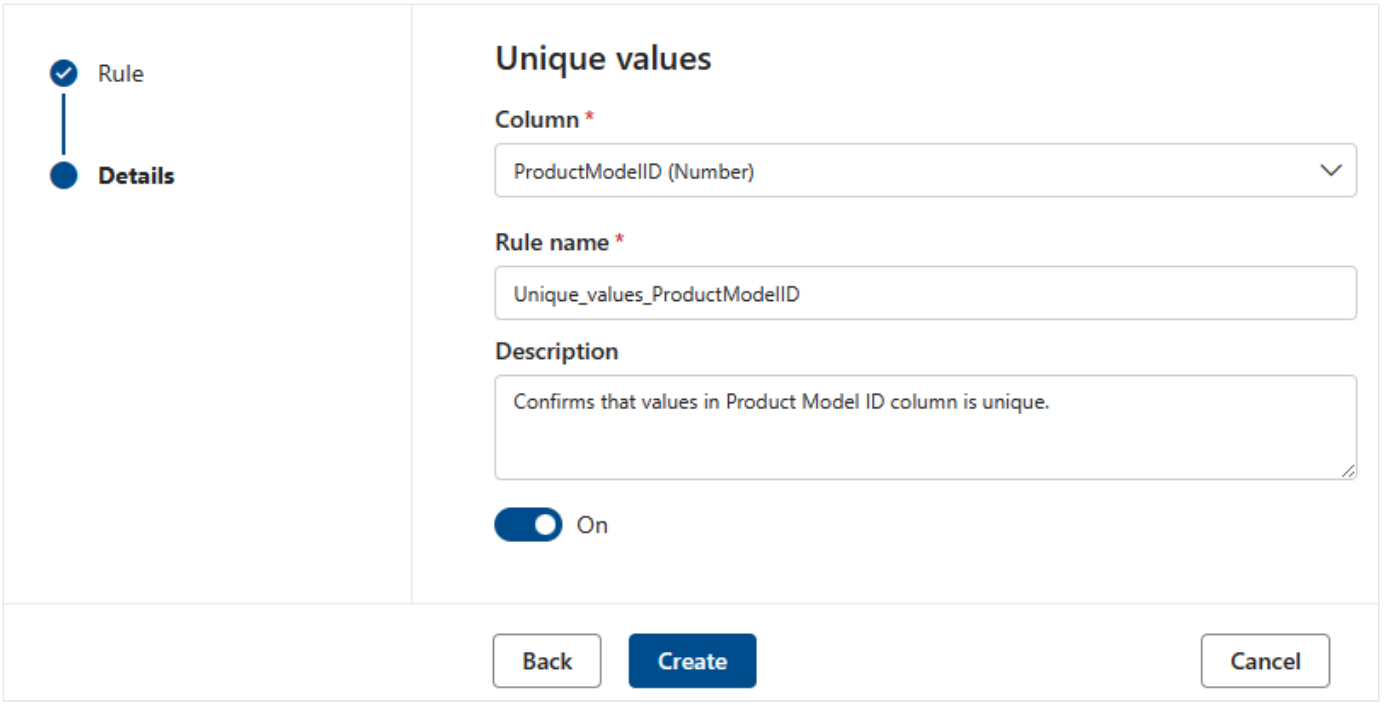
The **freshness rule** checks if the asset is updated within the expected time. Freshness is determined by the selection of *last modified dates*.



> **Note**
>
> The freshness rule score is either 100 (pass) or 0 (fail). The freshness rule isn't supported for Snowflake, Azure Databricks Unity Catalog, Google BigQuery, Synapse, and Microsoft Azure SQL.

Unique values

The **Unique values** rule states that all the values in the specified column must be unique. All values that are unique are treated as pass, and values that aren't unique are treated as fail. If the **Empty/blank fields** rule isn't defined on the column, then null or empty values are ignored for the purposes of this rule.

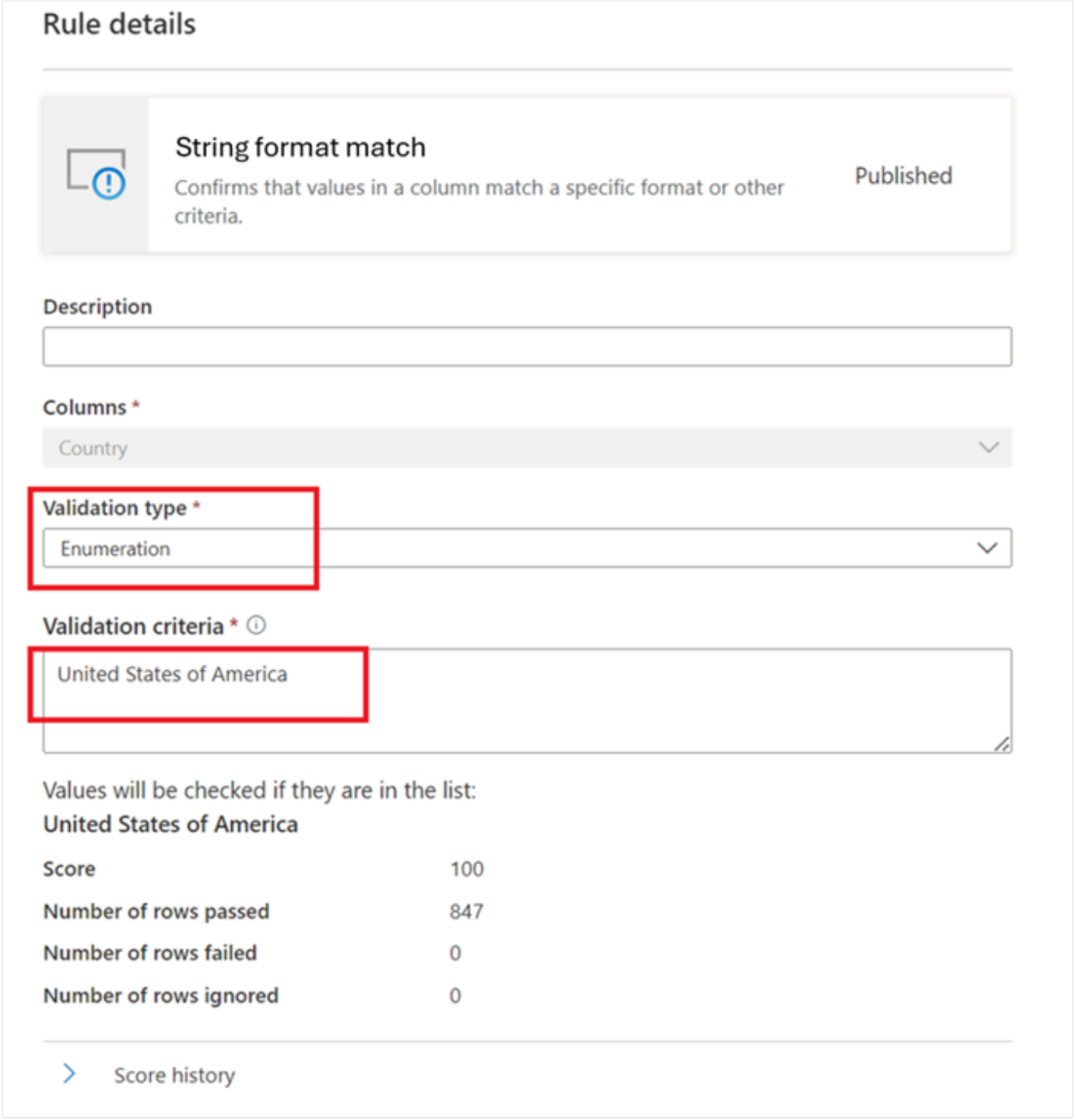


String format match

The **Format match** rule checks if all the values in the column are valid. If you don't define the **Empty/blank fields** rule on a column, the rule ignores null or empty values.

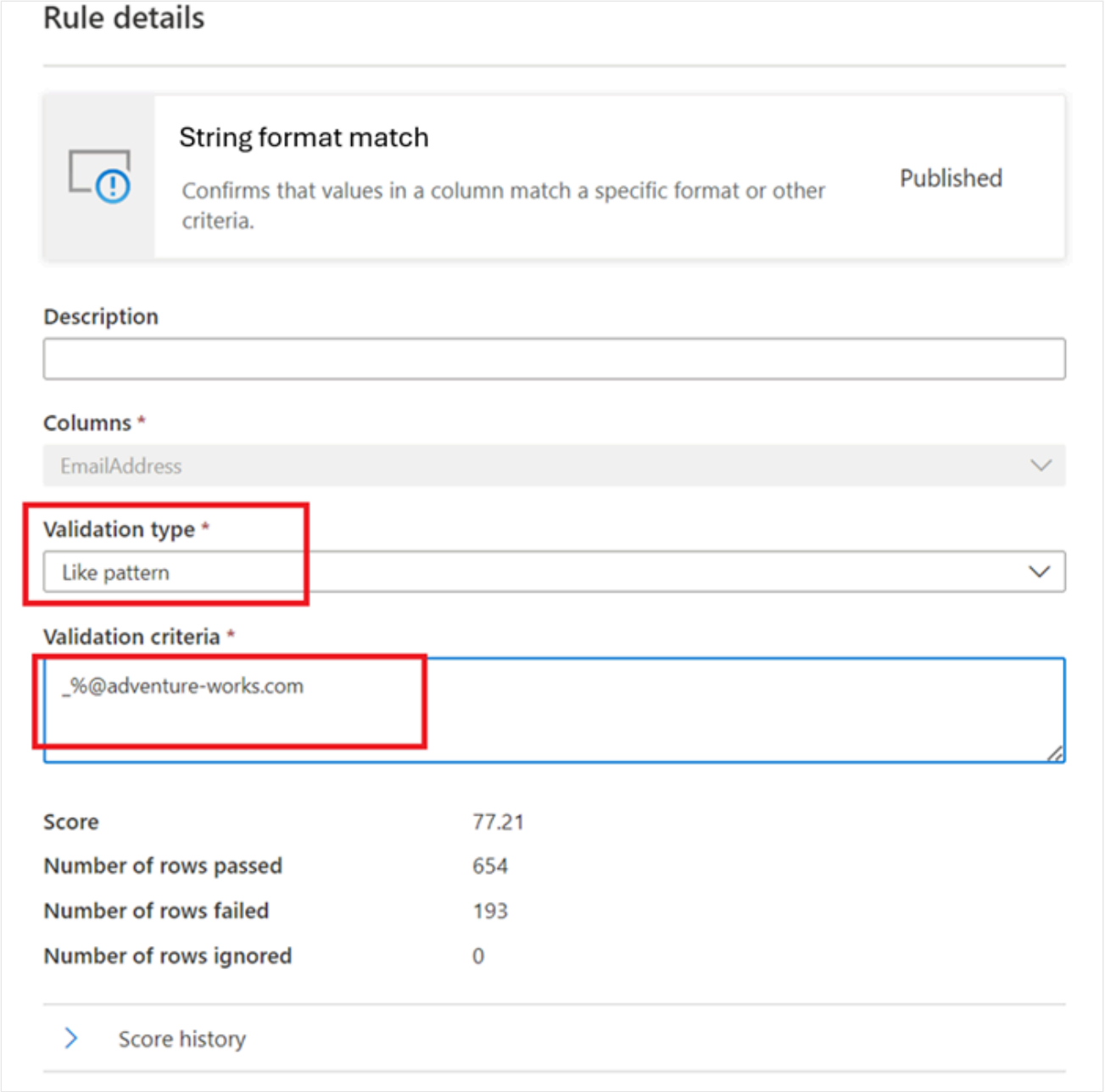
This rule can validate each value in the column by using three different approaches:

1. **Enumeration**: This approach uses a comma-separated list of values. If the value you evaluate doesn't match one of the listed values, it fails the check. You can escape commas and backslashes by using a backslash (`\`). So, `a \, b, c` contains two values: the first is `a , b` and the second is `c`.



1. **Like Pattern**: `like(<i><string></i> : string, <i><pattern match></i> : string) => boolean` The pattern is a string that the rule matches literally. The exceptions are the following special symbols: _ matches any one character in the input (similar to `.` in `posix` regular expressions) % matches zero or more characters in the input (similar to `.` in `posix` regular expressions). The escape character is `.` If an escape character precedes a special symbol or another escape character, the following character is matched literally. It's invalid to escape any other character.

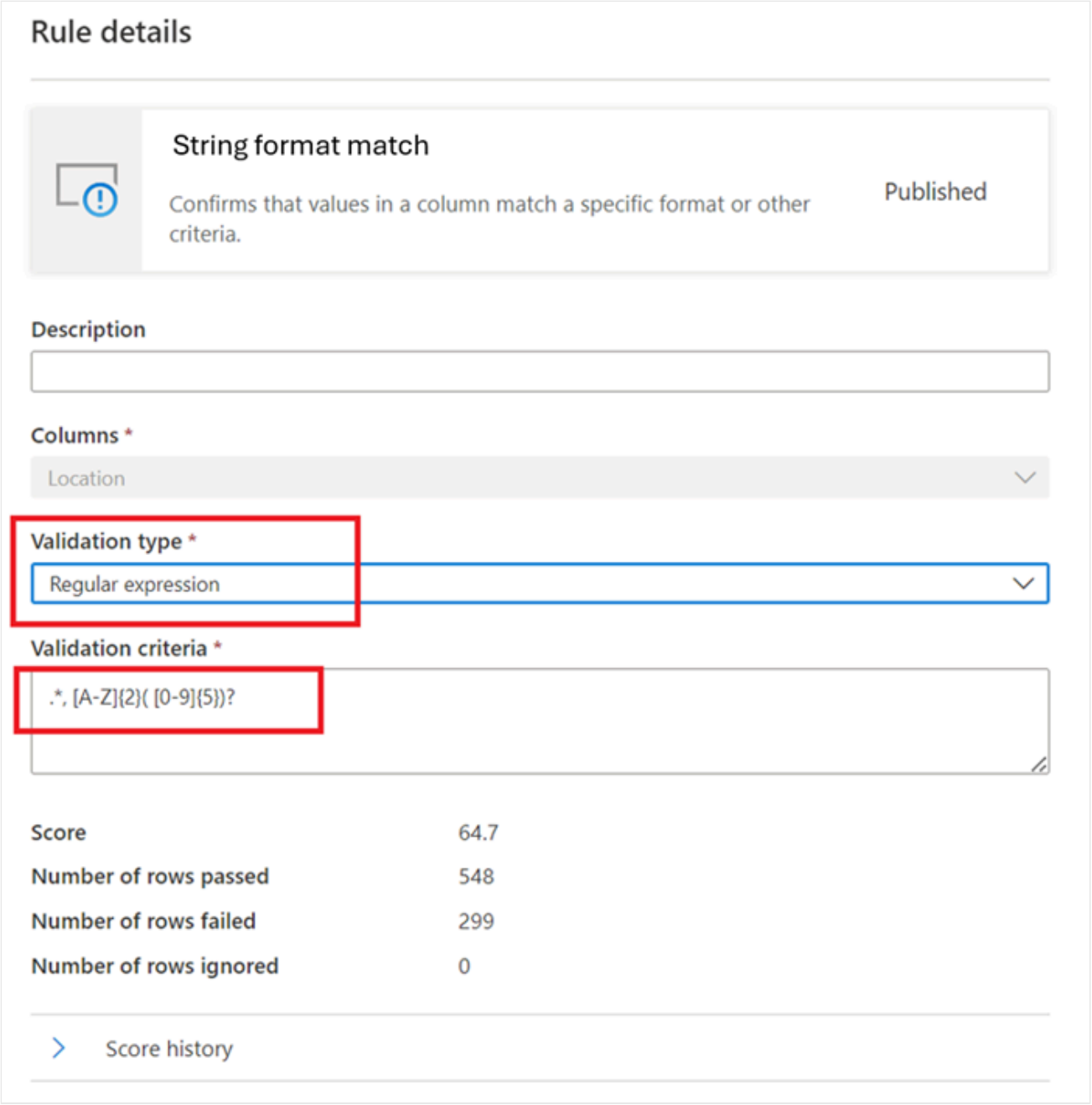
    - `like('icecream', 'ice%') -> true`



2. **Regular Expression**: `regexMatch(<i><string></i> : string, <i><regex to match></i> : string) => boolean`

   Checks if the string matches the given regex pattern. Use `<regex>` (back quote) to match a string without escaping.

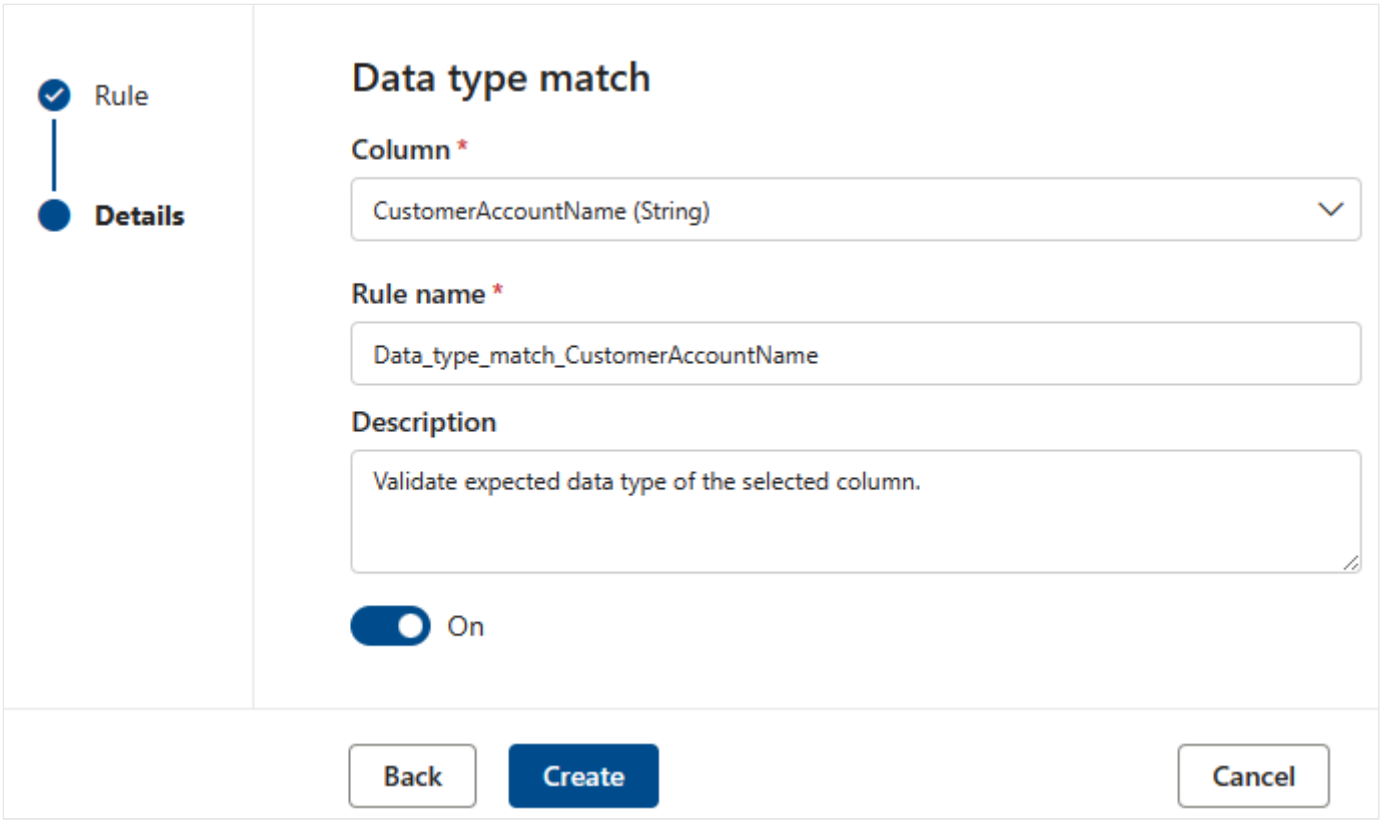
   - `regexMatch('200.50', '(\\d+).(\\d+)') -> true`
   - `regexMatch('200.50', `(\d+).(\d+)`) -> true`



Data type match

The **Data type match** rule specifies the expected data type for the associated column. Since the rule engine runs across many different data sources, it can't use native types like BIGINT or VARCHAR. Instead, it uses its own type system and translates native types into this system. This rule tells the quality scan engine which of its built-in types to use for the native type. The data type system comes from the Microsoft Azure Data Flow type system used in Azure Data Factory.

During a quality scan, the engine tests all native types against the data type match type. If it can't translate the native type into the data type match type, it treats that row as an error.

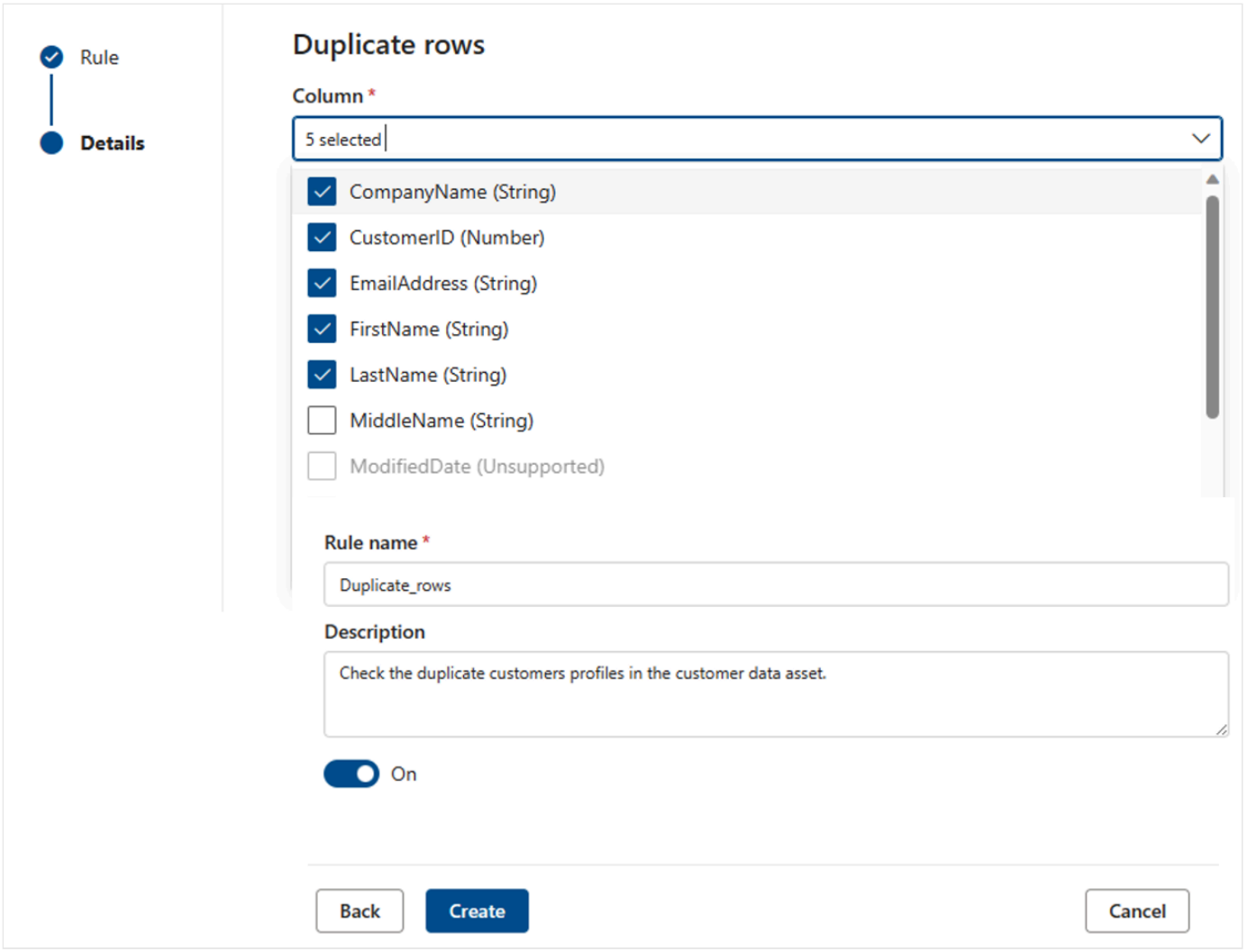


Duplicate rows

The **Duplicate rows** rule checks if the combination of the values in the column is unique for every row in the table.

In the following example, the expectation is that the concatenation of *CompanyName*, *CustomerID*, *EmailAddress*, *FirstName*, and *LastName* produces a value that's unique for all the rows in the table.

Each asset can have zero or one instance of this rule.



Empty/blank fields

The **Empty/blank fields** rule asserts that the identified columns shouldn't contain any null values. For strings, the rule also disallows empty or whitespace-only values. During a data quality scan, the engine treats any value in this column that isn't null as correct. This rule affects other rules such as the **Unique values** or **Format match** rules. If you don't define this rule on a column, those rules automatically ignore any null values when they run on that column. If you define this rule on a column, those rules examine null or empty values on that column and consider them for score purposes.

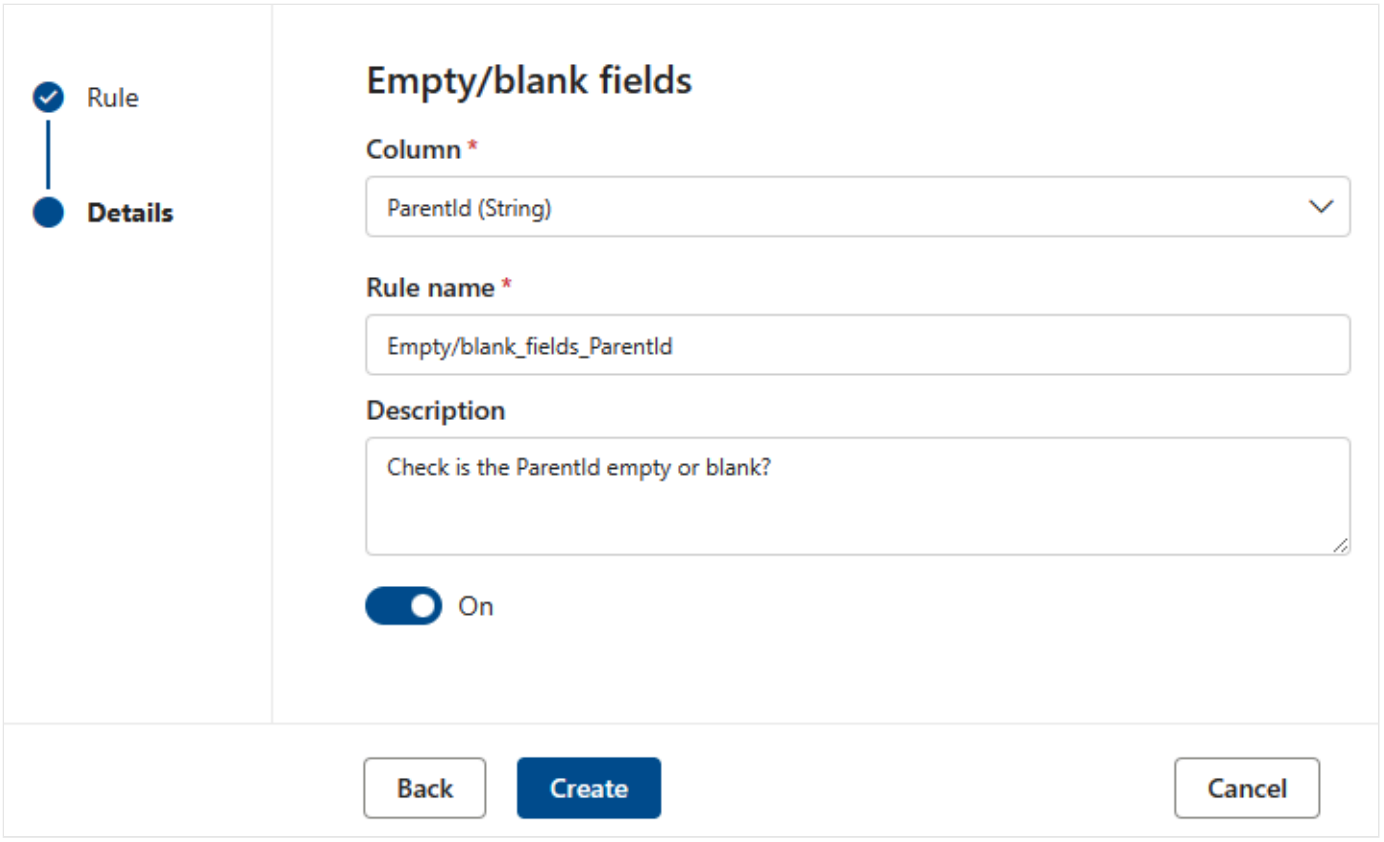
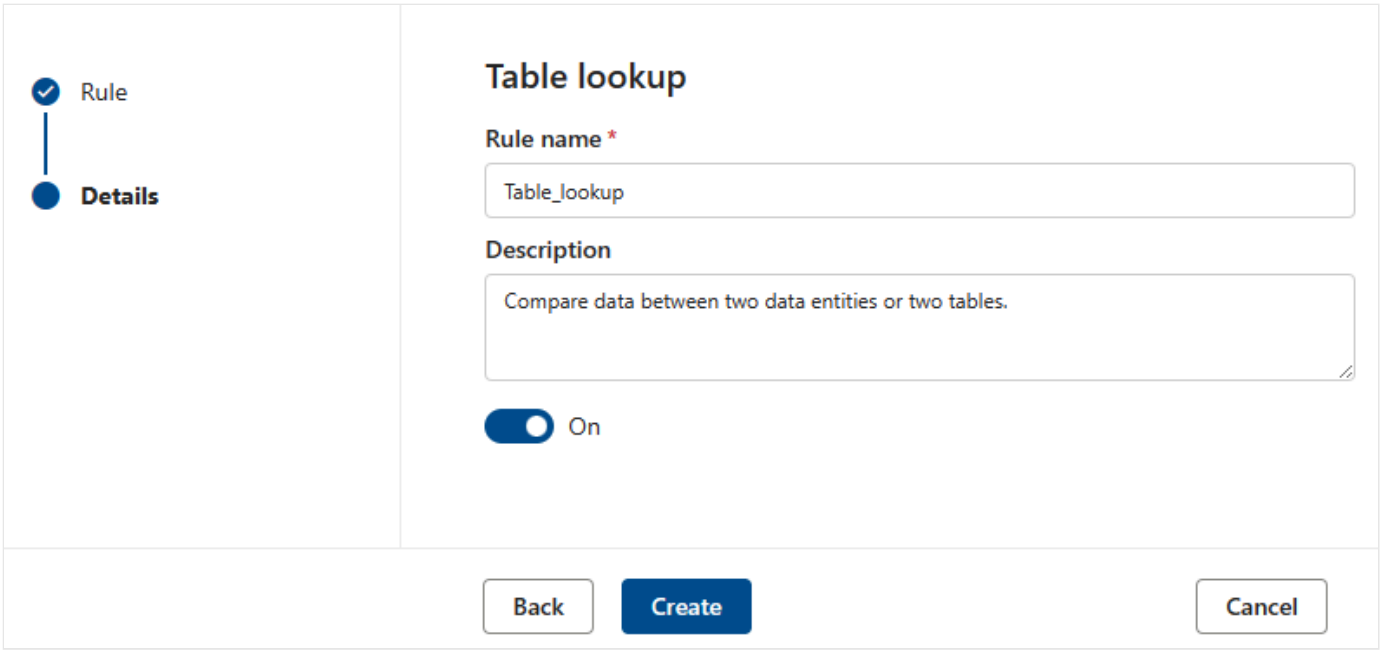


Table lookup

The **Table lookup** rule examines each value in the column where you define the rule and compares it to a reference table. For example, a primary table has a column called "location" that contains cities, states, and zip codes in the form "city, state zip". A reference table called "citystate" contains all the legal combinations of cities, states, and zip codes supported in the United States. The goal is to compare all the locations in the current column against that reference list to make sure that only legal combinations are used.

To set up this rule, enter the "citystatezip" name into the search assets dialog. Then select the desired asset and the column you want to compare against.



> **Note**
>
> The reference table or data asset must belong to the same governance domain. You can't compare a data asset across different governance domains.

Custom rules

The **Custom** rule enables you to specify rules that validate rows based on one or more values in that row. You can use regular expression language, Azure Data Factory expression, and SQL expression language to create custom rules.

A custom rule has three parts:

1. **Row expression**: This Boolean expression applies to each row that the filter expression approves. If this expression returns true, the row passes. If it returns false, the row fails.

2. **Filter expression**: This optional condition narrows down the dataset on which the row condition is evaluated. You activate it by selecting the **Use filter expression** checkbox. This expression returns a Boolean value. The filter expression applies to a row and if it returns true, then that row is considered for the rule. If the filter expression returns false for that row, then it means that row is ignored for the purposes of this rule. The default behavior of the filter expression is to pass all rows, so if you don't specify a filter expression, all rows are considered.

3. **Null expression**: Checks how NULL values should be handled. This expression returns a Boolean that handles cases where data is missing. If the expression returns true, the row

expression isn't applied.

Each part of the rule works similarly to existing Microsoft Purview Data Quality conditions. A rule only passes if the row expression evaluates to TRUE for the dataset that matches the filter expression and handles missing values as specified in the null expression.

Example: A rule to ensure "fareAmount" is positive and "tripDistance" is valid:

- Row expression: tripDistance > 0 AND fareAmount > 0
- Filter expression: paymentType = 'CRD'
- Null expression: tripDistance IS NULL

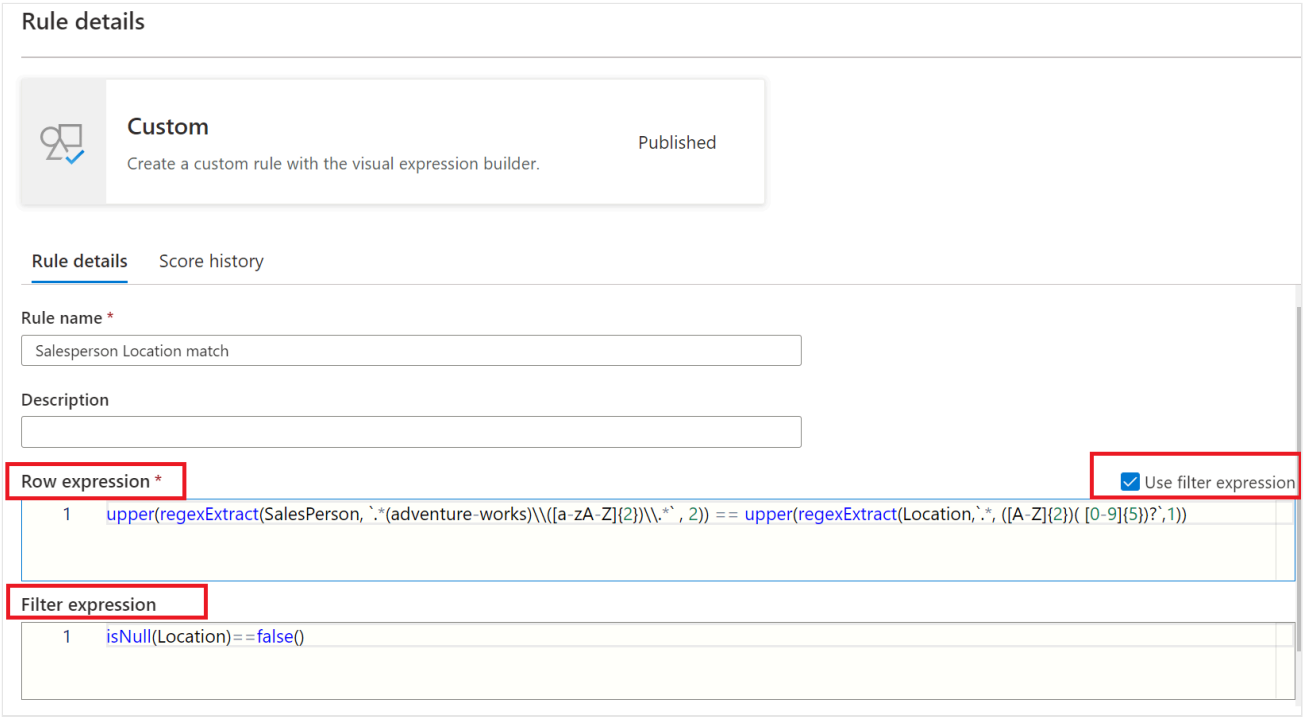
Create a custom rule

1. In Unified Catalog, go to **Health management** > **Data quality**.
2. Select a governance domain, select a data product, then select a data asset.
3. On the **Rules** tab, select **New rule**.

Create a custom rule using Azure Data Factory (ADF) expression

1. To create the rule by using regular expression or ADF expression, select **Custom** from the rule list of options, then select **Next**.

2. Add **Rule name** and **Description**, then select **Create**.



Examples of custom rules

Expand table

Scenario	Expressions
Validate if *state_id* is equal to California, and *aba_Routing_Number* matches a certain regex pattern, and date of birth falls in a certain range	`state_id=='California' && regexMatch(toString(aba_Routing_Number), '^((0[0-9])|(1[0-2])|(2[1-9])|(3[0-2])|(6[1-9])|(7[0-2])|80)([0-9]{7})$') && between(dateOfBirth,toDate('1968-12-13'),toDate('2020-12-13'))==true()`
Verify if *VendorID* is equal to 124	`{VendorID}=='124'`
Check if *fare_amount* is equal or greater than 100	`{fare_amount} >= "100"`
Validate if *fare_amount* is greater than 100 and *tolls_amount* isn't equal to 100	`{fare_amount} >= "100"||{tolls_amount} != "400"`
Check if *Rating* is less than 5	`Rating < 5`
Verify if number of digits in *year* is 4	`length(toString(year)) == 4`
Compare two columns *bbToLoanRatio* and *bankBalance* to check if their values are equal	`compare(variance(toLong(bbToLoanRatio)),variance(toLong(bankBalance)))<0`
Check if trimmed and concatenated number of characters in *firstName*, *lastName*, *LoanID*, *uuid* is greater than 20	`length(trim(concat(firstName,lastName,LoanID,uuid())))>20`
Verify if *aba_Routing_Number* matches certain regex pattern, and initial transaction date is greater than 2022-11-12, and *Disallow-Listed* is false, and average *bankBalance* is greater	`regexMatch(toString(aba_Routing_Number), '^((0[0-9])|(1[0-2])|(2[1-9])|(3[0-2])|(6[1-9])|(7[0-2])|80)([0-9]{7})$') && toDate(addDays(toTimestamp(initialTransaction, 'yyyy-MM-dd\'T\'HH:mm:ss'),15))>toDate('2022-11-12') && ({Disallow-Listed}=='false') && avg(toLong(bankBalance))>50000 && (state_id=='Massachusetts' || state_id=='Tennessee ' || state_id=='North Dakota' || state_id=='Alabama')`

Scenario	Expressions
than 50000, and *state_id* is equal to 'Massachusetts', 'Tennessee', 'North Dakota' or 'Alabama'	
Validate if *aba_Routing_Number* matches certain regex pattern, and *dateOfBirth* is between 1968-12-13 and 2020-12-13	`regexMatch(toString(aba_Routing_Number), '^((0[0-9])|(1[0-2])|(2[1-9])|(3[0-2])|(6[1-9])|(7[0-2])|80)([0-9]{7})$') && between(dateOfBirth,toDate('1968-12-13'),toDate('2020-12-13'))==true()`
Check if number of unique values in *aba_Routing_Number* is equal to 1,000,000, and number of unique values in *EMAIL_ADDR* is equal to 1,000,000	`approxDistinctCount({aba_Routing_Number})==1000000 && approxDistinctCount({EMAIL_ADDR})==1000000`

Both the filter expression and row expression are defined using the Azure Data Factory expression language, with the language defined here. However, not all the functions defined for the generic ADF expression language are available. The full list of available functions is in the Functions list available in the expression dialog. The following functions defined in here aren't supported: isDelete, isError, isIgnore, isInsert, isMatch, isUpdate, isUpsert, partitionId, cached lookup, and Window functions.

> **Note**
>
> `<regex>` (back quote) can be used in regular expressions included in custom rules to match string without escaping special characters. The regular expression language is based on Java. Learn about regular expressions and Java and understand the characters that need to be escaped.

Create a custom rule using SQL expression

Custom SQL rules in Microsoft Purview Data Quality provide a flexible way to define data quality checks using Spark SQL predicates. This feature allows users to author rules directly in Spark SQL for advanced validation scenarios. Only a row expression is required; filter and null expressions are optional for further customization. Use Custom SQL rules to address complex business

requirements and enhance data quality, using the full capabilities of Spark SQL. Custom SQL rules enable complex data validation that might not be possible with ADF expressions alone. By writing Spark SQL predicates, you can meet unique business needs and maintain high data quality standards.

1. To create the rule by using SQL expression language, select **Custom (SQL)** from the rule list of options, then select **Next**.

2. Add **Rule name** and **Description**, then select **Create**.

Expand table

Scenario	Expressions

Validates correct string patterns (for example, rateCodeId starting with '1' and numeric) and filters by valid payment types.	`Row: rateCodeId RLIKE '^1[0-9]+$'`
	`Filter: paymentType IN ('CRD', 'CSH')`
	`Null: rateCodeId IS NULL`

Ensures correct column comparisons between puLocationId and doLocationId, and fare compared to trip distance.	`Row: puLocationId > doLocationId AND fareAmount > tripDistance * 10'`
	`Filter: paymentType <> 'CSH''`
	`Null: tripDistance IS NULL`

Checks if the paymentType is in a given list (Card, Cash), filtering rows based on fare amounts.	`Row: paymentType IN ('CRD', 'CSH')'`
	`Filter: fareAmount >= 50`
	`Null: paymentType IS NULL`

Ensures distance is within an inclusive range (5-10 miles) while handling NULL and filtering for valid payment types.	`Row: tripDistance BETWEEN 5 AND 10`
	`Filter: paymentType <> 'CRD'`
	`Null: tripDistance IS NULL`

Ensures that the dataset doesn't exceed a 20% NULL value for fareAmount.	`Row: (SELECT avg(CASE WHEN fareAmount IS NULL THEN 1 ELSE 0 END) FROM nycyellowtaxidelta1BillionPartitioned) < 0.20'` `Filter: vendorID IN ('VTS', 'CMT')`
Verifies that there are at least 2 distinct paymentType values in the dataset.	`Row: (SELECT count(DISTINCT paymentType) FROM nycyellowtaxidelta1BillionPartitioned) >= 2` `Filter: vendorID IN ('1', '2')`
Ensures that the dataset's average fare amount falls within a specified range (80 <= avg <= 140).	`Row: (SELECT avg(fareAmount) FROM nycyellowtaxidelta1BillionPartitioned) BETWEEN 80 AND 140 '` `Filter: paymentType IN ('CRD', 'CSH')`

Scenario	Expressions
Ensures the maximum tripDistance in the dataset is <= 10 miles.	`Row: (SELECT max(tripDistance) FROM nycyellowtaxidelta1BillionPartitioned) <= 10.0` `Filter: vendorID IN ('VTS', 'CMT')`
Ensures that the standard deviation of fareAmount is below a certain threshold (< 30).	`Row: (SELECT stddev_samp(fareAmount) FROM nycyellowtaxidelta1BillionPartitioned) < 30.0` `Filter: vendorID IN ('VTS', 'CMT')`
Ensures that the dataset's median fare amount is within the specified threshold (<= 15).	`Row: (SELECT percentile_approx(fareAmount, 0.5) FROM nycyellowtaxidelta1BillionPartitioned) <= 15.0` `Filter: vendorID IN ('VTS', 'CMT')`
Ensures that vendorId is unique in the dataset within specific paymentType.	`Row: COUNT(1) OVER (PARTITION BY vendorID) = 1` `Filter: paymentType IN ('CRD', 'CSH','1', '2')` `Null: vendorID IS NULL`
Ensures that the combination of puLocationId and doLocationId is unique within the dataset.	`Row: COUNT(1) OVER (PARTITION BY puLocationId, doLocationId) = 1` `Filter: paymentType IN ('CRD', 'CSH')` `Null: puLocationId IS NULL OR doLocationId IS NULL`
Ensures that vendorId is unique per paymentType.	`Row: COUNT(1) OVER (PARTITION BY paymentType, vendorID) = 1 ,Filter: rateCodeId < 25, Null: vendorID IS NULL`
Ensures that the row's tpepPickupDateTime is greater than a given cutoff timestamp.	`Row: tpepPickupDateTime >= TIMESTAMP '2014-01-03 00:00:00'` `Filter: paymentType IN ('CRD', 'CSG', '1', '2')` `Null: tpepPickupDateTime IS NULL`
Each trip must be completed within 1 hour	`Row: (unix_timestamp(tpepDropoffDateTime) - unix_timestamp(tpepPickupDateTime)) <= 3600` `Filter: paymentType IN ('CRD', 'CSH', '1', '2')` `Null: tpepPickupDateTime IS NULL OR tpepDropoffDateTime IS NULL`
Keeps only highest fare trip per pickup location.	`Row: row_number() OVER (PARTITION BY puLocationId ORDER BY fareAmount DESC) = 1, Filter: paymentType IN ('CRD', 'CSH','1','2') AND tripDistance > 0, Null: fareAmount IS NULL OR puLocationId IS NULL`
All tied highest fares per pickup location pass (not just first by row_number).	`Row: rank() OVER (PARTITION BY puLocationId ORDER BY fareAmount DESC) = 1` `Filter: paymentType IN ('CRD', 'CSH','1','2') AND tripDistance > 0` `Null: fareAmount IS NULL OR puLocationId IS NULL`

Scenario	Expressions
Fare must not decrease over time for each payment type.	`Row: fareAmount >= lag(fareAmount) OVER (PARTITION BY paymentType ORDER BY tpepPickupDateTime)` `Null: tpepPickupDateTime IS NULL OR fareAmount IS NULL`

Each row's fare within 10 of group average by payment type.	`Row: abs(fareAmount - avg(fareAmount) OVER (PARTITION BY paymentType)) <= 10`
	`Filter: paymentType IN ('CRD', 'CSH','1','2')`
	`Null: fareAmount IS NULL`

Running total of trip distances must not exceed 20 miles.	`Row: sum(tripDistance) OVER (ORDER BY tpepPickupDateTime ROWS BETWEEN UNBOUNDED PRECEDING AND CURRENT ROW) <= 20`
	`Filter: paymentType = '1'`
	`Null: tripDistance IS NULL`

Checks if each trip's fare is above the global average for eligible vendors.	`Row: fareAmount > (SELECT avg(fareAmount) FROM nycyellowtaxidelta1BillionPartitioned)`
	`Filter: vendorID IN ('VTS', 'CMT')`
	`Null: fareAmount IS NULL`

Checks if each row's tripDistance is greater than the minimum for its paymentType (Card/Cash).	`Row: tripDistance > (SELECT min(u.tripDistance) FROM (SELECT tripDistance, paymentType AS pt FROM nycyellowtaxidelta1BillionPartitioned) u WHERE u.pt = paymentType)`
	`Filter: paymentType IN ('CRD', 'CSH')`
	`Null: tripDistance IS NULL`

For each trip, checks if the fare is above the average for its payment type	`Row: fareAmount > (SELECT avg(u.fareAmount) FROM (SELECT fareAmount, paymentType AS pt FROM nycyellowtaxidelta1BillionPartitioned) u WHERE u.pt = paymentType)`
	`Filter: paymentType IN ('CRD','CSH','1','2') AND vendorID IN ('VTS','CMT')`
	`Null: fareAmount IS NULL`

Validates if the fareAmount column (numeric) can be correctly represented as a string matching the numeric pattern (positive numbers with optional decimal). This uses casting as fareAmount is a numeric column.	`Row: CAST(fareAmount AS STRING) RLIKE '^[0-9]+(\.[0-9]+)?$'`
	`Filter: paymentType IN ('CRD', 'CSH')`
	`Null: fareAmount IS NULL`

Ensures that the tpepPickupDateTime is a valid timestamp in yyyy-MM-dd HH:mm:ss format. This column is already in DATETIME format	`Row: to_timestamp(tpepPickupDateTime, 'yyyy-MM-dd HH:mm:ss') IS NOT NULL`
	`Filter: paymentType IN ('CRD', 'CSH')`
	`Null: tpepPickupDateTime IS NULL`

Scenario	Expressions
Ensures that paymentType values are normalized to lowercase and have no leading or trailing spaces.	`Row: lower(trim(paymentType)) IN ('card','cash') AND length(trim(paymentType)) > 0` `Null: paymentType IS NULL OR trim(paymentType) = ''`
Safely calculates the ratio of fareAmount to tripDistance, ensuring that division by zero doesn't occur by first checking if tripDistance > 0.	`Row: CASE WHEN tripDistance > 0 THEN fareAmount / tripDistance ELSE NULL END >= 10` `Filter: tripDistance > 0 AND vendorID IN ('VTS', 'CMT')`
Demonstrates how coalesce can replace null values with default values (for example, 0.0), and ensures only valid rows are returned.	`Row: coalesce(fareAmount, 0.0) >= 5` `Filter: paymentType IN ('CRD','CSH')`

Best practices for writing custom SQL rules

- Keep expressions simple. Aim to write clear, straightforward expressions that are easy to maintain.
- Use built-in Spark SQL functions. Use Spark SQL's rich library of functions for string manipulation, date handling, and numeric operations to minimize errors and improve performance.
- Test with a small dataset first. Validate rules on a small dataset before applying them at scale to identify potential issues early.

Known limitations and considerations for SQL expressions rules

Ambiguous column references and column shadowing

- Problem: When a column appears in both the outer query and subquery (or across different parts of the query) with the same name, Spark SQL might not be able to resolve which column to use. This issue results in logical errors or incorrect query execution. This problem can arise in nested queries, subqueries, or joins, leading to ambiguity or shadowing.

- Ambiguity: Occurs when a column name is present in both the outer query and the subquery without clear qualification, causing Spark SQL to be uncertain about which column to reference.

- Shadowing: Refers to when a column in the outer query is "overridden" or "shadowed" by the same column in the subquery, causing the outer reference to be ignored.

  Example Expression:

SQL

```
distance_km >
(
SELECT min(distance_km)
FROM Tripdata t
WHERE t.payment_type = payment_type -- ambiguous outer reference
)
```

- Issue: Unqualified payment_type is resolved to the nearest scope that has a column of that name, that is, the inner t.payment_type - not the outer row's payment_type. That turns the predicate into t.payment_type = t.payment_type (always TRUE), so your subquery becomes a global min instead of a group min.
- Solution: To resolve this ambiguity and avoid column shadowing, rename the inner column in the subquery, ensuring that the outer query's payment_type remains unambiguous.

Corrected Expression:

SQL

```
distance_km >
(
SELECT min(u.distance_km)
FROM (
SELECT distance_km, payment_type AS pt
FROM Tripdata
) u
 WHERE u.pt = payment_type   -- this `payment_type` now binds to OUTER row
)
```

- In the subquery, the column payment_type is aliased as pt (that is, payment_type AS pt), and u.pt is used in the condition.
- In the outer query, the original payment_type can now be clearly referenced, and Spark SQL correctly resolves it as the outer payment_type.

Window operations (performance consideration)

- Window operations such as ROW_NUMBER() and RANK() can be expensive, especially for large datasets. Use them judiciously and test performance on smaller datasets before applying them at scale. Consider using PARTITION BY to reduce the data scope.

Column name escaping in Spark SQL

- If column names contain special characters (such as spaces, hyphens, or other non-alphanumeric characters), they must be escaped using backticks.
- Example if the column name is order-id and the rule needs to be that it should be greater than 10.
- Incorrect expression: order-id > 10
- Correct expression: `` `order-id` `` > 10

Data asset name referencing in expressions

When referencing your data asset in SQL expressions, you need to follow specific sanitization rules. The original data asset name doesn't need to be updated, but the data asset name referenced in SQL expressions must be sanitized to meet the following criteria:

Expand table

Rule	Description	Example- Original Name	Example- Sanitized Name
Allowable Characters	Only letters (A-Z, a-z), numbers (0-9), and underscores (_) are allowed. Special characters (spaces, hyphens, periods, etc.) must be removed.	my-dataset_v1+2023	mydataset_v12023
Trim Underscores	Underscores at the beginning or end of the name must be removed.	my_dataset_	my_dataset
Character Limit	The final, sanitized name must not exceed 64 characters.	[A long name exceeding 64 characters]	[The first 64 characters of the sanitized name]

If your data asset name already follows these guidelines (that is, it doesn't contain any special characters, leading/trailing underscores, and is within the 64-character limit), it can be used as-is in your SQL expressions without any modification.

How to sanitize a dataset name

Follow these steps to ensure your dataset name is valid for SQL expressions:

1. Remove special characters: Strip out all characters except letters, numbers, and underscores.
2. Trim underscores: Remove any leading or trailing underscores.
3. Truncate: If the resulting name exceeds 64 characters, truncate it to fit within the 64-character limit.

Example: Data asset name f07d724d-82c9-4c75-97c4-c5baf2cd12a4.parquet

1. Remove special characters: f07d724d82c94c7597c4c5baf2cd12a4parquet
2. Trim underscores: (N/A in this case, as there are no leading or trailing underscores.)
3. Truncate: The resulting name is 54 characters long, which is under the 64-character limit.

Final SQL reference name: f07d724d82c94c7597c4c5baf2cd12a4parquet

> ⓘ **Note**
>
> The original data asset name remains unchanged. Only the data asset name used in the SQL expressions needs to follow these rules. For column names that contain special characters, such as spaces or hyphens, you can escape them using back quotes in SQL expressions.

Joins aren't supported

Custom SQL rules in Microsoft Purview Data Quality don't support joins. The rules must operate on a single dataset. You can't join multiple tables or datasets when writing these custom rules.

Unsupported SQL operations (DML, DCL, and harmful SQL)

Custom SQL rules don't support Data Manipulation Language (DML) or Data Control Language (DCL) operations such as INSERT, UPDATE, DELETE, GRANT, and other harmful SQL operations like TRUNCATE, DROP, and ALTER. These operations aren't supported because they modify the data or the database's state.

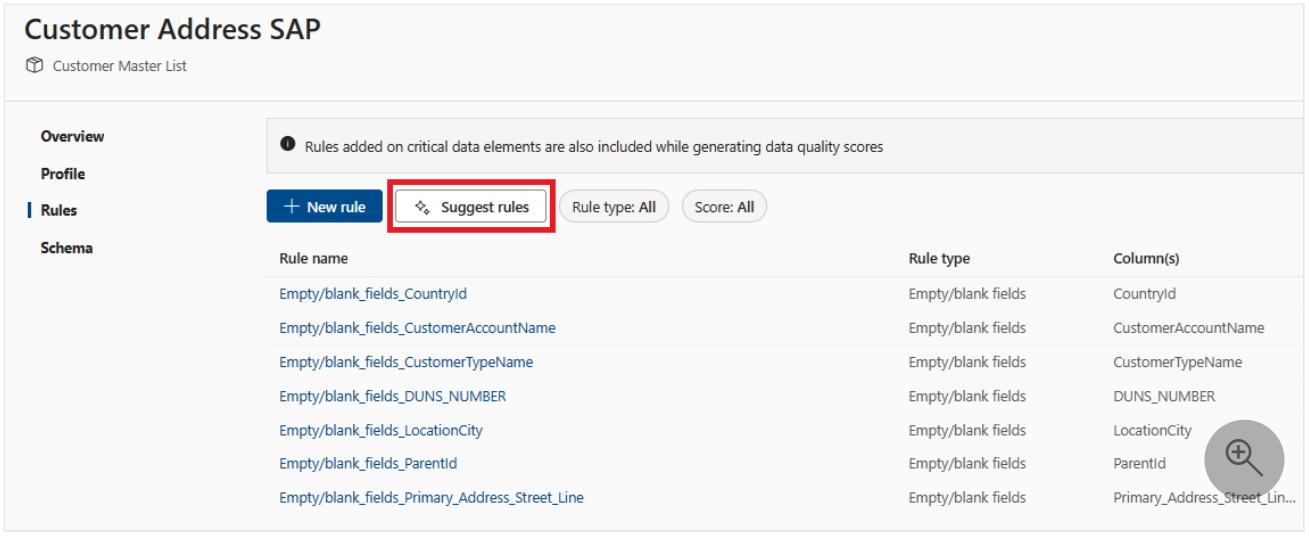
AI-assisted autogenerated rules

AI-assisted automated rule generation for data quality measurement uses artificial intelligence (AI) techniques to automatically create rules for assessing and improving the quality of data. Autogenerated rules are content specific. Most of the common rules are generated automatically so that you don't need to put much effort into creating custom rules.

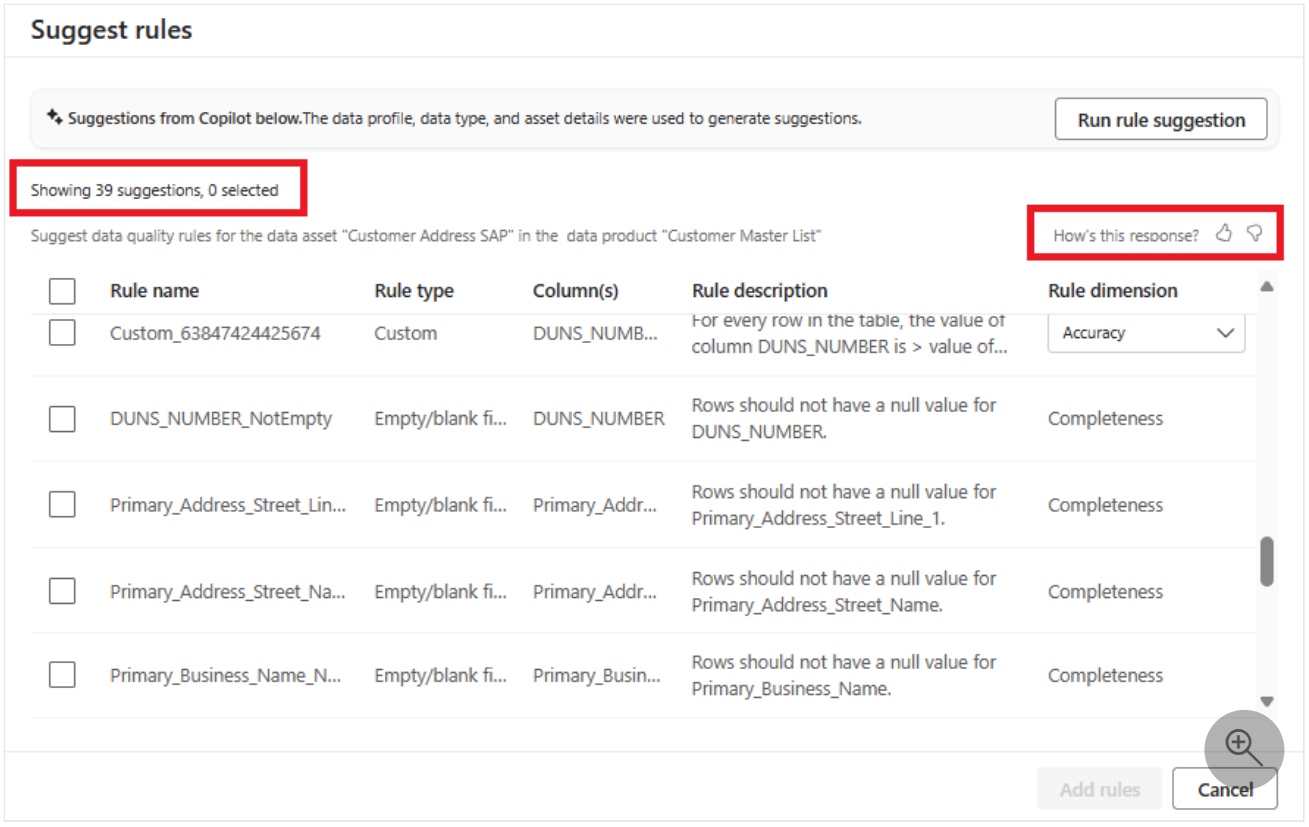
To browse and apply autogenerated rules:

1. On the **Rules** tab of a data asset, select **Suggest rules**.

2. Browse the list of suggested rules.



3. Select rules from the suggested rule list to apply to the data asset.



Limitation

You can apply a maximum of 200 data quality rules per data asset for a data quality scan. The scan fails if more than 200 rules are active for a single asset. If you need to define more than 200 rules for a data asset, they can do so, but only up to 200 rules can be in an active (ON) state at any given time.

To work around this limit:

- Add the same data asset to multiple data products to distribute rules across them.

- Activate and deactivate rules as needed to ensure no more than 200 rules are active at once.

Next steps

1. Configure and run a data quality scan on a data product to assess the quality of all supported assets in the data product.
2. Review your scan results to evaluate your data product's current data quality.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 07/17/2026

Data quality thresholds

A data quality threshold is a predefined limit that determines whether data meets an acceptable level of quality. In simple terms, it answers the question: How good is good enough?

For example, if you run a completeness rule on a column:

- Total records: 1,000
- Records with values: 920
- Completeness score: 92%

If your threshold is 90%, the data meets the expectation.

If your threshold is 95%, the data doesn't meet the expectation.

The threshold is the cutoff point that defines whether the data meets the expectation for specific use cases.

Why thresholds matter

Different types of data require different quality expectations. For example:

- Email address column: Might require 99–100% completeness.
- Description column: Perhaps 80–90% is acceptable.
- Financial transactions: Might require 100% accuracy.

Using one fixed threshold, such as 80% for everything, can be misleading, especially for critical or regulated data.

Types of thresholds

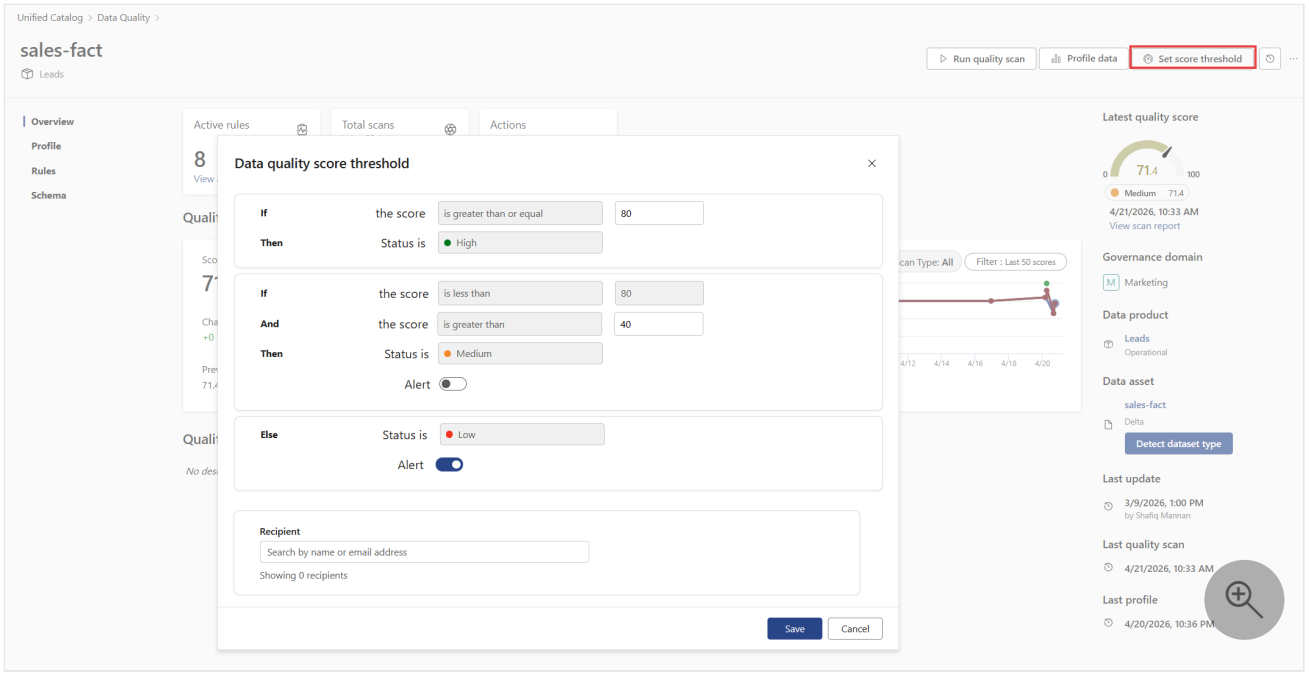
- Rule-level threshold: Set for each rule (for example, completeness must be ≥ 95%).
- Data asset level threshold: Different expectations for different columns.
- Data product level threshold: Overall score requirement.

Configurable data quality thresholds

You can configure Data Quality thresholds to set minimum acceptable quality scores at the rule and data asset levels. Define the threshold to align quality evaluation with business criticality.

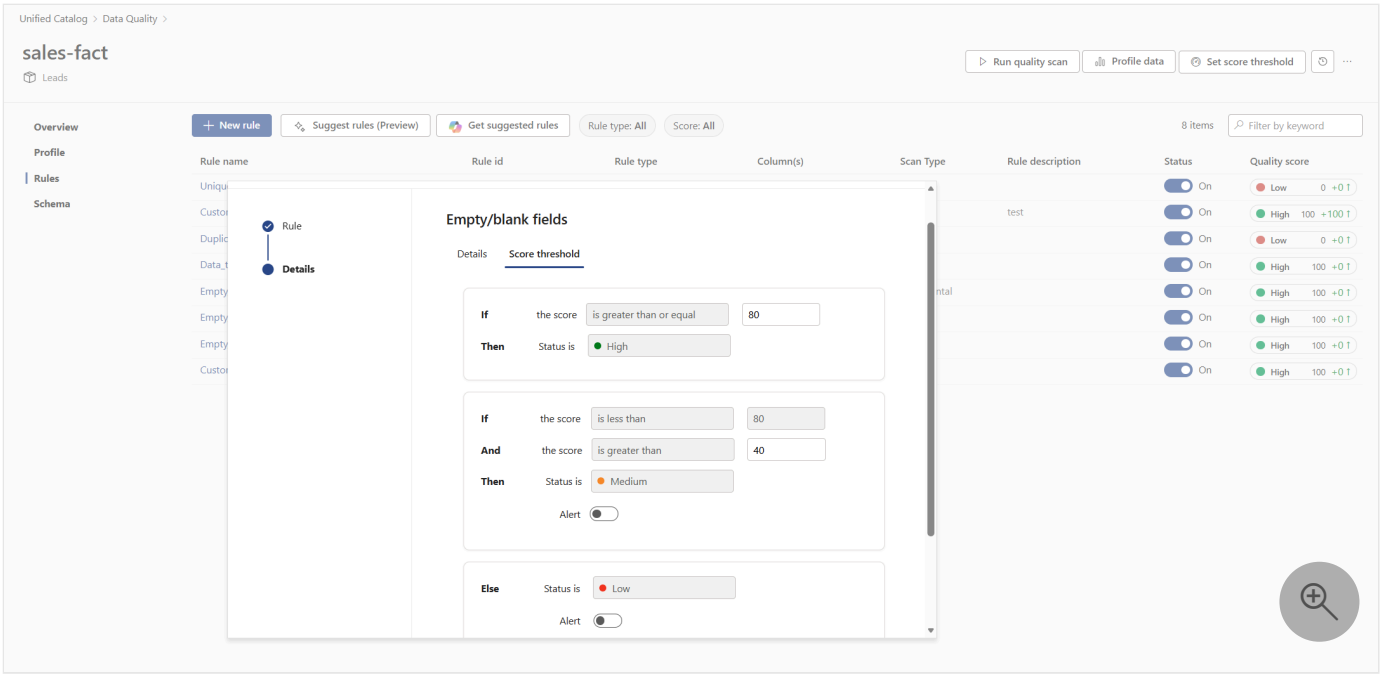
Configure or edit the default threshold for a data asset

1. In Microsoft Purview Unified Catalog, select **Health management**, and then select **Data quality**.
2. Select a governance domain, select a data product, and then select a data asset.
3. On the **Overview** page, select **Set score threshold**.
4. Change the default threshold.
5. Select **Save**.



Configure the threshold for a data quality rule

1. In Unified Catalog, select **Health management**, and then select **Data quality**.
2. Select a governance domain, select a data product, and then select a data asset.
3. On the **Rules** page, select **New rule**.
4. Select the out-of-the-box rule or the custom rule, and select **Next**.
5. Select **Score threshold** and change the default threshold value.
6. Select **Save**.



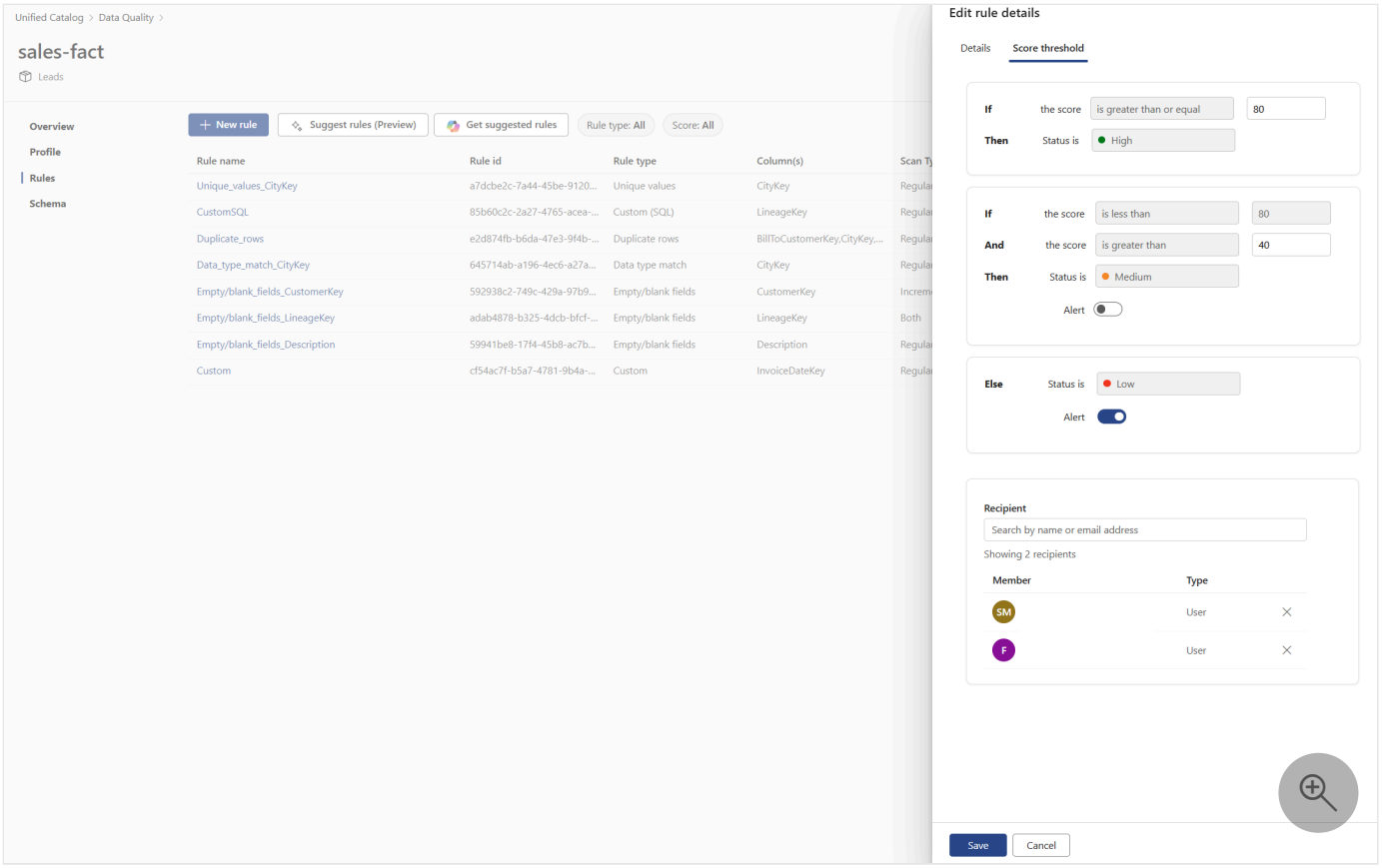
Edit a data quality rule threshold

1. On the **Rules** page of a data asset, select the edit icon for a rule.
2. From the overview, select **Score threshold**.
3. Edit the threshold value.
4. Select **Save**.

Configure alerts

Set an alert to notify you when the data quality score of a data asset or rule falls below a defined threshold. When you set up or edit a rule threshold, follow these steps:

- Set the **Alert** toggle to on.
- At **Recipient**, add one or more users.
- Select **Save**.



> **Note**
>
> If you don't configure a threshold, the default threshold values and colors are used. The default threshold score values are:
>
> - **Low (red)**: 0-40
> - **Medium (orange)**: 40-80
> - **High (green)**: 80-100

> **Important**
>
> Threshold metadata doesn't publish to Azure Data Lake Storage Gen2 or Fabric Lakehouse for self-serve analytics.

The benefits that an organization gets from the configurable data quality threshold feature are:

- Enables the enterprise to align their organization with quality governance.
- Improves trust in the data quality scores produced by Microsoft Purview Data Quality, enabling teams across the enterprise to confidently determine whether data is fit for their specific use cases.

- Helps organizations move beyond generic quality checks and toward business-driven data quality management.
- Supports regulatory and decision-critical use cases.
- Reduces user or organizational confusion caused by uniform threshold evaluation.

Setting rule-level thresholds creates a clear standard for what "good" looks like. Alerts ensure you act the moment data drifts below that standard. This has several practical benefits:

- **Early detection of issues**: Catch things like missing, inconsistent, or incorrect data before they impact reports, pipelines, or models.
- **Reduced business risk**: Prevents bad data from driving incorrect decisions, financial errors, or compliance issues.
- **Faster remediation**: Alerts enable teams to respond immediately instead of discovering issues days or weeks later.
- **Accountability and governance**: Clear thresholds define ownership and expectations for data quality.
- **Operational efficiency**: Eliminates manual monitoring and reduces reliance on ad hoc checks.
- **Trust in data**: Consistent enforcement of quality standards increases confidence in analytics and AI outputs.

In short, **thresholds** define acceptable quality, and alerts make that standard actionable.

Next steps

- Data quality scan

ⓘ **Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 05/10/2026

Configure a data quality scan for a data product

Data quality scans review your data assets based on their applied data quality rules and produce a score. Your data stewards can use that score to assess the data health and address any issues that might be lowering the quality of your data.

> **Note**
>
> There are two other types of data quality scans you can run:
>
> - Data quality scan for a standalone data asset (preview).
> - Incremental data quality scan (preview) using a time-based filter.

Prerequisites

- To run and schedule data quality assessment scans, users need the data quality steward role.
- Currently, you can set the Microsoft Purview account to allow public access or managed virtual network access so that data quality scans can run.

Data quality life cycle

Data quality scanning is the **seventh** step in the data quality life cycle for a data asset. The previous steps are:

1. Assign users data quality steward permissions in Microsoft Purview Unified Catalog so they can use all data quality features.
2. Register and scan a data source in Microsoft Purview Data Map.
3. Add your data asset to a data product.
4. Set up a data source connection to prepare your source for data quality assessment.
5. Configure and run data profiling for an asset in your data source. When profiling is complete, browse the results for each column in the data asset to understand your data's current structure and state.
6. Set up data quality rules based on the profiling results, and apply them to your data asset.

Supported multicloud data sources

Browse the supported data source document to view the list of supported data sources, including file formats for data profiling and data quality scanning, with and without virtual network support.

> **Important**
>
> Data quality for Parquet file is designed to support:
>
> 1. A directory with Parquet Part File. For example: **./Sales/{Parquet Part Files}**. The Fully Qualified Name must follow `https://(storage account).dfs.core.windows.net/(container)/path/path2/{SparkPartitions}`. Make sure there are no {n} patterns in directory/sub-directory structure. It must be a direct FQN leading to {SparkPartitions}.
> 2. A directory with Partitioned Parquet Files, partitioned by Columns within the dataset like sales data partitioned by year and month. For example: **./Sales/{Year=2018}/{Month=Dec}/{Parquet Part Files}.**
>
> Both of these essential scenarios, which present a consistent parquet dataset schema, are supported. **Limitation:** It isn't designed to or won't support N arbitrary Hierarchies of Directories with Parquet Files. **We recommend presenting data in (1) or (2) constructed structure.**

Supported authentication methods

Currently, Microsoft Purview can only run data quality scans by using Managed Identity as authentication option. Data quality services run on **Apache Spark 3.5** and **Delta Lake 3.2.1**. For more information about supported regions, see data quality overview.

> **Important**
>
> - If you update the schema on the data source, you need to import schema from data quality overview page by using **schema import feature** before running a data quality scan.
> - Virtual network isn't supported for Google BigQuery.

Run a data quality scan

1. Configure a data source connection to the assets you're scanning for data quality if you haven't already done so.
2. In Unified Catalog, select **Health Management**, then select **Data quality**.
3. Select a **governance domain** from the list.
4. Select a **data product** to assess the data quality of the **data assets** linked to that product.
5. Select the name of a data asset, which takes you to the data quality **Overview** page.
6. Browse the existing data quality rules and add new rules by selecting **Rules**. Browse the schema of the data asset by selecting **Schema**. Toggle on or off the rules you added.
7. Run the quality scan by selecting **Run quality scan** on the overview page.
8. While the scan is running, you can track its progress from the data quality monitoring page in the governance domain.

Schedule data quality scans

Although you can run data quality scans on an ad-hoc basis by selecting **Run quality scan**, in production scenarios the source data is likely to be constantly updated. You should regularly monitor data quality to detect any issues. Automating the scanning process helps you manage regular updates of quality scans.

1. In Unified Catalog, select **Health Management**, then select **Data quality**.
2. Select a **governance domain** from the list.
3. Select **Manage**, then select **Scheduled scans**.
4. Fill out the form on the **Create scheduled scan** page. Add a **name** and **description** for the source you're setting up the schedule.
5. Select **Continue**.
6. On the **Scope** tab, select individual data product and assets or all data products and data assets of the entire governance domain.
7. Select **Continue**.
8. Set a schedule based on your preferences and select **Continue**.

9. On the **Review** tab, select **Save** (or **Save and run** to test immediately) to complete scheduling the data quality assessment scan.

You can monitor scheduled scans on the data quality job monitoring page under the **Scans** tab.

> ⓘ Note
>
> - You can't add more than 30 assets across all data products in a single schedule. Create multiple schedules for 30 assets per batch. You can configure to run multiple schedules in the same time window.
> - Auto-scheduling to run a weekly scan is enabled at the asset level by default. If you want to disable auto-scheduling for a data asset, go to **Health management** > **Data quality** > **Run quality scan**, and then select the option to disable the scheduled scan.
> - When you manually run a data quality scan, a notification appears on the Data Quality page prompting you to configure alerts. Ignore this notification if you don't want to configure alerts.

Delete previous data quality scans and history

When you remove a data asset from a data product, if that data asset has a data quality score, you need to delete the data quality score first, then remove the data asset from the data product.

When you delete data quality history data, it removes the profile history, the data quality scan history, and data quality rules, but data quality actions aren't deleted.

Follow the steps below to delete previous data quality scans of a data asset:

1. In Unified Catalog, select **Health Management**, then select **Data quality**.
2. Select a **governance domain** from the list.
3. Select the **data product** from the list.
4. Select the **data asset** from the list to navigate the Data quality overview page.
5. Select the ellipsis (...) at the top right of the Data quality overview page.
6. Select **Delete data quality data** to delete the history of data quality runs.

> ⓘ Note
>
> - Use **Delete data quality data** for test runs, errored data quality runs, or if you're removing a data asset from a data product.

- The system stores up to 50 snapshots of data quality profiling and data quality assessment history. If you want to delete a specific snapshot, select the desired history run and select the delete icon.

Import schema

If the data type in a schema is undefined, incorrectly defined, or changed in the source, your data quality job might fail. If it fails, reimport the schema by using the schema import capability. Schema import is supported for data sources on both public networks and behind private endpoints. The supported data sources are listed at Data sources and file formats supported for data quality. To import a schema from your data sources, follow these steps:

- Select **Data quality** from **Health Management**.
- Select a business domain, then select a data product, then select a data asset from that data product. You arrive at the data quality overview page.
- Select **Schema**, then select the **Schema management** toggle.
- Select **Import schema** to import the schema.

Related content

- Data quality for Fabric data estate
- Data quality for Fabric Mirrored data sources
- Data quality for Fabric shortcut data sources
- Data quality for Azure Synapse serverless and data warehouses
- Data quality for Azure Databricks Unity Catalog
- Data quality for Snowflake data sources
- Data quality for Google BigQuery

Next steps

- Monitor data quality scan
- Review your scan results to evaluate your data product's current data quality.
- Configure alerts for data quality scan results

Last updated on 05/15/2026

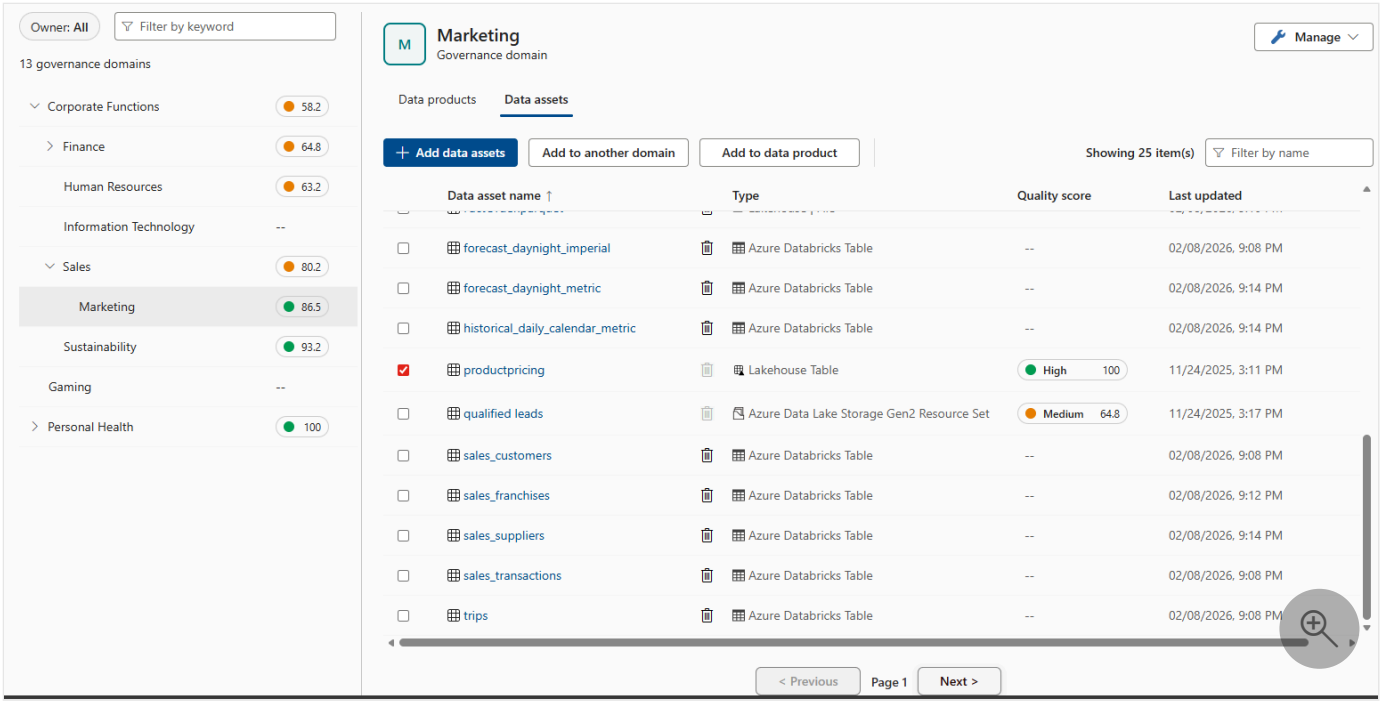
Configure a data quality scan for a data asset

A standalone data asset is a data asset that isn't associated with a data product in Microsoft Purview Unified Catalog. When you manage data quality for standalone data assets, you can measure, monitor, and improve the quality of these assets independently, without requiring a data product association.

Benefits of running data quality on a data asset

- **Helps you assess and improve the quality of your data before linking it to a data product**. Microsoft Purview Data Quality users can profile and evaluate the quality of their data before associating it with a data product. Users can standardize, clean, and resolve issues upfront.
- **Helps you decide which assets should be part of a data product**. By understanding the quality of standalone assets, organizations can make informed decisions about which assets are suitable for curation and governance within a data product.
- **Supports broader use cases beyond analytics.** Many organizations associate data products only with analytics use cases. Standalone data assets might serve other purposes such as:
  - Data monetization
  - AI grounding data
  - Operational workloads
- **Provides unified data quality tooling across all data**. Organizations can use a single Microsoft Purview Data Quality solution to assess both standalone data assets and data product–associated assets, including issue remediation.
- **Helps optimize data storage and reduce costs**. If standalone assets are low quality, incomplete, or unusable, organizations can archive them to lower-cost storage or apply data minimalism principles, which reduces unnecessary data retention and improving return on investment.
- **Accelerates data governance maturity**. Organizations can begin measuring and improving data quality immediately, without waiting for data product definitions or use cases. This significantly speeds up governance adoption.

Data quality scans review your data assets based on their applied data quality rules and produce a score. Your data stewards can use that score to assess the data health and address any issues that might be lowering the quality of your data. Here's an example of how data asset scores appear in a governance domain:



Prerequisites

- To run and schedule data quality assessment scans, you need the data quality steward role.
- Currently, you can set the Microsoft Purview account to allow public access or managed virtual network access so that data quality scans can run.
- Configure the data source connection if it isn't created yet for the data asset's source system.
- Configure the data quality error records storage if you want to store data quality failed records for review and correction.

Data quality life cycle for data assets

Data quality scanning is the **seventh** step in the data quality life cycle for a data asset. The previous steps are:

1. Register and scan a data source in Microsoft Purview Data Map.
2. Assign users data quality steward permissions in Unified Catalog so they can use all data quality features.
3. Set up a data source connection to prepare your source for data quality assessment.
4. Add your data asset from Data Map to a governance domain.

5. Configure a storage to store data quality error records if you want to store data quality error record for review and remediation.
6. Configure and run data profiling for an asset in your data source.. When profiling is complete, browse the results for each column in the data asset to understand your data's current structure and state to identify what rules need to be created to measure and monitor data quality of your data continuously.
7. Set up data quality rules based on the profiling results, and apply them to your data asset.
8. Review data quality error records and correct those failed records to improve data quality of your data.

Supported multicloud data sources

To see the list of supported data sources, including file formats for data profiling and data quality scanning with and without virtual network support, see the supported data sources.

> **Important**
>
> Data quality for Parquet file supports:
>
> 1. A directory with Parquet Part File. For example: **./Sales/{Parquet Part Files}**. The fully qualified name (FQN) must follow `https://(storage account).dfs.core.windows.net/(container)/path/path2/{SparkPartitions}`. Make sure there are no {n} patterns in directory or subdirectory structure. It must be a direct FQN leading to {SparkPartitions}.
> 2. A directory with Partitioned Parquet Files, partitioned by columns within the dataset like sales data partitioned by year and month. For example: **./Sales/{Year=2018}/{Month=Dec}/{Parquet Part Files}.**
>
> Both of these essential scenarios, which present a consistent parquet dataset schema, are supported. **Limitation:** The feature doesn't support N arbitrary hierarchies of directories with Parquet files. **Present data in (1) or (2) constructed structure.**

Supported authentication methods

Currently, Microsoft Purview can only run data quality scans by using Managed Identity as authentication option. Data quality services run on **Apache Spark 3.4** and **Delta Lake 2.4**. For more information about supported regions, see data quality overview.

> **Important**
>
> - If you update the schema on the data source, import the schema from the Data Quality overview page by using **schema import feature** before running a data quality scan.
> - Virtual network isn't supported for Google BigQuery.

Run a data quality scan for standalone data asset

1. In Unified Catalog, select **Health Management**, and then select **Data quality**.
2. Select a governance domain from the list.
3. Configure a data source connection to the assets you're scanning for data quality if you haven't already done so.
4. Select a **Data asset** and select **Add data assets** to add data assets from Data Map for a data quality assessment. Select the data assets you want from the Data Map assets list to add them to the governance domain to begin measuring their data quality.
5. Select the name of a data asset, which takes you to the data quality **Overview** page.
6. Profile the selected data asset if you want to know the distribution, min, max, uniqueness, completeness, standard deviation, and other statistical measures of your data assets.
7. To set up the scan, browse the existing data quality rules and add new rules by selecting **Rules**. Browse the schema of the data asset by selecting **Schema**. Toggle on or off the rules you added. If you toggle off a rule, that rule isn't included in the data quality scan.
8. Run the quality scan by selecting **Run quality scan** on the overview page.
9. While the scan is running, you can track its progress from the data quality monitoring page in the governance domain.
10. If you configured data quality error records in the governance domain level or data asset level, check the error records in the configured storage.

Add a data asset with a quality score to another governance domain

After a data asset has a data quality score, you can add it to another governance domain by cloning it. You can clone and associate the same data asset with any governance domain as

needed.

- Select a data asset with data quality score.
- Select **Add to another domain**, and then select the domain from the drop-down list.
- Select the domain to add the selected data asset.

Associate asset with a quality score to a data product

You can associate a data asset with its latest data quality score to a data product in a different governance domain.

- Select a data asset with data quality score.
- Select **Add to data product**, and then select the **Governance domain** and **Data product** from the drop-down list.
- Select the domain and data product to add the selected data asset.

When the same data asset is associated with multiple data products, the meaning and expectations of data quality for that asset can differ depending on the business context.

For example, a customer profile data asset might be associated with a Marketing data product for campaign use cases and a Finance data product for billing scenarios. In this case, billing address quality is critical for Finance but less important for Marketing, while customer contact email quality is critical for Marketing but less relevant for Finance.

Because data quality expectations vary by data product, you need to run data quality scans locally on the cloned or associated asset within each target data product to ensure that data quality is evaluated based on the specific use case.

> ⓘ **Note**
>
> - When you clone or associate a data asset with an existing data quality score from the Data Quality > Data Asset page to another domain or data product, you must run the data quality scan locally in the target domain or data product where the asset is cloned or associated.
> - For example, if you clone a data asset named Budget with an existing data quality score from the Marketing domain to the Sales domain, you will need to run a data quality scan for the Budget asset within the Sales domain to view the updated data quality score in that domain.

- A similar process must be followed for any asset cloned or associated with a data product.

Important

- Threshold and alert metadata don't publish to Azure Data Lake Storage Gen2 or Fabric Lakehouse for self-serve analytics.
- Standalone data assets and associated metadata, for example data quality score and data quality rules, don't publish to Azure Data Lake Storage Gen2 or Fabric Lakehouse for self-serve analytics. If a standalone data asset is associated with a data product, then the associated metadata is available for self-serve analytics with the data product association.

Roles and permissions

- You can't associate a standalone data asset with a data quality score to a data product that is in a published state. The data product must be in draft state.
- If you're not the data product owner, you can't add a standalone data asset to a data product.
- If you don't have permissions for a business domain, you can't clone a standalone data asset into that domain.
- A data product owner must have a local or global Domain Reader role or a Data Quality Reader role to browse the standalone data asset data quality page and to associate a data asset (with data quality scores and rules) to a data product.
- A data steward must have a local or global Domain Owner role to browse the standalone data asset data quality page.

Note

If you delete a data asset from the Data Map, it appears in read-only mode and you identify it by its GUID instead of the asset name. When you hover over the asset GUID, a tooltip message is displayed. You can't run data quality scans on deleted assets, so you can remove the asset from the list.

Schedule data quality scans

Although you can run data quality scans on an ad-hoc basis by selecting **Run quality scan**, in production scenarios, the source data is likely to be constantly updated. You should regularly monitor data quality to detect any issues. Automating the scanning process helps you manage regular updates of quality scans.

1. In Unified Catalog, select **Health Management**, and then select **Data quality**.
2. Select a **governance domain** from the list.
3. Select **Manage**, then select **Scheduled scans**.
4. Fill out the form on the **Create scheduled scan** page. Add a **name** and **description** for the source you're setting up the schedule.
5. Select **Continue**.
6. On the **Scope** tab, select data assets you want to set up schedule to run data quality as per configured schedule.
7. Select **Continue**.
8. Set a schedule based on your preferences and select **Continue**.
9. On the **Review** tab, select **Save** (or **Save and run** to test immediately) to complete scheduling the data quality assessment scan.

You can monitor scheduled scans on the data quality job monitoring page under the **Scans** tab.

> **Note**
>
> - You can't add more than 30 assets across all data products in a single schedule. Create multiple schedules for 30 assets per batch. You can configure to run multiple schedules in the same time window.
> - Auto-scheduling to run a weekly scan is enabled at the asset level by default. If you want to disable auto-scheduling for a data asset, go to **Health management** > **Data quality** > **Run quality scan**, and then select the option to disable the scheduled scan.
> - When you manually run a data quality scan, a notification appears on the Data Quality page prompting you to configure alerts. Ignore this notification if you don't want to configure alerts.

Delete previous data quality scans and history

To remove a data asset from the data asset list on the data quality page, first delete the data quality score if the data asset has one. Then, remove the data asset from the data quality data asset list.

When you delete data quality history data, you remove the profile history, the data quality scan history, and data quality rules. Data quality actions aren't deleted.

To delete previous data quality scans for a data asset, follow these steps:

1. In Unified Catalog, select **Health Management**, and then select **Data quality**.
2. Select a **governance domain** from the list.
3. Select the **data asset** from the list to go to the Data quality overview page.
4. Select the ellipsis (...) at the upper right of the Data quality overview page.
5. Select **Delete data quality data** to delete the history of data quality runs.

> ⓘ **Note**
>
> - Use **Delete data quality data** for test runs, errored data quality runs, or if you're removing a data asset from the data quality data asset list.
> - The system stores up to 50 snapshots of data quality profiling and data quality assessment history. To delete a specific snapshot, select the desired history run and select the delete icon.

Import schema

If the data type in a schema is undefined, incorrectly defined, or changed in the source, your data quality job might fail. If it fails, reimport the schema by using the schema import capability. You can import schemas for data sources on both public networks and behind private endpoints. For a list of supported data sources, see Data sources and file formats supported for data quality. To import a schema from your data sources, follow these steps:

- Select **Data quality** from **Health Management**.
- Select a governance domain, and then select a data asset from the list to go to the Data Quality overview page.
- Select **Schema**, and then select the **Schema management** toggle.
- Select **Import schema** to import the schema.

Related content

- Data quality for Fabric data estate
- Data quality for Fabric Mirrored data sources
- Data quality for Fabric shortcut data sources
- Data quality for Azure Synapse serverless and data warehouses
- Data quality for Azure Databricks Unity Catalog
- Data quality for Snowflake data sources
- Data quality for Google BigQuery

Next steps

- Monitor data quality scan.
- Review your scan results to evaluate your data product's current data quality.
- Configure alerts for data quality scan results.

ⓘ **Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 05/10/2026

Data quality for on-premises data sources (preview)

On-premises data source data quality refers to the ability to assess, monitor, score, and improve the quality of data that resides inside enterprise-owned infrastructure such as on-premises SQL Server and on-premises Oracle. These datasets are stored within a customer's private network environment rather than in cloud platforms.

Prerequisites

On-premises databases are scanned in on-premises infrastructure to ensure that data doesn't move beyond the customer premises. You need to set up a Kubernetes cluster to host the runtime that scans these databases. Before you use Microsoft Purview Data Quality for on-premises data sources:

1. Create a Kubernetes cluster.
2. Download the data integration runtime.
3. Register it with Microsoft Purview.

For prerequisites, installation, and set up, see:

- Set up DIRCTL tool for self-hosted integration runtime (preview)
- Kubernetes supported self-hosted data integration runtime for on-premises data sources (preview)

Before you run a data quality scan, you must scan the assets through a Data Map scan. The Data Map scan requires you to install and configure a self-hosted integration runtime.

Set up a connection to your on-premises databases

Create connections for on-premises data sources the same way you create connections for other data source types. In Unified Catalog, select **Health management**, then select **Data quality**. Select the governance domain where you created your data product with the on-premises data asset. Select **Manage** and select **Connection** to set up a connection for your on-premises databases. Supported on-premises data sources are:

- Oracle
- SQL Server

Create connections by associating them with a data integration runtime instance. In SQL Server, a single database can contain tables belonging to multiple schemas. So, you can use a single connection to scan all schemas in a single database. Therefore, a connection accepts only database information but not schema. For an Oracle database, you can create the connection at the schema level or the database level. If you want to restrict the connection to the schema level, add the schema name in the connection configuration page.

> **Note**
>
> - One Oracle database should have only one database level connection, but it can have many schema level connections for schema 1, schema 2, and schema 3.
> - Each schema level connection is used as a separate connection because they're separate schemas.
> - If an Oracle database has both schema level and database level connections, the system uses only the schema level connection.
> - If an Oracle database has two database level connections, the system uses only the first connection.

Data quality scan

After you complete the connection setup, follow the data quality scanning documentation to measure and monitor data quality of Oracle and SQL Server on-premises data sources.

- Data quality overview
- Supported data sources
- Set up data source connection
- Data product quality scan
- Data asset quality scan
- Incremental data quality scan

Limitations

- Profiling on-premises data sources isn't supported.
- Table lookup rules aren't supported if you don't publish the tables as data assets, or if the databases of the reference tables aren't running or hosting on the same server.
- Freshness rule isn't supported for SQL Server database.

- Custom expressions that are supported for cloud sources aren't supported for SQL Server database, but Custom SQL is supported.
- Regex rules aren't supported, but LIKE is supported.

Resolve common errors

- **Error**: *Value cannot be null. (Parameter 'Password property is missing')*. This error occurs when the key vault secret isn't found. The actual error message in the Kubernetes log is: "A secret with `(name/id)` KVSecretsOfficers-purview was not found in this key vault."

Last updated on 04/27/2026

Configure an incremental data quality scan

Microsoft Purview Data Quality supports incremental data quality scans by using time-based filtering. By using incremental scans, you can choose between full scans, incremental scans, or both when you run data quality rules on data assets.

While certain rules, such as duplicate detection, require a full scan of the entire dataset, you can efficiently evaluate many data quality checks on newly created or updated data. By using incremental scans, you can run these rules on a defined schedule, such as daily or weekly, which improves performance and reduces scan costs.

By using incremental data quality scans, you can:

- Create rules that run on full scans, incremental scans, or both.
- Flexibly assign existing rules to incremental or full scan execution.
- Configure schedules for data assets to run incremental scans at defined intervals.

Data quality scores are calculated across all rules and data assets, regardless of whether the scan is incremental or full. Microsoft Purview Data Quality computes a cumulative data quality score to help you understand overall data quality trends over time.

Benefits of running incremental scans

Data offices and data practitioners don't always need to assess data quality across the entire historical dataset, such as five to 10 years of data, on a frequent basis. You can evaluate historical data periodically, such as monthly or quarterly, while newly created or updated data often requires more frequent monitoring.

By using incremental scans, you can enable continuous data quality monitoring on recent data so that you can identify and address problems quickly, before they affect downstream systems and analytics.

Configure and run a data quality scan

Data quality scans review your data assets based on their applied data quality rules and produce a score. Your data stewards can use that score to assess the data health and address any problems that might lower the quality of your data.

Prerequisites

- To run and schedule data quality assessment scans, users need the data quality steward role.
- Currently, you can set the Microsoft Purview account to allow public access or managed virtual network access so that data quality scans can run.

Data quality life cycle

Data quality scanning is the **seventh** step in the data quality life cycle for a data asset. The previous steps are:

1. Register and scan a data source in Microsoft Purview Data Map.
2. Assign users data quality steward permissions in Unified Catalog so they can use all data quality features.
3. Add your data asset to a data product. This step is needed only if you want to run data quality for data assets associated to a data product to calculate data product data quality score.
4. If you want to measure data quality for a standalone data asset, add a data asset from the Data Map to a domain to run data quality scan for the asset without association to a data product.
5. Set up a data source connection to prepare your source for data quality assessment.
6. Configure a container or storage to store data quality error records for review and correct the failed records.
7. Configure and run data profiling for an asset in your data source.. When profiling is complete, browse the results for each column in the data asset to understand your data's current structure and state.
8. Set up data quality rules based on the profiling results, and apply them to your data asset.
9. Run data quality scan to understand and improve the quality of your data.

Supported multicloud data sources

Browse the supported data source document to view the list of supported data sources, including file formats for data profiling and data quality scanning, with and without virtual network support.

> **Important**
>
> Data quality for Parquet file is designed to support:

1. A directory with Parquet Part File. For example: **./Sales/{Parquet Part Files}**. The Fully Qualified Name must follow `https://(storage account).dfs.core.windows.net/(container)/path/path2/{SparkPartitions}`. Make sure there are no {n} patterns in directory/sub-directory structure. It must be a direct FQN leading to {SparkPartitions}.
2. A directory with Partitioned Parquet Files, partitioned by columns within the dataset like sales data partitioned by year and month. For example: **./Sales/{Year=2018}/{Month=Dec}/{Parquet Part Files}.**

Both of these essential scenarios, which present a consistent parquet dataset schema, are supported. **Limitation:** It isn't designed to or won't support N arbitrary hierarchies of directories with Parquet files. **We recommend presenting data in (1) or (2) constructed structure.**

Microsoft Purview Data Quality doesn't currently support specifying data quality rules on maps, lists, and structs in Parquet data structures.

Supported authentication methods

Currently, Microsoft Purview can only run data quality scans by using Managed Identity as authentication option. Data quality services run on **Apache Spark 3.4** and **Delta Lake 2.4**. For more information about supported regions, see data quality overview.

ⓘ Important

- If you update the schema on the data source, you need to import schema from data quality overview page by using **schema import feature** before running a data quality scan.
- Virtual network isn't supported for Google BigQuery.

Configure and run incremental data quality scan

1. In Unified Catalog, select **Health Management**, then select **Data quality**.
2. Select a **governance domain** from the list.
3. Configure a data source connection to the assets you're scanning for data quality if you haven't already done so.

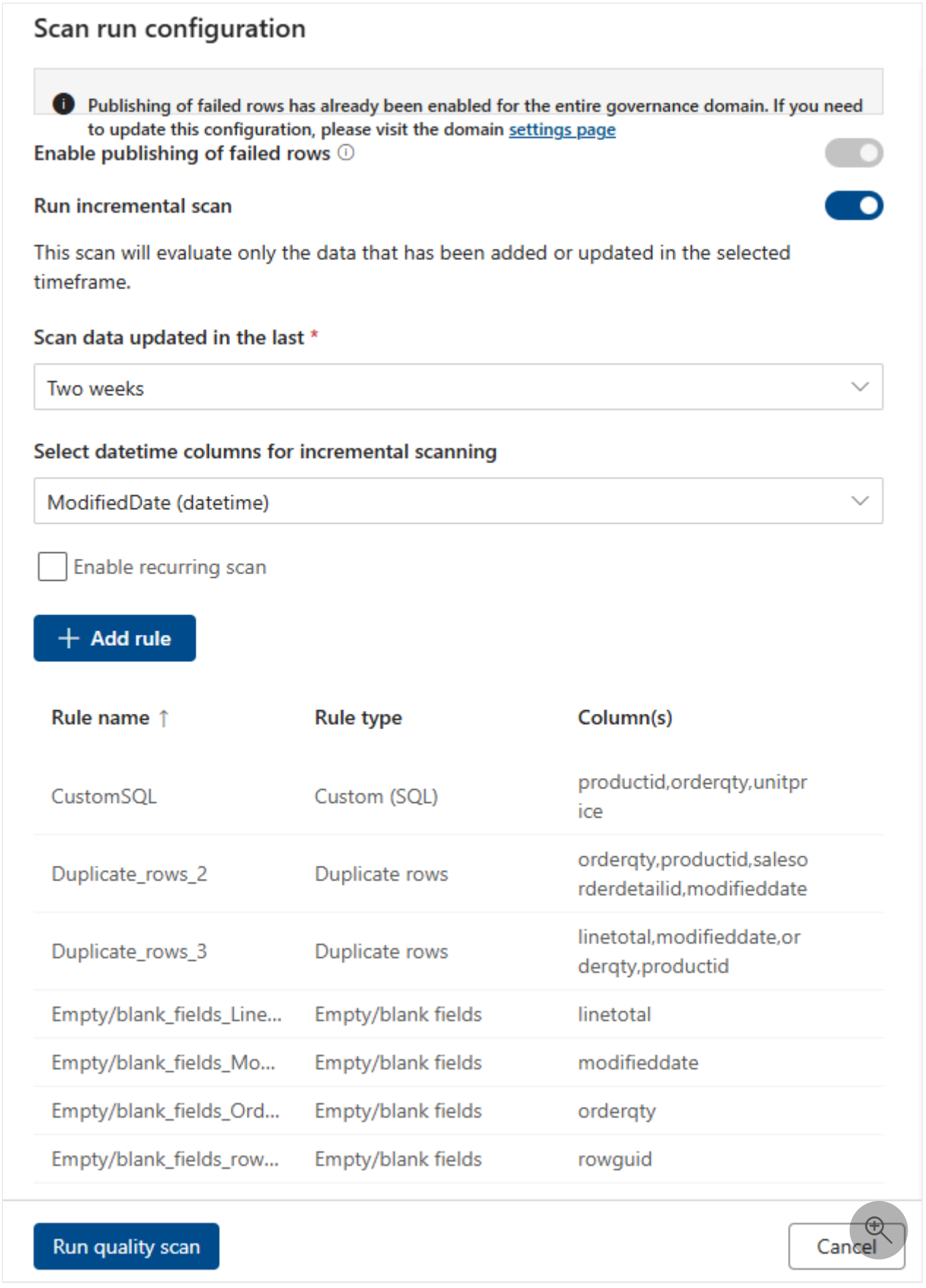
4. Select a **data product** to assess the data quality of the **data assets** associated with that product. If you want to scan a data asset without data product association, select **Add assets** to add a data asset to the asset list from Data Map.
5. Select the name of a data asset, which takes you to the data quality **Overview** page.
6. Browse the existing data quality rules and add new rules by selecting **Rules**. Create custom rules using SQL expression or ADF expression language if out of the box rules aren't enough for your use case. Toggle on or off the rules you added if you want to turn off or on any rule.
7. Select the scan recurrence option:

    - Incremental
    - Full
    - Both

8. Select **Run quality scan** from the upper-right corner on the overview page.
9. Enable the Run incremental scan toggle and provide the following details:

    - Choose the time window for updated data: last day, last week, last two weeks, or last month
    - Select the datetime column to use for incremental scanning
    - Use the **Add Rule** tab to select rules from the predefined rule list

10. Select **Run Quality Scan** to execute the incremental scan.
11. Schedule incremental scan. To set up a schedule to run data quality scan, enable scheduled execution by selecting **Enable recurring scan** and setting the desired schedule. If you configure a recurring schedule, the **Run Quality Scan** button is read-only on the Scan Run Configuration page.
12. While the scan is running, you can track its progress from the data quality monitoring page in the governance domain.



> **Note**
>
> - Both **datetime** and **date** columns are supported, but we recommend using a **datetime** column. If the data asset doesn't contain a timestamp column, incremental data quality scans can't be configured for daily scans. A date column prevents the system from

calculating the last 24 hours accurately, which is required to slice one day's data for data quality scans. The datetime data type is also more commonly used for created, updated, and other timestamp columns.

- Filter days: 90 days for three months, 180 days for six months, and 365 days for one year.
- Auto-scheduling to run a weekly scan is enabled at the asset level by default. If you want to disable auto-scheduling for a data asset, go to **Health management** > **Data quality** > **Run quality scan**, and then select the option to disable the scheduled scan.
- When you manually run a data quality scan, a notification appears on the Data Quality page prompting you to configure alerts. Ignore this notification if you don't want to configure alerts.
- Cumulative score is calculated as the average of the incremental scan score and the full (regular) scan score.
- Incremental data quality scan isn't supported for on-premises data sources.

> **Important**
>
> Scan type, incremental data quality score, and cumulative data quality score don't publish to Azure Data Lake Storage Gen2 or Fabric Lakehouse for self-serve analytics.

Delete previous data quality scans and history

When you remove a data asset from a data product, if that data asset has a data quality score, delete the data quality score first, and then remove the data asset from the data product.

When you delete data quality history data, you remove the profile history, the data quality scan history, and data quality rules. Data quality actions aren't deleted.

To delete previous data quality scans of a data asset, follow these steps:

1. In Unified Catalog, select **Health Management**, then select **Data quality**.
2. Select a **governance domain** from the list.
3. Select the **data product** from the list. Select the **data asset** from the list to go to the Data quality overview page. If you're using the standalone data asset data quality feature, select the data asset from the data asset list.
4. Select the ellipsis (...) at the upper right of the Data quality overview page.
5. Select **Delete data quality data** to delete the history of data quality runs.

> **Note**
>
> - Use **Delete data quality data** for test runs, errored data quality runs, or if you're removing a data asset from a data product.
> - The system stores up to 50 snapshots of data quality profiling and data quality assessment history. To delete a specific snapshot, select the desired history run and select the delete icon.

Import schema

If the data type in a schema is undefined, incorrectly defined, or changed in the source, your data quality job might fail. If it fails, reimport the schema by using the schema import capability. You can import schemas for data sources on both public networks and behind private endpoints. For a list of supported data sources, see Data sources and file formats supported for data quality. To import a schema from your data sources, follow these steps:

- Select **Data quality** from **Health Management**.
- Select a business domain, and then select a data product. Select a data asset from that data product. You arrive at the data quality overview page.
- Select **Schema**, then select the **Schema management** toggle.
- Select **Import schema** to import the schema.

Related content

- Data quality for Fabric data estate
- Data quality for Fabric Mirrored data sources
- Data quality for Fabric shortcut data sources
- Data quality for Azure Synapse serverless and data warehouses
- Data quality for Azure Databricks Unity Catalog
- Data quality for Snowflake data sources
- Data quality for Google BigQuery

Next steps

- Monitor data quality scan.
- Review your scan results to evaluate your data product's current data quality.
- Configure alerts for data quality scan results.

ⓘ **Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 05/10/2026

Data quality error records

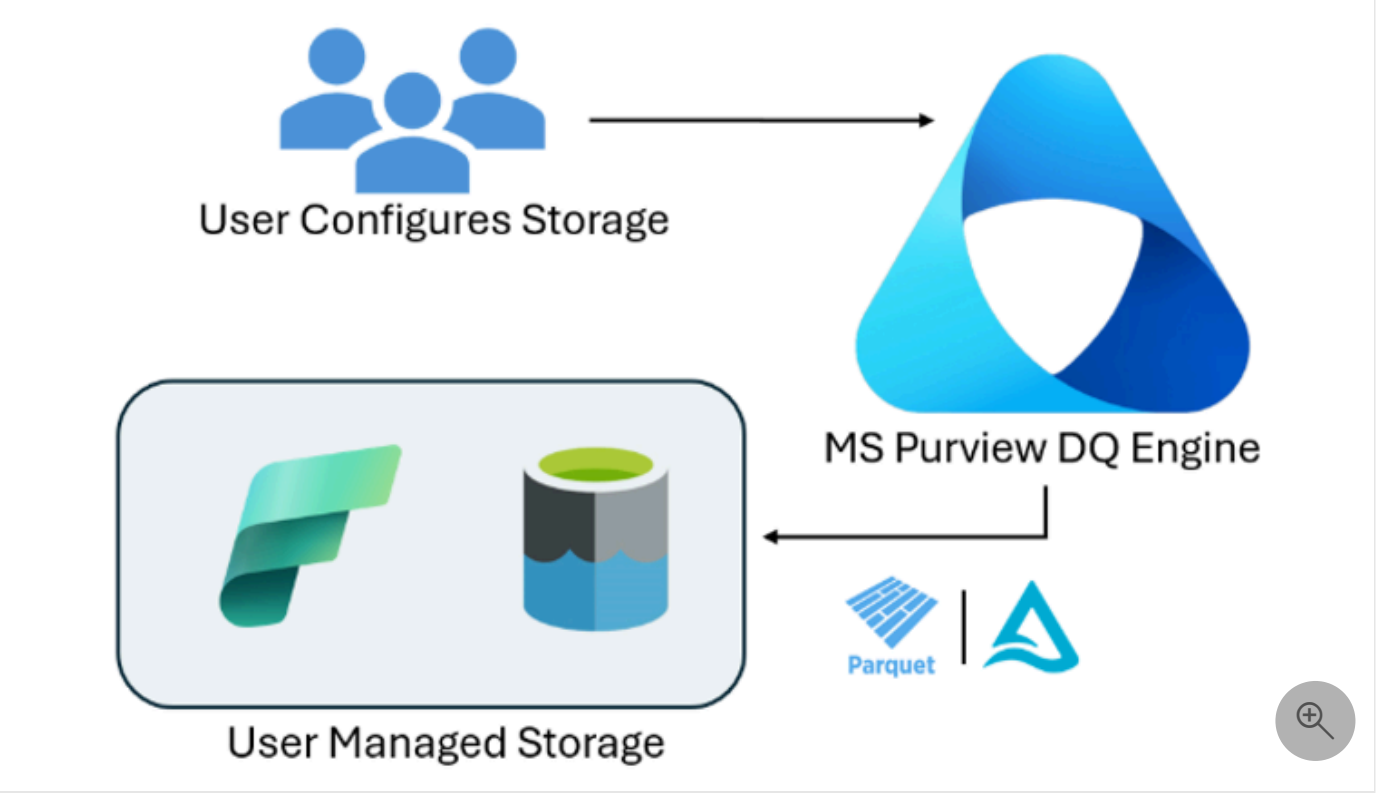
When you perform data quality checks in Microsoft Purview Unified Catalog, you can configure and publish data quality error records for rule exceptions. By using this setup, you can review and remediate these records directly in your managed storage.

For cloud data sources, you can store and manage data quality error records in Microsoft Fabric Lakehouse or Azure Data Lake Storage Gen2. For on-premises data sources, you can create an error records database in Oracle or SQL Server to store and manage data quality error records.

By using this feature, Microsoft Purview Data Quality users can identify, analyze, and address data quality problems. It helps ensure data accuracy, completeness, consistency, timeliness, and validity across systems and use cases.

Architecture

To publish data quality error records for rule exceptions, you need to set up your managed storage in Microsoft Purview. The Data Governance Administrator for your organization is responsible for configuring the storage details on the Microsoft Purview administration page. The Data Quality Steward can enable the toggle either from the Data Quality Manage console or the data quality overview page at **Unified Catalog** > **Health management** > **Data quality**.



Configuration storage location

- In Unified Catalog, go to **Health management** > **Data quality**.
- Select a governance domain from the list to open its details page.
- Select **Manage**, then select **Settings**.
- Select **Azure Region** from the dropdown list of all supported Azure regions.
- After selecting the Azure region, select **New** to configure storage details.
- Select **Storage Type**: Fabric or Data Lake Storage Gen2.
- Enter the **Location URL**.
- If you select **Fabric** as the storage type, enter the **Workspace ID** and **Lakehouse ID**. Grant contributor access to Microsoft Purview Manage Service Identity (MSI) to your Fabric workspace.
- If you select **Azure Data Lake Storage Gen2** as the storage type, enter following details:
  - Add **Azure subscription**. To find your Azure subscription name, go to `portal.azure.com`.
  - Add **Resource**. To find the resource name, select **Home** > select **Resource** from the **Resource list**.
  - Add **Container** name. To find the container name, select **Data Lake Storage** from the overview page in `portal.azure.com`
  - Add **Folder Path**. This is optional information.
  - Grant Storage Blob Data Contributor access to Microsoft Purview Manage Service Identity (MSI) to your Data Lake Storage Gen2 container.
- Test the connection.
- Save the configuration.
- You can **Save as draft** if you don't have all the required information so that you won't lose what you've entered.

Compute is a prerequisite for private endpoint or virtual network storage setup. For private endpoint approval request, select **Enable managed V-Net** checkbox. An approval request is sent to your Azure portal, which you must approve to enable virtual network access. This process is similar to setting up a Data Quality connection. You must deploy the compute in the same Azure region as the storage account.

If you configure your Fabric tenant to run on a virtual network or in private link, select the **Enable managed V-NET** checkbox on the **Configure Storage** page. Then add the Private Link Resource ID as listed here:

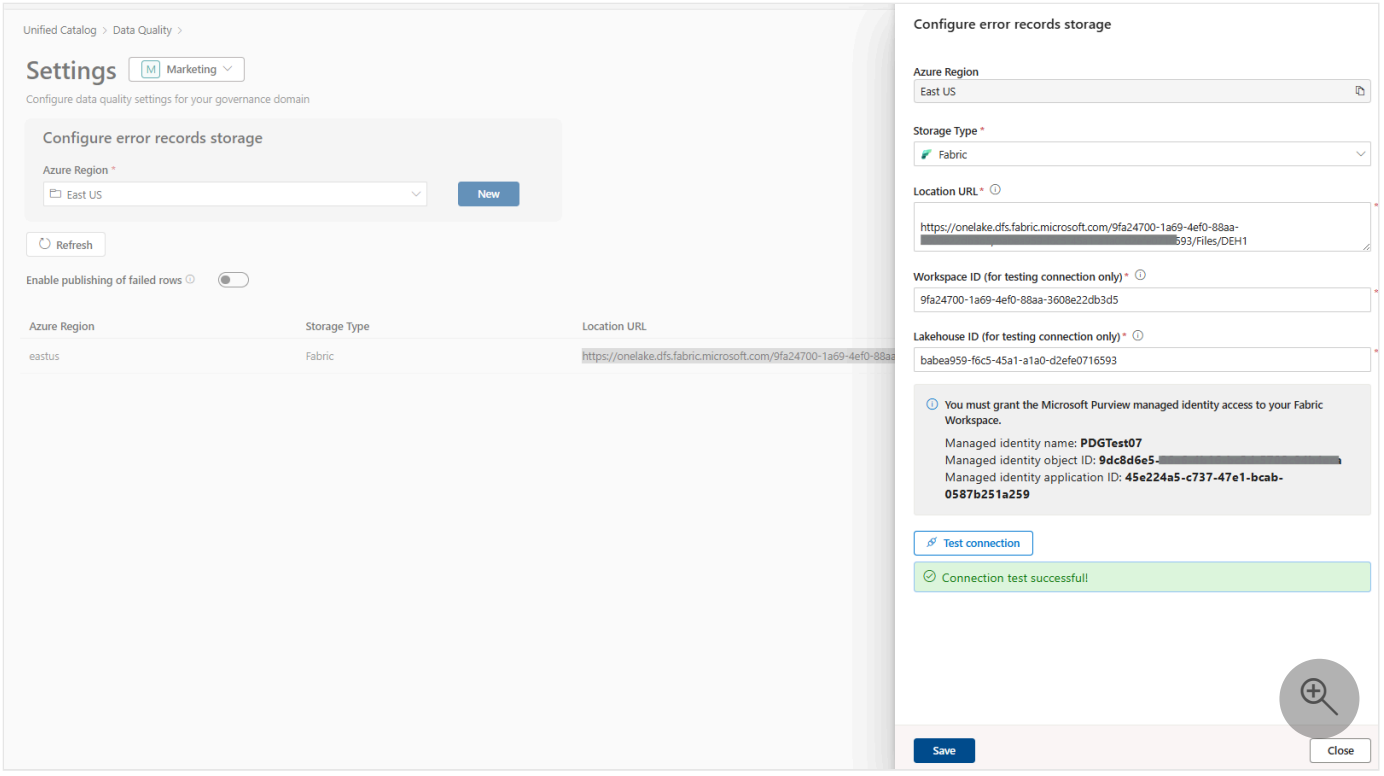
/subscriptions/07d669d6-83f2-4f15-8b5b
4a4b31a7050e/resourceGroups/pdgbugbashfabricvnet/providers/Microsoft.PowerBI/priva

teLinkServicesForPowerBI/fabricvnetpl.

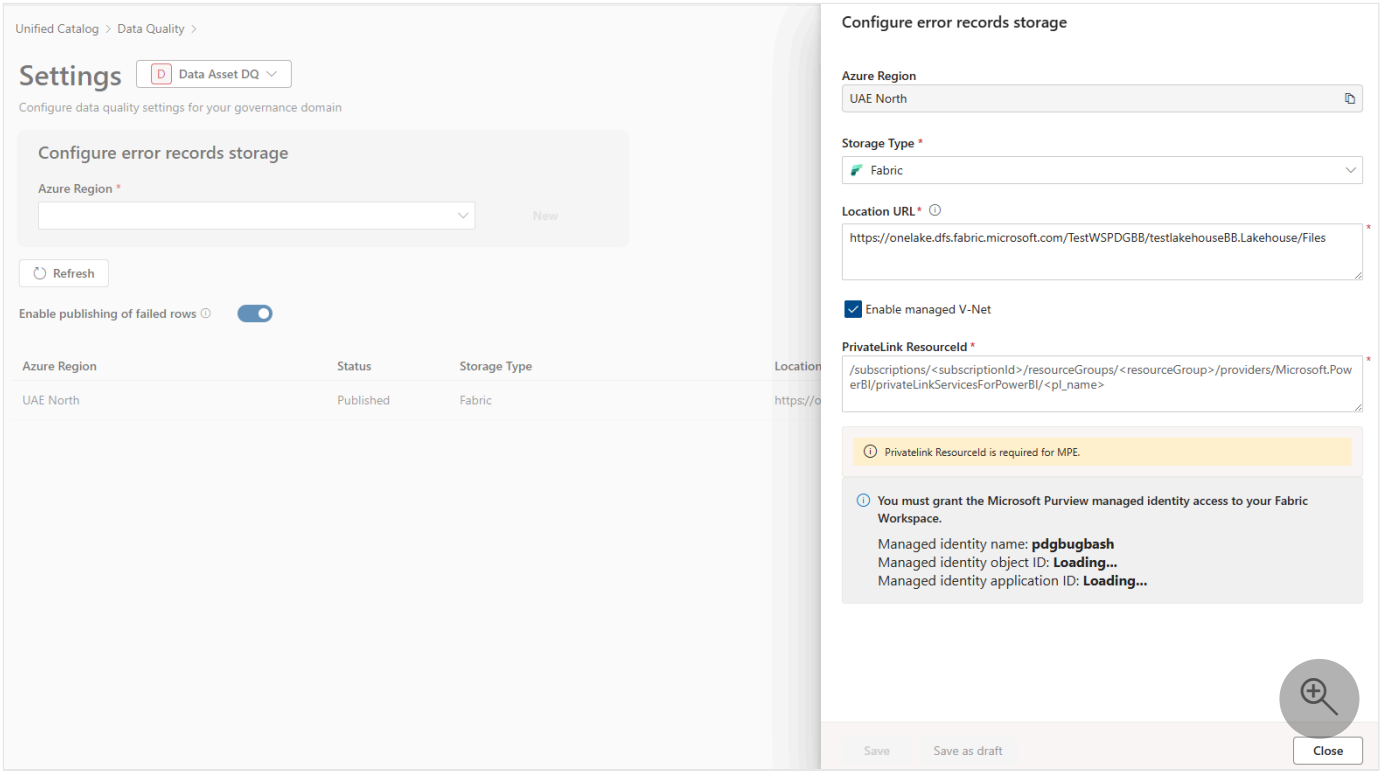
To set up compute for virtual networks, see Set up data quality for managed virtual networks. To set up and use private link for your Fabric tenant, follow the instructions to set up the private link for the Fabric tenant before configuring data quality virtual network connection and compute allocation.

> **Note**
>
> - Supported cloud data sources for Data Quality error records are: Azure SQL, Azure SQL MI, Azure Data Lake Storage G2, Azure Databricks, Microsoft Fabric, Azure Synapse, and Google BigQuery.
> - Supported on-premises data sources for error records are: Oracle and SQL Server.
> - If your data products include assets from multiple Azure regions, you must create a region-specific data quality error records folder for each region to meet data residency requirements. Microsoft Purview Data Quality doesn't copy data between Azure regions.
> - If you already created a storage location without specifying an Azure region, you must update it to include the region. Data quality failed/error records won't publish until a region is provided.
> - If you configure multiple region-specific folders for storing data quality failed/error records, Microsoft Purview Data Quality automatically routes the error records to the correct regional folder during the scan.
> - If you're using virtual networks, you can't test the connection. The **Test Connection** tab is disabled.
> - You can use error sink behind virtual networks only if you're also running the job in a virtual network connection. If not, error row publishing is skipped.
> - For a non-virtual network source connection to Azure Databricks Unity Catalog, Snowflake, Google BigQuery, Azure Synapse Analytics, on-premises Oracle Database, Fabric, and Azure SQL MI database, you need to set up error sink in the Microsoft Purview account region, not the data asset region.
> - For virtual network sources and all other types of non-virtual network sources, the error sink region has to match the asset region.

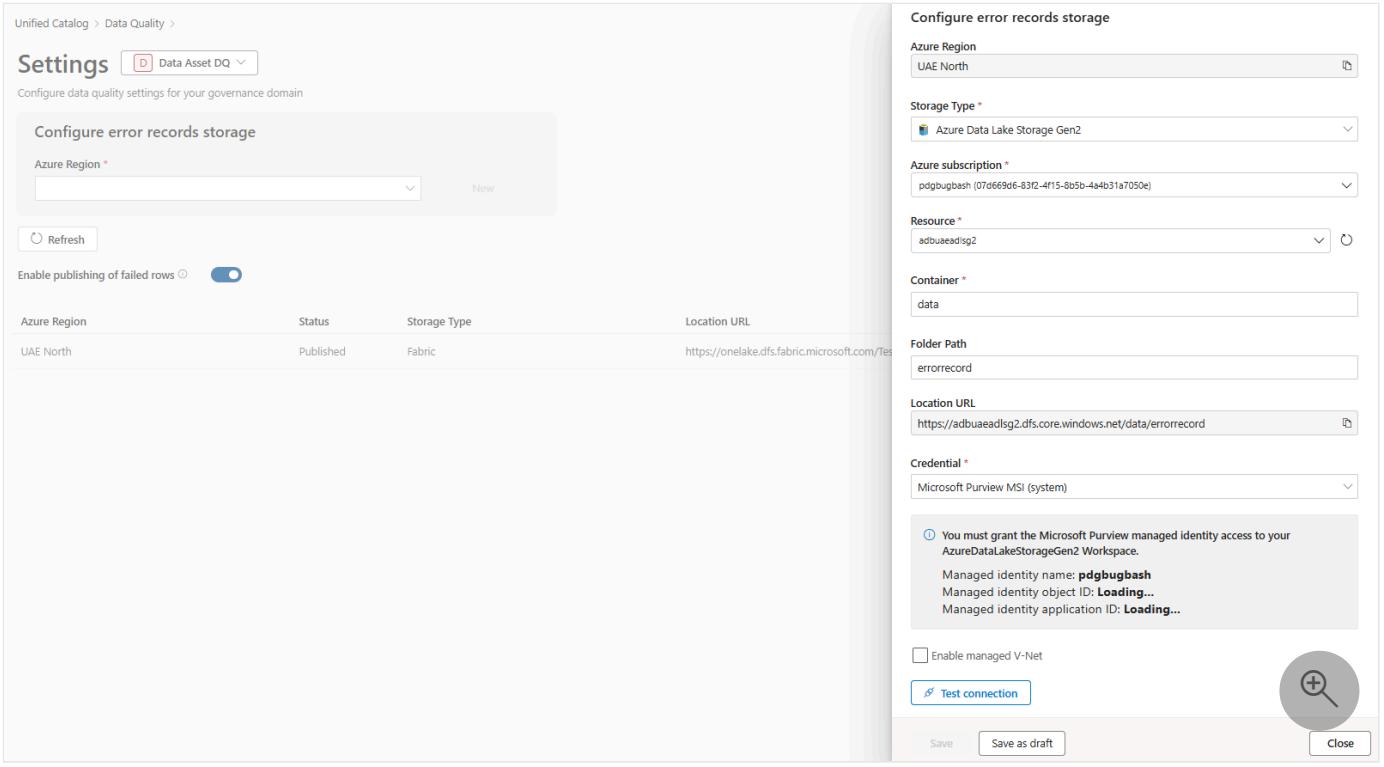
Data quality error records storage setting for non-virtual network storage - Fabric:



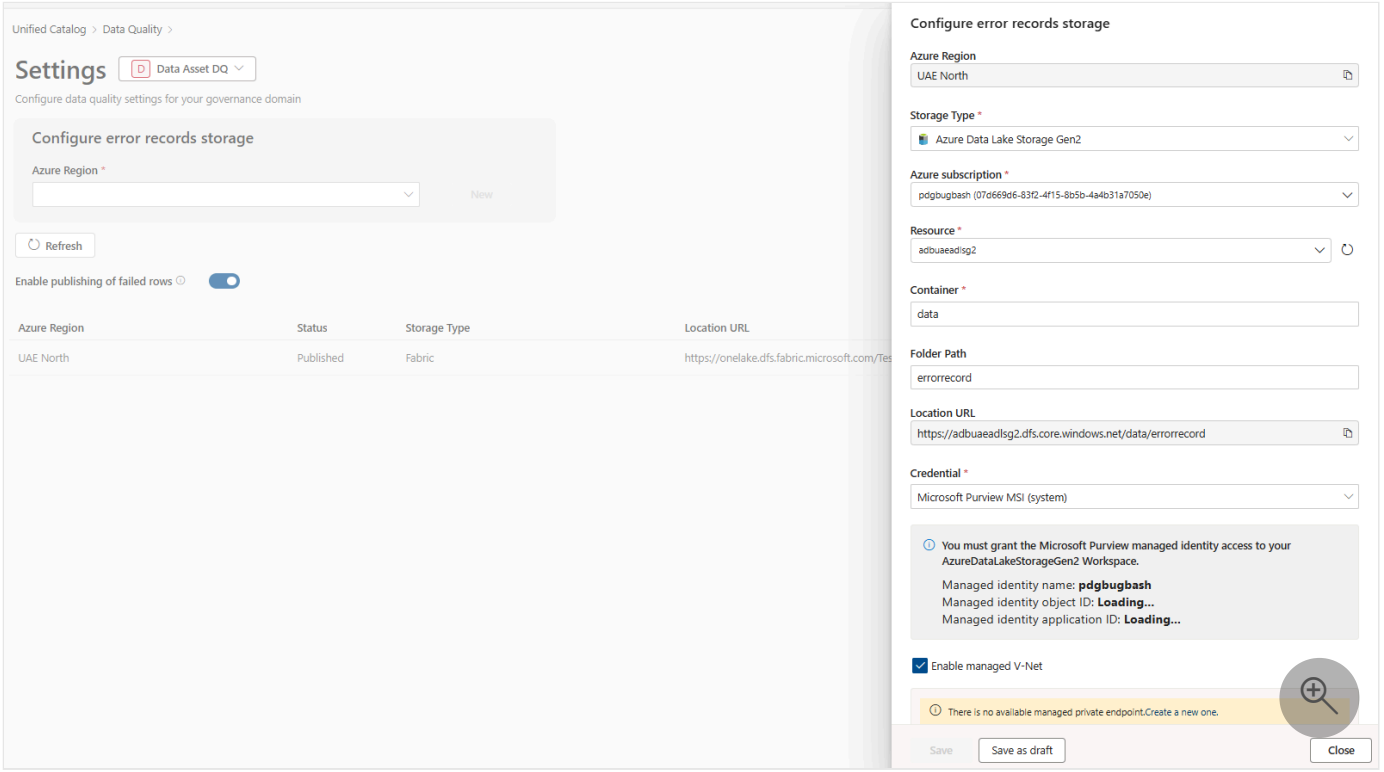
Data quality error records storage setting for virtual storage - Fabric:



Data quality error records storage setting for non-virtual storage - AdlsG2:



Data quality error records storage setting for virtual network storage - AdlsG2:



⊗ Caution

- If you don't have storage and a connection set up to store data quality error records in the asset region (as mentioned in the preceding section), the data quality scan job still succeeds, but error row publishing is skipped.

- After configuring the storage and connection, you can enable the data quality error record feature either by going to **Health management** > **Data quality** > **Manage** > **Settings** for the selected governance domain, or from the Data Quality overview page for a data asset.
- To analyze the published erroneous records, use the latest versions of Spark (>= 3.xx) on Synapse or Databricks.
- Error sink configurations in a **published** state will only be used for error sink output, *not* in a **draft** state. If the configuration is in a draft state, the system skips error row publishing to the configured storage.

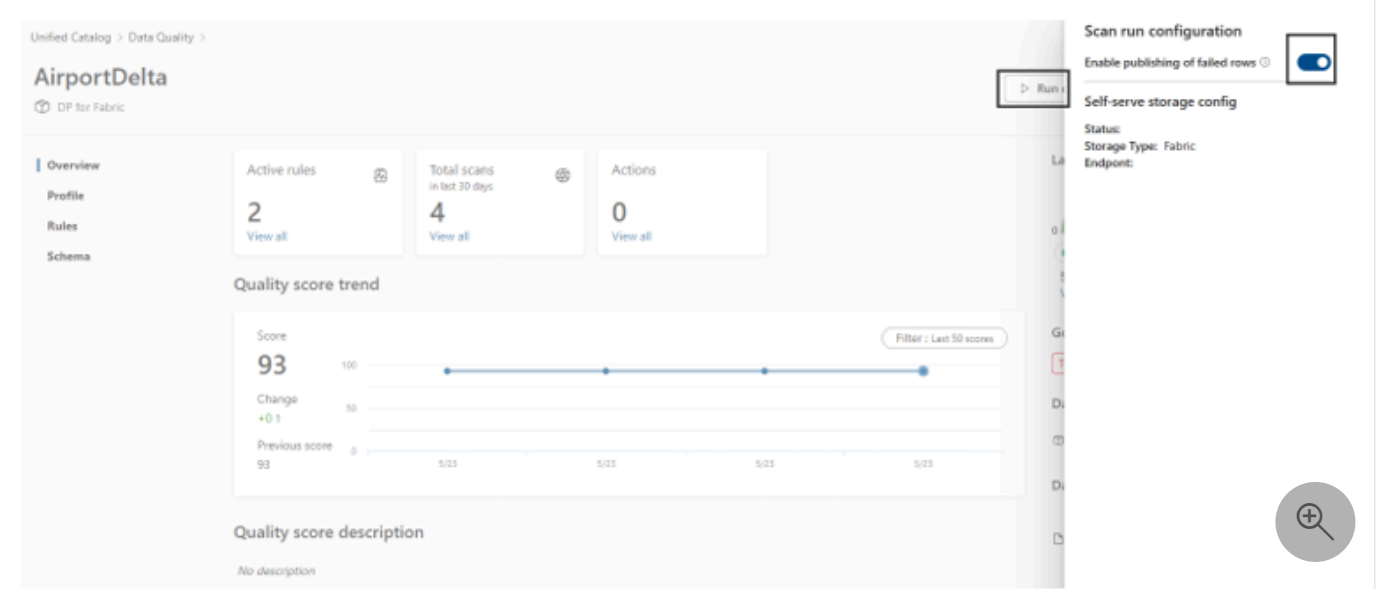
Data quality error detection

To detect data quality errors in your data, configure and set up a data quality scan. Select the data asset and configure rules for the critical columns that need a Data Quality measurement. You can use out-of-the-box rules, create custom rules, or use the rule suggestion feature to recommend rules for a data quality scan. For more information, see data quality rules.

Enable error recording

To get error records, enable the capture and storage of error records for every data quality job run by following these steps:

1. Follow the instructions to start a data quality scan on a data asset.
2. After selecting **Run quality scan** on the data asset's page, turn on the toggle for **Enable publishing of failed rows**.



You can also enable error recording for all data assets with data quality rules associated with data products in a **governance domain** level by following these steps:

- In Unified Catalog, go to **Health management** > **Data quality**.
- Select a governance domain from the list to open its details page.
- Select **Manage**, then select **Settings**.
- Turn on the toggle for **Enable publishing of failed rows**.

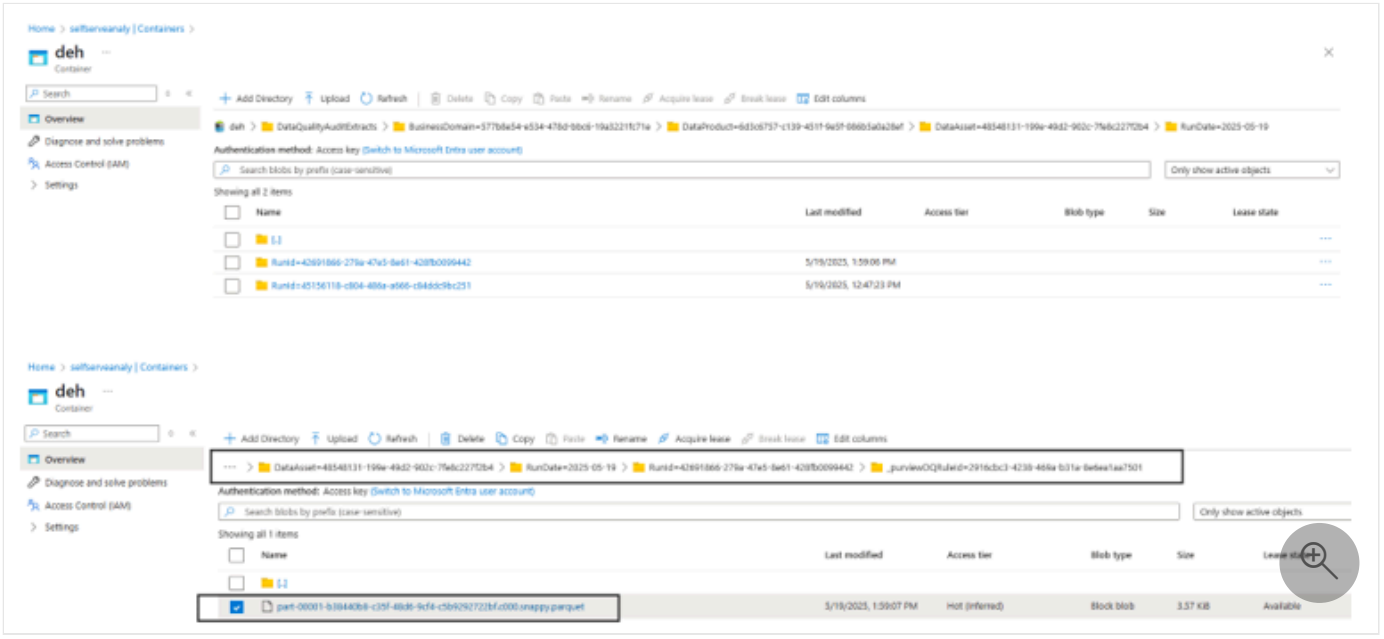
Set up schedule to run

Set up the schedule to run your data quality job and publish the error records into the configured storage. For more information, see configure and run data quality scans.

Though you can run data quality scans on an ad-hoc basis by selecting **Run quality scan** on a data asset page, in production scenarios it's likely that the source data is constantly updated. Regularly monitor data quality to detect any problems. Automating the scanning process helps you manage regular updates of quality scans.

Check the error records in the configured storage

1. Find the location of the published error rows from the data quality overview page.
2. Select **View Scan report**, in the top-right corner just below the **Latest quality score**.
3. Go to the path shown in your Data Lake Storage Gen2 or Fabric Lakehouse folder, as seen in the following example:



Data Lake Storage Gen2 folder hierarchy

Data Lake Storage Gen2 container (for example, folder name is DEH):

1. DataQualityAuditExtract
2. Governance domain (BusinessDomain)
3. DataProduct
4. DataAsset
5. RunDate
6. RunId
7. Purview DQRunId
8. Error record file

Read data quality error records from Data Lake Storage Gen2

Use a Synapse or Databricks notebook to read data from Data Lake Storage Gen2 container by using this PySpark script:

Python

```
#For complete error records of all jobs run on the asset
path =
"abfss://TestWSPDGBB@onelake.dfs.fabric.microsoft.com/TestSelfServeLakehouse.Lakehouse
/Files/SelfServeFolder/DataQualityAuditExtracts/BusinessDomain_577b8e54-e534-478d-
bbc6-19a3221fc71e/DataProduct_a04c82a2-2372-4b9e-9e0a-4dfd8959ee85/DataAsset_4b2b2644-
c94a-447f-9b2e-32961ac0170b"
df = spark.read.format("delta").load(path)
display(df)

#For All Runs on a particular Date
path =
"abfss://TestWSPDGBB@onelake.dfs.fabric.microsoft.com/TestSelfServeLakehouse.Lakehouse
/Files/SelfServeFolder/DataQualityAuditExtracts/BusinessDomain_577b8e54-e534-478d-
bbc6-19a3221fc71e/DataProduct_a04c82a2-2372-4b9e-9e0a-4dfd8959ee85/DataAsset_4b2b2644-
c94a-447f-9b2e-32961ac0170b"
df = spark.read.format("delta").load(path).where("_purviewDQRunDate = '2025-08-06'")
display(df)

#For a Specific run
path =
"abfss://TestWSPDGBB@onelake.dfs.fabric.microsoft.com/TestSelfServeLakehouse.Lakehouse
/Files/SelfServeFolder/DataQualityAuditExtracts/BusinessDomain_577b8e54-e534-478d-
bbc6-19a3221fc71e/DataProduct_a04c82a2-2372-4b9e-9e0a-4dfd8959ee85/DataAsset_4b2b2644-
c94a-447f-9b2e-32961ac0170b"
df = spark.read.format("delta").load(path).where("_purviewDQRunId = '647b8a8f-7140-
4158-9914-9c2409a5ec67'")
display(df)

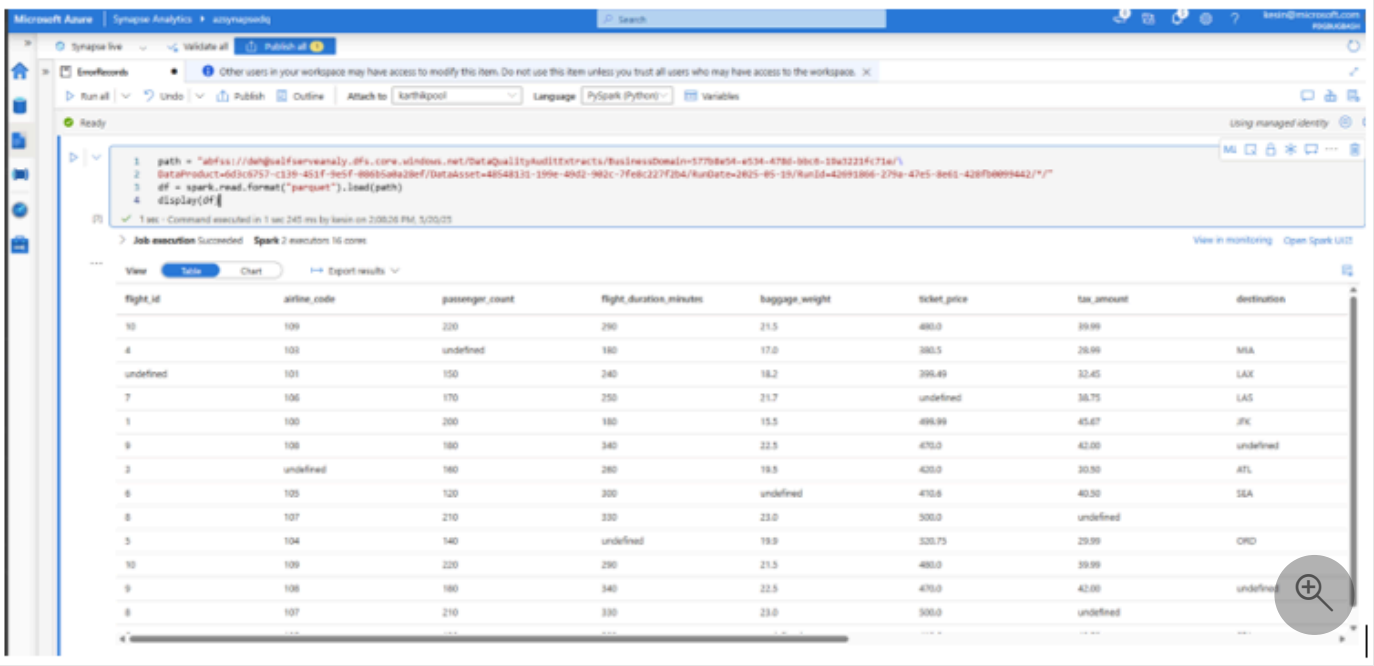
#For a specific rule of the run
path =
```

```
"abfss://TestWSPDGBB@onelake.dfs.fabric.microsoft.com/TestSelfServeLakehouse.Lakehouse
/Files/SelfServeFolder/DataQualityAuditExtracts/BusinessDomain_577b8e54-e534-478d-
bbc6-19a3221fc71e/DataProduct_a04c82a2-2372-4b9e-9e0a-4dfd8959ee85/DataAsset_4b2b2644-
c94a-447f-9b2e-32961ac0170b"
df = spark.read.format("delta").load(path).where("_purviewDQRunId = '647b8a8f-7140-
4158-9914-9c2409a5ec67' AND _purviewDQRuleId = '32d57f8d-87dc-4862-90c9-
c3bcad1051ad'")
display(df)

#For a specific rule across all runs
path =
"abfss://TestWSPDGBB@onelake.dfs.fabric.microsoft.com/TestSelfServeLakehouse.Lakehouse
/Files/SelfServeFolder/DataQualityAuditExtracts/BusinessDomain_577b8e54-e534-478d-
bbc6-19a3221fc71e/DataProduct_a04c82a2-2372-4b9e-9e0a-4dfd8959ee85/DataAsset_4b2b2644-
c94a-447f-9b2e-32961ac0170b"
df = spark.read.format("delta").load(path).where(" _purviewDQRuleId = '32d57f8d-87dc-
4862-90c9-c3bcad1051ad'")
display(df)

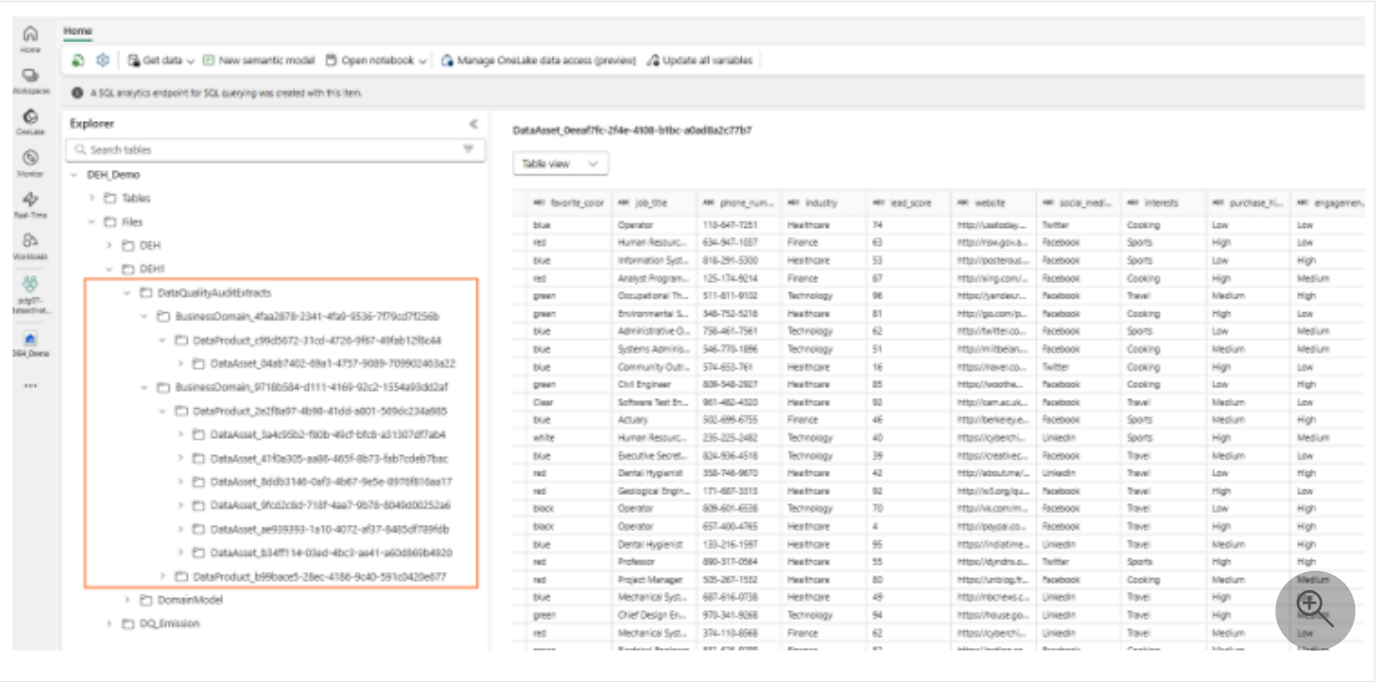
#For comparing two runs
path =
"abfss://TestWSPDGBB@onelake.dfs.fabric.microsoft.com/TestSelfServeLakehouse.Lakehouse
/Files/SelfServeFolder/DataQualityAuditExtracts/BusinessDomain_577b8e54-e534-478d-
bbc6-19a3221fc71e/DataProduct_a04c82a2-2372-4b9e-9e0a-4dfd8959ee85/DataAsset_4b2b2644-
c94a-447f-9b2e-32961ac0170b"
df = spark.read.format("delta").load(path).where("_purviewDQRunId = '647b8a8f-7140-
4158-9914-9c2409a5ec67' OR _purviewDQRunId = 'a508f960-254c-4eb1-9097-03307402475b'")
display(df)
```

Example dataset:



Read error records from Fabric Lakehouse

1. Go to the path shown in your Lakehouse, and browse all fail records published to Fabric Lakehouse.
2. Create a shortcut from Delta Parquet to Delta Table. Or you can use a Delta Parquet file to create a dashboard of your data quality error records, as seen in this example:



Use a notebook to create a shortcut table of failed data asset records. See the notebook script.

Output data model

You can view failed records at the level of a governed data asset. For each new assessment run that succeeds on the data asset, the process adds a new set of failed records for each rule evaluation to the failed records table.

Format: Delta

Output schema: Each row of output contains all the columns of an input data asset row that failed a rule evaluation, along with a set of metadata columns that you can use to identify and analyze the failed row.

Primary key: Composite key of input data asset columns + _purviewDQRunId + _purviewDQRuleId

Expand table

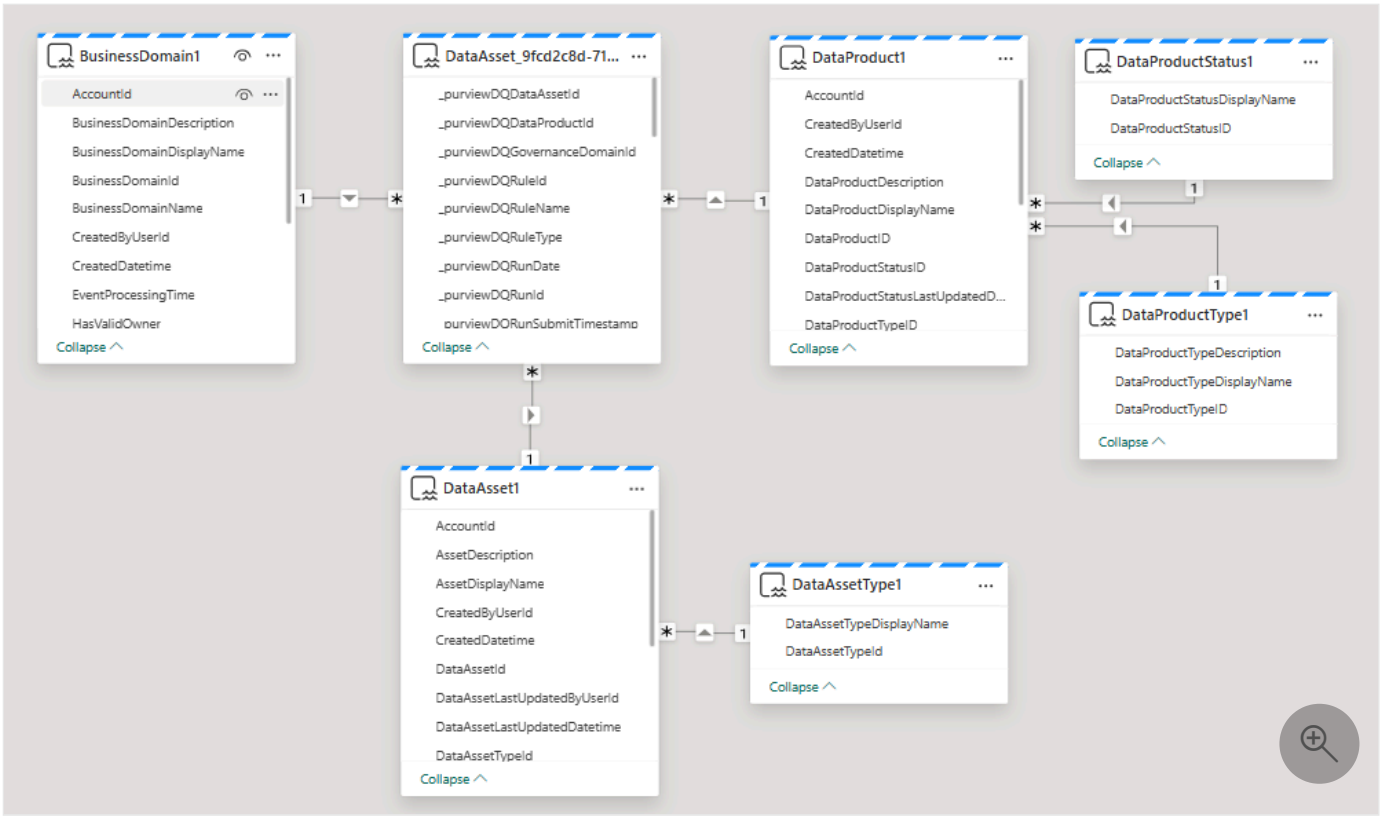
Metadata columns	Data type	Description
_purviewDQRunId	string	A GUID that represents the Run ID of the assessment scan available to the user at the time of run. Also, a Delta partition-column.
_purviewDQRunDate	string	The run date in YYYY-MM-DD format. Also, a Delta partition-column.
_purviewDQRuleId	string	The quality Rule ID that corresponds to the failed rule. Also, a Delta partition-column.
_purviewDQRuleName	string	The Rule Name at the time of job run.
_purviewDQRuleType	string	The Rule type (for example: UniqueCheck, NotNull, Regex).
_purviewDQGovernanceDomainId	string	The governance domain ID of the data asset that you ran.
_purviewDQDataProductId	string	The data product of the data asset that you ran.
_purviewDQDataAssetId	string	The data asset ID.
_purviewDQRunSubmitTimestamp	string	The exact timestamps of the run submit time in the default UTC time-zone per the ISO format.

Create a Power BI dashboard for your data quality error records

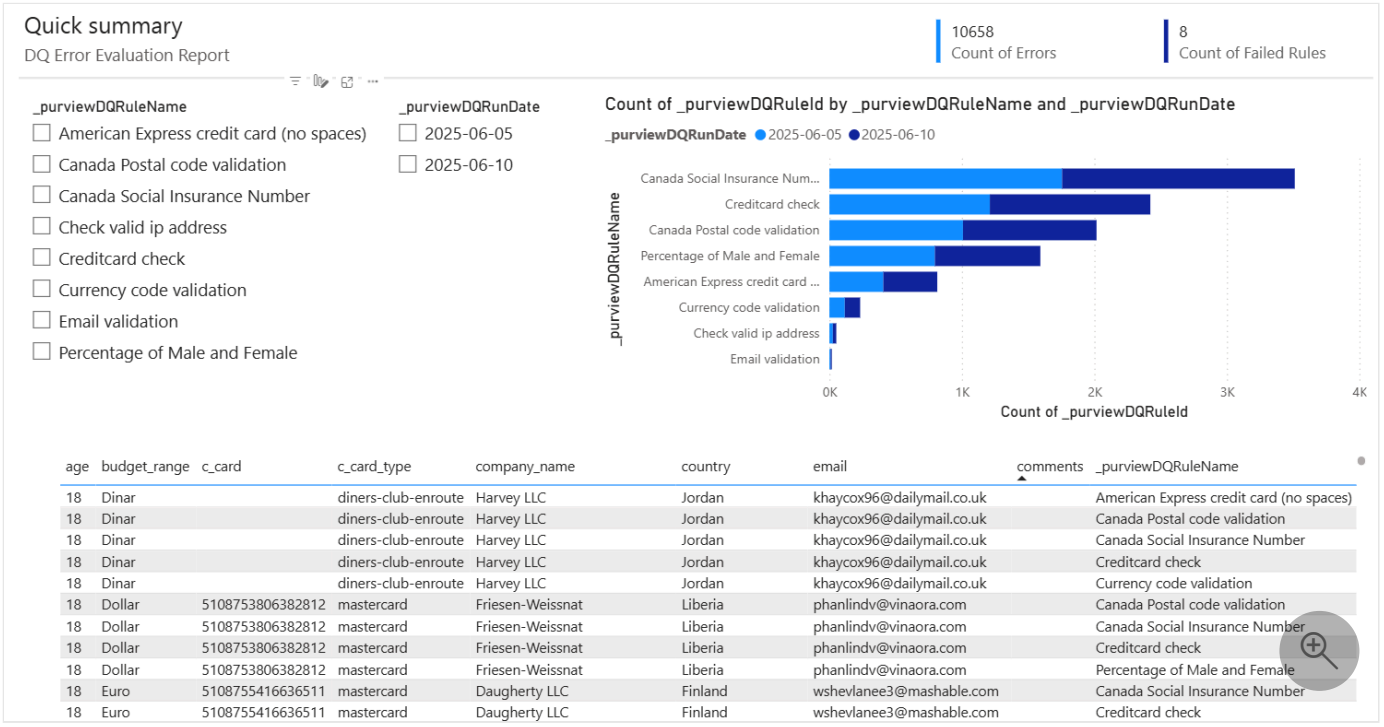
Use the Business Domain ID to link data products and their associated data assets for reporting data quality error records.

- One business domain can link to multiple data products.
- Each data product can link to multiple data assets.
- Each data asset can have multiple rules.
- Each rule can generate multiple error records.

This image illustrates a data model to create a basic report of your data quality error record:



This image illustrates a sample report created with the data model shown in the preceding image:



age	budget_range	c_card	c_card_type	company_name	country	email	comments	_purviewDQRuleName
18	Dinar		diners-club-enroute	Harvey LLC	Jordan	khaycox96@dailymail.co.uk		American Express credit card (no spaces)
18	Dinar		diners-club-enroute	Harvey LLC	Jordan	khaycox96@dailymail.co.uk		Canada Postal code validation
18	Dinar		diners-club-enroute	Harvey LLC	Jordan	khaycox96@dailymail.co.uk		Canada Social Insurance Number
18	Dinar		diners-club-enroute	Harvey LLC	Jordan	khaycox96@dailymail.co.uk		Creditcard check
18	Dinar		diners-club-enroute	Harvey LLC	Jordan	khaycox96@dailymail.co.uk		Currency code validation
18	Dollar	5108753806382812	mastercard	Friesen-Weissnat	Liberia	phanlindv@vinaora.com		Canada Postal code validation
18	Dollar	5108753806382812	mastercard	Friesen-Weissnat	Liberia	phanlindv@vinaora.com		Canada Social Insurance Number
18	Dollar	5108753806382812	mastercard	Friesen-Weissnat	Liberia	phanlindv@vinaora.com		Creditcard check
18	Dollar	5108753806382812	mastercard	Friesen-Weissnat	Liberia	phanlindv@vinaora.com		Percentage of Male and Female
18	Euro	5108755416636511	mastercard	Daugherty LLC	Finland	wshevlanee3@mashable.com		Canada Social Insurance Number
18	Euro	5108755416636511	mastercard	Daugherty LLC	Finland	wshevlanee3@mashable.com		Creditcard check

Limitations

- 100,000 failed records are published per rule for each run.
- Datasets up to 100 million rows with up to 40 rules have been benchmarked.
- As data is stored in organizations' managed storage, the organization owns role-based access control for the data. Microsoft publishes data quality error rows to an organization's

storage based on the organization's consent.

> **Note**
>
> Failed rows can't be accessed by selecting the URL. For security reasons, direct browsing of the error record database through the URL isn't supported.
>
> Example URL: `https://onelake.dfs.fabric.microsoft.com/xxxxxxx-4xxx-88aa-xxxe22db3d5/babea959-fxxxxxxxxx-d2efe0716593/Files/xxxx1/DataQualityAuditExtracts/BusinessDomain_xxxx-xxxx-ccc9-92c2-1554a93dd2af/DataProduct_abc33xxxxyu-28ec-4186-01fg-591c0420e677/DataAsset_xxxxx-583a-49ee-abd6-6543464654e4ss`

Script to create shortcut

You can automate the creation of table shortcuts from data quality audit extracts failed row data by using this script in a Fabric notebook. After creating the shortcut tables, your error records for new job runs on the asset shortcut table automatically refresh in Fabric.

Python

```
# NEW Script to automate the creation of table shortcuts from Data Quality audit extracts failed row data.

# Update these three values for your Fabric Workspace ID, Lakehouse ID and Purview BYOC Self-serve store path
workspaceId = "f28dc1c8-360c-4788-9f46-e69853b1c40d" #Example: f28dc1c8-360c-4788-9f46-e69853b1c40d
lakehouseId = "77d6df6b-64ab-4628-985f-9365591a85a8" #Example: 77d6df6b-64ab-4628-985f-9365591a85a8
dataqualityauditPath = "Files/SelfServeFolder/DataQualityAuditExtracts" #Example: "Files/DEH2/DataQualityAuditExtracts"

#Use for lakehouses with Lakehouse-Schema (Public preview feature)
DestinationShortcutUri = f"abfss://{workspaceId}@onelake.dfs.fabric.microsoft.com/{lakehouseId}/Tables/DataQualityAuditExtracts"

#Use for lakehouses without Lakehouse-Schema
#DestinationShortcutUri = f"abfss://{workspaceId}@onelake.dfs.fabric.microsoft.com/{lakehouseId}/Tables"

import sempy.fabric as fabric
from sempy.fabric.exceptions import FabricHTTPException
import fnmatch
```

```
import re
import os
from collections import deque

SourceUri =
f"abfss://{workspaceId}@onelake.dfs.fabric.microsoft.com/{lakehouseId}/{dataqualityau
ditPath}"

request_headers = {
            "Authorization": "Bearer " + mssparkutils.credentials.getToken("pbi"),
            "Content-Type": "application/json"
        }
print(request_headers)


def is_delta_table(uri: str):
    #print("Checking for uri:" + uri)
    delta_log_path = os.path.join(uri, "_delta_log")
    return mssparkutils.fs.exists(delta_log_path)

def extract_onelake_https_uri_components(uri):
    pattern = re.compile(r"abfss://([^@]+)@[^/]+/([^/]+)/(.*)")
    match = pattern.search(uri)
    if match:
        workspace_id, item_id, path = match.groups()
        return workspace_id, item_id, path
    else:
        return None, None, None

def is_valid_onelake_uri(uri: str) -> bool:
    workspace_id, item_id, path = extract_onelake_https_uri_components(uri)
    if "abfss://" not in uri or workspace_id is None or item_id is None or path is
None:
        return False
    return True

def get_onelake_shorcut(workspace_id: str, item_id: str, path: str, name: str):
    shortcut_uri = (
        f"v1/workspaces/{workspace_id}/items/{item_id}/shortcuts/{path}/{name}"
    )
    result = client.get(shortcut_uri).json()
    return result

def get_last_path_segment(uri: str):
    path = uri.split("/")  # Split the entire URI by '/'
    return path[-1] if path else None

def create_onelake_shorcut(SourceUri: str, dest_uri: str, result: list):
    src_workspace_id, src_item_id, src_path = extract_onelake_https_uri_components(
        SourceUri
    )

    dest_workspace_id, dest_item_id, dest_path = extract_onelake_https_uri_components(
        dest_uri
    )
```

```
    name = get_last_path_segment(SourceUri)
    dest_uri_joined = os.path.join(dest_uri, name)

    # If the destination path already exists, return without creating shortcut
    if mssparkutils.fs.exists(dest_uri_joined):
        print(f"Table already exists: {dest_uri_joined}")
        result.append(dest_uri_joined)
        return None

    request_body = {
        "name": name,
        "path": dest_path,
        "target": {
            "oneLake": {
                "itemId": src_item_id,
                "path": src_path,
                "workspaceId": src_workspace_id,
            }
        },
    }
    #print(request_body)

    shortcut_uri = f"v1/workspaces/{dest_workspace_id}/items/{dest_item_id}/shortcuts"
    print(f"Creating shortcut: {shortcut_uri}/{name}..")
    try:
        client.post(shortcut_uri, json=request_body, headers=request_headers)
    except FabricHTTPException as e:
        print(e)
        print(e.error_reason)
        return None

    return get_onelake_shorcut(dest_workspace_id, dest_item_id, dest_path, name)


client = fabric.FabricRestClient()

queue = deque([SourceUri])
result = []

while queue:
    current_uri = queue.popleft()

    #print(current_uri)

    if  is_delta_table(os.path.join("", current_uri)):
        #print(current_uri)
        shortcut = create_onelake_shorcut(os.path.join("", current_uri),
DestinationShortcutUri, result)
        if shortcut is not None:
            result.append(shortcut)
        continue;

    # List subitems in the current folder
    subitems = mssparkutils.fs.ls(current_uri)
```

```
    for item in subitems:
        if item.isDir:
            queue.append(item.path)


print(f"{len(result)}" + " error record shortcut tables present in the lakehouse: ")
for item in result:
    print(item)
```

Error records for on-premises data sources

Data quality error records for on-premises data sources are stored within the corresponding on-premises Oracle or SQL Server databases. Users must define a designated error database where data quality failed records will be stored. For each supported on-premises data source, users are required to create an error database in the same environment. For example, if users are performing data quality assessments on an Oracle database, they must create an Oracle error database to store the resulting failed or error records. Similarly, if users are assessing data quality on a SQL Server database, they must create a SQL Server database to store the corresponding data quality failed records.

Example of Oracle database data quality error records setup

Error record table naming convention

Table Name Format

`PURVIEW_ERRORRECORD_<SCHEMA>_<ASSET_NAME>`

Example

- Schema: ORACLENEW_DATAGOV
- Asset / Table Name: EMPLOYEE_DETAILS_ERRORRECORD

Final table name:

PURVIEW_ERRORRECORD_ORACLENEW_DATAGOV_EMPLOYEE_DETAILS_ERRORRECORD

This convention ensures:

- Clear traceability to the source schema
- Easy identification of error-record tables
- Alignment with Purview data governance standards

Data load scenarios

Scenario 1: First-Time Load (Table Creation)

- Table, columns, and indexes are created during the first execution
- Records inserted: 1,000,000
- Rules applied: 1

Timestamps:

- CreatedAt: 01/28/2026 07:37:55 PM
- LastModifiedAt: 01/28/2026 07:41:34 PM

Scenario 2: Load on existing table

- Table already exists
- Only data insertion occurs
- Records inserted: 1,000,000
- Rules applied: 1

Timestamps:

- CreatedAt: 01/28/2026 07:33:19 PM
- LastModifiedAt: 01/28/2026 07:36:54 PM

Failure recovery and restart logic

- If data insertion fails, the process resumes from the last successfully inserted record.
- Previously inserted records aren't duplicated.

This ensures:

- Idempotent behavior
- No data loss
- Minimal reprocessing

Querying error record tables

Source Table: MANUFACTURERS Error Record Table:
PURVIEW_ERRORRECORD_DBO_MANUFACTURERS

Sample query

```
SELECT TOP 5 *
FROM PURVIEW_ERRORRECORD_DBO_MANUFACTURERS WITH (NOLOCK)
WHERE _PurviewJobId = 'c7f7c92e-9a36-46c5-86bc-bc3c595d5ddd';
```

This query retrieves error records associated with a specific Microsoft Purview job execution.

SQL Server performance metrics

1 Million Record Insert Performance

Execution time:

- CreatedAt: 01/29/2026 04:45:05 PM
- LastModifiedAt: 01/29/2026 04:54:14 PM

Total Time Taken: ~9 minutes

This demonstrates efficient bulk insertion using parallel processing.

> **Note**
>
> - Proper Oracle and SQL Server permissions should be configured for data governance operations.
> - Use standardized naming convention for error record tables.
> - Large-scale inserts should be optimized using parallel batching.
> - Enabling failure recovery ensures safe restart without duplication.

Related content

- Data quality for Fabric Lakehouse data
- Data quality for Fabric mirrored data sources
- Data quality for Fabric shortcut data sources
- Data quality for Azure Synapse serverless and data warehouses
- Data quality for Azure Databricks Unity Catalog
- Data quality for Snowflake data sources
- Data quality for Google BigQuery
- Data quality native support for Iceberg data

Last updated on 05/12/2026

Monitor data quality jobs

Article • 04/21/2025

Data quality job monitoring enables data quality stewards to see the progress of data profiling, data quality rule generation, and data quality scanning jobs. Both manually configured and scheduled jobs' progress can be viewed from the data quality monitoring page.

Browse data quality job status

1. In Microsoft Purview Unified Catalog, select **Health Management**, then select **Data quality**.
2. Select a **governance domain**.
3. From the **Manage** dropdown menu, select **Monitoring**.
4. The **Activities** tab shows a summary of all activities. Use the filters to scope your view by submit time, status, activity type, scan type, or scan name. Select the activity for more details.
5. You can switch between governance domains from this menu, using the governance domain tab next to the page title.
6. Select the **Scan** tab to see the progress of scheduled scans.
7. Select and open a completed job for a data asset. Under **Scan details**, you can view the processing units consumed at **PU consumed**. To calculate the cost of the data quality job, multiply the **PU consumed** with SKU cost.

Next steps

- Review your scan results to evaluate your data product's current data quality.
- Configure alerts for data quality scan results.
- Run a data profile.
- Create a data quality scan.

Review data quality scores of data assets

After you create data quality rules and run a data quality scan, your data assets receive a data quality score based on results from your rules. This article explains how scores are calculated so you can better understand your data quality results and develop action items to improve your data's integrity.

Understand data quality scores

Data quality rules describe the state of the data. They show how far the data is from the ideal state described by the rules. Each rule produces a score that describes how close the data is to its desired state. Most rules are straightforward; they divide the total number of rows that pass the assessment by the total number of rows to arrive at the score.

The formula used to calculate the data quality score for a rule against data in a column is:

`[(total number of passed records)/(passed records + failed records + miscast records + empty records)]`

- Numerator = total number of passed records
- Denominator = total number of records (number of passed records + number of failed records + number of miscast records + number of empty records)

> ⓘ Note
>
> - **Passed**: Number of records that passed an applied rule.
> - **Unevaluable**: The columns required to evaluate this rule aren't evaluable.
> - **Failed**: Number of records that failed an applied rule.
> - **Miscast**: The data type of the asset and the type that customer listed it as not matching. It can't be converted to the expressed type.
> - **Empty**: Null or blank records.
> - **Ignored**: Rows didn't participate in the rule evaluation. Users can express rows to ignore; for example, ignore all rows that have email="n/a", or ignore all rows where departmentCode='test' or 'internal'. Ignored records aren't used for score calculation.

Microsoft Purview Data Quality then gives a sense for the state of each column by generating a column score. This score is the average of all the scores of the rules on that column. Once the column scores are calculated, the formula used to calculate average percentage data quality score for data products and governance domains is:

`[(Percentage 1 + Percentage 2) / (Sample size 1 + Sample size 2)] x 100`

The score is multiplied by 100 to make the scores more readable.

Example calculation

Let's imagine there's a column that doesn't have the **Empty/blank fields** rule defined on it. This implies that null values are allowed for this column. So certain rules, like the **Unique values** rule, will filter out null values in that case.

For example: If the asset has 10,000 rows in a table but 3,000 were null and 500 weren't unique then the score would be: `((10000 - 3000 - 500)/(10000 - 3000) )* 100 = 93`

The null rows are ignored when evaluating the data and determining a score.

Specific rule scores

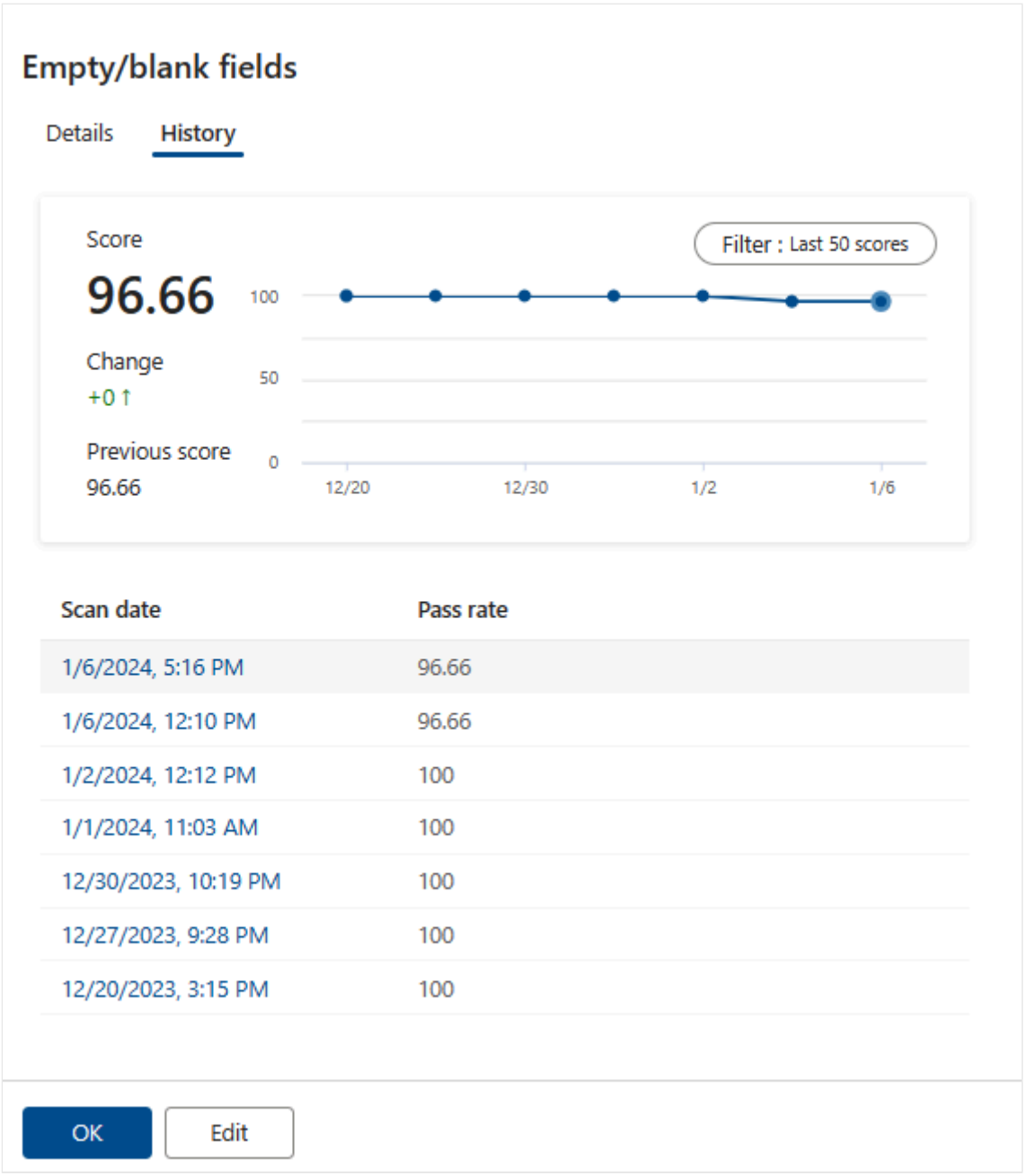
For **Custom** rules, there's a capability similar to what you might see for the unique values rule. But in this case, the filter isn't on nulls, but rather the filter expression.

Some rules, like the **Freshness** rule, are either pass or fail, so their scores are either 0 or 100. The Freshness rule is applied at the data asset level, not at the column level.

Rule details and history

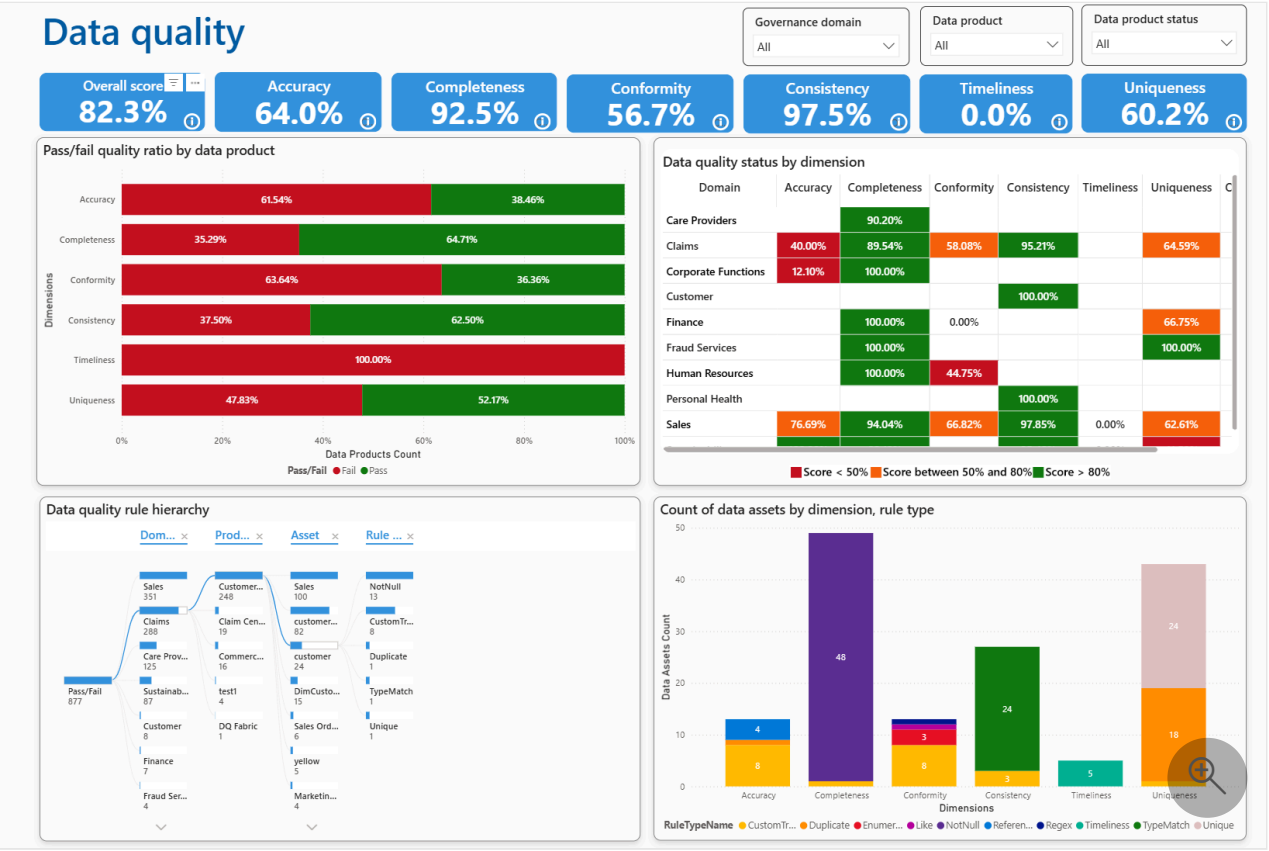
You can view the details and history of rule scores by selecting a rule. Selecting a specific rule name and navigating to the rule **History** tab, you see the trend of the different scan runs for the particular rule.

- Rule **Details** provides information on the number of rows passed, failed, and ignored for the various runs for the particular rule. Rules that are in **draft state** (OFF state) won't have their scores contribute to the global score. Rules in a draft state won't be run at all during quality scans and so won't have scores.



- Columns and rules have a many-to-many relationship, the same rule can be applied to many columns, and many rules can be applied to the same column. You can view the trend pattern of each rule by going to the **Schema** tab on the asset's data quality page and viewing the line in the **Trend** column.
- Asset level data quality score trends are available for the last 50 runs. This quality score trend helps data quality stewards to monitor data quality trend and fluctuations month over month. Data quality can also trigger alerts for every data quality scan if the quality score doesn't meet the threshold or business expectation.
- Global score is the average of the production rules defined on the asset. The asset level global score is also rolled up to the data product level and governance domain level. The global score is intended to be the official definition of the state of the data asset, data product, and governance domain in context of quality of data.

- A summary report is created for data quality dimension. This report contains data quality score for each data quality dimension. Global score for the governance domain is published also in this report. You can browse the quality score for each governance domain, data product, and data asset from this Power BI report.



> **Note**
>
> - Data quality dimensions are recognized terms used by data practitioners to describe a feature of data that can be measured or assessed against defined standards in order to quantify the quality level of data used to run a business.
> - The data quality score for an asset is the arithmetic average of the scores of the rules applied to its columns.
> - The data quality score for a data product is the arithmetic average of the data quality scores of the data assets associated with that data product.
> - The data quality score for a governance domain is the arithmetic average of the data quality scores of the data products associated with that domain.

Next steps

- Data quality scan

Last updated on 03/30/2026

Data quality improvement actions

The quality of data directly affects the quality of decision making. Good quality data provides business process insights, a better understanding of customers, and faster problem identification. Data quality actions identify problems that you should address to improve data quality in your data estate.

Required roles

- To see data quality actions in your environment, users need the data quality reader role.
- To assign and modify actions, users need the data quality steward role.

Data quality action types

The types of data quality actions generated are:

- A rule that falls below default thresholds.
- A global data quality score that falls below the threshold.
- A data quality scanning and data quality profiling job that fails.
- A data quality scanning and data quality profiling job that skips due to no changes on data since the last run.
- Outlier presence in profiled columns.
- Too many nulls in the profiled data asset.

Data quality actions severity

Data quality actions have a severity level and color indicator depending on the data quality score:

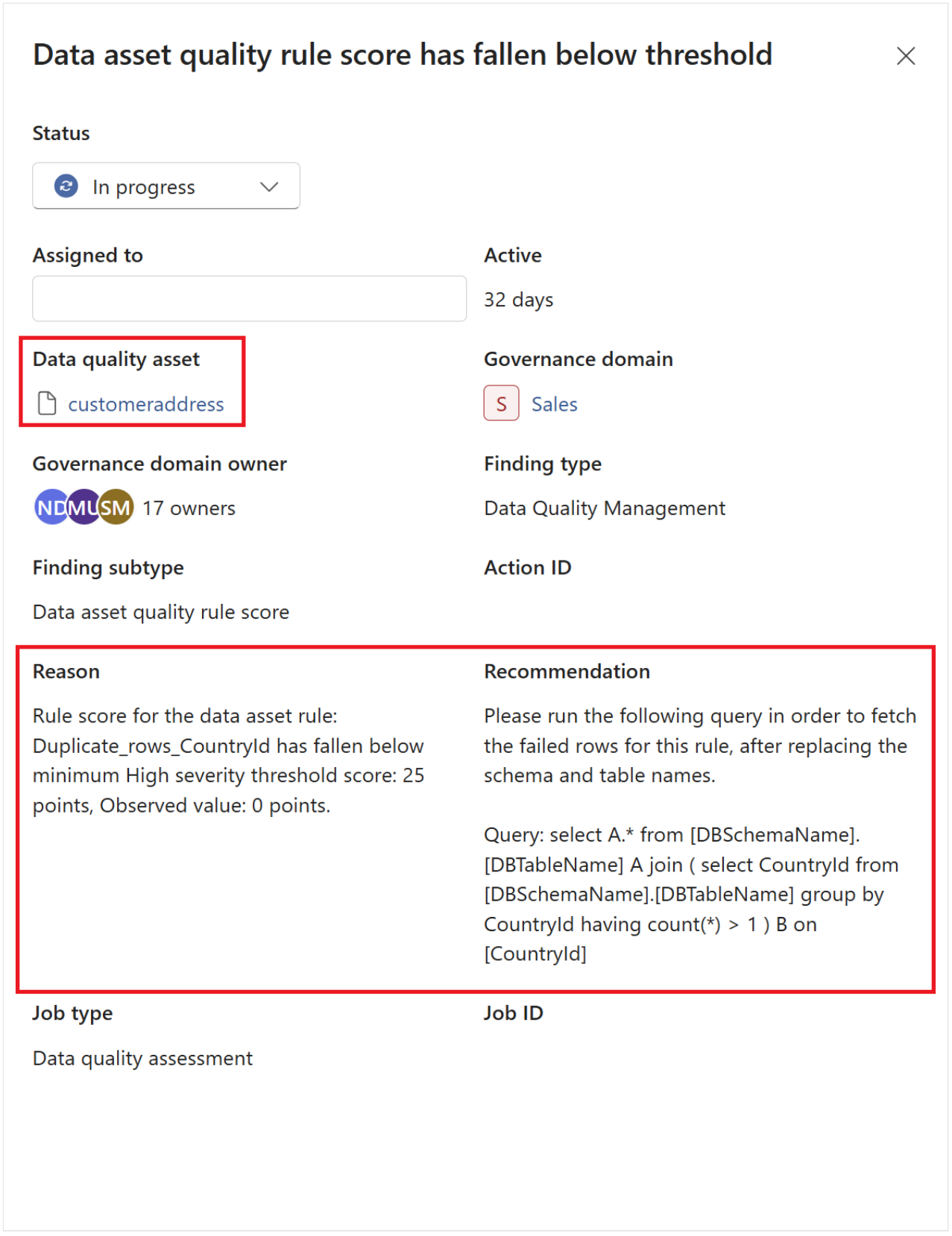
- **High (red)**: Data quality score is less than **25**.
- **Medium (orange)**: Data quality score is less than **50**.
- **Low (yellow)**: Data quality score is less than **75**.

View and manage data quality actions for governance domains

Data quality actions roll up from the data asset level to the governance domain level. To browse actions and assign them to users for further action:

1. In Microsoft Purview Unified Catalog, select **Health Management**, then select **Data quality**.
2. Select a governance domain from the list.
3. The **Action items** button shows the number of action items for the domain. Select **Action items** to view all data quality actions.
4. Select an action from the list of actions to view the action's details page.
5. From here, you can assign the action to a user to investigate and resolve. The **Recommendation** section suggests the action to take and provides queries to find the low quality data.

> **Note**
>
> - Not every action has a SQL query because not all actions are data related or related to a data quality rule. Some actions are system related, such as a failed or skipped data quality job.
> - If a custom rule exception occurs, which means that the data quality score generated from the custom rule doesn't meet the defined threshold, an action is created. However, a SQL query isn't suggested.

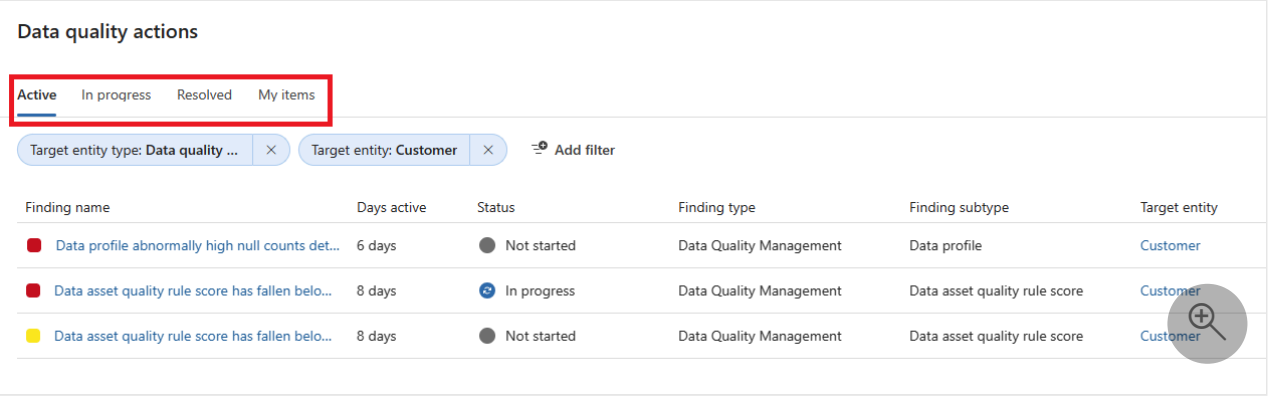


6. Update the *Status**, then select **Save**.

View and manage data quality actions for data assets

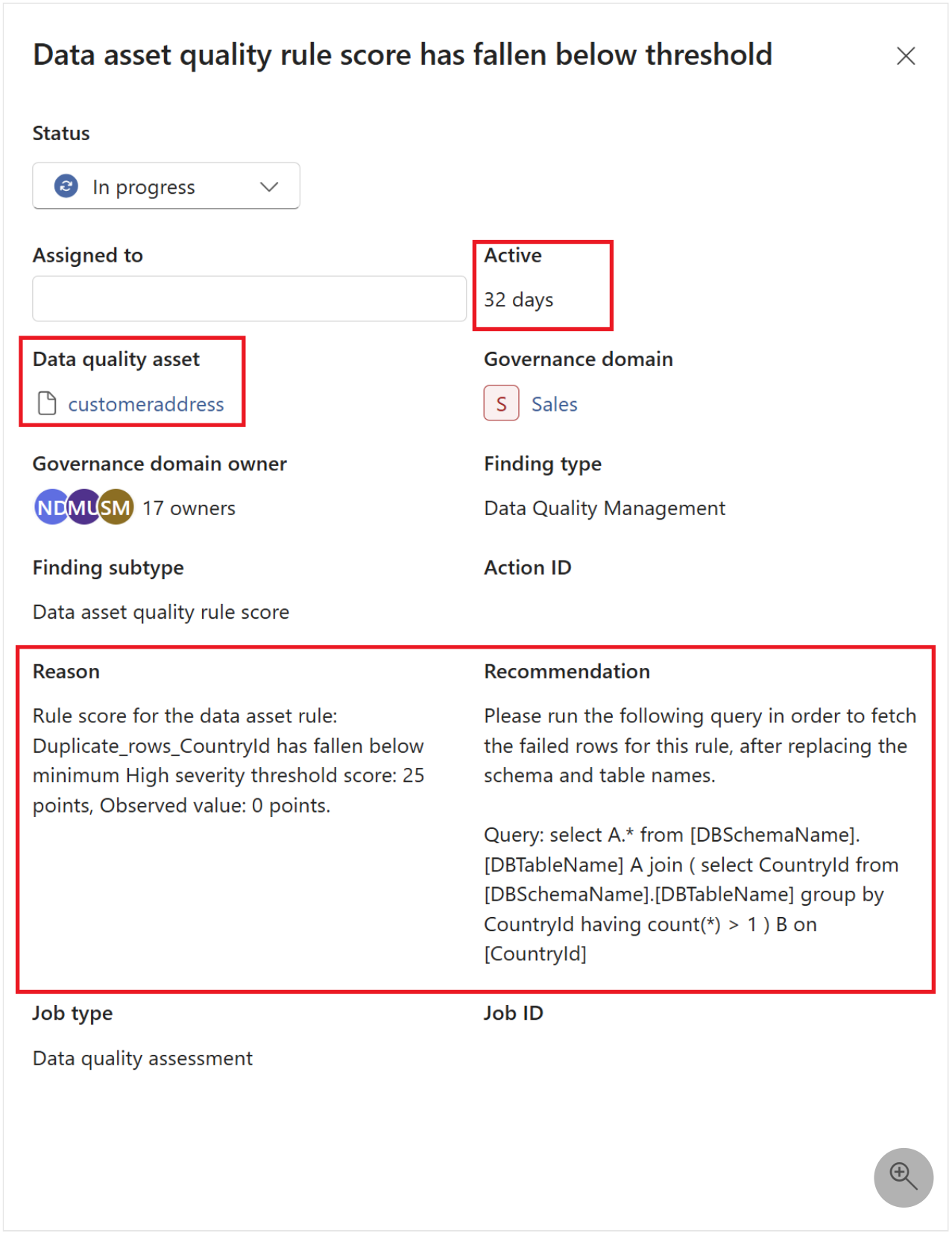
You can browse data quality actions for a data asset and assign them to users from the data quality overview page:

1. In Microsoft Purview Unified Catalog, select the **Health Management** menu and **Data quality** submenu.
2. Select a governance domain from the list.
3. Select a **data product** from the list, then select a **data asset** from the list of assets.
4. On the data asset's details page, select **View all** on the **Actions** tile to see all actions for the data asset.
5. Browse through the list of **Active**, **In progress**, **Resolved**, and **My items** actions.



6. Select an action to see its details and assign it to a user to investigate and resolve. an individual for remedy. The **Recommendation** section suggests the action to take and provides queries to find the low quality data.

> **Note**
>
> - Not every action has a SQL query because not all actions are data related or related to a data quality rule. Some actions are system related, such as a failed or skipped data quality job.
> - If a custom rule exception occurs, which means the data quality score generated from the custom rule doesn't meet the defined threshold, an action is created. However, a SQL query isn't suggested.



7. Update the **Status**, then select **Save**.

Resources

- Best practices for data quality.

Last updated on 11/13/2025

Set up data quality alerts

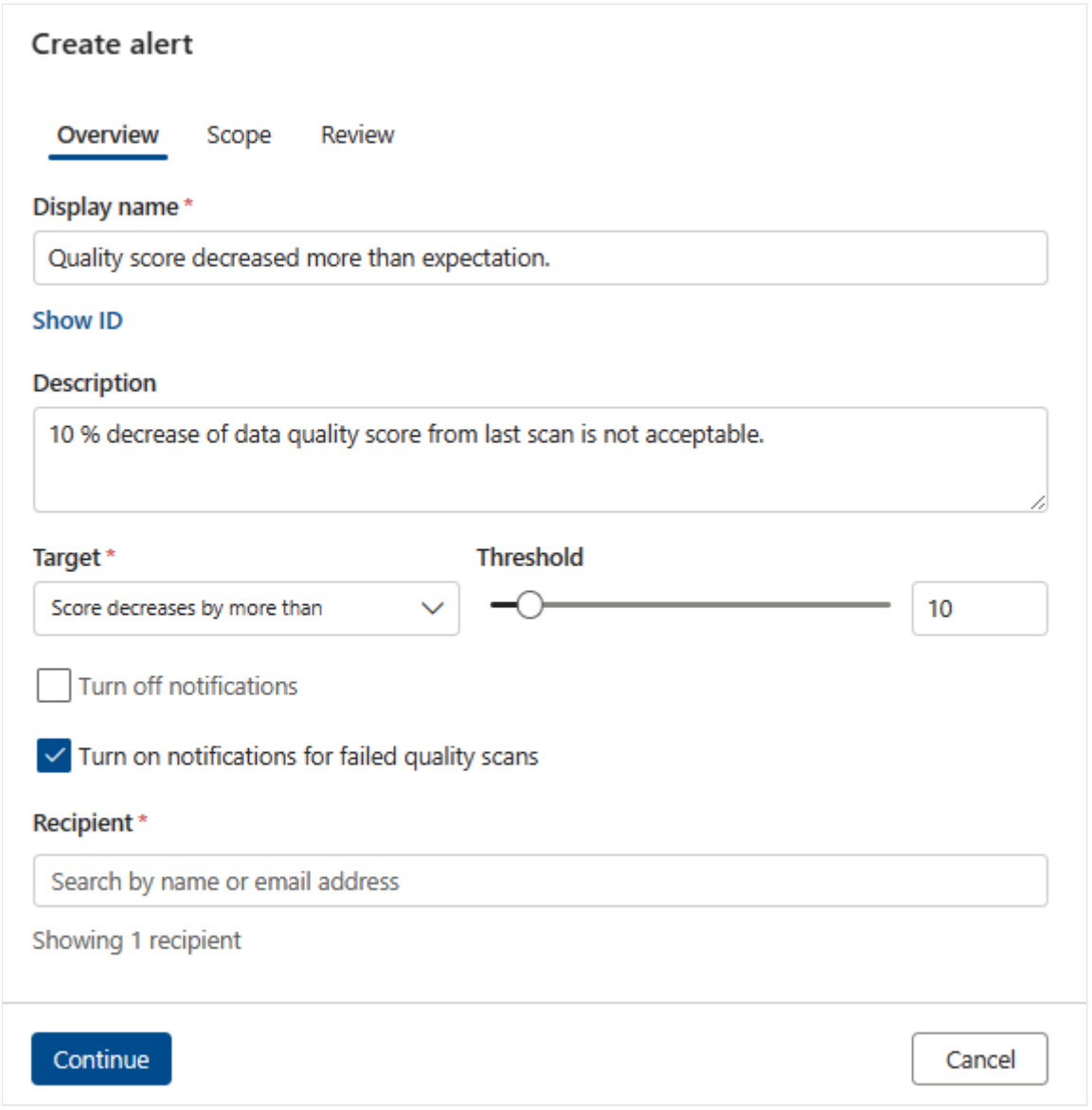
Data quality alerts notify you about important events or unexpected behavior detected around the quality of the data. When you create alerts for data assets, you receive email notifications about data quality scores.

Required roles

To create data quality alerts, you need the data quality steward role.

Set data quality alerts

1. In Microsoft Purview Unified Catalog, select **Health management**, then select **Data quality**.
2. On the **Data Quality** page, select a governance domain from the list.
3. Select **Manage**, then select **Alerts**.
4. Select **New** to set up a new alert for the data products and data assets in your selected governance domain.
5. On the **Overview** tab, enter a **Display name** and a **Description**.
6. For **Target**, select any of the following options:
    - **Score less than**: Alerts you if the assets' data quality score is less than certain percentage.
    - **Score decreased by more than**: Alerts you if the current quality score drops by certain percentage after a new scan is run.
7. Select the threshold that triggers an alert.
8. Select **Turn on notifications** for failed quality scan to enable alert notifications.
9. Add a **Recipient** to send the notification email to the recipient alias.



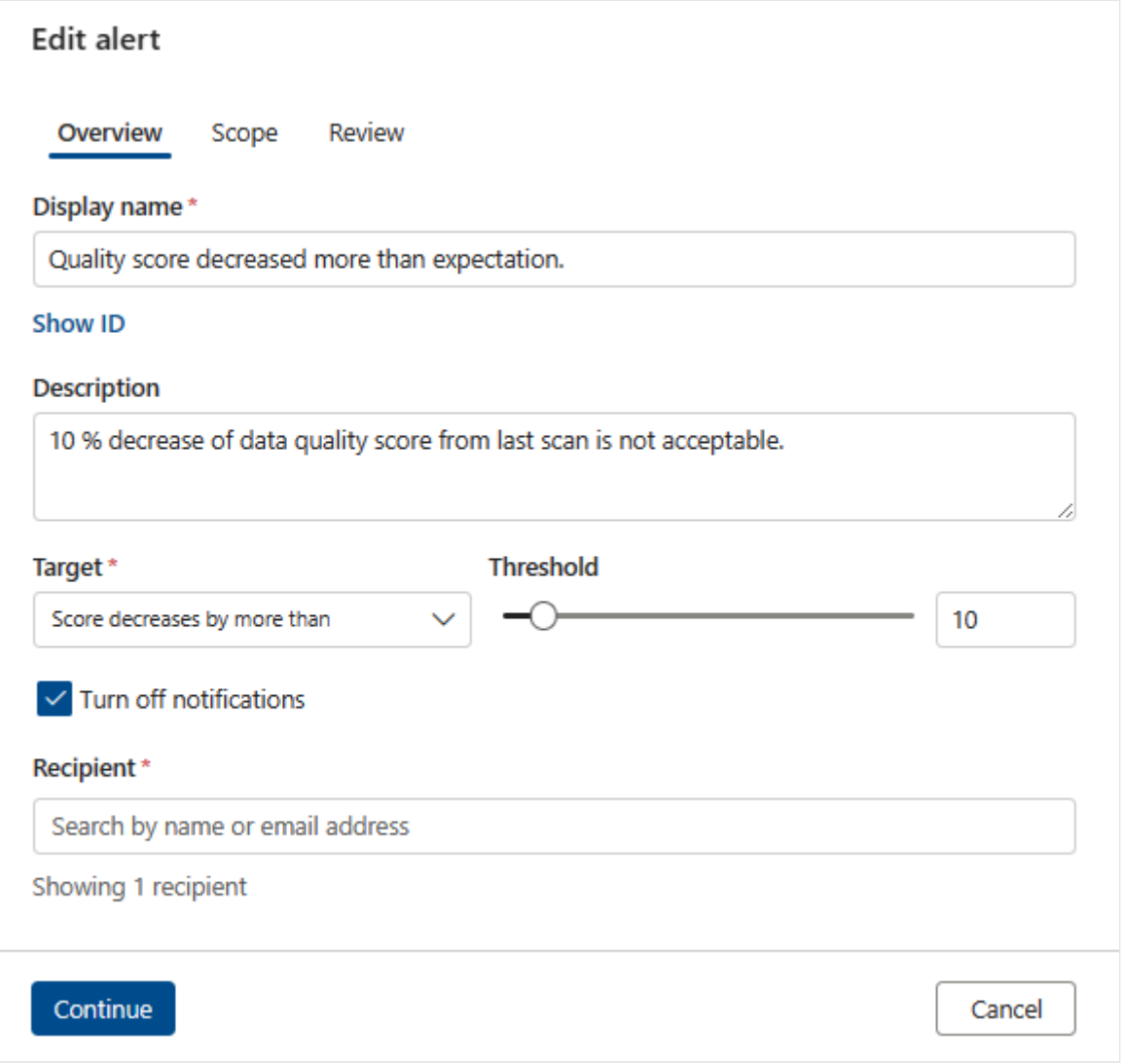
10. Select **Continue** to define the scope of alerts.
11. On the **Scope** tab, select data products and data assets that the alert will monitor, then select **Save**.
12. Select **Continue** to review.
13. Check alert details on the **Review** tab, then select **Submit** to create the alert.

Disable data quality alerts

Alerts are enabled by default. To disable an alert:

1. In Unified Catalog, select **Health management**, then select **Data quality**.
2. On the **Data Quality** page, select a governance domain from the list.
3. Select **Manage**, then select **Alerts**.

4. Select the **Notifications** toggle next to the alert, or select the alert and select the **Turn off notifications** checkbox.



> **Note**
>
> Disabling an alert stops the sending of notification emails related to the assets within the alert's scope.

Next steps

- Set up data quality rules based on the profiling results, and apply them to your data asset.
- Configure and run a data quality scan on a data product to assess the quality of all supported assets in the data product.
- Review your scan results to evaluate your data product's current data quality.

Last updated on 11/13/2025

Data quality troubleshooting

Summarize this article for me

This article describes common questions and symptoms of errors, along with possible resolutions:

Why do I see an invalid source error from the scanning job?

- You might see this error for two reasons:
  - The delta table doesn't exist in the location.
  - The data in the file aren't in a valid delta format.

I'm setting up data quality scans for my Microsoft Fabric delta tables. I see all data assets in the data quality view, I select one asset and apply rules for data quality scanning, but the scan fails.

- Your data quality scanning might fail for several reasons:
  - The data in tables aren't in the correct delta format. Make sure that your data are in delta format.
    - The Data Map scan didn't run successfully. Rerun the Data Map scan.
    - You need to delete any previous data quality runs for the data asset.

Why do I see the error message, "No connection can be used. Try to create connection first."?

- To profile data or to run data quality scanning, you first need to configure a data source connection. This alert disappears after you create a data source connection.

Why is the total count of profiled data less than the total count in my Microsoft Azure Data Lake Storage Gen2 delta table?

- Microsoft Purview Data Quality uses a **1 million** sample size for profiling. This sample is taken **randomly**. If your delta table has more than 1 million records, the total count doesn't match.

Why do I see an action about data quality score missing for a data product, but I see the score in the data product when I browsed the data product view?

- When you created the action, there wasn't any data quality score for that data product. Data quality scanning ran after you created the action and the score published for the data product. Close the action once the remediation is done to avoid confusion.

Data quality rule creation from "Suggest rules" throws an error about a "date" column when trying to add all 30 suggested rules

- This error occurs because the schema data type is **unsupported** in the data quality schema view. You can change the **data type** to **date** by selecting the **schema management** toggle and saving it. After you change the data type, you should be able to add the rule.

When trying to add all suggested rules, why do I see the error "ObserverId already exists"?

- Most likely, the same or identical rule is already added to a column. When you try to add the same or identical rule to a column, the application throws this error message.

Why is my scheduled job skipping instead of running? I see Skipped for data quality scanning jobs.

- The data quality job checks and runs data quality only if there are changes since the last run. It checks the delta history. **Skipped** means there are no changes in the data since last run, and the spark run for data quality isn't performed. **Skipped** doesn't mean failed.

When I select profile data tab, I see number of columns preselected. Can I change the selected columns?

- Microsoft Purview Data Quality uses an AI assisted profiling solution. The Microsoft Purview Data Profiling AI selects the preselected columns. You can deselect those columns and reselect based on criticality of the columns, and select **Save and run** to run profiling.

Why can't I select some of the data assets from data quality asset list page to profile and scan?

- There are a few reasons:
  - Those data assets are published from unsupported data sources.
  - The file format of those data assets isn't supported.

Why is my profiling job failing for the supported data sources?

- Check the schema to make sure that there's no column name with spaces. The current version doesn't support column names with spaces.

Why can't I run data quality scanning and data profiling for CSV, TSV, and text files?

- Data Quality currently supports the Delta format of Parquet, Delta, Iceberg ORC, and Iceberg AVRO. Purview Data Quality doesn't support CSV, TSV, and text files.

Why don't I see the data quality freshness rule in the rule list?

- Microsoft Azure SQL tables don't support data quality freshness. If your data asset is an Azure SQL table, the freshness rule doesn't appear in the list for selection and application to the data asset.

Why do I see datatype Undefined for some columns of a data asset schema in a data quality schema page?

- It seems the data type for all columns isn't identified correctly. You can import the schema to resolve the issue (to update the datatype). Select the schema menu item from the data quality overview page, select the **Schema management** toggle, and select **Import schema**. After importing the schema, select the schema management toggle again to save the updated schema.

- Microsoft Azure SQL tables don't support data quality freshness. If your data asset is an Azure SQL table, the freshness rule doesn't appear in the list for selection and application to the data asset.

My data quality scan job failed. I see an error message 'Internal service error occurred, please retry, or contact Microsoft support.' What should I do to troubleshoot?

- Many reasons can cause the scan to fail with this error message:
  - You're not authorized to perform the current operation for the workspace that you're trying to access for the data quality scan.
  - Error code 403, meaning access to data sources is forbidden temporarily.
  - Granted access to the data source for your managed identity (MSI) expired.
    - Microsoft Purview managed identity (MSI) needs contributor access to the Fabric workspace. If you don't provide the contributor access for the Microsoft Purview

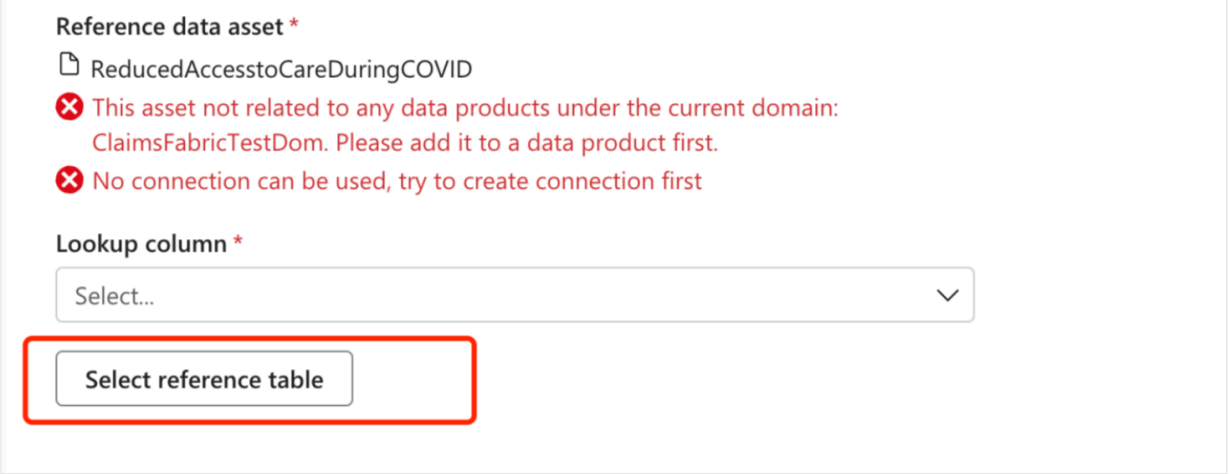
MSI to the Fabric workspace, the data quality scan fails.

Why am I getting delta format error even though I'm using delta format?

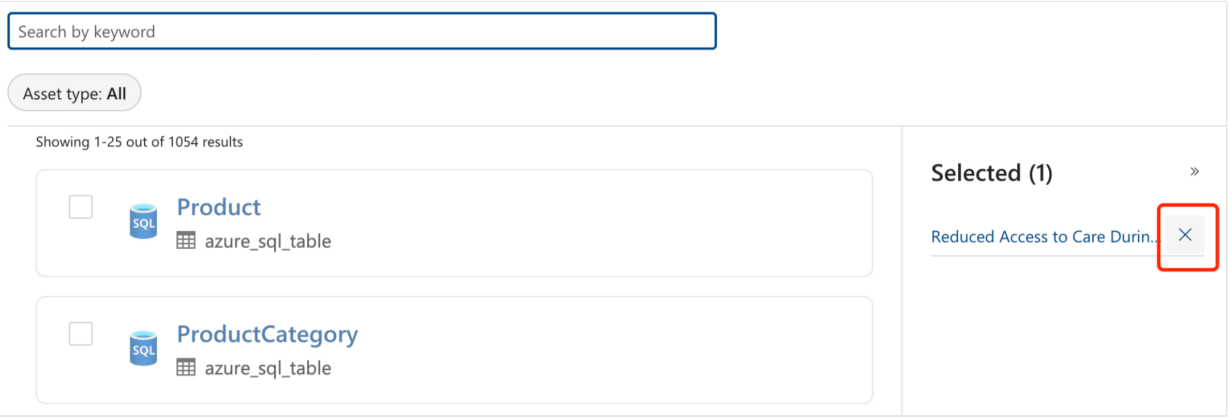
- We support Spark 3.4 Delta 2.4. Make sure that you're using delta lake version 2.4.

Why am I seeing the error when I selected a reference data asset to configure Table lookup rule?

- You selected a data asset that isn't part linked or referred to a data product under the same governance domain. To select the right data asset:
  - Select **select reference table**, as indicated in this image:



  - Cancel current selection, as shown in the following image:



  - After canceling the current selection, select another asset.

How can I configure access to data source for Microsoft Purview MSI?

- Here's MSI configuration guide. You find the details in this document.

All our data sources are behind the private end point (in virtual network), Can Microsoft Purview access data in virtual network for data quality scanning?

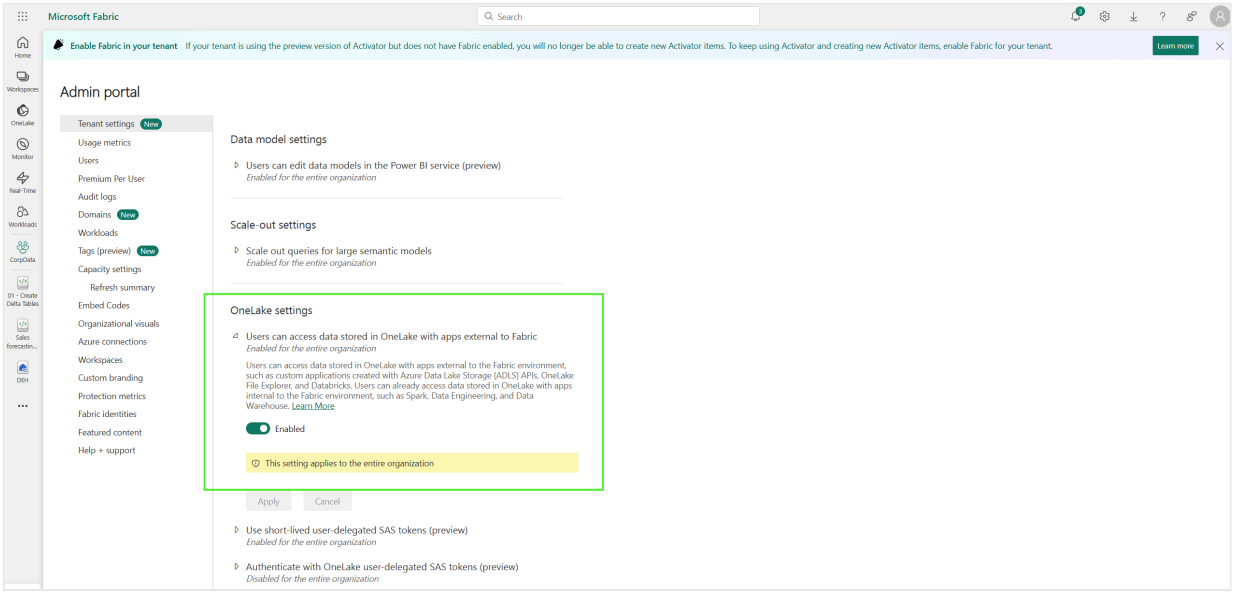
- Yes, Microsoft Purview supports Managed Virtual Network for data quality scanning. See the Microsoft Purview managed virtual network configuration article.

Where can I find documentation about expression function to create custom rules?

- Find information and examples at Create data quality rules.

Why is my data quality scan for Fabric Lakehouse table failing?

- Many reasons can cause the failure:
  - Make sure that your lakehouse tables are discoverable in Data Map with schema.
  - Make sure that you're using SPN for Data Map scan and MSI for data quality scan.
    - Make sure that you configured the data quality connection with MSI.
    - Make sure that Microsoft Purview MSI has contributor access to your Fabric workspace.
    - Enable Microsoft OneLake setting: Users can access data stored in OneLake with apps external to Fabric.



  - Learn how to configure data quality for Fabric Lakehouse.
  - Learn how to configure a Data Map scan setup for Fabric.

Why can't users enable managed virtual network (VNET) for data quality configuration on Microsoft Purview for Databricks or Synapse, or any other supported data sources?

- **Problem**: After selecting the **Enable managed Vnet** checkbox and receiving approval, no changes occur on the Microsoft Purview side. Subsequent attempts to create a connection fail with no request sent to the sources (for example, Azure Databricks, Synpase, Data Lake Storage Gen2) for approval.
- **Resolution**: This problem occurs when you delete the Managed Private Endpoint (MPE) after configuring it. Currently, if you update or delete the MPE directly from the source after initial approval, the Data Quality system can't track the MPE status change. To resolve this problem, contact Microsoft support.

Schema is blank or columns are missing

- Import the schema from the **Data Quality** page to ensure that all columns, including any missing ones, are properly loaded and to resolve the issue. After importing the schema, perform a hard refresh to avoid reloading cached or outdated information.
- Depending on the extent of the schema changes, you might need to remove the asset from the data asset data quality page or the data product page and then re-add it from the Data Map to fully resolve the issue.
- Rule bindings use the technical column name as the key, not just the DisplayName. If the column’s name changed, or the asset’s schema cached in data quality still references the old technical name, subsequent runs fail because the rule can’t find the column.
- Data product assets don’t auto refresh when you update metadata in Data Map (owner, description, lineage, and relevant here, schema). You need to manually refresh the asset inside catalog management so that data quality sees the latest column set.

Last updated on 02/24/2026

Overview of health management reports in Unified Catalog

Data health management reports in Microsoft Purview Unified Catalog enable your users to explore detailed information about your data health management and create customized reports for your organization.

Report types

Classic reports

- Classic assets: An overview of assets by type and collection, and their curation status.
- Classic catalog adoption: Helps you see at a glance how the classic Microsoft Purview Data Catalog is being used.
- Classic classifications: An overview of assets classified and the types of classifications.
- Classic data stewardship: Helps governance and quality focused users, like data stewards and chief data officers, understand governance health gaps like asset curation and asset ownership.
- Classic glossary: An overview of the health and use of glossary terms.
- Classic sensitivity labels: An overview of assets that have sensitivity labels applied and the types of labels applied.

> **Note**
>
> Classic reports refresh weekly. You can't change the schedule. Classic reports don't have any impact to the data governance billing model. When you consent to using Unified Catalog, the billing model starts, but classic reports have no charge.

Data governance reports

- Data governance (preview): Helps you understand your data estate governance and performance of data governance controls.
  - The data governance (preview) report depends on data health controls. If you don't use data health controls, the report isn't refreshed. Make sure that you set up the

schedule to run data health controls and refresh the health controls scores and data governance report.

- Data quality health: Includes assessments of various data quality dimensions, and metrics to help stakeholders understand the accuracy, completeness, consistency, reliability, and freshness of your data estate.
  - The data quality health report depends on data health controls and the Microsoft Purview metadata self-serve analytics model. If you don't use data health controls and don't subscribe Unified Catalog metadata, the data quality health report isn't refreshed. You either need to use data health controls or need to subscribe Unified Catalog metadata for self-serve analytics to refresh the data quality health report.
  - If you don't use the data quality feature, the data quality health report shows as blank because the report is created using data generated from data quality scans.

> **Note**
>
> - Data governance (preview) and Data quality health reports are part of the **data governance billing model**. If you don't want to use these reports, then you need to disable **self-serve analytics metadata subscription**, and you also need to disable the **health controls refresh schedule**. For details, see **self-serve analytics and insights**.

Last updated on 11/14/2025

Understand the classic assets report in Unified Catalog

This article describes how to access, view, and filter data asset insight reports for your data.

In this article, you learn how to:

- ✓ View data estate insights from your Microsoft Purview account.
- ✓ Get a bird's eye view of your data.
- ✓ Drill down for more asset count details.

Prerequisites

Before getting started with Microsoft Purview Data Estate Insights, make sure that you've completed the following steps:

- Set up a storage resource and populated the account with data.
- Set up and completed a scan your storage source. For more information to create and complete a scan, see Manage data sources in Microsoft Purview.
- Permissions depending on your current experience:
  - If using the classic Microsoft Purview Catalog, you need the data curator role or insight reader role.
  - If using the new Microsoft Purview Unified Catalog, you need data health reader permissions.

Understand your asset inventory in Data Estate Insights

In Microsoft Purview Data Estate Insights, you can get an overview of the assets that you scanned into the Data Map. You can also view key gaps that governance stakeholders can close for better governance of the data estate.

1. Open the Microsoft Purview portal.

2. If you're using the classic Microsoft Purview portal, select **Data estate insights**. If you're using the new Microsoft Purview portal open **Unified Catalog**, select **Health Management**, then select **Reports**.

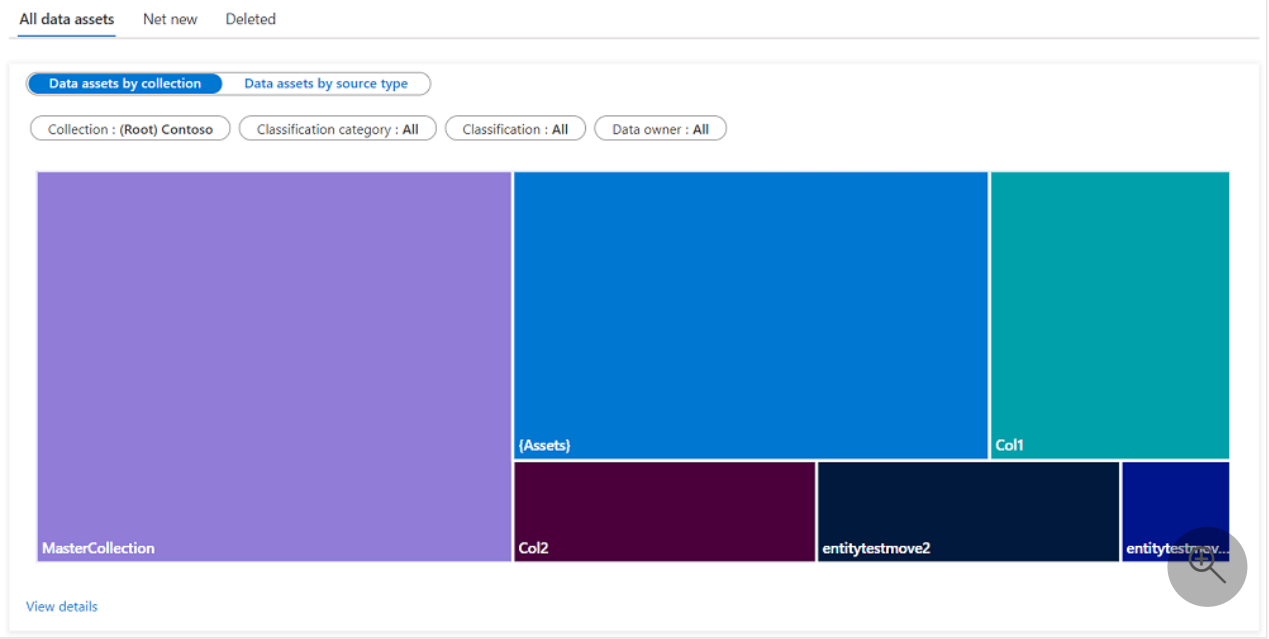
3. Select the **Classic assets** report.

4. The **Assets Summary** report provides several high-level KPIs, with these graphs:

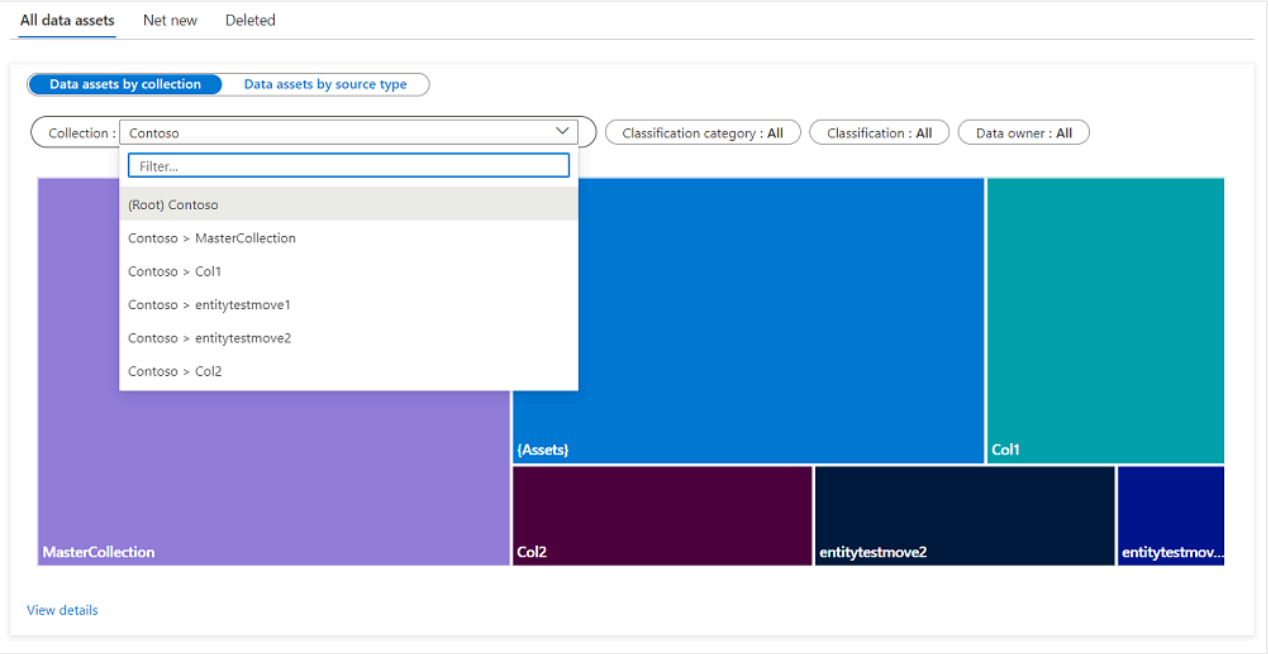
- **Unclassified assets**: Assets with no system or custom classification on the entity or its columns.
- **Unassigned data owner**: Assets that have the owner attribute within "Contacts" tab as blank.
- **Net new assets in last 30 days**: Assets that you added to the Purview account, via data scan or Atlas API pushes.
- **Deleted assets in last 30 days**: Assets that you deleted from the Purview account, as a result of deletion from data sources.



5. Below these KPIs, you can also view your data asset distribution by collection.

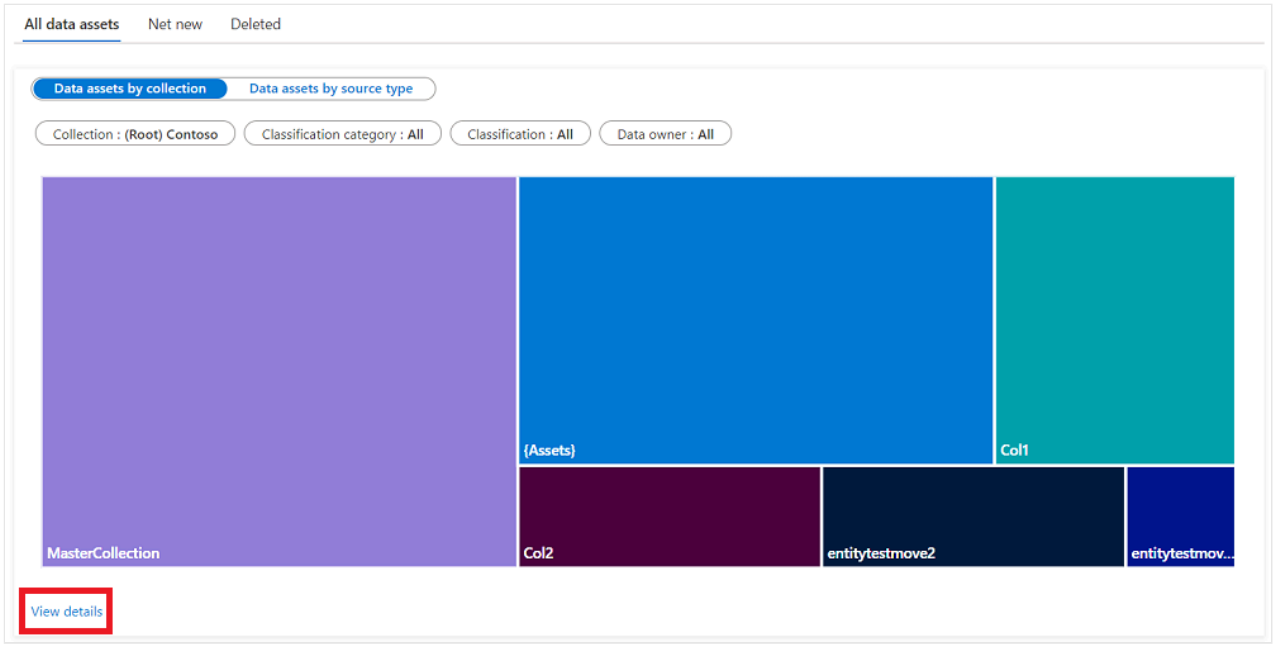


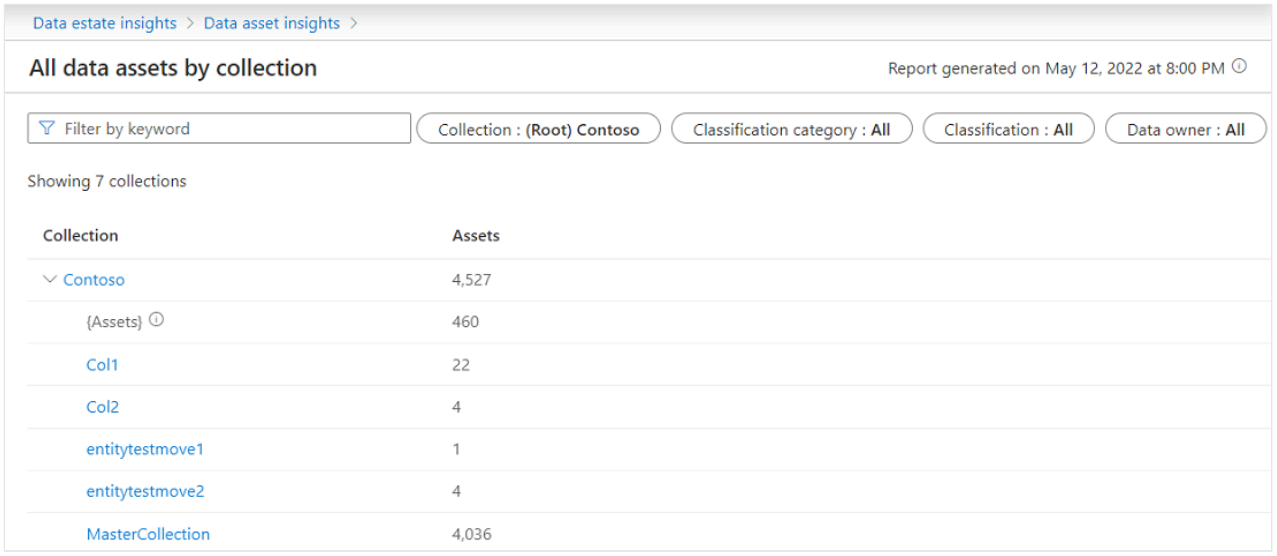
6. Using filters you can drill down to assets within a specific collection or classification category.



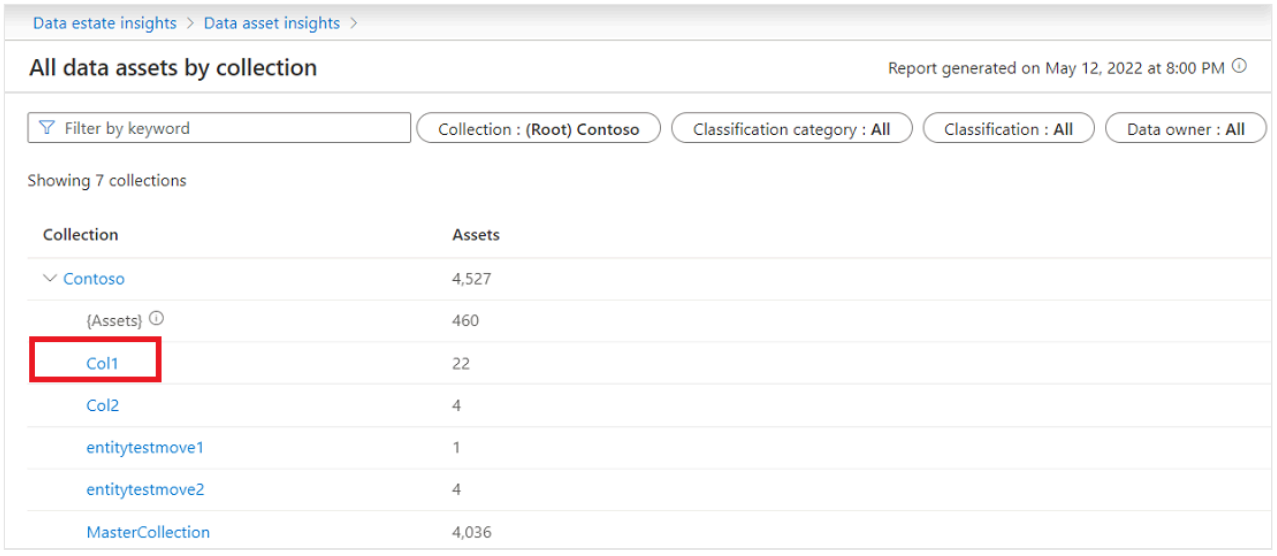
> ⓘ Note
>
> ***Each classification filter has some common values:***
>
> - **Applied**: Any filter value is applied
> - **Not Applied**: No filter value is applied. For example, if you pick a classification filter with value as "Not Applied", the graph shows all assets with no classification.
> - **All**: Filter values are cleared. Meaning the graph shows all assets, with or without classification.
> - **Specific**: You pick a specific classification from the filter, and only that classification is shown.

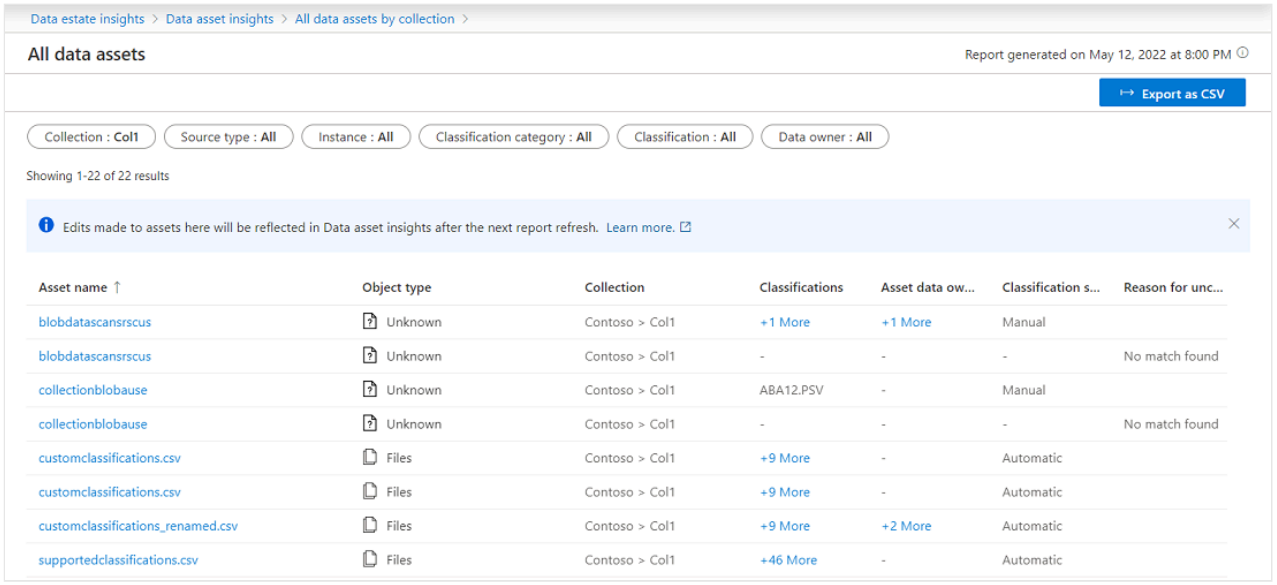
7. To learn more about which specific assets are shown in the graph, select **View details**.



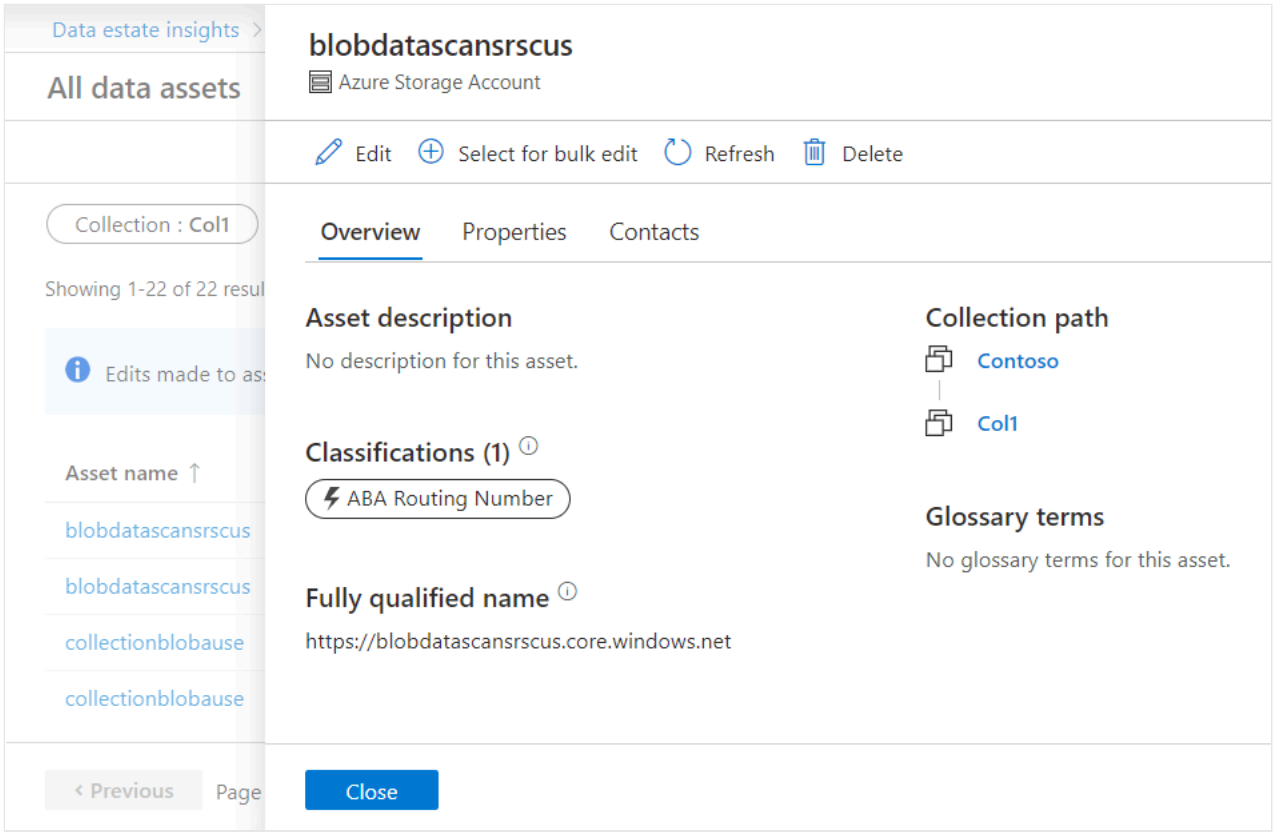


8. You can select any collection to view the collection's asset list.





9. You can also select an asset to edit without leaving the Data Estate Insights App.



File-based source types

The next graphs in asset insights show a distribution of file-based source types. The first graph, **Size trend (GB) of file type within source types**, shows top file type size trends over the last 30 days.

1. Pick your source type to view the file type within the source.

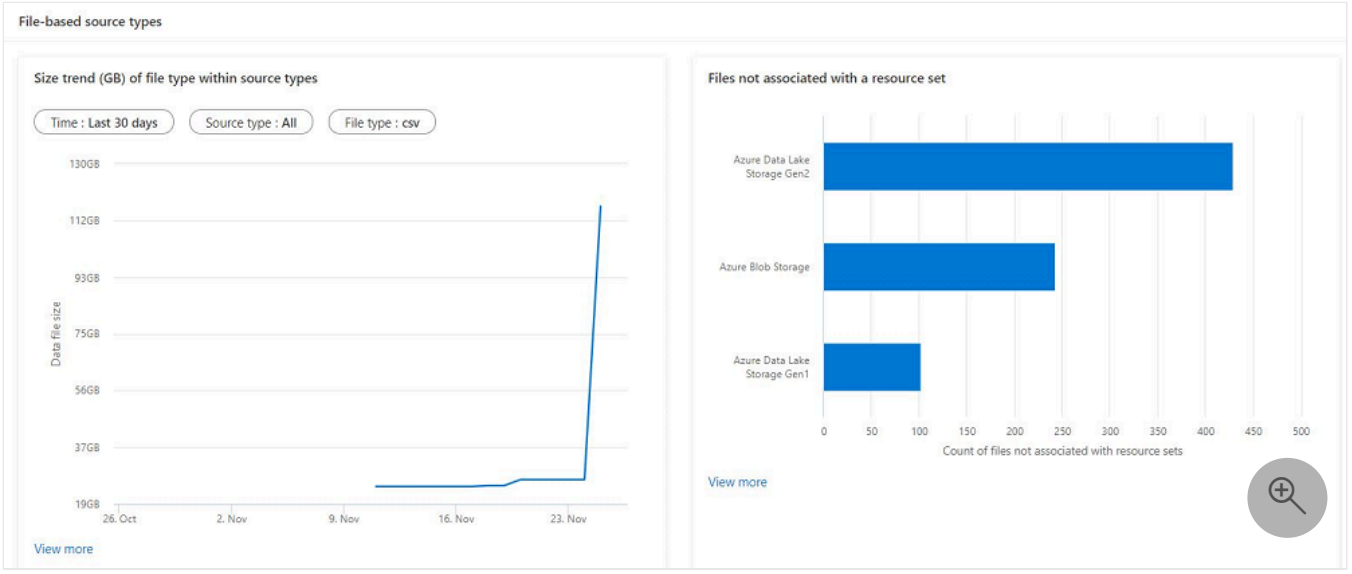
2. Select **View details** to see the current data size, change in size, current asset count, and change in asset count.

> **Note**
>
> If the scan runs only once in the last 30 days, or any catalog change such as classification addition or removal happens only once in 30 days, then the change information appears blank.

3. See the top folders with change top asset counts when you select source type.

4. Select the path to see the asset list.

The second graph in file-based source types is **Files not associated with a resource set**. If you expect that all files should roll up into a resource set, this graph can help you understand which assets aren't rolled up. Missing assets can be an indication of the wrong file-pattern in the folder. You can select **View details** below the graph for more information.



Next steps

Learn how to use Data Estate Insights with the following resources:

- Learn how to use data stewardship insights.
- Learn how to use classification insights.
- Learn how to use glossary insights.
- Learn how to use label insights.

Last updated on 11/14/2025

Understand the classic catalog adoption report in Unified Catalog

As described in the insights concepts, the catalog adoption report is part of the "Health" section of classic Microsoft Purview Data Estate Insights. This report offers a one-stop shop experience for administrators to determine if and how the classic Microsoft Purview Data Catalog is being used. It helps answer questions like:

- What are my users searching for?
- How many people used the catalog last month?
- What are the most used data assets?

Prerequisites

Before getting started with Data Estate Insights, make sure that you complete the following steps:

- Set up a storage resource and populate the account with data.
- Set up and complete a scan of your storage source. For more information about creating and completing a scan, see manage data sources in Microsoft Purview.
- Have the right permissions depending on your current experience:
  - If using the classic Microsoft Purview Data Catalog, you need the data curator role or insight reader role.
  - If using the new Microsoft Purview Unified Catalog, you need data health reader permissions.

Understand your data estate and catalog health in Data Estate Insights

In Data Estate Insights, you can get an overview of all assets inventoried in the Data Map, and any key gaps that governance stakeholders can close for better governance of the data estate.

1. Open the Microsoft Purview portal.

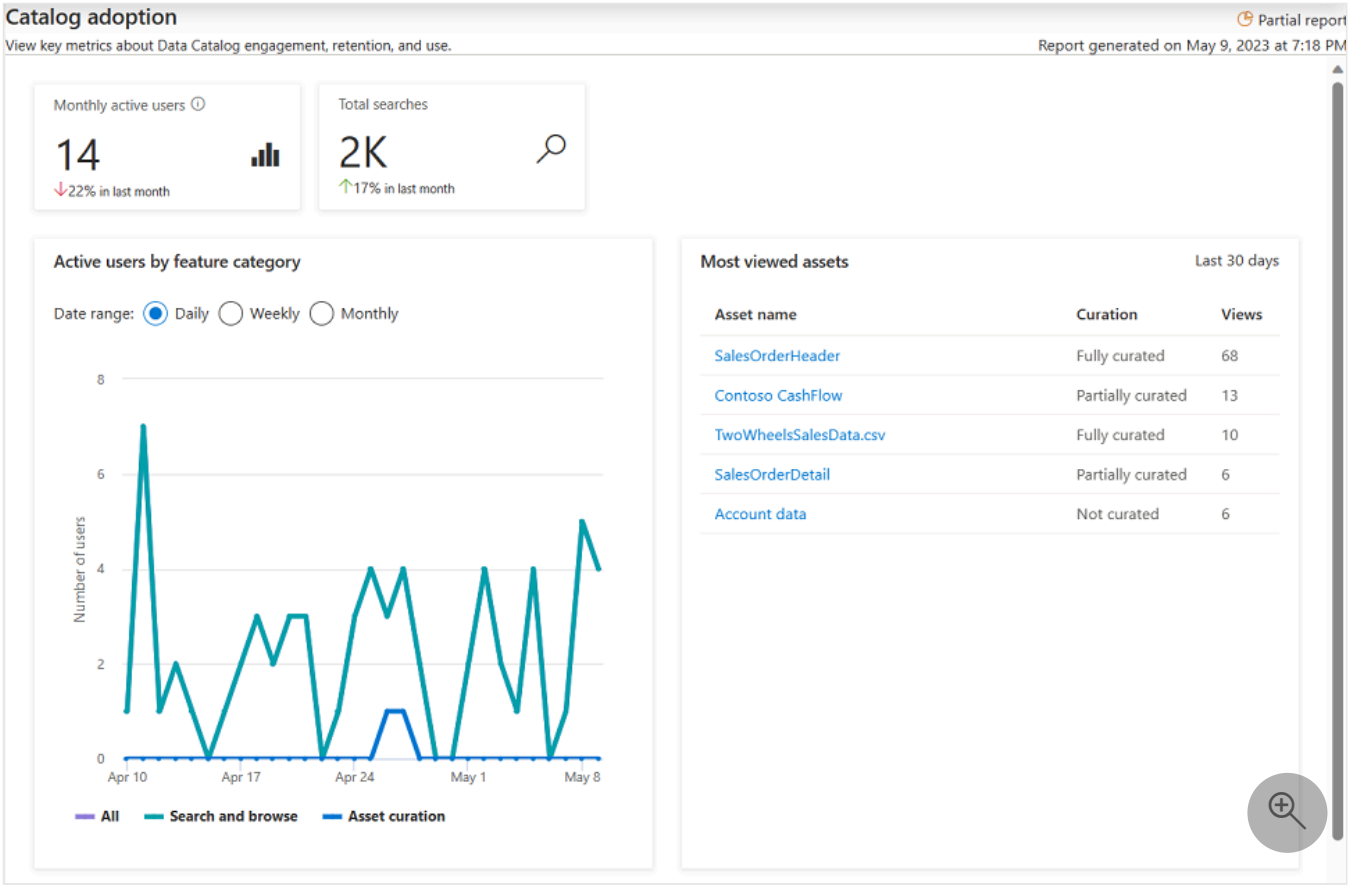
2. If you're using the classic Microsoft Purview portal, select **Data estate insights**. If you're using the new Microsoft Purview portal open **Unified Catalog**, select **Health Management**, then select **Reports**.

3. Select **Classic catalog adoption**.

View catalog adoption dashboard

The catalog adoption dashboard has several curated tiles and charts to identify:

- How many active users your catalog had in the last month.
- How many total searches were performed in the last month.
- Which catalog features are being used.
- Most viewed assets.
- Top searched keywords in catalog.



Monthly active users

The **Monthly active users** count is the number of unique users who perform any action, such as search, browse, or curation.

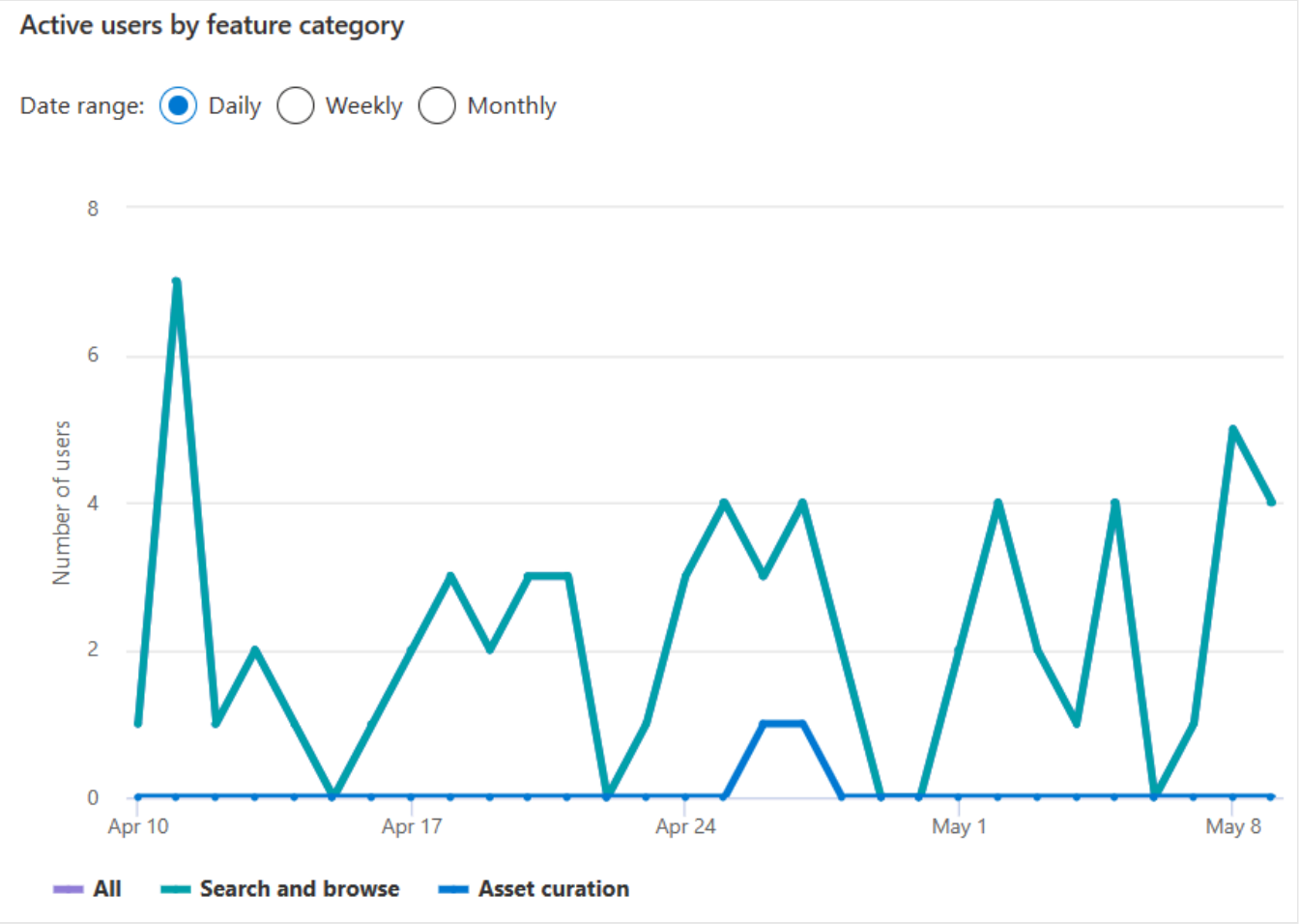
The **Monthly active users** tile shows how many users took at least one action in the catalog in the last 30 days. An action in the catalog includes searching for a term, browsing in the catalog, and updating an asset in the catalog. This tile also includes an indicator of a percentage increase or decrease in users from the previous month.

Total searches

The **Total searches** tile shows a count of all searches performed in the catalog over the last 30 days. The tile also shows the percentage increase or decrease in the number of searches from the previous month.

Active users by feature category

The **Active users by feature category** chart helps you monitor user activity trends in the catalog. Here's an example of the chart:



At the top of the chart, select your date range to view user activity on a daily, weekly, or monthly basis.

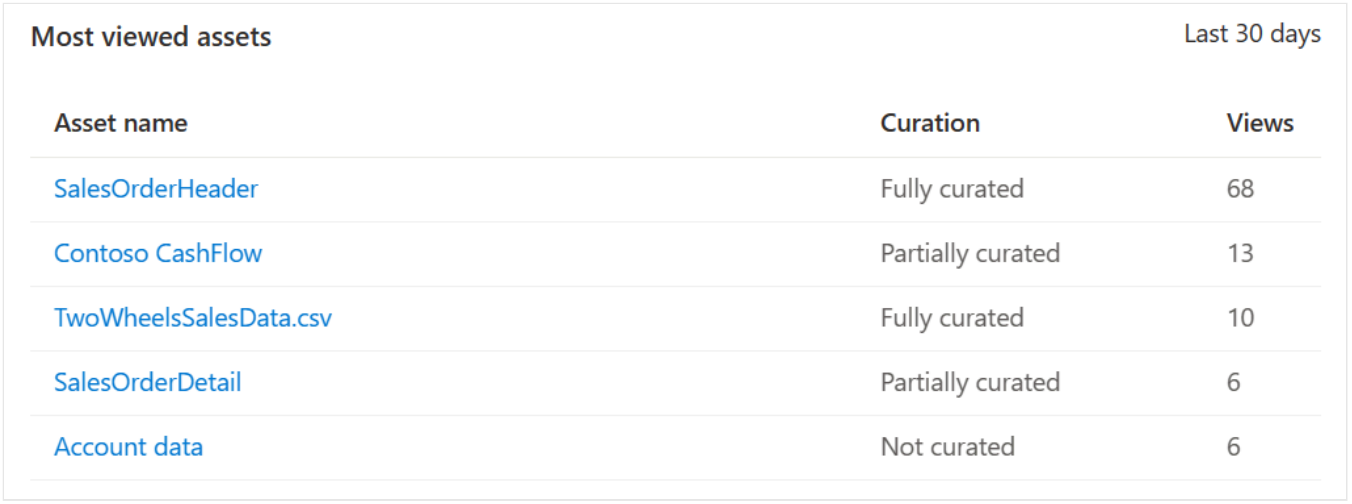
The chart shows the number of users in the date range who perform actions in one of three categories:

- **Search and browse**: The count of users per action category (search or browse). A user may appear in both counts if they perform both actions. For example, if a user performs both search and browse, the same user is counted twice in this number (once under search, once under browse). One user can contribute two counts in Search/Browse, but still counts as one in Monthly Active Users. This is why **Search and browse** can be higher than **Monthly active users**.

- **Asset curation**: The count of users who performed at least any one of these actions on an asset:
  - Added, removed, or updated a rating
  - Added or removed a tag
  - Added or removed a glossary term
  - Added or removed a classification
  - Edited an asset's name or description
  - Added or removed a certification
  - Added or removed a column level classification, glossary term, description, name, or data type
  - Added or removed manual lineage
  - Added or removed a contact
- **All**: The count of users who performed both a search or browse action, *and* at least one asset curation action within the date range.

Most viewed assets

The most viewed assets table shows the top assets in the catalog, by sum of views in the last 30 days.

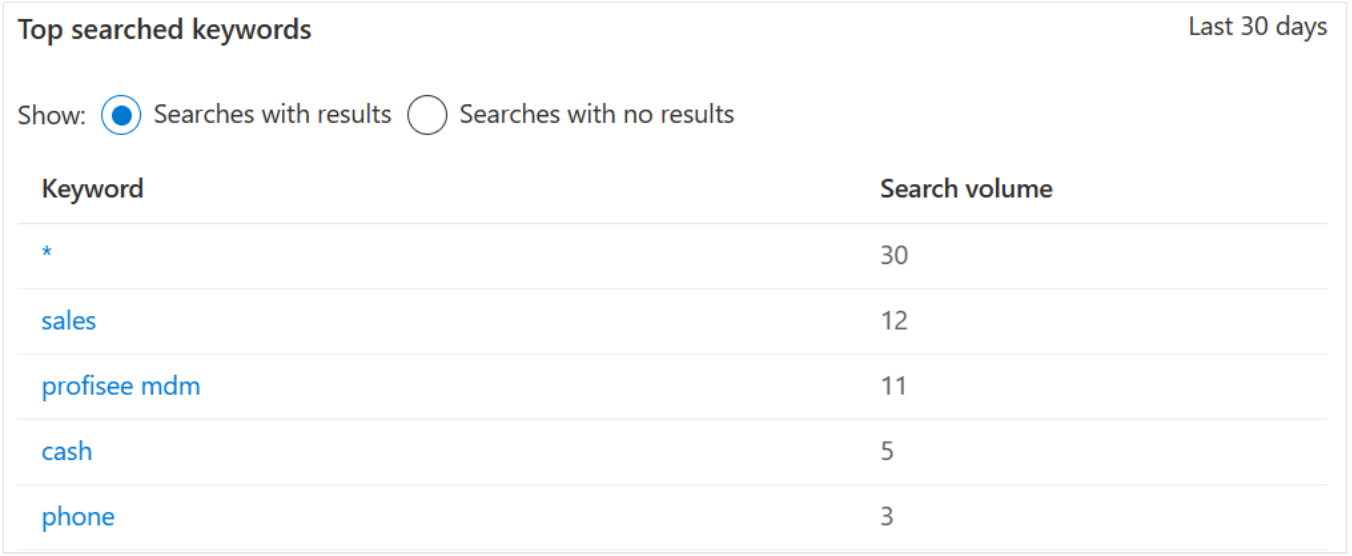


This table allows you to select through to your most viewed assets for more information, but the table also provides these other details:

- Curation status - There are three possible statuses: **Fully curated**, **Partially curated**, and **Not curated**, based on certain attributes of assets being present. An asset is **Fully curated** if it has at least one classification tag, an assigned Data Owner, and a description. If any of these attributes is missing, but not all, then the asset is categorized as **Partially curated** and if all of them are missing, then it's **Not curated**. For more information about the curation status of your data assets, see the data stewardship dashboard.
- Views - count of the number of views the asset received in the last 30 days.

Top searched keywords

The top searched keywords table shows your top keywords both for searches that produced results, and searches that didn't. That way you can know what users are finding with the catalog, and what they're still looking for.



At the top of the table, you can select one of the two radio buttons to select whether to show keywords for searches with results, or searches without results.

You can select any of the keywords to run the search in the catalog and see the results for yourself as well.

The table also provides search volume, which is the number of times that keyword was searched in the last 30 days.

Report generated timestamp

The timestamp displayed, such as, "Report generated on December 4, 2025 at 11:37 AM", indicates when the page was loaded or refreshed. It doesn't represent the time when the backend actually computed the KPI data. The actual calculation time comes from the backend timePoint field (for example, 2025-11-05T06:59:48) from the API response.

Monthly aggregation behavior for usage metrics

The **Monthly active users** and **Total searches** metrics display data from the previous fully completed month. For example, December views show November data. The KPI is updated during the job run after the month changes, even though the job runs every week.

Next steps

Learn more about Data Estate Insights:

- Data Estate Insights Concepts
- Data stewardship insights

Last updated on 01/15/2026

Understand the classic classifications report in Unified Catalog

This article describes how to access, view, and filter the classic classifications report for your data.

In this article, you learn how to:

- ✓ Launch your Microsoft Purview account from Microsoft Azure.
- ✓ View classification insights on your data.
- ✓ Drill down for more classification details on your data.

Prerequisites

Before getting started with Data Estate Insights, make sure you complete the following steps:

- Set up a storage resource and populate the account with data.
- Set up and complete a scan on the data in each data source. For more information, see Manage data sources in Microsoft Purview and Create a scan rule set.
- Permissions depending on your current experience:
  - If using the classic Microsoft Purview Catalog, you need the data curator role or insight reader role.
  - If using the new Microsoft Purview Unified Catalog, you need data health reader permissions.

Use Data Estate Insights for classifications

In Microsoft Purview, classifications are similar to subject tags, and are used to mark and identify data of a specific type that's found within your data estate during scanning.

Microsoft Purview uses the same sensitive information types as Microsoft 365, allowing you to stretch your existing security policies and protection across your entire data estate.

> ⓘ **Note**
>
> After you scan your source types, wait a couple of hours for the classic classifications report to reflect the new assets.

View classification insights

1. Open the Microsoft Purview portal.
2. If you're using the classic Microsoft Purview portal, select **Data estate insights**. If you're using the new Microsoft Purview portal open **Unified Catalog**, select **Health management**, then select **Reports**.
3. Select the **Classic classifications** report. The **Classification insights** page displays the following areas:

Expand table

Area	Description
Overview of sources with classifications	Displays tiles that provide: - The number of subscriptions found in your data - The number of unique classifications found in your data - The number of classified sources found - The number of classified files found - The number of classified tables found
Top sources with classified data (last 30 days)	Shows the trend, over the past 30 days, of the number of sources found with classified data.
Top classification categories by sources	Shows the number of sources found by classification category, such as **Financial** or **Government**.
Top classifications for files	Shows the top classifications applied to files in your data, such as credit card numbers or national/regional identification numbers.
Top classifications for tables	Shows the top classifications applied to tables in your data, such as personal identifying information.
Classification activity (files and tables)	Displays separate graphs for files and tables, each showing the number of files or tables classified over the selected timeframe. **Default**: 30 days Select the **Time** filter above the graphs to select a different time frame to display.

Classification insights drilldown

In any of the following **Classification insights** graphs, select the **View details** link to drill down for more details:

- **Top classification categories by sources**
- **Top classifications for files**
- **Top classifications for tables**

- **Classification activity > Classification data**

To learn more, use the following options:

Expand table

Option	Description
Filter your data	Use the filters above the grid to filter the data shown, including the classification name, subscription name, or source type. If you're not sure of the exact classification name, enter part or all of the name in the **Filter by keyword** box.
Sort the grid	Select a column header to sort the grid by that column.
Edit columns	To display more or fewer columns in your grid, select **Edit Columns**, then select the columns you want to view or change the order.
Drill down further	To drill down to a specific classification, select a name in the **Classification** column to view the **Classification by source** report. This report displays data for the selected classification, including the source name, source type, subscription ID, and the numbers of classified files and tables.
Browse assets	To browse through the assets found with a specific classification or source, select a classification or source, depending on the report you're viewing, then select **Browse assets** above the filters. The search results display all of the classified assets found for the selected filter. For more information, see Search the Microsoft Purview Unified Catalog.

Next steps

Learn how to use Data Estate Insights with the following resources:

- Learn how to use Asset insights.
- Learn how to use Data Stewardship.
- Learn how to use Classification insights.
- Learn how to use Glossary insights.
- Learn how to use Label insights.

Last updated on 11/14/2025

Understand the classic data stewardship report in Unified Catalog

As described in the insights concepts, the data stewardship report is part of the "Health" section of Data Estate Insights. This report offers a one-stop shop experience for data, governance, and quality focused users like chief data officers and data stewards to get actionable insights into key areas of gap in their data estate.

In this guide, you learn how to:

- ✓ Navigate and view data stewardship report from your Microsoft Purview account.
- ✓ Drill down for more asset count details.

Prerequisites

Before getting started with Data Estate Insights, make sure that you complete the following steps:

- Set up a storage resource and populate the account with data.
- Set up and complete a scan your storage source. For more information on how to create and complete a scan, see Manage data sources in Microsoft Purview.
- Permissions depending on your current experience:
  - If using the classic Microsoft Purview Catalog, you need the data curator role or insight reader role.
  - If using the new Microsoft Purview Unified Catalog, you need data health reader permissions.

Understand your data estate and catalog health in Data Estate Insights

In Data Estate Insights, you can get an overview of all assets inventoried in Data Map, and any key gaps that governance stakeholders can close for better governance of the data estate.

1. Open the Microsoft Purview portal.

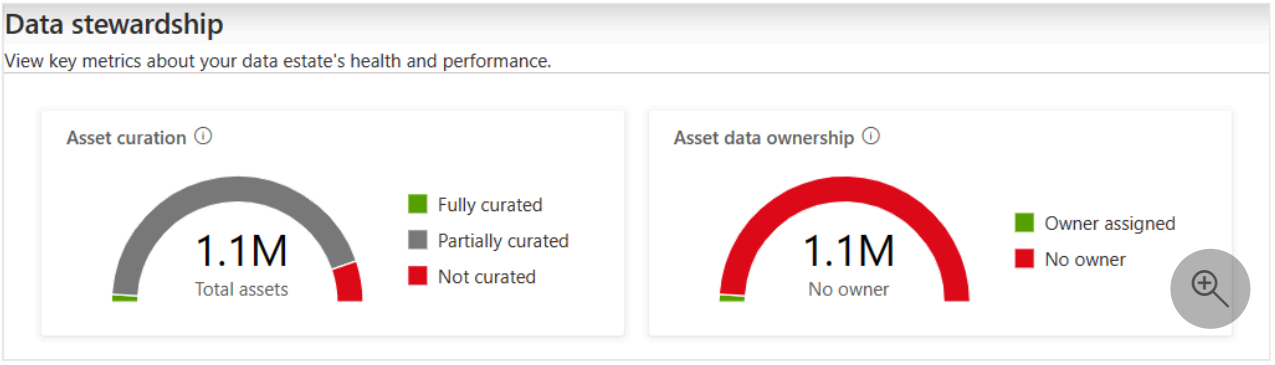
2. If you're using the classic Microsoft Purview portal, select **Data estate insights**. If you're using the new Microsoft Purview portal open **Unified Catalog**, select **Health Management**, then select **Reports**.

3. Select **Classic data stewardship**.

View data stewardship dashboard

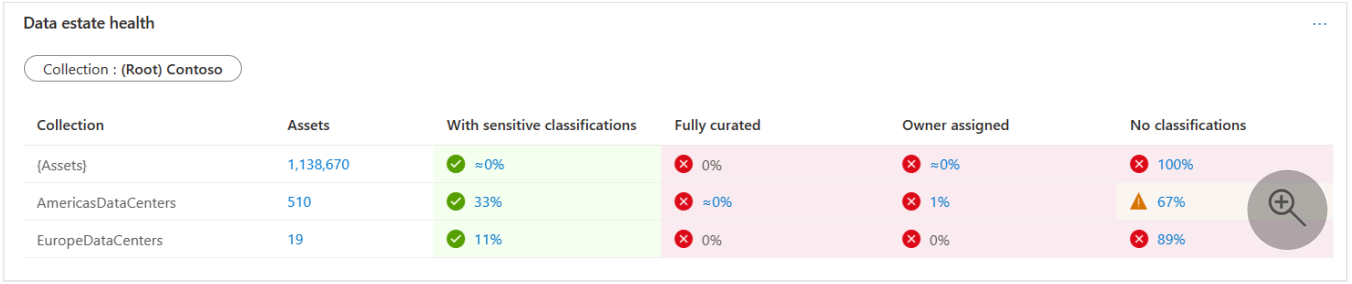
The dashboard is purpose-built for governance and quality-focused users, like data stewards and chief data officers, to understand the health management of their organization. The dashboard shows high-level KPIs that help reduce governance risks:

- **Asset curation**: The dashboard categorizes all data assets into three areas: "Fully curated," "Partially curated," and "Not curated," based on certain attributes of assets being present. An asset is "Fully curated" if it has at least one classification tag, an assigned data owner, and a description. If any of these attributes is missing, but not all, then the asset is categorized as "Partially curated" and if all of them are missing, then it's "Not curated."
- **Asset data ownership**: The dashboard categorizes assets that have the owner attribute within the "Contacts" tab as blank as "No owner". Otherwise, it categorizes them as "Owner assigned".



Health management

Health management is a scorecard view that helps management and governance focused users, like chief data officers, understand critical governance metrics that can be looked at by collection hierarchy.

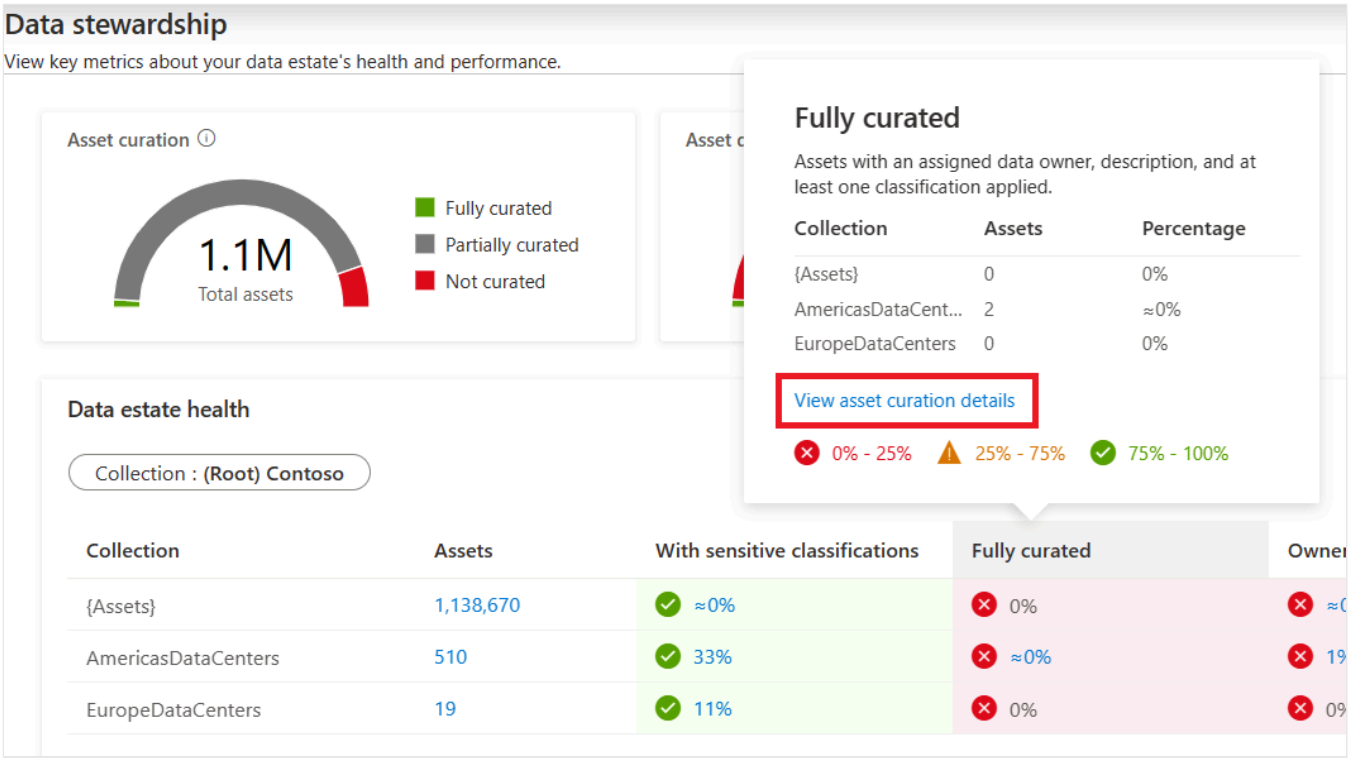


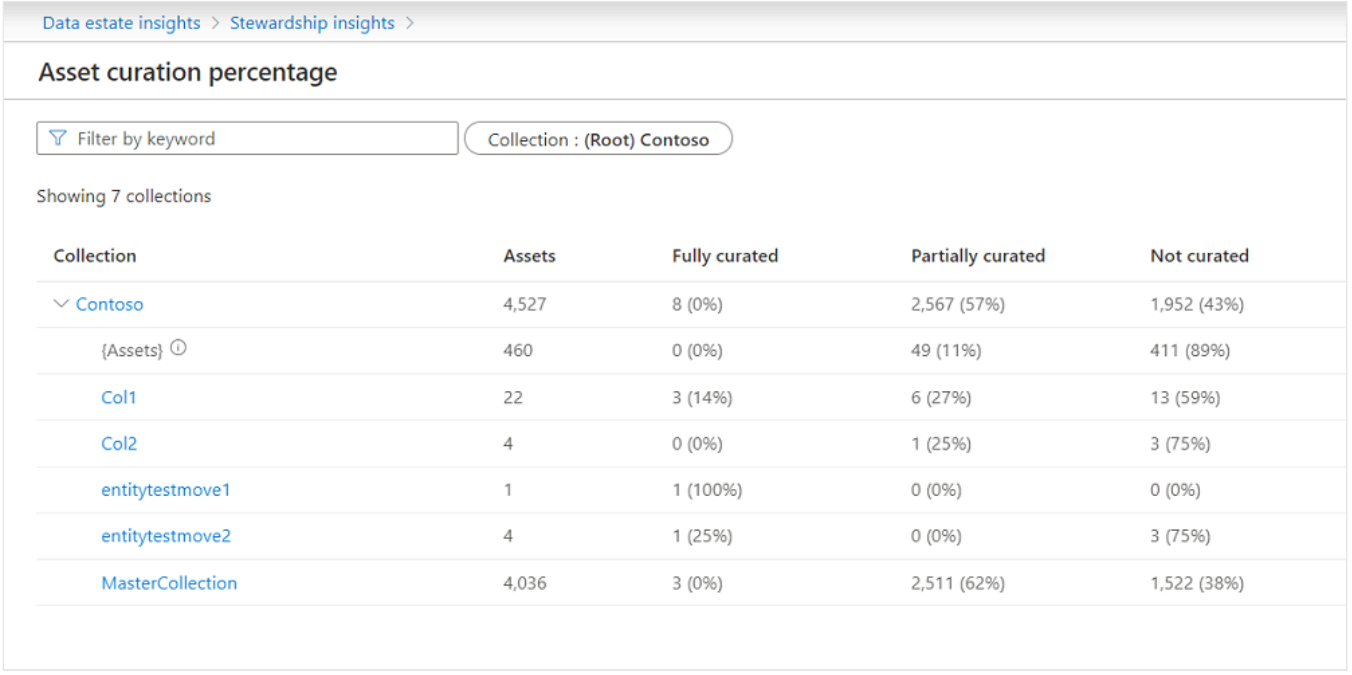
You can view the following metrics:

- **Assets**: Count of assets by collection drill-down.
- **With sensitive classifications**: Count of assets with any system classification applied.
- **Fully curated assets**: Count of assets that have a data owner, at least one classification, and a description.

- **Owner assigned**: Count of assets with data owner assigned on them.
- **No classifications**: Count of assets with no classification tag.
- **Out of date**: Percentage of assets that haven't been updated in over 365 days.
- **New**: Count of new assets pushed in the Data Map in the last 30 days.
- **Updated**: Count of assets updated in the Data Map in the last 30 days.
- **Deleted**: Count of deleted assets from the Data Map in the last 30 days.

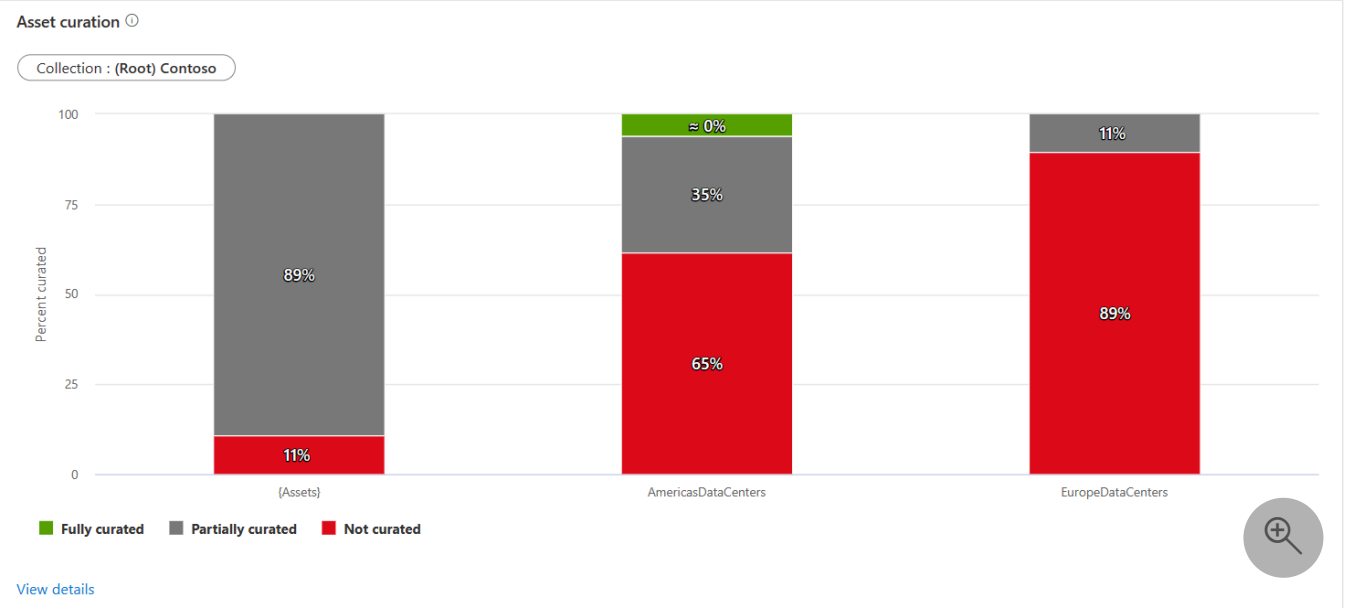
You can also drill down by collection paths. As you hover on each column name, it provides description of the column, provides recommended percentage ranges, and takes you to the detailed graph for further drill-down.





Asset curation

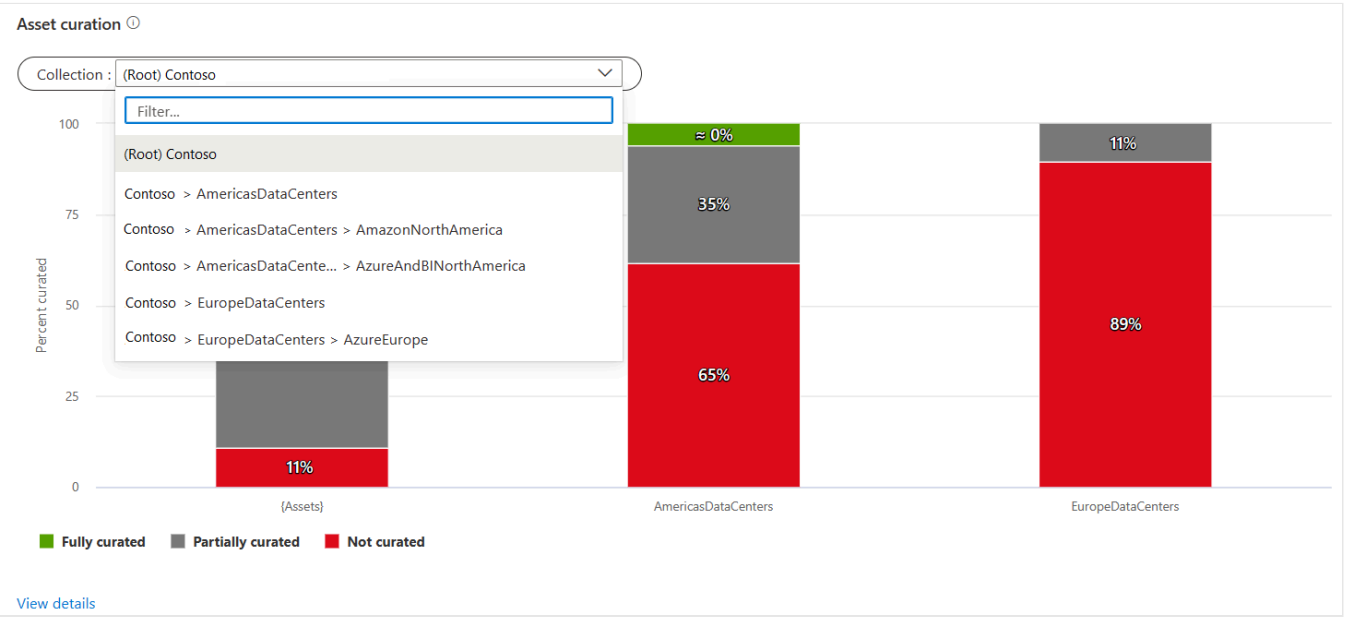
The system categorizes all data assets into three areas: **Fully curated**, **Partially curated**, and **Not curated**, based on whether assets have certain attributes.



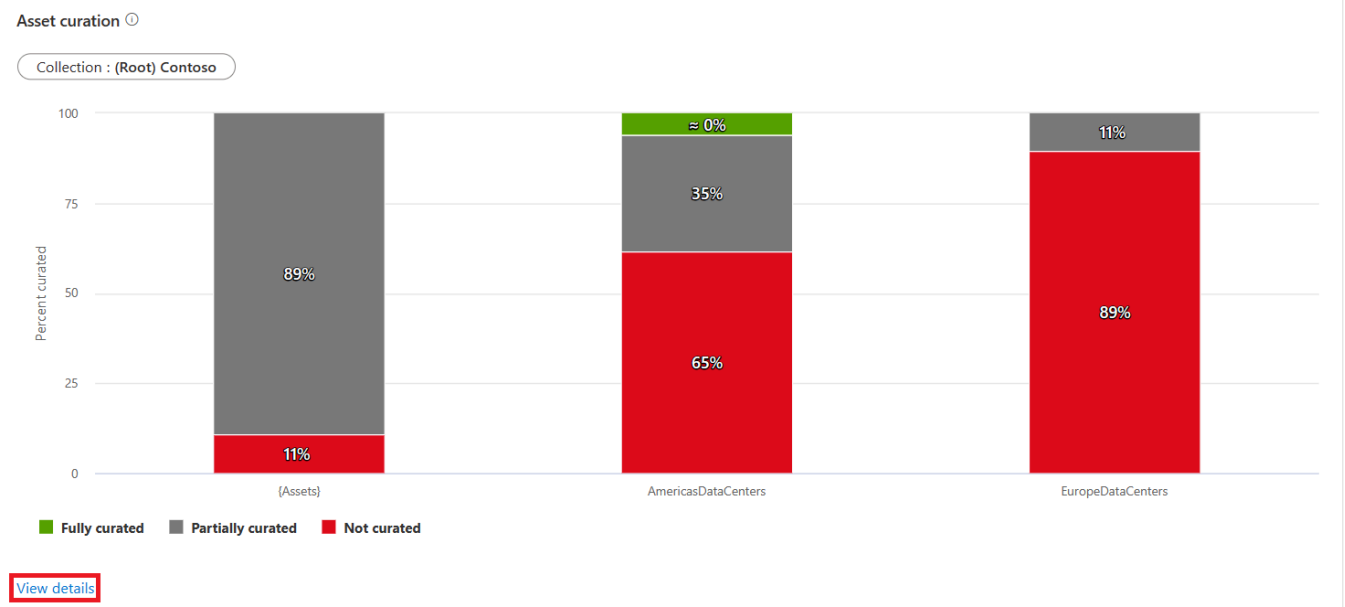
An asset is ***Fully curated*** if it has at least one classification tag, an assigned data owner, and a description.

If an asset is missing any of these attributes, but not all, the system categorizes it as **Partially curated**. If an asset is missing all of these attributes, the system lists it as **Not curated**.

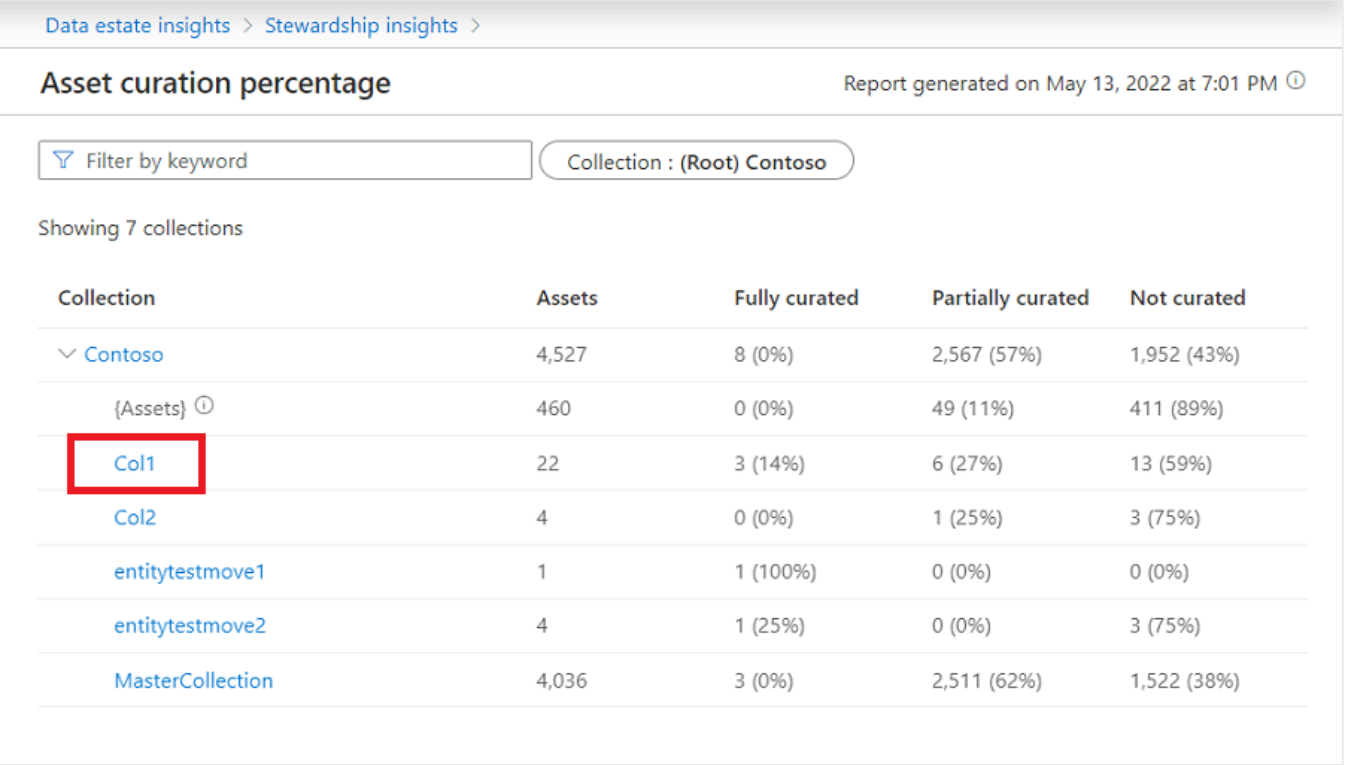
You can drill down by collection hierarchy.



For further information about which assets aren't fully curated, you can select **View details** link that takes you into the deeper view.



In the **View details** page, if you select a specific collection, it lists all assets with attribute values or blanks that make up the ***fully curated*** assets.



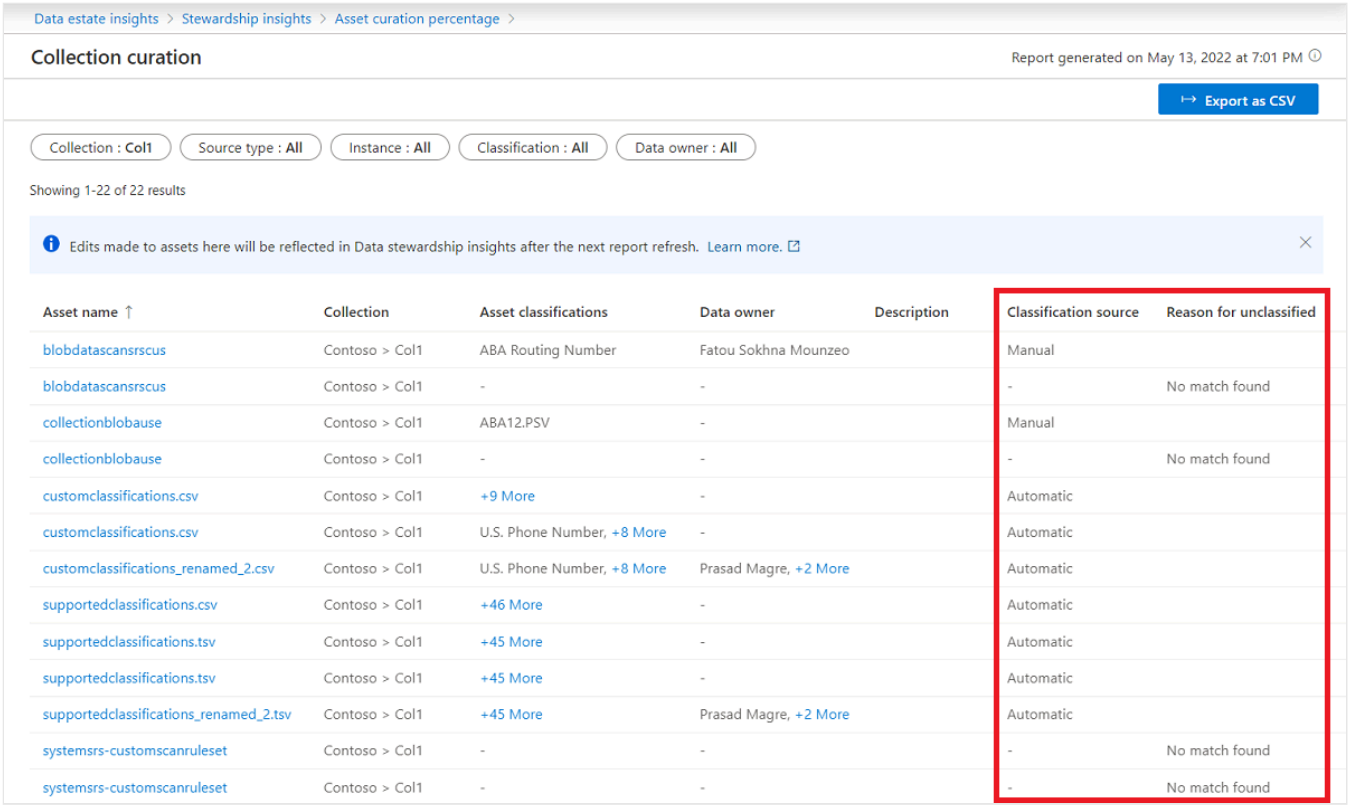
The detail page shows two key pieces of information:

First, it tells you what the ***classification source*** is, if the asset is classified. It's **Manual** if a curator or data owner applies the classification manually. It's **Automatic** if the system classifies the asset during scanning. This page only provides the last applied classification state.

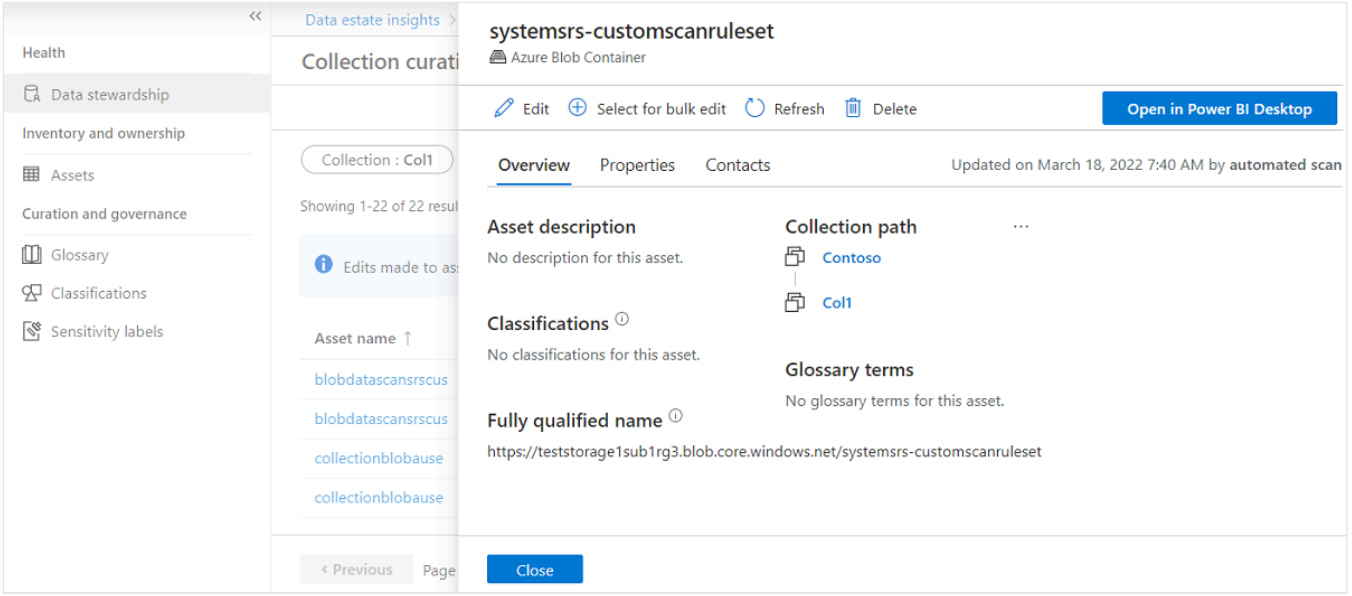
Second, if an asset is unclassified, it tells us why it's not classified, in the column ***Reasons for unclassified***. Currently, Data Estate Insights can provide one of the following reasons:

- No match found

- Low confidence score
- Not applicable



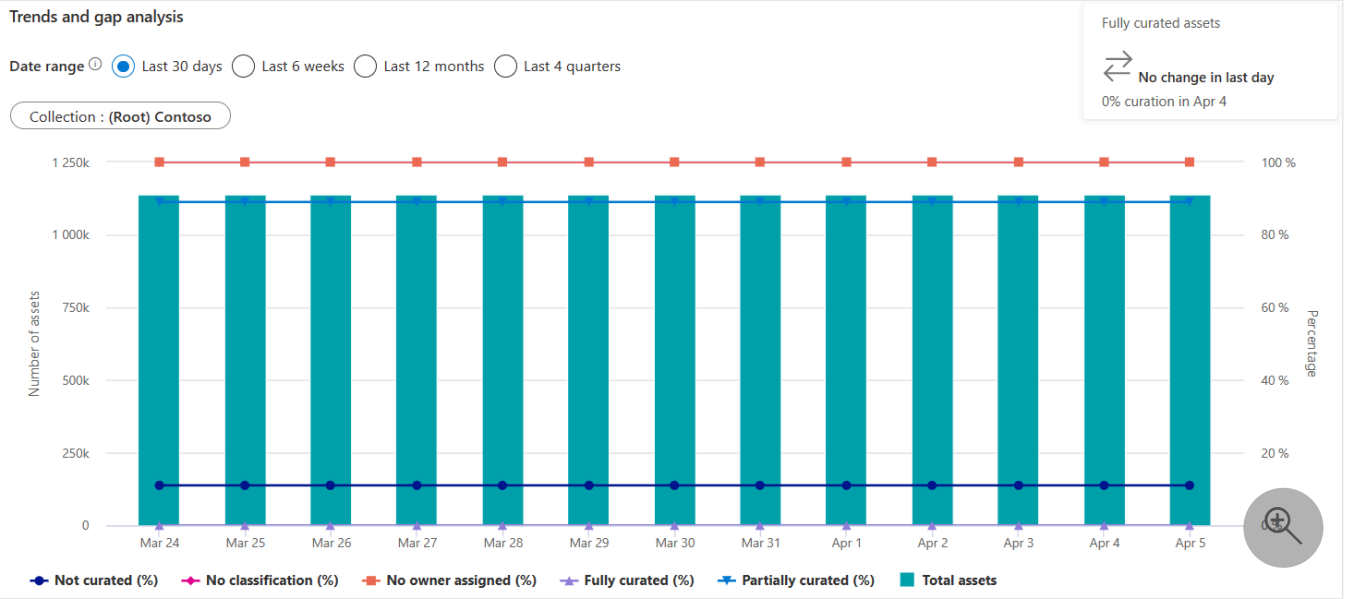
You can select any asset and add missing attributes, without leaving the **Data estate insights** app.



Trends and gap analysis

This graph shows how the assets and key metrics have been trending over different time periods:

- Last 30 days: The graph uses the last run of the day or the recording of the last run across days as a data point.
- Last six weeks: The graph uses the last run of the week, with the week ending on Sunday. If there's no run on Sunday, the graph uses the last recorded run.
- Last 12 months: The graph uses the last run of the month.
- Last four quarters: The graph uses the last run of the calendar quarter.



Next steps

Learn more about Classic catalog adoption insights.

Last updated on 11/14/2025

Understand the classic glossary report in Unified Catalog

This article describes how to access, view, and filter Microsoft Purview glossary insight reports for your data. In this how-to guide, you learn how to:

- ✓ Find Data Estate Insights.
- ✓ Get a bird's eye view of your data.

Prerequisites

Before getting started with Data Estate Insights glossary insights, make sure that you complete the following steps:

- Set up a storage resource and populate the account with data.
- Set up and complete a scan your storage source.
- Set up at least one business glossary term and attach it to an asset.
- Permissions depending on your current experience:
  - If using the classic Microsoft Purview Catalog, you need the data curator role or insight reader role.
  - If using the new Microsoft Purview Unified Catalog, you need data health reader permissions.

Use Microsoft Purview glossary insights

In Microsoft Purview, you can create glossary terms and attach them to assets. As you make use of these terms in your data map, you can view the glossary distribution in glossary insights. These insights give you the state of your glossary based on:

- Number of terms attached to assets.
- Status of terms.
- Distribution of roles by users.

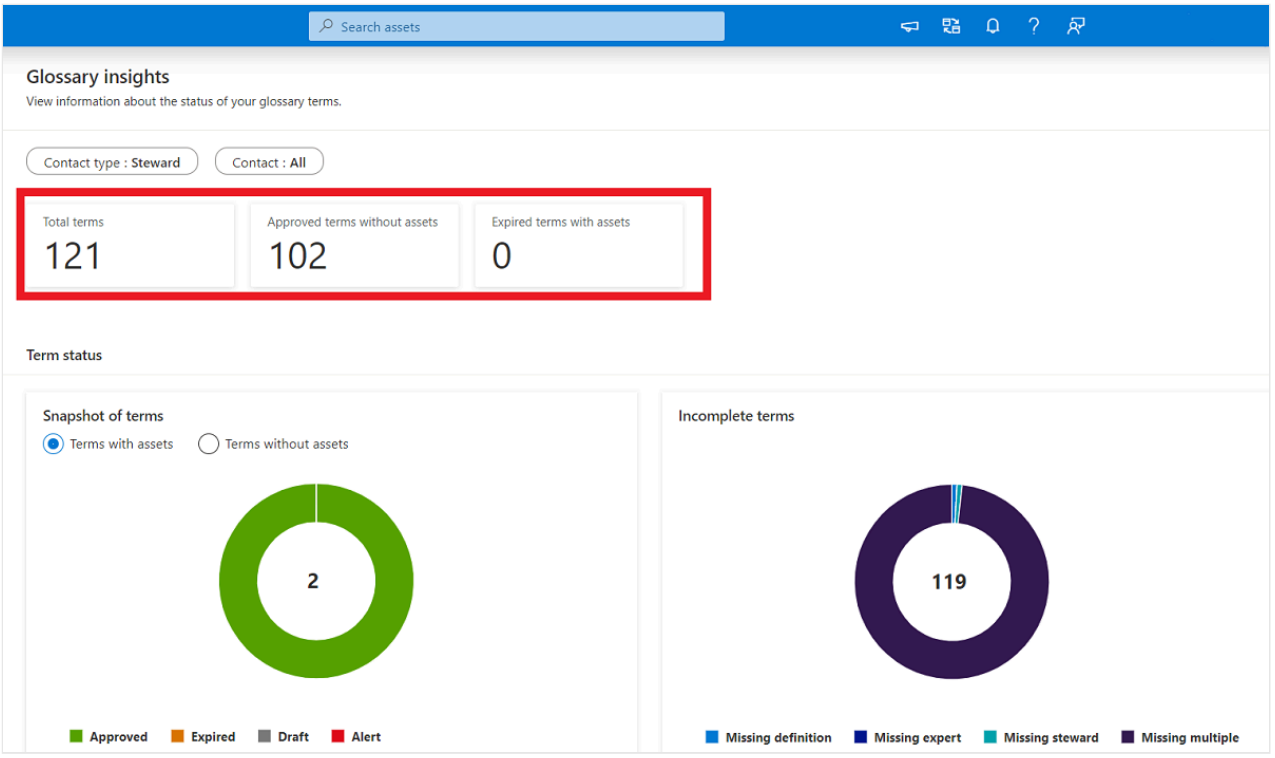
View glossary insights

1. Open the Microsoft Purview portal.

2. If you're using the classic Microsoft Purview portal, select **Data estate insights**. If you're using the new Microsoft Purview portal open **Unified Catalog**, select **Health Management**, then select **Reports**.

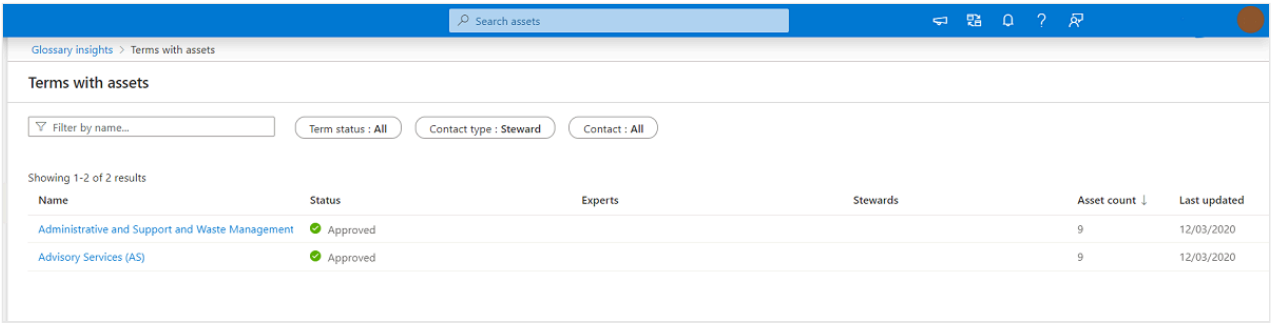
3. Select the **Classic glossary** report.

4. The report starts with **High-level KPIs** that shows ***Total terms*** in your Microsoft Purview account, ***Approved terms without assets***, and ***Expired terms with assets***. Each of these values helps you understand the current health of your Glossary.

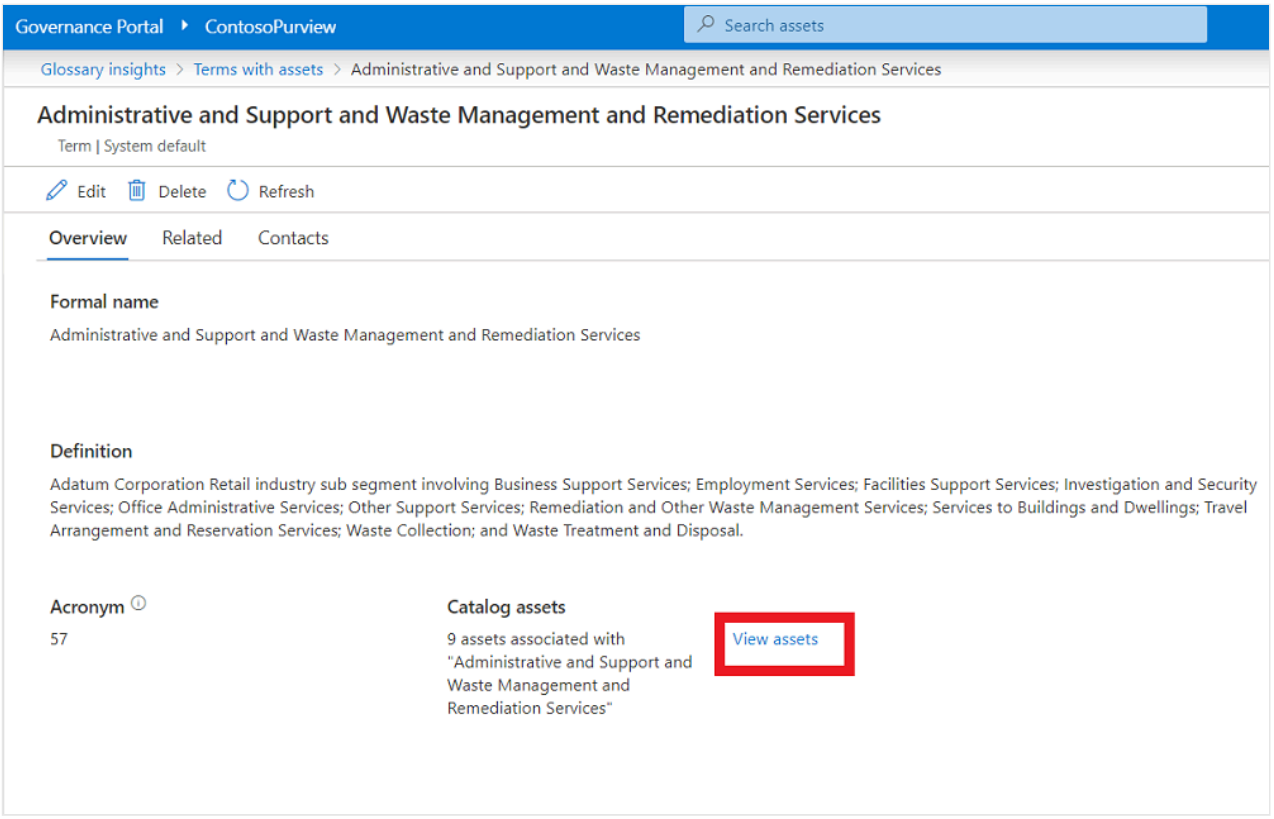


5. The **Snapshot of terms** section (displayed above) shows the term status as ***Draft***, ***Approved***, ***Alert***, and ***Expired*** for terms with assets and terms without assets.

6. Select **View details** to see the term names with various status and more details about ***Stewards*** and ***Experts***.



7. When you select "View more" for ***Approved terms with assets***, Data Estate Insights allows you to navigate to the **Glossary** term detail page, from where you can further navigate to the list of assets with the attached terms.



8. In Glossary insights page, view a distribution of **Incomplete terms** by type of information missing. The graph shows count of terms with ***Missing definition***, ***Missing expert***, ***Missing steward***, and ***Missing multiple*** fields.

9. Select ***View more*** from **Incomplete terms**, to view the terms that have missing information. You can navigate to Glossary term detail page to input the missing information and ensure the glossary term is complete.

Last updated on 11/14/2025

Understand the classic sensitivity labels report in Unified Catalog

This article describes how to access, view, and filter security insights provided by sensitivity labels applied to your data.

Supported data sources include: Azure Blob Storage, Azure Data Lake Storage (ADLS) GEN 1, Azure Data Lake Storage (ADLS) GEN 2, SQL Server, Azure SQL Database, Azure SQL Managed Instance, Amazon S3 buckets, Amazon RDS databases (public preview), Power BI.

In this article, you learn how to:

- ✓ Launch your Microsoft Purview account from Microsoft Azure.
- ✓ View sensitivity labeling insights on your data.
- ✓ Drill down for more sensitivity labeling details on your data.

Prerequisites

Before getting started with Data Estate Insights, make sure that you complete the following steps:

- Set up your Azure resources and populate the relevant accounts with test data.
- Extended sensitivity labels to assets in Microsoft Purview Data Map, and create or select the labels you want to apply to your data.
- Set up and complete a scan on the test data in each data source. For more information, see Manage data sources in Microsoft Purview and Create a scan rule set.
- Permissions depending on your current experience:
  - If using the classic Microsoft Purview Catalog, you need the data curator role or insight reader role.
  - If using the new Microsoft Purview Unified Catalog, you need data health reader permissions.

For more information, see Manage data sources in Microsoft Purview and Automatically label your data in Microsoft Purview.

Use Data Estate Insights for sensitivity labels

Classifications are similar to subject tags. Use them to mark and identify data of a specific type that you find within your data estate during scanning.

Sensitivity labels enable you to state how sensitive certain data is in your organization. For example, a specific project name might be highly confidential within your organization, while that same term isn't confidential to other organizations.

Classifications match directly, such as a social security number, which has a classification of **Social Security Number**.

In contrast, sensitivity labels are applied when one or more classifications and conditions are found together. In this context, conditions refer to all the parameters that you can define for unstructured data, such as **proximity to another classification**, and **% confidence**.

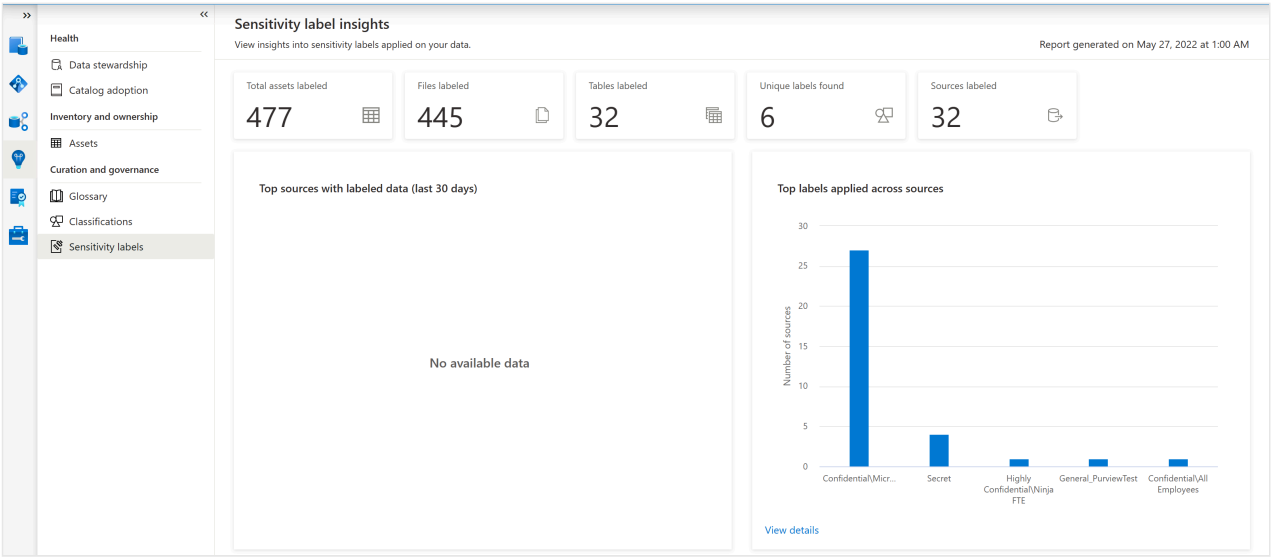
Data Estate Insights uses the same classifications, also known as sensitive information types, as those used with Microsoft 365 apps and services. This enables you to extend your existing sensitivity labels to assets in Data Map.

> ⓘ **Note**
>
> After you scan your source types, wait a couple of hours for the **Classic sensitivity labels** report to reflect the new assets.

View sensitivity labeling insights

1. Open the Microsoft Purview portal.
2. If you're using the classic Microsoft Purview portal, select **Data estate insights**. If you're using the new Microsoft Purview portal open **Unified Catalog**, select **Health Management**, then select **Reports**.
3. Select the **Classic sensitivity labels** report.

   > ⓘ **Note**
   >
   > If this report is empty, you might not have extended your sensitivity labels to Data Map. For more information, see Learn about sensitivity labels in Data Map.



The main **Sensitivity labeling insights** page displays the following areas:

Expand table

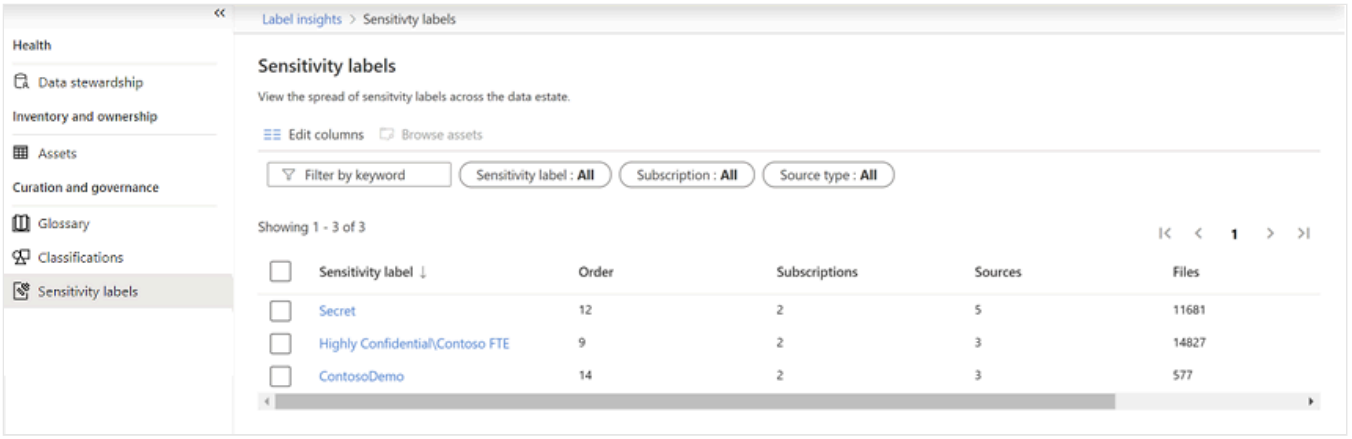
Area	Description
Overview of sources with sensitivity labels	Displays tiles that provide: - The number of subscriptions found in your data. - The number of unique sensitivity labels applied on your data - The number of sources with sensitivity labels applied - The number of files and tables found with sensitivity labels applied
Top sources with labeled data (last 30 days)	Shows the trend, over the past 30 days, of the number of sources with sensitivity labels applied.
Top labels applied across sources	Shows the top labels applied across all of your Microsoft Purview data resources.
Top labels applied on files	Shows the top sensitivity labels applied to files in your data.
Top labels applied on tables	Shows the top sensitivity labels applied to database tables in your data.
Labeling activity	Displays separate graphs for files and tables, each showing the number of files or tables labeled over the selected time frame. **Default**: 30 days Select the **Time** filter above the graphs to select a different time frame to display.

Sensitivity labeling insights drilldown

In any of the following **Sensitivity labeling insights** graphs, select the **View more** link to drill down for more details:

- **Top labels applied across sources**
- **Top labels applied on files**
- **Top labels applied on tables**
- **Labeling activity > Labeled data**

For example:



To learn more, use the following options:

Expand table

Option	Description
Filter your data	Use the filters above the grid to filter the data shown, including the label name, subscription name, or source type. If you're not sure of the exact label name, enter part or all of the name in the **Filter by keyword** box.
Sort the grid	Select a column header to sort the grid by that column.
Edit columns	To display more or fewer columns in your grid, select **Edit Columns**, then select the columns you want to view or change the order. Select a column header to sort the grid by that column.
Drill down further	To drill down to a specific label, select a name in the **Sensitivity label** column to view the **Label by source** report. This report displays data for the selected label, including the source name, source type, subscription ID, and the numbers of classified files and tables.
Browse assets	To browse through the assets found with a specific label or source, select one or more labels or sources, depending on the report you're viewing, then select **Browse assets** above the filters.

Option	Description
	The search results display all of the labeled assets found for the selected filter. For more information, see Search the Microsoft Purview Unified Catalog.

Sensitivity label integration with Microsoft Purview Information Protection

Close integration with Microsoft Purview Information Protection means that you have direct ways to extend visibility into your data estate, and classify and label your data. For more information about sensitivity labels, see: Learn about sensitivity labels.

To extend sensitivity labels to your assets in Data Map, you must turn on this capability in the Microsoft Purview portal.

For more information, see Apply sensitivity labels to your data in Data Map.

Next steps

Learn how to use Data Estate Insights with sources below:

- Learn how to use Asset insights.
- Learn how to use Data Stewardship.
- Learn how to use Classification insights.
- Learn how to use Glossary insights.

Last updated on 11/14/2025

Understand the data governance report in Unified Catalog (preview)

The data governance health report allows your team to track your health progress at a glance, and identify areas that need more work. This article covers how you can access this report and what the provided information means for your data governance health.

> **Note**
>
> If you don't use data health controls, the data governance health report doesn't refresh. Ensure that you set up the schedule to run data health controls for refreshing the health controls scores and the data governance health report.

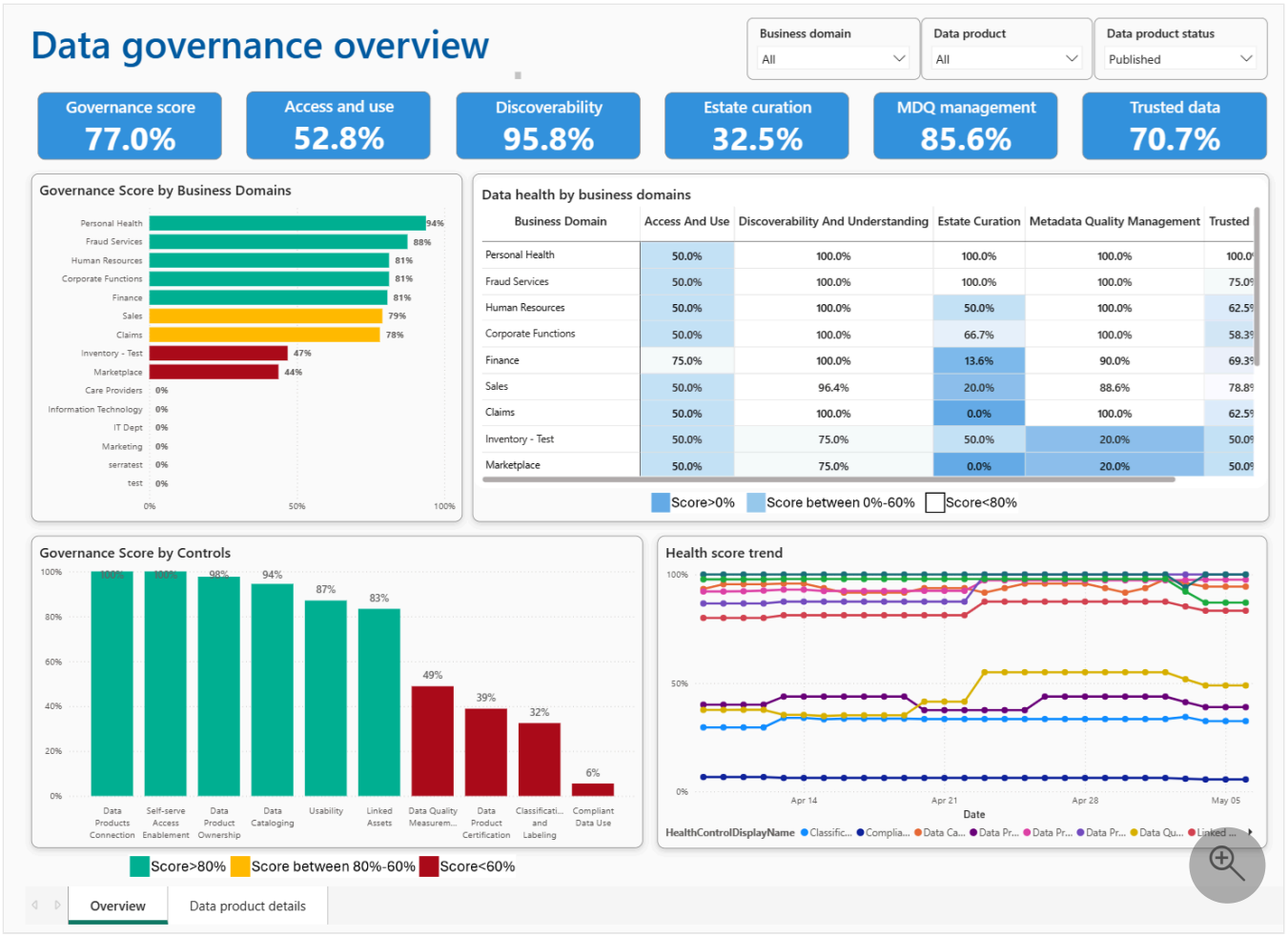
Prerequisites

- You need data health reader permissions to view Data Estate Health information.

View data governance health report

1. Open the Microsoft Purview portal and select **Unified Catalog**.
2. At **Health management**, select **Reports**.
3. Select the **Data governance (preview)** report.

The overview page of the report covers the overall health of your governance domains, and provides details about your data products.



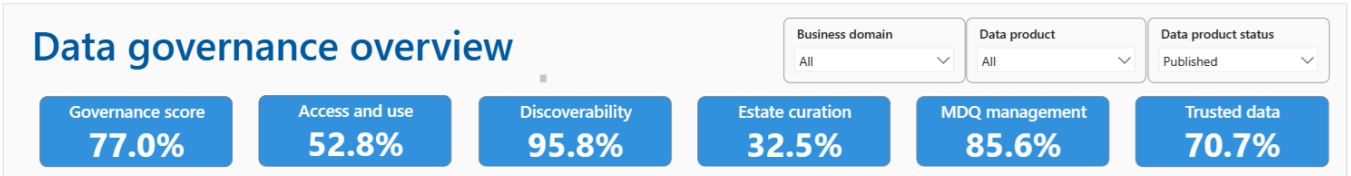
Use the filters to see information for specific governance domains, data products, or data products in a certain status (for example: draft).

Overview

The main overview covers the controls for your governance domains.

Controls and actions

The top controls and actions help you understand your overall governance health at a glance. They cover both controls and actions.



Expand table

Control or action	Description	Example
Governance score	Average of the ownership, cataloging, and classification controls.	
Access and use	Average of the self-service access enablement and compliant data use controls.	(Self-serv score is **100** + Compliaant score is **5.6**) / 2 = **52.8%** is the Access and use score control group score.
Discoverability	Average of these two checks: number of governance domains with Unified Catalogs and the number data products that have data assets mapped.	
Estate curation	Percentage of data assets that have classifications applied.	Estate curation score is **32.5%** because only 32.5% of data assets are classified during scan.
MDQ management	Average of data product useability (products with long descriptions and attached glossary terms), and percentage of data products with linked assets.	(MDQ data product usability is **87.9** + MDQ Linked asset is **83.3**) /2 = **85.6%** is the MDQ management score.
Trusted data	Average of three controls: Percentage of data products with owners, percentage of data products that have data quality scoring enabled, percentage of data products that are certified.	(% of data products with owner is **97.6** + % data product with data quality score is **68.4** + % of data product certified is **46.2**) = **70.7%** is the Trusted data score.

> **Tip**
>
> If you use the filters, these KPIs reflect scores for the governance domains or data products you select.

Data health by governance domain

The data health by governance domain chart gives you individual control scores for your governance domains without needing to use the filters. At a glance, you can see which of your governance domains are healthy and which need more attention.

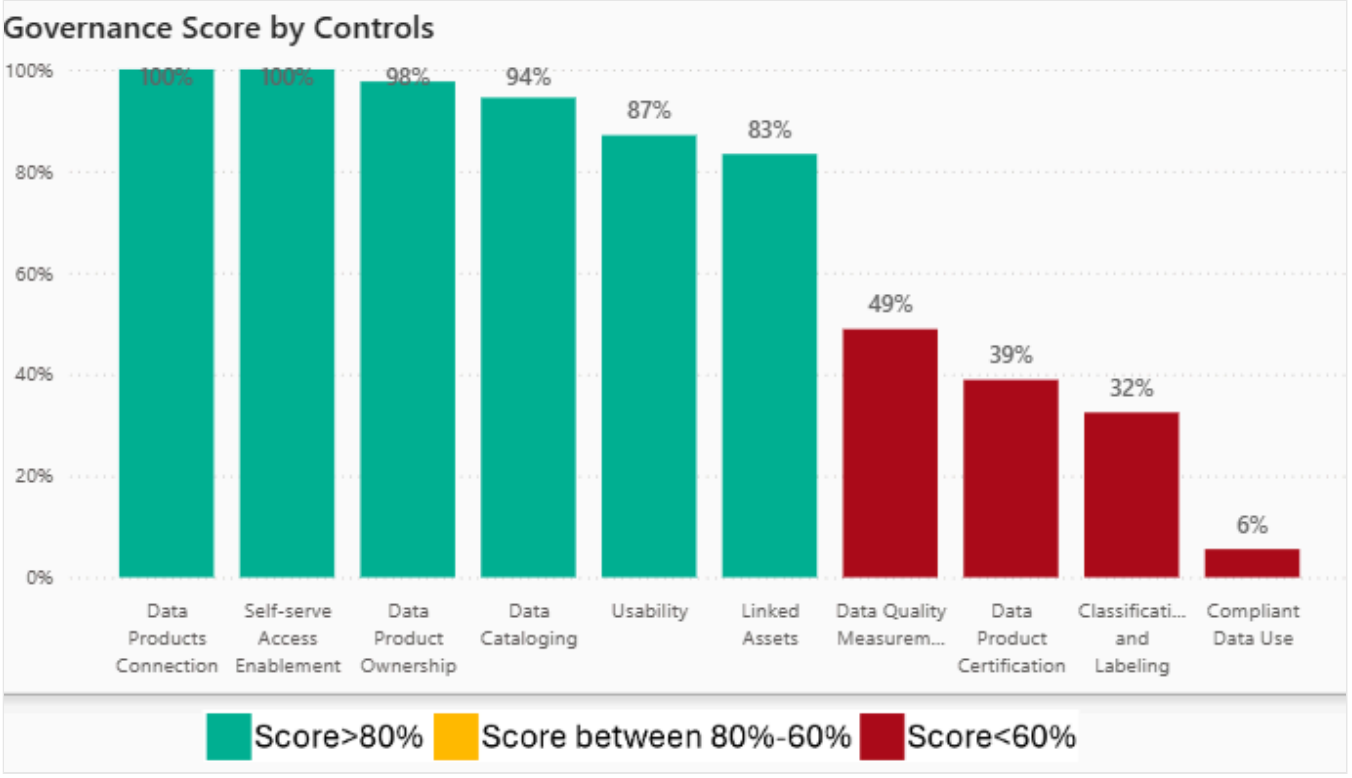
Data health by business domains

Business Domain	Access And Use	Discoverability And Understanding	Estate Curation	Metadata Quality Management	Trusted
Personal Health	50.0%	100.0%	100.0%	100.0%	100.0
Fraud Services	50.0%	100.0%	100.0%	100.0%	75.0
Human Resources	50.0%	100.0%	50.0%	100.0%	62.5
Corporate Functions	50.0%	100.0%	66.7%	100.0%	58.3
Finance	75.0%	100.0%	13.6%	90.0%	69.3
Sales	50.0%	96.4%	20.0%	88.6%	78.8
Claims	50.0%	100.0%	0.0%	100.0%	62.5
Inventory - Test	50.0%	75.0%	50.0%	20.0%	50.0
Marketplace	50.0%	75.0%	0.0%	20.0%	50.0

Score>0% Score between 0%-60% Score<80%

Governance score by controls

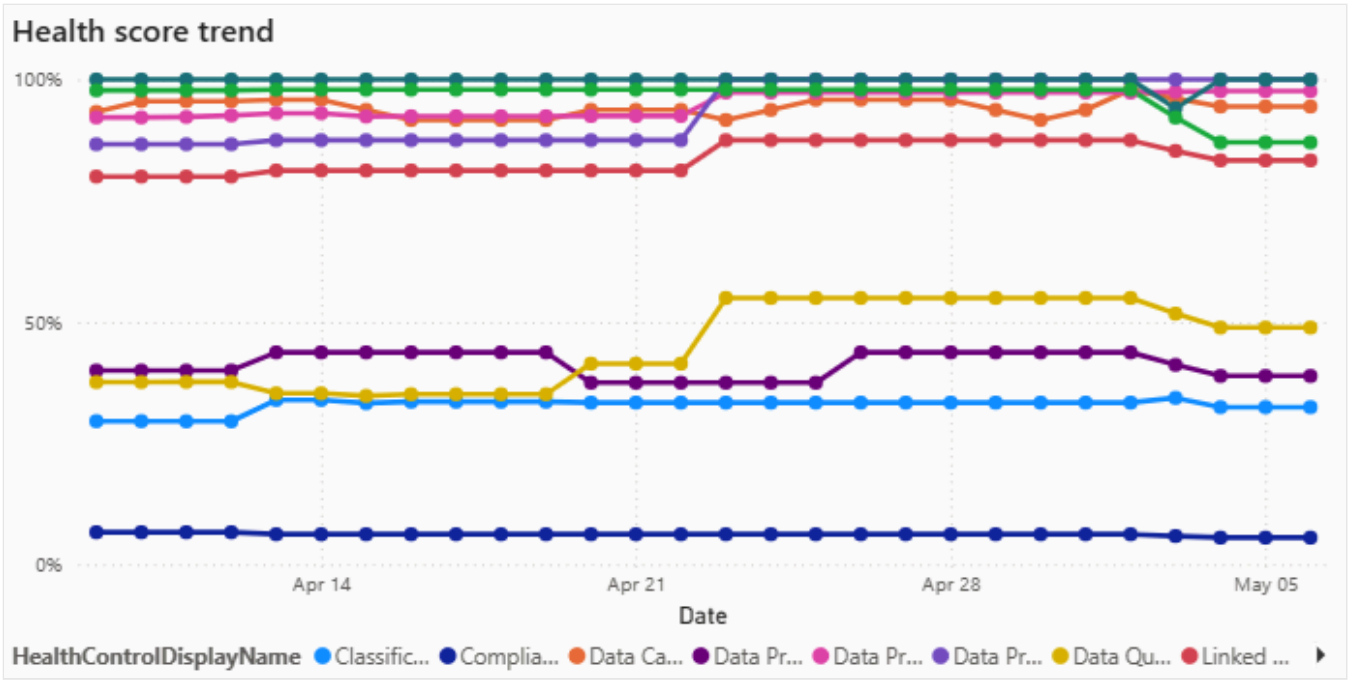
This column chart highlights your governance strengths and weaknesses by showing overall score for each control in Unified Catalog. Prioritize low or flagging scores to increase your data governance coverage across your data estate.



Health score trend

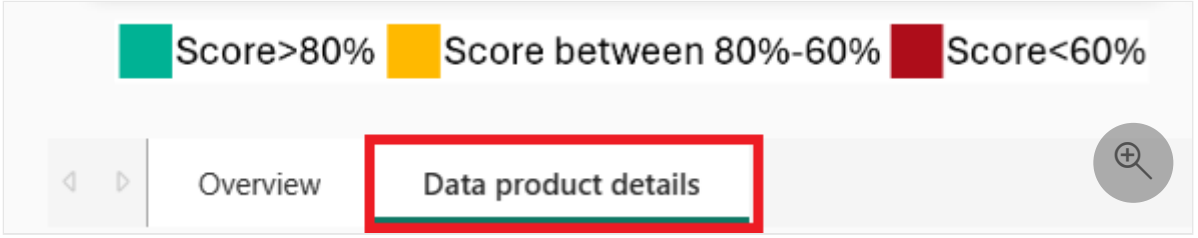
This table shows your daily health progression for all of the main health controls.

Use this table to track how well your team manages your data estate over time and to track your progress toward healthy data governance.



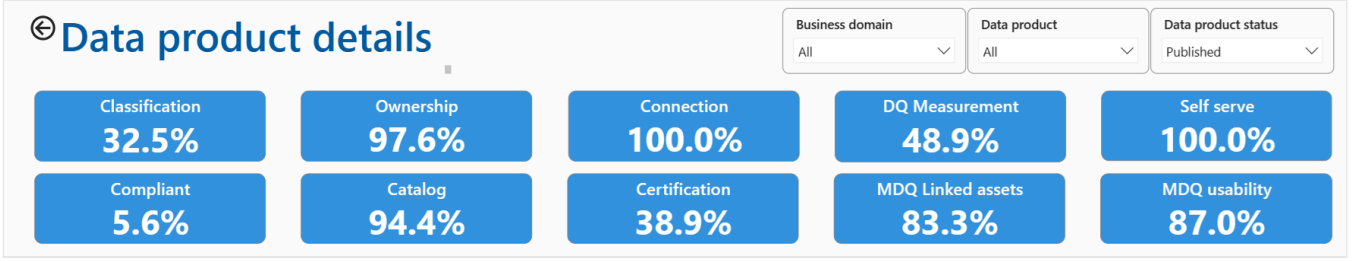
Data product details

Select the **Data product details** tab at the bottom of the page to see action and control information for your data products.



Data product controls and actions

The top controls and actions help you understand your overall governance health at a glance. They cover both controls and actions.



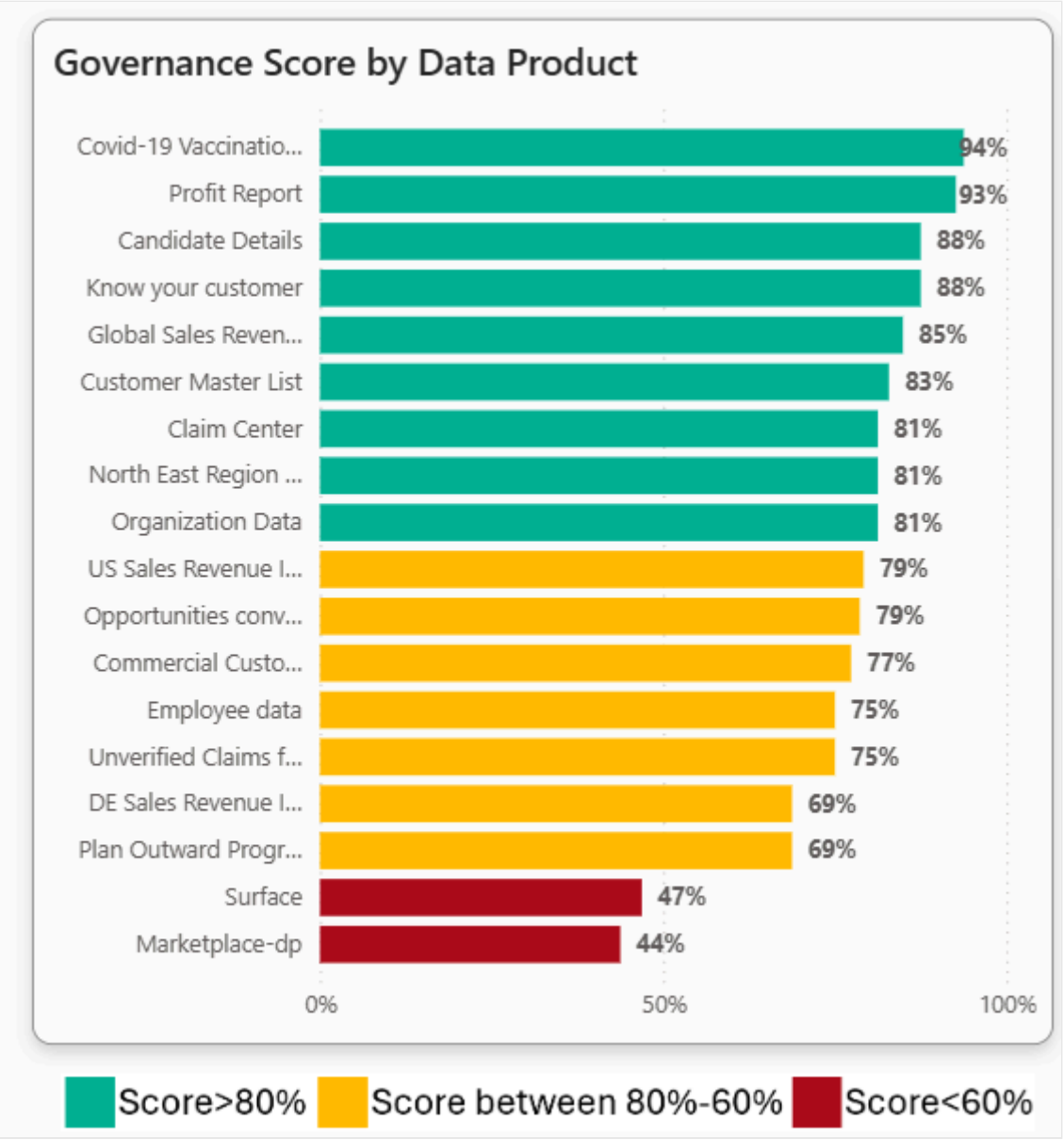
Expand table

Control or action	Description	Example in the image above
Classification	Percentage of data products that have classified assets attached.	**32.5%** of data assets are classified by Data Map scans.
Ownership	Percentage of data products that have an active owner.	**97.6%** of data products have at least one owner.
Connection	Percentage of data products that have at least one asset attached.	**100%** of data products have at least one data asset attached.
DQ measurement	Aggregation of data quality scores across the data products.	**48.9%** of data assets have a data quality score.
Self serve	Percentage of data products with self-service access setup.	**100%** of data products have a self-serve access policy attached.
Compliant	Percentage of data products with the purpose of use policy added.	**5.6%** of data products have a compliant policy attached.
Catalog	Percentage of data products with a description, published glossary terms, and a use case.	**94.4%** of data products have descriptions, use cases, and at least one glossary term associated.
Certification	Percentage of data products that have been verified.	**38.9%** of data products are verified.
MDQ linked assets	Percentage of data products with at least one asset and at least one published glossary term.	**83.3%** of data products have at least one data asset and one glossary term linked.
MDQ usability	Percentage of data products that have a description over 100 characters long, the parent domain has a description 100 characters long, and the associated term has a description 100 characters long.	**87%** of data products have descriptions and use cases that are at least 100 characters long.

Use the filters to narrow your view to only certain governance domains, data products, or data products with a certain status.

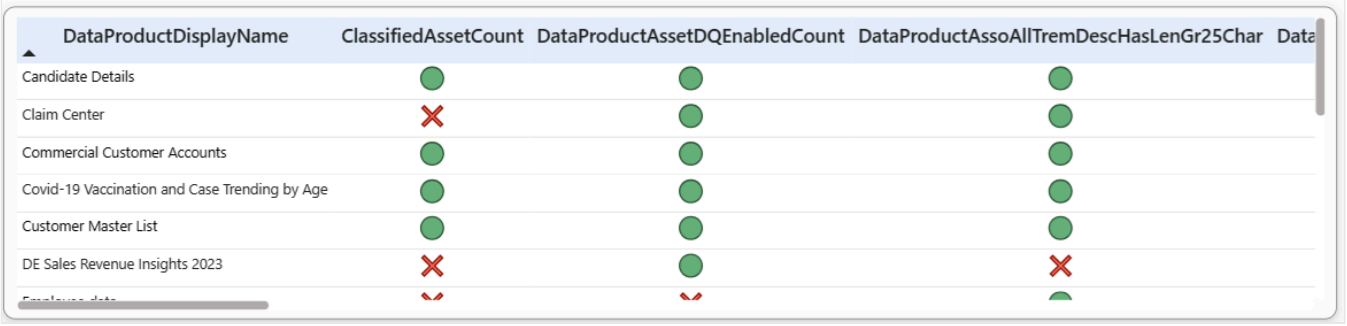
Governance score by data product

This bar chart highlights your strongest and weakest data products by showing the overall governance score for each data product. Prioritize low-scoring but high-use data products to bring the most value to Unified Catalog.



Individual control checks

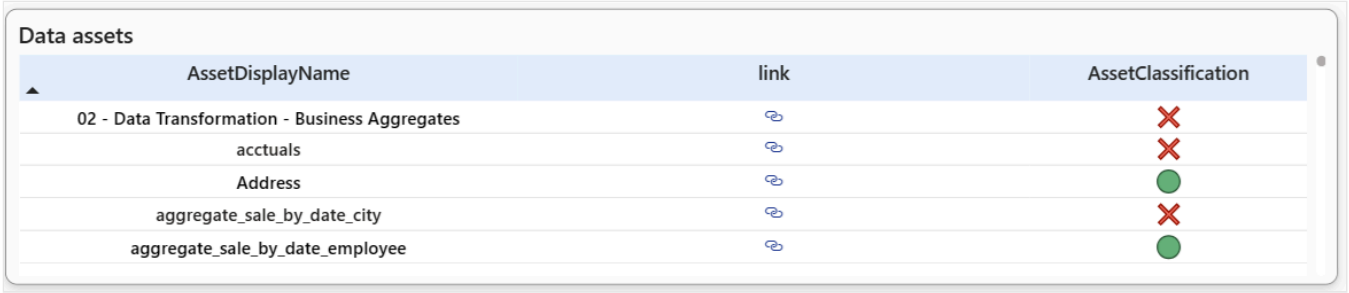
This table shows the pass or fail status for each individual control check available to your data product. You can check controls for low-scoring data products to see which areas need improvement.



DataProductDisplayName	ClassifiedAssetCount	DataProductAssetDQEnabledCount	DataProductAssoAllTremDescHasLenGr25Char	Data
Candidate Details	●	●	●	
Claim Center	✕	●	●	
Commercial Customer Accounts	●	●	●	
Covid-19 Vaccination and Case Trending by Age	●	●	●	
Customer Master List	●	●	●	
DE Sales Revenue Insights 2023	✕	●	✕	

Data assets

This table shows the pass or fail status for individual assets in Unified Catalog. If multiple data products reference a single asset, resolving any outstanding governance issues for that asset raises your score overall.

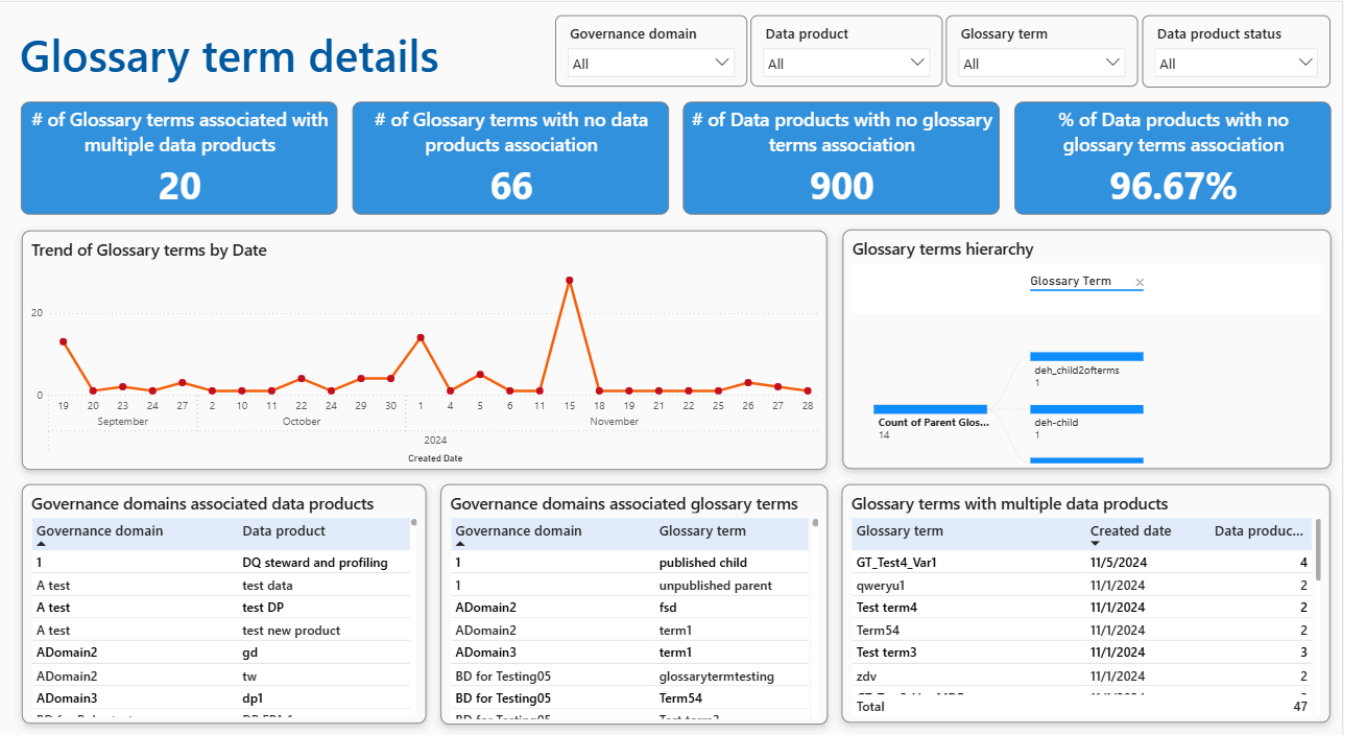

Data assets

AssetDisplayName	link	AssetClassification
02 - Data Transformation - Business Aggregates		✗
acctuals		✗
Address		●
aggregate_sale_by_date_city		✗
aggregate_sale_by_date_employee		●

Glossary terms report (preview)

Glossary terms provide vocabulary for business users. These terms allow users to discover and work with data in the vocabulary that's familiar to them versus using abstract technical terminology inherited from physical data sources. Terms are created under governance domains to create context specific to each part of your organization. Both sales and marketing might use the same term to mean different things, and your governance domains help your team to differentiate between those meanings. Once created, terms map to data products, to provide context for those data products, and to provide specific data governance based on business context. Terms are an active part of Microsoft Purview that provide data governance by now containing policies. The glossary terms report (preview) helps data stewards to browse:

- Glossary terms trends.
- Glossary terms hierarchy.
- Data product associated with glossary terms.
- Governance domain association with glossary terms.



Next steps

- Actions
- Controls
- Reports

Last updated on 11/14/2025

Understand the quality report in Unified Catalog

The data quality health report evaluates and summarizes the quality of data within an organization or system. It includes assessments of various data quality dimensions and metrics to help stakeholders understand the accuracy, completeness, consistency, reliability, and timeliness of their data. This data quality report allows your team to track your health management progress at a glance and identify areas that need more work to improve the quality of data in your data estate.

This article covers how you can access this report and what the provided data quality measure means for your health management.

Report purposes

- **Monitoring and governance**: To continuously monitor and manage the quality of data, ensuring it meets the organization's standards and regulatory requirements.
- **Decision support**: To provide stakeholders with reliable data for making informed business decisions.
- **Identifying issues**: To detect and document data quality issues, enabling timely remediation.
- **Improving data management**: To enhance data management practices by identifying root causes of data quality issues and implementing corrective measures.
- **Performance measurement**: To measure the effectiveness of data quality initiatives and track improvements over time.
- **Stakeholder communication**: To communicate data quality status and progress to stakeholders, including management, data product owners, data stewards, and IT teams. By providing a clear and comprehensive view of the state of data quality, these reports play a crucial role in maintaining the integrity and usefulness of data within an organization.

Prerequisites

You need the Data Health Reader role to view data health information.

View data quality health report

1. In the Microsoft Purview portal, open **Unified Catalog**.
2. Go to **Health management** > **Reports**.
3. Select the **DQ health** report.

Data quality dimension reporting

The overview page of the report shows data quality dimension scores, data quality rule hierarchy, data quality status by dimension, and data quality dimensions and rule types used for different data assets. The top controls help you understand your overall health management at a glance.



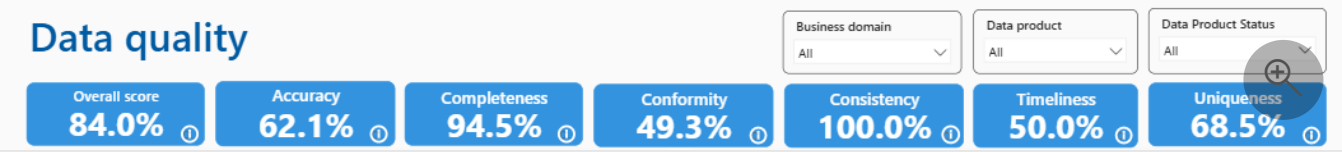
Use the filters to see information for specific governance domains, data products, or data products in a certain status (for example: draft).

Expand table

Data Quality Dimension	Description
Accuracy	Data should accurately represent real-world entities. Context matters. For example, if you're storing customer addresses, ensure they match the actual locations.
Completeness	The objective of this rule is to identify the empty, null, or missing data. This rule validates that all values are present (though not necessarily correct).
Conformity	This rule ensures that the data follows data formatting standards such as representation of dates, addresses, and allowed values.
Consistency	This rule checks that different values of the same record conform with a given rule and that there are no contradictions. Data consistency ensures that the same information is represented uniformly across different records. For example, if you have a product catalog, consistent product names and descriptions are crucial.
Timeliness	This rule aims to ensure that the data is accessible in as short a time as possible. It ensures that the data is up to date.
Uniqueness	This rule checks that values aren't duplicated. For example, if there's supposed to be only one record per customer, then there aren't multiple records for the same customer. Each customer, product, or transaction should have a unique identifier.

Data quality overall score and dimension scores help data practitioners and data estate owners understand how complete, accurate, consistent, and trustworthy their data is. These scores also indicate what improvement actions need to be taken to enhance the quality of their data estate.

How the scores are calculated



> **Tip**
>
> If you use the filters, these KPIs show scores for the governance domains or data products you select.

Data quality status by dimensions

Overall score % = Percentage of data quality dimension score of all measured dimensions divided by total number of dimensions with score. **Exception**: If any dimension score is **Blank**,

the calculation excludes the **Blank** score from the overall score; for example:

Expand table

Data Quality Dimension	Score %
Accuracy	100
Completeness	80
Consistency	100
Conformity	Blank
Uniqueness	100
Timeliness	50

Overall score % = ((100+80+100+100+50)/5)*100 = 86%

As conformity score is **Blank**, the calculation excludes it from the overall score. If the conformity score is zero (0) instead of **Blank**, the calculation includes the zero score. In this scenario, the overall score = ((100+80+100+0+100+50)/6)*100 = 72% (71.66%).

The score for each dimension is calculated as follows:

Data quality score for a dimension (%) = (Total number of records that pass all applied rules for the selected dimension across all data assets of a data product) ÷ (Total number of records of all data assets of the selected data product measured against the selected dimension) × 100

Example:

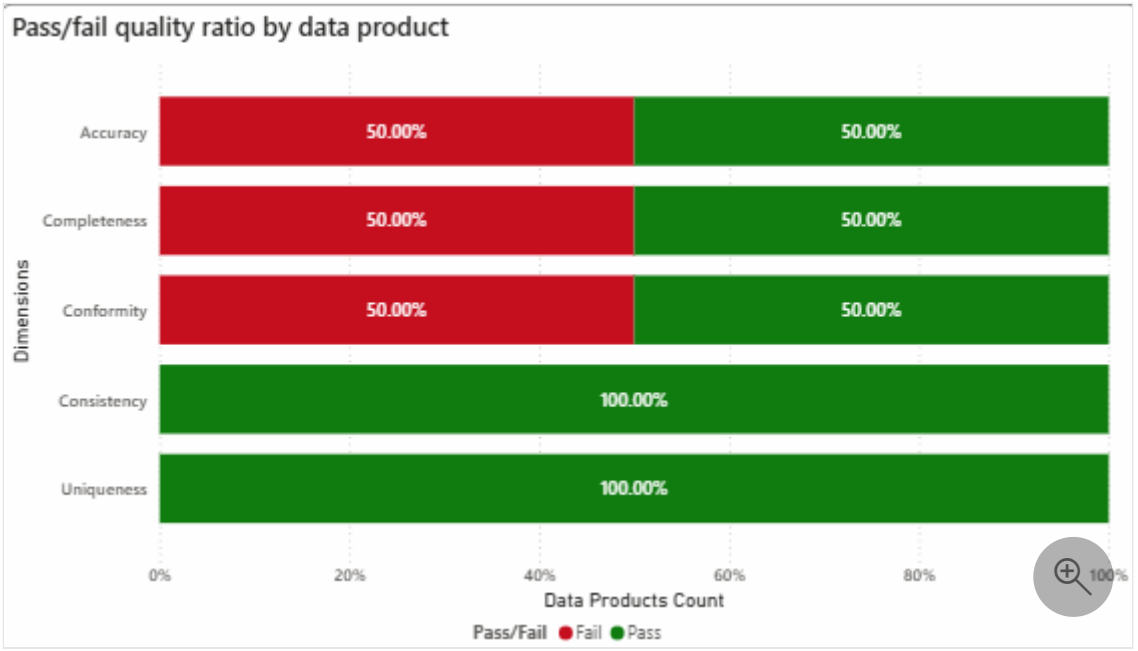
Total number of assets associated with a data product is 4. Total passed records for asset one are 1,000, for asset two are 200, for asset three are 4,000, and for asset four are 100. Total number of records for all four assets is 6,000. Data quality score (%) for the selected dimension = ((1,000 + 200 + 4,000 + 100) / 6,000) * 100 = 88%.

Data quality dimension scores are calculated for each governance domain. Dimensions are mapped with rules, and the score rolls up from data asset columns to data asset, and from data asset to data product and governance domain level for each industry standard dimension. You can filter dimension level scores per governance domain to investigate more details.

Domain	Accuracy	Completeness	Conformity	Consistency	Timeliness	Uniqueness	Overall Score
Care Providers		90.35%					90.35%
Claims	0.00%	94.75%	14.44%			0.00%	56.96%
Corporate Functions		99.35%				79.38%	86.54%
Finance		98.23%		100.00%		66.72%	93.21%
Fraud Services		100.00%				100.00%	100.00%
Human Resources		84.34%				2.68%	81.34%
Personal Health		100.00%		100.00%			100.00%
Sales	70.95%	92.87%	75.46%	100.00%	50.00%	78.09%	89.43%
Test-cap		89.98%				0.00%	15.00%

Data quality rules pass and fail ratio

The pass and fail ratio of data quality rules is measured for each data quality dimension for data products. This measure helps data owners and data practitioners understand what percentage of data in a data product is inaccurate, inconsistent, incomplete, duplicate, or not fresh enough as expected. This measure also helps to investigate and understand whether the applied rules are incorrect or the data is incorrect.



Example scenario 1

You create rules to measure accuracy and completeness of a data asset. After a data quality scan is completed and the data health report is refreshed, if the Accuracy score is 100% and Completeness score is 16.6%, the report displays one data product count.

- Accuracy as 100%, with a green bar.
- Completeness as 16.6%, which is below the threshold (100%), as Failed with a red bar.

Example scenario 2

You apply two accuracy dimension rules. After a data quality scan, you see one accuracy rule score is 100%, which is visible in the report as Passed, shown with a green bar. Another accuracy rule dimension for same data product has a score of 15.98%, which is visible in the report as Failed, shown with a red bar. The report shows 50% Passed and 50% Failed, because Accuracy has two rule dimensions.

Expand table

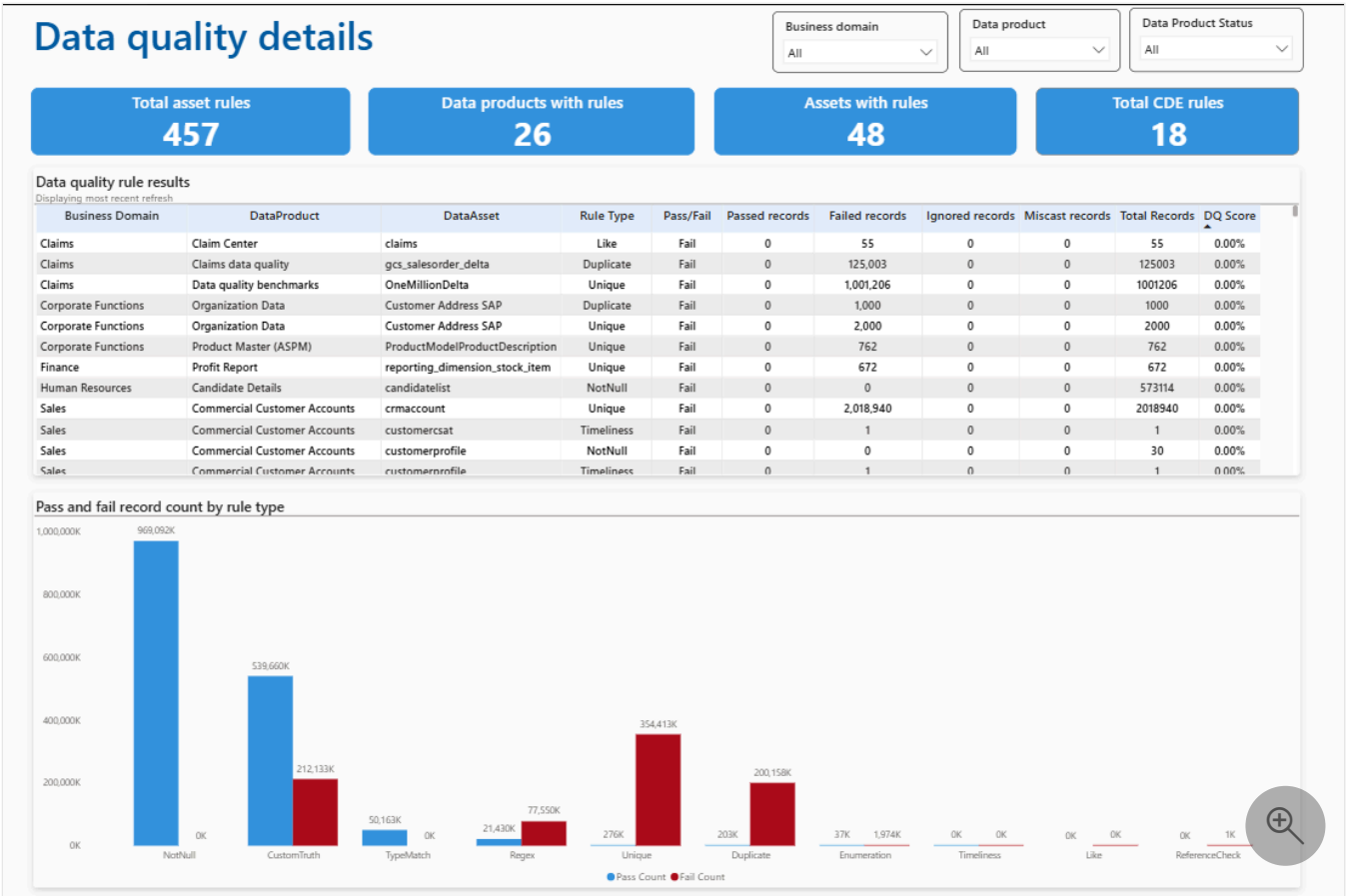
Accuracy rule	Score	Result
Rule 1	100%	Passed
Rule 2	15.98%	Failed

Calculation:

- Passed rule = 1
- Failed rule = 1
- Total rules = 2
- Accuracy score pass fail ratio = (1/2) x 100 = 50%, so you see 50% Pass, and 50% Fail.

Data quality details report

In the data quality health report, select the **Details** tab to see how many rules apply to data products, data assets, and critical data elements. These rules help measure and monitor the quality of the entire data estate of your organization. You can drill down to see how many records of a data asset failed for a rule type, which rule type is performing better, and which governance domain and data products are publishing and maintaining trustworthy data. You can filter the measures by governance domain and data product to understand the current state and to plan improvement actions. Here's an example of the details view:



Historical trend

In the data quality health report, select the **Historical Trend** tab to help you understand the historical trend of the data quality of governed data assets. You can filter governance domain and data product level trends for the associated data assets. You can also use the selector to select the top or bottom 10 data assets to browse the data quality trends of each data quality dimension. Here's an example of the historical trend view:



> **Note**
>
> - The data quality health report depends on data health controls and Microsoft Purview metadata self-serve analytics model. If you don't use the data health controls and don't subscribe Unified Catalog metadata, the data quality health report doesn't refresh. You need to use data health controls or subscribe Microsoft Purview metadata for self-serve analytics to refresh the data quality health report.
> - If you don't use the data quality feature, the data quality health report appears blank because the report is created using data generated from the data quality scan feature.

Next steps

- Actions
- Controls
- Reports

Last updated on 04/07/2026

Self-serve analytics and insights

Self-serve analytics and insights refer to data, tools, and platforms that enable business users to access, analyze, and generate insights from data independently. Microsoft Purview data governance publishes a domain model of metadata into Microsoft Fabric OneLake and Microsoft Azure Data Lake Storage Gen2, so you can analyze and generate insights by bringing your own tools and compute. Self-serve analytics of data governance metadata is valuable in driving continuous improvement of your data estate health management and fostering a data-driven culture across your organization by democratizing access to data estate insights and health management.

Key components

- **Data Model:** 3NF Model with domains and dimensions details.
- **Metadata:** Data governance metadata that includes:
  - Governance domains
  - Data products
  - Data assets
  - Glossary terms
  - Subscription request
  - Data quality rules
  - Dimensions
  - Data quality facts (pass and fail counts)
  - Data health actions (including data quality actions)
  - And many more

Benefits

- **Empowerment:** Empowers data practitioners, data product owners, data stewards, and analysts to explore data governance metadata and link metadata from various sources to derive insights.
- **Flexibility and efficiency:** You can create custom reports in addition to out-of-the-box reports in health management.
- **Agility:** Allows your organization to respond more swiftly to health management issues and remediation.

- **Cost-effective:** Reduces the need for setting up platforms and build tools. All data is available in OneLake and you can use available tools (Fabric semantic model, Power BI reporting, data flow, and notebook) in OneLake.

Available reports (out-of-the-box)

Here are the available out-of-the-box reports. You can't customize these reports.

- Classic assets - an overview of assets by type and collection, and their curation status.
- Classic catalog adoption - to understand at a glance how Unified Catalog is being used.
- Classic classifications - an overview of assets classified and the types of classifications.
- Classic data stewardship - an overview of assets classified and the types of classifications.
- Classic glossary - health and use of glossary terms.
- Classic sensitivity labels - an overview of assets that have sensitivity labels applied and the types of labels applied.
- Data governance - The data governance health report allows your team to track your health progress at a glance, and identify areas that need more work.
- Data quality health - Data quality dimensions and data quality rules performance reporting.

Data model for self-serve analytics metadata

The 3NF domain model is part of the normalization process in relational database design. It ensures that the database is free of redundancy and update anomalies. A database schema is in the Third Normal Form if it meets the requirements of the First and Second Normal Forms and all its attributes are functionally dependent only on the primary key.

Use the 3NF domain model to structure data in a way that minimizes duplication and ensures data integrity. It focuses on breaking down data into smaller, related tables where each piece of information is stored only once.

Characteristics:

- Elimination of transitive dependencies: Nonkey attributes shouldn't depend on other nonkey attributes.
- Logical grouping: Data is logically grouped into domains based on its function or meaning.
- Entity-relationship diagrams (ERDs): Commonly used to represent 3NF domain models, showing how entities relate to each other.

Expand table

Table name	What's stored in the table	Relationship keys
Access Policy Provisioning State	Information about provisioning state	ProvisioningStateId
Access Policy Resource Type	Access policy resource information	ResourceTypeId
Access Policy Set	Overall information about access policy, policy use case details, and where the policy is applied	AccessPolicySetId (UniqueId), ResourceTypeId (FK), ProvisioningStateId (FK)
Business Domain	Business domain (called **governance domain** in Unified Catalog) name, description, status, and ownership details	Parent Business Domain ID (FK), Created By User ID (FK), Last Modified By User ID (FK)
Classification	Data asset classification information	ClassificationId
Critical Data Column	Store glossary terms associated to data asset column in this field	Critical Data Column Id, Business Domain Id, Asset Id
Critical Data Element	Critical data elements of the data assets	Critical Data Element Id
Custom Access Use Case	Access use case information	AccessPolicySetId
Data Asset	Data asset name, description, and source information	DataAssetId (UniqueId), DataAssetTypeId (FK), CreatedByUserId (FK), LastModifiedByUserId (FK)
Data Asset Column	Data asset column name, column description, and references	DataAssetId (FK), ColumnId (Unique), DataAssetTypeId (FK), DataTypeId (FK), Created By User Id (FK), Last Modified By User Id (FK)
Data Asset Column Classification Assignment	Data classification assignment related reference keys	DataAssetId (FK), ColumnId (FK), ClassificationId (FK)
Data Asset Column Critical Data Element Assignment	Data asset column associating with critical data element	ColumnId
Data Asset Domain Assignment	Data asset governance domain assignment related information	DataAssetId (FK), BusinessDomainId (FK), AssignedByUserId (FK)
Data Asset Owner	Data asset owner information	DataAssetOwnerId

Table name	What's stored in the table	Relationship keys
Data Asset Owner Assignment	Data asset owner assignment information	DataAssetId, DataAssetOwnerId
Data Asset Type Data Type	Data asset type information	DataTypeId (UniqueId), DataAssetTypeId (FK)
Data Product	Data product name, description, use cases, status, and other relevant information	DataProductId (UniqueId), DataProductTypeId (FK), DataProductStatusId (FK), UpdateFrequencyId (FK), CreatedByUserId (FK), LastUpdatedByUserId (FK)
Data Product Asset Assignment	Data product and data asset assignment information	DataProductId, DataAssetId
Data Product Business Domain Assignment	Data product and governance domain assignment information	DataProductId (FK), BusinessDomainId (FK), AssignedByUserId (FK)
Data Product Critical Data Element Assignment	Critical data element assignment to data product	DataProductId
Data Product Documentation	Data product documentation reference information is stored in this table.	DataProductId, DocumentationId
Data Product OKR Assignment	Assigned objectives and key results to data product	Data Product Id, Objective Id, Key Result Id
Data Product Owner	Data product owner information	DataProductId, DataProductOwnerId
Data Product Status	Data product status (like published or draft) related information	DataProductStatusId
Data Product Terms Of Use	Data product usage terms information	DataProductId, TermOfUsedId, DataAssetId
Data Product Type	The information about data product types, such as master, reference, or operational	DataProductTypeId
Data Product Update Frequency	The information about how often this data product's data are updated	UpdateFrequencyId

Table name	What's stored in the table	Relationship keys
Data Quality Asset Rule Execution	Data quality scanning results, including custom rule column list and rule column IDs.	RuleId (FK), DataAssetId (FK), JobExecutionId (FK), RuleColumnList, RuleColumnIdList
Data Quality Job Execution	Data quality job execution status	JobExecutionId (UniqueId)
Data Quality Rule	Information about data quality rules	RuleId (UniqueId), RuleTypeId (FK), BusinessDomainId (FK), DataProductId (FK), DataAssetId (FK), JobTypeDisplayName (FK), RuleOriginDisplayName (FK), RuleTargetObjectType (FK), CreatedByUserId (FK), LastUpdatedByUserId (FK)
Data Quality Rule Column Execution	Information about data quality rules pass and fail count, data quality score at the column level, and data quality job execution details. Only out-of-the-box rule data are available. For custom rule column list and column ID list, see the Data Quality Asset Rule Execution table.	RuleId (FK), DataAssetId (FK), ColumnId (FK), JobExecutionId (FK)
Data Quality Rule Type	Data quality rule type and associated dimensions	RuleTypeId (UniqueId), DimensionDisplayName (FK)
Data Subscription Request	Information about data subscribers, policies applied, subscription request status, and other relevant information	SubscriberRequestId (UniqueId), SubscriberIdentityTypeDisplayName (FK), RequestorIdentityTypeDisplayName (FK), RequestorStatusDisplayName (FK)
Glossary Term	Information about glossary term, description, and overall status of the glossary term	GlossaryTermId (UniqueId), ParentGlossaryTermId (FK), CreatedByUserId (FK), LastModifiedByUserId (FK)
Glossary Term Business Domain Assignment	Information about glossary term governance domain assignment and statuses	GlossaryTermId (FK), BusinessDomainId (FK), AssignedByUserId (FK), GlossaryTermStatusId (FK), CreatedByUserId (FK), LastUpdatedByUserId (FK)
Glossary Term Critical Data Element Assignment	Associating glossary term with critical data element	GlossaryTermId
Glossary Term Data Product Assignment	Information about glossary term data product assignment	GlossaryTermId (FK), DataProductId (FK), AssignedByUserId (FK), GlossaryTermStatusId (FK), CreatedByUserId (FK), LastUpdatedByUserId (FK)

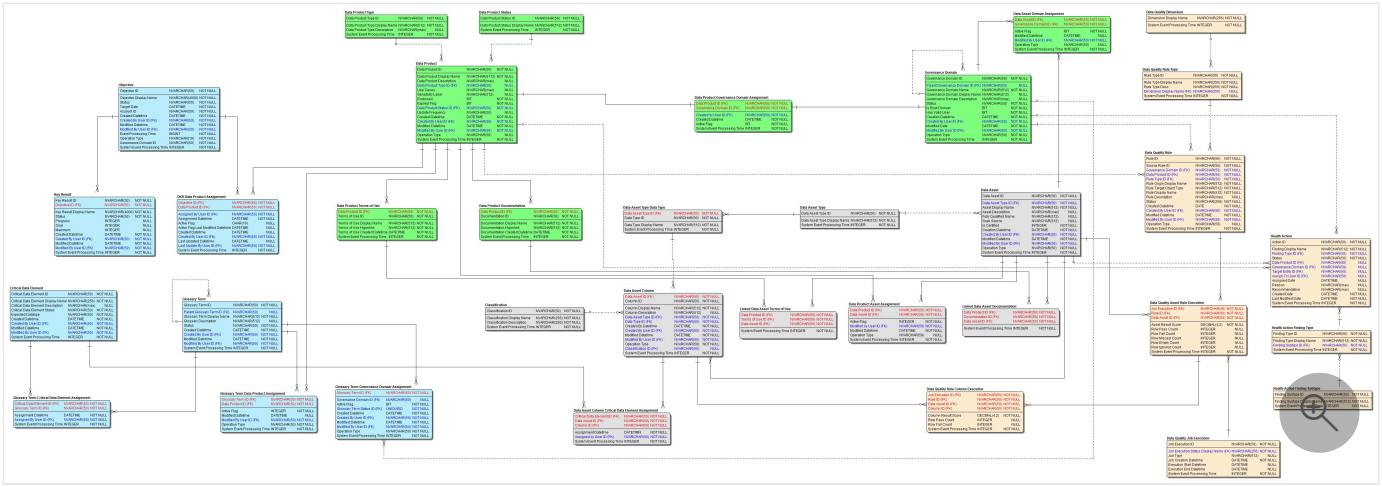
Table name	What's stored in the table	Relationship keys
Health Action	Information about data governance and data quality actions	ActionId, TargetEntityId, CreatedByUserId
Health Action Finding Type	Data health actions **Finding Types**	FindingTypeId
Health Action Finding Sub Type	Data health actions **Finding Sub Types**	FindingSubTypeId, FindingTypeId
Health Action User Assignment	Data health actions **User Assignment** information	ActionId, AssignedToUserId
Key Result	Key result of data product objective, including details description	Key Result Id, Objective Id
Objective	Objective of the data product contains objective Id, objective description, and objective status	Objective Id
Policy Set Approver	Policy set and the approver information	SubscriberRequestId (FK), AccessPolicySetId (FK), ApproverUserId (FK)
Relationship	Information about source type and target information. The relationship of glossary terms association with assets and columns is available in this relationship table.	AccountId, SourceId, TargetId

Note

- Business Domain is called **governance domain** in Unified Catalog, but it's still called **Business Domain** in the physical model in the database. The name difference doesn't affect the usage scenario described here for Business Domain.
- When you delete an asset from Data Map and later recreate it with the same name, the asset is treated as new and gets a new ID. As a result, the deleted asset is no longer valid for Data Quality rules column and column ID association. This scenario isn't currently supported. So, the `RuleColumnIdList` in the Data Quality Asset Rule Execution table stays blank for the deleted data asset in Data Map. You must run Data Quality

scans again on the newly created data asset, even if the asset name and schema are identical to the deleted one.

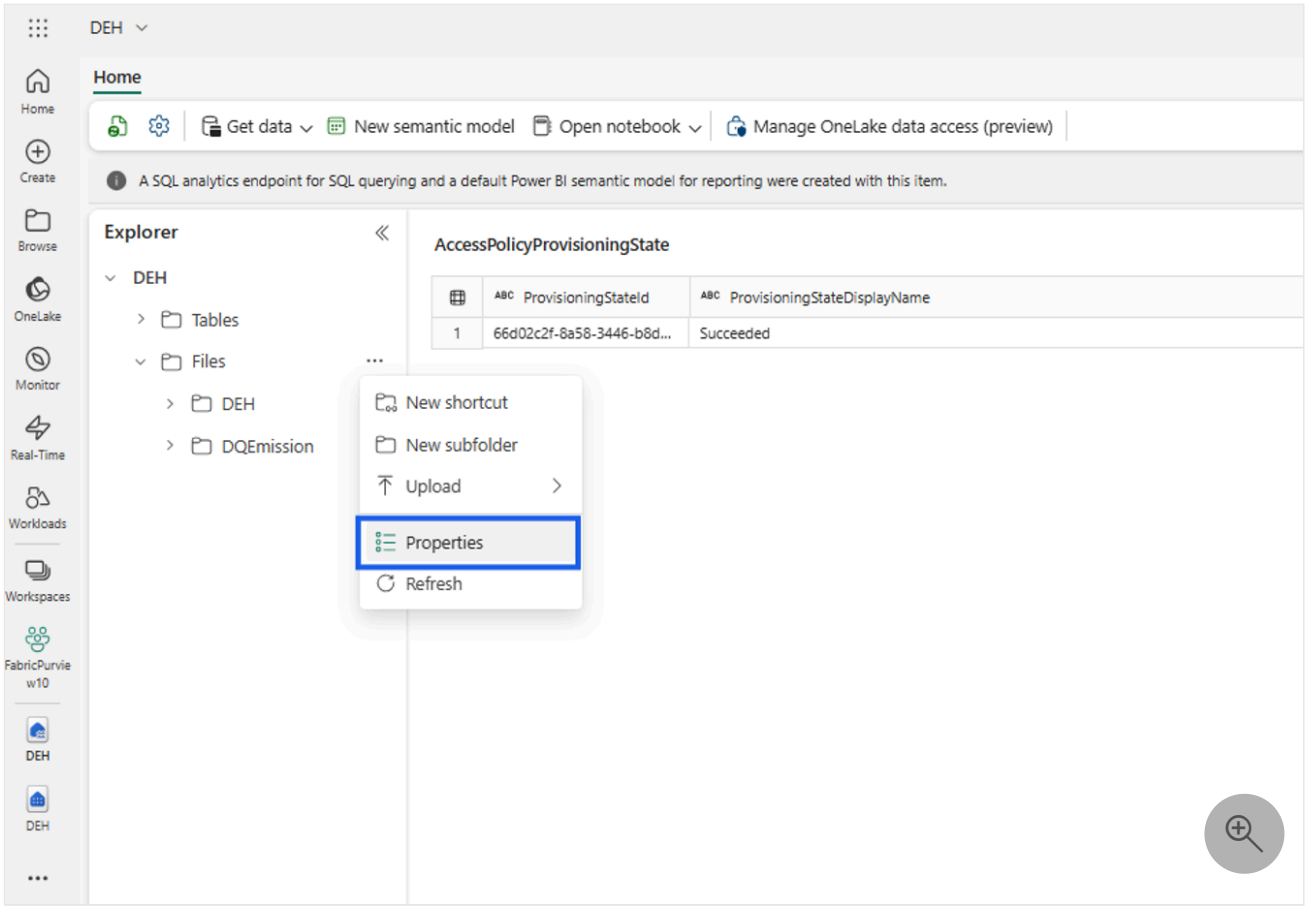
This diagram shows the entity relationship for the domain model described in the data model table:



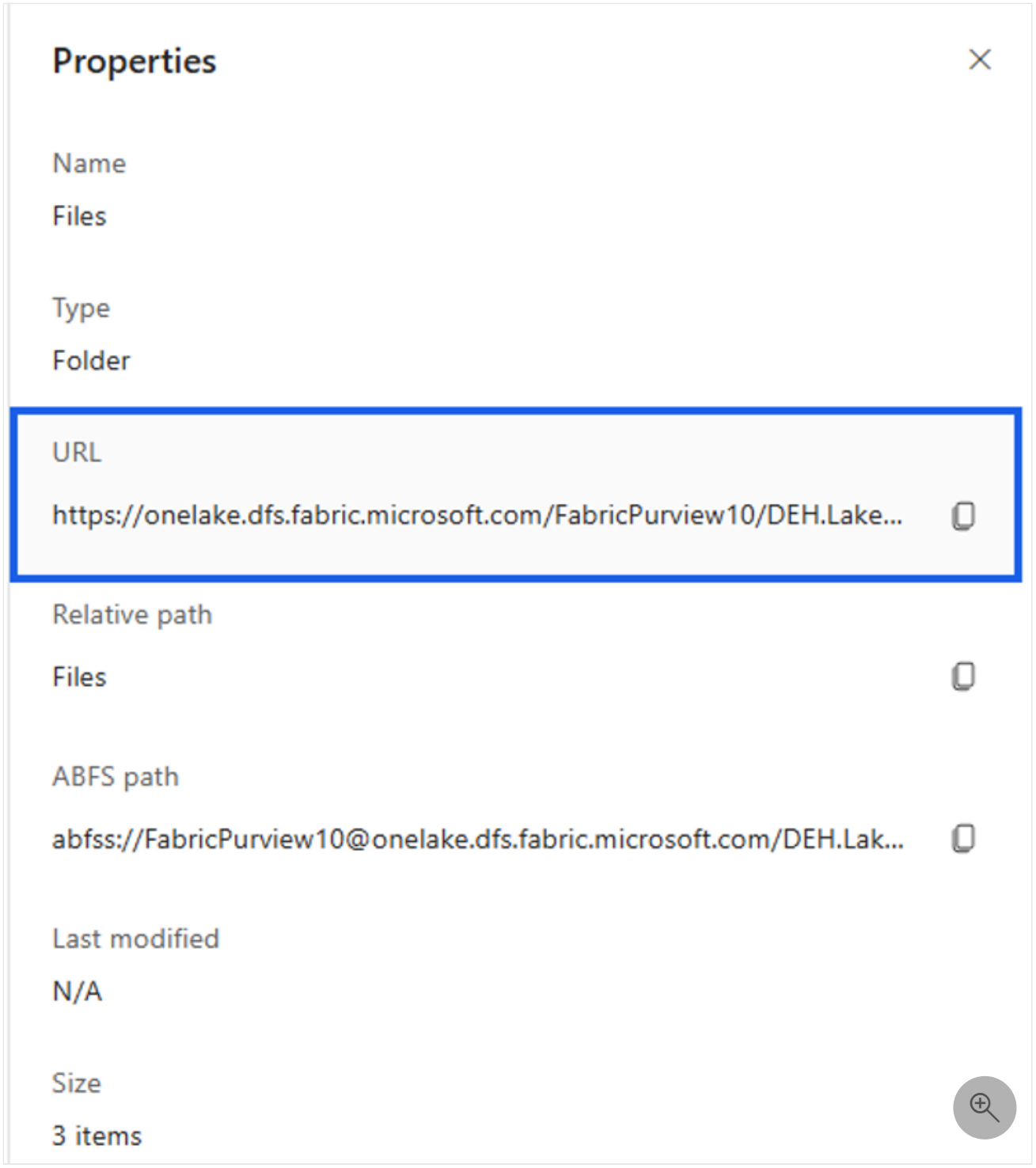
Subscribe Unified Catalog metadata to Fabric OneLake

Subscribe to Purview data governance metadata for analytics and get insights by following these steps:

1. In the Purview portal, select **Settings**, then select **Unified Catalog**, then select **Solution integrations (preview)**.
2. Make sure that you have the Data Governance Administrator role (learn more about roles and permissions).
3. Select **Edit**.
4. Add **Storage type** and **Enable** the setup.
5. Add **Location URL** (example: `https://onelake.dfs.fabric.microsoft.com/workspace name/lakehouse name/Files/purviewmetadata`)
6. Select **Properties** to copy the URL.



7. Copy the URL from the **Properties** page.



8. Add the folder name at the end of the URL. For example, add `/DEH` to the copied URL, as shown in the previous screenshot.

9. Grant contributor access to Purview Managed Service Identity (MSI) for your Fabric workspace.

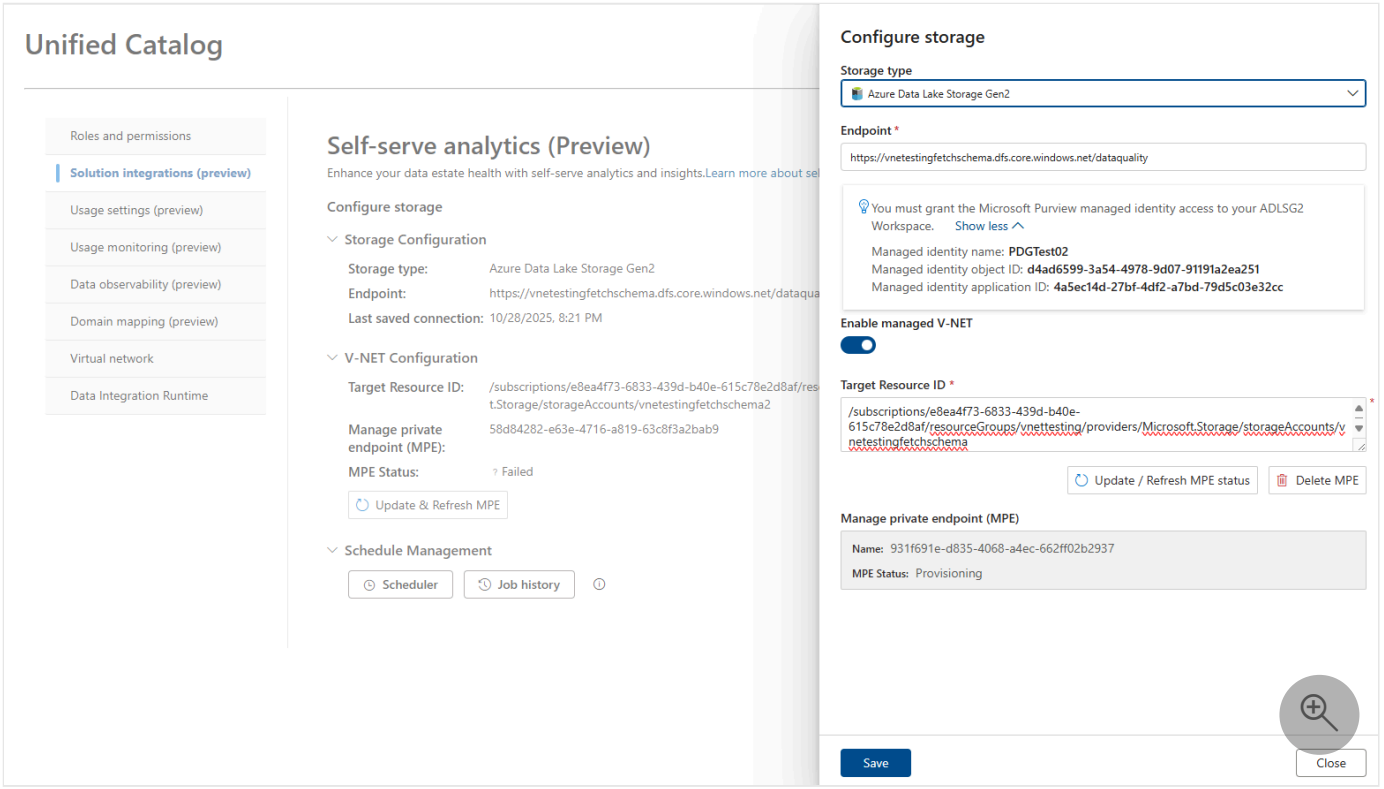
10. Test the connection.

11. Select **Save** to save the configuration and publish Purview metadata to your OneLake workspace.

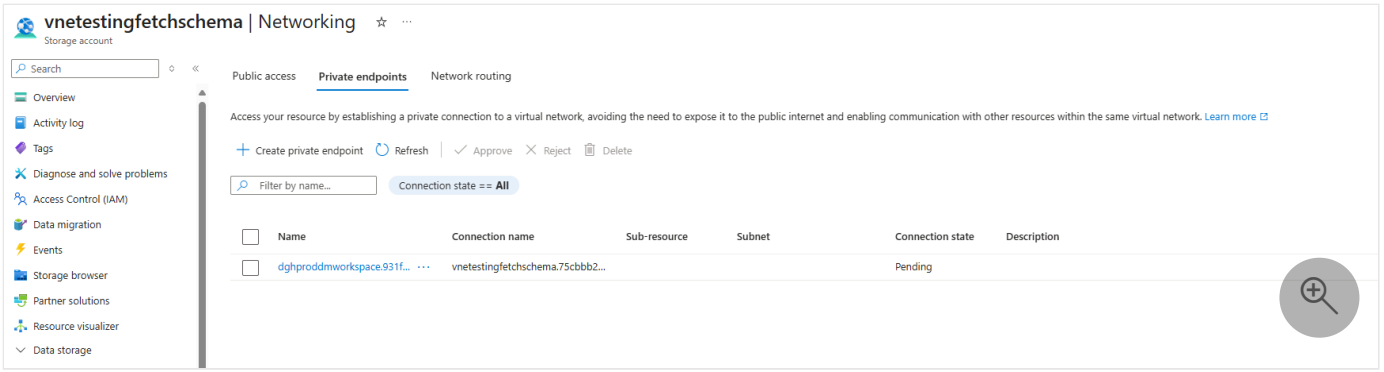
Enable managed V-NET

To enable or disable managed virtual network, follow these steps:

1. In the Purview portal, select **Settings**.
2. Under **Solution settings**, select Unified Catalog, and then select **Solution integrations (preview)**.
3. On **Configure storage**, select the **Enable managed V-NET** toggle to enable managed virtual network for your storage.



1. Enter the **Target Resource ID**. For example, the target resource ID is: `/subscriptions/e2223ea4f73-6833-439d-b40e615c78e2d8af/resourceGroups/vnettesting/providers/`. To copy a resource ID of your storage, select the JSON View of your storage and copy the JSON string.
2. Select **Update/Refresh MPE status** to see the update. The status changes from **Provisioning** to **Pending**.
3. You need to approve the MPE request in the Azure portal to complete the setup. After you approve the request, the MPE status changes from **Pending** to **Approved**.



To disable the virtual network setup, select the **Enable managed V-NET** toggle to turn it off. If you disable it, you need to change non-virtual network storage endpoint.

Remove a configured virtual network

To remove a managed virtual network setup, select **Delete MPE** on **Configure storage**. This action deletes the virtual network configuration and removes it from the **Configure storage** pane. The **Test connection** area reappears, which you can use to test a non-virtual network connection. **Test connection** is only enabled for non-virtual storage setup.

> **Note**
>
> Publishing Unified Catalog metadata to a Fabric workspace isn't supported in Fabric tenants with Managed Private Endpoints enabled. Use a Fabric tenant without Managed Private Endpoints to publish Unified Catalog metadata.

Helpful information about virtual network setup

- Use the Location URL or Endpoint from virtual network storage when you enable managed virtual network to store Purview metadata.
- Use JSON View to copy **Target Resource ID** from your Azure portal to avoid using an incorrect resource group.
- If you deleted the configured virtual network, then make sure that you change the endpoint or Location URL to use non-virtual-network storage and test the connection for non-virtual-network storage.

> **Note**
>
> To store data quality error records, create separate folder locations for each domain. This approach organizes and stores domain-specific data quality error records in dedicated

folders. You must configure data quality error records storage separately. To create data quality error record storage, go to **Data Quality** > **Manage** > **Settings**, select an Azure region, and set up the storage connection. You can't use self-serve storage connection setup for data quality error records storage even if you use identical or same storage location.

Create a semantic model in OneLake

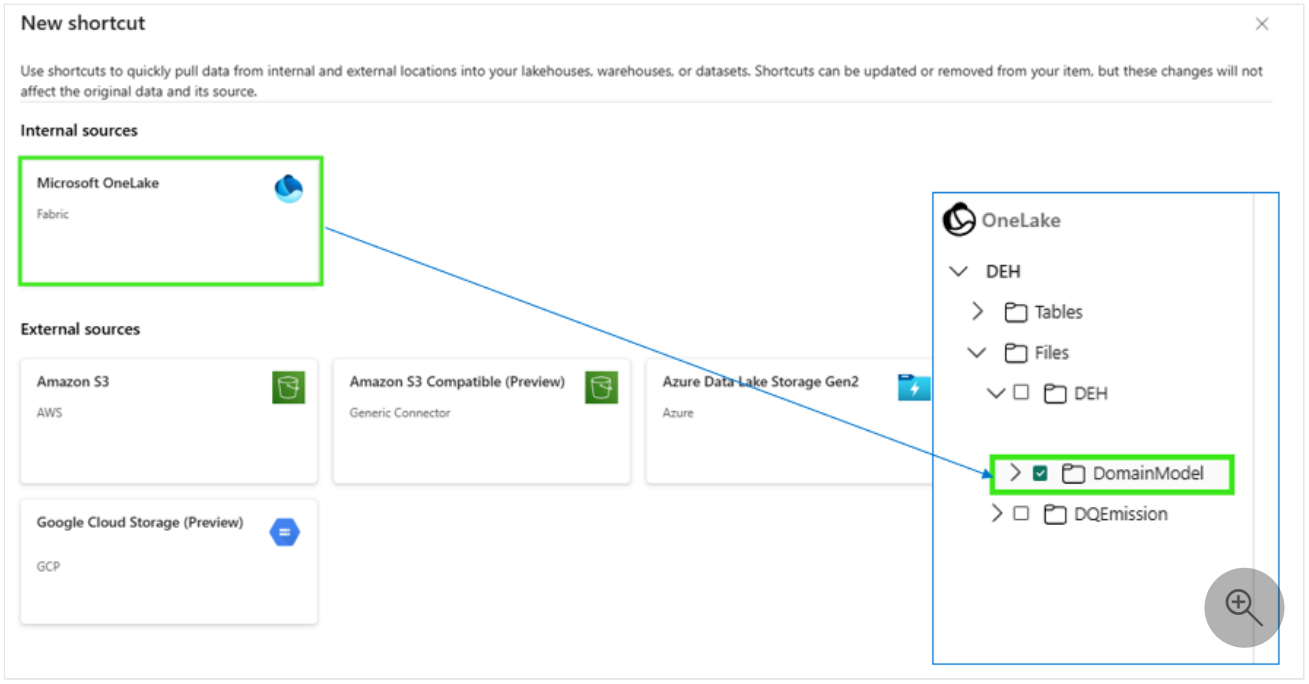
In the context of data and analytics, a semantic model is a structured representation of data that defines the meaning, relationships, and rules within a specific domain. It provides a layer of abstraction that helps users understand and interact with complex data by making it more intuitive and accessible, especially in the context of business intelligence (Power BI) and analytics platforms.

You always need a semantic model before you can build any reports. Within the warehouse, you can add warehouse objects - tables or views - to your default Power BI semantic model. You can also add other semantic modeling properties, such as hierarchies and descriptions. Use these properties to create the Power BI semantic model's tables. You can also remove objects from the default Power BI semantic model.

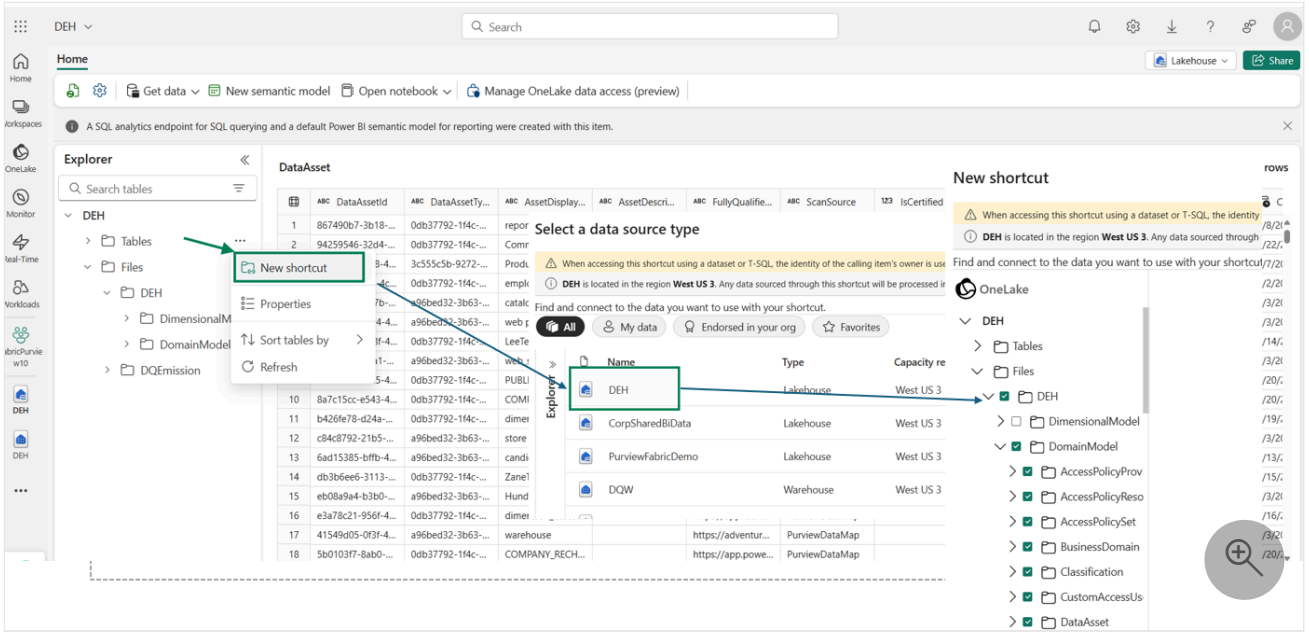
To create a semantic model from the Data Governance metadata domain model:

1. Open the Lakehouse in your Fabric workspace.

2. Use a **shortcut** to create a shortcut of the domain model from OneLake and within OneLake.
   a. Select the **Tables** ellipsis button (...).
   b. Select **New Shortcut** and select OneLake from New Shortcut sources page.
   c. Select the **DomainModel** table to shortcut.

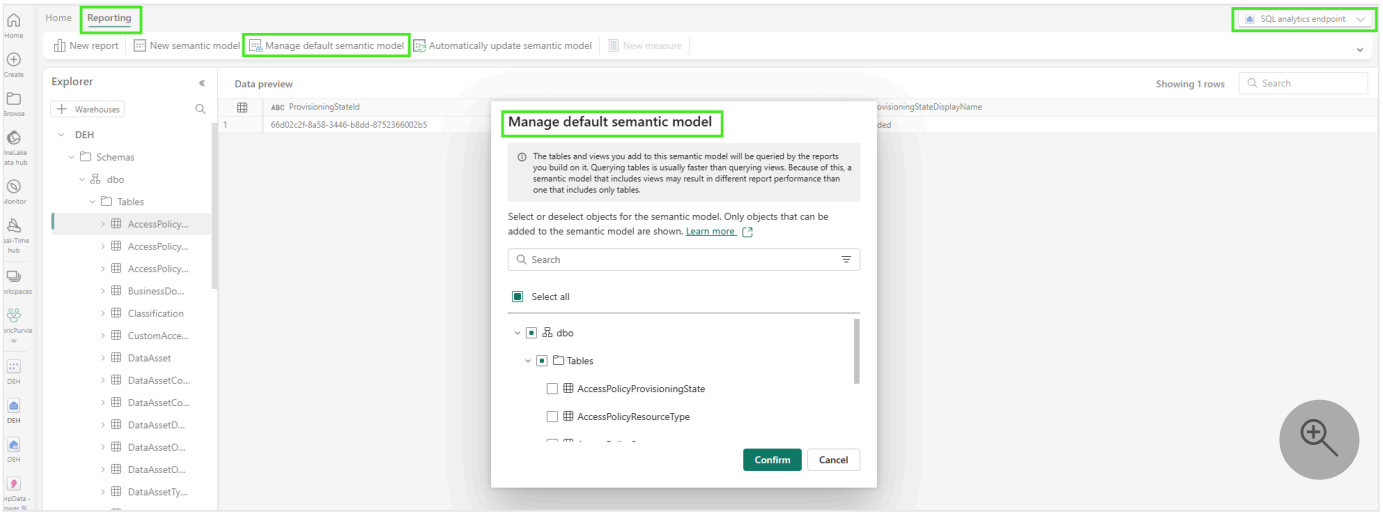


3. Create a shortcut directly at the table level for all files, which eliminates duplicate records.

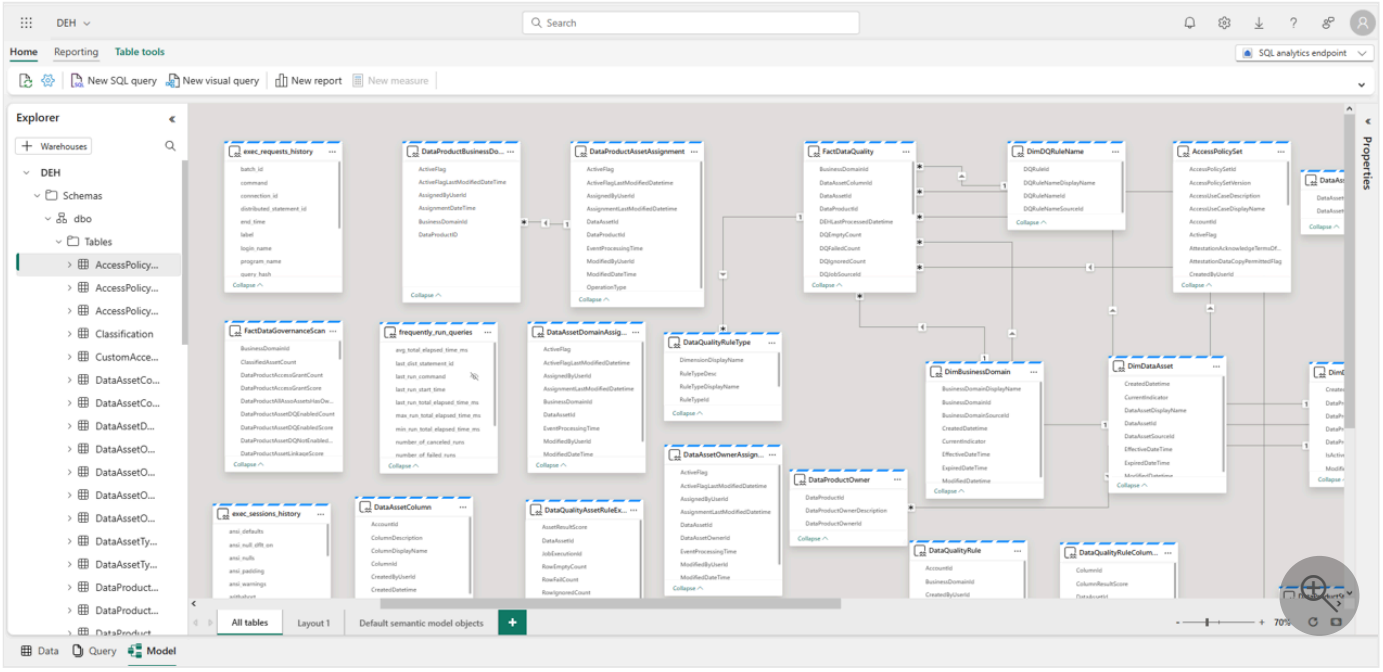


After you publish all files to delta tables via shortcut, you can add the delta tables to semantic model.

1. Switch to SQL analytics endpoint page from the Lakehouse page.
2. Select **Reporting** from top left corner of the **SQL analytics endpoint** page.
3. Select **Manage default semantic model**.
4. Select the tables from dbo > Tables that you want to add to the semantic model for reporting.



To add objects such as tables or views to the default Power BI semantic model, select **Automatically update semantic model**.

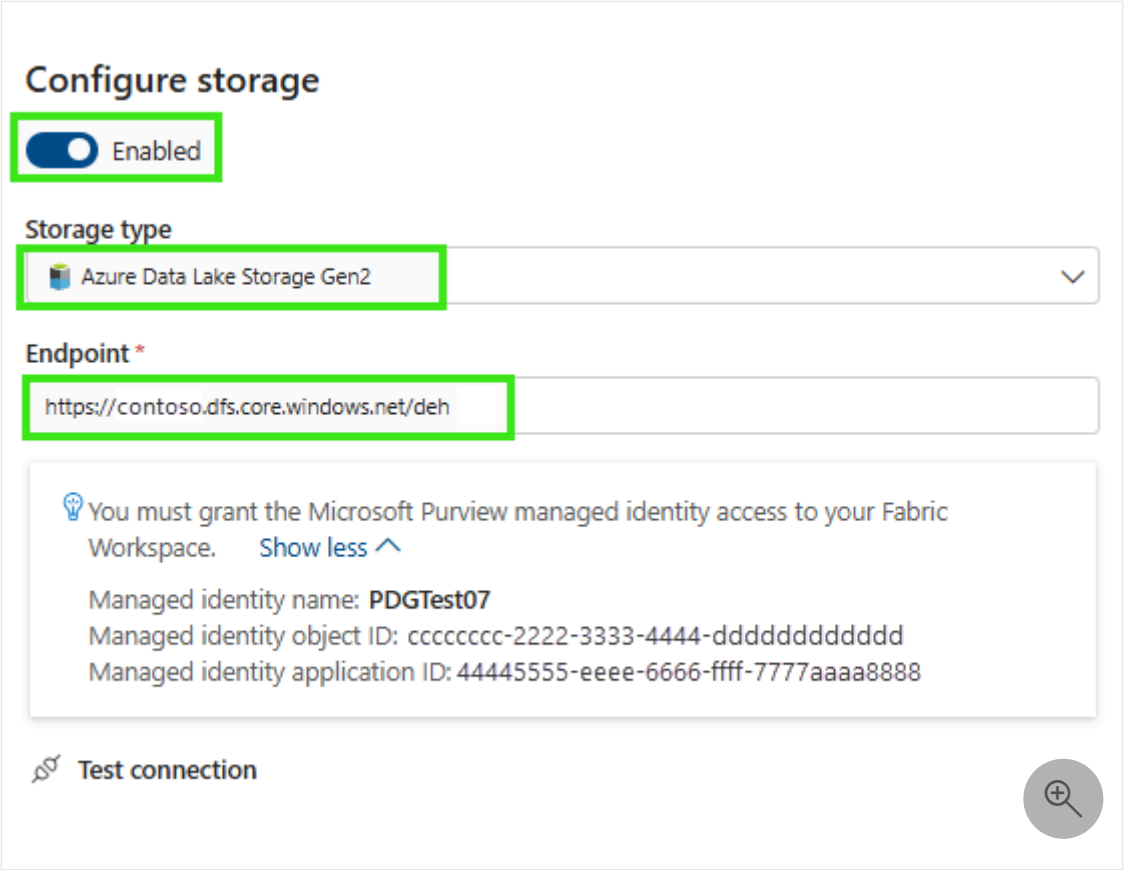


> **Note**
>
> - You can right-click to add related tables to create a relationship.
> - If you prefer not to subscribe Purview metadata for self-serve analytics, you can manually disable self-serve analytics (metadata subscription): go to Solution settings > Unified Catalog > Solution Integrations > Self-serve analytics and select the toggle to disable it. You need the Data Governance Administrator role to enable and disable this feature.

Subscribe to Purview catalog metadata to Data Lake Storage Gen2 storage

To subscribe to Purview's data governance metadata and publish it to your Data Lake Storage Gen2 storage for analytics and insights, follow these steps:

1. In the Purview portal, open **Unified Catalog**.
2. Select **Settings** in the left navigation. Under **Solution settings**, select **Unified Catalog**.
3. Select **Solution integration**, and then select **Edit**.
4. Select **Storage type** and **Enable** the setup.
5. Add **Location URL**. This URL needs to be the Data Lake Storage Gen2 path plus the container name, formatted as Data Lake Storage Gen2 endpoint + "/(container name)".
    a. Go to portal.azure.com.
        i. Select your Data Lake Storage Gen2 storage.
    b. Go to **Settings** > **Endpoints** and select the primary endpoint of your Data Lake Storage Gen2 storage.
6. Grant **Storage Blob Data Contributor** access to Purview Managed Service Identity (MSI) for your Data Lake Storage Gen2 container.
7. Test the connection.

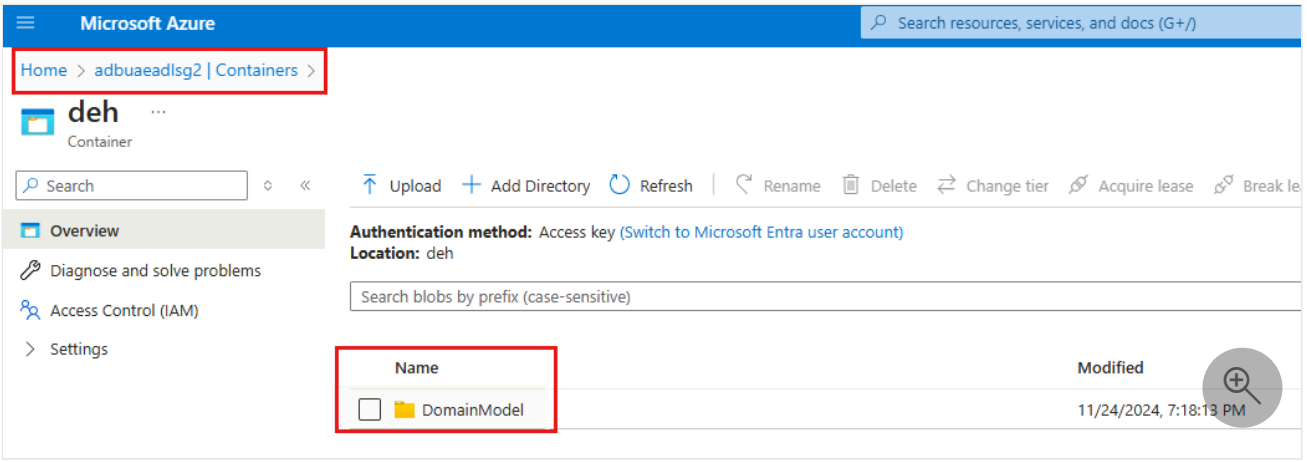


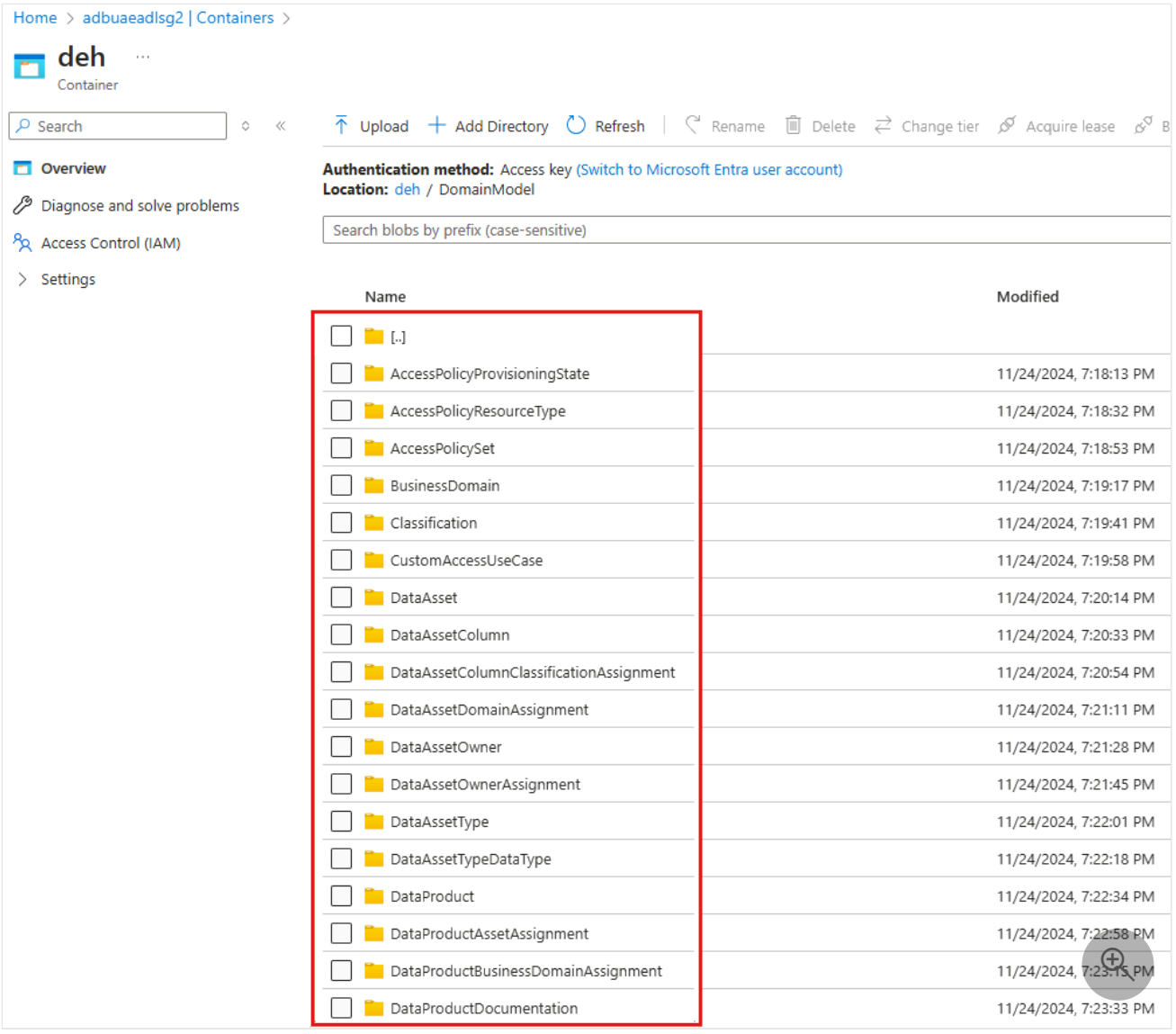
8. Select **Save** to save the configuration and publish the domain model to your Data Lake Storage Gen2 storage.

> **Note**
>
> To store data quality error records, create separate folder locations for each domain. This approach organizes and stores domain-specific data quality error records in dedicated folders.

Review published model and data

1. Open portal.azure.com.
2. Select your Data Lake Storage Gen2 storage.
3. Select the container that you added with the Data Lake Storage Gen2 endpoint in Purview.
4. Browse the list of Delta Parquet files published in the container.
5. Browse published model and metadata (see the following images).





Create a Power BI report

Power BI is natively integrated in the whole Fabric experience. This native integration includes a unique mode, called DirectLake, for accessing the data from the lakehouse to provide the most performant query and reporting experience. DirectLake is a groundbreaking new capability that allows you to analyze large semantic models in Power BI.

With DirectLake, you load parquet-formatted files directly from a data lake without needing to query a data warehouse or lakehouse endpoint, and without needing to import or duplicate data into a Power BI semantic model. DirectLake is a fast path to load the data from the data lake straight into the Power BI engine, ready for analysis.

In traditional DirectQuery mode, the Power BI engine directly queries the data from the source for each query execution, and the query performance depends on the data retrieval speed. DirectQuery eliminates the need to copy data, ensuring that any changes in the source are immediately reflected in query results.

For more information, see how to create a Power BI report in Microsoft Fabric.

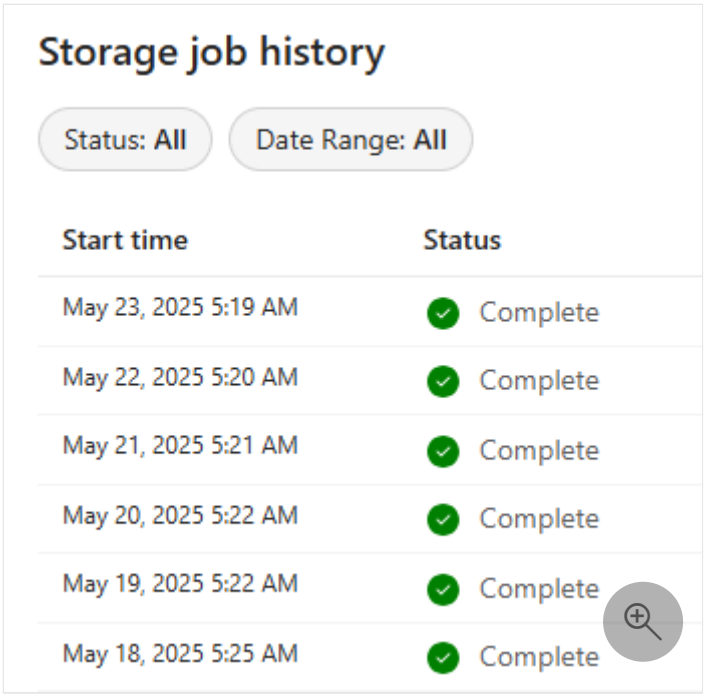
Set up the schedule

Data Governance Administrators can set up the self-serve analytics metadata refreshing schedule by following these steps:

1. In the Purview portal, open **Unified Catalog**.
2. Select **Settings** in the left navigation. Under **Solution settings**, select **Unified Catalog**.
3. Select **Solution integration**, then select the scheduler icon.
4. On **Schedule self-serve analytics**, turn on the **Enabled** toggle, and set the start date, frequency, and end date.
5. Select **Save**.

Job history

You can browse **Job history** by selecting the **Job history** icon on the self-serve analytics page. You can filter job status by Complete or Failed for a date range.



ⓘ Important

- Purview MSI needs contributor access to your Fabric workspace if you're subscribing Purview metadata to publish into your Fabric workspace.
- Purview MSI needs **Storage Blob Data Contributor** access to your Data Lake Storage Gen2 if you're subscribing Purview metadata to publish into your Data Lake Storage Gen2 container.
- Standalone data assets with a Data Quality score aren't published for self-serve analytics. If you associate a standalone data asset with a data product and run Data Quality for it, Purview publishes the associated data asset's metadata for self-serve analytics.
- If a user deletes the Data Quality history for a data asset, the deleted history continues to exist in the self-serve analytics storage (Fabric and Data Lake Storage Gen2). Purview doesn't publish a delete indicator or deletion flag to self-serve analytics storage, making it impossible to distinguish deleted Data Quality scan history from active history records. As a result, consumers of the self-serve analytics data might continue to see historical Data Quality records even after they're deleted from the product experience.
- Purview only publishes governed assets data. Data assets associated with a data product are categorized as governed assets. Data Map assets that aren't governed won't appear in the self-serve analytics data asset table.
- The Catalog implements role-based access control (RBAC), ensuring that not all users can view all domains or data products. However, for self-serve analytics, Purview

publishes all data, allowing anyone with access to this data to view the entire catalog. Access control for self-serve metadata depends on where the data is stored - either the Fabric workspace owner or the Data Lake Storage Gen2 storage owner can manage access.

- If you prefer not to subscribe Purview metadata for self-serve analytics, you can manually disable self-serve analytics (metadata subscription): go to Solution settings > Unified Catalog > Solution Integrations > Self-serve analytics and select the toggle to disable it. You need the Data Governance Administrator role to enable and disable this feature.

Related content

- OneLake shortcuts
- Fabric Data Activator
- What is data activator
- Build automation with PowerAPP
- Entity relationship diagram for domain model

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/23/2026

Disaster recovery for Unified Catalog (manual)

This article provides guidance on backup and recovery strategy when using Microsoft Purview Unified Catalog in production deployment. The scope of this article is to cover manual business continuity and disaster recovery (BCDR) methods that you can set up using APIs.

Features covered

The following areas are covered in this BCDR strategy:

- Unified Catalog foundation: Governance Domains, Glossary Terms, and Data Assets — the foundational catalog that organizes and describes an organization's data estate.
- Data products: Curated, business-ready data sets with ownership, documentation, terms of use, and audience metadata.
- Critical data elements: Regulated or high-value data elements (for example, personal data, financial identifiers) with lifecycle management (Draft → Published → Expired).
- Data quality: Quality rules and asset metadata that track data health and conformance across the estate.
- Data access: Access request workflows, multi-stage approval chains, policy sets, and permitted use-case definitions for governed data sharing.
- Data estate health: CDMC control templates and analytics schedules that continuously measure governance maturity and compliance posture.
- Workflows: Configurable multi-stage approval workflows including manager, privacy, and approver stages for data access governance.

BCDR workflow

Step 1. Create an Enterprise tier account

Create a Purview Data Governance account at the **Enterprise** tier. This serves as the **primary account** with full governance capabilities.

Step 2. Request Classic Purview provisioning in a secondary region

Create a **support ticket** through the Customer Support channel requesting enablement of Classic Purview account provisioning in a **secondary (paired) region** for BCDR purposes.

Step 3. Create a Classic Purview account

After Classic provisioning is enabled in the secondary region, create a **Classic Purview account**. This account serves as the **secondary (backup) account** with limited capabilities.

Step 4. Copy data from the Enterprise account to the Classic account

Using the available BCDR APIs (described in the API details section), copy governance data from the **Enterprise account (primary)** to the **Classic account (secondary)**. This includes:

- Governance domains
- Data products
- Glossary terms
- Critical data elements
- Data assets
- Data Quality metadata
- Data estate health configuration
- Data access workflows and policies

Step 5. Manage GUIDs during backup

You're responsible for **managing and mapping globally unique identifiers (GUIDs)** during the backup process. Because the secondary account generates new identifiers, you must maintain a mapping between primary and secondary GUIDs to ensure data consistency and traceability.

Step 6. Validate the secondary account

Use the **GET APIs** for each feature area to validate that all governance data is replicated to the secondary account:

- `GET /catalog/businessdomains`
- `GET /catalog/dataProducts`
- `GET /catalog/terms`
- `GET /catalog/criticalDataElements`
- `GET /catalog/dataAssets`
- `GET /quality/bulk-asset-metadata`

- `GET /health/analytics/schedule`
- `GET /dataaccess/workflows`

Step 7. Fail over during a disaster recovery (DR) event

When a disaster recovery event occurs, create a **support ticket** requesting:

1. **Promote** the Secondary (Classic) account to **Enterprise** tier
2. **Demote** the Primary (Enterprise) account to **Secondary**

This completes the failover, and the previously secondary account becomes the new primary with full Enterprise capabilities.

API details

1. Account and provisioning

Get current policies

Verify the BCDR account is operational by listing the current user's policies.

HTTP

```
GET https://<accountName>.purview.azure.com/datagovernance/catalog/policies
```

> ⓘ Note
>
> Account provisioning is performed with **Is BCDR = true**. Acquire tokens by using standard Microsoft Entra ID/OAuth flows for Purview REST APIs.

---

2. Governance domains

Create governance domain

Create a governance domain:

HTTP

```
POST https://<accountName>.purview.azure.com/datagovernance/catalog/businessdomains
```

Request body:

JSON

```
{
  "name": "BCDR - My Business Domain",
  "type": "DataDomain",
  "status": "Draft",
  "description": "Optional description up to 200k chars",
  "isStagingDomain": false,
  "managedAttributes": []
}
```

Response body:

JSON

```
{
  "status": "Draft",
  "type": "DataDomain",
  "name": "BCDR - My Business Domain",
  "description": "Optional description up to 200k chars",
  "isRestricted": false,
  "isStagingDomain": false
}
```

Get governance domains

List the governance domains to confirm replication:

HTTP

```
GET https://<accountName>.purview.azure.com/datagovernance/catalog/businessdomains
```

---

3. Data products

Create data product

Create a data product under a governance domain:

HTTP

```
POST https://<accountName>.purview.azure.com/datagovernance/catalog/dataProducts
```

Request body:

JSON

```
{
  "name": "BCDR - Sales Analytics",
  "domain": "<businessDomainId>",
  "status": "Draft",
  "type": "Analytical",
  "endorsed": false,
  "contacts": {
    "owner": [
      {
        "id": "<guidObjectId>",
        "info": "<email>"
      }
    ]
  },
  "description": "Aggregated sales data across regions",
  "businessUse": "Used for quarterly reporting",
  "audience": [
    "DataAnalyst",
    "BusinessUser"
  ],
  "updateFrequency": "Monthly",
  "termsOfUse": [
    {
      "name": "Data Policy",
      "url": "https://contoso.com/policy"
    }
  ],
  "documentation": [
    {
      "name": "User Guide",
      "url": "https://contoso.com/guide"
    }
  ]
}
```

Response body:

JSON

```
{
  "status": "Draft",
  "type": "Analytical",
  "name": "BCDR - Sales Analytics",
  "description": "Aggregated sales data across regions",
  "businessUse": "Used for quarterly reporting",
```

```
  "updateFrequency": "Monthly",
  "endorsed": false
}
```

Get data products

List the data products:

HTTP

```
GET https://<accountName>.purview.azure.com/datagovernance/catalog/dataProducts
```

---

4. Glossary terms

Create term

Create a glossary term under a governance domain:

HTTP

```
POST https://<accountName>.purview.azure.com/datagovernance/catalog/terms
```

Request body:

JSON

```
{
  "name": "BCDR - Customer Revenue",
  "domain": "<businessDomainId>",
  "status": "Draft",
  "description": "Total revenue attributed to a customer",
  "acronyms": [
    "CR",
    "CustRev"
  ],
  "contacts": {
    "owner": [
      {
        "id": "<guidObjectId>",
        "info": "<email>"
      }
    ],
    "expert": [
      {
        "id": "<guidObjectId>",
        "info": "<email>"
      }
```

```
    ]
  },
  "resources": [
    {
      "name": "Wiki",
      "url": "https://contoso.com/wiki/customer-revenue"
    }
  ]
}
```

Response body:

JSON

```
{
  "status": "Draft",
  "name": "BCDR - Customer Revenue",
  "description": "Total revenue attributed to a customer",
  "contacts": {
    "owner": [],
    "expert": []
  },
  "isMigrated": false
}
```

> **Note**
>
> The API response may return empty `contacts` arrays even when contacts are provided in the request. Contacts are linked asynchronously and will appear on subsequent GET calls.

Get terms

List the glossary terms:

HTTP

```
GET https://<accountName>.purview.azure.com/datagovernance/catalog/terms
```

---

5. Critical data elements (CDE)

Create critical data element

Create a critical data element under a governance domain:

HTTP

```
POST
https://<accountName>.purview.azure.com/datagovernance/catalog/criticalDataElements
```

Request body:

JSON

```
{
  "name": "BCDR - Customer SSN",
  "domain": "<businessDomainId>",
  "status": "Draft",
  "dataType": "DateTime",
  "description": "Social Security Number – regulated PII element",
  "contacts": {
    "owner": [
      {
        "id": "<guidObjectId>",
        "info": "<email>"
      }
    ],
    "expert": [
      {
        "id": "<guidObjectId>",
        "info": "<email>"
      }
    ]
  }
}
```

Response body:

JSON

```
{
  "status": "Draft",
  "dataType": "DateTime",
  "name": "BCDR - Customer SSN",
  "description": "Social Security Number - regulated PII element"
}
```

> **CDE Data Types:** Boolean, DateTime, Number, Text **CDE Status Lifecycle:** Draft → Published → Expired

Get critical data elements

List the critical data elements:

HTTP

```
GET
https://<accountName>.purview.azure.com/datagovernance/catalog/criticalDataElements
```

---

6. Data assets

Create data asset

Create a data asset that references a source:

HTTP

```
POST https://<accountName>.purview.azure.com/datagovernance/catalog/dataAssets
```

Request body:

JSON

```
{
  "name": "gold",
  "type": "ADLSGen2Path",
  "source": {
    "type": "PurviewDataMap",
    "accountName": "BCDRSecondary",
    "assetId": "<assetGuid>",
    "assetType": "azure_datalake_gen2_filesystem",
    "fqn": "https://<storageAccount>.dfs.core.windows.net/gold"
  },
  "typeProperties": {
    "serverEndpoint": "<storageAccount>.dfs.core.windows.net",
    "container": "gold"
  },
  "contacts": {
    "owner": [
      {
        "id": "<guidObjectId>",
        "info": "<email>"
      }
    ]
  }
}
```

Response body:

JSON

```
{
  "name": "gold",
  "status": "ACTIVE",
  "displayText": "gold",
  "source": {
    "type": "PurviewDataMap",
    "assetId": "<assetGuid>"
  }
}
```

Get data assets

HTTP

```
GET https://<accountName>.purview.azure.com/datagovernance/catalog/dataAssets
```

---

7. Data estate health

CDMC control template

CDMC (Cloud Data Management Capabilities) control templates are **auto-created** when the BCDR account is provisioned. A global analytics schedule is enabled by default (daily frequency).

Create or update analytics schedule

Create or update the schedule that controls how often data estate health is measured:

HTTP

```
PUT https://<accountName>.purview.azure.com/datagovernance/health/analytics/schedule?api-version=2024-02-01-preview
```

Request body:

JSON

```
{
  "frequency": "Hour",
  "interval": 6,
  "startTime": "2026-03-18T00:00:00Z",
  "status": "Enabled",
  "timeZone": "UTC"
}
```

Response body:

JSON

```
{
  "frequency": "Hour",
  "interval": 6,
  "status": "Enabled",
  "timeZone": "UTC"
}
```

Get analytics schedule

Retrieve the current analytics schedule:

HTTP

```
GET https://<accountName>.purview.azure.com/datagovernance/health/analytics/schedule?api-version=2024-02-01-preview
```

---

8. Data access

Get default access request template

Retrieve the default data access request template:

HTTP

```
GET https://<accountName>.purview.azure.com/datagovernance/dataaccess/workflows/templates/dataAccessRequestTemplate
```

Create data access workflow

Create or update a data access workflow:

HTTP

```
PUT https://<accountName>.purview.azure.com/datagovernance/dataaccess/workflows/{workflowId}
```

The workflow request body is a **large JSON directed acyclic graph (DAG)** that defines multi-stage approvals including:

- `DataSubscriptionWebhook`: subscription trigger
- `DataAccessRequestSubmitted`: entry trigger
- **Manager approval** stage
- **Privacy / compliance approval** stage
- **Approver chain** stages
- **Approve / Reject** terminal actions

Apply policy sets to data product

Apply policy sets to a data product to govern access:

HTTP

```
PUT
https://<accountName>.purview.azure.com/datagovernance/dataaccess/dataProducts/{dataProductId}/policySets/applied?api-version=2023-10-01-preview
```

Request body:

JSON

```
{
  "policies": {
    "approvers": [
      {
        "identityType": "User",
        "objectId": "<userObjectId>",
        "tenantId": "<tenantId>"
      }
    ],
    "permittedUseCases": [
      {
        "title": "Assessing Fit",
        "description": "Use to determine if data is useful"
      },
      {
        "title": "Critical Reporting",
        "description": "Use to create high quality reports"
      },
      {
        "title": "Directional Insights",
        "description": "Use to gather directional insights"
      }
    ],
    "managerApprovalRequired": false,
```

```
    "privacyComplianceApprovalRequired": false,
    "dataCopyPermitted": false,
    "attestations": [],
    "skipWorkflow": false,
    "maximumAccessDuration": {
      "durationType": "Years",
      "length": 1
    }
  }
}
```

List workflows

List the data access workflows:

HTTP

```
GET https://<accountName>.purview.azure.com/datagovernance/dataaccess/workflows
```

Get approvals

Retrieve pending data access approvals:

HTTP

```
GET https://<accountName>.purview.azure.com/datagovernance/dataaccess/approvals?
approvalStatus=Pending&api-version=2023-10-01-preview
```

Get workflow runs

Retrieve the data access workflow runs:

HTTP

```
GET https://<accountName>.purview.azure.com/datagovernance/dataaccess/workflowRuns?
ignoreActionStatuses=true&api-version=2023-10-01-preview
```

---

9. Data Quality

Create bulk asset metadata

Create bulk asset metadata for data quality:

```
```

HTTP

```
PUT https://<accountName>.purview.azure.com/datagovernance/quality/bulk-asset-metadata
```

> **Note**
>
> The request body references an external JSON file (`BCDR-DQ-Create Metadata.json`). Contact your support representative for the complete payload.

Get bulk asset metadata

Retrieve the bulk asset metadata:

HTTP

```
GET https://<accountName>.purview.azure.com/datagovernance/quality/bulk-asset-metadata
```

Related content

- Disaster recovery and migration best practices for Microsoft Purview data governance (classic).

Note: The author created this article with assistance from AI. Learn more

Last updated on 06/29/2026

Application card: Microsoft Purview Unified Catalog

What is an Application or Platform Card?

Microsoft's Application and Platform cards are intended to help you understand how our AI technology works, the choices application owners can make that influence application performance and behavior, and the importance of considering the whole application, including the technology, the people, and the environment. Application cards are created for AI applications and platform cards are created for AI platform services. These resources can support the development or deployment of your own applications and can be shared with users or stakeholders impacted by them.

As part of its commitment to responsible AI, Microsoft adheres to six core principles: fairness, reliability and safety, privacy and security, inclusiveness, transparency, and accountability. These principles are embedded in the Responsible AI Standard, which guides teams in designing, building, and testing AI applications. Application and Platform Cards play a key role in operationalizing these principles by offering transparency around capabilities, intended uses, and limitations. For further insight, readers are encouraged to explore Microsoft's Responsible AI Transparency Report and Code of Conduct, which outline how enterprise customers and individuals can engage with AI responsibly.

Overview

Microsoft Purview Unified Catalog is a data governance solution that enables business experts and technical data owners to collaborate and contribute to a shared understanding of data. Unified Catalog enables inventory of metadata and provides a framework for the governance of metadata and the use of the underlying data. Unified Catalog helps organizations create value from data, manage access to data, improve data quality, and monitor data health. Unified Catalog is intended for enterprise customers and their downstream users, including business users, analysts, data scientists, stewards, and governance leaders. Learn about Unified Catalog capabilities and implementation guidance.

AI capabilities are embedded in selected capabilities to improve discovery and curation efficiency. These capabilities include natural language data product search (preview), suggested data asset mapping (preview), and suggested data quality rules (preview).

Key terms

The following list provides a glossary of key terms related to Unified Catalog:

- **Data asset**: A technical data resource indexed in the catalog, such as a table, file, or report. Data assets are the building blocks that can be grouped and governed through higher-level business concepts.
- **Data product**: A curated grouping of related data assets with a defined business use case, ownership, and governance context. Data products are designed to improve discoverability, reuse, and controlled access.
- **Data quality rule**: A configurable rule that evaluates data quality dimensions such as completeness, consistency, conformity, accuracy, freshness, and uniqueness. Rules can be out-of-box, custom, or AI-suggested in supported preview workflows.
- **Governance domain**: A business-aligned boundary used to organize data and governance responsibilities. Domains help scale stewardship and simplify discovery by aligning data with business context.
- **Metadata**: Descriptive information about data, including technical properties, business definitions, lineage, classifications, and usage context. Unified Catalog uses metadata to power search, governance workflows, and health insights.
- **Natural language search (preview)**: An AI-enabled search mode that allows users to describe the data they need in plain language. The feature returns ranked data product results based on metadata and business context.

Key features or capabilities

The key features and capabilities outlined here describe what Unified Catalog is designed to do and how it performs across supported tasks.

- **Business-aligned data organization**: Unified Catalog organizes information through governance domains, data products, glossary terms, and related concepts. This structure helps users understand data in business context rather than only by technical source or schema. It also supports distributed stewardship so teams can manage data where domain expertise exists.
- **Natural language data product discovery (preview)**: Users can search for data products by describing intent in plain language, rather than relying only on exact keywords. The feature currently supports data products in English, French, and Spanish, with corresponding search terms. This improves data discovery for semi-technical users who may not know internal naming conventions.

- **AI-assisted data product curation (preview)**: During data product creation, Unified Catalog can suggest relevant data assets based on metadata such as names, descriptions, and use cases. Suggestions are prioritized so owners can review the most relevant candidates first. The experience accelerates curation while keeping final selection under human control.
- **AI-assisted data quality rule generation (preview)**: Data quality stewards can use AI-generated rule suggestions as a starting point for rule authoring. Suggested rules are informed by profiling metadata and are meant to be reviewed and tested before use. This reduces setup time for quality programs while preserving expert oversight.
- **Integrated access governance workflows**: Data product access policies support configurable approvals and attestations. Policy inheritance from governance domains, glossary terms, and critical data elements helps maintain consistent control at scale. This balances data accessibility with compliance and right-use expectations.
- **Data health and monitoring capabilities**: Unified Catalog provides health management actions, data quality scoring, and reporting to track governance progress. Teams can monitor quality, identify gaps, and prioritize remediation. These controls support continuous improvement instead of one-time governance exercises.

Intended uses

Unified Catalog can be used in multiple scenarios across a variety of industries. Some examples of use cases include:

- **Enterprise analytics readiness in financial services**: A bank can create curated data products for risk, finance, and customer analytics teams so analysts can find trusted data faster. Governance domains and access policies help enforce right-use and approval paths for regulated data. Natural language search can reduce the time it takes to discover relevant data products when users do not know technical table names. The result is quicker analysis with stronger governance controls.
- **Cross-functional reporting in government organizations**: A public sector organization can organize data products by service area, such as benefits, operations, and citizen support. Teams can use Unified Catalog to discover approved datasets, request access through policy workflows, and understand business definitions before use. This improves consistency across departments while preserving accountability and review requirements. It also helps reduce duplicated data preparation effort.
- **Data quality improvement in education systems**: A university can apply data quality rules to critical operational datasets used for enrollment, course planning, and student outcomes reporting. AI-suggested rules can speed up initial setup, while stewards validate and tune

rules for institutional standards. Quality scores and actions provide visibility into where remediation is needed. This supports more reliable reporting and decision making.

- **Consumer goods demand planning and operations**: A retail organization can publish data products for sales, inventory, and supply chain teams with clear ownership and terms of use. Teams can reuse curated data instead of repeatedly rebuilding similar datasets. Governance and quality controls help maintain trust as data is shared across planning, merchandising, and operations functions. The approach supports both day-to-day reporting and longer-term planning.

Models and training data

Unified Catalog leverages a variety of AI models to power the experience that users see. Some examples include Azure OpenAI GPT and embedding models offered by Foundry Models sold by Azure. To learn more about the data used to train the foundation models behind Unified Catalog, refer to the model details in Foundry.

- Foundry Models sold by Azure
- Responsible AI practices at Microsoft

Performance

Unified Catalog is designed to perform reliably in enterprise data governance scenarios where organizations have scanned and curated metadata in Microsoft Purview Data Map and have configured role-based access correctly. In this environment, users can discover data products, review business context, and request access through policy workflows with clear ownership and governance boundaries.

The application's intended input modalities are primarily structured metadata and user-entered text. Structured inputs include data asset metadata, governance domains, glossary mappings, access policy settings, and data quality definitions. Text inputs include keyword search, natural language search prompts (preview), business descriptions, and use-case narratives entered by stewards and owners. Expected outputs include ranked discovery results, suggested data assets and quality rules in supported preview flows, policy workflow states, and governance health indicators such as quality scores and remediation actions.

For multilingual capabilities, natural language data product search (preview) is designed and documented for English, French, and Spanish data products and corresponding search terms. Organizations operating outside these supported language conditions should validate effectiveness for their specific content and user patterns before broad deployment.

Performance also depends on governance maturity and metadata quality. Experiences such as discovery relevance and recommendation usefulness generally improve when data products are well described, business concepts are curated, and ownership is clearly defined. Organizations should expect better outcomes when they apply iterative stewardship practices and maintain up-to-date metadata.

Limitations

Understanding Unified Catalog's limitations is crucial to determine if it is used within safe and effective boundaries. While we encourage customers to leverage Unified Catalog in their innovative solutions or applications, it's important to note that Unified Catalog was not designed for every possible scenario. We encourage users to refer to either the Microsoft Enterprise AI Services Code of Conduct (for organizations) or the Code Conduct section in the Microsoft Services Agreement (for individuals) as well as the following considerations when choosing a use case:

- **Domain scope constraints**: AI features in Unified Catalog are designed for data governance scenarios, not general-purpose knowledge tasks. Using these features outside governance context can reduce response relevance or completeness. Teams should align prompts and workflows to cataloged metadata and business governance objectives. This helps avoid misuse and improves reliability.
- **Language and input constraints**: Natural language data product search (preview) supports data products created in English, French, and Spanish, and search terms in the corresponding languages. Results for unsupported languages or highly ambiguous prompts can be less reliable. Long text inputs might also be difficult for some AI-assisted workflows to process effectively. Users should provide clear, scoped prompts and validate outputs.
- **Human review requirement for AI suggestions**: AI-generated data asset mappings and suggested data quality rules are intended as starting points. They can accelerate setup, but they don't replace expert review and testing. Applying suggestions without validation can introduce quality or governance gaps. Human oversight is required before operational use.

Evaluations

Performance and safety evaluations assess whether AI applications are operating reliably and securely by examining factors like groundedness, relevance, and coherence while identifying the risks of generating harmful content. The following evaluations were conducted with safety components already in place, which are also described in Safety components and mitigations.

Performance and quality evaluations

Performance evaluations for AI applications are essential to improving their reliability in real-world applications. Metrics like groundedness, relevance, and coherence help assess the accuracy and consistency of AI-generated outputs, so that they are factually supported in grounded content scenarios, contextually appropriate, and logically structured. For Unified Catalog, we conducted performance evaluations for the following metrics, which are available through Microsoft Foundry:

- Groundedness
- Coherence
- Fluency
- Similarity

Risk and safety evaluations

Evaluating potential risks associated with AI-generated content is essential for safeguarding against content risks with varying degrees of severity. This includes evaluating an AI application's predisposition towards generating harmful content or testing vulnerabilities to jailbreak attacks. For [application/platform], we conducted risk and safety evaluations for the following metrics available through Microsoft Foundry:

- Hate and unfairness
- Sexual
- Violence
- Self-harm
- Protected material
- Indirect jailbreak
- Direct jailbreak
- Code vulnerability
- Ungrounded attributes

Evaluation data for safety and quality

Our evaluation data is custom-built to assess AI application performance across key areas of **safety** and **quality**, simulating real-world scenarios and risks. We begin by identifying relevant evaluation aspects of concern based on multi-disciplinary research and expert input. These concerns are translated into targeted evaluation objectives and guide formulation of evaluation metrics. For **safety**, we create adversarial prompts to elicit undesirable or edge-case responses,

which are then scored using AI-assisted annotators trained to assess alignment with Microsoft's safety standards. For **quality**, we craft rubric-based prompts relevant to scenarios including evaluating retrieval-augmented generation (RAG) applications and agents. Datasets are curated from diverse sources including synthetic and public datasets to simulate real-world user scenarios. Using the curated datasets, both evaluations undergo iterative refinement and human alignment to improve metric efficacy and reliability. This methodology forms the foundation of repeatable, rigorous assessments that reflect how customers use evaluations to build better and safer AI.

Custom evaluations

AI features in Unified Catalog underwent various testing before release, including red teaming to identify failure modes and off-scope behaviors. These evaluations focus on whether AI-assisted experiences remain aligned to intended data governance uses and whether they avoid generating responses that conflict with Microsoft AI principles. For current documented AI capabilities, the primary evaluated modality is text.

Evaluation methods include scenario-based testing of search and recommendation quality, plus risk-focused testing of edge-case prompts. Ideal outcomes are contextually relevant, governance-aligned outputs that help users complete discovery and curation tasks with clear next steps. Suboptimal outcomes include irrelevant recommendations, incomplete responses, or outputs that require substantial correction.

Safety components and mitigations

- **Human-in-the-loop validation for AI features**: AI-generated suggestions in preview are intended to be reviewed and tested by users before adoption. This mitigation lowers the risk of overreliance on automated recommendations and supports accountable decision making. Stewards and owners remain responsible for final curation and policy decisions.
- **Feedback and issue reporting loops**: Users can provide feedback on AI outputs that are inaccurate, incomplete, or unclear, including unacceptable results. Feedback helps Microsoft improve AI application.

Best practices for deploying and adopting Unified Catalog

Responsible AI is a shared commitment between Microsoft and its customers. While Microsoft builds AI applications with safety, fairness, and transparency at the core, customers play a critical

role in deploying and using these technologies responsibly within their own contexts. To support this partnership, we offer the following best practices for deployers and end users to help customers implement responsible AI effectively.

Deployers and end-users (admins) should:

- **Exercise caution and evaluate outcomes when using Unified Catalog for consequential decisions or in sensitive domains**: Consequential decisions are those that may have a legal or significant impact on a person's access to education, employment, financial platforms, government benefits, healthcare, housing, insurance, legal platforms, or that could result in physical, psychological, or financial harm. Sensitive domains—such as financial platforms, healthcare, and housing—require particular care due to the potential for disproportionate impact on different groups of people. When using AI for decisions in these areas, make sure that impacted stakeholders can understand how decisions are made, appeal decisions, and update any relevant input data.

- **Evaluate legal and regulatory considerations**: Customers need to evaluate potential specific legal and regulatory obligations when using any AI platforms and solutions, which may not be appropriate for use in every industry or scenario. Additionally, AI platforms or solutions are not designed for and may not be used in ways prohibited in applicable terms of service and relevant codes of conduct.

End-users should:

- **Exercise human oversight when appropriate**: Human oversight is an important safeguard when interacting with AI applications. While we continuously improve our AI applications, AI might still make mistakes. The outputs generated may be inaccurate, incomplete, biased, misaligned, or irrelevant to your intended goals. This could happen due to various reasons, such as ambiguity in the inputs or limitations of the underlying models. As such, users should review the responses generated by Unified Catalog and verify that they match their expectations and requirements.

- **Be aware of the risk of overreliance**: Overreliance on AI happens when users accept incorrect or incomplete AI outputs, mainly because mistakes in AI outputs may be hard to detect. For the end-user, overreliance could result in decreased productivity, loss of trust, application abandonment, financial loss, psychological harm, physical harm, among others. (e.g. a doctor accepts an incorrect AI output).

Learn more about Unified Catalog

For additional guidance or to learn more about the responsible use of Unified Catalog, we recommend reviewing the following documentation:

- Unified Catalog overview
- Plan for Unified Catalog with best practices
- Search for data products in Unified Catalog
- Create and manage data products in Unified Catalog
- Manage data product access policies
- Overview of data quality in Unified Catalog
- Microsoft Purview privacy statement
- Microsoft Trust Center

Learn more about responsible AI

- Microsoft AI principles
- Microsoft responsible AI resources
- Microsoft Azure Learning courses on responsible AI

ⓘ **Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/01/2026

Microsoft Purview Data Map overview

Microsoft Purview Data Map is the foundation for data discovery and data governance in Microsoft Purview. It captures metadata about data in analytics, software-as-a-service (SaaS), and operational systems across hybrid, on-premises, and multicloud environments. Built-in scanning and classification keep the Data Map up to date.

This overview helps Microsoft Purview admins plan capacity and monitor Data Map usage and costs. Every Microsoft Purview account has a Data Map that starts at one capacity unit and can scale based on request load and the amount of metadata stored.

Capacity and autoscaling

The Data Map has two capacity components: metadata storage and operation throughput. A capacity unit (CU) represents both components. By default, a Microsoft Purview account starts with one capacity unit and scales based on usage up to a maximum of 10 capacity units. Each capacity unit includes throughput of 25 operations per second and 10 GB of metadata storage.

The Data Map includes the following capabilities:

- **Elastic Data Map with autoscale**: Start with one capacity unit and automatically scale based on load up to 10 capacity units.
- **Enhanced scanning and ingestion**: Track and control how scanning and ingestion populate data assets, classifications, and lineage. This capability affects pricing.

Operation throughput

Operations measure the throughput of Microsoft Purview Data Map. They include create, read, write, update, and delete operations on metadata stored in the Data Map. Examples include:

- Creating an asset in the Data Map.
- Adding a relationship to an asset, such as an owner, steward, parent, or lineage.
- Editing an asset to add business metadata, such as a description or glossary term.
- Running a keyword search that returns results to the search results page.

Metadata storage

Storage includes technical, business, operational, and semantic metadata.

Technical metadata includes schemas, data types, and columns discovered through Microsoft Purview scanning. Business metadata includes automated metadata, such as metadata promoted from Microsoft Power BI datasets or descriptions from SQL tables, and manually added descriptions and glossary terms. Semantic metadata includes collection mappings to data sources or classifications. Operational metadata includes data factory copy and data flow activity run statuses and run times.

Example scenario

Claudia is a Microsoft Azure admin at Contoso who wants to create a Microsoft Purview account from the Azure portal. She doesn't know the Data Map size that the future state of the platform requires. However, she knows that Data Map billing uses capacity units and is affected by storage and operations throughput. She wants to start with the smallest Data Map to keep costs low and grow it elastically based on consumption.

Claudia can create a Microsoft Purview account with the default Data Map size of one capacity unit, which can automatically scale up and down. Autoscaling also allows capacity to adjust for intermittent or planned data bursts during specific periods. Claudia then configures the network settings and completes the account creation.

On the Microsoft Purview account's **Metrics** page in the Azure portal, Claudia can view Data Map storage and operations throughput consumption. She can also create an alert for when storage or operations throughput reaches a specified limit.

Data Map billing

Billing for classic Microsoft Purview accounts

You pay for one capacity unit, which includes 25 operations per second and 10 GB of metadata storage. Extra billing is based on the consumption of each extra capacity unit, calculated hourly. Data Map operations scale in increments of 25 operations per second, and metadata storage scales in increments of 10 GB. Data Map can automatically scale up and down within the elasticity window. You can increase the elasticity window only for classic Microsoft Purview accounts. To request an increase, review the current limits and create a support request.

Data Map capacity units come with a cap on operations throughput and storage. If storage exceeds the current capacity unit, you pay for the next capacity unit even if you don't use the operations throughput. The following table shows the Data Map capacity unit ranges. Contact support if the Data Map capacity unit goes beyond 10 capacity units.

Expand table

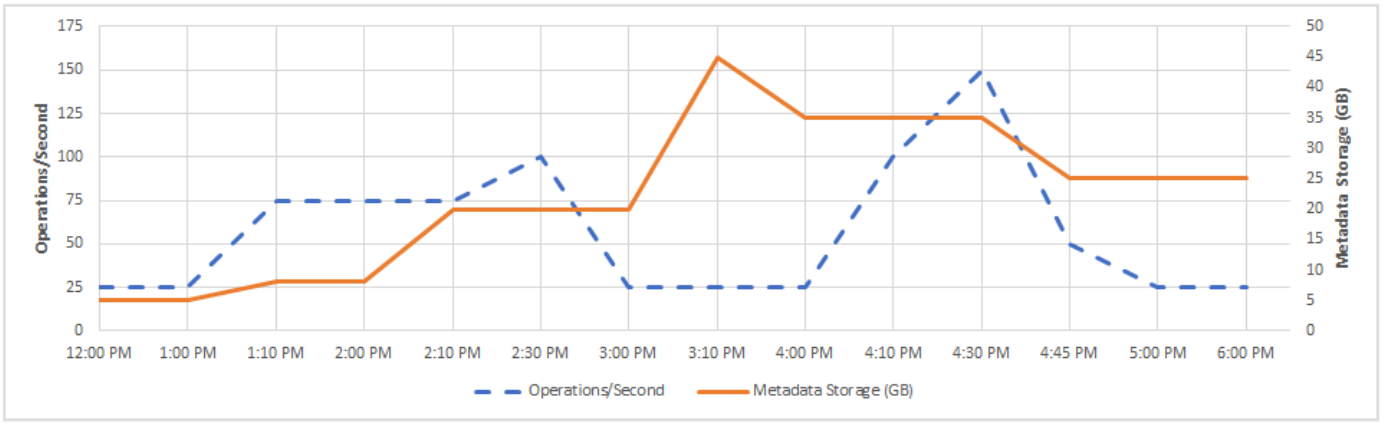
Data Map capacity unit	Operations-per-second throughput	Storage capacity in GB
1	25	10
2	50	20
3	75	30
4	100	40
5	125	50
6	150	60
7	175	70
8	200	80
9	225	90
10	250	100

Billing examples

- The operation throughput for the hour is less than or equal to 25 operations per second, and the storage size is 1 GB. You pay for one capacity unit.
- The operation throughput for the hour is less than or equal to 25 operations per second, and the storage size is 15 GB. You pay for two capacity units.
- The operation throughput for the hour is 50 operations per second, and the storage size is 15 GB. You pay for two capacity units.
- The operation throughput for the hour is 50 operations per second, and the storage size is 25 GB. You pay for three capacity units.
- The operation throughput for the hour is 250 operations per second, and the storage size is 15 GB. You pay for 10 capacity units.

Detailed billing example

The Data Map billing example shows growing metadata storage and variable operations per second over a six-hour window from 12 PM to 6 PM. The dashed line shows operations per second, and the solid line shows metadata storage:

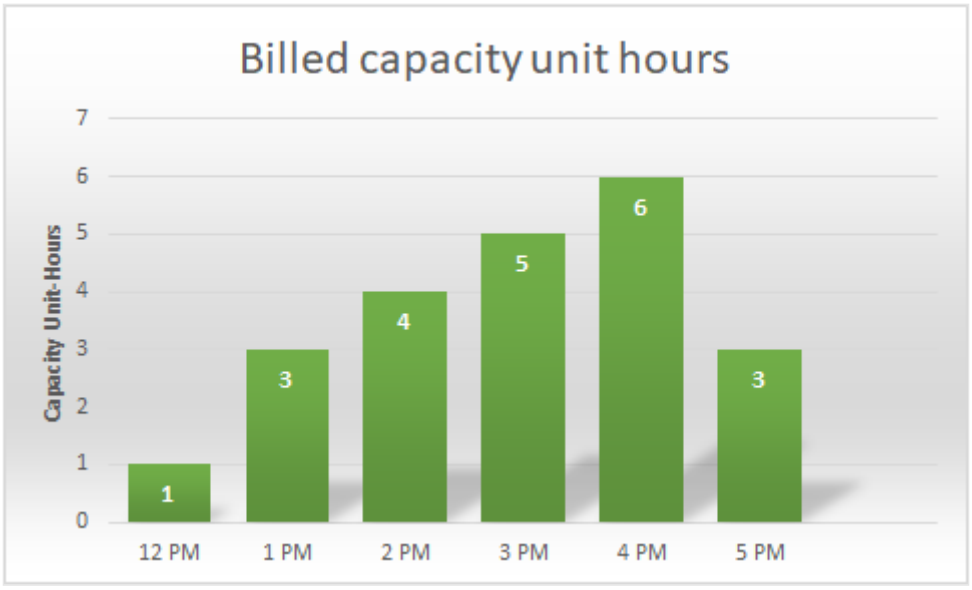


Each Data Map capacity unit supports 25 operations per second and 10 GB of metadata storage. The Data Map is billed hourly. The billing process considers the maximum Data Map capacity units needed within the hour, with a minimum of one capacity unit. At times, you might need more operations per second within the hour, and more operations increase the number of capacity units needed within that hour. At other times, your operations per second usage could be low, but you might still need a large volume of metadata storage. The metadata storage determines how many capacity units you need within the hour.

The table shows the maximum number of operations per second and metadata storage used per hour for this billing example:

Period	1	2	3	4	5	6
Start Time	12 PM	1 PM	2 PM	3 PM	4 PM	5 PM
End Time	1 PM	2 PM	3 PM	4 PM	5 PM	6 PM
Max ops/sec	25	75	100	25	150	25
Max Storage (GB)	5	8	20	45	35	25
Billed CU-hours	1	3	4	5	6	3

Based on the operations-per-second and metadata storage consumption during this period, the Data Map is billed for 22 capacity-unit hours over six hours (1 + 3 + 4 + 5 + 6 + 3):



> **Important**
>
> Data Map can automatically scale up and down within the elasticity window. To increase the elasticity window, review the current limits and create a support request.

Operations throughput limits

The default throughput limit is 10 capacity units. If you're working with a large Microsoft Purview environment and need higher throughput, you can request a larger elasticity window by creating a quota request. Select **Data map capacity unit** as the quota type. Include relevant information about your environment and the extra capacity that you need. You can increase the elasticity window only for classic Microsoft Purview accounts.

> **Important**
>
> There's no default limit for metadata storage. As you add more metadata to Data Map, it elastically increases.

When you increase the operations throughput limit, you also increase the minimum number of capacity units. For example, if you increase the throughput limit to 20, you pay for a minimum of two capacity units. The following table shows the possible throughput options. The number that you enter in the quota request is the minimum number of capacity units on the account.

Expand table

Minimum capacity units	Operations throughput limit
1	10 (default)
2	20
3	30
4	40
5	50
6	60
7	70
8	80
9	90
10	100

Optimize capacity units and operations throughput

Use these best practices to optimize capacity-unit and operations-throughput usage:

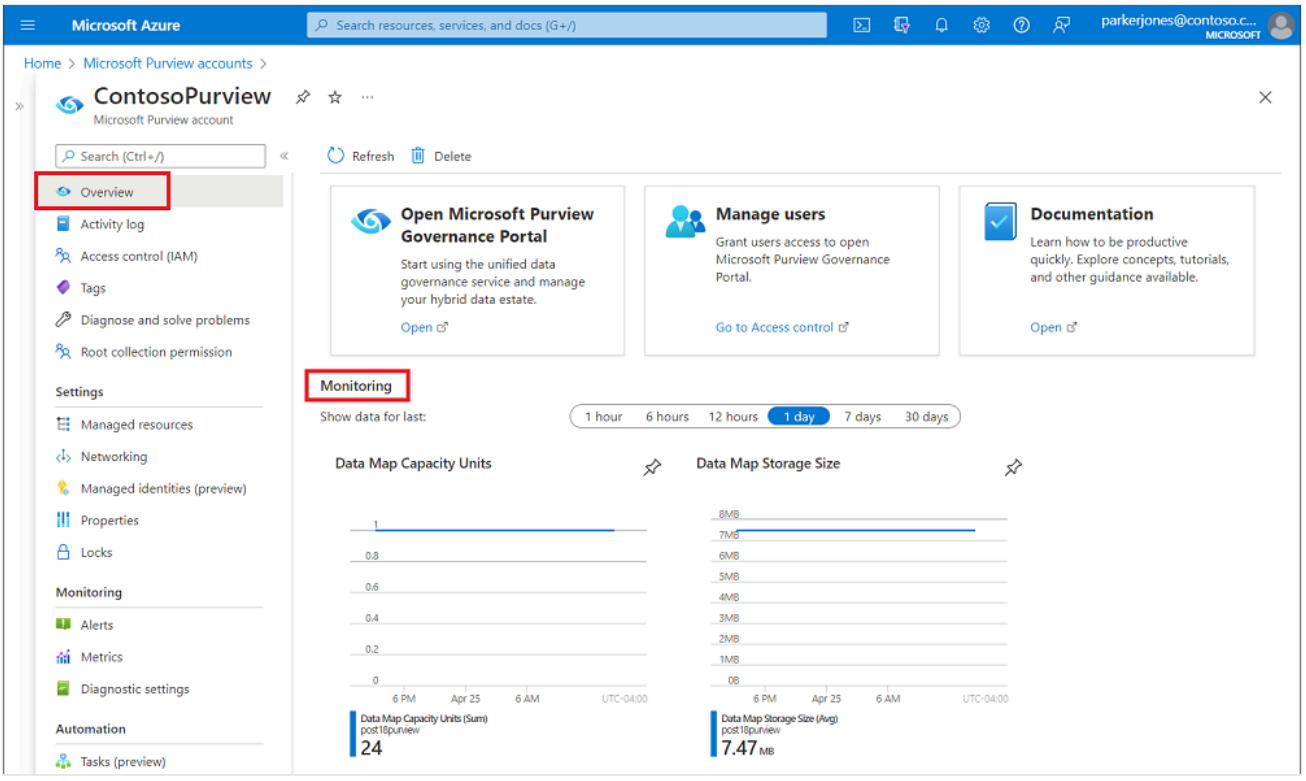
- Use incremental scans after the initial full scan. Incremental scans process only objects changed since the previous scan and typically provide the greatest reduction.
- Limit the scan scope and, where available, use **Advanced mode** to include or exclude specific paths instead of scanning the entire source.
- Stagger scan schedules. Avoid running several scans simultaneously; run them sequentially and outside peak usage periods.
- Reduce scan frequency. For example, change hourly scans to daily scans if the governance freshness requirement allows it.
- Disable lineage extraction if it is not required.
- Use separate, nonoverlapping scans.
- Clean up historical lineage regularly from Microsoft Purview storage.
- Manually remove nonessential or stale data from Microsoft Purview storage.

Monitoring and metrics

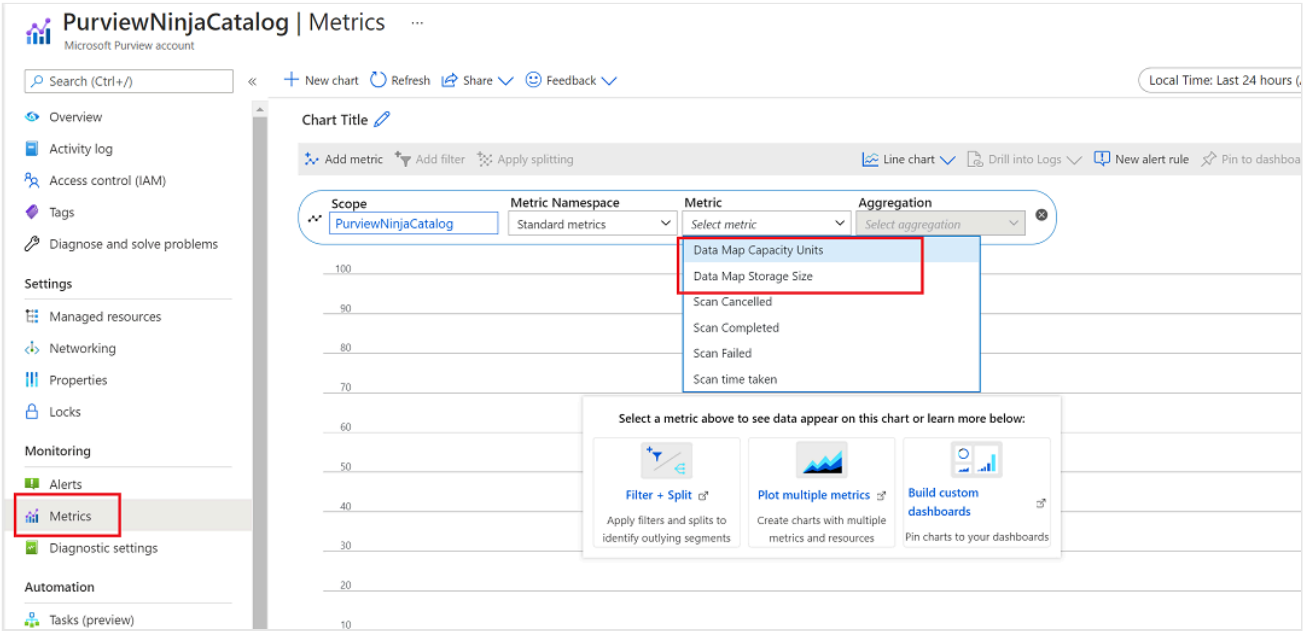
Monitor the **Data Map Capacity Units** and **Data Map Storage Size** metrics to understand the size and billing of your data estate.

1. In the Azure portal, go to **Microsoft Purview accounts**, and then select your Microsoft Purview account.

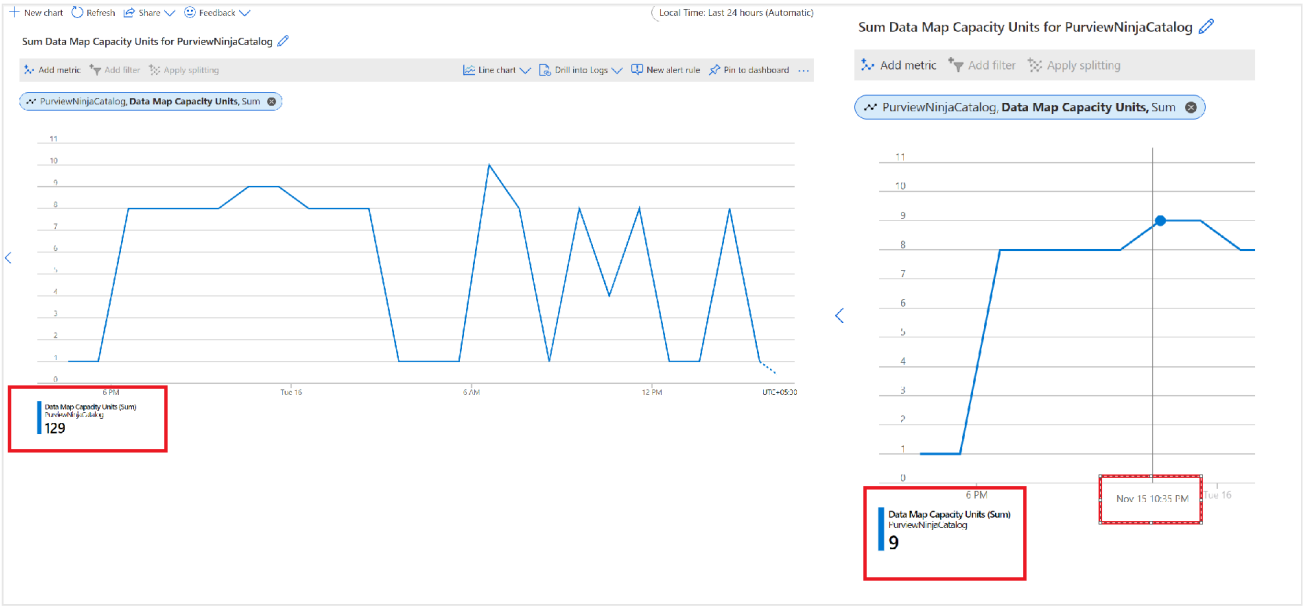
2. On the **Overview** page, scroll to the **Monitoring** section to view the **Data Map Capacity Units** and **Data Map Storage Size** metrics over different time periods.



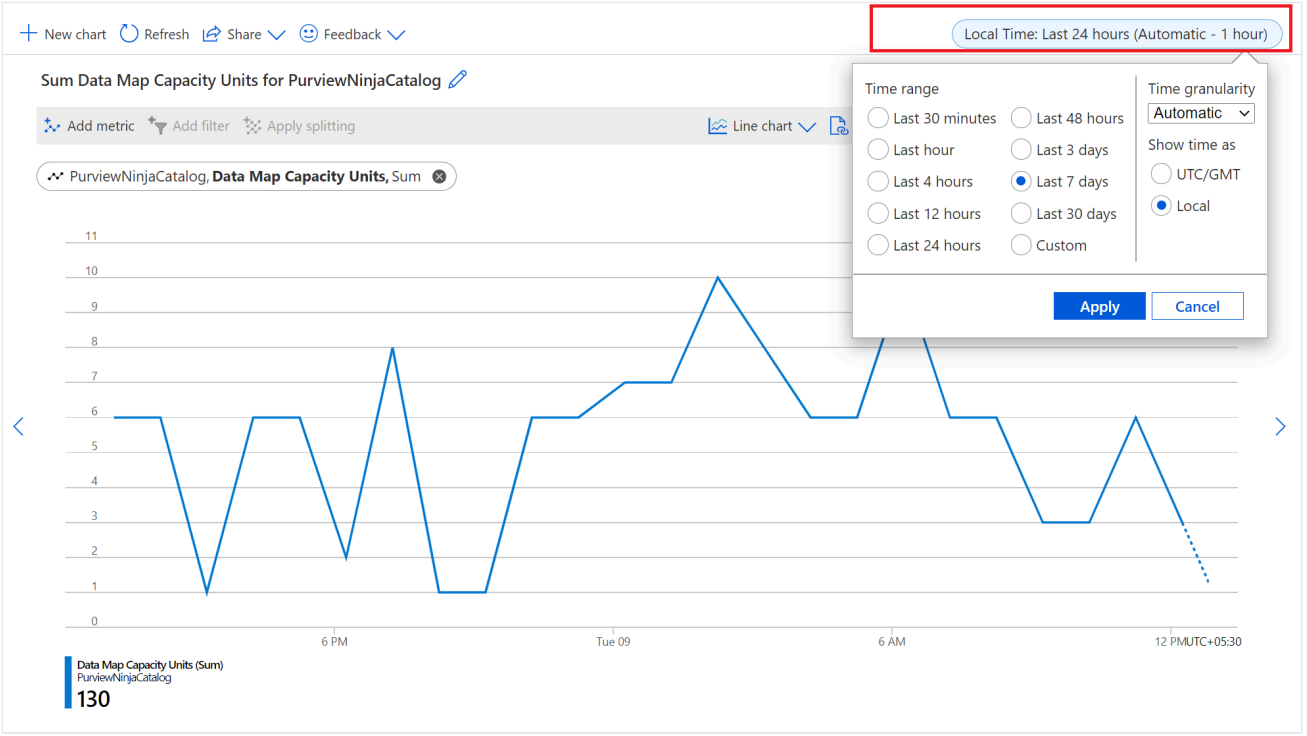
3. For more settings, go to **Monitoring** > **Metrics** to view **Data Map Capacity Units** and **Data Map Storage Size**.

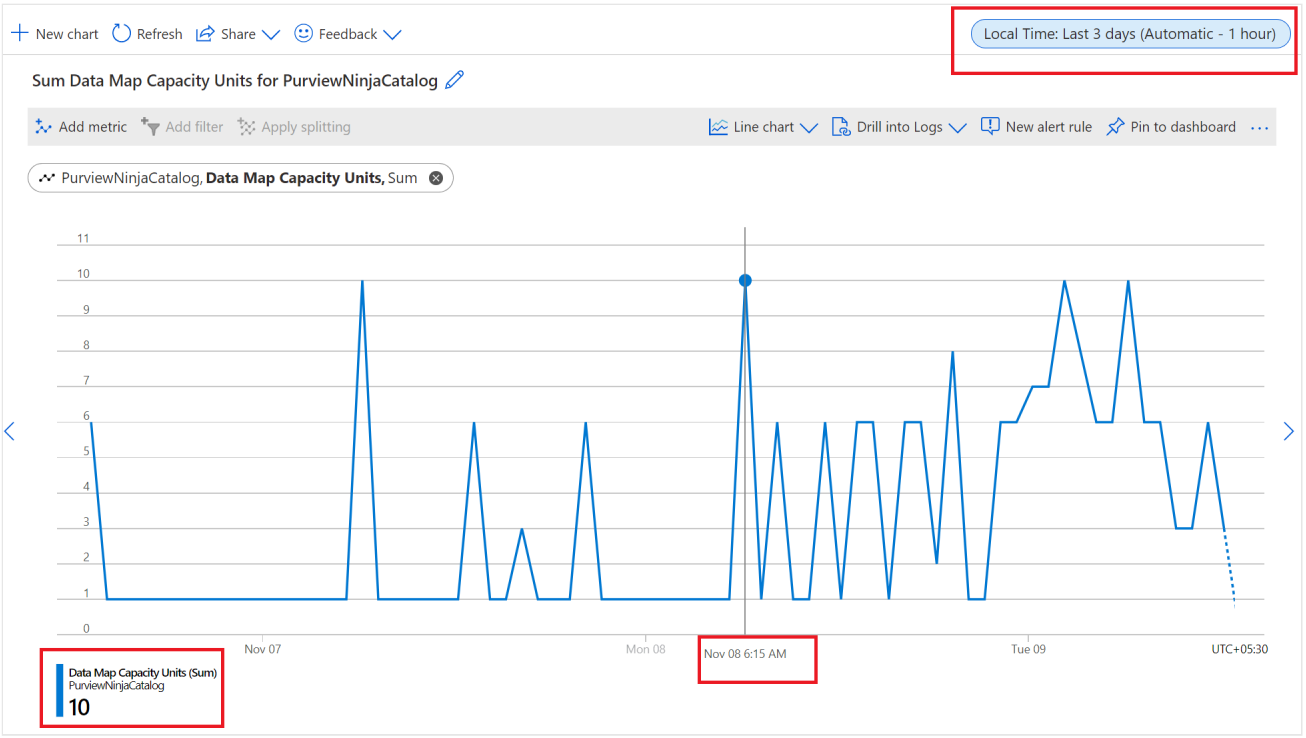


4. Select **Data Map Capacity Units** to view capacity unit usage over the last 24 hours. Hover over the line graph to view the Data Map capacity units consumed at a specific time.

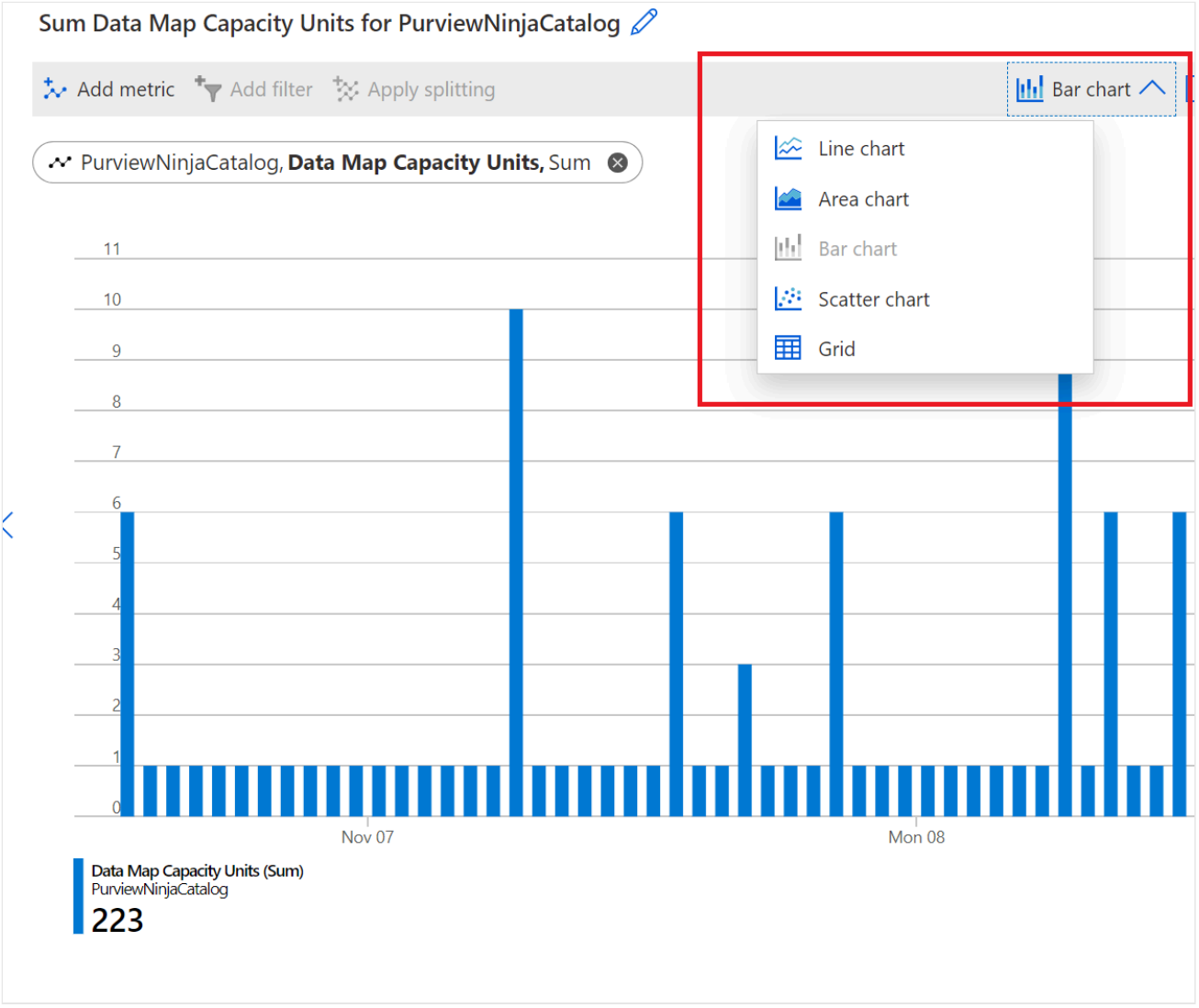


5. At the top of the page, select **Local Time: Last 24 hours (Automatic - 1 hour)** to change the time range displayed in the graph.

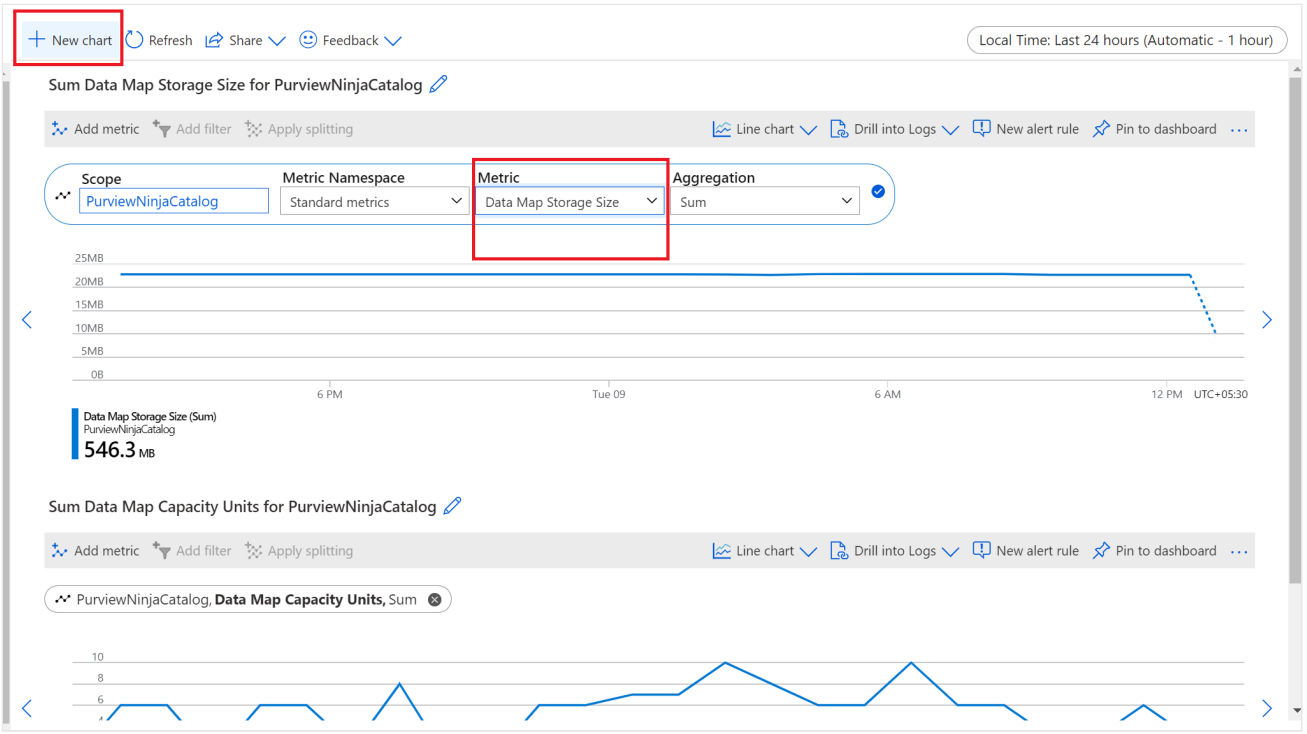




6. Select a graph type to customize the visualization.



7. Select **New chart** to add a **Data Map Storage Size** chart.



Next steps

- Create a Microsoft Purview account
- Learn about scans and ingestion in Microsoft Purview Data Map
- Monitor Microsoft Purview with Azure Monitor
- Review Data Map pricing
- Manage and increase quotas for classic data governance resources

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 10/01/2026

Domains in Microsoft Purview Data Map

> **Tip**
>
> This article covers domains in the Microsoft Purview Data Map. For governance domains in Microsoft Purview Unified Catalog, see the governance domains article.

A domain is a construct that you use to distribute organizational responsibility in your instance of Purview Data Map. Domains provide structure in Data Map so that domain administrators can maintain the isolation of collections, data sources, scans, and assigned roles for access control. Previously, only collections filled this role, but we're introducing domains as top-level nodes above the rest of the hierarchy to address a few other scenarios:

- Decentralize and distribute responsibility by assigning roles at the domain level.
- Create the possibility for hard logical separations within your Data Map.
- Replace multiple accounts within a tenant to multiple domains within a single Microsoft Purview resource.

Domains also serve as the container for collections and assets.

> **Note**
>
> Domains are only available to Microsoft Purview accounts using a tenant-level account (https://purview.microsoft.com).

Default domain

Every Microsoft Purview Data Map starts with a default domain.

When an account is upgraded to the new experience, the primary account's root collection becomes the default domain.

If you haven't upgraded to the new experience yet, take these points into consideration when choosing which account to elevate as your default domain:

- **Data Assets**: Choose the account with the most valuable or frequently used data assets, as this will become the default domain after the upgrade.
- **Account Usage**: Evaluate how each account is currently being used and its role within your organization. Accounts that are considered "production" would be most viable to select for this purpose.

- **Permissions and Access Controls**: Consider the existing permissions and access controls for each account, as they'll be carried over to the upgraded environment (as an isolated Domain). The permissions are applied in the new environment exactly as they are in the original environment. No more permissions are granted.

For more information about the new experience, see our guide on governance in the new Microsoft Purview experience.

Custom domains

You can also create up to four more custom domains in your Data Map to better organize and govern your assets in Data Map. Get more information on how to create and manage domains.

Domain level roles

Alongside all the permissions you can assign at a collection level, a new role can be assigned at the domain level to manage domains, collections, and glossaries:

- Domain admin - Can assign permissions within a domain and manage its resources.

For more information, see how to assign permissions in collections and on a domain level.

Data Map structure

Capabilities and responsibilities are distributed between these three new levels in Data Map's structure: tenant/organization level, domain level, and collection level.

- Tenant level - Classifications, search, browse, metamodel, managed attributes, integration runtimes, lineage, workflows, insights, private endpoints.
- Domain level - Credentials, term templates, custom scan rule sets, advanced resource sets, pattern rules, policies, assets.
- Collection level - Data sources, scans, assets.

> **Important**
>
> Data sources and assets can't be duplicated in a tenant.

Domains don't change permissions in Microsoft Purview. Search and browse across a tenant are still managed by the level a user's assigned permissions. In Unified Catalog, a user only has access to metadata if they have data-level access or collection-level access.

For more information, see permissions for Microsoft Purview Data Map and Unified Catalog.

Next steps

Merge secondary accounts into Microsoft Purview as domains.

Last updated on 11/24/2025

Manage domains and collections in Microsoft Purview Data Map

06/25/2025

> **Important**
>
> These features are only available if you're using the Microsoft Purview portal.
>
> If you're using the classic governance experience (https://web.purview.azure.com/), see the classic create and manage collections article instead.

Domains and collections in Microsoft Purview Data Map can be used to organize assets and sources by your business's flow. They're also the tool used to manage access across the Microsoft Purview portal. This guide takes you through the creation and management of domains and collections, as well as cover steps about how to register sources and add assets into your collections.

Find more information about domains and the structure of Data Map.

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- Your own Microsoft Entra tenant.
- A Microsoft Purview account using the Microsoft Purview portal experience.
- You need to be a **Microsoft Purview administrator** within the Microsoft Purview portal to manage domains. To check or set these roles, see our permissions documentation.

Permissions to manage domains and collections

To manage domains, and permissions in domains, a user needs to be a Purview administrator or a domain admin.

To manage collections, a user needs to be at least a **Domain Admin** or **Collection Admin** within the Microsoft Purview portal.

To manage collections, a user needs to be at least a **Domain Admin** or **Collection Admin** within the Microsoft Purview portal.

To check these permissions in the Microsoft Purview portal:

1. Open **Data Map**, then select **Domains** to open the domains management page.
2. Select your default domain, which is currently the only available domain.
3. Select the **Role assignments** tab.
4. Search under these roles for your user:
    - Domain admins - to be able to manage the domain or permissions in it.
    - Collection admins - to be able to create or manage collections in the domain.

> **Tip**
>
> If you don't have either of the needed permissions, contact the domain admin to grant you permission.

Default domain

Every Microsoft Purview Data Map starts with a default domain.

When an account is upgraded to the new experience, the primary account's root collection becomes the default domain.

If you haven't upgraded to the new experience yet, take these points into consideration when choosing which account to elevate as your default domain:

- **Data Assets**: Choose the account with the most valuable or frequently used data assets, as this will become the default domain after the upgrade.
- **Account Usage**: Evaluate how each account is currently being used and its role within your organization. Accounts that are considered "production" would be most viable to select for this purpose.
- **Permissions and Access Controls**: Consider the existing permissions and access controls for each account, as they'll be carried over to the upgraded environment (as an isolated Domain). The permissions are applied in the new environment exactly as they are in the original environment. No more permissions are granted.

For more information about the new experience, see our guide on governance in the new Microsoft Purview experience.

Custom domains

You can create up to four custom domains in your Data Map.

Create custom domains

1. In **Data Map**, select **Domains**.
2. Select **+ New domain**.
3. On the **New domain** pane, enter a name for the domain, a description, and assign one or more domain administrators.
4. Select **Create**.

Edit custom domains

You can modify the description or the administrators of your custom domains by editing them.

1. In **Data Map**, select **Domains**.
2. Select the domain you want to edit, and on its details page, select **Edit**.

Delete custom domains

> **Important**
>
> - Before you delete a custom domain, you must remove everything under it first. Including: subcollections, data sources, scans, assets, glossaries, terms, credentials, Azure Key Vaults, etc.
> - Default domains can't be deleted.

To remove a domain, you need to have either Microsoft Purview administrator or domain administrator permissions.

1. Navigate to https://purview.microsoft.com.
2. Open the **Data Map** solution.
3. Select **Domains** from the left pane to open the domains management page.
4. Select the domain you want to delete.
5. Select the **Delete** button and confirm the deletion.

Collection management

Create a collection

You need to be a collection admin or domain admin within a domain order to create a collection. If you aren't sure, see how to check permissions.

1. In **Data Map**, select **Domains**.
2. Use the dropdown next to your default domain to select your default domain, or select an existing collection where you want to make a child collection.
3. Select **+ New collection**. Only domain and collection admins can manage collections.
4. In the new collection window, enter the collection display name and description. If needed, you can also add users or groups as collection admins to the new collection.
5. Select **Create**.
6. The new collection's information reflects in the collection list under your domain, and you're taken to your collection's detail page.

Edit a collection

1. Select **Edit** either from the collection detail page, or from the collection's dropdown list.

2. Make changes, then select **Save**.

View Collections

1. On the **Domains** page in Data Map, select the arrow next to the collection's name to expand or collapse the collection hierarchy. Select the collection name to navigate.
2. Enter a name in the **Filter by name** box to filter collections.
3. Select **Refresh** in collection detail page to reload the single collection.

Delete a collection

You need to be a domain admin or a collection admin in order to delete a collection. If you aren't sure, check permissions. Collections can be deleted only if no child collections, assets, data sources, or scans are associated with it.

1. On the collection's details page, select **Delete**.
2. A confirmation window appears. Select **Confirm** to continue the deletion.
3. Verify deletion of the collection from your Data Map.

> ⓘ **Note**
>
> Collections can be deleted only if no child collections, assets, data sources, or scans are associated with it.

Move registered sources between collections

You can move registered sources from one collection to another you have access to. For steps, see Manage data sources.

Add roles and restrict access

Since permissions are managed through domains and collections in Data Map, it's important to understand the roles and what permissions they give to your users.

The roles are assigned and inherited in domains and collections the same way. A user granted permissions on a domain or collection will have access to sources and assets associated with that collection, and inherit permissions to its subcollections. Inheritance can be restricted, but is allowed by default.

The following guide discusses the roles, how to manage them, and permissions inheritance.

Roles

All assigned roles apply to sources, assets, and other objects within the domain or collection where the role is applied.

- **Domain admin** (domain level only) - Can assign permissions within a domain and manage its resources.
- **Collection administrator** - a role for users that will need to assign roles to other users in the Microsoft Purview governance portal or manage collections. Collection admins can add users to roles on collections where they're admins. They can also edit collections, their details, and add subcollections. A collection administrator on the root collection also automatically has permission to the Microsoft Purview governance portal. If your **root collection administrator** ever needs to be changed, you can follow the steps in the section below.
- **Data curators** - a role that provides access to the Microsoft Purview Unified Catalog to manage assets, configure custom classifications, create and manage glossary terms, and view data estate insights. Data curators can create, read, modify, move, and delete assets. They can also apply annotations to assets.
- **Data readers** - a role that provides read-only access to data assets, classifications, classification rules, collections and glossary terms.
- **Data source administrator** - a role that allows a user to manage data sources and scans. If a user is granted only to **Data source admin** role on a given data source, they can run new scans using an existing scan rule. To create new scan rules, the user must be also granted as either **Data reader** or **Data curator** roles.

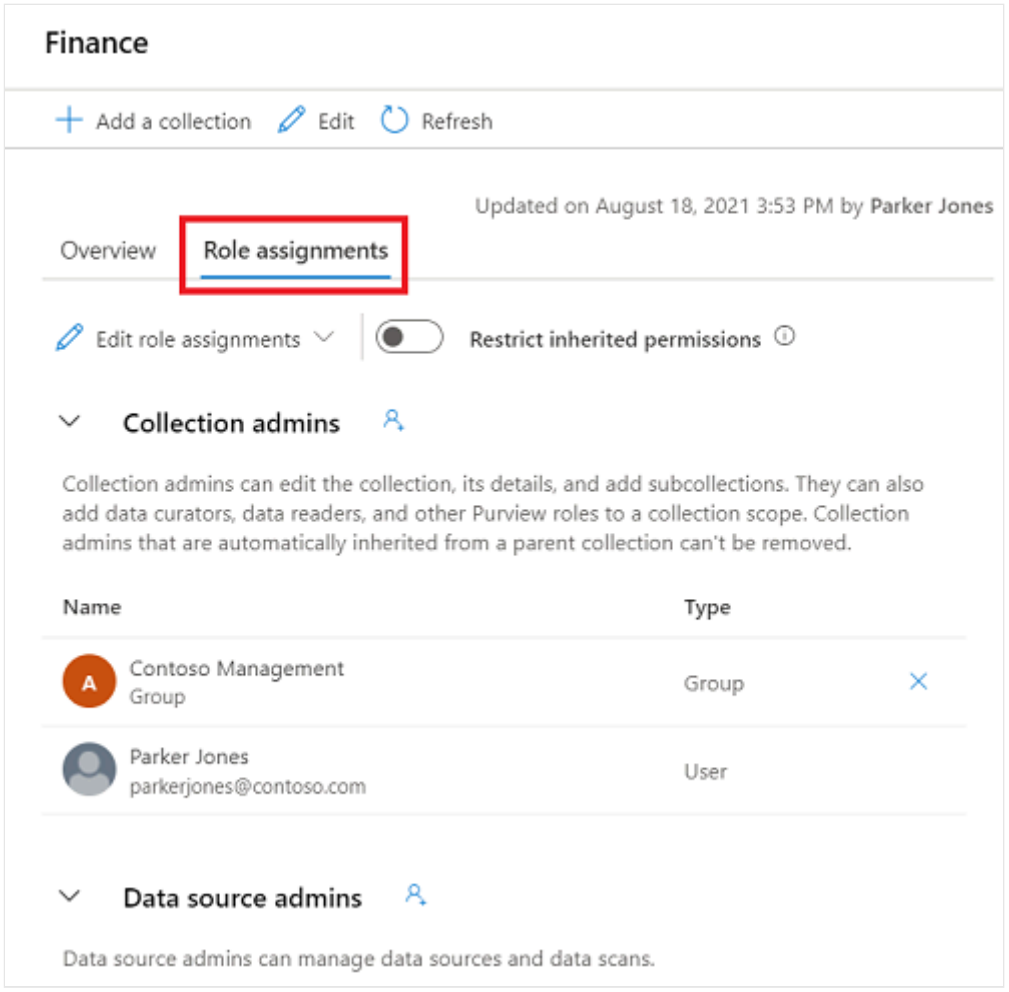
- **Insights reader** - a role that provides read-only access to insights reports for collections where the insights reader also has at least the **Data reader** role. For more information, see insights permissions.
- **Policy author** - a role that allows a user to view, update, and delete Microsoft Purview policies through the Data policy app within Microsoft Purview.
- **Workflow administrator** - a role that allows a user to access the workflow authoring page in the Microsoft Purview governance portal, and publish workflows on collections where they have access permissions. Workflow administrator only has access to authoring, and so will need at least Data reader permission on a collection to be able to access the Purview governance portal.

> **Note**
>
> At this time, Microsoft Purview policy author role is not sufficient to create policies. The Microsoft Purview data source admin role is also required.

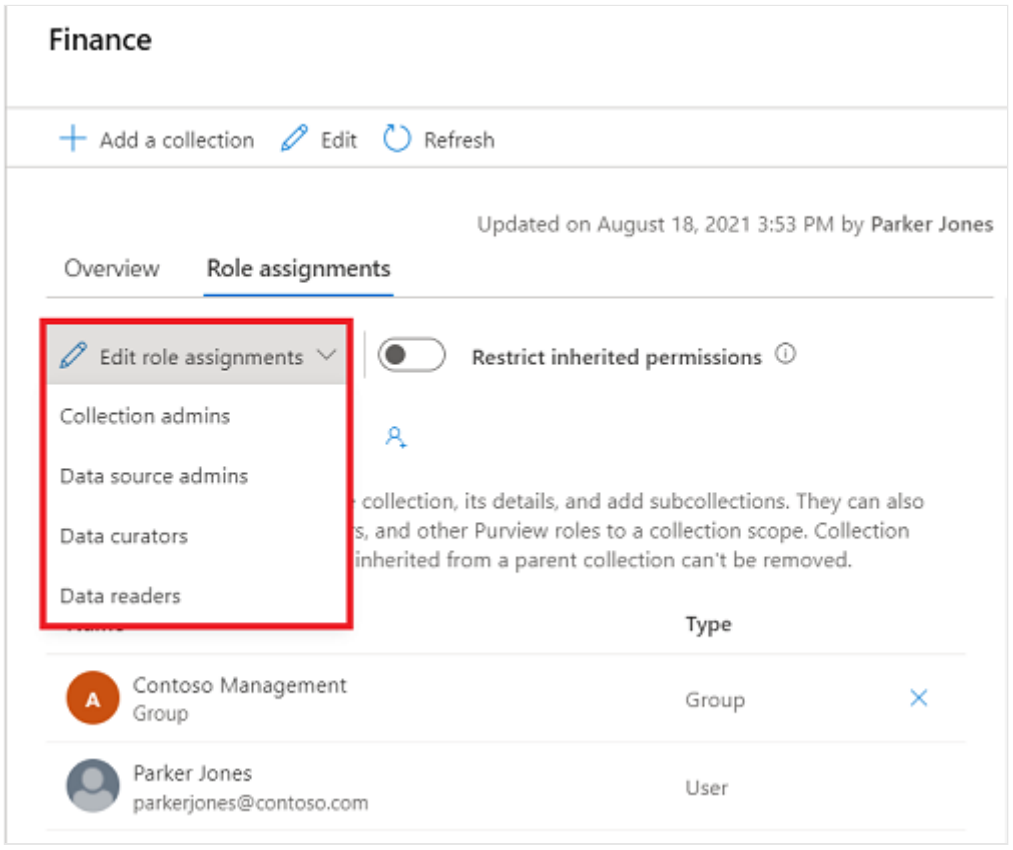
> **Important**
>
> The user that created the account is automatically assigned domain admin on the default domain and collection admin on the root collection.

Add role assignments

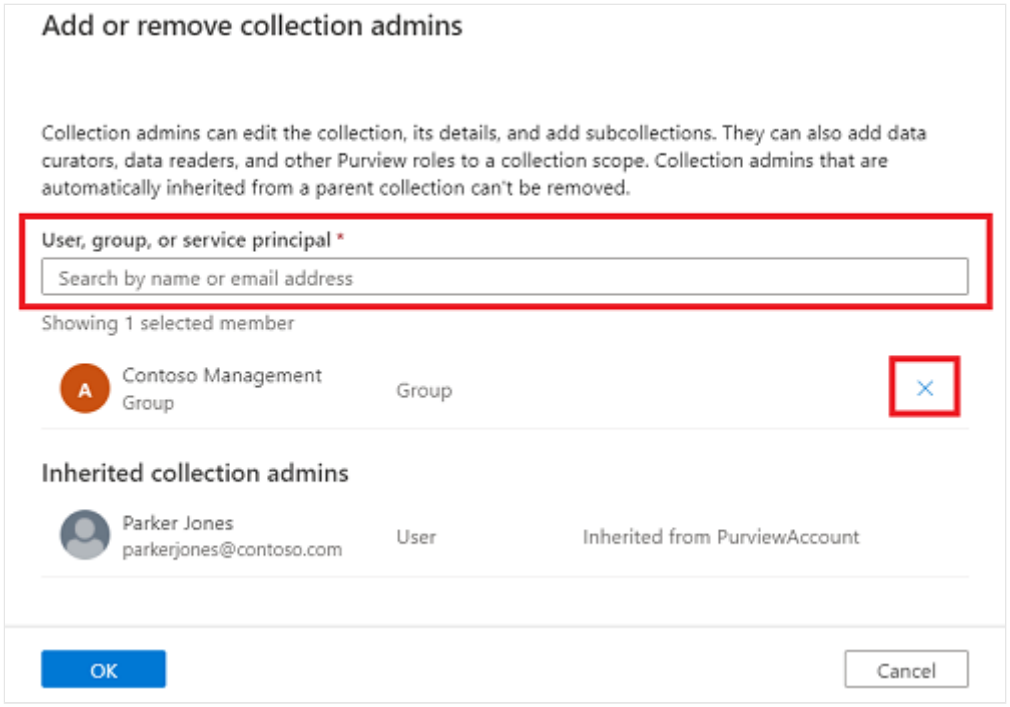
1. Open the Microsoft Purview Data Map.
2. Select the domain or collection where you want to add your role assignment.
3. Select the **Role assignments** tab to see all the roles in a collection or a domain. Only a collection admin or domain admin can manage role assignments.



4. Select **Edit role assignments** or the person icon to edit each role member.



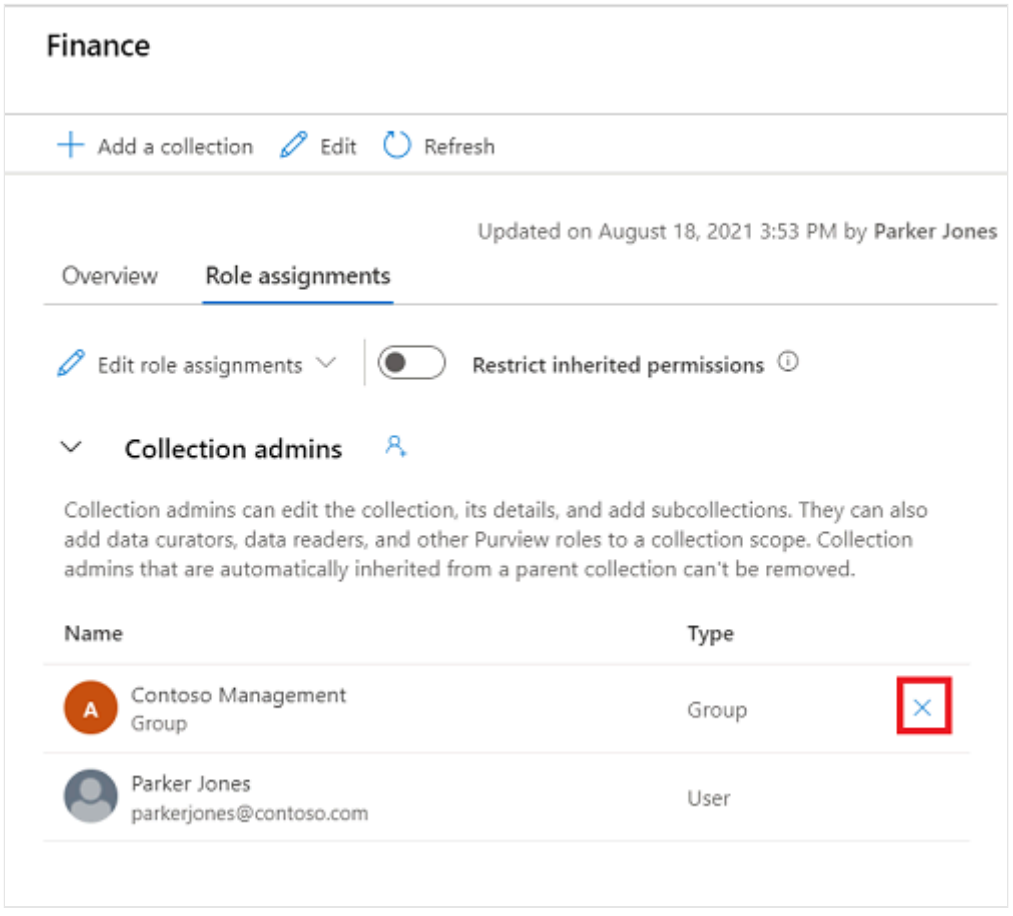
5. Type in the textbox to search for users you want to add to the role member. Select **X** to remove members you don't want to add.



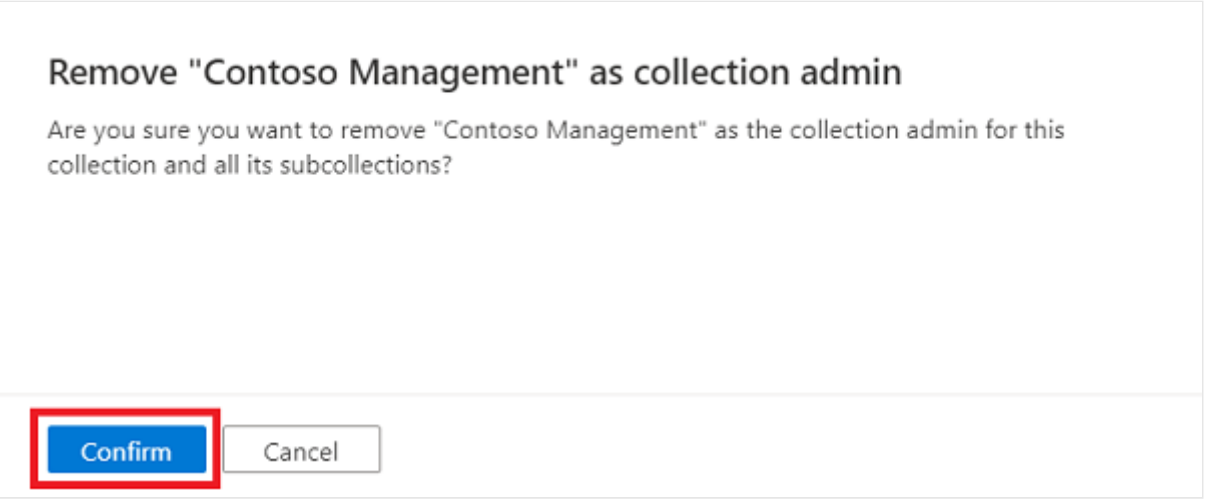
6. Select **OK** to save your changes, and you'll see the new users reflected in the role assignments list.

Remove role assignments

1. Select **X** button next to a user's name to remove a role assignment.



2. Select **Confirm** if you're sure to remove the user.



Restrict inheritance

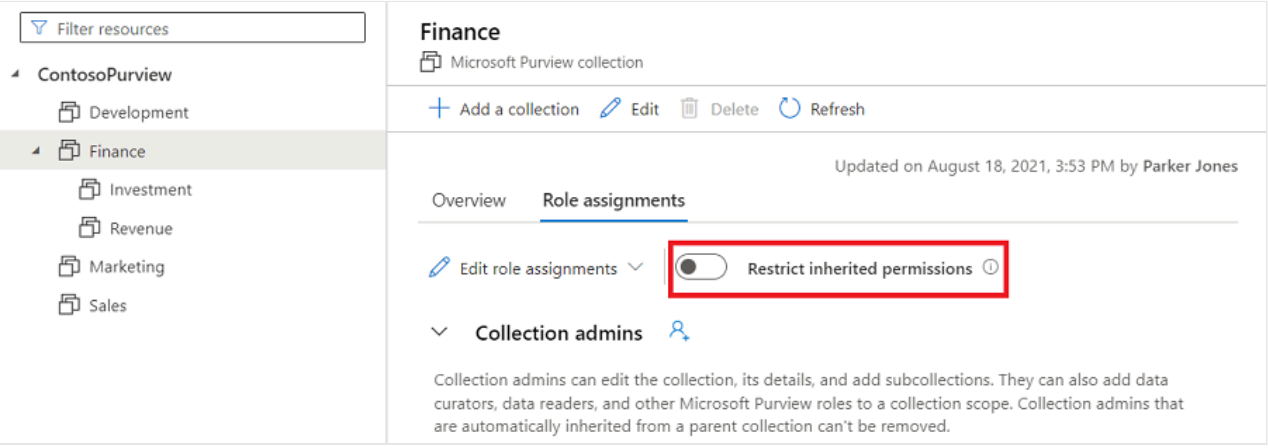
Collection permissions are inherited automatically from the parent collection. You can restrict inheritance from a parent collection at any time, using the restrict inherited permissions option.

> **Note**
>
> Currently permissions from the default domain cannot be restricted. Any permissions assigned at the default domain will be inherited by the domain's direct subcollections.

Once you restrict inheritance, you'll need to add users directly to the restricted collection to grant them access.

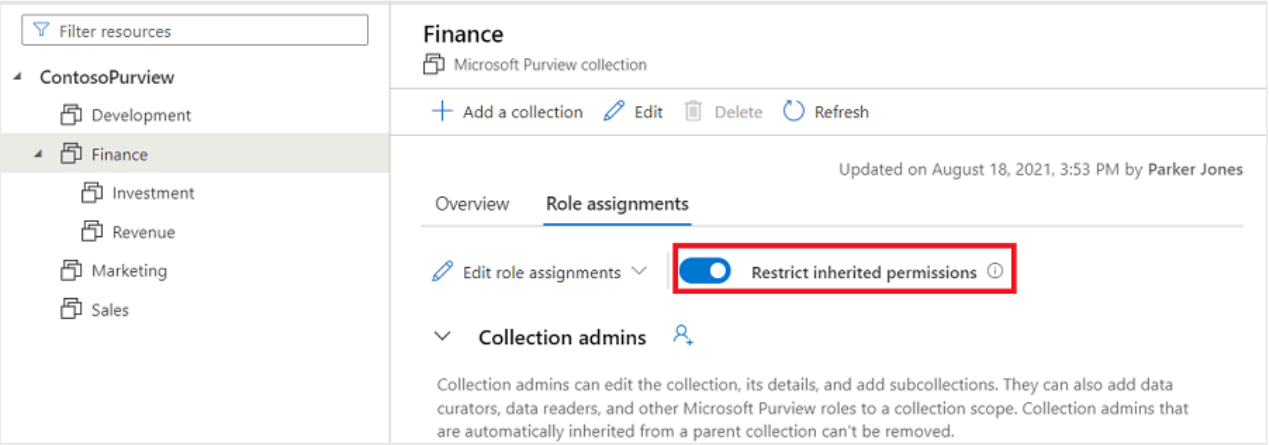
1. Navigate to the collection where you want to restrict inheritance and select the **Role assignments** tab.

2. Select **Restrict inherited permissions** and select **Restrict access** in the popup dialog to remove inherited permissions from this collection and any subcollections. Collection admin permissions won't be affected.



3. After restriction, inherited members are removed from the roles expect for collection admin.

4. Select the **Restrict inherited permissions** toggle button again to revert.



Register and scan source in a domain or collection

> **Important**
>
> Every source has its own prerequisites and information for registration and scanning. Get detailed registration and scanning steps for each data source.

1. Navigate to https://purview.microsoft.com.

2. Open the **Data Map** solution.

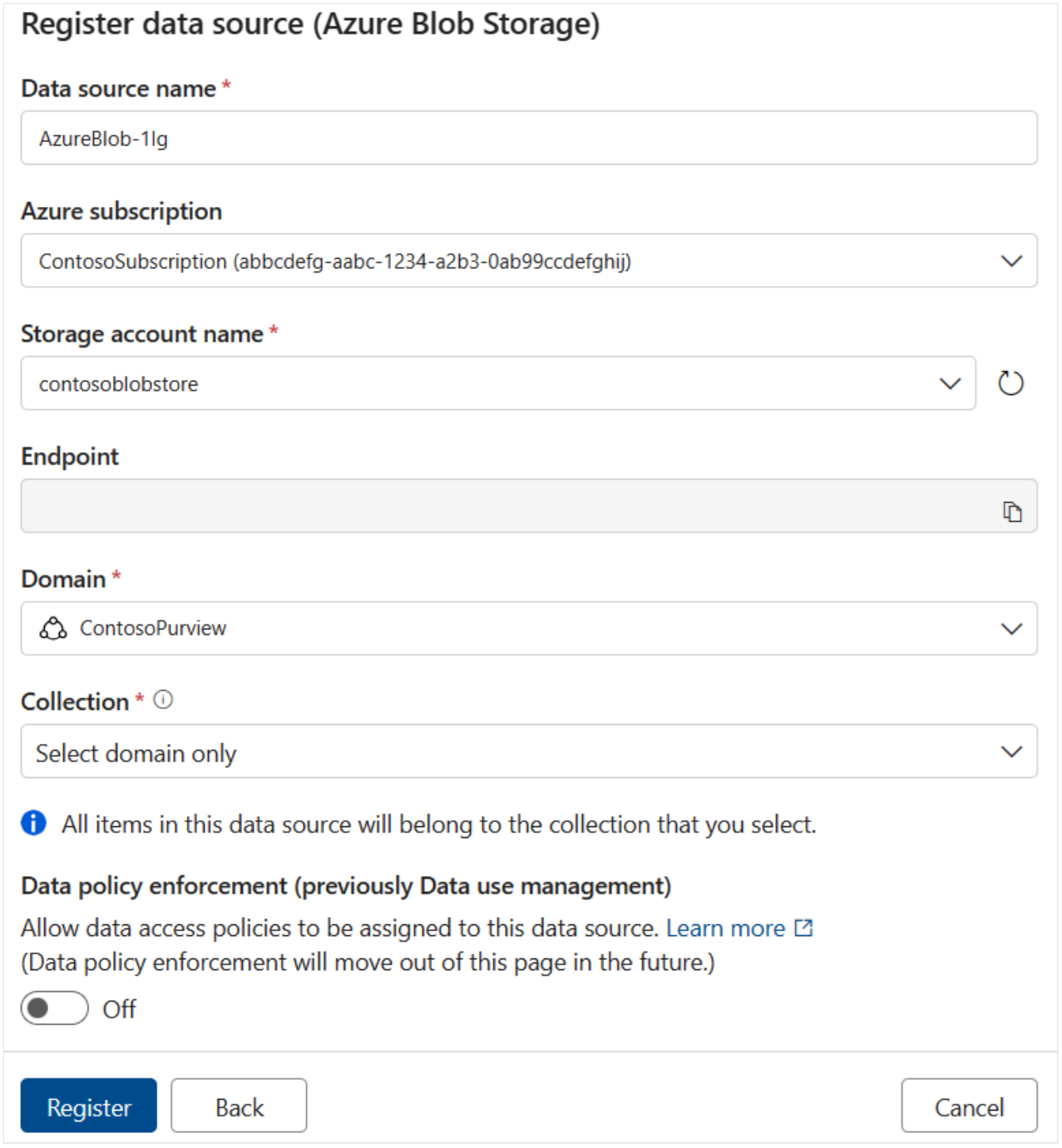
3. Select **Data sources** from the left pane to open the data sources map page.

4. Select **Register** or the register icon on a domain or collection node to register a data source. Only data source admins can register sources.

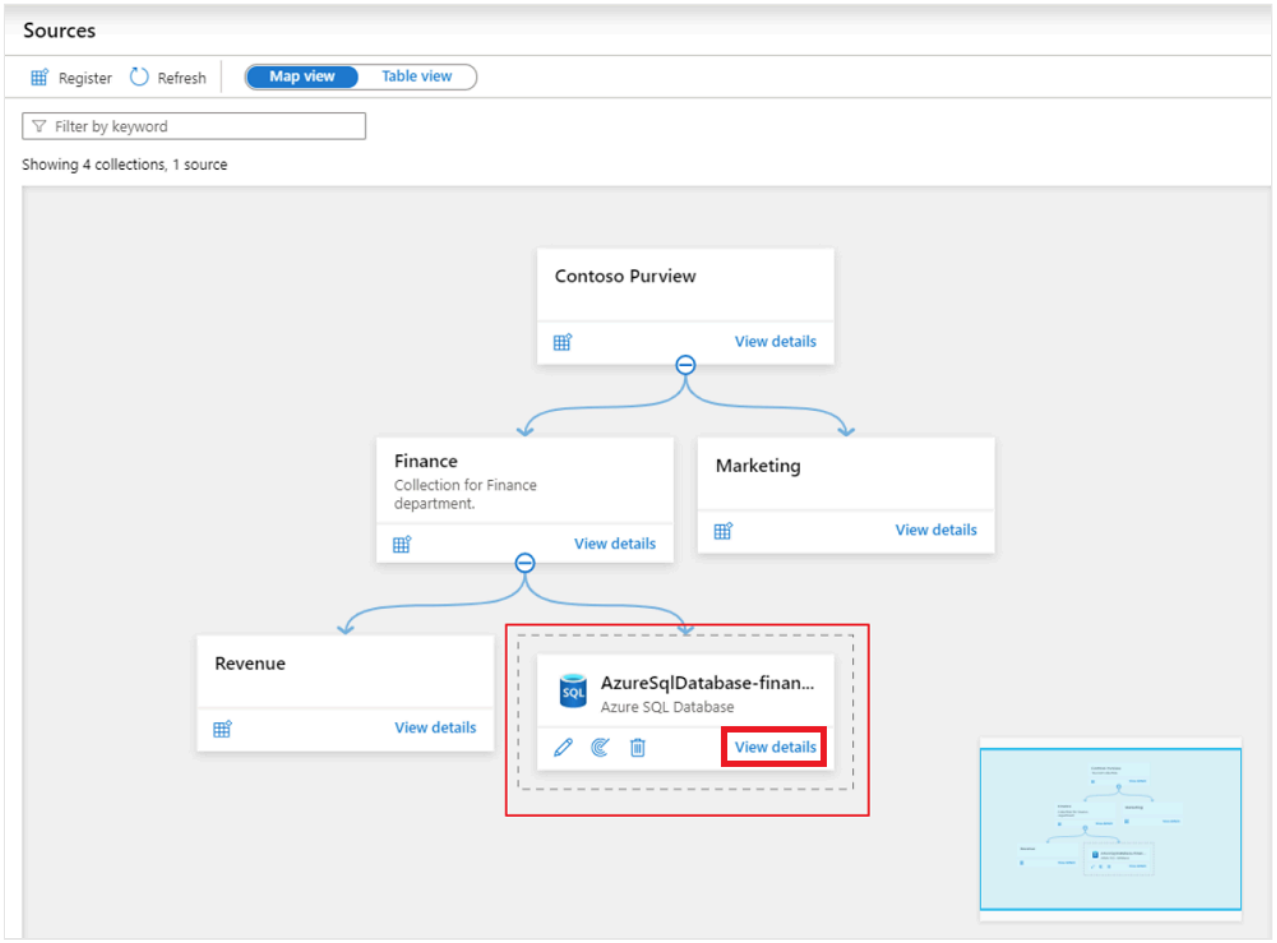
5. Fill in the data source name, and other source information.

6. Select a domain.

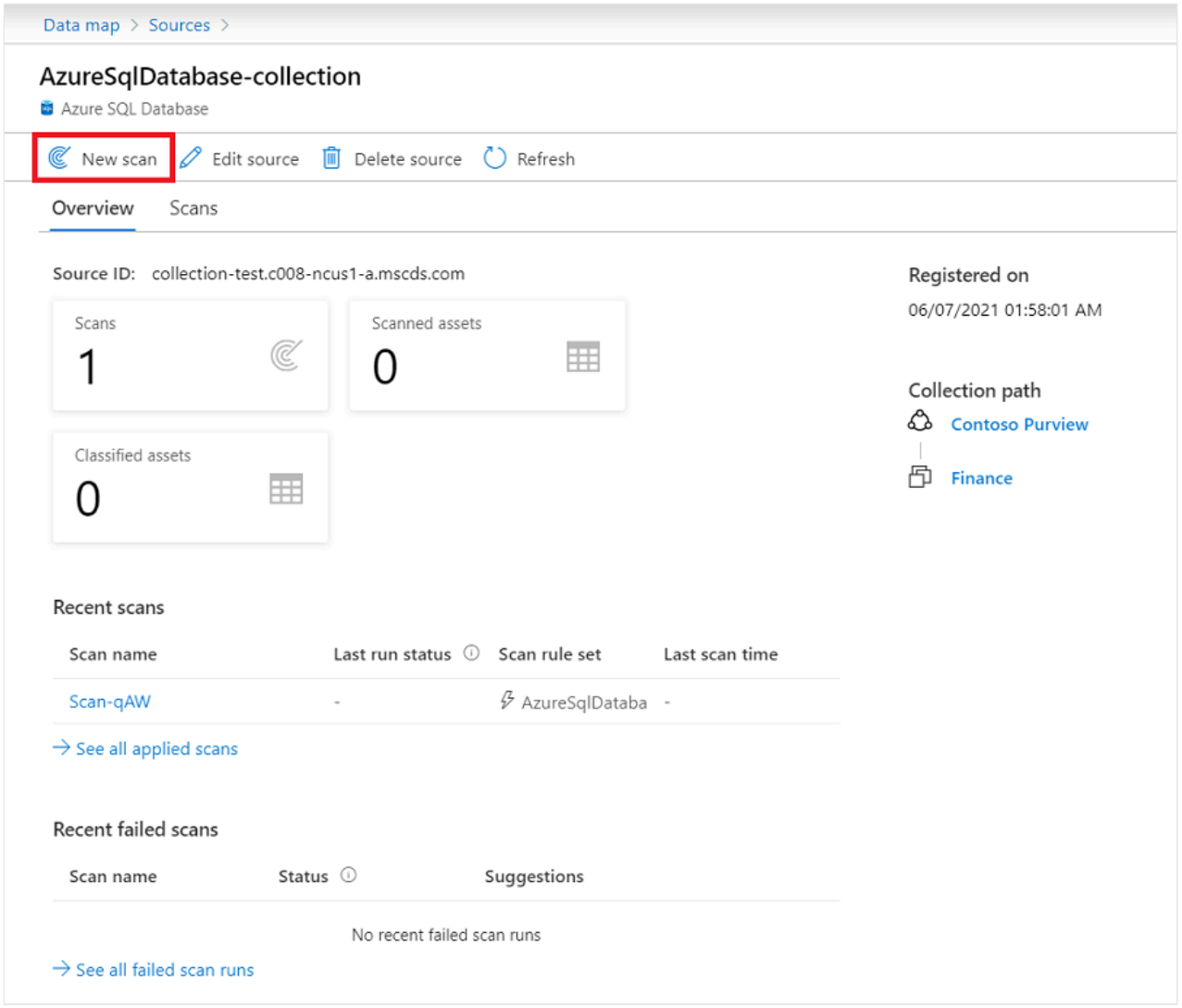
7. Select a collection. To register the resource in the domain, choose the **Select domain only** option. All assets under this source belong to the domain or collection you select.



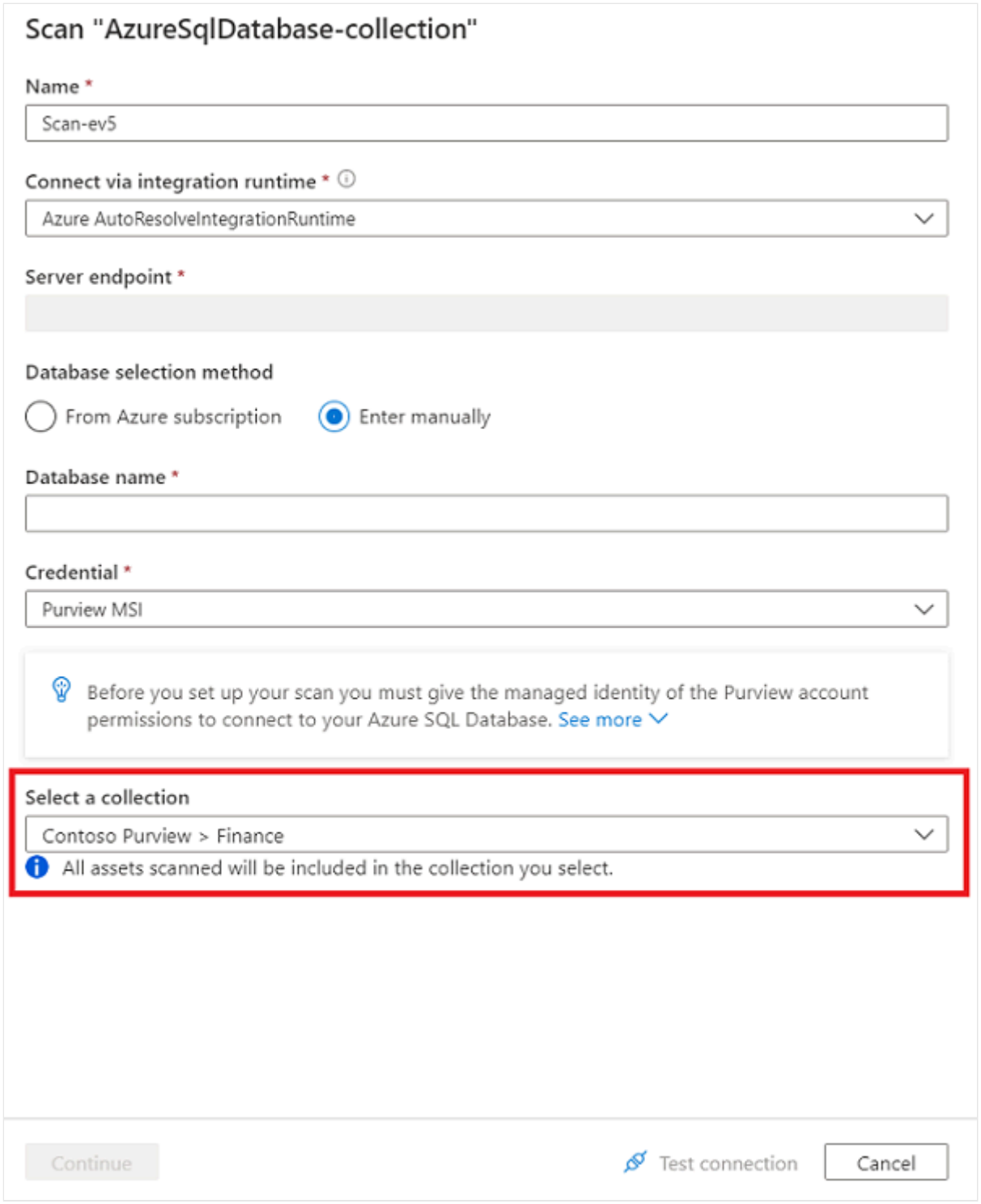
8. The created data source is put under the selected domain or collection. Select **View details** to see the data source.



9. Select **New scan** to create scan under the data source.



10. When creating a scan, you can select a collection and all assets scanned will be included in that collection. Collections listed in a scan are subcollections of the data source collection.

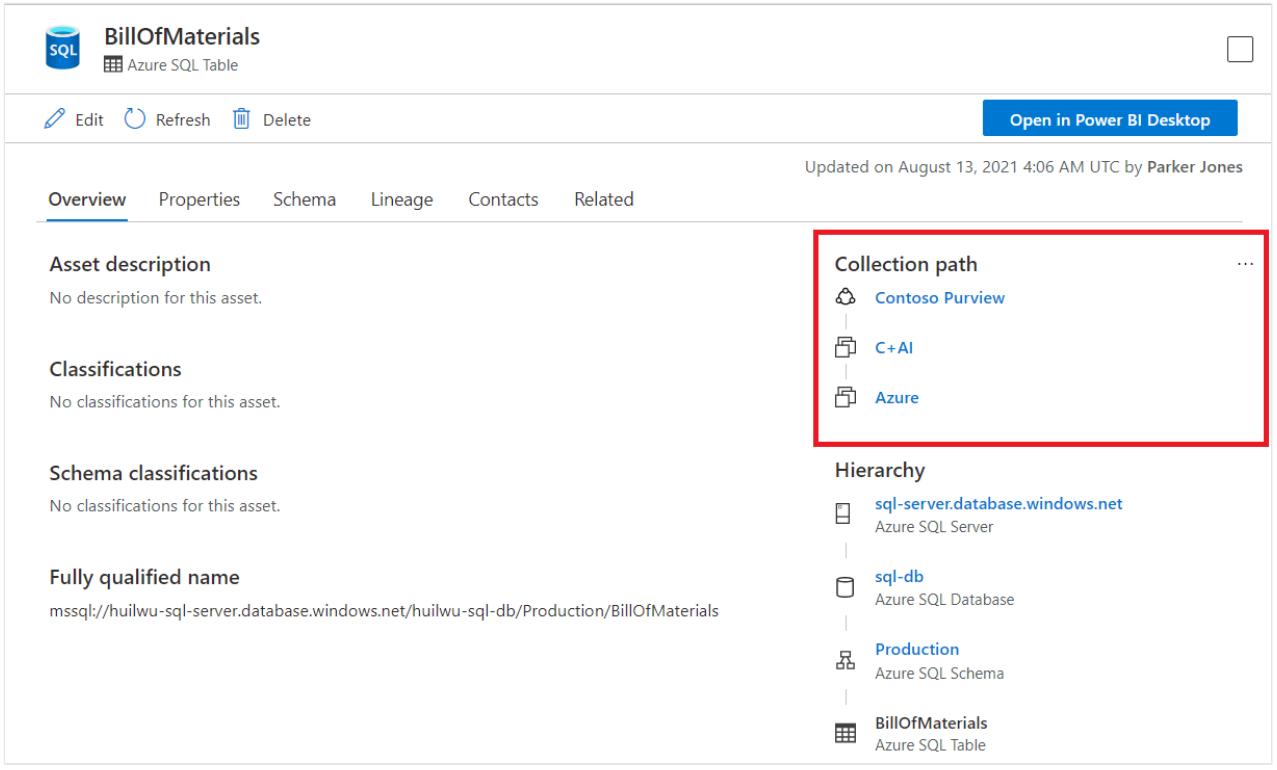


11. After your scan, you'll see the data sources linked to the collection or domain on their data sources card.

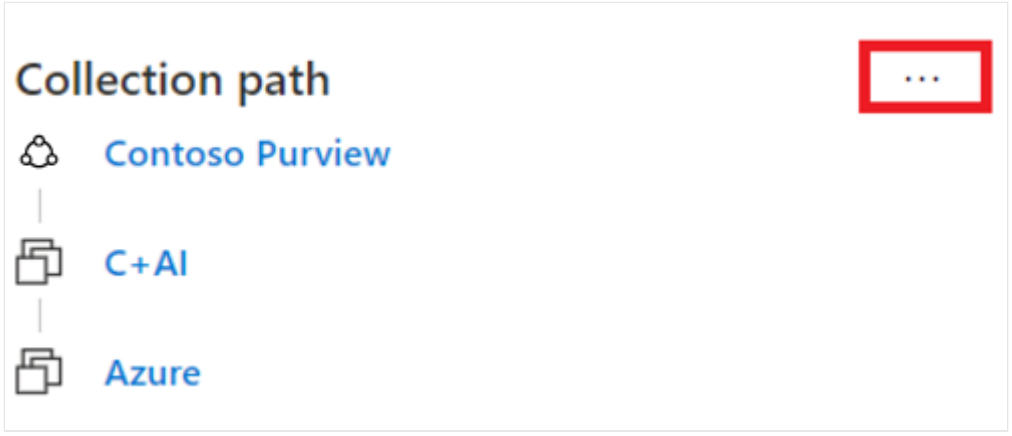
Add assets to collection and domains

Assets and sources are also associated with domains and collections. During a scan, if the scan was associated with a domain or collection, the assets are automatically added to that resource, but the assets can also be manually moved to any subcollections (if you have the write permissions on those subcollections.)

1. Check the domain and collection information in asset details. You can find information in the **Collection path** section on right-top corner of the asset details page.

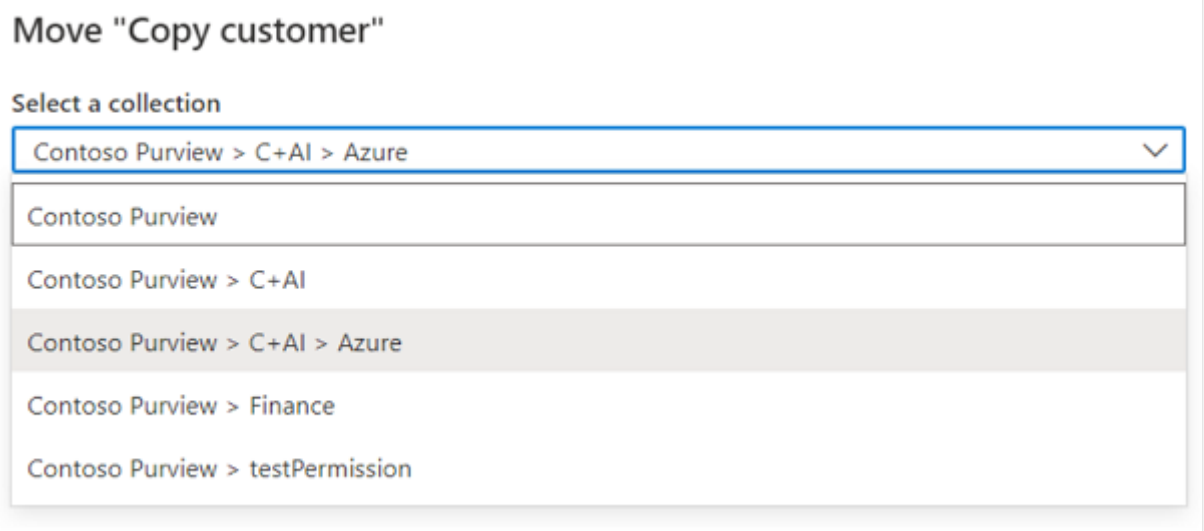


2. Select the ellipsis button on the right-top corner of the **Collection path** section.



3. Select the **Move to another collection** button.

4. In the right side panel, choose the target collection you want move to. You can only see the collections where you have write permissions. The asset can also only be added to the subcollections of the data source collection.



5. Select **Move** button on the bottom of the window to move the asset.

Rename collections

1. In Data Map, navigate to the collection that you want to rename.
2. Select **Edit** either from the collection detail page, or from the collection's dropdown menu.
3. Enter a new friendly name and select **Save**.

Move collections

1. In the Microsoft Purview portal, navigate to the collection that you want to move to a different parent collection.
2. Select the ellipsis button on the right-top corner of collection path section and select **Move to another collection**.
3. From **Select parent collection** dropdown list, select a new parent for the collection and select **OK**.

Best practices

- Domains and collections architecture best practices
- Data Map domains and governance domains best practices

Next steps

Now that you have organized your domains and collections, you can follow these guides below to add resources and scan:

- Manage data sources
- Supported data sources
- Scan and ingestion

Merge a classic Microsoft Purview account to a domain

Article • 05/30/2025

This guide describes how to merge an existing classic Microsoft Purview account into your Microsoft Purview tenant account as a domain. Merge is a one-time process to *copy* the information in your secondary classic Microsoft Purview account into your tenant level account. All the information is incorporated into a new domain. Your existing classic account remains as-is, so you can migrate your managed identity and deactivate and remove the secondary account once you have validated the merge.

Throughout this document, we use the terms **primary account** and **secondary account**:

- **Primary account** - refers to your tenant-level Microsoft Purview instance.
- **Secondary account** - refers to the Microsoft Purview instance you're migrating into your primary account as a domain.

Prerequisites

- Tenant with multiple Microsoft Purview accounts.
- Public network access enabled on your **secondary account** for the merging process.
- A Microsoft Purview account upgraded to the new experience as a primary account.

Limitations

1. Secondary accounts can only be merged once. Any values added to your secondary account after merging will have to be manually added to your primary account. It's recommended you stop using your secondary account after merging and either deactivate or delete the account.
2. Any values or assets added to a secondary account during an active merge aren't guaranteed to be migrated to the primary account. It's recommended you stop update and scanning activities on your secondary account during a merge.

There are also some current limitations that could prevent you from merging your existing Microsoft Purview account as a domain. Review these limitations and the behavior changes to confirm that merging is the correct option for you. The limitations could be updated in the future.

Expand table

Category	Limitation	Potential resolution
	Primary account can only have five domains.	Contact support to provide your business requirement and increase the domain number.
	Either primary account or secondary account has more than 1,000,000 assets.	Remove assets to the allowed limit, choose to filter out the asset category, or contact support to provide your business requirements.
	Secondary account's collection number + primary account's collection number > 400.	Remove or consolidate collections to the allowed limit.
Scan limitations	Secondary account has a user-assigned managed identity (UAMI) scan.	Remove all scheduled UAMI scans from any data sources or filter out the user assigned managed identity category during merge.
	Scan run histories aren't migrated to the primary account.	No resolution.
	Secondary account has private endpoints enabled.	Remove all private endpoints from secondary account or filter out the private endpoint category during merge.
	Secondary account has a self-hosted integration runtimes.	Remove all self-hosted integration runtimes from secondary account or filter out the self-hosted integration runtime category during merge.
	Secondary account has managed virtual networks.	Remove all managed virtual networks from secondary account or filter out the managed virtual networks category during merge.
	Secondary account has "bring your own Event Hubs" configured.	Remove all Kafka configurations from secondary accounts.
	Secondary account has managed Event Hubs enabled and it's actively used in last 30 days.	Disable managed event hub.
	Secondary account has a data factory connection.	Remove data factory connections.
	Secondary account has a data share connection.	Remove all data share connections.
Asset limitations	Secondary account has lineage scans.	Remove all lineage scans from any data sources.
Source limitations	Secondary account has a Website data source registered.	Unregister all Website data sources

Behavior changes

1. Any secondary accounts' scans that use the managed identity (MSI) to scan a data source will keep using the same MSI after migrating to the primary account. So, if the secondary account is deleted after merge, these migrated scans won't work unless you follow the managed identity switchover process after merging.

2. Any secondary accounts' key vault connections that use managed identity (MSI) to connect will keep using the same MSI after migrating to the primary account. So if the secondary account is deleted after merge, these migrated scans won't work unless you follow the managed identity switchover process after merging.

3. If secondary account has a glossary named *Glossary*, the migrated glossary is named *Glossary-{secondary account name}* in the primary account. This behavior could cause glossary conflicts during assessment:

    - Secondary account already has a glossary named *Glossary-{secondary account name}*
    - Primary account has a glossary named *Glossary-{secondary account name}*

4. If the secondary account was created before 7/12/2022 and the primary account was created after this date, after the merge, the scan will populate hierarchical schema types(parquet_schema, json_schema, xml_schema, and delimited_text_schema) rather than tabular_schema.

Request account merge

You can request an account merge from your primary account or from the secondary account you want to merge.

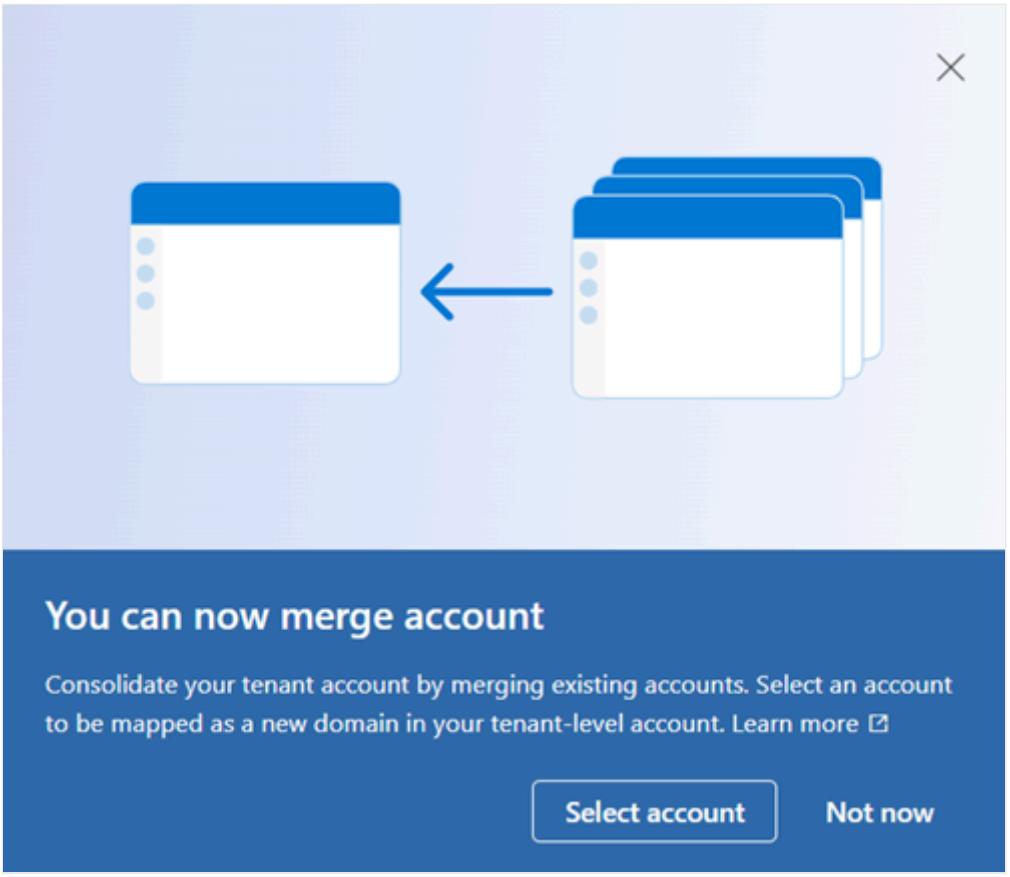
> **Tip**
>
> It's recommended to **assess your merge** before requesting a full merge to check for any conflicts or limitations.

Request account merge from the primary account

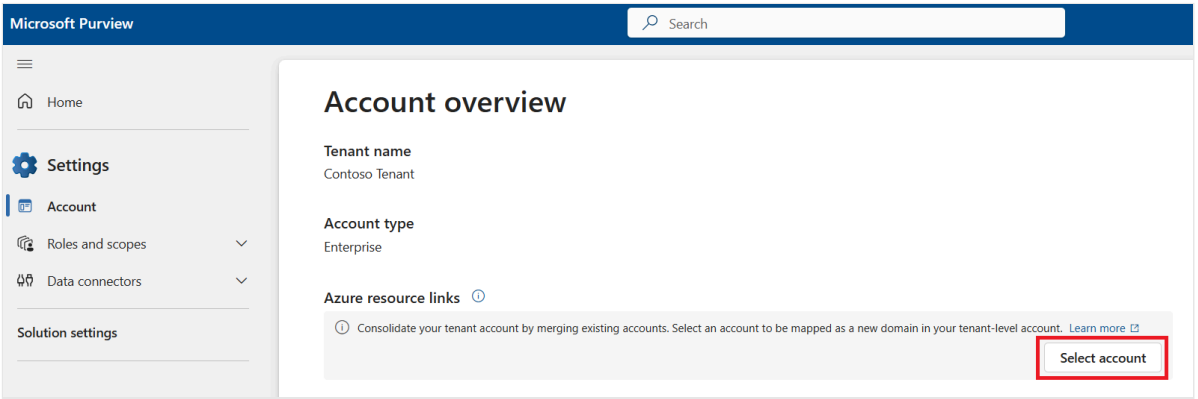
> **Important**

To request account merge from the primary account, requestors need to be a member of the Purview Administrator role group.

1. Open the Microsoft Purview portal.
2. To select an existing account to be mapped as a new domain in primary account. You can either:
   - Select the **Select account** button in the pop-up.

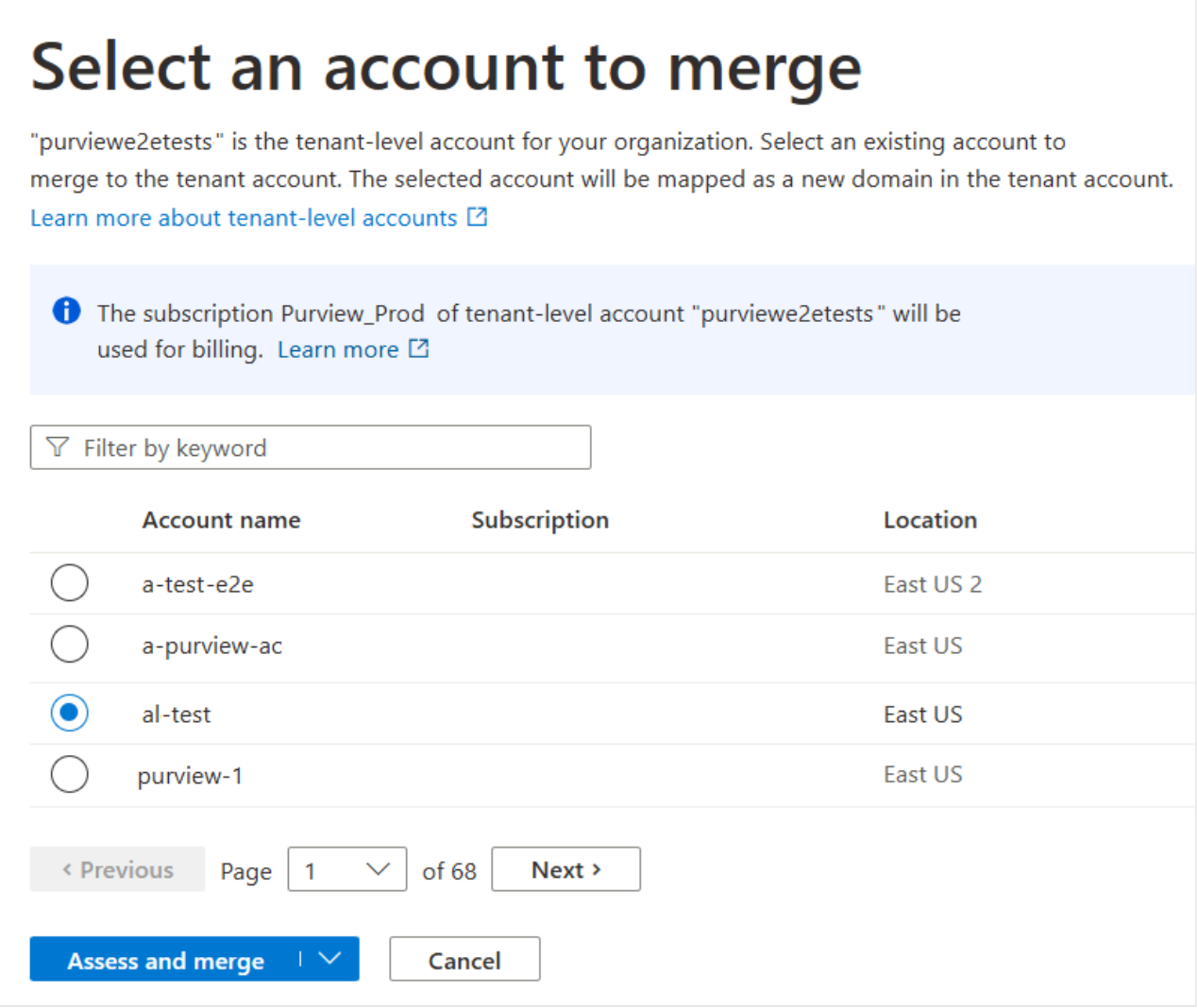


   - If you're using the classic Microsoft Purview governance portal, open the Management center, navigate to the **Overview** in the General section. If you're using the Microsoft Purview portal, open Settings solution card, select **Account** and then select **Select account** in the alert box.



> **Note**
>
> - If a merge is currently in progress within the tenant, then the **Submit request** pop-up and alert box won't appear.
> - If the logged-in user doesn't have the proper permissions to request a merge, the **Submit request** pop-up and alert box won't appear.

3. Select an account to merge. (It's recommended that you perform an assessment before you merge.)



4. Specify categories for the merge filter, and select whether you want to allow auto-resolve for conflicts in that category.

5. Select **Submit**.

6. An assessment runs to confirm if there are any conflicts or limitations, and if the assessment is successful, the merge process begins.

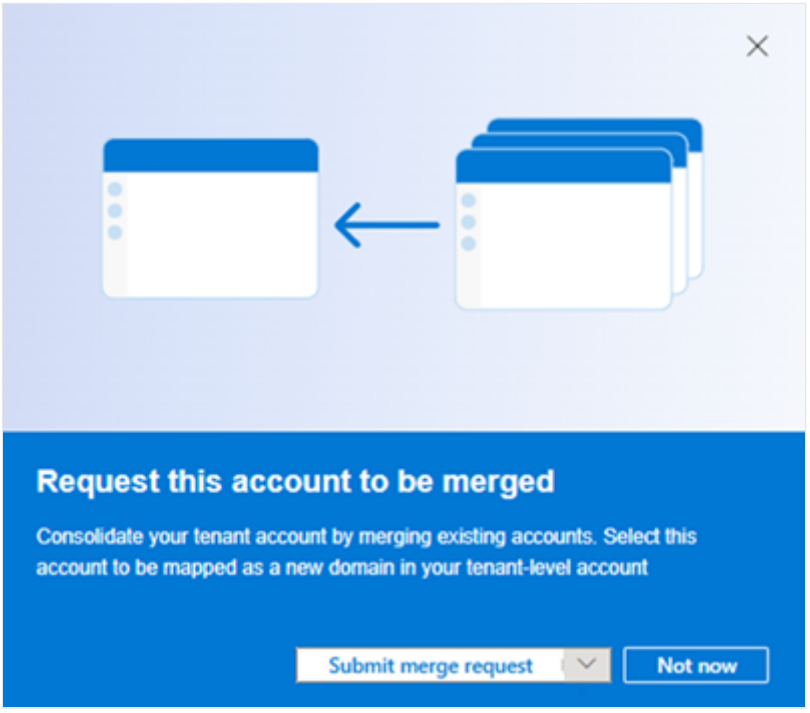
7. You can monitor the merge process from the report page.

> **Important**
>
> Failed merges start the cleanup process to revert everything back to its previous state.

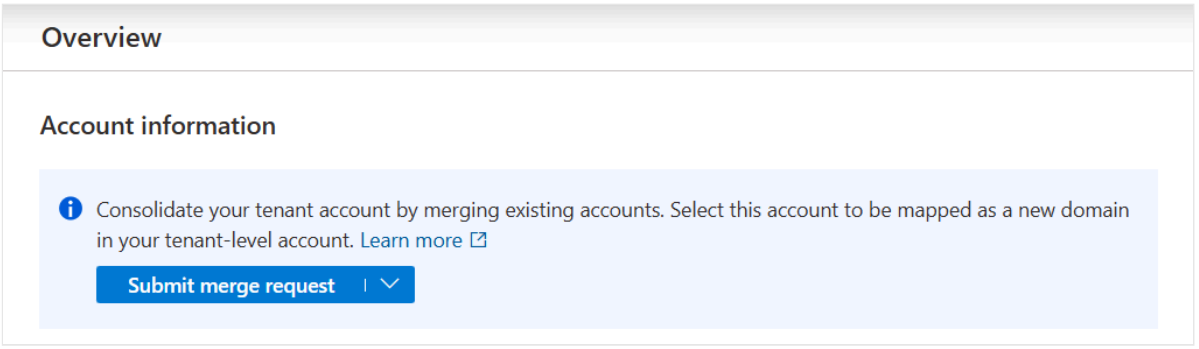
Request account merge from secondary account

> **Important**
>
> To request account merge from the secondary account, requestors need to be either be a Collection Admin on the root collection or an **ARM Resource Owner**, and then they also need to be assigned one of these roles: Compliance Manager Administration, Role Management, Case Management .

1. Open the secondary account's Microsoft Purview portal
2. To select this account to be mapped as a new domain in primary account, you can either:
   - Select **Submit merge request** in the pop-up.



   - Or in the classic Microsoft Purview governance portal, open the Management center, navigate to the **Overview** in the General section, then select the alert box.



> **Note**
>
> - If a merge is currently in progress within the tenant, then the **Submit request** pop-up and alert box won't appear.
> - If the logged-in user doesn't have the proper permissions to request a merge, the **Submit request** pop-up and alert box won't appear.

3. Specify categories for the merge filter, and select whether you want to allow auto-resolve for conflicts in that category.
4. Select **Submit**.
5. The request is sent through an approval process where a Purview Administrator needs to approve the request.
6. After approval, an assessment will run to confirm if there are any conflicts or limitations, and if the assessment is successful, the merge process will begin.
7. You can monitor the merge process from the report page.

> **Important**
>
> Failed merges start the cleanup process to revert everything back to its previous state.

Merge filter

Whenever you merge an account, you can select which objects to migrate and which objects to exclude. In many cases, the filter impacts objects in its own category. However, there are potential behavior changes when you unselect some objects that might apply to associated objects. Some of these categories and the associated behavior changes are listed below:

Expand table

Unselected category	Behavior change
Contact assignment	If unselected, contact assignments configured to glossary terms or assets are removed.
Credential	If unselected, scans that use credentials to scan data sources aren't migrated.
Custom classification rule	If unselected, scans that applied custom classification rules aren't migrated.
Custom scan rule set	If unselected, scans that use applied custom scan rule sets aren't migrated.
Custom type	If unselected: - All custom type assets aren't migrated - All relationships for custom types and affected assets aren't migrated. - All classification instances of custom types attached to glossary terms or assets aren't migrated. - Managed attributes attached to glossary terms or assets are removed.
Glossary	If unselected, workflows applied to glossary or glossary term aren't migrated, and assets won't have associated glossaries.
Key vault connection	If unselected, scans that use key vault credentials to scan data sources aren't migrated.
Managed virtual network	If unselected, scans that use managed virtual network to scan data sources aren't migrated.
Self-hosted integration runtime	If unselected, scans that use a self-hosted integration runtime to scan data sources aren't migrated.
User assigned managed identity	If unselected, scans that use a user assigned managed identity to scan data sources aren't migrated
Resource set rule	If unselected, future assets scanned will have no resource set rule applied util users configure new resource set rules in the primary account.
Contact assignment	If unselected, contact assignments configured to assets and glossary terms are removed.

Approval for account merges

If the merge is requested from the secondary account, it needs to be approved by a Purview Administrator before proceeding with the merge. Purview Administrators receive an email notification to approve the merge request. Alternatively, the merge request can be approved within primary account's Microsoft Purview portal.

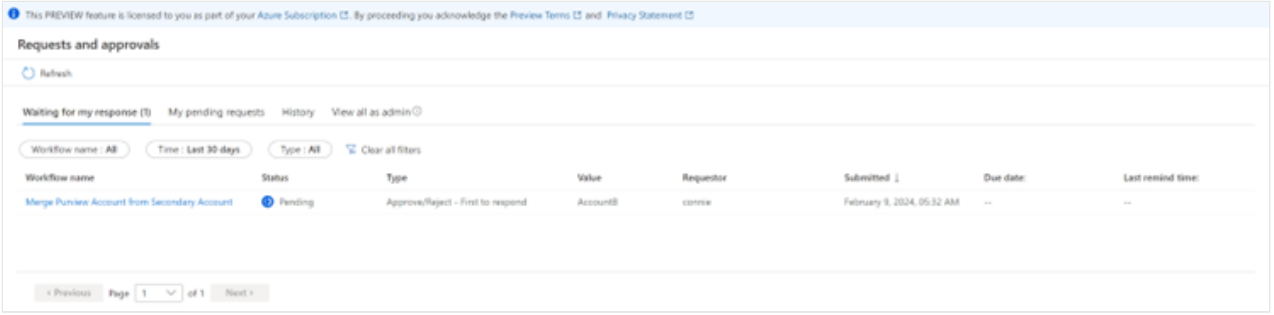
> ⓘ **Important**

If the outlook account and Microsoft Purview accounts are in different tenants, approving from the email might fail. In this case, use the Microsoft Purview portal to approve.

1. Open primary account's Microsoft Purview portal.

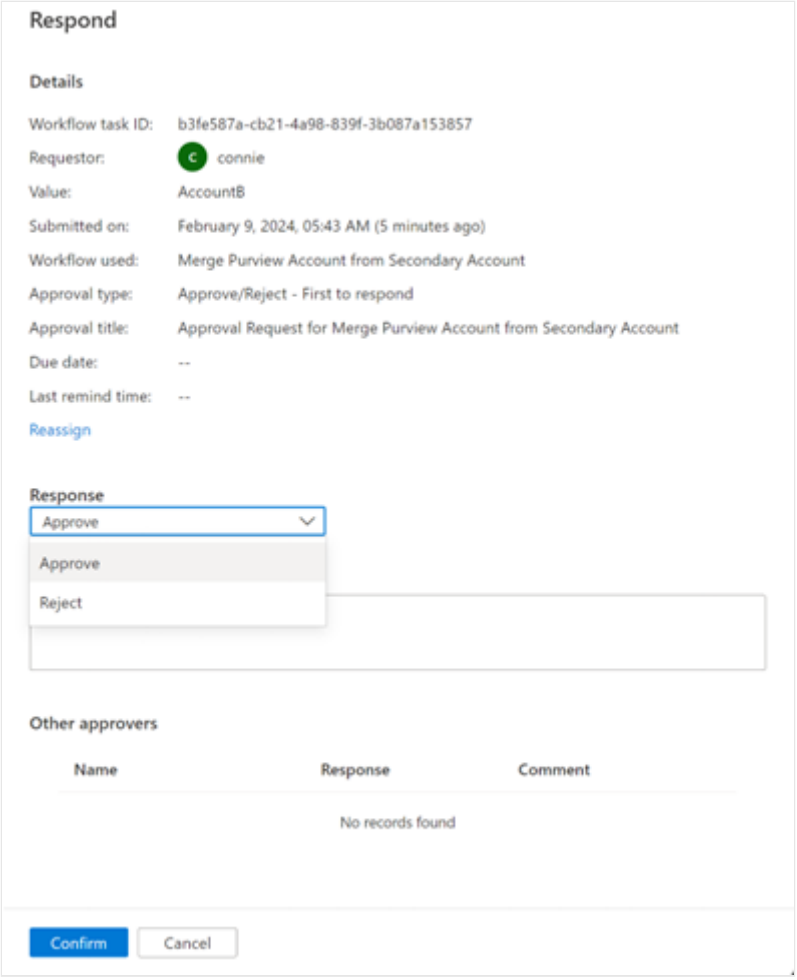
2. Depending on which portal you're using:

   - If you're using **the Microsoft Purview governance portal**:
     a. Open the Management center
     b. Select **Request and approvals** in the Workflow section.
   - If you're using the **Microsoft Purview portal**:
     a. open Workflows solution card
     b. select **Requests and approvals**.



3. Select the request.

4. Choose your response and select **Confirm**.



Assess conflicts and limitations

Before you run a merge, Microsoft Purview will assess a secondary account to see if it meets the prerequisites for merging with the primary account. If conflicts/limitations are detected and validation fails, manual intervention is required to proceed with the merge. After assessment, you'll have a report of any conflicts, and you can use our conflicts and resolutions to help resolve them.

Auto-resolve

During the Account Merge process, if the auto-resolve rule is set for a category, the system will automatically resolve any conflicts encountered according to the rule. Currently, only the **skip** rule is supported, which means the object causing the conflict is skipped and not moved from the secondary account to the primary account. Any other objects that depend on this object will also be skipped.

For example: If an auto-resolve rule is set for the **Glossary** category, and a Glossary term encounters a conflict, the system will automatically skip the merge of that term, and also skip the merge of any of that term's children.

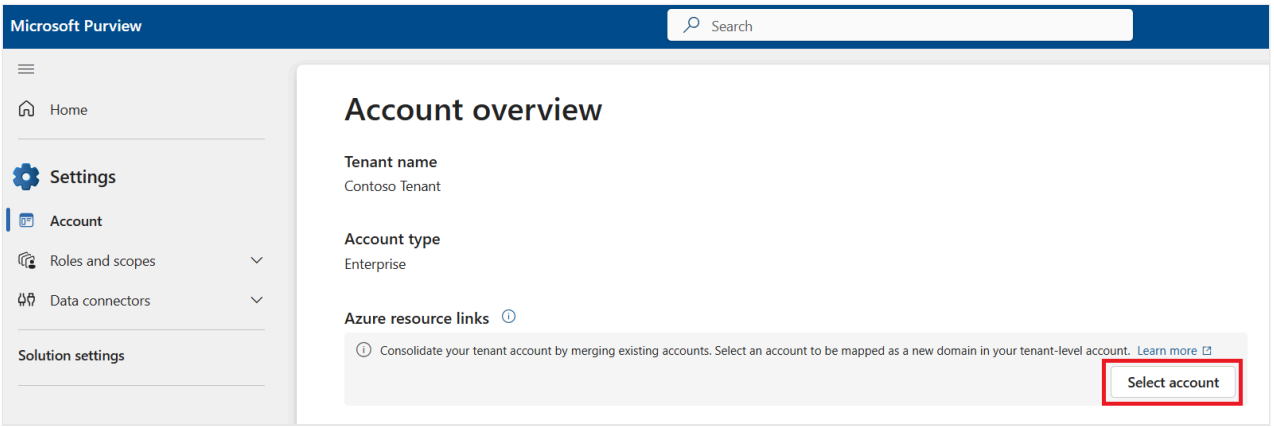
Any conflicts encountered when auto-resolve is enabled won't be shown as errors in the assessment or merge process, because they'll be automatically resolved according to the rule.

Assessment only

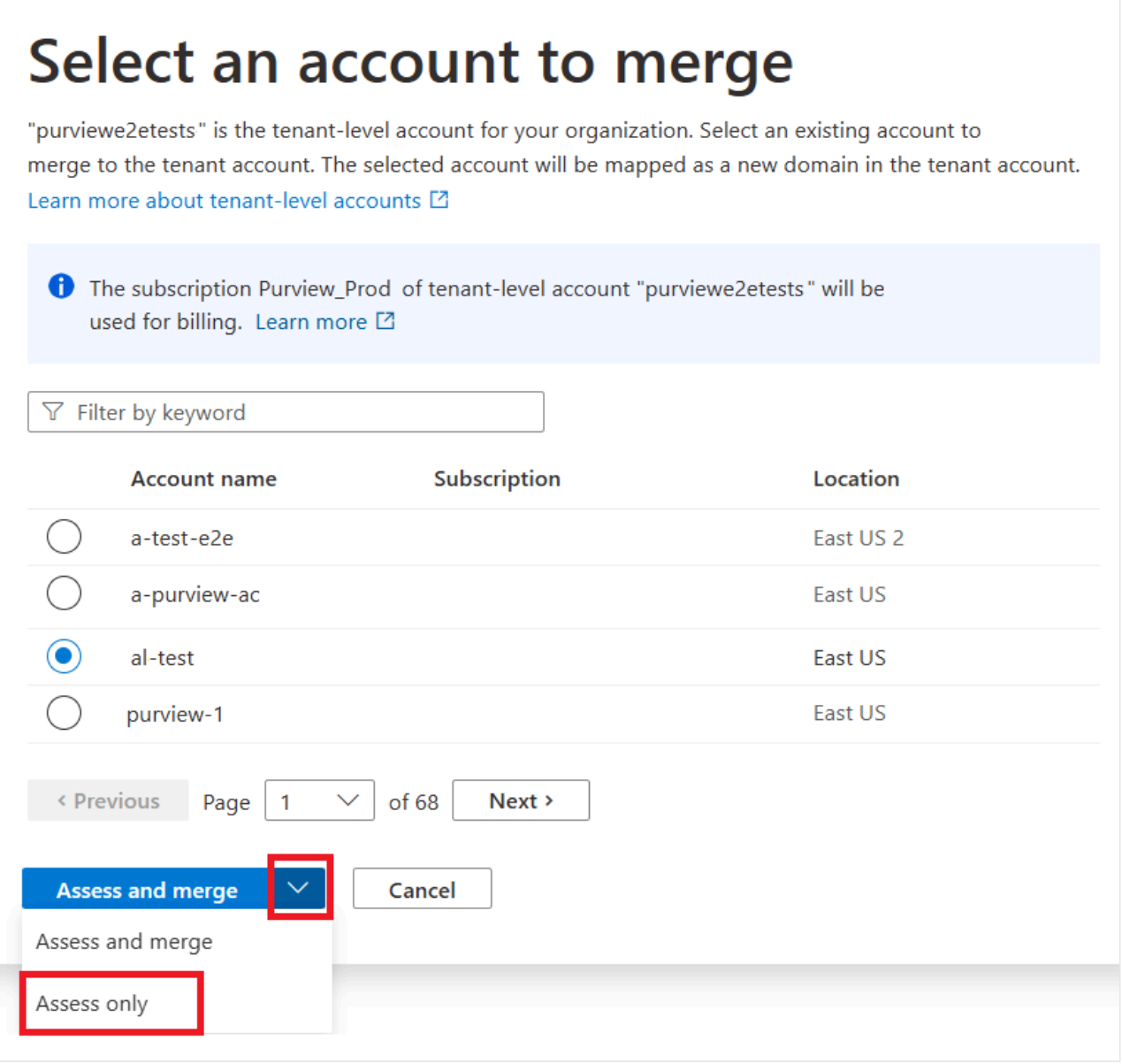
You can submit an account for a merge assessment without merging it, to check for any conflicts.

To run an assessment from your primary account:

1. Open the Microsoft Purview portal.
2. To assess an existing account for merging, if you're using the classic Microsoft Purview governance portal, open the Management center, navigate to the **Overview** in the General section. If you're using the Microsoft Purview portal, open Settings solution card, select **Account** and then select **Select account** in the alert box.

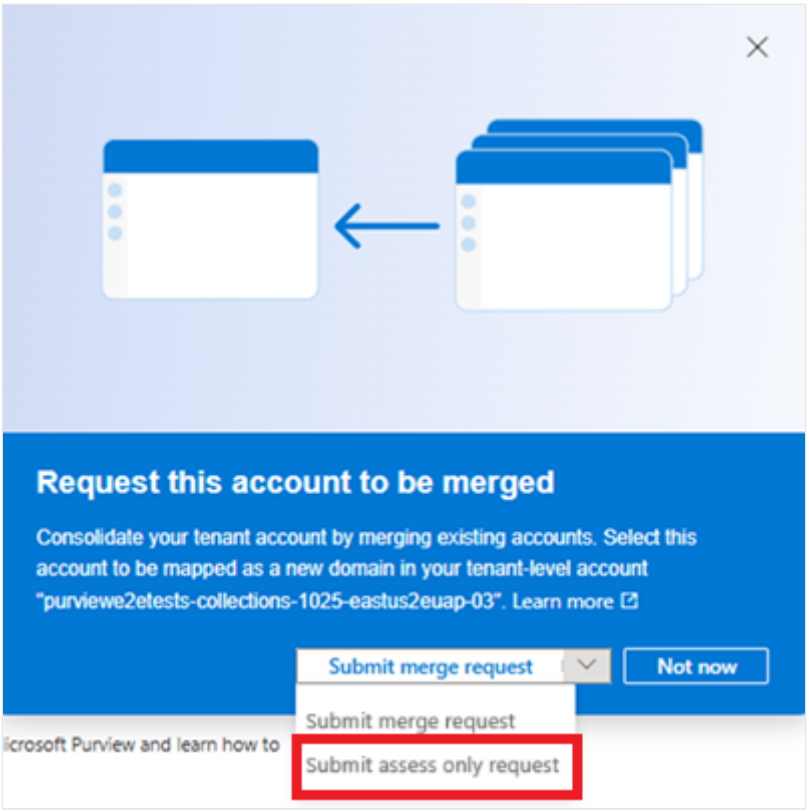


3. Select an account to assess.
4. Select the drop-down on the **Assess and merge** button and select **Assess only**. Only an assessment is run and the merge process won't start afterwards, regardless of the assessment result.



To run an assessment from your secondary account:

1. Select the dropdown in the pop-up and select **Submit assess only request**.



After assessment, you'll have a report of any conflicts, and you can use our conflicts and resolutions to help resolve them.

Conflicts and resolutions

> **Tip**
>
> For many conflicts, three potential options to resolve the issue are:
>
> 1. Remove the conflicting asset from either the primary or secondary account.
> 2. Choose to **filter out** that asset category during merge.
> 3. Select **auto-resolve** to skip migration for conflicting assets.

Here's a table of conflicts and their possible resolutions:

Expand table

Conflict	Potential resolution
Asset with the same qualified name exists in both the primary and secondary accounts.	There are a few options: - Remove the conflicting asset from either primary or secondary account - Choose to filter out the assets category - Select auto-resolve to skip migration when initiating a merge.

Conflict	Potential resolution
Glossary with the same name exists in both the primary and secondary accounts.	Remove the conflict glossary from either the primary or secondary account or filter out the glossary category during merge.
Glossary term with the same qualified name exists in both the primary and secondary accounts.	Remove the conflict glossary term from either the primary or secondary account.
Term template with the same name exists in both the primary and secondary accounts.	Remove the conflict term template from either the primary or secondary account.
Classification rule with the same name exists in both the primary and secondary accounts.	Remove the conflict classification rule from either the primary or secondary account.
Workflow with the same name exists in both the primary and secondary account.	Remove the conflict workflow from either the primary or secondary account.
Key vault connection with the same name exists in both the primary and secondary accounts.	Remove the conflict key vault connection from either the primary or secondary account.
Credential with the same name exists in both the primary and secondary accounts.	Remove the conflict credential from either the primary or secondary account.
Data source with the same data source name registered in both the primary and secondary accounts.	Unregister the conflict data source from either the primary or secondary account.
The same data source registered in both the primary and secondary account.	Unregister the conflict data source from either the primary or secondary account.
Data policy with the same policy name exists in both the primary and secondary accounts.	Remove the conflict policy from either the primary or secondary account: DevOps policy, data owner policy, self-service policy.
Scan rule set with the same name exists in both the primary and secondary accounts.	Remove the conflict scan rule set from either the primary or secondary account.
The same contact assignment exists in both the primary and secondary accounts.	Remove the conflict contact assignment from either the primary or secondary account.
Secondary account has a glossary named 'Glossary' and another glossary named 'Glossary-{secondary account name}'	Remove either 'Glossary' or 'Glossary-{secondary account name}' from the secondary account.
Secondary account has a glossary named 'Glossary' and primary account has a glossary named 'Glossary-{secondary account name}'	Remove 'Glossary' from secondary account or remove 'Glossary-{secondary account name}' from the primary account.
Collections with the same name (not display name) exists in both the primary and secondary accounts.	Remove the conflicting collection from either the secondary account or the primary account.

View merge and assessment reports

You can track the merge progress in real time during the merging process. To view merge report, users need to be a member of Purview Administrator role group.

1. Open the Microsoft Purview portal.
2. Open the Data Map and select **Monitoring**.
3. The merge reports are located on the **Account merging** tab.
4. Select any account name to see full information about its merge or assessment. Select the account name again to collapse its details and see the full account list.

On the report page, you can see the merge statistics summaries:

- Total accounts: Number of accounts that have attempted to merge
- Merging in progress: Accounts currently undergoing merging
- Completed merges: accounts that have successfully merged
- Failed merges: accounts that failed to merge

For any completed assessment, you can rerun an assessment, or rerun an assessment and merge, by selecting **Rerun assessment the assessment details.

Merge and assessment summary

On the account merging page, select an account, and you can view the merge or assessment summary and the detailed progress for that account.

- Assessment: the step that identifies conflicts and limitations between the primary and secondary accounts.

  > **Note**
  >
  > The discovered conflicts and limitations are blocking the merge from proceeding. You must resolve all of them before selecting **Rerun**.

- Migration: the step that copies the objects from the secondary account to the primary account

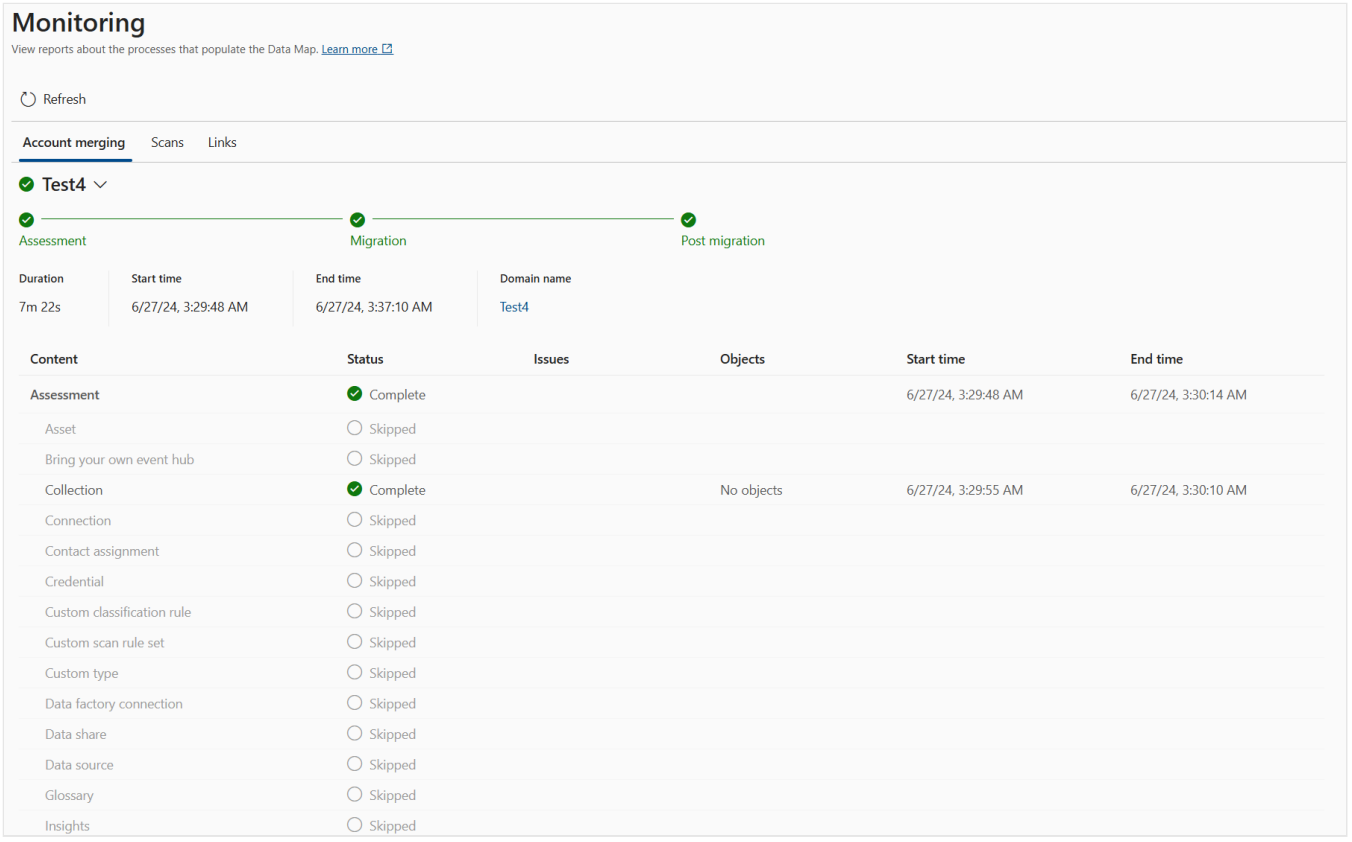
  > **Note**
  >
  > - Migrated objects in the primary account are invisible, and their associated functions are disabled during this step. For example, scan schedules are disabled

in the primary account during migration and will be enabled before merge complete.

  - Since both the primary and secondary accounts remain operational during the merge, there might be new conflicts and limitations. In such cases, migration will fail and automatic cleanup be triggered to roll back the state of the primary account back to its state before the merge.

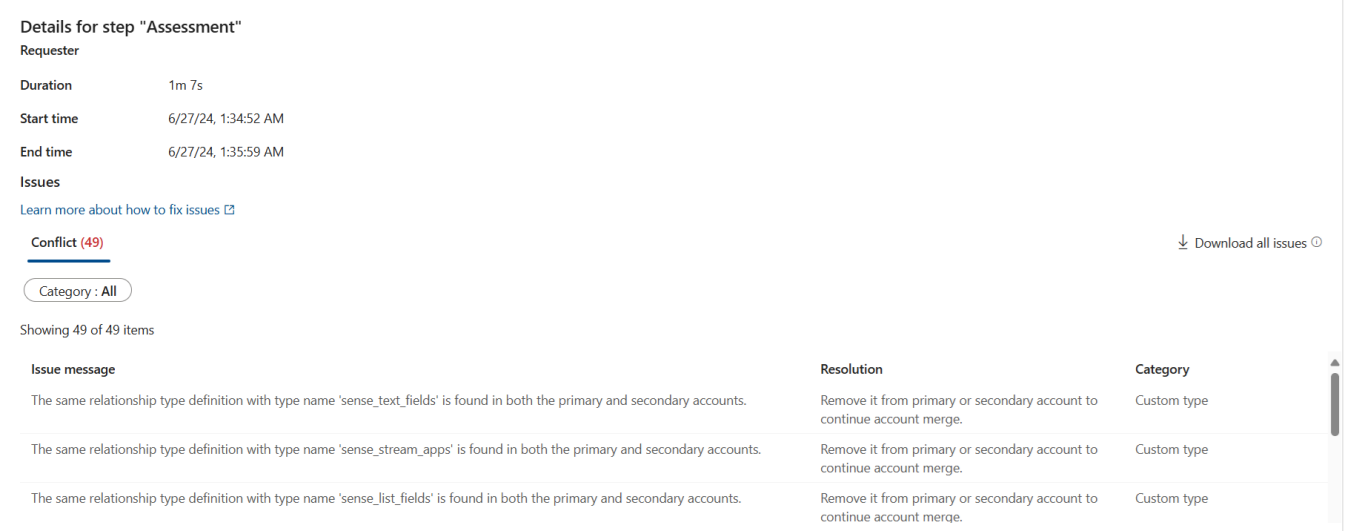
- Cleanup: the step that is automatically triggered to clean up migrated objects from the primary account in case the prior step **Migration** fails.
- PostMigration: the step involves preparation work before merge completes.

You can also view detailed merge progress at object category level under a step (Assessment/Migration/PostMigration/Cleanup).



> **Tip**
>
> Objects that were filtered to not be merged will show as **Skipped** in the report.

You can view issues details by selecting **View details**. It lists all conflicts and limitations detected during the assessment, as seen in this example:



Use filters to examine conflicts and limitations within selected categories.

Download report

You can download all conflict details for any failed assessment.

1. Go to **Data Map** -> **Monitoring** -> **Account merging**.
2. Select the secondary account's assessment.
3. Select the link in the **Issues** column.
4. Select **Download all issues** to download the conflict detail report that has **Description**, **Object category**, **Resolution**, and **Type** columns.

> **Tip**
>
> Objects that were **filtered to not be merged** will show as **Skipped** in the report.

After merge completion

After the merge has successfully completed, a new domain will appear in the primary account. It could take a few minutes before the domain appears. Its display name is the root collection's friendly name of the secondary account. This contains all the domain level data migrated from the secondary account.

Next, you should switch over your managed identity connections from your secondary to your primary account using the managed identity switchover process.

After merger completion, the merged account still incurs billing. To prevent billing through the merged account, you can either:

- Delete the account, or

- Deactivate the account.

Hard deletion of merged secondary account

Hard delete the Purview account will also delete System Assigned Managed Identity associated with the secondary account. To ensure your scans and resources will connect after deletion of the secondary account, you can use the managed identity switchover process to migrate your resources to the primary account's managed identity.

1. Open the Azure portal and select your Microsoft Purview account.
2. Select the **Delete** button and confirm the request.
3. Wait for the deletion to successfully complete.

After hard deleting the account:

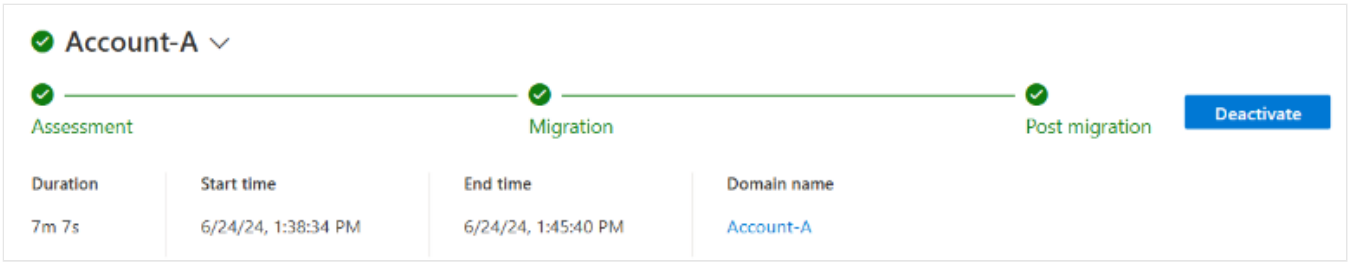
- Migrated scans that use secondary account's system assigned managed identity to access data source won't work in the Primary account.
- Migrated key vault connections that use secondary account's system assigned managed identity to access key vaults won't work in the Primary account.

Deactivate merged secondary account

If you want to stop billing on your secondary account, or you want to use secondary account's system assigned managed identity to access data sources and key vaults in the primary account, you can choose to deactivate the secondary account after merge complete. After the account is deactivated, the data plane operations for this account will be blocked and billing will stop. To switch your sources and scans to the primary managed identity, use the managed identity switchover process.

> ⊗ **Caution**
>
> Once an account is deactivated, it can't be reactivated.

1. Open secondary account's Microsoft Purview portal.
2. Open Data Map and select **Monitoring**.
3. Select the **Account merging** tab.
4. Select **Deactivate**.



Managed identity switchover

When a secondary account is merged into the tenant-level account, switching your managed identity to the new tenant account is a required step to use the self-hosted integration runtime (SHIR), and is still highly recommended even if you're not. Once the secondary account is deleted, any scans or key vaults that use its managed identity will fail.

To switch your scans and resources from your secondary accounts' managed identity to your primary account's managed identity, follow the steps below. Users must hold the Domain admin role in order to perform these functions.

1. Ensure that your primary/tenant-level Microsoft Purview account's managed identity has access to the migrated sources and connections.
2. You can start the switchover from:
    a. The domain overview in your secondary account
        i. If your secondary account was using its own managed identity, on the domain overview page there's a **Switch managed identity button**. Select it.
        ii. Select **Confirm**
        iii. Once the switchover is completed, you receive a notification.
    b. On the create/edit scan page of your secondary account
        i. Once a merge has been completed, on the secondary account's scan pages there's a notification that asks you to confirm switching your managed identity.
        ii. Select **Confirm**
        iii. Once the switchover is completed, you receive a notification.
    c. On the create/edit key vault connection page in your secondary account
        i. Once a merge has been completed, on the secondary account's key vault connection page there's a notification that asks you to confirm switching your managed identity.
        ii. Select **Confirm**
        iii. Once the switchover is completed, you receive a notification.

Billing

After merge completion, billing for the secondary account will be consolidated with the primary account and will be charged to the primary account's subscription. However, until you deactivate the secondary account, its billing will still be applicable to its subscription.

> **Tip**
>
> Once you've merged your secondary account, **deactivate** and then delete it to reduce billing.

Related content

- Manage domains and collections

Create and manage collections in Microsoft Purview Data Map (classic)

Article • 05/30/2025

> **Tip**
>
> If you're using the new **Microsoft Purview portal experience**, refer instead to **how to create and manage domains and collections** .

Collections in the Microsoft Purview Data Map can be used to organize assets and sources by your business's flow. They're also the tool used to manage access across the classic Microsoft Purview governance portal. This guide will take you through the creation and management of these collections, as well as cover steps about how to register sources and add assets into your collections.

Prerequisites

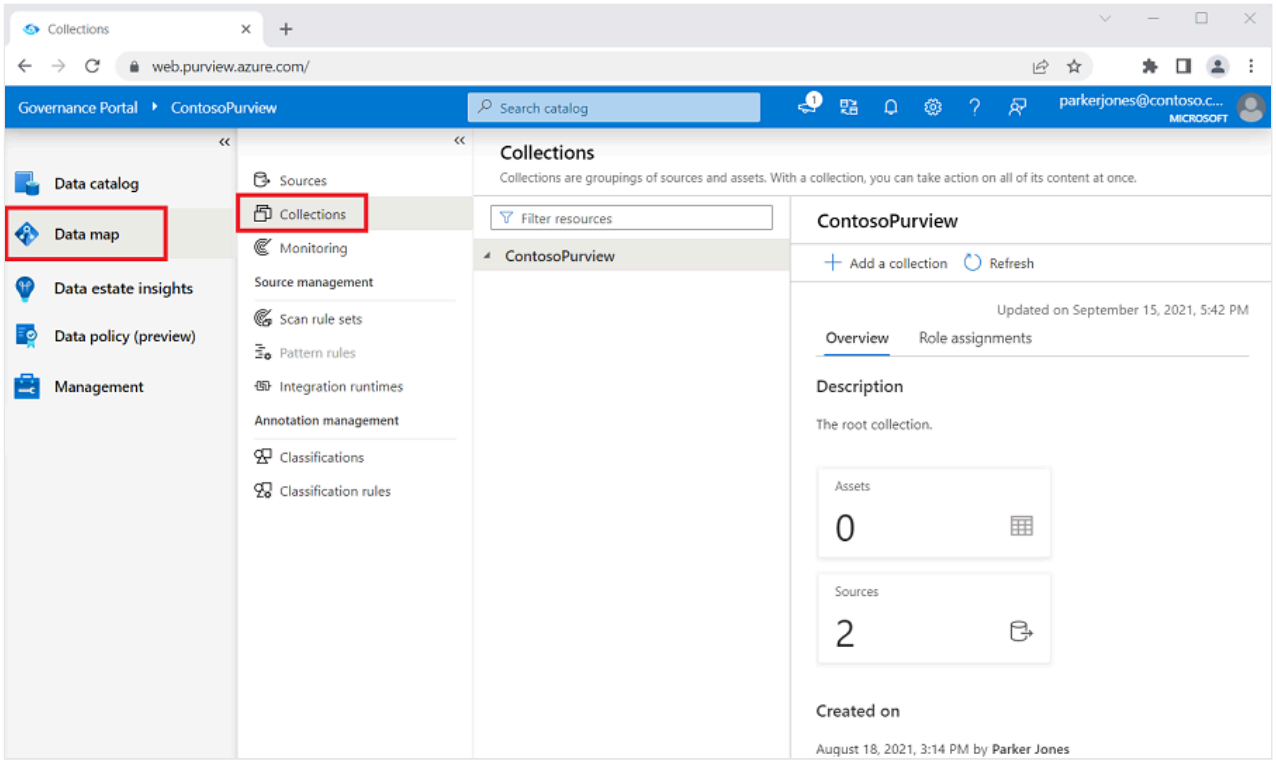
- An Azure account with an active subscription. Create an account for free .
- Your own Microsoft Entra tenant.
- An active Microsoft Purview (formerly Azure Purview) account.

Check permissions

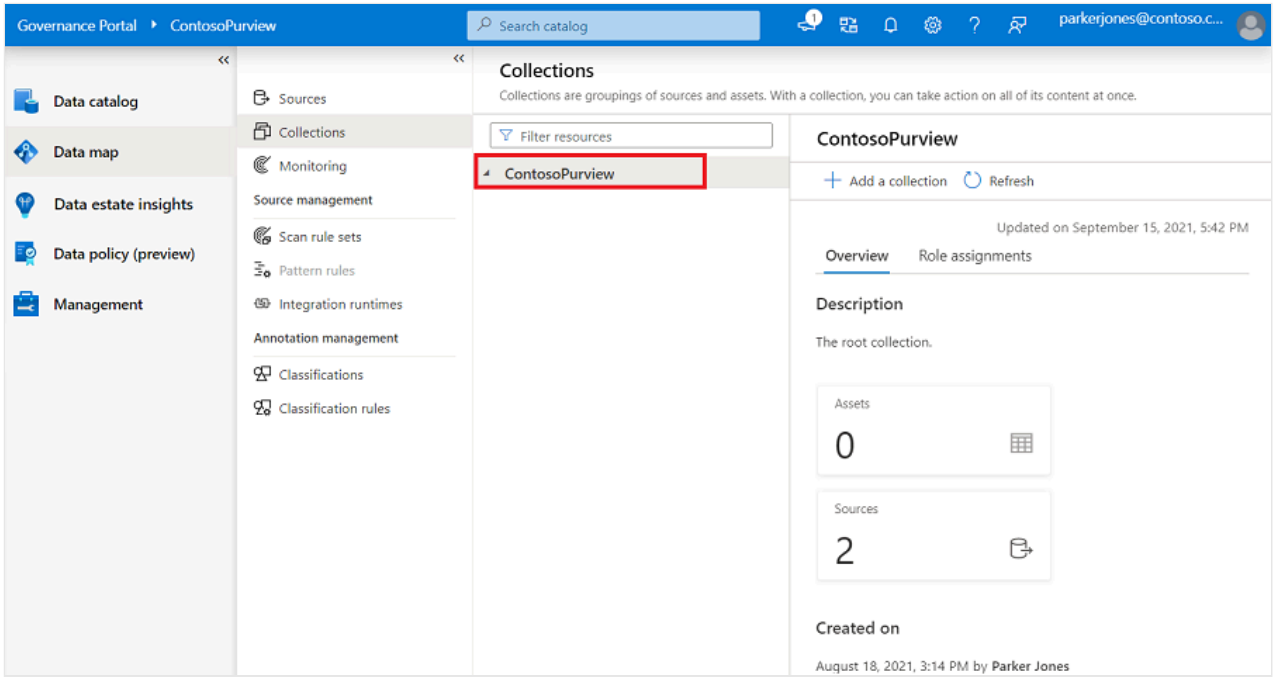
In order to create and manage collections in the Data Map, you'll need to be a **Collection Admin** within the classic Microsoft Purview governance portal. We can check these permissions in the classic Microsoft Purview governance portal . You can find the portal by:

- Browsing directly to https://web.purview.azure.com and selecting your Microsoft Purview account.
- Opening the Azure portal , searching for and selecting the Microsoft Purview account. Selecting the **Microsoft Purview governance portal** button.

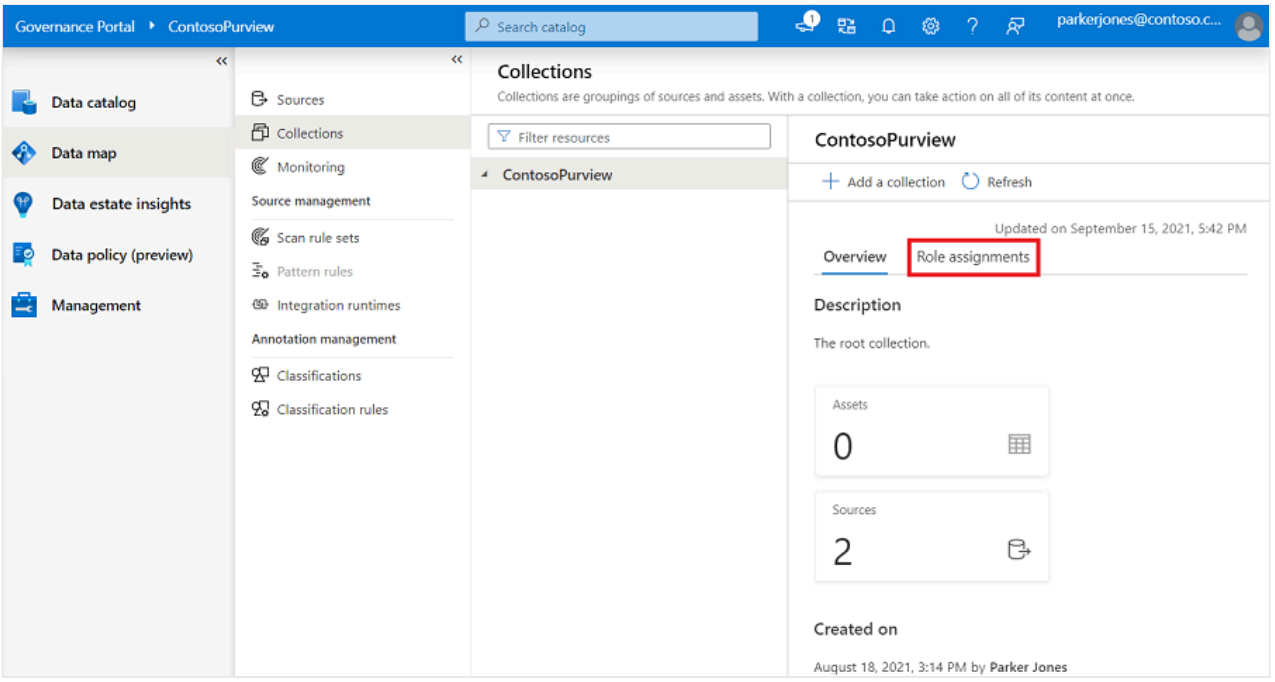
1. Select Data Map > Collections from the left pane to open collection management page.



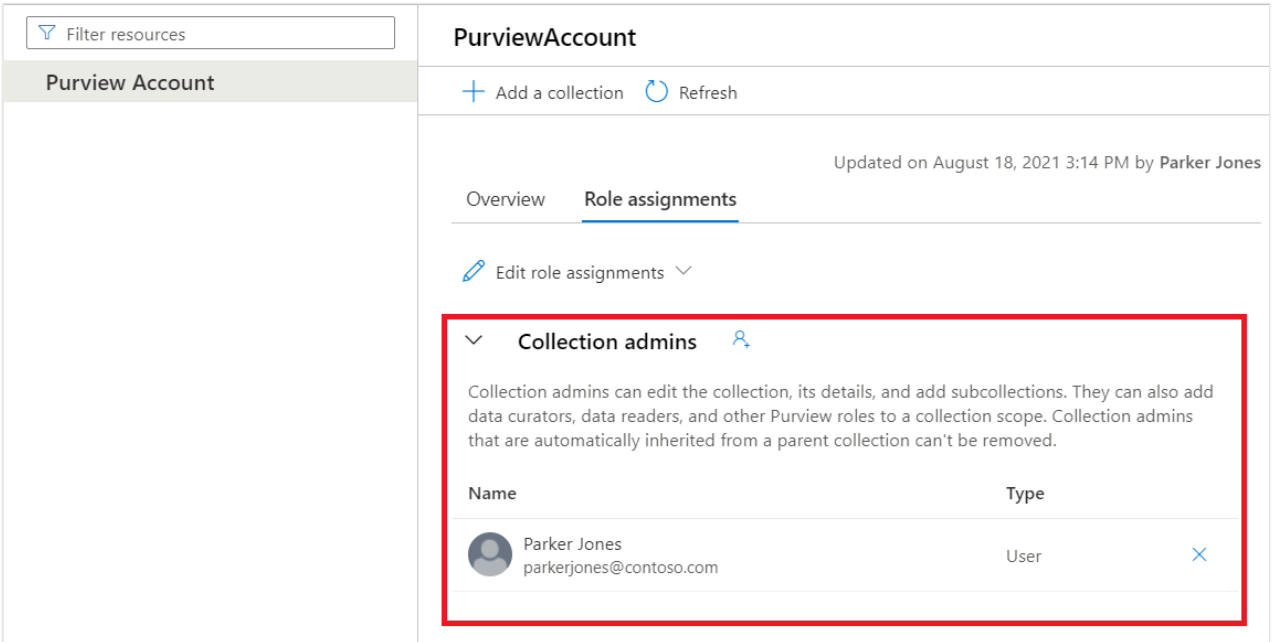
2. Select your root collection. This is the top collection in your collection list and will have the same name as your account. In the following example, it's called Contoso Microsoft Purview. Alternatively, if collections already exist you can select any collection where you want to create a subcollection.



3. Select **Role assignments** in the collection window.



4. To create a collection, you'll need to be in the collection admin list under role assignments. If you created the account, you should be listed as a collection admin under the root collection already. If not, you'll need to contact the collection admin to grant your permission.

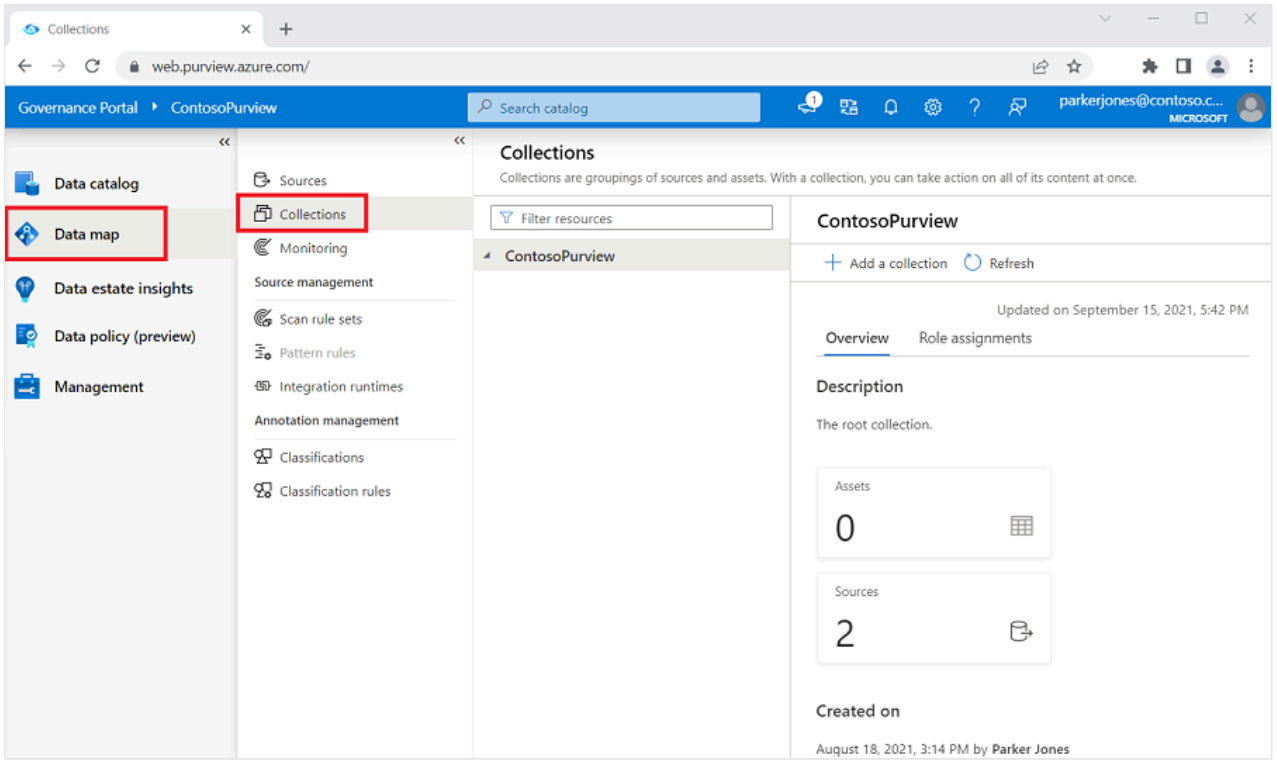


Collection management

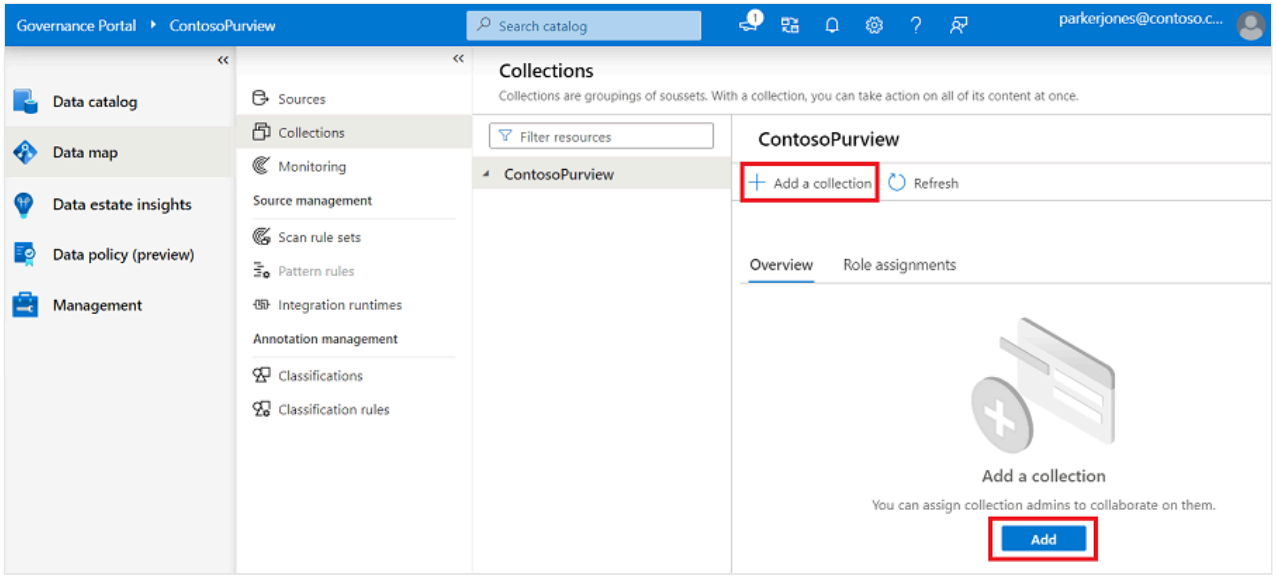
Create a collection

You'll need to be a collection admin in order to create a collection. If you aren't sure, follow the guide above to check permissions.

1. Select Data Map > Collections from the left pane to open collection management page.

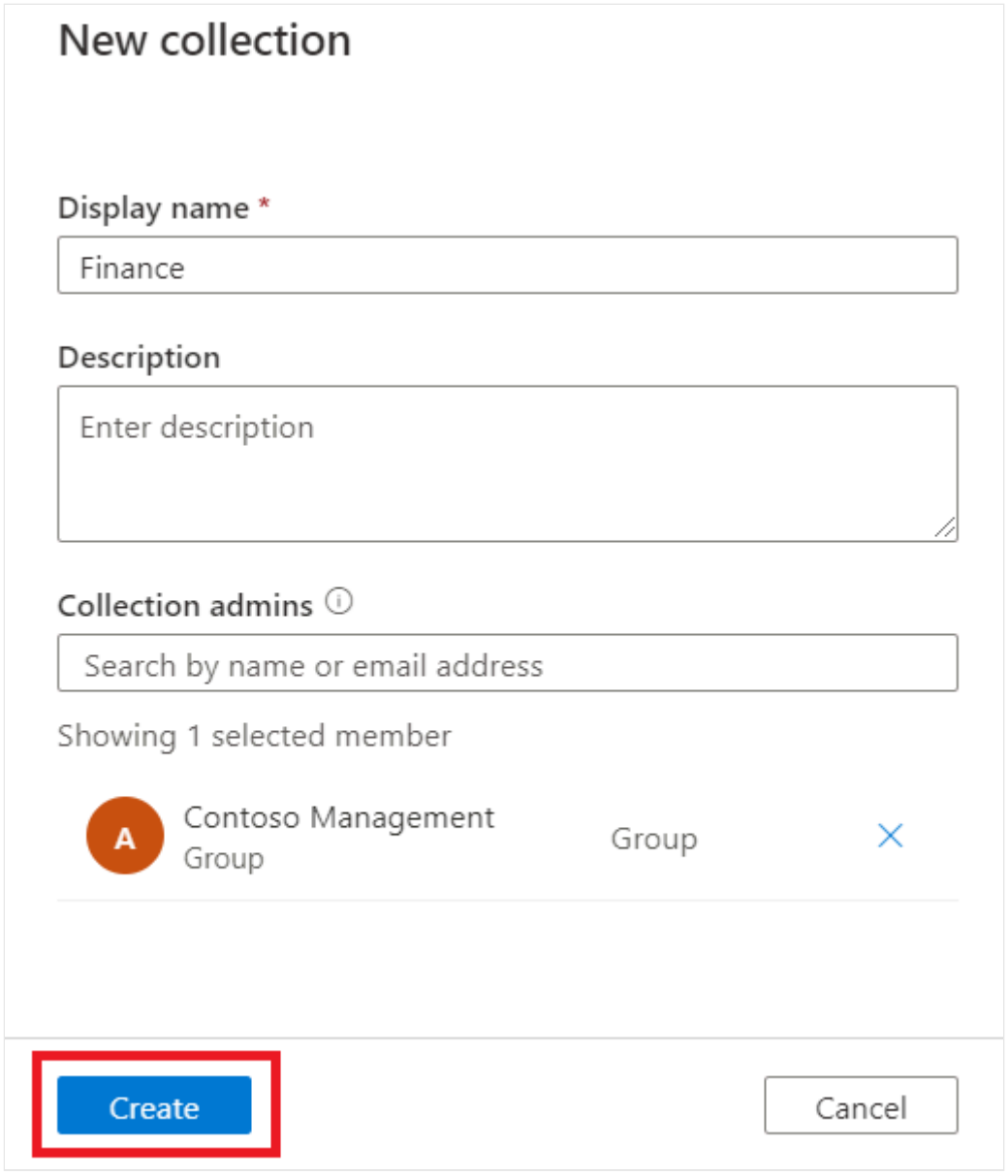


2. Select + **Add a collection**. Again, note that only collection admins can manage collections.

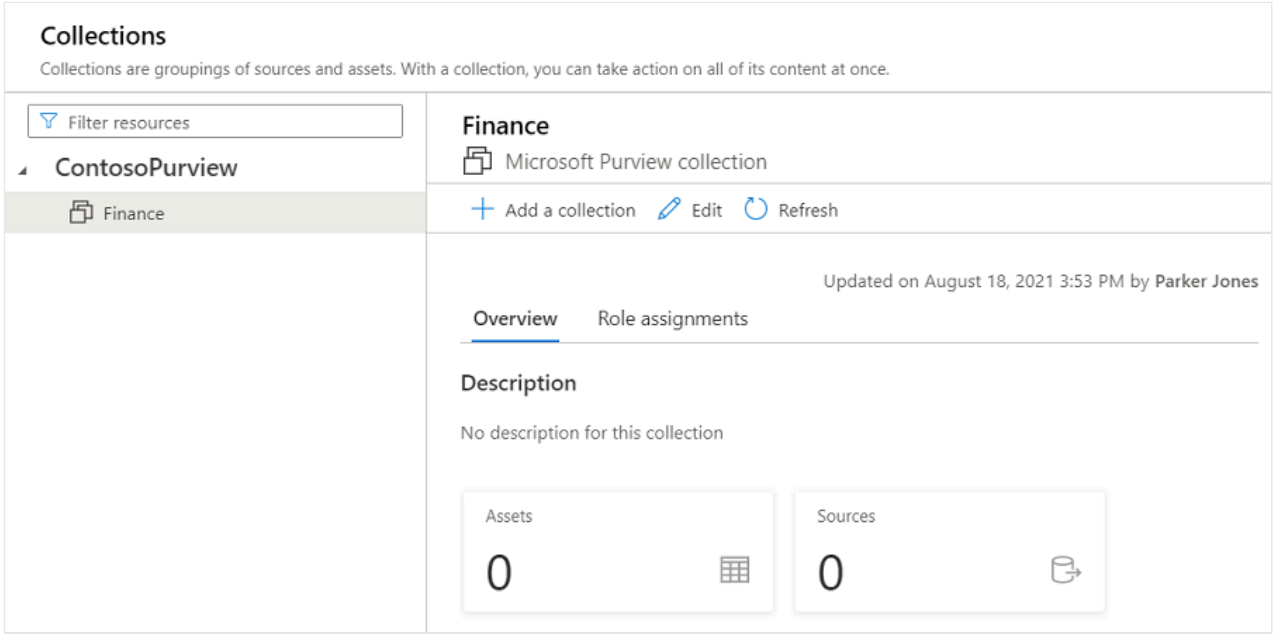


3. In the right panel, enter the collection name and description. If needed you can also add users or groups as collection admins to the new collection.

4. Select **Create**.

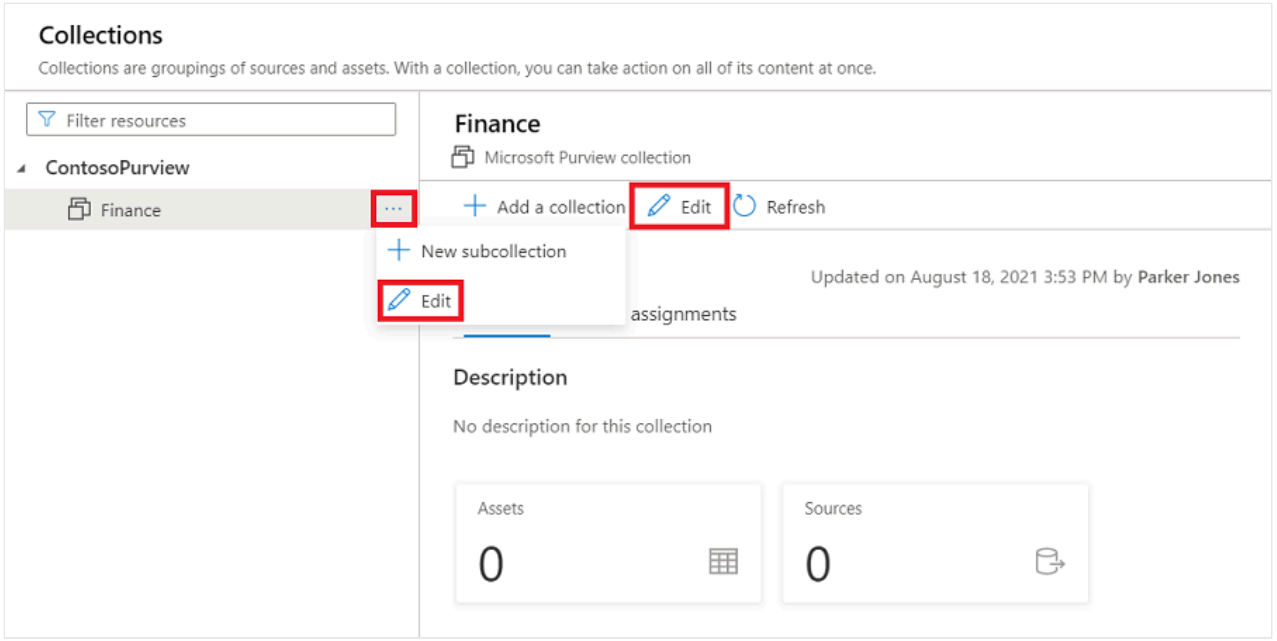


5. The new collection's information will reflect on the page.

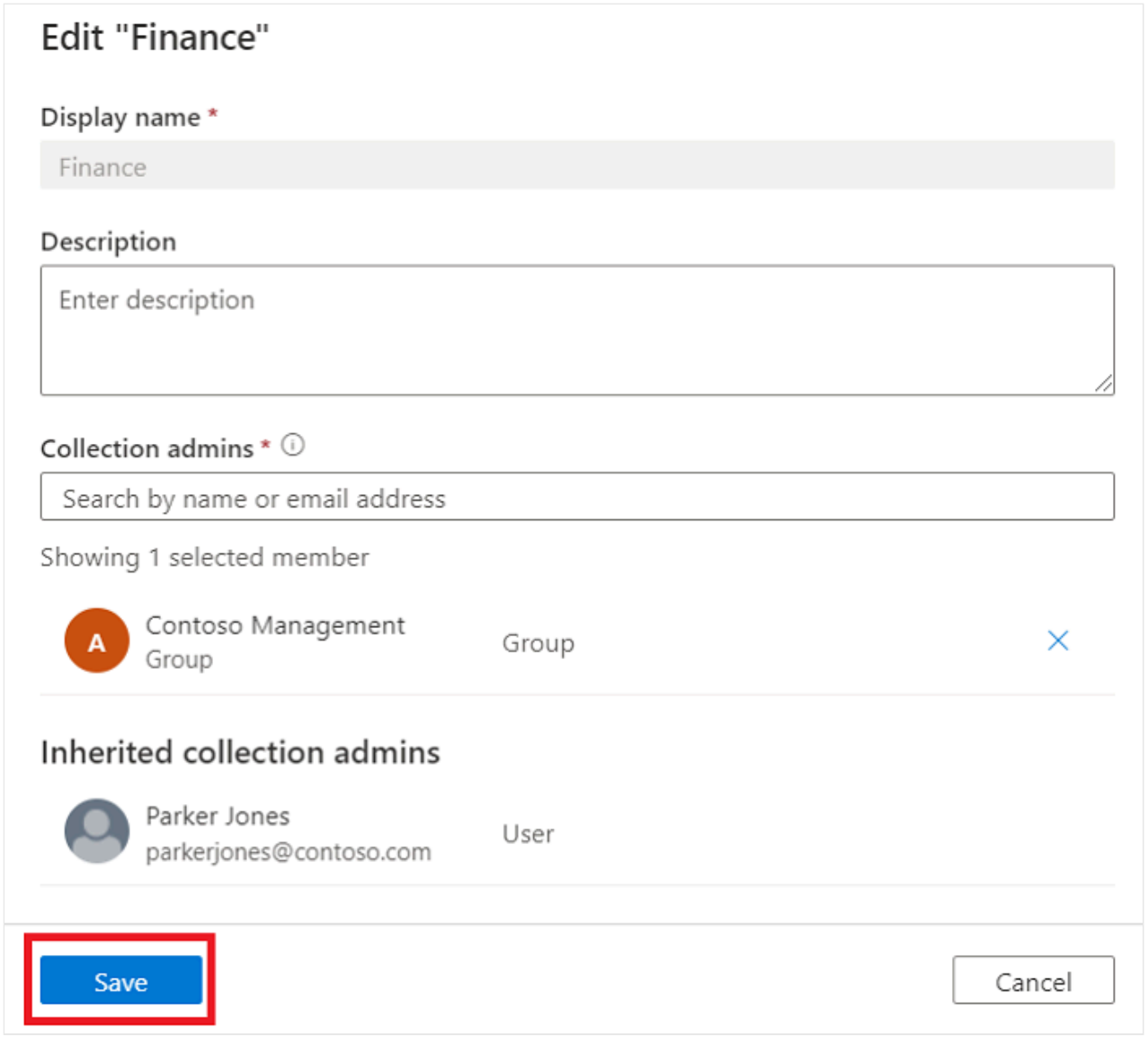


Edit a collection

1. Select **Edit** either from the collection detail page, or from the collection's dropdown menu.

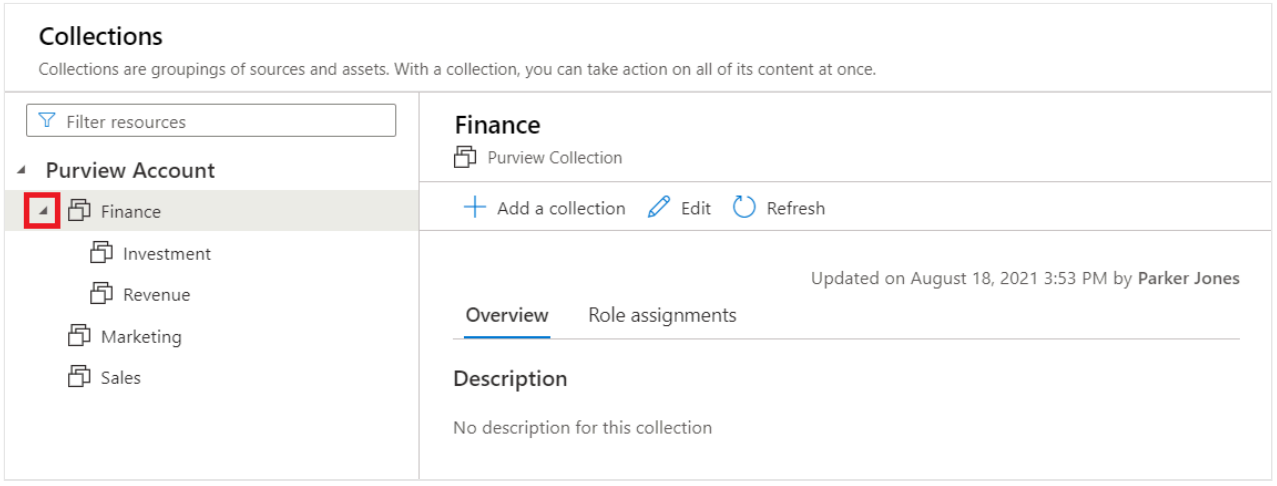


2. Currently collection description and collection admins can be edited. Make any changes, then select **Save** to save your change.

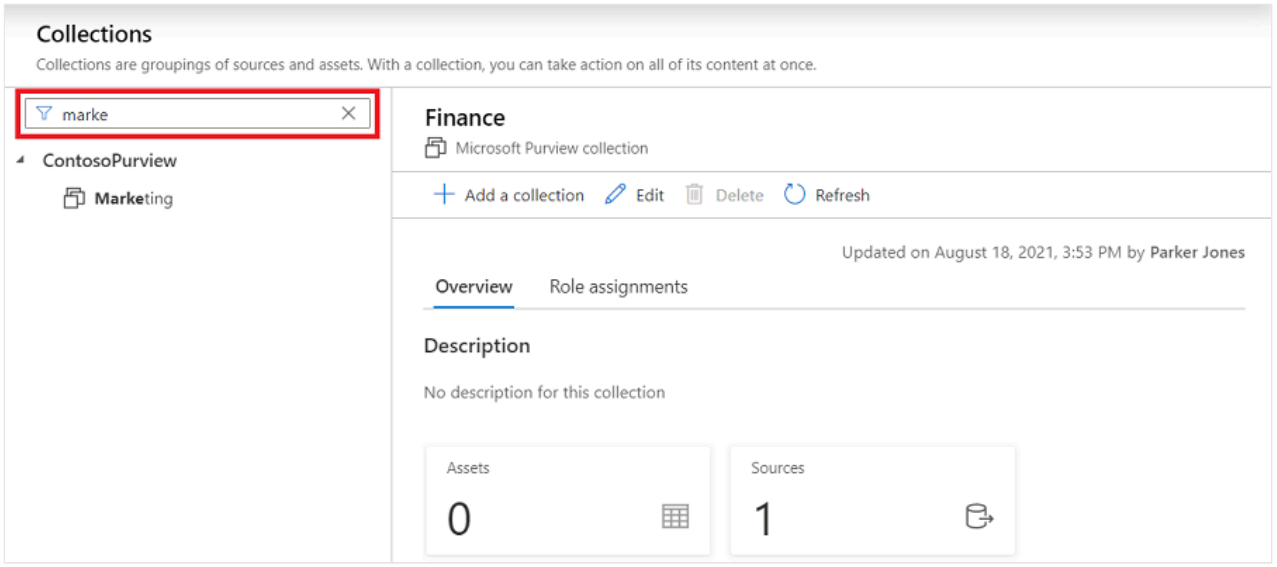


View Collections

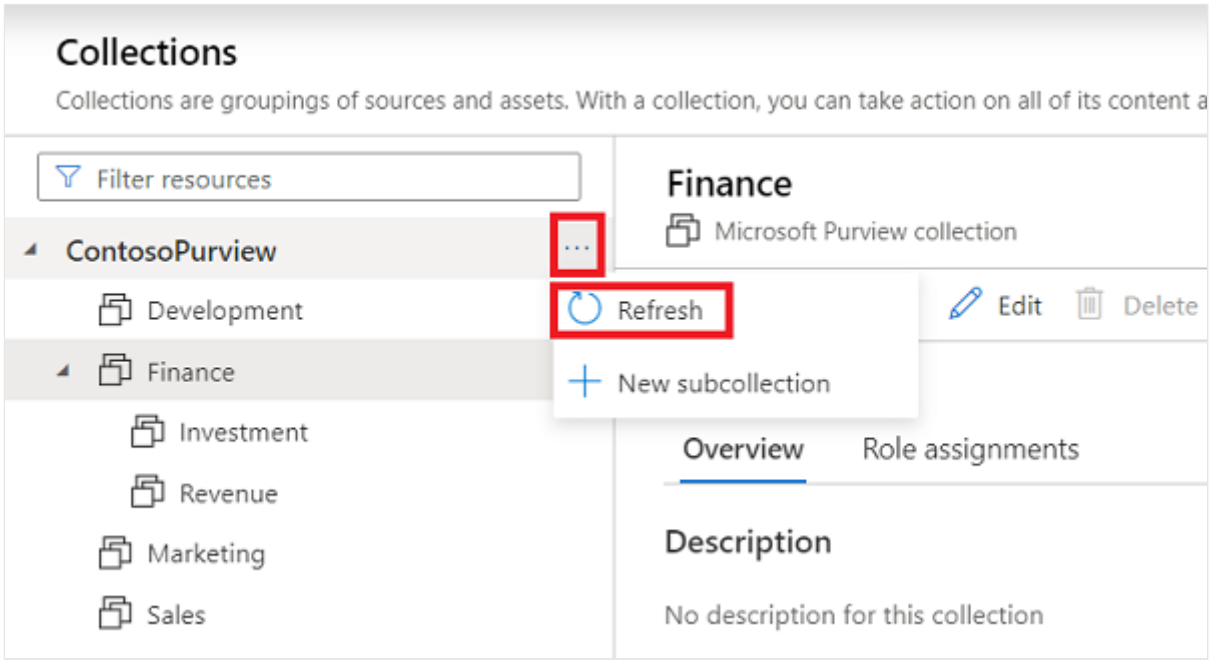
1. Select the triangle icon beside the collection's name to expand or collapse the collection hierarchy. Select the collection names to navigate.



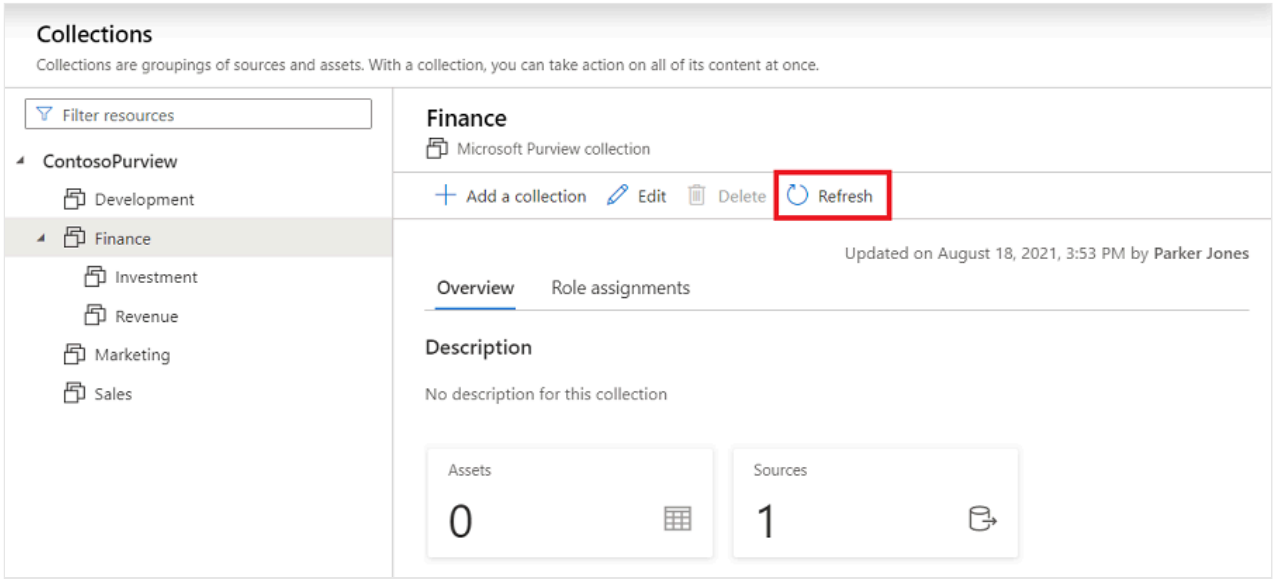
2. Type in the filter box at the top of the list to filter collections.



3. Select **Refresh** in Root collection's contextual menu to reload the collection list.



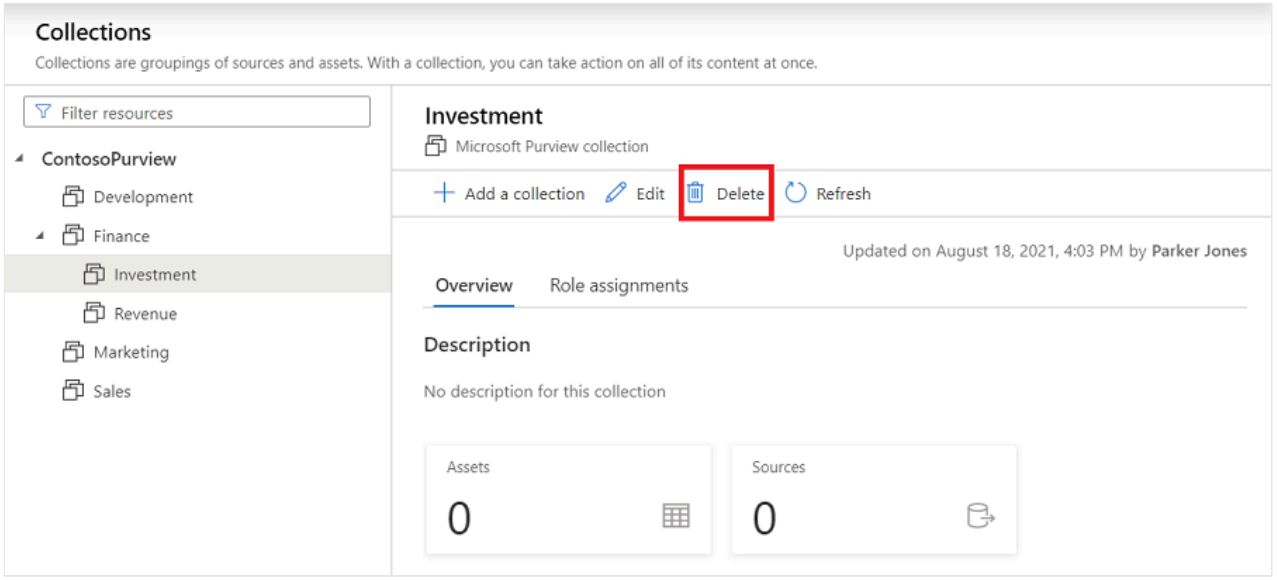
4. Select **Refresh** in collection detail page to reload the single collection.



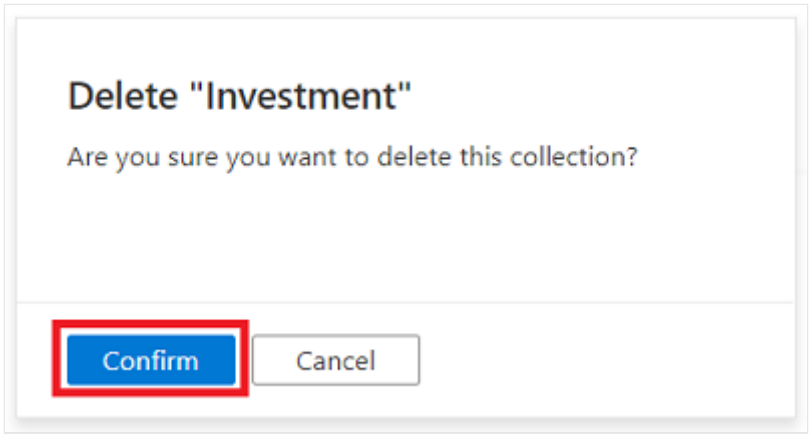
Delete a collection

You'll need to be a collection admin in order to delete a collection. If you aren't sure, follow the guide above to check permissions. Collection can be deleted only if no child collections, assets, data sources or scans are associated with it.

1. Select **Delete** from the collection detail page.



2. Select **Confirm** when prompted, **Are you sure you want to delete this collection?**



3. Verify deletion of the collection from your Data Map.

Move registered sources between collections

You can move registered sources from one collection to another you have access to. For steps, see Manage data sources.

Add roles and restrict access through collections

Since permissions are managed through collections in Data Map, it's important to understand the roles and what permissions they'll give your users. A user granted permissions on a collection will have access to sources and assets associated with that collection, and inherit permissions to subcollections. Inheritance can be restricted, but is allowed by default.

The following guide will discuss the roles, how to manage them, and permissions inheritance.

Roles

- **Domain admin** (domain level only) - Can assign permissions within a domain and manage its resources.
- **Collection administrator** - a role for users that will need to assign roles to other users in the Microsoft Purview governance portal or manage collections. Collection admins can add users to roles on collections where they're admins. They can also edit collections, their details, and add subcollections. A collection administrator on the root collection also automatically has permission to the Microsoft Purview governance portal. If your **root collection administrator** ever needs to be changed, you can follow the steps in the section below.
- **Data curators** - a role that provides access to the Microsoft Purview Unified Catalog to manage assets, configure custom classifications, create and manage glossary terms, and view data estate insights. Data curators can create, read, modify, move, and delete assets. They can also apply annotations to assets.
- **Data readers** - a role that provides read-only access to data assets, classifications, classification rules, collections and glossary terms.
- **Data source administrator** - a role that allows a user to manage data sources and scans. If a user is granted only to **Data source admin** role on a given data source, they can run new scans using an existing scan rule. To create new scan rules, the user must be also granted as either **Data reader** or **Data curator** roles.
- **Insights reader** - a role that provides read-only access to insights reports for collections where the insights reader also has at least the **Data reader** role. For more information, see insights permissions.
- **Policy author** - a role that allows a user to view, update, and delete Microsoft Purview policies through the Data policy app within Microsoft Purview.
- **Workflow administrator** - a role that allows a user to access the workflow authoring page in the Microsoft Purview governance portal, and publish workflows on collections where they have access permissions. Workflow administrator only has access to authoring, and so will need at least Data reader permission on a collection to be able to access the Purview governance portal.

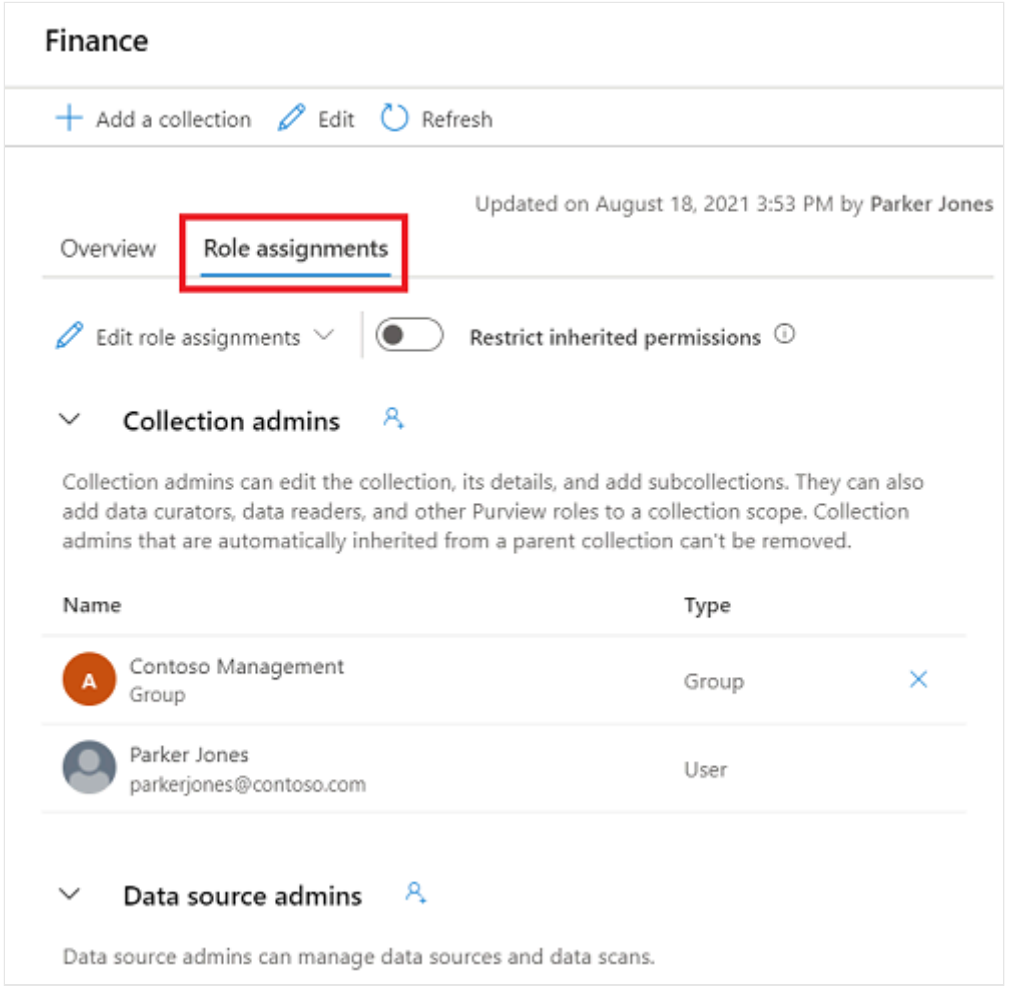
> ⓘ **Note**

At this time, Microsoft Purview policy author role is not sufficient to create policies. The Microsoft Purview data source admin role is also required.

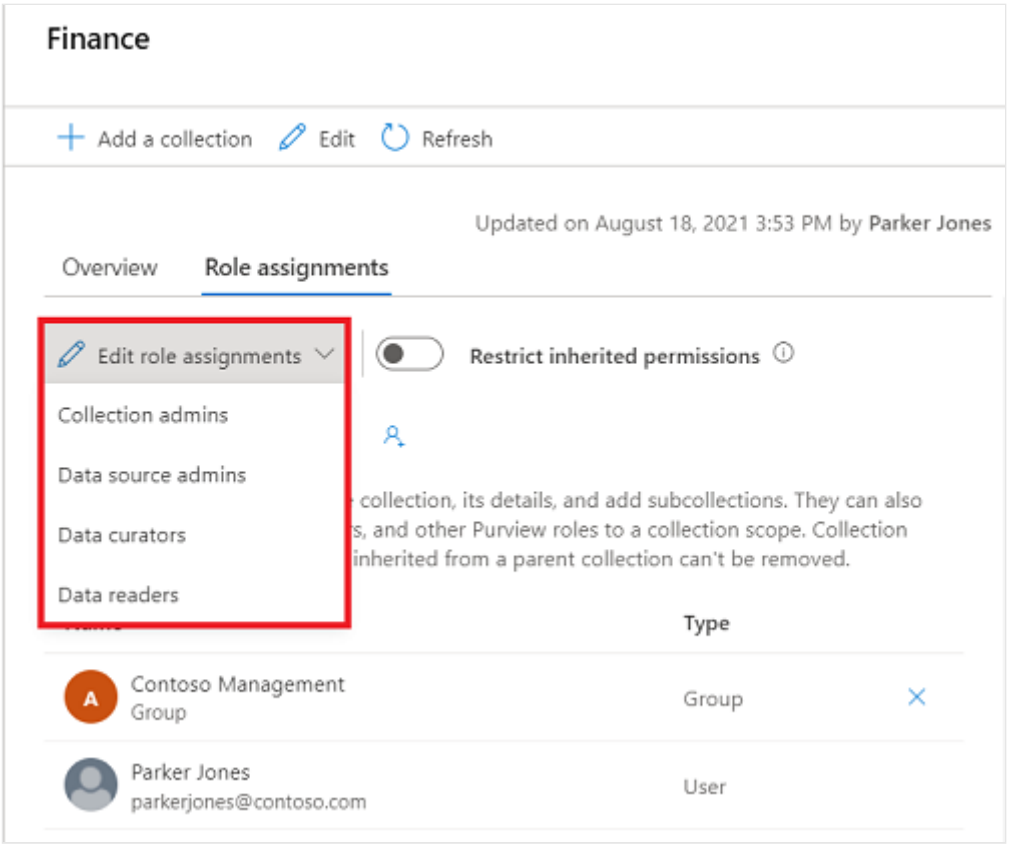
> **Important**
>
> The user that created the account is automatically assigned domain admin on the default domain and collection admin on the root collection.

Add role assignments

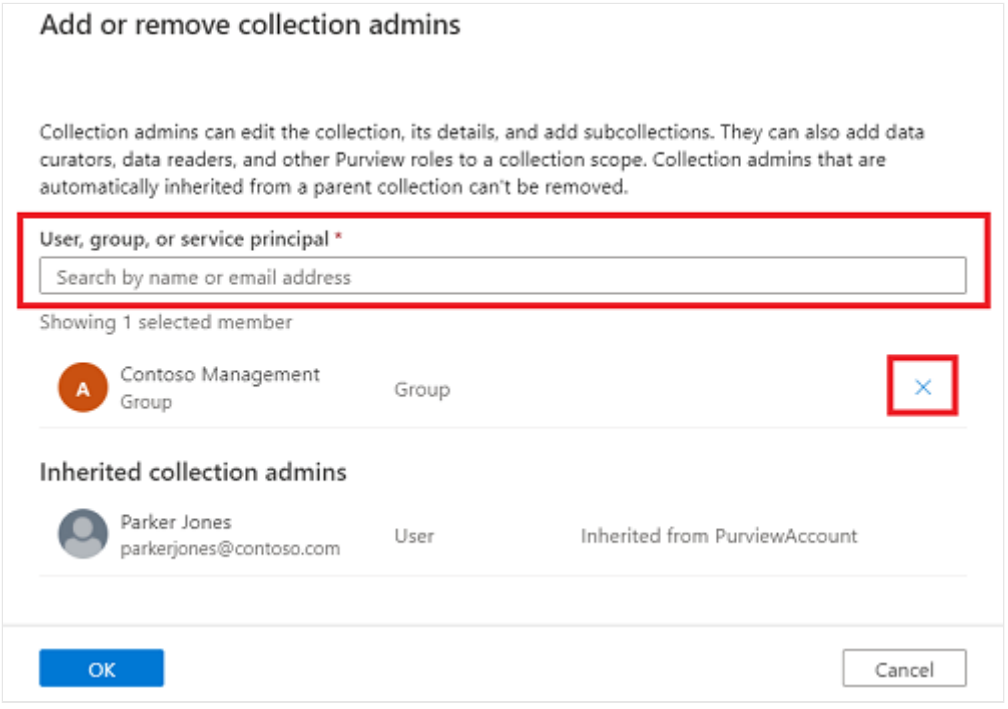
1. Open the Microsoft Purview Data Map.
2. Select the domain or collection where you want to add your role assignment.
3. Select the **Role assignments** tab to see all the roles in a collection or a domain. Only a collection admin or domain admin can manage role assignments.



4. Select **Edit role assignments** or the person icon to edit each role member.



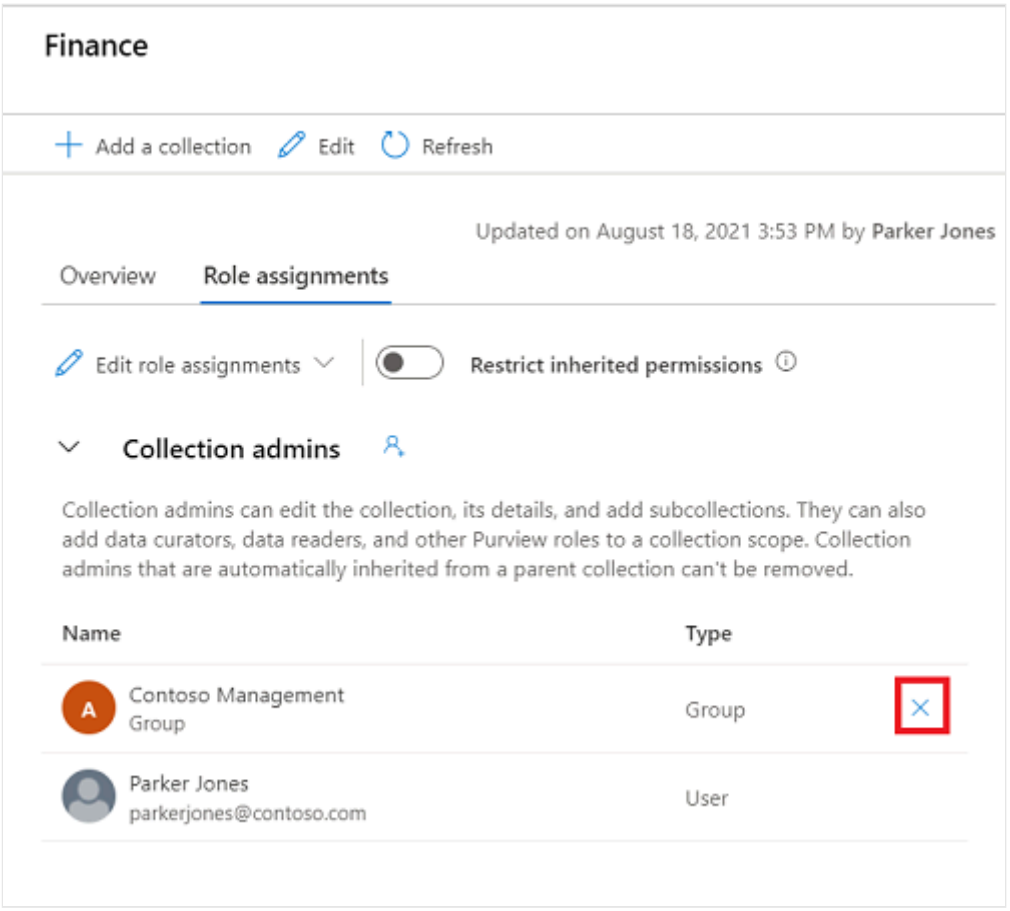
5. Type in the textbox to search for users you want to add to the role member. Select **X** to remove members you don't want to add.



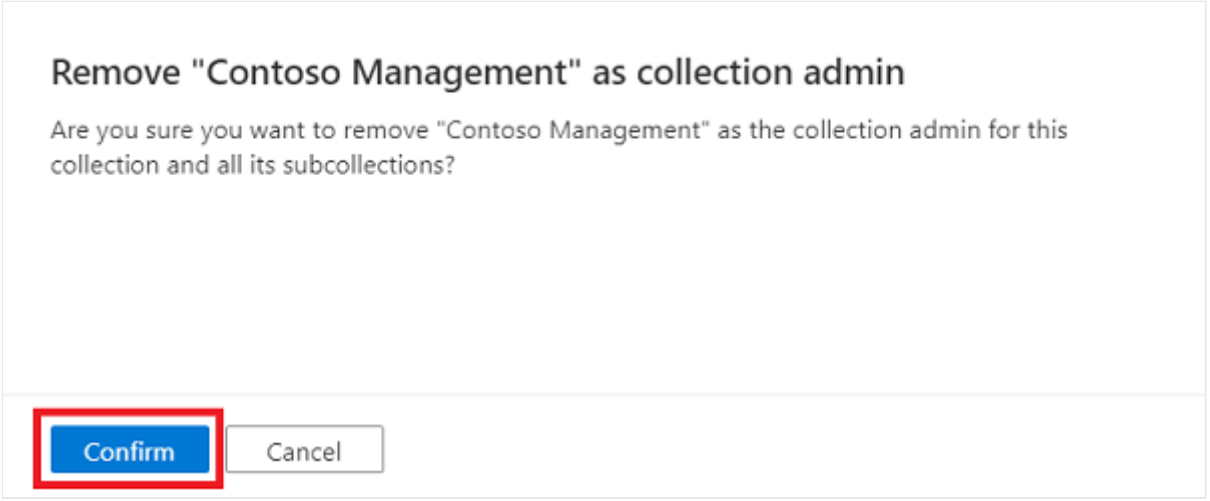
6. Select **OK** to save your changes, and you'll see the new users reflected in the role assignments list.

Remove role assignments

1. Select **X** button next to a user's name to remove a role assignment.



2. Select **Confirm** if you're sure to remove the user.



Restrict inheritance

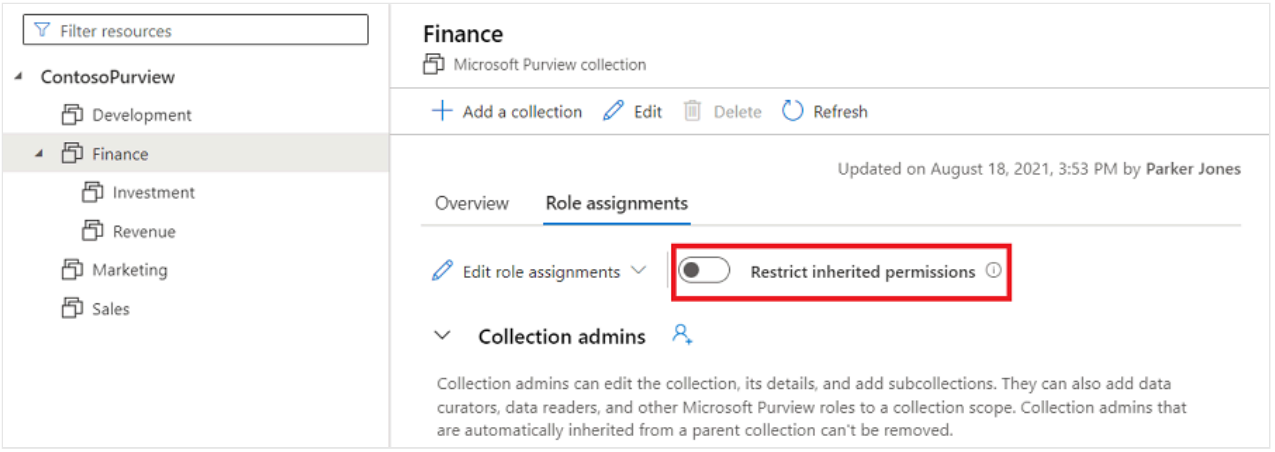
Collection permissions are inherited automatically from the parent collection. You can restrict inheritance from a parent collection at any time, using the restrict inherited permissions option.

> **Note**
>
> Currently permissions from the default domain cannot be restricted. Any permissions assigned at the default domain will be inherited by the domain's direct subcollections.

Once you restrict inheritance, you'll need to add users directly to the restricted collection to grant them access.

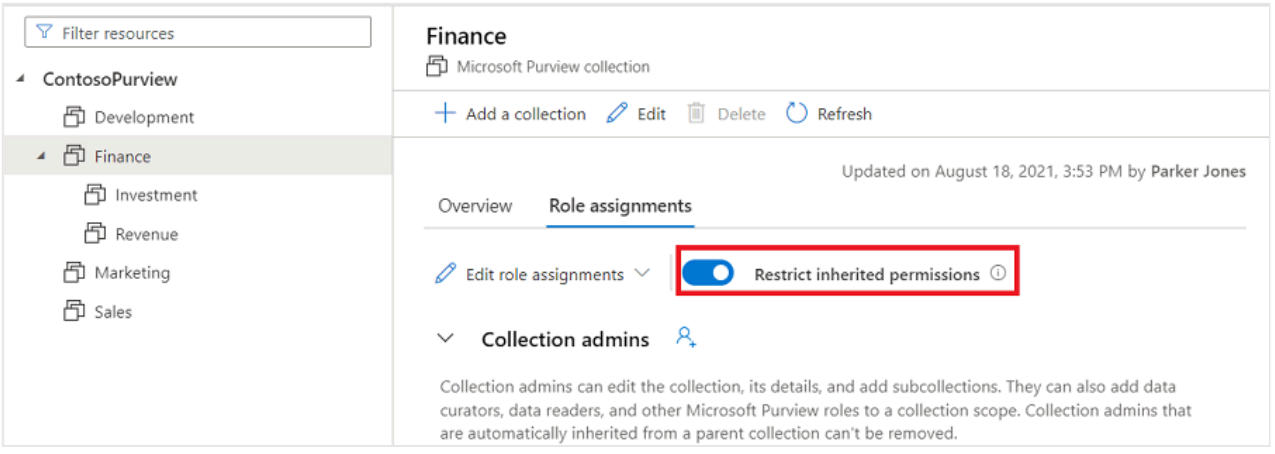
1. Navigate to the collection where you want to restrict inheritance and select the **Role assignments** tab.

2. Select **Restrict inherited permissions** and select **Restrict access** in the popup dialog to remove inherited permissions from this collection and any subcollections. Collection admin permissions won't be affected.



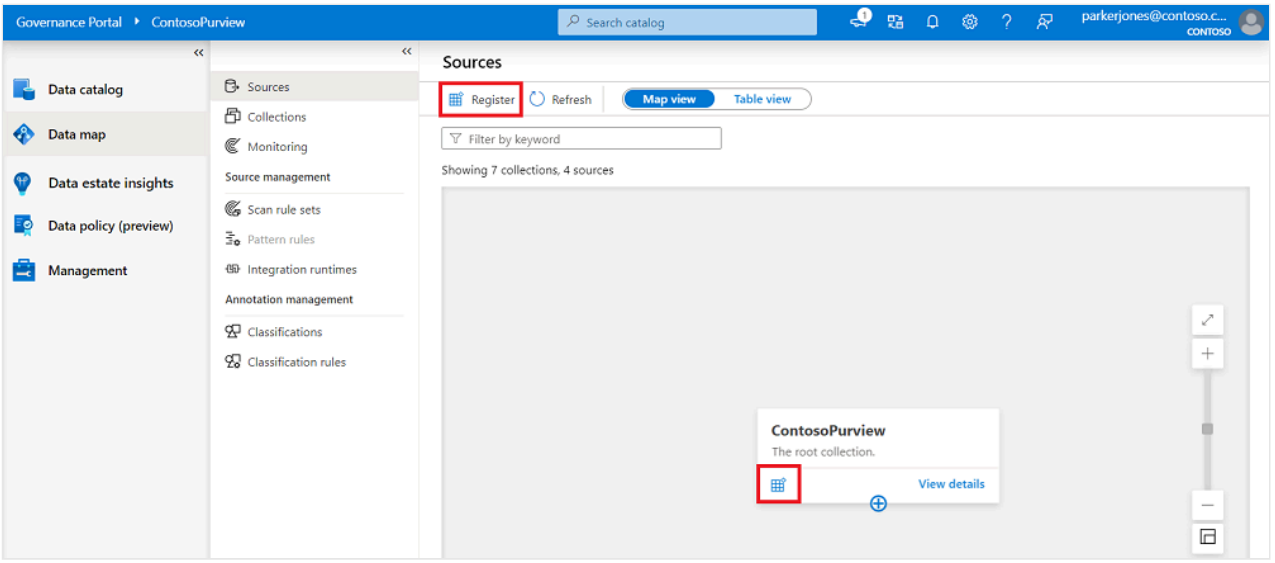
3. After restriction, inherited members are removed from the roles expect for collection admin.

4. Select the **Restrict inherited permissions** toggle button again to revert.

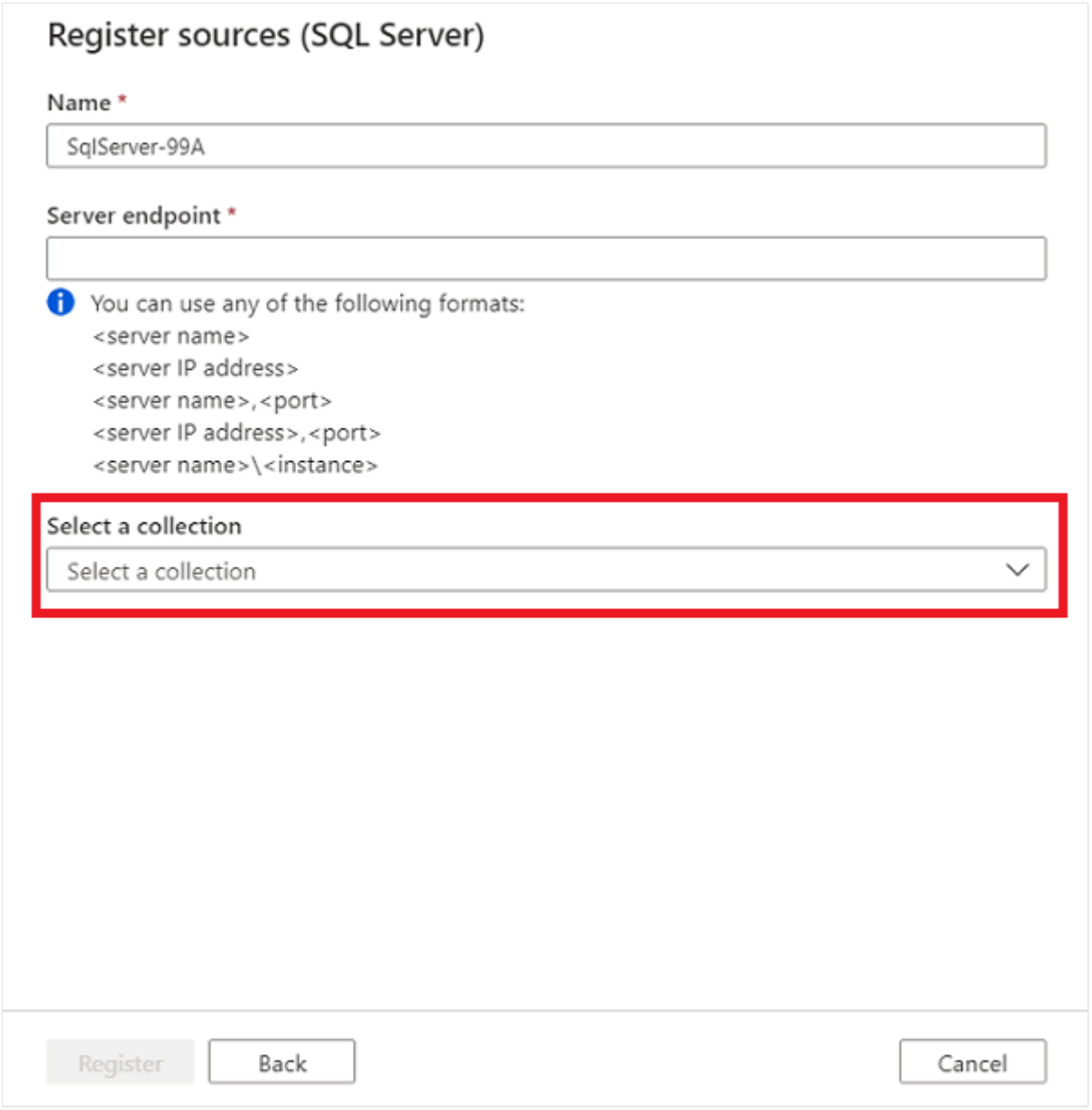


Register source to a collection

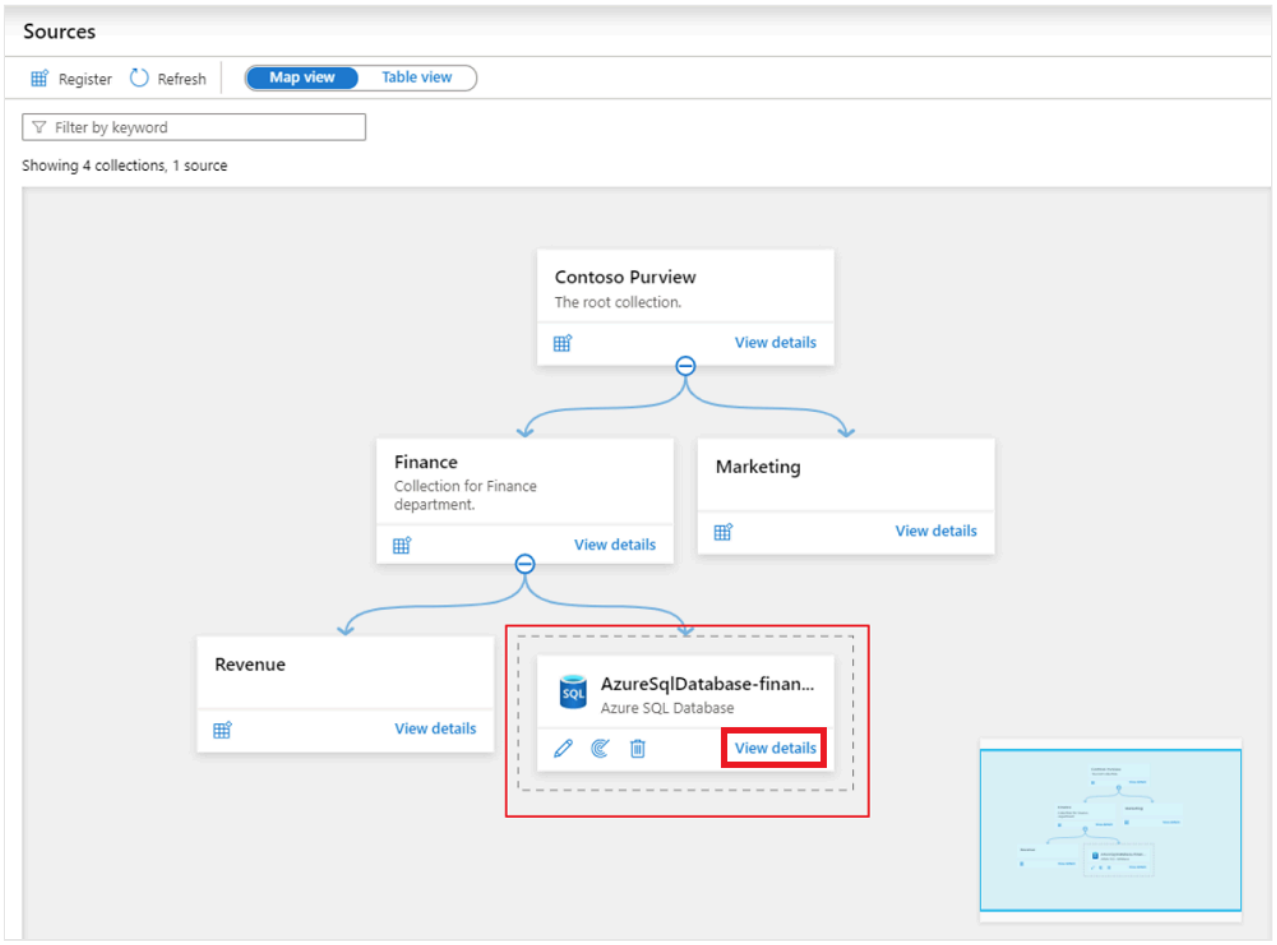
1. Select **Register** or register icon on collection node to register a data source. Only a data source admin can register sources.



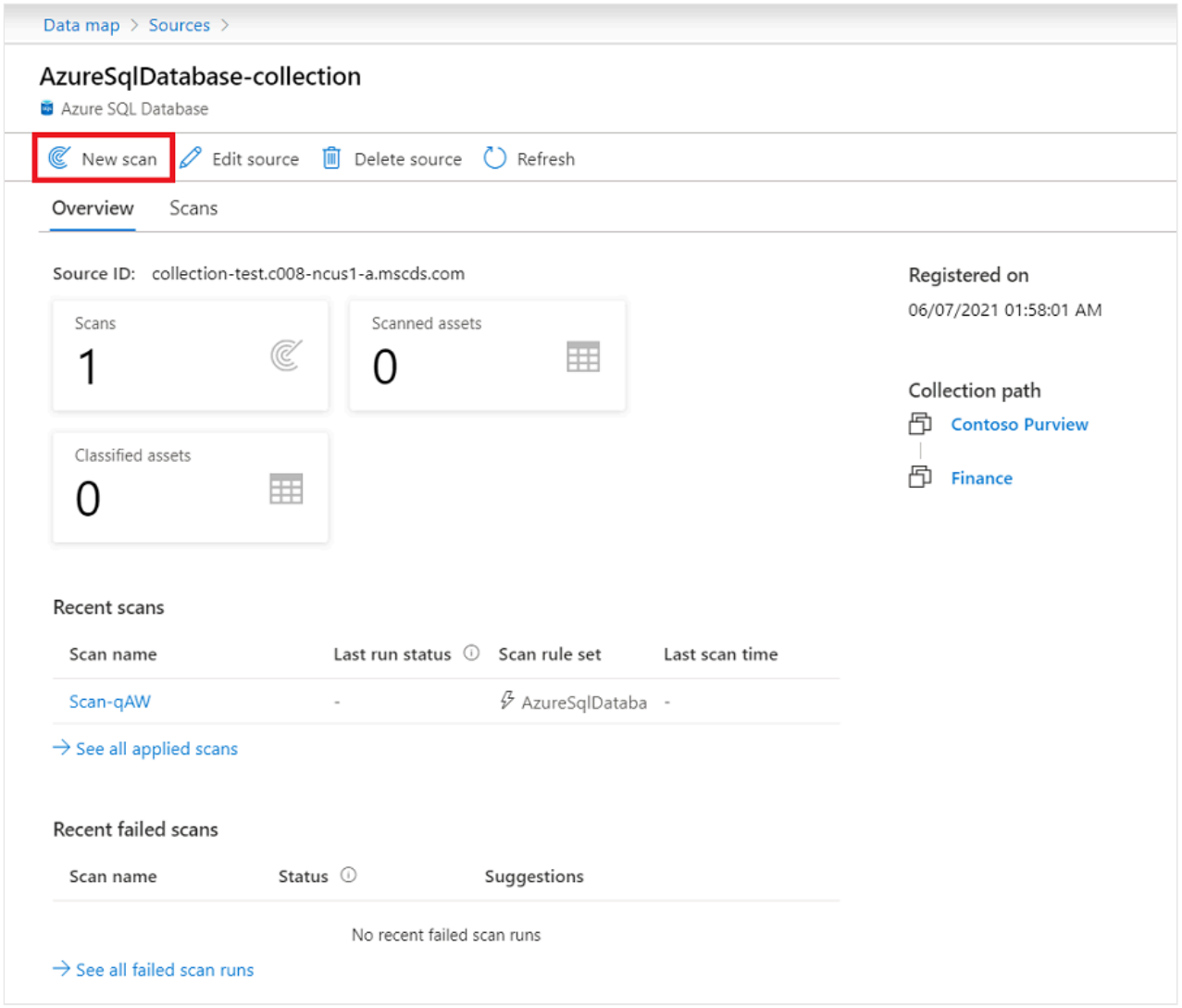
2. Fill in the data source name, and other source information. It lists all the collections where you have scan permission on the bottom of the form. You can select one collection. All assets under this source will belong to the collection you select.



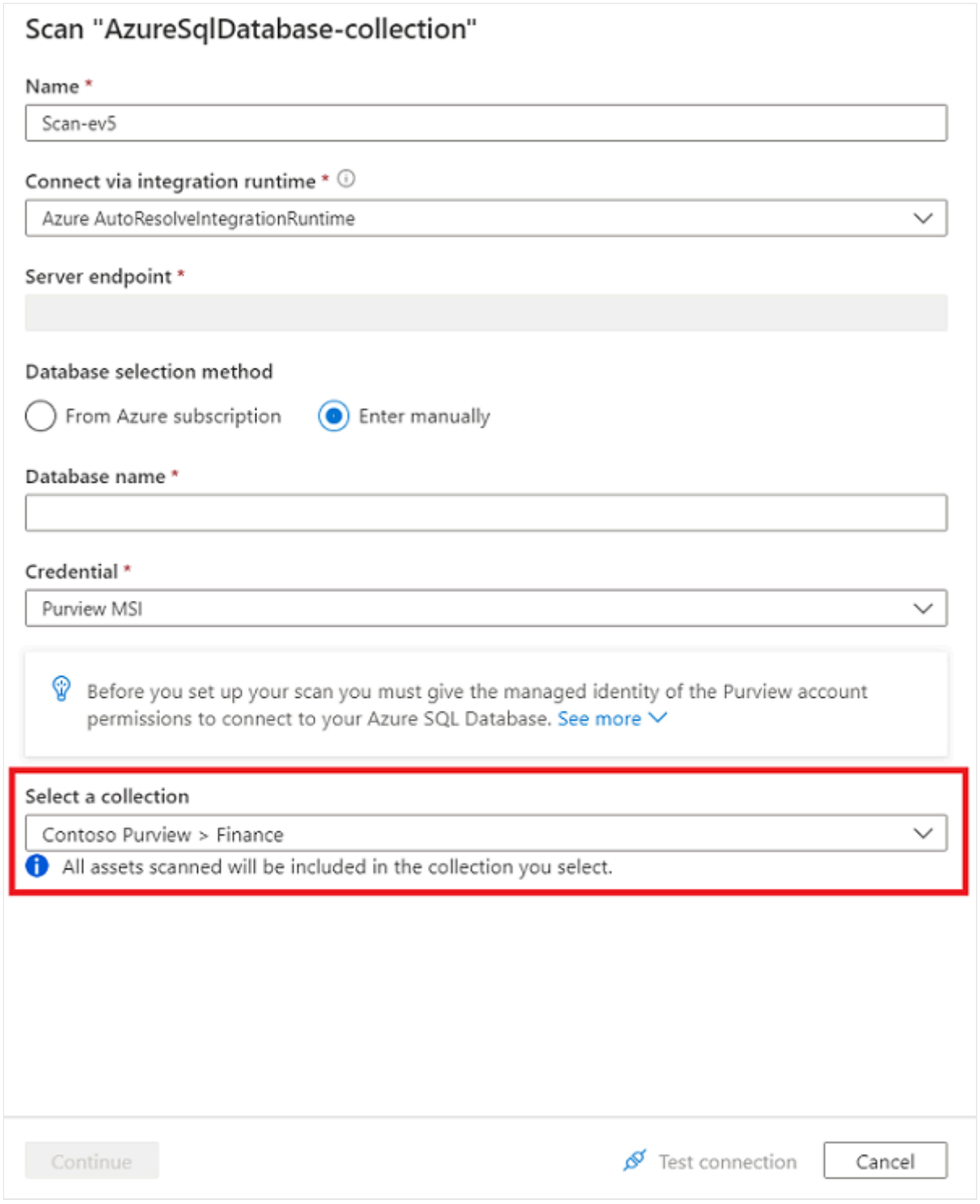
3. The created data source will be put under the selected collection. Select **View details** to see the data source.



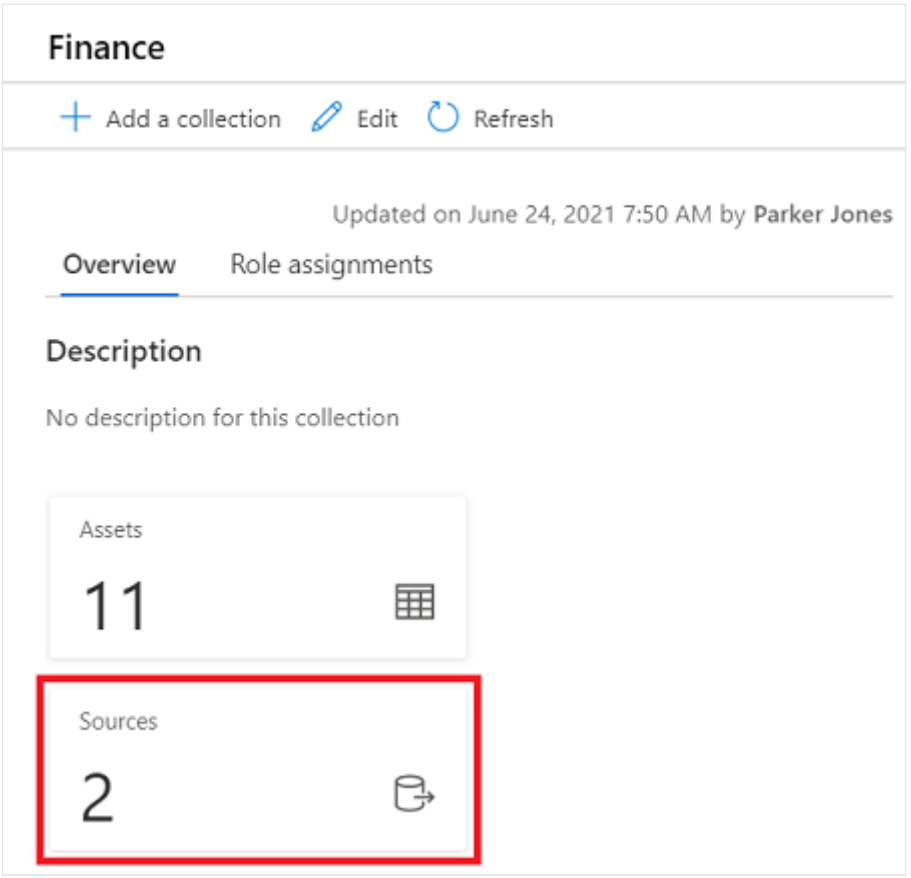
4. Select **New scan** to create scan under the data source.



5. Similarly, at the bottom of the form, you can select a collection, and all assets scanned will be included in the collection. The collections listed here are restricted to subcollections of the data source collection.



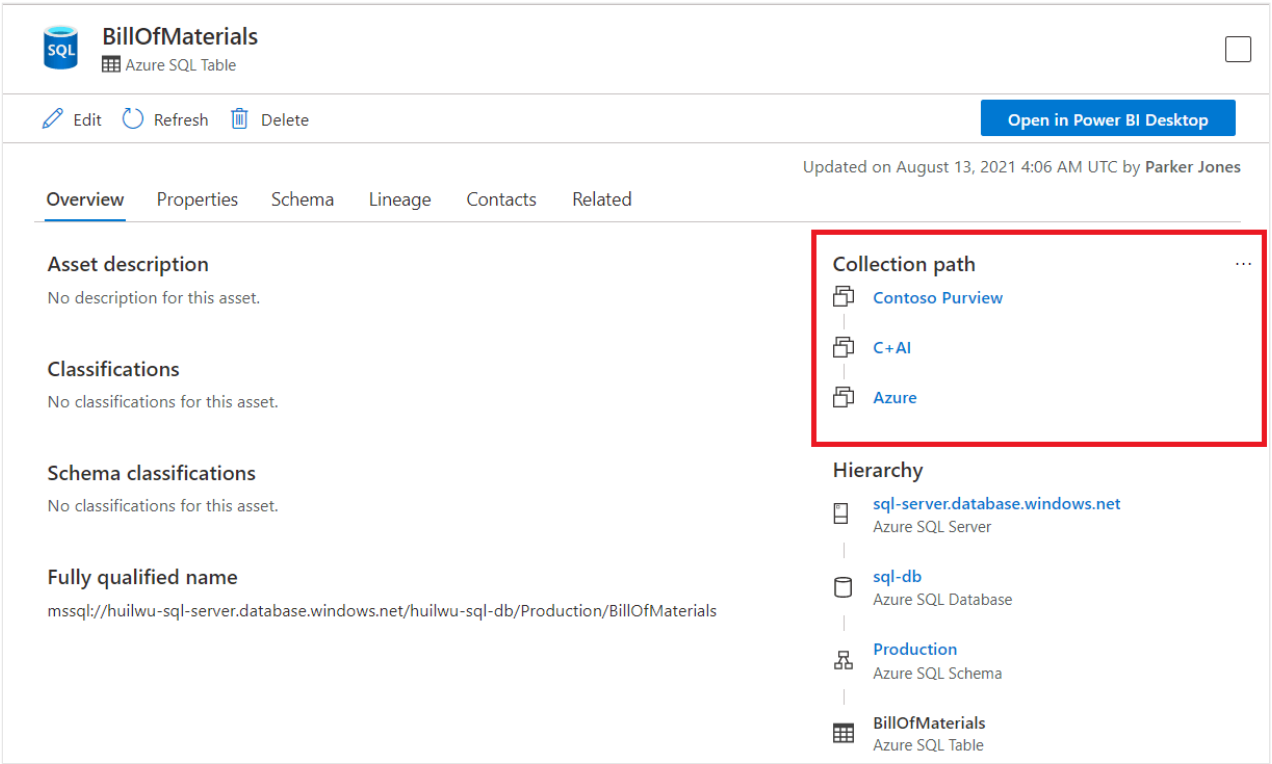
6. Back in the collection window, you'll see the data sources linked to the collection on the sources card.



Add assets to collections

Assets and sources are also associated with collections. During a scan, if the scan was associated with a collection the assets will be automatically added to that collection, but can also be manually added to any subcollections.

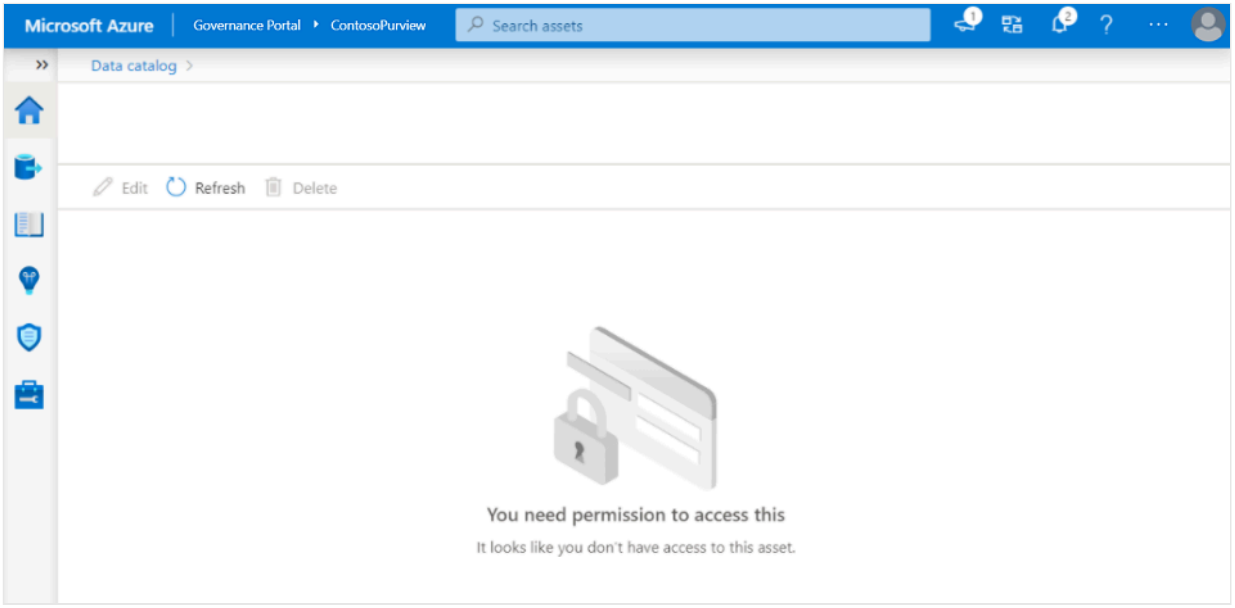
1. Check the collection information in asset details. You can find collection information in the **Collection path** section on right-top corner of the asset details page.



2. Permissions in asset details page:

   a. Check the collection-based permission model by following the add roles and restricting access on collections guide above.

   b. If you don't have read permission on a collection, the assets under that collection won't be listed in search results. If you get the direct URL of one asset and open it, you'll see the no access page. Contact your collection admin to grant you the access. You can select the **Refresh** button to check the permission again.

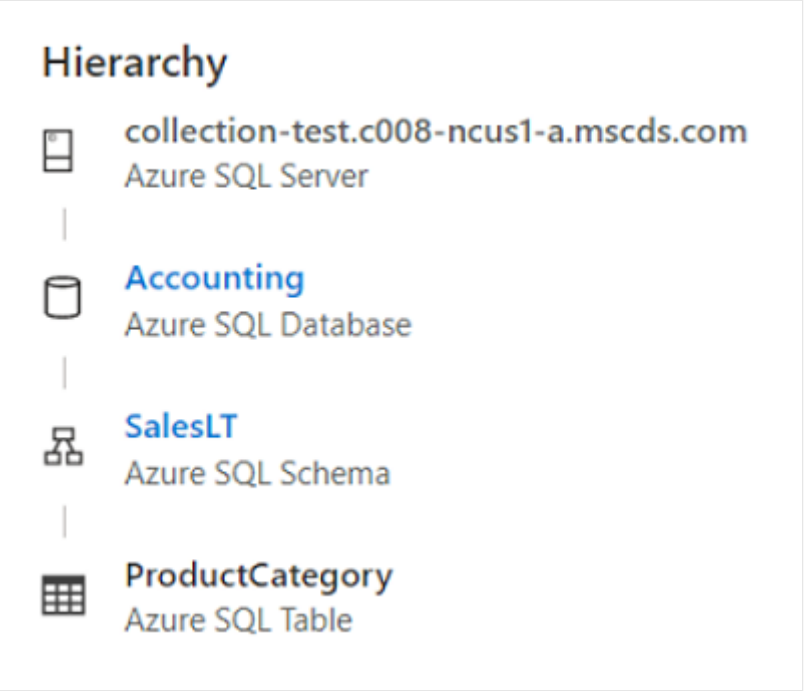


   c. If you have the read permission to one collection but don't have the write permission, you can browse the asset details page, but the following operations are disabled:

      - Edit the asset. The **Edit** button will be disabled.

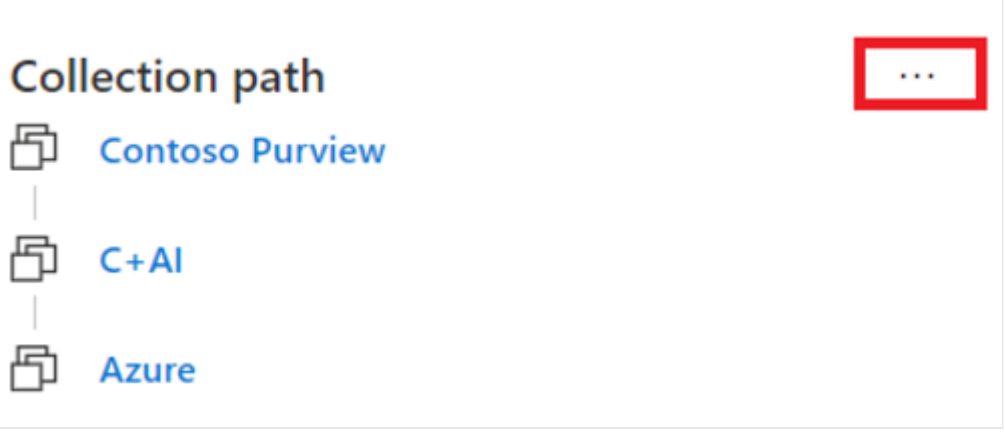
- Delete the asset. The **Delete** button will be disabled.
- Move asset to another collection. The ellipsis button on the right-top corner of Collection path section will be hidden.

d. The assets in **Hierarchy** section are also affected by permissions. Assets without read permission will be grayed.



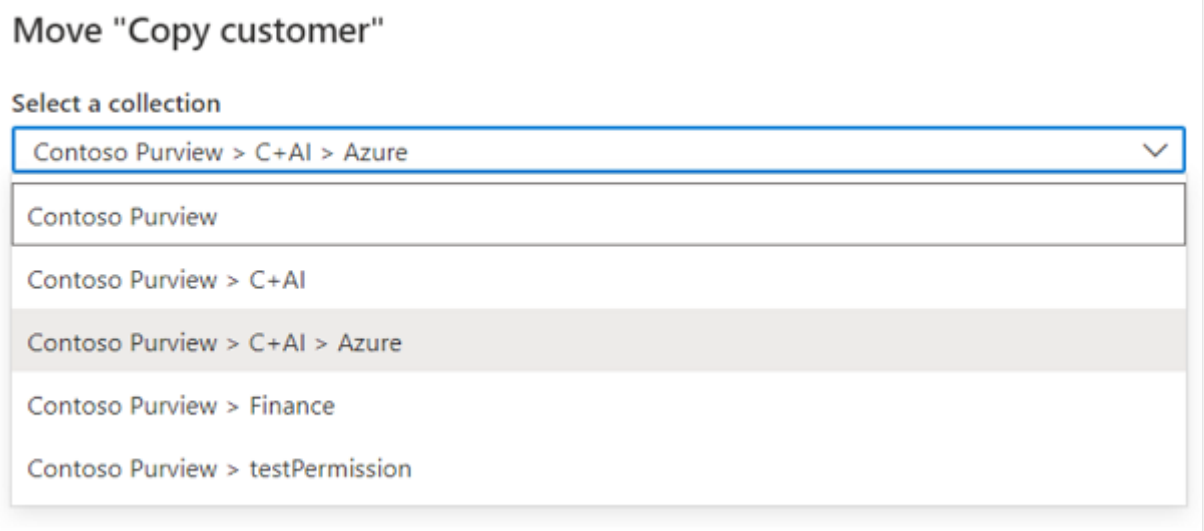
Move asset to another collection

1. Select the ellipsis button on the right-top corner of Collection path section.



2. Select the **Move to another collection** button.

3. In the right side panel, choose the target collection you want move to. You can only see the collections where you have write permissions. The asset can also only be added to the subcollections of the data source collection.

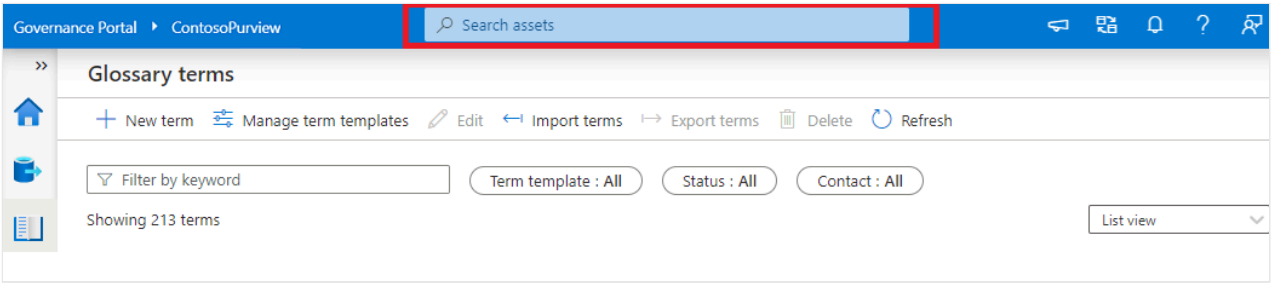


4. Select **Move** button on the bottom of the window to move the asset.

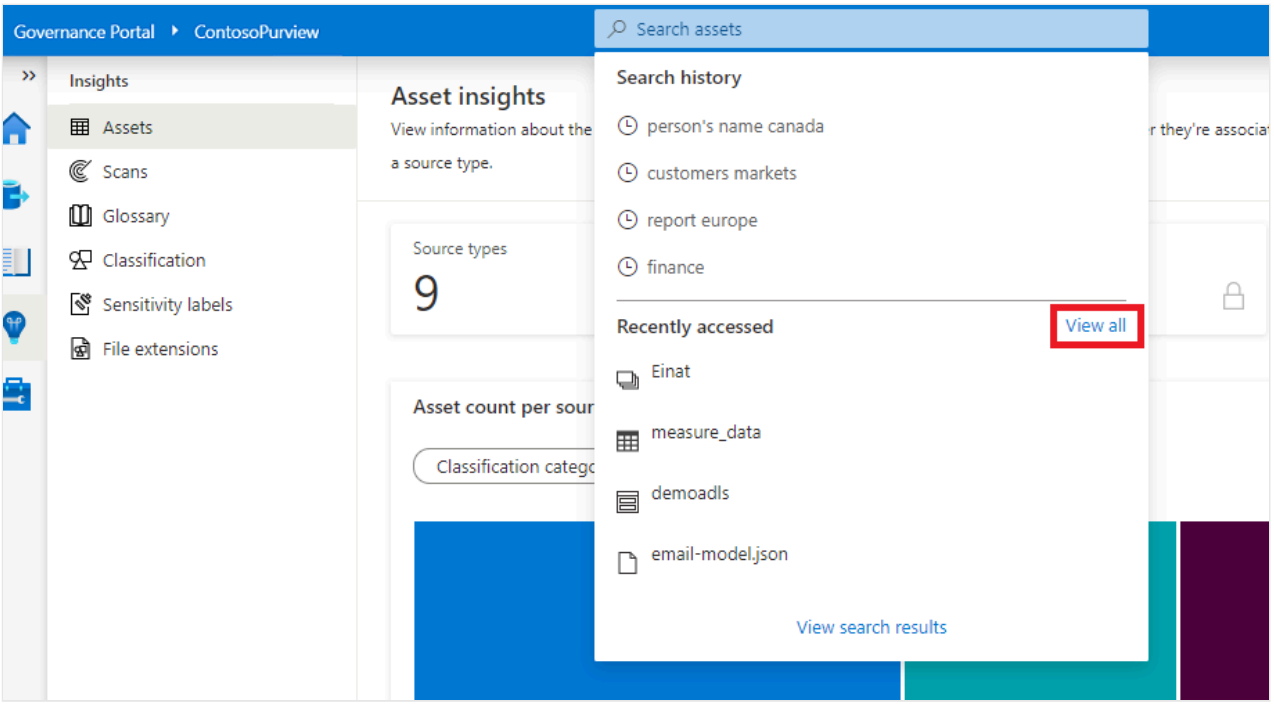
Search and browse by collections

Search by collection

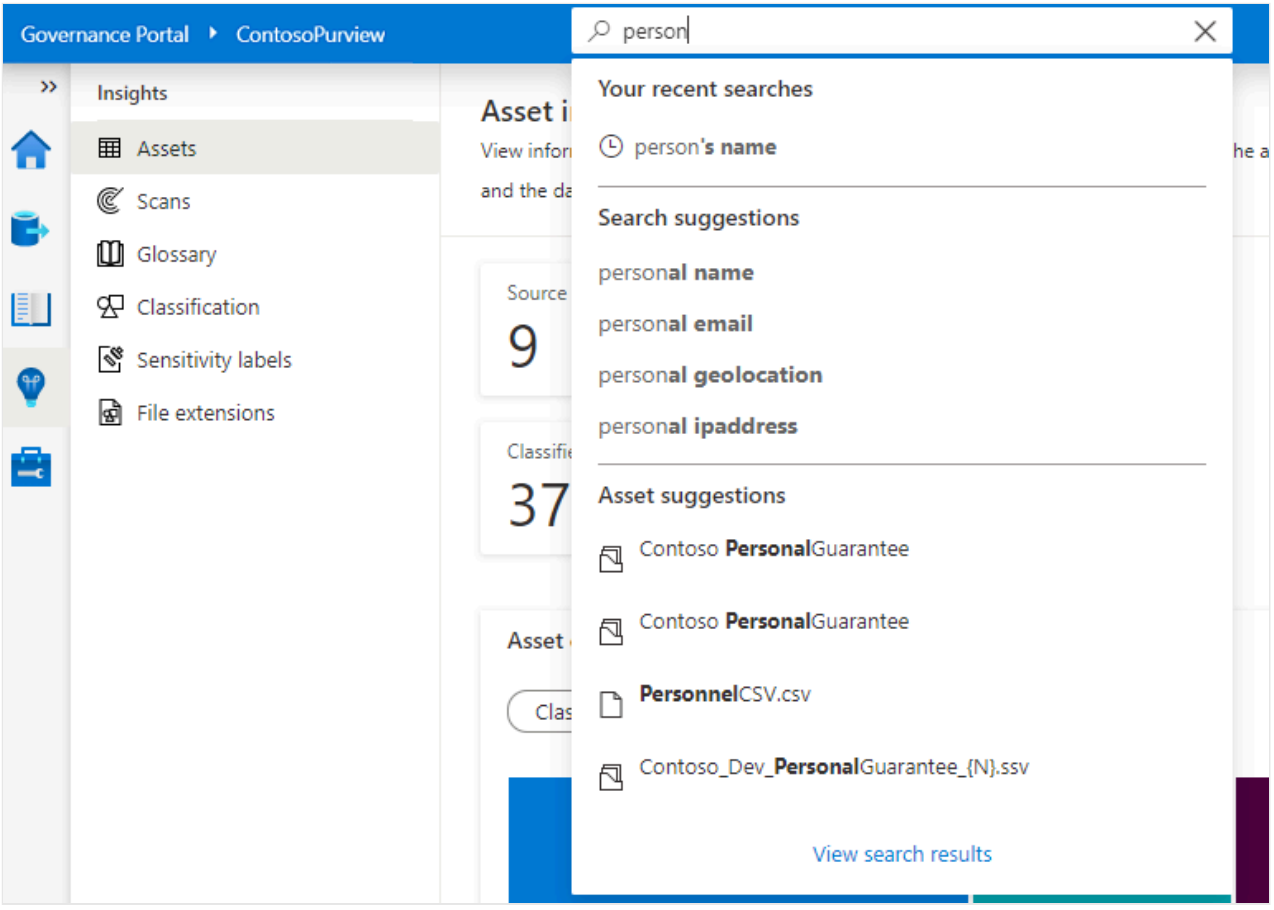
1. In the classic Microsoft Purview governance portal, the search bar is located at the top of the portal window.



2. When you select the search bar, you can see your recent search history and recently accessed assets. Select **View all** to see all of the recently viewed assets.

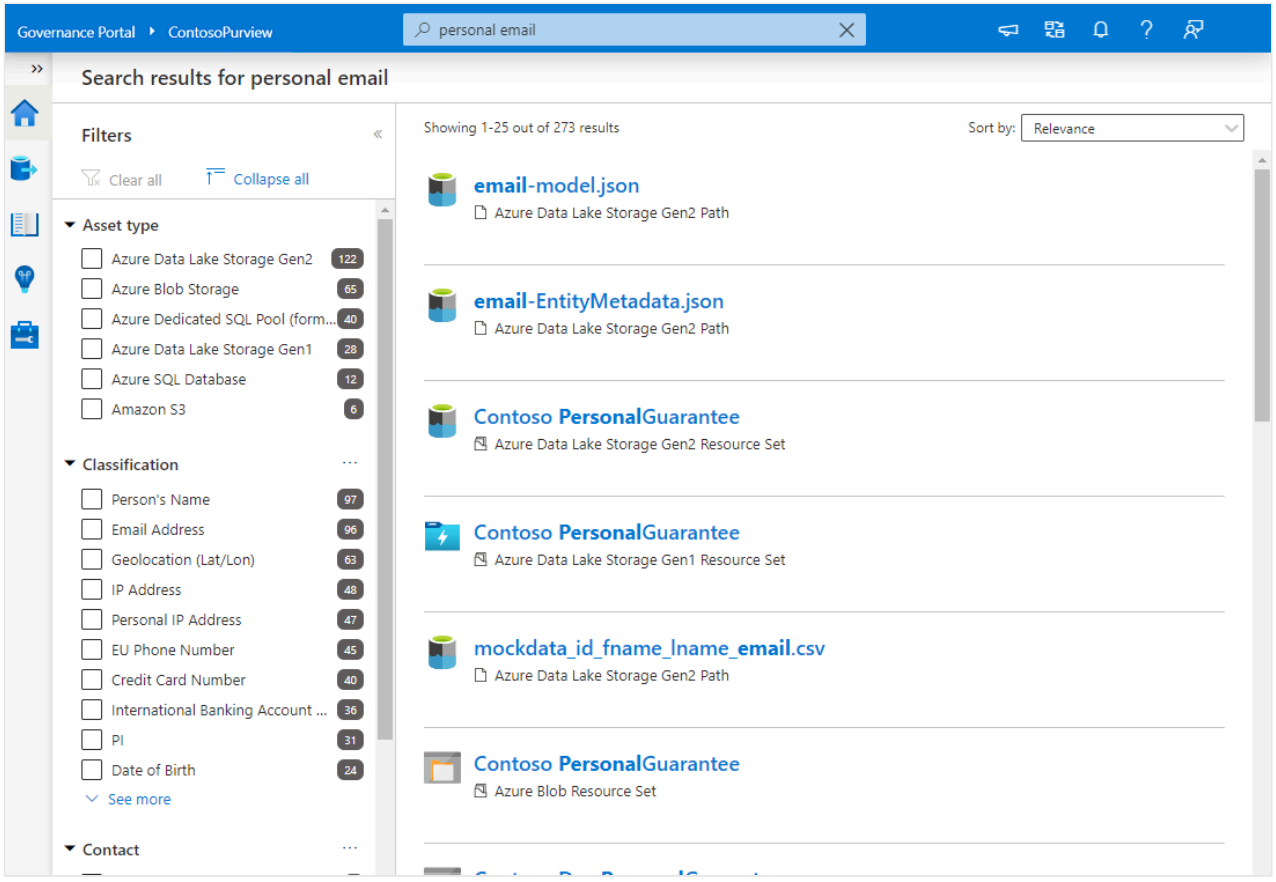


3. Enter in keywords that help identify your asset such as its name, data type, classifications, and glossary terms. As you enter in keywords relating to your desired asset, the classic Microsoft Purview governance portal displays suggestions on what to search and potential asset matches. To complete your search, select **View search results** or press **Enter**.

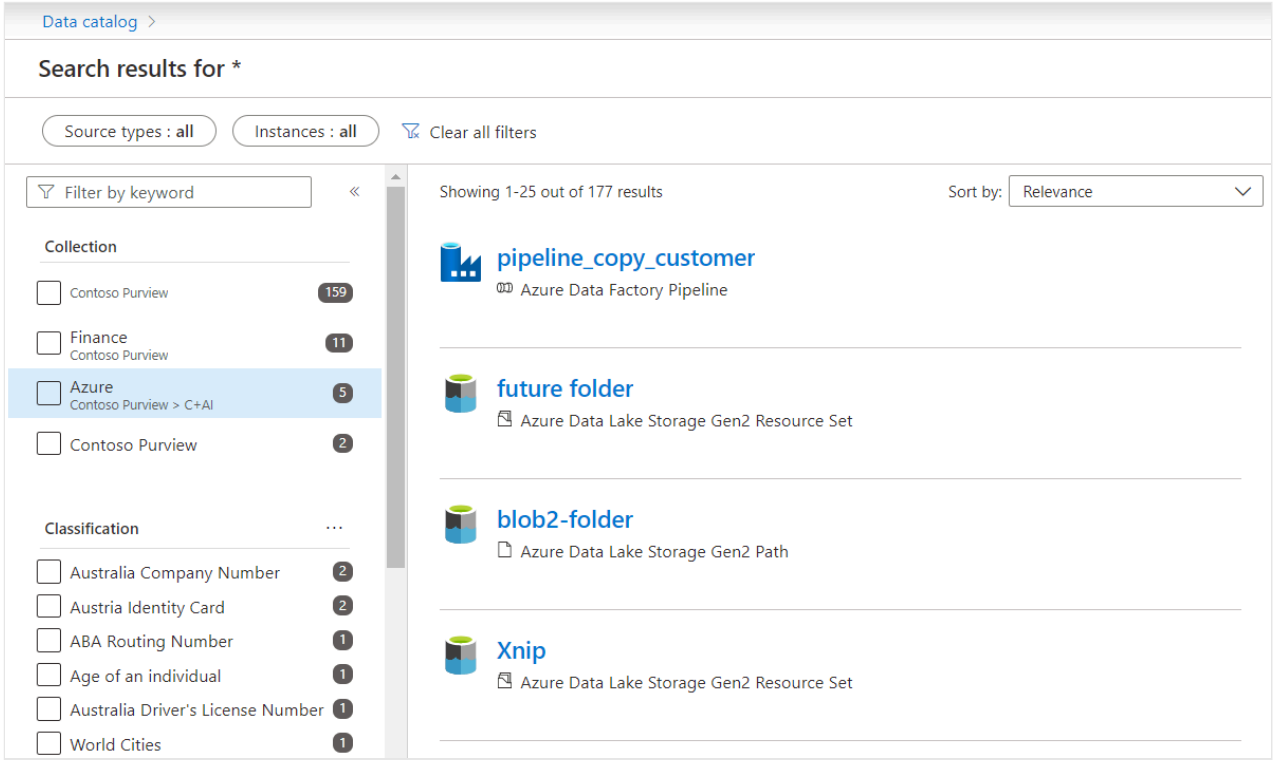


4. The search results page shows a list of assets that match the keywords provided in order of relevance. There are various factors that can affect the relevance score of an asset. You

can filter down the list more by selecting specific collections, data stores, classifications, contacts, labels, and glossary terms that apply to the asset you're looking for.

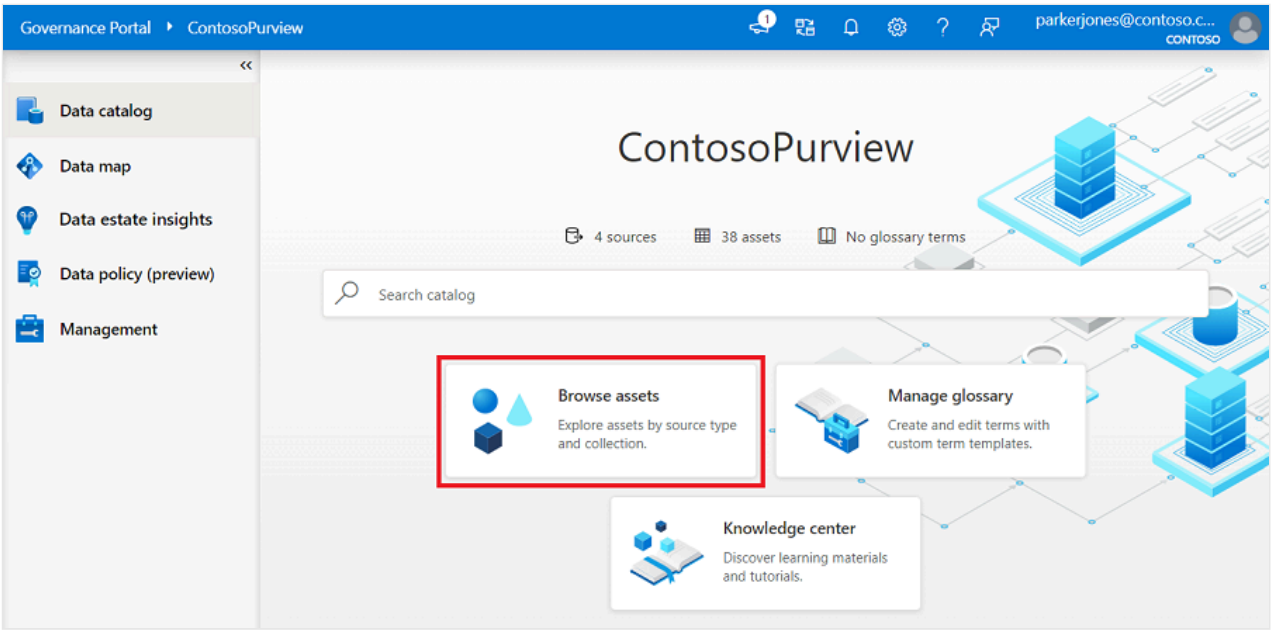


5. Select your desired asset to view the asset details page where you can view properties including schema, lineage, and asset owners.

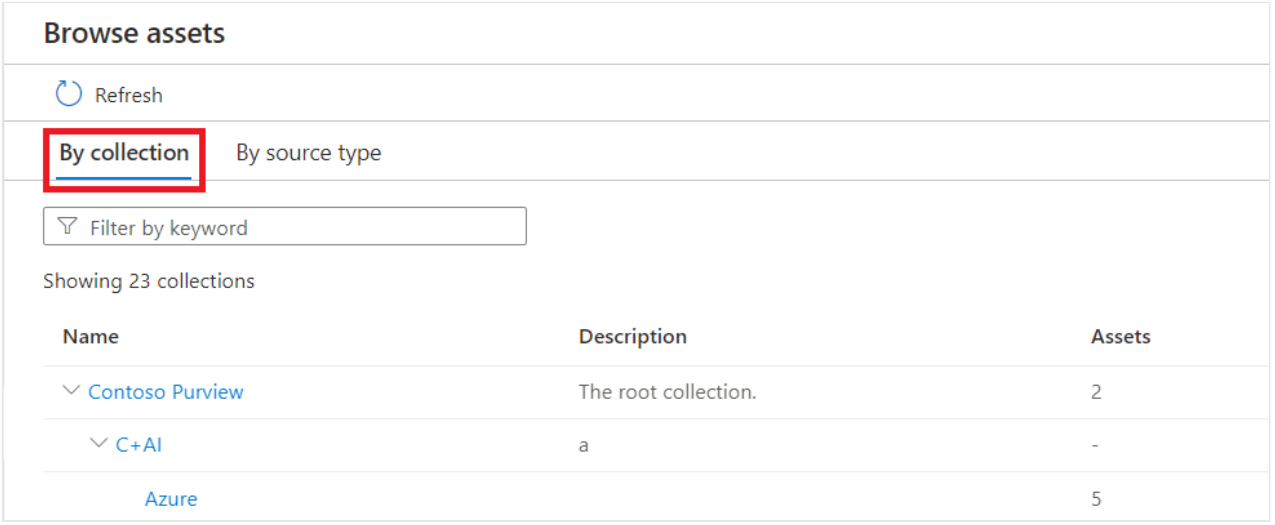


Browse by collection

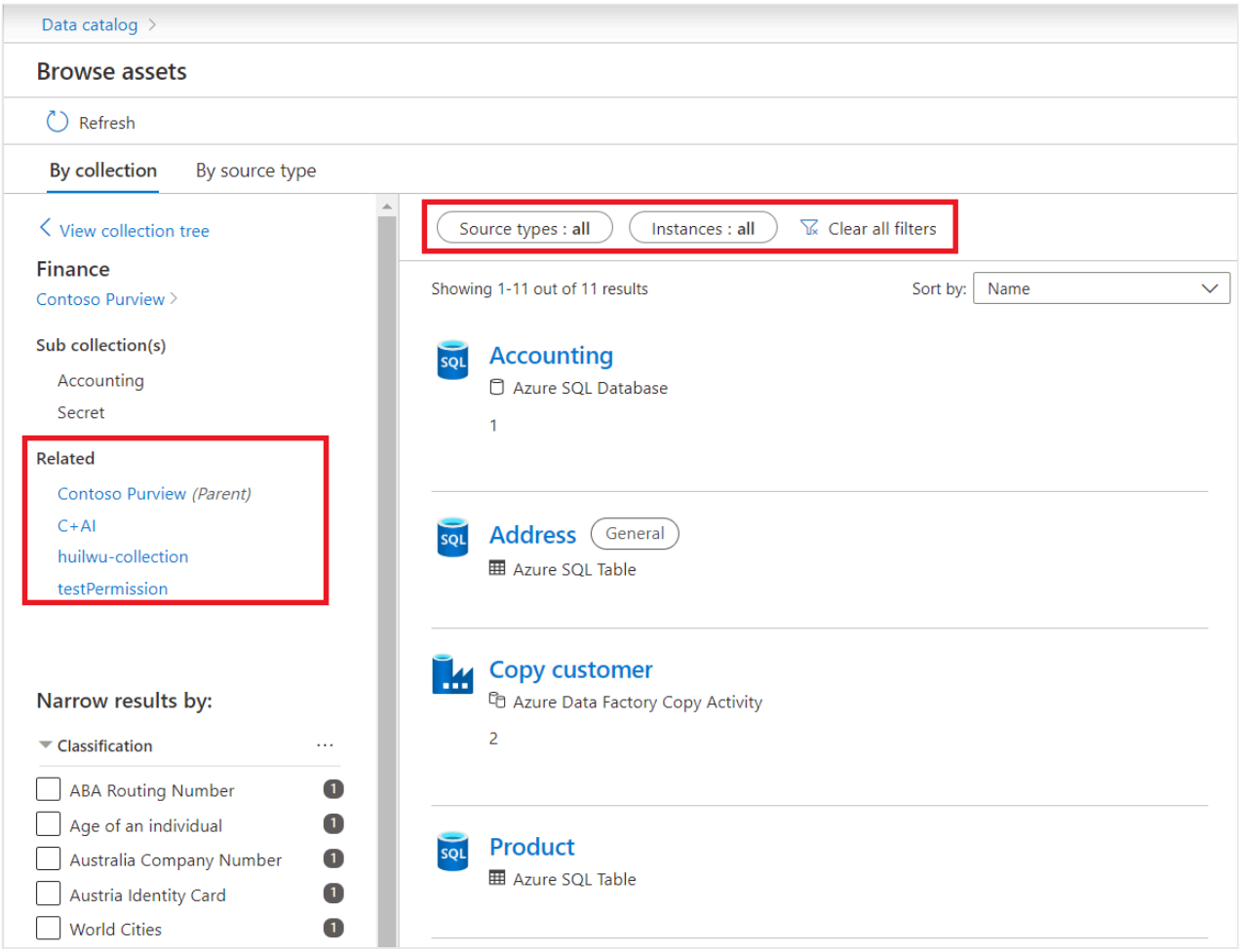
1. You can browse data assets, by selecting the **Browse assets** on the homepage.



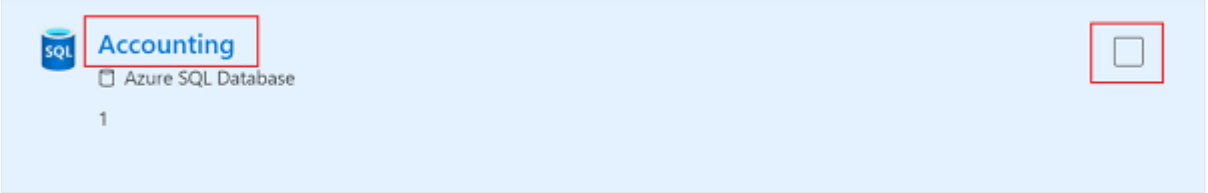
2. On the Browse asset page, select **By collection** pivot. Collections are listed with hierarchical table view. To further explore assets in each collection, select the corresponding collection name.



3. On the next page, the search results of the assets under selected collection will be shown. You can narrow the results by selecting the facet filters. Or you can see the assets under other collections by selecting the sub/related collection names.



4. To view the details of an asset, select the asset name in the search result. Or you can check the assets and bulk edit them.

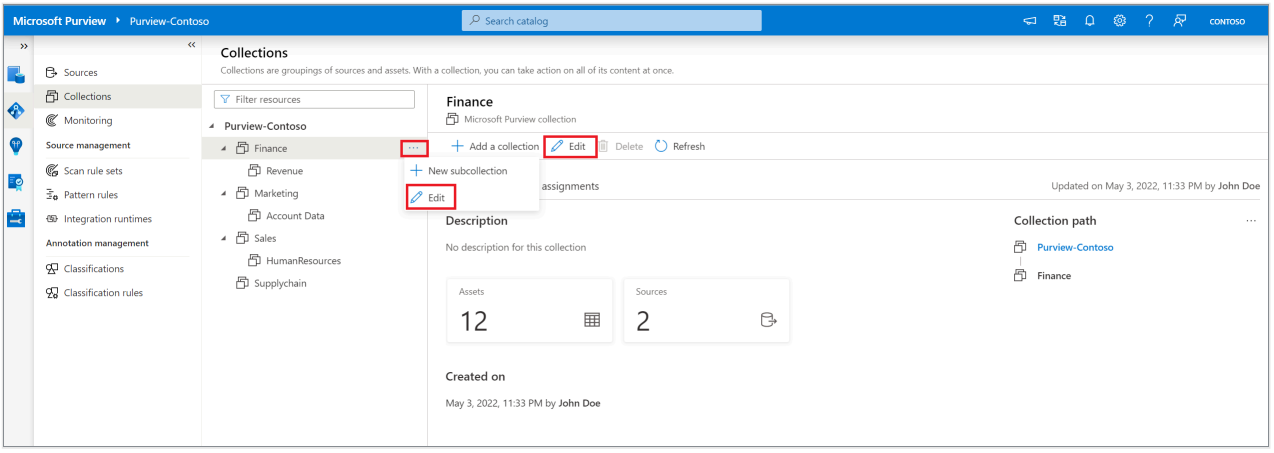


Rename collections (preview)

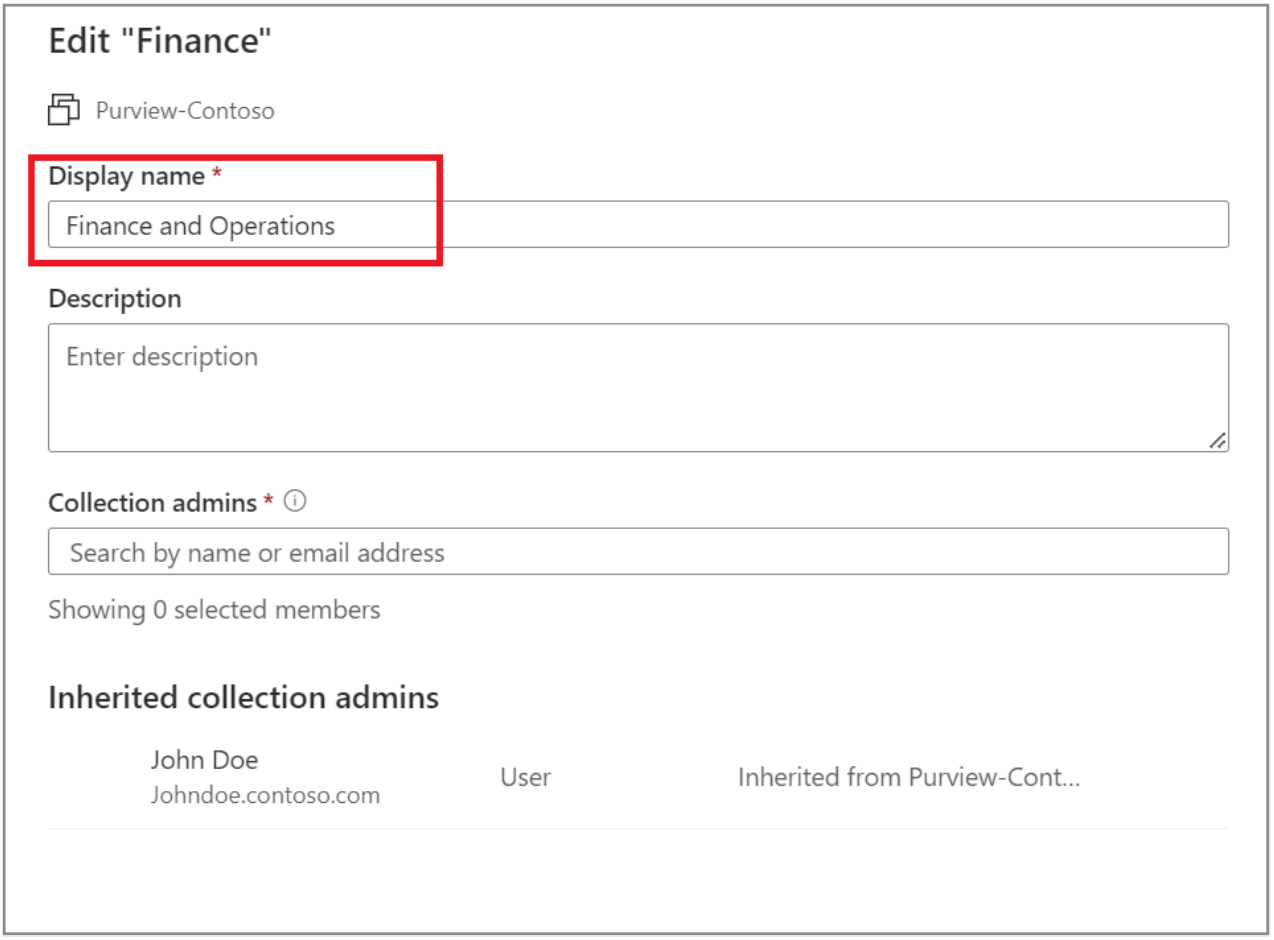
> **Important**
>
> This feature is currently in preview. The Supplemental Terms of Use for Microsoft Azure Previews include additional legal terms that apply to Azure features that are in beta, in preview, or otherwise not yet released into general availability.

1. In the classic Microsoft Purview governance portal, navigate to the collection that you want to rename.

2. Select **Edit** either from the collection detail page, or from the collection's dropdown menu.



3. Provide a new friendly name and select **Save** to save your change.



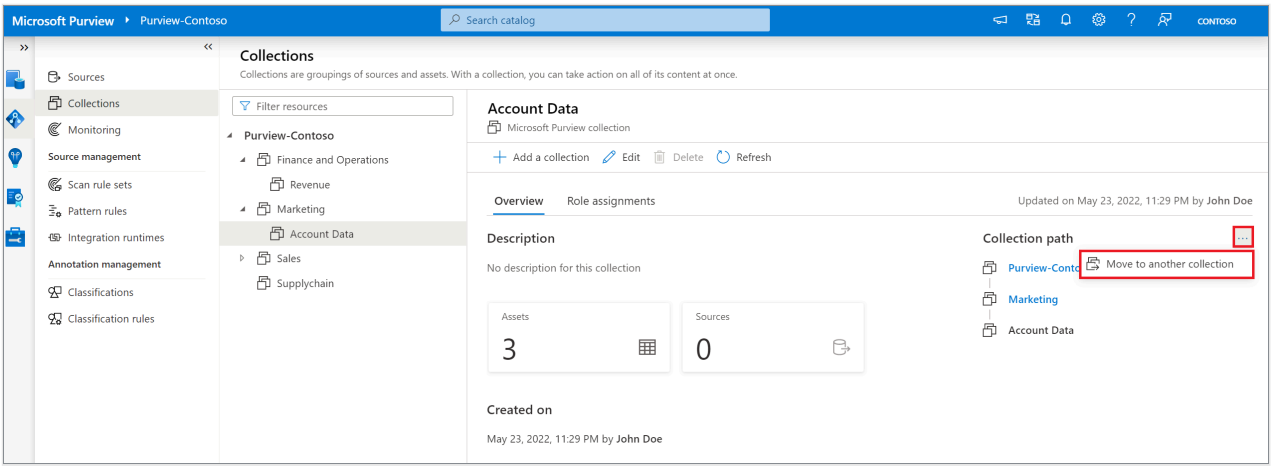
Move collections (preview)

> **Important**

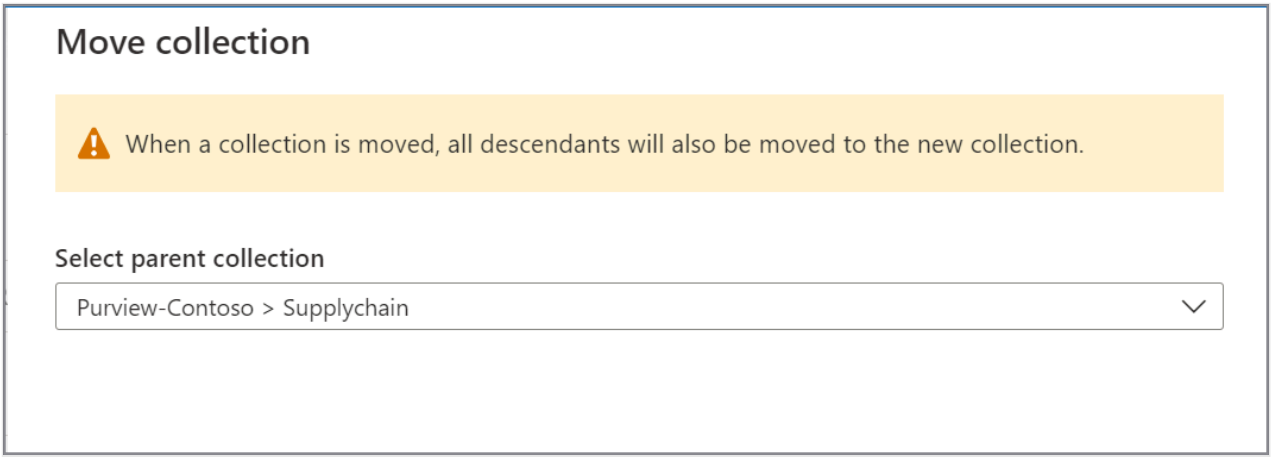
This feature is currently in preview. The [Supplemental Terms of Use for Microsoft Azure Previews] include additional legal terms that apply to Azure features that are in beta, in preview, or otherwise not yet released into general availability.

1. In the classic Microsoft Purview governance portal, navigate to the collection that you want to move to a different parent collection.

2. Select the ellipsis button on the right-top corner of collection path section and select **Move to another collection**.



3. From **Select parent collection** dropdown list, select a new parent for the collection and select **OK**.



Next steps

Now that you have a collection, you can follow these guides below to add resources and scan.

- Manage data sources
- Supported data sources
- Scan and ingestion

Quickstart: Create a collection and assign permissions in Microsoft Purview Data Map (classic)

Article • 05/30/2025

Collections are the Microsoft Purview Data Map's tool to manage ownership and access control across assets, sources, and information. They also organize your sources and assets into categories that are customized to match your management experience with your data. This guide takes you through setting up your first collection and collection admin to prepare your Microsoft Purview environment for your organization.

Prerequisites

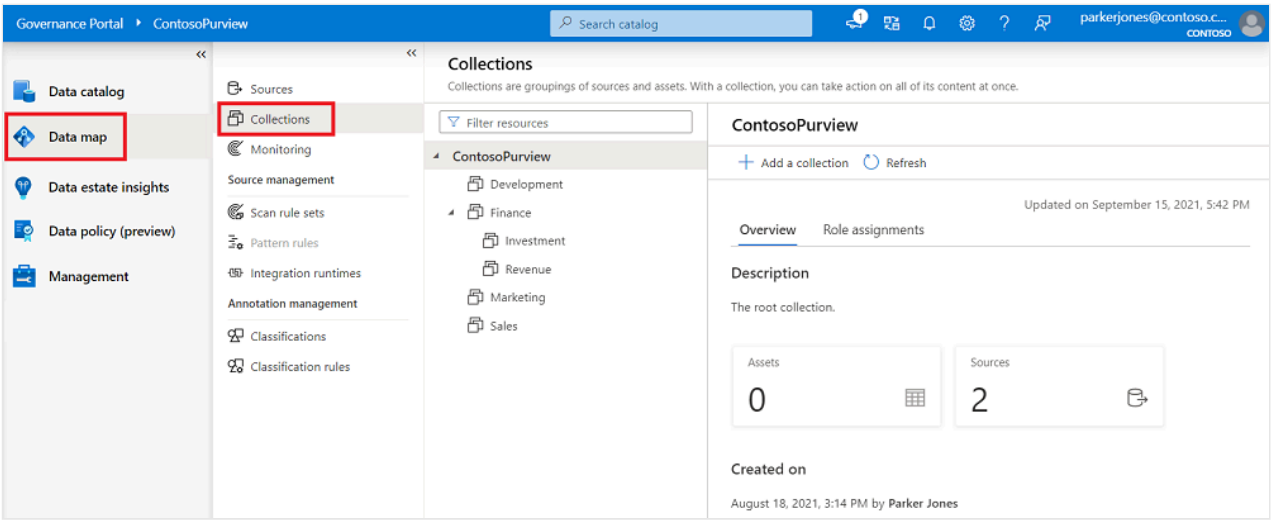
- An Azure account with an active subscription. Create an account for free.
- Your own Microsoft Entra tenant.
- An active Microsoft Purview account.

Check permissions

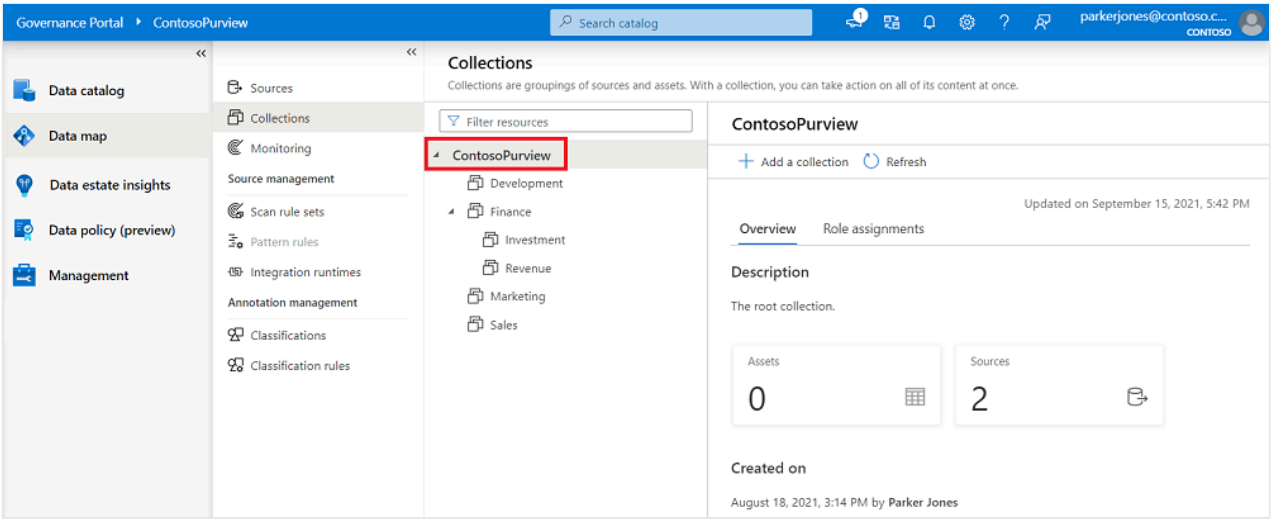
In order to create and manage collections in Data Map, you need to be a **Collection Admin** within the classic Microsoft Purview governance portal. We can check these permissions in the governance portal. You can find the governance portal by:

- Browsing directly to https://web.purview.azure.com and selecting your Microsoft Purview account.
- Open the Azure portal, search for and select the Microsoft Purview account you want to use to receive the share. Open the classic Microsoft Purview governance portal.

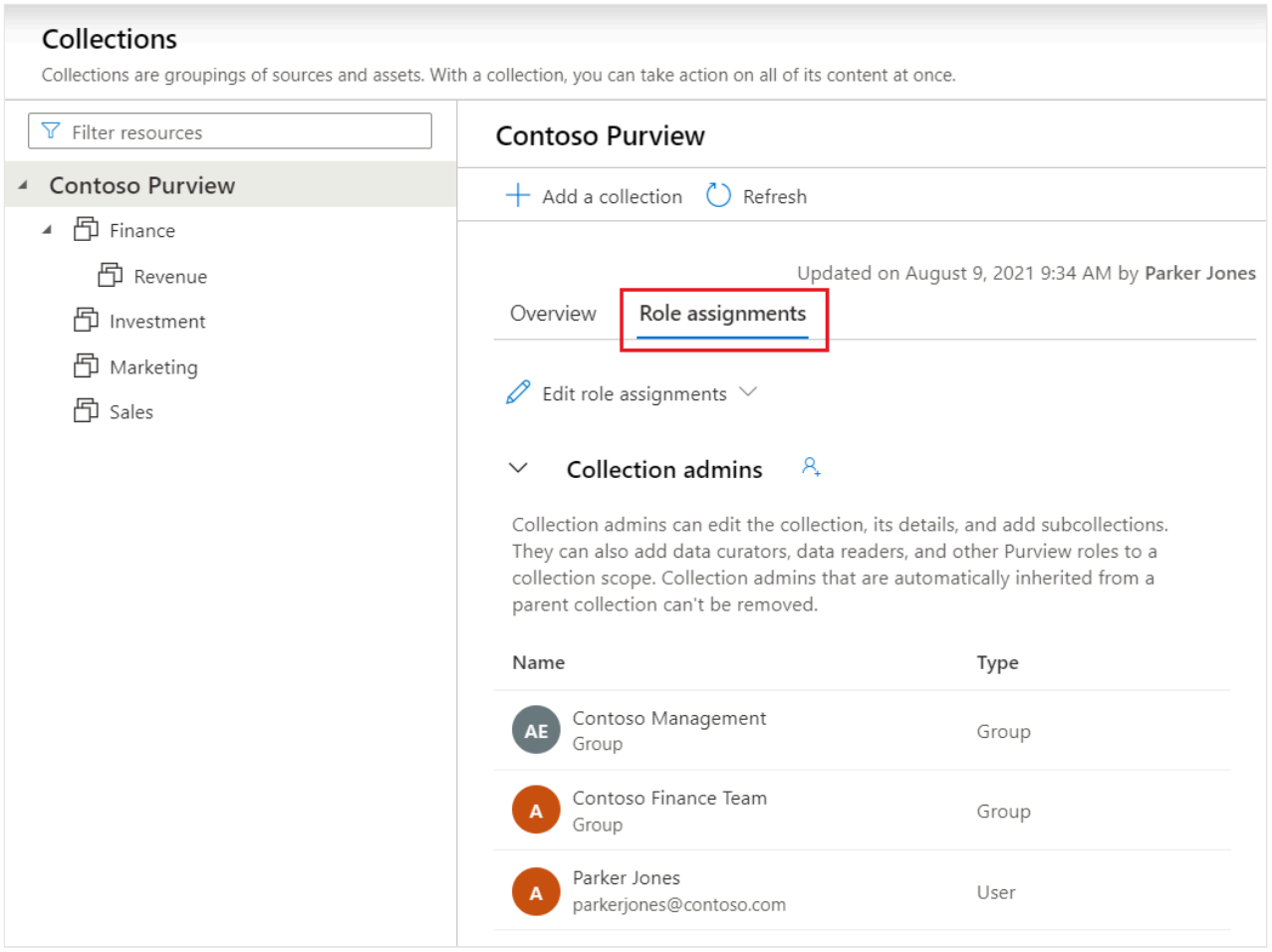
1. Select Data Map > Collections from the left pane to open collection management page.



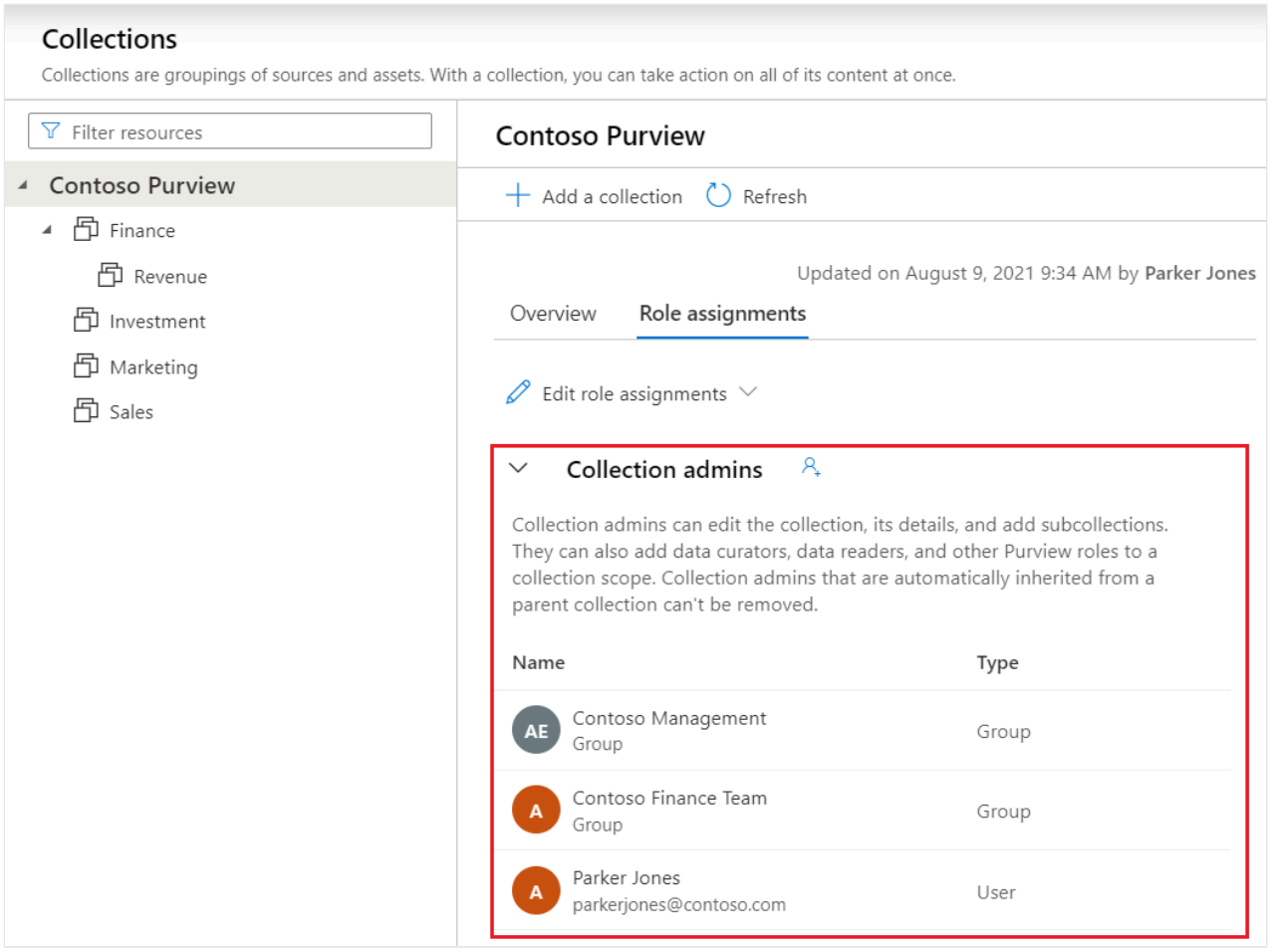
2. Select your root collection. This is the top collection in your collection list and will have the same name as your Microsoft Purview account. In our example below, it's called ContosoPurview.



3. Select role assignments in the collection window.



4. To create a collection, you need to be in the collection admin list under role assignments. If you created the account, you should be listed as a collection admin under the root collection already. If not, you need to contact the collection admin to grant you permission.

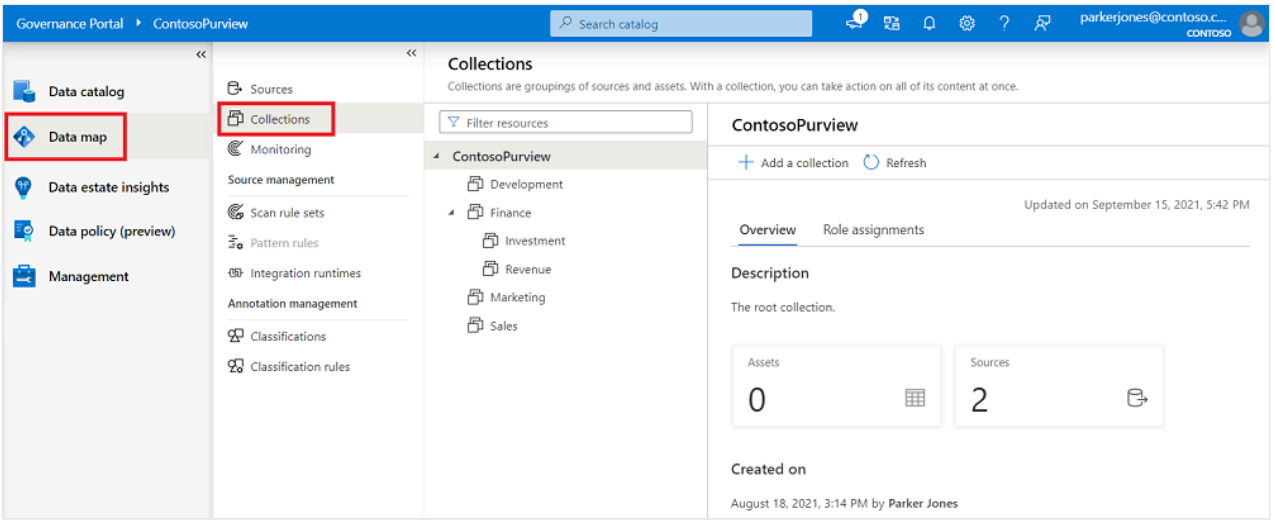


Create a collection in the portal

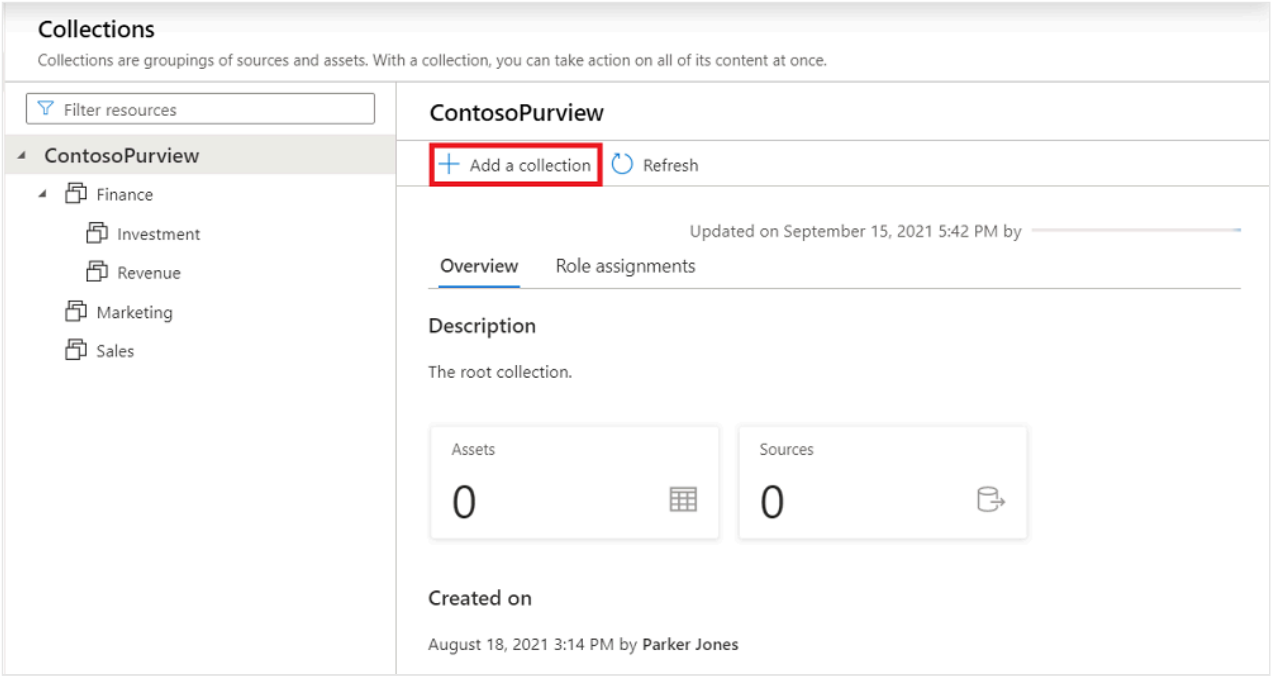
To create your collection, start in the classic Microsoft Purview governance portal. You can find the portal by:

- Browsing directly to https://web.purview.azure.com and selecting your Microsoft Purview account.
- Opening the Azure portal, searching for and selecting the Microsoft Purview account. Select the classic **Microsoft Purview governance portal** button.

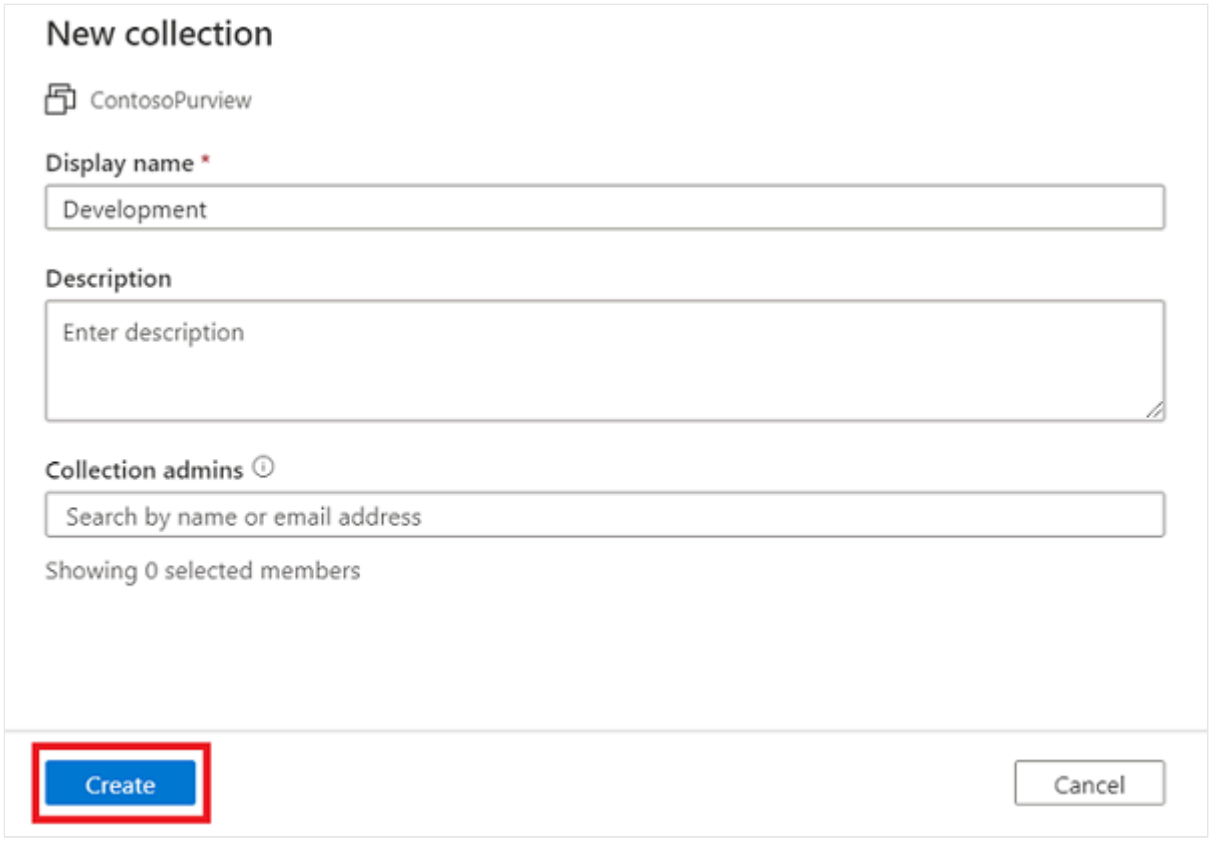
1. Select Data Map > Collections from the left pane to open collection management page.



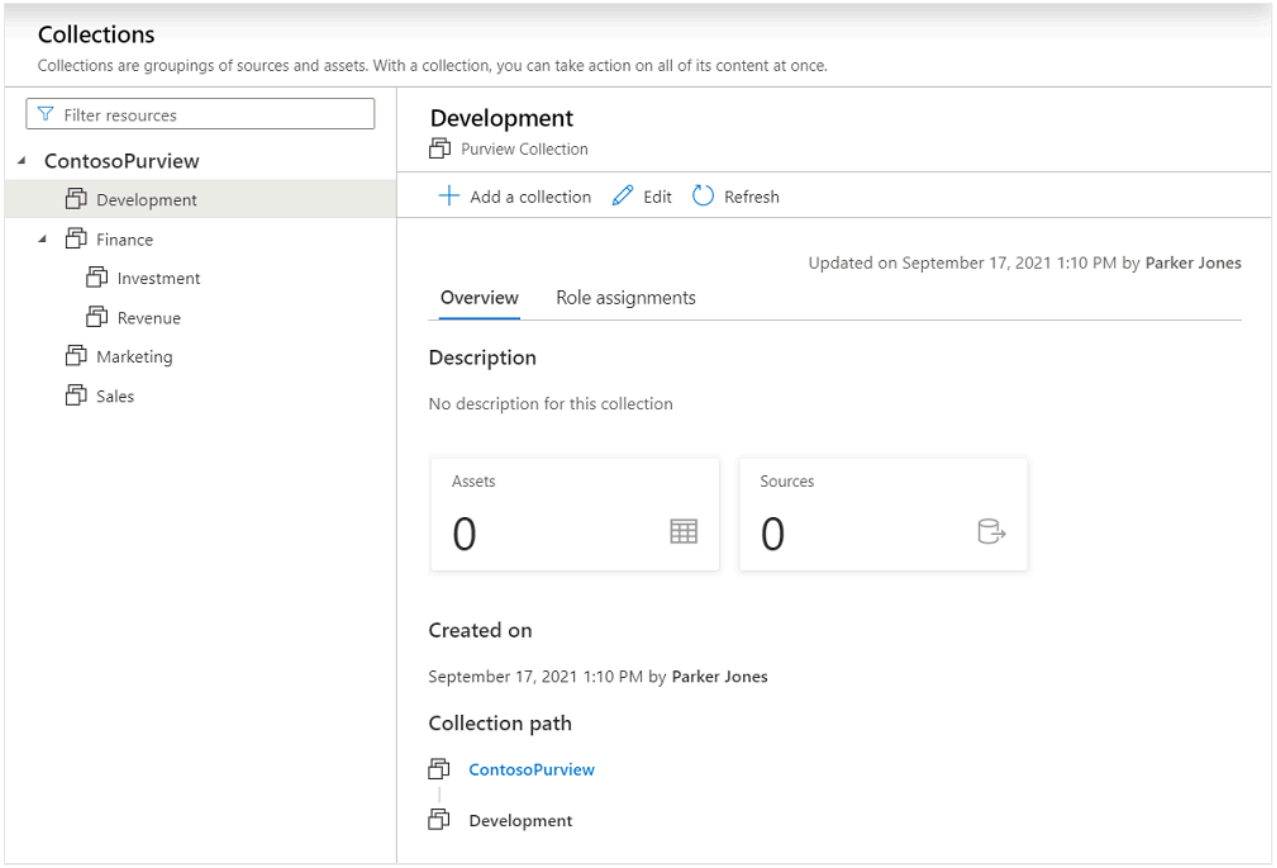
2. Select + **Add a collection**.



3. In the right panel, enter the collection name, description, and search for users to add them as collection admins.



4. Select **Create**. The collection information displays on the page.



Assign permissions to collection

Now that you have a collection, you can assign permissions to this collection to manage your users access to the classic Microsoft Purview governance portal.

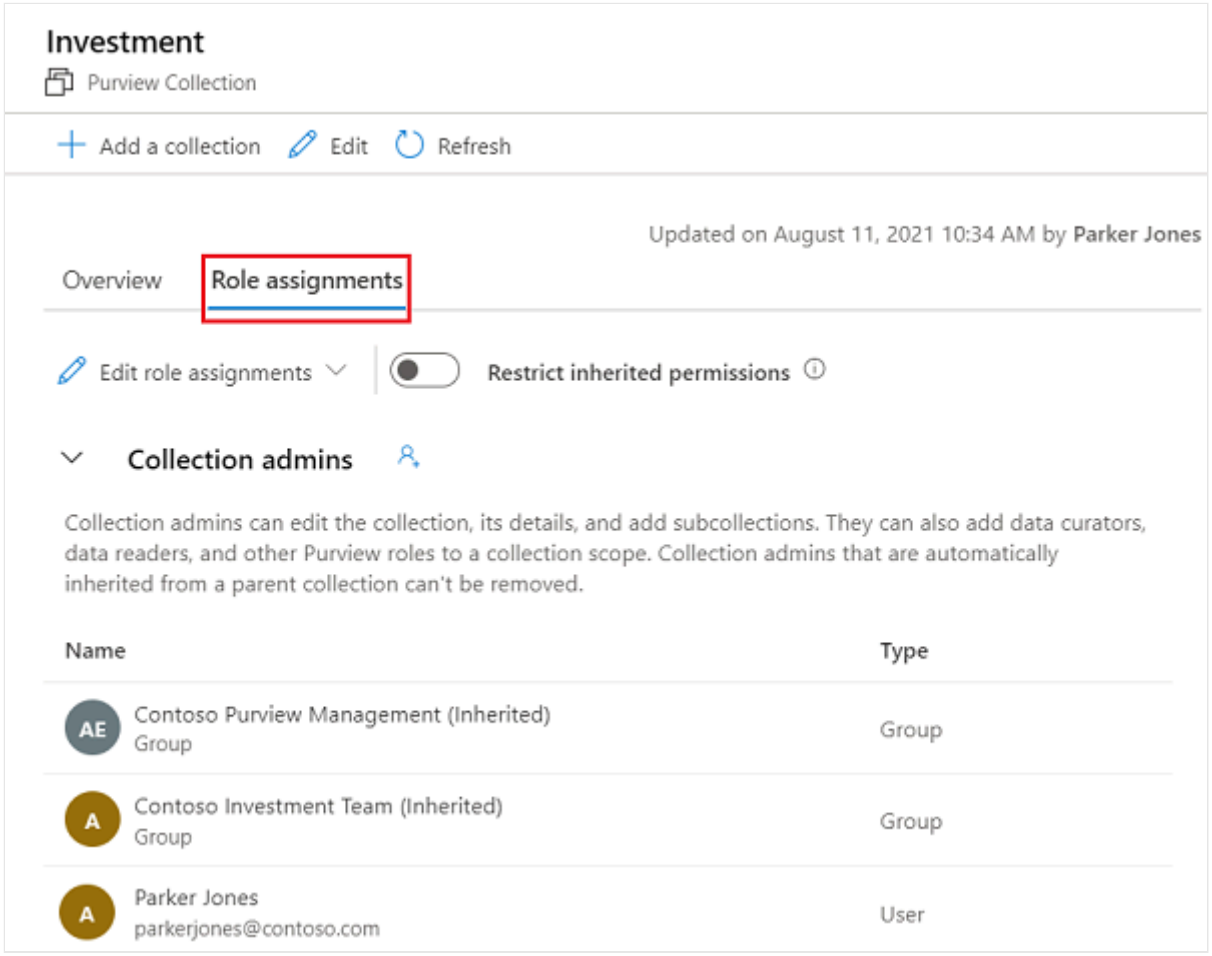
Roles

All assigned roles apply to sources, assets, and other objects within the collection where the role is applied.

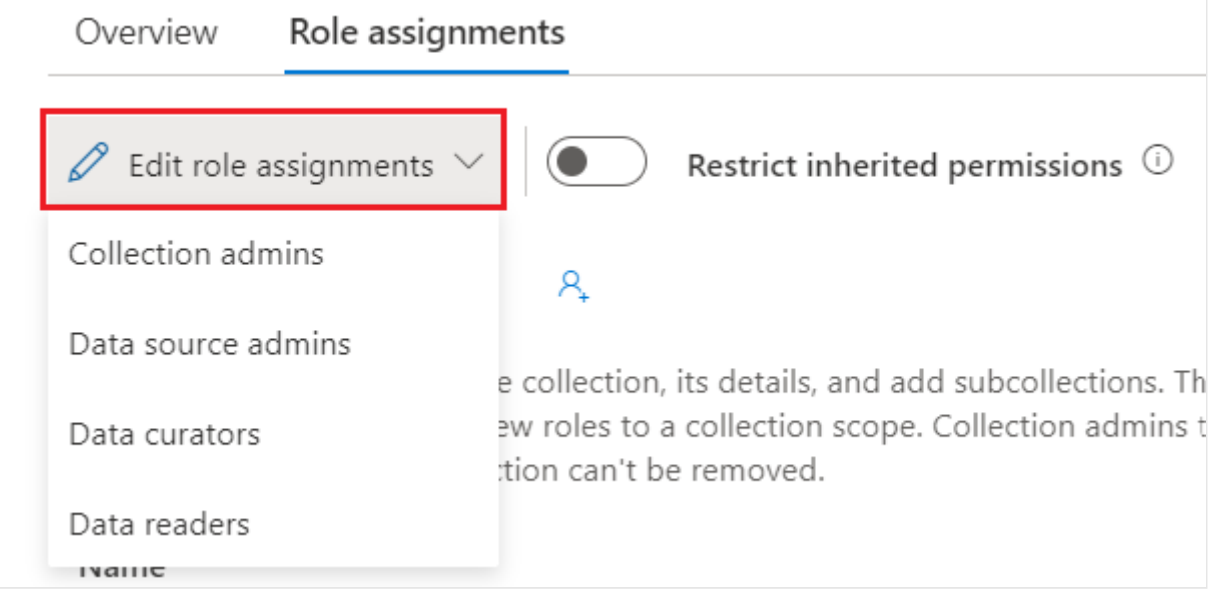
- **Collection admins** - can edit a collection, its details, manage access in the collection, and add subcollections.
- **Data source admins** - can manage data sources and data scans.
- **Data curators** - can create, read, modify, and delete actions on catalog data objects.
- **Data readers** - can access but not modify catalog data objects.

Assign permissions

1. Select **Role assignments** tab to see all the roles in a collection.



2. Select **Edit role assignments** or the person icon to edit each role member.



3. Type in the textbox to search for users you want to add to the role member. Select **OK** to save the change.

Next steps

Now that you have a collection, you can follow these guides to add resources, scan, and manage your collections.

- Register source to collection
- Access management through collections

Data sources that connect to Data Map

This article lists the supported data sources, file types, and scanning concepts in Microsoft Purview Data Map.

Data source listing by type

The tables below show all data sources that have technical metadata available in Microsoft Purview Data Map, along with other supported capabilities. Select a data source name in the **Data source** column for instructions on connecting that source to Data Map.

- Microsoft Azure
- Database
- File
- Services and apps

Azure

Azure resources are only available in the same tenant as your Microsoft Purview account, unless noted otherwise on each data source's page.

Expand table

Data source	Can automatically apply classifications	Can apply sensitivity labels to Data Map assets	Can apply policies	Data lineage	Accessible in live view
Select link for connection and scanning instructions.	*Select **Yes** for scanning instructions. Learn how classifications are applied during scanning.*	*Learn about sensitivity labeling (preview).*	*Select **Yes** to see supported policies; for example, data owner, self-service access, or protection.*	*Select **Yes** for details.*	*Learn about live view.*
Multiple sources	Yes	Source dependent	Yes	No	Limited
Azure Blob Storage	Yes	Yes	Yes (preview)	Limited*	Yes
Azure Cosmos	Yes	Yes	No	No*	No

Data source	Can automatically apply classifications	Can apply sensitivity labels to Data Map assets	Can apply policies	Data lineage	Accessible in live view
DB for SQL API					
Azure Data Explorer	Yes	Yes	No	No*	No
Azure Data Factory	No	No	No	Yes	No
Azure Data Lake Storage Gen2	Yes	Yes	Yes (preview)	Limited*	Yes
Azure Data Share	No	No	No	Yes	No
Azure Database for MySQL	Yes	Yes	No	No*	No
Azure Database for PostgreSQL	Yes	Yes	No	No*	No
Azure Databricks Hive Metastore	Yes**	No	No	Yes	No
Azure Databricks Unity Catalog	Yes***	No	No	Yes	No
Azure Dedicated SQL pool (formerly SQL DW)	Yes	No	No	No*	No
Azure Files	Yes	Yes	No	Limited*	No
Azure Machine Learning	No	No	No	Yes	No
Azure SQL Database	Yes	Yes	Yes^	Yes (Preview)	Yes
Azure SQL Managed Instance	Yes	Yes	Yes	No*	No
Azure Synapse Analytics	Yes	Yes	No	Yes - Synapse	No

Data source	Can automatically apply classifications	Can apply sensitivity labels to Data Map assets	Can apply policies	Data lineage	Accessible in live view
(Workspace)				pipelines	

* Besides the lineage on assets within the data source, lineage is also supported if dataset is used as a source/sink in Data Factory or Synapse pipeline.

** Classification support for only Unity Catalog in Azure Databricks data source.

*** Custom classification isn't supported for Azure Databricks Unity Catalog source.

^ Protection policies aren't supported for Azure SQL Database.

Database

Expand table

Data source	Can automatically apply classifications	Can apply sensitivity labels to Data Map assets	Can apply policies	Data lineage	Accessible in live view
Select link for connection and scanning instructions.	*Select **Yes** for scanning instructions. Learn how classifications are applied during scanning.*	*Learn about sensitivity labeling (preview).*	*Select **Yes** to see supported policies; for example, data owner, self-service access, or protection.*	*Select **Yes** for details.*	*Learn about live view.*
Amazon RDS	Yes	No	No	No	No
Amazon Redshift	No	No	No	No	No
Cassandra	No	No	No	Yes	No
Db2	No	No	No	Yes	No
Google BigQuery	No	No	No	Yes	No
Hive Metastore Database	No	No	No	Yes*	No
MongoDB	No	No	No	No	No

Data source	Can automatically apply classifications	Can apply sensitivity labels to Data Map assets	Can apply policies	Data lineage	Accessible in live view
MySQL	No	No	No	Yes	No
Oracle	Yes	No	No	Yes*	No
PostgreSQL	No	No	No	Yes	No
SAP Business Warehouse	No	No	No	No	No
SAP HANA	No	No	No	No	No
Snowflake	Yes	Yes	No	Yes*	No
SQL Server	Yes	Yes	No	No*	No
SQL Server on Azure-Arc	Yes	No	Yes	No*	No
Teradata	Yes	No	No	Yes*	No

* Besides the lineage on assets within the data source, lineage is also supported if dataset is used as a source/sink in Data Factory or Synapse pipeline.

File

Expand table

Data source	Can automatically apply classifications	Can apply sensitivity labels to Data Map assets	Can apply policies	Data lineage	Accessible in live view
Select link for connection and scanning instructions.	*Select **Yes** for scanning instructions. Learn how classifications are applied during scanning.*	*Learn about sensitivity labeling (preview).*	*Select **Yes** to see supported policies; for example, data owner, self-service access, or protection.*	*Select **Yes** for details.*	*Learn about live view.*
Amazon S3	Yes	No	No	Limited*	No
Hadoop Distributed File System (HDFS)	Yes	No	No	No	No

* Besides the lineage on assets within the data source, lineage is also supported if dataset is used as a source/sink in Data Factory or Synapse pipeline.

Services and apps

Expand table

Data source	Can automatically apply classifications	Can apply sensitivity labels to Data Map assets	Can apply policies	Data lineage	Accessible in live view
Select link for connection and scanning instructions.	*Select **Yes** for scanning instructions. Learn how classifications are applied during scanning.*	*Learn about sensitivity labeling (preview).*	*Select **Yes** to see supported policies; for example, data owner, self-service access, or protection.*	*Select **Yes** for details.*	*Learn about live view.*
Airflow	No	No	No	Yes	No
Dataverse	Yes	Yes	No	No	No
Erwin	No	No	No	Yes	No
Fabric	No	No	No	Yes	Yes
Looker	No	No	No	Yes	No
Power BI	No	No	No	Yes	Yes**
Qlik Sense	No	No	No	No	No
Salesforce	No	No	No	No	No
SAP ECC	No	No	No	Yes*	No
SAP S/4HANA	No	No	No	Yes*	No
Tableau	No	No	No	No	No

* Besides the lineage on assets within the data source, lineage is also supported if dataset is used as a source/sink in Data Factory or Synapse pipeline.

** Power BI items in a Fabric tenant are available using live view.

> **Note**

Currently, the Microsoft Purview Data Map can't scan an asset that has `/`, `\`, or `#` in its name. To scope your scan and avoid scanning assets that have those characters in the asset name, use the example in Register and scan an Azure SQL Database.

> **Important**
>
> If you plan on using a self-hosted integration runtime, scanning some data sources requires extra setup on the self-hosted integration runtime machine. For example, JDK, Microsoft Visual C++ Redistributable, or specific driver. For your source, refer to each source article for prerequisite details. Any requirements are listed in the **Prerequisites** section.

Data Map scanner regions

The following list shows the Azure data source (data center) regions where the Data Map scanner runs. If your Azure data source is in a region outside of this list, the scanner runs in the region of your Microsoft Purview instance.

- Australia East
- Australia Southeast
- Brazil South
- Canada Central
- Canada East
- Central India
- China North 3
- East Asia
- East US
- East US 2
- France Central
- Germany West Central
- Japan East
- Korea Central
- North Central US
- North Europe
- Qatar Central
- South Africa North
- South Central US
- Southeast Asia

- Switzerland North
- UAE North
- UK South
- USGov Virginia
- West Central US
- West Europe
- West US
- West US 2
- West US 3

File types supported for scanning

The file types listed in the following section support scanning, schema extraction, and classification where applicable. Additionally, Data Map supports custom file extensions and custom parsers.

Structured file formats supported by extension include scanning, schema extraction, and asset and column level classification:

- AVRO
- CSV
- GZIP
- JSON
- ORC
- PARQUET*
- PSV
- SSV
- TSV
- TXT
- XML

*For noncompressed PARQUET files, all Parquet formats are supported. For compressed PARQUET files, only snappy Parquet format is supported.

Document file formats supported by extension include scanning and asset level classification:

- DOC
- DOCM
- DOCX

- DOT
- ODP
- ODS
- ODT
- PDF
- POT
- PPS
- PPSX
- PPT
- PPTM
- PPTX
- XLC
- XLS
- XLSB
- XLSM
- XLSX
- XLT

> ⓘ **Note**
>
> **Known limitations:**
>
> - The Microsoft Purview Data Map scanner only supports schema extraction for the structured file types listed in the previous section.
> - For AVRO, ORC, and PARQUET file types, the scanner doesn't support schema extraction for files that contain complex data types (for example, MAP, LIST, STRUCT).
> - For noncompressed PARQUET files, all Parquet formats are supported. For compressed PARQUET files, only snappy Parquet format is supported for schema extraction and classification.
> - For GZIP file types, the GZIP must be mapped to a single CSV file within. GZIP files are subject to system and custom classification rules. The scanner currently doesn't support scanning a GZIP file mapped to multiple files within, or any file type other than CSV.
> - For Parquet files, if you're using a self-hosted integration runtime, you need to install the **64-bit JRE 11 (Java Runtime Environment) or OpenJDK** on your IR machine. See the Java runtime installation guide.

- The Delta format isn't supported. If you're scanning the Delta format directly from storage data source like Microsoft Azure Data Lake Storage Gen2, the set of Parquet files from the delta format are parsed and handled as resource set as described in Understanding resource sets. The columns used for partitioning aren't recognized as part of the schema for the resource set.

- Unsupported Parquet encodings: Microsoft Purview doesn't support schema extraction for Parquet files that use run-length encoding (RLE) in Microsoft Fabric Lakehouse. Tables containing RLE-encoded columns might fail to extract schema metadata, including tables, columns, and data types. The process still discovers asset-level metadata (L1). To enable full schema extraction, re-create the affected tables by using a supported encoding format, such as dictionary or plain encoding.

For delimited file types (CSV, PSV, SSV, TSV, TXT):

- Delimited files with only one column can't be determined to be CSV files and have no schema.
- Data type detection isn't supported. The data type is listed as "string" for all columns.
- The only supported delimiters are comma(','), semicolon(';'), vertical bar('|'), and tab('\t').
- Delimited files with less than three rows can't be determined to be CSV files if they're using a custom delimiter. For example, files with ~ delimiter and less than three rows can't be determined to be CSV files.
- If a field contains double quotes, the double quotes can only appear at the beginning and end of the field and must be matched. Double quotes that appear in the middle of the field or appear at the beginning and end but aren't matched are recognized as bad data and no schema is parsed from the file. Rows that have different number of columns than the header row are judged as error rows. The numbers of error rows divided by the numbers of rows sampled must be less than 0.1.

Schema extraction

For data sources that support schema extraction during scan, the number of columns doesn't directly truncate the asset schema.

Nested data

Nested data is only supported for JSON content. For all system supported file types, if there's nested JSON content in a column, then the scanner parses the nested JSON data and surfaces it

within the schema tab of the asset.

Nested data, or nested schema parsing, isn't supported in SQL. A column with nested data will be reported and classified as is, and subdata won't be parsed.

Sampling data for classification

In Data Map terminology,

- L1 scan: Extracts basic information and metadata like file name, size, and fully qualified name.
- L2 scan: Extracts schema for structured file types and database tables.
- L3 scan: Extracts schema where applicable and subjects the sampled file to system and custom classification rules.

Learn more about customizing the scan levels.

For all structured file formats, the Microsoft Purview Data Map scanner samples files in the following way:

- For structured file types, it samples the top 128 rows in each column or the first 1 MB, whichever is lower.
- For document file formats, it samples the first 20 MB of each file.
- If a document file is larger than 20 MB, the scanner doesn't perform a deep scan (subject to classification). In that case, Microsoft Purview captures only basic metadata like file name and fully qualified name.
- For **tabular data sources (SQL)**, it samples the top 128 rows.
- For **Azure Cosmos DB for NoSQL**, up to 300 distinct properties from the first 10 documents in a container are collected for the schema. For each property, the scanner samples values from up to 128 documents or the first 1 MB.
- For unstructured data formats, such as DOCX files, there's no requirement to meet a condition of eight distinct values. The presence of a single relevant keyword is sufficient for classification.

Resource set file sampling

If a folder or group of partition files matches a system resource set policy or a customer-defined resource set policy, Data Map detects it as a *resource set*. If the scanner detects a resource set, it samples each folder it contains. For more information about resource sets, see Resource sets in Microsoft Purview Data Map.

File sampling for resource sets by file types:

- **Delimited files (CSV, PSV, SSV, TSV)**: The scanner samples 1 in 100 files (L3 scan) within a folder or group of partition files that are considered as a resource set.
- **Data Lake file types (Parquet, Avro, Orc)**: The scanner samples 1 in 18,446,744,073,709,551,615 (long max) files (L3 scan) within a folder or group of partition files that are considered as a resource set.
- **Other structured file types (JSON, XML, TXT)**: The scanner samples 1 in 100 files (L3 scan) within a folder or group of partition files that are considered as a resource set.
- **SQL objects and Azure Cosmos DB entities**: The scanner L3 scans each file.
- **Document file types**: The scanner L3 scans each file. Resource set patterns don't apply to these file types.

Next steps

- Discover and govern Azure Blob storage source
- Scans and ingestion in Microsoft Purview
- Manage data sources in Microsoft Purview

Last updated on 07/01/2026

Credentials for source authentication in Microsoft Purview Data Map

10/24/2025

This article describes how you can create credentials in Microsoft Purview Data Map. These saved credentials let you quickly reuse and apply saved authentication information to your data source scans.

Prerequisites

- A Microsoft Azure Key Vault. To learn how to create one, see Quickstart: Create a key vault using the Azure portal.

Introduction

A credential is authentication information that Microsoft Purview can use to authenticate to your registered data sources. A credential object can be created for various types of authentication scenarios, such as Basic Authentication requiring username/password. Credential capture specific information required to authenticate, based on the chosen type of authentication method. Credentials use your existing Key Vault secrets for retrieving sensitive authentication information during the Credential creation process.

In Microsoft Purview, there are few options to use as authentication method to scan data sources such as the following options. Learn from each data source article for its supported authentication.

- Microsoft Purview system-assigned managed identity
- User-assigned managed identity (preview)
- Account Key (using Key Vault)
- SQL Authentication (using Key Vault)
- Service Principal (using Key Vault)
- Consumer Key (using Key Vault)
- And more

Before creating any credentials, consider your data source types and networking requirements to decide which authentication method you need for your scenario.

Use Microsoft Purview system-assigned managed identity to set up scans

If you're using the Microsoft Purview system-assigned managed identity (SAMI) to set up scans, you won't need to create a credential and link your key vault to Microsoft Purview to store them. For detailed instructions on adding the Microsoft Purview SAMI to have access to scan your data sources, refer to the data source-specific authentication sections below:

- Azure Blob Storage
- Azure Data Lake Storage Gen1
- Azure Data Lake Storage Gen2
- Azure SQL Database
- Azure SQL Managed Instance
- Azure Synapse Workspace
- Azure Synapse dedicated SQL pools (formerly SQL DW)

Grant Microsoft Purview access to your Azure Key Vault

To give Microsoft Purview access to your Azure Key Vault, there are two things you need to confirm:

- Firewall access to the Azure Key Vault
- Microsoft Purview permissions on the Azure Key Vault

Firewall access to Azure Key Vault

If your Azure Key Vault has disabled public network access, you have two options to allow access for Microsoft Purview.

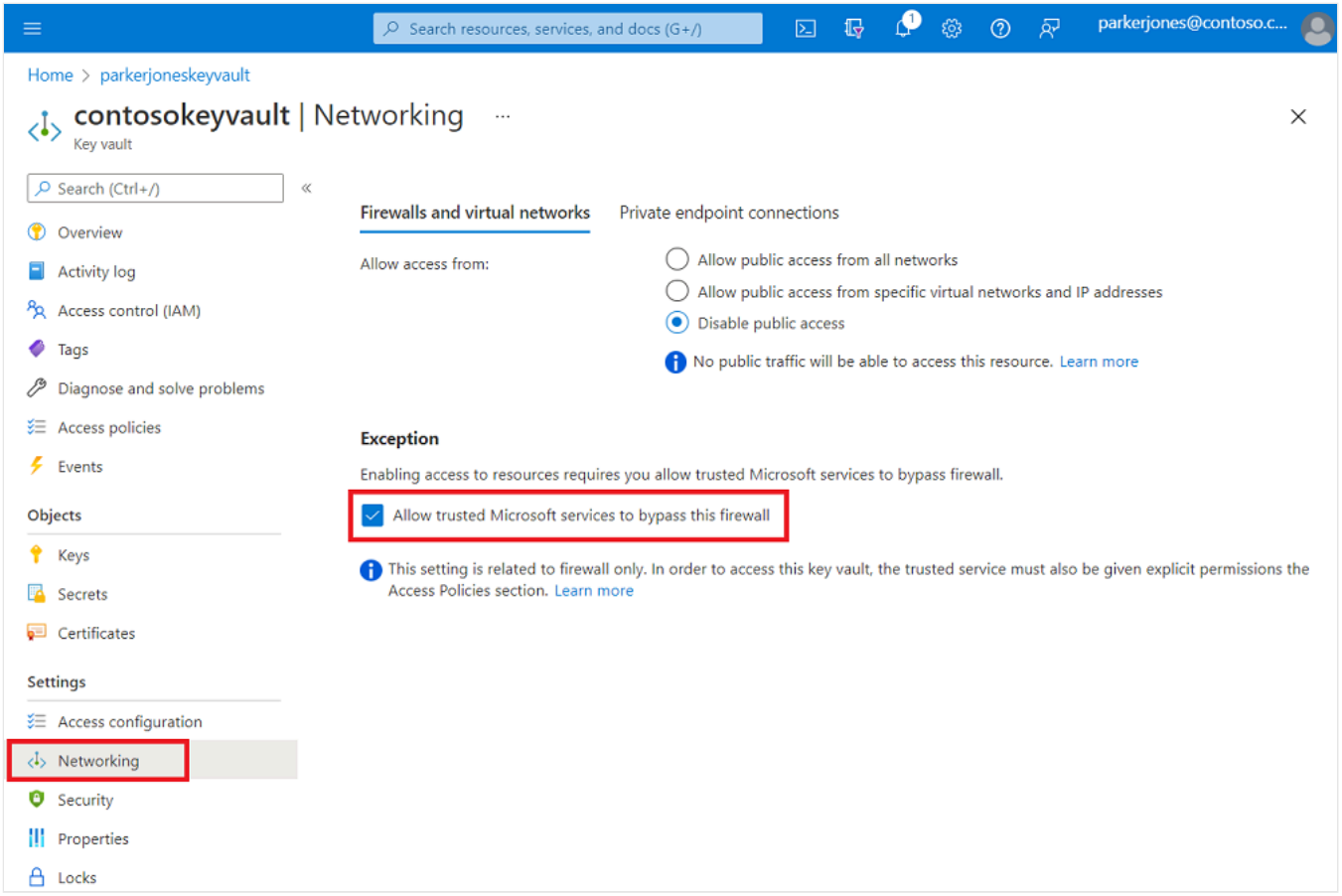
- Trusted Microsoft services
- Private endpoint connections

Trusted Microsoft services

Microsoft Purview is listed as one of Azure Key Vault's trusted services, so if public network access is disabled on your Azure Key Vault you can enable access only to trusted Microsoft services, and Microsoft Purview will be included.

You can enable this setting in your Azure Key Vault under the **Networking** tab.

At the bottom of the page, under Exception, enable the **Allow trusted Microsoft services to bypass this firewall** feature.



Private endpoint connections

To connect to Azure Key Vault with private endpoints, follow Azure Key Vault's private endpoint documentation.

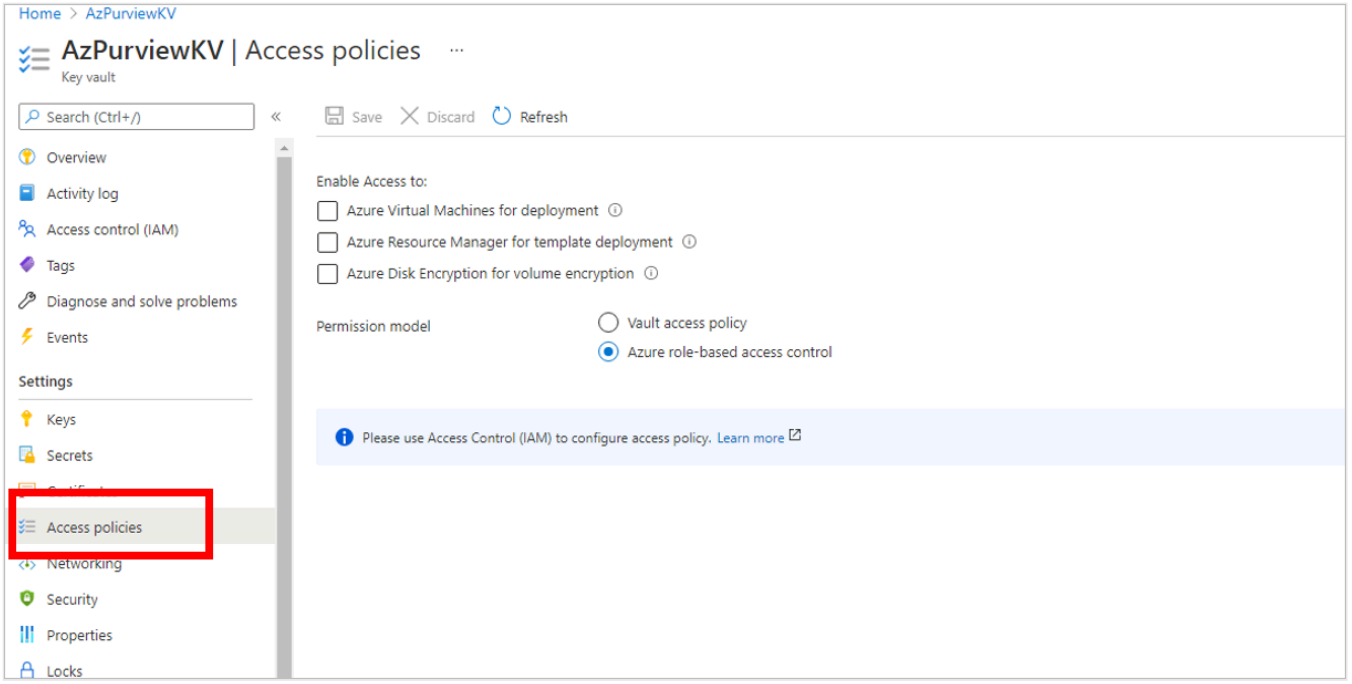
> **Note**
>
> Private endpoint connection option is supported when using Azure integration runtime in managed virtual network to scan the data sources. For self-hosted integration runtime, you need to enable trusted Microsoft services.

Microsoft Purview permissions on the Azure Key Vault

Currently Azure Key Vault supports two permission models:

- Option 1 - Access Policies
- Option 2 - Role-based Access Control

Before assigning access to the Microsoft Purview system-assigned managed identity (SAMI), first identify your Azure Key Vault permission model from Key Vault resource **Access Policies** in the menu. Follow steps below based on relevant the permission model.



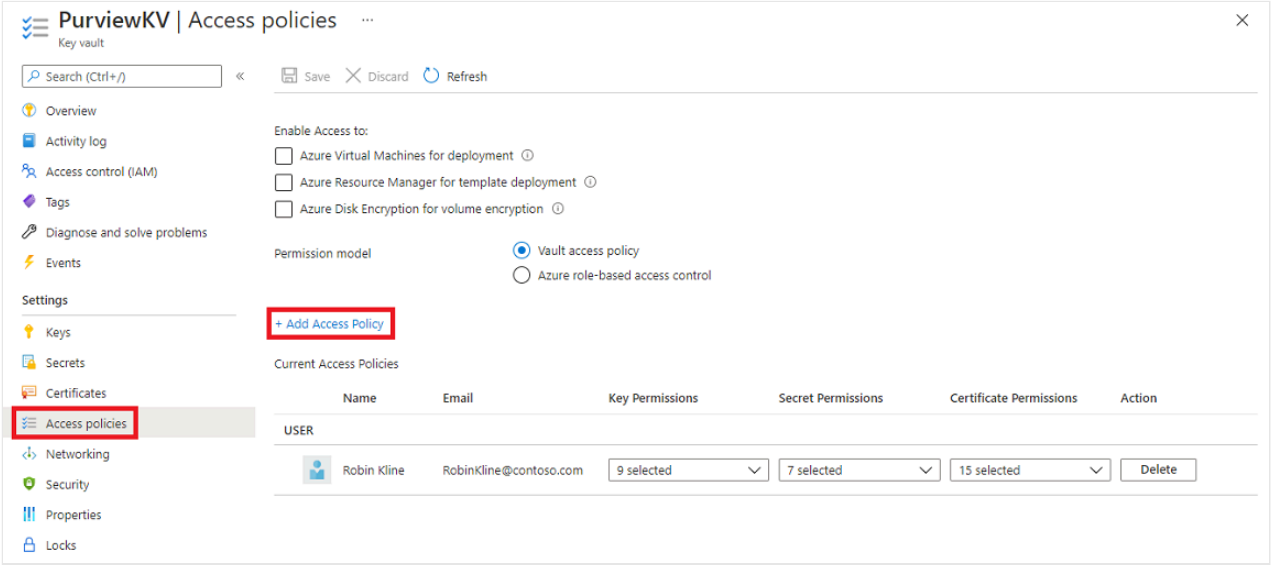
Option 1 - Assign access using Key Vault Access Policy

Follow these steps only if permission model in your Azure Key Vault resource is set to **Vault Access Policy**:

1. Navigate to your Azure Key Vault.

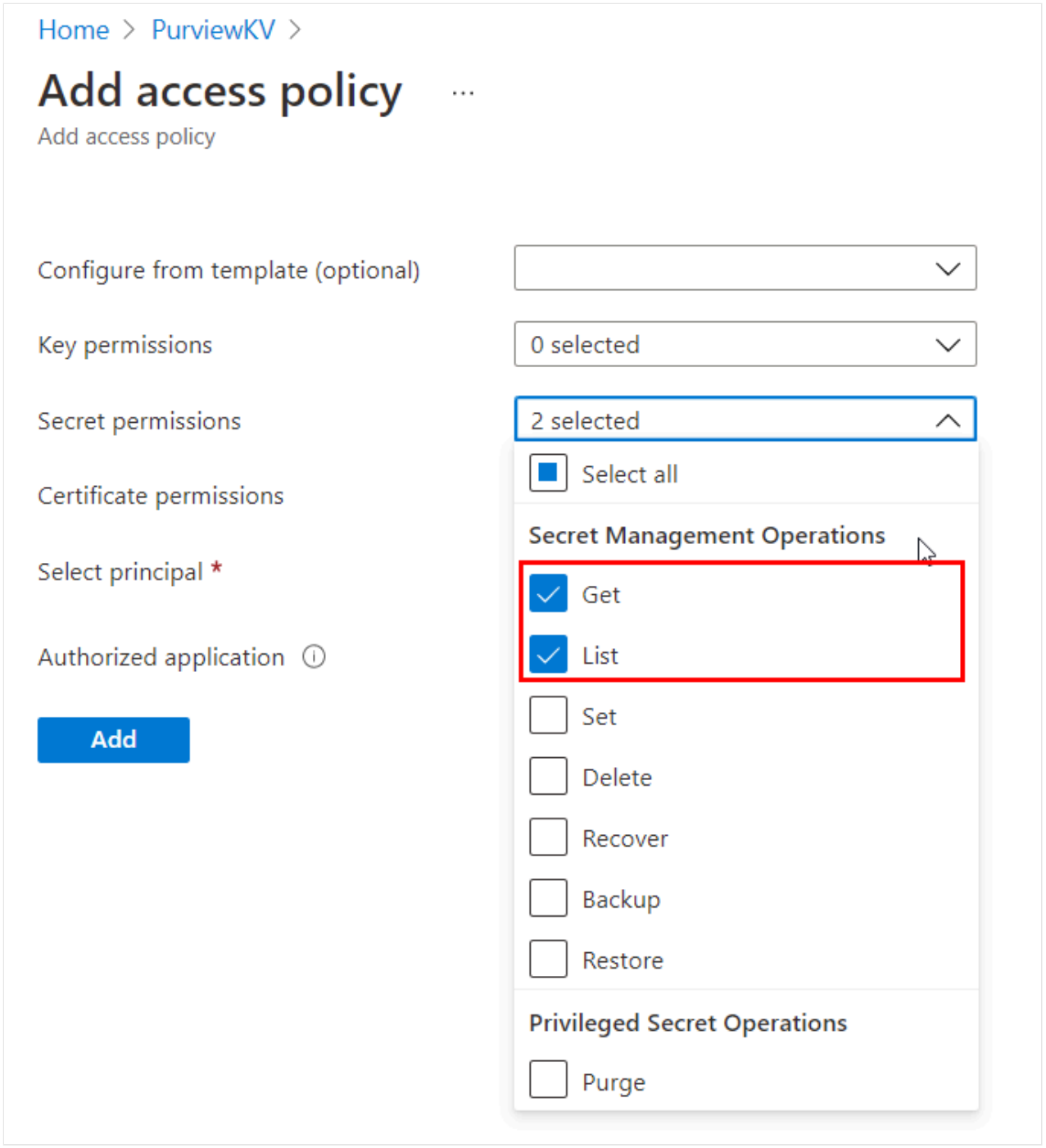
2. Select the **Access policies** page.

3. Select **Add Access Policy**.



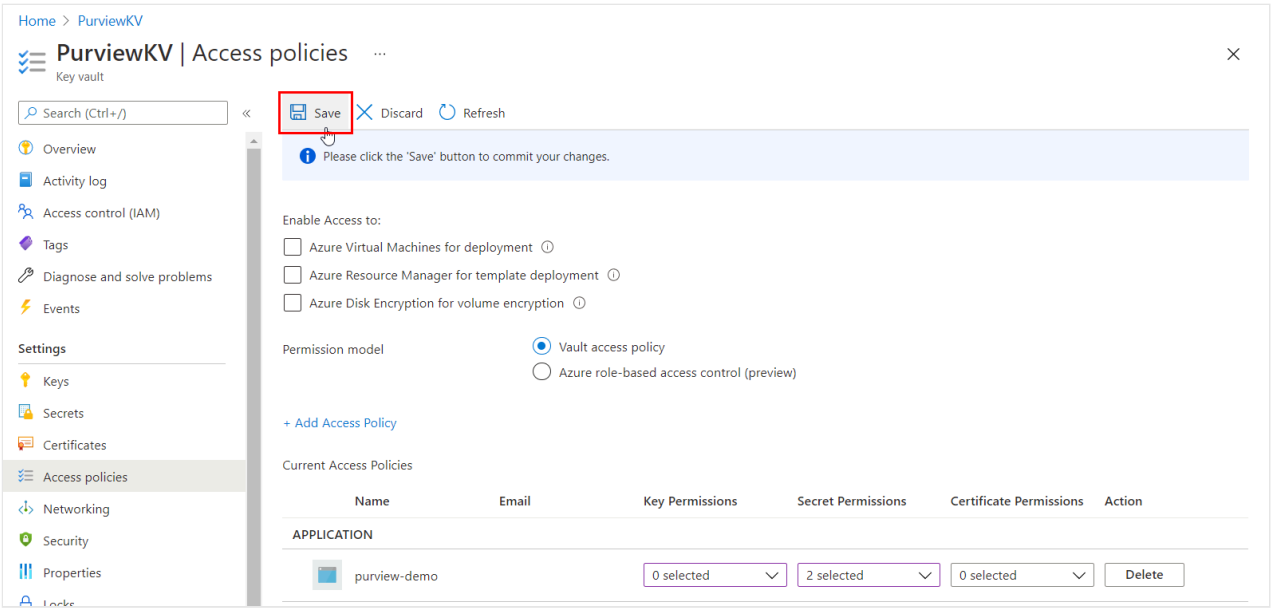
4. In the **Secrets permissions** dropdown, select **Get** and **List** permissions.

5. For **Select principal**, choose the Microsoft Purview system managed identity. You can search for the Microsoft Purview SAMI using either the Microsoft Purview instance name **or** the managed identity application ID. We don't currently support compound identities (managed identity name + application ID).



6. Select **Add**.

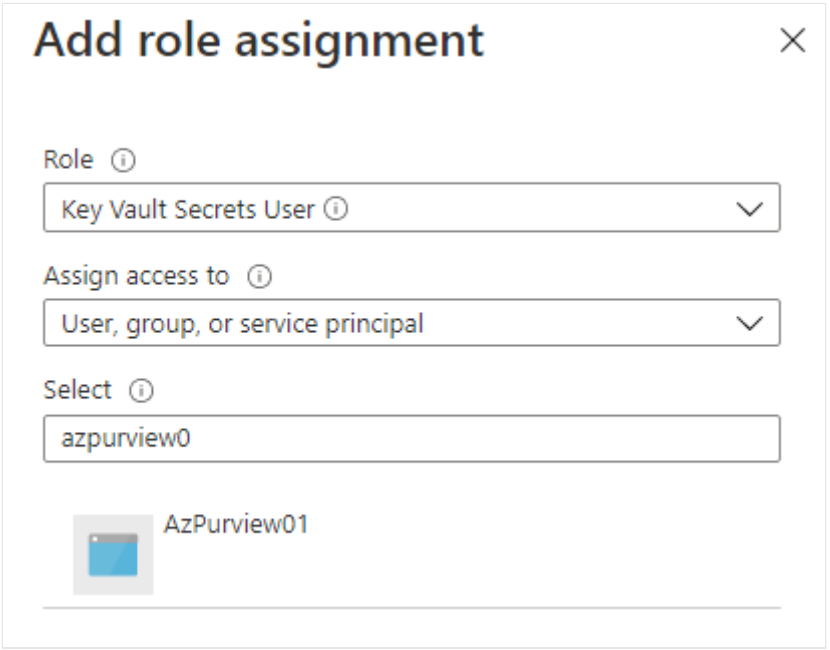
7. Select **Save** to save the Access policy.



Option 2 - Assign access using Key Vault Azure role-based access control

Follow these steps only if permission model in your Azure Key Vault resource is set to **Azure role-based access control**:

1. Navigate to your Azure Key Vault.
2. Select **Access Control (IAM)** from the left navigation menu.
3. Select **+ Add**.
4. Set the **Role** to **Key Vault Secrets User** and enter your Microsoft Purview account name under **Select** input box. Then, select Save to give this role assignment to your Microsoft Purview account.

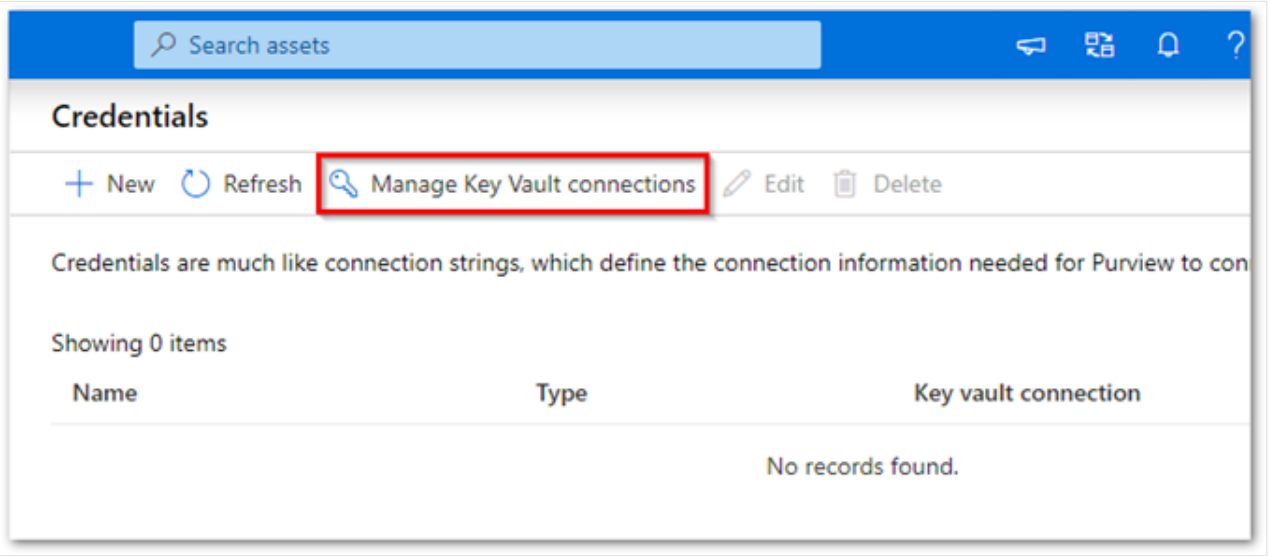


Create Azure Key Vaults connections in your Microsoft Purview account

Before you can create a Credential, first associate one or more of your existing Azure Key Vault instances with your Microsoft Purview account.

1. If you're using the classic Microsoft Purview portal, navigate to the **Management Center** in the studio and then navigate to **Credentials**. If you're using the new Microsoft Purview portal, open **Data Map**, select **Source management**, then select **Credentials**.

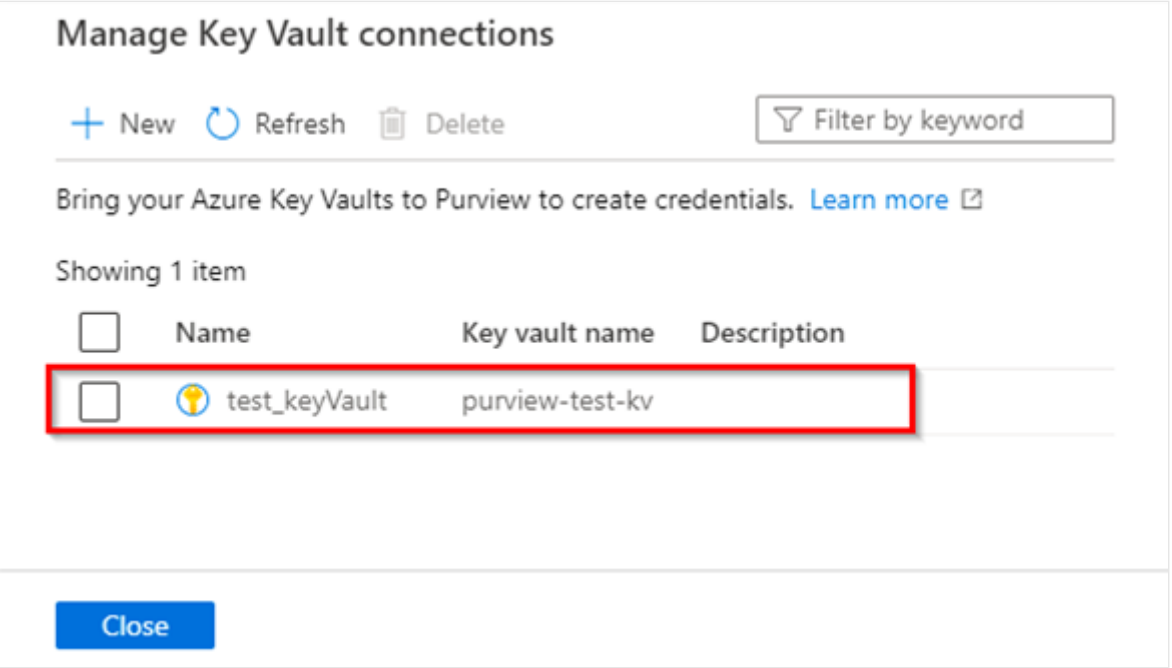
2. From the **Credentials** page, select **Manage Key Vault connections**.



3. Select + **New** from the Manage Key Vault connections page.

4. Provide the required information, then select **Create**.

5. Confirm that your Key Vault has been successfully associated with your Microsoft Purview account as shown in this example:



Create a new credential

These credential types are supported in Microsoft Purview:

- Basic authentication: You add the **password** as a secret in key vault.
- Service Principal: You add the **service principal key** as a secret in key vault.
- SQL authentication: You add the **password** as a secret in key vault.
- Windows authentication: You add the **password** as a secret in key vault.
- Account Key: You add the **account key** as a secret in key vault.
- Role ARN: For an Amazon S3 data source, add your **role ARN** in AWS.
- Consumer Key: For Salesforce data sources, you can add the **password** and the **consumer secret** in key vault.
- User-assigned managed identity (preview): You can add user-assigned managed identity credentials. For more information, see the create a user-assigned managed identity section below.

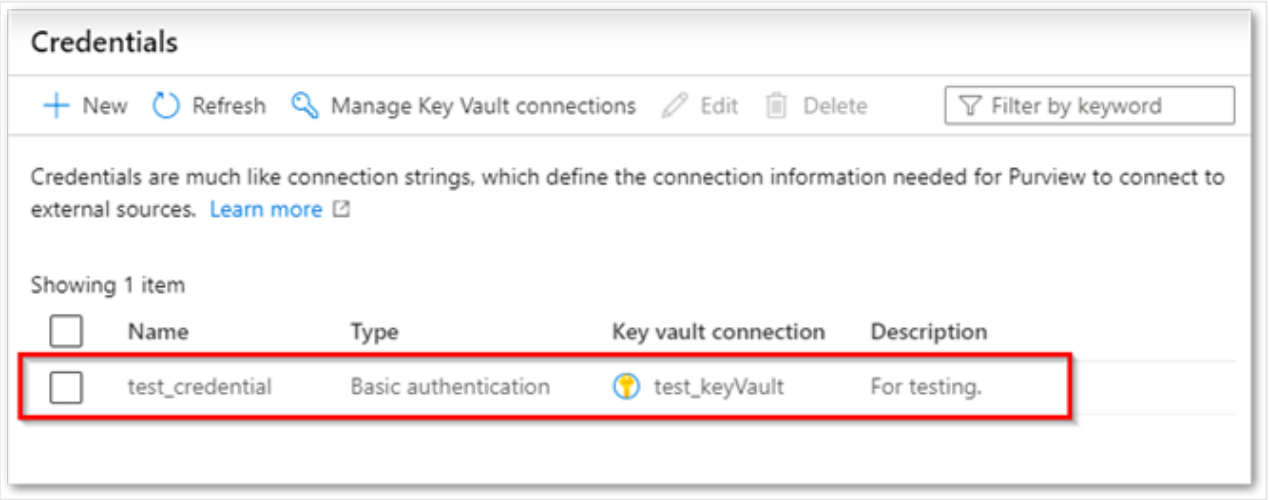
For more information, see Add a secret to Key Vault and Create a new AWS role for Microsoft Purview.

After storing your secrets in the key vault:

1. Provide 'Reader access' to the user at the Azure subscription level.
2. In Microsoft Purview, go to the Credentials page.
3. Create your new Credential by selecting + **New**.
4. Provide the required information. Select the **Authentication method** and a **Key Vault connection** from which to select a secret from.

5. Once all the details have been filled in, select **Create**.

6. Verify that your new credential shows up in the list view and is ready to use.



Manage your key vault connections

1. Search for Key Vault connections by entering a name in the search box.

2. Delete one or more Key Vault connections by selecting the checkbox next to the Key Vault's name, then selecting **Delete**.

Manage your credentials

1. Search/find Credentials by name.

2. Select and make updates to an existing Credential.

3. Delete one or more Credentials.

Create a user-assigned managed identity

User-assigned managed identities (UAMI) enable Azure resources to authenticate directly with other resources using Microsoft Entra authentication, without the need to manage those credentials. They allow you to authenticate and assign access just like you would with a system assigned managed identity, Microsoft Entra user, Microsoft Entra group, or service principal. User-assigned managed identities are created as their own resource (rather than being connected to a preexisting resource). For more information about managed identities, see the managed identities for Azure resources documentation.

The following steps explain how to create a UAMI for Microsoft Purview to use.

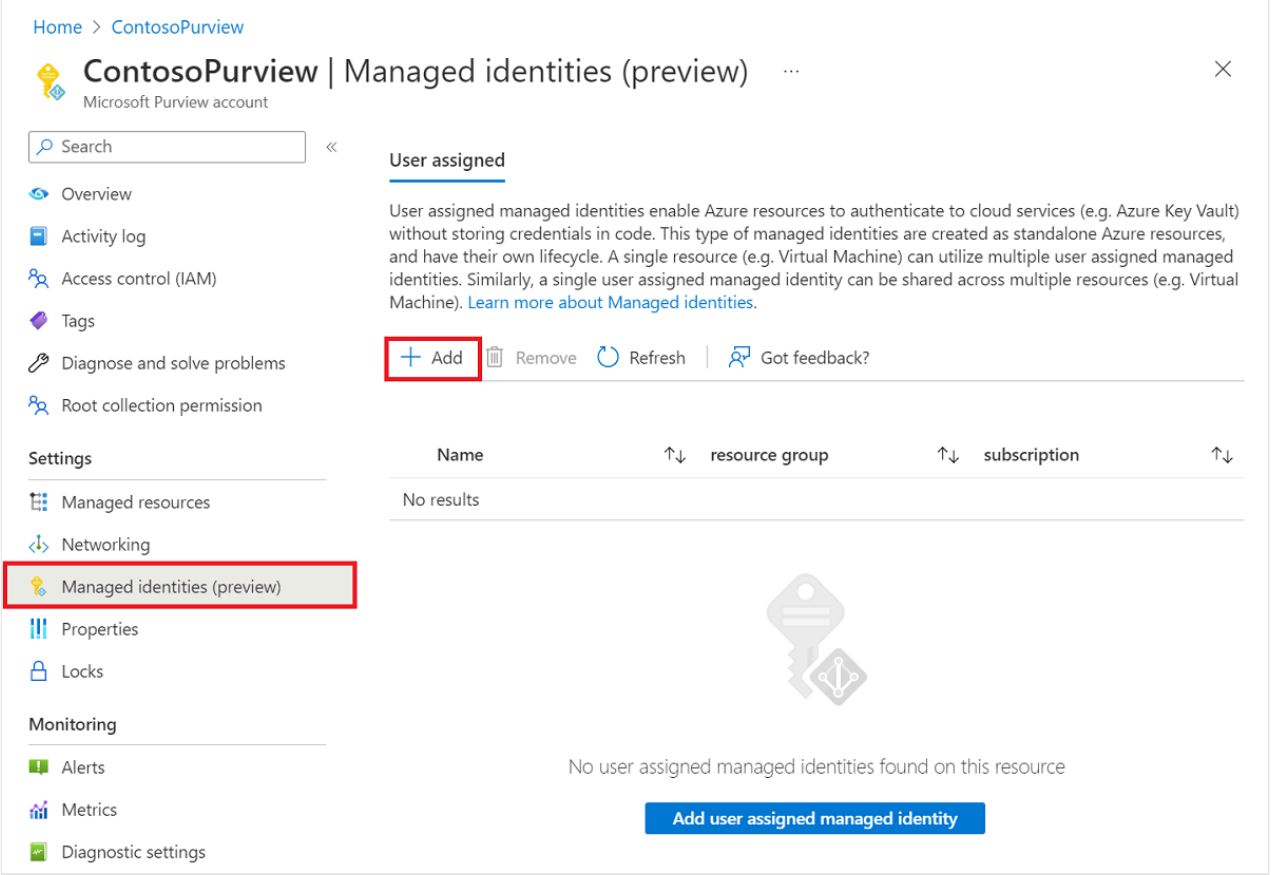
Supported data sources for UAMI

- Azure Data Lake Gen 1
- Azure Data Lake Gen 2
- Azure SQL Database
- Azure SQL Managed Instance
- Azure SQL Dedicated SQL pools
- Azure Blob Storage

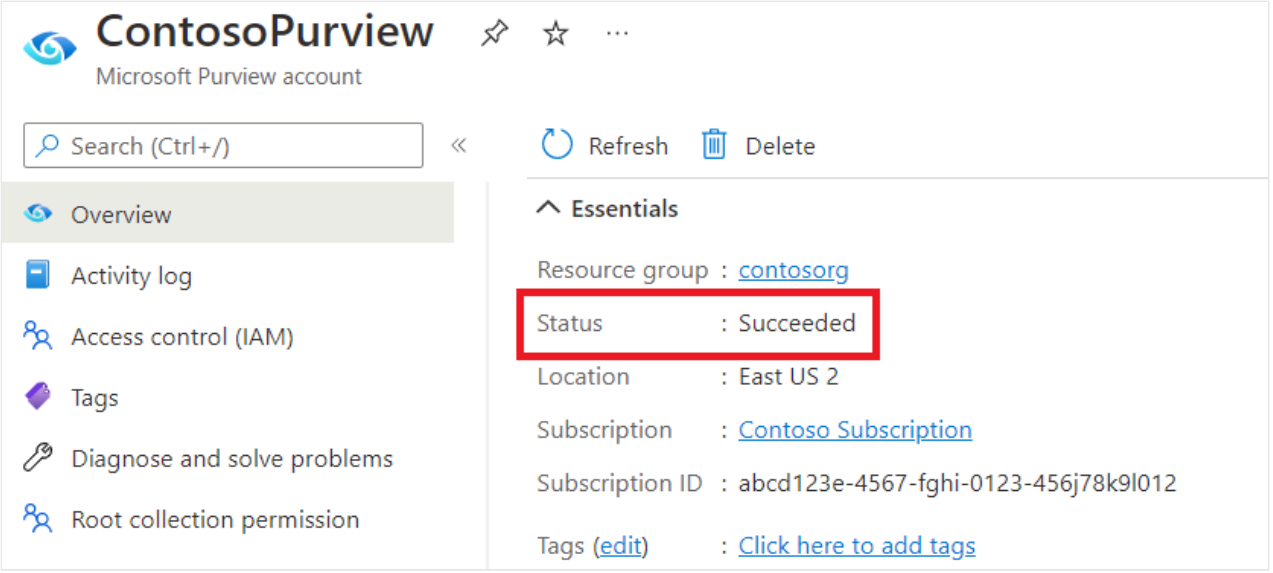
Create a user-assigned managed identity

1. In the Azure portal, navigate to your Microsoft Purview account.

2. In the **Managed identities** section on the left menu, select the **+ Add** button to add user assigned managed identities.



3. After finishing the setup, go back to your Microsoft Purview account in the Azure portal. If the managed identity is successfully deployed, you'll see the Microsoft Purview account's status as **Succeeded**.



4. Once the managed identity is successfully deployed, navigate to the classic Microsoft Purview governance portal by selecting **Open Microsoft Purview governance portal**.

5. In the classic Microsoft Purview governance portal, navigate to the Management Center in the studio and then navigate to the Credentials section.

6. Create a user-assigned managed identity by selecting **+New**.

7. Select the Managed identity authentication method from the **Authentication method** dropdown list, and make a selection from the **User assigned managed identities** dropdown list.

   > **Note**
   >
   > If the portal was open during creation of your user assigned managed identity, you need to refresh the Microsoft Purview portal to load the settings finished in the Azure portal.

8. When done filling in all the information, select **Create**.

Delete a user-assigned managed identity

1. In the Azure portal navigate to your Microsoft Purview account.

2. In the **Managed identities** section on the left menu, select the identity you want to delete.

3. Select the **Remove** button.

4. Once the managed identity is successfully removed, navigate to the classic Microsoft Purview governance portal by selecting **Open Microsoft Purview governance portal**.
5. Navigate to the Management Center in the studio and then navigate to the Credentials section.
6. Select the identity you want to delete, and then select the **Delete** button.

> **Note**
>
> If you have deleted a user-assigned managed identity in the Azure portal, you need to delete the original identity and create a new one in the Microsoft Purview portal.

Next steps

Create a scan rule set.

Create a service principal for use with Microsoft Purview

You can create a new or use an existing service principal in your Microsoft Entra tenant to authenticate Microsoft Purview with other services. This article explains how to prepare a service principal for Purview to use to authenticate with other services.

App registration

1. Follow the steps in this linked article section to create a service principal: Register an application with Microsoft Entra ID and create a service principal.

   > **Tip**
   >
   > For the Redirect URL, if you have an authentication endpoint for your organization you want to use, add it. Otherwise, `https://contoso.com/auth` works.

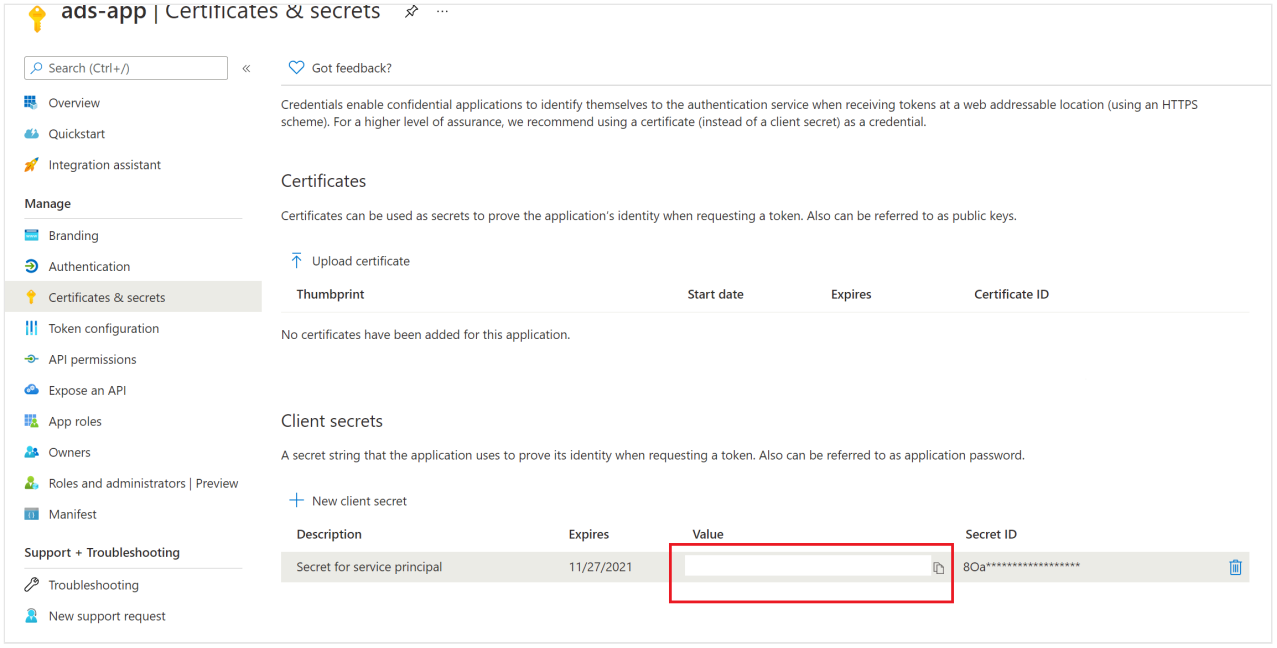
2. Once you register the app, copy the **Application (client) ID** value. Use this value later to create a credential in Microsoft Purview.



Add a secret to the client credentials

1. Follow the steps in this linked article section to create a new client secret: Create a new client secret.

2. Copy the **Secret value**. Use this value later to create a secret in Azure Key Vault (KV).

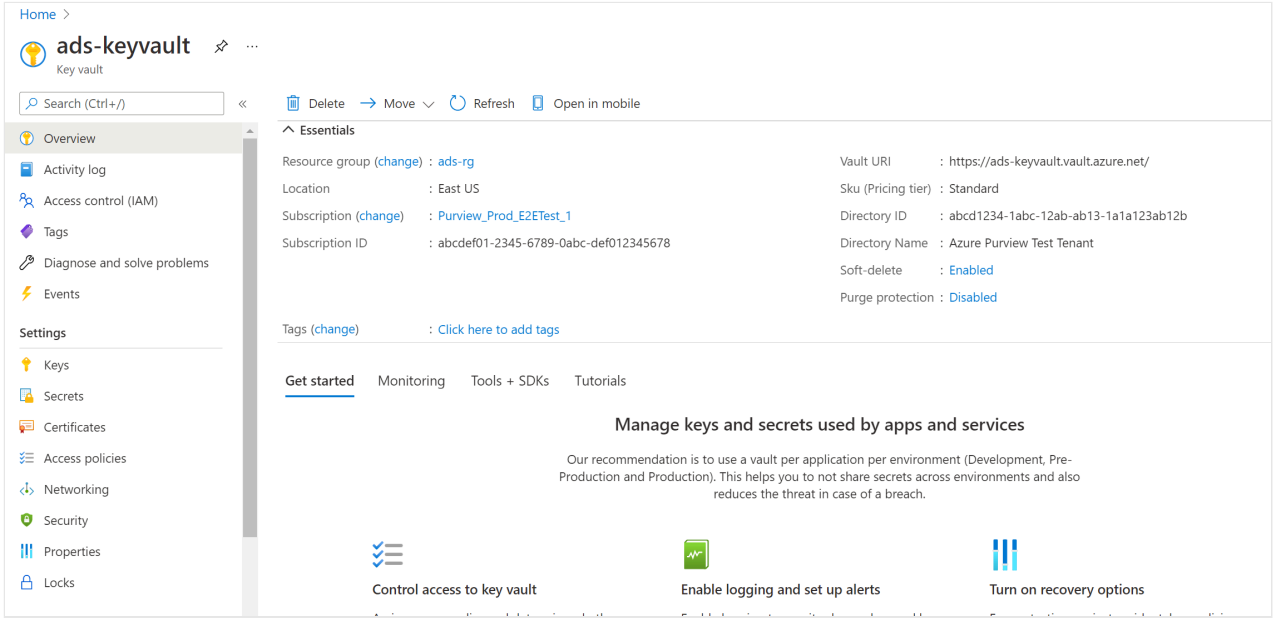


Add the secret to your Azure Key Vault

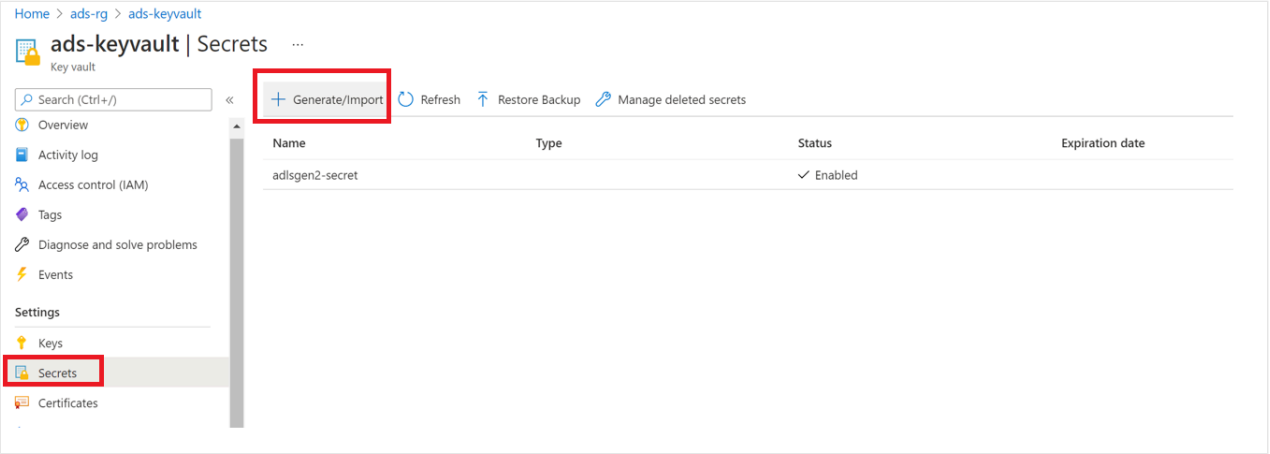
To allow Microsoft Purview to use this service principal to authenticate with other services, store this credential in Azure Key Vault.

- If you need an Azure Key vault, you can follow these steps to create one.
- To grant your Microsoft Purview account access to the Azure Key Vault, you can follow these steps.

1. Go to your **Key Vault**.

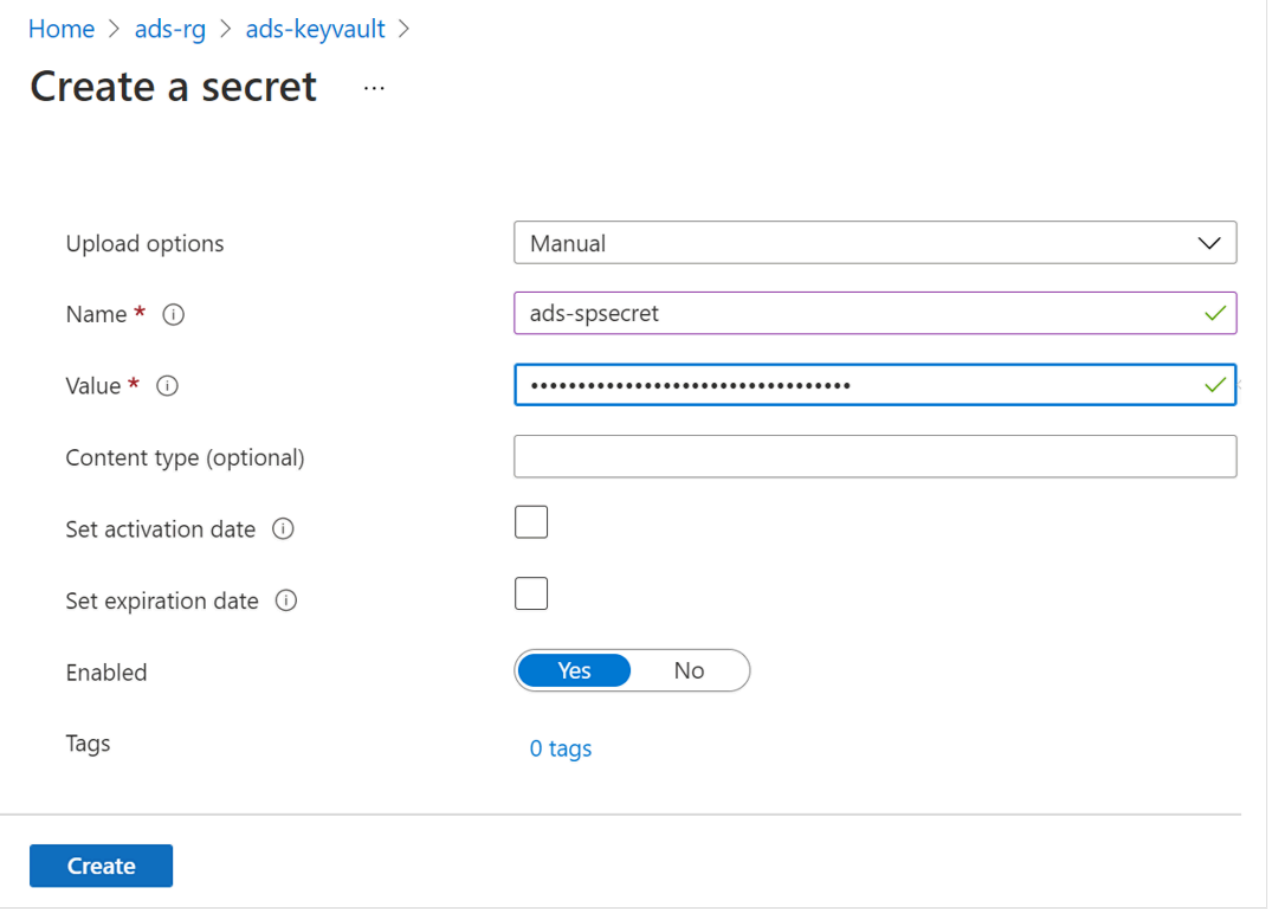


2. Select **Settings** > **Secrets** > **+ Generate/Import**.



3. Enter the **Name** of your choice, and save it to create a credential in Microsoft Purview.

4. Enter the **Value** as the **Secret value** from your Service Principal.



5. Select **Create** to complete.

Create a credential for your secret in Microsoft Purview

To enable Microsoft Purview to use this service principal to authenticate with other services, follow these three steps:

1. Connect your Azure Key Vault to Microsoft Purview
2. Grant your service principal authentication on your source. Every source is different, so select your specific source page from the list of supported sources and follow instructions to assign permissions.
3. Create a new credential in Microsoft Purview - Use the service principal's application (client) ID and the name of the secret you created in your Azure Key Vault.

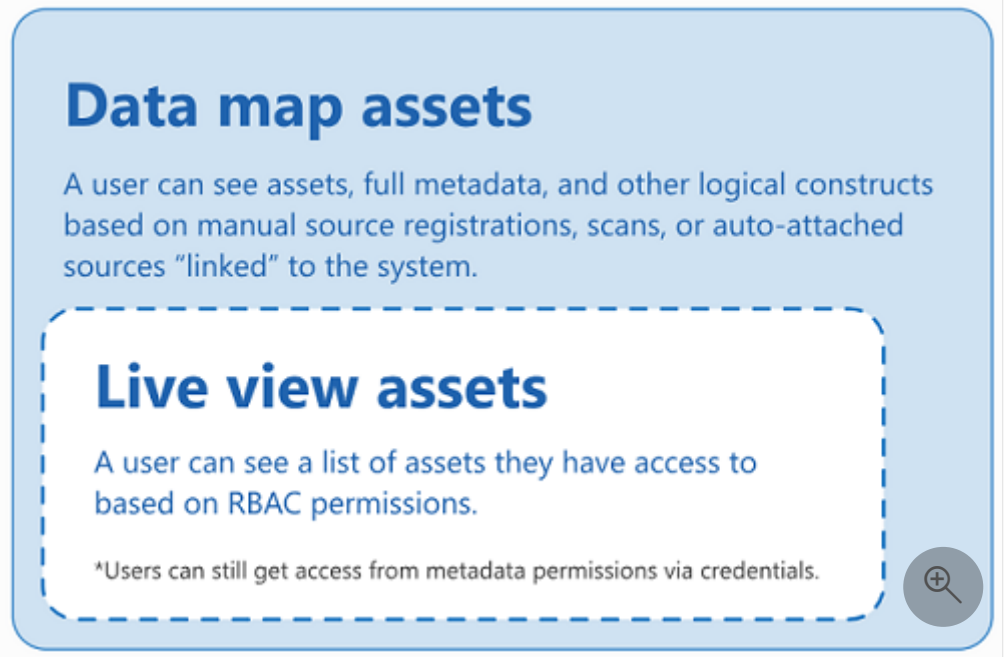
Last updated on 12/08/2025

Live view in Microsoft Purview (preview)

06/02/2025

> **Important**
>
> This feature is currently in preview. The Supplemental Terms of Use for Microsoft Azure Previews include additional legal terms that apply to Azure features that are in beta, in preview, or otherwise not yet released into general availability.

Some Azure resources are available to be seen in live view. Live view means that users with data access permissions can find these resources and their data assets in the catalog without setup or scanning. These resources and their metadata are immediately available when you access the free or enterprise versions of Microsoft Purview, so you can start your governance journey quickly.



In this article we'll cover:

- What resources have live view available
- What metadata is available in live view
- Who has permissions to live view resources
- Next steps to govern live view resources

Limitations

- Live view for Microsoft Fabric limitations:
  - Personal workspaces aren't supported yet in live view.

- Only Microsoft Fabric item level metadata are available in live view.
- Power BI Datasets don't support showing schema in live view if the item hasn't been scanned.
- Dataflow Gen2 (Preview) items are displayed as Dataflow items.

What resources have live view available?

Live view resources in the same tenant as your Microsoft Purview account are automatically accessible for users in your catalog. These resources have live view available:

Expand table

Resource	Level of view available
Azure Blob Storage	Storage account, Azure Blob Service, Container
Azure Data Lake Storage Gen 2	Storage account, Azure Data Lake Storage Gen2 Service, File system
Azure SQL Database	Server, Database, Schema, Tables, Columns
Azure Subscriptions	Live view assets in the subscription
Microsoft Fabric	Workspace, Item

Users with permissions can browse and search for these assets in the catalog without needing to set up a scan.

Metadata available in live view

Resources in live view in the catalog automatically have this metadata available:

- Name
- Properties
- Schema
- Lineage

> **Note**
>
> Live view resources won't automatically have classifications applied. To automatically apply classifications to these resources, you need to upgrade to the enterprise version of Microsoft Purview, register and scan the resource.

Data owners can also update the metadata on live view resources to add more information. Including manually adding description, contacts, etc.

Live view permissions

Users that have permissions on the available resources are automatically able to access these resources in the catalog as either data readers or data owners. No other permissions in Microsoft Purview are required.

Data readers

These are the permissions needed on the resources for a user to be able to see it as a reader in live view in Microsoft Purview.

Expand table

Data source	Reader permission
Azure SQL Database	Reader role in Azure
Azure Blob Storage	Reader role in Azure
Azure Data Lake Storage Gen2	Reader role in Azure
Azure Subscription	Read permission on subscription

Data owners

These are the permissions needed on the resources for a user to be able to see it as an owner and curate its metadata in live view in Microsoft Purview.

Expand table

Data source	Owner permission
Azure SQL Database	"Microsoft.Sql/servers/write", "Microsoft.Sql/servers/databases/write", "Microsoft.Authorization/roleAssignments/write"
Azure Blob Storage	"Microsoft.Storage/storageAccounts/write", "Microsoft.Authorization/roleAssignments/write"
Azure Data Lake Storage Gen2	"Microsoft.Storage/storageAccounts/write", "Microsoft.Authorization/roleAssignments/write"

Extend live view permissions

You can expand discoverability of these resources to users that don't have permissions already by enabling live view for all users. This can only be enabled with an enterprise version of Microsoft Purview. Extending live view also deepens the level of information available.

Expand table

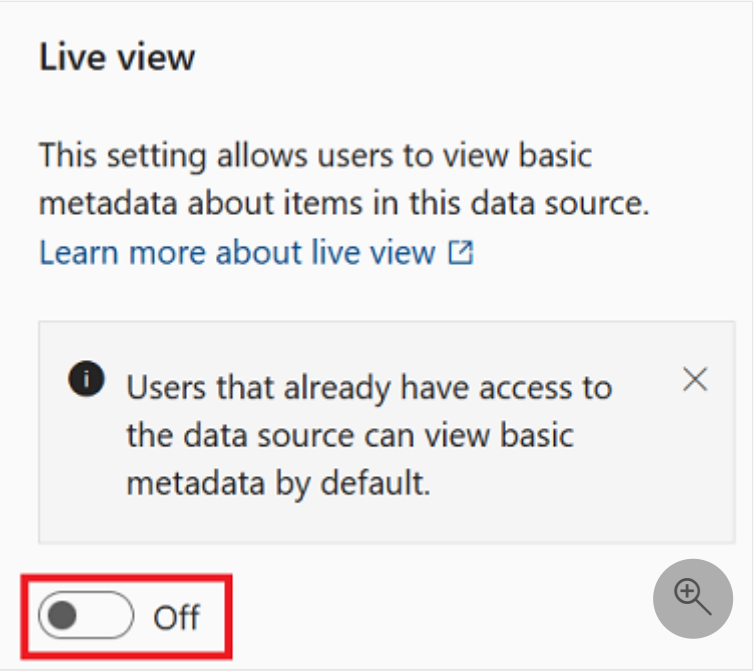
Resource	Level of view available
Azure Blob Storage	Storage account, Azure Blob Service, Container, Folders, Files
Azure Data Lake Storage Gen 2	Storage account, Azure Data Lake Storage Gen2 Service, File system, Folders, Files
Azure SQL Database	Server, Database, Schema, Tables, Columns, Views, Stored procedures

> **Note**
>
> - For Azure Blob storage and Azure Data Lake Storage Gen 2, file schema is provided for parquet files.
> - You can't extend live view permissions for Microsoft Fabric sources.

To enable these permissions, the user must be a Data Source Administrator and Reader on the collection where they're going to host the resource.

To enable:

1. Browse the catalog for the source where you want to expand live view to users who don't already have data access permissions.
2. Select the data source.
3. Switch the **Live view** toggle under **Source management** to enable it.



4. Add a credential for your resource. Select an existing credential or add a new one. Your available options depend on which resources you're using. For more information about each credential, follow these links:

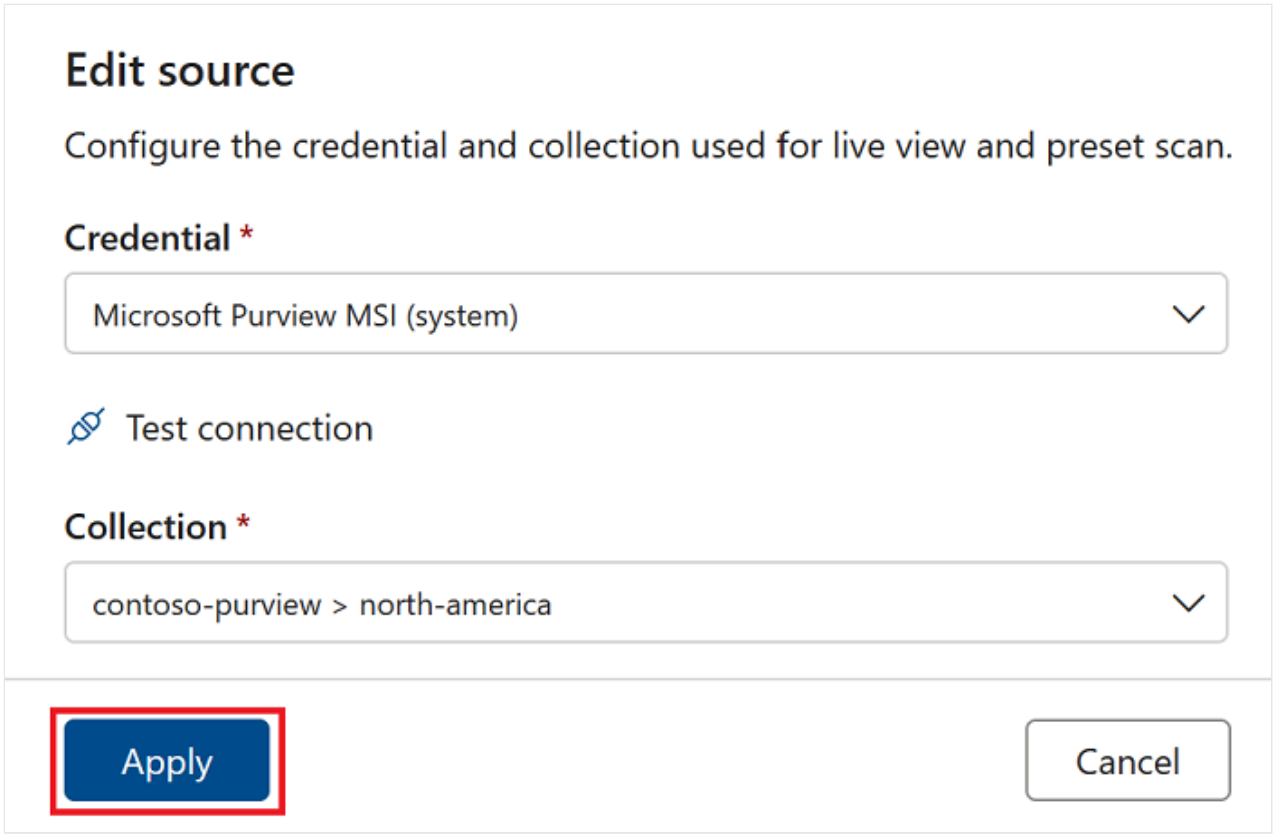
    - Azure Blob Storage credential
    - Azure Data Lake Gen 2 credential
    - Azure SQL Database credential

5. Select **Test connection** to test your credential. You can only extend live view with a successful connection. Follow this guide to troubleshoot connections.

6. Choose a collection for your resource.

    > **Tip**
    >
    > Follow this link for more information about permissions in collections.

7. Select **Apply** to register your source and extend your permissions.



Preset scan

> **Note**
>
> Microsoft Fabric resources don't currently support a preset scan.

Assets with a live view allow data source administrators to perform preset scans. A live view asset has only some of it's metadata automatically available to users. A preset scan gathers the rest of that metadata and makes it available in Microsoft Purview.

A preset scan can only be run in the enterprise version of Microsoft Purview, and runs a scan with these default options:

- Full scan - all data assets are scanned and updated
- Default classifications - Microsoft Purview will check for default classifications for the source
- Scheduled - the scan is scheduled to run weekly by default, but the schedule can be updated.

> **Tip**

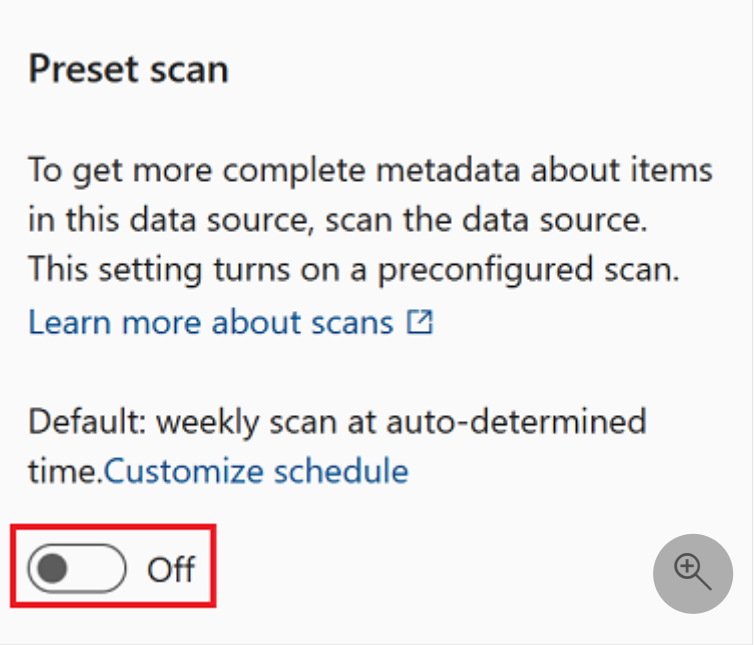
To run a scan with any other options available to your source and organization, like custom classifications, perform a regular scan.

To run a preset scan, a user must be a Data Source Administrator on the collection where they're going to host the resource.

1. Browse the catalog for the source where you want to run a preset scan to enhance the metadata.

2. Select the data source.

3. Select the **Preset scan** toggle under **Source management** to scan.

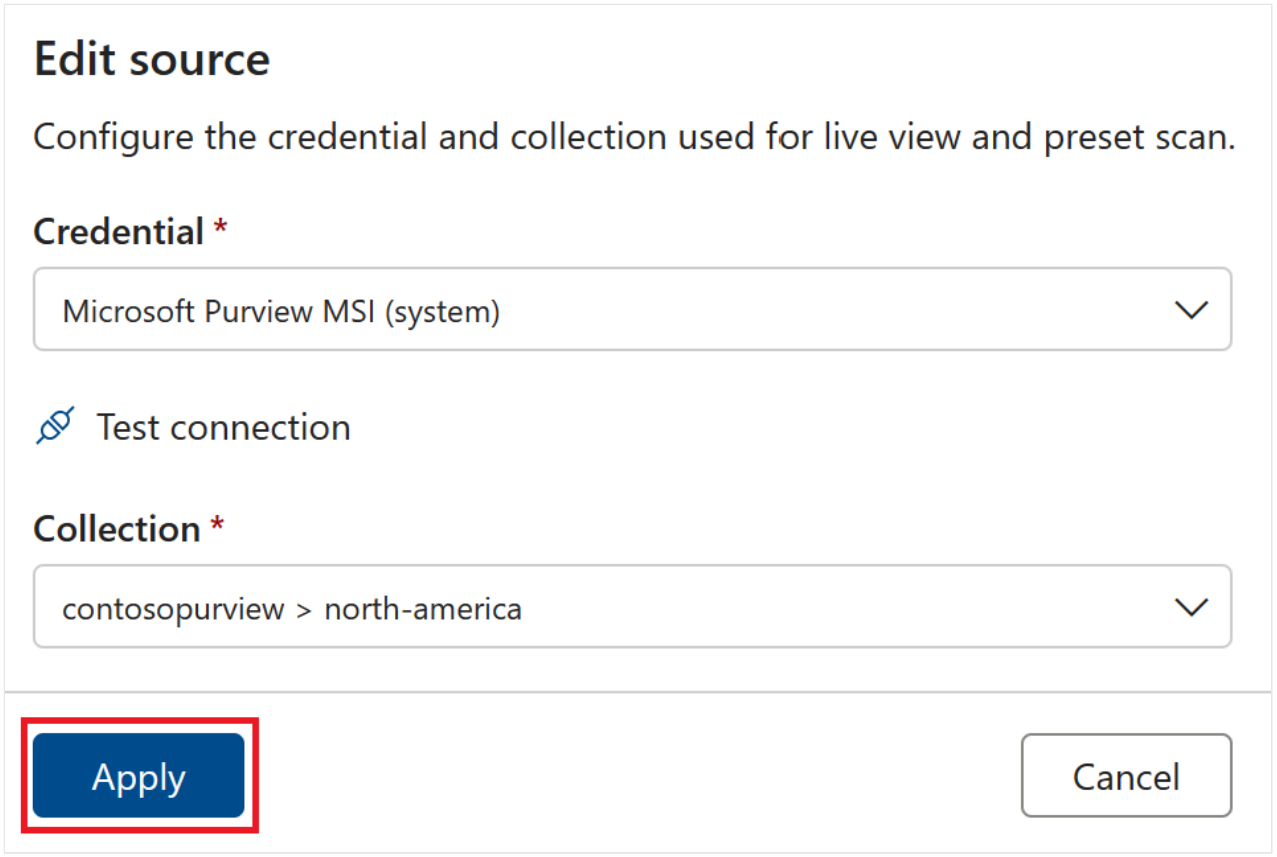


4. If live view permissions have already been extended, the scan starts. If live view hasn't been enabled, you need to submit a credential and select a collection to house the metadata:

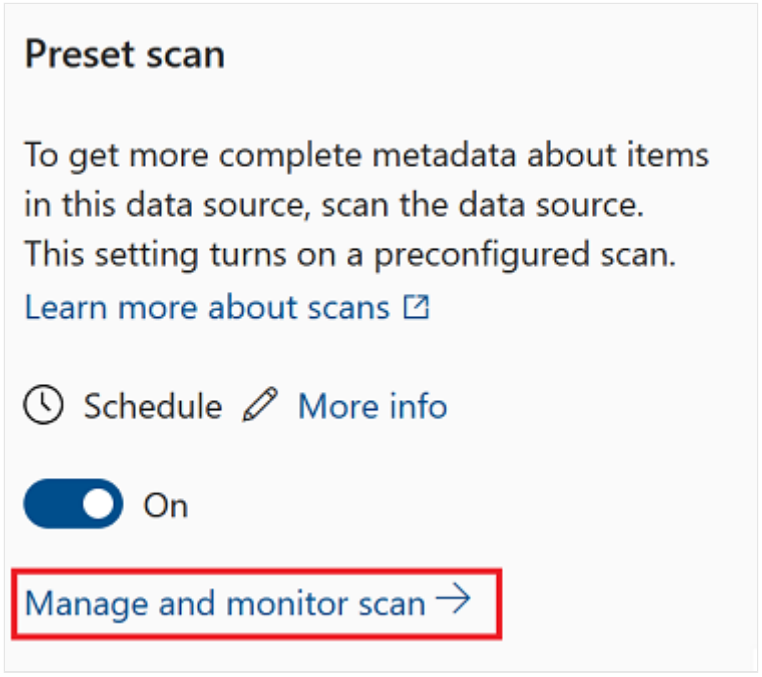
    - Azure Blob Storage credential
    - Azure Data Lake Gen 2 credential
    - Azure SQL Database credential

5. Select **Test connection** to test your credential. You can only scan with a successful connection. Follow this guide to troubleshoot connections.

6. Select **Apply** to run your scan.



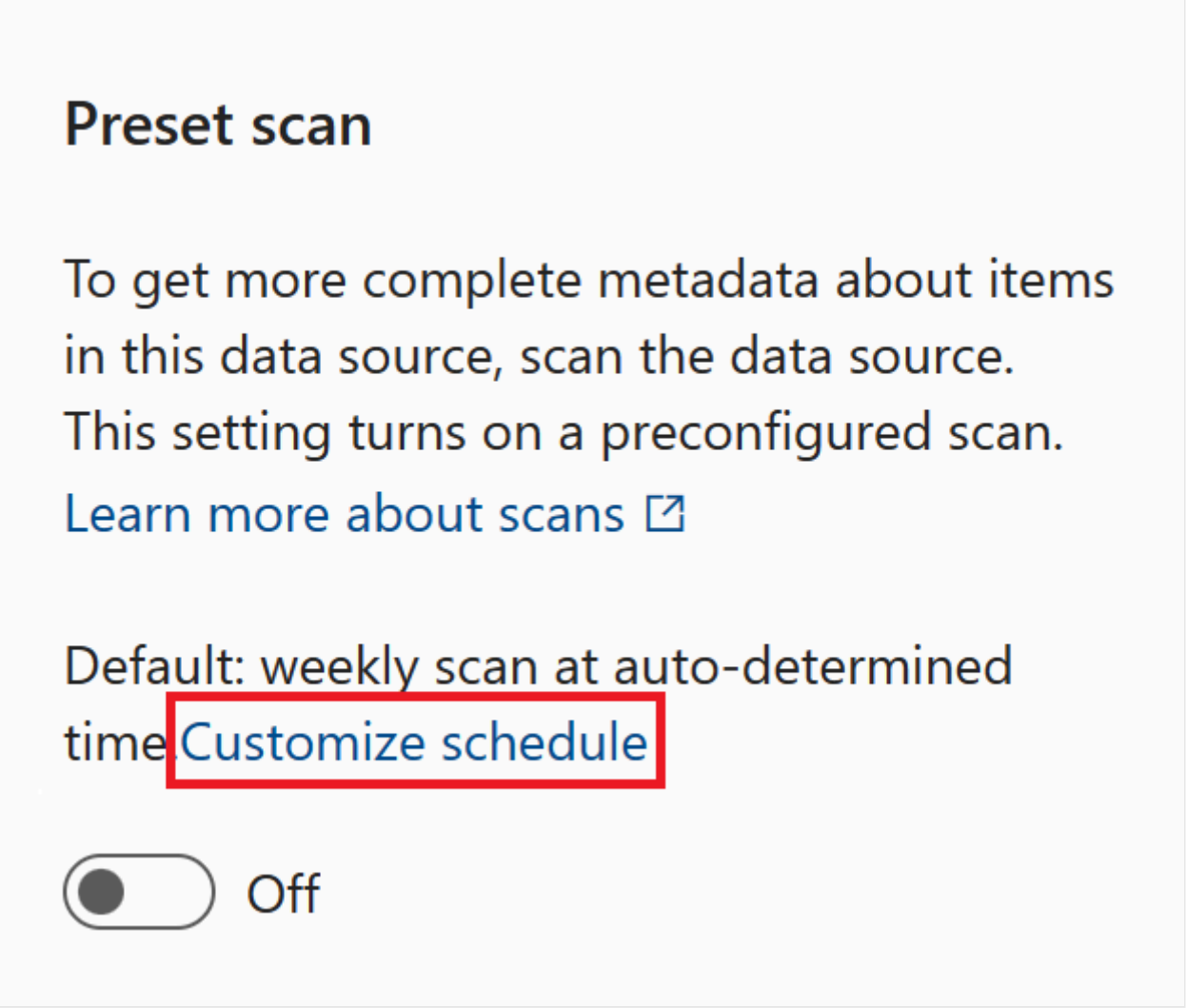
7. To monitor the scan, select **Manage scans** in the **Source management** window.



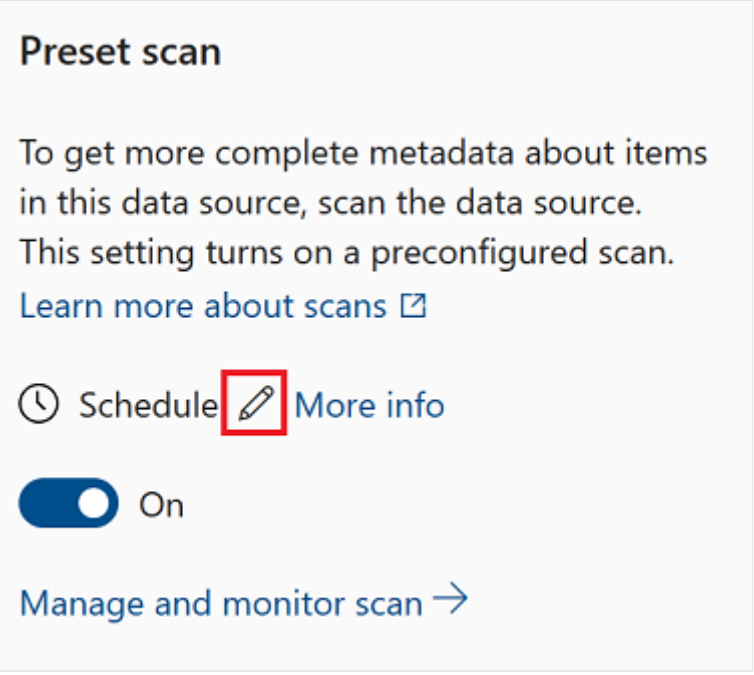
Update schedule

To update the schedule of a preset scan either:

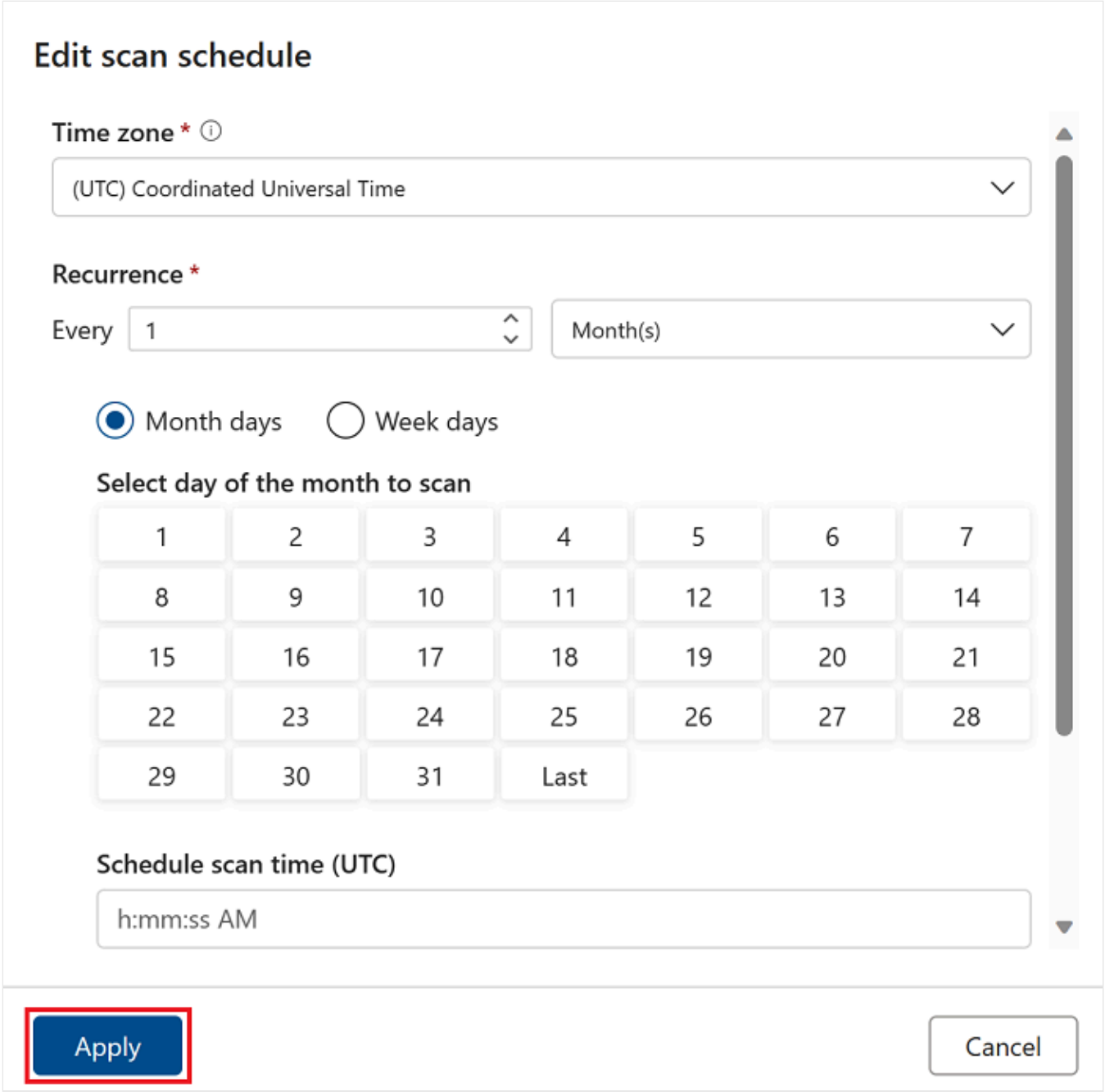
- Select **Customize schedule** before selecting the toggle. You'll set your scan schedule and then the preset scan will be automatically enabled.



- Select the **Edit pencil** next to the schedule label after enabling the preset scan



Then you can set your schedule to run monthly or weekly on specified days. Select **Apply** to apply the schedule to the scan.



Manage live view credentials

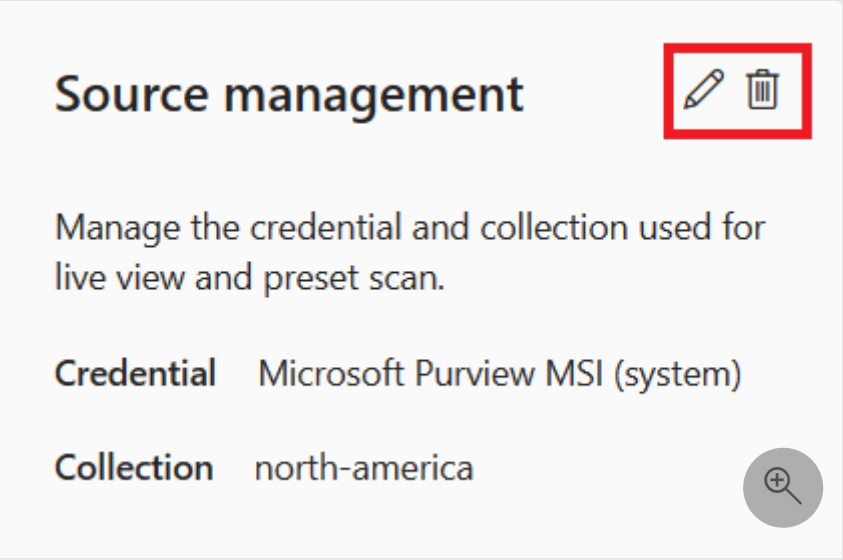
If you have extended live view permissions or run a preset scan your live view resource will have credentials that allow Microsoft Purview to access it.

If you need to update or remove those credentials, you can use the **Source management** window.

1. Browse the catalog for the source where you want to run a preset scan to enhance the metadata.

2. Select the data source.

3. Under source management, you can see your credential and collection listed, and the buttons to edit or delete the credentials.



4. Select the **Edit** pencil icon to update the credentials.

> ⓘ **Note**
>
> You can't change the collection the source is mapped to.

5. Select **Delete** to remove the credentials.

> ⓘ **Note**
>
> Deleting credentials will also disable extended live view permissions or preset scans.

Enterprise governance of your data estate

In free version of governance solutions, users can only view live resources and manually update some metadata.

For access to all Microsoft Purview's applications, sources, automation, and to fully govern data assets in Microsoft Purview, upgrade to the enterprise version of Microsoft Purview.

Next steps

- Metadata curation
- Use the new Microsoft Purview portal
- Roles and permissions

Manage data sources in Microsoft Purview

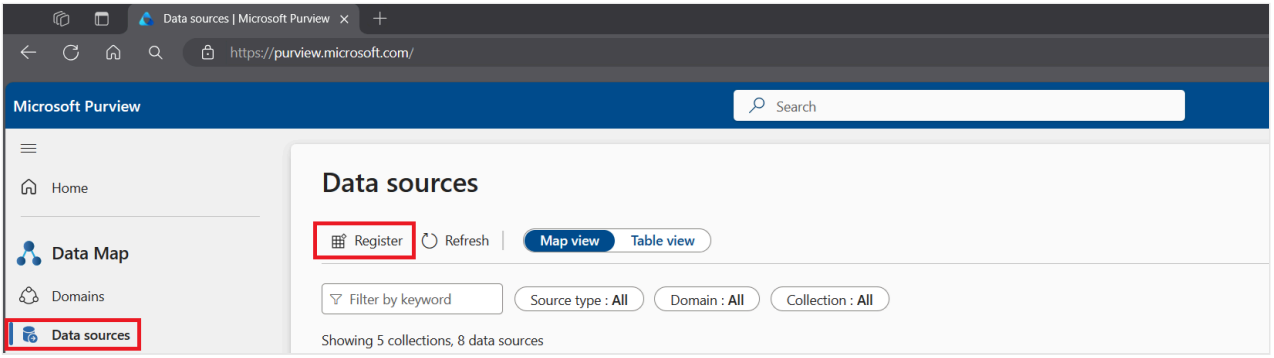
In this article, you learn how to register new data sources, manage domains and collections of data sources, view sources, and move sources between collections in the Microsoft Purview Data Map.

Register a new source

> **Note**
>
> To register and manage a source in Data Map, you need to be a Data Source Admin and have one of the other Data Map roles, such as Data Reader. For more information, see assigning data governance roles and permissions.

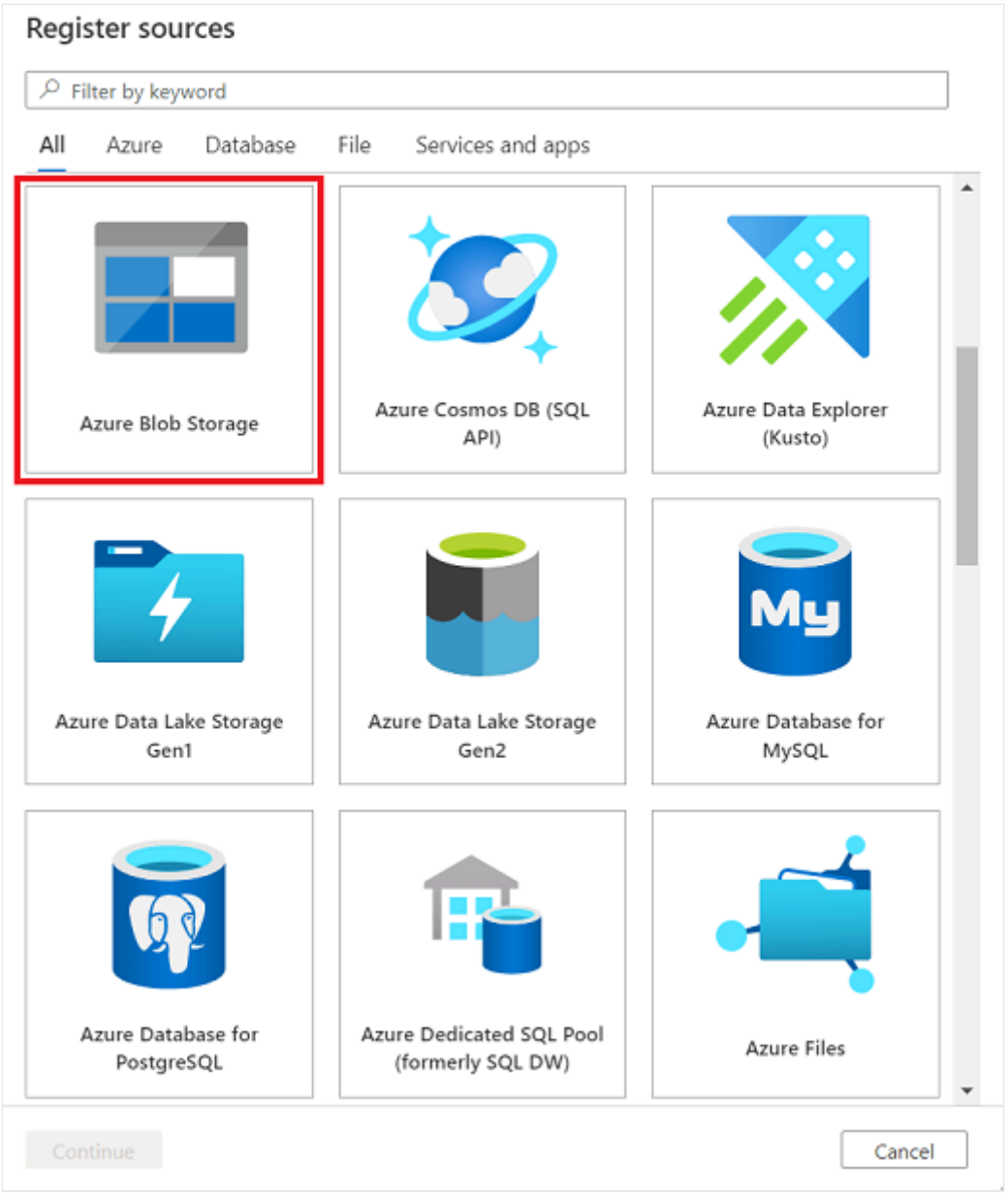
Use the following steps to register a new source:

1. In the Microsoft Purview portal, navigate to **Data Map**, **Data sources**, and select **Register**.



2. If you're using the classic experience, open the classic Microsoft Purview governance portal, go to **Data Map**, **Sources**, and select **Register**.

3. Select a source type. This example uses Azure Blob Storage. Select **Continue**.



4. Fill out the form on the **Register sources** page. Select a name for your source and enter the relevant information. If you choose **From Azure subscription** as your account selection method, the sources in your subscription appear in a dropdown list.

5. Select **Register**.

> **Important**
>
> Most data sources have prerequisites to register and scan them in Microsoft Purview. For a list of all available sources, and links to source-specific instructions to register and scan, see Data sources that connect to Data Map.

View sources

You can view all registered sources by going to the Microsoft Purview portal > **Data Map** > **Data sources**.

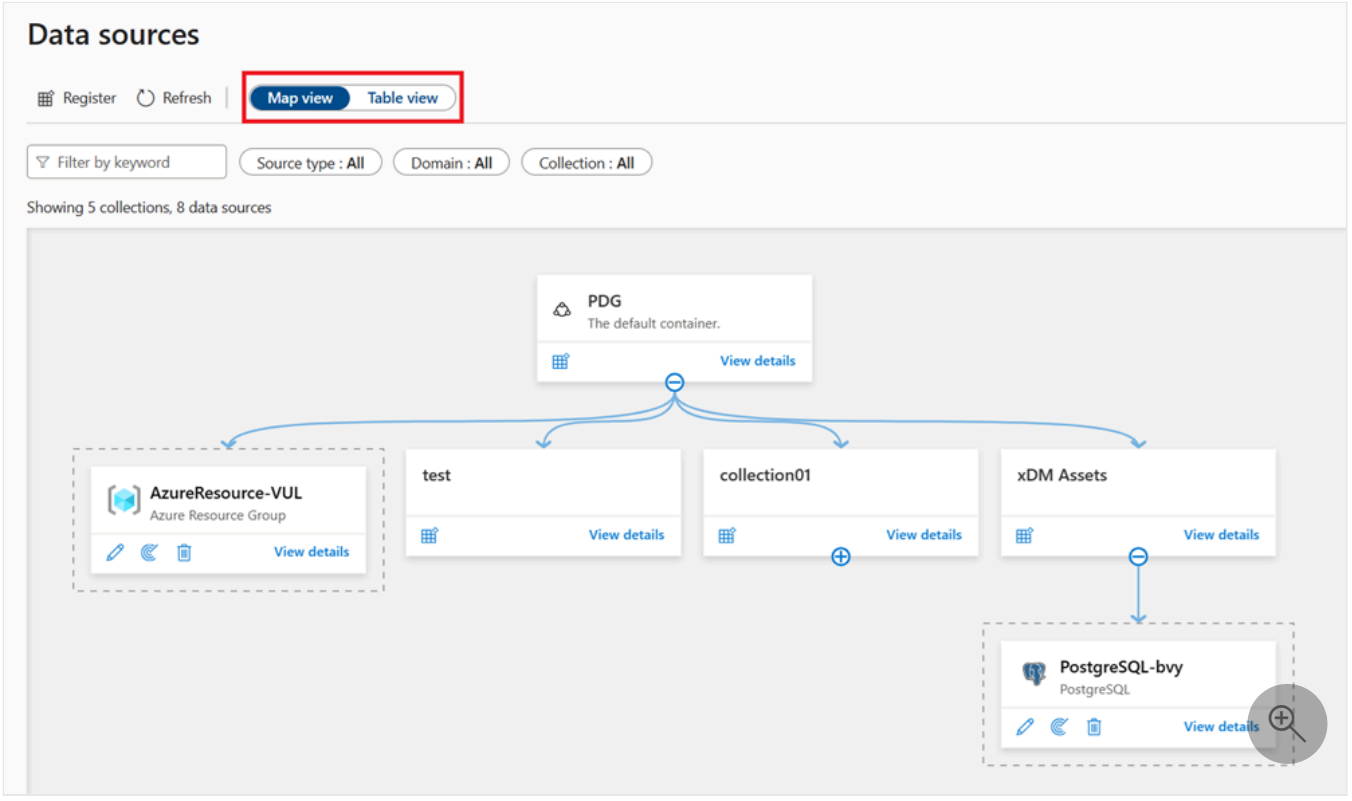
There are two view types:

- The map view
- The table view

Map view

To find the map view, open Data Map and select **Data sources**. Make sure the toggle at the top of the page is set to **Map view**.

In **Map view**, you see all of your sources and collections, and you can filter by sources, domains, and collections. In the following screenshot, you see the default domain at the top called PDG. It has three collections: test, collection01, and xDM Assets. The collection xDM Assets has a PostgreSQL source registered to it. An Azure resource group source is also housed under the root domain.



You can open and manage each of the subcollections from the map view by selecting the + button. You can also register a new source by selecting the register source button, or view details by selecting **View details**.

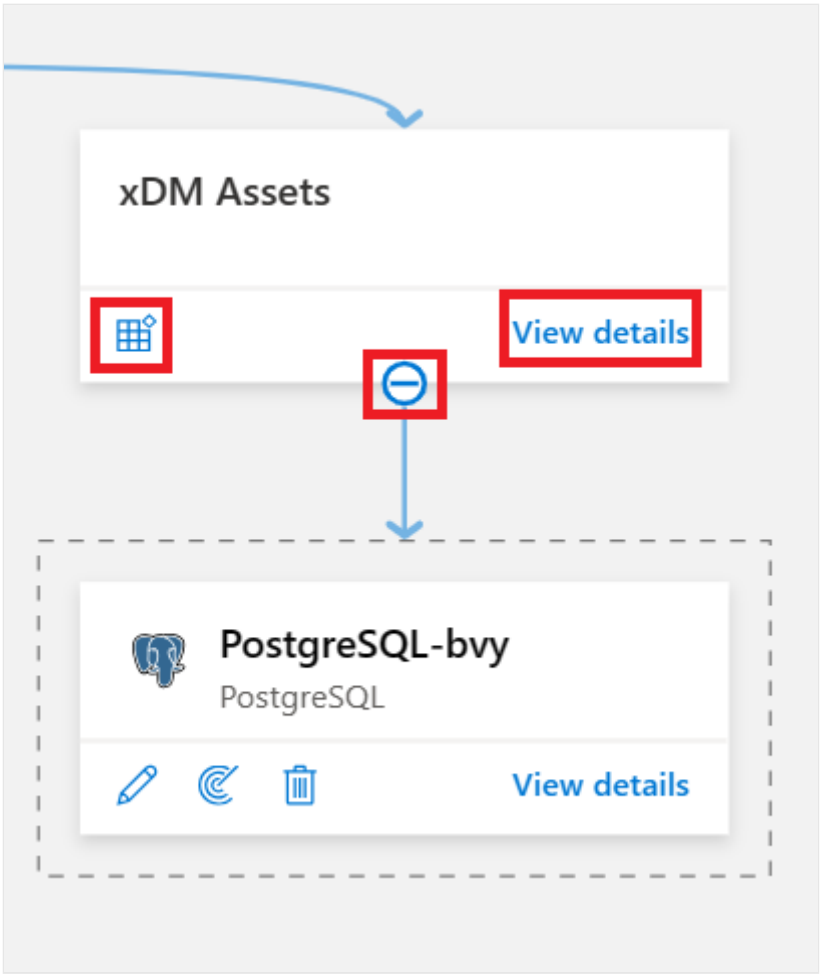
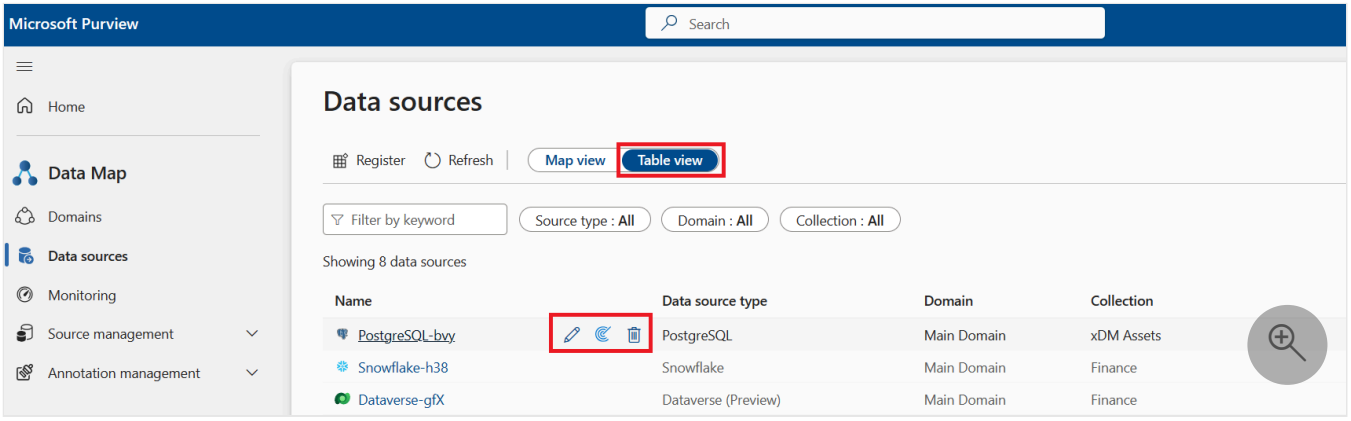


Table view

In the table view, you see a sortable list of sources. Hover over the source for options to edit, begin a new scan, or delete.



Manage collections and domains

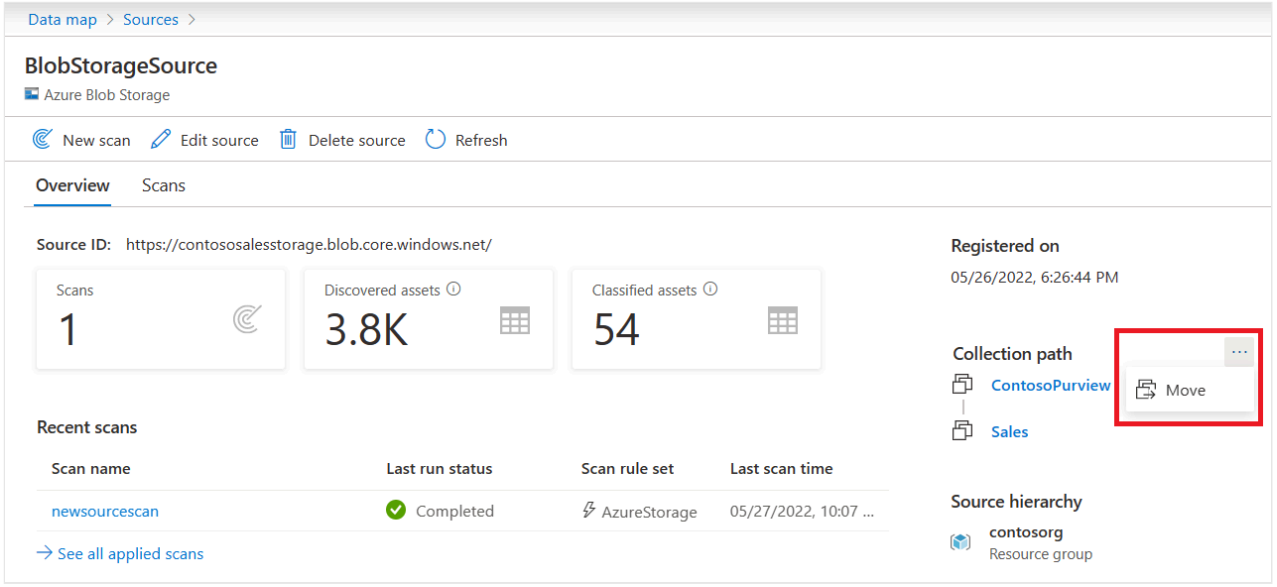
To manage collections and domains, and assign permissions based on your collections and domains, see Create and manage collections and domains.

Move sources between collections

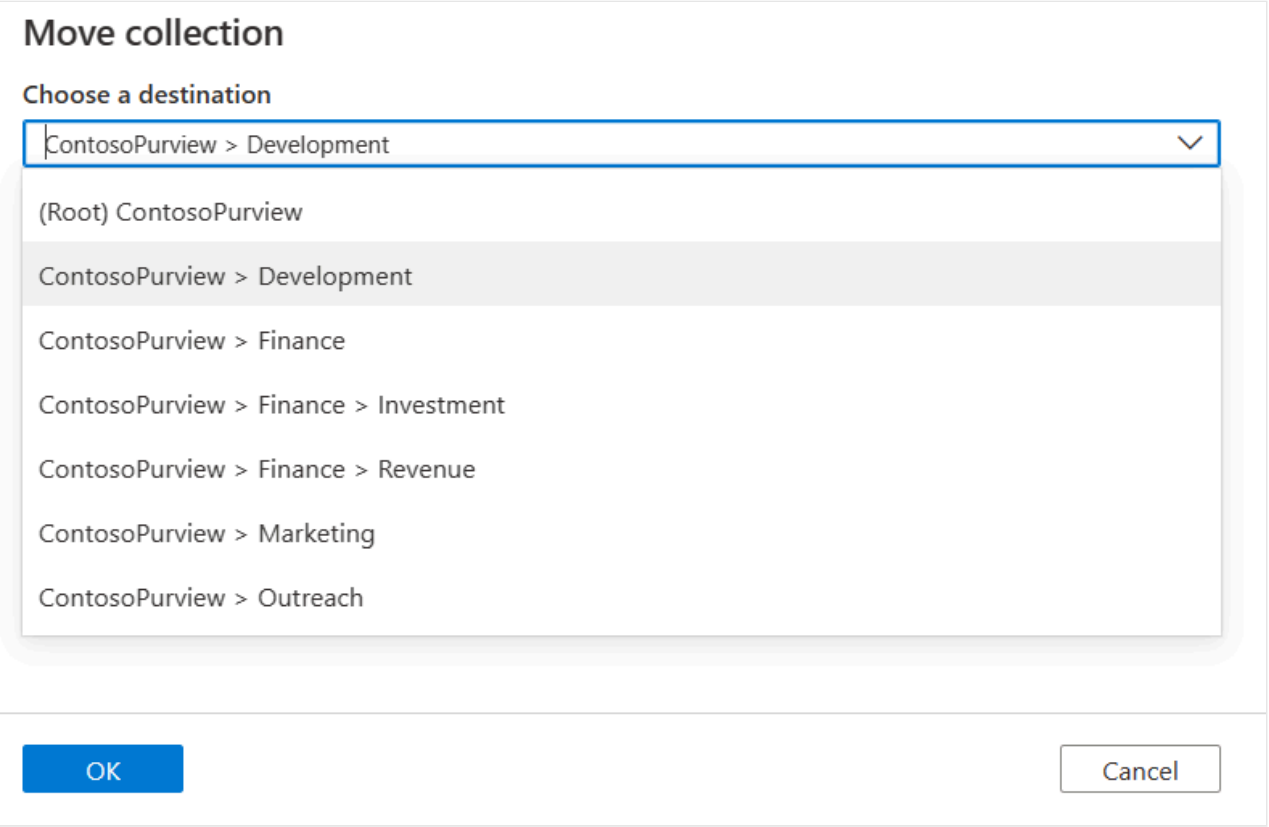
After you register your source, you can move it to another collection that you have access to within the same domain.

> **Important**
>
> Data source connections that register multiple sources can't be moved between collections:
>
> - Azure (Multiple)
> - AWS account (Multiple)
> - Azure Synapse Analytics (Multiple)

1. Find your source in Data Map and select it.
2. Beside the **Collection Path** list, select the ellipsis (...) button and select **Move**.



3. In the **Move collection** menu that appears, select your collection from the drop-down and then select **Ok**.



4. Your data source has been moved. It can take up to an hour for results to be fully seen across your Microsoft Purview environment. Your scans will move with your resource, but assets will remain in their original collection until your next scan, then they'll move to the new collection.

> **Note**
>
> - When a source moves to a new collection, its scans move with it, but assets don't appear in the new collection until your next scan.
> - If you manually moved any assets from your source to a different collection before migrating the source, the scan doesn't take them to the new collection. They remain in the collection you moved them to.

Next steps

Learn how to discover and govern various data sources:

- Azure Data Lake Storage Gen 2
- Power BI tenant
- Azure SQL Database

Last updated on 12/08/2025

Check Azure data source readiness to register and scan in Microsoft Purview

06/03/2025

To scan data sources, Microsoft Purview requires access to them. It uses credentials to obtain this access. A *credential* is the authentication information that Microsoft Purview can use to authenticate to your registered data sources. There are a few ways to set up the credentials for Microsoft Purview, including:

- The managed identity assigned to the Microsoft Purview account.
- Secrets stored in Azure Key Vault.
- Service principals.

In this two-part tutorial series, we'll help you verify and configure required Azure role assignments and network access for various Azure data sources across your Azure subscriptions at scale. You can then register and scan your Azure data sources in Microsoft Purview.

Run the Microsoft Purview data sources readiness checklist script after you deploy your Microsoft Purview account and before you register and scan your Azure data sources.

In part 1 of this tutorial series, you'll:

- ✓ Locate your data sources and prepare a list of data source subscriptions.
- ✓ Run the readiness checklist script to find any missing role-based access control (RBAC) or network configurations across your data sources in Azure.
- ✓ In the output report, review missing network configurations and role assignments required by Microsoft Purview Managed Identity (MSI).
- ✓ Share the report with data Azure subscription owners so they can take suggested actions.

Prerequisites

- Azure subscriptions where your data sources are located. If you don't have an Azure subscription, create a free account before you begin.
- A Microsoft Purview account.
- An Azure Key Vault resource in each subscription that has data sources like Azure SQL Database, Azure Synapse Analytics, or Azure SQL Managed Instance.
- The Microsoft Purview data sources readiness checklist script.

> ⓘ **Note**

The Microsoft Purview data sources readiness checklist is available only for Windows. This readiness checklist script is currently supported for Microsoft Purview MSI.

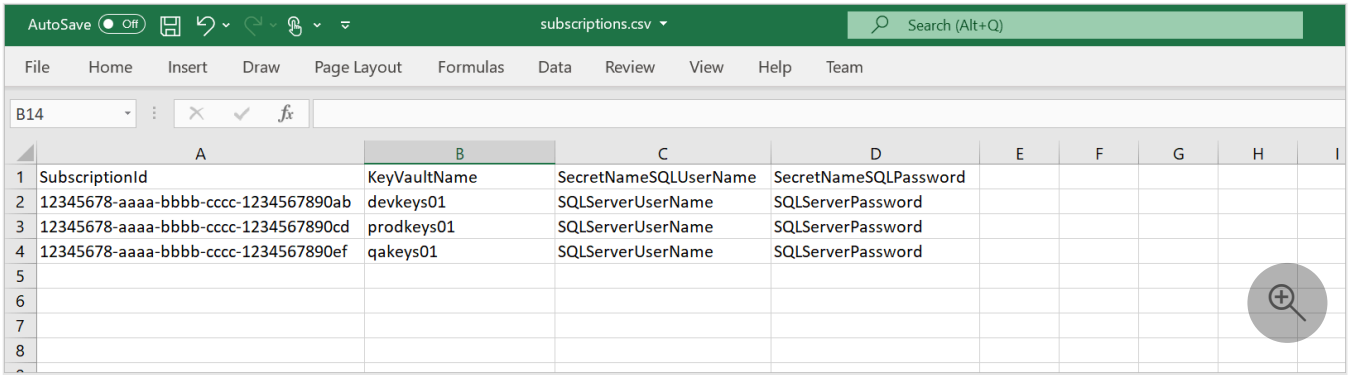
Prepare Azure subscriptions list for data sources

Before running the script, create a .csv file (for example, C:\temp\Subscriptions.csv) with four columns:

Expand table

Column name	Description	Example
`SubscriptionId`	Azure subscription IDs for your data sources.	12345678-aaaa-bbbb-cccc-1234567890ab
`KeyVaultName`	Name of existing key vault that's deployed in the data source subscription.	ContosoDevKeyVault
`SecretNameSQLUserName`	Name of an existing Azure Key Vault secret that contains a Microsoft Entra user name that can sign in to Azure Synapse, Azure SQL Database, or Azure SQL Managed Instance by using Microsoft Entra authentication.	ContosoDevSQLAdmin
`SecretNameSQLPassword`	Name of an existing Azure Key Vault secret that contains a Microsoft Entra user password that can sign in to Azure Synapse, Azure SQL Database, or Azure SQL Managed Instance by using Microsoft Entra authentication.	ContosoDevSQLPassword

Sample .csv file:



> **Note**
>
> You can update the file name and path in the code, if you need to.

Run the script and install the required PowerShell modules

Follow these steps to run the script from your Windows computer:

1. Download the Microsoft Purview data sources readiness checklist script to the location of your choice.
2. On your computer, enter **PowerShell** in the search box on the Windows taskbar. In the search list, select and hold (or right-click) **Windows PowerShell** and then select **Run as administrator**.
3. In the PowerShell window, enter the following command. (Replace `<path-to-script>` with the folder path of the extracted script file.)

   ```PowerShell
   dir -Path <path-to-script> | Unblock-File
   ```

4. Enter the following command to install the Azure cmdlets:

   ```PowerShell
   Install-Module -Name Az -AllowClobber -Scope CurrentUser
   ```

5. If you see the prompt *NuGet provider is required to continue*, enter **Y**, and then select **Enter**.
6. If you see the prompt *Untrusted repository*, enter **A**, and then select **Enter**.
7. Repeat the previous steps to install the `Az.Synapse` and `AzureAD` modules.

It might take up to a minute for PowerShell to install the required modules.

Collect other data needed to run the script

Before you run the PowerShell script to verify the readiness of data source subscriptions, obtain the values of the following arguments to use in the scripts:

- `AzureDataType`: Choose any of the following options as your data-source type to check the readiness for the data type across your subscriptions:
  - `BlobStorage`

- `AzureSQLMI`
- `AzureSQLDB`
- `ADLSGen2`
- `ADLSGen1`
- `Synapse`
- `All`

- `PurviewAccount`: Your existing Microsoft Purview account resource name.
- `PurviewSub`: Subscription ID where the Microsoft Purview account is deployed.

Verify your permissions

Make sure your user has the following roles and permissions:

Expand table

Role or permission	Scope
Global Reader	Microsoft Entra tenant
Reader	Azure subscriptions where your Azure data sources are located
Reader	Subscription where your Microsoft Purview account was created
SQL Admin (Microsoft Entra authentication)	Azure Synapse dedicated pools, Azure SQL Database instances, Azure SQL managed instances
Access to your Azure key vault	Access to get/list key vault's secret or Azure Key Vault secret user

Run the client-side readiness script

Run the script by completing these steps:

1. Use the following command to go to the script's folder. Replace `<path-to-script>` with the folder path of the extracted file.

```
PowerShell

cd <path-to-script>
```

2. Run the following command to set the execution policy for the local computer. Enter **A** for *Yes to All* when you're prompted to change the execution policy.

   ```PowerShell
   Set-ExecutionPolicy -ExecutionPolicy Unrestricted
   ```

3. Run the script with the following parameters. Replace the `DataType`, `PurviewName`, and `SubscriptionID` placeholders.

   ```PowerShell
   .\purview-data-sources-readiness-checklist.ps1 -AzureDataType <DataType> -PurviewAccount <PurviewName> -PurviewSub <SubscriptionID>
   ```

   When you run the command, a pop-up window might appear twice prompting you to sign in to Azure and Microsoft Entra ID by using your Microsoft Entra credentials.

It can take several minutes to create the report, depending on the number of Azure subscriptions and resources in the environment.

After the process completes, review the output report, which demonstrates the detected missing configurations in your Azure subscriptions or resources. The results can appear as *Passed*, *Not Passed*, or *Awareness*. You can share the results with the corresponding subscription admins in your organization so they can configure the required settings.

More information

What data sources are supported by the script?

Currently, the following data sources are supported by the script:

- Azure Blob Storage (BlobStorage)
- Azure Data Lake Storage Gen2 (ADLSGen2)
- Azure Data Lake Storage Gen1 (ADLSGen1)
- Azure SQL Database (AzureSQLDB)
- Azure SQL Managed Instance (AzureSQLMI)
- Azure Synapse (Synapse) dedicated pool

You can choose all or any of these data sources as the input parameter when you run the script.

What checks are included in the results?

Azure Blob Storage (BlobStorage)

- RBAC. Check whether Microsoft Purview MSI is assigned the **Storage Blob Data Reader** role in each of the subscriptions below the selected scope.
- RBAC. Check whether Microsoft Purview MSI is assigned the **Reader** role on the selected scope.
- Service endpoint. Check whether service endpoint is on, and check whether **Allow trusted Microsoft services to access this storage account** is enabled.
- Networking: Check whether private endpoint is created for storage and enabled for Blob Storage.

Azure Data Lake Storage Gen2 (ADLSGen2)

- RBAC. Check whether Microsoft Purview MSI is assigned the **Storage Blob Data Reader** role in each of the subscriptions below the selected scope.
- RBAC. Check whether Microsoft Purview MSI is assigned the **Reader** role on the selected scope.
- Service endpoint. Check whether service endpoint is on, and check whether **Allow trusted Microsoft services to access this storage account** is enabled.
- Networking: Check whether private endpoint is created for storage and enabled for Blob Storage.

Azure Data Lake Storage Gen1 (ADLSGen1)

- Networking. Check whether service endpoint is on, and check whether **Allow all Azure services to access this Data Lake Storage Gen1 account** is enabled.
- Permissions. Check whether Microsoft Purview MSI has Read/Execute permissions.

Azure SQL Database (AzureSQLDB)

- SQL Server instances:
  - Network. Check whether public endpoint or private endpoint is enabled.
  - Firewall. Check whether **Allow Azure services and resources to access this server** is enabled.
  - Microsoft Entra administration. Check whether Azure SQL Server has Microsoft Entra authentication.
  - Microsoft Entra administration. Populate the Azure SQL Server Microsoft Entra admin user or group.

- SQL databases:
  - SQL role. Check whether Microsoft Purview MSI is assigned the **db_datareader** role.

Azure SQL Managed Instance (AzureSQLMI)

- SQL Managed Instance servers:
  - Network. Check whether public endpoint or private endpoint is enabled.
  - ProxyOverride. Check whether Azure SQL Managed Instance is configured as Proxy or Redirect.
  - Networking. Check whether NSG has an inbound rule to allow AzureCloud over required ports:
    - Redirect: 1433 and 11000-11999
      or
    - Proxy: 3342
  - Microsoft Entra administration. Check whether Azure SQL Server has Microsoft Entra authentication.
  - Microsoft Entra administration. Populate the Azure SQL Server Microsoft Entra admin user or group.

- SQL databases:
  - SQL role. Check whether Microsoft Purview MSI is assigned the **db_datareader** role.

Azure Synapse (Synapse) dedicated pool

- RBAC. Check whether Microsoft Purview MSI is assigned the **Storage Blob Data Reader** role in each of the subscriptions below the selected scope.

- RBAC. Check whether Microsoft Purview MSI is assigned the **Reader** role on the selected scope.

- SQL Server instances (dedicated pools):
  - Network: Check whether public endpoint or private endpoint is enabled.
  - Firewall: Check whether **Allow Azure services and resources to access this server** is enabled.
  - Microsoft Entra administration: Check whether Azure SQL Server has Microsoft Entra authentication.
  - Microsoft Entra administration: Populate the Azure SQL Server Microsoft Entra admin user or group.

- SQL databases:
  - SQL role. Check whether Microsoft Purview MSI is assigned the **db_datareader** role.

Next steps

In this tutorial, you learned how to:

- ✓ Run the Microsoft Purview readiness checklist to check, at scale, whether your Azure subscriptions are missing configuration, before you register and scan them in Microsoft Purview.

Go to the next tutorial to learn how to identify the required access and set up required authentication and network rules for Microsoft Purview across Azure data sources:

Configure access to data sources for Microsoft Purview MSI at scale

Configure access to data sources for Microsoft Purview MSI at scale

06/03/2025

To scan data sources, Microsoft Purview requires access to them. This tutorial is intended for Azure subscription owners and Microsoft Purview Data Source Administrators. It will help you identify required access and set up required authentication and network rules for Microsoft Purview across Azure data sources.

In part 2 of this tutorial series, you'll:

- ✓ Locate your data sources and prepare a list of data source subscriptions.
- ✓ Run a script to configure any missing role-based access control (RBAC) or required network configurations across your data sources in Azure.
- ✓ Review the output report.

Prerequisites

- Azure subscriptions where your data sources are located. If you don't have an Azure subscription, create a free account before you begin.
- An Microsoft Purview account.
- An Azure Key Vault resource in each subscription that has data sources like Azure SQL Database, Azure Synapse Analytics, or Azure SQL Managed Instance.
- The Microsoft Purview MSI Configuration script.

> **Note**
>
> The Microsoft Purview MSI Configuration script is available only for Windows. This script is currently supported for Microsoft Purview Managed Identity (MSI).

> **Important**
>
> We strongly recommend that you test and verify all the changes the script performs in your Azure environment before you deploy it into your production environment.

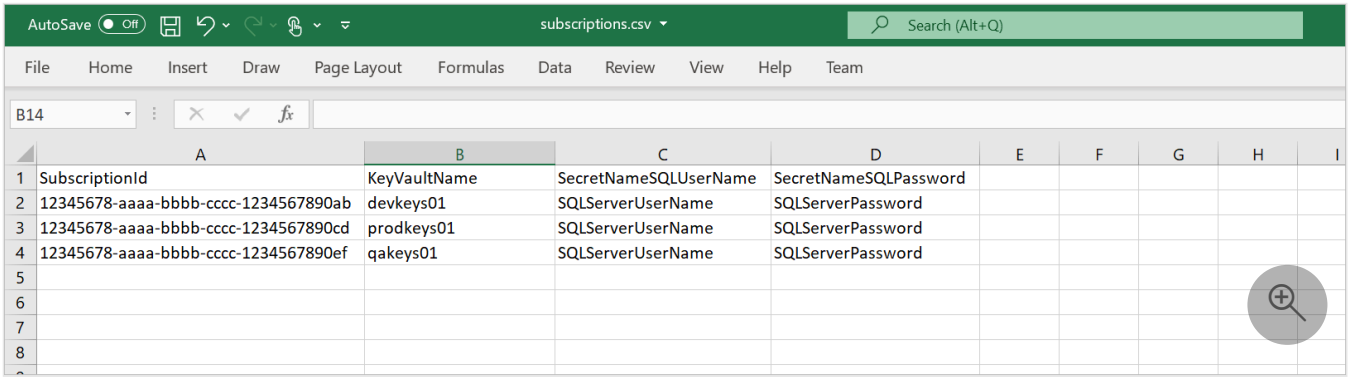
Prepare Azure subscriptions list for data sources

Before you run the script, create a .csv file (for example, "C:\temp\Subscriptions.csv) with four columns:

Expand table

Column name	Description	Example
`SubscriptionId`	Azure subscription IDs for your data sources.	12345678-aaaa-bbbb-cccc-1234567890ab
`KeyVaultName`	Name of existing key vault that's deployed in the data source subscription.	ContosoDevKeyVault
`SecretNameSQLUserName`	Name of an existing Azure Key Vault secret that contains a Microsoft Entra user name that can sign in to Azure Synapse, Azure SQL Database, or Azure SQL Managed Instance by using Microsoft Entra authentication.	ContosoDevSQLAdmin
`SecretNameSQLPassword`	Name of an existing Azure Key Vault secret that contains a Microsoft Entra user password that can sign in to Azure Synapse, Azure SQL Database, or Azure SQL Managed Instance by using Microsoft Entra authentication.	ContosoDevSQLPassword

Sample .csv file:



> **Note**
>
> You can update the file name and path in the code, if you need to.

Run the script and install the required PowerShell modules

Follow these steps to run the script from your Windows computer:

1. Download Microsoft Purview MSI Configuration script to the location of your choice.
2. On your computer, enter **PowerShell** in the search box on the Windows taskbar. In the search list, select and hold (or right-click) **Windows PowerShell** and then select **Run as administrator**.
3. In the PowerShell window, enter the following command. (Replace `<path-to-script>` with the folder path of the extracted script file.)

   ```PowerShell
   dir -Path <path-to-script> | Unblock-File
   ```

4. Enter the following command to install the Azure cmdlets:

   ```PowerShell
   Install-Module -Name Az -AllowClobber -Scope CurrentUser
   ```

5. If you see the prompt *NuGet provider is required to continue*, enter **Y**, and then select **Enter**.
6. If you see the prompt *Untrusted repository*, enter **A**, and then select **Enter**.
7. Repeat the previous steps to install the `Az.Synapse` and `AzureAD` modules.

It might take up to a minute for PowerShell to install the required modules.

Collect other data needed to run the script

Before you run the PowerShell script to verify the readiness of data source subscriptions, obtain the values of the following arguments to use in the scripts:

- `AzureDataType`: Choose any of the following options as your data-source type to check the readiness for the data type across your subscriptions:
  - `BlobStorage`
  - `AzureSQLMI`
  - `AzureSQLDB`
  - `ADLSGen2`

- `ADLSGen1`
- `Synapse`
- `All`

- `PurviewAccount`: Your existing Microsoft Purview account resource name.
- `PurviewSub`: Subscription ID where the Microsoft Purview account is deployed.

Verify your permissions

Make sure your user has the following roles and permissions:

At a minimum, you need the following permissions to run the script in your Azure environment:

Expand table

Role	Scope	Why is it needed?
Global Reader	Microsoft Entra tenant	To read Azure SQL Admin user group membership and Microsoft Purview MSI
Application Administrator	Microsoft Entra tenant	To assign the **Directory Reader** role to Azure SQL managed instances
Contributor	Subscription or resource group where your Microsoft Purview account is created	To read the Microsoft Purview account resource and create a Key Vault resource and secret
Owner or User Access Administrator	Management group or subscription where your Azure data sources are located	To assign RBAC
Contributor	Management group or subscription where your Azure data sources are located	To set up network configuration
SQL Admin (Microsoft Entra authentication)	Azure SQL Server instances or Azure SQL managed instances	To assign the **db_datareader** role to Microsoft Purview
Access to your Azure key vault	Access to get/list Key Vault secret for Azure SQL Database, Azure SQL Managed Instance, or Azure Synapse authentication	

Run the client-side readiness script

Run the script by completing these steps:

1. Use the following command to go to the script's folder. Replace `<path-to-script>` with the folder path of the extracted file.

   ```
   PowerShell

   cd <path-to-script>
   ```

2. Run the following command to set the execution policy for the local computer. Enter **A** for *Yes to All* when you're prompted to change the execution policy.

   ```
   PowerShell

   Set-ExecutionPolicy -ExecutionPolicy Unrestricted
   ```

3. Run the script with the following parameters. Replace the `DataType`, `PurviewName`, and `SubscriptionID` placeholders.

   ```
   PowerShell

   .\purview-msi-configuration.ps1 -AzureDataType <DataType> -PurviewAccount
   <PurviewName> -PurviewSub <SubscriptionID>
   ```

   When you run the command, a pop-up window might appear twice prompting you to sign in to Azure and Microsoft Entra ID by using your Microsoft Entra credentials.

It can take several minutes to create the report, depending on the number of Azure subscriptions and resources in the environment.

You might be prompted to sign in to your Azure SQL Server instances if the credentials in the key vault don't match. You can provide the credentials or select **Enter** to skip the specific server.

After the process completes, view the output report to review the changes.

More information

What data sources are supported by the script?

Currently, the following data sources are supported by the script:

- Azure Blob Storage (BlobStorage)

- Azure Data Lake Storage Gen2 (ADLSGen2)
- Azure Data Lake Storage Gen1 (ADLSGen1)
- Azure SQL Database (AzureSQLDB)
- Azure SQL Managed Instance (AzureSQLMI)
- Azure Synapse (Synapse) dedicated pool

You can choose all or any of these data sources as input parameter when you run the script.

What configurations are included in the script?

This script can help you automatically complete the following tasks:

Azure Blob Storage (BlobStorage)

- RBAC. Assign the Azure RBAC **Reader** role to Microsoft Purview MSI on the selected scope. Verify the assignment.
- RBAC. Assign the Azure RBAC **Storage Blob Data Reader** role to Microsoft Purview MSI in each of the subscriptions below the selected scope. Verify the assignments.
- Networking. Report whether private endpoint is created for storage and enabled for Blob Storage.
- Service endpoint. If private endpoint is off, check whether service endpoint is on, and enable **Allow trusted Microsoft services to access this storage account**.

Azure Data Lake Storage Gen2 (ADLSGen2)

- RBAC. Assign the Azure RBAC **Reader** role to Microsoft Purview MSI on the selected scope. Verify the assignment.
- RBAC. Assign the Azure RBAC **Storage Blob Data Reader** role to Microsoft Purview MSI in each of the subscriptions below the selected scope. Verify the assignments.
- Networking. Report whether private endpoint is created for storage and enabled for Blob Storage.
- Service endpoint. If private endpoint is off, check whether service endpoint is on, and enable **Allow trusted Microsoft services to access this storage account**.

Azure Data Lake Storage Gen1 (ADLSGen1)

- Networking. Verify that service endpoint is on, and enable **Allow all Azure services to access this Data Lake Storage Gen1 account** on Data Lake Storage.
- Permissions. Assign Read/Execute access to Microsoft Purview MSI. Verify the access.

Azure SQL Database (AzureSQLDB)

- SQL Server instances:
  - Network. Report whether public endpoint or private endpoint is enabled.
  - Firewall. If private endpoint is off, verify firewall rules and enable **Allow Azure services and resources to access this server**.
  - Microsoft Entra administration. Enable Microsoft Entra authentication for Azure SQL Database.

- SQL databases:
  - SQL role. Assign the **db_datareader** role to Microsoft Purview MSI.

Azure SQL Managed Instance (AzureSQLMI)

- SQL Managed Instance servers:

  - Network. Verify that public endpoint or private endpoint is on. Report if public endpoint is off.

  - ProxyOverride. Verify that Azure SQL Managed Instance is configured as Proxy or Redirect.

  - Networking. Update NSG rules to allow AzureCloud inbound access to SQL Server instances over required ports:
    - Redirect: 1433 and 11000-11999

    or
    - Proxy: 3342

    Verify this access.

  - Microsoft Entra administration. Enable Microsoft Entra authentication for Azure SQL Managed Instance.

- SQL databases:
  - SQL role. Assign the **db_datareader** role to Microsoft Purview MSI.

Azure Synapse (Synapse) dedicated pool

- RBAC. Assign the Azure RBAC **Reader** role to Microsoft Purview MSI on the selected scope. Verify the assignment.

- RBAC. Assign the Azure RBAC **Storage Blob Data Reader** role to Microsoft Purview MSI in each of the subscriptions below the selected scope. Verify the assignments.

- SQL Server instances (dedicated pools):
  - Network. Report whether public endpoint or private endpoint is on.
  - Firewall. If private endpoint is off, verify firewall rules and enable **Allow Azure services and resources to access this server**.
  - Microsoft Entra administration. Enable Microsoft Entra authentication for Azure SQL Database.

- SQL databases:
  - SQL role. Assign the **db_datareader** role to Microsoft Purview MSI.

Next steps

In this tutorial, you learned how to:

- ✓ Identify required access and set up required authentication and network rules for Microsoft Purview across Azure data sources.

Go to the next tutorial to learn how to Register and scan multiple sources in Microsoft Purview.

Discover and govern multiple Azure sources in Microsoft Purview

Article • 05/21/2025

This article outlines how to register multiple Azure sources and how to authenticate and interact with them in Microsoft Purview. For more information about Microsoft Purview, read the introductory article.

Supported capabilities

Scanning capabilities

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan
Yes	Yes	Yes	Yes

Other capabilities

For **classifications**, **sensitivity labels**, **policies**, **data lineage**, and **live view**, see the list of supported capabilities.

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- An active Microsoft Purview account.
- You'll need to be a Data Source Administrator and Data Reader to register a source and manage it in the Microsoft Purview governance portal. See our Microsoft Purview Permissions page for details.

Register

This section describes how to register multiple Azure sources in Microsoft Purview using the Microsoft Purview governance portal.

Prerequisites for registration

Microsoft Purview needs permissions to be able to list resources under a subscription or resource group.

1. Go to the subscription or the resource group in the Azure portal.
2. Select **Access Control (IAM)** from the left menu.
3. Select **+Add**.
4. In the **Select input** box, select the **Reader** role and enter your Microsoft Purview account name (which represents its MSI file name).
5. Select **Save** to finish the role assignment. This will allow Microsoft Purview to list resources under a subscription or resource group.

Authentication for registration

There are two ways to set up authentication for multiple sources in Azure:

- Managed identity
- Service principal

You must set up authentication on each resource within your subscription or resource group that you want to register and scan. Azure Storage resource types (Azure Blob Storage and Azure Data Lake Storage Gen2) make it easy by allowing you to add the MSI file or service principal at the subscription or resource group level as a storage blob data reader. The permissions then trickle down to each storage account within that subscription or resource group. For all other resource types, you must apply the MSI file or service principal on each resource, or create a script to do so.

To learn how to add permissions on each resource type within a subscription or resource group, see the following resources:

- Azure Blob Storage
- Azure Data Explorer
- Azure Data Lake Storage Gen1
- Azure Data Lake Storage Gen2
- Azure Dedicated SQL pool (formerly SQL DW)
- Azure SQL Database
- Azure SQL Managed Instance
- Azure Synapse Analytics

Steps to register

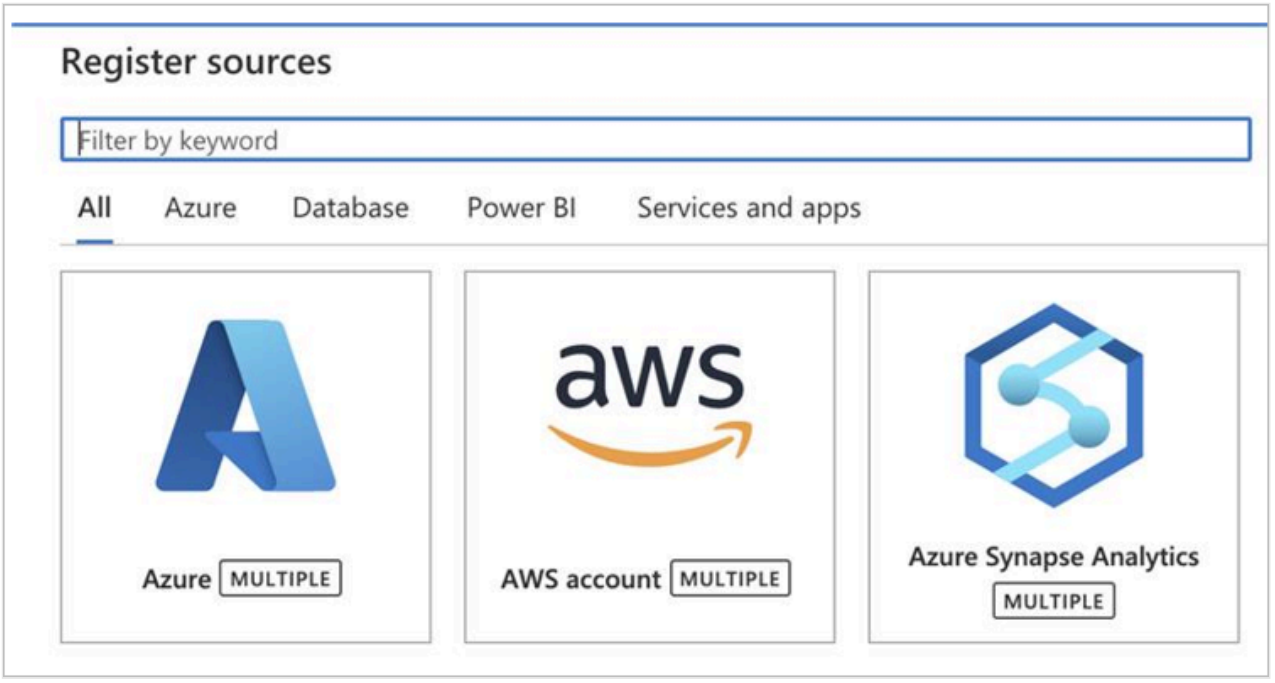
1. Open the Microsoft Purview governance portal by:

- Browsing directly to https://web.purview.azure.com and selecting your Microsoft Purview account.
- Opening the Azure portal, searching for and selecting the Microsoft Purview account. Selecting the **the Microsoft Purview governance portal** button.

2. Select **Data Map** on the left menu.

3. Select **Register**.

4. On **Register sources**, select **Azure (multiple)**.



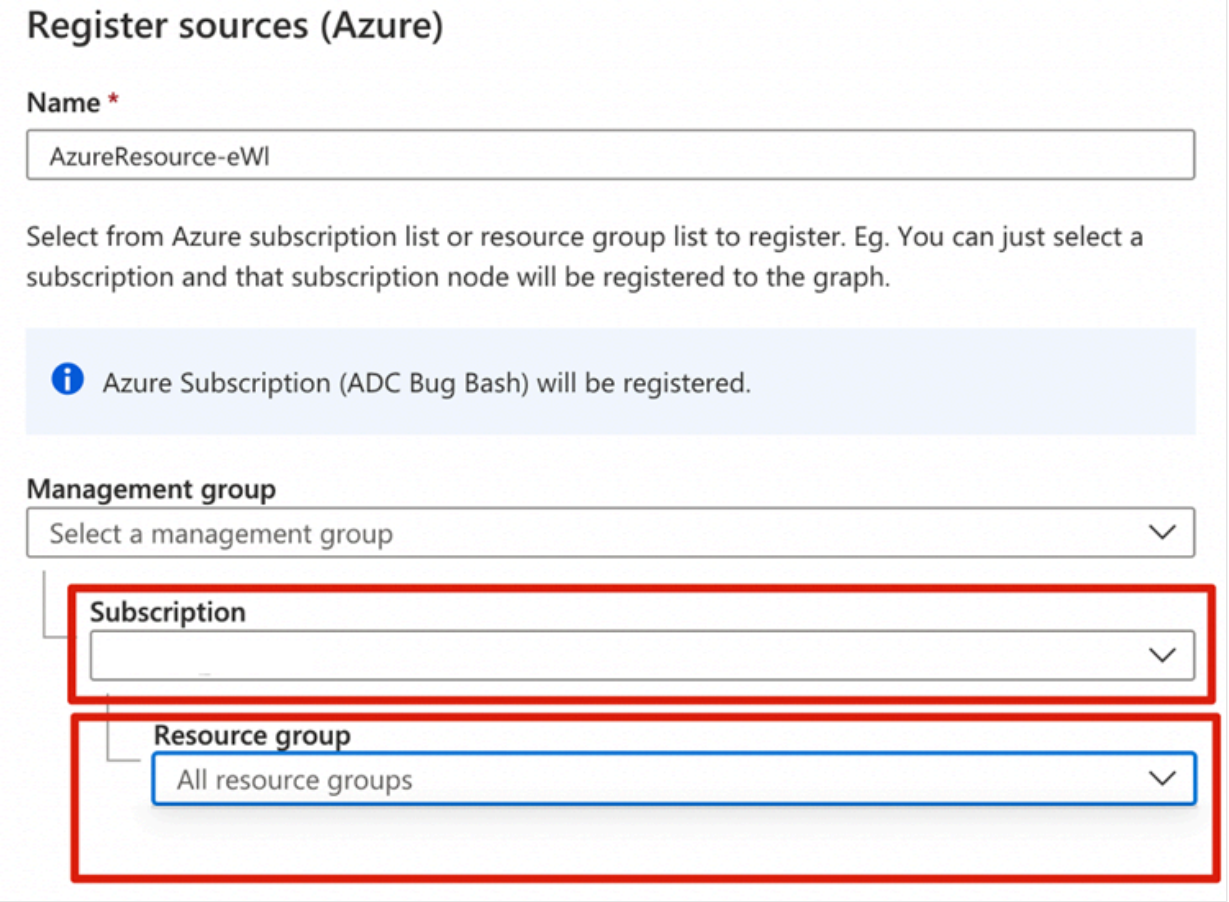
5. Select **Continue**.

6. On the **Register sources (Azure)** screen, do the following:

   a. In the **Name** box, enter a name that the data source will be listed with in the catalog.

   b. In the **Management group** box, optionally choose a management group to filter down to.

   c. In the **Subscription** and **Resource group** dropdown list boxes, select a subscription or a specific resource group, respectively. The registration scope will be set to the selected subscription or resource group.



d. Select a collection from the list.

e. Select **Register** to register the data sources.

Scan

> **Important**
>
> Currently, scanning multiple Azure sources is only supported using Azure integration runtime, therefore, only Microsoft Purview accounts that allow public access on the firewall can use this option.

Follow the steps below to scan multiple Azure sources to automatically identify assets and classify your data. For more information about scanning in general, see our introduction to scans and ingestion.

Create and run scan

To create and run a new scan, do the following:

1. Select the **Data Map** tab on the left pane in the Microsoft Purview governance portal.

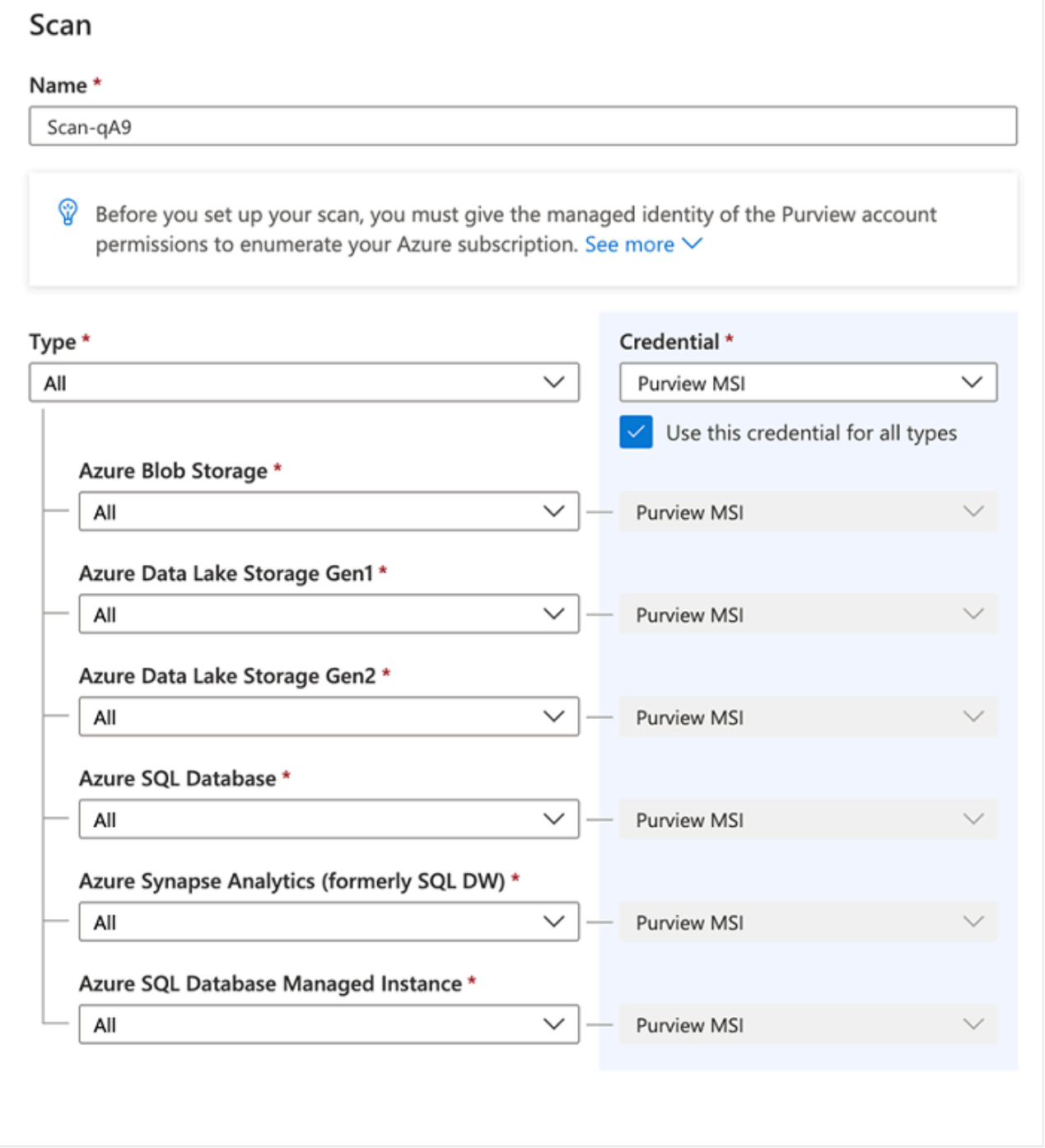
2. Select the data source that you registered.

3. Select **View details** > **+ New scan**, or use the **Scan** quick-action icon on the source tile.

4. For **Name**, fill in the name.

5. For **Type**, select the types of resources that you want to scan within this source. Choose one of these options:

    - Leave it as **All**. This selection includes future resource types that might not currently exist within that subscription or resource group.
    - Use the boxes to specifically select resource types that you want to scan. If you choose this option, future resource types that might be created within this subscription or resource group won't be included for scans, unless the scan is explicitly edited in the future.



6. Select the credential to connect to the resources within your data source:

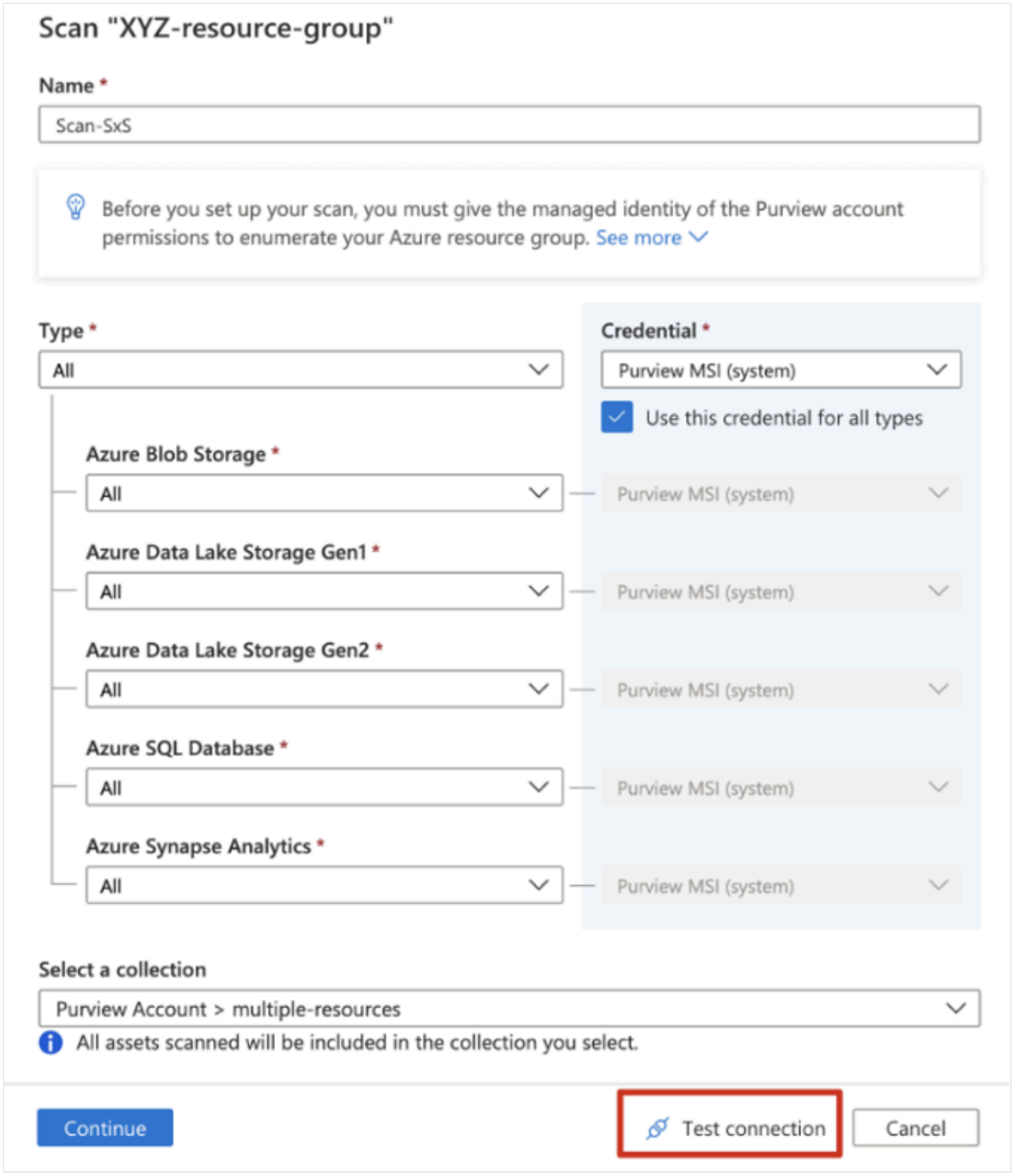
    - You can select a credential at the parent level as an MSI file, or you can select a credential for a particular service principal type. You can then use that credential for all the resource types under the subscription or resource group.
    - You can specifically select the resource type and apply a different credential for that resource type.

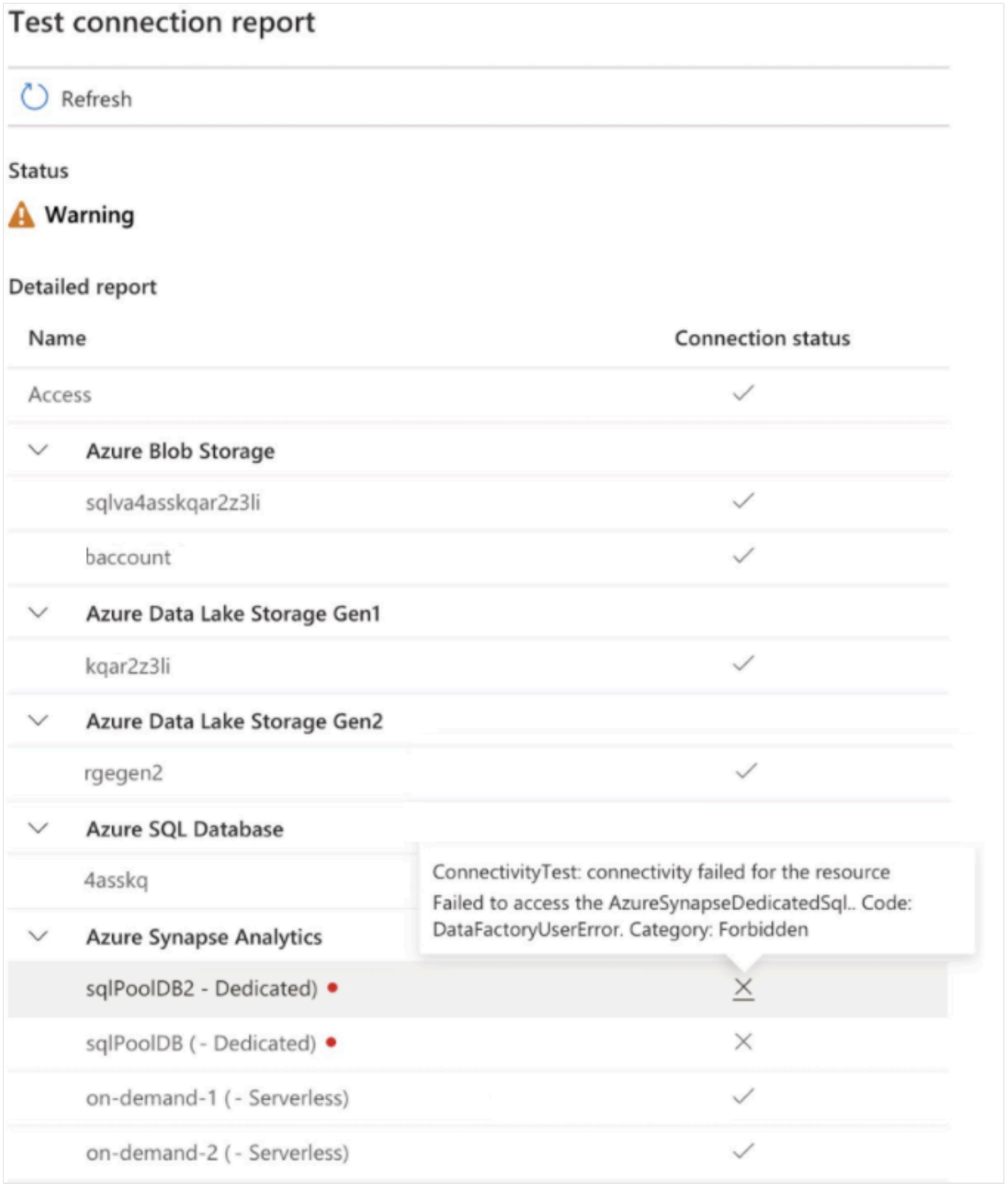
    Each credential will be considered as the method of authentication for all the resources under a particular type. You must set the chosen credential on the resources in order to successfully scan them, as described earlier in this article.

7. Within each type, you can select to either scan all the resources or scan a subset of them by name:

- If you leave the option as **All**, then future resources of that type will also be scanned in future scan runs.
- If you select specific storage accounts or SQL databases, then future resources of that type created within this subscription or resource group won't be included for scans, unless the scan is explicitly edited in the future.

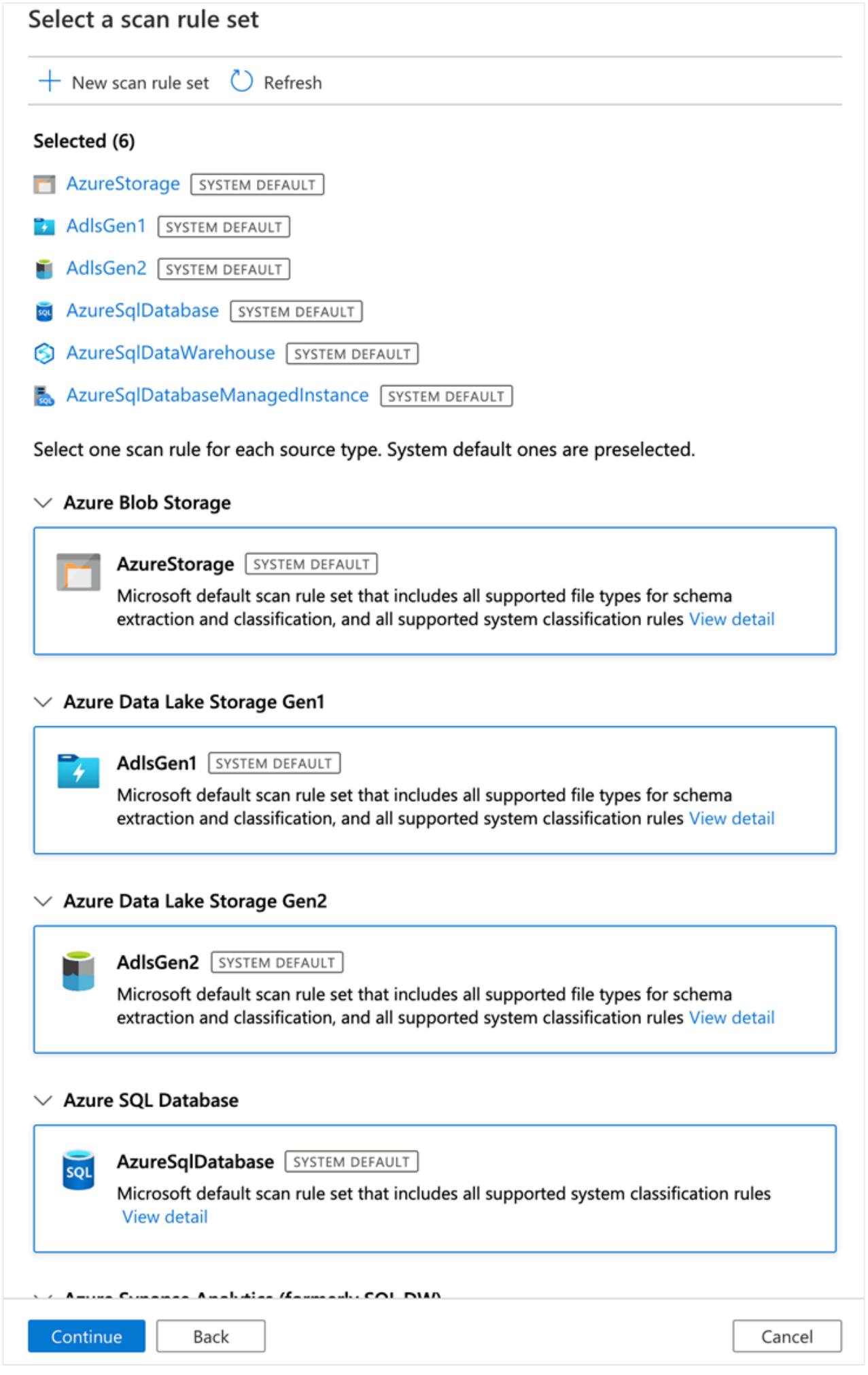
8. Select **Test connection**. This will first test access to check if you've applied the Microsoft Purview MSI file as a reader on the subscription or resource group. If you get an error message, follow these instructions to resolve it. Then it will test your authentication and connection to each of your selected sources and generate a report. The number of sources selected will impact the time it takes to generate this report. If failed on some resources, hovering over the **X** icon will display the detailed error message.


Scan "XYZ-resource-group"
Name *
Scan-SxS
Before you set up your scan, you must give the managed identity of the Purview account permissions to enumerate your Azure resource group. See more
Type *
All
Credential *
Purview MSI (system)
Use this credential for all types
Azure Blob Storage *
All
Purview MSI (system)
Azure Data Lake Storage Gen1 *
All
Purview MSI (system)
Azure Data Lake Storage Gen2 *
All
Purview MSI (system)
Azure SQL Database *
All
Purview MSI (system)
Azure Synapse Analytics *
All
Purview MSI (system)
Select a collection
Purview Account > multiple-resources
All assets scanned will be included in the collection you select.
Continue
Test connection
Cancel



9. After your test connection has passed, select **Continue** to proceed.

10. Select scan rule sets for each resource type that you chose in the previous step. You can also create scan rule sets inline.

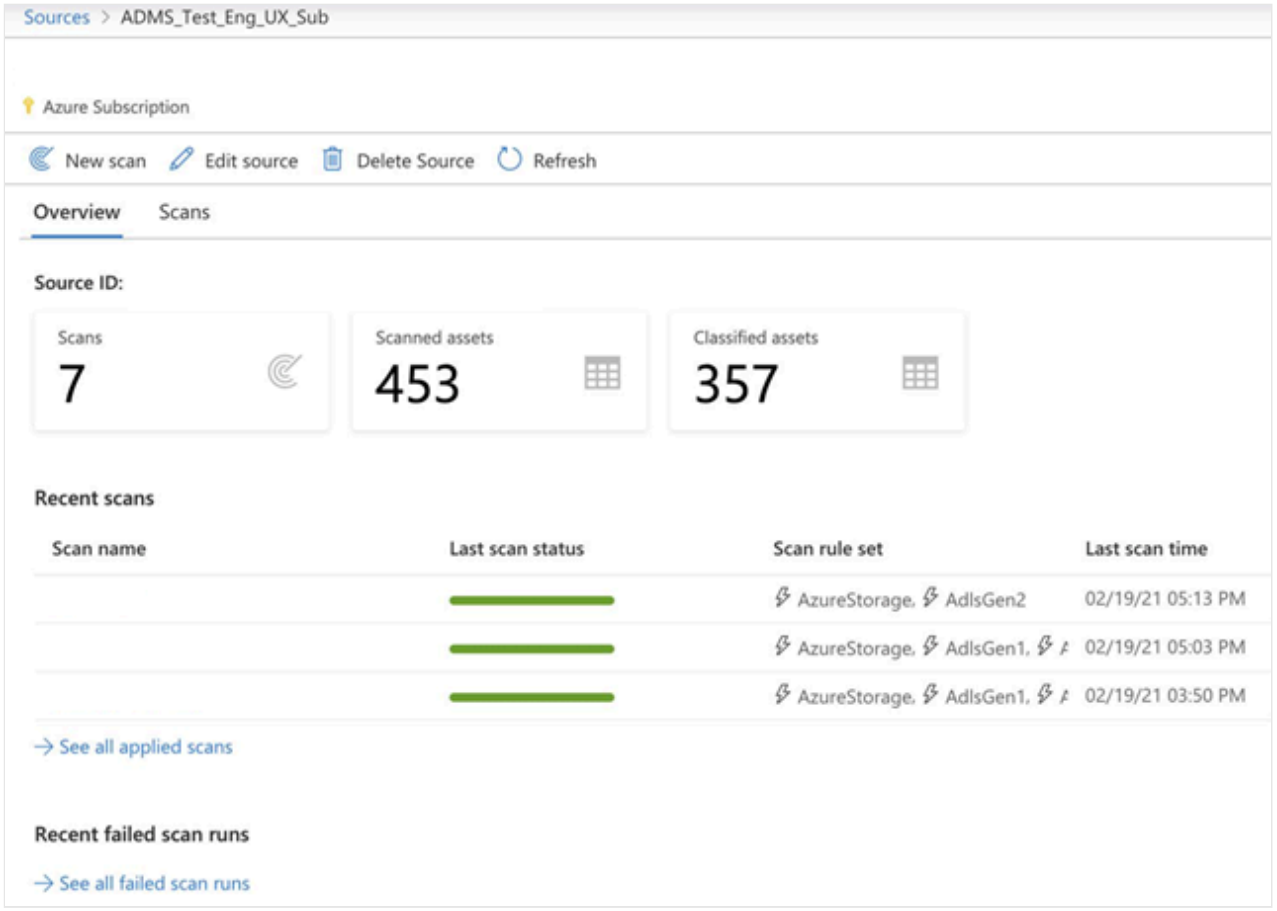


11. Choose your scan trigger. You can set up a schedule or run the scan once.

12. Review your scan and select **Save** to complete setup.

View your scans and scan runs

1. View source details by selecting **View details** on the tile under the **Data Map** section.

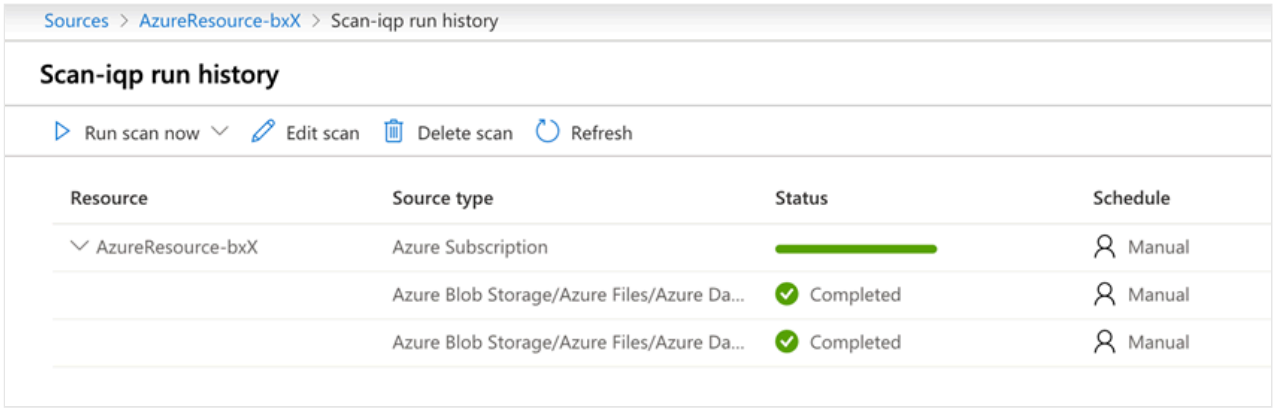


2. View scan run details by going to the **Scan details** page.

   The *status bar* is a brief summary of the running status of the child resources. It's displayed on the subscription level or resource group level. The colors have the following meanings:

   - Green: The scan was successful.
   - Red: The scan failed.
   - Gray: The scan is still in progress.

   You can select each scan to view finer details.



3. View a summary of recent failed scan runs at the bottom of the source details. You can also view more granular details about these runs.

Manage your scans: edit, delete, or cancel

To manage a scan, do the following:

1. Go to the management center.
2. Select **Data sources** under the **Sources and scanning** section, and then select the desired data source.
3. Select the scan that you want to manage. Then:
    - You can edit the scan by selecting **Edit**.
    - You can delete the scan by selecting **Delete**.
    - If the scan is running, you can cancel it by selecting **Cancel**.

Policies

The following types of policies are supported on this data resource from Microsoft Purview:

- DevOps policies
- Data owner policies

Access policy pre-requisites on Azure Storage accounts

To be able to enforce policies from Microsoft Purview, data sources under a resource group or subscription need to be configured first. Instructions vary based on the data source type. Please review whether they support Microsoft Purview policies, and if so, the specific instructions to enable them, under the Access Policy link in the Microsoft Purview connector document.

Configure the Microsoft Purview account for policies

Register the data source in Microsoft Purview

Before a policy can be created in Microsoft Purview for a data resource, you must register that data resource in Microsoft Purview Studio. You will find the instructions related to registering the data resource later in this guide.

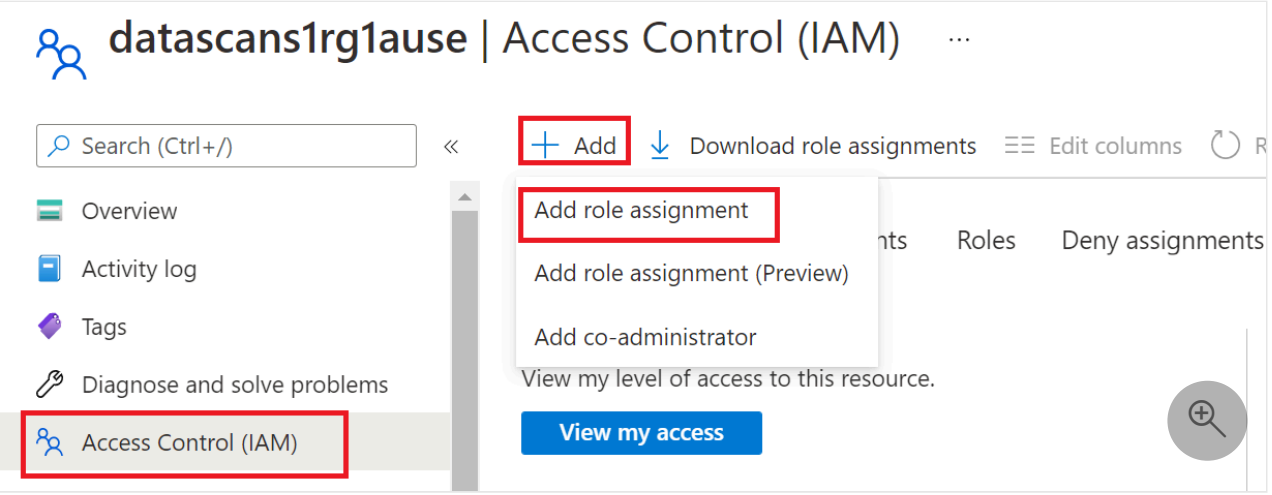
> ⓘ **Note**
>
> Microsoft Purview policies rely on the data resource ARM path. If a data resource is moved to a new resource group or subscription it will need to be de-registered and then registered again in Microsoft Purview.

Configure permissions to enable Data policy enforcement on the data source

Once a resource is registered, but before a policy can be created in Microsoft Purview for that resource, you must configure permissions. A set of permissions are needed to enable the **Data policy enforcement**. This applies to data sources, resource groups, or subscriptions. To enable *Data policy enforcement*, you must have *both* specific Identity and Access Management (IAM) privileges on the resource as well as specific Microsoft Purview privileges:

- You must have *either one* of the following IAM role combinations on the resource's Azure Resource Manager path or any parent of it (that is, using IAM permission inheritance):
  - IAM Owner
  - Both IAM Contributor and IAM User Access Administrator

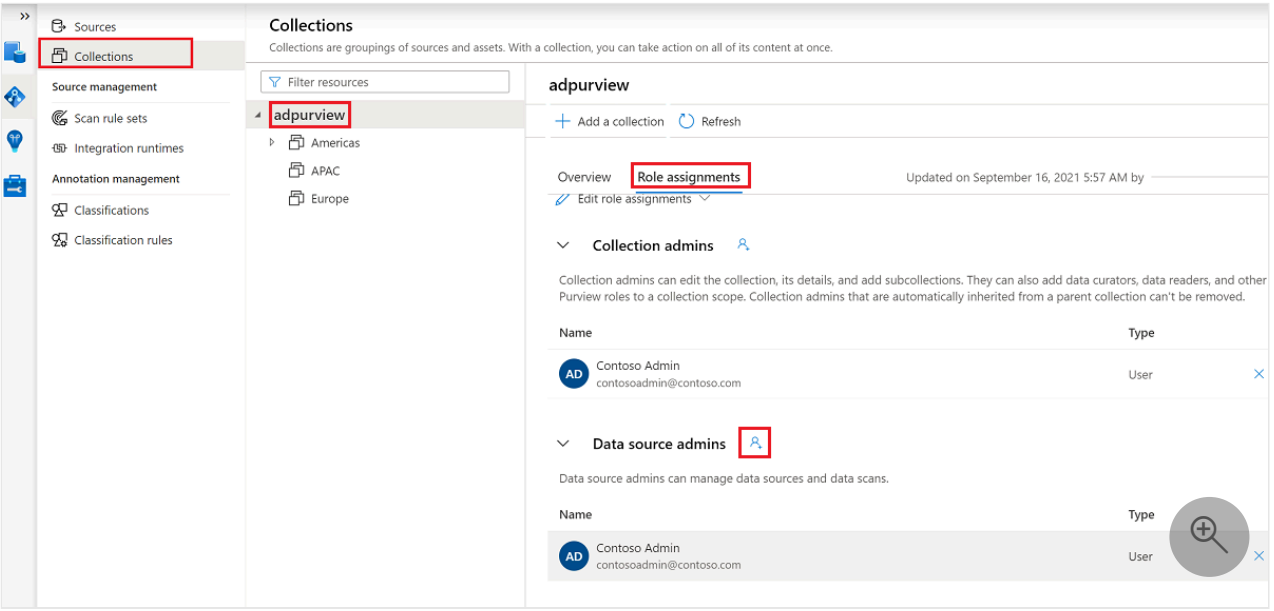
  To configure Azure role-based access control (RBAC) permissions, follow this guide. The following screenshot shows how to access the **Access Control** section in the Azure portal for the data resource to add a role assignment.



> **Note**
>
> The *IAM Owner* role for a data resource can be inherited from a parent resource group, a subscription, or a subscription management group. Check which Microsoft Entra users, groups, and service principals hold or are inheriting the *IAM Owner* role for the resource.

- You also need to have the Microsoft Purview *Data source admin* role for the collection or a parent collection (if inheritance is enabled). For more information, see the guide on managing Microsoft Purview role assignments.

  The following screenshot shows how to assign the *Data source admin* role at the root collection level.



Configure Microsoft Purview permissions to create, update, or delete access policies

To create, update or delete policies, you need to get the *Policy author* role in Microsoft Purview at *root collection level*:

- The *Policy author* role can create, update, and delete DevOps and Data Owner policies.
- The *Policy author* role can delete self-service access policies.

For more information about managing Microsoft Purview role assignments, see Create and manage collections in the Microsoft Purview Data Map.

> **Note**
>
> Policy author role must be configured at the root collection level.

In addition, to easily search Microsoft Entra users or groups when creating or updating the subject of a policy, you can greatly benefit from getting the Directory Readers permission in Microsoft Entra ID. This is a common permission for users in an Azure tenant. Without the Directory Reader permission, the Policy Author will have to type the complete username or email for all the principals included in the subject of a data policy.

Configure Microsoft Purview permissions for publishing Data Owner policies

Data Owner policies allow for checks and balances if you assign the Microsoft Purview *Policy author* and *Data source admin* roles to different people in the organization. Before a Data owner policy takes effect, a second person (*Data source admin*) must review it and explicitly approve it by publishing it. This does not apply to DevOps or Self-service access policies as publishing is automatic for them when those policies are created or updated.

To publish a Data owner policy you need to get the *Data source admin* role in Microsoft Purview at *root collection level*.

For more information about managing Microsoft Purview role assignments, see Create and manage collections in the Microsoft Purview Data Map.

> **Note**
>
> To publish Data owner policies, the Data source admin role must be configured at the root collection level.

Delegate access provisioning responsibility to roles in Microsoft Purview

After a resource has been enabled for **Data policy enforcement**, any Microsoft Purview user with the *Policy author* role at the root collection level can provision access to that data source from Microsoft Purview.

> ⓘ **Note**
>
> Any Microsoft Purview root *Collection admin* can assign new users to root *Policy author* roles. Any *Collection admin* can assign new users to a *Data source admin* role under the collection. Minimize and carefully vet the users who hold Microsoft Purview *Collection admin*, *Data source admin*, or *Policy author* roles.

If a Microsoft Purview account with published policies is deleted, such policies will stop being enforced within an amount of time that depends on the specific data source. This change can have implications on both security and data access availability. The Contributor and Owner roles in IAM can delete Microsoft Purview accounts. You can check these permissions by going to the **Access control (IAM)** section for your Microsoft Purview account and selecting **Role Assignments**. You can also use a lock to prevent the Microsoft Purview account from being deleted through Resource Manager locks.

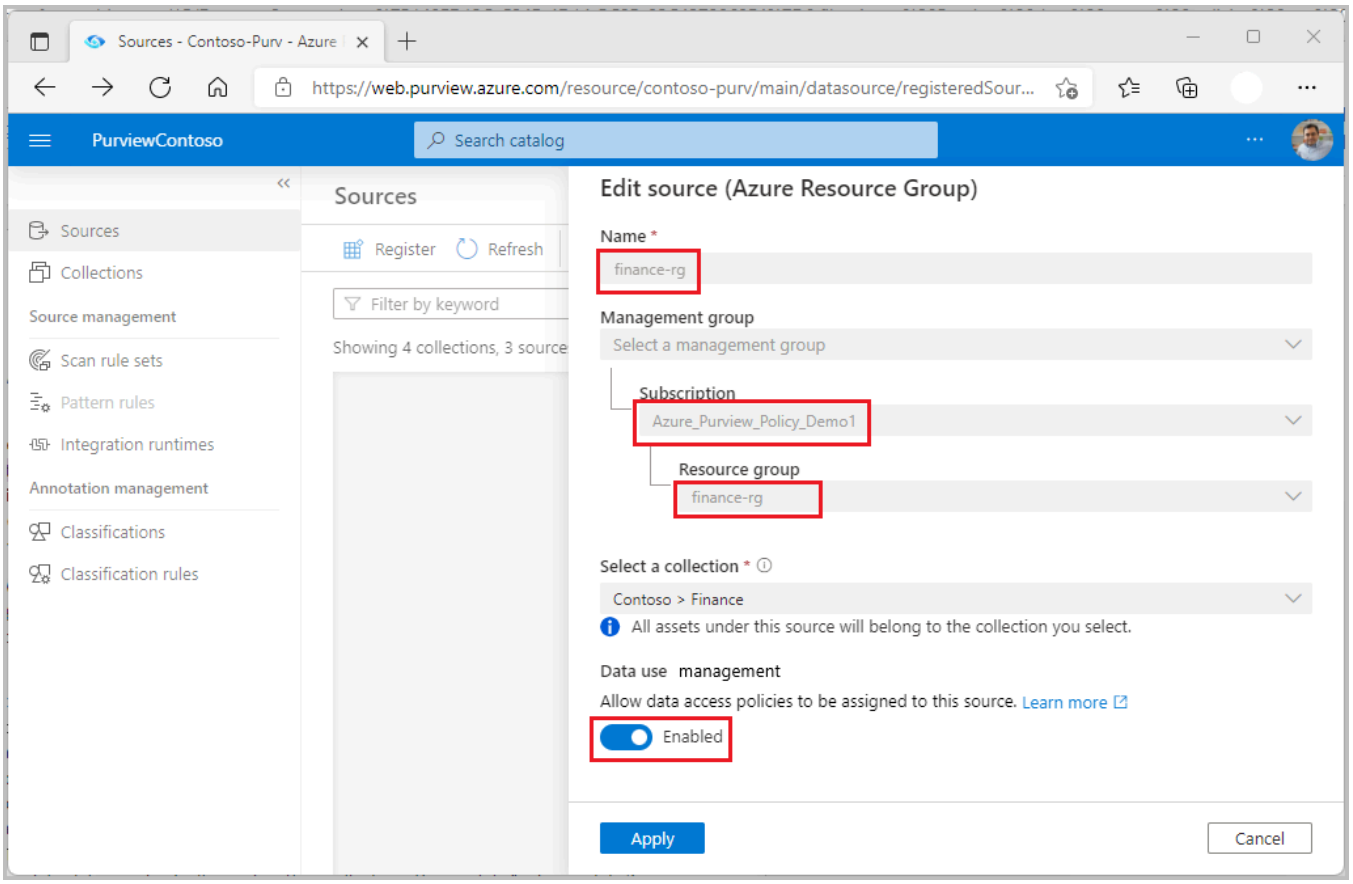
Register the data source in Microsoft Purview for Data Policy Enforcement

The Azure subscription or resource group needs to be registered first with Microsoft Purview before you can create access policies. To register your resource, follow the **Prerequisites** and **Register** sections of this guide:

- Register multiple sources in Microsoft Purview

After you've registered the data resource, you'll need to enable Data Policy Enforcement. This is a pre-requisite before you can create policies on the data resource. Data Policy Enforcement can impact the security of your data, as it delegates to certain Microsoft Purview roles managing access to the data sources. **Go through the secure practices related to Data Policy Enforcement in this guide**: How to enable Data Policy Enforcement

Once your data source has the **Data Policy Enforcement** option set to **Enabled**, it will look like this screenshot:



Create a policy

To create an access policy on an entire Azure subscription or resource group, follow these guides:

- DevOps policy covering all sources in a subscription or resource group
- Provision read/modify access to all sources in a subscription or resource group

Next steps

Now that you've registered your source, follow the below guides to learn more about Microsoft Purview and your data.

- Devops policies in Microsoft Purview
- Data Estate Insights in Microsoft Purview
- Lineage in Microsoft Purview
- Search Unified Catalog

Connect to and manage Azure Synapse Analytics workspaces in Microsoft Purview

Summarize this article for me

This article outlines how to register Azure Synapse Analytics workspaces. It also describes how to authenticate and interact with Azure Synapse Analytics workspaces in Microsoft Purview. Learn more about Microsoft Purview.

> **Note**
>
> There are dedicated SQL pools (formerly SQL DW) and dedicated SQL pools in Azure Synapse workspace. For the detailed difference between the two, see What is the difference between dedicated SQL pools (SQL DW) and dedicated SQL pools in Azure Synapse workspaces. Currently Microsoft Purview provides separate data sources for dedicated SQL pools (formerly SQL DW) and dedicated SQL pools:
>
> - To register and scan a dedicated SQL pool (formerly SQL DW), see Connect to and manage dedicated SQL pools in Microsoft Purview.
> - To register and scan a dedicated SQL pool (formerly SQL DW) that has enabled Azure Synapse workspace features, as documented in Enable Azure Synapse workspace features for a dedicated SQL pool (formerly SQL DW), see Connect to and manage Azure Synapse Analytics workspaces in Microsoft Purview.
> - To register and scan a dedicated SQL pool or a serverless SQL pool in Azure Synapse workspace, see Connect to and manage Azure Synapse Analytics workspaces in Microsoft Purview.

Supported capabilities

Scanning capabilities

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan
Yes	Yes	Yes	No

Azure Synapse Analytics lake databases aren't currently supported. You might observe some inconsistencies in the scanned assets for these databases.

For external tables, Azure Synapse Analytics doesn't currently capture the relationship of those tables to their original files.

Other capabilities

For **classifications**, **sensitivity labels**, **policies**, **data lineage**, and **live view**, see the list of supported capabilities.

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- An active Microsoft Purview account.
- Data source administrator and data reader permissions to register a source and manage it in the Microsoft Purview governance portal. For details, see Access control in the Microsoft Purview governance portal.

Register

The following procedure describes how to register Azure Synapse Analytics workspaces in Microsoft Purview by using the Microsoft Purview governance portal.

To register an Synapse Analytics workspace, you need at least the data reader role on the Synapse Analytics workspace and be a data source administrator in Microsoft Purview.

1. Open the Microsoft Purview governance portal and select your Microsoft Purview account.

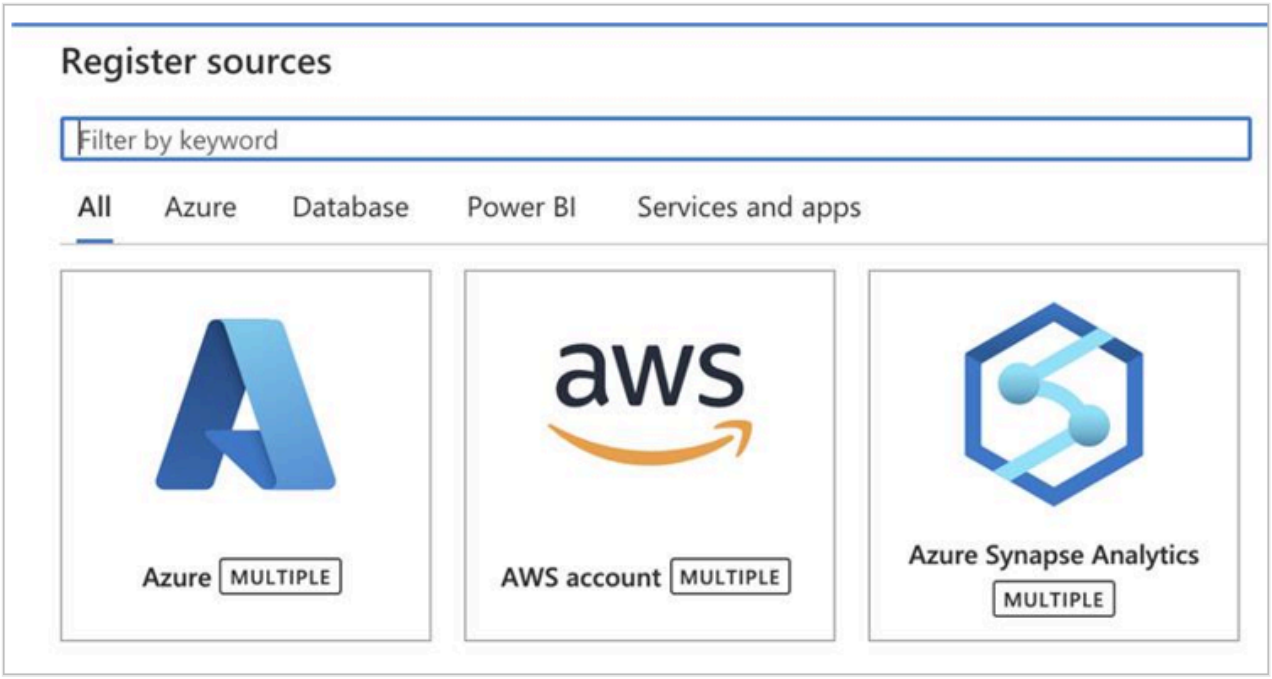
   Alternatively, go to the Azure portal, search for and select the Microsoft Purview account, and then select the **Microsoft Purview governance portal** button.

2. On the left pane, select **Sources**.

3. Select **Register**.

4. Under **Register sources**, select **Azure Synapse Analytics (multiple)**.

5. Select **Continue**.



6. On the **Register sources (Azure Synapse Analytics)** page, do the following:

   a. For **Name**, enter a name for the data source to be listed in Microsoft Purview Unified Catalog.

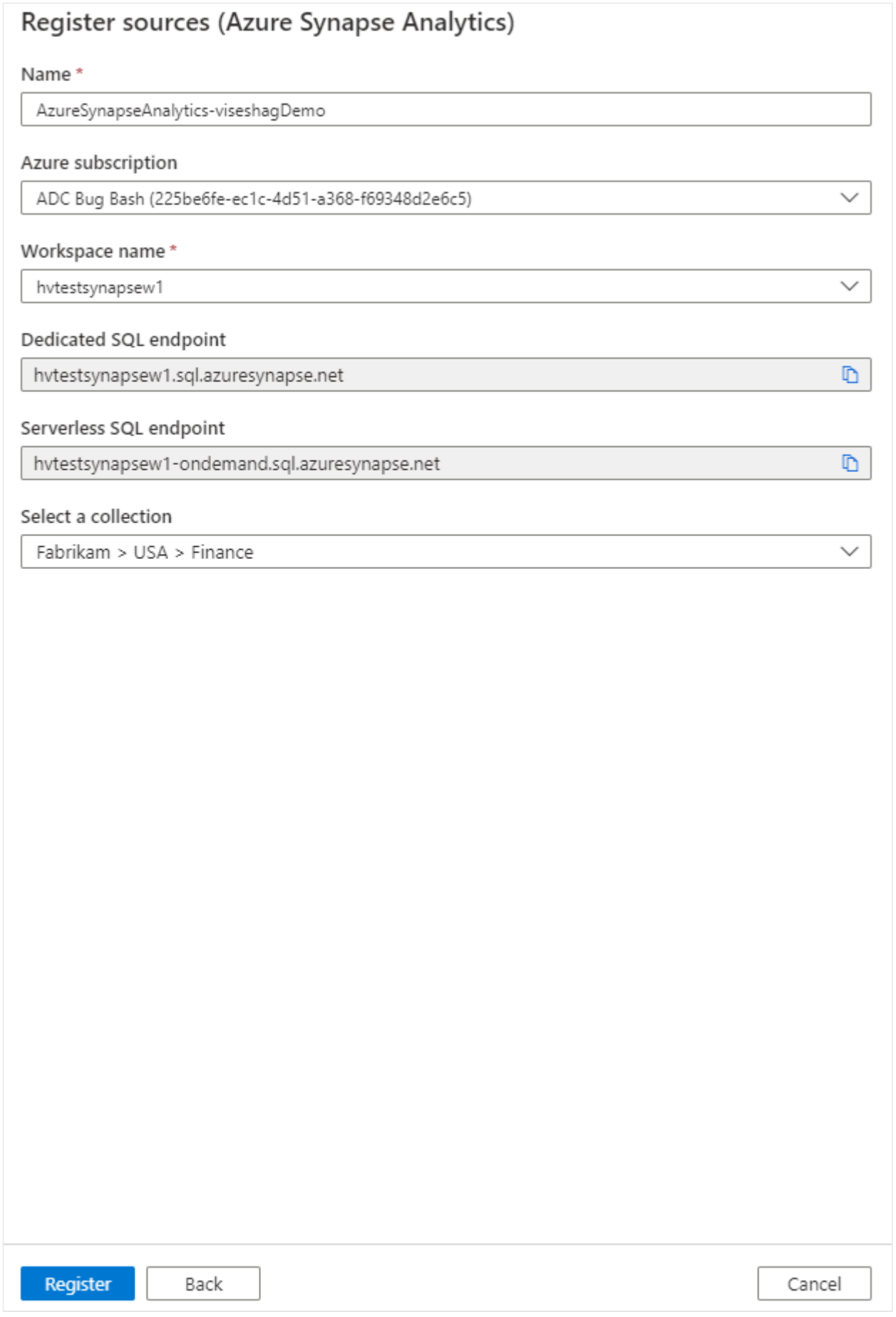
   b. Optionally, for **Azure subscription**, choose a subscription to filter down to.

   c. For **Workspace name**, select the workspace that you're working with.

      The boxes for the SQL endpoints are automatically filled in based on your workspace selection.

   d. Select a collection from the list.

   e. Select **Register** to finish registering the data source.


Register sources (Azure Synapse Analytics)
Name *
AzureSynapseAnalytics-viseshagDemo
Azure subscription
ADC Bug Bash (225be6fe-ec1c-4d51-a368-f69348d2e6c5)
Workspace name *
hvtestsynapsew1
Dedicated SQL endpoint
hvtestsynapsew1.sql.azuresynapse.net
Serverless SQL endpoint
hvtestsynapsew1-ondemand.sql.azuresynapse.net
Select a collection
Fabrikam > USA > Finance
Register
Back
Cancel

Scan

Use the following steps to scan Azure Synapse Analytics workspaces to automatically identify assets and classify your data. For more information about scanning in general, see Scans and ingestion in Microsoft Purview.

1. Set up authentication for enumerating your dedicated or serverless resources. By setting up authentication, Microsoft Purview can enumerate your workspace assets and perform scans.
2. Apply permissions to scan the contents of the workspace.
3. Confirm that your network is set up to allow access for Microsoft Purview.

Enumeration authentication

Use the following procedures to set up authentication. You must be an owner or a user access administrator to add the specified roles.

Authentication for enumerating dedicated SQL database resources

1. In the Azure portal, go to the Azure Synapse Analytics workspace resource.
2. On the left pane, select **Access Control (IAM)**.
3. Select the **Add** button.
4. Set the **Reader** role and enter your Microsoft Purview account name, which represents its managed service identity (MSI).
5. Select **Save** to finish assigning the role.

If you want to scan a dedicated SQL pool (formerly SQL DW) that has enabled Azure Synapse Analytics workspace features as documented in Enable Azure Synapse workspace features for a dedicated SQL pool (formerly SQL DW), follow these additional configuration steps for the dedicated SQL pool (formerly SQL DW):

1. In the Azure portal, go to the SQL Server resource associated with the dedicated SQL pool (formerly SQL DW).
2. On the left pane, select **Access Control (IAM)**.
3. Select the **Add** button.
4. Set the **Reader** role and enter your Microsoft Purview account name, which represents its managed service identity (MSI).
5. Select **Save** to finish assigning the role.

> **Note**

If you're planning to register and scan multiple Azure Synapse Analytics workspaces in your Microsoft Purview account, you can also assign the role from a higher level, such as a resource group or a subscription.

Authentication for enumerating serverless SQL Database resources

Set authentication in three places to allow Microsoft Purview to enumerate your serverless SQL Database resources.

To set authentication for the Azure Synapse Analytics workspace:

1. In the Azure portal, go to the Azure Synapse Analytics workspace resource.
2. On the left pane, select **Access Control (IAM)**.
3. Select the **Add** button.
4. Set the **Reader** role and enter your Microsoft Purview account name, which represents its MSI.
5. Select **Save** to finish assigning the role.

To set authentication for the storage account:

1. In the Azure portal, go to the resource group or subscription that contains the storage account associated with the Azure Synapse Analytics workspace.
2. On the left pane, select **Access Control (IAM)**.
3. Select the **Add** button.
4. Set the **Storage blob data reader** role and enter your Microsoft Purview account name (which represents its MSI) in the **Select** box.
5. Select **Save** to finish assigning the role.

To set authentication for the Azure Synapse Analytics serverless database:

1. Go to your Azure Synapse Analytics workspace and open Synapse Studio.
2. On the left pane, select **Data**.
3. Select the ellipsis (...) next to one of your databases, and then start a new SQL script.
4. Run the following command in your SQL script to add the Microsoft Purview account MSI (represented by the account name) on the serverless SQL databases:

```SQL
CREATE LOGIN [PurviewAccountName] FROM EXTERNAL PROVIDER;
```

Apply permissions to scan the contents of the workspace

You must set up authentication on each SQL Database that you want to register and scan from your Azure Synapse Analytics workspace. For steps to apply permissions, see the following scenarios.

> **Important**
>
> The following steps for serverless databases *do not* apply to replicated databases. In Azure Synapse Analytics, serverless databases that are replicated from Spark databases are currently read-only. For more information, see Operation isn't allowed for a replicated database.

Managed identity

Use a managed identity for dedicated SQL databases

> **Important**
>
> If you use a self-hosted integration runtime to connect to your resource on a private network, managed identities don't work. You need to use service principal authentication or SQL authentication.

To run the commands in the following procedure, you must be an Azure Synapse administrator on the workspace. For more information about Azure Synapse Analytics permissions, see Set up access control for your Azure Synapse Analytics workspace.

1. Go to your Azure Synapse Analytics workspace.
2. Go to the **Data** section, and then look for one of your dedicated SQL databases.
3. Select the ellipsis (...) next to the database name, and then start a new SQL script.
4. Run the following command in your SQL script to add the Microsoft Purview account MSI (represented by the account name) as `db_datareader` on the dedicated SQL database:

```
SQL

CREATE USER [PurviewAccountName] FROM EXTERNAL PROVIDER
GO
```

```
EXEC sp_addrolemember 'db_datareader', [PurviewAccountName]
GO
```

5. Run the following command in your SQL script to verify the addition of the role:

```
SQL

SELECT p.name AS UserName, r.name AS RoleName
FROM sys.database_principals p
LEFT JOIN sys.database_role_members rm ON p.principal_id =
rm.member_principal_id
LEFT JOIN sys.database_principals r ON rm.role_principal_id =
r.principal_id
WHERE p.authentication_type_desc = 'EXTERNAL'
ORDER BY p.name;
```

Follow the same steps for each database that you want to scan.

Use a managed identity for serverless SQL databases

1. Go to your Azure Synapse Analytics workspace.
2. Go to the **Data** section, and select one of your SQL databases.
3. Select the ellipsis (...) next to the database name, and then start a new SQL script.
4. Run the following command in your SQL script to add the Microsoft Purview account MSI (represented by the account name) as `db_datareader` on the serverless SQL databases:

```
SQL

CREATE USER [PurviewAccountName] FOR LOGIN [PurviewAccountName];
ALTER ROLE db_datareader ADD MEMBER [PurviewAccountName];
```

5. Run the following command in your SQL script to verify the addition of the role:

```
SQL

SELECT p.name AS UserName, r.name AS RoleName
FROM sys.database_principals p
LEFT JOIN sys.database_role_members rm ON p.principal_id =
rm.member_principal_id
LEFT JOIN sys.database_principals r ON rm.role_principal_id =
r.principal_id
```

```
WHERE p.authentication_type_desc = 'EXTERNAL'
ORDER BY p.name;
```

Follow the same steps for each database that you want to scan.

Grant permission to use credentials for external tables

If the Synapse Analytics workspace has any external tables, you must give the Microsoft Purview managed identity References permission on the external table's scoped credentials. By using the References permission, Microsoft Purview can read data from external tables.

1. Run the following command in your SQL script to get the list of database scoped credentials:

   SQL

   ```
   Select name, credential_identity
   from sys.database_scoped_credentials;
   ```

2. To grant access to database scoped credentials, run the following command. Replace `scoped_credential` with the name of the database scoped credential.

   SQL

   ```
   GRANT REFERENCES ON DATABASE SCOPED CREDENTIAL::[scoped_credential] TO
   [PurviewAccountName];
   ```

3. To verify the permission assignment, run the following command in your SQL script:

   SQL

   ```
   SELECT dp.permission_name, dp.grantee_principal_id, p.name AS
   grantee_principal_name
   FROM sys.database_permissions AS dp
   JOIN sys.database_principals AS p ON dp.grantee_principal_id =
   p.principal_id
   JOIN sys.database_scoped_credentials AS c ON dp.major_id =
   c.credential_id;
   ```

Set up firewall access for the Azure Synapse Analytics workspace

1. In the Azure portal, go to the Azure Synapse Analytics workspace.
2. On the left pane, select **Networking**.
3. For **Allow Azure services and resources to access this workspace**, select **ON**.
4. Select **Save**.

If you want to scan a dedicated SQL pool (formerly SQL DW) that has enabled Azure Synapse Analytics workspace features as documented in Enable Azure Synapse workspace features for a dedicated SQL pool (formerly SQL DW), follow these additional configuration steps for the dedicated SQL pool (formerly SQL DW):

1. In the Azure portal, go to the SQL Server resource associated with the dedicated SQL pool (formerly SQL DW).
2. On the left pane, select **Networking**.
3. For **Allow Azure services and resources to access this server**, select **ON**.
4. Select **Save**.

> **Important**
>
> If you can't enable **Allow Azure services and resources to access this workspace** on your Azure Synapse Analytics workspaces, you'll get a serverless database enumeration failure when you set up a scan in the Microsoft Purview governance portal. In this case, you can choose the **Enter manually** option to specify the database names that you want to scan, and then proceed or **set up a scan using an API**.

Create and run a scan

1. In the Microsoft Purview governance portal, on the left pane, select **Data Map**.
2. Select the data source that you registered.
3. Select **View details**, and then select **New scan**. Alternatively, you can select the **Scan quick action** icon on the source tile.
4. On the **Scan** details pane, in the **Name** box, enter a name for the scan.

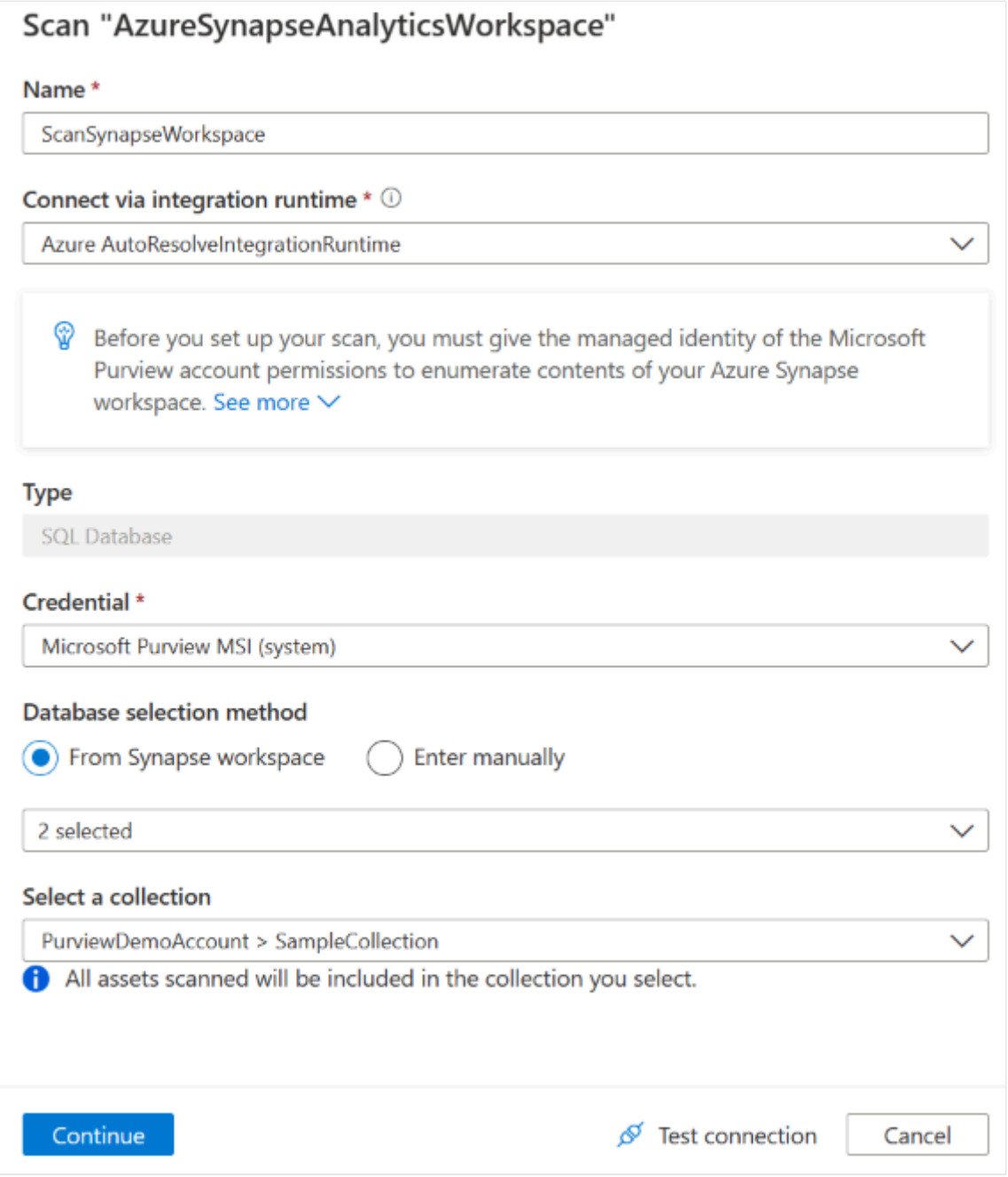
> **Note**

For the integration runtime, if you're using **Managed VNet Runtime**, make sure you create the required Managed Private Endpoints:

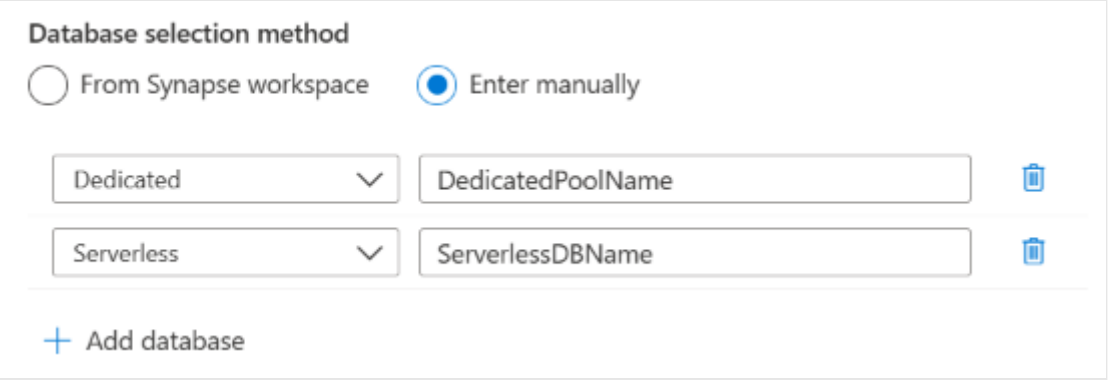
- To scan Serverless pools, create a managed private endpoint of *sqlOnDemand* sub-resource type for your Synapse workspace.
- To scan Dedicated pools, create a managed private endpoint of *sql* sub-resource type for your Synapse workspace.
- If you're scanning both Serverless and Dedicated pools, you must create both managed private endpoints and in the wizard select one.

1. In the **Credential** dropdown list, select the credential to connect to the resources within your data source.

2. For **Database selection method**, select **From Synapse workspace** or **Enter manually**. By default, Microsoft Purview tries to enumerate the databases under the workspace, and you can select the ones that you want to scan.



If you get an error that says Microsoft Purview failed to load the serverless databases, you can select **Enter manually** to specify the type of database (dedicated or serverless) and the corresponding database name.



3. Select **Test connection** to validate the settings. If you get any error, on the report page, hover over the connection status to see details.

4. Select **Continue**.
5. Select **Scan rule sets** of type **Azure Synapse SQL**. You can also create scan rule sets inline.
6. Choose your scan trigger. You can set up a schedule or run the scan once.
7. Review your scan, and then select **Save** to complete the setup.

View your scans and scan runs

To view existing scans:

1. Go to the Microsoft Purview portal. On the left pane, select **Data map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that has results you want to view. The pane shows you all the previous scan runs, along with the status and metrics for each scan run.
4. Select the run ID to check the scan run details.

Manage your scans

To edit, cancel, or delete a scan:

1. Go to the Microsoft Purview portal. On the left pane, select **Data Map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that you want to manage. You can then:
    - Edit the scan by selecting **Edit scan**.
    - Cancel an in-progress scan by selecting **Cancel scan run**.
    - Delete your scan by selecting **Delete scan**.

> ⓘ **Note**
>
> - Deleting your scan does not delete catalog assets created from previous scans.

Set up a scan by using an API

Here's an example of creating a scan for a serverless database by using the Microsoft Purview REST API. Replace the placeholders in braces (`{}`) with your actual settings. To learn more, see

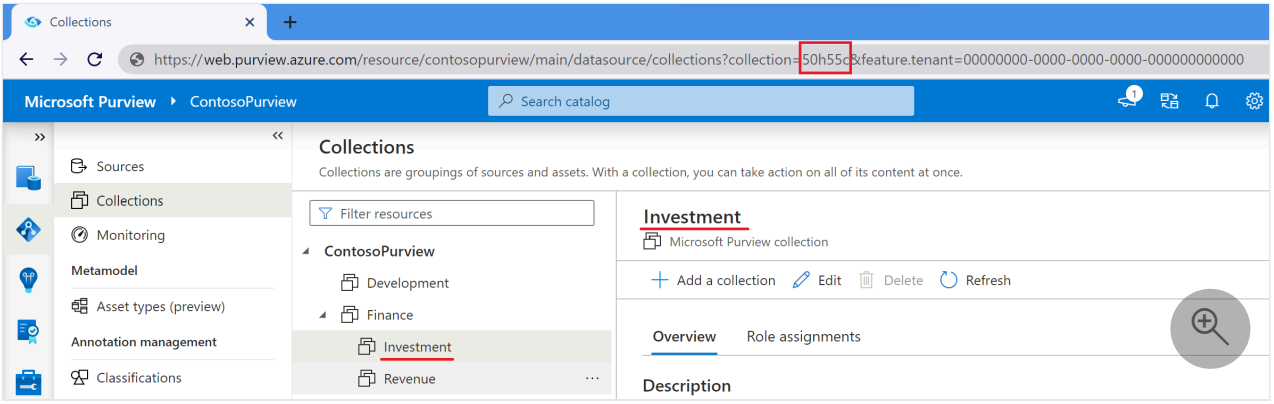
Scans - Create Or Update.

HTTP

```
PUT
https://{purview_account_name}.purview.azure.com/scan/datasources/<data_source_name
>/scans/{scan_name}?api-version=2022-02-01-preview
```

In the following code, `collection_id` isn't the friendly name for the collection. It's a five-character ID. For the root collection, `collection_id` is the name of the collection. For all subcollections, it's the ID that you can find in one of these places:

- The URL in the Microsoft Purview governance portal. Select the collection, and check the URL to find where it says **collection**=. That's your ID. In the following example, the **Investment** collection has the ID **50h55c**.



- You can list child collection names of the root collection to list the collections, and then use the name instead of the friendly name.

JSON

```
{
    "properties":{
        "resourceTypes":{
            "AzureSynapseServerlessSql":{
                "scanRulesetName":"AzureSynapseSQL",
                "scanRulesetType":"System",
                "resourceNameFilter":{
                    "resources":[ "{serverless_database_name_1}", "
{serverless_database_name_2}", ...]
                }
            }
        },
        "credential":{
            "referenceName":"{credential_name}",
            "credentialType":"SqlAuth | ServicePrincipal | ManagedIdentity (if UAMI
authentication)"
        },
```

```
        "collection":{
            "referenceName":"{collection_id}",
            "type":"CollectionReference"
        },
        "connectedVia":{
            "referenceName":"{integration_runtime_name}",
            "integrationRuntimeType":"SelfHosted (if self-hosted IR) | Managed (if
VNet IR)"
        }
    },
    "kind":"AzureSynapseWorkspaceCredential | AzureSynapseWorkspaceMsi (if system-
assigned managed identity authentication)"
}
```

To schedule the scan, create a trigger for it after scan creation. For more information, see Triggers - Create Trigger.

Troubleshooting

If you have any problems with scanning:

- Confirm that you followed all prerequisites.
- Confirm that you set up enumeration authentication for your resources.
- Confirm that you set up authentication.
- Check the network by confirming firewall settings.
- Review the scan troubleshooting documentation.

Next steps

After you register your source, use the following guides to learn more about Microsoft Purview and your data:

- Data Estate Insights in Microsoft Purview
- Lineage in Microsoft Purview
- Search the Microsoft Purview Unified Catalog

Last updated on 02/24/2026

Connect to and manage Azure Arc-enabled SQL Server in Microsoft Purview

Article • 03/21/2024

This article shows how to register an Azure Arc-enabled SQL Server instance. It also shows how to authenticate and interact with Azure Arc-enabled SQL Server in Microsoft Purview. For more information about Microsoft Purview, read the introductory article.

Supported capabilities

Scanning capabilities

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan
Yes	Yes (preview)	Yes (preview)	Yes (preview)

The supported SQL Server versions are 2012 and later. SQL Server Express LocalDB isn't supported.

When you're scanning Azure Arc-enabled SQL Server, Microsoft Purview supports extracting the following technical metadata:

- Instances
- Databases
- Schemas
- Tables, including the columns
- Views, including the columns

When you're setting up a scan, you can choose to specify the database name to scan one database. You can further scope the scan by selecting tables and views as needed. The whole Azure Arc-enabled SQL Server instance will be scanned if you don't provide a database name.

Other capabilities

View a table listing the supported capabilities for classifications, sensitivity labels, policies, data lineage, and live view.

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- An active Microsoft Purview account.
- Data Source Administrator and Data Reader permissions to register a source and manage it in the Microsoft Purview governance portal. See Access control in the Microsoft Purview governance portal for details.
- Set up the right integration runtime for your scenario:
  - **To use a self-hosted integration runtime:** Follow the article to create and configure a self-hosted integration runtime.
  - **To use a kubernetes supported self-hosted integration runtime:** Follow the article to create and configure a kubernetes supported integration runtime. (SQL authentication only.)

Networking

If your Microsoft Purview instance has disabled all public network access, to access your Arc-Enabled SQL Server, the virtual machine must be added to the virtual network that is connected to your Microsoft Purview instance.

Register

This section describes how to register an Azure Arc-enabled SQL Server instance in Microsoft Purview by using the Microsoft Purview governance portal.

Authentication for registration

There are two ways to set up authentication for scanning Azure Arc-enabled SQL Server with a self-hosted integration runtime:

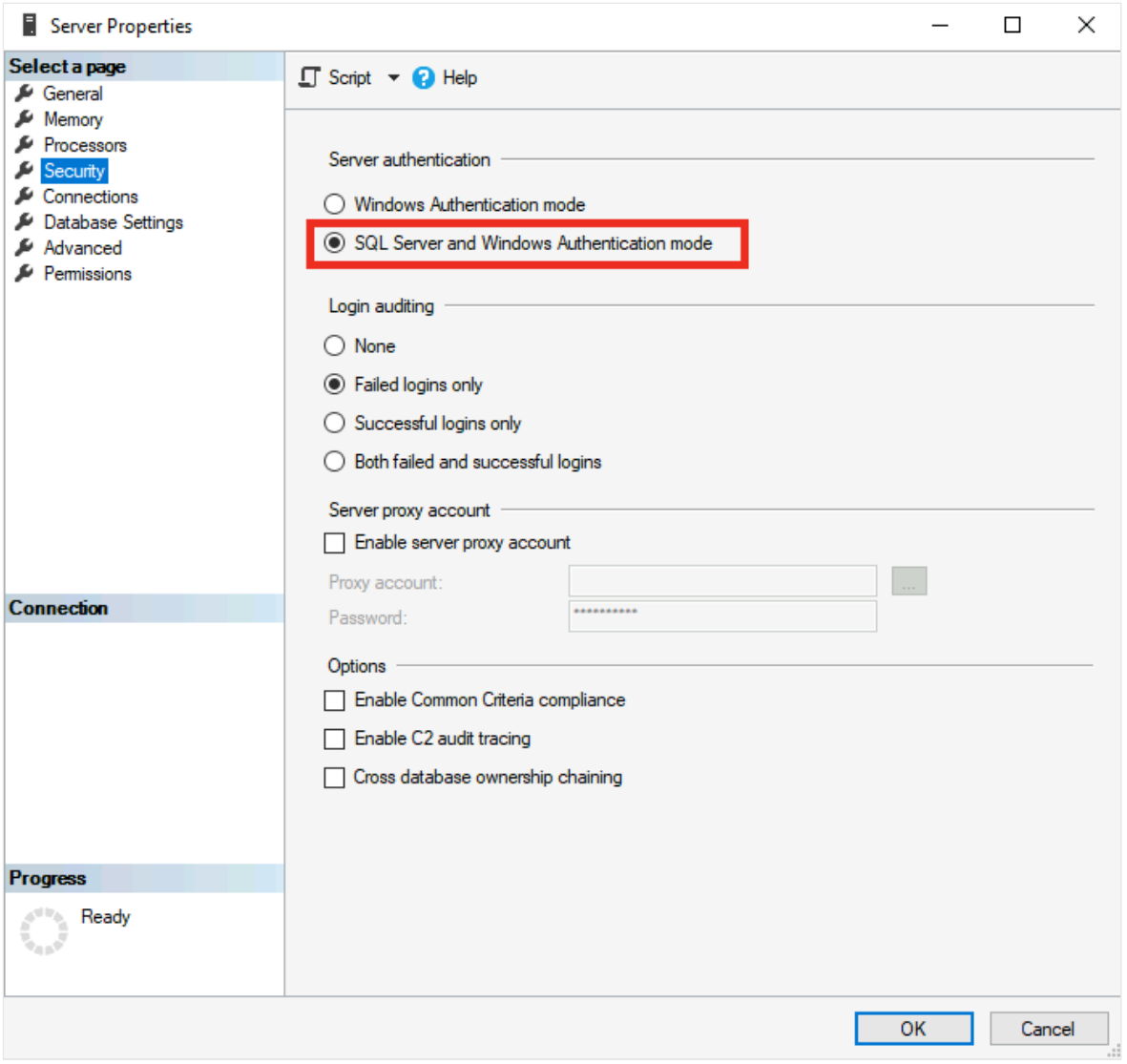
- SQL Server authentication
- Windows authentication - Not supported by a Kubernetes SHIR.

To configure authentication for the SQL Server deployment:

1. In SQL Server Management Studio (SSMS), go to **Server Properties**, and then select **Security** on the left pane.

2. Under **Server authentication**:

    - For Windows authentication, select either **Windows Authentication mode** or **SQL Server and Windows Authentication mode**.
    - For SQL Server authentication, select **SQL Server and Windows Authentication mode**.



A change to the server authentication requires you to restart the SQL Server instance and SQL Server Agent. In SSMS, go to the SQL Server instance and select **Restart** on the right-click options pane.

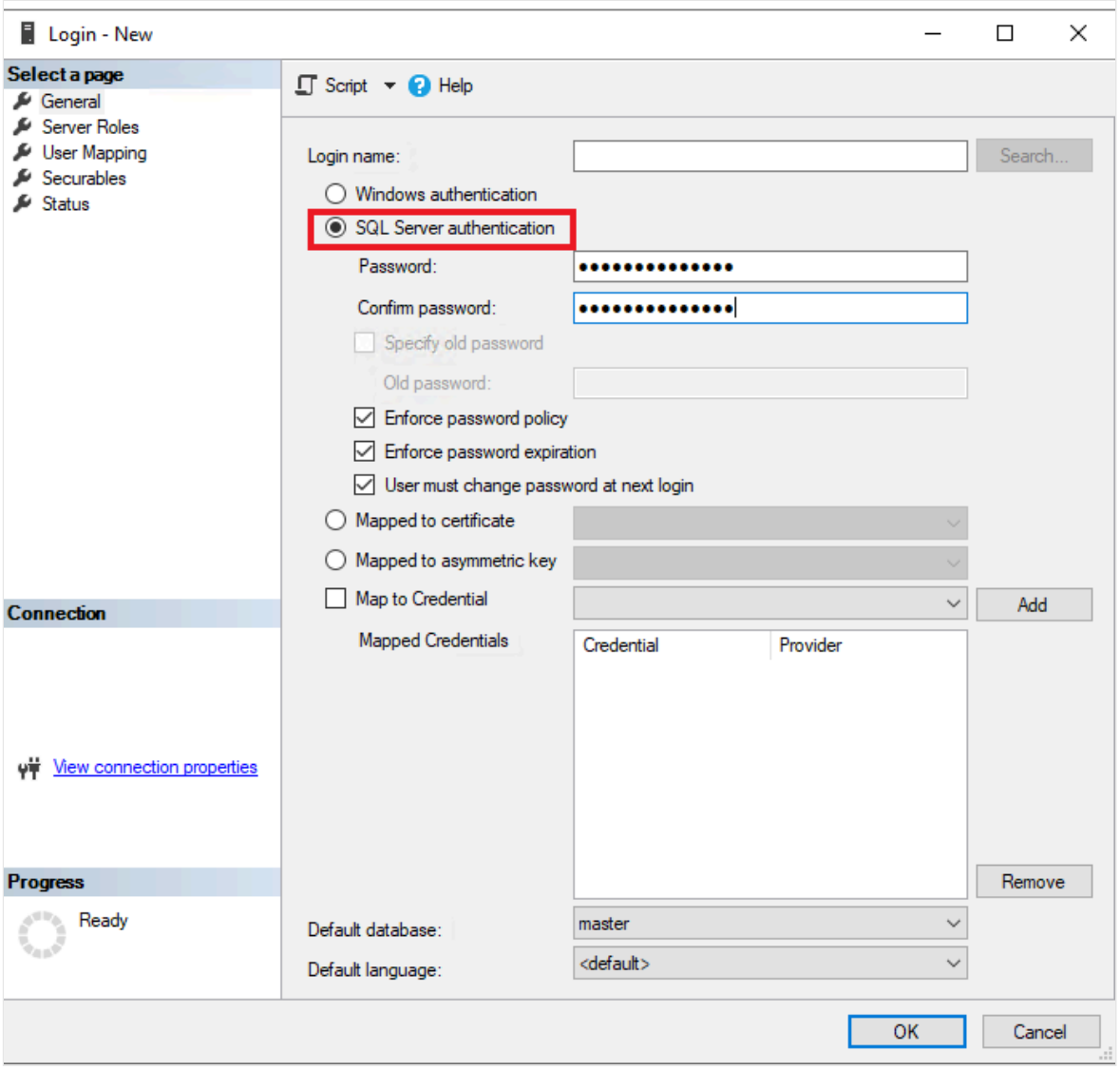
Create a new login and user

If you want to create a new login and user to scan your SQL Server instance, use the following steps.

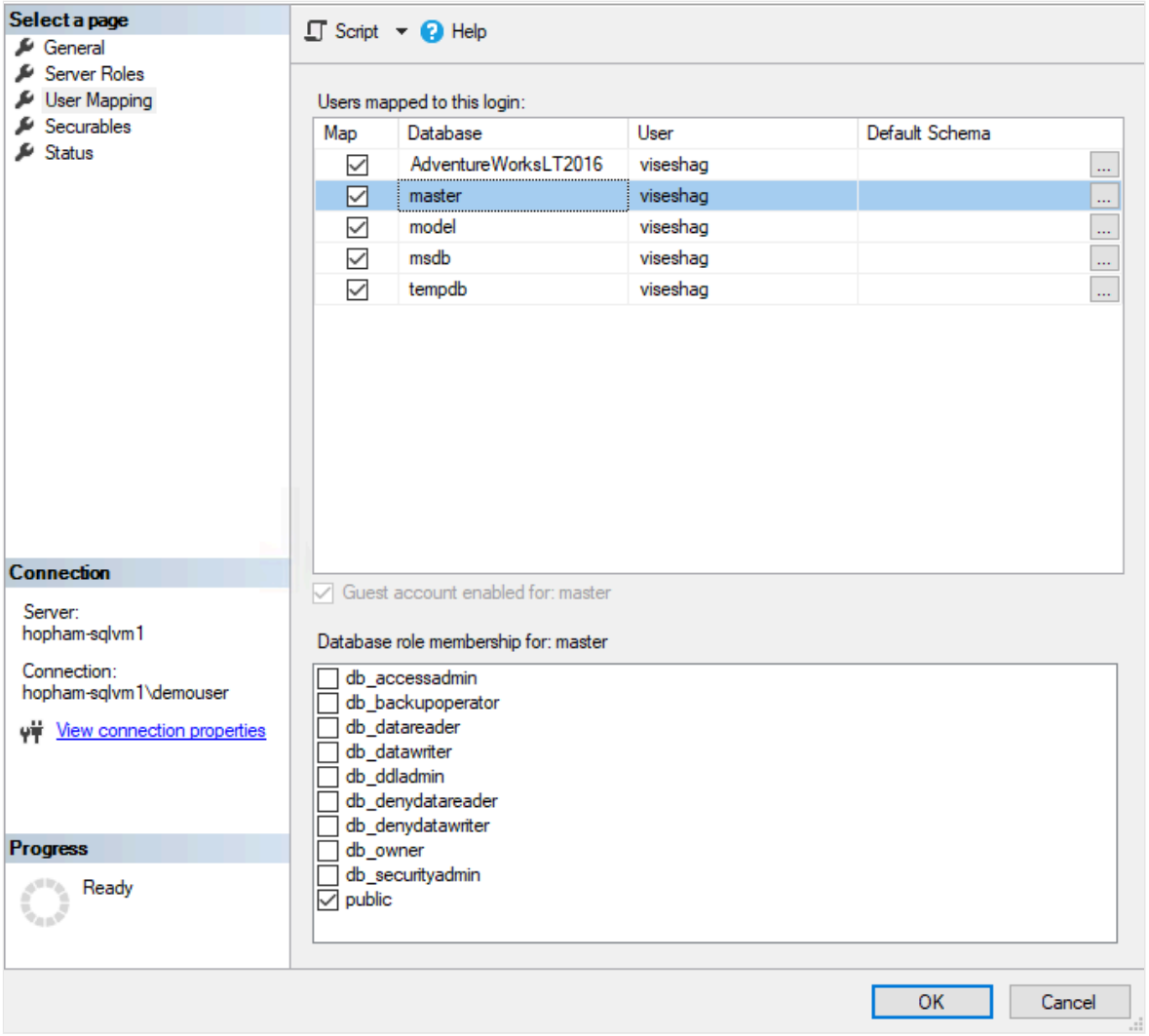
The account must have access to the master database, because `sys.databases` is in the master database. The Microsoft Purview scanner needs to enumerate `sys.databases` in order to find all the SQL databases on the server.

> **Note**
>
> You can run all the following steps by using this code.

1. Go to SSMS, connect to the server, and then select **Security** on the left pane.
2. Select and hold (or right-click) **Login**, and then select **New login**. If Windows authentication is applied, select **Windows authentication**. If SQL Server authentication is applied, select **SQL Server authentication**.

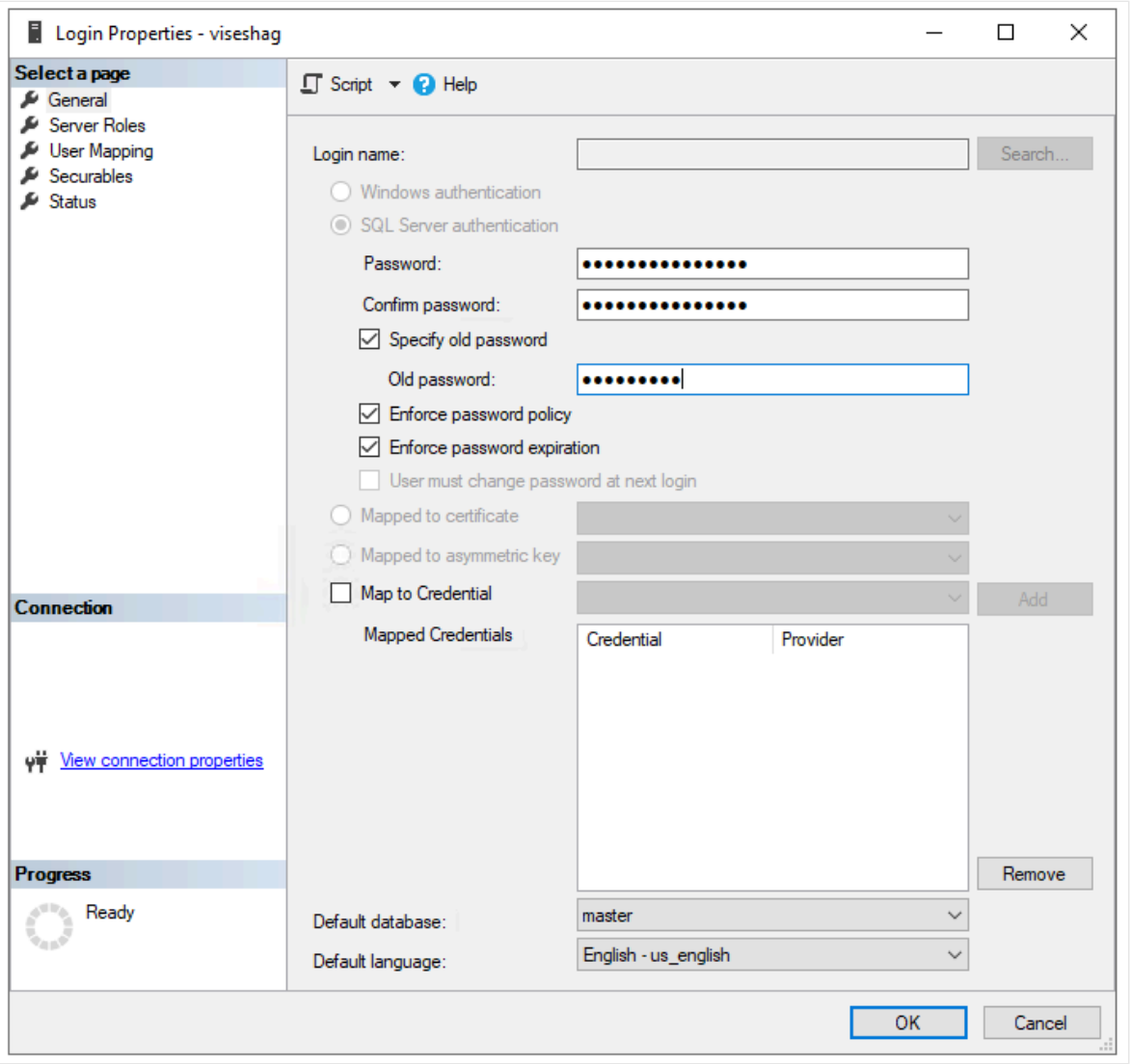


3. Select **Server Roles** on the left pane, and ensure that a public role is assigned.
4. Select **User Mapping** on the left pane, select all the databases in the map, and then select the **db_datareader** database role.



5. Select **OK** to save.

6. If SQL Server authentication is applied, you must change your password as soon as you create a new login:
   a. Select and hold (or right-click) the user that you created, and then select **Properties**.
   b. Enter a new password and confirm it.
   c. Select the **Specify old password** checkbox and enter the old password.
   d. Select **OK**.



Store your SQL Server login password in a key vault and create a credential in Microsoft Purview

1. Go to your key vault in the Azure portal. Select **Settings** > **Secrets**.

2. Select **+ Generate/Import**. For **Name** and **Value**, enter the password from your SQL Server login.

3. Select **Create**.

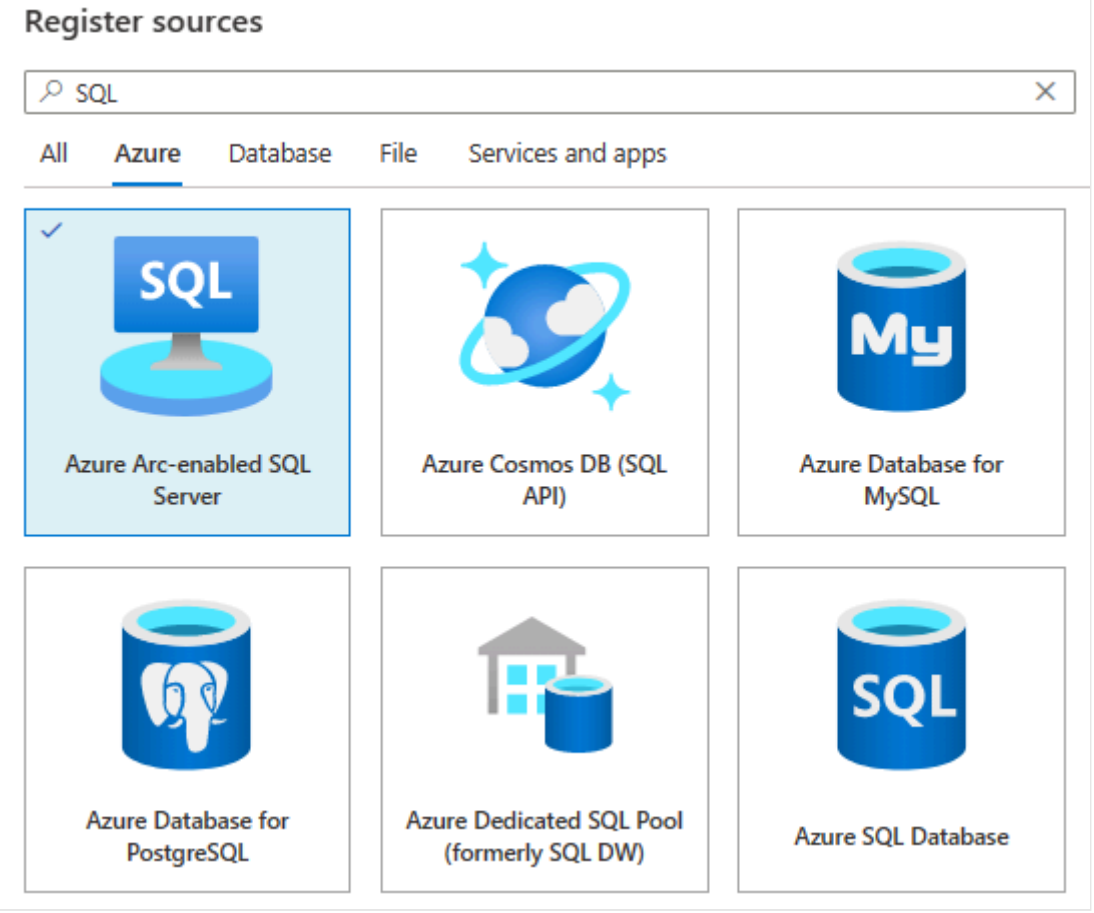
4. If your key vault isn't connected to Microsoft Purview yet, create a new key vault connection.

5. Create a new credential by using the username and password to set up your scan.

   Be sure to select the right authentication method when you're creating a new credential. If Windows authentication is applied, select **Windows authentication**. If SQL Server authentication is applied, select **SQL Server authentication**.

Steps to register

1. Go to your Microsoft Purview account.
2. Under **Sources and scanning** on the left pane, select **Integration runtimes**. Make sure that a self-hosted integration runtime is set up. If it isn't set up, follow the steps to create a self-hosted integration runtime for scanning on an on-premises or Azure virtual machine that has access to your on-premises network.
3. Select **Data Map** on the left pane.
4. Select **Register**.
5. Select **Azure Arc-enabled SQL Server**, and then select **Continue**.



6. Provide a friendly name, which is a short name that you can use to identify your server. Also provide the server endpoint.
7. Select **Finish** to register the data source.

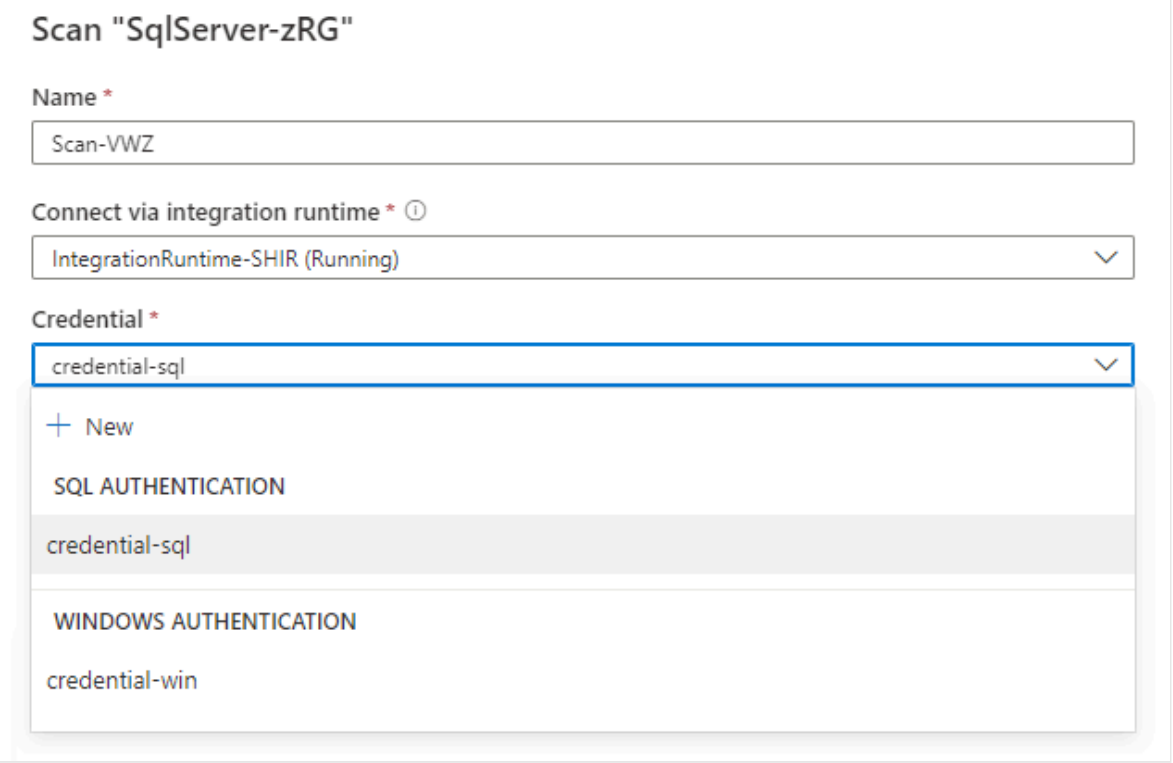
Scan

Use the following steps to scan Azure Arc-enabled SQL Server instances to automatically identify assets and classify your data. For more information about scanning in general,

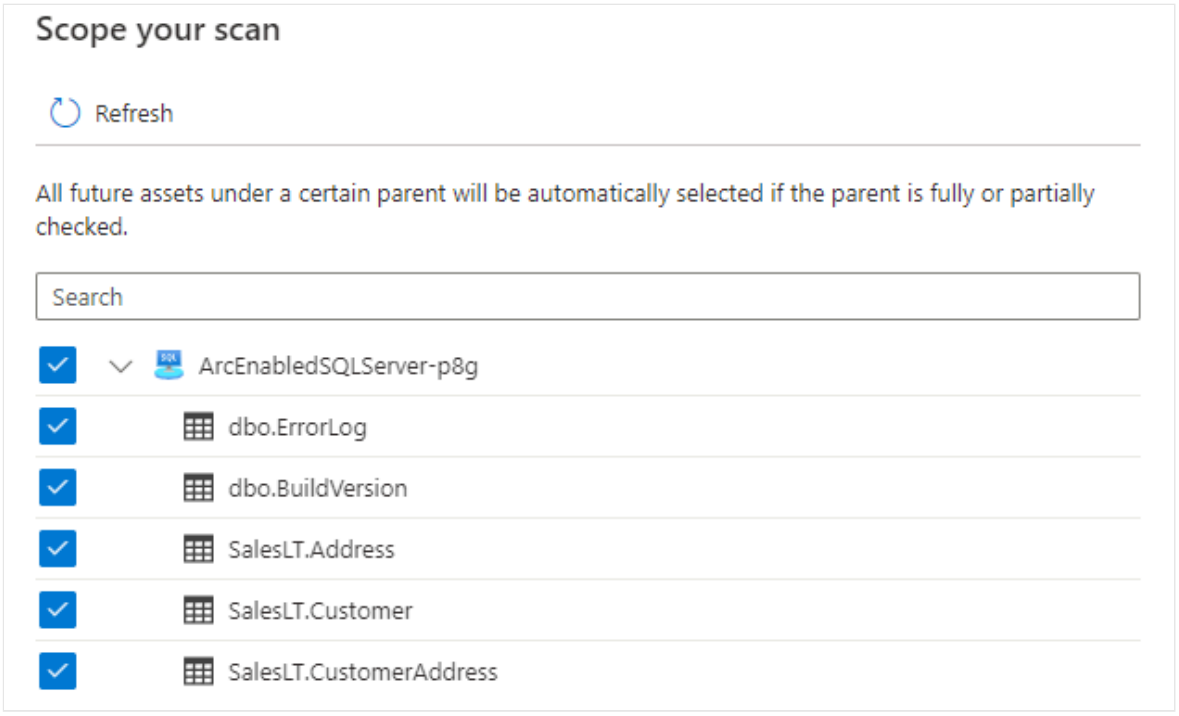
see Scans and ingestion in Microsoft Purview.

To create and run a new scan:

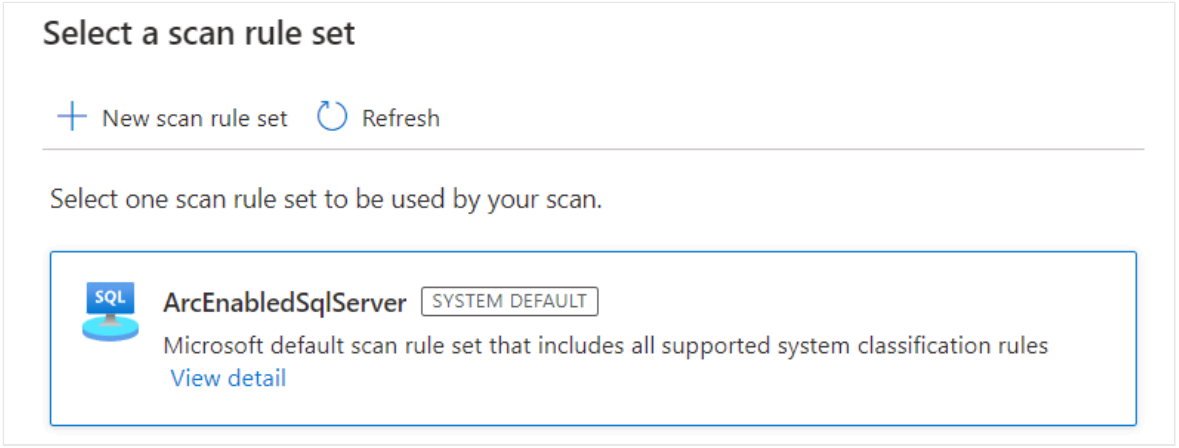
1. In the Microsoft Purview governance portal, select the **Data Map** tab on the left pane.
2. Select the Azure Arc-enabled SQL Server source that you registered.
3. Select **New scan**.
4. Select the credential to connect to your data source. Credentials are grouped and listed under the authentication methods.



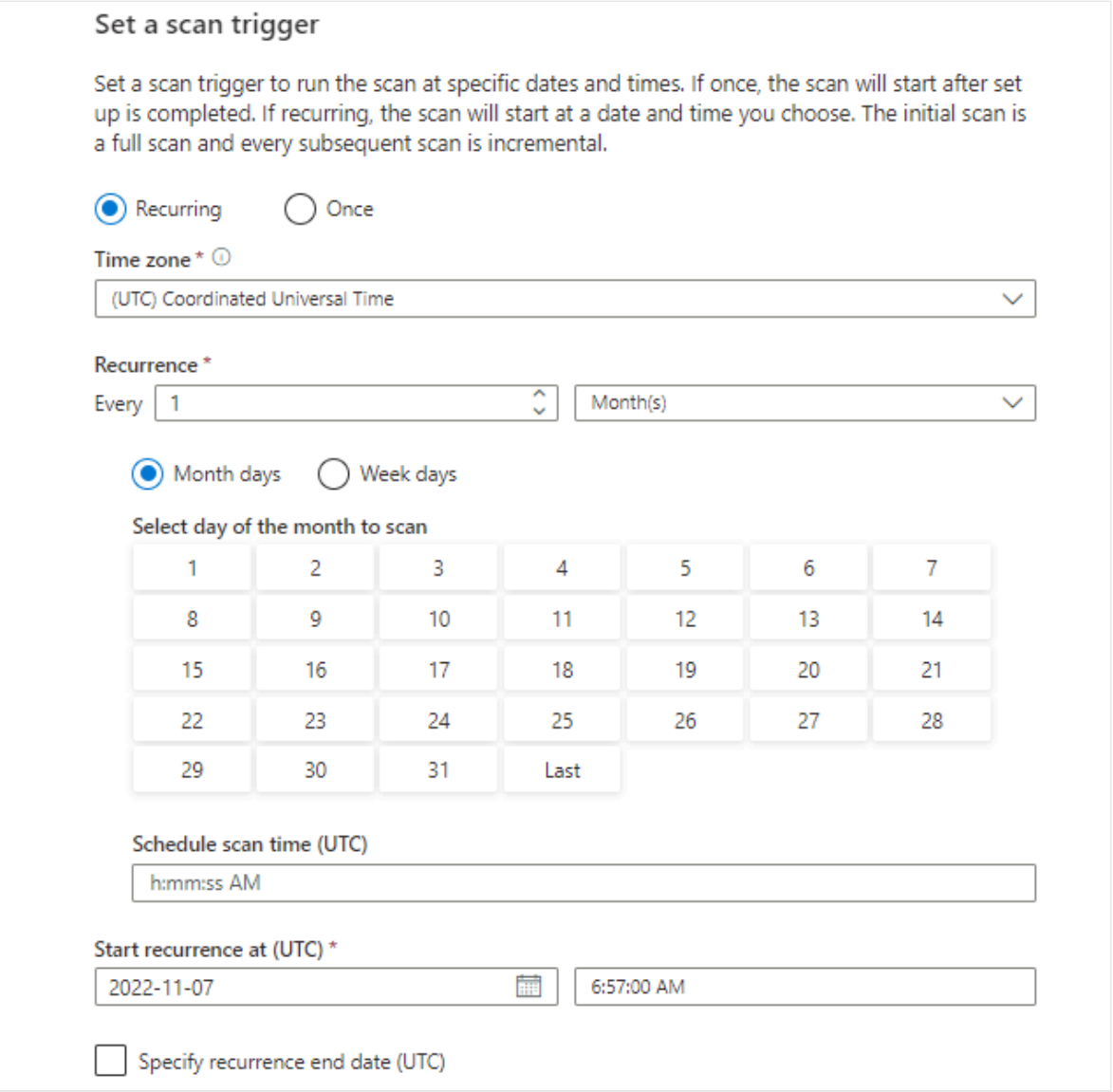
5. You can scope your scan to specific tables by choosing the appropriate items in the list.



6. Select a scan rule set. You can choose between the system default, existing custom rule sets, or creation of a new rule set inline.



7. Choose your scan trigger. You can set up a schedule or run the scan once.



8. Review your scan, and then select **Save and run**.

View your scans and scan runs

To view existing scans:

1. Go to the Microsoft Purview portal. On the left pane, select **Data map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that has results you want to view. The pane shows you all the previous scan runs, along with the status and metrics for each scan run.
4. Select the run ID to check the scan run details.

Manage your scans

To edit, cancel, or delete a scan:

1. Go to the Microsoft Purview portal. On the left pane, select **Data Map**.

2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.

3. Select the scan that you want to manage. You can then:

    - Edit the scan by selecting **Edit scan**.
    - Cancel an in-progress scan by selecting **Cancel scan run**.
    - Delete your scan by selecting **Delete scan**.

> ⓘ **Note**
>
> - Deleting your scan does not delete catalog assets created from previous scans.

Access policy

Supported policies

The following types of policies are supported on this data resource from Microsoft Purview:

- DevOps policies
- Data Owner policies (preview)

Access policy prerequisites on Azure Arc-enabled SQL Server

- Get SQL Server on-premises version 2022 and install it. Versions 2022 or more recent are supported on both Windows and Linux. You can try the free Developer edition.
- Configure your permissions and then register a list of resource providers in the subscription you'll use to onboard the SQL Server instance to Azure Arc.
- Complete prerequisites and onboard the SQL Server instance with Azure Arc. A simple configuration for Windows SQL Server is here. An alternate configuration for Linux SQL Server is here.
- Enable Microsoft Entra authentication in SQL Server. For a simpler setup, complete the prerequisites and process outlined in this article.
- Remember to grant application permissions and grant admin consent
- You need to configure a Microsoft Entra admin for the SQL Server instance, but it isn't necessary to set up other Microsoft Entra logins or users. You'll grant access to

those users by using Microsoft Purview policies.

Region support

Data policy enforcement is available in all Microsoft Purview regions except:

- West US2
- East Asia
- US Gov Virginia
- China North 3

Security considerations for Azure Arc-enabled SQL Server

- The server admin can turn off the Microsoft Purview policy enforcement.
- Azure Arc admin and server admin permissions provide the ability to change the Azure Resource Manager path of the server. Because mappings in Microsoft Purview use Resource Manager paths, it can lead to wrong policy enforcements.
- A SQL Server admin (database admin) can gain the power of a server admin and can tamper with the cached policies from Microsoft Purview.
- The recommended configuration is to create a separate app registration for each SQL server instance. This configuration prevents the second SQL Server instance from reading the policies meant for the first SQL Server instance, in case a rogue admin in the second SQL Server instance tampers with the Resource Manager path.

Verify the prerequisites

1. Sign in to the Azure portal through this link
2. Navigate to **SQL servers** on the left pane. You'll see a list of SQL Server instances on Azure Arc.
3. Select the SQL Server instance that you want to configure.
4. Go to **Microsoft Entra ID** on the left pane.
5. Ensure that Microsoft Entra authentication is configured with an admin login. If not, refer to the access policy prerequisites section in this guide.
6. Ensure that a certificate has been provided to for SQL Server to authenticate to Azure. If not, refer to the access policy prerequisites section in this guide.

7. Ensure that an app registration has been entered to create a trust relationship between SQL Server and Microsoft Entra ID. If not, refer to the access policy prerequisites section in this guide.

8. If you made any changes, select the **Save** button to save the configuration and wait until the operation completes successfully. This could take a few minutes. The message "*Saved successfully*" will be displayed at the top of the page in green background. You might need to scroll up to see it.

Configure the Microsoft Purview account for policies

Register the data source in Microsoft Purview

Before a policy can be created in Microsoft Purview for a data resource, you must register that data resource in Microsoft Purview Studio. You will find the instructions related to registering the data resource later in this guide.

> ⓘ **Note**
>
> Microsoft Purview policies rely on the data resource ARM path. If a data resource is moved to a new resource group or subscription it will need to be de-registered and then registered again in Microsoft Purview.

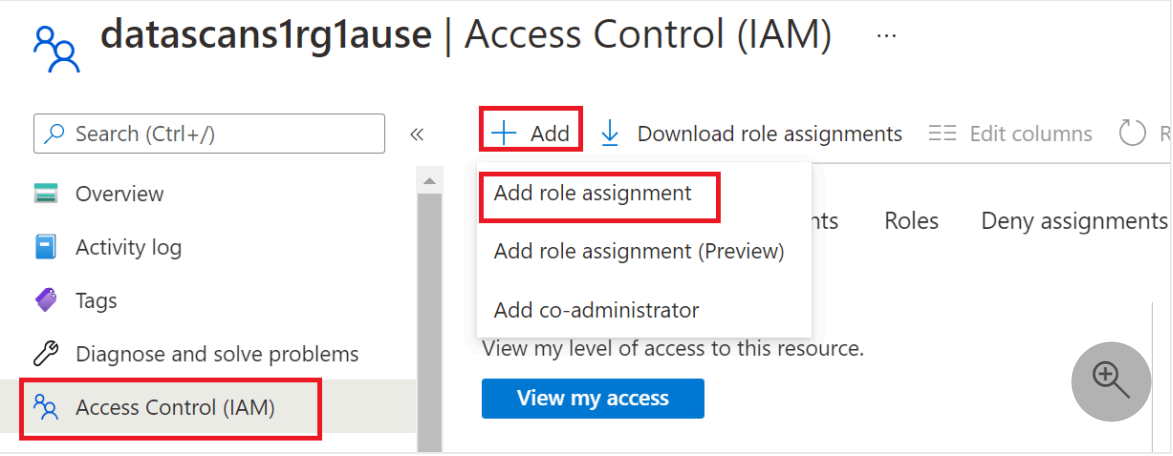
Configure permissions to enable Data policy enforcement on the data source

Once a resource is registered, but before a policy can be created in Microsoft Purview for that resource, you must configure permissions. A set of permissions are needed to enable the **Data policy enforcement**. This applies to data sources, resource groups, or subscriptions. To enable *Data policy enforcement*, you must have *both* specific Identity and Access Management (IAM) privileges on the resource as well as specific Microsoft Purview privileges:

- You must have *either one* of the following IAM role combinations on the resource's Azure Resource Manager path or any parent of it (that is, using IAM permission inheritance):
  - IAM Owner
  - Both IAM Contributor and IAM User Access Administrator

  To configure Azure role-based access control (RBAC) permissions, follow this guide. The following screenshot shows how to access the **Access Control** section in the

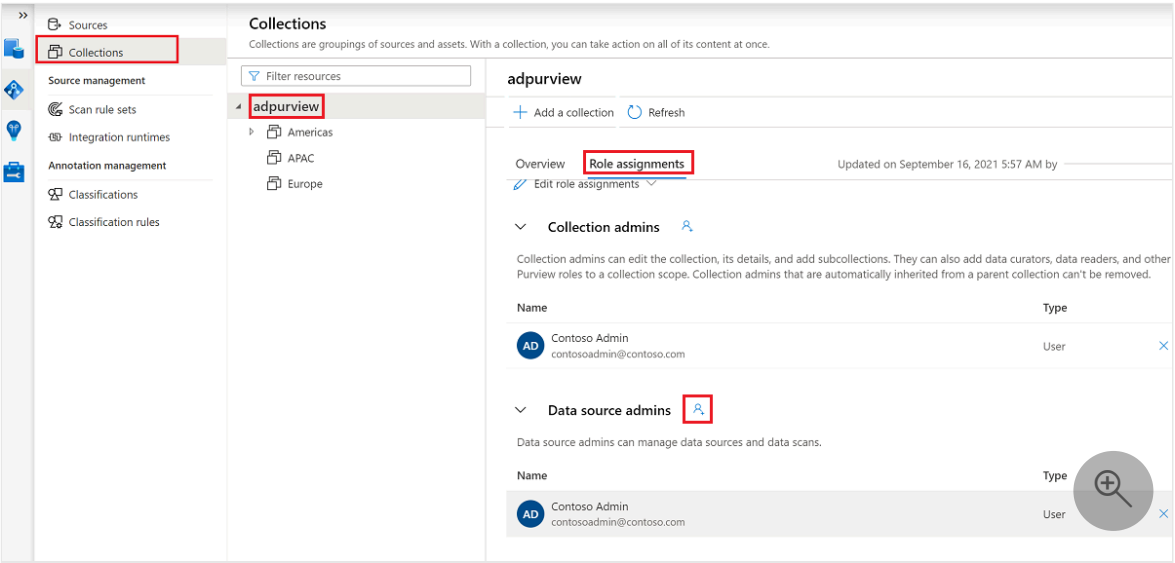
Azure portal for the data resource to add a role assignment.



> **Note**
>
> The *IAM Owner* role for a data resource can be inherited from a parent resource group, a subscription, or a subscription management group. Check which Microsoft Entra users, groups, and service principals hold or are inheriting the *IAM Owner* role for the resource.

- You also need to have the Microsoft Purview *Data source admin* role for the collection or a parent collection (if inheritance is enabled). For more information, see the guide on managing Microsoft Purview role assignments.

  The following screenshot shows how to assign the *Data source admin* role at the root collection level.



Configure Microsoft Purview permissions to create, update, or delete access policies

To create, update or delete policies, you need to get the *Policy author* role in Microsoft Purview at *root collection level*:

- The *Policy author* role can create, update, and delete DevOps and Data Owner policies.
- The *Policy author* role can delete self-service access policies.

For more information about managing Microsoft Purview role assignments, see Create and manage collections in the Microsoft Purview Data Map.

> ⓘ **Note**
>
> Policy author role must be configured at the root collection level.

In addition, to easily search Microsoft Entra users or groups when creating or updating the subject of a policy, you can greatly benefit from getting the Directory Readers permission in Microsoft Entra ID. This is a common permission for users in an Azure tenant. Without the Directory Reader permission, the Policy Author will have to type the complete username or email for all the principals included in the subject of a data policy.

Configure Microsoft Purview permissions for publishing Data Owner policies

Data Owner policies allow for checks and balances if you assign the Microsoft Purview *Policy author* and *Data source admin* roles to different people in the organization. Before a Data owner policy takes effect, a second person (*Data source admin*) must review it and explicitly approve it by publishing it. This does not apply to DevOps or Self-service access policies as publishing is automatic for them when those policies are created or updated.

To publish a Data owner policy you need to get the *Data source admin* role in Microsoft Purview at *root collection level*.

For more information about managing Microsoft Purview role assignments, see Create and manage collections in the Microsoft Purview Data Map.

> ⓘ **Note**
>
> To publish Data owner policies, the Data source admin role must be configured at the root collection level.

Delegate access provisioning responsibility to roles in Microsoft Purview

After a resource has been enabled for **Data policy enforcement**, any Microsoft Purview user with the *Policy author* role at the root collection level can provision access to that data source from Microsoft Purview.

> **Note**
>
> Any Microsoft Purview root *Collection admin* can assign new users to root *Policy author* roles. Any *Collection admin* can assign new users to a *Data source admin* role under the collection. Minimize and carefully vet the users who hold Microsoft Purview *Collection admin*, *Data source admin*, or *Policy author* roles.

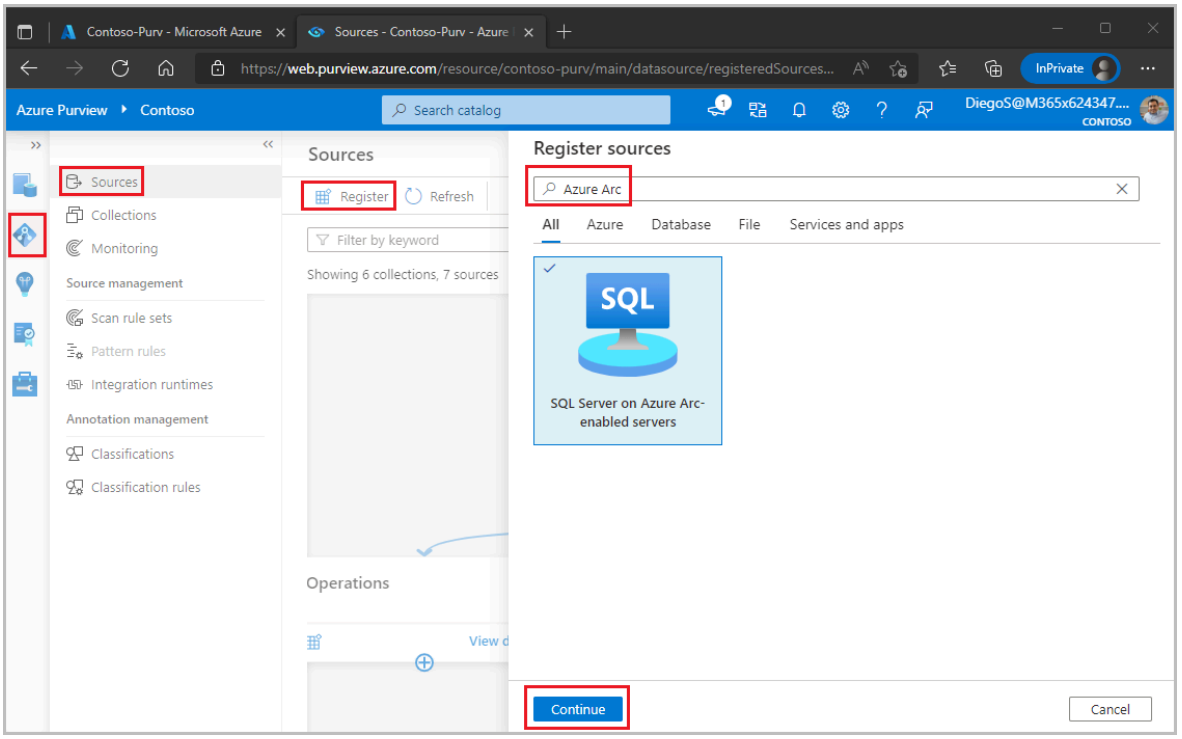
If a Microsoft Purview account with published policies is deleted, such policies will stop being enforced within an amount of time that depends on the specific data source. This change can have implications on both security and data access availability. The Contributor and Owner roles in IAM can delete Microsoft Purview accounts. You can check these permissions by going to the **Access control (IAM)** section for your Microsoft Purview account and selecting **Role Assignments**. You can also use a lock to prevent the Microsoft Purview account from being deleted through Resource Manager locks.

Register the data source and enable Data policy enforcement

Before you can create policies, you must register the Azure Arc-enabled SQL Server data source with Microsoft Purview:

1. Sign in to Microsoft Purview Studio.

2. Go to **Data map** on the left pane, select **Sources**, and then select **Register**. Enter **Azure Arc** in the search box and select **SQL Server on Azure Arc**. Then select **Continue**.



3. For **Name**, enter a name for this registration. It's best practice to make the name of the registration the same as the server name in the next step.

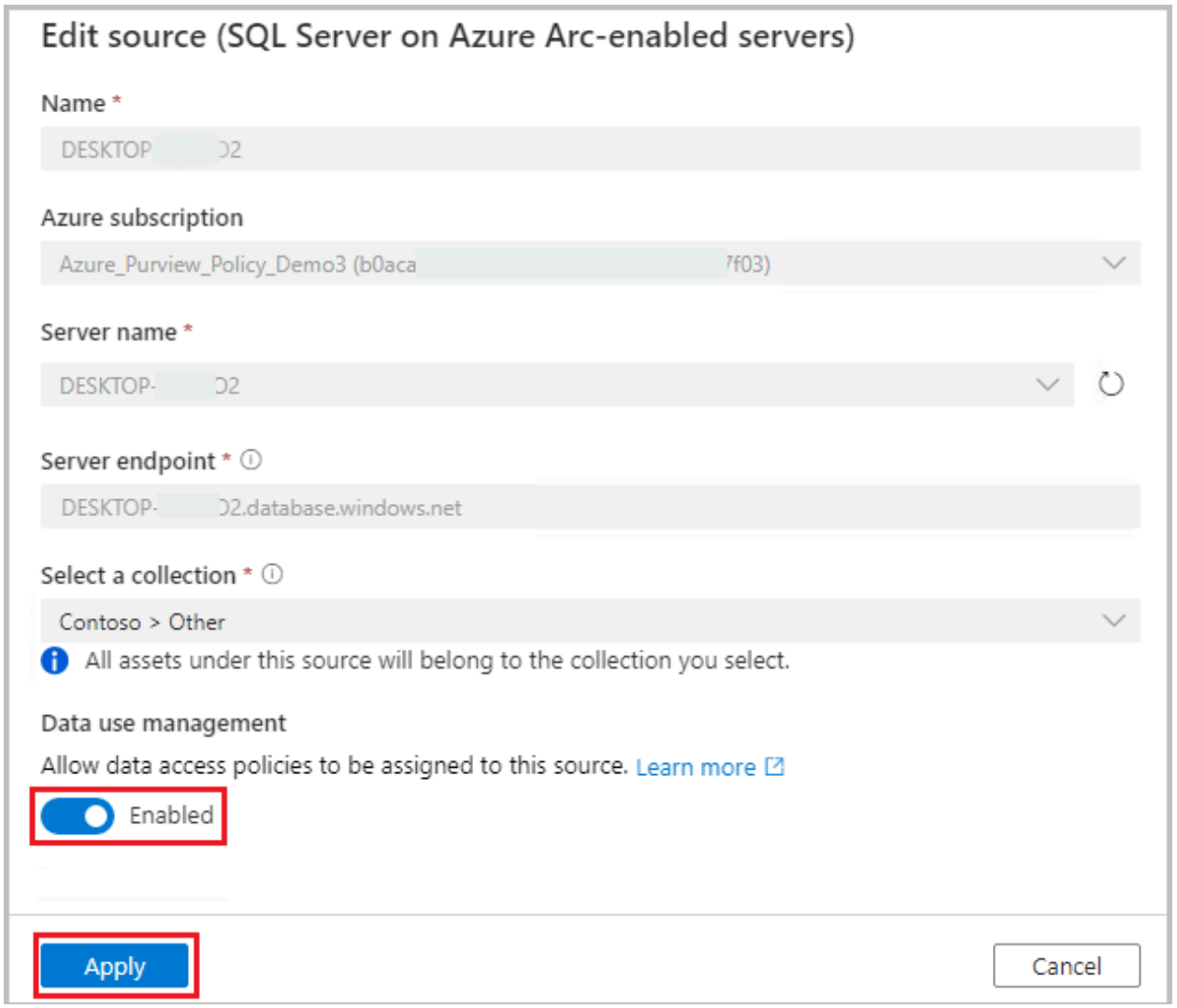
4. Select values for **Azure subscription**, **Server name**, and **Server endpoint**.

5. For **Select a collection**, choose a collection to put this registration in.

6. Enable **Data policy enforcement**. **Data policy enforcement** needs certain permissions and can affect the security of your data, because it delegates to certain Microsoft Purview roles to manage access to the data sources. Go through the secure practices related to **Data policy enforcement** in this guide: Enable Data policy enforcement on your Microsoft Purview sources.

7. Upon enabling Data Policy Enforcement, Microsoft Purview will automatically capture the **Application ID** of the App Registration related to this Azure Arc-enabled SQL Server if one has been configured. Come back to this screen and hit the refresh button on the side of it to refresh, in case the association between the Azure Arc-enabled SQL Server and the App Registration changes in the future.

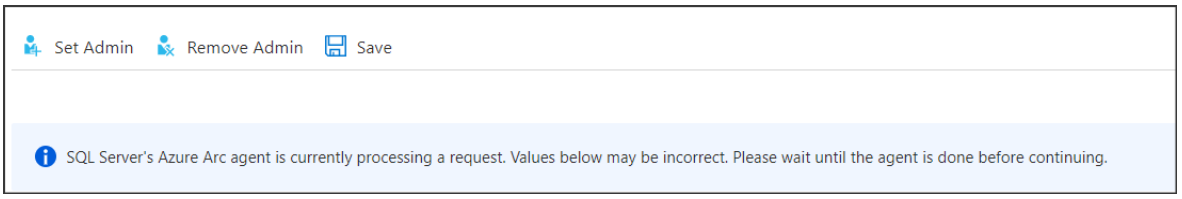
8. Select **Register** or **Apply**.



Enable policies in Azure Arc-enabled SQL Server

This section describes the steps to configure SQL Server on Azure Arc to use Microsoft Purview. Execute these steps after you enable the *Data policy enforcement* option for this data source in the Microsoft Purview account.

1. Sign in to the Azure portal through this link
2. Navigate to **SQL servers** on the left pane. You'll see a list of SQL Server instances on Azure Arc.
3. Select the SQL Server instance that you want to configure.
4. Go to **Microsoft Entra ID** on the left pane.
5. Scroll down to **Microsoft Purview access policies**.
6. Select the button to **Check for Microsoft Purview Governance**. Wait while the request is processed. While that happens, this message will be displayed at the top of the page. You might need to scroll up to see it.



7. At the bottom of the page, confirm that the Microsoft Purview Governance Status shows `Governed`. Note that **it may take up to 30 minutes** for the correct status to be reflected. Continue doing a browser refresh until that happens.

8. Confirm that the Microsoft Purview Endpoint points to the Microsoft Purview account where you registered this data source and enabled the *data policy enforcement*

Create a policy

To create an access policy for Azure Arc-enabled SQL Server, follow these guides:

- Provision access to system health, performance and audit information in SQL Server 2022
- Provision read/modify access on a single SQL Server 2022

To create policies that cover all data sources inside a resource group or Azure subscription, see Discover and govern multiple Azure sources in Microsoft Purview.

Next steps

Now that you've registered your source, use the following guides to learn more about Microsoft Purview and your data:

- Concepts for Microsoft Purview Data owner policies
- Concepts for Microsoft Purview DevOps policies
- Data Estate Insights in Microsoft Purview
- Lineage in Microsoft Purview
- Search the Unified Catalog

Feedback

Was this page helpful? Yes No

Connect to Azure Blob storage in Microsoft Purview

This article outlines the process to register and govern Azure Blob Storage accounts in Microsoft Purview including instructions to authenticate and interact with the Azure Blob Storage source

Supported capabilities

Scanning capabilities

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan
Yes	Yes	Yes	Yes

For file types such as csv, tsv, psv, ssv, the schema is extracted when the following logics are in place:

- First row values are non-empty
- First row values are unique
- First row values aren't a date or a number

Other capabilities

For **classifications**, **sensitivity labels**, **policies**, **data lineage**, and **live view**, see the list of supported capabilities.

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- An active Microsoft Purview account.
- You'll need to be a Data Source Administrator and Data Reader to register a source and manage it in the Microsoft Purview governance portal. See our Microsoft Purview Permissions page for details.

** Lineage is supported if dataset is used as a source/sink in Data Factory Copy activity

> **Note**
>
> Avoid using an FNS file system design with Blob Storage, as it causes only folders to be ingested as blobs and excludes subfolders and files. Instead, use Microsoft Azure Data Lake Storage Gen2 with hierarchical namespace to ensure complete ingestion of folders and their contents.

Register

This section will enable you to register the Azure Blob storage account for scan and data share in Purview.

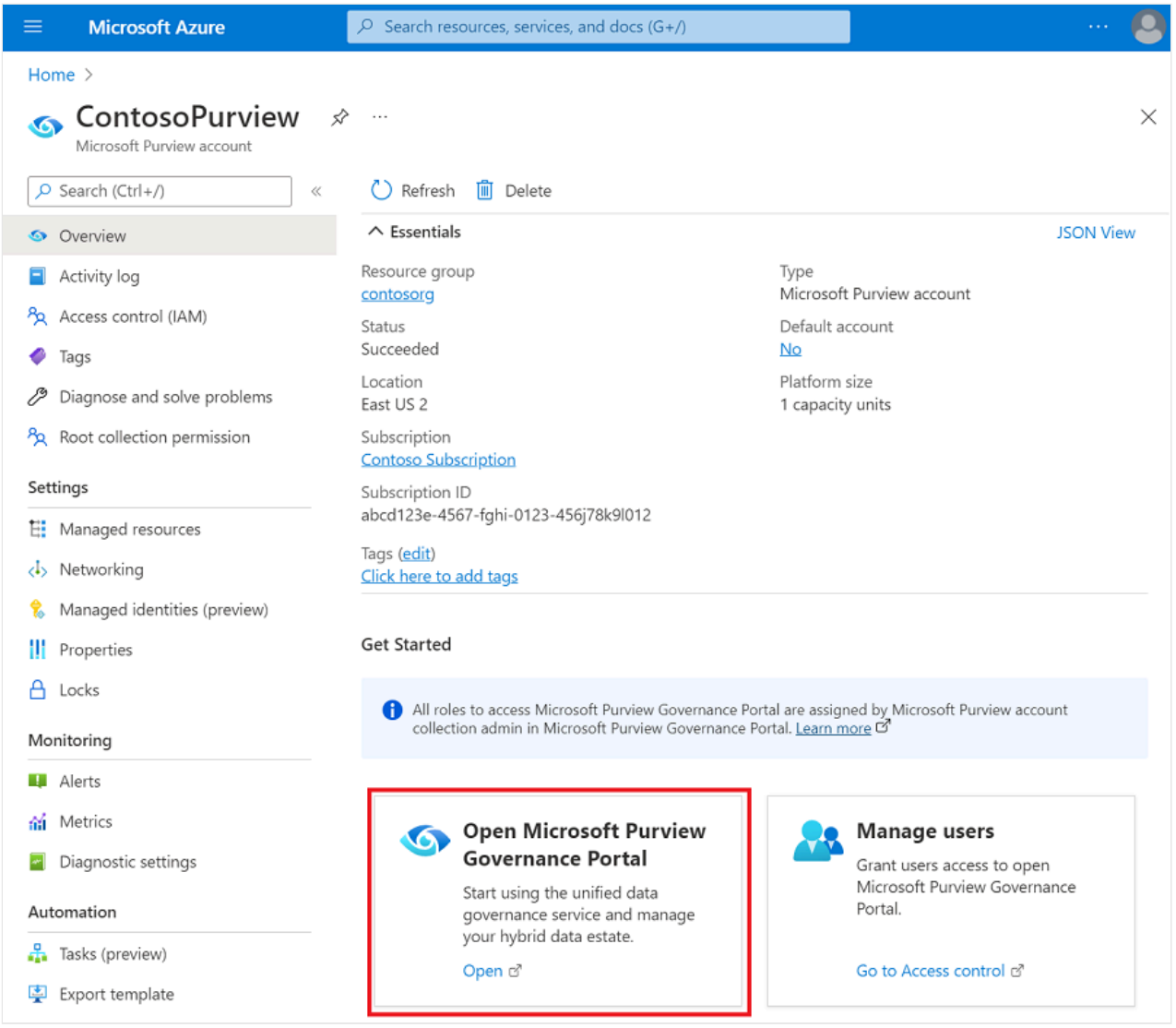
Prerequisites for register

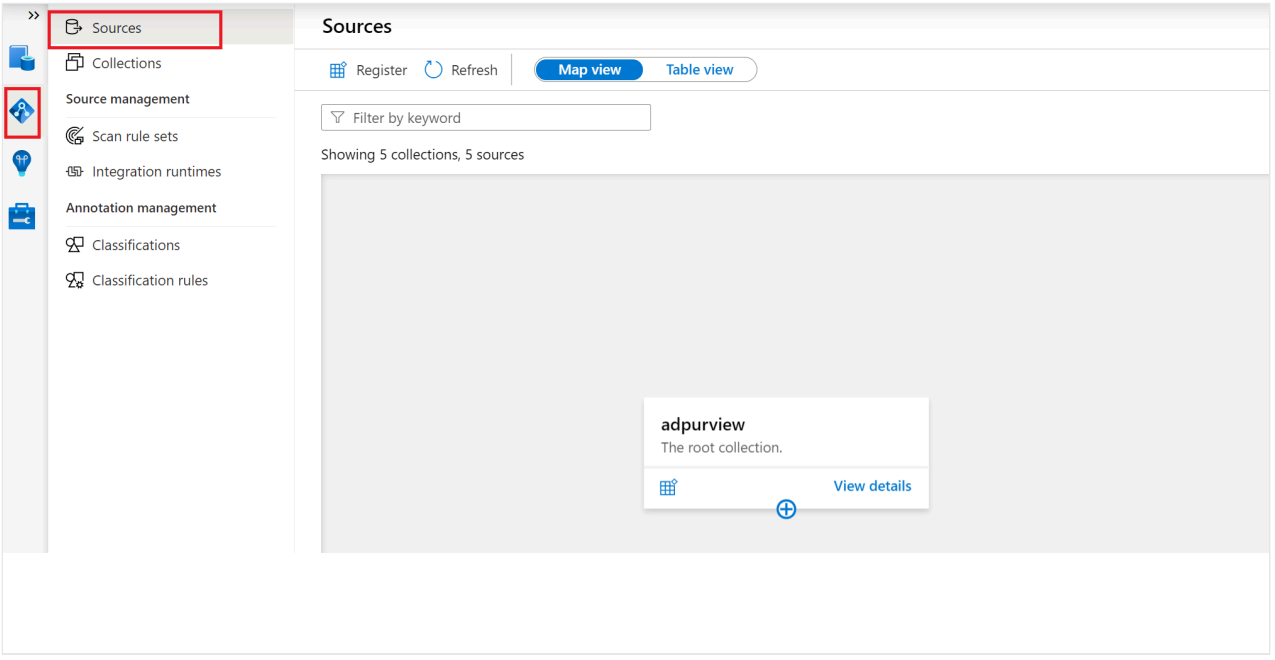
- You'll need to be a Data Source Admin and one of the other Purview roles (for example, Data Reader or Data Share Contributor) to register a source and manage it in the Microsoft Purview governance portal. See our Microsoft Purview Permissions page for details.

Steps to register

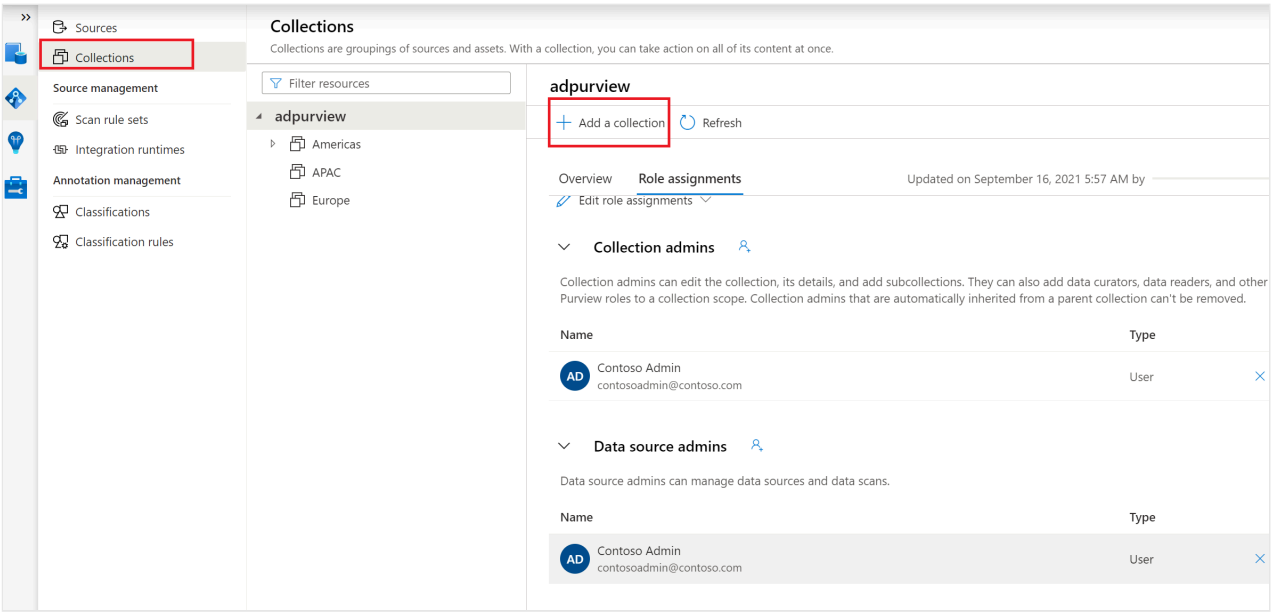
It is important to register the data source in Microsoft Purview prior to setting up a scan for the data source.

1. Go to the Microsoft Purview governance portal by:
    - Browsing directly to https://web.purview.azure.com and selecting your Microsoft Purview account.
    - Opening the Azure portal, searching for and selecting the Microsoft Purview account. Select the **the Microsoft Purview governance portal** button.
2. Navigate to the **Data Map --> Sources**

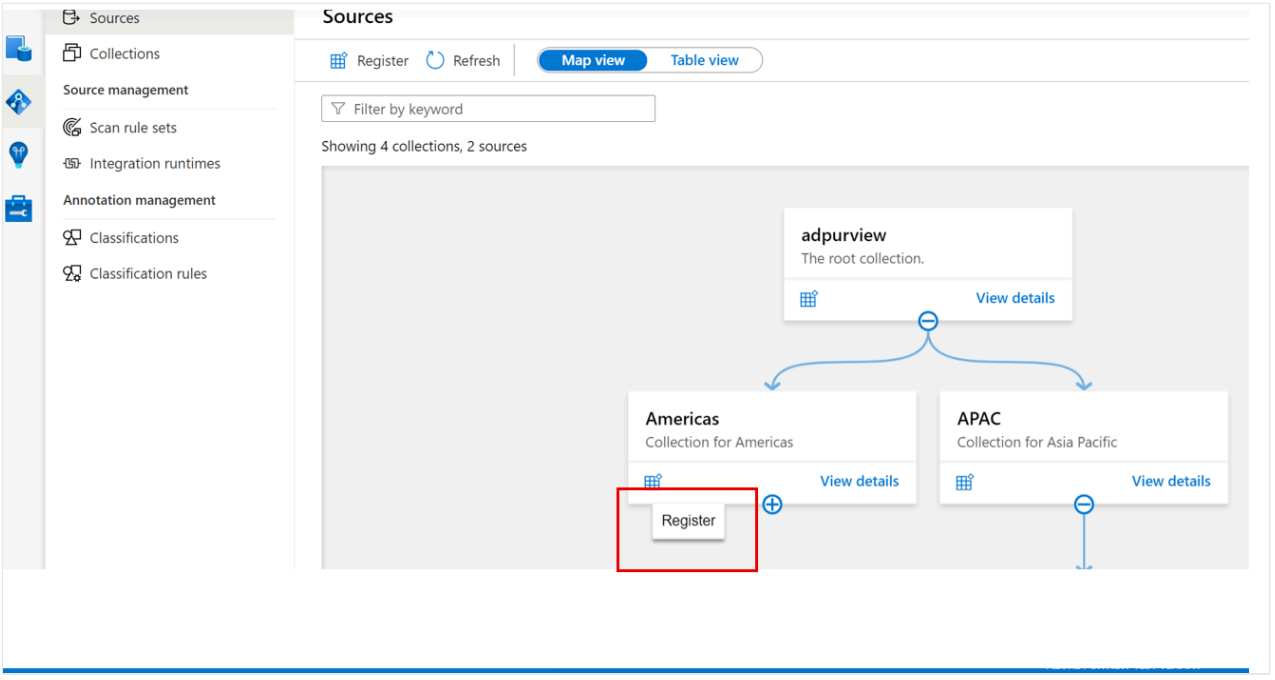




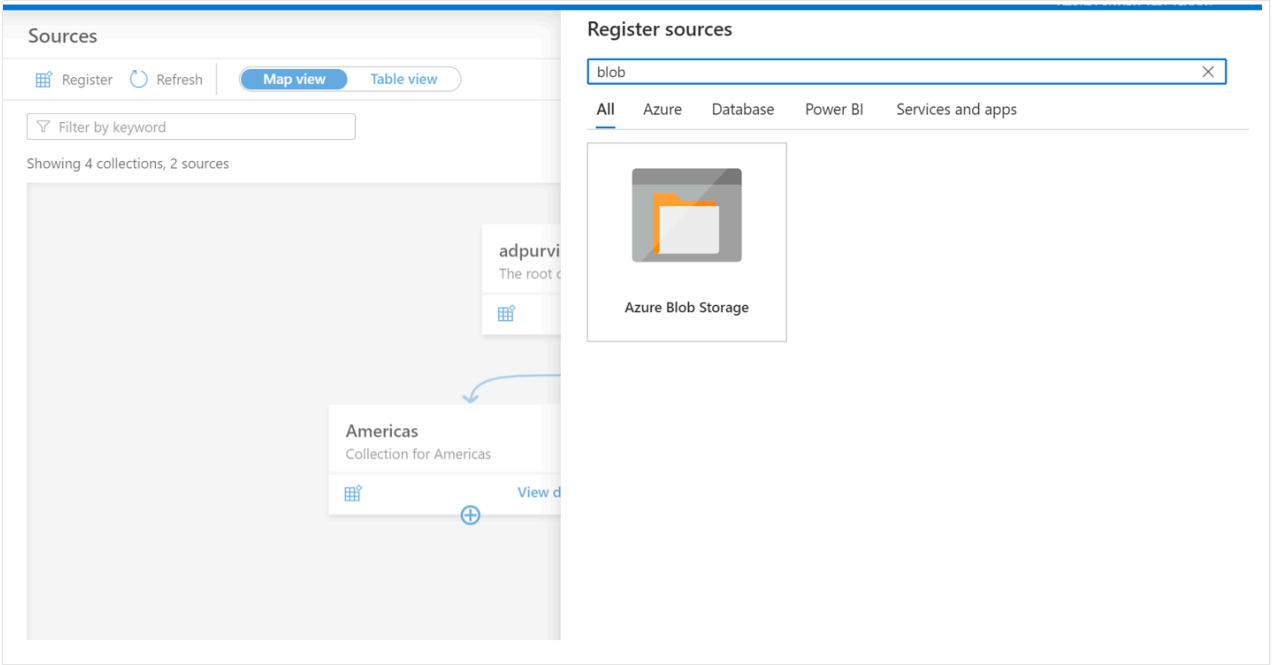
3. Create the Collection hierarchy using the **Collections** menu and assign permissions to individual subcollections, as required



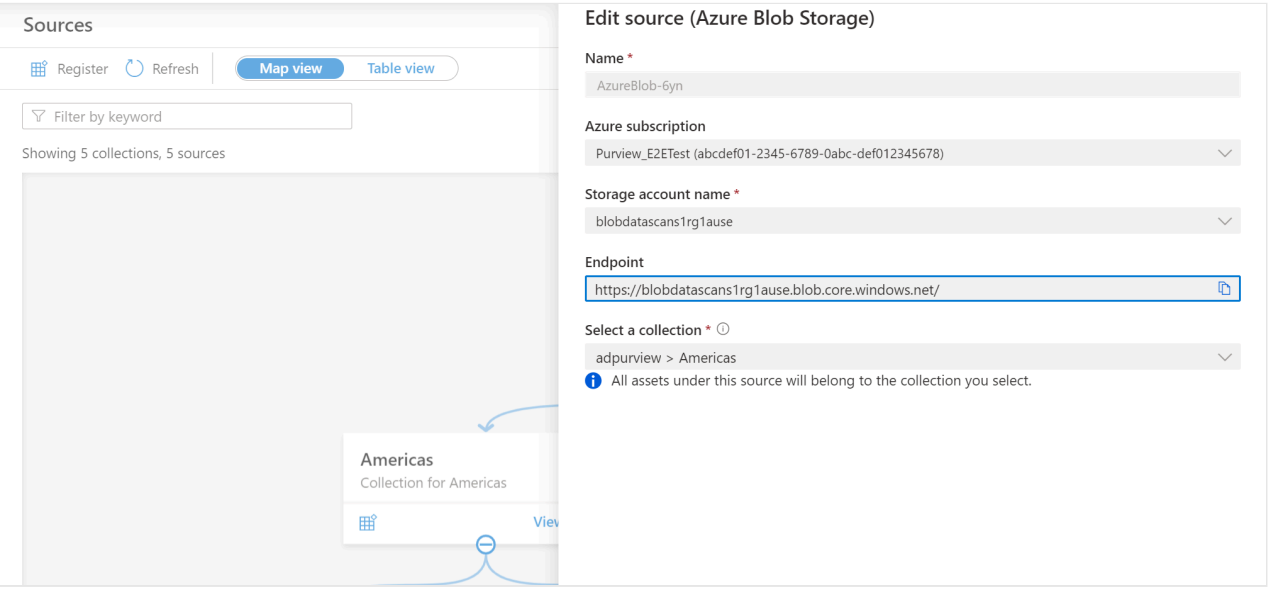
4. Navigate to the appropriate collection under the **Sources** menu and select the **Register** icon to register a new Azure Blob data source



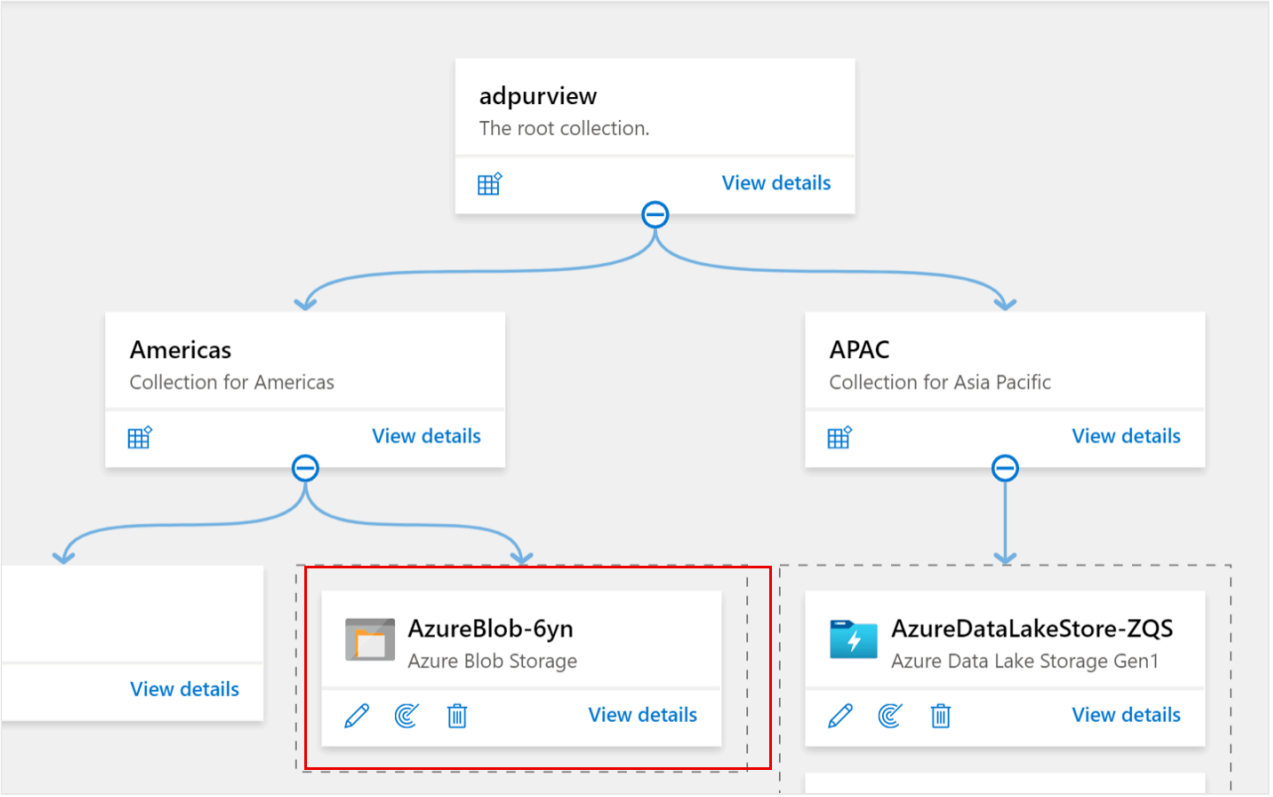
5. Select the **Azure Blob Storage** data source and select **Continue**



6. Provide a suitable **Name** for the data source, select the relevant **Azure subscription**, existing **Azure Blob Storage account name** and the **collection** and select **Apply**. Leave the **Data Policy Enforcement** toggle on the **disabled** position until you have a chance to carefully go over this document.



7. The Azure Blob storage account will be shown under the selected Collection



Scan

For file types such as csv, tsv, psv, ssv, the schema is extracted when the following logics are in place:

- First row values are non-empty
- First row values are unique
- First row values are not a date or a number

Authentication for a scan

Your Azure network may allow for communications between your Azure resources, but if you've set up firewalls, private endpoints, or virtual networks within Azure, you'll need to follow one of these configurations below.

Expand table

Networking constraints	Integration runtime type	Available credential types
No private endpoints or firewalls	Azure IR	Managed identity (Recommended), service principal, or account key
Firewall enabled but no private endpoints	Azure IR	Managed identity

Networking constraints	Integration runtime type	Available credential types
Private endpoints enabled	*Self-Hosted IR	Service principal, account key

*To use a self-hosted integration runtime, you'll first need to choose the right one for your scenario, create one, and confirm your network settings for Microsoft Purview.

Using a system or user assigned managed identity for scanning

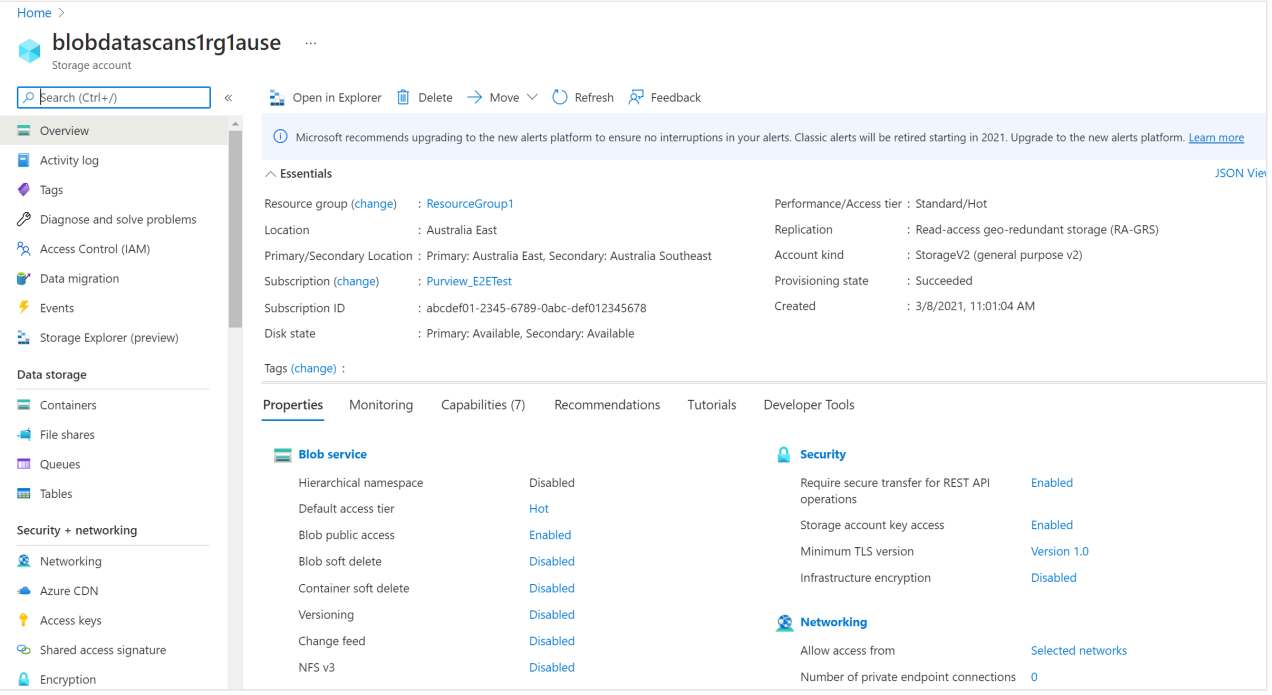
There are two types of managed identity you can use:

- **System-assigned managed identity (Recommended)** - As soon as the Microsoft Purview Account is created, a system-assigned managed identity (SAMI) is created automatically in Microsoft Entra tenant. Depending on the type of resource, specific RBAC role assignments are required for the Microsoft Purview system-assigned managed identity (SAMI) to perform the scans.
- **User-assigned managed identity** (preview) - Similar to a system managed identity, a user-assigned managed identity (UAMI) is a credential resource that can be used to allow Microsoft Purview to authenticate against Microsoft Entra ID. For more information, you can see our User-assigned managed identity guide. It's important to give your Microsoft Purview account the permission to scan the Azure Blob data source. You can add access for the SAMI or UAMI at the Subscription, Resource Group, or Resource level, depending on what level scan permission is needed.

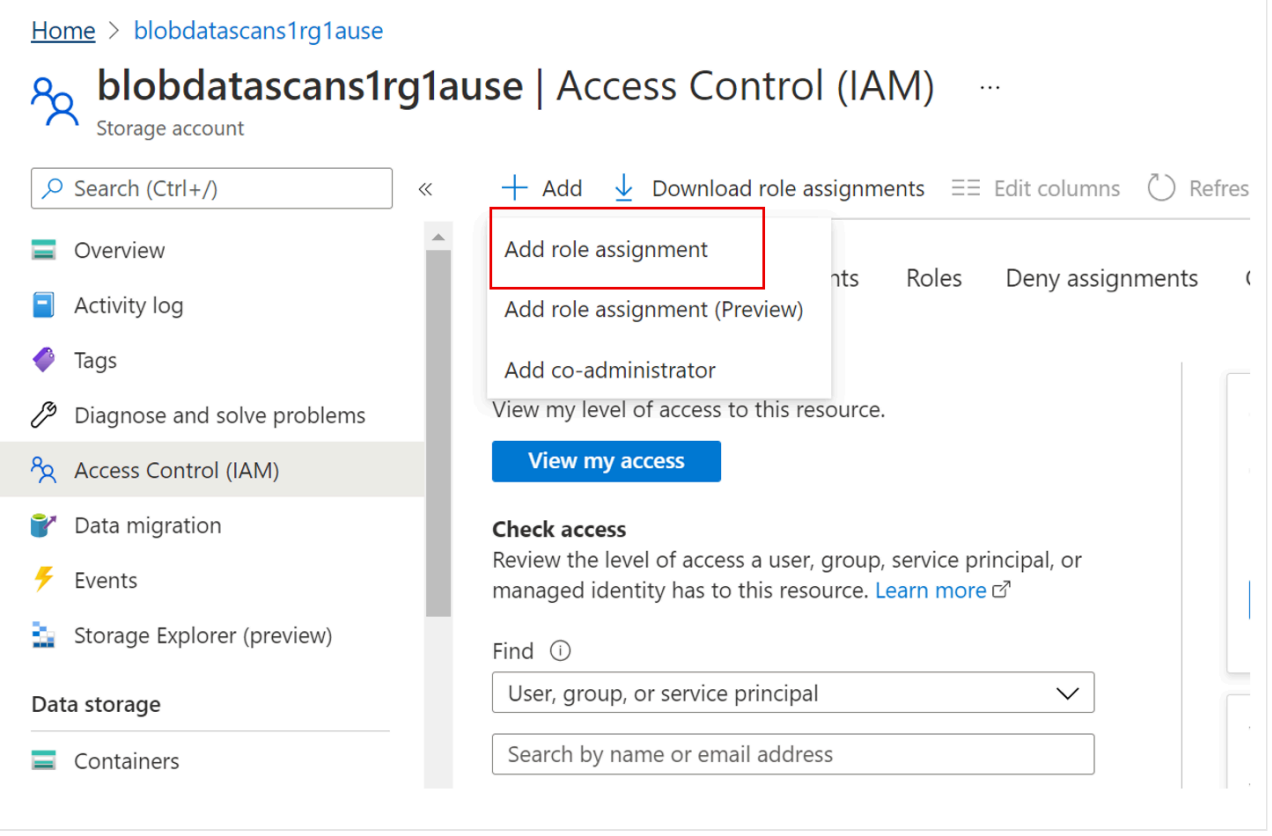
> **Note**
>
> If you have firewall enabled for the storage account, you must use **managed identity** authentication method when setting up a scan.

> **Note**
>
> You need to be an owner of the subscription to be able to add a managed identity on an Azure resource.

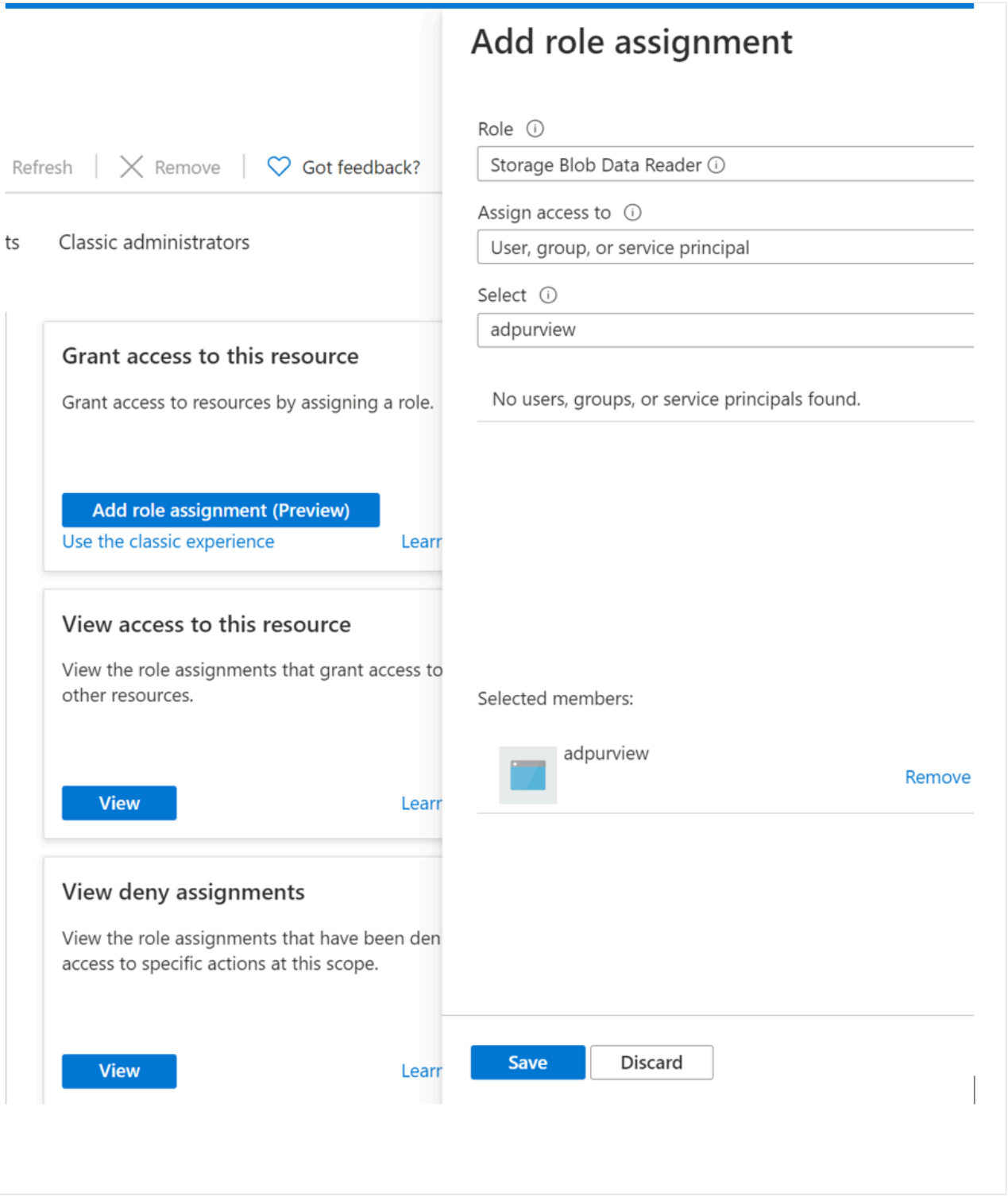
1. From the Azure portal, find either the subscription, resource group, or resource (for example, an Azure Blob storage account) that you would like to allow the catalog to scan.



2. Select **Access Control (IAM)** in the left navigation and then select **+ Add** --> **Add role assignment**



3. Set the **Role** to **Storage Blob Data Reader** and enter your *Microsoft Purview account name* or *user-assigned managed identity* under **Select** input box. Then, select **Save** to give this role assignment to your Microsoft Purview account.

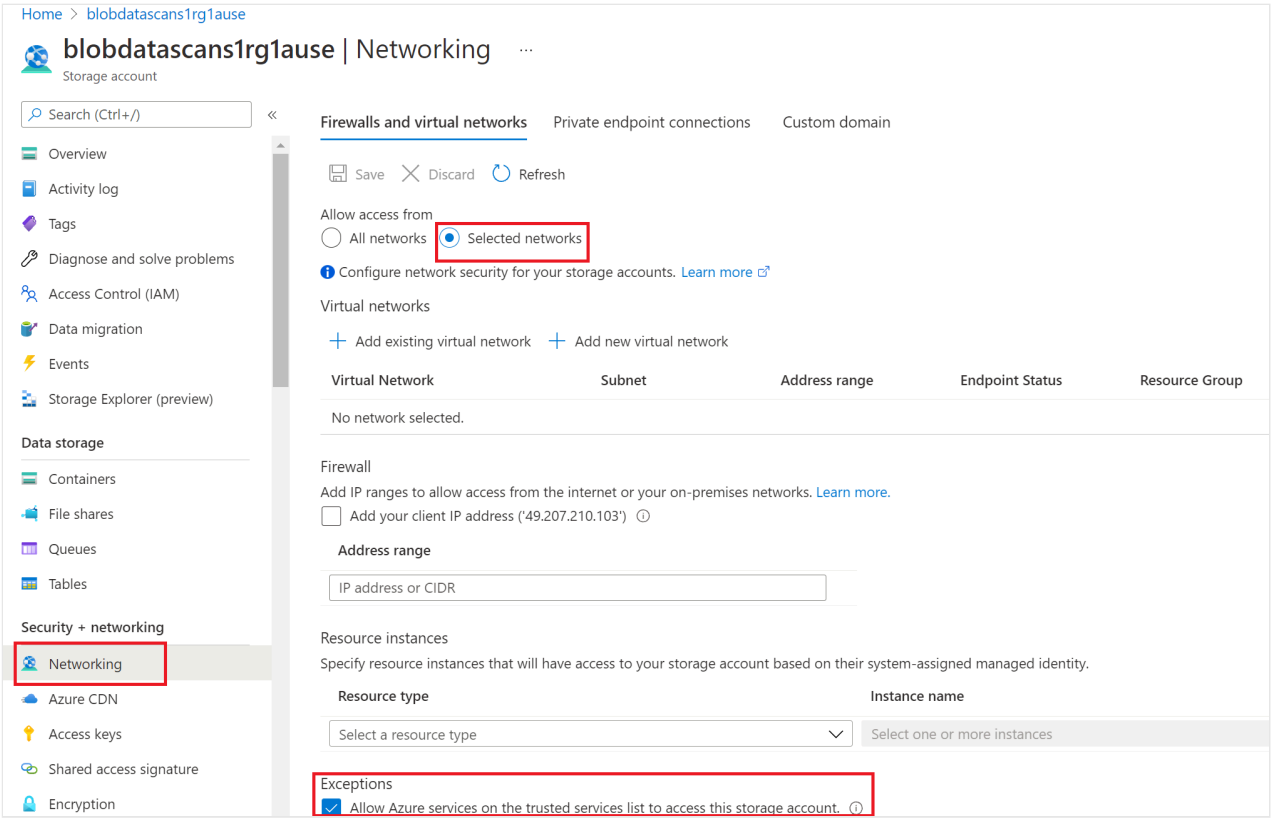


4. Go into your Azure Blob storage account in Azure portal

5. Navigate to **Security + networking** > **Networking**

6. Choose **Selected Networks** under **Allow access from**

7. In the **Exceptions** section, select **Allow trusted Microsoft services to access this storage account** and hit **Save**

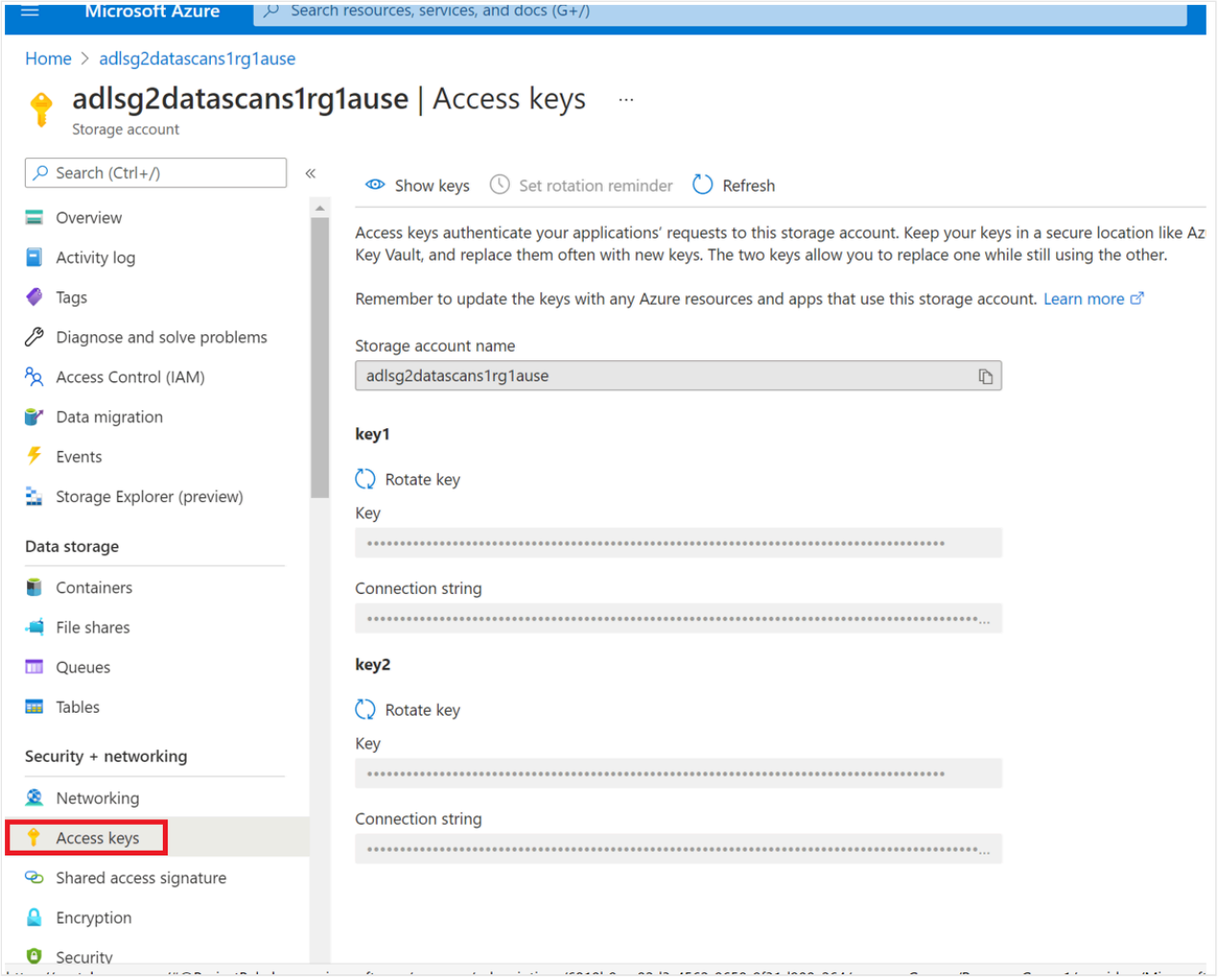


> **Note**
>
> For more details, please see steps in Authorize access to blobs and queues using Microsoft Entra ID

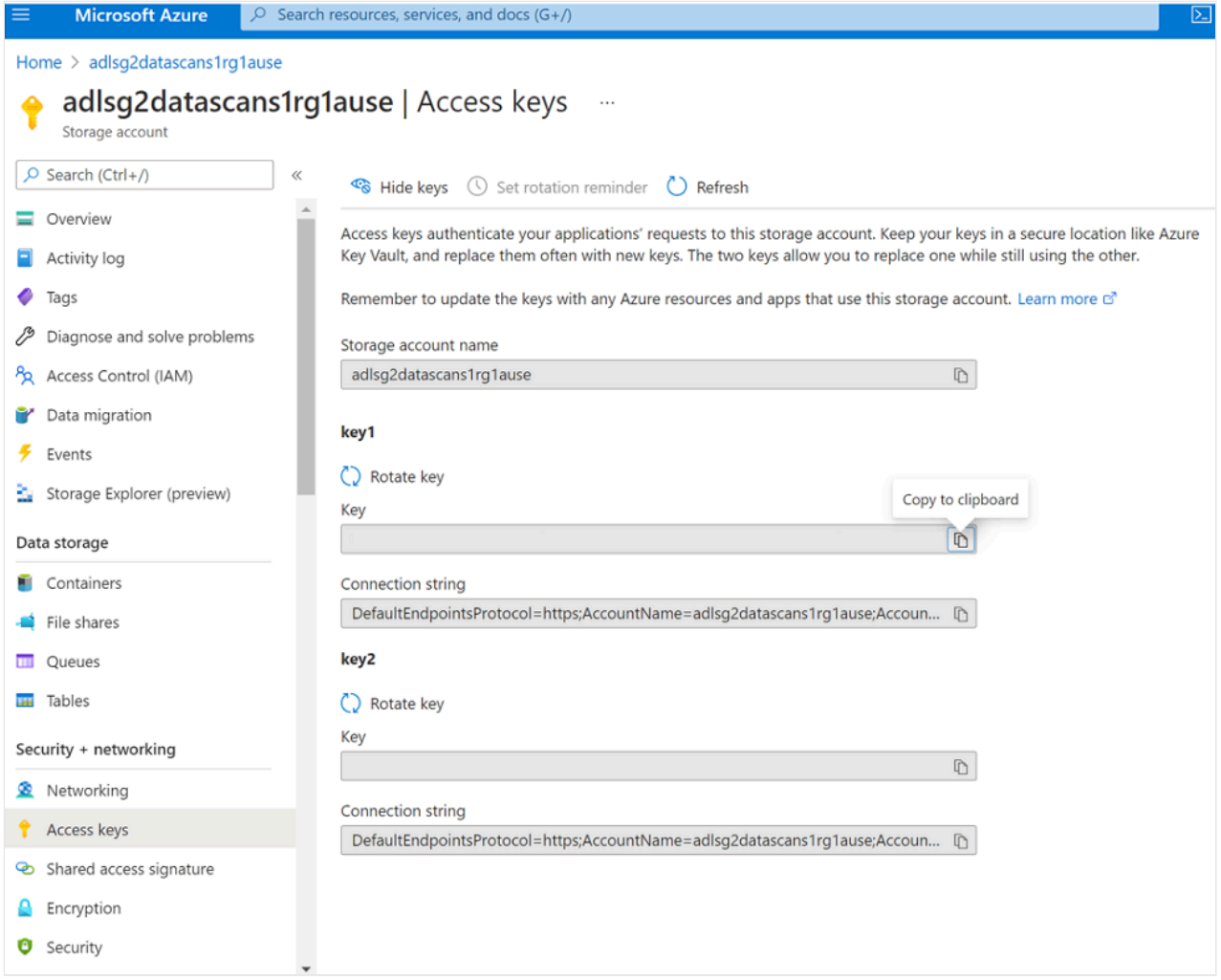
Using Account Key for scanning

When authentication method selected is **Account Key**, you need to get your access key and store in the key vault:

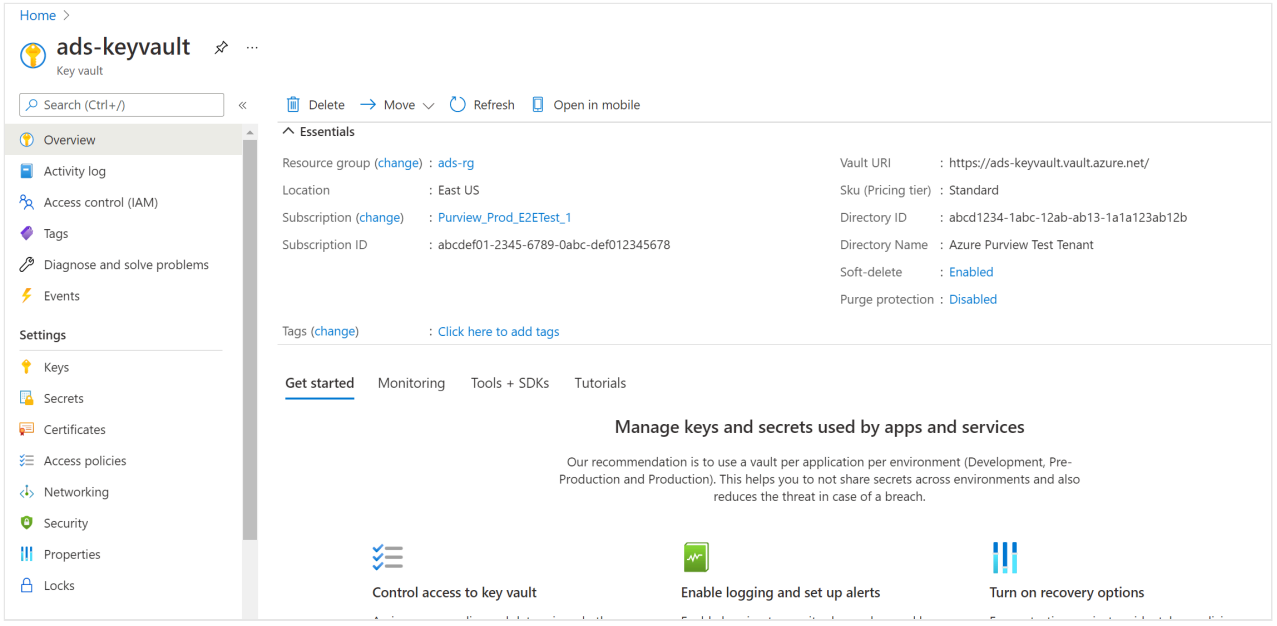
1. Navigate to your Azure Blob storage account
2. Select **Security + networking** > **Access keys**



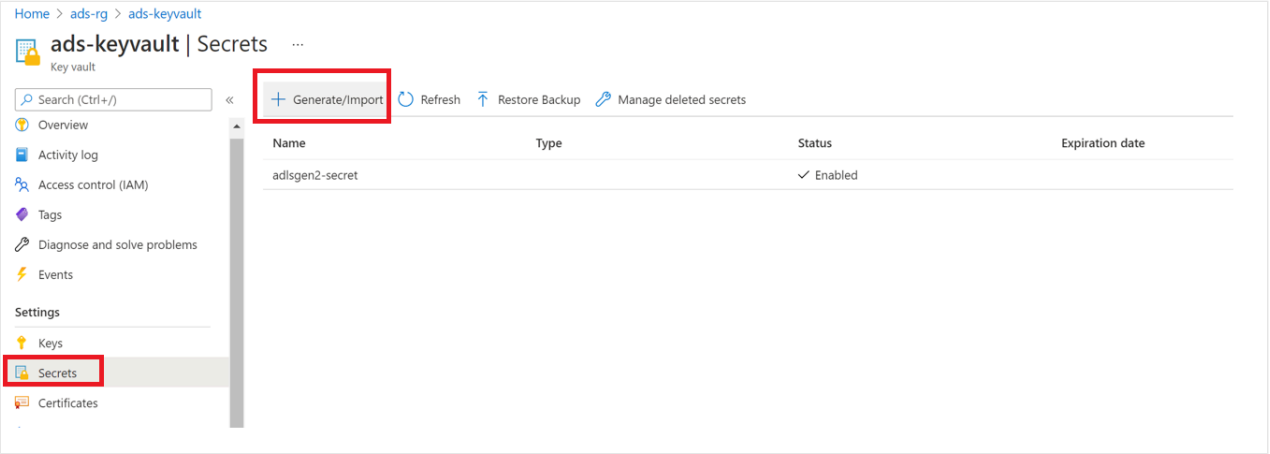
3. Copy your *key* and save it separately for the next steps



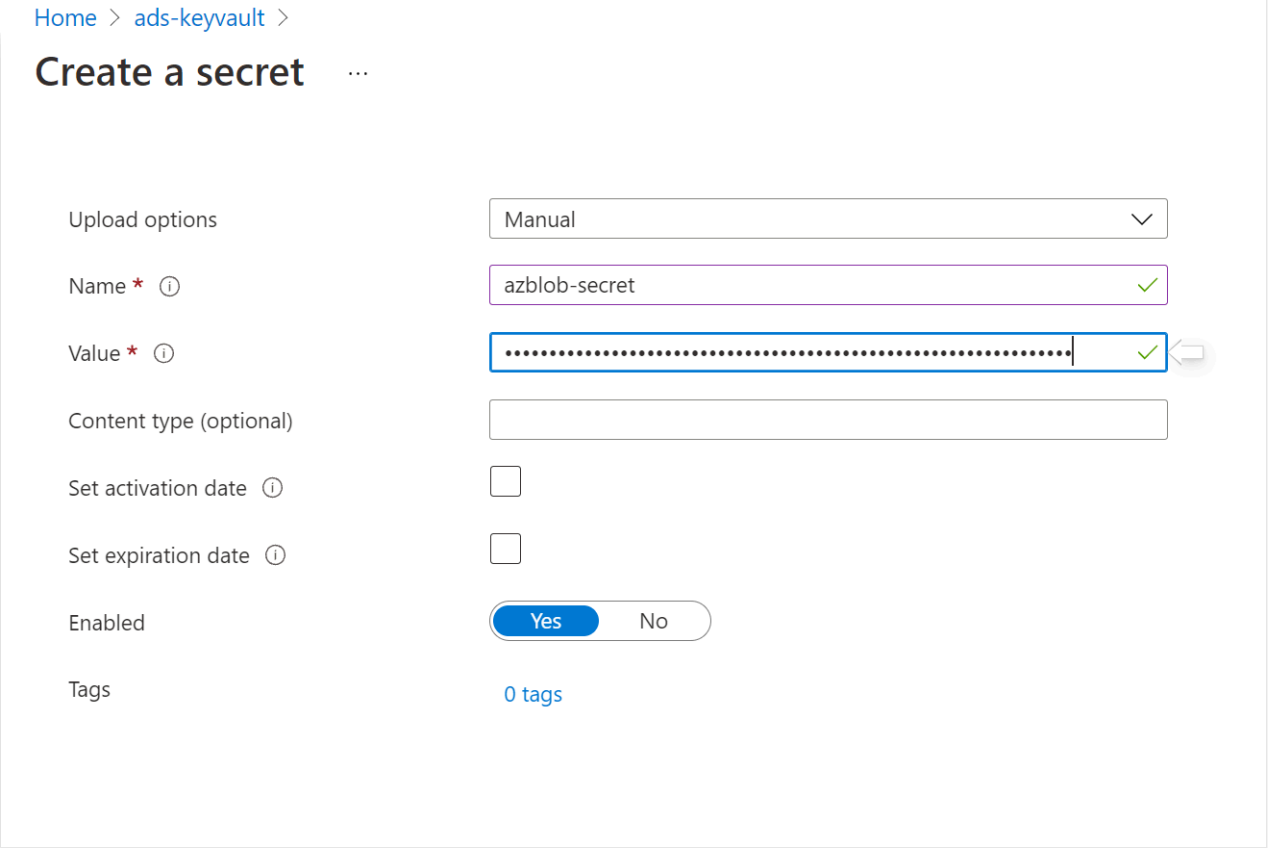
4. Navigate to your key vault



5. Select **Settings** > **Secrets** and select **+ Generate/Import**



6. Enter the **Name** and **Value** as the *key* from your storage account



7. Select **Create** to complete

8. If your key vault isn't connected to Microsoft Purview yet, you'll need to create a new key vault connection

9. Finally, create a new credential using the key to set up your scan

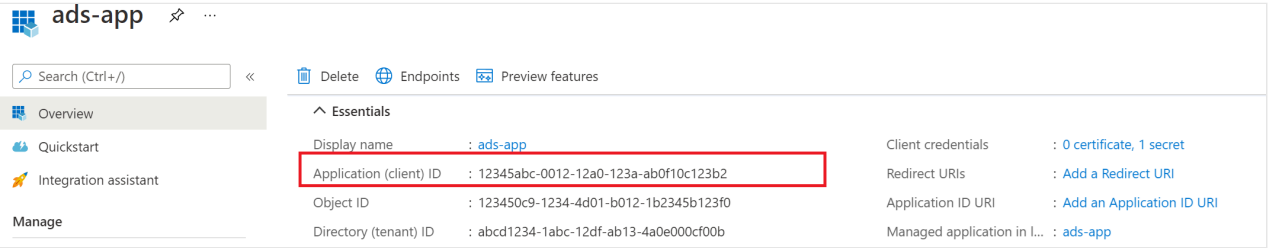
Using Service Principal for scanning

Creating a new service principal

If you need to Create a new service principal, it's required to register an application in your Microsoft Entra tenant and provide access to Service Principal in your data sources. Your Microsoft Entra Application Administrator can perform this operation.

Getting the Service Principal's Application ID

1. Copy the **Application (client) ID** present in the **Overview** of the *Service Principal* already created

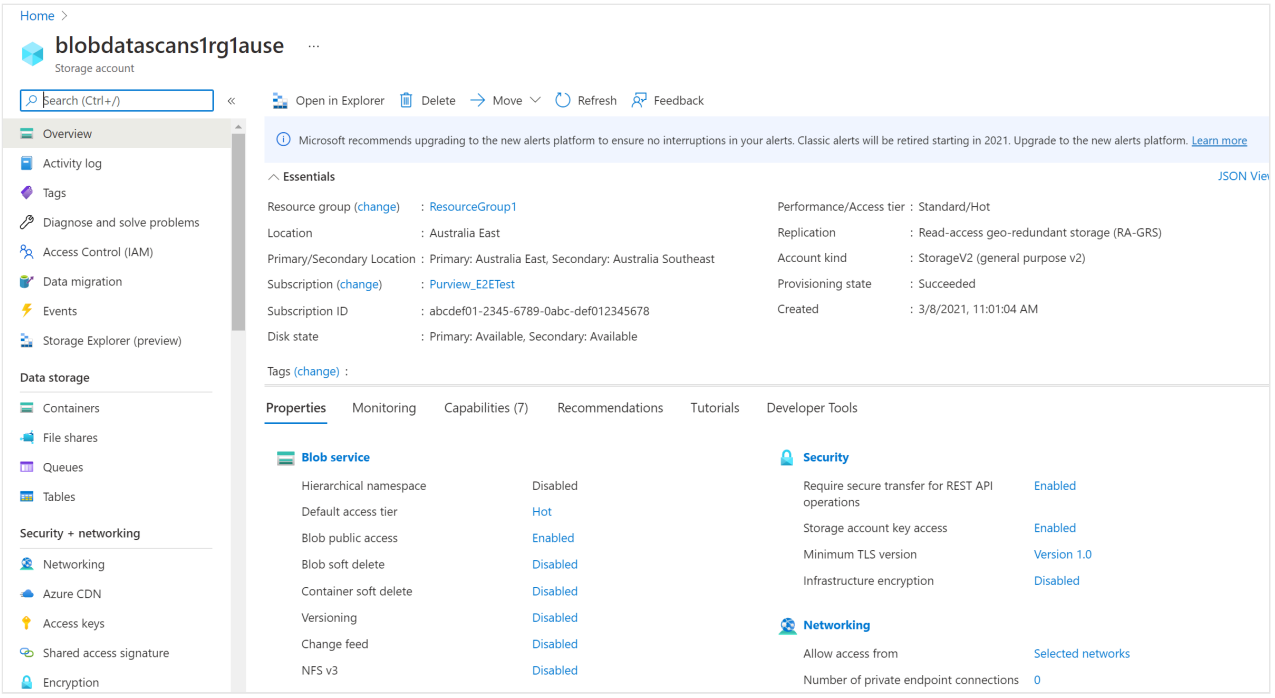


Granting the Service Principal access to your Azure Blob account

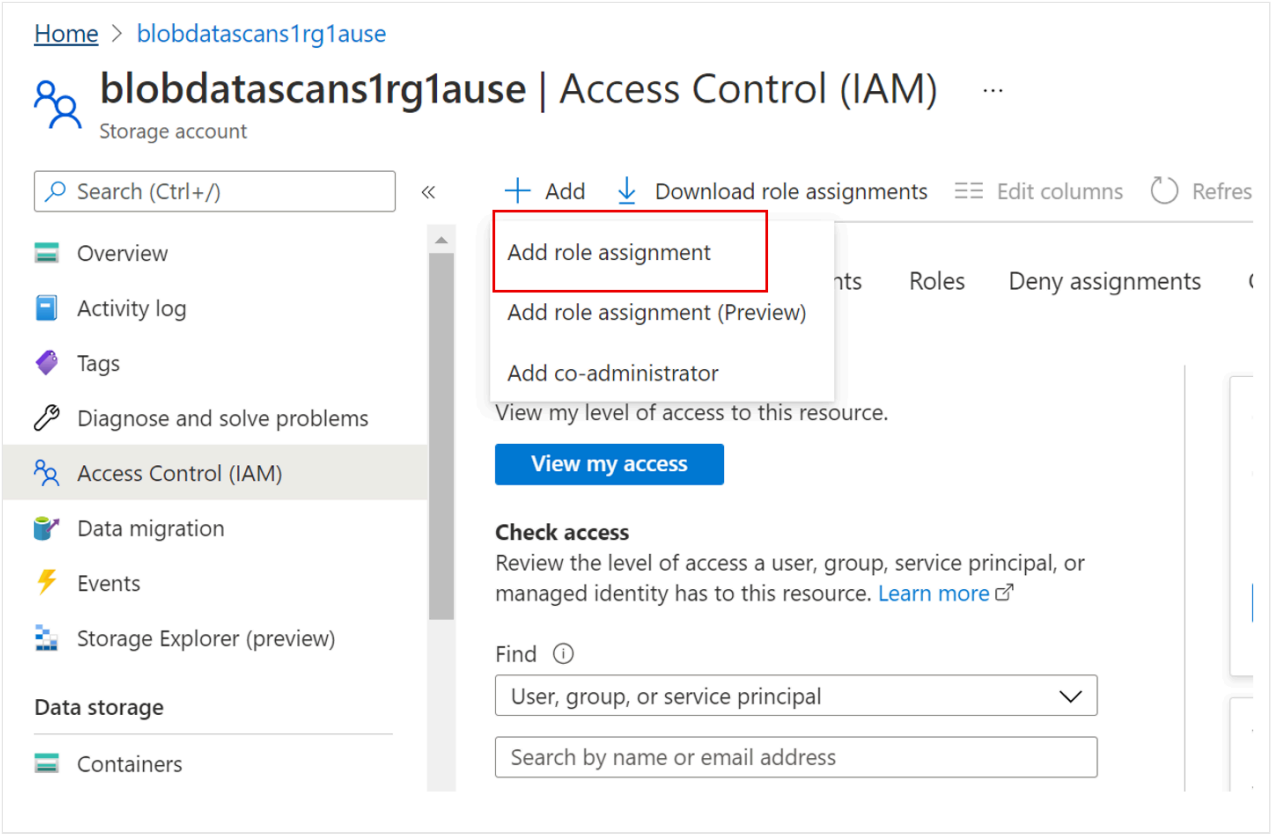
It's important to give your service principal the permission to scan the Azure Blob data source. You can add access for the service principal at the Subscription, Resource Group, or Resource level, depending on what level scan access is needed.

> **Note**
>
> You need to be an owner of the subscription to be able to add a service principal on an Azure resource.

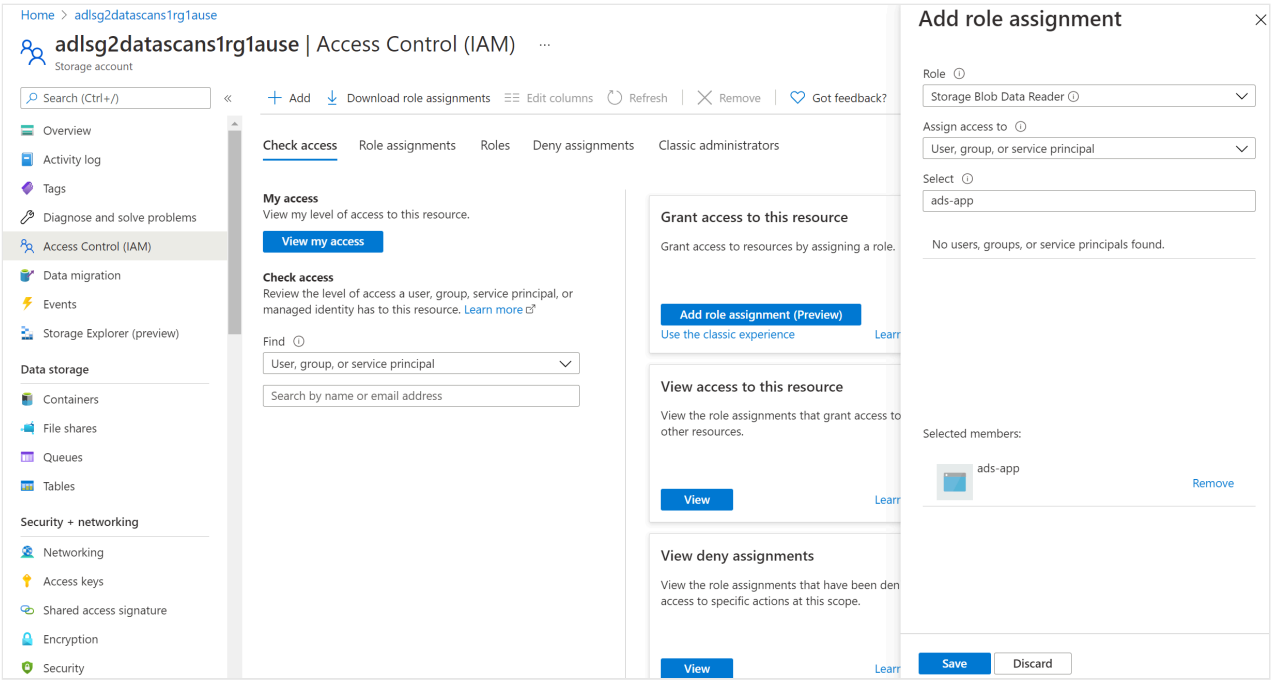
1. From the Azure portal, find either the subscription, resource group, or resource (for example, an Azure Blob Storage storage account) that you would like to allow the catalog to scan.



2. Select **Access Control (IAM)** in the left navigation and then select **+ Add** --> **Add role assignment**

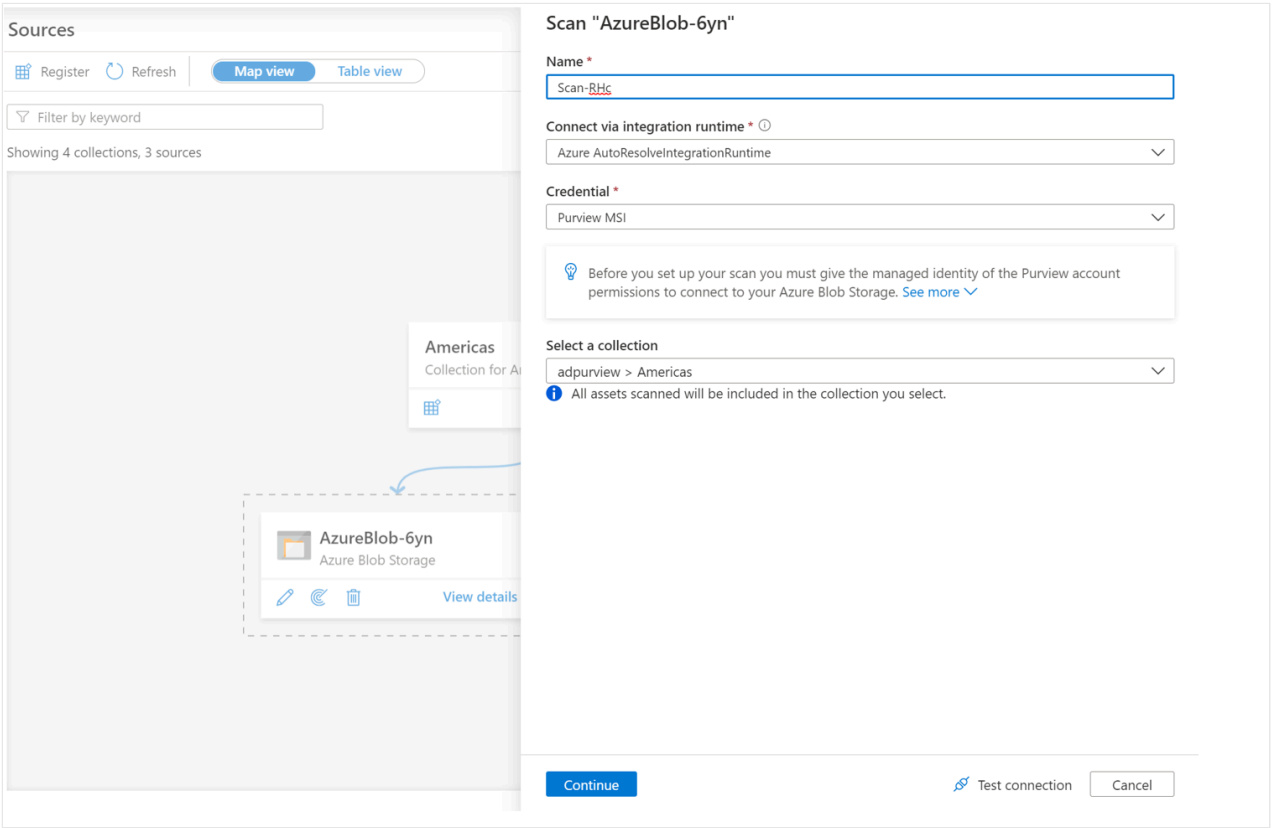


3. Set the **Role** to **Storage Blob Data Reader** and enter your *service principal* under **Select** input box. Then, select **Save** to give this role assignment to your Microsoft Purview account.



Creating the scan

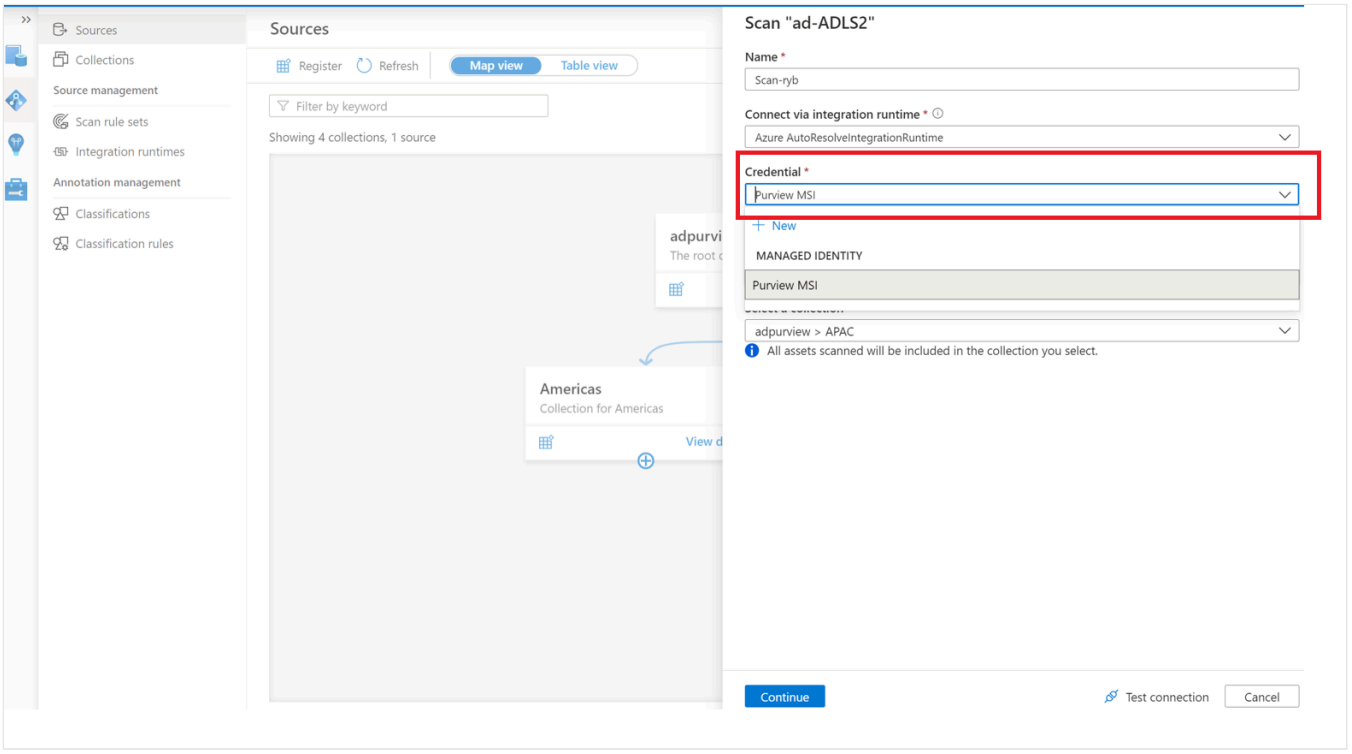
1. Open your **Microsoft Purview account** and select the **Open Microsoft Purview governance portal**
2. Navigate to the **Data map** --> **Sources** to view the collection hierarchy
3. Select the **New Scan** icon under the **Azure Blob data source** registered earlier



4. Choose either the Azure integration runtime if your source is publicly accessible, a managed virtual network integration runtime if using a managed virtual network, or a self-hosted integration runtime if your source is in a private virtual network. For more information about which integration runtime to use, see the choose the right integration runtime configuration article.

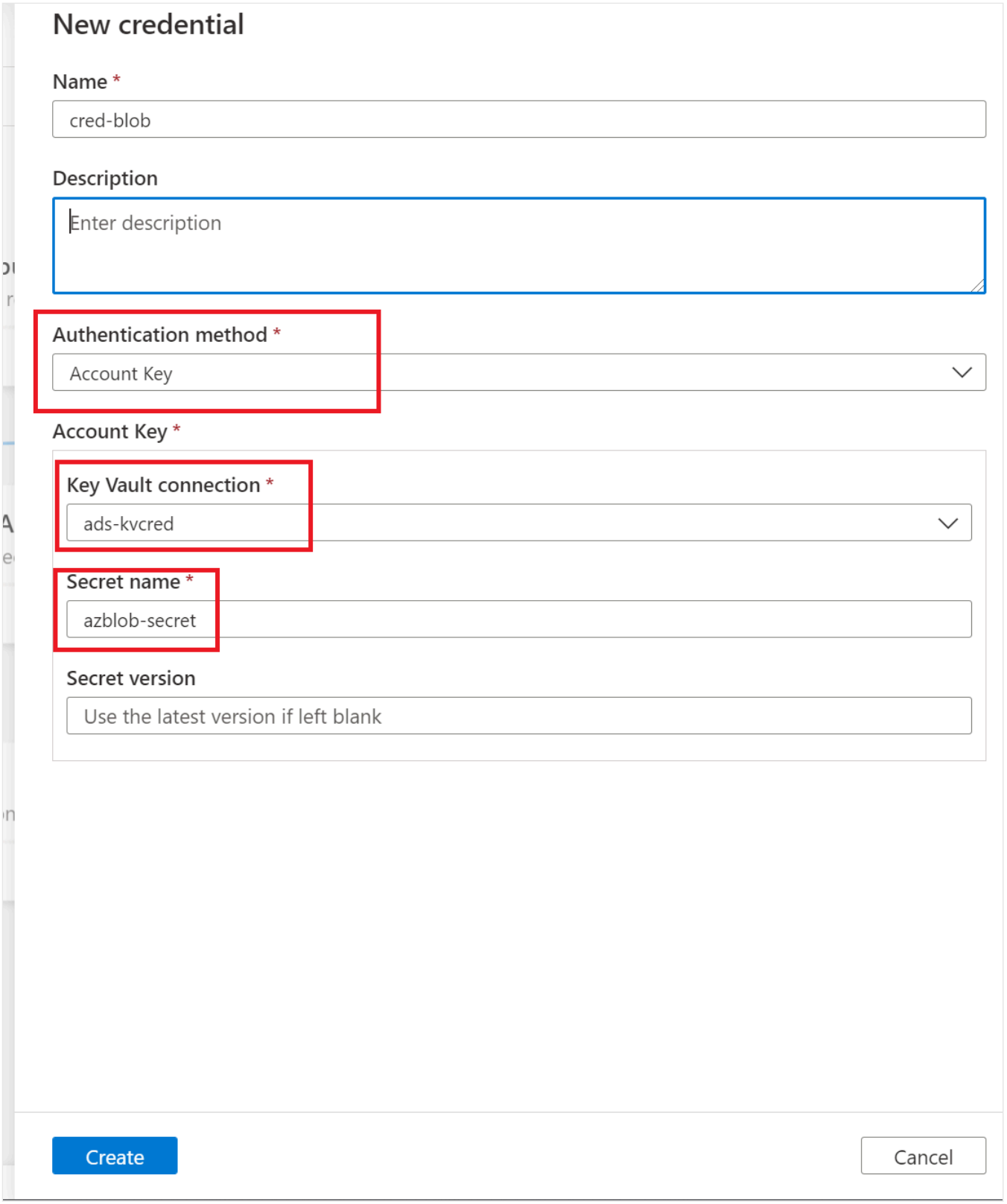
If using a system or user assigned managed identity

Provide a **Name** for the scan, select the Microsoft Purview accounts SAMI or UAMI under **Credential**, choose the appropriate collection for the scan, and select **Test connection**. On a successful connection, select **Continue**



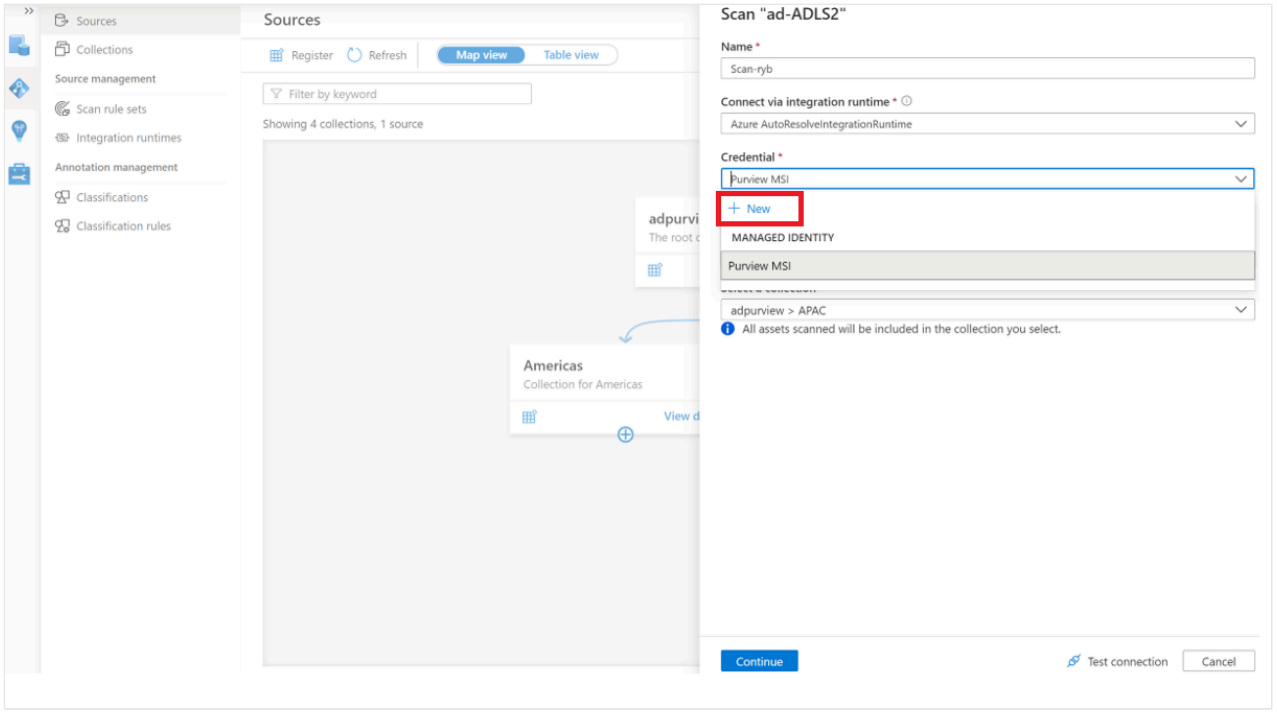
If using Account Key

Provide a **Name** for the scan, select the Azure IR or your Self-Hosted IR depending on your configuration, choose the appropriate collection for the scan, and select **Authentication method** as *Account Key* and select **Create**

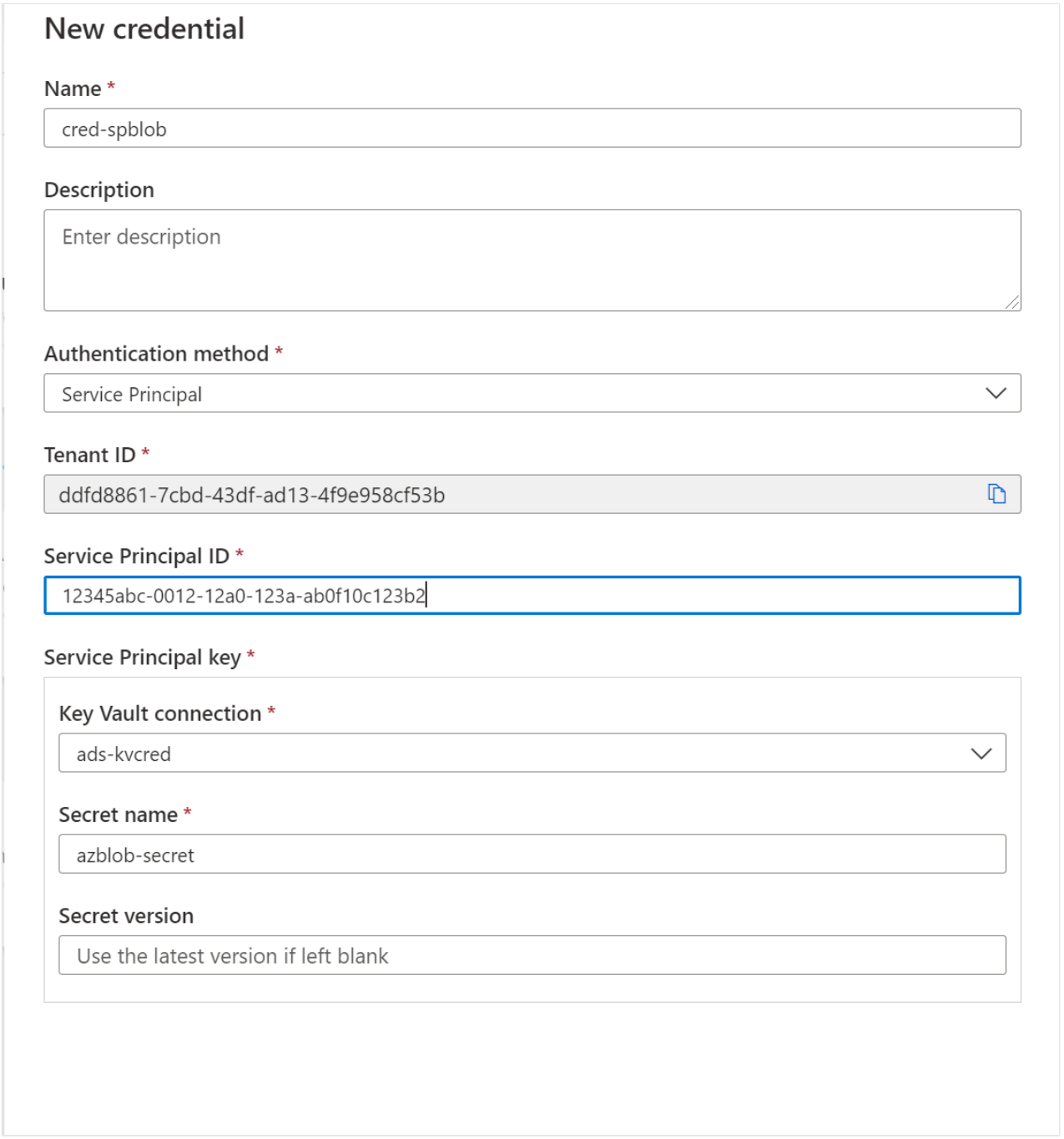


If using Service Principal

1. Provide a **Name** for the scan, select the Azure IR or your Self-Hosted IR depending on your configuration, choose the appropriate collection for the scan, and select the **+ New** under **Credential**



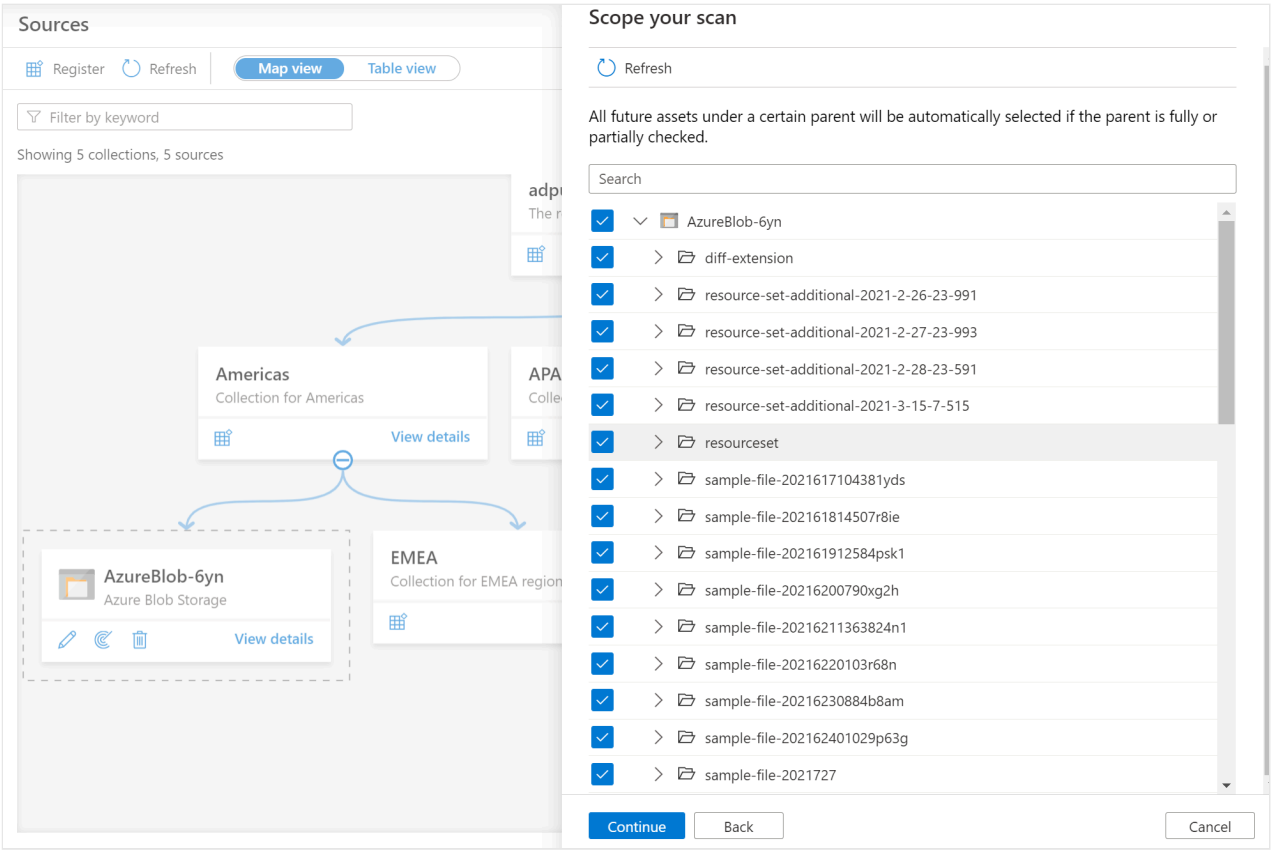
2. Select the appropriate **Key vault connection** and the **Secret name** that was used while creating the *Service Principal*. The **Service Principal ID** is the **Application (client) ID** copied earlier



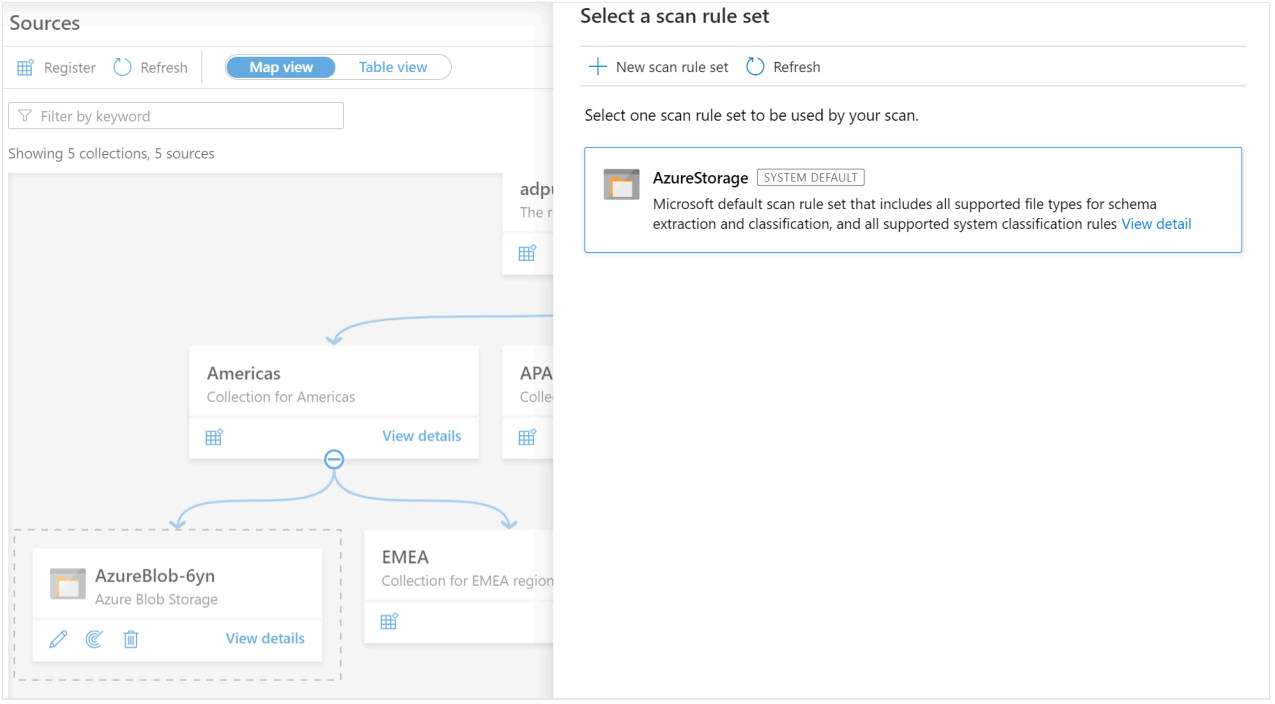
3. Select **Test connection**. On a successful connection, select **Continue**

Scoping and running the scan

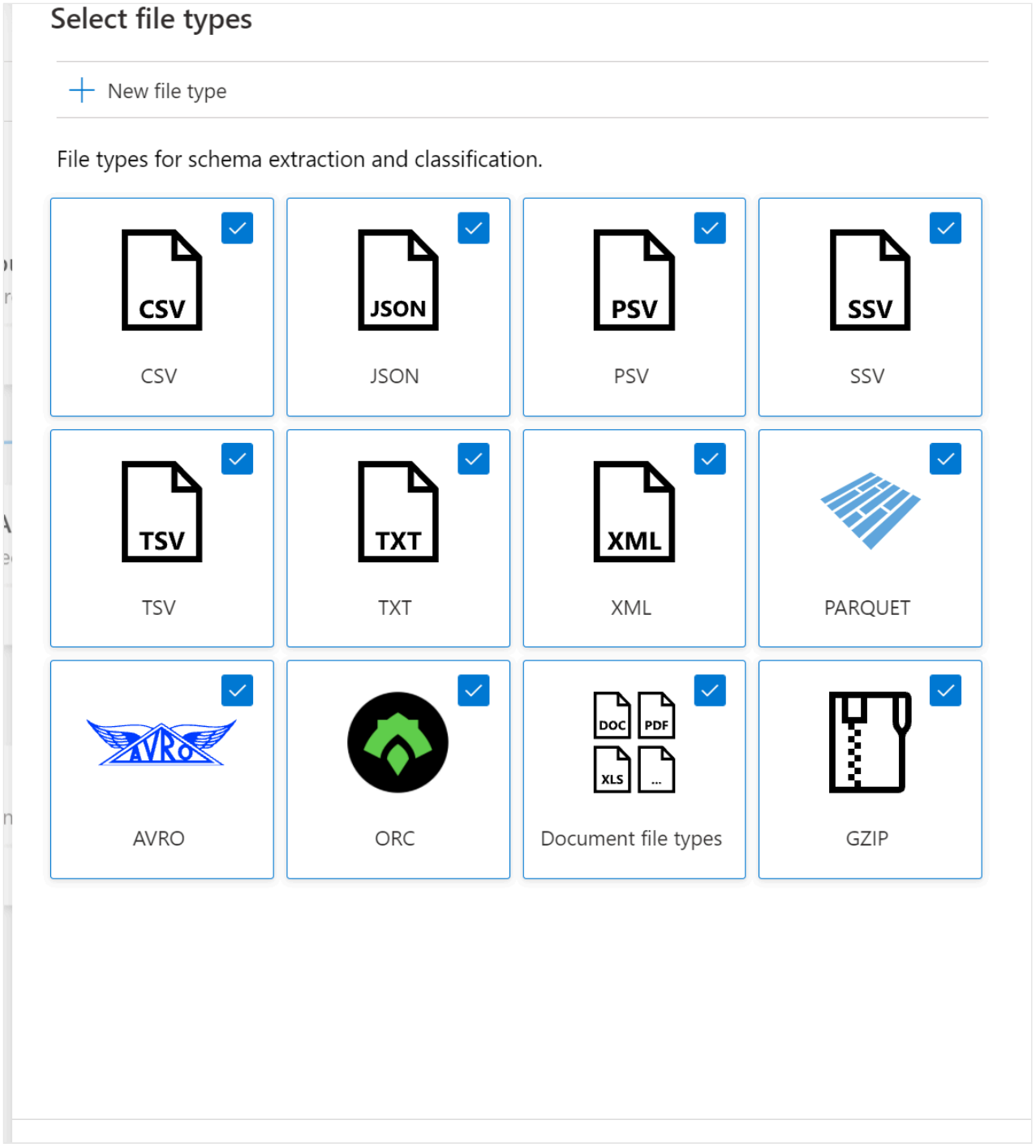
1. You can scope your scan to specific folders and subfolders by choosing the appropriate items in the list.



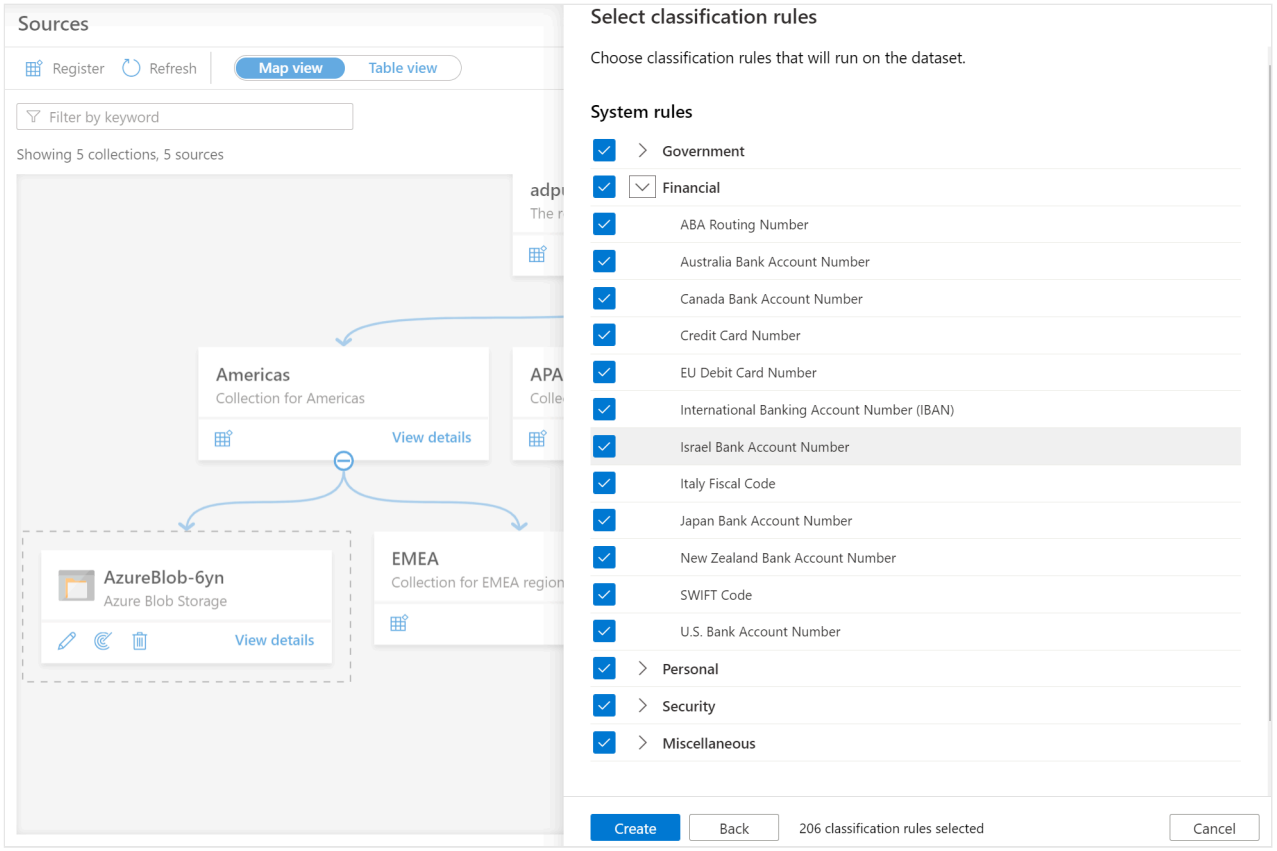
2. Then select a scan rule set. You can choose between the system default, existing custom rule sets, or create a new rule set inline.

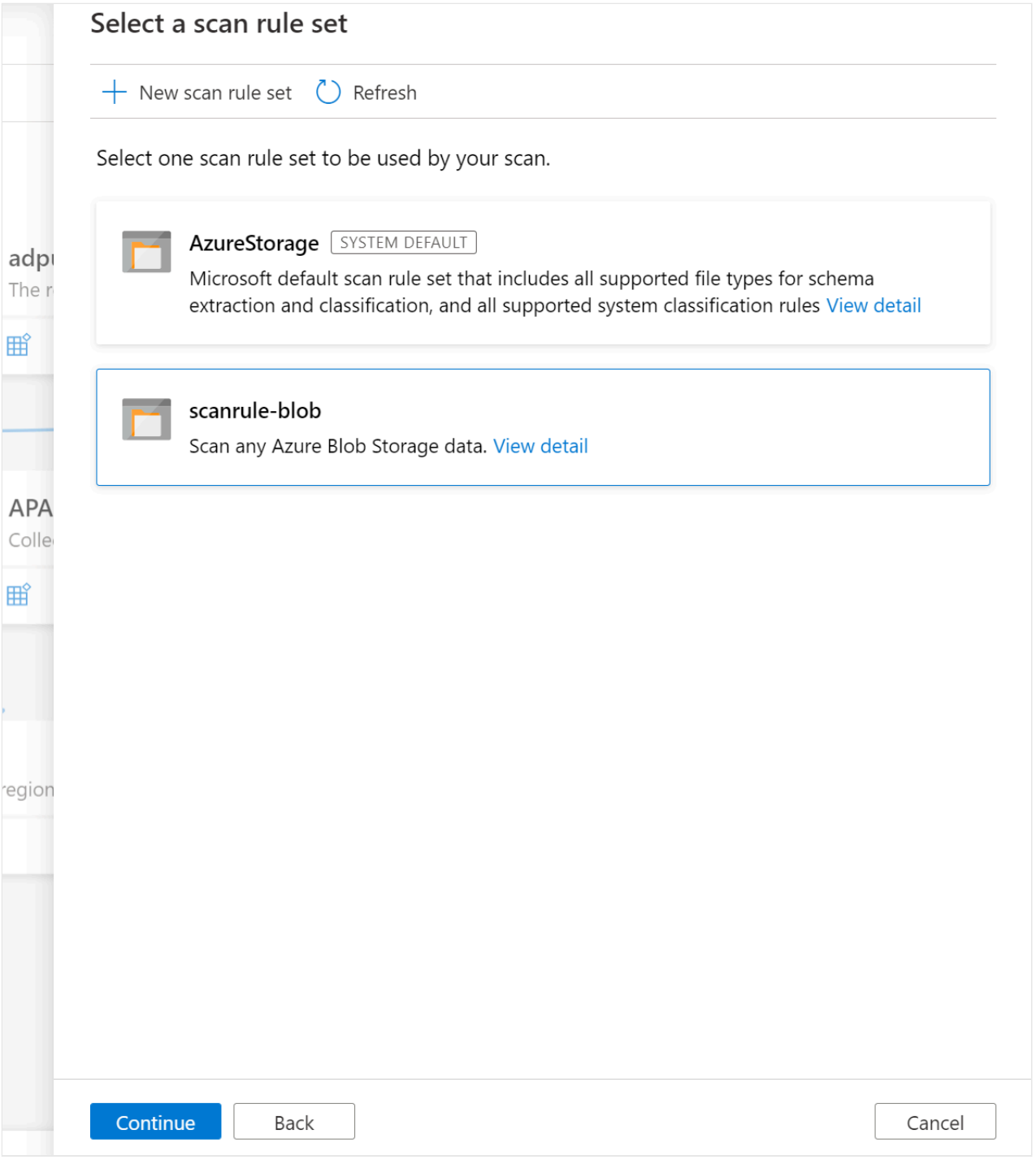


3. If creating a new *scan rule set*, select the **file types** to be included in the scan rule.

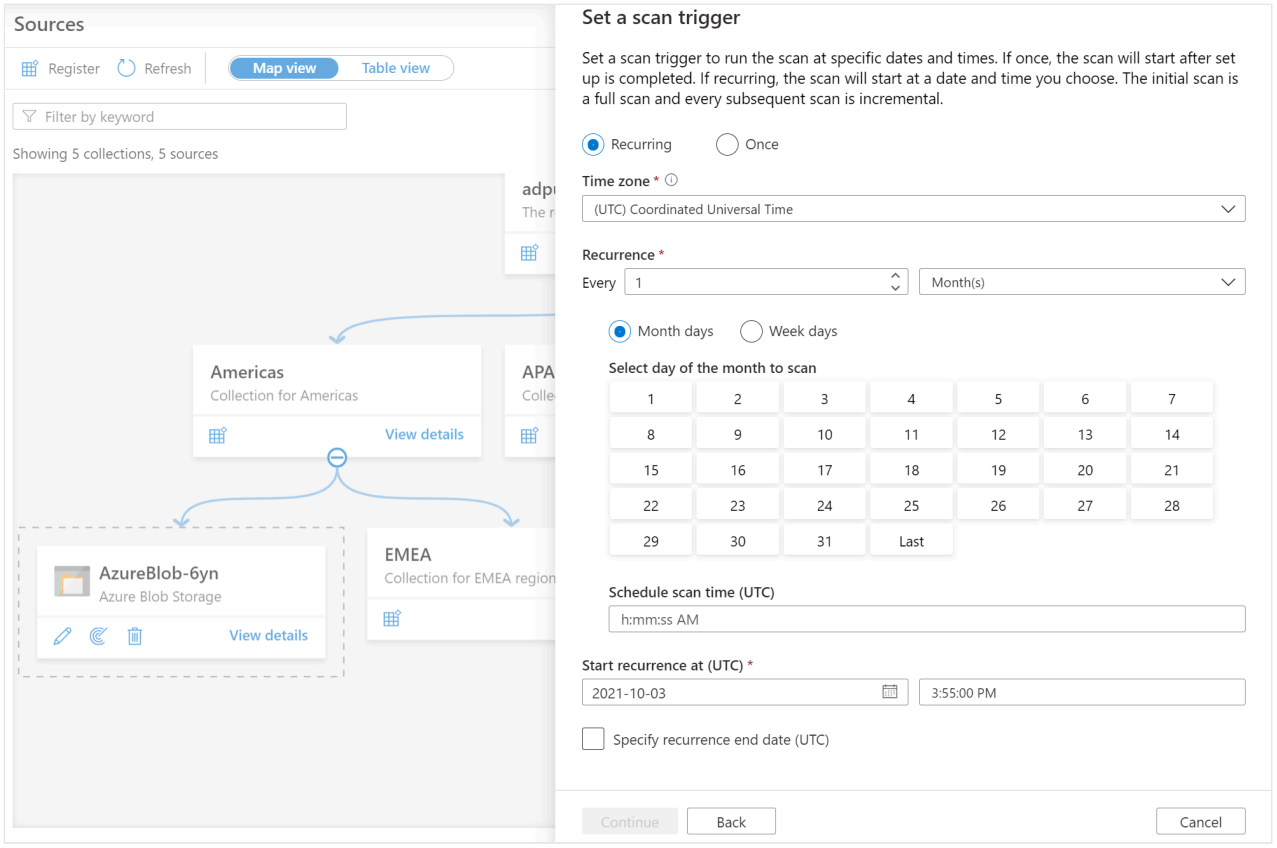


4. You can select the **classification rules** to be included in the scan rule

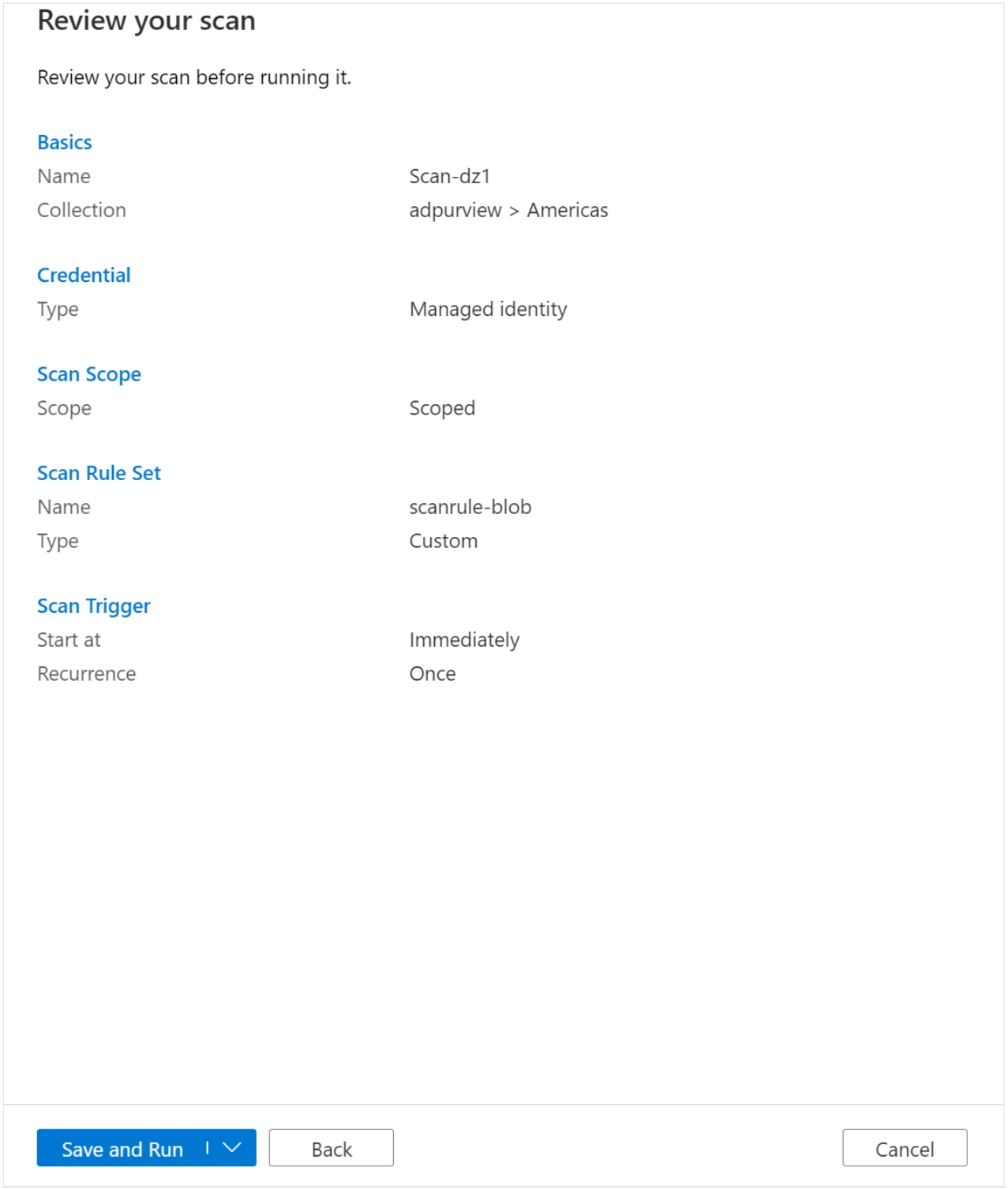

Sources
Register
Refresh
Map view
Table view
Filter by keyword
Showing 5 collections, 5 sources
Americas
Collection for Americas
View details
EMEA
Collection for EMEA region
AzureBlob-6yn
Azure Blob Storage
View details
Select classification rules
Choose classification rules that will run on the dataset.
System rules
Government
Financial
ABA Routing Number
Australia Bank Account Number
Canada Bank Account Number
Credit Card Number
EU Debit Card Number
International Banking Account Number (IBAN)
Israel Bank Account Number
Italy Fiscal Code
Japan Bank Account Number
New Zealand Bank Account Number
SWIFT Code
U.S. Bank Account Number
Personal
Security
Miscellaneous
Create
Back
206 classification rules selected
Cancel



5. Choose your scan trigger. You can set up a schedule or run the scan once.

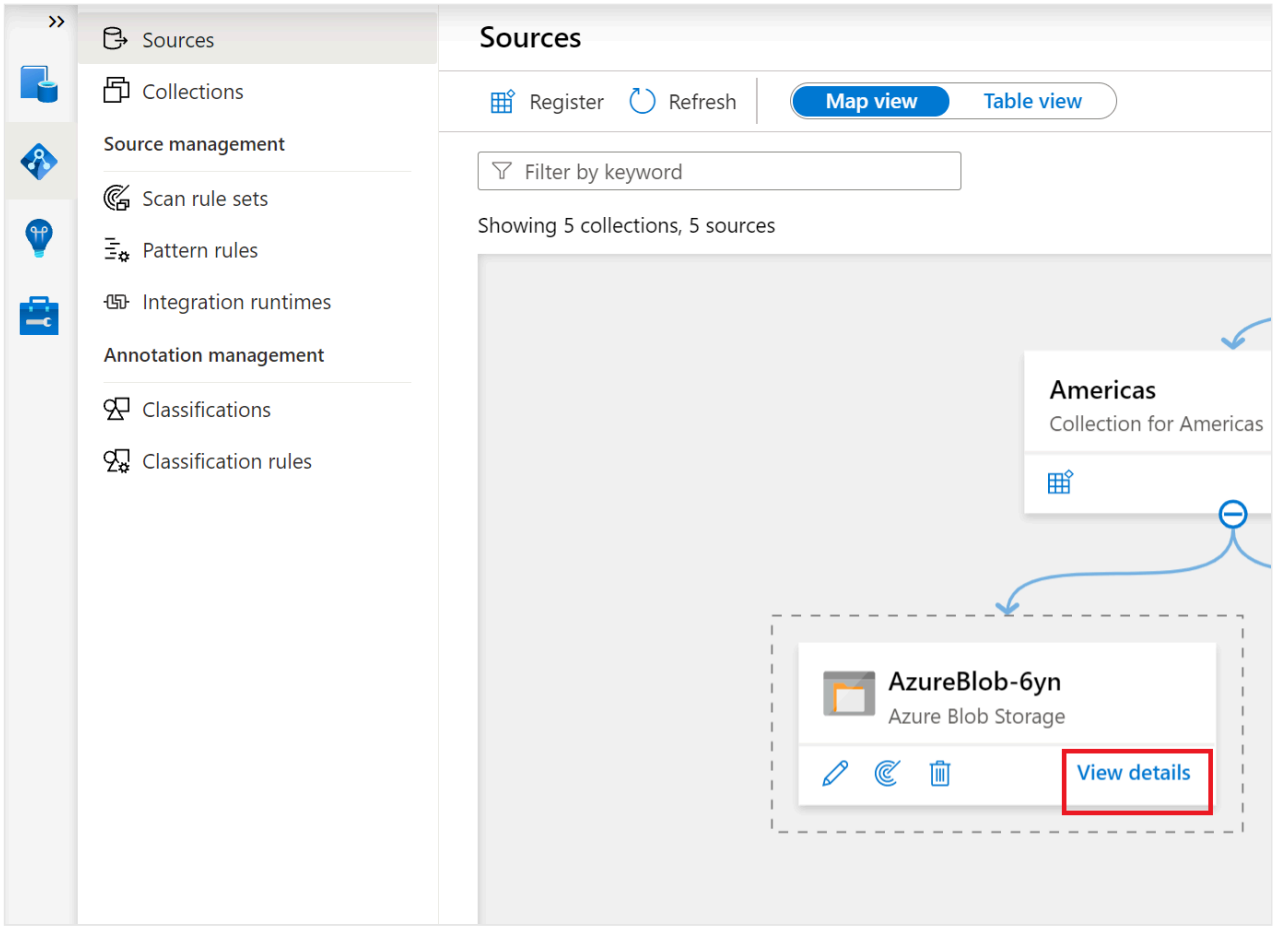


6. Review your scan and select **Save and run**.

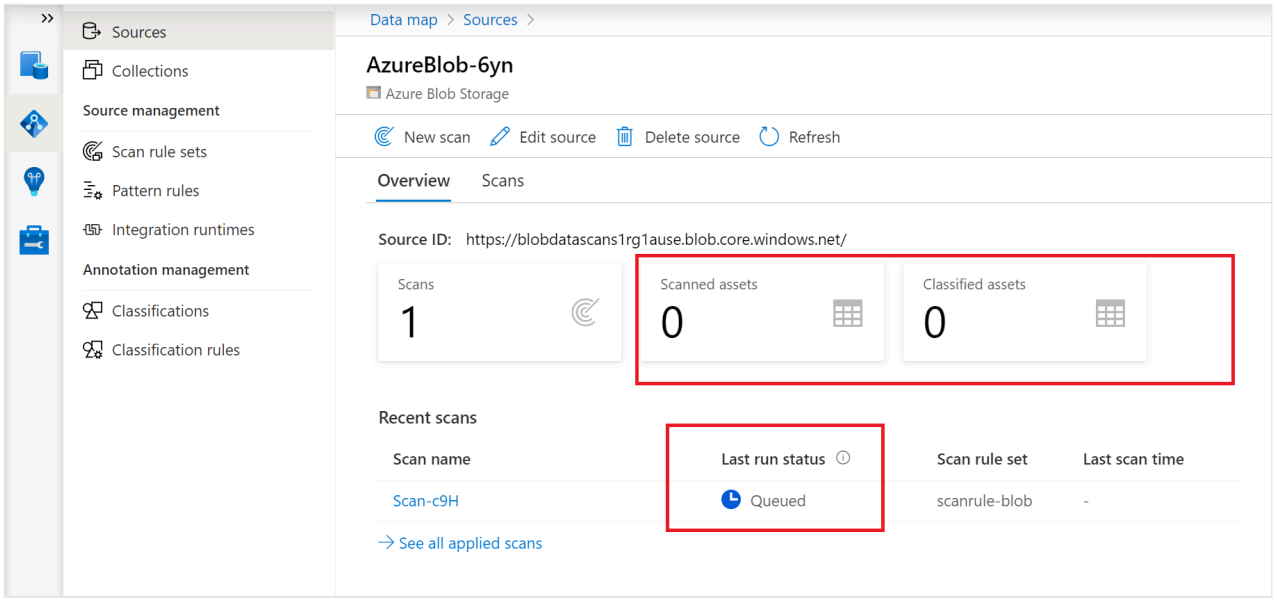


Viewing Scan

1. Navigate to the *data source* in the *Collection* and select **View Details** to check the status of the scan



2. The scan details indicate the progress of the scan in the **Last run status** and the number of assets *scanned* and *classified*



3. The **Last run status** will be updated to **In progress** and then **Completed** once the entire scan has run successfully

AzureBlob-6yn

Azure Blob Storage

New scan | Edit source | Delete source | Refresh

Overview | Scans

Source ID: https://blobdatascans1rg1ause.blob.core.windows.net/

Scans	Scanned assets	Classified assets
1	20	9

Recent scans

Scan name	Last run status	Scan rule set
Scan-c9H	In progress	scanrule-blob

See all applied scans

Data map > Sources >

AzureBlob-6yn

Azure Blob Storage

New scan | Edit source | Delete source | Refresh

Overview | Scans

Source ID: https://blobdatascans1rg1ause.blob.core.windows.net/

Scans	Scanned assets	Classified assets
1	20	9

Recent scans

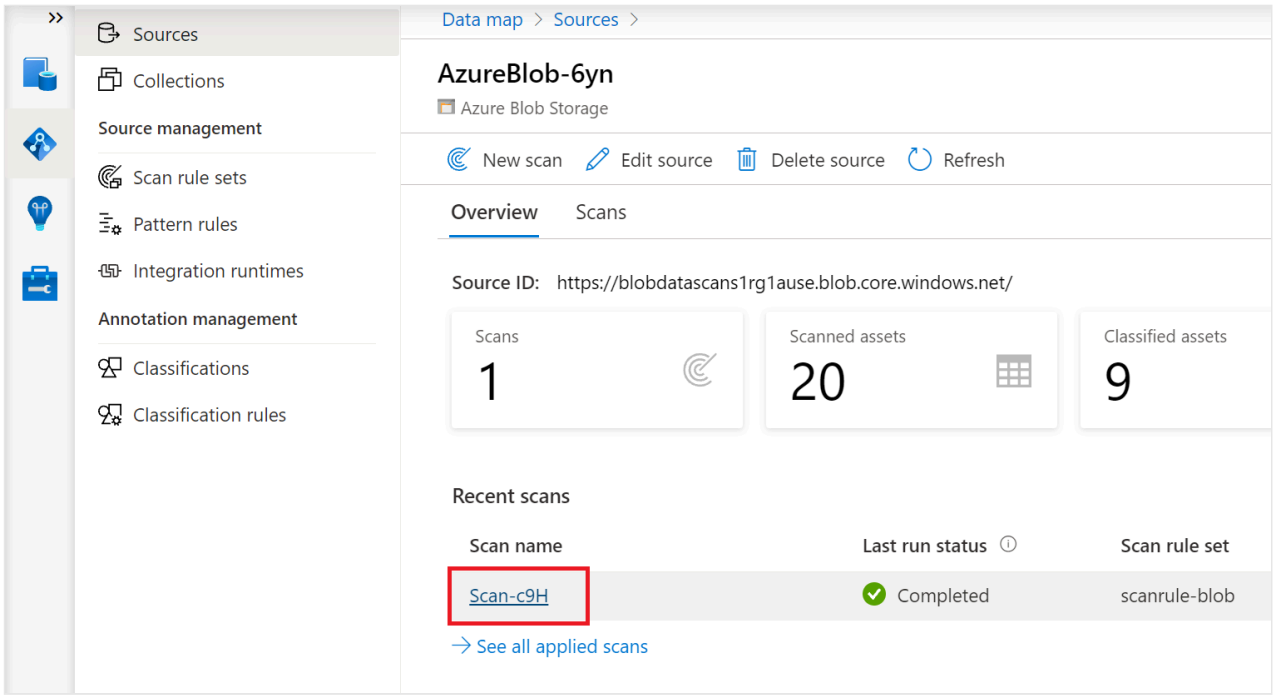
Scan name	Last run status	Scan rule set
Scan-c9H	Completed	scanrule-blob

See all applied scans

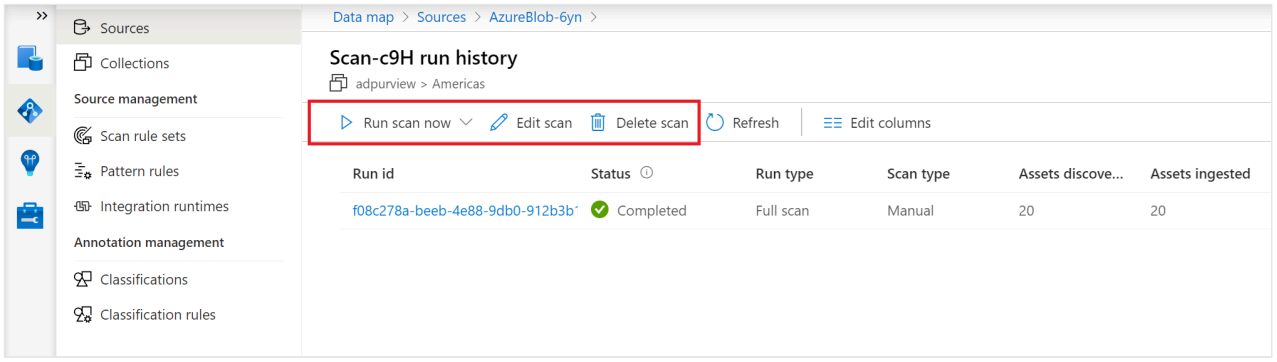
Managing Scan

Scans can be managed or run again on completion

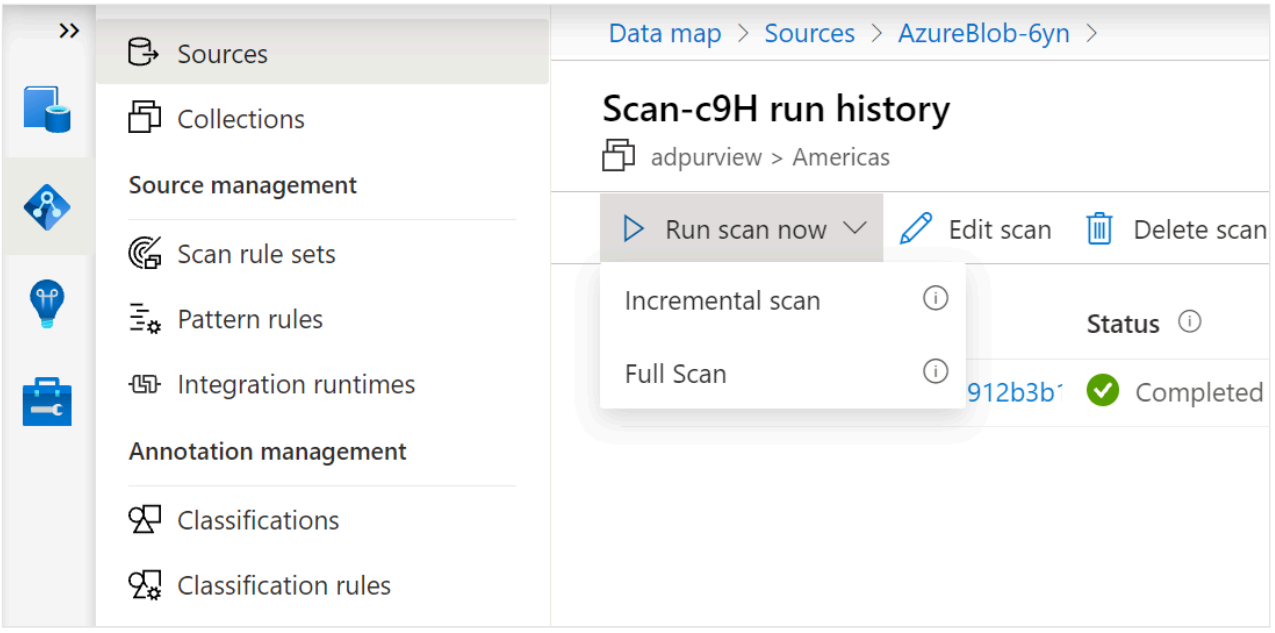
1. Select the **Scan name** to manage the scan



2. You can *run the scan* again, *edit the scan*, *delete the scan*



3. You can *run an incremental scan* or a *full scan* again.



Data sharing

Microsoft Purview Data Sharing (preview) enables sharing of data in-place from Azure Blob storage account to Azure Blob storage account. This section provides details about the specific requirements for sharing and receiving data in-place between Azure Blob storage accounts. Refer to How to share data and How to receive share for step by step guides on how to use data sharing.

Storage accounts supported for in-place data sharing

The following storage accounts are supported for in-place data sharing:

- Regions: Canada Central, Canada East, UK South, UK West, Australia East, Japan East, Korea South, and South Africa North
- Redundancy options: LRS, GRS, RA-GRS
- Tiers: Hot, Cool

Only use storage accounts without production workload for the preview.

> **Note**
>
> Source and target storage accounts must be in the same region as each other. They don't need to be in the same region as the Microsoft Purview account.

Storage account permissions required to share data

To add or update a storage account asset to a share, you need ONE of the following permissions:

- **Microsoft.Authorization/roleAssignments/write** - This permission is available in the *Owner* role.
- **Microsoft.Storage/storageAccounts/blobServices/containers/blobs/modifyPermissions/** - This permission is available in the *Blob Storage Data Owner* role.

Storage account permissions required to receive shared data

To map a storage account asset in a received share, you need ONE of the following permissions:

- **Microsoft.Storage/storageAccounts/write** - This permission is available in the *Contributor* and *Owner* role.
- **Microsoft.Storage/storageAccounts/blobServices/containers/write** - This permission is available in the *Contributor*, *Owner*, *Storage Blob Data Contributor* and *Storage Blob Data Owner* role.

Update shared data in source storage account

Updates you make to shared files or data in the shared folder from source storage account will be made available to recipient in target storage account in near real time. When you delete subfolder or files within the shared folder, they will disappear for recipient. To delete the shared folder, file or parent folders or containers, you need to first revoke access to all your shares from the source storage account.

Access shared data in target storage account

The target storage account enables recipient to access the shared data read-only in near real time. You can connect analytics tools such as Synapse Workspace and Databricks to the shared data to perform analytics. Cost of accessing the shared data is charged to the target storage account.

Service limit

Source storage account can support up to 20 targets, and target storage account can support up to 100 sources. If you require an increase in limit, please contact Support.

Policies

The following types of policies are supported on this data resource from Microsoft Purview:

- Data owner policies - a set of policy statements that allow you to grant users and groups access to data sources.
- Self-service access policies - policy that allows users to request access to data sources registered to Microsoft Purview.
- Protection policies - denies access to data tagged with sensitivity labels to all users except those specified by the policy.

Access policy pre-requisites on Azure Storage accounts

Region support

- All Microsoft Purview regions are supported.
- Storage accounts in the following regions are supported without the need for additional configuration. However, zone-redundant storage (ZRS) accounts are not supported.
  - Australia Central
  - Australia East
  - Australia Southeast
  - Brazil South
  - Canada Central
  - Canada East
  - Central India
  - Central US
  - East Asia
  - East US 2
  - East US
  - France Central
  - Germany West Central
  - Japan East
  - Japan West
  - Korea Central
  - North Central US
  - North Europe
  - Norway East
  - Poland Central
  - Qatar Central
  - South Central US
  - South Africa North
  - Southeast Asia

    - South India
    - Sweden Central
    - Switzerland North
    - West Central US
    - West Europe
    - West US
    - West US 2
    - West US 3
    - UAE North
    - UK South
    - UK West
- Storage accounts in other regions in Public Cloud are supported after setting feature flag *AllowPurviewPolicyEnforcement*, as outlined in the next section. Newly created ZRS Storage accounts are supported, if created after setting the feature flag *AllowPurviewPolicyEnforcement*.

If needed, you can create a new Storage account by following this guide.

Configure the subscription where the Azure Storage account resides for policies from Microsoft Purview

This step is only necessary in certain regions (see prior section). To enable Microsoft Purview to manage policies for one or more Azure Storage accounts, execute the following PowerShell commands in the subscription where you'll deploy your Azure Storage account. These PowerShell commands will enable Microsoft Purview to manage policies on all Azure Storage accounts in that subscription.

If you're executing these commands locally, be sure to run PowerShell as an administrator. Alternatively, you can use the Azure Cloud Shell in the Azure portal: https://shell.azure.com.

PowerShell

```
# Install the Az module
Install-Module -Name Az -Scope CurrentUser -Repository PSGallery -Force
# Login into the subscription
Connect-AzAccount -Subscription <SubscriptionID>
# Register the feature
Register-AzProviderFeature -FeatureName AllowPurviewPolicyEnforcement -
ProviderNamespace Microsoft.Storage
```

If the output of the last command shows *RegistrationState* as *Registered*, then your subscription is enabled for access policies. If the output is *Registering*, wait at least 10 minutes, and then retry the command. **Do not continue unless the RegistrationState shows as *Registered*.**

Configure the Microsoft Purview account for policies

Register the data source in Microsoft Purview

Before a policy can be created in Microsoft Purview for a data resource, you must register that data resource in Microsoft Purview Studio. You will find the instructions related to registering the data resource later in this guide.

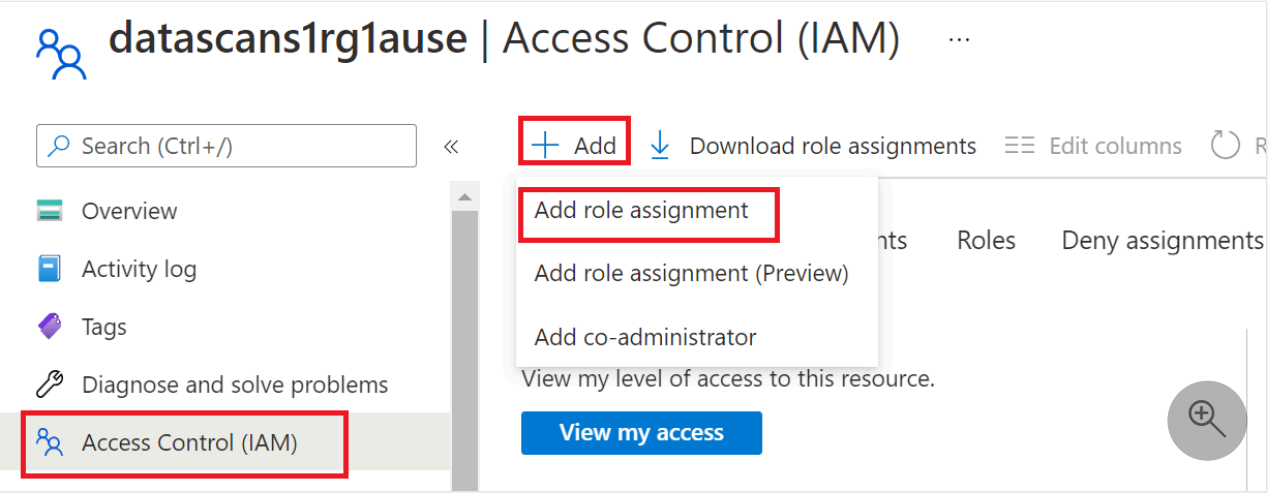
> ⓘ **Note**
>
> Microsoft Purview policies rely on the data resource ARM path. If a data resource is moved to a new resource group or subscription it will need to be de-registered and then registered again in Microsoft Purview.

Configure permissions to enable Data policy enforcement on the data source

Once a resource is registered, but before a policy can be created in Microsoft Purview for that resource, you must configure permissions. A set of permissions are needed to enable the **Data policy enforcement**. This applies to data sources, resource groups, or subscriptions. To enable *Data policy enforcement*, you must have *both* specific Identity and Access Management (IAM) privileges on the resource as well as specific Microsoft Purview privileges:

- You must have *either one* of the following IAM role combinations on the resource's Azure Resource Manager path or any parent of it (that is, using IAM permission inheritance):
  - IAM Owner
  - Both IAM Contributor and IAM User Access Administrator

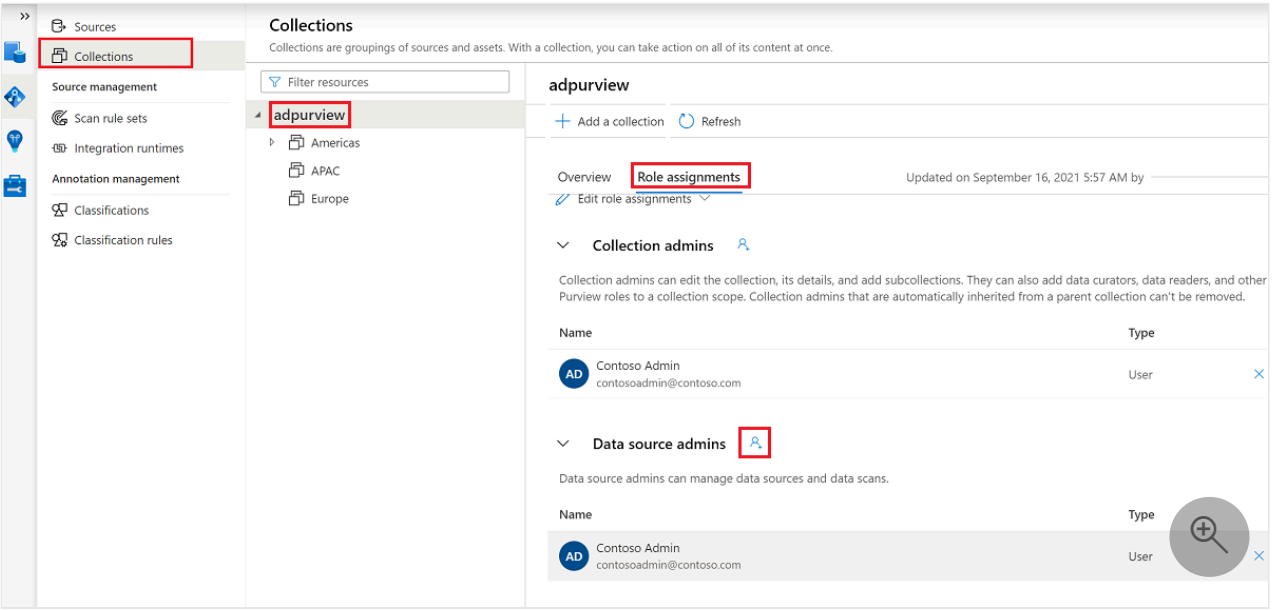
  To configure Azure role-based access control (RBAC) permissions, follow this guide. The following screenshot shows how to access the **Access Control** section in the Azure portal for the data resource to add a role assignment.



> **Note**
>
> The *IAM Owner* role for a data resource can be inherited from a parent resource group, a subscription, or a subscription management group. Check which Microsoft Entra users, groups, and service principals hold or are inheriting the *IAM Owner* role for the resource.

- You also need to have the Microsoft Purview *Data source admin* role for the collection or a parent collection (if inheritance is enabled). For more information, see the guide on managing Microsoft Purview role assignments.

  The following screenshot shows how to assign the *Data source admin* role at the root collection level.



Configure Microsoft Purview permissions to create, update, or delete access policies

To create, update or delete policies, you need to get the *Policy author* role in Microsoft Purview at *root collection level*:

- The *Policy author* role can create, update, and delete DevOps and Data Owner policies.
- The *Policy author* role can delete self-service access policies.

For more information about managing Microsoft Purview role assignments, see Create and manage collections in the Microsoft Purview Data Map.

> **Note**
>
> Policy author role must be configured at the root collection level.

In addition, to easily search Microsoft Entra users or groups when creating or updating the subject of a policy, you can greatly benefit from getting the Directory Readers permission in Microsoft Entra ID. This is a common permission for users in an Azure tenant. Without the Directory Reader permission, the Policy Author will have to type the complete username or email for all the principals included in the subject of a data policy.

Configure Microsoft Purview permissions for publishing Data Owner policies

Data Owner policies allow for checks and balances if you assign the Microsoft Purview *Policy author* and *Data source admin* roles to different people in the organization. Before a Data owner policy takes effect, a second person (*Data source admin*) must review it and explicitly approve it by publishing it. This does not apply to DevOps or Self-service access policies as publishing is automatic for them when those policies are created or updated.

To publish a Data owner policy you need to get the *Data source admin* role in Microsoft Purview at *root collection level*.

For more information about managing Microsoft Purview role assignments, see Create and manage collections in the Microsoft Purview Data Map.

> **Note**
>
> To publish Data owner policies, the Data source admin role must be configured at the root collection level.

Delegate access provisioning responsibility to roles in Microsoft Purview

After a resource has been enabled for **Data policy enforcement**, any Microsoft Purview user with the *Policy author* role at the root collection level can provision access to that data source from Microsoft Purview.

> ⓘ **Note**
>
> Any Microsoft Purview root *Collection admin* can assign new users to root *Policy author* roles. Any *Collection admin* can assign new users to a *Data source admin* role under the collection. Minimize and carefully vet the users who hold Microsoft Purview *Collection admin*, *Data source admin*, or *Policy author* roles.

If a Microsoft Purview account with published policies is deleted, such policies will stop being enforced within an amount of time that depends on the specific data source. This change can have implications on both security and data access availability. The Contributor and Owner roles in IAM can delete Microsoft Purview accounts. You can check these permissions by going to the **Access control (IAM)** section for your Microsoft Purview account and selecting **Role Assignments**. You can also use a lock to prevent the Microsoft Purview account from being deleted through Resource Manager locks.

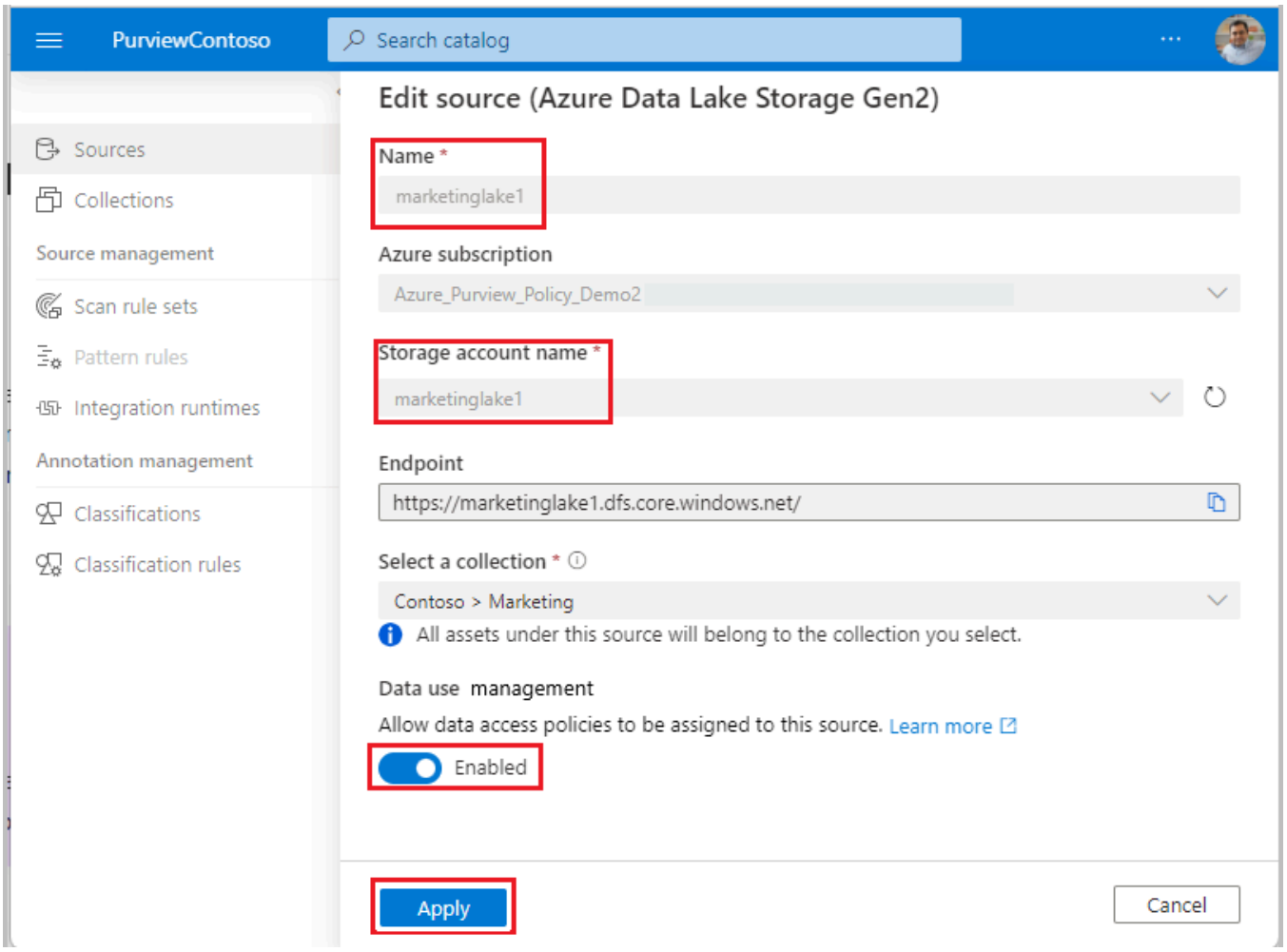
Register the data source in Microsoft Purview for Data Policy Enforcement

The Azure Storage resource needs to be registered first with Microsoft Purview before you can create access policies. To register your resource, follow the **Prerequisites** and **Register** sections of this guide:

- Register Azure Blob Storage in Microsoft Purview

After you've registered the data source, you'll need to enable Data Policy Enforcement. This is a pre-requisite before you can create policies on the data source. Data Policy Enforcement can impact the security of your data, as it delegates to certain Microsoft Purview roles managing access to the data sources. **Go through the secure practices related to Data Policy Enforcement in this guide**: How to enable Data Policy Enforcement

Once your data source has the **Data Policy Enforcement** option set to **Enabled**, it will look like this screenshot:



Create a policy

To create an access policy for Azure Blob Storage, follow this guide: Provision read/modify access on a single storage account.

To create policies that cover all data sources inside a resource group or Azure subscription you can refer to this section.

Protection policy

Protection access control policies (protection policies) enable organizations to automatically protect sensitive data across data sources. Microsoft Purview already scans data assets and identifies sensitive data elements, and this new feature allows you to automatically restrict access to that data using sensitivity labels from Microsoft Purview Information Protection.

Follow this documentation to create a protection policy: How to create a Microsoft Purview Information Protection policy.

Next steps

Follow the below guides to learn more about Microsoft Purview and your data.

- Data owner policies in Microsoft Purview
- Data Estate Insights in Microsoft Purview
- Data Sharing in Microsoft Purview
- Lineage in Microsoft Purview
- Search Unified Catalog

Last updated on 01/22/2026

Connect to and manage Azure Files in Microsoft Purview

Article • 05/21/2025

This article outlines how to register Azure Files, and how to authenticate and interact with Azure Files in Microsoft Purview. For more information about Microsoft Purview, read the introductory article.

Supported capabilities

Scanning capabilities

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan
Yes	Yes	Yes	Yes

Azure Files supports full and incremental scans to capture the metadata and classifications, based on system default and custom classification rules.

For file types such as .csv, .tsv, .psv, and .ssv, the schema is extracted when the following logics are in place:

1. First row values are nonempty.
2. First row values are unique.
3. First row values are not a date or a number.

Other capabilities

For **classifications**, **sensitivity labels**, **policies**, **data lineage**, and **live view**, see the list of supported capabilities.

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- An active Microsoft Purview account.
- You'll need to be a Data Source Administrator and Data Reader to register a source and manage it in the Microsoft Purview governance portal. See our Microsoft Purview

Permissions page for details.

Register

This section describes how to register Azure Files in Microsoft Purview using the Microsoft Purview governance portal.

Authentication for registration

Currently there's only one way to set up authentication for Azure file shares:

- Account Key

Account Key to register

When authentication method selected is **Account Key**, you need to get your access key and store in the key vault:

1. Navigate to your storage account
2. Select **Security + networking** > **Access keys**
3. Copy your *key* and save it somewhere for the next steps
4. Navigate to your key vault
5. Select **Objects** > **Secrets**
6. Select **+ Generate/Import** and enter the **Name** and **Value** as the *key* from your storage account
7. Select **Create** to complete
8. If your key vault isn't connected to Microsoft Purview yet, you'll need to create a new key vault connection
9. Finally, create a new credential using the key to set up your scan

Steps to register

To register a new Azure Files account in Microsoft Purview Unified Catalog, follow these steps:

1. Open the Microsoft Purview governance portal by:

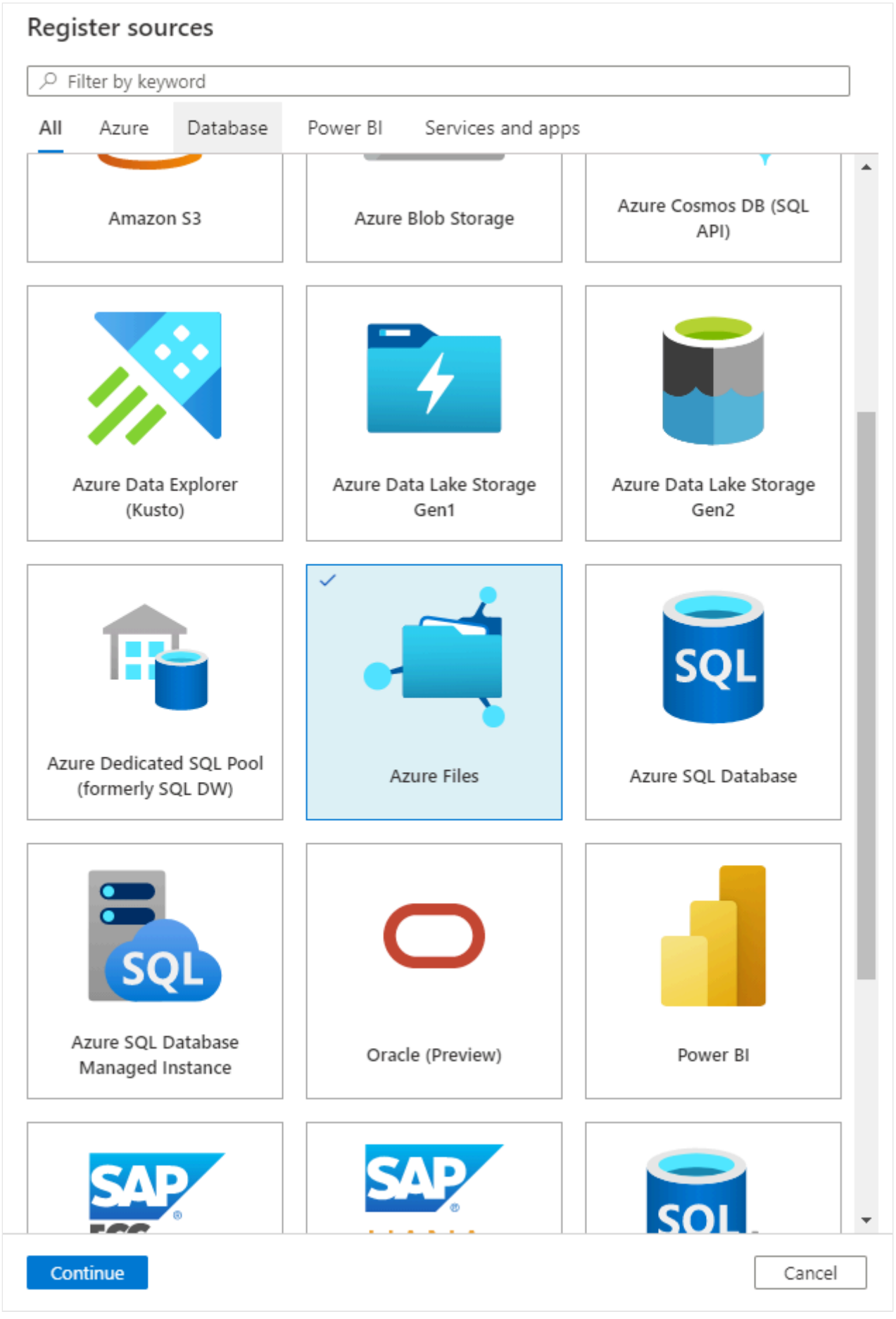
    - Browsing directly to https://web.purview.azure.com and selecting your Microsoft Purview account.
    - Opening the Azure portal, searching for and selecting the Microsoft Purview account. Selecting the **the Microsoft Purview governance portal** button.

2. Select **Data Map** on the left navigation.

3. Select **Register**

4. On **Register sources**, select **Azure Files**

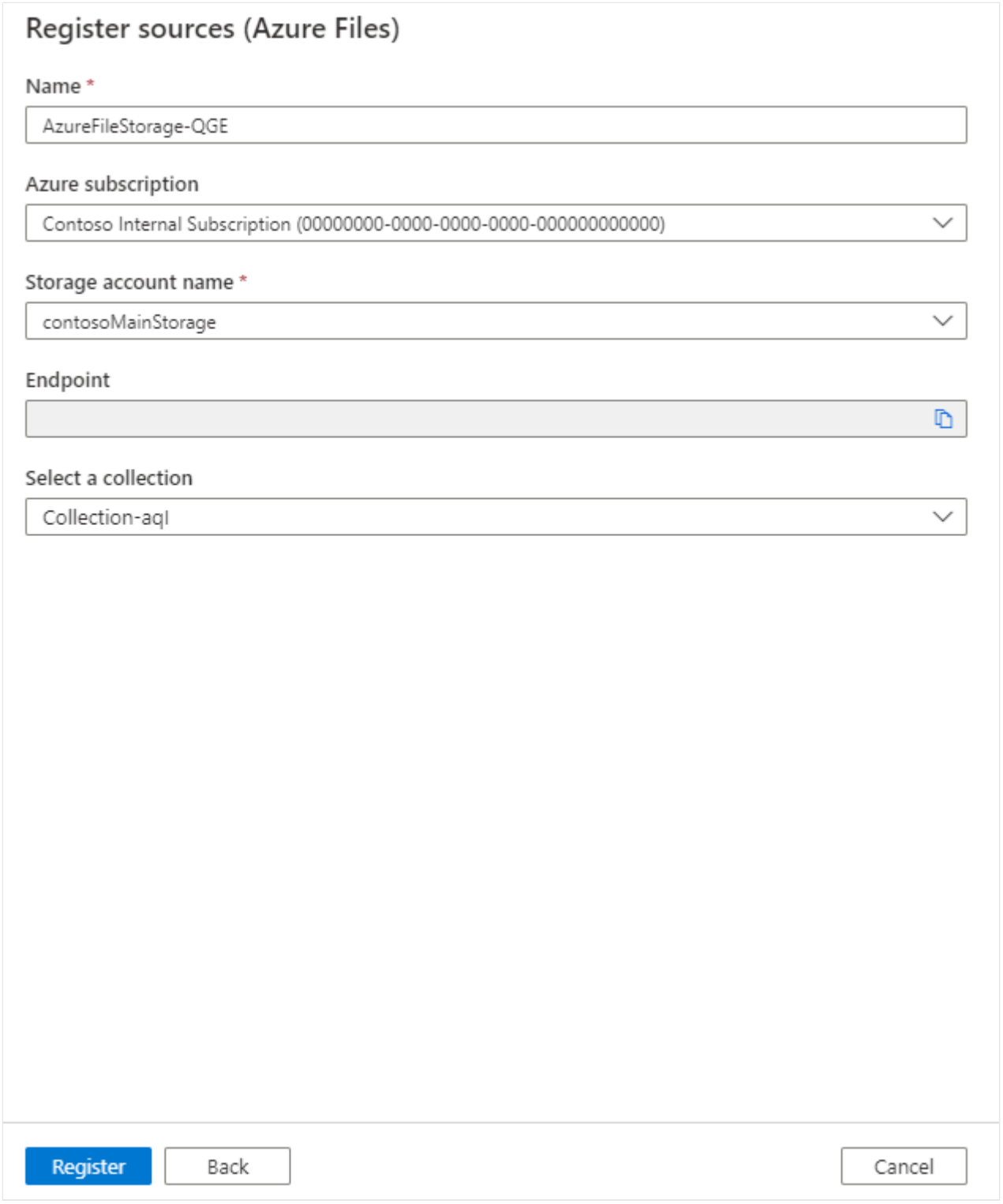
5. Select **Continue**



On the **Register sources (Azure Files)** screen, follow these steps:

1. Enter a **Name** that the data source will be listed with in the Catalog.
2. Choose your Azure subscription to filter down Azure Storage Accounts.

3. Select an Azure Storage Account.
4. Select a collection from the list.
5. Select **Register** to register the data source.



Scan

Follow the steps below to scan Azure Files to automatically identify assets and classify your data. For more information about scanning in general, see our introduction to scans and ingestion

Create and run scan

To create and run a new scan, follow these steps:

1. Select the **Data Map** tab on the left pane in the Microsoft Purview governance portal.

2. Select the Azure Files source that you registered.

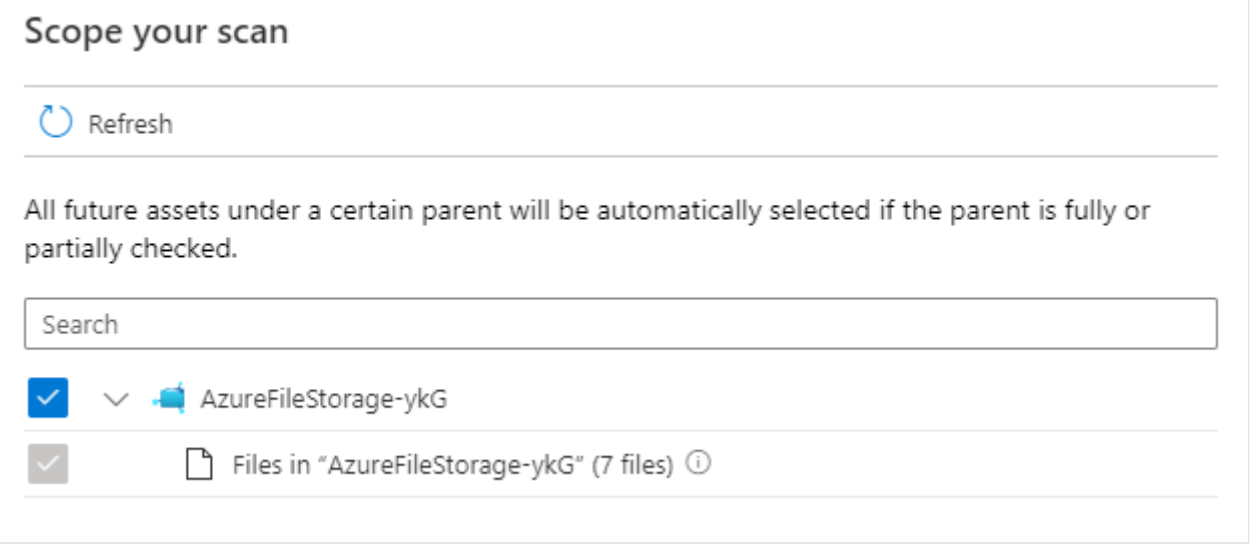
3. Select **New scan**

4. Choose either the Azure integration runtime if your source is publicly accessible, a managed virtual network integration runtime if using a managed virtual network, or a self-hosted integration runtime if your source is in a private virtual network. For more information about which integration runtime to use, see the choose the right integration runtime configuration article.

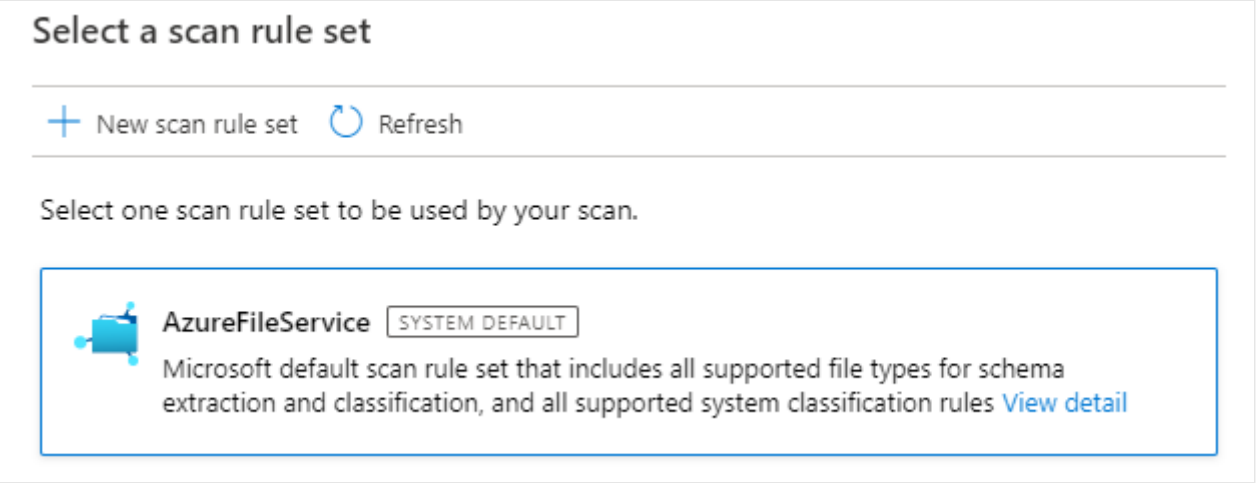
5. Select the account key credential to connect to your data source.



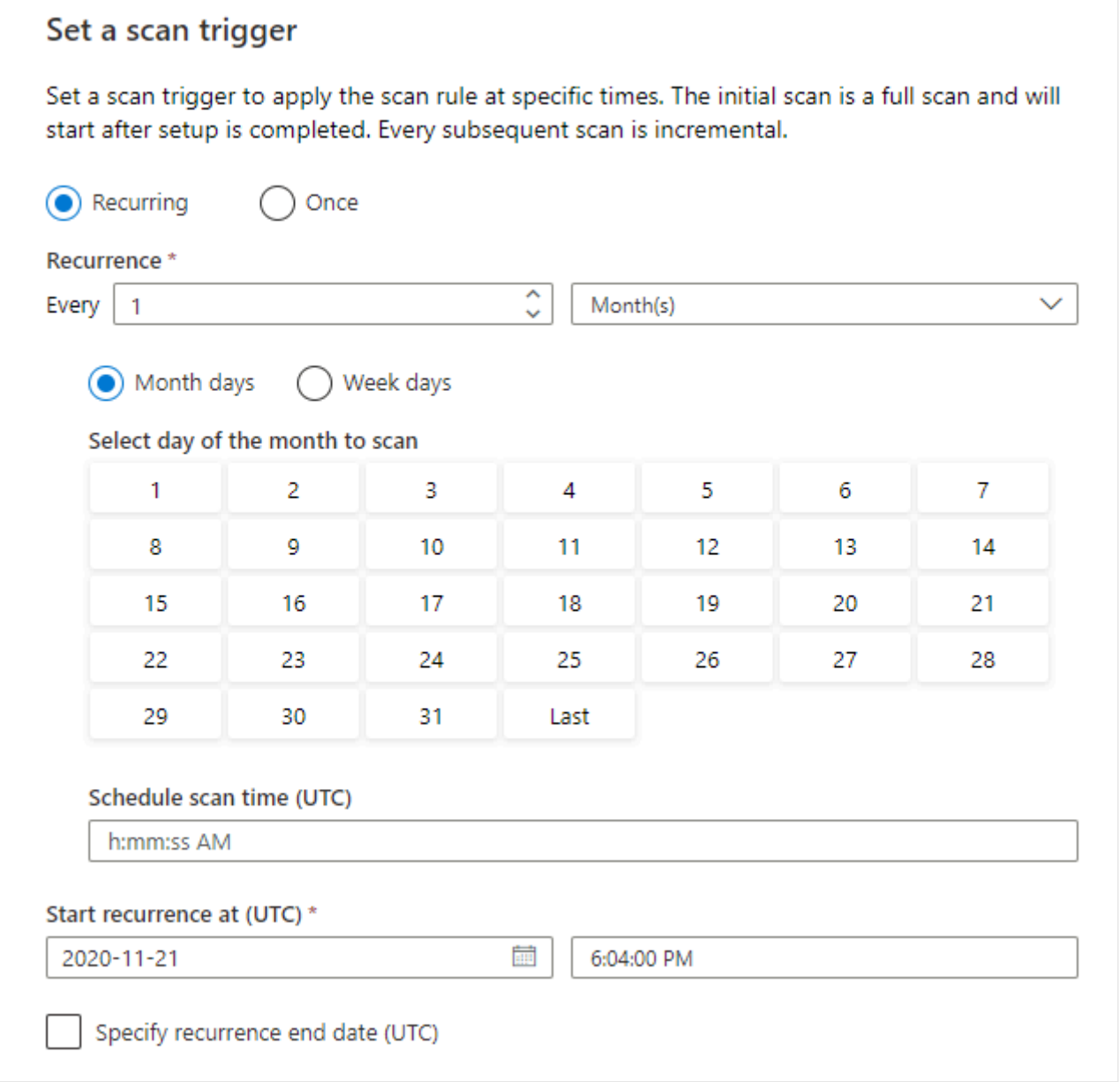
6. You can scope your scan to specific databases by choosing the appropriate items in the list.



7. Then select a scan rule set. You can choose between the system default, existing custom rule sets, or create a new rule set inline.



8. Choose your scan trigger. You can set up a schedule to reoccur, or run the scan once.



9. Review your scan and select **Save and run**.

View your scans and scan runs

To view existing scans:

1. Go to the Microsoft Purview portal. On the left pane, select **Data map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that has results you want to view. The pane shows you all the previous scan runs, along with the status and metrics for each scan run.
4. Select the run ID to check the scan run details.

Manage your scans

To edit, cancel, or delete a scan:

1. Go to the Microsoft Purview portal. On the left pane, select **Data Map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that you want to manage. You can then:
    - Edit the scan by selecting **Edit scan**.
    - Cancel an in-progress scan by selecting **Cancel scan run**.
    - Delete your scan by selecting **Delete scan**.

> ⓘ **Note**
>
> - Deleting your scan does not delete catalog assets created from previous scans.

Next steps

Now that you have registered your source, follow the below guides to learn more about Microsoft Purview and your data.

- Data Estate Insights in Microsoft Purview
- Lineage in Microsoft Purview
- Search Unified Catalog

Connect to Azure Cosmos DB for SQL API in Microsoft Purview

Article • 05/21/2025

This article outlines the process to register and scan Azure Cosmos DB for SQL API instance in Microsoft Purview, including instructions to authenticate and interact with the Azure Cosmos DB database source

Supported capabilities

Scanning capabilities

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan
Yes	Yes	No	Yes

Other capabilities

For **classifications**, **sensitivity labels**, **policies**, **data lineage**, and **live view**, see the list of supported capabilities.

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- An active Microsoft Purview account.
- You will need to be a Data Source Administrator and Data Reader to register a source and manage it in the Microsoft Purview governance portal. See our Microsoft Purview Permissions page for details.

Register

This section will enable you to register the Azure Cosmos DB for SQL API instance and set up an appropriate authentication mechanism to ensure successful scanning of the data source.

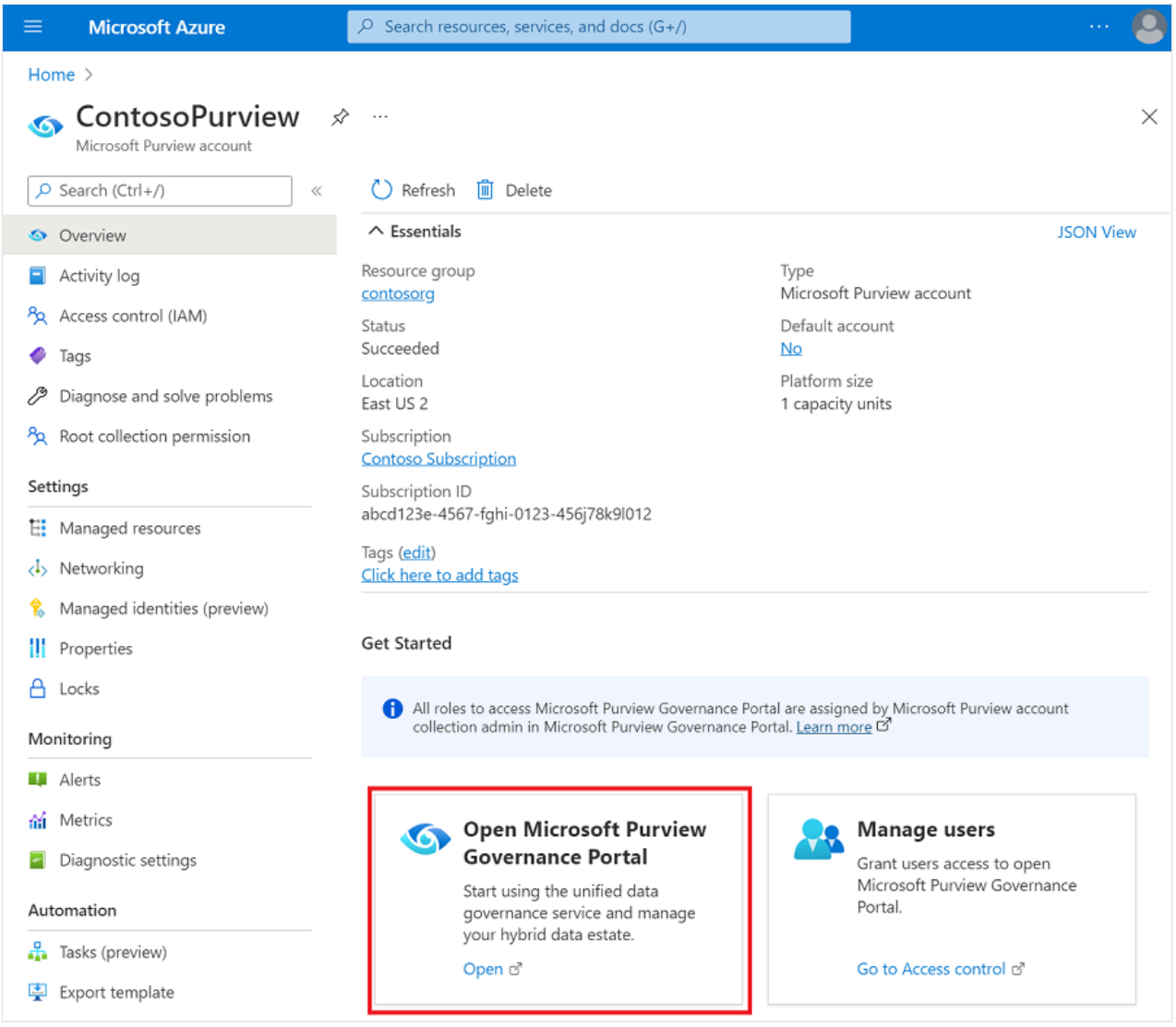
Steps to register

It is important to register the data source in Microsoft Purview prior to setting up a scan for the data source.

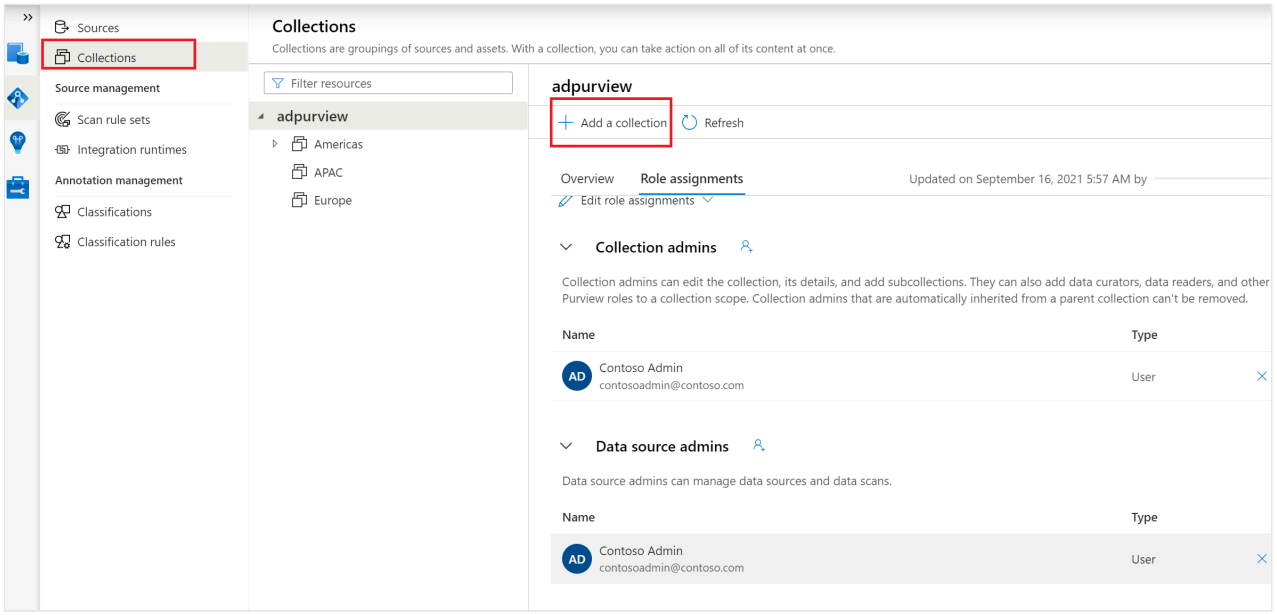
1. Open the Microsoft Purview governance portal by:

    - Browsing directly to https://web.purview.azure.com and selecting your Microsoft Purview account.
    - Opening the Azure portal, searching for and selecting the Microsoft Purview account. Select the **the Microsoft Purview governance portal** button.

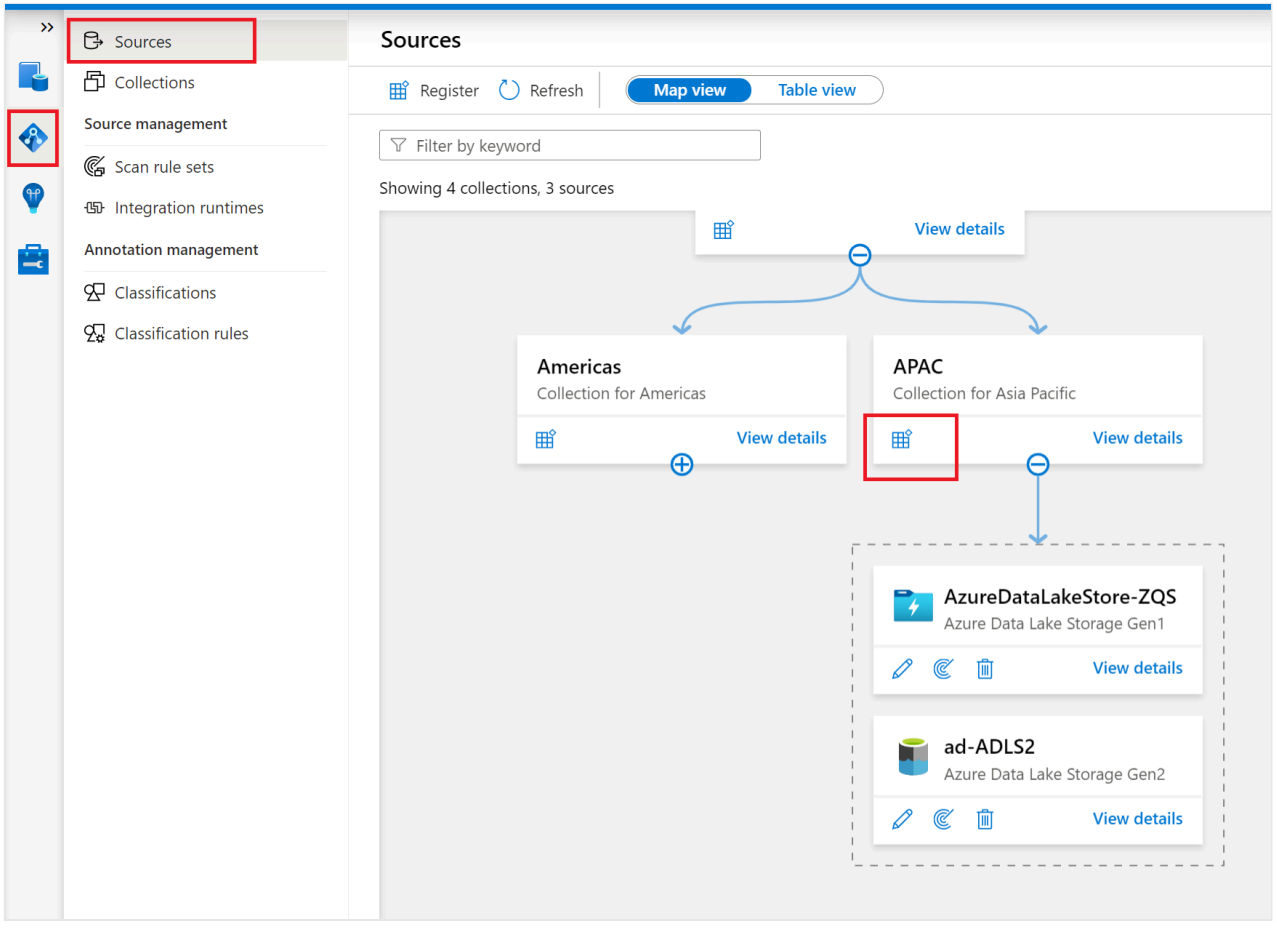
2. Navigate to the **Data Map --> Collections**



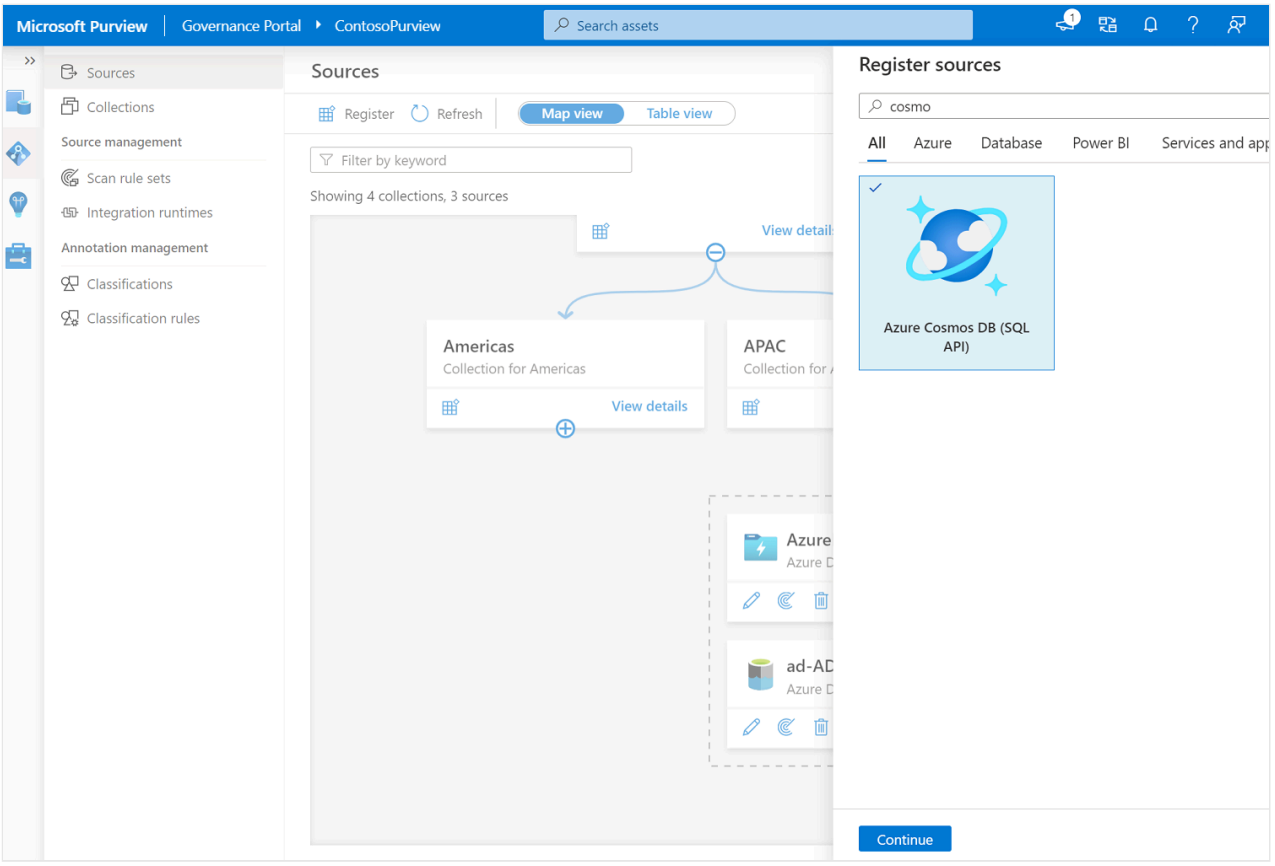
3. Create the Collection hierarchy using the **Collections** menu and assign permissions to individual subcollections, as required



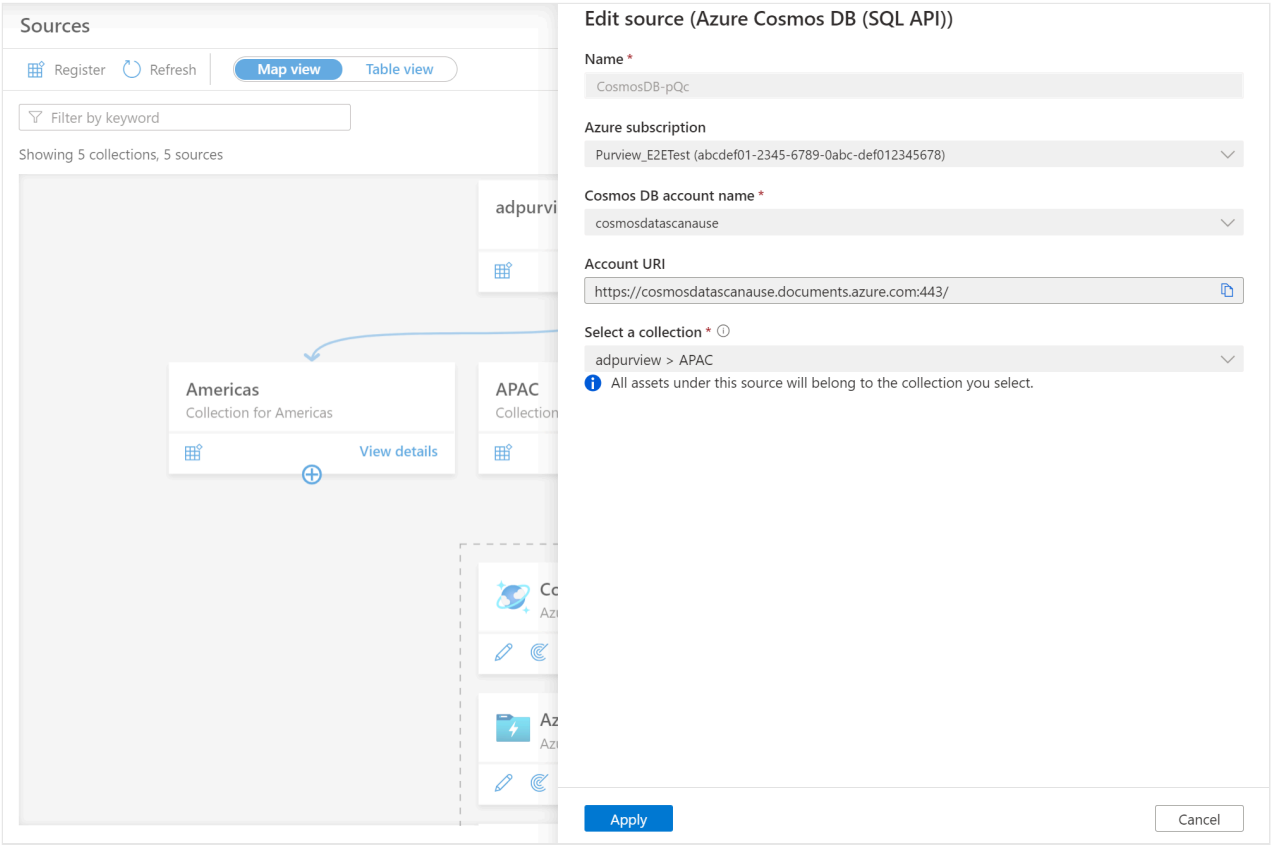
4. Navigate to the appropriate collection under the **Sources** menu and select the **Register** icon to register a new Azure Cosmos DB database



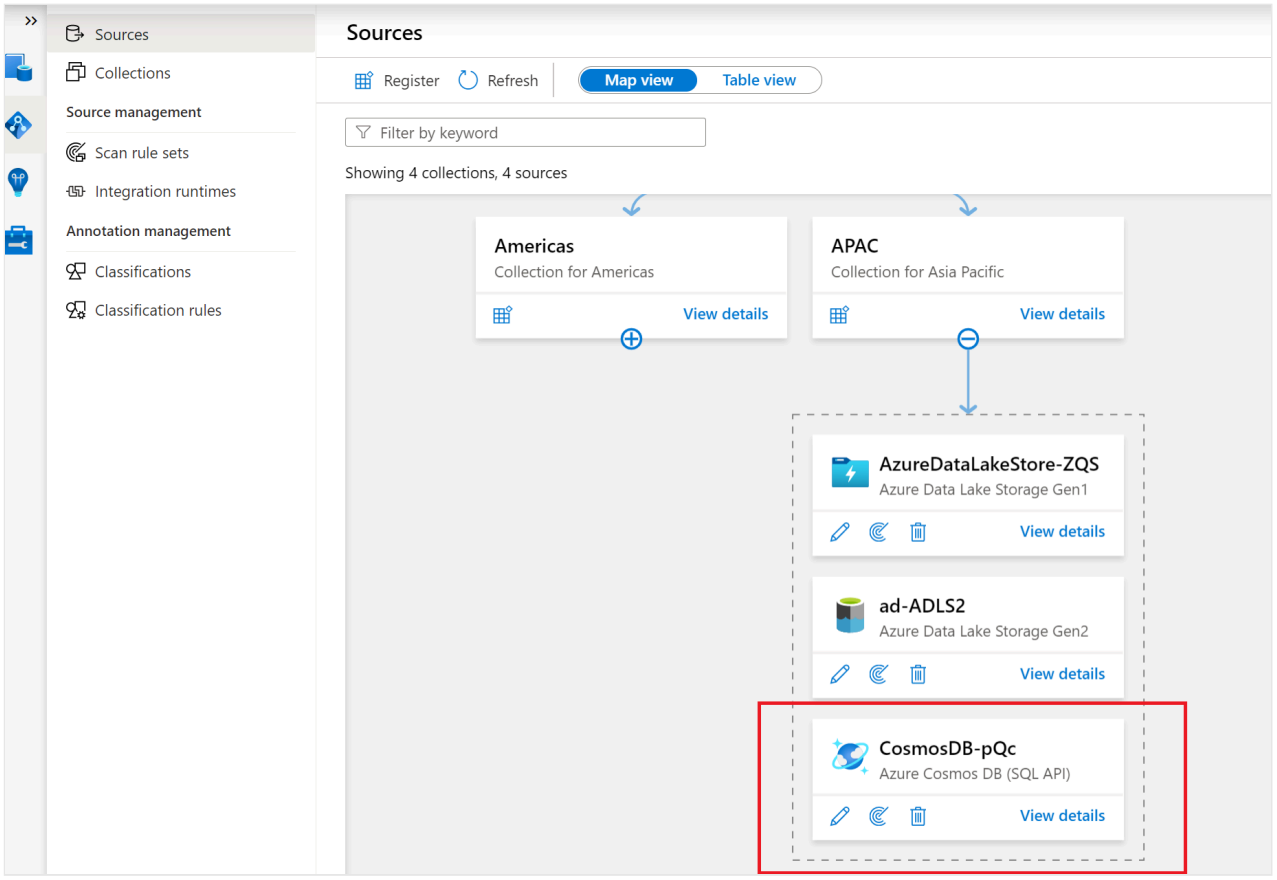
5. Select the **Azure Cosmos DB for SQL API** data source and select **Continue**



6. Provide a suitable **Name** for the data source, select the relevant **Azure subscription**, **Cosmos DB account name** and the **collection** and select **Apply**



7. The *Azure Cosmos DB database* storage account will be shown under the selected Collection



Scan

Authentication for a scan

In order to have access to scan the data source, an authentication method in the Azure Cosmos DB database Storage account needs to be configured.

There is only one way to set up authentication for Azure Cosmos DB Database:

Account Key - Secrets can be created inside an Azure Key Vault to store credentials in order to enable access for Microsoft Purview to scan data sources securely using the secrets. A secret can be a storage account key, SQL login password or a password.

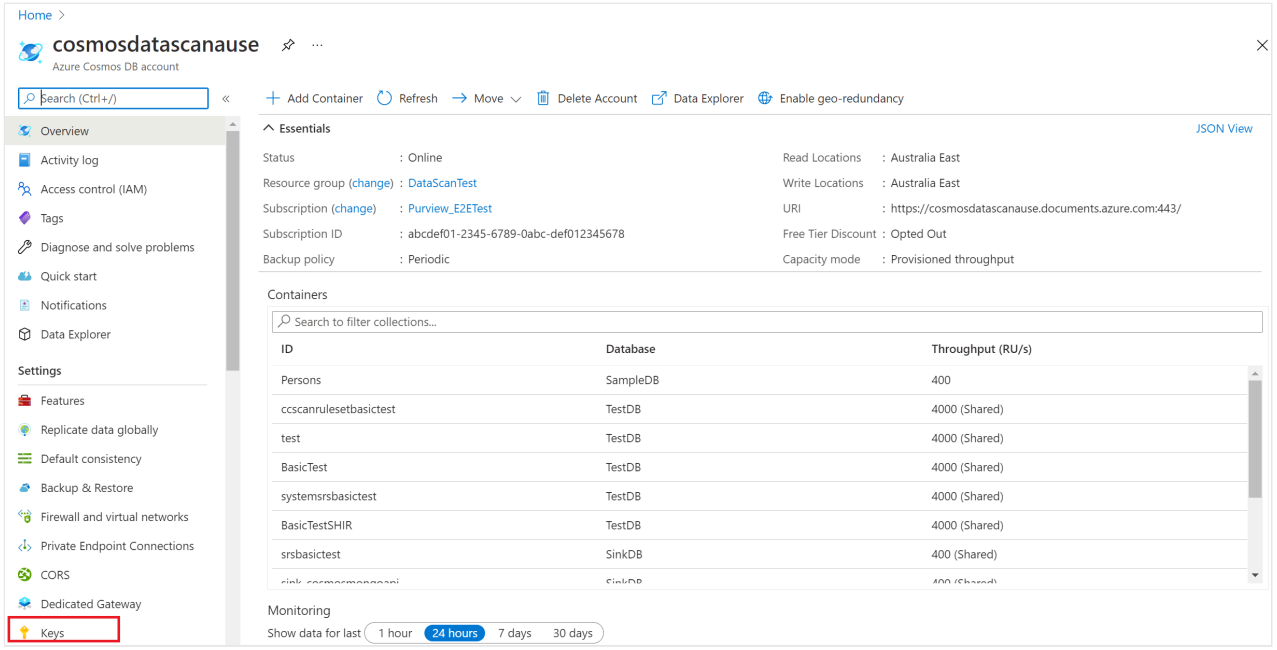
> ⓘ **Note**
>
> You need to deploy an *Azure key vault* resource in your subscription and assign *Microsoft Purview account's* MSI with required access permission to secrets inside *Azure key vault*.

Using Account Key for scanning

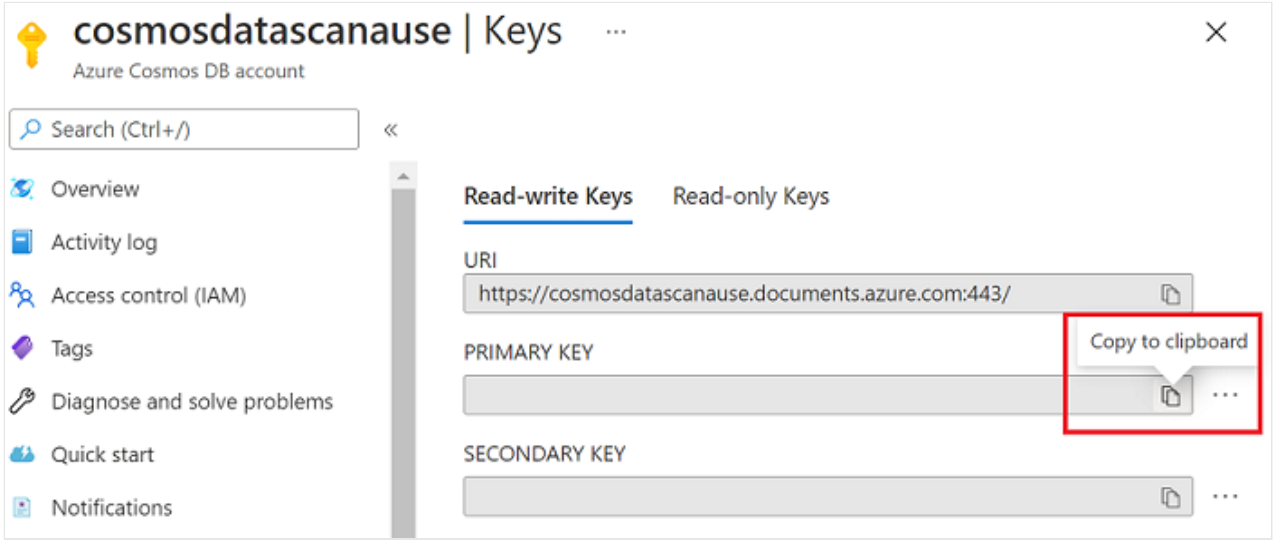
You need to get your access key and store in the key vault:

1. Navigate to your Azure Cosmos DB database storage account

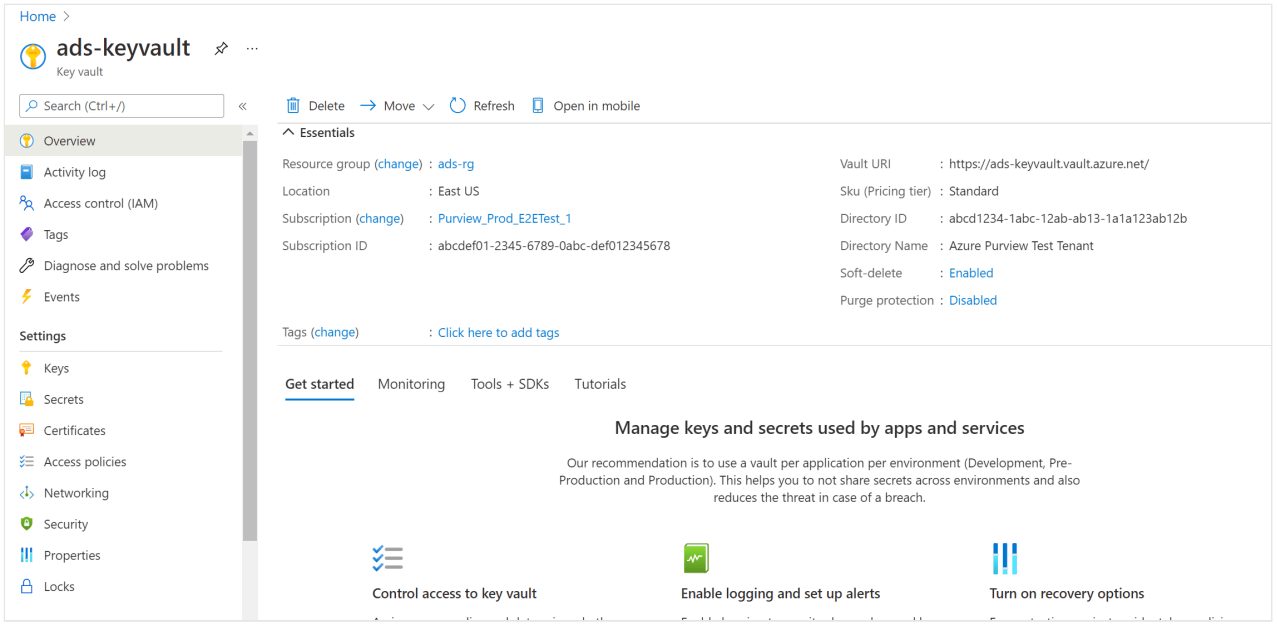
2. Select **Settings** > **Keys**



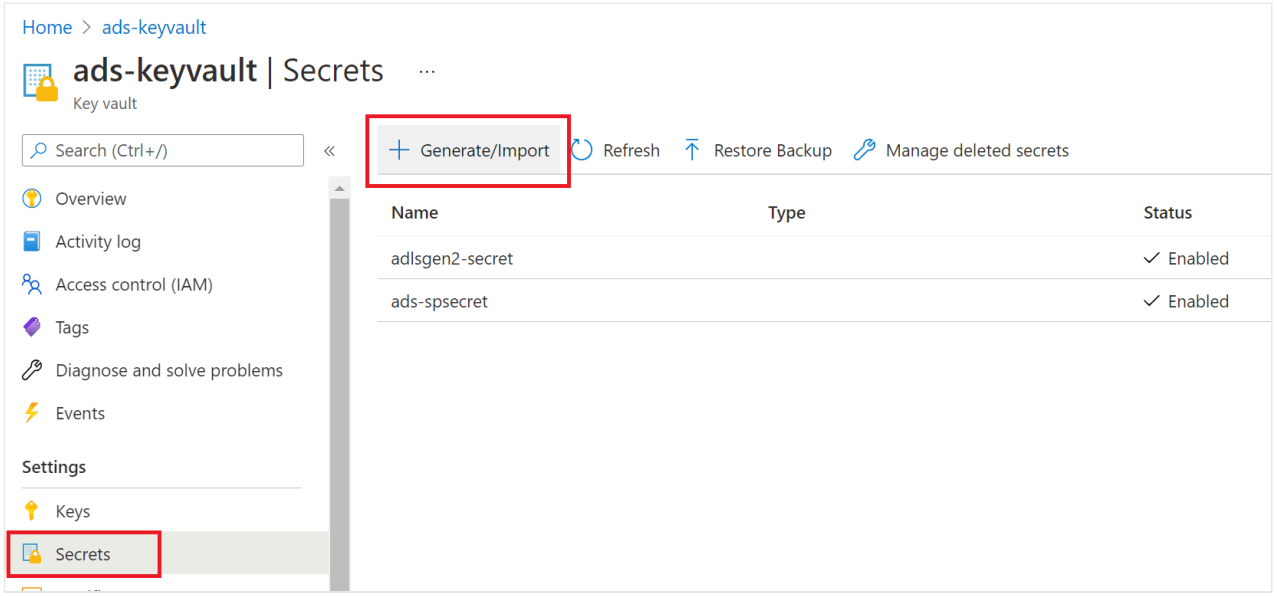
3. Copy your *key* and save it separately for the next steps



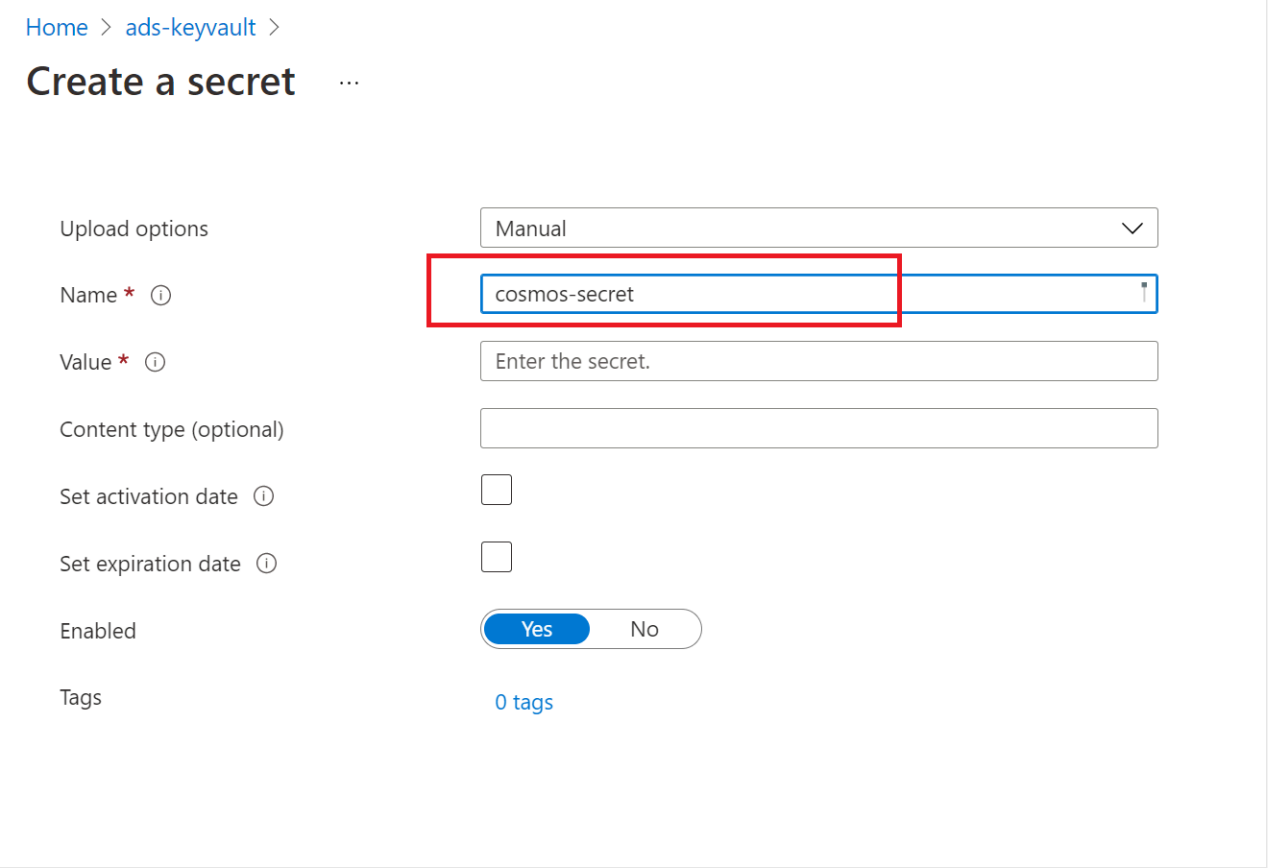
4. Navigate to your key vault



5. Select **Settings** > **Secrets** and select **+ Generate/Import**



6. Enter the **Name** and **Value** as the *key* from your storage account and Select **Create** to complete



7. If your key vault is not connected to Microsoft Purview yet, you will need to create a new key vault connection

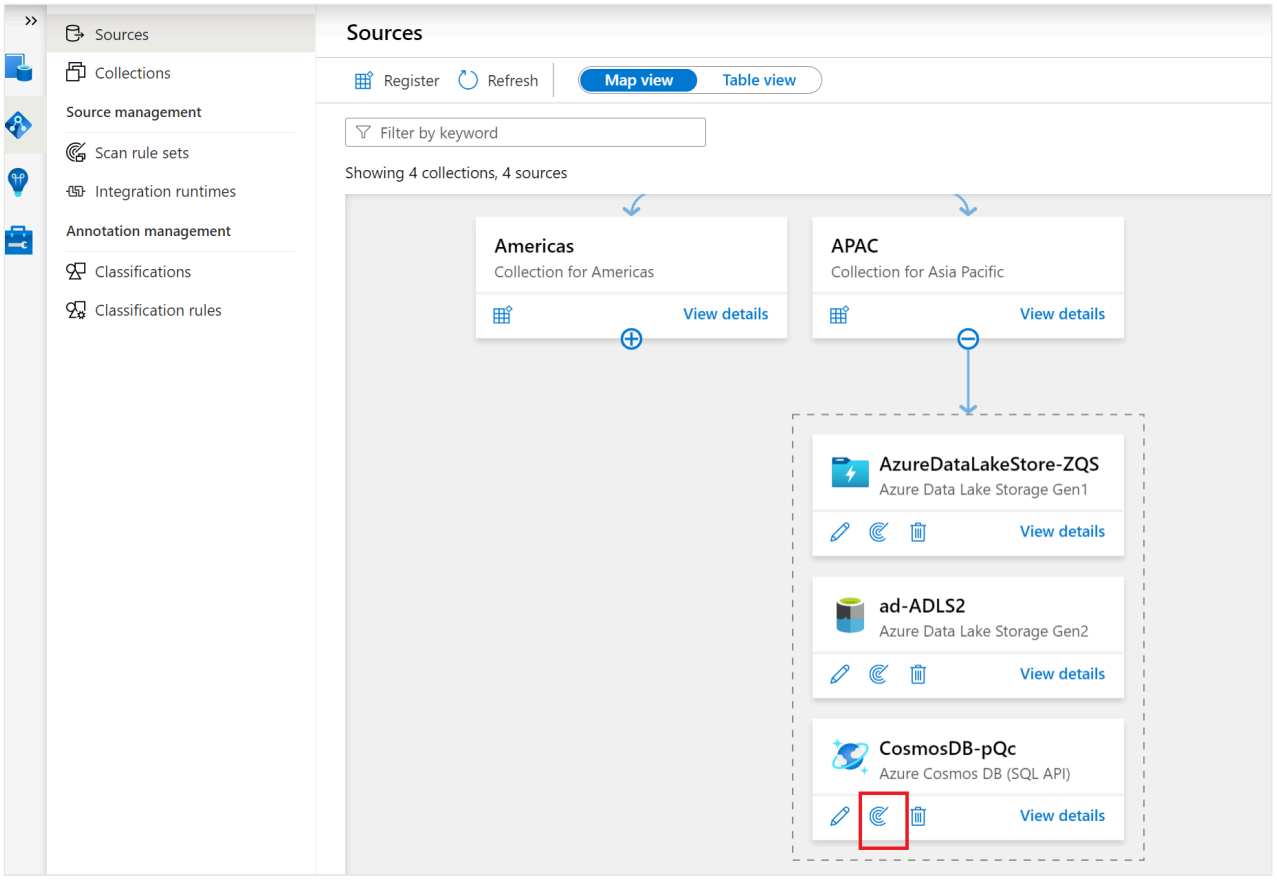
8. Finally, create a new credential using the key to set up your scan.

Creating the scan

1. Open your **Microsoft Purview account** and select the **Open Microsoft Purview governance portal**

2. Navigate to the **Data map** --> **Sources** to view the collection hierarchy

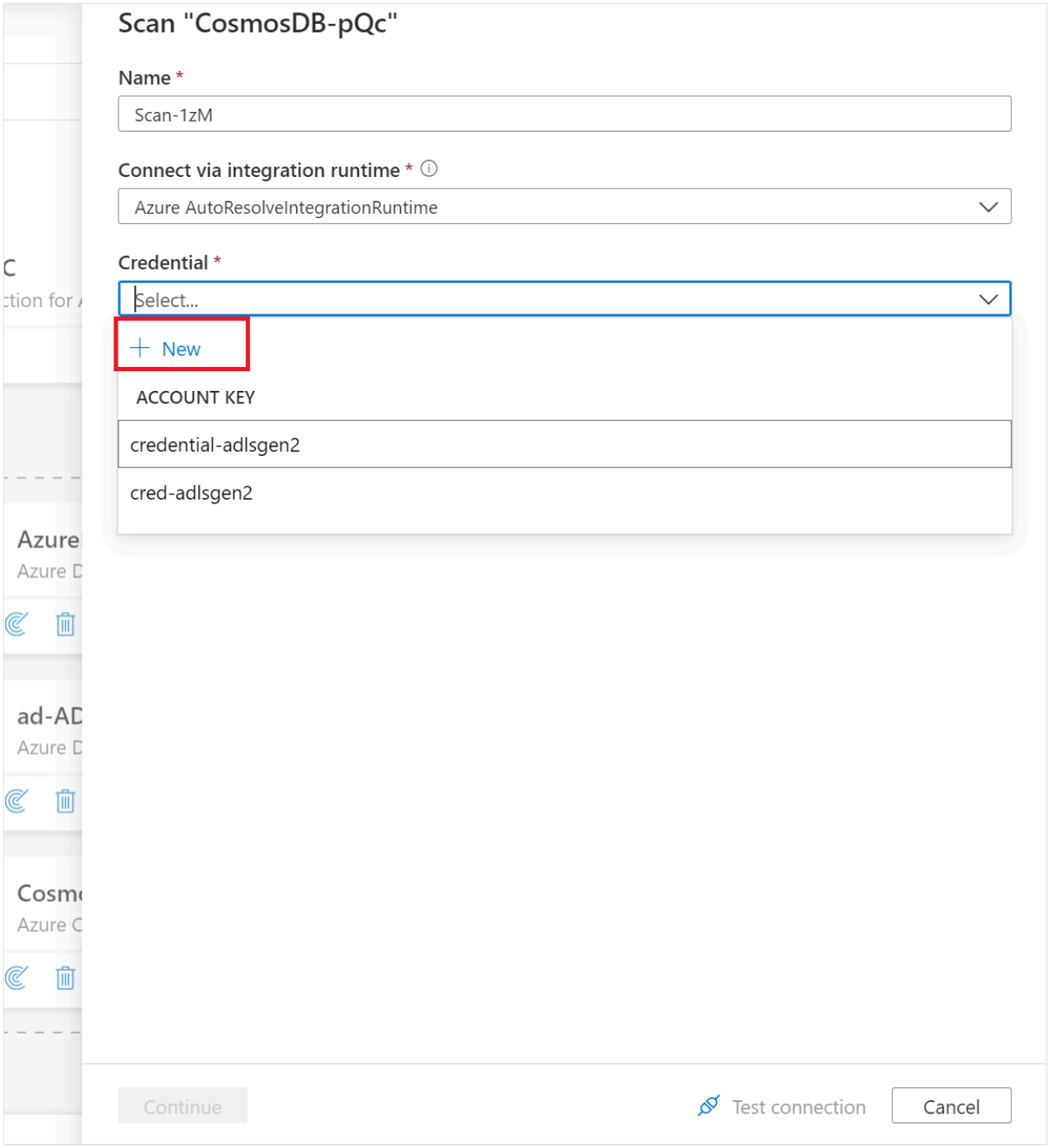
3. Select the **New Scan** icon under the **Azure Cosmos database** registered earlier



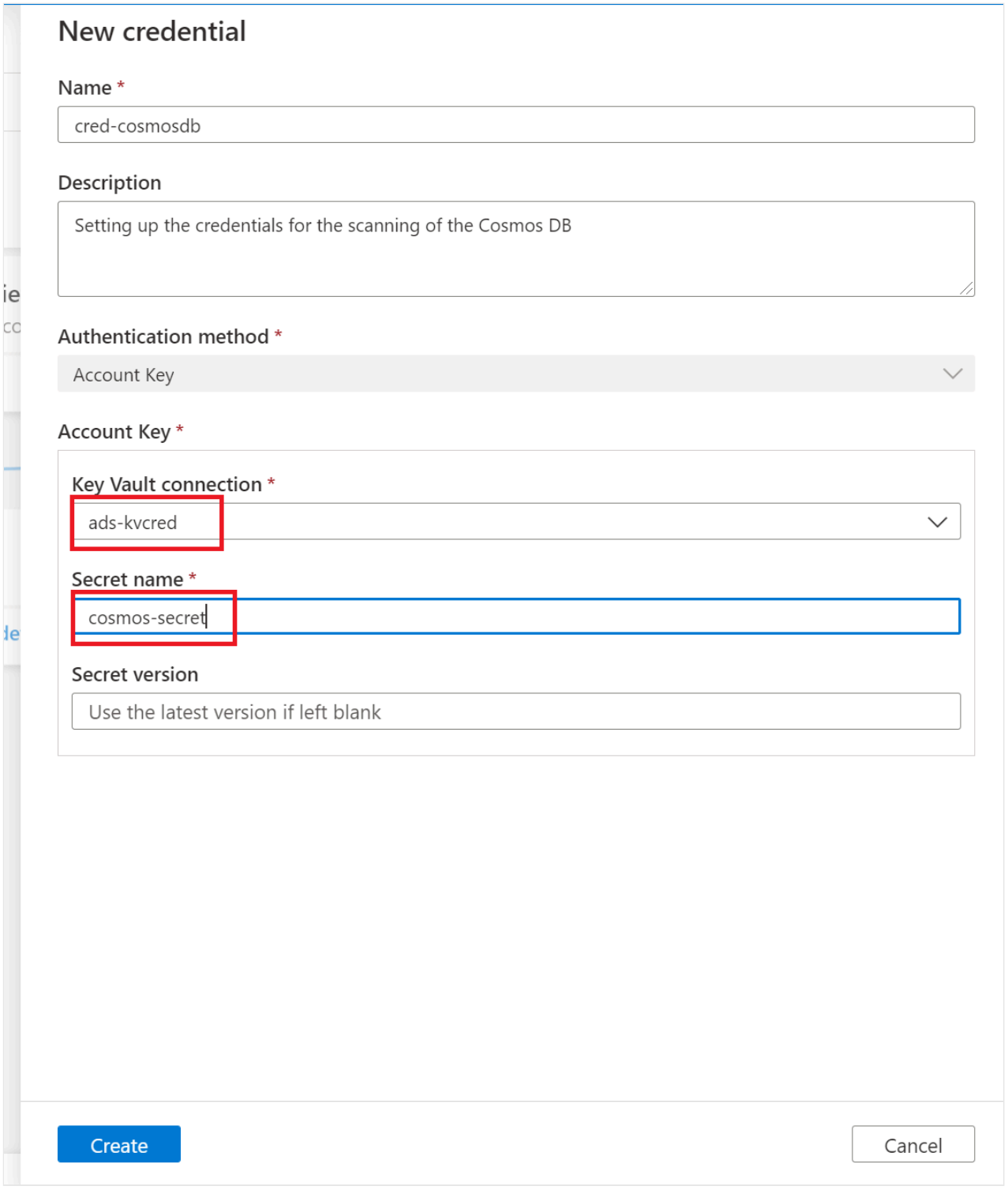
4. Provide a **Name** for the scan.

5. Choose either the Azure integration runtime if your source is publicly accessible, a managed virtual network integration runtime if using a managed virtual network, or a self-hosted integration runtime if your source is in a private virtual network. For more information about which integration runtime to use, see the choose the right integration runtime configuration article.

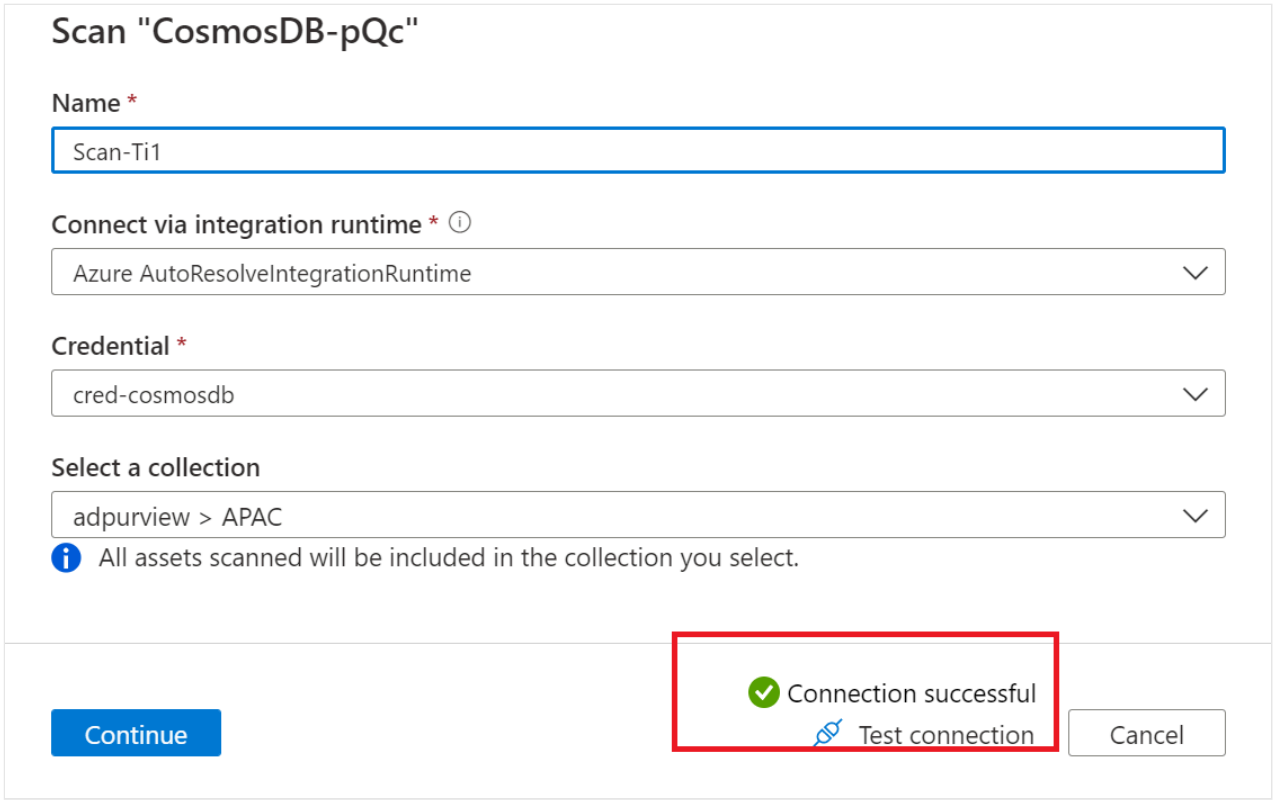
6. Choose the appropriate collection for the scan and select + **New** under **Credential**



7. Select the appropriate **Key vault connection** and the **Secret name** that was used while creating the *Account Key*. Choose **Authentication method** as *Account Key*

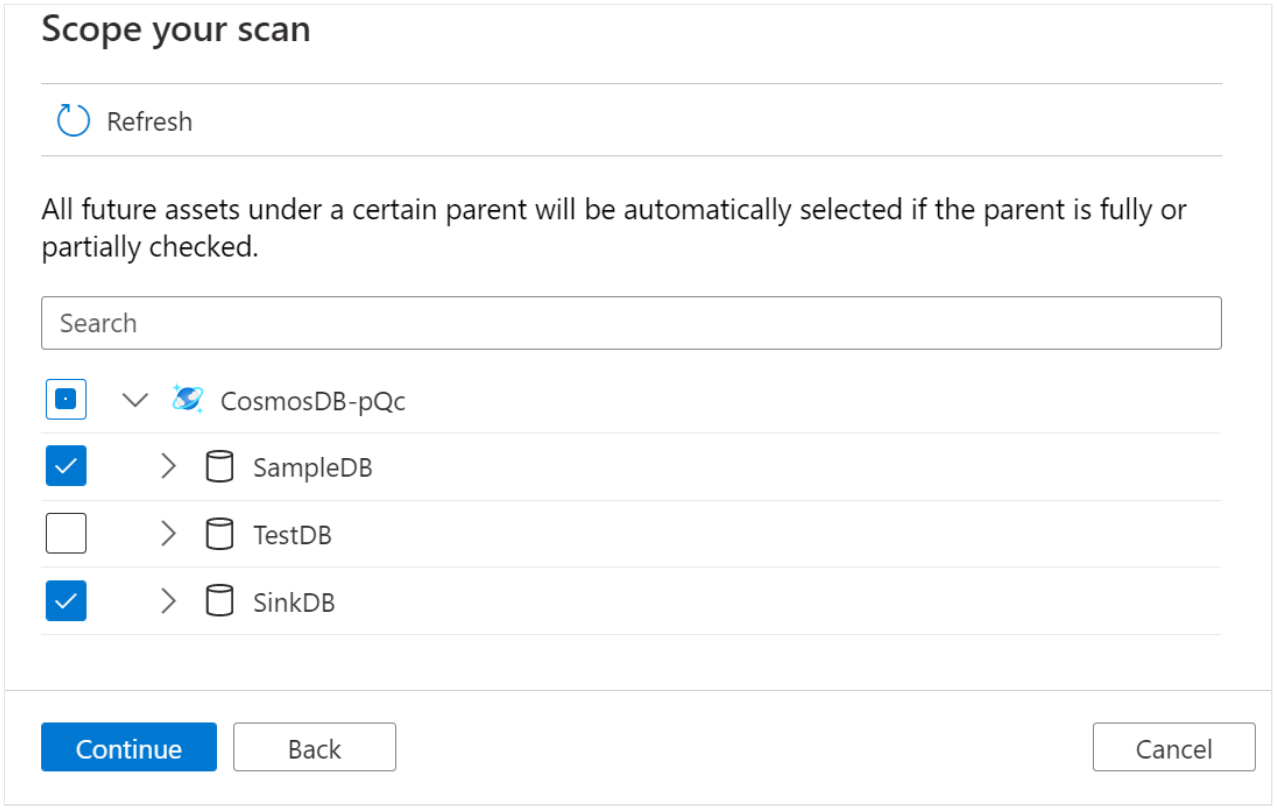


8. Select **Test connection**. On a successful connection, select **Continue**

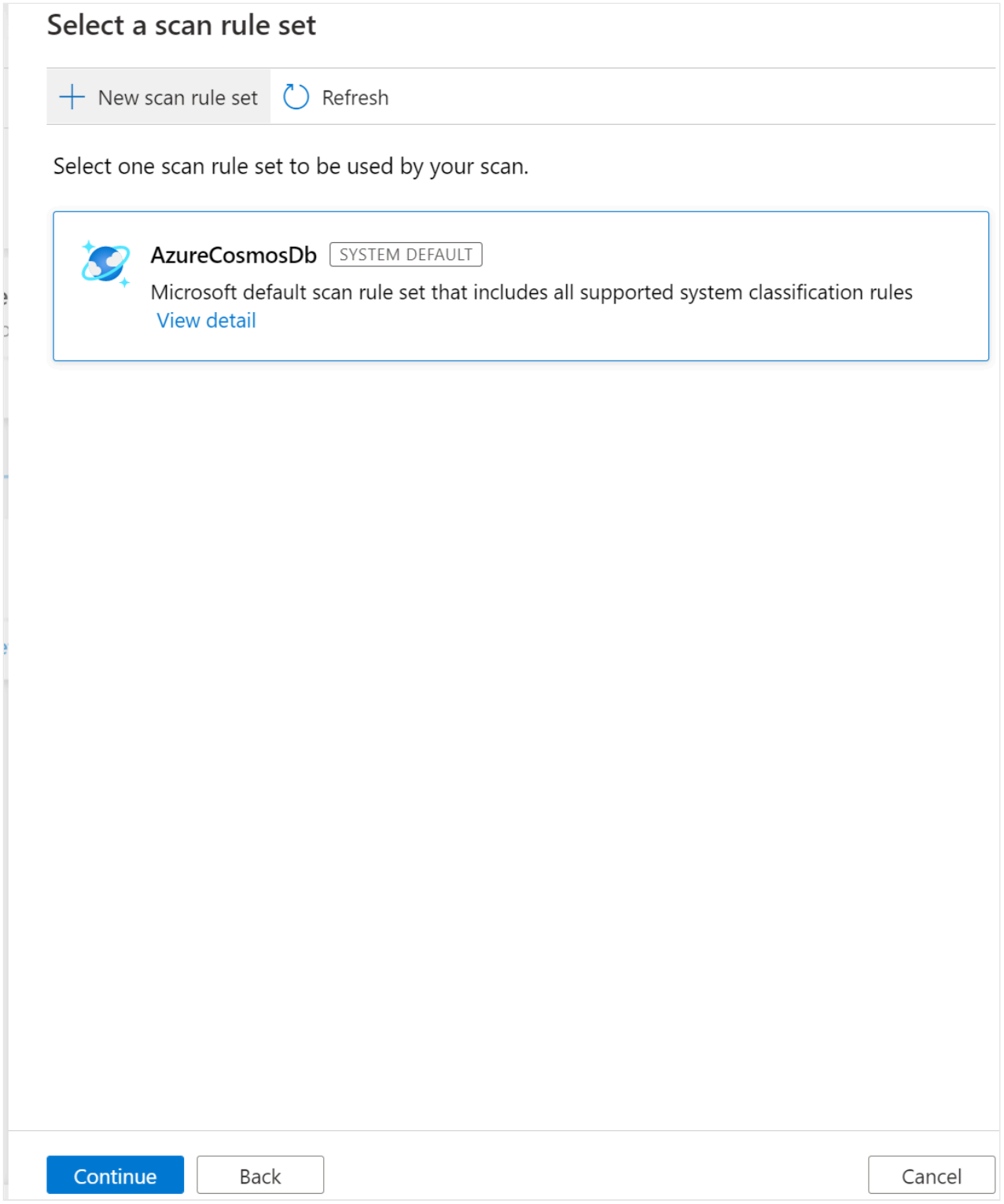


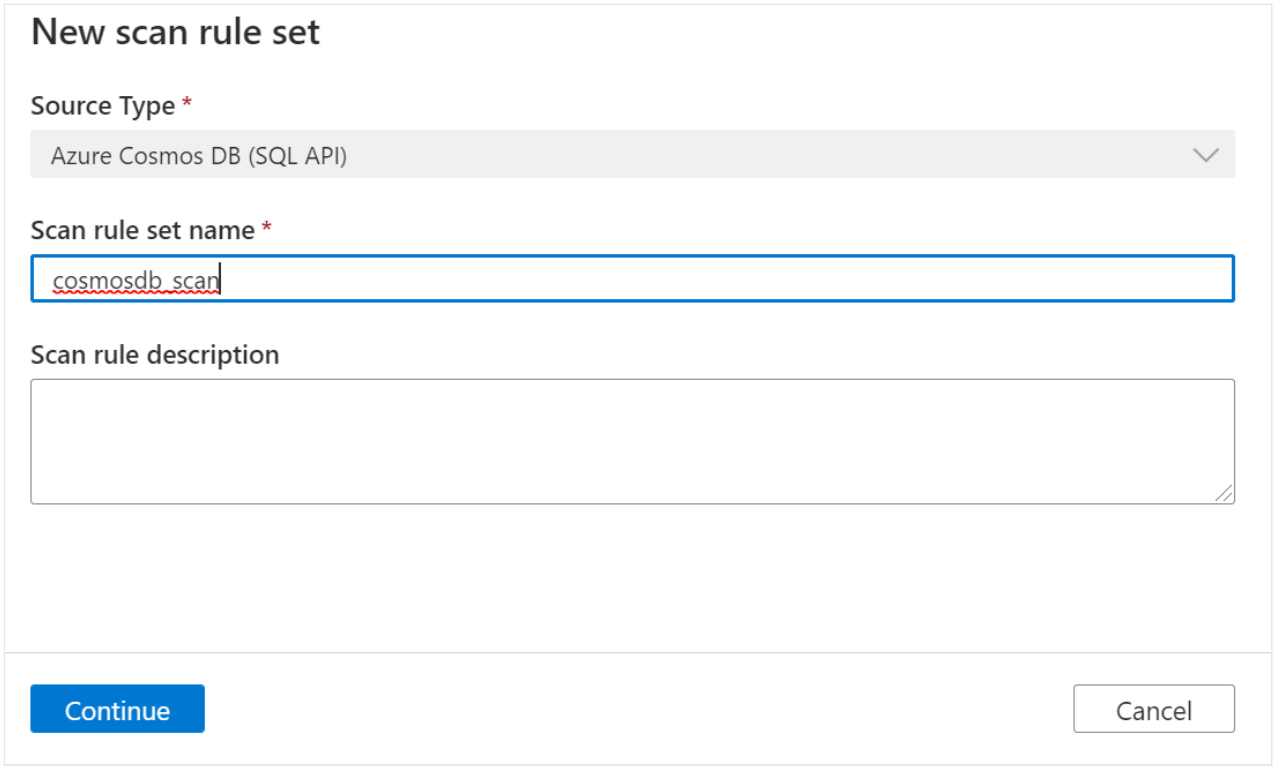
Scoping and running the scan

1. You can scope your scan to specific folders and subfolders by choosing the appropriate items in the list.



2. Then select a scan rule set. You can choose between the system default, existing custom rule sets, or create a new rule set inline.


Select a scan rule set
New scan rule set
Refresh
Select one scan rule set to be used by your scan.
AzureCosmosDb
SYSTEM DEFAULT
Microsoft default scan rule set that includes all supported system classification rules
View detail
Continue
Back
Cancel



3. You can select the **classification rules** to be included in the scan rule

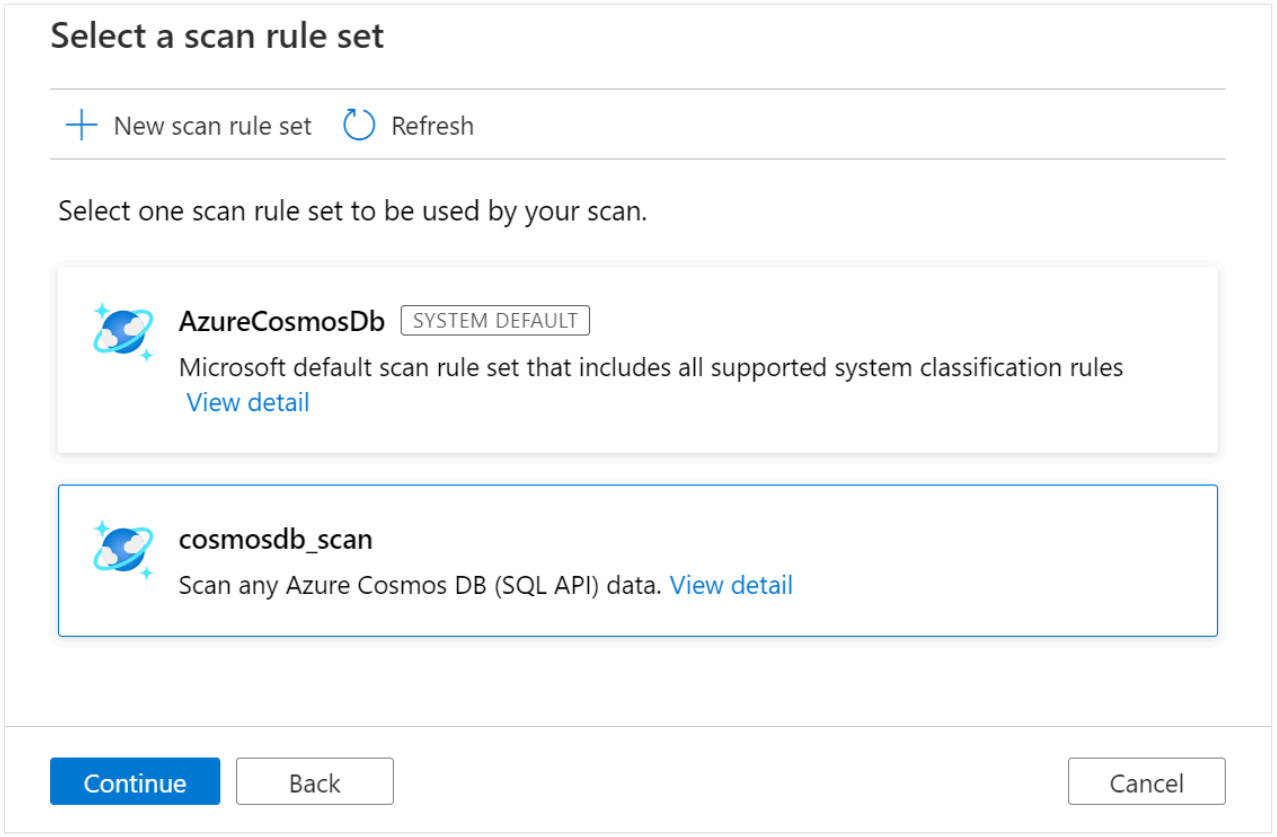
Select classification rules

Choose classification rules that will run on the dataset.

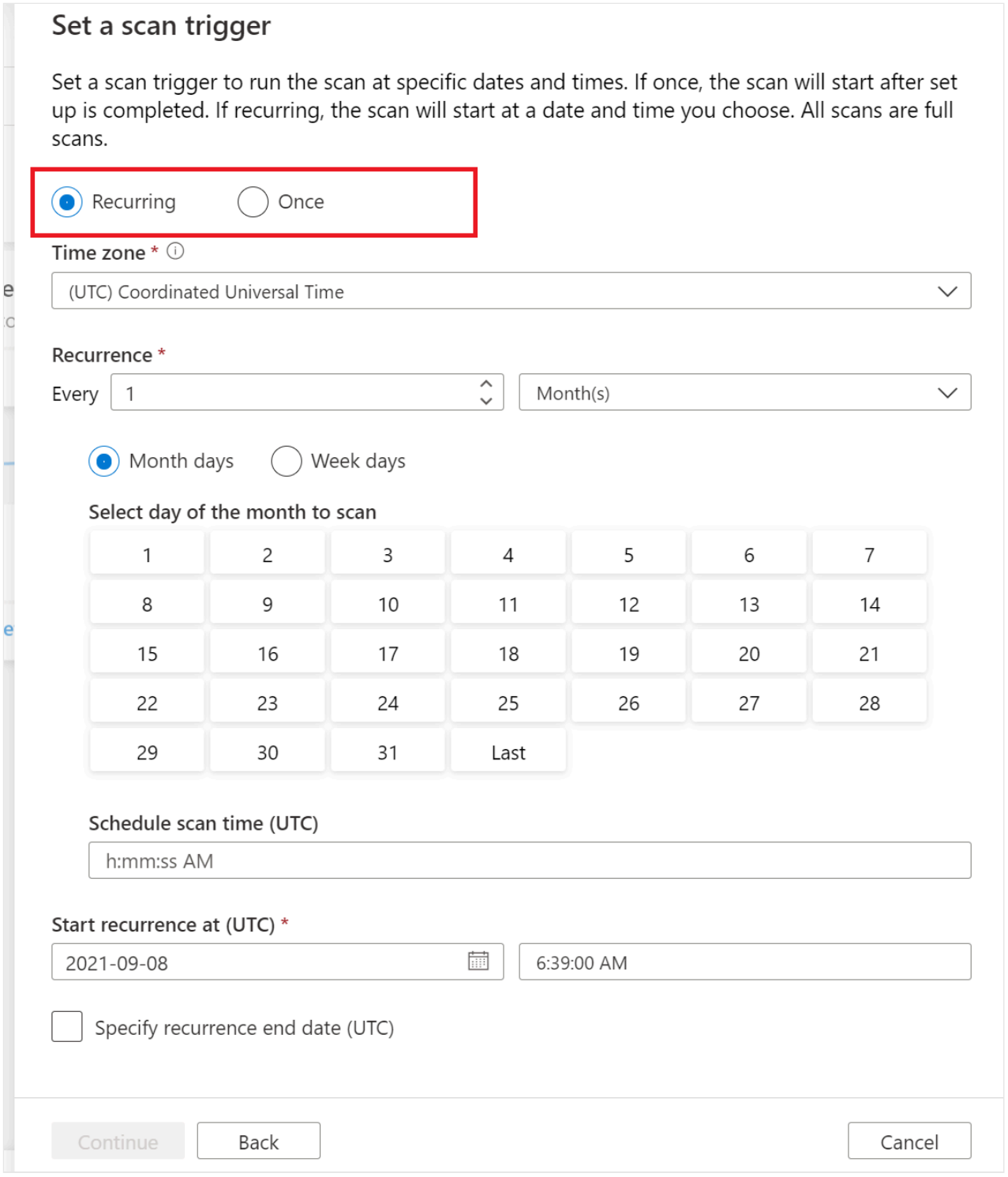
System rules

- [x] Government
- [x] Financial
- [-] Personal
  - [x] Age of an individual
  - [x] Date of Birth
  - [x] Email Address
  - [ ] Ethnic Group
  - [ ] EU GPS Coordinates
  - [ ] EU Mobile Phone Number
  - [ ] EU Phone Number
  - [ ] Geolocation (Lat/Lon)
  - [ ] Japanese My Number – Corporate
  - [ ] Japanese My Number – Personal
  - [ ] Person's Name
  - [x] Personal IP Address
  - [x] U.S. Phone Number
- [x] Security
- [x] Miscellaneous

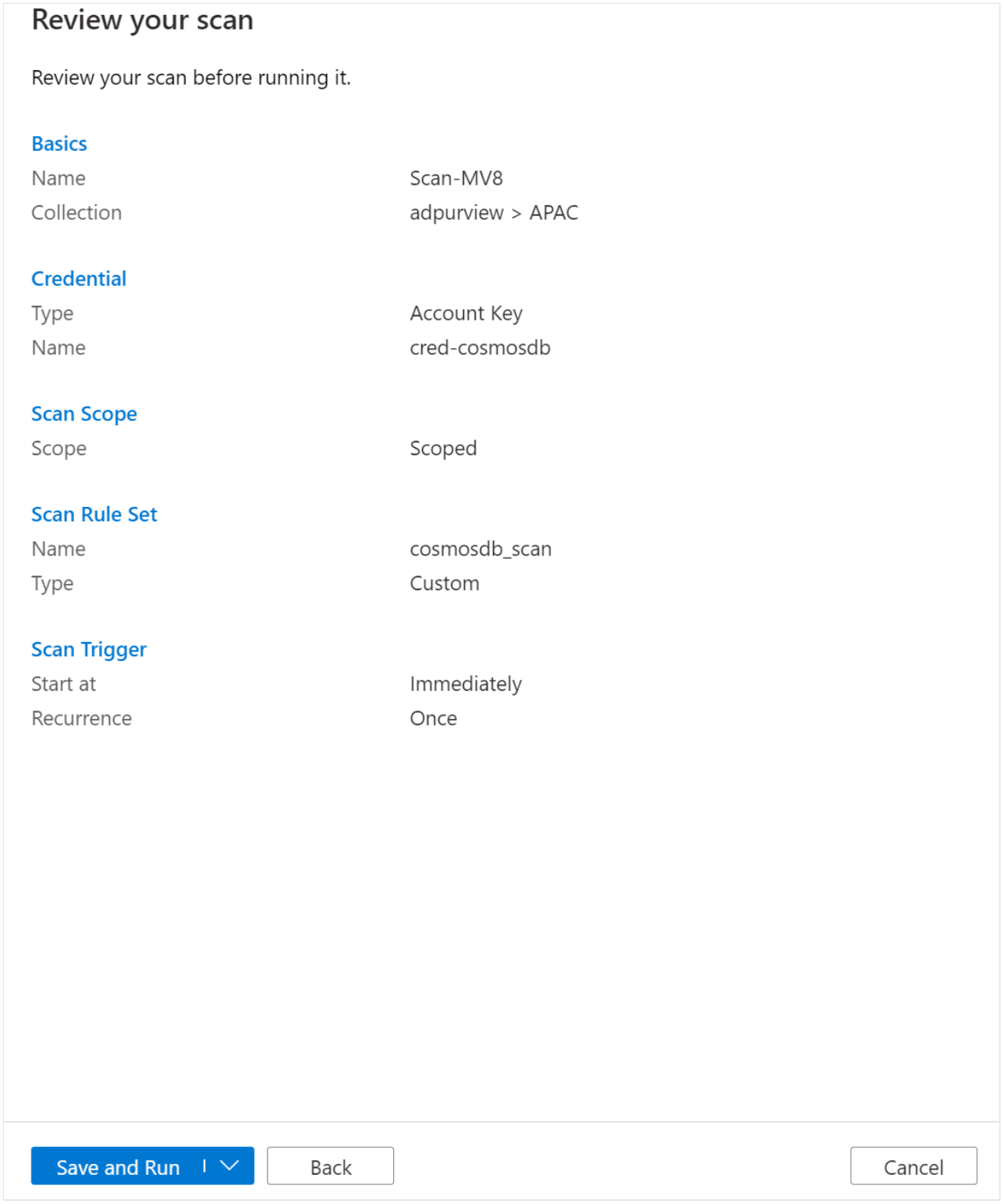
Create | Back | 198 classification rules selected | Cancel



4. Choose your scan trigger. You can set up a schedule or run the scan once.

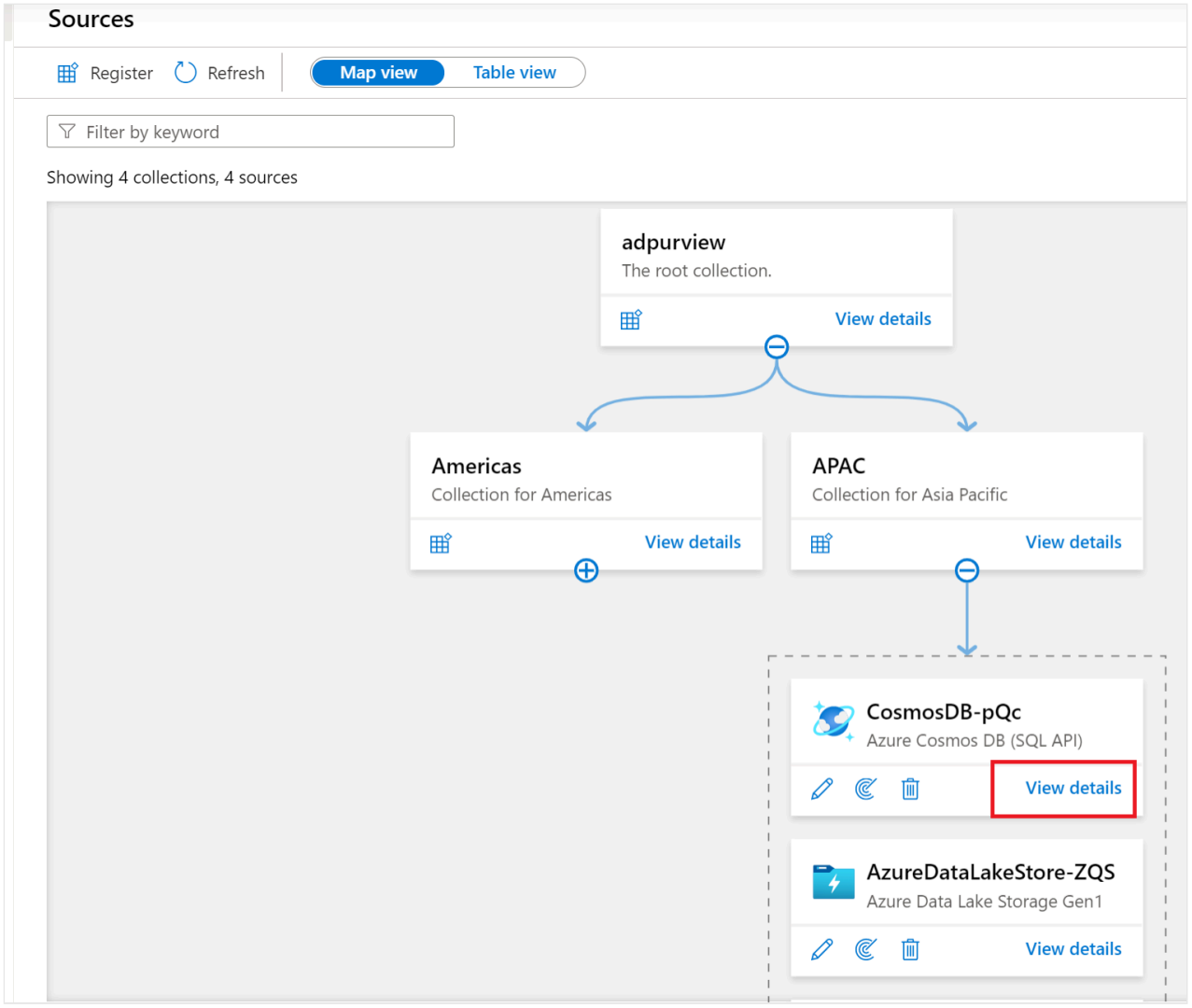


5. Review your scan and select **Save and run**.

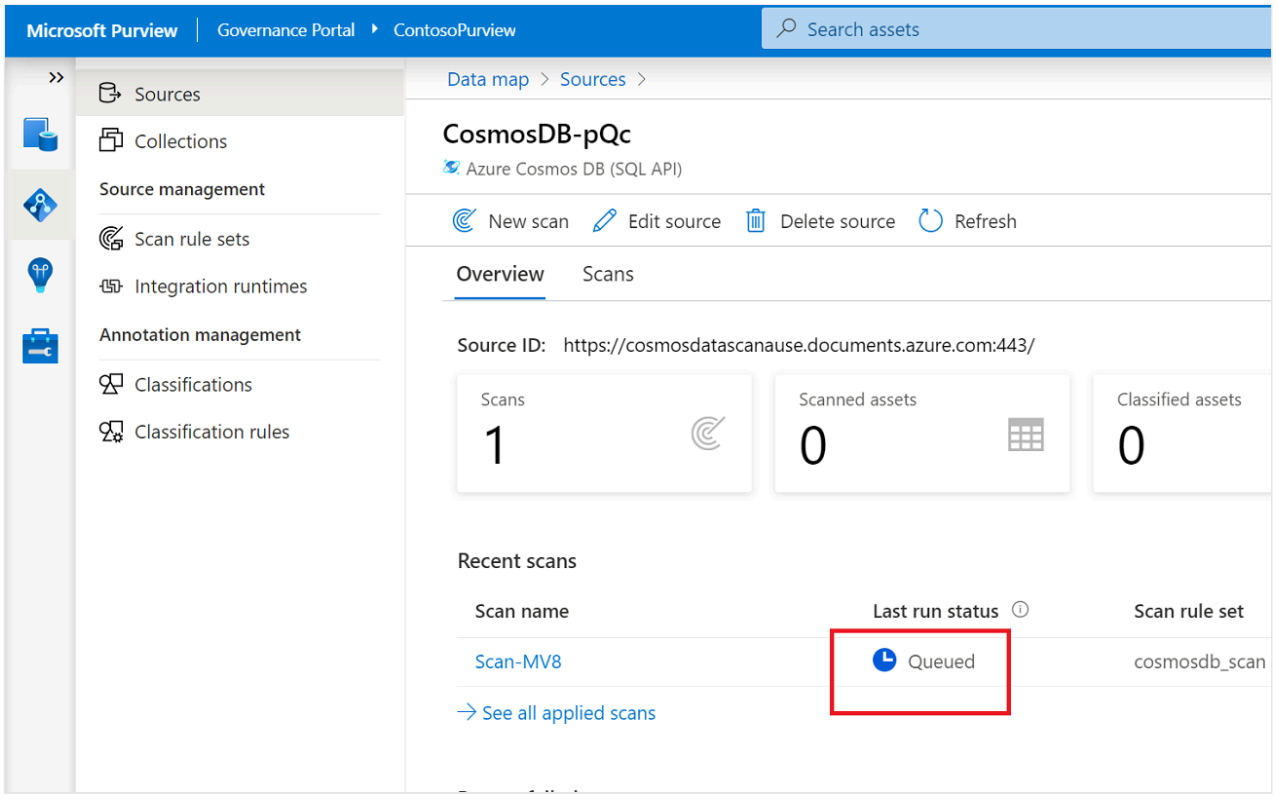


Viewing Scan

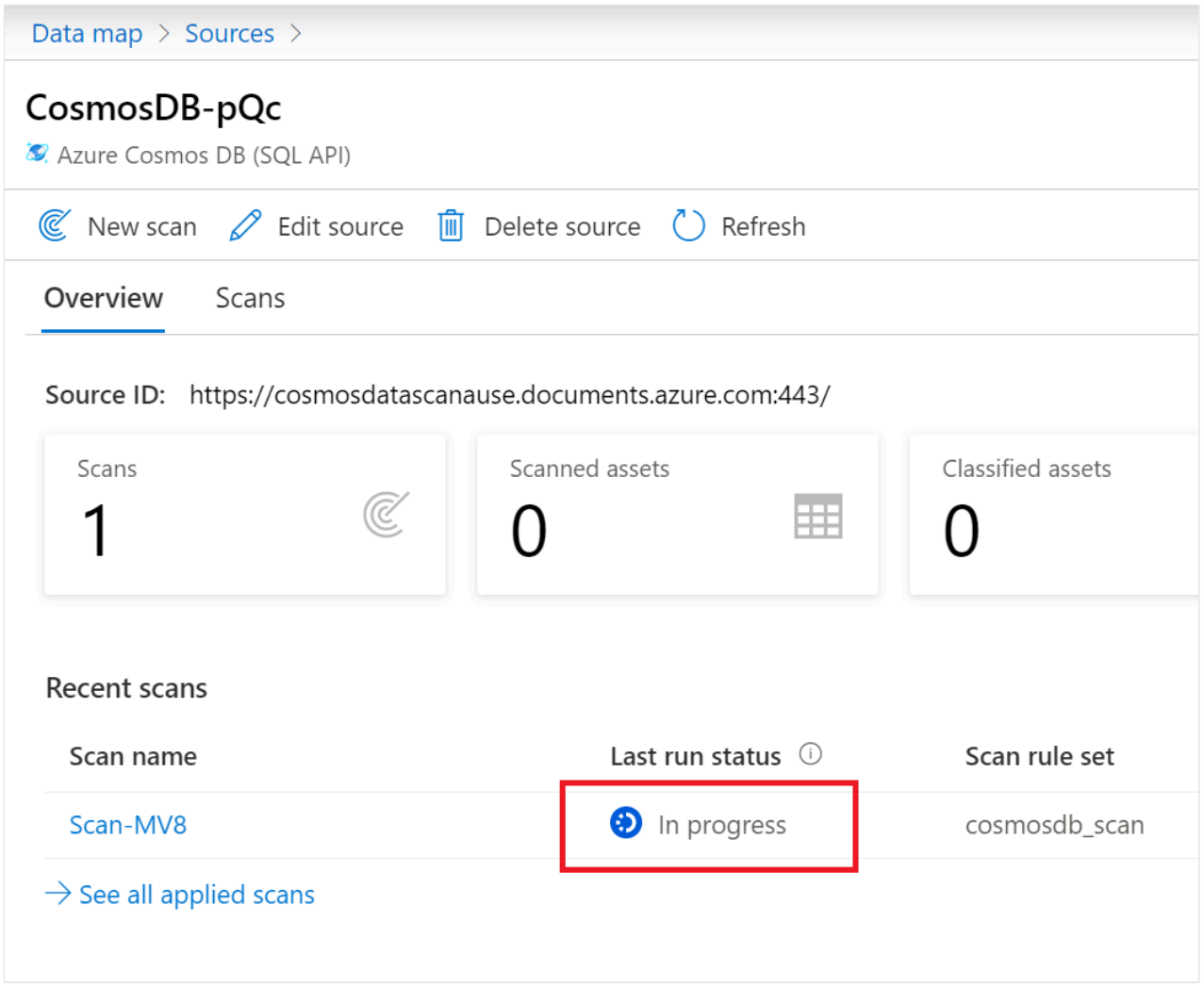
1. Navigate to the *data source* in the *Collection* and select **View Details** to check the status of the scan

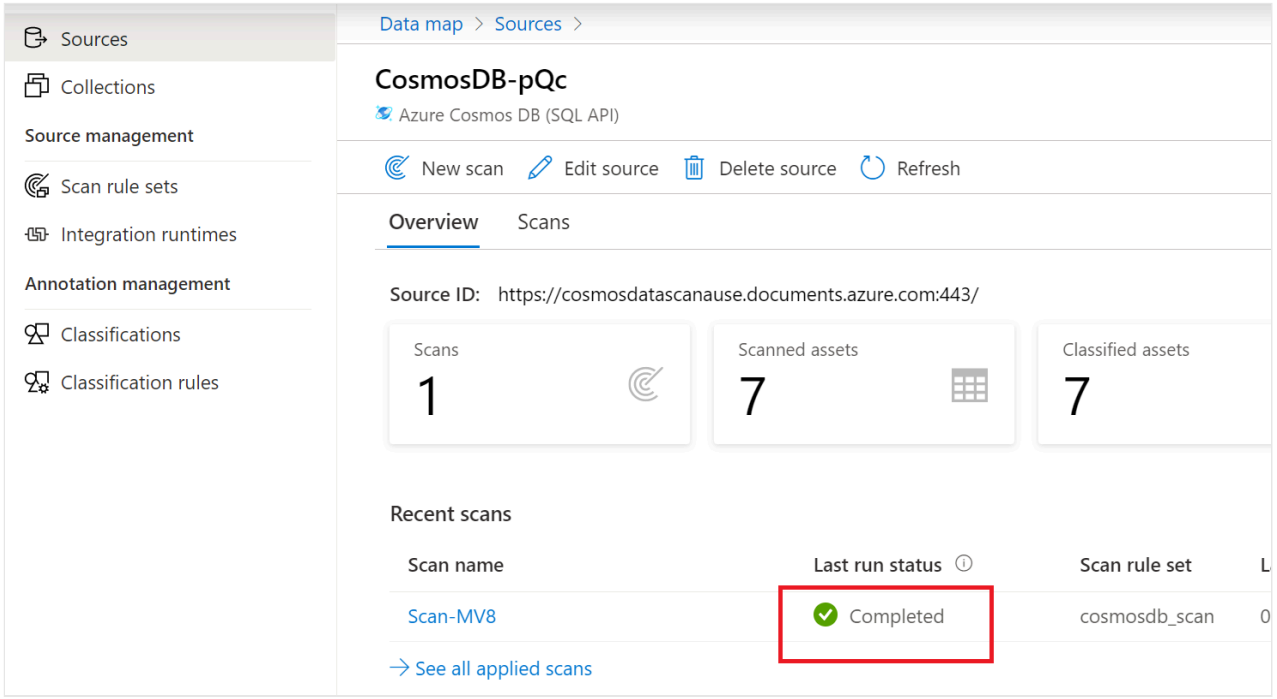


2. The scan details indicate the progress of the scan in the **Last run status** and the number of assets *scanned* and *classified*



3. The **Last run status** will be updated to **In progress** and then **Completed** once the entire scan has run successfully

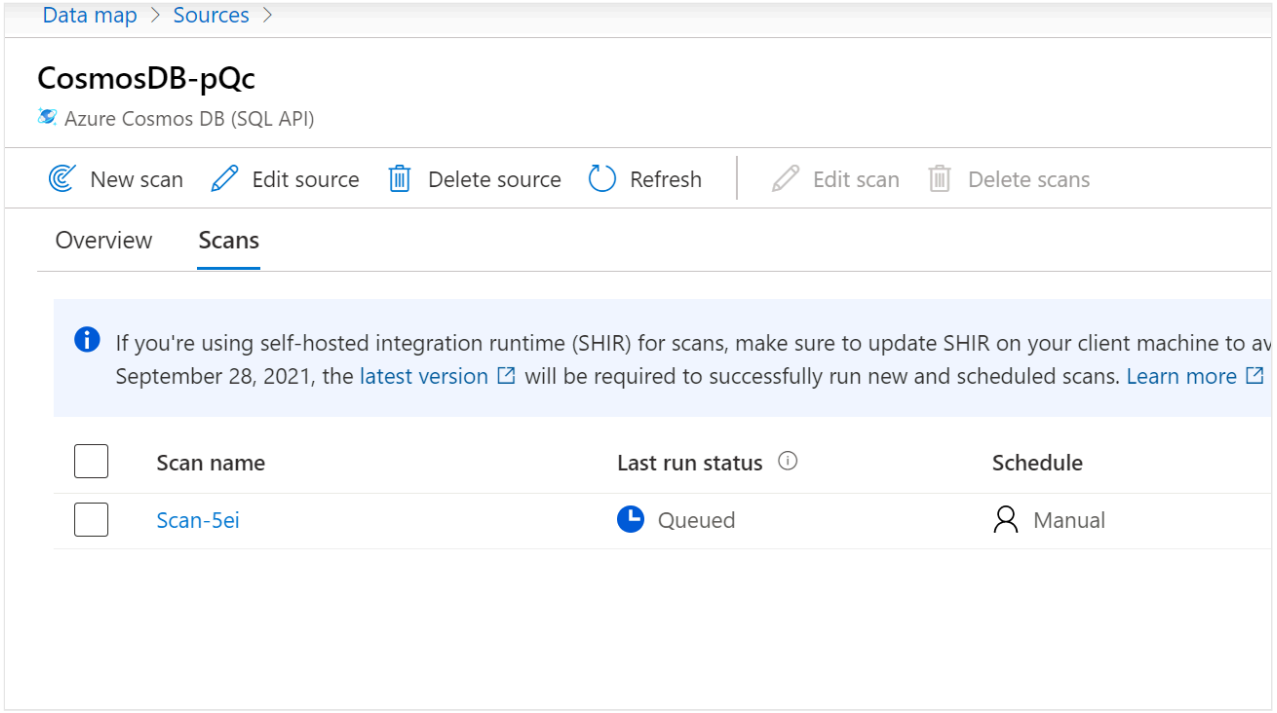




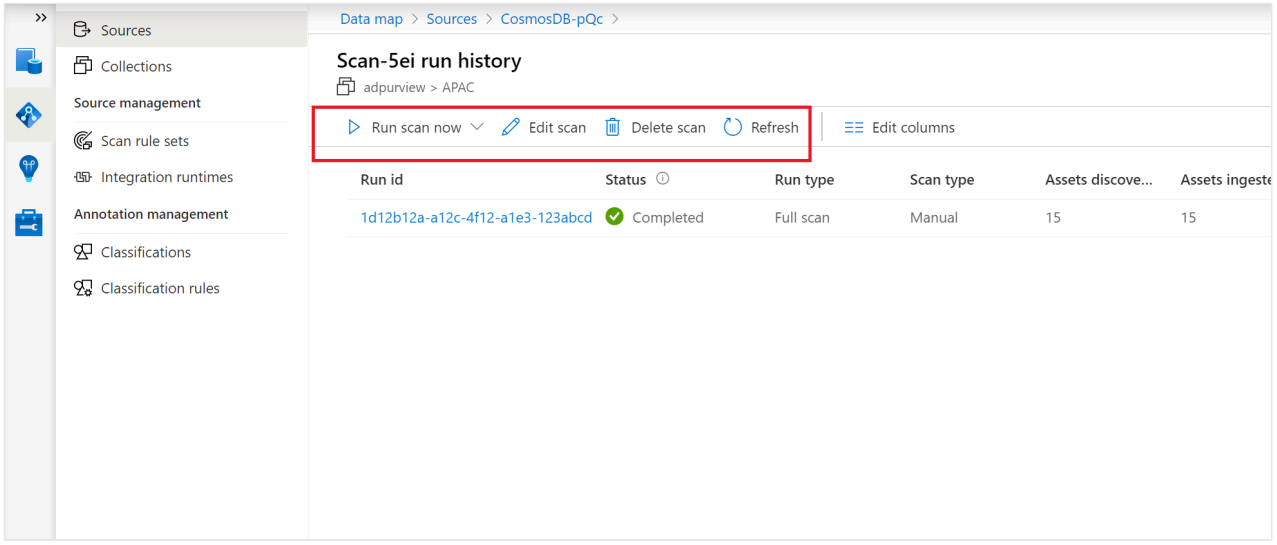
Managing Scan

Scans can be managed or run again on completion.

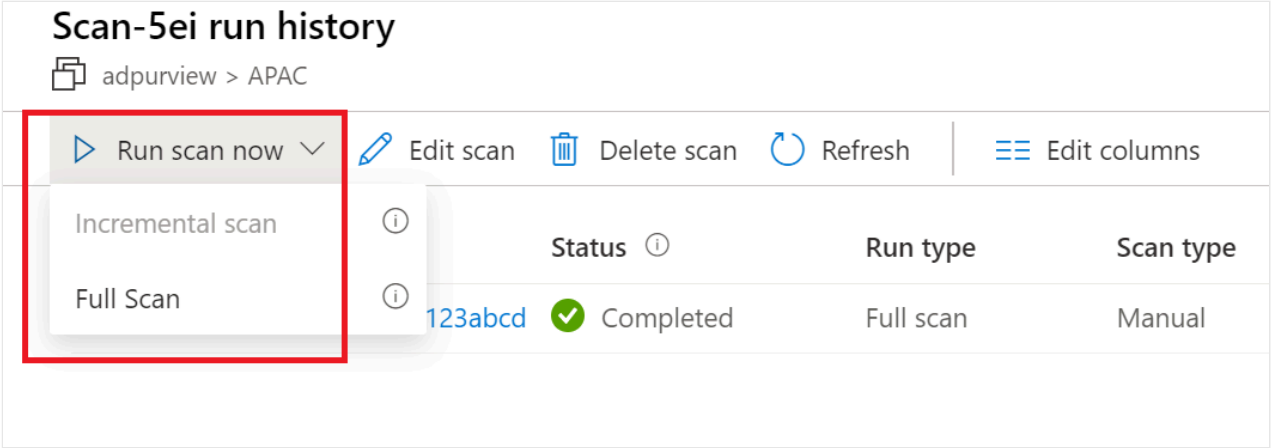
1. Select the **Scan name** to manage the scan



2. You can *run the scan* again, *edit the scan*, *delete the scan*



3. You can run a *Full Scan* again



Next steps

Now that you have registered your source, follow the below guides to learn more about Microsoft Purview and your data.

- Data Estate Insights in Microsoft Purview
- Lineage in Microsoft Purview
- Search Unified Catalog

Connect to and manage Azure Data Explorer in Microsoft Purview

Article • 05/21/2025

This article outlines how to register Azure Data Explorer, and how to authenticate and interact with Azure Data Explorer in Microsoft Purview. For more information about Microsoft Purview, read the introductory article.

Supported capabilities

Scanning capabilities

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan
Yes	Yes	Yes	Yes

Other capabilities

For **classifications**, **sensitivity labels**, **policies**, **data lineage**, and **live view**, see the list of supported capabilities.

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- An active Microsoft Purview account.
- You will need to be a Data Source Administrator and Data Reader to register a source and manage it in the Microsoft Purview governance portal. See our Microsoft Purview Permissions page for details.

Register

This section describes how to register Azure Data Explorer in Microsoft Purview using the Microsoft Purview governance portal.

Authentication for registration

There are several methods available for authentication for Azure Data Explorer:

- Service Principal
- System-assigned managed identity (SAMI)
- User-assigned managed identity (UAMI)

Service Principal to register

To use service principal authentication for scans, you can use an existing one or create a new one.

> **Note**
>
> If you have to create a new Service Principal, please follow these steps:
>
> 1. Navigate to the Azure portal.
> 2. Select **Microsoft Entra ID** from the left-hand side menu.
> 3. Select **App registrations**.
> 4. Select **+ New application registration**.
> 5. Enter a name for the **application** (the service principal name).
> 6. Select **Accounts in this organizational directory only**.
> 7. For Redirect URI select **Web** and enter any URL you want; it doesn't have to be real or work.
> 8. Then select **Register**.

It is required to get the Service Principal's application ID and secret:

1. Navigate to your Service Principal in the Azure portal
2. Copy the values the **Application (client) ID** from **Overview** and **Client secret** from **Certificates & secrets**.
3. Navigate to your key vault
4. Select **Settings** > **Secrets**
5. Select **+ Generate/Import** and enter the **Name** of your choice and **Value** as the **Client secret** from your Service Principal
6. Select **Create** to complete
7. If your key vault is not connected to Microsoft Purview yet, you will need to create a new key vault connection
8. Finally, create a new credential using the Service Principal to set up your scan

Granting the Service Principal access to your Azure data explorer instance

1. Navigate to the Azure portal. Then navigate to your Azure data explorer instance.
2. Add the service principal to the **AllDatabasesViewer** role in the **Permissions** tab.

System or user assigned managed identity to register

- **System-assigned managed identity** - As soon as the Microsoft Purview Account is created, a system-assigned managed identity (SAMI) is created automatically in Microsoft Entra tenant. It has the same name as your Microsoft Purview account.
- **User-assigned managed identity** (preview) - Similar to a system-managed identity, a user-assigned managed identity (UAMI) is a credential resource that can be used to allow Microsoft Purview to authenticate against Microsoft Entra ID. For more information, you can see our User-assigned managed identity guide.

To register using either of these managed identities, follow these steps:

1. If you would like to use a user-assigned managed identity and have not created one, follow the steps to create the identity in the User-assigned managed identity guide.
2. Navigate to the Azure portal. Then navigate to your Azure data explorer instance.
3. Select the **Permissions** tab on the left pane.
4. Add the SAMI or UAMI to the **AllDatabasesViewer** role in the **Permissions** tab.

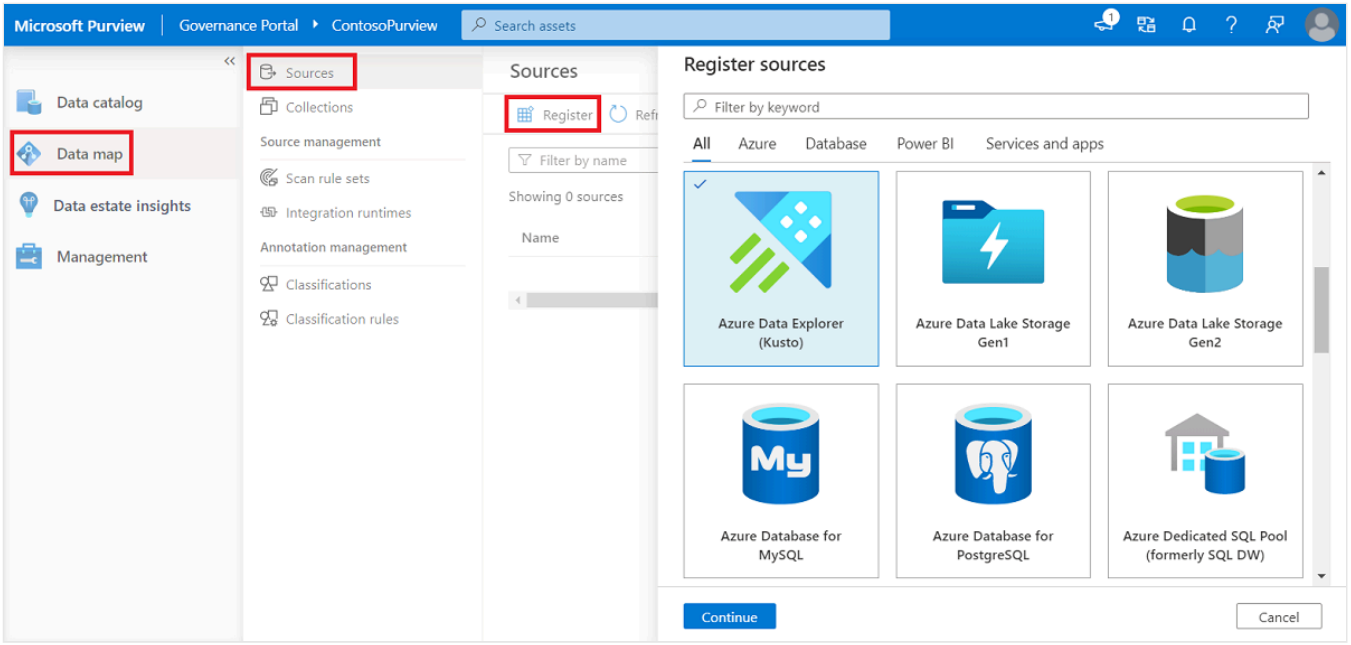
Steps to register

To register a new Azure Data Explorer (Kusto) account in Microsoft Purview Unified Catalog, follow these steps:

1. Open the Microsoft Purview governance portal by:
    - Browsing directly to https://web.purview.azure.com and selecting your Microsoft Purview account.
    - Opening the Azure portal, searching for and selecting the Microsoft Purview account. Selecting the **the Microsoft Purview governance portal** button.
2. Select **Data Map** on the left navigation.
3. Select **Register**

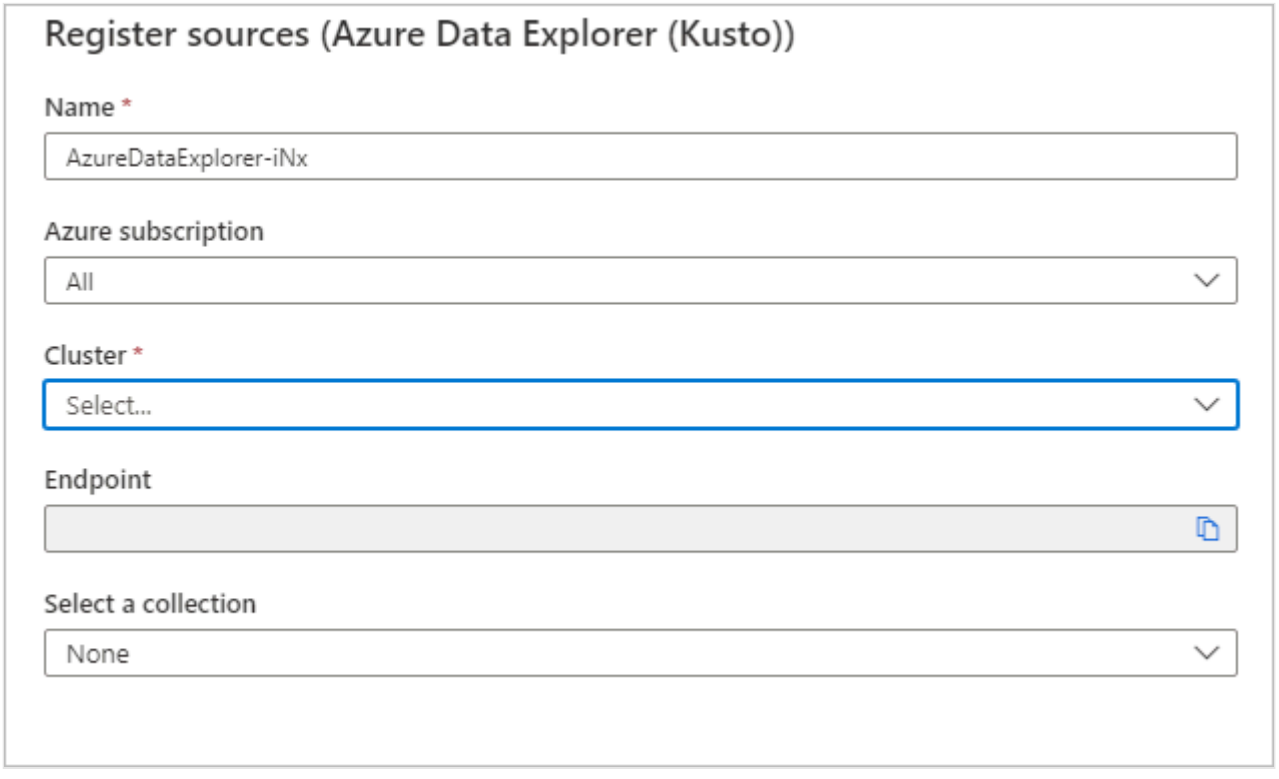
4. On **Register sources**, select **Azure Data Explorer**

5. Select **Continue**



On the **Register sources (Azure Data Explorer (Kusto))** screen, do the following:

1. Enter a **Name** that the data source will be listed with in the Catalog.
2. Choose your Azure subscription to filter down Azure Data Explorer.
3. Select an appropriate cluster.
4. Select a collection from the list.
5. Select **Register** to register the data source.



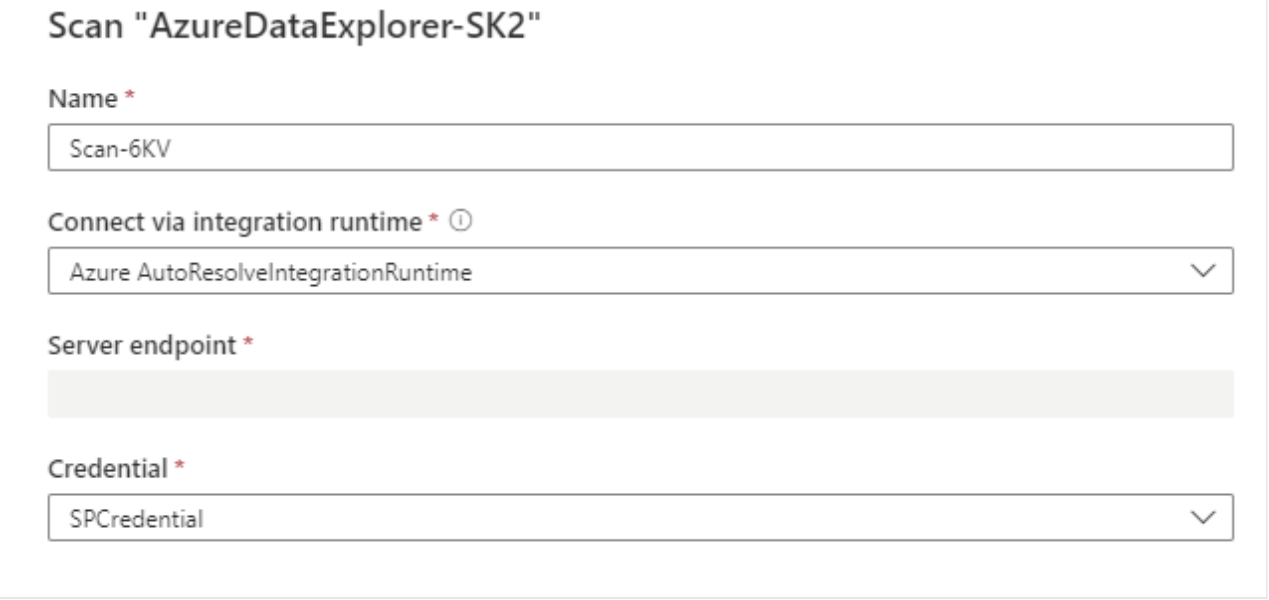
Scan

Follow the steps below to scan Azure Data Explorer to automatically identify assets and classify your data. For more information about scanning in general, see our introduction to scans and ingestion.

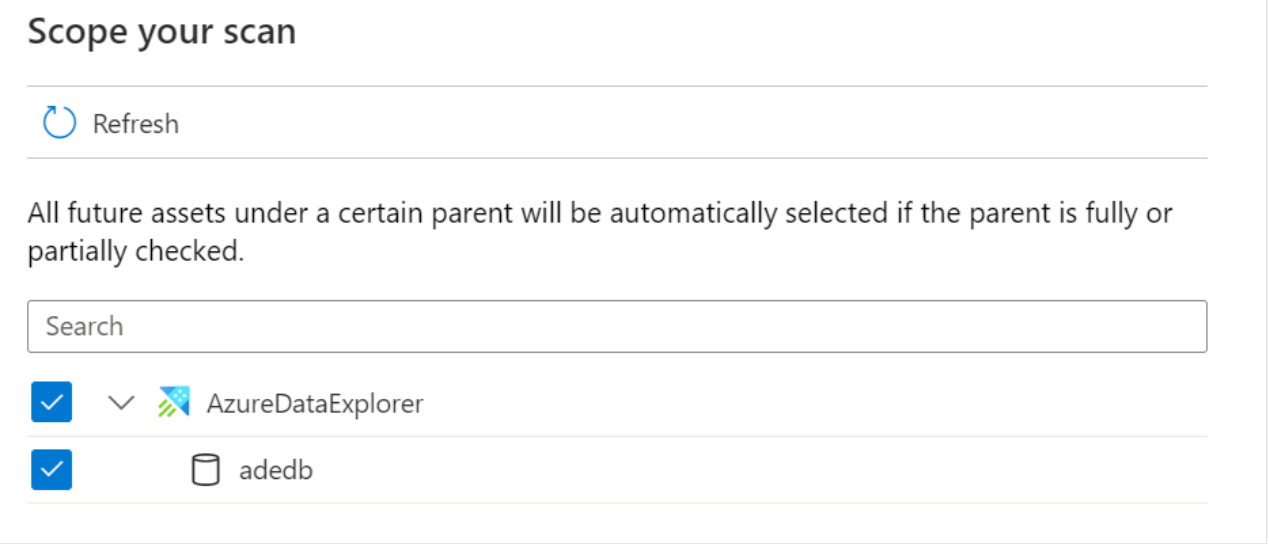
Create and run scan

To create and run a new scan, follow these steps:

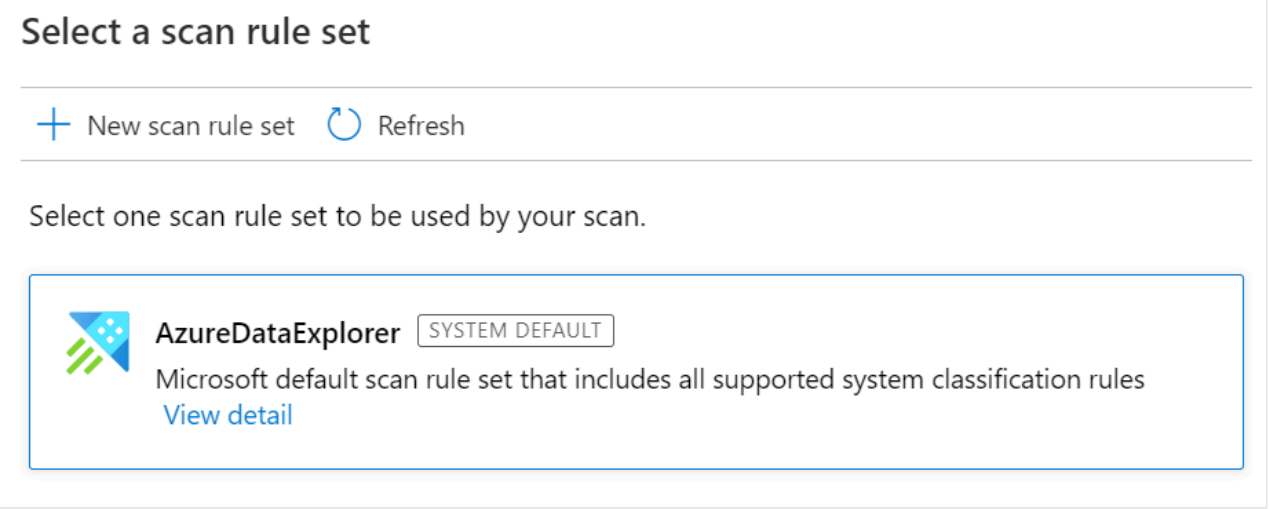
1. Select the **Data Map** tab on the left pane in the Microsoft Purview governance portal.
2. Select the Azure Data Explorer source that you registered.
3. Select **New scan**
4. Choose either the Azure integration runtime if your source is publicly accessible, a managed virtual network integration runtime if using a managed virtual network, or a self-hosted integration runtime if your source is in a private virtual network. For more information about which integration runtime to use, see the choose the right integration runtime configuration article.
5. Select the credential to connect to your data source.



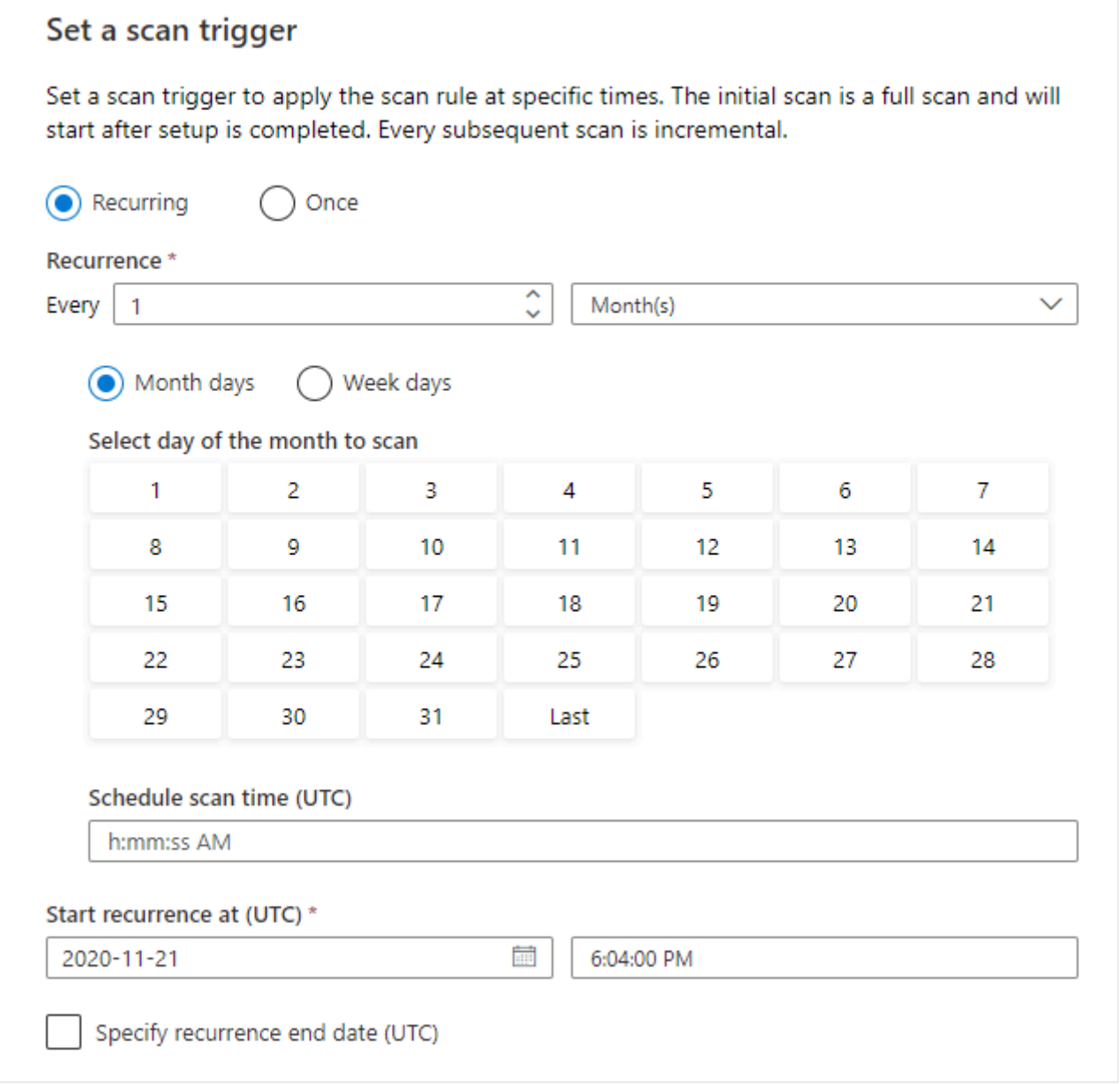
6. You can scope your scan to specific databases by choosing the appropriate items in the list.



7. Then select a scan rule set. You can choose between the system default, existing custom rule sets, or create a new rule set inline.



8. Choose your scan trigger. You can set up a schedule or run the scan once.



9. Review your scan and select **Save and run**.

View your scans and scan runs

To view existing scans:

1. Go to the Microsoft Purview portal. On the left pane, select **Data map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that has results you want to view. The pane shows you all the previous scan runs, along with the status and metrics for each scan run.
4. Select the run ID to check the scan run details.

Manage your scans

To edit, cancel, or delete a scan:

1. Go to the Microsoft Purview portal. On the left pane, select **Data Map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that you want to manage. You can then:
    - Edit the scan by selecting **Edit scan**.
    - Cancel an in-progress scan by selecting **Cancel scan run**.
    - Delete your scan by selecting **Delete scan**.

> ⓘ **Note**
>
> - Deleting your scan does not delete catalog assets created from previous scans.

Next steps

Now that you have registered your source, follow the below guides to learn more about Microsoft Purview and your data.

- Data Estate Insights in Microsoft Purview
- Lineage in Microsoft Purview
- Search Unified Catalog

Connect to Azure Data Lake Storage in Microsoft Purview

Article • 05/21/2025

This article outlines the process to register and govern an Azure Data Lake Storage (ADLS Gen2) data source in Microsoft Purview, including instructions to authenticate and interact with the ADLS Gen2 source.

Supported capabilities

Scanning capabilities

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan
Yes	Yes	Yes	Yes

When scanning Azure Data Lake Storage Gen2 source, Microsoft Purview supports extracting technical metadata including:

- Storage account
- Data Lake Storage Gen2 Service
- File system (container)
- Folders
- Files
- Resource sets

When setting up scan, you can choose to scan the entire ADLS Gen2 or selective folders. Learn about the supported file format here.

Other capabilities

For **classifications**, **sensitivity labels**, **policies**, **data lineage**, and **live view**, see the list of supported capabilities.

Prerequisites

- An Azure account with an active subscription. Create an account for free.

- An active Microsoft Purview account.
- You'll need to be a Data Source Administrator and Data Reader to register a source and manage it in the Microsoft Purview governance portal. See our Microsoft Purview Permissions page for details.
- You need to have at least Reader permission on the ADLS Gen 2 account to be able to register it.

Register

This section will enable you to register the ADLS Gen2 data source for scan and data share in Purview.

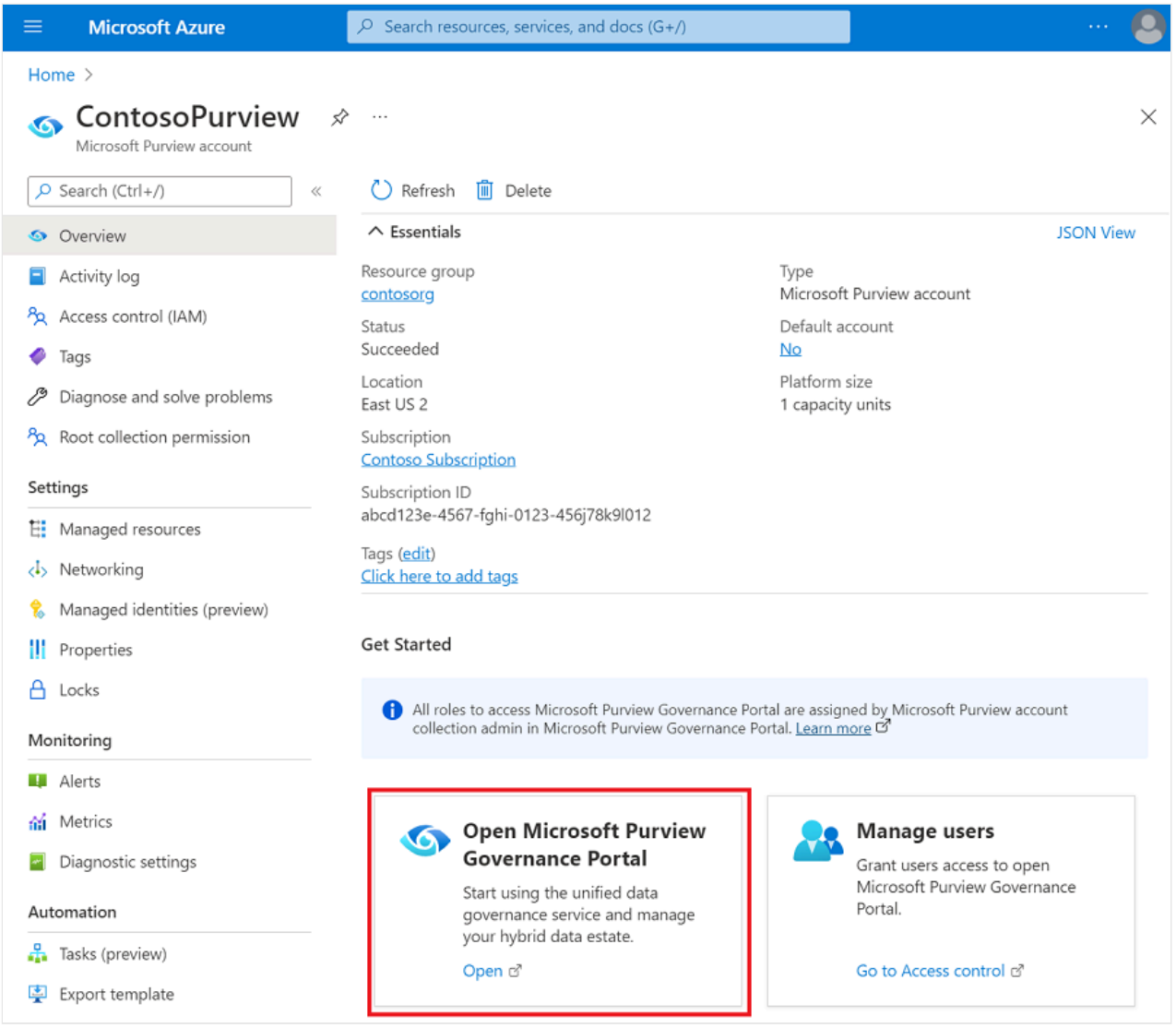
Prerequisites for register

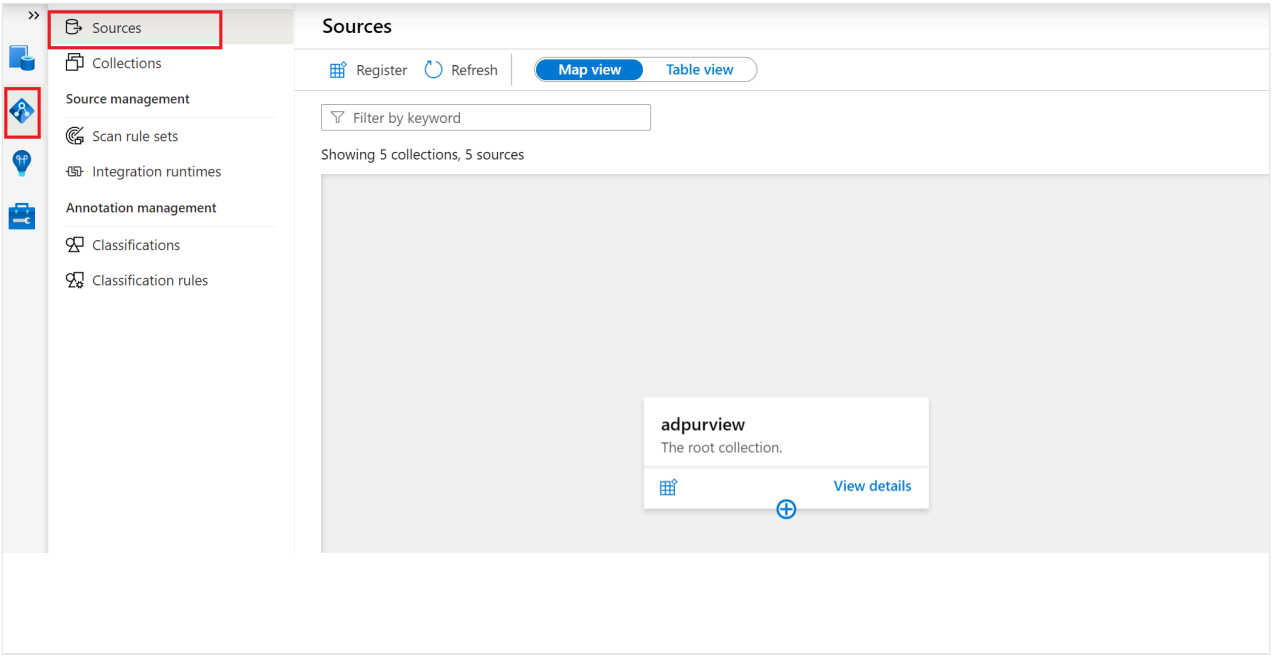
- You'll need to be a Data Source Admin and one of the other Purview roles (for example, Data Reader or Data Share Contributor) to register a source and manage it in the Microsoft Purview governance portal. See our Microsoft Purview Permissions page for details.

Steps to register

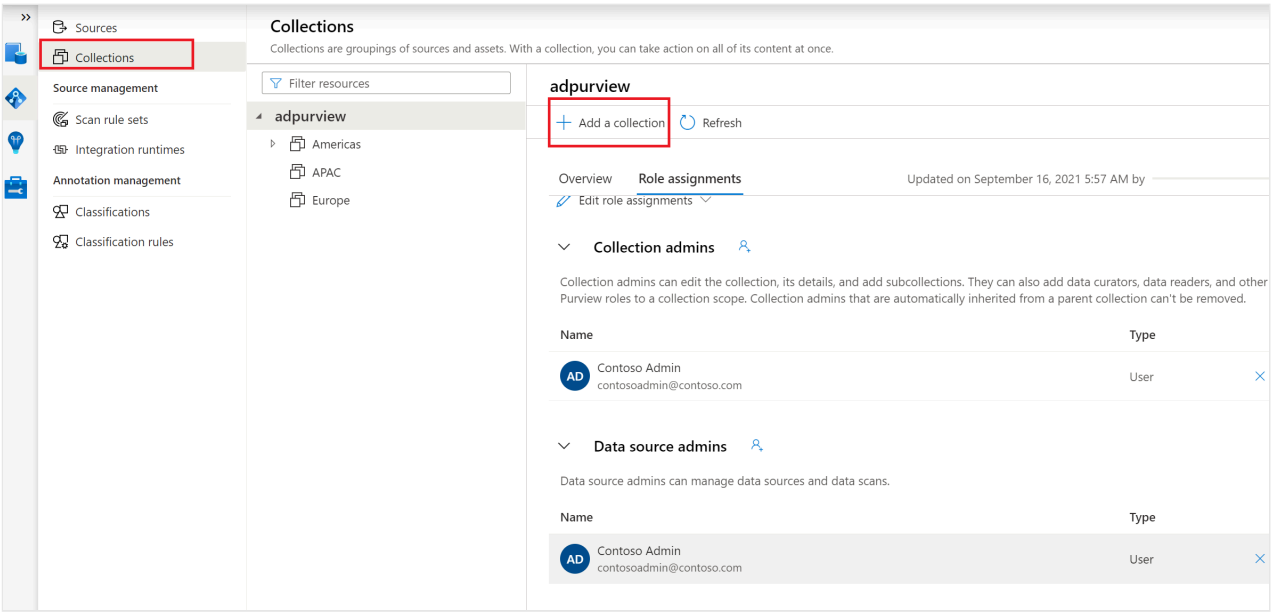
It's important to register the data source in Microsoft Purview prior to setting up a scan for the data source.

1. Go to the Microsoft Purview governance portal by:
    - Browsing directly to https://web.purview.azure.com and selecting your Microsoft Purview account.
    - Opening the Azure portal, searching for and selecting the Microsoft Purview account. Select the **the Microsoft Purview governance portal** button.
2. Navigate to the **Data Map --> Sources**

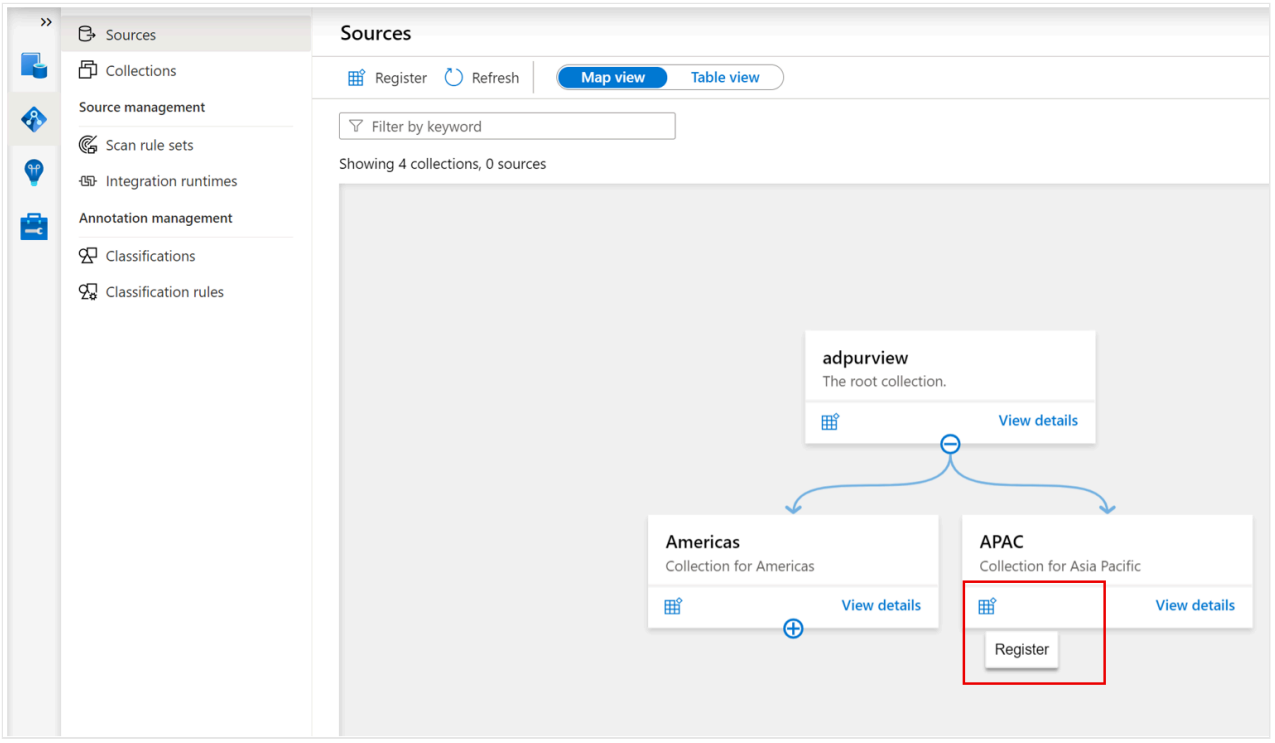




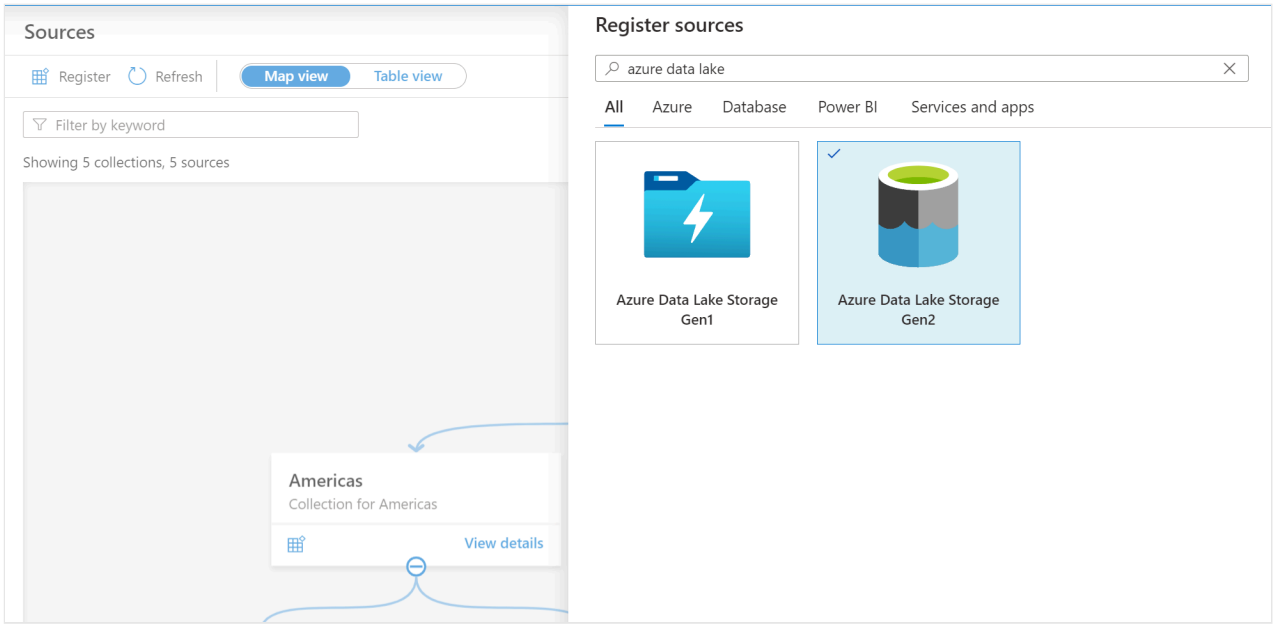
3. Create the Collection hierarchy using the **Collections** menu and assign permissions to individual subcollections, as required



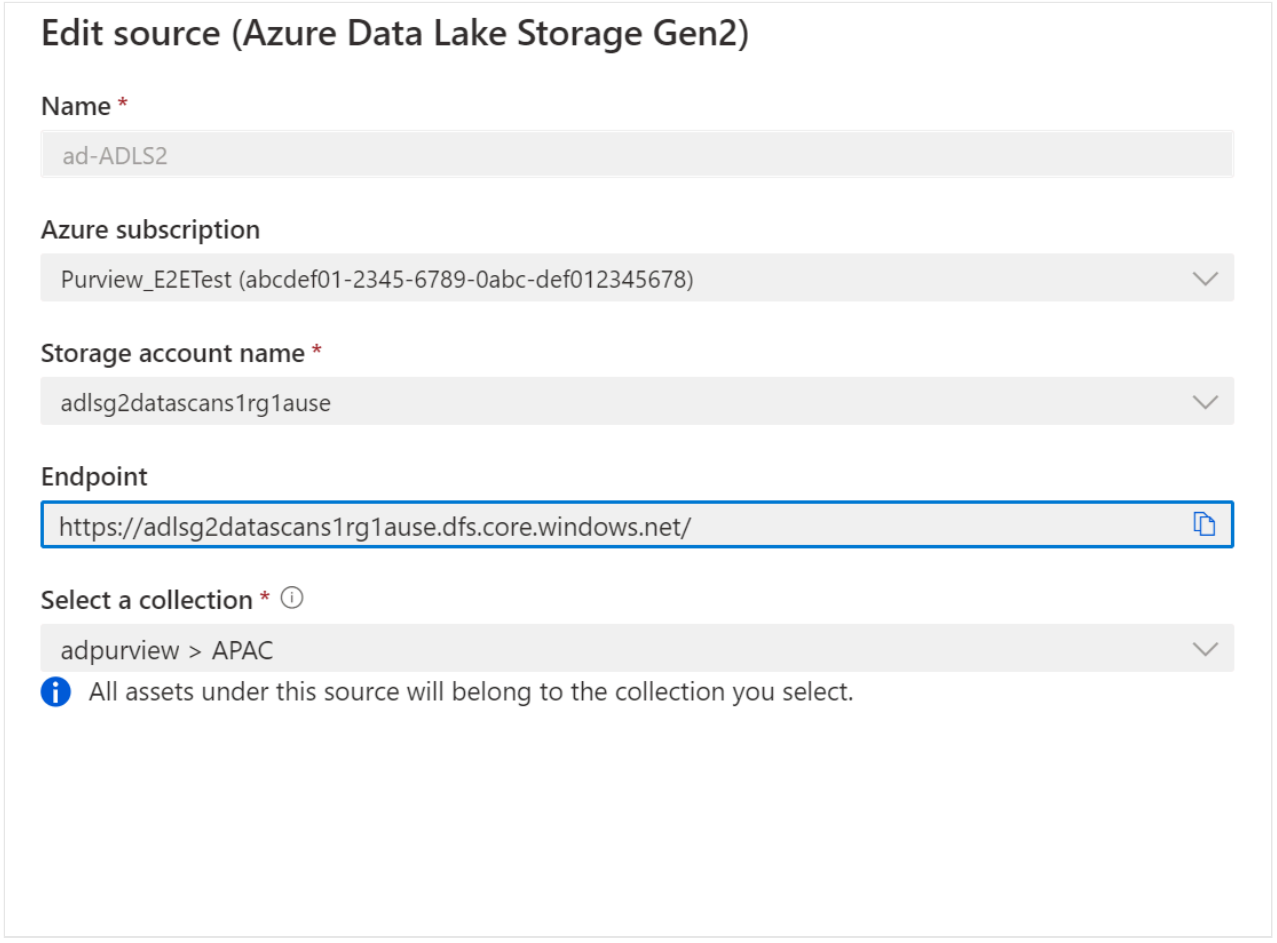
4. Navigate to the appropriate collection under the **Sources** menu and select the **Register** icon to register a new ADLS Gen2 data source



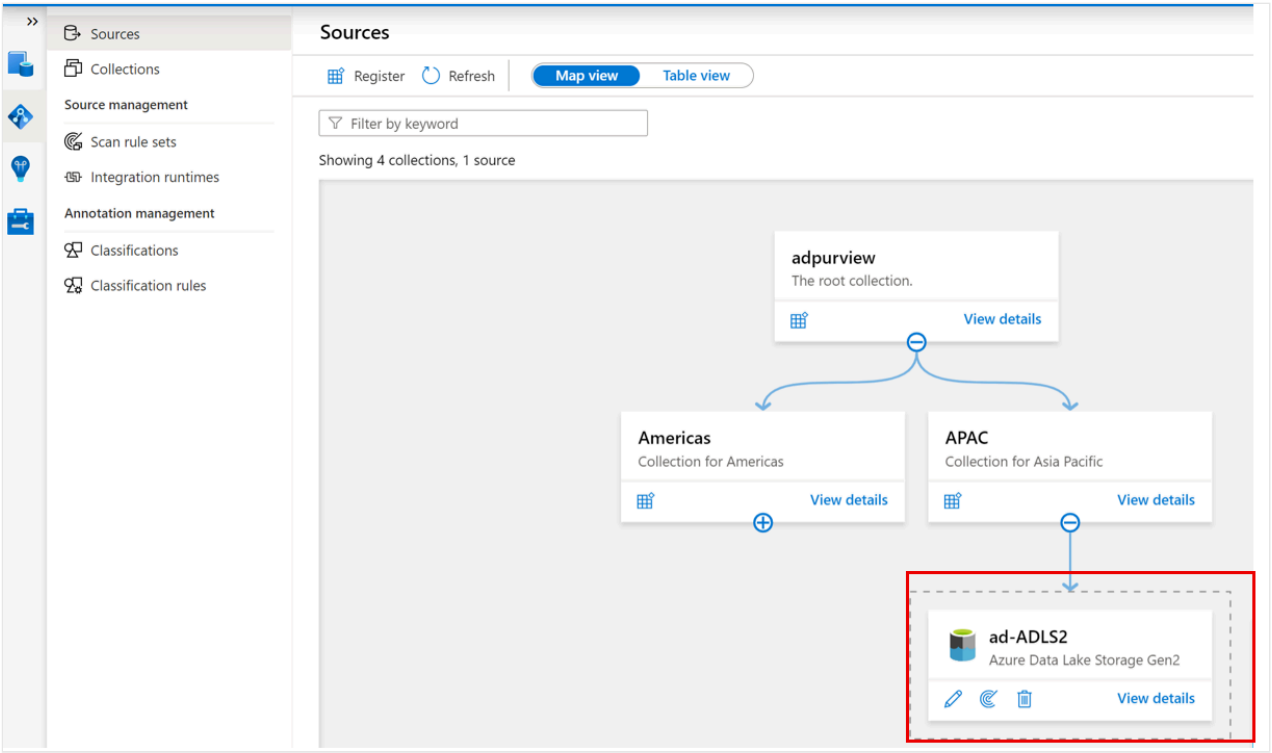
5. Select the **Azure Data Lake Storage Gen2** data source and select **Continue**



6. Provide a suitable **Name** for the data source, select the relevant **Azure subscription**, existing **Data Lake Store account name** and the **collection** and select **Apply**. Leave the **Data Policy Enforcement** toggle on the **disabled** position until you have a chance to carefully go over this document.



7. The ADLS Gen2 storage account will be shown under the selected Collection



Scan

> **Tip**
>
> To troubleshoot any issues with scanning:
>
> 1. Confirm you have properly set up authentication for scanning
> 2. Review our scan troubleshooting documentation.

Authentication for a scan

Your Azure network may allow for communications between your Azure resources, but if you've set up firewalls, private endpoints, or virtual networks within Azure, you'll need to follow one of these configurations below.

Expand table

Networking constraints	Integration runtime type	Available credential types
No private endpoints or firewalls	Azure IR	Managed identity (Recommended), service principal, or account key
Firewall enabled but no private endpoints	Azure IR	Managed identity

Networking constraints	Integration runtime type	Available credential types
Private endpoints enabled	*Self-Hosted IR	Service principal, account key

> **Important**
>
> 1. *To use a self-hosted integration runtime, you'll first need to choose the right one for your scenario, create one, and confirm your network settings for Microsoft Purview.
> 2. If you use the Self-Hosted Integration runtime to scan parquet files, you need to install the **64-bit JRE 8 (Java Runtime Environment) or OpenJDK** on your IR machine. You can check our Java Runtime Environment installation guide.

System or user assigned managed identity

Using a system or user assigned managed identity for scanning

There are two types of managed identity you can use:

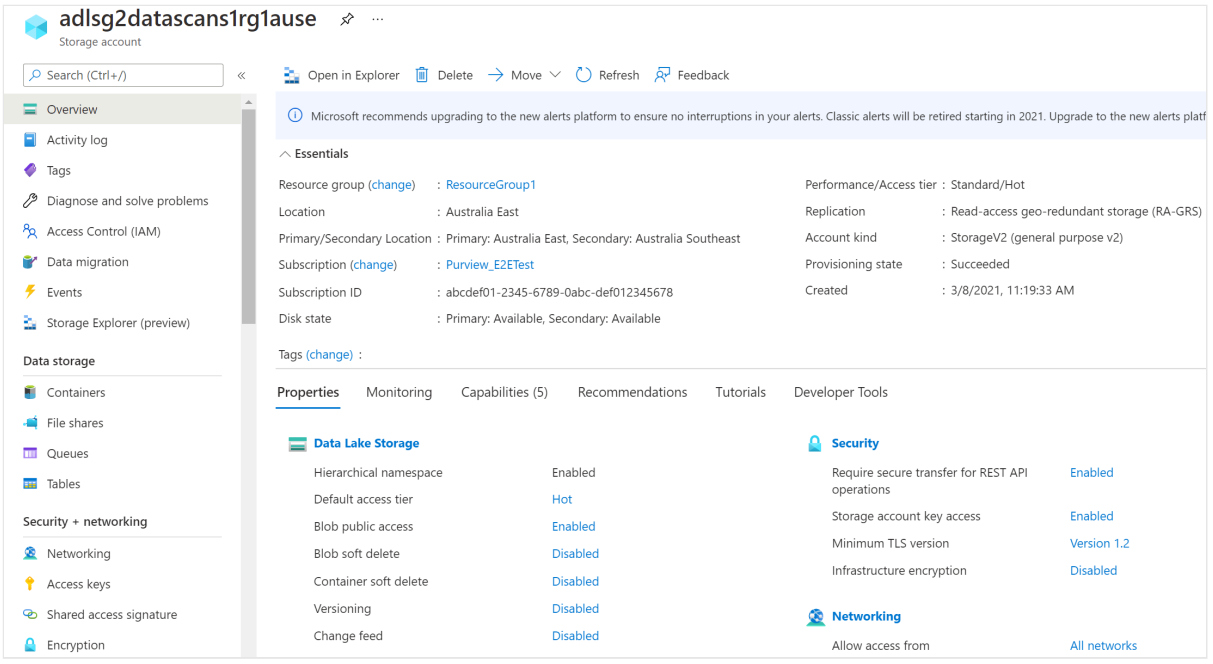
- **System-assigned managed identity (Recommended)** - As soon as the Microsoft Purview Account is created, a system-assigned managed identity (SAMI) is created automatically in Microsoft Entra tenant. Depending on the type of resource, specific RBAC role assignments are required for the Microsoft Purview system-assigned managed identity (SAMI) to perform the scans.
- **User-assigned managed identity** (preview) - Similar to a system managed identity, a user-assigned managed identity (UAMI) is a credential resource that can be used to allow Microsoft Purview to authenticate against Microsoft Entra ID. For more information, you can see our User-assigned managed identity guide.

It's important to give your Microsoft Purview account or user-assigned managed identity (UAMI) the permission to scan the ADLS Gen2 data source. You can add your Microsoft Purview account's system-assigned managed identity (which has the same name as your Microsoft Purview account) or UAMI at the Subscription, Resource Group, or Resource level, depending on what level scan permissions are needed.

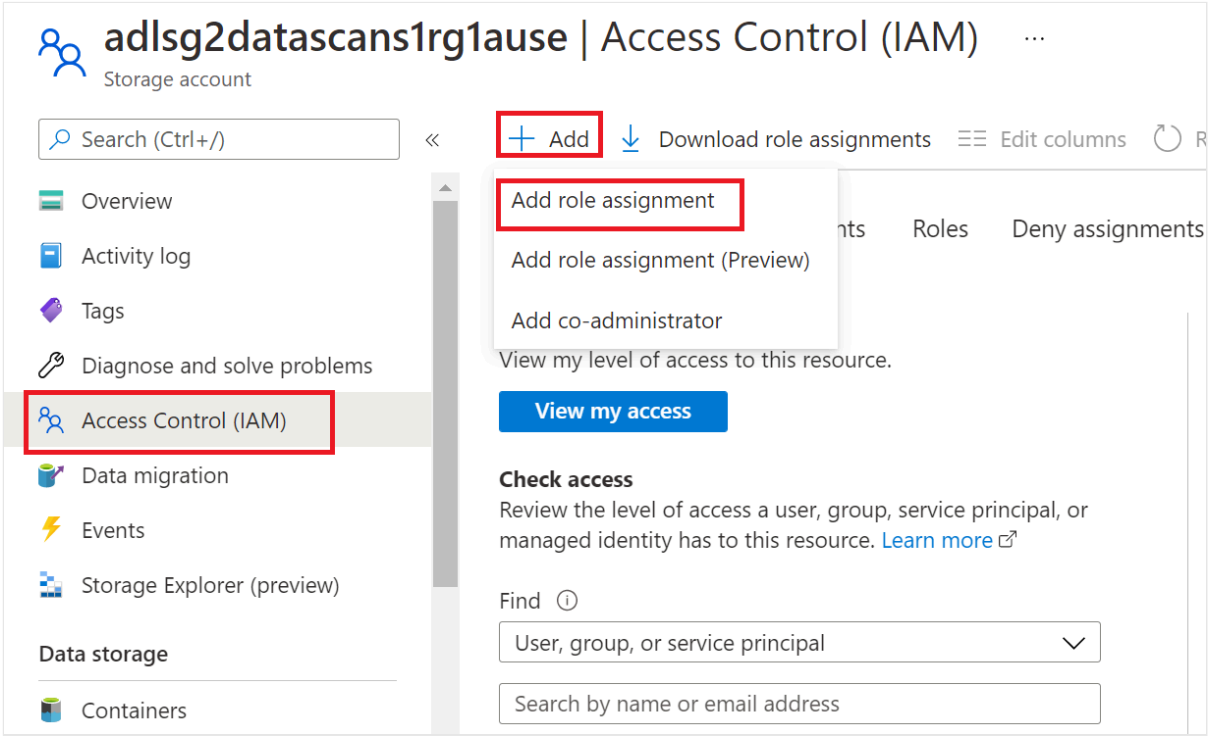
> **Note**

You need to be an owner of the subscription to be able to add a managed identity on an Azure resource.

1. From the Azure portal, find either the subscription, resource group, or resource (for example, an Azure Data Lake Storage Gen2 storage account) that you would like to allow the catalog to scan.



2. Select **Access Control (IAM)** in the left navigation and then select **+ Add** --> **Add role assignment**



3. Set the **Role** to **Storage Blob Data Reader** and enter your *Microsoft Purview account name* or *user-assigned managed identity* under the **Select** input box. Then, select

Save to give this role assignment to your Microsoft Purview account.

Add role assignment

Role ⓘ

Storage Blob Data Reader ⓘ

Assign access to ⓘ

User, group, or service principal

Select ⓘ

adpurview

No users, groups, or service principals found.

Selected members:

adpurview

Remove

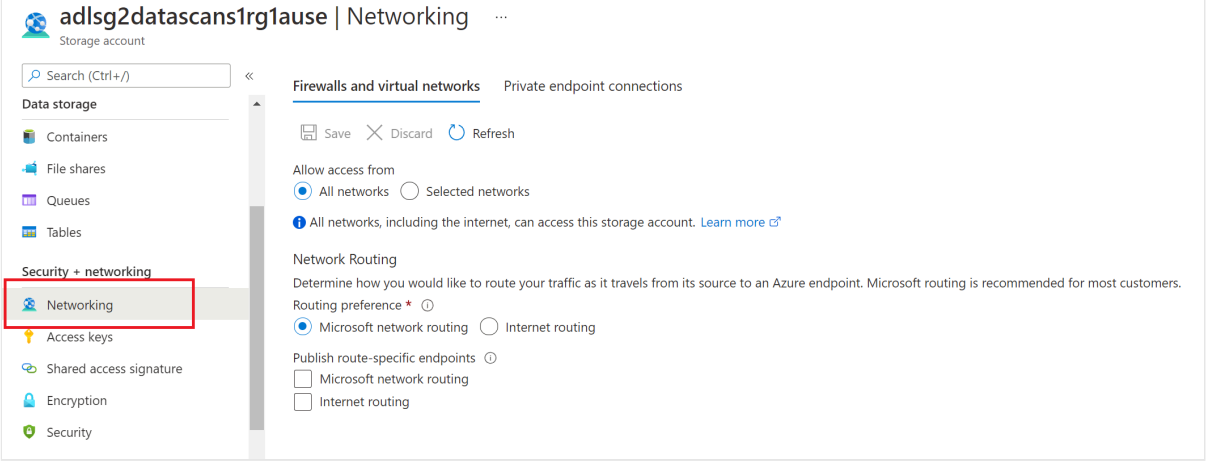
Save Discard

> **Note**
>
> For more details, please see steps in Authorize access to blobs and queues using Microsoft Entra ID

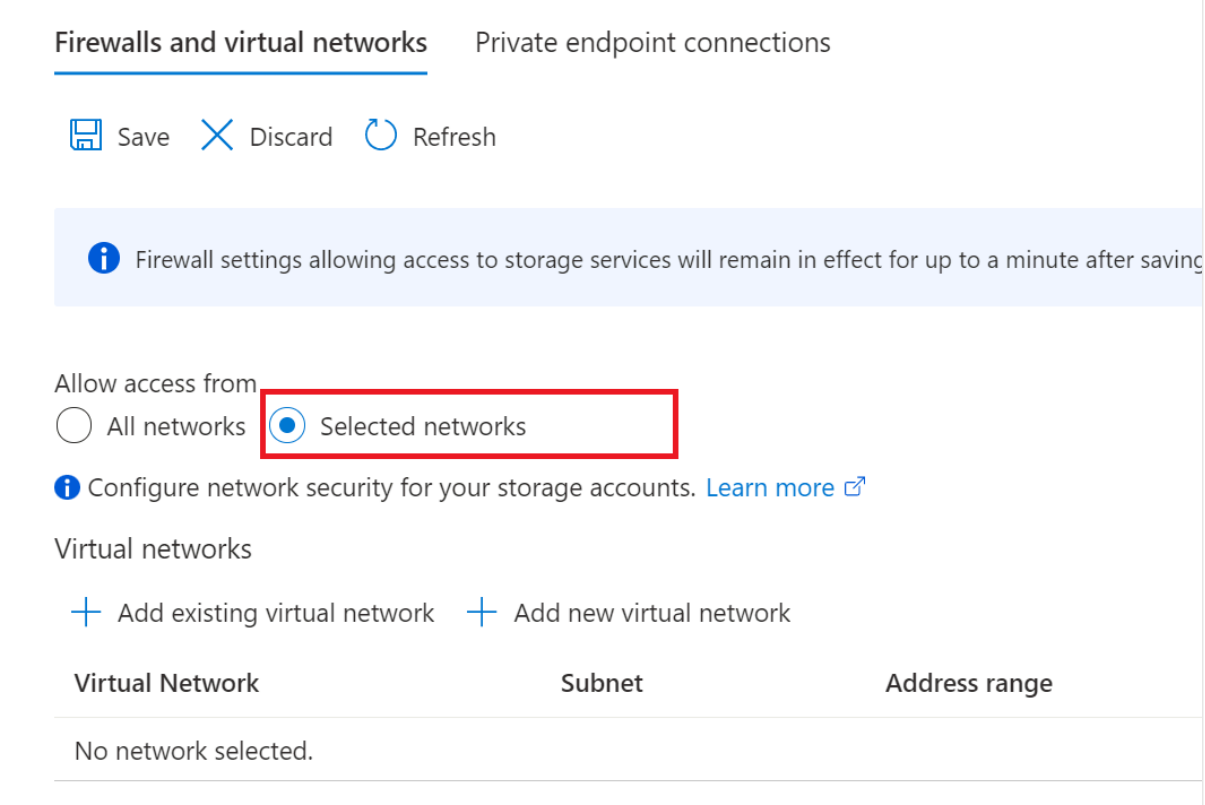
> **Note**
>
> If you have firewall enabled for the storage account, you must use **managed identity** authentication method when setting up a scan.

4. Go into your ADLS Gen2 storage account in Azure portal

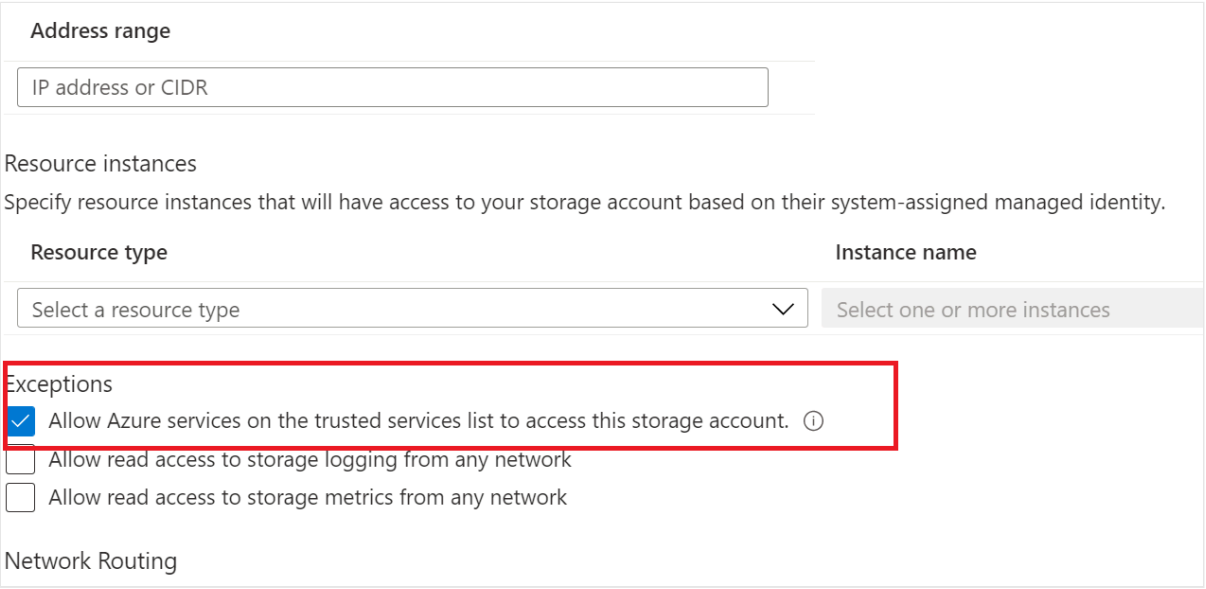
5. Navigate to **Security + networking** > **Networking**



6. Choose **Selected Networks** under **Allow access from**

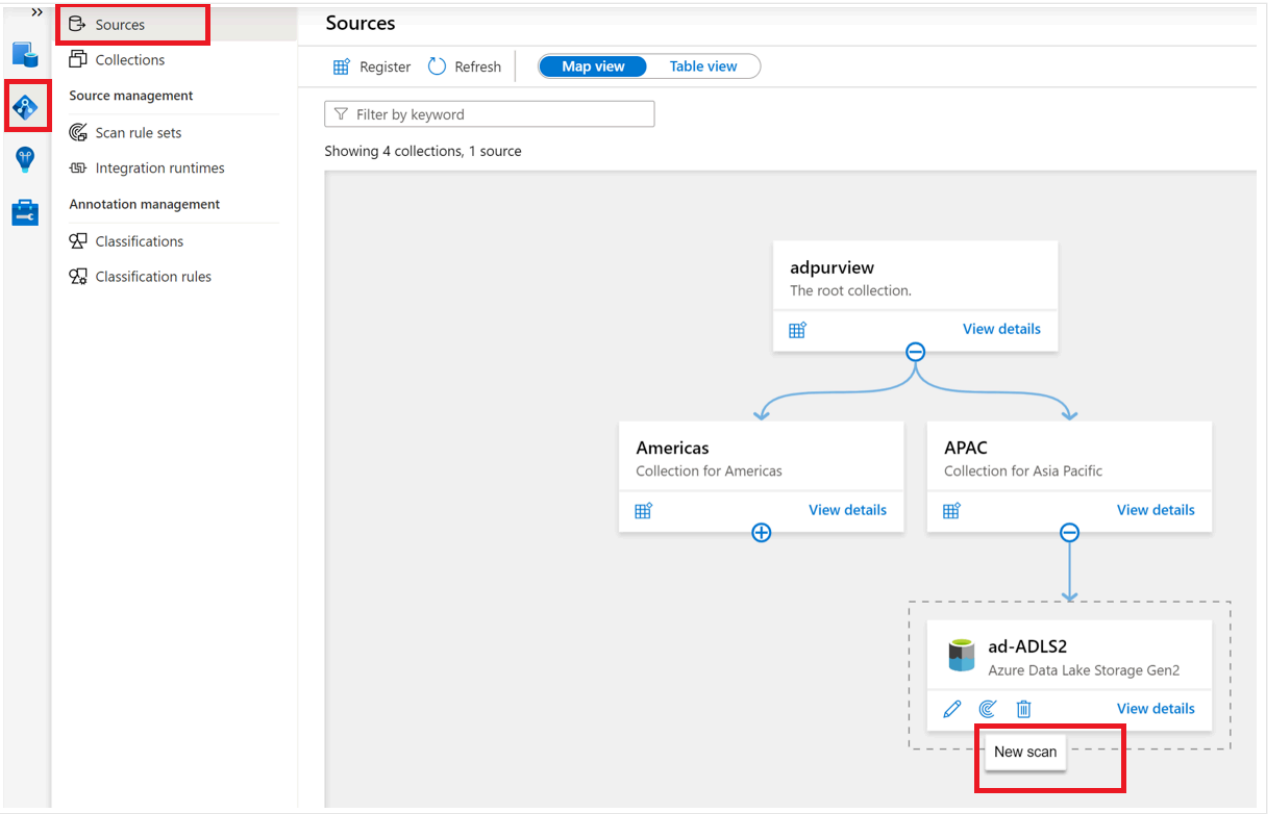


7. In the **Exceptions** section, select **Allow trusted Microsoft services to access this storage account** and hit **Save**



Create the scan

1. Open your **Microsoft Purview account** and select the **Open Microsoft Purview governance portal**
2. Navigate to the **Data map** --> **Sources** to view the collection hierarchy
3. Select the **New Scan** icon under the **ADLS Gen2 data source** registered earlier

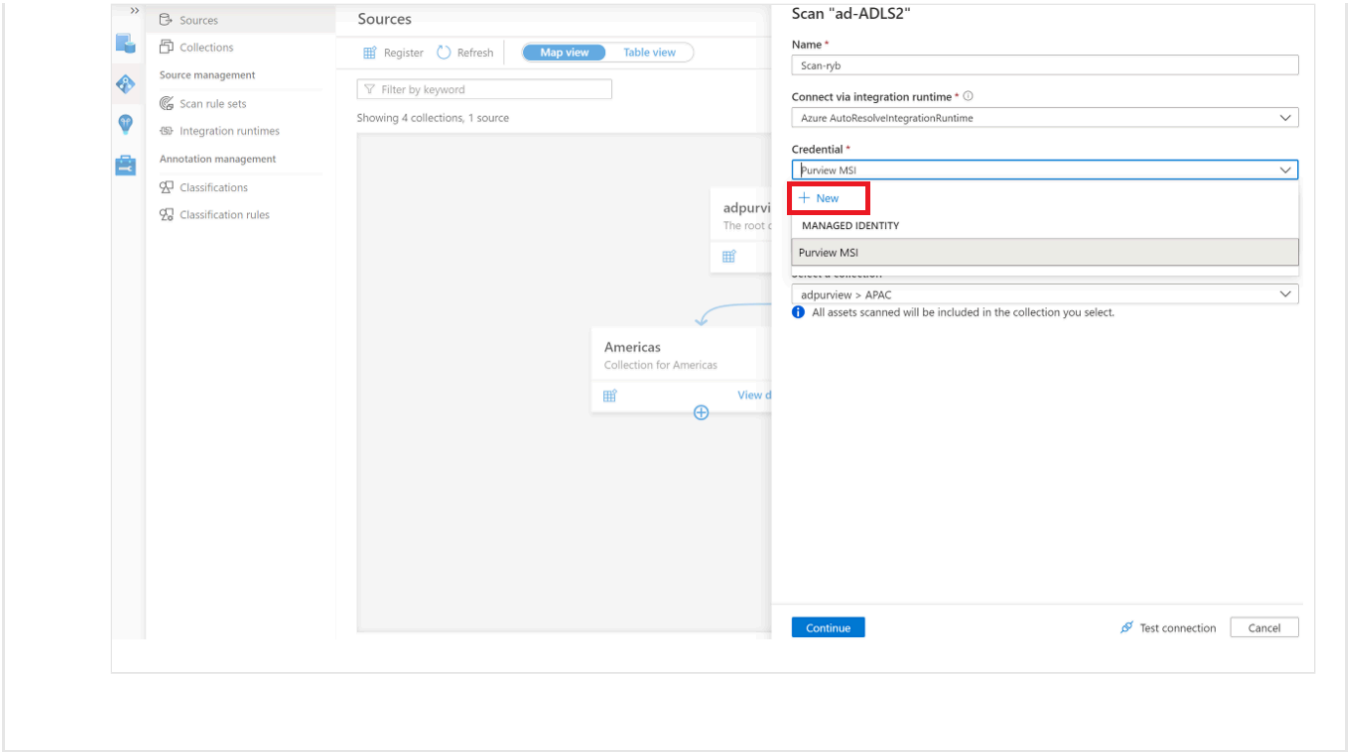


4. Choose either the Azure integration runtime if your source is publicly accessible, a managed virtual network integration runtime if using a managed virtual network, or a self-hosted integration runtime if your source is in a private virtual network. For more information about which integration runtime to use, see the choose the right integration runtime configuration article.

System or user assigned managed identity

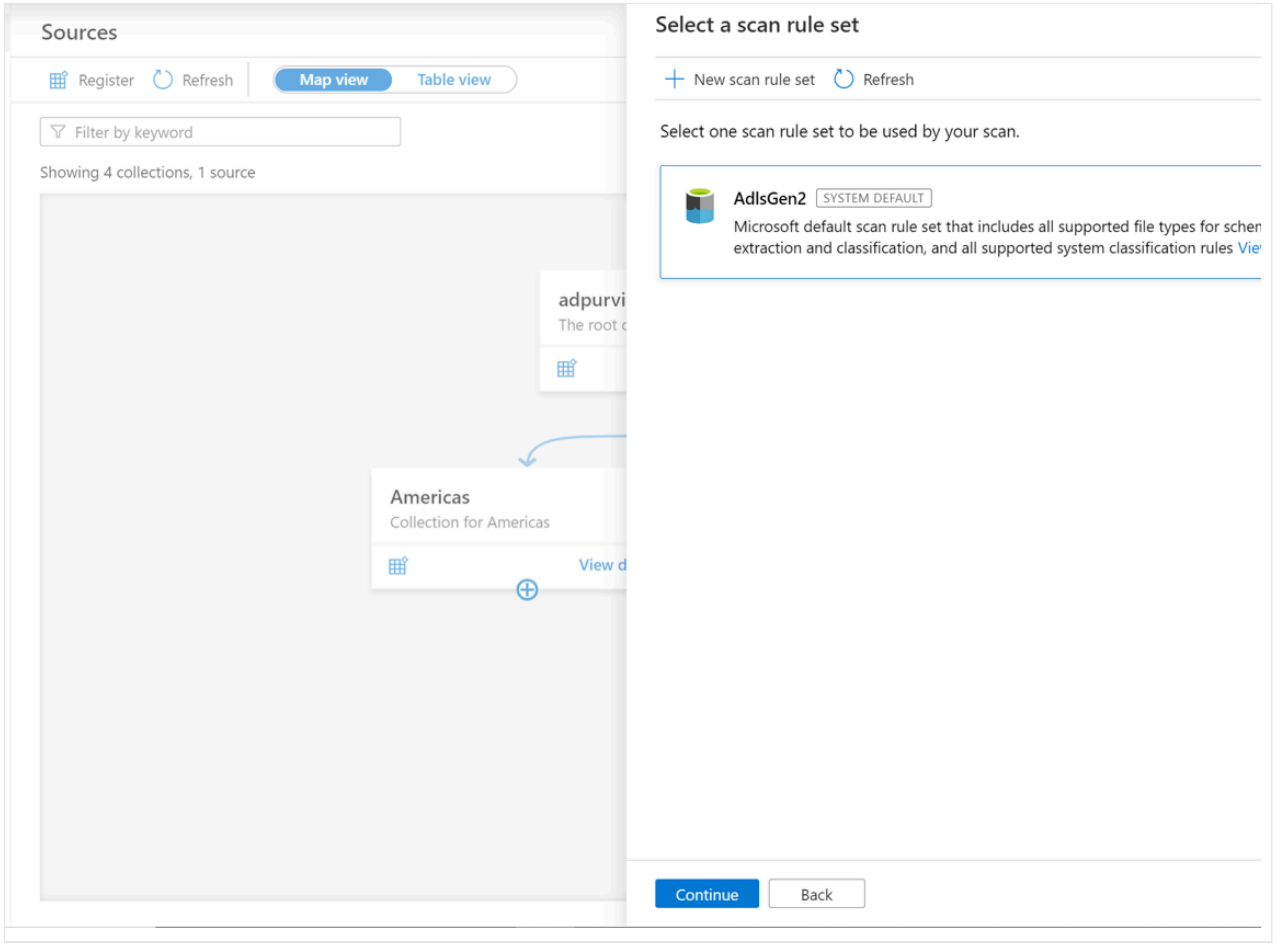
If using a system or user assigned managed identity

1. Provide a **Name** for the scan, select the system-assigned or user-assigned managed identity under **Credential**, choose the appropriate collection for the scan, and select **Test connection**. On a successful connection, select **Continue**.

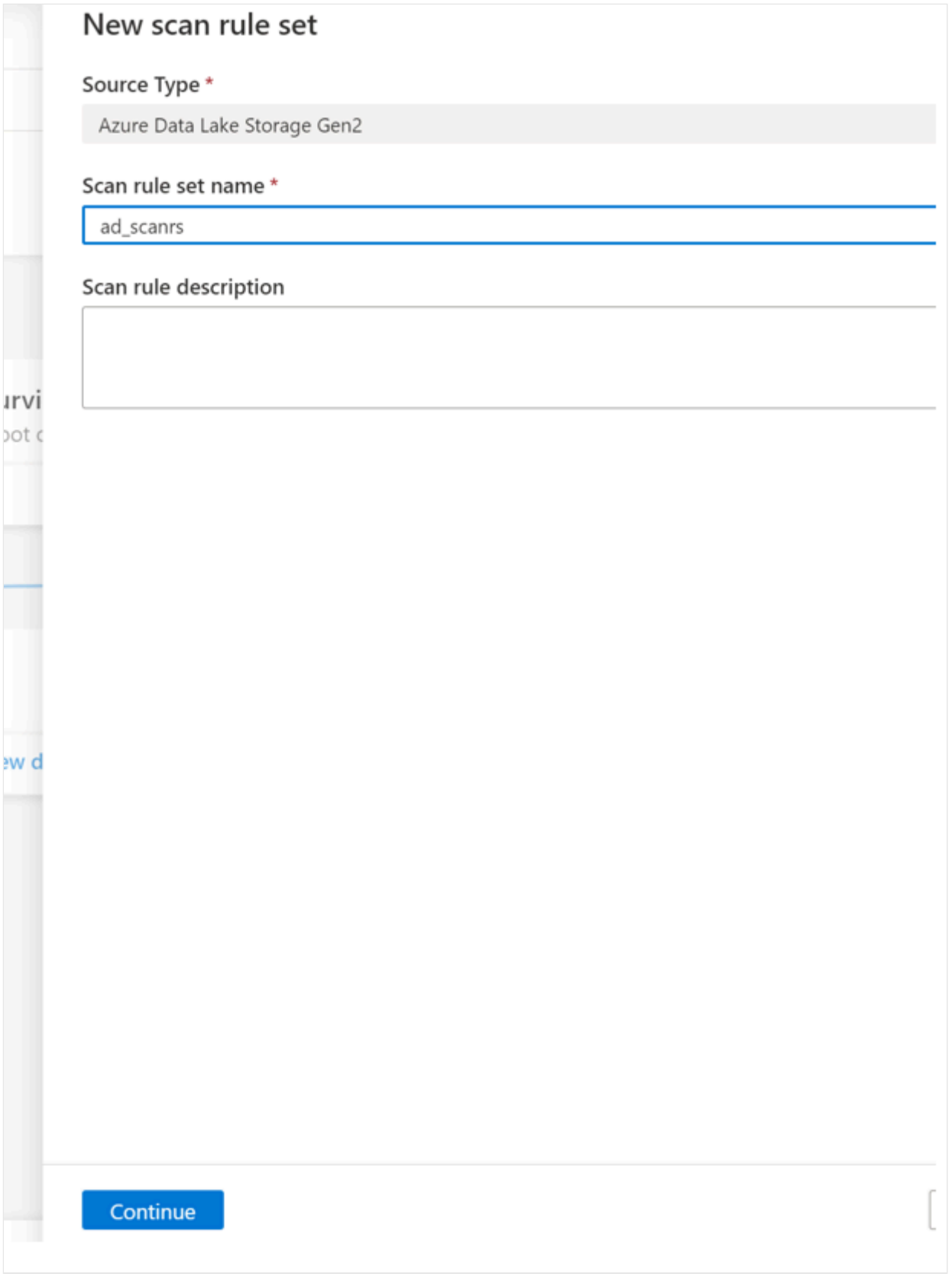


Scope and run the scan

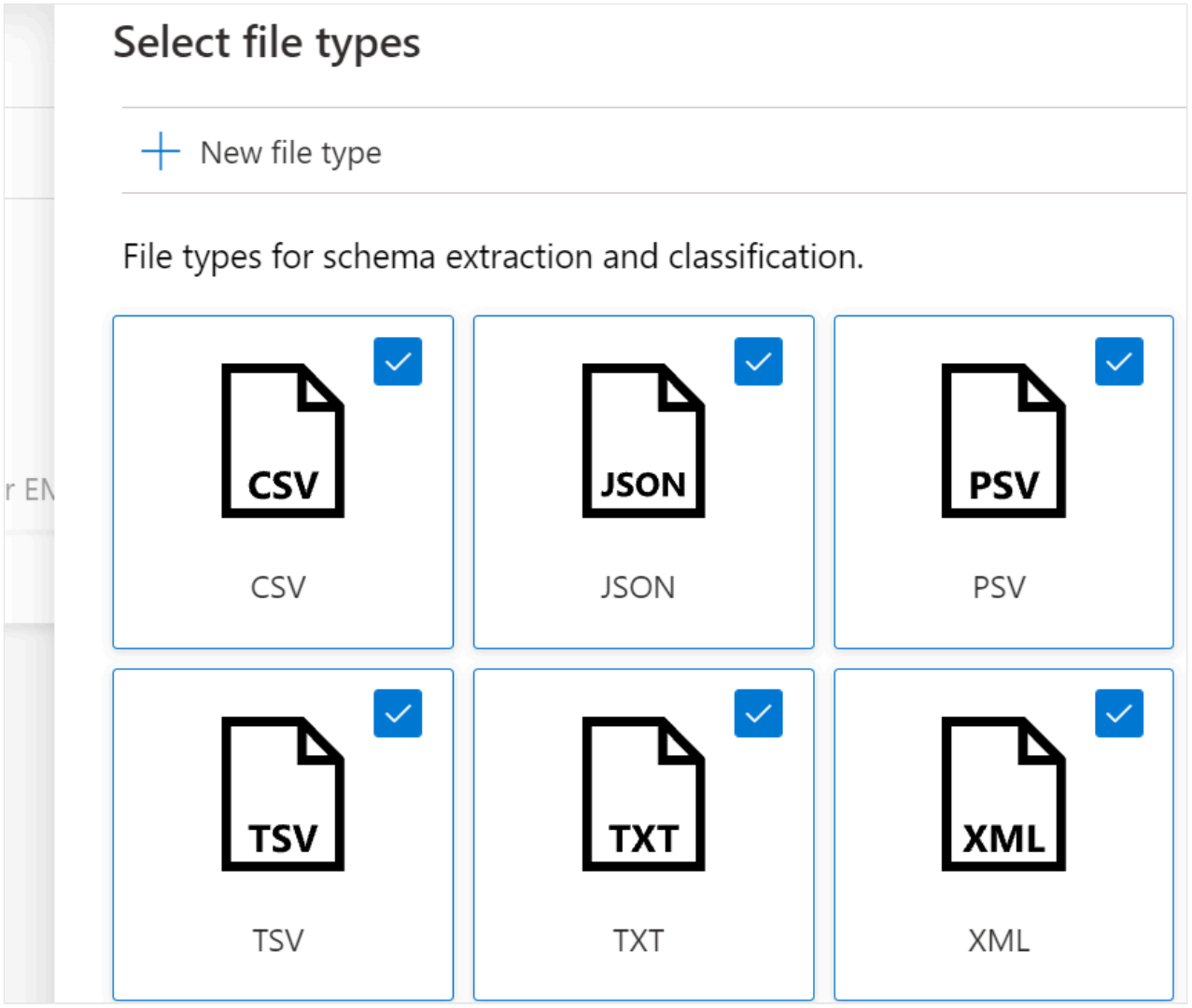
1. You can scope your scan to specific folders and subfolders by choosing the appropriate items in the list.



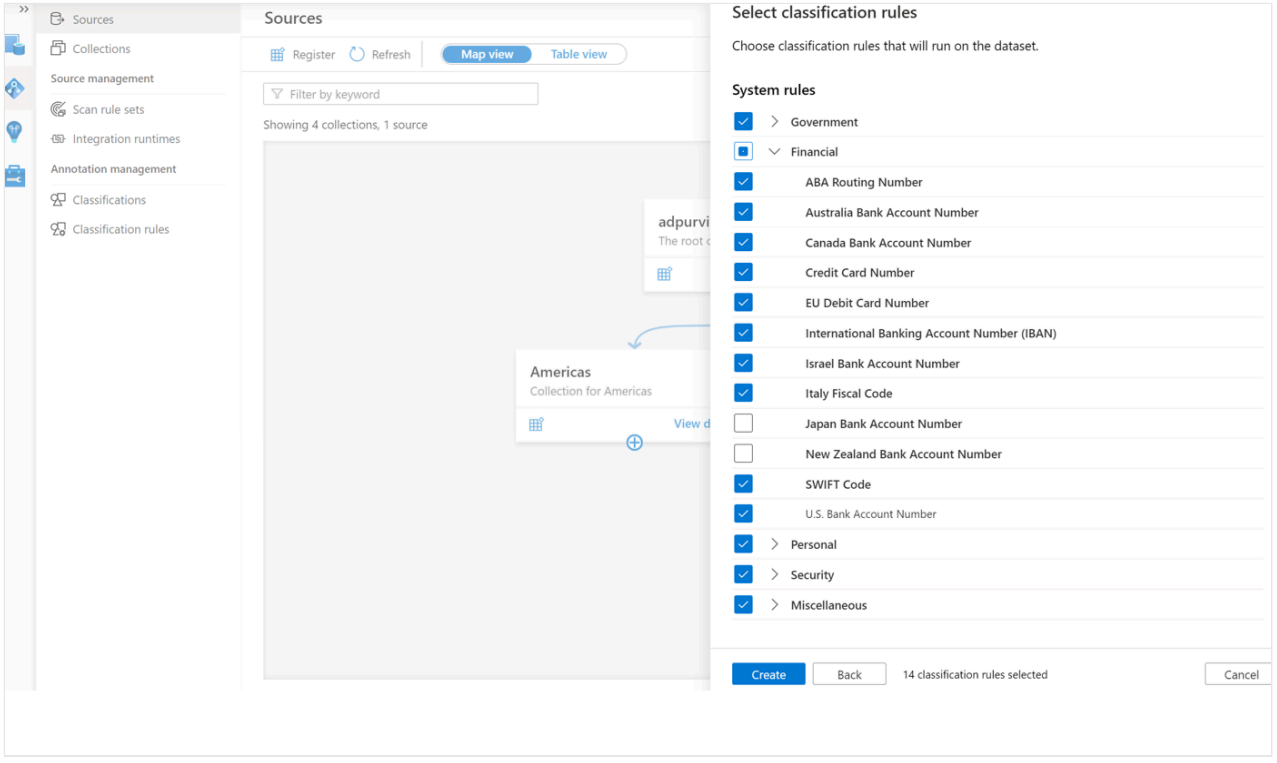
2. Then select a scan rule set. You can choose between the system default, existing custom rule sets, or create a new rule set inline.

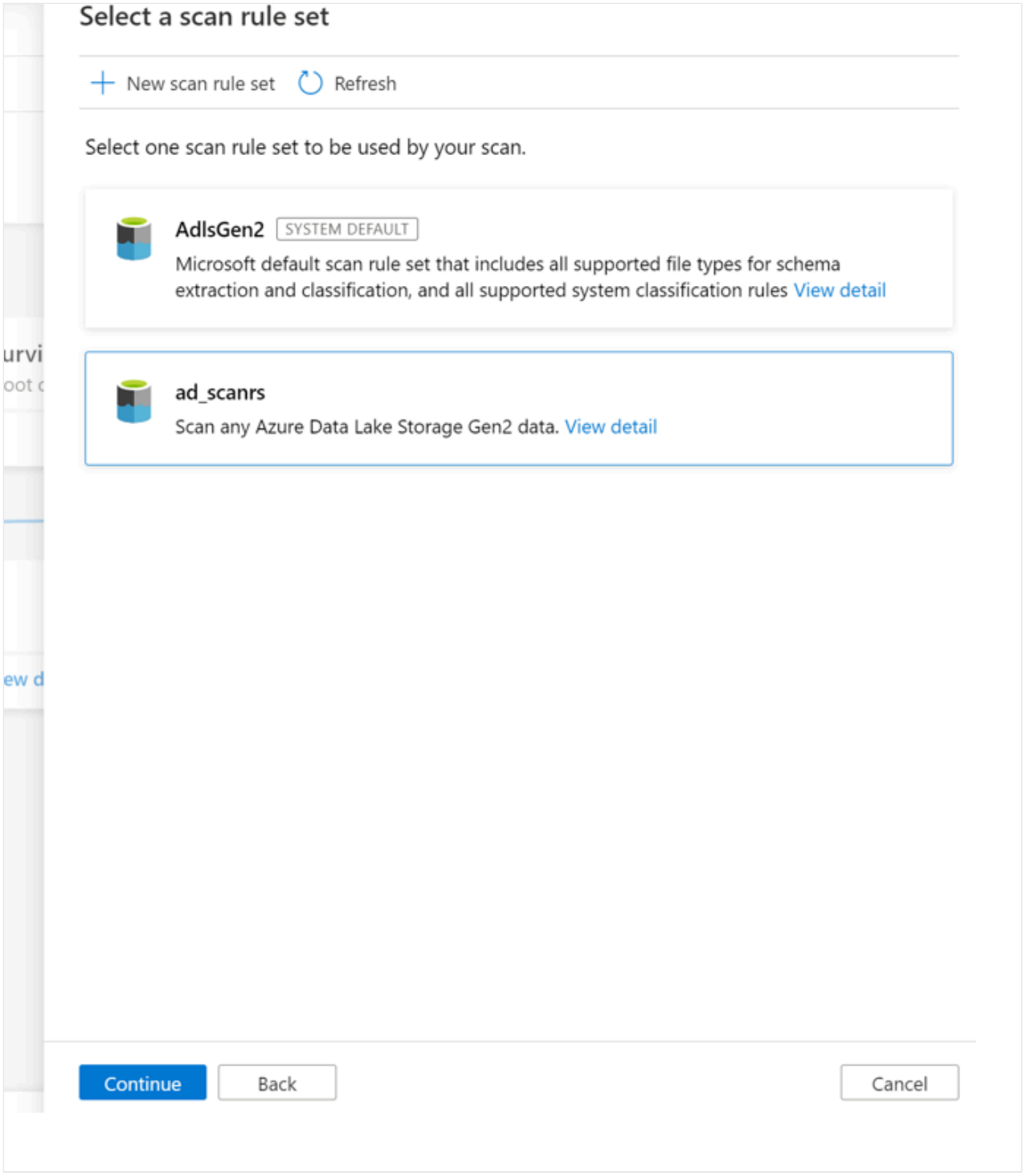


3. If creating a new *scan rule set*, select the **file types** to be included in the scan rule.

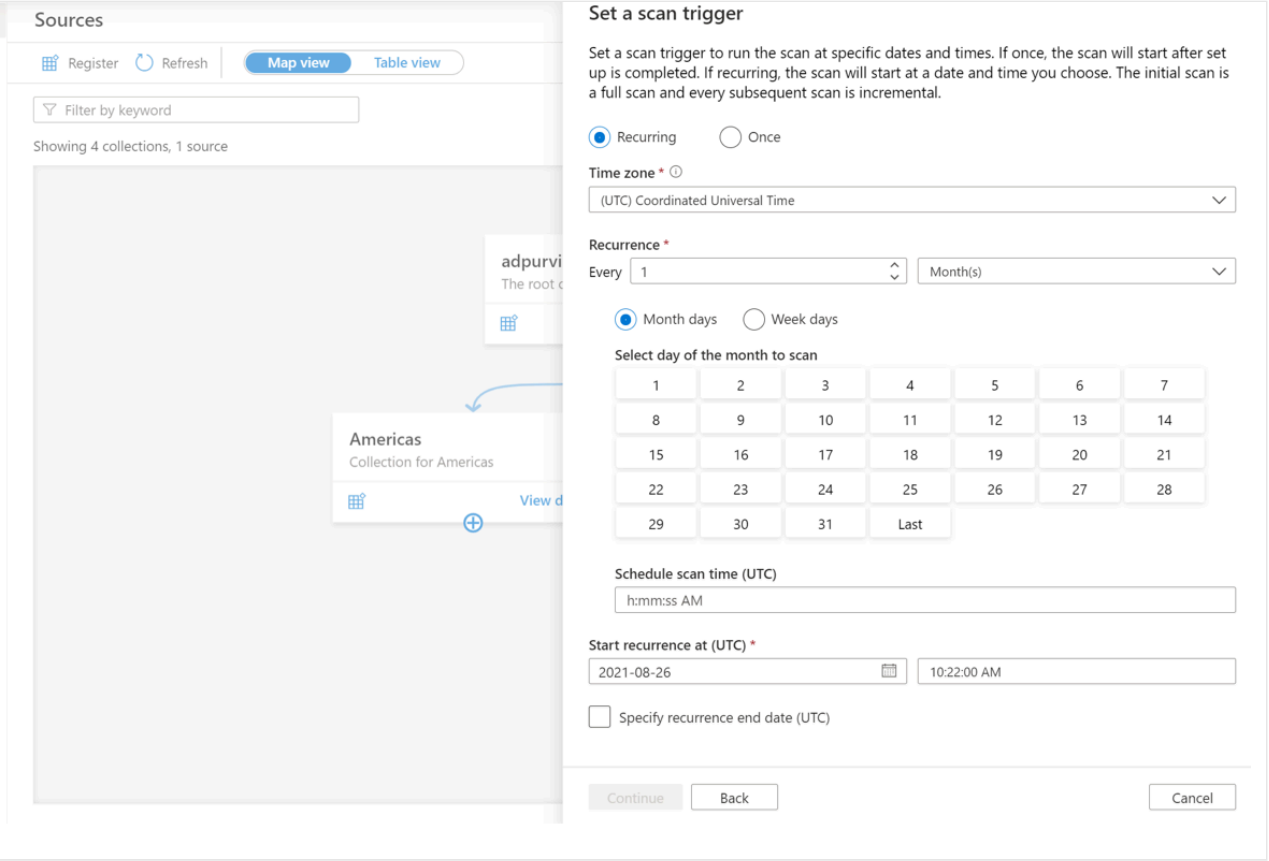


4. You can select the **classification rules** to be included in the scan rule

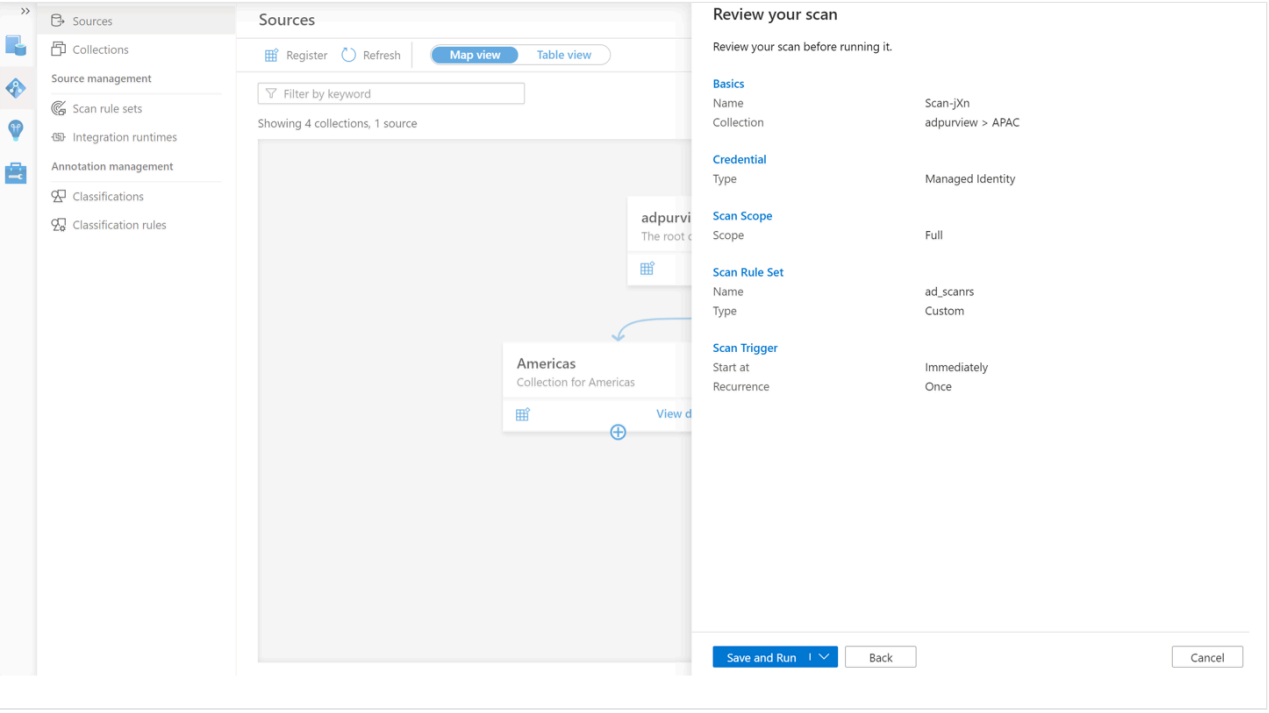




5. Choose your scan trigger. You can set up a schedule or run the scan once.



6. Review your scan and select **Save and run**.



View your scans and scan runs

To view existing scans:

1. Go to the Microsoft Purview portal. On the left pane, select **Data map**.

2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that has results you want to view. The pane shows you all the previous scan runs, along with the status and metrics for each scan run.
4. Select the run ID to check the scan run details.

Manage your scans

To edit, cancel, or delete a scan:

1. Go to the Microsoft Purview portal. On the left pane, select **Data Map**.

2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.

3. Select the scan that you want to manage. You can then:

    - Edit the scan by selecting **Edit scan**.
    - Cancel an in-progress scan by selecting **Cancel scan run**.
    - Delete your scan by selecting **Delete scan**.

> ⓘ Note
>
> - Deleting your scan does not delete catalog assets created from previous scans.

Data sharing

Microsoft Purview Data Sharing (preview) enables sharing of data in-place from ADLS Gen2 to ADLS Gen2. This section provides details about the ADLS Gen2 specific requirements for sharing and receiving data in-place. Refer to How to share data and How to receive share for step by step guide on how to use data sharing.

Storage accounts supported for in-place data sharing

The following storage accounts are supported for in-place data sharing:

- Regions: Canada Central, Canada East, UK South, UK West, Australia East, Japan East, Korea South, and South Africa North
- Redundancy options: LRS, GRS, RA-GRS
- Tiers: Hot, Cool

Only use storage accounts without production workload for the preview.

> **Note**
>
> Source and target storage accounts must be in the same region as each other. They don't need to be in the same region as the Microsoft Purview account.

Storage account permissions required to share data

To add or update a storage account asset to a share, you need ONE of the following permissions:

- **Microsoft.Authorization/roleAssignments/write** - This permission is available in the *Owner* role.
- **Microsoft.Storage/storageAccounts/blobServices/containers/blobs/modifyPermissions/** - This permission is available in the *Blob Storage Data Owner* role.

Storage account permissions required to receive shared data

To map a storage account asset in a received share, you need ONE of the following permissions:

- **Microsoft.Storage/storageAccounts/write** - This permission is available in the *Contributor* and *Owner* role.
- **Microsoft.Storage/storageAccounts/blobServices/containers/write** - This permission is available in the *Contributor*, *Owner*, *Storage Blob Data Contributor* and *Storage Blob Data Owner* role.

Update shared data in source storage account

Updates you make to shared files or data in the shared folder from source storage account will be made available to recipient in target storage account in near real time. When you delete subfolder or files within the shared folder, they'll disappear for recipient. To delete the shared folder, file or parent folders or containers, you need to first revoke access to all your shares from the source storage account.

Access shared data in target storage account

The target storage account enables recipient to access the shared data read-only in near real time. You can connect analytics tools such as Synapse Workspace and Databricks to the shared data to perform analytics. Cost of accessing the shared data is charged to the target storage account.

Service limit

Source storage account can support up to 20 targets, and target storage account can support up to 100 sources. If you require an increase in limit, contact Support.

Policies

The following types of policies are supported on this data resource from Microsoft Purview:

- Data owner policies - a set of policy statements that allow you to grant users and groups access to data sources.
- Self-service access policies - policy that allows users to request access to data sources registered to Microsoft Purview.
- Protection policies - denies access to data tagged with sensitivity labels to all users except those specified by the policy.

Access policy pre-requisites on Azure Storage accounts

Region support

- All Microsoft Purview regions are supported.
- Storage accounts in the following regions are supported without the need for additional configuration. However, zone-redundant storage (ZRS) accounts are not supported.
  - Australia Central
  - Australia East
  - Australia Southeast
  - Brazil South
  - Canada Central
  - Canada East
  - Central India
  - Central US
  - East Asia
  - East US 2
  - East US
  - France Central
  - Germany West Central
  - Japan East
  - Japan West
  - Korea Central
  - North Central US

    - North Europe
    - Norway East
    - Poland Central
    - Qatar Central
    - South Central US
    - South Africa North
    - Southeast Asia
    - South India
    - Sweden Central
    - Switzerland North
    - West Central US
    - West Europe
    - West US
    - West US 2
    - West US 3
    - UAE North
    - UK South
    - UK West
- Storage accounts in other regions in Public Cloud are supported after setting feature flag *AllowPurviewPolicyEnforcement*, as outlined in the next section. Newly created ZRS Storage accounts are supported, if created after setting the feature flag *AllowPurviewPolicyEnforcement*.

If needed, you can create a new Storage account by following this guide.

Configure the subscription where the Azure Storage account resides for policies from Microsoft Purview

This step is only necessary in certain regions (see prior section). To enable Microsoft Purview to manage policies for one or more Azure Storage accounts, execute the following PowerShell commands in the subscription where you'll deploy your Azure Storage account. These PowerShell commands will enable Microsoft Purview to manage policies on all Azure Storage accounts in that subscription.

If you're executing these commands locally, be sure to run PowerShell as an administrator. Alternatively, you can use the Azure Cloud Shell in the Azure portal: https://shell.azure.com.

PowerShell

```
# Install the Az module
Install-Module -Name Az -Scope CurrentUser -Repository PSGallery -Force
# Login into the subscription
Connect-AzAccount -Subscription <SubscriptionID>
```

```
# Register the feature
Register-AzProviderFeature -FeatureName AllowPurviewPolicyEnforcement -
ProviderNamespace Microsoft.Storage
```

If the output of the last command shows *RegistrationState* as *Registered*, then your subscription is enabled for access policies. If the output is *Registering*, wait at least 10 minutes, and then retry the command. **Do not continue unless the RegistrationState shows as *Registered***.

Configure the Microsoft Purview account for policies

Register the data source in Microsoft Purview

Before a policy can be created in Microsoft Purview for a data resource, you must register that data resource in Microsoft Purview Studio. You will find the instructions related to registering the data resource later in this guide.

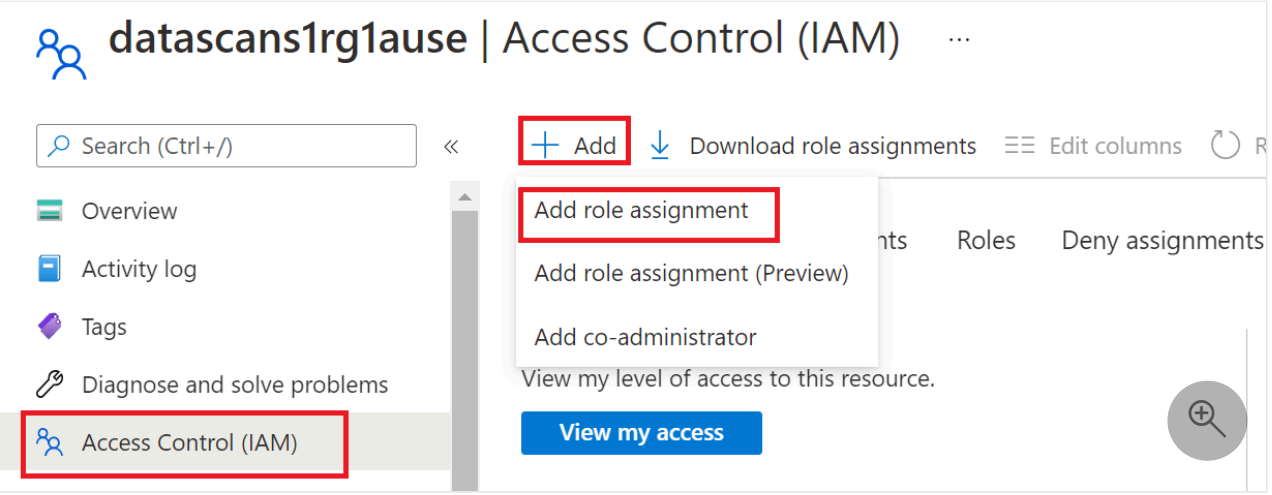
> **Note**
>
> Microsoft Purview policies rely on the data resource ARM path. If a data resource is moved to a new resource group or subscription it will need to be de-registered and then registered again in Microsoft Purview.

Configure permissions to enable Data policy enforcement on the data source

Once a resource is registered, but before a policy can be created in Microsoft Purview for that resource, you must configure permissions. A set of permissions are needed to enable the **Data policy enforcement**. This applies to data sources, resource groups, or subscriptions. To enable *Data policy enforcement*, you must have *both* specific Identity and Access Management (IAM) privileges on the resource as well as specific Microsoft Purview privileges:

- You must have *either one* of the following IAM role combinations on the resource's Azure Resource Manager path or any parent of it (that is, using IAM permission inheritance):
  - IAM Owner
  - Both IAM Contributor and IAM User Access Administrator

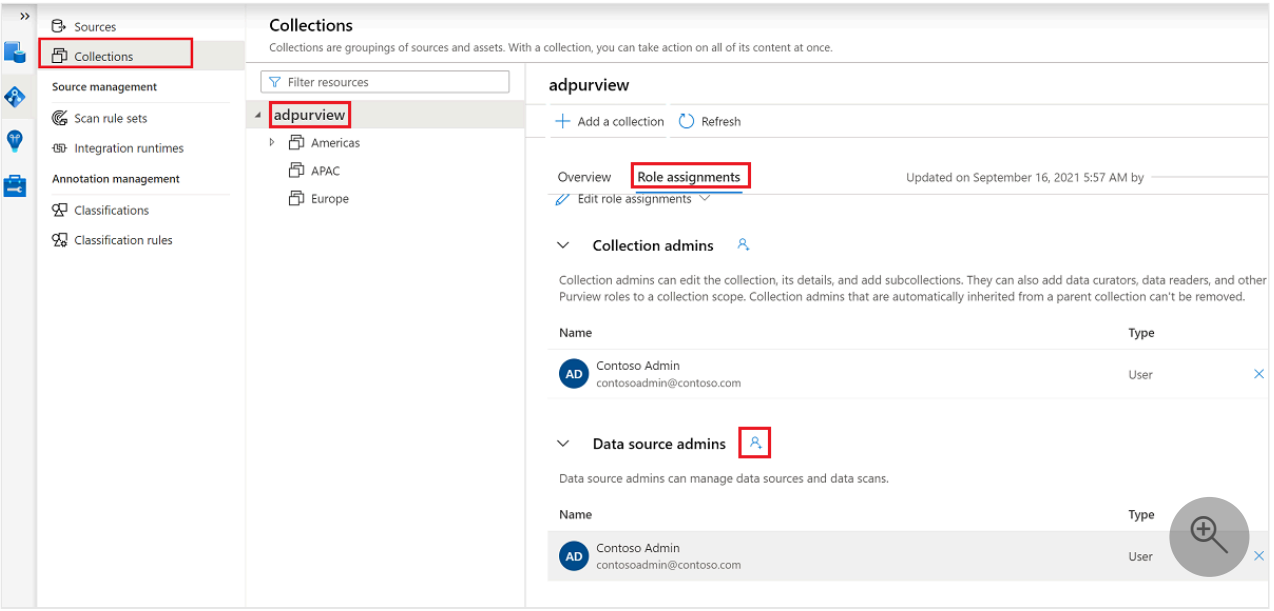
  To configure Azure role-based access control (RBAC) permissions, follow this guide. The following screenshot shows how to access the **Access Control** section in the Azure portal for the data resource to add a role assignment.



> **Note**
>
> The *IAM Owner* role for a data resource can be inherited from a parent resource group, a subscription, or a subscription management group. Check which Microsoft Entra users, groups, and service principals hold or are inheriting the *IAM Owner* role for the resource.

- You also need to have the Microsoft Purview *Data source admin* role for the collection or a parent collection (if inheritance is enabled). For more information, see the guide on managing Microsoft Purview role assignments.

  The following screenshot shows how to assign the *Data source admin* role at the root collection level.



Configure Microsoft Purview permissions to create, update, or delete access policies

To create, update or delete policies, you need to get the *Policy author* role in Microsoft Purview at *root collection level*:

- The *Policy author* role can create, update, and delete DevOps and Data Owner policies.
- The *Policy author* role can delete self-service access policies.

For more information about managing Microsoft Purview role assignments, see Create and manage collections in the Microsoft Purview Data Map.

> **Note**
>
> Policy author role must be configured at the root collection level.

In addition, to easily search Microsoft Entra users or groups when creating or updating the subject of a policy, you can greatly benefit from getting the Directory Readers permission in Microsoft Entra ID. This is a common permission for users in an Azure tenant. Without the Directory Reader permission, the Policy Author will have to type the complete username or email for all the principals included in the subject of a data policy.

Configure Microsoft Purview permissions for publishing Data Owner policies

Data Owner policies allow for checks and balances if you assign the Microsoft Purview *Policy author* and *Data source admin* roles to different people in the organization. Before a Data owner policy takes effect, a second person (*Data source admin*) must review it and explicitly approve it by publishing it. This does not apply to DevOps or Self-service access policies as publishing is automatic for them when those policies are created or updated.

To publish a Data owner policy you need to get the *Data source admin* role in Microsoft Purview at *root collection level*.

For more information about managing Microsoft Purview role assignments, see Create and manage collections in the Microsoft Purview Data Map.

> **Note**
>
> To publish Data owner policies, the Data source admin role must be configured at the root collection level.

Delegate access provisioning responsibility to roles in Microsoft Purview

After a resource has been enabled for **Data policy enforcement**, any Microsoft Purview user with the *Policy author* role at the root collection level can provision access to that data source from Microsoft Purview.

> ⓘ **Note**
>
> Any Microsoft Purview root *Collection admin* can assign new users to root *Policy author* roles. Any *Collection admin* can assign new users to a *Data source admin* role under the collection. Minimize and carefully vet the users who hold Microsoft Purview *Collection admin*, *Data source admin*, or *Policy author* roles.

If a Microsoft Purview account with published policies is deleted, such policies will stop being enforced within an amount of time that depends on the specific data source. This change can have implications on both security and data access availability. The Contributor and Owner roles in IAM can delete Microsoft Purview accounts. You can check these permissions by going to the **Access control (IAM)** section for your Microsoft Purview account and selecting **Role Assignments**. You can also use a lock to prevent the Microsoft Purview account from being deleted through Resource Manager locks.

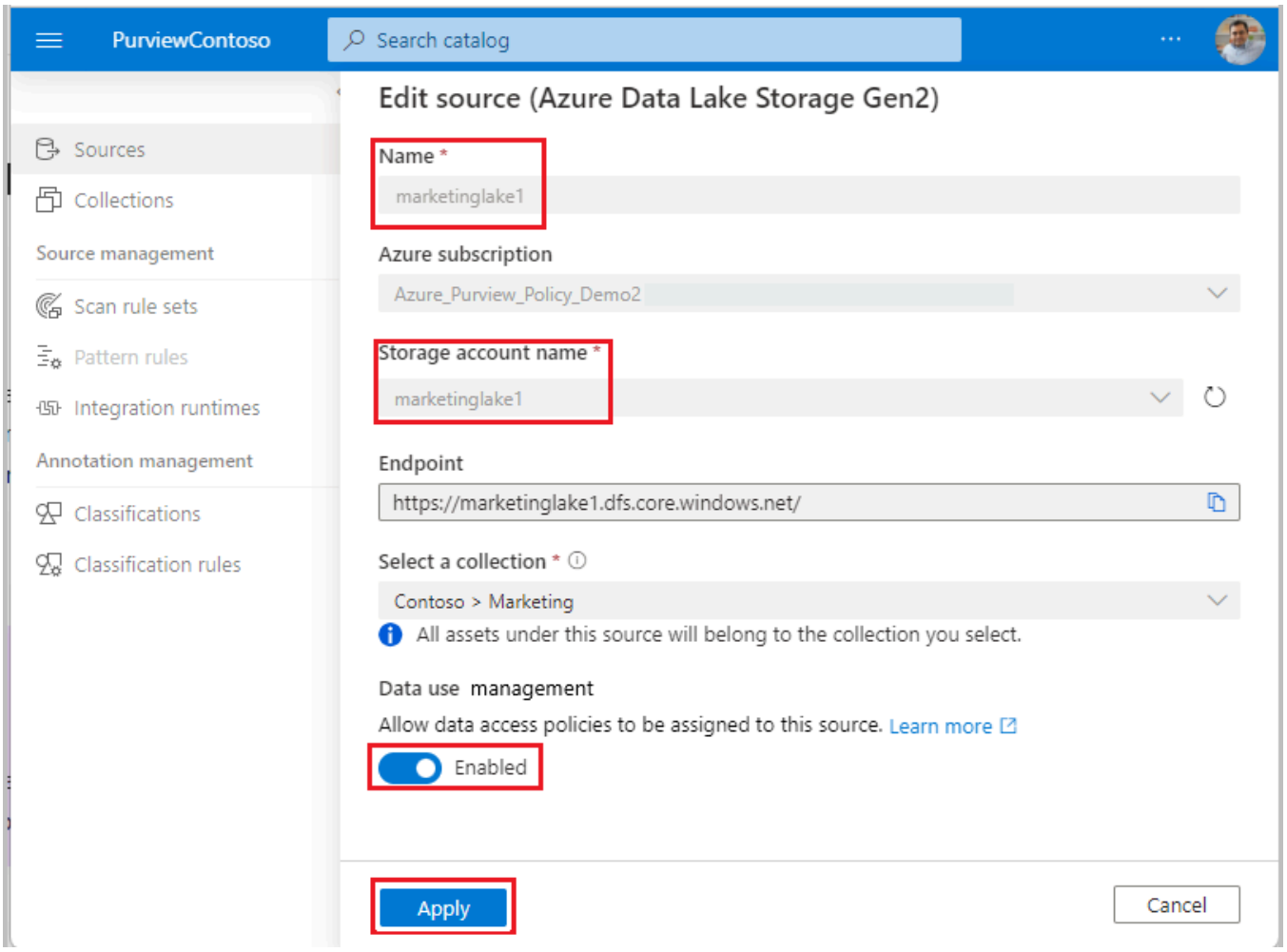
Register the data source in Microsoft Purview for Data Policy Enforcement

The Azure Storage resource needs to be registered first with Microsoft Purview before you can create access policies. To register your resource, follow the **Prerequisites** and **Register** sections of this guide:

- Register Azure Data Lake Storage (ADLS) Gen2 in Microsoft Purview

After you've registered the data source, you'll need to enable Data Policy Enforcement. This is a pre-requisite before you can create policies on the data source. Data Policy Enforcement can impact the security of your data, as it delegates to certain Microsoft Purview roles managing access to the data sources. **Go through the secure practices related to Data Policy Enforcement in this guide**: How to enable Data Policy Enforcement

Once your data source has the **Data Policy Enforcement** option set to **Enabled**, it will look like this screenshot:



Create a policy

To create an access policy for Azure Data Lake Storage Gen2, follow this guide:

- Provision read/modify access on a single storage account

To create policies that cover all data sources inside a resource group or Azure subscription you can refer to this section.

Protection policy

Protection access control policies (protection policies) enable organizations to automatically protect sensitive data across data sources. Microsoft Purview already scans data assets and identifies sensitive data elements, and this new feature allows you to automatically restrict access to that data using sensitivity labels from Microsoft Purview Information Protection.

Follow this documentation to create a protection policy: How to create a Microsoft Purview Information Protection policy.

Next steps

Follow the below guides to learn more about Microsoft Purview and your data.

- Data owner policies in Microsoft Purview
- Data Estate Insights in Microsoft Purview
- Lineage in Microsoft Purview
- Data share in Microsoft Purview

Connect to and manage Azure Database for MySQL in Microsoft Purview

Article • 04/01/2025

This article outlines how to register a database in Azure Database for MySQL, and how to authenticate and interact with Azure Database for MySQL in Microsoft Purview. For more information about Microsoft Purview, read the introductory article.

Supported capabilities

Scanning capabilities

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan
Yes	Yes	Yes*	Yes

* Microsoft Purview relies on UPDATE_TIME metadata from Azure Database for MySQL for incremental scans. In some cases, this field might not persist in the database and a full scan is performed. For more information, see The INFORMATION_SCHEMA TABLES Table for MySQL.

> **Important**
>
> Microsoft Purview only supports single server deployment option for Azure Database for MySQL.

Other capabilities

View a table listing the supported capabilities for classifications, sensitivity labels, policies, data lineage, and live view.

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- An active Microsoft Purview account.

- You'll need to be a Data Source Administrator and Data Reader to register a source and manage it in the Microsoft Purview governance portal. See our Microsoft Purview Permissions page for details.

Register

This section describes how to register an Azure Database for MySQL in Microsoft Purview using the Microsoft Purview governance portal.

Authentication for registration

You'll need **username** and **password** for the next steps.

Follow the instructions in CREATE DATABASES AND USERS to create a login for your Azure Database for MySQL.

1. Navigate to your key vault in the Azure portal
2. Select **Settings** > **Secrets**
3. Select + **Generate/Import** and enter the **Name** and **Value** as the *password* from your Azure SQL Database
4. Select **Create** to complete
5. If your key vault isn't connected to Microsoft Purview yet, you'll need to create a new key vault connection
6. Finally, create a new credential of type SQL authentication using the **username** and **password** to set up your scan.

Steps to register

To register a new Azure Database for MySQL in Microsoft Purview Unified Catalog, do the following:

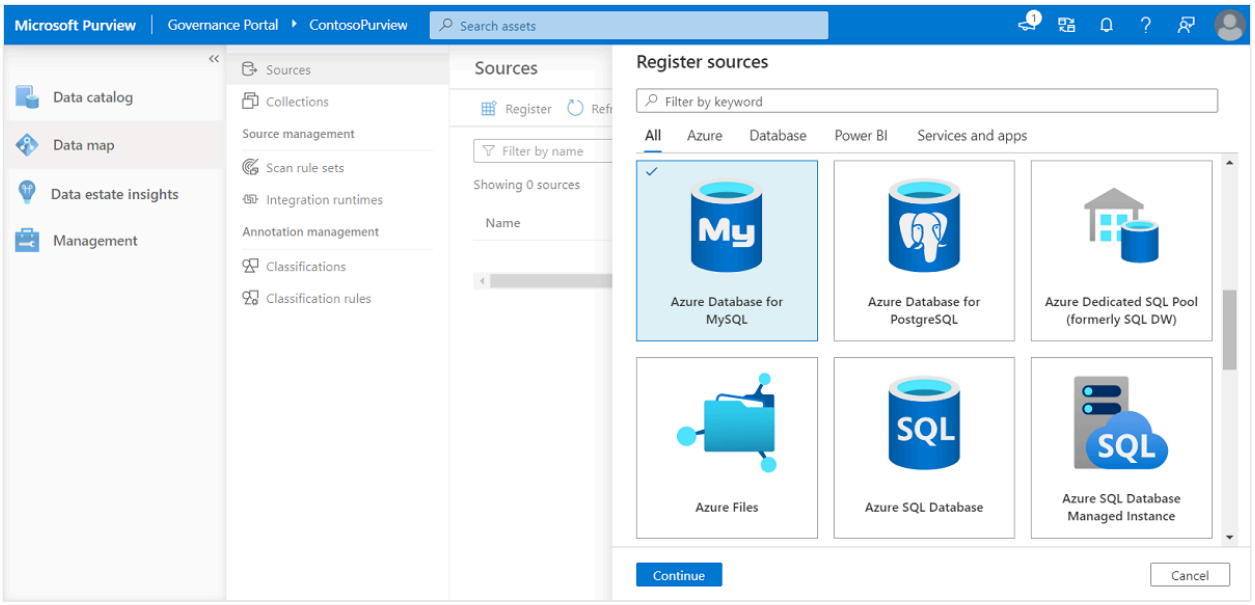
1. Open the Microsoft Purview governance portal by:

    - Browsing directly to https://web.purview.azure.com and selecting your Microsoft Purview account.
    - Opening the Azure portal, searching for and selecting the Microsoft Purview account. Selecting the **the Microsoft Purview governance portal** button.

2. Select **Data Map** on the left navigation.

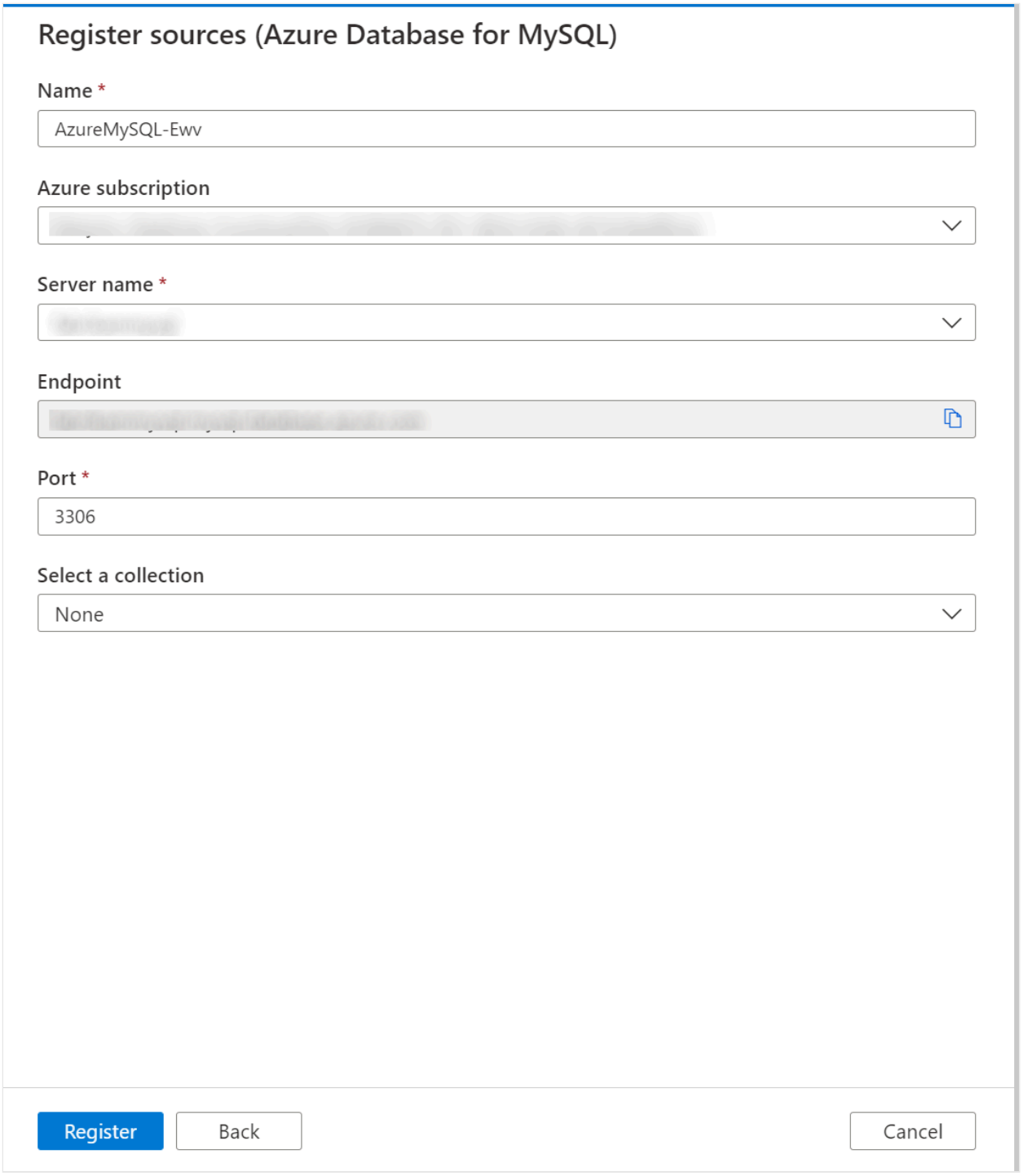
3. Select **Register**.

4. On **Register sources**, select **Azure Database for MySQL**. Select **Continue**.



On the **Register sources (Azure Database for MySQL)** screen, do the following:

1. Enter a **Name** for your data source. This will be the display name for this data source in your Catalog.
2. Select **From Azure subscription**, select the appropriate subscription from the **Azure subscription** drop-down box and the appropriate server from the **Server name** drop-down box.
3. Select **Register** to register the data source.



Scan

Follow the steps below to scan Azure Database for MySQL to automatically identify assets and classify your data. For more information about scanning in general, see our introduction to scans and ingestion

Create and run scan

To create and run a new scan, do the following:

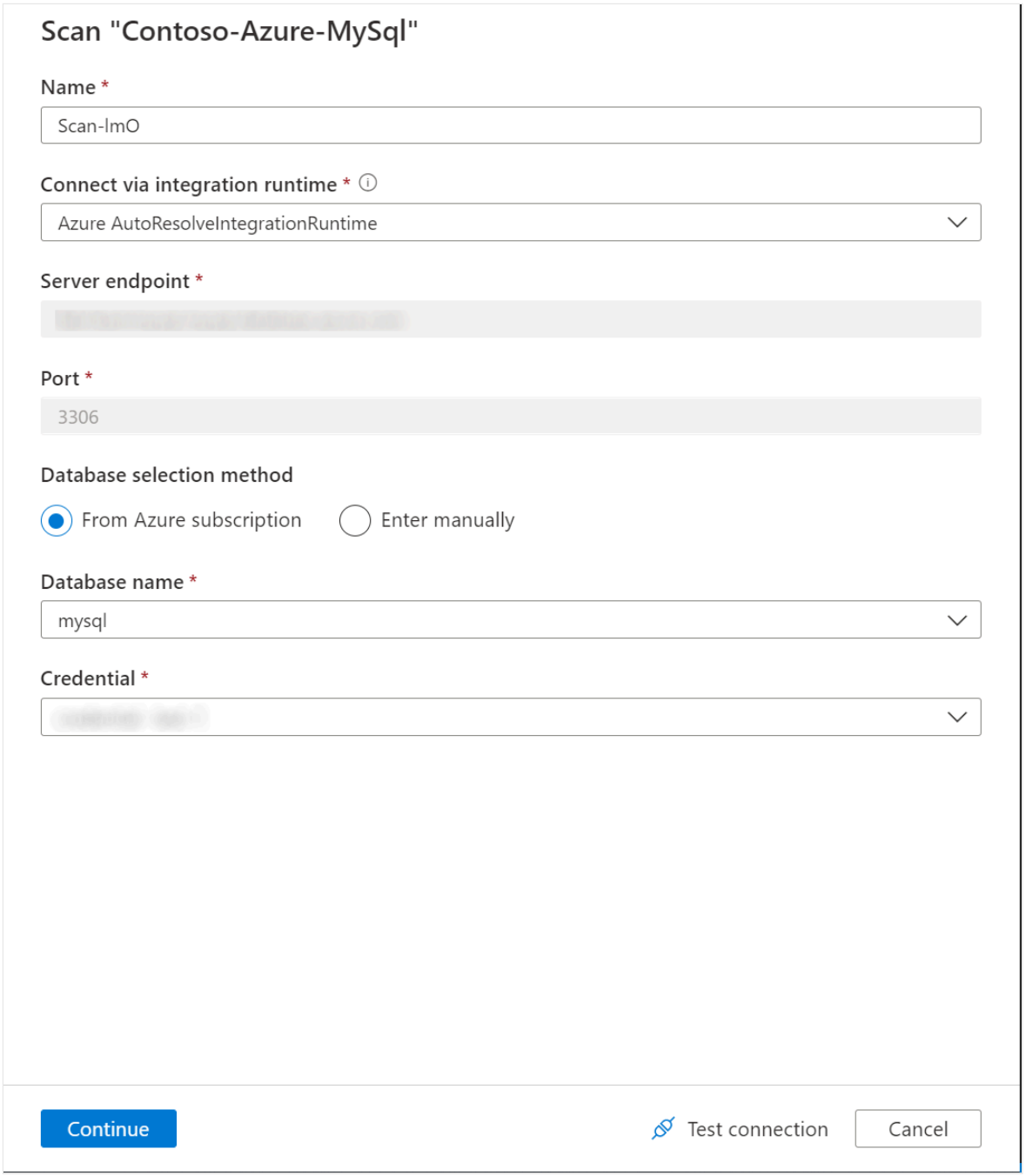
1. Select the **Data Map** tab on the left pane in the Microsoft Purview governance portal.

2. Select the Azure Database for MySQL source that you registered.

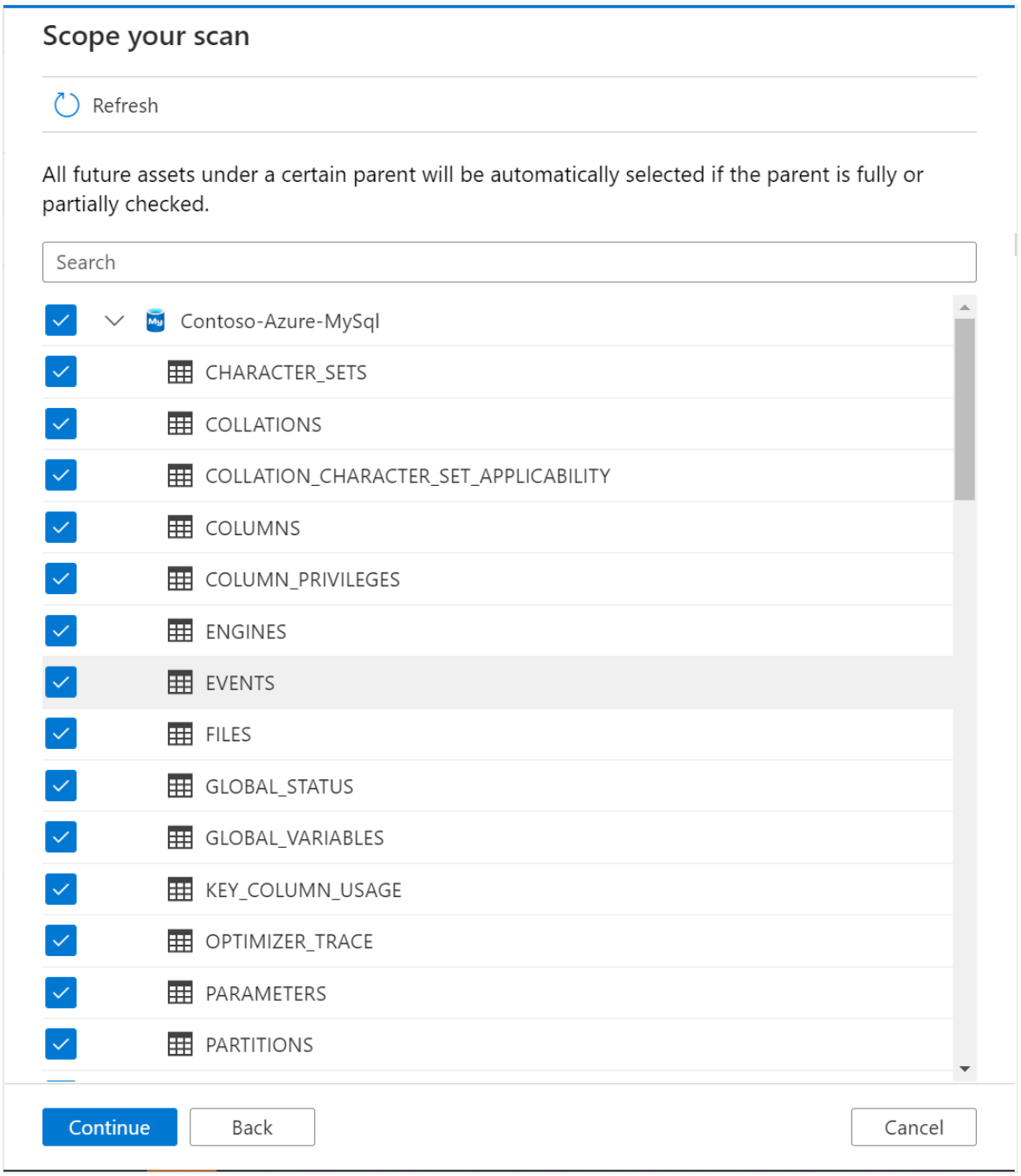
3. Select **New scan**

4. Choose either the Azure integration runtime if your source is publicly accessible, a managed virtual network integration runtime if using a managed virtual network, or a self-hosted integration runtime if your source is in a private virtual network. For more information about which integration runtime to use, see the choose the right integration runtime configuration article.

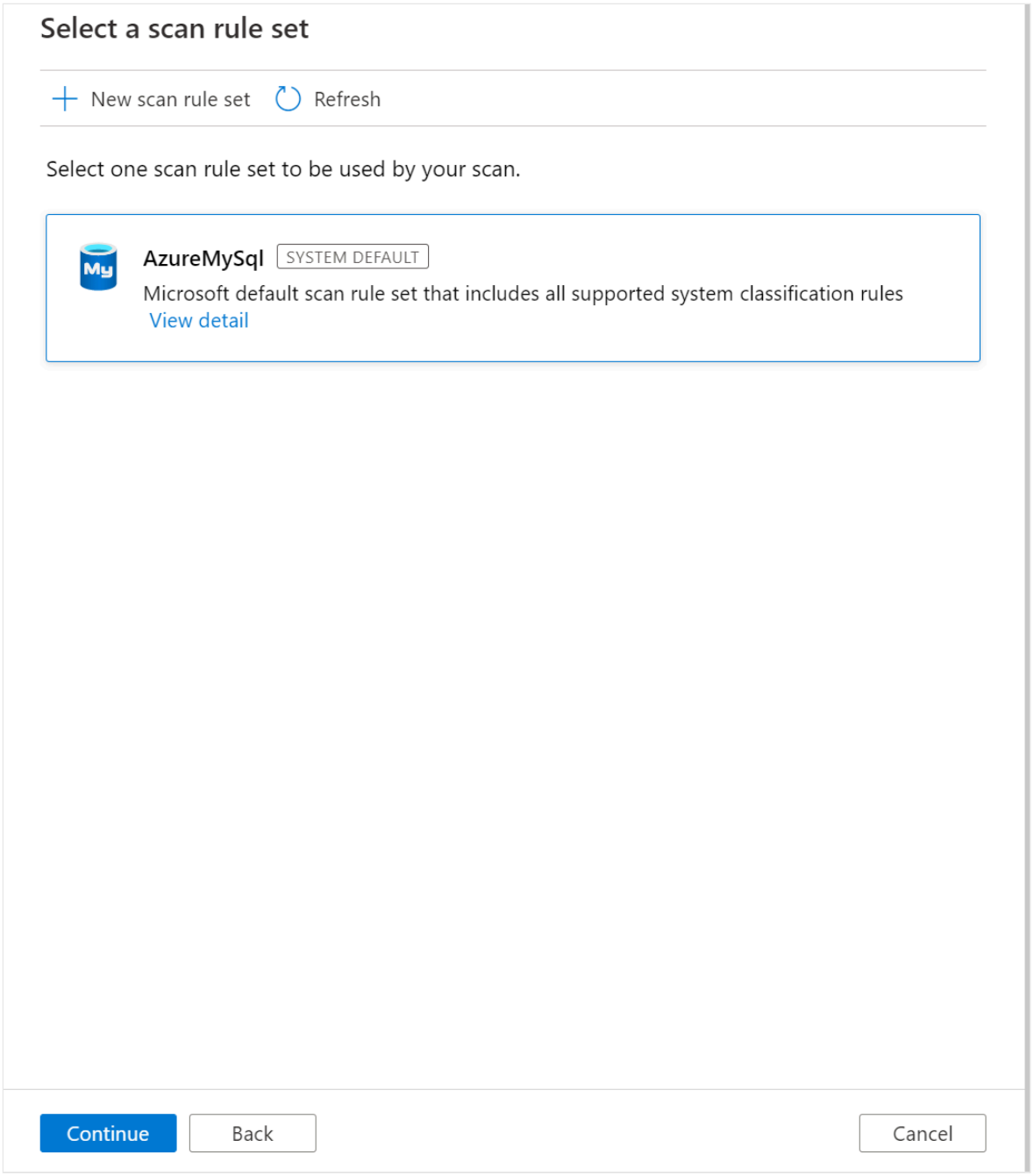
5. Select the credential to connect to your data source and check your connection to make sure your credential is properly configured.



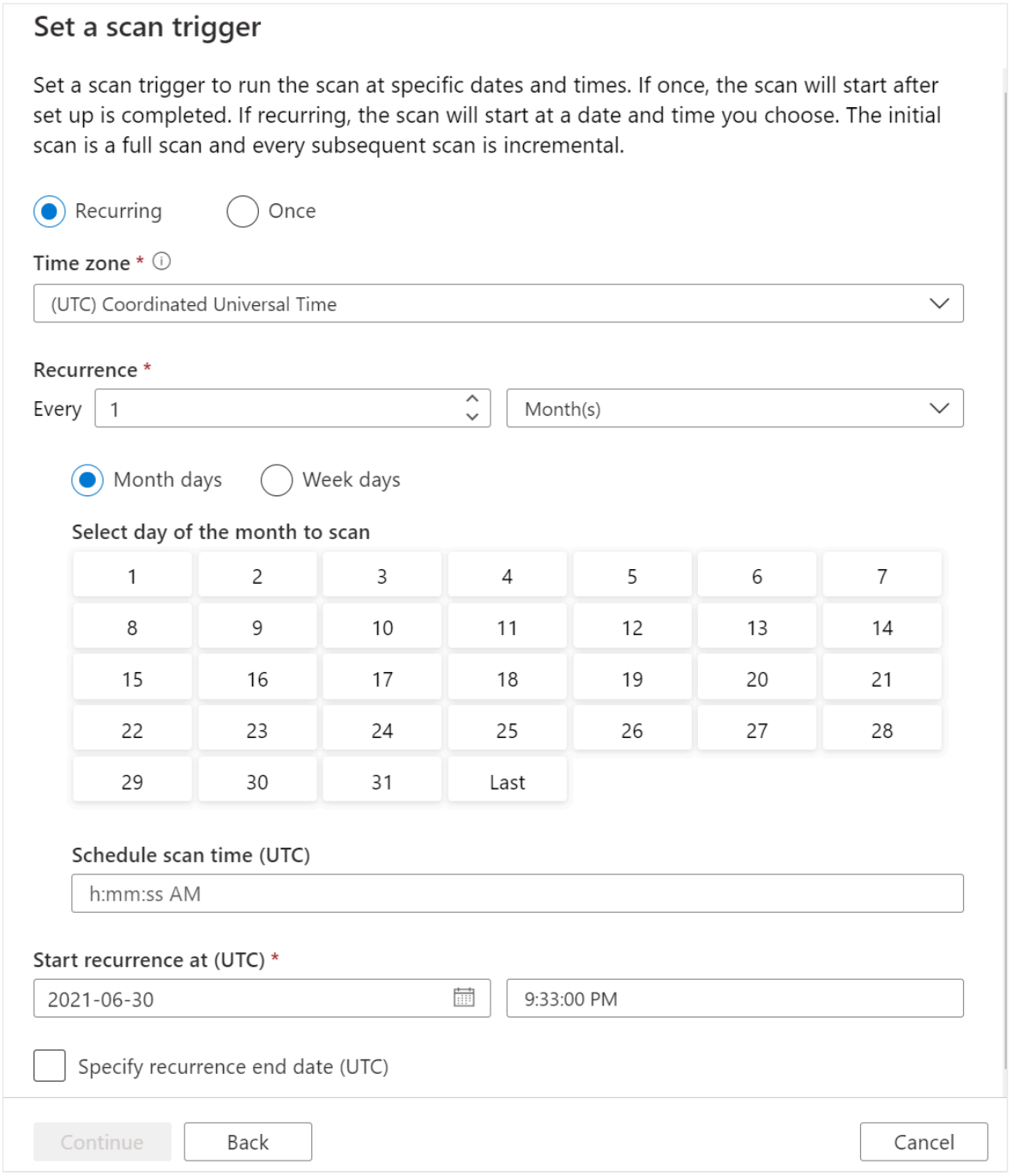
6. You can scope your scan to specific folders or subfolders by choosing the appropriate items in the list.



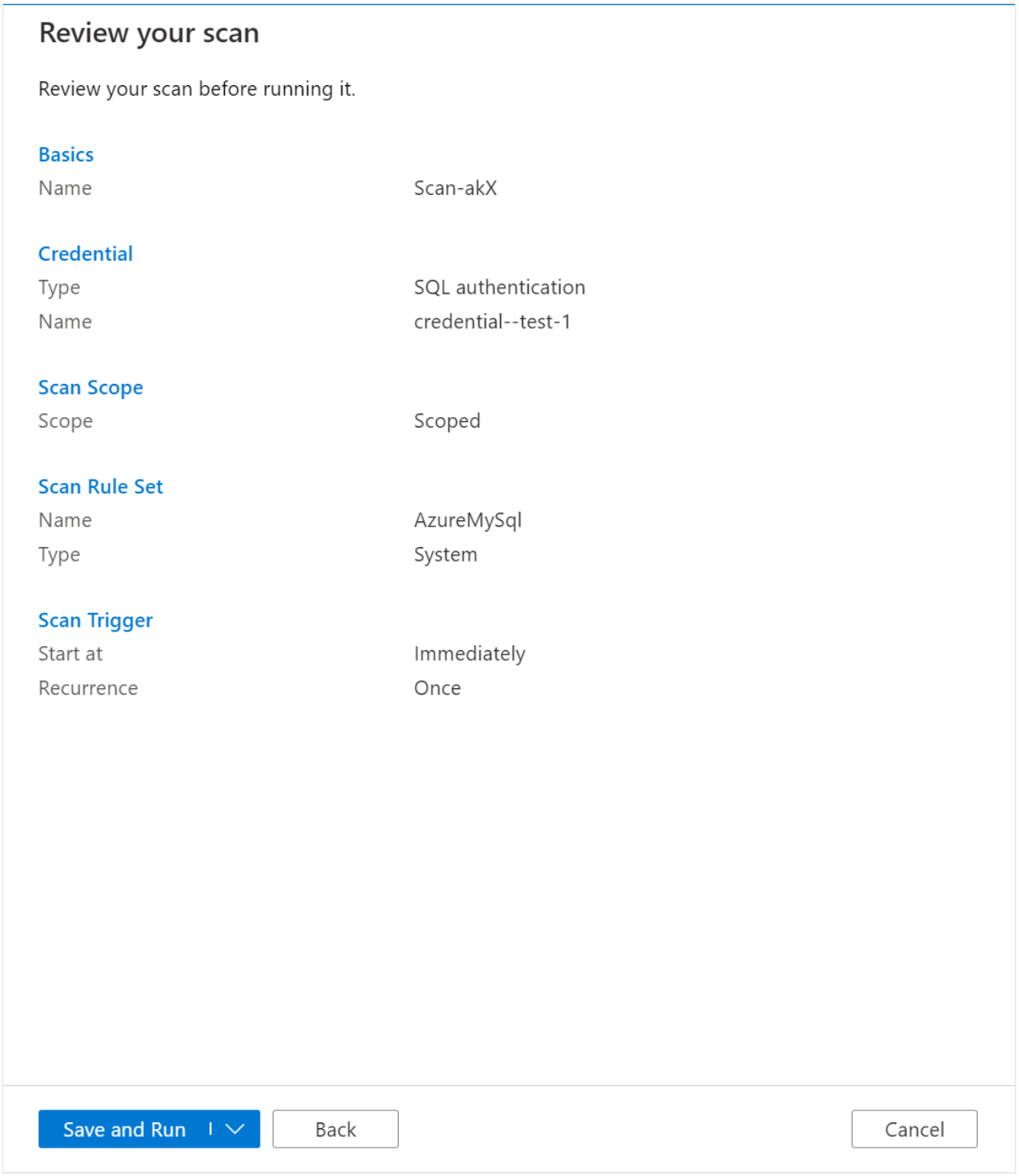
7. Then select a scan rule set. You can choose between the system default, existing custom rule sets, or create a new rule set inline.



8. Choose your scan trigger. You can set up a schedule or run the scan once.



9. Review your scan and select **Save and run**.-->



View your scans and scan runs

To view existing scans:

1. Go to the Microsoft Purview portal. On the left pane, select **Data map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that has results you want to view. The pane shows you all the previous scan runs, along with the status and metrics for each scan run.
4. Select the run ID to check the scan run details.

Manage your scans

To edit, cancel, or delete a scan:

1. Go to the Microsoft Purview portal. On the left pane, select **Data Map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that you want to manage. You can then:
    - Edit the scan by selecting **Edit scan**.
    - Cancel an in-progress scan by selecting **Cancel scan run**.
    - Delete your scan by selecting **Delete scan**.

> **Note**
>
> - Deleting your scan does not delete catalog assets created from previous scans.

Next steps

Now that you've registered your source, follow the below guides to learn more about Microsoft Purview and your data.

- Data Estate Insights in Microsoft Purview
- Lineage in Microsoft Purview
- Search Unified Catalog

Feedback

Was this page helpful? Yes No

Connect to and manage an Azure Database for PostgreSQL in Microsoft Purview (single server source retired in March 2025)

This article outlines how to register an Azure Database for PostgreSQL deployed with single server deployment option, and how to authenticate and interact with an Azure Database for PostgreSQL in Microsoft Purview. For more information about Microsoft Purview, read the introductory article.

Supported capabilities

Scanning capabilities

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan
Yes	Yes	No	Yes

Important

Microsoft Purview only supports single server deployment option for Azure Database for PostgreSQL. Versions 8.x to 12.x.

> **Azure Database for PostgreSQL single server retired in March 2025. We are working on bringing support for Azure Database for PostgreSQL Flex server.**

Other capabilities

For **classifications**, **sensitivity labels**, **policies**, **data lineage**, and **live view**, see the list of supported capabilities.

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- An active Microsoft Purview account.

- You'll need to be a Data Source Administrator and Data Reader to register a source and manage it in the Microsoft Purview governance portal. See our Microsoft Purview Permissions page for details.

Register

This section describes how to register an Azure Database for PostgreSQL in Microsoft Purview using the Microsoft Purview governance portal.

Authentication for registration

Currently, to be able to manage and interact with an Azure Database for PostgreSQL single server, only SQL Authentication is supported.

SQL Authentication

Connecting to an Azure Database for PostgreSQL database requires the fully qualified server name and login credentials. You can follow the instructions in CONNECT AND QUERY to create a login for your Azure Database for PostgreSQL if you don't have this available. You'll need **username** and **password** for the next steps.

1. If you don't have an Azure Key vault already, follow this guide to create an Azure Key Vault.
2. Navigate to your key vault in the Azure portal
3. Select **Settings** > **Secrets**
4. Select **+ Generate/Import** and enter the **Name** and **Value** as the *password* from your Azure PostgreSQL Database
5. Select **Create** to complete
6. If your key vault isn't connected to Microsoft Purview yet, you'll need to create a new key vault connection
7. Finally, create a new credential of type SQL authentication using the **username** and **password** to set up your scan

Steps to register

To register a new Azure Database for PostgreSQL in Microsoft Purview Unified Catalog, do the following:

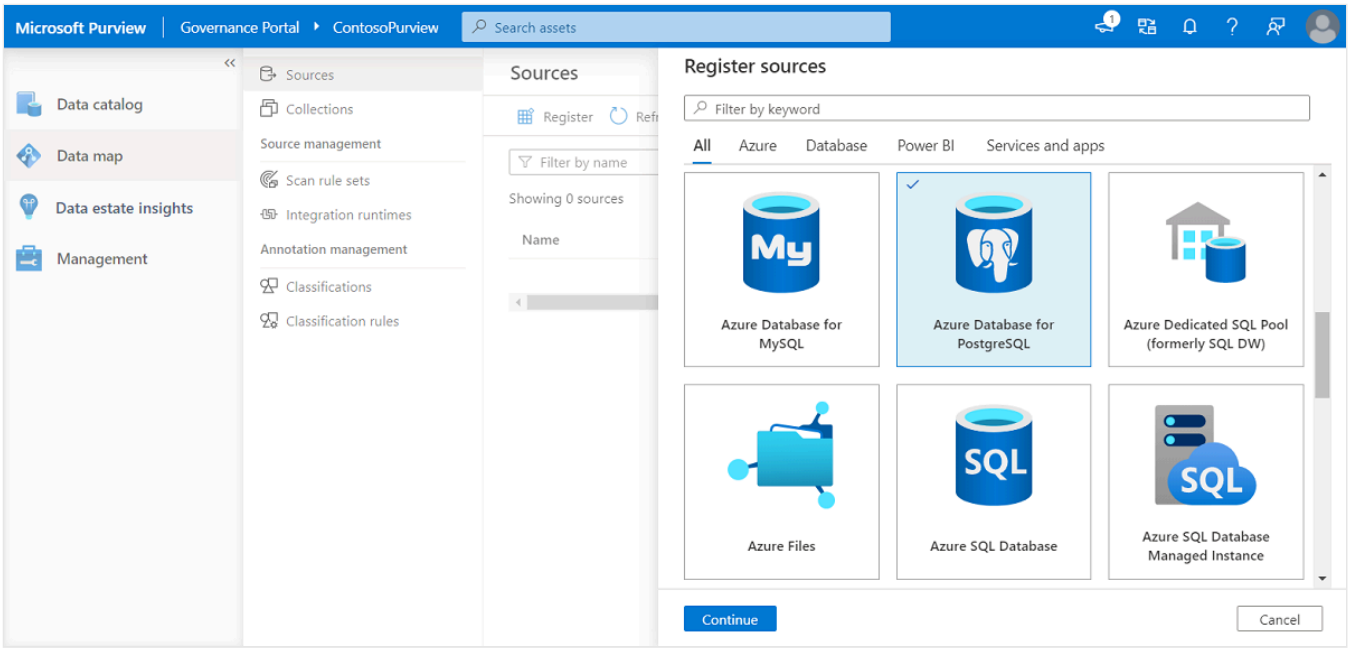
1. Open the Microsoft Purview governance portal by:

- Browsing directly to https://web.purview.azure.com and selecting your Microsoft Purview account.
- Opening the Azure portal, searching for and selecting the Microsoft Purview account. Selecting the **the Microsoft Purview governance portal** button.

2. Select **Data Map** on the left navigation.

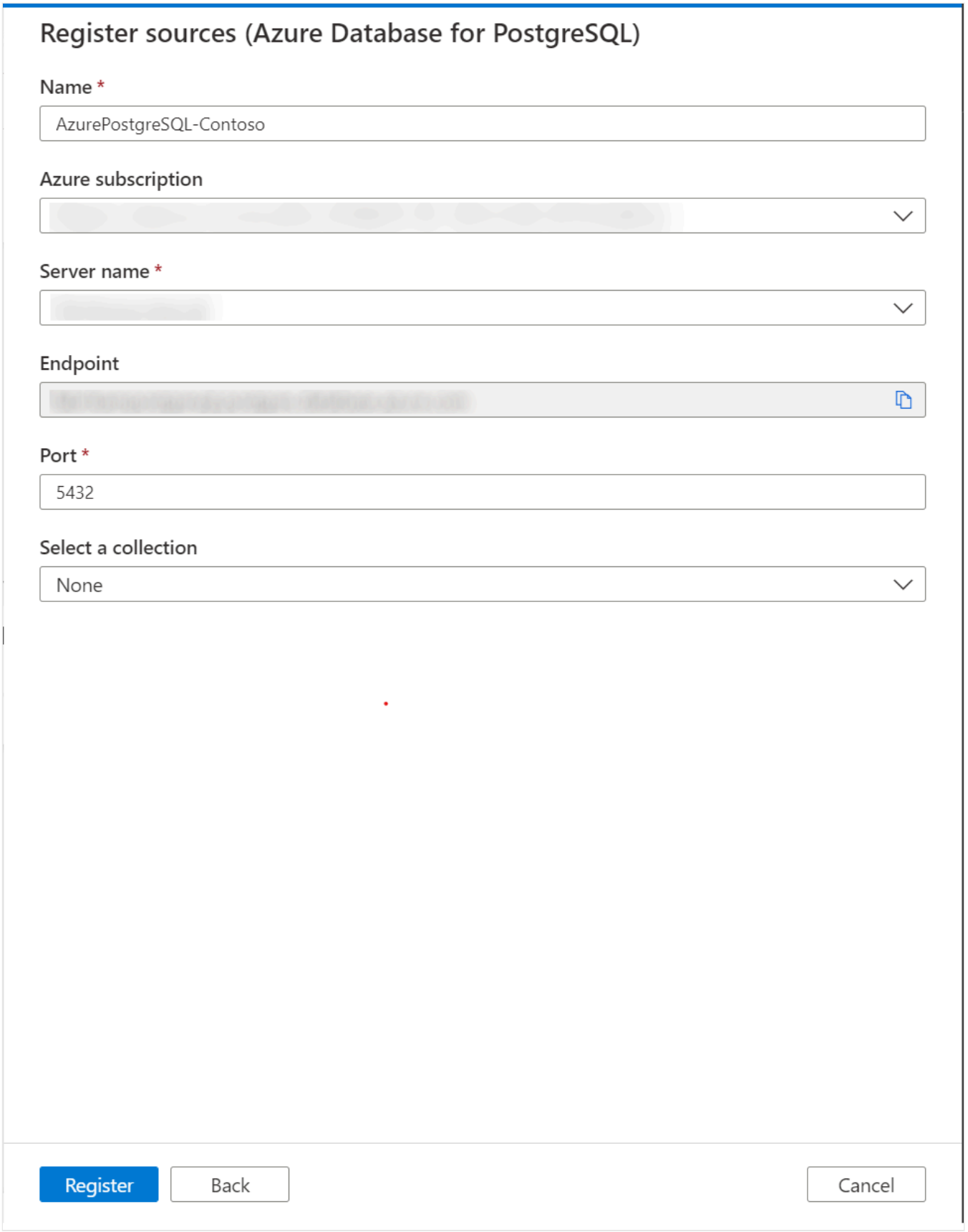
3. Select **Register**

4. On **Register sources**, select **Azure Database for PostgreSQL**. Select **Continue**.



On the **Register sources Azure Database for PostgreSQL** screen, do the following:

1. Enter a **Name** for your data source. This will be the display name for this data source in your Catalog.
2. Select **From Azure subscription**, select the appropriate subscription from the **Azure subscription** drop-down box and the appropriate server from the **Server name** drop-down box.
3. Select **Register** to register the data source.



Scan

Follow the steps below to scan an Azure Database for PostgreSQL database to automatically identify assets and classify your data. For more information about scanning in general, see our introduction to scans and ingestion

Create and run scan

To create and run a new scan, do the following:

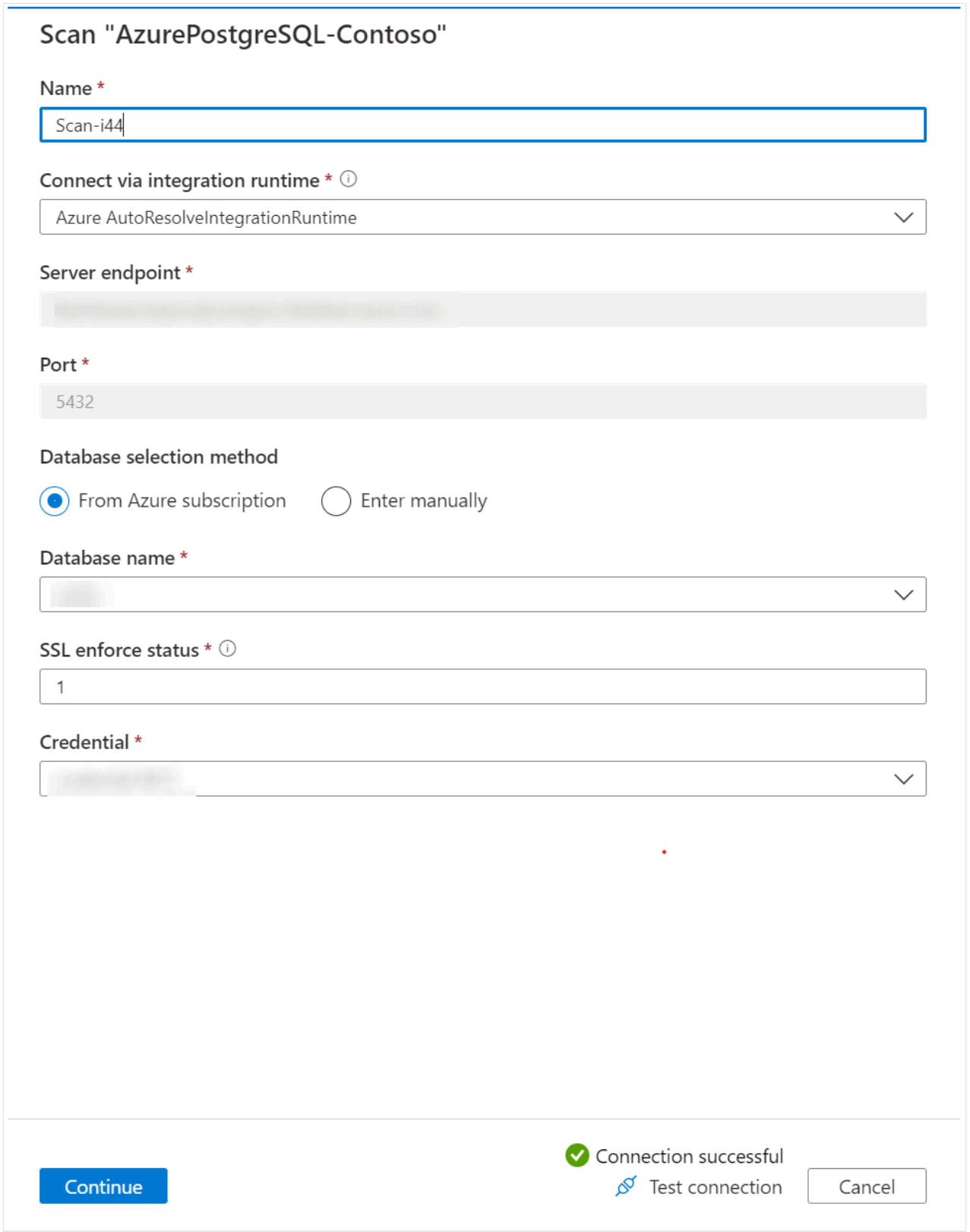
1. Select the **Data Map** tab on the left pane in the Microsoft Purview governance portal.

2. Select the Azure Database for PostgreSQL source that you registered.

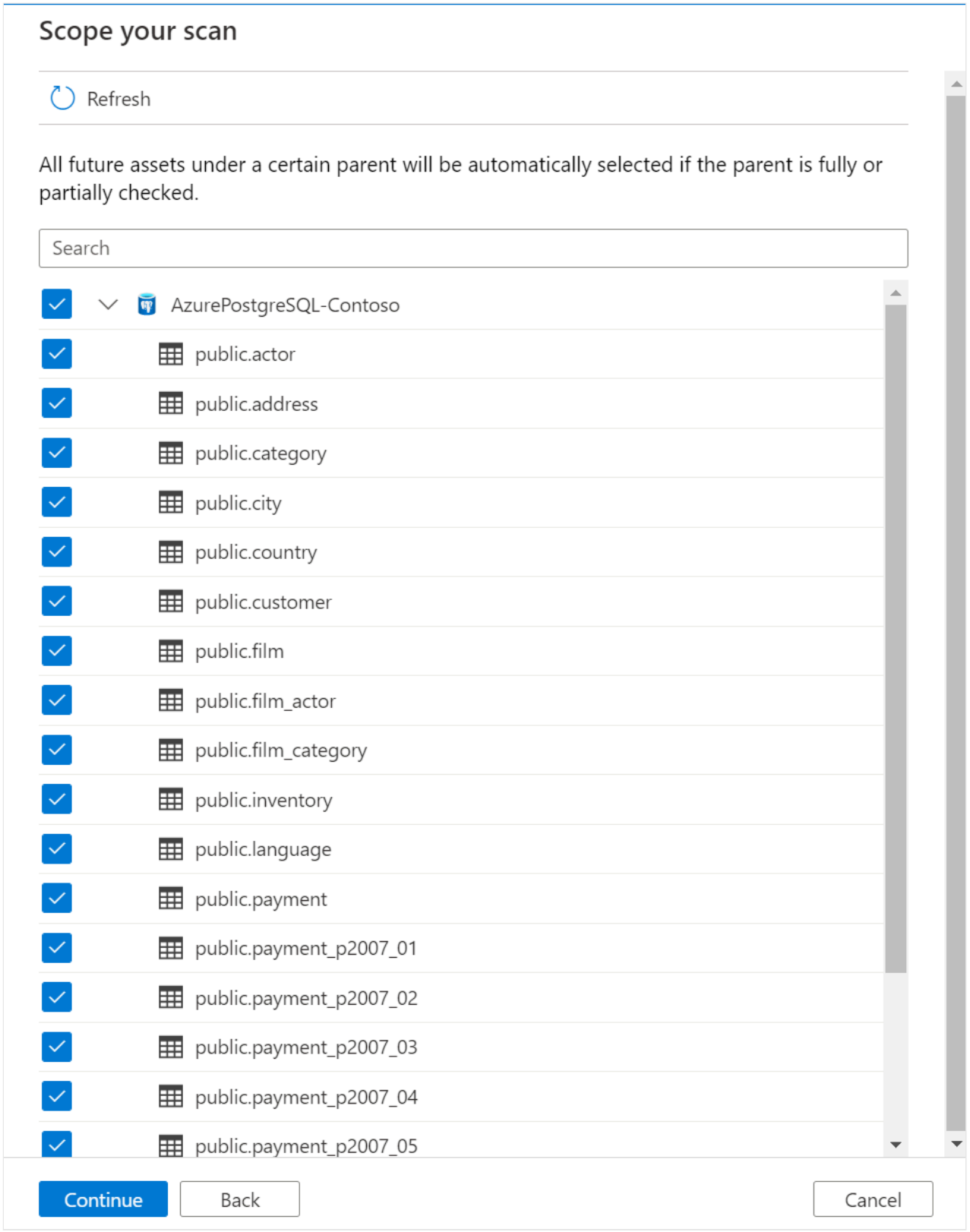
3. Select **New scan**

4. Choose either the Azure integration runtime if your source is publicly accessible, a managed virtual network integration runtime if using a managed virtual network, or a self-hosted integration runtime if your source is in a private virtual network. For more information about which integration runtime to use, see the choose the right integration runtime configuration article.

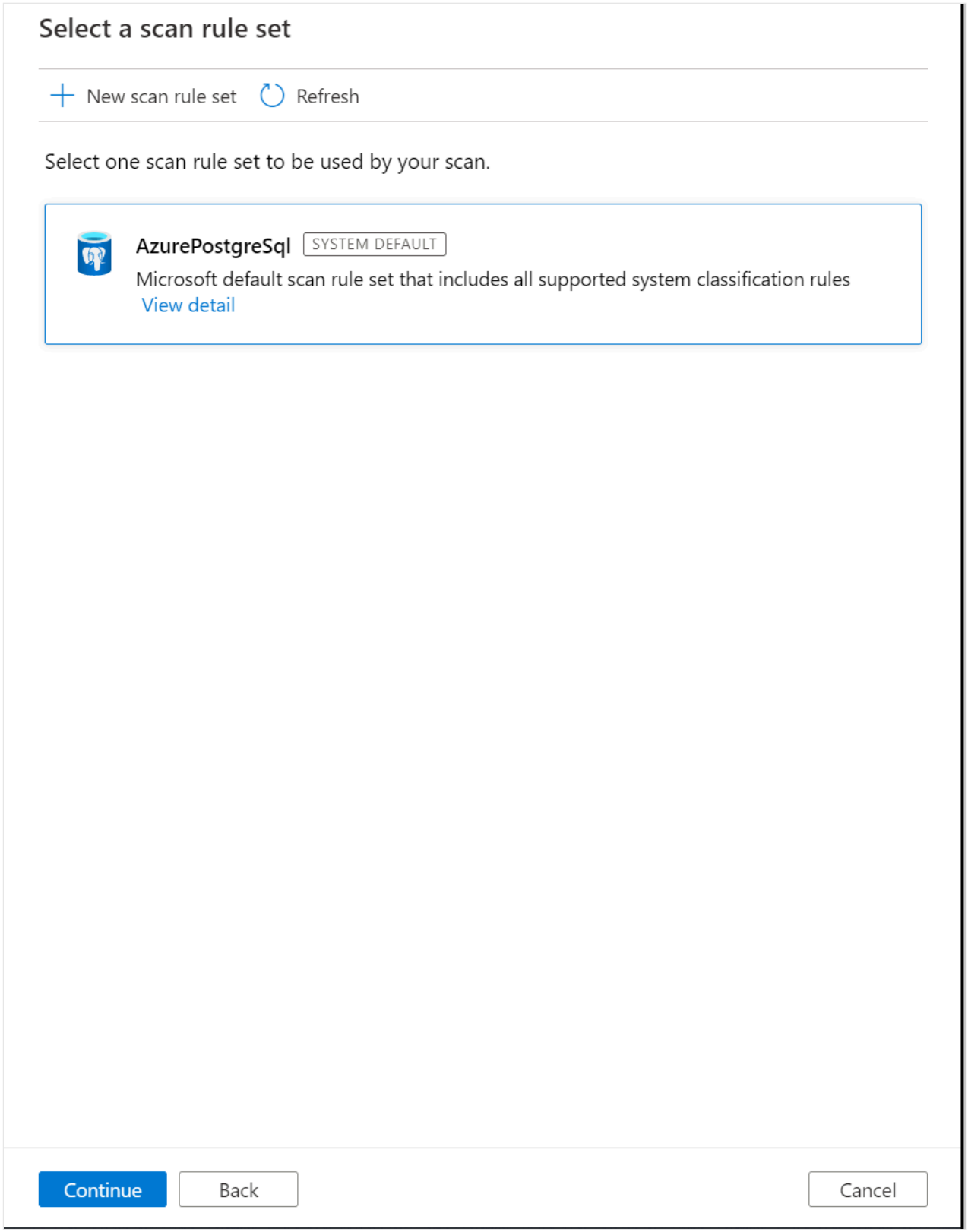
5. Select the credential to connect to your data source.



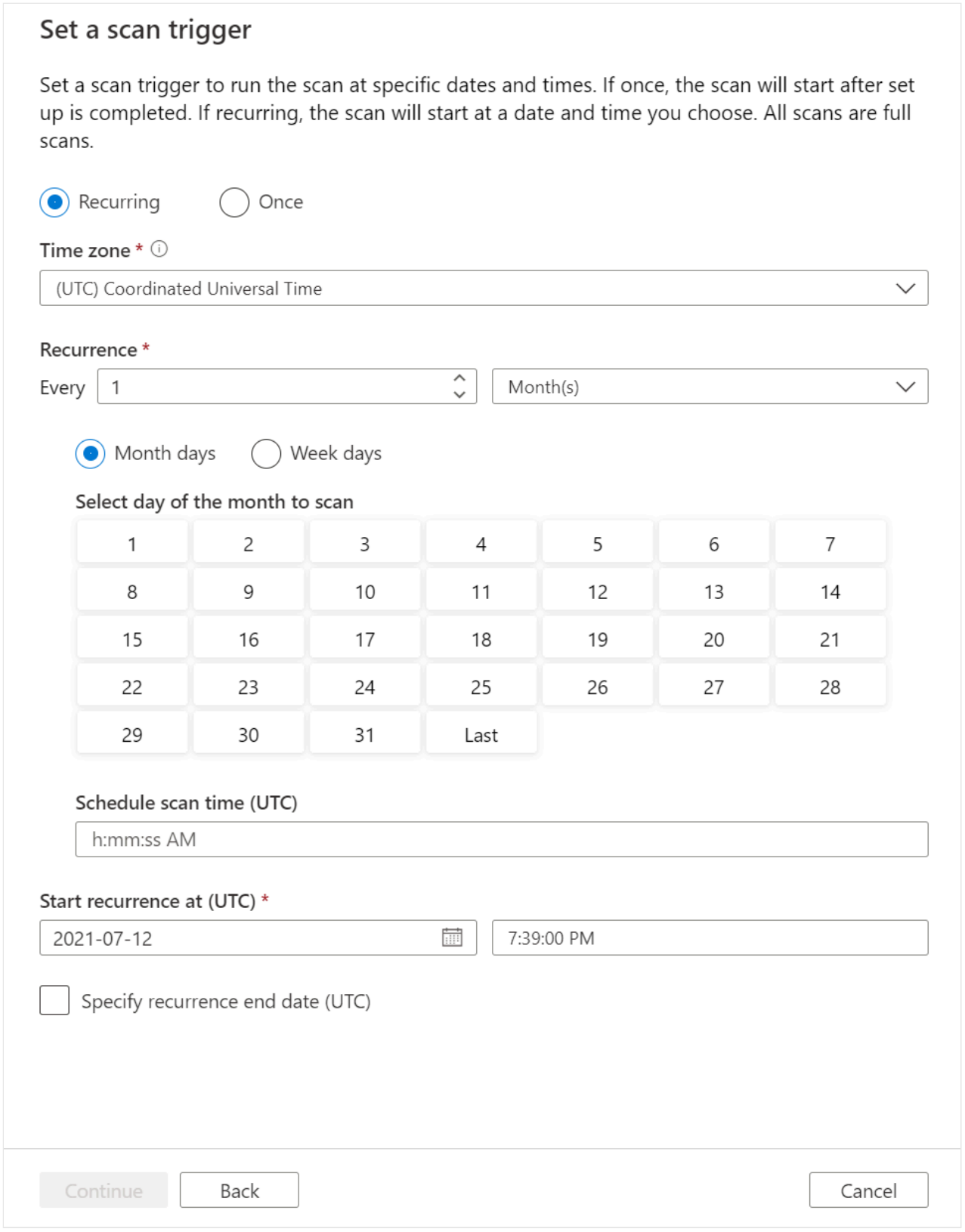
6. You can scope your scan to specific folders or subfolders by choosing the appropriate items in the list.



7. Then select a scan rule set. You can choose between the system default, existing custom rule sets, or create a new rule set inline.



8. Choose your scan trigger. You can set up a schedule or run the scan once.



9. Review your scan and select **Save and run**.

View your scans and scan runs

To view existing scans:

1. Go to the Microsoft Purview portal. On the left pane, select **Data map**.

2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that has results you want to view. The pane shows you all the previous scan runs, along with the status and metrics for each scan run.
4. Select the run ID to check the scan run details.

Manage your scans

To edit, cancel, or delete a scan:

1. Go to the Microsoft Purview portal. On the left pane, select **Data Map**.

2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.

3. Select the scan that you want to manage. You can then:

    - Edit the scan by selecting **Edit scan**.
    - Cancel an in-progress scan by selecting **Cancel scan run**.
    - Delete your scan by selecting **Delete scan**.

> ⓘ **Note**
>
> - Deleting your scan does not delete catalog assets created from previous scans.

Next steps

Now that you've registered your source, follow the below guides to learn more about Microsoft Purview and your data.

- Data Estate Insights in Microsoft Purview
- Lineage in Microsoft Purview
- Search Unified Catalog

Last updated on 03/16/2026

Connect to and manage Azure Databricks in Microsoft Purview

This article outlines how to register Azure Databricks, and how to authenticate and interact with Azure Databricks in Microsoft Purview. For more information about Microsoft Purview, read the introductory article.

Supported capabilities

Scanning capabilities

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan
Yes	Yes	No	Yes

> **Note**
>
> This connector brings metadata from Azure Databricks workspace-scoped Hive metastore. To scan metadata in Azure Databricks Unity Catalog, refer to Azure Databricks Unity Catalog connector.

When scanning Azure Databricks Hive metastore, Microsoft Purview supports:

- Extracting technical metadata, including:
  - Azure Databricks workspace
  - Hive server
  - Databases
  - Tables including the columns, foreign keys, unique constraints, and storage description
  - Views including the columns and storage description
- Fetching relationship between external tables and Azure Data Lake Storage Gen2/Azure Blob assets (external locations).
- Fetching static lineage between tables and views based on the view definition.

When setting up scan, you can choose to scan the entire Hive metastore, or scope the scan to a subset of schemas.

Compared to scanning via generic Hive Metastore connector if you use it to scan Azure Databricks earlier:

- You can directly set up scan for Azure Databricks workspaces without direct HMS access. It uses Databricks personal access token for authentication and connects to a cluster to perform scan.
- The Databricks workspace info is captured.
- The relationship between tables and storage assets is captured.

Other capabilities

For **classifications**, **sensitivity labels**, **policies**, **data lineage**, and **live view**, see the list of supported capabilities.

Known limitations

When you delete an object from the data source, the subsequent scan doesn't automatically remove the corresponding asset in Purview.

Prerequisites

- You must have an Azure account with an active subscription. Create an account for free.
- You must have an active Microsoft Purview account.
- You need an Azure Key Vault, and to grant Microsoft Purview permissions to access secrets.
- You need Data Source Administrator and Data Reader permissions to register a source and manage it in the Microsoft Purview governance portal. For more information about permissions, see Access control in Microsoft Purview.
- Set up the latest self-hosted integration runtime. For more information, see Create and configure a self-hosted integration runtime. The minimal supported self-hosted Integration Runtime version is 5.20.8227.2.
    - Ensure JDK 11 is installed on the machine where the self-hosted integration runtime is installed. Restart the machine after you newly install the JDK for it to take effect.
    - Ensure that Visual C++ Redistributable (version Visual Studio 2012 Update 4 or newer) is installed on the machine where the self-hosted integration runtime is running. If you don't have this update installed, download it now.

- In your Azure Databricks workspace:
  - Generate a personal access token, and store it as a secret in Azure Key Vault.
  - Create a cluster. Note the cluster ID - you can find it in Azure Databricks workspace -> Compute -> your cluster -> Tags -> Automatically added tags -> `ClusterId`.
  - Make sure the user has the following permissions so as to connect to the Azure Databricks cluster:
    - **Can Attach To** permission to connect to the running cluster.
    - **Can Restart** permission to automatically trigger the cluster to start if its state is terminated when connecting.

Register

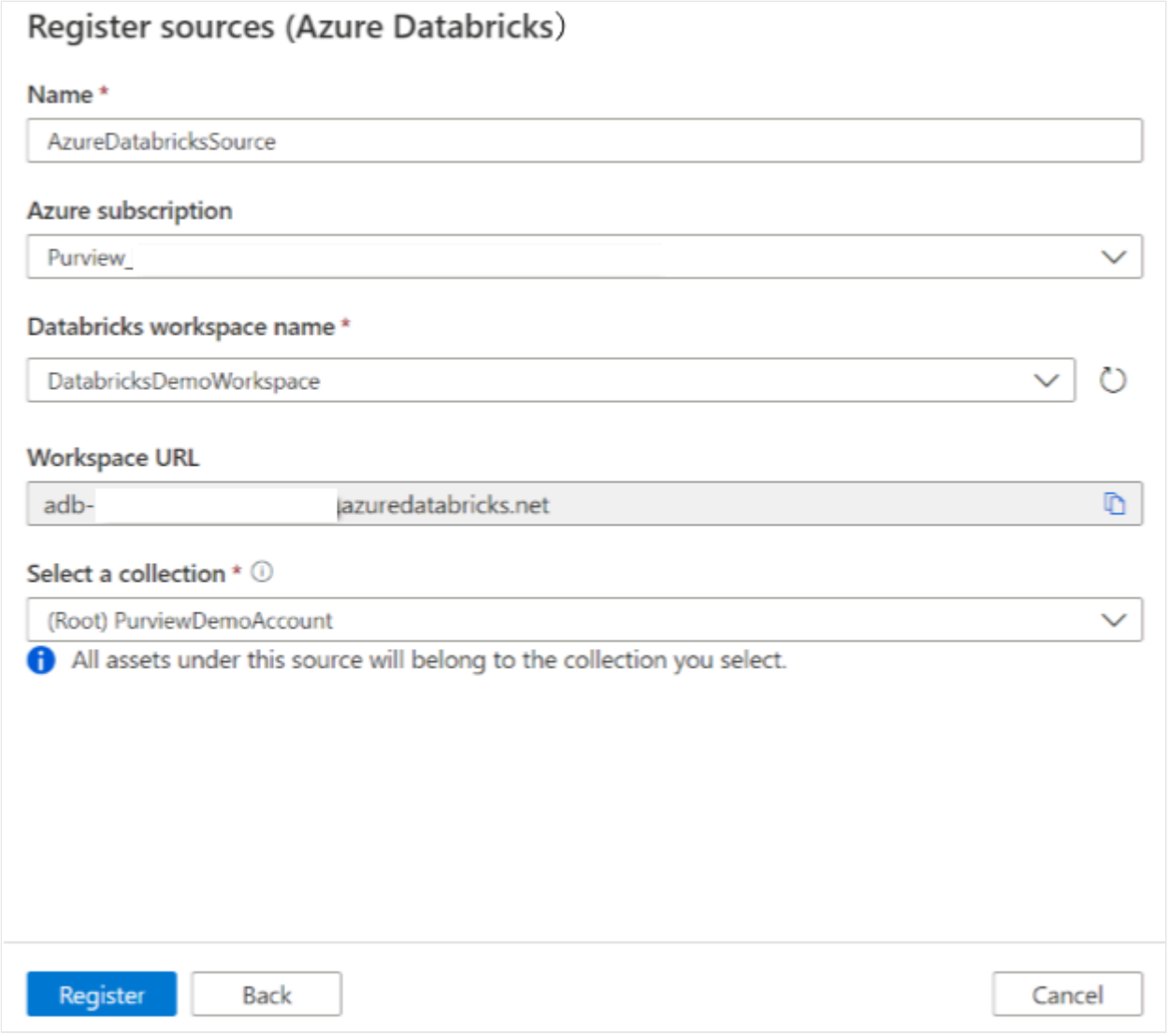
This section describes how to register an Azure Databricks workspace in Microsoft Purview by using the Microsoft Purview governance portal.

1. Go to your Microsoft Purview account.
2. Select **Data Map** on the left pane.
3. Select **Register**.
4. In **Register sources**, select **Azure Databricks** > **Continue**.
5. On the **Register sources (Azure Databricks)** screen, complete the following steps:

   a. For **Name**, enter a name that Microsoft Purview lists as the data source.

   b. For **Azure subscription** and **Databricks workspace name**, select the subscription and workspace that you want to scan from the dropdown. The Databricks workspace URL is automatically populated.

   c. Select a collection from the list.



6. Select **Finish**.

Scan

> **Tip**
>
> To troubleshoot any issues with scanning:
>
> 1. Confirm you followed all prerequisites.
> 2. Review the scan troubleshooting documentation.

Use the following steps to scan Azure Databricks to automatically identify assets. For more information about scanning in general, see Scans and ingestion in Microsoft Purview.

1. In the Management Center, select integration runtimes. Make sure that a self-hosted integration runtime is set up. If it isn't set up, use the steps in Create and manage a self-hosted integration runtime.

2. Go to **Sources**.

3. Select the registered Azure Databricks.

4. Select + **New scan**.

5. Provide the following details:

   a. **Name**: Enter a name for the scan.

   b. **Extraction method:** Indicate to extract metadata from Hive Metastore or Unity Catalog. Select **Hive Metastore**.

   c. **Connect via integration runtime**: Select the configured self-hosted integration runtime.

   d. **Credential**: Select the credential to connect to your data source. Make sure to:

      - Select **Access Token Authentication** while creating a credential.
      - Provide secret name of the personal access token that you created in Prerequisites in the appropriate box.

      For more information, see Credentials for source authentication in Microsoft Purview.

   e. **Cluster ID**: Specify the cluster ID that Microsoft Purview connects to and powers the scan. You can find it in Azure Databricks workspace -> Compute -> your cluster -> Tags -> Automatically added tags -> `ClusterId`.

      i. **Mount points**: Provide the mount point and Azure Storage source location string when you have external storage manually mounted to Databricks. Use the format `/mnt/<path>=abfss://<container>@<adls_gen2_storage_account>.dfs.core.windows.net/;/mnt/<path>=wasbs://<container>@<blob_storage_account>.blob.core.windows.net`. It's used to capture the relationship between tables and the corresponding storage assets in Microsoft Purview. This setting is optional. If you don't specify it, the relationship isn't retrieved.

   You can get the list of mount points in your Databricks workspace by running the following Python command in a notebook:

   ```
   dbutils.fs.mounts()
   ```

   It prints all the mount points like below:

   ```
   ```

```
[MountInfo(mountPoint='/databricks-datasets', source='databricks-datasets',
encryptionType=''),
MountInfo(mountPoint='/mnt/ADLS2',
source='abfss://samplelocation1@azurestorage1.dfs.core.windows.net/',
encryptionType=''),
MountInfo(mountPoint='/databricks/mlflow-tracking',
source='databricks/mlflow-tracking', encryptionType=''),
MountInfo(mountPoint='/mnt/Blob',
source='wasbs://samplelocation2@azurestorage2.blob.core.windows.net',
encryptionType=''),
MountInfo(mountPoint='/databricks-results', source='databricks-results',
encryptionType=''),
MountInfo(mountPoint='/databricks/mlflow-registry',
source='databricks/mlflow-registry', encryptionType=''),
MountInfo(mountPoint='/', source='DatabricksRoot', encryptionType='')]
```

In this example, specify the following as mount points:

`/mnt/ADLS2=abfss://samplelocation1@azurestorage1.dfs.core.windows.net/;/mnt/Blob=wasbs://samplelocation2@azurestorage2.blob.core.windows.net`

f. **Schema**: The subset of schemas to import expressed as a semicolon separated list of schemas. For example, `schema1;schema2`. All user schemas are imported if that list is empty. All system schemas and objects are ignored by default.

Acceptable schema name patterns can be static names or contain wildcard %. For example: `A%;%B;%C%;D`

- Start with A or
- End with B or
- Contain C or
- Equal D

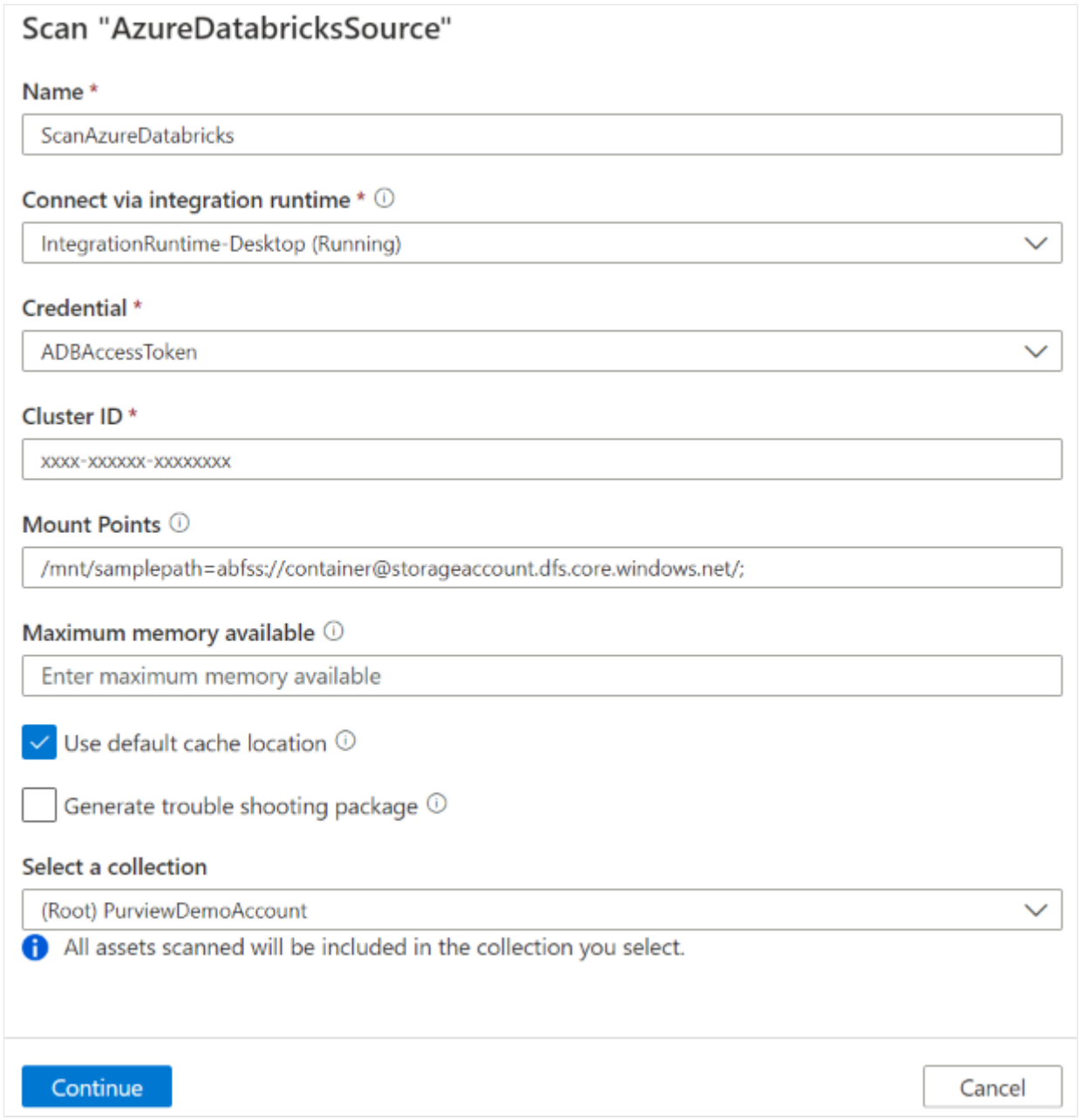
Usage of NOT and special characters aren't acceptable.

> ⓘ **Note**
>
> This schema filter is supported on Self-hosted Integration Runtime version 5.32.8597.1 and above.

g. **Maximum memory available**: Maximum memory (in gigabytes) available on the customer's machine for the scanning processes to use. This value is dependent on the size of Azure Databricks to be scanned.

> ⓘ **Note**

As a thumb rule, please provide 1GB memory for every 1000 tables.



6. Select **Continue**.

7. For **Scan trigger**, choose whether to set up a schedule or run the scan once.

8. Review your scan and select **Save and Run**.

Once the scan successfully completes, see how to browse and search Azure Databricks assets.

View your scans and scan runs

To view existing scans:

1. Go to the Microsoft Purview portal. On the left pane, select **Data map**.

2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that has results you want to view. The pane shows you all the previous scan runs, along with the status and metrics for each scan run.
4. Select the run ID to check the scan run details.

Manage your scans

To edit, cancel, or delete a scan:

1. Go to the Microsoft Purview portal. On the left pane, select **Data Map**.

2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.

3. Select the scan that you want to manage. You can then:

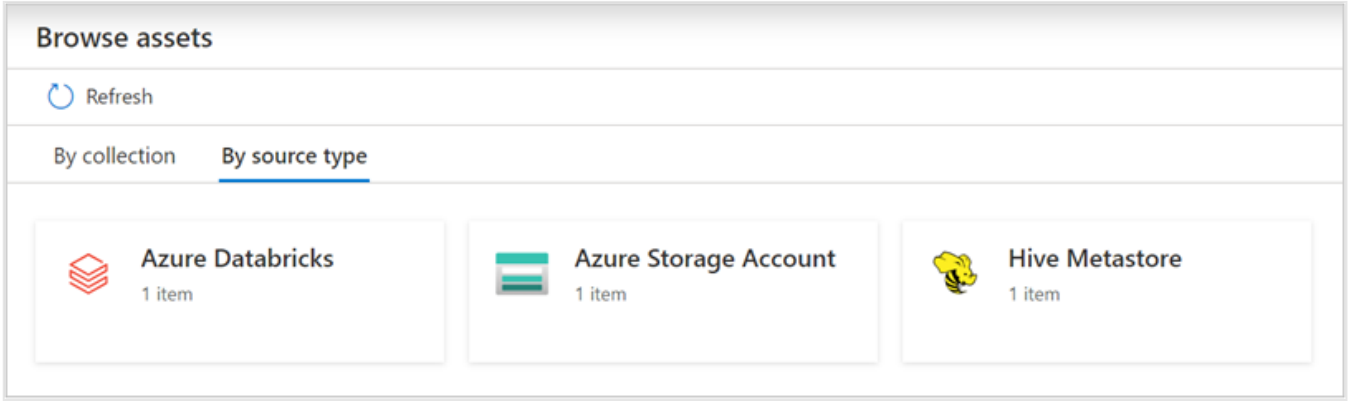
    - Edit the scan by selecting **Edit scan**.
    - Cancel an in-progress scan by selecting **Cancel scan run**.
    - Delete your scan by selecting **Delete scan**.

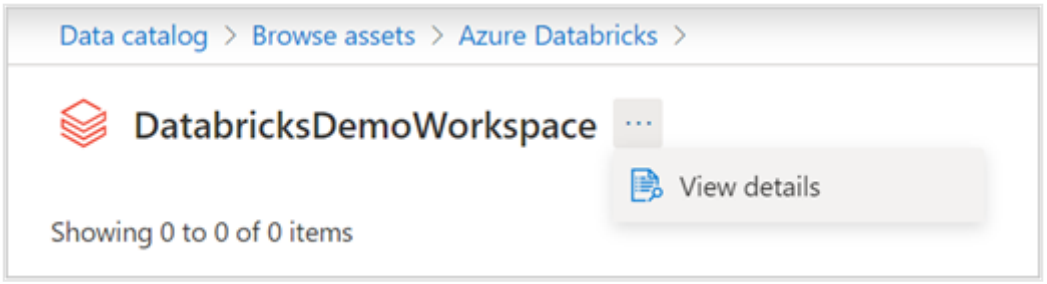
> ⓘ Note
>
> - Deleting your scan does not delete catalog assets created from previous scans.

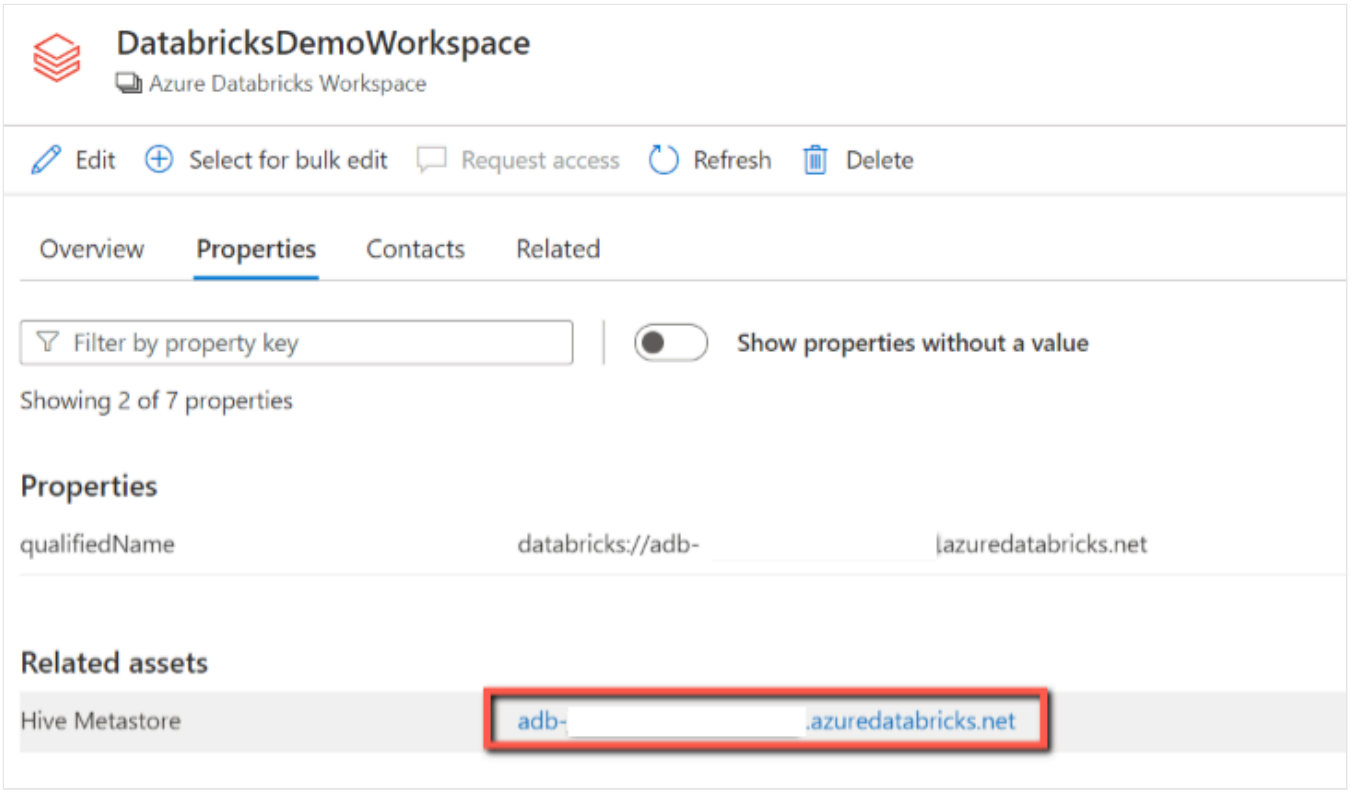
Browse and search assets

After scanning your Azure Databricks, you can browse Unified Catalog or search Unified Catalog to view the asset details.

From the Databricks workspace asset, you can find the associated Hive Metastore and the tables/views, reversed applies too.



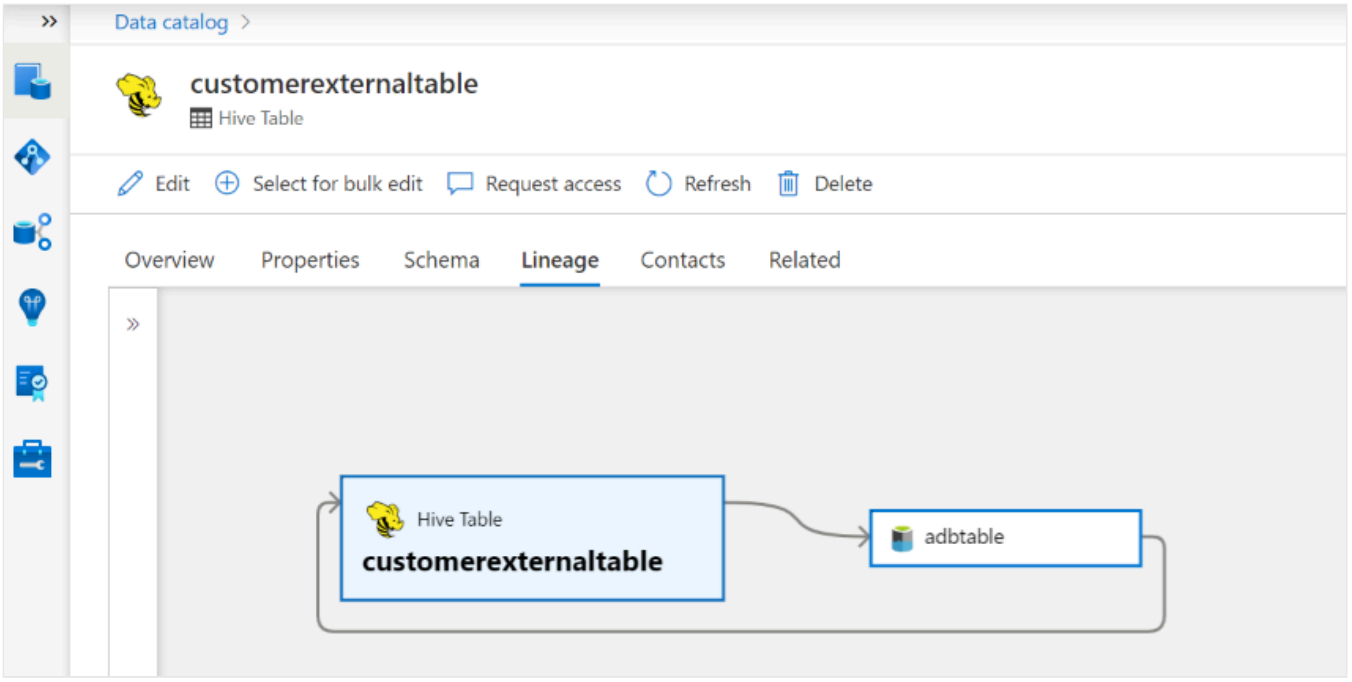




Lineage

Refer to the supported capabilities section on the supported Azure Databricks scenarios. For more information about lineage in general, see data lineage and lineage user guide.

Go to the Hive table/view asset -> lineage tab, you can see the asset relationship when applicable. For relationship between table and external storage assets, you see Hive table asset and the storage asset are directly connected bi-directionally, as they mutually impact each other. If you use mount point in create table statement, you need to provide the mount point information in scan settings to extract such relationship.



Next steps

Now that you've registered your source, use the following guides to learn more about Microsoft Purview and your data:

- Data Estate Insights in Microsoft Purview
- Lineage in Microsoft Purview
- Search the Unified Catalog

Last updated on 03/31/2026

Connect to and manage Azure Databricks Unity Catalog in Microsoft Purview

This article outlines how to register Azure Databricks, and how to authenticate and interact with Azure Databricks Unity Catalog in Microsoft Purview. For more information about Microsoft Purview, see the introductory article.

Supported capabilities

Scanning capabilities

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan
Yes	Yes	Yes	No

When scanning Azure Databricks Unity Catalog, Microsoft Purview supports:

- Extracting technical metadata, including:
  - Metastore
  - Catalogs
  - Schemas
  - Tables, including the columns. Purview supports external tables metadata, but lineage isn't supported.
  - Views, including the columns
  - Tags from Unity Catalog (tags that represent labels, classification equivalent in Unity Catalog are scanned in Microsoft Purview in properties)
- Fetching lineage on assets relationships between tables, views, and columns during notebook runs.

When setting up a scan, you can choose to scan the entire Unity Catalog, or scope the scan to a subset of catalogs.

Other capabilities

For **classifications**, **sensitivity labels**, **policies**, **data lineage**, and **live view**, see the list of supported capabilities.

> **Note**
>
> This connector brings metadata from Azure Databricks Unity Catalog. To scan Azure Databricks workspace-scoped metadata, see Azure Databricks Hive Metastore connector.

Known limitations

- In Microsoft Purview, Databricks notebook names appear as numeric IDs instead of readable names. This limitation exists because Databricks doesn't expose notebook names in the Unity Catalog system table.
- You might encounter errors if scan results from Azure Databricks exceed 1 MB and Azure Databricks-managed blob storage denies public network access. To prevent this problem, ensure that Microsoft Purview has access to the internal DBFS storage location of the Azure Databricks workspace being scanned. To learn more, see Cloud fetch in JDBC.
- Incremental scan is available only for the *Azure Databricks Unity Catalog* data source.
- Scoped scan is available only for the Unity Catalog option under *Azure Databricks* data source.
- You can add managed private endpoints only for the Unity Catalog option under *Azure Databricks* data source.
- When you delete an object from the data source, the subsequent scan doesn't automatically remove the corresponding asset in Microsoft Purview.
- Lineage information isn't available in Azure Databricks workspaces in the China region. This limitation exists because Azure Databricks system tables aren't supported in this region. Microsoft Purview uses these tables to extract lineage, so it can't retrieve lineage in this region.
- Set the Databricks table column comment to an empty string if you don't want the column description displayed in Microsoft Purview.
- Scan of external tables or dynamic tables isn't supported.
- Tags aren't currently supported for scans done using a Kubernetes SHIR option.

For more information about other limitations related to native Azure Databricks lineage, see Azure Databricks documentation.

Prerequisites

- You must have an Azure account with an active subscription. Create an account for free.
- You must have an active Microsoft Purview account.
- You need an Azure Key Vault, and to grant Microsoft Purview permissions to access secrets.
- You need Data Source Administrator and Data Reader permissions to register a source and manage it in the classic Microsoft Purview governance portal. For more information about permissions, see Access control in Microsoft Purview.
- To scan Azure Databricks Unity Catalog, Microsoft Purview connects to a SQL Warehouse in your workspace, and uses a Personal Access Token for authentication. You need to have an Azure Databricks workspace that is Unity Catalog enabled and attached to the metastore you want to scan. In your Azure Databricks workspace:
    - Create a SQL Warehouse. You can use the autocreated Starter warehouse as well if applicable.
    - Note down the HTTP path. You can find it in Azure Databricks workspace > SQL Warehouses > your warehouse > Connection details > HTTP path.
    - Make sure the user has the **Can Use** permission so they can connect to the Azure Databricks SQL warehouse. Learn more at SQL warehouse access control.
    - The SQL warehouse in Azure Databricks must be running to establish a connection with Microsoft Purview before setting up the scan in Microsoft Purview.
- To fetch lineage from Azure Databricks using Microsoft Purview, the following prerequisites must be in place:
    - Enable the system schema: The system schema **system.access** must be enabled in your Unity Catalog. This requirement exists because lineage information is stored in system tables, and enabling this schema allows access to those tables. Learn more about monitoring usage with system tables.
    - User privileges:
        - The user account you use for scanning needs to have SELECT privileges on the following system tables:
            - system.access.table_lineage

    - system.access.column_lineage

    These permissions are required because lineage data is read directly from the system tables, and without the necessary access, Microsoft Purview can't retrieve the lineage information.

    The user account you use for scanning also needs to have these permissions:
    - USE CATALOG on system
    - USE SCHEMA on system.access

- If your Azure Databricks workspace doesn't allow access from public network or if your Microsoft Purview account doesn't enable access from all networks, you can use the Managed Virtual Network Integration Runtime or a kubernetes supported self-hosted integration runtime to scan. You can set up a managed private endpoint for Azure Databricks as needed to establish private connectivity.

Data source setup (Azure Databricks vs. Azure Databricks Unity Catalog)

You can set up an Azure Databricks Unity Catalog connection to Microsoft Purview through two sources:

- Azure Databricks (old source)
- Azure Databricks Unity Catalog

Differences in capabilities between data sources

The capabilities of the Azure Databricks Unity Catalog connector differ based on the source used for the connection, as outlined in the following table:

Expand table

Capability	Old source	New source
Source name	Azure Databricks	Azure Databricks Unity Catalog
Source support	Hive & Unity Catalog	Unity Catalog
Authentication: Personal Access Token	Supported	Supported
Authentication: Service Principal	No	Supported

Capability	Old source	New source
Authentication: Managed Identity	No	Supported (system managed identity)
Integration Runtime: Azure IR	Supported	Supported
Integration Runtime: Managed Virtual Network IR	Supported	Yes**
Integration Runtime: Kubernetes Self-Hosted IR	Supported	Supported
Scoped scan	Yes - catalog level	No
Incremental scan	No	Yes
Lineage	Supported	Supported

**Scanning Azure Databricks Unity Catalog is supported through Managed Virtual Network Integration Runtime. However, Managed Private Endpoint isn't supported in this case.

Which source should I use?

Choose either source based on your organization's needs, or use both in parallel. Switching between sources or using both simultaneously **doesn't cause asset duplication** in Microsoft Purview. This flexibility allows you to start with the option that best fits your current setup and adjust over time as needed.

Authentication for a scan

To scan Azure Databricks Unity Catalog, use personal access tokens, managed identities, or service principal authentication methods.

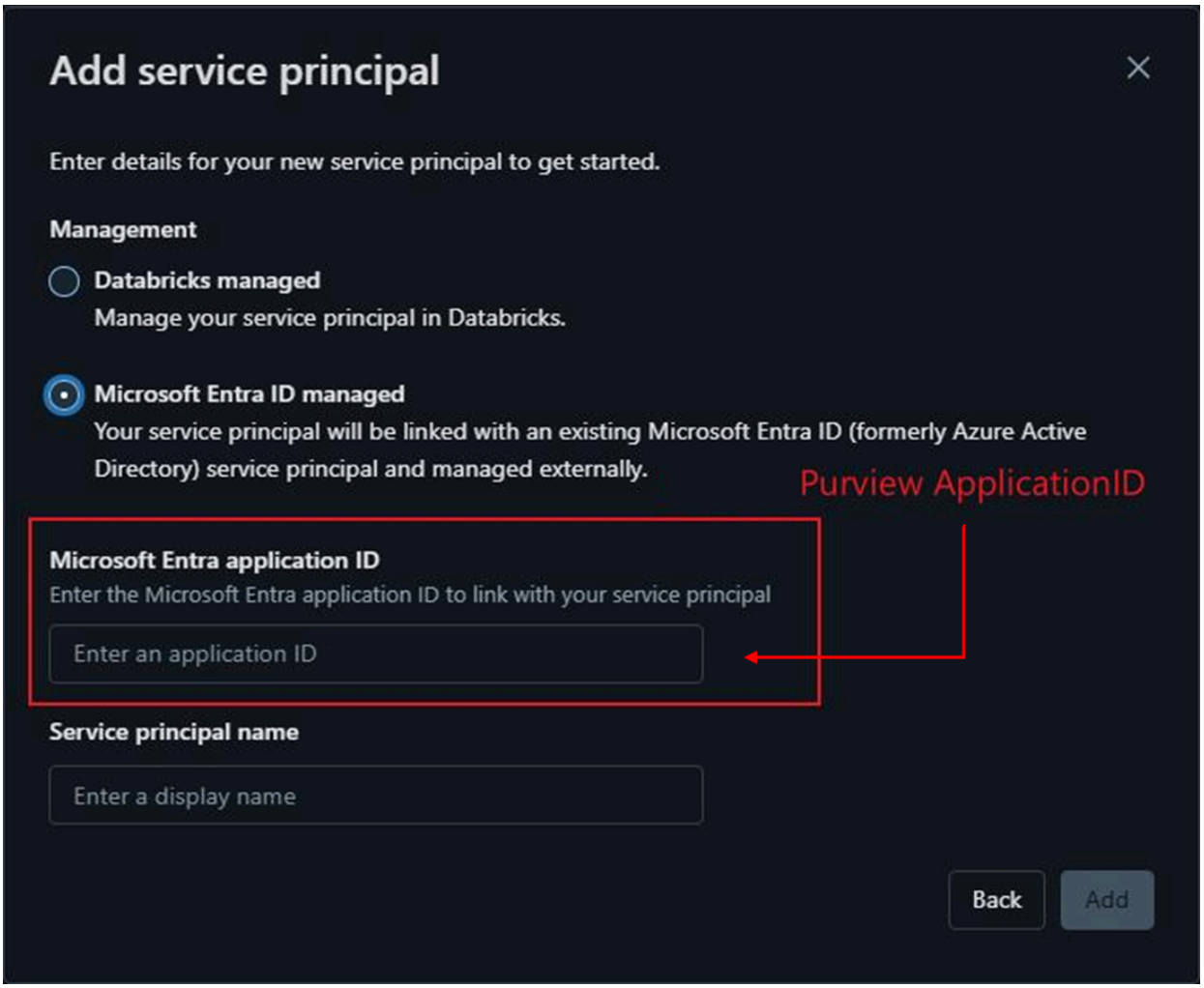
System assigned managed identity

If you use a system assigned managed identity

In Azure Databricks

1. Go to **Settings** > **Workspace admin** > **Identity and access**.
2. Select **Add service principal**.
3. Select **Add new**.

4. Select **Microsoft Entra ID managed**. Enter the **Application ID** for Microsoft Purview. You can find the Application ID under Microsoft Purview account resource details in the Azure portal.



In Microsoft Purview

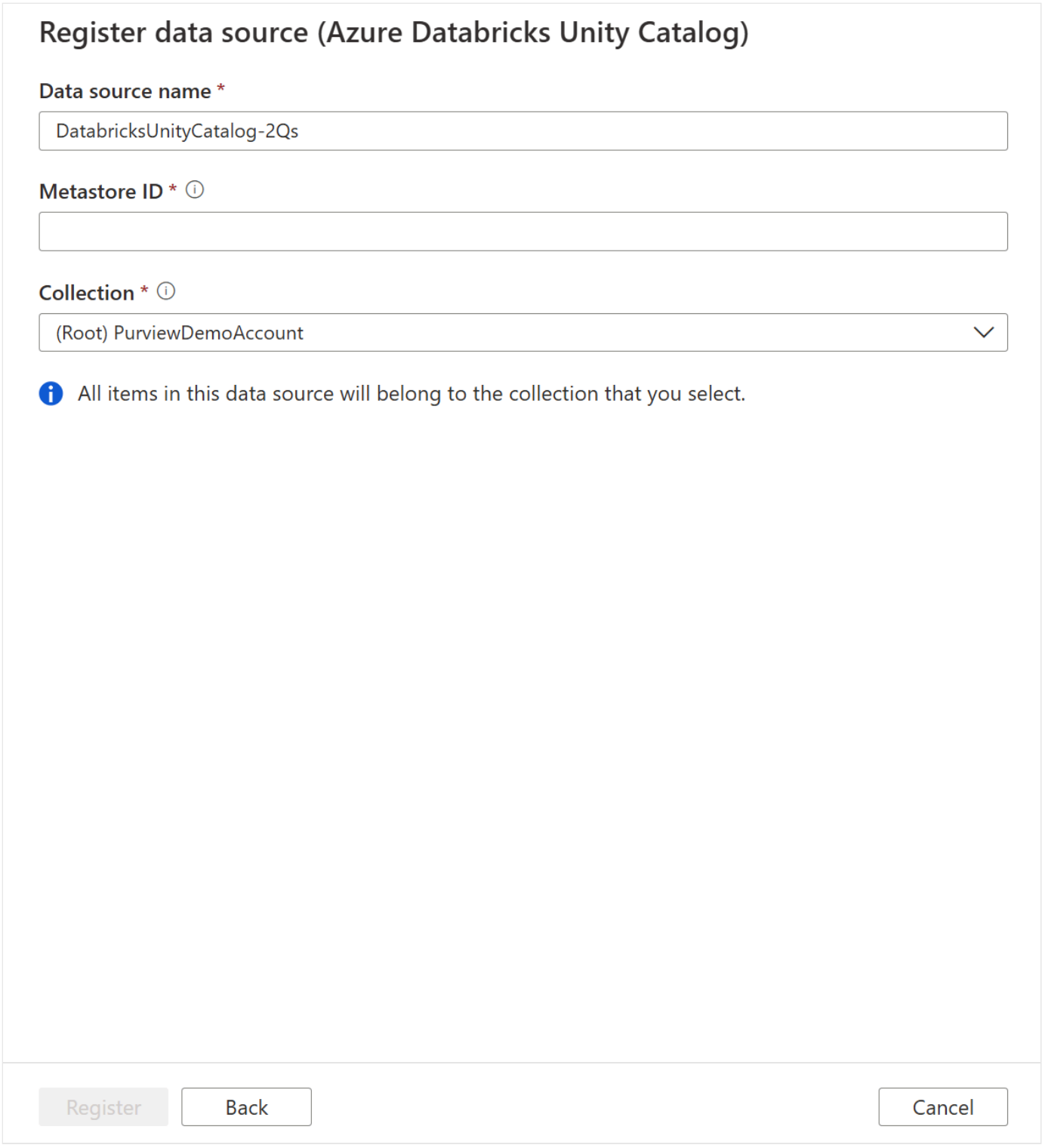
1. Select the system-assigned managed identity under **Credential**.

- For all the objects that you want to bring into Microsoft Purview, the user or service principal needs to have at least **SELECT** privilege on tables/views, **USE CATALOG** on the object's catalog, and **USE SCHEMA** on the object's schema.
    - To scan all the objects in a Unity Catalog metastore, use a user or service principal with metastore admin role. For more information, see Manage privileges in Unity Catalog and Unity Catalog privileges and securable objects.
        - For classification, the user also needs to have **SELECT** privilege on the tables/views to retrieve sample data.

Register

This section describes how to register an Azure Databricks workspace in Microsoft Purview by using the classic Microsoft Purview governance portal.

1. Go to your Microsoft Purview account.
2. Select **Data Map** on the left pane.
3. Select **Register**.
4. In **Register sources**, select **Azure Databricks Unity Catalog** > **Continue**.
5. On the **Register sources (Azure Databricks Unity Catalog)** screen, complete the following steps:
6. For **Name**, enter a name that Microsoft Purview will list as the data source.
    a. For **Metastore ID**, enter the metastore ID for the Azure Databricks Unity Catalog metastore that you want to scan.
    b. Select a collection from the list.



7. Select **Finish**.

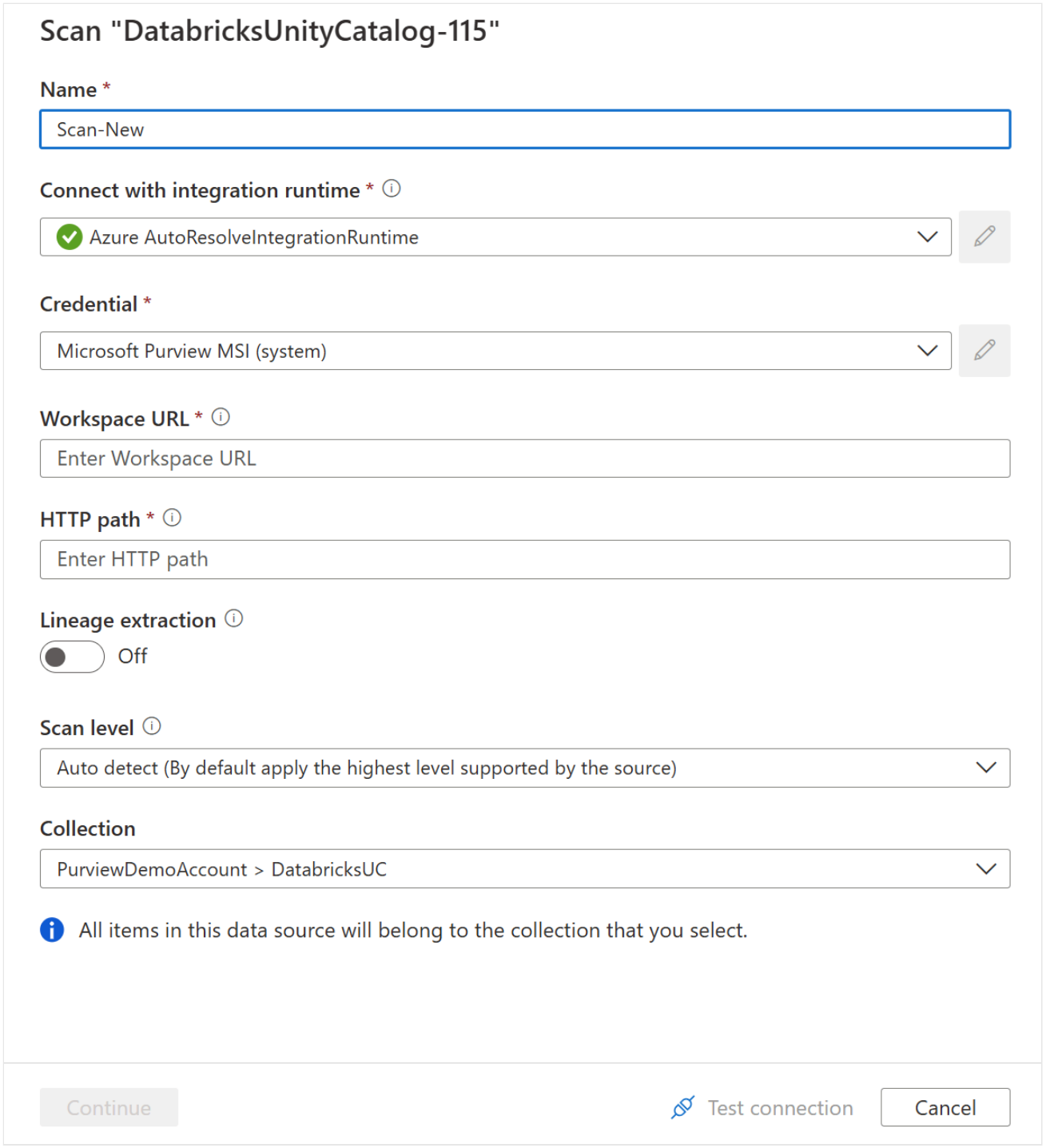
Scan

> **Tip**
>
> To troubleshoot any problems with scanning:
>
> 1. Confirm you meet all prerequisites.

2. Review the **scan troubleshooting documentation**.

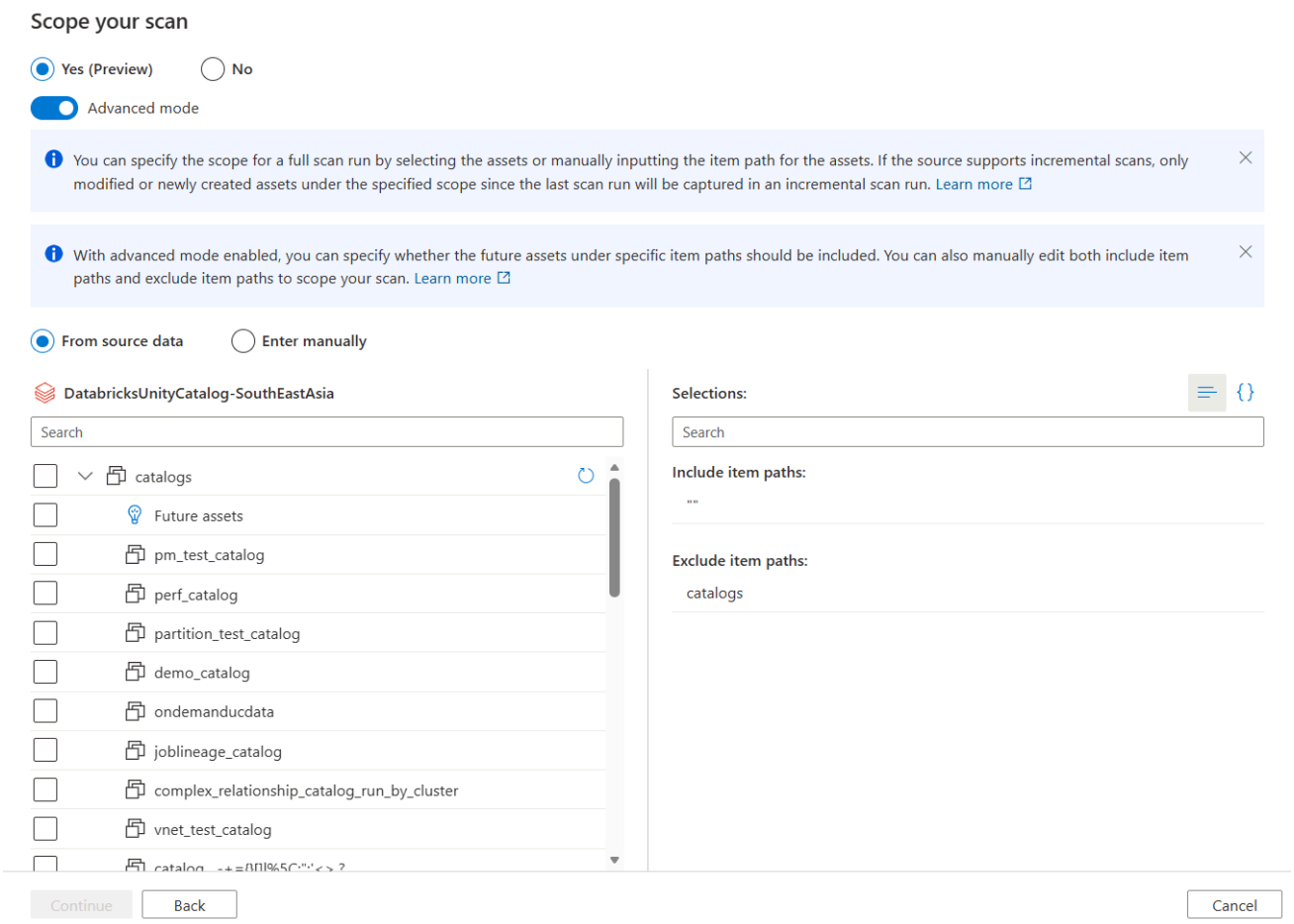
Use the following steps to scan Azure Databricks and automatically identify assets. For more information about scanning, see Scans and ingestion in Microsoft Purview.

1. Go to **Sources**.
2. Select the registered Azure Databricks.
3. Select **+ New scan**.
4. Provide the following details:
    a. **Name**: Enter a name for the scan.
    b. **Connect via integration runtime**: Choose the default Azure integration runtime, Managed Virtual Network IR, or a Kubernetes supported self-hosted integration runtime you created.
    c. **Credential**: Select the credential to connect to your data source. Make sure to:
        - Select **Access Token**, **Managed Identity**, or **Service Principal**.
        - You can create a new Access Token or Service Principal credential while registering a scan. For more information, see Credentials for source authentication in Microsoft Purview.
    d. **Workspace URL:** Provide the URL for the workspace that you want to scan.
        i. **HTTP path:** Specify the Databricks SQL Warehouse's HTTP path that Microsoft Purview connects to and performs the scan; for example, `/sql/1.0/endpoints/xxxxxxxxxxxxxxxx`. You can find it in Azure Databricks workspace -> SQL Warehouses -> your warehouse -> Connection details -> HTTP path.
    e. **Lineage extraction:** Toggle lineage extraction to **On** to fetch lineage of the scanned assets.
5. Select **Test connection** to validate the settings.



6. Select **Continue**.

7. On the **Scope your scan** page, select one or more catalogs on which you want to limit your scan. Use **Advanced mode** (preview) to select specific item paths to include or exclude from the scan.



8. For **Scan trigger**, choose whether to set up a schedule or run the scan once.

9. Review your scan and select **Save and Run**.

After the scan finishes successfully, see how to browse and search assets.

View your scans and scan runs

To view existing scans:

1. Go to the Microsoft Purview portal. On the left pane, select **Data map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that has results you want to view. The pane shows you all the previous scan runs, along with the status and metrics for each scan run.
4. Select the run ID to check the scan run details.

Manage your scans

To edit, cancel, or delete a scan:

1. Go to the Microsoft Purview portal. On the left pane, select **Data Map**.

2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.

3. Select the scan that you want to manage. You can then:

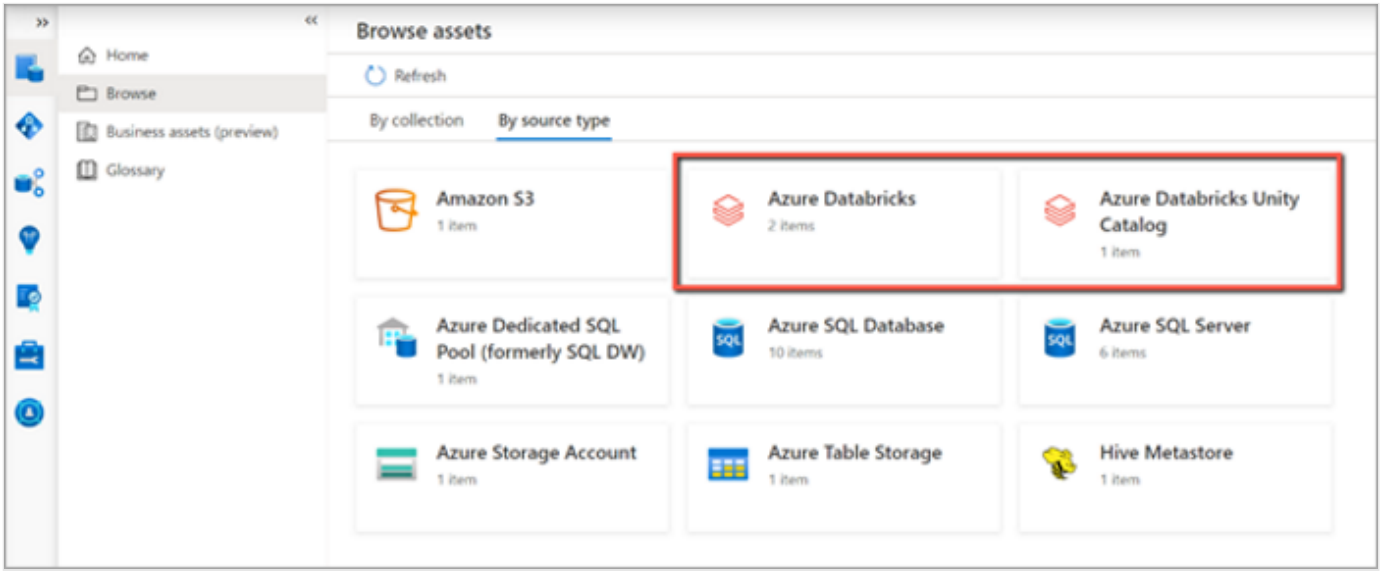
    - Edit the scan by selecting **Edit scan**.
    - Cancel an in-progress scan by selecting **Cancel scan run**.
    - Delete your scan by selecting **Delete scan**.

> ⓘ Note
>
> - Deleting your scan does not delete catalog assets created from previous scans.

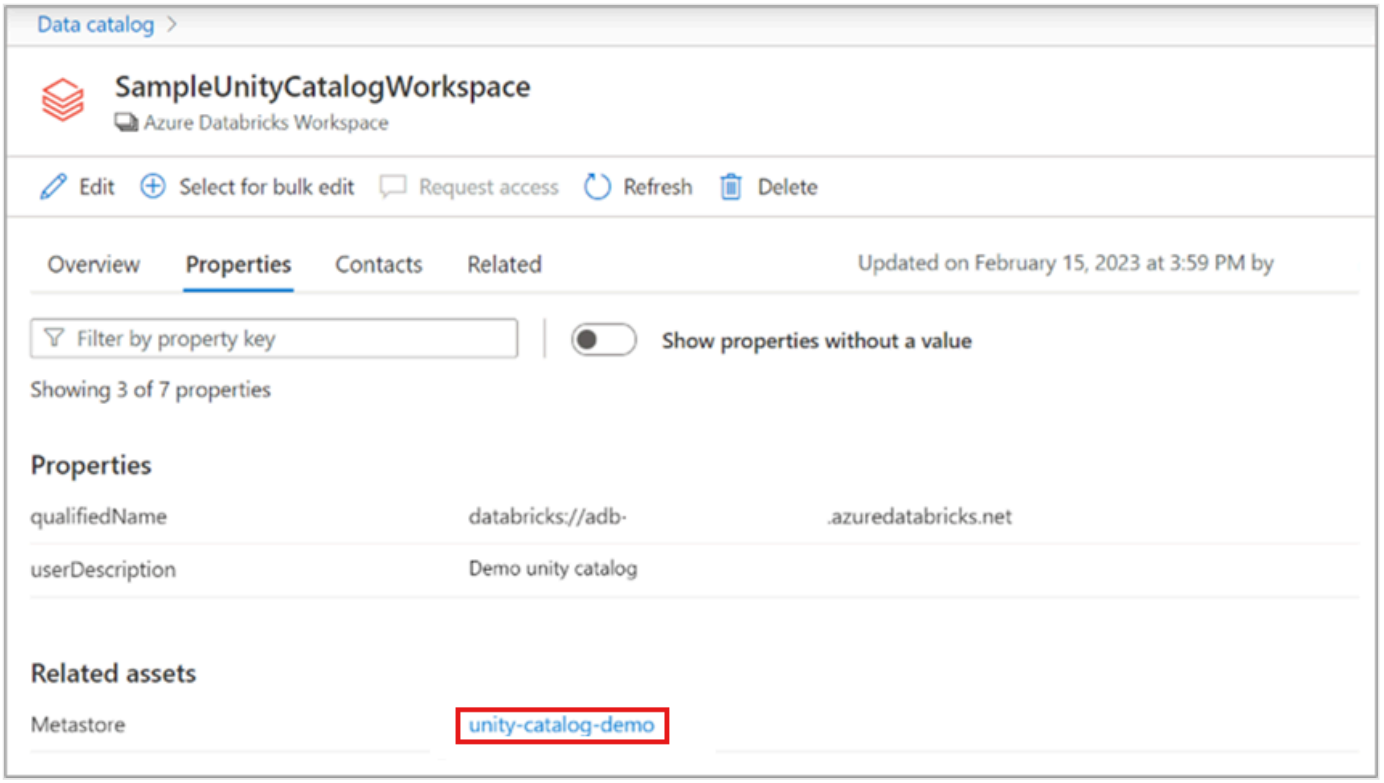
Browse and search assets

After scanning your Azure Databricks, you can browse Unified Catalog or search Unified Catalog to view the asset details and lineage.

When you browse by source types, you see two entries for **Azure Databricks Unity Catalog** and **Azure Databricks**. The **Azure Databricks Unity Catalog** entry contains the Unity Catalog artifacts, including the metastore and its catalogs, schemas, tables, and views. The **Azure Databricks** entry contains the workspace artifacts.



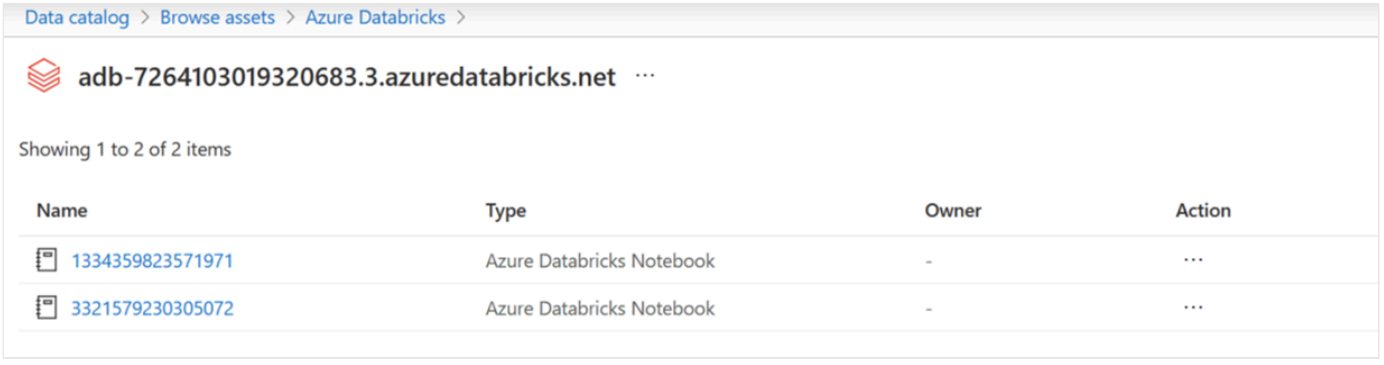
From the Azure Databricks workspace asset, you can find the associated Unity Catalog under the **Properties** tab. The association also works in reverse.

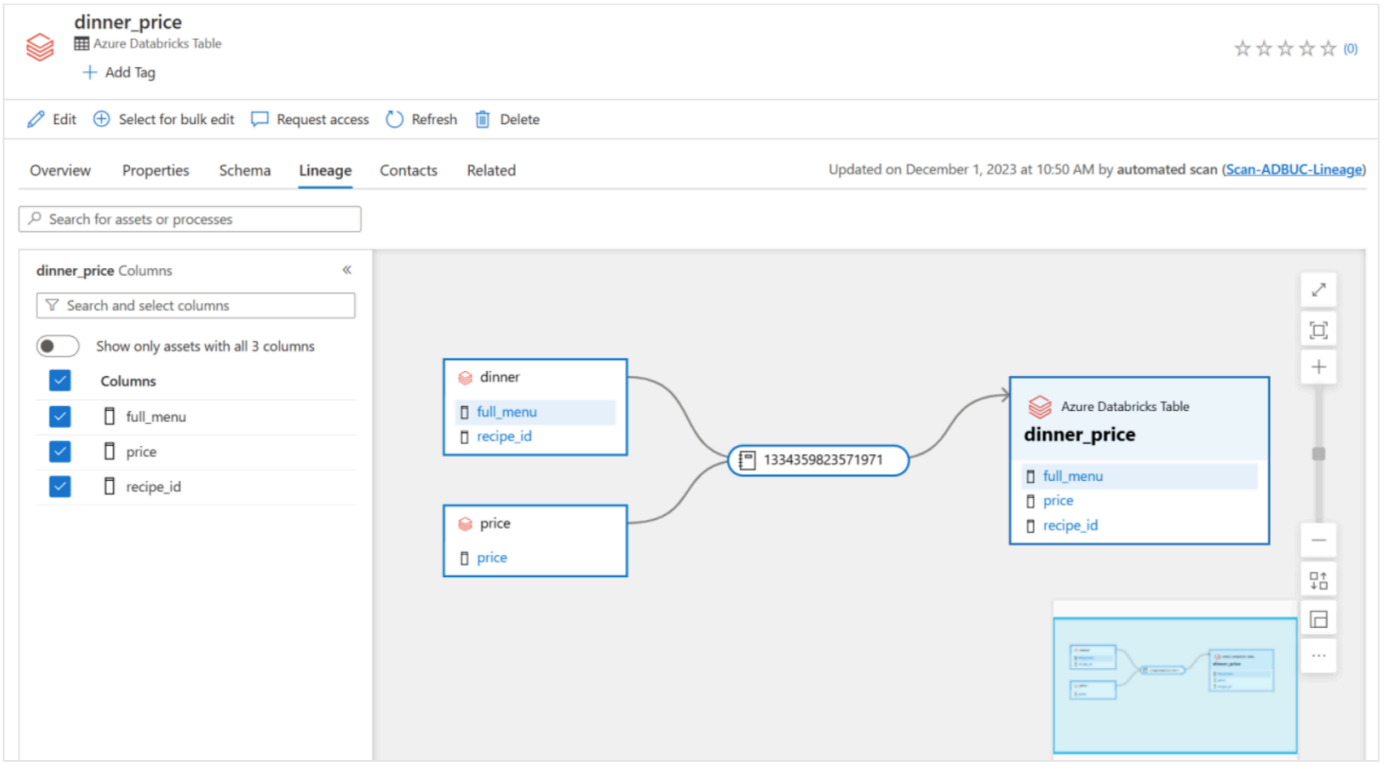


Lineage

When you browse a particular Azure Databricks asset, you can see the notebooks that have captured lineage.

Go to the asset and select the lineage tab. You can see the lineage on the Azure Databricks Notebook asset or table/view asset when applicable.





Supported lineage scenarios

- Lineage is supported for tables and views when they're connected via Databricks notebooks within the Unity Catalog.
- Lineage appears only for objects scanned through Microsoft Purview. All related assets must be scanned to form a complete lineage graph. Ensure all Databricks workspaces with relevant notebooks are included in Microsoft Purview scans.

Lineage limitations

- When you execute notebooks through Databricks jobs, column-level lineage might not be captured.
- Microsoft Purview shows only partial lineage if you don't scan all objects involved in a data flow. For example, if a notebook in Workspace A writes data to a table in Workspace B, but you only scan Workspace A by using Microsoft Purview, the lineage shows the notebook but not the destination table, resulting in incomplete lineage.
- If an external service triggers a notebook (for example, Azure Data Factory [ADF] pipeline calling a Databricks job), the lineage in Microsoft Purview doesn't reflect that dependency.
  - In this case, lineage between ADF datasets and Databricks assets isn't shown.
  - Only lineage built within Databricks notebooks is captured.

Lineage is missing scenario

Lineage extraction is passive; Microsoft Purview only ingests what Unity Catalog system table logs and makes accessible.

- Refer to the preceding sections to ensure that your lineage scenario is supported.
- Validate that Unity Catalog lineage system tables (`system.access.table_lineage`, `system.access.column_lineage`) are populated correctly.
- Raise a support ticket if you still encounter problems.

Refer to the supported capabilities section for supported Databricks Unity Catalog lineage scenarios. For more information about lineage in general, see data lineage and lineage user guide.

Frequently asked questions (FAQ)

Does Microsoft Purview capture column-level lineage from Unity Catalog?

Microsoft Purview captures lineage at both the Unity Catalog table and view level, and at the column level.

Why don't I see column-level lineage?

Column-level lineage appears when you run your notebook from a cluster. SQL warehouses don't generate column-level lineage.

How do I fix a timeout error?

Your scan might not finish if your workspace contains a large number of assets. To fix this problem, scope your scan to a few catalogs at a time. This approach reduces the number of assets in each scan and helps your scans complete.

Why didn't Microsoft Purview fetch the lineage after I ran my notebook?

After you run your notebook, Databricks might take a few minutes to update the lineage information in its system tables. Microsoft Purview can fetch the lineage after the system tables are updated.

Next steps

After you register your source, use the following guides to learn more about Microsoft Purview and your data:

- Search the Unified Catalog
- Data Estate Insights in Microsoft Purview

Last updated on 05/19/2026

Connect to and manage dedicated SQL pools in Microsoft Purview

Article • 05/21/2025

This article outlines how to register dedicated SQL pools (formerly SQL DW), and how to authenticate and interact with dedicated SQL pools in Microsoft Purview. For more information about Microsoft Purview, read the introductory article.

> **Note**
>
> There are dedicated SQL pools (formerly SQL DW) and dedicated SQL pools in Azure Synapse workspace and you can refer to What is the difference between dedicated SQL pools (SQL DW) and dedicated SQL pools in Azure Synapse workspaces for the detail difference between the two. Currently Microsoft Purview provides separate data sources for dedicated SQL pools (formerly SQL DW) and dedicated SQL pools:
>
> - if you are looking to register and scan a dedicated SQL pool (formerly SQL DW), you can follow instructions in Connect to and manage dedicated SQL pools in Microsoft Purview.
> - if you are looking to register and scan a dedicated SQL pool (formerly SQL DW) which has enabled Azure Synapse workspace features as documented in Enable Azure Synapse workspace features for a dedicated SQL pool (formerly SQL DW), you can also follow instructions in Connect to and manage Azure Synapse Analytics workspaces in Microsoft Purview.
> - if you are looking to register and scan a dedicated SQL pool or a serverless SQL pool in Azure Synapse workspace, you must follow instructions in Connect to and manage Azure Synapse Analytics workspaces in Microsoft Purview.

Supported capabilities

Scanning capabilities

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan
Yes	Yes	Yes	Yes

When scanning dedicated SQL pool (formerly SQL DW) source, Microsoft Purview supports extracting technical metadata including:

- Server
- Dedicated SQL pools
- Schemas
- Tables including columns
- Views including columns

When setting up the scan, you can further scope it after providing the dedicated SQL pool name by selecting tables and views as needed.

Other capabilities

For **classifications**, **sensitivity labels**, **policies**, **data lineage**, and **live view**, see the list of supported capabilities.

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- An active Microsoft Purview account.
- You will need to be a Data Source Administrator and Data Reader to register a source and manage it in the Microsoft Purview governance portal. See our Microsoft Purview Permissions page for details.

Register

This section describes how to register dedicated SQL pools in Microsoft Purview using the Microsoft Purview governance portal.

Authentication for registration

There are three ways to set up authentication:

- System or user assigned managed identity (Recommended)
- Service Principal
- SQL authentication

> **Note**
>
> Only the server-level principal login (created by the provisioning process) or members of the `loginmanager` database role in the master database can create new logins. It takes about 15 minutes after granting permission, the Microsoft Purview account should have the appropriate permissions to be able to scan the resource(s).

System or user assigned managed identity to register

You can use either your Microsoft Purview system-assigned managed identity (SAMI), or a User-assigned managed identity (UAMI) to authenticate. Both options allow you to assign authentication directly to Microsoft Purview, like you would for any other user, group, or service principal. The Microsoft Purview SAMI is created automatically when the account is created. A UAMI is a resource that can be created independently, and to create one you can follow our user-assigned managed identity guide. Create a Microsoft Entra user in the dedicated SQL pool using your managed identity object name by following the prerequisites and tutorial on Create Microsoft Entra users using Microsoft Entra applications.

Example SQL syntax to create user and grant permission:

```SQL
CREATE USER [PurviewManagedIdentity] FROM EXTERNAL PROVIDER
GO

EXEC sp_addrolemember 'db_datareader', [PurviewManagedIdentity]
GO
```

The authentication must have permission to get metadata for the database, schemas, and tables. It must also be able to query the tables to sample for classification. The recommendation is to assign `db_datareader` permission to the identity.

Service Principal to register

To use service principal authentication for scans, you can use an existing one or create a new one.

If you need to create a new Service Principal, follow these steps:

1. Navigate to the Azure portal.
2. Select **Microsoft Entra ID** from the left-hand side menu.
3. a. Select **App registrations**.

4. Select + **New application registration**.
5. Enter a name for the **application** (the service principal name).
6. Select **Accounts in this organizational directory only**.
7. For Redirect URI, select **Web** and enter any URL you want; it doesn't have to be real or work.
8. Then select **Register**.

It is required to get the Service Principal's application ID and secret:

1. Navigate to your Service Principal in the Azure portal
2. Copy the values the **Application (client) ID** from **Overview** and **Client secret** from **Certificates & secrets**.
3. Navigate to your key vault
4. Select **Settings** > **Secrets**
5. Select + **Generate/Import** and enter the **Name** of your choice and **Value** as the **Client secret** from your Service Principal
6. Select **Create** to complete
7. If your key vault is not connected to Microsoft Purview yet, you will need to create a new key vault connection
8. Finally, create a new credential using the Service Principal to set up your scan.

Granting the Service Principal access

In addition, you must also create a Microsoft Entra user in the dedicated pool by following the prerequisites and tutorial on Create Microsoft Entra users using Microsoft Entra applications. Example SQL syntax to create user and grant permission:

```SQL
CREATE USER [ServicePrincipalName] FROM EXTERNAL PROVIDER
GO

ALTER ROLE db_datareader ADD MEMBER [ServicePrincipalName]
GO
```

> **Note**
>
> Microsoft Purview will need the **Application (client) ID** and the **client secret** in order to scan.

SQL authentication to register

You can follow the instructions in CREATE LOGIN to create a login for your dedicated SQL pool (formerly SQL DW) if you don't already have one.

When authentication method selected is **SQL Authentication**, you need to get your password and store in the key vault:

1. Get the password for your SQL login
2. Navigate to your key vault
3. Select **Settings** > **Secrets**
4. Select + **Generate/Import** and enter the **Name** and **Value** as the *password* for your SQL login
5. Select **Create** to complete
6. If your key vault is not connected to Microsoft Purview yet, you will need to create a new key vault connection
7. Finally, create a new credential using the key to set up your scan.

Steps to register

To register a new SQL dedicated pool in Microsoft Purview, complete the following steps:

1. Open the Microsoft Purview governance portal by:

    - Browsing directly to https://web.purview.azure.com and selecting your Microsoft Purview account.
    - Opening the Azure portal, searching for and selecting the Microsoft Purview account. Select the **the Microsoft Purview governance portal** button.

2. Select **Data Map** on the left navigation.

3. Select **Register**

4. On **Register sources**, select **Azure Dedicated SQL Pool (formerly SQL DW)**.

5. Select **Continue**

On the **Register sources** screen, complete the following steps:

1. Enter a **Name** that the data source will be listed with in the Catalog.
2. Choose your Azure subscription to filter down dedicated SQL pools.
3. Select your dedicated SQL pool.
4. Select a collection from the list.
5. Select **Register** to register the data source.

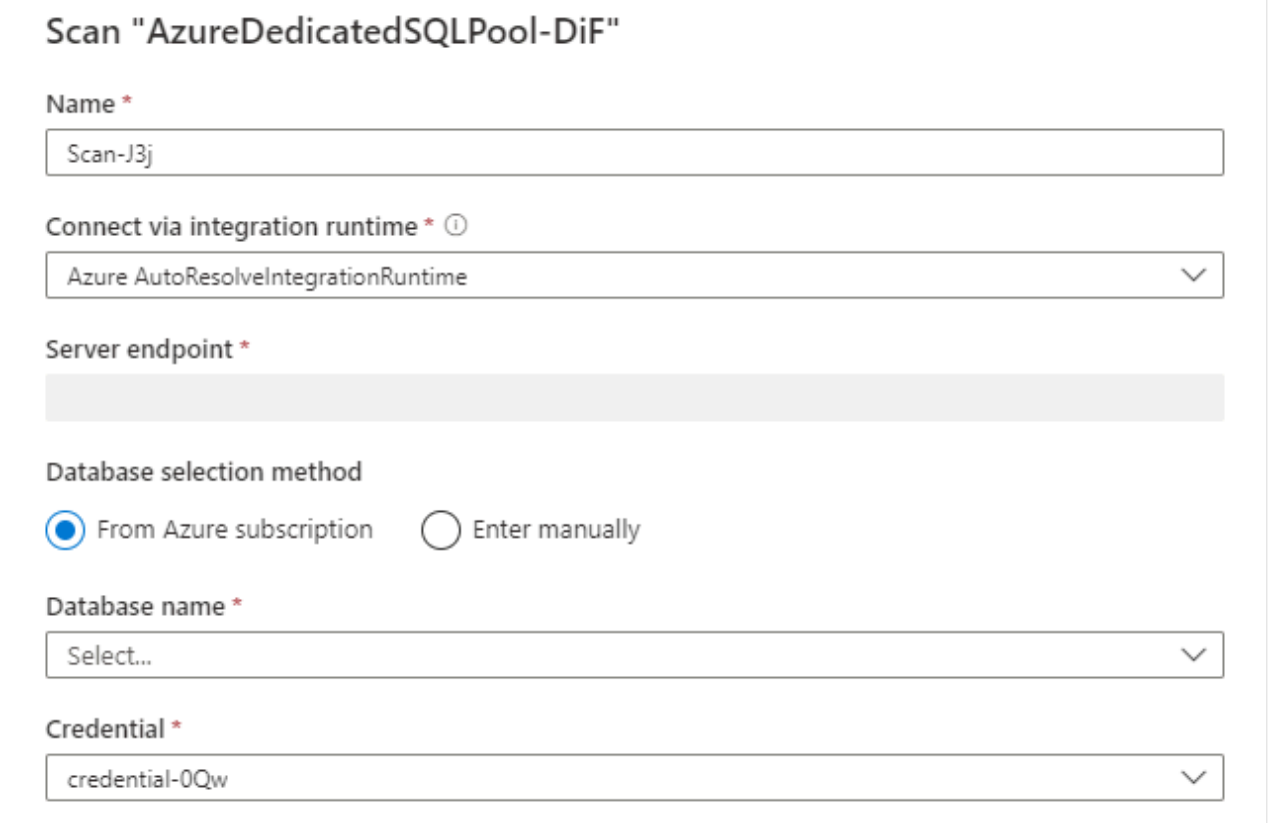
Scan

Follow the steps below to scan dedicated SQL pools to automatically identify assets and classify your data. For more information about scanning in general, see our introduction to scans and ingestion

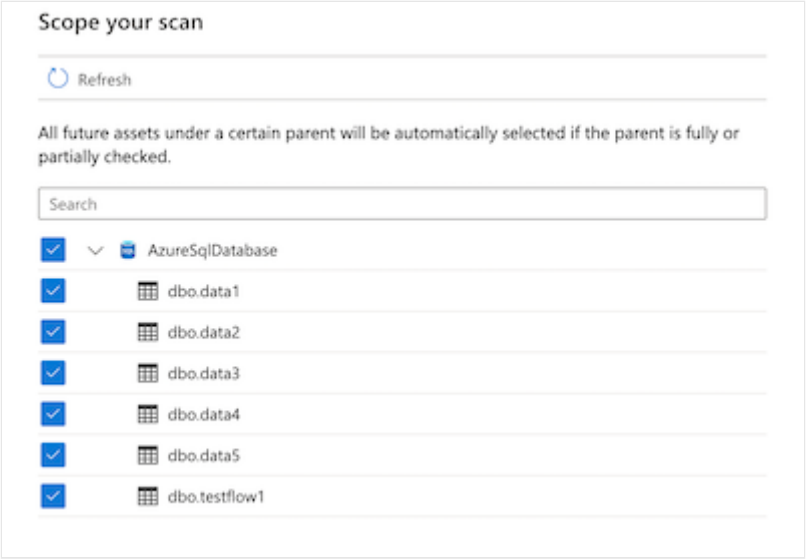
Create and run scan

To create and run a new scan, complete the following steps:

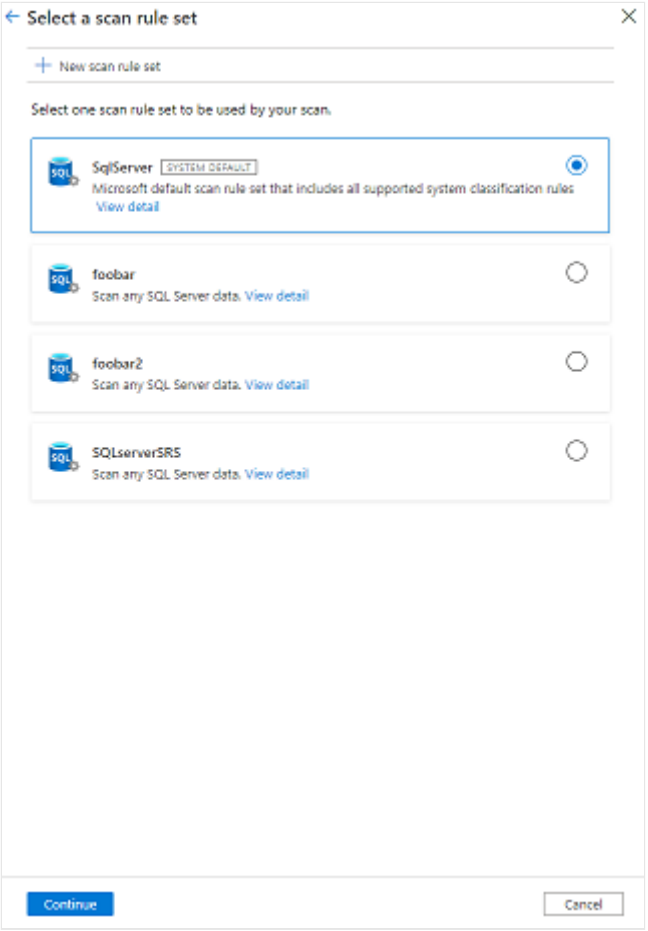
1. Select the **Data Map** tab on the left pane in the Microsoft Purview governance portal.
2. Select the SQL dedicated pool source that you registered.
3. Select **New scan**
4. Choose either the Azure integration runtime if your source is publicly accessible, a managed virtual network integration runtime if using a managed virtual network, or a self-hosted integration runtime if your source is in a private virtual network. For more information about which integration runtime to use, see the choose the right integration runtime configuration article.
5. Select the credential to connect to your data source.



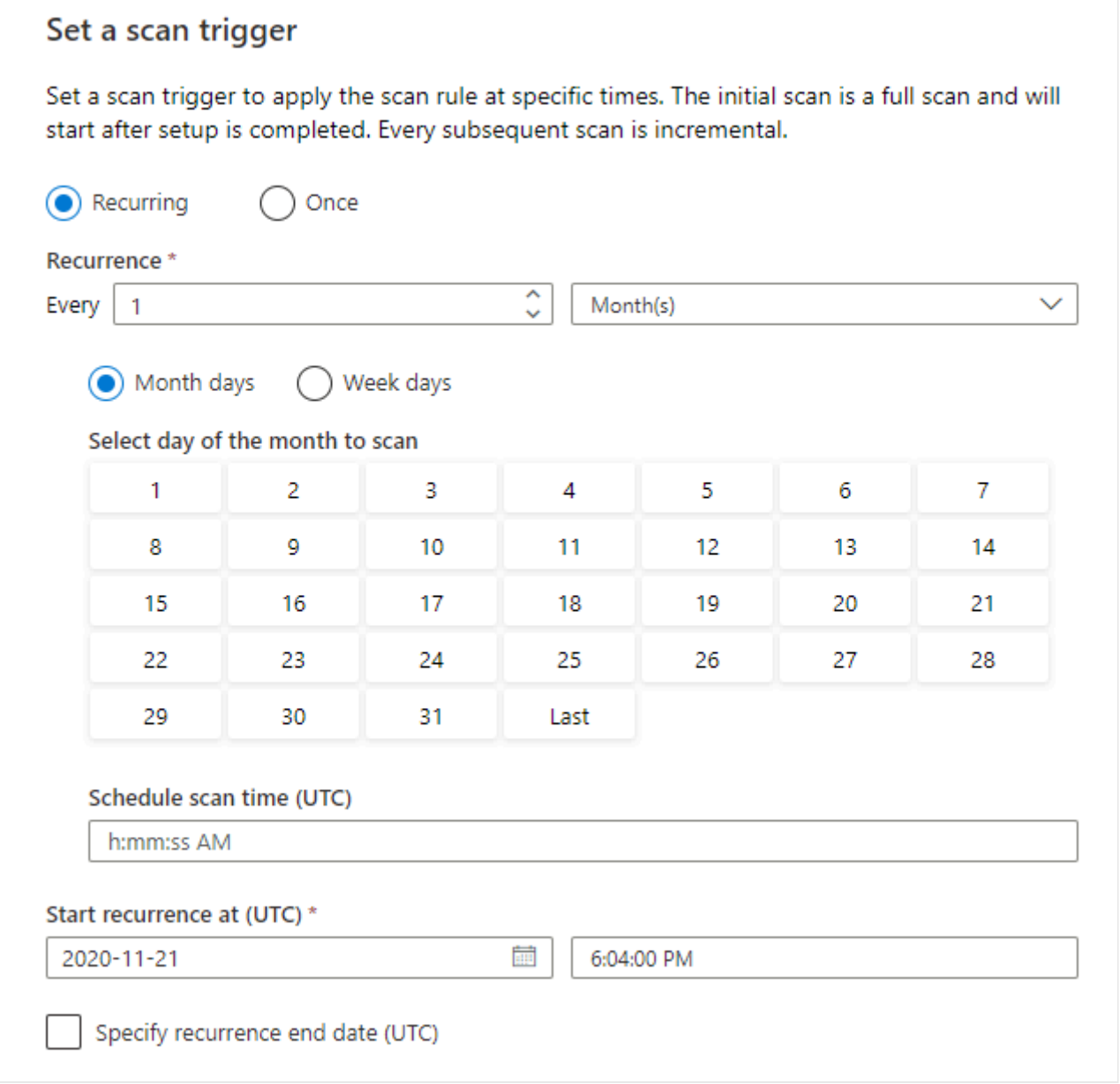
6. You can scope your scan to specific tables by choosing the appropriate items in the list.



7. Then select a scan rule set. You can choose between the system default, existing custom rule sets, or create a new rule set inline.



8. Choose your scan trigger. You can set up a schedule or run the scan once.



9. Review your scan and select **Save and run**.

View your scans and scan runs

To view existing scans:

1. Go to the Microsoft Purview portal. On the left pane, select **Data map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that has results you want to view. The pane shows you all the previous scan runs, along with the status and metrics for each scan run.
4. Select the run ID to check the scan run details.

Manage your scans

To edit, cancel, or delete a scan:

1. Go to the Microsoft Purview portal. On the left pane, select **Data Map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that you want to manage. You can then:
    - Edit the scan by selecting **Edit scan**.
    - Cancel an in-progress scan by selecting **Cancel scan run**.
    - Delete your scan by selecting **Delete scan**.

> ⓘ **Note**
>
> - Deleting your scan does not delete catalog assets created from previous scans.

Next steps

Now that you have registered your source, follow the below guides to learn more about Microsoft Purview and your data.

- Data Estate Insights in Microsoft Purview
- Lineage in Microsoft Purview
- Search Unified Catalog

Connect to and manage Azure Machine Learning in Microsoft Purview (preview)

Article • 05/21/2025

This article outlines how to register Azure Machine Learning and how to authenticate and interact with Azure Machine Learning in Microsoft Purview. For more information about Microsoft Purview, read the introductory article.

This integration between Azure Machine Learning and Microsoft Purview applies an auto push model that, once the Azure Machine Learning workspace has been registered in Microsoft Purview, the metadata from workspace is pushed to Microsoft Purview automatically on a daily basis. It isn't necessary to manually scan to bring metadata from the workspace into Microsoft Purview.

> **Important**
>
> This feature is currently in preview. The Supplemental Terms of Use for Microsoft Azure Previews include additional legal terms that apply to Azure features that are in beta, in preview, or otherwise not yet released into general availability.

Supported capabilities

Scanning capabilities

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan
Yes	Yes	Yes	No

When scanning the Azure Machine Learning source, Microsoft Purview supports:

- Extracting technical metadata from Azure Machine Learning, including:
  - Workspace
  - Models
  - Datasets
  - Jobs

> **Note**

1. AML workspaces don't currently support pushing metadata through a private endpoint to Microsoft Purview.
2. You must register assets in your AML workspace for them to appear in Microsoft Purview.

Other capabilities

For **classifications**, **sensitivity labels**, **policies**, **data lineage**, and **live view**, see the list of supported capabilities.

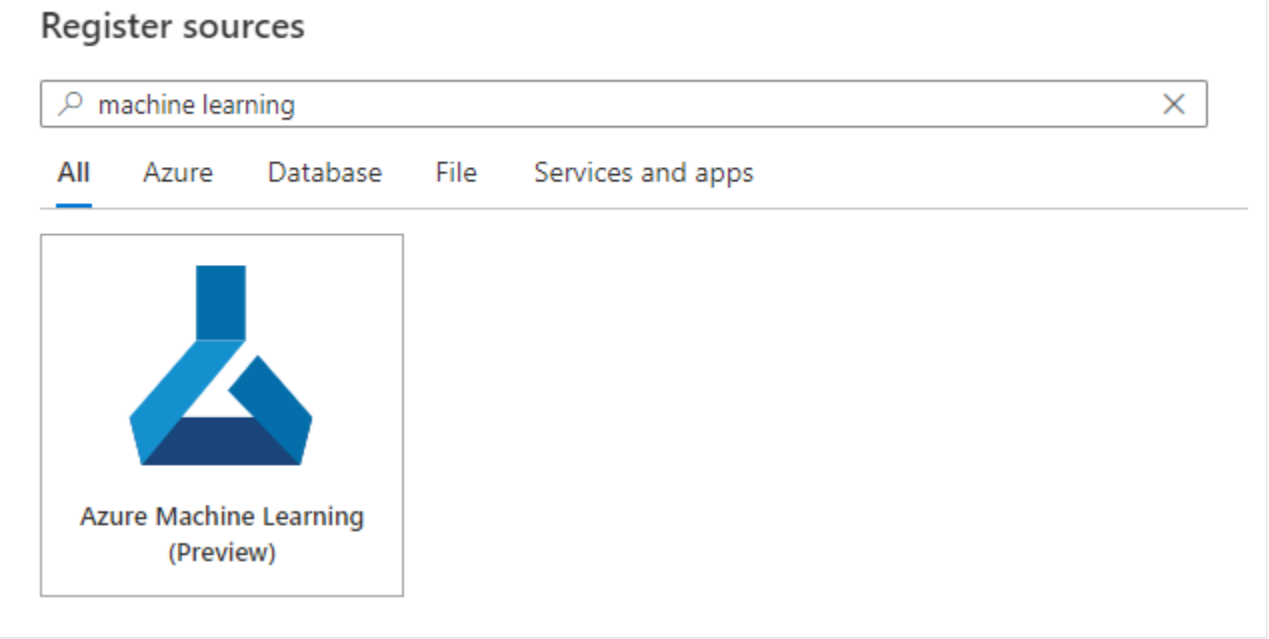
Prerequisites

- You must have an Azure account with an active subscription. Create an account for free.
- You must have an active Microsoft Purview account.
- You need Data Source Administrator and Data Reader permissions to register a source and manage it in the Microsoft Purview governance portal. For more information about permissions, see Access control in Microsoft Purview.
- An active Azure Machine Learning workspace
- A user needs the Contributor role in the Azure Machine Learning workspace to enable auto push from Azure Machine Learning workspace.

Register

This section describes how to register an Azure Machine Learning workspace in Microsoft Purview by using the Microsoft Purview governance portal.

1. Go to your Microsoft Purview account.
2. Select **Data Map** on the left pane.
3. Select **Register**.
4. In **Register sources**, select **Azure Machine Learning (Preview)** > **Continue**.



5. On the **Register sources (Azure Machine Learning)** screen, do the following:

    a. For **Name**, enter a friendly name that Microsoft Purview lists as the data source for the workspace.

    b. For **Azure subscription** and **Workspace name**, select the subscription and workspace that you want to push from the dropdown. The Azure Machine Learning workspace URL is automatically populated.

    c. Select a collection from the list.

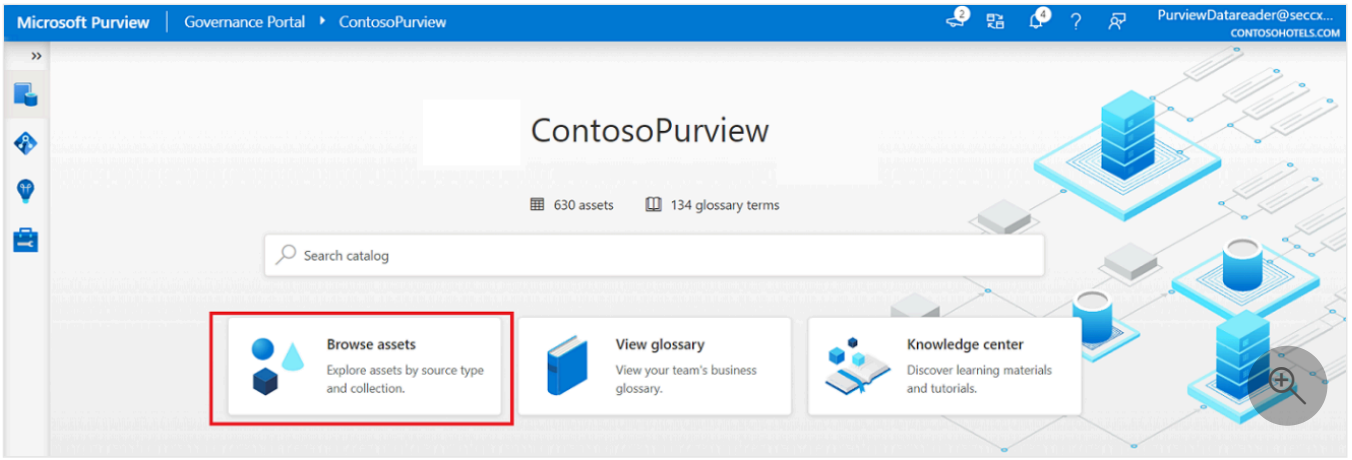
6. Select **Register** to register the source.

Scan

After you register your Azure Machine Learning workspace, the metadata will be automatically pushed to Microsoft Purview on a daily basis.

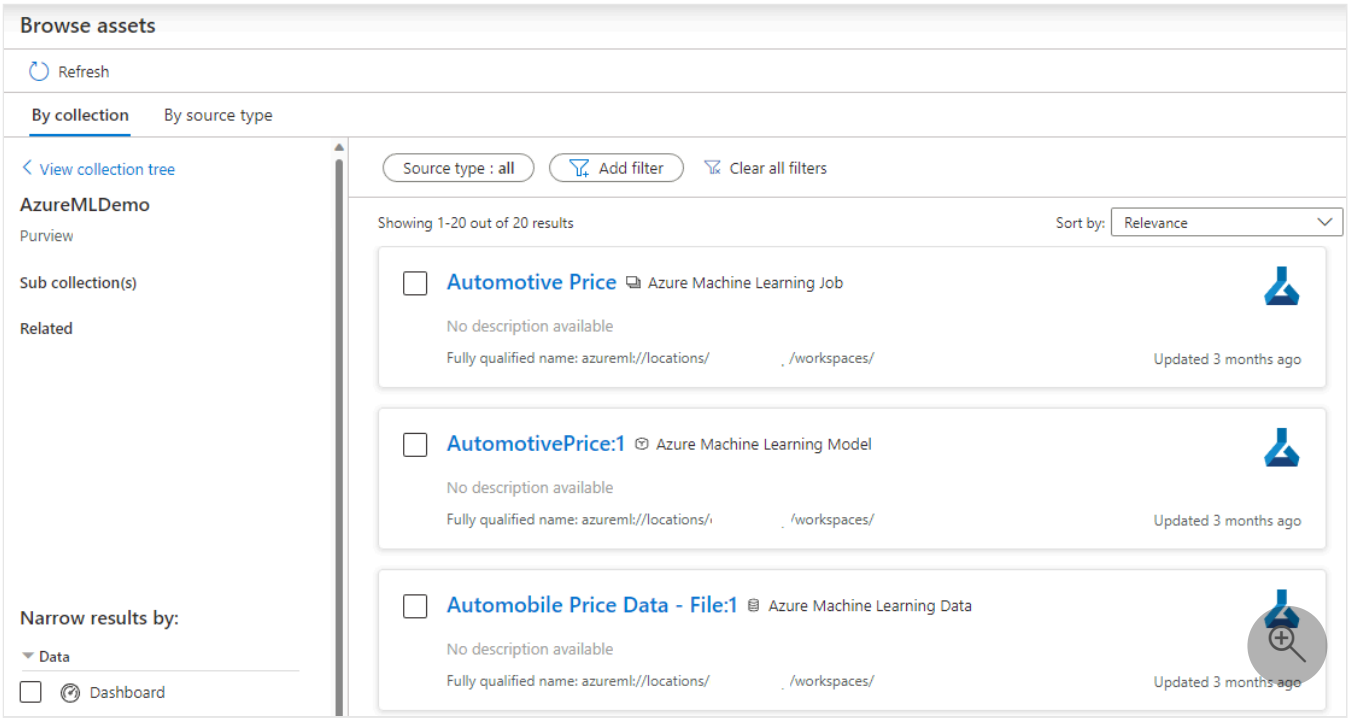
Browse and discover

To access the browse experience for data assets from your Azure Machine Learning workspace, select **Browse Assets**.



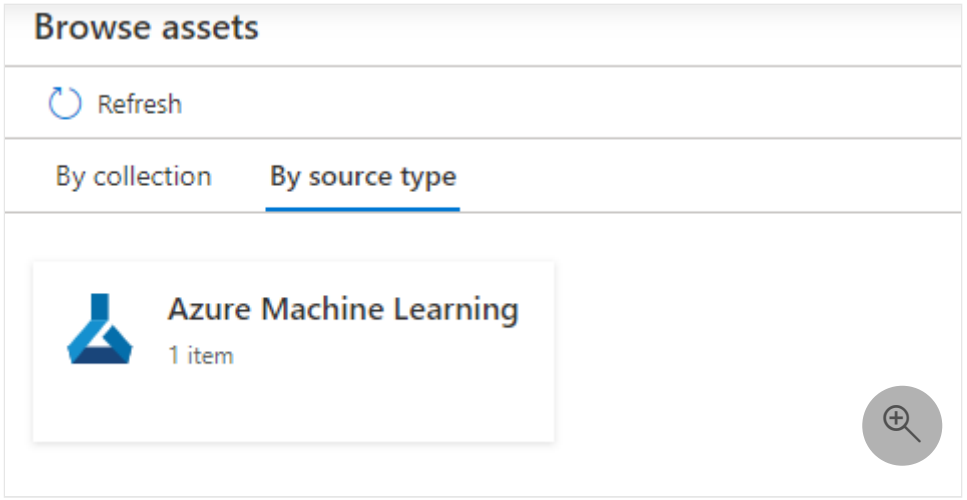
Browse by collection

Browse by collection allows you to explore the different collections you're a data reader or curator for.

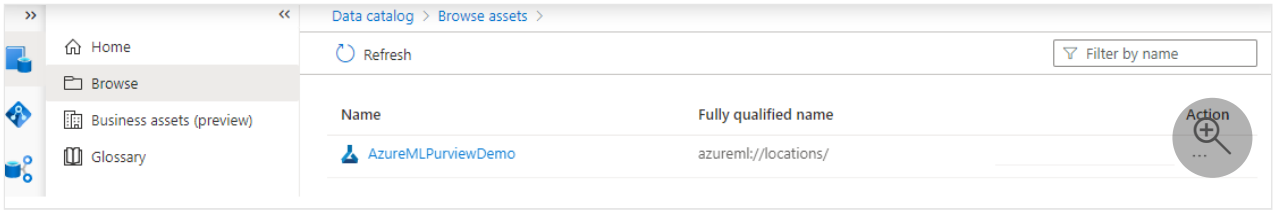


Browse by source type

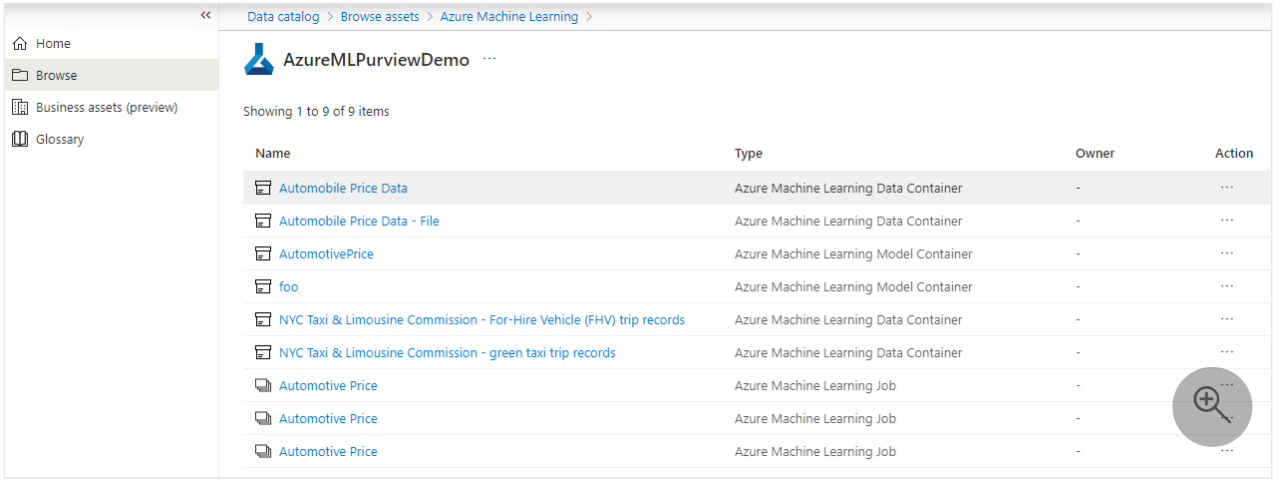
1. On the browse by source types page, select **Azure Machine Learning**.



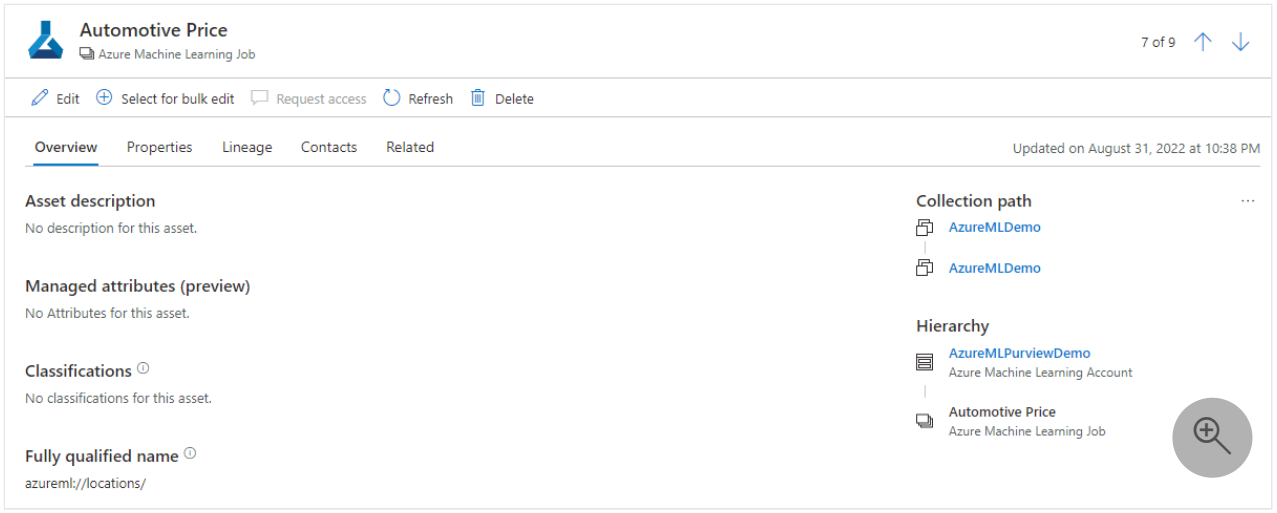
2. The top-level assets under your selected data type are listed. Pick one of the assets to further explore its contents. For example, after selecting Azure Machine Learning, you'll see a list of workspaces with assets in Unified Catalog.



3. Selecting one of the workspaces displays the child assets.



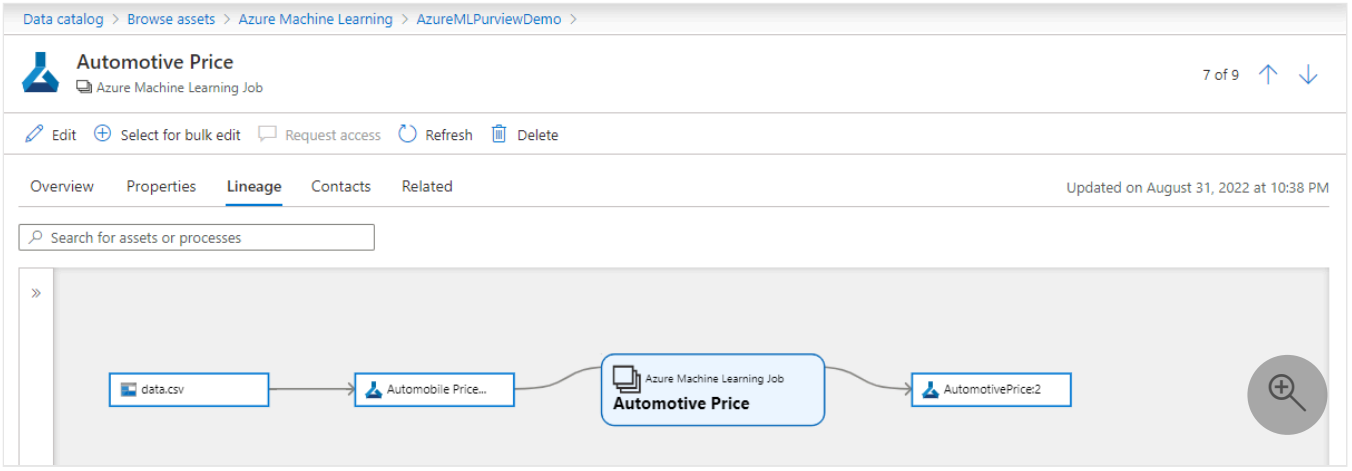
4. From the list, you can select on any of the asset items to view details. For example, selecting one of the Azure Machine Learning job assets displays the details of the job.



Lineage

To view lineage information, select an asset and then select the **Lineage** tab. From the lineage tab, you can see the asset's relationships when applicable. You can see what source data was used (if registered in Purview), the data asset created in Azure Machine Learning, any jobs, and finally the resulting machine learning model. In more advanced scenarios, you can see:

- If multiple data sources were used
- Multiple stages of training on multiple data assets
- If multiple models were created from the same data sources



For more information on lineage in general, see data lineage and lineage users guide.

Next steps

Now that you've registered your source, use the following guides to learn more about Microsoft Purview and your data:

- Data Estate Insights in Microsoft Purview
- Lineage in Microsoft Purview

- Search the Unified Catalog

Discover and govern Azure SQL Database in Microsoft Purview

This article outlines the process to register an Azure SQL database source in Microsoft Purview. It includes instructions to authenticate and interact with the SQL database.

Supported capabilities

Scanning capabilities

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan
Yes	Yes	Yes	Yes

When you're scanning Azure SQL Database, Microsoft Purview supports extracting technical metadata from these sources:

- Server
- Database
- Schemas
- Tables, including columns
- Views, including columns
- Stored procedures (with lineage extraction enabled)*
- Stored procedure runs (with lineage extraction enabled)*

When you're setting up a scan, you can further scope it after providing the database name by selecting tables and views as needed.

* Stored procedure lineage extraction scan has some limitations. Refer to the Known limitations section on this page for details.

Other capabilities

For **classifications**, **sensitivity labels**, **policies**, **data lineage**, and **live view**, see the list of supported capabilities.

Known limitations

For stored procedures (SP) lineage extraction:

- SP lineage extraction scan isn't supported if your logical server in Azure disables public access or doesn't allow Azure services to access it.
- SP lineage extraction scan is currently not supported if your stored procedures contain an INSERT or a DROP query.
- SP lineage extraction scan isn't supported if your Microsoft Purview account disables public access.
  - SP lineage extraction scan is scheduled to run every six hours by default. The frequency can't be changed.
  - Lineage is captured only when the stored procedure execution transfers data from one table to another. And it's not supported for temporary tables.
  - Lineage extraction isn't supported for functions or triggers.
  - Due to the following limitations, you could see duplicate assets in the catalog if you have scenarios such as:
    - The object names in assets and fully qualified names follow the case used in stored procedure statements, which might not align with the object case in original data source.
    - When SQL views are referenced in stored procedures, they're currently captured as SQL tables.

For scanned assets:

- For a full scan, instead of INT and NUMERIC data type, DECIMAL is displayed as a data type for SQL asset properties. If length of the field is MAX, it is displayed as zero (0).
- For **nchar** data type, column length is displayed in character count.

> **Note**
>
> Lineage is also supported if Azure SQL tables or views are used as a source/sink in Azure Data Factory Copy and Data Flow activities.

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- An active Microsoft Purview account.

- Data Source Administrator and Data Reader permissions, so you can register a source and manage it in the Microsoft Purview governance portal. For details, see Access control in the Microsoft Purview governance portal.

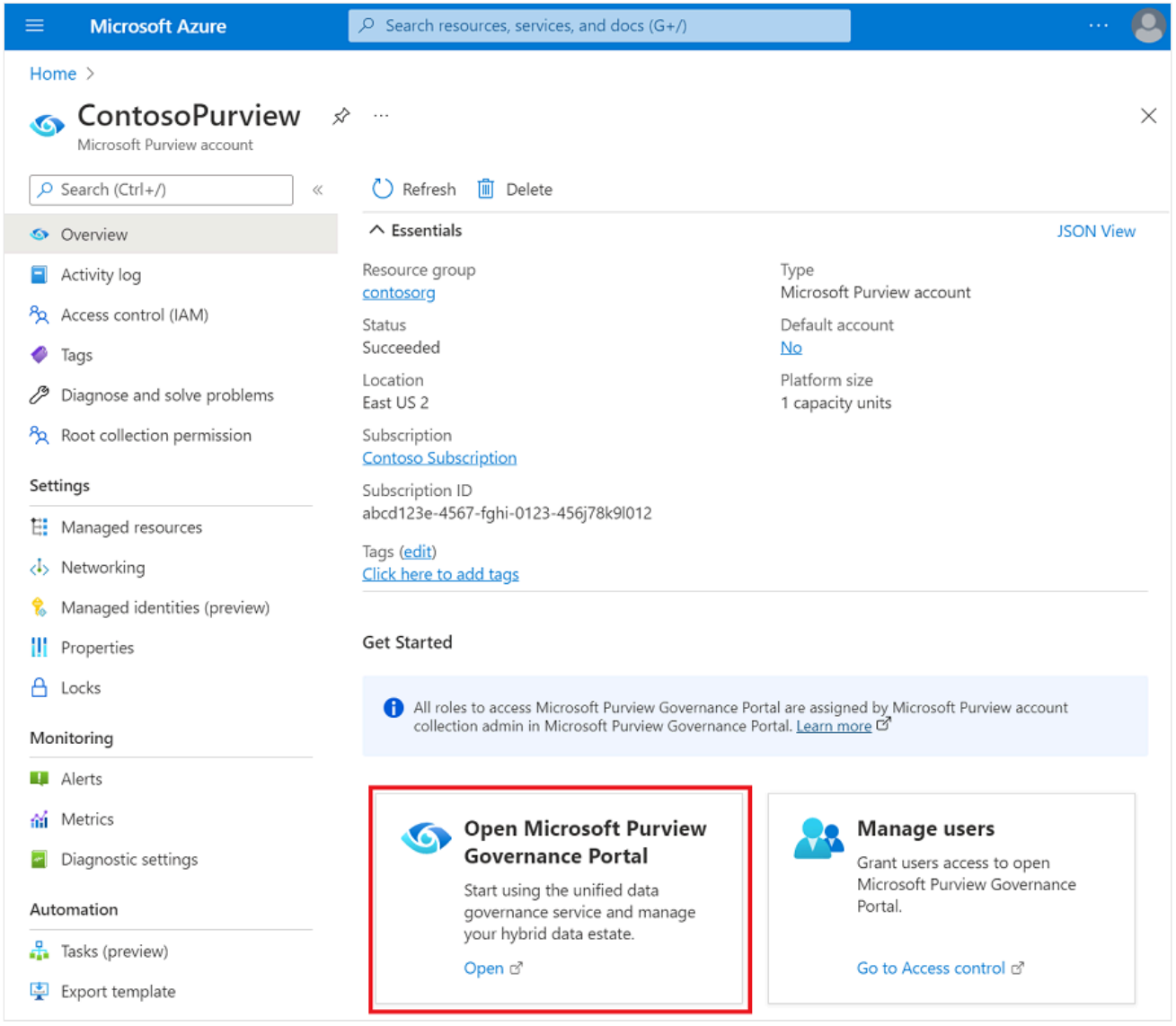
Register the data source

Before you scan, it's important to register the data source in Microsoft Purview:

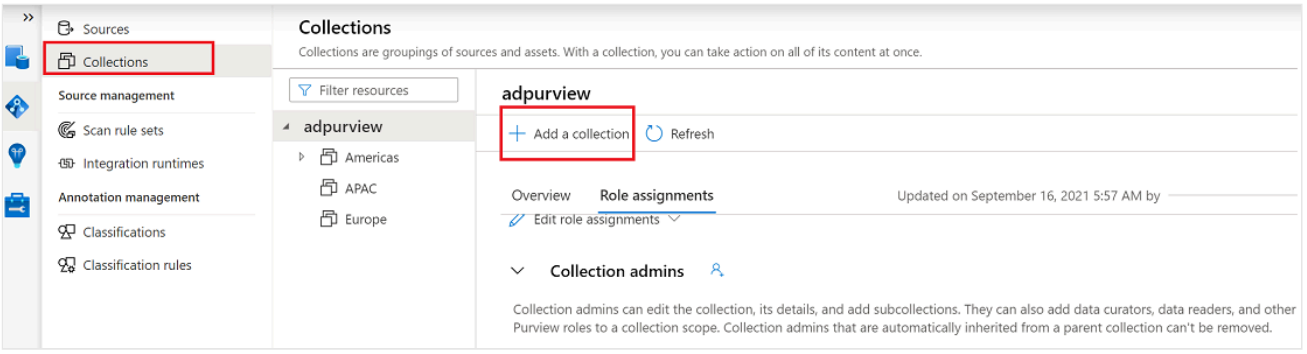
1. Open the Microsoft Purview governance portal by:

    - Browsing directly to https://web.purview.azure.com and selecting your Microsoft Purview account.
    - Opening the Azure portal, searching for and selecting the Microsoft Purview account. Select the **the Microsoft Purview governance portal** button.

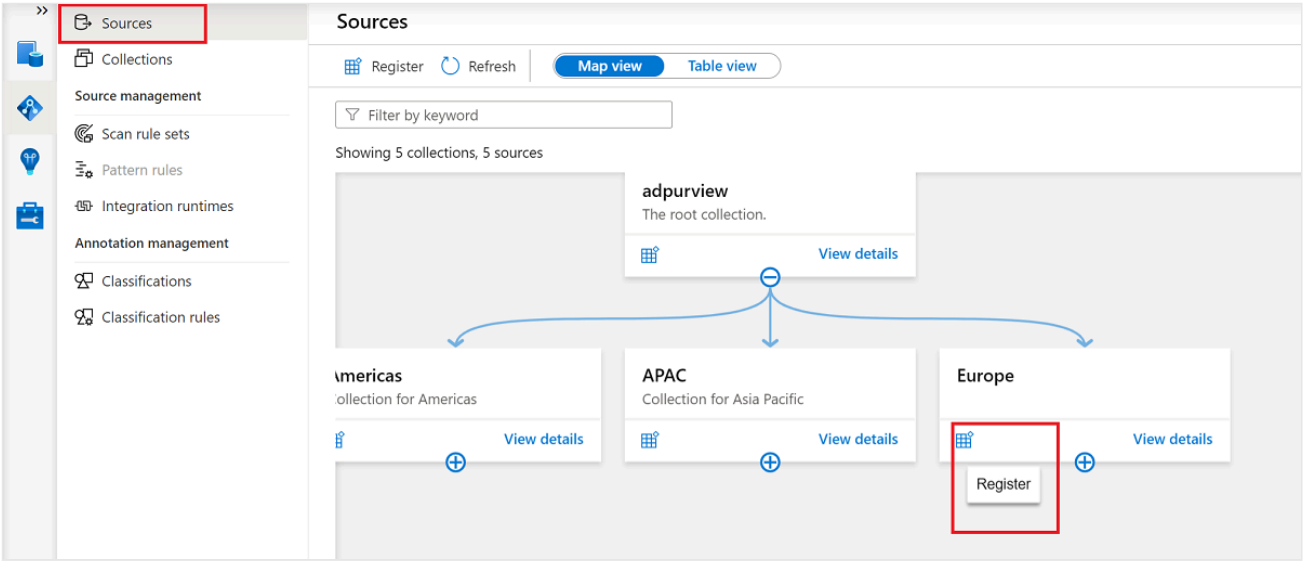
2. Navigate to the **Data Map**.



3. Create the collection hierarchy by going to **Collections** and then selecting **Add a collection**. Assign permissions to individual subcollections as required.

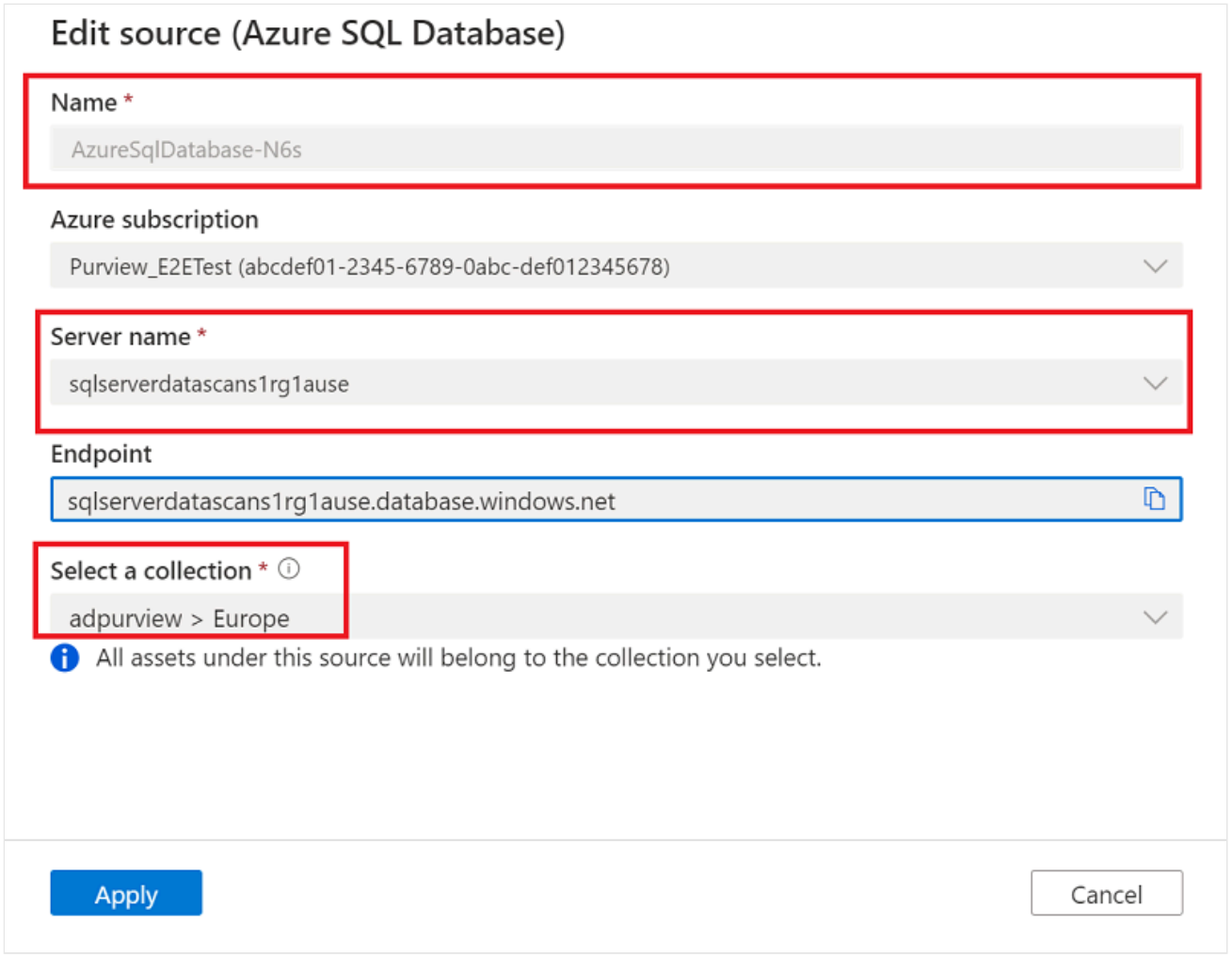


4. Go to the appropriate collection under **Sources**, and then select the **Register** icon to register a new SQL database.

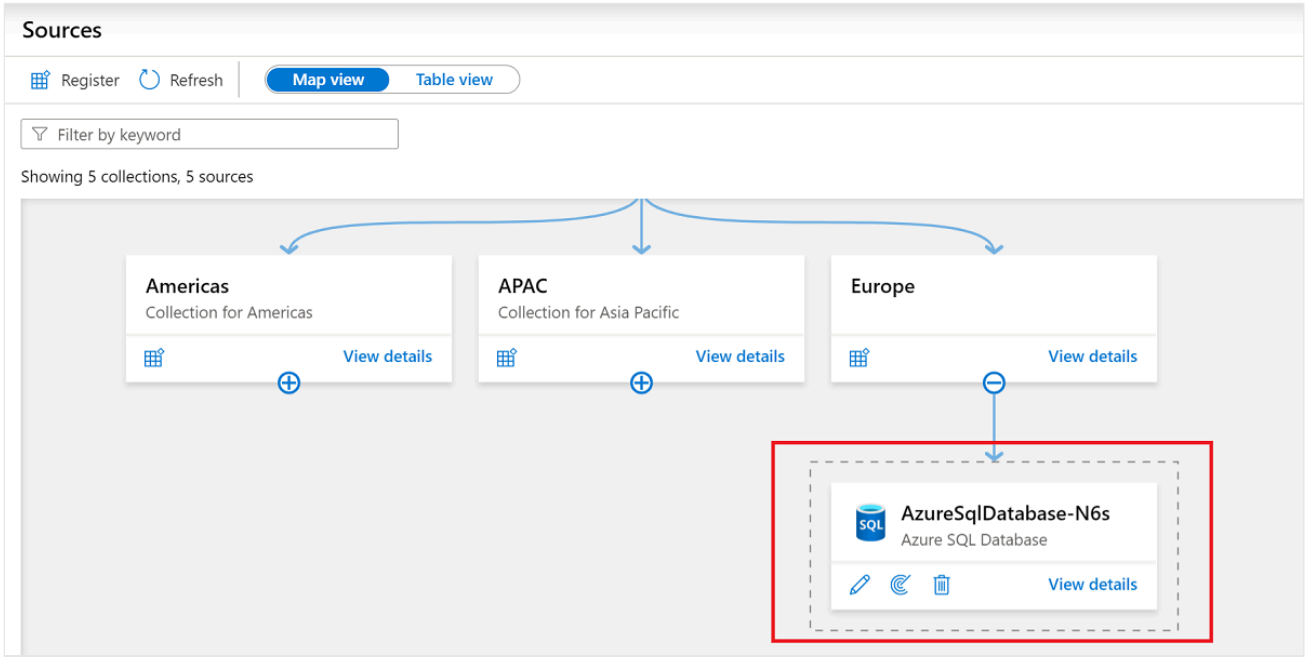


5. Select the **Azure SQL Database** data source, and then select **Continue**.

6. For **Name**, provide a suitable name for the data source. Select relevant names for **Azure subscription**, **Server name**, and **Select a collection**, and then select **Apply**.



7. Confirm that the SQL database appears under the selected collection.



Update firewall settings

If your database server has a firewall enabled, you need to update the firewall to allow access in one of the following ways:

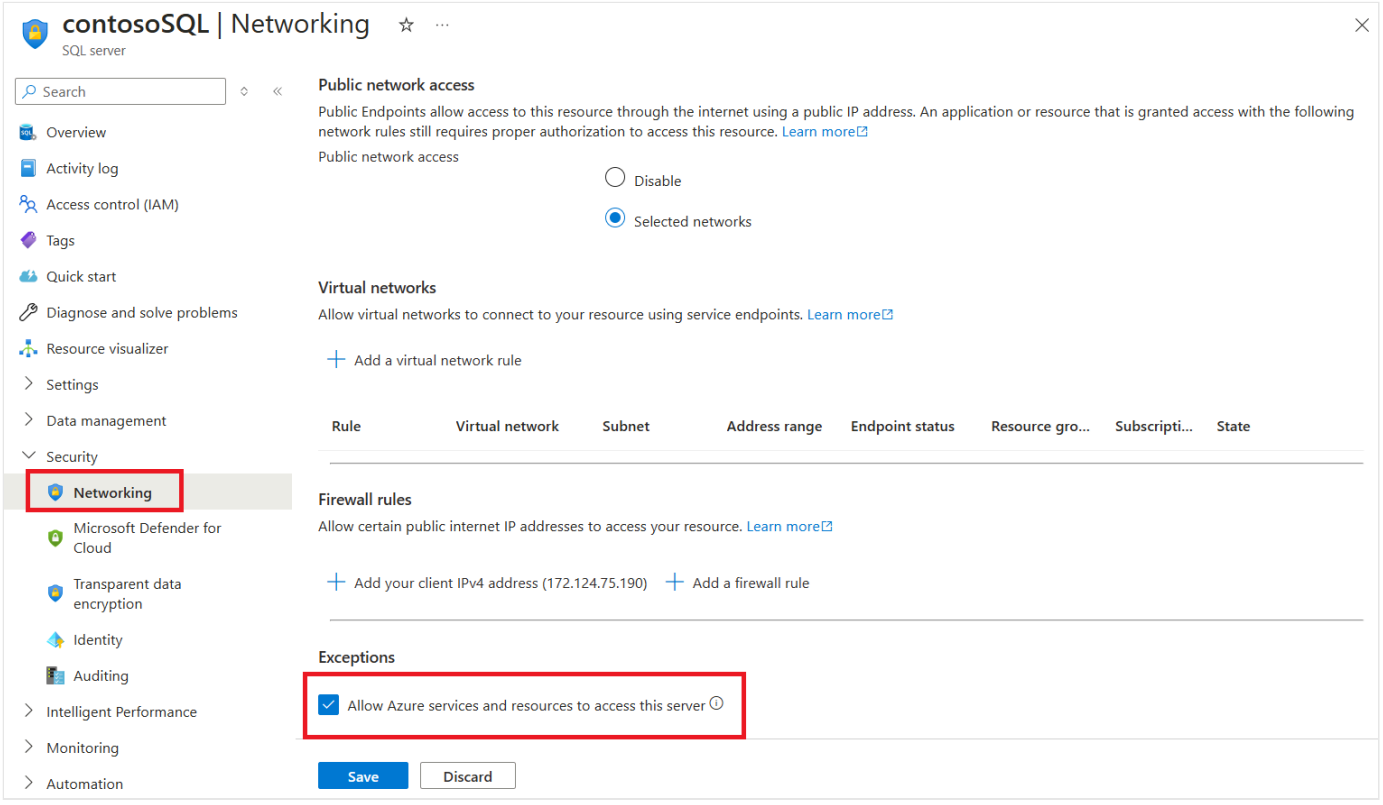
- Allow Azure connections through the firewall. This is a straightforward option to route traffic through Azure networking, without needing to manage virtual machines.
- Install a self-hosted integration runtime on a machine in your network and give it access through the firewall. If you have a private virtual network set up within Azure, or have any other closed network set up, using a self-hosted integration runtime on a machine within that network will allow you to fully manage traffic flow and utilize your existing network.
- Use a managed virtual network. Setting up a managed virtual network with your Microsoft Purview account allows you to connect to Azure SQL by using the Azure integration runtime in a closed network.

For more information about the firewall, see the Azure SQL Database firewall documentation.

Allow Azure connections

Enabling Azure connections allows Microsoft Purview to connect to the server without requiring you to update the firewall itself.

1. Go to your database account.
2. On the **Overview** page, select the server name.
3. Select **Security** > **Networking**.
4. For **Allow Azure services and resources to access this server**, select **Yes**.



For more information about allowing connections from inside Azure, see the how-to guide.

Install a self-hosted integration runtime

You can install a self-hosted integration runtime on a machine to connect with a resource in a private network:

1. Choose the right integration runtime for your scenario
2. Create and install your integration runtime:

    - **To use a self-hosted integration runtime:** Follow the article to create and configure a self-hosted integration runtime.
    - **To use a kubernetes supported self-hosted integration runtime:** Follow the article to create and configure a kubernetes supported integration runtime.

3. Check your database server's networking configuration to confirm that a private endpoint is accessible to the machine that contains the self-hosted integration runtime. Add the IP address of the machine if it doesn't already have access.
4. If your logical server is behind a private endpoint or in a virtual network, you can use an ingestion private endpoint to ensure end-to-end network isolation.

Configure authentication for a scan

To scan your data source, you need to configure an authentication method in Azure SQL Database.

> **Important**
>
> If you're using a **self-hosted integration runtime** to connect to your resource, system-assigned and user-assigned managed identities won't work. You need to use service principal authentication or SQL authentication.

Microsoft Purview supports the following options:

- **System-assigned managed identity (SAMI)** (recommended). This is an identity that is associated directly with your Microsoft Purview account. It allows you to authenticate directly with other Azure resources without needing to manage a go-between user or credential set.

  The SAMI is created when your Microsoft Purview resource is created. The SAMI is managed by Azure and uses your Microsoft Purview account's name. The SAMI can't currently be used with a self-hosted integration runtime for Azure SQL.

For more information, see the managed identity overview.

- **User-assigned managed identity (UAMI)** (preview). Similar to a SAMI, a UAMI is a credential resource that allows Microsoft Purview to authenticate against Microsoft Entra ID.

  The UAMI is managed by users in Azure, rather than by Azure itself, which gives you more control over security. The UAMI can't currently be used with a self-hosted integration runtime for Azure SQL.

  For more information, see the guide for user-assigned managed identities.

- **Service principal**. A service principal is an application that can be assigned permissions like any other group or user, without being associated directly with a person. Authentication for service principals has an expiration date, so it can be useful for temporary projects.

  For more information, see the service principal documentation.

- **SQL authentication**. Connect to the SQL database with a username and password. For more information, see the SQL authentication documentation.

  If you need to create a login, follow this guide to query a SQL database. Use this guide to create a login by using T-SQL.

  > **Note**
  >
  > Be sure to select the **Azure SQL Database** option on the page.

For steps to authenticate with your SQL database, select your chosen method of authentication from the following tabs.

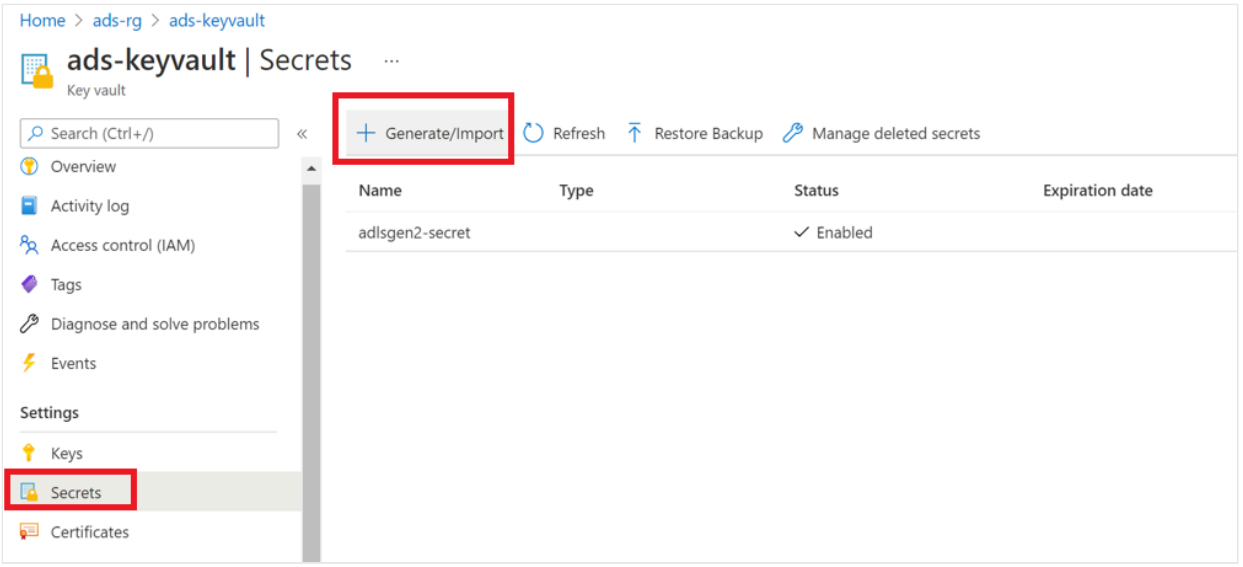
SQL authentication

> **Note**
>
> Only the server-level principal login (created by the provisioning process) or members of the `loginmanager` database role in the master database can create new logins. The Microsoft Purview account should be able to scan the resources about 15 minutes after it gets permissions.

1. You need a SQL login with at least `db_datareader` permissions to be able to access the information that Microsoft Purview needs to scan the database. You can follow the instructions in CREATE LOGIN to create a login for Azure SQL Database. Save the username and password for the next steps.

2. Go to your key vault in the Azure portal.

3. Select **Settings** > **Secrets**, and then select + **Generate/Import**.

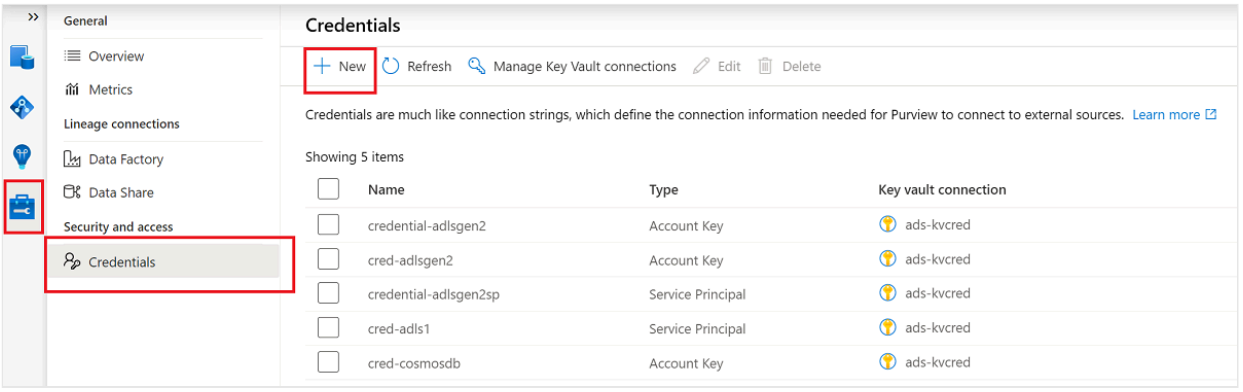


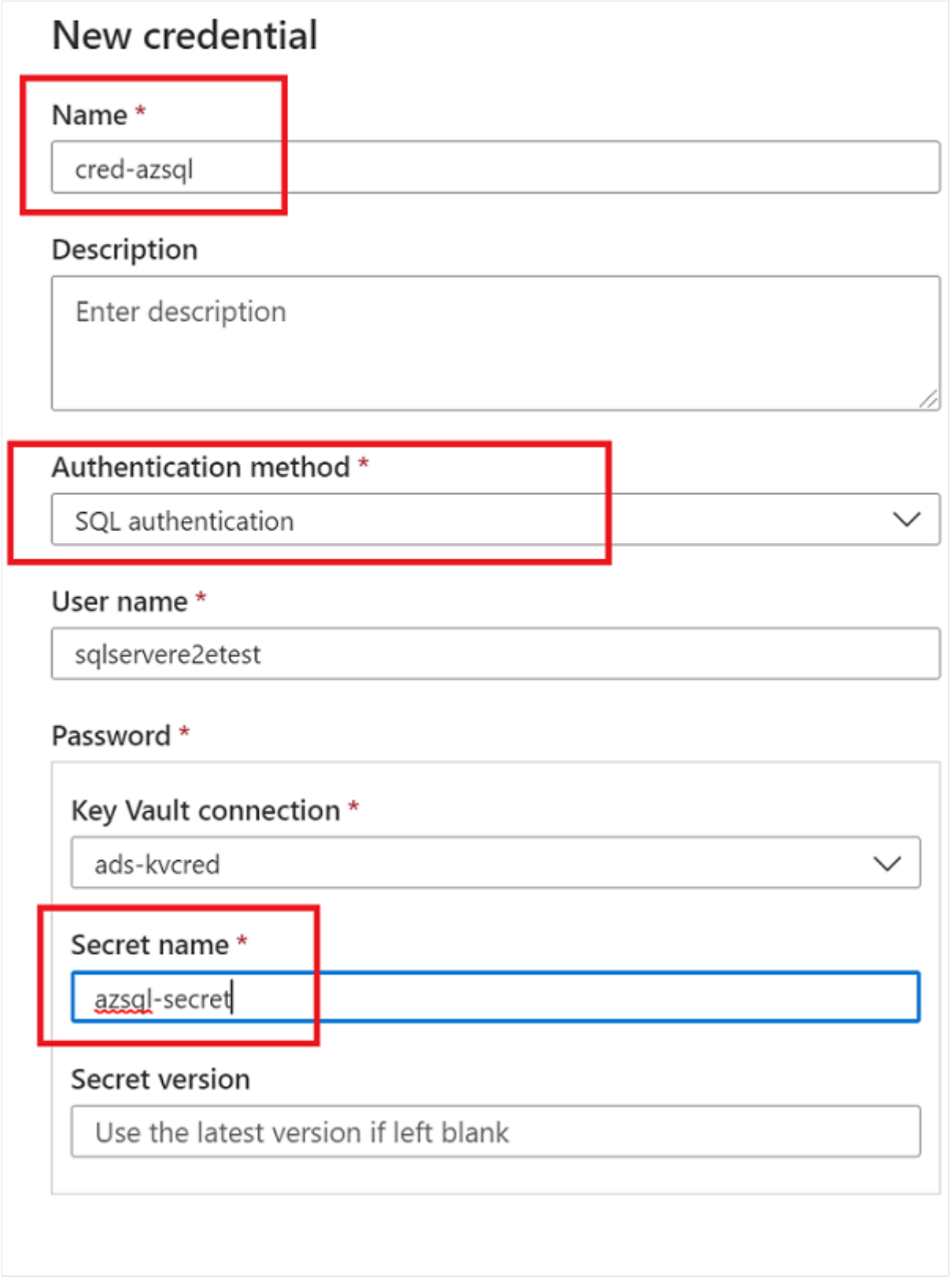
4. For **Name** and **Value**, use the username and password (respectively) from your SQL database.

5. Select **Create**.

6. If your key vault isn't connected to Microsoft Purview yet, create a new key vault connection.

7. Create a new credential by using the key to set up your scan.



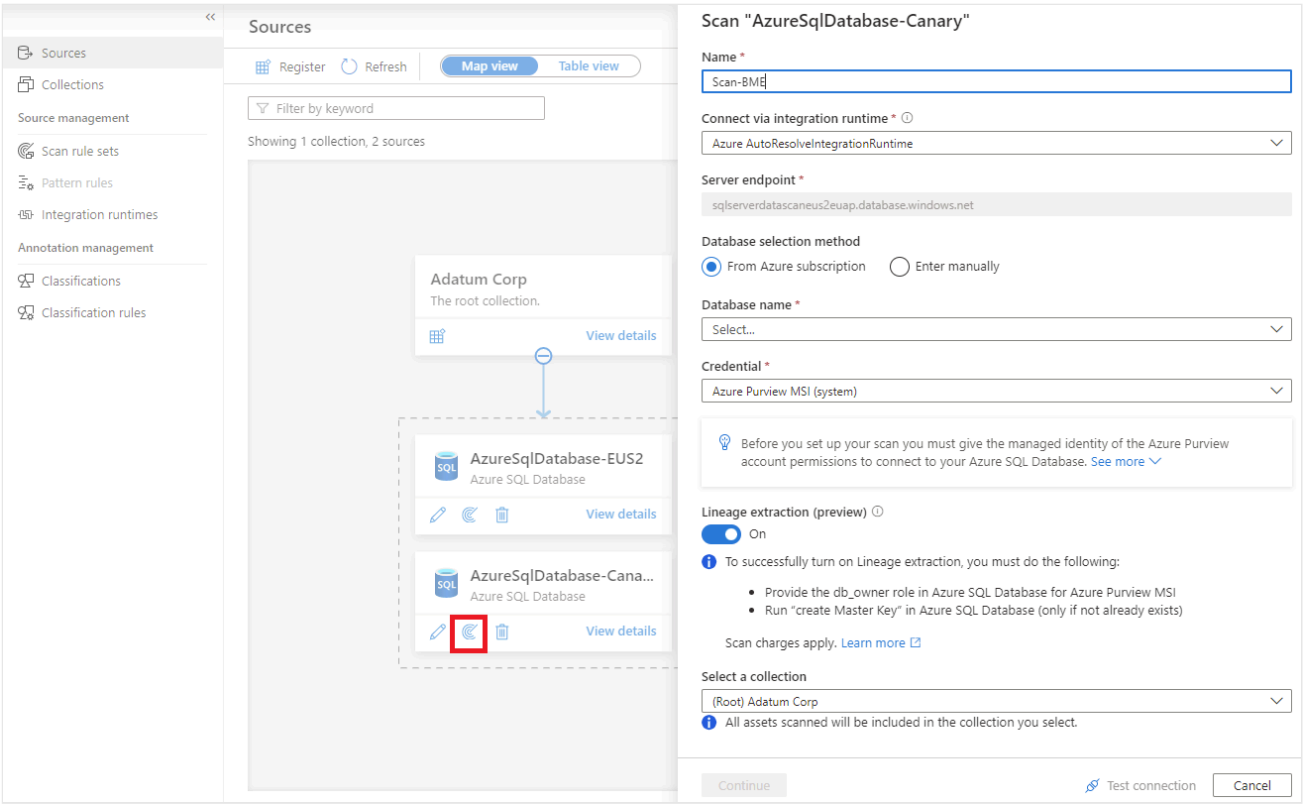


Create the scan

1. Open your Microsoft Purview account and select **Open Microsoft Purview governance portal**.

2. Go to **Data map** > **Sources** to view the collection hierarchy.

3. Select the **New Scan** icon under the SQL database that you registered earlier.



To learn more about stored procedure data lineage in Azure SQL Database, see the Extract lineage (preview) section of this article.

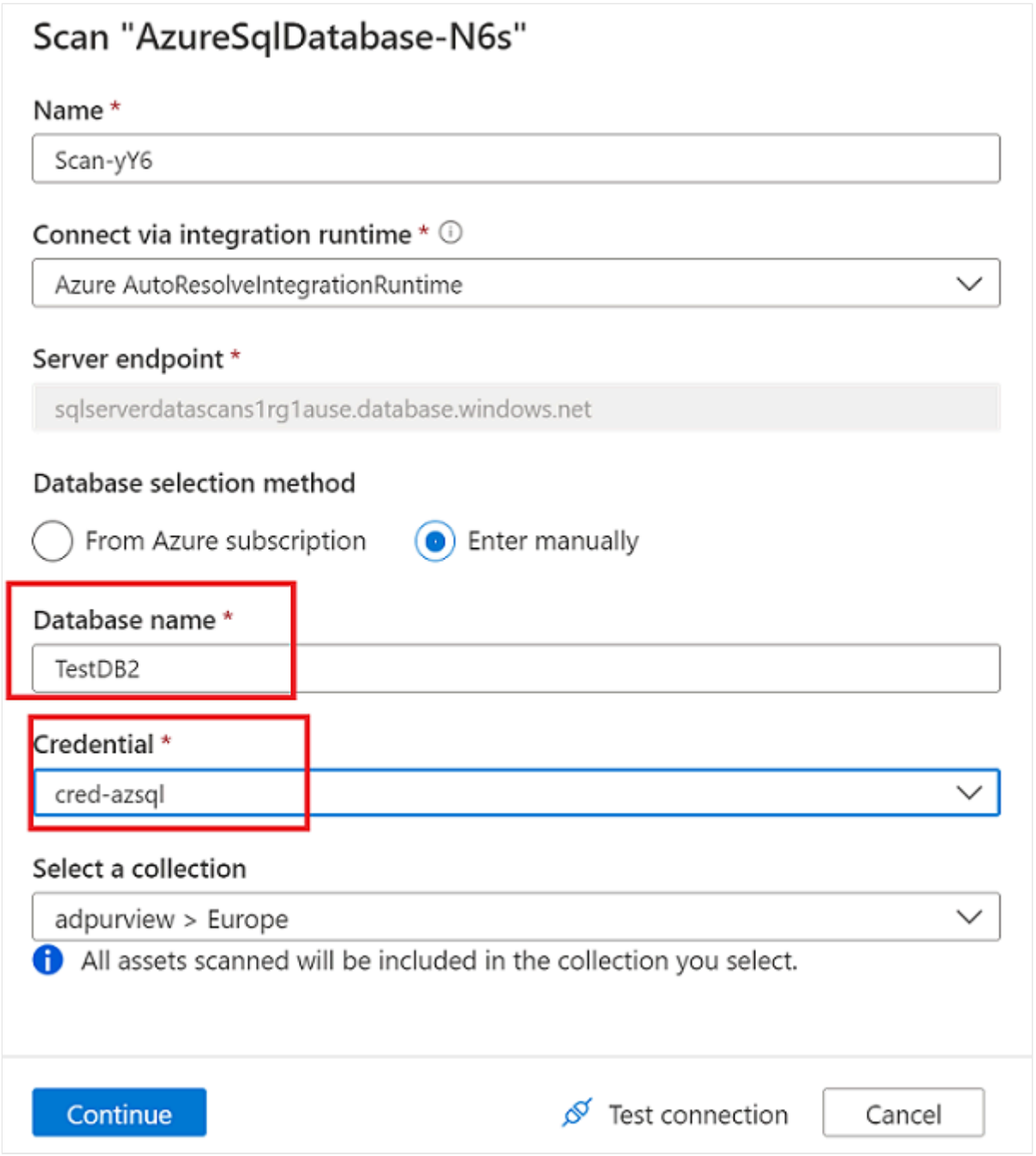
For scanning steps, select your method of authentication from the following tabs.

SQL authentication

1. For **Name**, provide a name for the scan.

2. For **Database selection method**, select **Enter manually**.

3. For **Database name** and **Credential**, enter the values that you created earlier.

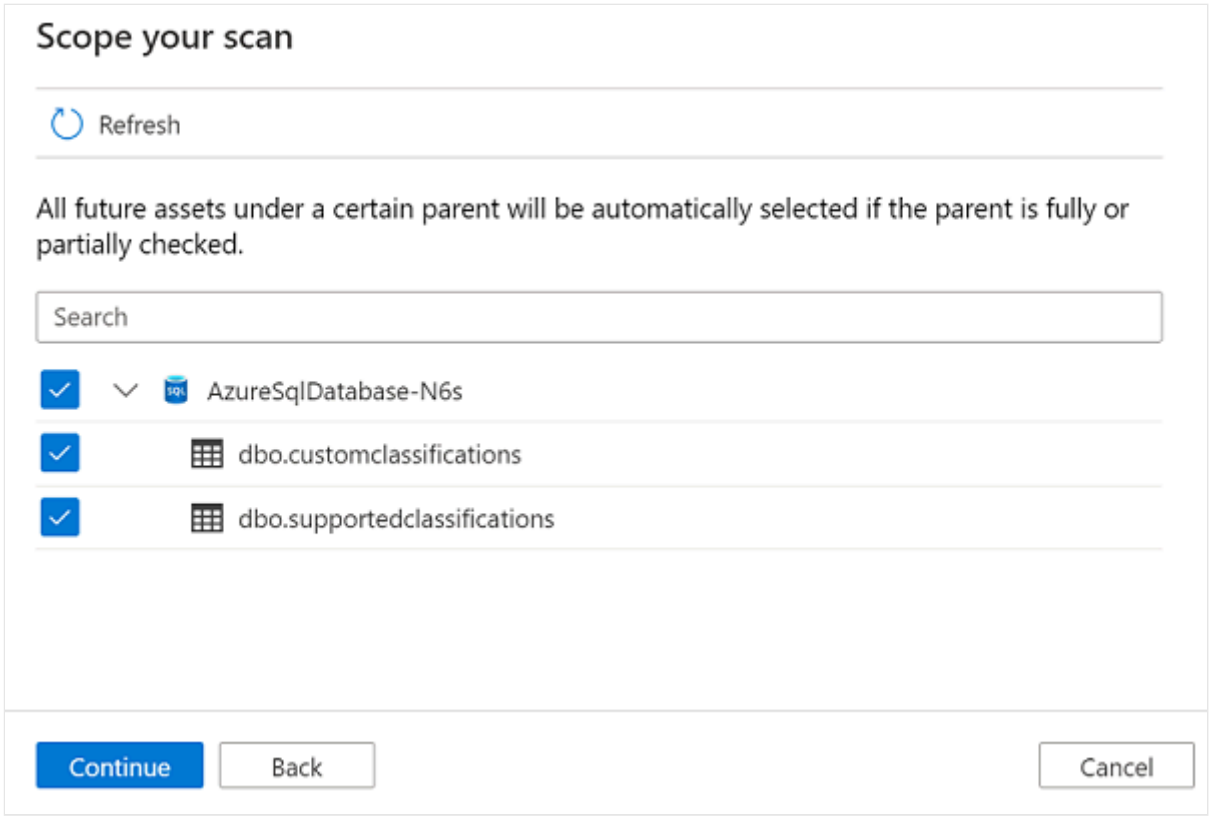


4. For **Select a connection**, choose the appropriate collection for the scan.

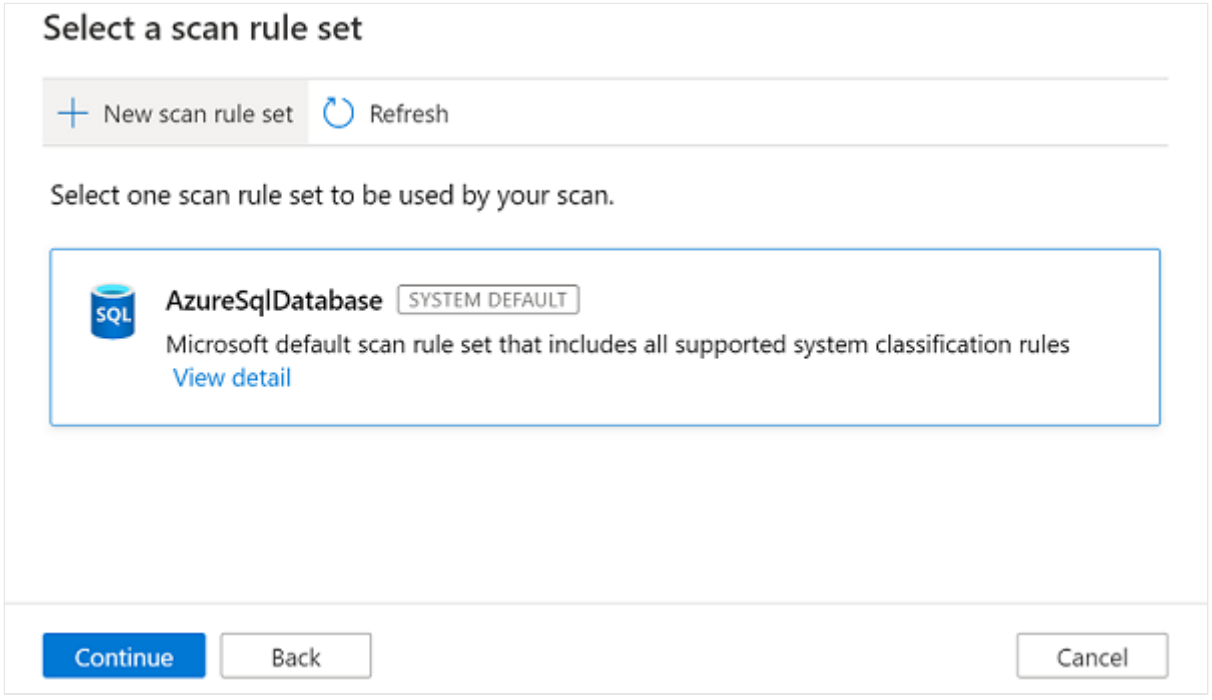
5. Select **Test connection** to validate the connection. After the connection is successful, select **Continue**.

Scope and run the scan

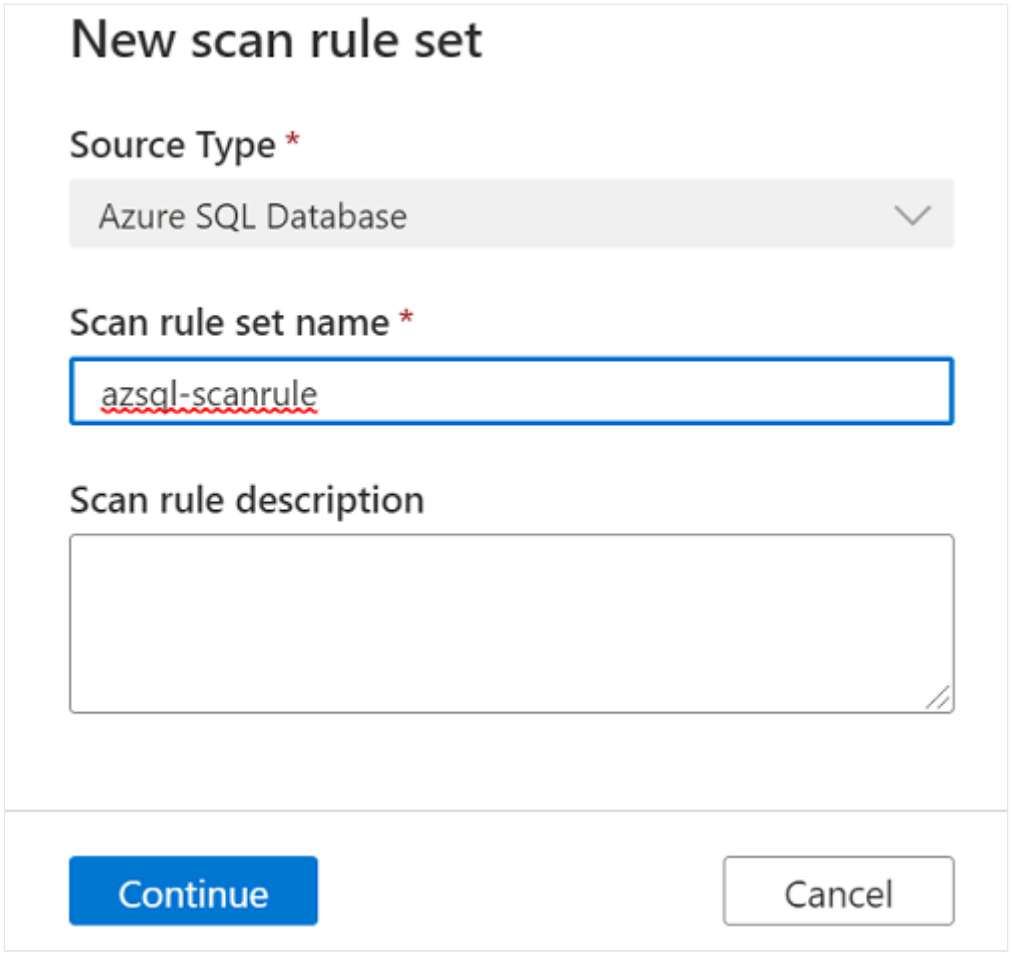
1. You can scope your scan to specific database objects by choosing the appropriate items in the list.



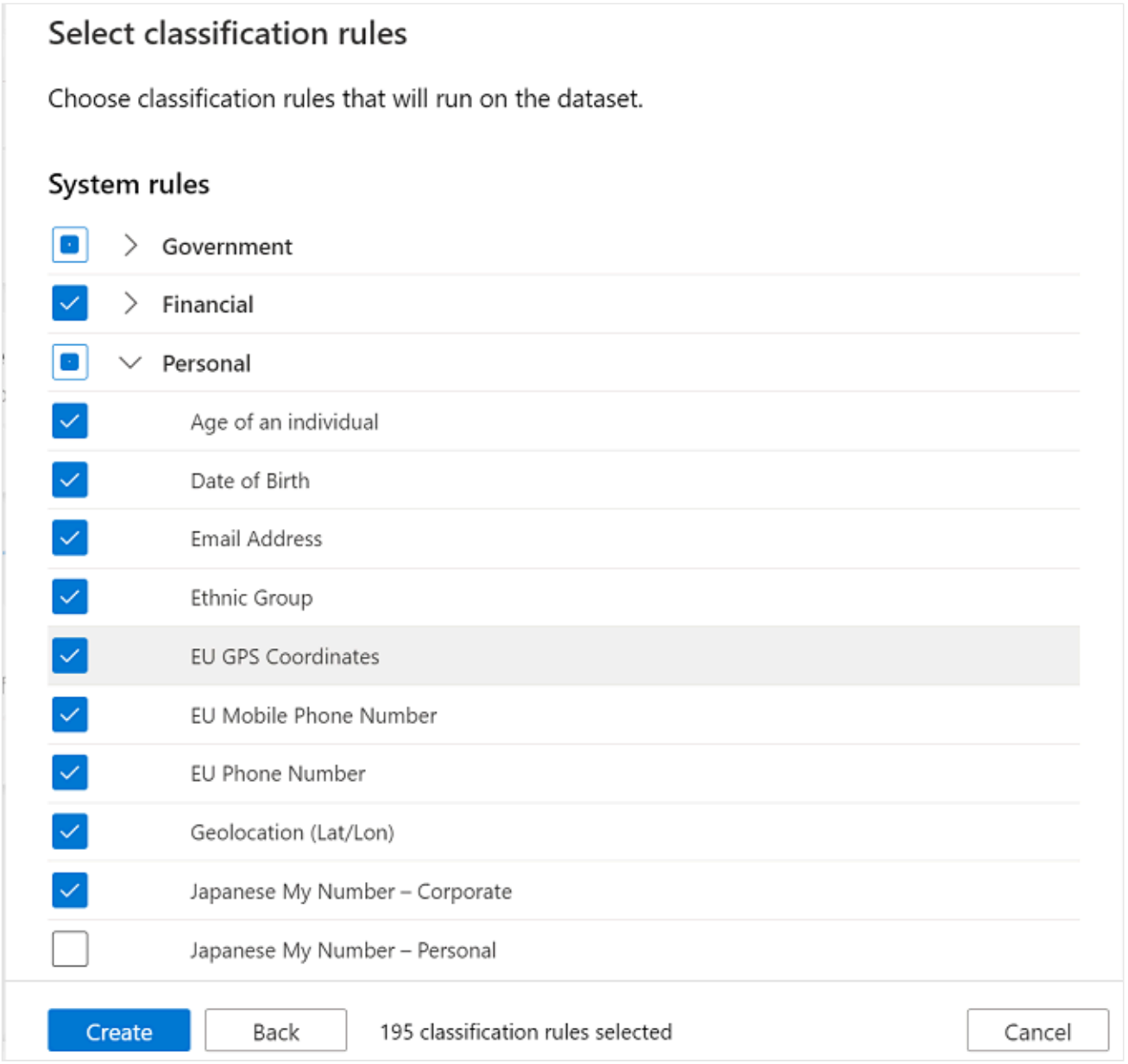
2. Select a scan rule set. You can use the system default, choose from existing custom rule sets, or create a new rule set inline. Select **Continue** when you're finished.



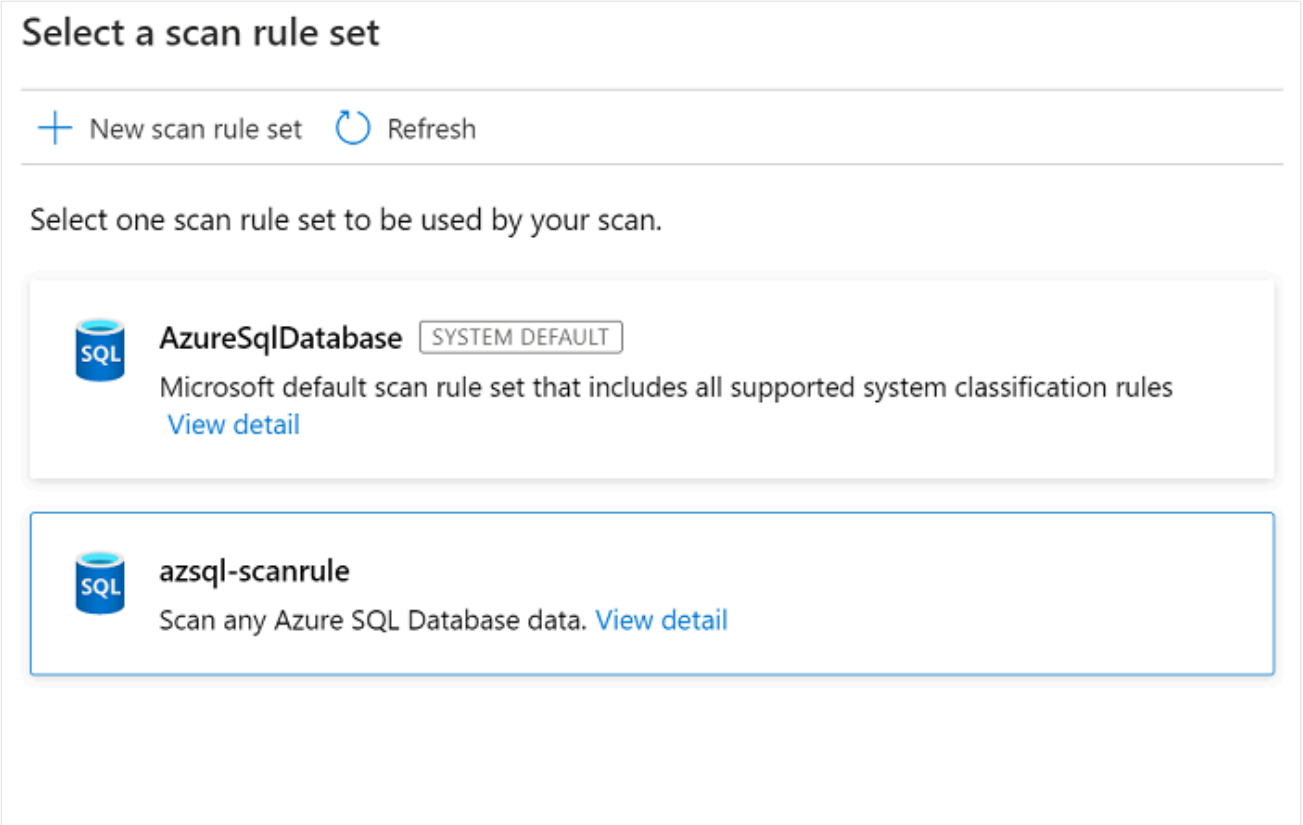
If you select **New scan rule set**, a pane opens so that you can enter the source type, the name of the rule set, and a description. Select **Continue** when you're finished.



For **Select classification rules**, choose the classification rules that you want to include in the scan rule set, and then select **Create**.



The new scan rule set then appears in the list of available rule sets.

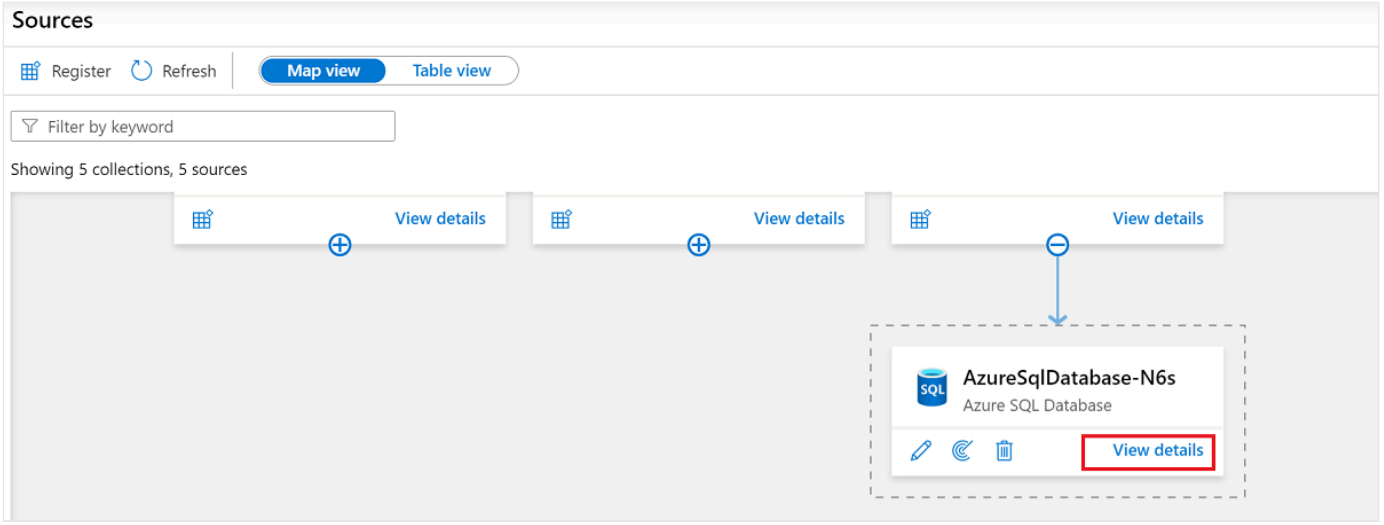


3. Choose your scan trigger. You can set up a schedule or run the scan once.

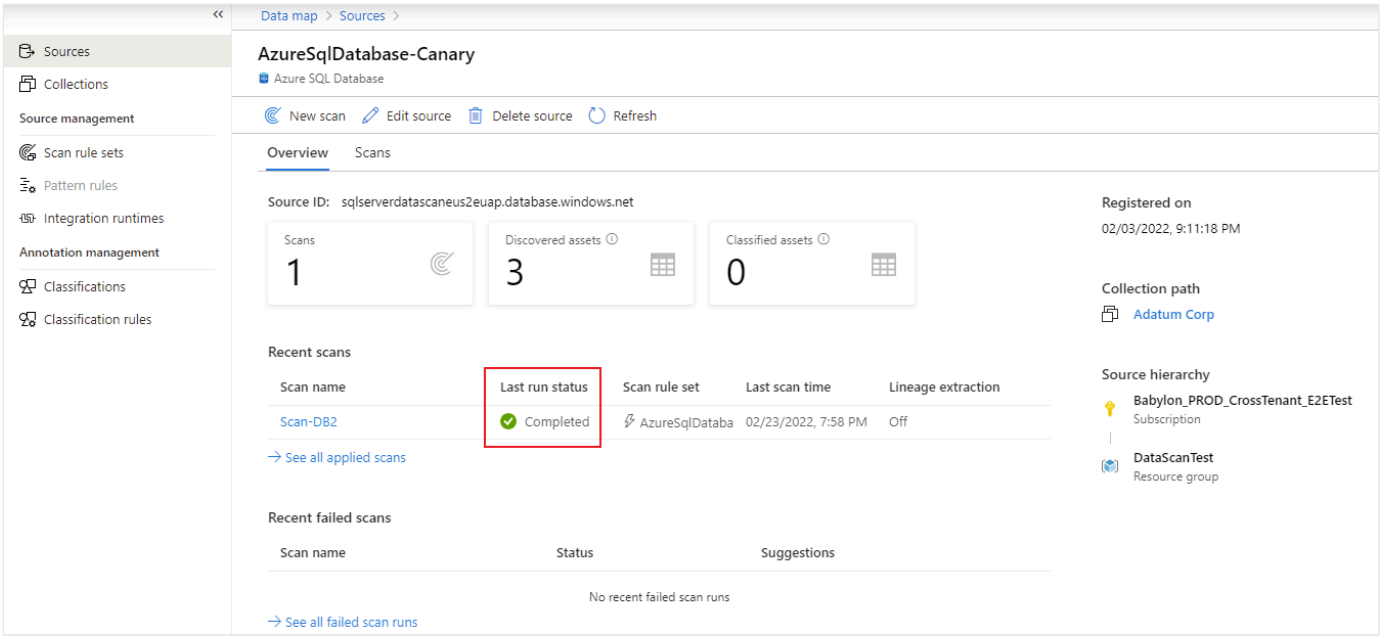
4. Review your scan, and then select **Save and run**.

View a scan

To check the status of a scan, go to the data source in the collection, and then select **View details**.



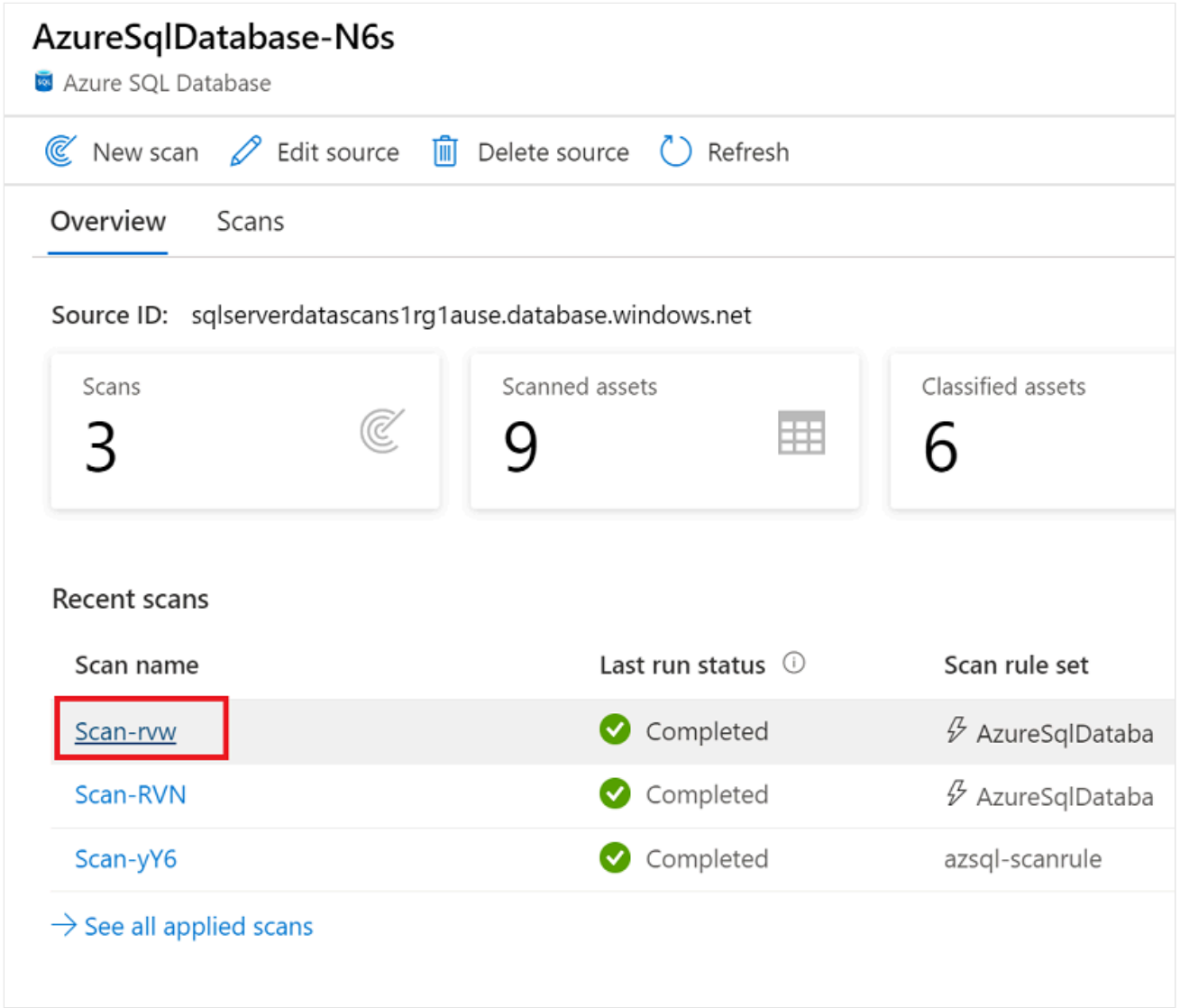
The scan details indicate the progress of the scan in **Last run status**, along with the number of assets scanned and classified. **Last run status** is updated to **In progress** and then **Completed** after the entire scan has run successfully.



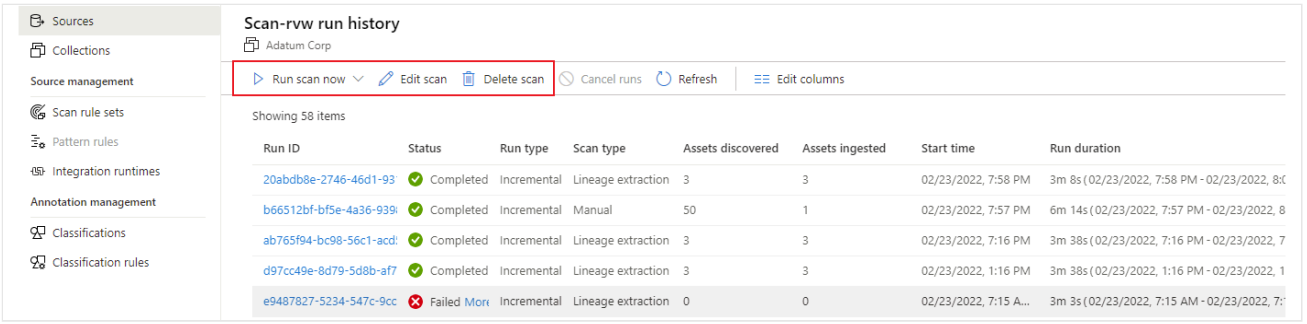
Manage a scan

After you run a scan, you can use the run history to manage it:

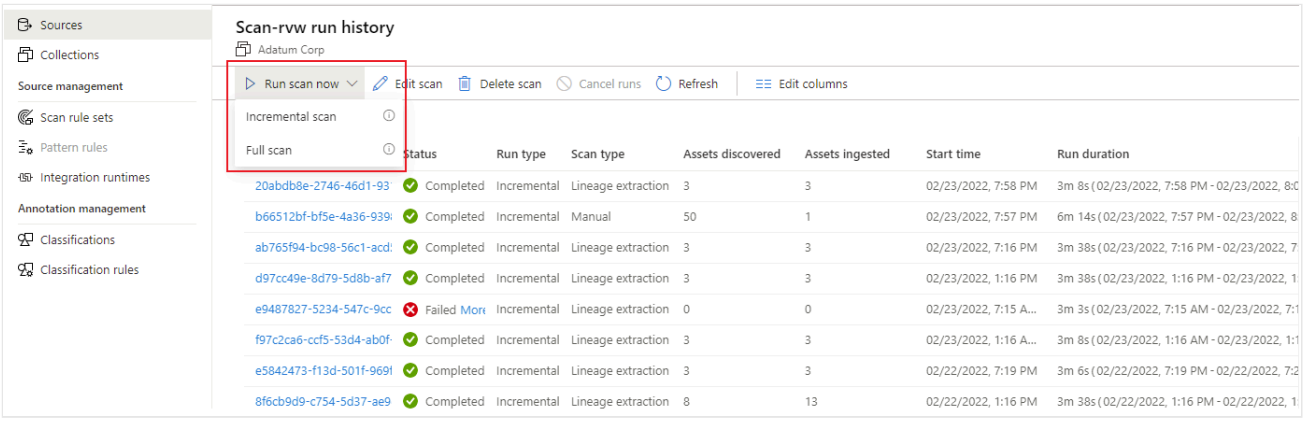
1. Under **Recent scans**, select a scan.



2. In the run history, you have options for running the scan again, editing it, or deleting it.



If you select **Run scan now** to rerun the scan, you can then choose either **Incremental scan** or **Full scan**.



Troubleshoot scanning

If you have problems with scanning, try these tips:

- Confirm that you followed all prerequisites.
- Check the network by confirming firewall, Azure connections, or integration runtime settings.
- Confirm that authentication is properly set up.

For more information, review Troubleshoot your connections in Microsoft Purview.

Set up policies

The following types of Microsoft Purview policies are supported on this data resource:

- Data owner policies - a set of policy statements that allow you to grant users and groups access to data sources.
- Self-service policies - policy that allows users to request access to data sources registered to Microsoft Purview.

- DevOps policies - grants access to database system metadata across multiple sources. They simplify access provisioning for IT operations and security auditing personnel. They only grant access, and don't deny access.

Access policy prerequisites on Azure SQL Database

- Create a new Azure SQL Database instance, or use an existing one, in one of the currently available regions for this feature. You can follow this guide to create an Azure SQL Database instance.

Region support

All Microsoft Purview regions are supported.

Enforcement of Microsoft Purview policies is available only in the following regions for Azure SQL Database:

Public cloud:

- East US
- East US2
- South Central US
- West Central US
- West US3
- Canada Central
- Brazil South
- West Europe
- North Europe
- France Central
- UK South
- South Africa North
- Central India
- South East Asia
- East Asia
- Australia East

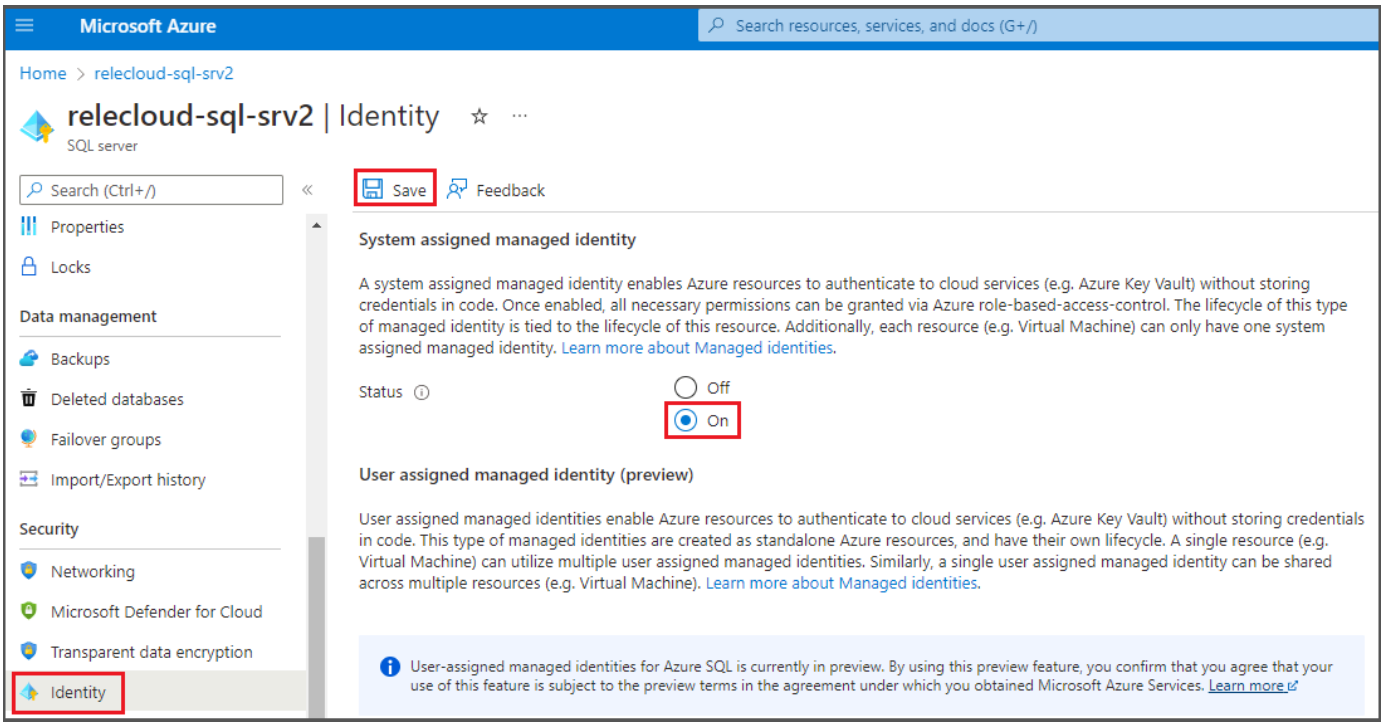
Sovereign clouds:

- USGov Virginia
- China North 3

Configure the Azure SQL Database instance for policies from Microsoft Purview

For the logical server associated with Azure SQL Database to honor policies from Microsoft Purview, you need to configure a Microsoft Entra admin. In the Azure portal, go to the logical server that hosts the Azure SQL Database instance. On the side menu, select **Microsoft Entra ID**. Set an admin name to any Microsoft Entra user or group that you prefer, and then select **Save**.

Then, on the side menu, select **Identity**. Under **System assigned managed identity**, turn the status to **On**, and then select **Save**.



Configure the Microsoft Purview account for policies

Register the data source in Microsoft Purview

Before a policy can be created in Microsoft Purview for a data resource, you must register that data resource in Microsoft Purview Studio. You will find the instructions related to registering the data resource later in this guide.

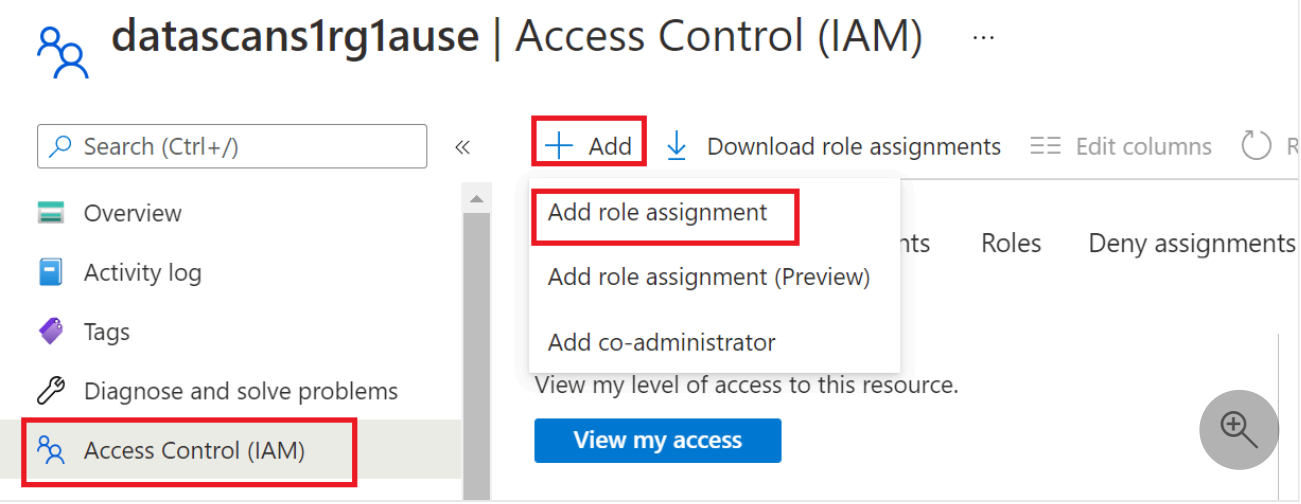
> **Note**
>
> Microsoft Purview policies rely on the data resource ARM path. If a data resource is moved to a new resource group or subscription it will need to be de-registered and then registered again in Microsoft Purview.

Configure permissions to enable Data policy enforcement on the data source

Once a resource is registered, but before a policy can be created in Microsoft Purview for that resource, you must configure permissions. A set of permissions are needed to enable the **Data policy enforcement**. This applies to data sources, resource groups, or subscriptions. To enable *Data policy enforcement*, you must have *both* specific Identity and Access Management (IAM) privileges on the resource as well as specific Microsoft Purview privileges:

- You must have *either one* of the following IAM role combinations on the resource's Azure Resource Manager path or any parent of it (that is, using IAM permission inheritance):
  - IAM Owner
  - Both IAM Contributor and IAM User Access Administrator

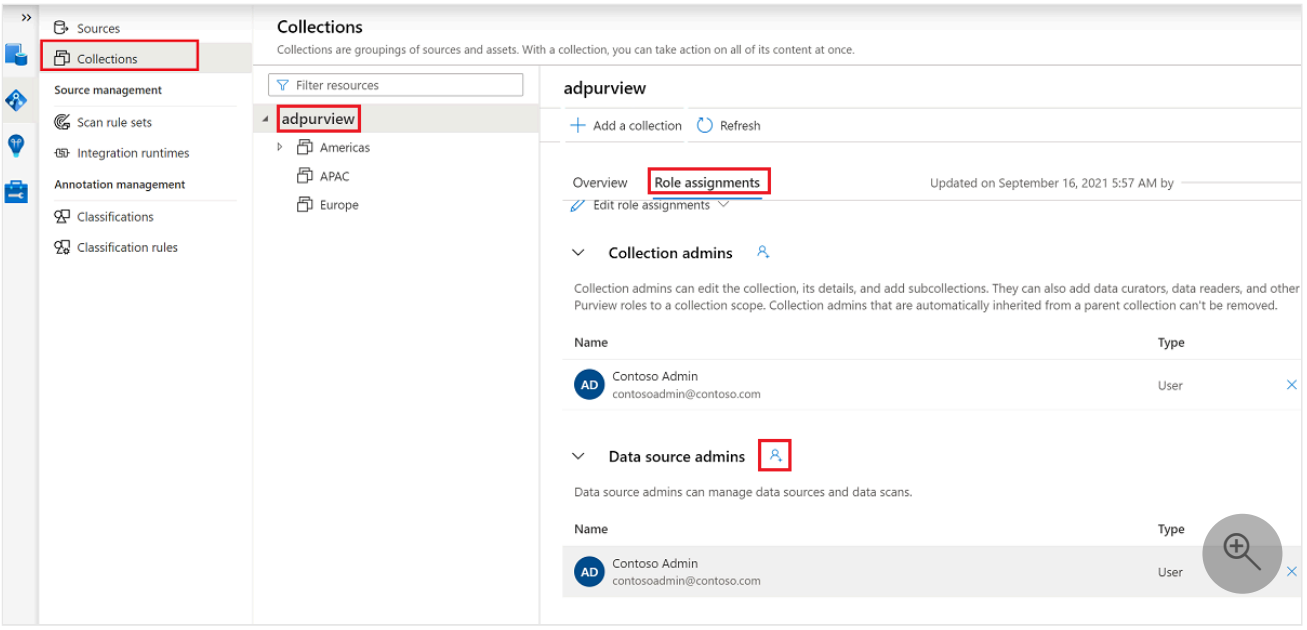
  To configure Azure role-based access control (RBAC) permissions, follow this guide. The following screenshot shows how to access the **Access Control** section in the Azure portal for the data resource to add a role assignment.



> **Note**
>
> The *IAM Owner* role for a data resource can be inherited from a parent resource group, a subscription, or a subscription management group. Check which Microsoft Entra users, groups, and service principals hold or are inheriting the *IAM Owner* role for the resource.

- You also need to have the Microsoft Purview *Data source admin* role for the collection or a parent collection (if inheritance is enabled). For more information, see the guide on managing Microsoft Purview role assignments.

The following screenshot shows how to assign the *Data source admin* role at the root collection level.



Configure Microsoft Purview permissions to create, update, or delete access policies

To create, update or delete policies, you need to get the *Policy author* role in Microsoft Purview at *root collection level*:

- The *Policy author* role can create, update, and delete DevOps and Data Owner policies.
- The *Policy author* role can delete self-service access policies.

For more information about managing Microsoft Purview role assignments, see Create and manage collections in the Microsoft Purview Data Map.

> **Note**
>
> Policy author role must be configured at the root collection level.

In addition, to easily search Microsoft Entra users or groups when creating or updating the subject of a policy, you can greatly benefit from getting the Directory Readers permission in Microsoft Entra ID. This is a common permission for users in an Azure tenant. Without the Directory Reader permission, the Policy Author will have to type the complete username or email for all the principals included in the subject of a data policy.

Configure Microsoft Purview permissions for publishing Data Owner policies

Data Owner policies allow for checks and balances if you assign the Microsoft Purview *Policy author* and *Data source admin* roles to different people in the organization. Before a Data owner policy takes effect, a second person (*Data source admin*) must review it and explicitly approve it by publishing it. This does not apply to DevOps or Self-service access policies as publishing is automatic for them when those policies are created or updated.

To publish a Data owner policy you need to get the *Data source admin* role in Microsoft Purview at *root collection level*.

For more information about managing Microsoft Purview role assignments, see Create and manage collections in the Microsoft Purview Data Map.

> **Note**
>
> To publish Data owner policies, the Data source admin role must be configured at the root collection level.

Delegate access provisioning responsibility to roles in Microsoft Purview

After a resource has been enabled for **Data policy enforcement**, any Microsoft Purview user with the *Policy author* role at the root collection level can provision access to that data source from Microsoft Purview.

> **Note**
>
> Any Microsoft Purview root *Collection admin* can assign new users to root *Policy author* roles. Any *Collection admin* can assign new users to a *Data source admin* role under the collection. Minimize and carefully vet the users who hold Microsoft Purview *Collection admin*, *Data source admin*, or *Policy author* roles.

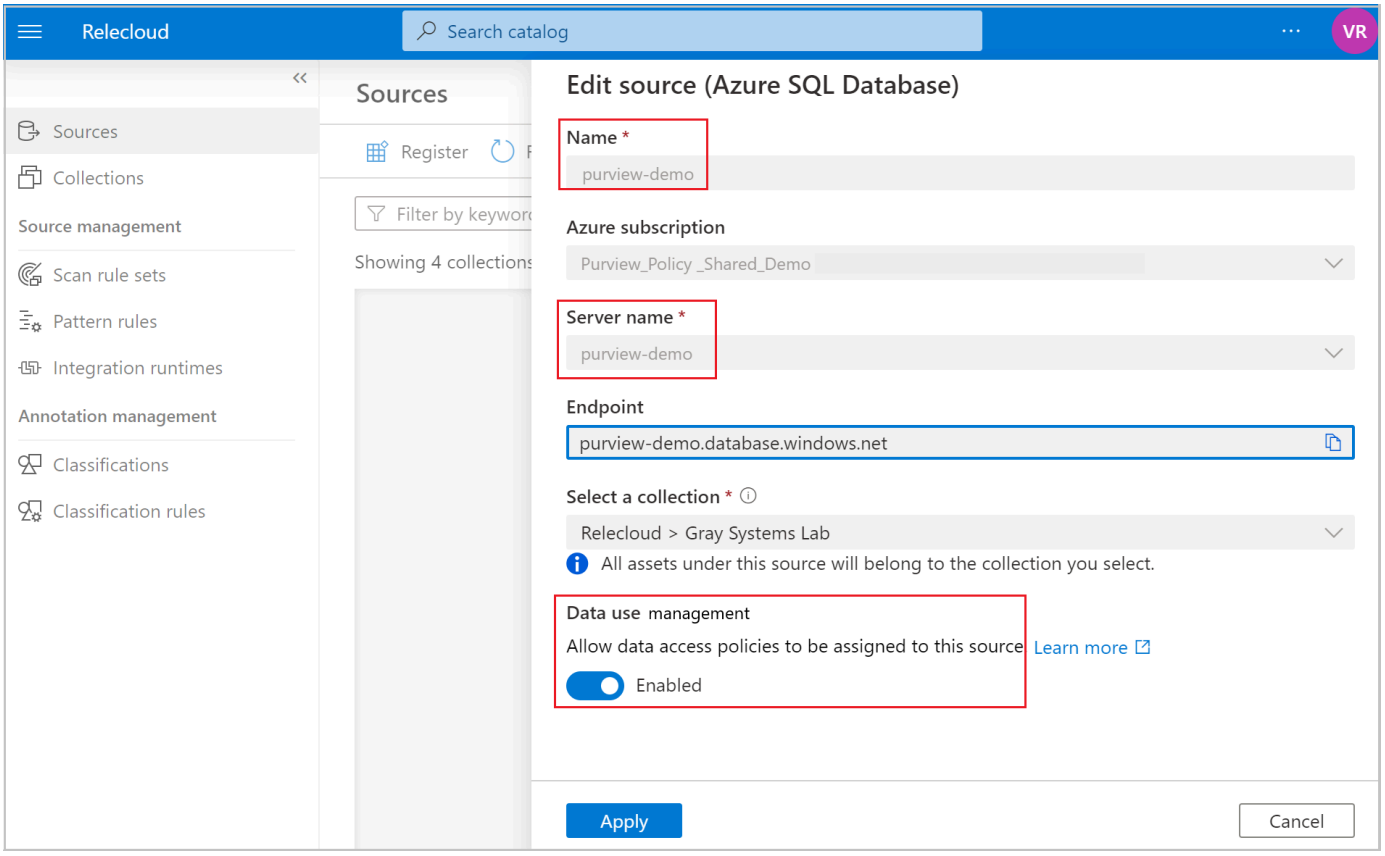
If a Microsoft Purview account with published policies is deleted, such policies will stop being enforced within an amount of time that depends on the specific data source. This change can have implications on both security and data access availability. The Contributor and Owner roles in IAM can delete Microsoft Purview accounts. You can check these permissions by going to the **Access control (IAM)** section for your Microsoft Purview account and selecting **Role Assignments**. You can also use a lock to prevent the Microsoft Purview account from being deleted through Resource Manager locks.

Register the data source and enable Data policy enforcement

The Azure SQL Database resource needs to be registered with Microsoft Purview before you can create access policies. To register your resources, follow the "Prerequisites" and "Register the data source" sections in Enable Data policy enforcement on your Microsoft Purview sources.

After you register the data source, you need to enable **Data policy enforcement**. This is a prerequisite before you can create policies on the data source. **Data policy enforcement** can affect the security of your data, because it delegates to certain Microsoft Purview roles that manage access to the data sources. Go through the security practices in Enable Data policy enforcement on your Microsoft Purview sources.

After your data source has the **Data policy enforcement** option set to **Enabled**, it will look like this screenshot:



Return to the Azure portal for Azure SQL Database to verify it's now governed by Microsoft Purview:

1. Sign in to the Azure portal through this link.
2. Select the Azure SQL Server that you want to configure.
3. Go to **Microsoft Entra ID** on the left pane.

4. Scroll down to **Microsoft Purview access policies**.

5. Select the button to **Check for Microsoft Purview Governance**. Wait while the request is processed, which could take a few minutes.



6. Confirm that the Microsoft Purview Governance Status shows **Governed**. Note that **it could take a few minutes** after you enable *Data policy enforcement* in Microsoft Purview for the correct status to be reflected.

> **Note**
>
> If you disable data policy enforcement for this Azure SQL Database data source in Microsoft Purview, select **Check for Microsoft Purview Governance** for the Microsoft Purview Governance Status to be updated to **Not Governed**. This step must be done for the impacted data source each time data policy enforcement status is changed for an existing or new access policy in Microsoft Purview. Before you enable *Data policy enforcement* for the data source in another Microsoft Purview account, ensure that the Purview Governance Status shows as **Not Governed**. Then repeat the steps above with the new Microsoft Purview account.

Create an access policy

To create an access policy for Azure SQL Database, follow these guides:

- Provision access to system health, performance, and audit information in Azure SQL Database. Use this guide to apply a DevOps policy on a single SQL database.
- Provision read/modify access on a single Azure SQL Database. Use this guide to provision access on a single SQL database account in your subscription.
- Self-service access policies for Azure SQL Database. Use this guide to allow data consumers to request access to data assets by using a self-service workflow.

To create policies that cover all data sources inside a resource group or Azure subscription, see Discover and govern multiple Azure sources in Microsoft Purview.

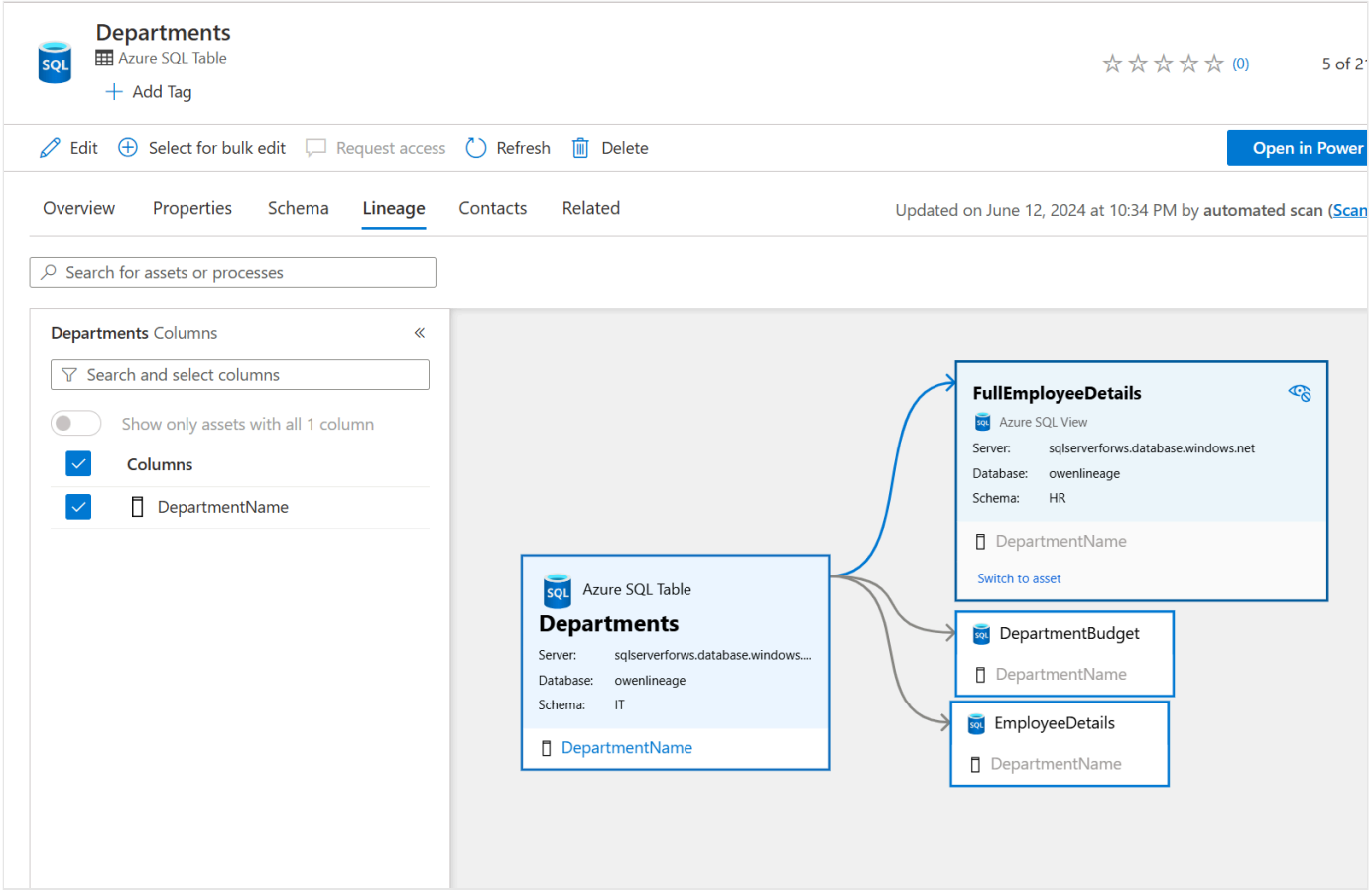
Extract lineage (preview)

Microsoft Purview supports lineage for views and stored procedures from Azure SQL Database. While lineage for views is supported as part of scanning, you'll need to turn on the **Lineage extraction** toggle to extract stored procedure lineage when you're setting up a scan.

> **Note**
>
> Lineage isn't supported using a self-hosted integration runtime or managed VNET runtime and an Azure SQL private endpoint. You need to enable Azure services to access the server under network settings for your Azure SQL Database, and your Microsoft Purview account needs to allow public access. Learn more about the known limitations in lineage extraction scan.

Lineage for SQL DB views

SQL DB metadata scan includes lineage extraction for views. Only new scans include the view lineage extraction. Lineage is extracted at all scan levels (L1/L2/L3). In case of an incremental scan, whatever metadata is scanned as part of incremental scan, the corresponding static lineage for tables/views will be extracted.



Prerequisites for setting up a scan with SP lineage extraction

1. Follow the steps in the Configure authentication for a scan section of this article to authorize Microsoft Purview to scan your SQL database.

2. Sign in to Azure SQL Database with your Microsoft Entra account, and assign `db_owner` permissions to the Microsoft Purview managed identity.

   > **Note**
   >
   > The 'db_owner' permissions is needed because lineage is based on XEvent sessions. So Microsoft Purview needs the permission to manage the XEvent sessions in SQL.

   Use the following example SQL syntax to create a user and grant permission. Replace `<purview-account>` with your account name.

   ```
   SQL

   Create user <purview-account> FROM EXTERNAL PROVIDER
   GO
   EXEC sp_addrolemember 'db_owner', <purview-account>
   GO
   ```

3. Run the following command on your SQL database to create a master key:

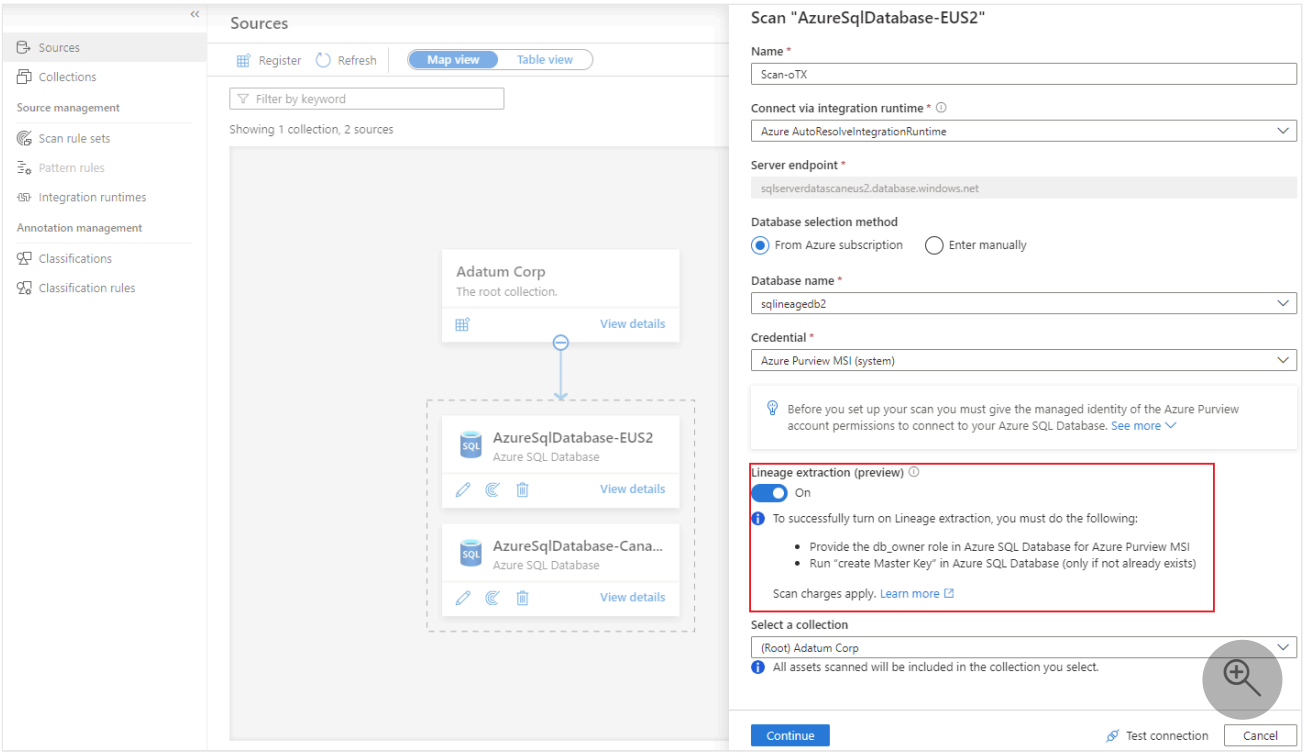
   ```
   SQL

   Create master key
   Go
   ```

4. Ensure that **Allow Azure services and resources to access this server** is enabled under networking/firewall for your Azure SQL resource.

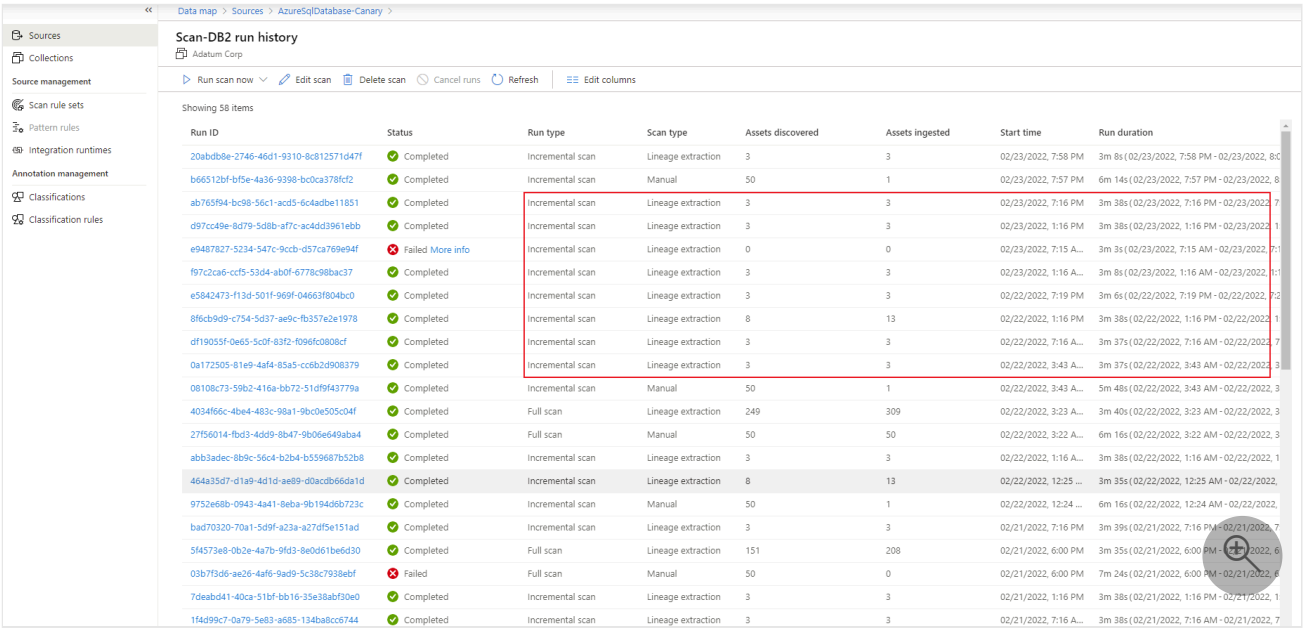
Create a scan with lineage extraction turned on

1. On the pane for setting up a scan, turn on the **Enable lineage extraction** toggle.



2. Select your method of authentication by following the steps in the Create the scan section of this article.

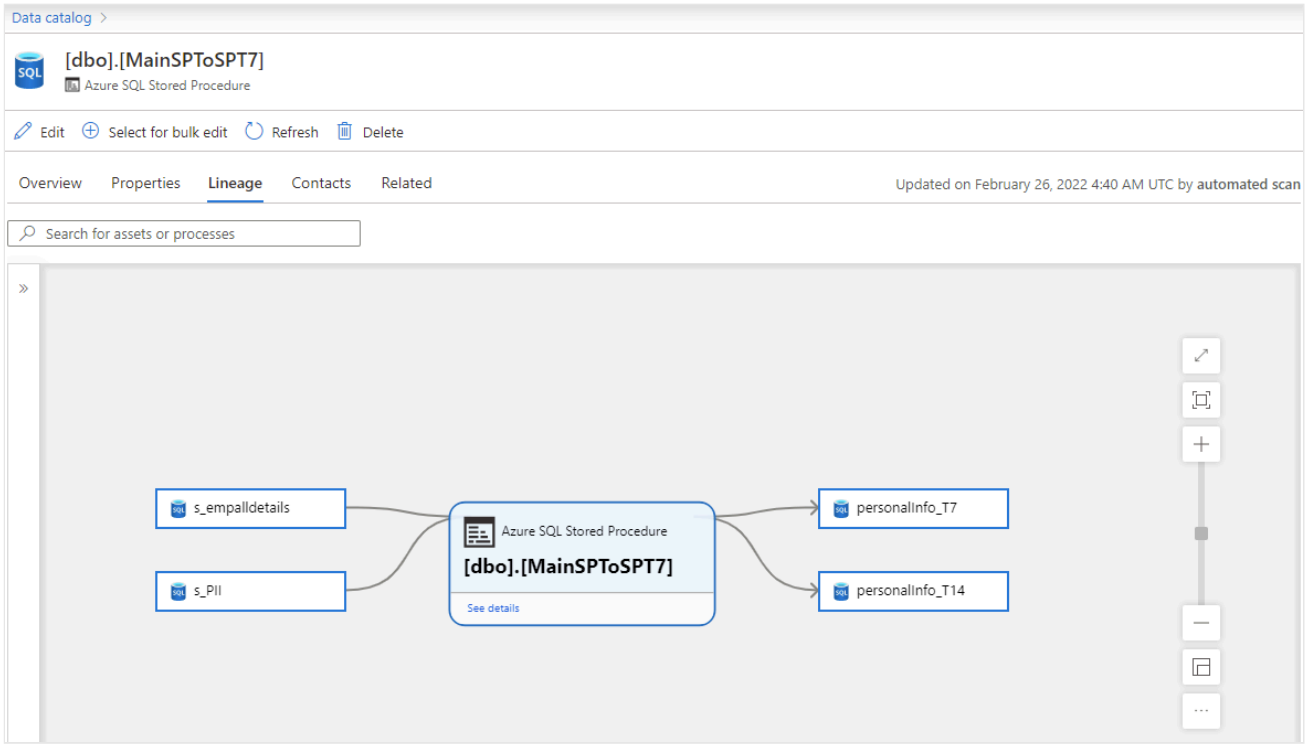
3. After you successfully set up the scan, a new scan type called **Lineage extraction** will run incremental scans **every six hours** to extract lineage from Azure SQL Database. Lineage is extracted based on the stored procedure runs in the SQL database.



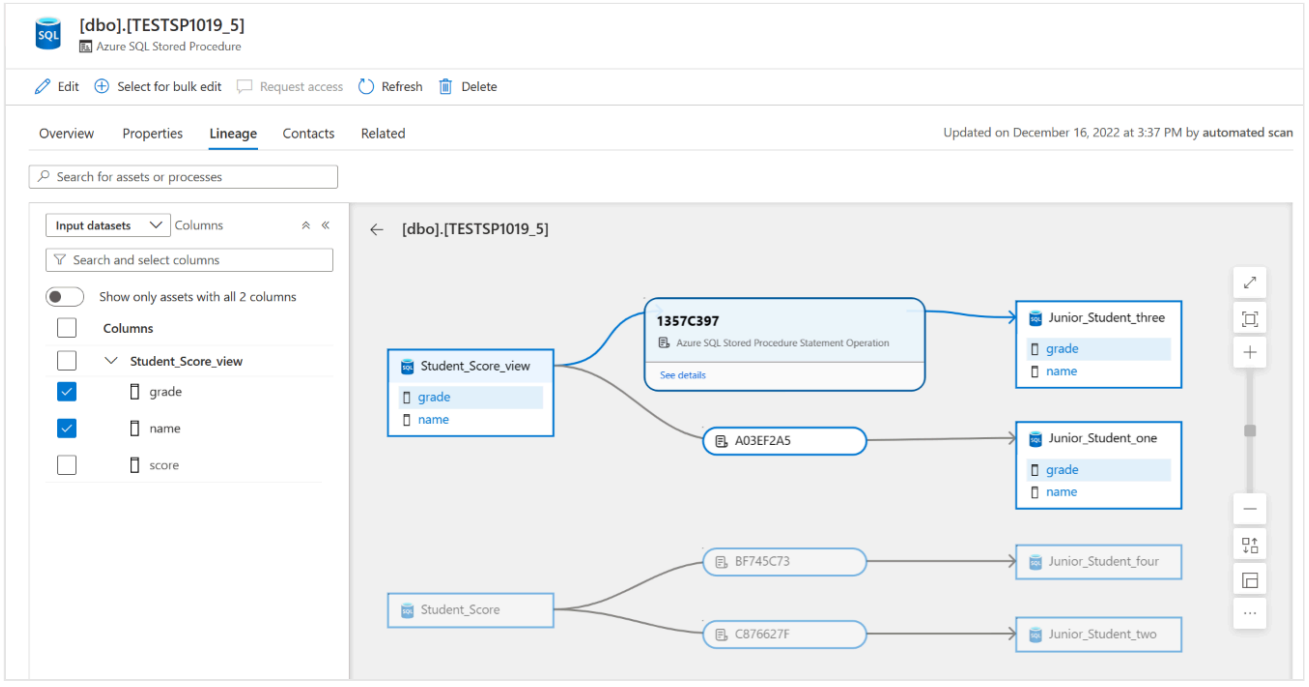
Search Azure SQL Database assets and view runtime lineage

You can browse through the Unified Catalog or search the Unified Catalog to view asset details for Azure SQL Database. The following steps describe how to view runtime lineage details:

1. Go to the **Lineage** tab for the asset. When applicable, the asset lineage appears here.

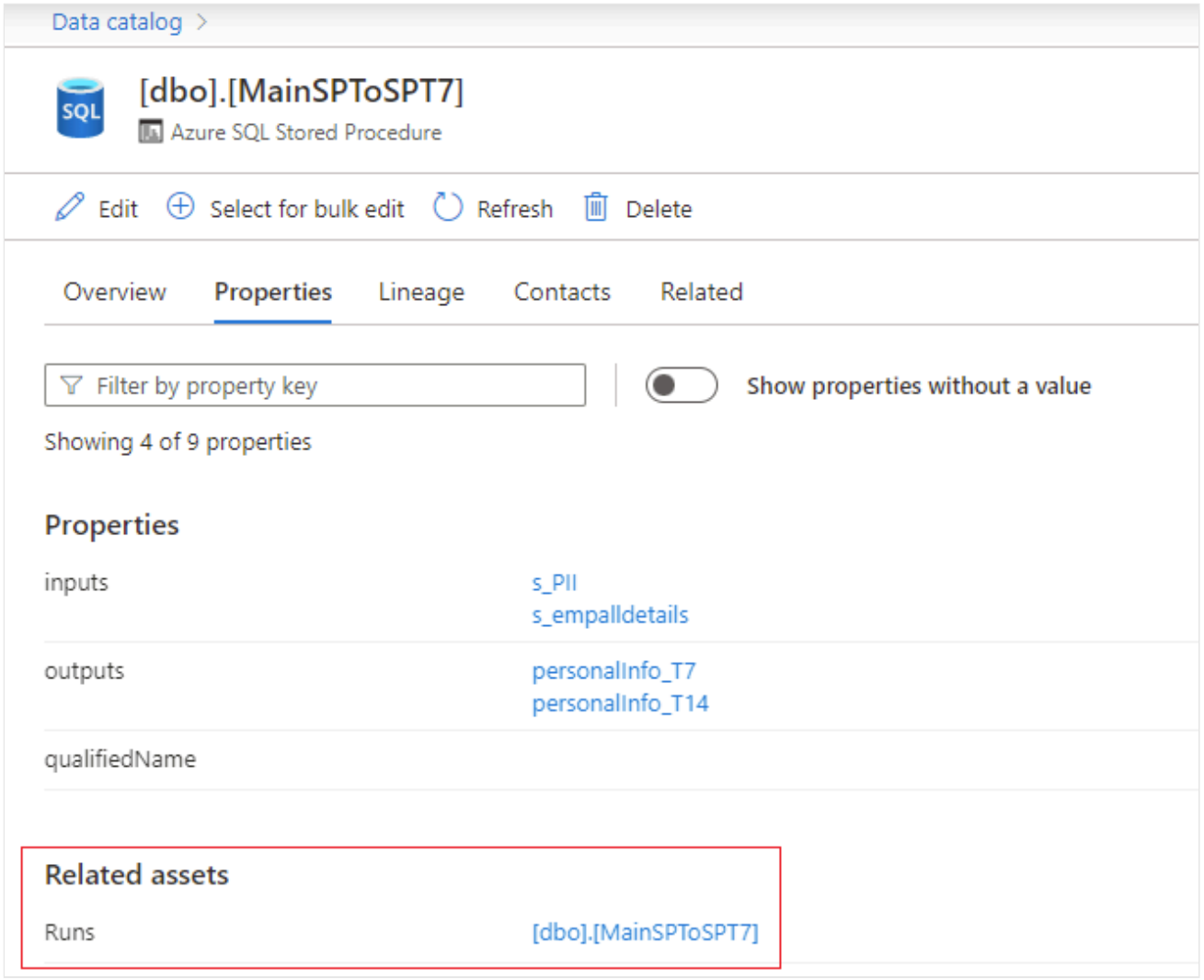


When applicable, you can further drill down to see the lineage at SQL statement level within a stored procedure, along with column level lineage. When using Self-hosted Integration Runtime for scan, retrieving the lineage drilldown information during scan is supported since version 5.25.8374.1.

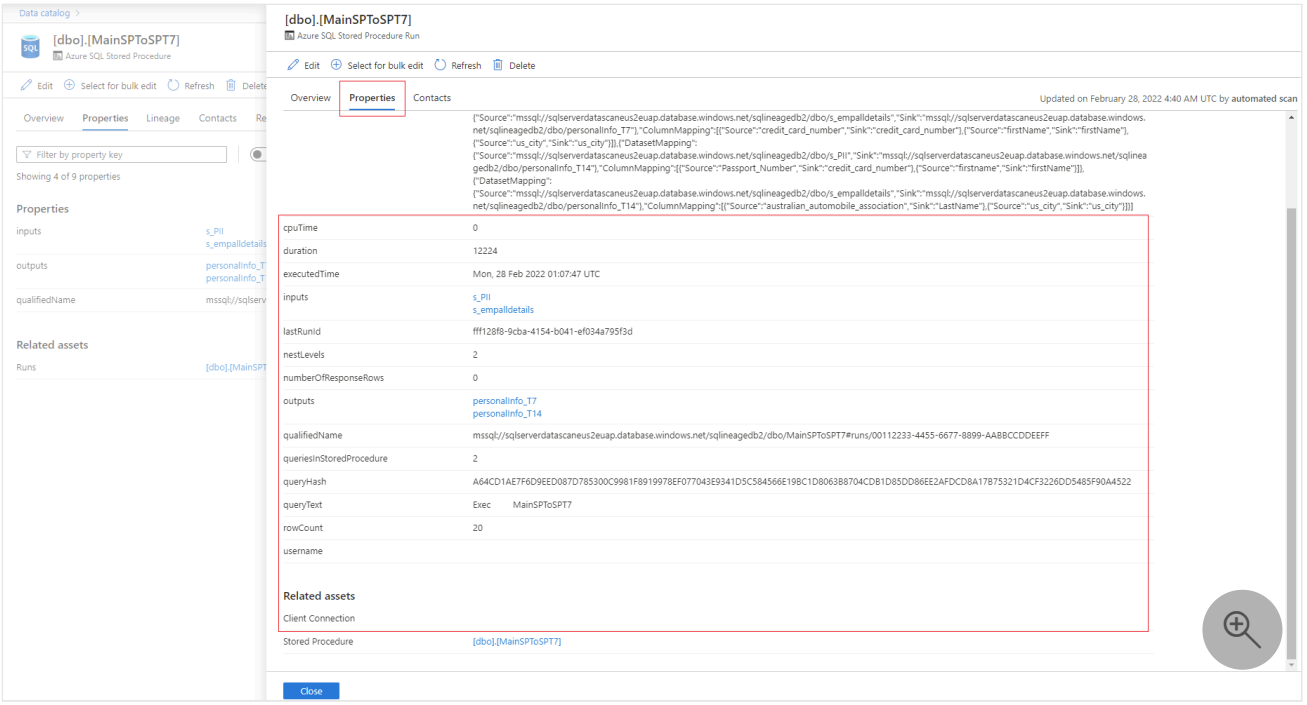


For information about supported Azure SQL Database lineage scenarios, refer to the Supported capabilities section of this article. For more information about lineage in general, see Data lineage in Microsoft Purview and Microsoft Purview lineage user guide.

2. Go to the stored procedure asset. On the **Properties** tab, go to **Related assets** to get the latest run details of stored procedures.



3. Select the stored procedure hyperlink next to **Runs** to see the **Azure SQL Stored Procedure Run** overview. Go to the **Properties** tab to see enhanced runtime information from the stored procedure, such as **executedTime**, **rowCount**, and **Client Connection**.



Troubleshoot lineage extraction for Stored Procedures

The following tips can help you solve problems related to lineage:

- If no lineage is captured after a successful **Lineage extraction** run, it's possible that no stored procedures have run at least once since you set up the scan.
- Lineage is captured for stored procedure runs that happen after a successful scan is set up. Lineage from past stored procedure runs isn't captured.
- If your database is processing heavy workloads with lots of stored procedure runs, lineage extraction filters only the most recent runs. Stored procedure runs early in the six-hour window, or the run instances that create heavy query load, won't be extracted. Contact support if you're missing lineage from any stored procedure runs.
- If a stored procedure contains drop or create statements, they aren't currently captured in lineage

Next steps

To learn more about Microsoft Purview and your data, use these guides:

- Concepts for Microsoft Purview Data owner policies
- Concepts for Microsoft Purview DevOps policies
- Understand the Microsoft Purview Data Estate Insights application
- Microsoft Purview lineage user guide
- Search the Microsoft Purview Unified Catalog

Last updated on 07/01/2026

Connect to and manage an Azure SQL Managed Instance in Microsoft Purview

Article • 06/06/2024

This article outlines how to register and Azure SQL Managed Instance, as well as how to authenticate and interact with the Azure SQL Managed Instance in Microsoft Purview. For more information about Microsoft Purview, read the introductory article.

Supported capabilities

Scanning capabilities

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan
Yes	Yes	Yes	Yes

Other capabilities

View a table listing the supported capabilities for classifications, sensitivity labels, policies, data lineage, and live view.

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- An active Microsoft Purview account.
- You'll need to be a Data Source Administrator and Data Reader to register a source and manage it in the Microsoft Purview governance portal. See our Microsoft Purview Permissions page for details.
- Configure public endpoint in Azure SQL Managed Instance

> **Note**
>
> We now support scanning Azure SQL Managed Instances over the private connection using Microsoft Purview ingestion private endpoints and a self-

hosted integration runtime VM. For more information related to prerequisites, see Connect to your Microsoft Purview and scan data sources privately and securely

Register

This section describes how to register an Azure SQL Managed Instance in Microsoft Purview using the Microsoft Purview governance portal.

Authentication for registration

If you need to create new authentication, you need to authorize database access to SQL Database SQL Managed Instance. There are three authentication methods that Microsoft Purview supports today:

- System or user assigned managed identity
- Service Principal
- SQL authentication

System or user assigned managed identity to register

> **Note**
>
> If you're using Private Endpoints to connect to Microsoft Purview, managed identity isn't supported.

You can use either your Microsoft Purview system-assigned managed identity (SAMI), or a user-assigned managed identity (UAMI) to authenticate. Both options allow you to assign authentication directly to Microsoft Purview, like you would for any other user, group, or service principal. The Microsoft Purview system-assigned managed identity is created automatically when the account is created and has the same name as your Microsoft Purview account. A user-assigned managed identity is a resource that can be created independently. To create one, you can follow our user-assigned managed identity guide.

You can find your managed identity Object ID in the Azure portal by following these steps:

For Microsoft Purview account's system-assigned managed identity:

1. Open the Azure portal, and navigate to your Microsoft Purview account.
2. Select the **Properties** tab on the left side menu.
3. Select the **Managed identity object ID** value and copy it.

For user-assigned managed identity (preview):

1. Open the Azure portal, and navigate to your Microsoft Purview account.
2. Select the Managed identities tab on the left side menu
3. Select the user assigned managed identities, select the intended identity to view the details.
4. The object (principal) ID is displayed in the overview essential section.

Either managed identity will need permission to get metadata for the database, schemas and tables, and to query the tables for classification.

- Create a Microsoft Entra user in Azure SQL Managed Instance by following the prerequisites and tutorial on Create contained users mapped to Microsoft Entra identities
- Assign `db_datareader` permission to the identity.

Service Principal to register

There are several steps to allow Microsoft Purview to use service principal to scan your Azure SQL Managed Instance.

Create or use an existing service principal

To use a service principal, you can use an existing one or create a new one. If you're going to use an existing service principal, skip to the next step. If you have to create a new Service Principal, follow these steps:

1. Navigate to the Azure portal.
2. Select **Microsoft Entra ID** from the left-hand side menu.
3. Select **App registrations**.
4. Select **+ New application registration**.
5. Enter a name for the **application** (the service principal name).
6. Select **Accounts in this organizational directory only**.
7. For Redirect URI, select **Web** and enter any URL you want; it doesn't have to be real or work.
8. Then select **Register**.

Configure Microsoft Entra authentication in the database account

The service principal must have permission to get metadata for the database, schemas, and tables. It must also be able to query the tables to sample for classification.

- Configure and manage Microsoft Entra authentication with Azure SQL
- Create a Microsoft Entra user in Azure SQL Managed Instance by following the prerequisites and tutorial on Create contained users mapped to Microsoft Entra identities
- Assign `db_datareader` permission to the identity.

Add service principal to key vault and Microsoft Purview's credential

It's required to get the service principal's application ID and secret:

1. Navigate to your Service Principal in the Azure portal
2. Copy the values the **Application (client) ID** from **Overview** and **Client secret** from **Certificates & secrets**.
3. Navigate to your key vault
4. Select **Settings** > **Secrets**
5. Select + **Generate/Import** and enter the **Name** of your choice and **Value** as the **Client secret** from your Service Principal
6. Select **Create** to complete
7. If your key vault isn't connected to Microsoft Purview yet, you'll need to create a new key vault connection
8. Finally, create a new credential using the Service Principal to set up your scan.

SQL authentication to register

> **Note**
>
> Only the server-level principal login (created by the provisioning process) or members of the `loginmanager` database role in the master database can create new logins. It takes about **15 minutes** after granting permission, the Microsoft Purview account should have the appropriate permissions to be able to scan the resource(s).

You can follow the instructions in CREATE LOGIN to create a login for Azure SQL Managed Instance if you don't have this login available. You'll need **username** and **password** for the next steps.

1. Navigate to your key vault in the Azure portal

2. Select **Settings** > **Secrets**
3. Select + **Generate/Import** and enter the **Name** and **Value** as the *password* from your Azure SQL Managed Instance
4. Select **Create** to complete
5. If your key vault isn't connected to Microsoft Purview yet, you'll need to create a new key vault connection
6. Finally, create a new credential using the **username** and **password** to set up your scan.

Steps to register

1. Open the Microsoft Purview governance portal by:

    - Browsing directly to https://web.purview.azure.com and selecting your Microsoft Purview account.
    - Opening the Azure portal, searching for and selecting the Microsoft Purview account. Select the **the Microsoft Purview governance portal** button.

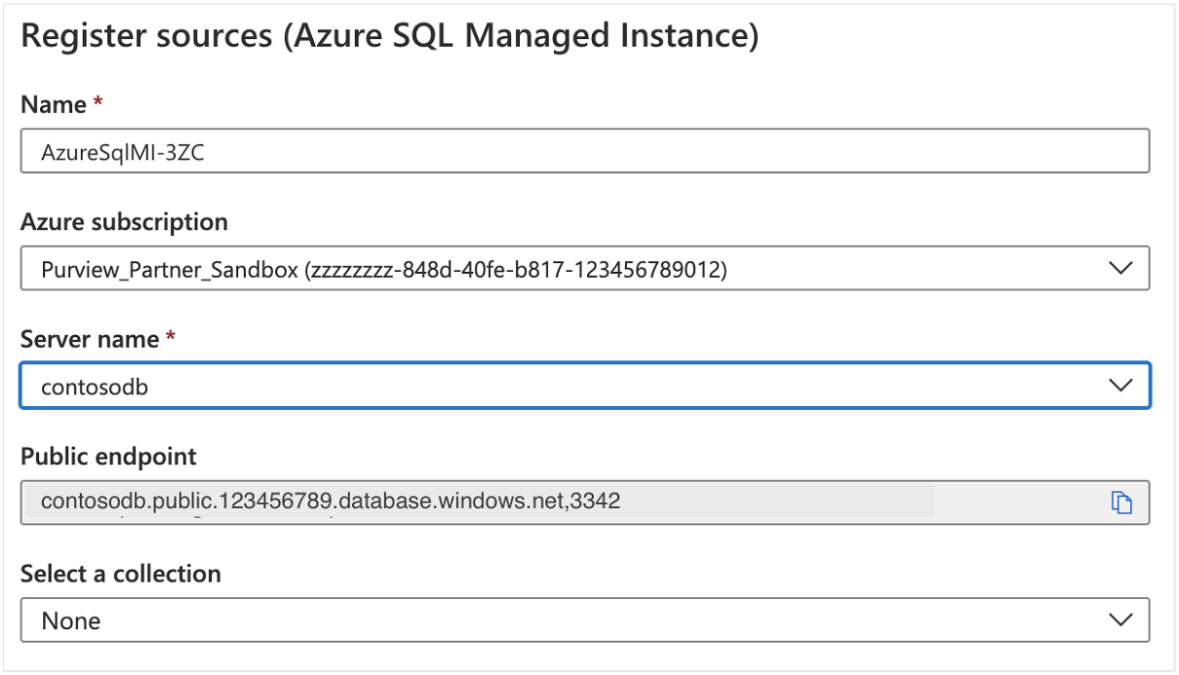
2. Navigate to the **Data Map**.

3. Select **Register**

4. Select **Azure SQL Managed Instance** and then **Continue**.

5. Select **From Azure subscription**, select the appropriate subscription from the **Azure subscription** drop-down box and the appropriate server from the **Server name** drop-down box.

6. Provide the **public endpoint fully qualified domain name** and **port number**. Then select **Register** to register the data source.



For Example: `foobar.public.123.database.windows.net,3342`

Scan

Follow the steps below to scan an Azure SQL Managed Instance to automatically identify assets and classify your data. For more information about scanning in general, see our introduction to scans and ingestion.

Create and run scan

To create and run a new scan, complete the following steps:

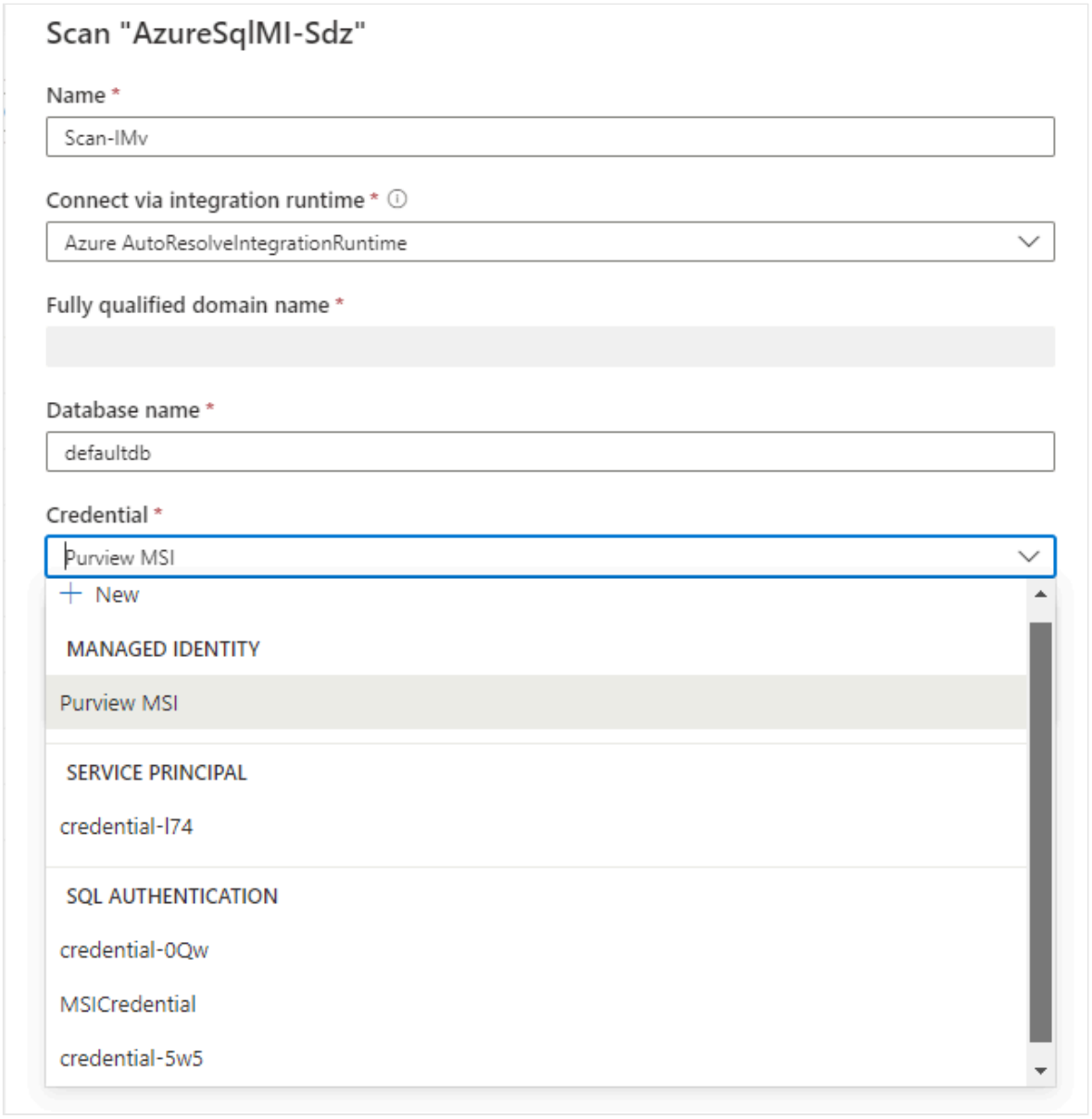
1. Select the **Data Map** tab on the left pane in the Microsoft Purview governance portal.

2. Select the Azure SQL Managed Instance source that you registered.

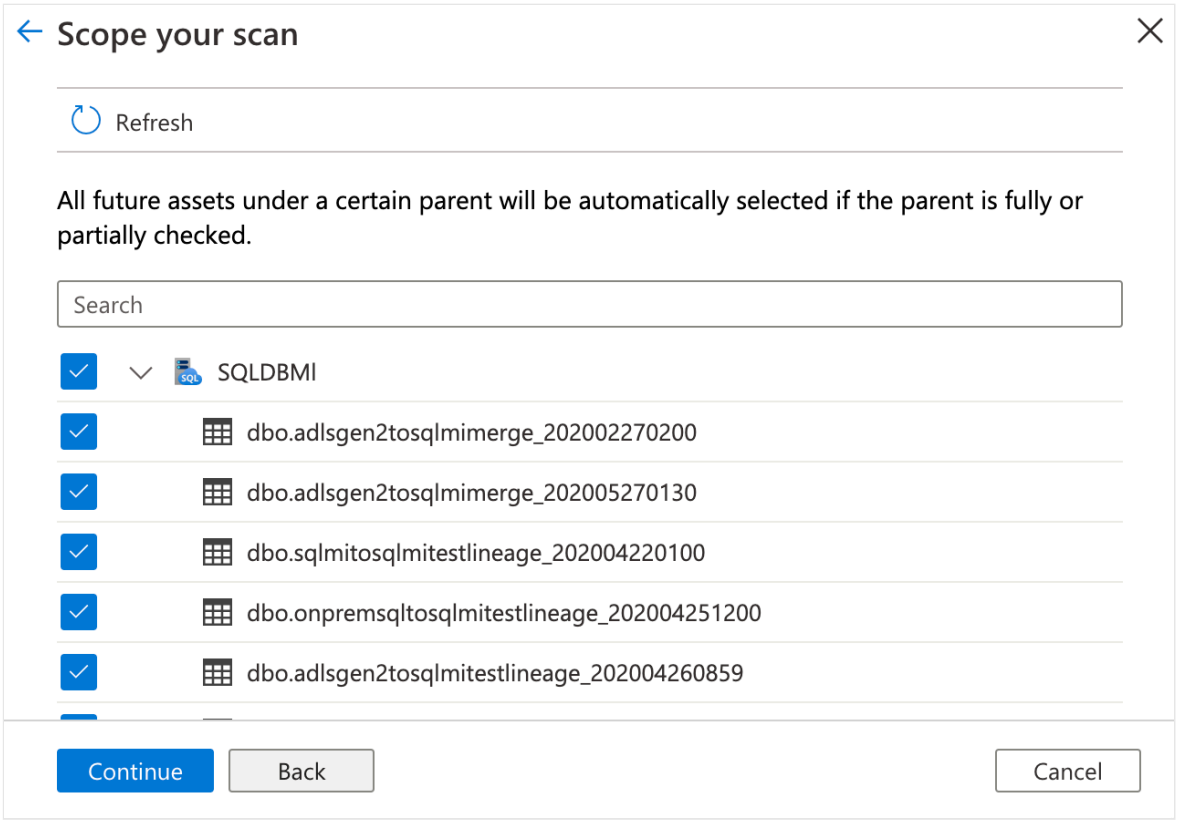
3. Select **New scan**

4. Choose either the Azure integration runtime if your source is publicly accessible, a managed virtual network integration runtime if using a managed virtual network, or a standard or kubernetes supported self-hosted integration runtime if your source is in a private virtual network. For more information about which integration runtime to use, see the choose the right integration runtime configuration article.

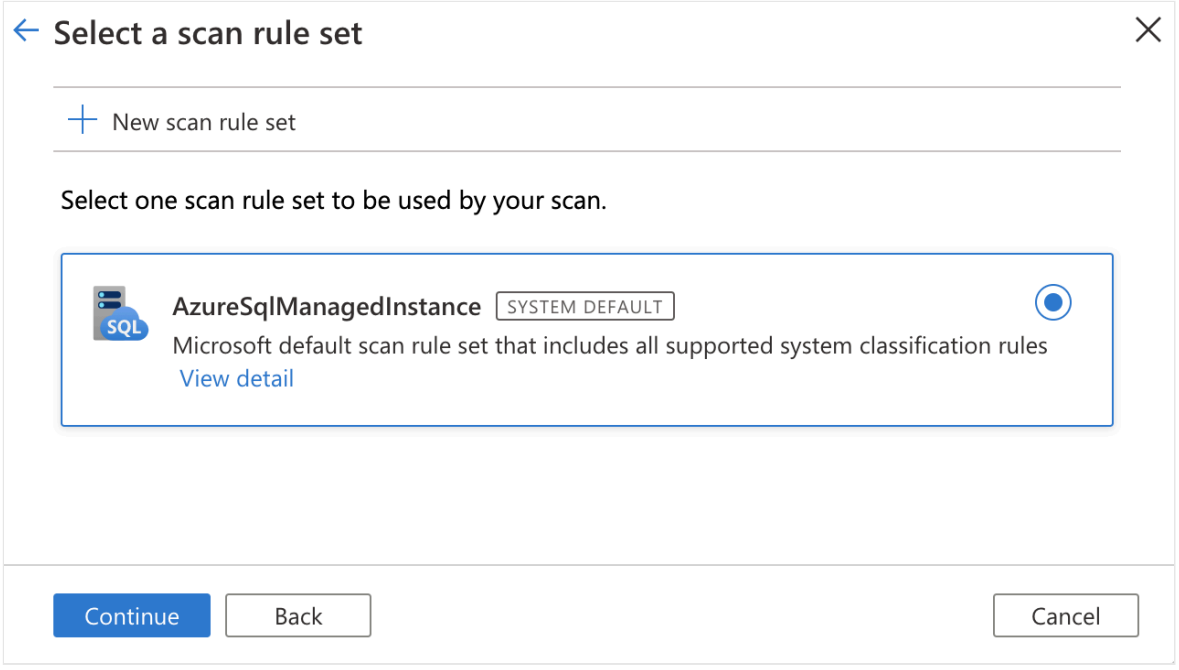
5. Select the credential to connect to your data source.



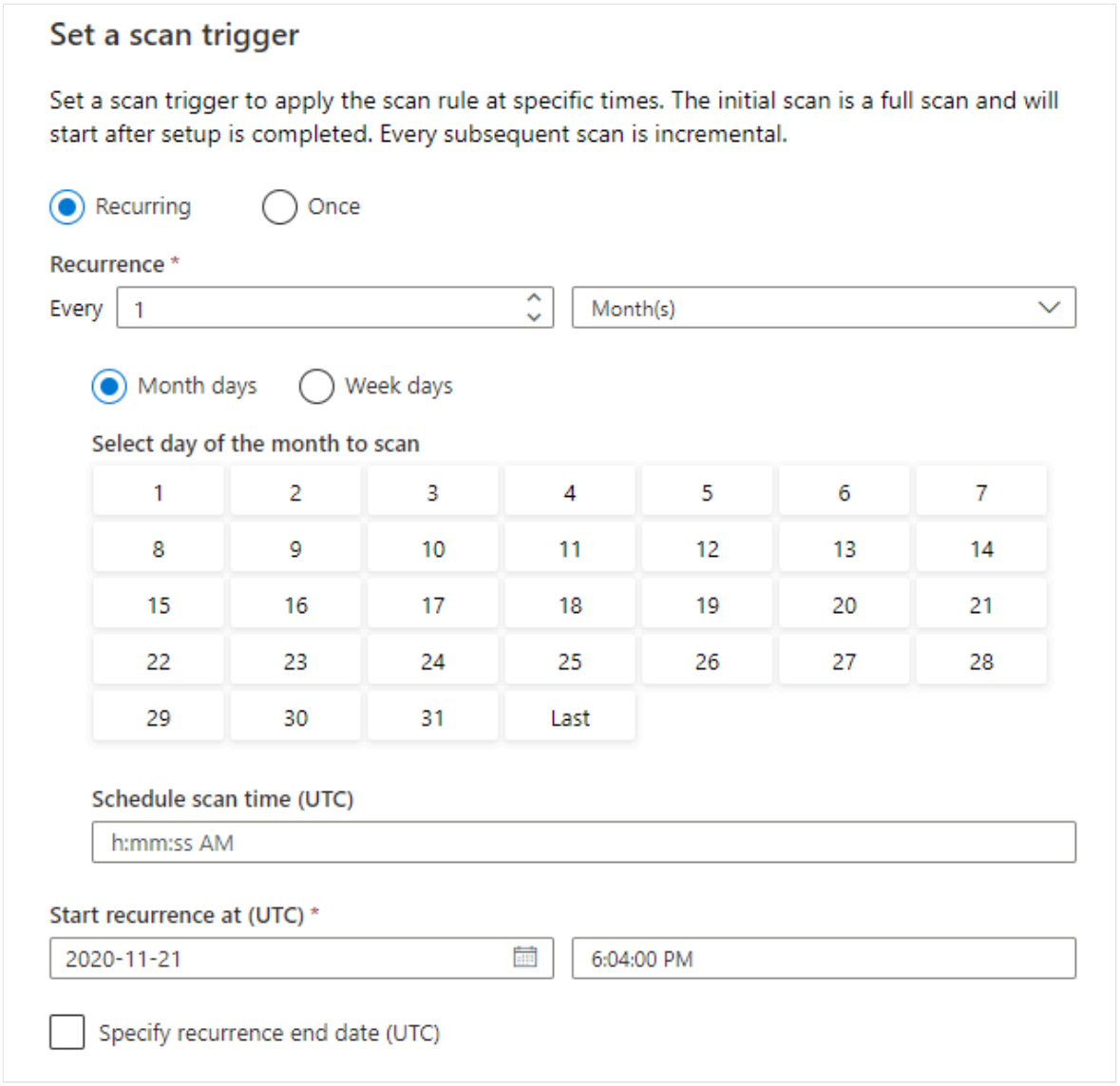
6. You can scope your scan to specific tables by choosing the appropriate items in the list.



7. Then select a scan rule set. You can choose between the system default, existing custom rule sets, or create a new rule set inline.



8. Choose your scan trigger. You can set up a schedule or run the scan once.



9. Review your scan and select **Save and run**.

If you're having trouble connecting to your data source, or running your scan, you scan see our troubleshooting guide for scans and connections.

View your scans and scan runs

To view existing scans:

1. Go to the Microsoft Purview portal. On the left pane, select **Data map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that has results you want to view. The pane shows you all the previous scan runs, along with the status and metrics for each scan run.
4. Select the run ID to check the scan run details.

Manage your scans

To edit, cancel, or delete a scan:

1. Go to the Microsoft Purview portal. On the left pane, select **Data Map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that you want to manage. You can then:
    - Edit the scan by selecting **Edit scan**.
    - Cancel an in-progress scan by selecting **Cancel scan run**.
    - Delete your scan by selecting **Delete scan**.

> ⓘ **Note**
>
> - Deleting your scan does not delete catalog assets created from previous scans.

Set up access policies

The following types of Microsoft Purview policies are supported on this data resource:

- DevOps policies

Access policy prerequisites on Azure SQL MI

- Create a new Azure SQL MI or use an existing one in one of the currently available regions for this feature. You can follow this guide to create a new Azure SQL MI.

Region support

All Microsoft Purview regions are supported.

Enforcement of Microsoft Purview policies is available only in the following regions for Azure SQL MI:

Public cloud:

- East US
- East US2
- South Central US
- West Central US

- West US3
- Canada Central
- Brazil South
- West Europe
- North Europe
- France Central
- UK South
- South Africa North
- Central India
- South East Asia
- East Asia
- Australia East

Azure SQL MI configuration

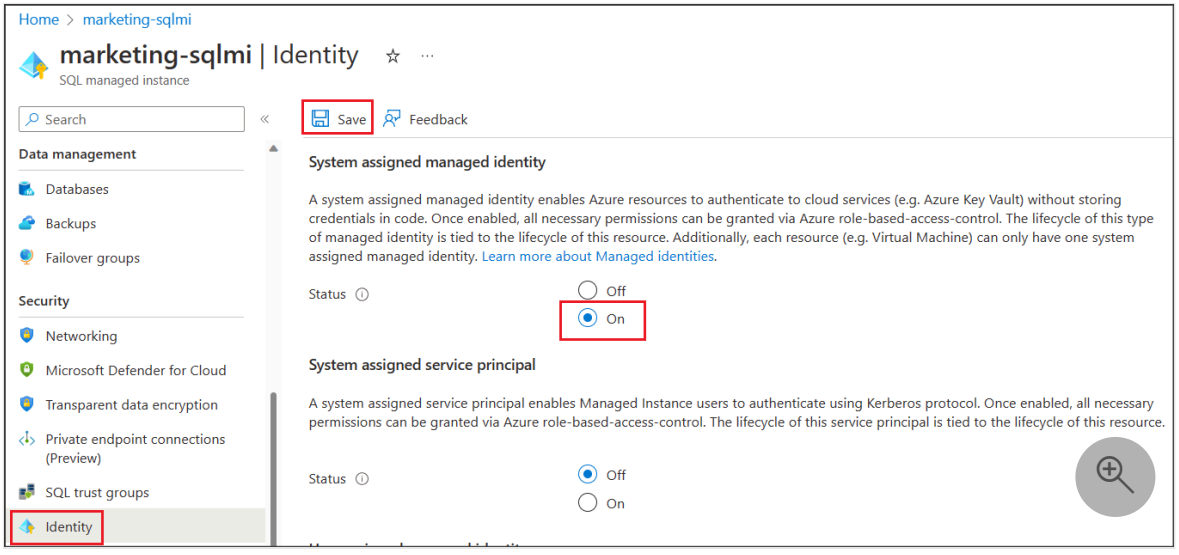
This section explains how you can set up Azure SQL MI to honor policies from Microsoft Purview. Check first whether Azure SQL MI is configured for public or private endpoint. This guide explains how to do so.

Configuration for SQL MI public endpoint

If Azure SQL MI is configured for public endpoint, follow these steps:

- Configure a Microsoft Entra Admin. In Azure portal navigate to the Azure SQL MI then navigate to Microsoft Entra ID on the side menu (formerly called Active Directory admin). Set an Admin name and then select **Save**.

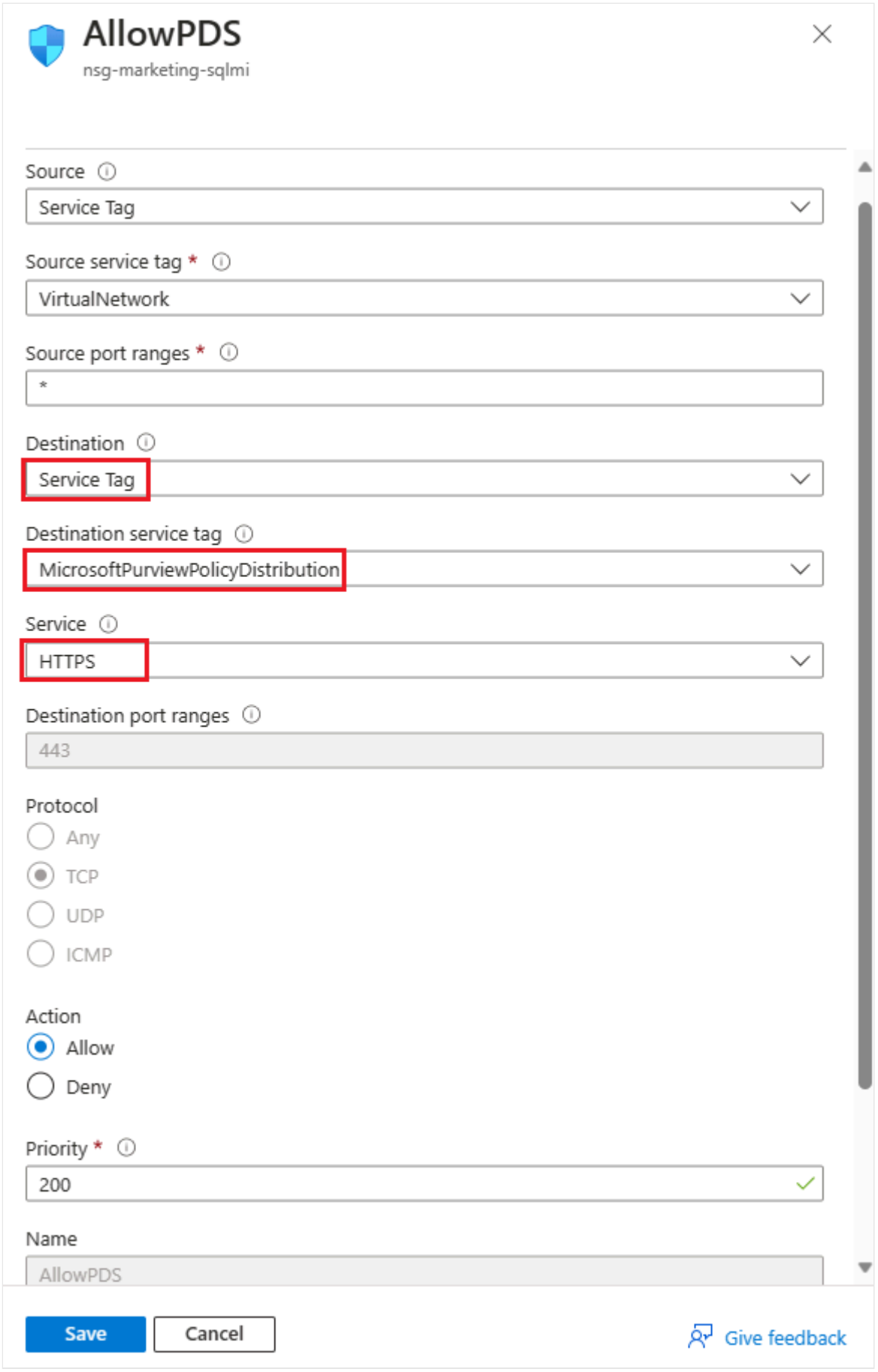
- Then navigate to Identity on the side menu. Under System assigned managed identity check status to *On* and then select **Save**. See screenshot:



Configuration for SQL MI private endpoint

If Azure SQL MI is configured to use private endpoint, execute the same steps outlined in the configuration for public endpoint, and in addition, do the following:

- Navigate to the Network Security Group (NSG) associated to your Azure SQL MI.
- Add an outbound security rule similar to the one in the following screenshot. *Destination = Service Tag, Destination service tag = MicrosoftPurviewPolicyDistribution, Service = HTTPS, Action = Allow.* Ensure also that this rule's priority is lower than the one for the *deny_all_outbound* rule.



Configure the Microsoft Purview account for policies

Register the data source in Microsoft Purview

Before a policy can be created in Microsoft Purview for a data resource, you must register that data resource in Microsoft Purview Studio. You will find the instructions related to registering the data resource later in this guide.

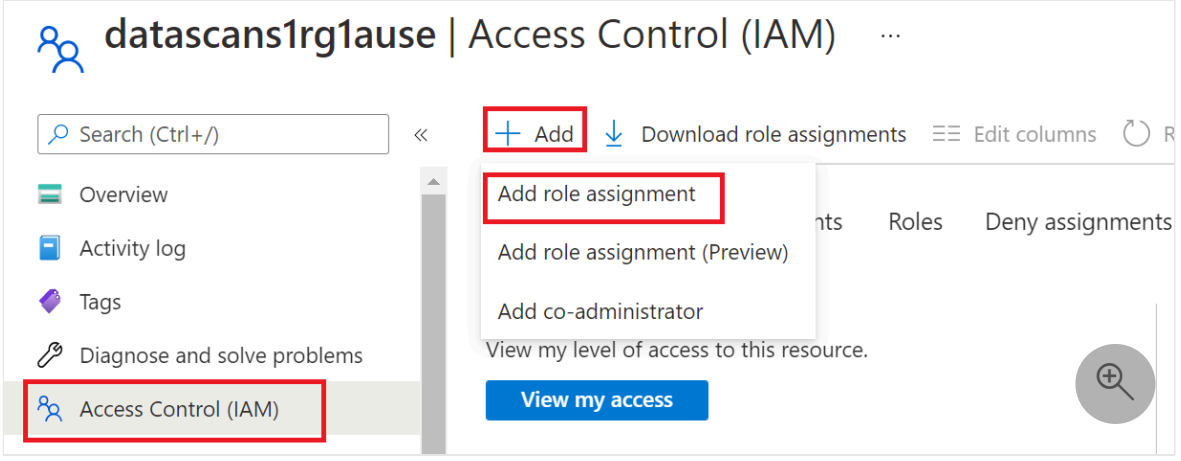
> ⓘ Note
>
> Microsoft Purview policies rely on the data resource ARM path. If a data resource is moved to a new resource group or subscription it will need to be de-registered and then registered again in Microsoft Purview.

Configure permissions to enable Data policy enforcement on the data source

Once a resource is registered, but before a policy can be created in Microsoft Purview for that resource, you must configure permissions. A set of permissions are needed to enable the **Data policy enforcement**. This applies to data sources, resource groups, or subscriptions. To enable *Data policy enforcement*, you must have *both* specific Identity and Access Management (IAM) privileges on the resource as well as specific Microsoft Purview privileges:

- You must have *either one* of the following IAM role combinations on the resource's Azure Resource Manager path or any parent of it (that is, using IAM permission inheritance):
  - IAM Owner
  - Both IAM Contributor and IAM User Access Administrator

  To configure Azure role-based access control (RBAC) permissions, follow this guide. The following screenshot shows how to access the **Access Control** section in the Azure portal for the data resource to add a role assignment.

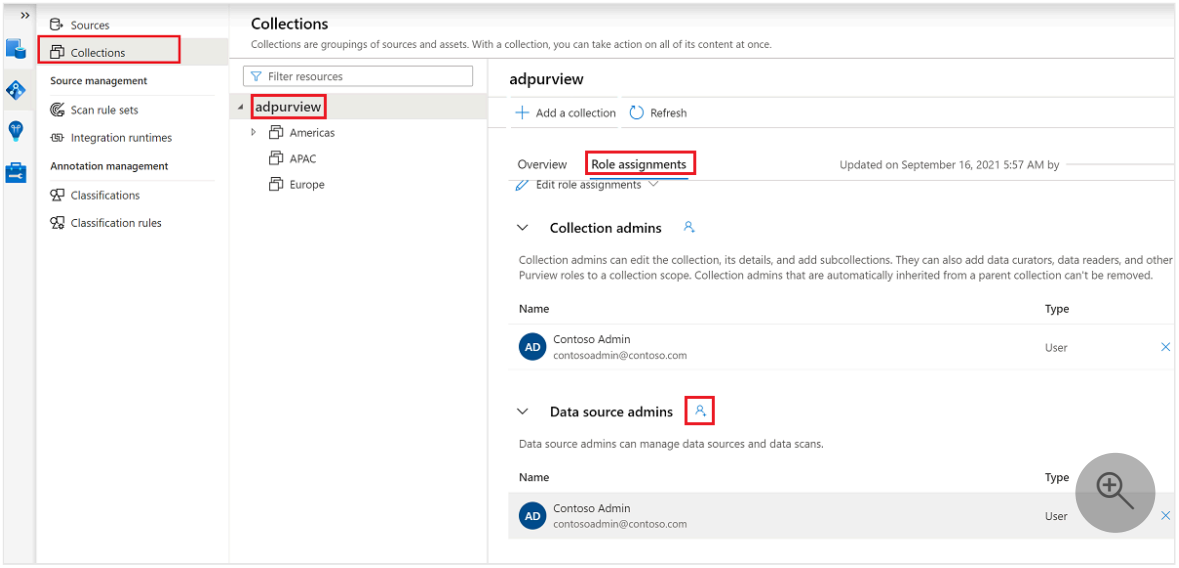


> ⓘ Note

The *IAM Owner* role for a data resource can be inherited from a parent resource group, a subscription, or a subscription management group. Check which Microsoft Entra users, groups, and service principals hold or are inheriting the *IAM Owner* role for the resource.

- You also need to have the Microsoft Purview *Data source admin* role for the collection or a parent collection (if inheritance is enabled). For more information, see the guide on managing Microsoft Purview role assignments.

  The following screenshot shows how to assign the *Data source admin* role at the root collection level.



Configure Microsoft Purview permissions to create, update, or delete access policies

To create, update or delete policies, you need to get the *Policy author* role in Microsoft Purview at *root collection level*:

- The *Policy author* role can create, update, and delete DevOps and Data Owner policies.
- The *Policy author* role can delete self-service access policies.

For more information about managing Microsoft Purview role assignments, see Create and manage collections in the Microsoft Purview Data Map.

> **Note**
>
> Policy author role must be configured at the root collection level.

In addition, to easily search Microsoft Entra users or groups when creating or updating the subject of a policy, you can greatly benefit from getting the Directory Readers permission in Microsoft Entra ID. This is a common permission for users in an Azure tenant. Without the Directory Reader permission, the Policy Author will have to type the complete username or email for all the principals included in the subject of a data policy.

Configure Microsoft Purview permissions for publishing Data Owner policies

Data Owner policies allow for checks and balances if you assign the Microsoft Purview *Policy author* and *Data source admin* roles to different people in the organization. Before a Data owner policy takes effect, a second person (*Data source admin*) must review it and explicitly approve it by publishing it. This does not apply to DevOps or Self-service access policies as publishing is automatic for them when those policies are created or updated.

To publish a Data owner policy you need to get the *Data source admin* role in Microsoft Purview at *root collection level*.

For more information about managing Microsoft Purview role assignments, see Create and manage collections in the Microsoft Purview Data Map.

> ⓘ **Note**
>
> To publish Data owner policies, the Data source admin role must be configured at the root collection level.

Delegate access provisioning responsibility to roles in Microsoft Purview

After a resource has been enabled for **Data policy enforcement**, any Microsoft Purview user with the *Policy author* role at the root collection level can provision access to that data source from Microsoft Purview.

> ⓘ **Note**
>
> Any Microsoft Purview root *Collection admin* can assign new users to root *Policy author* roles. Any *Collection admin* can assign new users to a *Data source admin*

role under the collection. Minimize and carefully vet the users who hold Microsoft Purview *Collection admin*, *Data source admin*, or *Policy author* roles.

If a Microsoft Purview account with published policies is deleted, such policies will stop being enforced within an amount of time that depends on the specific data source. This change can have implications on both security and data access availability. The Contributor and Owner roles in IAM can delete Microsoft Purview accounts. You can check these permissions by going to the **Access control (IAM)** section for your Microsoft Purview account and selecting **Role Assignments**. You can also use a lock to prevent the Microsoft Purview account from being deleted through Resource Manager locks.

Register the data sources in Microsoft Purview

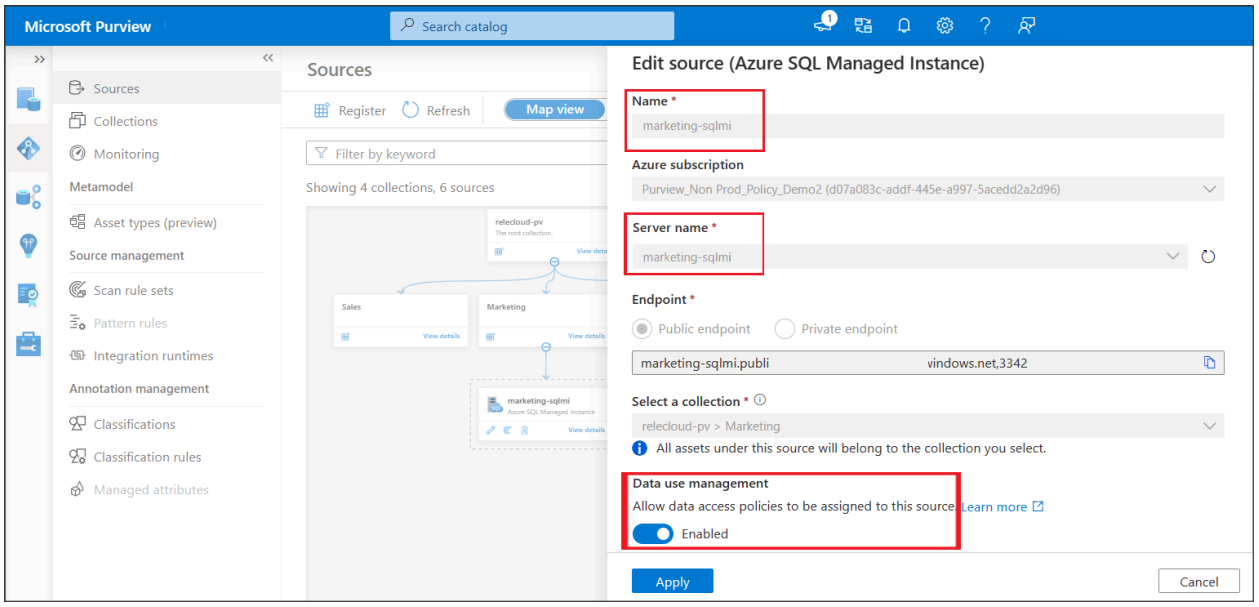
The Azure SQL Managed Instance data source needs to be registered first with Microsoft Purview, before you can create access policies. You can follow the "Prerequisites" and "Register the data source" sections in this guide:

Register and scan Azure SQL MI

After you've registered your resources, you'll need to enable data policy enforcement (formerly Data Use Management). Data policy enforcement needs certain permissions and can affect the security of your data, as it delegates to certain Microsoft Purview roles the ability to manage access to data sources. **Go through the secure practices related to data policy enforcement in this guide**: How to enable data policy enforcement

Once your data source has the **Data policy enforcement** toggle *Enabled*, it will look like this screenshot. This will enable the access policies to be used with the given data

source



Return to the Azure portal for Azure SQL Database to verify it's now governed by Microsoft Purview:

1. Sign in to the Azure portal through this link.
2. Select the Azure SQL Server that you want to configure.
3. Go to **Microsoft Entra ID** on the left pane.
4. Scroll down to **Microsoft Purview access policies**.
5. Select the button to **Check for Microsoft Purview Governance**. Wait while the request is processed, which could take a few minutes.



6. Confirm that the Microsoft Purview Governance Status shows **Governed**. Note that **it could take a few minutes** after you enable *Data policy enforcement* in Microsoft Purview for the correct status to be reflected.

> **Note**
>
> If you disable data policy enforcement for this Azure SQL Database data source in Microsoft Purview, select **Check for Microsoft Purview Governance** for the Microsoft Purview Governance Status to be updated to **Not Governed**. This step must be done for the impacted data source each time data policy enforcement status is changed for an existing or new access policy in Microsoft Purview. Before

you enable *Data policy enforcement* for the data source in another Microsoft Purview account, ensure that the Purview Governance Status shows as **Not Governed**. Then repeat the steps above with the new Microsoft Purview account.

Create a policy

To create access policies for Azure SQL MI, follow these guides:

- To create a DevOps policy on a single Azure SQL MI see Provision access to system health, performance and audit information in Azure SQL MI.
- To create policies that cover all data sources inside a resource group or Azure subscription, see Discover and govern multiple Azure sources in Microsoft Purview.

Next steps

Now that you've registered your source, follow the below guides to learn more about Microsoft Purview and your data.

- Concepts for Microsoft Purview DevOps policies
- Data Estate Insights in Microsoft Purview
- Lineage in Microsoft Purview
- Search Unified Catalog

Feedback

Was this page helpful? Yes No

Connect to and manage Cassandra in Microsoft Purview

Article • 06/27/2024

This article outlines how to register Cassandra, and how to authenticate and interact with Cassandra in Microsoft Purview. For more information about Microsoft Purview, read the introductory article.

Supported capabilities

Scanning capabilities

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan
Yes	Yes	No	Yes

The supported Cassandra server versions are 3.*x* or 4.*x*.

When scanning Cassandra source, Microsoft Purview supports:

- Extracting technical metadata including:
  - Cluster
  - Keyspaces
  - Tables including the columns and indexes
  - Materialized views including the columns
- Fetching static lineage on assets relationships among tables and materialized views.

When setting up scan, you can choose to scan an entire Cassandra instance, or scope the scan to a subset of keyspaces matching the given name(s) or name pattern(s).

Other capabilities

View a table listing the supported capabilities for classifications, sensitivity labels, policies, data lineage, and live view.

Known limitations

When object is deleted from the data source, currently the subsequent scan won't automatically remove the corresponding asset in Microsoft Purview.

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- An active Microsoft Purview account.
- You need Data Source Administrator and Data Reader permissions to register a source and manage it in the Microsoft Purview governance portal. For more information about permissions, see Access control in Microsoft Purview.

> **Note**
>
> **If your data store is not publicly accessible** (if your data store limits access from on-premises network, private network or specific IPs, etc.), **you will need to configure a self hosted integration runtime to connect to it**.

- If your data source isn't publicly accessible, set up the latest self-hosted integration runtime.
  - Choose the right integration runtime for your scenario:
    - **To use a self-hosted integration runtime:**
      - Follow the article to create and configure a self-hosted integration runtime.
      - Ensure JDK 11 is installed on the machine where the self-hosted integration runtime is installed. Restart the machine after you newly install the JDK for it to take effect.
      - Ensure that Visual C++ Redistributable (version Visual Studio 2012 Update 4 or newer) is installed on the machine where the self-hosted integration runtime is running. If you don't have this update installed, download it now.
    - **To use a kubernetes supported self-hosted integration runtime:**
      - Follow the article to create and configure a kubernetes supported integration runtime.

Register

This section describes how to register Cassandra in Microsoft Purview using the Microsoft Purview governance portal.

Steps to register

To register a new Cassandra server in Microsoft Purview Unified Catalog:

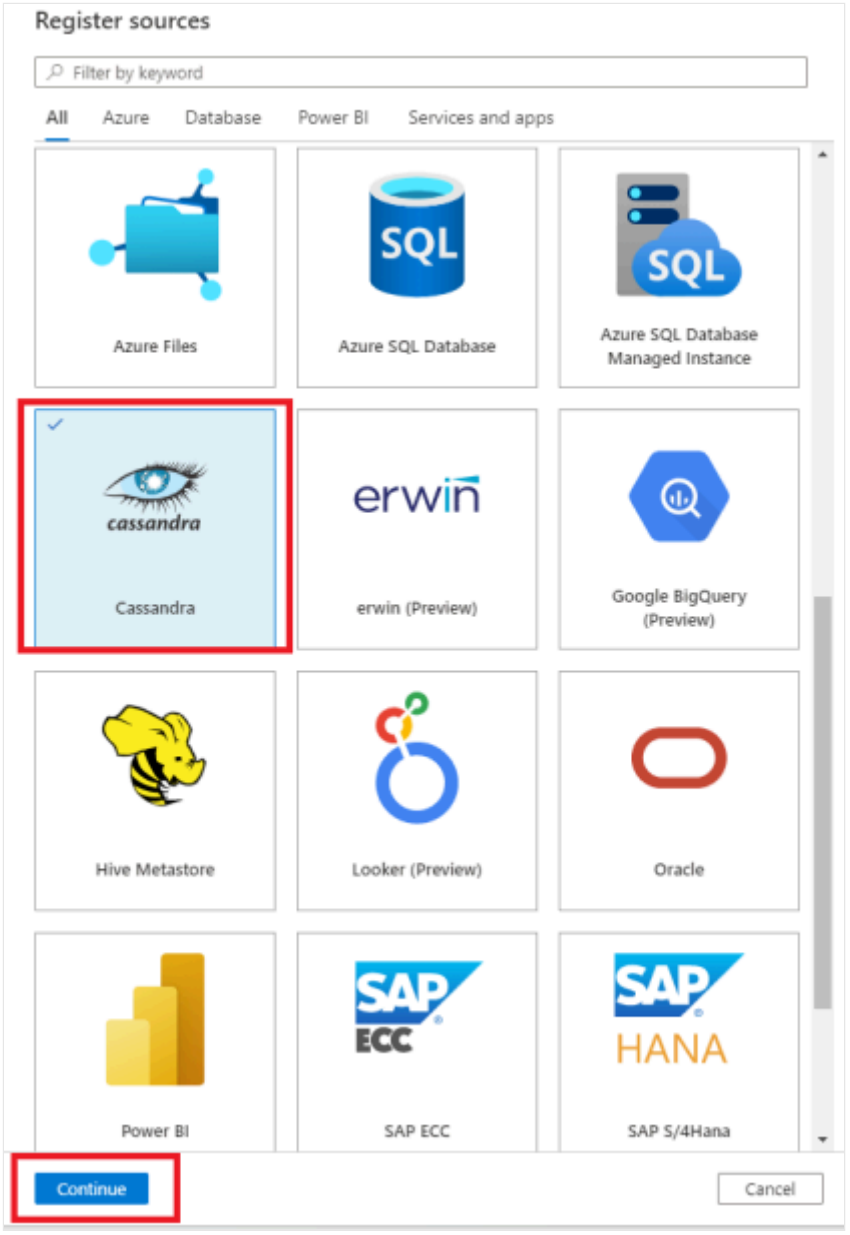
1. Open the Microsoft Purview governance portal by:

    - Browsing directly to https://web.purview.azure.com and selecting your Microsoft Purview account.
    - Opening the Azure portal, searching for and selecting the Microsoft Purview account. Selecting the **the Microsoft Purview governance portal** button.

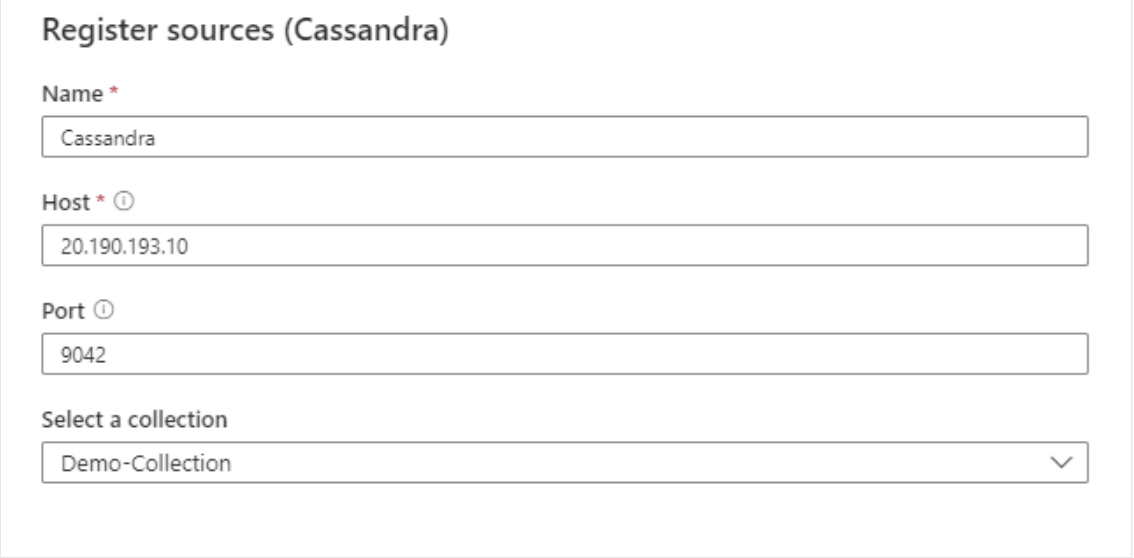
2. Select **Data Map** on the left pane.

3. Select **Register**.

4. On the **Register sources** screen, select **Cassandra**, and then select **Continue**:



5. On the **Register sources (Cassandra)** screen:
   a. Enter a **Name**. The data source will use this name in the catalog.
   b. In the **Host** box, enter the server address where the Cassandra server is running. For example, 20.190.193.10.
   c. In the **Port** box, enter the port used by the Cassandra server.
   d. Select a collection from the list.



   e. Select **Register**.

Scan

Follow the steps below to scan Cassandra to automatically identify assets. For more information about scanning in general, see our introduction to scans and ingestion

Create and run scan

To create and run a new scan:

1. If your server is publicly accessible, skip to step two. Otherwise, you'll need to make sure your self-hosted integration runtime is configured:
   a. In the Microsoft Purview governance portal, got to the Management Center, and select **Integration runtimes**.
   b. Make sure a self-hosted integration runtime is available. If one isn't set up, use the steps mentioned in prerequisites to set up a self-hosted integration runtime.

2. In the Microsoft Purview governance portal, navigate to **Sources**.

3. Select the registered Cassandra server.

4. Select **New scan**.

5. Provide the following details.

a. **Name**: Specify a name for the scan.

b. **Connect via integration runtime**: Select the Azure auto-resolved integration runtime if your server is publicly accessible, or your configured self-hosted integration runtime if it isn't publicly available.

c. **Credential**: When you configure the Cassandra credentials, be sure to:

   - Select **Basic Authentication** as the authentication method.
   - In the **User name** box, provide the name of the user you're making the connection for.
   - In the key vault's secret, save the password of the Cassandra user you're making the connection for.

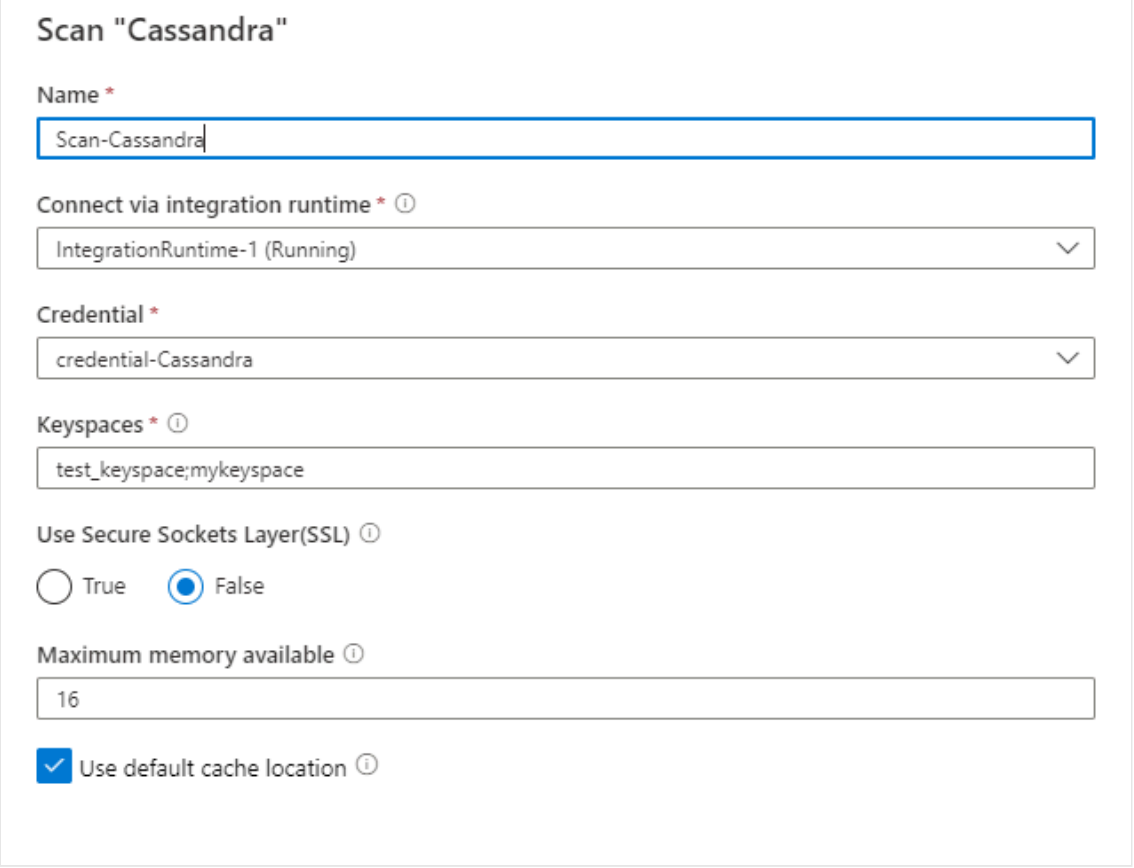
   For more information, see Credentials for source authentication in Microsoft Purview.

d. **Keyspaces**: Specify a list of Cassandra keyspaces to import. Multiple keyspaces must be separated with semicolons. For example, `keyspace1;keyspace2`. When the list is empty, all available keyspaces are imported.

e. **Use Secure Sockets Layer(SSL)**: Select **True** or **False** to specify whether to use Secure Sockets Layer (SSL) when connecting to the Cassandra server. By default, this option is set to **False**.

f. **Maximum memory available** (applicable when using self-hosted integration runtime): Specify the maximum memory (in GB) available on your VM to be used for scanning processes. This value depends on the size of Cassandra server

to be scanned.



6. Select **Test connection** to validate the settings.

7. Select **Continue**.

8. Select a **scan trigger**. You can set up a schedule or run the scan once.

9. Review your scan, and then select **Save and Run**.

View your scans and scan runs

To view existing scans:

1. Go to the Microsoft Purview portal. On the left pane, select **Data map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that has results you want to view. The pane shows you all the previous scan runs, along with the status and metrics for each scan run.
4. Select the run ID to check the scan run details.

Manage your scans

To edit, cancel, or delete a scan:

1. Go to the Microsoft Purview portal. On the left pane, select **Data Map**.

2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.

3. Select the scan that you want to manage. You can then:

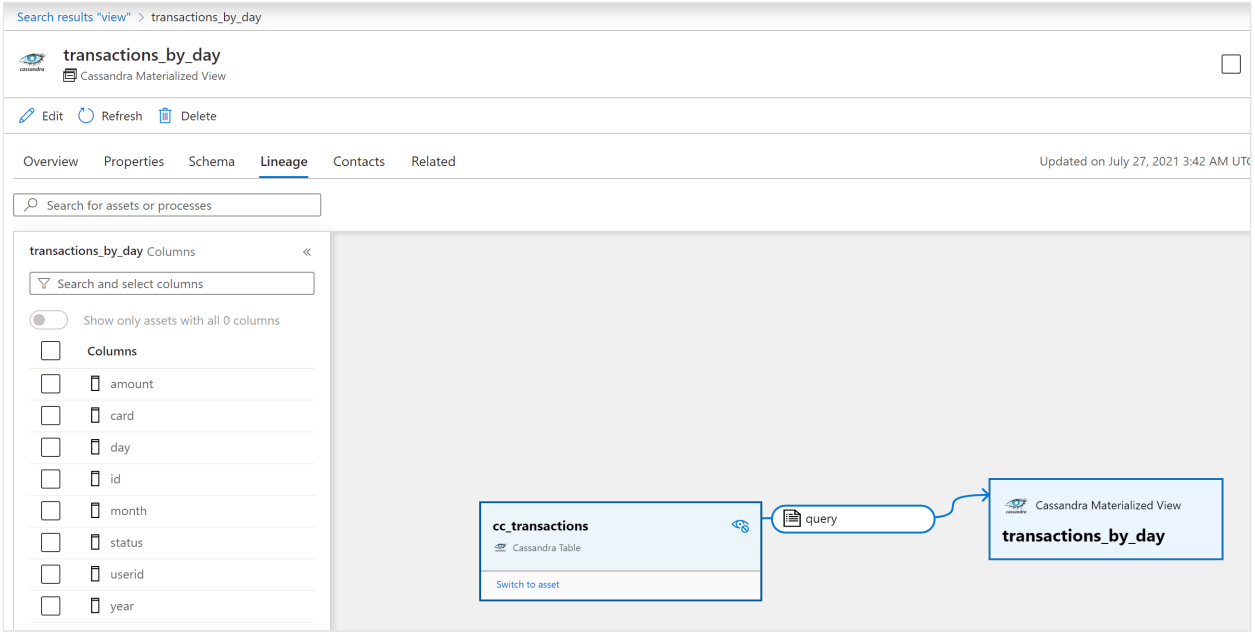
    - Edit the scan by selecting **Edit scan**.
    - Cancel an in-progress scan by selecting **Cancel scan run**.
    - Delete your scan by selecting **Delete scan**.

> **Note**
>
> - Deleting your scan does not delete catalog assets created from previous scans.

Lineage

After scanning your Cassandra source, you can browse Unified Catalog or search Unified Catalog to view the asset details.

Go to the asset -> lineage tab, you can see the asset relationship when applicable. Refer to the supported capabilities section on the supported Cassandra lineage scenarios. For more information about lineage in general, see data lineage and lineage user guide.



Next steps

Now that you've registered your source, follow the below guides to learn more about Microsoft Purview and your data.

- Data Estate Insights in Microsoft Purview
- Lineage in Microsoft Purview
- Search Unified Catalog

Feedback

Was this page helpful? Yes No

Connect to and manage ChatGPT Enterprise AI interactions in Microsoft Purview (preview)

This article outlines how to register ChatGPT Enterprise workspaces as a data source in Microsoft Purview. This connector allows you to discover and govern interactions with ChatGPT Enterprise AI across your organization. Before you start, ensure you have the required permissions, billing configuration, and a ChatGPT Enterprise plan as described in the Prerequisites section.

> **Important**
>
> Any scans run are subject to the new Microsoft Purview billing model. Be sure to enable pay-as-you-go billing in your organization.

Find more information about how Microsoft Purview solutions support AI interactions with ChatGPT Enterprise.

Supported capabilities

This section summarizes which scan modes and governance features are supported for ChatGPT Enterprise sources.

Choose to scan all user interactions from the current date or pick a specific start date from past when configuring your scan.

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan	Classification	Labeling	Access Policy	Lineage	Data Sharing	Live view
Yes	Yes	Yes	No	Yes	No	No	No	No	No

When scanning ChatGPT Enterprise sources, Microsoft Purview supports:

- Extracting technical metadata, including:
  - Text prompts
  - Text responses

> **Note**
>
> Due to OpenAI's current API limitations, conversations are ingested 24 hours after they occur.

Prerequisites

Make sure you have the following accounts, permissions, and configurations in place:

- An Azure account with an active subscription.
- An active Enterprise Microsoft Purview account.
  - If you create a classic Microsoft Purview instance for the first time in your organization, the instance is automatically upgraded to the enterprise version of Microsoft Purview.
- *Data Source Administrator* and *Data Reader* Microsoft Purview Data Governance classic permissions must be assigned to register a source and manage it in the Microsoft Purview governance portal.
- A ChatGPT Enterprise plan.
- Have pay-as-you-go billing enabled in your organization.
- Have a collection policy allowing for the ingestion of prompts and responses. There are two ways you can set up the policy:
  - Set up a one-click policy to capture interactions for enterprise AI apps.
  - Set up a collection policy for enterprise AI apps in Microsoft Purview Data Loss Prevention.
- At present, scope for API key is not defined while creating the API key on the OpenAI API Platform Portal, but it is a prerequisite for Microsoft Purview to run a scan and fetch metadata. Complete the following steps to define the scope:
  - Send an email to `support@openai.com` with:
    1. The last 4 digits of the API key.
    2. The Key Name.
    3. The Created By Name.
    4. The requested scope (**read**, **delete**, or both). Microsoft Purview needs at least the **read** scope to run a scan.

- The Open AI team will verify the key and grant the requested Compliance API scopes.
- Administrators can now use this key or pass it to a partner for use with the Compliance API.

Add permissions

Add the necessary permissions for other Microsoft Purview solutions used as part of this scanning process:

- Azure Key Vault
- Microsoft Purview AI hub
- Microsoft Purview Communication Compliance
- Microsoft Purview Data Lifecycle Management
- Microsoft Purview Data Map
- Microsoft Purview eDiscovery
- Microsoft Purview Insider Risk Management
- Microsoft Purview Records Management

Microsoft Purview uses the OAuth 2.0 protocol for accessing ChatGPT Enterprise workspaces. To set up the credential, create an API key credential mapped to your Azure Key Vault when you configure the scan. For detailed steps, see the Scan section.

Register a ChatGPT Enterprise workspace

This section describes how to register a ChatGPT Enterprise workspace in Microsoft Purview using the Microsoft Purview governance portal.

Complete the following steps to register:

1. Open the Azure portal and complete the following steps:
   a. Search for *Key Vault* to create/manage the secret to use for this connector.
   b. Create a Key Vault.
   c. Grant role assignments to your Key Vault. Assign Key Vault administrators and Key Vault secret users for the required members. For other roles, check the Key Vault roles in the **Required permissions** section.
   d. Create a secret for your Key Vault.
   e. For your Key Vault, assign access policy.

2. Open the Microsoft Purview governance portal and complete the following steps:

a. Search for Microsoft Purview accounts, select the account you want to use and select **Open Microsoft Purview governance portal**.
b. Select **Data Map** in left navigation and select the *Data source*.
c. Select **Register data source**.
d. On register sources, select **ChatGPT Enterprise** and select **Continue**.

3. Open the ChatGPT Enterprise register source and complete the following steps:
   a. Enter a name for the data source within the catalog.
   b. Enter the WorkspaceID for your Enterprise workspace.
   c. Select a domain and collection.
   d. Select **Register**.

4. Open the Microsoft Purview portal and complete the following steps:
   a. Select **Source management** under **Data Map**, and select **Credentials**.
   b. Select **New**.
   c. Create and manage credentials for scans in Microsoft Purview Data Map
   d. Enter the name, description, and select the domain.
   e. For authentication method, select **API key**.
   f. For Key Vault connection, select the Azure key vault created for *ChatGPT Enterprise*.
   g. Enter the secret name and secret version.

Scan a ChatGPT Enterprise workspace

Complete the following steps to scan a ChatGPT Enterprise workspace to automatically identify messages. For more information about scanning in general, see Scans and ingestion in Microsoft Purview.

1. Navigate to **Sources**.
2. Select the registered ChatGPT Enterprise workspace.
3. Select **New scan** and complete the following:
   a. **Name**: The name of the scan.
   b. **Capture data since**: Select if the scan should capture data from the current date of the new scan or choose a specific date from the past.
   c. **Credential**: Select the credential mapping to the correct Key Vault for this connector.
   d. **Domain**: Select a domain from the existing list or choose default domain.
   e. **Collection**: Select a collection from your domain.
4. Select **Test connection** to ensure the connection is established successfully. Select **Continue**.
5. Review your scan and select **Save** and **Run**.

For future scans, we recommend running an incremental scan. Avoid creating a new scan to avoid duplicate ingestion of data.

View scan status and results

To view existing scans, complete the following steps:

1. Open the Microsoft Purview portal and select **Data Map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that has results you want to view. The pane shows you all the previous scans, along with the status and metrics for each scan.
4. Select the scan ID to check the scan details. During scanning, the organization is considered an asset, but isn't classified. The number of assets classified in the scan details are always one less than the number of assets ingested.
5. Review added data to Microsoft Purview in Data Security Posture Management (DSPM) for AI.
    a. View analytics for AI interactions, sensitive data, and insider risk in DSPM for AI reports.
    b. View the AI interactions for prompts and responses in Activity explorer.
    c. View the sensitive info type data identified in prompts and responses in Activity explorer.
    \

Use other Microsoft Purview solutions with this data:

- Communication Compliance
- Data Lifecycle Management
- eDiscovery
- Insider Risk Management
- Records Management

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/24/2026

Connect to and manage Db2 in Microsoft Purview

Article • 05/21/2025

This article outlines how to register Db2, and how to authenticate and interact with Db2 in Microsoft Purview. For more information about Microsoft Purview, read the introductory article.

Supported capabilities

Scanning capabilities

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan
Yes	Yes	No	Yes

The supported IBM Db2 versions are Db2 for LUW 9.7 to 11.x. Db2 for z/OS (mainframe) and iSeries (AS/400) aren't supported.

When you scan an IBM Db2 source, Microsoft Purview supports:

- Extracting technical metadata including:
  - Server
  - Databases
  - Schemas
  - Tables including the columns, foreign keys, indexes, and constraints
  - Views including the columns
  - Triggers
- Fetching static lineage on assets relationships among tables and views.

You can scan an entire Db2 database, or scope the scan to a subset matching the given names or name patterns.

Other capabilities

For **classifications**, **sensitivity labels**, **policies**, **data lineage**, and **live view**, see the list of supported capabilities.

Known limitations

When object is deleted from the data source, currently the subsequent scan won't remove the corresponding asset in Microsoft Purview.

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- An active Microsoft Purview account.
- You need Data Source Administrator and Data Reader permissions to register a source and manage it in the Microsoft Purview governance portal. For more information about permissions, see Access control in Microsoft Purview.
- Set up the right integration runtime for your scenario:
  - **To use a self-hosted integration runtime:**
    - Follow the article to create and configure a self-hosted integration runtime.
    - Ensure JDK 11 is installed on the machine where the self-hosted integration runtime is installed. Restart the machine after you newly install the JDK for it to take effect.
    - Ensure that Visual C++ Redistributable (version Visual Studio 2012 Update 4 or newer) is installed on the machine where the self-hosted integration runtime is running. If you don't have this update installed, download it now.
    - Download the Db2 JDBC driver on the machine where your self-hosted integration runtime is running. Note down the folder path so you can use it to set up the scan.
  - **To use a kubernetes supported self-hosted integration runtime:**
    - Follow the article to create and configure a kubernetes supported integration runtime.
    - Download the Db2 JDBC driver on the machine where your self-hosted integration runtime is running. Note down the folder path so you can use it to set up the scan.

  > **Note**
  >
  > The driver should be accessible by the self-hosted integration runtime. By default, self-hosted integration runtime uses local service account "NT SERVICE\DIAHostService". Make sure it has "Read and execute" and "List folder contents" permission to the driver folder.

- The Db2 user must have the CONNECT permission. Microsoft Purview connects to the syscat tables in IBM Db2 environment when importing metadata.

Register

This section describes how to register Db2 in Microsoft Purview using the Microsoft Purview governance portal.

Steps to register

To register a new Db2 source in Microsoft Purview Unified Catalog, do the following steps:

1. Navigate to your Microsoft Purview account in the Microsoft Purview governance portal.
2. Select **Data Map** on the left navigation.
3. Select **Register**
4. On Register sources, select **Db2**. Select **Continue**.

On the **Register sources (Db2)** screen, do the following steps:

1. Enter a **Name** that the data source is within the Catalog.

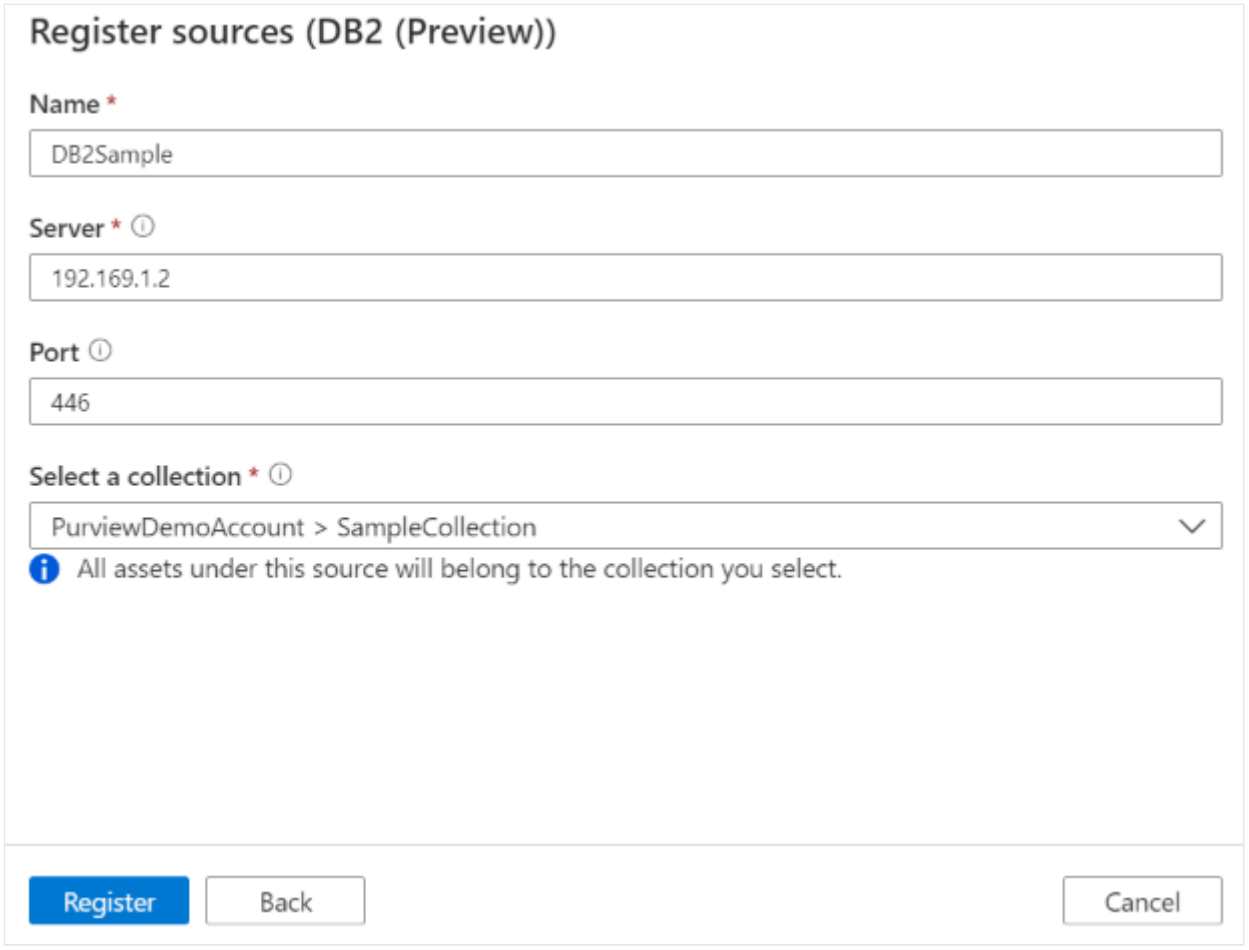
2. Enter the **Server** name to connect to a Db2 source. This name can either be:

    - A host name used to connect to the database server. For example: `MyDatabaseServer.com`
    - An IP address. For example: `192.169.1.2`

3. Enter the **Port** used to connect to the database server (446 by default for Db2).

4. Select a collection from the list.

5. Finish to register the data source.



Scan

Follow the steps to scan Db2 to automatically identify assets. For more information about scanning in general, see our introduction to scans and ingestion.

Authentication for a scan

The supported authentication type for a Db2 source is **Basic authentication**.

Create and run scan

To create and run a new scan, follow these steps:

1. In the Management Center, select Integration runtimes. Make sure a self-hosted integration runtime is set up. If it isn't set up, use the steps mentioned in prerequisites to create a self-hosted integration runtime.

2. Navigate to **Sources**.

3. Select the registered Db2 source.

4. Select + **New scan**.

5. Provide the below details:

   a. **Name**: The name of the scan

   b. **Connect via integration runtime**: Select the configured self-hosted integration runtime

   c. **Credential**: Select the credential to connect to your data source. Make sure to:

      - Select **Basic Authentication** while creating a credential.
      - Provide the user name used to connect to the database server in the User name input field.
      - Store the user password used to connect to the database server in the secret key.

   d. **Database**: The name of the database instance to import.

   e. **Schema**: List subset of schemas to import expressed as a semicolon separated list. For example, `schema1;schema2`. All user schemas are imported if that list is empty. All system schemas (for example, SysAdmin) and objects are ignored by default.

      Acceptable schema name patterns can be static names or contain wildcard %. For example: `A%;%B;%C%;D`

      - Start with A or
      - End with B or
      - Contain C or
      - Equal D

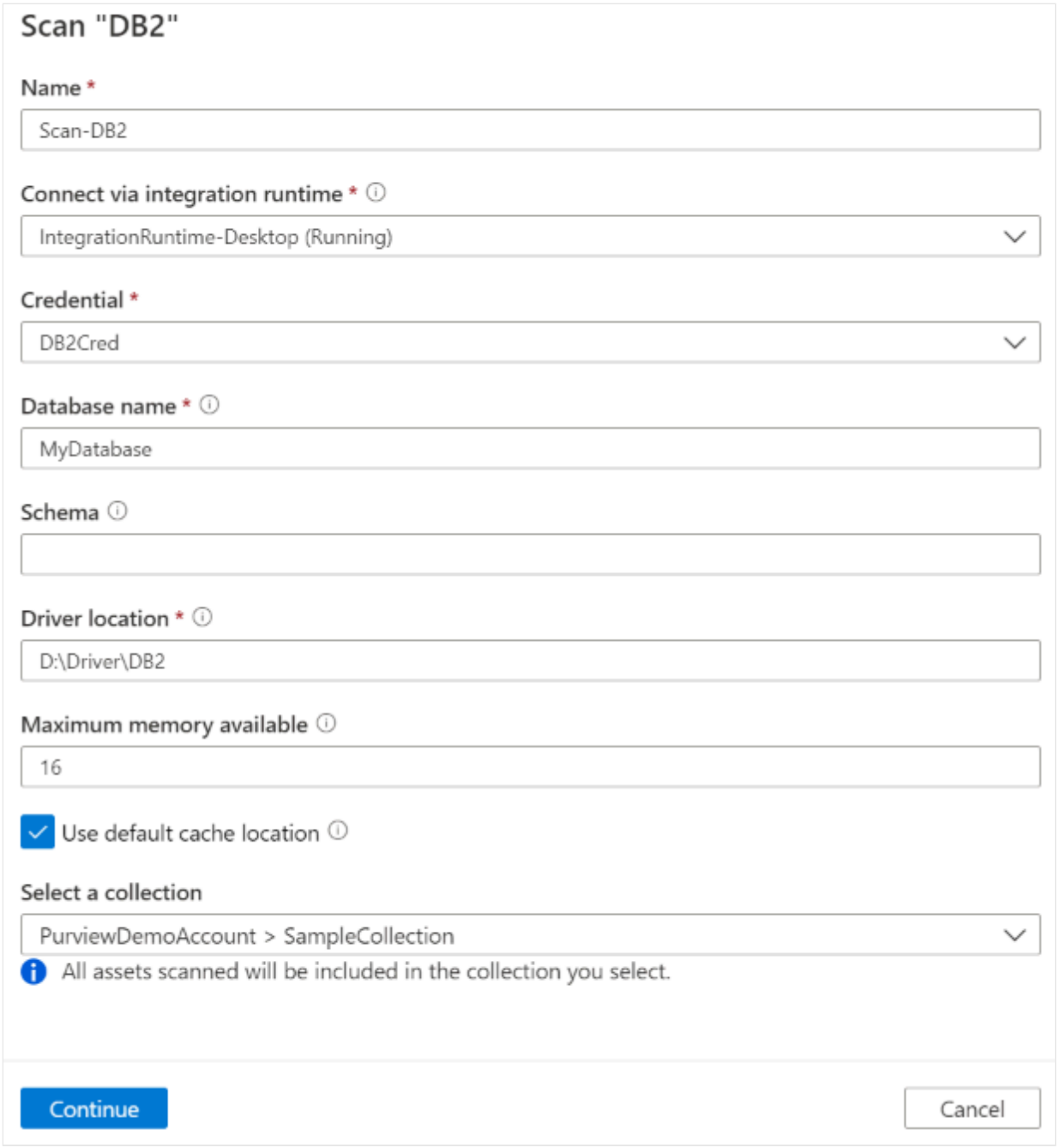
      Usage of NOT and special characters aren't acceptable.

   f. **Driver location**: Specify the path to the JDBC driver location in your machine where self-host integration runtime is running. For example, `D:\Drivers\Db2`.

      i. For Kubernetes-supported self-hosted integration runtime: `./drivers/Db2`. It's the path to valid JAR folder location. The value must be a valid relative file path. Please refer to the documentation to set up a scan with external drivers for uploading drivers in advance.

   g. **Maximum memory available**: Maximum memory (in GB) available on customer's VM to be used by scanning processes. The memory needed is dependent on the size of Db2 source to be scanned.

      > ⓘ **Note**
      >
      > As a thumb rule, please provide 1GB memory for every 1000 tables



6. Select **Continue**.

7. Choose your **scan trigger**. You can set up a schedule or ran the scan once.

8. Review your scan and select **Save and Run**.

View your scans and scan runs

To view existing scans:

1. Go to the Microsoft Purview portal. On the left pane, select **Data map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that has results you want to view. The pane shows you all the previous scan runs, along with the status and metrics for each scan run.

4. Select the run ID to check the scan run details.

Manage your scans

To edit, cancel, or delete a scan:

1. Go to the Microsoft Purview portal. On the left pane, select **Data Map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that you want to manage. You can then:
    - Edit the scan by selecting **Edit scan**.
    - Cancel an in-progress scan by selecting **Cancel scan run**.
    - Delete your scan by selecting **Delete scan**.

> ⓘ Note
>
> - Deleting your scan does not delete catalog assets created from previous scans.

Lineage

After scanning your Db2 source, you can browse Unified Catalog or search Unified Catalog to view the asset details.

Go to the asset -> lineage tab, you can see the asset relationship when applicable. Refer to the supported capabilities section on the supported Db2 lineage scenarios. For more information about lineage in general, see data lineage and lineage user guide.

Next steps

Now that you registered your source, follow the below guides to learn more about Microsoft Purview and your data.

- Data Estate Insights in Microsoft Purview
- Lineage in Microsoft Purview
- Search Unified Catalog

Connect to and manage Microsoft Dataverse in Microsoft Purview

This article outlines the process to register and govern a Microsoft Dataverse environment data source in Microsoft Purview including instructions to authenticate and interact with the Microsoft Dataverse source.

Supported capabilities

Scanning capabilities

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan
Yes	Yes	No	Yes

When scanning Dataverse source, Microsoft Purview supports extracting technical metadata including:

- Environment
- Tables, including columns

When setting up a scan, you can choose to scope the scan by selecting tables as needed.

Other capabilities

For **classifications**, **sensitivity labels**, **policies**, **data lineage**, and **live view**, see the list of supported capabilities.

Known limitations

- When object is deleted from the data source, currently the subsequent scan won't automatically remove the corresponding asset in Microsoft Purview.
- Dataverse integration with Microsoft Purview Data Map isn't available in sovereign clouds.

Prerequisites

- An Azure account with an active subscription. Create an account for free.

- An active Microsoft Purview account.
- You need to be a Data Source Administrator and Data Reader to register a source and manage it in the classic Microsoft Purview governance portal. See our Microsoft Purview Permissions page for details.
- You need to have the permission in your Dataverse environment to create an application user and assign the Service Reader role for the user.

Register

This section describes how to register Dataverse in Microsoft Purview. Before you create a scan, it's necessary to register the data source in Microsoft Purview.

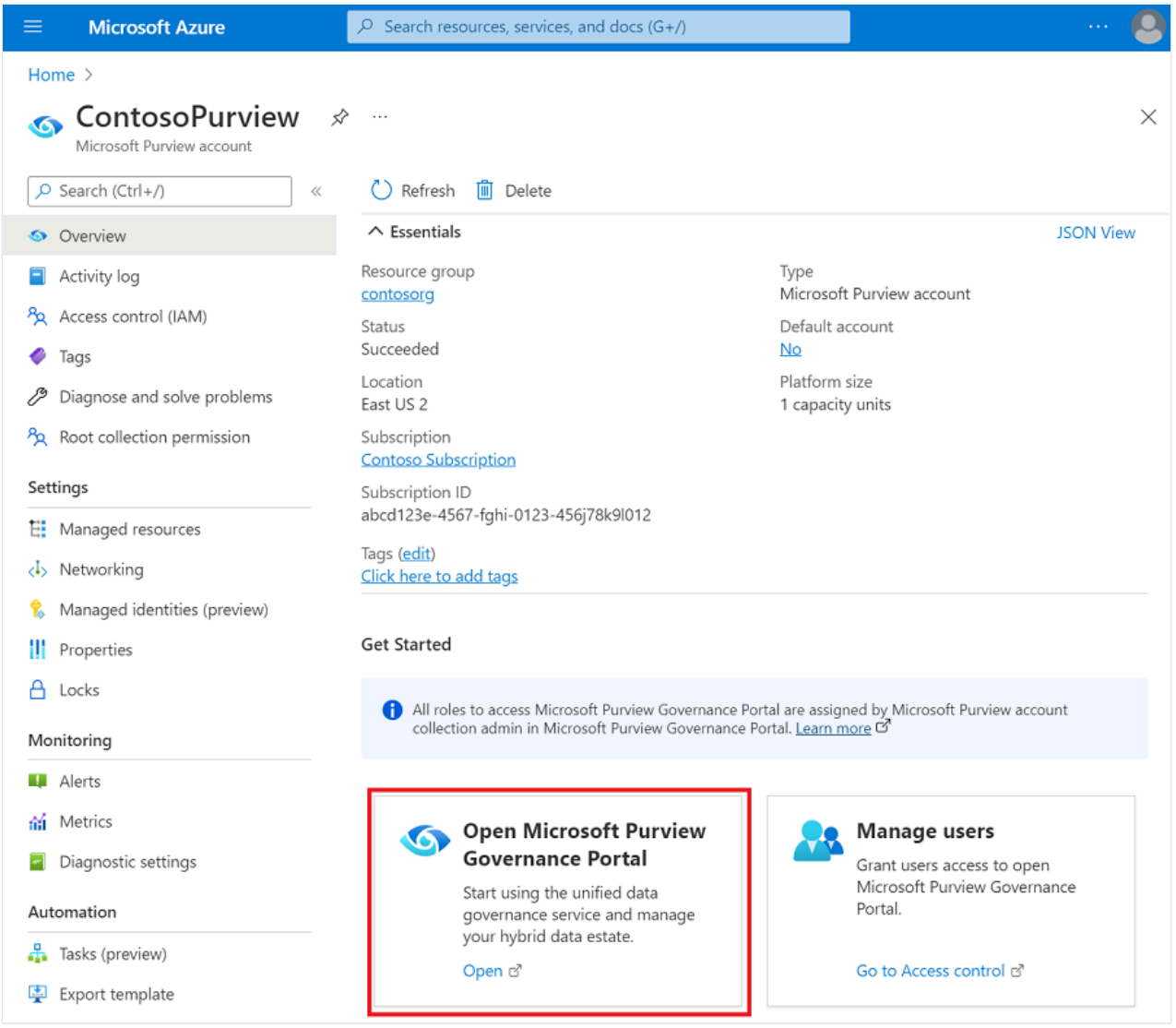
Prerequisites for register

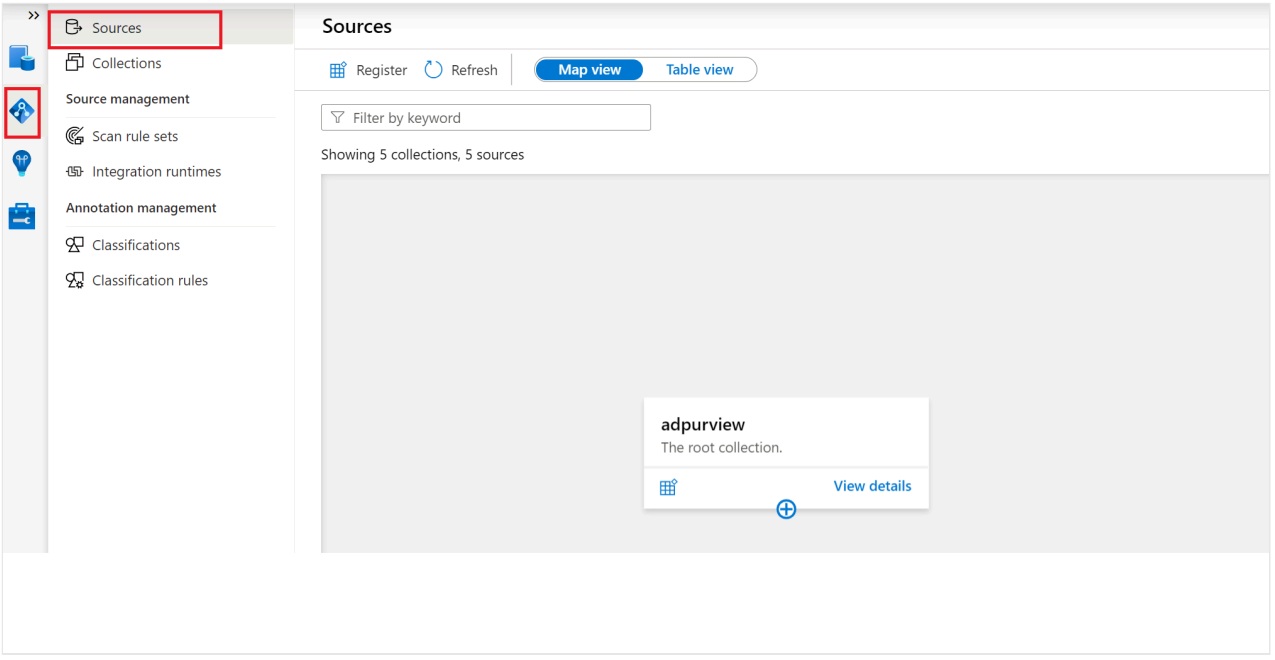
- You need to be a Data Source Admin and one of the other Purview roles (for example, Data Reader or Data Share Contributor) to register a source and manage it in the classic Microsoft Purview governance portal. See our Microsoft Purview Permissions page for details.

Steps to register

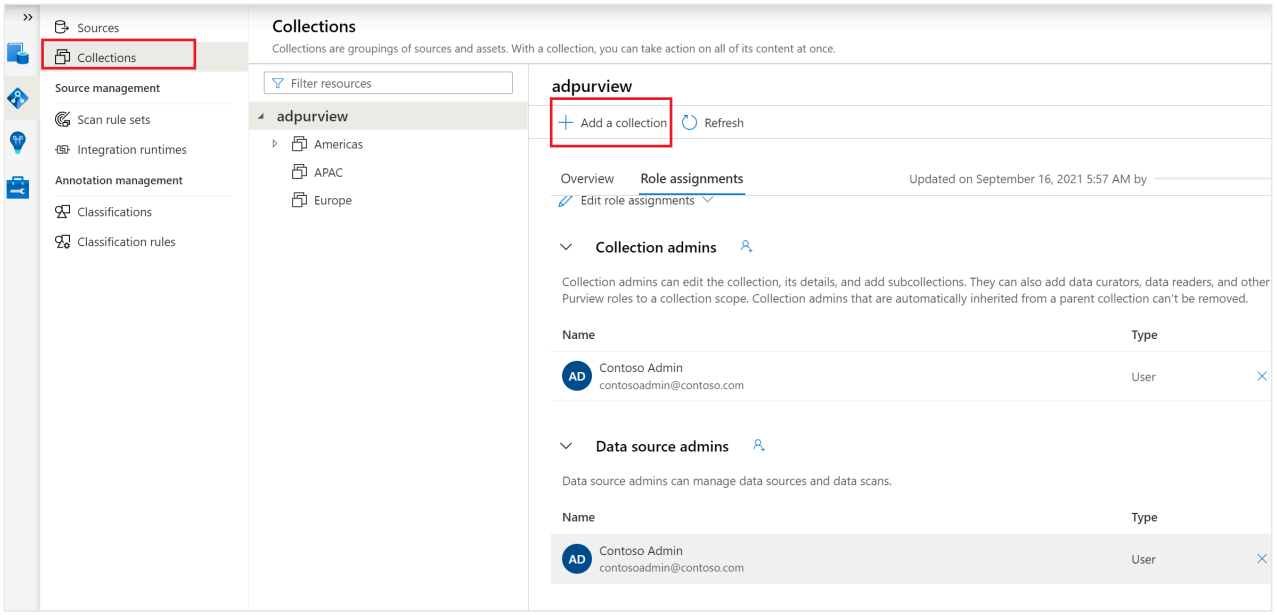
It's important to register the data source in Microsoft Purview prior to setting up a scan for the data source.

1. Go to the classic Microsoft Purview governance portal by:
    - Browsing directly to https://web.purview.azure.com and selecting your Microsoft Purview account.
    - Opening the Azure portal, searching for and selecting the Microsoft Purview account. Select the classic **Microsoft Purview governance portal**.
2. Navigate to the **Data Map --> Sources**

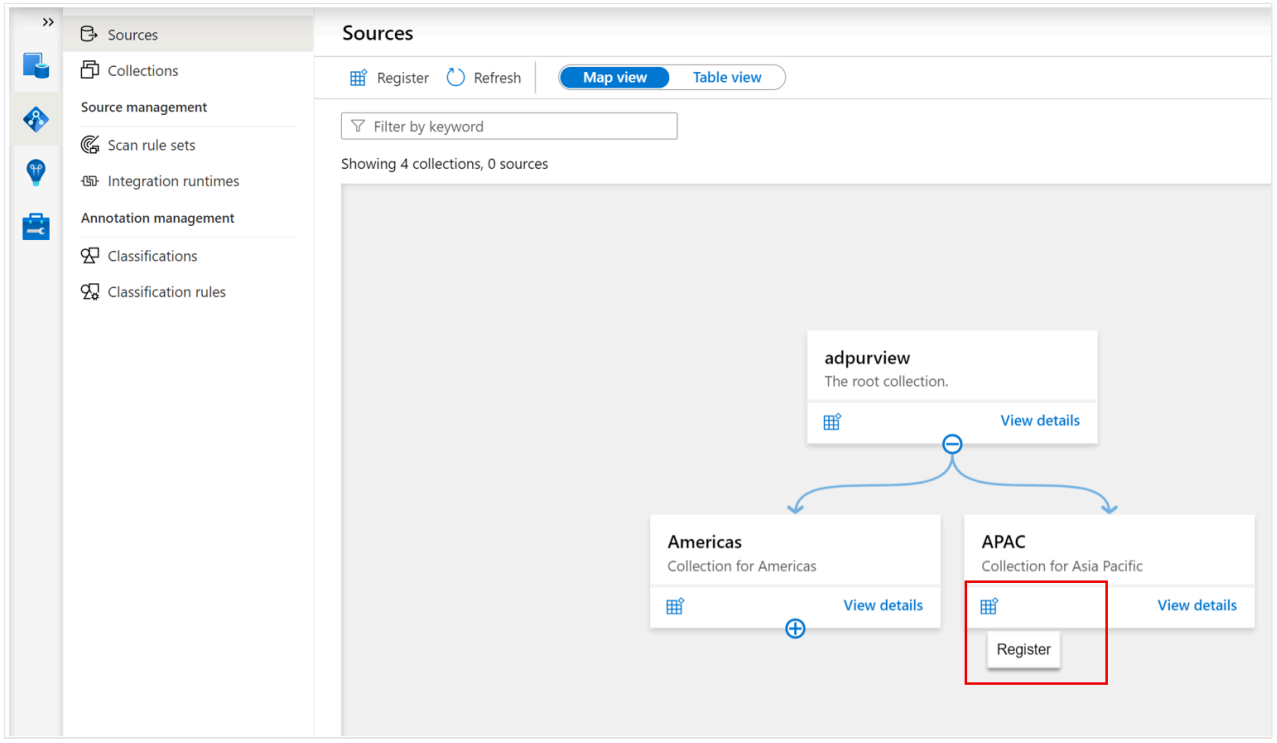




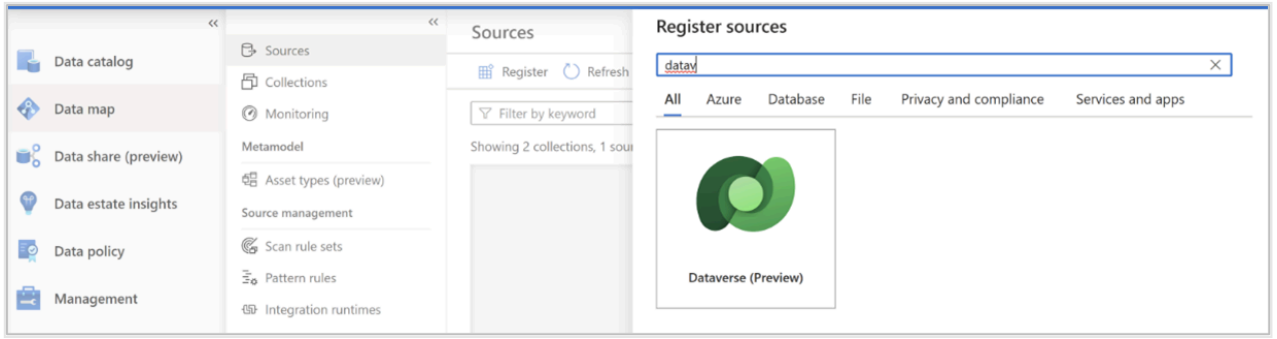
3. Create the Collection hierarchy using the **Collections** menu and assign permissions to individual subcollections, as required



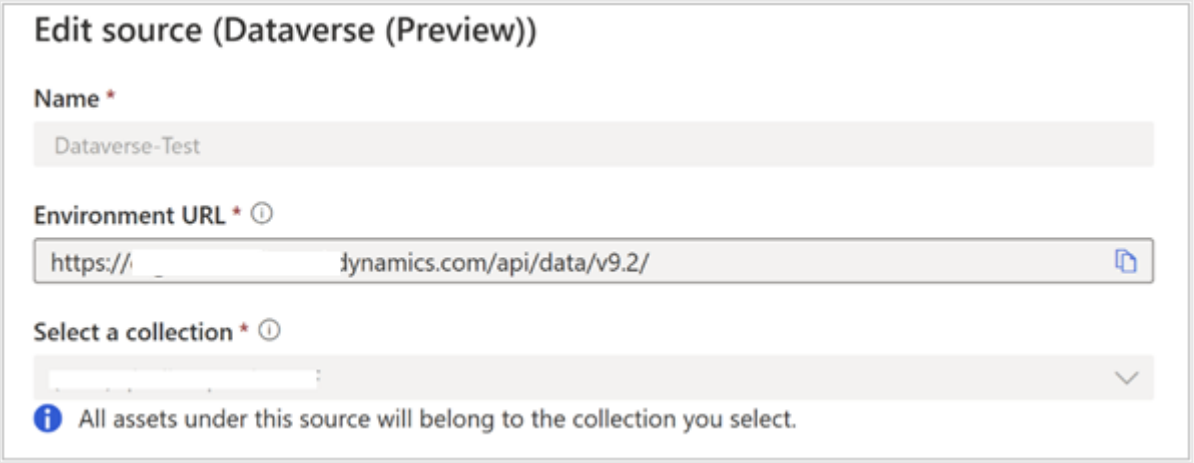
4. Navigate to the appropriate collection under the **Sources** menu and select the **Register** icon to register a new Dataverse data source



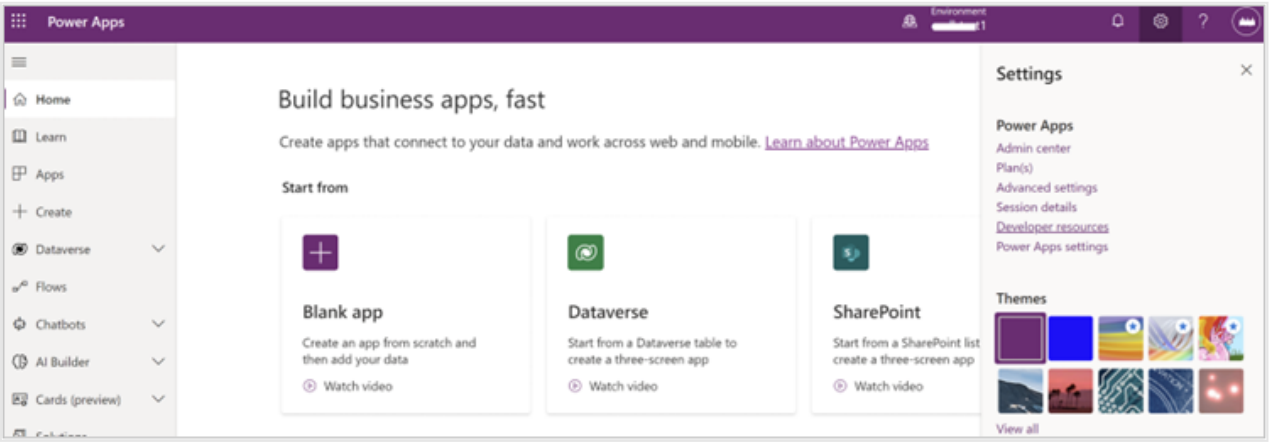
5. Select the **Dataverse** data source and select **Continue**

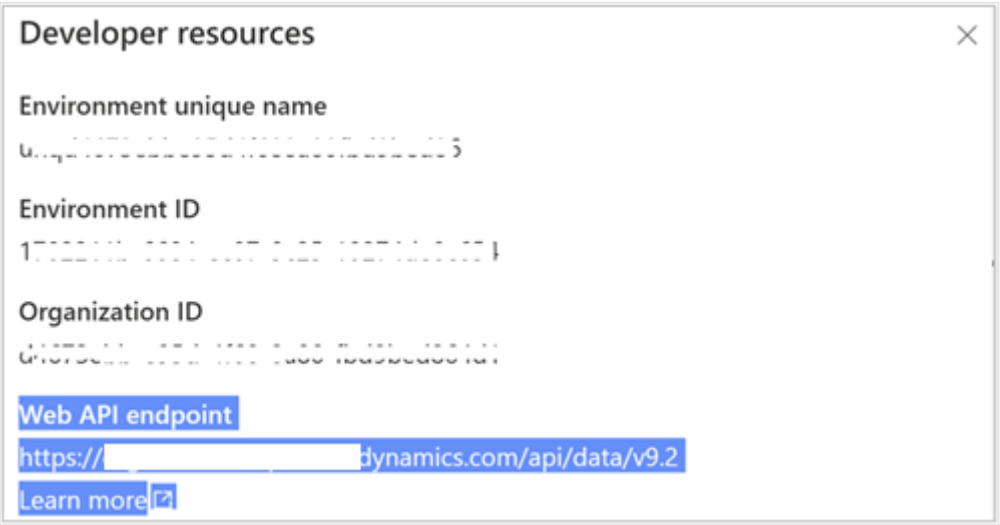


6. Provide a suitable **Name** for the data source. Fill in the Web API endpoint for your Dataverse environment as **Environment URL** and select a **collection** and select **Apply**.



You can get the Web API endpoint from Developer resources from the Microsoft Power Apps maker portal.





7. The Dataverse source is shown under the selected Collection

Scan

Authentication method

Currently Microsoft Purview supports the following options to scan Dataverse:

- **System assigned managed identity (SAMI)**. This is an identity that's associated directly with your Microsoft Purview account. The SAMI is created when your Microsoft Purview resource is created. It's managed by Azure and uses your Microsoft Purview account's name.
- **Service principal**. A service principal is an application that can be assigned permissions like any other group or user, without being associated directly with a person.

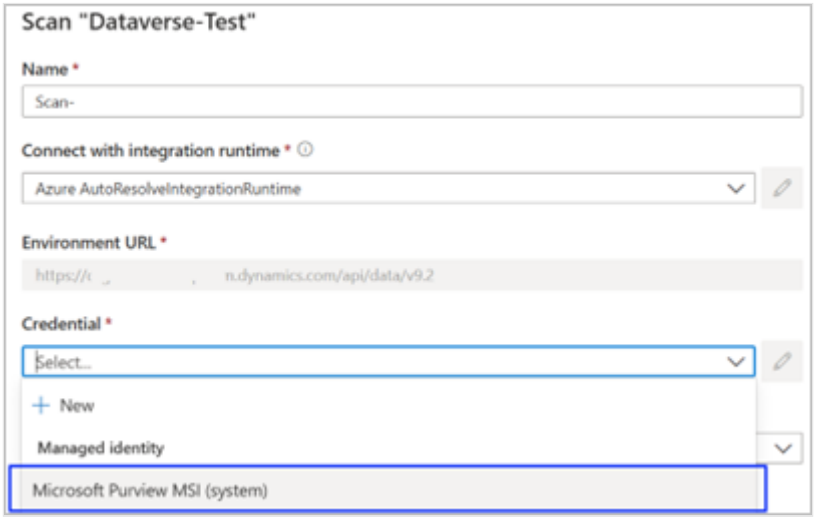
System assigned managed identity

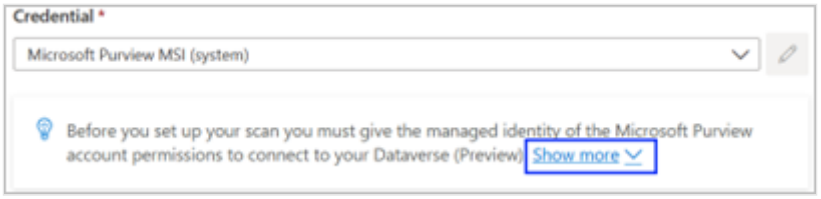
Using a system assigned managed identity for scanning

The system-assigned managed identity needs permission to get metadata from Dataverse environments. It must also be authorized to query the Dataverse tables to sample for classification. There are several steps to allow Microsoft Purview to use system-assigned managed identity to scan your Dataverse environment.

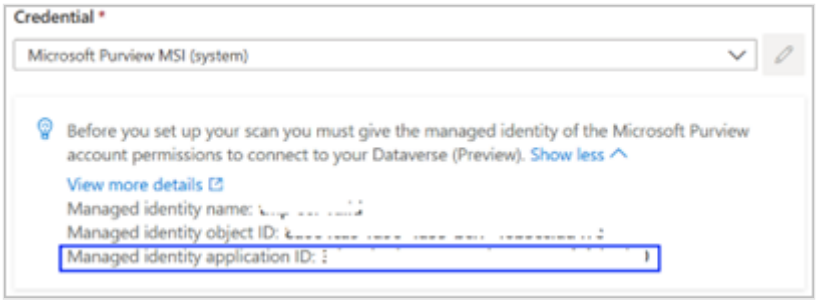
Get the application ID for the system-assigned managed identity

1. Go to **Data Map** > **Sources** to view the collection hierarchy.
2. Select the **New Scan** icon under the Dataverse source that you registered earlier.
3. For **Credential**, select **Microsoft Purview MSI (system)** from the drop-down list then select **Show more** to view details about the system assigned managed identity.



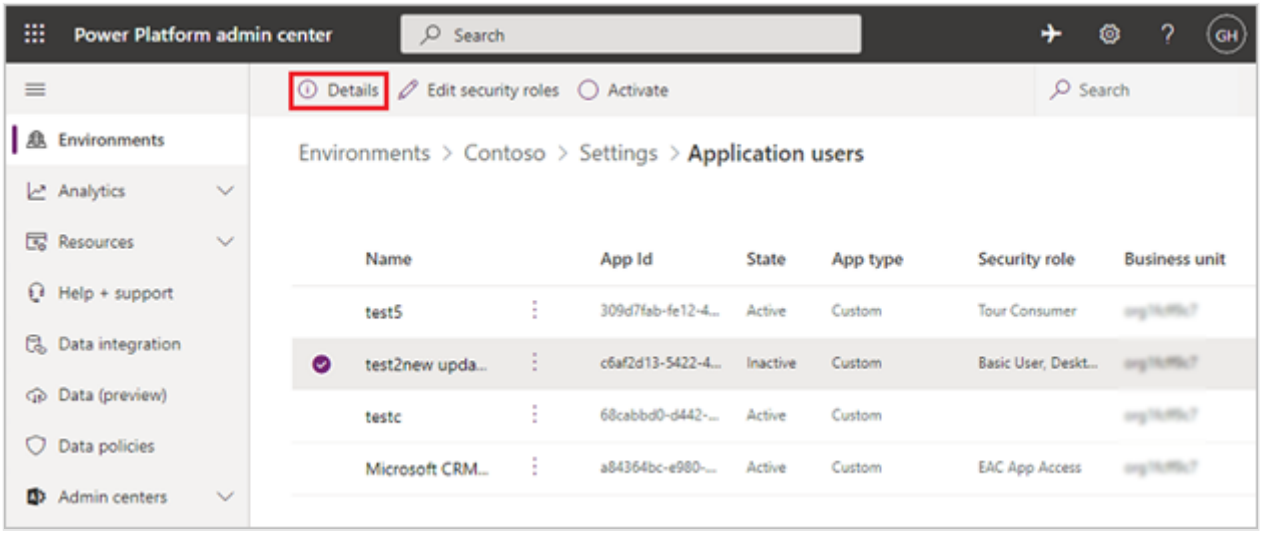


4. Record the managed identity application ID.



Create and enable an application user in Power Platform Admin Center

An application user is needed to be created in the Dataverse environment that you want to scan in Microsoft Purview. Follow the steps in Create An Application User to create the user with the application ID you recorded and assign the security role **Service Reader** to it.



You can check the application user in the Details page. Now the application user is bound with the system assigned managed identity created with your Microsoft Purview resource.

Create the scan

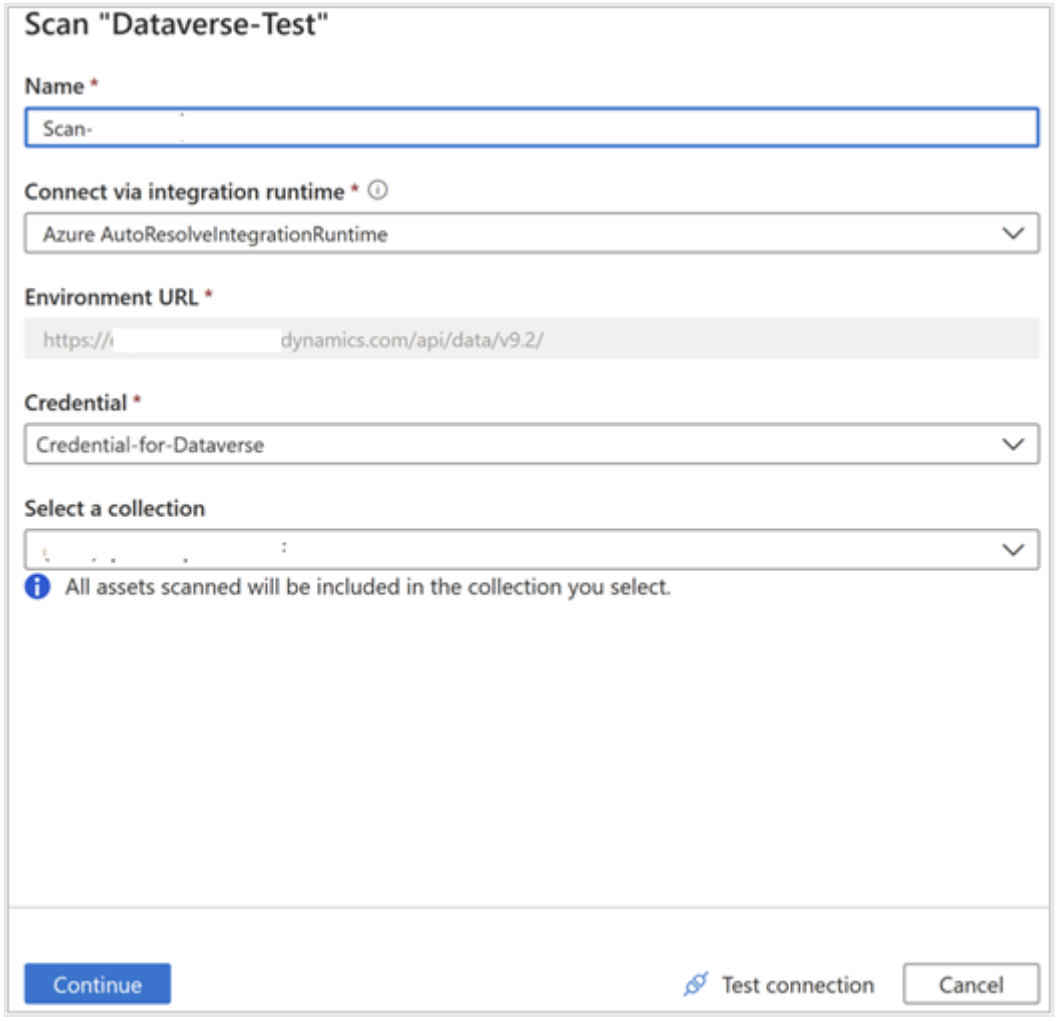
1. Go to **Data Map** > **Sources** to view the collection hierarchy.
2. Select the **New Scan** icon under the Dataverse source that you registered earlier.
3. For **Name**, provide a name for the scan.

4. For **Integration runtime**, see the list of supported data sources for scans to choose right integration runtime according to the scenario.
5. For **Credential**, select the credential you prepared earlier.

- If system-assigned managed identity is configured for authentication, then select **Microsoft Purview MSI (system)** from the drop-down list.
- If service principal is configured for authentication, then select the credential you created before from the drop-down list.

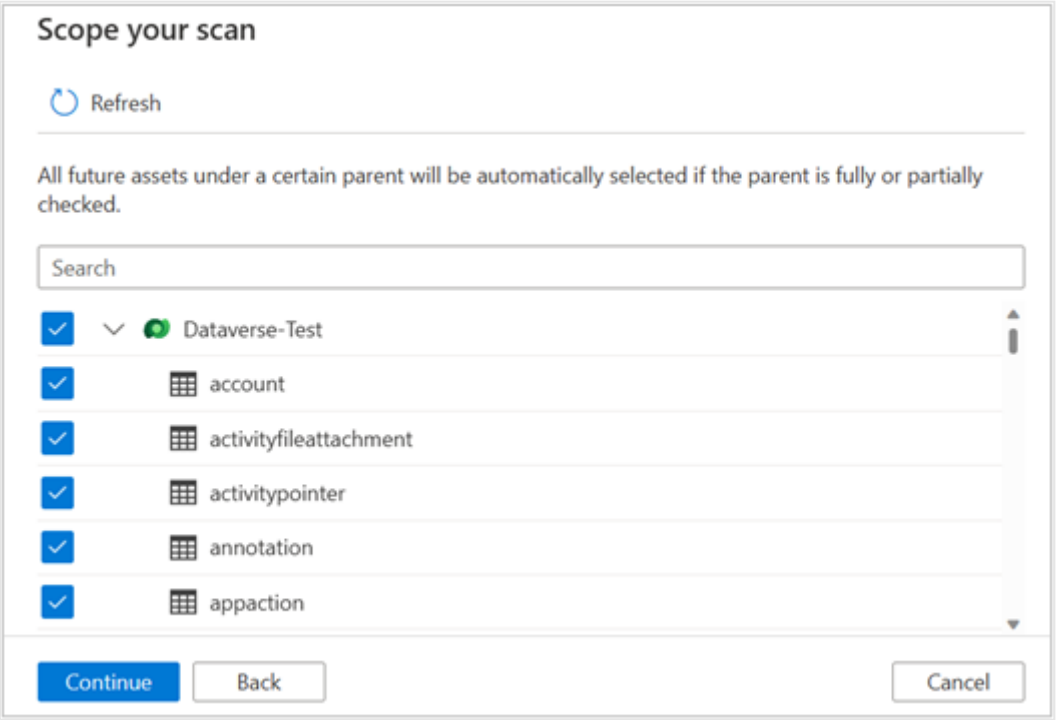
1. For **Select a collection**, choose the appropriate collection for the scan.

2. Select **Test connection** to validate the connection. After the connection is successful, select **Continue**.



Scope and run the scan

1. You can scope your scan to specific Dataverse tables by choosing the appropriate items in the list.



2. Choose your scan trigger. You can set up a schedule or run the scan once.

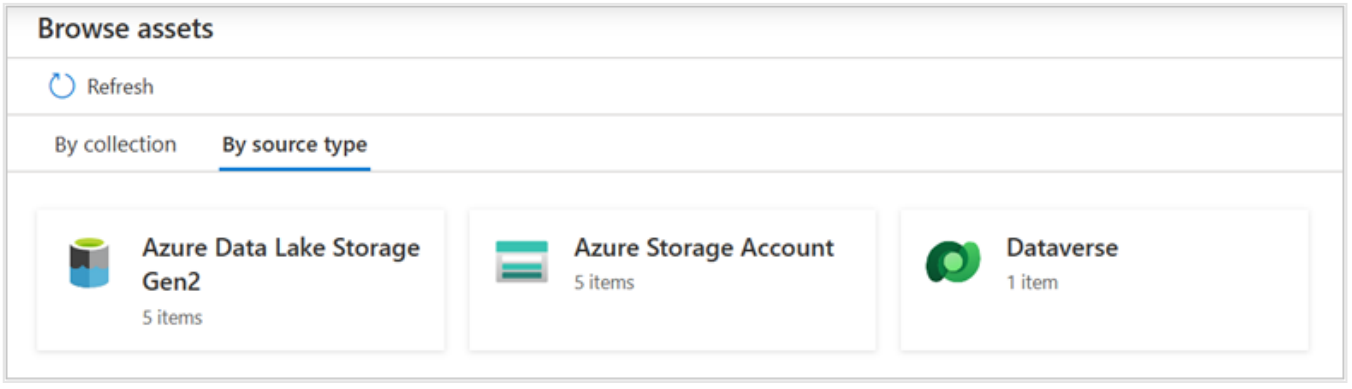
3. Review your scan, and then select **Save and run**.

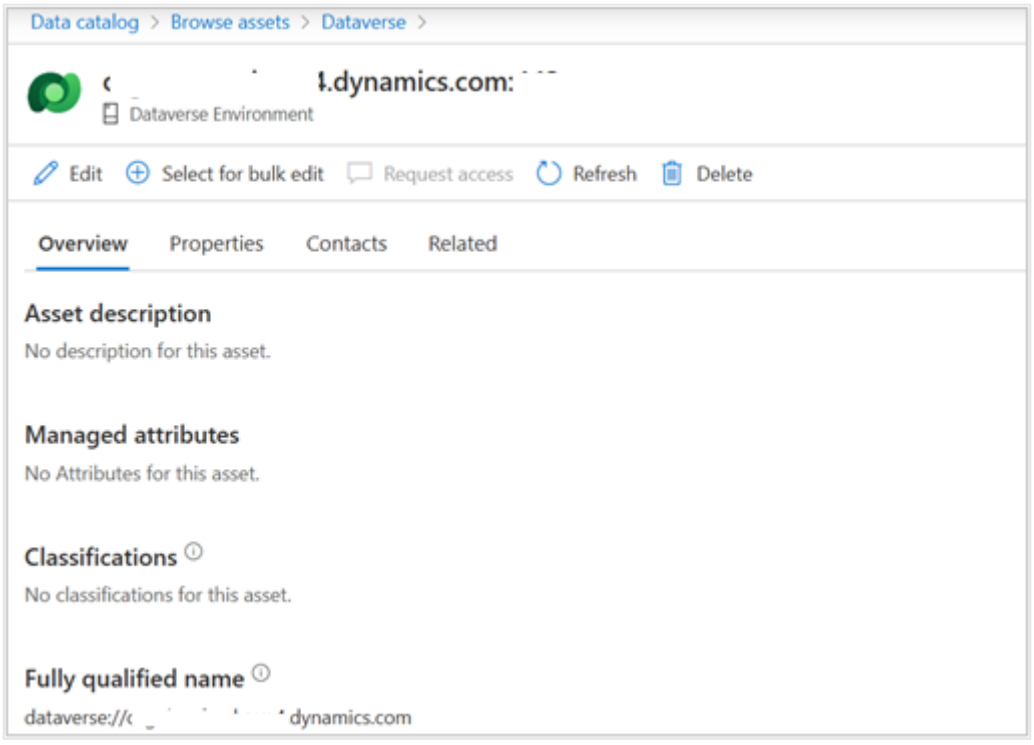
View and manage a scan

To check the status of a scan, go to the data source in the collection, and then select **View details**. The scan details indicate the progress of the scan in **Last run status**, along with the number of assets scanned and classified. **Last run status** is updated to **In progress** and then **Completed** after the entire scan runs successfully. After you run a scan, you can use the run history to manage it. In the run history, you have options for running the scan again, editing it, or deleting it.

Browse, search, and view Dataverse assets

Once the scan is completed, you can go to Microsoft Purview Unified Catalog to browse or search the Dataverse related assets, and view related details.





Last updated on 02/02/2026

Connect to and manage erwin Mart servers in Microsoft Purview

Article • 03/28/2025

This article outlines how to register erwin Mart servers, and how to authenticate and interact with erwin Mart Servers in Microsoft Purview. For more information about Microsoft Purview, read the introductory article.

Supported capabilities

Scanning capabilities

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan
Yes	Yes	No	Yes

The supported erwin Mart versions are 9.x to 2021.

When scanning erwin Mart source, Microsoft Purview supports:

- Extracting technical metadata including:
  - Mart
  - Libraries
  - Models
  - Entities including the attributes, foreign keys, indexes, index members, candidate keys, and triggers
  - Default values
  - Synonyms
  - Sequences
  - Domains
  - Subject areas
  - Relationships
  - Validation rules including the valid values
  - ER diagrams
  - Views including the attributes
  - Stored procedures including the parameters
  - Schemas

    - Subtype relationships
    - View relationship
    - User defined properties

- Fetching static lineage on assets relationships among entities, views and stored procedures.

When setting up scan, you can choose to scan an entire erwin Mart server, or scope the scan to a list of models matching the given name(s).

Other capabilities

View a table listing the supported capabilities for classifications, sensitivity labels, policies, data lineage, and live view.

Known limitations

When object is deleted from the data source, currently the subsequent scan won't automatically remove the corresponding asset in Microsoft Purview.

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- An active Microsoft Purview account.
- You need Data Source Administrator and Data Reader permissions to register a source and manage it in the Microsoft Purview governance portal. For more information about permissions, see Access control in Microsoft Purview.
- Set up the latest self-hosted integration runtime. For more information, see the create and configure a self-hosted integration runtime guide.

  > **Important**
  >
  > Make sure to install the self-hosted integration runtime and the Erwin Data Modeler software on the same machine where erwin Mart instance is running.

    - Ensure JDK 11 is installed on the machine where the self-hosted integration runtime is installed. Restart the machine after you newly install the JDK for it to take effect.

- Ensure Visual C++ Redistributable (version Visual Studio 2012 Update 4 or newer) is installed on the self-hosted integration runtime machine. If you don't have this update installed, you can download it here.

Register

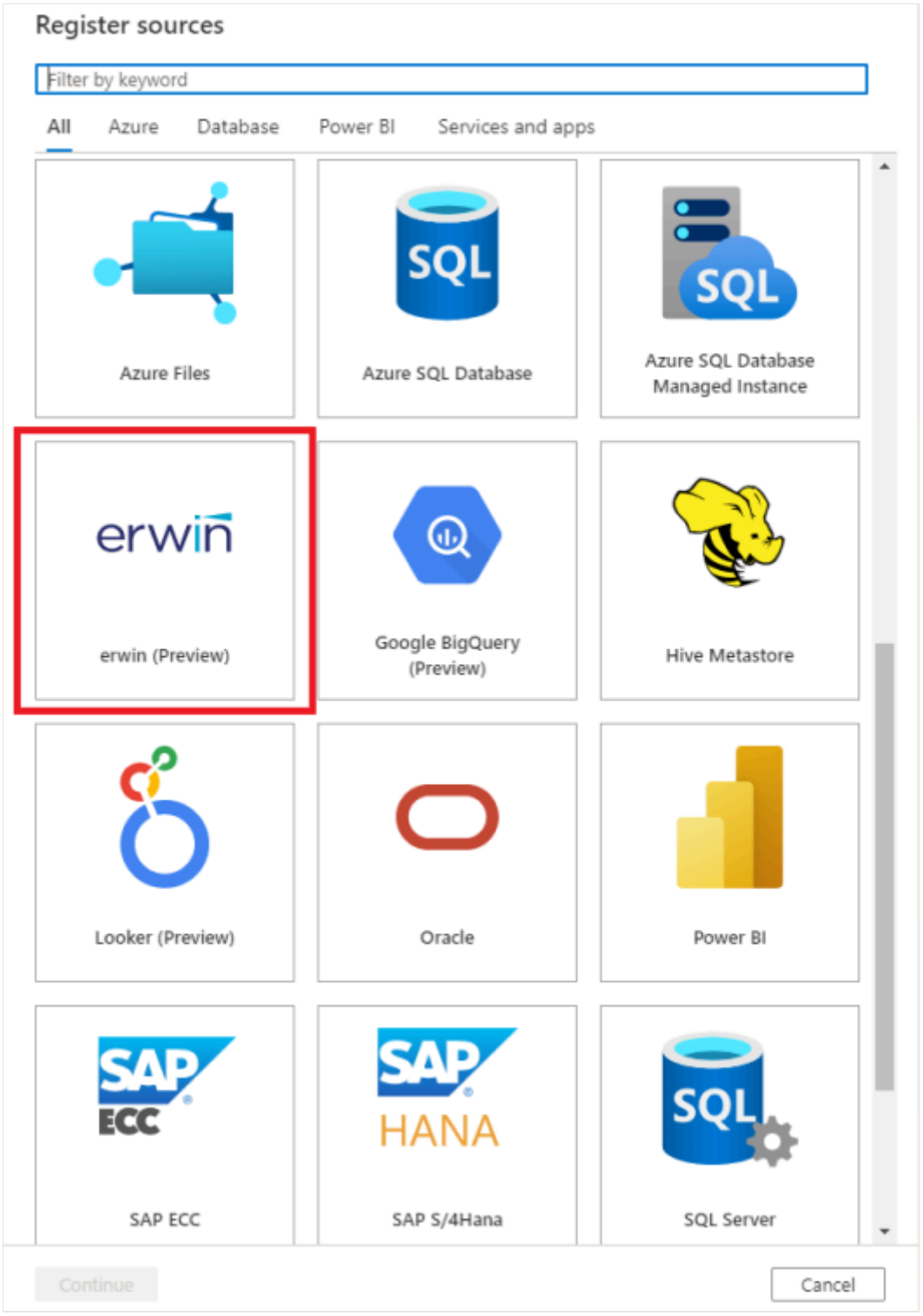
This section describes how to register erwin Mart servers in Microsoft Purview using the Microsoft Purview governance portal.

The only supported authentication for an erwin Mart source is **Server Authentication** in the form of username and password.

Steps to register

1. Navigate to your Microsoft Purview account in the Microsoft Purview governance portal.
2. Select **Data Map** on the left navigation.
3. Select **Register**

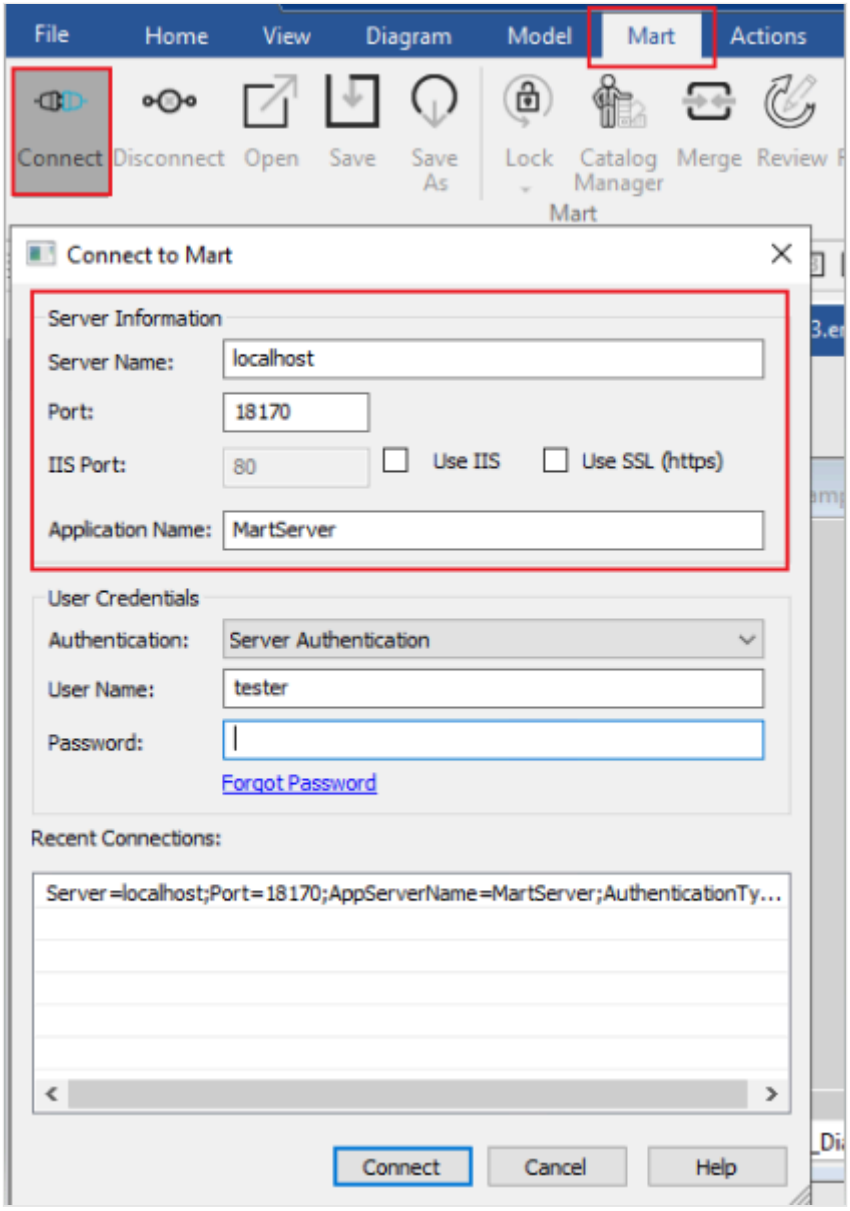
4. On Register sources, select **erwin**. Select **Continue.**



On the Register sources (erwin) screen, do the following:

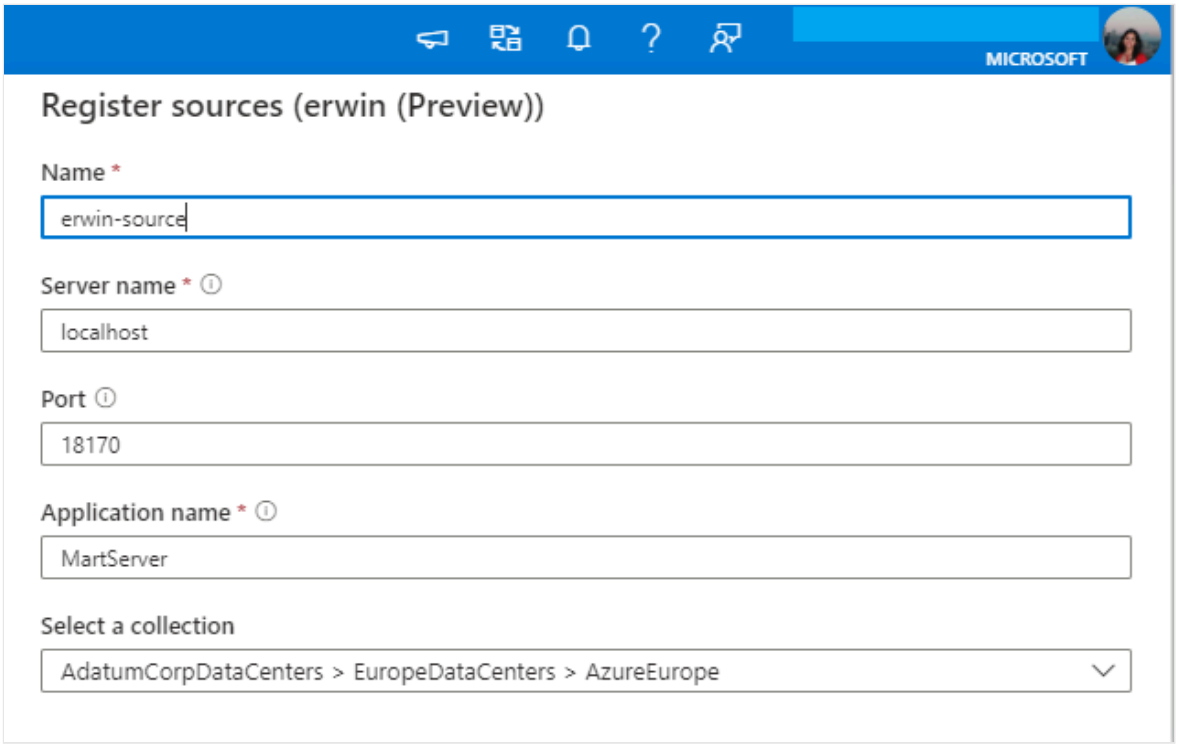
1. Enter a **Name** that the data source will be listed within the Catalog.
2. Enter the erwin Mart **Server name.** This is the network host name used to connect to the erwin Mart server. For example, localhost
3. Enter the **Port** number used when connecting to erwin Mart. By default, this value is 18170.
4. Enter the **Application name**

> ⓘ **Note**
>
> The above details can be found by navigating to your erwin Data Modeler. Select Mart -> Connect to see details related to server name, port and application name.



5. Select a collection from the list.

6. Finish to register the data source.



Scan

Follow the steps below to scan erwin Mart servers to automatically identify assets. For more information about scanning in general, see our introduction to scans and ingestion

Create and run scan

To create and run a new scan, do the following:

1. In the Management Center, select Integration runtimes. Make sure a self-hosted integration runtime is set up on the VM where erwin Mart instance is running. If it isn't set up, use the steps mentioned here to set up a self-hosted integration runtime.

2. Navigate to **Sources**.

3. Select the registered **erwin** Mart.

4. Select + **New scan**.

5. Provide the below details:

    a. **Name**: The name of the scan

    b. **Connect via integration runtime**: Select the configured self-hosted integration runtime.

c. **Server name, Port** and **Application name** are auto populated based on the values entered during registration.

d. **Credential:** Select the credential configured to connect to your erwin Mart server. While creating a credential, make sure to:

   - Select **Basic Authentication** as the Authentication method
   - Provide your erwin Mart server's username in the User name field.
   - Save your user password for server authentication in the key vault's secret.

   To understand more on credentials, refer to the link here.

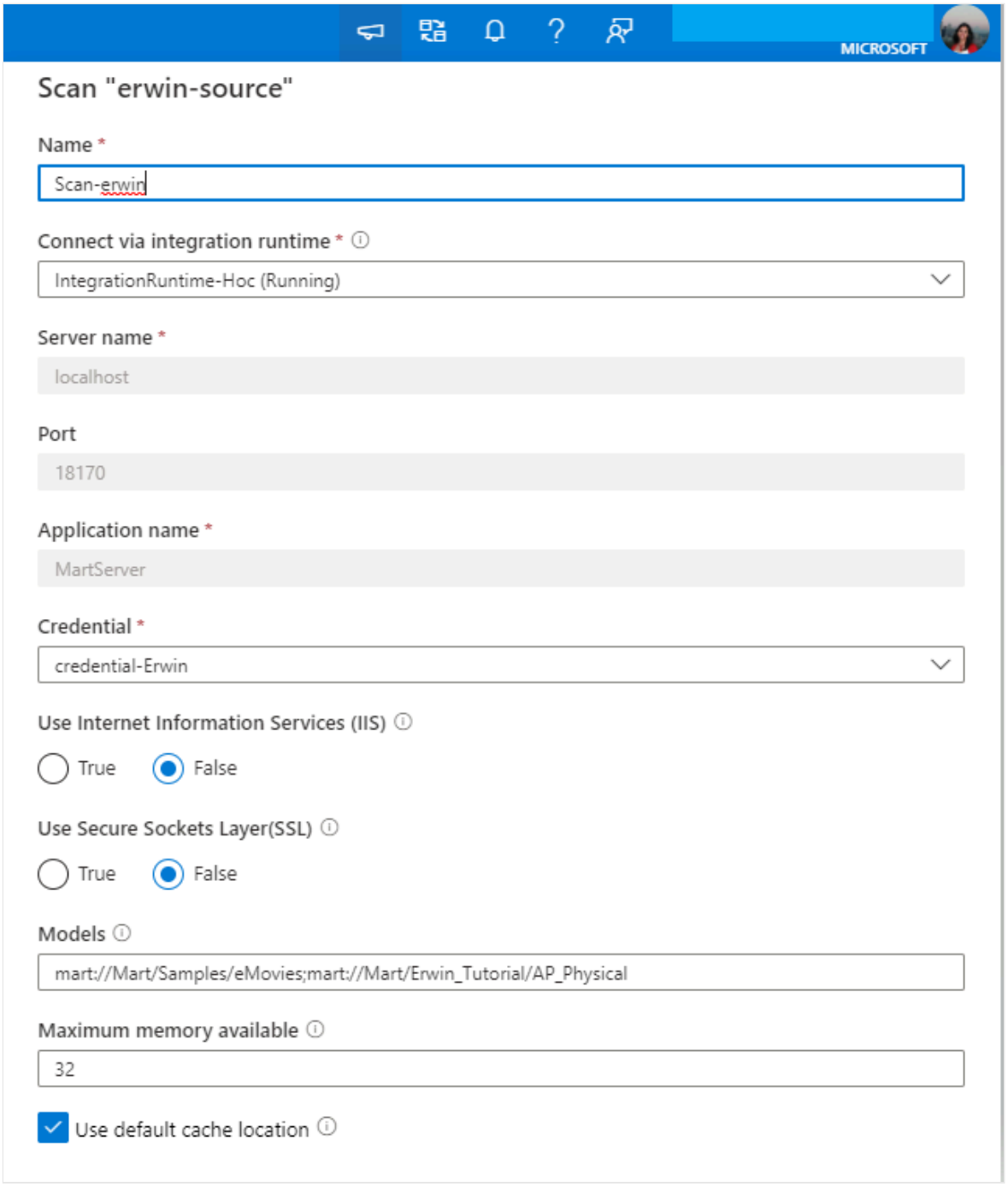
e. **Use Internet Information Services (IIS)** - Select True or False to notify if Microsoft Internet Information Services (IIS) must be used when connecting to the erwin Mart server. By default this value is set to False.

f. **Use Secure Sockets Layer (SSL)** - Select True or False to Notify if Secure Sockets Layer (SSL) must be used when connecting to the erwin Mart server. By default, this value is set to False.

> ⓘ **Note**
>
> This parameter is only applicable for erwin Mart version 9.1 or later.

g. **Models** - Scope your scan by providing a semicolon separated list of erwin model locator strings. For example, mart://Mart/Samples/eMovies;mart://Mart/Erwin_Tutorial/AP_Physical

h. **Maximum memory available**: Maximum memory (in GB) available on customer's VM to be used by scanning processes. This is dependent on the size of erwin Mart to be scanned.



6. Select **Test connection.**

7. Select **Continue**.

8. Choose your **scan trigger**. You can set up a schedule or ran the scan once.

9. Review your scan and select **Save and Run**.

View your scans and scan runs

To view existing scans:

1. Go to the Microsoft Purview portal. On the left pane, select **Data map**.

2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that has results you want to view. The pane shows you all the previous scan runs, along with the status and metrics for each scan run.
4. Select the run ID to check the scan run details.

Manage your scans

To edit, cancel, or delete a scan:

1. Go to the Microsoft Purview portal. On the left pane, select **Data Map**.

2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.

3. Select the scan that you want to manage. You can then:

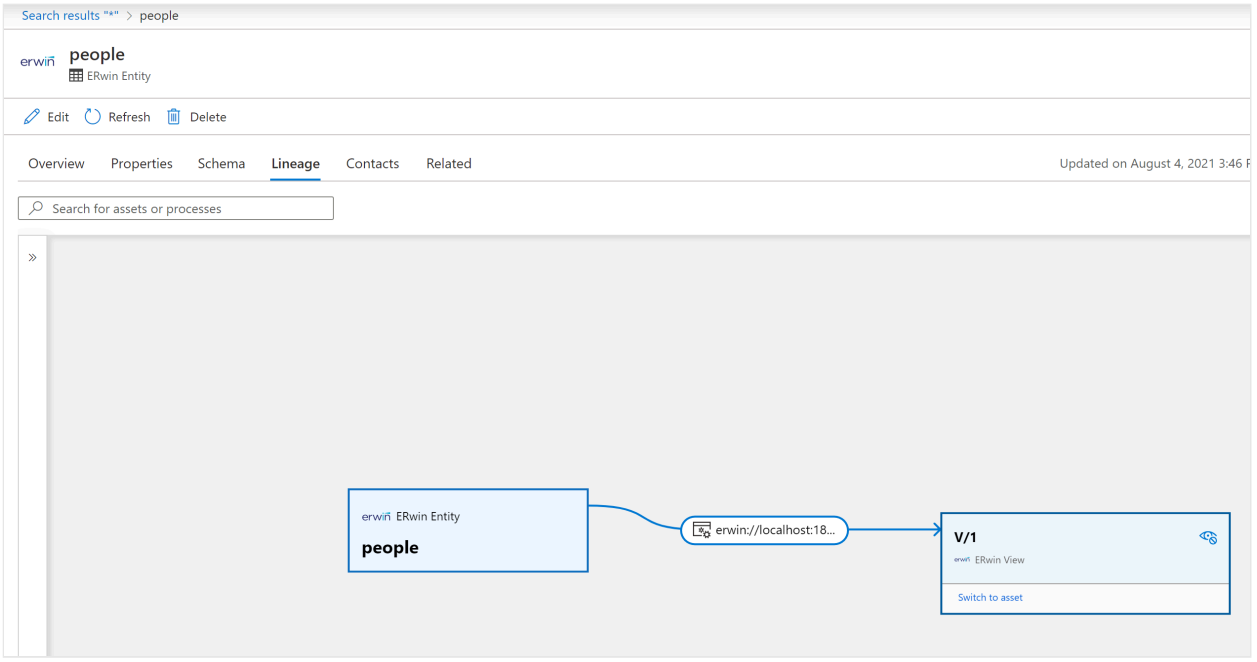
    - Edit the scan by selecting **Edit scan**.
    - Cancel an in-progress scan by selecting **Cancel scan run**.
    - Delete your scan by selecting **Delete scan**.

> ⓘ **Note**
>
> - Deleting your scan does not delete catalog assets created from previous scans.

Lineage

After scanning your erwin source, you can browse Unified Catalog or search Unified Catalog to view the asset details.

Go to the asset -> lineage tab, you can see the asset relationship when applicable. Refer to the supported capabilities section on the supported erwin lineage scenarios. For more information about lineage in general, see data lineage and lineage user guide.



Next steps

Now that you've registered your source, follow the below guides to learn more about Microsoft Purview and your data.

- Data Estate Insights in Microsoft Purview
- Lineage in Microsoft Purview
- Search Unified Catalog

Feedback

Was this page helpful? Yes No

Connect to your Microsoft Fabric tenant in the same tenant as Microsoft Purview

This article outlines how to register a Microsoft Fabric tenant that's in **the same tenant** as your Microsoft Purview resource, and how to authenticate and interact with the Fabric source in Microsoft Purview. For more information about Microsoft Purview in general, read the introductory article.

> **Note**
>
> Scanning a Fabric tenant brings in metadata and lineage from Fabric items including Power BI. To learn more, see scan a Power BI tenant. The experience of registering a Fabric tenant and setting up a scan is similar to Power BI tenant and shared among all Fabric items.

Supported capabilities

Scanning capabilities

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan
Yes	Yes	Yes	Yes

Other capabilities

For **classifications**, **sensitivity labels**, **policies**, **data lineage**, and **live view**, see the list of supported capabilities.

Expand table

Experiences	Fabric items	Available in scan	Available in live view (same tenant only)
Real-Time Analytics	Kusto Query Language (KQL) Database	Yes	Yes
	KQL Queryset	Yes	Yes

Experiences	Fabric items	Available in scan	Available in live view (same tenant only)
Data Science	Experiment	Yes	Yes
	ML model	Yes	Yes
Data Factory	Data pipeline	Yes	Yes
	Dataflow Gen2	Yes	Yes
Data Engineering	Lakehouse	Yes	Yes
	Notebook	Yes	Yes
	Spark Job Definition	Yes	Yes
	SQL analytics endpoint	Yes	Yes
Data Warehouse	Warehouse	Yes	Yes
Power BI	Dashboard	Yes	Yes
	Dataflow	Yes	Yes
	Datamart	Yes	Yes
	Semantic model (Dataset)	Yes	Yes
	Report	Yes	Yes
	Paginated report	Yes	

Supported scenarios for Fabric scans

Expand table

Scenarios	Microsoft Purview public access allowed/denied	Fabric public access allowed /denied	Runtime option	Authentication option	Deployment checklist
Public access with Azure IR	Allowed	Allowed	Azure Runtime	Managed Identity/ Delegated authentication / Service principal	Review deployment checklist
Public access with	Allowed	Allowed	SHIR or Kubernetes SHIR	Service principal	Review deployment checklist

Scenarios	Microsoft Purview public access allowed/denied	Fabric public access allowed /denied	Runtime option	Authentication option	Deployment checklist
Self-hosted IR					
Private access	Denied	Allowed	Managed Virtual Network IR (v2 only)	Managed Identity / Delegated authentication / Service principal	Review deployment checklist

> **Note**
>
> The supported scenarios above apply for non-Power BI items in Fabric. See the **Power BI documentation** for supported scenarios that apply to Power BI items.
>
> Support for scanning Fabric Lakehouses by using Managed Identity is now available. Hence, you can scan Fabric Lakehouses by using Service Principal or Managed Identity. Previously, only service principal was supported for Fabric Lakehouses. There's no change for other Fabric items.
>
> Support for scanning Fabric Lakehouses by using Managed Virtual Network IR is now available.

Known limitations

- Currently for all Fabric items besides Power BI, only item level metadata and lineage can be scanned. For Lakehouse tables and files, sub-item level metadata scanning is available but sub-item level lineage isn't supported.
- For self-hosted integration runtime, the standard self-hosted integration runtime with minimal version 5.47.9073.1 or Kubernetes supported self-hosted integration runtime is supported.
- Empty workspaces are skipped.
- Microsoft Purview doesn't support Fabric DAX expressions as part of metadata scanning.
- You might get a **Too many requests** error while running multiple Fabric or Power BI scans. If so, reduce the concurrent scans and scatter the scans over different times to help resolve

the error.

- In the incremental scan, you might see a non-zero count of assets ingested and assets discovered, even if there's no change in the data source. This behavior should be ignored. The data catalog and Data Map assets continue to reflect the true state of the data source.
- Fabric with tenant-level or workspace-level private links aren't supported.

> **Note**
>
> For Fabric Lakehouse tables with schema-enabled databases, the schema name and table name are combined into a single segment in the qualified name using `%252F` (double-URL-encoded forward slash) as a separator. For example, a table at the OneLake path `Tables/dbo/my_table` will appear in Data Map with the qualified name `.../tables/dbo%252Fmy_table`, while the entity display name will show only the table name (for example, `my_table`) without the schema prefix.

Prerequisites

Before you start, make sure you have the following prerequisites:

- An Azure account with an active subscription. Create an account for free.
- An active Microsoft Purview account.

Authentication options

- Managed Identity
- *Service Principal
- Delegated Authentication

Deployment checklist

Use any of the following deployment checklists during the setup or for troubleshooting purposes, based on your scenario:

Public access with Azure IR

Scan same-tenant Fabric using Azure IR and Managed Identity in public network

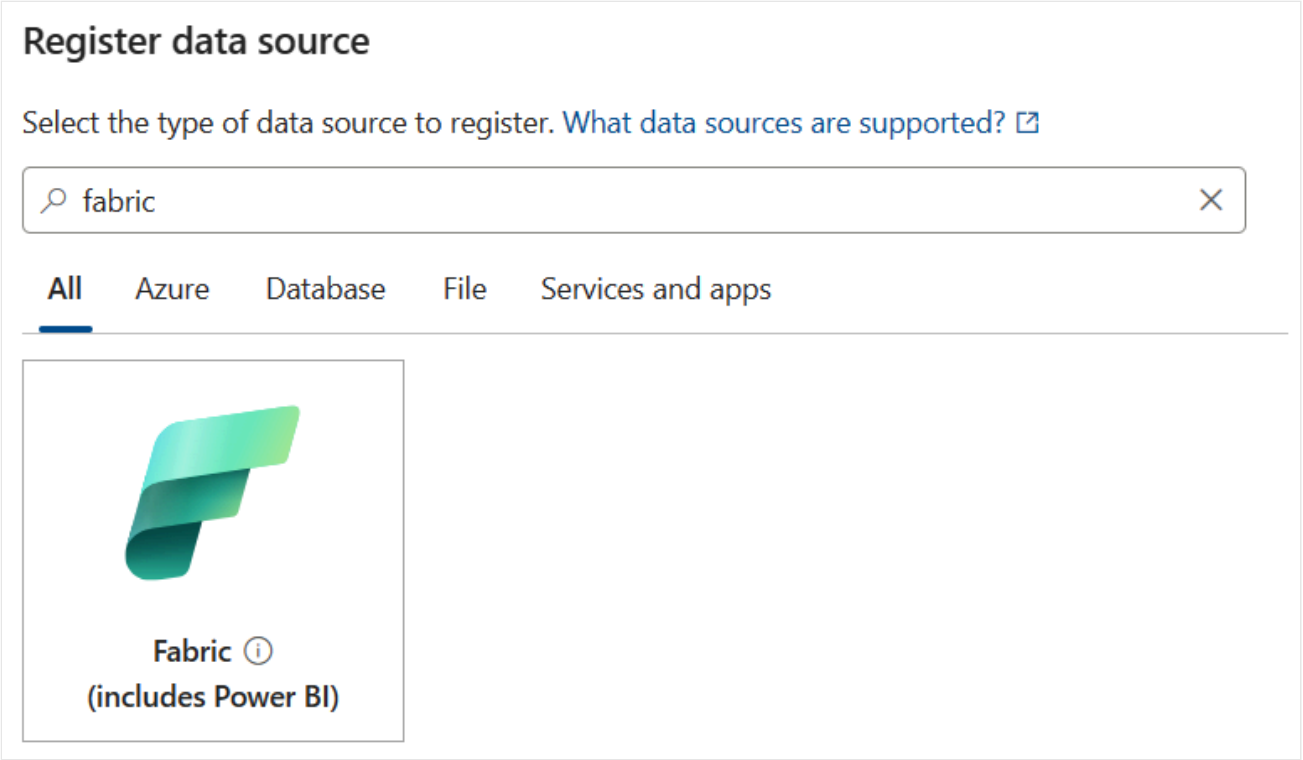
1. Make sure Fabric and Microsoft Purview accounts are in the same tenant.
2. Make sure you enter the Fabric tenant ID correctly during the registration.
3. Make sure your Fabric Metadata model is up to date by enabling metadata scanning.
4. From Azure portal, validate if Microsoft Purview account Network is set to public access.
5. From Fabric tenant Admin Portal, make sure Fabric tenant is configured to allow public network.
6. In Microsoft Entra tenant, create a security group.
7. From Microsoft Entra tenant, make sure Microsoft Purview account MSI is member of the new security group.
8. On the Fabric Tenant Admin portal, validate if Allow service principals to use read-only admin APIs is enabled for the new security group.
9. On the Fabric Tenant Admin portal, validate if **Tenant Settings** -> **OneLake Settings** -> Users can access data stored in OneLake with apps external to Fabric is enabled.

Register Fabric tenant

This section describes how to register a Fabric tenant in Microsoft Purview for same-tenant scenario.

1. Select the **Data Map** on the left navigation.
2. Select **Register**.

   Select **Fabric** as your data source.



3. Enter a friendly name for your Fabric instance and select a collection.

   The name must be between 3-63 characters long and can only contain letters, numbers, underscores, and hyphens. Don't use spaces.

   By default, the system finds the Fabric tenant that exists in the same Microsoft Entra tenant.

Authentication to scan

To scan your Fabric tenant and review its metadata, you first need to create a way to authenticate by using the Fabric tenant. Then, provide Microsoft Purview with the authentication information.

Authenticate to Fabric tenant

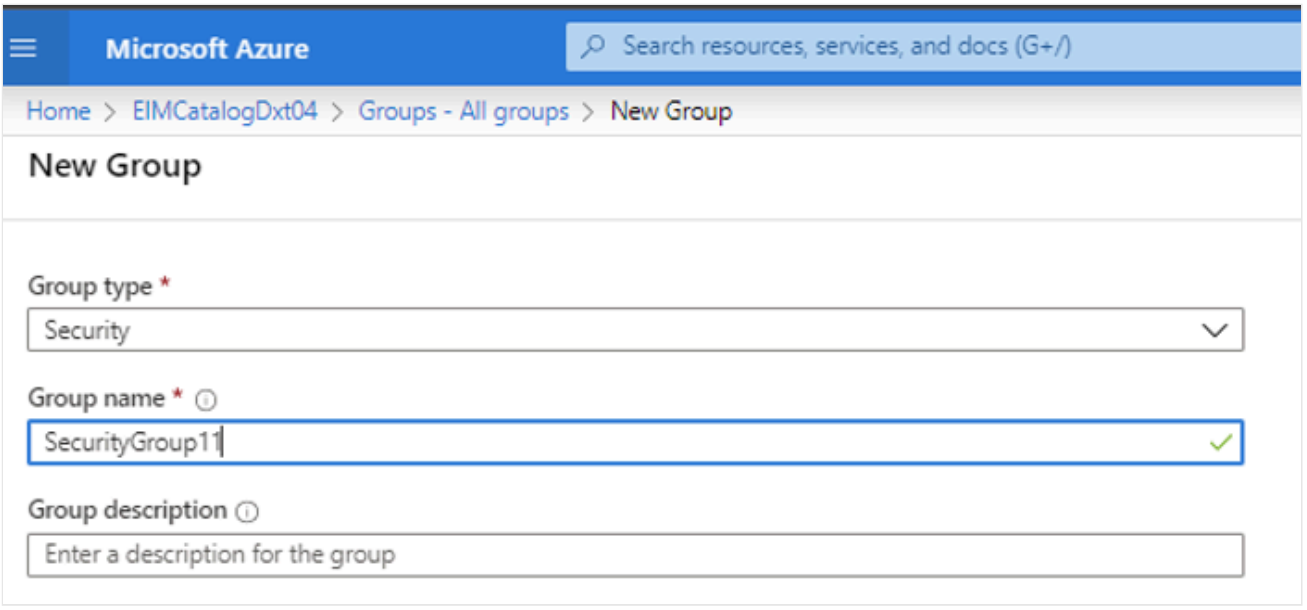
In the Microsoft Entra tenant where the Fabric tenant is located:

1. In the Azure portal, search for **Microsoft Entra ID**.

2. Create a new security group in your Microsoft Entra ID, by following Create a basic group and add members using Microsoft Entra ID.

   > **Tip**
   >
   > You can skip this step if you already have a security group you want to use.

3. Select **Security** as the **Group Type**.

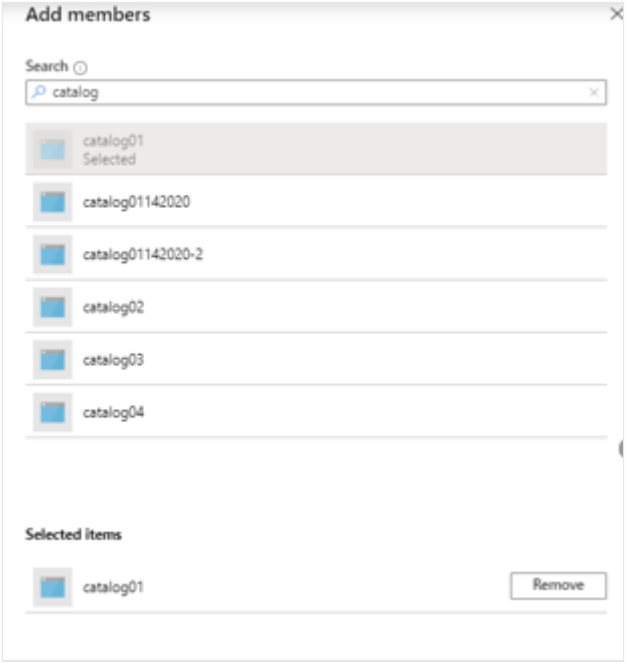


4. Add any relevant users to the security group:

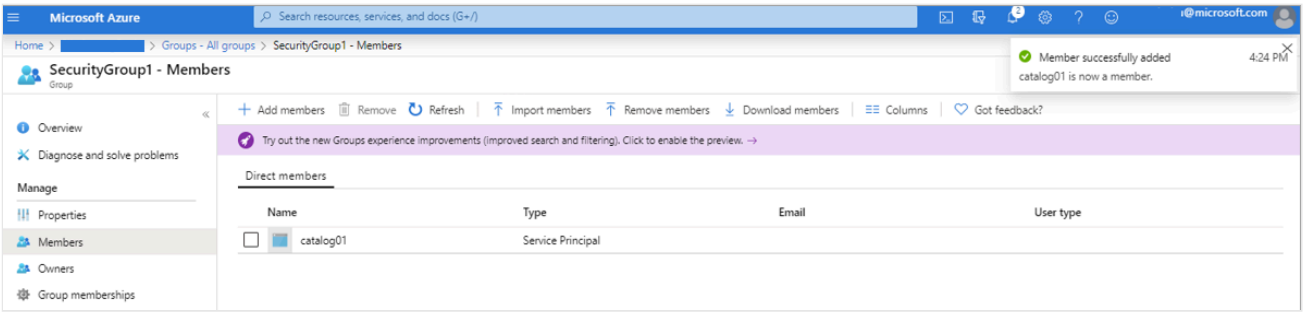
   - If you're using **Managed Identity** as the authentication method, add your Microsoft Purview managed identity to this security group. Select **Members**, and then select **Add members**.

   - If you're using **delegated authentication** or **service principal** as the authentication method, add your **service principal** to this security group. Select **Members**, and then select **Add members**.

5. Search for your Microsoft Purview managed identity or service principal and select it.



You see a success notification that shows you it was added.



Associate the security group with Fabric tenant

1. Sign in to the Fabric admin portal.

2. Select the **Tenant settings** page.

   > **Important**
   >
   > You need to be a Fabric Admin to see the tenant settings page.

3. Select **Admin API settings** > **Allow service principals to use read-only admin APIs**.

4. Select **Specific security groups**.

5. Select **Admin API settings** > **Enhance admin APIs responses with detailed metadata** and **Enhance admin APIs responses with DAX and mashup expressions** > Enable the toggle to allow Microsoft Purview Data Map automatically discover the detailed metadata of Fabric datasets as part of its scans.

   > **Important**
   >
   > After you update the Admin API settings on your Fabric tenant, wait about 15 minutes before registering a scan and testing connection.

   > **Caution**
   >
   > When you allow the security group you created (that has your Microsoft Purview managed identity as a member) to use read-only admin APIs, you also allow it to access the metadata (such as dashboard and report names, owners, and descriptions) for all of your Fabric artifacts in this tenant. Once the metadata is pulled into Microsoft

Purview, Microsoft Purview's permissions, not Fabric permissions, determine who can see that metadata.

> **Note**
>
> - You can remove the security group from your developer settings, but the metadata previously extracted won't be removed from the Microsoft Purview account. You can delete it separately if you wish.
> - If your Fabric item (such as lakehouse or mirrored catalog) has OneLake security enabled, then you must create a OneLake security role with **Read** permission assigned to the Microsoft Purview user. See Get started with OneLake Security for step-by-step instructions.

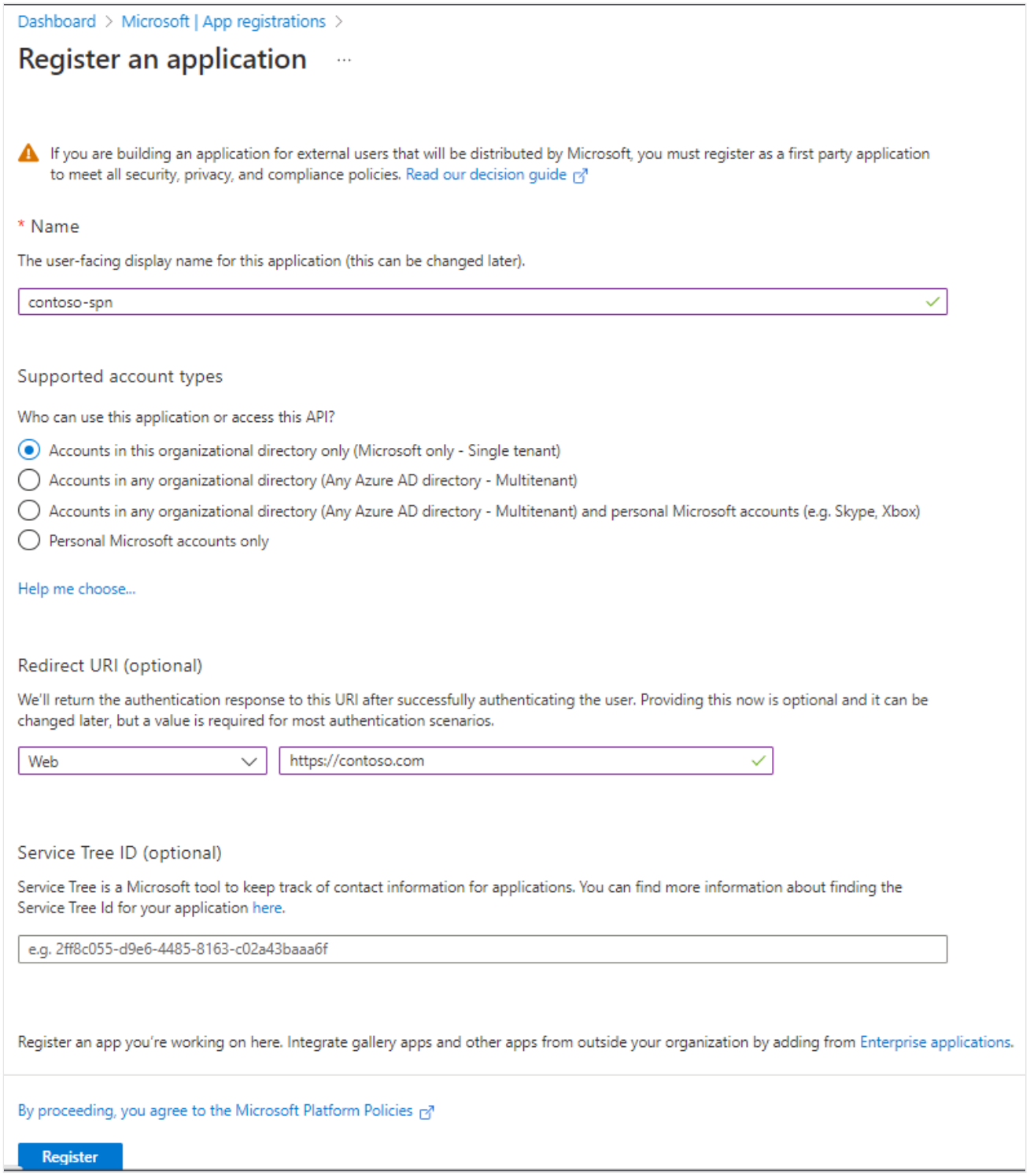
Configure credentials for scans in Microsoft Purview

Managed identity

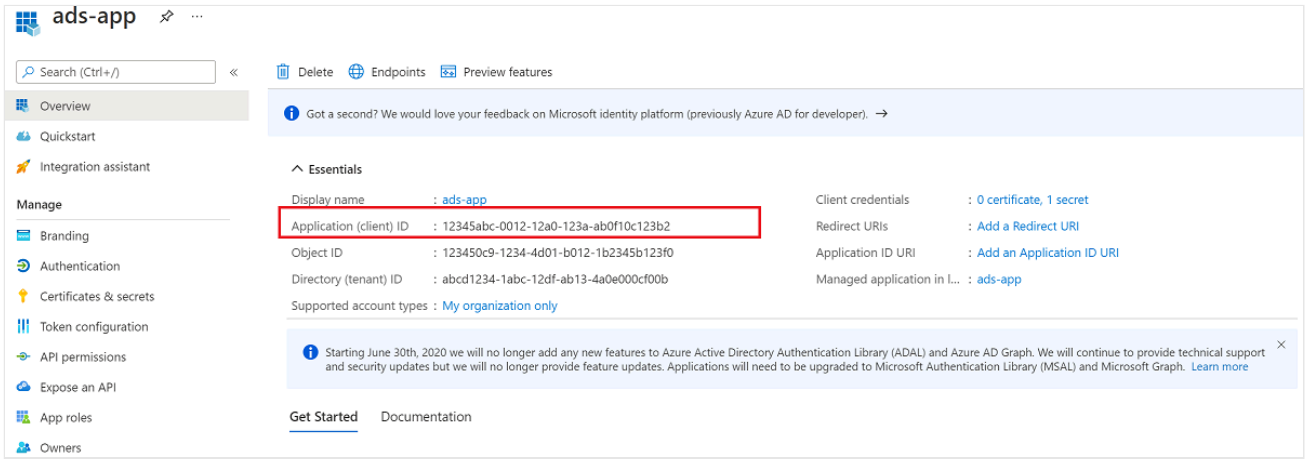
If you follow all the steps in authenticate Microsoft Purview to Fabric, you don't need any other authentication steps for managed identities.

Service principal

1. In the Azure portal, select **Microsoft Entra ID** and create an App Registration in the tenant. Provide a web URL in the **Redirect URI**. For information about the Redirect URI see this documentation from Microsoft Entra ID.

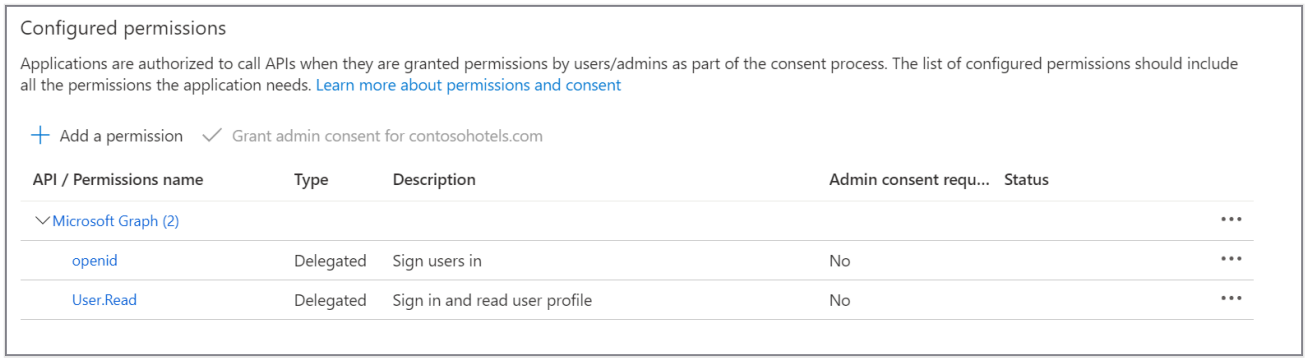


2. Take note of Client ID(App ID).



3. From Microsoft Entra dashboard, select newly created application and then select **App registration**. From **API Permissions**, assign the application the following delegated permissions:

    - Microsoft Graph openid
    - Microsoft Graph User.Read

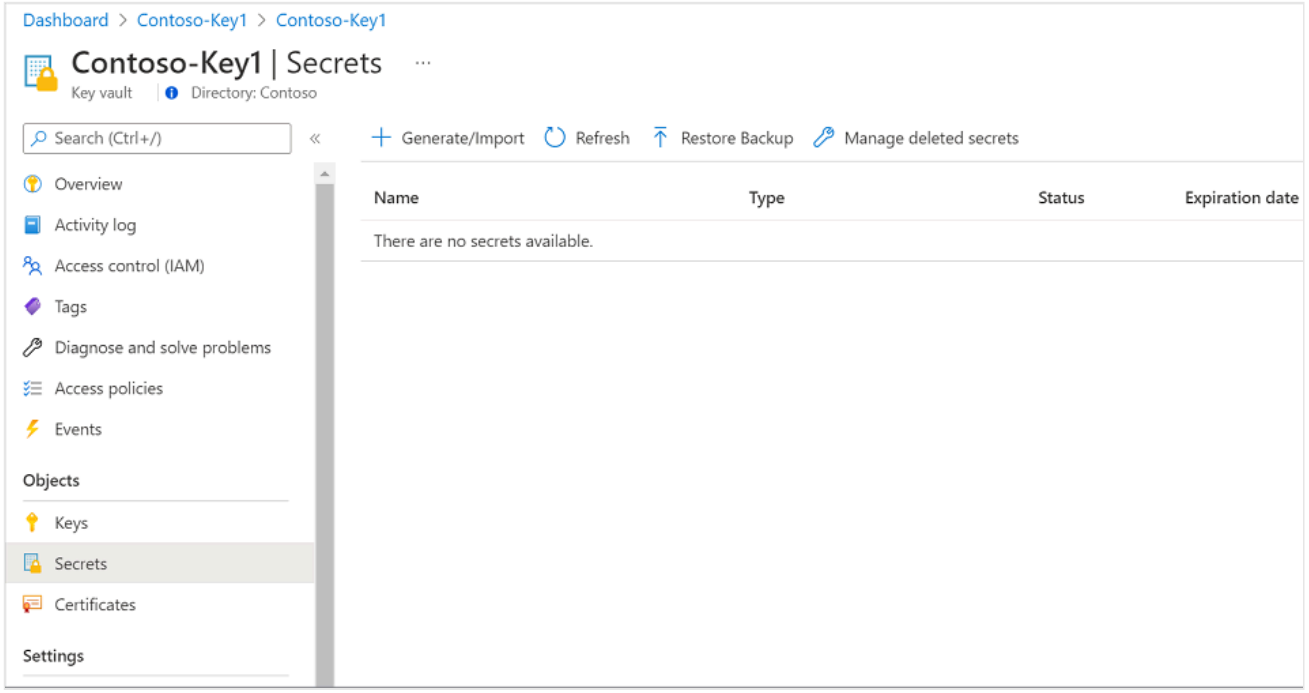


4. Under **Advanced settings**, enable **Allow Public client flows**.

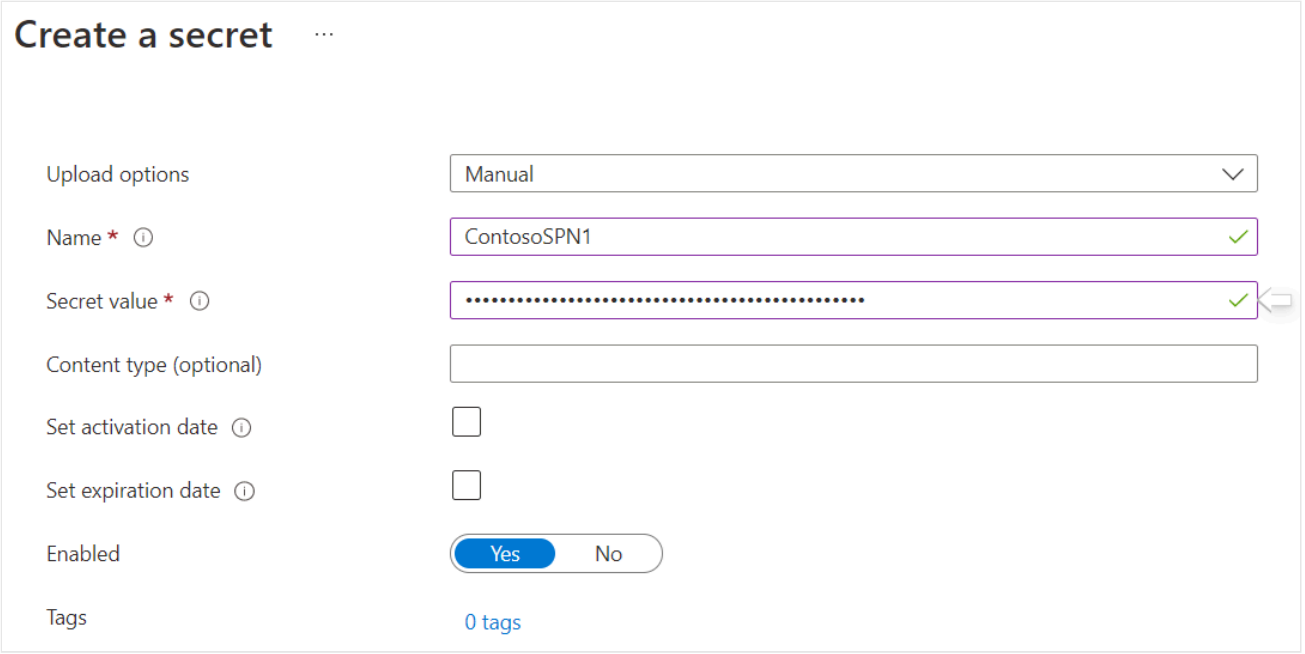
5. Under **Certificates & secrets**, create a new secret and save it securely for next steps.

6. In Azure portal, navigate to your Key Vault.

7. Select **Settings** > **Secrets** and select **+ Generate/Import**.



8. Enter a name for the secret and for **Value**, type the newly created secret for the App registration. Select **Create** to complete.



9. If your key vault isn't connected to Microsoft Purview yet, you need to create a new key vault connection.

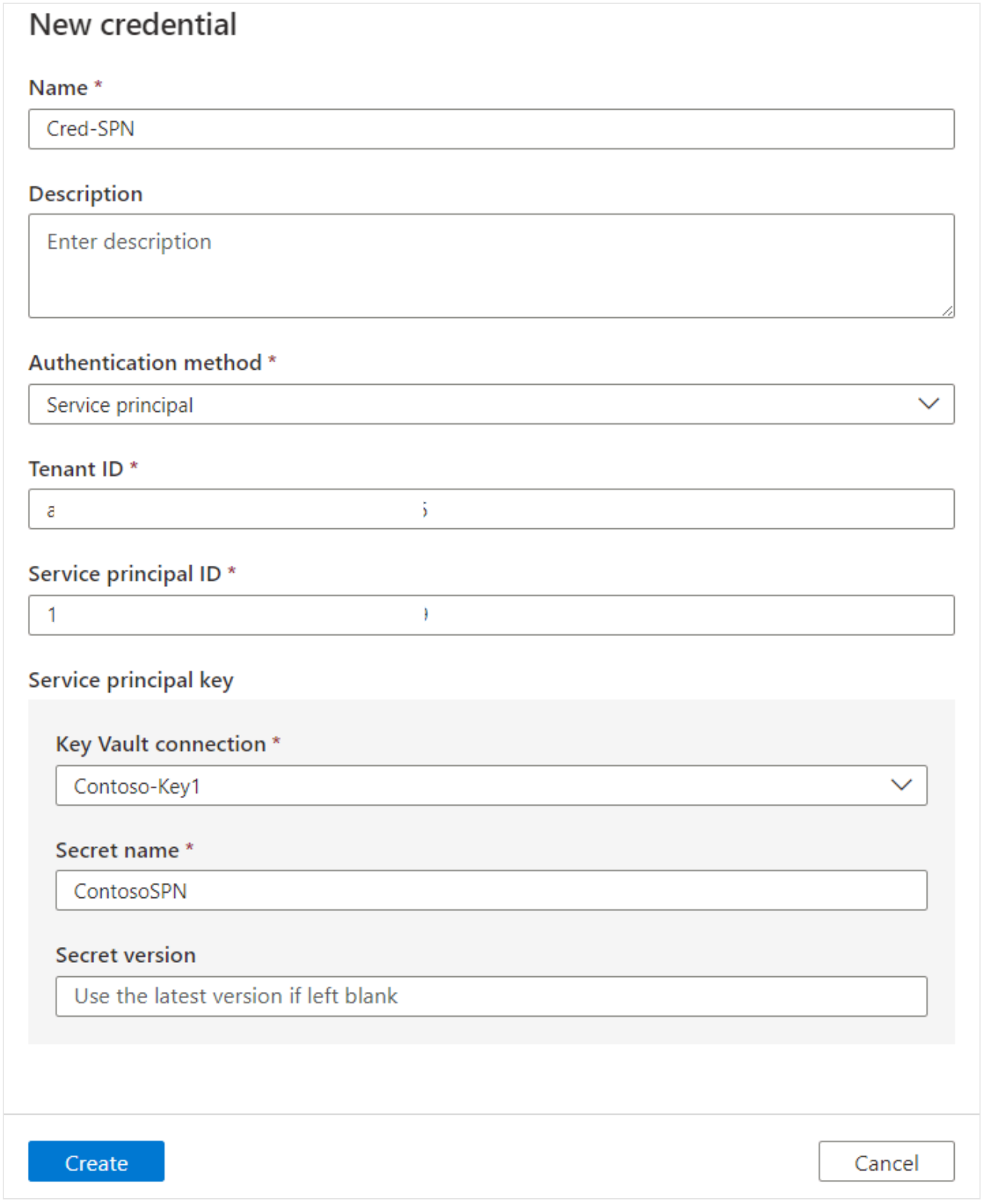
10. Now that you created your secret, in Microsoft Purview, go to the **Credentials** page under **Management**.

    > **Tip**
    >
    > Alternatively, you can create a new credential during the scanning process.

11. Select **New** to create your new credential.

12. Provide the required parameters:

    - **Name**: Provide a unique name for the credential.
    - **Authentication method**: Service principal
    - **Tenant ID**: Your Fabric tenant ID
    - **Client ID**: Use Service Principal Client ID (App ID) you created earlier.
    - **Key Vault Connection**: the Microsoft Purview connection to the Key Vault where you created your secret earlier.
    - **Secret name**: the name of the secret you created earlier.



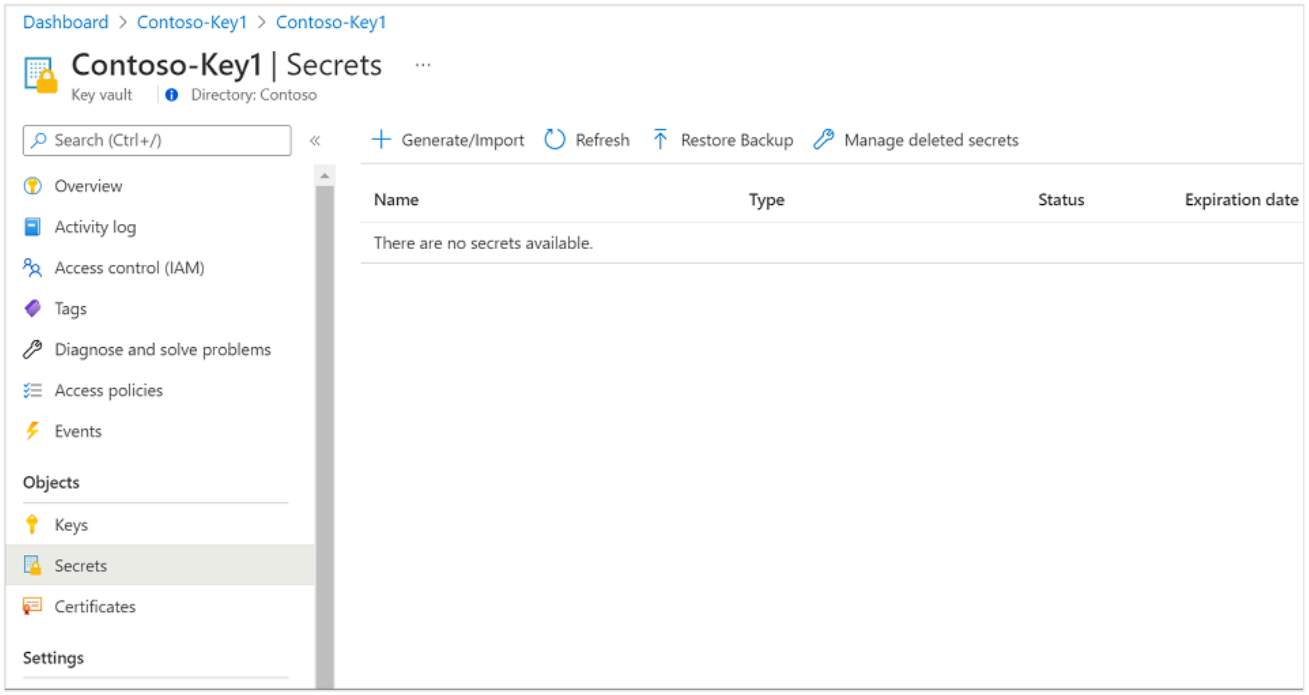
13. Once you fill in all the details, select **Create**.

Delegated authentication

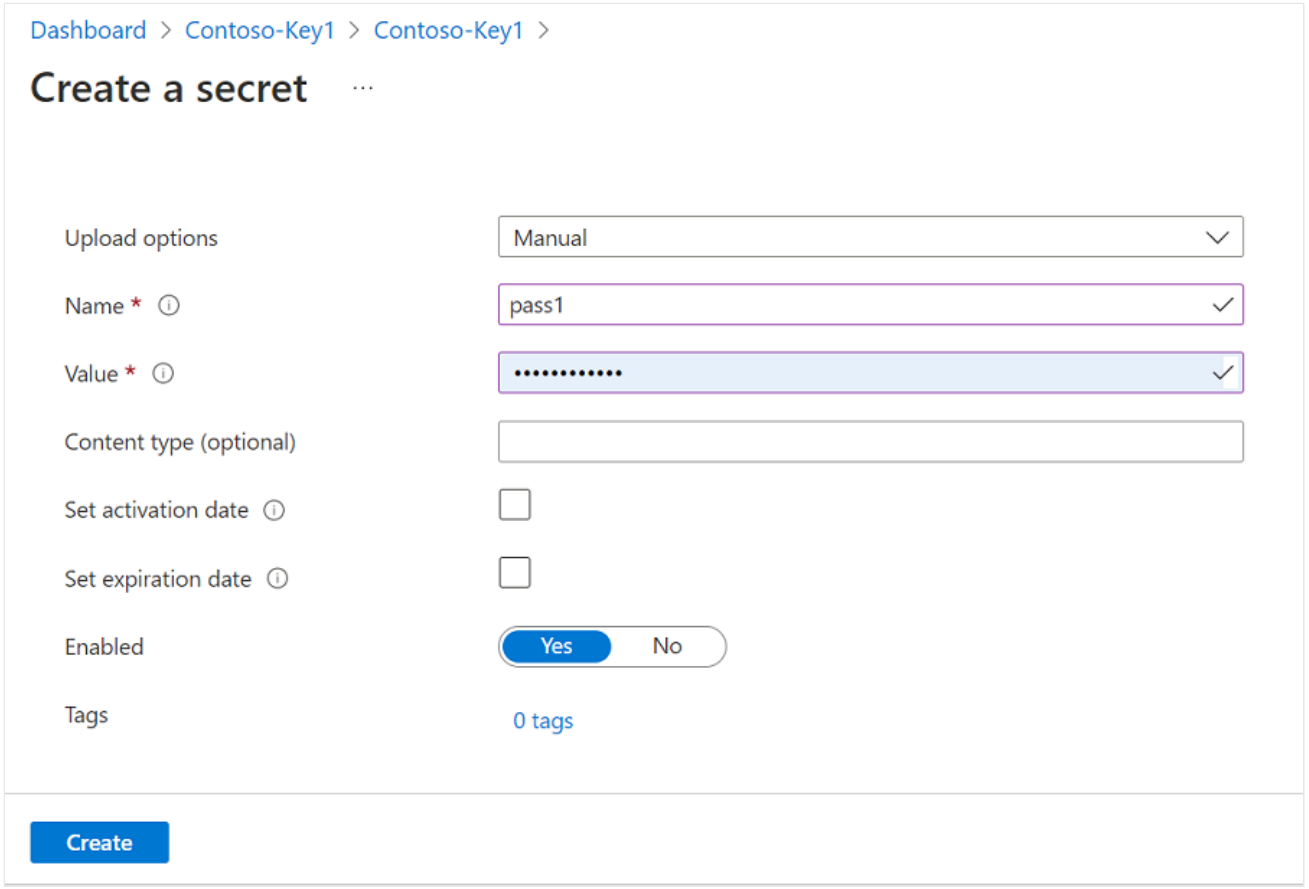
1. Create a user account in Microsoft Entra tenant and assign the user to Microsoft Entra role, **Fabric Administrator**. Take note of username and sign in to change the password.

2. Go to your Key Vault.

3. Select **Settings** > **Secrets** and select **+ Generate/Import**.

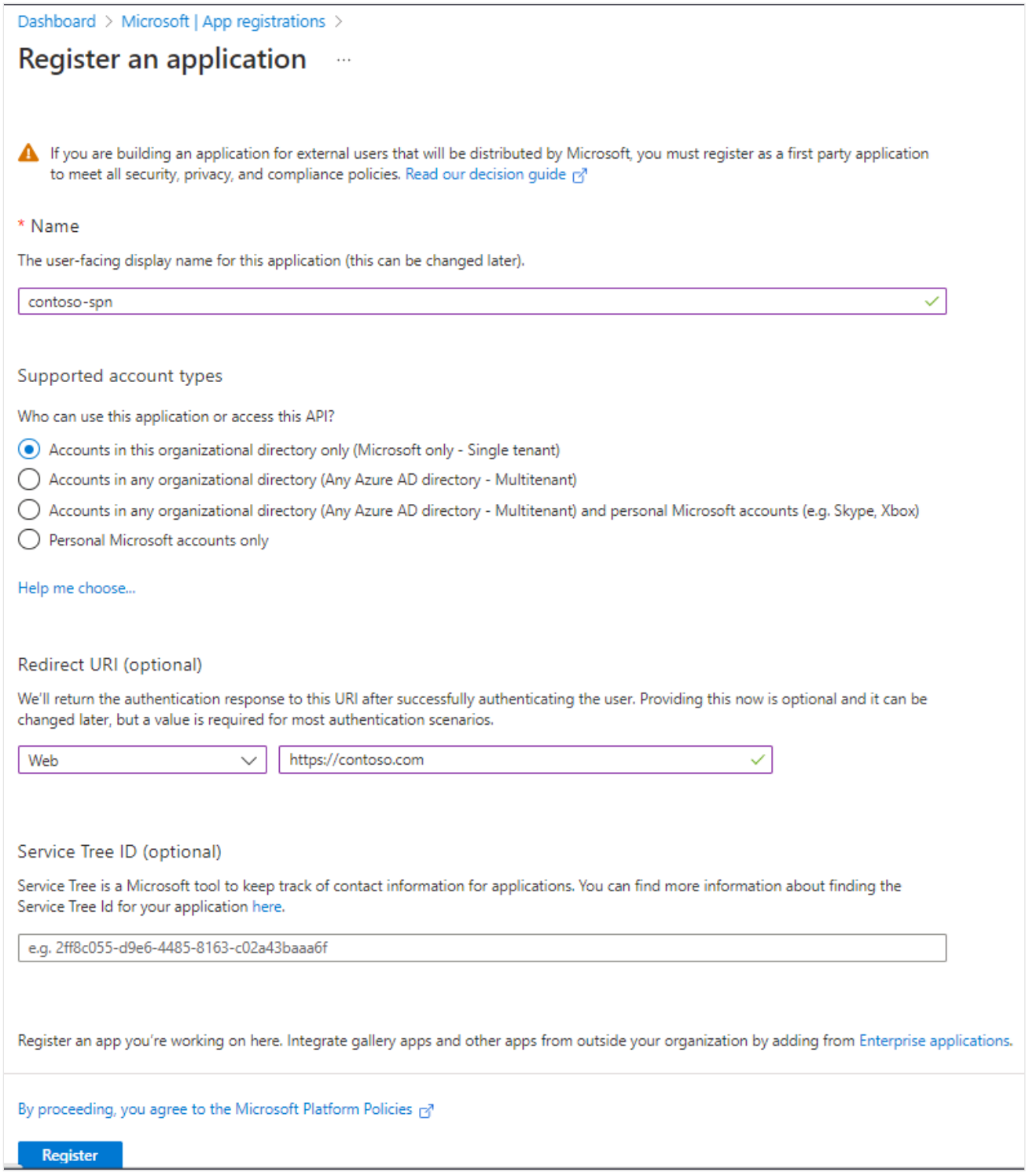


4. Enter a name for the secret and for **Value**, type the newly created password for the Microsoft Entra user. Select **Create** to complete.

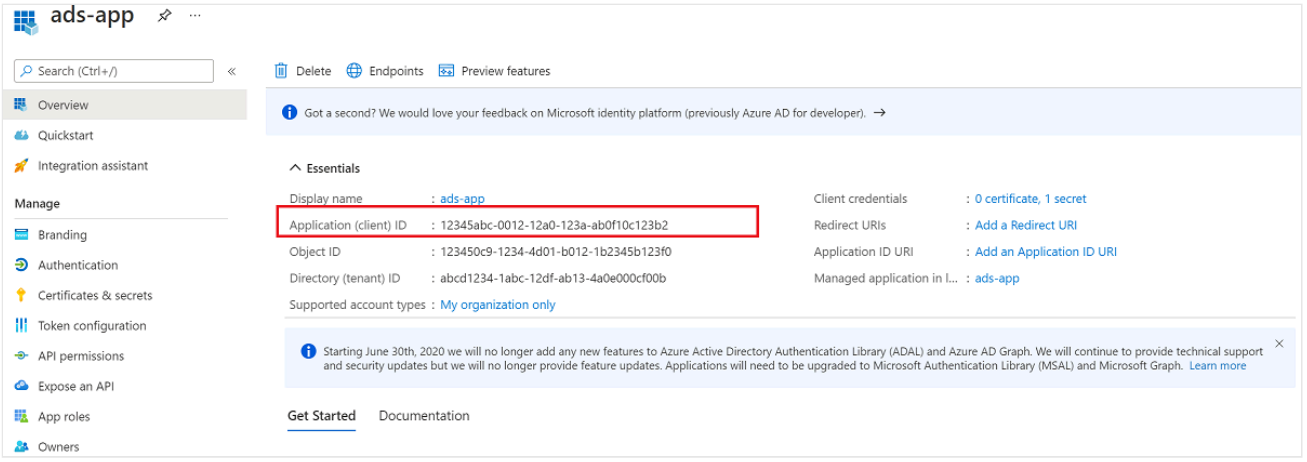


5. If your key vault isn't connected to Microsoft Purview yet, you need to create a new key vault connection

6. Create an app registration in your Microsoft Entra tenant. Provide a web URL in the **Redirect URI**.

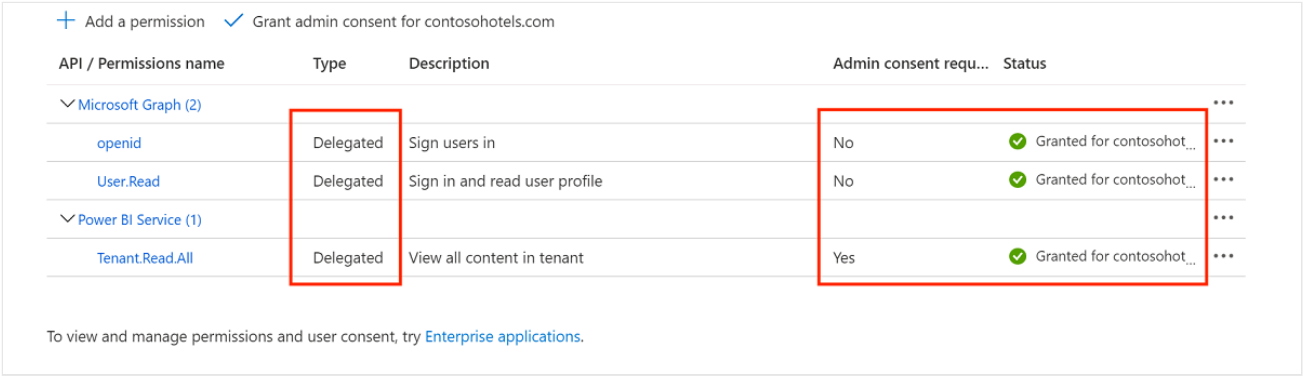


7. Take note of Client ID(App ID).



8. From Microsoft Entra dashboard, select newly created application and then select **API permissions**. Assign the application the following delegated permissions, and grant admin consent for the tenant:

    - Fabric Service Tenant.Read.All
    - Microsoft Graph openid
    - Microsoft Graph User.Read



9. Under **Advanced settings**, enable **Allow Public client flows**.

10. Next, in Microsoft Purview, go to the **Credentials** page under **Management** to create a new credential.

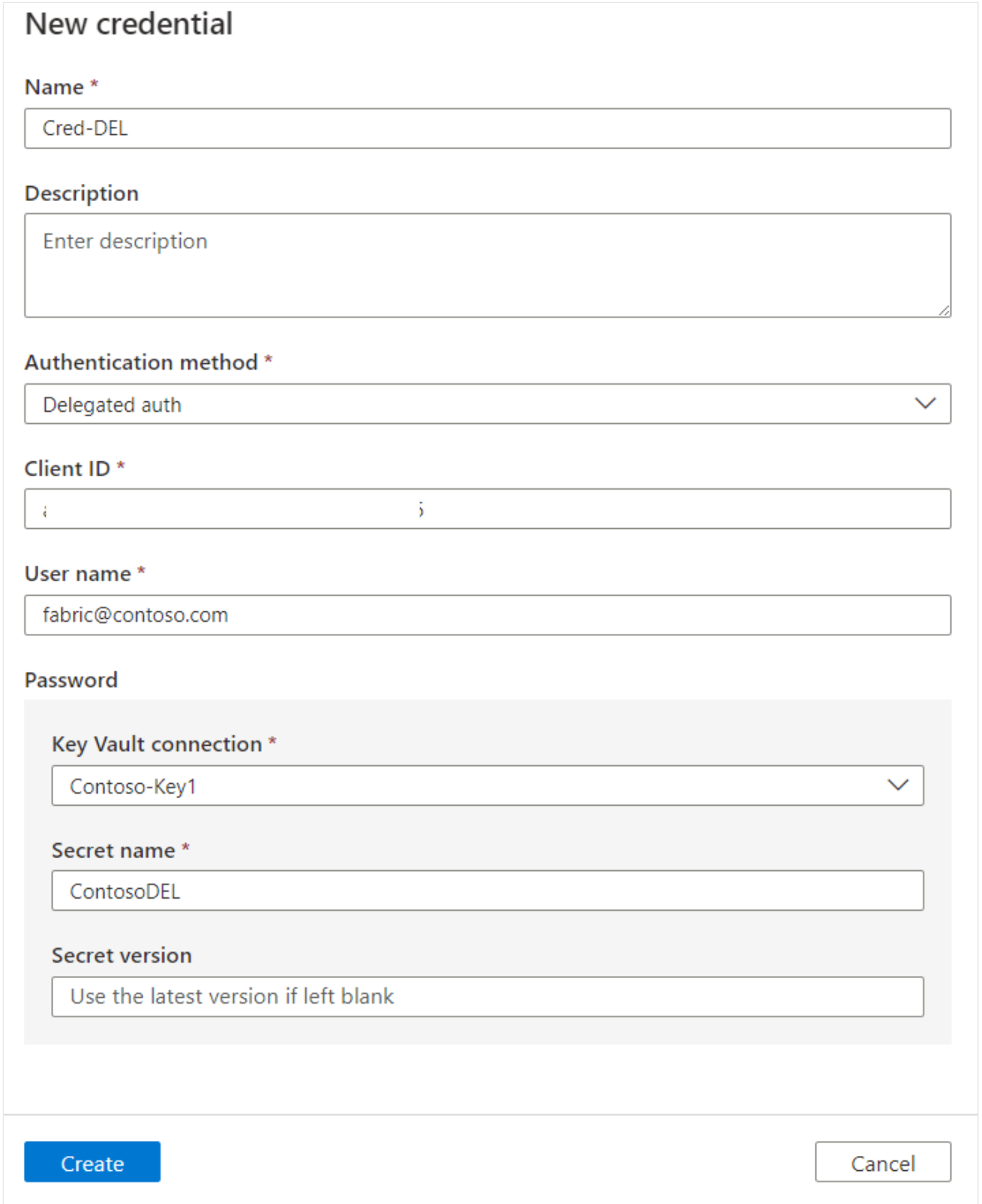
    > **Tip**
    >
    > Alternatively, you can create a new credential during the scanning process.

11. Select **New** to create your new credential.

12. Provide the required parameters:

    - **Name**: Provide a unique name for the credential.
    - **Authentication method**: Delegated auth.

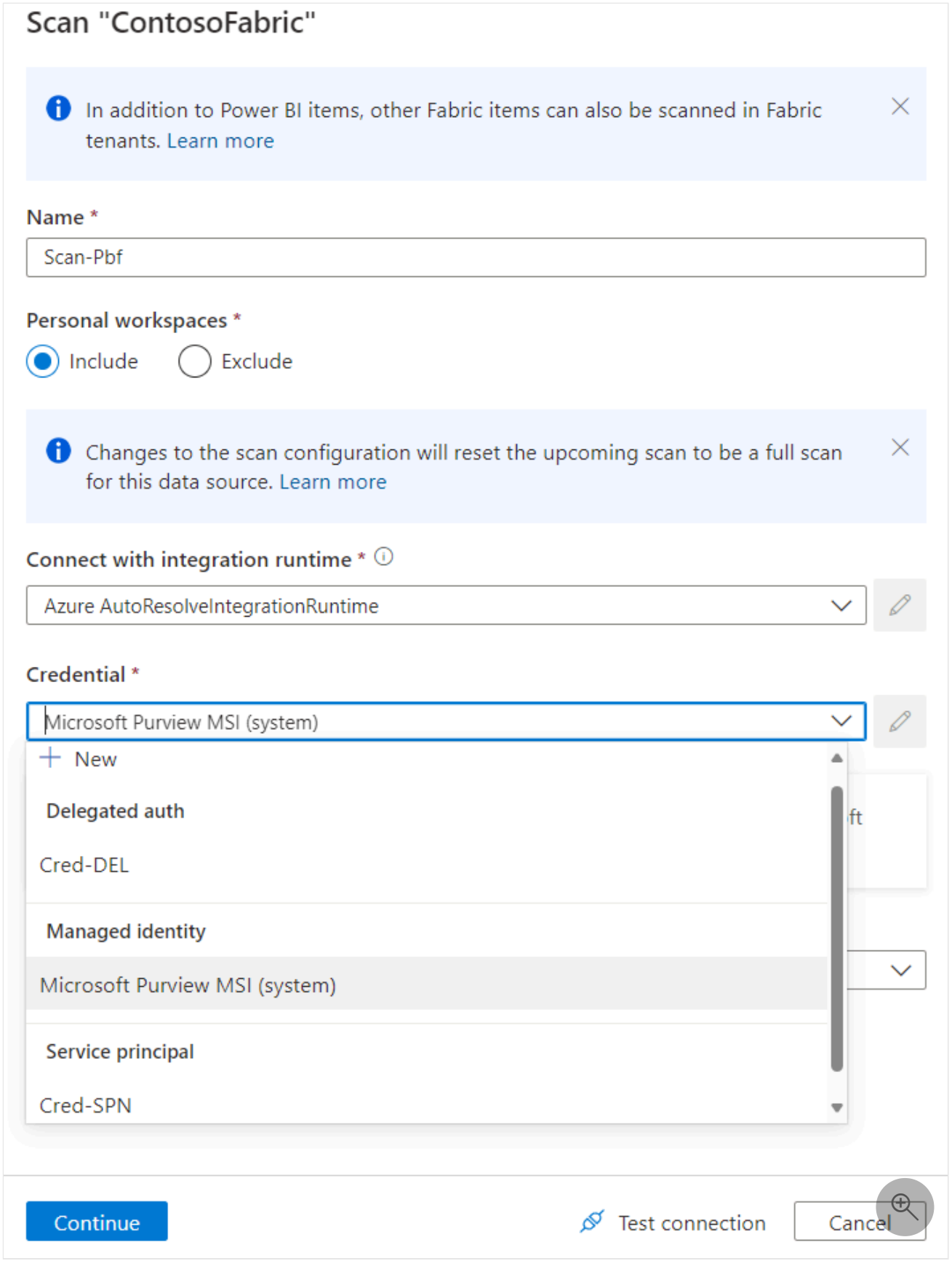
- **Client ID**: Use Service Principal Client ID (App ID) you created earlier.
- **User name**: Provide the username of Fabric Administrator you created earlier.
- **Key Vault Connection**: the Microsoft Purview connection to the Key Vault where you created your secret earlier.
- **Secret name**: the name of the secret you created earlier.



13. Once you fill in all the details, select **Create**.

Create scan

1. In the Microsoft Purview portal, go to **Data Map**.
2. Go to **Sources**.
3. Select the registered Fabric source.
4. Select **+ New scan**.
5. Enter a name for your scan.
6. Select either your managed identity credential or the credential you created for your service principal or delegated authentication.



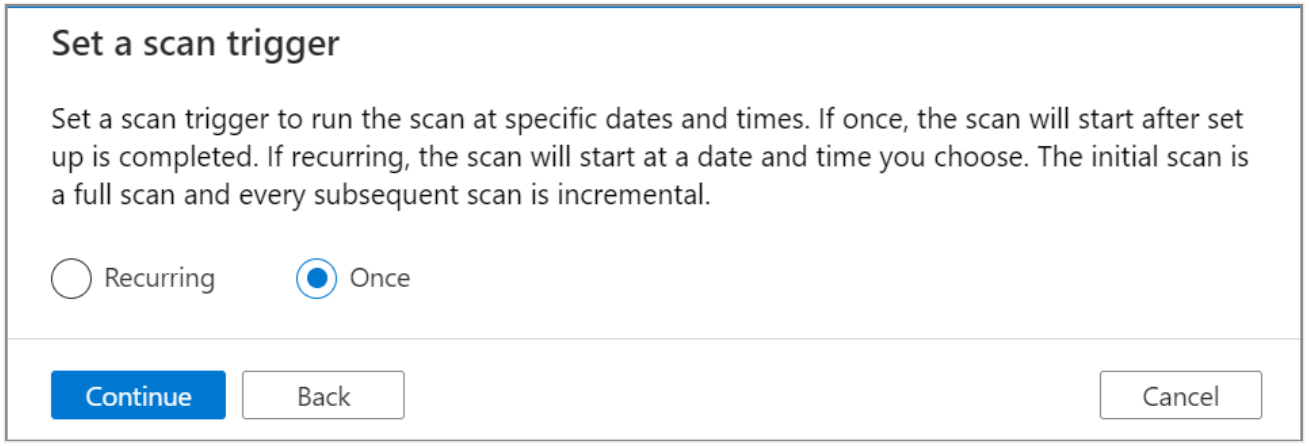
7. Select **Test Connection** before continuing to next steps. If **Test Connection** fails, select **View Report** to see the detailed status and troubleshoot the problem. 1. Access - Failed status means the user authentication failed. Scans that use managed identity always pass because no user authentication is required. If you use service principal or delegated authentication, make sure your Key Vault credential is set up correctly and that Microsoft Purview has

access to the key vault. 1. Assets (+ lineage) - Failed status means the Microsoft Purview - Fabric authorization failed. Make sure the Microsoft Purview managed identity is added to the security group in the Fabric admin portal.

a. Detailed metadata (Enhanced) - Failed status means the Fabric admin portal is disabled for the following setting - **Enhance admin APIs responses with detailed metadata**

> **Tip**
>
> For more troubleshooting, see the deployment checklist to make sure you covered every step in your scenario.

8. Set up a scan trigger. Your options are **Recurring** and **Once**.



9. On **Review new scan**, select **Save and run** to launch your scan.

View your scans and scan runs

To view existing scans:

1. Go to the Microsoft Purview portal. On the left pane, select **Data map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that has results you want to view. The pane shows you all the previous scan runs, along with the status and metrics for each scan run.
4. Select the run ID to check the scan run details.

Manage your scans

To edit, cancel, or delete a scan:

1. Go to the Microsoft Purview portal. On the left pane, select **Data Map**.

2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.

3. Select the scan that you want to manage. You can then:

    - Edit the scan by selecting **Edit scan**.
    - Cancel an in-progress scan by selecting **Cancel scan run**.
    - Delete your scan by selecting **Delete scan**.

> ⓘ Note
>
> - Deleting your scan does not delete catalog assets created from previous scans.

Glossary

Items and **sub-items** in Microsoft Fabric Lakehouse:

- **Items**: In Fabric, items are top-level workspace artifacts such as Lakehouse, Warehouse, Notebook, Pipeline, and Semantic Model. Each item acts as a single entity for lifecycle operations, including creation, deletion, and CI/CD tracking. A Lakehouse itself is an item. Tables (Delta and non-Delta), files, folders, and sometimes shortcuts within a Lakehouse are considered sub-items.
- **Sub-items**: These elements represent the actual data layer. You can view and query them through Spark and SQL endpoints. However, you can't manage or deploy them independently.

Next steps

Now that you registered your source, follow these guides to learn more about Microsoft Purview and your data.

- Data Estate Insights in Microsoft Purview
- Search Unified Catalog

ⓘ **Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 04/28/2026

Connect to your Microsoft Fabric tenant from Microsoft Purview in a different tenant

This article outlines how to register a Microsoft Fabric tenant that's in **a different tenant** from your Microsoft Purview resource, and how to authenticate and interact with the Fabric source in Microsoft Purview. For more information about Microsoft Purview in general, read the introductory article.

> **Note**
>
> Scanning a Microsoft Fabric tenant brings in metadata and lineage from Fabric items, including Power BI. Learn how to scan a Power BI tenant. The experience of registering a Fabric tenant and setting up a scan is similar to Power BI tenant and shared among all Fabric items.

Supported capabilities

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan	Classification	Labeling	Access Policy	Lineage	Data Sharing	Live view
Yes	Yes	Yes	Yes	No	No	No	Yes	No	No

Expand table

Experiences	Fabric items	Available in scan	Available in live view (same tenant only)
Real-Time Analytics	Kusto Query Language (KQL) Database	Yes	Yes
	KQL Queryset	Yes	Yes
Data Science	Experiment	Yes	Yes
	ML model	Yes	Yes
Data Factory	Data pipeline	Yes	Yes

Experiences	Fabric items	Available in scan	Available in live view (same tenant only)
	Dataflow Gen2	Yes	Yes
Data Engineering	Lakehouse	Yes	Yes
	Notebook	Yes	Yes
	Spark Job Definition	Yes	Yes
	SQL analytics endpoint	Yes	Yes
Data Warehouse	Warehouse	Yes	Yes
Power BI	Dashboard	Yes	Yes
	Dataflow	Yes	Yes
	Datamart	Yes	Yes
	Semantic model (Dataset)	Yes	Yes
	Report	Yes	Yes
	Paginated report	Yes	

Supported scenarios for Fabric scans

Expand table

Scenarios	Microsoft Purview public access allowed/denied	Fabric public access allowed /denied	Runtime option	Authentication option	Deployment checklist
Public access with Azure IR	Allowed	Allowed	Azure Runtime	Delegated authentication / Service principal	Review deployment checklist
Public access with Self-hosted IR	Allowed	Allowed	SHIR or Kubernetes SHIR	Service principal	Review deployment checklist
Private access	Denied	Allowed	Managed Virtual Network IR (v2 only)	Delegated authentication / Service principal	Review deployment checklist

> **Note**
>
> The supported scenarios above apply for non-Power BI items in Fabric. See the **Power BI documentation** for supported scenarios applicable for Power BI items.

Known limitations

- Currently for all Fabric items besides Power BI, only item level metadata and lineage will be scanned. For Lakehouse tables and files, sub-item level metadata scanning is available but sub-item level lineage is still not supported.
- For self-hosted integration runtime, the standard self-hosted integration runtime with minimal version 5.40.8836.1 or Kubernetes supported self-hosted integration runtime is supported.
- Empty workspaces are skipped.

> **Note**
>
> For Fabric Lakehouse tables with schema-enabled databases, the schema name and table name are combined into a single segment in the qualified name using `%252F` (double-URL-encoded forward slash) as a separator. For example, a table at the OneLake path `Tables/dbo/my_table` will appear in Data Map with the qualified name `.../tables/dbo%252Fmy_table`, while the entity display name will show only the table name (for example, `my_table`) without the schema prefix.

Prerequisites

Before you start, make sure you have the following:

- An Azure account with an active subscription. Create an account for free.
- An active Microsoft Purview account.

Authentication options

- Service Principal
- Delegated Authentication

Deployment checklist

The deployment checklist is a summary of all the steps you need to take to set up a cross-tenant Fabric source. You can use it during setup or for troubleshooting, to confirm you've followed all the necessary steps to connect.

Public access with Azure IR

Scan cross-tenant Fabric by using delegated authentication in a public network

1. Make sure the Fabric tenant ID is entered correctly during the registration. By default, the Fabric tenant ID that exists in the same Microsoft Entra instance as Microsoft Purview will be populated.

2. Make sure your Fabric metadata model is up to date by enabling metadata scanning.

3. From the Azure portal, validate if the Microsoft Purview account network is set to **public access**.

4. From the Fabric tenant admin portal, make sure the Fabric tenant is configured to allow a public network.

5. Check your instance of Azure Key Vault to make sure:
   a. There are no typos in the password or secret.
   b. Microsoft Purview managed identity has **get** and **list** access to secrets.

6. Review your credential to validate that the:
   a. Client ID matches the *Application (Client) ID* of the app registration.
   b. For **delegated auth**, username includes the user principal name, such as `johndoe@contoso.com`.

7. In the Fabric Microsoft Entra tenant, validate the following Fabric admin user settings:
   a. The user is assigned to the Fabric administrator role.
   b. If the user is recently created, sign in with the user at least once, to make sure that the password is reset successfully, and the user can successfully initiate the session.
   c. There are no multifactor authentication or conditional access policies enforced on the user.

8. In the Fabric Microsoft Entra tenant, validate the following app registration settings:

a. The app registration exists in your Microsoft Entra tenant where the Fabric tenant is located.
b. If service principal is used, under **API permissions**, the following **delegated permissions** are assigned with read for the following APIs:

- Microsoft Graph openid
- Microsoft Graph User.Read

c. If delegated authentication is used, under **API permissions**, the following **delegated permissions** and **grant admin consent for the tenant** is set up with read for the following APIs:

- Power BI Service Tenant.Read.All
- Microsoft Graph openid
- Microsoft Graph User.Read

d. Under **Authentication**:
   i. **Supported account types** > **Accounts in any organizational directory (Any Microsoft Entra directory - Multitenant)** is selected.
   ii. **Implicit grant and hybrid flows** > **ID tokens (used for implicit and hybrid flows)** is selected.
   iii. **Allow public client flows** is enabled.

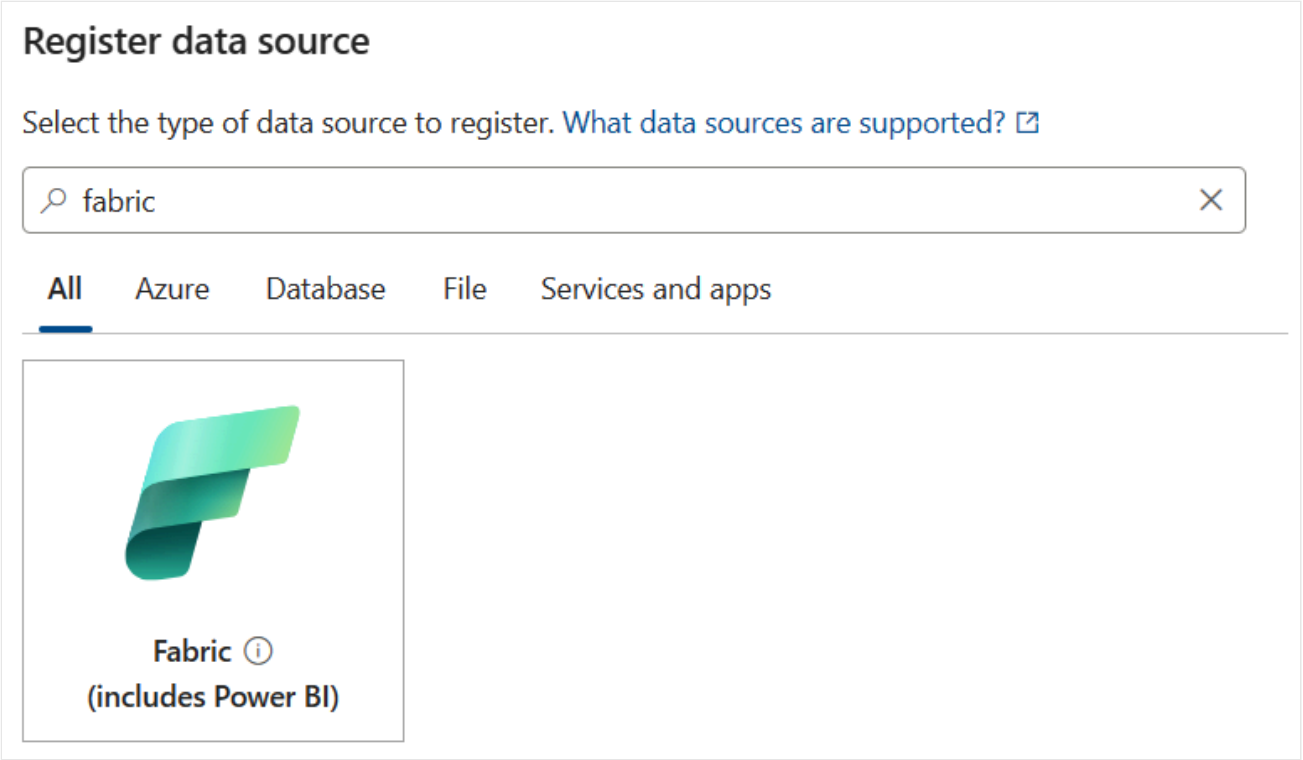
9. In Fabric tenant, from Microsoft Entra tenant, make sure Service Principal is member of the new security group.

10. On the Fabric Tenant Admin portal, validate if Allow service principals to use read-only admin APIs is enabled for the new security group.

Register the Fabric tenant

1. From the options on the left, select **Data Map**.

2. Select **Register**, and then select **Fabric** as your data source.



3. Give your Fabric instance a friendly name. The name must be 3 to 63 characters long, and must contain only letters, numbers, underscores, and hyphens. Spaces aren't allowed.

4. Edit the **Tenant ID** field, to replace with the tenant for the cross-tenant Fabric that you want to register and scan. By default, Microsoft Purview's tenant ID is populated.



Authentication to scan

To be able to scan your Microsoft Fabric tenant and review its metadata, you'll first need to create a way to authenticate with the Fabric tenant, and then give Microsoft Purview the authentication information.

Authenticate to Fabric tenant

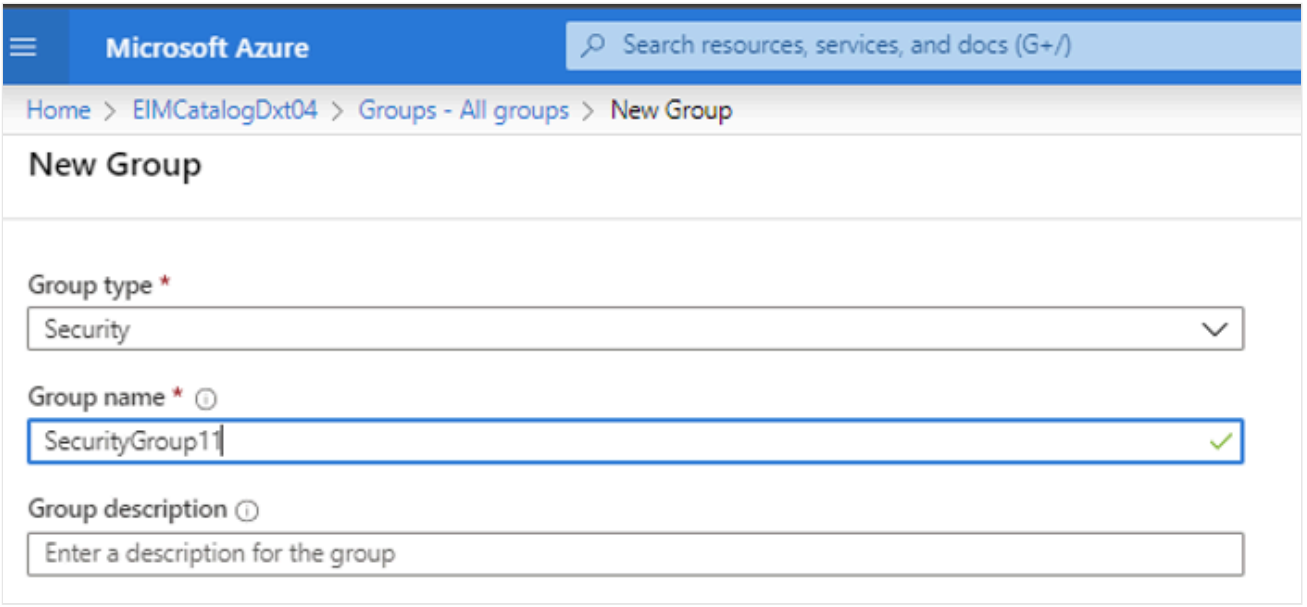
In Microsoft Entra tenant, where Fabric tenant is located:

1. In the Azure portal, search for **Microsoft Entra ID**.

2. Create a new security group in your Microsoft Entra ID, by following Create a basic group and add members using Microsoft Entra ID.

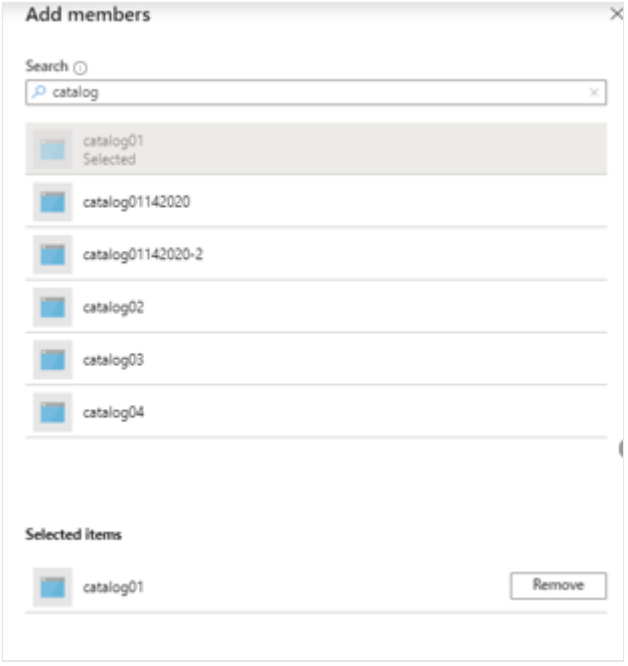
> **Tip**
>
> You can skip this step if you already have a security group you want to use.

3. Select **Security** as the **Group Type**.

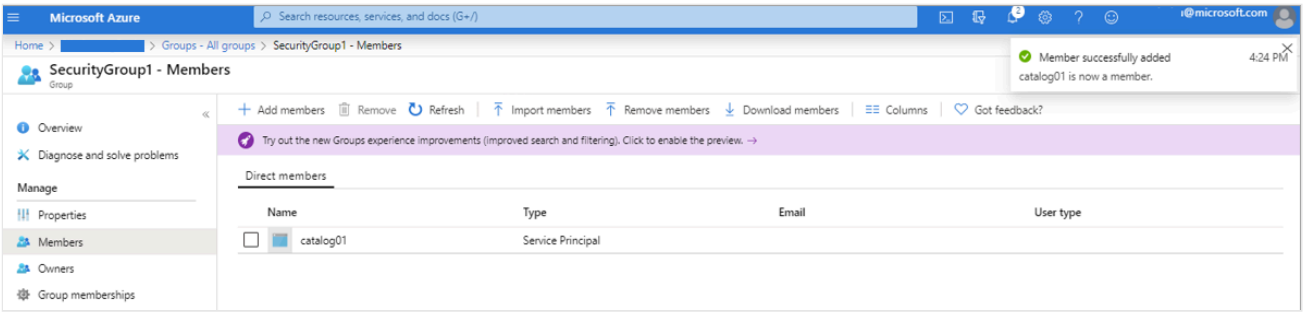


4. Select **Members**, then select **+ Add members**.

5. Search for your Microsoft Purview managed identity or service principal and select it.



You should see a success notification showing you that it was added.



Associate the security group with Fabric tenant

1. Log into the Fabric admin portal.

2. Select the **Tenant settings** page.

   > **Important**
   >
   > You need to be a Fabric Admin to see the tenant settings page.

3. Select **Admin API settings** > **Allow service principals to use read-only admin APIs**.

4. Select **Specific security groups**.

5. Select **Admin API settings** > **Enhance admin APIs responses with detailed metadata** and **Enhance admin APIs responses with DAX and mashup expressions** > Enable the toggle to allow Microsoft Purview Data Map automatically discover the detailed metadata of Fabric datasets as part of its scans.

   > **Important**
   >
   > After you update the Admin API settings on your Fabric tenant, wait around 15 minutes before registering a scan and test connection.

   > **Caution**
   >
   > When you allow the security group you created to use read-only admin APIs, you also allow it to access the metadata (such as dashboard and report names, owners, and descriptions) for all of your Fabric artifacts in this tenant. Once the metadata has been

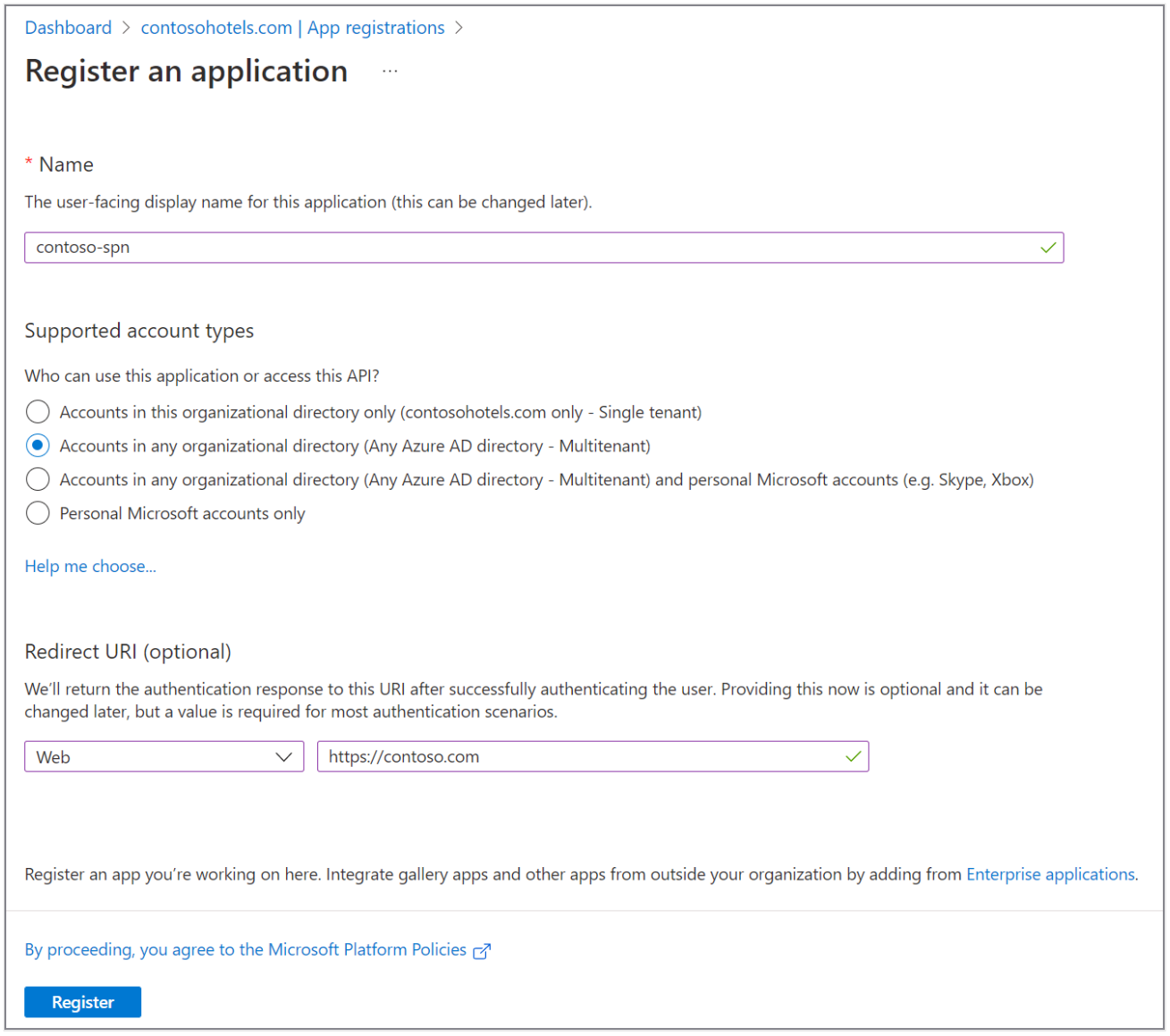
pulled into the Microsoft Purview, Microsoft Purview's permissions, not Fabric permissions, determine who can see that metadata.

> **Note**
>
> - You can remove the security group from your developer settings, but the metadata previously extracted won't be removed from the Microsoft Purview account. You can delete it separately if you wish.
> - If your Fabric item (such as lakehouse or mirrored catalog) has OneLake security enabled, then you must create a OneLake security role with **Read** permission assigned to the Microsoft Purview user. See Get started with OneLake Security for step-by-step instructions.

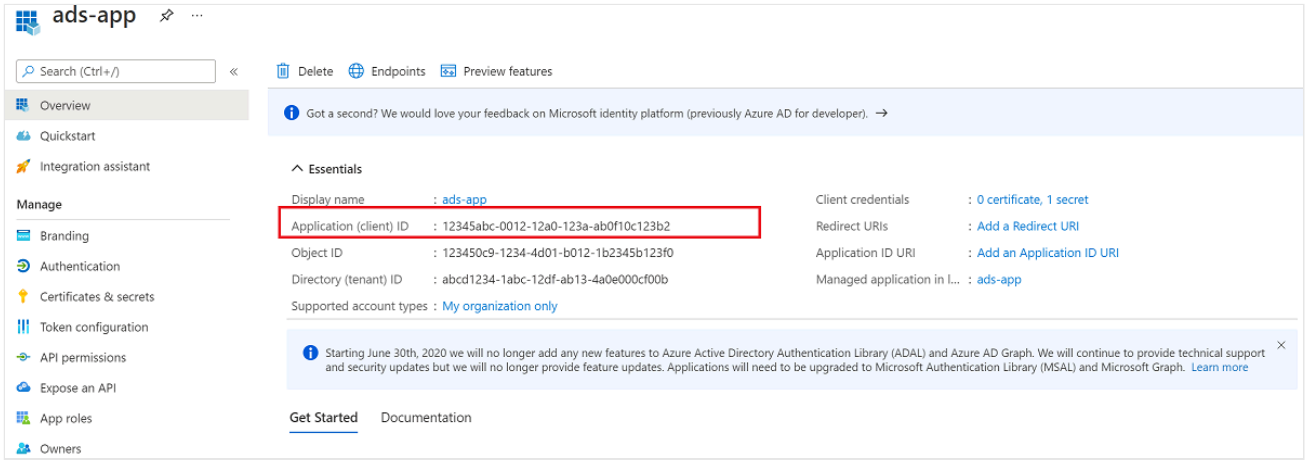
Configure credentials for scans in Microsoft Purview

Service principal

1. Create an app registration in your Microsoft Entra tenant where Fabric is located. Provide a web URL in the **Redirect URI**.

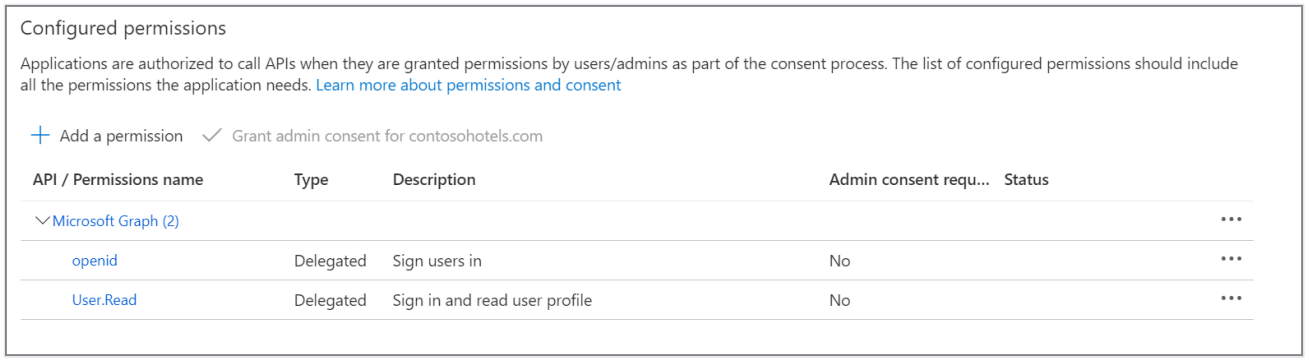


2. Take note of the client ID (app ID).

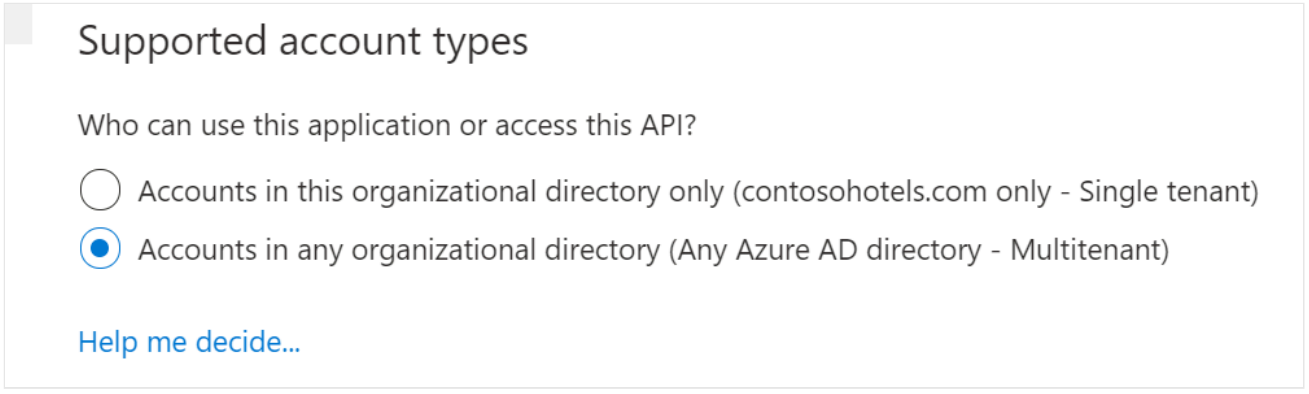


3. From the Microsoft Entra dashboard, select the newly created application, and then select **App permissions**. Assign the application the following delegated permissions:

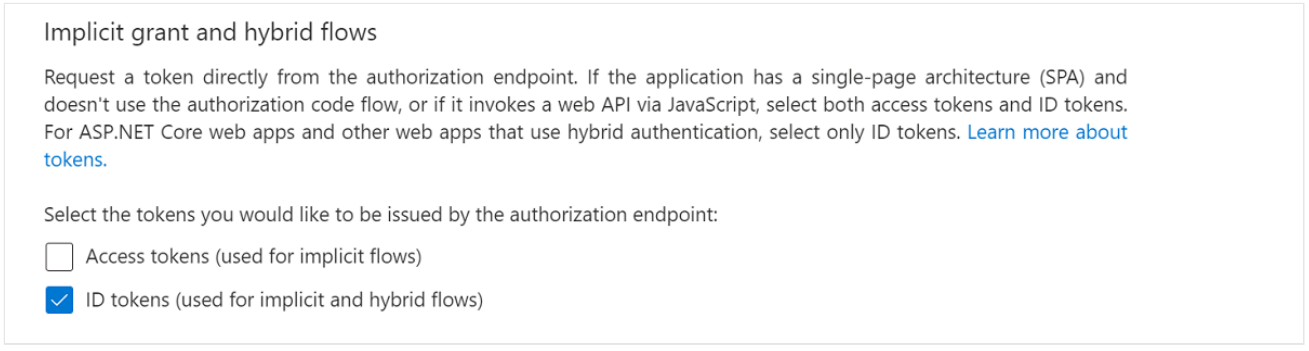
   - Microsoft Graph openid
   - Microsoft Graph User.Read



4. From the Microsoft Entra dashboard, select the newly created application, and then select **Authentication**. Under **Supported account types**, select **Accounts in any organizational directory (Any Microsoft Entra directory - Multitenant)**.



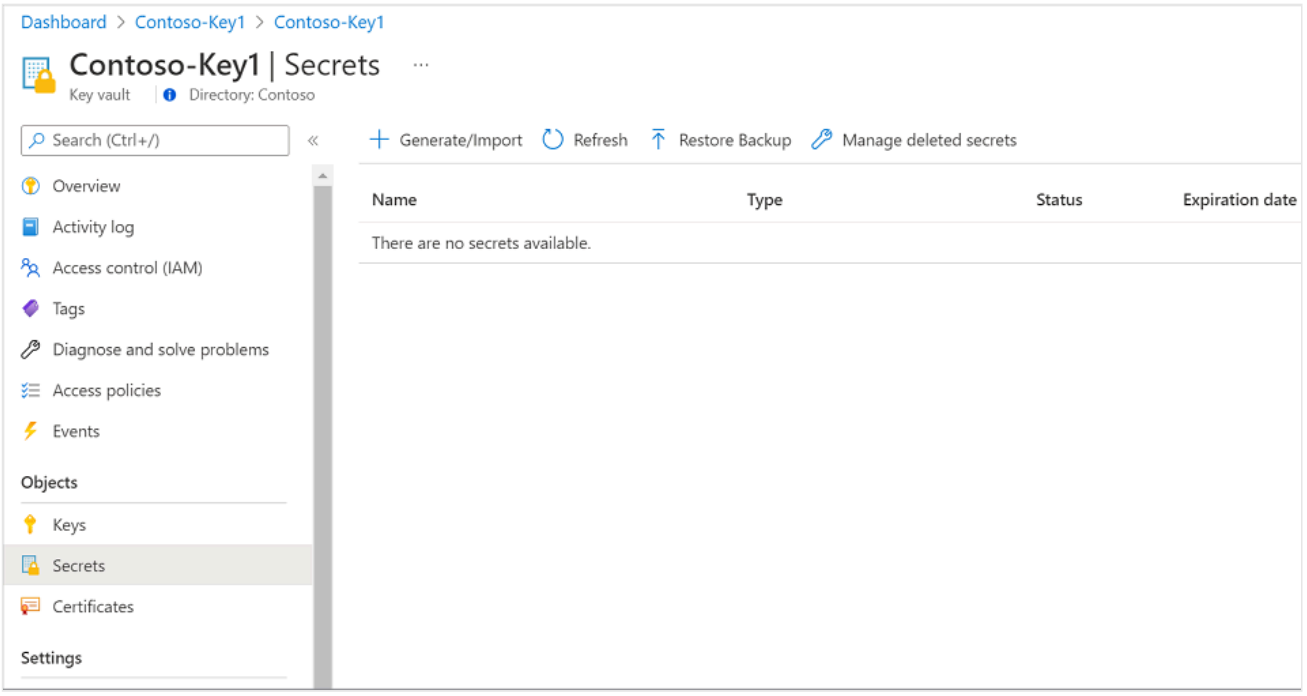
5. Under **Implicit grant and hybrid flows**, select **ID tokens (used for implicit and hybrid flows)**.



6. Under **Advanced settings**, enable **Allow Public client flows**.

7. In the tenant where Microsoft Purview is created go to the instance of Azure Key Vault.

8. Select **Settings** > **Secrets**, and then select **+ Generate/Import**.

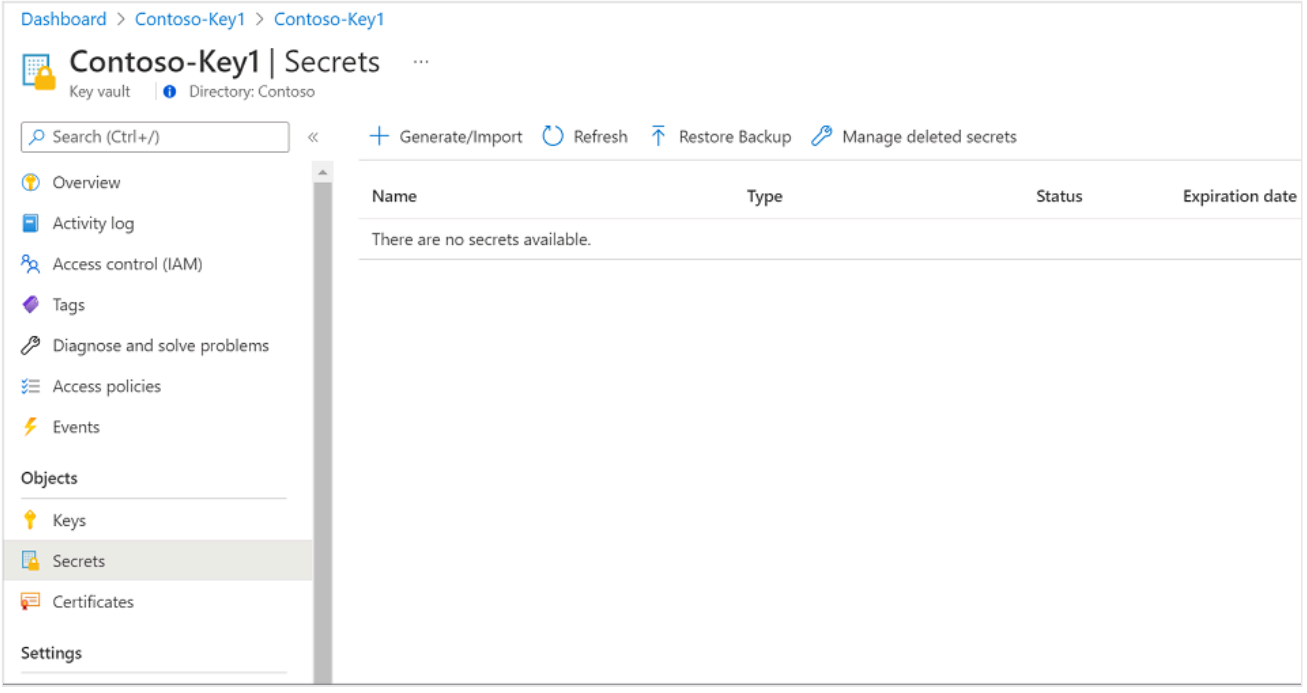


9. Enter a name for the secret. For **Value**, type the newly created secret for the App registration. Select **Create** to complete.

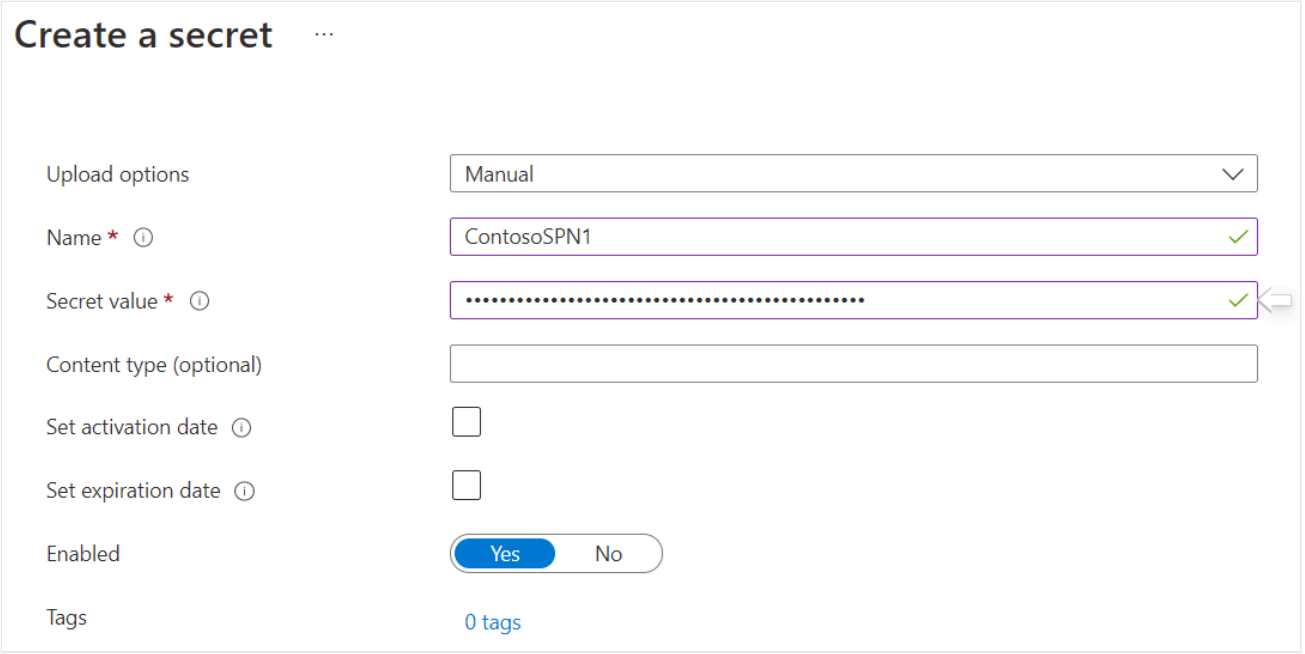
10. Under **Certificates & secrets**, create a new secret and save it securely for next steps.

11. In Azure portal, navigate to your Azure key vault.

12. Select **Settings** > **Secrets** and select **+ Generate/Import**.



13. Enter a name for the secret and for **Value**, type the newly created secret for the App registration. Select **Create** to complete.



14. If your key vault isn't connected to Microsoft Purview yet, you need to create a new key vault connection.

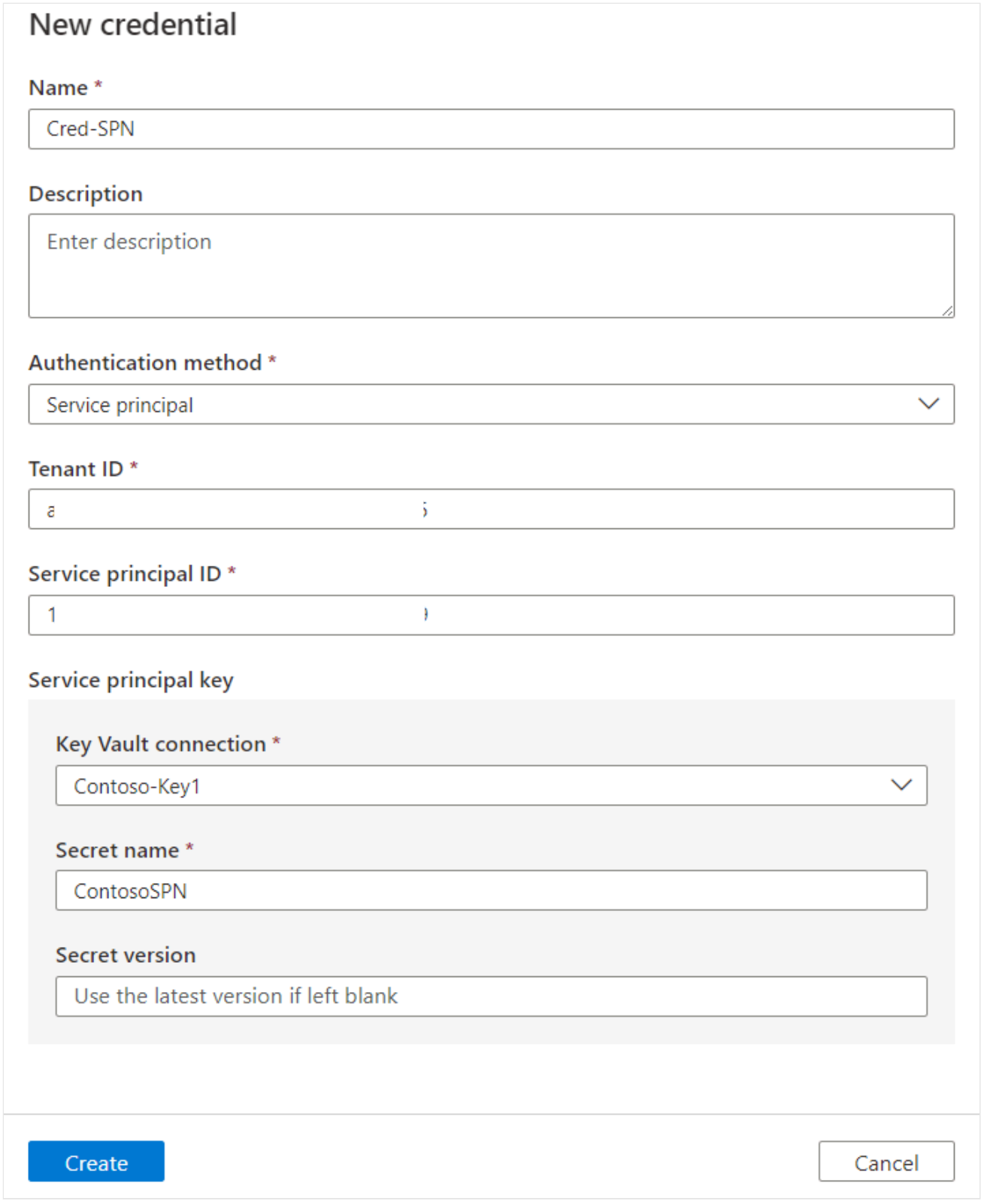
15. Next, in Microsoft Purview, go to the **Credentials** page under **Management** to create a new credential.

> **Tip**
>
> Alternatively, you can create a new credential during the scanning process.

16. Create your new Credential by selecting **+ New**.

17. Provide required parameters:

    - **Name**: Provide a unique name for credential
    - **Authentication method**: Service principal
    - **Tenant ID**: Your Fabric tenant ID
    - **Client ID**: Use Service Principal Client ID (App ID) you created earlier
    - **Key Vault Connection**: the Microsoft Purview connection to the Key Vault where you created your secret earlier.
    - **Secret name**: the name of the secret you created earlier.



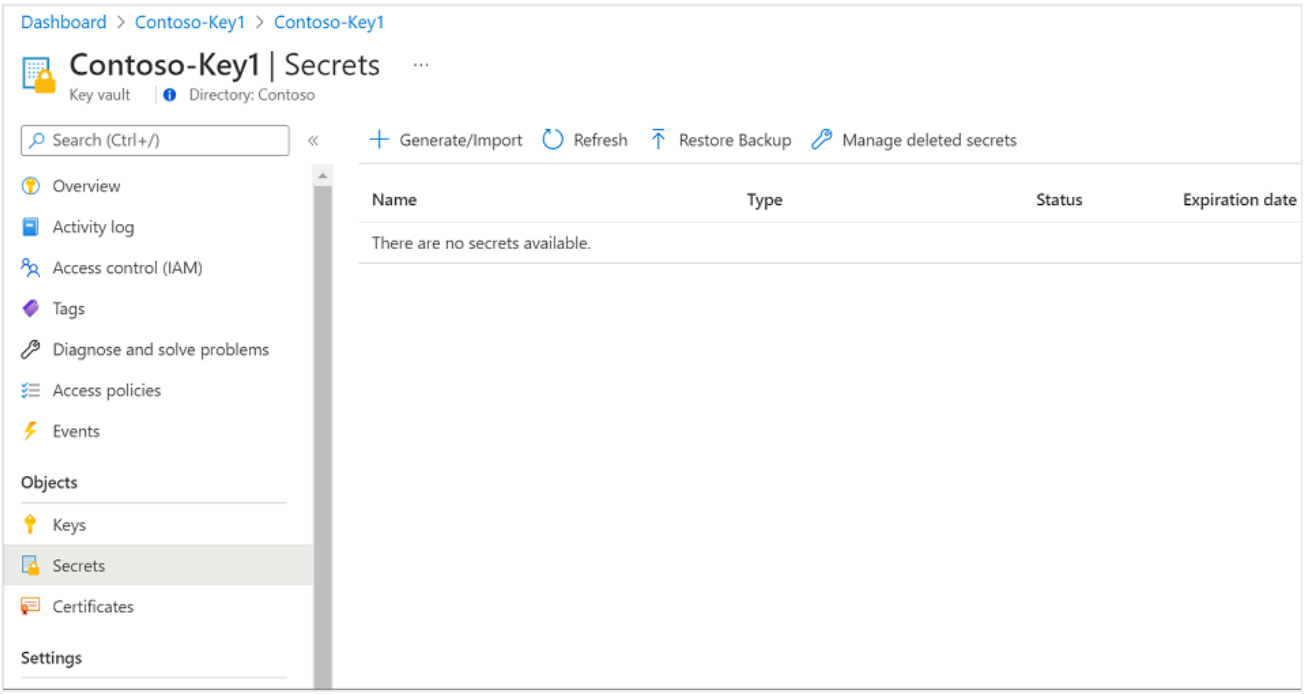
18. Once all the details have been filled in, select **Create**.

Delegated authentication

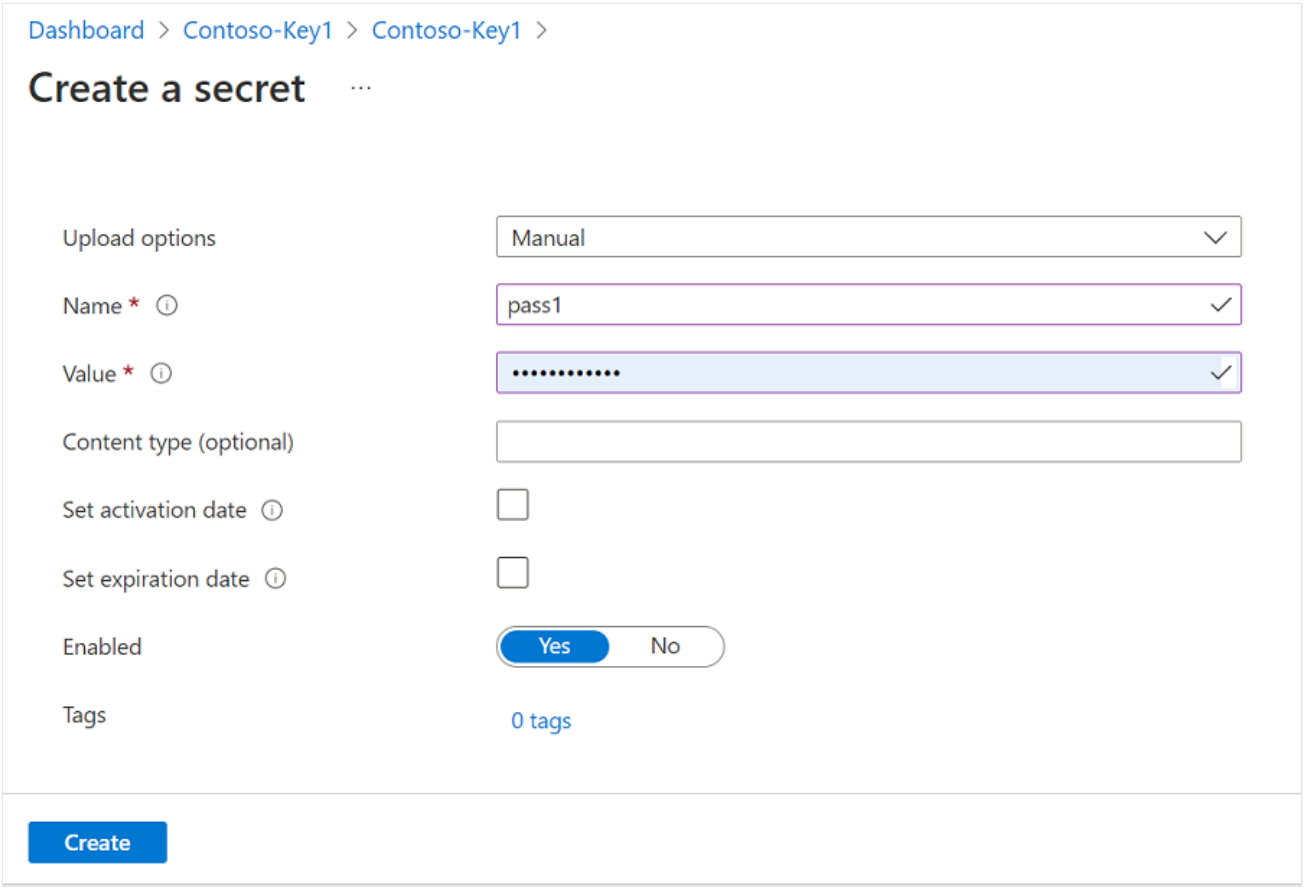
1. Create a user account in the Microsoft Entra tenant where the Fabric tenant is located, and assign the user to this role: **Fabric Administrator**. Take note of the username and sign in to change the password.

2. Go to the instance of Azure Key Vault in the tenant where Microsoft Purview is created.

3. Select **Settings** > **Secrets**, and then select **+ Generate/Import**.

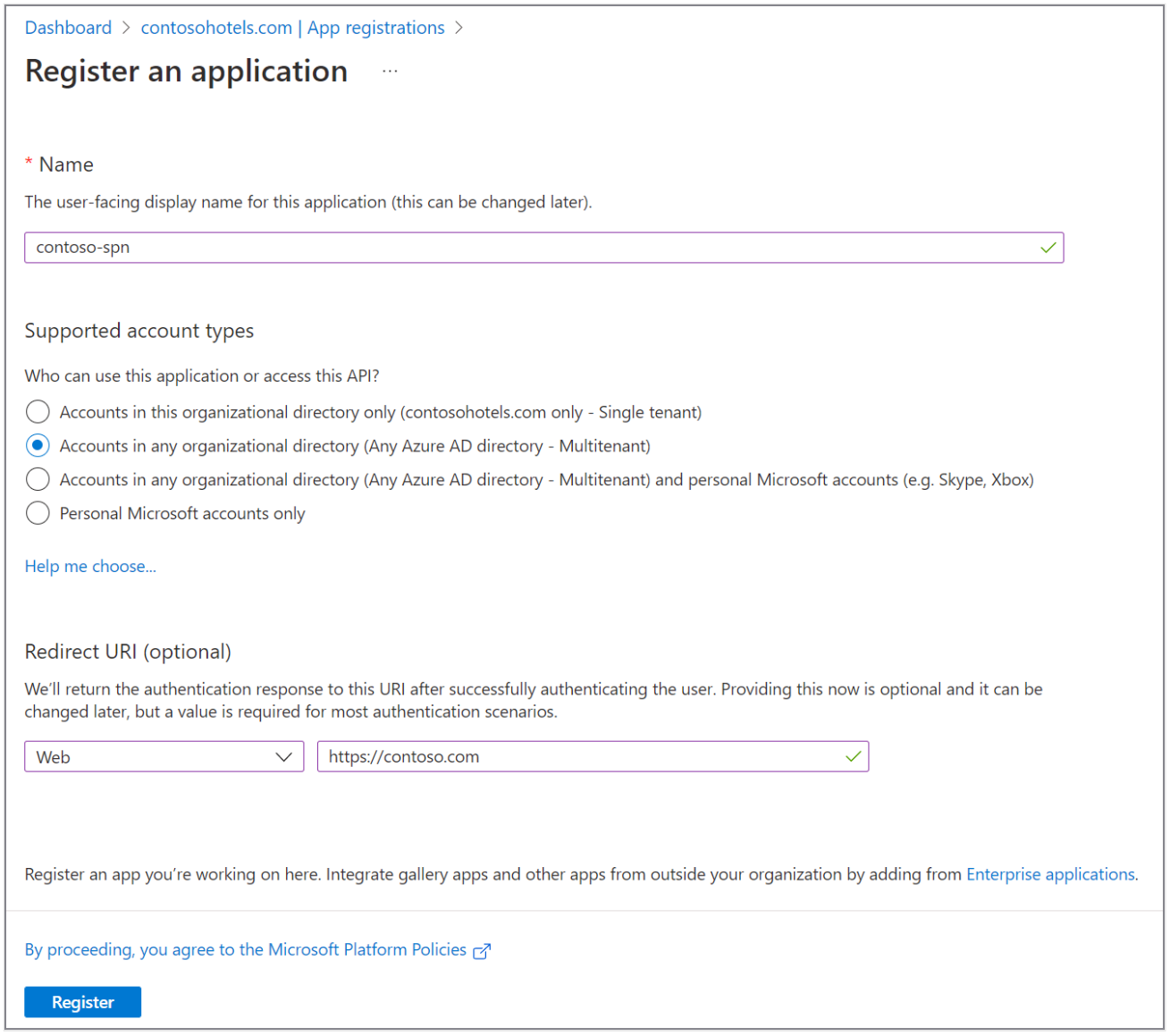


4. Enter a name for the secret. For **Value**, type the newly created password for the Microsoft Entra user. Select **Create** to complete.

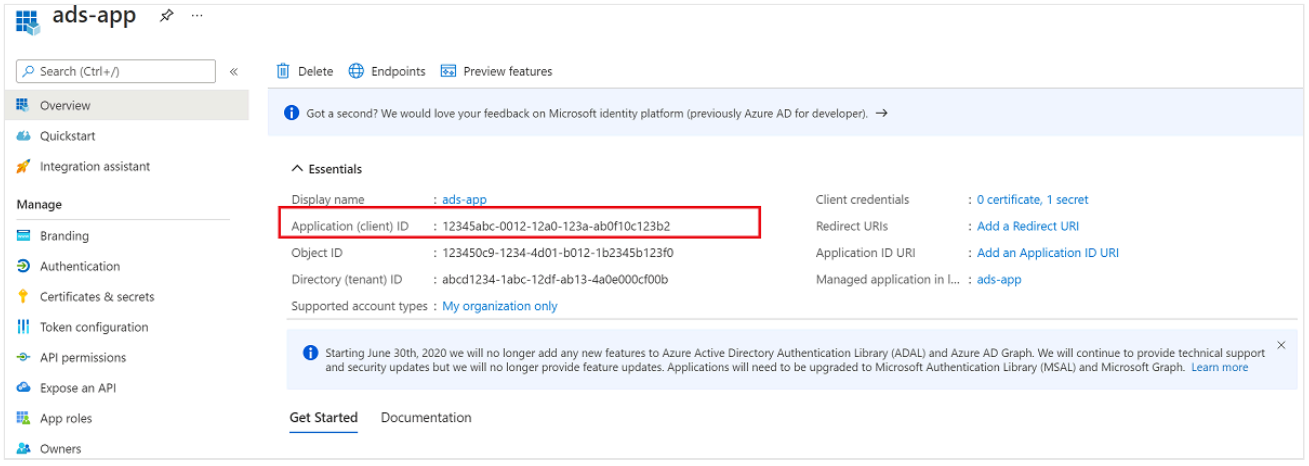


5. If your key vault isn't connected to Microsoft Purview yet, you need to create a new key vault connection.

6. Create an app registration in your Microsoft Entra tenant where Fabric is located. Provide a web URL in the **Redirect URI**.



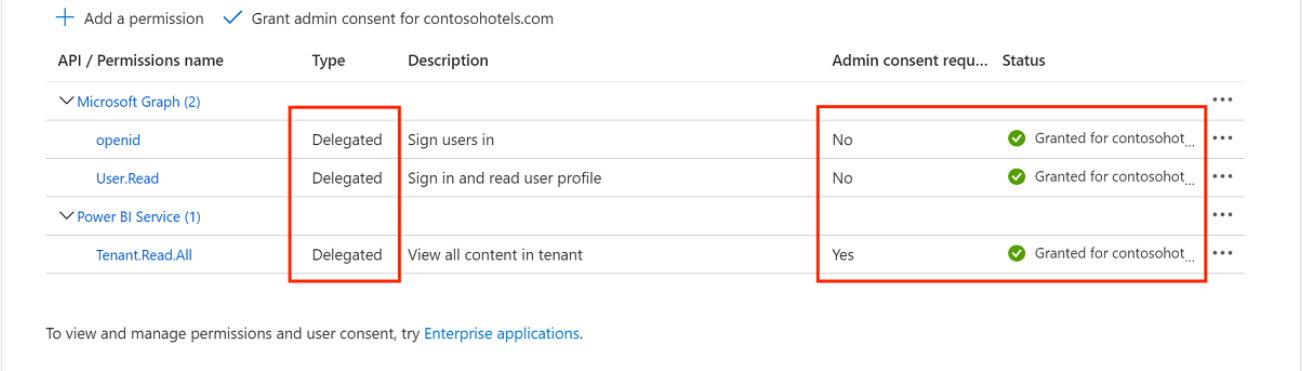
7. Take note of the client ID (app ID).



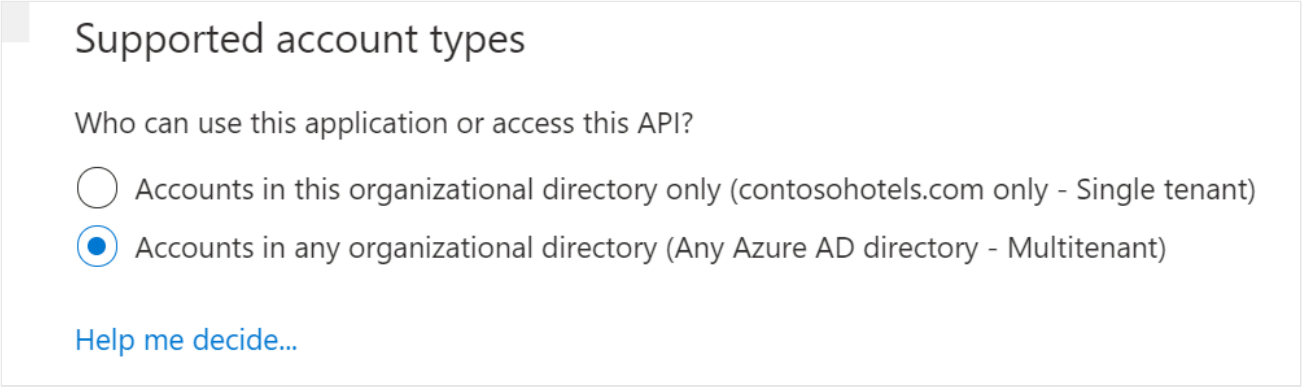
8. From the Microsoft Entra dashboard, select the newly created application, and then select **App permissions**. Assign the application the following delegated permissions, and grant

admin consent for the tenant:

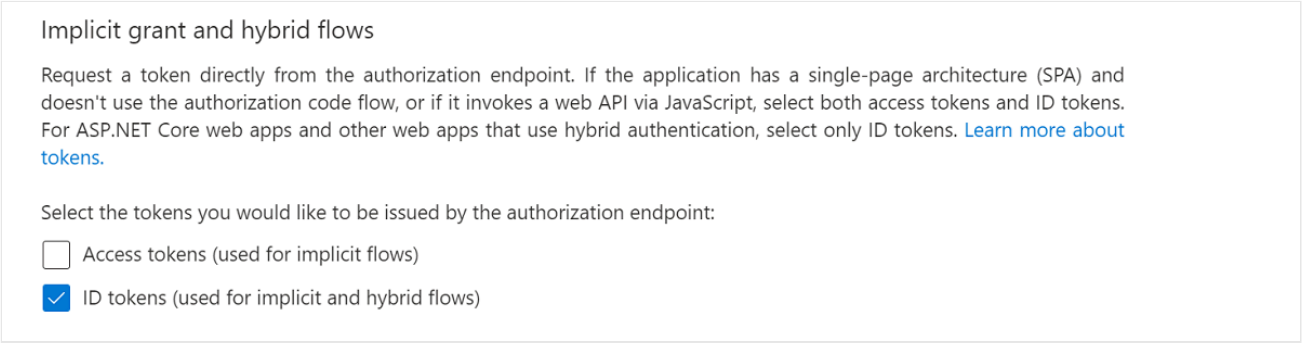
- Fabric Service Tenant.Read.All
- Microsoft Graph openid
- Microsoft Graph User.Read



9. From the Microsoft Entra dashboard, select the newly created application, and then select **Authentication**. Under **Supported account types**, select **Accounts in any organizational directory (Any Microsoft Entra directory - Multitenant)**.



10. Under **Implicit grant and hybrid flows**, select **ID tokens (used for implicit and hybrid flows)**.



11. Under **Advanced settings**, enable **Allow Public client flows**.

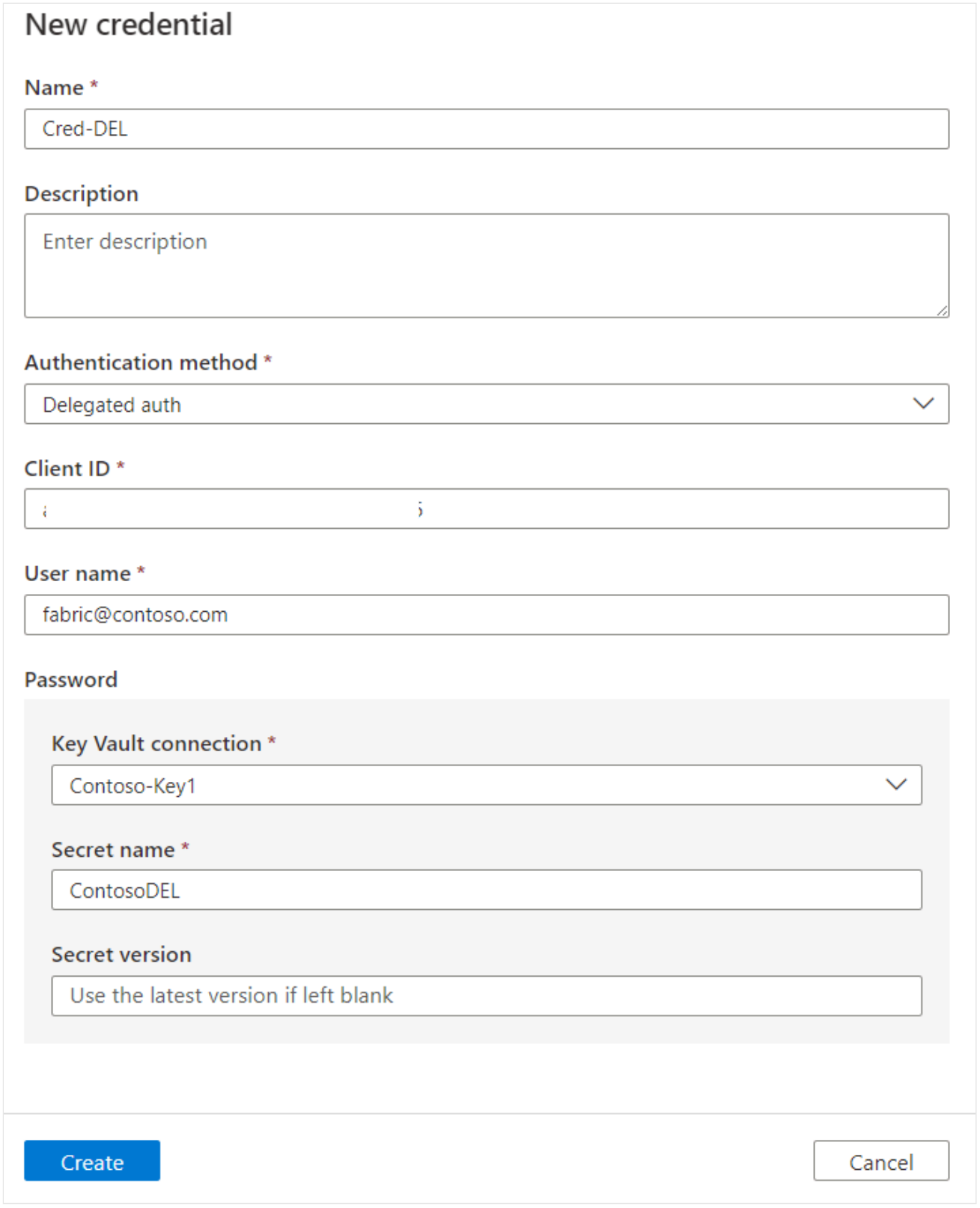
12. Next, in Microsoft Purview, go to the **Credentials** page under **Management** to create a new credential.

> **Tip**
>
> Alternatively, you can create a new credential during the scanning process.

13. Create your new Credential by selecting **+ New**.

14. Provide required parameters:

    - **Name**: Provide a unique name for credential
    - **Authentication method**: Delegated auth
    - **Client ID**: Use Service Principal Client ID (App ID) you created earlier
    - **User name**: Provide the username of Fabric Administrator you created earlier
    - **Key Vault Connection**: the Microsoft Purview connection to the Key Vault where you created your secret earlier.
    - **Secret name**: the name of the secret you created earlier.



15. Once all the details have been filled in, select **Create**.

Create scan

1. In the Microsoft Purview Studio, go to the **Data map** in the left menu. Go to **Sources**.

2. Select the registered Fabric source from cross-tenant.

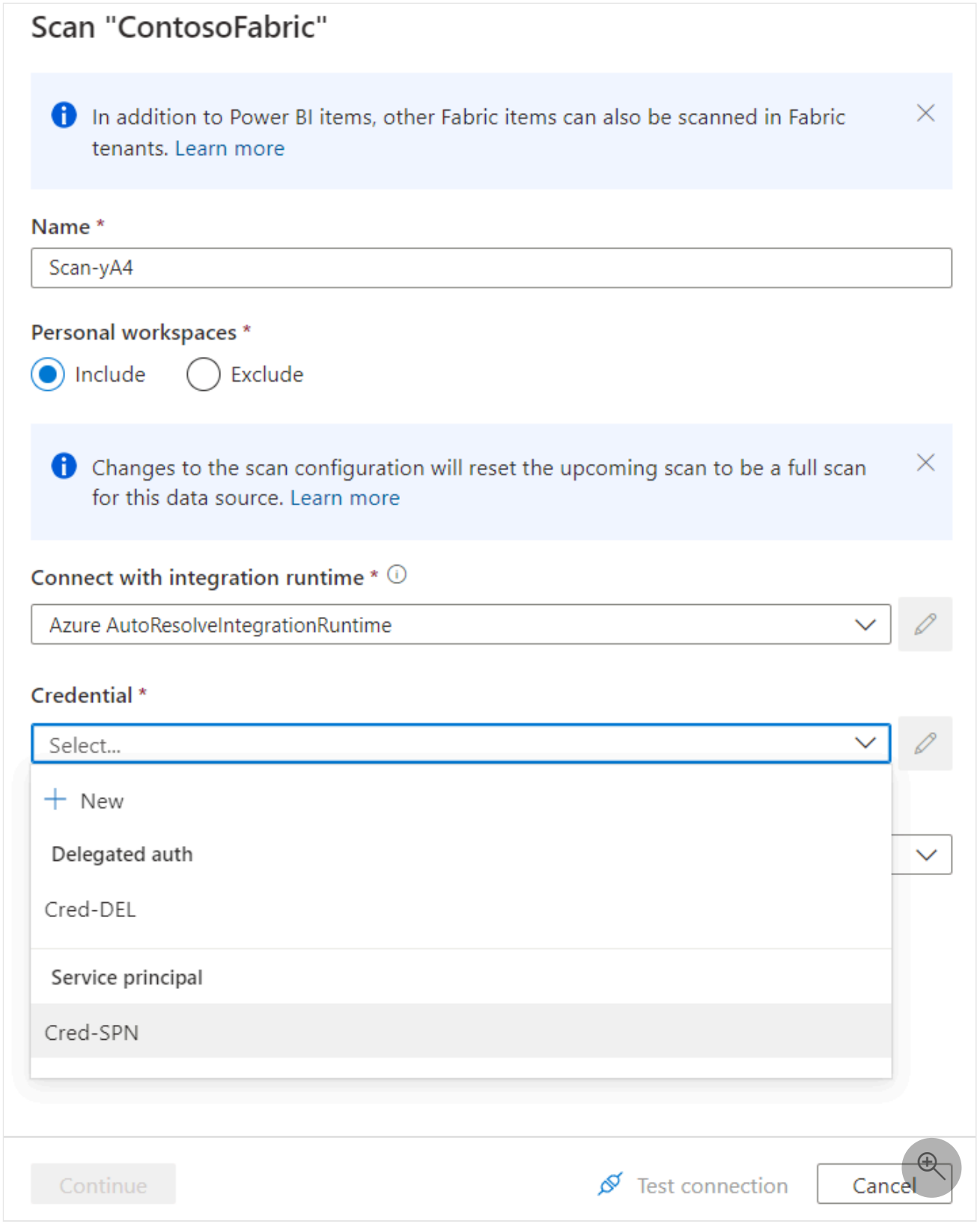
3. Select + **New scan**.

4. Give your scan a name. Then select the option to include or exclude the personal workspaces.

   > **Note**
   >
   > If you switch the configuration of a scan to include or exclude a personal workspace, you trigger a full scan of the Fabric source.

5. Select **Azure AutoResolveIntegrationRuntime** from the dropdown list.

6. For the **Credential**, select the credential you created for your service principal or delegated authentication.



7. Select **Test connection** before continuing to the next steps.

   If the test fails, select **View Report** to see the detailed status and troubleshoot the problem:

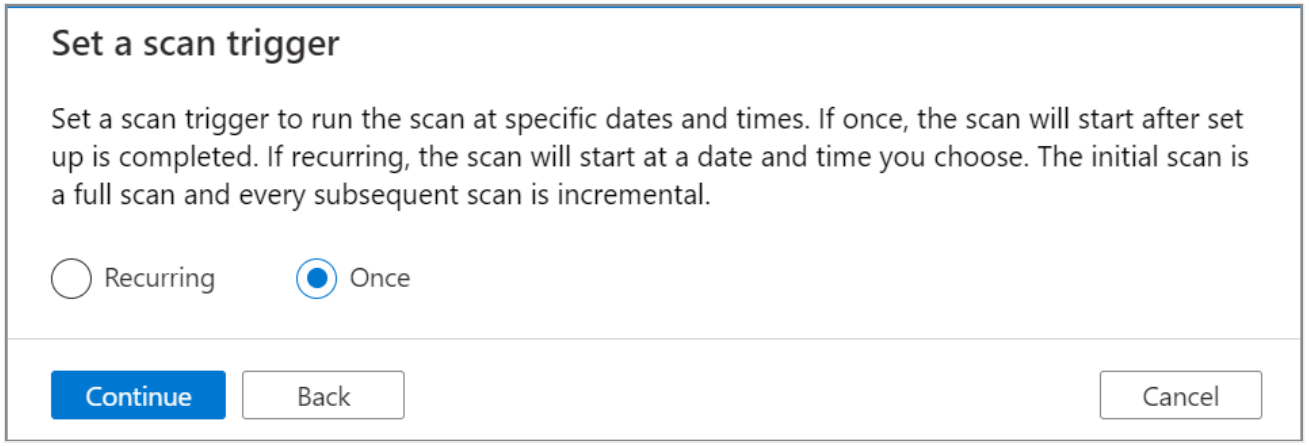
   a. *Access - Failed* status means that the user authentication failed. Validate if the username and password are correct. Review if the credential contains the correct client (app) ID from the app registration.

   b. *Assets (+ lineage) - Failed* status means that the authorization between Microsoft Purview and Fabric has failed. Make sure that the user is added to the Fabric administrator role.

c. *Detailed metadata (Enhanced) - Failed* status means that the Fabric admin portal is disabled for the following setting: **Enhance admin APIs responses with detailed metadata**.

> **Tip**
>
> For more troubleshooting, see the deployment checklist to make sure you've covered every step in your scenario.

8. Set up a scan trigger. Your options are **Recurring** or **Once**.



9. On **Review new scan**, select **Save and run** to launch your scan.

View your scans and scan runs

To view existing scans:

1. Go to the Microsoft Purview portal. On the left pane, select **Data map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that has results you want to view. The pane shows you all the previous scan runs, along with the status and metrics for each scan run.
4. Select the run ID to check the scan run details.

Manage your scans

To edit, cancel, or delete a scan:

1. Go to the Microsoft Purview portal. On the left pane, select **Data Map**.

2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.

3. Select the scan that you want to manage. You can then:

    - Edit the scan by selecting **Edit scan**.
    - Cancel an in-progress scan by selecting **Cancel scan run**.
    - Delete your scan by selecting **Delete scan**.

> ⓘ Note
>
> - Deleting your scan does not delete catalog assets created from previous scans.

Next steps

Now that you registered your source, see these guides to learn more about Microsoft Purview and your data.

- Data estate insights in Microsoft Purview
- Search Unified Catalog

ⓘ **Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 04/28/2026

Connect to and manage Google BigQuery projects in Microsoft Purview

Article • 12/12/2024

This article outlines how to register Google BigQuery projects, and how to authenticate and interact with Google BigQuery in Microsoft Purview. For more information about Microsoft Purview, read the introductory article.

Supported capabilities

Scanning capabilities

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan
Yes	Yes	No	Yes

When scanning Google BigQuery source, Microsoft Purview supports:

- Extracting technical metadata including:
  - Projects
  - Datasets
  - Tables including the columns
  - Views including the columns
- Fetching static lineage on assets relationships among tables and views.

When setting up scan, you can choose to scan an entire Google BigQuery project, or scope the scan to a subset of datasets matching the given name(s) or name pattern(s).

Other capabilities

View a table listing the supported capabilities for classifications, sensitivity labels, policies, data lineage, and live view.

Known limitations

- Currently, Microsoft Purview only supports scanning Google BigQuery datasets in US multi-regional location. If the specified dataset is in other location e.g. us-east1

or EU, you will observe scan completes but no assets shown up in Microsoft Purview.
- When object is deleted from the data source, currently the subsequent scan won't automatically remove the corresponding asset in Microsoft Purview.

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- An active Microsoft Purview account.
- You need Data Source Administrator and Data Reader permissions to register a source and manage it in the Microsoft Purview governance portal. For more information about permissions, see Access control in Microsoft Purview.
- Set up the right integration runtime for your scenario:
  - **To use a self-hosted integration runtime:**
    - Follow the article to create and configure a self-hosted integration runtime.
    - Ensure JDK 11 is installed on the machine where the self-hosted integration runtime is installed. Restart the machine after you newly install the JDK for it to take effect.
    - Ensure that Visual C++ Redistributable (version Visual Studio 2012 Update 4 or newer) is installed on the machine where the self-hosted integration runtime is running. If you don't have this update installed, download it now.
    - Download and unzip the BigQuery JDBC driver on the machine where your self-hosted integration runtime is running. Note down the folder path which you will use to set up the scan.
  - **To use a kubernetes supported self-hosted integration runtime:**
    - Follow the article to create and configure a kubernetes supported integration runtime.
    - Download and unzip the BigQuery JDBC driver on the machine where your self-hosted integration runtime is running. Note down the folder path which you will use to set up the scan.

> **Note**
>
> The driver should be accessible by the self-hosted integration runtime. By default, self-hosted integration runtime uses local service account "NT SERVICE\DIAHostService". Make sure it has "Read and execute" and "List folder contents" permission to the driver folder.

Required permissions for scan

The Google BigQuery service account you use for scan needs to have both BigQuery Metadata Viewer and BigQuery Job User IAM roles on the project(s) you want to scan. These permissions are required because Microsoft Purview extracts the metadata by reading the Google BigQuery database system tables (e.g. INFORMATION_SCHEMA). And the underlying Google BigQuery JDBC driver creates a BigQuery job when it needs to read from these system tables.

Microsoft Purview uses Oauth 2.0 protocol for accessing Google BigQuery service. Set up the credential by following the instruction in Create and run scan section.

Register

This section describes how to register a Google BigQuery project in Microsoft Purview using the Microsoft Purview governance portal.

Steps to register

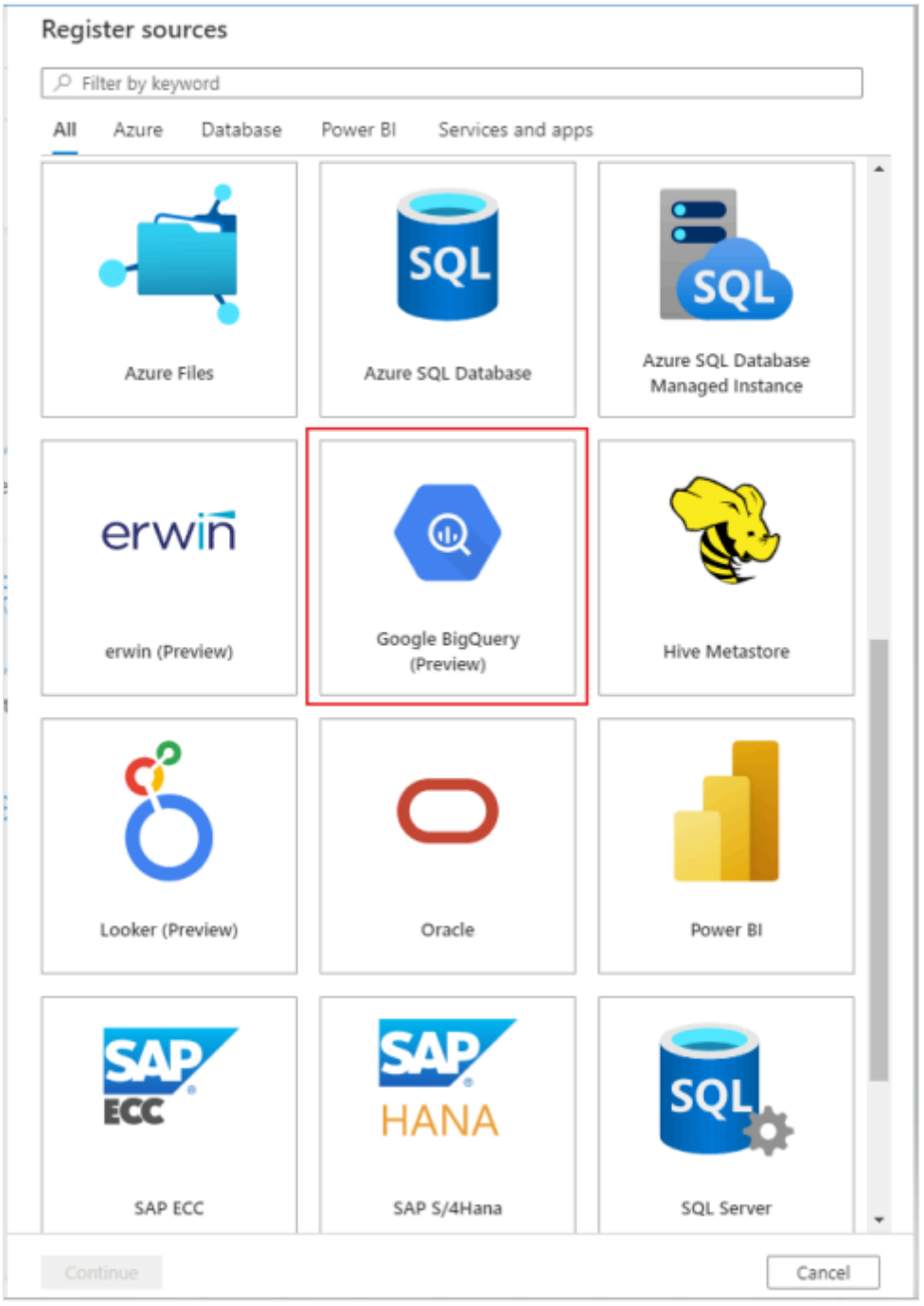
1. Open the Microsoft Purview governance portal by:

    - Browsing directly to https://web.purview.azure.com and selecting your Microsoft Purview account.
    - Opening the Azure portal, searching for and selecting the Microsoft Purview account. Selecting the **the Microsoft Purview governance portal** button.

2. Select **Data Map** on the left navigation.

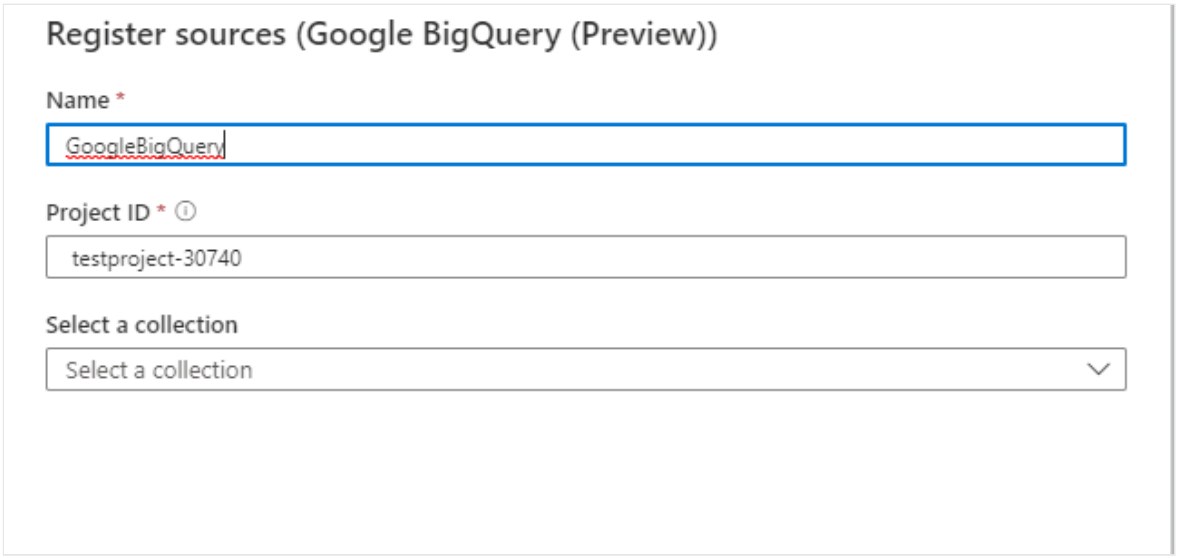
3. Select **Register.**

4. On Register sources, select **Google BigQuery** . Select **Continue.**



On the Register sources (Google BigQuery) screen, do the following:

1. Enter a **Name** that the data source will be listed within the Catalog.
2. Enter the **ProjectID.** This should be a fully qualified project ID. For example, mydomain.com:myProject
3. Select a collection from the list.
4. Select **Register**.



Scan

Follow the steps below to scan a Google BigQuery project to automatically identify assets. For more information about scanning in general, see our introduction to scans and ingestion.

Create and run scan

1. In the Management Center, select Integration runtimes. Make sure a self-hosted integration runtime is set up. If it isn't set up, use the steps mentioned in prerequisites.

2. Navigate to **Sources**.

3. Select the registered **BigQuery** project.

4. Select **+ New scan**.

5. Provide the below details:

    a. **Name**: The name of the scan

    b. **Connect via integration runtime**: Select the configured self-hosted integration runtime

    c. **Credential**: While configuring BigQuery credential, make sure to:

    - Select **Basic Authentication** as the Authentication method
    - Provide the email ID of the service account in the User name field. For example, `xyz\@developer.gserviceaccount.com`

- Follow below steps to generate the private key, copy the entire JSON key file then store it as the value of a Key Vault secret.

To create a new private key from Google's cloud platform:

i. In the navigation menu, select IAM & Admin -> Service Accounts -> Select a project ->

ii. Select the email address of the service account that you want to create a key for.

iii. Select the **Keys** tab.

iv. Select the **Add key** drop-down menu, then select Create new key.

v. Choose JSON format.

> ⓘ **Note**
>
> The contents of the private key are saved in a temp file on the VM when scanning processes are running. This temp file is deleted after the scans are successfully completed. In the event of a scan failure, the system will continue to retry until success. Please make sure access is appropriately restricted on the VM where SHIR is running.

To understand more on credentials, refer to the link here.

d. **Driver location**: Specify the path to the JDBC driver location in your machine where self-host integration runtime is running. For example: `D:\Drivers\GoogleBigQuery`.

i. For self-hosted integration runtime on a local machine: `D:\Drivers\GoogleBigQuery`. It's the path to valid JAR folder location. The value must be a valid absolute file path and doesn't contain space. Make sure the driver is accessible by the self-hosted integration runtime;; learn more from prerequisites section.

ii. For Kubernetes-supported self-hosted integration runtime: `./drivers/GoogleBigQuery`. It's the path to valid JAR folder location. The value must be a valid relative file path. Please refer to the documentation to set up a scan with external drivers for uploading drivers in advance.

e. **Dataset**: Specify a list of BigQuery datasets to import. For example, `dataset1;dataset2`. When the list is empty, all available datasets are imported. Acceptable dataset name patterns can be static names or contain wildcard %.

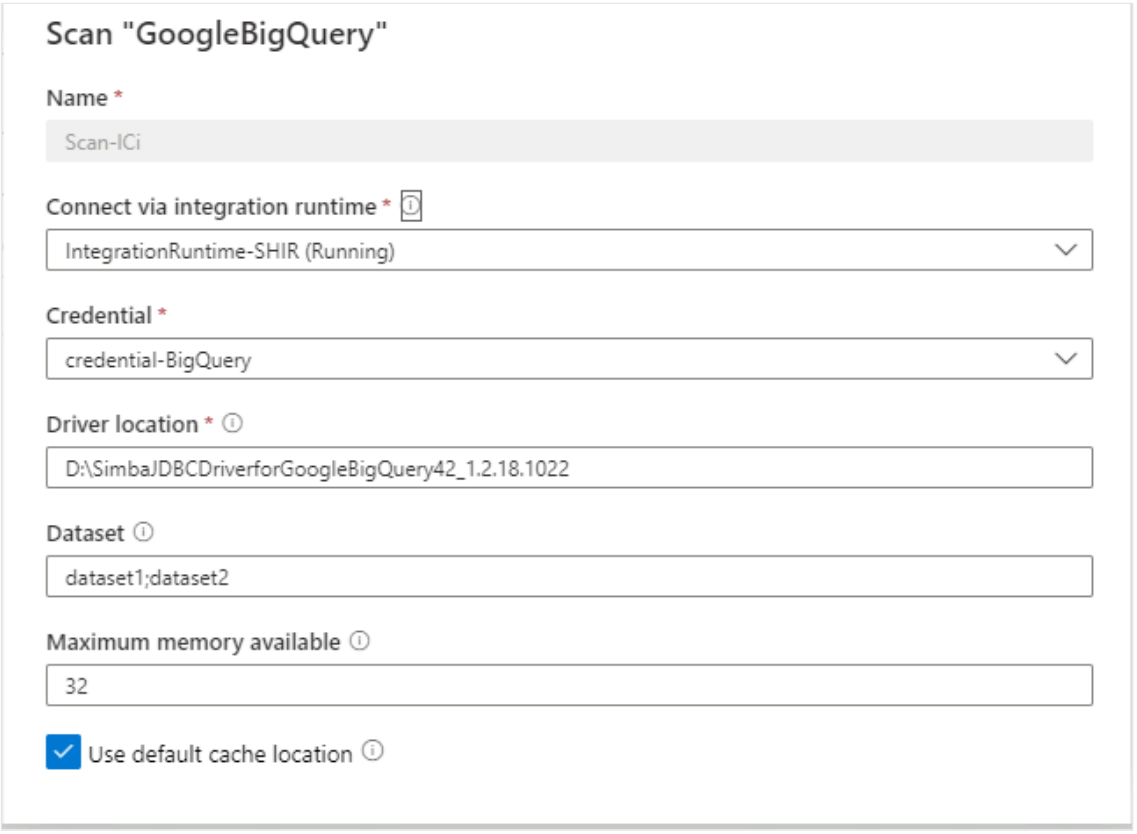
Example: `A%;%B;%C%;D`

- Start with A or

- end with B or
- contain C or
- equal D

Usage of NOT and special characters aren't acceptable.

f. **Maximum memory available**: Maximum memory (in GB) available on your VM to be used by scanning processes. This is dependent on the size of Google BigQuery project to be scanned.



6. Select **Test connection.**

7. Select **Continue**.

8. Choose your **scan trigger**. You can set up a schedule or ran the scan once.

9. Review your scan and select **Save and Run**.

View your scans and scan runs

To view existing scans:

1. Go to the Microsoft Purview portal. On the left pane, select **Data map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.

3. Select the scan that has results you want to view. The pane shows you all the previous scan runs, along with the status and metrics for each scan run.
4. Select the run ID to check the scan run details.

Manage your scans

To edit, cancel, or delete a scan:

1. Go to the Microsoft Purview portal. On the left pane, select **Data Map**.

2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.

3. Select the scan that you want to manage. You can then:

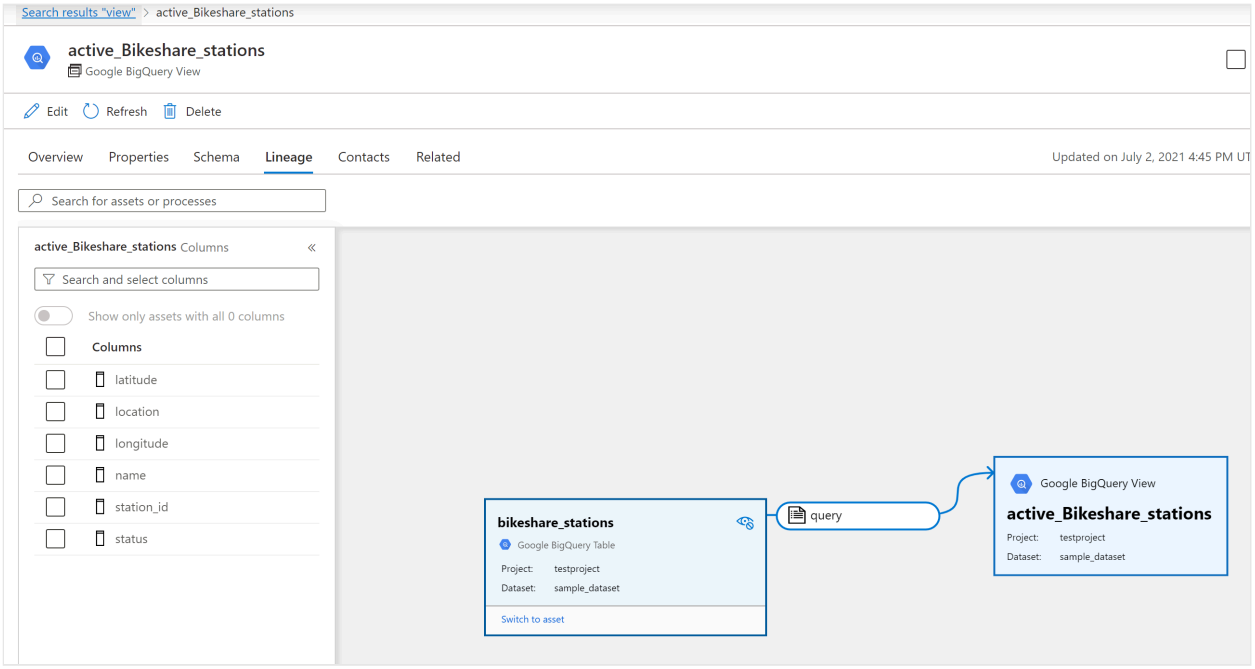
    - Edit the scan by selecting **Edit scan**.
    - Cancel an in-progress scan by selecting **Cancel scan run**.
    - Delete your scan by selecting **Delete scan**.

> ⓘ **Note**
>
> - Deleting your scan does not delete catalog assets created from previous scans.

Lineage

After scanning your Google BigQuery source, you can browse Unified Catalog or search Unified Catalog to view the asset details.

Go to the asset -> lineage tab, you can see the asset relationship when applicable. Refer to the supported capabilities section on the supported Google BigQuery lineage scenarios. For more information about lineage in general, see data lineage and lineage user guide.



Next steps

Now that you've registered your source, follow the below guides to learn more about Microsoft Purview and your data.

- Data Estate Insights in Microsoft Purview
- Lineage in Microsoft Purview
- Search Unified Catalog

Feedback

Was this page helpful? Yes No

Connect to and manage HDFS in Microsoft Purview

Article • 06/27/2024

This article outlines how to register Hadoop Distributed File System (HDFS), and how to authenticate and interact with HDFS in Microsoft Purview. For more information about Microsoft Purview, read the introductory article.

Supported capabilities

Scanning capabilities

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan
Yes	Yes	Yes	Yes

When scanning HDFS source, Microsoft Purview supports extracting technical metadata including HDFS:

- Namenode
- Folders
- Files
- Resource sets

When setting up scan, you can choose to scan the entire HDFS or selective folders. Learn about the supported file format here.

The connector uses *webhdfs* protocol to connect to HDFS and retrieve metadata. MapR Hadoop distribution is not supported.

Other capabilities

View a table listing the supported capabilities for classifications, sensitivity labels, policies, data lineage, and live view.

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- An active Microsoft Purview account.
- You need Data Source Administrator and Data Reader permissions to register a source and manage it in the Microsoft Purview governance portal. For more information about permissions, see Access control in Microsoft Purview.
- Set up the right integration runtime for your scenario:
  - **To use a self-hosted integration runtime:**
    - Follow the article to create and configure a self-hosted integration runtime.
    - Ensure JDK 11 is installed on the machine where the self-hosted integration runtime is installed. Restart the machine after you newly install the JDK for it to take effect.
    - Ensure that Visual C++ Redistributable (version Visual Studio 2012 Update 4 or newer) is installed on the machine where the self-hosted integration runtime is running. If you don't have this update installed, download it now.
    - To set up your environment to enable Kerberos authentication, see the Use Kerberos authentication for the HDFS connector section.
  - **To use a kubernetes supported self-hosted integration runtime:**
    - Follow the article to create and configure a kubernetes supported integration runtime.
    - To set up your environment to enable Kerberos authentication, see the Use Kerberos authentication for the HDFS connector section.

Register

This section describes how to register HDFS in Microsoft Purview using the Microsoft Purview governance portal.

Steps to register

To register a new HDFS source in Microsoft Purview Unified Catalog, follow these steps:

1. Navigate to your Microsoft Purview account in the Microsoft Purview governance portal.
2. Select **Data Map** on the left navigation.
3. Select **Register**
4. On Register sources, select **HDFS**. Select **Continue**.

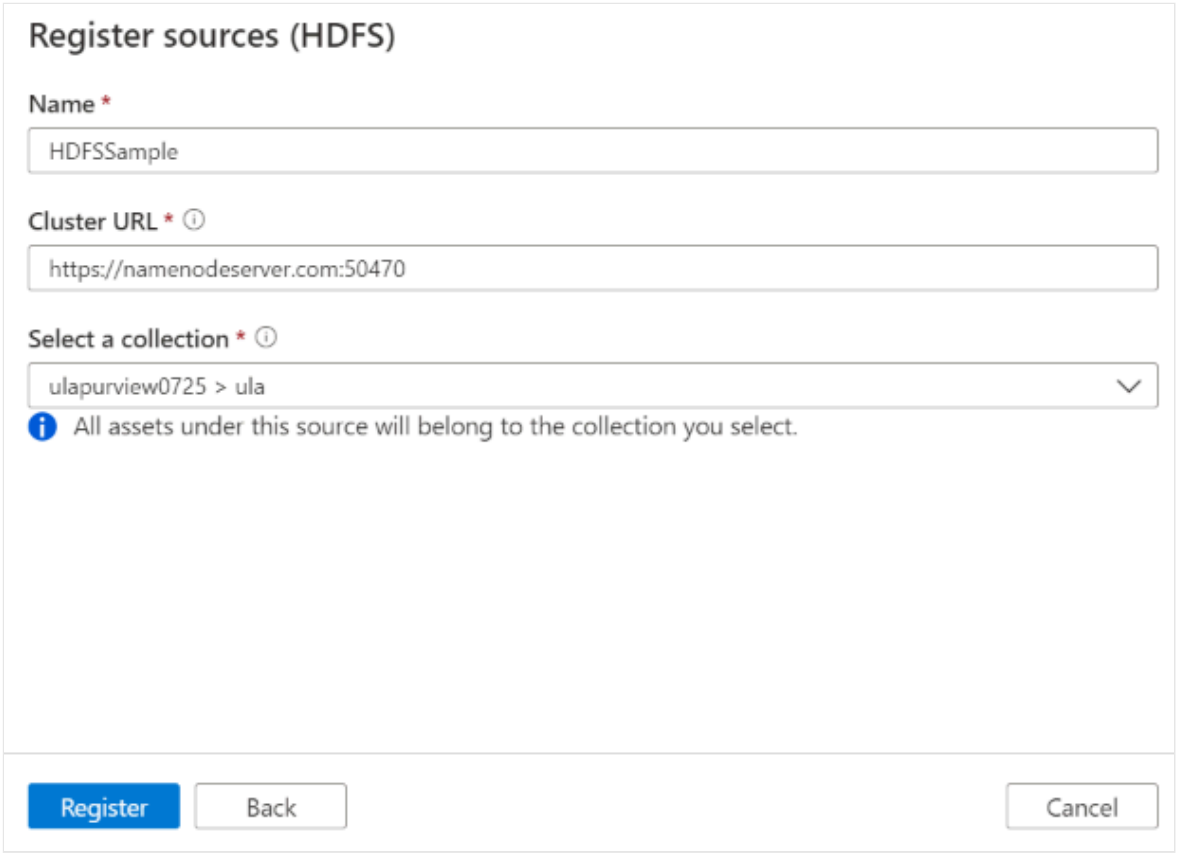
On the **Register sources (HDFS)** screen, follow these steps:

1. Enter a **Name** that the data source will be listed within the Catalog.

2. Enter the **Cluster URL** of the HDFS NameNode in the form of `https://<namenode>:<port>` or `http://<namenode>:<port>`, e.g. `https://namenodeserver.com:50470` or `http://namenodeserver.com:50070`.

3. Select a collection from the list.

4. Finish to register the data source.



Scan

Follow the steps below to scan HDFS to automatically identify assets. For more information about scanning in general, see our introduction to scans and ingestion.

Authentication for a scan

The supported authentication type for an HDFS source is **Kerberos authentication**.

Create and run scan

To create and run a new scan, follow these steps:

1. Make sure a self-hosted integration runtime is set up. If it isn't set up, use the steps mentioned in prerequisites to create a self-hosted integration runtime.

2. Navigate to **Sources**.

3. Select the registered HDFS source.

4. Select + **New scan**.

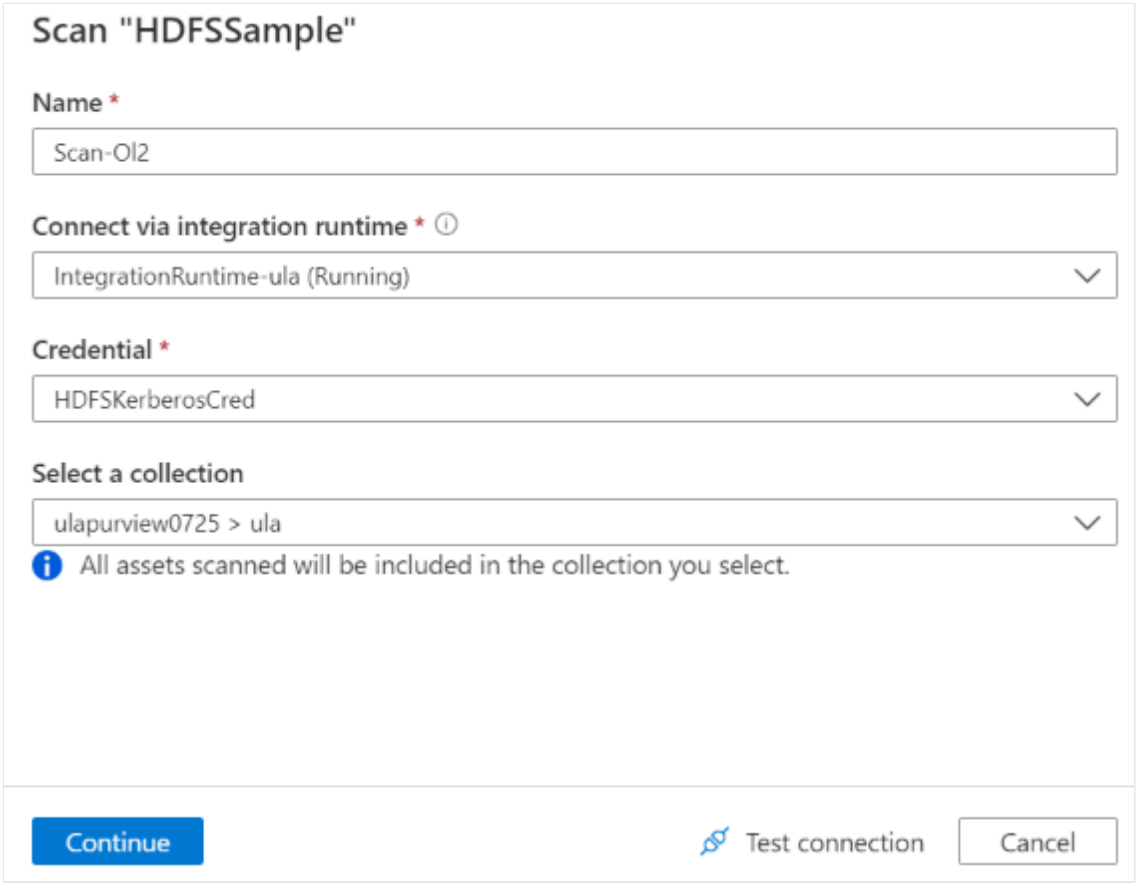
5. On "**Scan *source_name***"" page, provide the below details:

   a. **Name**: The name of the scan

   b. **Connect via integration runtime**: Select the configured self-hosted integration runtime. See setup requirements in Prerequisites section.

   c. **Credential**: Select the credential to connect to your data source. Make sure to:

   - Select **Kerberos Authentication** while creating a credential.
   - Provide the user name in the format of `<username>@<domain>.com` in the User name input field. Learn more from Use Kerberos authentication for the HDFS connector.
   - Store the user password used to connect to HDFS in the secret key.



6. Select **Test connection**.

7. Select **Continue**.

8. On "**Scope your scan**" page, select the path(s) that you want to scan.

9. On "**Select a scan rule set**" page, select the scan rule set you want to use for schema extraction and classification. You can choose between the system default, existing custom rule sets, or create a new rule set inline. Learn more from Create a scan rule set.

10. On "**Set a scan trigger**" page, choose your **scan trigger**. You can set up a schedule or ran the scan once.

11. Review your scan and select **Save and Run**.

View your scans and scan runs

To view existing scans:

1. Go to the Microsoft Purview portal. On the left pane, select **Data map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that has results you want to view. The pane shows you all the previous scan runs, along with the status and metrics for each scan run.
4. Select the run ID to check the scan run details.

Manage your scans

To edit, cancel, or delete a scan:

1. Go to the Microsoft Purview portal. On the left pane, select **Data Map**.

2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.

3. Select the scan that you want to manage. You can then:

    - Edit the scan by selecting **Edit scan**.
    - Cancel an in-progress scan by selecting **Cancel scan run**.
    - Delete your scan by selecting **Delete scan**.

> ⓘ **Note**
>
> - Deleting your scan does not delete catalog assets created from previous scans.

Use Kerberos authentication for the HDFS connector

There are two options for setting up the on-premises environment to use Kerberos authentication for the HDFS connector. You can choose the one that better fits your situation.

- Option 1: Join a self-hosted integration runtime machine in the Kerberos realm
- Option 2: Enable mutual trust between the Windows domain and the Kerberos realm

For either option, make sure you turn on webhdfs for Hadoop cluster:

1. Create the HTTP principal and keytab for webhdfs.

   > **Important**
   >
   > The HTTP Kerberos principal must start with "**HTTP/**" according to Kerberos HTTP SPNEGO specification. Learn more from here.

   Bash

   ```
   Kadmin> addprinc -randkey HTTP/<namenode hostname>@<REALM.COM>
   Kadmin> ktadd -k /etc/security/keytab/spnego.service.keytab
   HTTP/<namenode hostname>@<REALM.COM>
   ```

2. HDFS configuration options: add the following three properties in `hdfs-site.xml`.

   XML

   ```
   <property>
       <name>dfs.webhdfs.enabled</name>
       <value>true</value>
   </property>
   <property>
       <name>dfs.web.authentication.kerberos.principal</name>
       <value>HTTP/_HOST@<REALM.COM></value>
   </property>
   <property>
       <name>dfs.web.authentication.kerberos.keytab</name>
       <value>/etc/security/keytab/spnego.service.keytab</value>
   </property>
   ```

Option 1: Join a self-hosted integration runtime machine in the Kerberos realm

Requirements

- The self-hosted integration runtime machine needs to join the Kerberos realm and can't join any Windows domain.

How to configure

On the KDC server:

Create a principal, and specify the password.

> **Important**
>
> The username should not contain the hostname.

```
Bash

Kadmin> addprinc <username>@<REALM.COM>
```

On the self-hosted integration runtime machine:

1. Run the Ksetup utility to configure the Kerberos Key Distribution Center (KDC) server and realm.

   The machine must be configured as a member of a workgroup, because a Kerberos realm is different from a Windows domain. You can achieve this configuration by setting the Kerberos realm and adding a KDC server by running the following commands. Replace *REALM.COM* with your own realm name.

   ```
   Windows Command Prompt

   C:> Ksetup /setdomain REALM.COM
   C:> Ksetup /addkdc REALM.COM <your_kdc_server_address>
   ```

   After you run these commands, restart the machine.

2. Verify the configuration with the `Ksetup` command. The output should be like:

   ```
   Windows Command Prompt
   ```

```
C:> Ksetup
default realm = REALM.COM (external)
REALM.com:
    kdc = <your_kdc_server_address>
```

In your Purview account:

- Configure a credential with Kerberos authentication type with your Kerberos principal name and password to scan the HDFS. For configuration details, check the credential setting part in Scan section.

Option 2: Enable mutual trust between the Windows domain and the Kerberos realm

Requirements

- The self-hosted integration runtime machine must join a Windows domain.
- You need permission to update the domain controller's settings.

How to configure

> **Note**
>
> Replace REALM.COM and AD.COM in the following tutorial with your own realm name and domain controller.

On the KDC server:

1. Edit the KDC configuration in the *krb5.conf* file to let KDC trust the Windows domain by referring to the following configuration template. By default, the configuration is located at */etc/krb5.conf*.

```
config
```

```
[logging]
 default = FILE:/var/log/krb5libs.log
 kdc = FILE:/var/log/krb5kdc.log
 admin_server = FILE:/var/log/kadmind.log

[libdefaults]
 default_realm = REALM.COM
 dns_lookup_realm = false
 dns_lookup_kdc = false
```

```
 ticket_lifetime = 24h
 renew_lifetime = 7d
 forwardable = true

[realms]
 REALM.COM = {
  kdc = node.REALM.COM
  admin_server = node.REALM.COM
 }
AD.COM = {
 kdc = windc.ad.com
 admin_server = windc.ad.com
}

[domain_realm]
 .REALM.COM = REALM.COM
 REALM.COM = REALM.COM
 .ad.com = AD.COM
 ad.com = AD.COM

[capaths]
 AD.COM = {
  REALM.COM = .
 }
```

After you configure the file, restart the KDC service.

2. Prepare a principal named *krbtgt/REALM.COM@AD.COM* in the KDC server with the following command:

```
Windows Command Prompt

Kadmin> addprinc krbtgt/REALM.COM@AD.COM
```

3. In the *hadoop.security.auth_to_local* HDFS service configuration file, add `RULE:[1:$1@$0](.*\@AD.COM)s/\@.*//`.

On the domain controller:

1. Run the following `Ksetup` commands to add a realm entry:

```
Windows Command Prompt

C:> Ksetup /addkdc REALM.COM <your_kdc_server_address>
C:> ksetup /addhosttorealmmap HDFS-service-FQDN REALM.COM
```

2. Establish trust from the Windows domain to the Kerberos realm. [password] is the password for the principal *krbtgt/REALM.COM@AD.COM*.

Windows Command Prompt

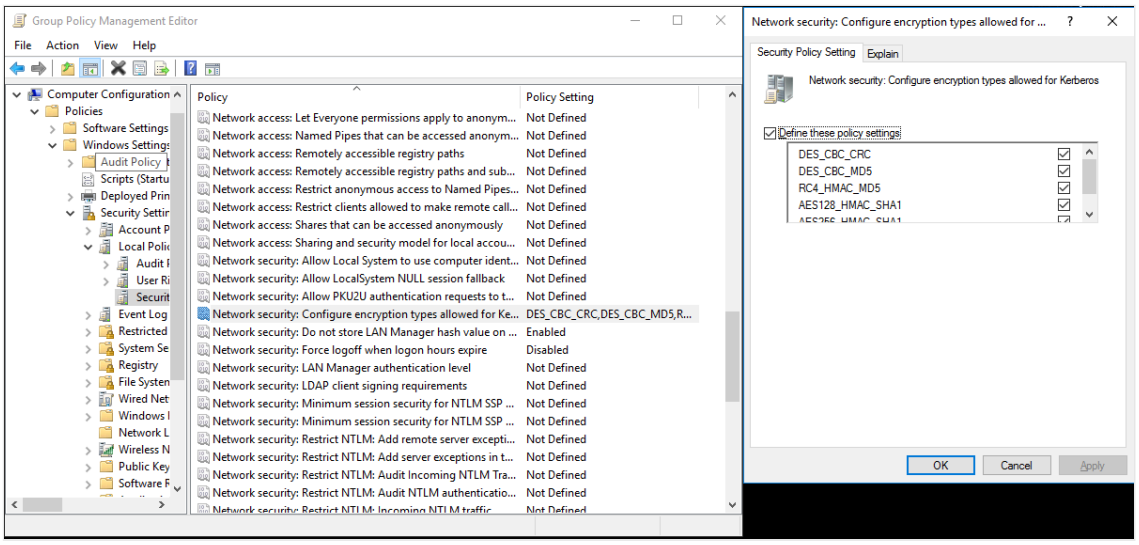
```
C:> netdom trust REALM.COM /Domain: AD.COM /add /realm /password:
[password]
```

3. Select the encryption algorithm that's used in Kerberos.

   a. Select **Server Manager** > **Group Policy Management** > **Domain** > **Group Policy Objects** > **Default or Active Domain Policy**, and then select **Edit**.

   b. On the **Group Policy Management Editor** pane, select **Computer Configuration** > **Policies** > **Windows Settings** > **Security Settings** > **Local Policies** > **Security Options**, and then configure **Network security: Configure Encryption types allowed for Kerberos**.

   c. Select the encryption algorithm you want to use when you connect to the KDC server. You can select all the options.



   d. Use the `Ksetup` command to specify the encryption algorithm to be used on the specified realm.

Windows Command Prompt

```
C:> ksetup /SetEncTypeAttr REALM.COM DES-CBC-CRC DES-CBC-MD5 RC4-
HMAC-MD5 AES128-CTS-HMAC-SHA1-96 AES256-CTS-HMAC-SHA1-96
```

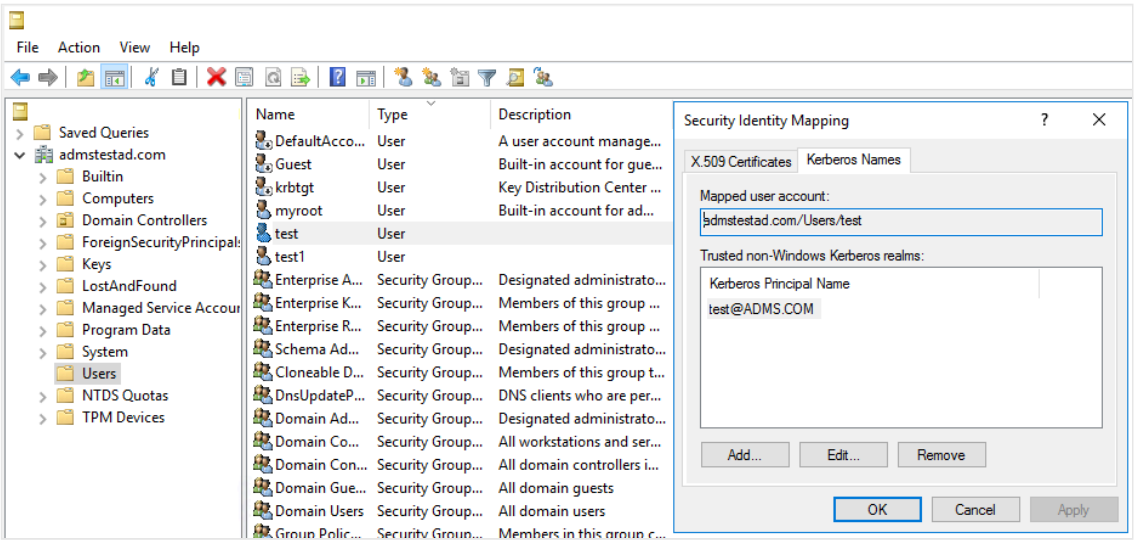
4. Create the mapping between the domain account and the Kerberos principal, so that you can use the Kerberos principal in the Windows domain.

   a. Select **Administrative tools** > **Active Directory Users and Computers**.

   b. Configure advanced features by selecting **View** > **Advanced Features**.

c. On the **Advanced Features** pane, right-click the account to which you want to create mappings and, on the **Name Mappings** pane, select the **Kerberos Names** tab.

d. Add a principal from the realm.



On the self-hosted integration runtime machine:

- Run the following `Ksetup` commands to add a realm entry.

```
Windows Command Prompt

C:> Ksetup /addkdc REALM.COM <your_kdc_server_address>
C:> ksetup /addhosttorealmmap HDFS-service-FQDN REALM.COM
```

In your Purview account:

- Configure a credential with Kerberos authentication type with your Kerberos principal name and password to scan the HDFS. For configuration details, check the credential setting part in Scan section.

Known limitations

Currently, HDFS connector doesn't support custom resource set pattern rule for advanced resource set, the built-in resource set patterns will be applied.

Sensitivity label is not yet supported.

Next steps

Now that you've registered your source, follow the below guides to learn more about Microsoft Purview and your data.

- Search Unified Catalog
- Data Estate Insights in Microsoft Purview

Feedback

Was this page helpful? Yes No

Connect to and manage Hive Metastore databases in Microsoft Purview

Article • 12/12/2024

This article outlines how to register Hive Metastore databases, and how to authenticate and interact with Hive Metastore databases in Microsoft Purview. For more information about Microsoft Purview, read the introductory article.

Supported capabilities

Scanning capabilities

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan
Yes	Yes	No	Yes

The supported Hive versions are 2.x to 3.x. The supported platforms are Apache Hadoop, Cloudera, and Hortonworks. If you want to scan Azure Databricks, you're suggested to use the Azure Databricks connector, which is more compatible and user friendly.

When scanning Hive metastore source, Microsoft Purview supports:

- Extracting technical metadata including:
  - Server
  - Databases
  - Tables including the columns, foreign keys, unique constraints, and storage description
  - Views including the columns and storage description
- Fetching static lineage on assets relationships among tables and views.

When setting up scan, you can choose to scan an entire Hive metastore database, or scope the scan to a subset of schemas matching the given name(s) or name pattern(s).

Other capabilities

View a table listing the supported capabilities for classifications, sensitivity labels, policies, data lineage, and live view.

Known limitations

When object is deleted from the data source, currently the subsequent scan won't automatically remove the corresponding asset in Microsoft Purview.

Prerequisites

- You must have an Azure account with an active subscription. Create an account for free.
- You must have an active Microsoft Purview account.
- You need Data Source Administrator and Data Reader permissions to register a source and manage it in the Microsoft Purview governance portal. For more information about permissions, see Access control in Microsoft Purview.
- If your data source isn't publicly accessible, set up the latest self-hosted integration runtime.
  - Choose the right integration runtime for your scenario:
    - **To use a self-hosted integration runtime:**
      - Follow the article to create and configure a self-hosted integration runtime.
      - Ensure JDK 11 is installed on the machine where the self-hosted integration runtime is installed. Restart the machine after you newly install the JDK for it to take effect.
      - Ensure that Visual C++ Redistributable (version Visual Studio 2012 Update 4 or newer) is installed on the machine where the self-hosted integration runtime is running. If you don't have this update installed, download it now.
      - Download the Hive Metastore database's JDBC driver on the machine where your self-hosted integration runtime is running. For example, if the database is *mssql*, download Microsoft's JDBC driver for SQL Server. Note down the folder path that you'll use to set up the scan.
    - **To use a kubernetes supported self-hosted integration runtime:**
      - Follow the article to create and configure a kubernetes supported integration runtime.
      - Download the Hive Metastore database's JDBC driver on the machine where your self-hosted integration runtime is running. For example, if the

database is *mssql*, download Microsoft's JDBC driver for SQL Server. Note down the folder path that you'll use to set up the scan.

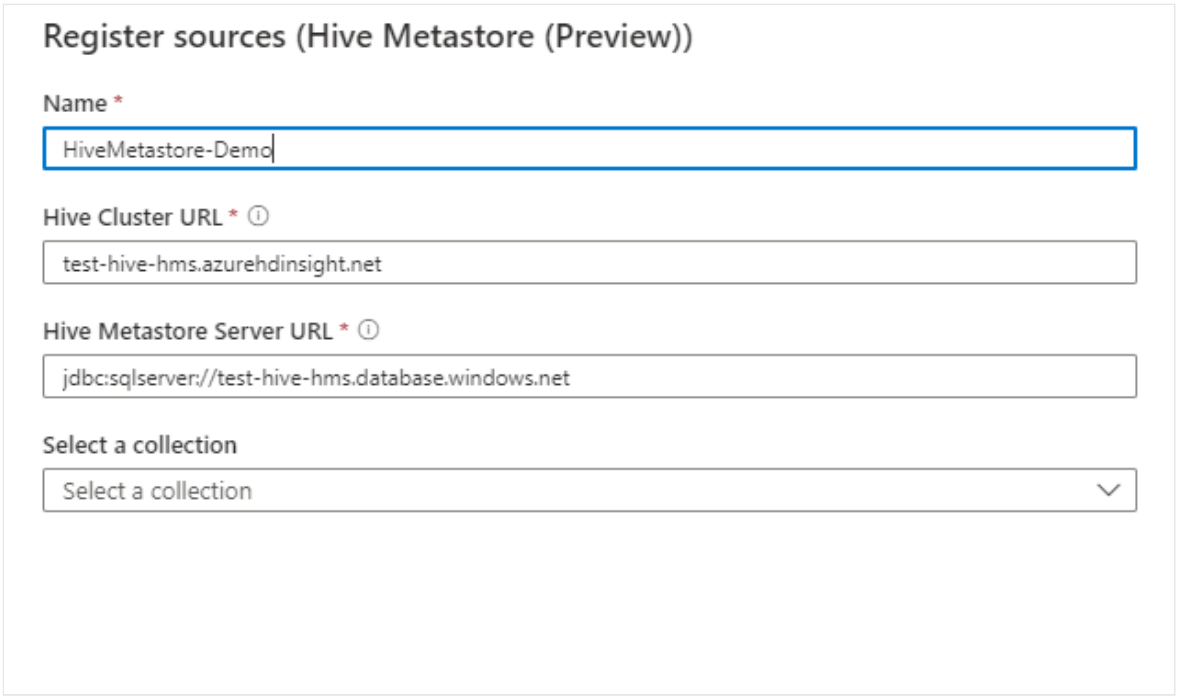
> **Note**
>
> The JDBC driver should be accessible by the self-hosted integration runtime. By default, self-hosted integration runtime uses local service account "NT SERVICE\DIAHostService". Make sure it has "Read and execute" and "List folder contents" permission to the driver folder.

Register

This section describes how to register a Hive Metastore database in Microsoft Purview by using the Microsoft Purview governance portal.

The only supported authentication for a Hive Metastore database is Basic Authentication.

1. Open the Microsoft Purview governance portal by:
   - Browsing directly to https://web.purview.azure.com and selecting your Microsoft Purview account.
   - Opening the Azure portal, searching for and selecting the Microsoft Purview account. Selecting the **the Microsoft Purview governance portal** button.
2. Select **Data Map** on the left pane.
3. Select **Register**.
4. In **Register sources**, select **Hive Metastore** > **Continue**.
5. On the **Register sources (Hive Metastore)** screen, do the following:
   a. For **Name**, enter a name that Microsoft Purview will list as the data source.
   b. For **Hive Cluster URL**, enter a value that you get from the Ambari URL. For example, enter **hive.azurehdinsight.net**.
   c. For **Hive Metastore Server URL**, enter a URL for the server. For example, enter **sqlserver://hive.database.windows.net**.
   d. Select a collection from the list.



6. Select **Finish**.

Scan

> **Tip**
>
> To troubleshoot any issues with scanning:
>
> 1. Confirm you have followed all **prerequisites**.
> 2. Review our **scan troubleshooting documentation**.

Use the following steps to scan Hive Metastore databases to automatically identify assets. For more information about scanning in general, see Scans and ingestion in Microsoft Purview.

1. In the Management Center, select integration runtimes. Make sure that a self-hosted integration runtime is set up. If it isn't set up, use the steps in prerequisites.

2. Go to **Sources**.

3. Select the registered Hive Metastore database.

4. Select + **New scan**.

5. Provide the following details:

    a. **Name**: Enter a name for the scan.

b. **Connect via integration runtime**: Select the configured self-hosted integration runtime.

c. **Credential**: Select the credential to connect to your data source. Make sure to:

- Select Basic Authentication while creating a credential.
- Provide the Metastore username in the appropriate box.
- Store the Metastore password in the secret key.

For more information, see Credentials for source authentication in Microsoft Purview.

d. **Metastore JDBC Driver Location**: Specify the path to the JDBC driver location in your machine where self-host integration runtime is running. For example, `D:\Drivers\HiveMetastore`.

i. For self-hosted integration runtime on a local machine: `D:\Drivers\HiveMetastore`. It's the path to valid JAR folder location. The value must be a valid absolute file path and doesn't contain space. Make sure the driver is accessible by the self-hosted integration runtime;; learn more from prerequisites section.

ii. For Kubernetes-supported self-hosted integration runtime: `./drivers/HiveMetastore`. It's the path to valid JAR folder location. The value must be a valid relative file path. Please refer to the documentation to set up a scan with external drivers for uploading drivers in advance.

e. **Metastore JDBC Driver Class**: Provide the class name for the connection driver. For example, enter **\com.microsoft.sqlserver.jdbc.SQLServerDriver**.

f. **Metastore JDBC URL**: Provide the connection URL value and define the connection to the URL of the Metastore database server. For example: `jdbc:sqlserver://hive.database.windows.net;database=hive;encrypt=true;trustServerCertificate=true;create=false;loginTimeout=300`.

> **Note**
>
> When you copy the URL from *hive-site.xml*, remove `amp;` from the string or the scan will fail.
>
> Download the SSL certificate to the self-hosted integration runtime machine, then update the path to the SSL certificate's location on your machine in the URL.

When you enter local file paths in the scan configuration, change the Windows path separator character from a backslash (`\`) to a forward slash (`/`). For example, if you place the SSL certificate at local file path *D:\Drivers\SSLCert\BaltimoreCyberTrustRoot.crt.pem*, change the `serverSslCert` parameter value to *D:/Drivers/SSLCert/BaltimoreCyberTrustRoot.crt.pem*.

The **Metastore JDBC URL** value will look like this example:

```
jdbc:mariadb://samplehost.mysql.database.azure.com:3306/XXXXXXXXXXXXXXXX?useSSL=true&enabledSslProtocolSuites=TLSv1,TLSv1.1,TLSv1.2&serverSslCert=D:/Drivers/SSLCert/BaltimoreCyberTrustRoot.crt.pem
```

g. **Metastore database name**: Provide the name of the Hive Metastore database.

h. **Schema**: Specify a list of Hive schemas to import. For example: **schema1; schema2**.

All user schemas are imported if that list is empty. All system schemas (for example, SysAdmin) and objects are ignored by default.

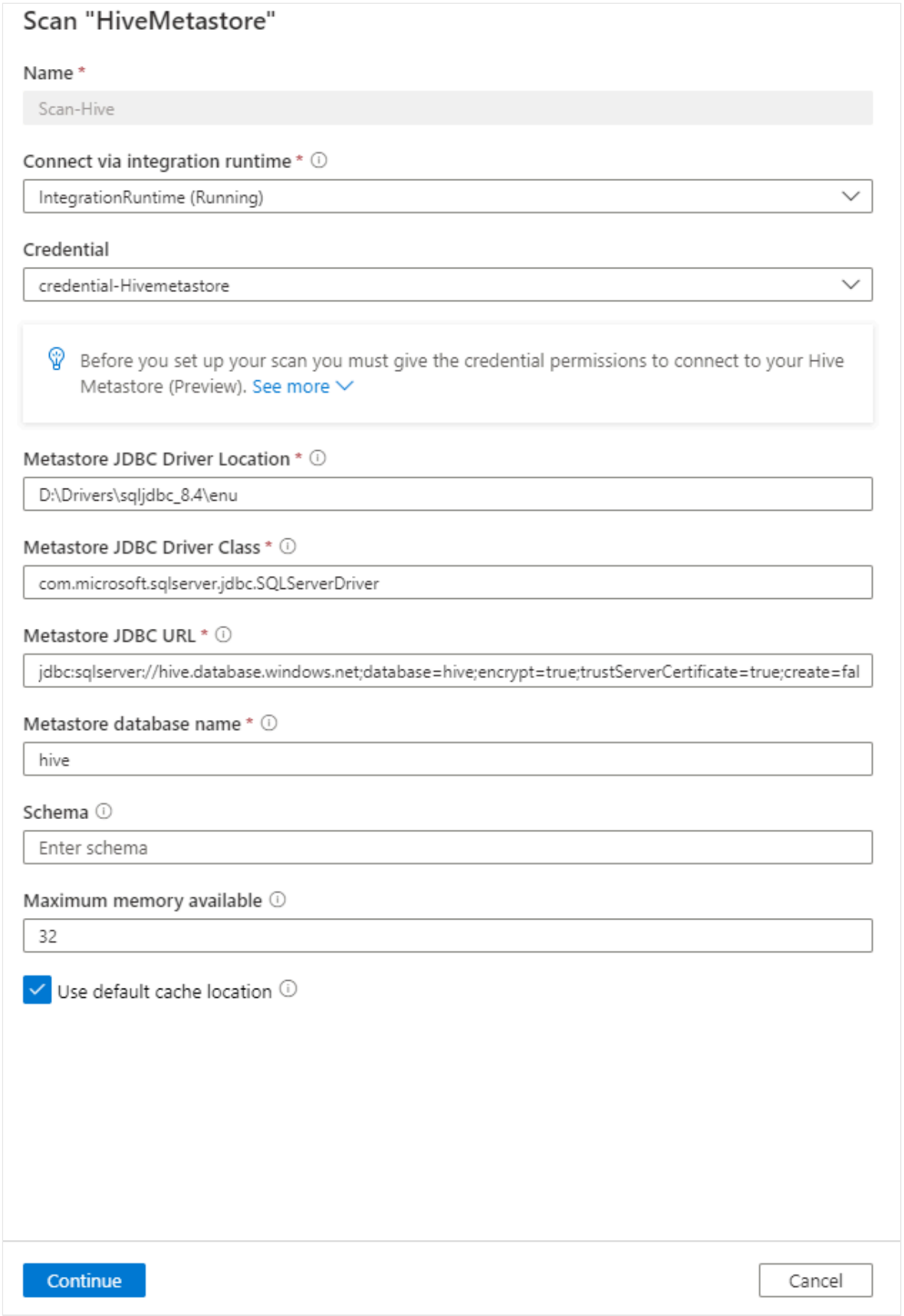
Acceptable schema name patterns that use SQL `LIKE` expression syntax include the percent sign (%). For example, `A%; %B; %C%; D` means:

- Start with A or
- End with B or
- Contain C or
- Equal D

Usage of `NOT` and special characters isn't acceptable.

i. **Maximum memory available**: Maximum memory (in gigabytes) available on the customer's machine for the scanning processes to use. This value is dependent on the size of Hive Metastore database to be scanned.

> ⓘ **Note**
>
> As a thumb rule, please provide 1GB memory for every 1000 tables.



6. Select **Continue**.

7. For **Scan trigger**, choose whether to set up a schedule or run the scan once.

8. Review your scan and select **Save and Run**.

View your scans and scan runs

To view existing scans:

1. Go to the Microsoft Purview portal. On the left pane, select **Data map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that has results you want to view. The pane shows you all the previous scan runs, along with the status and metrics for each scan run.
4. Select the run ID to check the scan run details.

Manage your scans

To edit, cancel, or delete a scan:

1. Go to the Microsoft Purview portal. On the left pane, select **Data Map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that you want to manage. You can then:
    - Edit the scan by selecting **Edit scan**.
    - Cancel an in-progress scan by selecting **Cancel scan run**.
    - Delete your scan by selecting **Delete scan**.

> **Note**
>
> - Deleting your scan does not delete catalog assets created from previous scans.

Lineage

After scanning your Hive Metastore source, you can browse Unified Catalog or search Unified Catalog to view the asset details.

Go to the asset -> lineage tab, you can see the asset relationship when applicable. Refer to the supported capabilities section on the supported Hive Metastore lineage scenarios. For more information about lineage in general, see data lineage and lineage user guide.

Next steps

Now that you've registered your source, use the following guides to learn more about Microsoft Purview and your data:

- Data Estate Insights in Microsoft Purview
- Lineage in Microsoft Purview
- Search the Unified Catalog

Feedback

Was this page helpful? Yes No

Connect to and manage Looker in Microsoft Purview (preview)

Article • 06/27/2024

This article outlines how to register Looker, and how to authenticate and interact with Looker in Microsoft Purview. For more information about Microsoft Purview, read the introductory article.

> **Important**
>
> This feature is currently in preview. The Supplemental Terms of Use for Microsoft Azure Previews include additional legal terms that apply to Azure features that are in beta, in preview, or otherwise not yet released into general availability.

Supported capabilities

Scanning capabilities

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan
Yes	Yes	No	Yes

The supported Looker server version is 7.2.

When scanning Looker source, Microsoft Purview supports:

- Extracting technical metadata including:
  - Server
  - Folders
  - Projects
  - Models
  - Dashboards
  - Looks
  - Explore diagrams including the joins
  - Views including the dimensions, measures, parameters, and filters
  - Layouts including the chart layouts, table layouts, text, and fields

- Fetching static lineage on assets relationships among views and layouts.

When setting up scan, you can choose to scan an entire Looker server, or scope the scan to a subset of Looker projects matching the given name(s).

Other capabilities

View a table listing the supported capabilities for classifications, sensitivity labels, policies, data lineage, and live view.

Known limitations

When object is deleted from the data source, currently the subsequent scan won't automatically remove the corresponding asset in Microsoft Purview.

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- An active Microsoft Purview account.
- You need Data Source Administrator and Data Reader permissions to register a source and manage it in the Microsoft Purview governance portal. For more information about permissions, see Access control in Microsoft Purview.

> **Note**
>
> **If your data store is not publicly accessible** (if your data store limits access from on-premises network, private network or specific IPs, etc.), **you will need to configure a self hosted integration runtime to connect to it**.

- If your data source isn't publicly accessible, set up the latest self-hosted integration runtime.
  - Choose the right integration runtime for your scenario:
    - **To use a self-hosted integration runtime:**
      - Follow the article to create and configure a self-hosted integration runtime.
      - Ensure JDK 11 is installed on the machine where the self-hosted integration runtime is installed. Restart the machine after you newly install the JDK for it to take effect.
      - Ensure that Visual C++ Redistributable (version Visual Studio 2012 Update 4 or newer) is installed on the machine where the self-hosted integration

runtime is running. If you don't have this update installed, download it now.

- **To use a kubernetes supported self-hosted integration runtime:**
  - Follow the article to create and configure a kubernetes supported integration runtime.

Register

This section describes how to register Looker in Microsoft Purview using the Microsoft Purview governance portal.

Authentication for registration

An API3 key is required to connect to the Looker server. The API3 key consists in a public client_id and a private client_secret and follows an OAuth2 authentication pattern.

Steps to register

To register a new Looker server in Microsoft Purview Unified Catalog, follow these steps:

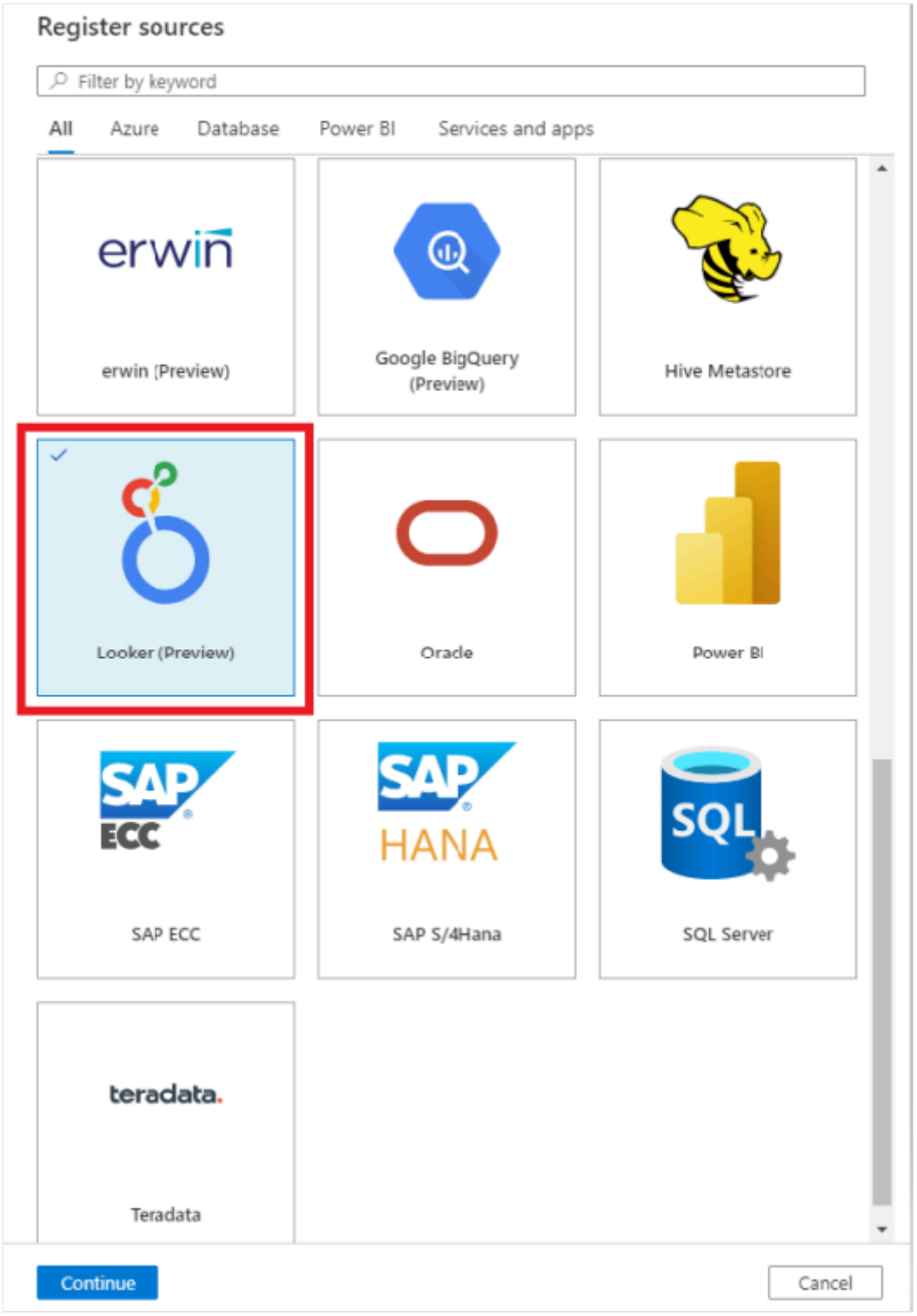
1. Open the Microsoft Purview governance portal by:

   - Browsing directly to https://web.purview.azure.com and selecting your Microsoft Purview account.
   - Opening the Azure portal, searching for and selecting the Microsoft Purview account. Selecting the **the Microsoft Purview governance portal** button.

2. Select **Data Map** on the left navigation.

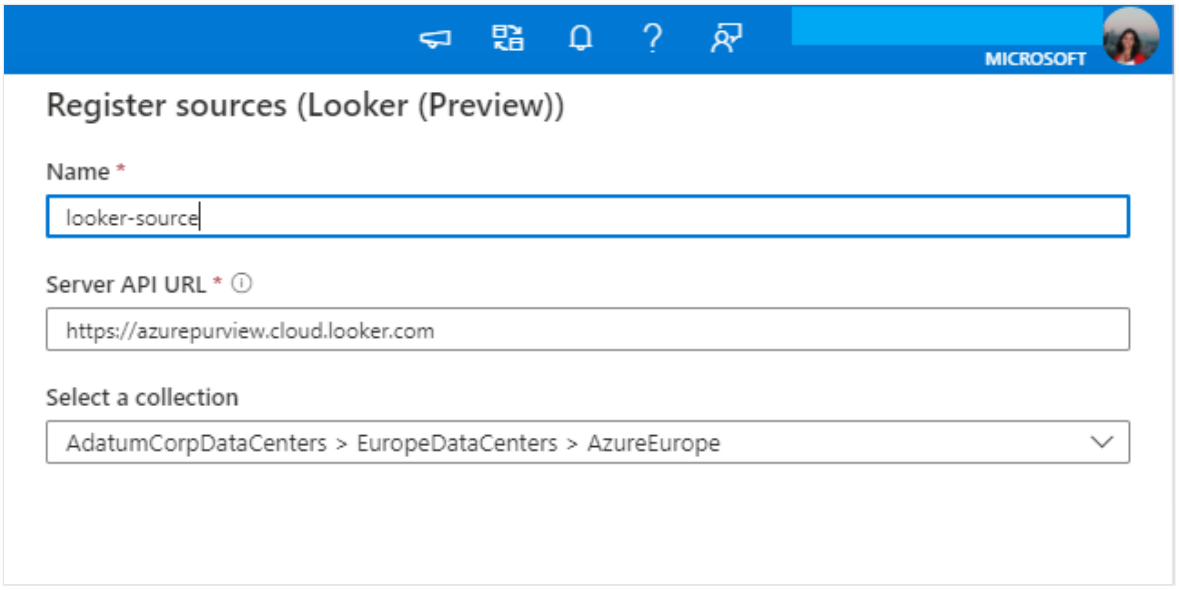
3. Select **Register.**

4. On Register sources, select **Looker**. Select **Continue.**



On the Register sources (Looker) screen, follow these steps:

1. Enter a **Name** that the data source will be listed within the Catalog.
2. Enter the Looker API URL in the **Server API URL** field. The default port for API requests is port 19999. Also, all Looker API endpoints require an HTTPS connection. For example: 'https://azurepurview.cloud.looker.com'
3. Select a collection from the list.
4. Finish to register the data source.



Scan

Follow the steps below to scan Looker to automatically identify assets. For more information about scanning in general, see our introduction to scans and ingestion

Create and run scan

To create and run a new scan, follow these steps:

1. If your server is publicly accessible, skip to step two. Otherwise, you'll need to make sure your self-hosted integration runtime is configured:
   a. In the Microsoft Purview governance portal, got to the Management Center, and select **Integration runtimes**.
   b. Make sure a self-hosted integration runtime is available. If one isn't set up, use the steps mentioned in prerequisites to set up a self-hosted integration runtime.

2. In the Microsoft Purview governance portal, navigate to **Sources**.

3. Select the registered **Looker** server.

4. Select + **New scan**.

5. Provide the below details:

   a. **Name**: The name of the scan

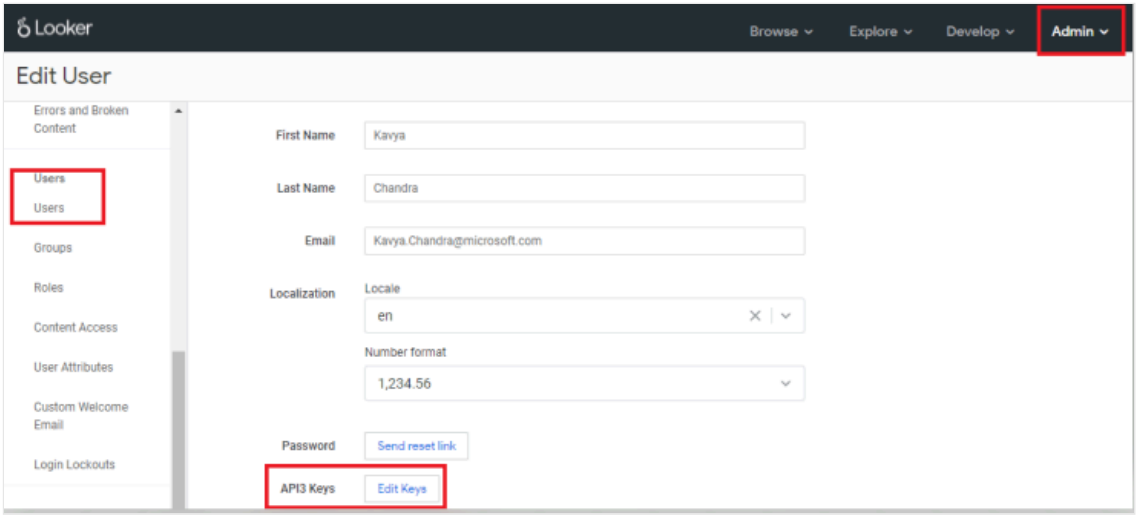
   b. **Connect via integration runtime**: Select the Azure auto-resolved integration runtime if your server is publicly accessible, or your configured self-hosted integration runtime if it isn't publicly available.

c. **Server API URL** is auto populated based on the value entered during registration.

d. **Credential:** While configuring Looker credential, make sure to:

- Select **Basic Authentication** as the Authentication method
- Provide your Looker API3 key's client ID in the User name field
- Save your Looker API3 key's client secret in the key vault's secret.

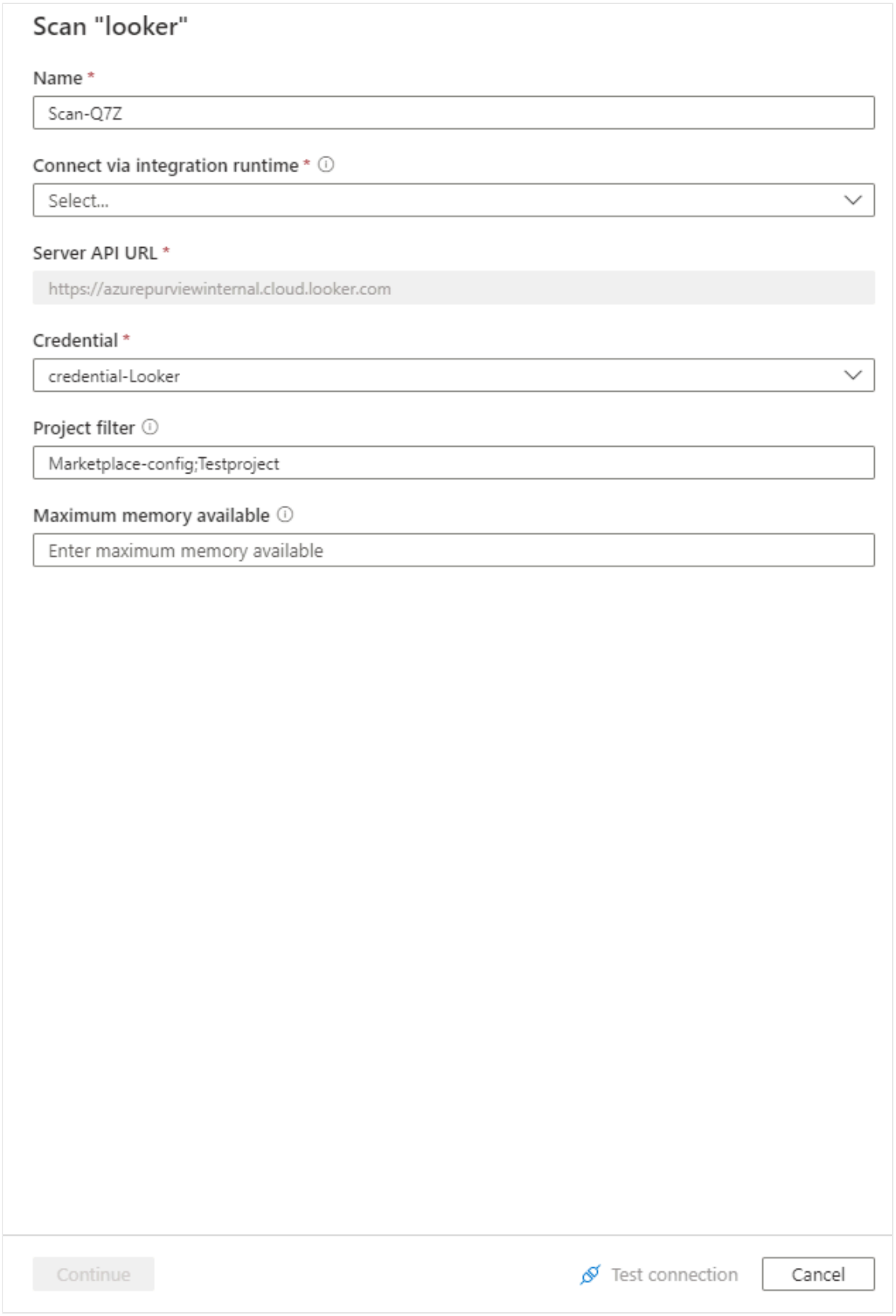
To access client ID and client secret, navigate to Looker ->Admin -> Users -> Select **Edit** on a user -> Select **EditKeys** -> Use the Client ID and Client Secret or create a new one.



To understand more on credentials, refer to the link here

e. **Model filter** - Scope your scan by providing a semicolon separated list of Looker models. This option is used to select looks and dashboards by their parent model.

f. **Maximum memory available** (applicable when using self-hosted integration runtime): Maximum memory (in GB) available on customer's VM to be used by scanning processes. This is dependent on the size of erwin Mart to be scanned.



6. Select **Test connection** to validate the settings.

7. Select **Continue**.

8. Choose your **scan trigger**. You can set up a schedule or ran the scan once.

9. Review your scan and select on **Save and Run**.

View your scans and scan runs

To view existing scans:

1. Go to the Microsoft Purview portal. On the left pane, select **Data map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that has results you want to view. The pane shows you all the previous scan runs, along with the status and metrics for each scan run.
4. Select the run ID to check the scan run details.

Manage your scans

To edit, cancel, or delete a scan:

1. Go to the Microsoft Purview portal. On the left pane, select **Data Map**.

2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.

3. Select the scan that you want to manage. You can then:

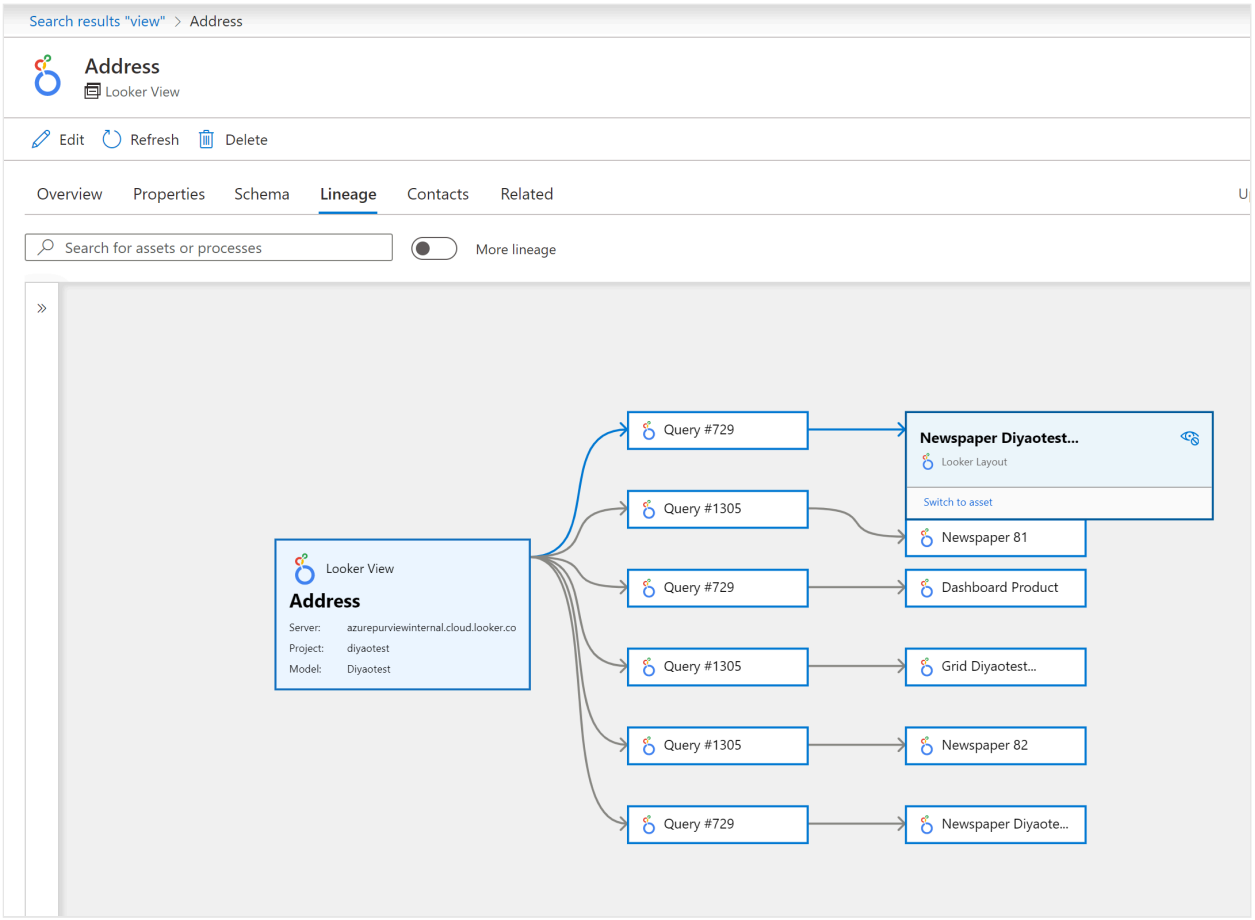
    - Edit the scan by selecting **Edit scan**.
    - Cancel an in-progress scan by selecting **Cancel scan run**.
    - Delete your scan by selecting **Delete scan**.

> ⓘ **Note**
>
> - Deleting your scan does not delete catalog assets created from previous scans.

Lineage

After scanning your Looker source, you can browse Unified Catalog or search Unified Catalog to view the asset details.

Go to the asset -> lineage tab, you can see the asset relationship when applicable. Refer to the supported capabilities section on the supported Looker lineage scenarios. For more information about lineage in general, see data lineage and lineage user guide.



Next steps

Now that you've registered your source, follow the below guides to learn more about Microsoft Purview and your data.

- Data Estate Insights in Microsoft Purview
- Lineage in Microsoft Purview
- Search Unified Catalog

Feedback

Was this page helpful? Yes No

Connect to and manage MongoDB in Microsoft Purview

This article outlines how to register MongoDB, and how to authenticate and interact with MongoDB in Microsoft Purview. For more information about Microsoft Purview, read the introductory article.

Supported capabilities

Scanning capabilities

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan
Yes	Yes	No	Yes

The supported MongoDB versions are 3.6 to 7.0. You can use this connector to scan MongoDB Atlas as well.

When scanning MongoDB source, Microsoft Purview supports extracting technical metadata including:

- Server
- Databases
- Collections including the schema
- Views including the schema

During scan, Microsoft Purview retrieves and analyzes sample documents to infer the collection/view schema. The sample size is configurable.

When setting up scan, you can choose to scan one or more MongoDB databases entirely, or further scope the scan to a subset of collections matching the given names or name patterns.

Other capabilities

For **classifications**, **sensitivity labels**, **policies**, **data lineage**, and **live view**, see the list of supported capabilities.

Known limitations

When object is deleted from the data source, currently the subsequent scan won't automatically remove the corresponding asset in Microsoft Purview.

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- An active Microsoft Purview account.
- You need Data Source Administrator and Data Reader permissions to register a source and manage it in the Microsoft Purview governance portal. For more information about permissions, see Access control in Microsoft Purview.
- If your data source isn't publicly accessible, set up the latest self-hosted integration runtime.
  - Choose the right integration runtime for your scenario:
    - **To use a self-hosted integration runtime:**
      - Follow the article to create and configure a self-hosted integration runtime.
      - Ensure JDK 11 is installed on the machine where the self-hosted integration runtime is installed. Restart the machine after you newly install the JDK for it to take effect.
      - Ensure that Visual C++ Redistributable (version Visual Studio 2012 Update 4 or newer) is installed on the machine where the self-hosted integration runtime is running. If you don't have this update installed, download it now.
    - **To use a kubernetes supported self-hosted integration runtime:**
      - Follow the article to create and configure a kubernetes supported integration runtime.

Register

This section describes how to register MongoDB in Microsoft Purview using the classic Microsoft Purview governance portal.

Steps to register

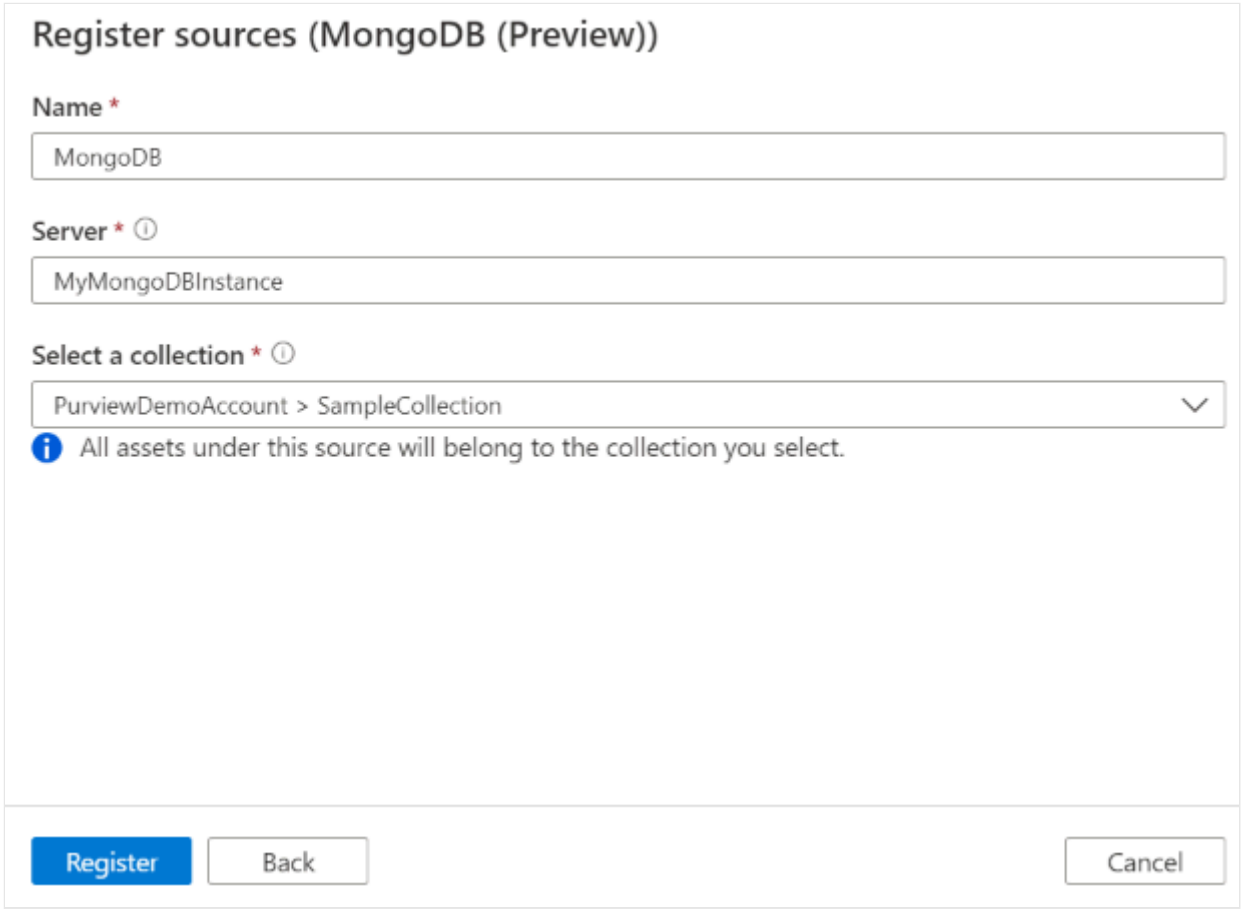
To register a new MongoDB source in Microsoft Purview Unified Catalog, do the following:

1. Navigate to your Microsoft Purview account in the classic Microsoft Purview governance portal.
2. Select **Data Map** on the left navigation.
3. Select **Register**

4. On Register sources, select **MongoDB**. Select **Continue**.

On the **Register sources (MongoDB)** screen, do the following:

1. Enter a **Name** that the data source will be listed within the Catalog.
2. Enter the **server** name. Specify a name to uniquely identify your MongoDB instance in your company. For example, `host` for standalone deployment, `MyReplicaSetName` for replica set, `MyClusterName` for sharded cluster. This value is used in asset qualified name and can't be changed.
3. Select a collection from the list.
4. Finish to register the data source.



Scan

Follow the steps below to scan MongoDB to automatically identify assets. For more information about scanning in general, see our introduction to scans and ingestion.

Authentication for a scan

The supported authentication type for a MongoDB source is **Basic authentication**.

Create and run scan

To create and run a new scan, do the following:

1. In the Management Center, select Integration runtimes. Make sure a self-hosted integration runtime is set up. If it isn't set up, use the steps mentioned in prerequisites to create a self-hosted integration runtime.

2. Navigate to **Sources**.

3. Select the registered MongoDB source.

4. Select + **New scan**.

5. Provide the below details:

    a. **Name**: The name of the scan

    b. **Connect via integration runtime**: Select the self-hosted integration runtime used to perform scan.

    c. **Credential**: Select the credential to connect to your data source. Make sure to:

    - Select **Basic Authentication** while creating a credential.
    - Provide the user name used to connect to MongoDB in the User name input field.
    - Store the user password used to connect to MongoDB in the secret key.

    d. **Connection string**: Specify the MongoDB connection string used to connect to your MongoDB, excluding the username and password. For example, `mongodb://mongodb0.example.com:27017,mongodb1.example.com:27017/?replicaSet=myRepl`, `mongodb+srv://mongodb0.example.com/?authSource=admin&replicaSet=myRepl`.

    e. **Databases**: Specify a list of MongoDB databases to be imported. The list can have one or more database names separated by semicolon (;), for example, `database1;database2`.

    f. **Collections**: The subset of collections to import expressed as a semicolon separated list of collections, for example, `collection1;collection2`. All collections are imported if the list is empty.

    Acceptable collection name patterns can be static names or contain wildcard %. For example: `A%;%B;%C%;D`:

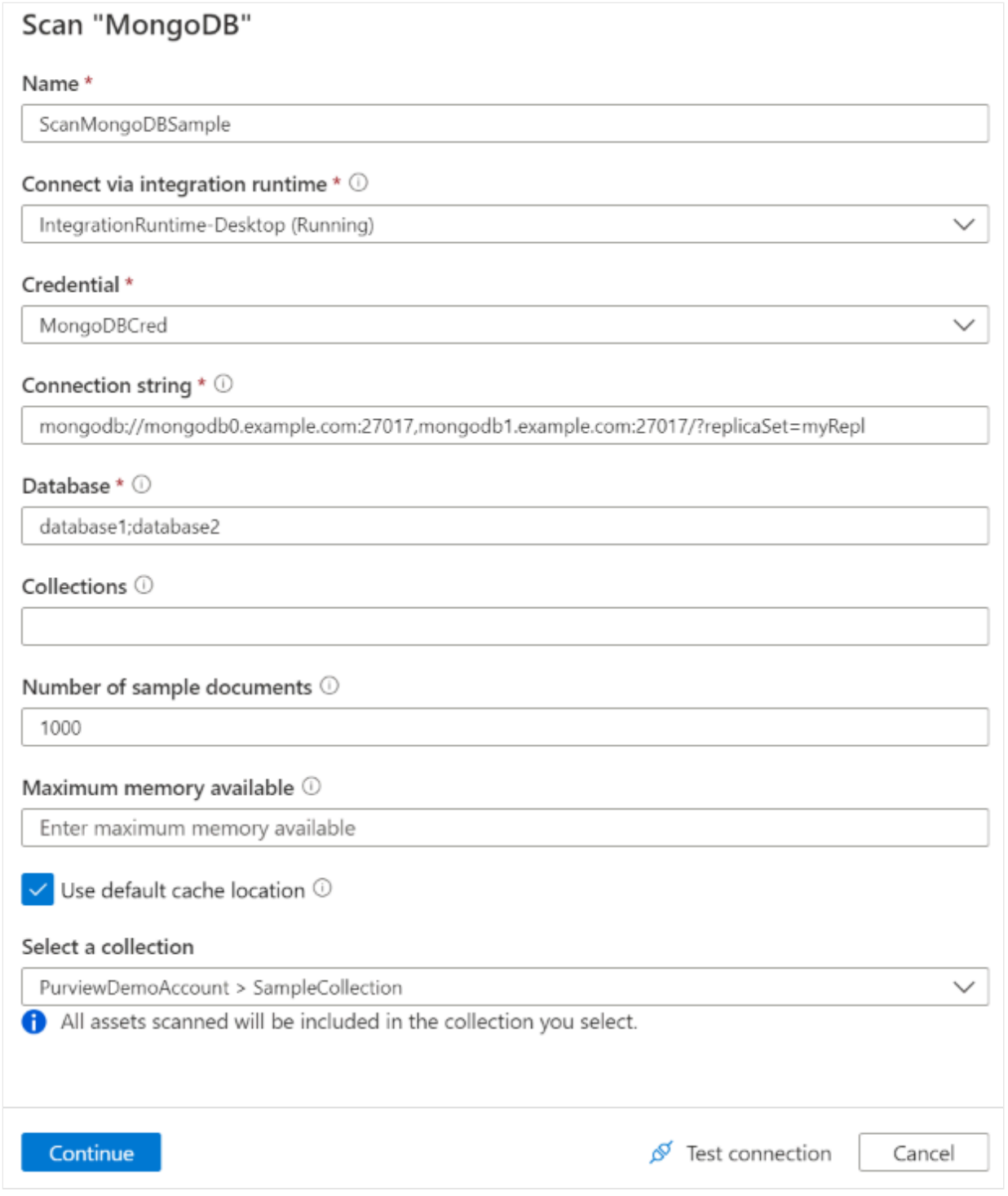
    - Start with A or

- End with B or
- Contain C or
- Equal D

Usage of NOT and special characters aren't acceptable.

g. **Number of sample documents**: Number of sample documents to be analyzed for schema extraction. Default is 10.

h. **Maximum memory available** (applicable when using self-hosted integration runtime): Maximum memory (in GB) available on customer's VM to be used by scanning processes. It's dependent on the size of MongoDB source to be scanned.



6. Select **Test connection** to validate the configurations.

   > ⓘ Note
   >
   > After you select **Test connection**, this message might appear: "The legacy MongoDB connector is deprecated." If you get this message, it doesn't affect the scan. You can skip the test connection and continue scanning Mongo DB in Microsoft Purview.

7. Select **Continue**.

8. Choose your **scan trigger**. You can set up a schedule or ran the scan once.

9. Review your scan and select **Save and Run**.

View your scans and scan runs

To view existing scans:

1. Go to the Microsoft Purview portal. On the left pane, select **Data map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that has results you want to view. The pane shows you all the previous scan runs, along with the status and metrics for each scan run.
4. Select the run ID to check the scan run details.

Manage your scans

To edit, cancel, or delete a scan:

1. Go to the Microsoft Purview portal. On the left pane, select **Data Map**.

2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.

3. Select the scan that you want to manage. You can then:

   - Edit the scan by selecting **Edit scan**.
   - Cancel an in-progress scan by selecting **Cancel scan run**.
   - Delete your scan by selecting **Delete scan**.

> ⓘ Note

- Deleting your scan does not delete catalog assets created from previous scans.

Next steps

Now that you've registered your source, visit these articles to learn more about Microsoft Purview and your data.

- Data Estate Insights in Microsoft Purview
- Search Unified Catalog

Last updated on 01/20/2026

Connect to and manage MySQL in Microsoft Purview

Article • 03/28/2025

This article outlines how to register MySQL, and how to authenticate and interact with MySQL in Microsoft Purview. For more information about Microsoft Purview, read the introductory article.

Supported capabilities

Scanning capabilities

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan
Yes	Yes	No	Yes

The supported MySQL server versions are 5.7 to 8.x.

When scanning MySQL source, Microsoft Purview supports:

- Extracting technical metadata including:
  - Server
  - Databases
  - Tables including the columns
  - Views including the columns
- Fetching static lineage on assets relationships among tables and views.

When setting up scan, you can choose to scan an entire MySQL server, or scope the scan to a subset of databases matching the given name(s) or name pattern(s).

Other capabilities

View a table listing the supported capabilities for classifications, sensitivity labels, policies, data lineage, and live view.

Known limitations

When object is deleted from the data source, currently the subsequent scan won't automatically remove the corresponding asset in Microsoft Purview.

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- An active Microsoft Purview account.
- You need Data Source Administrator and Data Reader permissions to register a source and manage it in the Microsoft Purview governance portal. For more information about permissions, see Access control in Microsoft Purview.

> **Note**
>
> **If your data store is not publicly accessible** (if your data store limits access from on-premises network, private network or specific IPs, etc.), **you will need to configure a self hosted integration runtime to connect to it**.

- If your data source isn't publicly accessible, set up the latest self-hosted integration runtime.
  - Choose the right integration runtime for your scenario:
    - **To use a self-hosted integration runtime:**
      - Follow the article to create and configure a self-hosted integration runtime.
      - Ensure JDK 11 is installed on the machine where the self-hosted integration runtime is installed. Restart the machine after you newly install the JDK for it to take effect.
      - Ensure that Visual C++ Redistributable (version Visual Studio 2012 Update 4 or newer) is installed on the machine where the self-hosted integration runtime is running. If you don't have this update installed, download it now.
    - **To use a kubernetes supported self-hosted integration runtime:**
      - Follow the article to create and configure a kubernetes supported integration runtime.

Required permissions for scan

The MySQL user must have the SELECT, SHOW VIEW and EXECUTE permissions for each target MySQL schema that contains metadata.

Register

This section describes how to register MySQL in Microsoft Purview using the Microsoft Purview governance portal.

Steps to register

To register a new MySQL source in Microsoft Purview Unified Catalog, follow these steps:

1. Navigate to your Microsoft Purview account in the Microsoft Purview governance portal.
2. Select **Data Map** on the left navigation.
3. Select **Register**
4. On Register sources, select **MySQL**. Select **Continue**.

On the **Register sources (MySQL)** screen, follow these steps:

1. Enter a **Name** that the data source will be listed within the Catalog.

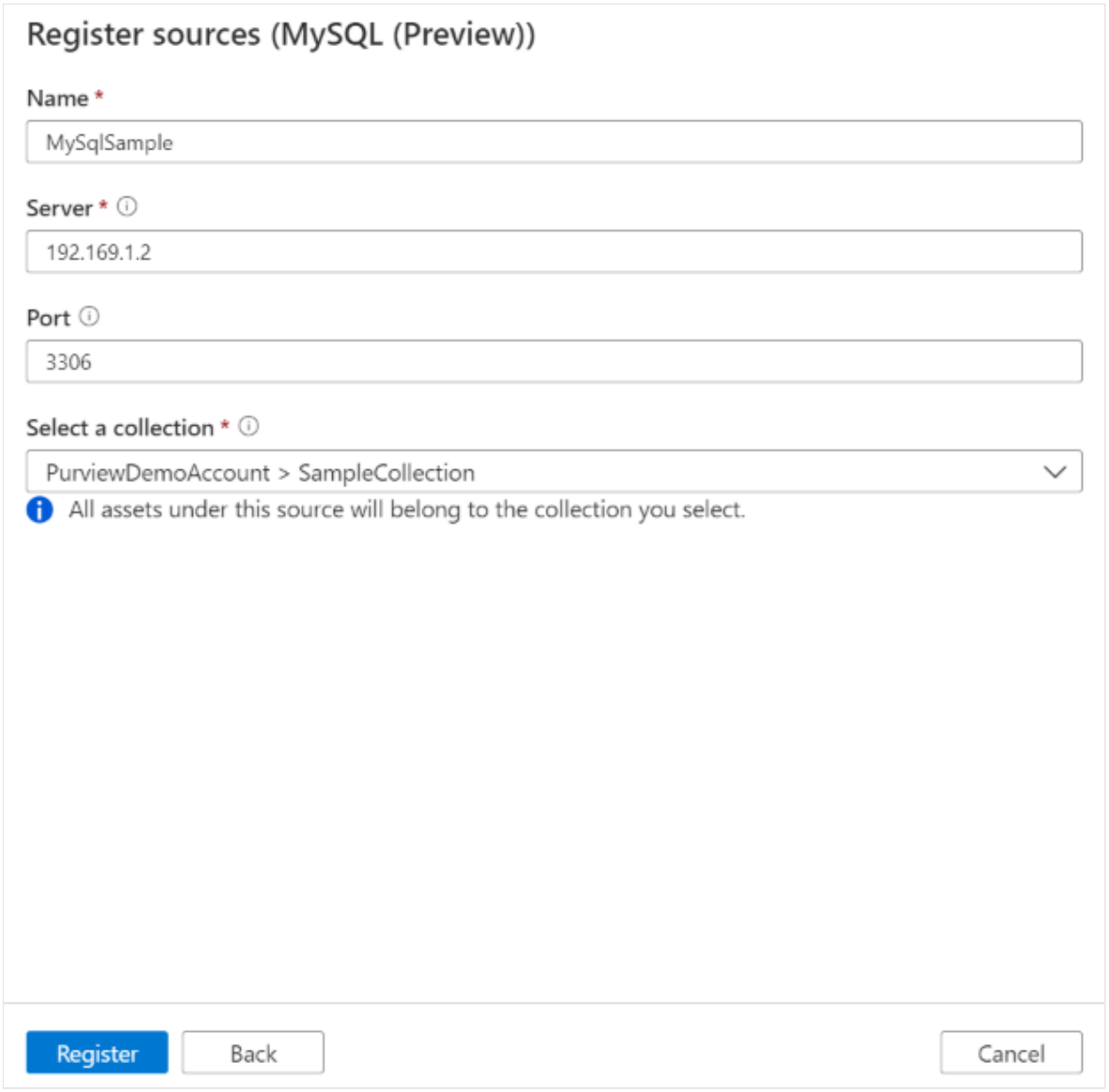
2. Enter the **Server** name to connect to a MySQL source. This can either be:

    - A host name used to connect to the database server. For example: `MyDatabaseServer.com`
    - An IP address. For example: `192.169.1.2`

3. Enter the **Port** used to connect to the database server (3306 by default for MySQL).

4. Select a collection from the list.

5. Finish to register the data source.



Scan

Follow the steps below to scan MySQL to automatically identify assets. For more information about scanning in general, see our introduction to scans and ingestion.

Authentication for a scan

The supported authentication type for a MySQL source is **Basic authentication**.

Create and run scan

To create and run a new scan, follow these steps:

1. If your server is publicly accessible, skip to step two. Otherwise, you'll need to make sure your self-hosted integration runtime is configured:
   a. In the Microsoft Purview governance portal, got to the Management Center, and select **Integration runtimes**.

b. Make sure a self-hosted integration runtime is available. If one isn't set up, use the steps mentioned in prerequisites to set up a self-hosted integration runtime.

2. In the Microsoft Purview governance portal, navigate to **Sources**.

3. Select the registered MySQL source.

4. Select + **New scan**.

5. Provide the below details:

   a. **Name**: The name of the scan

   b. **Connect via integration runtime**: Select the Azure auto-resolved integration runtime if your server is publicly accessible, or your configured self-hosted integration runtime if it isn't publicly available.

   c. **Credential**: Select the credential to connect to your data source. Make sure to:

      - Select **Basic Authentication** while creating a credential.
      - Provide the user name used to connect to the database server in the User name input field.
      - Store the user password used to connect to the database server in the secret key.

   d. **Database**: List subset of databases to import expressed as a semicolon separated list. For example, `database1;database2`. All user databases are imported if the list is empty. All system databases (for example, SysAdmin) are ignored by default.

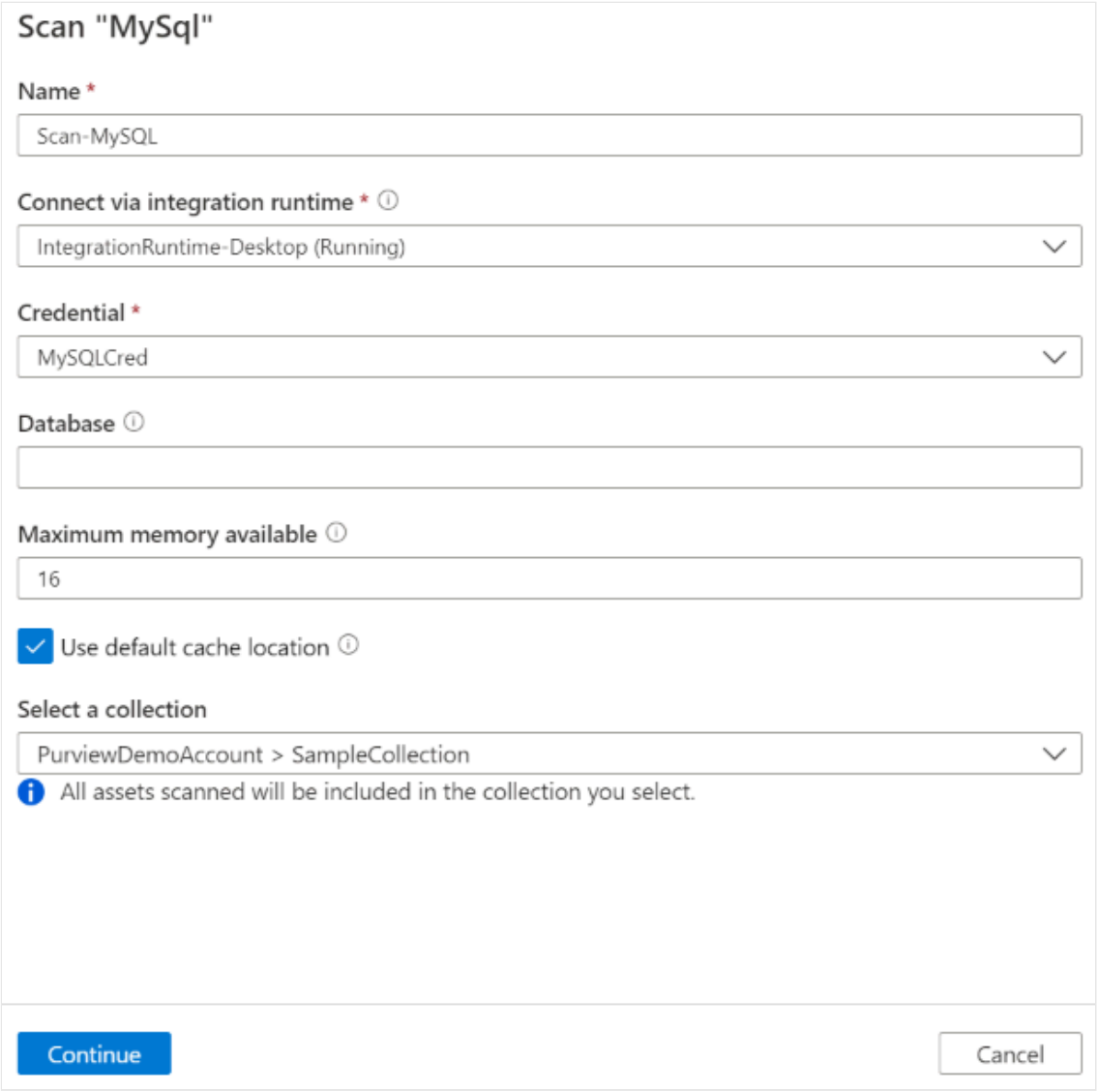
      Acceptable schema name patterns can be static names or contain wildcard %. For example: `A%;%B;%C%;D`

      - Start with A or
      - End with B or
      - Contain C or
      - Equal D

      Usage of NOT and special characters aren't acceptable.

   e. **Maximum memory available** (applicable when using self-hosted integration runtime): Maximum memory (in GB) available on customer's VM to be used by scanning processes. This is dependent on the size of MySQL source to be scanned.

> **Note**
>
> As a rule of thumb, please provide 1GB memory for every 1000 tables



6. Select **Test connection** to validate the settings (available when using Azure Integration Runtime).

7. Select **Continue**.

8. Choose your **scan trigger**. You can set up a schedule or ran the scan once.

9. Review your scan and select **Save and Run**.

View your scans and scan runs

To view existing scans:

1. Go to the Microsoft Purview portal. On the left pane, select **Data map**.

2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that has results you want to view. The pane shows you all the previous scan runs, along with the status and metrics for each scan run.
4. Select the run ID to check the scan run details.

Manage your scans

To edit, cancel, or delete a scan:

1. Go to the Microsoft Purview portal. On the left pane, select **Data Map**.

2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.

3. Select the scan that you want to manage. You can then:

    - Edit the scan by selecting **Edit scan**.
    - Cancel an in-progress scan by selecting **Cancel scan run**.
    - Delete your scan by selecting **Delete scan**.

> ⓘ **Note**
>
> - Deleting your scan does not delete catalog assets created from previous scans.

Lineage

After scanning your MySQL source, you can browse Unified Catalog or search Unified Catalog to view the asset details.

Go to the asset -> lineage tab, you can see the asset relationship when applicable. Refer to the supported capabilities section on the supported MySQL lineage scenarios. For more information about lineage in general, see data lineage and lineage user guide.

Next steps

Now that you've registered your source, follow the below guides to learn more about Microsoft Purview and your data.

- Data Estate Insights in Microsoft Purview

- Lineage in Microsoft Purview
- Search Unified Catalog

Feedback

Was this page helpful? Yes No

Connect to and manage an on-premises SQL server instance in Microsoft Purview

This article outlines how to register on-premises SQL server instances, and how to authenticate and interact with an on-premises SQL server instance in Microsoft Purview. For more information about Microsoft Purview, read the introductory article.

Supported capabilities

Scanning capabilities

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan
Yes	Yes	Yes	Yes

The supported SQL Server versions are 2005 and above. SQL Server Express LocalDB isn't supported.

When you scan an on-premises SQL server, Microsoft Purview supports:

- Extracting technical metadata including:
  - Instance
  - Databases
  - Schemas
  - Tables including the columns
  - Views including the columns

When setting up scan, you can choose to specify the database name to scan one database, and you can further scope the scan by selecting tables and views as needed. The whole SQL Server instance will be scanned if database name isn't provided.

Other capabilities

For **classifications**, **sensitivity labels**, **policies**, **data lineage**, and **live view**, see the list of supported capabilities.

Prerequisites

- The enterprise version of Microsoft Purview or an active Microsoft Purview account using the classic governance portal.
- You'll need to be a Data Source Administrator and Data Reader to register a source in the Microsoft Purview Data Map. Get details about domain and collection permissions.
- Set up the right integration runtime for your scenario:
    - **To use a self-hosted integration runtime:** Use these instructions to create and configure a self-hosted integration runtime. In this case, classification of SqlBigDecimal isn't supported.
        - **To use a kubernetes supported self-hosted integration runtime:** Follow the article to create and configure a kubernetes supported integration runtime. (SQL authentication only.)

Register

This section describes how to register an on-premises SQL server instance in Microsoft Purview using the Microsoft Purview portal.

1. Accessing the Microsoft Purview Data Map depends on which portal experience you're using.

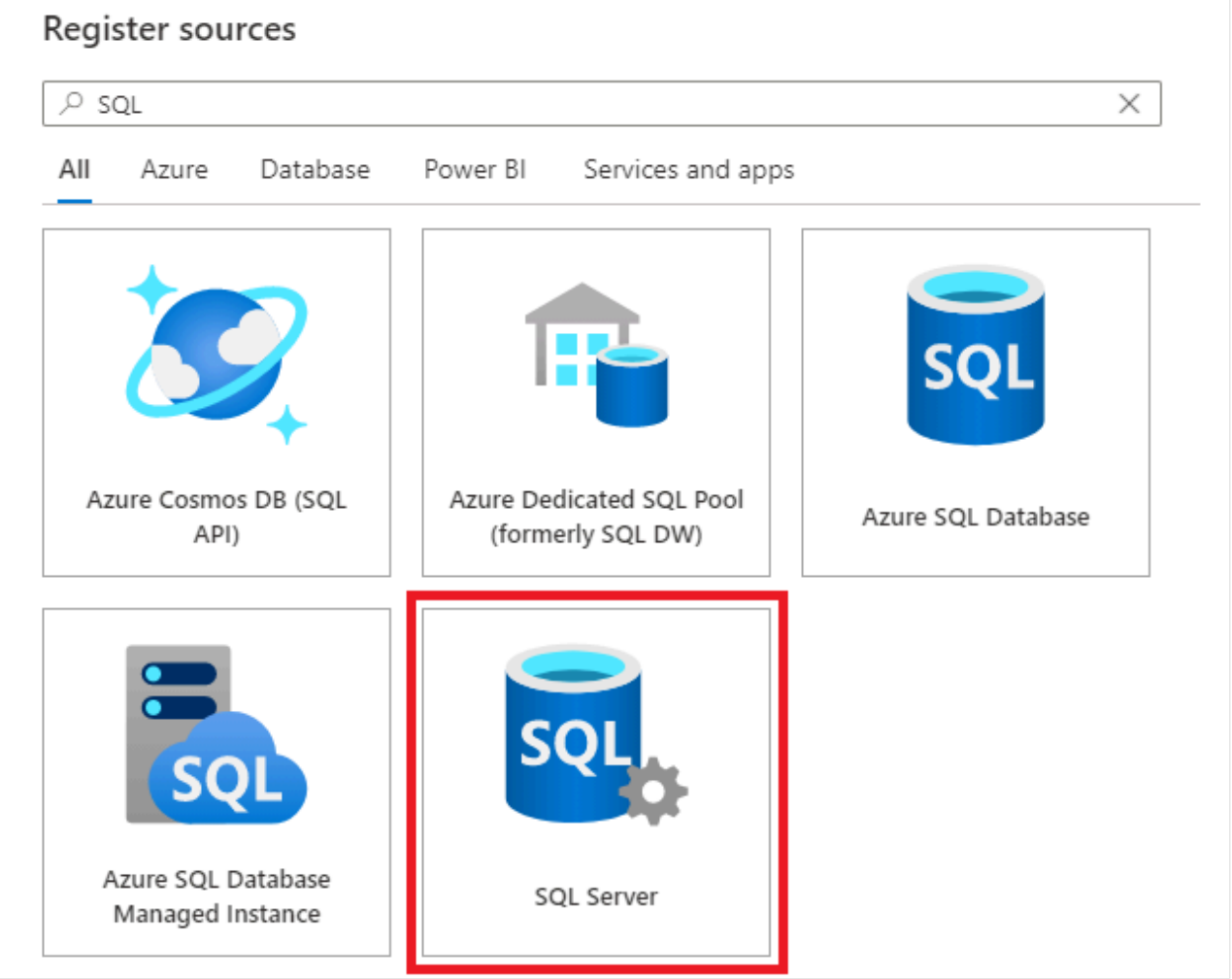
    - If you're using the Microsoft Purview portal, navigate directly to the portal (https://purview.microsoft.com) and select the **Data Map** solution.
    - If you're using the classic Microsoft Purview governance portal, navigate directly to the portal (https://web.purview.azure.com), select your Microsoft Purview account, and select **Data map** from the left menu.

2. Under **Source management**, select **Integration runtimes**. Make sure a self-hosted integration runtime is set up. If it isn't set up, follow the steps mentioned in prerequisites to create a self-hosted integration runtime for scanning from an on-premises or Azure VM that has access to your on-premises network.

3. Select **Data sources** on the left navigation.

4. Select **Register**.

5. Select **SQL Server**, then **Continue9**.



6. Provide a friendly name, which will be a short name you can use to identify your server, and the server endpoint.

7. Select **Finish** to register the data source.

Scan

Follow the steps below to scan on-premises SQL server instances to automatically identify assets and classify your data. For more information about scanning in general, see our introduction to scans and ingestion

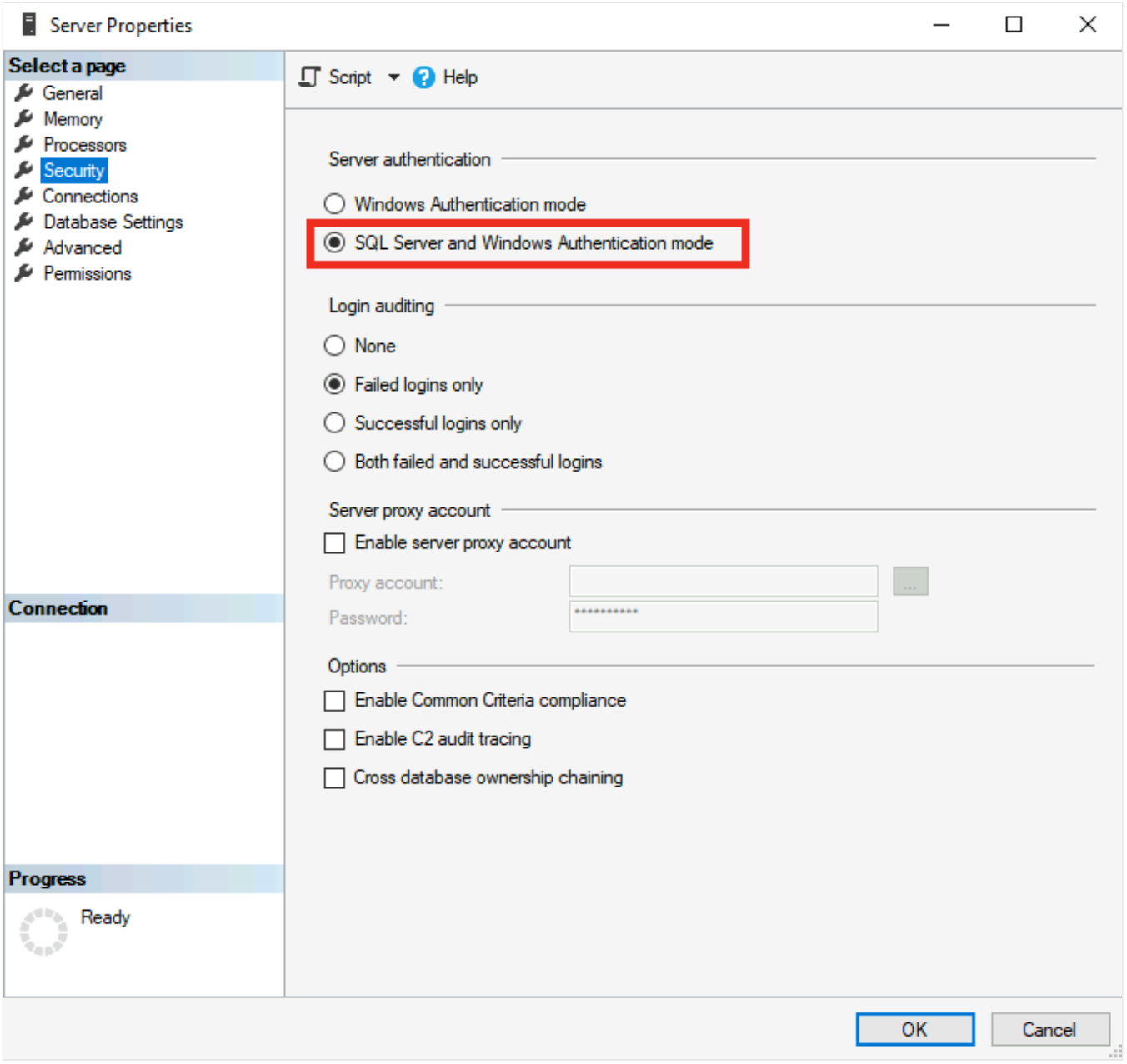
Authentication for scanning

There are two authentication methods supported for SQL server on-premises:

- SQL Authentication
- Windows Authentication - Not supported by a Kubernetes SHIR.

If SQL Authentication is applied, ensure the SQL Server deployment is configured to allow SQL Server and Windows Authentication.

To enable this, within SQL Server Management Studio (SSMS), navigate to **Server Properties** and change from **Windows Authentication Mode** to **SQL Server and Windows Authentication mode**.



If Windows Authentication is applied, configure the SQL Server deployment to use Windows Authentication mode.

A change to the Server Authentication will require a restart of the SQL Server Instance and SQL Server Agent, this can be triggered within SSMS by navigating to the SQL Server instance and selecting "Restart" within the right-click options pane.

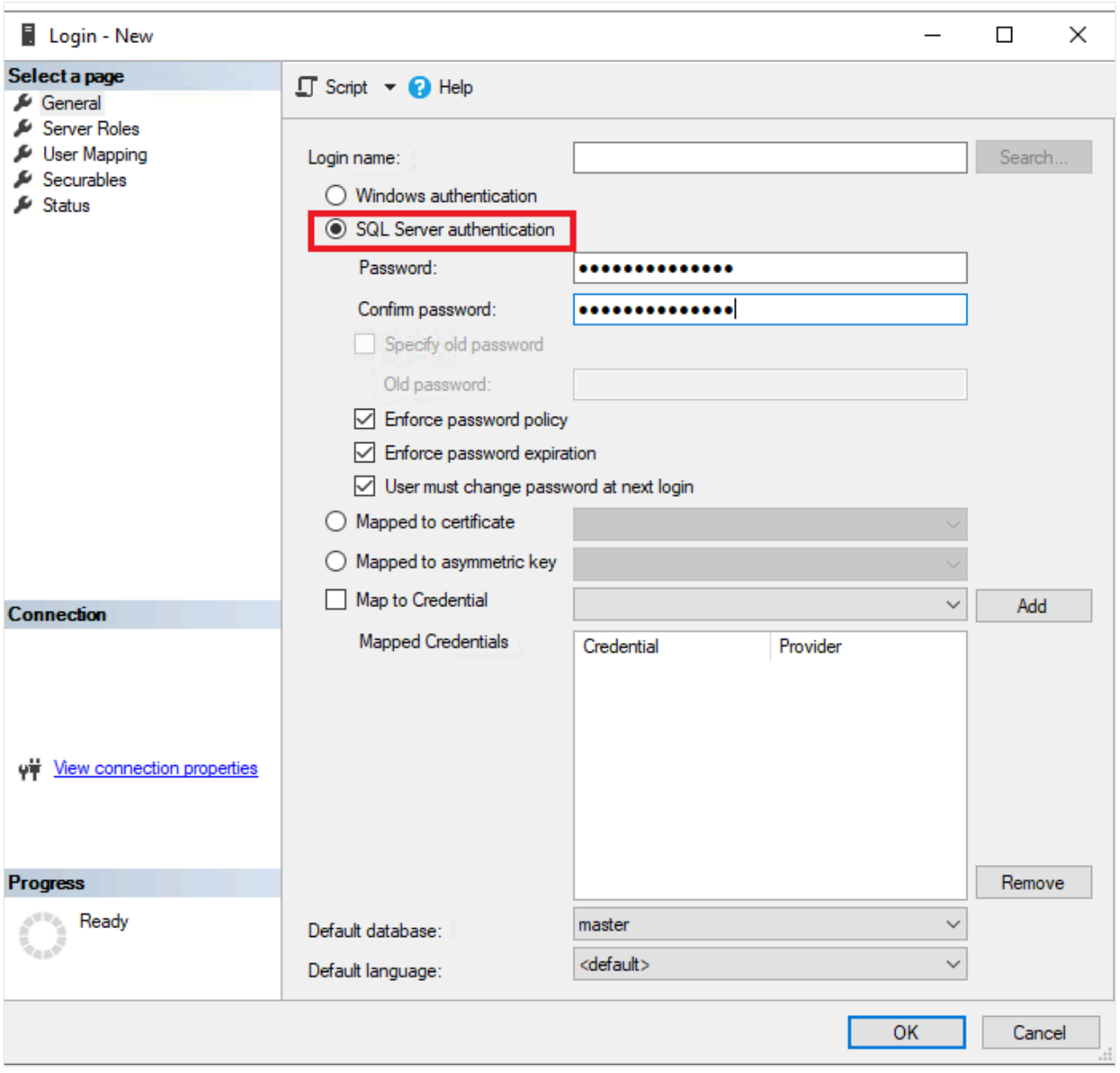
Creating a new login and user

If you would like to create a new login and user to be able to scan your SQL server, follow the steps below:

The account must have access to the **master** database. This is because the `sys.databases` is in the master database. The Microsoft Purview scanner needs to enumerate `sys.databases` in order to find all the SQL databases on the server.

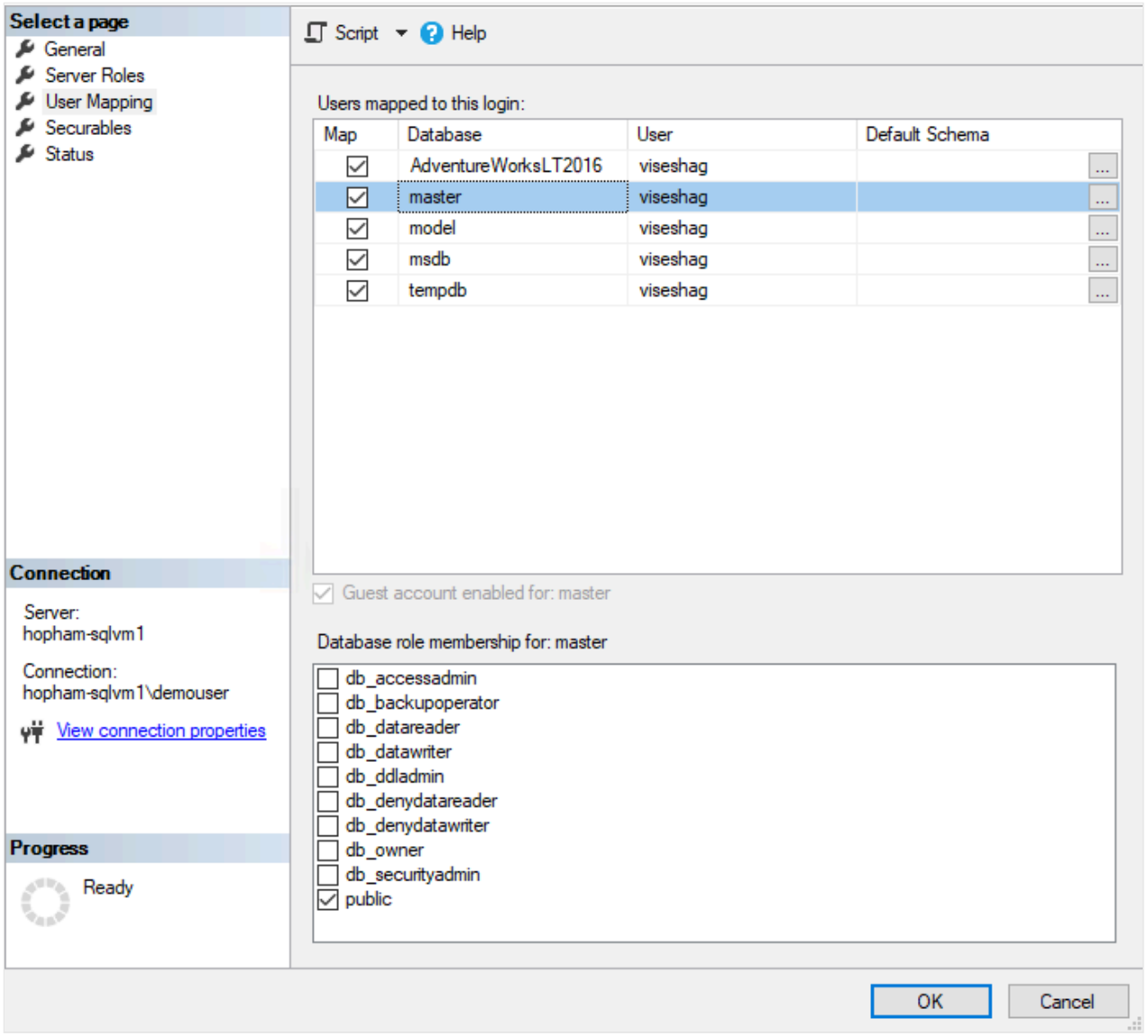
> **Note**
>
> All the steps below can be executed using the code provided here

1. Navigate to SQL Server Management Studio (SSMS), connect to the server, navigate to security, select and hold (or right-click) on login and create New login. If Windows Authentication is applied, select **Windows authentication**. If SQL Authentication is applied, make sure to select **SQL authentication**.



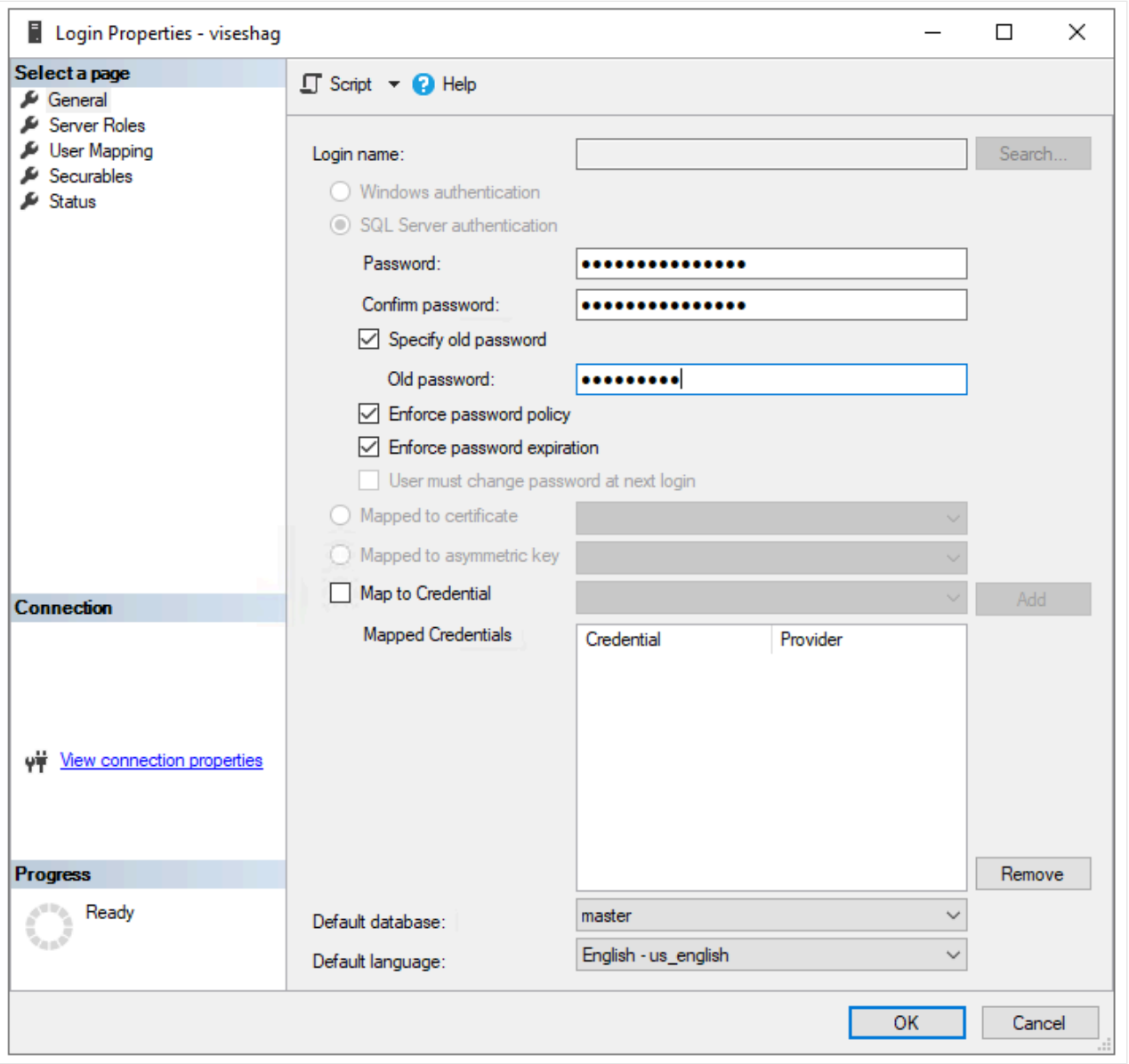
2. Select Server roles on the left navigation and ensure that public role is assigned.

3. Select User mapping on the left navigation, select all the databases in the map and select the Database role: **db_datareader**.



4. Select OK to save.

5. If SQL Authentication is applied, navigate again to the user you created, by selecting and holding (or right-clicking) and selecting **Properties**. Enter a new password and confirm it. Select the 'Specify old password' and enter the old password. **It is required to change your password as soon as you create a new login.**



Storing your SQL login password in a key vault and creating a credential in Microsoft Purview

1. Navigate to your key vault in the Azure portal. Select **Settings** > **Secrets**
2. Select **+ Generate/Import** and enter the **Name** and **Value** as the *password* from your SQL server login
3. Select **Create** to complete
4. If your key vault isn't connected to Microsoft Purview yet, you'll need to create a new key vault connection
5. Finally, create a new credential using the **username** and **password** to set up your scan. Make sure the right authentication method is selected when creating a new credential. If SQL Authentication is applied, select "SQL authentication" as the authentication method. If Windows Authentication is applied, then select **Windows authentication**.

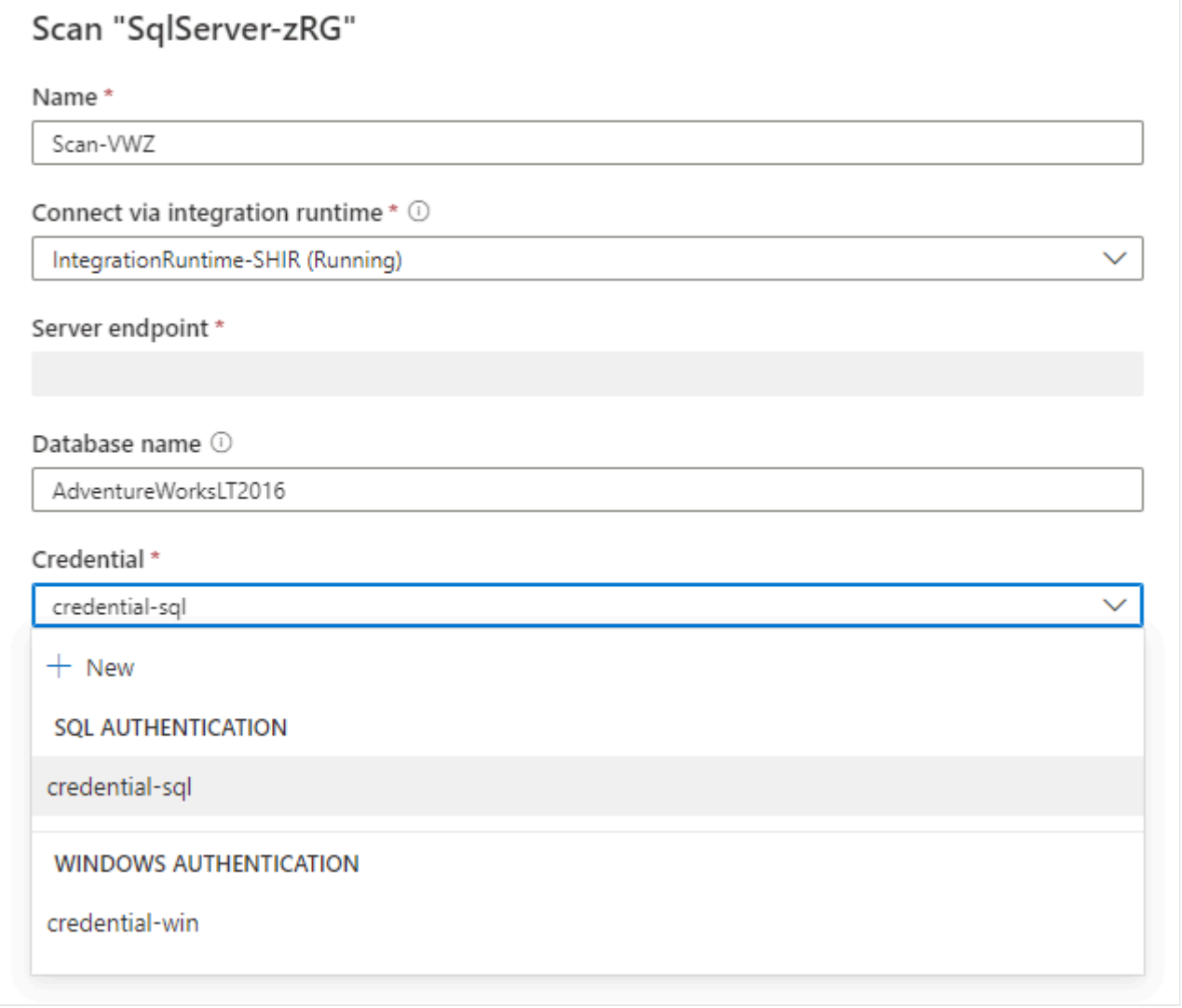
Set up a self-hosted integration runtime

If you haven't already, you need to download and install the latest self-hosted integration runtime. For more information, see the create and configure a self-hosted integration runtime guide.

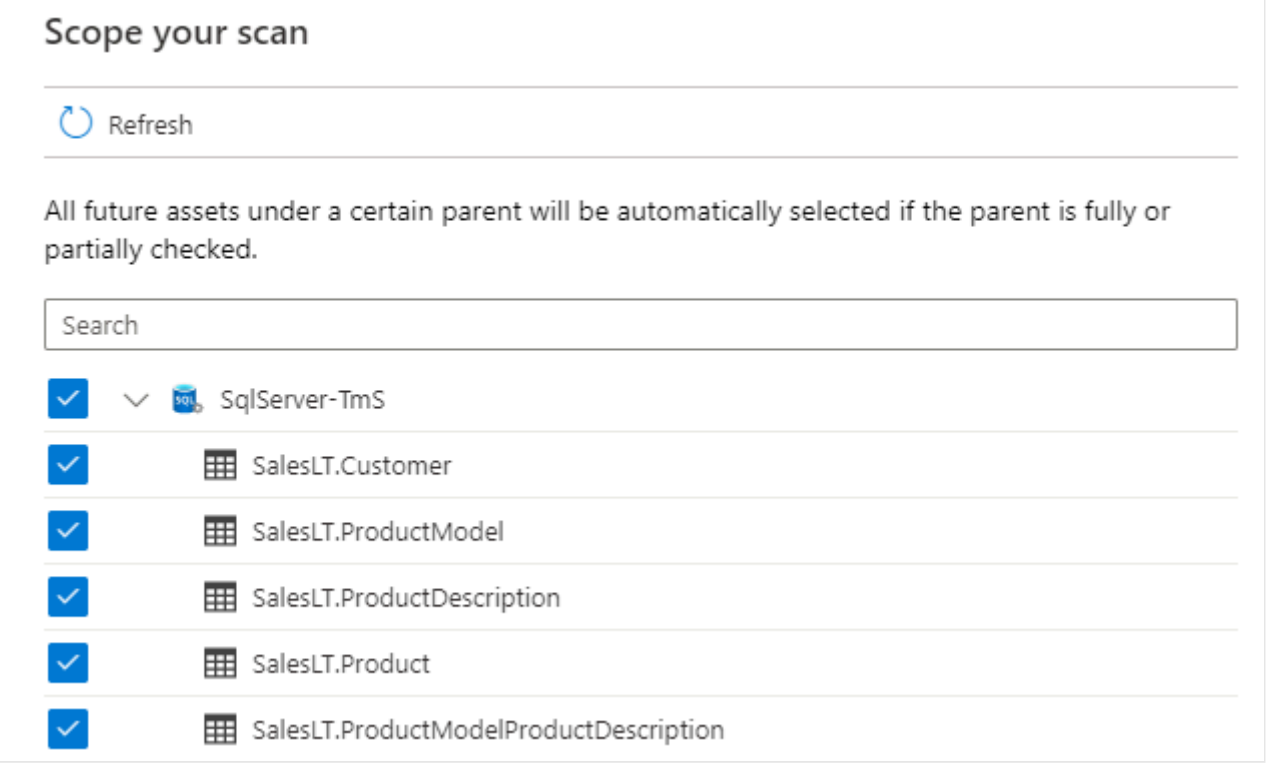
Create and run scan

Now that you've set up authentication and a self-hosted integration

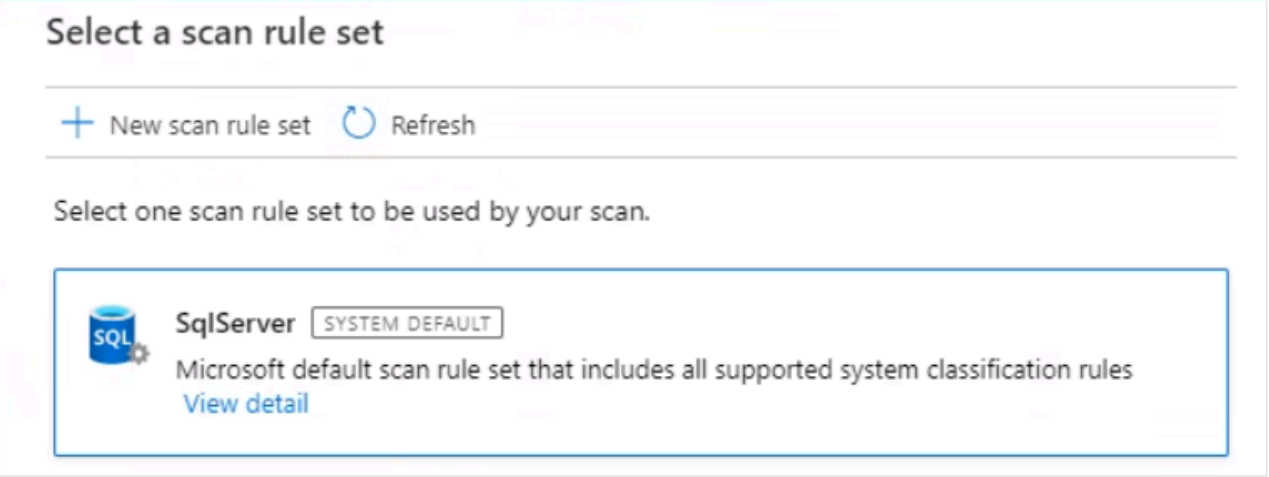
1. Open the Microsoft Purview portal and select the **Data Map**.
2. Select **Data sources**.
3. Select the SQL Server source that you registered.
4. Select **New scan**
5. Select the credential to connect to your data source. The credentials are grouped and listed under different authentication methods.



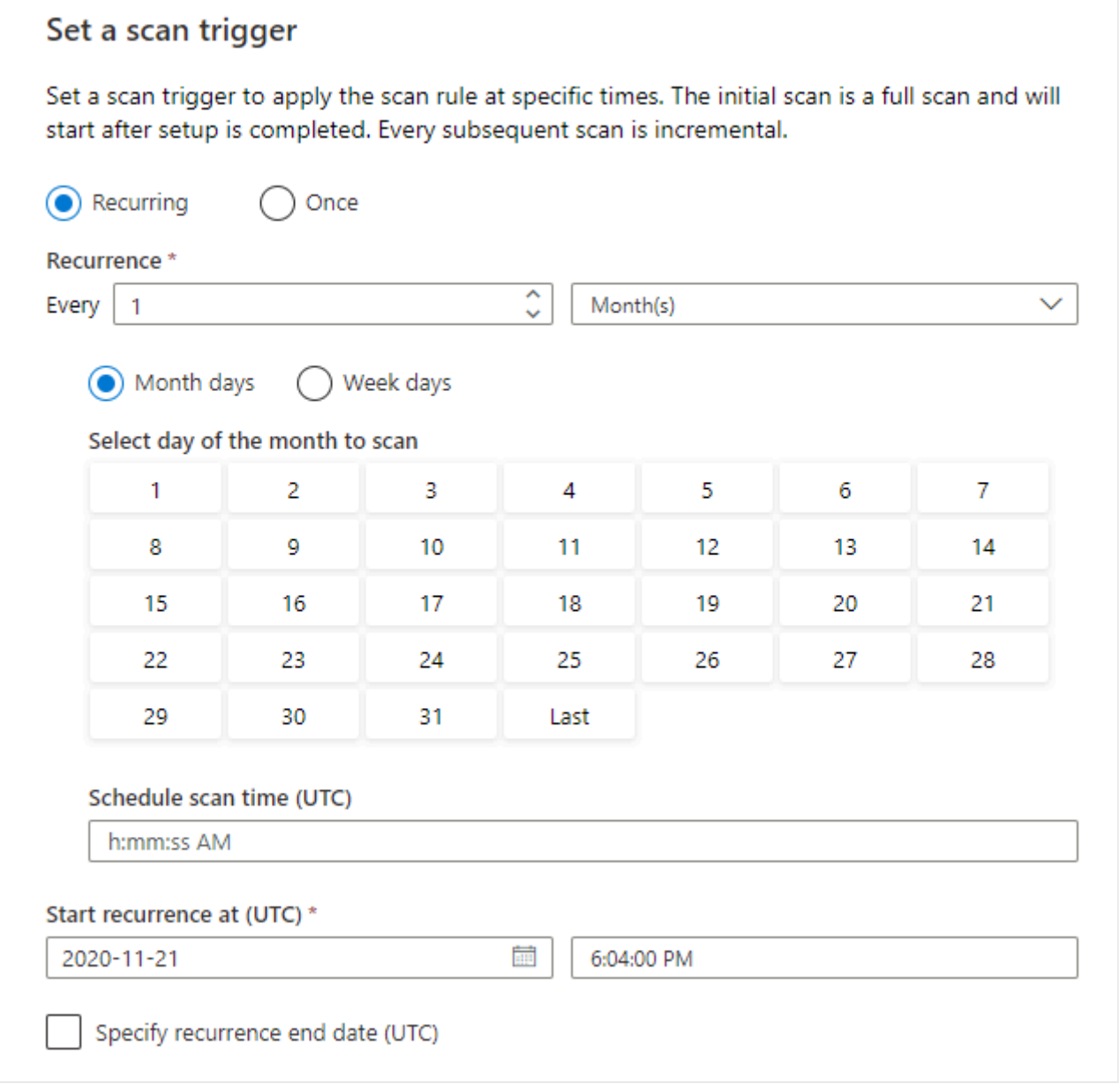
6. You can scope your scan to specific tables by choosing the appropriate items in the list after enter Database name.



7. Then select a scan rule set. You can choose between the system default, existing custom rule sets, or create a new rule set inline.



8. Choose your scan trigger. You can set up a schedule or run the scan once.



9. Review your scan and select **Save and run**.

View your scans and scan runs

To view existing scans:

1. Go to the Microsoft Purview portal. On the left pane, select **Data map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that has results you want to view. The pane shows you all the previous scan runs, along with the status and metrics for each scan run.
4. Select the run ID to check the scan run details.

Manage your scans

To edit, cancel, or delete a scan:

1. Go to the Microsoft Purview portal. On the left pane, select **Data Map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that you want to manage. You can then:
    - Edit the scan by selecting **Edit scan**.
    - Cancel an in-progress scan by selecting **Cancel scan run**.
    - Delete your scan by selecting **Delete scan**.

> **Note**
>
> - Deleting your scan does not delete catalog assets created from previous scans.

Next steps

Now that you have registered your source, follow the below guides to learn more about Microsoft Purview and your data.

- Data Estate Insights in Microsoft Purview
- Lineage in Microsoft Purview
- Search Unified Catalog

Last updated on 03/20/2026

Connect to and manage Oracle in Microsoft Purview

This article outlines how to register Oracle, and how to authenticate and interact with Oracle in Microsoft Purview. For more information about Microsoft Purview, read the introductory article.

Supported capabilities

Scanning capabilities

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan
Yes	Yes	No	Yes

The supported Oracle server versions are 6i to 19c. Oracle proxy server isn't supported when scanning Oracle source.

When scanning Oracle source, Microsoft Purview supports:

- Extracting technical metadata, including:
  - Server
  - Schemas
  - Packages
  - Tables including the columns, foreign keys, indexes, triggers, and unique constraints
  - Views including the columns and triggers
  - Stored procedures including the parameter dataset and result set
  - Functions including the parameter dataset
  - Sequences
  - Synonyms
  - Types including the type attributes
- Fetching static lineage on assets relationships among tables and views.
- Stored procedure lineage.

When setting up scan, you can choose to scan an entire Oracle server, or scope the scan to a subset of schemas matching one or more given names or name patterns.

Other capabilities

For **classifications**, **sensitivity labels**, **policies**, **data lineage**, and **live view**, see the list of supported capabilities.

Known limitations

- Currently, the Oracle service name isn't captured in the metadata or hierarchy.
- When object is deleted from the data source, currently the subsequent scan won't automatically remove the corresponding asset in Microsoft Purview.
- Stored procedures directly under a schema or not within a package aren't supported.

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- The enterprise version of Microsoft Purview or an active Microsoft Purview account using the classic governance portal.
- You need to be a Data Source Administrator and Data Reader to register a source in the Microsoft Purview Data Map. Get details about domain and collection permissions.
- Set up the latest self-hosted integration runtime. For more information, see the create and configure a self-hosted integration runtime guide.
    - Ensure OpenJDK 8 with OJDBC driver 8 is installed on the machine where the self-hosted integration runtime is installed. Restart the machine after you install the JDK for it to take effect.
        - Ensure Visual C++ Redistributable (version Visual Studio 2012 Update 4 or newer) is installed on the self-hosted integration runtime machine. If you don't have this update installed, you can download it here.
        - Download the Oracle JDBC driver on the machine where your self-hosted integration runtime is running. Note down the folder path that you'll use to set up the scan.

          > **Note**

The driver should be accessible by the self-hosted integration runtime. By default, self-hosted integration runtime uses local service account "NT SERVICE\DIAHostService". Make sure it has "Read and execute" and "List folder contents" permission to the driver folder.

Required permissions for scan

Microsoft Purview supports basic authentication (username and password) for scanning Oracle. The Oracle user must have read access to system tables in order to access advanced metadata and classification.

The user should have permission to create a session and role SELECT_CATALOG_ROLE assigned. Alternatively, the user could have SELECT permission granted for every individual system table that this connector queries metadata from:

SQL

```
grant create session to [user];
grant select on all_users to [user];
grant select on dba_objects to [user];
grant select on dba_tab_comments to [user];
grant select on dba_external_locations to [user];
grant select on dba_directories to [user];
grant select on dba_mviews to [user];
grant select on dba_clu_columns to [user];
grant select on dba_tab_columns to [user];
grant select on dba_col_comments to [user];
grant select on dba_constraints to [user];
grant select on dba_cons_columns to [user];
grant select on dba_indexes to [user];
grant select on dba_ind_columns to [user];
grant select on dba_procedures to [user];
grant select on dba_synonyms to [user];
grant select on dba_views to [user];
grant select on dba_source to [user];
grant select on dba_triggers to [user];
grant select on dba_arguments to [user];
grant select on dba_sequences to [user];
grant select on dba_dependencies to [user];
grant select on dba_type_attrs to [user];
grant select on V_$INSTANCE to [user];
grant select on v_$database to [user];
```

For classification, the user must have reader (SELECT) permissions on the relevant tables.

For Individual Tables Use the following SQL command to grant access to a specific table:

SQL

```
GRANT SELECT ON [table_name] TO [purview_user];
```

For All Tables at Once To grant access across all tables in the database, use:

SQL

```
GRANT SELECT ANY TABLE TO [purview_user];
```

Register

This section describes how to register Oracle in Microsoft Purview using the Microsoft Purview portal.

Steps to register

To register a new Oracle source in your data map, do the following:

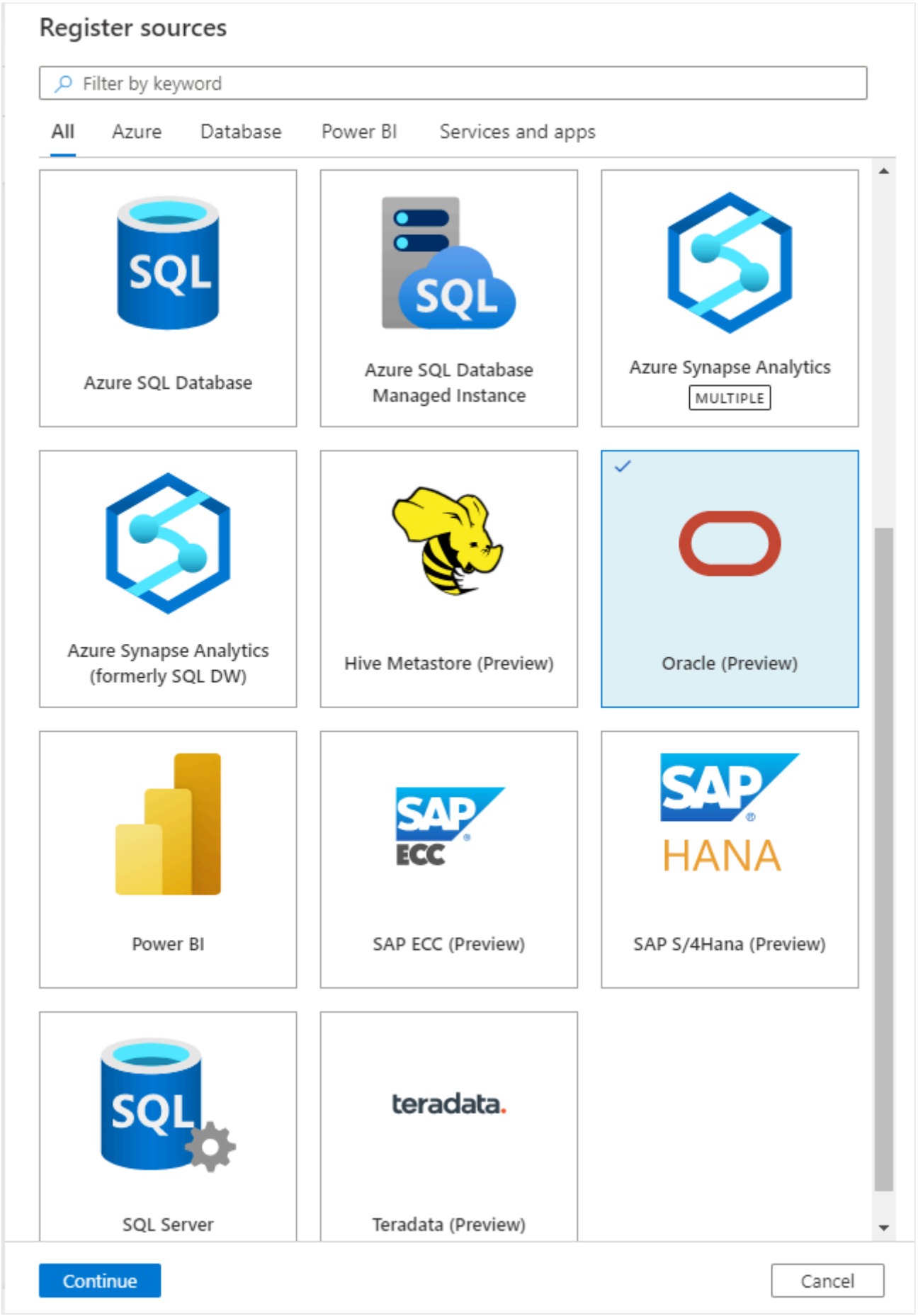
1. Accessing the Microsoft Purview Data Map depends on which portal experience you're using.

    - If you're using the Microsoft Purview portal, navigate directly to the portal (https://purview.microsoft.com) and select the **Data Map** solution.
    - If you're using the classic Microsoft Purview governance portal, navigate directly to the portal (https://web.purview.azure.com), select your Microsoft Purview account, and select **Data map** from the left menu.

2. Select **Data sources**

3. Select **Register**

4. On Register sources, select **Oracle**. Select **Continue**.



On the **Register sources (Oracle)** screen, do the following:

1. Enter a **Name** that the data source will be listed within the Catalog.

2. Enter the **Host** name to connect to an Oracle source. This can either be:

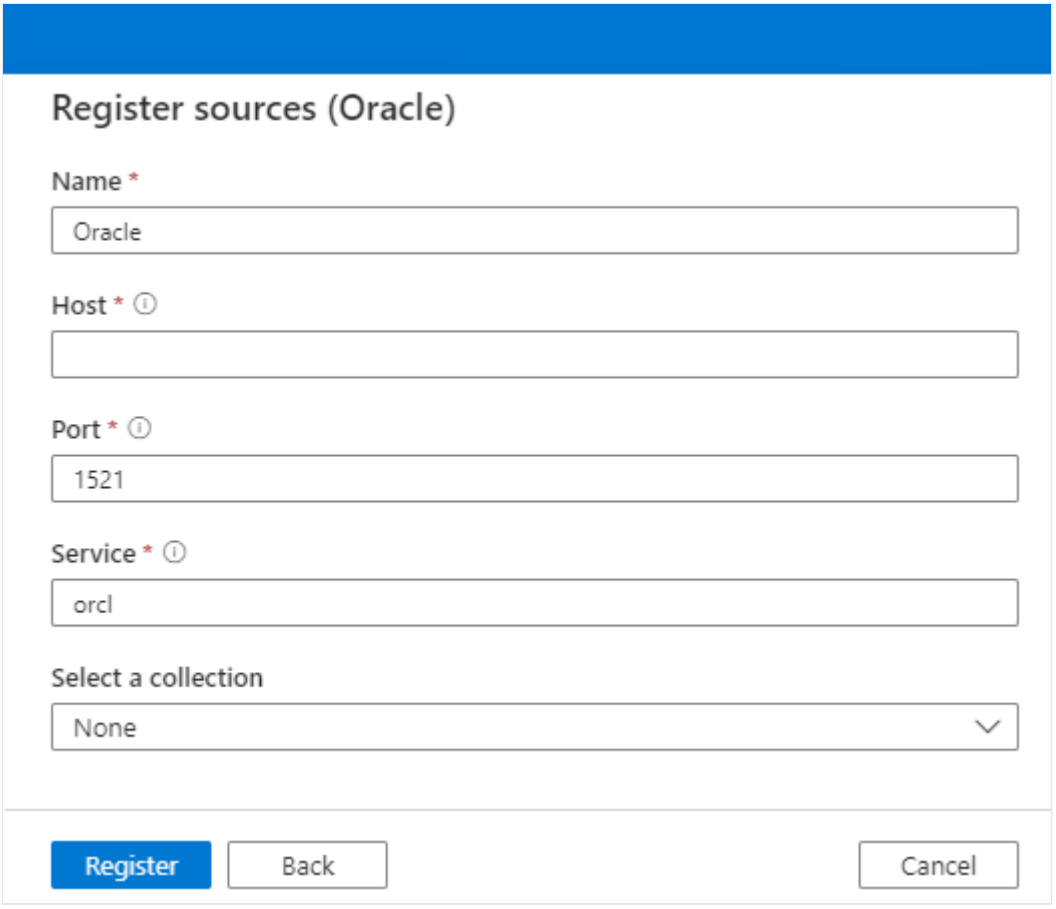
    - A host name used to connect to the database server. For example: `MyDatabaseServer.com`
    - An IP address. For example: `192.169.1.2`

3. Enter the **Port number** used to connect to the database server (1521 by default for Oracle).

4. Enter the **Oracle service name** (not Oracle UID) used to connect to the database server.

5. Select a collection from the list.

6. Finish to register the data source.



Scan

Follow the steps below to scan Oracle and automatically classify assets. For more information about scanning in general, see our introduction to scans and ingestion.

> **Tip**
>
> To troubleshoot any issues with scanning:

1. Confirm you have followed all **prerequisites**.
2. Review our **scan troubleshooting documentation**.

Create and run scan

To create and run a new scan, do the following:

1. Under **Source management** in the data map, select **Integration runtimes**. Make sure a self-hosted integration runtime is set up. If it isn't set up, use the steps mentioned here to create a self-hosted integration runtime.

2. Navigate to **Sources**.

3. Select the registered Oracle source.

4. Select **+ New scan**.

5. Provide the below details:

   a. **Name**: The name of the scan

   b. **Connect via integration runtime**: Select the configured self-hosted integration runtime

   c. **Credential**: Select the credential to connect to your data source. Make sure to:

      - Select Basic Authentication while creating a credential.
      - Provide the user name used to connect to the database server in the User name input field.
      - Store the user password used to connect to the database server in the secret key.

   d. **Schema**: List subset of schemas to import expressed as a semicolon separated list in **case-sensitive** manner. For example, `schema1;schema2`. All user schemas are imported if that list is empty. All system schemas (for example, SysAdmin) and objects are ignored by default.

      Acceptable schema name patterns can be static names or contain wildcard %. For example: `A%;%B;%C%;D`

      - Start with A or
      - End with B or
      - Contain C or
      - Equal D

Usage of NOT and special characters aren't acceptable.

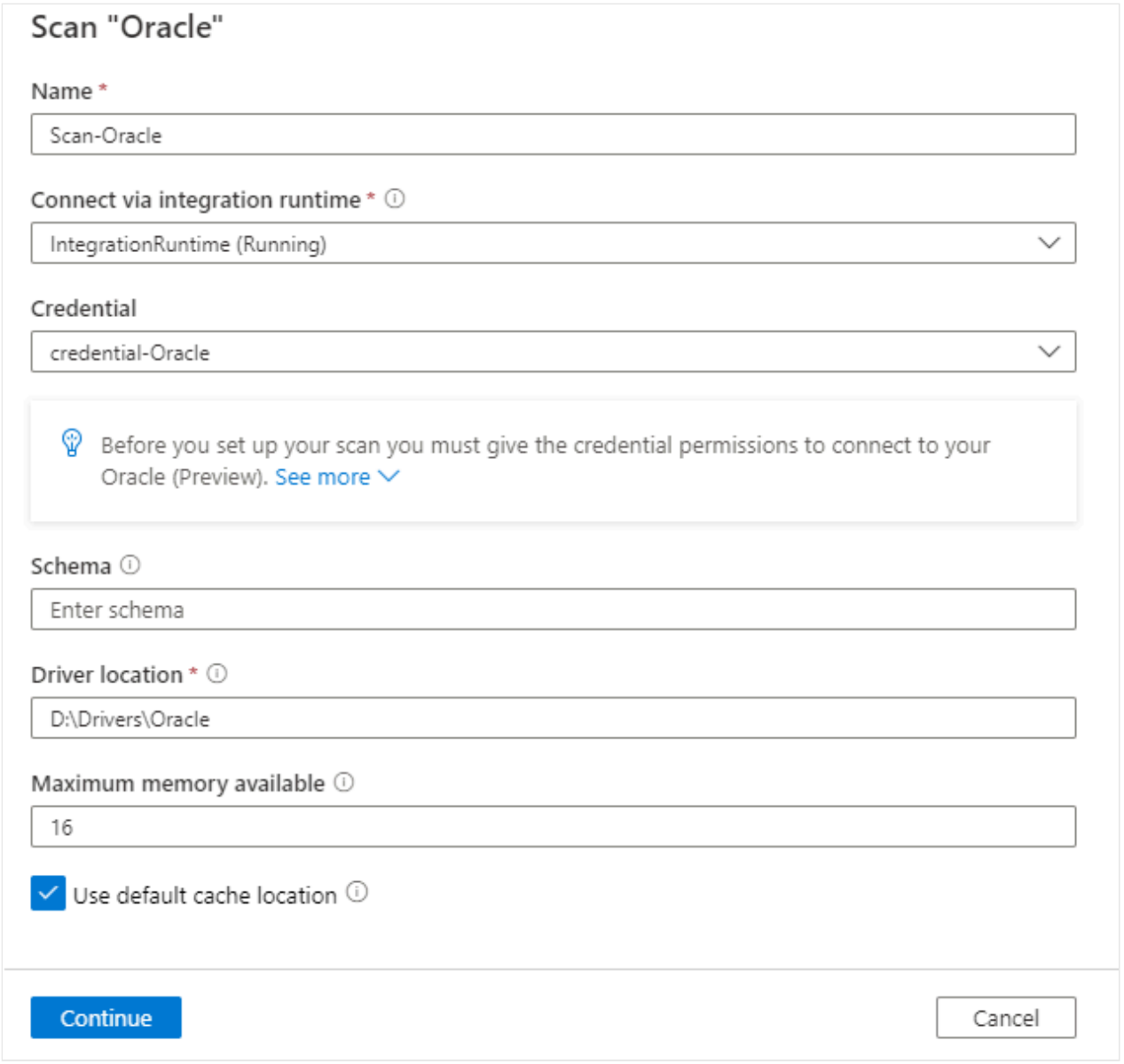
e. **Driver location**: Specify the path to the JDBC driver location in your machine where self-host integration runtime is running. For example: `D:\Drivers\Oracle`.
   i. For self-hosted integration runtime on a local machine: `D:\Drivers\Oracle`. It's the path to valid JAR folder location. The value must be a valid absolute file path and doesn't contain space. Make sure the driver is accessible by the self-hosted integration runtime; learn more at prerequisites.

f. **Stored procedure details**: Controls the number of details imported from stored procedures:

   - Signature: The name and parameters of stored procedures.
   - Code, signature: The name, parameters, and code of stored procedures.
   - Lineage, code, signature: The name, parameters and code of stored procedures, and the data lineage derived from the code.
   - None: Stored procedure details aren't included.

g. **Maximum memory available**: Maximum memory (in GB) available on customer's VM to be used by scanning processes. This is dependent on the size of Oracle source to be scanned.

   > ⓘ **Note**
   >
   > As a rule of thumb, please provide 1GB memory for every 1000 tables



6. Select **Test connection**.

> **Note**
>
> Use the "Test connection" button in the scan setup UI to test the connection. The "Test Connection" in self-hosted integration runtime configuration manager UI -> Diagnostics tab does not fully validate the connectivity.

7. Select **Continue**.

8. Select a **scan rule set** for classification. You can choose between the system default, existing custom rule sets, or create a new rule set inline.

9. Choose your **scan trigger**. You can set up a schedule or ran the scan once.

10. Review your scan and select **Save and Run**.

View your scans and scan runs

To view existing scans:

1. Go to the Microsoft Purview portal. On the left pane, select **Data map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that has results you want to view. The pane shows you all the previous scan runs, along with the status and metrics for each scan run.
4. Select the run ID to check the scan run details.

Manage your scans

To edit, cancel, or delete a scan:

1. Go to the Microsoft Purview portal. On the left pane, select **Data Map**.

2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.

3. Select the scan that you want to manage. You can then:

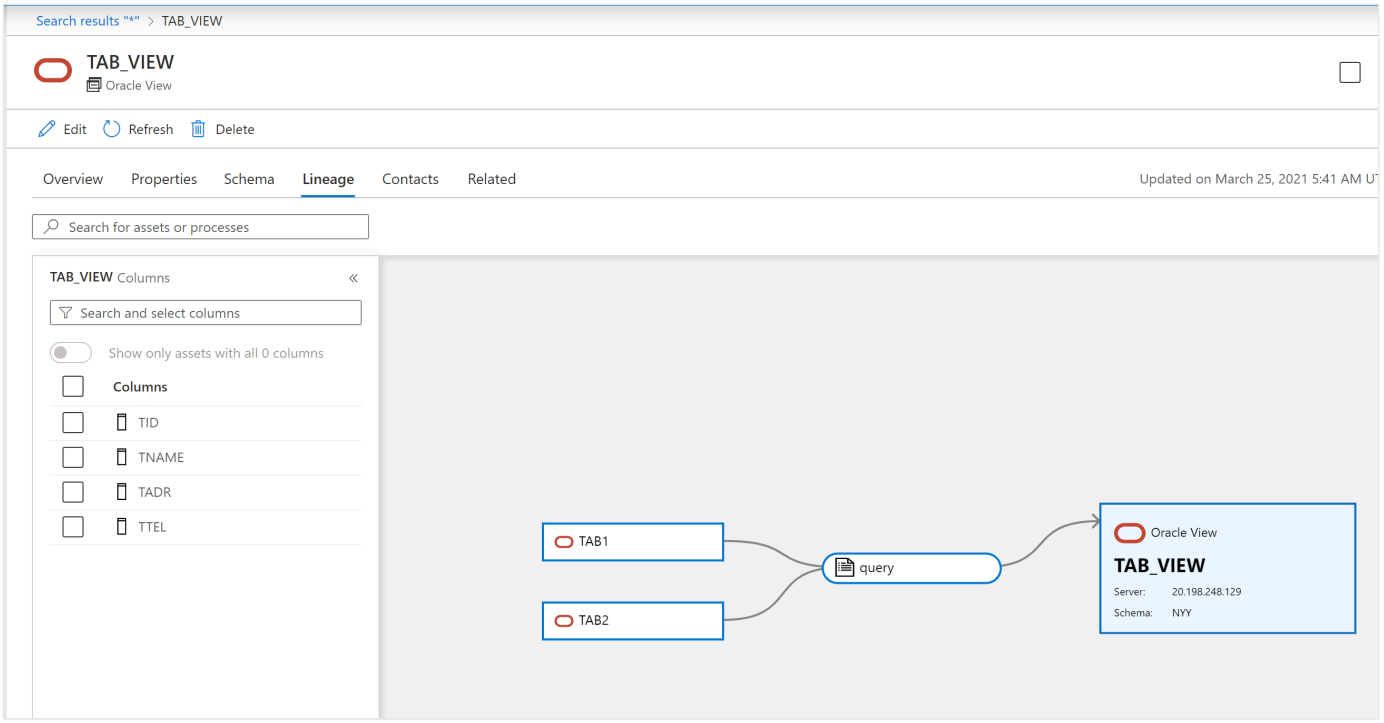
    - Edit the scan by selecting **Edit scan**.
    - Cancel an in-progress scan by selecting **Cancel scan run**.
    - Delete your scan by selecting **Delete scan**.

> ⓘ **Note**
>
> - Deleting your scan does not delete catalog assets created from previous scans.

Lineage

After scanning your Oracle source, you can browse Unified Catalog or search Unified Catalog to view the asset details.

Go to the asset -> lineage tab, you can see the asset relationship when applicable. Refer to the supported capabilities section on the supported Oracle lineage scenarios. For more information about lineage in general, see data lineage and lineage user guide.



Next steps

Now that you've registered your source, follow the below guides to learn more about Microsoft Purview and your data.

- Data Estate Insights in Microsoft Purview
- Lineage in Microsoft Purview
- Search Unified Catalog

Last updated on 04/20/2026

Connect to and manage PostgreSQL in Microsoft Purview

This article outlines how to register PostgreSQL, and how to authenticate and interact with PostgreSQL in Microsoft Purview. For more information about Microsoft Purview, read the introductory article.

Supported capabilities

Scanning capabilities

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan
Yes	Yes	No	Yes

The supported PostgreSQL server versions are 8.x to 16.x.

When scanning PostgreSQL source, Microsoft Purview supports:

- Extracting technical metadata including:
  - Server
  - Databases
  - Schemas
  - Tables including the columns
  - Views including the columns
- Fetching static lineage on assets relationships among tables and views.

When setting up scan, you can choose to scan an entire PostgreSQL database, or scope the scan to a subset of schemas matching the given name(s) or name pattern(s).

Other capabilities

For **classifications**, **sensitivity labels**, **policies**, **data lineage**, and **live view**, see the list of supported capabilities.

Known limitations

When object is deleted from the data source, currently the subsequent scan won't automatically remove the corresponding asset in Microsoft Purview.

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- An active Microsoft Purview account.
- You need Data Source Administrator and Data Reader permissions to register a source and manage it in the Microsoft Purview governance portal. For more information about permissions, see Access control in Microsoft Purview.

> **Note**
>
> **If your data store is not publicly accessible** (if your data store limits access from on-premises network, private network or specific IPs, etc.), **you will need to configure a self hosted integration runtime to connect to it**.

- If your data source isn't publicly accessible, set up the latest self-hosted integration runtime.
  - Choose the right integration runtime for your scenario:
    - **To use a self-hosted integration runtime:**
      - Follow the article to create and configure a self-hosted integration runtime.
      - Ensure JDK 11 is installed on the machine where the self-hosted integration runtime is installed. Restart the machine after you newly install the JDK for it to take effect.
      - Ensure that Visual C++ Redistributable (version Visual Studio 2012 Update 4 or newer) is installed on the machine where the self-hosted integration runtime is running. If you don't have this update installed, download it now.
    - **To use a kubernetes supported self-hosted integration runtime:**
      - Follow the article to create and configure a kubernetes supported integration runtime.

Required permissions for scan

- The PostgreSQL user needs read access to system tables in order to access advanced metadata.
- The user must be granted CONNECT permission on the default postgres database (Command: GRANT CONNECT ON DATABASE postgres TO [user];).

Register

This section describes how to register PostgreSQL in Microsoft Purview using the Microsoft Purview governance portal.

Steps to register

To register a new PostgreSQL source in Microsoft Purview Unified Catalog, follow these steps:

1. Navigate to your Microsoft Purview account in the Microsoft Purview governance portal.
2. Select **Data Map** on the left navigation.
3. Select **Register**
4. On Register sources, select **PostgreSQL**. Select **Continue**.

On the **Register sources (PostgreSQL)** screen, follow these steps:

1. Enter a **Name** that the data source will be listed within the Catalog.

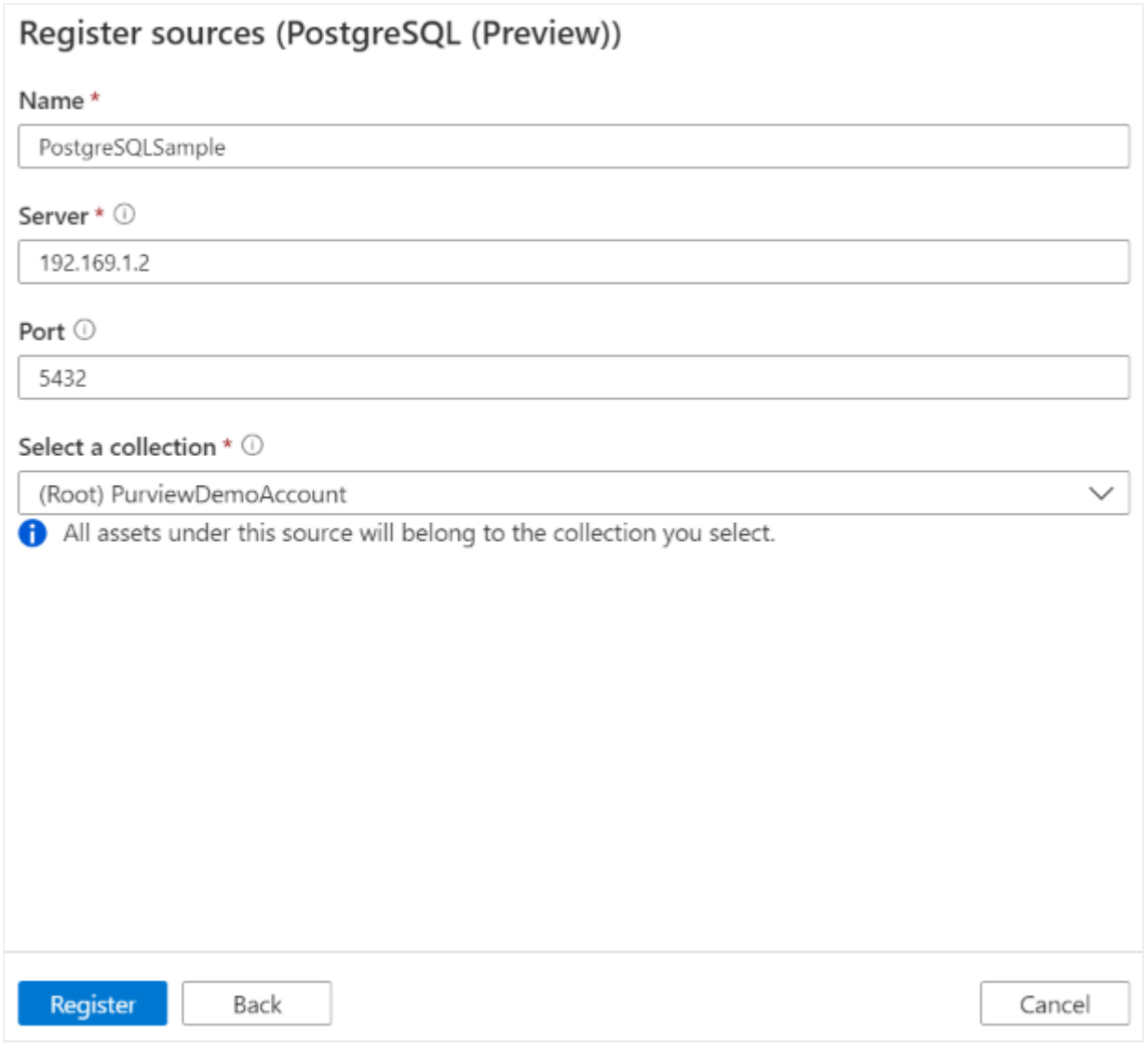
2. Enter the **Server** name to connect to a PostgreSQL source. This can either be:

    - A host name used to connect to the database server. For example: `MyDatabaseServer.com`
    - An IP address. For example: `192.169.1.2`

3. Enter the **Port** used to connect to the database server (5432 by default for PostgreSQL).

4. Select a collection from the list.

5. Finish to register the data source.



Scan

Follow the steps below to scan PostgreSQL to automatically identify assets. For more information about scanning in general, see our introduction to scans and ingestion.

Authentication for a scan

The supported authentication type for a PostgreSQL source is **Basic authentication**.

Create and run scan

To create and run a new scan, follow these steps:

1. If your server is publicly accessible, skip to step two. Otherwise, you'll need to make sure your self-hosted integration runtime is configured:
   a. In the Microsoft Purview governance portal, got to the Management Center, and select **Integration runtimes**.

b. Make sure a self-hosted integration runtime is available. If one isn't set up, use the steps mentioned in prerequisites to set up a self-hosted integration runtime.

2. In the Microsoft Purview governance portal, navigate to **Sources**.

3. Select the registered PostgreSQL source.

4. Select + **New scan**.

5. Provide the below details:

   a. **Name**: The name of the scan

   b. **Connect via integration runtime**: Select the Azure auto-resolved integration runtime if your server is publicly accessible, or your configured self-hosted integration runtime if it isn't publicly available.

   c. **Credential**: Select the credential to connect to your data source. Make sure to:

      - Select **Basic Authentication** while creating a credential.
      - Provide the user name used to connect to the database server in the User name input field.
      - Store the user password used to connect to the database server in the secret key.

   d. **Database**: Specify the name of the database instance to import.

   e. **Schema**: List subset of schemas to import expressed as a semicolon separated list. For example, `schema1;schema2`. All user schemas are imported if that list is empty. All system schemas (for example, SysAdmin) and objects are ignored by default.

      Acceptable schema name patterns can be static names or contain wildcard %. For example: `A%;%B;%C%;D`

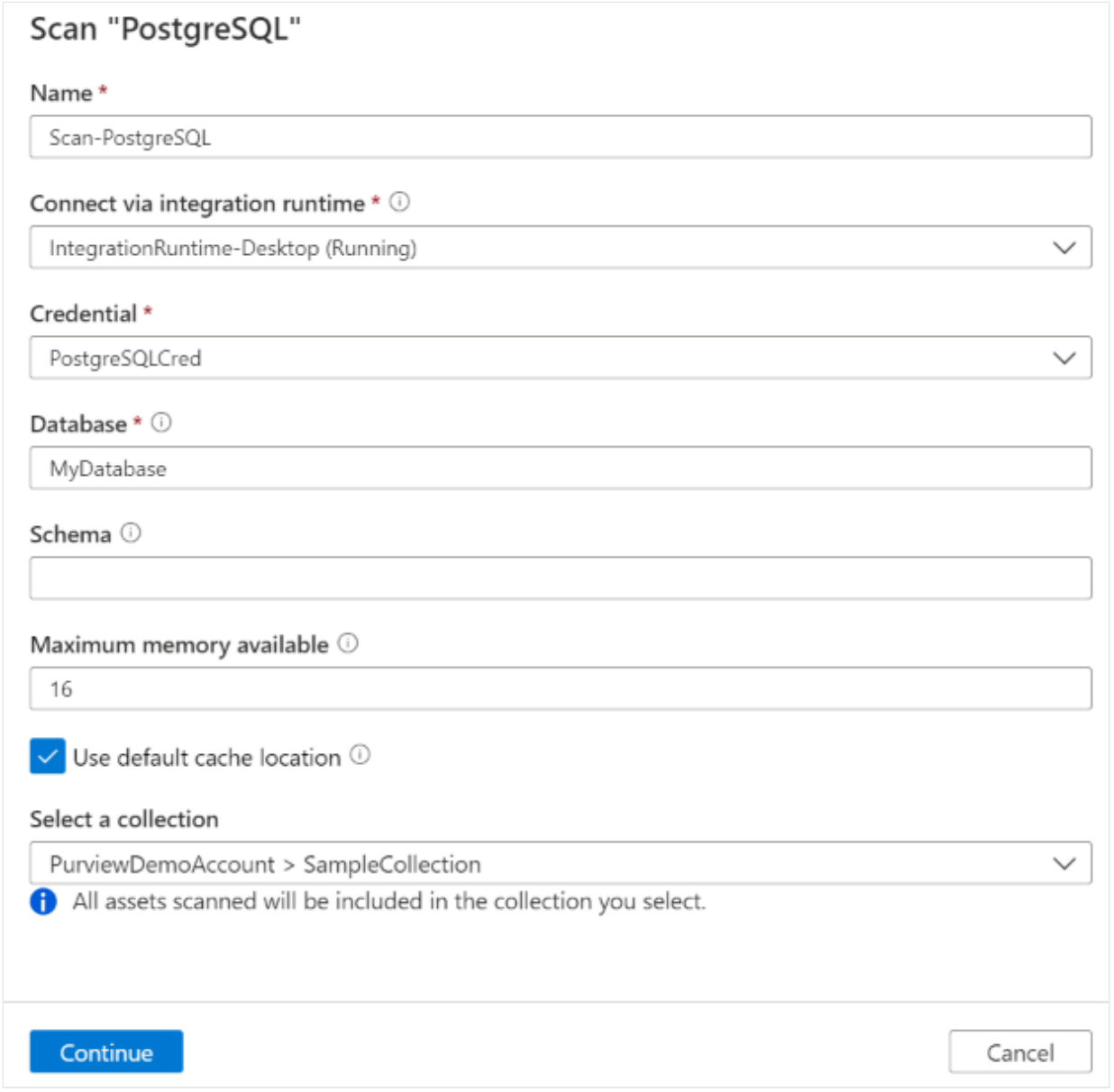
      - Start with A or
      - End with B or
      - Contain C or
      - Equal D

      Usage of NOT and special characters aren't acceptable.

   f. **Maximum memory available** (applicable when using self-hosted integration runtime): Maximum memory (in GB) available on customer's VM to be used by scanning processes. This is dependent on the size of PostgreSQL source to be scanned.

      > ⓘ **Note**

As a rule of thumb, please provide 1GB memory for every 1000 tables



6. Select **Test connection** to validate the settings (available when using Azure Integration Runtime).

7. Select **Continue**.

8. Choose your **scan trigger**. You can set up a schedule or ran the scan once.

9. Review your scan and select **Save and Run**.

View your scans and scan runs

To view existing scans:

1. Go to the Microsoft Purview portal. On the left pane, select **Data map**.

2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that has results you want to view. The pane shows you all the previous scan runs, along with the status and metrics for each scan run.
4. Select the run ID to check the scan run details.

Manage your scans

To edit, cancel, or delete a scan:

1. Go to the Microsoft Purview portal. On the left pane, select **Data Map**.

2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.

3. Select the scan that you want to manage. You can then:

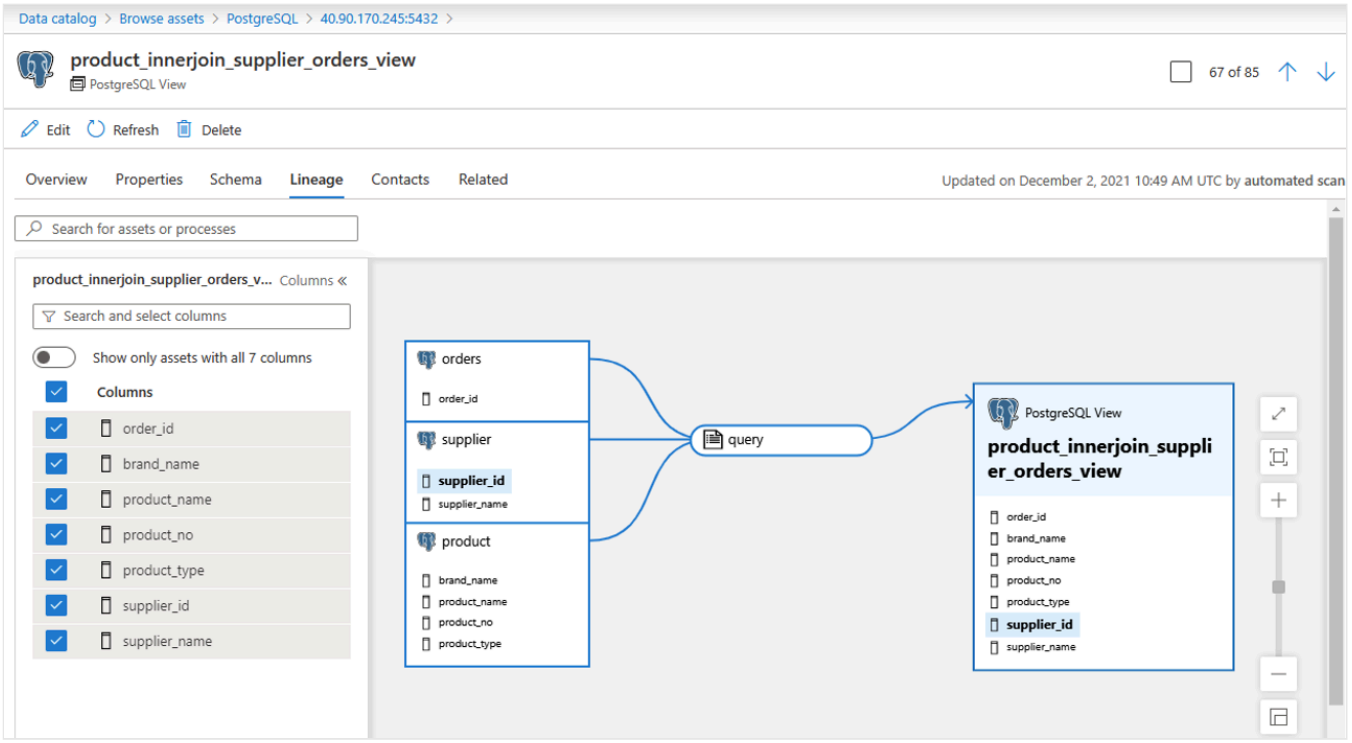
    - Edit the scan by selecting **Edit scan**.
    - Cancel an in-progress scan by selecting **Cancel scan run**.
    - Delete your scan by selecting **Delete scan**.

> ⓘ **Note**
>
> - Deleting your scan does not delete catalog assets created from previous scans.

Lineage

After scanning your PostgreSQL source, you can browse Unified Catalog or search Unified Catalog to view the asset details.

Go to the asset -> lineage tab, you can see the asset relationship when applicable. Refer to the supported capabilities section on the supported PostgreSQL lineage scenarios. For more information about lineage in general, see data lineage and lineage user guide.



Next steps

Now that you've registered your source, follow the below guides to learn more about Microsoft Purview and your data.

- Data Estate Insights in Microsoft Purview
- Lineage in Microsoft Purview
- Search Unified Catalog

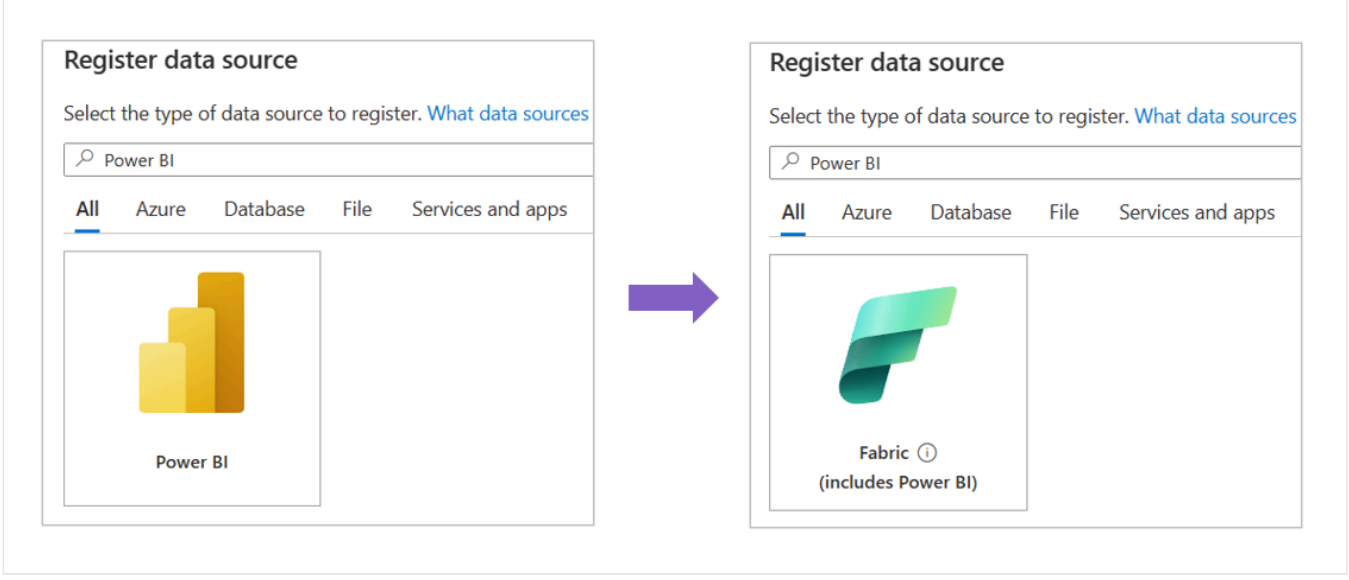
Last updated on 01/22/2026

Connect to and manage a Power BI tenant in Microsoft Purview (same tenant)

Article • 05/21/2025

This article outlines how to register a Power BI tenant in a **same-tenant scenario**, and how to authenticate and interact with the tenant in Microsoft Purview. For more information about Microsoft Purview, read the introductory article.

> **Note**
>
> The Power BI data source in Microsoft Purview has been renamed to Fabric in all regions. You can still use "Power BI" as the keyword to quickly locate the source to register a Power BI tenant and set up a scan.
>
> Starting from December 13th 2023, scanning Fabric tenants registered with the Fabric data source in Microsoft Purview will capture metadata and lineage from Fabric items including Power BI. There are no other configuration steps needed to enable scanning Fabric items besides Power BI for existing scans. Fabric tenants and Power BI tenants will share the same data source and same experience to set up scans. For scanning a Fabric tenant, see our Fabric documentation.



Supported capabilities

Scanning capabilities

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan
Yes	Yes	Yes	Yes

When scanning Power BI source, Microsoft Purview supports:

- Extracting technical metadata including:
  - Workspaces
  - Dashboards
  - Reports
  - Datasets including the tables and columns
  - Dataflows
  - Datamarts
- Fetching static lineage on assets relationships among above Power BI artifacts and external data source assets. Learn more from Power BI lineage.

See a list of metadata available for Power BI.

Other capabilities

For **classifications**, **sensitivity labels**, **policies**, **data lineage**, and **live view**, see the list of supported capabilities.

Supported scenarios for Power BI scans

Expand table

Scenarios	Microsoft Purview public access allowed/denied	Power BI public access allowed /denied	Runtime option	Authentication option	Deployment checklist
Public access with Azure IR	Allowed	Allowed	Azure Runtime	Managed Identity / Delegated authentication / Service principal	Review deployment checklist
Public access with Self-hosted IR	Allowed	Allowed	Self-hosted runtime	Delegated authentication / Service principal	Review deployment checklist

Scenarios	Microsoft Purview public access allowed/denied	Power BI public access allowed /denied	Runtime option	Authentication option	Deployment checklist
Private access	Denied	Allowed	Managed VNet IR (v2 only)	Managed Identity / Delegated authentication / Service principal	Review deployment checklist
Private access	Allowed	Denied	Self-hosted runtime	Delegated authentication / Service principal	Review deployment checklist
Private access	Denied	Allowed	Self-hosted runtime	Delegated authentication / Service principal	Review deployment checklist
Private access	Denied	Denied	Self-hosted runtime	Delegated authentication / Service principal	Review deployment checklist

Known limitations

- If Power BI tenant is protected behind a private endpoint, standard or kubernetes supported self-hosted runtimes are the only options to scan.
- Delegated authentication and service principal are the only supported authentication options when a self-hosted integration runtime is used during the scan.
- If Power BI dataset schema isn't shown after scan, it's due to one of the current limitations with Power BI Metadata scanner.
- Empty workspaces are skipped.
- Other limits please refer to Microsoft Purview limits.

Prerequisites

Before you start, make sure you have the following prerequisites:

- An Azure account with an active subscription. Create an account for free.
- An active Microsoft Purview account.

Authentication options

- Managed Identity
- Delegated Authentication
- Service Principal

Deployment checklist

Use any of the following deployment checklists during the setup or for troubleshooting purposes, based on your scenario:

Public access with Azure IR

Scan same-tenant Power BI using Azure IR and Managed Identity in public network

1. Make sure Power BI and Microsoft Purview accounts are in the same tenant.
2. Make sure Power BI tenant ID is entered correctly during the registration.
3. Make sure your Power BI Metadata model is up to date by enabling metadata scanning.
4. From Azure portal, validate if Microsoft Purview account Network is set to public access.
5. From Power BI tenant Admin Portal, make sure Power BI tenant is configured to allow public network.
6. In Microsoft Entra tenant, create a security group.
7. From Microsoft Entra tenant, make sure Microsoft Purview account MSI is member of the new security group.
8. On the Power BI Tenant Admin portal, validate if Allow service principals to use read-only Power BI admin APIs is enabled for the new security group.

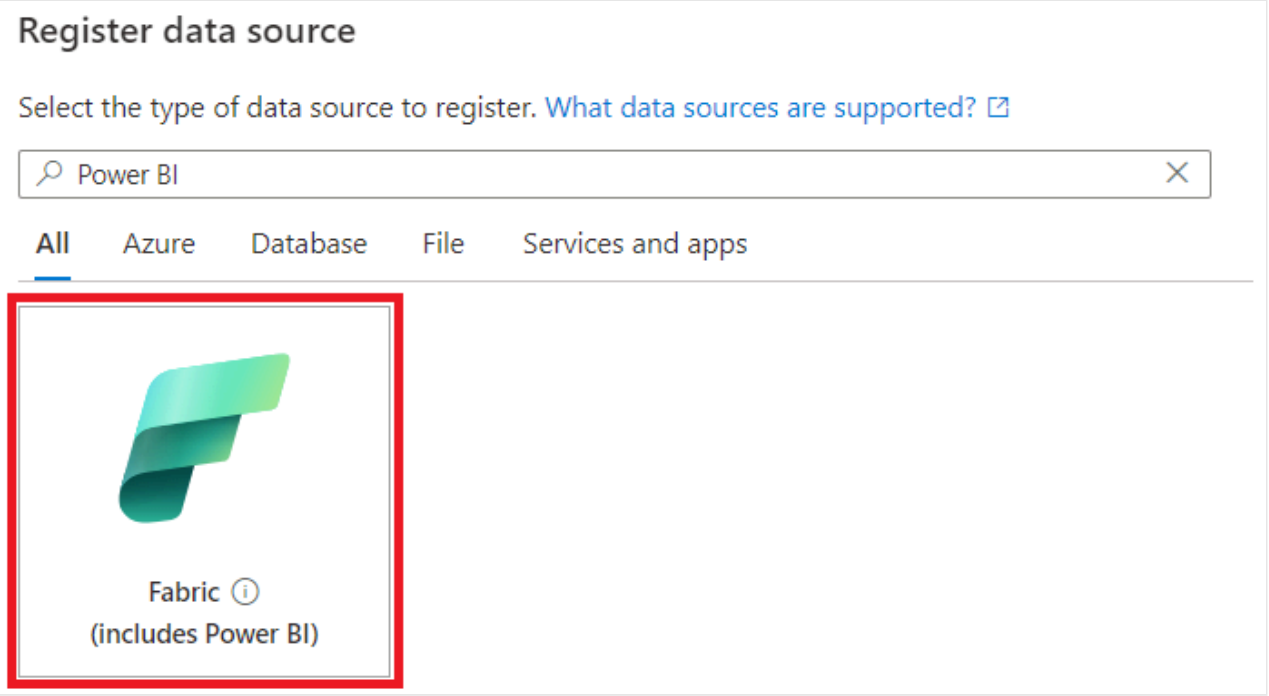
Register Power BI tenant

This section describes how to register a Power BI tenant in Microsoft Purview for same-tenant scenario.

1. Select the **Data Map** on the left navigation.

2. Then select **Register**.

   Select **Fabric** as your data source. It includes Power BI sources and assets.



3. Give your data source a name.

   The name must be between 3-63 characters long and must contain only letters, numbers, underscores, and hyphens. Spaces aren't allowed.

   By default, the system will find the Fabric tenant that exists in the same Microsoft Entra tenant.

4. Select the collection where you want to register the source.

5. Select **Register**.

Scan same-tenant Power BI

> **Tip**
>
> To troubleshoot any issues with scanning:
>
> 1. Confirm you have completed the deployment checklist for your scenario.
> 2. Review our scan troubleshooting documentation.

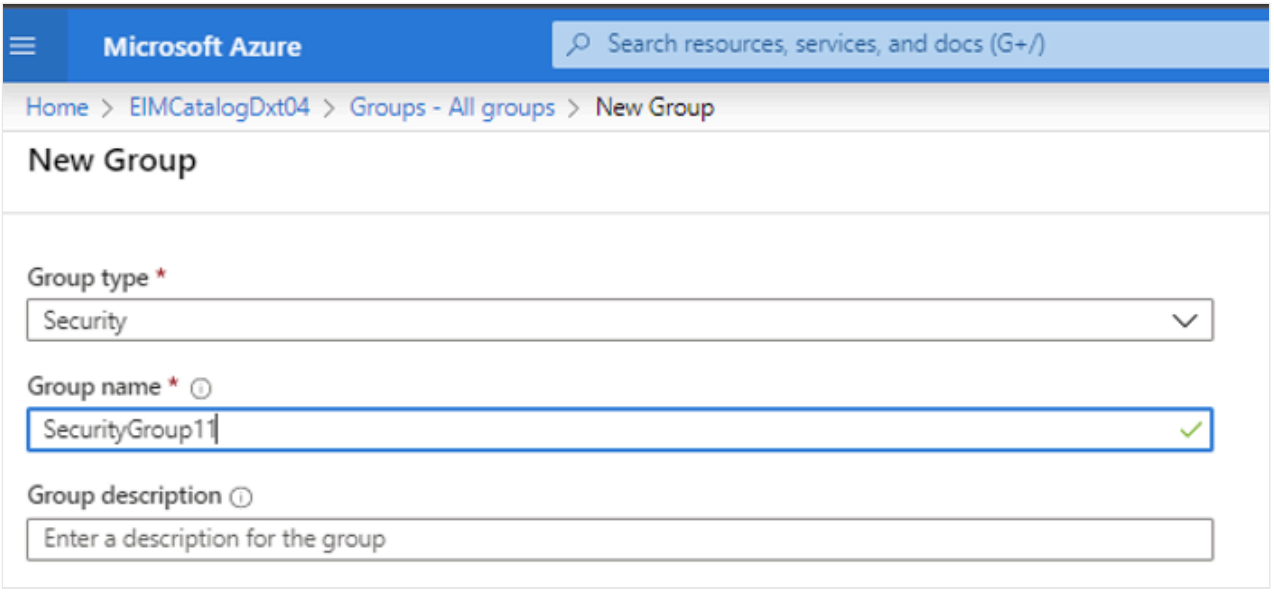
Authenticate to Power BI tenant

In Microsoft Entra tenant, where Power BI tenant is located:

1. In the Azure portal, search for **Microsoft Entra ID**.
2. Create a new security group in your Microsoft Entra ID, by following Create a basic group and add members using Microsoft Entra ID.

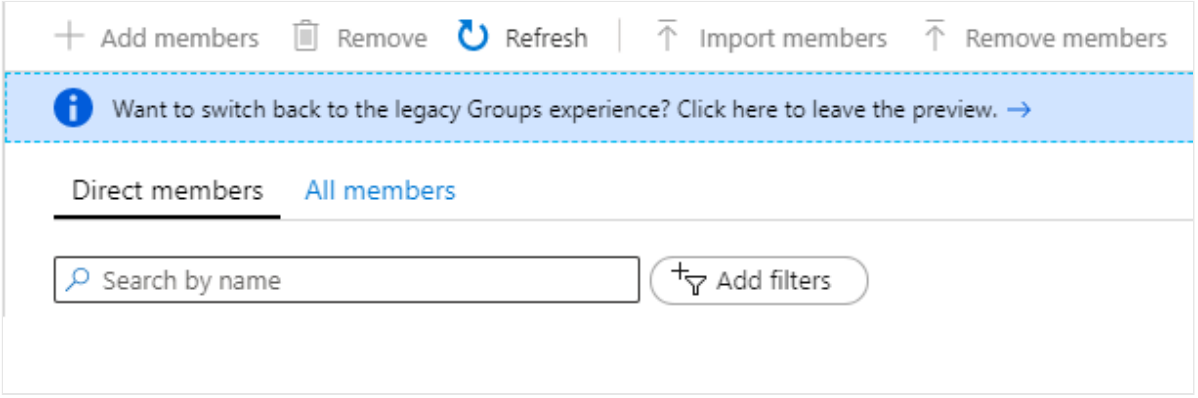
   > **Tip**
   >
   > You can skip this step if you already have a security group you want to use.

3. Select **Security** as the **Group Type**.



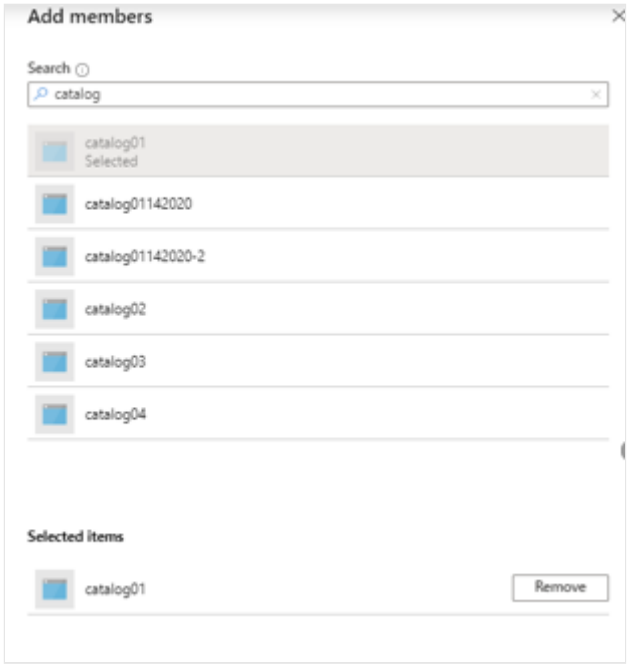
4. Add relevant user to the security group:

   - If you're using **Managed Identity** as authentication method, add your Microsoft Purview managed identity to this security group. Select **Members**, then select **+ Add members**.

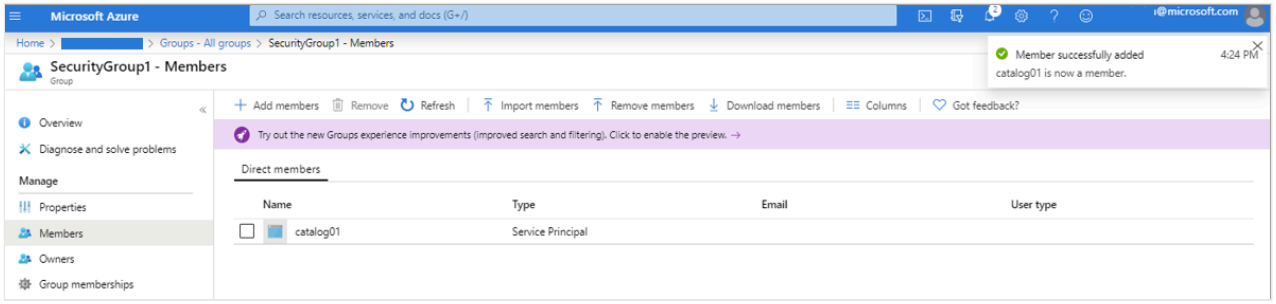


- If you're using **delegated authentication** or **service principal** as authentication method, add your **service principal** to this security group. Select **Members**, then select **+ Add members**.

5. Search for your Microsoft Purview managed identity or service principal and select it.



You should see a success notification showing you that it was added.



Associate the security group with Power BI tenant

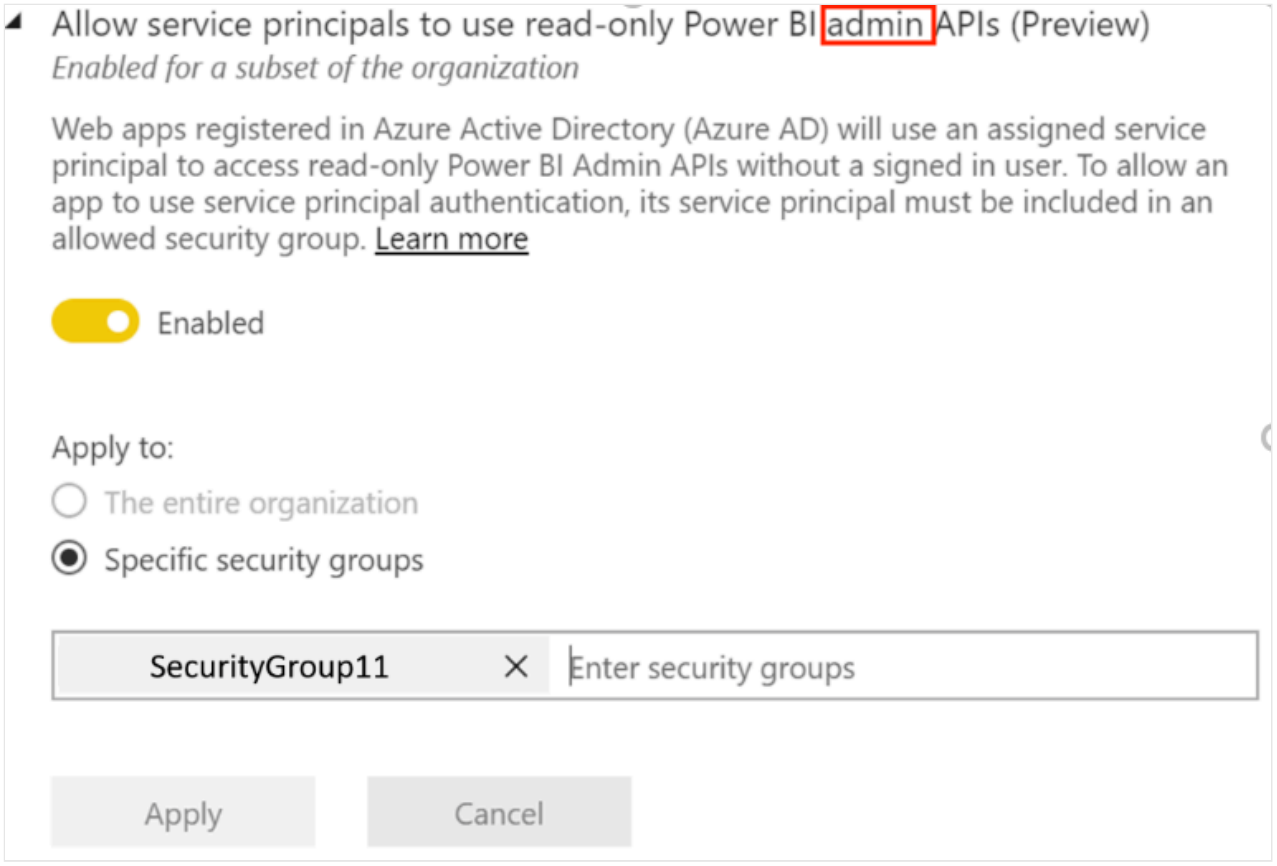
1. Log into the Power BI admin portal.

2. Select the **Tenant settings** page.

> **Important**
>
> You need to be a Power BI Admin to see the tenant settings page.

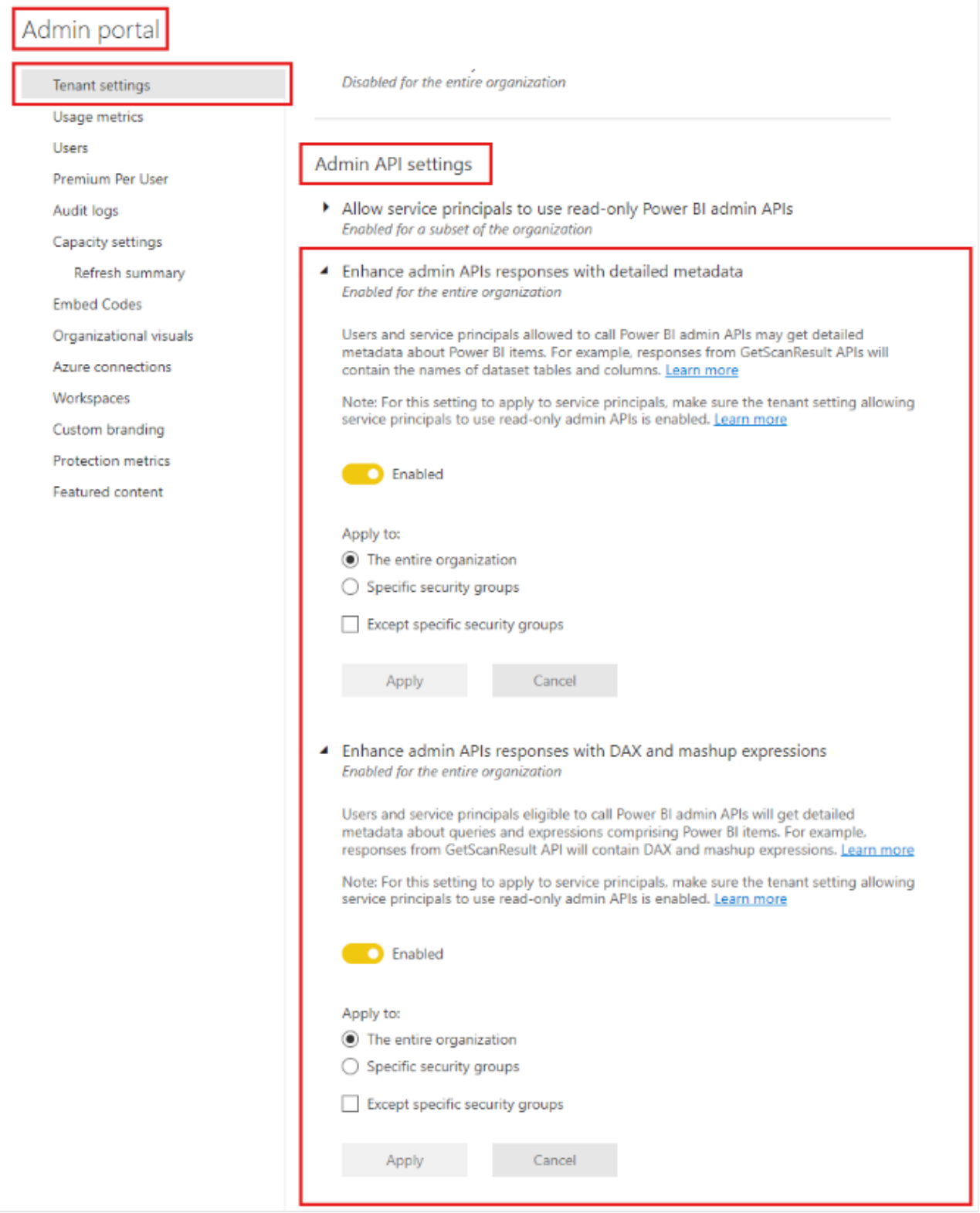
3. Select **Admin API settings** > **Allow service principals to use read-only Power BI admin APIs (Preview)**.

4. Select **Specific security groups**.



5. Select **Admin API settings** > **Enhance admin APIs responses with detailed metadata** and **Enhance admin APIs responses with DAX and mashup expressions** > Enable the toggle to allow Microsoft Purview Data Map automatically discover the detailed metadata of Power BI datasets as part of its scans.

> **ⓘ Important**
>
> After you update the Admin API settings on your power bi tenant, wait around 15 minutes before registering a scan and test connection.



⊗ Caution

When you allow the security group you created (that has your Microsoft Purview managed identity as a member) to use read-only Power BI admin APIs, you also allow it to access the metadata (e.g. dashboard and report names, owners, descriptions, etc.) for all of your Power BI artifacts in this tenant. Once the metadata has been pulled into the Microsoft Purview, Microsoft Purview's permissions, not Power BI permissions, determine who can see that metadata.

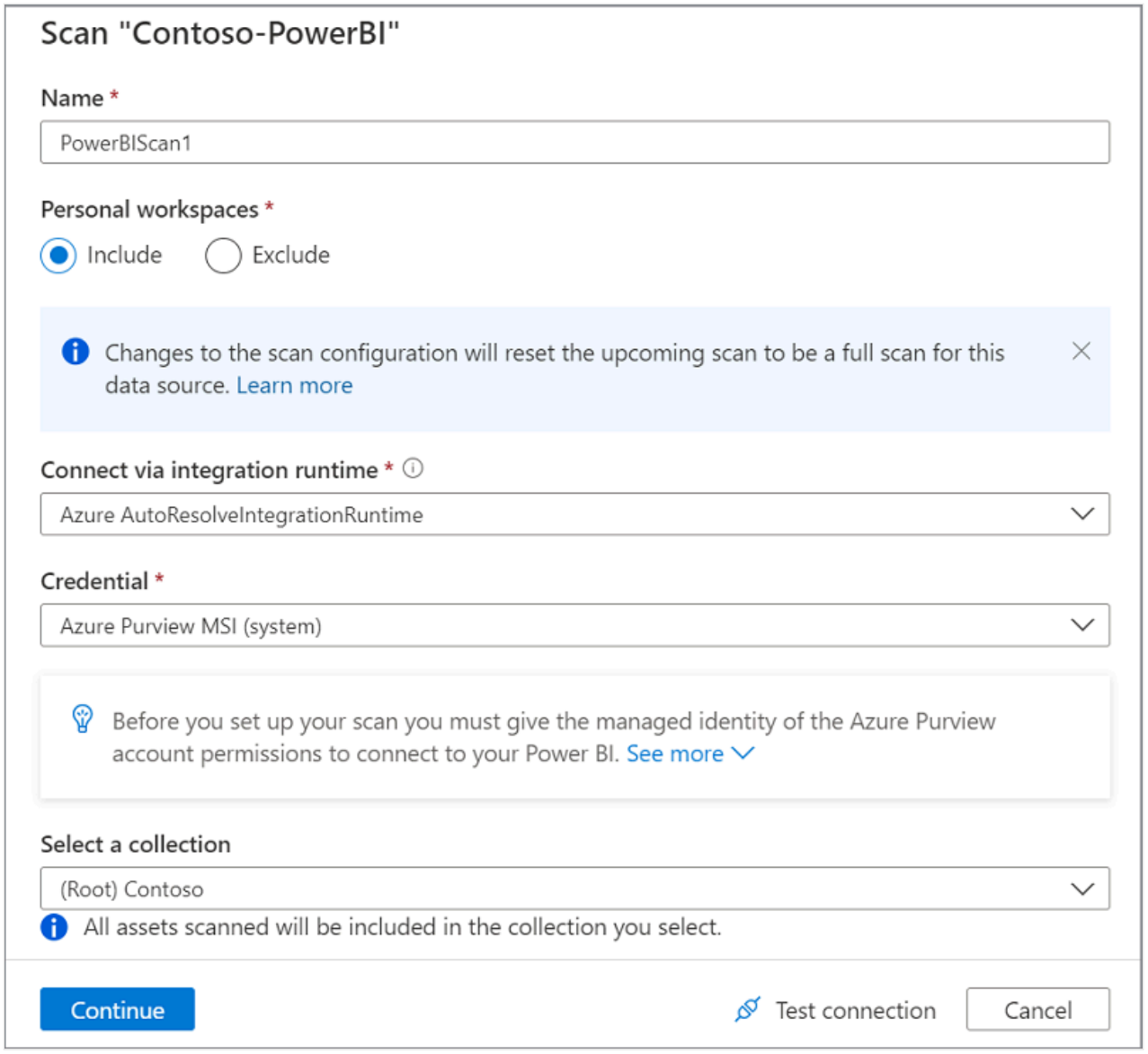
> **Note**
>
> You can remove the security group from your developer settings, but the metadata previously extracted won't be removed from the Microsoft Purview account. You can delete it separately, if you wish.

Create scan for same-tenant Power BI using Azure IR and Managed Identity

This is a suitable scenario, if both Microsoft Purview and Power BI tenant are configured to allow public access in the network settings.

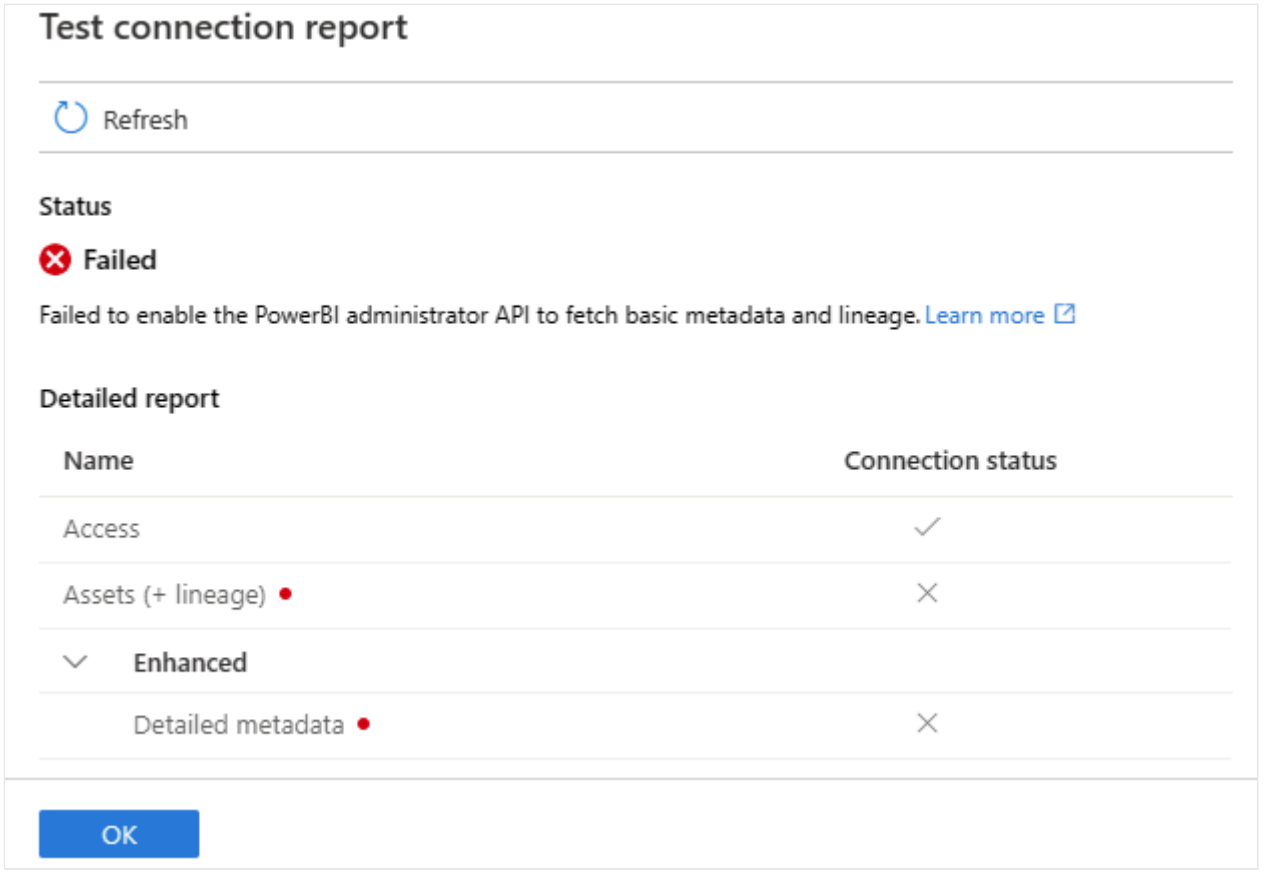
To create and run a new scan, do the following:

1. In the Microsoft Purview Studio, navigate to the **Data map** in the left menu.
2. Navigate to **Sources**.
3. Select the registered Power BI source.
4. Select + **New scan**.
5. Give your scan a name. Then select the option to include or exclude the personal workspaces.

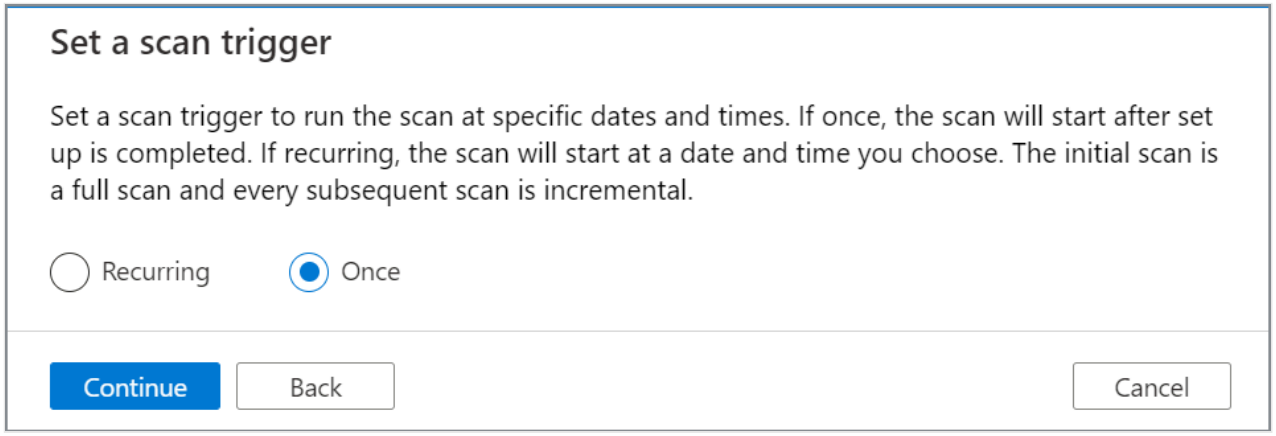


> **Note**
>
> Switching the configuration of a scan to include or exclude a personal workspace will trigger a full scan of Power BI source.

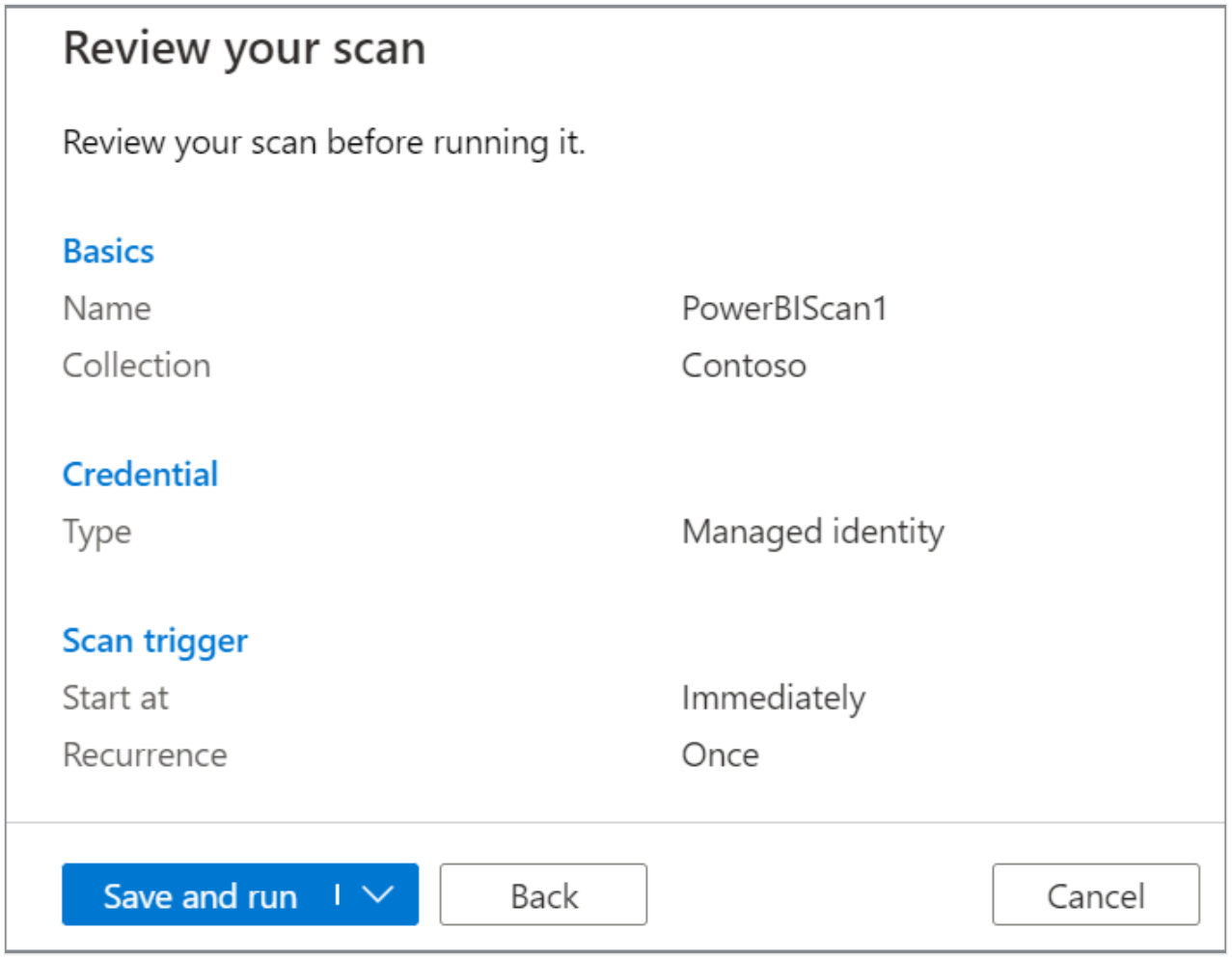
6. Select **Test Connection** before continuing to next steps. If **Test Connection** failed, select **View Report** to see the detailed status and troubleshoot the problem.
   a. Access - Failed status means the user authentication failed. Scans using managed identity will always pass because no user authentication required.
   b. Assets (+ lineage) - Failed status means the Microsoft Purview - Power BI authorization has failed. Make sure the Microsoft Purview managed identity is added to the security group associated in Power BI admin portal.
   c. Detailed metadata (Enhanced) - Failed status means the Power BI admin portal is disabled for the following setting - **Enhance admin APIs responses with detailed metadata**



7. Set up a scan trigger. Your options are **Recurring**, and **Once**.



8. On **Review new scan**, select **Save and run** to launch your scan.



Create scan for same-tenant using self-hosted IR with service principal

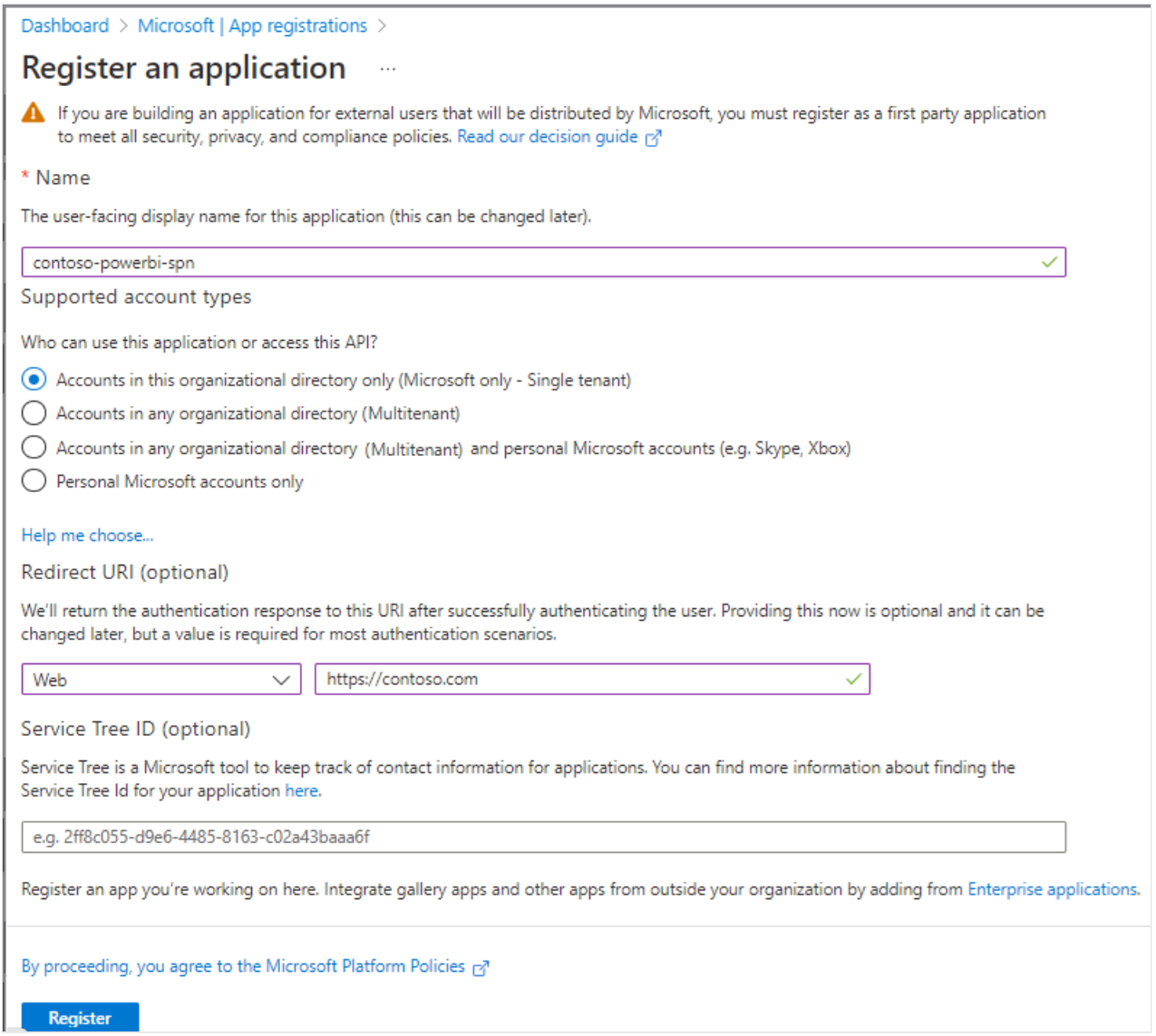
This scenario can be used when Microsoft Purview and Power BI tenant or both, are configured to use private endpoint and deny public access. Additionally, this option is also applicable if Microsoft Purview and Power BI tenant are configured to allow public access.

For more information related to Power BI network, see How to configure private endpoints for accessing Power BI.

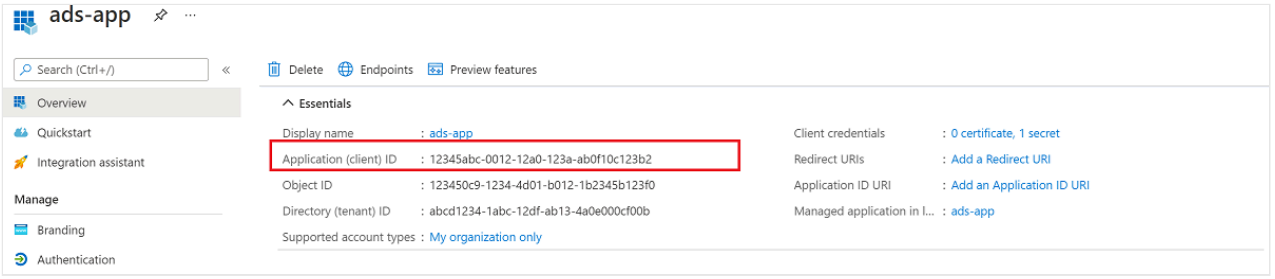
For more information about Microsoft Purview network settings, see Use private endpoints for your Microsoft Purview account.

To create and run a new scan, do the following:

1. In the Azure portal, select **Microsoft Entra ID** and create an App Registration in the tenant. Provide a web URL in the **Redirect URI**. For information about the Redirect URI see this documentation from Microsoft Entra ID.

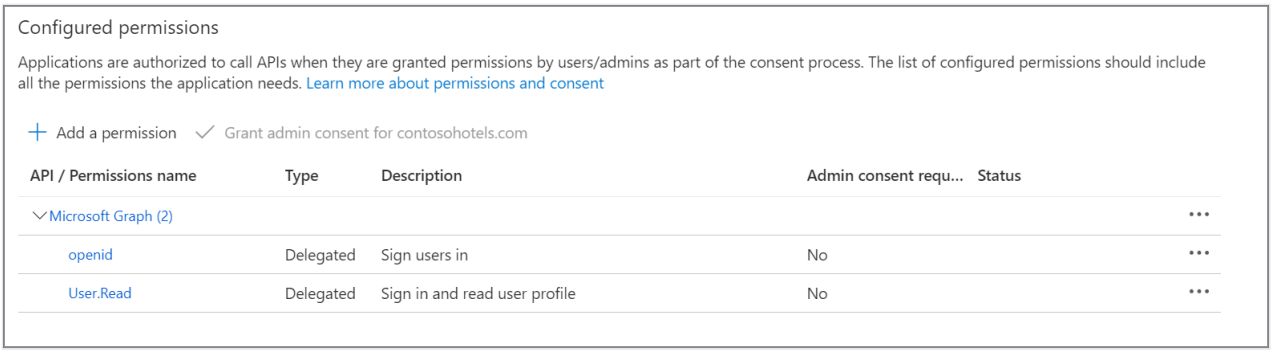


2. Take note of Client ID(App ID).



3. From Microsoft Entra dashboard, select newly created application and then select **App registration**. From **API Permissions**, assign the application the following delegated permissions:

   - Microsoft Graph openid
   - Microsoft Graph User.Read

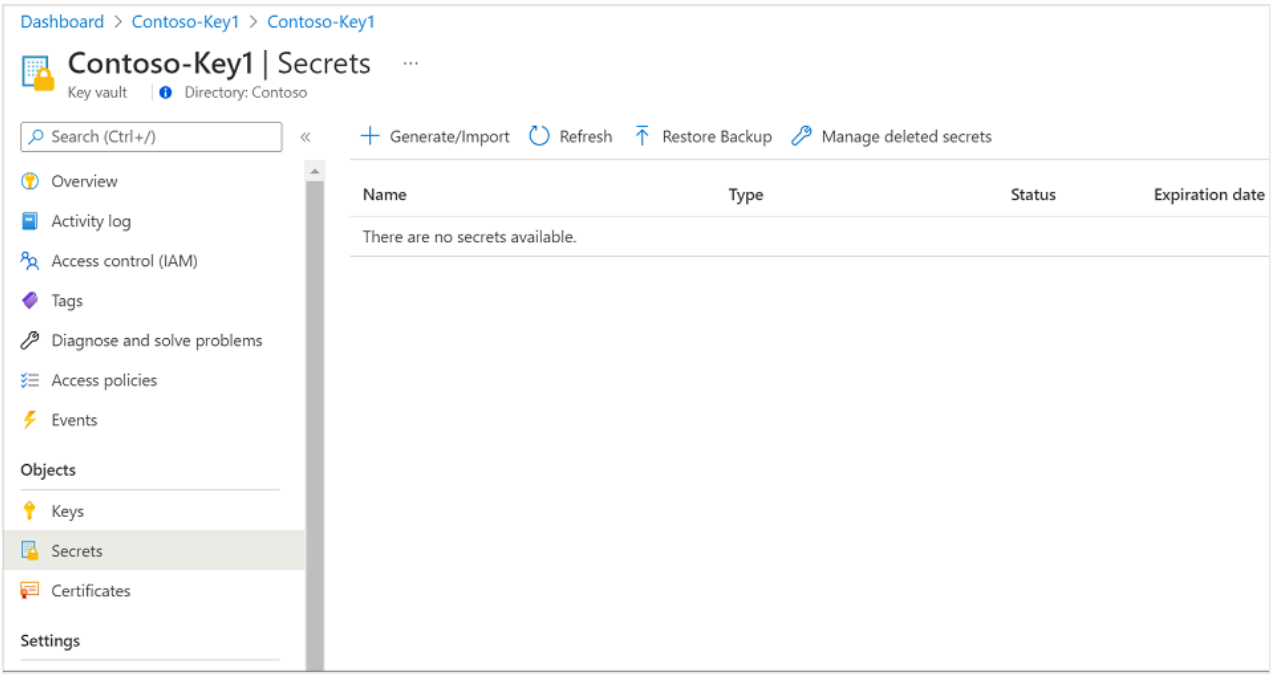


4. Under **Advanced settings**, enable **Allow Public client flows**.

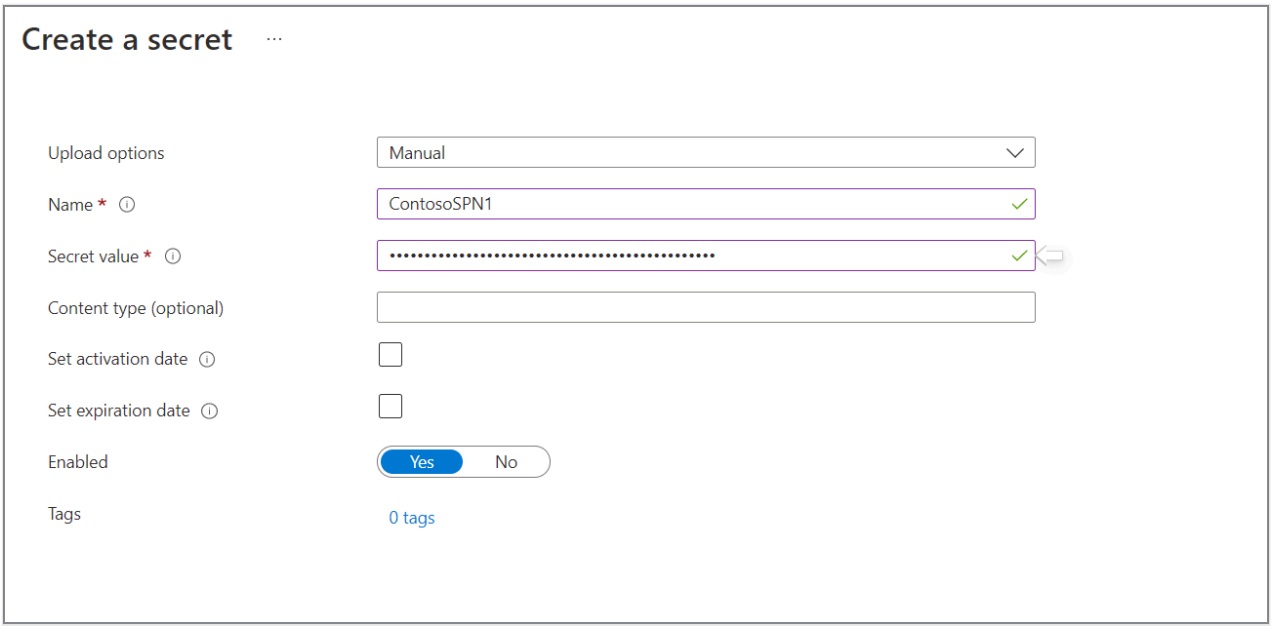
5. Under **Certificates & secrets**, create a new secret and save it securely for next steps.

6. In Azure portal, navigate to your Azure key vault.

7. Select **Settings** > **Secrets** and select **+ Generate/Import**.



8. Enter a name for the secret and for **Value**, type the newly created secret for the App registration. Select **Create** to complete.



9. If your key vault isn't connected to Microsoft Purview yet, you'll need to create a new key vault connection

10. In the Microsoft Purview Studio, navigate to the **Data map** in the left menu.

11. Navigate to **Sources**.

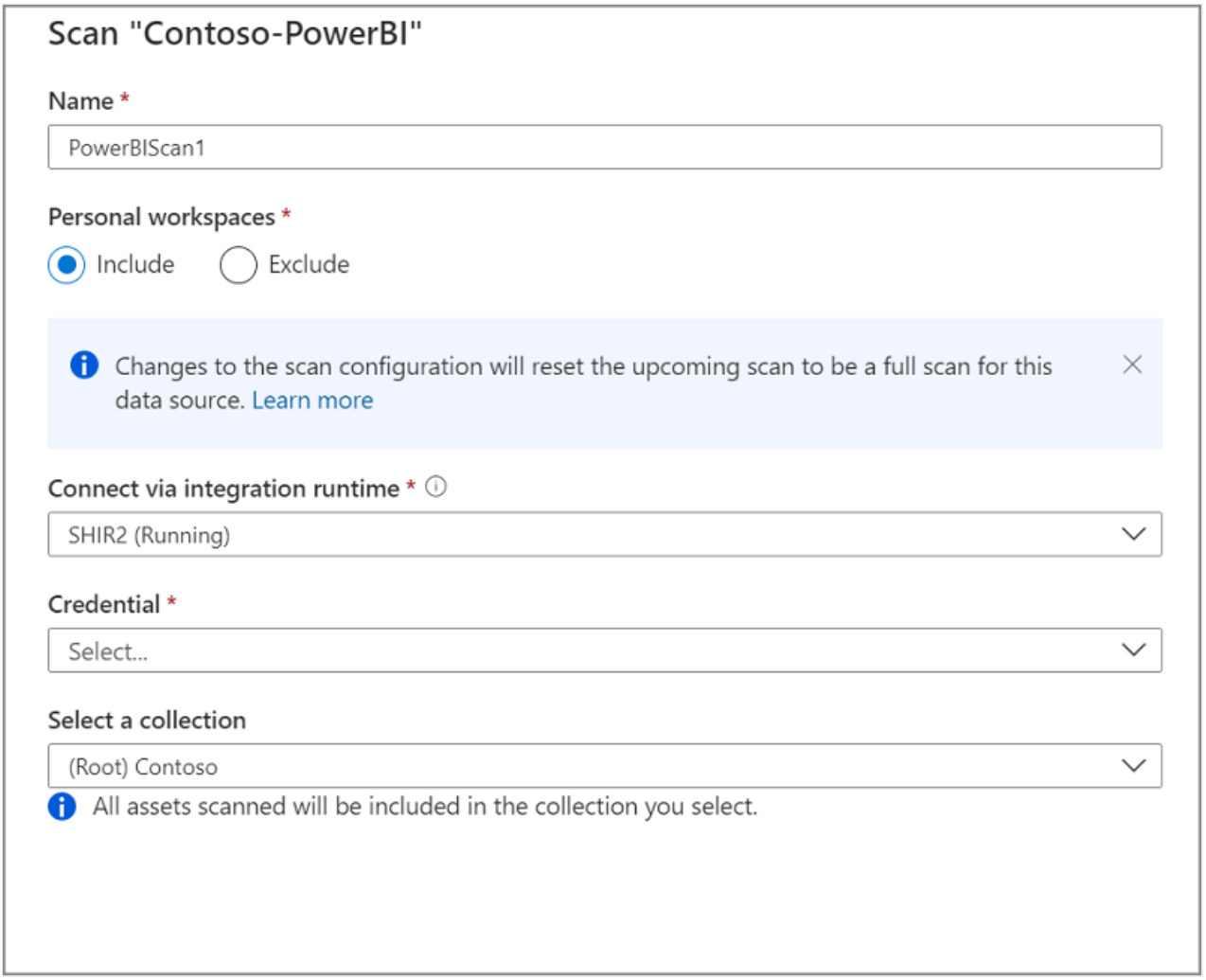
12. Select the registered Power BI source.

13. Select + **New scan**.

14. Give your scan a name. Then select the option to include or exclude the personal workspaces.

    > **Note**
    >
    > Switching the configuration of a scan to include or exclude a personal workspace will trigger a full scan of Power BI source.

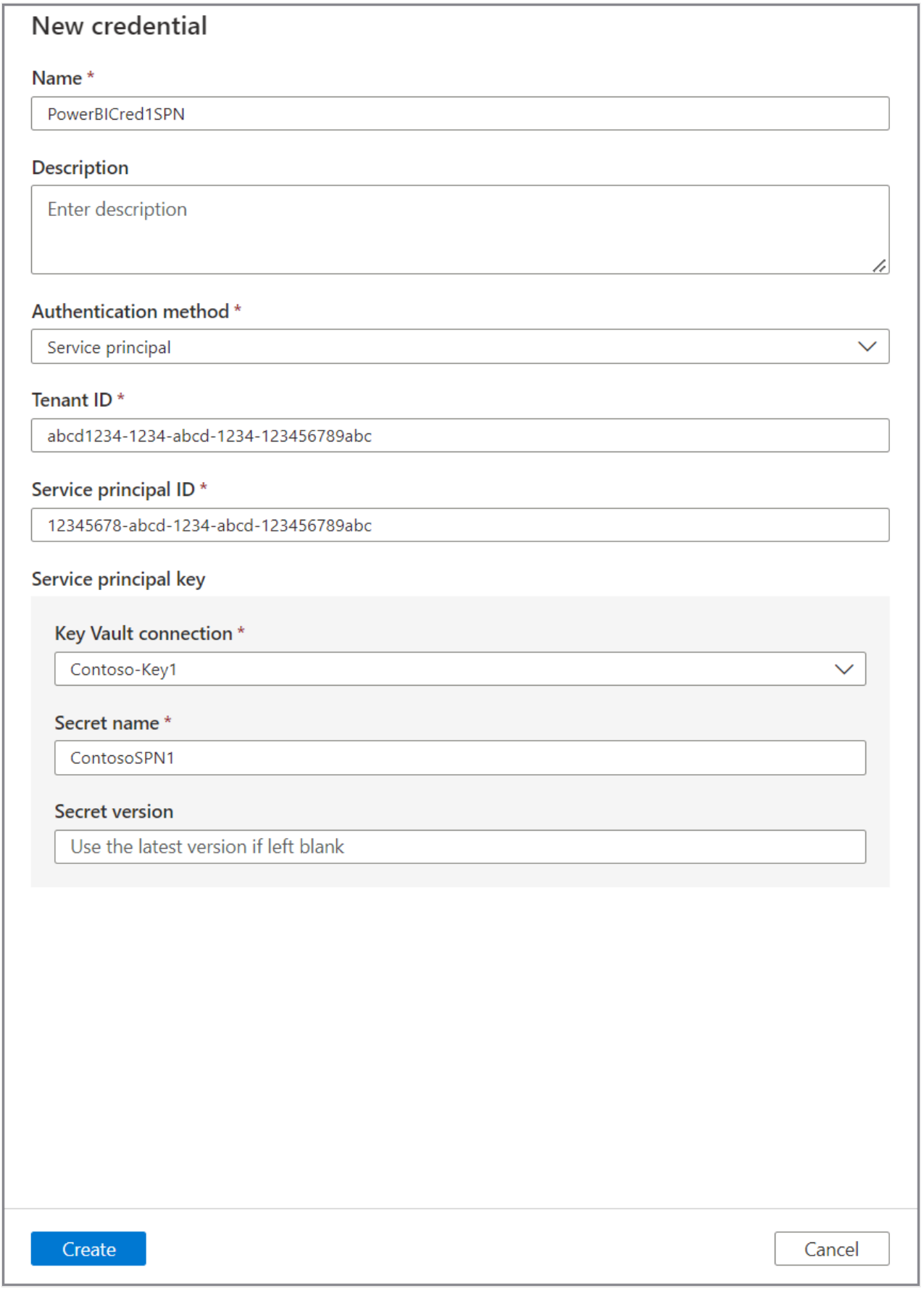
15. Select your self-hosted integration runtime from the drop-down list.



16. For the **Credential**, select **service principal** and select **+ New** to create a new credential.

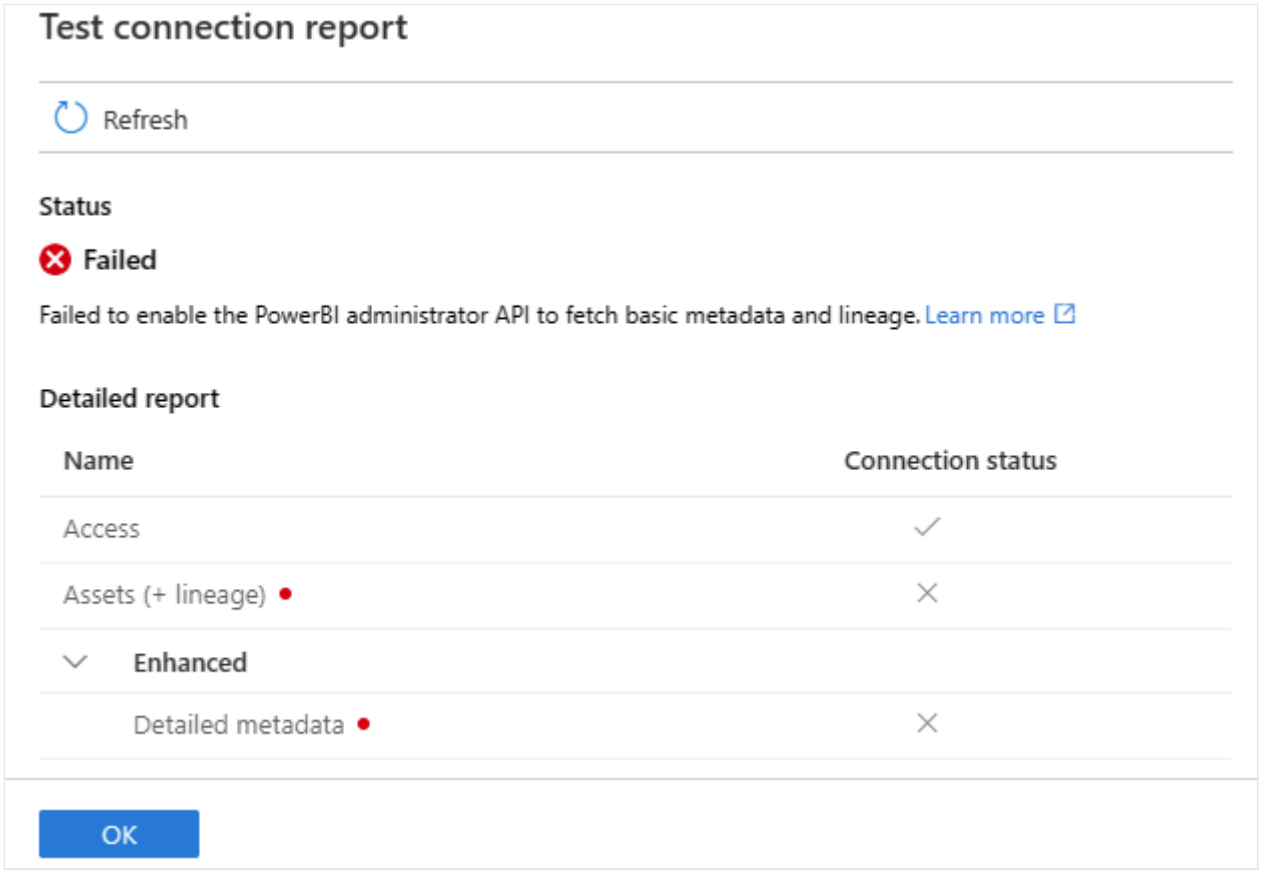
17. Create a new credential and provide required parameters:

    - **Name**: Provide a unique name for credential
    - **Authentication method**: Service principal
    - **Tenant ID**: Your Power BI tenant ID
    - **Client ID**: Use Service Principal Client ID (App ID) you created earlier

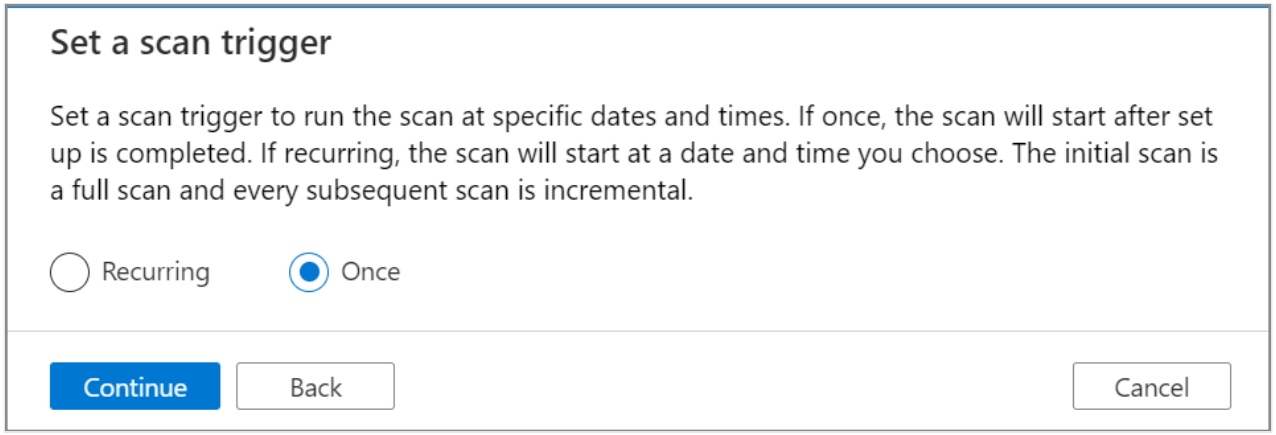


18. Select **Test Connection** before continuing to next steps. If **Test Connection** failed, select **View Report** to see the detailed status and troubleshoot the problem
    a. Access - Failed status means the user authentication failed. Scans using managed identity will always pass because no user authentication required.

b. Assets (+ lineage) - Failed status means the Microsoft Purview - Power BI authorization has failed. Make sure the Microsoft Purview managed identity is added to the security group associated in Power BI admin portal.
c. Detailed metadata (Enhanced) - Failed status means the Power BI admin portal is disabled for the following setting - **Enhance admin APIs responses with detailed metadata**



19. Set up a scan trigger. Your options are **Recurring**, and **Once**.



20. On **Review new scan**, select **Save and run** to launch your scan.

Create scan for same-tenant using self-hosted IR with delegated authentication

This scenario can be used when Microsoft Purview and Power BI tenant or both, are configured to use private endpoint and deny public access. Additionally, this option is also applicable if Microsoft Purview and Power BI tenant are configured to allow public access.

For more information related to Power BI network, see How to configure private endpoints for accessing Power BI.

For more information about Microsoft Purview network settings, see Use private endpoints for your Microsoft Purview account.

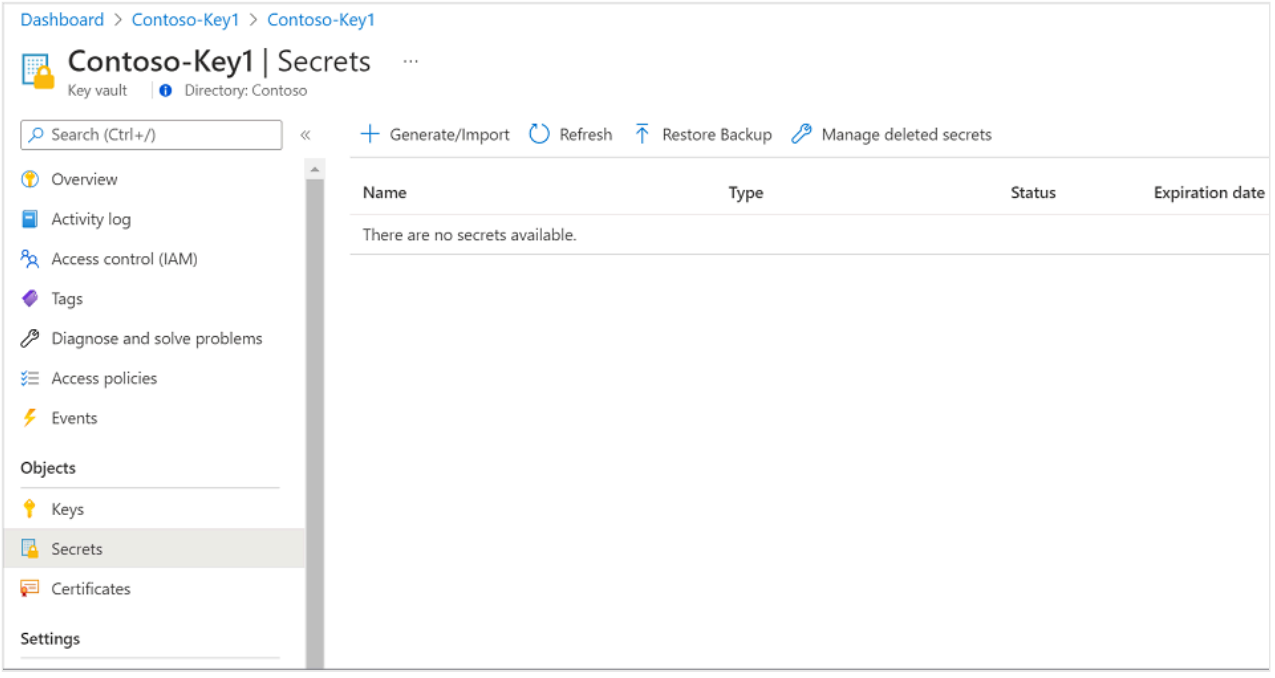
To create and run a new scan, do the following:

1. Create a user account in Microsoft Entra tenant and assign the user to Microsoft Entra role, **Fabric Administrator**. Take note of username and sign in to change the password.

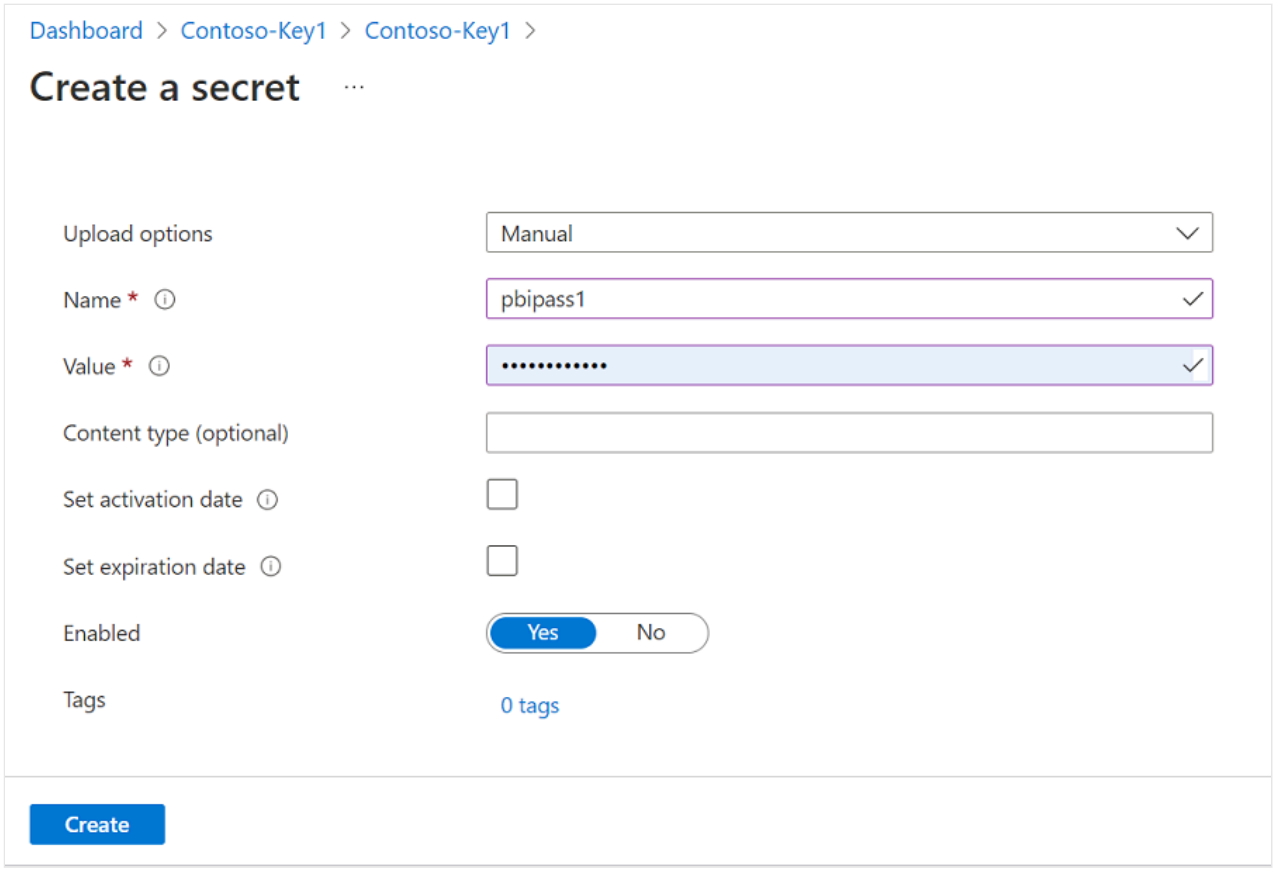
2. Assign proper Power BI license to the user.

3. Navigate to your Azure key vault.

4. Select **Settings** > **Secrets** and select **+ Generate/Import**.

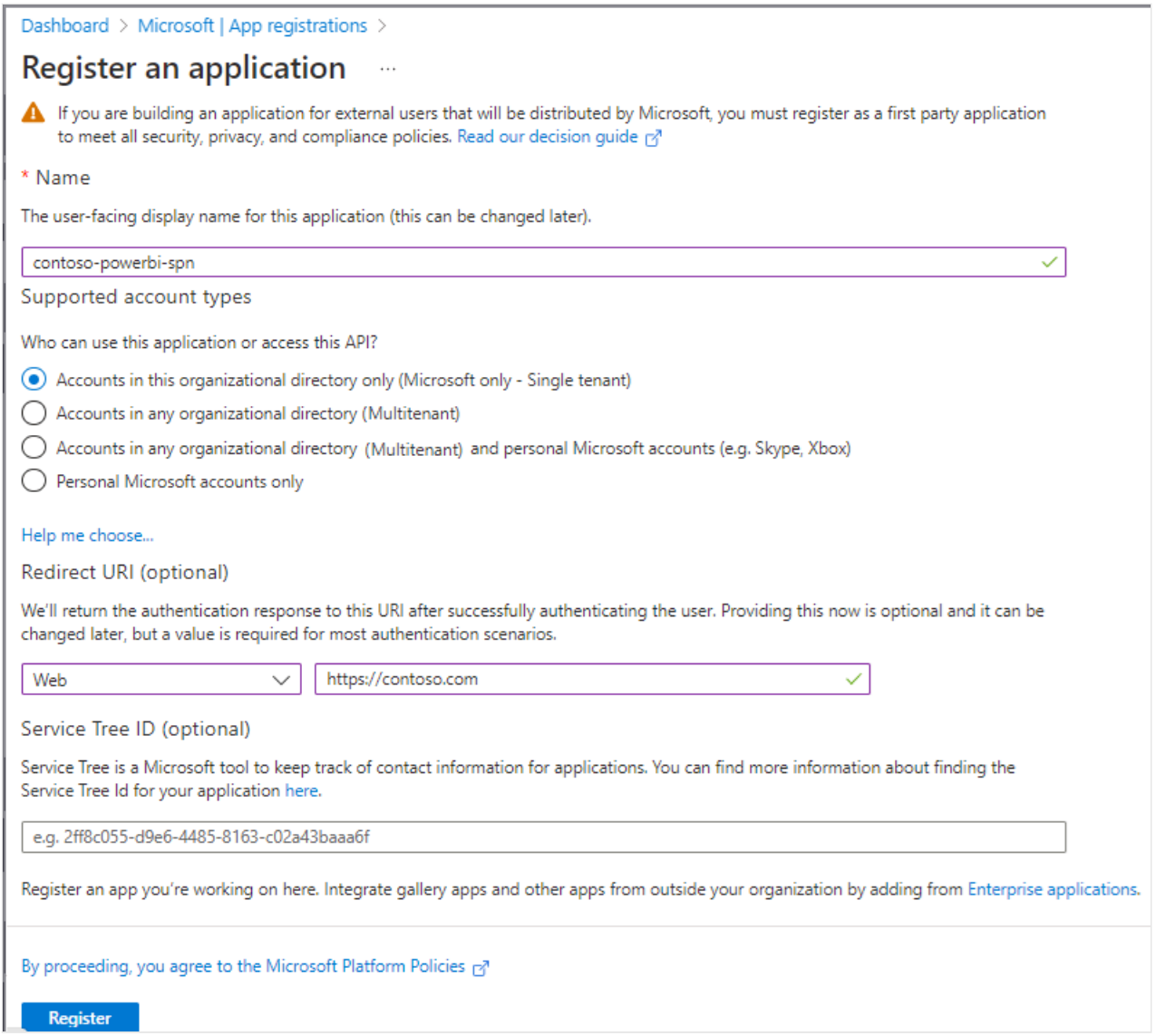


5. Enter a name for the secret and for **Value**, type the newly created password for the Microsoft Entra user. Select **Create** to complete.

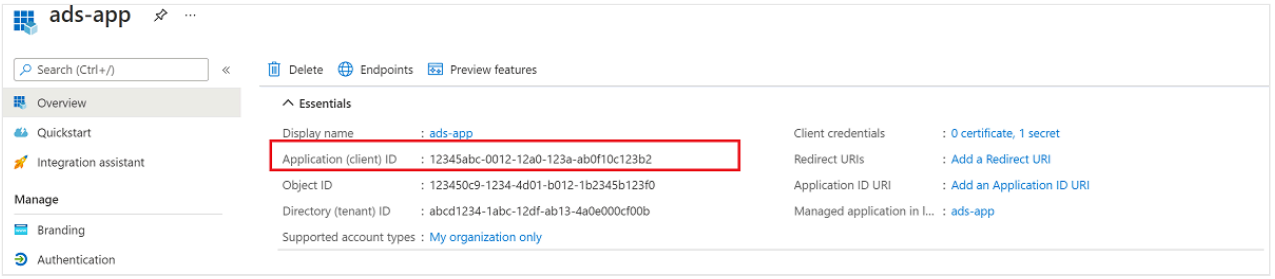


6. If your key vault isn't connected to Microsoft Purview yet, you'll need to create a new key vault connection

7. Create an App Registration in your Microsoft Entra tenant. Provide a web URL in the **Redirect URI**.

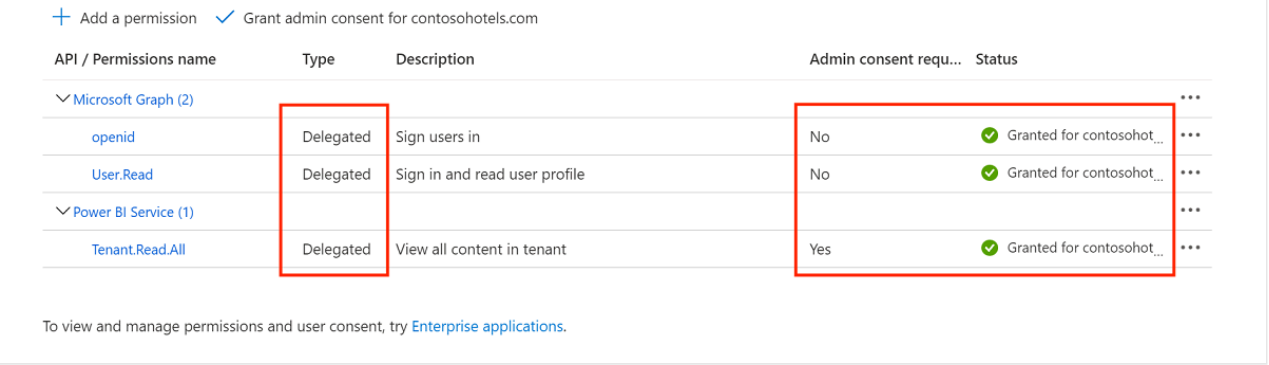


8. Take note of Client ID(App ID).



9. From Microsoft Entra dashboard, select newly created application and then select **App registration**. Assign the application the following delegated permissions, and grant admin consent for the tenant:

   - Power BI Service Tenant.Read.All
   - Microsoft Graph openid
   - Microsoft Graph User.Read



10. Under **Advanced settings**, enable **Allow Public client flows**.

11. In the Microsoft Purview Studio, navigate to the **Data map** in the left menu.

12. Navigate to **Sources**.

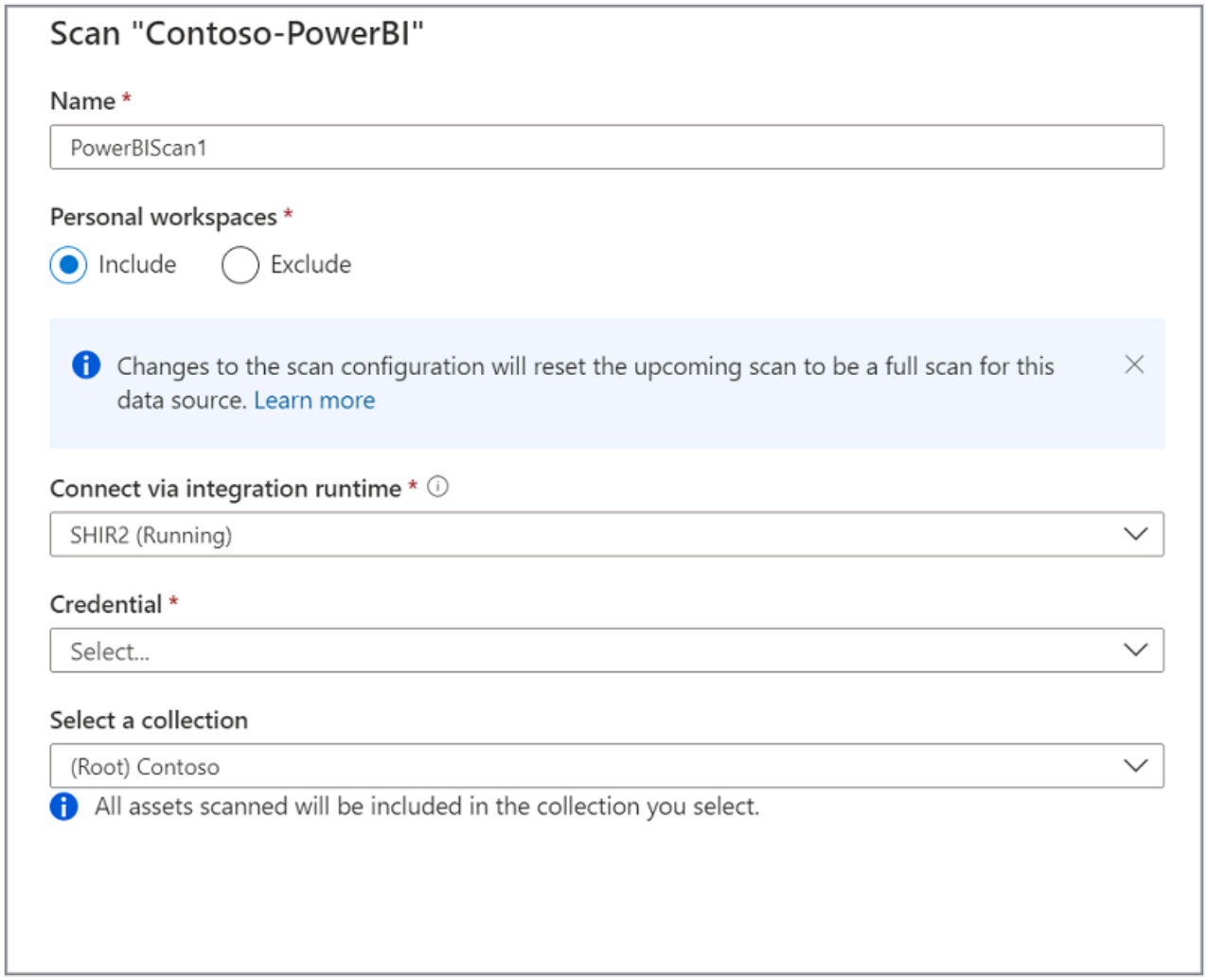
13. Select the registered Power BI source.

14. Select **+ New scan**.

15. Give your scan a name. Then select the option to include or exclude the personal workspaces.

    > **Note**
    >
    > Switching the configuration of a scan to include or exclude a personal workspace will trigger a full scan of Power BI source.

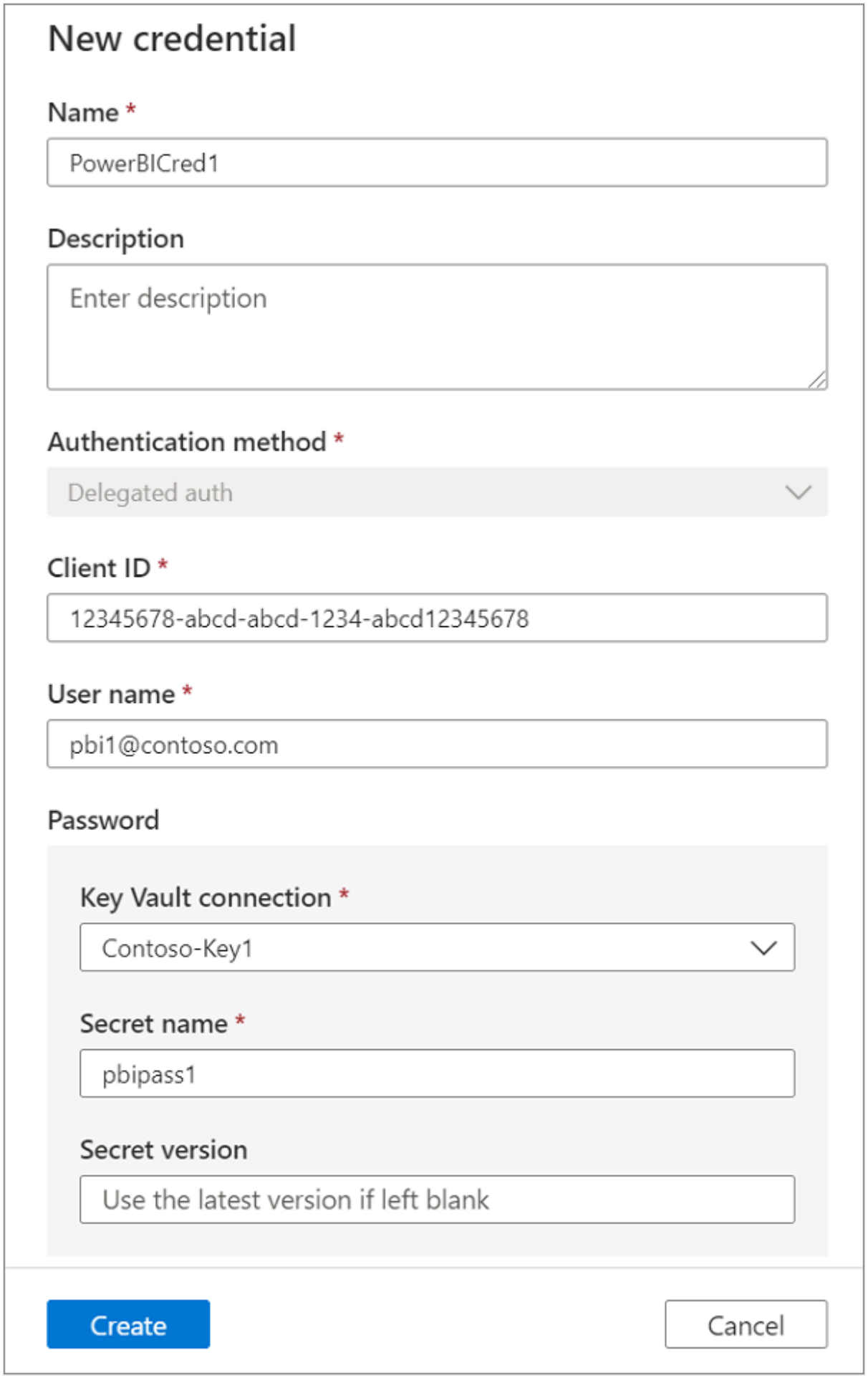
16. Select your self-hosted integration runtime from the drop-down list.



17. For the **Credential**, select **Delegated authentication** and select **+ New** to create a new credential.

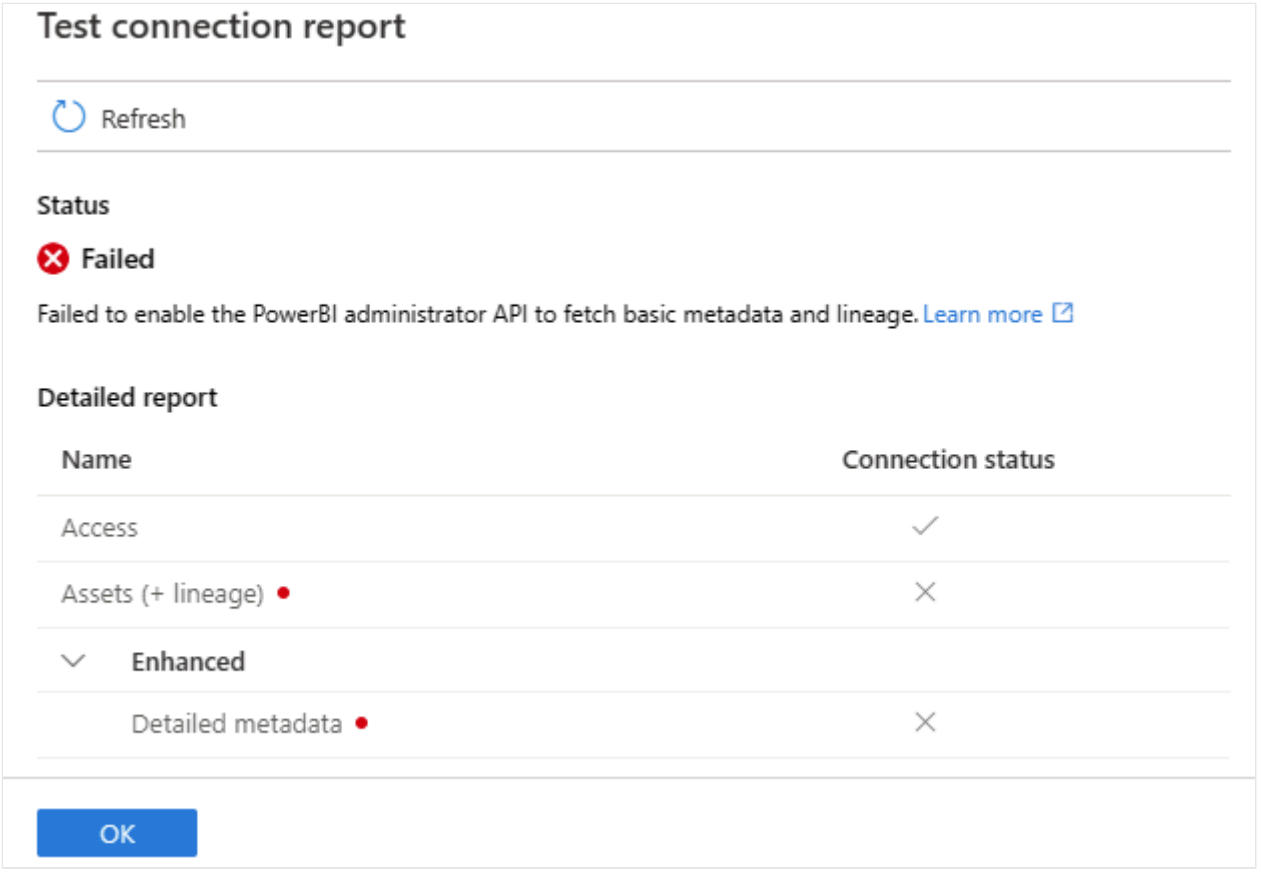
18. Create a new credential and provide required parameters:

    - **Name**: Provide a unique name for credential
    - **Authentication method**: Delegated auth
    - **Client ID**: Use Service Principal Client ID (App ID) you created earlier
    - **User name**: Provide the username of Fabric Administrator you created earlier
    - **Password**: Select the appropriate Key vault connection and the **Secret name** where the Power BI account password was saved earlier.

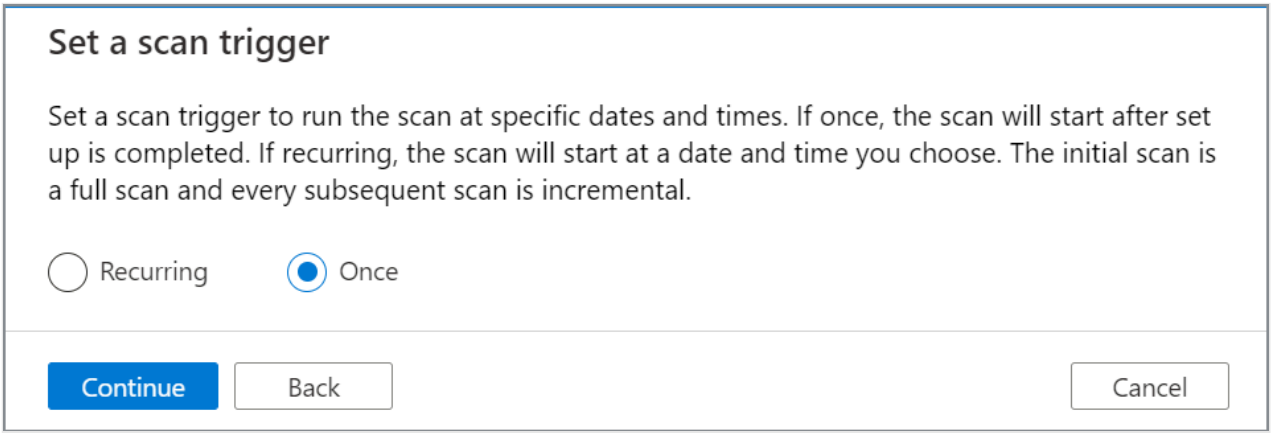


19. Select **Test Connection** before continuing to next steps. If **Test Connection** failed, select **View Report** to see the detailed status and troubleshoot the problem

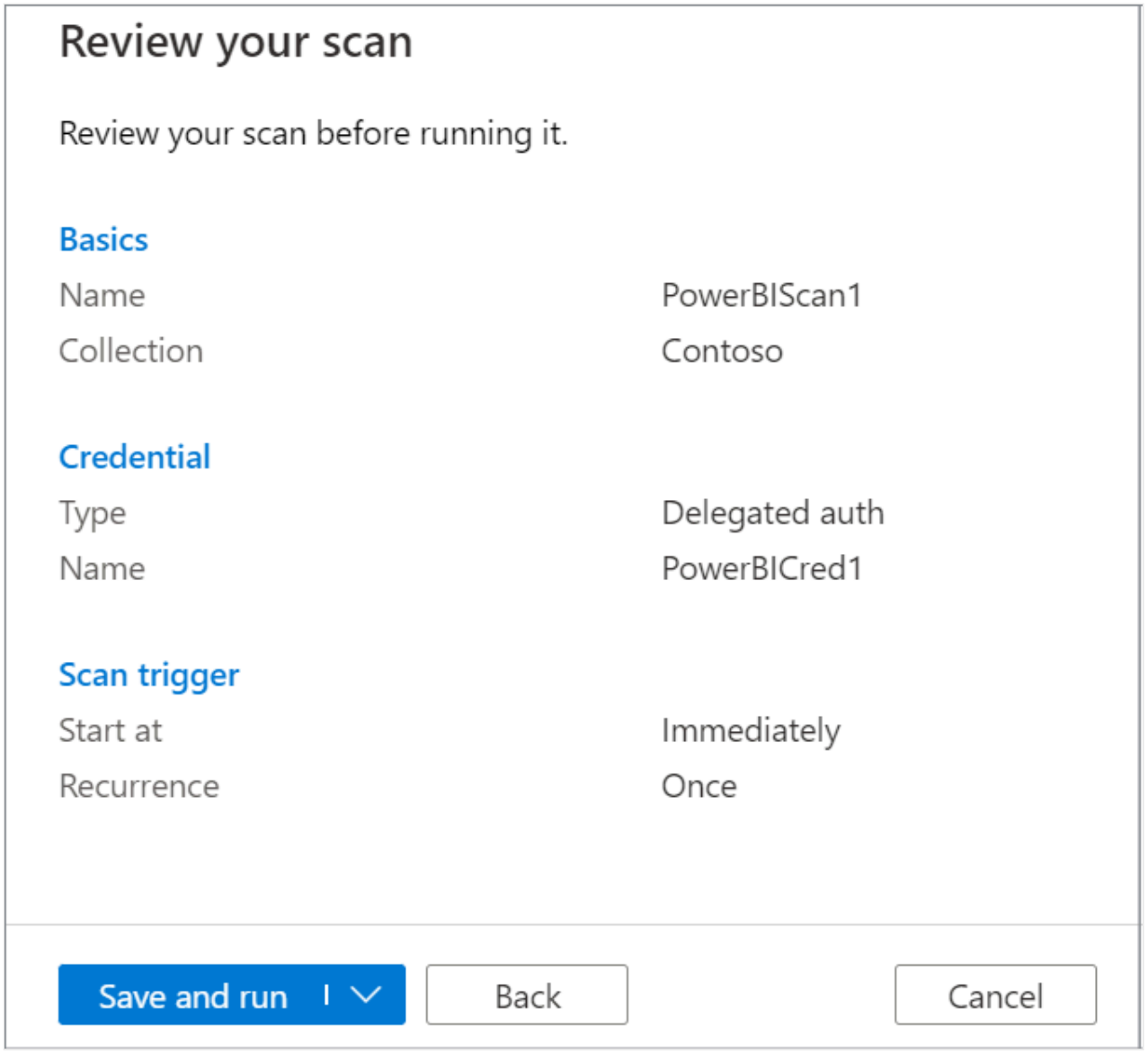
a. Access - Failed status means the user authentication failed. Scans using managed identity will always pass because no user authentication required.
b. Assets (+ lineage) - Failed status means the Microsoft Purview - Power BI authorization has failed. Make sure the Microsoft Purview managed identity is added to the security group associated in Power BI admin portal.
c. Detailed metadata (Enhanced) - Failed status means the Power BI admin portal is disabled for the following setting - **Enhance admin APIs responses with detailed metadata**



20. Set up a scan trigger. Your options are **Recurring**, and **Once**.



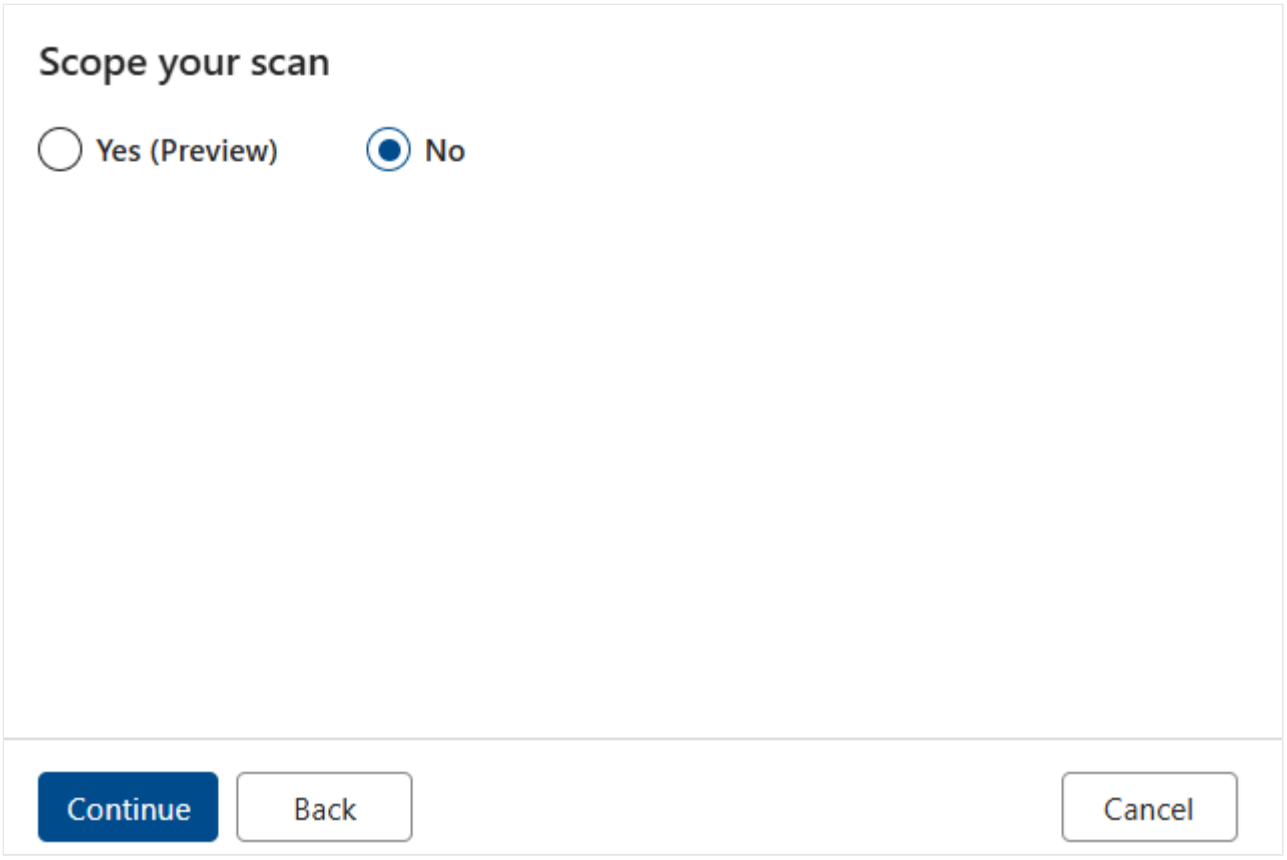
21. On **Review new scan**, select **Save and run** to launch your scan.



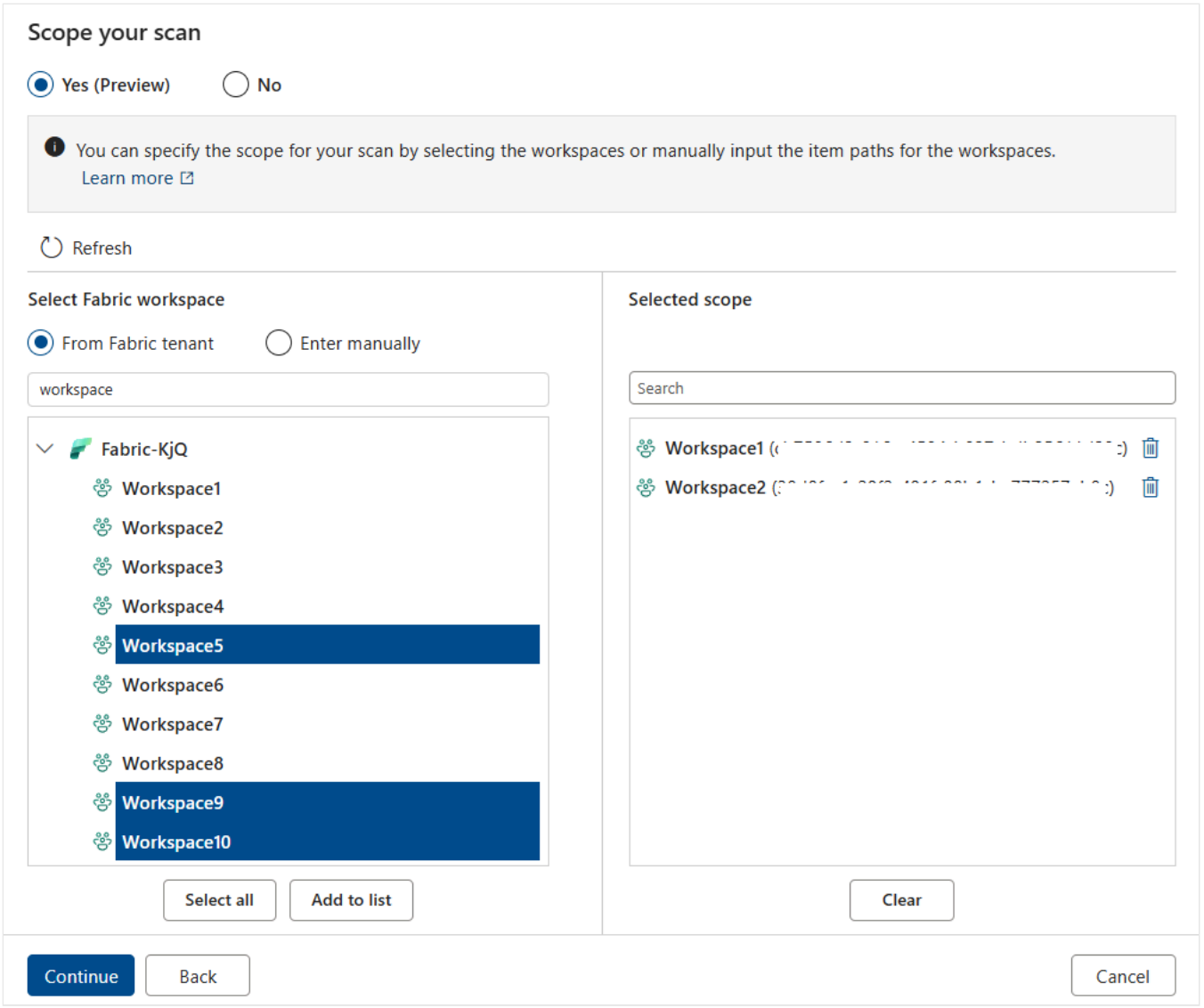
Scope your scan

This feature is currently in preview. The Supplemental Terms of Use for Microsoft Azure Previews include additional legal terms that apply to Azure features that are in beta, in preview, or otherwise not yet released into general availability.

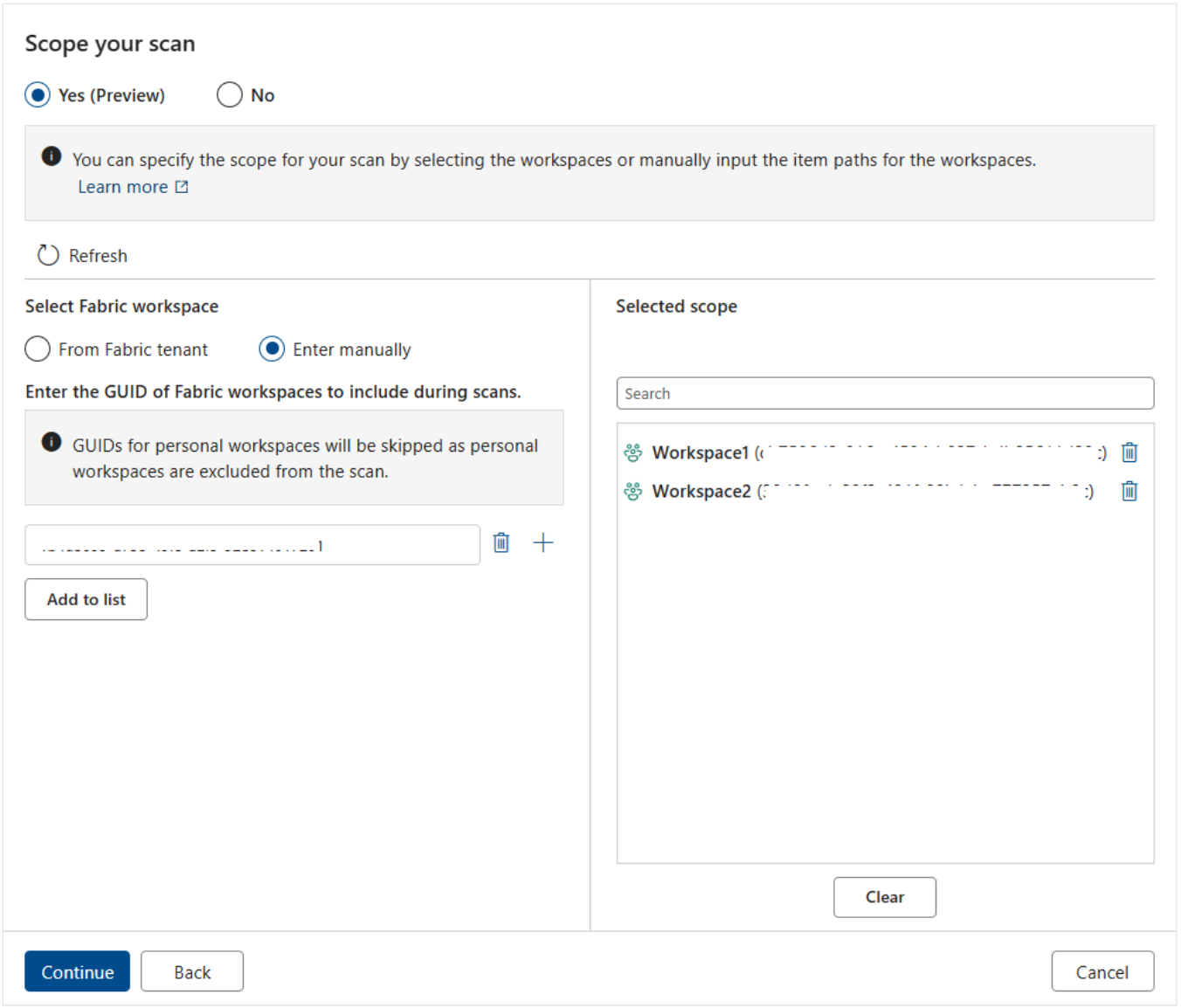
In the Scope your scan stage, data source administrators can scope the scans by specifying the workspaces as needed. Data source administrators can either select from the workspace list under the tenant or directly input the workspace GUIDs to scope the scan. No additional permission is required to use the scoped scan. There are two options available and by default the option "No" is selected and you can directly click on "Continue" to proceed to run a full scan or select the "Yes" to enable the scoped scan experience.



If the scoped scan is enabled, you will see a list of existing workspaces under the tenant on the left hand side (currently there's a limitation that the UI selection experience only supports less than 5000 workspaces) and you can select the workspaces and add to the list on the right hand side. If personal workspace is configured as "Include", personal workspaces will show in the workspace list with a prefix "PersonalWorkspace ". You can directly click on the workspace(s) and click on the button "Add to list" to put the workspaces to the list of selected workspaces. You can input keyword of the workspace name or workspace GUID in the search box to filter out the workspace(s) for selection.



You can also directly input the workspace GUID and add it to the list of selected workspaces.



> **Note**
>
> - Scoping scan by selecting from Microsoft Fabric or Power BI tenant is only supported when the total number of workspaces is less than 5000 (include or exclude personal workspaces will be considered), or you need to switch to manual input of workspace GUIDs to scope your scan.
> - You can switch between selecting workspaces from Microsoft Fabric or Power BI tenant, or entering workspace GUID manually and the input will be merged into the same list of selected workspaces.
> - Microsoft Purview will check whether the GUID input by user has a correct format but will not check whether the GUID represents a valid workspace in the Microsoft Fabric or Power BI tenant. If invalid GUIDs are included in scoped scan, the scan will complete with exceptions and you can find those invalid GUIDs in the scan log.
> - If personal workspace is configured as "Exclude", the GUIDs for personal workspaces can still be added to the list of selected workspaces but will be skipped in the scan

and these GUIDs will be included in the scan log.

- If self-hosted integration runtime is used, version 5.40.8836.1 and above is required and only manual input of workspace GUIDs is supported in the scoped scan.
- If managed VNet integration runtime (v2) is used, only manual input of workspace GUIDs is supported in the scoped scan.

Next steps

Now that you've registered your source, follow the below guides to learn more about Microsoft Purview and your data.

- Data Estate Insights in Microsoft Purview
- Lineage in Microsoft Purview
- Search Unified Catalog

Connect to and manage a Power BI tenant in Microsoft Purview (cross-tenant)

Article • 05/21/2025

This article outlines how to register a Power BI tenant in a cross-tenant scenario, and how to authenticate and interact with the tenant in Microsoft Purview. If you're unfamiliar with the service, see Learn about Microsoft Purview.

> **Note**
>
> For scanning a Microsoft Fabric tenant, see our Microsoft Fabric documentation.

Supported capabilities

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan	Classification	Labeling	Access Policy	Lineage	Data Sharing	Live view
Yes	Yes	Yes	Yes	No	No	No	Yes	No	No

* Power BI items in a Microsoft Fabric tenant are available using live view.

When scanning Power BI source, Microsoft Purview supports:

- Extracting technical metadata including:
  - Workspaces
  - Dashboards
  - Reports
  - Datasets including the tables and columns
  - Dataflows
  - Datamarts
- Fetching static lineage on assets relationships among above Power BI artifacts and external data source assets. Learn more from Power BI lineage.

Supported scenarios for Power BI scans

Expand table

Scenario	Microsoft Purview public access	Power BI public access	Runtime option	Authentication option	Deployment checklist
Public access with Azure integration runtime	Allowed	Allowed	Azure runtime	Delegated authentication / Service principal	Deployment checklist
Public access with self-hosted integration runtime	Allowed	Allowed	Self-hosted runtime	Delegated authentication / Service principal	Deployment checklist
Private access	Denied	Allowed	Managed VNet IR (v2 only)	Delegated authentication / Service principal	Review deployment checklist
Private access	Allowed	Denied	Self-hosted runtime	Delegated authentication / Service principal	Deployment checklist
Private access	Denied	Allowed	Self-hosted runtime	Delegated authentication / Service principal	Deployment checklist
Private access	Denied	Denied	Self-hosted runtime	Delegated authentication / Service principal	Deployment checklist

Known limitations

- If Power BI tenant is protected behind a private endpoint, standard or kubernetes supported self-hosted runtimes are the only options to scan.
- For the cross-tenant scenario, delegated authentication and service principal are the only supported authentication options for scanning.
- If the Power BI dataset schema isn't shown after the scan, it's due to one of the current limitations with the Power BI metadata scanner.
- Empty workspaces are skipped.

Prerequisites

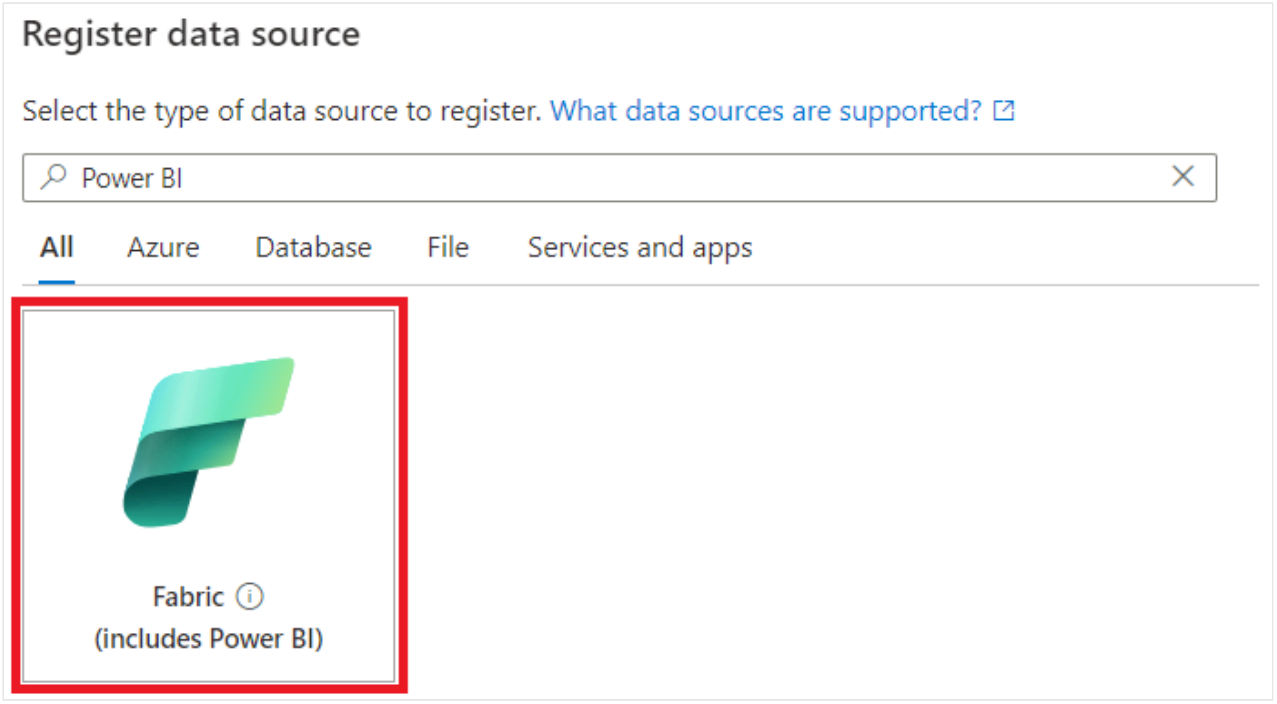
Before you start, make sure you have the following:

- An Azure account with an active subscription. Create an account for free.
- An active Microsoft Purview account.

Register the Power BI tenant

1. From the options on the left, select **Data Map**.

2. Select **Register**, and then select **Fabric** as your data source. It includes Power BI sources and assets.



3. Give your data source a friendly name. The name must be 3 to 63 characters long, and must contain only letters, numbers, underscores, and hyphens. Spaces aren't allowed.

4. Edit the **Tenant ID** field, to replace with the tenant for the cross-tenant Power BI that you want to register and scan. By default, Microsoft Purview's tenant ID is populated.



5. Select **Register**.

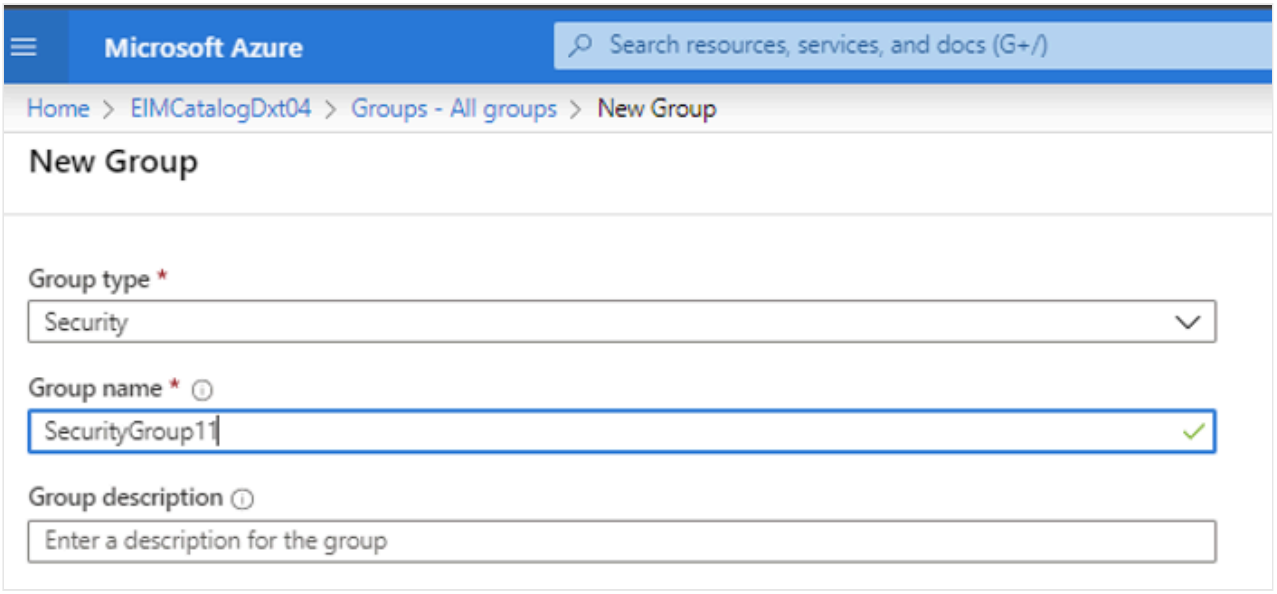
Authenticate to Power BI tenant

In Microsoft Entra tenant, where Power BI tenant is located:

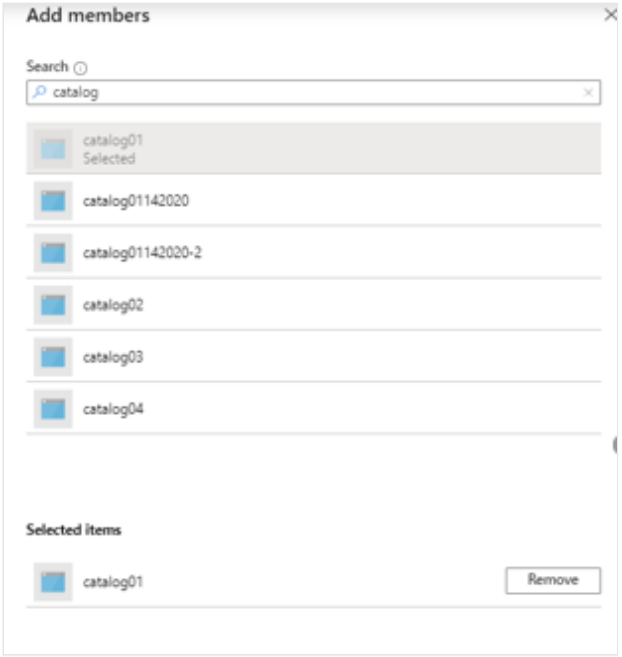
1. In the Azure portal, search for **Microsoft Entra ID**.
2. Create a new security group in your Microsoft Entra ID, by following Create a basic group and add members using Microsoft Entra ID.

   > **Tip**
   >
   > You can skip this step if you already have a security group you want to use.

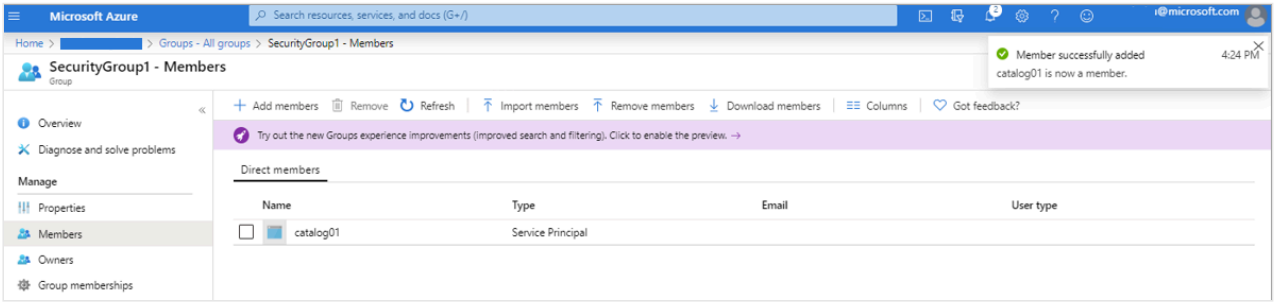
3. Select **Security** as the **Group Type**.



4. Select **Members**, then select **+ Add members**.
5. Search for your Microsoft Purview managed identity or service principal and select it.



You should see a success notification showing you that it was added.



Associate the security group with Power BI tenant

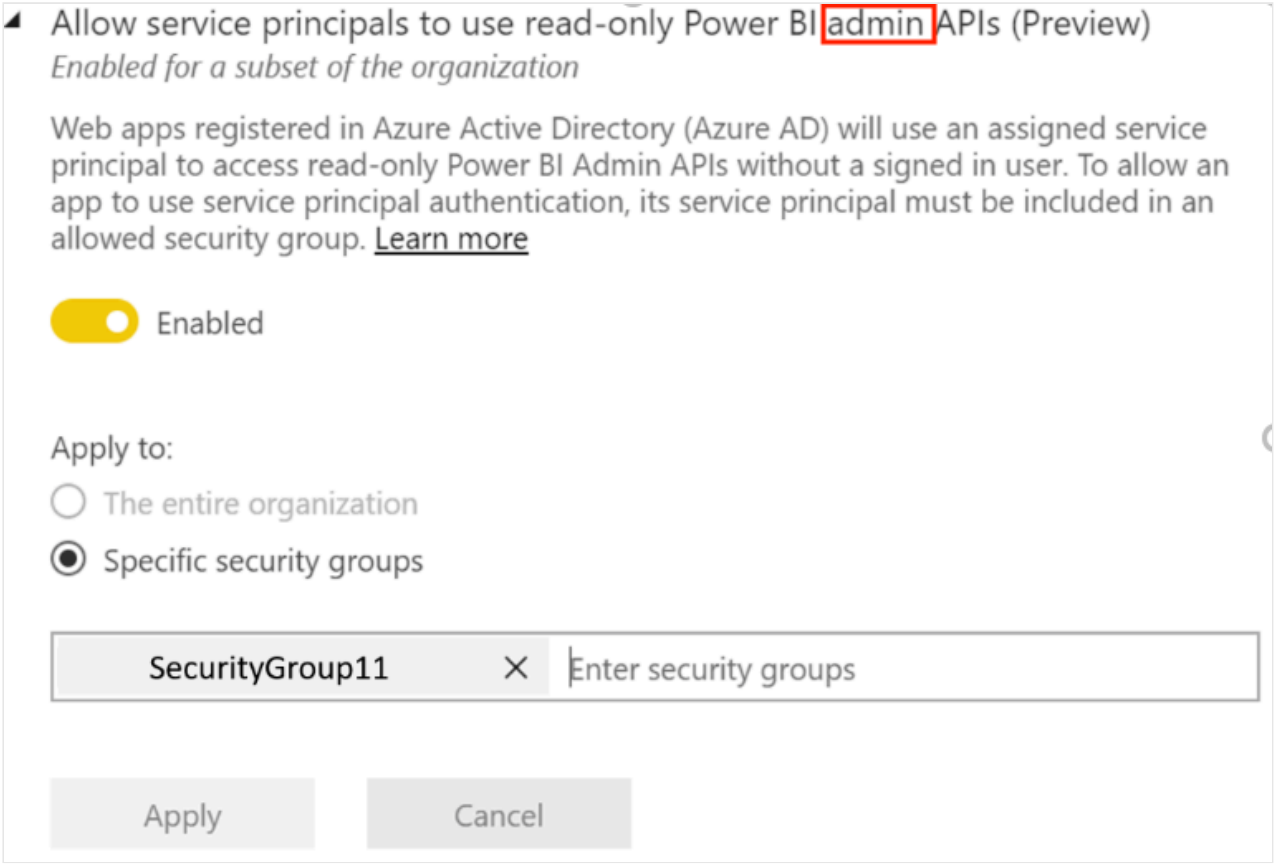
1. Log into the Power BI admin portal.

2. Select the **Tenant settings** page.

> **Important**
>
> You need to be a Power BI Admin to see the tenant settings page.

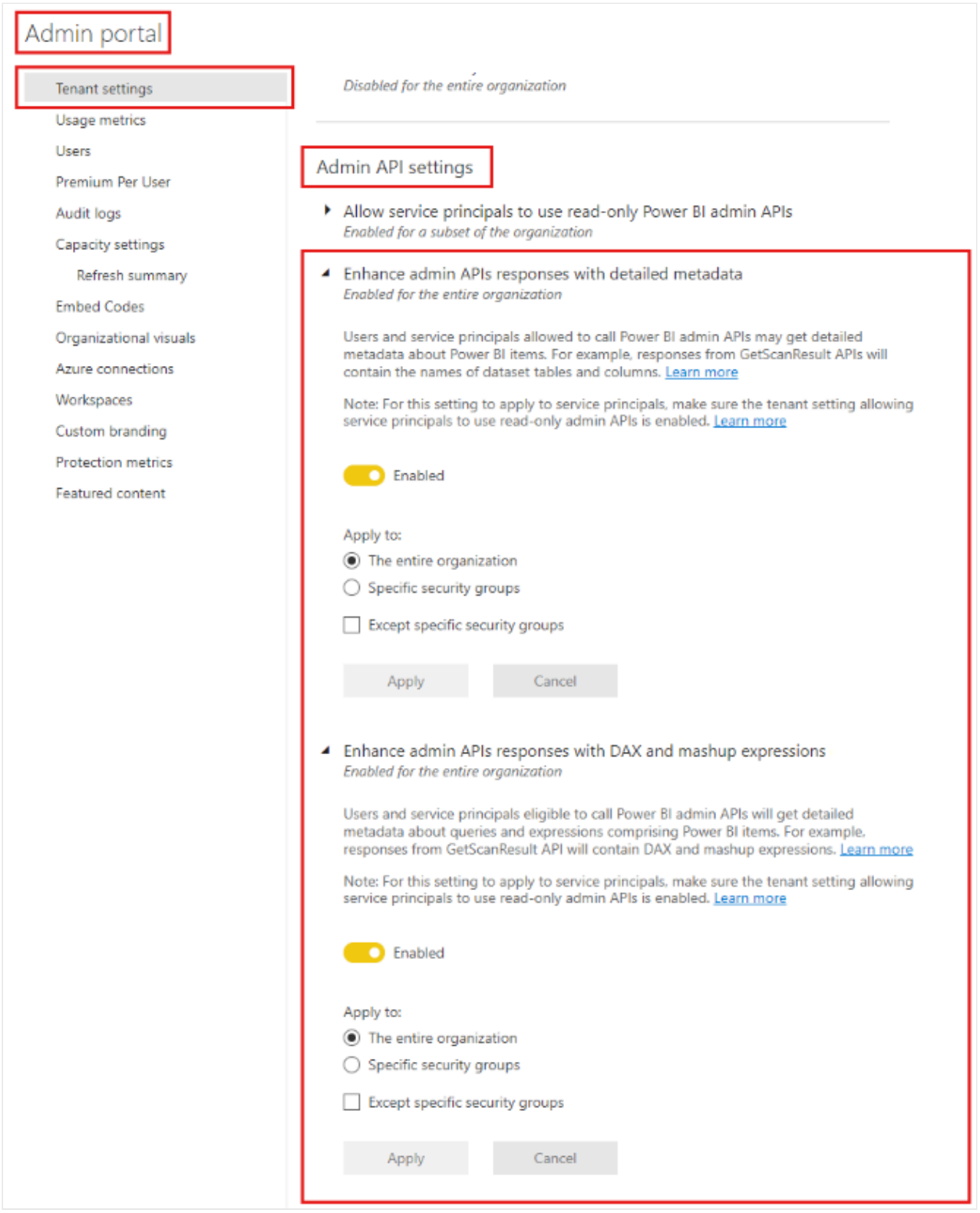
3. Select **Admin API settings** > **Allow service principals to use read-only Power BI admin APIs (Preview)**.

4. Select **Specific security groups**.



5. Select **Admin API settings** > **Enhance admin APIs responses with detailed metadata** and **Enhance admin APIs responses with DAX and mashup expressions** > Enable the toggle to allow Microsoft Purview Data Map automatically discover the detailed metadata of Power BI datasets as part of its scans.

   > **Important**
   >
   > After you update the Admin API settings on your power bi tenant, wait around 15 minutes before registering a scan and test connection.



⊗ Caution

When you allow the security group you created (that has your Microsoft Purview managed identity as a member) to use read-only Power BI admin APIs, you also allow it to access the metadata (e.g. dashboard and report names, owners, descriptions, etc.) for all of your Power BI artifacts in this tenant. Once the metadata has been pulled into the Microsoft Purview, Microsoft Purview's permissions, not Power BI permissions, determine who can see that metadata.

> **Note**
>
> You can remove the security group from your developer settings, but the metadata previously extracted won't be removed from the Microsoft Purview account. You can delete it separately, if you wish.

Scan cross-tenant Power BI

Delegated authentication and service principals are the only supported options for cross-tenant scanning. You can scan using either:

- The Azure integration runtime
- A self-hosted integration runtime

Create scan for cross-tenant using Azure IR with delegated authentication

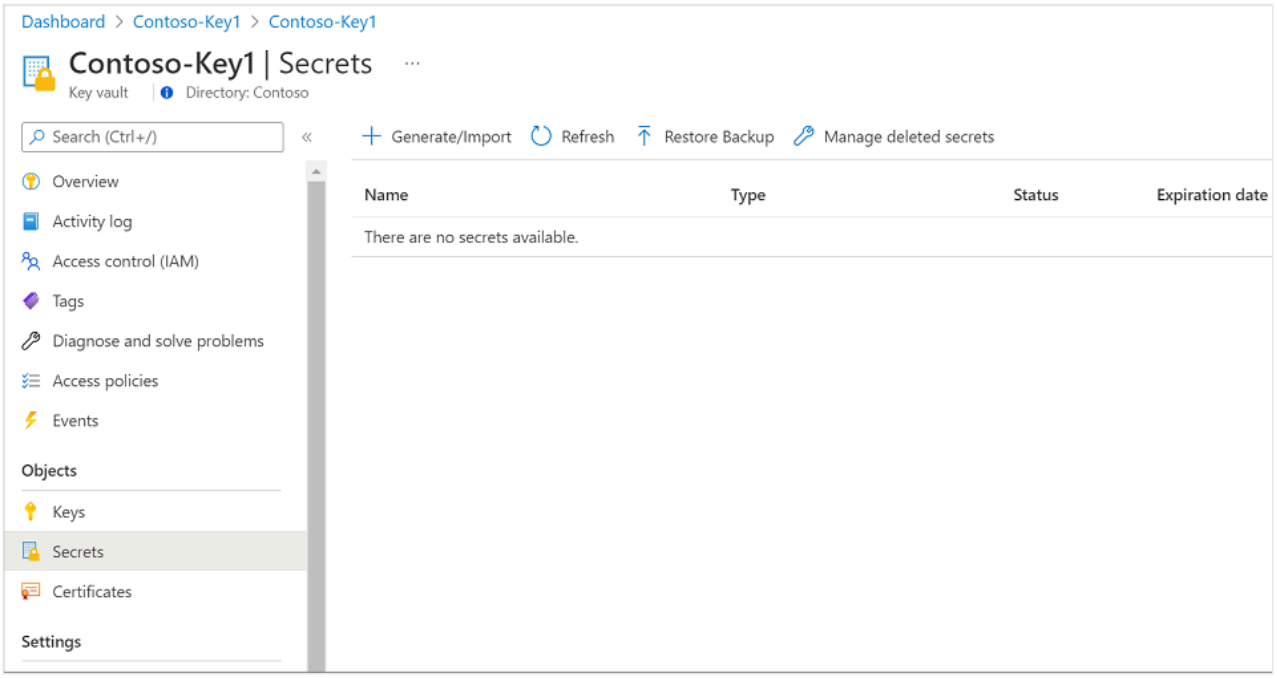
To create and run a new scan by using the Azure runtime, perform the following steps:

1. Create a user account in the Microsoft Entra tenant where the Power BI tenant is located, and assign the user to this role: **Fabric Administrator**. Take note of the username and sign in to change the password.

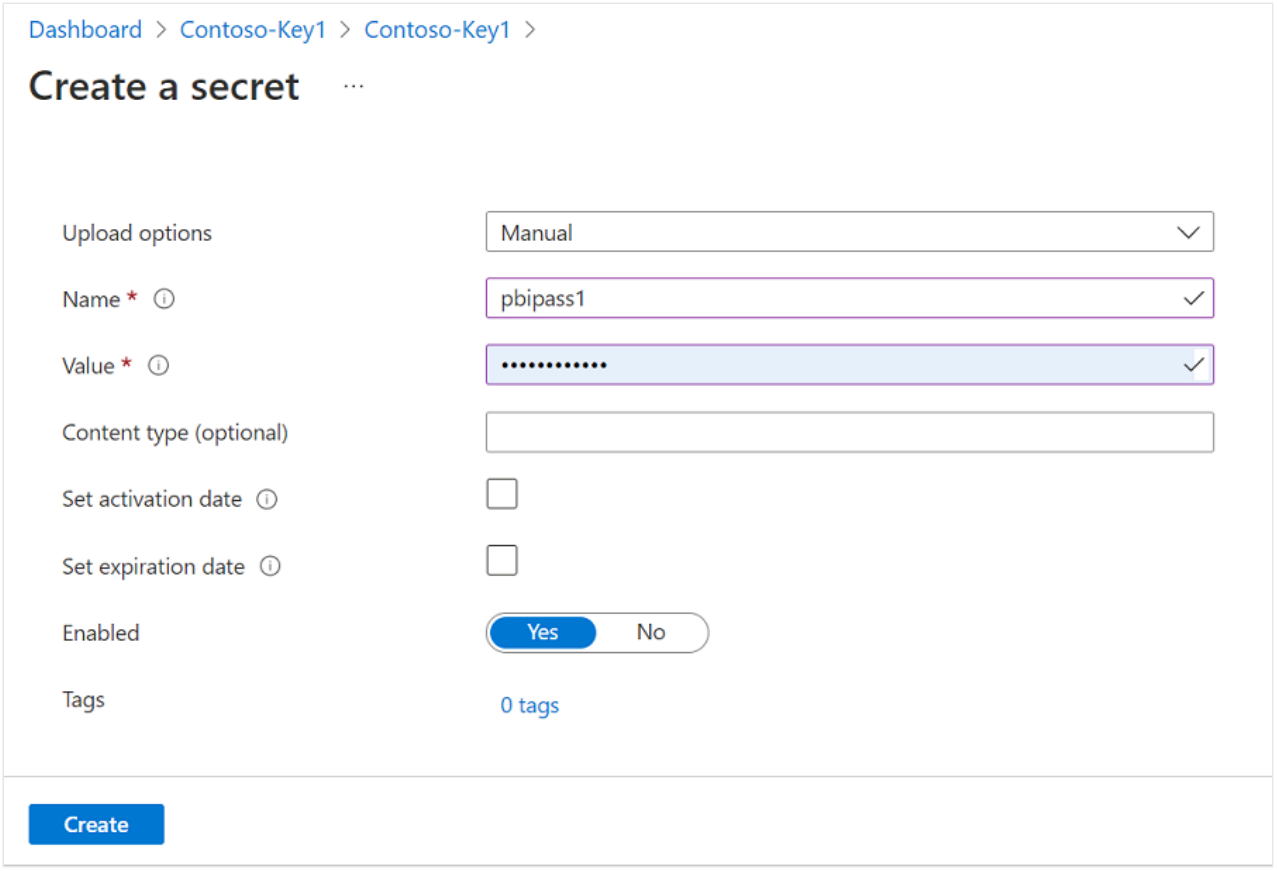
2. Assign the proper Power BI license to the user.

3. Go to the instance of Azure Key Vault in the tenant where Microsoft Purview is created.

4. Select **Settings** > **Secrets**, and then select **+ Generate/Import**.

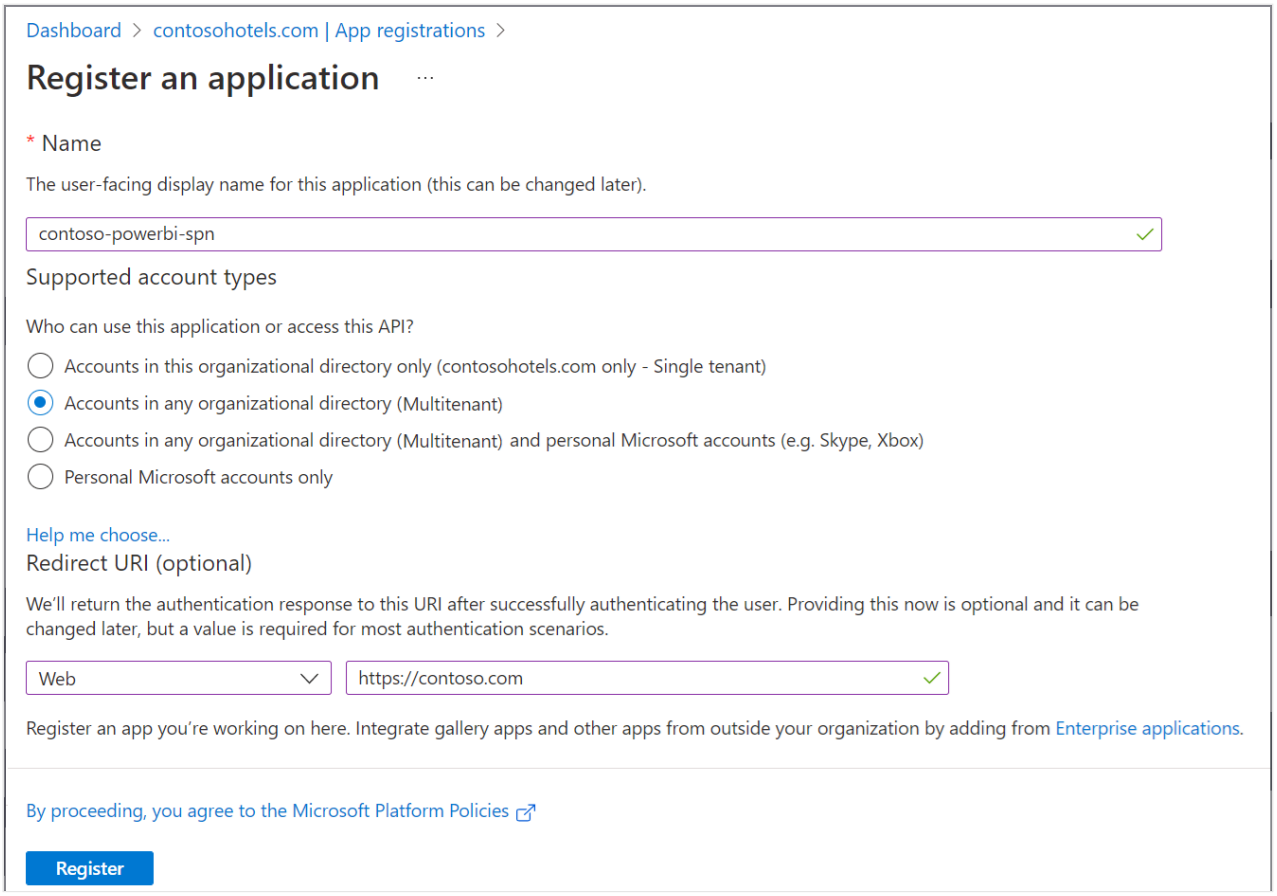


5. Enter a name for the secret. For **Value**, type the newly created password for the Microsoft Entra user. Select **Create** to complete.

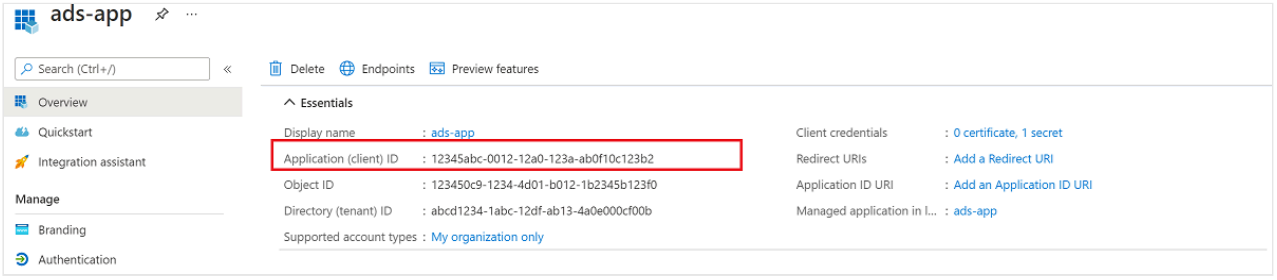


6. If your key vault isn't connected to Microsoft Purview yet, you need to create a new key vault connection.

7. Create an app registration in your Microsoft Entra tenant where Power BI is located. Provide a web URL in the **Redirect URI**.

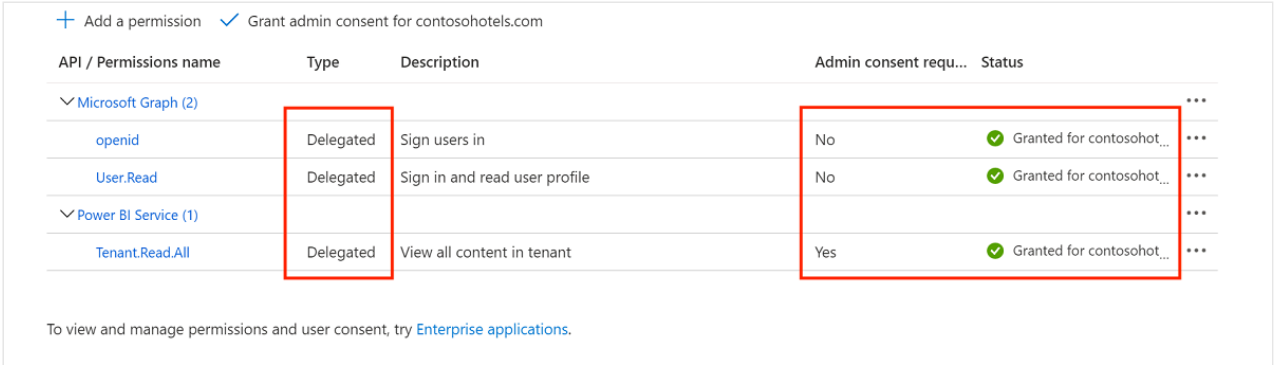


8. Take note of the client ID (app ID).

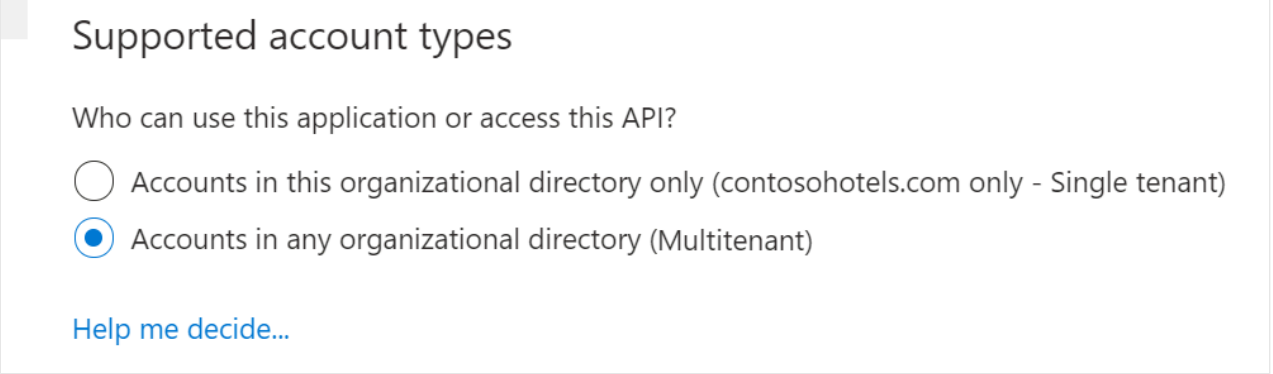


9. From the Microsoft Entra dashboard, select the newly created application, and then select **App permissions**. Assign the application the following delegated permissions, and grant admin consent for the tenant:

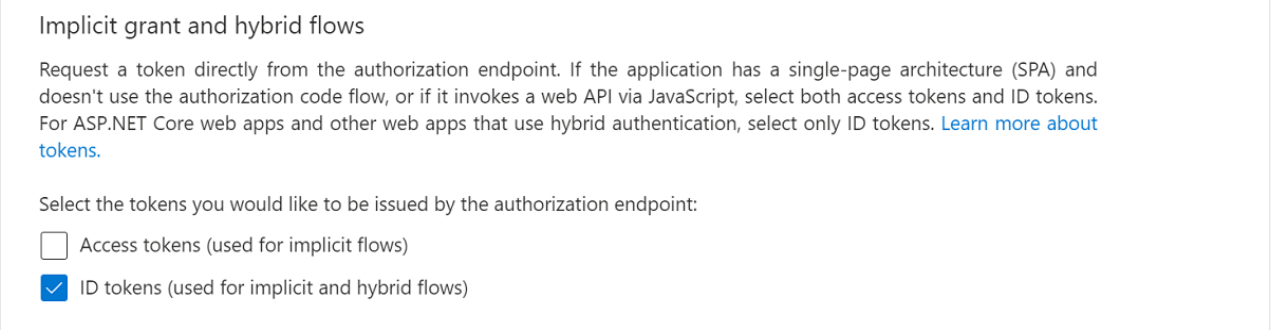
   - Power BI Service Tenant.Read.All
   - Microsoft Graph openid
   - Microsoft Graph User.Read



10. From the Microsoft Entra dashboard, select the newly created application, and then select **Authentication**. Under **Supported account types**, select **Accounts in any organizational directory (Any Microsoft Entra directory - Multitenant)**.



11. Under **Implicit grant and hybrid flows**, select **ID tokens (used for implicit and hybrid flows)**.



12. Under **Advanced settings**, enable **Allow Public client flows**.

13. In the Microsoft Purview Studio, go to the **Data map** in the left menu. Go to **Sources**.

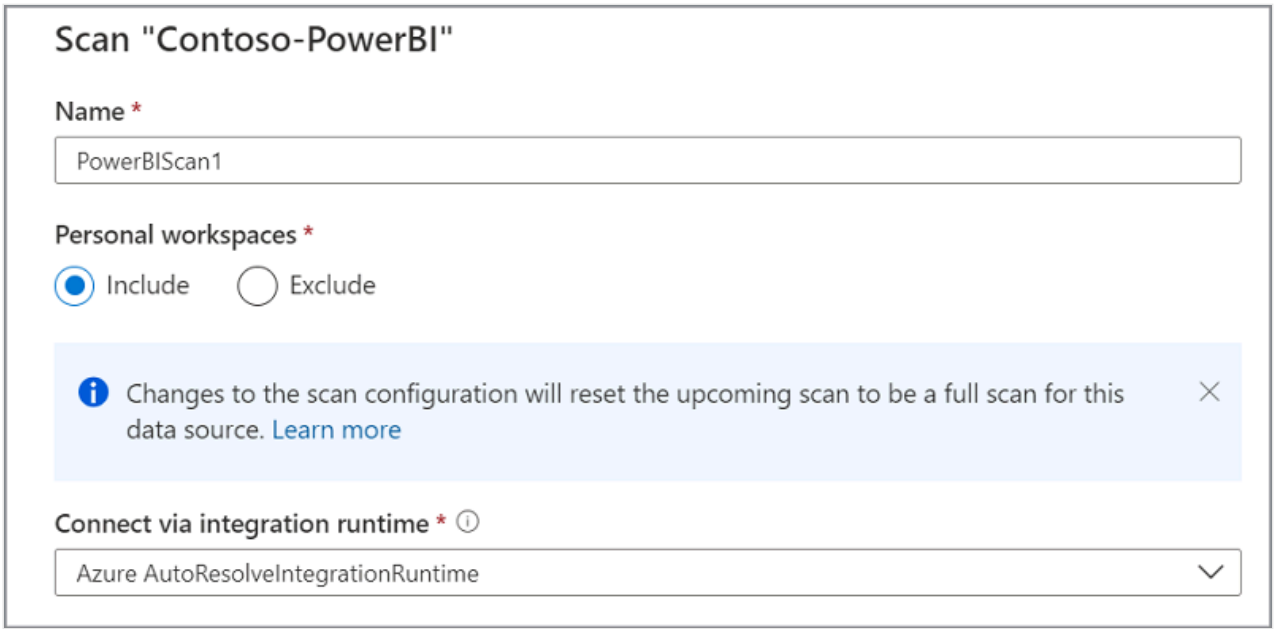
14. Select the registered Power BI source from cross-tenant.

15. Select **+ New scan**.

16. Give your scan a name. Then select the option to include or exclude the personal workspaces.

> **Note**
>
> If you switch the configuration of a scan to include or exclude a personal workspace, you trigger a full scan of the Power BI source.

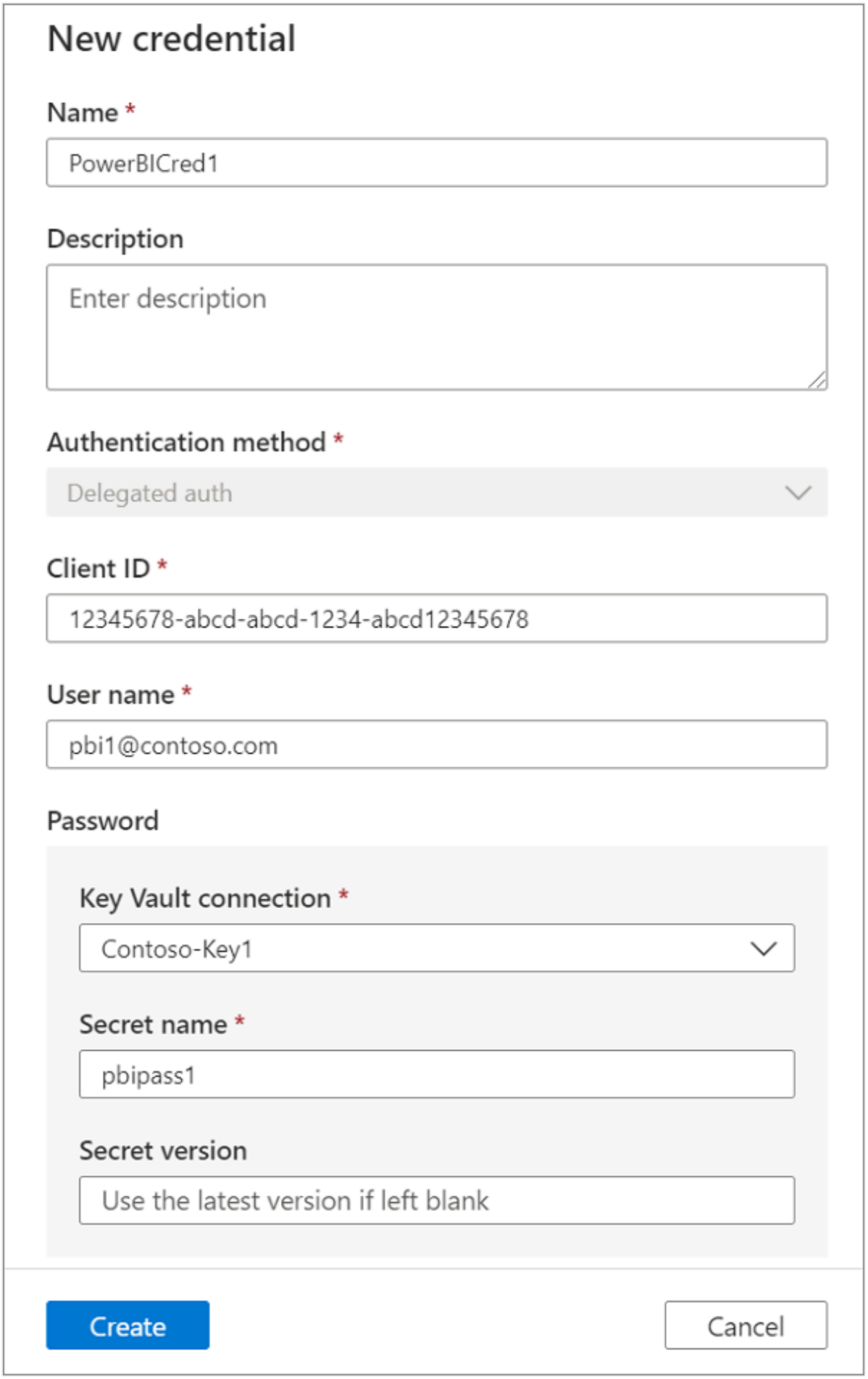
17. Select **Azure AutoResolveIntegrationRuntime** from the dropdown list.



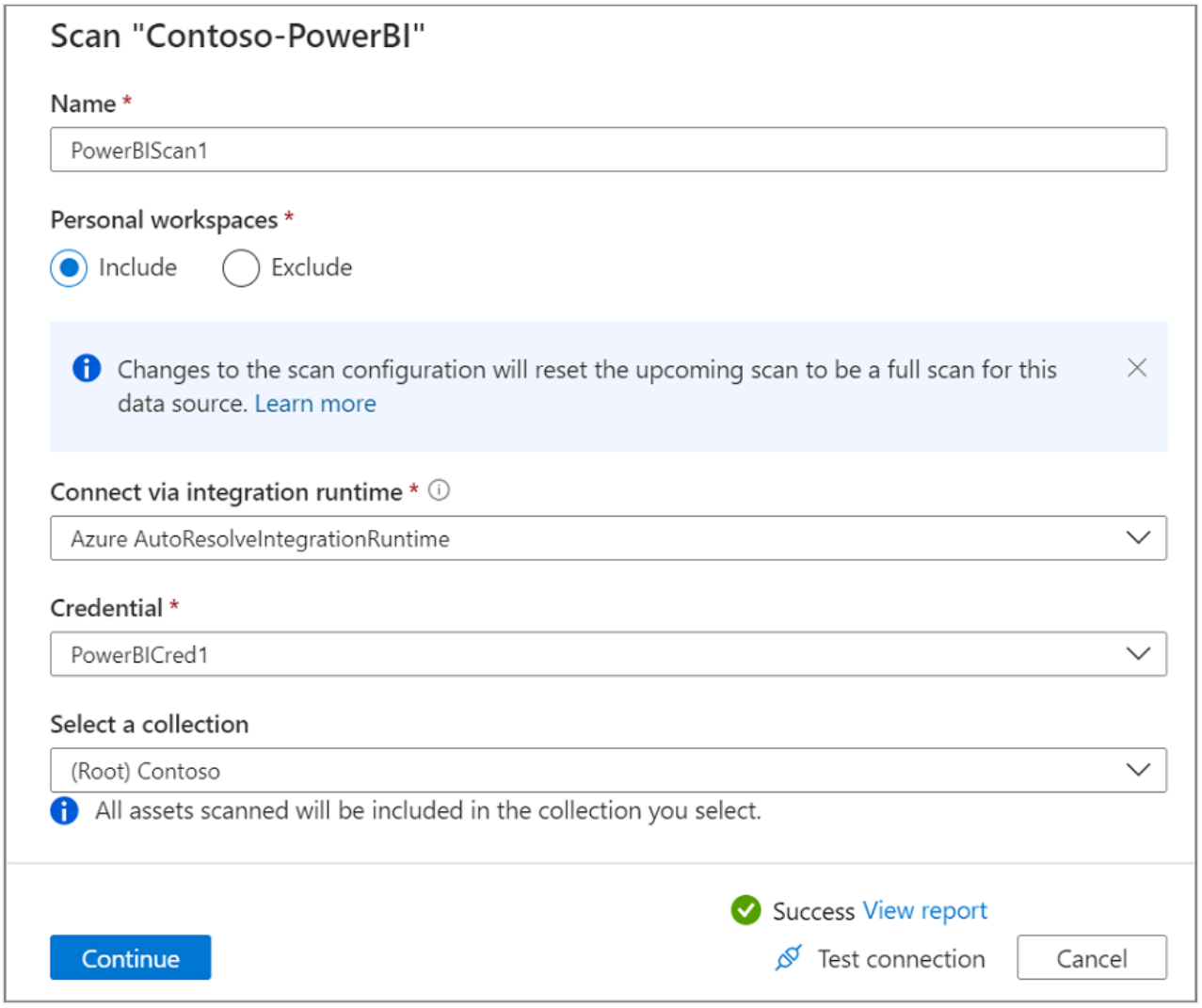
18. For the **Credential**, select **Delegated authentication**, and then select **+ New** to create a new credential.

19. Create a new credential and provide the following required parameters:

    - **Name**: Provide a unique name for the credential.
    - **Client ID**: Use the service principal client ID (app ID) that you created earlier.
    - **User name**: Provide the username of the Fabric administrator that you created earlier.
    - **Password**: Select the appropriate Key Vault connection, and the **Secret name** where the Power BI account password was saved earlier.



20. Select **Test connection** before continuing to the next steps.



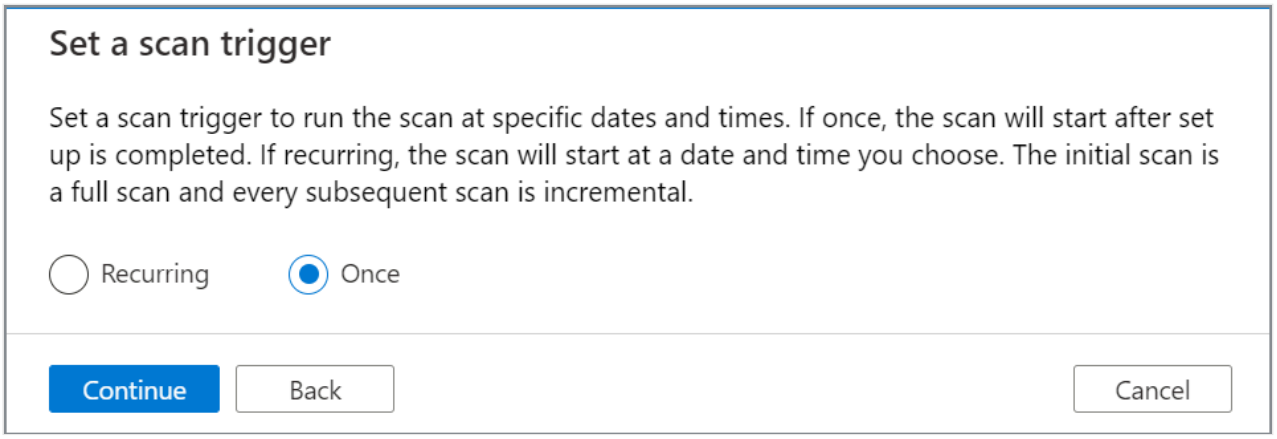
If the test fails, select **View Report** to see the detailed status and troubleshoot the problem:

a. *Access - Failed* status means that the user authentication failed. Validate if the username and password are correct. Review if the credential contains the correct client (app) ID from the app registration.

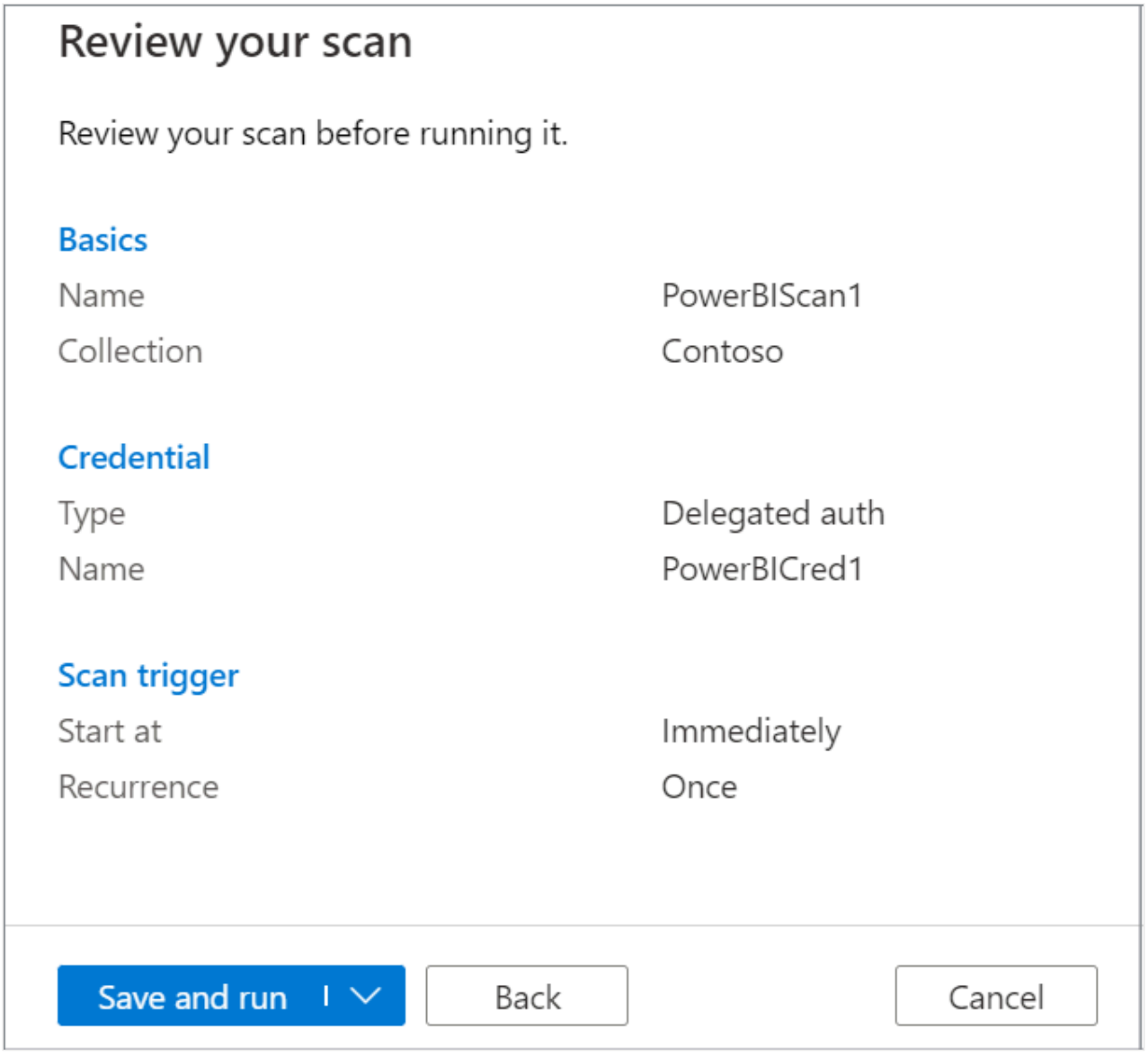
b. *Assets (+ lineage) - Failed* status means that the authorization between Microsoft Purview and Power BI has failed. Make sure that the user is added to the Fabric Administrator role (former Power BI Administrator role), and has the proper Power BI license assigned.

c. *Detailed metadata (Enhanced) - Failed* status means that the Power BI admin portal is disabled for the following setting: **Enhance admin APIs responses with detailed metadata**.

21. Set up a scan trigger. Your options are **Recurring** or **Once**.



22. On **Review new scan**, select **Save and run** to launch your scan.



Tip

To troubleshoot any issues with scanning:

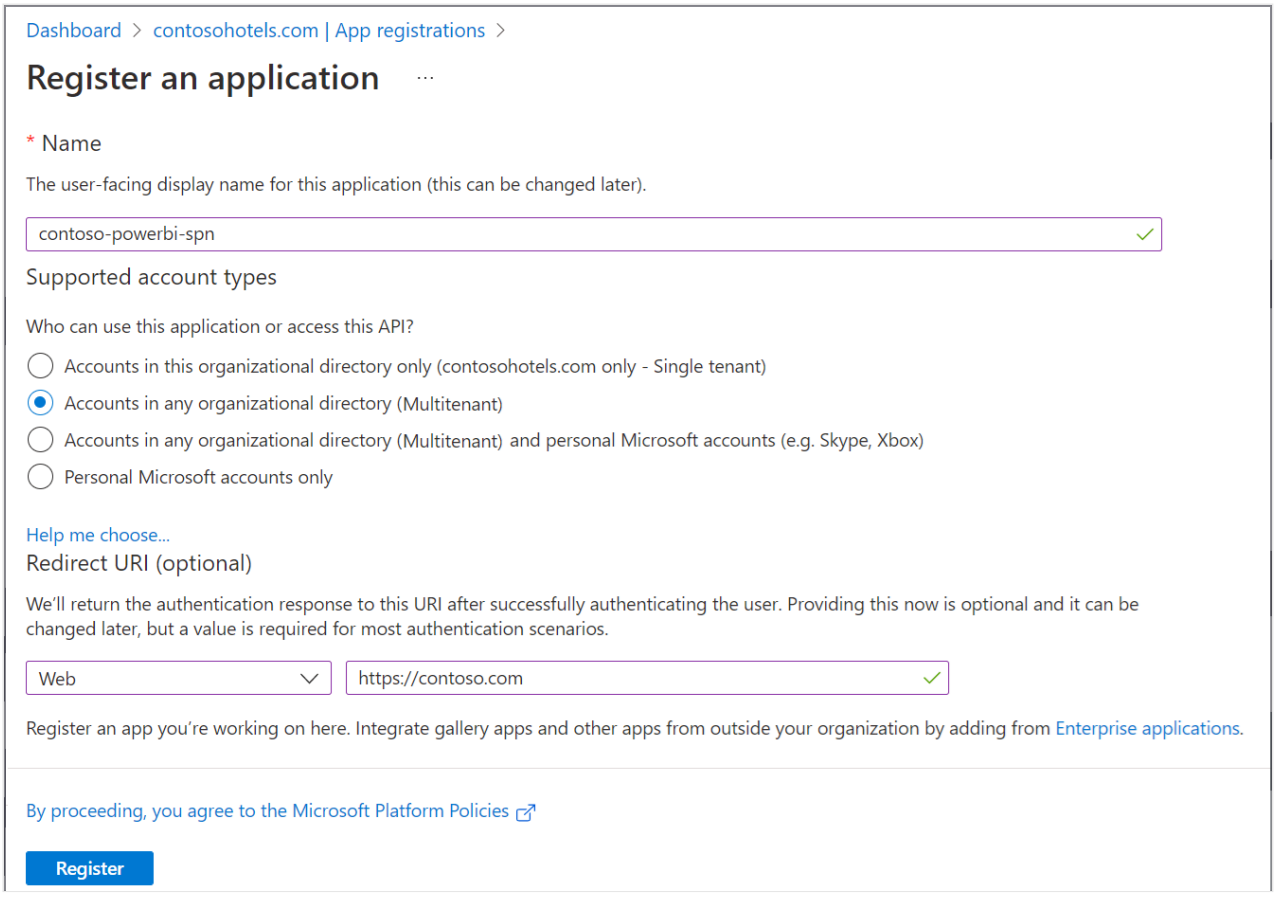
1. Confirm you have completed the deployment checklist for your scenario.

2. Review our scan troubleshooting documentation.

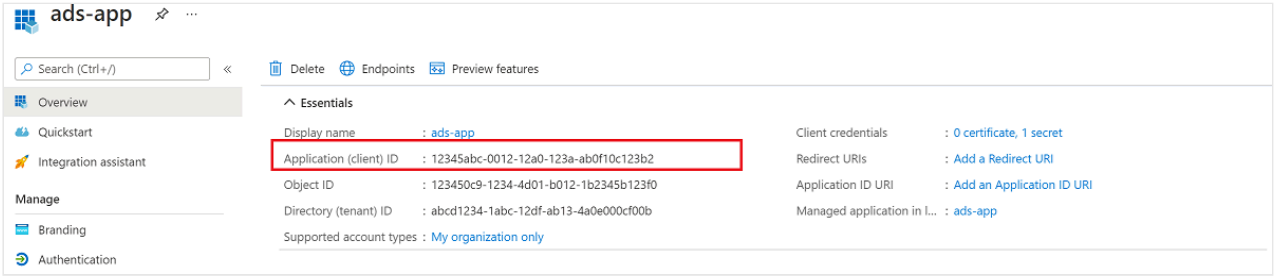
Create scan for cross-tenant using self-hosted IR with service principal

To create and run a new scan by using the self-hosted integration runtime, perform the following steps:

1. Create an app registration in your Microsoft Entra tenant where Power BI is located. Provide a web URL in the **Redirect URI**.



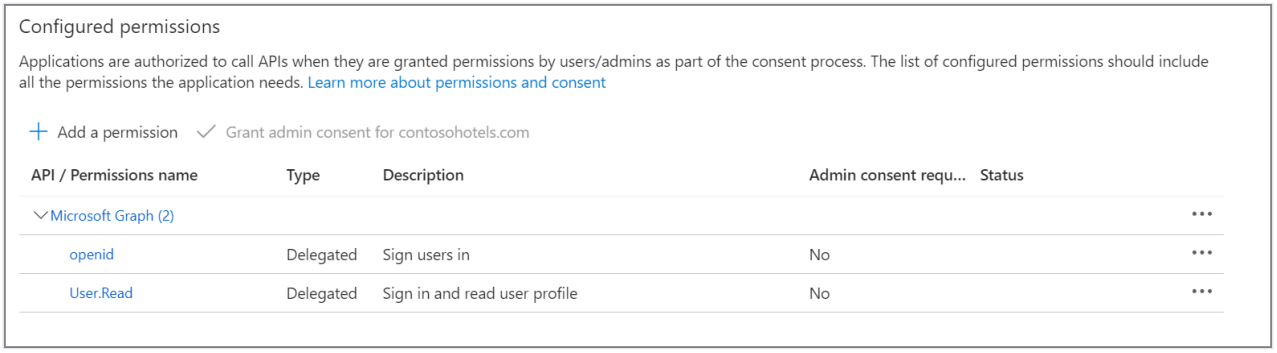
2. Take note of the client ID (app ID).



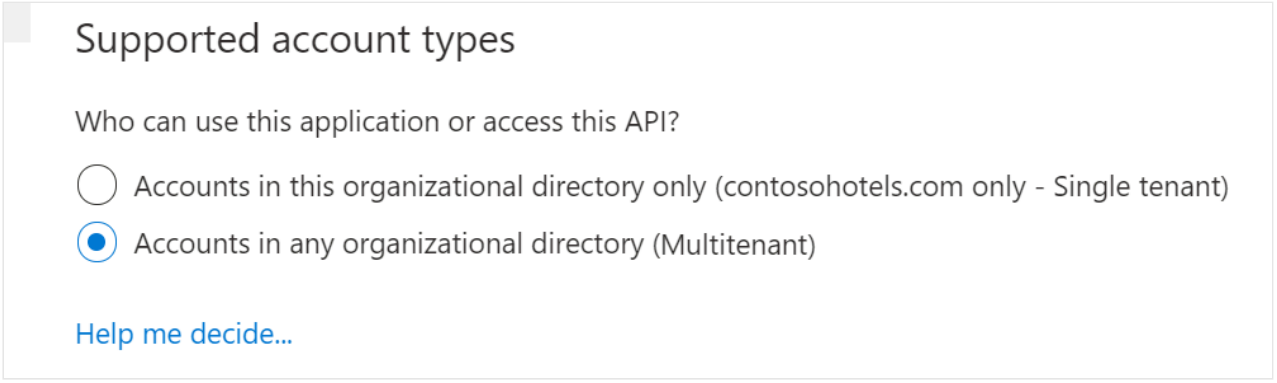
3. From the Microsoft Entra dashboard, select the newly created application, and then select **App permissions**. Assign the application the following delegated permissions:

   - Microsoft Graph openid

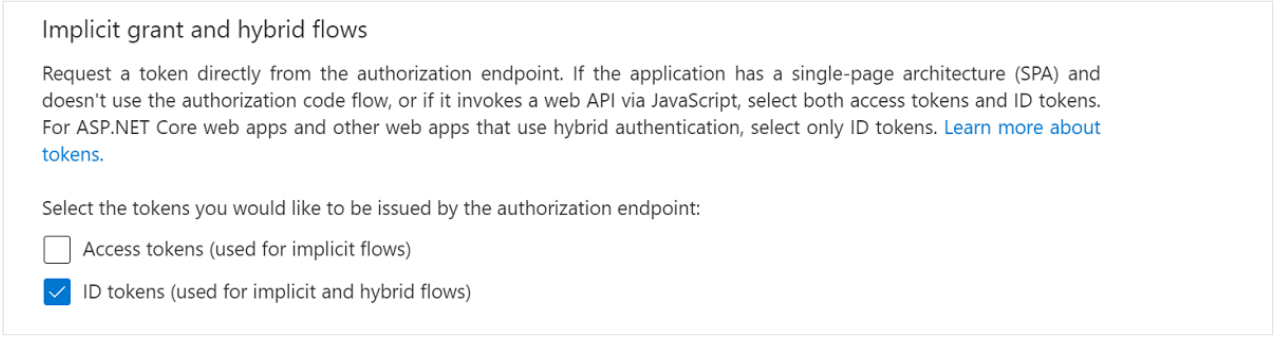
- Microsoft Graph User.Read



4. From the Microsoft Entra dashboard, select the newly created application, and then select **Authentication**. Under **Supported account types**, select **Accounts in any organizational directory (Any Microsoft Entra directory - Multitenant)**.



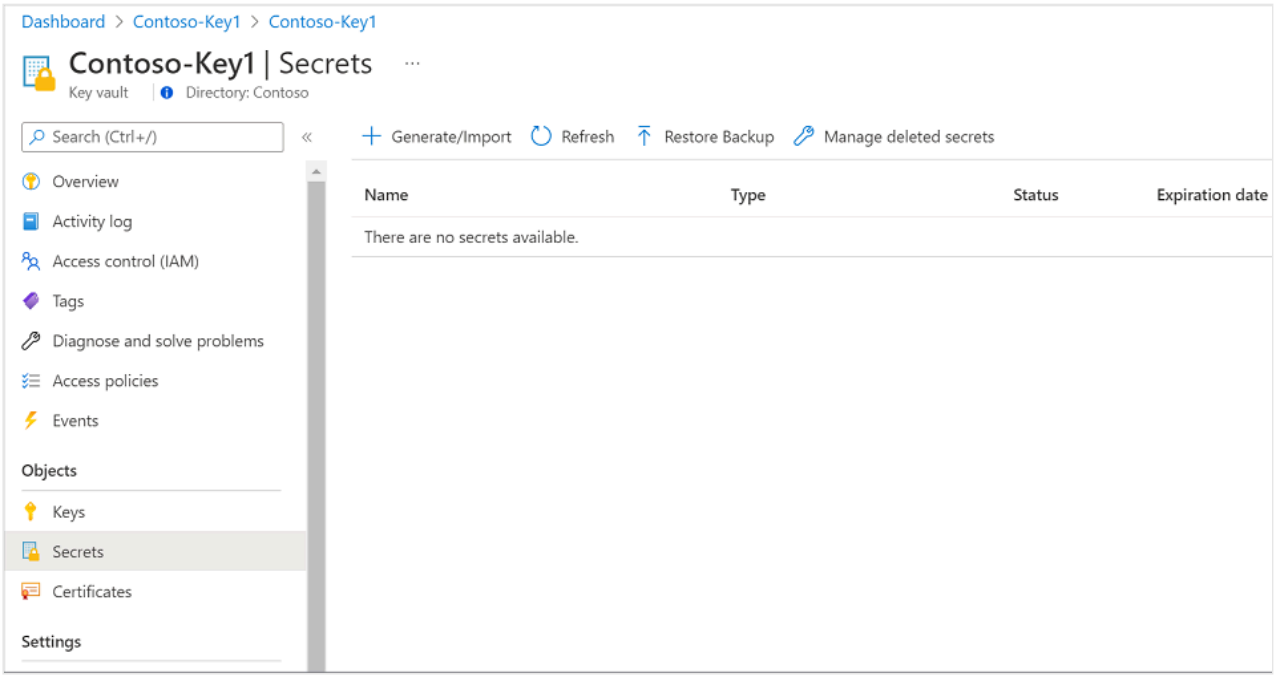
5. Under **Implicit grant and hybrid flows**, select **ID tokens (used for implicit and hybrid flows)**.



6. Under **Advanced settings**, enable **Allow Public client flows**.

7. In the tenant where Microsoft Purview is created go to the instance of Azure Key Vault.

8. Select **Settings** > **Secrets**, and then select **+ Generate/Import**.

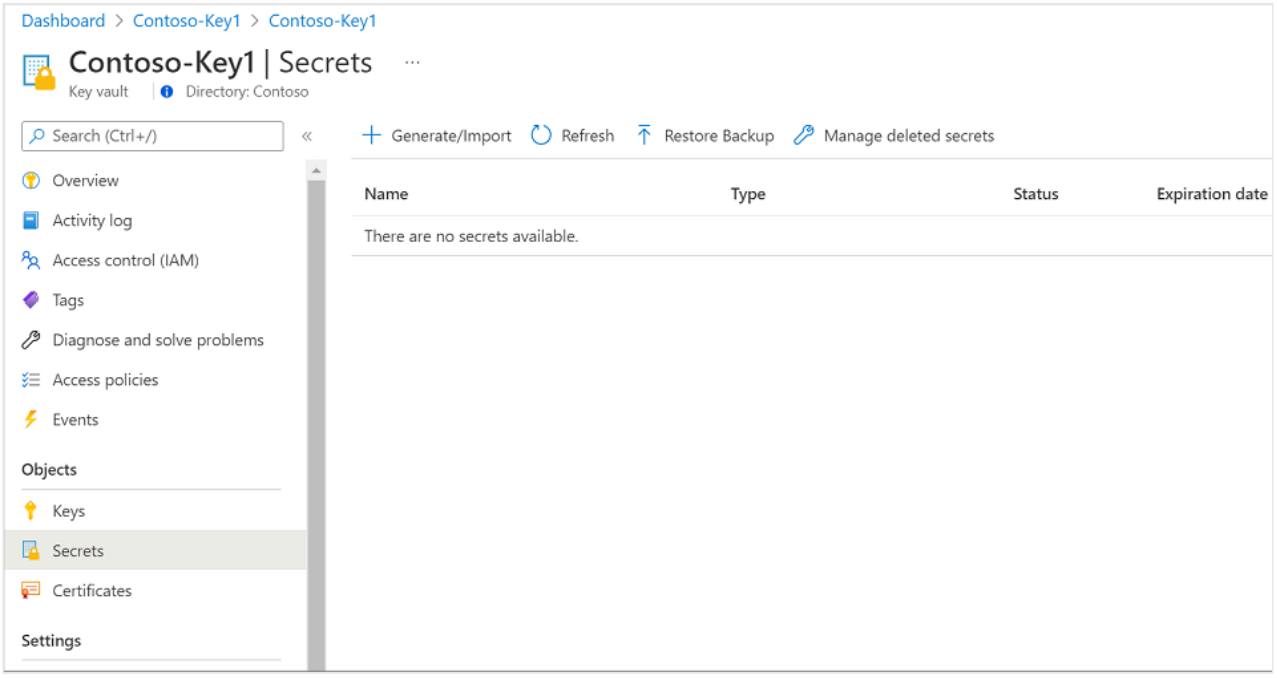


9. Enter a name for the secret. For **Value**, type the newly created secret for the App registration. Select **Create** to complete.

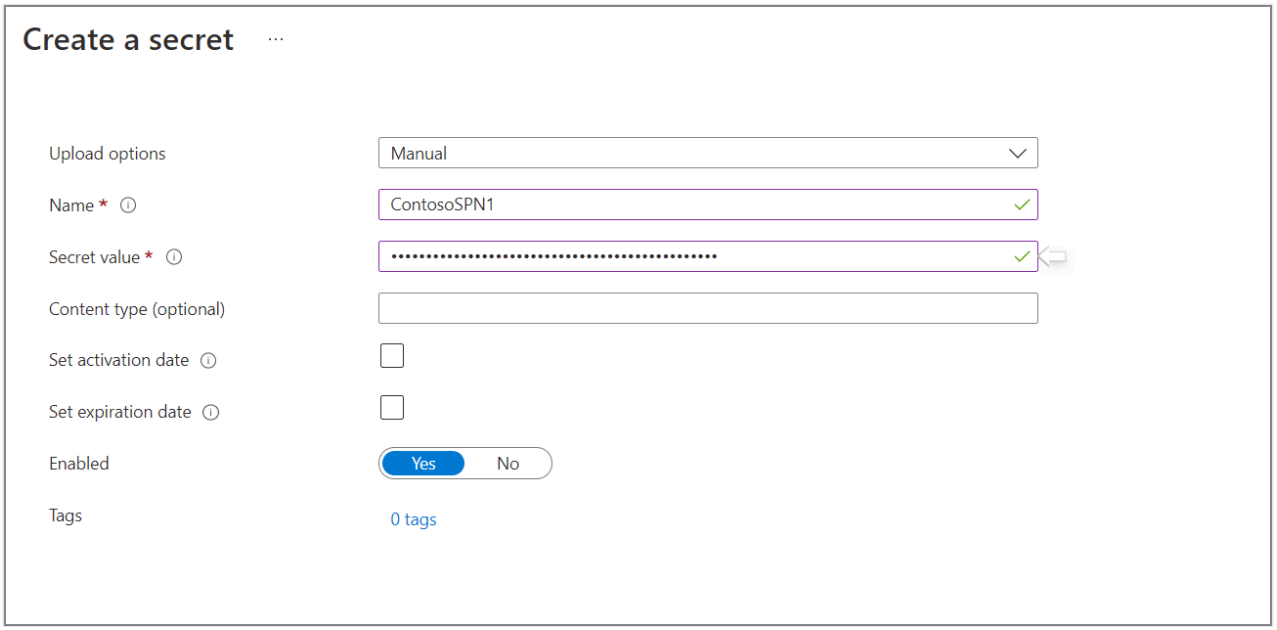
10. Under **Certificates & secrets**, create a new secret and save it securely for next steps.

11. In Azure portal, navigate to your Azure key vault.

12. Select **Settings** > **Secrets** and select + **Generate/Import**.



13. Enter a name for the secret and for **Value**, type the newly created secret for the App registration. Select **Create** to complete.



14. If your key vault isn't connected to Microsoft Purview yet, you need to create a new key vault connection.

15. In the Microsoft Purview Studio, go to the **Data map** in the left menu. Go to **Sources**.

16. Select the registered Power BI source from cross-tenant.

17. Select + **New scan**.

18. Give your scan a name. Then select the option to include or exclude the personal workspaces.

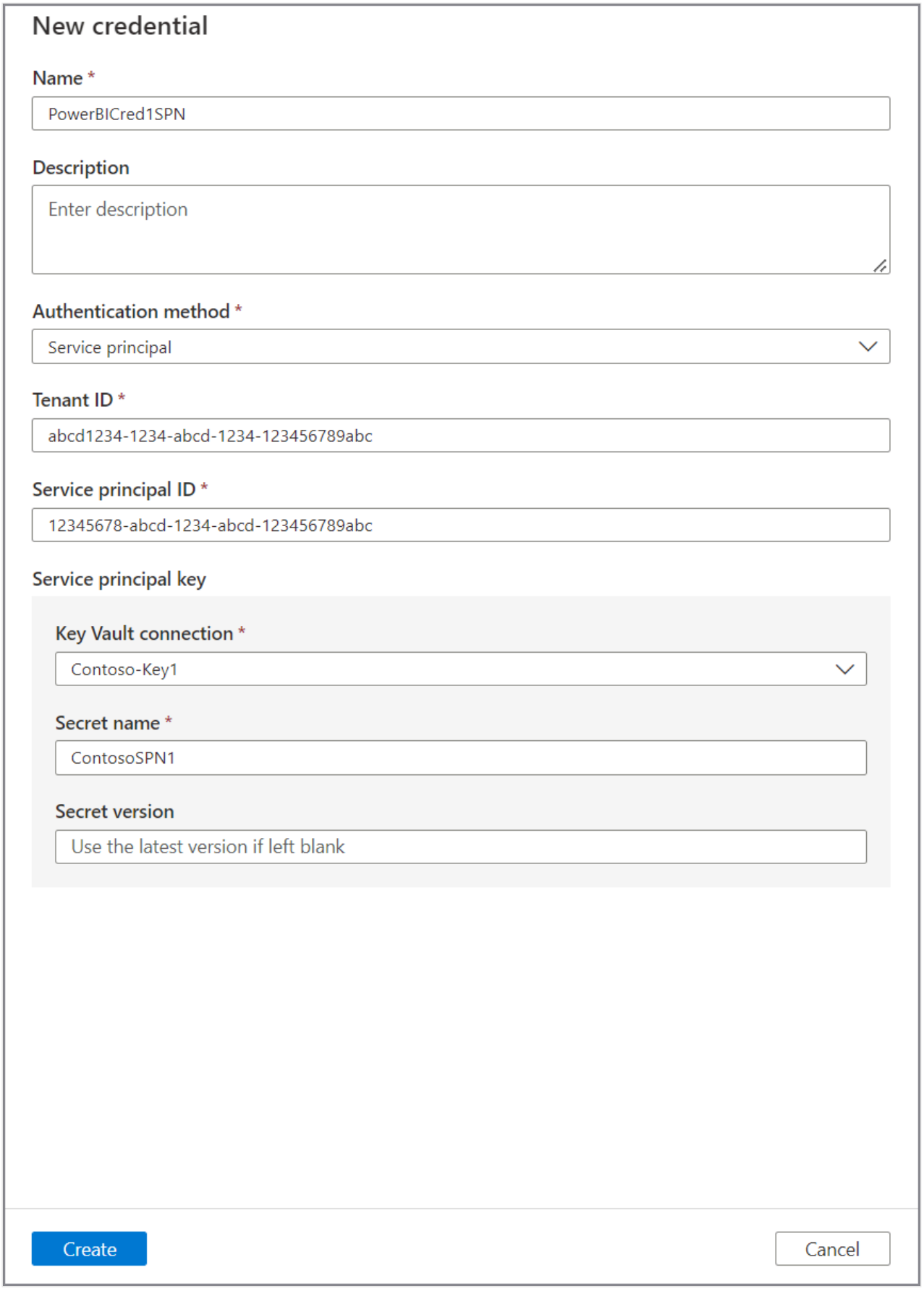
    > **Note**
    >
    > If you switch the configuration of a scan to include or exclude a personal workspace, you trigger a full scan of the Power BI source.

19. Select your self-hosted integration runtime from the drop-down list.

20. For the **Credential**, select **Service Principal**, and then select + **New** to create a new credential.

21. Create a new credential and provide the following required parameters:

    - **Name**: Provide a unique name for credential
    - **Authentication method**: Service principal
    - **Tenant ID**: Your Power BI tenant ID
    - **Client ID**: Use Service Principal Client ID (App ID) you created earlier

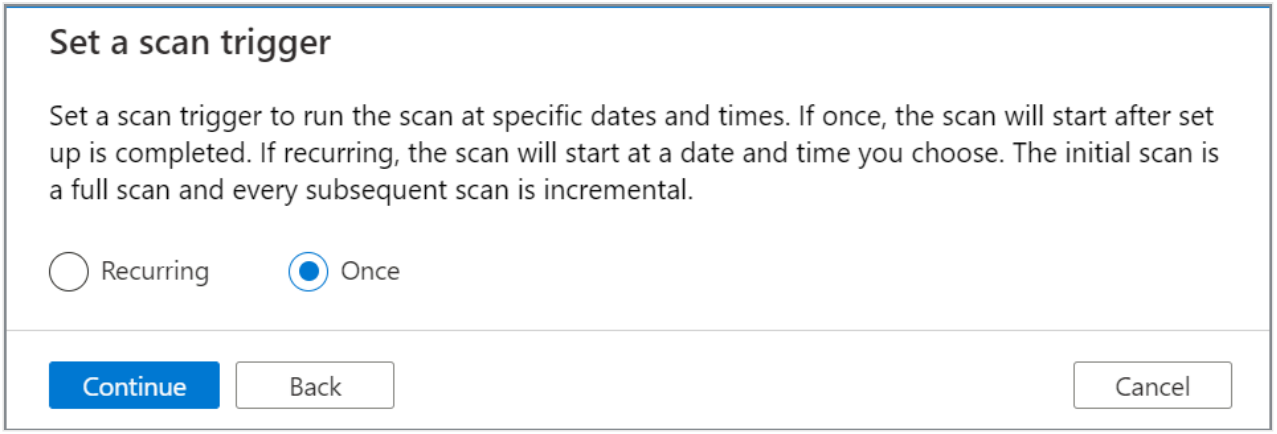


22. Select **Test connection** before continuing to the next steps.

    If the test fails, select **View Report** to see the detailed status and troubleshoot the problem:

a. *Access - Failed* status means that the user authentication failed. Validate if the App ID and secret are correct. Review if the credential contains the correct client (app) ID from the app registration.
b. *Assets (+ lineage) - Failed* status means that the authorization between Microsoft Purview and Power BI has failed. Make sure that the user is added to the Fabric Administrator role (former Power BI Administrator role), and has the proper Power BI license assigned.
c. *Detailed metadata (Enhanced) - Failed* status means that the Power BI admin portal is disabled for the following setting: **Enhance admin APIs responses with detailed metadata**.

23. Set up a scan trigger. Your options are **Recurring** or **Once**.



24. On **Review new scan**, select **Save and run** to launch your scan.

> **Tip**
>
> To troubleshoot any issues with scanning:
>
> 1. Confirm you have completed the deployment checklist for your scenario.
> 2. Review our scan troubleshooting documentation.

Deployment checklist

The deployment checklist is a summary of all the steps you'll need to take to set up a cross-tenant Power BI source. You can use it during setup or for troubleshooting, to confirm you've followed all the necessary steps to connect.

Public access with Azure IR

Scan cross-tenant Power BI by using delegated authentication in a public network

1. Make sure the Power BI tenant ID is entered correctly during the registration. By default, the Power BI tenant ID that exists in the same Microsoft Entra instance as Microsoft Purview will be populated.

2. Make sure your Power BI metadata model is up to date by enabling metadata scanning.

3. From the Azure portal, validate if the Microsoft Purview account network is set to **public access**.

4. From the Power BI tenant admin portal, make sure the Power BI tenant is configured to allow a public network.

5. Check your instance of Azure Key Vault to make sure:
   a. There are no typos in the password or secret.
   b. Microsoft Purview managed identity has **get** and **list** access to secrets.

6. Review your credential to validate that the:
   a. Client ID matches the *Application (Client) ID* of the app registration.
   b. For **delegated auth**, username includes the user principal name, such as `johndoe@contoso.com`.

7. In the Power BI Microsoft Entra tenant, validate the following Power BI admin user settings:
   a. The user is assigned to the Fabric Administrator role (former Power BI Administrator role).
   b. At least one Power BI license is assigned to the user.
   c. If the user is recently created, sign in with the user at least once, to make sure that the password is reset successfully, and the user can successfully initiate the session.
   d. There are no multifactor authentication or conditional access policies enforced on the user.

8. In the Power BI Microsoft Entra tenant, validate the following app registration settings:
   a. The app registration exists in your Microsoft Entra tenant where the Power BI tenant is located.
   b. If service principal is used, under **API permissions**, the following **delegated permissions** are assigned with read for the following APIs:

      - Microsoft Graph openid

- Microsoft Graph User.Read

c. If delegated authentication is used, under **API permissions**, the following **delegated permissions** and **grant admin consent for the tenant** is set up with read for the following APIs:

- Power BI Service Tenant.Read.All
- Microsoft Graph openid
- Microsoft Graph User.Read

d. Under **Authentication**:

i. **Supported account types** > **Accounts in any organizational directory (Any Microsoft Entra directory - Multitenant)** is selected.

ii. **Implicit grant and hybrid flows** > **ID tokens (used for implicit and hybrid flows)** is selected.

iii. **Allow public client flows** is enabled.

9. In Power BI tenant, from Microsoft Entra tenant, make sure Service Principal is member of the new security group.

10. On the Power BI Tenant Admin portal, validate if Allow service principals to use read-only Power BI admin APIs is enabled for the new security group.

Next steps

Now that you've registered your source, see the following guides to learn more about Microsoft Purview and your data.

- Data estate insights in Microsoft Purview
- Lineage in Microsoft Purview
- Search Unified Catalog

Troubleshoot Power BI tenant scans in Microsoft Purview

Article • 05/21/2025

This article explores common troubleshooting methods for scanning Power BI tenants in Microsoft Purview.

Supported scenarios for Power BI scans

Same-tenant

Expand table

Scenarios	Microsoft Purview public access allowed/denied	Power BI public access allowed /denied	Runtime option	Authentication option	Deployment checklist
Public access with Azure IR	Allowed	Allowed	Azure Runtime	Microsoft Purview Managed Identity	Review deployment checklist
Public access with self-hosted IR	Allowed	Allowed	Self-hosted runtime	Delegated authentication / Service principal	Review deployment checklist
Private access	Allowed	Denied	Self-hosted runtime	Delegated authentication / Service principal	Review deployment checklist
Private access	Denied	Allowed	Self-hosted runtime	Delegated authentication / Service principal	Review deployment checklist
Private access	Denied	Denied	Self-hosted runtime	Delegated authentication / Service principal	Review deployment checklist

Cross-tenant

Expand table

Scenarios	Microsoft Purview public access allowed/denied	Power BI public access allowed /denied	Runtime option	Authentication option	Deployment checklist
Public access with Azure IR	Allowed	Allowed	Azure runtime	Delegated Authentication	Deployment checklist
Public access with Self-hosted IR	Allowed	Allowed	Self-hosted runtime	Delegated authentication / Service principal	Deployment checklist

Troubleshooting tips

If delegated auth is used:

- Check your key vault. Make sure there are no typos in the password.
- Assign proper Power BI license to Power BI administrator user.
- Validate if user is assigned to Power BI Administrator role.
- If user is recently created, make sure password is reset successfully and user can successfully initiate the session.

My schema is not showing up after scanning

It can take some time for schema to finish the scanning and ingestion process, depending on the size of your Power BI Tenant. Currently if you have a large PowerBI tenant, this process could take a few hours.

Error code: Test connection failed - AASDST50079

- **Message**: `Failed to get access token with given credential to access Power BI tenant. Authentication type PowerBIDelegated Message: AASDST50079 Due to a configuration change made by your administrator or because you moved to a new location, you must enroll in multi-factor authentication.`
- **Cause**: Authentication is interrupted, due multi-factor authentication requirement for the Power BI admin user.
- **Recommendation**: Disable multi-factor authentication requirement and exclude user from conditional access policies. Login with the user to Power BI dashboard to validate if user can successfully login to the application.

Error code: Test connection failed - AASTS70002

- **Message**: `Failed to access token with given credential to access Power BI tenant. Authentication type: PowerBiDelegated Message AASTS70002: The request body must contain the following parameter: 'client_assertion' or 'client_secret'.`
- **Cause**: If Delegated Authentication is used, the problem could be a misconfiguration on the app registration.
- **Recommendation**: Review Power BI deployment checklist based on your scenario.

Error code: Test connection failed - Detailed metadata

- **Message**: `Failed to enable the PowerBI administrator API to fetch basic metadata and lineage.`
- **Cause**: **Allow service principals to use read-only Power BI admin APIs** is disabled.
- **Recommendation**: Under Power BI Admin portal, enable **Allow service principals to use read-only Power BI admin APIs**.

Issue: Test Connection succeeded. No assets discovered.

- **Message**: N/A
- **Cause**: This problem can occur in same-tenant or cross-tenant scenarios, due problem with networking or authentication issues.
- **Recommendation**:
  - If Delegated Authentication is used, validate Power BI Admin user sign in logs in Microsoft Entra logs to make sure user sign in is successful. Login with the user to Power BI dashboard to validate if user can successfully login to the application.
  - Review your network configurations. Private endpoint is required for **both** Power BI tenant and Purview account, if one of these services (Power BI tenant or Microsoft Purview) is configured to block public access.

Next steps

Follow the below guides to learn more about Microsoft Purview and your data.

- Data Estate Insights in Microsoft Purview
- Lineage in Microsoft Purview
- Search Unified Catalog

Available metadata for Power BI in Unified Catalog

Article • 05/21/2025

This article lists the metadata that is available for a Power BI tenant in the Microsoft Purview Unified Catalog.

Power BI

Expand table

Metadata	Population method	Source of truth	Asset type	Editable	Upstream metadata
Classification	Manual	Microsoft Purview	All types	Yes	N/A
Sensitivity Labels	Automatic	Microsoft Purview	All types	No	
Glossary terms	Manual	Microsoft Purview	All types	Yes	N/A
Collection	Automatic	Microsoft Purview	All types	Yes	N/A
Hierarchy	Automatic	Microsoft Purview	All types	No	N/A
qualifiedName	Automatic	Microsoft Purview	All types	No	N/A
Asset Description	Automatic/Manual*	Microsoft Purview	All types	Yes	N/A
Contacts - Expert	Manual	Microsoft Purview	All types	Yes	N/A
Contacts - Owner	Manual	Microsoft Purview	All types	Yes	N/A
name	Automatic	Power BI	Power BI Dashboard	Yes	dashboard.DisplayName
isReadOnly	Automatic	Power BI	Power BI Dashboard	No	dashboard.IsReadOnly
embedUrl	Automatic	Power BI	Power BI Dashboard	No	dashboard.EmbedUrl
tileNames	Automatic	Power BI	Power BI	No	TileTitles

Metadata	Population method	Source of truth	Asset type	Editable	Upstream metadata
			Dashboard		
users	Automatic	Power BI	Power BI Dashboard	No	dashboard.Users
Lineage	Automatic	Power BI	Power BI Dashboard	No	N/A
name	Automatic	Power BI	Power BI Dataflow	Yes	dataflow.Name
description	Automatic	Power BI	Power BI Dataflow	Yes	dataflow.Description
configuredBy	Automatic	Power BI	Power BI Dataflow	No	dataflow.ConfiguredBy
modifiedBy	Automatic	Power BI	Power BI Dataflow	No	dataflow.ModifiedBy
modifiedDateTime	Automatic	Power BI	Power BI Dataflow	No	dataflow.ModifiedDateTime
Endorsement	Automatic	Power BI	Power BI Dataflow	No	dataflow.EndorsementDetails
users	Automatic	Power BI	Power BI Dataflow	No	dataflow.Users
Lineage	Automatic	Microsoft Purview	Power BI Dataflow	No	
name	Automatic	Power BI	Power BI Datamart	Yes	datamart.Name
description	Automatic	Power BI	Power BI Datamart	Yes	datamart.Description
configuredBy	Automatic	Power BI	Power BI Datamart	No	datamart.ConfiguredBy
modifiedBy	Automatic	Power BI	Power BI Datamart	No	datamart.ModifiedBy
modifiedDateTime	Automatic	Power BI	Power BI Datamart	No	datamart.ModifiedDateTime
Endorsement	Automatic	Power BI	Power BI Datamart	No	datamart.EndorsementDetails
users	Automatic	Power BI	Power BI Datamart	No	datamart.Users

Metadata	Population method	Source of truth	Asset type	Editable	Upstream metadata
Lineage	Automatic	Microsoft Purview	Power BI Datamart	No	
name	Automatic	Power BI	Power BI Dataset	Yes	dataset.Name
description	Automatic	Power BI	Power BI Dataset	Yes	dataset.Description
isRefreshable	Automatic	Power BI	Power BI Dataset	No	dataset.IsRefreshable
configuredBy	Automatic	Power BI	Power BI Dataset	No	dataset.ConfiguredBy
contentProviderType	Automatic	Power BI	Power BI Dataset	No	dataset.ContentProviderType
createdDate	Automatic	Power BI	Power BI Dataset	No	dataset.CreatedDateTime
targetStorageMode	Automatic	Power BI	Power BI Dataset	No	dataset.TargetStorageMode
Schema	Automatic/Manual	Power BI	Power BI Dataset		tables & columns
Endorsement	Automatic	Power BI	Power BI Dataset	No	dataset.EndorsementDetails
users	Automatic	Power BI	Power BI Dataset	No	dataset.Users
Lineage	Automatic	Microsoft Purview	Power BI Dataset	No	
name	Automatic	Power BI	Power BI Report	Yes	report.Name
description	Automatic	Power BI	Power BI Report	Yes	report.Description
createdDateTime	Automatic	Power BI	Power BI Report	No	report.CreatedDateTime
webUrl	Automatic	Power BI	Power BI Report	No	report.WebUrl
embedUrl	Automatic	Power BI	Power BI Report	No	report.EmbedUrl
PBIDatasetId	Automatic	Power BI	Power BI Report	No	report.DatasetId

Metadata	Population method	Source of truth	Asset type	Editable	Upstream metadata
modifiedBy	Automatic	Power BI	Power BI Report	No	report.ModifiedBy
modifiedDateTime	Automatic	Power BI	Power BI Report	No	report.ModifiedDateTime
reportType	Automatic	Power BI	Power BI Report	No	report.ReportType
Endorsement	Automatic	Power BI	Power BI Report	No	report.EndorsementDetails
users	Automatic	Power BI	Power BI Report	No	report.Users
Lineage	Automatic	Microsoft Purview	Power BI Report	No	N/A
name	Automatic	Power BI	Power BI Workspace	Yes	workspace.Name
description	Automatic	Power BI	Power BI Workspace	Yes	workspace.Description
state	Automatic	Power BI	Power BI Workspace	No	workspace.State
type	Automatic	Power BI	Power BI Workspace	No	ResourceType.Workspace
IsReadOnly	Automatic	Power BI	Power BI Workspace	No	workspace.IsReadOnly
IsOnDedicatedCapacity	Automatic	Power BI	Power BI Workspace	No	workspace.IsOnDedicatedCapacity
users	Automatic	Power BI	Power BI Workspace	No	workspace.Users

Next steps

- Connect to and manage a Power BI tenant
- Connect to and manage Power BI across tenants
- Connect to and manage Power BI troubleshooting

Connect to and manage Qlik Sense in Microsoft Purview

Article • 03/28/2025

This article outlines how to register Qlik Sense, and how to authenticate and interact with Qlik Sense in Microsoft Purview. For more information about Microsoft Purview, read the introductory article.

Supported capabilities

Scanning capabilities

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan
Yes	Yes	No	No

The supported Qlik Sense version is 11.11 to 14.x.

When scanning Qlik Sense, Microsoft Purview supports extracting technical metadata including:

- Server
- Folders
- Streams
- Applications
- Stories
- Dimensions
- Measures
- Expressions
- QVD (QlikView Data)
- Tables and the columns
- Sheets and the lists, charts, report tables, pivot tables, axes, text boxes, text fields

Other capabilities

View a table listing the supported capabilities for classifications, sensitivity labels, policies, data lineage, and live view.

Known limitations

- When object is deleted from the data source, currently the subsequent scan won't automatically remove the corresponding asset in Microsoft Purview.

Prerequisites

- You must have an Azure account with an active subscription. Create an account for free.
- You must have an active Microsoft Purview account.
- You need an Azure Key Vault, and to grant Microsoft Purview permissions to access secrets.
- You need Data Source Administrator and Data Reader permissions to register a source and manage it in the Microsoft Purview governance portal. For more information about permissions, see Access control in Microsoft Purview.
- Set up the right integration runtime for your scenario. If your data source isn't publicly accessible, set up the latest kubernetes supported self-hosted integration runtime.

Required permissions for scan

Microsoft Purview extracts metadata via Qlik Engine API (JSON over WebSocket). It supports **Qlik Sense certificate authentication** for scanning Qlik Sense. Make sure:

- User account must have read access granted.
- This may be accomplished by creating security rules for a specific application and user.
- Another alternative is if the application is published to a specific stream and the user has rights to read from this stream.

The user must have appropriate permissions:

- Read all application objects (especially script).
- Read all data connections (or at least data connections specified in application).

Create Qlik Sense certificate

You need to export the certificate from Qlik Sense. Learn more from Exporting certificates through the QMC.

1. Launch Qlik Management Console (QMC): https://your_QPS_server_name/qmc.
2. Under **CONFIGURE SYSTEM**, click **Certificates**.
3. Enter the machine name, and optionally input the certificate password or leave it empty.
4. Keep the export file format as "Windows format".
5. Note the location where the certificate will be exported to, and then click Export certificates.
6. In File Explorer, navigate to where the certificate was exported to, e.g. `C:\ProgramData\Qlik\Sense\Repository\Exported Certificates`.

Next, store the **client.pfx** certificate you exported in your Azure Key Vault. You can choose to store it as a certificate or as a secret.

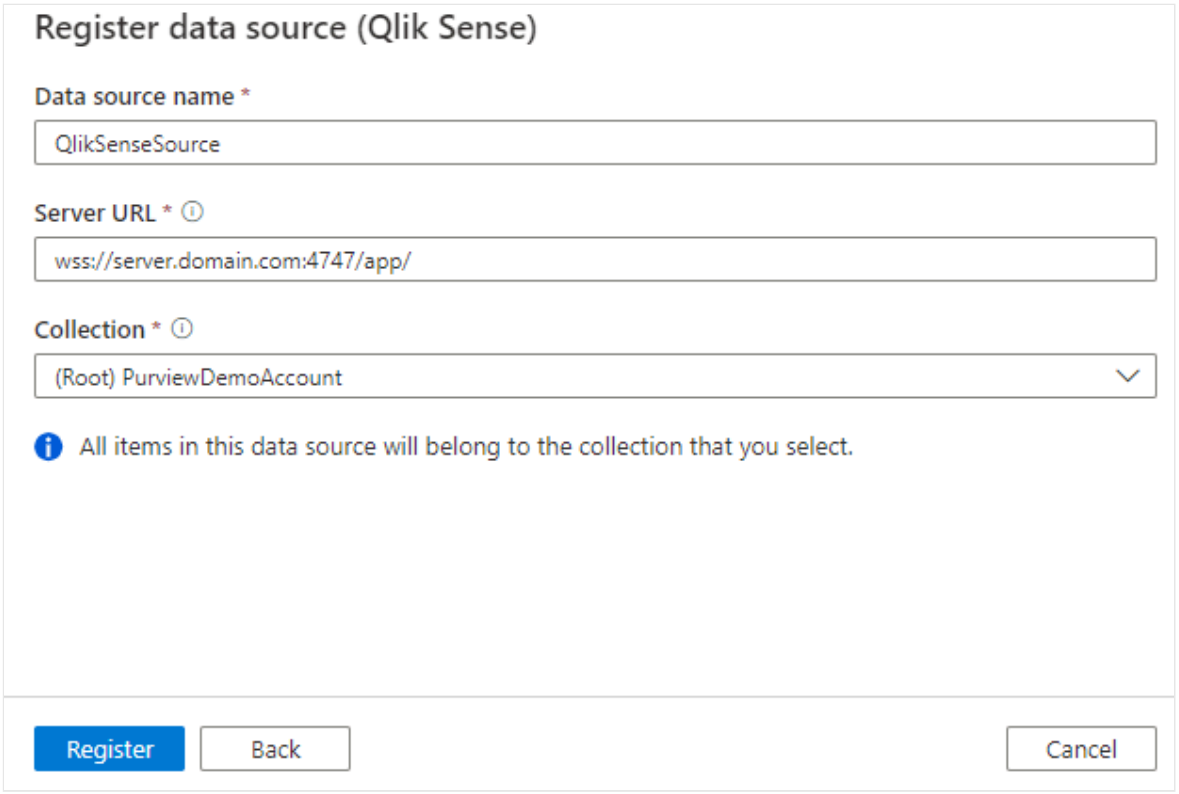
- To store it as a certificate, follow the steps in Tutorial: Import a certificate in Azure Key Vault. If your certificate is protected by a password, note after imported in Key Vault, its associated password isn't saved. The password is required only once during the import operation.
- To store it as a secret, you can convert the content of client.pfx to Base64 format using https://www.base64encode.org/, and put the resulted string as secret value. If your certificate is protected by a password, also save the password as another secret.

Register

This section describes how to register a Qlik Sense server in Microsoft Purview by using the Microsoft Purview governance portal.

1. Go to your Microsoft Purview account.
2. Select **Data Map** on the left pane.
3. Select **Register**.
4. In **Register sources**, select **Qlik Sense** > **Continue**.
5. On the **Register sources (Qlik Sense)** screen, do the following:
   a. For **Name**, enter a name that Microsoft Purview will list as the data source.
   b. For **Server URL**, enter the URL of your Qlik Sense server. In order to use the Qlik Engine API, you need to open a WebSocket to the engine using one of the following URIs: `wss://server.domain.com:4747/app/` or `wss://server.domain.com[/virtual proxy]/app/`.

c. Select a collection from the list.



6. Select **Register**.

Scan

> **Tip**
>
> To troubleshoot any issues with scanning:
>
> 1. Confirm you have followed all **prerequisites**.
> 2. Review our **scan troubleshooting documentation**.

Use the following steps to scan Qlik Sense server to automatically identify assets. For more information about scanning in general, see Scans and ingestion in Microsoft Purview.

1. Go to **Sources**.

2. Select the registered Qlik Sense source.

3. Select + **New scan**.

4. Provide the following details:

   a. **Name**: Enter a name for the scan.

b. **Connect via integration runtime**: Choose the proper type of integration runtime. Learn more from Choose the right integration runtime configuration for your scenario.

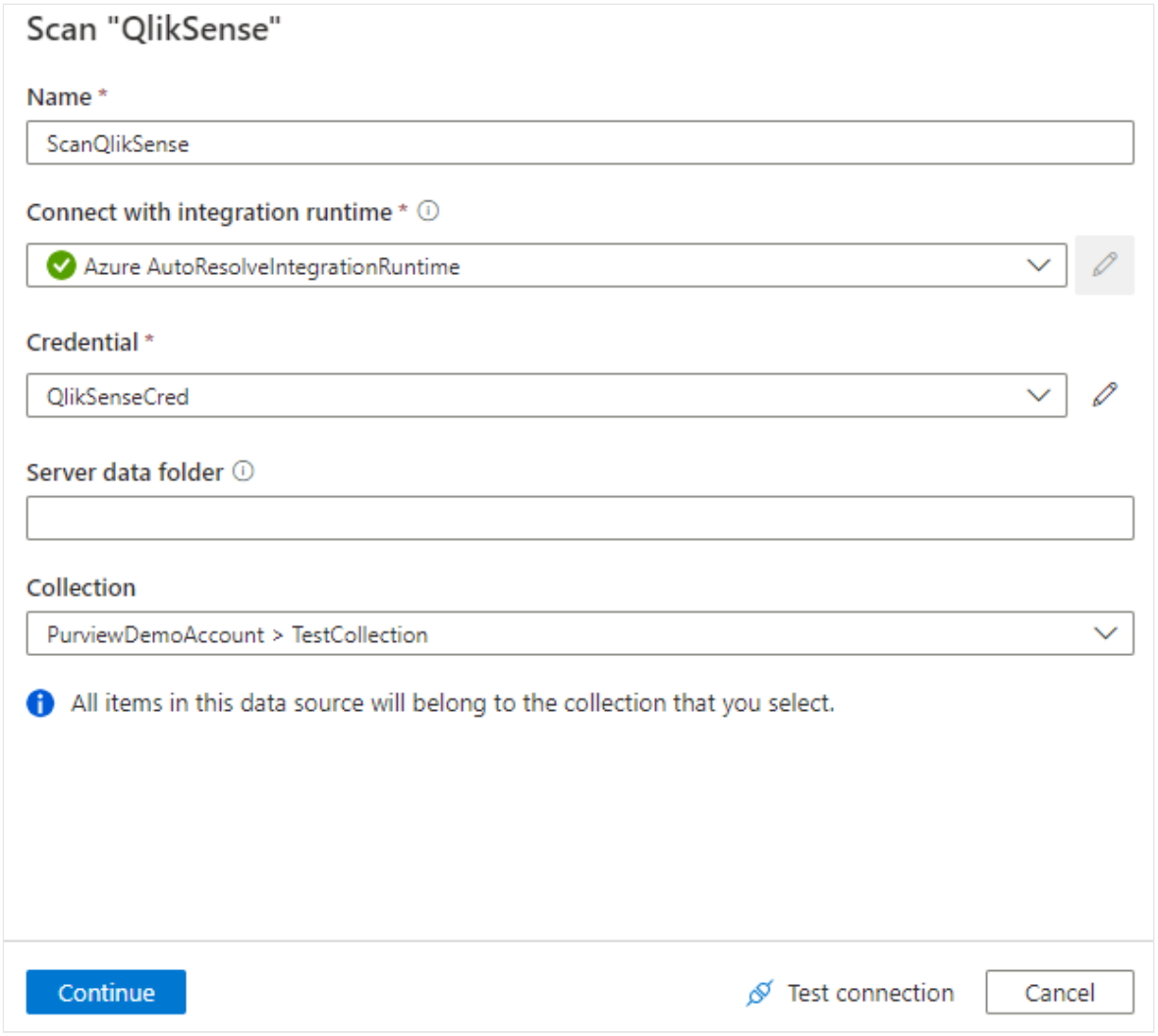
c. **Credential**: Select the credential to connect to your data source. Follow the steps to Create Qlik Sense certificate and store it in Azure Key Vault. When creating the credential, make sure to:

   - Select **Qlik Sense certificate Authentication** while creating a credential.
   - Specify the user name.
   - Specify the user directory. See the "Users" page in the Qlik Management Console (https://your_QPS_server_name/qmc).
   - For the certificate, specify the Key Vault certificate or secret name.
   - If your certificate is protected by a password and you store the certificate as a Key Vault secret, also specify the Key Vault secret name for the password.

   For more information, see Credentials for source authentication in Microsoft Purview.

d. **Server data folder**: The network shared location of Qlik Sense Server's Qlik View Data (QVD) folder, e.g. `\\QlikSense\qliksenseshare`. If specified, Microsoft Purview tries to access the network shared location and scan the QVD files.

5. Click **Test connection** to validate the settings.



6. Select **Continue**.

7. For **Scan trigger**, choose whether to set up a schedule or run the scan once.

8. Review your scan and select **Save and Run**.

View your scans and scan runs

To view existing scans:

1. Go to the Microsoft Purview portal. On the left pane, select **Data map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that has results you want to view. The pane shows you all the previous scan runs, along with the status and metrics for each scan run.
4. Select the run ID to check the scan run details.

Manage your scans

To edit, cancel, or delete a scan:

1. Go to the Microsoft Purview portal. On the left pane, select **Data Map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that you want to manage. You can then:
    - Edit the scan by selecting **Edit scan**.
    - Cancel an in-progress scan by selecting **Cancel scan run**.
    - Delete your scan by selecting **Delete scan**.

> **Note**
>
> - Deleting your scan does not delete catalog assets created from previous scans.

Browse, search, and view assets

After successfully scanning your Qlik Sense server, you can browse Unified Catalog or search Unified Catalog to view the asset details.

Next steps

Now that you've registered your source, use the following guides to learn more about Microsoft Purview and your data:

- Search the Unified Catalog

Feedback

Was this page helpful? Yes No

Connect to and manage Salesforce in Microsoft Purview

Article • 03/28/2025

This article outlines how to register Salesforce, and how to authenticate and interact with Salesforce in Microsoft Purview. For more information about Microsoft Purview, read the introductory article.

Supported capabilities

Scanning capabilities

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan
Yes	Yes	No	Yes

When scanning Salesforce source, Microsoft Purview supports extracting technical metadata including:

- Organization
- Objects including the fields, foreign keys, and unique_constraints

When setting up scan, you can choose to scan an entire Salesforce organization, or scope the scan to a subset of objects matching the given names or name patterns.

Other capabilities

View a table listing the supported capabilities for classifications, sensitivity labels, policies, data lineage, and live view.

Known limitations

When object is deleted from the data source, currently the subsequent scan won't automatically remove the corresponding asset in Microsoft Purview.

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- An active Microsoft Purview account.
- You need Data Source Administrator and Data Reader permissions to register a source and manage it in the Microsoft Purview governance portal. For more information about permissions, see Access control in Microsoft Purview.
- A Salesforce connected app, which will be used to access your Salesforce information.
  - If you need to create a connected app, you can follow the Salesforce documentation.
  - You'll need to enable OAuth for your Salesforce application.

> **Note**
>
> **If your data store is not publicly accessible** (if your data store limits access from on-premises network, private network or specific IPs, etc.), **you will need to configure a self hosted integration runtime to connect to it**.

- If your data source isn't publicly accessible, set up the latest self-hosted integration runtime.
  - Choose the right integration runtime for your scenario:
    - **To use a self-hosted integration runtime:**
      - Follow the article to create and configure a self-hosted integration runtime.
      - Ensure JDK 11 is installed on the machine where the self-hosted integration runtime is installed. Restart the machine after you newly install the JDK for it to take effect.
      - Ensure that Visual C++ Redistributable (version Visual Studio 2012 Update 4 or newer) is installed on the machine where the self-hosted integration runtime is running. If you don't have this update installed, download it now.
      - Ensure the self-hosted integration runtime machine's IP is within the trusted IP ranges for your organization set on Salesforce.
    - **To use a kubernetes supported self-hosted integration runtime:**
      - Follow the article to create and configure a kubernetes supported integration runtime.
      - Ensure the self-hosted integration runtime machine's IP is within the trusted IP ranges for your organization set on Salesforce.

Required permissions for scan

If users will be submitting Salesforce Documents, certain security settings must be configured to allow this access on Standard Objects and Custom Objects. To configure permissions:

- Within Salesforce, select Setup and then select Manage Users.
- Under the Manage Users tree select Profiles.
- Once the Profiles appear on the right, select which Profile you want to edit and select the Edit link next to the corresponding profile.

For Standard Objects, ensure that the "Documents" section has the Read permissions selected. For Custom Objects, ensure that the Read permissions selected for each custom objects.

Register

This section describes how to register Salesforce in Microsoft Purview using the Microsoft Purview governance portal.

Steps to register

To register a new Salesforce source in Microsoft Purview Unified Catalog, follow these steps:

1. Navigate to your Microsoft Purview account in the Microsoft Purview governance portal.
2. Select **Data Map** on the left navigation.
3. Select **Register**
4. On Register sources, select **Salesforce**. Select **Continue**.

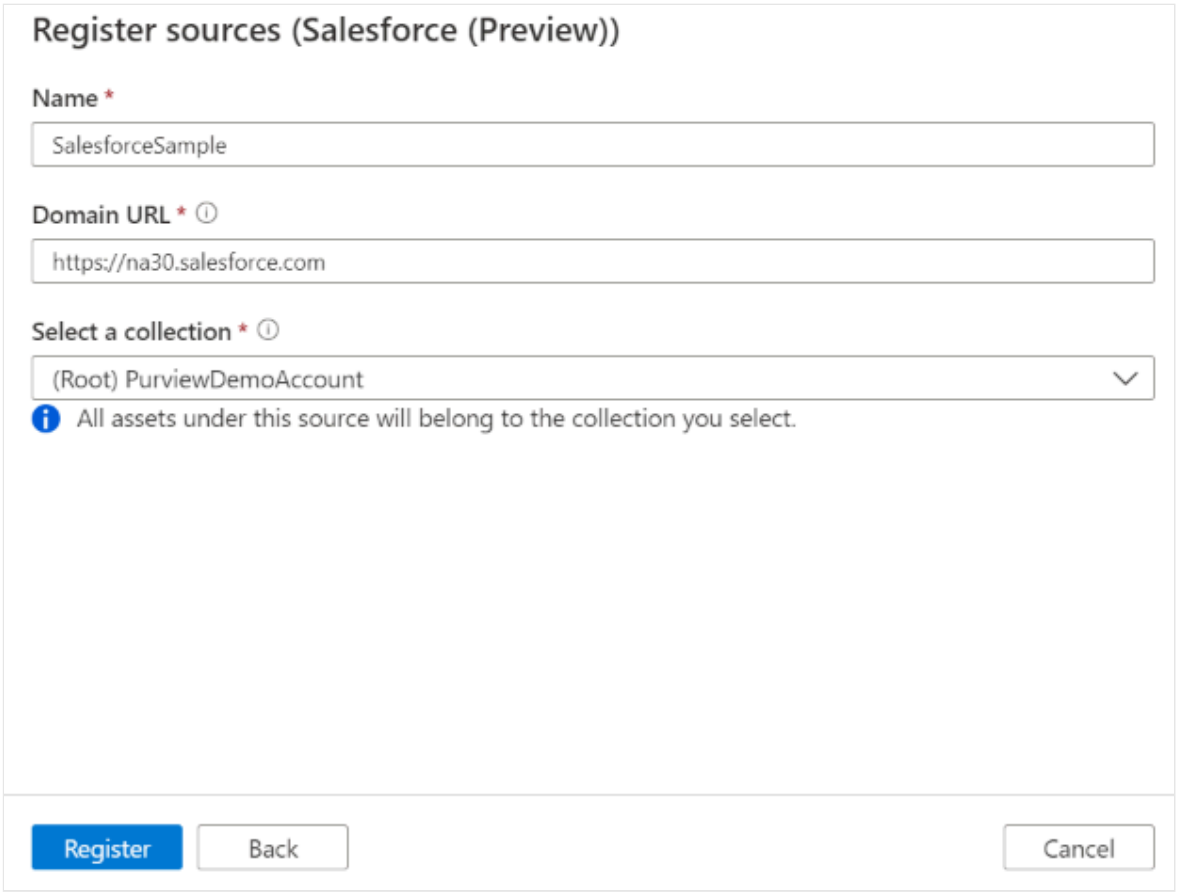
On the **Register sources (Salesforce)** screen, follow these steps:

1. Enter a **Name** that the data source will be listed within the Catalog.

2. Enter the Salesforce sign in endpoint URL as **Domain URL**. For example, `https://login.salesforce.com`. You can use your company's instance URL (such as `https://na30.salesforce.com`) or My Domain URL (such as `https://myCompanyName.my.salesforce.com/`) or Sandbox URL (such as `https://myCompanyName.sandbox.my.salesforce.com/`).

3. Select a collection from the list.

4. Finish to register the data source.



Scan

Follow these steps to scan Salesforce to automatically identify assets. For more information about scanning in general, see our introduction to scans and ingestion.

Microsoft Purview uses Salesforce REST API version 41.0 to extract metadata, including REST requests like 'Describe Global' URI (/v41.0/sobjects/),'sObject Basic Information' URI (/v41.0/sobjects/sObject/), and 'SOQL Query' URI (/v41.0/query?).

Authentication for a scan

The supported authentication type for a Salesforce source is **Consumer key authentication**.

Create and run scan

To create and run a new scan, follow these steps:

1. If your server is publicly accessible, skip to step two. Otherwise, you'll need to make sure your self-hosted integration runtime is configured:
   a. In the Microsoft Purview governance portal, got to the Management Center, and select **Integration runtimes**.

b. Make sure a self-hosted integration runtime is available. If one isn't set up, use the steps mentioned in prerequisites to set up a self-hosted integration runtime.

2. In the Microsoft Purview governance portal, navigate to **Sources**.

3. Select the registered Salesforce source.

4. Select + **New scan**.

5. Provide the below details:

   a. **Name**: The name of the scan

   b. **Connect via integration runtime**: Select the Azure autoresolved integration runtime if your server is publicly accessible, or your configured self-hosted integration runtime if it isn't publicly available.

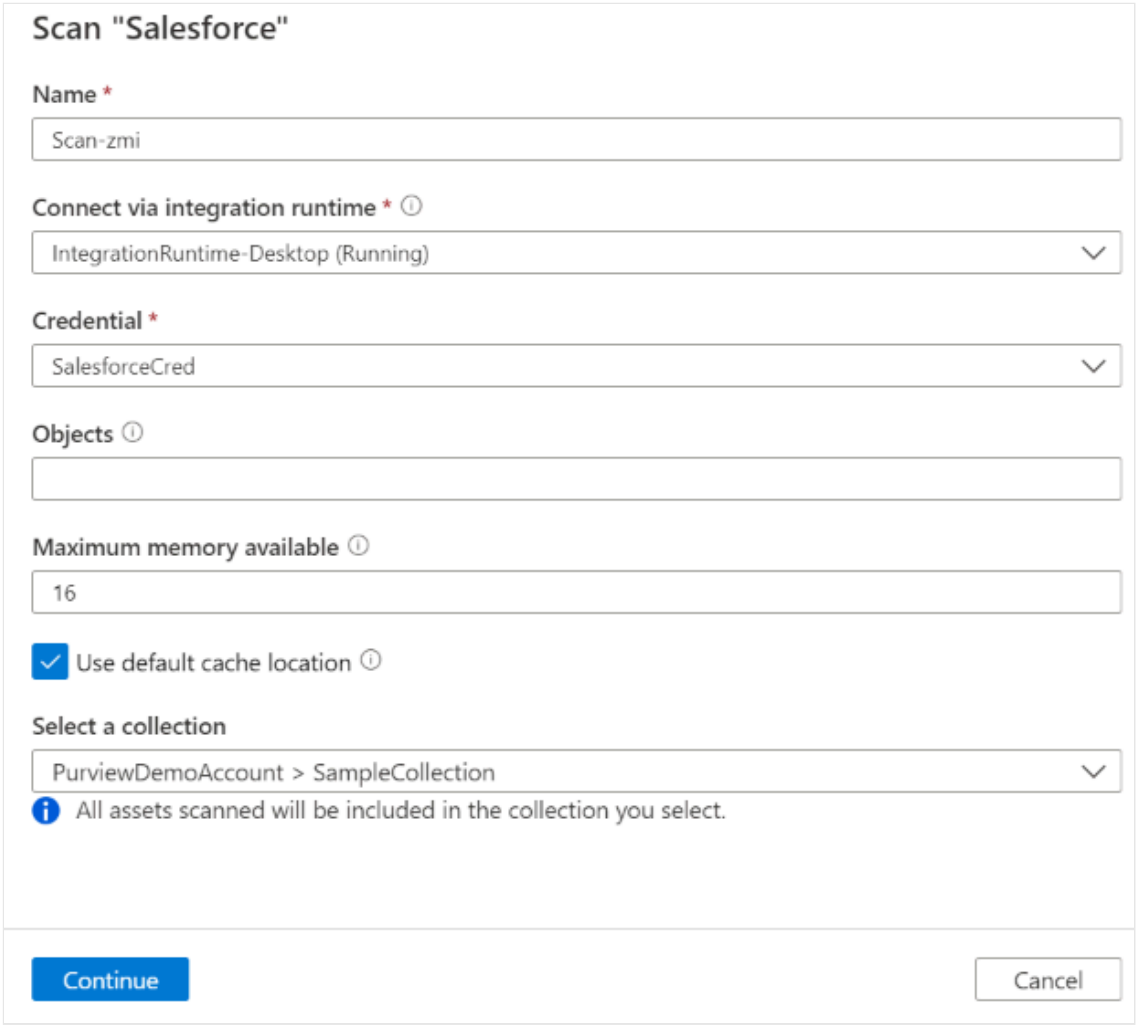
   c. **Credential**: Select the credential to connect to your data source. Make sure to:

      - Select **Consumer key** while creating a credential.
      - Provide the username of the user that the connected app is imitating in the User name input field.
      - Store the password of the user that the connected app is imitating in an Azure Key Vault secret.
        - If your self-hosted integration runtime machine's IP is within the trusted IP ranges for your organization set on Salesforce, provide just the password of the user.
        - Otherwise, **concatenate the password and security token as the value of the secret**. The security token is an automatically generated key that must be added to the end of the password when logging in to Salesforce from an untrusted network. Learn more about how to get or reset a security token.
      - Provide the consumer key from the connected app definition. You can find it on the connected app's Manage Connected Apps page or from the connected app's definition.
      - Stored the consumer secret from the connected app definition in an Azure Key Vault secret. You can find it along with consumer key.

   d. **Objects**: Provide a list of object names to scope your scan. For example, `object1; object2`. An empty list means it's retrieving all available objects. You can specify object names as a wildcard pattern. For example, `topic?`, `*topic*`, or `topic_?,*topic*`.

e. **Maximum memory available** (applicable when using self-hosted integration runtime): Maximum memory (in GB) available on customer's VM to be used by scanning processes. This is dependent on the size of Salesforce source to be scanned.

> **Note**
>
> As a rule of thumb, please provide 1GB memory for every 1000 tables



6. Select **Test connection** to validate the settings (available when using Azure Integration Runtime).

7. Select **Continue**.

8. Choose your **scan trigger**. You can set up a schedule or ran the scan once.

9. Review your scan and select **Save and Run**.

View your scans and scan runs

To view existing scans:

1. Go to the Microsoft Purview portal. On the left pane, select **Data map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that has results you want to view. The pane shows you all the previous scan runs, along with the status and metrics for each scan run.
4. Select the run ID to check the scan run details.

Manage your scans

To edit, cancel, or delete a scan:

1. Go to the Microsoft Purview portal. On the left pane, select **Data Map**.

2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.

3. Select the scan that you want to manage. You can then:

    - Edit the scan by selecting **Edit scan**.
    - Cancel an in-progress scan by selecting **Cancel scan run**.
    - Delete your scan by selecting **Delete scan**.

> **Note**
>
> - Deleting your scan does not delete catalog assets created from previous scans.

Next steps

Now that you've registered your source, follow the below guides to learn more about Microsoft Purview and your data.

- Data Estate Insights in Microsoft Purview
- Lineage in Microsoft Purview
- Search Unified Catalog

Feedback

Was this page helpful? Yes No

Connect to and manage SAP Business Warehouse in Microsoft Purview

This article outlines how to register SAP Business Warehouse (BW), and how to authenticate and interact with SAP BW in Microsoft Purview. For more information about Microsoft Purview, read the introductory article.

Supported capabilities

Scanning capabilities

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan
Yes	Yes	No	No

Purview supports SAP BW versions 7.3 to 7.5, and SAP BW/4HANA.

When you scan an SAP BW source, Microsoft Purview extracts technical metadata, including:

- Instance
- InfoArea
- InfoSet
- InfoSet query
- Classic InfoSet
- InfoObject, including unit of measurement, time characteristic, navigation attribute, data packet characteristic, currency, characteristic, field, and key figure
- Data store object (DSO), including its schema of InfoObject attributes and primary keys
- Advanced data store object (ADSO), including its schema of InfoObject attributes and primary keys
- Aggregation level
- Open hub destination
- Query, including the query condition
- Query view
- HybridProvider
- MultiProvider

- InfoCube
- Aggregate
- Dimension
- Time dimension

Other capabilities

For **classifications**, **sensitivity labels**, **policies**, **data lineage**, and **live view**, see the list of supported capabilities.

Known limitations

- When you delete an object from the data source, the subsequent scan doesn't automatically remove the corresponding asset in Purview.
- You can extract only the `sap_bw_query_condition` under the `sap_bw_query`.

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- An active Microsoft Purview resource.
- You need Data Source Administrator and Data Reader permissions to register a source and manage it in the classic Microsoft Purview governance portal. For more information about permissions, see Access control in Microsoft Purview.
- When you create a scan for the first time, you need to create a function module in your SAP server. Microsoft Purview provides some code in the asset page in the Microsoft Purview Data Map to create the module. To download this code, you need to have Data Source Administrator permissions on the root collection in your data map.
- Set up the right integration runtime for your scenario:
  - **To use a self-hosted integration runtime:**
    - Follow the article to create and configure a self-hosted integration runtime.
    - Ensure JDK 11 is installed on the machine where the self-hosted integration runtime is installed. Restart the machine after you newly install the JDK for it to take effect.
    - Ensure that Visual C++ Redistributable (version Visual Studio 2012 Update 4 or newer) is installed on the machine where the self-hosted integration runtime is running. If you don't have this update installed, download it now.

    - The connector reads metadata from SAP using the SAP Java Connector (JCo) 3.0 API. Make sure the Java Connector is available on your machine where self-hosted integration runtime is installed. Make sure that you use the correct JCo distribution for your environment, and the **sapjco3.jar** and **sapjco3.dll** files are available.
  - **To use a kubernetes supported self-hosted integration runtime:**
    - Follow the article to create and configure a kubernetes supported integration runtime.
    - The connector reads metadata from SAP using the SAP Java Connector (JCo) 3.0 API. Make sure the Java Connector is available on your machine where self-hosted integration runtime is installed. Make sure that you use the correct JCo distribution for your environment, and the **sapjco3.jar** and **sapjco3.dll** files are available.

    > **Note**
    >
    > The driver should be accessible to all accounts in the machine. Don't put it in a path under user account.

  - Self-hosted integration runtimes communicate with the SAP server over dispatcher port 32NN and gateway port 33NN, where NN is your SAP instance number from 00 to 99. Make sure the outbound traffic is allowed on your firewall.

- Deploy the metadata extraction ABAP function module on the SAP server by following the steps mentioned in ABAP functions deployment guide. You need to download a separate ABAP module for SAP BW/4 HANA version from the scan page. You need an ABAP developer account to create the RFC function module on the SAP server. For scan execution, the user account requires sufficient permissions to connect to the SAP server and execute the following RFC function modules:

- STFC_CONNECTION (check connectivity)
  - RFC_SYSTEM_INFO (check system information)
  - OCS_GET_INSTALLED_COMPS (check software versions)
  - Z_MITI_BW_DOWNLOAD (main metadata import, the function module you create following the Purview guide)

  The underlying SAP Java Connector (JCo) libraries might call more RFC function modules, such as. RFC_PING, RFC_METADATA_GET. Refer to SAP support note 460089 for details.

Register

This section describes how to register SAP BW in Microsoft Purview using the classic Microsoft Purview governance portal.

Authentication for registration

The only supported authentication method for an SAP BW source is **Basic authentication**.

Steps to register

1. Open the classic Microsoft Purview governance portal by:

    - Browsing directly to https://web.purview.azure.com and selecting your Microsoft Purview account.
    - Opening the Azure portal, searching for and selecting the Microsoft Purview account. Selecting the **the Microsoft Purview governance portal** button.

2. Select **Data Map** on the left navigation.

3. Select **Register**.

4. In **Register sources**, select **SAP BW** > **Continue**.

On the **Register sources (SAP BW)** screen, complete the following steps:

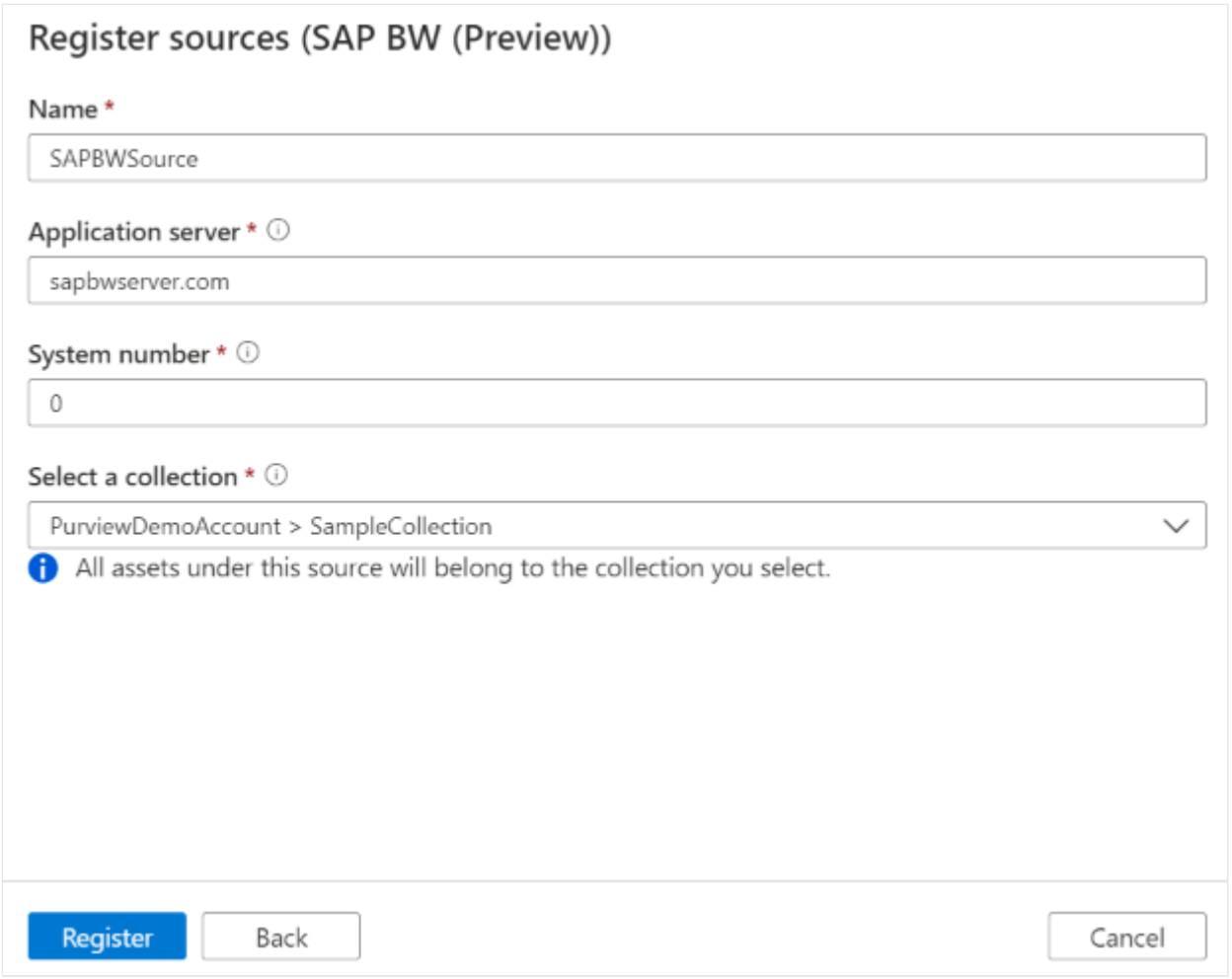
1. Enter a **Name** for the data source to list it within the catalog.

2. Enter the **Application server** name to connect to the SAP BW source. You can also enter an IP address for the SAP application server host.

3. Enter the SAP **System number**. It's an integer between 0 and 99.

4. Select a collection from the list.

5. Finish the registration of the data source.



Scan

Follow the steps in this section to scan SAP BW and automatically identify assets. For more information about scanning, see introduction to scans and ingestion.

> **Important**
>
> When you create a scan for the first time, you need to create a function module in your SAP server. Microsoft Purview provides some code in the asset page in the Microsoft Purview Data Map to create the module. To download this code, you need to have Data Source Administrator permissions on the root collection in your data map.

Create and run scan

1. In the Management Center, select **Integration runtimes**. Make sure you set up a self-hosted integration runtime. If you don't have a self-hosted integration runtime, use the steps mentioned in prerequisites to create one.

2. Go to **Sources**.

3. Select the registered SAP BW source.

4. Select **New scan**.

5. Enter the following details:

   a. **Name**: The name of the scan.

   b. **Connect via integration runtime**: Select the configured self-hosted integration runtime.

   c. **Credential**: Select the credential to connect to your data source. Make sure to:

      - Select Basic Authentication while creating a credential.
      - Provide a user ID to connect to SAP server in the User name input field.
      - Store the user password used to connect to SAP server in the secret key.

   d. **Client ID**: Enter the SAP Client ID. It's a three-digit numeric number from 000 to 999.

   e. **SNC Mode (Optional)**: Turn the toggle **on** if you prefer to use Secure Network Communications (SNC) mode for securing your technical connections to SAP RFCs through a certificate-based authentication mechanism. SNC Mode is **off** by default. Follow the steps to set up the SAP personal secure environment with SNC certificate.

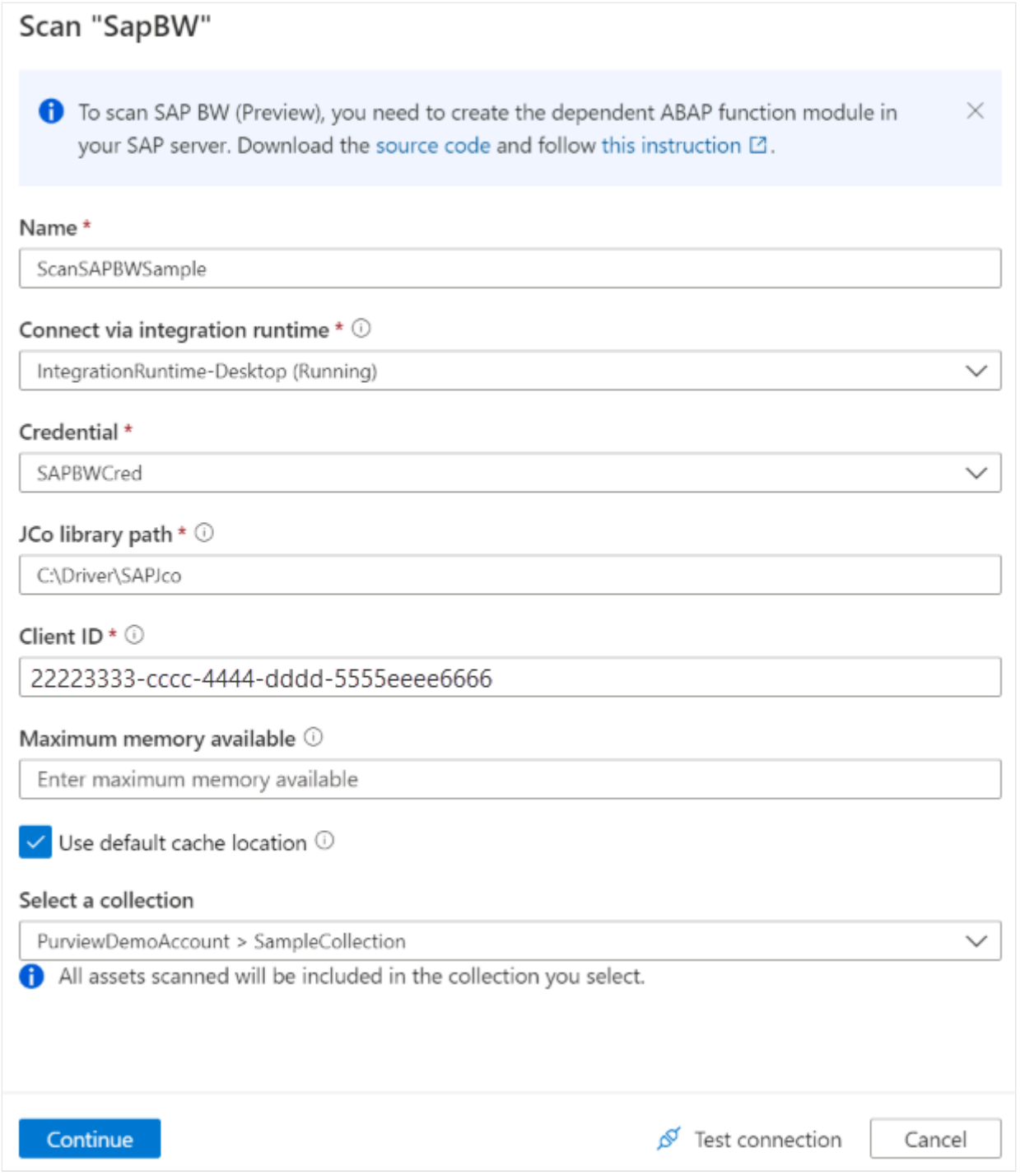
   f. **JCo library path**: Specify the directory path where the JCo libraries are located, for example: `D:\Drivers\SAPJCo`. Make sure the path is accessible by the self-hosted integration runtime, learn more from prerequisites section.

      - For self-hosted integration runtime on a local machine: `D:\Drivers\SAPJCo`. It's the path to a valid JAR folder location. The value must be a valid absolute file path and doesn't contain space. Make sure the driver is accessible by the self-hosted integration runtime. To learn more, see the prerequisites section.

      - For Kubernetes-supported self-hosted integration runtime: `./drivers/SAPJCo`. It's the path to a valid JAR folder location. The value must be a valid relative file path. Refer to the documentation to set up a scan with external drivers for uploading drivers in advance.

   g. **Maximum memory available:** Maximum memory (in GB) available on the self-hosted integration runtime machine to be used by scanning processes. This value depends on

the size of SAP BW source to be scanned.



6. Select **Test connection**.

7. Select **Continue**.

8. Choose your **scan trigger**. You can set up a schedule or run the scan once.

9. Review your scan and select **Save and run**.

View your scans and scan runs

To view existing scans:

1. Go to the Microsoft Purview portal. On the left pane, select **Data map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that has results you want to view. The pane shows you all the previous scan runs, along with the status and metrics for each scan run.
4. Select the run ID to check the scan run details.

Manage your scans

To edit, cancel, or delete a scan:

1. Go to the Microsoft Purview portal. On the left pane, select **Data Map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that you want to manage. You can then:
    - Edit the scan by selecting **Edit scan**.
    - Cancel an in-progress scan by selecting **Cancel scan run**.
    - Delete your scan by selecting **Delete scan**.

> ⓘ **Note**
>
> - Deleting your scan does not delete catalog assets created from previous scans.

Set up SAP personal secure environment with SNC certificate

Get SAPCAR

- Go to the SAP Software Download Center and sign in with your SAP credentials.
- Search for SAPCAR and select the latest nonarchived version.
- Select your operating system.
- Download the `.EXE file to C:\sap\SAR`.

Get SAP Common Crypto Library

- In the SAP Software Download Center, search for "COMMONCRYPTOLIB" and select the latest version.
- Select your operating system.

- Download the .SAR file with the most recent release date to `C:\sap\SAR`.

Extract SAP Common Crypto Library

- Open PowerShell and go to `C:\sap\SAR`.
- Enter the following command, replacing xxxx with your values: `.\SAPCAR_xxxx.EXE -xvf .\SAPCRYPTOLIBP_xxxx.SAR -R .\..\libs\sapcryptolib`.
- Confirm that `sapgenpse.exe` is in the `C:\sap\libs\sapcryptolib` directory.

Generate certificates

> **Note**
>
> This method is for demonstration purposes only and isn't recommended for production systems. For production systems, consult your internal PKI guidance or security team.

- Set up the folder structure:
  - `mkdir rootCA`
  - `mkdir sncCert`
- Create the necessary serial and index files if they don't exist:
  - `if (-Not (Test-Path "rootCA\index.txt")) { New-Item -Path "rootCA\index.txt" ItemType File }`
  - `if (-Not (Test-Path "rootCA\serial")) { Set-Content -Path "rootCA\serial" -Value "01" }`
- Generate a Root CA:
  - `openssl genpkey -algorithm RSA -out rootCA/ca.key.pem -pkeyopt rsa_keygen_bits:2048`
  - `openssl req -x509 -new -key rootCA/ca.key.pem -days 7305 -sha256 -extensions v3_ca -out rootCA/ca.cert.pem -subj "/O=Contoso/CN=Root CA"`
- Generate the SNC certificate
  - `openssl genrsa -out sncCert/snc.key.pem 2048`
  - `openssl req -key sncCert/snc.key.pem -new -sha256 -out sncCert/snc.csr.pem subj "/O=Contoso/CN=SNC"`
- Create an OpenSSL configuration file, `sncCert/extensions.cnf`, for signing:
  - `subjectKeyIdentifier = hash`
  - `authorityKeyIdentifier = keyid,issuer`
  - `basicConstraints = critical,CA:false`

- `keyUsage = critical,digitalSignature,keyEncipherment,dataEncipherment`
- `extendedKeyUsage = clientAuth,emailProtection`

- Sign the SNC certificate with the Root CA:
  - `openssl x509 -req in sncCert/snc.csr.pem CA rootCA/ca.cert.pem CAkey rootCA/ca.key.pem CAcreateserial out sncCert/snc.cert.pem days 3650 sha256 extfile sncCert\extensions.cnf extensions v3_leaf`

Create a personal secure environment

Create a personal secure environment (PSE) for the on-premises data gateway. The NCo library looks for the SNC certificate inside the PSE.

- Create a PKCS#12 container:
  - `openssl pkcs12 -export -out snc.p12 -inkey sncCert\snc.key.pem -in sncCert\snc.cert.pem -certfile rootCA\ca.cert.pem`
- Create the SECUDIR environment variable.
- Open **System Properties**: In **File Explorer**, right-click **This PC**, and then select **Properties** > **Advanced system settings**.
- Select **Environment Variables**.
- Under **System variables**, select **New**.
- Set the variable name to **SECUDIR**.
- Set the value to `C:\sapsecudir`.
- Select **OK**.

Import the PKCS#12 container into a PSE: `C:\sap\libs\sapcryptolib\sapgenpse.exe import_p12 -p SAPSNCSKERB.pse C:\pki certs\snc.p12`

Allow your SHIR process to use your SAP PSE

Verify which user or service is being used by your SNC client to obtain the certificate to communicate with SAP. The Microsoft Purview SHIR uses the service user **NT SERVICE\DIAHostService**.

- Add a credential to allow the certificate retrieval request from the PSE: `.\sapgenpse.exe seclogin -p C:\sapsecudir\SAPSNCSKERB.pse -x your-pse-pin -O "NT`

`SERVICE\DIAHostService"`

- Verify credentials like this: `.\sapgenpse.exe seclogin -l -O "NT SERVICE\DIAHostService"`
- You can delete them like this: `.\sapgenpse.exe seclogin -d -O "NT SERVICE\DIAHostService"`
- Use the -h parameter to get help with the sapgenpse command line tool or check the command reference here.

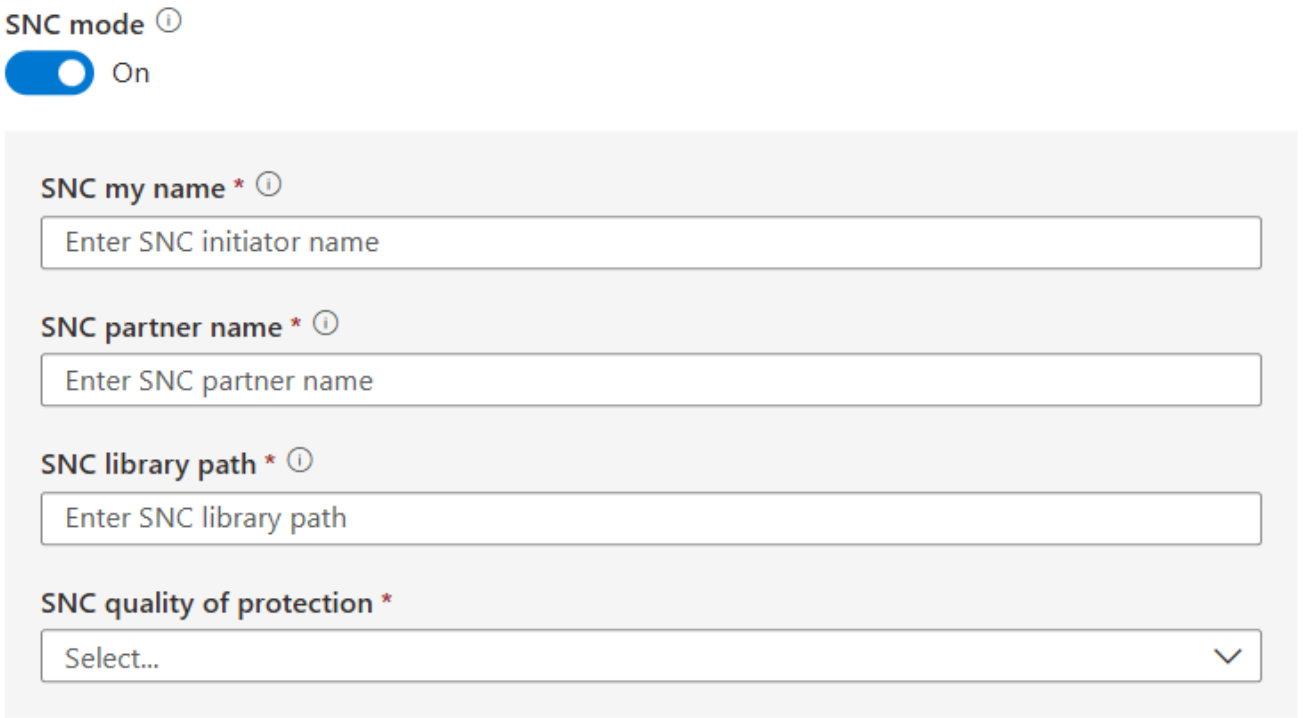
Give permission to access the SAP JCo library to SHIR process

Verify if Microsoft Purview SHIR user **NT SERVICE\DIAHostService** has **Read/Write/Execute/List** permissions over these folders:

- SNC Library Path: The folder that has the SNC Library, `sapcrypto.dll`. Example: `C:\Users\shir-admin\Desktop\snc\SAPCRYPTOLIBP_8557-20011729\sapcrypto.dll`
- JCo Library Path: The folder that has the SAP Java Connector jar file. Example: `C:\Users\shir-admin\Desktop\snc\sapjco3-ntamd64-3.1\sapjco3-ntamd64-3.1.3`

SAP NCo Library should also be installed on the virtual machine for test connection to work properly, as it still uses NCo library. This isn't necessary, and Test Connection can be skipped.

Enter SCN details for **my name**, **partner name**, **library path**, and **quality of protection**, as seen below, and initiate the scan:



Next steps

After registering your source, explore the following guides to learn more about Purview and your data:

- Search Unified Catalog
- Data Estate Insights in Microsoft Purview
- Supported data sources and file types

Last updated on 04/27/2026

Connect to and manage SAP ECC in Microsoft Purview

06/25/2025

This article outlines how to register SAP ECC, and how to authenticate and interact with SAP ECC in Microsoft Purview. For more information about Microsoft Purview, read the introductory article.

Supported capabilities

Scanning capabilities

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan
Yes	Yes	No	No

When scanning SAP ECC source, Microsoft Purview supports:

- Extracting technical metadata including:
  - Instance
  - Application components
  - Packages
  - Tables including the fields, foreign keys, indexes, and index members
  - Views including the fields
  - Transactions
  - Programs
  - Classes
  - Function groups
  - Function modules
  - Domains including the domain values
  - Data elements
- Fetching static lineage on assets relationships among tables and views.

Other capabilities

For **classifications**, **sensitivity labels**, **policies**, **data lineage**, and **live view**, see the list of supported capabilities.

Known limitations

When object is deleted from the data source, currently the subsequent scan won't automatically remove the corresponding asset in Microsoft Purview.

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- An active Microsoft Purview account.
- You need Data Source Administrator and Data Reader permissions to register a source and manage it in the Microsoft Purview governance portal. For more information about permissions, see Access control in Microsoft Purview.
- Set up the right integration runtime for your scenario:
  - **To use a self-hosted integration runtime:**
    - Follow the article to create and configure a self-hosted integration runtime.
    - Ensure JDK 11 is installed on the machine where the self-hosted integration runtime is installed. Restart the machine after you newly install the JDK for it to take effect.
    - Ensure that Visual C++ Redistributable (version Visual Studio 2012 Update 4 or newer) is installed on the machine where the self-hosted integration runtime is running. If you don't have this update installed, download it now.
    - Download the 64-bit SAP Connector for Microsoft .NET 3.0 from SAP's website and install it on the self-hosted integration runtime machine. During installation, make sure you select the **Install Assemblies to GAC** option in the **Optional setup steps** window.
    - The connector reads metadata from SAP using the SAP Java Connector (JCo) 3.0 API. Make sure the Java Connector is available on your virtual machine where self-hosted integration runtime is installed. Make sure that you're using the correct JCo distribution for your environment. For example: on a Microsoft Windows machine, make sure the sapjco3.jar and sapjco3.dll files are available. Note down the folder path which you will use to set up the scan.
    - Self-hosted integration runtime communicates with the SAP server over dispatcher port 32NN and gateway port 33NN, where NN is your SAP instance number from 00 to 99. Make sure the outbound traffic is allowed on your firewall.
  - **To use a kubernetes supported self-hosted integration runtime:**
    - Follow the article to create and configure a kubernetes supported integration runtime.

- Download the 64-bit SAP Connector for Microsoft .NET 3.0 from SAP's website and install it on the self-hosted integration runtime machine. During installation, make sure you select the **Install Assemblies to GAC** option in the **Optional setup steps** window.
- The connector reads metadata from SAP using the SAP Java Connector (JCo) 3.0 API. Make sure the Java Connector is available on your virtual machine where self-hosted integration runtime is installed. Make sure that you're using the correct JCo distribution for your environment. For example: on a Microsoft Windows machine, make sure the sapjco3.jar and sapjco3.dll files are available. Note down the folder path which you'll use to set up the scan.
- Self-hosted integration runtime communicates with the SAP server over dispatcher port 32NN and gateway port 33NN, where NN is your SAP instance number from 00 to 99. Make sure the outbound traffic is allowed on your firewall.

> **Note**
>
> Scanning SAP ECC is a memory intensive operation, you're recommended to install Self-hosted Integration Runtime on a machine with at least 128 GB RAM.

- Deploy the metadata extraction ABAP function module on the SAP server by following the steps mentioned in ABAP functions deployment guide. You need an ABAP developer account to create the RFC function module on the SAP server. For scan execution, the user account requires sufficient permissions to connect to the SAP server and execute the following RFC function modules:
  - STFC_CONNECTION (check connectivity)
  - RFC_SYSTEM_INFO (check system information)
  - OCS_GET_INSTALLED_COMPS (check software versions)
  - Z_MITI_DOWNLOAD (main metadata import, the function module you create following the Purview guide)

  The underlying SAP Java Connector (JCo) libraries might call more RFC function modules; for example, RFC_PING, RFC_METADATA_GET, etc. See SAP support note 460089 for details.

Register

This section describes how to register SAP ECC in Microsoft Purview using the Microsoft Purview governance portal.

Authentication for registration

The only supported authentication for SAP ECC source is **Basic authentication**.

Steps to register

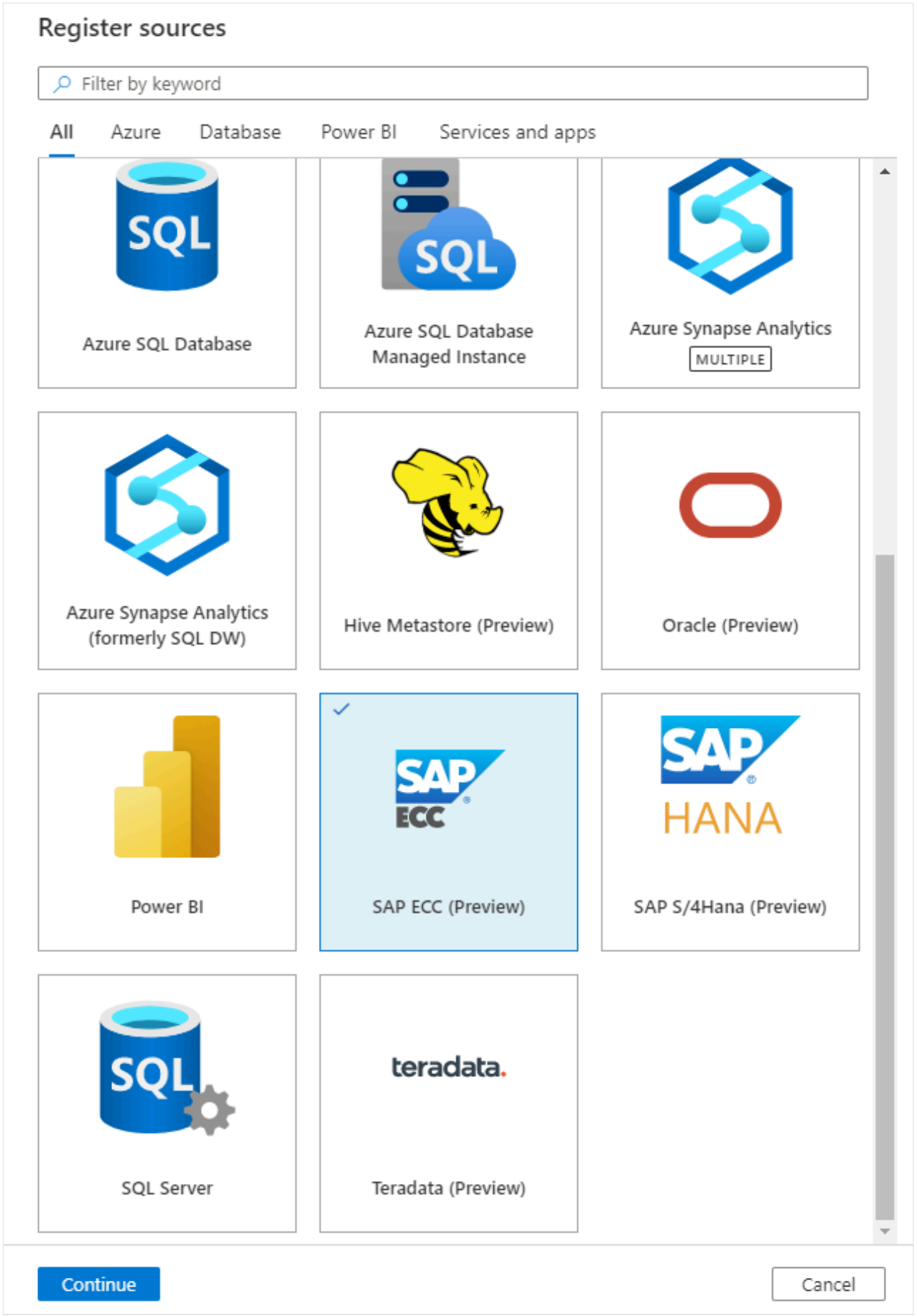
1. Open the Microsoft Purview governance portal by:

    - Browsing directly to https://web.purview.azure.com and selecting your Microsoft Purview account.
    - Opening the Azure portal, searching for and selecting the Microsoft Purview account. Selecting the **the Microsoft Purview governance portal** button.

2. Select **Data Map** on the left navigation.

3. Select **Register**

4. On Register sources, select **SAP ECC**. Select **Continue.**



On the **Register sources (SAP ECC)** screen, do the following:

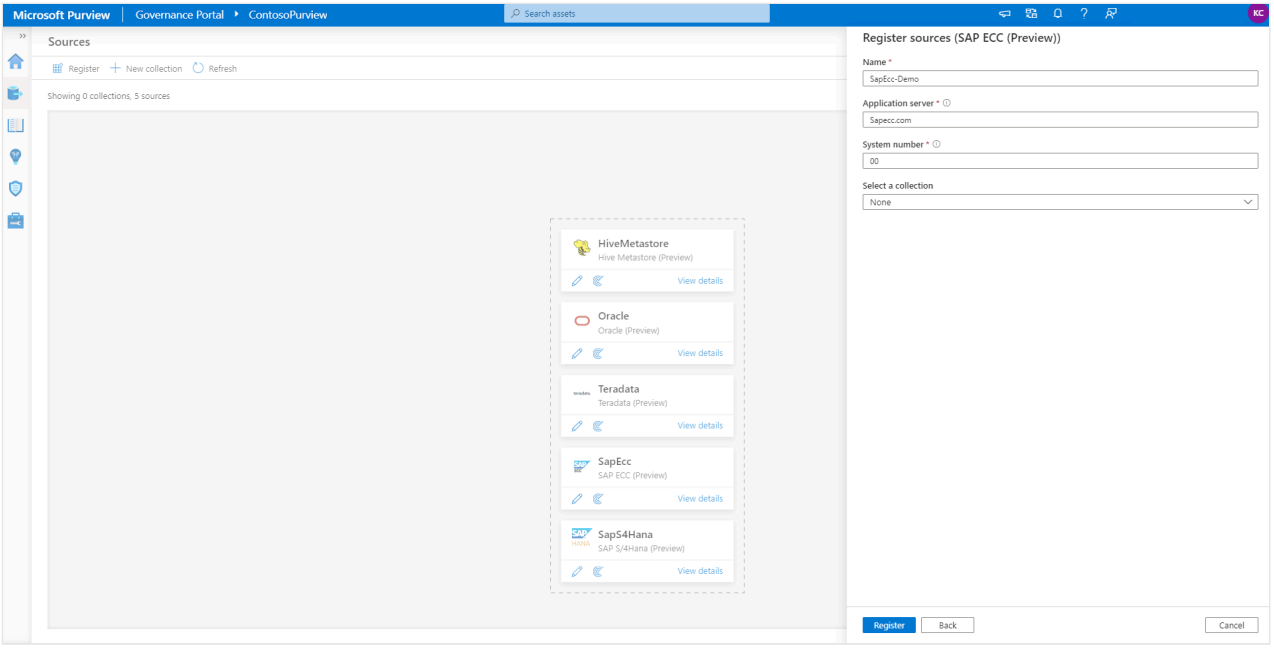
1. Enter a **Name** that the data source will be listed within the Catalog.

2. Enter the **Application server** name to connect to SAP ECC source. It can also be an IP address of the SAP application server host.

3. Enter the SAP **System number**. This is a two-digit integer between 00 and 99.

4. Select a collection from the list.

5. Finish to register the data source.



Scan

Follow the steps below to scan SAP ECC to automatically identify assets. For more information about scanning in general, see our introduction to scans and ingestion.

Create and run scan

1. In the Management Center, select Integration runtimes. Make sure a self-hosted integration runtime is set up. If it isn't set up, use the steps mentioned in prerequisites to create a self-hosted integration runtime.

2. Navigate to **Sources**

3. Select the registered SAP ECC source.

4. Select + **New scan**

5. Provide the below details:

   a. **Name**: The name of the scan

b. **Connect via integration runtime**: Select the configured self-hosted integration runtime.

c. **Credential**: Select the credential to connect to your data source. Make sure to:

- Select Basic Authentication while creating a credential.
- Provide a user ID to connect to SAP server in the User name input field.
- Store the user password used to connect to SAP server in the secret key.

d. **Client ID**: Enter the SAP Client ID. It's a three-digit numeric number from 000 to 999.

e. **SNC Mode (Optional)**: Turn the toggle **on** if you prefer to use Secure Network Communications (SNC) mode for securing your technical connections to SAP RFCs through a certificate-based authentication mechanism. SNC Mode is **off** by default.

Follow these steps to set up the SAP Personal Secure Environment with SNC Certificate:

Get SAPCAR

- Go to the SAP Software Download Center and sign in with your SAP credentials.
- Search for SAPCAR and select the latest nonarchived version.
- Select your operating system.
- Download the `.EXE file to C:\sap\SAR`.

Get SAP Common Crypto Library

- In the SAP Software Download Center, search for "COMMONCRYPTOLIB" and select the latest version.
- Select your operating system.
- Download the .SAR file with the most recent release date to `C:\sap\SAR`.

Extract SAP Common Crypto Library

- Open PowerShell and navigate to `C:\sap\SAR`.
- Enter the following command, replacing xxxx with your values: `.\SAPCAR_xxxx.EXE -xvf .\SAPCRYPTOLIBP_xxxx.SAR -R .\..\libs\sapcryptolib`.
- Confirm that `sapgenpse.exe` is in the `C:\sap\libs\sapcryptolib` directory.

Generate certificates

Note: This method is for demonstration purposes only and isn't recommended for production systems. For production systems, consult your internal PKI guidance or security team.

- Set up the folder structure:

    - `mkdir rootCA`
    - `mkdir sncCert`

- Create the necessary serial and index files if they don't exist:
    - `if (-Not (Test-Path "rootCA\index.txt")) { New-Item -Path "rootCA\index.txt" ItemType File }`
    - `if (-Not (Test-Path "rootCA\serial")) { Set-Content -Path "rootCA\serial" -Value "01" }`

- Generate a Root CA:
    - `openssl genpkey -algorithm RSA -out rootCA/ca.key.pem -pkeyopt rsa_keygen_bits:2048`
    - `openssl req -x509 -new -key rootCA/ca.key.pem -days 7305 -sha256 -extensions v3_ca -out rootCA/ca.cert.pem -subj "/O=Contoso/CN=Root CA"`

- Generate the SNC certificate
    - `openssl genrsa -out sncCert/snc.key.pem 2048`
    - `openssl req -key sncCert/snc.key.pem -new -sha256 -out sncCert/snc.csr.pem subj "/O=Contoso/CN=SNC"`

- Create an OpenSSL configuration file, `sncCert/extensions.cnf`, for signing:
    - `subjectKeyIdentifier = hash`
    - `authorityKeyIdentifier = keyid,issuer`
    - `basicConstraints = critical,CA:false`
    - `keyUsage = critical,digitalSignature,keyEncipherment,dataEncipherment`
    - `extendedKeyUsage = clientAuth,emailProtection`

- Sign the SNC certificate with the Root CA:
    - `openssl x509 -req in sncCert/snc.csr.pem CA rootCA/ca.cert.pem CAkey rootCA/ca.key.pem CAcreateserial out sncCert/snc.cert.pem days 3650 sha256 extfile sncCert\extensions.cnf extensions v3_leaf`

Create a Personal Secure Environment

Create a Personal Secure Environment (PSE) for the on-premises data gateway. The NCo library looks for the SNC certificate inside the PSE.

- Create a PKCS#12 container:
    - `openssl pkcs12 -export -out snc.p12 -inkey sncCert\snc.key.pem -in sncCert\snc.cert.pem -certfile rootCA\ca.cert.pem`

- Create the SECUDIR environment variable.

- Open **System Properties**: In **File Explorer**, right-click **This PC**, and then select **Properties** > **Advanced system settings**.
- Select **Environment Variables**.
- Under **System variables**, select **New**.
- Set the variable name to **SECUDIR**.
- Set the value to `C:\sapsecudir`.
- Select **OK**.

Import the PKCS#12 container into a PSE: `C:\sap\libs\sapcryptolib\sapgenpse.exe import_p12 -p SAPSNCSKERB.pse C:\pki certs\snc.p12`

Allow your SHIR process to use your SAP PSE

Verify which user or service is being used by your SNC client to obtain the certificate to communicate with SAP. The Microsoft Purview SHIR uses the service user **NT SERVICE\DIAHostService**.

- Add a credential to allow the certificate retrieval request from the PSE: `.\sapgenpse.exe seclogin -p C:\sapsecudir\SAPSNCSKERB.pse -x your-pse-pin -O "NT SERVICE\DIAHostService"`
- Verify credentials like this: `.\sapgenpse.exe seclogin -l -O "NT SERVICE\DIAHostService"`
- You can delete them like this: `.\sapgenpse.exe seclogin -d -O "NT SERVICE\DIAHostService"`
- Use the -h parameter to get help with the sapgenpse command line tool or check the command reference here.

Give permission to access the SAP JCo library to SHIR Process

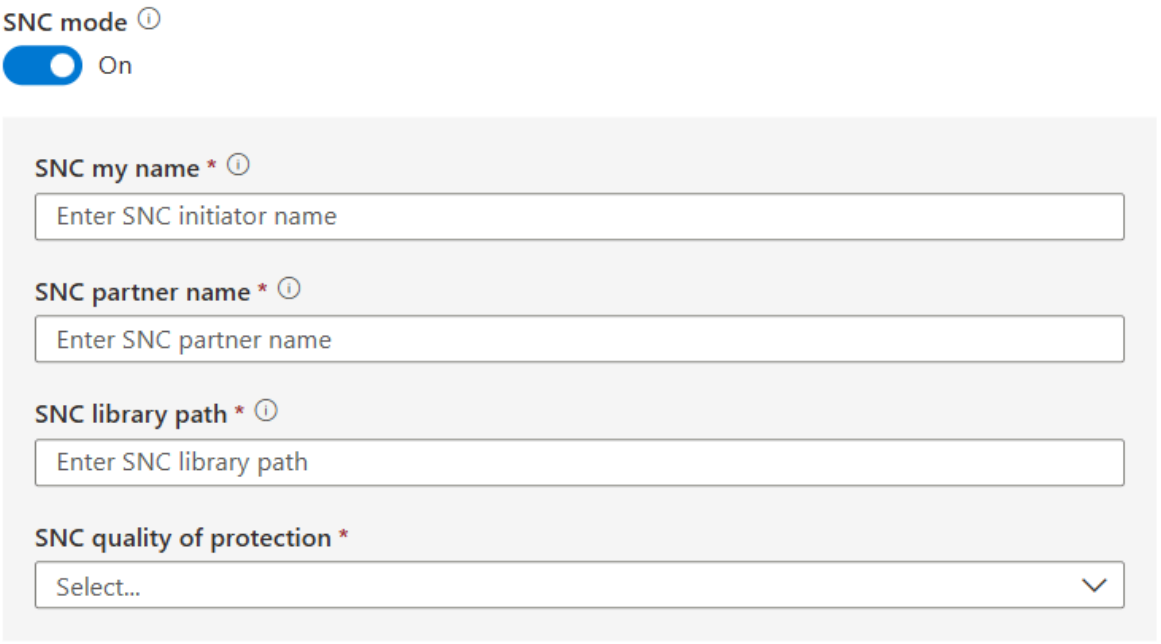
Verify if Microsoft Purview SHIR user **NT SERVICE\DIAHostService** has **Read/Write/Execute/List** permissions over these folders:

- SNC Library Path: The folder that has the SNC Library, `sapcrypto.dll`. Example: `C:\Users\shir-admin\Desktop\snc\SAPCRYPTOLIBP_8557-20011729\sapcrypto.dll`
- JCo Library Path: The folder that has the SAP Java Connector jar file. Example: `C:\Users\shir-admin\Desktop\snc\sapjco3-ntamd64-3.1\sapjco3-ntamd64-3.1.3`

SAP NCo Library should also be installed on the virtual machine for test connection to work properly, as it still uses NCo library. This isn't necessary, and Test Connection can

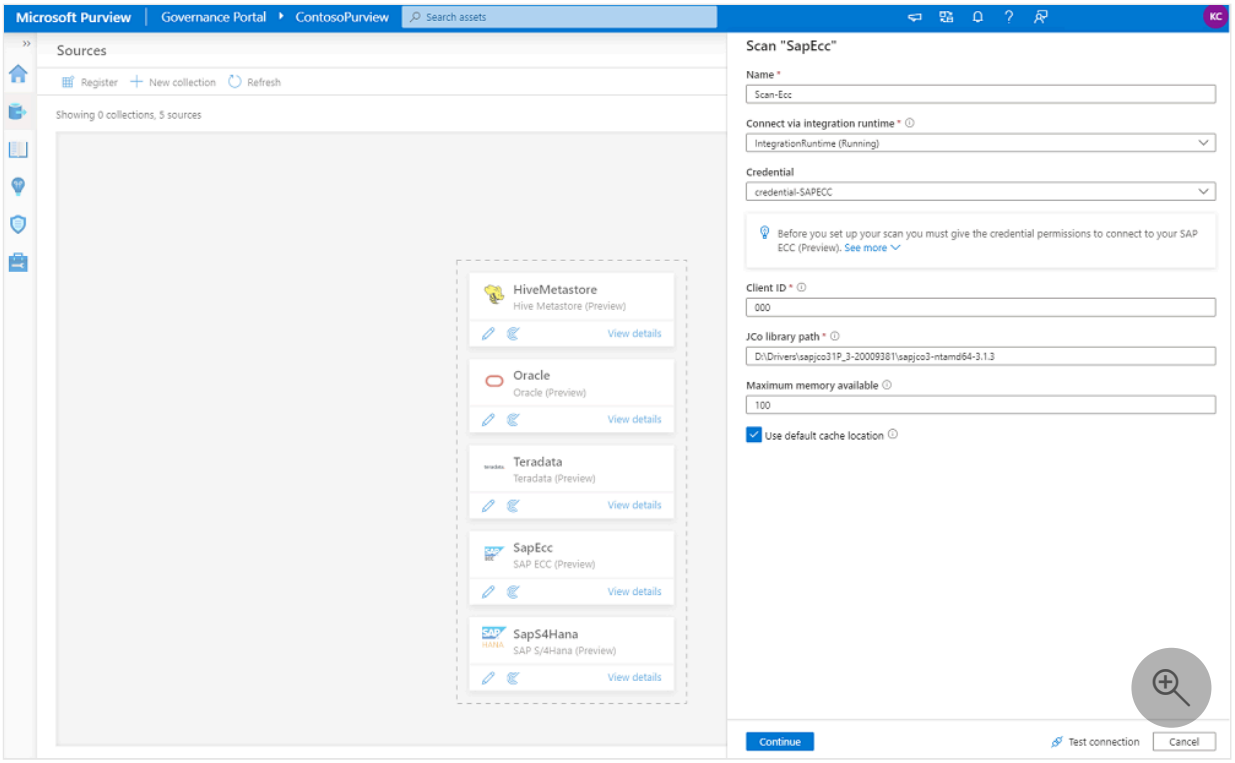
be skipped.

Enter SCN details for **my name**, **partner name**, **library path**, and **quality of protection**, as seen below, and initiate the scan:



f. **JCo library path**: Specify the directory path where the JCo libraries are located, for example: `D:\Drivers\SAPJCo`. Make sure the path is accessible by the self-hosted integration runtime, learn more from prerequisites section.
   i. For self-hosted integration runtime on a local machine: `D:\Drivers\SAPJCo`. It's the path to valid JAR folder location. The value must be a valid absolute file path and doesn't contain space. Make sure the driver is accessible by the self-hosted integration runtime; learn more from prerequisites section.
   ii. For Kubernetes-supported self-hosted integration runtime: `./drivers/SAPJCo`. It's the path to valid JAR folder location. The value must be a valid relative file path. Refer to the documentation to set up a scan with external drivers for uploading drivers in advance.

g. **Maximum memory available:** Maximum memory (in GB) available on the Self-hosted Integration Runtime machine to be used by scanning processes. This is dependent on the size of SAP ECC source to be scanned. It's recommended to provide large available memory, for example, 100.



6. Select **Continue**.

7. Choose your **scan trigger**. You can set up a schedule or ran the scan once.

8. Review your scan and select **Save and Run**.

View your scans and scan runs

To view existing scans:

1. Go to the Microsoft Purview portal. On the left pane, select **Data map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that has results you want to view. The pane shows you all the previous scan runs, along with the status and metrics for each scan run.
4. Select the run ID to check the scan run details.

Manage your scans

To edit, cancel, or delete a scan:

1. Go to the Microsoft Purview portal. On the left pane, select **Data Map**.

2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.

3. Select the scan that you want to manage. You can then:

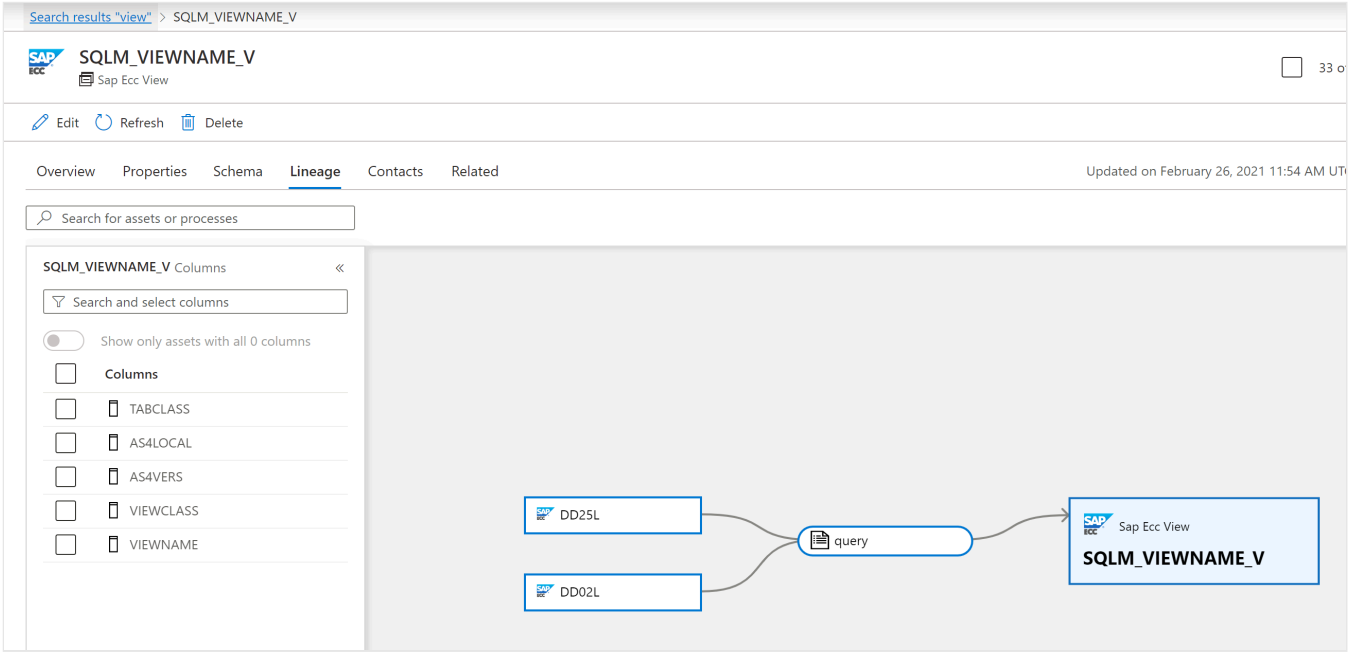
- Edit the scan by selecting **Edit scan**.
- Cancel an in-progress scan by selecting **Cancel scan run**.
- Delete your scan by selecting **Delete scan**.

> **Note**
>
> - Deleting your scan does not delete catalog assets created from previous scans.

Lineage

After scanning your SAP ECC source, you can browse Unified Catalog or search Unified Catalog to view the asset details.

Go to the asset -> lineage tab, you can see the asset relationship when applicable. Refer to the supported capabilities section on the supported SAP ECC lineage scenarios. For more information about lineage in general, see data lineage and lineage user guide.



Next steps

Now that you've registered your source, follow the below guides to learn more about Microsoft Purview and your data.

- Data Estate Insights in Microsoft Purview
- Lineage in Microsoft Purview
- Search Unified Catalog

Connect to and manage SAP HANA in Microsoft Purview

This article outlines how to register SAP HANA, and how to authenticate and interact with SAP HANA in Microsoft Purview. For more information about Microsoft Purview, read the introductory article.

Supported capabilities

Scanning capabilities

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan
Yes	Yes	No	Yes

When scanning SAP HANA source, Microsoft Purview supports extracting technical metadata including:

- Server
- Databases
- Schemas
- Tables, including the columns, foreign keys, indexes, and unique constraints.
- Views, including the columns.
- Calculation views. These are scanned as a part of _SIS_BIC schema. For scoped scan scenarios, this schema needs to be included.
- Stored procedures including the parameter dataset and result set.
- Functions, including the parameter dataset.
- Sequences
- Synonyms

When setting up scan, you can choose to scan an entire SAP HANA database, or scope the scan to a subset of schemas matching the given names or name patterns.

Other capabilities

For **classifications**, **sensitivity labels**, **policies**, **data lineage**, and **live view**, see the list of supported capabilities.

Known limitations

When object is deleted from the data source, currently the subsequent scan won't automatically remove the corresponding asset in Microsoft Purview.

Prerequisites

- You must have an Azure account with an active subscription. Create an account for free.
- You must have an active Microsoft Purview account.
- You need Data Source Administrator and Data Reader permissions to register a source and manage it in the Microsoft Purview portal. For more information about permissions, see Access control in Microsoft Purview.
- Set up the right integration runtime for your scenario:
  - **To use a self-hosted integration runtime:**
    - Follow the article to create and configure a self-hosted integration runtime.
    - Ensure JDK 11 is installed on the machine where the self-hosted integration runtime is installed. Restart the machine after you newly install the JDK for it to take effect.
    - Ensure that Visual C++ Redistributable (version Visual Studio 2012 Update 4 or newer) is installed on the machine where the self-hosted integration runtime is running. If you don't have this update installed, download it now.
    - Download the SAP HANA JDBC driver (JAR ngdbc) on the machine where your self-hosted integration runtime is running. Note down the folder path that you'll use to set up the scan.
  - **To use a kubernetes supported self-hosted integration runtime:**
    - Follow the article to create and configure a kubernetes supported integration runtime.
    - Download the SAP HANA JDBC driver (JAR ngdbc) on the machine where your self-hosted integration runtime is running. Note down the folder path that you'll use to set up the scan.

> **Note**
>
> The driver should be accessible by the self-hosted integration runtime. By default, self-hosted integration runtime uses local service account "NT SERVICE\DIAHostService".

Make sure it has "Read and execute" and "List folder contents" permission to the driver folder.

Required permissions for scan

Microsoft Purview supports basic authentication (username and password) for scanning SAP HANA.

The SAP HANA user you specified must have the permission to select metadata of the schemas you want to import.

SQL

```
CREATE USER <user> PASSWORD <password> NO FORCE_FIRST_PASSWORD_CHANGE;
GRANT SELECT METADATA ON SCHEMA <schema1> TO <user>;
GRANT SELECT METADATA ON SCHEMA <schema2> TO <user>;
```

And the user must have the permission to select on system table _SYS_REPO.ACTIVE_OBJECT and on system schemas _SYS_BI and _SYS_BIC.

SQL

```
GRANT SELECT ON _SYS_REPO.ACTIVE_OBJECT TO <user>;
GRANT SELECT ON SCHEMA _SYS_BI TO <user>;
GRANT SELECT ON SCHEMA _SYS_BIC TO <user>;
```

Register

This section describes how to register an SAP HANA in Microsoft Purview by using the classic Microsoft Purview governance portal.

1. Open the classic Microsoft Purview governance portal by:

   - Browsing directly to https://web.purview.azure.com and selecting your Microsoft Purview account.
   - Opening the Azure portal, searching for and selecting the Microsoft Purview account. Select the classic Microsoft Purview governance portal button.

2. Select **Data Map** on the left pane.

3. Select **Register**.

4. In **Register sources**, select **SAP HANA** > **Continue**.

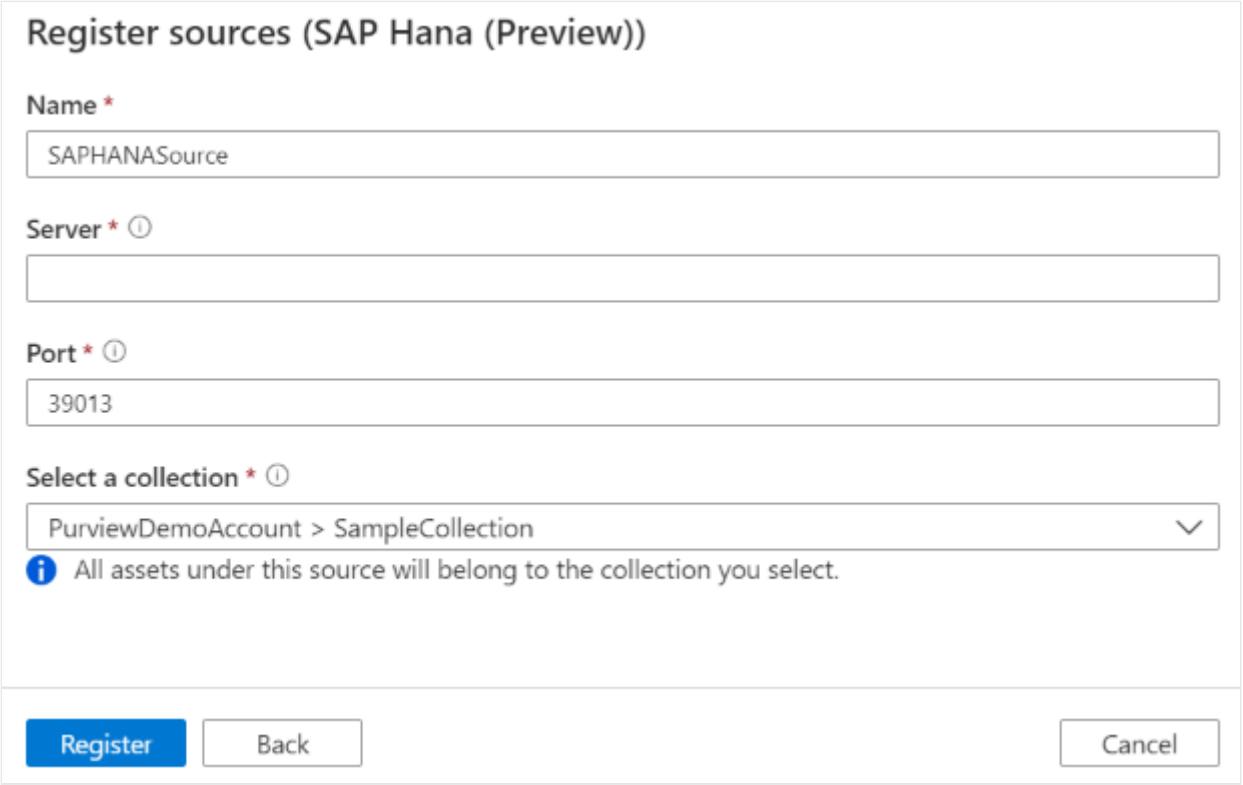
5. On the **Register sources (SAP HANA)** screen, do the following:

   a. For **Name**, enter a name that Microsoft Purview will list as the data source.

   b. For **Server**, enter the host name or IP address used to connect to an SAP HANA source. For example, `MyDatabaseServer.com` or `192.169.1.2`.

   c. For **Port**, enter the port number used to connect to the database server (39013 by default for SAP HANA).

   d. Select a collection from the list.



6. Select **Finish**.

Scan

Use the following steps to scan SAP HANA databases to automatically identify assets. For more information about scanning in general, see Scans and ingestion in Microsoft Purview.

Authentication for a scan

The supported authentication type for an SAP HANA source is **Basic authentication**.

Create and run scan

1. In the Management Center, select integration runtimes. Make sure that a self-hosted integration runtime is set up. If it isn't set up, use the steps in in prerequisites.

2. Go to **Sources**.

3. Select the registered SAP HANA source.

4. Select **+ New scan**.

5. Provide the following details:

   a. **Name**: Enter a name for the scan.

   b. **Connect via integration runtime**: Select the configured self-hosted integration runtime.

   c. **Credential**: Select the credential to connect to your data source. Make sure to:

      - Select **Basic Authentication** while creating a credential.
      - Provide the user name used to connect to the database server in the User name input field.
      - Store the user password used to connect to the database server in the secret key.

      For more information, see Credentials for source authentication in Microsoft Purview.

   d. **Database**: Specify the name of the database instance to import.

   e. **Schema**: List subset of schemas to import expressed as a semicolon separated list. For example, `schema1; schema2`. All user schemas are imported if that list is empty. All system schemas and objects are ignored by default.

      Acceptable schema name patterns that use SQL `LIKE` expression syntax include the percent sign (%). For example, `A%; %B; %C%; D` means:

      - Start with A or
      - End with B or
      - Contain C or
      - Equal D

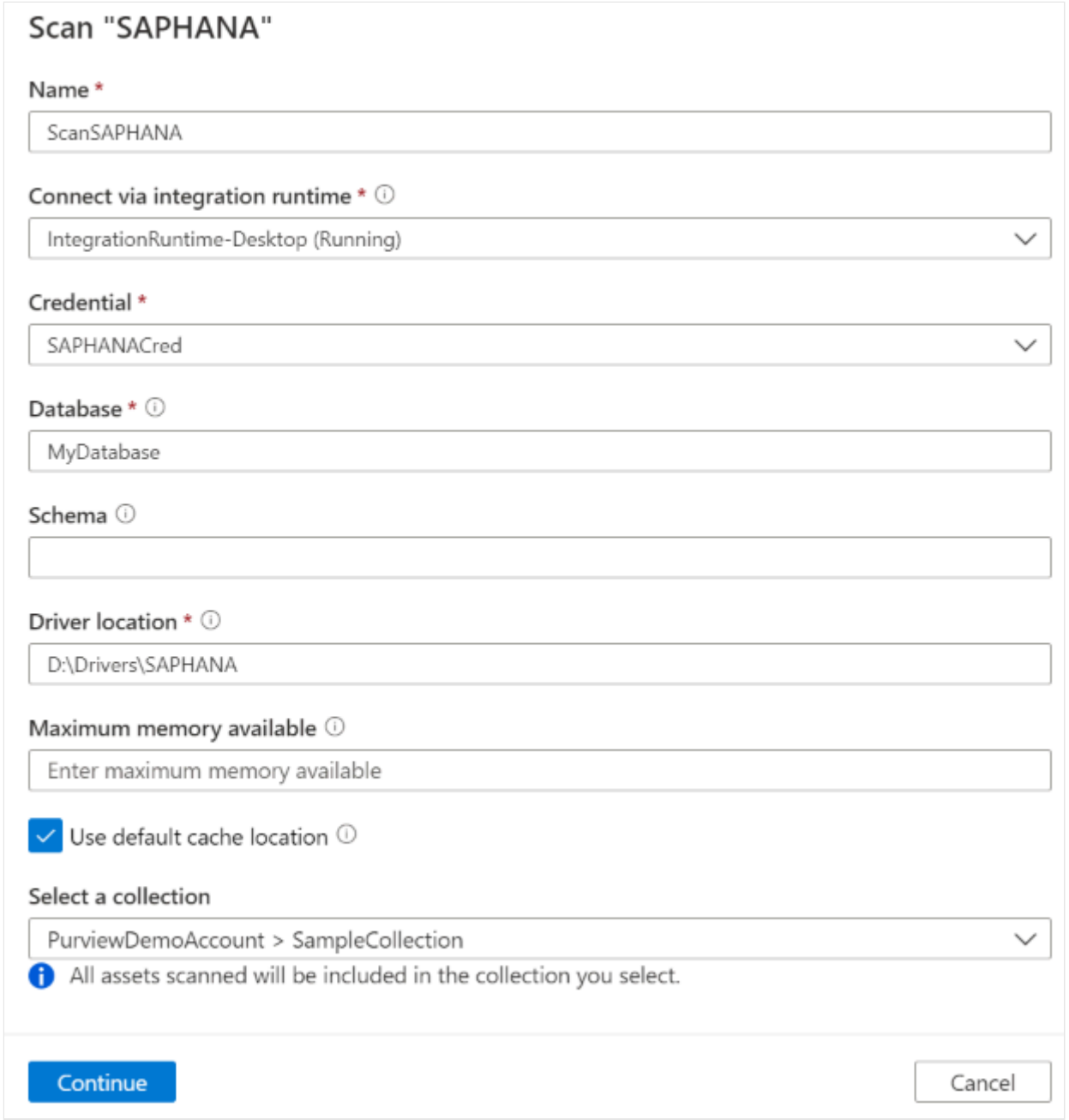
      Usage of NOT and special characters aren't acceptable.

   f. **Driver location**: Specify the path to the JDBC driver location in your machine where self-host integration runtime is running. For example: `D:\Drivers\SAPHANA`.

i. For self-hosted integration runtime on a local machine: `D:\Drivers\SAPHANA`. It's the path to valid JAR folder location. The value must be a valid absolute file path and doesn't contain space. Make sure the driver is accessible by the self-hosted integration runtime; learn more from prerequisites section.

ii. For Kubernetes-supported self-hosted integration runtime: `./drivers/SAPHANA`. It's the path to valid JAR folder location. The value must be a valid relative file path. Refer to the documentation to set up a scan with external drivers for uploading drivers in advance.

g. **Maximum memory available**: Maximum memory (in gigabytes) available on the customer's machine for the scanning processes to use. This value is dependent on the size of SAP HANA database to be scanned.



6. Select **Continue**.

7. For **Scan trigger**, choose whether to set up a schedule or run the scan once.

8. Review your scan and select **Save and Run**.

View your scans and scan runs

To view existing scans:

1. Go to the Microsoft Purview portal. On the left pane, select **Data map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that has results you want to view. The pane shows you all the previous scan runs, along with the status and metrics for each scan run.
4. Select the run ID to check the scan run details.

Manage your scans

To edit, cancel, or delete a scan:

1. Go to the Microsoft Purview portal. On the left pane, select **Data Map**.

2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.

3. Select the scan that you want to manage. You can then:

    - Edit the scan by selecting **Edit scan**.
    - Cancel an in-progress scan by selecting **Cancel scan run**.
    - Delete your scan by selecting **Delete scan**.

> ⓘ **Note**
>
> - Deleting your scan does not delete catalog assets created from previous scans.

Next steps

Now that you've registered your source, use the following guides to learn more about Microsoft Purview and your data:

- Data Estate Insights in Microsoft Purview

- Lineage in Microsoft Purview
- Search the Unified Catalog

Last updated on 04/27/2026

Connect to and manage SAP S/4HANA in Microsoft Purview

06/25/2025

This article outlines how to register SAP S/4HANA, and how to authenticate and interact with SAP S/4HANA in Microsoft Purview. For more information about Microsoft Purview, read the introductory article.

Supported capabilities

Scanning capabilities

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan
Yes	Yes	No	No

When scanning SAP S/4HANA source, Microsoft Purview supports:

- Extracting technical metadata including:
  - Instance
  - Application components
  - Packages
  - Tables including the fields, foreign keys, indexes, and index members
  - Views including the fields
  - Transactions
  - Programs
  - Classes
  - Function groups
  - Function modules
  - Domains including the domain values
  - Data elements
- Fetching static lineage on assets relationships among tables and views.

Other capabilities

For **classifications**, **sensitivity labels**, **policies**, **data lineage**, and **live view**, see the list of supported capabilities.

Known limitations

When object is deleted from the data source, currently the subsequent scan won't automatically remove the corresponding asset in Microsoft Purview.

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- An active Microsoft Purview account.
- You need Data Source Administrator and Data Reader permissions to register a source and manage it in the Microsoft Purview governance portal. For more information about permissions, see Access control in Microsoft Purview.
- Set up the right integration runtime for your scenario:
  - **To use a self-hosted integration runtime:**
    - Follow the article to create and configure a self-hosted integration runtime.
    - Ensure JDK 11 is installed on the machine where the self-hosted integration runtime is installed. Restart the machine after you newly install the JDK for it to take effect.
    - Ensure that Visual C++ Redistributable (version Visual Studio 2012 Update 4 or newer) is installed on the machine where the self-hosted integration runtime is running. If you don't have this update installed, download it now.
    - Download the 64-bit SAP Connector for Microsoft .NET 3.0 from SAP's website and install it on the self-hosted integration runtime machine. During installation, make sure you select the **Install Assemblies to GAC** option in the **Optional setup steps** window.
    - The connector reads metadata from SAP using the SAP Java Connector (JCo) 3.0 API. Make sure the Java Connector is available on your virtual machine where self-hosted integration runtime is installed. Make sure that you're using the correct JCo distribution for your environment. For example: on a Microsoft Windows machine, make sure the sapjco3.jar and sapjco3.dll files are available. Note down the folder path which you'll use to set up the scan.
    - Self-hosted integration runtime communicates with the SAP server over dispatcher port 32NN and gateway port 33NN, where NN is your SAP instance number from 00 to 99. Make sure the outbound traffic is allowed on your firewall.
  - **To use a kubernetes supported self-hosted integration runtime:**
    - Follow the article to create and configure a kubernetes supported integration runtime.

- Download the 64-bit SAP Connector for Microsoft .NET 3.0 from SAP's website and install it on the self-hosted integration runtime machine. During installation, make sure you select the **Install Assemblies to GAC** option in the **Optional setup steps** window.
- The connector reads metadata from SAP using the SAP Java Connector (JCo) 3.0 API. Make sure the Java Connector is available on your virtual machine where self-hosted integration runtime is installed. Make sure that you're using the correct JCo distribution for your environment. For example: on a Microsoft Windows machine, make sure the sapjco3.jar and sapjco3.dll files are available. Note down the folder path which you'll use to set up the scan.
- Self-hosted integration runtime communicates with the SAP server over dispatcher port 32NN and gateway port 33NN, where NN is your SAP instance number from 00 to 99. Make sure the outbound traffic is allowed on your firewall.

> **Note**
>
> Scanning SAP ECC is a memory intensive operation, you're recommended to install Self-hosted Integration Runtime on a machine with at least 128 GB RAM.

- Deploy the metadata extraction ABAP function module on the SAP server by following the steps mentioned in ABAP functions deployment guide. You need an ABAP developer account to create the RFC function module on the SAP server. For scan execution, the user account requires sufficient permissions to connect to the SAP server and execute the following RFC function modules:
  - STFC_CONNECTION (check connectivity)
  - RFC_SYSTEM_INFO (check system information)
  - OCS_GET_INSTALLED_COMPS (check software versions)
  - Z_MITI_DOWNLOAD (main metadata import, the function module you create following the Purview guide)

  The underlying SAP Java Connector (JCo) libraries might call more RFC function modules; for example, RFC_PING, RFC_METADATA_GET, etc. See SAP support note 460089 for details.

Register

This section describes how to register SAP S/4HANA in Microsoft Purview using the Microsoft Purview governance portal.

Authentication for registration

The only supported authentication for SAP S/4HANA source is **Basic authentication**.

Steps to register

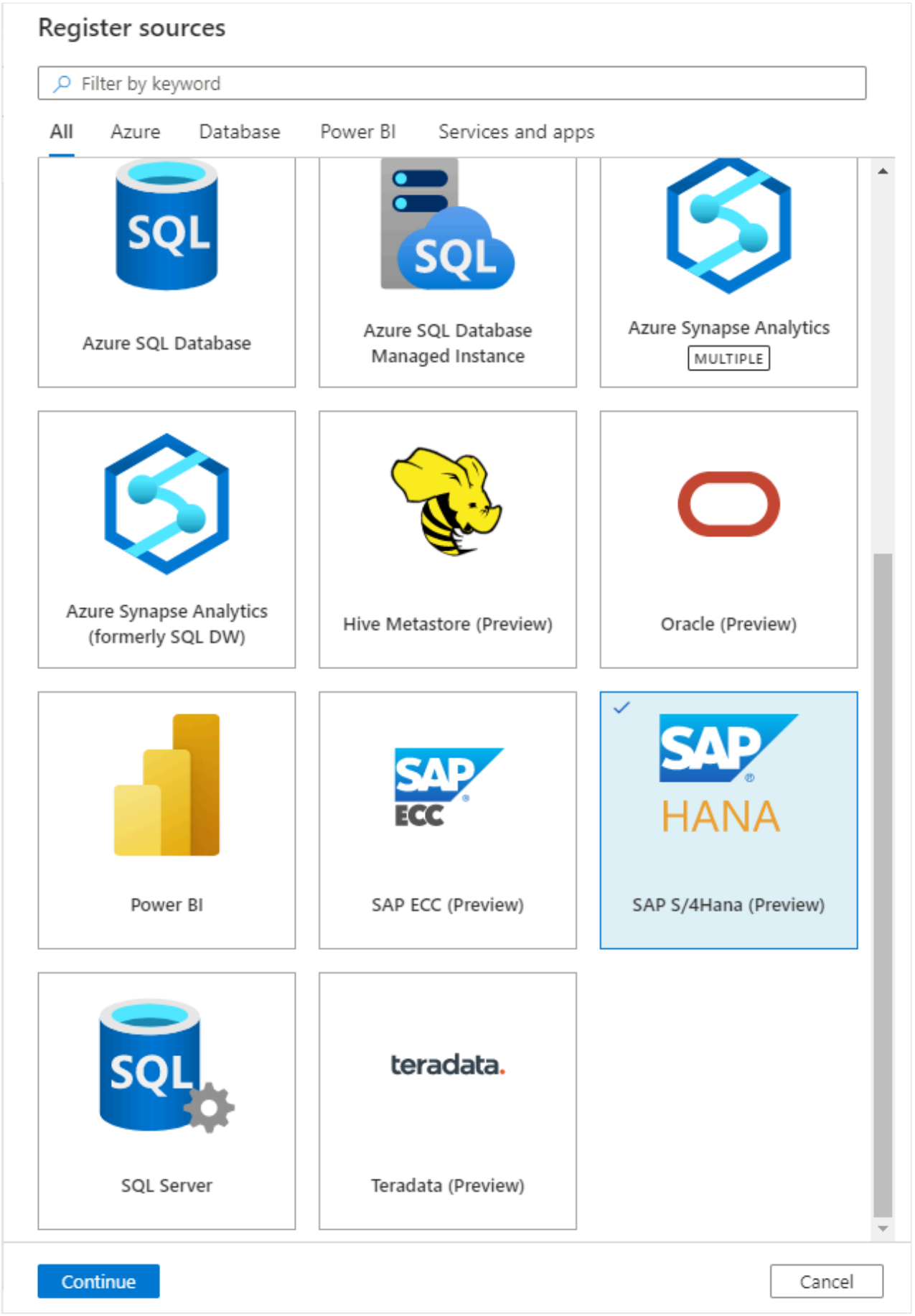
1. Open the Microsoft Purview governance portal by:

    - Browsing directly to https://web.purview.azure.com and selecting your Microsoft Purview account.
    - Opening the Azure portal, searching for and selecting the Microsoft Purview account. Selecting the **the Microsoft Purview governance portal** button.

2. Select **Data Map** on the left navigation.

3. Select **Register**

4. On Register sources, select **SAP S/4HANA.** Select **Continue**



On the **Register sources (SAP S/4HANA)** screen, do the following:

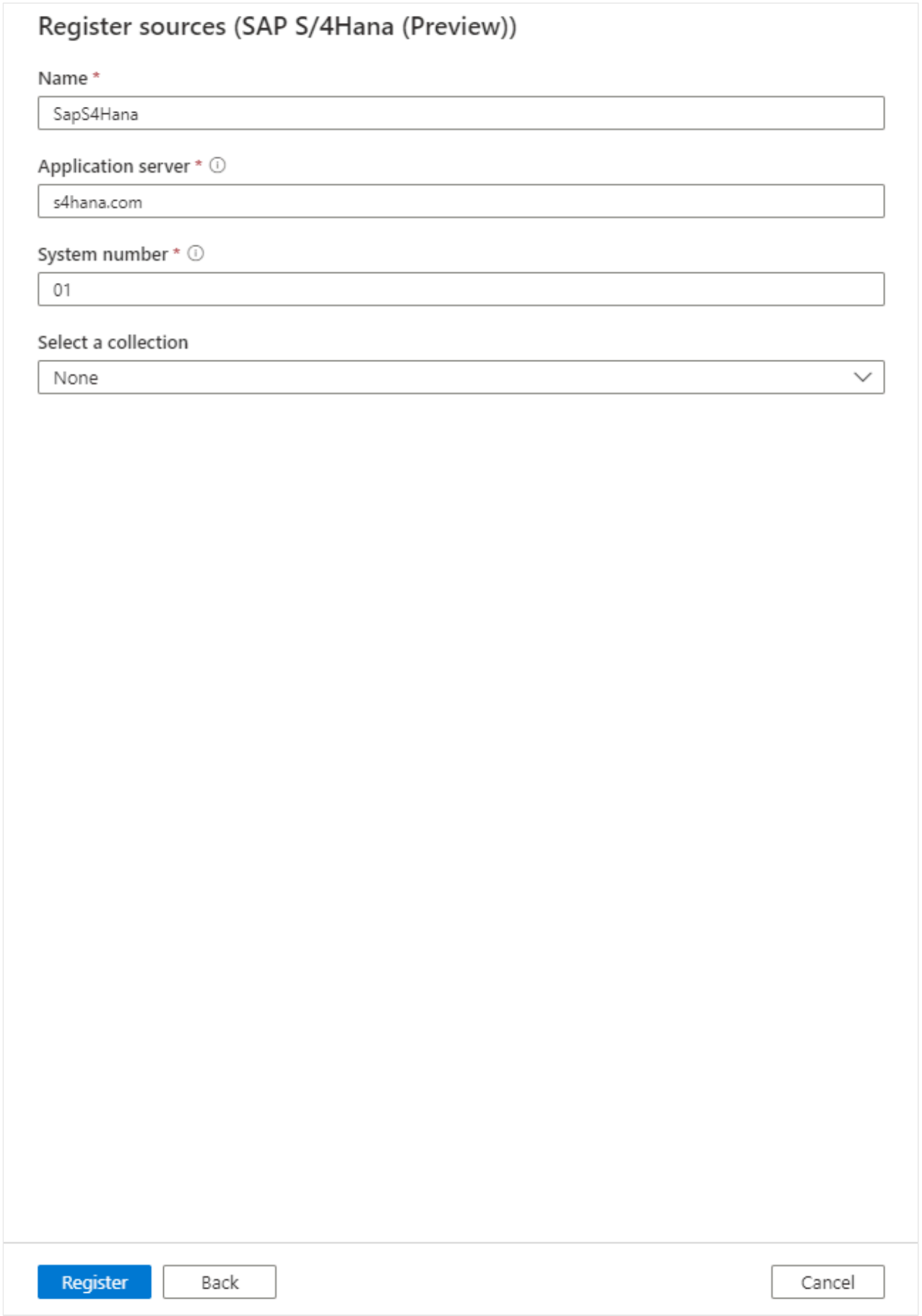
1. Enter a **Name** that the data source will be listed within the Catalog.

2. Enter the **Application server** name to connect to SAP S/4HANA source. It can also be an IP address of the SAP application server host.

3. Enter the SAP **System number**. This is a two-digit integer between 00 and 99.

4. Select a collection from the list.

5. Finish to register the data source.


Register sources (SAP S/4Hana (Preview))
Name *
SapS4Hana
Application server *
s4hana.com
System number *
01
Select a collection
None
Register
Back
Cancel

Scan

Follow the steps below to scan SAP S/4HANA to automatically identify assets. For more information about scanning in general, see our introduction to scans and ingestion.

Create and run scan

1. In the Management Center, select Integration runtimes. Make sure a self-hosted integration runtime is set up. If it isn't set up, use the steps mentioned in prerequisites to create a self-hosted integration runtime

2. Navigate to **Sources.**

3. Select the registered SAP S/4HANA source.

4. Select **+ New scan**

5. Provide the below details:

    a. **Name**: The name of the scan

    b. **Connect via integration runtime**: Select the configured self-hosted integration runtime.

    c. **Credential**: Select the credential to connect to your data source. Make sure to:

    - Select Basic Authentication while creating a credential.
    - Provide a user ID to connect to SAP server in the User name input field.
    - Store the user password used to connect to SAP server in the secret key.

    d. **Client ID**: Enter the SAP Client ID. It's a three-digit numeric number from 000 to 999.

    e. **SNC Mode (Optional)**: Turn the toggle **on** if you prefer to use Secure Network Communications (SNC) mode for securing your technical connections to SAP RFCs through a certificate-based authentication mechanism. SNC Mode is **off** by default.

    Follow these steps to set up the SAP Personal Secure Environment with SNC Certificate:

    Get SAPCAR

    - Go to the SAP Software Download Center and sign in with your SAP credentials.
    - Search for SAPCAR and select the latest nonarchived version.
    - Select your operating system.
    - Download the `.EXE file to C:\sap\SAR`.

    Get SAP Common Crypto Library

- In the SAP Software Download Center, search for "COMMONCRYPTOLIB" and select the latest version.
- Select your operating system.
- Download the .SAR file with the most recent release date to `C:\sap\SAR`.

Extract SAP Common Crypto Library

- Open PowerShell and navigate to `C:\sap\SAR`.
- Enter the following command, replacing xxxx with your values: `.\SAPCAR_xxxx.EXE -xvf .\SAPCRYPTOLIBP_xxxx.SAR -R .\..\libs\sapcryptolib`.
- Confirm that `sapgenpse.exe` is in the `C:\sap\libs\sapcryptolib` directory.

Generate certificates

Note: This method is for demonstration purposes only and isn't recommended for production systems. For production systems, consult your internal PKI guidance or security team.

- Set up the folder structure:
  - `mkdir rootCA`
  - `mkdir sncCert`
- Create the necessary serial and index files if they don't exist:
  - `if (-Not (Test-Path "rootCA\index.txt")) { New-Item -Path "rootCA\index.txt" ItemType File }`
  - `if (-Not (Test-Path "rootCA\serial")) { Set-Content -Path "rootCA\serial" -Value "01" }`
- Generate a Root CA:
  - `openssl genpkey -algorithm RSA -out rootCA/ca.key.pem -pkeyopt rsa_keygen_bits:2048`
  - `openssl req -x509 -new -key rootCA/ca.key.pem -days 7305 -sha256 -extensions v3_ca -out rootCA/ca.cert.pem -subj "/O=Contoso/CN=Root CA"`
- Generate the SNC certificate
  - `openssl genrsa -out sncCert/snc.key.pem 2048`
  - `openssl req -key sncCert/snc.key.pem -new -sha256 -out sncCert/snc.csr.pem subj "/O=Contoso/CN=SNC"`
- Create an OpenSSL configuration file, `sncCert/extensions.cnf`, for signing:
  - `subjectKeyIdentifier = hash`
  - `authorityKeyIdentifier = keyid,issuer`

- `basicConstraints = critical,CA:false`
- `keyUsage = critical,digitalSignature,keyEncipherment,dataEncipherment`
- `extendedKeyUsage = clientAuth,emailProtection`

- Sign the SNC certificate with the Root CA:
  - `openssl x509 -req in sncCert/snc.csr.pem CA rootCA/ca.cert.pem CAkey rootCA/ca.key.pem CAcreateserial out sncCert/snc.cert.pem days 3650 sha256 extfile sncCert\extensions.cnf extensions v3_leaf`

Create a Personal Secure Environment

Create a Personal Secure Environment (PSE) for the on-premises data gateway. The NCo library looks for the SNC certificate inside the PSE.

- Create a PKCS#12 container:
  - `openssl pkcs12 -export -out snc.p12 -inkey sncCert\snc.key.pem -in sncCert\snc.cert.pem -certfile rootCA\ca.cert.pem`

- Create the SECUDIR environment variable.
- Open **System Properties**: In **File Explorer**, right-click **This PC**, and then select **Properties** > **Advanced system settings**.
- Select **Environment Variables**.
- Under **System variables**, select **New**.
- Set the variable name to **SECUDIR**.
- Set the value to `C:\sapsecudir`.
- Select **OK**.

Import the PKCS#12 container into a PSE: `C:\sap\libs\sapcryptolib\sapgenpse.exe import_p12 -p SAPSNCSKERB.pse C:\pki certs\snc.p12`

Allow your SHIR process to use your SAP PSE

Verify which user or service is being used by your SNC client to obtain the certificate to communicate with SAP. The Microsoft Purview SHIR uses the service user **NT SERVICE\DIAHostService**.

- Add a credential to allow the certificate retrieval request from the PSE: `.\sapgenpse.exe seclogin -p C:\sapsecudir\SAPSNCSKERB.pse -x your-pse-pin -O "NT SERVICE\DIAHostService"`

- Verify credentials like this: `.\sapgenpse.exe seclogin -l -O "NT SERVICE\DIAHostService"`
- You can delete them like this: `.\sapgenpse.exe seclogin -d -O "NT SERVICE\DIAHostService"`
- Use the -h parameter to get help with the sapgenpse command line tool or check the command reference here.

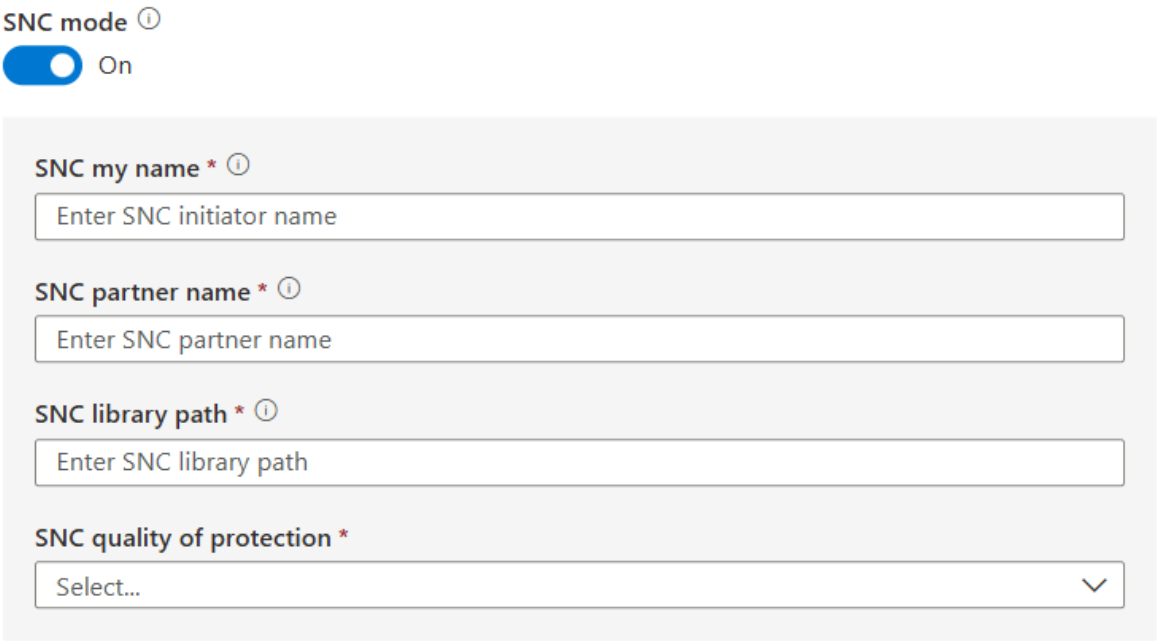
Give permission to access the SAP JCo library to SHIR Process

Verify if Microsoft Purview SHIR user **NT SERVICE\DIAHostService** has **Read/Write/Execute/List** permissions over these folders:

- SNC Library Path: The folder that has the SNC Library, `sapcrypto.dll`. Example: `C:\Users\shir-admin\Desktop\snc\SAPCRYPTOLIBP_8557-20011729\sapcrypto.dll`
- JCo Library Path: The folder that has the SAP Java Connector jar file. Example: `C:\Users\shir-admin\Desktop\snc\sapjco3-ntamd64-3.1\sapjco3-ntamd64-3.1.3`

SAP NCo Library should also be installed on the virtual machine for test connection to work properly, as it still uses NCo library. This isn't necessary, and Test Connection can be skipped.

Enter SCN details for **my name**, **partner name**, **library path**, and **quality of protection**, as seen below, and initiate the scan:

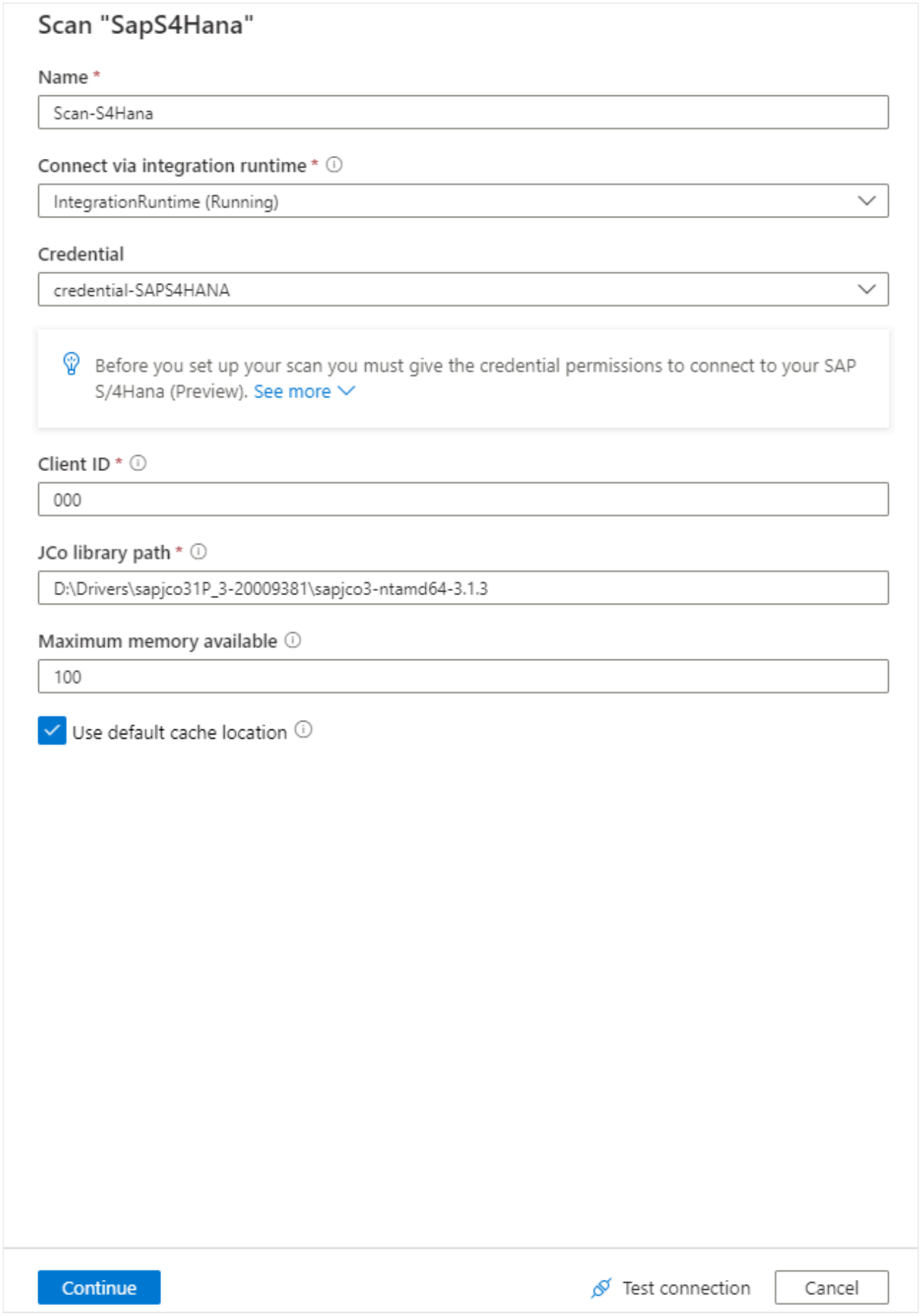


f. **JCo library path**: Specify the directory path where the JCo libraries are located, for example: `D:\Drivers\SAPJCo`. Make sure the path is accessible by the self-hosted integration runtime, learn more from prerequisites section.

i. For self-hosted integration runtime on a local machine: `D:\Drivers\SAPJCo`. It's the path to valid JAR folder location. The value must be a valid absolute file path and doesn't contain space. Make sure the driver is accessible by the self-hosted integration runtime; learn more from prerequisites section.

ii. For Kubernetes-supported self-hosted integration runtime: `./drivers/SAPJCo`. It's the path to valid JAR folder location. The value must be a valid relative file path. Refer to the documentation to set up a scan with external drivers for uploading drivers in advance.

g. **Maximum memory available:** Maximum memory (in GB) available on customer's VM to be used by scanning processes. This is dependent on the size of SAP S/4HANA source to be scanned. It's recommended to provide large available memory, for example, 100.



6. Select **Continue**.

7. Choose your **scan trigger**. You can set up a schedule or ran the scan once.

8. Review your scan and select **Save and Run**.

View your scans and scan runs

To view existing scans:

1. Go to the Microsoft Purview portal. On the left pane, select **Data map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that has results you want to view. The pane shows you all the previous scan runs, along with the status and metrics for each scan run.
4. Select the run ID to check the scan run details.

Manage your scans

To edit, cancel, or delete a scan:

1. Go to the Microsoft Purview portal. On the left pane, select **Data Map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that you want to manage. You can then:
    - Edit the scan by selecting **Edit scan**.
    - Cancel an in-progress scan by selecting **Cancel scan run**.
    - Delete your scan by selecting **Delete scan**.

> ⓘ **Note**
>
> - Deleting your scan does not delete catalog assets created from previous scans.

Lineage

After scanning your SAP S/4HANA source, you can browse Unified Catalog or search Unified Catalog to view the asset details.

Go to the asset -> lineage tab, you can see the asset relationship when applicable. Refer to the supported capabilities section on the supported SAP S/4HANA lineage scenarios. For more information about lineage in general, see data lineage and lineage user guide.

Next steps

Now that you've registered your source, follow the below guides to learn more about Microsoft Purview and your data.

- Data Estate Insights in Microsoft Purview
- Lineage in Microsoft Purview
- Search Unified Catalog

SAP ABAP function module deployment guide

Article • 04/18/2024

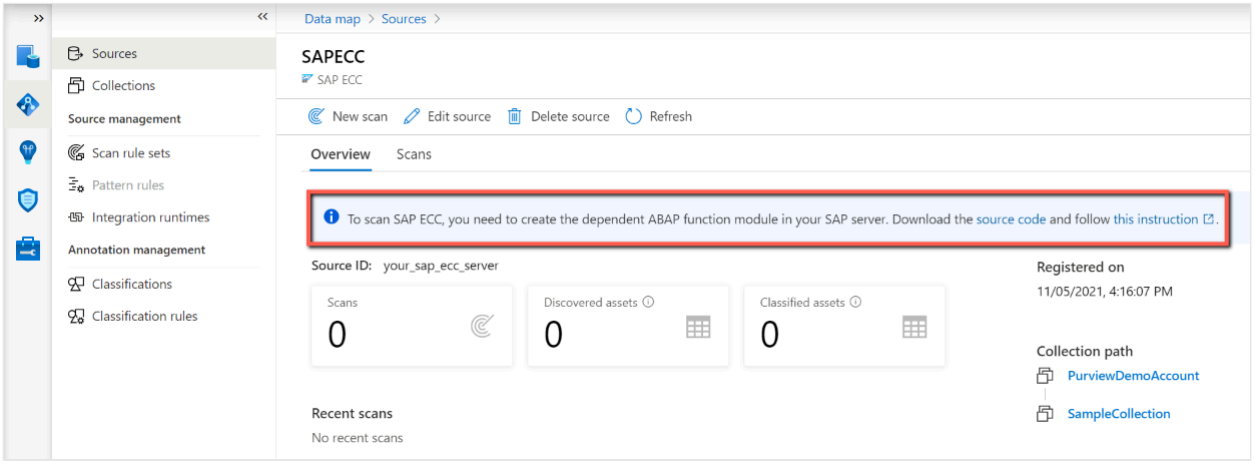
When you scan SAP ECC, SAP S/4HANA, and SAP BW sources in Microsoft Purview, you need to create the dependent ABAP function module in your SAP server. Microsoft Purview invokes this function module to extract the metadata from your SAP system during scan.

This article describes the steps required to deploy this module.

> **Note**
>
> The following instructions were compiled based on the SAP GUI v. 7.2.

Prerequisites

Download the SAP ABAP function module source code from the Microsoft Purview governance portal. After you register a source for SAP ECC, SAP S/4HANA, or SAP BW, you can find a download link on top as shown in the following image. You can also see the link when you create a new scan or edit a scan. To download this code, you need to have Data Source Administrator permissions on the root collection in your data map.



Deploy the module

Follow the instructions to deploy a module.

Create a package

This step is optional, and an existing package can be used.

1. Sign in to the SAP server and open **Object Navigator** (SE80 transaction).
2. Select **Package** from the list and enter a name for the new package. For example, use **Z_MITI**. Then select **Display**.
3. In the **Create Package** window, select **Yes**. In the **Package Builder: Create Package** window, enter a value in the **Short Description** box. Select the **Continue** icon.
4. In the **Prompt for local Workbench request** window, select **Own Requests**. Select the **development** request.

Create a function group

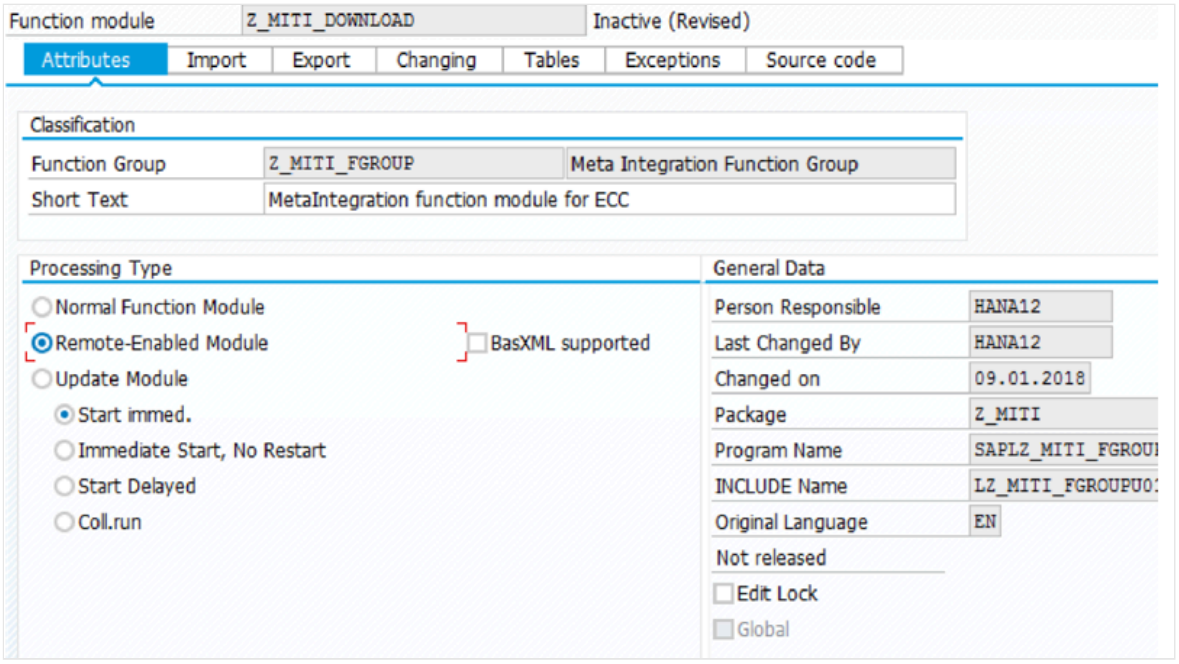
1. In **Object Navigator**, select **Function Group** from the list and enter a name in the input box. For example, use **Z_MITI_FGROUP**. Select the **View** icon.
2. In the **Create Object** window, select **yes** to create a new function group.
3. Enter a description in the **Short Text** box and select **Save**.
4. Select a package that was prepared in the **Create a Package** step, and select **Save**.
5. Confirm a request by selecting **Continue**.
6. Activate the function group.

Create the ABAP function module

1. After the function group is created, select it.
2. Select and hold (or right-click) the function group name in the repository browser. Select **Create** and then select **Function Module**.
3. In the **Function module** box, enter **Z_MITI_DOWNLOAD** in case of SAP ECC or S/4HANA and **Z_MITI_BW_DOWNLOAD** in case of SAP BW. Enter a description in the **Short Text** box.

After the module is created, specify the following information:

1. Go to the **Attributes** tab.
2. Under **Processing Type**, select **Remote-Enabled Module**.



3. Go to the **Source code** tab. There are two ways to deploy code for the function:

   a. On the main menu, upload the text file you downloaded from the Microsoft Purview governance portal as described in Prerequisites. To do so, select **Utilities** > **More Utilities** > **Upload/Download** > **Upload**.

   b. Alternatively, open the file and copy and paste the contents in the **Source code** area.

4. Go to the **Import** tab and create the following parameters:

   a. *P_AREA TYPE DD02L-TABNAME* (**Optional** = True)

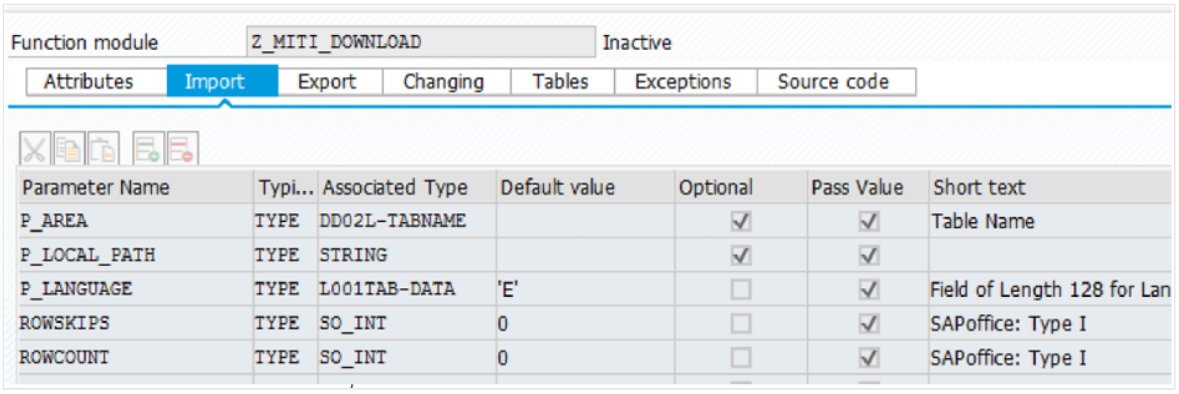
   b. *P_LOCAL_PATH TYPE STRING* (**Optional** = True)

   c. *P_LANGUAGE TYPE L001TAB-DATA DEFAULT 'E'*

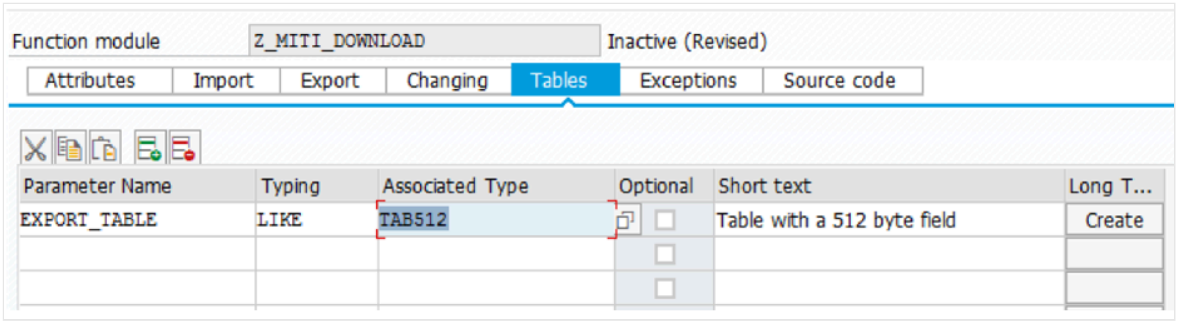
   d. *ROWSKIPS TYPE SO_INT DEFAULT 0*

   e. *ROWCOUNT TYPE SO_INT DEFAULT 0*

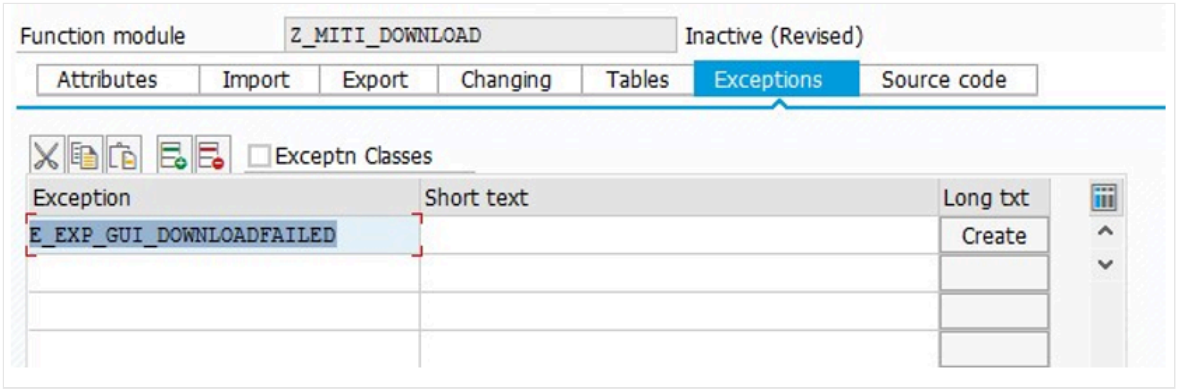
> ⓘ **Note**
>
> Select the **Pass Value** checkbox for all the parameters.



5. Go to the **Tables** tab and define *EXPORT_TABLE LIKE TAB512*.



6. Go to the **Exceptions** tab and define the exception *E_EXP_GUI_DOWNLOADFAILED*.



7. Save the function by selecting **Ctrl+S**. Or select **Function module** and then select **Save** on the main menu.

8. Select the **Activate** icon on the toolbar and then select **Continue**. You can also select **Ctrl+F3**. If prompted, select the generated includes to be activated along with the main function module.

Test the function

After you finish the previous steps, test the function:

1. Open the **Z_MITI_DOWNLOAD** or **Z_MITI_BW_DOWNLOAD** function module you created.

2. On the main menu, select **Function Module** > **Test** > **Test Function Module**. You can also select **F8**.

3. Enter a path to the folder on the local file system in the parameter *P_LOCAL_PATH*. Then select the **Execute** icon on the toolbar. You can also select **F8**.

4. Enter the name of the area of interest in the **P_AREA** field if a file with metadata must be downloaded or updated. After the function finishes working, the folder indicated in the *P_LOCAL_PATH* parameter must contain several files with metadata inside. The names of files mimic areas that can be specified in the **P_AREA** field.

The function finishes its execution and metadata is downloaded much faster if it's launched on the machine that has a high-speed network connection with the SAP server.

Next steps

- Connect to and manage SAP ECC source
- Connect to and manage SAP S/4HANA source
- Connect to and manage SAP Business Warehouse (BW) source

Feedback

Was this page helpful? Yes No

Connect to and manage Snowflake in Microsoft Purview

This article outlines how to register Snowflake, and how to authenticate and interact with Snowflake in Microsoft Purview. For more information about Microsoft Purview, read the Microsoft Purview overview.

> **Important**
>
> Snowflake announced the deprecation of basic authentication and plans to disable it by August 2026. To align with the change and avoid any service disruptions, Microsoft Purview customers using the Snowflake Data Map connector should transition to Key Pair (SHIR) authentication.
>
> **Recommended action (complete by July 2026)**: To provide customers with more time, Snowflake allows service users of type **LEGACY_SERVICE** to continue using basic authentication until July 2026. This policy helps ensure there's no disruption to your service while you transition to the new authentication method. You can continue to use basic authentication for a limited period of time, while at the same time setting up and validating the new authentication mechanism Key Pair.
>
> We recommend that Microsoft Purview customers update their users to the **LEGACY_SERVICE** type. This change ensures continued access and provides more time to complete the transition. **This update must be completed before the end of July 2026 by running the following command in Snowflake**:
>
> ```
> *ALTER USER <purview_user_name> SET TYPE = LEGACY_SERVICE;
> ```
>
> The creation of new users with the LEGACY_SERVICE type isn't permitted beginning August 2026.

Supported capabilities

The following sections describe the capabilities available when scanning Snowflake in Microsoft Purview.

Scanning capabilities

The following table summarizes the scanning capabilities supported for Snowflake in Microsoft Purview.

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan
Yes	Yes	No	Yes

When scanning a Snowflake source, Microsoft Purview supports:

- Extracting technical metadata including:
  - Server
  - Databases
  - Schemas
  - Tables including the columns, foreign keys, and unique constraints
  - Views including the columns
  - Stored procedures including the parameter dataset and result set
  - Functions including the parameter dataset
  - Pipes
  - Stages
  - Streams including the columns
  - Tasks
  - Sequences
- Fetching static lineage on assets relationships among tables, views, streams, and stored procedures.

For stored procedures, you can choose the level of detail to extract in the Create and run scan section. Stored procedure lineage is supported for Snowflake Scripting (SQL) and JavaScript languages, and generated based on the procedure definition.

When setting up a scan, you can choose to scan one or more Snowflake databases entirely based on the given names or name patterns, or further scope the scan to a subset of schemas matching the given names or name patterns.

Other capabilities

For **classifications**, **sensitivity labels**, **policies**, **data lineage**, and **live view**, see the list of supported capabilities.

Known limitations

Be aware of the following limitations when scanning Snowflake in Microsoft Purview.

- When you delete an object from the data source, the subsequent scan doesn't automatically remove the corresponding asset in Microsoft Purview.
- Stored procedure lineage isn't supported for the following patterns:
  - Stored procedure defined in Java, Python, and Scala languages.
  - Stored procedure using SQL EXECUTE IMMEDIATE with static SQL query as variable.

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- An active Microsoft Purview account.
- You need Data Source Administrator and Data Reader permissions to register a source and manage it in the Microsoft Purview portal. For more information about permissions, see Access control in Microsoft Purview.
- Choose the right integration runtime configuration for your scenario.
  - To use Managed Virtual Network Integration Runtime to connect to Snowflake via private link, follow the steps in Connect to Snowflake via Managed Private Endpoint to set up the private endpoint.
  - To use Self-hosted Integration Runtime for scan, set up the latest self-hosted integration runtime. For more information, see the create and configure a self-hosted integration runtime guide.
    - Ensure JDK 11 is installed on the machine where the self-hosted integration runtime is installed. Restart the machine after you newly install the JDK for it to take effect.
    - Ensure Visual C++ Redistributable (version Visual Studio 2012 Update 4 or newer) is installed on the self-hosted integration runtime machine. If you don't have this update installed, download the Visual C++ Redistributable.

Required permissions for scan

Microsoft Purview supports basic authentication (username and password) for scanning Snowflake. The default role of the given user will be used to perform the scan. The Snowflake

user must have usage rights on a warehouse and the databases to be scanned, and read access to system tables to access advanced metadata.

Here's a sample walkthrough to create a user specifically for Microsoft Purview scan and set up the permissions. If you choose to use an existing user, make sure it has adequate rights to the warehouse and database objects.

1. Set up a `purview_reader` role. You need *ACCOUNTADMIN* rights to do this.

```SQL
USE ROLE ACCOUNTADMIN;

--create role to allow read only access - this will later be assigned to the Microsoft Purview user
CREATE OR REPLACE ROLE purview_reader;

--make sysadmin the parent role
GRANT ROLE purview_reader TO ROLE sysadmin;
```

2. Create a warehouse for Microsoft Purview to use and grant rights.

```SQL
--create warehouse - account admin required
CREATE OR REPLACE WAREHOUSE purview_wh WITH
    WAREHOUSE_SIZE = 'XSMALL'
    WAREHOUSE_TYPE = 'STANDARD'
    AUTO_SUSPEND = 300
    AUTO_RESUME = TRUE
    MIN_CLUSTER_COUNT = 1
    MAX_CLUSTER_COUNT = 2
    SCALING_POLICY = 'STANDARD';

--grant rights to the warehouse
GRANT USAGE ON WAREHOUSE purview_wh TO ROLE purview_reader;
```

3. Create a user `purview` for Microsoft Purview scan.

```SQL
CREATE OR REPLACE USER purview
    PASSWORD = '<password>';

--note the default role will be used during scan
ALTER USER purview SET DEFAULT_ROLE = purview_reader;
```

```
--add user to purview_reader role
GRANT ROLE purview_reader TO USER purview;
```

4. Grant reader rights to the database objects.

SQL

```
GRANT DATABASE ROLE SNOWFLAKE.OBJECT_VIEWER TO USER purview_reader;
GRANT USAGE ON DATABASE <your_database_name> TO purview_reader;

--grant reader access to all the database structures that purview can currently
scan
GRANT USAGE ON ALL SCHEMAS IN DATABASE <your_database_name> TO role
purview_reader;
GRANT USAGE ON ALL FUNCTIONS IN DATABASE <your_database_name> TO role
purview_reader;
GRANT USAGE ON ALL PROCEDURES IN DATABASE <your_database_name> TO role
purview_reader;
GRANT SELECT ON ALL TABLES IN DATABASE <your_database_name> TO role
purview_reader;
GRANT SELECT ON ALL VIEWS IN DATABASE <your_database_name> TO role
purview_reader;
GRANT USAGE, READ on ALL STAGES IN DATABASE <your_database_name> TO role
purview_reader;

--grant reader access to any future objects that could be created
GRANT USAGE ON FUTURE SCHEMAS IN DATABASE <your_database_name> TO role
purview_reader;
GRANT USAGE ON FUTURE FUNCTIONS IN DATABASE <your_database_name> TO role
purview_reader;
GRANT USAGE ON FUTURE PROCEDURES IN DATABASE <your_database_name> TO role
purview_reader;
GRANT SELECT ON FUTURE TABLES IN DATABASE <your_database_name> TO role
purview_reader;
GRANT SELECT ON FUTURE VIEWS IN DATABASE <your_database_name> TO role
purview_reader;
GRANT USAGE, READ ON FUTURE STAGES IN DATABASE <your_database_name> TO role
purview_reader;
```

Register

This section describes how to register Snowflake in Microsoft Purview using the classic Microsoft Purview governance portal.

Steps to register

To register a new Snowflake source in Microsoft Purview Unified Catalog, follow these steps:

1. Navigate to your Microsoft Purview account in the classic Microsoft Purview governance portal.
2. Select **Data Map** on the left navigation.
3. Select **Register.**
4. On Register sources, select **Snowflake**. Select **Continue**.

On the **Register sources (Snowflake)** screen, follow these steps:

1. Enter a **Name** that the data source will be listed within the Catalog.

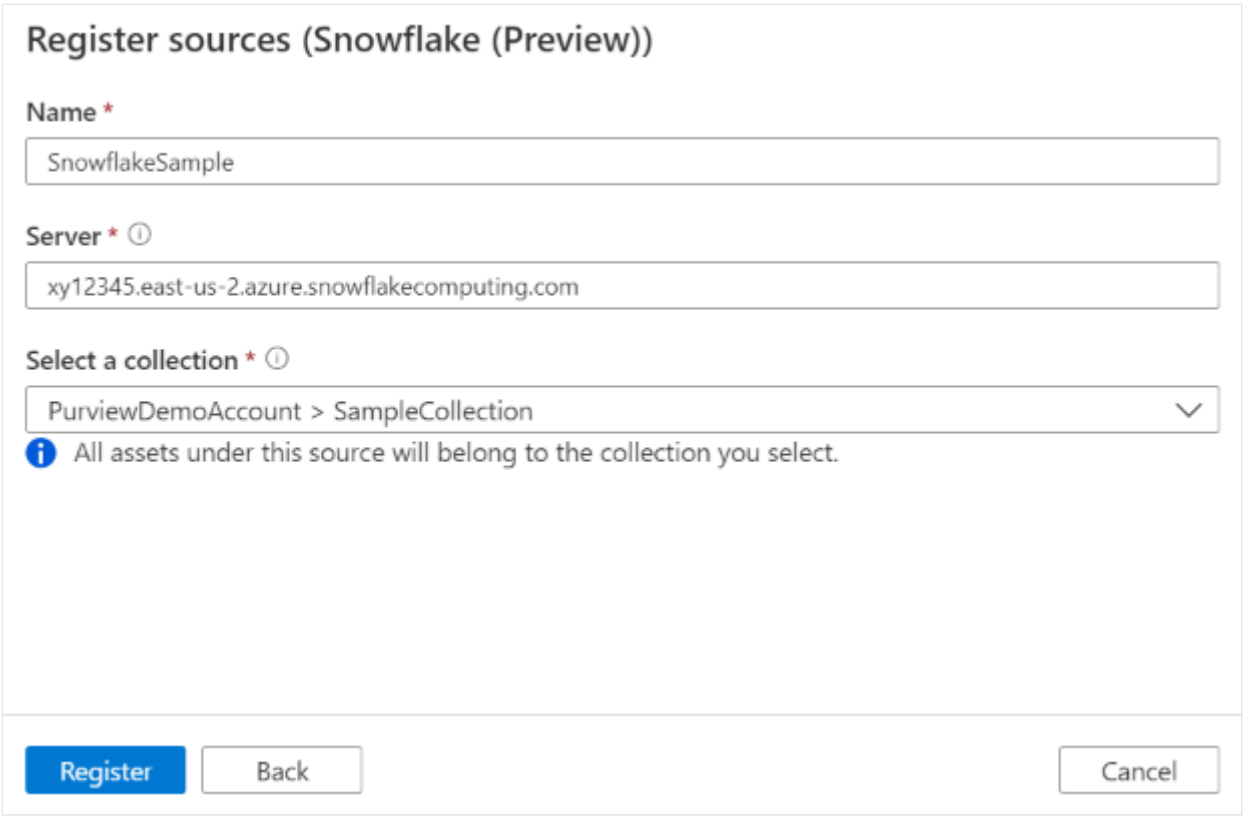
2. Enter the **server** URL in the form of `<account_identifier>.snowflakecomputing.com`, for example, `orgname-accountname.snowflakecomputing.com`. Learn more about Snowflake account identifier. Note this URL is used as part of the Snowflake assets' fully qualified name, and is the default endpoint for Microsoft Purview to connect to Snowflake during scan.

3. Add **additional hosts** if applicable. Specify it when you want the scan operations to connect to a Snowflake endpoint other than the server URL. You can choose the host for connection during scan setup.

   > **Tip**
   >
   > If you already scanned Snowflake but want to switch to use another endpoint, for example, from public endpoint to private endpoint, you can add another host in data source and choose that host for connection in scan to ensure Microsoft Purview generate the assets with the same fully qualified name as previous.
   >
   > When registering a data source, Microsoft Purview performs uniqueness check that the server URL and other hosts doesn't overlap with any existing sources.

4. Select a collection from the list.

5. Finish to register the data source.



Scan

Follow these steps to scan Snowflake to automatically identify assets. For more information about scanning in general, see our introduction to scans and ingestion.

Authentication for a scan

The supported authentication types for a Snowflake source are **Basic authentication** and **Key Pair**.

Follow this process to configure a Key Pair credential.

Prerequisites for Key Pair Authentication

Before starting, ensure you have:

1. **Snowflake Access**

   a. Snowflake account with admin privileges

   b. User account for Microsoft Purview scanning (for example, purview_scanner)

   c. Access to execute ALTER USER commands

2. **Tools & Software**

a. OpenSSL installed (for key generation)

b. Azure CLI or PowerShell (for Key Vault operations)

c. Access to Microsoft Purview Portal

d. Ensure you have Microsoft Open JDK 11. Download Microsoft Open JDK 11. Other JDK variants might have issues with advanced encrypted algorithms.

Step by Step Configuration

Part 1: Snowflake Configuration

This step involves creating public and private keys and then assigning the public key to the purview user ('purview_scanner' in below example) that connects to Snowflake to perform the scan.

Step 1.1: Generate RSA Key Pair

- Generate Private Key

Open a terminal and generate the private key:

```text
--For Encrypted Private Key (Recommended):

openssl genrsa 2048 | openssl pkcs8 -topk8 -v2 aes-256-cbc -v2prf hmacWithSHA512 -
inform PEM -out rsa_key.p8

--Note: You'll be prompted to enter a passphrase. Remember this passphrase!
```

OR

```text
--For Unencrypted Private Key:

openssl genrsa 2048 | openssl pkcs8 -topk8 -v2 aes-256-cbc -v2prf hmacWithSHA512 -
inform PEM -out rsa_key.p8 -nocrypt
```

Note: Use the above commands to generate the key

- Generate Public Key

```text
```

```
openssl rsa -in rsa_key.p8 -pubout -out rsa_key.pub
```

- Extract Public Key Content

View and copy the public key content (excluding headers):

text

```
cat rsa_key.pub
```

text

```
-----BEGIN PUBLIC KEY-----
MIIBIjANBgkqhkiG9w0BAQEFAAOCAQ8AMIIBCgKCAQEA...
... -----END PUBLIC KEY-----
```

Copy everything between the BEGIN and END lines (don't include the delimiter lines).

Step 1.2: Assign Public Key to Snowflake User

Connect to Snowflake and execute:

SQL

```
-- Create user if it doesn't exist
CREATE USER IF NOT EXISTS purview_scanner
  DEFAULT_ROLE = 'PUVRIEW_READER'
  TYPE = 'SERVICE'  ;
 -- Assign the public key (paste your key without the BEGIN/END lines)
ALTER USER purview_scanner
SET RSA_PUBLIC_KEY='MIIBIjANBgkqhkiG9w0BAQEFAAOCAQ8AMIIBCgKCAQEA...';
 -- Grant necessary privileges
GRANT ROLE purview_reader TO USER purview_scanner;
 -- Verify the fingerprint
DESC USER purview_scanner;
```

Step 1.3: Verify Key Configuration [optional step to validate]

Verify the public key fingerprint matches:

In Snowflake:

SQL

```
DESC USER purview_scanner ->> SELECT SUBSTR(
(SELECT "value" FROM $1
```

```
WHERE "property" = 'RSA_PUBLIC_KEY_FP'),
LEN('SHA256:') + 1) AS key;
```

In Command Line:

```
text
openssl rsa -pubin -in rsa_key.pub -outform DER | openssl dgst -sha256 -binary |
openssl
enc -base64
```

Note that both should produce the same fingerprint hash.

Part 2: Azure Key Vault Setup

Step 2.1: Prepare Private Key Content

View your private key:

```
text
cat rsa_key.p8
```

Copy the entire content including the -----BEGIN ENCRYPTED PRIVATE KEY----- and -----END ENCRYPTED PRIVATE KEY----- lines.

Step 2.2: Store Secrets in Azure Key Vault

You need to store TWO secrets in Azure Key Vault:

Secret 1: Private Key

CRITICAL REQUIREMENT: The private key **MUST be stored using Azure CLI or Azure Cloud Shell**.

DO NOT use the Azure Portal UI for creating the private key secret because:

- The Azure Portal text input doesn't preserve multi-line formatting
- It converts line breaks to spaces, corrupting the PEM format
- **Snowflake connections will FAIL** with malformed key errors

The password/passphrase secret CAN be created via Portal (single-line value).

Method 1: Azure CLI (Recommended)

```text
# Replace with your values
KEY_VAULT_NAME="your-keyvault-name"
CREDENTIAL_NAME="snowflake-purview-cred"

# Store the private key from file (preserves multi-line format)
az keyvault secret set \
 --vault-name $KEY_VAULT_NAME \
 --name "${CREDENTIAL_NAME}-privateKey" \
 --file rsa_key.p8
```

OR

Method 2: Azure Cloud Shell

1. Open Azure Cloud Shell (https://shell.azure.com)
2. Upload your rsa_key.p8 file using the upload button (folder icon)
3. Run the command:

```text
az keyvault secret set \
  --vault-name your-keyvault-name \
  --name "snowflake-purview-cred-privateKey" \
  --file rsa_key.p8
```

OR

Method 3: PowerShell with Azure CLI

```PowerShell
# Read the private key preserving line breaks
$privateKeyContent = Get-Content -Path "rsa_key.p8" -Raw

# Store in Key Vault using Azure CLI
az keyvault secret set `
  --vault-name "your-keyvault-name" `
  --name "snowflake-purview-cred-privateKey" `
  --value $privateKeyContent
```

OR

Verify the Private Key Format After Upload: [optional step]

text

```
# Retrieve and check the secret maintains proper formatting
az keyvault secret show \
  --vault-name $KEY_VAULT_NAME \
  --name "${CREDENTIAL_NAME}-privateKey" \
  --query value -o tsv

# Expected output: Multiple lines like this
# -----BEGIN ENCRYPTED PRIVATE KEY-----
# MIIFHDBOBgkqhkiG9w0BBQ0wQTApBgkqhkiG9w0BBQwwHAQI...
# (multiple lines of base64 encoded content)
# ...
# -----END ENCRYPTED PRIVATE KEY-----
```

Correct: Multiple lines with line breaks preserved

Incorrect: Single long line or text with spaces instead of newlines

Secret 2: Private Key Passphrase

The passphrase **can be created using either Azure Portal or Azure CLI** (it's a single-line value).

text

```
# Store the passphrase (if your key is encrypted)
az keyvault secret set \
 --vault-name $KEY_VAULT_NAME \
 --name "${CREDENTIAL_NAME}-password" \
 --value "your_passphrase_here"
```

Important: If you used an **unencrypted** private key, you still need to create the .password secret, but set it to an empty string or a placeholder value.

Using Azure Portal:

1. Go to **Secrets** → **Generate/Import**
2. Name: snowflake-purview-cred.password
3. Value: Your passphrase (or empty string for unencrypted keys)
4. Select **Create**

Create and run scan

To create and run a new scan, follow these steps:

1. In the classic Microsoft Purview governance portal, navigate to **Sources**.

2. Select the registered Snowflake source.

3. Select + **New scan**.

4. Provide the following details:

   a. **Name**: The name of the scan

   b. **Connect via integration runtime**: Select the Azure autoresolved integration runtime, Managed Virtual Network IR or SHIR according to your scenario. Learn more from Choose the right integration runtime configuration for your scenario. To use Managed Virtual Network IR to connect to Snowflake via private link, follow the steps in Connect to Snowflake via Managed Private Endpoint to set up the private endpoint first.

   c. **Host for connection**: Choose the endpoint used to establish connection to Snowflake during scan. You can choose from the server URL or the other hosts that you configured in Register.

   d. **Credential**: Select the credential to connect to your data source.

      If using **Basic Authentication** while creating a credential.

      - Provide the user name used to connect to Snowflake in the User name input field.
      - Store the user password used to connect to Snowflake in the secret key.

      If using **Key Pair** while creating a credential

      - Provide User name: purview_scanner (your Snowflake user)
      - Password Key Vault Reference:
      - Key Vault Connection: Select your Key Vault
      - Secret Name: snowflake-purview-cred.password
      - Private Key Vault Reference:
      - Key Vault Connection: Select your Key Vault
      - Secret Name: snowflake-purview-cred.privateKey

   e. **Warehouse**: Specify the name of the warehouse instance used to empower scan in uppercase. The default role assigned to the user specified in the credential must have

USAGE rights on this warehouse.

f. **Databases**: Specify one or more database instance names to import in uppercase. Separate the names in the list with a semi-colon (;). For example, `DB1;DB2`. The default role assigned to the user specified in the credential must have adequate rights on the database objects.

> ⓘ **Note**
>
> Classification is not applied to tables when more than one database instance names are specified.

Acceptable database name patterns can be static names or contain wildcard %. For example: `A%;%B;%C%;D`:

- Start with A or
- End with B or
- Contain C or
- Equal D

g. **Schema**: List subset of schemas to import expressed as a semicolon separated list. For example, `schema1;schema2`. All user schemas are imported if that list is empty. All system schemas and objects are ignored by default.

Acceptable schema name patterns using can be static names or contain wildcard %. For example: `A%;%B;%C%;D`:

- Start with A or
- End with B or
- Contain C or
- Equal D

Usage of NOT and special characters aren't acceptable.

h. **Stored procedure details**: Controls the number of details imported from stored procedures:

- Signature (default): The name and parameters of stored procedures.
- Code, signature: The name, parameters, and code of stored procedures.
- Lineage, code, signature: The name, parameters and code of stored procedures, and the data lineage derived from the code.

- None: Stored procedure details aren't included.

> ⓘ **Note**
>
> If you use Self-hosted Integration Runtime for scan, customized setting other than the default Signature is supported since version 5.30.8541.1. The earlier versions always extract the name and parameters of stored procedures.

i. **Maximum memory available** (applicable when using self-hosted integration runtime): Maximum memory (in GB) available on customer's VM to be used by scanning processes. It's dependent on the size of Snowflake source to be scanned.

> ⓘ **Note**
>
> As a rule of thumb, provide 1GB memory for every 1000 tables.

5. Select **Test connection** to validate the settings (available when using Azure Integration Runtime).

6. Select **Continue**.

7. Select a **scan rule set** for classification. You can choose between the system default, existing custom rule sets, or create a new rule set inline. For more information, see Apply classifications automatically in Microsoft Purview.

> ⓘ **Note**
>
> Classification won't be applied to tables or views when the table name, view name, schema name, or database name contain special characters.

> ⓘ **Note**
>
> If you're using self-hosted runtime, you need to upgrade to version 5.26.404.1 or higher to use Snowflake classification. Download the latest Microsoft Integration Runtime.

8. Choose your **scan trigger**. You can set up a schedule or ran the scan once.

9. Review your scan and select **Save and Run**.

View your scans and scan runs

To view existing scans:

1. Go to the Microsoft Purview portal. On the left pane, select **Data map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that has results you want to view. The pane shows you all the previous scan runs, along with the status and metrics for each scan run.
4. Select the run ID to check the scan run details.

Manage your scans

To edit, cancel, or delete a scan:

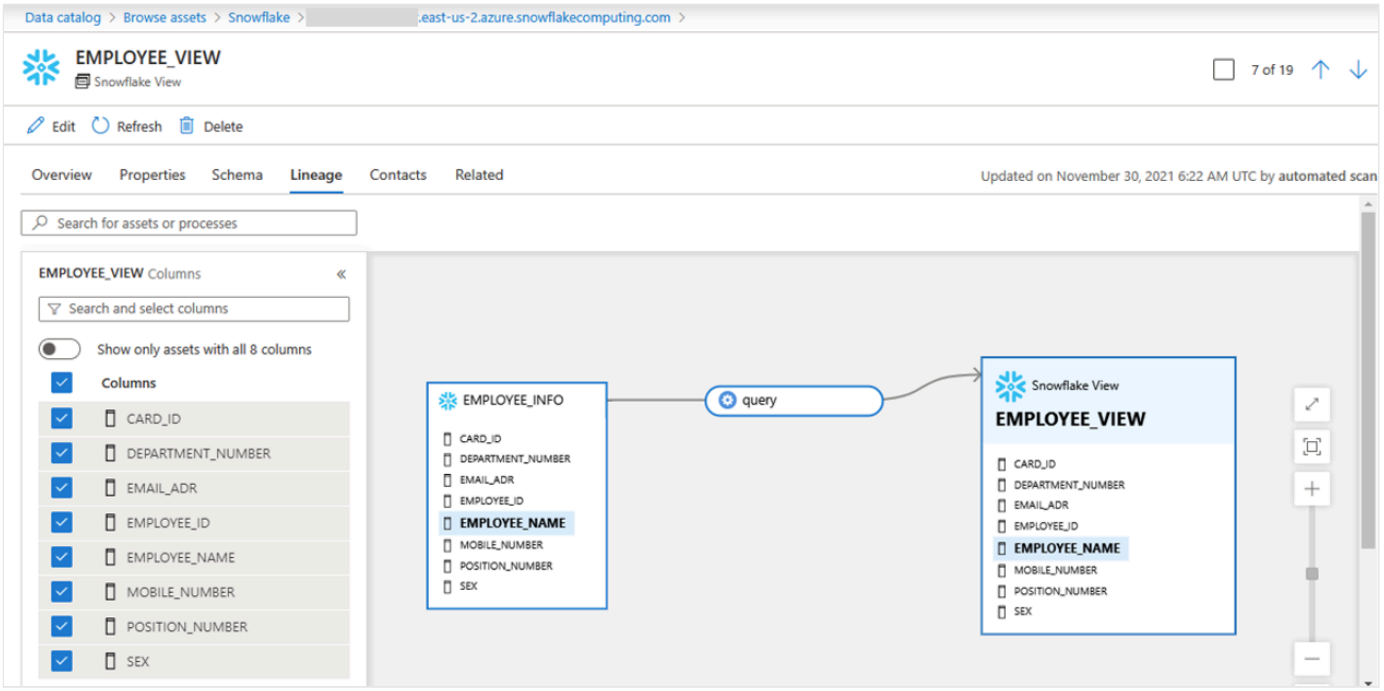
1. Go to the Microsoft Purview portal. On the left pane, select **Data Map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that you want to manage. You can then:
    - Edit the scan by selecting **Edit scan**.
    - Cancel an in-progress scan by selecting **Cancel scan run**.
    - Delete your scan by selecting **Delete scan**.

> ⓘ **Note**
>
> - Deleting your scan does not delete catalog assets created from previous scans.

Lineage

After scanning your Snowflake source, you can browse Unified Catalog or search Unified Catalog to view the asset details.

Go to the asset -> lineage tab, you can see the asset relationship when applicable. Refer to the supported capabilities section on the supported Snowflake lineage scenarios. For more information about lineage in general, see data lineage and lineage user guide.



> **Note**
>
> If a view was created by tables from different databases or different schemas, scan all databases and schemas simultaneously using the names in the semicolon (;) list.

Connect to Snowflake via Managed Private Endpoint

This section describes the required steps to set up a Managed Private Endpoint (PE) from Microsoft Purview to Snowflake on Azure. Learn more about Azure Private Link and Snowflake.

1. In your Snowflake account, retrieve the target endpoints and resource ID. Run the system function SYSTEM$GET_PRIVATELINK_CONFIG() with Account Admin role, and note down the values of the following properties:

    - privatelink-pls-id
    - privatelink-account-url
    - regionless-privatelink-account-url
    - privatelink_ocsp-url

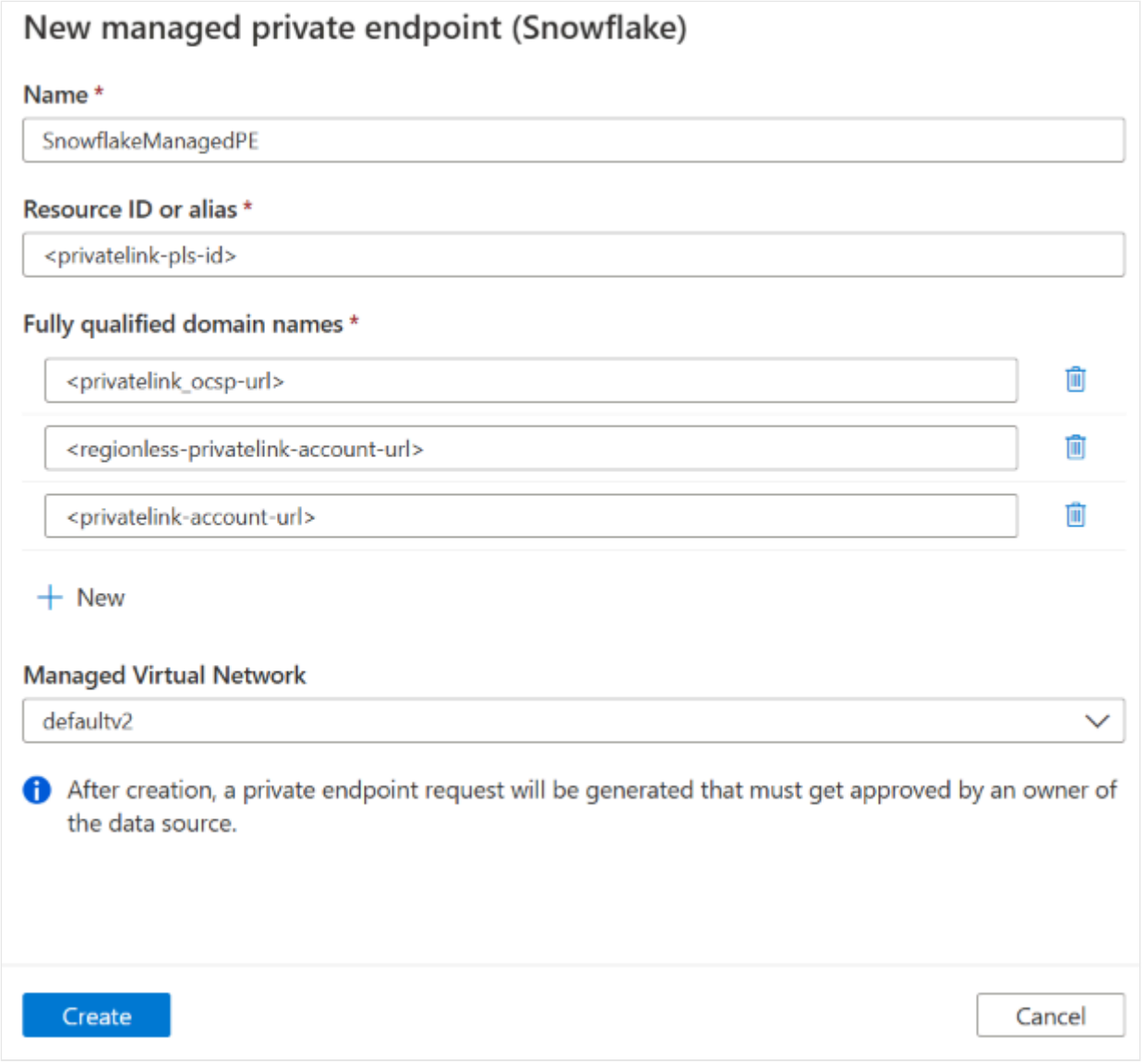
    SQL

    ```
    use role accountadmin;
    select key, value::varchar from
    table(flatten(input=>parse_json(SYSTEM$GET_PRIVATELINK_CONFIG())));
    ```

2. In Microsoft Purview, create a managed private endpoint to Snowflake.

a. Follow the steps to Create a Managed Virtual Network Integration Runtime. If you already have one, make sure it's in version 2, which supports Snowflake PE.
b. Create a managed private endpoint for Snowflake. Navigate to **Managed private endpoint** -> **+ New** -> **Snowflake**, and fill in the following information.

- **Resource ID or alias**: Fill in the value of **privatelink-pls-id** you retrieved.
- **Fully qualified domain names**: Add the **privatelink_ocsp-url**, **regionless-privatelink-account-url** and **privatelink-account-url**.



3. After creating the PE, you can see its provisioning state as **Succeeded** and approval state as **Pending**. Open the PE and find the **Managed private endpoint resource ID** in details page.

4. Contact Snowflake Technical Support and provide the resource ID to let Snowflake approve this PE.

5. Once Snowflake Support confirms the PE is approved, you should see your managed private endpoint approval state as **Approved** in Microsoft Purview.

6. Register the data source and set up a scan. When setting up scan, choose the Managed Virtual Network IR associated with the Snowflake PE.

Troubleshooting tips

If you encounter issues while registering or scanning Snowflake, review the following tips.

- Check your account identifier in the source registration step. Don't include `https://` at the beginning.
- Make sure the warehouse name and database name are in uppercase on the scan setup page.
- Check your key vault. Make sure there are no typos in the password.
- Check the credential you set up in Microsoft Purview. The user you specify must have a default role with the necessary access rights to both the warehouse and the database you're trying to scan. See Required permissions for scan. USE `DESCRIBE USER;` to verify the default role of the user you specified for Microsoft Purview.
- Use Query History in Snowflake to see if any activity is coming across.
  - If there's a problem with the account identifier or password, you won't see any activity.
  - If there's a problem with the default role, you should at least see a `USE WAREHOUSE . . .` statement.
  - You can use the QUERY_HISTORY_BY_USER table function to identify what role is being used by the connection. Setting up a dedicated Microsoft Purview user makes troubleshooting easier.

Troubleshooting for Key Pair authentication

This section covers common connection issues when using Snowflake key pair authentication in Microsoft Purview.

Issue 1: "Test Connection Failed"

Symptoms: Connection test fails in Purview

Possible causes and solutions:

1. **Public Key Not Assigned**

   a. Verify in Snowflake: DESC USER purview_scanner;

   b. Check RSA_PUBLIC_KEY_FP is populated

2. **Key Vault Access Issues**

    a. Verify Managed Identity has Get and List permissions

    b. Check Key Vault firewall settings

3. **Incorrect Secret Names**

    a. Verify: {credentialName}.password exists

    b. Verify: {credentialName}.privateKey exists

Issue 2: "Invalid Private Key Format" or "Connection Failed with Key Error"

Symptoms:

- Authentication fails with "invalid private key format" error
- Connection test fails immediately
- Error messages about malformed PEM or key parsing errors

Common cause: Private key was stored via Azure Portal (corrupted multi-line format)

Solution:

1. Verify if the key format is corrupted:

```text
# Retrieve the secret and check format
az keyvault secret show \
 --vault-name your-keyvault \
 --name "snowflake-purview-cred-privateKey" \
 --query value -o tsv | head -5

# Should show multiple lines starting with:
# -----BEGIN ENCRYPTED PRIVATE KEY-----
# MIIFHDBOBgkqhkiG9w0BBQ0wQTApBgkqhkiG9w...
# (NOT a single long line)
```

2. If corrupted, DELETE and recreate using CLI:

```text
# Delete the corrupted secret az keyvault secret delete \
 --vault-name your-keyvault \
```

```
 --name "snowflake-purview-cred-privateKey"

# Wait for soft-delete purge (or purge immediately if needed)
az keyvault secret purge \
 --vault-name your-keyvault \
 --name "snowflake-purview-cred-privateKey"

# Recreate using CLI with
 --file parameter az keyvault secret set \
 --vault-name your-keyvault \
 --name "snowflake-purview-cred-privateKey" \
 --file rsa_key.p8
```

3. Additional checks:

   a. Ensure private key is in PKCS#8 PEM format

   b. Include -----BEGIN ENCRYPTED PRIVATE KEY----- and -----END ENCRYPTED PRIVATE KEY----- lines

   c. Verify key locally first:

```
text
openssl rsa -in rsa_key.p8 -check
```

Remember: Always use Azure CLI or Cloud Shell for private key secrets, NEVER the Azure Portal UI.

Issue 3: "Public Key Fingerprint Mismatch"

Symptom: Authentication succeeds but wrong key is being used

Solution:

```
text
# Regenerate public key from private key
openssl rsa -in rsa_key.p8 -pubout -out rsa_key_new.pub

# Verify fingerprint
openssl rsa -pubin -in rsa_key_new.pub -outform DER | openssl dgst sha256 -binary |
openssl enc -base64

# Update in Snowflake
ALTER USER purview_scanner SET RSA_PUBLIC_KEY='';
```

Issue 4: "Passphrase Incorrect"

Symptom: Key decryption fails

Solution:

1. Verify passphrase in Key Vault matches the one used during key generation
2. Test passphrase locally:

```text
openssl rsa -in rsa_key.p8 -check
# Enter passphrase when prompted
```

Issue 5: "Missing Snowflake Permissions"

Symptom: Connection succeeds but scan fails

Solution:

Grant required privileges to the Snowflake user:

```SQL
-- Create role for Purview
CREATE ROLE IF NOT EXISTS purview_reader;

-- Grant usage on warehouse
GRANT USAGE ON WAREHOUSE COMPUTE_WH TO ROLE purview_reader;

-- Grant usage on database
GRANT USAGE ON DATABASE my_database TO ROLE purview_reader;

-- Grant usage on all schemas
GRANT USAGE ON ALL SCHEMAS IN DATABASE my_database TO ROLE  purview_reader;

-- Grant select on all tables
GRANT SELECT ON ALL TABLES IN DATABASE my_database TO ROLE  purview_reader;
GRANT SELECT ON FUTURE TABLES IN DATABASE my_database TO ROLE  purview_reader;

-- Grant select on all views
GRANT SELECT ON ALL VIEWS IN DATABASE my_database TO ROLE  purview_reader;
GRANT SELECT ON FUTURE VIEWS IN DATABASE my_database TO ROLE  purview_reader;

-- Assign role to user
GRANT ROLE purview_reader TO USER purview_scanner;

-- Set as default role
ALTER USER purview_scanner SET DEFAULT_ROLE = purview_reader;
```

Related content

Now that you've registered your source, use the following resources to learn more about Microsoft Purview and your data.

- Data Estate Insights in Microsoft Purview
- Lineage in Microsoft Purview
- Search Unified Catalog

Last updated on 06/01/2026

Connect to and manage Tableau in Microsoft Purview (Preview)

Article • 03/28/2025

This article outlines how to register Tableau, and how to authenticate and interact with Tableau in Microsoft Purview. For more information about Microsoft Purview, read the introductory article.

> **Important**
>
> This feature is currently in preview. The Supplemental Terms of Use for Microsoft Azure Previews include additional legal terms that apply to Azure features that are in beta, in preview, or otherwise not yet released into general availability.

Supported capabilities

Scanning capabilities

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan
Yes	Yes	No	Yes

The supported Tableau Server version is 8.x to 2023.x.

When scanning Tableau, Microsoft Purview supports extracting technical metadata including:

- Server
- Projects
- Workbooks
- Dashboards
- Data sources
- Stories
- Worksheets and the fields/axes
- Workbook data sources
- Tables and the columns

When setting up scan, you can choose to scan an entire Tableau server, or if you have multiple sites installed, to scope the scan to a given site.

Other capabilities

View a table listing the supported capabilities for classifications, sensitivity labels, policies, data lineage, and live view.

Known limitations

- Currently, scan isn't supported if your Tableau server doesn’t allow access from public network.
- When object is deleted from the data source, currently the subsequent scan won't automatically remove the corresponding asset in Microsoft Purview.

Prerequisites

- You must have an Azure account with an active subscription. Create an account for free.
- You must have an active Microsoft Purview account.
- You need an Azure Key Vault, and to grant Microsoft Purview permissions to access secrets.
- You need Data Source Administrator and Data Reader permissions to register a source and manage it in the Microsoft Purview governance portal. For more information about permissions, see Access control in Microsoft Purview.
- Set up the right integration runtime for your scenario. If your data source isn't publicly accessible, set up the latest kubernetes supported self-hosted integration runtime.

Required permissions for scan

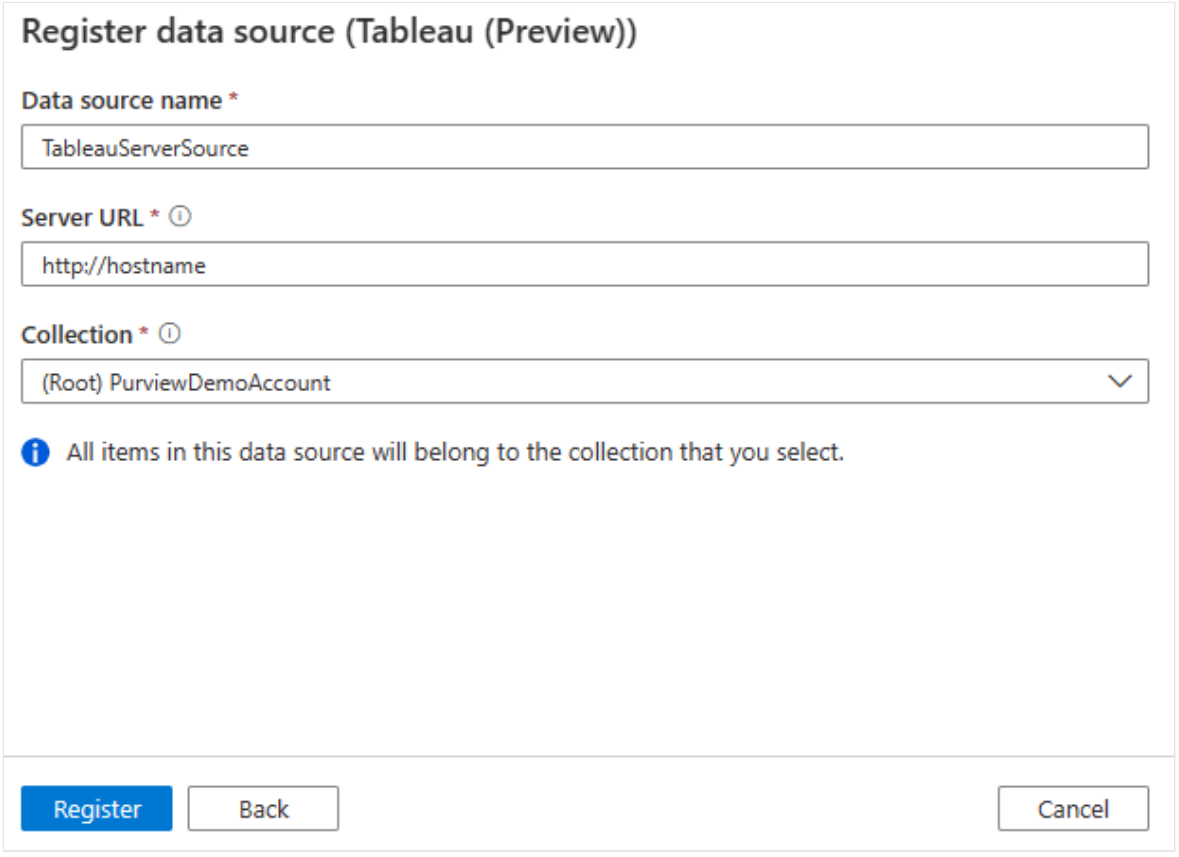
Microsoft Purview supports basic authentication (username and password) for scanning Tableau server.

The user account must be licensed as "Interactor" (renamed to "Explorer" in the latest versions) on the Tableau server, and must have 'View' and 'Download' rights for all of the Workbooks, Data Sources and Projects that are about to import. It is preferable to use an account that has "Site Administrator" security role.

Register

This section describes how to register a Tableau server in Microsoft Purview by using the Microsoft Purview governance portal.

1. Go to your Microsoft Purview account.
2. Select **Data Map** on the left pane.
3. Select **Register**.
4. In **Register sources**, select **Tableau** > **Continue**.
5. On the **Register sources (Tableau)** screen, do the following:
   a. For **Name**, enter a name that Microsoft Purview will list as the data source.
   b. For **Host**, enter the URL of your Tableau server. If SSL is used, the URL must start by https and possibly end with a custom port if the SSL default port of 443 is not used.
   c. Select a collection from the list.



6. Select **Register**.

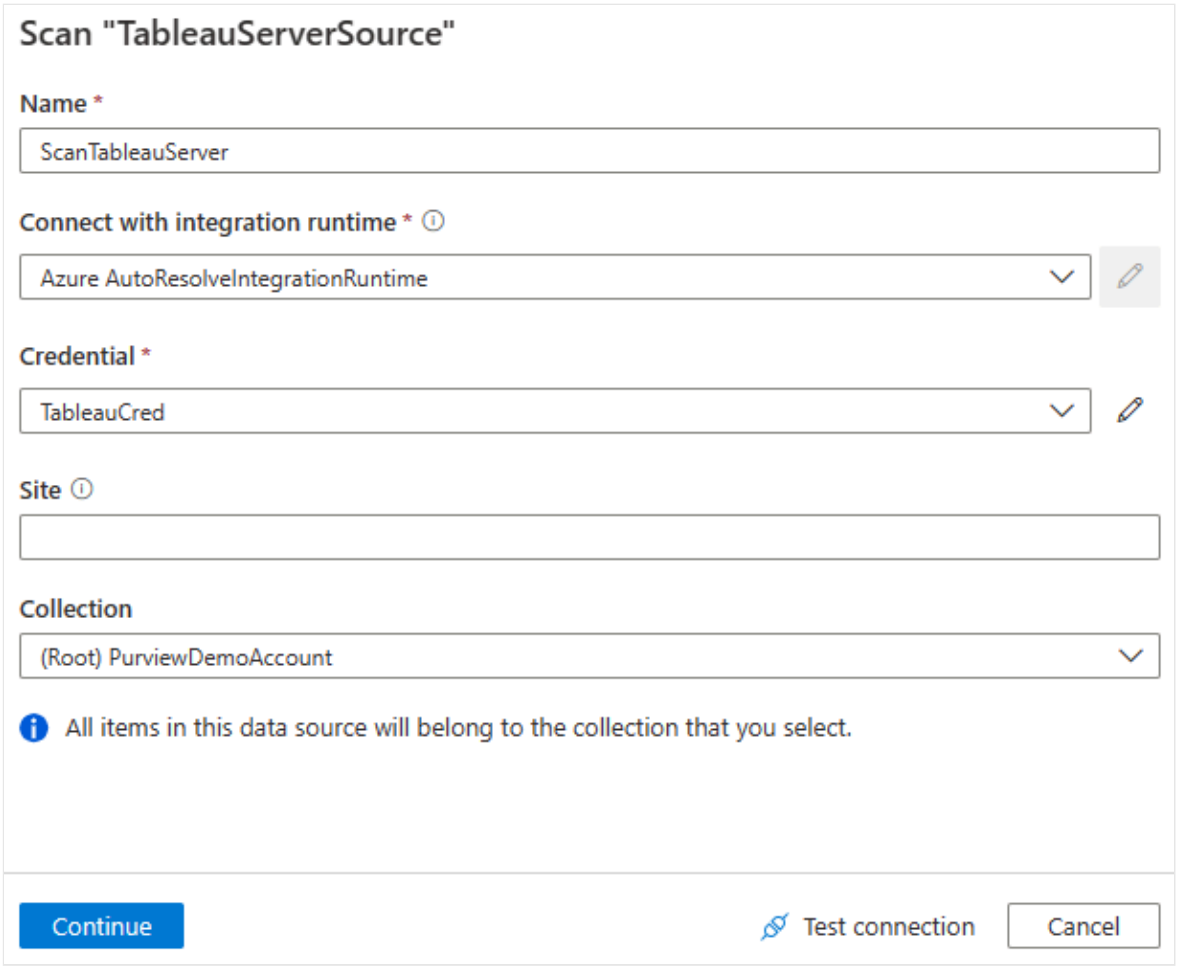
Scan

> **Tip**
>
> To troubleshoot any issues with scanning:
>
> 1. Confirm you have followed all **prerequisites**.
> 2. Review our **scan troubleshooting documentation**.

Use the following steps to scan Tableau server to automatically identify assets. For more information about scanning in general, see Scans and ingestion in Microsoft Purview.

1. Go to **Sources**.
2. Select the registered Tableau source.
3. Select + **New scan**.
4. Provide the following details:
    a. **Name**: Enter a name for the scan.
    b. **Connect via integration runtime**: Choose the default Azure integration runtime, Managed VNet IR, or a Kubernetes supported self-hosted integration runtime you created.
    c. **Credential**: Select the credential to connect to your data source. Make sure to:
        - Select **Basic Authentication** while creating a credential.
        - Provide the user name used to connect to Tableau server in the User name input field.
        - Store the user password used to connect to Tableau server in the secret key.

        For more information, see Credentials for source authentication in Microsoft Purview.
    d. **Site**: Site is only required when connecting to an online Tableau server with multiple sites installed. The value is case sensitive, and must not contain any space as it is used in the URL to connect to Tableau.
5. Click **Test connection** to validate the settings.



6. Select **Continue**.

7. For **Scan trigger**, choose whether to set up a schedule or run the scan once.

8. Review your scan and select **Save and Run**.

View your scans and scan runs

To view existing scans:

1. Go to the Microsoft Purview portal. On the left pane, select **Data map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that has results you want to view. The pane shows you all the previous scan runs, along with the status and metrics for each scan run.
4. Select the run ID to check the scan run details.

Manage your scans

To edit, cancel, or delete a scan:

1. Go to the Microsoft Purview portal. On the left pane, select **Data Map**.

2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.

3. Select the scan that you want to manage. You can then:

    - Edit the scan by selecting **Edit scan**.
    - Cancel an in-progress scan by selecting **Cancel scan run**.
    - Delete your scan by selecting **Delete scan**.

> **Note**
>
> - Deleting your scan does not delete catalog assets created from previous scans.

Browse, search, and view assets

After successfully scanning your Tableau server, you can browse Unified Catalog or search Unified Catalog to view the asset details.

Next steps

Now that you've registered your source, use the following guides to learn more about Microsoft Purview and your data:

- Search the Unified Catalog
- Data Estate Insights in Microsoft Purview

Feedback

Was this page helpful? Yes No

Connect to and manage Teradata in Microsoft Purview

Article • 03/28/2025

This article outlines how to register Teradata, and how to authenticate and interact with Teradata in Microsoft Purview. For more information about Microsoft Purview, read the introductory article.

Supported capabilities

Scanning capabilities

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan
Yes	Yes	No	Yes

The supported Teradata database versions are 12.x to 17.x.

When scanning Teradata source, Microsoft Purview supports:

- Extracting technical metadata including:
  - Server
  - Databases
  - Tables including the columns, foreign keys, indexes, and constraints
  - Views including the columns
  - Stored procedures including the parameter dataset and result set
  - Functions including the parameter dataset
- Fetching static lineage on assets relationships among tables and views.

When setting up scan, you can choose to scan an entire Teradata server, or scope the scan to a subset of databases matching the given name(s) or name pattern(s).

Other capabilities

View a table listing the supported capabilities for classifications, sensitivity labels, policies, data lineage, and live view.

Known limitations

When object is deleted from the data source, currently the subsequent scan won't automatically remove the corresponding asset in Microsoft Purview.

Required permissions for scan

Microsoft Purview supports basic authentication (username and password) for scanning Teradata. The user should have SELECT permission granted for every individual system table listed below:

SQL

```
grant select on dbc.tvm to [user];
grant select on dbc.dbase to [user];
grant select on dbc.tvfields to [user];
grant select on dbc.udtinfo to [user];
grant select on dbc.idcol to [user];
grant select on dbc.udfinfo to [user];
```

To retrieve data types of view columns, Microsoft Purview issues a prepare statement for `select * from <view>` for each of the view queries and parse the metadata that contains the data type details for better performance. It requires the SELECT data permission on views. If the permission is missing, view column data types will be skipped.

For classification, user also needs to have read permission on the tables/views to retrieve sample data.

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- An active Microsoft Purview account.
- You need Data Source Administrator and Data Reader permissions to register a source and manage it in the Microsoft Purview governance portal. For more information about permissions, see Access control in Microsoft Purview.
- Set up the latest self-hosted integration runtime. For more information, see the create and configure a self-hosted integration runtime guide.
    - Ensure JDK 11 is installed on the machine where the self-hosted integration runtime is installed. Restart the machine after you newly install the JDK for it to take effect.

- Ensure Visual C++ Redistributable (version Visual Studio 2012 Update 4 or newer) is installed on the self-hosted integration runtime machine. If you don't have this update installed, you can download it here.
- Download the Teradata JDBC driver on the machine where your self-hosted integration runtime is running. Note down the folder path which you will use to set up the scan.

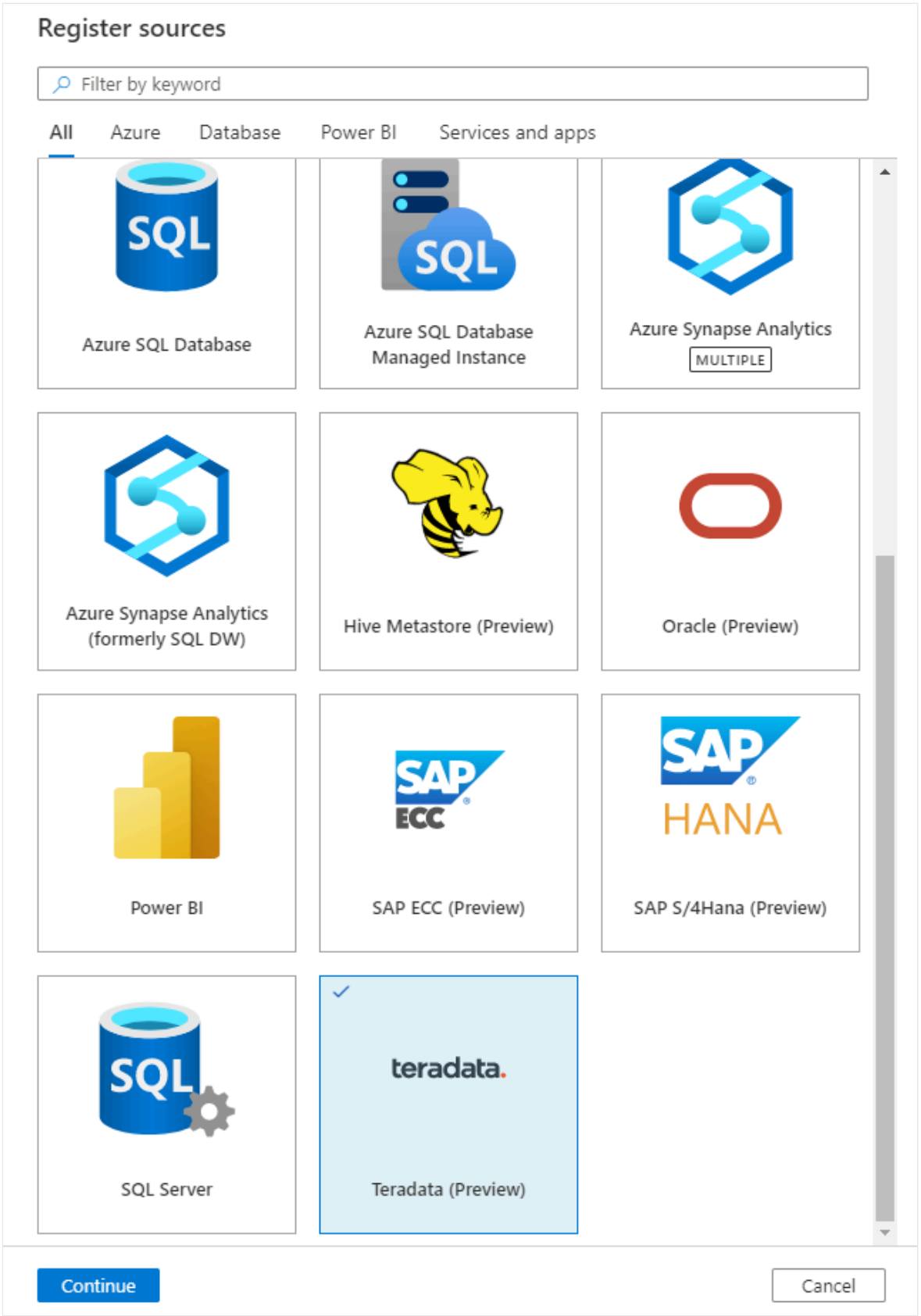
  > **Note**
  >
  > The driver should be accessible by the self-hosted integration runtime. By default, self-hosted integration runtime uses local service account "NT SERVICE\DIAHostService". Make sure it has "Read and execute" and "List folder contents" permission to the driver folder.

Register

This section describes how to register Teradata in Microsoft Purview using the Microsoft Purview governance portal.

Steps to register

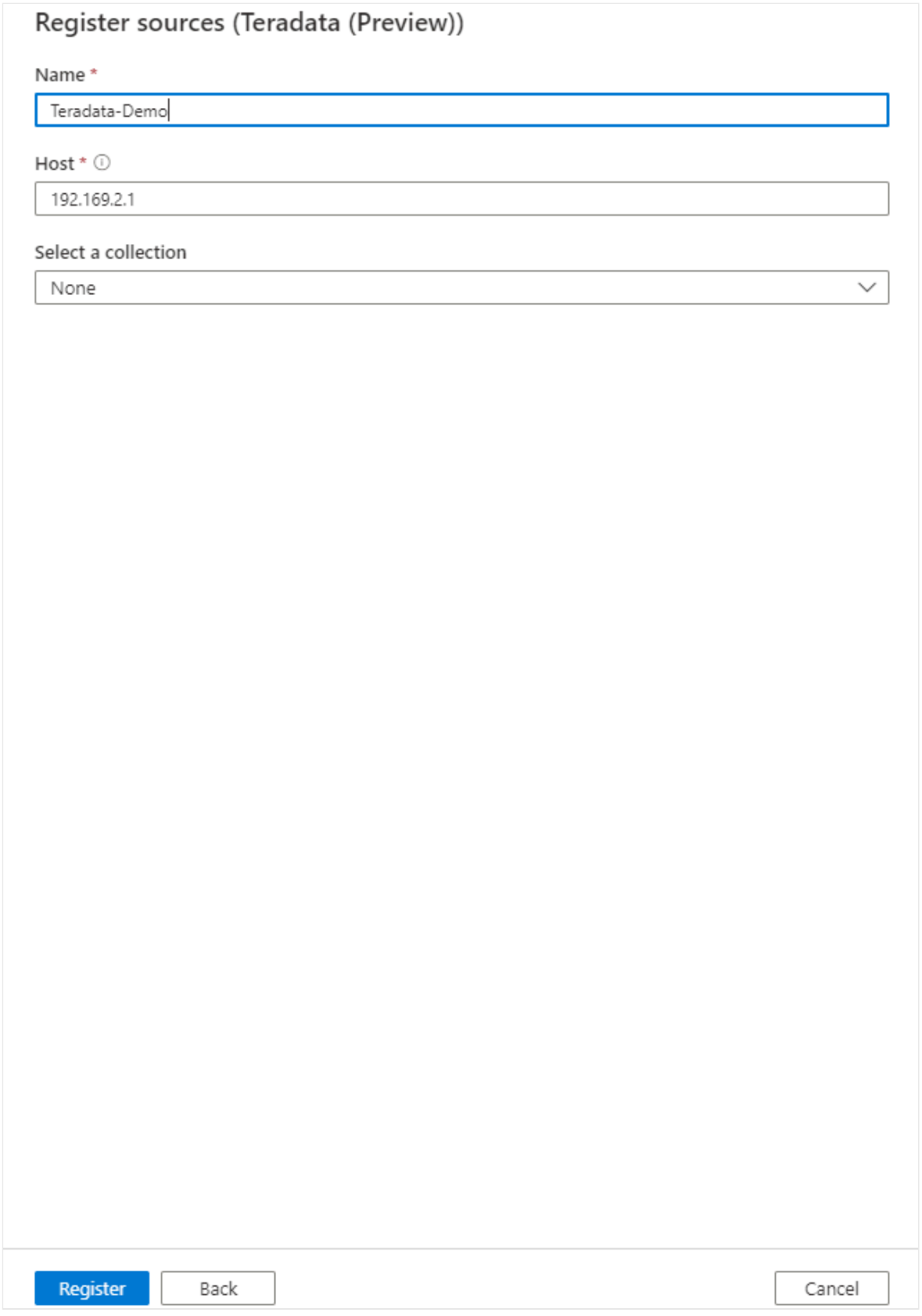
1. Open the Microsoft Purview governance portal by:
   - Browsing directly to https://web.purview.azure.com and selecting your Microsoft Purview account.
   - Opening the Azure portal, searching for and selecting the Microsoft Purview account. Selecting the **the Microsoft Purview governance portal** button.
2. Select **Data Map** on the left navigation.
3. Select **Register**
4. On Register sources, select **Teradata**. Select **Continue**



On the **Register sources (Teradata)** screen, do the following:

1. Enter a **Name** that the data source will be listed with in the Catalog.
2. Enter the **Host** name to connect to a Teradata source. It can also be an IP address of the server.
3. Select a collection from the list.

4. Finish to register the data source.



Scan

Follow the steps below to scan Teradata to automatically identify assets. For more information about scanning in general, see our introduction to scans and ingestion.

Create and run scan

1. In the Management Center, select **Integration runtimes**. Make sure a self-hosted integration runtime is set up. If it isn't set up, use the steps mentioned here to set up a self-hosted integration runtime

2. Select the **Data Map** tab on the left pane in the Microsoft Purview governance portal.

3. Select the registered Teradata source.

4. Select **New scan**

5. Provide the below details:

   a. **Name**: The name of the scan

   b. **Connect via integration runtime**: Select the configured self-hosted integration runtime.

   c. **Credential**: Select the credential to connect to your data source. Make sure to:

      - Select Basic Authentication while creating a credential.
      - Provide a user name to connect to database server in the User name input field
      - Store the database server password in the secret key.

      To understand more on credentials, refer to the link here

   d. **Schema**: List subset of databases to import expressed as a semicolon separated list. For Example: `schema1;schema2`. All user databases are imported if that list is empty. All system databases (for example, SysAdmin) and objects are ignored by default.

      Acceptable database name patterns can be static names or contain wildcard %. For example: `A%;%B;%C%;D`

      - Start with A or
      - End with B or
      - Contain C or
      - Equal D

      Usage of NOT and special characters aren't acceptable

   e. **Driver location**: Specify the path to the JDBC driver location in your machine where self-host integration runtime is running. For example:

`D:\Drivers\Teradata`.

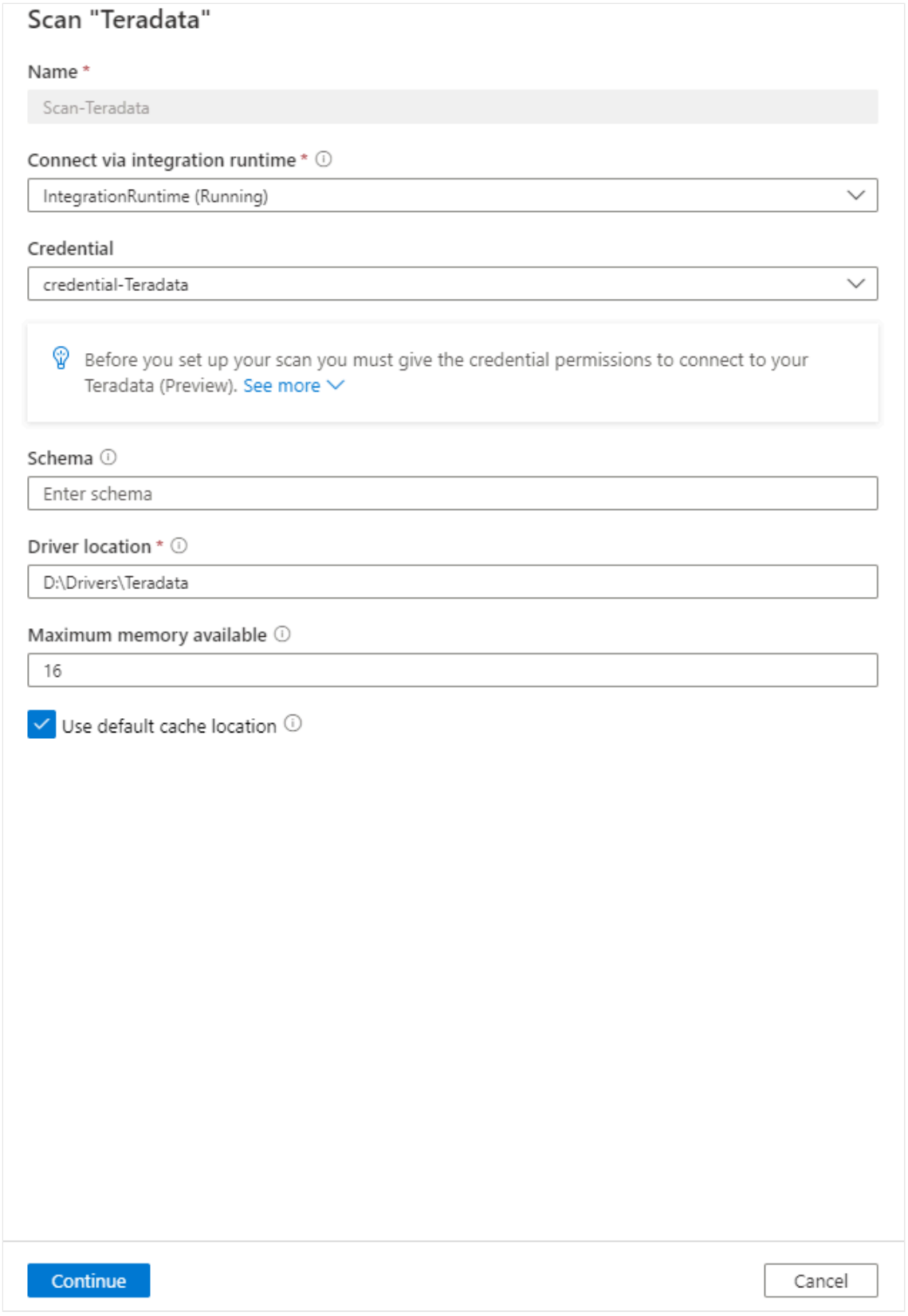
i. For self-hosted integration runtime on a local machine: `D:\Drivers\Teradata`. It's the path to valid JAR folder location. The value must be a valid absolute file path and doesn't contain space. Make sure the driver is accessible by the self-hosted integration runtime;; learn more from prerequisites section.

f. **Stored procedure details**: Controls the number of details imported from stored procedures:

- Signature: The name and parameters of stored procedures.
- Code, signature: The name, parameters and code of stored procedures.
- Lineage, code, signature: The name, parameters and code of stored procedures, and the data lineage derived from the code.
- None: Stored procedure details aren't included.

g. **Maximum memory available:** Maximum memory (in GB) available on customer's VM to be used by scanning processes. This is dependent on the size of Teradata source to be scanned.

> ⓘ **Note**
>
> As a rule of thumb, please provide 2GB memory for every 1000 tables



6. Select **Continue**.

7. Select a **scan rule set** for classification. You can choose between the system default, existing custom rule sets, or create a new rule set inline.

8. Choose your **scan trigger**. You can set up a schedule or ran the scan once.

9. Review your scan and select **Save and Run**.

View your scans and scan runs

To view existing scans:

1. Go to the Microsoft Purview portal. On the left pane, select **Data map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that has results you want to view. The pane shows you all the previous scan runs, along with the status and metrics for each scan run.
4. Select the run ID to check the scan run details.

Manage your scans

To edit, cancel, or delete a scan:

1. Go to the Microsoft Purview portal. On the left pane, select **Data Map**.

2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.

3. Select the scan that you want to manage. You can then:

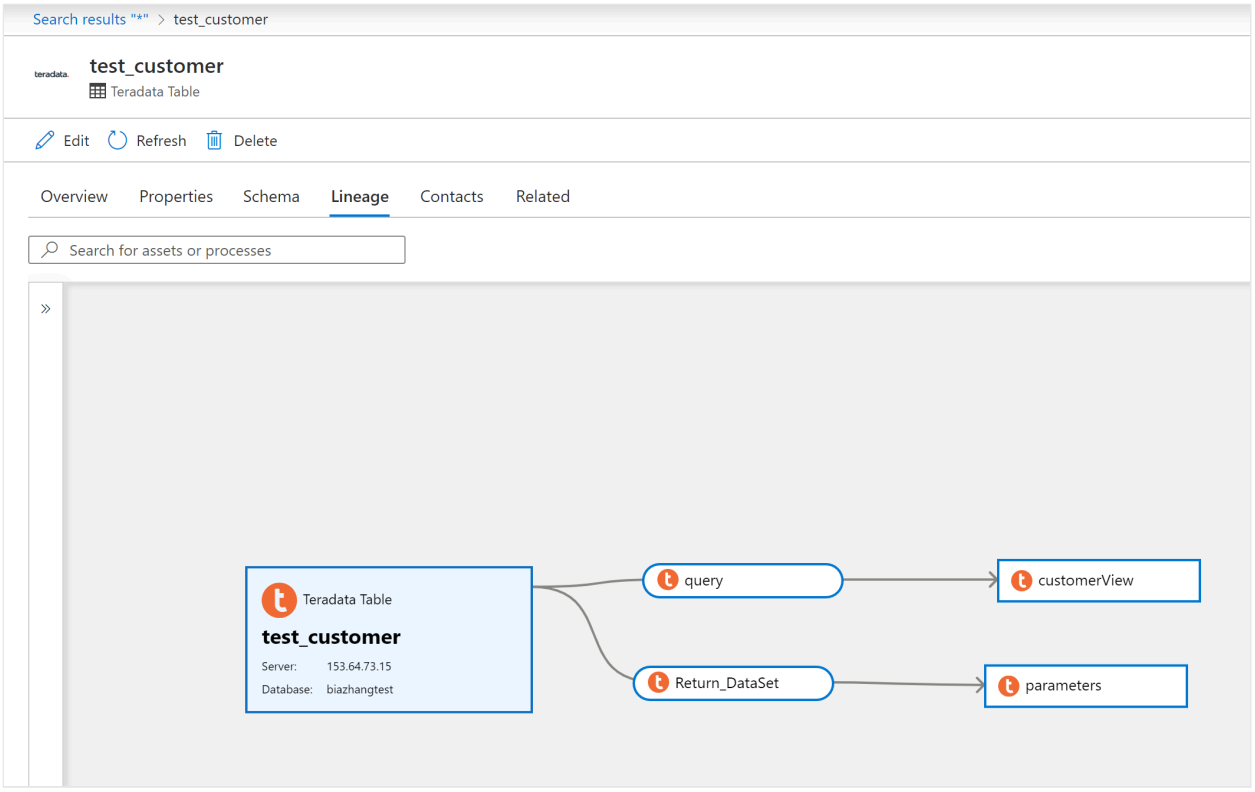
    - Edit the scan by selecting **Edit scan**.
    - Cancel an in-progress scan by selecting **Cancel scan run**.
    - Delete your scan by selecting **Delete scan**.

> ⓘ **Note**
>
> - Deleting your scan does not delete catalog assets created from previous scans.

Lineage

After scanning your Teradata source, you can browse Unified Catalog or search Unified Catalog to view the asset details.

Go to the asset -> lineage tab, you can see the asset relationship when applicable. Refer to the supported capabilities section on the supported Teradata lineage scenarios. For more information about lineage in general, see data lineage and lineage user guide.



Next steps

Now that you've registered your source, follow the below guides to learn more about Microsoft Purview and your data.

- Data Estate Insights in Microsoft Purview
- Lineage in Microsoft Purview
- Search Unified Catalog

Feedback

Was this page helpful? Yes No

Amazon RDS Multicloud Scanning Connector for Microsoft Purview (preview)

The Multicloud Scanning Connector for Microsoft Purview allows you to explore your organizational data across cloud providers, including Amazon Web Services, in addition to Azure storage services.

> **Important**
>
> This feature is currently in preview. The Supplemental Terms of Use for Microsoft Azure Previews include additional legal terms that apply to Azure features that are in beta, in preview, or otherwise not yet released into general availability.

This article describes how to use Microsoft Purview to scan your structured data currently stored in Amazon RDS, including both Microsoft SQL and PostgreSQL databases, and discover what types of sensitive information exist in your data. You'll also learn how to identify the Amazon RDS databases where the data is currently stored for easy information protection and data compliance.

> **Important**
>
> The Multicloud Scanning Connectors for Microsoft Purview are separate add-ons to Microsoft Purview. The terms and conditions for the Multicloud Scanning Connectors for Microsoft Purview are contained in the agreement under which you obtained Microsoft Azure Services. For more information, see Microsoft Azure Legal Information at https://azure.microsoft.com/support/legal/.

Microsoft Purview scope for Amazon RDS

- **Supported database engines**: Amazon RDS structured data storage supports multiple database engines. Microsoft Purview supports Amazon RDS with/based on Microsoft SQL and PostgreSQL.
- **Supported regions**: For private databases that use Kubernetes supported self-hosted integration runtime, there's no region limitation for Amazon RDS databases.

For public databases that use Amazon AutoResolveIntegrationRuntime, Microsoft Purview only supports Amazon RDS databases that are located in the following AWS regions:

- US East (Ohio)
- US East (N. Virginia)
- US West (N. California)
- US West (Oregon)
- Canada (Central)
- Africa (Cape Town)
- Asia Pacific (Hong Kong Special Administrative Region)
- Asia Pacific (Mumbai)
- Asia Pacific (Osaka-Local)
- Asia Pacific (Seoul)
- Asia Pacific (Tokyo)
- Asia Pacific (Singapore)
- Asia Pacific (Sydney)
- Europe (Frankfurt)
- Europe (Ireland)
- Europe (London)
- Europe (Paris)
- Europe (Milan)
- Europe (Stockholm)
- Middle East (Bahrain)
- South America (São Paulo)

Known issues: The following functionality isn't currently supported:

- The **Test connection** button. The scan status messages indicate any errors related to connection setup.
- Selecting specific tables in your database to scan.
- Data lineage.

For more information, see:

- Manage and increase quotas for resources with Microsoft Purview
- Supported data sources and file types in Microsoft Purview
- Use private endpoints for your Microsoft Purview account

Prerequisites

Ensure that you've performed the following prerequisites before adding your Amazon RDS database as Microsoft Purview data sources and scanning your RDS data.

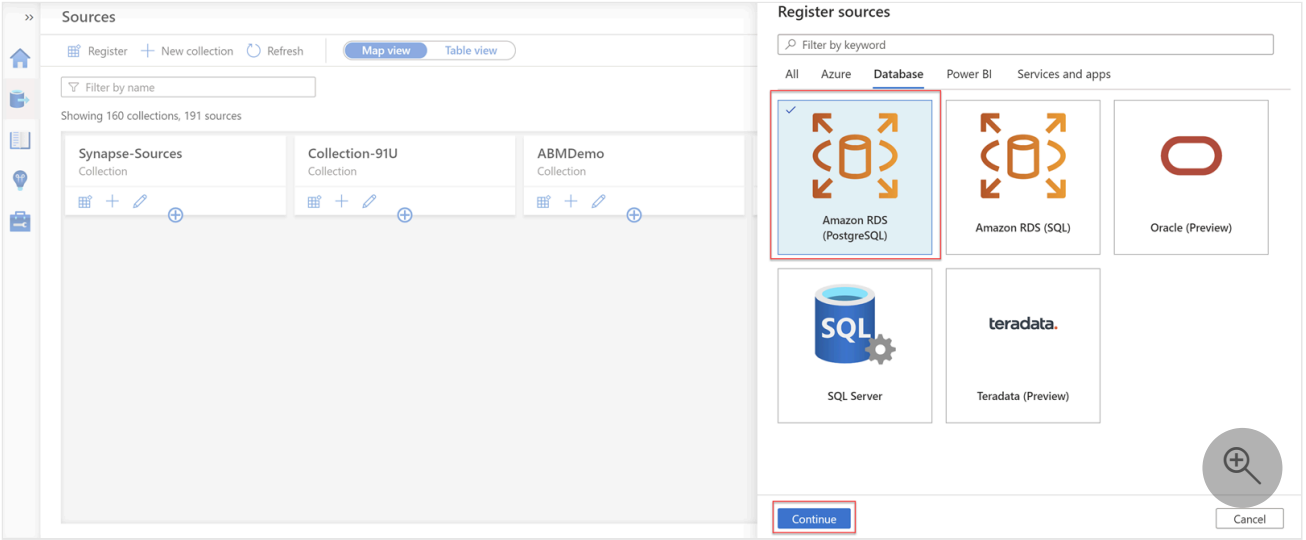
- ✔ You need to be a Microsoft Purview Data Source Admin.
- ✔ You need a Microsoft Purview account. Create a Microsoft Purview account instance, if you don't yet have one.
- ✔ You need an Amazon RDS PostgreSQL or Microsoft SQL database, with data.
- ✔ You need to create a Self-hosted Integration Runtime if you are connecting to a private database. Choose the right integration runtime configuration for your scenario.
  - To use Self-hosted Integration Runtime for scan, set up a Kubernetes supported self-hosted integration runtime.

Register an Amazon RDS data source

To add your Amazon RDS server as a Microsoft Purview data source:

1. In Microsoft Purview, navigate to the **Data Map** page, and select **Register** .

2. On the **Sources** page, select **Register.** On the **Register sources** page that appears on the right, select the **Database** tab, and then select **Amazon RDS (PostgreSQL)** or **Amazon RDS (SQL)**.



3. Enter the details for your source:

Expand table

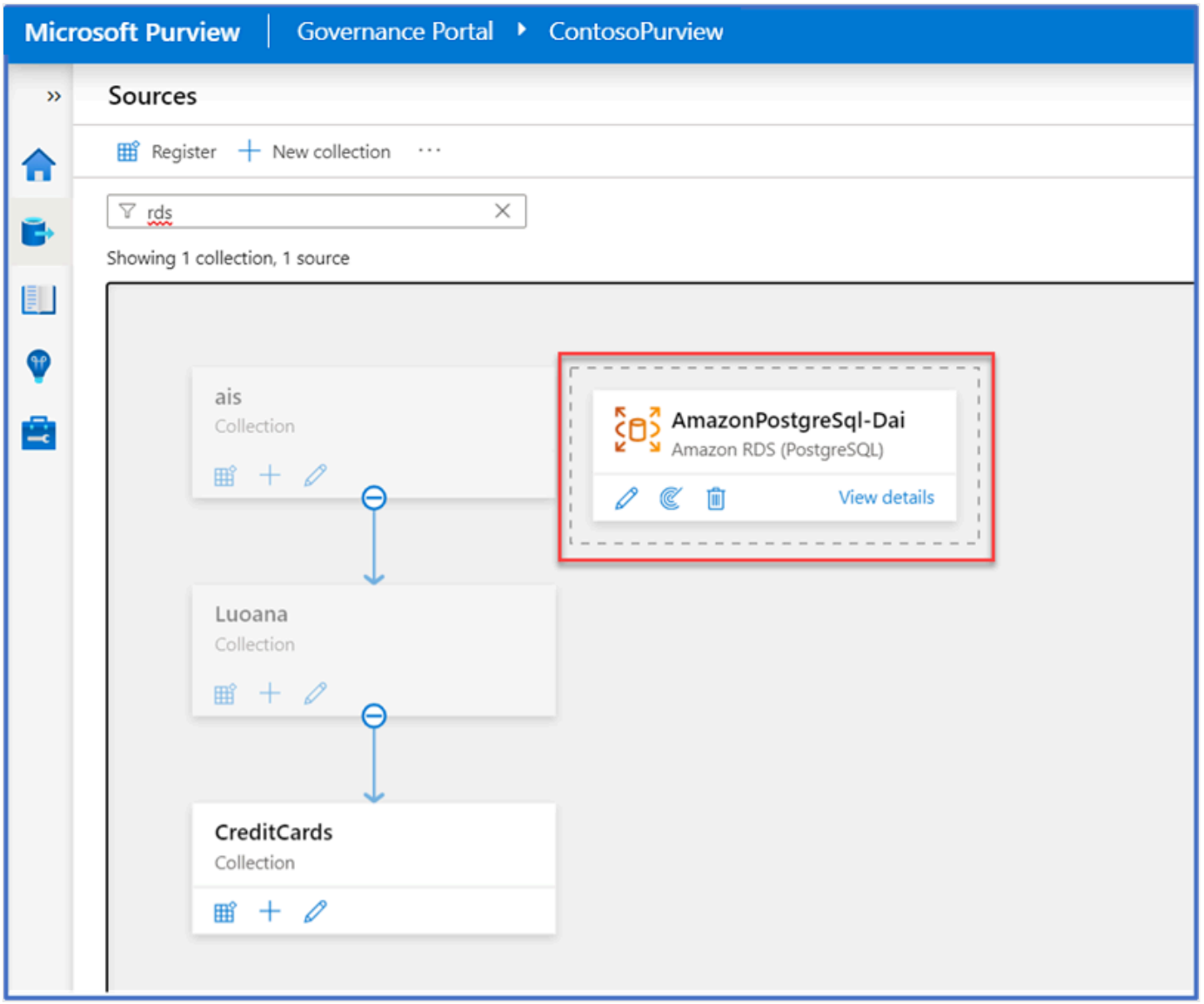
Field	Description
Name	Enter a meaningful name for your source, such as `AmazonPostgreSql-Contoso`

Field	Description
Server name	Enter the name of your RDS database in the following syntax: `<instance identifier>.<xxxxxxxxxxxx>.<region>.rds.amazonaws.com` We recommend that you copy this URL from the Amazon RDS portal, and make sure that the URL includes the AWS region.
Port	Enter the port used to connect to the RDS database: - PostgreSQL: `5432` - Microsoft SQL: `1433`
Collection (optional)	Select a collection to add your data source to. For more information, see Manage data sources in Microsoft Purview (Preview).

4. Select **Register** when you're ready to continue.

> **Note**
>
> You can't register a data source with a name that already exists for a given server name. You must use a different name for the data source.

Your RDS data source appears in the Sources map or list. For example:



Create Microsoft Purview credentials for your RDS scan

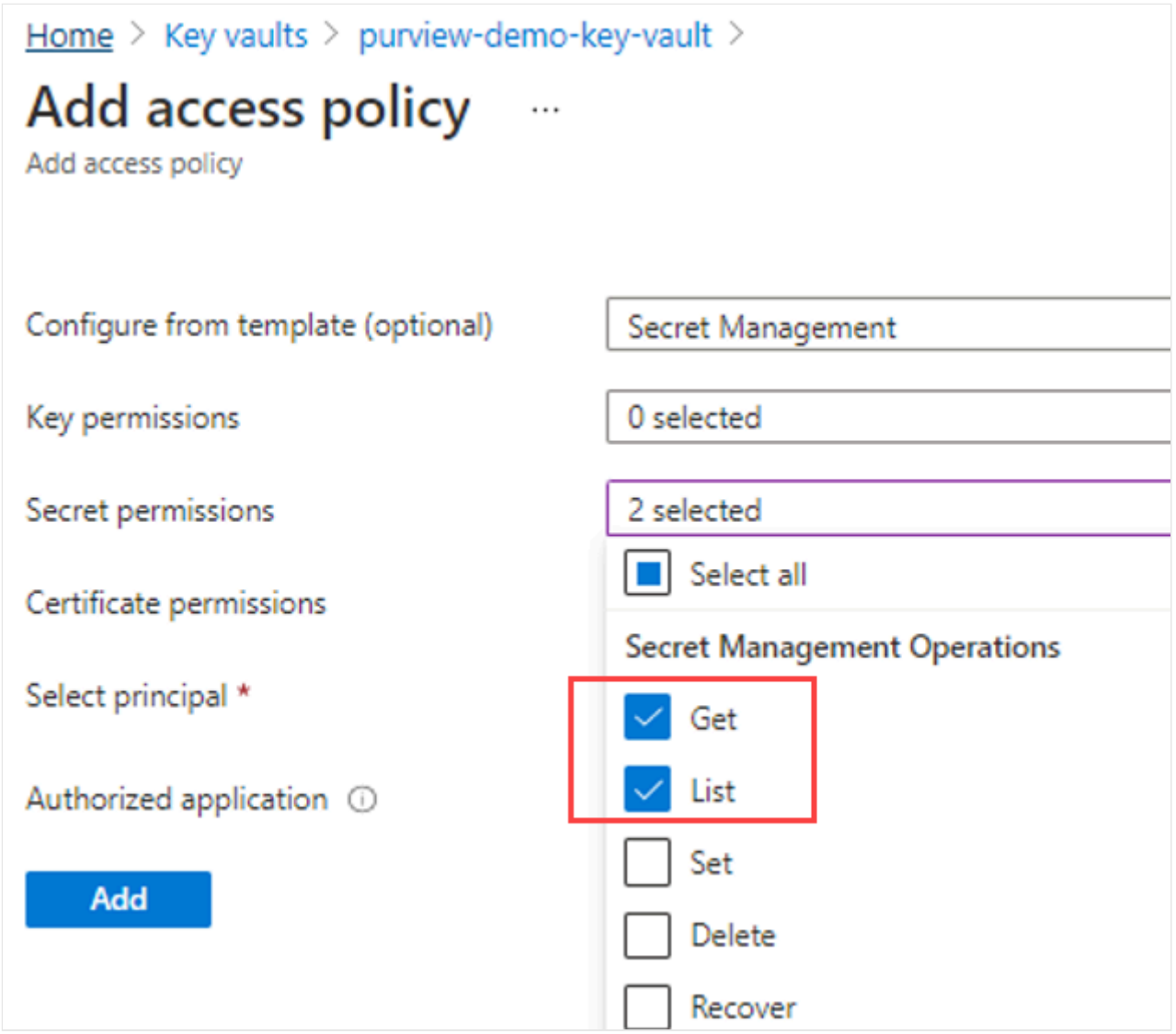
Credentials supported for Amazon RDS data sources include username/password authentication only, with a password stored in an Azure KeyVault secret.

> **Note**
>
> Because Amazon RDS connections in Microsoft Purview require SQL authentication (username and password), we recommend the following security practices:
>
> - Store credentials in Azure Key Vault, as described below.
> - Use strong, unique passwords and rotate them regularly.
> - Restrict network access to your RDS instance using security groups or private connectivity.

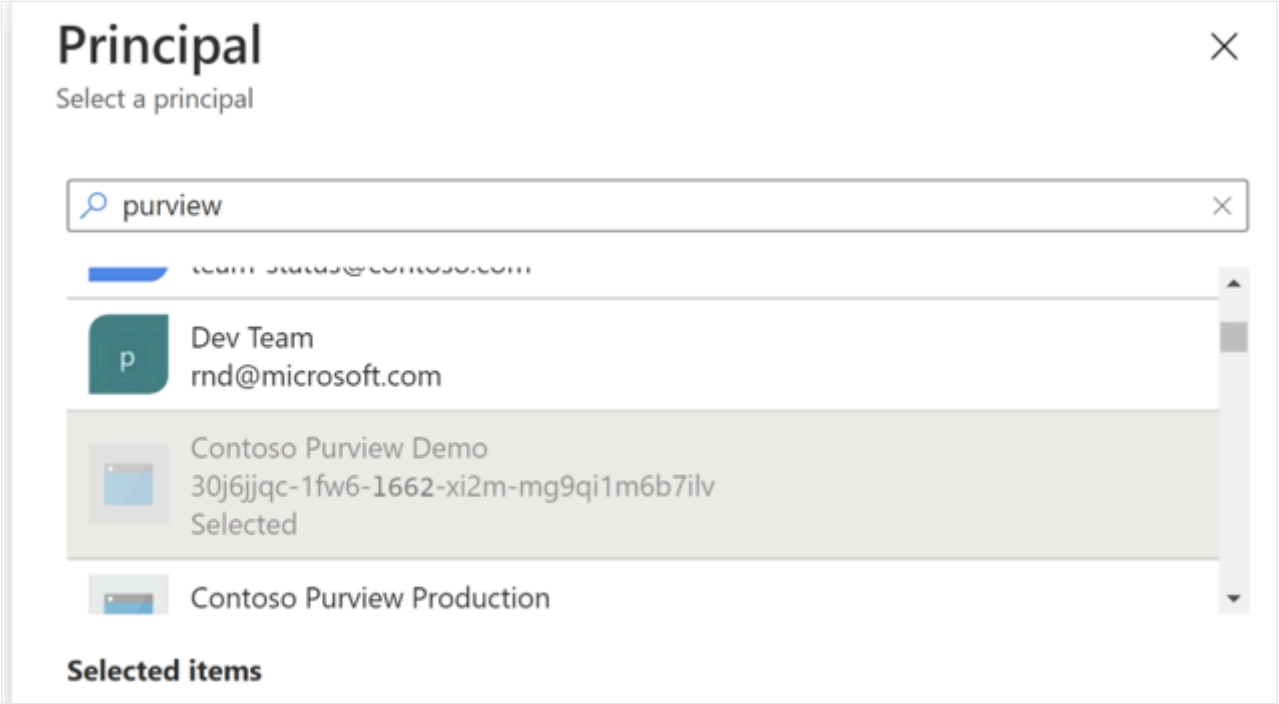
Create a secret for your RDS credentials to use in Microsoft Purview

1. Add your password to an Azure KeyVault as a secret. For more information, see Set and retrieve a secret from Key Vault using Azure portal.

2. Add an access policy to your KeyVault with **Get** and **List** permissions. For example:



When defining the principal for the policy, select your Microsoft Purview account. For example:



Select **Save** to save your Access Policy update. For more information, see Assign an Azure Key Vault access policy.

3. In Microsoft Purview, add a KeyVault connection to connect the KeyVault with your RDS secret to Microsoft Purview. For more information, see Credentials for source authentication in Microsoft Purview.

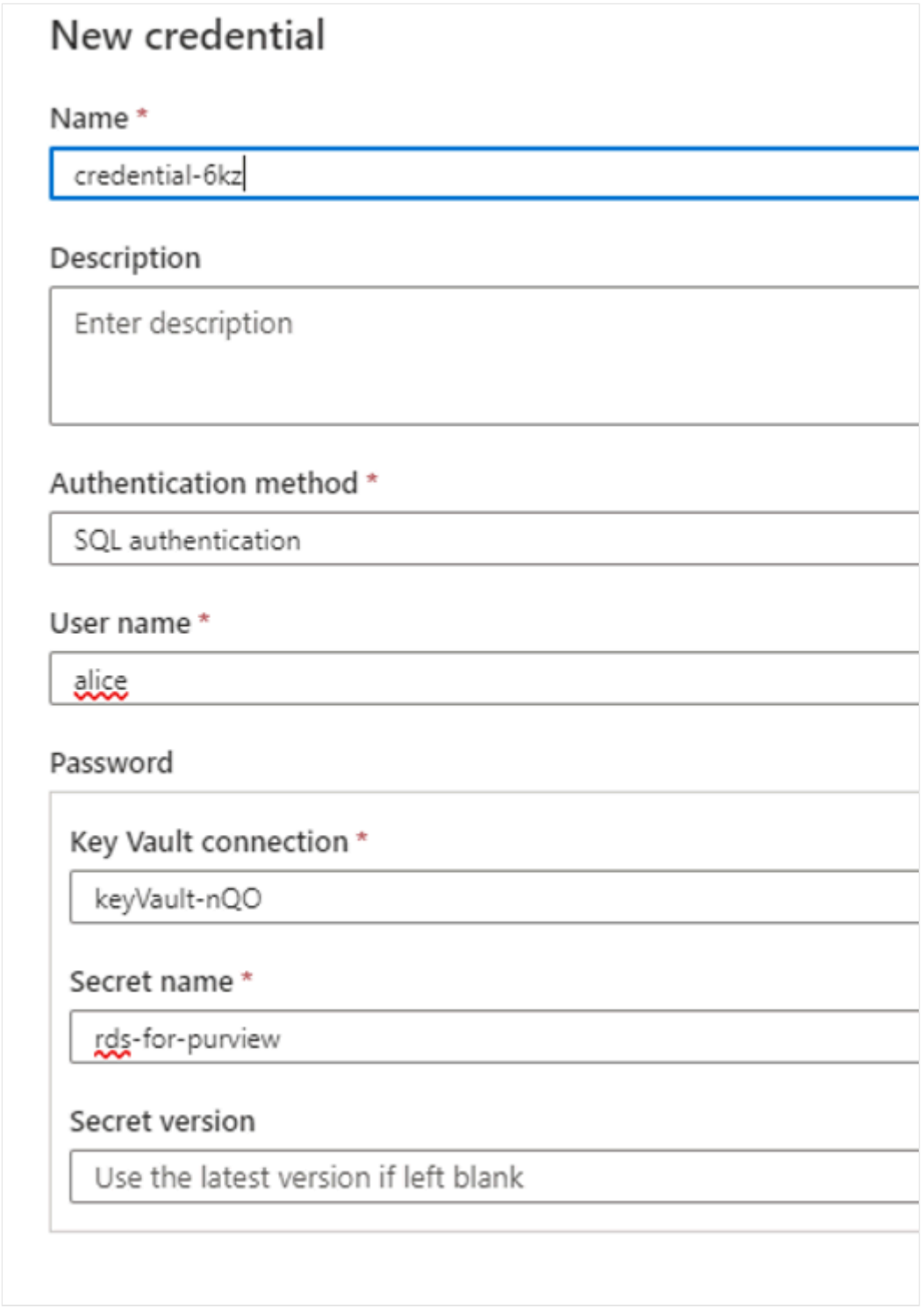
Create your Microsoft Purview credential object for RDS

In Microsoft Purview, create a credentials object to use when scanning your Amazon RDS account.

1. In the Microsoft Purview **Management** area, select **Security and access** > **Credentials** > **New**.

2. Select **SQL authentication** as the authentication method. Then, enter details for the Key Vault where your RDS credentials are stored, including the names of your Key Vault and secret.

   For example:



For more information, see Credentials for source authentication in Microsoft Purview.

Scan an Amazon RDS database

To configure a Microsoft Purview scan for your RDS database:

1. From the Microsoft Purview **Sources** page, select the Amazon RDS data source to scan.
2. Select **New scan** to start defining your scan. In the pane that opens on the right, enter the following details, and then select **Continue**.

- **Name**: Enter a meaningful name for your scan.
- **Connect with integration runtime**: Choose an integration runtime based on your database type.
  - Amazon AutoResolveIntegrationRuntime: Select for a public database.
  - Self-Hosted Integration Runtime: Create new Kubernetes-based SHIR and use for a private database.
- **Database name**: Enter the name of the database you want to scan. You need to find the names available from outside Microsoft Purview, and create a separate scan for each database in the registered RDS server.
- **Credential**: Select the credential you created earlier for the Multicloud Scanning Connectors for Microsoft Purview to access the RDS database.

3. On the **Select a scan rule set** pane, select the scan rule set you want to use, or create a new one. For more information, see Create a scan rule set.

4. On the **Set a scan trigger** pane, select whether you want to run the scan once, or at a recurring time, and then select **Continue**.

5. On the **Review your scan** pane, review the details and then select **Save and Run**, or **Save** to run it later.

While you run your scan, select **Refresh** to monitor the scan progress.

> ⓘ **Note**
>
> When working with Amazon RDS PostgreSQL databases, only full scans are supported. Incremental scans aren't supported as PostgreSQL doesn't have a **Last Modified Time** value.

Explore scanning results

After a Microsoft Purview scan is complete on your Amazon RDS databases, drill down in the Microsoft Purview **Data Map** area to view the scan history. Select a data source to view its details, and then select the **Scans** tab to view any currently running or completed scans.

Use the other areas of Microsoft Purview to find out details about the content in your data estate, including your Amazon RDS databases:

- **Explore RDS data in the catalog**. The Microsoft Purview catalog shows a unified view across all source types, and RDS scanning results are displayed in a similar way to Azure SQL. You

can browse the catalog using filters or browse the assets and navigate through the hierarchy. For more information, see:
    - Tutorial: Browse assets in Microsoft Purview (preview) and view their lineage
    - Search the Microsoft Purview Unified Catalog
    - Register and scan an Azure SQL Database

- **View Insight reports** to view statistics for the classification, sensitivity labels, file types, and more details about your content.

  All Microsoft Purview Insight reports include the Amazon RDS scanning results, along with the rest of the results from your Azure data sources. When relevant, an **Amazon RDS** asset type is added to the report filtering options.

  For more information, see the Understand Data Estate Insights in Microsoft Purview.

- **View RDS data in other Microsoft Purview features**, such as the **Scans** and **Glossary** areas. For more information, see:
    - Create a scan rule set
    - Tutorial: Create and import glossary terms in Microsoft Purview (preview)

RDS errors

The following errors may appear in Microsoft Purview:

- `Unknown database`. In this case, the database defined doesn't exist. Check to see that the configured database name is correct

- `Failed to login to the Sql data source. The given auth credential does not have permission on the target database.` In this case, your username and password is incorrect. Check your credentials and update them as needed.

Legacy AWS RDS Scan Configuration

> **Note**
>
> The method described below will be deprecated soon and should only be used for reference. We strongly recommend using the new AWS RDS Scan configuration method.

Previously, configuring Microsoft Purview to connect to your RDS VPC involved setting up a direct connection using a VPN or other network configuration methods. This approach required

following steps for setup of network peering, security groups, and routing configurations.

1. **Set Up VPN or Direct Connect**: Establish a secure connection between your AWS VPC and the Microsoft Purview scanning infrastructure.
2. **Configure Network Peering**: Create VPC peering connections between your VPC and the VPC used by Microsoft Purview.
3. **Security Group Adjustments**: Modify security groups to allow traffic from Microsoft Purview IP ranges to your RDS instances.
4. **Route Table Updates**: Update route tables to ensure proper routing of traffic between the VPCs.

By transitioning to the new AWS Kubernetes-based Integration Runtime configuration, you can achieve a more secure, reliable, and simpler connection setup, ensuring better integration with Microsoft Purview.

Next steps

Learn more about Microsoft Purview Insight reports:

Understand Data Estate Insights in Microsoft Purview

ⓘ **Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 05/07/2026

Connect to Amazon Redshift in Microsoft Purview

Article • 03/28/2025

Scanning capabilities

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan
Yes	Yes	No	Yes

> **Important**
>
> This feature is currently in preview. The Supplemental Terms of Use for Microsoft Azure Previews include additional legal terms that apply to Azure features that are in beta, in preview, or otherwise not yet released into general availability.

This document introduces the preview for scanning Amazon Redshift in Microsoft Purview.

Supported capabilities

Other capabilities

View a table listing the supported capabilities for classifications, sensitivity labels, policies, data lineage, and live view.

When scanning Amazon Redshift, Microsoft Purview supports extracting technical metadata including:

- Server
- Databases
- Schemas
- Tables including the columns, foreign keys and unique constraints
- Views including the columns
- Stored procedures including the parameter dataset
- Functions including the parameter dataset

When setting up scan, you can choose to scan an entire Amazon Redshift database, or scope the scan to a subset of schemas matching the given name(s) or name pattern(s).

Known limitations

- When an object is deleted from the data source, currently the subsequent scan won't automatically remove the corresponding asset in Microsoft Purview.

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- An active Microsoft Purview account.
- You'll need to be a Data Source Administrator and Data Reader to register a source and manage it in the Microsoft Purview governance portal. See our Microsoft Purview Permissions page for details.
- Set up the right integration runtime for your scenario. If your data source isn't publicly accessible, set up the latest kubernetes supported self-hosted integration runtime.

Register

This section will enable you to register the Amazon Redshift data source for scan and data share in Purview.

Prerequisites for register

- You'll need to be a Data Source Admin and one of the other Purview roles (for example, Data Reader or Data Share Contributor) to register a source and manage it in the Microsoft Purview governance portal. See our Microsoft Purview Permissions page for details.

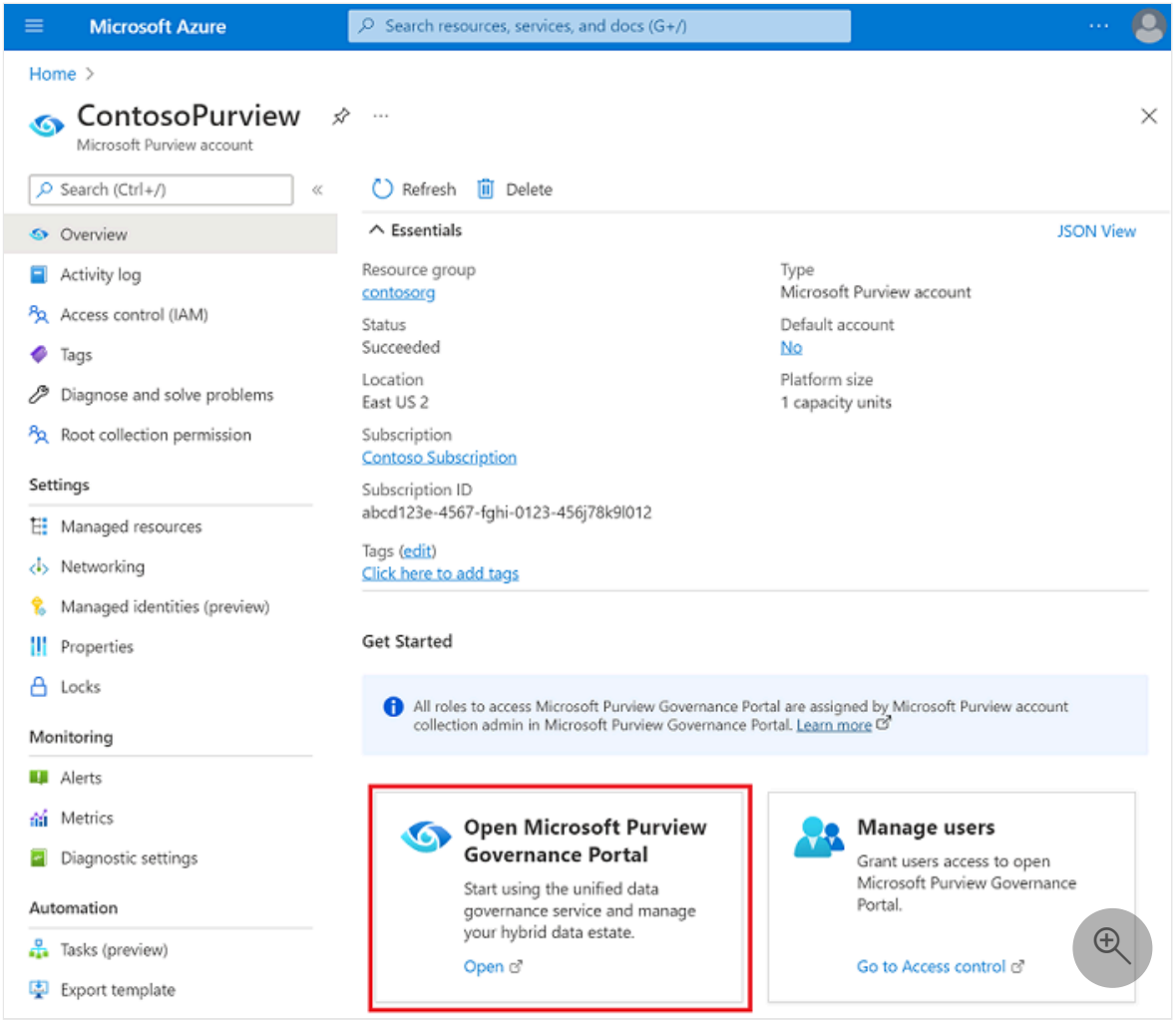
Steps to register

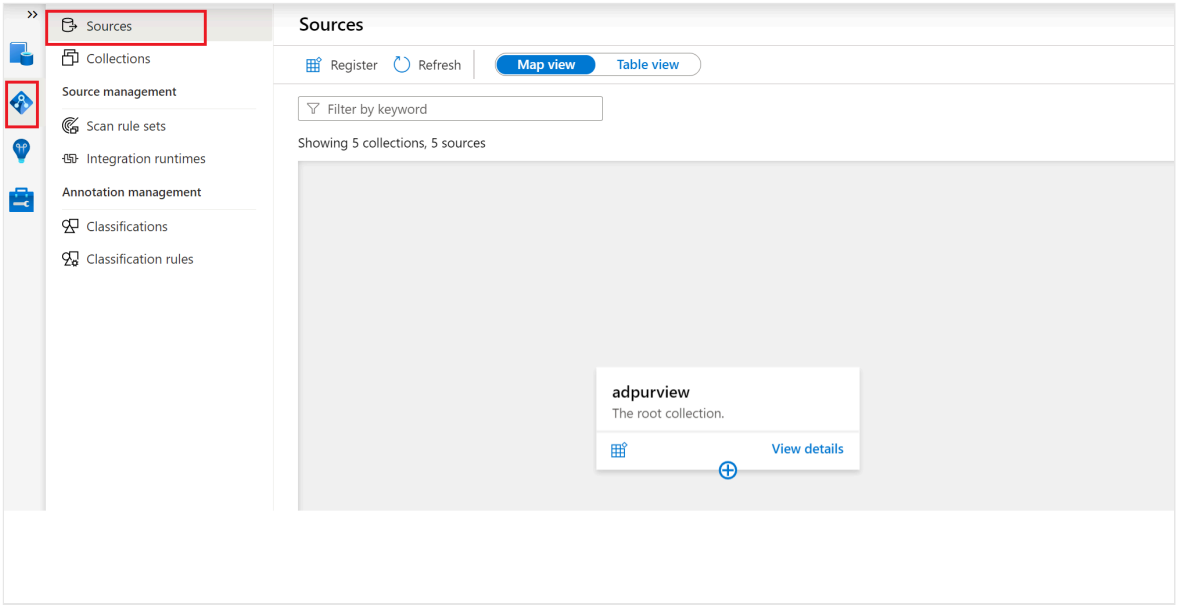
It's important to register the data source in Microsoft Purview prior to setting up a scan for the data source.

1. Go to the Microsoft Purview governance portal by:

- Browsing directly to https://web.purview.azure.com and selecting your Microsoft Purview account.
- Opening the Azure portal, searching for and selecting the Microsoft Purview account. Select the **the Microsoft Purview governance portal** button.

2. Navigate to the **Data Map --> Sources**

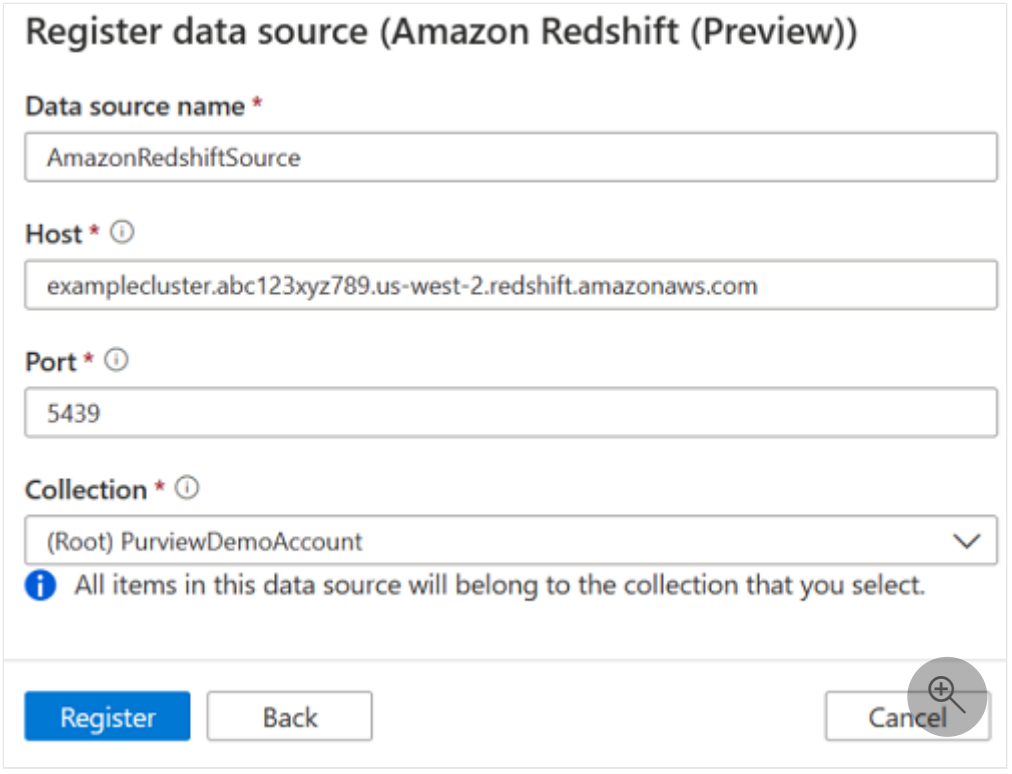




3. Navigate to the appropriate collection under the **Sources** menu and select the **Register** icon to register a new Amazon Redshift data source.

4. Select the **Amazon Redshift** data source and select **Continue**.

5. Provide a suitable **Name** for the data source, and provide these details:
   a. Endpoint - The endpoint of your Amazon Redshift cluster. For example: `examplecluster.abc123xyz789.us-west-2.redshift.amazonaws.com`
   b. Port - The port number that you specified when you launched the cluster. The default is 5439.



6. The Amazon Redshift cluster is shown under the selected Collection

Scan

> **Tip**
>
> To troubleshoot any issues with scanning:
>
> 1. Confirm you have properly set up authentication for scanning.
> 2. Review our scan troubleshooting documentation.

Authentication for a scan

Microsoft Purview supports basic authentication (username and password) for scanning Amazon Redshift.

The user should have **SELECT** permission granted for every individual system table for Microsoft Purview to query metadata from:

- svv_external_tables
- svv_external_columns
- svv_table_info
- information_schema.routines
- information_schema.parameters
- pg_views
- pg_database
- pg_description

And user should have **EXECUTE** permission granted for system function for Microsoft Purview to query metadata from:

- pg_get_late_binding_view_cols

Create a credential

Microsoft Purview uses **Azure Key Vault** to safely store the credentials it uses to authenticate with sources.

- If you haven't set up an Azure Key Vault to store credentials, first create an Azure Key Vault and then follow these steps to connect your Microsoft Purview account to your Azure Key Vault.
- Once you have an Azure Key Vault, store your Amazon Redshift password as a secret in the Key Vault, and then follow these steps to create a credential with this information:
  - Basic authentication credential
  - Add your user name in the input field
  - Add your Key Vault connection and the name of the secret where your password is stored

Create the scan

1. Open your **Microsoft Purview account** and select the **Open Microsoft Purview governance portal**

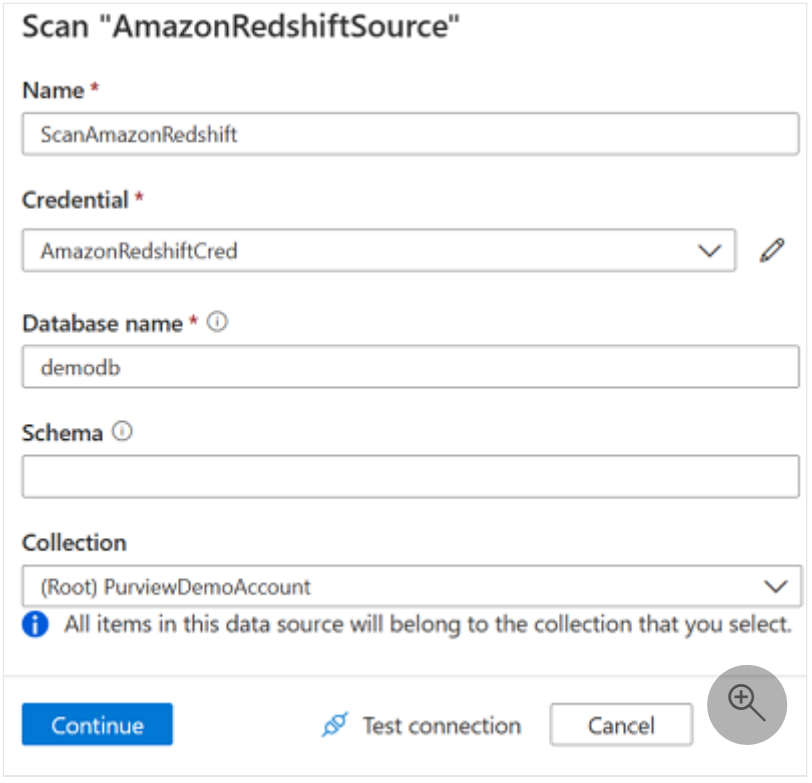
2. Navigate to the **Data map** --> **Sources** to view the collection hierarchy

3. Select the **New Scan** icon under the **Amazon Redshift data source** registered earlier.

4. Select your integration runtime.

5. Provide these details:
   a. Name - a name for your scan
   b. Credential - choose the credential you created earlier
   c. Database - name of the database instance to scan
   d. Schema - The subset of schemas to import expressed as a semicolon separated list of schemas. For example: "schema1;schema2". All user schemas are imported if that list is empty. All system schemas and objects are ignored by default. Acceptable schema name patterns can be static names or contain wildcard %. For example: "A%;%B;%C%;D":

      - Start with A or
      - End with B or
      - Contain C or
      - Equal D



6. Select **Test connection** to validate the settings.

7. Select **Continue**.

8. For **Scan trigger** choose whether to set up a schedule or run the scan once.

9. Review your scan and select **Save and Run**.

View your scans and scan runs

To view existing scans:

1. Go to the Microsoft Purview portal. On the left pane, select **Data map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that has results you want to view. The pane shows you all the previous scan runs, along with the status and metrics for each scan run.
4. Select the run ID to check the scan run details.

Manage your scans

To edit, cancel, or delete a scan:

1. Go to the Microsoft Purview portal. On the left pane, select **Data Map**.
2. Select the data source. You can view a list of existing scans on that data source under **Recent scans**, or you can view all scans on the **Scans** tab.
3. Select the scan that you want to manage. You can then:
    - Edit the scan by selecting **Edit scan**.
    - Cancel an in-progress scan by selecting **Cancel scan run**.
    - Delete your scan by selecting **Delete scan**.

> ⓘ **Note**
>
> - Deleting your scan does not delete catalog assets created from previous scans.

Browse, search, and view assets

After successfully scanning your Amazon Redshift clusters, you can browse the Unified Catalog or search Unified Catalog to view the asset details.

Next steps

Follow the below guides to learn more about Microsoft Purview and your data.

- Browse or Search Microsoft Purview Unified Catalog.

- Data Estate Insights in Microsoft Purview

Feedback

Was this page helpful? Yes No

Amazon S3 Multicloud Scanning Connector for Microsoft Purview

The Multicloud Scanning Connector for Microsoft Purview allows you to explore your organizational data across cloud providers, including Amazon Web Services in addition to Azure storage services.

This article describes how to use Microsoft Purview to scan your unstructured data currently stored in Amazon S3 standard buckets, and discover what types of sensitive information exist in your data. This how-to guide also describes how to identify the Amazon S3 Buckets where the data is currently stored for easy information protection and data compliance.

For this service, use Microsoft Purview to provide a Microsoft account with secure access to AWS, where the Multicloud Scanning Connector for Microsoft Purview will run. The Multicloud Scanning Connector for Microsoft Purview uses this access to your Amazon S3 buckets to read your data, and then reports the scanning results, including only the metadata and classification, back to Azure. Use the Microsoft Purview classification and labeling reports to analyze and review your data scan results.

> **Important**
>
> The Multicloud Scanning Connector for Microsoft Purview is a separate add-on to Microsoft Purview. The terms and conditions for the Multicloud Scanning Connector for Microsoft Purview are contained in the agreement under which you obtained Microsoft Azure Services. For more information, see Microsoft Azure Legal Information at https://azure.microsoft.com/support/legal/.

Supported capabilities

Scanning capabilities

Expand table

Metadata Extraction	Full Scan	Incremental Scan	Scoped Scan
Yes	Yes	Yes	Yes

Other capabilities

For **classifications**, **sensitivity labels**, **policies**, **data lineage**, and **live view**, see the list of supported capabilities.

Known limitations

When scanning Amazon S3 of Glacier storage classes, schema extraction, classification, and sensitivity labels aren't supported.

Microsoft Purview private endpoints aren't supported when scanning Amazon S3.

For more information about Microsoft Purview limits, see:

- Manage and increase quotas for resources with Microsoft Purview
- Supported data sources and file types in Microsoft Purview

Storage and scanning regions

The Microsoft Purview connector for the Amazon S3 service is currently deployed in specific regions only. The following table maps the regions where your data is stored to the region where it would be scanned by Microsoft Purview.

> **Important**
>
> Customers are charged for all related data transfer charges according to the region of their bucket.

Expand table

Storage region	Scanning region
US East (Ohio)	US East (Ohio)
US East (N. Virginia)	US East (N. Virginia)
US West (N. California)	US West (N. California)
US West (Oregon)	US West (Oregon)
Africa (Cape Town)	Europe (Frankfurt)
Asia Pacific (Hong Kong Special Administrative Region)	Asia Pacific (Tokyo)
Asia Pacific (Mumbai)	Asia Pacific (Singapore)
Asia Pacific (Osaka-Local)	Asia Pacific (Tokyo)

Storage region	Scanning region
Asia Pacific (Seoul)	Asia Pacific (Tokyo)
Asia Pacific (Singapore)	Asia Pacific (Singapore)
Asia Pacific (Sydney)	Asia Pacific (Sydney)
Asia Pacific (Tokyo)	Asia Pacific (Tokyo)
Canada (Central)	US East (Ohio)
China (Beijing)	Not supported
China (Ningxia)	Not supported
Europe (Frankfurt)	Europe (Frankfurt)
Europe (Ireland)	Europe (Ireland)
Europe (London)	Europe (London)
Europe (Milan)	Europe (Paris)
Europe (Paris)	Europe (Paris)
Europe (Stockholm)	Europe (Frankfurt)
Middle East (Bahrain)	Europe (Frankfurt)
South America (São Paulo)	US East (Ohio)

Prerequisites

Ensure that you've performed the following prerequisites before adding your Amazon S3 buckets as Microsoft Purview data sources and scanning your S3 data.

- ✓ You need to be a Microsoft Purview Data Source Admin.
- ✓ Create a Microsoft Purview account if you don't yet have one
- ✓ Create a new AWS role for use with Microsoft Purview
- ✓ Create a Microsoft Purview credential for your AWS bucket scan
- ✓ Configure scanning for encrypted Amazon S3 buckets, if relevant
- ✓ Make sure that your bucket policy does not block the connection. For more information, see Bucket policy requirements and SCP policy requirements. For these items, you may need to consult with an AWS expert to ensure that your policies allow required access.
- ✓ When adding your buckets as Microsoft Purview resources, you'll need the values of your AWS ARN, bucket name, and sometimes your AWS account ID.

Create a Microsoft Purview account

- **If you already have a Microsoft Purview account,** you can continue with the configurations required for AWS S3 support. Start with Create a Microsoft Purview credential for your AWS bucket scan.
- **If you need to create a Microsoft Purview account,** follow the instructions in Create a Microsoft Purview account instance. After creating your account, return here to complete configuration and begin using Microsoft Purview connector for Amazon S3.

Create a new AWS role for Microsoft Purview

The Microsoft Purview scanner is deployed in a Microsoft account in AWS. To allow the Microsoft Purview scanner to read your S3 data, you must create a dedicated role in the AWS portal, in the IAM area, to be used by the scanner.

This procedure describes how to create the AWS role, with the required Microsoft Account ID and External ID from Microsoft Purview, and then enter the Role ARN value in Microsoft Purview.

To locate your Microsoft Account ID and External ID:

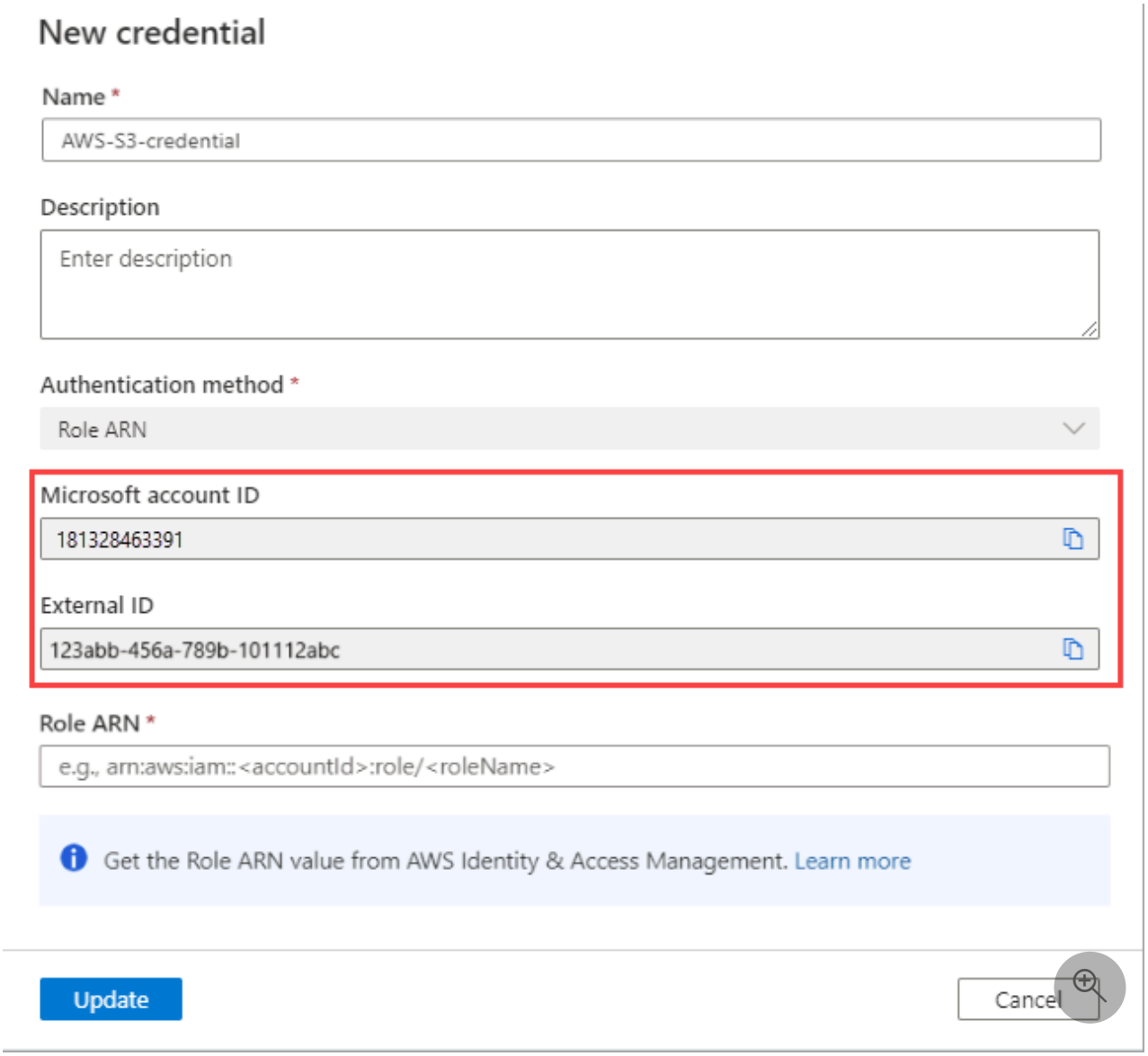
1. To locate credentials:

    - In the Microsoft Purview portal: go to **Data Map** > **Source management** > **Credentials**.
    - In the classic Microsoft Purview governance portal go to the **Management Center** > **Security and access** > **Credentials**.

2. Select **New** to create a new credential.

    In the **New credential** pane that appears, in the **Authentication method** dropdown, select **Role ARN**.

    Then copy the **Microsoft account ID** and **External ID** values that appear to a separate file, or have them handy for pasting into the relevant field in AWS. For example:



To create your AWS role for Microsoft Purview:

1. Open your **Amazon Web Services** console, and under **Security, Identity, and Compliance**, select **IAM**.

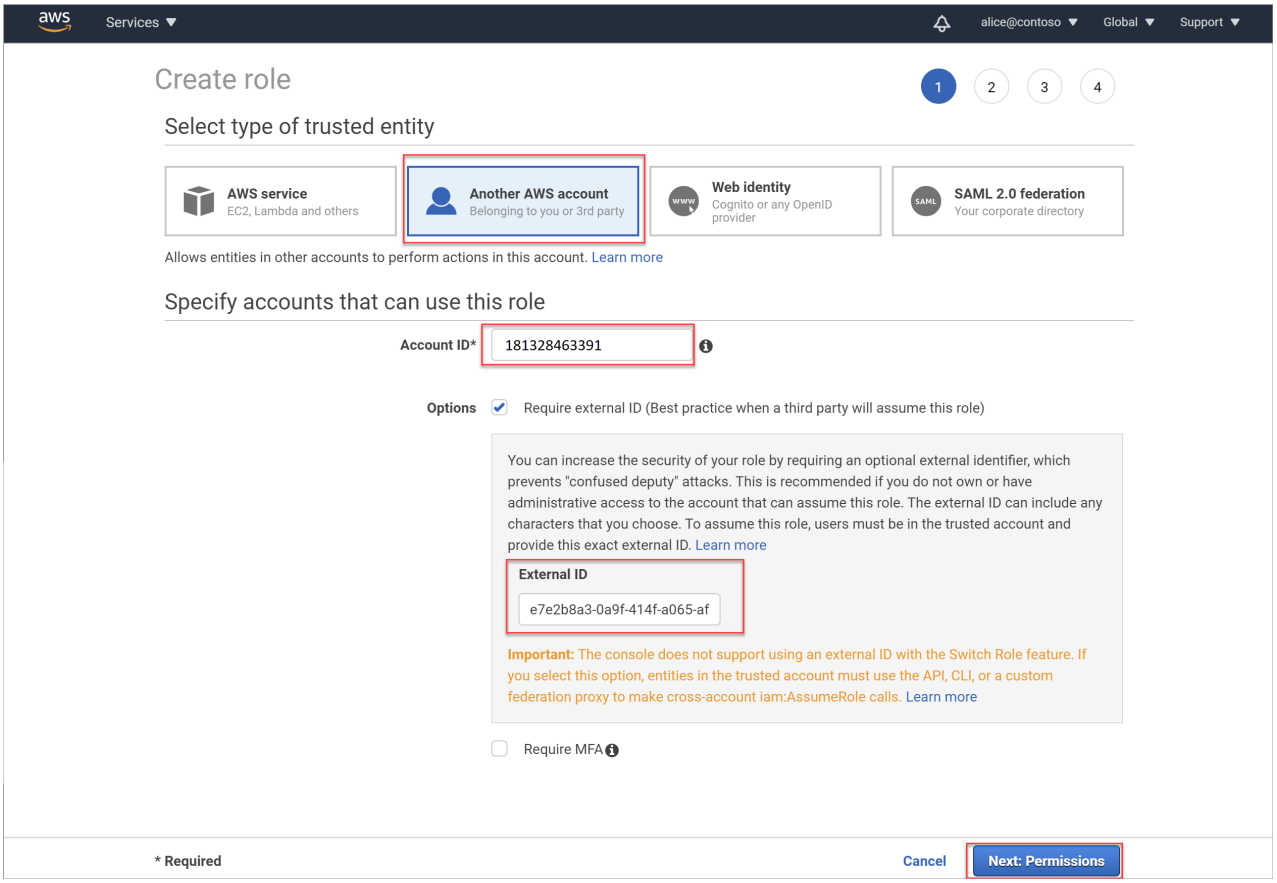
2. Select **Roles** and then **Create role**.

3. Select **Another AWS account**, and then enter the following values:

Expand table

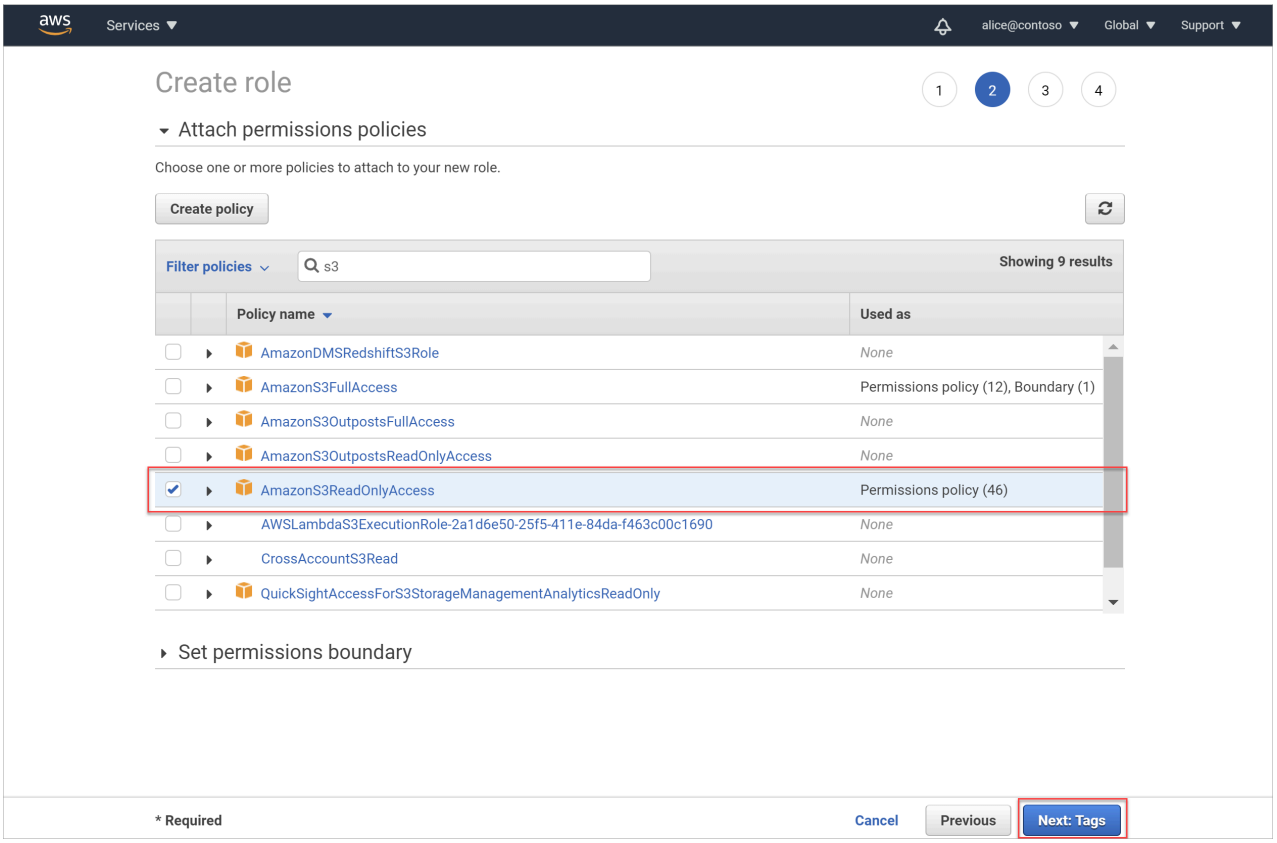
Field	Description
Account ID	Enter your Microsoft Account ID. For example: `181328463391`
External ID	Under options, select **Require external ID...**, and then enter your External ID in the designated field. For example: `e7e2b8a3-0a9f-414f-a065-afaf4ac6d994`

For example:



4. In the **Create role** > **Attach permissions policies** area, filter the permissions displayed to **S3**. Select **AmazonS3ReadOnlyAccess**. If you also want to be able to apply protection policies, add these permissions also:

    - GetBucketLocation
    - GetBucketPublicAccessBlock
    - GetObject
    - PutBucketPolicy
    - PutObjectTagging
    - DeleteBucketPolicy
    - List Bucket



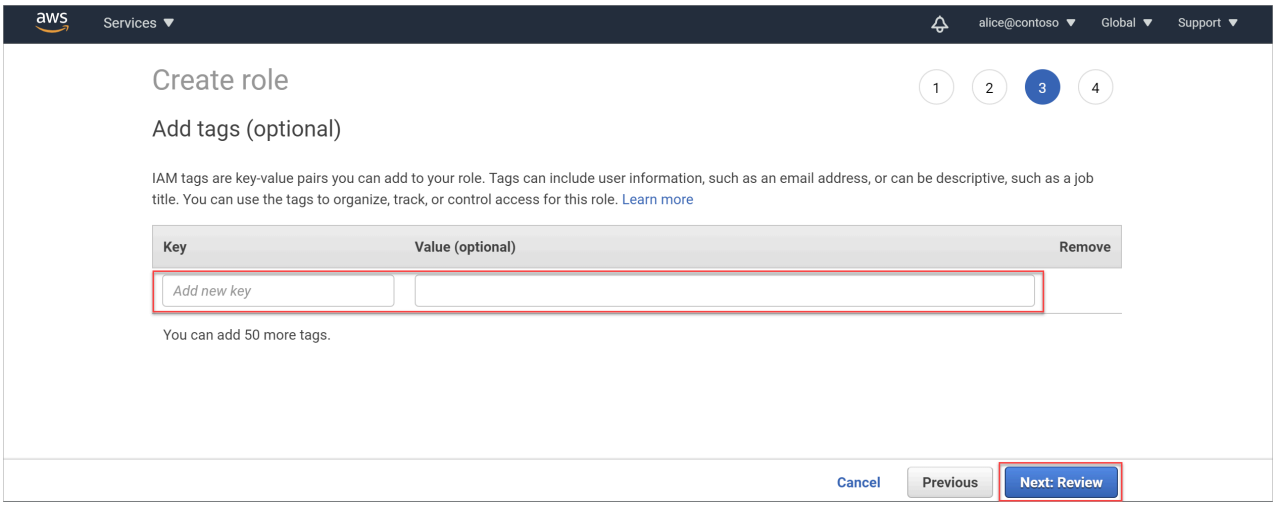
> **ⓘ Important**
>
> The **AmazonS3ReadOnlyAccess** policy provides minimum permissions required for scanning your S3 buckets, and may include other permissions as well.
>
> To apply only the minimum permissions required for scanning your buckets, create a new policy with the permissions listed in Minimum permissions for your AWS policy, depending on whether you want to scan a single bucket or all the buckets in your account.
>
> Apply your new policy to the role instead of **AmazonS3ReadOnlyAccess.**

5. Select **Next: Tags.**

6. In the **Add tags (optional)** area, you can optionally choose to create a meaningful tag for this new role. Useful tags enable you to organize, track, and control access for each role you create.

   If you also want to be able to apply protection policies, add a tag with the key: `msftpurview_allowlisted` and the value `true`.

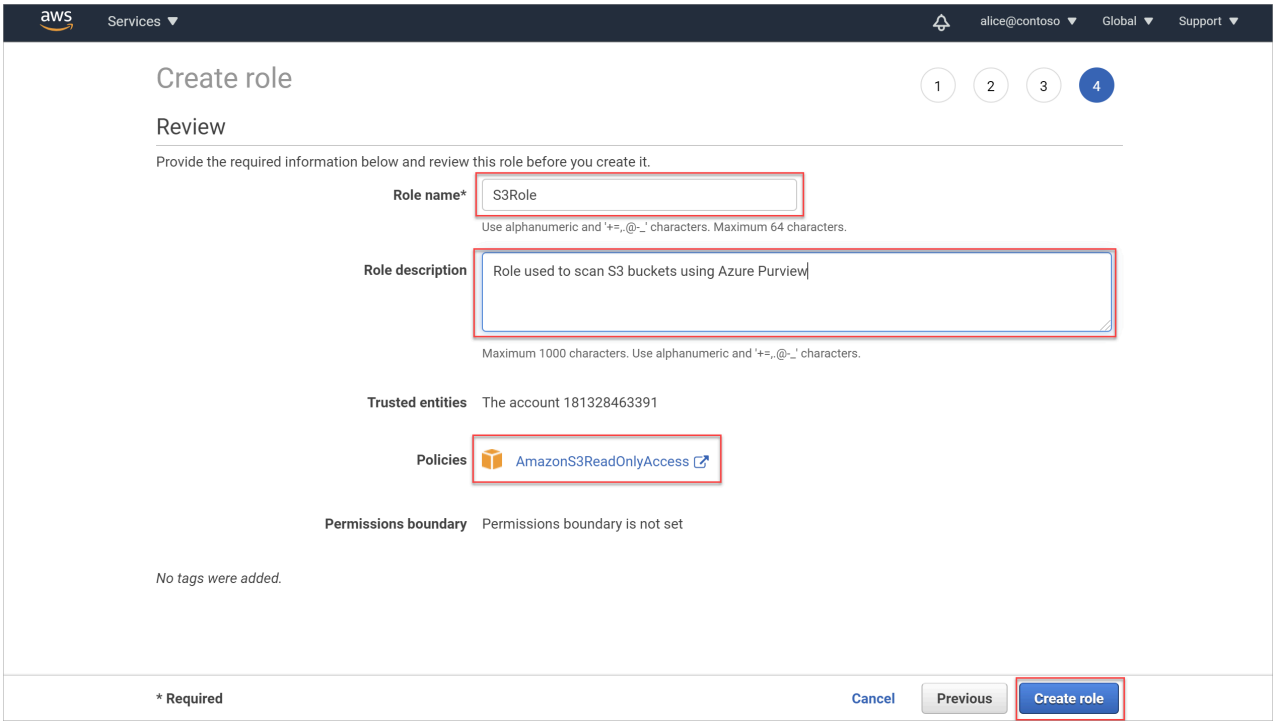
   Enter a new key and value for your tag as needed. When you're done, or if you want to skip this step, select **Next: Review** to review the role details and complete the role creation.



7. In the **Review** area, do the following:

    - In the **Role name** field, enter a meaningful name for your role
    - In the **Role description** box, enter an optional description to identify the role's purpose
    - In the **Policies** section, confirm that the correct policy (**AmazonS3ReadOnlyAccess**) is attached to the role.

    Then select **Create role** to complete the process. For example:



Extra required configurations:

- For buckets that use **AWS-KMS** encryption, special configuration is required to enable scanning.

- Make sure that your bucket policy doesn't block the connection. For more information, see:

- Confirm your bucket policy access
- Confirm your SCP policy access

Create a Microsoft Purview credential for your AWS S3 scan

This procedure describes how to create a new Microsoft Purview credential to use when scanning your AWS buckets.

> **Tip**
>
> If you're continuing directly on from Create a new AWS role for Microsoft Purview, you may already have the **New credential** pane open in Microsoft Purview.
>
> You can also create a new credential in the middle of the process, while configuring your scan. In that case, in the **Credential** field, select **New**.

1. In Microsoft Purview, go to the **Management Center**, and under **Security and access**, select **Credentials**.

2. Select **New**, and in the **New credential** pane that appears on the right, use the following fields to create your Microsoft Purview credential:

   Expand table

Field	Description
Name	Enter a meaningful name for this credential.
Description	Enter an optional description for this credential, such as `Used to scan the tutorial S3 buckets`
Authentication method	Select **Role ARN**, since you're using a role ARN to access your bucket.
Role ARN	Once you've created your Amazon IAM role, navigate to your role in the AWS IAM area, copy the **Role ARN** value, and enter it here. For example: `arn:aws:iam::181328463391:role/S3Role`. For more information, see Retrieve your new Role ARN.

   The **Microsoft account ID** and the **External ID** values are used when creating your Role ARN in AWS..

3. Select **Create** when you're done to finish creating the credential.

For more information about Microsoft Purview credentials, see Credentials for source authentication in Microsoft Purview.

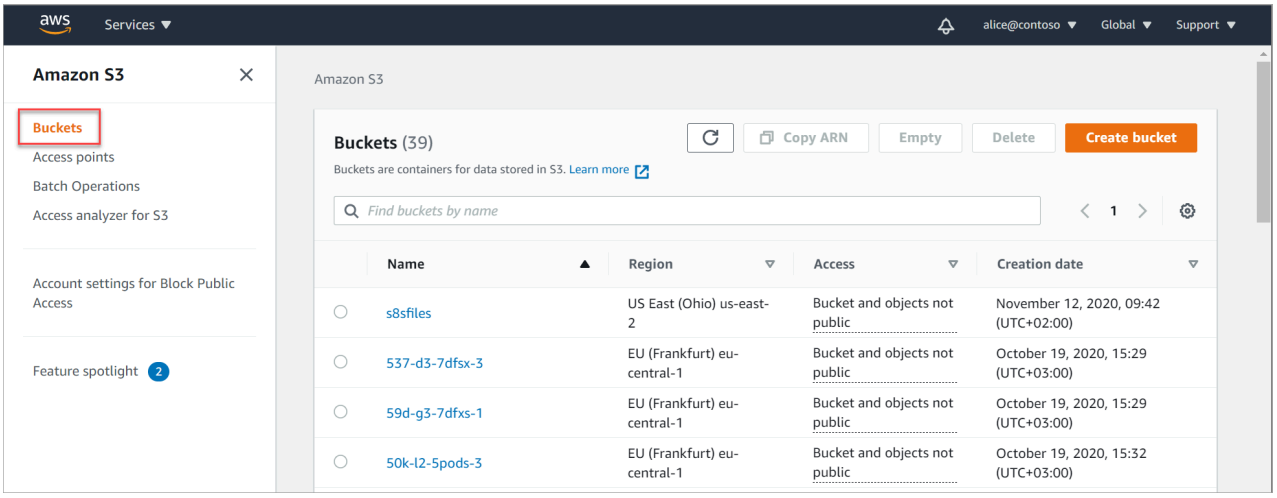
Configure scanning for encrypted Amazon S3 buckets

AWS buckets support multiple encryption types. For buckets that use **AWS-KMS** encryption, special configuration is required to enable scanning.

> **Note**
>
> For buckets that use no encryption, AES-256 or AWS-KMS S3 encryption, skip this section and continue to Retrieve your Amazon S3 bucket name.

To check the type of encryption used in your Amazon S3 buckets:

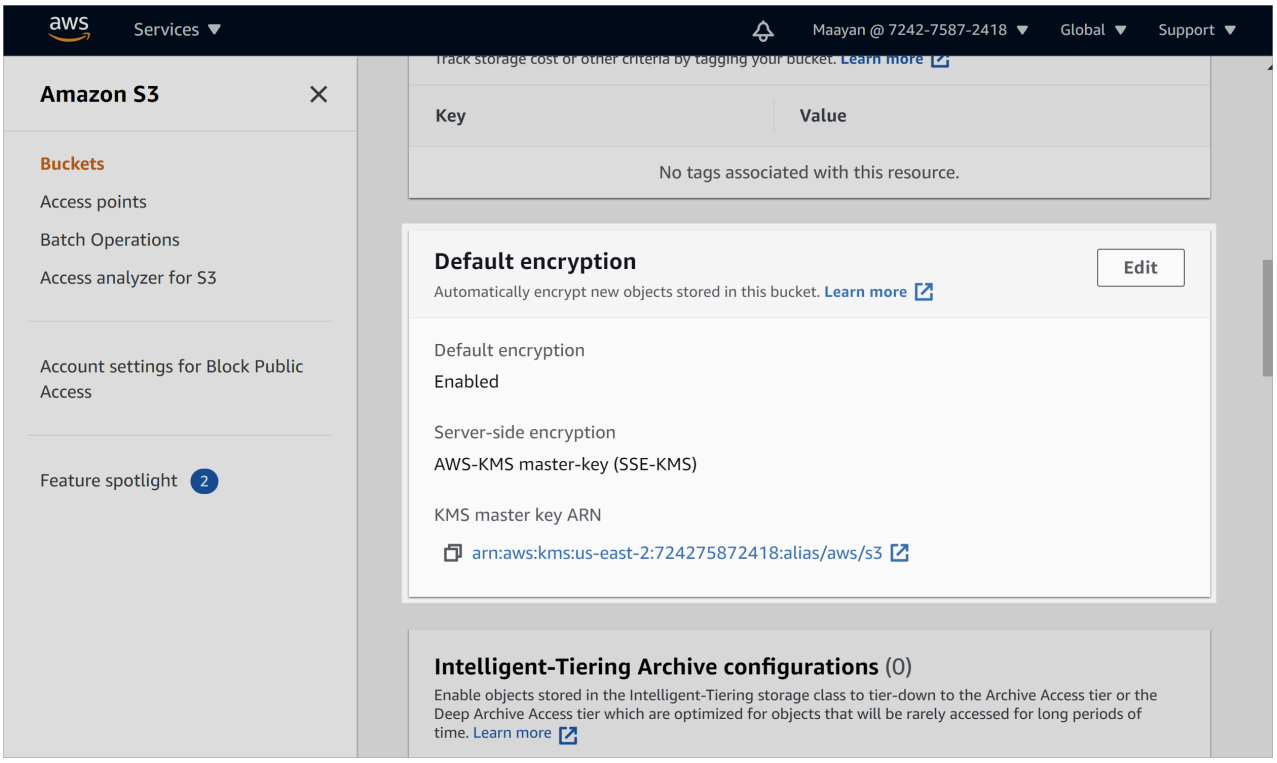
1. In AWS, navigate to **Storage** > **S3** > and select **Buckets** from the menu on the left.



2. Select the bucket you want to check. On the bucket's details page, select the **Properties** tab and scroll down to the **Default encryption** area.

    - If the bucket you selected is configured for anything but **AWS-KMS** encryption, including if default encryption for your bucket is **Disabled**, skip the rest of this procedure and continue with Retrieve your Amazon S3 bucket name.
    - If the bucket you selected is configured for **AWS-KMS** encryption, continue as described below to add a new policy that allows for scanning a bucket with custom **AWS-KMS** encryption.

    For example:



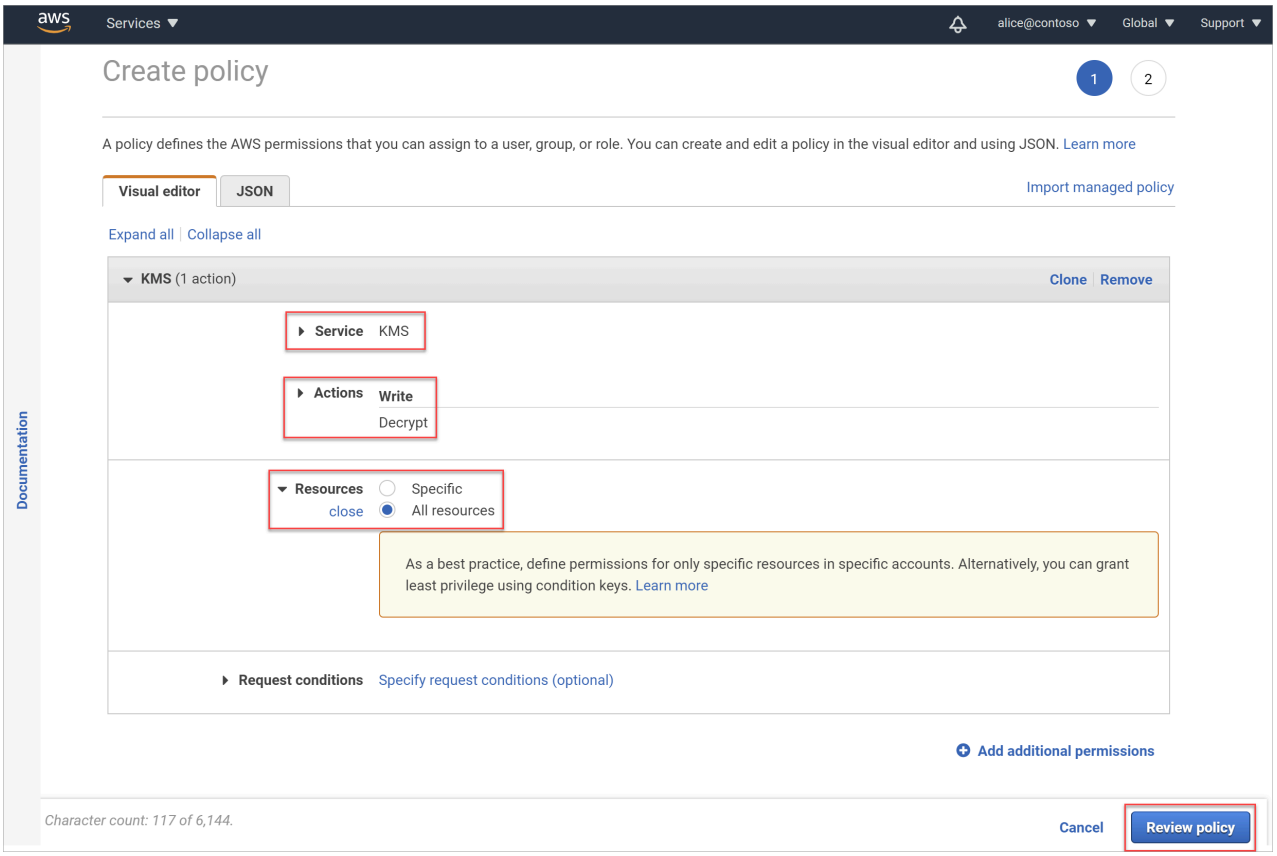
To add a new policy to allow for scanning a bucket with custom AWS-KMS encryption:

1. In AWS, navigate to **Services** > **IAM** > **Policies**, and select **Create policy**.
2. On the **Create policy** > **Visual editor** tab, define your policy with the following values:

Expand table

Field	Description
Service	Enter and select **KMS**.
Actions	Under **Access level**, select **Write** to expand the **Write** section. Once expanded, select only the **Decrypt** option.
Resources	Select a specific resource or **All resources**.

When you're done, select **Review policy** to continue.



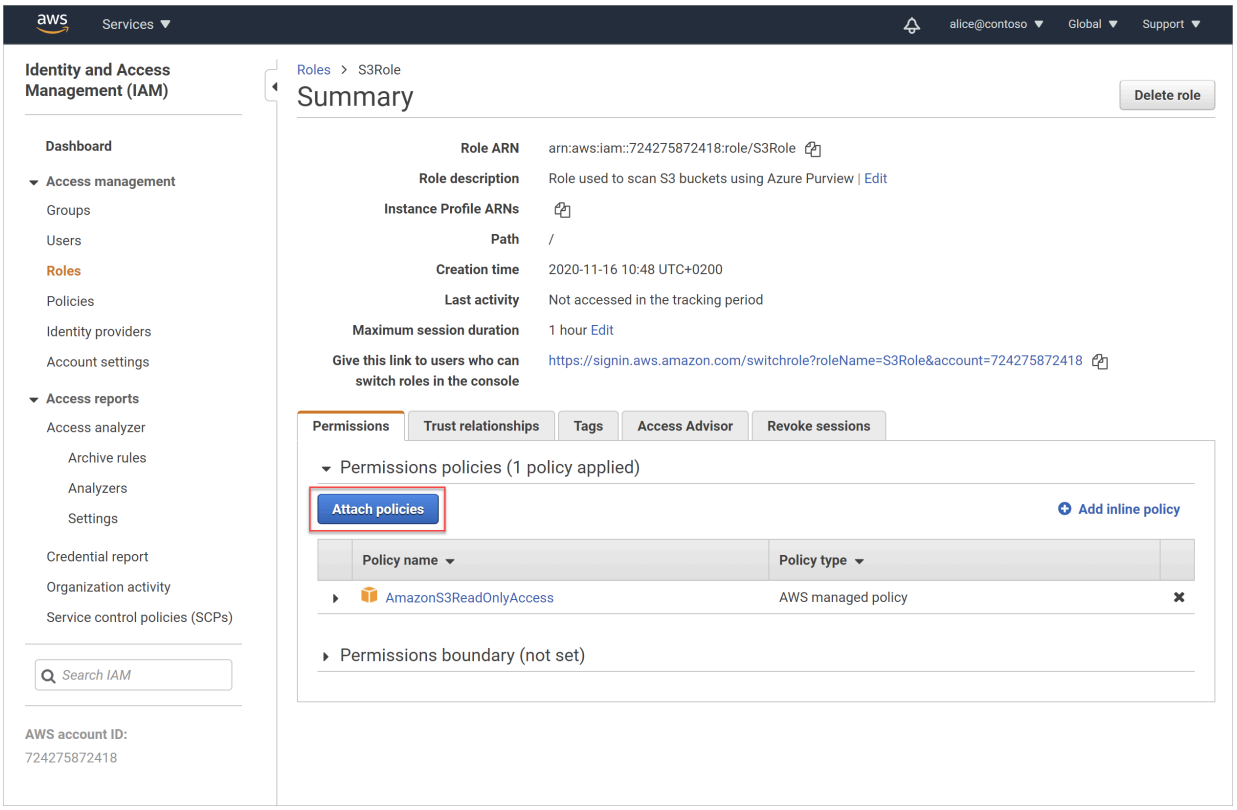
3. On the **Review policy** page, enter a meaningful name for your policy and an optional description, and then select **Create policy**.

   The newly created policy is added to your list of policies.

4. Attach your new policy to the role you added for scanning.

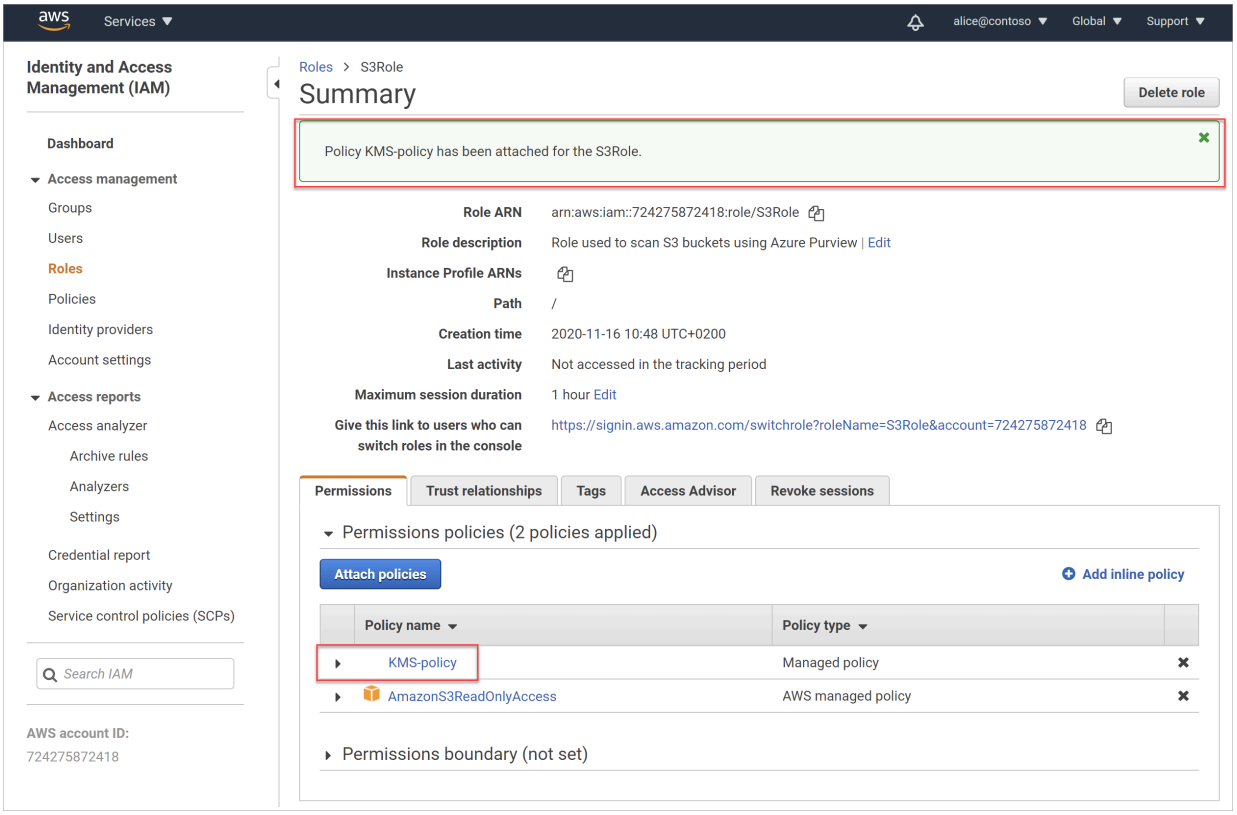
   a. Navigate back to the **IAM** > **Roles** page, and select the role you added earlier.

   b. On the **Permissions** tab, select **Attach policies**.



c. On the **Attach Permissions** page, search for and select the new policy you created above. Select **Attach policy** to attach your policy to the role.

The **Summary** page is updated, with your new policy attached to your role.



Confirm your bucket policy access

Make sure that the S3 bucket policy doesn't block the connection:

1. In AWS, navigate to your S3 bucket, and then select the **Permissions** tab > **Bucket policy**.
2. Check the policy details to make sure that it doesn't block the connection from the Microsoft Purview scanner service.

Confirm your SCP policy access

Make sure that there's no SCP policy that blocks the connection to the S3 bucket.

For example, your SCP policy might block read API calls to the AWS Region where your S3 bucket is hosted.

- Required API calls, which must be allowed by your SCP policy, include: `AssumeRole`, `GetBucketLocation`, `GetObject`, `ListBucket`, `GetBucketPublicAccessBlock`.
- Your SCP policy must also allow calls to the **us-east-1** AWS Region, which is the default Region for API calls. For more information, see the AWS documentation.

Follow the SCP documentation, review your organization's SCP policies, and make sure all the permissions required for the Microsoft Purview scanner are available.

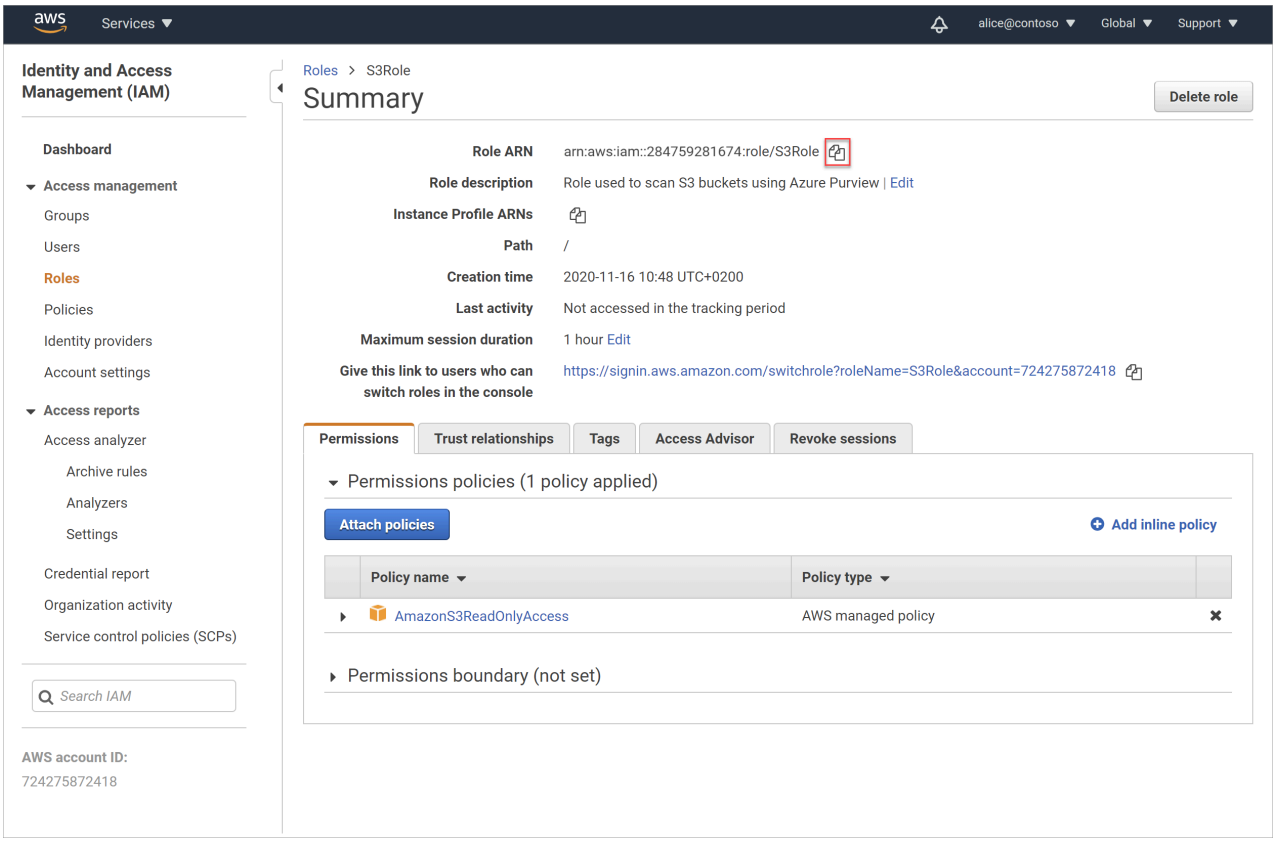
Retrieve your new Role ARN

You need to record your AWS Role ARN and copy it in to Microsoft Purview when creating a scan for your Amazon S3 bucket.

To retrieve your role ARN:

1. In the AWS **Identity and Access Management (IAM)** > **Roles** area, search for and select the new role you created for Microsoft Purview.

2. On the role's **Summary** page, select the **Copy to clipboard** button to the right of the **Role ARN** value.



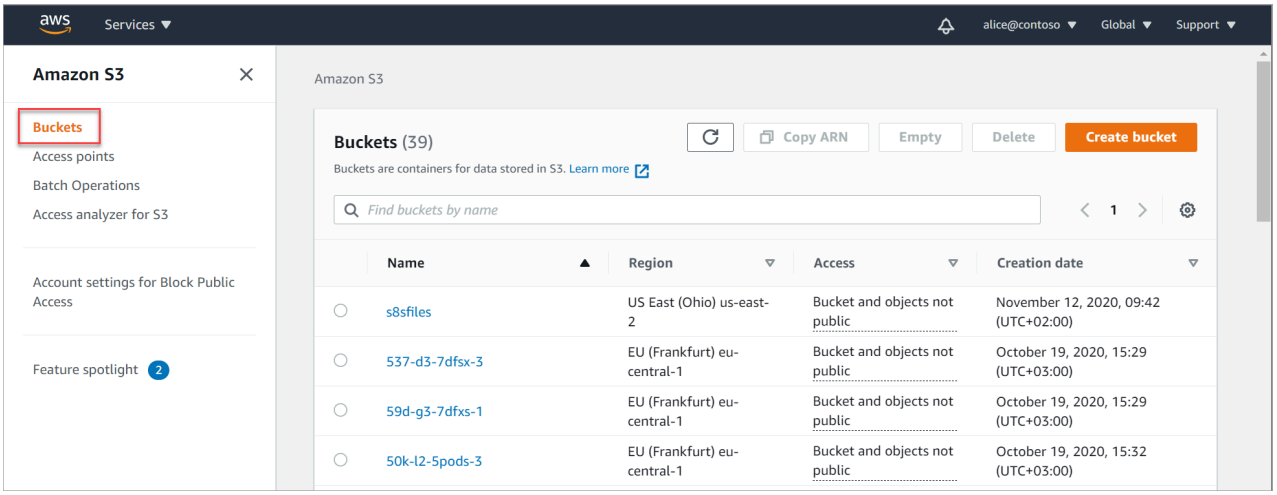
In Microsoft Purview, you can edit your credential for AWS S3, and paste the retrieved role in the **Role ARN** field. For more information, see Create a scan for one or more Amazon S3 buckets.

Retrieve your Amazon S3 bucket name

You need the name of your Amazon S3 bucket to copy it in to Microsoft Purview when creating a scan for your Amazon S3 bucket

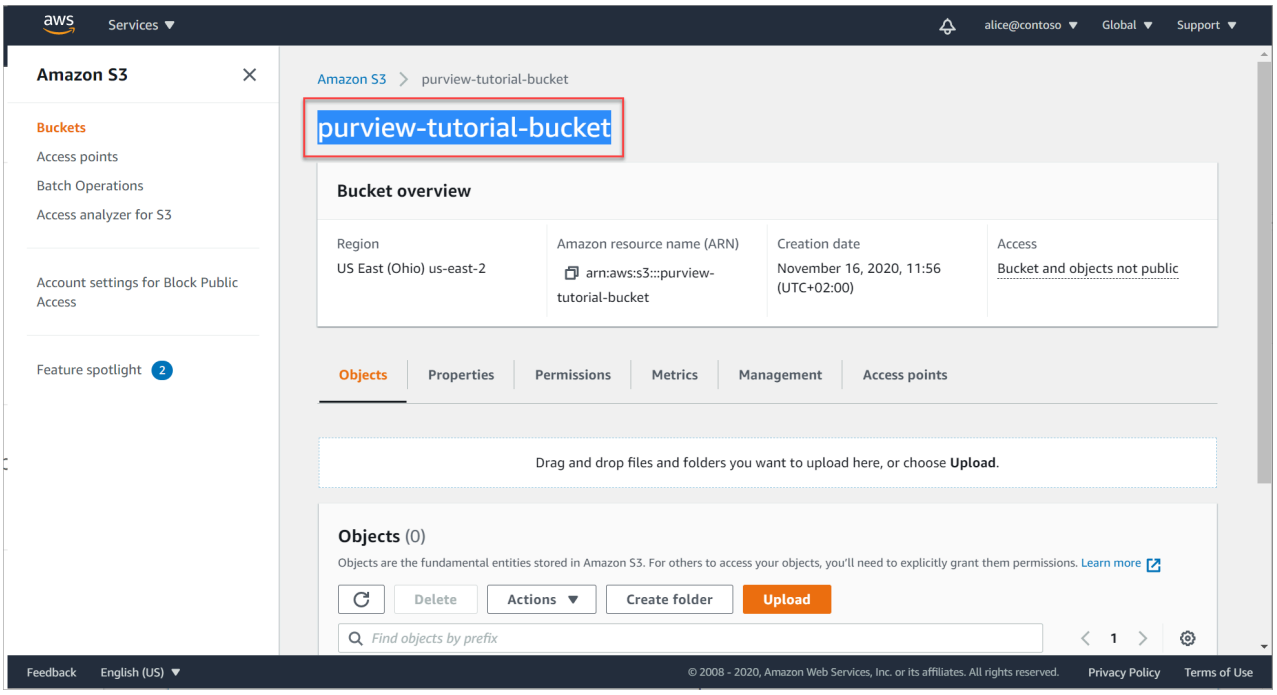
To retrieve your bucket name:

1. In AWS, navigate to **Storage** > **S3** > and select **Buckets** from the menu on the left.



2. Search for and select your bucket to view the bucket details page, and then copy the bucket name to the clipboard.

   For example:



   Paste your bucket name in a secure file, and add an `s3://` prefix to it to create the value you need to enter when configuring your bucket as a Microsoft Purview account.

   For example: `s3://purview-tutorial-bucket`

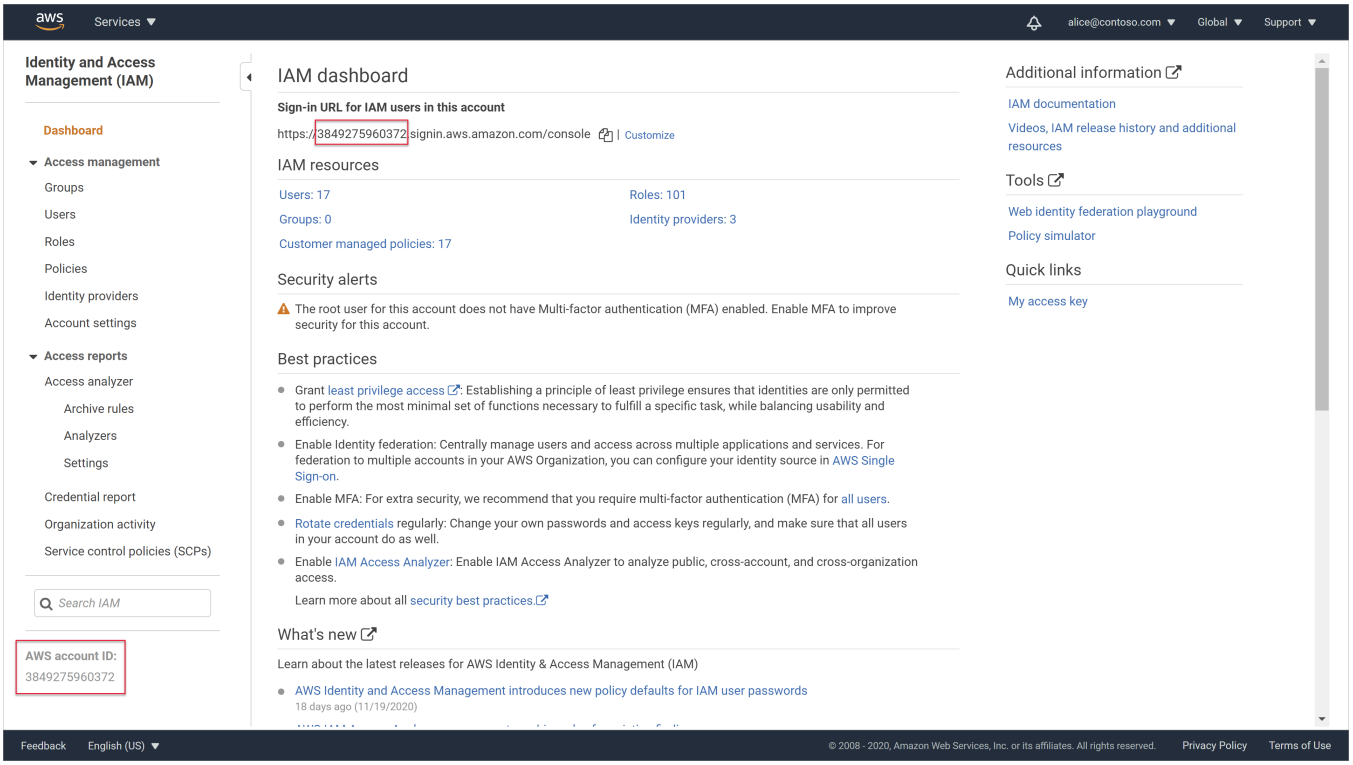
> **Tip**
>
> Only the root level of your bucket is supported as a Microsoft Purview data source. For example, the following URL, which includes a sub-folder is *not* supported: `s3://purview-tutorial-bucket/view-data`
>
> However, if you configure a scan for a specific S3 bucket, you can select one or more specific folders for your scan. For more information, see the step to scope your scan.

Locate your AWS account ID

You need your AWS account ID to register your AWS account as a Microsoft Purview data source, together with all of its buckets.

Your AWS account ID is the ID you use to sign in to the AWS console. You can also find it once you're logged in on the IAM dashboard, on the left under the navigation options, and at the top, as the numerical part of your sign-in URL:

For example:

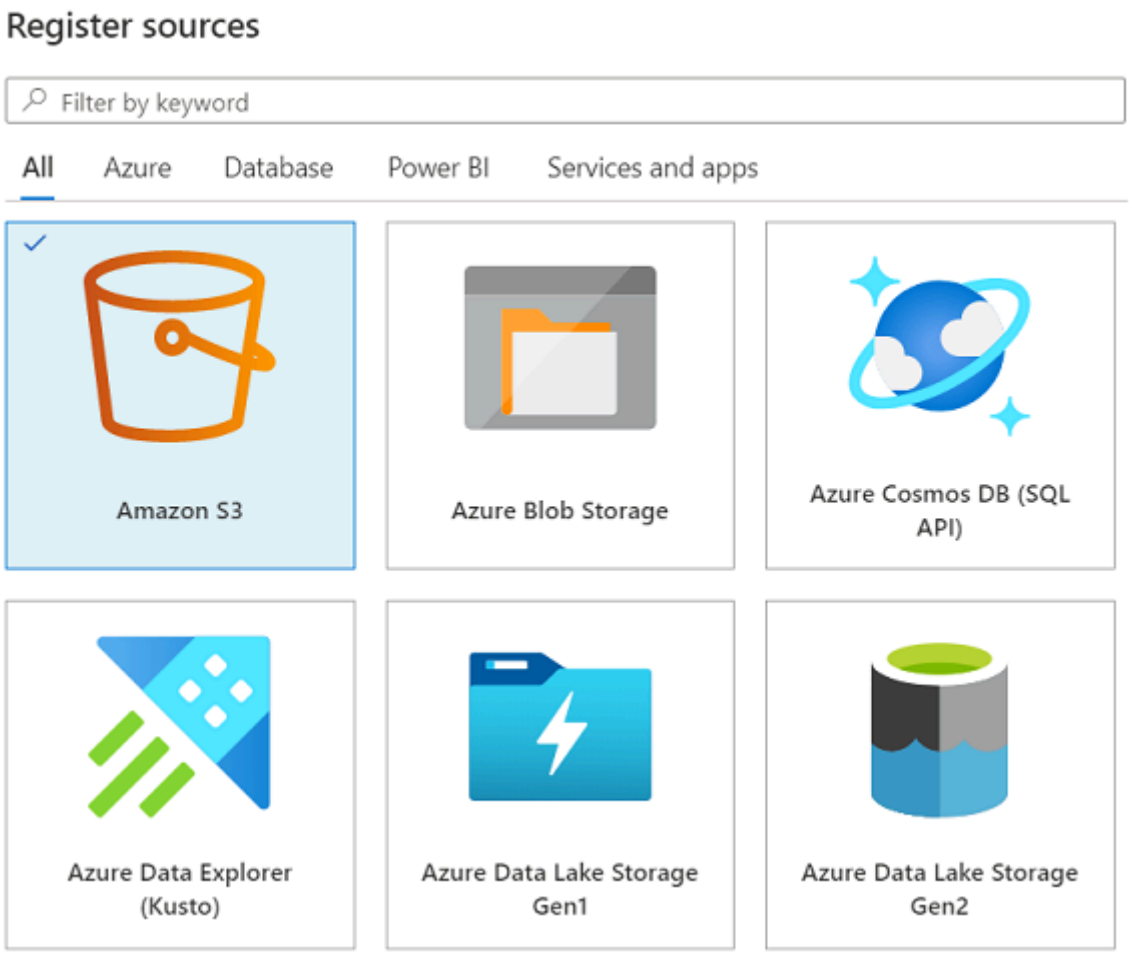


Add a single Amazon S3 bucket as a Microsoft Purview account

Use this procedure if you only have a single S3 bucket that you want to register to Microsoft Purview as a data source, or if you have multiple buckets in your AWS account, but don't want to register all of them to Microsoft Purview.

To add your bucket:

1. In Microsoft Purview, go to the **Data Map**, and select **Data sources** > **Register** > **Amazon S3** > **Continue**.



> **Tip**
>
> If you have multiple collections and want to add your Amazon S3 to a specific collection, select the **Map view** at the top right, and then select the **Register** button inside your collection.

2. In the **Register sources (Amazon S3)** pane that opens, enter the following details:

Expand table

Field	Description
Name	Enter a meaningful name, or use the default provided.
Bucket URL	Enter your AWS bucket URL, using the following syntax: `s3://<bucketName>` **Note**: Make sure to use only the root level of your bucket. For more information, see Retrieve your Amazon S3 bucket name.
Select a collection	If you selected to register a data source from within a collection, that collection already listed. Select a different collection as needed, **None** to assign no collection, or **New** to

Field	Description
	create a new collection now. For more information about Microsoft Purview collections, see Manage data sources in Microsoft Purview.

When you're done, select **Finish** to complete the registration.

Continue with Create a scan for one or more Amazon S3 buckets..

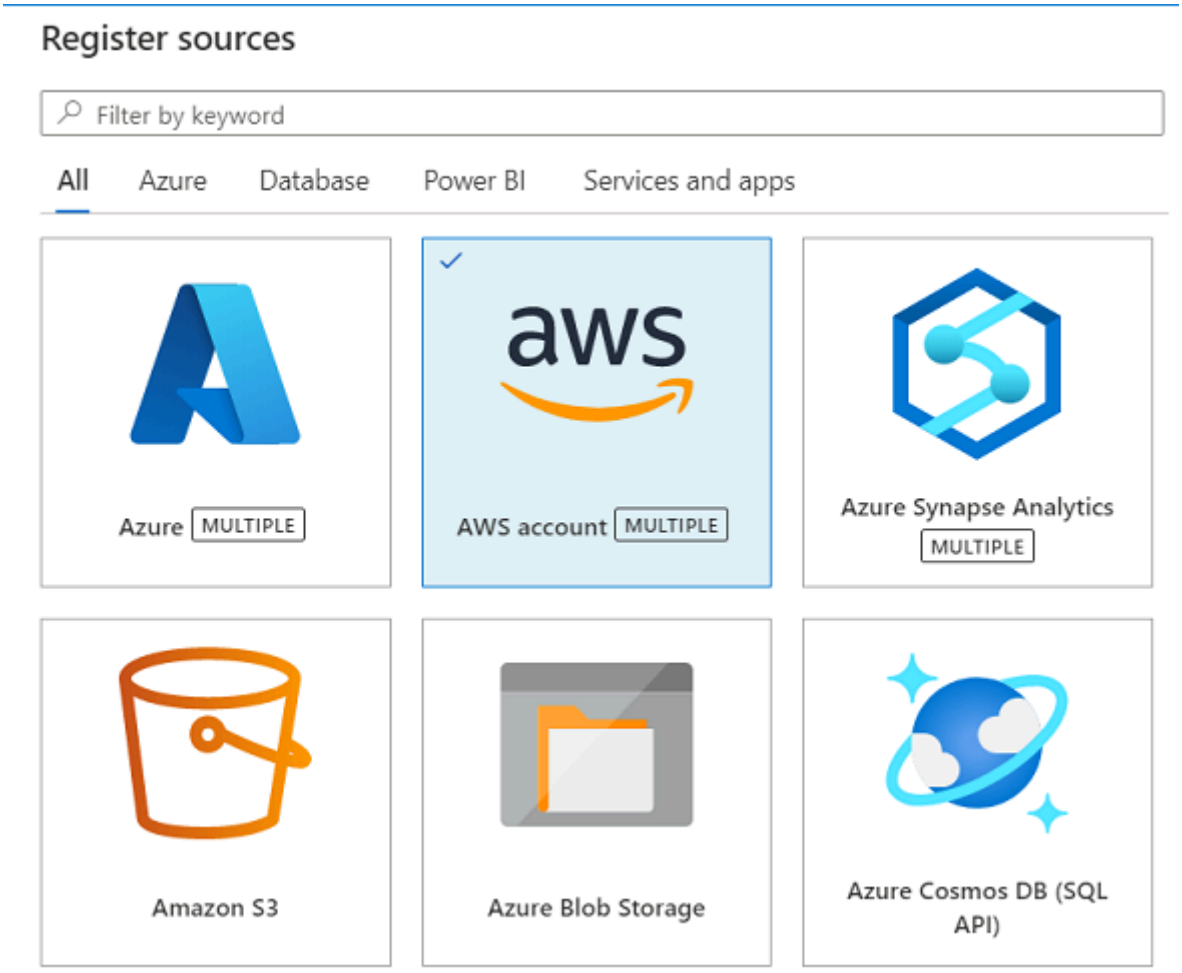
Add an AWS account as a Microsoft Purview account

Use this procedure if you have multiple S3 buckets in your Amazon account, and you want to register all of them as Microsoft Purview data sources.

When configuring your scan, you are able to select the specific buckets you want to scan, if you don't want to scan all of them together.

To add your Amazon account:

1. In Microsoft Purview, go to the **Data Map** page, and select **Data sources** > **Register** > **Amazon accounts** > **Continue**.



> **Tip**
>
> If you have multiple collections and want to add your Amazon S3 to a specific collection, select the **Map view** at the top right, and then select the **Register** button inside your collection.

2. In the **Register sources (Amazon S3)** pane that opens, enter the following details:

Expand table

Field	Description
Name	Enter a meaningful name, or use the default provided.
AWS account ID	Enter your AWS account ID. For more information, see Locate your AWS account ID
Select a collection	If you selected to register a data source from within a collection, that collection already listed. Select a different collection as needed, **None** to assign no collection, or **New** to create a new collection now.

Field	Description
	For more information about Microsoft Purview collections, see Manage data sources in Microsoft Purview.

When you're done, select **Finish** to complete the registration.

Continue with Create a scan for one or more Amazon S3 buckets.

Create a scan for one or more Amazon S3 buckets

Once you've added your buckets as Microsoft Purview data sources, you can configure a scan to run at scheduled intervals or immediately.

1. Select the **Data Map** in the Microsoft Purview portal:

   - In the **Map view**, select **New scan** in your data source box.
   - In the **List view**, hover over the row for your data source, and select **New scan**.

2. On the **Scan...** pane that opens on the right, define the following fields and then select **Continue**:

Expand table

Field	Description
Name	Enter a meaningful name for your scan or use the default.
Type	Displayed only if you've added your AWS account, with all buckets included. Current options include only **All** > **Amazon S3**. Stay tuned for more options to select as Microsoft Purview's support matrix expands.
Credential	Select a Microsoft Purview credential with your role ARN. **Tip**: If you want to create a new credential at this time, select **New**. For more information, see Create a Microsoft Purview credential for your AWS bucket scan.
Amazon S3	Displayed only if you've added your AWS account, with all buckets included. Select one or more buckets to scan, or **Select all** to scan all the buckets in your account.

Microsoft Purview automatically checks that the role ARN is valid, and that the buckets and objects within the buckets are accessible, and then continues if the connection succeeds.

For **Integration runtime**, see the list of supported data sources for scans to choose the right integration runtime according to the scenario.

> **Tip**
>
> To enter different values and test the connection yourself before continuing, select **Test connection** at the bottom right before selecting **Continue**.

3. On the **Scope your scan** pane, select the specific buckets or folders you want to include in your scan.

   When creating a scan for an entire AWS account, you can select specific buckets to scan. When creating a scan for a specific AWS S3 bucket, you can select specific folders to scan.

4. On the **Select a scan rule set** pane, either select the **AmazonS3** default rule set, or select **New scan rule set** to create a new custom rule set. Once you have your rule set selected, select **Continue**.

   If you select to create a new custom scan rule set, use the wizard to define the following settings:

Expand table

Pane	Description
New scan rule set / **Scan rule description**	Enter a meaningful name and an optional description for your rule set
Select file types	Select all the file types you want to include in the scan, and then select **Continue**. To add a new file type, select **New file type**, and define the following: - The file extension you want to add - An optional description - Whether the file contents have a custom delimiter, or are a system file type. Then, enter your custom delimiter, or select your system file type. Select **Create** to create your custom file type.
Select classification rules	Navigate to and select the classification rules you want to run on your dataset.

   Select **Create** when you're done to create your rule set.

5. On the **Set a scan trigger** pane, select one of the following, and then select **Continue**:

    - **Recurring** to configure a schedule for a recurring scan
    - **Once** to configure a scan that starts immediately

6. On the **Review your scan** pane, check your scanning details to confirm that they're correct, and then select **Save** or **Save and Run** if you selected **Once** in the previous pane.

    > **Note**
    >
    > Once started, scanning can take up to 24 hours to complete. You'll be able to review your **Insight Reports** and search the catalog 24 hours after you started each scan.

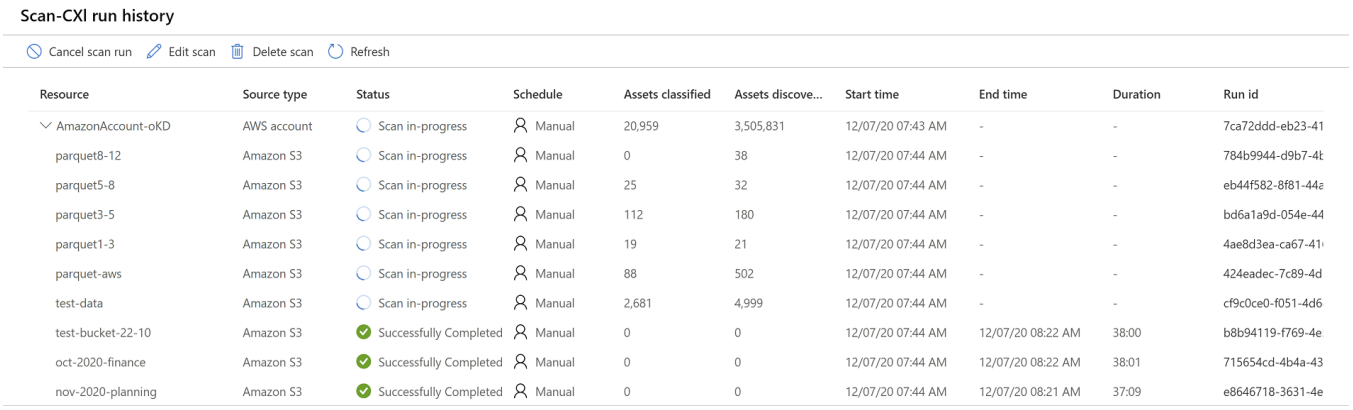
For more information, see Explore Microsoft Purview scanning results.

Explore Microsoft Purview scanning results

Once a Microsoft Purview scan is complete on your Amazon S3 buckets, drill down in the Microsoft Purview **Data Map** area to view the scan history.

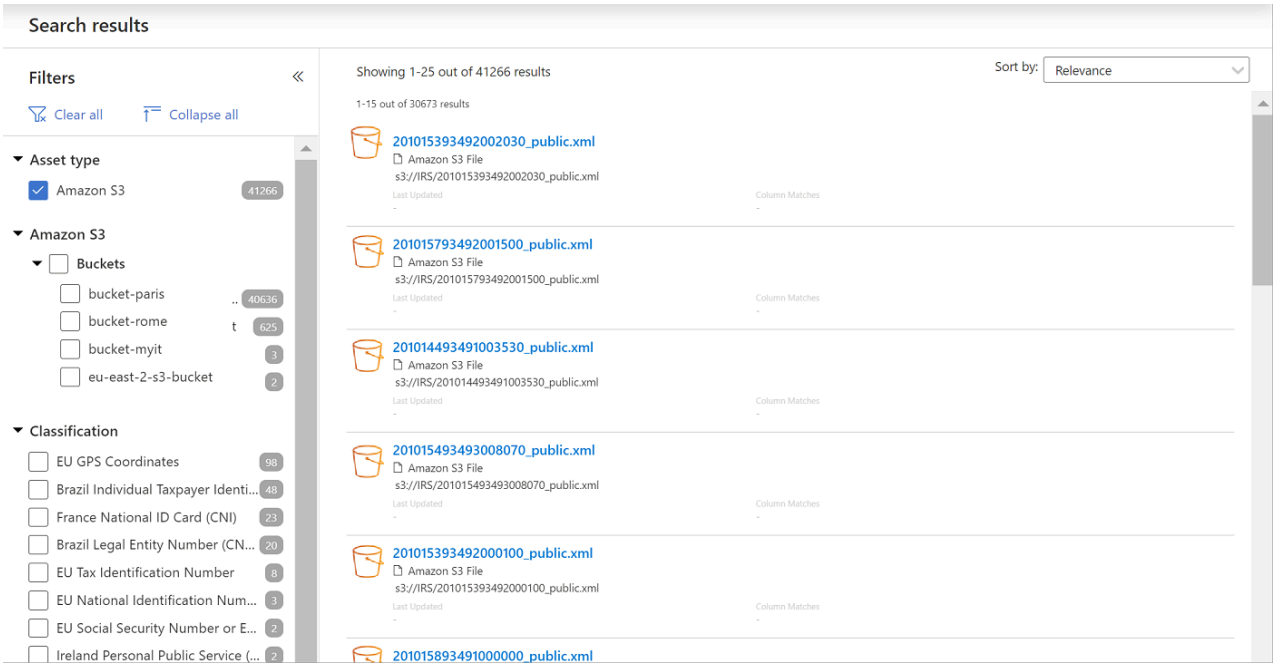
Select a data source to view its details, and then select the **Scans** tab to view any currently running or completed scans. If you've added an AWS account with multiple buckets, the scan history for each bucket is shown under the account.

For example:



Use the other areas of Microsoft Purview to find out details about the content in your data estate, including your Amazon S3 buckets:

- **Search the Microsoft Purview Unified Catalog,** and filter for a specific bucket. For example:



- **View Insight reports** to view statistics for the classification, sensitivity labels, file types, and more details about your content.

  All Microsoft Purview Insight reports include the Amazon S3 scanning results, along with the rest of the results from your Azure data sources. When relevant, another **Amazon S3** asset type was added to the report filtering options.

  For more information, see the Understand Data Estate Insights in Microsoft Purview.

Minimum permissions for your AWS policy

The default procedure for creating an AWS role for Microsoft Purview to use when scanning your S3 buckets uses the **AmazonS3ReadOnlyAccess** policy.

The **AmazonS3ReadOnlyAccess** policy provides minimum permissions required for scanning your S3 buckets, and might include other permissions as well.

To apply only the minimum permissions required for scanning your buckets, create a new policy with the permissions listed in the following sections, depending on whether you want to scan a single bucket or all the buckets in your account.

Apply your new policy to the role instead of **AmazonS3ReadOnlyAccess.**

Individual buckets

When scanning individual S3 buckets, minimum AWS permissions include:

- `GetBucketLocation`

- `GetBucketPublicAccessBlock`
- `GetObject`
- `ListBucket`

Make sure to define your resource with the specific bucket name. For example:

JSON

```
{
"Version": "2012-10-17",
"Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "s3:GetBucketLocation",
                "s3:GetBucketPublicAccessBlock",
                "s3:GetObject",
                "s3:ListBucket"
            ],
            "Resource": "arn:aws:s3:::<bucketname>"
        },
        {
            "Effect": "Allow",
            "Action": [
                "s3:GetObject"
            ],
            "Resource": "arn:aws:s3::: <bucketname>/*"
        }
    ]
}
```

All buckets in your account

When scanning all the buckets in your AWS account, minimum AWS permissions include:

- `GetBucketLocation`
- `GetBucketPublicAccessBlock`
- `GetObject`
- `ListAllMyBuckets`
- `ListBucket`.

Make sure to define your resource with a wildcard. For example:

JSON

```
{
"Version": "2012-10-17",
"Statement": [
```

```
        {
            "Effect": "Allow",
            "Action": [
                "s3:GetBucketLocation",
                "s3:GetBucketPublicAccessBlock",
                "s3:GetObject",
                "s3:ListAllMyBuckets",
                "s3:ListBucket"
            ],
            "Resource": "*"
        },
        {
            "Effect": "Allow",
            "Action": [
                "s3:GetObject"
            ],
            "Resource": "*"
        }
    ]
}
```

Troubleshooting

Scanning Amazon S3 resources requires creating a role in AWS IAM to allow the Microsoft Purview scanner service running in a Microsoft account in AWS to read the data.

Configuration errors in the role can lead to connection failure. This section describes some examples of connection failures that might occur while setting up the scan, and the troubleshooting guidelines for each case.

If all of the items described in the following sections are properly configured, and scanning S3 buckets still fails with errors, contact Microsoft support.

> **Note**
>
> For policy access issues, make sure that neither your bucket policy, nor your SCP policy are blocking access to your S3 bucket from Microsoft Purview.
>
> For more information, see Confirm your bucket policy access and Confirm your SCP policy access.

Bucket is encrypted with KMS

Make sure that the AWS role has **KMS Decrypt** permissions. For more information, see Configure scanning for encrypted Amazon S3 buckets.

AWS role is missing an external ID

Make sure that the AWS role has the correct external ID:

1. In the AWS IAM area, select the **Role** > **Trust relationships** tab.
2. Follow the steps in Create a new AWS role for Microsoft Purview again to verify your details.

Error found with the role ARN

This is a general error that indicates an issue when using the Role ARN. For example, you may want to troubleshoot as follows:

- Make sure that the AWS role has the required permissions to read the selected S3 bucket. Required permissions include `AmazonS3ReadOnlyAccess` or the minimum read permissions, and `KMS Decrypt` for encrypted buckets.

- Make sure that the AWS role has the correct Microsoft account ID. In the AWS IAM area, select the **Role** > **Trust relationships** tab and then follow the steps in Create a new AWS role for Microsoft Purview again to verify your details.

For more information, see Can't find the specified bucket,

Can't find the specified bucket

Make sure that the S3 bucket URL is properly defined:

1. In AWS, navigate to your S3 bucket, and copy the bucket name.
2. In Microsoft Purview, edit the Amazon S3 data source, and update the bucket URL to include your copied bucket name, using the following syntax: `s3://<BucketName>`

Protection policy

Protection access control policies (protection policies) enable organizations to automatically protect sensitive data across data sources. Microsoft Purview already scans data assets and identifies sensitive data elements, and this new feature allows you to automatically restrict access to that data using sensitivity labels from Microsoft Purview Information Protection.

To set up protection policies for Amazon S3 sources, see: How to create protection policies for Amazon S3.

Next steps

Learn more about Microsoft Purview Insight reports:

Understand Data Estate Insights in Microsoft Purview

Last updated on 02/02/2026

Scans and ingestion in Data Map

This article provides an overview of the scanning and ingestion features in Microsoft Purview Data Map. These features connect your Microsoft Purview account to your sources to populate Data Map and Unified Catalog so you can begin exploring and managing your data through Microsoft Purview.

- **Scanning** captures metadata from data sources and brings it to Microsoft Purview.
- **Ingestion** processes metadata and stores it in Unified Catalog from both:
  - Data source scans - scanned metadata is added to Data Map.
  - Lineage connections - transformation resources add metadata about their sources, outputs, and activities to Data Map.

Scanning

After you register data sources in your Microsoft Purview account, scan the data sources. The scanning process connects to the data source and captures technical metadata like names, file size, columns, and more. It also extracts schema for structured data sources, applies classifications on schemas, and applies sensitivity labels if your Data Map is connected to Microsoft Purview portal. You can trigger the scanning process to run immediately or schedule it to run periodically to keep your Microsoft Purview account up to date.

For each scan, you can customize the process so that you only scan the information you need, rather than the whole source.

Choose an authentication method for your scans

Microsoft Purview is secure by default. It doesn't store passwords or secrets directly, so you need to choose an authentication method for your sources. You can authenticate your Microsoft Purview account in several ways, but not all methods are supported for each data source.

- Managed Identity
- Service Principal
- SQL Authentication
- Windows Authentication
- Role ARN
- Delegated Authentication
- Consumer Key
- Account Key or Basic Authentication

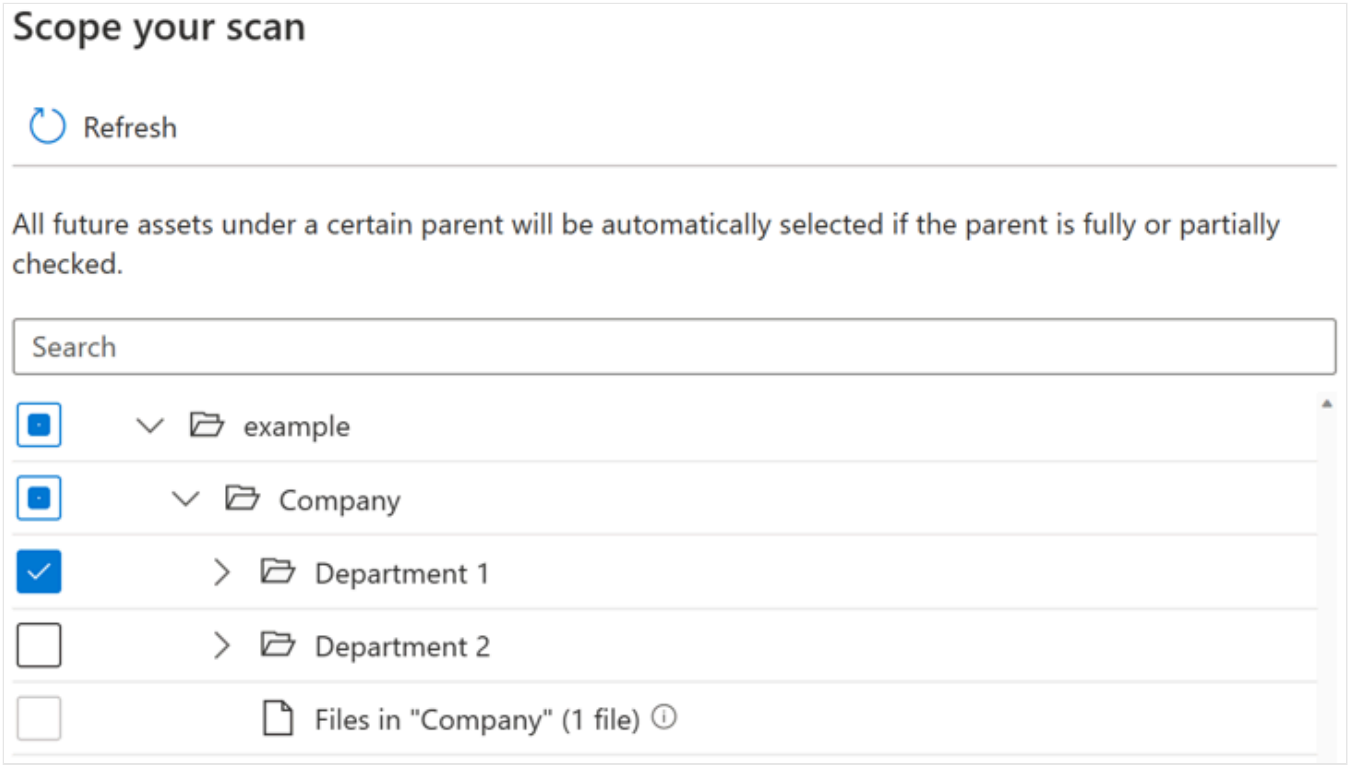
Whenever possible, use a Managed Identity because it eliminates the need for storing and managing credentials for individual data sources. This method can greatly reduce the time you and your team spend setting up and troubleshooting authentication for scans. When you enable a managed identity for your Microsoft Purview account, an identity is created in Microsoft Entra ID (Entra ID) and is tied to the lifecycle of your account.

Scope your scan

When scanning a source, you can scan the entire data source or choose only specific entities (folders or tables) to scan. The available options depend on the source you're scanning. You can define these options for both one-time and scheduled scans.

For example, when creating and running a scan for an Azure SQL Database, you can choose which tables to scan, or select the entire database.

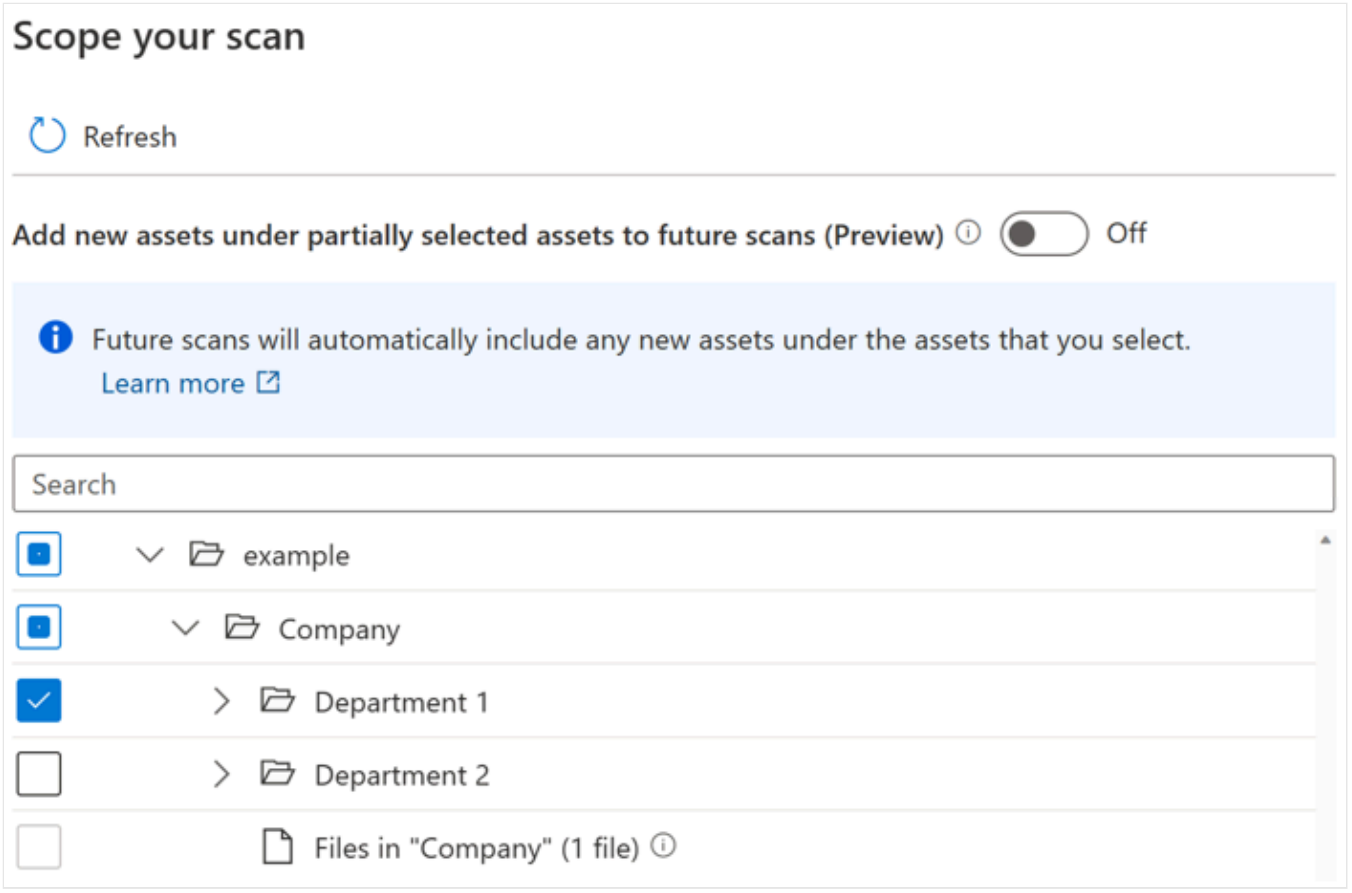
For each entity (folder or table), three selection states exist: fully selected, partially selected, and not selected. In the following example, if you select **Department 1** on the folder hierarchy, **Department 1** is considered as fully selected. The parent entities for **Department 1**, which are **Company** and **example**, are considered as partially selected, since the other entities under the same parent (for example, **Department 2**) aren't selected. Different icons are used on the UI for entities with different selection states.



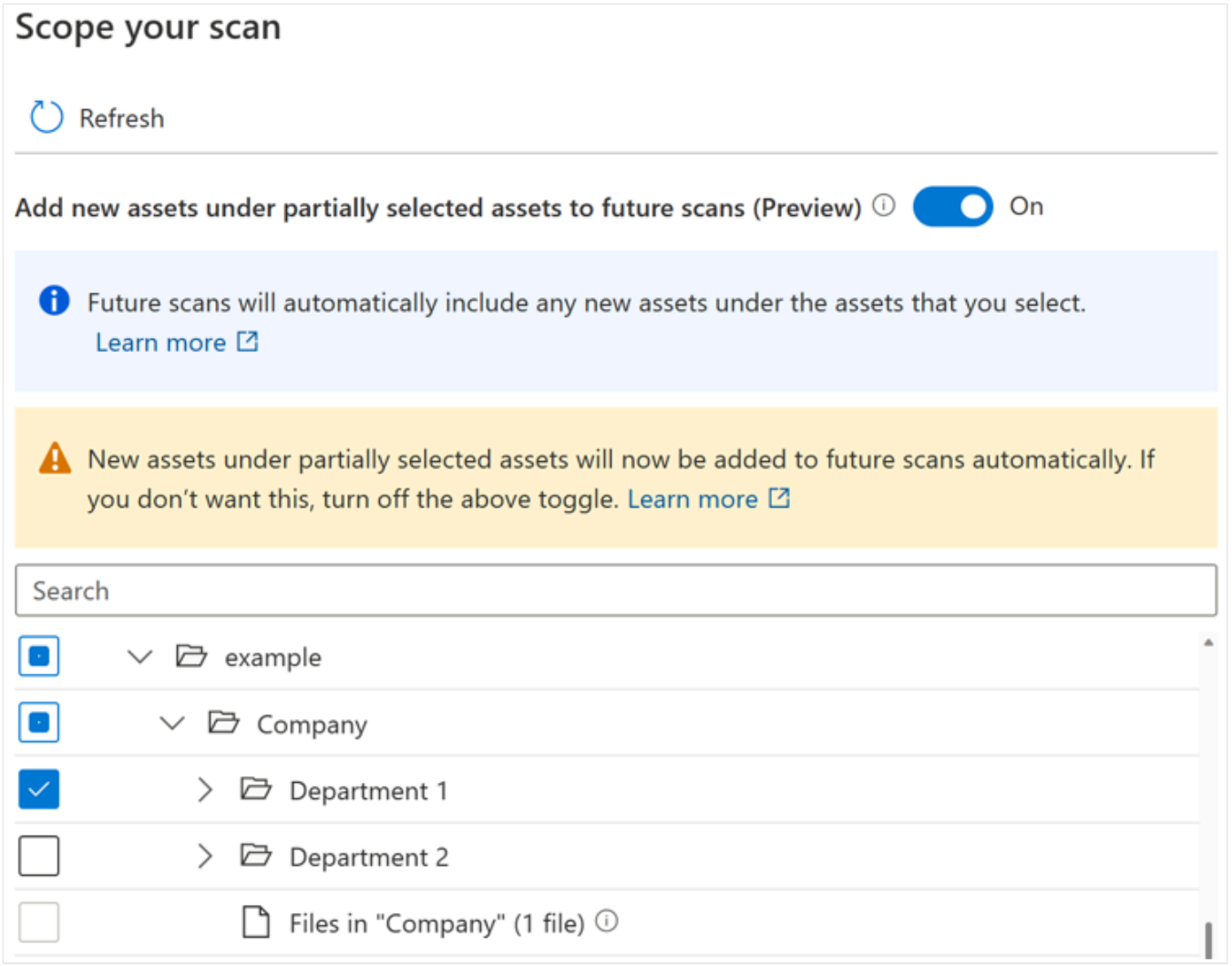
After you run the scan, it's likely that new assets will be added in the source system. By default, future assets under a certain parent are automatically selected if the parent is fully or partially selected when you run the scan again. In the preceding example, after you select **Department 1**

and run the scan, any new assets under folder **Department 1** or under **Company** and **example** are included when you run the scan again.

A toggle button, as seen in the image below, allows users to control the automatic inclusion for new assets under partially selected parent. By default, the toggle is turned off and the automatic inclusion behavior for a partially selected parent is disabled. With the toggle turned off, any new assets under partially selected parents like **Company** and **example** aren't included when you run the scan again; only new assets under **Department 1** are included in future scan.



If the toggle is turned on, the new assets under a parent are automatically selected if the parent is fully or partially selected when you run the scan again. The inclusion behavior is the same as before the toggle is introduced.



> **Note**
>
> - The availability of the toggle button depends on the data source type. Currently it's available in public preview for sources including Azure Blob Storage, Azure Data Lake Storage Gen 1, Azure Data Lake Storage Gen 2, Azure Files, and Azure Dedicated SQL pool (formerly SQL DW).
> - For any scans created or scheduled before the toggle is introduced, the toggle state is set as **On** and can't be changed. For any scans created or scheduled after the toggle is introduced, the toggle state can't be changed after the scan is saved. You need to create a new scan to change the toggle state.
> - When the toggle is turned off, for sources of storage type like Azure Data Lake Storage Gen 2, it can take up to four hours before the browse by source type experience becomes fully available after your scan job is completed.

Known limitations

When the toggle is turned off:

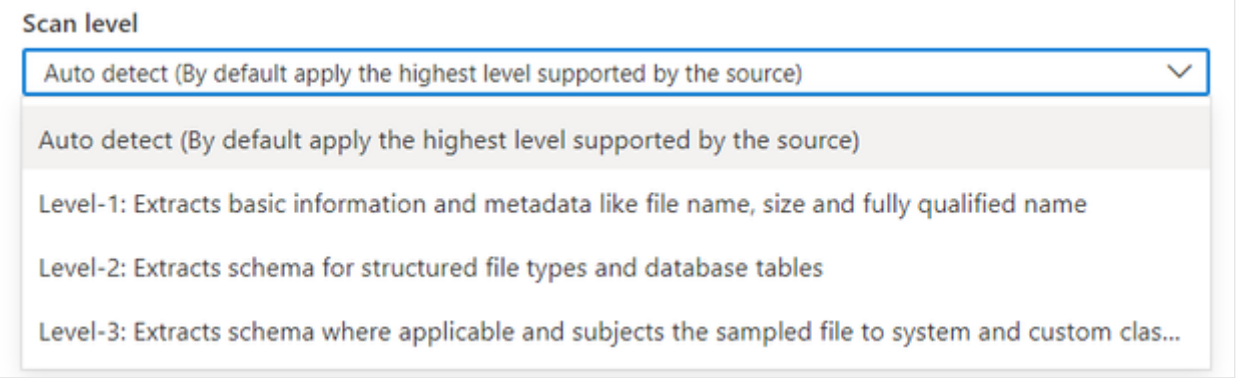
- The file entities under a partially selected parent aren't scanned.
- If all existing entities under a parent are explicitly selected, the parent is considered fully selected, and any new assets under the parent are included when you run the scan again.

Customize scan level

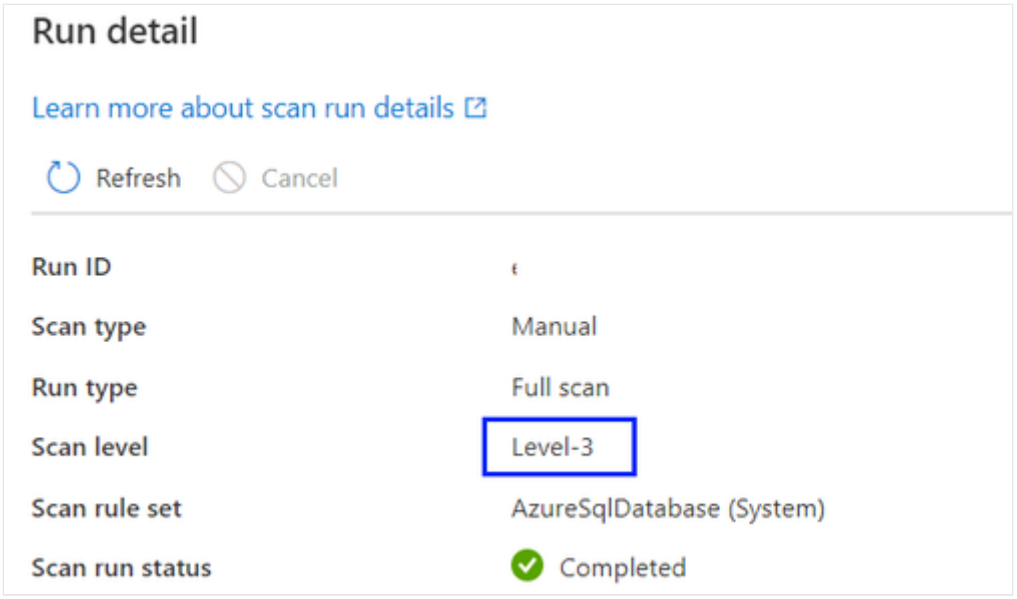
In Data Map terminology, three different levels of scanning exist based on the metadata scope and functionalities:

- L1 scan: Extracts basic information and metadata like file name, size, and fully qualified name
- L2 scan: Extracts schema for structured file types and database tables
- L3 scan: Extracts schema where applicable and subjects the sampled file to the system and custom classification rules

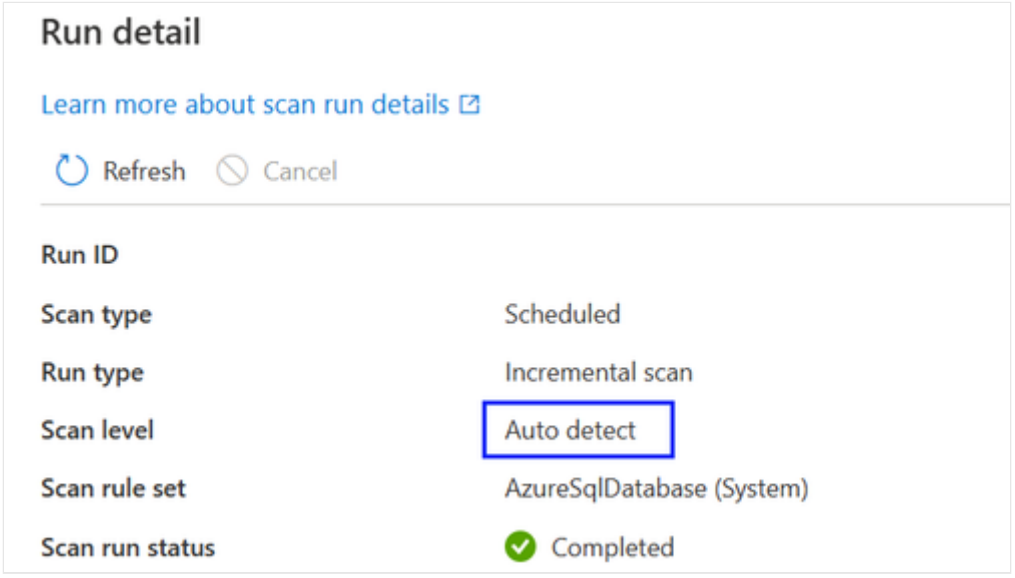
When you set up a new scan or edit an existing scan, you can customize the scan level for scanning data sources that support the scan level configuration.



By default, the "Auto detect" will be selected, which means Microsoft Purview applies the highest scan level available for this data source. Take Azure SQL Database as an example, the "Auto detect" will be resolved as “Level-3” when the scan is executed as the data source has already supported classification in Microsoft Purview. The scan level in the scan run detail shows the actual level applied.



For all scan runs in the scan history that completed before customizing scan level as a new feature, the scan level is set and displayed as **Auto detect**.



- When a higher scan level becomes available for a data source, the saved or scheduled scans that have scan level set to **Auto detect** automatically apply the new scan level. For example, if classification as a new feature is enabled for a given data source, all existing scans on this data source automatically apply classification.
- The scan level setting shows in the scan monitoring interface for each scan run.
- If you select **Level-1**, scanning returns only basic technical metadata like asset name, asset size, modified timestamp, and more, based on the existing metadata availability of a specific data source. For SQL Database, the process creates asset entities like tables in Data Map but doesn't extract the table schema. (Note: users can still see the table schema via live view if they have necessary permissions in the source system).
- If you select **Level-2**, scanning returns table schemas and basic technical metadata, but it doesn't perform data sampling and classification. For Azure SQL Database, table asset

entities capture the table schema without classification information.

- If you select **Level-3**, scanning performs the data sampling and classification. This is a standard configuration for Azure SQL Database scanning before scan level as a new feature is introduced.
- If you set a scheduled scan to a lower scan level and later modify it to a higher scan level, the next scan run automatically performs a full scan and updates all existing data assets from the data source with metadata introduced by a higher scan level setting. For example, when you change a scheduled scan set with **Level-2** on an Azure SQL Database to **Level-3**, the next scan run is a full scan and updates all existing Azure SQL Database table and view assets with classification information. All scans thereafter resume as incremental scans set with **Level-3**.
- If you set a scheduled scan to a higher scan level and later modify it to a lower scan level, the next scan run continues to perform an incremental scan and all new data assets from the data source only have metadata introduced by a lower scan level setting. For example, when you change a scheduled scan set with **Level-3** on an Azure SQL Database to **Level-2**, the next scan run is an incremental scan and all new Azure SQL Database table and view assets added in Data Map have no classification information. All existing data assets still keep the classification information generated from the previous scan set with **Level-3**.

> ⓘ **Note**
>
> - Customizing scan level is currently available for the following data sources: Azure SQL Database, Azure SQL Managed Instance, Azure Cosmos DB for NoSQL, Azure Database for PostgreSQL, Azure Database for MySQL, Azure Data Lake Storage Gen2, Azure Blob Storage, Azure Files, Azure Synapse Analytics, Azure Dedicated SQL pool (formerly SQL DW), Azure Data Explorer, Dataverse, Azure Multiple (Azure Subscription), Azure Multiple (Azure Resource Group), Snowflake, Azure Databricks Unity Catalog
> - Currently the feature is only available on Azure integration runtime and Managed Virtual Network integration runtime v2.

Scan rule set

A scan rule set determines the kinds of information a scan looks for when it's running against one of your sources. Available rules depend on the kind of source you're scanning, but include things like the file types you should scan, and the kinds of classifications you need.

Many data source types already have system scan rule sets, but you can also create your own scan rule sets to tailor your scans to your organization.

Schedule your scan

Microsoft Purview gives you a choice of scanning daily, weekly, or monthly at a specific time you choose. Learn more about the supported schedule options. Daily or weekly scans might be appropriate for data sources with structures that are actively under development or frequently change. Monthly scanning is more appropriate for data sources that change infrequently. Work with the administrator of the source you want to scan to identify a time when compute demands on the source are low.

How scans detect deleted assets

A Microsoft Purview catalog is only aware of the state of a data store when it runs a scan. For the catalog to know if a file, table, or container is deleted, it compares the last scan output against the current scan output. For example, suppose that the last time you scanned an Azure Data Lake Storage Gen2 account, it included a folder named *folder1*. When the same account is scanned again, *folder1* is missing. Therefore, the catalog assumes the folder is deleted.

> **Tip**
>
> Because of how deleted files are detected, you might need multiple successful scans to detect and resolve deleted assets. If Unified Catalog isn't registering deletions for a scoped scan, try multiple full scans to resolve the issue.

Detecting deleted files

The logic for detecting missing files works for multiple scans by the same user and by different users. For example, suppose a user runs a one-time scan on a Data Lake Storage Gen2 data store on folders A, B, and C. Later, a different user in the same account runs a different one-time scan on folders C, D, and E of the same data store. Because folder C was scanned twice, the catalog checks it for possible deletions. Folders A, B, D, and E, however, were scanned only once, and the catalog doesn't check them for deleted assets.

To keep deleted files out of your catalog, it's important to run regular scans. The scan interval is important, because the catalog can't detect deleted assets until another scan is run. So, if you run scans once a month on a particular store, the catalog can't detect any deleted data assets in that store until you run the next scan a month later.

When you enumerate large data stores like Data Lake Storage Gen2, there are multiple ways (including enumeration errors and dropped events) to miss information. A particular scan might miss that a file was created or deleted. So, unless the catalog is certain a file is deleted, it doesn't delete it from the catalog. This strategy means there can be errors when a file that doesn't exist in the scanned data store still exists in the catalog. In some cases, a data store might need to be scanned two or three times before it catches certain deleted assets.

> ⓘ **Note**
>
> - Assets that are marked for deletion are deleted after a successful scan. Deleted assets might continue to be visible in your catalog for some time before they're processed and removed.
> - Deletion detection is supported only for these sources within Microsoft Purview: Azure Synapse Analytics workspaces, Azure Arc-enabled SQL Server, Azure Blob Storage, Azure Files, Azure Cosmos DB, Azure Data explorer, Azure Database for MySQL, Azure Database for PostgreSQL, Azure Dedicated SQL pool, Azure Machine Learning, Azure SQL Database, and Azure SQL Managed instance. For these sources, when an asset is deleted from the data source, subsequent scans will automatically remove the corresponding metadata and lineage in Microsoft Purview.

Ingestion

Ingestion is the process that populates Data Map with metadata gathered through its various processes.

> ⓘ **Note**
>
> The combined count of all child objects (referred entities) and contacts (owner, expert) must not exceed 20,000 entities.

Ingestion from scans

The scanning process identifies technical metadata or classifications and sends them to ingestion. Ingestion analyzes the input from the scan, applies resource set patterns, populates available lineage information, and then automatically loads Data Map. You can discover or curate assets and schemas only after ingestion is complete. If your scan completes but you

don't see your assets in the Data Map or catalog, you need to wait for the ingestion process to finish.

Ingestion from lineage connections

You can connect resources like Azure Data Factory and Azure Synapse to Microsoft Purview to bring data source and lineage information into Data Map. For example, when a copy pipeline runs in an Azure Data Factory that you connected to Microsoft Purview, the service ingests metadata about the input sources, the activity, and the output sources. The information is added to Data Map.

If you add a data source to Data Map through a scan, the lineage information about the activity is added to the existing source. If you haven't yet added the data source to Data Map, the lineage ingestion process adds it to the root collection with its lineage information.

For more information about the available lineage connections, see the lineage user guide.

Next steps

For more information, or for specific instructions for scanning sources, follow the links below.

- Understanding resource sets
- How to govern an Azure SQL Database
- Lineage in Microsoft Purview

Last updated on 12/16/2025

Create a scan rule set in Data Map

In Microsoft Purview Data Map, you can create scan rule sets to customize data source scans in your organization.

A scan rule set is a container that groups a set of scan rules together. You can easily associate a scan rule set with a scan. For example, you might create a default scan rule set for each of your data source types, and then use these scan rule sets by default for all scans within your company. You might also want users with the right permissions to create other scan rule sets with different configurations based on business need.

Steps to create a scan rule set

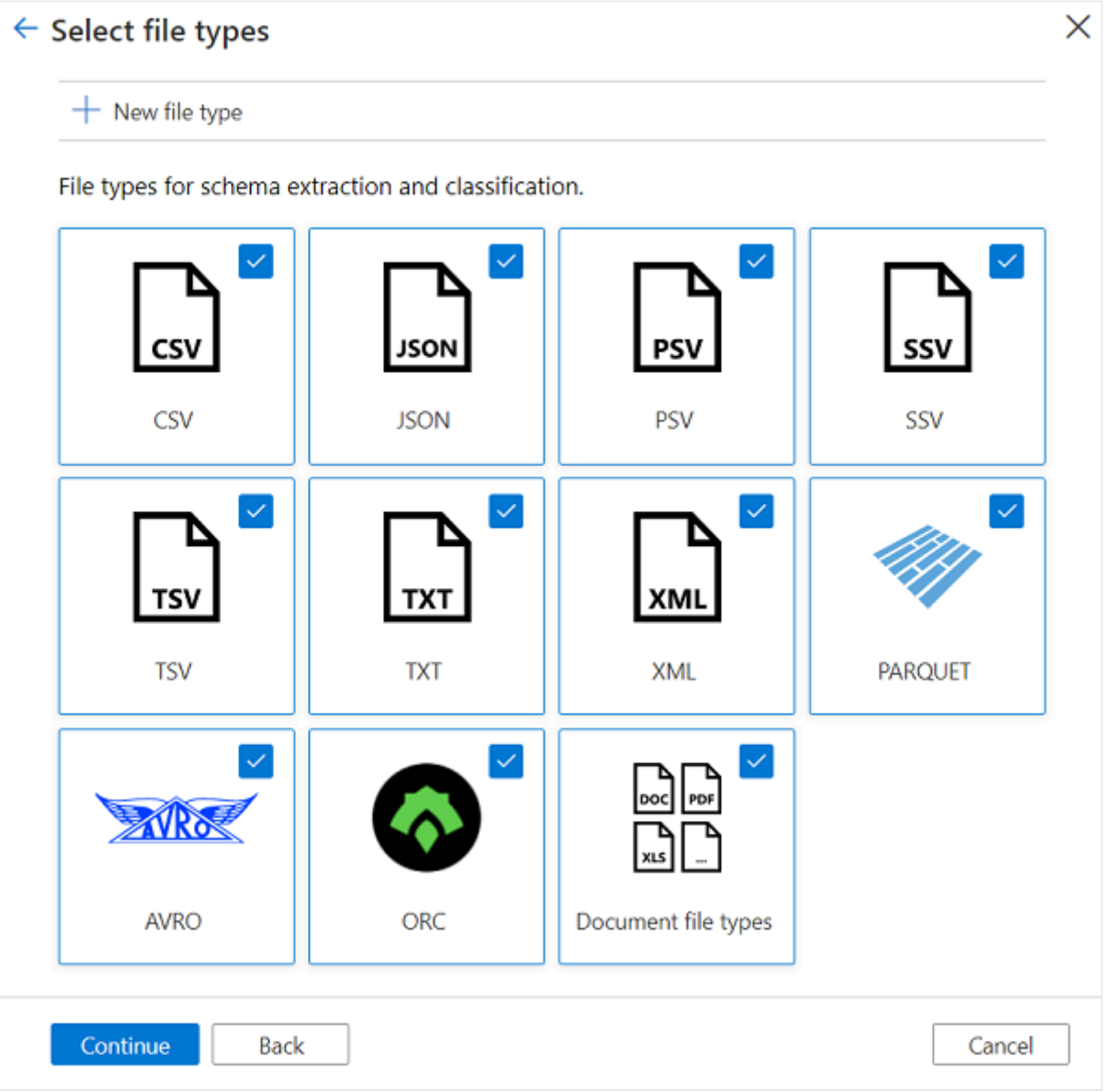
To create a scan rule set:

1. In the Microsoft Purview portal, select the **Data Map** solution.
2. Under the **Source management** section, select **Scan rule sets**, and then select **New**.
3. From the **New scan rule set** page, select the data sources that the catalog scanner supports from the **Source Type** dropdown list. You can create a scan rule set for each type of data source you intend to scan.
4. Enter a **Name** for the scan rule set. The maximum length is 63 characters, with no spaces allowed. Optionally, enter a **Description**. The maximum length is 256 characters.
5. Select a domain.

   > ⓘ **Important**
   >
   > You can only use the scan rule set in the domain where you create it.

6. Select **Continue**.

   The **Select file types** page appears. Notice that the file type options on this page vary based on the data source type that you chose on the previous page. All the file types are enabled by default.



The **Document file types** selection on this page allows you to include or exclude the following Microsoft Office file types: `.doc; .docm; .docx; .dot; .odp; .ods; .odt; .pdf; .pot; .pps; .ppsx; .ppt; .pptm; .pptx; .xlc; .xls; .xlsb; .xlsm; .xlsx; xlt`.

7. Enable or disable a file type for schema and classification by selecting or clearing its check box. If you choose a Data Lake type data source (for example, Azure Data Lake Storage Gen2 or Azure Blob), enable the file types for which you want to have schema extracted and classified.

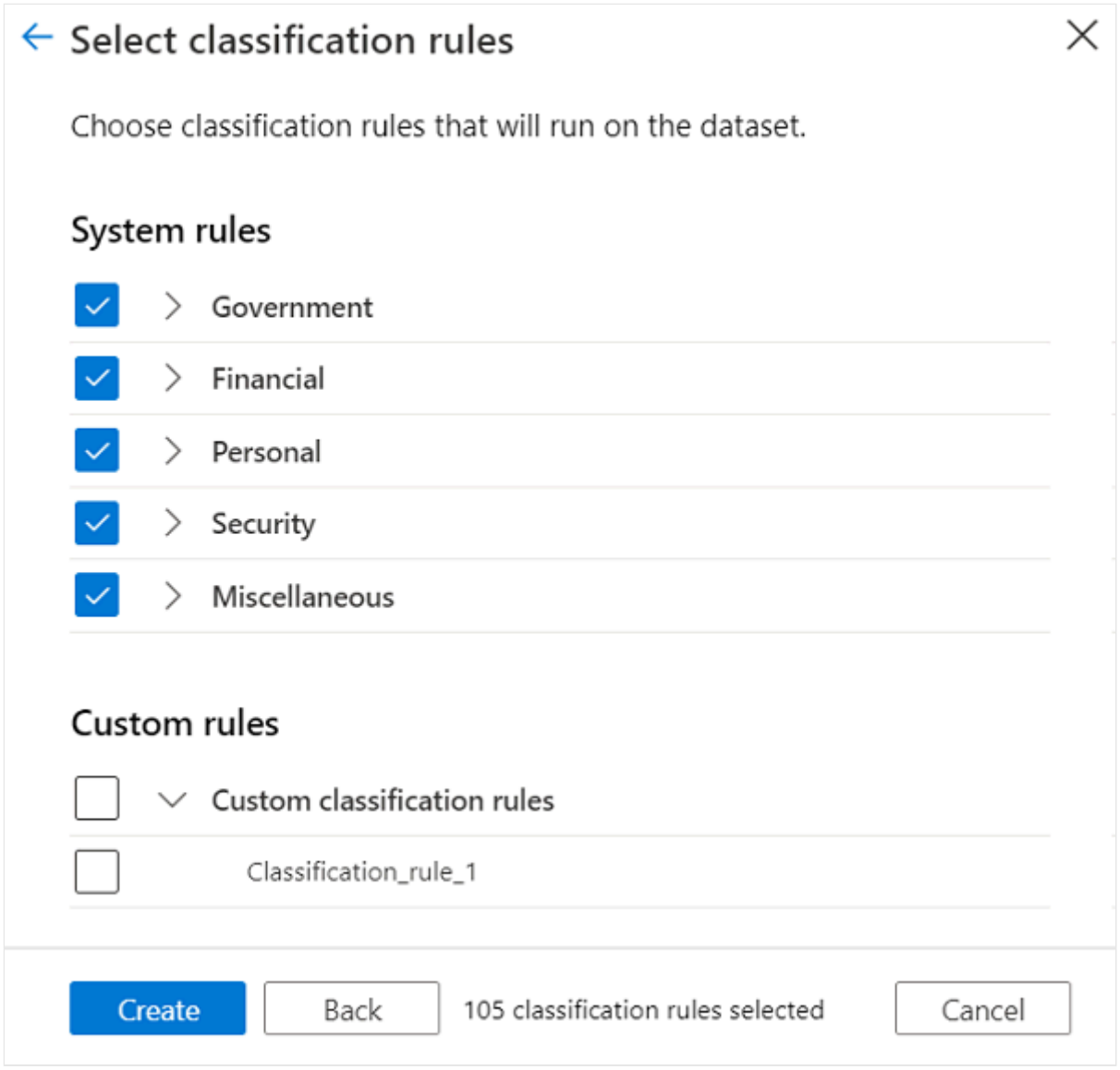
   > **ⓘ Important**
   >
   > Deselecting a file type **doesn't** prevent ingestion. The file's information is still ingested, but its schema and classifications aren't extracted.

8. For certain data source types, you can also Create a custom file type.

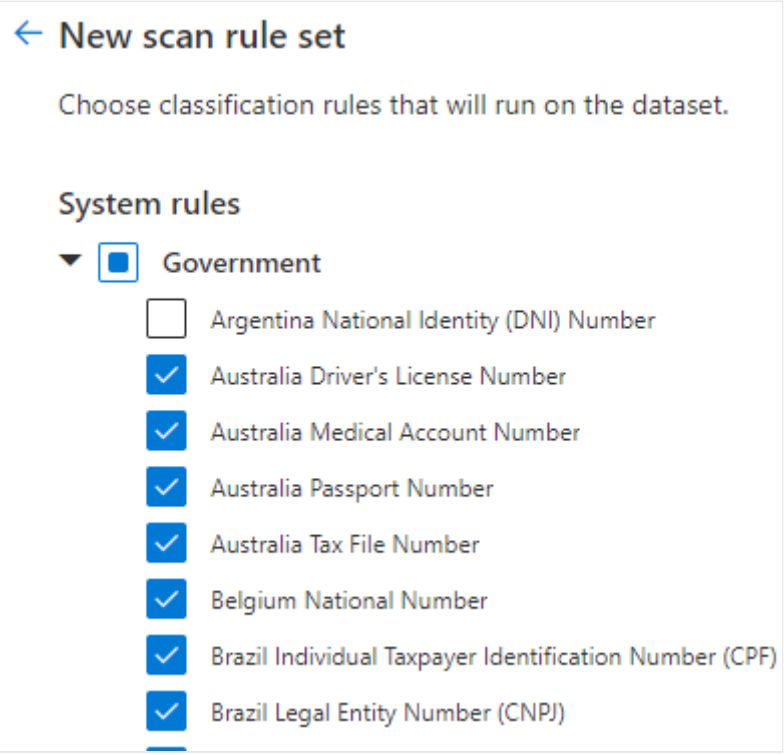
9. Select **Continue**.

The **Select classification rules** page appears. This page displays the selected **System rules** and **Custom rules**, and the total number of classification rules selected. By default, all the **System rules** check boxes are selected.

10. For the rules you want to include or exclude, select or clear the **System rules** classification rule checkboxes globally by category.



11. You can expand the category node and select or clear individual check boxes. For example, if the rule for **Argentina.DNI Number** has high false positives, you can clear that specific check box.



12. Select **Create** to finish creating the scan rule set.

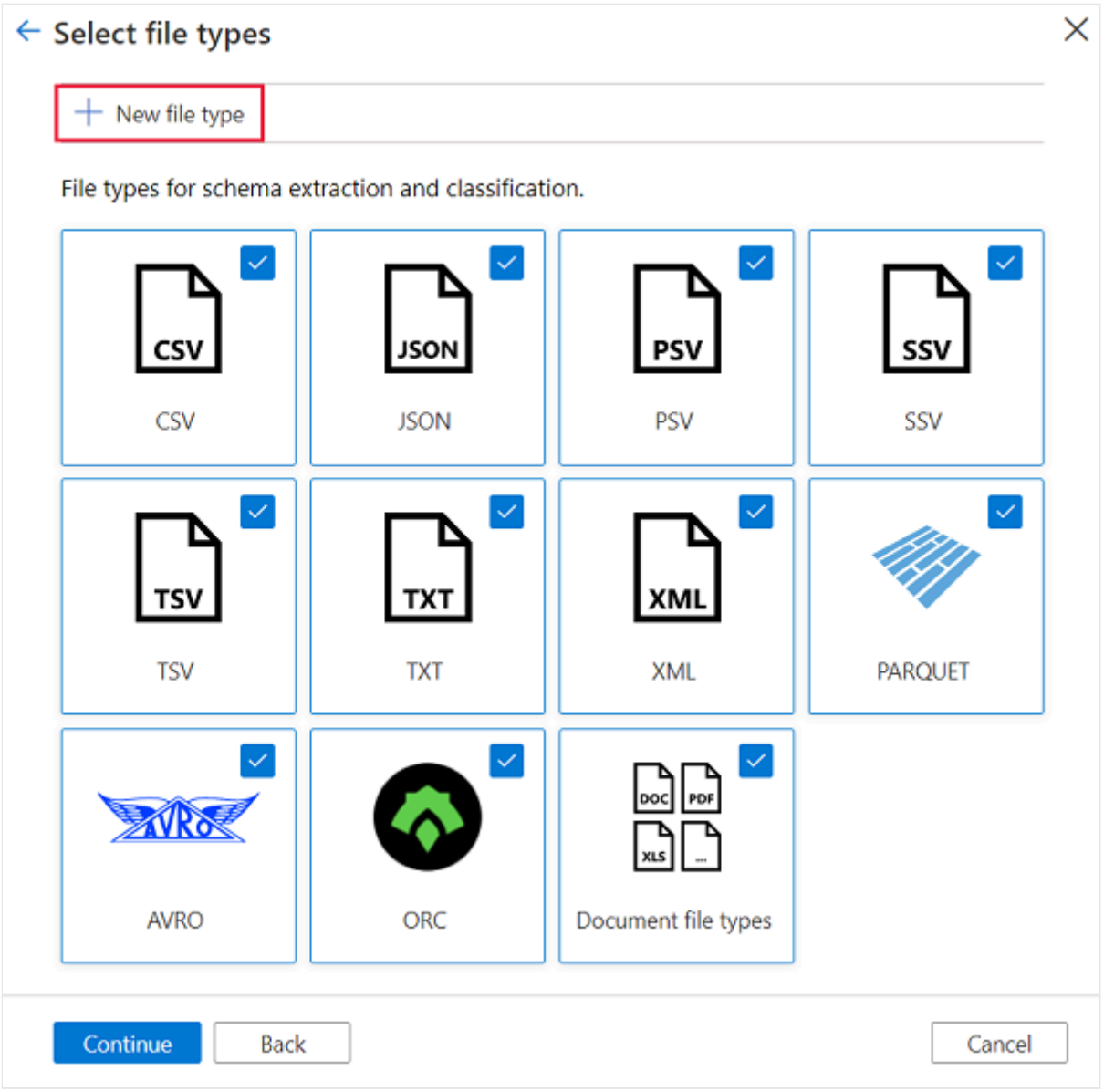
Create a custom file type

Microsoft Purview supports adding a custom extension and defining a custom column delimiter in a scan rule set.

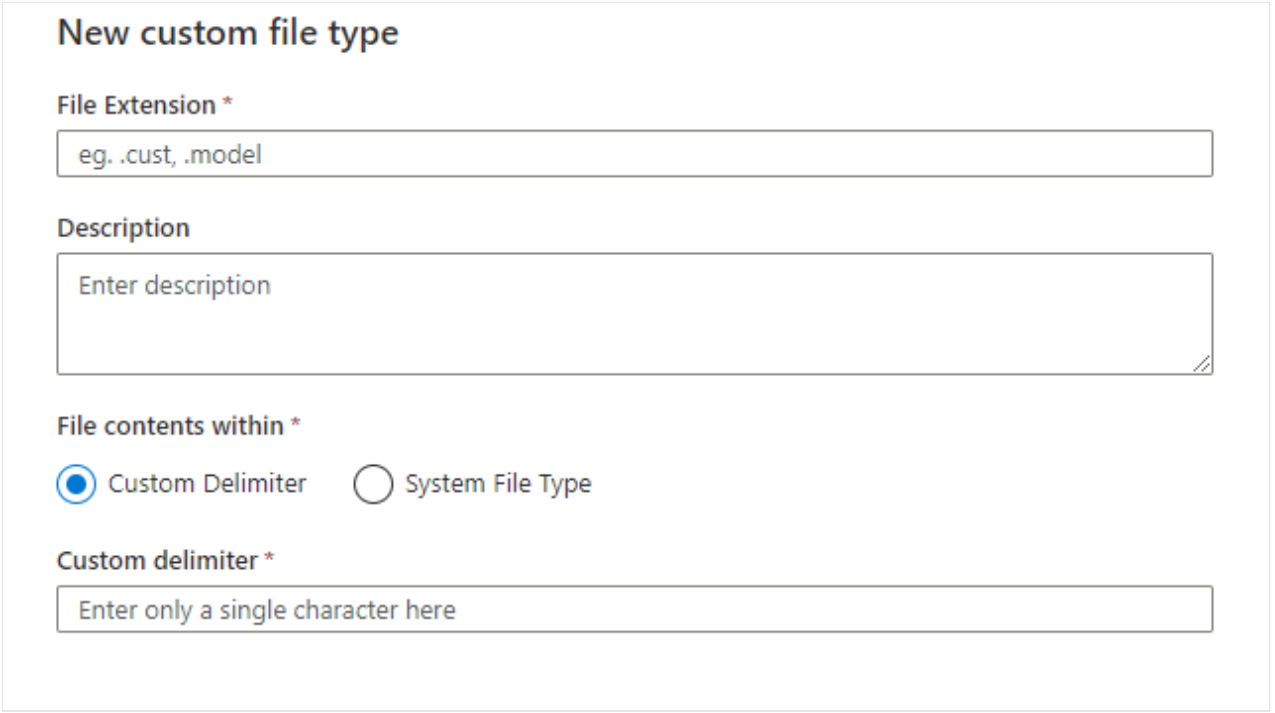
To create a custom file type:

1. Follow steps 1–5 in Steps to create a scan rule set or edit an existing scan rule set.

2. On the **Select file types** page, select **New file type** to create a new custom file type.



3. Enter a **File Extension** and an optional **Description**.

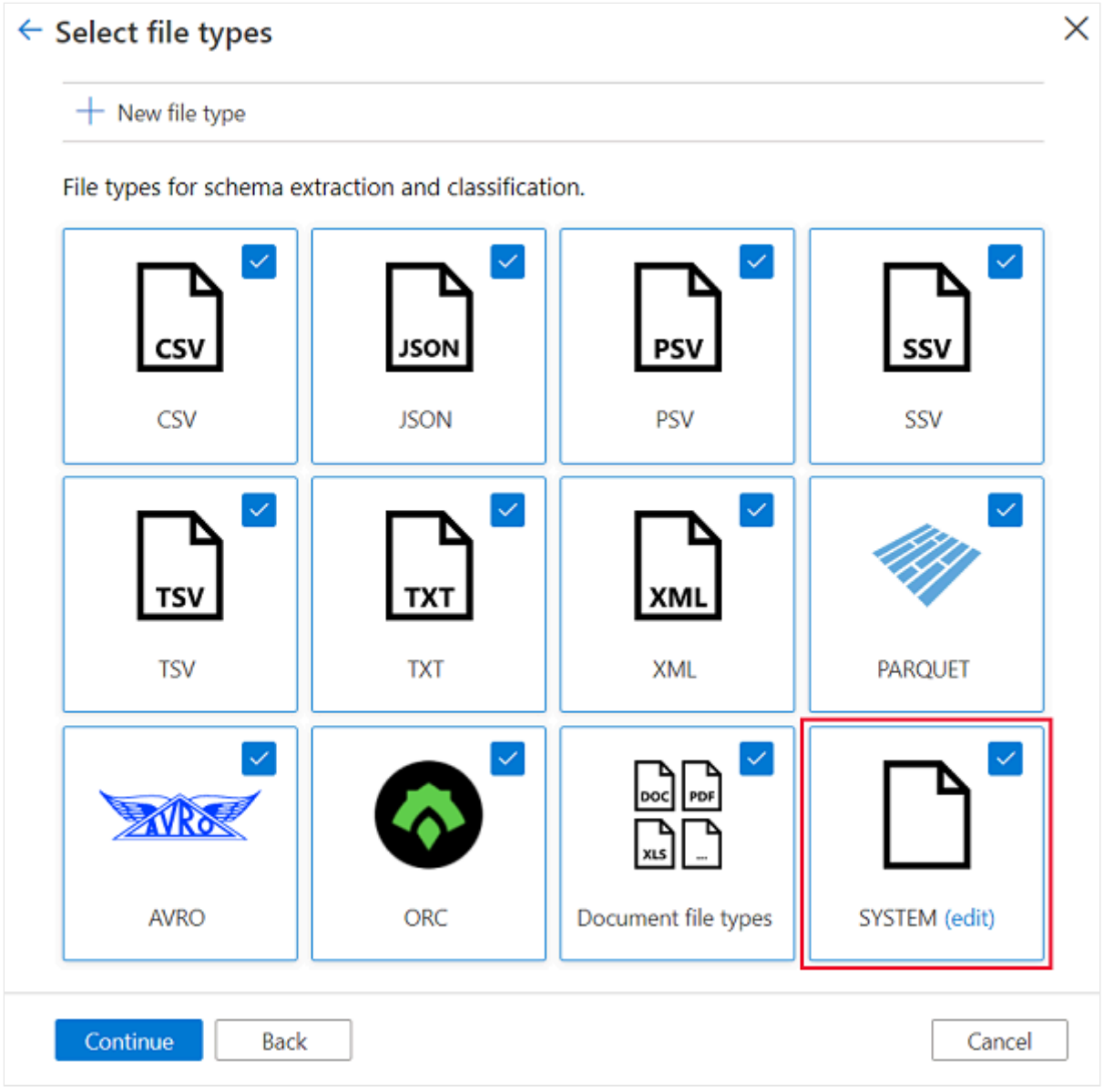


4. Make one of the following selections for **File contents within** to specify the type of file contents within your file:

    - Select **Custom Delimiter** and enter your own **Custom delimiter** (single character only).
    - Select **System File Type** and choose a system file type (for example XML) from the **System file type** drop-down list.

5. Select **Create** to save the custom file.

    The system returns to the **Select file types** page and inserts the new custom file type as a new tile.



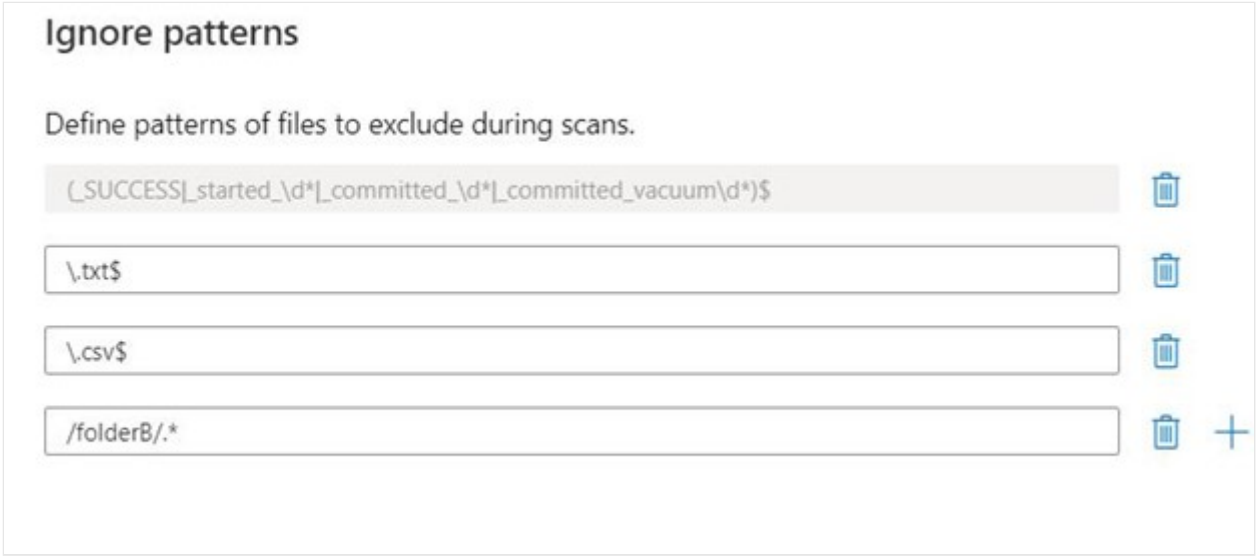
6. Select **Edit** in the new file type tile if you want to change or delete it.

7. Select **Continue** to finish configuring the scan rule set.

Ignore patterns

Microsoft Purview supports defining regular expressions (regex) to exclude assets during scanning. During scanning, Purview compares the asset's URL against these regular expressions. Microsoft Purview ignores all assets that match any of the regexes.

The **Ignore patterns** pane prepopulates one regex for spark transaction files. You can remove the preexisting pattern if it isn't required. You can define up to 10 ignore patterns.

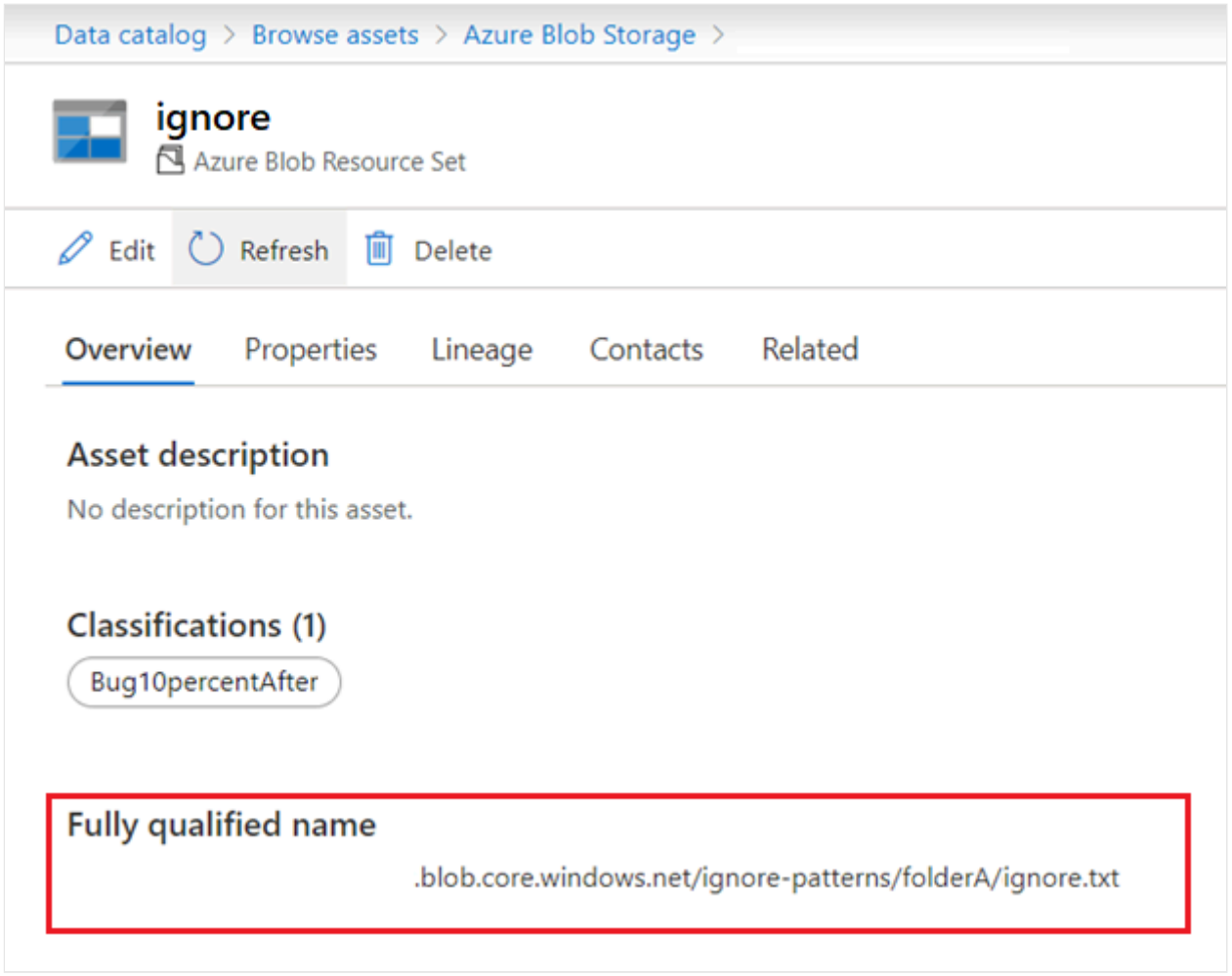


In the preceding example:

- Regexes 2 and 3 ignore all files ending with .txt and .csv during scanning.
- Regex 4 ignores /folderB/ and all its contents during scanning.

Here are some more tips you can use to ignore patterns:

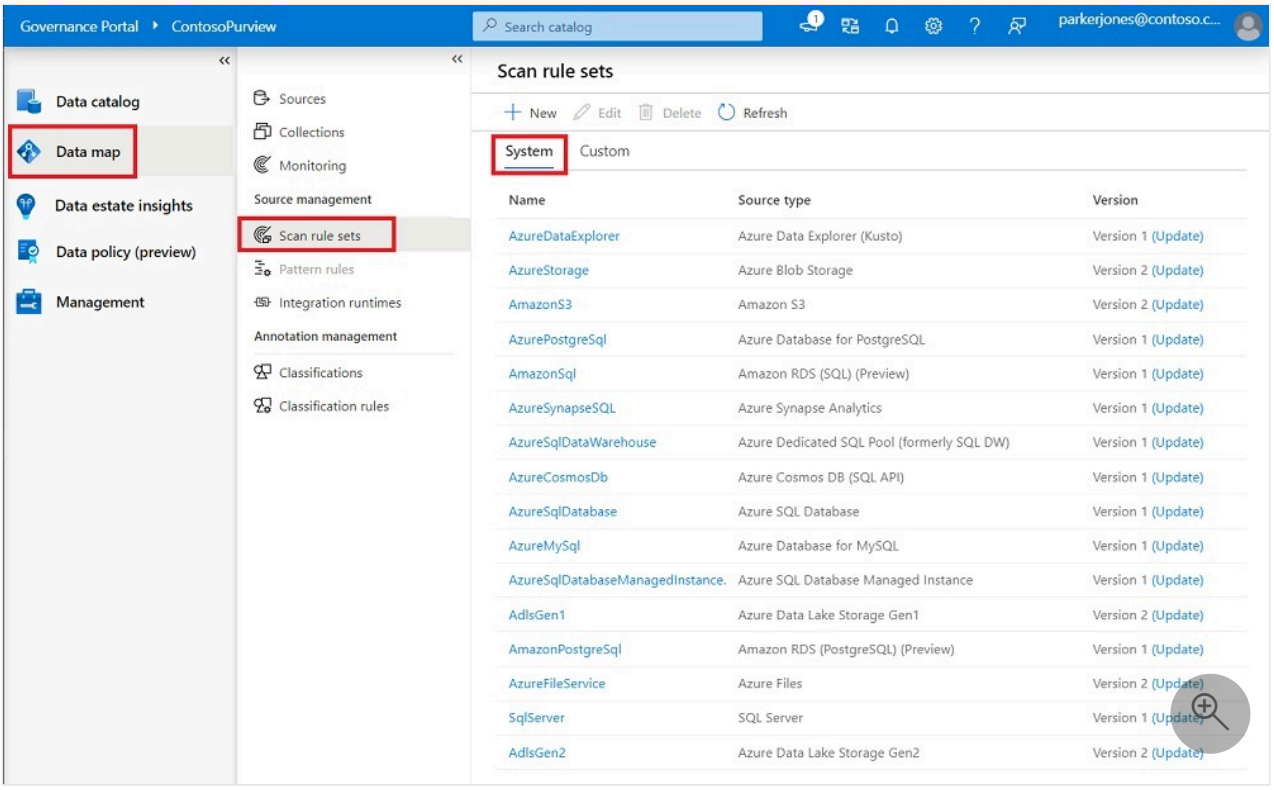
- While processing the regex, Microsoft Purview adds $ to the regex by default.
- To understand what URL the scanning agent compares with your regular expression, browse through Microsoft Purview Unified Catalog, find the asset you want to ignore in the future, and see its fully qualified name (FQN) in the **Overview** tab.



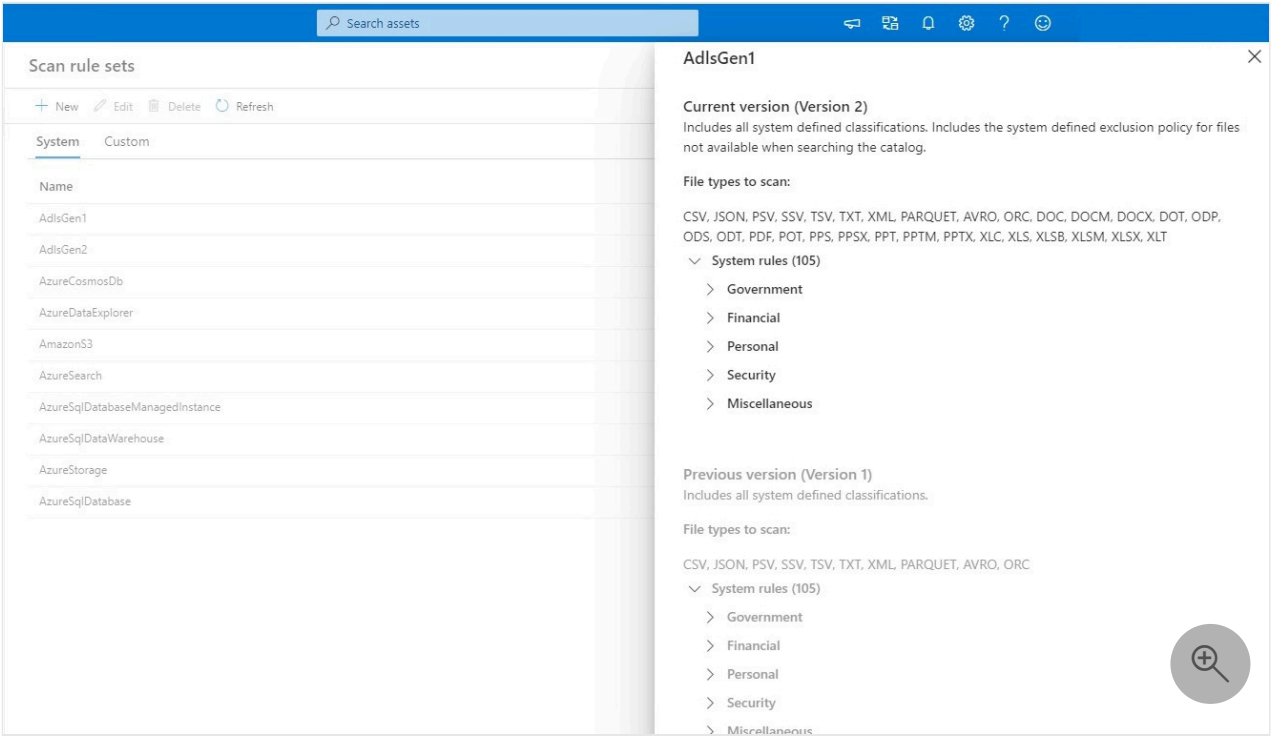
System scan rule sets

System scan rule sets are Microsoft-defined scan rule sets that are automatically created for each Microsoft Purview catalog. Each system scan rule set is associated with a specific data source type. When you create a scan, you can associate it with a system scan rule set. Every time Microsoft makes an update to these system rule sets, you can update them in your catalog, and apply the update to all the associated scans.

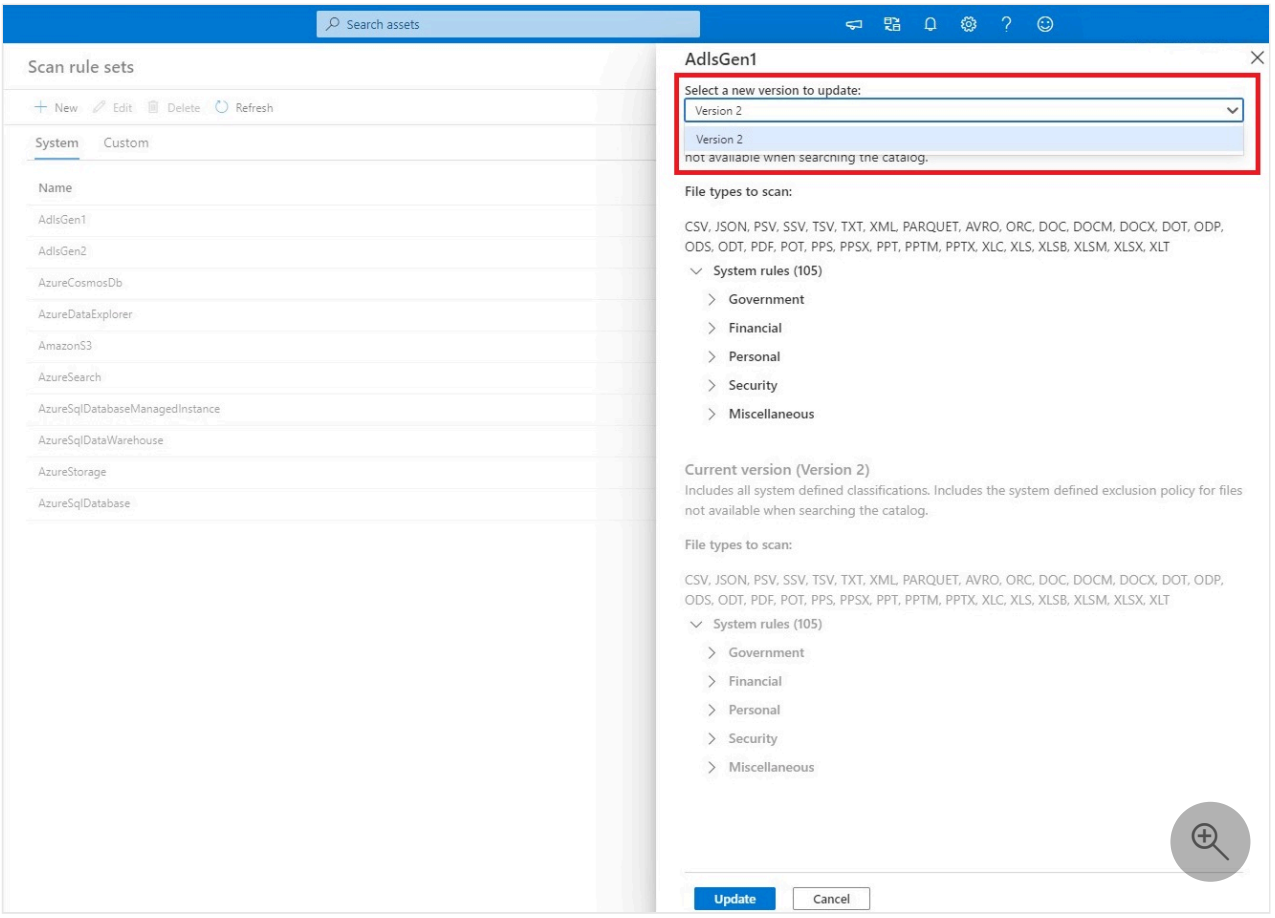
1. To view the list of system scan rule sets, select **Scan rule sets** in the **Management Center** and choose the **System** tab.



2. Each system scan rule set has a **Name**, **Source type**, and a **Version**. If you select the version number of a scan rule set in the **Version** column, you see the rules associated with the current version and the previous versions (if any).



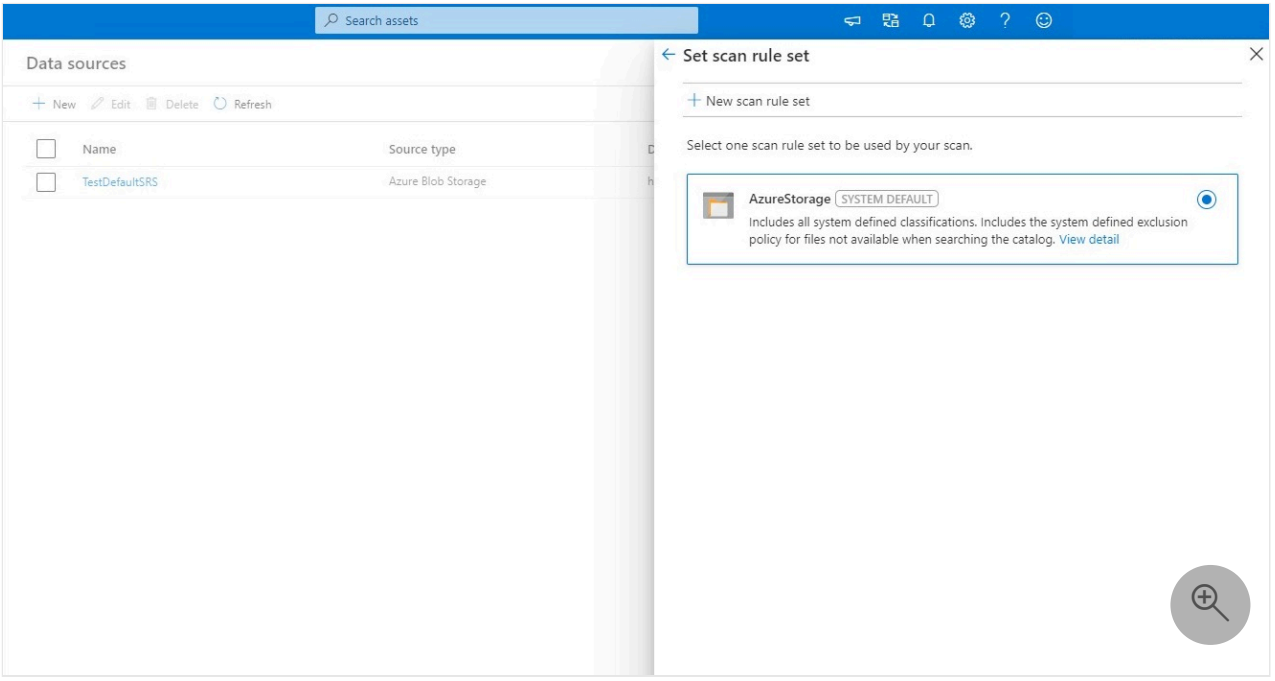
3. If an update is available for a system scan rule set, you can select **Update** in the **Version** column. In the system scan rule page, choose a version from the **Select a new version to update** drop-down list. The page provides a list of system classification rules associated with the new version and current version.



Associate a scan with a system scan rule set

When you create a scan, choose to associate it with a system scan rule set as follows:

1. On the **Select a scan rule set** page, select the system scan rule set.



2. Select **Continue**, and then select **Save and Run**.

Last updated on 12/09/2025

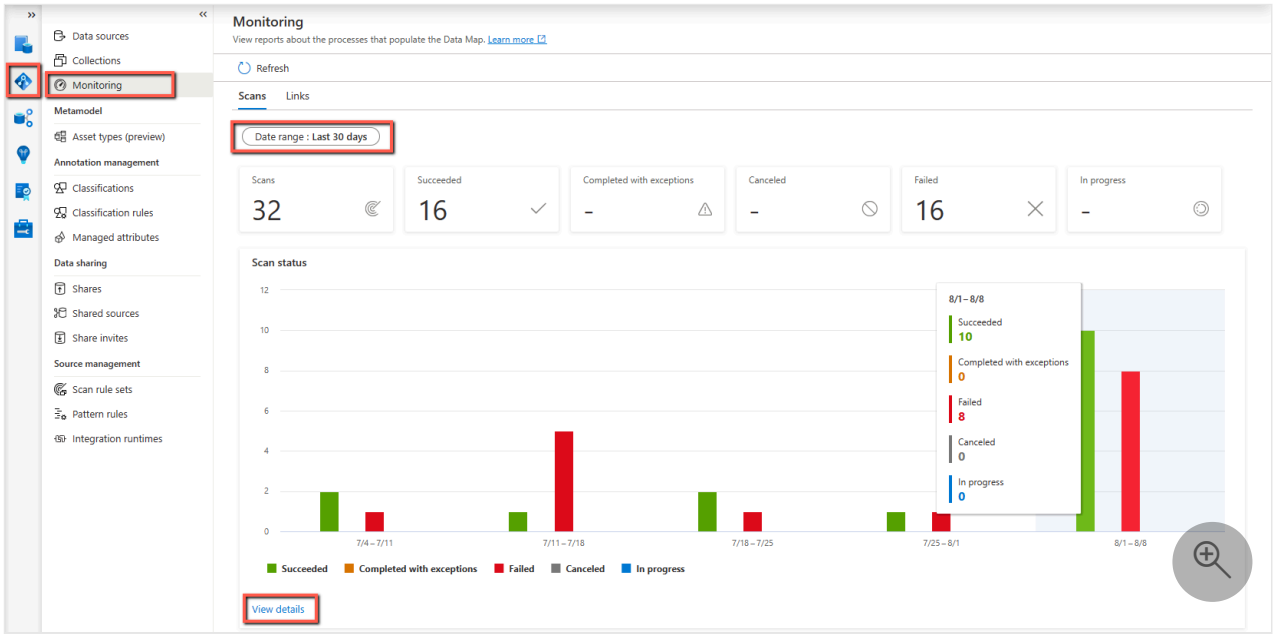
Monitor Data Map population in Microsoft Purview

In Microsoft Purview, you can scan various types of data sources and view the scan status over time. You can also connect other services with Microsoft Purview and view the trends of the ingested assets and relationships. This article outlines how to monitor and get a bird's eye view of the Purview Data Map population.

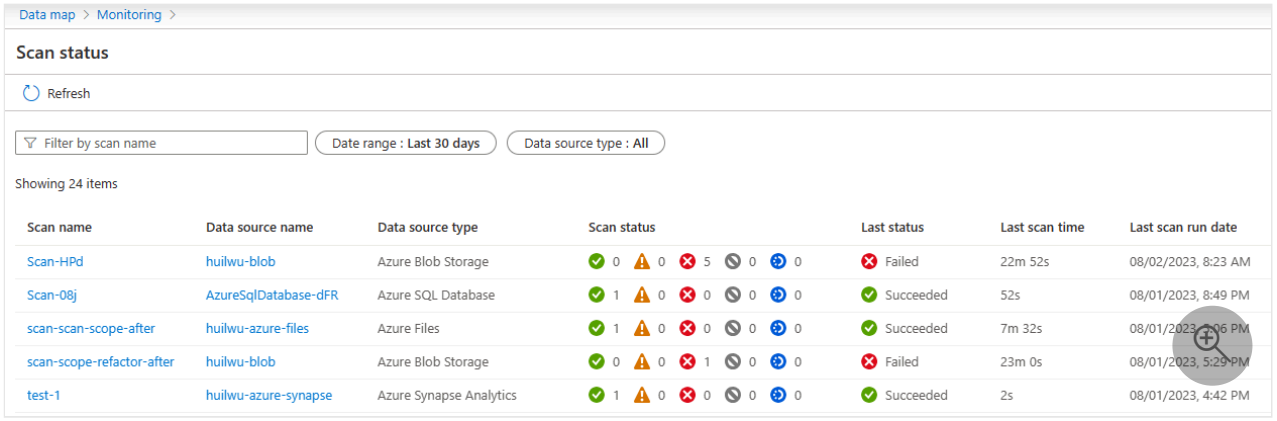
Monitor scan runs

1. Open your Microsoft Purview account and select **Data Map** -> **Monitoring**. You need to have the **Data source admin** role on any collection to access this page. You can see the scan runs that belong to the collections on which you have data source admin privilege.

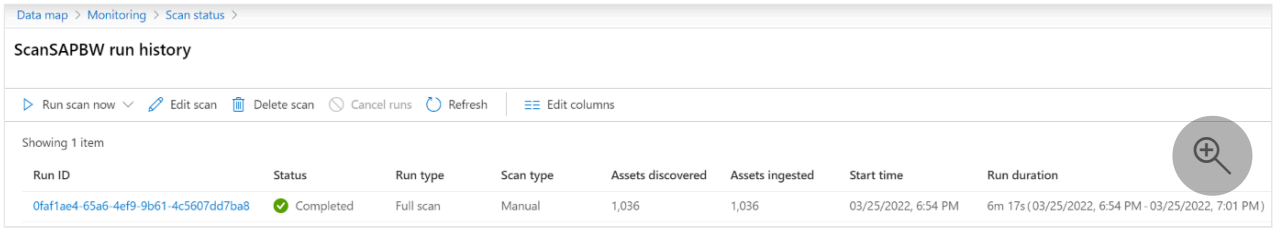
2. The high-level KPIs show total scan runs within a period. The default time period is the last 30 days. You can also choose to select the last seven days. Based on the time filter you select, you can see the distribution of succeeded, completed with exceptions, failed, canceled, and in progress scan runs by week or by the day in the graph. It's the scan run end-to-end status including the discovery and ingestion phases. For more information about the status details, see the scan run details section.



3. At the bottom of the graph, there's a **View more** link for you to explore further. The link opens the **Scan status** page. Here you can see a scan name and the number of times it has status of succeeded, completed with exceptions, failed, or canceled in the time period. You can also filter the list by source types.



4. You can explore a specific scan further by selecting the **scan name**. It connects you to the scan history page, where you can find the list of scan runs with more execution details.



5. Select **run ID** to check more about the scan run details.

Known limitations

Currently, this monitoring experience doesn't include the following information. You can check the corresponding scan run details.

- Lineage extraction scan runs for Azure SQL Database aren't covered.
- For scans of Azure Synapse Analytics workspace, Azure subscription/resource group, and Amazon Web Services (AWS) account, it only captures the parent scan run's status/count but not child scan runs' per resource.

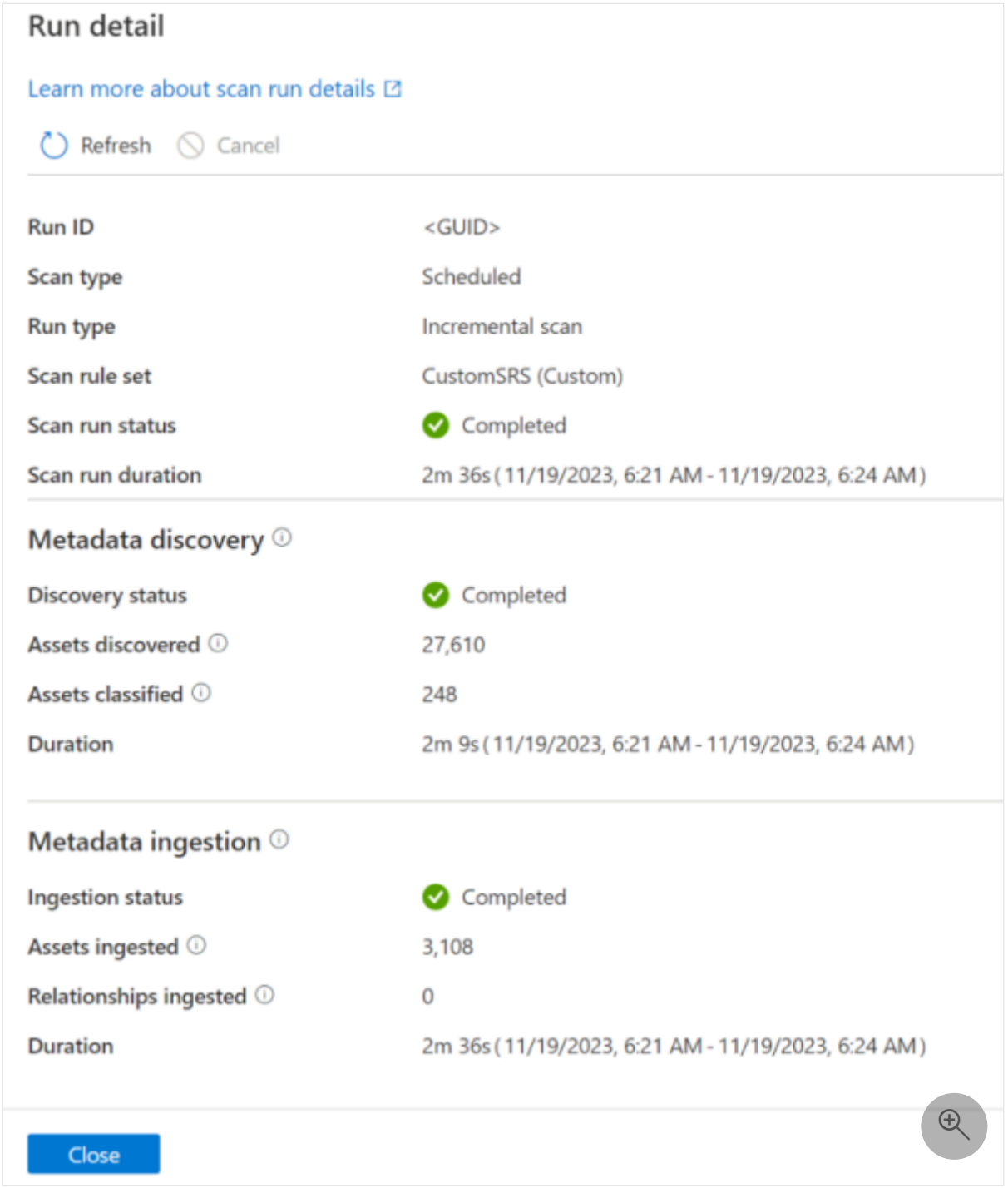
Scan run details

You can view the scan run history for a given scan from different places:

- Go to **Data Map** -> **Monitoring** as described in the Monitor scan runs section.
- Go to **Data Map** -> **Sources** -> select the desired data source -> see **Scans**, **Recent scans**, or **Recent failed scans**.
- Go to **Data Map** -> **Collections** -> select the desired collection -> **Scans** -> select the scan name that you want to view.

The retention of the scan run history is 90 days.

Select **run ID** to check more about the scan run details:



- The section on top provides an overview of the scan run, including the following information:
  - **Run ID**: A GUID used to identify this particular scan run.
  - **Scan type**: Manual or scheduled scan.
  - **Run type**: Full or incremental scan.
  - **Scan rule set**: The scan rule set used in the run, if applicable.
  - **Scan run status**: Overall scan status. It combines the metadata discovery and ingestion phases' status.
  - **Scan run duration**: The end-to-end execution duration and the start/end time.

- **Metadata discovery** section summarizes the metrics for discovery phase that Purview connects to the source, extracts the metadata/lineage and classifies the data.

- **Discovery status**:

Expand table

Status	Description
Completed	The metadata discovery succeeds.
Completed with exceptions	The metadata discovery completes, while some underlying operation fails that might cause asset, metadata, or classification missing in the scan result. You can view details in the exception logs.
Failed	The metadata discovery fails. You can check the error details by selecting the "More info" link next to it.
Canceled	The scan run is canceled by user.
In Progress	The metadata discovery is running in progress.
Queued	The metadata discovery is waiting for available integration runtime resource. If you use self-hosted integration runtime, note each node can run many concurrent scans at the same time depending on your machine specification (CPU and memory). More scans are in Queued status.
Throttled	The metadata discovery is being throttled. It means this Microsoft Purview account at the moment has more ongoing scan runs than the allowed max concurrent count. Learn more about the limit here. This particular scan run is waiting and will start once your other ongoing scans finish.

  The scan run isn't charged during "Throttled" or "Queued" status.
  - **Assets discovered**: The number of assets enumerated from the source. For both full and incremental scans, it includes all assets in the configured scope, regardless of whether they're existing assets or newly created or updated assets since the last scan run. For incremental scan, Microsoft Purview extracts detailed metadata only for newly created or updated assets.
  - **Assets classified**: The number of assets sampled to classify the data, regardless of whether the assets have any matching classification or not. It's a subset of the discovered assets based on the sampling mechanism. For incremental scan, only newly created or updated assets might be selected for classification.
  - **Duration**: The discovery phase duration and the start/end time.
- **Metadata ingestion** section summarizes the metrics for ingestion phase that populates Data Map with the identified metadata and relationship.
  - **Ingestion status**:

Expand table

Status	Description
Completed	All of the assets and relationships are ingested into Data Map successfully.
Completed with exceptions (previous "Partially completed")	Partial of the assets and relationships are ingested into Data Map successfully, while some fail. You can view details in the exception logs.
Failed	The ingestion phase fails.
Canceled	The scan run is canceled by user, thus the ingestion is canceled along.
In Progress	The ingestion is running in progress.
Queued	The ingestion is waiting for available service resource or waiting for scan to discover metadata.

- **Assets ingested**: The number of assets ingested into Data Map. For incremental scan, it only includes the newly created or updated assets, in which case might be less than the "assets discovered" count. When scanning file-based source, it's the raw assets count before resource set aggregation.
- **Relationships ingested**: The number of relationships ingested into Data Map. It includes lineage and other relationships like foreign key relationships.
- **Duration**: The ingestion duration, and the start and end times.

View the exception log (preview)

> **Important**
>
> This feature is currently in preview. The Supplemental Terms of Use for Microsoft Azure Previews include additional legal terms that apply to Azure features that are in beta, in preview, or otherwise not yet released into general availability.

When some assets or relationships fail to be discovered from the source or fail to ingest into Data Map during a scan run, the scan status shows **completed with exceptions**. You can view the exception log by following these steps:

1. Select the scan.
2. Select the run ID of a scan with failures or exceptions.

3. Select **Download log** in the scan run details panel. It provides you with the exception log files that capture the details of the failures.

Discovery phase log

The following table shows the schema of a discovery phase log file.

Expand table

Column	Description
TimeStamp	The UTC timestamp when the discovery operation happens.
ErrorCode	Error code of the exception.
OperationName	It contains the following values: - **Enumerate**: Indicates the operation of enumerating the child objects under a given scope; for example, list the folders/files under a container, list the tables under a database. When there's a failure, the operation item's child objects aren't included in the scan result. - **GetMetadata**: Indicates the operation of extracting the metadata for an object, such as a table or file. When there's a failure, the asset of the operation item might not contain full metadata. - **SampleResourceSet**: Indicates the operation of sampling the latest file matching the resource set pattern when you scan file-based data sources. When there's a failure, the resource set that the operation item belongs to might miss classification or schema. - **ReadData**: Indicates the operation of retrieving sample data from the source. When there's a failure, the operation item isn't classified. If it's a file-based source, it might also miss schema. - **Classify**: Indicates the operation of applying the classifications. When there's a failure, the operation item might miss classification. - **GenerateAsset**: Indicates the operation of converting the discovered metadata into Purview asset. When there's a failure, usually due to system error, the asset of the operation item doesn't show up in Data Map at last.
OperationItem	Identifier for the failed asset or relationship, usually uses the fully qualified name.
Message	More information on which asset or relationship failed to discover and the reason for the failure.

Known limitations:

- The discovery phase log is supported when you use Azure integration runtime or Managed Virtual Network integration runtime for scan.
- The discovery phase log is supported when you scan Azure or Power BI data sources.
- Currently, it might not capture all the exceptions that happen during a scan run.

Ingestion phase log

The following table shows the schema of an ingestion phase log file.

Expand table

Column	Description
TimeStamp	The UTC timestamp when the ingestion operation happens.
ErrorCode	Error code of the exception.
OperationItem	Identifier for the failed asset or relationship, usually uses the fully qualified name.
Message	More information on which asset or relationship failed to ingest and the reason for the failure. If there's an ingestion failure for a resource set, it might apply to multiple assets that match the same naming pattern, and the message includes the impacted count.

Monitor links

You can connect other services with Microsoft Purview to establish a "link", which makes the metadata and lineage of that service's assets available to Microsoft Purview. Currently, link is supported for Azure Data Factory and Azure Synapse Analytics.

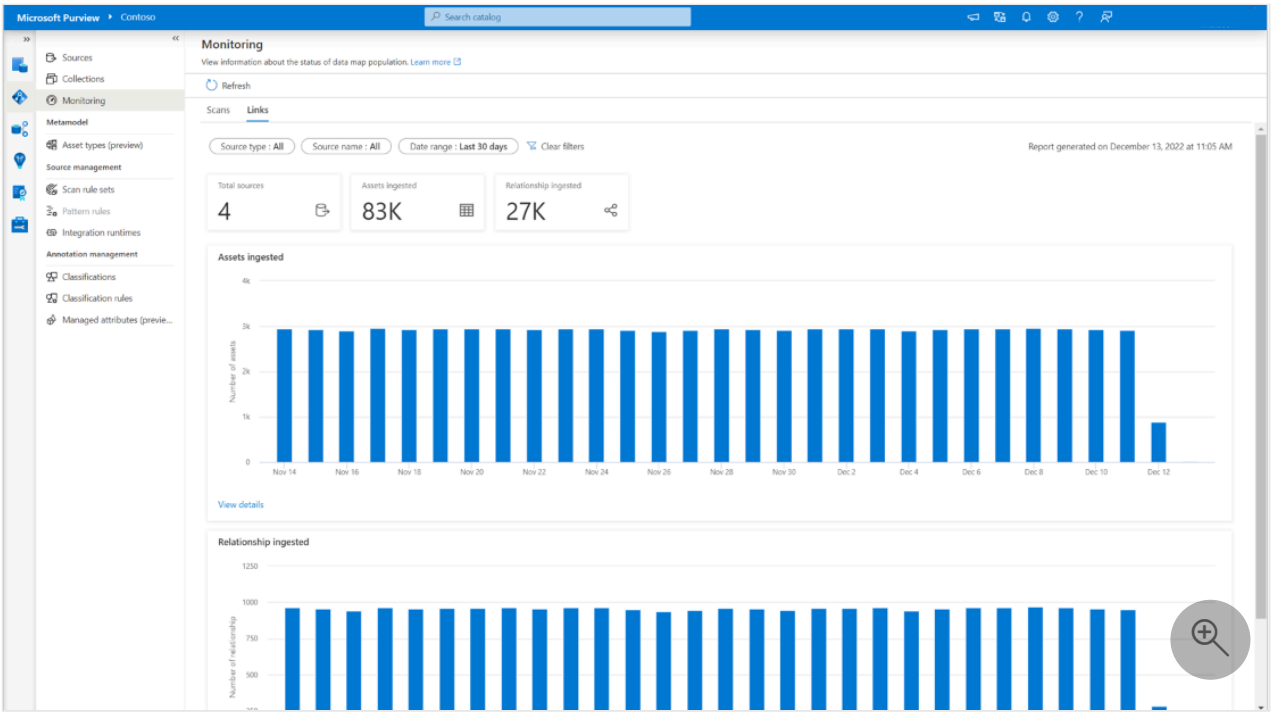
To monitor the assets and relationship ingested over the links:

1. Go to your Microsoft Purview account -> open **Microsoft Purview governance portal** -> **Data Map** -> **Monitoring** -> **Links**. You need to have **Data source admin** role on any collection to access the Monitoring tab. And you can see the results that belong to the collections on which you have data source admin privilege. Permission on root collection is needed to monitor Azure Data Factory and Azure Synapse Analytics links.

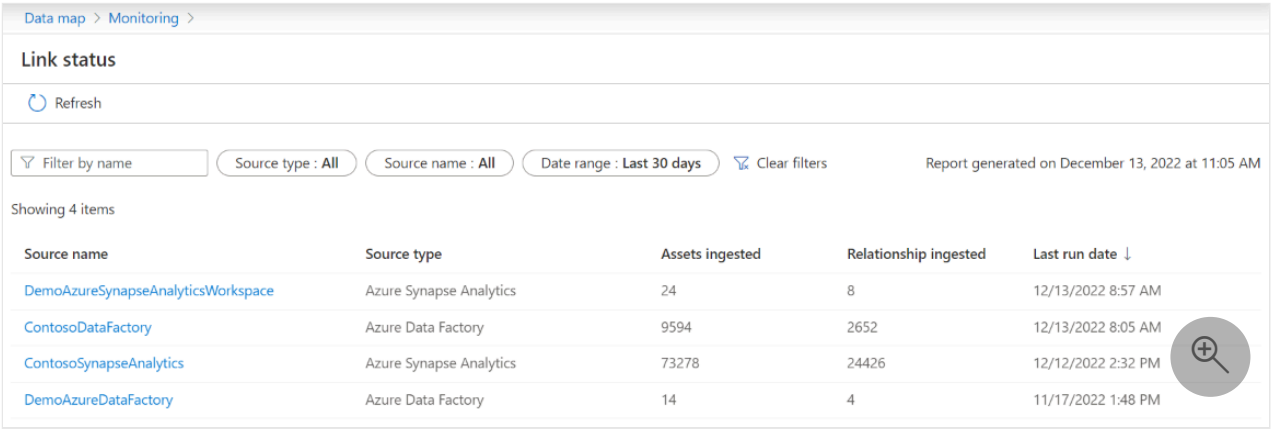
2. You can see the high-level KPIs including total number of sources, number of ingested assets and relationship (lineage), followed by trending charts over time. You can apply more filters on the following to narrow down the results:

   - Source type
   - Source name
   - Date range: Default is 30 days. You can also choose last seven days or a custom date range. The retention is 45 days.

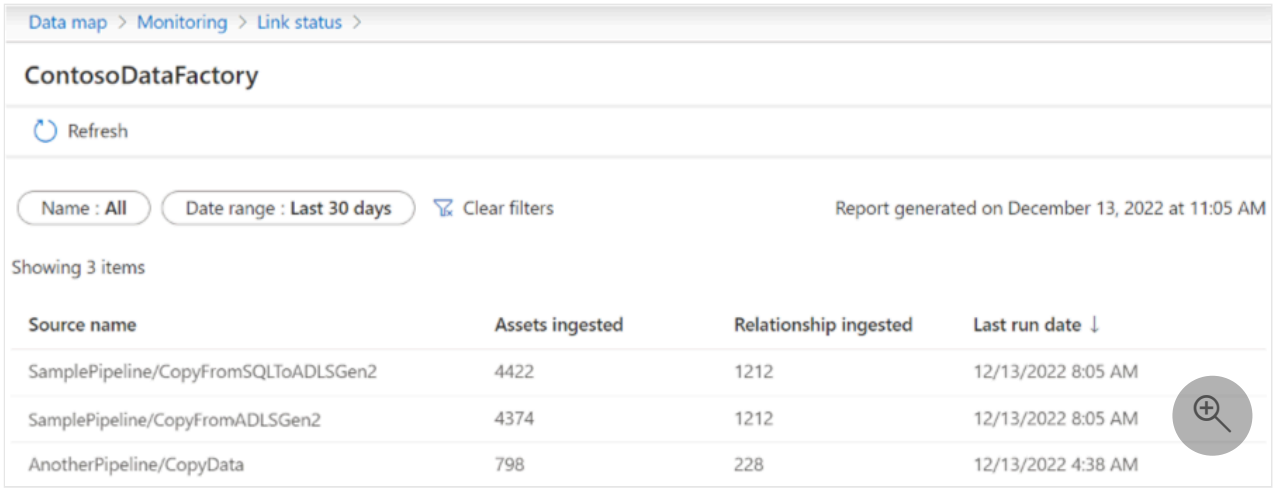
   The metrics are reported till the date time shown at the top right corner. And the aggregation happens hourly.



3. At the bottom of the graph, there's a **View more** link for you to explore further. In the **Link status** page, you can see a list of source names along with the source type, assets ingested, relationship ingested and the last run date time. The filters in the previous page are carried over, and you can further filter the list by source type, source name, and date range.



4. You can drill down to each source to see the next level details by clicking the source name. For example, for Azure Data Factory, it shows how each pipeline activity reports the assets and relationship to Microsoft Purview, with the name in the format of `<pipeline_name>/<activity_name>`.



Known limitations

- For Azure Data Factory and Azure Synapse Analytics, currently this link monitoring captures the assets and relationship generated from copy activity, but not data flow and SSIS activities.
- The aggregation and date filter are in UTC time.

Scans no longer run

If your Microsoft Purview scan used to run successfully but now fails, check these things:

1. Check the error message for failure details.
2. Did you change or rotate credentials to your resource? If so, update your scan with the correct credentials.
3. Is an Azure Policy preventing **updates to Storage accounts**? If so, follow the Microsoft Purview exception tag guide to create an exception for Purview accounts.
4. Are you using a self-hosted integration runtime? Check that it's up to date with the latest software and that it's connected to your network.

Next steps

- Microsoft Purview supported data sources and file types
- Manage data sources
- Scan and ingestion

Last updated on 12/09/2025

Understanding resource sets

This article helps you understand how resource sets are used in Microsoft Purview to map data assets to logical resources.

> **Note**
>
> Now rolling out, the advanced resource sets capability is available to all customers in Microsoft Purview Unified Catalog. Pricing for advanced resource sets is consistent with existing rates for classic Microsoft Purview data governance.

Background info

At-scale data processing systems typically store a single table in storage as multiple files. In the Microsoft Purview Unified Catalog, this concept is represented by using resource sets. A resource set is a single object in the catalog that represents a large number of assets in storage.

For example, suppose your Spark cluster has persisted a DataFrame into an Azure Data Lake Storage (ADLS) Gen2 data source. Although in Spark the table looks like a single logical resource, on the disk there are likely thousands of Parquet files, each of which represents a partition of the total DataFrame's contents. IoT data and web log data have the same challenge. Imagine you have a sensor that outputs log files several times a second. It won't take long until you have hundreds of thousands of log files from that single sensor.

How Microsoft Purview detects resource sets

Microsoft Purview supports detecting resource sets in Azure Blob Storage, ADLS Gen1, ADLS Gen2, Azure Files, and Amazon S3.

Microsoft Purview automatically detects resource sets when scanning. This feature looks at all of the data that's ingested via scanning and compares it to a set of defined patterns.

For example, suppose you scan a data source whose URL is `https://myaccount.blob.core.windows.net/mycontainer/machinesets/23/foo.parquet`. Microsoft Purview looks at the path segments and determines if they match any built-in patterns. It has built-in patterns for GUIDs, numbers, date formats, localization codes (for example, en-us), and so on. In this case, the number pattern matches *23*. Microsoft Purview assumes that this file is part of a resource set named `https://myaccount.blob.core.windows.net/mycontainer/machinesets/{N}/foo.parquet`.

Or, for a URL like `https://myaccount.blob.core.windows.net/mycontainer/weblogs/en_au/23.json`, Microsoft Purview matches both the localization pattern and the number pattern, producing a resource set named `https://myaccount.blob.core.windows.net/mycontainer/weblogs/{LOC}/{N}.json`.

Using this strategy, Microsoft Purview would map the following resources to the same resource set, `https://myaccount.blob.core.windows.net/mycontainer/weblogs/{LOC}/{N}.json`:

- `https://myaccount.blob.core.windows.net/mycontainer/weblogs/cy_gb/1004.json`
- `https://myaccount.blob.core.windows.net/mycontainer/weblogs/cy_gb/234.json`
- `https://myaccount.blob.core.windows.net/mycontainer/weblogs/de_Ch/23434.json`

File types that Microsoft Purview will not detect as resource sets

Microsoft Purview intentionally doesn't try to classify most document file types like Word, Excel, or PDF as Resource Sets. The exception is CSV format since that is a common partitioned file format.

How Microsoft Purview scans resource sets

When Microsoft Purview detects resources that it thinks are part of a resource set, it switches from a full scan to a sample scan. A sample scan opens only a subset of the files that it thinks are in the resource set. For each file it opens, it uses its schema and runs its classifiers. Microsoft Purview then finds the newest resource among the opened resources and uses that resource's schema and classifications in the entry for the entire resource set in the catalog.

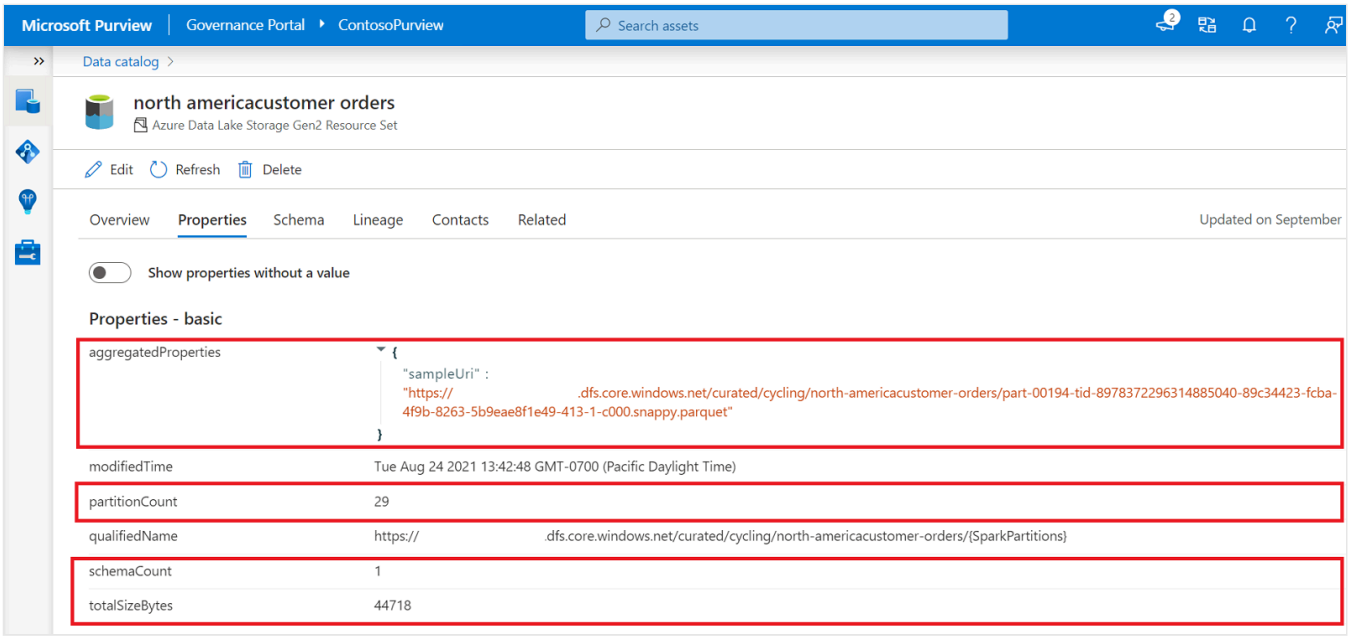
Advanced resource sets

Microsoft Purview can customize and further enrich your resource set assets through the **Advanced Resource Sets** capability. Advanced resource sets allow Microsoft Purview to understand the underlying partitions of data ingested and enables the creation of resource set pattern rules that customize how Microsoft Purview groups resource sets during scanning.

When Advanced Resource Sets are enabled, Microsoft Purview runs extra aggregations to compute the following information about resource set assets:

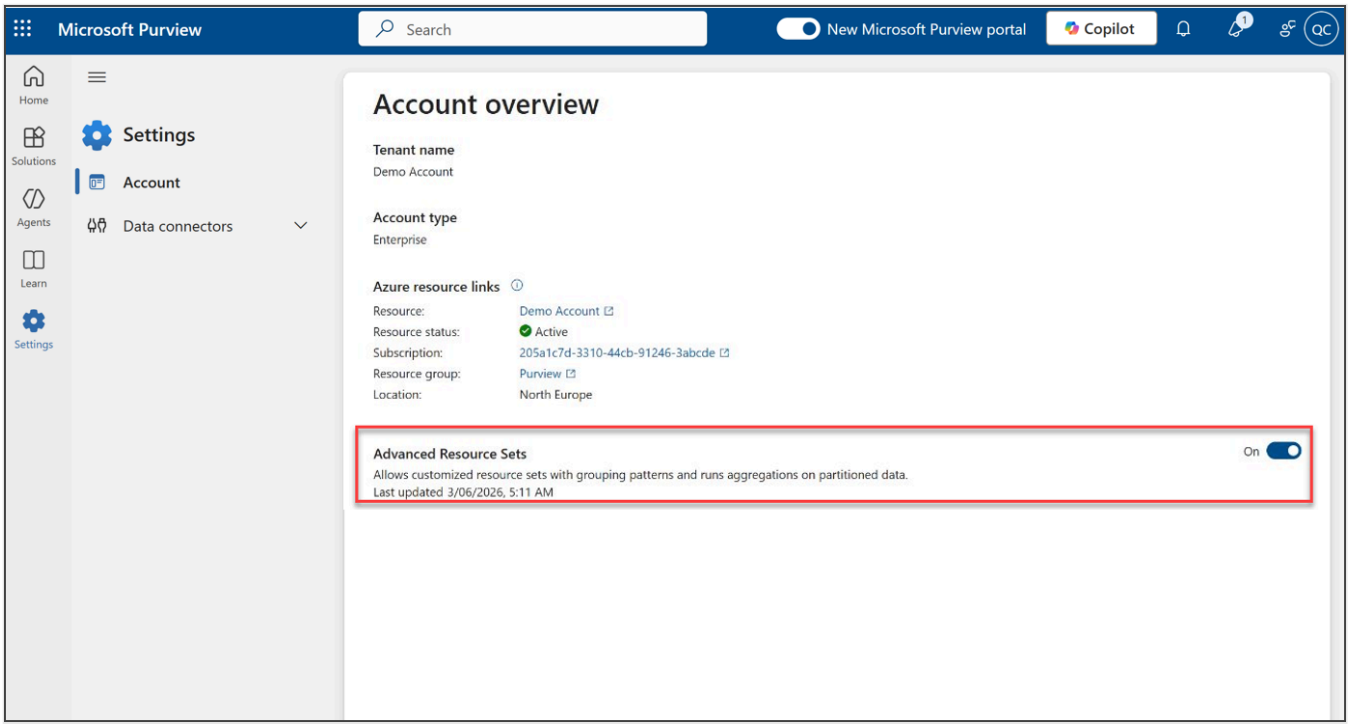
- A sample path from a file that comprises the resource set.
- A partition count that shows how many files make up the resource set.
- The total size of all files that comprise the resource set.

These properties can be found on the asset details page of the resource set.



Turning on advanced resource sets

Advanced resource sets are off by default in all new Microsoft Purview instances. Only users with the Data Curator role at root collection can manage Advanced Resource Sets settings. To enable advanced resource sets, go to **Microsoft Purview portal** > **Settings** > **Account**, and select the toggle for **Advanced Resources Sets**, as seen in this image:



After enabling advanced resource sets, the extra enrichments will occur on all newly ingested assets. These enrichments could take up to **12 hours** to be available on your assets after ingestion. The Microsoft Purview team recommends waiting an hour before scanning in new data lake data after toggling on the feature.

> **Important**
>
> Enabling advanced resource sets affects the refresh rate of asset and classification insights. When advanced resource sets are on, asset and classification insights update twice a day.
>
> Also, when you enable advanced resource sets, it could take up to **12 hours** to see schema updates.

Built-in resource set patterns

Microsoft Purview supports the following resource set patterns. These patterns can appear as a name in a directory or as part of a file name.

Regex-based patterns

Expand table

Pattern Name	Display Name	Description
Guid	{GUID}	A globally unique identifier as defined in RFC 4122
Number	{N}	One or more digits
Date/Time Formats	{Year}{Month}{Day}{N}	We support various date/time formats but all are represented with {Year}[delimiter]{Month}[delimiter]{Day} or series of {N}s.
4ByteHex	{HEX}	A four-digit HEX number.
Localization	{LOC}	A language tag as defined in BCP 47, both - and _ names are supported (for example, en_ca and en-ca)

Complex patterns

Expand table

Pattern Name	Display Name	Description
SparkPath	{SparkPartitions}	Spark partition file identifier
Date(yyyy/mm/dd)InPath	{Year}/{Month}/{Day}	Year/month/day pattern spanning multiple folders

How resource sets are displayed in Unified Catalog

When Microsoft Purview matches a group of assets into a resource set, it attempts to extract the most useful information to use as a display name in the catalog. Some examples of the default naming convention applied:

Example 1

Qualified name: `https://myblob.blob.core.windows.net/sample-data/name-of-spark-output/{SparkPartitions}`

Display name: "name of spark output"

Example 2

Qualified name: `https://myblob.blob.core.windows.net/my-partitioned-data/{Year}-{Month}-{Day}/{N}-{N}-{N}-{N}/{GUID}`

Display name: "my partitioned data"

Example 3

Qualified name: `https://myblob.blob.core.windows.net/sample-data/data{N}.csv`

Display name: "data"

Customizing resource set grouping using pattern rules

When scanning a storage account, Microsoft Purview uses a set of defined patterns to determine if a group of assets is a resource set. In some cases, Microsoft Purview's resource set grouping might not accurately reflect your data estate. These issues can include:

- Incorrectly marking an asset as a resource set.
- Putting an asset into the wrong resource set.
- Incorrectly marking an asset as not being a resource set.

To customize or override how Microsoft Purview detects which assets are grouped as resource sets and how they're displayed within the catalog, you can define pattern rules in the management center. For step-by-step instructions and syntax, see resource set pattern rules.

Known limitations with resource sets

- By default, resource set assets will only be deleted by a scan if Advanced Resource sets are enabled. If this capability is off, resource set assets can only be deleted manually or via API.

Next steps

To get started with Microsoft Purview, see Quickstart: Create a Microsoft Purview account.

Last updated on 03/31/2026

Create resource set pattern rules

06/06/2025

At-scale data processing systems typically store a single table in storage as multiple files. This concept is represented in Microsoft Purview by using resource sets. A resource set is a single object in Microsoft Purview Unified Catalog that represents a large number of assets in storage. To learn more, see Understanding resource sets.

When scanning a storage account, Microsoft Purview uses a set of defined patterns to determine if a group of assets is a resource set. In some cases, the resource set grouping might not accurately reflect your data estate. Resource set pattern rules allow you to customize or override how Microsoft Purview detects which assets are grouped as resource sets and how they're displayed within the catalog.

Pattern rules are currently supported in the following source types:

- Azure Data Lake Storage Gen2
- Azure Blob Storage
- Azure Files
- Amazon S3

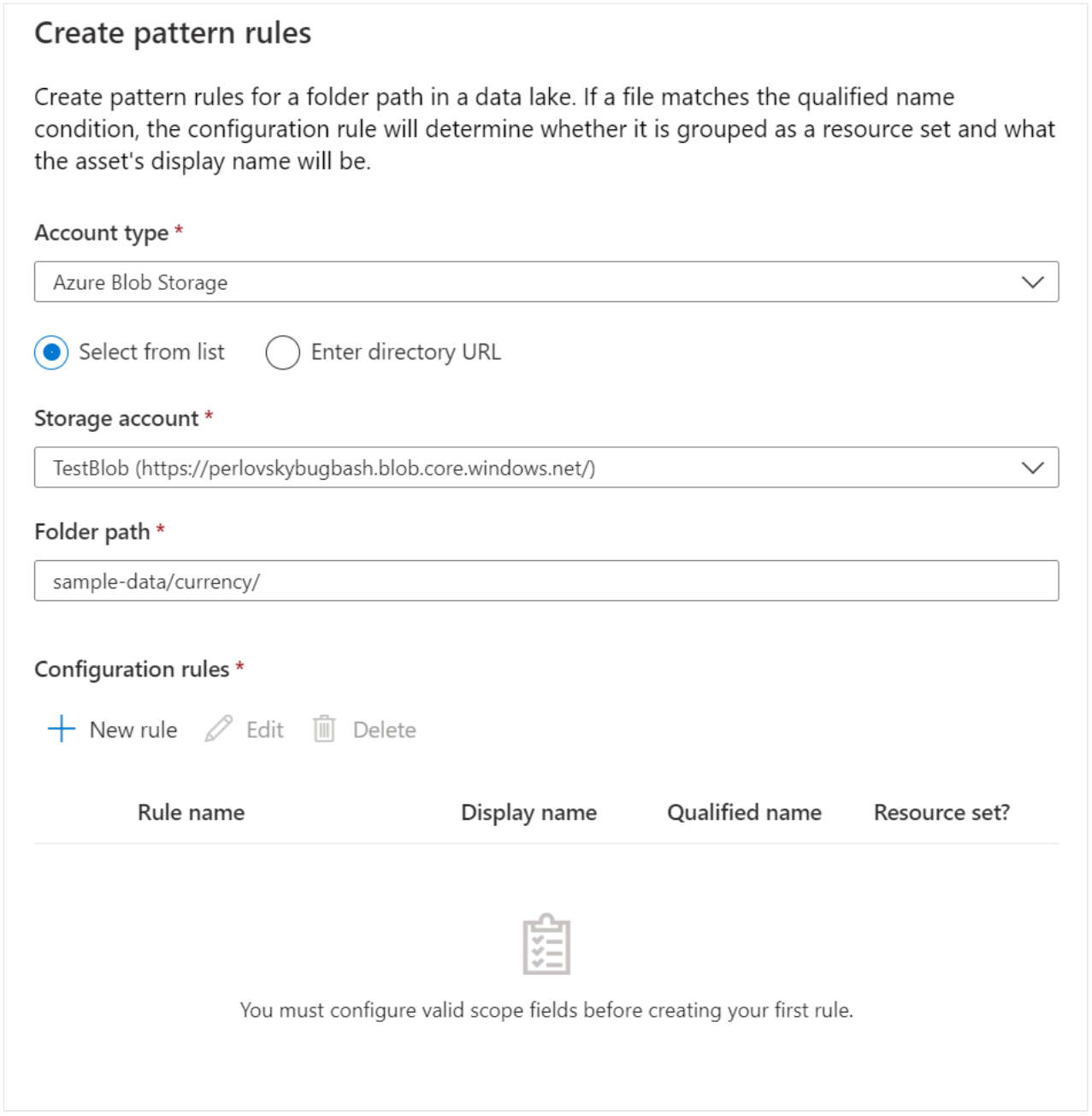
The Advanced resource set feature set must be enabled to create resource set pattern rules. To learn more, see Understanding Advanced resource sets.

How to create a resource set pattern rule

Follow the steps below to create a new resource set pattern rule:

1. Go to Data Map and select **Source management**, select **Pattern rules**, then select **New** to create a new rule set.

2. Enter the scope of your resource set pattern rule. Select your storage account type and the name of the storage account you wish to create a rule set on. Each set of rules is applied relative to a folder path scope specified in the **Folder path** field.



3. To enter a rule for a configuration scope, select **+ New Rule**.

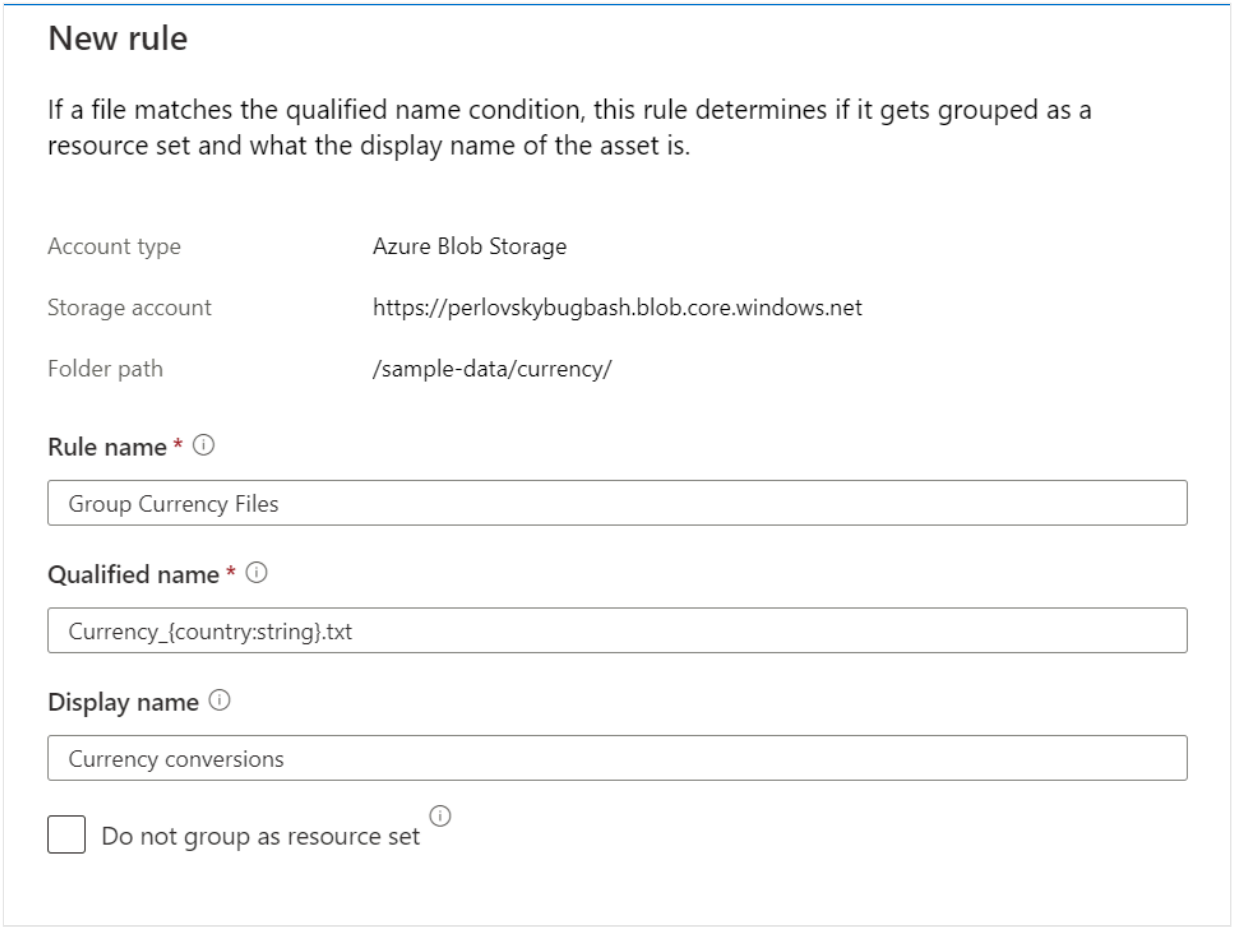
4. Enter in the following fields to create a rule:

   a. **Rule name:** The name of the configuration rule. This field has no effect on the assets the rule applies to.

   b. **Qualified name:** A qualified path that uses a combination of text, dynamic replacers, and static replacers to match assets to the configuration rule. This path is relative to the scope of the configuration rule. See the syntax section below for detailed instructions on how to specify qualified names.

   c. **Display name:** The display name of the asset. This field is optional. Use plain text and static replacers to customize how an asset is displayed in the catalog. For more detailed instructions, see the syntax section below.

d. **Do not group as resource set:** If enabled, matched resource won't be grouped into a resource set.



5. Save the rule by selecting **Add**.

> **Note**
>
> After a pattern rule is created, all new scans will apply the rule during ingestion. Existing assets in Unified Catalog will be updated via a background process which can take up to a few hours.

Pattern rule syntax

When creating resource set pattern rules, use the following syntax to specify which assets rules apply to.

Dynamic replacers (single brackets)

Single brackets are used as **dynamic replacers** in a pattern rules. Specify a dynamic replacer in the qualified name using format `{<replacerName>:<replacerType>}`. If matched, dynamic replacers are used as a grouping condition that indicate assets should be represented as a

resource set. If the assets are grouped into a resource set, the resource set qualified path would contain `{replacerName}` where the replacer was specified.

For example, If two assets `folder1/file-1.csv` and `folder2/file-2.csv` matched to rule `{folder:string}/file-{NUM:int}.csv`, the resource set would be a single entity `{folder}/file-{NUM}.csv`.

Special case: Dynamic replacers when not grouping into resource set

If *Don't group as resource set* is enabled for a pattern rule, the replacer name is an optional field. `{:<replacerType>}` is valid syntax. For example, `file-{:int}.csv` would successfully match for `file-1.csv` and `file-2.csv` and create two different assets instead of a resource set.

Static replacers (double brackets)

Double brackets are used as **static replacers** in the qualified name of a pattern rule. Specify a static replacer in the qualified name using format `{{<replacerName>:<replacerType>}}`. If matched, each set of unique static replacer values creates different resource set groupings.

For example, If two assets `folder1/file-1.csv` and `folder2/file-2.csv` matched to rule `{{folder:string}}/file-{NUM:int}.csv`, two resource sets would be created `folder1/file-{NUM}.csv` and `folder2/file-{NUM}.csv`.

Static replacers can be used to specify the display name of an asset matching to a pattern rule. Using `{{<replacerName>}}` in the display name of a rule uses the matched value in the asset name.

Available replacement types

Below are the available types that can be used in static and dynamic replacers:

Expand table

Type	Structure
string	A series of one or more Unicode characters including delimiters like spaces.
int	A series of 1 or more 0-9 ASCII characters, it can be 0 prefixed (for example, 0001).
guid	A series of 32 or 8-4-4-4-12 string representation of an UUID as defined in RFC 4122.
date	A series of 6 or 8 0-9 ASCII characters with optionally separators: yyyymmdd, yyyy-mm-dd, yymmdd, yy-mm-dd, specified in RFC 3339.

Type	Structure
time	A series of 4 or 6 0-9 ASCII characters with optionally separators: HHmm, HH:mm, HHmmss, HH:mm:ss specified in RFC 3339.
timestamp	A series of 12 or 14 0-9 ASCII characters with optionally separators: yyyy-mm-ddTHH:mm, yyyymmddhhmm, yyyy-mm-ddTHH:mm:ss, yyyymmddHHmmss specified in RFC 3339.
boolean	Can contain 'true' or 'false', case insensitive.
number	A series of 0 or more 0-9 ASCII characters, it can be 0 prefixed (for example, 0001) followed by optionally a dot '.' and a series of 1 or more 0-9 ASCII characters, it can be 0 postfixed (for example, .100)
hex	A series of one or more ASCII characters from the set 0-1 and A-F, the value can be 0 prefixed
locale	A string that matches the syntax specified in RFC 5646.

Order of resource set pattern rules getting applied

Below is the order of operations for applying pattern rules:

1. More specific scopes take priority if an asset matches to two rules. For example, rules in a scope `container/folder` will apply before rules in scope `container`.
2. Order of rules within a specific scope. This can be edited in the UX.
3. If an asset doesn't match to any specified rule, the default resource set heuristics apply.

Examples

Example 1

SAP data extraction into full and delta loads

Inputs

Files:

- `https://myazureblob.blob.core.windows.net/bar/customer/full/2020/01/13/saptable_customer_20200101_20200102_01.txt`
- `https://myazureblob.blob.core.windows.net/bar/customer/full/2020/01/13/saptable_customer_20200101_20200102_02.txt`

- `https://myazureblob.blob.core.windows.net/bar/customer/delta/2020/01/15/saptable_customer_20200101_20200102_01.txt`
- `https://myazureblob.blob.core.windows.net/bar/customer/full/2020/01/17/saptable_customer_20200101_20200102_01.txt`
- `https://myazureblob.blob.core.windows.net/bar/customer/full/2020/01/17/saptable_customer_20200101_20200102_02.txt`

Pattern rule

Scope: `https://myazureblob.blob.core.windows.net/bar/`

Display name: 'External Customer'

Qualified Name: `customer/{extract:string}/{year:int}/{month:int}/{day:int}/saptable_customer_{date_from:date}_{date_to:time}_{sequence:int}.txt`

Resource Set: true

Output

One resource set asset

Display Name: External Customer

Qualified Name: `https://myazureblob.blob.core.windows.net/bar/customer/{extract}/{year}/{month}/{day}/saptable_customer_{date_from}_{date_to}_{sequence}.txt`

Example 2

IoT data in avro format

Inputs

Files:

- `https://myazureblob.blob.core.windows.net/bar/raw/machinename-89/01-01-2020/22:33:22-001.avro`
- `https://myazureblob.blob.core.windows.net/bar/raw/machinename-89/01-01-2020/22:33:22-002.avro`

- `https://myazureblob.blob.core.windows.net/bar/raw/machinename-89/02-01-2020/22:33:22-001.avro`
- `https://myazureblob.blob.core.windows.net/bar/raw/machinename-90/01-01-2020/22:33:22-001.avro`

Pattern rules

Scope: `https://myazureblob.blob.core.windows.net/bar/`

Rule 1

Display name: 'machine-89'

Qualified Name: `raw/machinename-89/{date:date}/{time:time}-{id:int}.avro`

Resource Set: true

Rule 2

Display name: 'machine-90'

Qualified Name: `raw/machinename-90/{date:date}/{time:time}-{id:int}.avro`

Resource Set: true

Outputs

Two resource sets

Resource Set 1

Display Name: machine-89

Qualified Name: `https://myazureblob.blob.core.windows.net/bar/raw/machinename-89/{date}/{time}-{id}.avro`

Resource Set 2

Display Name: machine-90

Qualified Name: `https://myazureblob.blob.core.windows.net/bar/raw/machinename-90/{date}/{time}-{id}.avro`

Example 3

IoT data in avro format

Inputs

Files:

- `https://myazureblob.blob.core.windows.net/bar/raw/machinename-89/01-01-2020/22:33:22-001.avro`
- `https://myazureblob.blob.core.windows.net/bar/raw/machinename-89/01-01-2020/22:33:22-002.avro`
- `https://myazureblob.blob.core.windows.netbar/raw/machinename-89/02-01-2020/22:33:22-001.avro`
- `https://myazureblob.blob.core.windows.net/bar/raw/machinename-90/01-01-2020/22:33:22-001.avro`

Pattern rule

Scope: `https://myazureblob.blob.core.windows.net/bar/`

Display name: 'Machine-{{machineid}}'

Qualified Name: `raw/machinename-{{machineid:int}}/{date:date}/{time:time}-{id:int}.avro`

Resource Set: true

Outputs

Resource Set 1

Display name: machine-89

Qualified Name: `https://myazureblob.blob.core.windows.net/bar/raw/machinename-89/{date}/{time}-{id}.avro`

Resource Set 2

Display name: machine-90

Qualified Name: `https://myazureblob.blob.core.windows.net/bar/raw/machinename-90/{date}/{time}-{id}.avro`

Example 4

Don't group into resource sets

Inputs

Files:

- `https://myazureblob.blob.core.windows.net/bar/raw/machinename-89/01-01-2020/22:33:22-001.avro`
- `https://myazureblob.blob.core.windows.net/bar/raw/machinename-89/01-01-2020/22:33:22-002.avro`
- `https://myazureblob.blob.core.windows.net/bar/raw/machinename-89/02-01-2020/22:33:22-001.avro`
- `https://myazureblob.blob.core.windows.net/bar/raw/machinename-90/01-01-2020/22:33:22-001.avro`

Pattern rule

Scope: `https://myazureblob.blob.core.windows.net/bar/`

Display name: `Machine-{{machineid}}`

Qualified Name: `raw/machinename-{{machineid:int}}/{{:date}}/{{:time}}-{{:int}}.avro`

Resource Set: false

Outputs

Four individual assets

Asset 1

Display name: machine-89

Qualified Name: `https://myazureblob.blob.core.windows.net/bar/raw/machinename-89/01-01-2020/22:33:22-001.avro`

Asset 2

Display name: machine-89

Qualified Name: `https://myazureblob.blob.core.windows.net/bar/raw/machinename-89/01-01-2020/22:33:22-002.avro`

Asset 3

Display name: machine-89

Qualified Name: `https://myazureblob.blob.core.windows.net/bar/raw/machinename-89/02-01-2020/22:33:22-001.avro`

Asset 4

Display name: machine-90

Qualified Name: `https://myazureblob.blob.core.windows.net/bar/raw/machinename-90/01-01-2020/22:33:22-001.avro`

Next steps

Get started by registering and scanning an Azure Data Lake Gen2 storage account.

Scan data sources in Data Map

In Microsoft Purview, after you register your data source, you can scan your source to capture technical metadata, extract schema, and apply classifications to your data.

- For more information about scanning, see Scans and ingestion in Data Map.
- Review scanning best practices.

In this article, you learn the basic steps for scanning any data source.

> **Tip**
>
> Each source has its own instructions and prerequisites for scanning. For the most complete scanning instructions, select your source from the **supported sources list** and review its scanning instructions.

Prerequisites

Review the list of sources that are available to register and scan in Microsoft Purview.

Before you can scan your data source, complete these steps:

1. Register your data source. This step essentially gives Microsoft Purview the address of your data source and maps it to a collection or domain in Data Map.
2. Consider your network and choose the right integration runtime configuration for your scenario.
3. Consider what credentials you use to connect to your source. All source pages have a **Scan** section that includes details about what authentication types are available.

Create a scan

In the following steps, use Azure Blob Storage as an example, and authenticate by using the Microsoft Purview Managed Identity.

> **Important**
>
> These steps describe how to create a scan. For source-specific prerequisites and scanning instructions, see **Data sources that connect to Data Map**.

1. Open the Microsoft Purview portal and go to **Data Map** > **Data sources**. You can view your registered sources in either a map or table view.

> **Tip**
>
> If your Data Map has a large number of registered sources, the table view might perform better.

2. Find your source and select **New Scan**.

3. Enter a **Name** for the scan.

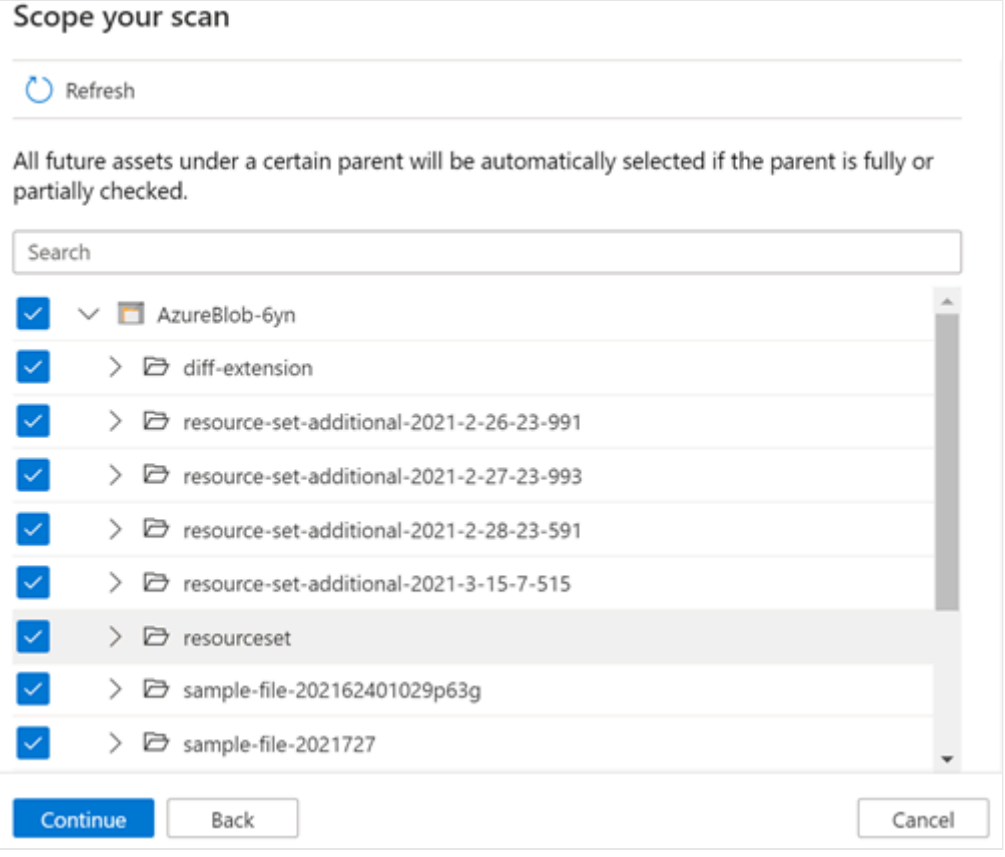
4. For **Credential**, select your authentication method.

5. Choose the current domain, collection, or a subcollection for the scan. The collection or domain you choose is where the scan stores the discovered metadata.

   > **Note**
   >
   > The scan is always in the same domain as the registered source, but you can select a subcollection.

6. Select **Test connection**. If the connection is successful, select **Continue**. See troubleshooting if the connection isn't successful.

7. Depending on the source, you can scope your scan to a specific subset of data. For Azure Blob Storage, select folders and subfolders by choosing the appropriate items in the list.



8. Select a scan rule set. The scan rule set contains the kinds of data classifications your scan checks. You can choose the system default (which contains all classifications available for the source), existing custom rule sets made by others in your organization, or create a new rule set inline.

   > **Note**
   >
   > You can only select the credentials and scan rule sets associated with the domain where you registered your source.

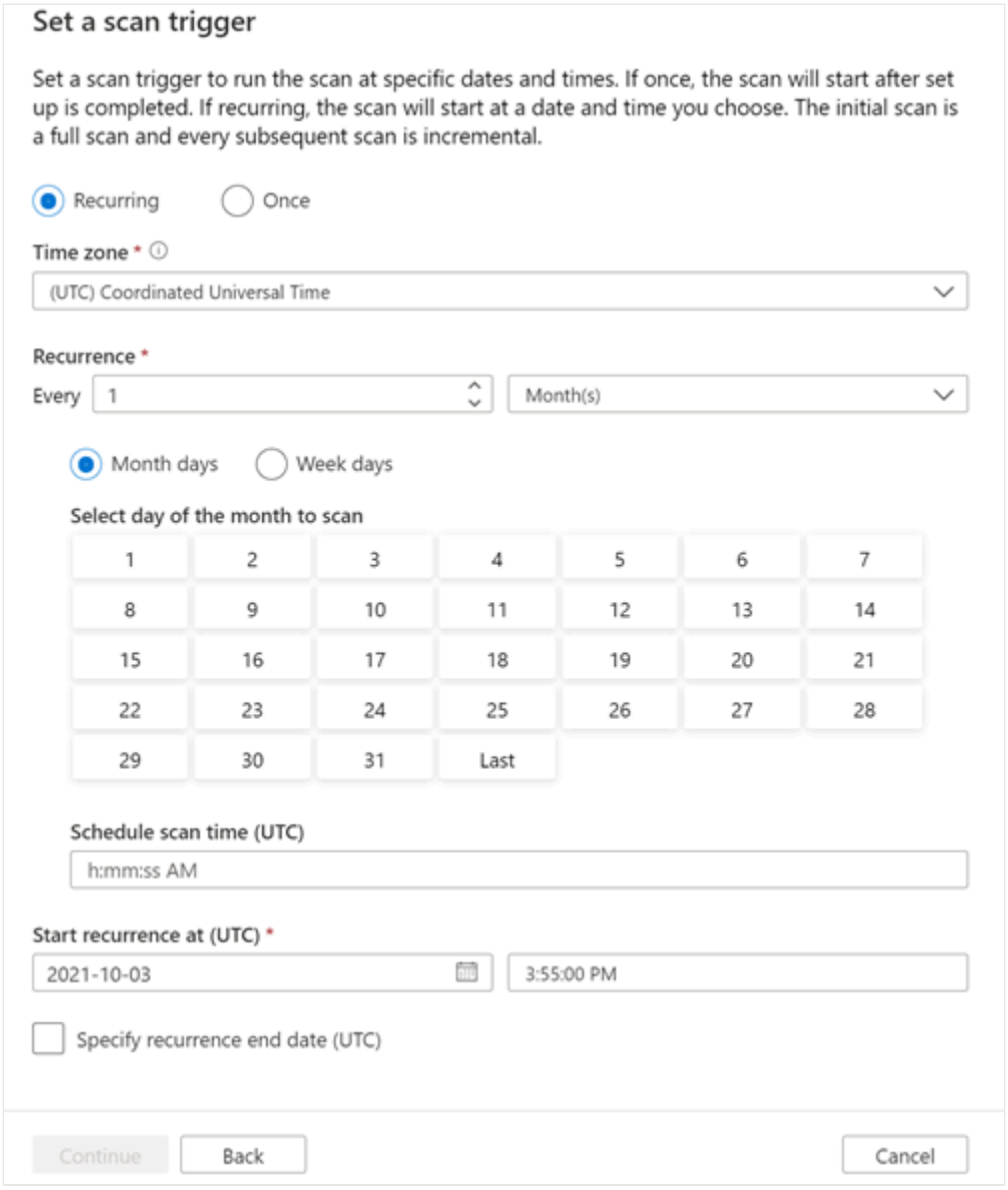
9. Choose your scan trigger. You can set up a schedule or run the scan once. Learn more about the supported schedule options.

10. Review your scan and select **Save and run**.

Schedule a scan

When you set up the scan, choose to run it once or on demand, or set up a recurring schedule. You can configure the following schedule options:

- **Time zone**: Select the time zone you'd like to align your scan schedule with. If the time zone you select observes daylight savings, the trigger autoadjusts for the difference.
- **Recurrence**: Select a **daily**, **weekly**, or **monthly** scan recurrence.
  - **Daily recurrence**: Set recurrence to every X days, and specify the scan start time of the day.
  - **Weekly recurrence**: Set recurrence to every X weeks, select one or multiple days of the week, and specify the scan start time of the day.
  - **Monthly recurrence**: Set recurrence to every X months, choose between by month days or by weekdays, select one or multiple days or weekdays of the month, and specify the scan start time of the day.
- **Start recurrence at**: Set when the scan schedule begins.
- **Specify recurrence end date** (optional): If you want to stop the scan after a certain amount of time, enable this option by selecting the check box and provide an end date.

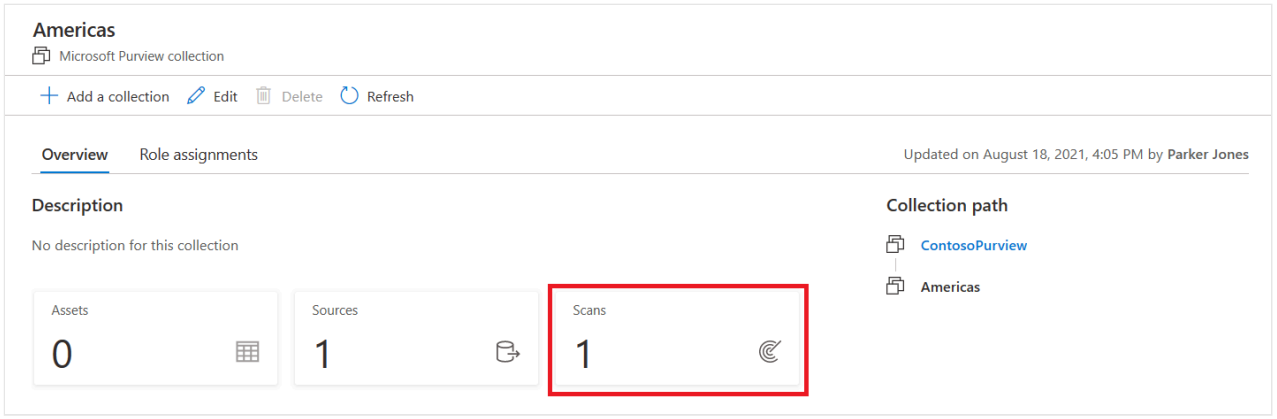


View a scan

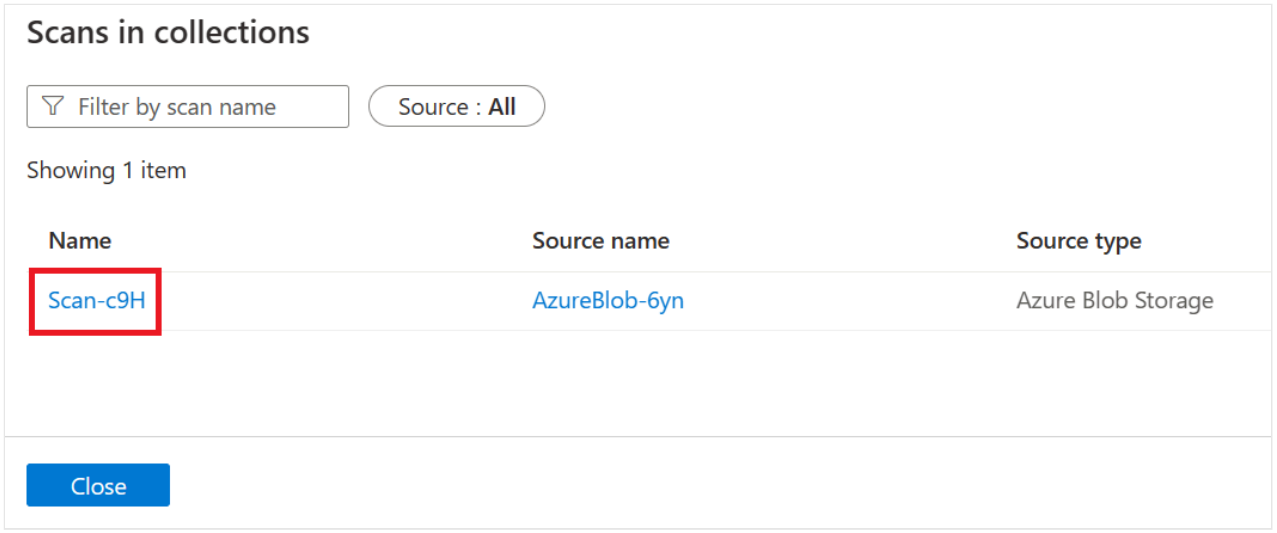
Depending on the amount of data in your data source, a scan can take some time to run. Here's how you can check on progress and see results when the scan is complete.

1. You can view your scan from the collection, domain, or from the source itself.

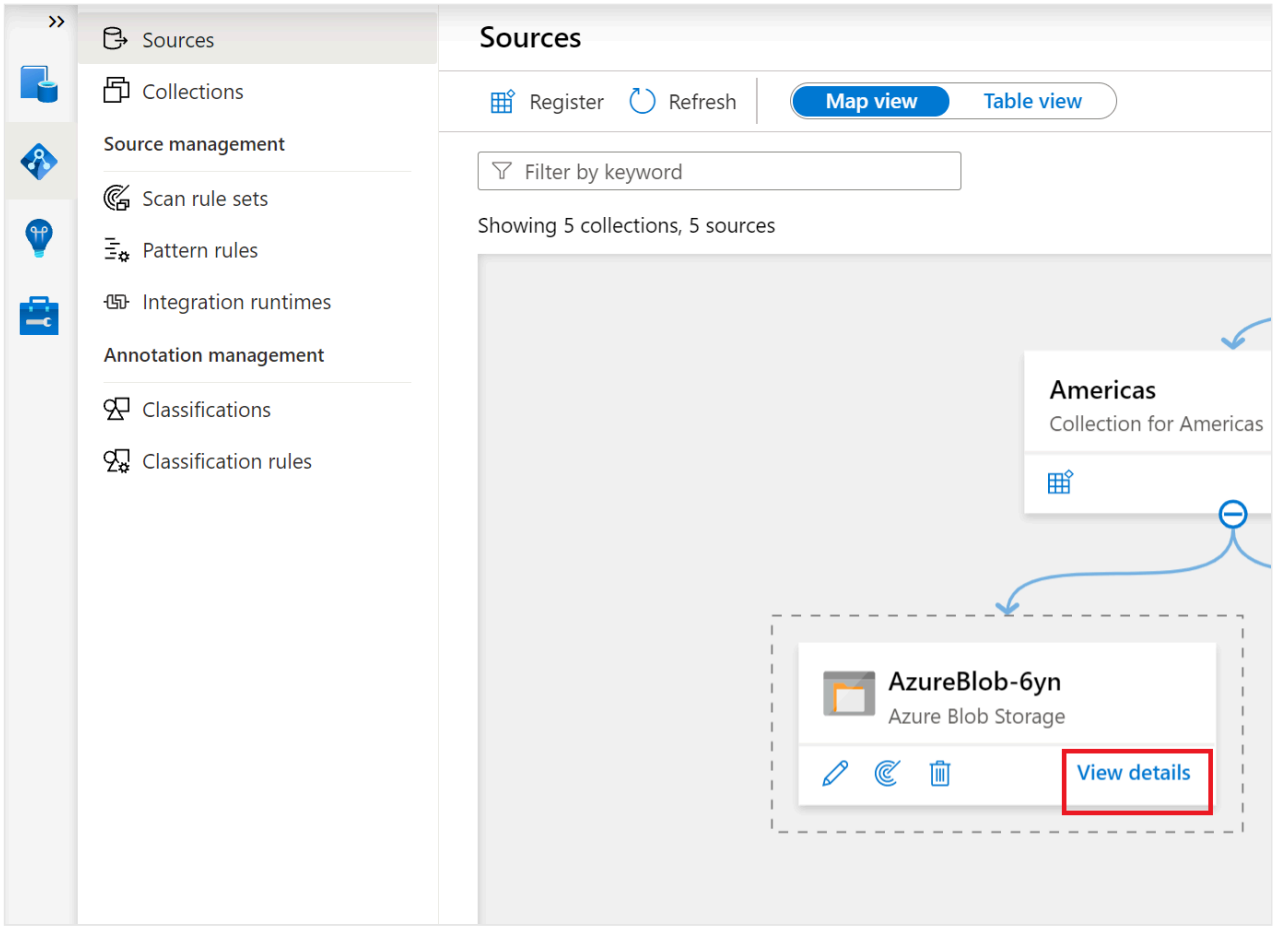
2. To view from the collection or domain, go to your *Collection* or *Domain* in Data Map and select **Scans**.



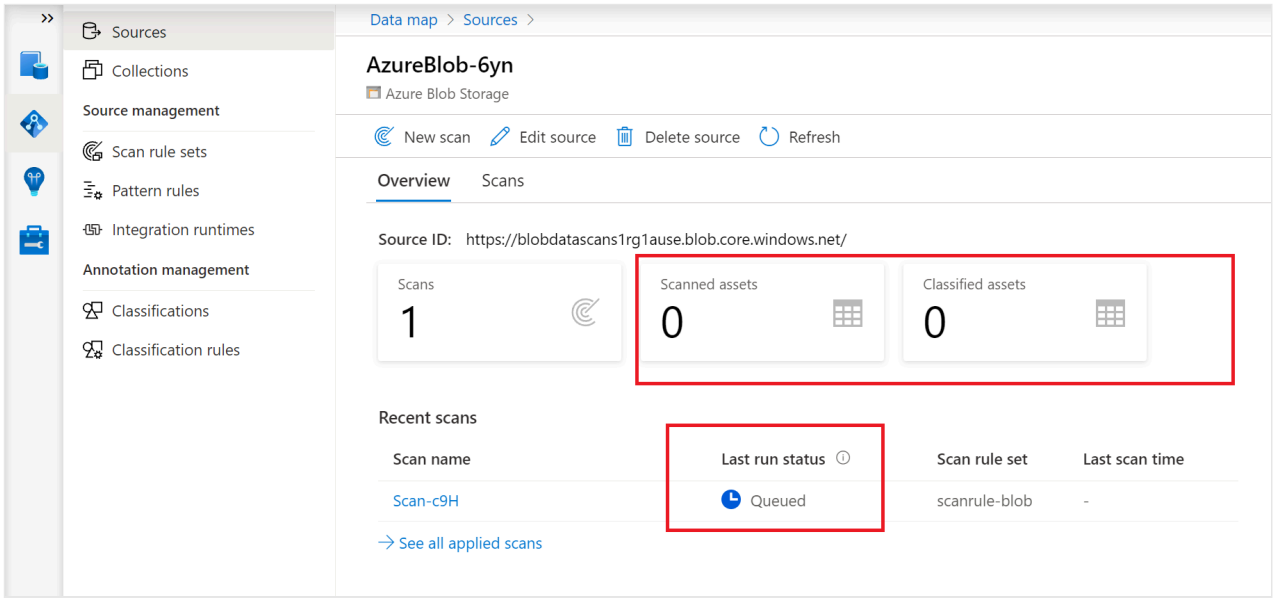
3. Select your scan name to see details.



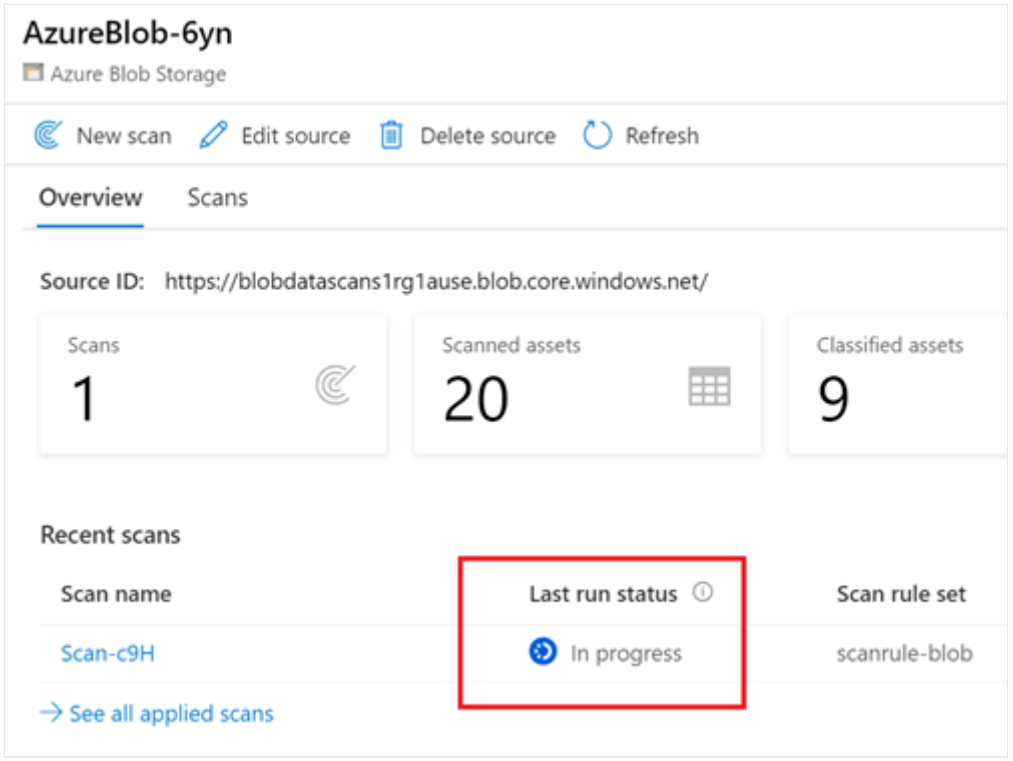
4. Or, you can go directly to the *data source* in its *Collection* or *Domain* and select **View Details** to check the status of the scan.

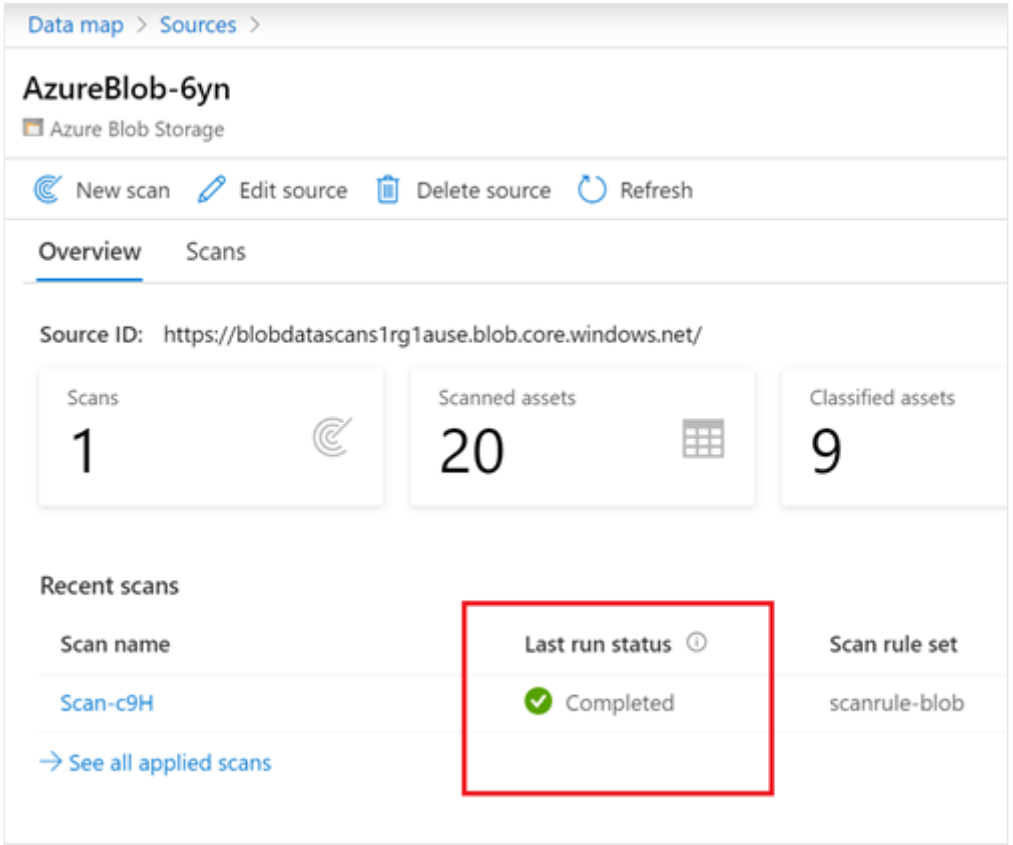


5. The scan details show the progress of the scan in the **Last run status** and the number of assets *scanned* and *classified*.



6. The **Last run status** updates to **In progress** and then **Completed** once the entire scan runs successfully.

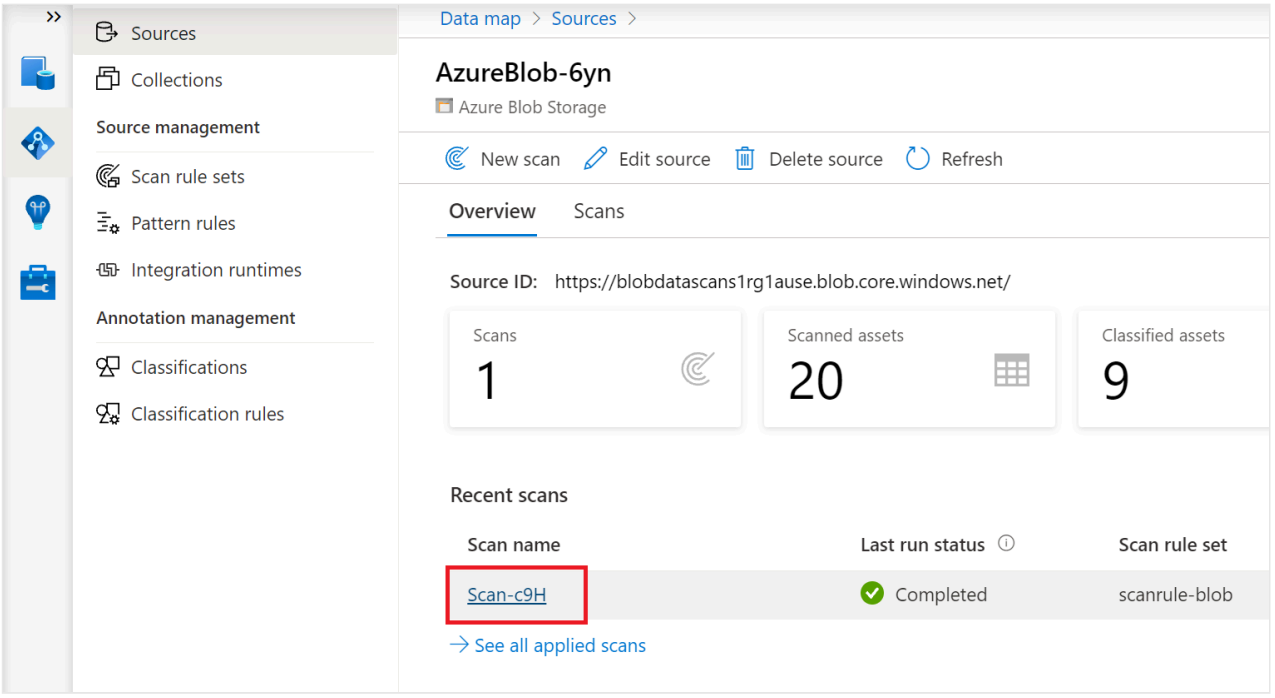




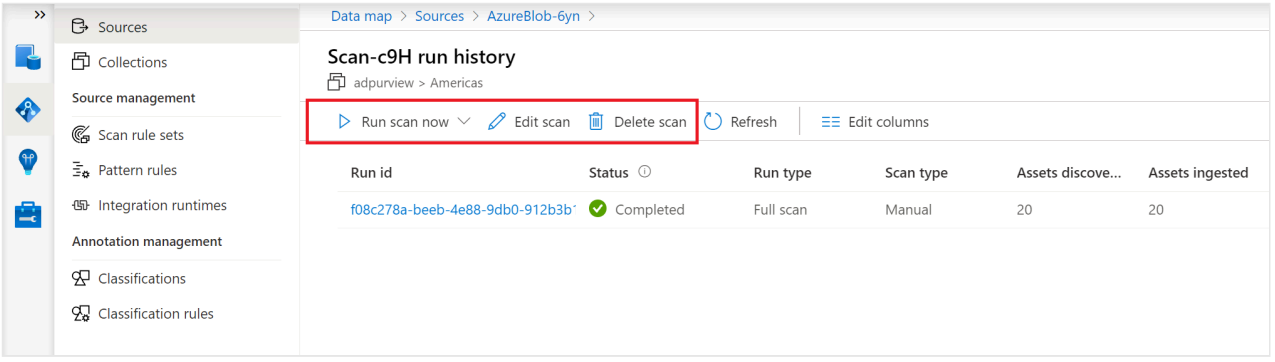
Manage a scan

After a scan finishes, you can manage it or run it again.

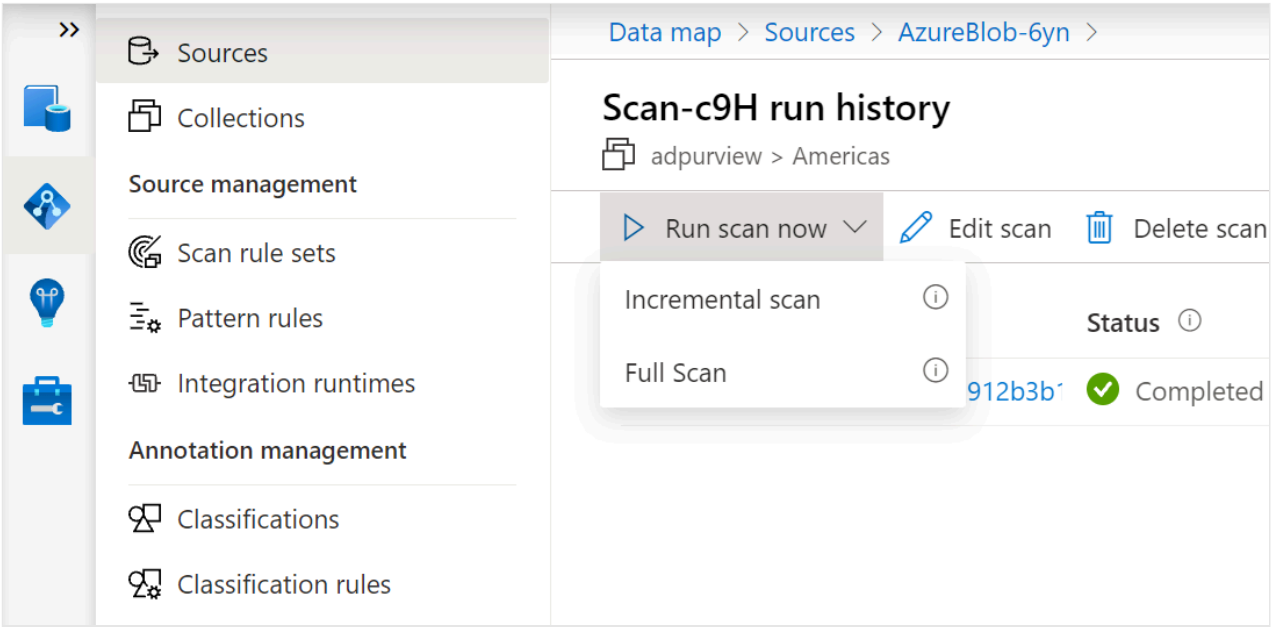
1. Select the **Scan name** from either the collections list or the source page to manage the scan.



2. You can run the scan again, edit the scan, or delete the scan.



3. You can run a full scan, which scans all the content in your scope. Some sources also have an **Incremental scan** option. An **Incremental scan** scans only those resources that are updated since the last scan. Check the **supported capabilities** table in your source page to see if incremental scan is available for your source after the first scan.



Troubleshooting

Setting up the connection for your scan can be complex, since it's a custom set up for your network and your credentials.

If you can't connect to your source, follow these steps:

1. Review your source page prerequisites to make sure you didn't miss anything.
2. Review your authentication option in the **Scan** section of your source page to confirm you set up the authentication method correctly.
3. Review troubleshoot connections.
4. Create a support request, so the support team can help you troubleshoot your specific environment.

Next steps

- Scanning best practices
- Azure Data Lake Storage Gen 2
- Power BI tenant
- Azure SQL Database

Last updated on 12/11/2025

Data Map scanning best practices

Microsoft Purview Data Map supports automated scanning of on-premises, multicloud, and software as a service (SaaS) data sources.

When you run a scan, the process starts to ingest metadata from the registered data sources. At the end of the scan and curation process, you get curated metadata that includes technical metadata. This metadata can include data asset names such as table names or file names, file size, columns, and data lineage. For structured data sources, schema details are also captured. A relational database management system is an example of this type of source.

The curation process applies automated classification labels on the schema attributes based on the scan rule set you configure. If your Microsoft Purview account is connected to the Microsoft Purview portal, sensitivity labels are applied.

> **Important**
>
> If you have any Azure Policies that prevent updates to storage accounts, these policies cause errors for the Microsoft Purview scanning process. See Create an Azure policy exclusion for Microsoft Purview to create an exception for Microsoft Purview accounts.

Why do you need best practices to manage data sources?

Best practices help you:

- Optimize cost.
- Build operational excellence.
- Improve security compliance.
- Gain performance efficiency.

Register a source and establish a connection

The following design considerations and recommendations help you register a source and establish a connection.

Design considerations

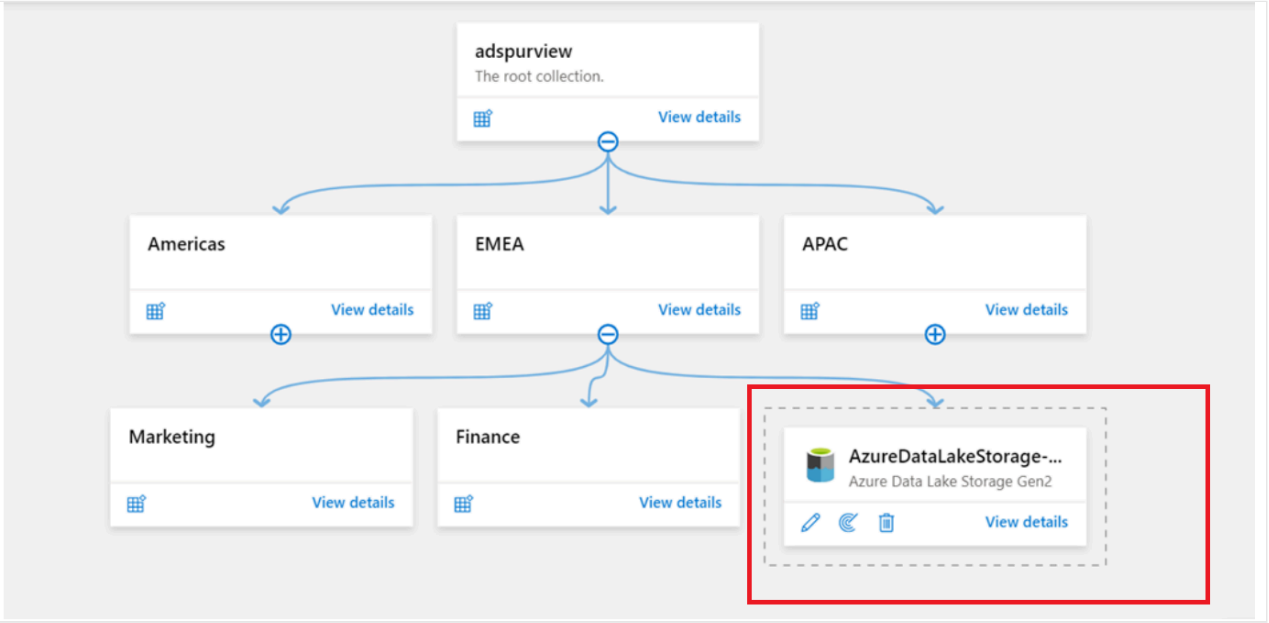
- Use collections to create the hierarchy that aligns with the organization's strategy, such as geographical location, business function, or source of data. The hierarchy defines the data

sources to register and scan.

- By design, you can't register data sources multiple times in the same Microsoft Purview account. This architecture helps you avoid the risk of assigning different access control to the same data source.

Design recommendations

- If multiple teams consume the metadata of the same data source, register and manage the data source at a parent collection. Then, create corresponding scans under each subcollection. In this way, relevant assets appear under each child collection. The map view groups sources without parents in a dotted box. No arrows link them to parents.



- Use the **Azure Multiple** option if you need to register multiple sources, such as Azure subscriptions or resource groups, in the cloud. For more information, see the following documentation:
  - Scan multiple sources in Microsoft Purview
  - Check data source readiness at scale
  - Configure access to data sources for Microsoft Purview MSI at scale

- After you register a data source, you can scan the same source multiple times. Different teams or business units might use the same source in different ways.

For more information on how to define a hierarchy for registering data sources, see Best practices on collections architecture.

Scanning

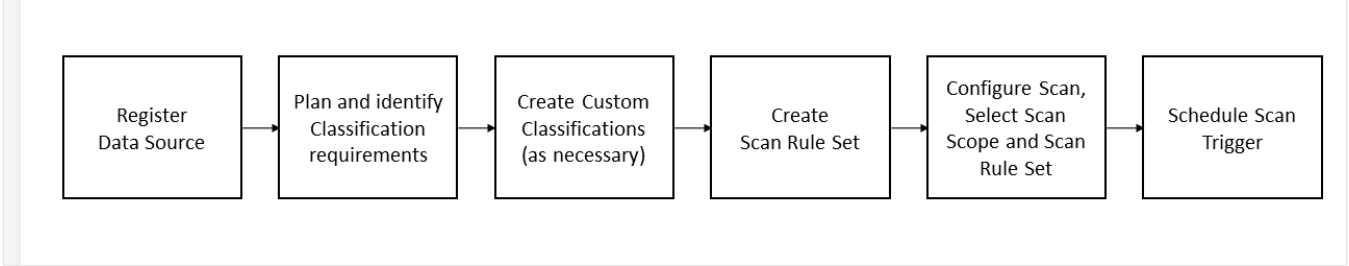
The following design considerations and recommendations are organized based on the key steps involved in the scanning process.

Design considerations

- After you register the data source, set up a scan to manage automated and secure metadata scanning and curation.
- Scan setup includes configuring the name of the scan, scope of scan, integration runtime, scan trigger frequency, scan rule set, and resource set uniquely for each data source per scan frequency.
- Before you create any credentials, consider your data source types and networking requirements. This information helps you decide which authentication method and integration runtime you need for your scenario.

Design recommendations

After you register your source in the relevant collection, plan and follow the order shown in this example when you set up the scan. This process order helps you avoid unexpected costs and rework.



1. Identify your classification requirements from the system in-built classification rules. Or create specific custom classification rules, as necessary. Base them on specific industry, business, or regional requirements, which aren't available out of the box:

    - See the classification best practices.
    - See how to create a custom classification and classification rule.

2. Create scan rule sets before you configure the scan. When you create the scan rule set, ensure the following points:

    - Verify if the system default scan rule set is sufficient for the data source you're scanning. Otherwise, define your custom scan rule set.

    - The custom scan rule set can include both system default and custom rules, so clear those options that aren't relevant for the data assets you're scanning.

- Where necessary, create a custom rule set to exclude unwanted classification labels. For example, the system rule set contains generic government code patterns for the planet, not just the United States. Your data might match the pattern of some other type, such as "Belgium Driver's License Number."
- Limit custom classification rules to *most important* and *relevant* labels to avoid clutter. You don't want to have too many labels tagged to the asset.
- If you modify the custom classification or scan rule set, a full scan is triggered. Configure the classification and scan rule set appropriately to avoid rework and costly full scans.

  > ⓘ **Note**
  >
  > When you scan a storage account, Microsoft Purview uses a set of defined patterns to determine if a group of assets forms a resource set. You can use resource set pattern rules to customize or override how Microsoft Purview detects which assets are grouped as resource sets. The rules also determine how the assets are displayed within the catalog. For more information, see Create resource set pattern rules. This feature has cost considerations. For information, see the Microsoft Purview pricing site.

3. Set up a scan for the registered data sources. - **Scan name**: By default, Microsoft Purview uses the naming convention **SCAN-[A-Z][a-z][a-z]**, which isn't helpful when you're trying to identify a scan that you ran. Use a meaningful naming convention. For instance, you could name the scan *environment-source-frequency-time* as DEVODS-Daily-0200. This name represents a daily scan at 0200 hours.

   - **Authentication**: Microsoft Purview offers various authentication methods for scanning data sources, depending on the type of source. It could be Azure cloud or on-premises or non-Microsoft sources. Follow the least-privilege principle for the authentication method in this order of preference:
     - Microsoft Purview MSI - Managed Service Identity (for example, for Azure Data Lake Storage Gen2 sources)
     - User-assigned managed identity
     - Service principal
     - SQL authentication (for example, for on-premises or Azure SQL sources)
     - Account key or basic authentication (for example, for SAP S/4HANA sources)

     For more information, see the how-to guide to manage credentials.

     > ⓘ **Note**

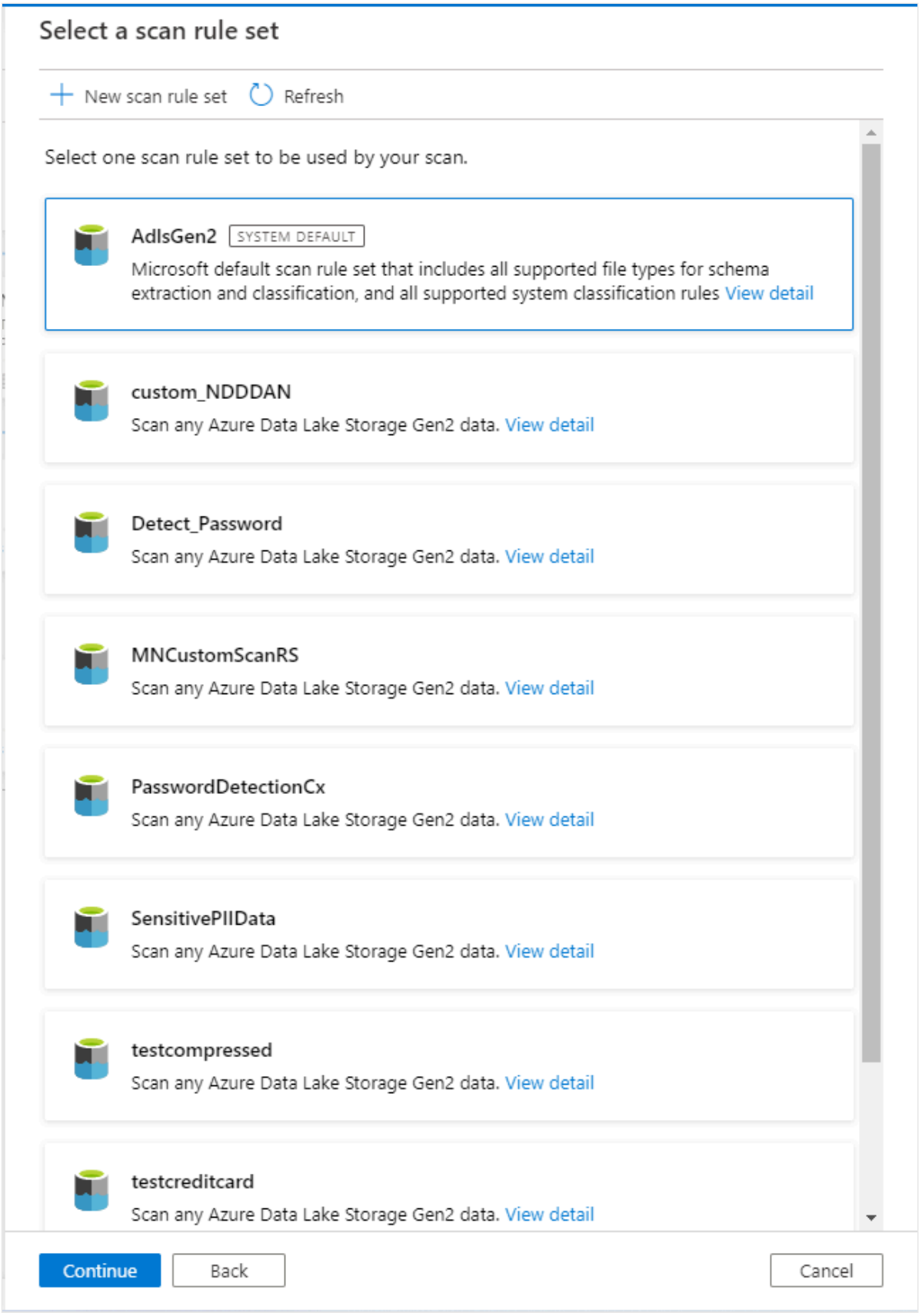
> If you enable a firewall for the storage account, you must use the managed identity authentication method when you set up a scan. When you set up a new credential, the credential name can only contain *letters, numbers, underscores, and hyphens*.

- **Integration runtime**
  - For more information, see Network architecture best practices.
  - If self-hosted integration runtime (SHIR) is deleted, any ongoing scans that rely on it fail.
  - When you use SHIR, make sure that the memory is sufficient for the data source you're scanning. For example, when you use SHIR for scanning an SAP source, if you see "out of memory error":
    - Ensure the SHIR machine has enough memory. The recommended amount is 128 GB.
    - In the scan setting, set the maximum memory available as some appropriate value, for example, 100.
    - For more information, see the prerequisites in Scan to and manage SAP ECC Microsoft Purview.

- **Scope scan**
  - When you set up the scope for the scan, select only the assets that are relevant at a granular level or parent level. This practice ensures that the scan cost is optimal and performance is efficient. All future assets under a certain parent are automatically selected if the parent is fully or partially checked.
  - Some examples for some data sources:
    - For Azure SQL Database or Data Lake Storage Gen2, you can scope your scan to specific parts of the data source. Select the appropriate items in the list, such as folders, subfolders, collections, or schemas.
    - For Oracle, Hive Metastore Database, and Teradata sources, you can specify a specific list of schemas to be exported through semicolon-separated values or schema name patterns.
    - For Google Big query, you can specify a specific list of datasets to be exported through semicolon-separated values.
    - When you create a scan for an entire AWS account, you can select specific buckets to scan. When you create a scan for a specific AWS S3 bucket, you can select specific folders to scan.
    - For Erwin, you can scope your scan by providing a semicolon-separated list of Erwin model locator strings.
    - For Cassandra, you can specify a specific list of key spaces to be exported through semicolon-separated values or through key spaces name patterns.

    - For Looker, you can scope your scan by providing a semicolon-separated list of Looker projects.
    - For Power BI tenant, you might only specify whether to include or exclude personal workspace.
  - In general, use ignore patterns where they're supported, based on wild cards (for example, for data lakes) to exclude temp, config files, RDBMS system tables, or backup or STG tables.
  - When you scan documents or unstructured data, avoid scanning a huge number of such documents. The scan processes the first 20 MB of such documents and might result in longer scan duration.

- **Scan rule set**
  - When you select the scan rule set, make sure to configure the relevant system or custom scan rule set that you created earlier. - You can create custom file types and fill in the details accordingly. Currently, Microsoft Purview supports only one character in Custom Delimiter. If you use custom delimiters, such as ~, in your actual data, you need to create a new scan rule set.



- **Scan type and schedule**
  - You can configure the scan process to run full or incremental scans.
  - Run the scans during nonbusiness or off-peak hours to avoid any processing overload on the source.
  - The initial scan is a full scan, and every subsequent scan is incremental. You can schedule subsequent scans as periodic incremental scans. Learn more about the

supported schedule options.
    - The frequency of scans should align with the change management schedule of the data source or business requirements. For example:
        - If the source structure could potentially change weekly, the scan frequency should be in sync. Changes include new assets or fields within an asset that are added, modified, or deleted.
        - If the classification or sensitivity labels need to be up to date on a weekly basis, perhaps for regulatory reasons, the scan frequency should be weekly. - If partition files are added every week in a source data lake, you might schedule monthly scans. You don't need to schedule weekly scans because there's no change in metadata. This suggestion assumes there are no new classification scenarios.
        - The maximum duration that the scan can run is seven days, possibly because of memory issues. This time period excludes the ingestion process. If progress isn't updated after seven days, the scan is marked as failed. The ingestion (into catalog) process currently doesn't have any such limitation.

- **Canceling scans**
    - Currently, you can cancel or pause scans only if the status of the scan transitions into an "In Progress" state from "Queued" after you trigger the scan.
    - Canceling an individual child scan isn't supported.

Points to note

- If you remove a field, column, table, or file from the source system after a scan runs, Microsoft Purview only shows the removal after the next scheduled full or incremental scan.
- You can delete an asset from a Microsoft Purview catalog by selecting **Delete** under the asset name. This action doesn't remove the object in the source. If you run a full scan on the same source, the scan reingests the object in the catalog. If you run an incremental scan, the deleted asset isn't picked up unless the object is modified at the source. For example, if a column is added or removed from the table.
- To understand the behavior of subsequent scans after *manually* editing a data asset or an underlying schema through the classic Microsoft Purview governance portal, see classic catalog asset details.
- For more information, see how to view, edit, and delete assets.

Next steps

Manage data sources

Last updated on 12/11/2025

Troubleshoot your scans and connections in the Microsoft Purview Data Map

Article • 04/30/2024

This article describes how to troubleshoot connection errors while setting up scans on data sources in Microsoft Purview, or errors that could occur with your scans.

Permission the credential on the data source

If you're using a managed identity or service principal as a method of authentication for scans, you'll have to allow these identities to have access to your data source.

There are specific instructions for each source type.

> **Important**
>
> Verify that you have followed all prerequisite and authentication steps for the source you are connecting to. You can find all available sources listed in the Microsoft Purview supported sources article.

Verifying Azure Role-based Access Control to enumerate Azure resources in the Microsoft Purview governance portal

Registering single Azure data source

To register a single data source in Microsoft Purview, such as an Azure Blob Storage or an Azure SQL Database, you must be granted at least **Reader** role on the resource or inherited from higher scope such as resource group or subscription. Some Azure RBAC roles, such as Security Admin, don't have read access to view Azure resources in control plane.

Verify this by following the steps below:

1. From the Azure portal, navigate to the resource that you're trying to register in Microsoft Purview. If you can view the resource, it's likely, that you already have at

least reader role on the resource.
2. Select **Access control (IAM)** > **Role Assignments**.
3. Search by name or email address of the user who is trying to register data sources in Microsoft Purview.
4. Verify if any role assignments, such as Reader, exist in the list or add a new role assignment if needed.

Scanning multiple Azure data sources

1. From the Azure portal, navigate to the subscription or the resource group.
2. Select **Access Control (IAM)** from the left menu.
3. Select **+Add**.
4. In the **Select input** box, select the **Reader** role and enter your Microsoft Purview account name (which represents its MSI name).
5. Select **Save** to finish the role assignment.
6. Repeat the steps above to add the identity of the user who is trying to create a new scan for multiple data sources in Microsoft Purview.

Scanning data sources using Private Link

If public endpoint is restricted on your data sources, to scan Azure data sources using Private Link, you need to set up a Self-hosted integration runtime and create a credential.

> **Important**
>
> Scanning multiple data sources which contain databases as Azure SQL database with *Deny public network access*, would fail. To scan these data sources using private Endpoint, instead use registering single data source option.

For more information about setting up a self-hosted integration runtime, see the article about ingestion private endpoints and scanning sources.

For more information how to create a new credential in Microsoft Purview, see the article about credentials for source authentication in Microsoft Purview.

Storing your credential in your key vault and using the right secret name and version

You must also store your credential in your Azure Key Vault instance and use the right secret name and version.

Verify this by following the steps below:

1. Navigate to your Key Vault.
2. Select **Settings** > **Secrets**.
3. Select the secret you're using to authenticate against your data source for scans.
4. Select the version that you intend to use and verify that the password or account key is correct by selecting **Show Secret Value**.

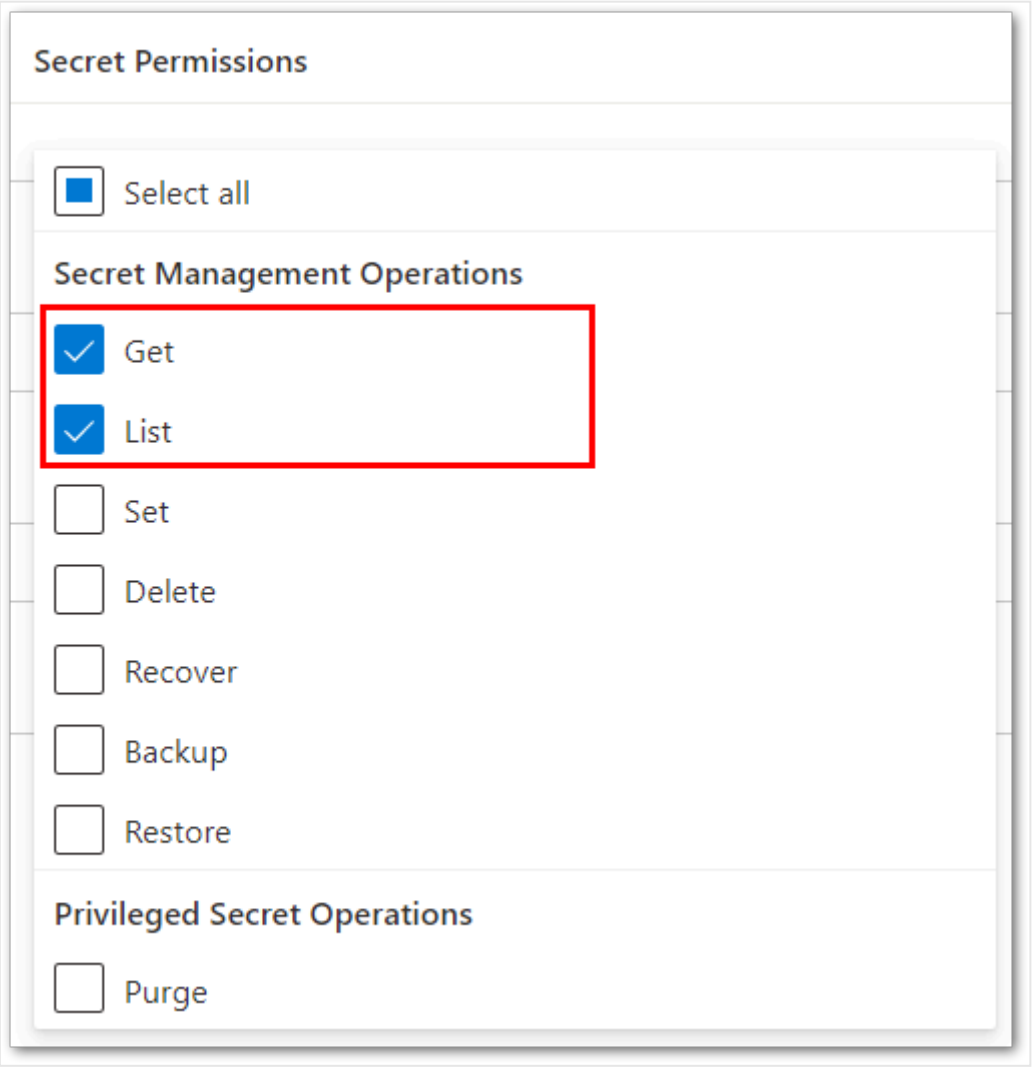
Verify permissions for the Microsoft Purview managed identity on your Azure Key Vault

Verify that the correct permissions have been configured for the Microsoft Purview managed identity to access your Azure Key Vault.

To verify this, do the following steps:

1. Navigate to your key vault and to the **Access policies** section

2. Verify that your Microsoft Purview managed identity shows under the *Current access policies* section with at least **Get** and **List** permissions on Secrets



If you don't see your Microsoft Purview managed identity listed, then follow the steps in Create and manage credentials for scans to add it.

Scans no longer run

If your Microsoft Purview scan used to successfully run, but are now failing, check these things:

1. Have credentials to your resource changed or been rotated? If so, you'll need to update your scan to have the correct credentials.
2. Is an Azure Policy preventing **updates to Storage accounts**? If so follow the Microsoft Purview exception tag guide to create an exception for Microsoft Purview accounts.
3. Are you using a self-hosted integration runtime? Check that it's up to date with the latest software and that it's connected to your network.

Test connection passes but scan fails with connection error

Are you using private endpoints or virtual networks? Confirm your network settings, paying attention to your Network Security Group (NSG) rules.

Receive JDK error when using a self-hosted integration runtime

Many sources that require a self-hosted integration runtime also require a certain version of Java to be installed on the machine where the self-hosted integration runtime is hosted. If you receive a JDK error, check the prerequisites on your specific data source's page and install the recommended JDK version.

Next steps

- Browse the Microsoft Purview Unified Catalog
- Search the Microsoft Purview Unified Catalog

Feedback

Was this page helpful? Yes No

Data classification in Data Map

Article • 03/04/2025

Data classification in Microsoft Purview Data Map is a way of categorizing data assets by assigning unique logical tags or classes to the data assets. Classification is based on the business context of the data. For example, you might classify assets by *Passport Number*, *Driver's License Number*, *Credit Card Number*, *SWIFT Code*, *Person's Name*, and so on.

When you classify data assets, you make them easier to understand, search, and govern. Classifying data assets also helps you understand the risks associated with them. This in turn can help you implement measures to protect sensitive or important data from ungoverned proliferation and unauthorized access across the data estate.

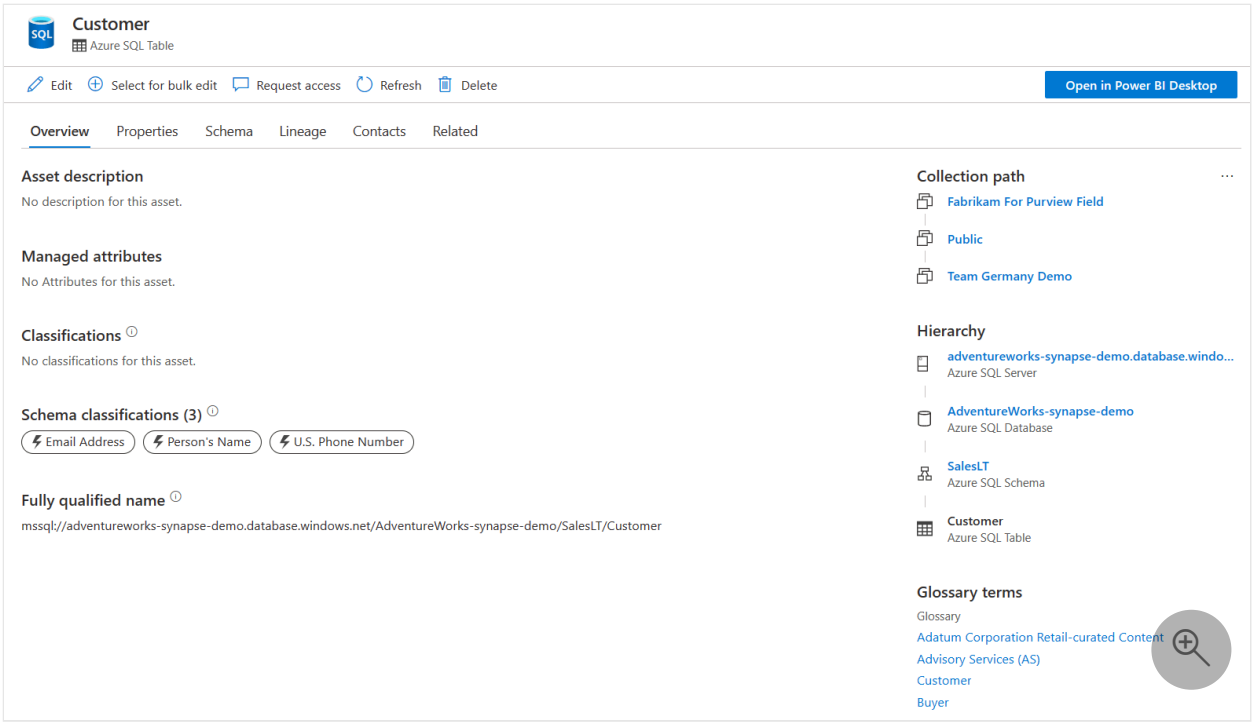
Data Map provides an automated classification capability while you scan your data sources. You get more than 200+ built-in system classifications and the ability to create custom classifications for your data. You can classify assets automatically when they're ingested as part of a configured scan, or you can edit them manually in the Microsoft Purview governance portal after they're scanned and ingested.

Uses of classification

Classification is the process of organizing data into *logical categories* that make the data easy to retrieve, sort, and identify for future use. This can be important for data governance. Among other reasons, classifying data assets is important because it helps you:

- Narrow down the search for data assets that you're interested in.
- Organize and understand the variety of data classes that are important in your organization and where they're stored.
- Understand the risks associated with your most important data assets and then take appropriate measures to mitigate them.

Following image shows classification applied while scanning on the *Customer* table in Azure SQL Database.



Types of classification

The Microsoft Purview governance portal supports both system and custom classifications.

- **System classifications**: 200+ system classifications supported are out of the box. See the full list of available system classifications.

  In the example in the preceding image, *Person's Name* is a system classification. System classification has the thunderbolt icon along with classification name. Hovering over the classification itself provides more details on the type of classification, and more details on how it was applied

- **Custom classifications**: You can create custom classifications when you want to classify assets based on a pattern or a specific column name that's unavailable as a system classification. Custom classification rules can be based on a *regular expression* pattern or *dictionary*.

  Let's say that the *Employee ID* column follows the EMPLOYEE{GUID} pattern (for example, EMPLOYEE9c55c474-9996-420c-a285-0d0fc23f1f55). You can create your own custom classification by using a regular expression, such as `\^Employee\[A-Za-z0-9\]{8}-\[A-Za-z0-9\]{4}-\[A-Za-z0-9\]{4}-\[A-Za-z0-9\]{4}-\[A-Za-z0-9\]{12}\$`.

> **Note**

Sensitivity labels are different from classifications. Sensitivity labels categorize assets in the context of data security and privacy, such as *Highly Confidential*, *Restricted*, *Public*, and so on. To use sensitivity labels in the Microsoft Purview Data Map, you'll need at least one Microsoft 365 license or account within the same Microsoft Entra tenant as your Microsoft Purview Data Map. For more information about the differences between sensitivity labels and classifications, see **sensitivity labels in the Microsoft Purview governance portal FAQ**.

Next steps

- Read about classification best practices
- Create custom classifications
- Automatically apply classifications
- Manually apply classifications
- Use the Microsoft Purview governance portal

Feedback

Was this page helpful? Yes No

Automatically apply classifications on assets in Data Map

Article • 03/25/2025

After data sources are registered in Microsoft Purview Data Map, the next step is to scan the data sources. The scanning process establishes a connection to the data source, captures technical metadata, and can automatically classify data using either the supported system classifications or rules for your custom classifications.

For example, if you have a file named *multiple.docx* and it has a National ID number in its content, during the scanning process Data Map adds the classification **EU National Identification Number** to the file asset's detail page.

These classifications help you and your team identify the kinds of data you have across your data estate. For example: if files or tables contain credit card numbers, or addresses. Then you can more easily search for certain kinds of information, like customer IDs, or prioritize security for sensitive data types.

Classifications can be automatically applied on file and column assets during scanning.

In this article we'll discuss:

- The steps to automatically apply classifications
- How Microsoft Purview classifies assets

Automatically apply classifications

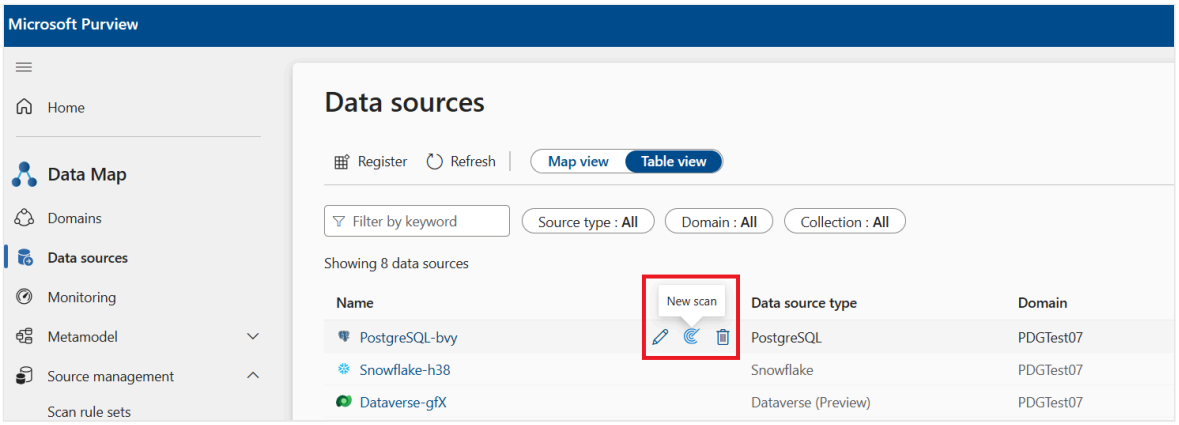
> **Note**
>
> Table assets are not automatically assigned classifications, because the classifications are assigned to their columns, but you can manually apply classifications to table assets.

After data sources are registered, you can automatically classify data in that source's data assets by running a scan.

1. Check the **Scan** section of the source article for your data source to confirm any prerequisites or authentication are set up and ready for a scan.

2. Search the Microsoft Purview Data Map the registered source that has the data assets (files and columns), you want to classify.

3. Select the **New Scan** icon under the resource.

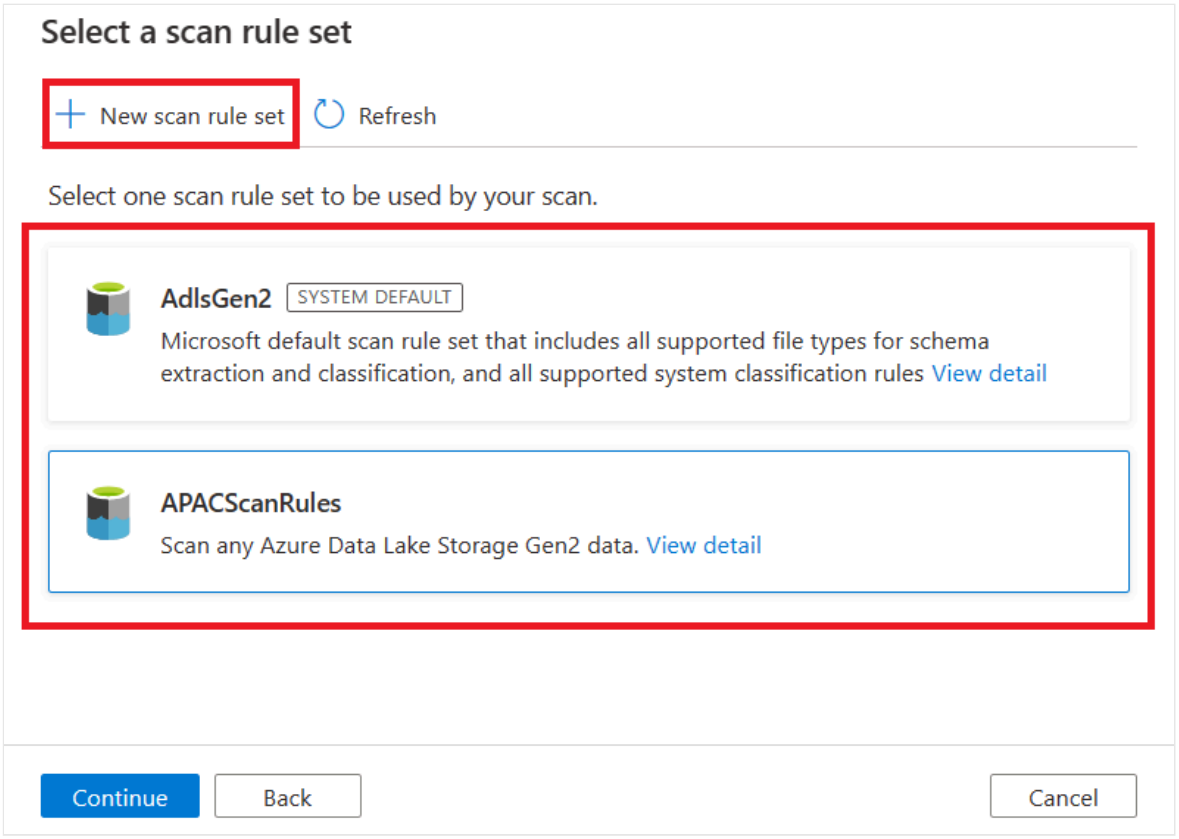


> **Tip**
>
> If you don't see the New Scan button, you may not have correct permissions. To run a scan, you'll need at least data source administrator permissions on the collection where the source is registered.

4. Select your credential and authenticate with your source. (For more information about authenticating with your source, see the **prerequisite** and **scan** sections of your specific source source article.) Select **Continue**.

5. If necessary, select the assets in the source you want to scan. You can scan all assets, or a subset of folders, files, or tables depending on the source.

6. Select your scan rule set. You'll see a list of available scan rule sets and can select one, or you can choose to create a new scan rule set using the **New scan rule set** button at the top. The scan rule set will determine which classifications will be compared and applied to your data. For more information, see how Microsoft Purview classifies assets.



> **Tip**
>
> For more information about the options available when creating a scan rule set, start at step 4 of these steps to create a scan rule set.

7. Schedule your scan.

8. Save and run your scan. Applicable classifications in your scan rule set will be automatically applied to the assets you scan. You'll be able to view and manage them once the scan is complete.

View classification details

Microsoft Purview captures important details like who applied a classification and when it was applied. To view the details, hover over the classification to revel the Classification details card. The classification details card shows the following information:

- Classification name - Name of the classification applied on the asset or column.
- Applied by - Who applied the classification. Possible values are scan and user name.
- Applied time - Local timestamp when the classification was applied via scan or manually.
- Classification type - System or custom.

Users with Data Curator role can see more details for classifications that were applied automatically through a scan. These details include sample count that the scanner read to classify the data, and distinct data count in the sample that the scanner found.

Impact of rescanning on existing classifications

Classification tags are automatically applied the first time based on sampling your data and matching those samples against predefined RegEx patterns or dictionaries. Subsequent rescans can preserve, remove, or incrementally add classification tags depending on several factors:

- A scan run doesn't remove classification tags that were manually applied.
- A scan run doesn't reapply classification tags that were manually removed.
- A later run of a scan that continues to use the same scan rule can update the classification tags.
- For a scan run, classification tags from the previous scans can get removed or retained based on the criteria listed in this table:

Expand table

File type	File size	Classification tags from previous scan will be:
File types with no size (for example, SQL)	All	Retained
.doc, .docm, .docx, .dot, .odp, .ods, .odt, .pdf, .pot, .pps, .ppsx, .ppt, .pptm, .pptx, .txt, .xlc, .xls, .xlsb, .xlsm, .xlsx, .xlt	>20 MB	Retained
.doc, .docm, .docx, .dot, .odp, .ods, .odt, .pdf, .pot, .pps, .ppsx, .ppt, .pptm, .pptx, .txt, .xlc, .xls, .xlsb, .xlsm, .xlsx, .xlt	<20 MB	Removed
GZ	>400 KB	Retained
GZ	<400 KB	Removed
Files with no extension or structured file types	> 1 MB	Retained
Files with no extension or structured file types	< 1 MB	Removed

How Microsoft Purview classifies assets

When a data source is scanned, Microsoft Purview compares data in the asset to a list of possible classifications called a scan rule set.

There are system scan rule sets already available for each data source that contains every currently available system classification for that data source. Or, you can create a custom scan rule set to make a list of classifications tailored to your data set.

Making a custom rule sets for your data can be a good idea if your data is limited to specific kinds of information, or regions, as comparing your data to fewer classification types will speed up the scanning process. For example, if your dataset only contains European data, you could create a custom scan rule set that excludes identification for other regions.

You might also make a custom rule set if you've created custom classifications and classification rules, so that your custom classifications can be automatically applied during scanning.

For more information about the available system classifications and how your data is classified, see the list of system classifications.

Next steps

- Learn how to create a custom classification.
- Learn how to manually apply classifications.

Feedback

Was this page helpful? Yes No

Custom classifications in Data Map

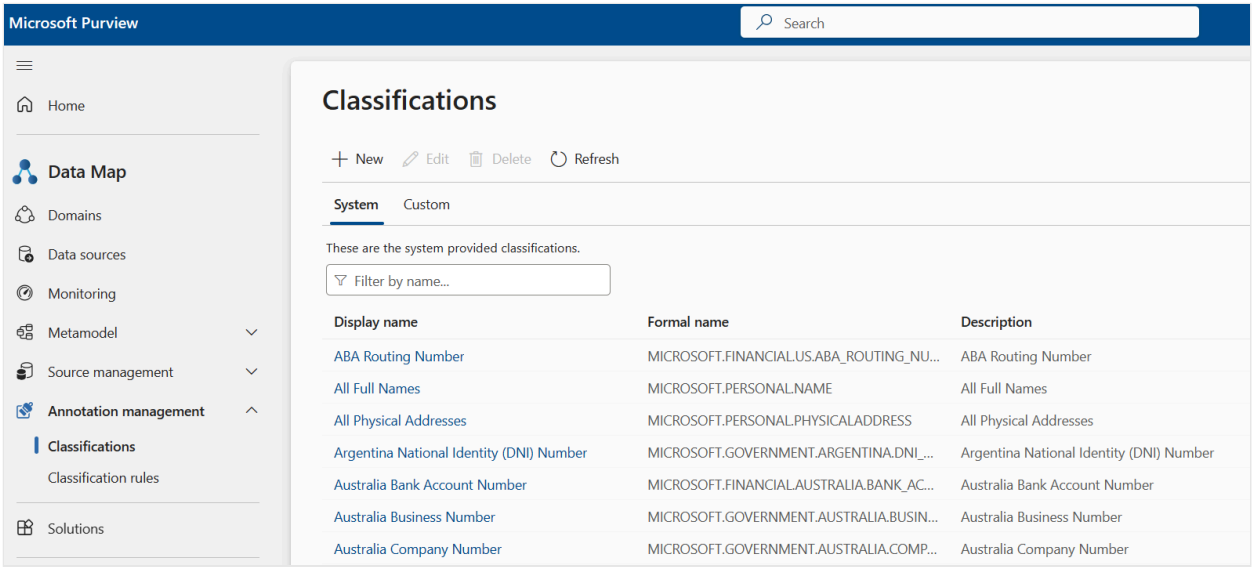
Article • 03/04/2025

This article describes how you can create custom classifications in Microsoft Purview Data Map to define data types in your data estate that are unique to your organization. It also describes the creation of custom classification rules that let you find specified data throughout your data estate.

> **Important**
>
> To create a custom classification you need either **data curator** or **data source administrator** permission on a domain or collection. Permissions at any collection level are sufficient. For more information about permissions, see Microsoft Purview data governance permissions.

Default system classifications

Data Map provides a large set of default system classifications that represent typical personal data types that you might have in your data estate. See the list of available system classifications in Data Map.



You also have the ability to create custom classifications, if any of the default classifications don't meet your needs.

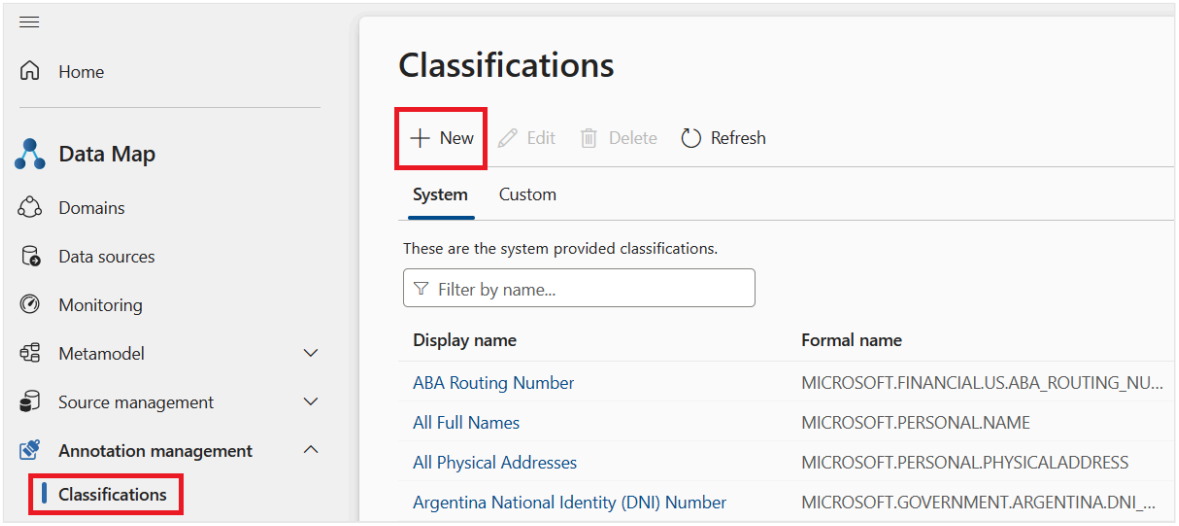
> **Note**

- Our data sampling rules are applied to both system and custom classifications.
- Custom classifications are only applied to structured data sources like SQL and CosmosDB, and to structured file types like CSV, JSON, and Parquet. Custom classification isn't applied to unstructured data file types like DOC, PDF, and XLSX.

Steps to create a custom classification

To create a custom classification, follow these steps:

1. You'll need **data curator** or **data source administrator** permissions on any collection to be able to create a custom classification.
2. In the Microsoft Purview portal, open the **Data map** solution.
3. Select **Annotation management** select **Classifications**.
4. Select + **New**.



The **Add new classification** pane opens, where you can give your classification a name and a description. It's good practice to use a name-spacing convention, such as `your company name.classification name`.

The Microsoft system classifications are grouped under the reserved `MICROSOFT.` namespace. An example is **MICROSOFT.GOVERNMENT.US.SOCIAL_SECURITY_NUMBER**.

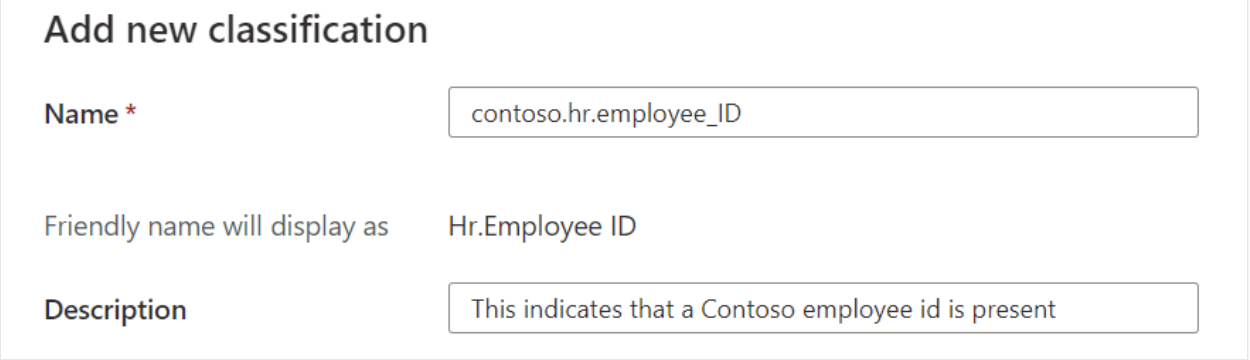
The name of your classification must start with a letter followed by a sequence of letters, numbers, and period (.) or underscore characters. As you type, the UX automatically

generates a friendly name. This friendly name is what users see when you apply it to an asset in the catalog.

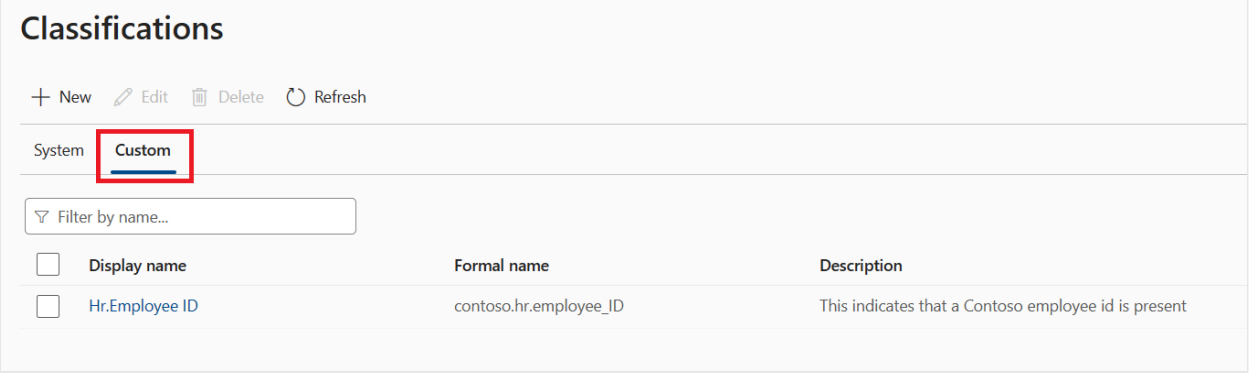
To keep the name short, the system creates the friendly name based on the following logic:

- All but the last two segments of the namespace are trimmed.
- The casing is adjusted so that the first letter of each word is capitalized.
- All underscores (_) are replaced with spaces.

As an example, if you named your classification **contoso.hr.employee_ID**, the friendly name is stored in the system as **Hr.Employee ID**.

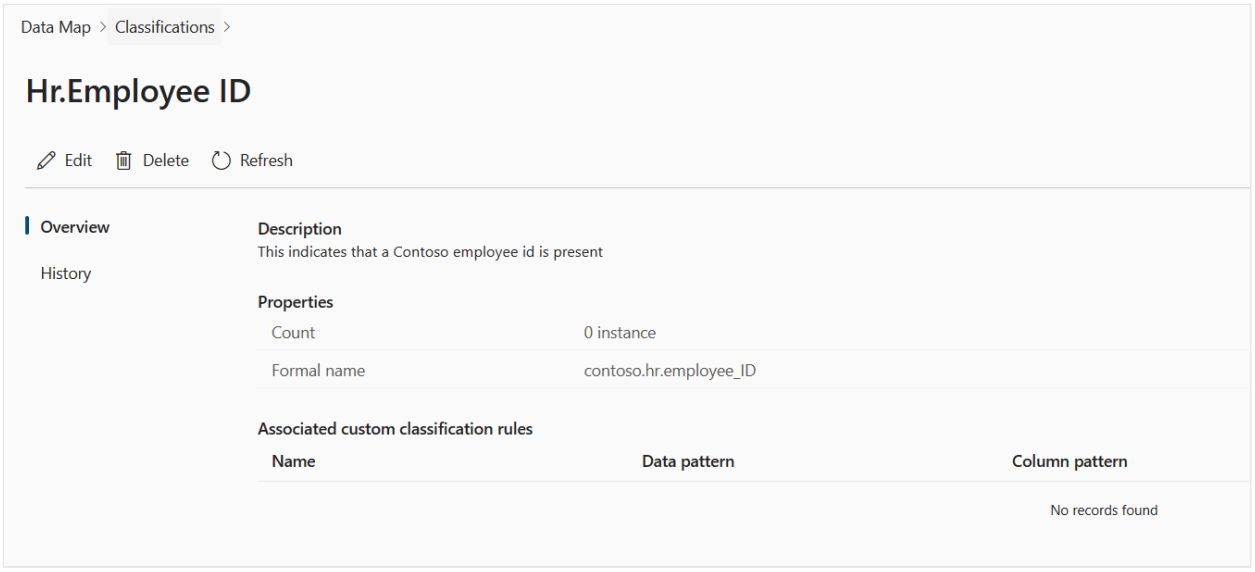


Select **OK**, and your new classification is added to your **Custom** classification list.



Selecting the classification in the list opens the classification details page. Here, you find all the details about the classification.

These details include the count of how many instances there are, the formal name, associated classification rules (if any), and the owner name.



Custom classification rules

The data map provides a set of default classification rules that are used by the scanner to automatically detect certain data types. You can also add your own custom classification rules to detect other types of data that you might be interested in finding across your data estate. This capability can be powerful when you're trying to find data within your data estate.

> **Note**
>
> Custom classification rules are only supported in the English language.

As an example, let's say that a company named Contoso has employee IDs that are standardized throughout the company with the word "Employee" followed by a GUID to create EMPLOYEE{GUID}. For example, one instance of an employee ID looks like `EMPLOYEE9c55c474-9996-420c-a285-0d0fc23f1f55`.

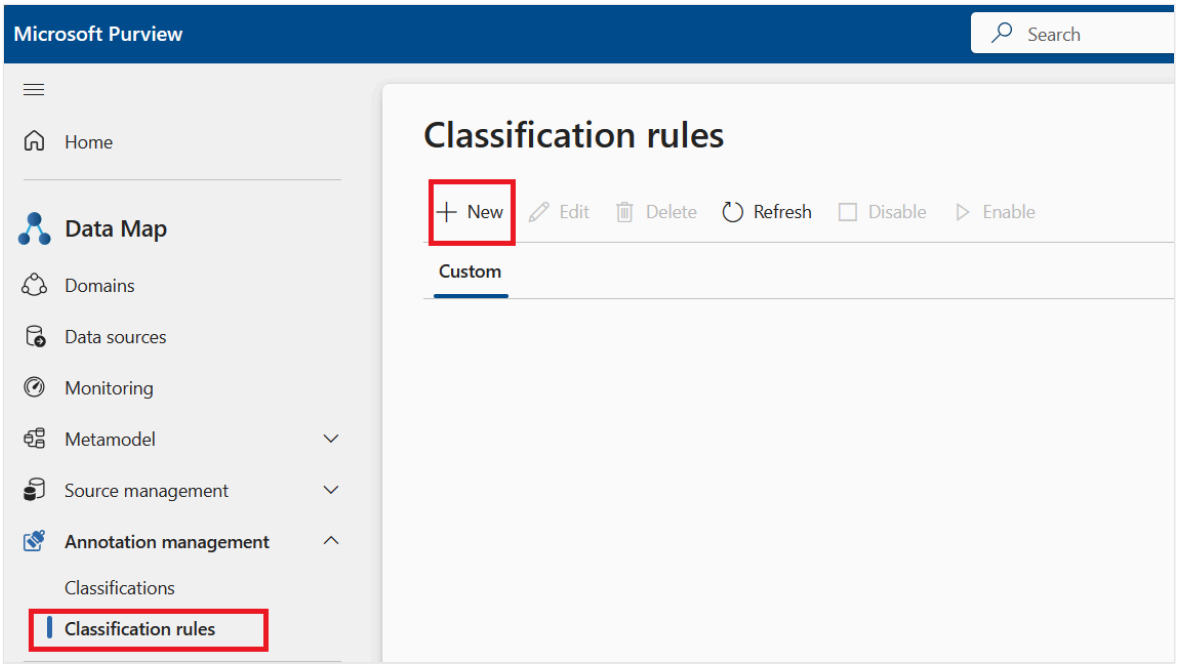
Contoso can configure the scanning system to find instances of these IDs by creating a custom classification rule. They can supply a regular expression that matches the data pattern, in this case, `\^Employee\[A-Za-z0-9\]{8}-\[A-Za-z0-9\]{4}-\[A-Za-z0-9\]{4}-\[A-Za-z0-9\]{4}-\[A-Za-z0-9\]{12}\$`. Optionally, if the data usually is in a column that they know the name of, such as Employee_ID or EmployeeID, they can add a column pattern regular expression to make the scan even more accurate. An example regex is Employee_ID|EmployeeID.

The scanning system can then use this rule to examine the actual data in the column and the column name to try to identify every instance of where the employee ID pattern is found.

Steps to create a custom classification rule

Create a custom classification that following the instructions in the previous section. You'll add this custom classification in the classification rule configuration so that the system applies it when it finds a match in the column.

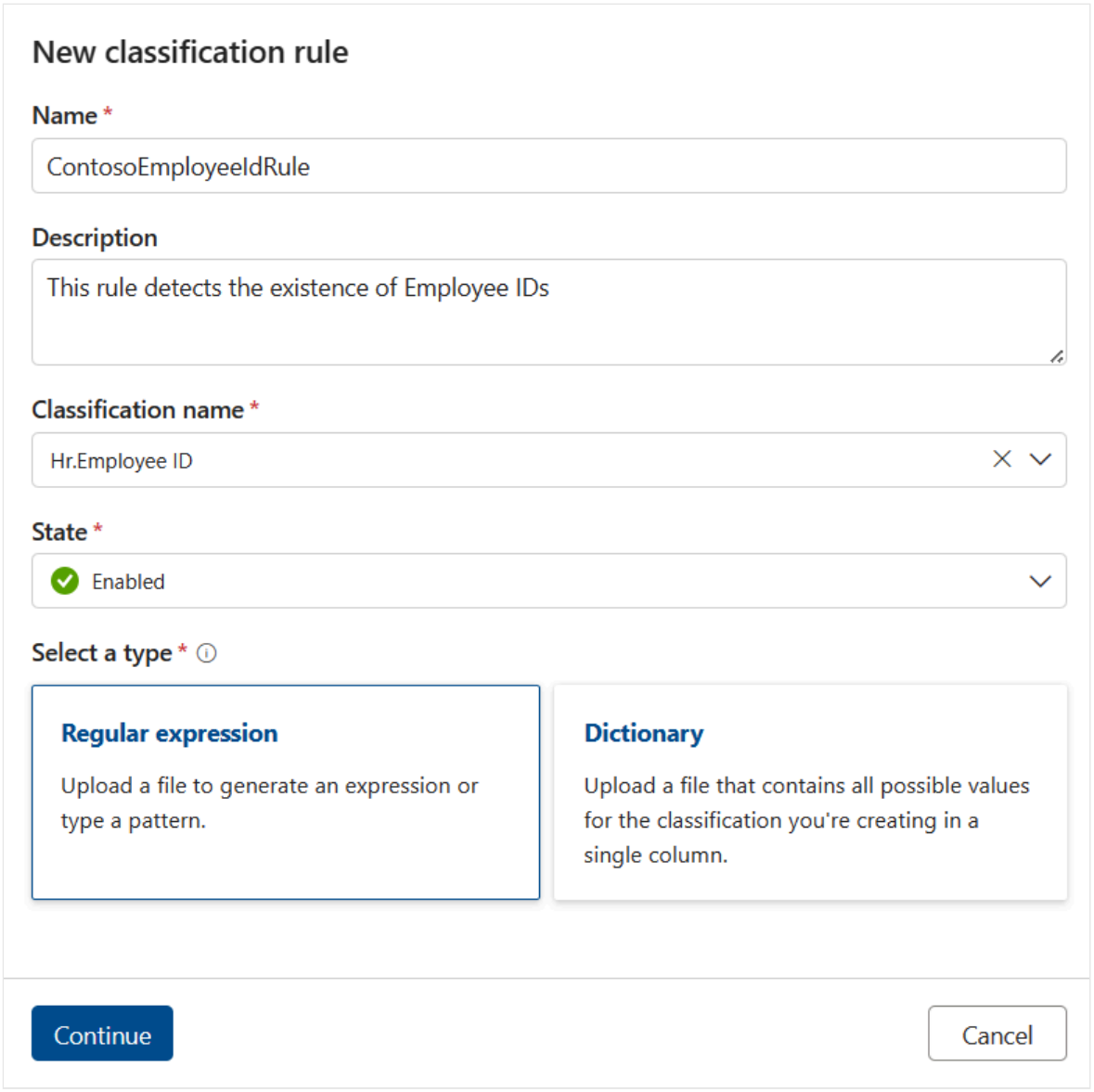
1. In the Microsoft Purview portal, open the **Data map** solution.
2. Under **Annotation management**, select the **Classifications rules** section.
3. Select **New**.



4. The **New classification rule** dialog box opens. Fill in the fields and decide whether to create a **regular expression rule** or a **dictionary rule**.

Expand table

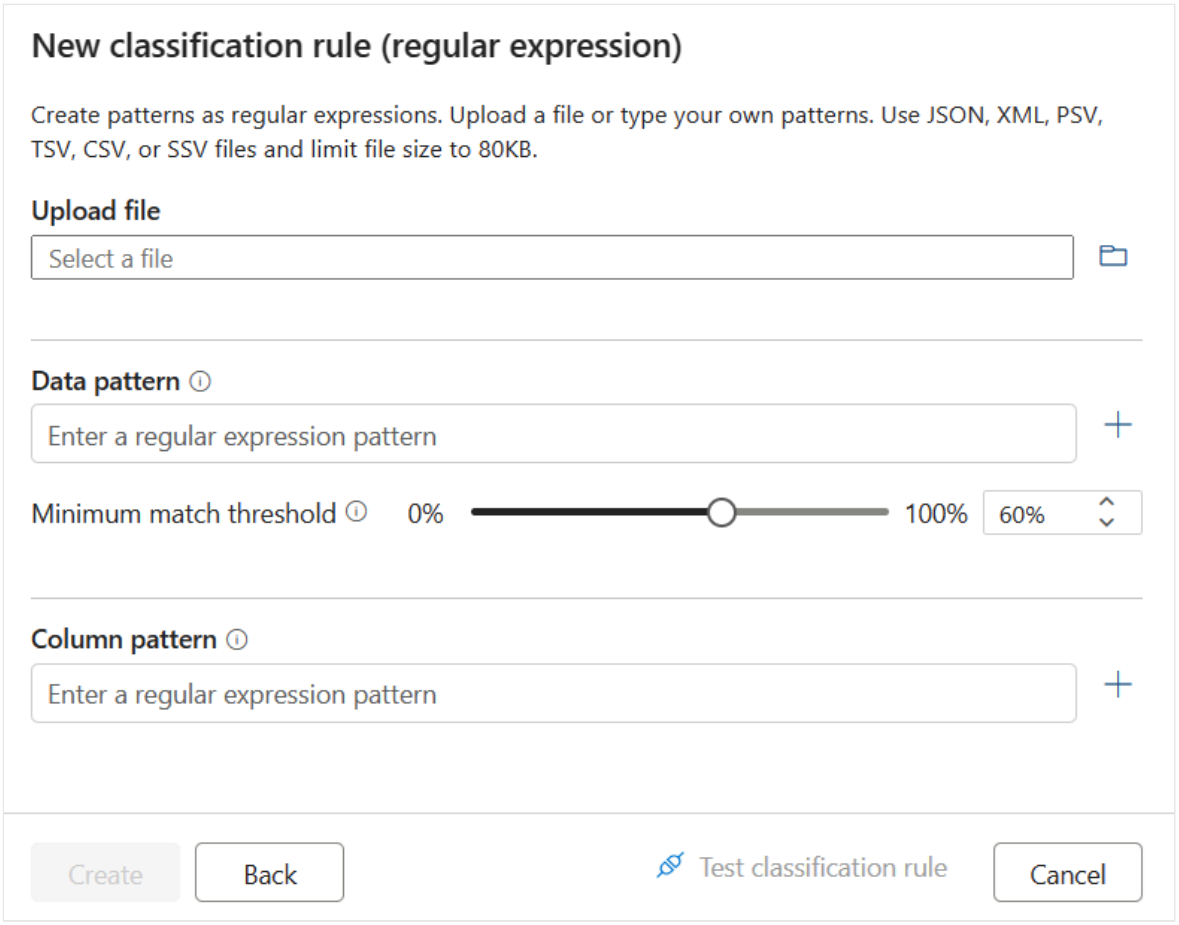
Field	Description
Name	Required. The maximum is 100 characters.
Description	Optional. The maximum is 256 characters.
Classification Name	Required. Select the name of the classification from the drop-down list to tell the scanner to apply it if a match is found.
State	Required. The options are enabled or disabled. Enabled is the default.



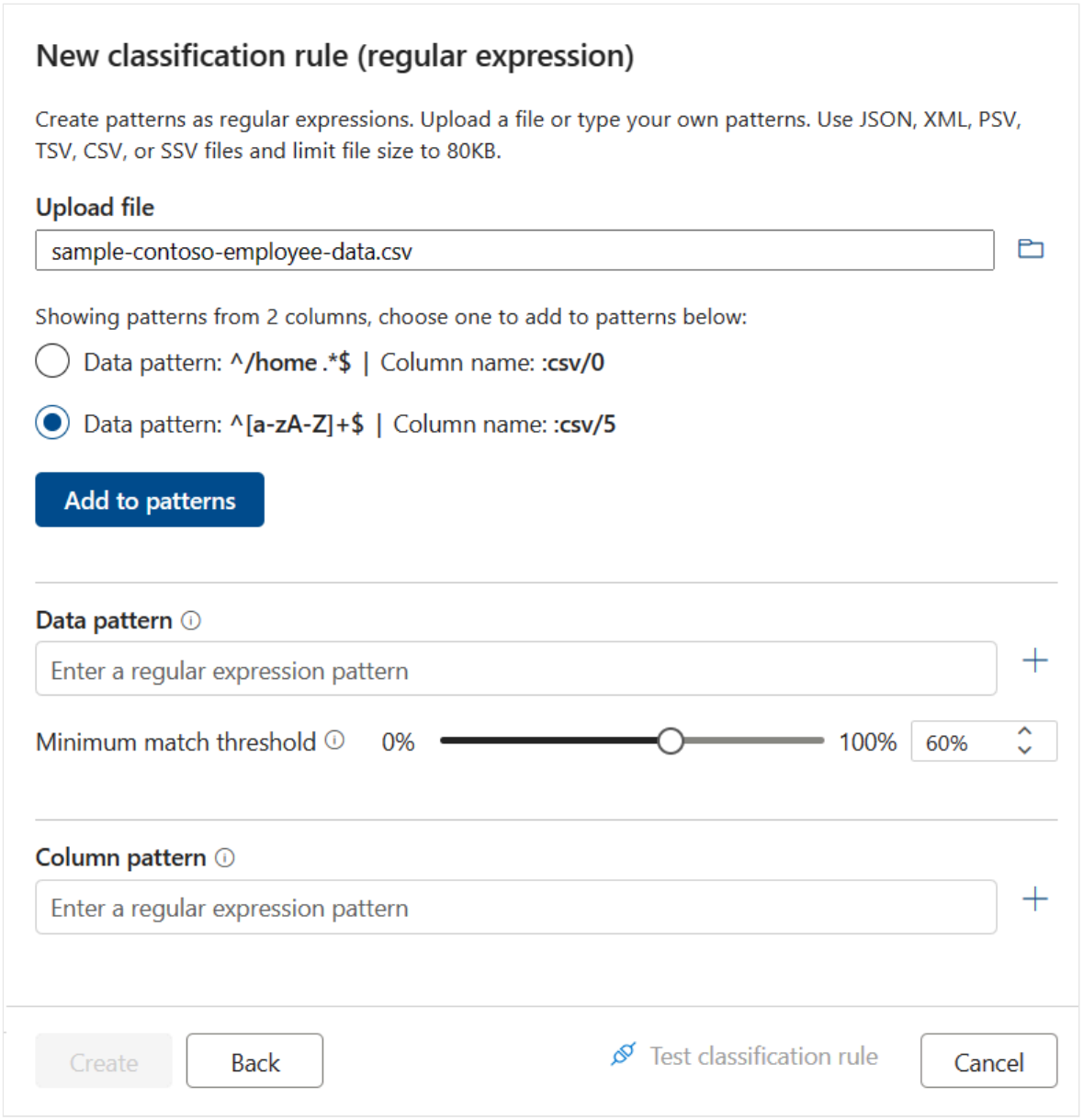
Creating a Regular Expression Rule

> **ⓘ Important**
>
> Regular expressions in custom classifications are case insensitive.

1. If creating a regular expression rule, you'll see the following screen. You could optionally upload a file that will be used to **generate suggested regex patterns** for your rule. Only English language rules are supported.



2. If you decide to generate a suggested regex pattern, after uploading a file, select one of the suggested patterns and select **Add to Patterns** to use the suggested data and column patterns. You can tweak the suggested patterns or you may also type your own patterns without uploading a file.



Expand table

Field	Description
Data Pattern	Optional. A regular expression that represents the data that's stored in the data field. The limit is large. In the previous example, the data patterns test for an employee ID that's literally the word `Employee{GUID}`.
Column Pattern	Optional. A regular expression that represents the column names that you want to match. The limit is large.

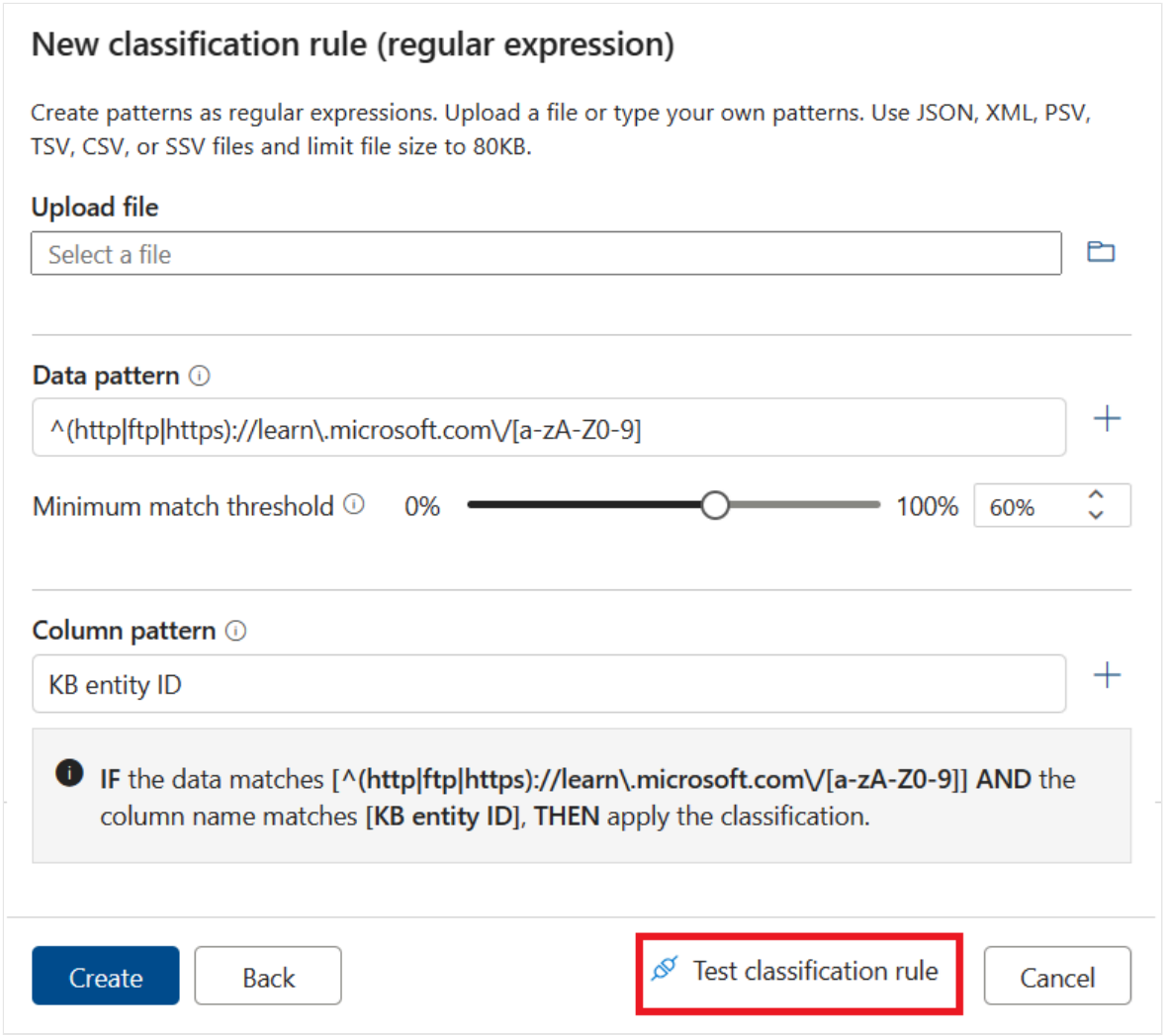
3. Under **Data Pattern** you can use the **Minimum match threshold** to set the minimum percentage of the distinct data value matches in a column that must be found by the scanner for the classification to be applied. The suggested value is 60%. If you specify multiple data patterns, this setting is disabled and the value is fixed at 60%.

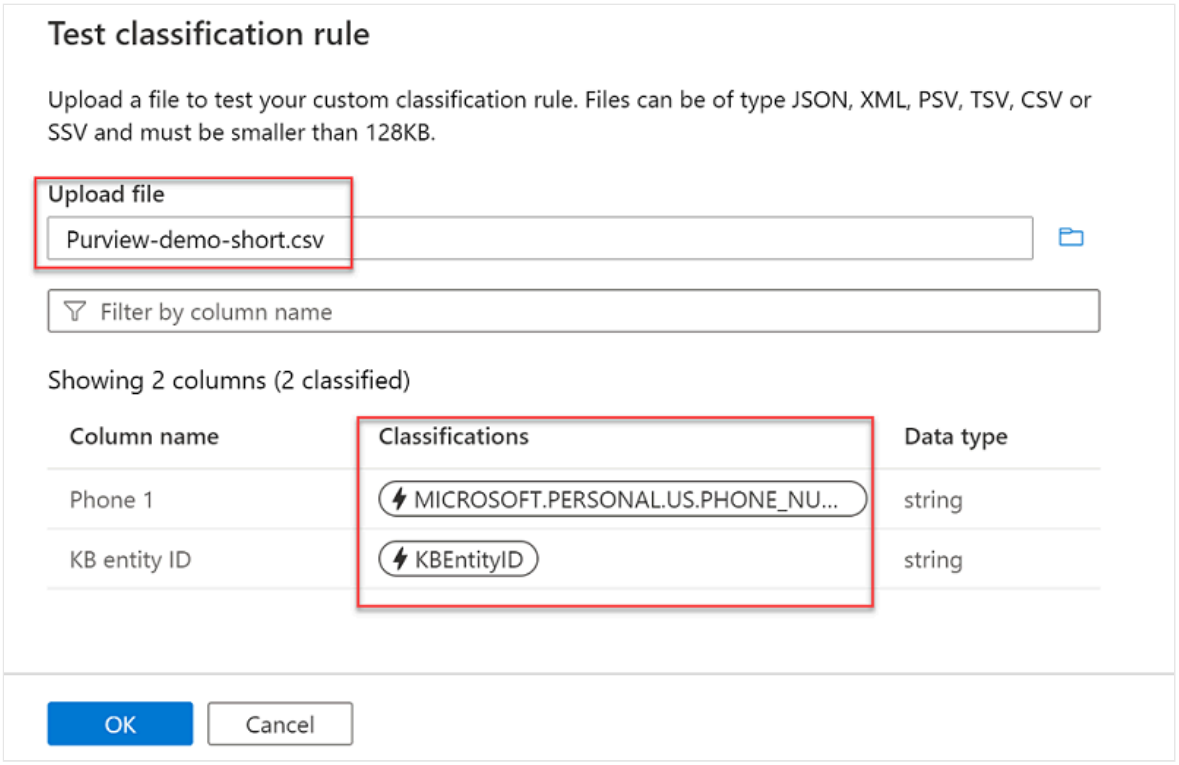
> **Note**
>
> The Minimum match threshold must be at least 1%.

4. You can now verify your rule and **create** it.

5. Test the classification rule before completing the creation process to validate that it will apply tags to your assets. The classifications in the rule will be applied to the sample data uploaded just as it would in a scan. This means all of the system classifications and your custom classification will be matched to the data in your file.

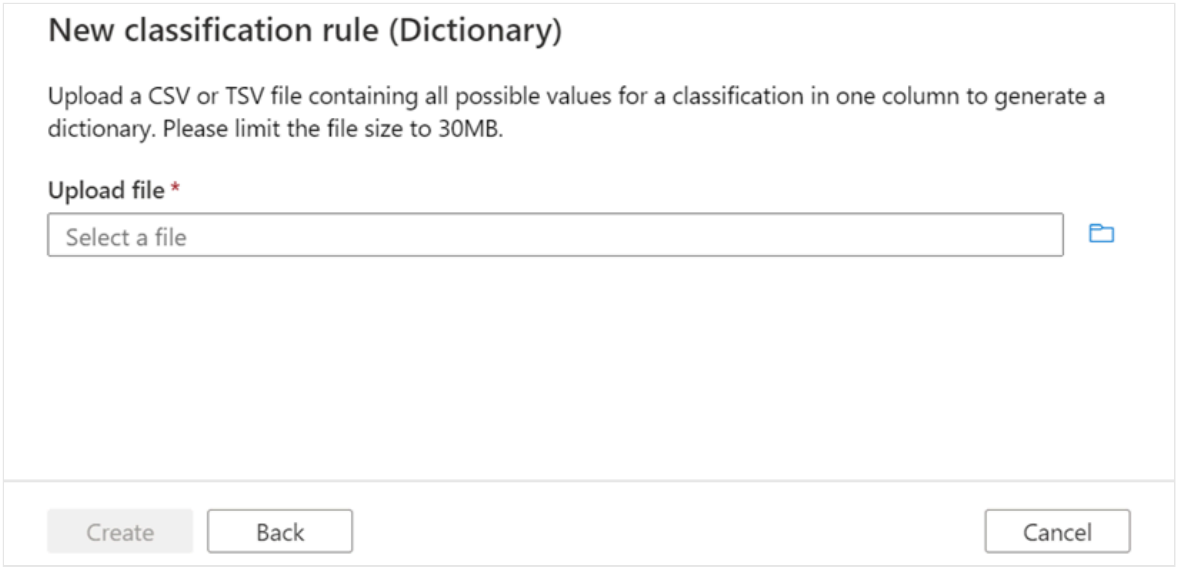
   Input files could include delimited files (CSV, PSV, SSV, TSV), JSON, or XML content. The content will be parsed based on the file extension of the input file. Delimited data might have a file extension that matches any of the mentioned types. For example, TSV data can exist in a file named MySampleData.csv. Delimited content must also have a minimum of three columns.



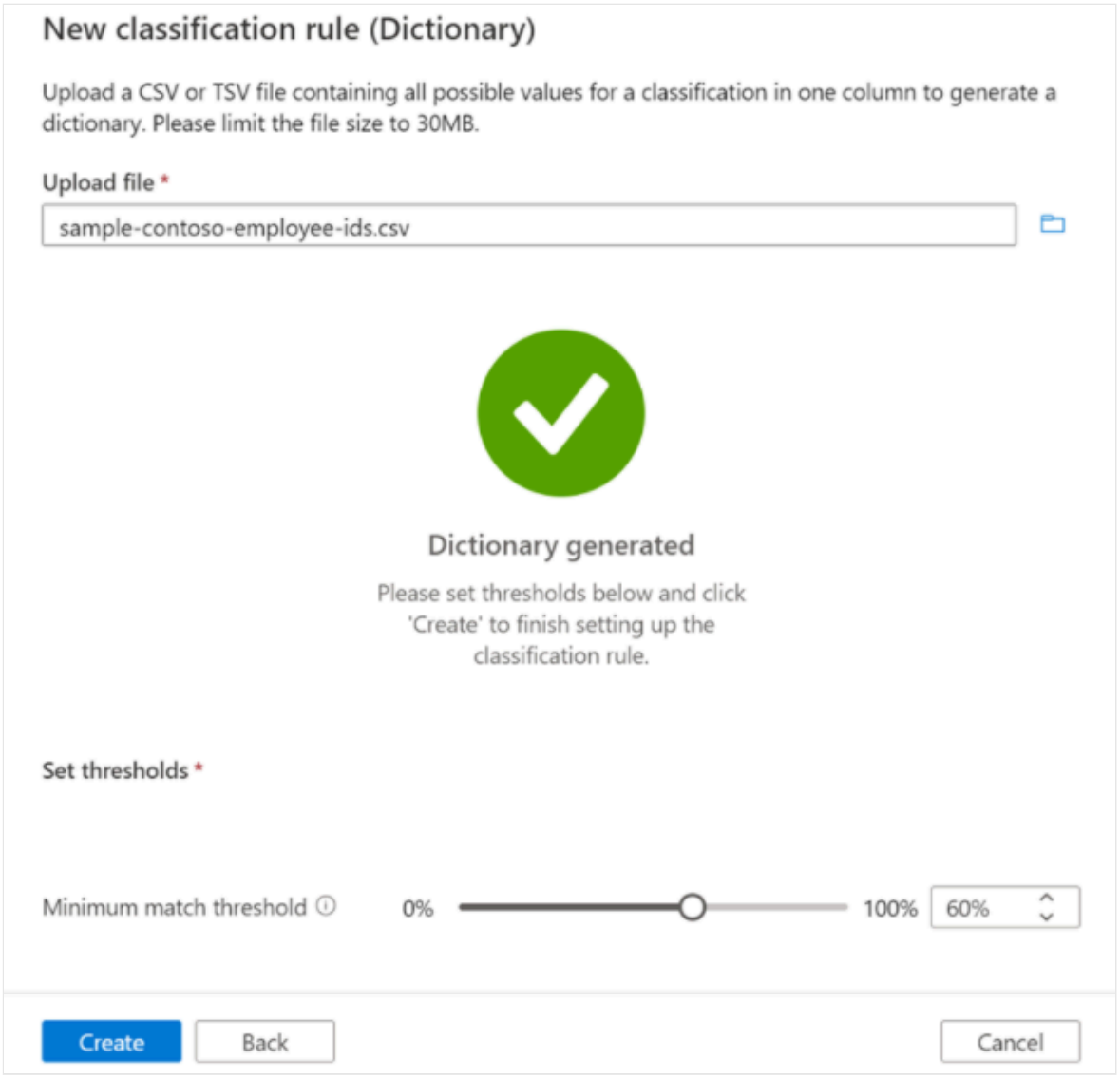


Creating a Dictionary Rule

1. If creating a dictionary rule, you'll see the following screen. Upload a file that contains all possible values for the classification you're creating in a single column. Only English language rules are supported.



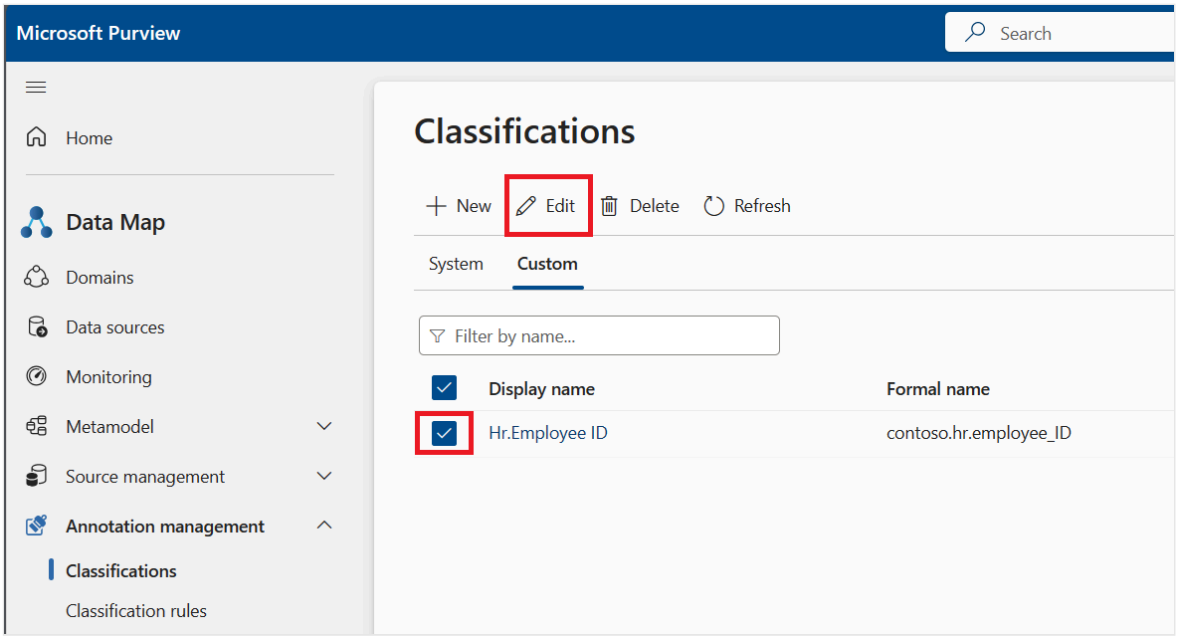
2. After the dictionary is generated, you can adjust the minimum match threshold and submit the rule.



Edit or delete a custom classification

To update or edit a custom classification, follow these steps:

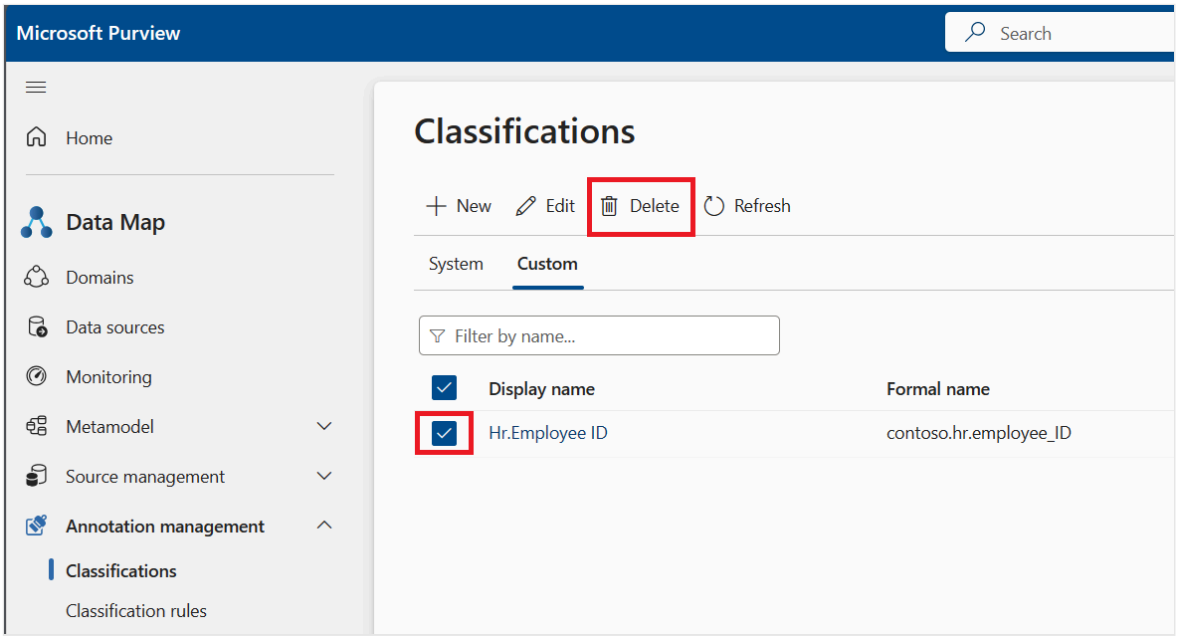
1. In the Microsoft Purview portal, open the **Data map** solution.
2. Under **Annotation management** select **Classifications**.
3. Select the **Custom** tab.
4. Select the classification you want to edit, then select the **Edit** button.



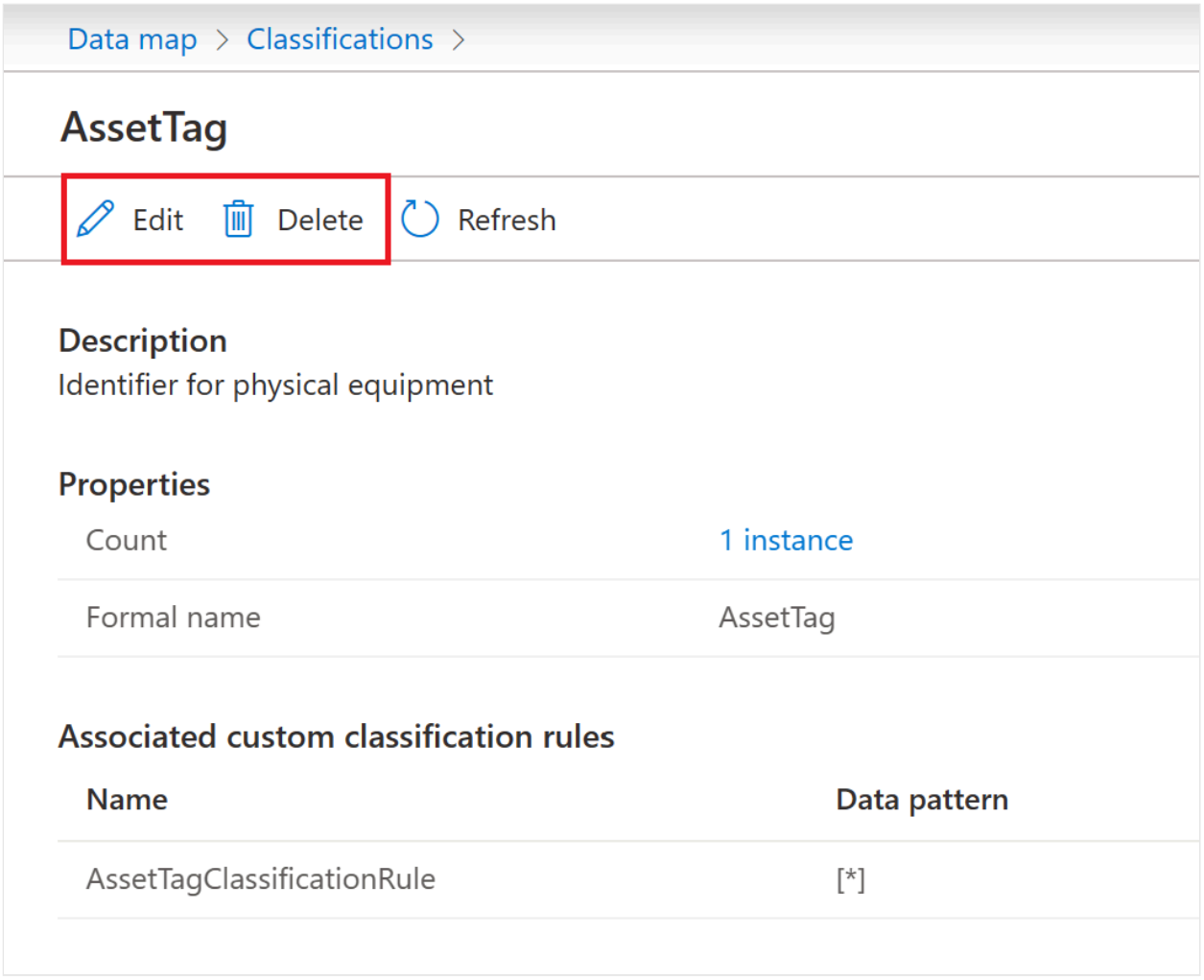
5. Now can edit the description of this custom classification. Select the **Ok** button when you're finished to save your changes.

To delete a custom classification:

1. After opening the **Data map**, and then **Classifications**, select the **Custom** tab.
2. Select the classification you want to delete, or multiple classifications you want to delete, and then select the **Delete** button.

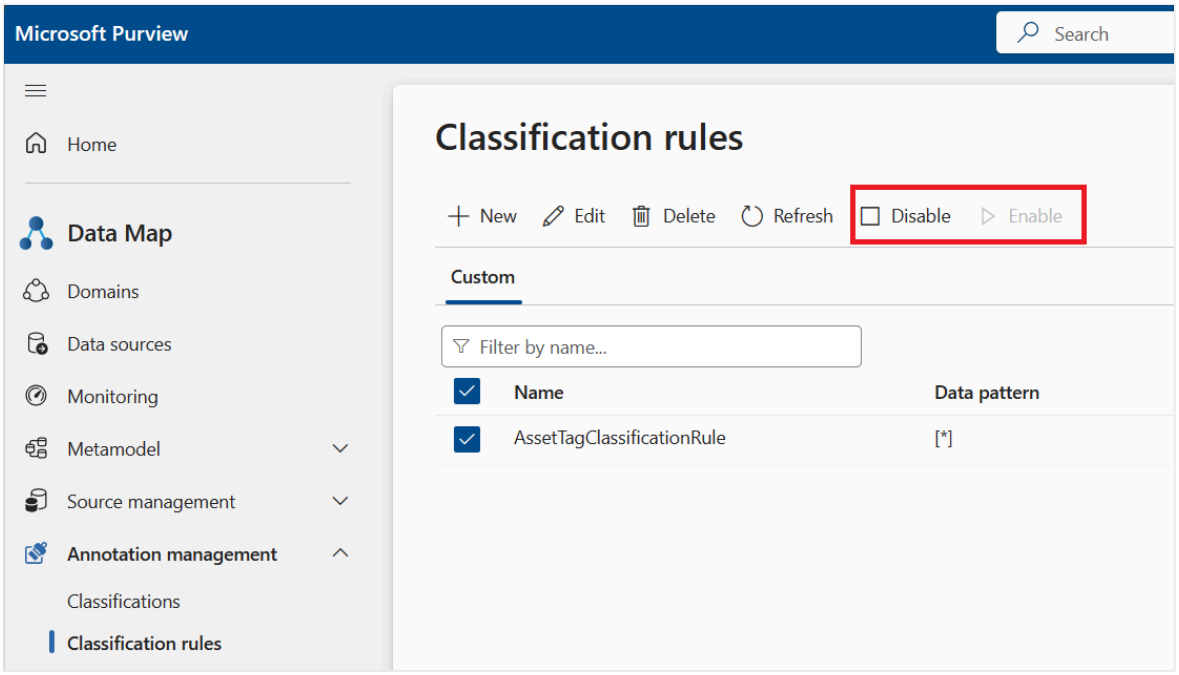


You can also edit or delete a classification from inside the classification itself. Just select your classification, then select the **Edit** or **Delete** buttons in the top menu.



Enable or disable classification rules

1. In your Microsoft Purview account, select the **Data map**, and then **Classification rules**.
2. Select the **Custom** tab.
3. You can check the current status of a classification rule by looking at the **Status** column in the table.
4. Select the classification rule, or multiple classification rules, that you want to enable or disable.
5. Select either the **Enable** or **Disable** buttons in the top menu.



You can also update the status of a rule when editing the rule.

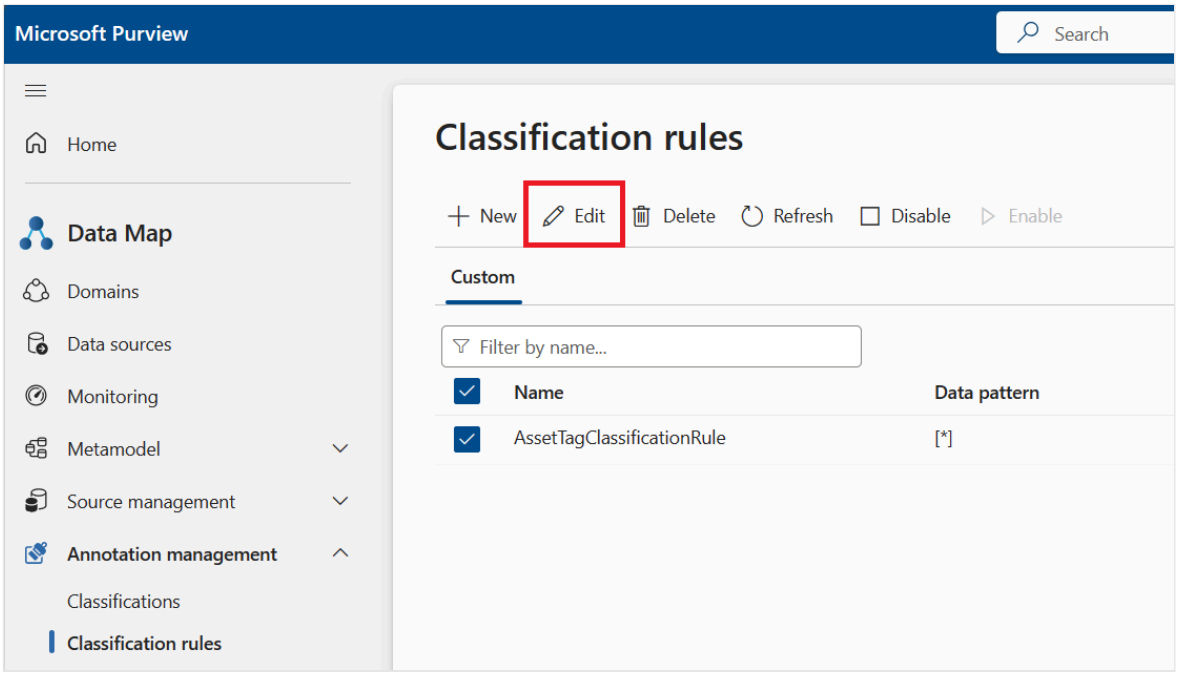
Edit or delete a classification rule

To update or edit a custom classification rule, follow these steps:

1. In your Microsoft Purview account, select the **Data map**, and then **Classification rules**.

2. Select the **Custom** tab.

3. Select the classification rule you want to edit, then select the **Edit** button.



4. Now you can edit the state, the description, and the associated classification rule.

5. Select the **Continue** button.

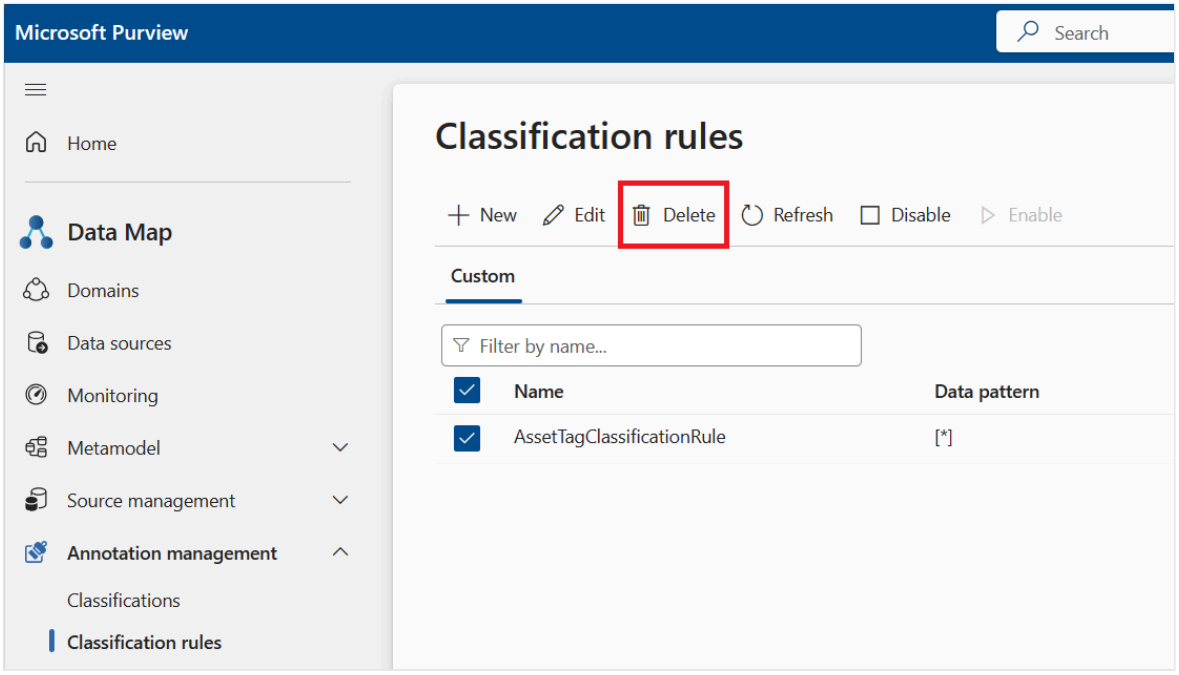
6. You can upload a new file for your regular expression or dictionary rule to match against, and update your match threshold and column pattern match.

7. Select **Apply** to save your changes. Scans will need to be rerun with the new rule to apply changes across your assets.

To delete a custom classification:

1. After opening the **Data map**, and then **Classification rules**, select the **Custom** tab.

2. Select the classification rule you want to delete, and then select the **Delete** button.



Next steps

Now that you've created your classification rule, it's ready to be added to a scan rule set so that your scan uses the rule when scanning. For more information, see Create a scan rule set.

Feedback

Was this page helpful? Yes No

Manually apply classifications on assets in Data Map

Article • 03/04/2025

This article discusses how to manually apply classifications on assets in Microsoft Purview Data Map.

Classifications are logical labels to help you and your team identify the kinds of data you have across your data estate. For example: if files or tables contain credit card numbers, or addresses.

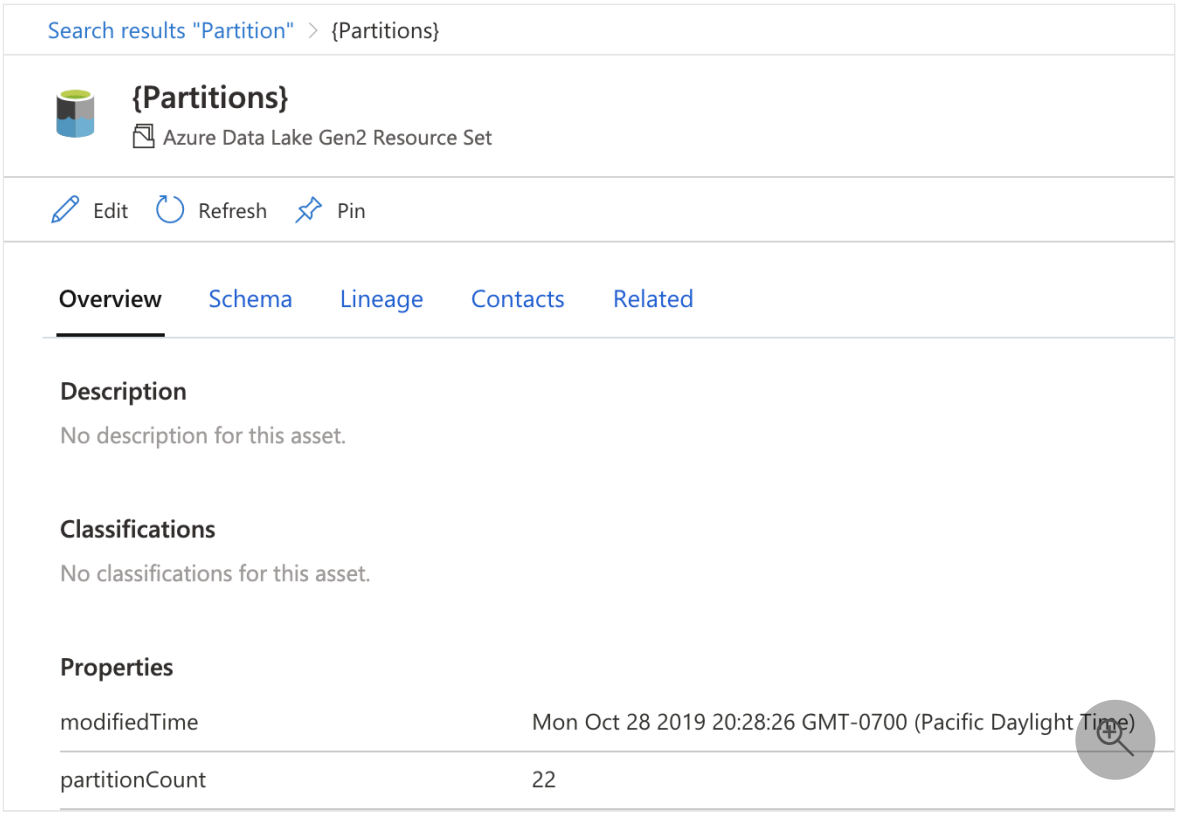
Microsoft Purview automatically applies classifications to some assets during the scanning process, but there are some scenarios when you may want to manually apply more classifications. For example, Microsoft Purview doesn't automatically apply classifications to table assets (only their columns), or you might want to apply custom classifications, or add classifications to assets grouped a resource set.

> ⓘ **Note**
>
> Some custom classifications can be automatically applied after setting up a custom classification rule.

Follow the steps in this article to manually apply classifications to file, table, and column assets.

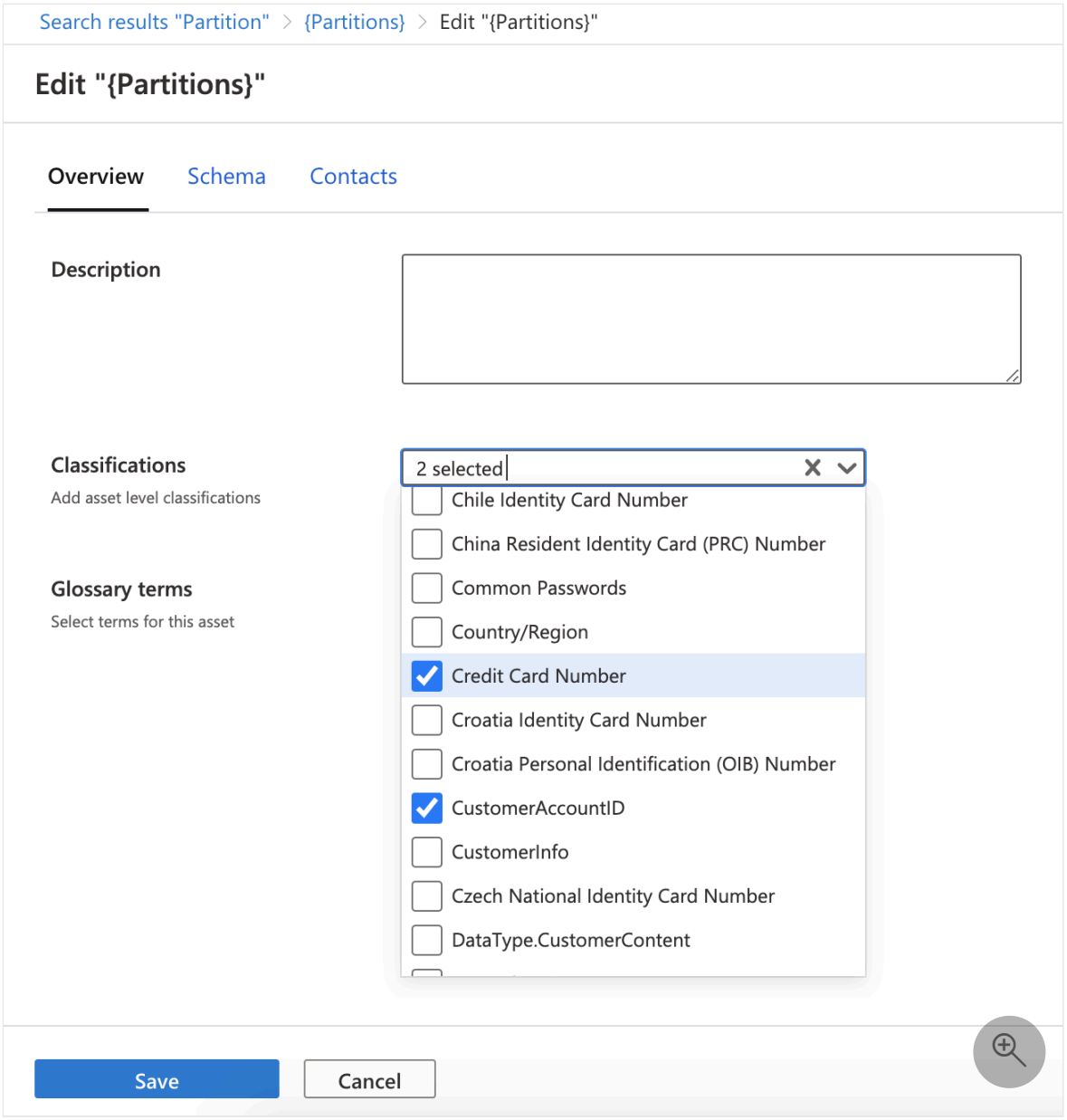
Manually apply classification to a file asset

1. Search or browse Microsoft Purview Unified Catalog for the file you're interested in and navigate to the asset detail page.



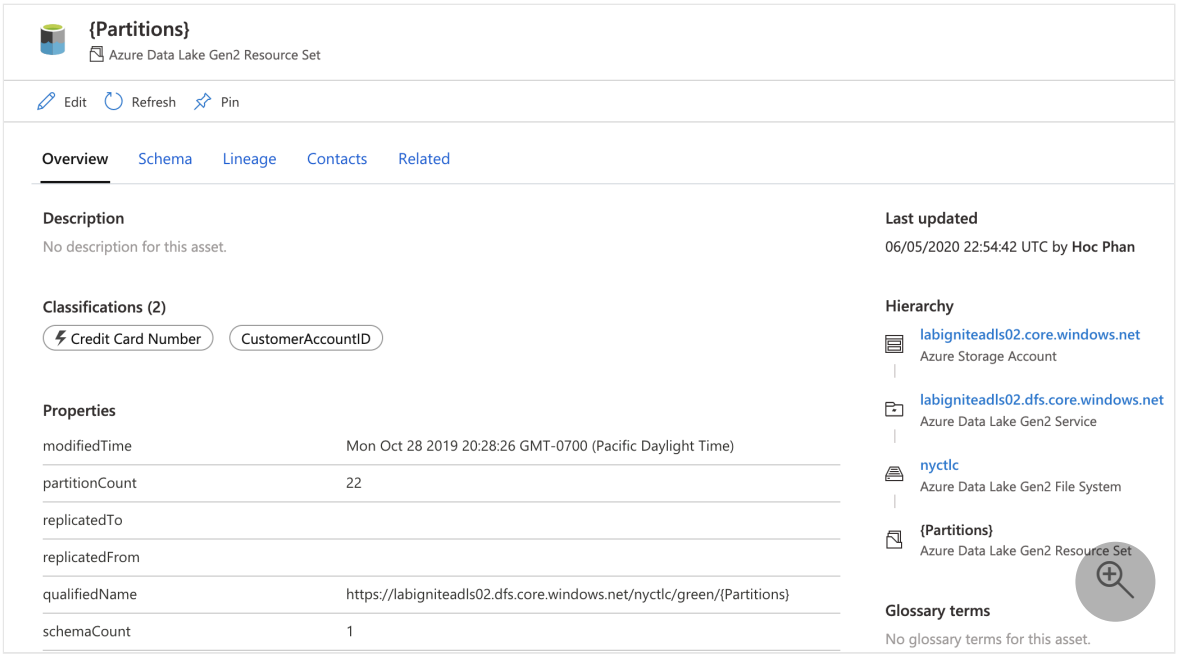
2. On the **Overview** tab, view the **Classifications** section to see if there are any existing classifications. Select **Edit**.

3. From the **Classifications** drop-down list, select the specific classifications you're interested in. In our example, we're adding **Credit Card Number**, which is a system classification and **CustomerAccountID**, which is a custom classification.



4. Select **Save**.

5. On the **Overview** tab, confirm that the classifications you selected appear under the **Classifications** section.



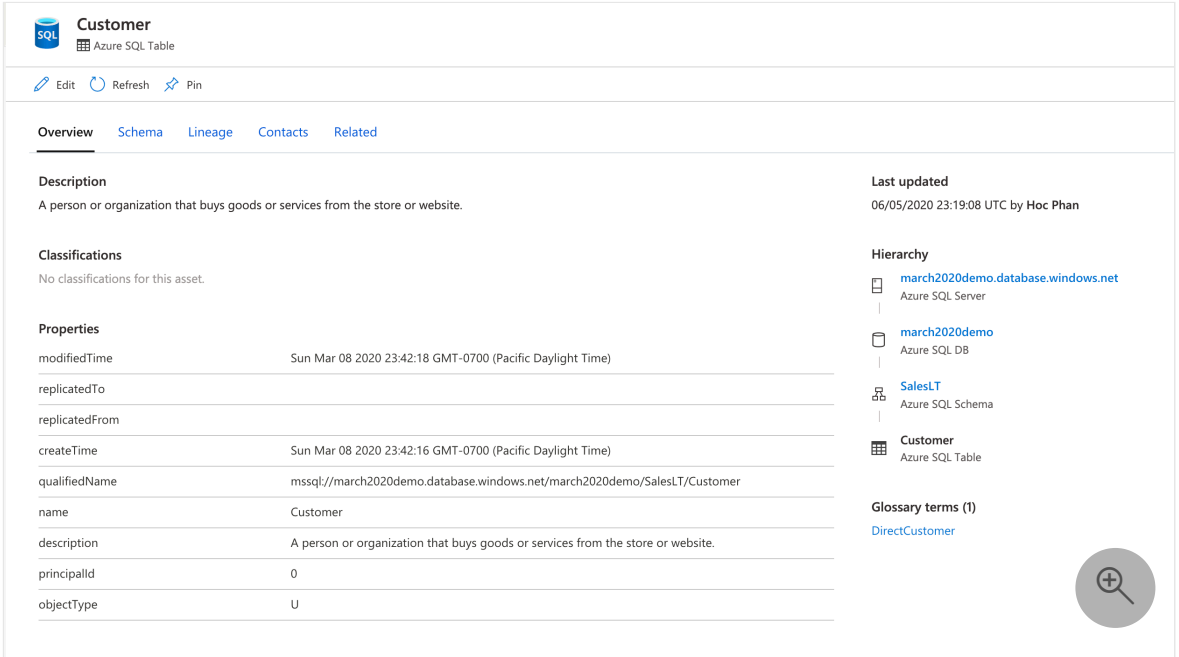
Manually apply classification to a table asset

When Microsoft Purview scans your data sources, it doesn't automatically assign classifications to table assets (only on columns). For a table asset to have classifications, you must add them manually:

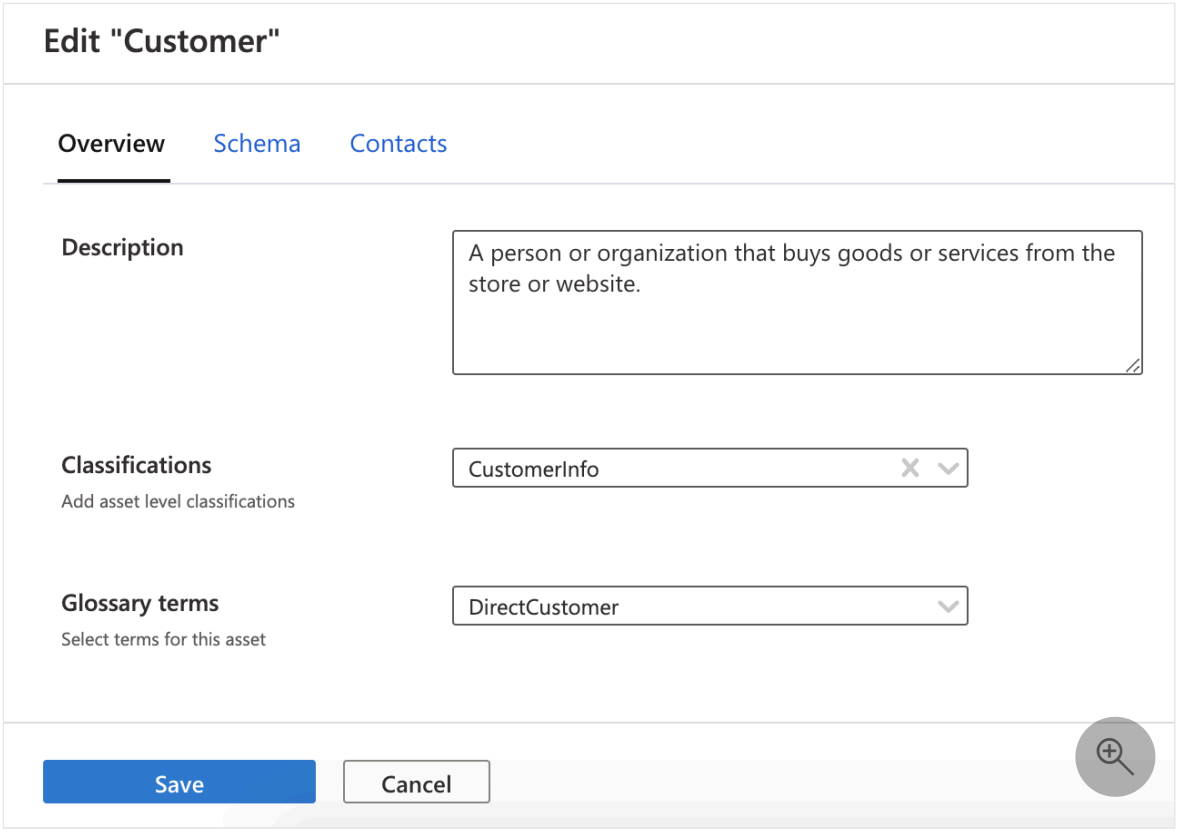
To add a classification to a table asset:

1. Search or browse Unified Catalog for the table asset that you're interested in. For example, **Customer** table.

2. Confirm that no classifications are assigned to the table. Select **Edit**.

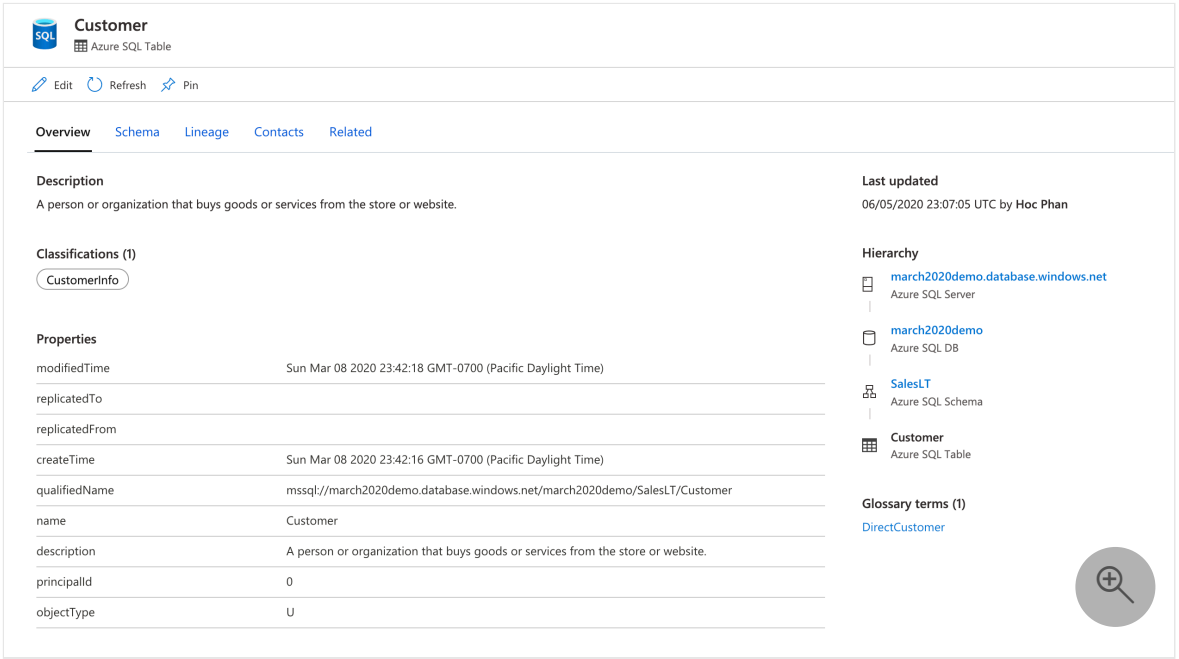


3. From the **Classifications** drop-down list, select one or more classifications. This example uses a custom classification named **CustomerInfo**, but you can select any classifications for this step.



4. Select **Save** to save the classifications.

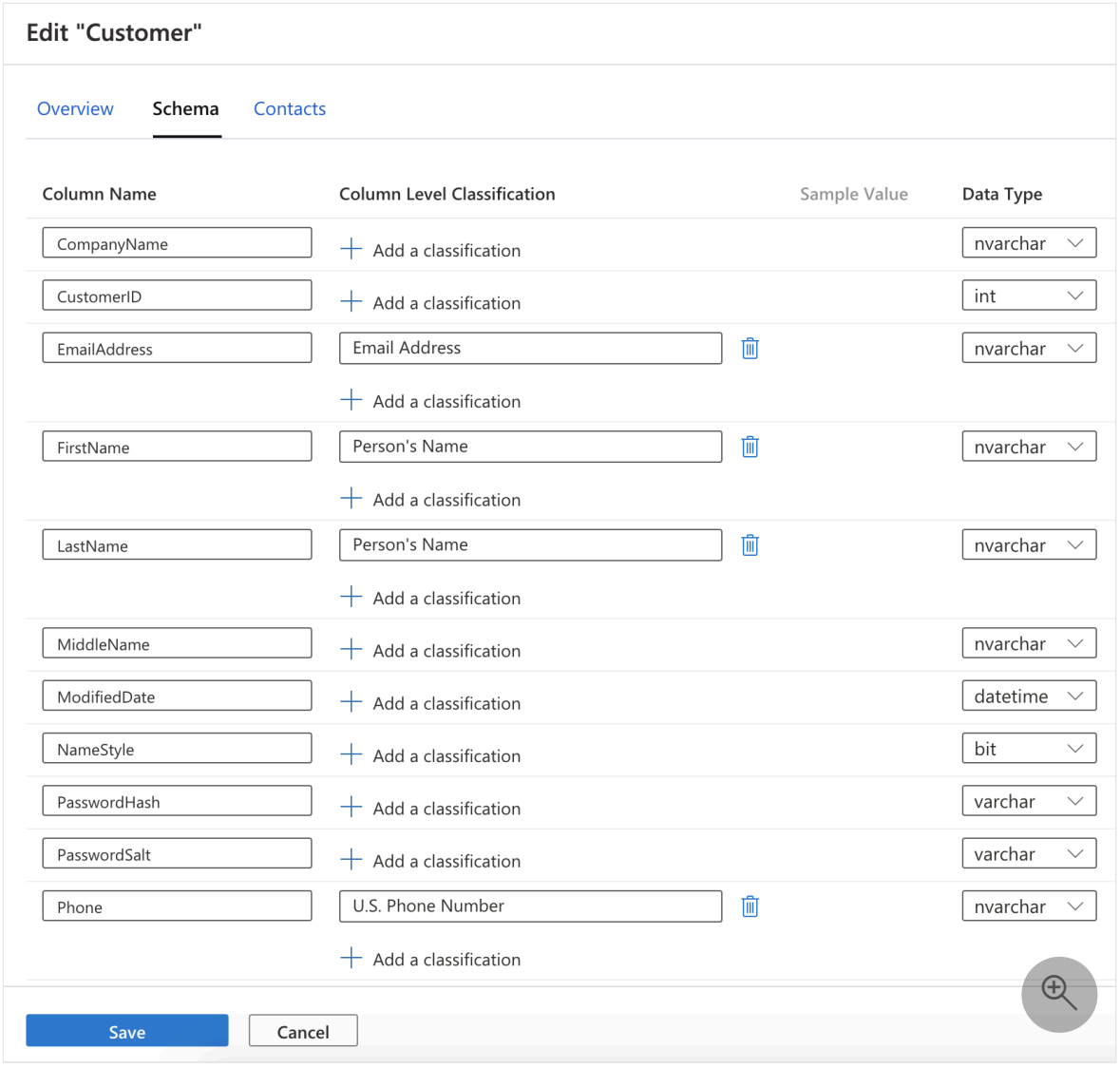
5. On the **Overview** page, verify that Microsoft Purview added your new classifications.



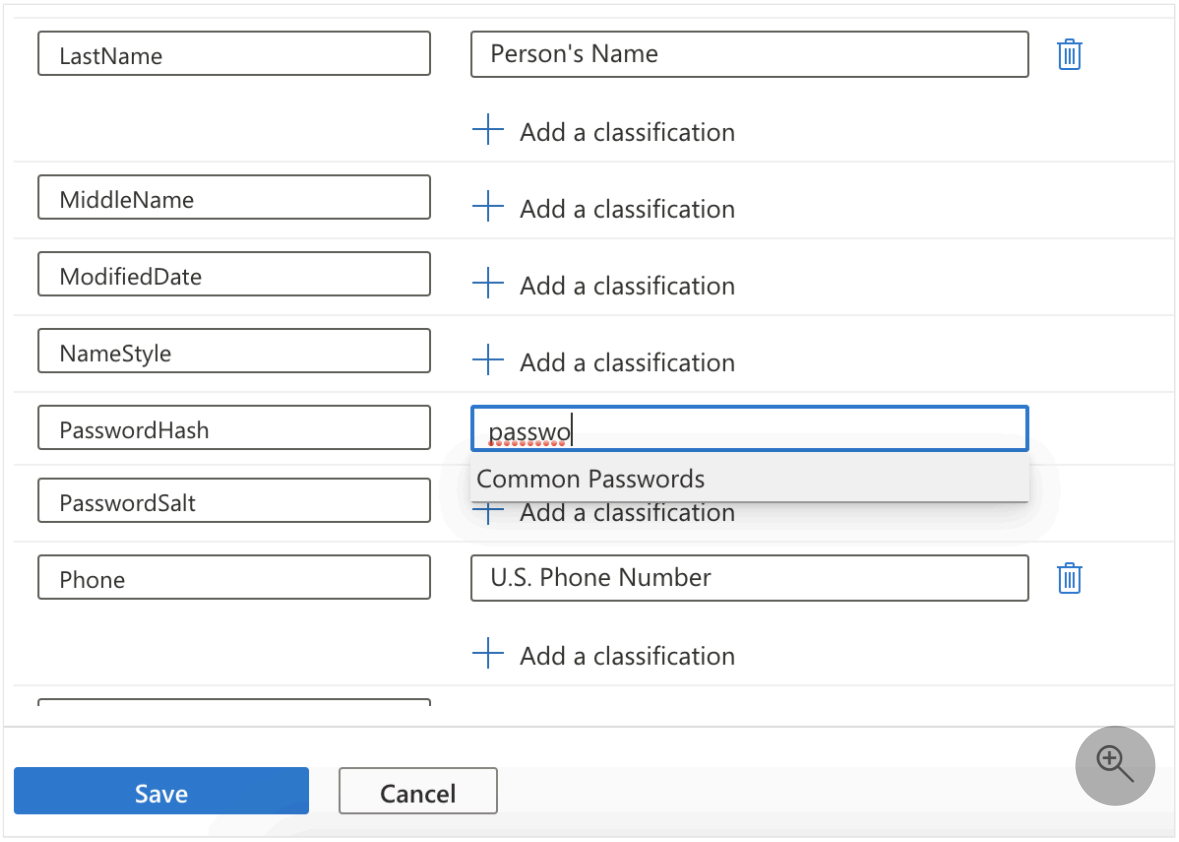
Manually add classification to a column asset

Microsoft Purview automatically scans and adds classifications to all column assets. However, if you want to change the classification, you can do so at the column level:

1. Search or browse Unified Catalog for the table asset that contains the column you want to update.
2. Select **Edit** from the **Overview** tab.
3. Select the **Schema** tab.

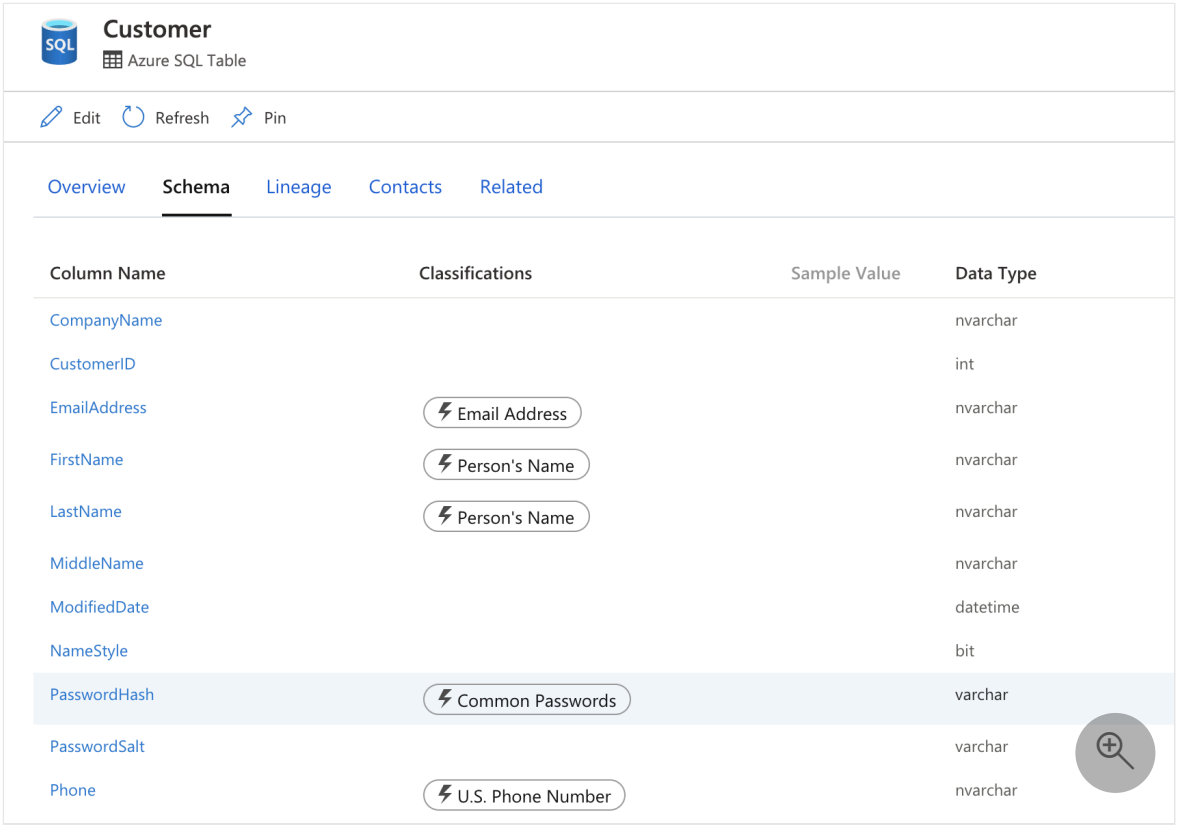


4. Identify the columns you're interested in and select **Add a classification**. This example adds a **Common Passwords** classification to the **PasswordHash** column.



5. Select **Save**.

6. Select the **Schema** tab and confirm that the classification has been added to the column.



View classification details

Microsoft Purview captures important details like who applied a classification and when it was applied. To view the details, hover over the classification to revel the Classification details card. The classification details card shows the following information:

- Classification name - Name of the classification applied on the asset or column.
- Applied by - Who applied the classification. Possible values are scan and user name.
- Applied time - Local timestamp when the classification was applied via scan or manually.
- Classification type - System or custom.

Users with Data Curator role can see more details for classifications that were applied automatically through a scan. These details include sample count that the scanner read to classify the data, and distinct data count in the sample that the scanner found.

Impact of rescanning on existing classifications

Classification tags are automatically applied the first time based on sampling your data and matching those samples against predefined RegEx patterns or dictionaries. Subsequent rescans can preserve, remove, or incrementally add classification tags depending on several factors:

- A scan run doesn't remove classification tags that were manually applied.
- A scan run doesn't reapply classification tags that were manually removed.
- A later run of a scan that continues to use the same scan rule can update the classification tags.
- For a scan run, classification tags from the previous scans can get removed or retained based on the criteria listed in this table:

Expand table

File type	File size	Classification tags from previous scan will be:
File types with no size (for example, SQL)	All	Retained
.doc, .docm, .docx, .dot, .odp, .ods, .odt, .pdf, .pot, .pps, .ppsx, .ppt, .pptm, .pptx, .txt, .xlc, .xls, .xlsb, .xlsm, .xlsx, .xlt	>20 MB	Retained
.doc, .docm, .docx, .dot, .odp, .ods, .odt, .pdf, .pot, .pps, .ppsx, .ppt, .pptm, .pptx, .txt, .xlc, .xls, .xlsb,	<20 MB	Removed

File type	File size	Classification tags from previous scan will be:
.xlsm, .xlsx, .xlt		
GZ	>400 KB	Retained
GZ	<400 KB	Removed
Files with no extension or structured file types	> 1 MB	Retained
Files with no extension or structured file types	< 1 MB	Removed

Next steps

- Learn how to create a custom classification.
- Learn how to automatically apply classifications.

Feedback

Was this page helpful? Yes No

System classifications supported in Data Map

Article • 03/04/2025

This article lists the supported system classifications in Microsoft Purview Data Map. To learn more about classification, see Data classification in Data Map.

Microsoft Purview classifies data by using RegEx, Bloom Filter and Machine Learning models. The following lists describe the format, pattern, and keywords for the Microsoft Purview defined system classifications. Each classification name is prefixed by *MICROSOFT*.

> **Note**
>
> Microsoft Purview can classify both structured (CSV, TSV, JSON, SQL Table etc.) as well as unstructured data (DOC, PDF, TXT etc.). However, there are certain classifications that are only applicable to structured data. Here is the list of classifications that Microsoft Purview doesn't apply on unstructured data - City Name, Country Name, Date Of Birth, Email, Ethnic Group, GeoLocation, Person Name, U.S. Phone Number, U.S. States, U.S. ZipCode

> **Note**
>
> **Minimum match threshold**: It is the minimum percentage of data value matches in a column that must be found by the scanner for the classification to be applied. For system classification minimum match threshold value is set at 60% and cannot be changed. For custom classification, this value is configurable.

Bloom Filter based classifications

World Cities, Country

The City and Country classifier identifies the data based on their full names as well as short codes.

Keywords

Keywords for City

- burg
- city
- cities
- city names
- cosmopolis
- metropolis
- municipality
- place
- town

Keywords for Country

- country
- countries
- country names
- nation
- nationality

Machine Learning based classifications

> ⓘ **Note**
>
> Machine learning based classifiers are only supported for structured data like tabular or columnar data sources.

Person's Name

Person Name machine learning model has been trained using global datasets of names in English language. Microsoft Purview classifies full names stored in the same column as well as first and last names in separate columns.

Person's Address

Person's address classification is used to detect full address stored in a single column containing the following elements: House number, Street Name, City, State,

Country/Region, Zip Code. Person's Address classifier uses machine learning model that is trained on the global addresses data set in English language.

Supported formats

Currently the address model supports the following formats in the same column:

- number, street, city
- name, street, pincode or zipcode
- number, street, area, pincode or zipcode
- street, city, pincode or zipcode
- landmark, city

Person's Gender

Person's Gender machine learning model has been trained using US Census data and other public data sources in English language. It supports classifying 50+ genders out of the box.

Keywords

- sex
- gender
- sexual
- orientation

Person's Age

Person's Age machine learning model detects age of an individual specified in various different formats. The qualifiers for days, months, and years must be in English language.

Keywords

- age
- ages

Supported formats

- {%y} y, {%m} m
- {%y} years {%m} months
- {%y} years and {%m} months
- {%y} years {%w} weeks
- {%y} years and {%w} weeks
- {%y} y, {%d} d
- {%y} y, {%w} w
- {%y} years, {%d} days
- {%y} years and {%d} days
- "{%y} years, {%m} months and {%d} days
- {%y} months and {%d} days
- {%y} yr
- {%y}.{%yd} yr
- {%y} years
- {%y} years old
- {%y}.{%yd} years
- age {%y}
- {%y} to {%y2}
- {%y} to {%y2} yrs
- {%y} years to {%y2} years
- {%m} months to {%y} years
- {%m} m to {%y} years
- {%y}-{%y2} yrs
- {%y}-{%y2}
- {%y} - {%y2}
- {%y}+
- {%m}-{%m2} mos
- {%y} and over
- {%y} and under
- below {%y}
- above {%y}
- month {%m}
- week {%w}
- {%y}

Unsupported formats

- {%y}y {%m}m
- {%y}y {%d}d
- {%y}y {%w}w

- {%y}.{%m}
- {%y}.{%yd}

RegEx Classifications

ABA routing number

Format

Nine digits that can be in a formatted or unformatted pattern.

Pattern

- two digits in the ranges 00-12, 21-32, 61-72, or 80
- two digits
- an optional hyphen
- four digits
- an optional hyphen
- a digit

Checksum

Yes

Keywords

Keyword_aba_routing

- aba number
- aba#
- aba
- abarouting#
- abaroutingnumber
- americanbankassociationrouting#
- americanbankassociationroutingnumber
- bankrouting#
- bankroutingnumber
- routing #

- routing no
- routing number
- routing transit number
- routing#
- RTN

Argentina national identity (DNI) number

Format

Eight digits with or without periods

Pattern

Eight digits:

- two digits
- an optional period
- three digits
- an optional period
- three digits

Checksum

No

Keywords

Keyword_argentina_national_id

- Argentina National Identity number
- cedula
- cédula
- dni
- documento nacional de identidad
- documento número
- documento numero
- registro nacional de las personas
- rnp

Australia bank account number

Format

six to 10 digits with or without a bank state branch number

Pattern

Account number is 6 to 10 digits.

Australia bank state branch number:

- three digits
- a hyphen
- three digits

Checksum

No

Keywords

Keyword_australia_bank_account_number

- swift bank code
- correspondent bank
- base currency
- usa account
- holder address
- bank address
- information account
- fund transfers
- bank charges
- bank details
- banking information
- full names
- iaea

Australia business number

Format

11 digits with optional delimiters

Pattern

11 digits with optional delimiters:

- two digits
- an optional hyphen or space
- three digits
- an optional hyphen or space
- three digits
- an optional hyphen or space
- three digits

Checksum

Yes

Keywords

Keyword_australia_business_number

- australia business no
- business number
- abn#
- businessid#
- business id
- abn
- businessno#

Australia company number

Format

nine digits with delimiters

Pattern

nine digits with delimiters:

- three digits
- a space
- three digits
- a space
- three digits

Checksum

Yes

Keywords

Keyword_australia_company_number

- acn
- australia company no
- australia company no#
- australia company number
- australian company no
- australian company no#
- australian company number

Australia driver's license number

Format

nine letters and digits

Pattern

nine letters and digits:

- two digits or letters (not case-sensitive)
- two digits
- five digits or letters (not case-sensitive)

OR

- one to two optional letters (not case-sensitive)
- four to nine digits

OR

- nine digits or letters (not case-sensitive)

Checksum

No

Keywords

Keyword_australia_drivers_license_number

- international driving permits
- australian automobile association
- international driving permit
- DriverLicence
- DriverLicences
- Driver Lic
- Driver Licence
- Driver Licences
- DriversLic
- DriversLicence
- DriversLicences
- Drivers Lic
- Drivers Lics
- Drivers Licence
- Drivers Licences
- Driver'Lic
- Driver'Lics
- Driver'Licence
- Driver'Licences
- Driver' Lic
- Driver' Lics
- Driver' Licence
- Driver' Licences
- Driver'sLic

- Driver'sLics
- Driver'sLicence
- Driver'sLicences
- Driver's Lic
- Driver's Lics
- Driver's Licence
- Driver's Licences
- DriverLic#
- DriverLics#
- DriverLicence#
- DriverLicences#
- Driver Lic#
- Driver Lics#
- Driver Licence#
- Driver Licences#
- DriversLic#
- DriversLics#
- DriversLicence#
- DriversLicences#
- Drivers Lic#
- Drivers Lics#
- Drivers Licence#
- Drivers Licences#
- Driver'Lic#
- Driver'Lics#
- Driver'Licence#
- Driver'Licences#
- Driver' Lic#
- Driver' Lics#
- Driver' Licence#
- Driver' Licences#
- Driver'sLic#
- Driver'sLics#
- Driver'sLicence#
- Driver'sLicences#
- Driver's Lic#
- Driver's Lics#
- Driver's Licence#
- Driver's Licences#

Keyword_australia_drivers_license_number_exclusions

- aaa
- DriverLicense
- DriverLicenses
- Driver License
- Driver Licenses
- DriversLicense
- DriversLicenses
- Drivers License
- Drivers Licenses
- Driver'License
- Driver'Licenses
- Driver' License
- Driver' Licenses
- Driver'sLicense
- Driver'sLicenses
- Driver's License
- Driver's Licenses
- DriverLicense#
- DriverLicenses#
- Driver License#
- Driver Licenses#
- DriversLicense#
- DriversLicenses#
- Drivers License#
- Drivers Licenses#
- Driver'License#
- Driver'Licenses#
- Driver' License#
- Driver' Licenses#
- Driver'sLicense#
- Driver'sLicenses#
- Driver's License#
- Driver's Licenses#

Australia medical account number

Format

10-11 digits

Pattern

10-11 digits:

- First digit is in the range 2-6
- Ninth digit is a check digit
- Tenth digit is the issue digit
- Eleventh digit (optional) is the individual number

Checksum

Yes

Keywords

Keyword_Australia_Medical_Account_Number

- bank account details
- medicare payments
- mortgage account
- bank payments
- information branch
- credit card loan
- department of human services
- local service
- medicare

Australia passport number

Format

eight or nine alphanumeric characters

Pattern

- one letter (N, E, D, F, A, C, U, X) followed by seven digits

or

- Two letters (PA, PB, PC, PD, PE, PF, PU, PW, PX, PZ) followed by seven digits.

Checksum

No

Keywords

Keyword_australia_passport_number

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers
- passport details
- immigration and citizenship
- commonwealth of australia
- department of immigration
- national identity card
- travel document
- issuing authority

Australia tax file number

Format

eight to nine digits

Pattern

eight to nine digits typically presented with spaces as follows:

- three digits
- an optional space
- three digits
- an optional space
- two to three digits where the last digit is a check digit

Checksum

Yes

Keywords

Keyword_australia_tax_file_number

- australian business number
- marginal tax rate
- medicare levy
- portfolio number
- service veterans
- withholding tax
- individual tax return
- tax file number
- tfn

Austria driver's license number

Format

eight digits without spaces and delimiters

Pattern

eight digits

Checksum

No

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic

- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#

- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses
- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_austria_eu_driver's_license_number

- fuhrerschein
- führerschein
- Führerscheine
- Führerscheinnummer
- Führerscheinnummern

Austria identity card

Format

A 24-character combination of letters, digits, and special characters

Pattern

24 characters:

- 22 letters (not case-sensitive), digits, backslashes, forward slashes, or plus signs
- two letters (not case-sensitive), digits, backslashes, forward slashes, plus signs, or equal signs

Checksum

Not applicable

Keywords

Keywords_austria_eu_national_id_card

- identity number
- national id
- personalausweis republik österreich

Austria passport number

Format

One letter followed by an optional space and seven digits

Pattern

A combination of one letter, seven digits, and one space:

- one letter (not case-sensitive)
- one space (optional)
- seven digits

Checksum

not applicable

Keywords

Keywords_eu_passport_number

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers

Keywords_austria_eu_passport_number

- reisepassnummer
- reisepasse
- No-Reisepass
- Nr-Reisepass
- Reisepass-Nr
- Passnummer
- reisepässe

Keywords_eu_passport_date

- date of issue
- date of expiry

Austria social security number

Format

10 digits in the specified format

Pattern

10 digits:

- three digits that correspond to a serial number
- one check digit
- six digits that correspond to the birth date (DDMMYY)

Checksum

Yes

Keywords

Keywords_austria_eu_ssn_or_equivalent

- austrian ssn
- ehic number
- ehic no
- insurance code
- insurancecode#
- insurance number
- insurance no
- krankenkassennummer
- krankenversicherung
- socialsecurityno
- socialsecurityno#
- social security no
- social security number
- social security code

- sozialversicherungsnummer
- sozialversicherungsnummer#
- soziale sicherheit kein
- sozialesicherheitkein#
- ssn#
- ssn
- versicherungscode
- versicherungsnummer
- zdravstveno zavarovanje

Austria tax identification number

Format

nine digits with optional hyphen and forward slash

Pattern

nine digits with optional hyphen and forward slash:

- two digits
- a hyphen (optional)
- three digits
- a forward slash (optional)
- four digits

Checksum

Yes

Keywords

Keywords_austria_eu_tax_file_number

- österreich
- st.nr.
- steuernummer
- tax id
- tax identification no

- tax identification number
- tax no#
- tax no
- tax number
- tax registration number
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id
- tin no
- tin#
- tax number

Austria value added tax

Format

11-character alphanumeric pattern

Pattern

11-character alphanumeric pattern:

- Or a
- T or t
- Optional space
- U or u
- optional space
- two or three digits
- optional space
- four digits
- optional space
- one or two digits

Checksum

Yes

Keywords

Keyword_austria_value_added_tax

- vat number
- vat#
- austrian vat number
- vat no.
- vatno#
- value added tax number
- austrian vat
- mwst
- umsatzsteuernummer
- mwstnummer
- ust.-identifikationsnummer
- umsatzsteuer-identifikationsnummer
- vat identification number
- atu number
- uid number

Belgium driver's license number

Format

10 digits without spaces and delimiters

Pattern

10 digits

Checksum

No

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense

- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#

- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses
- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_belgium_eu_driver's_license_number

- rijbewijs
- rijbewijsnummer
- führerschein
- führerscheinnummer
- füehrerscheinnummer
- fuhrerschein
- fuehrerschein
- fuhrerscheinnummer
- fuehrerscheinnummer
- permis de conduire
- numéro permis conduire

Belgium national number

Format

11 digits plus optional delimiters

Pattern

11 digits plus delimiters:

- six digits and two optional periods in the format YY.MM.DD for date of birth
- An optional delimiter from dot, dash, space
- three sequential digits (odd for males, even for females)
- An optional delimiter from dot, dash, space
- two check digits

Checksum

Yes

Keywords

Keyword_belgium_national_number

- belasting aantal
- bnn#
- bnn

- carte d’identité
- identifiant national
- identifiantnational#
- identificatie
- identification
- identifikation
- identifikationsnummer
- identifizierung
- identité
- identiteit
- identiteitskaart
- identity
- inscription
- national number
- national register
- nationalnumber#
- nationalnumber
- nif#
- nif
- numéro d'assuré
- numéro de registre national
- numéro de sécurité
- numéro d'identification
- numéro d'immatriculation
- numéro national
- numéronational#
- personal id number
- personalausweis
- personalidnumber#
- registratie
- registration
- registrationsnumme
- registrierung
- social security number
- ssn#
- ssn
- steuernummer
- tax id
- tax identification no
- tax identification number
- tax no#

- tax no
- tax number
- tax registration number
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id
- tin no
- tin#

Belgium passport number

Format

two letters followed by six digits with no spaces or delimiters

Pattern

two letters and followed by six digits

Checksum

not applicable

Keywords

Keywords_eu_passport_number

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number

- passportnumbers
- passport numbers

Keywords_belgium_eu_passport_number

- numéro passeport
- paspoort nr
- paspoort-nr
- paspoortnummer
- paspoortnummers
- Passeport carte
- Passeport livre
- Pass-Nr
- Passnummer
- reisepass kein

Keywords_eu_passport_date

- date of issue
- date of expiry

Belgium value added tax number

Format

12-character alphanumeric pattern

Pattern

12-character alphanumeric pattern:

- a letter B or b
- a letter E or e
- a digit 0
- a digit from 1 to 9
- an optional dot or Hyphen or space
- four digits
- an optional dot or Hyphen or space
- four digits

Checksum

Yes

Keywords

Keyword_belgium_value_added_tax_number

- nº tva
- vat number
- vat no
- numéro t.v.a
- umsatzsteuer-identifikationsnummer
- umsatzsteuernummer
- btw
- btw#
- vat#

Brazil CPF number

Format

11 digits that include a check digit and can be formatted or unformatted

Pattern

Formatted:

- three digits
- a period
- three digits
- a period
- three digits
- a hyphen
- two digits that are check digits

Unformatted:

- 11 digits where the last two digits are check digits

Checksum

Yes

Keywords

Keyword_brazil_cpf

- CPF
- Identification
- Registration
- Revenue
- Cadastro de Pessoas Físicas
- Imposto
- Identificação
- Inscrição
- Receita

Brazil legal entity number (CNPJ)

Format

14 digits that include a registration number, branch number, and check digits, plus delimiters

Pattern

14 digits, plus delimiters:

- two digits
- a period
- three digits
- a period
- three digits (these first eight digits are the registration number)
- a forward slash
- four-digit branch number
- a hyphen
- two digits that are check digits

Checksum

Yes

Keywords

Keyword_brazil_cnpj

- CNPJ
- CNPJ/MF
- CNPJ-MF
- National Registry of Legal Entities
- Taxpayers Registry
- Legal entity
- Legal entities
- Registration Status
- Business
- Company
- CNPJ
- Cadastro Nacional da Pessoa Jurídica
- Cadastro Geral de Contribuintes
- CGC
- Pessoa jurídica
- Pessoas jurídicas
- Situação cadastral
- Inscrição
- Empresa

Brazil national identification card (RG)

Format

Registro Geral (old format): Nine digits

Registro de Identidade (RIC) (new format): 11 digits

Pattern

Registro Geral (old format):

- two digits
- a period
- three digits
- a period
- three digits
- a hyphen
- one digit that is a check digit

Registro de Identidade (RIC) (new format):

- 10 digits
- a hyphen
- one digit that is a check digit

Checksum

Yes

Keywords

Keyword_brazil_rg

- Cédula de identidade
- identity card
- national id
- número de rregistro
- registro de lidentidade
- registro geral
- RG (this keyword is case-sensitive)
- RIC (this keyword is case-sensitive)

Bulgaria driver's license number

Format

nine digits without spaces and delimiters

Pattern

nine digits

Checksum

No

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences

- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#

- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses

- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_bulgaria_eu_driver's_license_number

- свидетелство за управление на мпс
- свидетелство за управление на моторно превозно средство
- сумпс
- шофьорска книжка
- шофьорски книжки

Bulgaria uniform civil number

Format

10 digits without spaces and delimiters

Pattern

10 digits without spaces and delimiters

- six digits that correspond to the birth date (YYMMDD)
- two digits that correspond to the birth order
- one digit that corresponds to gender: An even digit for male and an odd digit for female
- one check digit

Checksum

Yes

Keywords

Keywords_bulgaria_eu_national_id_card

- bnn#
- bnn
- bucn#
- bucn
- edinen grazhdanski nomer
- egn#
- egn
- identification number
- national id
- national number
- nationalnumber#
- nationalnumber
- personal id
- personal no
- personal number
- personalidnumber#
- social security number
- ssn#
- ssn
- uniform civil id
- uniform civil no
- uniform civil number
- uniformcivilno#
- uniformcivilno
- uniformcivilnumber#
- uniformcivilnumber
- unique citizenship number
- егн#
- егн
- единен граждански номер
- идентификационен номер
- личен номер
- лична идентификация
- лично не
- национален номер
- номер на гражданството
- униформ id
- униформ граждански id
- униформ граждански не
- униформ граждански номер
- униформгражданскиid#

- уніформгражданскине.#

Bulgaria passport number

Format

nine digits without spaces and delimiters

Pattern

nine digits

Checksum

No

Keywords

Keywords_eu_passport_number

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers

Keywords_bulgaria_eu_passport_number

- номер на паспорта
- номер на паспорт
- паспорт №

Keywords_eu_passport_date

- date of issue
- date of expiry

Canada bank account number

Format

7 or 12 digits

Pattern

A Canada Bank Account Number is 7 or 12 digits.

A Canada bank account transit number is:

- five digits
- a hyphen
- three digits

or

- a zero "0"
- eight digits

Checksum

No

Keywords

Keyword_canada_bank_account_number

- canada savings bonds
- canada revenue agency
- canadian financial institution
- direct deposit form
- canadian citizen
- legal representative
- notary public
- commissioner for oaths

- child care benefit
- universal child care
- canada child tax benefit
- income tax benefit
- harmonized sales tax
- social insurance number
- income tax refund
- child tax benefit
- territorial payments
- institution number
- deposit request
- banking information
- direct deposit

Canada driver's license number

Format

Varies by province

Pattern

Various patterns covering:

- Alberta
- British Columbia
- Manitoba
- New Brunswick
- Newfoundland/Labrador
- Nova Scotia
- Ontario
- Prince Edward Island
- Quebec
- Saskatchewan

Checksum

No

Keywords

Keyword_[province_name]_drivers_license_name

- The province abbreviation, for example AB
- The province name, for example Alberta

Keyword_canada_drivers_license

- DL
- DLS
- CDL
- CDLS
- DriverLic
- DriverLics
- DriverLicense
- DriverLicenses
- DriverLicence
- DriverLicences
- Driver Lic
- Driver Lics
- Driver License
- Driver Licenses
- Driver Licence
- Driver Licences
- DriversLic
- DriversLics
- DriversLicence
- DriversLicences
- DriversLicense
- DriversLicenses
- Drivers Lic
- Drivers Lics
- Drivers License
- Drivers Licenses
- Drivers Licence
- Drivers Licences
- Driver'Lic
- Driver'Lics
- Driver'License
- Driver'Licenses

- Driver'Licence
- Driver'Licences
- Driver' Lic
- Driver' Lics
- Driver' License
- Driver' Licenses
- Driver' Licence
- Driver' Licences
- Driver'sLic
- Driver'sLics
- Driver'sLicense
- Driver'sLicenses
- Driver'sLicence
- Driver'sLicences
- Driver's Lic
- Driver's Lics
- Driver's License
- Driver's Licenses
- Driver's Licence
- Driver's Licences
- Permis de Conduire
- id
- ids
- idcard number
- idcard numbers
- idcard #
- idcard #s
- idcard card
- idcard cards
- idcard
- identification number
- identification numbers
- identification #
- identification #s
- identification card
- identification cards
- identification
- DL#
- DLS#
- CDL#
- CDLS#

- DriverLic#
- DriverLics#
- DriverLicense#
- DriverLicenses#
- DriverLicence#
- DriverLicences#
- Driver Lic#
- Driver Lics#
- Driver License#
- Driver Licenses#
- Driver License#
- Driver Licences#
- DriversLic#
- DriversLics#
- DriversLicense#
- DriversLicenses#
- DriversLicence#
- DriversLicences#
- Drivers Lic#
- Drivers Lics#
- Drivers License#
- Drivers Licenses#
- Drivers Licence#
- Drivers Licences#
- Driver'Lic#
- Driver'Lics#
- Driver'License#
- Driver'Licenses#
- Driver'Licence#
- Driver'Licences#
- Driver' Lic#
- Driver' Lics#
- Driver' License#
- Driver' Licenses#
- Driver' Licence#
- Driver' Licences#
- Driver'sLic#
- Driver'sLics#
- Driver'sLicense#
- Driver'sLicenses#
- Driver'sLicence#

- Driver'sLicences#
- Driver's Lic#
- Driver's Lics#
- Driver's License#
- Driver's Licenses#
- Driver's Licence#
- Driver's Licences#
- Permis de Conduire#
- id#
- ids#
- idcard card#
- idcard cards#
- idcard#
- identification card#
- identification cards#
- identification#

Canada health service number

Format

10 digits

Pattern

10 digits

Checksum

No

Keywords

Keyword_canada_health_service_number

- personal health number
- patient information
- health services
- speciality services

- automobile accident
- patient hospital
- psychiatrist
- workers compensation
- disability

Canada passport number

Format

two uppercase letters followed by six digits

Pattern

two uppercase letters followed by six digits

Checksum

No

Keywords

Keyword_canada_passport_number

- canadian citizenship
- canadian passport
- passport application
- passport photos
- certified translator
- canadian citizens
- processing times
- renewal application

Keyword_passport

- Passport Number
- Passport No
- Passport #
- Passport#

- PassportID
- Passportno
- passportnumber
- パスポート
- パスポート番号
- パスポートのNum
- パスポート＃
- Numéro de passeport
- Passeport n °
- Passeport Non
- Passeport #
- Passeport#
- PasseportNon
- Passeportn °

Canada personal health identification number (PHIN)

Format

nine digits

Pattern

nine digits

Checksum

No

Keywords

Keyword_canada_phin

- social insurance number
- health information act
- income tax information
- manitoba health
- health registration
- prescription purchases

- benefit eligibility
- personal health
- power of attorney
- registration number
- personal health number
- practitioner referral
- wellness professional
- patient referral
- health and wellness

Keyword_canada_provinces

- Nunavut
- Quebec
- Northwest Territories
- Ontario
- British Columbia
- Alberta
- Saskatchewan
- Manitoba
- Yukon
- Newfoundland and Labrador
- New Brunswick
- Nova Scotia
- Prince Edward Island
- Canada

Canada social insurance number

Format

nine digits with optional hyphens or spaces

Pattern

Formatted:

- three digits
- a hyphen or space

- three digits
- a hyphen or space
- three digits

Unformatted: nine digits

Checksum

Yes

Keywords

Keyword_sin

- sin
- social insurance
- numero d'assurance sociale
- sins
- ssn
- ssns
- social security
- numero d'assurance social
- national identification number
- national id
- sin#
- soc ins
- social ins

Keyword_sin_collaborative

- driver's license
- drivers license
- driver's licence
- drivers licence
- DOB
- Birthdate
- Birthday
- Date of Birth

Chile identity card number

Format

seven to eight digits plus delimiters a check digit or letter

Pattern

seven to eight digits plus delimiters:

- one to two digits
- an optional period
- three digits
- an optional period
- three digits
- a dash
- one digit or letter (not case-sensitive) which is a check digit

Checksum

Yes

Keywords

Keyword_chile_id_card

- cédula de identidad
- identificación
- national identification
- national identification number
- national id
- número de identificación nacional
- rol único nacional
- rol único tributario
- RUN
- RUT
- tarjeta de identificación
- Rol Unico Nacional
- Rol Unico Tributario
- RUN#

- RUT#
- nationaluniqueroleID#
- nacional identidad
- número identificación
- identidad número
- numero identificacion
- identidad numero
- Chilean identity no.
- Chilean identity number
- Chilean identity #
- Unique Tax Registry
- Unique Tributary Role
- Unique Tax Role
- Unique Tributary Number
- Unique National Number
- Unique National Role
- National unique role
- Chile identity no.
- Chile identity number
- Chile identity #

China resident identity card (PRC) number

Format

18 digits

Pattern

18 digits:

- six digits that are an address code
- eight digits in the form YYYYMMDD, which are the date of birth
- three digits that are an order code
- one digit that is a check digit

Checksum

Yes

Keywords

Keyword_china_resident_id

- Resident Identity Card
- PRC
- National Identification Card
- 身份证
- 居民 身份证
- 居民身份证
- 鉴定
- 身分證
- 居民 身份證
- 鑑定

Credit card number

Format

14 to 16 digits that can be formatted or unformatted (ddddddddddddddd) and that must pass the Luhn test.

Pattern

Detects cards from all major brands worldwide, including Visa, MasterCard, Discover Card, JCB, American Express, gift cards, diner's cards, Rupay and China UnionPay.

Checksum

Yes, the Luhn checksum

Keywords

Keyword_cc_verification

- card verification
- card identification number
- cvn

- cid
- cvc2
- cvv2
- pin block
- security code
- security number
- security no
- issue number
- issue no
- cryptogramme
- numéro de sécurité
- numero de securite
- kreditkartenprüfnummer
- kreditkartenprufnummer
- prüfziffer
- prufziffer
- sicherheits Kode
- sicherheitscode
- sicherheitsnummer
- verfalldatum
- codice di verifica
- cod. sicurezza
- cod sicurezza
- n autorizzazione
- código
- codigo
- cod. seg
- cod seg
- código de segurança
- codigo de seguranca
- codigo de segurança
- código de seguranca
- cód. segurança
- cod. seguranca
- cod. segurança
- cód. seguranca
- cód segurança
- cod seguranca
- cod segurança
- cód seguranca
- número de verificação

- numero de verificacao
- ablauf
- gültig bis
- gültigkeitsdatum
- gultig bis
- gultigkeitsdatum
- scadenza
- data scad
- fecha de expiracion
- fecha de venc
- vencimiento
- válido hasta
- valido hasta
- vto
- data de expiração
- data de expiracao
- data em que expira
- validade
- valor
- vencimento
- transaction
- transaction number
- reference number
- セキュリティコード
- セキュリティ コード
- セキュリティナンバー
- セキュリティ ナンバー
- セキュリティ番号

Keyword_cc_name

- amex
- american express
- americanexpress
- americano espresso
- Visa
- mastercard
- master card
- mc
- mastercards
- master cards

- diner's Club
- diners club
- dinersclub
- discover
- discover card
- discovercard
- discover cards
- JCB
- BrandSmart
- japanese card bureau
- carte blanche
- carteblanche
- credit card
- cc#
- cc#:
- expiration date
- exp date
- expiry date
- date d’expiration
- date d'exp
- date expiration
- bank card
- bankcard
- card number
- card num
- cardnumber
- cardnumbers
- card numbers
- creditcard
- credit cards
- creditcards
- ccn
- card holder
- cardholder
- card holders
- cardholders
- check card
- checkcard
- check cards
- checkcards
- debit card

- debitcard
- debit cards
- debitcards
- atm card
- atmcard
- atm cards
- atmcards
- enroute
- en route
- card type
- Cardmember Acct
- cardmember account
- Cardno
- Corporate Card
- Corporate cards
- Type of card
- card account number
- card member account
- Cardmember Acct.
- card no.
- card no
- card number
- carte bancaire
- carte de crédit
- carte de credit
- numéro de carte
- numero de carte
- nº de la carte
- nº de carte
- kreditkarte
- karte
- karteninhaber
- karteninhabers
- kreditkarteninhaber
- kreditkarteninstitut
- kreditkartentyp
- eigentümername
- kartennr
- kartennummer
- kreditkartennummer
- kreditkarten-nummer

- carta di credito
- carta credito
- n. carta
- n carta
- nr. carta
- nr carta
- numero carta
- numero della carta
- numero di carta
- tarjeta credito
- tarjeta de credito
- tarjeta crédito
- tarjeta de crédito
- tarjeta de atm
- tarjeta atm
- tarjeta debito
- tarjeta de debito
- tarjeta débito
- tarjeta de débito
- nº de tarjeta
- no. de tarjeta
- no de tarjeta
- numero de tarjeta
- número de tarjeta
- tarjeta no
- tarjetahabiente
- cartão de crédito
- cartão de credito
- cartao de crédito
- cartao de credito
- cartão de débito
- cartao de débito
- cartão de debito
- cartao de debito
- débito automático
- debito automatico
- número do cartão
- numero do cartão
- número do cartao
- numero do cartao
- número de cartão

- numero de cartão
- número de cartao
- numero de cartao
- nº do cartão
- nº do cartao
- nº. do cartão
- no do cartão
- no do cartao
- no. do cartão
- no. do cartao
- rupay
- union pay
- unionpay
- diner's
- diners
- クレジットカード番号
- クレジットカードナンバー
- クレジットカード＃
- クレジットカード
- クレジット
- クレカ
- カード番号
- カードナンバー
- カード＃
- アメックス
- アメリカンエクスプレス
- アメリカン エクスプレス
- Visaカード
- Visa カード
- マスターカード
- マスター カード
- マスター
- ダイナースクラブ
- ダイナース クラブ
- ダイナース
- 有効期限
- 期限
- キャッシュカード
- キャッシュ カード
- カード名義人
- カードの名義人

- カードの名義
- デビット カード
- デビットカード
- 中国银联
- 银联

Croatia driver's license number

Format

eight digits without spaces and delimiters

Pattern

eight digits

Checksum

No

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence

- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#

- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#

- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses
- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_croatia_eu_driver's_license_number

- vozačka dozvola
- vozačke dozvole

Croatia identity card number

This entity is included in the EU National Identification Number sensitive information type. It's available as a stand-alone sensitive information type entity.

Format

nine digits

Pattern

nine consecutive digits

Checksum

No

Keywords

Keyword_croatia_id_card

- majstorski broj građana
- master citizen number
- nacionalni identifikacijski broj
- national identification number
- oib#
- oib
- osobna iskaznica
- osobni id
- osobni identifikacijski broj
- personal identification number
- porezni broj
- porezni identifikacijski broj
- tax id
- tax identification no
- tax identification number
- tax no#
- tax no
- tax number
- tax registration number
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id
- tin no
- tin#

Croatia passport number

Format

nine digits without spaces and delimiters

Pattern

nine digits

Checksum

No

Keywords

Keywords_eu_passport_number_common

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers

Keywords_croatia_eu_passport_number

- broj putovnice
- br. Putovnice
- br putovnice

Croatia personal identification (OIB) number

Format

11 digits

Pattern

11 digits:

- 10 digits
- final digit is a check digit

Checksum

Yes

Keywords

Keyword_croatia_oib_number

- majstorski broj građana
- master citizen number
- nacionalni identifikacijski broj
- national identification number
- oib#
- oib
- osobna iskaznica
- osobni id
- osobni identifikacijski broj
- personal identification number
- porezni broj
- porezni identifikacijski broj
- tax id
- tax identification no
- tax identification number
- tax no#
- tax no
- tax number
- tax registration number
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id

- tin no
- tin#

Cyprus drivers license number

Format

12 digits without spaces and delimiters

Pattern

12 digits

Checksum

No

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses

- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#

- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#

- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses
- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_cyprus_eu_driver's_license_number

- άδεια οδήγησης
- αριθμό άδειας οδήγησης
- άδειες οδήγησης

Cyprus identity card

Format

10 digits without spaces and delimiters

Pattern

10 digits

Checksum

not applicable

Keywords

Keywords_cyprus_eu_national_id_card

- id card number
- identity card number
- kimlik karti
- national identification number
- personal id number
- ταυτοτηταση

Cyprus passport number

Format

one letter followed by 6-8 digits with no spaces or delimiters

Pattern

one letter followed by six to eight digits

Checksum

No

Keywords

Keywords_eu_passport_number_common

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers

Keywords_cyprus_eu_passport_number

- αριθμό διαβατηρίου
- pasaportu
- Αριθμός Διαβατηρίου
- κυπριακό διαβατήριο
- διαβατήριο#
- διαβατήριο
- αριθμός διαβατηρίου
- Pasaport Kimliği
- pasaport numarası
- Pasaport no.
- Αρ. Διαβατηρίου

Keywords_cyprus_eu_passport_date

- expires on
- issued on

Cyprus tax identification number

Format

eight digits and one letter in the specified pattern

Pattern

eight digits and one letter:

- a "0" or "9"
- seven digits
- one letter (not case-sensitive)

Checksum

not applicable

Keywords

Keywords_cyprus_eu_tax_file_number

- tax id
- tax identification code
- tax identification no
- tax identification number
- tax no#
- tax no
- tax number
- tax registration number
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tic#
- tic
- tin id
- tin no
- tin#
- vergi kimlik kodu
- vergi kimlik numarası
- αριθμός φορολογικού μητρώου
- κωδικός φορολογικού μητρώου
- φορολογική ταυτότητα
- φορολογικού κωδικού

Czech Republic Driver's License Number

Format

two letters followed by six digits

Pattern

eight letters and digits:

- letter 'E' (not case-sensitive)
- a letter

- a space (optional)
- six digits

Checksum

No

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence

- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#

- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen

- driving licenses
- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_czech_republic_eu_driver's_license_number

- řidičský prúkaz
- řidičské průkazy
- číslo řidičského průkazu
- čísla řidičských průkazů

Czech passport number

Format

eight digits without spaces or delimiters

Pattern

eight digits without spaces or delimiters

Checksum

No

Keywords

Keywords_eu_passport_number_common

- passport#
- passport #
- passportid
- passports
- passportno
- passport no

- passportnumber
- passport number
- passportnumbers
- passport numbers

Keywords_czech_republic_eu_passport_number

- cestovní pas
- číslo pasu
- cestovní pasu
- passeport no
- čísla pasu

Keywords_eu_passport_date

- date of issue
- date of expiry

Czech National Identity Card Number

Format

nine digits with optional forward slash (old format) 10 digits with optional forward slash (new format)

Pattern

nine digits (old format):

- six digits that represent date of birth
- an optional forward slash
- three digits

10 digits (new format):

- six digits that represent date of birth
- an optional forward slash
- four digits where last digit is a check digit

Checksum

Yes

Keywords

Keyword_czech_id_card

- birth number
- czech republic id
- czechidno#
- daňové číslo
- identifikační číslo
- identity no
- identity number
- identityno#
- identityno
- insurance number
- national identification number
- nationalnumber#
- national number
- osobní číslo
- personalidnumber#
- personal id number
- personal identification number
- personal number
- pid#
- pid
- pojištění číslo
- rč
- rodne cislo
- rodné číslo
- ssn
- ssn#
- social security number
- tax id
- tax identification no
- tax identification number
- tax no#
- tax no
- tax number
- tax registration number

- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id
- tin no
- tin#
- unique identification number

Date Of Birth

Format

Any valid date

Checksum

Not applicable

Keywords

Keywords_date_of_birth

- dob
- birth day
- natal day
- any word containing the string *birth*

Denmark driver's license number

Format

eight digits without spaces and delimiters

Pattern

eight digits

Checksum

No

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences

- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#

- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses

- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_denmark_eu_driver's_license_number

- kørekort
- kørekortnummer

Denmark passport number

Format

nine digits without spaces and delimiters

Pattern

nine digits

Checksum

No

Keywords

Keywords_eu_passport_number_common

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers

- passport numbers

Keywords_denmark_eu_passport_number

- pasnummer
- Passeport n°
- pasnumre

Keywords_eu_passport_date

- date of issue
- date of expiry

Denmark personal identification number

Format

10 digits containing a hyphen

Pattern

10 digits:

- six digits in the format DDMMYY, which are the date of birth
- a hyphen
- four digits where the final digit is a check digit

Checksum

Yes

Keywords

Keyword_denmark_id

- centrale personregister
- civilt registreringssystem
- cpr
- cpr#

- gesundheitskarte nummer
- gesundheitsversicherungkarte nummer
- health card
- health insurance card number
- health insurance number
- identification number
- identifikationsnummer
- identifikationsnummer#
- identity number
- krankenkassennummer
- nationalid#
- nationalnumber#
- national number
- personalidnumber#
- personalidentityno#
- personal id number
- personnummer
- personnummer#
- reisekrankenversicherungskartenummer
- rejsesygesikringskort
- ssn
- ssn#
- skat id
- skat kode
- skat nummer
- skattenummer
- social security number
- sundhedsforsikringskort
- sundhedsforsikringsnummer
- sundhedskort
- sundhedskortnummer
- sygesikring
- sygesikringkortnummer
- tax code
- travel health insurance card
- uniqueidentityno#
- tax number
- tax registration number
- tax id
- tax identification number
- taxid#

- taxnumber#
- tax no
- taxno#
- taxnumber
- tax identification no
- tin#
- taxidno#
- taxidnumber#
- tax no#
- tin id
- tin no
- cpr.nr
- cprnr
- cprnummer
- personnr
- personregister
- sygesikringsbevis
- sygesikringsbevisnr
- sygesikringsbevisnummer
- sygesikringskort
- sygesikringskortnr
- sygesikringskortnummer
- sygesikringsnr
- sygesikringsnummer

Email

Format

Any valid email address that abides by RFC 5322

Checksum

Not applicable

Keywords

Keywords_email

- contact
- email
- electronic
- login
- mail
- online
- user
- webmail

Estonia driver's license number

Format

two letters followed by six digits

Pattern

two letters and six digits:

- the letters "ET" (not case-sensitive)
- six digits

Checksum

No

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license

- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#

- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#

- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses
- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_estonia_eu_driver's_license_number

- permis de conduire
- juhilubade numbrid
- juhiloa number
- juhiluba

Estonia Personal Identification Code

Format

11 digits without spaces and delimiters

Pattern

11 digits:

- one digit that corresponds to sex and century of birth (odd number male, even number female; 1-2: 19th century; 3-4: 20th century; 5-6: 21st century)
- six digits that correspond to date of birth (YYMMDD)
- three digits that correspond to a serial number separating persons born on the same date
- one check digit

Checksum

Yes

Keywords

Keywords_estonia_eu_national_id_card

- id-kaart
- ik
- isikukood#
- isikukood
- maksu id
- maksukohustuslase identifitseerimisnumber
- maksunumber
- national identification number
- national number
- personal code
- personal id number
- personal identification code
- personal identification number
- personalidnumber#
- tax id
- tax identification no
- tax identification number
- tax no#
- tax no

- tax number
- tax registration number
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id
- tin no
- tin#

Estonia passport number

Format

one letter followed by seven digits with no spaces or delimiters

Pattern

one letter followed by seven digits

Checksum

No

Keywords

Keywords_eu_passport_number_common

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers

- passport numbers

Keywords_estonia_eu_passport_number

- eesti kodaniku pass
- passi number
- passinumbrid
- document number
- document no
- dokumendi nr

Keywords_eu_passport_date

- date of issue
- date of expiry

Ethnic groups

Format

This classifier consists of the most common ethnic groups. For a reference list, see this article.

Checksum

Not applicable

Keywords

Keywords_ethnic_group

- ethnic
- ethnic groups
- ethnicity
- ethnicities
- nationality
- race

EU debit card number

Format

16 digits

Pattern

Complex and robust pattern

Checksum

Yes

Keywords

Keyword_eu_debit_card

- account number
- card number
- card no.
- security number
- cc#

Keyword_card_terms_dict

- acct nbr
- acct num
- acct no
- american express
- americanexpress
- americano espresso
- amex
- atm card
- atm cards
- atm kaart
- atmcard
- atmcards
- atmkaart
- atmkaarten

- bancontact
- bank card
- bankkaart
- card holder
- card holders
- card num
- card number
- card numbers
- card type
- cardano numerico
- cardholder
- cardholders
- cardnumber
- cardnumbers
- carta bianca
- carta credito
- carta di credito
- cartao de credito
- cartao de crédito
- cartao de debito
- cartao de débito
- carte bancaire
- carte blanche
- carte bleue
- carte de credit
- carte de crédit
- carte di credito
- carteblanche
- cartão de credito
- cartão de crédito
- cartão de debito
- cartão de débito
- cb
- ccn
- check card
- check cards
- checkcard
- checkcards
- chequekaart
- cirrus
- cirrus-edc-maestro

- controlekaart
- controlekaarten
- credit card
- credit cards
- creditcard
- creditcards
- debetkaart
- debetkaarten
- debit card
- debit cards
- debitcard
- debitcards
- debito automatico
- diners club
- dinersclub
- discover
- discover card
- discover cards
- discovercard
- discovercards
- débito automático
- edc
- eigentümername
- european debit card
- hoofdkaart
- hoofdkaarten
- in viaggio
- japanese card bureau
- japanse kaartdienst
- jcb
- kaart
- kaart num
- kaartaantal
- kaartaantallen
- kaarthouder
- kaarthouders
- karte
- karteninhaber
- karteninhabers
- kartennr
- kartennummer

- kreditkarte
- kreditkarten-nummer
- kreditkarteninhaber
- kreditkarteninstitut
- kreditkartennummer
- kreditkartentyp
- maestro
- master card
- master cards
- mastercard
- mastercards
- mc
- mister cash
- n carta
- carta
- no de tarjeta
- no do cartao
- no do cartão
- no. de tarjeta
- no. do cartao
- no. do cartão
- nr carta
- nr. carta
- numeri di scheda
- numero carta
- numero de cartao
- numero de carte
- numero de cartão
- numero de tarjeta
- numero della carta
- numero di carta
- numero di scheda
- numero do cartao
- numero do cartão
- numéro de carte
- nº carta
- nº de carte
- nº de la carte
- nº de tarjeta
- nº do cartao
- nº do cartão

- nº. do cartão
- número de cartao
- número de cartão
- número de tarjeta
- número do cartao
- scheda dell'assegno
- scheda dell'atmosfera
- scheda dell'atmosfera
- scheda della banca
- scheda di controllo
- scheda di debito
- scheda matrice
- schede dell'atmosfera
- schede di controllo
- schede di debito
- schede matrici
- scoprono la scheda
- scoprono le schede
- solo
- supporti di scheda
- supporto di scheda
- switch
- tarjeta atm
- tarjeta credito
- tarjeta de atm
- tarjeta de credito
- tarjeta de debito
- tarjeta debito
- tarjeta no
- tarjetahabiente
- tipo della scheda
- ufficio giapponese della
- scheda
- v pay
- v-pay
- visa
- visa plus
- visa electron
- visto
- visum
- vpay

Keyword_card_security_terms_dict

- card identification number
- card verification
- cardi la verifica
- cid
- cod seg
- cod seguranca
- cod segurança
- cod sicurezza
- cod. seg
- cod. seguranca
- cod. segurança
- cod. sicurezza
- codice di sicurezza
- codice di verifica
- codigo
- codigo de seguranca
- codigo de segurança
- crittogramma
- cryptogram
- cryptogramme
- cv2
- cvc
- cvc2
- cvn
- cvv
- cvv2
- cód seguranca
- cód segurança
- cód. seguranca
- cód. segurança
- código
- código de seguranca
- código de segurança
- de kaart controle
- geeft nr uit
- issue no
- issue number
- kaartidentificatienummer
- kreditkartenprufnummer

- kreditkartenprüfnummer
- kwestieaantal
- no. dell'edizione
- no. di sicurezza
- numero de securite
- numero de verificacao
- numero dell'edizione
- numero di identificazione della
- scheda
- numero di sicurezza
- numero van veiligheid
- numéro de sécurité
- nº autorizzazione
- número de verificação
- perno il blocco
- pin block
- prufziffer
- prüfziffer
- security code
- security no
- security number
- sicherheits kode
- sicherheitscode
- sicherheitsnummer
- speldblok
- veiligheid nr
- veiligheidsaantal
- veiligheidscode
- veiligheidsnummer
- verfalldatum

Keyword_card_expiration_terms_dict

- ablauf
- data de expiracao
- data de expiração
- data del exp
- data di exp
- data di scadenza
- data em que expira
- data scad

- data scadenza
- date de validité
- datum afloop
- datum van exp
- de afloop
- espira
- espira
- exp date
- exp datum
- expiration
- expire
- expires
- expiry
- fecha de expiracion
- fecha de venc
- gultig bis
- gultigkeitsdatum
- gültig bis
- gültigkeitsdatum
- la scadenza
- scadenza
- valable
- validade
- valido hasta
- valor
- venc
- vencimento
- vencimiento
- verloopt
- vervaldag
- vervaldatum
- vto
- válido hasta

EU driver's license number

These entities are in the EU Driver's License Number and are sensitive information types.

- Austria
- Belgium

- Bulgaria
- Croatia
- Cyprus
- Czech
- Denmark
- Estonia
- Finland
- France
- Germany
- Greece
- Hungary
- Ireland
- Italy
- Latvia
- Lithuania
- Luxemburg
- Malta
- Netherlands
- Poland
- Portugal
- Romania
- Slovakia
- Slovenia
- Spain
- Sweden
- U.K.

EU passport number

These entities are in the EU passport number and are sensitive information types. These entities are in the EU passport number bundle.

- Austria
- Belgium
- Bulgaria
- Croatia
- Cyprus
- Czech
- Denmark
- Estonia

- Finland
- France
- Germany
- Greece
- Hungary
- Ireland
- Italy
- Latvia
- Lithuania
- Luxemburg
- Malta
- Netherlands
- Poland
- Portugal
- Romania
- Slovakia
- Slovenia
- Spain
- Sweden
- U.K.

Finland driver's license number

Format

10 digits containing a hyphen

Pattern

10 digits containing a hyphen:

- six digits
- a hyphen
- three digits
- a digit or letter

Checksum

No

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic

- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#

- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses
- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_finland_eu_driver's_license_number

- ajokortti
- permis de conduire
- ajokortin numero
- kuljettaja lic.
- körkort
- körkortnummer
- förare lic.
- ajokortit
- ajokortin numerot

Finland european health insurance number

Format

20-digit number

Pattern

20-digit number:

- 10 digits - 8024680246
- an optional space or hyphen
- 10 digits

Checksum

No

Keywords

Keyword_finland_european_health_insurance_number

- ehic#
- ehic
- finlandehicnumber#
- finska sjukförsäkringskort
- health card

- health insurance card
- health insurance number
- hälsokort
- sairaanhoitokortin
- sairausvakuutuskortti
- sairausvakuutusnumero
- sjukförsäkring nummer
- sjukförsäkringskort
- suomen sairausvakuutuskortti
- terveyskortti

Finland national ID

Format

six digits plus a character indicating a century plus three digits plus a check digit

Pattern

Pattern must include all of the following:

- six digits in the format DDMMYY, which are a date of birth
- century marker (either '-', '+' or 'a')
- three-digit personal identification number
- a digit or letter (case insensitive) which is a check digit

Checksum

Yes

Keywords

- ainutlaatuinen henkilökohtainen tunnus
- henkilökohtainen tunnus
- henkilötunnus
- henkilötunnusnumero#
- henkilötunnusnumero
- hetu
- id no

- id number
- identification number
- identiteetti numero
- identity number
- idnumber
- kansallinen henkilötunnus
- kansallisen henkilökortin
- national id card
- national id no.
- personal id
- personal identity code
- personalidnumber#
- personbeteckning
- personnummer
- social security number
- sosiaaliturvatunnus
- tax id
- tax identification no
- tax identification number
- tax no#
- tax no
- tax number
- tax registration number
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id
- tin no
- tin#
- tunnistenumero
- tunnus numero
- tunnusluku
- tunnusnumero
- verokortti
- veronumero
- verotunniste
- verotunnus

Finland passport number

This entity is available in the EU Passport Number sensitive information type and is available as a stand-alone sensitive information type entity.

Format

combination of nine letters and digits

Pattern

combination of nine letters and digits:

- two letters (not case-sensitive)
- seven digits

Checksum

No

Keywords

Keywords_eu_passport_number

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers

Keyword_finland_passport_number

- suomalainen passi
- passin numero
- passin numero.#

- passin numero#
- passin numero.
- passi#
- passi number

Keywords_eu_passport_date

- date of issue
- date of expiry

France driver's license number

This entity is available in the EU Driver's License Number sensitive information type and is available as a stand-alone sensitive information type entity.

Format

12 digits

Pattern

12 digits with validation to discount similar patterns such as French telephone numbers

Checksum

No

Keywords

Keyword_french_drivers_license

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics

- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#

- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#

- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses
- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number
- permis de conduire
- licence number
- license number
- licence numbers
- license numbers
- numéros de licence

France health insurance number

Format

21-digit number

Pattern

21-digit number:

- 10 digits
- an optional space
- 10 digits
- an optional space
- a digit

Checksum

No

Keywords

Keyword_France_health_insurance_number

- insurance card
- carte vitale
- carte d'assuré social

France national id card (CNI)

Format

12 digits

Pattern

12 digits

Checksum

No

Keywords

Keywords_france_eu_national_id_card

- card number
- carte nationale d’identité
- carte nationale d'idenite no
- cni#
- cni
- compte bancaire
- national identification number
- national identity
- nationalidno#
- numéro d'assurance maladie
- numéro de carte vitale

France passport number

This entity is available in the EU Passport Number sensitive information type. It's also available as a stand-alone sensitive information type entity.

Format

nine digits and letters

Pattern

nine digits and letters:

- two digits
- two letters (not case-sensitive)
- five digits

Checksum

No

Keywords

Keywords_eu_passport_number

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers

Keywords_france_eu_passport_number

- numéro de passeport
- passeport n °
- passeport non
- passeport #
- passeport#
- passeportnon
- passeportn °
- passeport français
- passeport livre
- passeport carte
- numéro passeport
- passeport n°
- n° du passeport
- n° passeport

Keywords_eu_passport_date

- date of issue
- date of expiry

France social security number (INSEE)

Format

15 digits

Pattern

Must match one of two patterns:

- 13 digits followed by a space followed by two digits
  or
- 15 consecutive digits

Checksum

Yes

Keywords

Keyword_fr_insee

- code sécu
- d'identité nationale
- insee
- fssn#
- le numéro d'identification nationale
- le code de la sécurité sociale
- national id
- national identification
- no d'identité
- no. d'identité
- numéro d'assurance
- numéro d'identité
- numero d'identite
- numéro de sécu
- numéro de sécurité sociale
- no d'identite
- no. d'identite
- ssn
- ssn#
- sécurité sociale
- securité sociale
- securite sociale
- socialsecuritynumber
- social security number
- social security code

- social insurance number

France tax identification number

Format

13 digits

Pattern

13 digits

- One digit that must be 0, 1, 2, or 3
- One digit
- A space (optional)
- Two digits
- A space (optional)
- Three digits
- A space (optional)
- Three digits
- A space (optional)
- Three check digits

Checksum

Yes

Keywords

Keywords_france_eu_tax_file_number

- numéro d'identification fiscale
- tax id
- tax identification no
- tax identification number
- tax no#
- tax no
- tax number
- tax registration number

- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id
- tin no
- tin#

France value added tax number

Format

13 character alphanumeric pattern

Pattern

13 character alphanumeric pattern:

- two letters - FR (case insensitive)
- an optional space or hyphen
- two letters or digits
- an optional space, dot, hyphen, or comma
- three digits
- an optional space, dot, hyphen, or comma
- three digits
- an optional space, dot, hyphen, or comma
- three digits

Checksum

Yes

Keywords

Keyword_France_value_added_tax_number

- vat number

- vat no
- vat#
- value added tax
- siren identification no numéro d'identification taxe sur valeur ajoutée
- taxe valeur ajoutée
- taxe sur la valeur ajoutée
- n° tva
- numéro de tva
- numéro d'identification siren

Germany driver's license number

This sensitive information type entity is included in the EU Driver's License Number sensitive information type. It's also available as a stand-alone sensitive information type entity.

Format

combination of 11 digits and letters

Pattern

11 digits and letters (not case-sensitive):

- a digit or letter
- two digits
- six digits or letters
- a digit
- a digit or letter

Checksum

Yes

Keywords

Keyword_german_drivers_license_number

- ausstellungsdatum

- ausstellungsort
- ausstellende behöde
- ausstellende behorde
- ausstellende behoerde
- führerschein
- fuhrerschein
- fuehrerschein
- führerscheinnummer
- fuhrerscheinnummer
- fuehrerscheinnummer
- führerschein-
- fuhrerschein-
- fuehrerschein-
- führerscheinnummernr
- fuhrerscheinnummernr
- fuehrerscheinnummernr
- führerscheinnummerklasse
- fuhrerscheinnummerklasse
- fuehrerscheinnummerklasse
- nr-führerschein
- nr-fuhrerschein
- nr-fuehrerschein
- no-führerschein
- no-fuhrerschein
- no-fuehrerschein
- n-führerschein
- n-fuhrerschein
- n-fuehrerschein
- permis de conduire
- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences

- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#

- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#

- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses
- driving licence
- driving licences
- driving permit
- dlno

Germany identity card number

Format

since 1 November 2010: Nine letters and digits

from 1 April 1987 until 31 October 2010: 10 digits

Pattern

since 1 November 2010:

- one letter (not case-sensitive)
- eight digits

from 1 April 1987 until 31 October 2010:

- 10 digits

Checksum

No

Keywords

Keyword_germany_id_card

- ausweis
- gpid
- identification
- identifikation
- identifizierungsnummer
- identity card
- identity number
- id-nummer
- personal id
- personalausweis
- persönliche id nummer
- persönliche identifikationsnummer
- persönliche-id-nummer

Germany passport number

This entity is included in the EU Passport Number sensitive information type and is available as a stand-alone sensitive information type entity.

Format

10 digits or letters

Pattern

Pattern must include all of the following:

- first character is a digit or a letter from this set (C, F, G, H, J, K)
- three digits
- five digits or letters from this set (C, -H, J-N, P, R, T, V-Z)
- a digit

Checksum

Yes

Keywords

Keyword_german_passport

- reisepasse
- reisepassnummer
- No-Reisepass
- Nr-Reisepass
- Reisepass-Nr
- Passnummer
- reisepässe
- passeport no.
- passeport no

Keywords_eu_passport_number_common

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers

Germany tax identification number

Format

11 digits without spaces and delimiters

Pattern

11 digits

- Two digits
- An optional space
- Three digits
- An optional space
- Three digits
- An optional space
- Two digits
- one check digit

Checksum

Yes

Keywords

Keywords_germany_eu_tax_file_number

- identifikationsnummer
- steuer id
- steueridentifikationsnummer
- steuernummer
- tax id
- tax identification no
- tax identification number
- tax no#
- tax no
- tax number
- tax registration number
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id
- tin no
- tin#
- zinn#

- zinn
- zinnnummer

Germany value added tax number

Format

11 character alphanumeric pattern

Pattern

11-character alphanumeric pattern:

- a letter D or d
- a letter E or e
- an optional space
- three digits
- an optional space or comma
- three digits
- an optional space or comma
- three digits

Checksum

Yes

Keywords

Keyword_germany_value_added_tax_number

- vat number
- vat no
- vat#
- vat## mehrwertsteuer
- mwst
- mehrwertsteuer identifikationsnummer
- mehrwertsteuer nummer

Greece driver's license number

This entity is included in the EU Driver's License Number sensitive information type. It's also available as a stand-alone sensitive information type entity.

Format

nine digits without spaces and delimiters

Pattern

nine digits

Checksum

No

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic

- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#

- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic

- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses
- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_greece_eu_driver's_license_number

- δεια οδήγησης
- Adeia odigisis
- Άδεια οδήγησης
- Δίπλωμα οδήγησης

Greece national ID card

Format

Combination of 7-8 letters and numbers plus a dash

Pattern

Seven letters and numbers (old format):

- One letter (any letter of the Greek alphabet)
- A dash
- Six digits

Eight letters and numbers (new format):

- Two letters whose uppercase character occurs in both the Greek and Latin alphabets (ABEZHIKMNOPTYX)
- A dash
- Six digits

Checksum

No

Keywords

Keyword_greece_id_card

- greek id
- greek national id
- greek personal id card
- greek police id
- identity card
- tautotita
- ταυτότητα
- ταυτότητας

Greece passport number

Format

Two letters followed by seven digits with no spaces or delimiters

Pattern

Two letters followed by seven digits

Checksum

No

Keywords

Keywords_eu_passport_number

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers

Keywords_greece_eu_passport_number

- αριθμός διαβατηρίου
- αριθμούς διαβατηρίου
- αριθμός διαβατηριο

Greece Social Security Number (AMKA)

Format

11 digits without spaces and delimiters

Pattern

- Six digits as date of birth YYMMDD
- Four digits
- a check digit

Checksum

Yes

Keywords

Keywords_greece_eu_ssn_or_equivalent

- ssn
- ssn#
- social security no
- socialsecurityno#
- social security number
- amka
- a.m.k.a.
- Αριθμού Μητρώου Κοινωνικής Ασφάλισης

Greece tax identification number

Format

Nine digits without spaces and delimiters

Pattern

Nine digits

Checksum

Not applicable

Keywords

Keywords_greece_eu_tax_file_number

- afm#
- afm
- αφμ|αφμ αριθμός
- αφμ
- tax id
- tax identification no
- tax identification number
- tax no#
- tax no
- tax number
- tax registration number
- tax registry no

- tax registry number
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- taxregistryno#
- tin id
- tin no
- tin#
- αριθμός φορολογικού μητρώου
- τον αριθμό φορολογικού μητρώου
- φορολογικού μητρώου νο

Hong Kong identity card (HKID) number

Format

Combination of 8-9 letters and numbers plus optional parentheses around the final character

Pattern

Combination of 8-9 letters:

- 1-2 letters (not case-sensitive)
- Six digits
- The final character (any digit or the letter A), which is the check digit and is optionally enclosed in parentheses.

Checksum

Yes

Keywords

Keyword_hong_kong_id_card

- hkid
- hong kong identity card
- HKIDC
- id card
- identity card
- hk identity card
- hong kong id
- 香港身份證
- 香港永久性居民身份證
- 身份證
- 身份証
- 身分證
- 身分証
- 香港身份証
- 香港身分證
- 香港身分証
- 香港身份證
- 香港居民身份證
- 香港居民身份証
- 香港居民身分證
- 香港居民身分証
- 香港永久性居民身份証
- 香港永久性居民身分證
- 香港永久性居民身分証
- 香港永久性居民身份證
- 香港非永久性居民身份證
- 香港非永久性居民身份証
- 香港非永久性居民身分證
- 香港非永久性居民身分証
- 香港特別行政區永久性居民身份證
- 香港特別行政區永久性居民身份証
- 香港特別行政區永久性居民身分證
- 香港特別行政區永久性居民身分証
- 香港特別行政區非永久性居民身份證
- 香港特別行政區非永久性居民身份証
- 香港特別行政區非永久性居民身分證
- 香港特別行政區非永久性居民身分証

Hungary driver's license number

Format

Two letters followed by six digits

Pattern

Two letters and six digits:

- Two letters (not case-sensitive)
- Six digits

Checksum

No

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses

- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#

- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses

- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses
- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_hungary_eu_driver's_license_number

- vezetoi engedely
- vezetői engedély
- vezetői engedélyek

Hungary personal identification number

Format

11 digits

Pattern

11 digits:

- One digit that corresponds to gender, 1 for male, 2 for female. Other numbers are also possible for citizens born before 1900 or citizens with double citizenship.
- Six digits that correspond to birth date (YYMMDD)
- Three digits that correspond to a serial number
- One check digit

Checksum

Yes

Keywords

Keywords_hungary_eu_national_id_card

- id number
- identification number
- sz ig
- sz. ig.
- sz.ig.
- személyazonosító igazolvány
- személyi igazolvány

Hungary passport number

Format

Two letters followed by six or seven digits with no spaces or delimiters

Pattern

Two letters followed by six or seven digits

Checksum

No

Keywords

Keywords_eu_passport_number

- passport#
- passport #
- passportid
- passports
- passportno
- passport no

- passportnumber
- passport number
- passportnumbers
- passport numbers

Keywords_hungary_eu_passport_number

- útlevél száma
- Útlevelek száma
- útlevél szám

Keywords_eu_passport_date

- date of issue
- date of expiry

Hungary social security number (TAJ)

Format

Nine digits without spaces and delimiters

Pattern

Nine digits

Checksum

Yes

Keywords

Keywords_hungary_eu_ssn_or_equivalent

- hungarian social security number
- social security number
- socialsecuritynumber#
- hssn#
- socialsecuritynno

- hssn
- taj
- taj#
- ssn
- ssn#
- social security no
- áfa
- közösségi adószám
- általános forgalmi adó szám
- hozzáadottérték adó
- áfa szám
- magyar áfa szám

Hungary tax identification number

Format

10 digits with no spaces or delimiters

Pattern

10 digits:

- One digit that must be "8"
- Eight digits
- One check digit

Checksum

Yes

Keywords

Keywords_hungary_eu_tax_file_number

- adóazonosító szám
- adóhatóság szám
- adószám
- hungarian tin

- hungatiantin#
- tax authority no
- tax id
- tax identification no
- tax identification number
- tax no#
- tax no
- tax number
- tax registration number
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id
- tin no
- tin#
- vat number

Hungary value added tax number

Format

10 character alphanumeric pattern

Pattern

10 character alphanumeric pattern:

- two letters - HU or hu
- optional space
- eight digits

Checksum

Yes

Keywords

Keyword_Hungary_value_added_tax_number

- vat
- value added tax number
- vat#
- vatno#
- hungarianvatno#
- tax no.
- value added tax áfa
- közösségi adószám
- általános forgalmi adó szám
- hozzáadottérték adó
- áfa szám

India permanent account number (PAN)

Format

10 letters or digits

Pattern

10 letters or digits:

- Three letters (not case-sensitive)
- A letter in C, P, H, F, A, T, B, L, J, G (not case-sensitive)
- A letter
- Four digits
- A letter that is an alphabetic check digit

Checksum

No

Keywords

Keyword_india_permanent_account_number

- Permanent Account Number

- PAN

India unique identification (Aadhaar) number

Format

12 digits containing optional spaces or dashes

Pattern

12 digits:

- A digit that is not 0 or 1
- Three digits
- An optional space or dash
- Four digits
- An optional space or dash
- The final digit, which is the check digit

Checksum

Yes

Keywords

Keyword_india_aadhar

- aadhaar
- aadhar
- aadhar#
- uid
- आधार
- uidai

Indonesia identity card (KTP) number

Format

16 digits containing optional periods

Pattern

16 digits:

- Two-digit province code
- A period (optional)
- Two-digit regency or city code
- Two-digit subdistrict code
- A period (optional)
- Six digits in the format DDMMYY, which are the date of birth
- A period (optional)
- Four digits

Checksum

No

Keywords

Keyword_indonesia_id_card

- KTP
- Kartu Tanda Penduduk
- Nomor Induk Kependudukan

International banking account number (IBAN)

Format

Country code (two letters) plus check digits (two digits) plus bban number (up to 30 characters)

Pattern

Pattern must include all of the following:

- Two-letter country code

- Two check digits (followed by an optional space)
- 1-7 groups of four letters or digits (can be separated by spaces)
- 1-3 letters or digits

The format for each country is slightly different. The IBAN sensitive information type covers these 60 countries/regions:

- ad
- ae
- al
- at
- az
- ba
- be
- bg
- bh
- ch
- cr
- cy
- cz
- de
- dk
- do
- ee
- es
- fi
- fo
- fr
- gb
- ge
- gi
- gl
- gr
- hr
- hu
- ie
- il
- is
- it
- kw
- kz

- lb
- li
- lt
- lu
- lv
- mc
- md
- me
- mk
- mr
- mt
- mu
- nl
- no
- pl
- pt
- ro
- rs
- sa
- se
- si
- sk
- sm
- tn
- tr
- vg

Checksum

Yes

Keywords

None

IP address

Format

IPv4:

Complex pattern that accounts for formatted (periods) and unformatted (no periods) versions of the IPv4 addresses

IPv6:

Complex pattern that accounts for formatted IPv6 numbers (which include colons)

Pattern

- 000.00.000.00
- 000.000.00.00
- 00.000.000.00
- 000.000.000.00
- 000.000.000.000
- 0000:0000:00a0:0:00a:0b00:0a0a:0ea
- 0000:0000:0000::00:00:000:0
- 0a0a:a000:00a::
- 0000:0000:0000:00::0
- 0a00:00a0:a000:0000:a00a:a00a:00a:aaa0

Checksum

No

Keywords

Keyword_ipaddress

- IP (this keyword is case-sensitive)
- ip address
- ip addresses
- internet protocol
- כתובת ה-IP

Personal IP Address

Format

Complex pattern that accounts for formatted (periods) versions of the IPv4 addresses

Pattern

000.000.000.000

Checksum

No

Keywords

Keyword_ipaddress

- IP (case sensitive)
- ip address
- ip addresses
- internet protocol
- כתובת ה-IP

Ireland driver's license number

Format

Six digits followed by four letters

Pattern

Six digits and four letters:

- Six digits
- Four letters (not case-sensitive)

Checksum

No

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense

- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#

- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses
- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_ireland_eu_driver's_license_number

- ceadúnas tiomána
- ceadúnais tiomána

Ireland passport number

Format

Two letters or digits followed by seven digits with no spaces or delimiters

Pattern

Two letters or digits followed by seven digits:

- Two digits or letters (not case-sensitive)
- Seven digits

Checksum

No

Keywords

Keywords_eu_passport_number_common

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers

Keywords_ireland_eu_passport_number

- passeport numero
- uimhreacha pasanna

- uimhir pas
- uimhir phas
- uimhreacha pas
- uimhir cárta
- uimhir chárta

Keywords_eu_passport_date

- date of issue
- date of expiry

Ireland personal public service (PPS) number

Format

Old format (until 31 December 2012):

- seven digits followed by 1-2 letters

New format (1 January 2013 and after):

- seven digits followed by two letters

Pattern

Old format (until 31 December 2012):

- seven digits
- one to two letters (not case-sensitive)

New format (1 January 2013 and after):

- seven digits
- a letter (not case-sensitive) which is an alphabetic check digit
- An optional letter in the range A-I, or “W”

Checksum

Yes

Keywords

Keywords_ireland_eu_national_id_card

- client identity service
- identification number
- personal id number
- personal public service number
- personal service no
- phearsanta seirbhíse poiblí
- pps no
- pps number
- pps num
- pps service no
- ppsn
- ppsno#
- ppsno
- psp
- public service no
- publicserviceno#
- publicserviceno
- revenue and social insurance number
- rsi no
- rsi number
- rsin
- seirbhís aitheantais cliant
- uimh
- uimhir aitheantais chánach
- uimhir aitheantais phearsanta
- uimhir phearsanta seirbhíse poiblí
- tax id
- tax identification no
- tax identification number
- tax no#
- tax no
- tax number
- tax registration number
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber

- tin id
- tin no
- tin#

Israel bank account number

Format

13 digits

Pattern

Formatted:

- two digits
- a dash
- three digits
- a dash
- eight digits

Unformatted:

- 13 consecutive digits

Checksum

No

Keywords

Keyword_israel_bank_account_number

- Bank Account Number
- Bank Account
- Account Number
- מספר חשבון בנק

Israel national identification number

Format

nine digits

Pattern

nine consecutive digits

Checksum

Yes

Keywords

Keyword_Israel_National_ID

- מספר זהות
- מספר זיה וי
- מספר זיהוי ישר אלי
- זהותישר אלית
- هو ية اسرائيل ية عدد
- هوية إسرائ يلية
- رقم الهوية
- عدد هوية فريدة من نوعها
- idnumber#
- id number
- identity no
- identitynumber#
- identity number
- israeliidentitynumber
- personal id
- unique id

Italy driver's license number

This type entity is included in the EU Driver's License Number sensitive information type. It's also available as a stand-alone sensitive information type entity.

Format

a combination of 10 letters and digits

Pattern

a combination of 10 letters and digits:

- one letter (not case-sensitive)
- the letter "A" or "V" (not case-sensitive)
- seven digits
- one letter (not case-sensitive)

Checksum

No

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license

- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#

- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license

- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses
- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keyword_italy_drivers_license_number

- numero di patente
- patente di guida
- patente guida
- patenti di guida
- patenti guida

Italy fiscal code

Format

a 16-character combination of letters and digits in the specified pattern

Pattern

A 16-character combination of letters and digits:

- three letters that correspond to the first three consonants in the family name
- three letters that correspond to the first, third, and fourth consonants in the first name
- two digits that correspond to the last digits of the birth year

- one letter that corresponds to the letter for the month of birth—letters are used in alphabetical order, but only the letters A to E, H, L, M, P, R to T are used (so, January is A and October is R)
- two digits that correspond to the day of the month of birth in order to differentiate between genders, 40 is added to the day of birth for women
- four digits that correspond to the area code specific to the municipality where the person was born (country-wide codes are used for foreign countries/regions)
- one parity digit

Checksum

Yes

Keywords

Keywords_italy_eu_national_id_card

- codice fiscal
- codice fiscale
- codice id personale
- codice personale
- fiscal code
- numero certificato personale
- numero di identificazione fiscale
- numero id personale
- numero personale
- personal certificate number
- personal code
- personal id code
- personal id number
- personalcodeno#
- tax code
- tax id
- tax identification no
- tax identification number
- tax identity number
- tax no#
- tax no
- tax number
- tax registration number

- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id
- tin no
- tin#

Italy passport number

Format

two letters or digits followed by seven digits with no spaces or delimiters

Pattern

two letters or digits followed by seven digits:

- two digits or letters (not case-sensitive)
- seven digits

Checksum

not applicable

Keywords

Keywords_eu_passport_number_common

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number

- passportnumbers
- passport numbers

Keywords_italy_eu_passport_number

- italiana passaporto
- passaporto italiana
- passaporto numero
- numéro passeport
- numero di passaporto
- numeri del passaporto
- passeport italien

Keywords_eu_passport_date

- date of issue
- date of expiry

Italy value added tax number

Format

13 character alphanumeric pattern with optional delimiters

Pattern

13 character alphanumeric pattern with optional delimiters:

- I or i
- T or t
- optional space, dot, hyphen, or comma
- 11 digits

Checksum

Yes

Keywords

Keyword_italy_value_added_tax_number

- vat number
- vat no
- vat#
- iva
- iva#

Japan bank account number

Format

seven or eight digits

Pattern

bank account number:

- seven or eight digits
- bank account branch code:
- four digits
- a space or dash (optional)
- three digits

Checksum

No

Keywords

Keyword_jp_bank_account

- Checking Account Number
- Checking Account
- Checking Account #
- Checking Acct Number
- Checking Acct #
- Checking Acct No.
- Checking Account No.
- Bank Account Number

- Bank Account
- Bank Account #
- Bank Acct Number
- Bank Acct #
- Bank Acct No.
- Bank Account No.
- Savings Account Number
- Savings Account
- Savings Account #
- Savings Acct Number
- Savings Acct #
- Savings Acct No.
- Savings Account No.
- Debit Account Number
- Debit Account
- Debit Account #
- Debit Acct Number
- Debit Acct #
- Debit Acct No.
- Debit Account No.
- 口座番号
- 銀行口座
- 銀行口座番号
- 総合口座
- 普通預金口座
- 普通口座
- 当座預金口座
- 当座口座
- 預金口座
- 振替口座
- 銀行
- バンク

Keyword_jp_bank_branch_code

- 支店番号
- 支店コード
- 店番号

Japan driver's license number

Format

12 digits

Pattern

12 consecutive digits

Checksum

No

Keywords

Keyword_jp_drivers_license_number

- driverlicense
- driverslicense
- driver'slicense
- driverslicenses
- driver'slicenses
- driverlicenses
- dl#
- dls#
- lic#
- lics#
- 運転免許証
- 運転免許
- 免許証
- 免許
- 運転免許証番号
- 運転免許番号
- 免許証番号
- 免許番号
- 運転免許証ナンバー
- 運転免許ナンバー
- 免許証ナンバー
- 運転免許証no

- 運転免許no
- 免許証no
- 免許no
- 運転経歴証明書番号
- 運転経歴証明書
- 運転免許証No.
- 運転免許No.
- 免許証No.
- 免許No.
- 運転免許証#
- 運転免許#
- 免許証#
- 免許#

Japan My Number - Corporate

Format

13-digit number

Pattern

13-digit number:

- one digit from one to nine
- 12 digits

Checksum

Yes

Keywords

Keyword_japan_my_number_corporate

- corporate number
- マイナンバー
- 共通番号
- マイナンバーカード

- マイナンバーカード番号
- 個人番号カード
- 個人識別番号
- 個人識別ナンバー
- 法人番号
- 指定通知書

Japan My Number - Personal

Format

12-digit number

Pattern

12-digit number:

- four digits
- an optional space, dot, or hyphen
- four digits
- an optional space, dot, or hyphen
- four digits

Checksum

Yes

Keywords

Keyword_japan_my_number_personal

- my number
- マイナンバー
- 個人番号
- 共通番号
- マイナンバーカード
- マイナンバーカード番号
- 個人番号カード
- 個人識別番号

- 個人識別ナンバー
- 通知カード

Japan passport number

Format

two letters followed by seven digits

Pattern

two letters (not case-sensitive) followed by seven digits

Checksum

No

Keywords

Keyword_jp_passport

- Passport
- Passport Number
- Passport No.
- Passport #
- パスポート
- パスポート番号
- パスポートナンバー
- パスポート＃
- パスポート#
- パスポートNo.
- 旅券番号
- 旅券番号＃
- 旅券番号♯
- 旅券ナンバー

Japan residence card number

Format

12 letters and digits

Pattern

12 letters and digits:

- two letters (not case-sensitive)
- eight digits
- two letters (not case-sensitive)

Checksum

No

Keywords

Keyword_jp_residence_card_number

- Residence card number
- Residence card no
- Residence card #
- 在留カード番号
- 在留カード
- 在留番号

Japan resident registration number

Format

11 digits

Pattern

11 consecutive digits

Checksum

No

Keywords

Keyword_jp_resident_registration_number

- Resident Registration Number
- Residents Basic Registry Number
- Resident Registration No.
- Resident Register No.
- Residents Basic Registry No.
- Basic Resident Register No.
- 外国人登録証明書番号
- 証明書番号
- 登録番号
- 外国人登録証

Japan social insurance number (SIN)

Format

7-12 digits

Pattern

7-12 digits:

- four digits
- a hyphen (optional)
- six digits

or

- 7-12 consecutive digits

Checksum

No

Keywords

Keyword_jp_sin

- Social Insurance No.
- Social Insurance Num
- Social Insurance Number
- 健康保険被保険者番号
- 健保番号
- 基礎年金番号
- 雇用保険被保険者番号
- 雇用保険番号
- 保険証番号
- 社会保険番号
- 社会保険No.
- 社会保険
- 介護保険
- 介護保険被保険者番号
- 健康保険被保険者整理番号
- 雇用保険被保険者整理番号
- 厚生年金
- 厚生年金被保険者整理番号

Latvia driver's license number

Format

three letters followed by six digits

Pattern

three letters and six digits:

- three letters (not case-sensitive)
- six digits

Checksum

No

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic

- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#

- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses
- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_latvia_eu_driver's_license_number

- autovadītāja apliecība
- autovadītāja apliecības
- vadītāja apliecība

Latvia Personal Code

Format

11 digits and an optional hyphen

Pattern

Old format

11 digits and a hyphen:

- six digits that correspond to the birth date (DDMMYY)
- a hyphen
- one digit that corresponds to the century of birth ("0" for 19th century, "1" for 20th century, and "2" for 21st century)
- four digits, randomly generated

New format

11 digits

- Two digits "32"
- Nine digits

Checksum

Yes

Keywords

Keywords_latvia_eu_national_id_card

- administrative number
- alvas nē

- birth number
- citizen number
- civil number
- electronic census number
- electronic number
- fiscal code
- healthcare user number
- id#
- id-code
- identification number
- identifikācijas numurs
- id-number
- individual number
- latvija alva
- nacionālais id
- national id
- national identifying number
- national identity number
- national insurance number
- national register number
- nodokļa numurs
- nodokļu id
- nodokļu identifikācija numurs
- personal certificate number
- personal code
- personal id code
- personal id number
- personal identification code
- personal identifier
- personal identity number
- personal number
- personal numeric code
- personalcodeno#
- personas kods
- population identification code
- public service number
- registration number
- revenue number
- social insurance number
- social security number
- state tax code

- tax file number
- tax id
- tax identification no
- tax identification number
- tax no#
- tax no
- tax number
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id
- tin no
- tin#
- voter’s number

Latvia passport number

Format

two letters or digits followed by seven digits with no spaces or delimiters

Pattern

two letters or digits followed by seven digits:

- two digits or letters (not case-sensitive)
- seven digits

Checksum

No

Keywords

Keywords_eu_passport_number_common

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers

Keywords_latvia_eu_passport_number

- pase numurs
- pase numur
- pases numuri
- pases nr
- passeport no
- n° du Passeport

Keywords_eu_passport_date

- date of issue
- date of expiry

Lithuania driver's license number

Format

eight digits without spaces and delimiters

Pattern

eight digits

Checksum

No

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic

- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#

- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses
- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_lithuania_eu_driver's_license_number

- vairuotojo pažymėjimas
- vairuotojo pažymėjimo numeris
- vairuotojo pažymėjimo numeriai

Lithuania Personal Code

Format

11 digits without spaces and delimiters

Pattern

11 digits without spaces and delimiters:

- one digit (1-6) that corresponds to the person's gender and century of birth
- six digits that correspond to birth date (YYMMDD)
- three digits that correspond to the serial number of the date of birth
- one check digit

Checksum

Yes

Keywords

Keywords_lithuania_eu_national_id_card

- asmeninis skaitmeninis kodas
- asmens kodas
- citizen service number
- mokesčių id
- mokesčių identifikavimas numeris
- mokesčių identifikavimo numeris
- mokesčių numeris
- national identification number
- personal code
- personal numeric code

- piliečio paslaugos numeris
- tax id
- tax identification no
- tax identification number
- tax no#
- tax no
- tax number
- tax registration number
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id
- tin no
- tin#
- unikalus identifikavimo kodas
- unikalus identifikavimo numeris
- unique identification number
- unique identity number
- uniqueidentityno#

Lithuania passport number

Format

eight digits or letters with no spaces or delimiters

Pattern

eight digits or letters (not case-sensitive)

Checksum

not applicable

Keywords

Keywords_eu_passport_number

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers

Keywords_lithuania_eu_passport_number

- paso numeris
- paso numeriai
- paso nr

Keywords_eu_passport_date

- date of issue
- date of expiry

Location

Format

Longitude can range from -180.0 to 180.0. Latitude can range from -90.0 to 90.0.

Checksum

Not applicable

Keywords

- lat
- latitude
- long

- longitude
- coord
- coordinates
- geo
- geolocation
- loc
- location
- position

Luxemburg driver's license number

Format

six digits without spaces and delimiters

Pattern

six digits

Checksum

No

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences

- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#

- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#

- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses
- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_luxemburg_eu_driver's_license_number

- fahrerlaubnis
- Führerschäin

Luxemburg national identification number natural persons

Format

13 digits with no spaces or delimiters

Pattern

13 digits:

- 11 digits
- two check digits

Checksum

yes

Keywords

Keywords_luxemburg_eu_national_id_card

- eindeutige id
- eindeutige id-nummer
- eindeutigeid#
- id personnelle
- idpersonnelle#
- idpersonnelle
- individual code
- individual id
- individual identification
- individual identity
- numéro d'identification personnel
- personal id
- personal identification
- personal identity
- personalidno#
- personalidnumber#
- persönliche identifikationsnummer
- unique id
- unique identity
- uniqueidkey#

Luxemburg national identification number non-natural persons

Format

11 digits

Pattern

11 digits

- two digits
- an optional space
- three digits
- an optional space
- three digits
- an optional space
- two digits
- one check digit

Checksum

Yes

Keywords

Keywords_luxemburg_eu_tax_file_number

- carte de sécurité sociale
- étain non
- étain#
- identifiant d'impôt
- luxembourg tax identifikatiounsnummer
- numéro d'étain
- numéro d'identification fiscal luxembourgeois
- numéro d'identification fiscale
- social security
- sozialunterstützung
- sozialversécherung
- sozialversicherungsausweis
- steier id
- steier identifikatiounsnummer
- steier nummer
- steuer id
- steueridentifikationsnummer

- steuernummer
- tax id
- tax identification no
- tax identification number
- tax no#
- tax no
- tax number
- tax registration number
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id
- tin no
- tin#
- zinn#
- zinn
- zinnzahl

Luxemburg passport number

Format

eight digits or letters with no spaces or delimiters

Pattern

eight digits or letters (not case-sensitive)

Checksum

No

Keywords

Keywords_eu_passport_number

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers

Keywords_luxemburg_eu_passport_number

- ausweisnummer
- luxembourg pass
- luxembourg passeport
- luxembourg passport
- no de passeport
- no-reisepass
- nr-reisepass
- numéro de passeport
- pass net
- pass nr
- passnummer
- passeport nombre
- reisepässe
- reisepass-nr
- reisepassnummer

Keywords_eu_passport_date

- date of issue
- date of expiry

Malaysia identification card number

Format

12 digits containing optional hyphens

Pattern

12 digits:

- six digits in the format YYMMDD, which are the date of birth
- a dash (optional)
- two-letter place-of-birth code
- a dash (optional)
- three random digits
- one-digit gender code

Checksum

No

Keywords

Keyword_malaysia_id_card_number

- digital application card
- i/c
- i/c no
- ic
- ic no
- id card
- identification Card
- identity card
- k/p
- k/p no
- kad akuan diri
- kad aplikasi digital
- kad pengenalan malaysia
- kp
- kp no
- mykad
- mykas
- mykid
- mypr
- mytentera
- malaysia identity card
- malaysian identity card

- nric
- personal identification card

Malta driver's license number

Format

Combination of two characters and six digits in the specified pattern

Pattern

combination of two characters and six digits:

- two characters (digits or letters, not case-sensitive)
- a space (optional)
- three digits
- a space (optional)
- three digits

Checksum

No

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences

- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#

- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#

- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses
- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_malta_eu_driver's_license_number

- liċenzja tas-sewqan
- liċenzji tas-sewwieq

Malta identity card number

Format

seven digits followed by one letter

Pattern

seven digits followed by one letter:

- seven digits
- one letter in "M, G, A, P, L, H, B, Z" (case insensitive)

Checksum

Not applicable

Keywords

Keywords_malta_eu_national_id_card

- citizen service number
- id tat-taxxa
- identifika numru tal-biljett
- kodiċi numerali personali
- numru ta 'identifikazzjoni personali
- numru ta 'identifikazzjoni tat-taxxa
- numru ta 'identifikazzjoni uniku
- numru ta' identità uniku
- numru tas-servizz taċ-ċittadin
- numru tat-taxxa
- personal numeric code
- unique identification number
- unique identity number
- uniqueidentityno#

Malta passport number

Format

seven digits without spaces or delimiters

Pattern

seven digits

Checksum

No

Keywords

Keywords_eu_passport_number

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers

Keywords_malta_eu_passport_number

- numru tal-passaport
- numri tal-passaport
- Nru tal-passaport

Keywords_eu_passport_date

- date of issue
- date of expiry

Malta tax identification number

Format

For Maltese nationals:

- seven digits and one letter in the specified pattern

Non-Maltese nationals and Maltese entities:

- nine digits

Pattern

Maltese nationals: seven digits and one letter

- seven digits
- one letter (not case-sensitive)

Non-Maltese nationals and Maltese entities: nine digits

- nine digits

Checksum

Not applicable

Keywords

Keywords_malta_eu_tax_file_number

- citizen service number
- id tat-taxxa
- identifika numru tal-biljett
- kodiċi numerali personali
- numru ta 'identifikazzjoni personali
- numru ta 'identifikazzjoni tat-taxxa
- numru ta 'identifikazzjoni uniku
- numru ta' identità uniku
- numru tas-servizz taċ-ċittadin
- numru tat-taxxa
- personal numeric code
- tax id
- tax identification no
- tax identification number
- tax no#
- tax no
- tax number
- tax registration number
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#

- taxnumber
- tin id
- tin no
- tin#
- unique identification number
- unique identity number
- uniqueidentityno#

Netherlands citizen's service (BSN) number

Format

eight or nine digits containing optional spaces

Pattern

eight-nine digits:

- three digits
- a space (optional)
- three digits
- a space (optional)
- two-three digits

Checksum

Yes

Keywords

Keywords_netherlands_eu_national_id_card

- bsn#
- bsn
- burgerservicenummer
- citizen service number
- person number
- personal number
- personal numeric code

- person-related number
- persoonlijk nummer
- persoonlijke numerieke code
- persoonsgebonden
- persoonsnummer
- sociaal-fiscaal nummer
- social-fiscal number
- sofi
- sofinummer
- uniek identificatienummer
- uniek identiteitsnummer
- unique identification number
- unique identity number
- uniqueidentityno#

Netherlands driver's license number

Format

10 digits without spaces and delimiters

Pattern

10 digits

Checksum

No

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences

- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence

- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#

- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses
- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_netherlands_eu_driver's_license_number

- permis de conduire
- rijbewijs
- rijbewijsnummer
- rijbewijzen
- rijbewijs nummer
- rijbewijsnummers

Netherlands passport number

Format

nine letters or digits with no spaces or delimiters

Pattern

nine letters or digits

Checksum

not applicable

Keywords

Keywords_eu_passport_number

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers

Keywords_netherlands_eu_passport_number

- paspoort nummer
- paspoortnummers
- paspoortnummer
- paspoort nr

Netherlands tax identification number

Format

nine digits without spaces or delimiters

Pattern

nine digits

Checksum

Yes

Keywords

Keywords_netherlands_eu_tax_file_number

- btw nummer
- hollânske tax identification
- hulandes impuesto id number
- hulandes impuesto identification
- identificatienummer belasting
- identificatienummer van belasting
- impuesto identification number
- impuesto number
- nederlands belasting id nummer
- nederlands belasting identificatie
- nederlands belasting identificatienummer
- nederlands belastingnummer
- nederlandse belasting identificatie
- netherlands tax identification
- netherland's tax identification
- netherlands tin
- netherland's tin
- tax id
- tax identification no
- tax identification number
- tax identification tal
- tax no#
- tax no
- tax number

- tax registration number
- tax tal
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id
- tin no
- tin#

Netherlands value added tax number

Format

14 character alphanumeric pattern

Pattern

14-character alphanumeric pattern:

- N or n
- L or l
- optional space, dot, or hyphen
- nine digits
- optional space, dot, or hyphen
- B or b
- two digits

Checksum

Yes

Keywords

Keyword_netherlands_value_added_tax_number

- vat number

- vat no
- vat#
- wearde tafoege tax getal
- btw nûmer
- btw-nummer

New Zealand bank account number

Format

14-digit to 16-digit pattern with optional delimiter

Pattern

14-digit to 16-digit pattern with optional delimiter:

- two digits
- an optional hyphen or space
- three to four digits
- an optional hyphen or space
- seven digits
- an optional hyphen or space
- two to three digits
- an options hyphen or space

Checksum

Yes

Keywords

Keyword_new_zealand_bank_account_number

- account number
- bank account
- bank_acct_id
- bank_acct_branch
- bank_acct_nbr

New Zealand driver's license number

Format

eight character alphanumeric pattern

Pattern

eight character alphanumeric pattern

- two letters
- six digits

Checksum

Yes

Keywords

Keyword_new_zealand_drivers_license_number

- driverlicence
- driverlicences
- driver lic
- driver licence
- driver licences
- driverslic
- driverslicence
- driverslicences
- drivers lic
- drivers lics
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'licence
- driver'licences
- driver' lic
- driver' lics
- driver' licence

- driver' licences
- driver'slic
- driver'slics
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's licence
- driver's licences
- driverlic#
- driverlics#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver licence#
- driver licences#
- driverslic#
- driverslics#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's licence#
- driver's licences#

- international driving permit
- international driving permits
- nz automobile association
- new zealand automobile association

New Zealand inland revenue number

Format

eight or nine digits with optional delimiters

Pattern

eight or nine digits with optional delimiters

- two or three digits
- an optional space or hyphen
- three digits
- an optional space or hyphen
- three digits

Checksum

Yes

Keywords

Keyword_new_zealand_inland_revenue_number

- ird no.
- ird no#
- nz ird
- new zealand ird
- ird number
- inland revenue number

New Zealand ministry of health number

Format

three letters, a space (optional), and four digits

Pattern

- three letters (not case-sensitive) except 'I' and 'O'
- a space (optional)
- four digits

Checksum

Yes

Keywords

Keyword_nz_terms

- NHI
- New Zealand
- Health
- treatment
- National Health Index Number
- nhi number
- nhi no.
- NHI#
- National Health Index No.
- National Health Index Id
- National Health Index #

New Zealand social welfare number

Format

nine digits

Pattern

nine digits

- three digits
- an optional hyphen
- three digits
- an optional hyphen
- three digits

Checksum

Yes

Keywords

Keyword_new_zealand_social_welfare_number

- social welfare #
- social welfare#
- social welfare No.
- social welfare number
- swn#

Norway identification number

Format

11 digits

Pattern

11 digits:

- six digits in the format DDMMYY, which are the date of birth
- three-digit individual number
- two check digits

Checksum

Yes

Keywords

Keyword_norway_id_number

- Personal identification number
- Norwegian ID Number
- ID Number
- Identification
- Personnummer
- Fødselsnummer

Philippines unified multi-purpose identification number

Format

12 digits separated by hyphens

Pattern

12 digits:

- four digits
- a hyphen
- seven digits
- a hyphen
- one digit

Checksum

No

Keywords

Keyword_philippines_id

- Unified Multi-Purpose ID
- UMID
- Identity Card
- Pinag-isang Multi-Layunin ID

Poland driver's license number

Format

14 digits containing two forward slashes

Pattern

14 digits and two forward slashes:

- five digits
- a forward slash
- two digits
- a forward slash
- seven digits

Checksum

No

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence

- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#

- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#

- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses
- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_poland_eu_driver's_license_number

- prawo jazdy
- prawa jazdy

Poland identity card

Format

three letters and six digits

Pattern

three letters (not case-sensitive) followed by six digits

Checksum

Yes

Keywords

Keyword_poland_national_id_passport_number

- Dowód osobisty
- Numer dowodu osobistego
- Nazwa i numer dowodu osobistego
- Nazwa i nr dowodu osobistego
- Nazwa i nr dowodu tożsamości
- Dowód Tożsamości
- dow. os.

Poland national ID (PESEL)

Format

11 digits

Pattern

- six digits representing date of birth in the format YYMMDD
- four digits
- one check digit

Checksum

Yes

Keywords

Keyword_pesel_identification_number

- dowód osobisty
- dowódosobisty
- niepowtarzalny numer
- niepowtarzalnynumer

- nr.-pesel
- nr-pesel
- numer identyfikacyjny
- pesel
- tożsamości narodowej

Poland passport number

This sensitive information type entity is included in the EU Passport Number sensitive information type. It's also available as a stand-alone sensitive information type entity.

Format

two letters and seven digits

Pattern

Two letters (not case-sensitive) followed by seven digits

Checksum

Yes

Keywords

Keywords_eu_passport_number

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers

Keyword_polish_national_passport_number

- numer paszportu
- numery paszportów
- numery paszportowe
- nr paszportu
- nr. paszportu
- nr paszportów
- n° passeport
- passeport n°

Keywords_eu_passport_date

- date of issue
- date of expiry

Poland REGON number

Format

9-digit or 14-digit number

Pattern

nine digit or 14-digit number:

- nine digits

or

- nine digits
- hyphen
- five digits

Checksum

Yes

Keywords

Keywords_poland_regon_number

- regon id
- statistical number
- statistical id
- statistical no
- regon number
- regonid#
- regonno#
- company id
- companyid#
- companyidno#
- numer statystyczny
- numeru regon
- numerstatystyczny#
- numeruregon#

Poland tax identification number

Format

11 digits with no spaces or delimiters

Pattern

11 digits

Checksum

Yes

Keywords

Keywords_poland_eu_tax_file_number

- nip#
- nip
- numer identyfikacji podatkowej
- numeridentyfikacjipodatkowej#
- tax id
- tax identification no

- tax identification number
- tax no#
- tax no
- tax number
- tax registration number
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id
- tin no
- tin#
- vat id#
- vat id
- vat no
- vat number
- vatid#
- vatid
- vatno#

Portugal citizen card number

Format

eight digits

Pattern

eight digits

Checksum

No

Keywords

Keyword_portugal_citizen_card

- bilhete de identidade
- cartão de cidadão
- citizen card
- document number
- documento de identificação
- id number
- identification no
- identification number
- identity card no
- identity card number
- national id card
- nic
- número bi de portugal
- número de identificação civil
- número de identificação fiscal
- número do documento
- portugal bi number

Portugal driver's license number

Format

two patterns - two letters followed by 5-8 digits with special characters

Pattern

Pattern 1: Two letters followed by 5/6 with special characters:

- Two letters (not case-sensitive)
- A hyphen
- Five or Six digits
- A space
- One digit

Pattern 2: One letter followed by 6/8 digits with special characters:

- One letter (not case-sensitive)
- A hyphen
- Six or eight digits

- A space
- One digit

Checksum

No

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence

- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#

- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen

- driving licenses
- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_portugal_eu_driver's_license_number

- carteira de motorista
- carteira motorista
- carteira de habilitação
- carteira habilitação
- número de licença
- número licença
- permissão de condução
- permissão condução
- Licença condução Portugal
- carta de condução

Portugal passport number

Format

one letter followed by six digits with no spaces or delimiters

Pattern

one letter followed by six digits:

- one letter (not case-sensitive)
- six digits

Checksum

No

Keywords

Keywords_eu_passport_number

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers

Keywords_portugal_eu_passport_number

- número do passaporte
- portuguese passport
- portuguese passeport
- portuguese passaporte
- passaporte nº
- passeport nº
- números de passaporte
- portuguese passports
- número passaporte
- números passaporte

Keywords_eu_passport_date

- date of issue
- date of expiry

Portugal tax identification number

Format

nine digits with optional spaces

Pattern

- three digits

- an optional space
- three digits
- an optional space
- three digits

Checksum

Yes

Keywords

Keywords_portugal_eu_tax_file_number

- cpf#
- cpf
- nif#
- nif
- número de identificação fisca
- numero fiscal
- tax id
- tax identification no
- tax identification number
- tax no#
- tax no
- tax number
- tax registration number
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id
- tin no
- tin#

Romania driver's license number

Format

one character followed by eight digits

Pattern

one character followed by eight digits:

- one letter (not case-sensitive) or digit
- eight digits

Checksum

No

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses

- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#

- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses

- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses
- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_romania_eu_driver's_license_number

- permis de conducere
- permisului de conducere
- permisului conducere
- permisele de conducere
- permisele conducere
- permis conducere

Romania personal numeric code (CNP)

Format

13 digits without spaces and delimiters

Pattern

- one digit from 1-9
- six digits representing date of birth (YYMMDD)
- two digits, which can be 01-52 or 99
- four digits

Checksum

Yes

Keywords

Keywords_romania_eu_national_id_card

- cnp#
- cnp
- cod identificare personal
- cod numeric personal
- cod unic identificare
- codnumericpersonal#
- codul fiscal nr.
- identificarea fiscală nr#
- id-ul taxei
- insurance number
- insurancenumber#
- national id#
- national id
- national identification number
- număr identificare personal
- număr identitate
- număr personal unic
- număridentitate#
- număridentitate
- numărpersonalunic#
- numărpersonalunic
- număru de identificare fiscală
- numărul de identificare fiscală
- personal numeric code
- pin#
- pin
- tax file no
- tax file number
- tax id
- tax identification no
- tax identification number
- tax no#
- tax no
- tax number

- tax registration number
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id
- tin no
- tin#
- unique identification number
- unique identity number
- uniqueidentityno#
- uniqueidentityno

Romania passport number

Format

eight or nine digits without spaces and delimiters

Pattern

eight or nine digits

Checksum

No

Keywords

Keywords_eu_passport_number

- passport#
- passport #
- passportid
- passports
- passportno
- passport no

- passportnumber
- passport number
- passportnumbers
- passport numbers

Keywords_romania_eu_passport_number

- numărul pașaportului
- numarul pasaportului
- numerele pașaportului
- Pașaport nr

Keywords_eu_passport_date

- date of issue
- date of expiry

Russia passport number domestic

Format

10-digit number

Pattern

10-digit number:

- two digits
- an optional space or hyphen
- two digits
- an optional space
- six digits

Checksum

No

Keywords

Keyword_russia_passport_number_domestic

- passport number
- passport no
- passport #
- passport id
- passportno#
- passportnumber#
- паспорт нет
- паспорт id
- российской паспорт
- русский номер паспорта
- паспорт#
- паспортid#
- номер паспорта
- номерпаспорта#

Russia passport number international

Format

nine-digit number

Pattern

nine-digit number:

- two digits
- an optional space or hyphen
- seven digits

Checksum

No

Keywords

Keywords_russia_passport_number_international

- passport number
- passport no
- passport #
- passport id
- passportno#
- passportnumber#
- паспорт нет
- паспорт id
- российской паспорт
- русский номер паспорта
- паспорт#
- паспортid#
- номер паспорта
- номерпаспорта#

Saudi Arabia National ID

Format

10 digits

Pattern

10 consecutive digits

Checksum

No

Keywords

Keyword_saudi_arabia_national_id

- Identification Card
- I card number
- ID number
- الوطنية الهوية بطاقة رقم

Singapore national registration identity card (NRIC) number

Format

nine letters and digits

Pattern

- nine letters and digits:
- the letter "F", "G", "S", or "T" (not case-sensitive)
- seven digits
- an alphabetic check digit

Checksum

Yes

Keywords

Keyword_singapore_nric

- National Registration Identity Card
- Identity Card Number
- NRIC
- IC
- Foreign Identification Number
- FIN
- 身份证
- 身份證

Slovakia driver's license number

Format

one character followed by seven digits

Pattern

one character followed by seven digits

- one letter (not case-sensitive) or digit
- seven digits

Checksum

No

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license

- driver'licenses
- driver'licence
- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#

- drivers lics#
- drivers license#
- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen

- driving lic
- driving licen
- driving licenses
- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_slovakia_eu_driver's_license_number

- vodičský preukaz
- vodičské preukazy
- vodičského preukazu
- vodičských preukazov

Slovakia personal number

Format

nine or ten digits containing optional backslash

Pattern

- six digits representing date of birth
- optional slash (/)
- three digits
- one optional check digit

Checksum

Yes

Keywords

Keywords_slovakia_eu_national_id_card

- azonosító szám

- birth number
- číslo národnej identifikačnej karty
- číslo občianského preukazu
- daňové číslo
- id number
- identification no
- identification number
- identifikačná karta č
- identifikačné číslo
- identity card no
- identity card number
- národná identifikačná značka č
- national number
- nationalnumber#
- nemzeti személyazonosító igazolvány
- personalidnumber#
- rč
- rodne cislo
- rodné číslo
- social security number
- ssn#
- ssn
- személyi igazolvány szám
- személyi igazolvány száma
- személyigazolvány szám
- tax file no
- tax file number
- tax id
- tax identification no
- tax identification number
- tax no#
- tax no
- tax number
- tax registration number
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id

- tin no
- tin#

Slovakia passport number

Format

one digit or letter followed by seven digits with no spaces or delimiters

Pattern

one digit or letter (not case-sensitive) followed by seven digits

Checksum

No

Keywords

Keywords_eu_passport_number

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers

Keywords_slovakia_eu_passport_number

- číslo pasu
- čísla pasov
- pas č.
- Passeport n°
- n° Passeport

Keywords_eu_passport_date

- date of issue
- date of expiry

Slovenia driver's license number

Format

nine digits without spaces and delimiters

Pattern

nine digits

Checksum

No

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences

- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#

- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence

- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses
- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_slovenia_eu_driver's_license_number

- vozniško dovoljenje
- vozniška številka licence
- vozniških dovoljenj
- številka vozniškega dovoljenja
- številke vozniških dovoljenj

Slovenia Unique Master Citizen Number

Format

13 digits without spaces or delimiters

Pattern

13 digits in the specified pattern:

- seven digits that correspond to the birth date (DDMMLLL) where "LLL" corresponds to the last three digits of the birth year
- two digits that correspond to the area of birth "50"
- three digits that correspond to a combination of gender and serial number for persons born on the same day. 000-499 for male and 500-999 for female.
- one check digit

Checksum

Yes

Keywords

Keywords_slovenia_eu_national_id_card

- edinstvena številka glavnega državljana
- emšo
- enotna maticna številka obcana
- id card
- identification number
- identifikacijska številka
- identity card
- nacionalna id
- nacionalni potni list
- national id
- osebna izkaznica
- osebni koda
- osebni ne
- osebni številka
- personal code
- personal number
- personal numeric code
- številka državljana
- unique citizen number
- unique id number
- unique identity number
- unique master citizen number
- unique registration number
- uniqueidentityno #
- uniqueidentityno#

Slovenia passport number

Format

two letters followed by seven digits with no spaces or delimiters

Pattern

two letters followed by seven digits:

- the letter "P"
- one uppercase letter
- seven digits

Checksum

No

Keywords

Keywords_eu_passport_number

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers

Keywords_slovenia_eu_passport_number

- številka potnega lista
- potek veljavnosti
- potni list#
- datum rojstva

- potni list
- številke potnih listov

Keywords_eu_passport_date

- date of issue
- date of expiry

Slovenia tax identification number

Format

eight digits with no spaces or delimiters

Pattern

- one digit from 1-9
- six digits
- one check digit

Checksum

Yes

Keywords

Keywords_slovenia_eu_tax_file_number

- davčna številka
- identifikacijska številka davka
- številka davčne datoteke
- tax file no
- tax file number
- tax id
- tax identification no
- tax identification number
- tax no#
- tax no
- tax number

- tax registration number
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id
- tin no
- tin#

South Africa identification number

Format

13 digits that may contain spaces

Pattern

13 digits:

- six digits in the format YYMMDD, which are the date of birth
- four digits
- a single-digit citizenship indicator
- the digit "8" or "9"
- one digit, which is a checksum digit

Checksum

Yes

Keywords

Keyword_south_africa_identification_number

- Identity card
- ID
- Identification

South Korea resident registration number

Format

13 digits containing a hyphen

Pattern

13 digits:

- six digits in the format YYMMDD, which are the date of birth
- a hyphen
- one digit determined by the century and gender
- four-digit region-of-birth code
- one digit used to differentiate people for whom the preceding numbers are identical
- a check digit.

Checksum

Yes

Keywords

Keyword_south_korea_resident_number

- National ID card
- Citizen's Registration Number
- Jumin deungnok beonho
- RRN
- 주민등록번호

Spain driver's license number

Format

eight digits followed by one character

Pattern

eight digits followed by one character:

- eight digits
- one digit or letter (not case-sensitive)

Checksum

Yes

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics

- driver'license
- driver'licenses
- driver'licence
- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#

- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen

- driver's licen
- driving lic
- driving licen
- driving licenses
- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_spain_eu_driver's_license_number

- permiso de conducción
- permiso conducción
- licencia de conducir
- licencia conducir
- permiso conducir
- permiso de conducir
- permisos de conducir
- permisos conducir
- carnet conducir
- carnet de conducir
- licencia de manejo
- licencia manejo

Spain DNI

Format

eight digits followed by one character

Pattern

seven digits followed by one character

- eight digits
- An optional space or hyphen
- one check letter (not case-sensitive)

Checksum

Yes

Keywords

Keywords_spain_eu_national_id_card

- carné de identidad
- dni#
- dni
- dninúmero#
- documento nacional de identidad
- identidad único
- identidadúnico#
- insurance number
- national identification number
- national identity
- nationalid#
- nationalidno#
- nie#
- nie
- nienúmero#
- número de identificación
- número nacional identidad
- personal identification number
- personal identity no
- unique identity number
- uniqueid#

Spain passport number

Format

an eight- or nine-character combination of letters and numbers with no spaces or delimiters

Pattern

an eight- or nine-character combination of letters and numbers:

- two digits or letters
- one digit or letter (optional)
- six digits

Checksum

Not applicable

Keywords

Keywords_eu_passport_number

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers

Keywords_spain_eu_passport_number

- libreta pasaporte
- número pasaporte
- españa pasaporte
- números de pasaporte
- número de pasaporte
- números pasaporte
- pasaporte no
- Passeport n°
- n° Passeport
- pasaporte no.
- pasaporte n°
- spain passport

Keywords_eu_passport_date

- date of issue
- date of expiry

Spain social security number (SSN)

Format

11-12 digits

Pattern

11-12 digits:

- two digits
- a forward slash (optional)
- seven to eight digits
- a forward slash (optional)
- two digits

Checksum

Yes

Keywords

Keywords_spain_eu_passport_number

- ssn
- ssn#
- socialsecurityno
- social security no
- social security number
- número de la seguridad social

Spain tax identification number

Format

seven or eight digits and one or two letters in the specified pattern

Pattern

Spanish Natural Persons with a Spain National Identity Card:

- eight digits
- one uppercase letter (case-sensitive)

Non-resident Spaniards without a Spain National Identity Card

- one uppercase letter "L" (case-sensitive)
- seven digits
- one uppercase letter (case-sensitive)

Resident Spaniards under the age of 14 years without a Spain National Identity Card:

- one uppercase letter "K" (case-sensitive)
- seven digits
- one uppercase letter (case-sensitive)

Foreigners with a Foreigner's Identification Number

- one uppercase letter that is "X", "Y", or "Z" (case-sensitive)
- seven digits
- one uppercase letter (case-sensitive)

Foreigners without a Foreigner's Identification Number

- one uppercase letter that is "M" (case-sensitive)
- seven digits
- one uppercase letter (case-sensitive)

Checksum

Yes

Keywords

Keywords_spain_eu_tax_file_number

- cif

- cifid#
- cifnúmero#
- número de contribuyente
- número de identificación fiscal
- número de impuesto corporativo
- spanishcifid#
- spanishcifid
- spanishcifno#
- spanishcifno
- tax file no
- tax file number
- tax id
- tax identification no
- tax identification number
- tax no#
- tax no
- tax number
- tax registration number
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id
- tin no
- tin#

Sweden driver's license number

Format

10 digits containing a hyphen

Pattern

10 digits containing a hyphen:

- six digits

- a hyphen
- four digits

Checksum

No

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence

- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#

- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen

- driving licenses
- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_sweden_eu_driver's_license_number

- ajokortti
- permis de conducere
- ajokortin numero
- kuljettajat lic.
- drivere lic.
- körkort
- numărul permisului de conducere
- שאָפער דערלויבעניש נומער
- förare lic.
- דריווערס דערלויבעניש
- körkortsnummer

Sweden national ID

Format

10 or 12 digits and an optional delimiter

Pattern

10 or 12 digits and an optional delimiter:

- two digits (optional)
- Six digits in date format YYMMDD
- delimiter of "-" or "+" (optional)
- four digits

Checksum

Yes

Keywords

Keywords_swedish_national_identifier

- id no
- id number
- id#
- identification no
- identification number
- identifikationsnumret#
- identifikationsnumret
- identitetshandling
- identity document
- identity no
- identity number
- id-nummer
- personal id
- personnummer#
- personnummer
- skatteidentifikationsnummer

Sweden passport number

Format

eight digits

Pattern

eight consecutive digits

Checksum

No

Keywords

Keywords_eu_passport_number

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers

Keyword_sweden_passport

- alien registration card
- g3 processing fees
- multiple entry
- Numéro de passeport
- passeport n °
- passeport non
- passeport #
- passeport#
- passeportnon
- passeportn °
- passnummer
- pass nr
- schengen visa
- schengen visas
- single entry
- sverige pass
- visa requirements
- visa processing
- visa type

Keywords_eu_passport_date

- date of issue
- date of expiry

Sweden tax identification number

Format

10 digits and a symbol in the specified pattern

Pattern

10 digits and a symbol:

- six digits that correspond to the birth date (YYMMDD)
- a plus sign or minus sign
- three digits that make the identification number unique where:
  - for numbers issued before 1990, the seventh and eighth digit identify the county of birth or foreign-born people
  - the digit in the ninth position indicates gender by either odd for male or even for female
- one check digit

Checksum

Yes

Keywords

Keywords_sweden_eu_tax_file_number

- personal id number
- personnummer
- skatt id nummer
- skatt identifikation
- skattebetalarens identifikationsnummer
- sverige tin
- tax file
- tax id
- tax identification no
- tax identification number
- tax no#
- tax no
- tax number

- tax registration number
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id
- tin no
- tin#

SWIFT code

Format

four letters followed by 5-31 letters or digits

Pattern

four letters followed by 5-31 letters or digits:

- four-letter bank code (not case-sensitive)
- an optional space
- 4-28 letters or digits (the Basic Bank Account Number (BBAN))
- an optional space
- one to three letters or digits (remainder of the BBAN)

Checksum

No

Keywords

Keyword_swift

- international organization for standardization 9362
- iso 9362
- iso9362
- swift#

- swiftcode
- swiftnumber
- swiftroutingnumber
- swift code
- swift number #
- swift routing number
- bic number
- bic code
- bic #
- bic#
- bank identifier code
- Organisation internationale de normalisation 9362
- rapide #
- code SWIFT
- le numéro de swift
- swift numéro d'acheminement
- le numéro BIC
- # BIC
- code identificateur de banque
- SWIFTコード
- SWIFT番号
- BIC番号
- BICコード
- SWIFT コード
- SWIFT 番号
- BIC 番号
- BIC コード
- 金融機関識別コード
- 金融機関コード
- 銀行コード

Switzerland SSN AHV number

Format

13-digit number

Pattern

13-digit number:

- three digits - 756
- an optional dot
- four digits
- an optional dot
- four digits
- an optional dot
- two digits

Checksum

Yes

Keywords

Keyword_swiss_ssn_AHV_number

- ahv
- ssn
- pid
- insurance number
- personalidno#
- social security number
- personal id number
- personal identification no.
- insuranceno#
- uniqueidno#
- unique identification no.
- avs number
- personal identity no versicherungsnummer
- identifikationsnummer
- einzigartige identität nicht
- sozialversicherungsnummer
- identification personnelle id
- numéro de sécurité sociale

Taiwanese identification number

Format

one letter (in English) followed by nine digits

Pattern

one letter (in English) followed by nine digits:

- one letter (in English, not case-sensitive)
- the digit "1" or "2"
- eight digits

Checksum

Yes

Keywords

Keyword_taiwan_national_id

- 身份證字號
- 身份證
- 身份證號碼
- 身份證號
- 身分證字號
- 身分證
- 身分證號碼
- 身份證號
- 身分證統一編號
- 國民身分證統一編號
- 簽名
- 蓋章
- 簽名或蓋章
- 簽章

Taiwan passport number

Format

- biometric passport number: nine digits
- non-biometric passport number: nine digits

Pattern

biometric passport number:

- the character "3"
- eight digits

non-biometric passport number:

- nine digits

Checksum

No

Keywords

Keyword_taiwan_passport

- ROC passport number
- Passport number
- Passport no
- Passport Num
- Passport #
- 护照
- 中華民國護照
- Zhōnghuá Mínguó hùzhào

Taiwan-resident certificate (ARC/TARC) number

Format

10 letters and digits

Pattern

10 letters and digits:

- two letters (not case-sensitive)
- eight digits

Checksum

No

Keywords

Keyword_taiwan_resident_certificate

- Resident Certificate
- Resident Cert
- Resident Cert.
- Identification card
- Alien Resident Certificate
- ARC
- Taiwan Area Resident Certificate
- TARC
- 居留證
- 外僑居留證
- 台灣地區居留證

U.K. driver's license number

Format

Combination of 18 letters and digits in the specified format

Pattern

18 letters and digits:

- Five letters (not case-sensitive) or the digit "9" in place of a letter.
- One digit.
- Five digits in the date format MMDDY for date of birth. The seventh character is incremented by 50 if driver is female; for example, 51 to 62 instead of 01 to 12.
- Two letters (not case-sensitive) or the digit "9" in place of a letter.
- Five digits.

Checksum

Yes

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences
- driver' lic
- driver' lics

- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#
- drivers licence#
- drivers licences#

- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses
- driving licence
- driving licences

- driving permit
- dl no
- dlno
- dl number

U.K. electoral roll number

Format

two letters followed by 1-4 digits

Pattern

two letters (not case-sensitive) followed by 1-4 numbers

Checksum

No

Keywords

Keyword_uk_electoral

- council nomination
- nomination form
- electoral register
- electoral roll

U.K. national health service number

Format

10-17 digits separated by spaces

Pattern

10-17 digits:

- either 3 or 10 digits
- a space
- three digits
- a space
- four digits

Checksum

Yes

Keywords

Keyword_uk_nhs_number

- national health service
- nhs
- health services authority
- health authority

Keyword_uk_nhs_number1

- patient id
- patient identification
- patient no
- patient number

Keyword_uk_nhs_number_dob

- GP
- DOB
- D.O.B
- Date of Birth
- Birth Date

U.K. national insurance number (NINO)

This sensitive information type entity is included in the EU National Identification Number sensitive information type. It's also available as a stand-alone sensitive information type entity.

Format

seven characters or nine characters separated by spaces or dashes

Pattern

two possible patterns:

- two letters (valid NINOs use only certain characters in this prefix, which this pattern validates; not case-sensitive)
- six digits
- either 'A', 'B', 'C', or 'D' (like the prefix, only certain characters are allowed in the suffix; not case-sensitive)

OR

- two letters
- a space or dash
- two digits
- a space or dash
- two digits
- a space or dash
- two digits
- a space or dash
- either 'A', 'B', 'C', or 'D'

Checksum

No

Keywords

Keyword_uk_nino

- national insurance number
- national insurance contributions
- protection act
- insurance
- social security number
- insurance application
- medical application

- social insurance
- medical attention
- social security
- great britain
- NI Number
- NI No.
- NI #
- NI#
- insurance#
- insurancenumber
- nationalinsurance#
- nationalinsurancenumber

U.K. Unique Taxpayer Reference Number

Format

10 digits without spaces and delimiters

Pattern

10 digits

Checksum

No

Keywords

Keywords_uk_eu_tax_file_number

- tax number
- tax file
- tax id
- tax identification no
- tax identification number
- tax no#
- tax no
- tax registration number

- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id
- tin no
- tin#

U.S. bank account number

Format

6-17 digits

Pattern

6-17 consecutive digits

Checksum

No

Keywords

Keyword_usa_Bank_Account

- Checking Account Number
- Checking Account
- Checking Account #
- Checking Acct Number
- Checking Acct #
- Checking Acct No.
- Checking Account No.
- Bank Account Number
- Bank Account #
- Bank Acct Number
- Bank Acct #

- Bank Acct No.
- Bank Account No.
- Savings Account Number
- Savings Account.
- Savings Account #
- Savings Acct Number
- Savings Acct #
- Savings Acct No.
- Savings Account No.
- Debit Account Number
- Debit Account
- Debit Account #
- Debit Acct Number
- Debit Acct #
- Debit Acct No.
- Debit Account No.

U.S. driver's license number

Format

Depends on the state

Pattern

depends on the state - for example, New York:

- nine digits formatted like ddd ddd ddd will match.
- nine digits like ddddddddd won't match.

Checksum

No

Keywords

Keyword_us_drivers_license_abbreviations

- DL

- DLS
- CDL
- CDLS
- ID
- IDs
- DL#
- DLS#
- CDL#
- CDLS#
- ID#
- IDs#
- ID number
- ID numbers
- LIC
- LIC#

Keyword_us_drivers_license

- DriverLic
- DriverLics
- DriverLicense
- DriverLicenses
- Driver Lic
- Driver Lics
- Driver License
- Driver Licenses
- DriversLic
- DriversLics
- DriversLicense
- DriversLicenses
- Drivers Lic
- Drivers Lics
- Drivers License
- Drivers Licenses
- Driver'Lic
- Driver'Lics
- Driver'License
- Driver'Licenses
- Driver' Lic
- Driver' Lics
- Driver' License

- Driver' Licenses
- Driver'sLic
- Driver'sLics
- Driver'sLicense
- Driver'sLicenses
- Driver's Lic
- Driver's Lics
- Driver's License
- Driver's Licenses
- identification number
- identification numbers
- identification #
- id card
- id cards
- identification card
- identification cards
- DriverLic#
- DriverLics#
- DriverLicense#
- DriverLicenses#
- Driver Lic#
- Driver Lics#
- Driver License#
- Driver Licenses#
- DriversLic#
- DriversLics#
- DriversLicense#
- DriversLicenses#
- Drivers Lic#
- Drivers Lics#
- Drivers License#
- Drivers Licenses#
- Driver'Lic#
- Driver'Lics#
- Driver'License#
- Driver'Licenses#
- Driver' Lic#
- Driver' Lics#
- Driver' License#
- Driver' Licenses#
- Driver'sLic#

- Driver'sLics#
- Driver'sLicense#
- Driver'sLicenses#
- Driver's Lic#
- Driver's Lics#
- Driver's License#
- Driver's Licenses#
- id card#
- id cards#
- identification card#
- identification cards#

Keyword_[state_name]_drivers_license_name

- state abbreviation (for example, "NY")
- state name (for example, "New York")

U.S. individual taxpayer identification number (ITIN)

Format

nine digits that start with a "9" and contain a "7" or "8" as the fourth digit, optionally formatted with spaces or dashes

Pattern

formatted:

- the digit "9"
- two digits
- a space or dash
- a "7" or "8"
- a digit
- a space, or dash
- four digits

unformatted:

- the digit "9"
- two digits

- a "7" or "8"
- five digits

Checksum

No

Keywords

Keyword_itin

- taxpayer
- tax id
- tax identification
- itin
- i.t.i.n.
- ssn
- tin
- social security
- tax payer
- itins
- taxid
- individual taxpayer

U.S. phone number

Pattern

- 10 digit number, for example, +1 nxx-nxx-xxxx
- Optional area code: +1
- n can be any digit between 2-9
- x can be any digit between 0-9
- Optional paranthesis around the area code
- Optional space or - between area code, exchange code, and the last four digits
- Optional four digit extension

Checksum

Not applicable

Keywords

Keywords_us_phone_number

- cell
- cellphone
- contact
- landline
- mobile
- mob
- mob#
- ph#
- phone
- telephone
- tel#

U.S. social security number (SSN)

Format

nine digits, which may be in a formatted or unformatted pattern

> ⓘ **Note**
>
> If issued before mid-2011, an SSN has strong formatting where certain parts of the number must fall within certain ranges to be valid (but there's no checksum).

Pattern

four functions look for SSNs in four different patterns:

- Func_ssn finds SSNs with pre-2011 strong formatting that are formatted with dashes or spaces (ddd-dd-dddd OR ddd dd dddd)
- Func_unformatted_ssn finds SSNs with pre-2011 strong formatting that are unformatted as nine consecutive digits (ddddddddd)
- Func_randomized_formatted_ssn finds post-2011 SSNs that are formatted with dashes or spaces (ddd-dd-dddd OR ddd dd dddd)
- Func_randomized_unformatted_ssn finds post-2011 SSNs that are unformatted as nine consecutive digits (ddddddddd)

Checksum

No

Keywords

Keyword_ssn

- SSA Number
- social security number
- social security #
- social security#
- social security no
- Social Security#
- Soc Sec
- SSN
- SSNS
- SSN#
- SS#
- SSID

U.S. states

Format

Includes all 50 U.S. state names and the two digit short codes.

Checksum

Not applicable

Keywords

Keywords_us_states

- State

U.S. zipcode

Format

Five digit U.S. Zip code and an optional four digit code separated by a hyphen (-).

Checksum

Not applicable

Keywords

Keywords_us_zip_code

- zip
- zipcode
- postal
- postalcode

U.S. / U.K. passport number

Format

nine digits

Pattern

nine consecutive digits

Checksum

No

Keywords

Keywords_eu_passport_number

- passport#

- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers

Keywords_uk_eu_passport_number

- british passport
- uk passport

Ukraine passport domestic

Format

nine digits

Pattern

nine digits

Checksum

No

Keywords

Keyword_ukraine_passport_domestic

- ukraine passport
- passport number
- passport no
- паспорт України
- номер паспорта
- персональний

Ukraine passport international

Format

eight-character alphanumeric pattern

Pattern

eight-character alphanumeric pattern:

- two letters or digits
- six digits

Checksum

No

Keywords

Keyword_ukraine_passport_international

- ukraine passport
- passport number
- passport no
- паспорт України
- номер паспорта

Feedback

Was this page helpful? Yes No

Data Map history (preview)

06/06/2025

> **Important**
>
> This feature is currently in preview. The Supplemental Terms of Use for Microsoft Azure Previews include additional legal terms that apply to Azure features that are in beta, in preview, or otherwise not yet released into general availability.

When a user updates an asset in Microsoft Purview, Data Map history logs that change and makes the information available in the Microsoft Purview portal. Information about who added or edited an asset, what edits were made, and when the edit happened, is stored for reference later. This information helps data auditors and business owners to maintain compliance and security in their Microsoft Purview accounts. It also allows data stewards and admins to debug issues and curation in their data assets.

Data Map history is available to users in both the free and enterprise versions of Microsoft Purview.

Currently, you can view your assets' history either:

- In the Audit solution in the Microsoft Purview portal
- Within individual solutions
- Using the REST API

Prerequisites

Before you can use Data Map history, make sure that the audit log is turned on in your organization. If it's not, you need to turn it on to access your Data Map history.

Which assets have history

Currently, these kinds of assets have Data Map history available:

- Classifications*
- Data assets*
- Glossary terms*
- Business assets*
- Domains
- Collections

- Role assignments
- Data sources
- Scan
- Scan rule set
- Credential
- Policy

* In-solution experience is currently only available for classifications, data assets, business assets, and glossary terms.

How to use Data Map history

Data Map history allows you to see the operational history of changes to your Data Map assets in Microsoft Purview.

You're able to see the history for select operations that include:

- Creation
- Updates
- Assignment

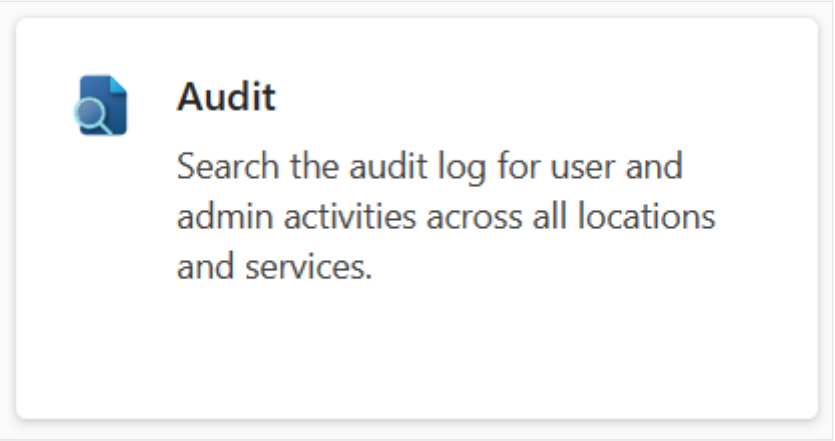
For a full list of available operations in Data Map History, see the audit log activities documentation.

> **Important**
>
> Data Map history entries are kept in Audit for 180 days, and are available for 30 days in the **in-solution experience**. To view audit history for longer, you can use **Microsoft Purview Audit (Premium)**.

Audit solution experience

To view the activities within the Audit solution, follow these instructions:

1. Navigate to the Audit solution within Microsoft Purview

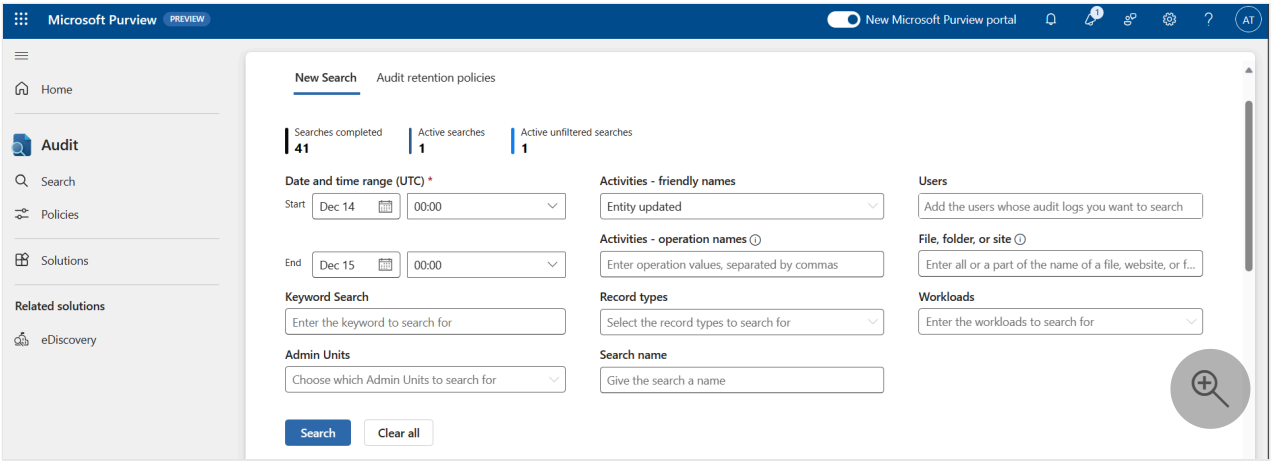


2. Select **Search** to search for audit records

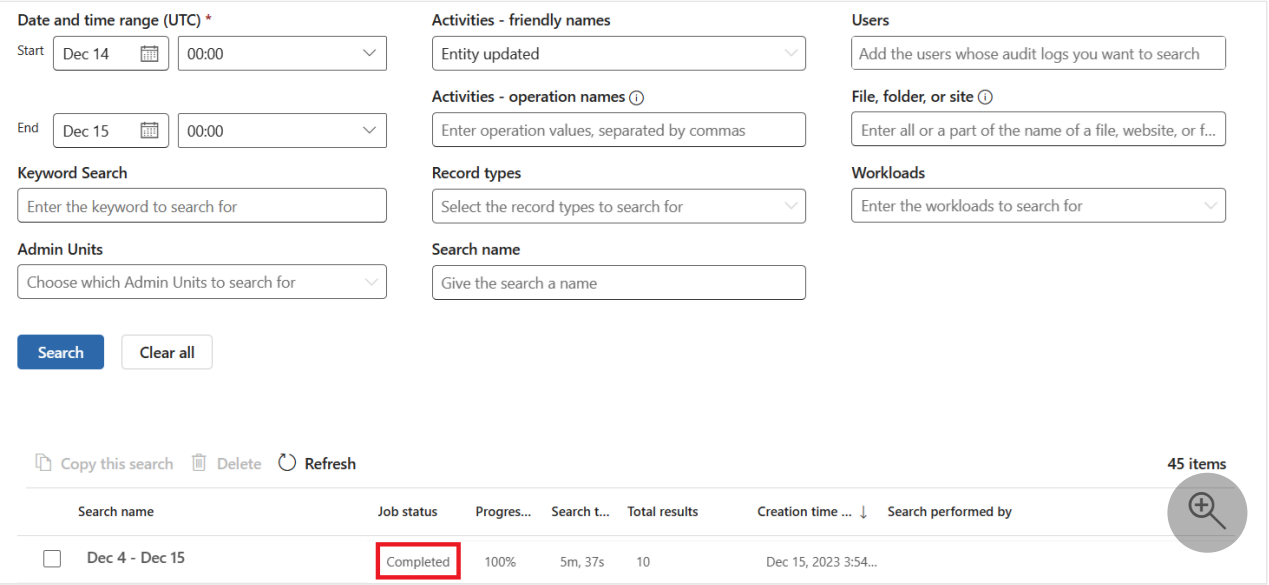
3. Select time range for activity.

4. Narrow the scope of your search using **Activities - friendly names**, **Activities - operation names**, or **Record types**.

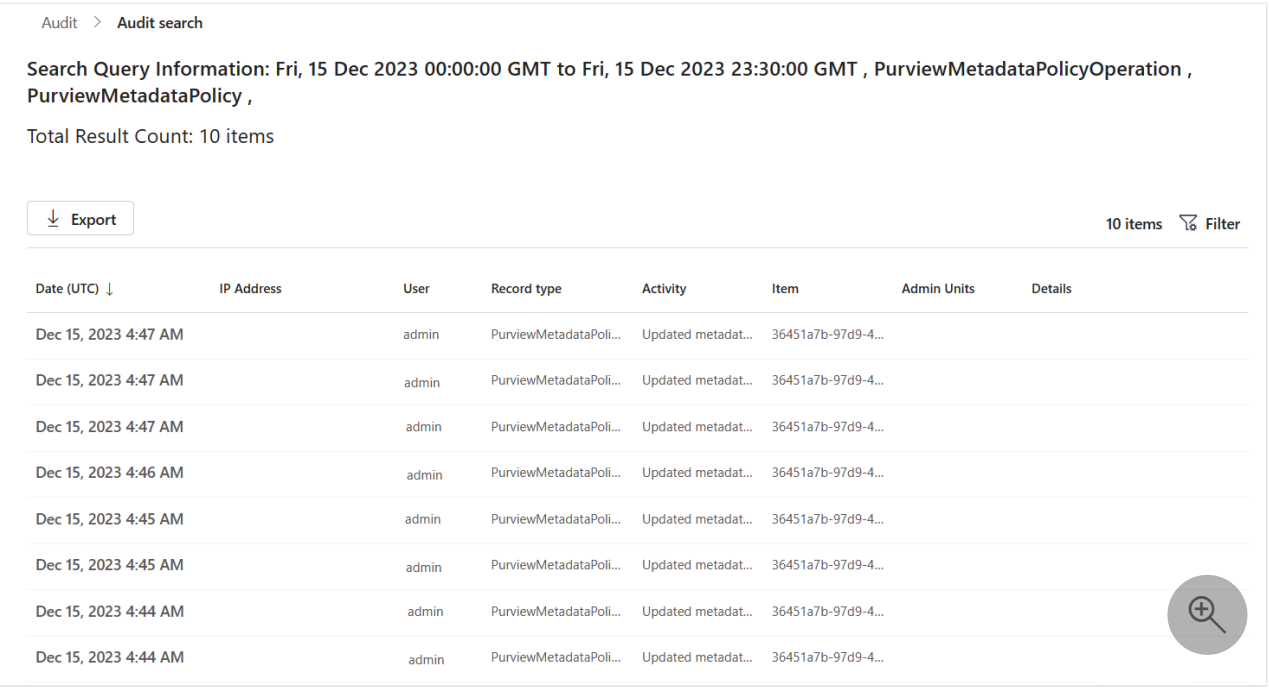
> **Tip**
>
> For the list of available activities, see the [audit log activities page]



5. Select **Search** and wait for job state to change to **Complete**.



6. Select your search from the list, and scroll through or filter the history table to find the operations you're looking for.



In-solution history experience

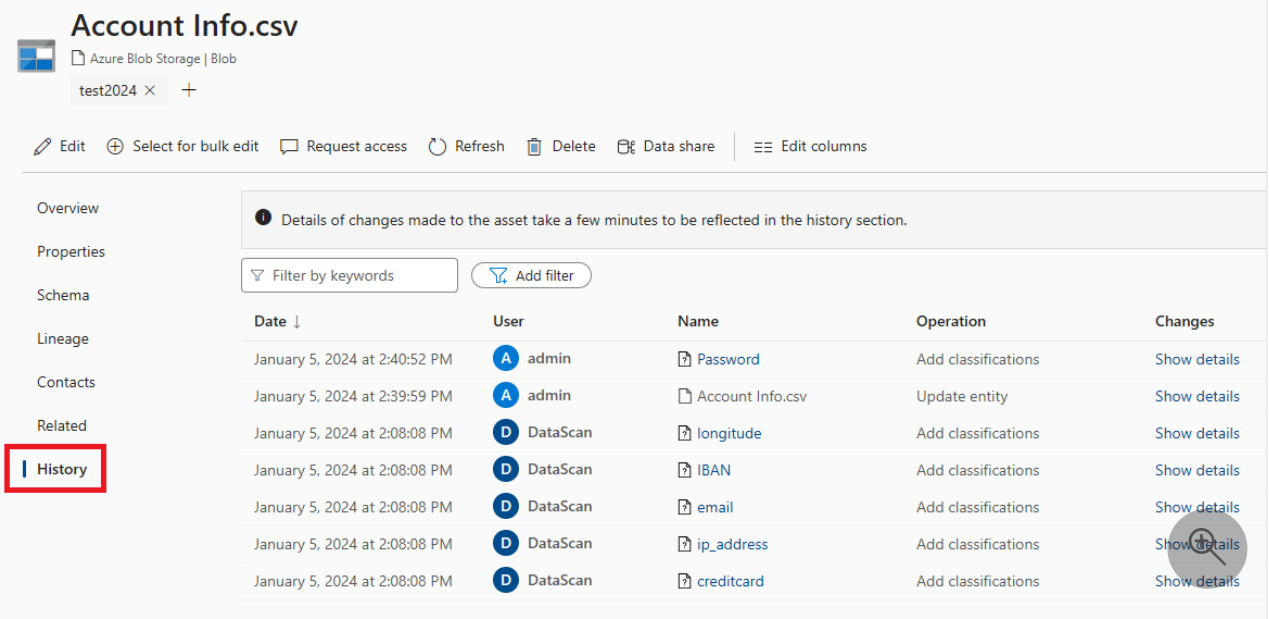
> **Note**
>
> In-solution experience is currently only available for classifications, data assets, business assets, and glossary terms.

To view the activities, follow these instructions:

1. Search for a data asset, glossary term, or classification in Microsoft Purview Unified Catalog.

2. Select the **History** tab.

3. Scroll through or filter the history table to find the operations you're looking for.



Query Audit log using the REST API

You also can use the Microsoft Purview REST API to access your Data Map history.

1. To authenticate with the REST API, follow the instructions on the Authenticate for APIs page.
2. Use the authentication token to POST to this request URL: `https://api.purview-service.microsoft.com/datamap/api/audit/query?api-version=2023-10-01-preview`

Here are a few example API calls you can use to get started:

- Retrieve an asset's history in a time range
- Retrieve the history of a specific operation
- Filter by user who edited an asset
- Filter by keywords
- Example with pagination

Retrieve an asset's history in a time range

Get all audit history of an asset within a specific time range.

JSON

```
{
  "category": "Asset",
  "guid": "f432d351-5442-4724-945b-516d5d501fc9",
```

```
  "startTime": "2024-01-01T00:00:00.000Z",
  "endTime": "2024-01-06T00:00:00.000Z",
  "sortBy": "CreationTime",
  "sortOrder": "Descending"
}
```

Retrieve the history of a specific operation

Get the history of specific operation type. For example, when classifications are applied to the asset.

> **Tip**
>
> For the list of available operations, see the audit log activities page.

JSON

```
{
  "category": "Asset",
  "guid": "f432d351-5442-4724-945b-516d5d501fc9",
  "operationType": "ClassificationAdded",
  "sortBy": "CreationTime",
  "sortOrder": "Descending"
}
```

Filter by user who edited an asset

JSON

```
{
  "category": "Asset",
  "guid": "f432d351-5442-4724-945b-516d5d501fc9",
  "userId": "contoso@microsoft.com",
  "sortBy": "CreationTime",
  "sortOrder": "Descending"
}
```

Filter by keywords

Filter by keywords such as a changed attribute name, attribute value, etc.

JSON

```
{
  "category": "Asset",
  "guid": "f432d351-5442-4724-945b-516d5d501fc9",
  "keywords": "userDescription",
  "sortBy": "CreationTime",
  "sortOrder": "Descending"
}
```

Example with pagination

After the first query, if the "lastPage" property in the response body is false, the use can use the "continuationToken" in the response body to fetch the next page.

JSON

```
{
  "continuationToken": "TY7BbsIwEAV%2fKDZr4oQ0KAcEVKo4klNvK…",
  "category": "Asset",
  "guid": "f432d351-5442-4724-945b-516d5d501fc9",
  "sortBy": "CreationTime",
  "sortOrder": "Descending",
  "pageSize": 10
}
```

Limitations

1. Currently, it takes a couple minutes for an operational item to appear in an assets **History** tab.
2. Currently, to see a new item in the **History** tab, you need to manually refresh the tab.

Next steps

- Metadata curation
- Use the new Microsoft Purview portal
- Roles and permissions

Choose the right integration runtime configuration for your scenario

The integration runtime (IR) is the compute infrastructure that Microsoft Purview uses to power data scan across different network environments. This article introduces the different types of integration runtime available in Microsoft Purview, and provides guidance on how to choose the right integration runtime configuration for your scenario.

Types of integration runtimes

Purview provides the following types of integration runtimes:

- **Azure integration runtime**: The Azure integration runtime is a fully managed and elastic compute that you can use to scan Azure or non-Azure data sources. The Azure IR supports connections to data stores and compute services with publicly accessible endpoints. It's the default integration runtime that you don't need to create anything to get started.

- **Managed Virtual Network integration runtime**: You can create a Managed Virtual Network integration runtime, which resides in a Purview Managed Virtual Network. It can use private endpoints to securely connect to and scan the supported data sources. Learn more at Managed Virtual Network and managed private endpoints.

- **Self-hosted integration runtime**: The self-hosted integration runtime can be used to scan data sources in an on-premises network or a virtual network. You can install it on an on-premises machine or a virtual machine inside your private network. Learn more at Create and manage self-hosted integration runtimes.

- **Kubernetes supported self-hosted integration runtime**: This integration runtime is hosted on a Kubernetes cluster and can be used to scan data sources in an on-premises network or a virtual network. Kubernetes support improves overall performance and allows the integration runtime to scale with the job. Learn more at Create and manage Kubernetes supported self-hosted integration runtimes.

- **AWS integration runtime**: The AWS integration runtime is a fully managed and elastic computed hosted by Microsoft Purview in AWS. It's applicable when scanning Amazon data sources like S3, RDS.

Choose the right integration runtime

Choose the right integration runtime for your needs. Consider your existing architecture and requirements for data integration. Also think about how to meet growing business needs and any future increase in workload.

The following considerations can help you make your decision:

1. **What data source types do you want to scan?**

   Check the supported data sources section to learn about the supported IR types for the data sources you want to scan.

2. **What's the network access control on your data source?**

   Different data sources have different network firewall settings to protect them from random access over the internet. These settings apply to on-premises, cloud, and SaaS data stores. The following table lists some common firewall options. Choose the supported IR type according to your scenario.

Expand table

Data source firewall	Azure IR	Managed Virtual Network IR	SHIR	Kubernetes supported SHIR
Allow public access	✓	✓	✓	✓
Allow Azure service or trusted service	✓	✓	✓	✓
Allow access from specific Azure virtual network		✓ (with managed private endpoint support)	✓	✓
Allow specific IP / IP range			✓	✓
Other on-premises or private network access			✓	✓

3. **What's the firewall setting of your Microsoft Purview?**

   Purview provides different network firewall options. Learn more from Configure Microsoft Purview firewall. Choose the supported IR type according to your scenario.

Expand table

Purview firewall	Azure IR	Managed Virtual Network IR	SHIR	Kubernetes supported SHIR
Enabled from all networks	✓	✓	✓	✓
Disabled from all networks		✓ (managed private endpoint required)	✓ (need to create private endpoint from your network)	✓ (need to create private endpoint from your network)

4. **What level of security do you require during data transmission?**

   The integration runtime location defines the location of its back-end compute and where the scan operations are performed. For data residency consideration:

   - When you use Azure IR, Purview automatically detects the data source's location and uses the IR in that region. If Purview can't detect the region, it uses the Purview account's region.
   - When you use Managed Virtual Network IR, it runs in the region you configure for the managed virtual network.
   - When you use SHIR, you can fully decide the location in your on-premises or Azure virtual machines.

     To protect against, for example, man-in-the-middle attacks during data transmission, use a private endpoint and private link to ensure data security.
   - You can create managed private endpoints to your data stores when using Managed Virtual Network IR. The Purview service maintains the private endpoints within the managed virtual network.
   - You can also create private endpoints in your virtual network and the SHIR can use them to access data stores.

5. **What level of maintenance are you able to provide?**

   Maintaining infrastructure, servers, and equipment is one of the important tasks of the IT department of an enterprise. It usually takes much time and effort.

   - When using Azure IR and Managed Virtual Network IR, you don't need to worry about maintenance such as updates, patches, and versions. The Purview service takes care of all the maintenance efforts.

- Because the SHIR is installed on your machines and the Kubernetes supported SHIR is on your Kubernetes clusters, you need to manage the maintenance.
  - The SHIR supports autoupdate to automatically get the latest version whenever there's' an update. Learn more from Self-hosted integration runtime autoupdate and expiration.
  - Currently, the Kubernetes supported self-hosted integration runtime only supports manual updates.

6. **Performance and scalability**

   Use the fully managed and autoscaled Azure IR, Managed Virtual Network IR, or the Kubernetes-supported self-hosted integration runtime whenever applicable. By using elasticity, they can provide you with better performance and scalability, especially when scanning large-scale data systems.

Hibernation of managed virtual network integration runtime

If the integration runtime is inactive (no scans on the integration runtime for more than 90 days), your Managed Virtual Network Integration Runtime automatically goes into hibernation. Its status shows as **Hibernated** when you select the integration runtime.

What this change means for you

1. When you run Test Connection on a hibernated integration runtime, the test connection fails. You see a message to try Test Connection after 15 minutes. By this time, your Managed Virtual Network returns to a normal state. After this, you can run your Test Connections and Scans normally.

2. When you run a scan directly by using **Run scan now** or **Edit Scan** options without running a Test Connection first from a hibernated Integration Runtime, or you run a scan through API, you see a message that this scan takes up to 15 extra minutes. This extra time is for the hibernated Integration Runtime to wake up and the scan process to begin. You see your scan status as **Queued_Waking Up IR** instead of **Queued** state that you see in case of normal scan. After the first scan, you can run all your following scans normally.

Supported data sources

The following table shows all the data sources that Purview scan supports, and the supported integration runtime types.

Expand table

Category	Supported data store	Azure IR/AWS IR	Managed Virtual Network IR	SHIR	Kubernetes SHIR
Azure	Multiple sources	✓			
	Azure Blob Storage	✓	✓ (including managed private endpoint)	✓	✓
	Azure Cosmos DB (API for NoSQL)	✓	✓ (including managed private endpoint)	✓	✓
	Azure Data Explorer	✓	✓ (v2 only)	✓	✓
	Azure Data Lake Storage Gen1	✓	✓ (v2 only)	✓	
	Azure Data Lake Storage Gen2	✓	✓ (including managed private endpoint)	✓	✓
	Azure Database for MySQL	✓	✓ (including managed private endpoint)	✓	✓
	Azure Database for PostgreSQL	✓	✓ (including managed private endpoint)	✓	✓
	Azure Databricks Hive Metastore			✓	✓
	Azure Databricks Unity Catalog	✓	✓ (v2 only, including managed private endpoint)		✓
	Azure Dedicated SQL pool (formerly SQL DW)	✓	✓ (including managed private endpoint)	✓	✓
	Azure Files	✓	✓ (including managed private endpoint)	✓	✓
	Azure SQL Database	✓	✓ (including managed private endpoint)	✓	✓
	Azure SQL Managed Instance	✓	✓ (including managed private endpoint)	✓	✓
	Azure Synapse Analytics (Workspace)	✓	✓ (including managed private endpoint)	✓	✓
Database	Amazon RDS	✓			✓
	Amazon Redshift	✓			✓

Category	Supported data store	Azure IR/AWS IR	Managed Virtual Network IR	SHIR	Kubernetes SHIR
	Cassandra	✓	✓ (v2 only)	✓	✓
	Db2			✓	✓
	Google BigQuery			✓	✓
	Hive Metastore Database			✓	✓
	MongoDB			✓	✓
	MySQL	✓	✓ (v2 only)	✓	✓
	Oracle			✓	
	PostgreSQL	✓	✓ (v2 only)	✓	✓
	SAP Business Warehouse			✓	✓
	SAP HANA			✓	✓
	Snowflake	✓	✓ (v2 only, including managed private endpoint)	✓	
	SQL Server			✓	✓
	SQL Server on Azure-Arc			✓	✓
	Teradata			✓	
File	Amazon S3	✓			
	HDFS			✓	✓
Services and apps	Dataverse	✓	✓ (v2 only)		✓
	Erwin			✓	
	Looker	✓	✓ (v2 only)	✓	✓
	Fabric	✓	✓ (v2 only)	✓	✓
	Power BI	✓	✓ (v2 only)	✓	✓
	Qlik Sense	✓	✓ (v2 only)		✓
	Salesforce	✓	✓ (v2 only)	✓	✓
	SAP ECC			✓	✓

Category	Supported data store	Azure IR/AWS IR	Managed Virtual Network IR	SHIR	Kubernetes SHIR
	SAP S/4HANA			✓	✓
	Tableau	✓	✓ (v2 only)		✓

Last updated on 05/15/2026

Create and manage a self-hosted integration runtime

09/30/2025

The integration runtime (IR) is the compute infrastructure that Microsoft Purview uses to power data scan across different network environments.

A self-hosted integration runtime (SHIR) can be used to scan data source in an on-premises network or a virtual network. The installation of a self-hosted integration runtime needs an on-premises machine or a virtual machine inside a private network.

This article covers both setup of a self-hosted integration runtime, and troubleshooting and management.

> **Important**
>
> Download the self-hosted integration runtime from: Microsoft Integration Runtime.

Expand table

To learn about	See
Setting up a new self-hosted integration runtime	Machine requirements
	Source-specific machine requirements are listed under prerequisites in each source article
	Set up guide
Networking	Networking requirements
	Proxy servers
	Private endpoints
	Troubleshoot proxy and firewall
	Troubleshoot connectivity
Management	General

> **Note**

The Microsoft Purview Integration Runtime can't be shared with an Azure Synapse Analytics or Azure Data Factory Integration Runtime on the same machine. It needs to be installed on a separated machine.

Prerequisites

- The supported versions of Windows are:
  - Windows 8.1
  - Windows 10
  - Windows 11
  - Windows Server 2012
  - Windows Server 2012 R2
  - Windows Server 2016
  - Windows Server 2019
  - Windows Server 2022
- Installation of the self-hosted integration runtime on a domain controller isn't supported.
- FIPS mode isn't currently supported for SHIR machines.

> **ⓘ Important**
>
> Scanning some data sources requires extra setup on the self-hosted integration runtime machine. For example, JDK, Visual C++ Redistributable, or specific driver. For your source, refer to each source article for prerequisite details. Any requirements will be listed in the **Prerequisites** section.

- To add and manage a SHIR in Microsoft Purview, you'll need data source administrator permissions in Microsoft Purview.
- Self-hosted integration runtime requires a 64-bit Operating System with .NET Framework 4.7.2 or above. See .NET Framework System Requirements for details.
- The recommended minimum configuration for the self-hosted integration runtime machine is a 2-GHz processor with 8 cores, 28 GB of RAM, and 80 GB of available hard drive space. Scanning some data sources may require higher machine specification based on your scenario. Also check the prerequisites in corresponding connector article.
- If the host machine hibernates, the self-hosted integration runtime doesn't respond to data requests. Configure an appropriate power plan on the computer before you install

the self-hosted integration runtime. If the machine is configured to hibernate, the self-hosted integration runtime installer prompts with a message.

- You must be an administrator on the machine to successfully install and configure the self-hosted integration runtime.
- Scan runs happen with a specific frequency per the schedule you've set up. Processor and RAM usage on the machine follows the same pattern with peak and idle times. Resource usage also depends heavily on the amount of data that is scanned. When multiple scan jobs are in progress, you see resource usage goes up during peak times.
- Enable Windows long path by default by following these steps.

> ⓘ **Important**
>
> If you use the Self-Hosted Integration runtime to scan Parquet files, you need to install the **64-bit JRE 8 (Java Runtime Environment) or OpenJDK** on your IR machine. Check our Java Runtime Environment section at the bottom of the page for an installation guide.

Considerations for using a self-hosted IR

- You can use a single self-hosted integration runtime for scanning multiple data sources.
- You can install only one instance of self-hosted integration runtime on any single machine. If you have two Microsoft Purview accounts that need to scan on-premises data sources, install the self-hosted IR on two machines, one for each Microsoft Purview account.
- The self-hosted integration runtime doesn't need to be on the same machine as the data source, unless specially called out as a prerequisite in the respective source article. Having the self-hosted integration runtime close to the data source reduces the time for the self-hosted integration runtime to connect to the data source.
- To optimize disk space, it's recommended to periodically clean up the numerical folders, along with numerical folders under the **Data Scan** and **MITIlib** folder that are older than seven days, in the Temp folder generated on your machine.

Setting up a self-hosted integration runtime

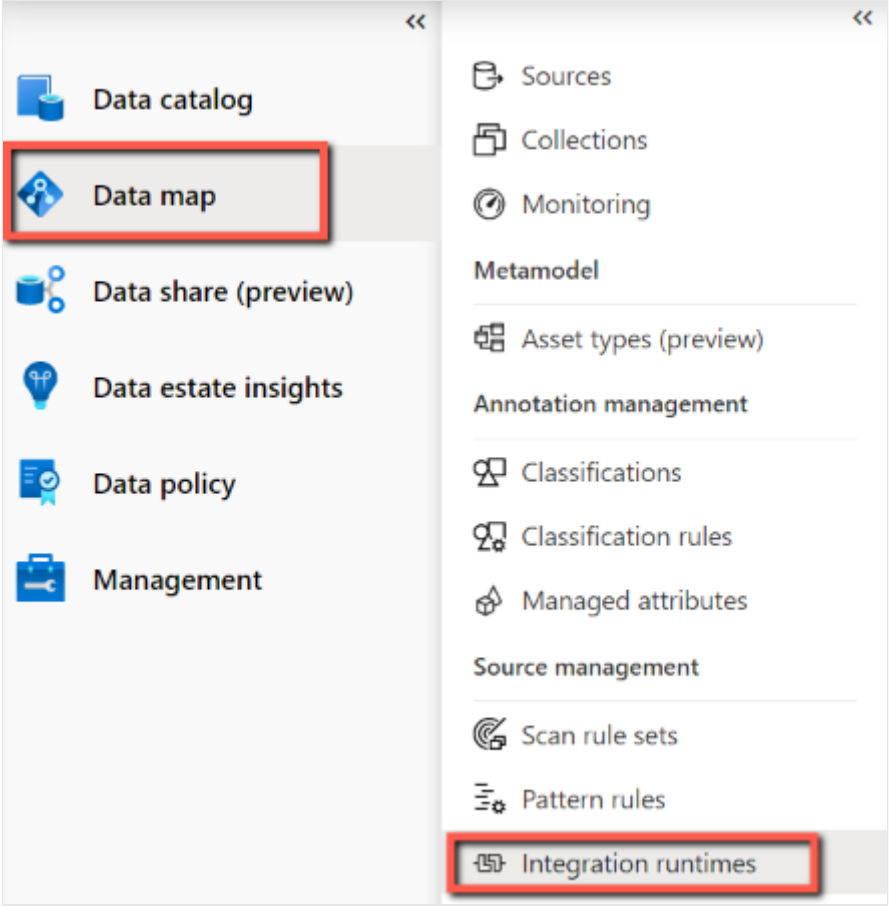
To create and set up a self-hosted integration runtime, use the following procedures.

Create a self-hosted integration runtime

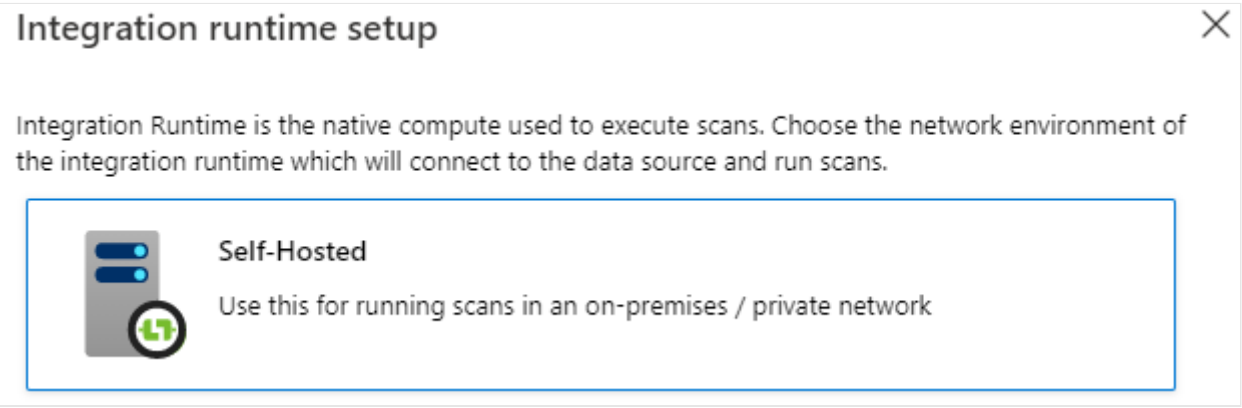
> **Note**
>
> To add or manage a SHIR in Microsoft Purview, you'll need data source administrator permissions in Microsoft Purview.

1. On the home page of the classic Microsoft Purview governance portal, select **Data Map** from the left navigation pane.

2. Under **Sources and scanning** on the left pane, select **Integration runtimes**, and then select + **New**.

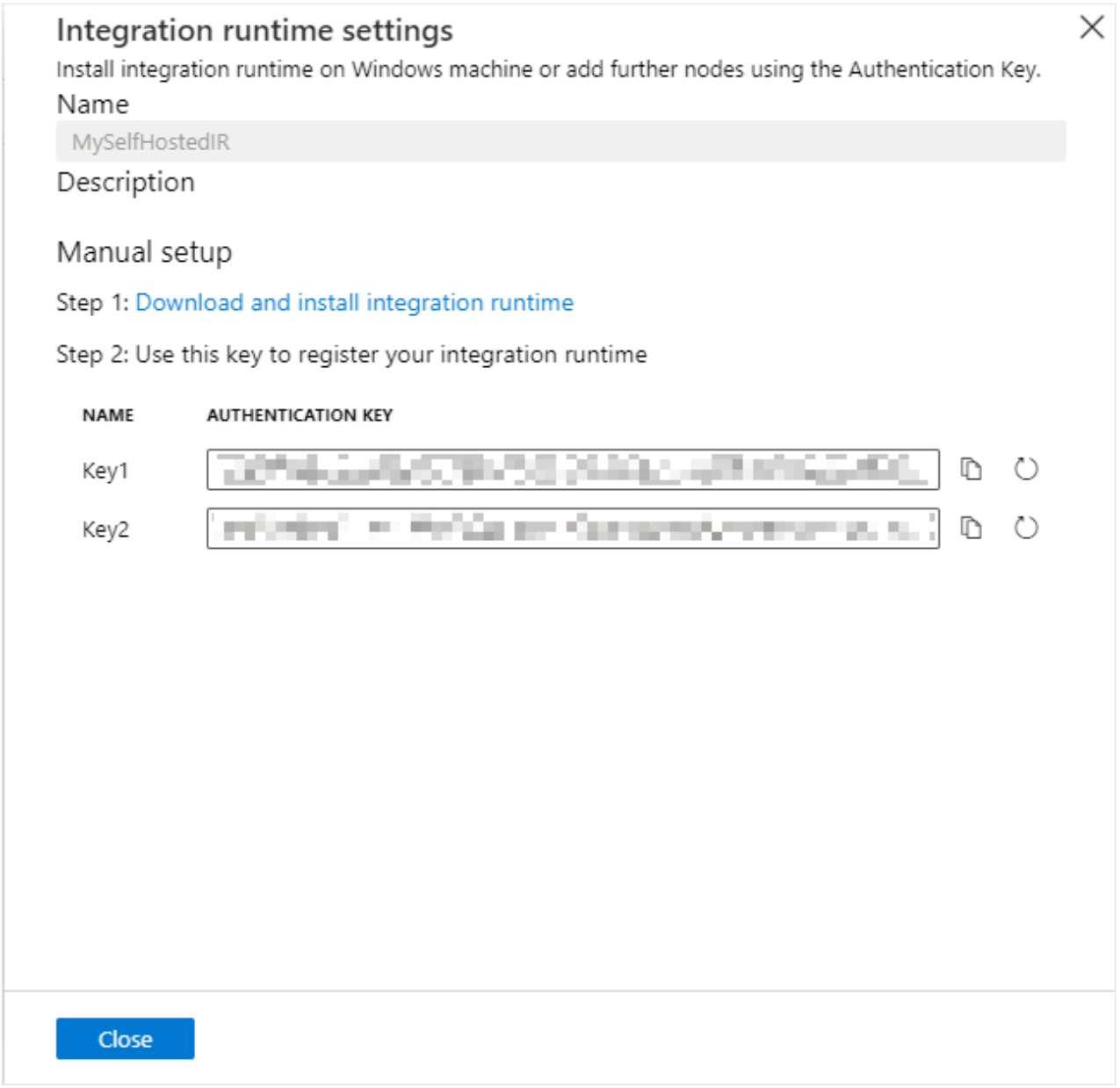


3. On the **Integration runtime setup** page, select **Self-Hosted** to create a self-Hosted IR, and then select **Continue**.

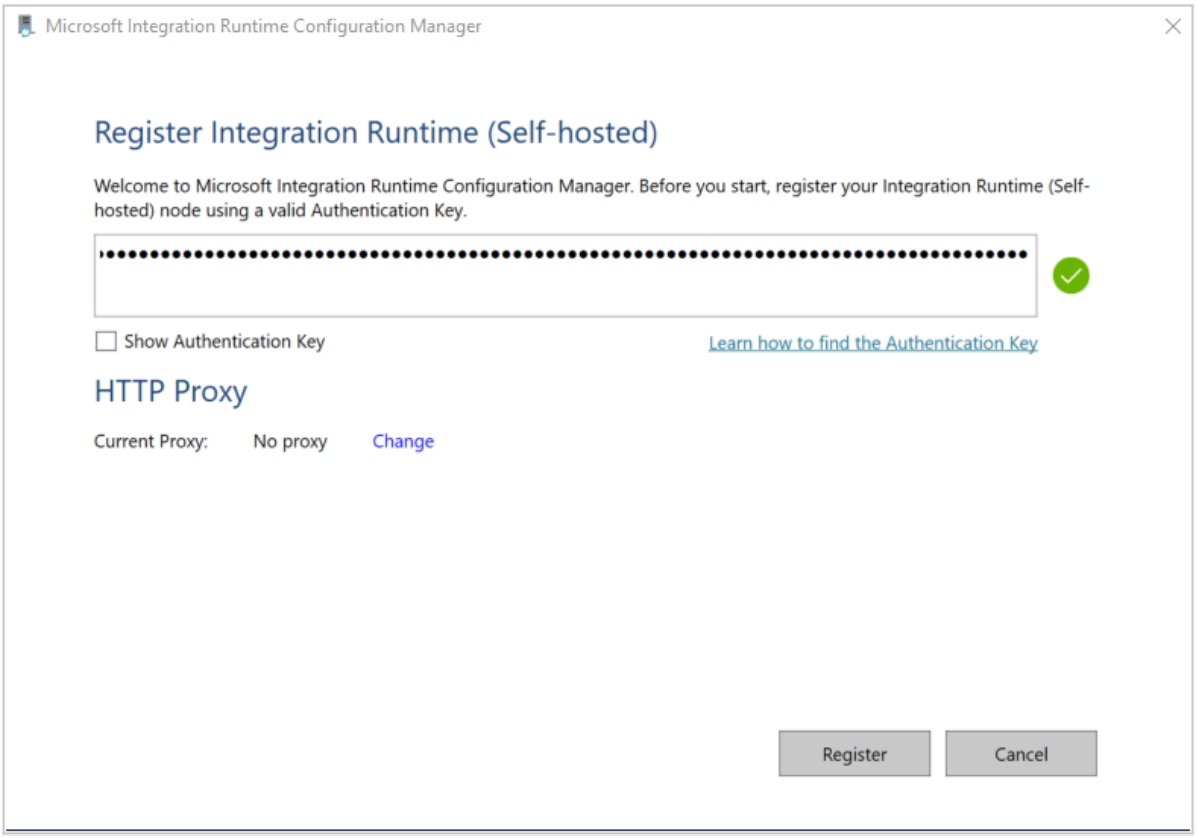


4. Enter a name for your IR, and select Create.

5. On the **Integration Runtime settings** page, follow the steps under the **Manual setup** section. You'll have to download the integration runtime from the download site onto a VM or machine where you intend to run it.

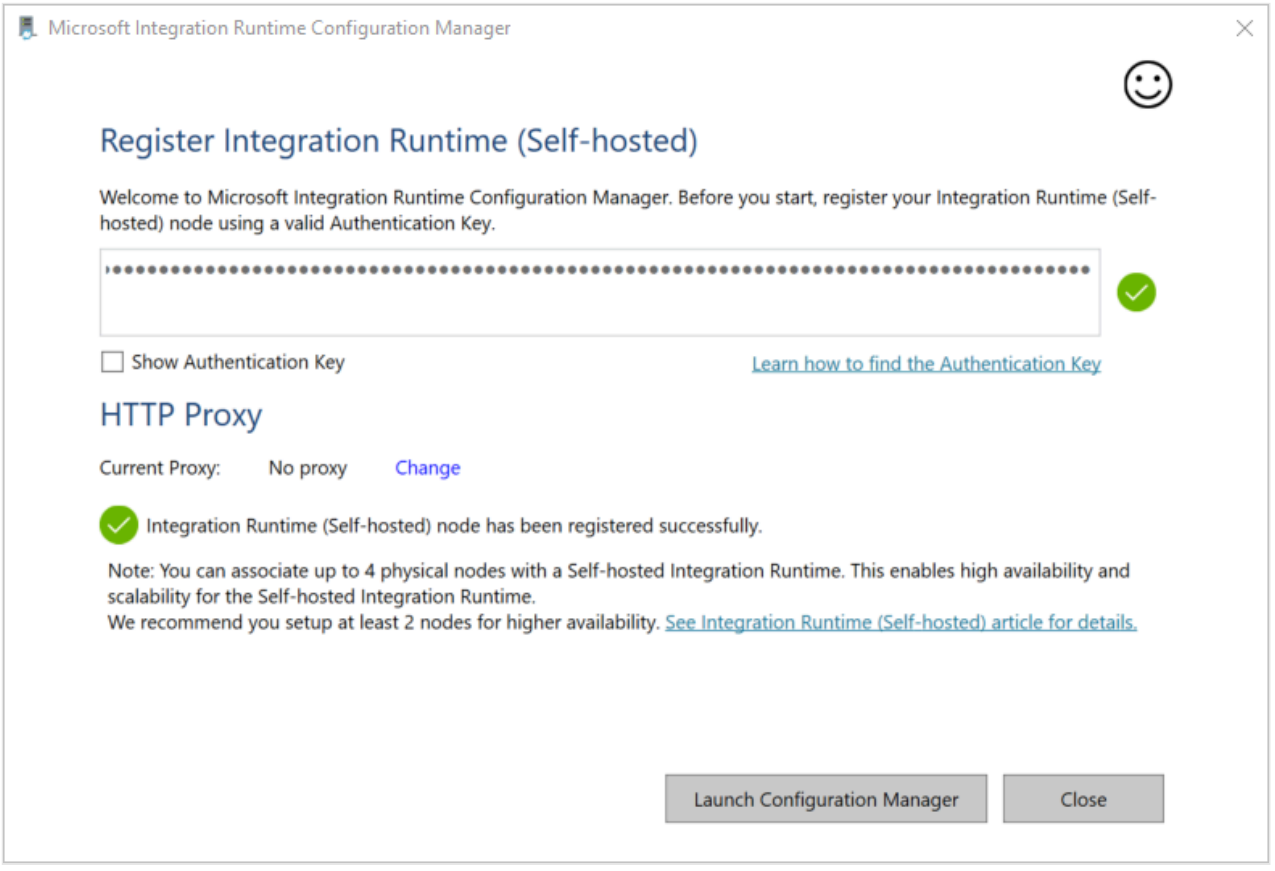


- Copy and paste the authentication key.
- Download the self-hosted integration runtime from Microsoft Integration Runtime on a local Windows machine. Run the installer. Self-hosted integration runtime versions such as 5.4.7803.1 and 5.6.7795.1 are supported.
- On the **Register Integration Runtime (Self-hosted)** page, paste one of the two keys you saved earlier, and select **Register**.



- On the **New Integration Runtime (Self-hosted) Node** page, select **Finish**.

6. After the Self-hosted integration runtime is registered successfully, you see the following window:

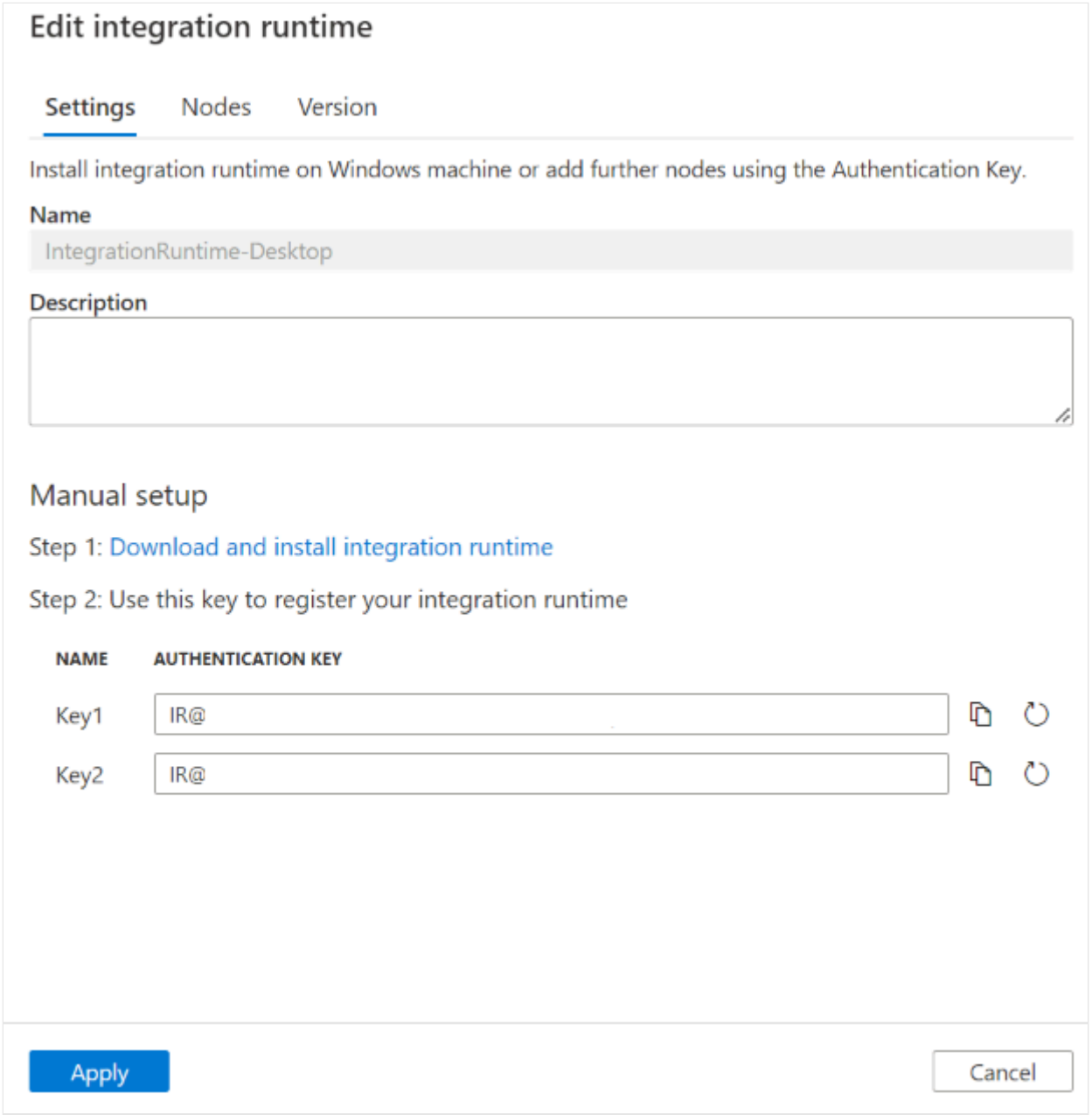


You can register multiple nodes for a self-hosted integration runtime using the same key. Learn more from High availability and scalability.

Manage a self-hosted integration runtime

You can edit a self-hosted integration runtime by navigating to **Integration runtimes** in the classic Microsoft Purview governance portal by hovering on the IR, then select **Edit**.

- In the **Settings** tab, you can update the description, copy the key, or regenerate new keys.
- In the **Nodes** tab, you can see a list of the registered nodes, along with the status, IP address, and the option of node deletion. Learn more from High availability and scalability.
- In the **Version** tab, you can see the IR version status. Learn more from Self-hosted integration runtime autoupdate and expire notification.



You can delete a self-hosted integration runtime by navigating to **Integration runtimes**, hover on the IR then select the **Delete** button.

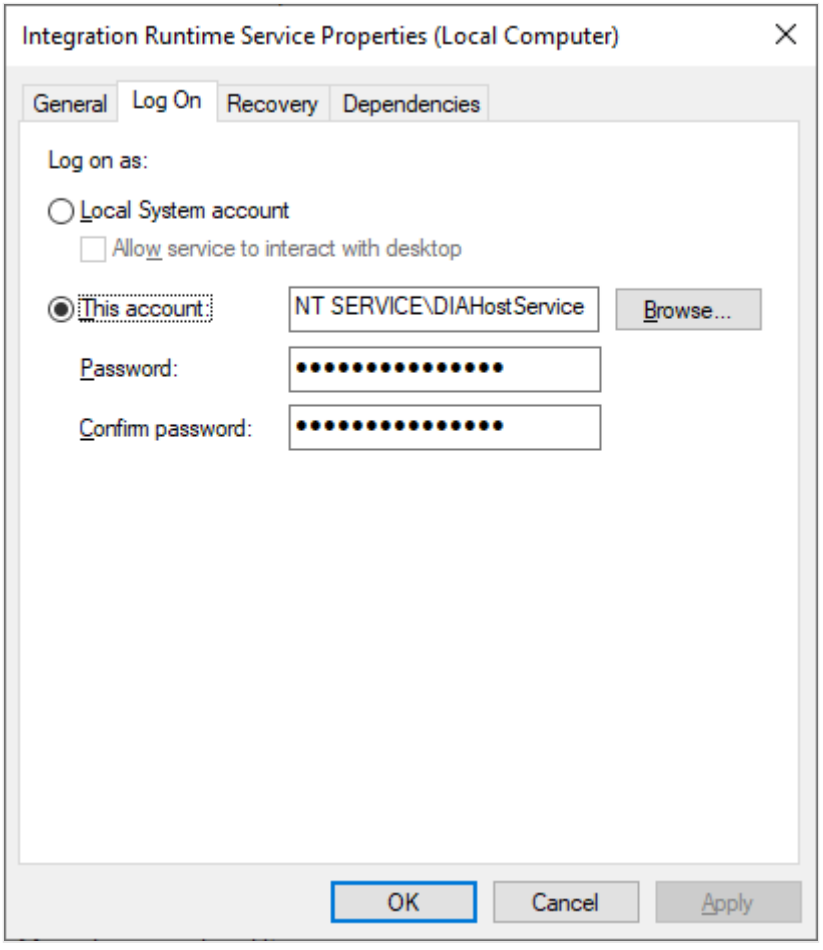
Notification area icons and notifications

If you move your cursor over the icon or message in the notification area, you can see details about the state of the self-hosted integration runtime.



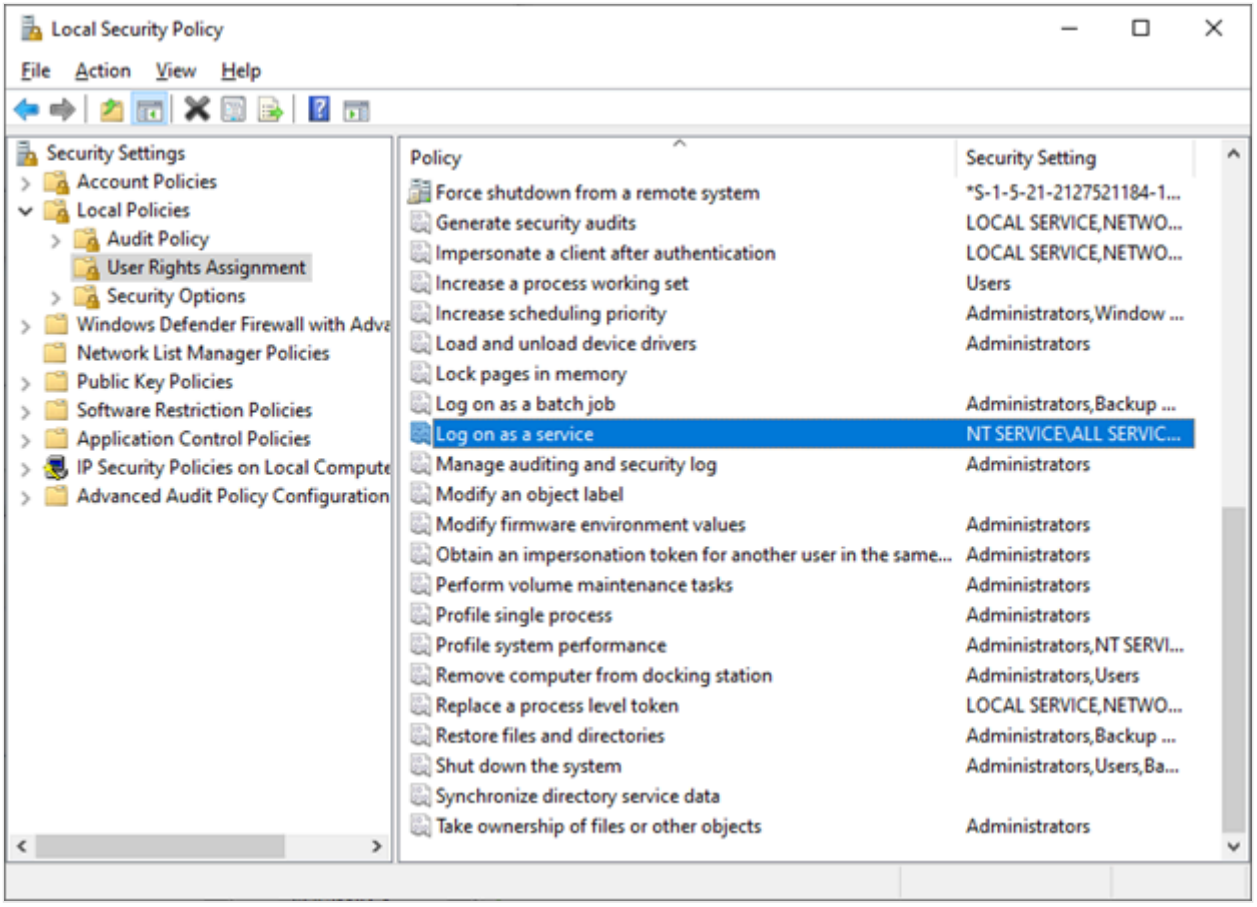
Service account for Self-hosted integration runtime

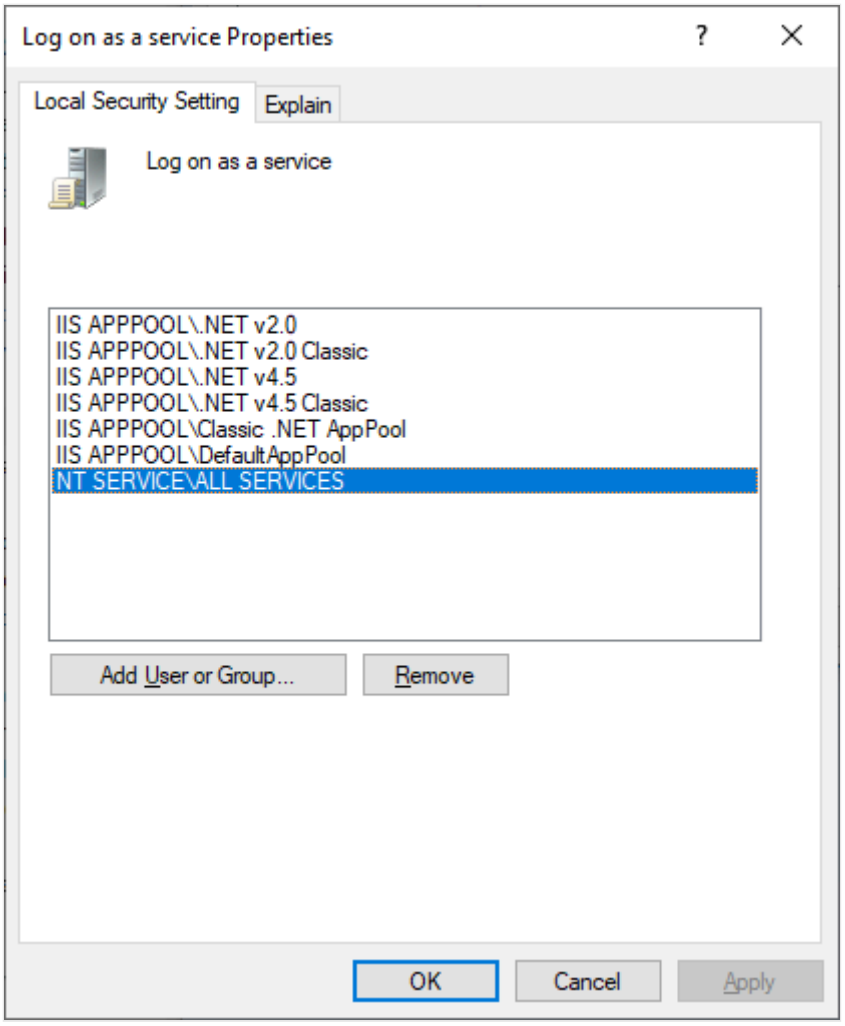
The default sign-in service account of self-hosted integration runtime is **NT SERVICE\DIAHostService**. You can see it in **Services** -> **Integration Runtime Service** -> **Properties** -> **Log on**.



Make sure the account has the permission of Log-on as a service. Otherwise self-hosted integration runtime can't start successfully. You can check the permission in **Local Security**

Policy -> Security Settings -> Local Policies -> User Rights Assignment -> Log on as a service





High availability and scalability

You can associate a self-hosted integration runtime with multiple on-premises machines or virtual machines in Azure. These machines are called nodes. You can have up to four nodes associated with a self-hosted integration runtime. The benefits of having multiple nodes are:

- Higher availability of the self-hosted integration runtime so that it's no longer the single point of failure for scan. This availability helps ensure continuity when you use up to four nodes.
- Run more concurrent scans. Each self-hosted integration runtime can empower many scan runs at the same time, auto determined based on the machine's CPU/memory. You can install more nodes if you have more concurrency need.
- When scanning sources like Azure Blob, Azure Data Lake Storage Gen1, Azure Data Lake Storage Gen2 and Azure Files, each scan run can use all those nodes to boost the scan performance. For other sources, scan will be executed on one of the nodes.

You can associate multiple nodes by installing the self-hosted integration runtime software from Download Center. Then, register it by using the same authentication key.

> **Note**
>
> Before you add another node for high availability and scalability, ensure that the **Remote access to intranet** option is enabled on the first node. To do so, select **Microsoft Integration Runtime Configuration Manager** > **Settings** > **Remote access to intranet**.

Networking requirements

Your self-hosted integration runtime machine needs to connect to several resources to work correctly:

- The Microsoft Purview services used to manage the self-hosted integration runtime.
- The data sources you want to scan using the self-hosted integration runtime.
- If your account was created before December 15, 2023, your integration runtime needs to be able to connect to the managed storage account created by Microsoft Purview. If your account is created after this date (or deployed using API version 2023-05-01-preview onwards), an ingestion storage account is used. Microsoft Purview uses this resource to ingest the results of the scan, among many other things.

There are two firewalls to consider:

- The *corporate firewall* that runs on the central router of the organization
- The *Windows Firewall* that is configured as a daemon on the local machine where the self-hosted integration runtime is installed

Here are the domains and outbound ports that you need to allow at both **corporate and Windows/machine firewalls**.

> **Tip**
>
> - For domains listed with '<managed_storage_account>', add the name of the managed resources associated with your Microsoft Purview account. You can find them from Azure portal -> your Microsoft Purview account -> **Settings** -> **Managed resources** tab.
> - If your account doesn't have a managed storage account, it's using ingestion storage. Refer to the domains with '<ingestion_storage_account>' in below table. You can find the storage information from Azure portal -> **Properties** -> **Ingestion storage ID**. To check the endpoint details, go to **Overview** -> **JSON View** -> "**primaryEndpoint**" property.

Expand table

Domain names	Outbound ports	Description
Public cloud: `*.frontend.clouddatahub.net` Azure Government: `*.frontend.datamovement.azure.us` China: `*.frontend.datamovement.azure.cn`	443	Required to connect to the Microsoft Purview service. Currently wildcard is required as there's no dedicated resource.
Public cloud: `*.servicebus.windows.net` Azure Government: `*.servicebus.usgovcloudapi.net` China: `*.servicebus.chinacloudapi.cn`	443	Required for setting up scan in the classic Microsoft Purview governance portal. This endpoint is used for interactive authoring from UI, for example, test connection, browse folder list and table list to scope scan. To avoid using wildcard, see Get URL of Azure Relay.
Public cloud: `<tenantId>-api.purview-service.microsoft.com` Azure Government: `<tenantId>-api.purview-service.microsoft.us` China: `<tenantId>-api.purview-service.microsoft.cn`	443	Required to connect to Microsoft Purview service. If you use Purview Private Endpoints, this endpoint is covered by *platform private endpoint*.
Public cloud: `<purview_account>.purview.azure.com` Azure Government: `<purview-account>.purview.azure.us` China: `<purview_account>.purview.azure.cn`	443	Required to connect to Microsoft Purview service. If you use Purview Private Endpoints, this endpoint is covered by *account private endpoint*.
Public cloud: `<managed_storage_account>.blob.core.windows.net` or `<ingestion_storage_account>.*.blob.storage.azure.net` Azure Government: `<managed_storage_account>.blob.core.usgovcloudapi.net` or `<ingestion_storage_account>.blob.core.usgovcloudapi.net` China:	443	Required to connect to the Microsoft Purview managed Azure Blob storage account. If you use Purview Private Endpoints, this endpoint is covered by *ingestion private endpoint*.

Domain names	Outbound ports	Description
`<managed_storage_account>.blob.core.chinacloudapi.cn` or `<ingestion_storage_account>.blob.core.chinacloudapi.cn`		
Public cloud: `<managed_storage_account>.queue.core.windows.net` or `<ingestion_storage_account>.*.queue.storage.azure.net` Azure Government: `<managed_storage_account>.queue.core.usgovcloudapi.net` or `<ingestion_storage_account>.queue.core.usgovcloudapi.net` China: `<managed_storage_account>.queue.core.chinacloudapi.cn` or `<ingestion_storage_account>.queue.core.chinacloudapi.cn`	443	Required to connect to the Microsoft Purview managed Azure Queue storage account. If you use Purview Private Endpoints, this endpoint is covered by *ingestion private endpoint*.
`download.microsoft.com`	443	Required to download the self-hosted integration runtime updates. If you have disabled autoupdate, you can skip configuring this domain.
Public cloud: `login.windows.net` and `login.microsoftonline.com` Azure Government: `login.microsoftonline.us` China: `login.partner.microsoftonline.cn`	443	Required to sign in to the Microsoft Entra ID.

> **Note**
>
> As currently Azure Relay doesn't support service tag, you have to use service tag AzureCloud or Internet in NSG rules for the communication to Azure Relay.

Depending on the sources you want to scan, you also need to allow other domains and outbound ports for other Azure or external sources. A few examples are provided here:

Expand table

Domain names	Outbound ports	Description
`<your_storage_account>.dfs.core.windows.net`	443	When scan Azure Data Lake Store Gen 2.
`<your_storage_account>.blob.core.windows.net`	443	When scan Azure Blob storage.

Domain names	Outbound ports	Description
`<your_sql_server>.database.windows.net`	1433	When scan Azure SQL Database.
`*.powerbi.com` and `*.analysis.windows.net`	443	When scan Power BI tenant.
`<your_ADLS_account>.azuredatalakestore.net`	443	When scan Azure Data Lake Store Gen 1.
Various domains	Dependent	Domains and ports for any other sources the SHIR will scan.

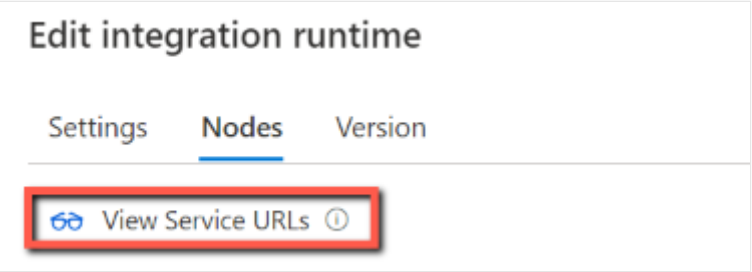
For some cloud data stores such as Azure SQL Database and Azure Storage, you may need to allow IP address of self-hosted integration runtime machine on their firewall configuration, or you can create private endpoint of the service in your self-hosted integration runtime's network.

> **Important**
>
> In most environments, you will also need to make sure that your DNS is correctly configured. To confirm, you can use **nslookup** from your SHIR machine to check connectivity to each of the domains. Each nslookup should return the IP of the resource. If you are using Private Endpoints, the private IP should be returned and not the Public IP. If no IP is returned, or if when using Private Endpoints the public IP is returned, you need to address your DNS/VNet association, or your Private Endpoint/VNet peering.

Get URL of Azure Relay

One required domain and port that need to be put in the allowlist of your firewall is for the communication to Azure Relay. The self-hosted integration runtime uses it for interactive authoring such as test connection and browse folder/table list. If you don't want to allow **.servicebus.windows.net** and would like to have more specific URLs, then you can see all the FQDNs that are required by your self-hosted integration runtime. Follow these steps:

1. Go to the classic Microsoft Purview governance portal -> Data Map -> Integration runtimes, and edit your self-hosted integration runtime.
2. In Edit page, select **Nodes** tab.
3. Select **View Service URLs** to get all FQDNs.

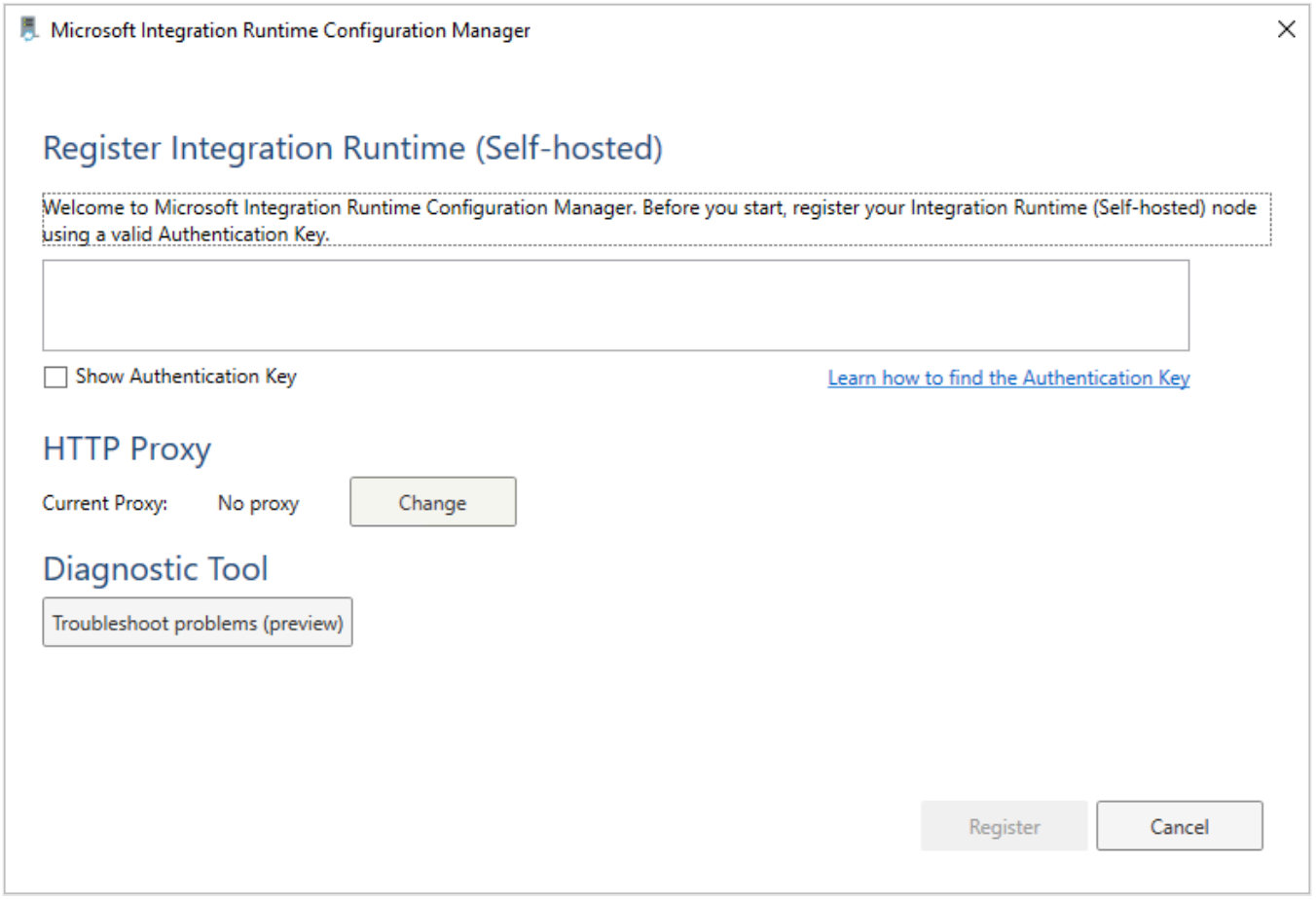


4. You can add these FQDNs in the allowlist of firewall rules.

> ⓘ **Note**
>
> For the details related to Azure Relay connections protocol, see Azure Relay Hybrid Connections protocol.

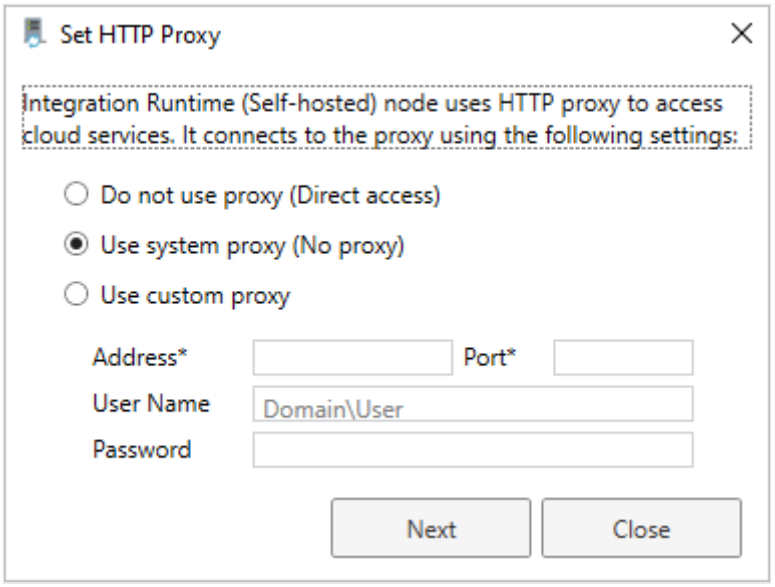
Proxy server considerations

If your corporate network environment uses a proxy server to access the internet, configure the self-hosted integration runtime to use appropriate proxy settings. You can set the proxy during the initial registration phase or after it's being registered.



When configured, the self-hosted integration runtime uses the proxy server to connect to the services that use HTTP or HTTPS protocol. This is why you select **Change link** during initial

setup.



There are two supported configuration options by Microsoft Purview:

- **Do not use proxy**: The self-hosted integration runtime doesn't explicitly use any proxy to connect to cloud services.
- **Use system proxy**: The self-hosted integration runtime uses the proxy setting that is configured in the executable's configuration files. If no proxy is specified in these files, the self-hosted integration runtime connects to the services directly without going through a proxy.
- **Use custom proxy**: Configure the HTTP proxy setting to use for the self-hosted integration runtime, instead of using configurations in diahost.exe.config and diawp.exe.config. **Address** and **Port** values are required. **User Name** and **Password** values are optional, depending on your proxy's authentication setting. All settings are encrypted with Windows DPAPI on the self-hosted integration runtime and stored locally on the machine.

> ⓘ **Note**
>
> Connecting to data sources via proxy isn't supported for connectors other than Azure data sources and Power BI.

The integration runtime host service restarts automatically after you save the updated proxy settings.

After you register the self-hosted integration runtime, if you want to view or update proxy settings, use Microsoft Integration Runtime Configuration Manager.

1. Open **Microsoft Integration Runtime Configuration Manager**.
2. Select the **Settings** tab.

3. Under **HTTP Proxy**, select the **Change** link to open the **Set HTTP Proxy** dialog box.
4. Select **Next**. You then see a warning that asks for your permission to save the proxy setting and restart the integration runtime host service.

> **Note**
>
> If you set up a proxy server with NTLM authentication, the integration runtime host service runs under the domain account. If you later change the password for the domain account, remember to update the configuration settings for the service and restart the service. Because of this requirement, we suggest that you access the proxy server by using a dedicated domain account that doesn't require you to update the password frequently.

If using system proxy, make sure your proxy server allow outbound traffic to the network rules.

Configure proxy server settings

If you select the **Use system proxy** option for the HTTP proxy, the self-hosted integration runtime uses the proxy settings in the following four files under the path C:\Program Files\Microsoft Integration Runtime\5.0\ to perform different operations:

- .\Shared\diahost.exe.config
- .\Shared\diawp.exe.config
- .\Gateway\DataScan\Microsoft.DataMap.Agent.exe.config
- .\Gateway\DataScan\DataTransfer\Microsoft.DataMap.Agent.Connectors.Azure.DataFactory.ServiceHost.exe.config

When no proxy is specified in these files, the self-hosted integration runtime connects to the services directly without going through a proxy.

The following procedure provides instructions for updating the **diahost.exe.config** file.

1. In File Explorer, make a safe copy of C:\Program Files\Microsoft Integration Runtime\5.0\Shared\diahost.exe.config as a backup of the original file.

2. Open Notepad running as administrator.

3. In Notepad, open the text file C:\Program Files\Microsoft Integration Runtime\5.0\Shared\diahost.exe.config.

4. Find the default **system.net** tag as shown in the following code:

```
XML
```

```
<system.net>
    <defaultProxy useDefaultCredentials="true" />
</system.net>
```

You can then add proxy server details as shown in the following example:

XML

```
<system.net>
  <defaultProxy>
    <proxy bypassonlocal="true" proxyaddress="<your proxy server e.g.
http://proxy.domain.org:8888/>" />
  </defaultProxy>
</system.net>
```

The proxy tag allows other properties to specify required settings like `scriptLocation`. See <proxy> Element (Network Settings) for syntax.

XML

```
<proxy autoDetect="true|false|unspecified"
bypassonlocal="true|false|unspecified" proxyaddress="uriString"
scriptLocation="uriString" usesystemdefault="true|false|unspecified "/>
```

5. Save the configuration file in its original location.

Repeat the same procedure to update **diawp.exe.config** and **Microsoft.DataMap.Agent.exe.config** files.

Then go to path C:\Program Files\Microsoft Integration Runtime\5.0\Gateway\DataScan\DataTransfer, create a file named "**Microsoft.DataMap.Agent.Connectors.Azure.DataFactory.ServiceHost.exe.config**", and configure the proxy setting as follows. You can also extend the settings as described above.

XML

```
<?xml version="1.0" encoding="utf-8"?>
<configuration>
  <system.net>
    <defaultProxy>
      <proxy bypassonlocal="true" proxyaddress="<your proxy server e.g.
http://proxy.domain.org:8888/>" />
    </defaultProxy>
  </system.net>
</configuration>
```

Local traffic must be excluded from proxy, for example if your Microsoft Purview account is behind private endpoints. In such cases, update the following four files under the path to include bypass list C:\Program Files\Microsoft Integration Runtime\5.0\ with required bypass list:

- .\Shared\diahost.exe.config
- .\Shared\diawp.exe.config
- .\Gateway\DataScan\Microsoft.DataMap.Agent.exe.config
- .\Gateway\DataScan\DataTransfer\Microsoft.DataMap.Agent.Connectors.Azure.DataFactory.ServiceHost.exe.config

An example for bypass list for scanning an Azure SQL Database and ADLS gen 2 Storage:

XML

```
 <system.net>
   <defaultProxy>
     <bypasslist>
       <add address="scaneastus4123.blob.core.windows.net" />
       <add address="scaneastus4123.queue.core.windows.net" />
       <add address="Atlas-abc12345-1234-abcd-a73c-
394243a566fa.servicebus.windows.net" />
       <add address="contosopurview123.purview.azure.com" />
       <add address="contososqlsrv123.database.windows.net" />
       <add address="contosoadls123.dfs.core.windows.net" />
       <add address="contosoakv123.vault.azure.net" />
     </bypasslist>
     <proxy proxyaddress=http://proxy.domain.org:8888 bypassonlocal="True" />
   </defaultProxy>
 </system.net>
```

Restart the self-hosted integration runtime host service, which picks up the changes. To restart the service, use the services applet from Control Panel. Or from Integration Runtime Configuration Manager, select the **Stop Service** button, and then select **Start Service**. If the service doesn't start, you likely added incorrect XML tag syntax in the application configuration file that you edited.

> **Important**
>
> Don't forget to update all four files mentioned above.

You also need to make sure that Microsoft Azure is in your company's allowlist. You can download the list of valid Azure IP addresses. IP ranges for each cloud, broken down by region and by the tagged services in that cloud are now available on MS Download:

- Public: https://www.microsoft.com/download/details.aspx?id=56519

Possible symptoms for issues related to the firewall and proxy server

If you see error messages like the following ones, the likely reason is improper configuration of the firewall or proxy server. Such configuration prevents the self-hosted integration runtime from connecting to Microsoft Purview services. To ensure that your firewall and proxy server are properly configured, refer to the previous section.

- When you try to register the self-hosted integration runtime, you receive the following error message: "Failed to register this Integration Runtime node! Confirm that the Authentication key is valid and the integration service host service is running on this machine."

- When you open Integration Runtime Configuration Manager, you see a status of **Disconnected** or **Connecting**. When you view Windows event logs, under **Event Viewer** > **Application and Services Logs** > **Microsoft Integration Runtime**, you see error messages like this one:

  Output

  ```
  Unable to connect to the remote server
  A component of Integration Runtime has become unresponsive and restarts
  automatically. Component name: Integration Runtime (Self-hosted)
  ```

Java Runtime Environment Installation

If you scan Parquet files using the self-hosted integration runtime with Microsoft Purview, you'll need to install either the Java Runtime Environment or OpenJDK on your self-hosted IR machine.

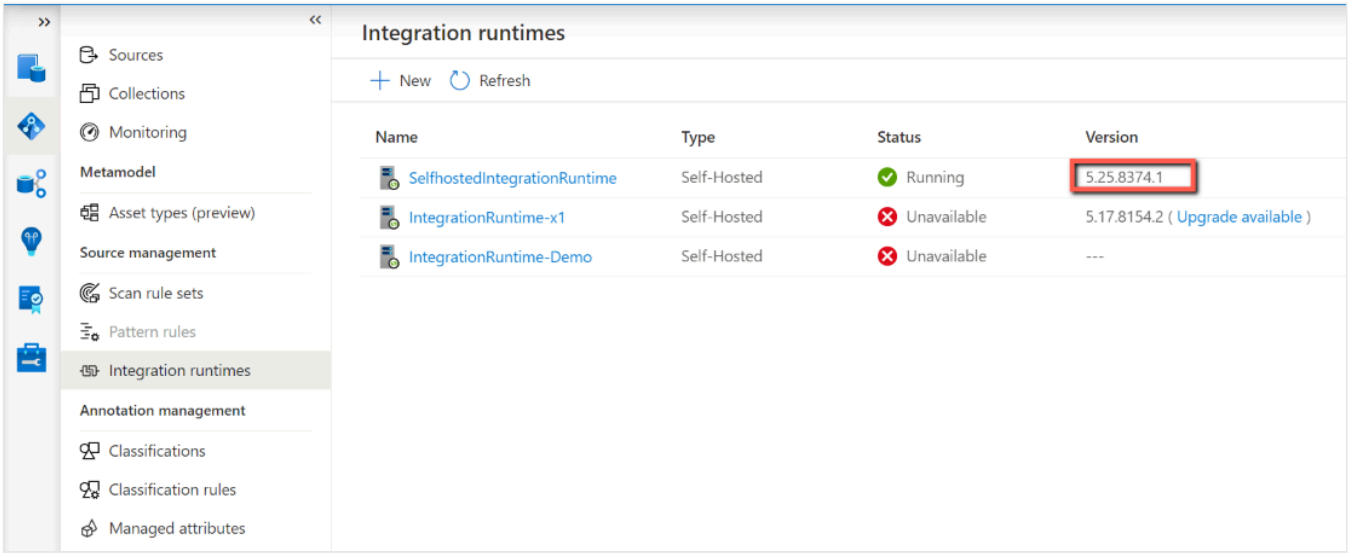
When scanning Parquet files using the self-hosted IR, the service locates the Java runtime by firstly checking the registry `(HKEY_LOCAL_MACHINE\SOFTWARE\JavaSoft\Java Runtime Environment\{Current Version}\JavaHome)` for JRE, if not found, secondly checking system variable `JAVA_HOME` for OpenJDK. You can set JAVA_HOME under System Settings, Environment Variables on your machine. Create or edit the JAVA_HOME variable to point to the Java jre on your machine. For example: `C:\Program Files\Java\jdk1.8\jre`

- **To use JRE**: The 64-bit IR requires 64-bit JRE. You can find it from here.
- **To use OpenJDK**: It's supported since IR version 3.13. Package the jvm.dll with all other required assemblies of OpenJDK into self-hosted IR machine, and set system environment

variable JAVA_HOME accordingly.

How to check your self-hosted integration runtime version

You can check the version of your self-hosted integration runtime in the classic Microsoft Purview governance portal -> Data Map -> Integration runtimes:

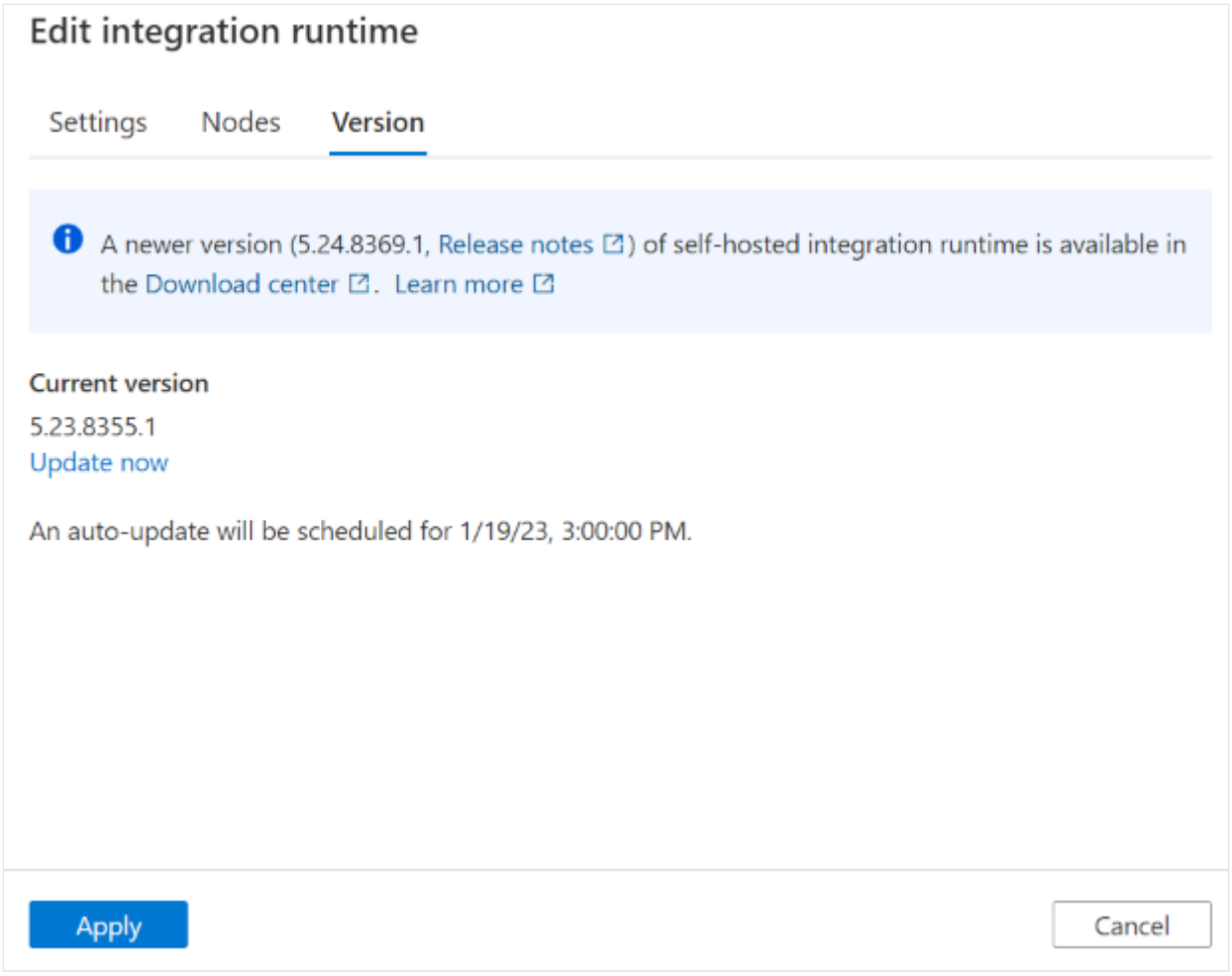


You can also check the version in your self-hosted integration runtime client -> Help tab.

Self-hosted Integration Runtime autoupdate

Auto update is enabled by default when you install a self-hosted integration runtime. You have two options to manage the version of self-hosted integration runtime: autoupdate or maintain manually. Typically, Microsoft Purview releases two new versions of self-hosted integration runtime every month, which includes new feature release, bug fix or enhancement. So we recommend users to update to newer version in order to get the latest feature and enhancement.

The self-hosted integration runtime is automatically updated to newer version. When new version is available while not yet scheduled for your instance, you can also trigger the update from the portal.



> **Note**
>
> If you have multiple self-hosted integration runtime nodes, there's no downtime during autoupdate. The autoupdate happens in one node first while others are working on tasks. When the first node finishes the update, it takes over the remaining tasks when other nodes are updating. If you only have one self-hosted integration runtime node, then it has some downtime during the autoupdate.

Autoupdate version vs latest version

To ensure the stability of self-hosted integration runtime, although we release two versions, we only push one version every month. So sometimes you find that the autoupdate version is the previous version of the actual latest version. If you want to get the latest version, you can go to download center and do so manually. Additionally, **auto-update** to a new version is managed by the service, and you can't change it.

The self-hosted integration runtime **Version** tab in the classic Microsoft Purview governance portal shows the newer version if current version is old. When your self-hosted integration runtime is online, this version is the autoupdate version and automatically updates your self-

hosted integration runtime in the scheduled time. But if your self-hosted integration runtime is offline, the page only shows the newer version.

If you have multiple nodes, and for some reasons that some of them aren't autoupdated successfully. Then these nodes roll back to the version, which was the same across all nodes before autoupdate.

Self-hosted Integration Runtime expiration

Each version of self-hosted integration runtime expires in one year. The expiring message is shown in the classic Microsoft Purview governance portal and the self-hosted integration runtime client **90 days** before expiration.

Next steps

- Microsoft Purview network architecture and best practices
- Use private endpoints with Microsoft Purview

Troubleshoot Microsoft Purview self-hosted integration runtime (SHIR)

Article • 09/26/2023

This article explores common troubleshooting methods for self-hosted integration runtime (SHIR) in Microsoft Purview.

The self-hosted integration runtime (SHIR) is also used by Azure Data Factory and Azure Synapse Analytics. Though many of the troubleshooting steps overlap, follow this guide to troubleshoot your SHIR for those products:

- Troubleshoot SHIR for Azure Data Factory and Azure Synapse Analytics

Gather Microsoft Purview specific SHIR self-hosted integration runtime logs

For failed Microsoft Purview scans that are running on a self-hosted IR, the service supports viewing and uploading error logs from the Windows Event Viewer.

You can look up any errors you see in the error guide below. To get support and troubleshooting guidance for SHIR issues, you may need to generate an error report ID and reach out to Microsoft support.

To generate the error report ID for Microsoft Support, follow these instructions:

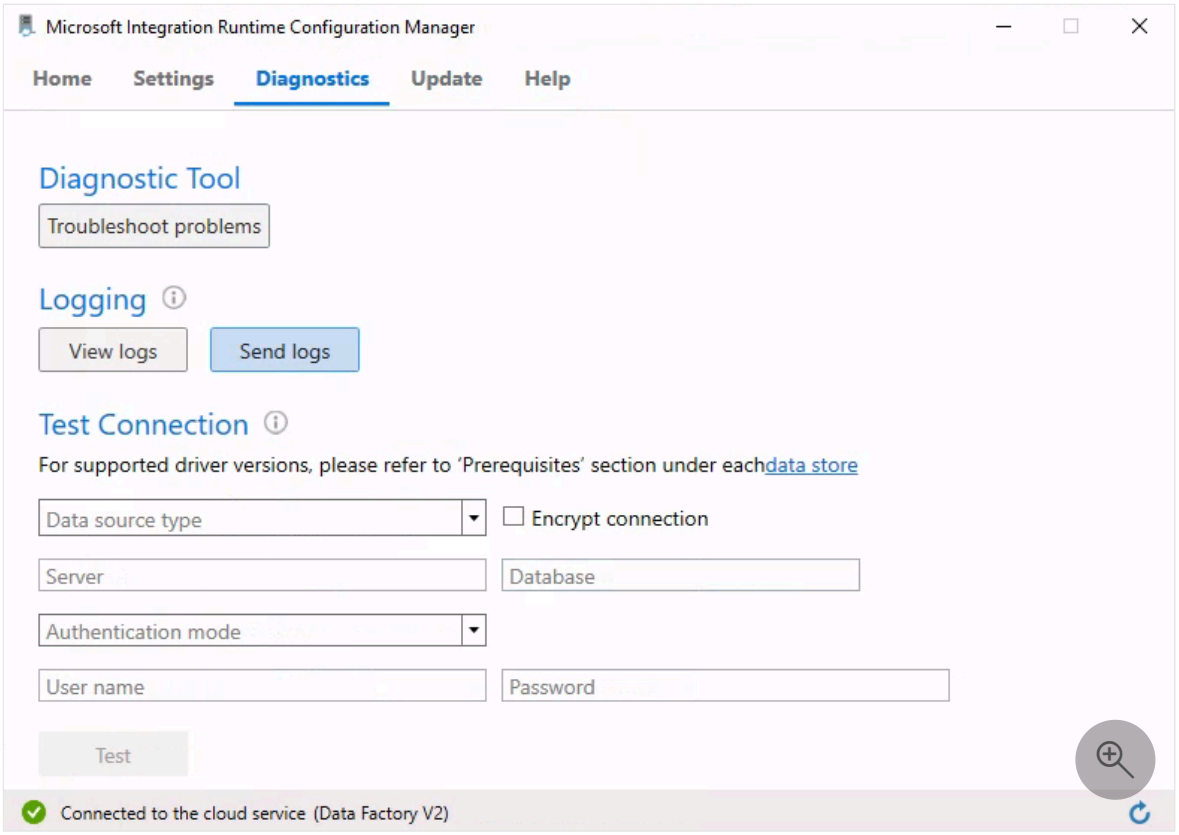
1. Before starting a scan in the Microsoft Purview governance portal:
   a. Navigate to the machine where the self-hosted integration runtime is installed and open the Windows Event Viewer.
   b. Clear the windows event viewer logs in the **Integration Runtime** section. Right-click on the logs and select the clear logs option.
   c. Navigate back to the Microsoft Purview governance portal and start the scan.

2. Once the scan shows status **Failed**, navigate back to the SHIR VM, or machine and refresh the event viewer in the **Integration Runtime** section.

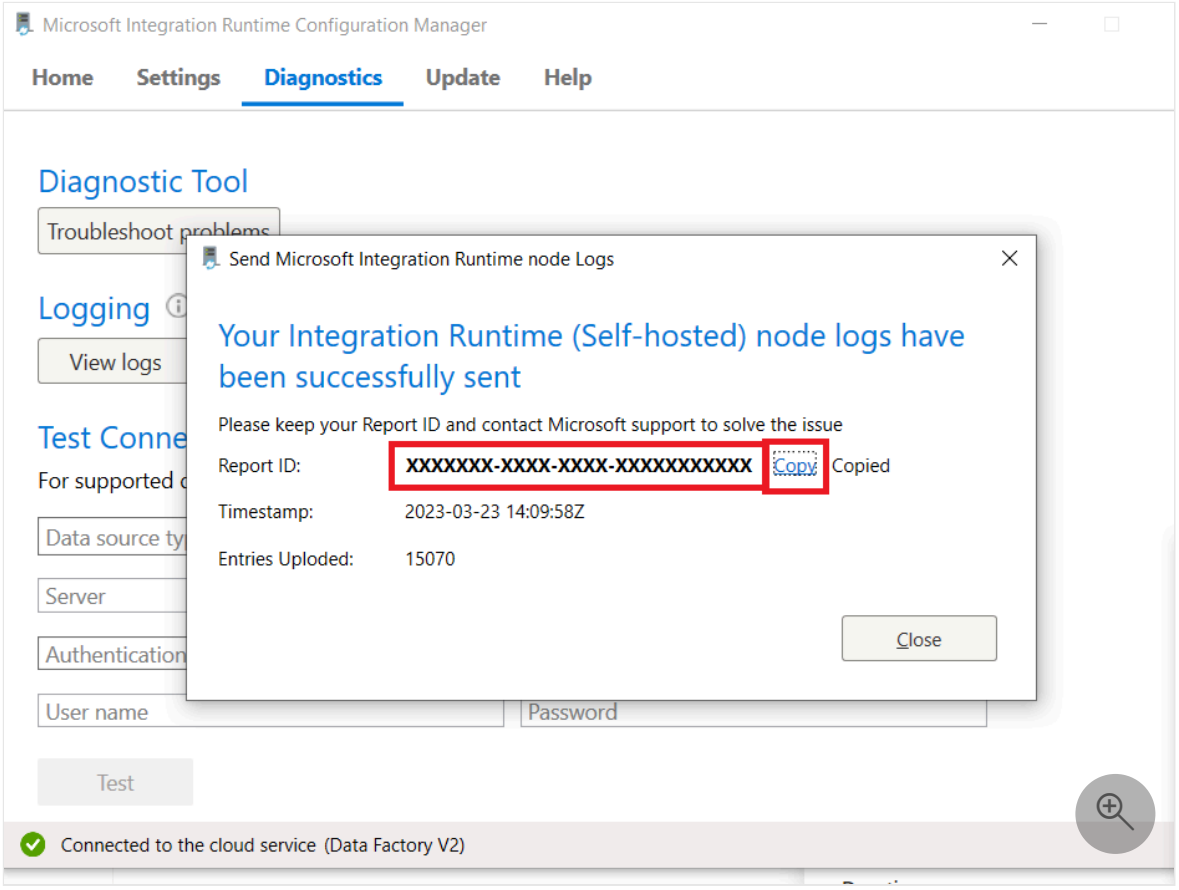
3. The activity logs are displayed for the failed scan run.

4. For further assistance from Microsoft, select **Send Logs**.

   The **Share the self-hosted integration runtime (SHIR) logs with Microsoft** window opens.



5. When the logs are uploaded, keep a record of the Report ID to use if you reach out to Microsoft support.



Self-hosted integration runtime SHIR general failure or error

There are lots of common errors, warnings, and issues between the Purview SHIR, Azure Data Factory, or Azure Synapse SHIR. The guide below covers many of the most common issues.

Self-hosted IR general failure or error

Out of memory issue

Symptoms

An OutOfMemoryException (OOM) error occurs when you try to run a scan with a self-hosted IR.

Cause

A new scan can throw an OOM error if the IR machine experiences momentary high memory usage. The issue might be caused by a large volume of concurrent activity or a low amount of memory and the error is by design.

Resolution

Check the resource usage to confirm if any other software is running at the same time, and confirm that your SHIR machine meets the recommended configurations.

Self-hosted IR could not load file or assembly

Symptoms

You get the following error message:

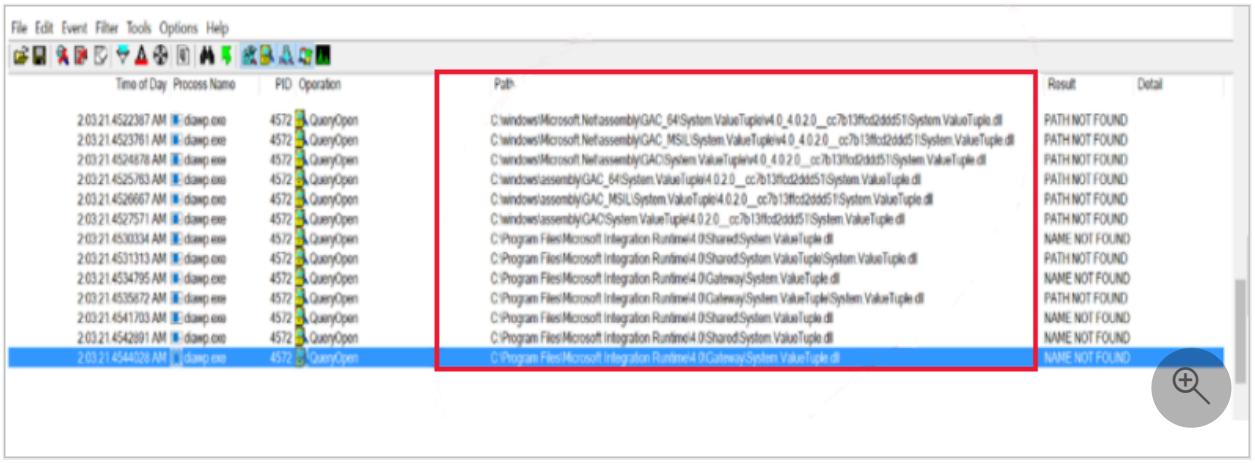
"Could not load file or assembly 'XXXXXXXXXXXXXXXX, Version=4.0.2.0, Culture=neutral, PublicKeyToken=XXXXXXXXX' or one of its dependencies. The system cannot find the file specified. Activity ID: 92693b45-b4bf-4fc8-89da-2d3dc56f27c3"

Here is a more specific error message:

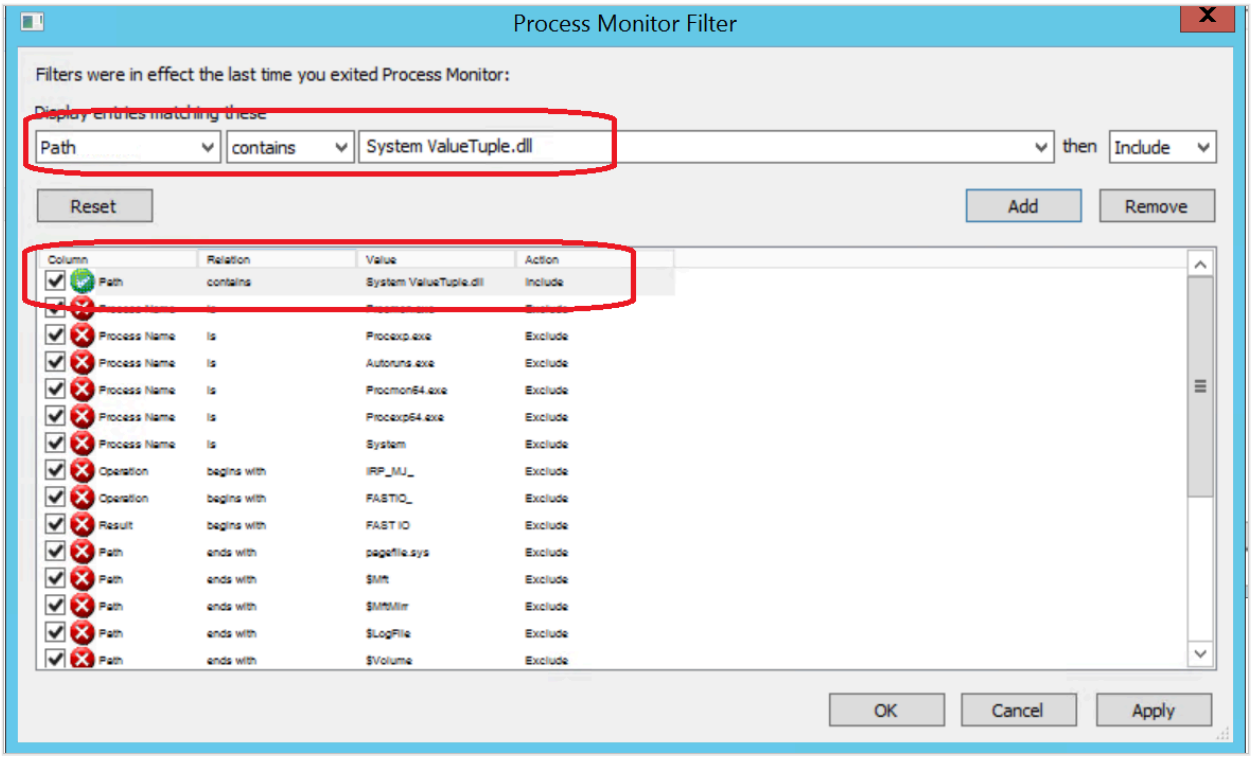
"Could not load file or assembly 'System.ValueTuple, Version=4.0.2.0, Culture=neutral, PublicKeyToken=XXXXXXXXX' or one of its dependencies. The system cannot find the file specified. Activity ID: 92693b45-b4bf-4fc8-89da-2d3dc56f27c3"

Cause

In Process Monitor, you can view the following result:



> **Tip**
>
> In Process Monitor, you can set filters as shown in following screenshot. The preceding error message says that the DLL System.ValueTuple is not located in the related *Global Assembly Cache* (GAC) folder, in the *C:\Program Files\Microsoft Integration Runtime\4.0\Gateway* folder, or in the *C:\Program Files\Microsoft Integration Runtime\4.0\Shared* folder. Basically, the process loads the DLL first from the *GAC* folder, then from the *Shared* folder, and finally from the *Gateway* folder. Therefore, you can load the DLL from any path that's helpful.



Resolution

You'll find the *System.ValueTuple.dll* file in the *C:\Program Files\Microsoft Integration Runtime\4.0\Gateway\DataScan* folder. To resolve the issue, copy the *System.ValueTuple.dll* file to the *C:\Program Files\Microsoft Integration Runtime\4.0\Gateway* folder.

You can use the same method to resolve other missing file or assembly issues.

More information

The reason why you see the *System.ValueTuple.dll* under *%windir%\Microsoft.NET\assembly* and *%windir%\assembly* is that this is a .NET behavior.

In the following error, you can clearly see that the *System.ValueTuple* assembly is missing. This issue arises when the application tries to check the *System.ValueTuple.dll* assembly.

> "<LogProperties><ErrorInfo>[{"Code":0,"Message":"The type initializer for 'Npgsql.PoolManager' threw an exception.","EventType":0,"Category":5,"Data":{},"MsgId":null,"ExceptionType":"System.TypeInitializationException","Source":"Npgsql","StackTrace":"","InnerEventInfos":[{"Code":0,"Message":"Could not load file or assembly 'System.ValueTuple, Version=4.0.2.0, Culture=neutral, PublicKeyToken=XXXXXXXXX' or one of its dependencies. The system cannot find

> the file specified.","EventType":0,"Category":5,"Data":{},"MsgId":null,"ExceptionType":"System.IO.FileNotFoundException","Source":"Npgsql","StackTrace":"","InnerEventInfos":[]}]}]</ErrorInfo></LogProperties>"

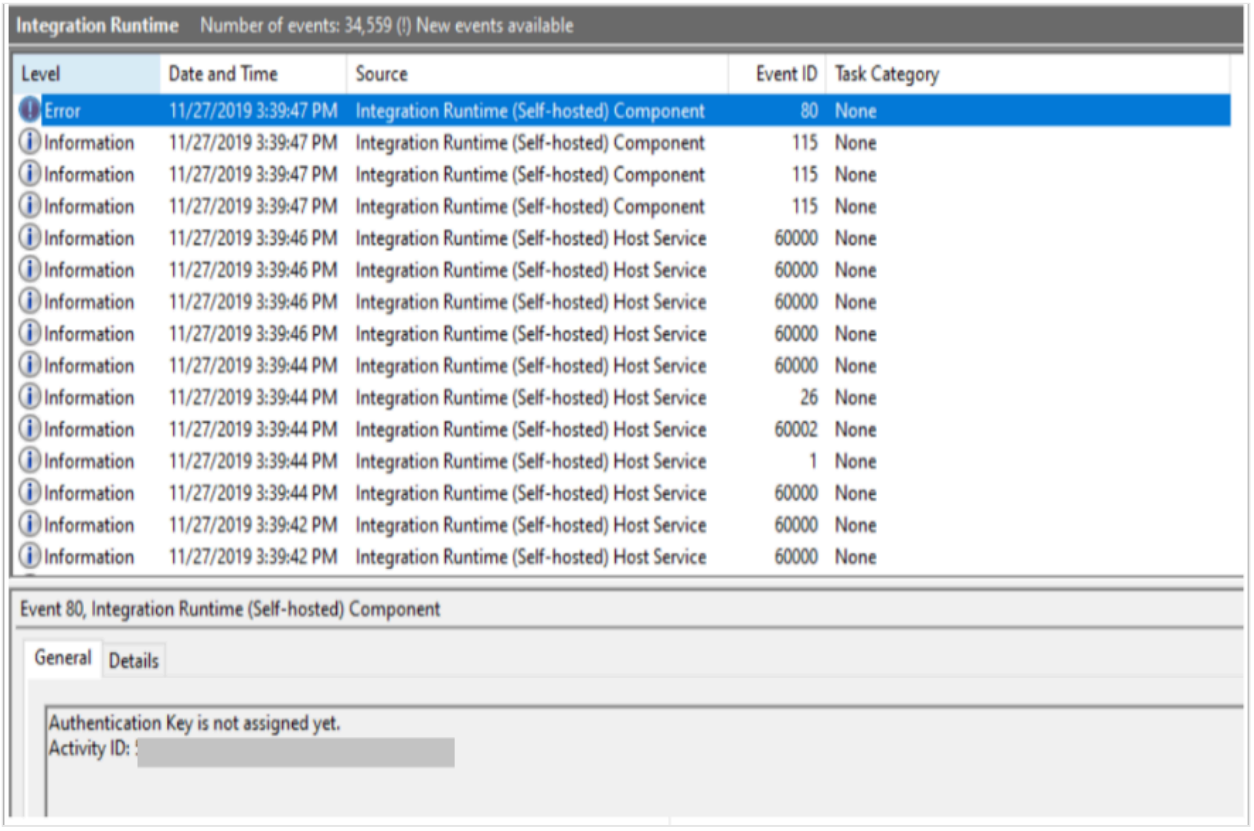
For more information about GAC, see Global Assembly Cache.

Self-hosted integration runtime Authentication Key is missing

Symptoms

The self-hosted integration runtime suddenly goes offline without an Authentication Key, and the Event Log displays the following error message:

"Authentication Key is not assigned yet"

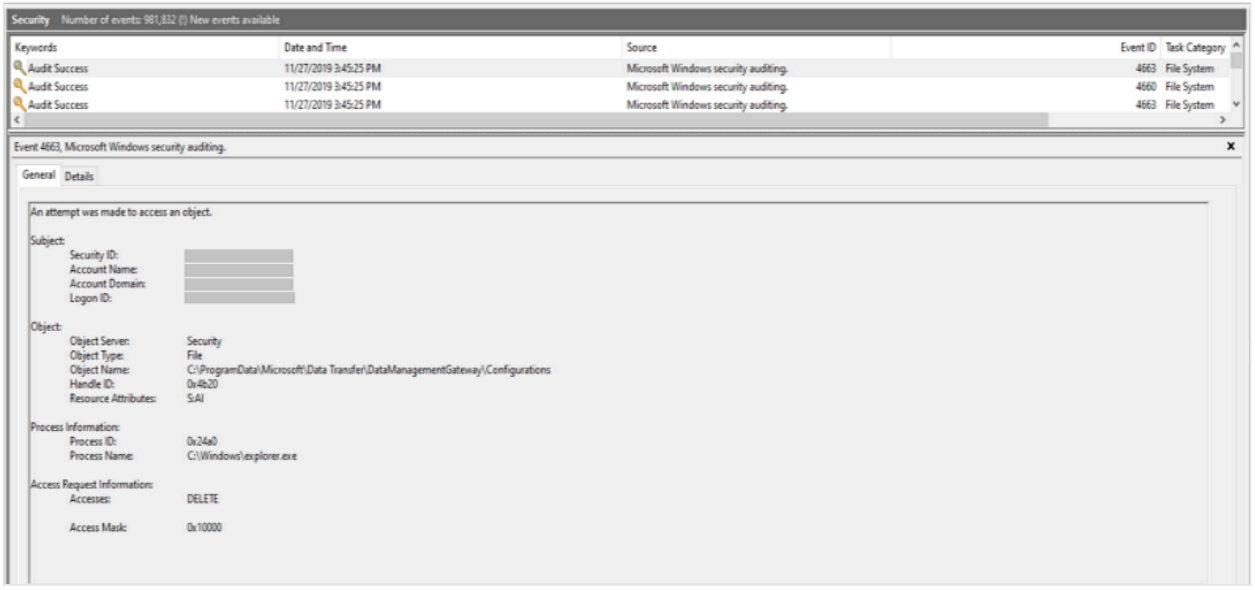


Cause

The self-hosted IR node or logical self-hosted IR in the Microsoft Purview portal was deleted, or a clean uninstall was performed.

Resolution

If neither of the preceding causes applies, you can go to the *%programdata%\Microsoft\Data Transfer\DataManagementGateway* folder to see whether the *Configurations* file has been deleted. If it was deleted, follow the instructions in the Netwrix article Detect who deleted a file from your Windows file servers.

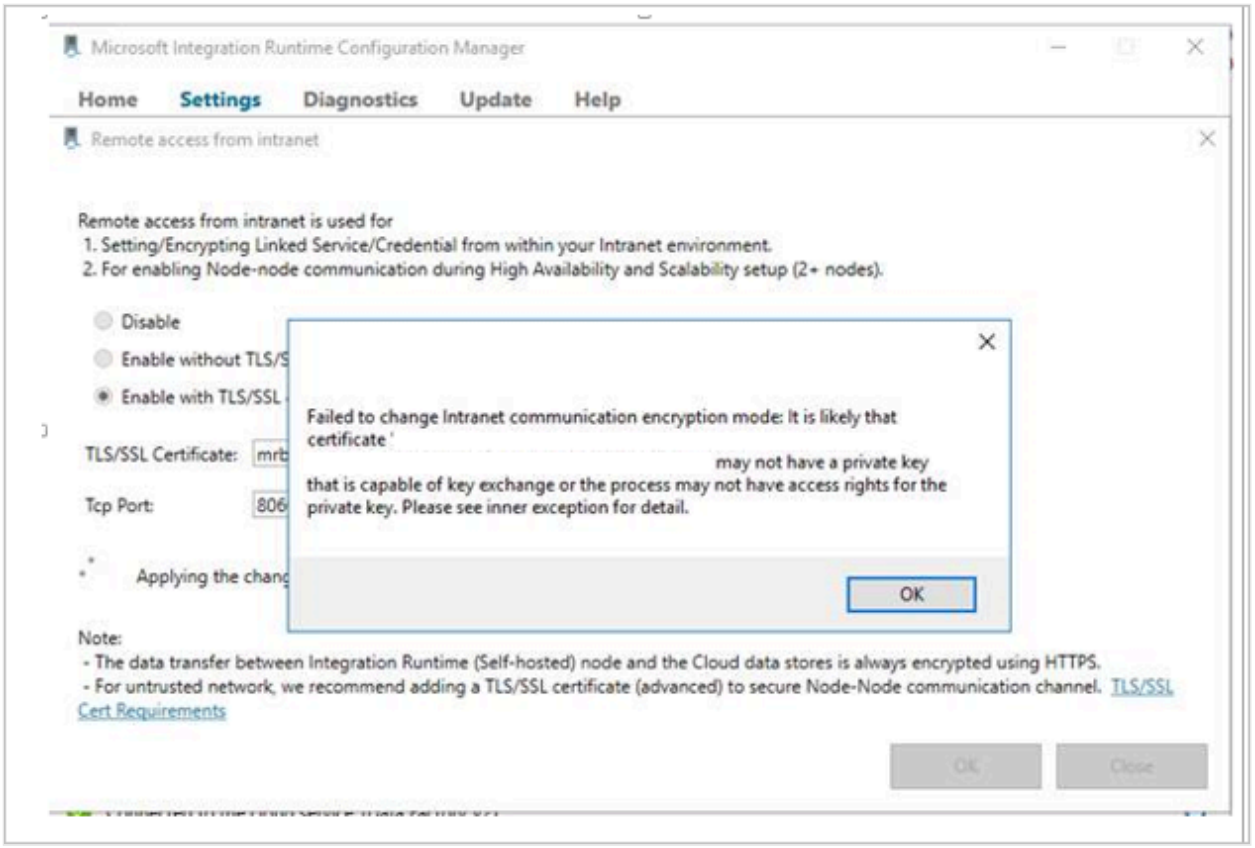


Can't choose the certificate because the private key is missing

Symptoms

- You've imported a PFX file to the certificate store.
- When you selected the certificate through the IR Configuration Manager UI, you received the following error message:

"Failed to change intranet communication encryption mode. It is likely that certificate '<*certificate name*>' might not have a private key that is capable of key exchange or the process might not have access rights for the private key. Please see inner exception for detail."



Cause

- The user account has a low privilege level and can't access the private key.
- The certificate was generated as a signature but not as a key exchange.

Resolution

- To operate the UI, use an account with appropriate privileges for accessing the private key.
- Import the certificate by running the following command:

```
certutil -importpfx FILENAME.pfx AT_KEYEXCHANGE
```

UserErrorJreNotFound error message when you run a scan

Symptoms

When you try to scan a data source using a Java-based tool or program (for example, Parquet format files), you receive an error message that resembles the following:

> ErrorCode=UserErrorJreNotFound,'Type=Microsoft.DataTransfer.Common.Shared.HybridDeliveryException,Message=Java Runtime Environment is not found. Go to `http://go.microsoft.com/fwlink/?LinkId=808605` to download and install on your Integration Runtime (Self-hosted) node machine. Note 64-bit Integration Runtime requires 64-bit JRE and 32-bit Integration Runtime requires 32-bit JRE.,Source=Microsoft.DataTransfer.Common,''Type=System.DllNotFoundException,Message=Unable to load DLL 'jvm.dll': The specified module could not be found. (Exception from HRESULT: 0x8007007E),Source=Microsoft.DataTransfer.Richfile.HiveOrcBridge

Cause

This issue occurs for either of the following reasons:

- Java Runtime Environment (JRE) isn't installed correctly on your Integration Runtime server.
- Your Integration Runtime server lacks the required dependency for JRE.

By default, Integration Runtime resolves the JRE path by using registry entries. Those entries should be automatically set during JRE installation.

Resolution

Follow the steps in this section carefully. Serious problems might occur if you modify the registry incorrectly. Before you modify it, back up the registry for restoration in case problems occur.

To fix this issue, follow these steps to verify the status of the JRE installation:

1. Make sure that Integration Runtime (Diahost.exe) and JRE are installed on the same platform. Check the following conditions:
    - 64-bit JRE for 64-bit ADF Integration Runtime should be installed in the folder: `C:\Program Files\Java\`

      > **Note**
      >
      > The folder is not `C:\Program Files (x86)\Java\`

    - JRE 11 is compatible for scanning. JRE 8 is also compatible, but versions that are earlier than JRE 8 have not been validated for this use.

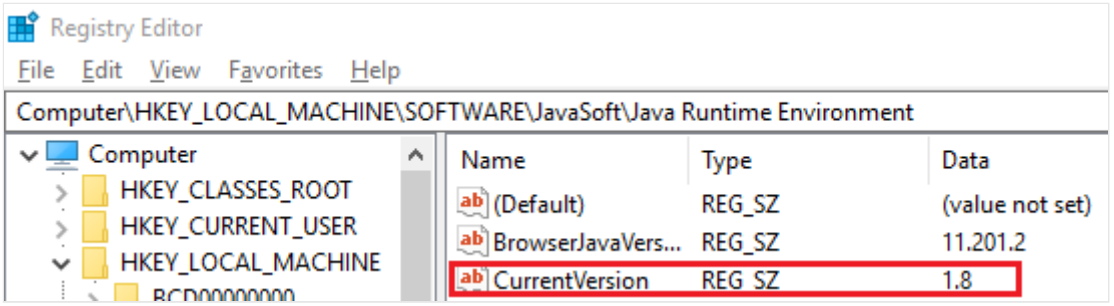
2. Check the registry for the appropriate settings. To do this, follow these steps:

3. In the **Run** menu, type **Regedit**, and then press Enter.

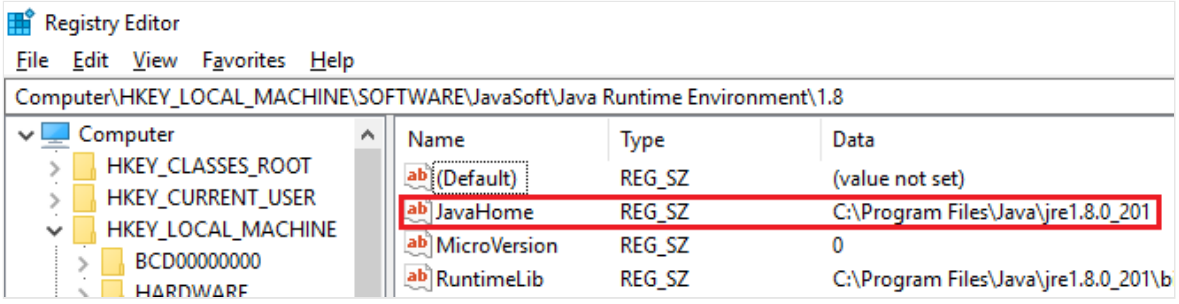
4. In the navigation pane, locate the following subkey:

   `HKEY_LOCAL_MACHINE\SOFTWARE\JavaSoft\Java Runtime Environment`.

   In the **Details** pane, there should be a Current Version entry that shows the JRE version (for example, 1.8).

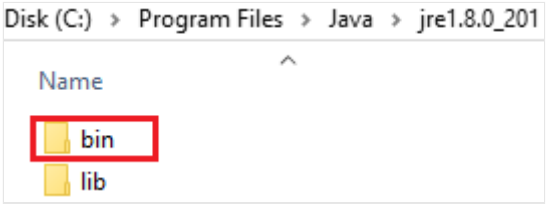


5. In the navigation pane, locate a subkey that is an exact match for the version (for example 1.8) under the JRE folder. In the details pane, there should be a **JavaHome** entry. The value of this entry is the JRE installation path.



6. Locate the bin\server folder in the following path:

   `C:\Program Files\Java\jre1.xx.xxx`



7. Check whether this folder contains a jvm.dll file. If it does not, check for the file in the `bin\client` folder.



> **Note**
>
> - If any of these configurations are not as described in these steps, use the JRE windows installer to fix the problems.
> - If all the configurations in these steps are correct as described, there might be a VC++ runtime library missing in the system. You can fix this problem by installing the VC++ 2010 Redistributable Package.

Self-hosted IR setup

Integration runtime registration error

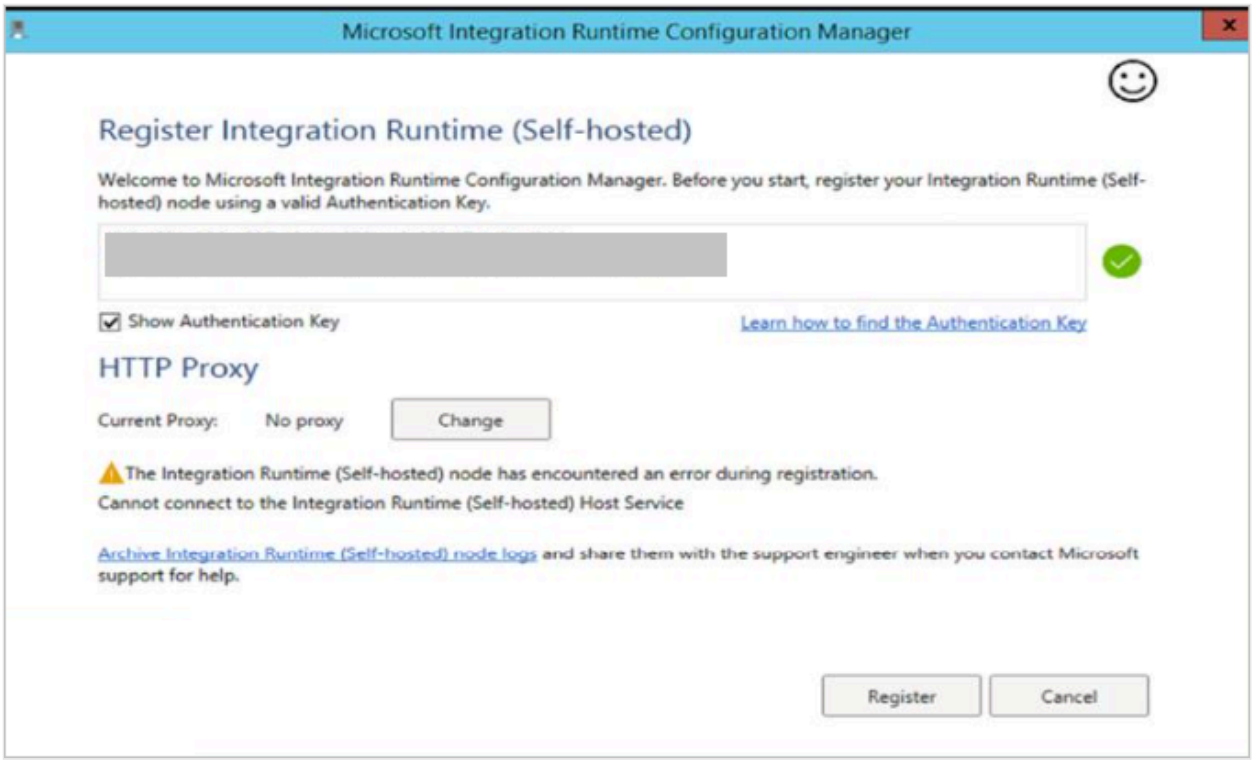
Symptoms

You might occasionally want to run a self-hosted IR in a different account for either of the following reasons:

- Company policy disallows the service account.
- Some authentication is required.

After you change the service account on the service pane, you might find that the integration runtime stops working, and you get the following error message:

"The Integration Runtime (Self-hosted) node has encountered an error during registration. Cannot connect to the Integration Runtime (Self-hosted) Host Service."

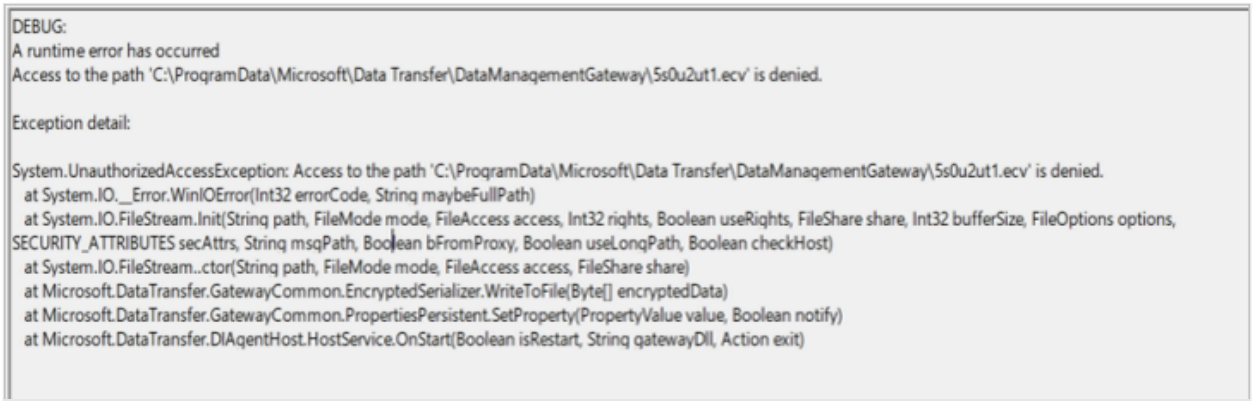


Cause

Many resources are granted only to the service account. When you change the service account to another account, the permissions of all dependent resources remain unchanged.

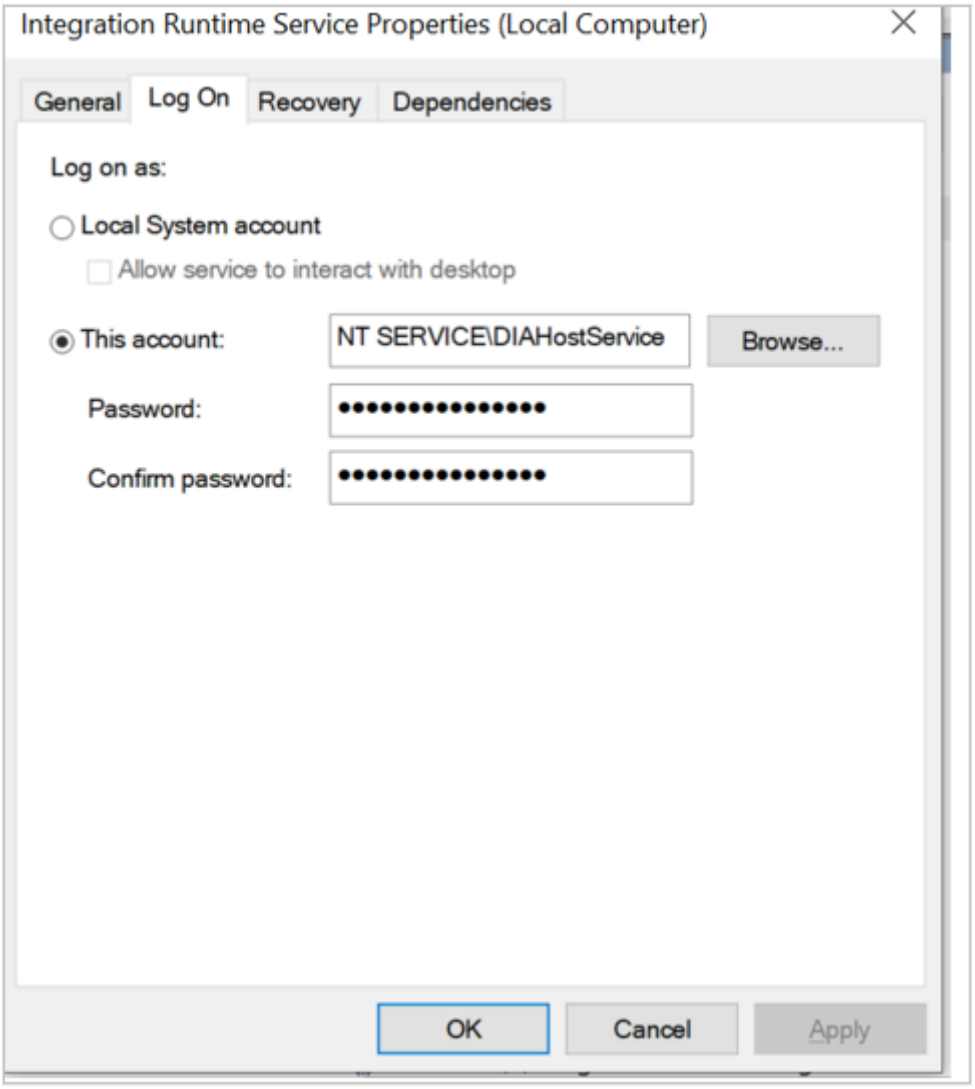
Resolution

Go to the integration runtime event log to check the error.



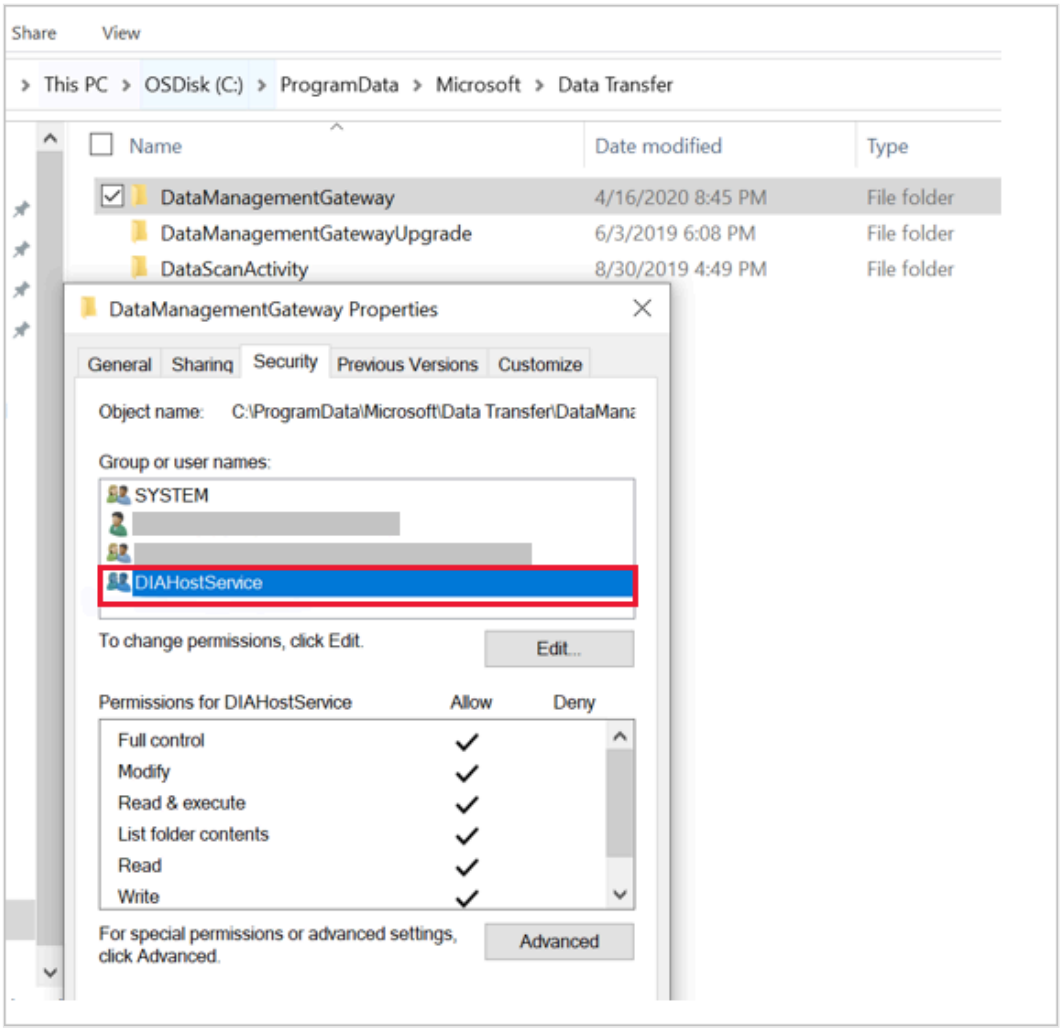
- If the error in the event log is "UnauthorizedAccessException," do the following:

    1. Check the *DIAHostService* sign-in service account in the Windows service panel.



2. Check to see whether the sign-in service account has read/write permissions for the *%programdata%\Microsoft\DataTransfer\DataManagementGateway* folder.

    - By default, if the service sign-in account hasn't been changed, it should have read/write permissions.

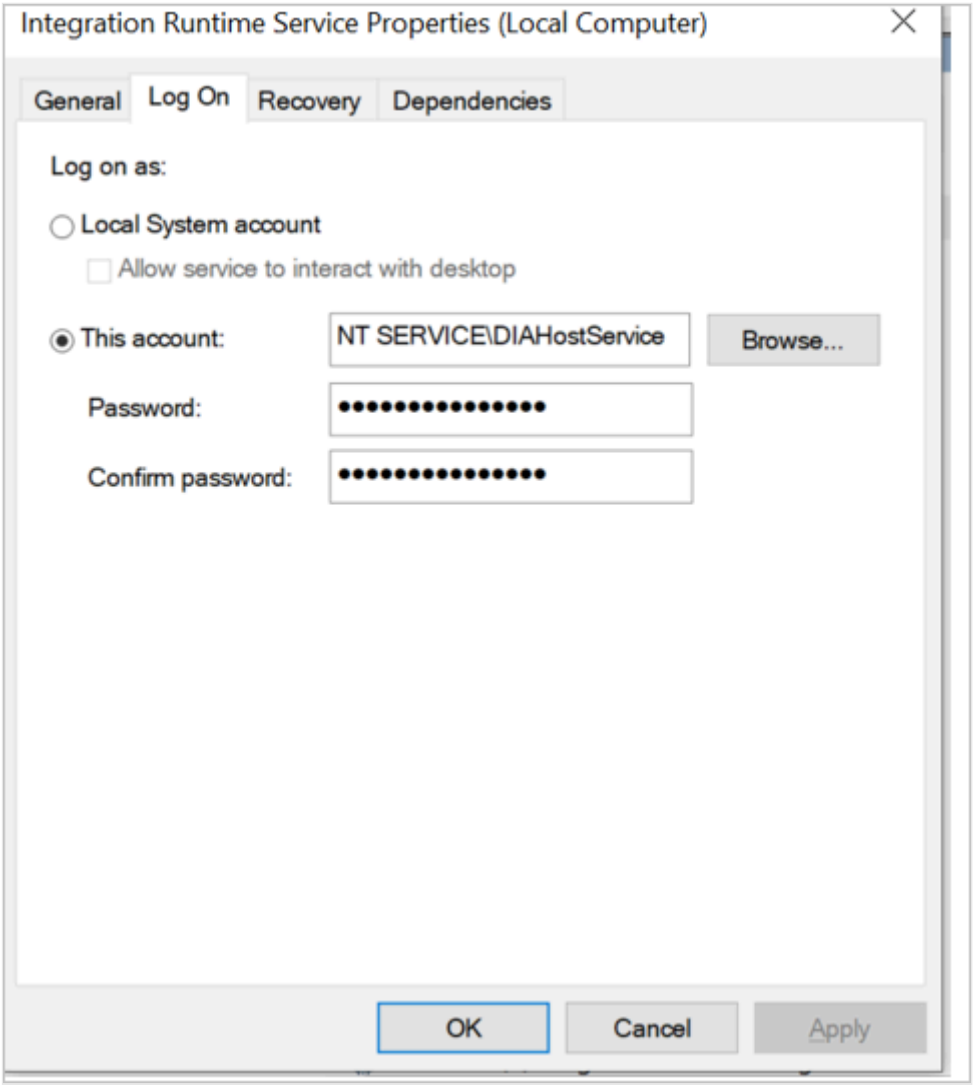


- If you've changed the service sign-in account, mitigate the issue by doing the following:
  a. Perform a clean uninstallation of the current self-hosted IR.
  b. Install the self-hosted IR bits.
  c. Change the service account by doing the following:
    i. Go to the self-hosted IR installation folder, and then switch to the *Microsoft Integration Runtime\4.0\Shared* folder.
    ii. Open a Command Prompt window by using elevated privileges. Replace <*user*> and <*password*> with your own username and password, and then run the following command: `dmgcmd.exe -SwitchServiceAccount "<user>" "<password>"`
    iii. If you want to change to the LocalSystem account, be sure to use the correct format for this account: `dmgcmd.exe -SwitchServiceAccount "NT Authority\System" ""`
    iv. Do *not* use this format: `dmgcmd.exe -SwitchServiceAccount "LocalSystem" ""`
    v. Optionally, because Local System has higher privileges than Administrator, you can also directly change it in "Services".
    vi. You can use a local/domain user for the IR service sign-in account.

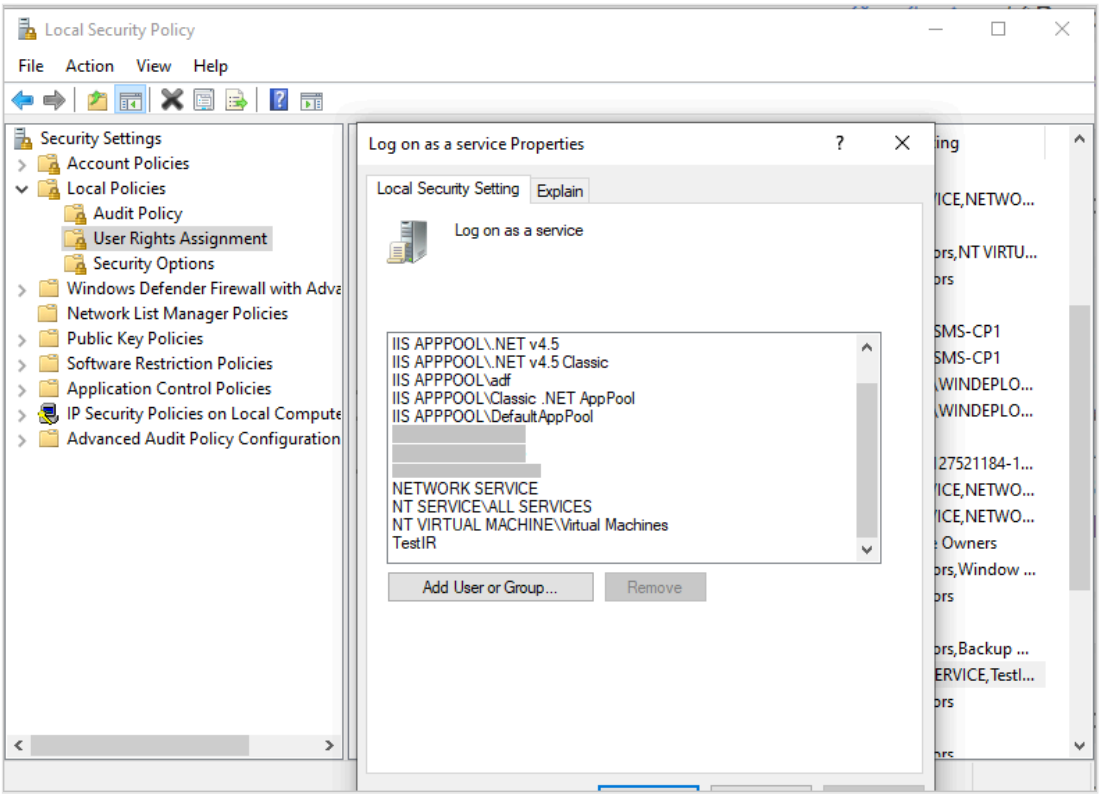
d. Register the integration runtime.

- If the error is "Service 'Integration Runtime Service' (DIAHostService) failed to start. Verify that you have sufficient privileges to start system services," do the following:

    1. Check the *DIAHostService* sign-in service account in the Windows service panel.



    2. Check to see whether the sign-in service account has **Log on as a service** permission to start the Windows service:



More information

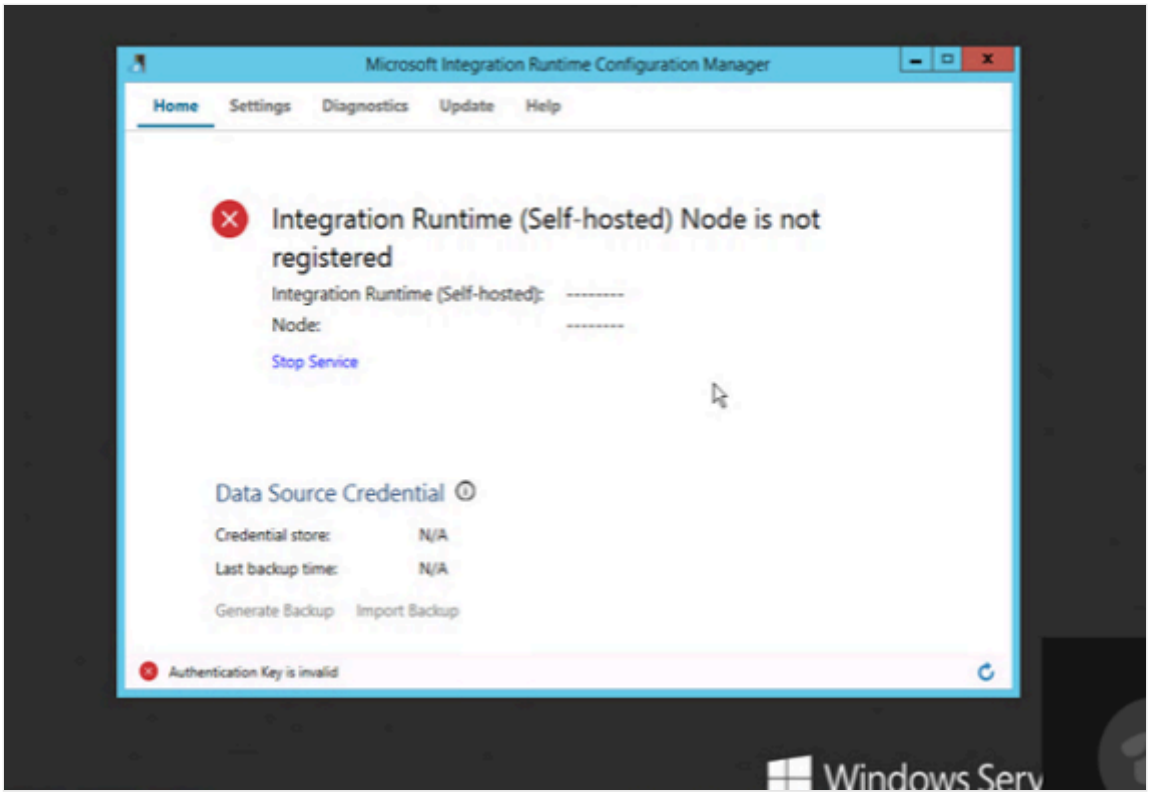
If neither of the preceding two resolution patterns applies in your case, try to collect the following Windows event logs:

- Applications and Services Logs > Integration Runtime
- Windows Logs > Application

Can't find the Register button to register a self-hosted IR

Symptoms

When you register a self-hosted IR, the **Register** button isn't displayed on the Configuration Manager pane.



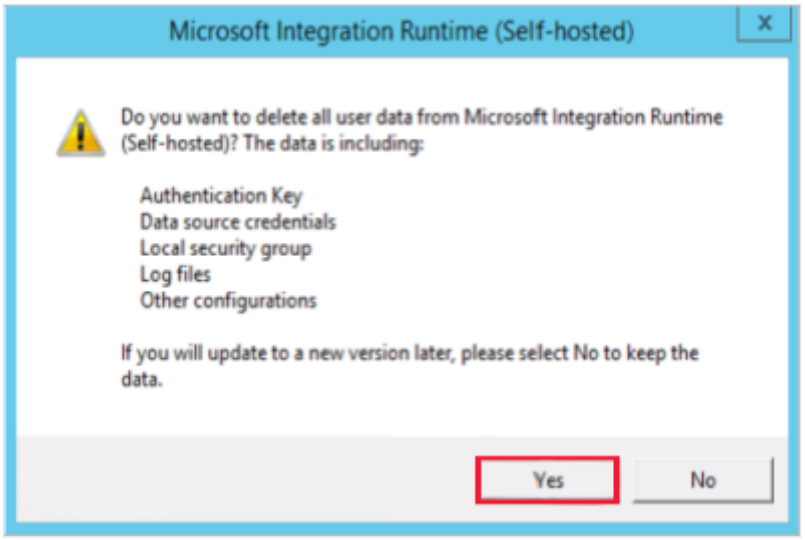
Cause

As of the release of Integration Runtime 3.0, the **Register** button on existing integration runtime nodes has been removed to enable a cleaner and more secure environment. If a node has been registered to an integration runtime, whether it's online or not, re-register it to another integration runtime by uninstalling the previous node, and then install and register the node.

Resolution

1. In Control Panel, uninstall the existing integration runtime.

   > ⓘ **Important**
   >
   > In the following process, select **Yes**. Do not keep data during the uninstallation process.



2. If you don't have the integration runtime installer MSI file, go to download center to download the latest integration runtime.

3. Install the MSI file, and register the integration runtime.

Unable to register the self-hosted IR because of localhost

Symptoms

You're unable to register the self-hosted IR on a new machine when you use get_LoopbackIpOrName.

Debug: A runtime error has occurred. The type initializer for 'Microsoft.DataTransfer.DIAgentHost.DataSourceCache' threw an exception. A nonrecoverable error occurred during a database lookup.

Exception detail: System.TypeInitializationException: The type initializer for 'Microsoft.DataTransfer.DIAgentHost.DataSourceCache' threw an exception. ---> System.Net.Sockets.SocketException: A nonrecoverable error occurred during a database lookup at System.Net.Dns.GetAddrInfo(String name).

Cause

The issue usually occurs when the localhost is being resolved.

Resolution

Use localhost IP address 127.0.0.1 to host the file and resolve the issue.

Self-hosted setup failed

Symptoms

You're unable to uninstall an existing IR, install a new IR, or upgrade an existing IR to a new IR.

Cause

The integration runtime installation depends on the Windows Installer service. You might experience installation problems for the following reasons:

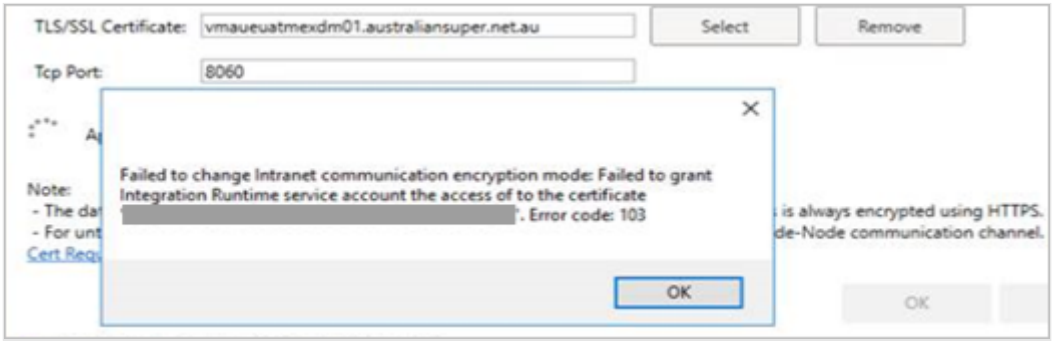
- Insufficient available disk space.
- Lack of permissions.
- The Windows NT service is locked.
- CPU utilization is too high.
- The MSI file is hosted in a slow network location.
- Some system files or registries were touched unintentionally.

The IR service account failed to fetch certificate access

Symptoms

When you install a self-hosted IR via Microsoft Integration Runtime Configuration Manager, a certificate with a trusted certificate authority (CA) is generated. The certificate couldn't be applied to encrypt communication between two nodes, and the following error message is displayed:

"Failed to change Intranet communication encryption mode: Failed to grant Integration Runtime service account the access of to the certificate '<*certificate name*>'. Error code 103"



Cause

The certificate is using key storage provider (KSP) storage, which is not supported yet. To date, self-hosted IR supports only cryptographic service provider (CSP) storage.

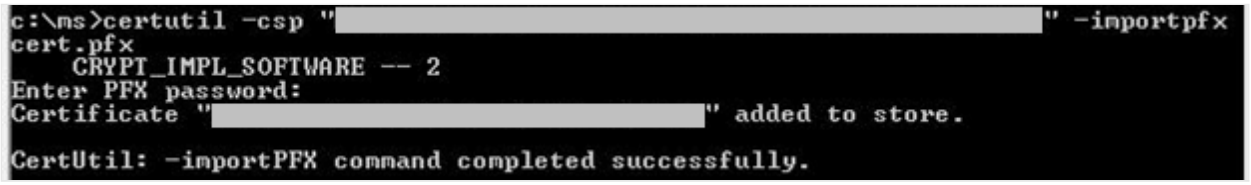
Resolution

We recommend that you use CSP certificates in this case.

Solution 1

To import the certificate, run the following command:

```
Certutil.exe -CSP "CSP or KSP" -ImportPFX FILENAME.pfx
```



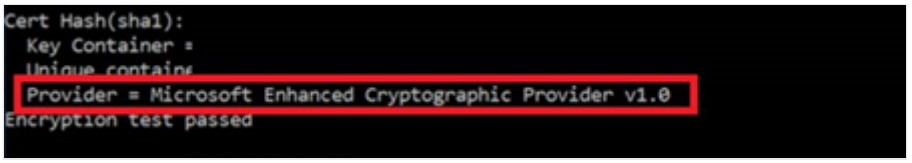
Solution 2

To convert the certificate, run the following commands:

```
openssl pkcs12 -in .\xxxx.pfx -out .\xxxx_new.pem -password pass: <EnterPassword>
openssl pkcs12 -export -in .\xxxx_new.pem -out xxxx_new.pfx
```

Before and after conversion:





Self-hosted integration runtime version 5.x

For the upgrade to version 5.x of the self-hosted integration runtime, we require **.NET Framework Runtime 4.7.2** or later. On the download page, you'll find download links for the latest 4.x version and the latest two 5.x versions.

For Azure Data Factory v2 and Azure Synapse customers:

- If automatic update is on and you've already upgraded your .NET Framework Runtime to 4.7.2 or later, the self-hosted integration runtime will be automatically upgraded to the latest 5.x version.
- If automatic update is on and you haven't upgraded your .NET Framework Runtime to 4.7.2 or later, the self-hosted integration runtime won't be automatically upgraded to the latest 5.x version. The self-hosted integration runtime will stay in the current 4.x version. You can see a warning for a .NET Framework Runtime upgrade in the portal and the self-hosted integration runtime client.
- If automatic update is off and you've already upgraded your .NET Framework Runtime to 4.7.2 or later, you can manually download the latest 5.x and install it on your machine.
- If automatic update is off and you haven't upgraded your .NET Framework Runtime to 4.7.2 or later. When you try to manually install self-hosted integration runtime 5.x and register the key, you will be required to upgrade your .NET Framework Runtime version first.

For Azure Data Factory v1 customers:

- Self-hosted integration runtime 5.X doesn't support Azure Data Factory v1.
- The self-hosted integration runtime will be automatically upgraded to the latest version of 4.x. And the latest version of 4.x won't expire.
- If you try to manually install self-hosted integration runtime 5.x and register the key, you'll be notified that self-hosted integration runtime 5.x doesn't support Azure Data Factory v1.

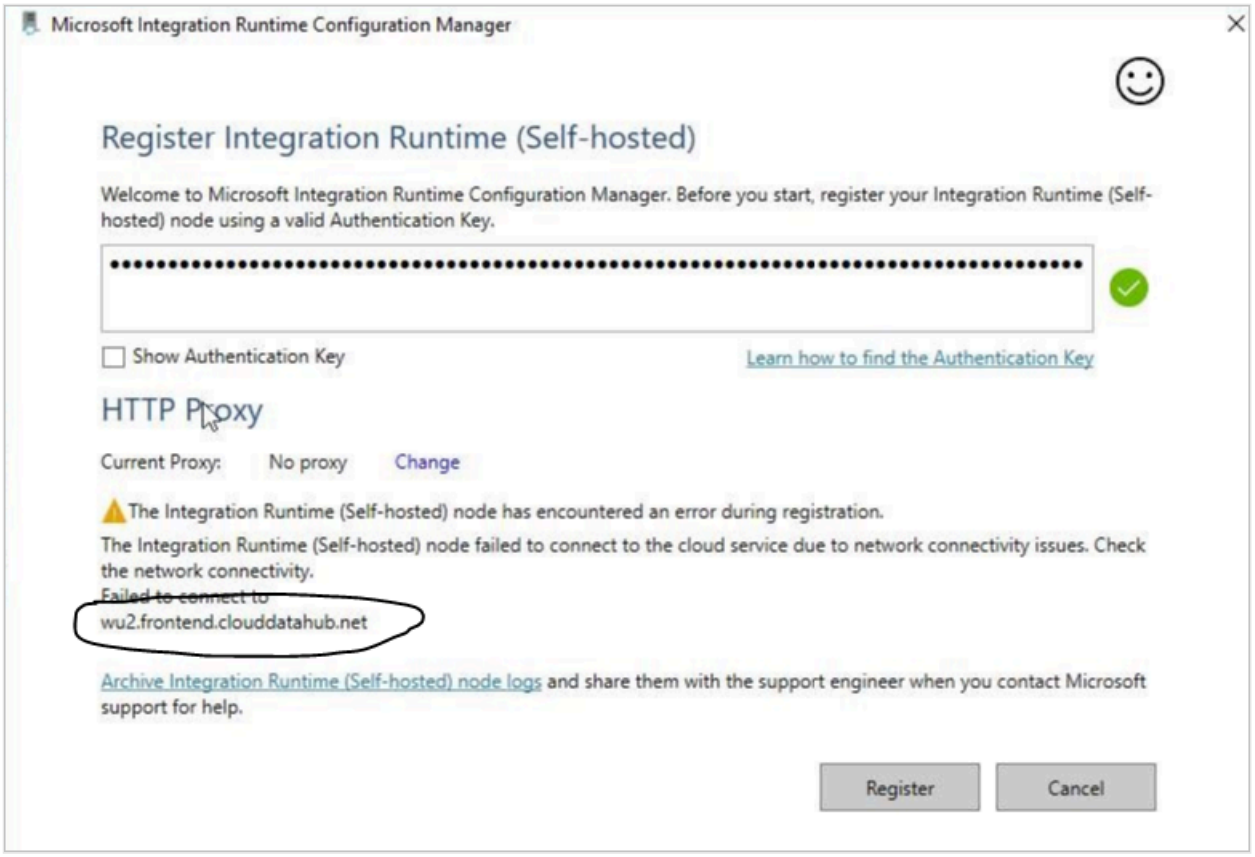
Self-hosted IR connectivity issues

Self-hosted integration runtime can't connect to the cloud service

Symptoms

When you attempt to register the self-hosted integration runtime, Configuration Manager displays the following error message:

"The Integration Runtime (Self-hosted) node has encountered an error during registration."

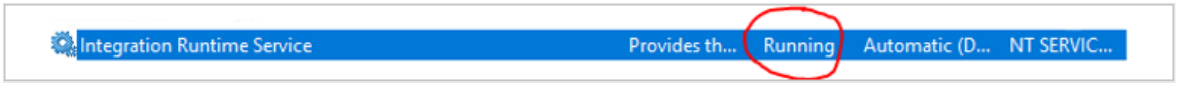


Cause

The self-hosted IR can't connect to the service back end. This issue is usually caused by network settings in the firewall.

Resolution

1. Check to see whether the integration runtime service is running. If it is, go to step 2.



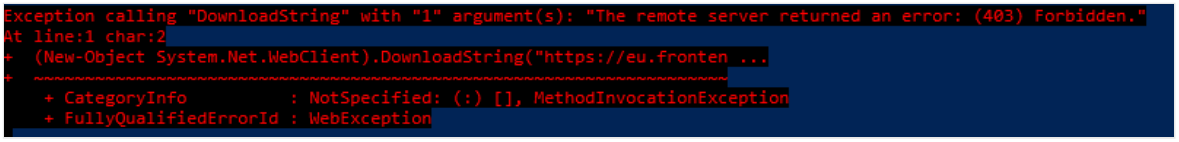
2. If no proxy is configured on the self-hosted IR, which is the default setting, run the following PowerShell command on the machine where the self-hosted integration runtime is installed:

```PowerShell
 (New-Object System.Net.WebClient).DownloadString("https://wu2.frontend.clouddatahub.net/")
```

> **Note**

> The service URL might vary, depending on the location of your data factory or Synapse workspace instance. To find the service URL, use the Manage page of the UI in your data factory or Azure Synapse instance to find **Integration runtimes** and select your self-hosted IR to edit it. There select the **Nodes** tab and select **View Service URLs**.

3. The following is the expected response:



4. If you don't receive the response you had expected, use one of the following methods, as appropriate:

   - If you receive a "Remote name could not be resolved" message, there's a Domain Name System (DNS) issue. Contact your network team to fix the issue.
   - If you received an "ssl/tls cert is not trusted" message, check the certificate (`https://wu2.frontend.clouddatahub.net/`) to see whether it's trusted on the machine, and then install the public certificate by using Certificate Manager. This action should mitigate the issue.
   - Go to **Windows** > **Event viewer (logs)** > **Applications and Services Logs** > **Integration Runtime**, and check for any failure that's caused by DNS, a firewall rule, or company network settings. If you find such a failure, forcibly close the connection. Because every company has its own customized network settings, contact your network team to troubleshoot these issues.

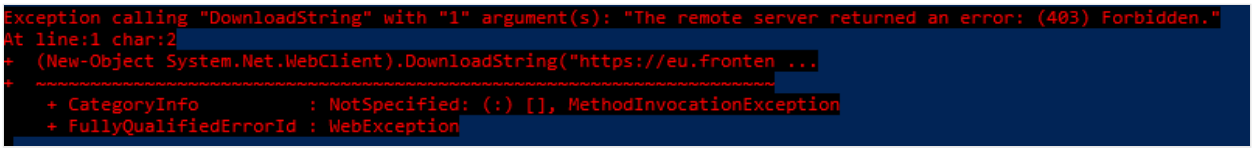
5. If "proxy" has been configured on the self-hosted integration runtime, verify that your proxy server can access the service endpoint. For a sample command, see PowerShell, web requests, and proxies.

PowerShell

```
$user = $env:username
$webproxy = (get-itemproperty
'HKCU:\Software\Microsoft\Windows\CurrentVersion\Internet
Settings').ProxyServer
$pwd = Read-Host "Password?" -assecurestring
$proxy = new-object System.Net.WebProxy
$proxy.Address = $webproxy
$account = new-object System.Net.NetworkCredential($user,
[Runtime.InteropServices.Marshal]::PtrToStringAuto([Runtime.InteropServ
ices.Marshal]::SecureStringToBSTR($pwd)), "")
$proxy.credentials = $account
```

```
$url = "https://wu2.frontend.clouddatahub.net/"
$wc = new-object system.net.WebClient
$wc.proxy = $proxy
$webpage = $wc.DownloadData($url)
$string = [System.Text.Encoding]::ASCII.GetString($webpage)
$string
```

The following is the expected response:



> ⓘ Note
>
> Proxy considerations:
>
> - Check to see whether the proxy server needs to be put on the Safe Recipients list. If so, make sure these domains are on the Safe Recipients list.
> - Check to see whether SSL/TLS certificate "wu2.frontend.clouddatahub.net/" is trusted on the proxy server.
> - If you're using Active Directory authentication on the proxy, change the service account to the user account that can access the proxy as "Integration Runtime Service."

Couldn't establish a trust relationship for the SSL/TLS secure channel

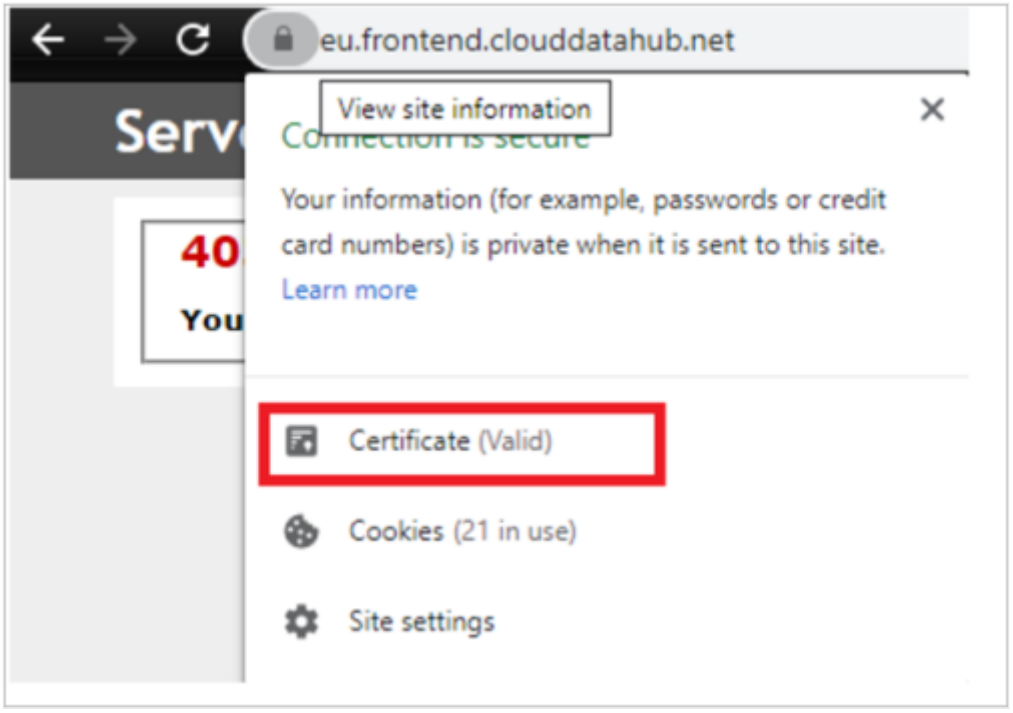
Symptoms

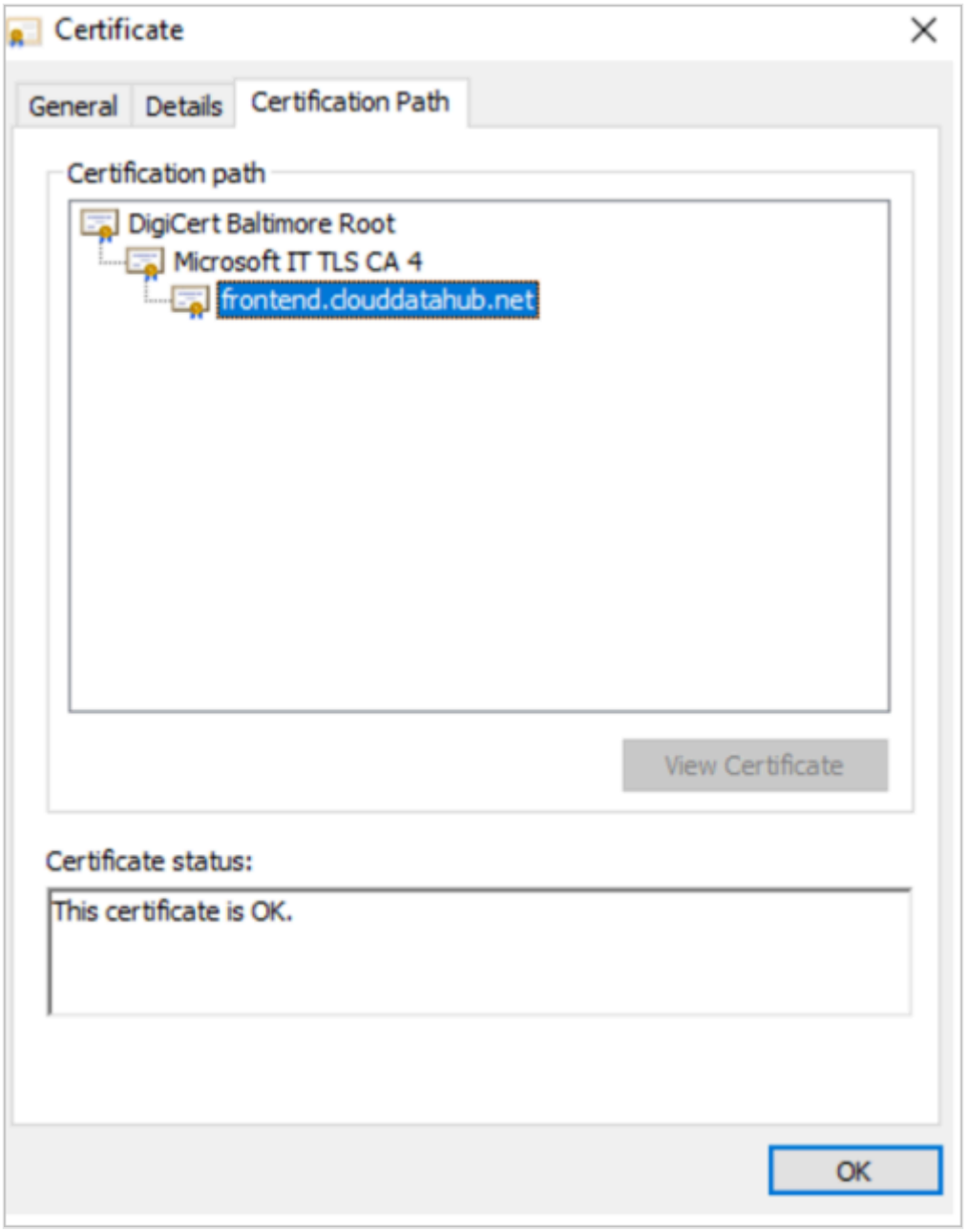
The self-hosted IR couldn't connect to Microsoft Purview.

When you check the self-hosted IR event log after going to **Windows** > **Event viewer (logs)** > **Applications and Services Logs** > **Integration Runtime**, you'll find the following error message:

"The underlying connection was closed: Could not establish trust relationship for the SSL/TLS secure channel. The remote certificate is invalid according to the validation procedure."

The simplest way to check the server certificate of the service is to open the service URL in your browser. For example, open the check server certificate link (`(https://eu.frontend.clouddatahub.net/`) on the machine where the self-hosted IR is installed, and then view the server certificate information.





Cause

There are two possible reasons for this issue:

- Reason 1: The root CA of the service's server certificate isn't trusted on the machine where the self-hosted IR is installed.
- Reason 2: You're using a proxy in your environment, the server certificate of the service is replaced by the proxy, and the replaced server certificate isn't trusted by the machine where the self-hosted IR is installed.

Resolution

- For reason 1: Make sure that the service's server certificate and its certificate chain are trusted by the machine where the self-hosted IR is installed.

- For reason 2: Either trust the replaced root CA on the self-hosted IR machine, or configure the proxy not to replace the service's server certificate.

For more information about trusting certificates on Windows, see Installing the trusted root certificate.

More information

We've rolled out a new SSL certificate, which is signed from DigiCert. Check to see whether the DigiCert Global Root G2 is in the trusted root CA.



If it isn't in the trusted root CA, download it here.

Next Steps

For more help with troubleshooting, try the following resources:

- Getting started with Microsoft Purview
- Create and Manage SHIR Self-hosted integration runtimes in Purview
- Data governance frequently asked questions

Feedback

Was this page helpful? Yes No

Create and manage a Kubernetes supported self-hosted integration runtime

07/15/2025

This article covers the details for the new SHIR feature that is Kubernetes-based for Linux that improves the underlying infrastructure to provide several benefits:

- **Scalability**: Ability to scale to hundreds of machines.
- **Performance**: Improved performance in scanning workloads.
- **Security** (containerized): Ability to have containerized security on a Kubernetes cluster, instead of hosting SHIR on a Windows machine directly

This article covers the details to install and manage a Kubernetes supported self-hosted integration runtime.

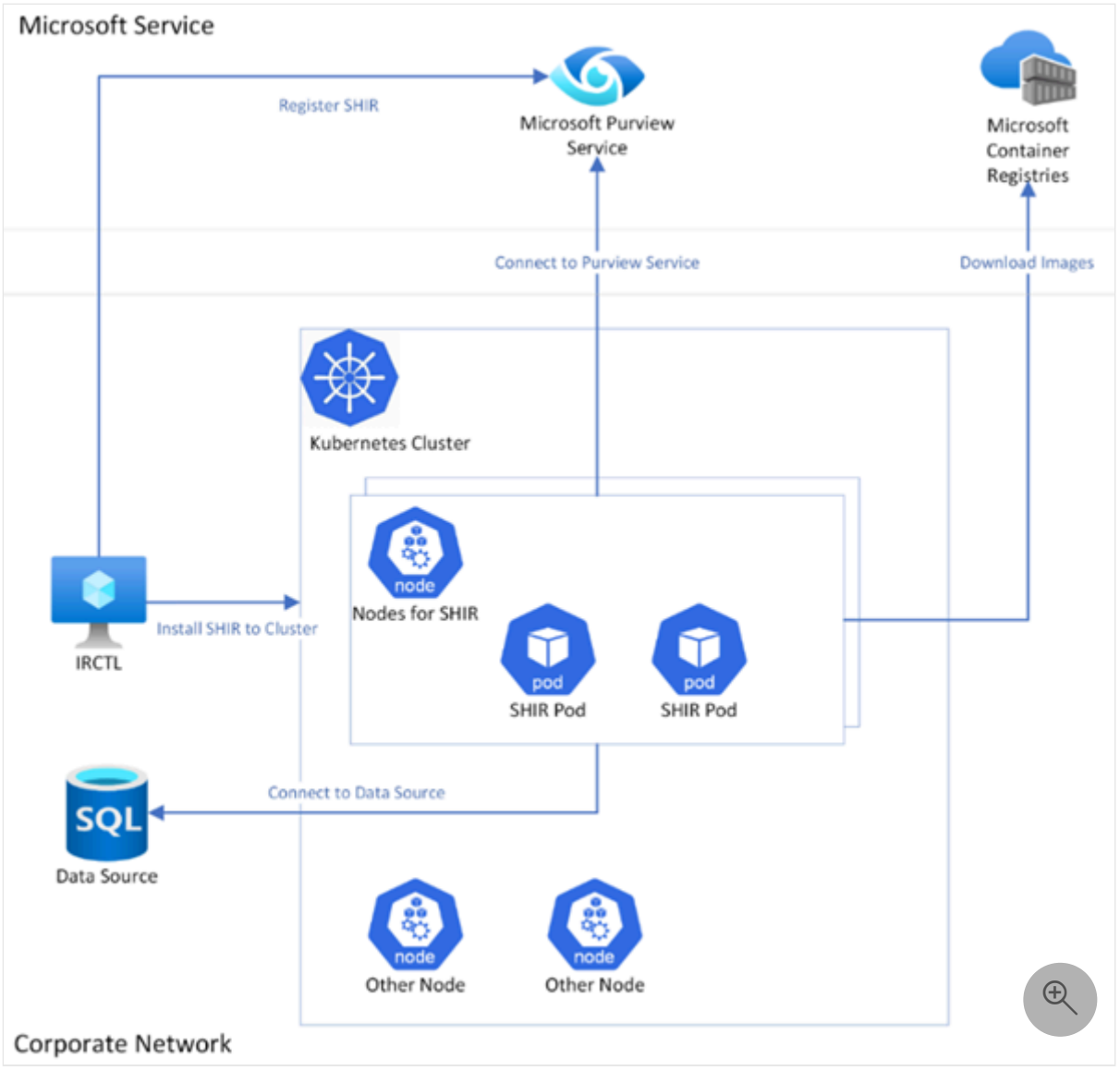
Supported data sources

For a list of all supported sources, see the supported data sources for each integration runtime table.

Architecture

At a high-level architectural view, when a Kubernetes based SHIR is installed, several pods get autocreated on the nodes of users' Kubernetes cluster. This installation can be triggered by a command line tool named IRCTL (more detail in following sections). IRCTL connects to the Microsoft Purview Service to register the SHIR and connect to the Kubernetes cluster to install the SHIR.

During the installation, SHIR images are downloaded from MCR (Microsoft Container Registries) to the SHIR pods. After installation is done, the pods in users' cluster will connect to the Microsoft Purview Service to pull scan jobs. As a scan job is pulled, it can connect users' on-premises Data Source for Data Scanning.



Prerequisites

- A Microsoft Purview account using enterprise data governance solutions.
- **Kubernetes cluster**: You need to have an existing Linux-based Kubernetes cluster or to prepare one. The nodes can be identified by node selector, which follows the definition of Kubernetes node selector. Minimum configuration:
  - Container type: Linux
  - Kubernetes version: 1.24.9 or above
  - Node OS: Linux based OS running on x86 architecture
  - Node spec: minimal eight cores CPU, 32-GB memory, and at least 80 GB of available hard disk space
  - Node count: >=1 (should be fixed, **not** enable cluster auto scaler)
  - Pod number per Node: >= 20 (max Pod number – count of other Pods not belonging to Self-Hosted IR)

> ⓘ **Note**

> The folder /var/irstorage/ of each Node is reserved for SHIR. It is readable and writable to SHIR. You can get logs being persisted from this folder or upload external drivers to this folder. It will be created by SHIR if it doesn't exist, and it will not be deleted after SHIR being deleted. The container images used by SHIR are managed by Kubernetes Garbage Collection, which won't be cleaned-up by SHIR. Configure the proper threshold for your Kubernetes cluster.

- **Kubernetes cluster network**: The Kubernetes cluster you have should be able to connect to the endpoint listed in networking requirements.
- **Integration runtime command line tool**: In order manage your Microsoft Purview Kubernetes SHIR locally, you need a command line tool named IRCTL. You can download this tool during the SHIR creation process. IRCTL is a command line tool to manage your Microsoft Purview SHIR. For more information, see the IRCTL documentation.
- **Kubernetes context**: Kubernetes context, which contains Kubernetes cluster information and user's permissions and credential for this cluster, is needed to talk to your Kubernetes cluster. To ease the configuration for the user's permissions for SHIR management, you can start with Kubernetes Admin role. This context is generated with the setup of your Kubernetes cluster and saved in a config file. Where and how you can get this file depends on your setting up the Kubernetes cluster.
  - If you use `kubeadm init` to set up the Kubernetes cluster, you can find the config file under `/etc/Kubernetes/admin.conf`.
  - If you use AKS, you can follow the guidance of AKS to use Az PowerShell module command to get credentials of this cluster to your local machine. The context can be merged to the config file under `$HOME/.kube/config` directly.
  - If you're using other tools setting up a Kubernetes cluster, refer to the Kubernetes documentation.
  - As you have the config file of the Kubernetes context, merge it to the config file, which is `$HOME/.kube/config`, on the machine you would like to run IRCTL command. Or you can set the config file of the Kubernetes context in an environment variable named KUBECONFIG as well. For more information about the Kubernetes context, see Configure Access to Multiple Clusters.

Create Kubernetes supported self-hosted integration runtime

To control and manage a Kubernetes SHIR, users can download a command line tool named IRCTL. The following are the steps to your Kubernetes supported self-hosted integration runtime.

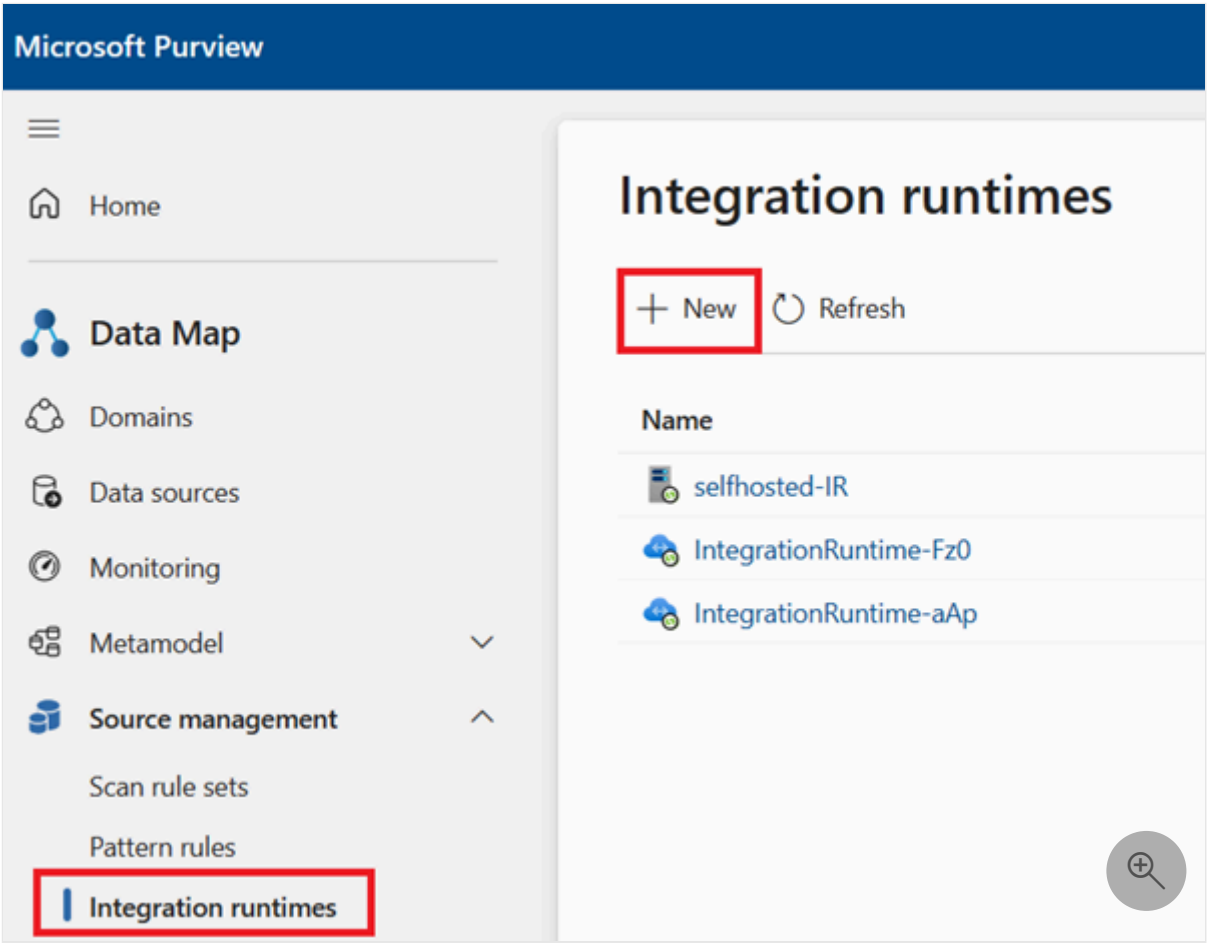
The steps take you through downloading IRCTL, but for direct links, see the IRCTL documentation.

Set up a Kubernetes supported self-hosted integration runtime

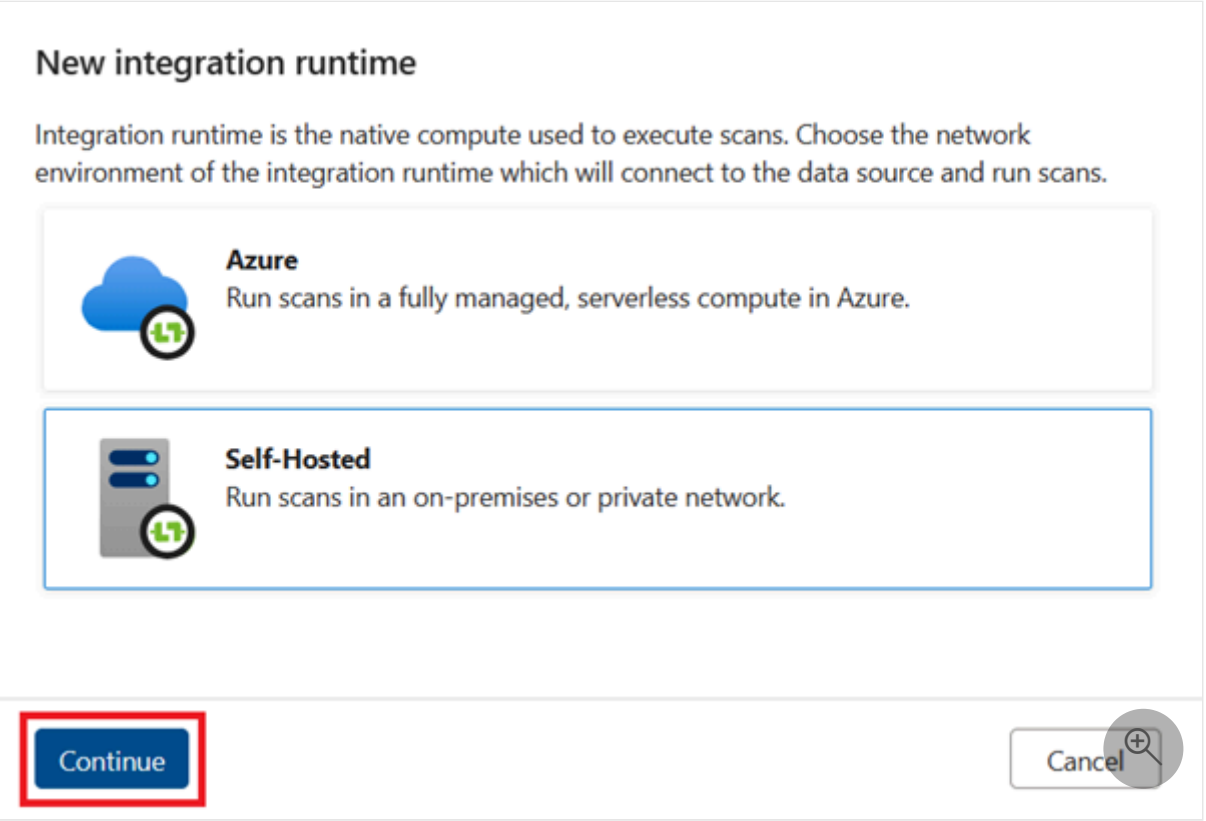
1. Open the **Integration runtimes** window in the Microsoft Purview Data Map

    - If you're using the new Microsoft Purview portal:
        a. Open the **Data Map**
        b. Select **Source management**
        c. Select **Integration runtimes**
    - If you're using the classic Microsoft Purview governance portal:
        a. Open the **Data Map**
        b. Select **Integration runtimes**

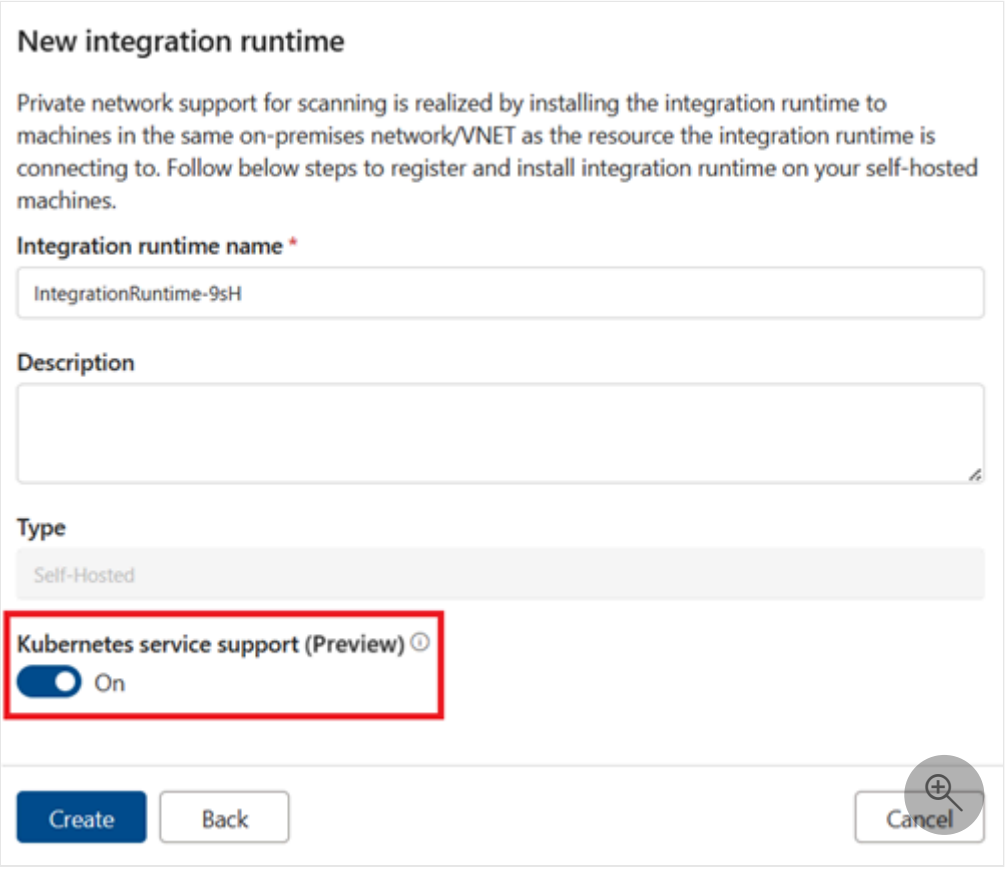
2. Select the + **New** button



3. Select **Self-hosted** and then select **Continue**

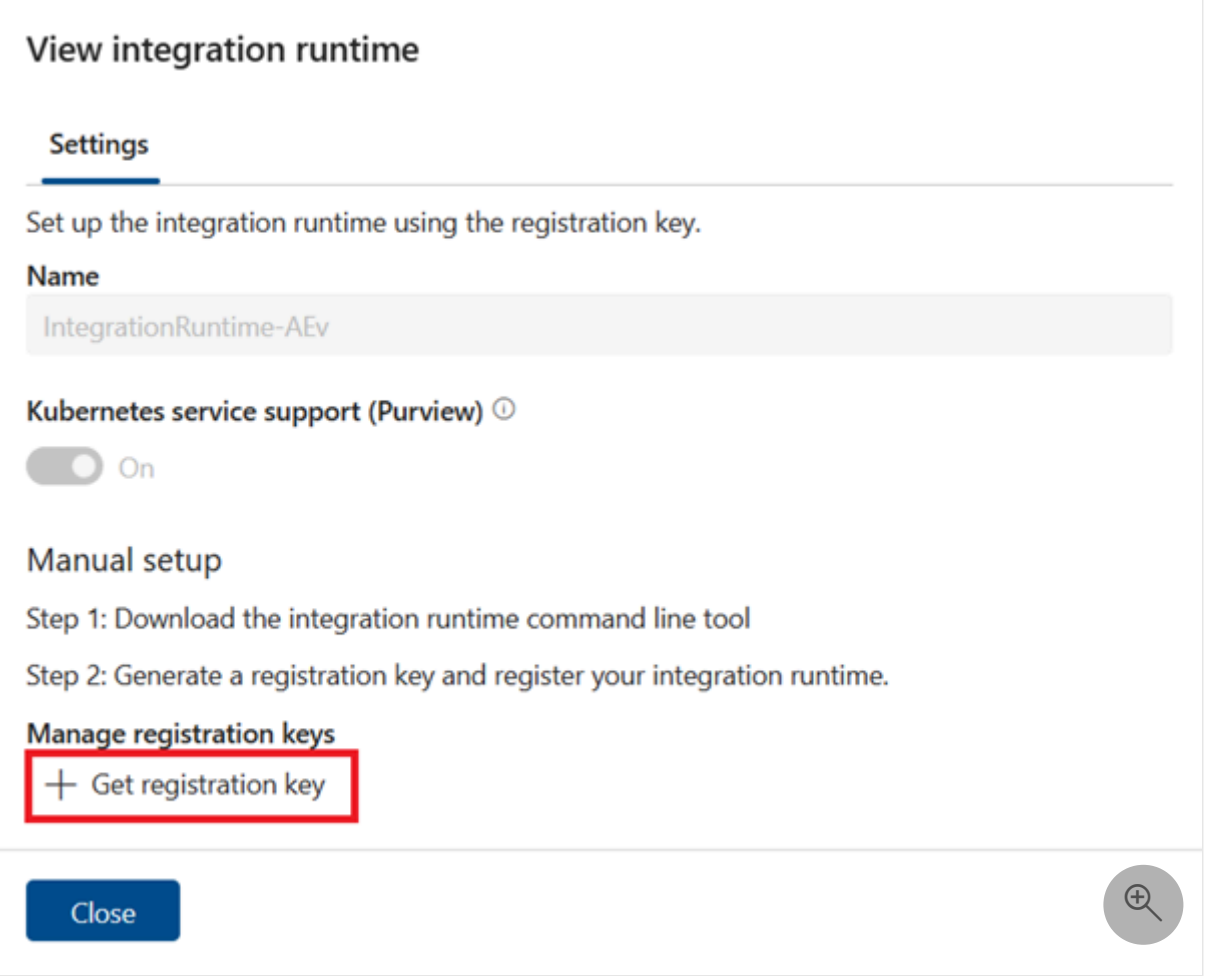


4. Give your runtime a name, then select the **Kubernetes service support** toggle to enable



5. Select **Create**

6. Select **Get registration key**



7. Copy the key value. You need it to run commands in IRCTL later.

> **Tip**
>
> If needed, you can regenerate a key or revoke a generated key.

8. Select the **Download IRCTL and install integration runtime** link to download the IRCTL tool. (You can also follow these steps to download IRCTL directly.)

9. On the machine where you want to run the IRCTL command line, install IRCTL from the download. IRCTL connects to your Kubernetes cluster by context of the Kube config. If context isn't specified, IRCTL uses the current context. You can set the context in one of two ways:

    - Run kubectl command line and execute this command to confirm the current context:

      Bash

      ```
      kubectl config get-contexts – List all contexts configured on the machine
      ```

```
Bash

kubectl config current-context – Get the current context name
```

```
Bash

kubectl config use-context <name of context>
```

- Run IRCTL and execute `--context` to specify the context in the Kube config

10. Run the IRCTL command line and execute this command with the registration key you copied.

```
Bash

./irctl create --registration-key <registration key copied from the portal>
```

> **Note**
>
> If node selector isn't specified, will use all nodes of the Kubernetes cluster. For AKS, we suggest using the label of AKS node pool as the node selector or you can customize different labels to the SHIR nodes.

11. You see this printout:

```
Bash

[Info] Start to create SHIR with Kubernetes context [your-context]......
[Info] Environment validation passed!
[Info] Registering SHIR[example-k8s-shir] for Microsoft Purview Account
[yourpurviewaccount]......
[Info] SHIR Registration done!
[Info] Provisioning SHIR, it may take about 5-30 minutes......done!
[Info] SHIR creation succeeded!
```

> **Tip**
>
> If the installation progress is broken by Ctrl-C or other reasons, the following command can be used to monitor the installation progress: `./irctl install status`

12. Once installation is complete, to check the current status of the SHIR, run this command:

```bash
./irctl describe
```

13. You can also check the status of your SHIR in the Microsoft Purview portal, on the **Integration runtimes** page.

Set up a scan with external drivers

When scanning some data sources, you need to install the corresponding driver on the machine where the SHIR is installed for Microsoft Purview to connect with the data source. Below is an example for Db2 scan. Refer to respective connector article for specific prerequisites.

> **Note**
>
> Data sources that need these external drivers have the information listed in their prerequisites.

In this example, we are installing the Db2 driver. Steps for other drivers are similar.

1. First, install the integration runtime.
2. Download the driver (each source has their individual driver listed.) For example, you can find the DB2 driver here: Connect to and manage Db2.
3. Upload the driver to each node for your integration runtime. You can use a command like this:

   ```bash
   ./irctl storage upload --source jdbc_sqlj/db2_driver --destination driver/db2
   ```

   A successful upload confirmation looks like this:

   ```bash
   ========== Context ==========
   Kubernetes Context             : k8s-shir-test-cluster
   Purview Account                : test-purview-1
   Self-hosted Intrgration Runtime: k8s-shir-demo
   ========== Progress ==========
   Processing 2/2 nodes...
   aks-shirpool-27141791-vmss000000: SUCCEEDED
   ```

```
aks-shirpool-27141791-vmss000001: SUCCEEDED
========== Results ==========
jdbc_sqlj/db2_driver -> /var/irstorage/driver/db2
```

> **Note**
>
> If you replace nodes or scale out to new nodes, you need to upload the external driver again.

4. Verify the files uploaded with this command:

```
Bash

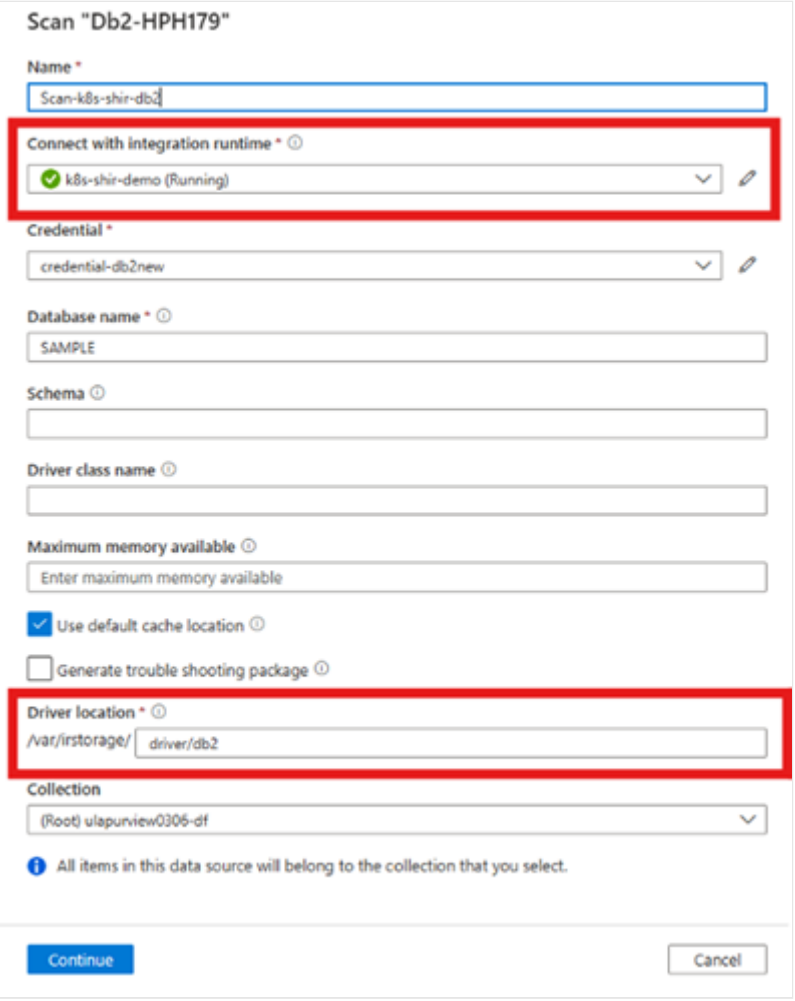
./irctl storage list driver/db2
```

You should see a response like this:

```
Bash

========== Context ==========
Kubernetes Context             : k8s-shir-test-cluster
Purview Account                : test-purview-1
Self-hosted Intrgration Runtime: k8s-shir-demo
========== Progress ==========
Processing 2/2 nodes...
aks-shirpool-27141791-vmss000000: SUCCEEDED
aks-shirpool-27141791-vmss000001: SUCCEEDED
========== Results ==========
Node: aks-shirpool-27141791-vmss000000 - Succeeded
/var/irstorage/driver/db2
total 9364
drwxr-xr-x    2 root     root          4096 May 15 14:23 .
drwxr-xr-x    3 root     root          4096 May 15 14:23 ..
-rwxrwxr-x    1 root     root       6568346 May 15 14:23 db2jcc4.jar
Node: aks-shirpool-27141791-vmss000001 - Succeeded
/var/irstorage/driver/db2
total 9364
drwxr-xr-x    2 root     root          4096 May 15 14:23 .
drwxr-xr-x    3 root     root          4096 May 15 14:23 ..
-rwxrwxr-x    1 root     root       6568346 May 15 14:23 db2jcc4.jar
```

5. Create scan with the value for **DriverLocation** with the **Destination** value from step 3.



High availability and scalability

You can assign multiple nodes of the Kubernetes cluster to have high availability by using the node-selector during the Kubernetes supported self-hosted integration runtime installation. The benefits of having multiple nodes are:

- Higher availability of the self-hosted integration runtime so that it's no longer the single point of failure for scans.
- Run more concurrent scans. Each node can empower many scan runs at the same time. You can manually scale out nodes of the Kubernetes cluster if you need more concurrent scans.
- When scanning some sources like Azure Blob, Azure Data Lake Storage Gen2, and Azure Files, each scan run can use multiple nodes to boost the scan performance. For other sources, scans are executed on only one of the nodes.

The capability of Kubernetes supported self-hosted integration runtime can be updated by manually scaling out/in nodes of the Kubernetes cluster.

> ⓘ **Note**

You must upload all necessary drivers for scanning on each new node.

Networking requirements

Expand table

Domain name	Outbound ports	Description
Public cloud: `<tenantID>-api.purview-service.microsoft.com` Azure Government: `<tenantID>-api.purview-service.microsoft.us` China: `<tenantID>-api.purview-service.microsoft.cn`	443	Required to connect to Microsoft Purview service. If you use Microsoft Purview Private Endpoints, this endpoint is covered by account private endpoint.
Public cloud: `<purview_account>.purview.azure.com` Azure Government: `<purview_account>.purview.azure.us` China: `<purview_account>.purview.azure.cn`	443	Required to connect to Microsoft Purview service. If you use Microsoft Purview Private Endpoints, this endpoint is covered by account private endpoint.
Public cloud: `<managed_storage_account>.blob.core.windows.net` or `<ingestion_storage_account>.*.blob.storage.azure.net` Azure Government: `<managed_storage_account>.blob.core.usgovcloudapi.net` or `<ingestion_storage_account>.blob.core.usgovcloudapi.net` China: `<managed_storage_account>.blob.core.chinacloudapi.cn` or `<ingestion_storage_account>.blob.core.chinacloudapi.cn`	443	Required to connect to the Microsoft Purview managed Azure Blob storage account.
Public cloud: `<managed_storage_account>.queue.core.windows.net` or `<ingestion_storage_account>.*.queue.storage.azure.net` Azure Government: `<managed_storage_account>.queue.core.usgovcloudapi.net` or `<ingestion_storage_account>.queue.core.usgovcloudapi.net` China: `<managed_storage_account>.queue.core.chinacloudapi.cn` or `<ingestion_storage_account>.queue.core.chinacloudapi.cn`	443	Required to connect to the Microsoft Purview managed Azure Queue storage account.
Public cloud: `*.compute.governance.azure.com` Azure Government: `*.compute.governance.azure.us` China: `*.compute.governance.azure.cn`	443	Required to connect to the Microsoft Purview service. Currently

Domain name	Outbound ports	Description
		wildcard is required as there's no dedicated resource.
mcr.microsoft.com	443	Required to download images.
*.data.mcr.microsoft.com	443	Required to download images.

> **Note**
>
> Depending on the sources users want to scan, they also need to allow other domains and outbound ports for other Azure or external sources.

Version

Typically, we release one new minor version of self-hosted integration runtime every month, which includes features, enhancements, and bug fixes.

Each version of the self-hosted integration runtime expires in one year.

How to check the current version

You can check the version of your Kubernetes self-hosted integration runtime either on the portal, or with the IRCTL.

Portal

1. In the Microsoft Purview portal, navigate to the Data Map.
2. Select **Integration runtimes**
3. The fourth column in your integration runtime's description line will be **Version**, and you can check the version there.

IRCTL (1.1.0 and above)

The describe command returns the integration runtime's version.

```
Bash
```

```
./irctl describe
```

Auto-update

Starting from version 1.1.0, the Kubernetes self-hosted integration runtime supports auto-update, which is enabled by default. This feature ensures your integration runtime is automatically upgraded to the latest Microsoft-managed version approximately once a month.

Opt-out

We recommend keeping auto-update enabled to benefit from the newest features and enhancements. However, you have the option to opt-out of auto-update using IRCTL. The auto-update configuration persists through reinstallation, so you don't need to disable it with each installation.

Bash

```
./irctl config set autoUpdate.enabled false
./irctl config view
```

Auto-update version vs latest version

To ensure stability, the auto-update is usually behind the latest version with a one-month delay. The auto-update version is managed by Microsoft.

If you would like to upgrade your integration runtime to newer versions, a manual upgrade should be performed with IRCTL of the specific version.

Next Steps

- Choose the right integration runtime configuration
- IRCTL instructions and details
- Troubleshoot kubernetes supported integration runtime

IRCTL for Kubernetes supported self-hosted integration runtime

07/15/2025

In order to manage your Microsoft Purview Kubernetes SHIR locally, you need a command line tool named IRCTL. You can download this tool during the SHIR creation process. IRCTL is a command line tool to manage your Microsoft Purview SHIR.

Prerequisites

The system requirements for IRCTL are:

- **Machine running IRCTL** - OS is Linux or Windows supporting TLS 1.2
- The machine to run IRCTL should be able to connect to the following endpoints:

Expand table

Domain name	Outbound ports	Description
Public cloud: `*.compute.governance.azure.com` Azure Government: `*.compute.governance.azure.us` China: `*.compute.governance.azure.cn`	443	Required to connect to Microsoft Purview service. If you use Microsoft Purview Private Endpoints, this endpoint is covered by account private endpoint.

- **Kubernetes context**: Kubernetes context, which contains Kubernetes cluster information and user's permissions and credential for this cluster, is needed to talk to your Kubernetes cluster. To ease the configuration for the user's permissions for SHIR management, you can start with Kubernetes Admin role. This context is generated with the setup of your Kubernetes cluster and saved in a config file. Where and how you can get this file depends on your setting up the Kubernetes cluster.
  - If you use `kubeadm init` to set up the Kubernetes cluster, you can find the config file under `/etc/Kubernetes/admin.conf`.
  - If you use AKS, you can follow the guidance of AKS to use Az PowerShell module command to get credentials of this cluster to your local machine. The context can be merged to the config file under `$HOME/.kube/config` directly.
  - If you're using other tools setting up a Kubernetes cluster, refer to the Kubernetes documentation.

- As you have the config file of the Kubernetes context, merge it to the config file, which is `$HOME/.kube/config`, on the machine you would like to run IRCTL command. Or you can set the config file of the Kubernetes context in an environment variable named KUBECONFIG as well. For more information about the Kubernetes context, see Configure Access to Multiple Clusters.

Install and run IRCTL

To control and manage a Kubernetes supported self-hosted integration runtime, users can download a command line tool named IRCTL. Users can also monitor the status on the Portal by IRCTL and kubectl command line. The following are the steps to your Kubernetes supported self-hosted integration runtime.

Windows

Download Latest IRCTL for Windows

Specific version

To download a specific version of the IRCTL, change the version segment in the URL: `https://aka.ms/purview-irctl/irctl-windows-<version>.tar.gz`

For example for version 1.0.0: `https://aka.ms/purview-irctl/irctl-windows-1.0.0.tar.gz`

Available IRCTL versions can be found in the IRCTL release notes.

After you download the tool:

1. Unzip and open the tool folder.
2. Move it to a location on your machine where you can access it from the command line.
3. Open the command line and navigate to the folder where the IRCTL application is housed.
4. Use the below commands to interact with IRCTL and follow these steps to install a Kubernetes supported integration runtime.

IRCTL command details

You can download the IRCTL command line tool from your integration runtime settings in Microsoft Purview. Select your integration runtime and under **Settings** you can see the **Download IRCTL and install integration runtime** link. Select it to download the tool.

You can use the IRCTL command line to set up or manage a Kubernetes supported self-hosted integration runtime. Use this syntax to run IRCTL commands from your terminal window:

```
Bash

./irctl [command][flags]...
```

- **Command** specifies the operation you want to perform
- **Flags** specifies parameters for the operation

Expand table

Command	Description
config	Manages IRCTL configurations
create	Create a new SHIR resource
delete	Delete an SHIR resource
describe	Describes the SHIR running on the Kubernetes cluster
help	Provides help about any command
storage	Manage local storage reserved for SHIR
log	Manage logs of SHIR
upgrade	Upgrade an SHIR to the same version with IRCTL

Config

Manage IRCTL configurations.

```
Bash

./irctl config [flags]
./irctl config [command]
```

Expand table

Command	Description
set	Modify one SHIR configuration.
view	View all SHIR configurations.

Expand table

Flag	Parameter type	Description
-h,--help		Help for config

Config set

Modify one self-hosted integration runtime configuration.

Bash

```
./irctl config set <config-name> <config-value> [flags]
```

Expand table

Flag	Parameter type	Description
-c, --context	string	Name of the Kubernetes context to use, if not specified, current context of Kube config is used.
-h,--help		Help for config

Config view

View all SHIR configurations.

Bash

```
./irctl
```

Expand table

Flag	Parameter type	Description
-c, --context	string	Name of the Kubernetes context to use, if not specified, current context of Kube config is used.
-h,--help		Help for config

Allowed SHIR Configurations

Expand table

Config name	Config type	Valid values	Default value (value unset)	Description	Persists through reinstallation
autoUpdate.enabled	Boolean	true/false	true	Controls whether the current SHIR should be autoupdated.	Yes

Create

Create a new SHIR resource.

Bash

```
./irctl create [flags]
```

Expand table

Flag	Parameter type	Description
-c,--context	string	Name of the Kubernetes context to use, if not specified, current context of Kube config is used
-h,--help		Help for create
-r,--registration-key	string	Key to register SHIR
-s,--selector	string	Node selector (label query) to filter on, supports '='. (for example,--selector key1=value1,key2=value2) Select nodes by given labels for SHIR

Flag	Parameter type	Description
		to run on. Default is empty, which means SHIR runs on all nodes.

Example:

```
Bash

./irctl create --registration-key 1234@keyvalue
```

Upgrade

Upgrade an SHIR to the same version with IRCTL.

```
Bash

./irctl upgrade [flags]
./irctl upgrade [command]
```

Expand table

Command	Description
status	Monitors the status the latest SHIR upgrade.

Expand table

Flag	Parameter type	Description
-c,--context	string	Name of the Kubernetes context to use, if not specified, current context of Kube config is used
-h,--help		Help for upgrade

Delete

Delete an SHIR resource.

```
Bash

./irctl delete [flags]
```

Expand table

Flag	Parameter type	Description
-c,--context	string	Name of the Kubernetes context to use, if not specified, current context of Kube config is used
-h,--help		Help for delete

Describe

Describe SHIR running on the Kubernetes cluster.

```
Bash

./irctl describe [flags]
```

Expand table

Flag	Parameter type	Description
-c,--context	string	Name of the Kubernetes context to use, if not specified, current context of Kube config is used
-h,--help		Help for describe

Storage

Manage files in SHIR shared storage.

```
Bash

./irctl storage [flags]
./irctl storage [command]
```

Expand table

Command	Description
upload	Upload files to SHIR shared storage.
list	List files in SHIR shared storage.

Command	Description
delete	Delete files in SHIR shared storage.

Expand table

Flag	Parameter type	Description
-h,--help		Help for storage

Storage upload

Upload files to SHIR shared storage.

Bash

```
./irctl storage upload -s <source> -d <destination> [flags]
```

Expand table

Flag	Parameter type	Description
-c,--context	string	Name of the Kubernetes context to use, if not specified, current context of Kube config is used.
-d,--destination string	string	Path to store file/folder uploaded in SHIR shared storage.
-h,--help		Help for delete
-s,--source	string	Path to the file/folder to be uploaded.

Storage list

List files in SHIR shared storage.

Bash

```
./irctl storage list <path> [flags]
```

Expand table

Flag	Parameter type	Description
-c,--context	string	Name of the Kubernetes context to use, if not specified, current context of Kube config is used.
-h,--help		Help for list.

Storage delete

Delete files in SHIR shared storage.

```
Bash

./irctl storage delete <path> [flags]
```

Expand table

Flag	Parameter type	Description
-c,--context	string	Name of the Kubernetes context to use, if not specified, current context of Kube config is used.
-h,--help		Help for delete.

Log

Log operations from SHIR.

```
Bash

./irctl log [flags]
./irctl log [command]
```

Expand table

Command	Description
upload	Upload logs from SHIR.
download	Download logs from SHIR.

Expand table

Flag	Parameter type	Description
-h,--help		Help for log.

Log upload

Upload logs from SHIR.

```
Bash

./irctl log upload [flags]
./irctl log upload [command]
```

Expand table

Command	Description
Status	Monitors the status of SHIR log uploading.

Expand table

Flag	Parameter type	Description
-c,--context	string	Name of the Kubernetes context to use, if not specified, current context of Kube config is used.
-h,--help		Help for upload.
-t,--timeout	int	Timeout for the log upload operation in seconds: [900, 3600]. Default is 900 seconds. (Default 900)

Log download

Upload logs from SHIR.

```
Bash

./irctl log download [flags]
```

Expand table

Flag	Parameter type	Description
-c,--context	string	Name of the Kubernetes context to use, if not specified, current context of Kube config is used.
-d,--destination	string	Local path to store logs downloaded from SHIR.
-h,--help		Help for download.

Next Steps

- Install and configure kubernetes supported integration runtime
- Choose the right integration runtime configuration
- Troubleshoot kubernetes supported integration runtime

Troubleshoot a Kubernetes supported self-hosted integration runtime

Article • 10/14/2024

This article covers the details for troubleshooting the new self-hosted integration runtime that is Kubernetes-based for Linux.

You can look up any errors you see in the error guide below. To get support and troubleshooting guidance for SHIR issues, you might need to generate a Log Upload ID and reach out to Microsoft support.

Gather Kubernetes self-hosted integration runtime logs

To generate the Log Upload ID for Microsoft Support:

1. Once a scan shows status Failed, navigate to the VM or machine where your IRCTL tool is installed.
2. Use the `./irctl log upload` command.
3. When the logs are uploaded, keep a record of the Log Upload ID that's printed.
4. If the self-hosted integration runtime fails to register, use the guide below to download logs to local and send it to Microsoft support

To collect the self-hosted integration runtime log and scan run log to troubleshoot your SHIR, use the `./irctl log download` command.

For example:

```Bash
./irctl log download --destination "C:\Users\user\logs\"
```

Logs will be downloaded to the destination path.

> **Note**
>
> The log will be reserved for 14 days. Please persist it by uploading it to Microsoft or downloading it to your local machine.

IRCTL connectivity error to Kubernetes

You might get a Kubernetes context configuration from your Kubernetes administrator, and the registration could fail with one of the following error messages:

- `Error: invalid flag context [] .kube/config: no such file or directory`
- `[Warning] Failed to create kube client with context [] with error`

Cause

When installing a self-hosted integration runtime, a correct Kubernetes config and a stable connectivity is required.

Resolution

1. Make sure the Kubernetes context file is located in the correct path.
2. Make sure the IRCTL machine can reach the Kubernetes cluster API server.

IRCTL permission error

When connecting you could see the following error messages:

- `[Error] Failed to list namespaces to get Running SHIR`
- `[Error] Failed to get configmap/ create job/ etc.`

Cause

When installing a self-hosted integration runtime, sufficient permissions to multiple Kubernetes resources are required.

Resolution

Regenerate the Kubernetes service-account-token with an Admin Role.

IRCTL connectivity error to Microsoft Purview service endpoint

When you attempt to register the Kubernetes supported self-hosted integration runtime, the IRCTL create command could return the following errors:

- `[Error] Failed to register SHIR with error: Post “https://[REGION].compute.governance.azure.com/purviewAccounts/[]/integrationruntimes/[]/registerselfhostedintegrationruntime: []”`

Cause

IRCTL can't connect to service backend. Such issue is usually caused by network settings or firewall.

Resolution

Review the network topology of the IRCTL host machine. Refer to these sections of the general integration runtime troubleshooting page: Firewall, DNSServer, SSL cert trusts, and http proxy.

Registration key isn't authorized

When you attempt to register the Kubernetes supported self-hosted integration runtime, the IRCTL create command could return the following errors:

- `[Error] failed to register SHIR with error: Request is not authorized.`

Cause

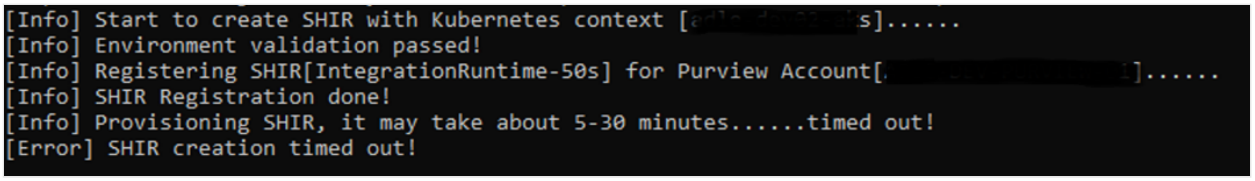
They had expired or was manually revoked.

Resolution

Regenerate the key from the integration runtime's page in the Microsoft Purview portal and register again.

Kubernetes self-hosted IR creation timed out error

When you attempt to register the Kubernetes supported self-hosted integration runtime, the IRCTL create command could run long until it eventually times out.



You can start with checking the status of Pods under the namespaces mentioned by the `irctl describe` command.

For example:

Bash

```
./irctl describe

K8s SHIR Name:shir-demo
Purview AccountName:   shirdemopurview
Installation ID:       00000000-0000-0000-0000-000000000000
Kubernetes Namespace:  shirdemopurview-shir-demo, compute-fleet-
system(control-plane)
K8s SHIR Version:      Unknown (Installation not completed)
Status:                Initializing
Healthiness:           Unhealthy

kubectl get pods --namespace shirdemopurview-shir-demo

NAME                                    READY   STATUS    RESTARTS   AGE
batch-defaultspec-4pbwx                 0/1     Pending   0          10m
batch-defaultspec-7t9bl                 0/1     Pending   0          10m
dynamic-config-provider-778c686fdc-9mkjb   0/1     Pending   0          10m
interactive-schemaprocess-bcrmf         0/1     Pending   0          10m
interactive-schemaprocess-fn66x         0/1     Pending   0          10m
logagent-ds-84jqn                       0/1     Pending   0          10m
logagent-ds-k7vw8                       0/1     Pending   0          10m
user-credential-proxy-579c899b64-d4q5v     0/1     Pending   0          10m
```

There are a couple potential causes, listed below:

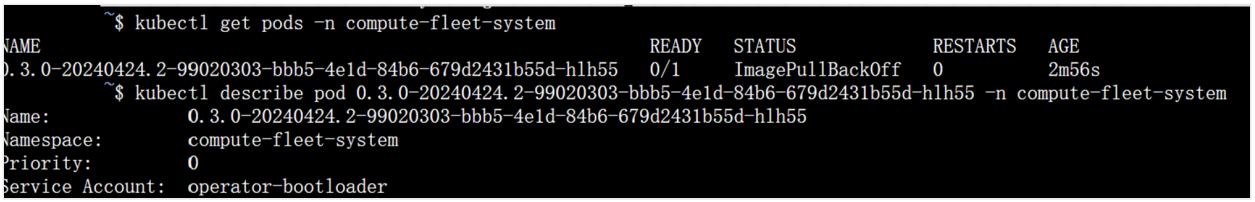
- Connectivity with the Microsoft Purview service endpoint
- Kubernetes node configuration error

Cause - Connectivity with the Microsoft Purview service endpoint

Kubernetes can't connect to MCR (mcr.microsoft.com). This error is usually caused by network settings or a firewall.

If seeing the status as 'ImagePullBackOff', it means your Kubernetes cannot connect to MCR (mcr.microsoft.com) to download Pod images. This error is usually caused by

network settings or a firewall.



Resolution - Connectivity with the Microsoft Purview service endpoint

Review the network topology for your Kubernetes cluster. For example, for Azure Kubernetes you should check:

- DNS Resolution
- Egress outbound
- VNet/Firewall

> **Note**
>
> The troubleshooting steps required will be different for each Kubernetes provider. Deployment location and individual network details will vary between networks. You'll need to review connectivity through your organization's network.

Review the network topology of the IRCTL host machine. Refer to these sections of the general integration runtime troubleshooting page: Firewall, DNSServer, SSL cert trusts, and http proxy.

Cause - Kubernetes node configuration error

If the status of some of the Pods stuck in 'Pending', use the describe pod command to see details of the Pod.

For example:

```Bash
kubectl describe pod batch-defaultspec-4pbwx  --namespace shirdemopurview-shir-demo

Events:
  Type     Reason            Age                From               Message

  ----     ------            ----               ----               -------
```

```
  Warning  FailedScheduling  13m                default-scheduler  0/5 nodes
are available: 1 Too many pods. preemption: 0/5 nodes are available: 5 No
preemption victims found for incoming pod..
```

The Events from the describe command can tell the reason why the Pod is pending. The FailedScheduling error with detailed message shows that the total number of Pods is exceeded the maximum number of the Pod on a Node. A new Pod cannot be scheduled to the selected Node.

> ⓘ **Note**
>
> If there are no Events seen under the description, please try to delete the Pod manually by `kubectl delete pod` command and track the newly created one.

Resolution - Kubernetes node configuration error

Reserve 20 pod number for the kubernetes integration runtime to support normal utilization and upgrade scenarios.

Kubernetes connectivity error with Microsoft Purview service endpoint

When you attempt to register the Kubernetes supported self-hosted integration runtime, the IRCTRL create command could run long until it eventually times out. Or, after a successful installation, the self-hosted integration runtime status shows as unhealthy or offline in the Microsoft Purview portal.

Check the logs using this command: `kubectl logs [podName] -n compute-fleet-system`

You could see one of these errors:

- `“TraceMessage”:”HttpRequestFailed”, “Host”: “fleet.[REGION].compute.governance.azure.com”`
- `Exception":"System.Net.Http.HttpRequestException: Connection refused fleet.[REGION].compute.governance.azure.com:443`
- `System.AggregateException: Failed to acquire identity token from https://fleet. [REGION].compute.governance.azure.com:443`

Cause

Kubernetes can't connect to the service backend. This error is usually caused by network settings or a firewall.

Resolution

Review the network topology for your Kubernetes cluster. For example, for Azure Kubernetes you should check:

- DNS Resolution
- Egress outbound
- VNet/Firewall

> **Note**
>
> The troubleshooting steps required will be different for each Kubernetes provider. Deployment location and individual network details will vary between networks. You'll need to review connectivity through your organization's network.

Review the network topology of the IRCTL host machine. Refer to these sections of the general integration runtime troubleshooting page: Firewall, DNSServer, SSL cert trusts, and http proxy.

Unregister a runtime whose local resource is unavailable

If your local self-hosted integration runtime is deleted from the Kubernetes cluster by accident, you can't delete it using the irctl delete command and you can't install it to another Kubernetes cluster.

Cause

A self-hosted integration runtime can only be installed on one Kubernetes cluster. After it's registered, it can't be installed on another cluster before it's unregistered.

Resolution

1. Check the self-hosted integration local status. You should see that no running self-hosted integration runtime is found.

   ```
   Bash
   ```

```
$./irctl describe
```

2. Check the self-hosted integration runtime in the Microsoft Purview portal. You should see an **Offline** status. (Though, there's a 1 hour latency for token expiration.)
3. Select **Unregister installation** next to the status and confirm the operation.
4. After the unregistration is complete, you can see the status shows as **Not registered**.
5. Select the integration runtime and get the registration key.
6. Reinstall the integration runtime.

Next Steps

- Install and configure kubernetes supported integration runtime
- Choose the right integration runtime configuration
- IRCTL instructions and details

Feedback

Was this page helpful? Yes No

Learn about sensitivity labels in Data Map (preview)

07/16/2025

> **Important**
>
> Labeling in the Microsoft Purview Data Map is currently in preview. The Supplemental Terms of Use for Microsoft Azure Previews include additional legal terms that apply to Azure features that are in beta, preview, or otherwise not yet released into general availability.

To get work done, people in your organization collaborate with others both inside and outside the organization. Data doesn't always stay in your cloud, and often roams everywhere, across devices, apps, and services. When your data roams, you still want it to be secure in a way that meets your organization's business and compliance policies.

Applying sensitivity labels to your content enables you to keep your data secure by stating how sensitive certain data is in your organization. It also abstracts the data itself, so you use labels to track the type of data, without exposing sensitive data on another platform.

For example, applying a sensitivity label 'highly confidential' to a document that contains social security number and credit card numbers helps you identify the sensitivity of the document without knowing the actual data in the document.

Benefits of labeling in Microsoft Purview

Microsoft Purview allows you to apply sensitivity labels to assets, enabling you to classify and protect your data.

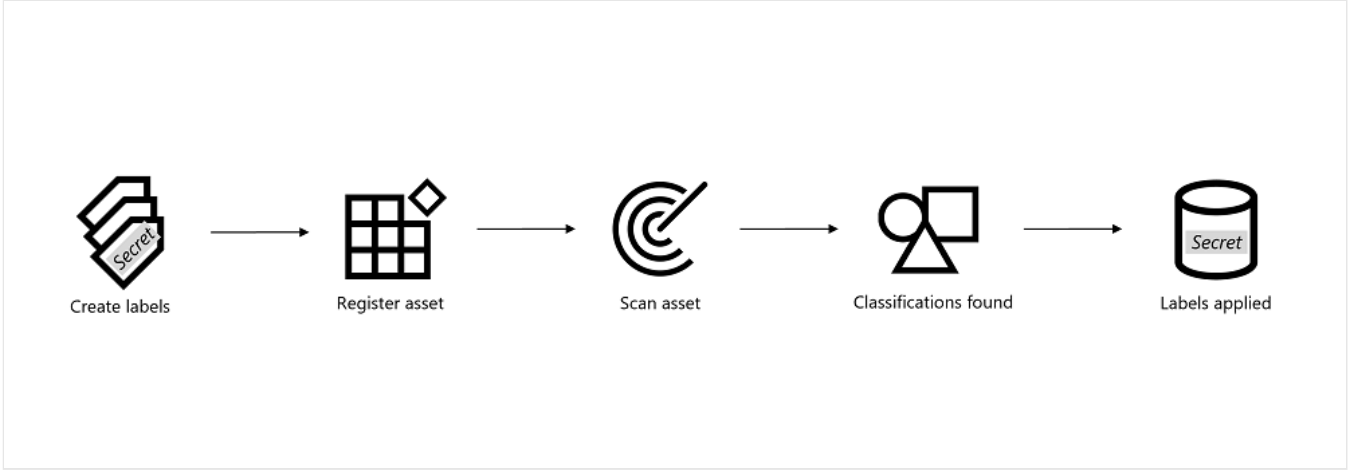
- **Label travels with the data:** The sensitivity labels created in Microsoft Purview Information Protection can also be extended to the Microsoft Purview Data Map, SharePoint, Teams, Power BI, and SQL. When you apply a label on an office document and then scan it into the Microsoft Purview Data Map, the label will be applied to the data asset. While the label is applied to the actual file in Microsoft Purview Information Protection, it's only added as metadata in the Microsoft Purview map. While there are differences in how a label is applied to an asset across various services/applications, labels travel with the data and is recognized by all the services you extend it to.
- **Overview of your data estate:** Microsoft Purview provides insights into your data through precanned reports. When you scan data into the Microsoft Purview Data Map, we hydrate

the reports with information on what assets you have, scan history, classifications found in your data, labels applied, glossary terms, etc.

- **Automatic labeling:** Labels can be applied automatically based on sensitivity of the data. When an asset is scanned for sensitive data, autolabeling policies are used to decide which sensitivity label to apply. You can create autolabeling policies for sensitivity labels, defining which classification/sensitive information types when found will apply a label automatically.
- **Apply labels to files and database columns:** Labels can be applied to files in storage such as Azure Data Lake or Azure Files and to table columns in Azure SQL Database.

Sensitivity labels are tags that you can apply on assets to classify and protect your data. Learn more about sensitivity labels here.

How to apply labels to assets in the Microsoft Purview Data Map



Being able to apply labels to your asset in the data map requires you to perform the following steps:

1. Create new or apply existing sensitivity labels. You can apply any sensitivity label with label scope of **Files & other data assets** to Microsoft Purview assets.
2. Register and scan your asset in the Microsoft Purview Data Map.
3. Microsoft Purview applies **classifications**: When you schedule a scan on an asset, Microsoft Purview scans the type of data in your asset and applies classifications to it in the data map. Application of classifications is done automatically by Microsoft Purview, there's no action for you.
4. Create an auto-labeling policy for non-Microsoft 365 workloads: Create an auto-labeling policy scoped to All or specific assets within supported non-Microsoft 365 locations. Define the out of the box classifications that once detected in the next data scan will result in the automatic labeling of the selected sensitivity label.

> **Note**
>
> Autolabeling policies are conditions that you specify, stating when a particular label should be applied. When the data is next scanned Purview data map will check if these conditions are met to automatically assign the label to the data. If you don't see a supported data source show up in the auto-labeling policy for selection, verify that the data source was correctly registered and shows up in data map. The locations shown for auto-labeling policy scoping depends on what assets have already been registered.

Supported data sources

You can find the data sources that support labeling on the supported sources page.

Labeling for SQL databases

Microsoft also supports labeling for SQL database columns using the SQL data classification in SQL Server Management Studio (SSMS). While Microsoft Purview uses the global sensitivity labels, SSMS only uses labels defined locally.

Labeling in Microsoft Purview and labeling in SSMS are separate processes that don't currently interact with each other. Therefore, **labels applied in SSMS are not shown in Microsoft Purview, and vice versa**. We recommend Microsoft Purview for labeling SQL databases, because the labels can be applied globally, across multiple platforms.

For more information, see the SQL data discovery and classification documentation.

Next steps

- How to automatically label your content
- Sensitivity label insights
- Labeling Frequently Asked Questions

Apply sensitivity labels to your data in Data Map (preview)

07/16/2025

> **Important**
>
> Labeling in the Microsoft Purview Data Map is currently in preview and involves pay-as-you-go billing model. The Supplemental Terms of Use for Microsoft Azure Previews include additional legal terms that apply to Azure features that are in beta, preview, or otherwise not yet released into general availability.

Create new or apply existing sensitivity labels in Data Map

If you don't already have sensitivity labels, you need to create them and make them scoped for **Files & other data assets**. Existing sensitivity labels from Microsoft Purview Information Protection can also be modified with this scope to make them available to Data Map.

Step 1: Licensing requirements

Sensitivity labels are created and managed in the Microsoft Purview Information Protection. To create sensitivity labels for use through Microsoft Purview, you must have an active Microsoft 365 license that offers the benefit of automatically applying sensitivity labels and set up of pay as you go billing model.

To understand pay as you go billing by assets protected, see this pay-as-you-go billing model. For information about the specific licenses required, see this information on sensitivity labels. Microsoft 365 E5 trial licenses can be attained for your tenant by navigating here from your environment. For the full list of licenses, see the Sensitivity labels in Microsoft Purview FAQ. If you don't already have the required license, you can sign up for a trial of Microsoft 365 E5.

Step 2: Create or modify existing label to be scoped to "Files & other data assets" so they're available for Data Map

To create new sensitivity labels or modify existing labels, follow these instructions:

1. Sign in to the Microsoft Purview portal > **Information Protection** card > **Sensitivity labels**. If the **Information Protection** solution card isn't displayed, select **View all**

solutions and then select **Information Protection** from the **Data Security** section.

2. Select + **Create a label**.
3. Name the label. Then, under **Define the scope for this label**, select **Files & other data assets**.
4. Follow the rest of the prompts to configure the label settings. For more information about configuration options, see What sensitivity labels can do in the Microsoft 365 documentation.
5. To modify existing labels, browse to **Information Protection** > **Labels**, and select your label. Then select **Edit label** to open the **Edit sensitivity label** configuration, with all of the settings you defined when you created the label.
6. When you're done creating all of your labels, make sure to view your label order, and reorder them as needed.

   To change the order of a label, select **...** > **More actions** > **Move up** or **Move down.**

   For more information, see the documentation for label priority (order matters).

Step 3: Create auto-labeling policies scoped to non-Microsoft 365 workloads

To create an auto-labeling policy for non-Microsoft 365 workloads, follow these instructions:

1. Sign in to the Microsoft Purview portal. Go to **Information Protection** > **Policies** > **Auto-labeling policies**.
2. Select + **Create auto-labeling policy**.
3. At **Choose info you want this label applied to**, you can select any **non-enhanced** templated or Custom.
4. Name your auto-labeling policy and give it an optional description.
5. Select a sensitivity label with the **Files & other data assets** scope.
6. Keep the admin units set as **Full directory**, then select **Next**.
7. Select the non-Microsoft 365 locations you want to label. Select to label all assets within a location or specify which assets in each location you want to label. If you don't see a supported data source show up in the auto-labeling policy for selection, verify that the data source was correctly registered and shows up in Data Map. The locations displayed

for auto-labeling policy scoping depend on what assets have already been registered. Check the list of supported data sources for auto-labeling.

8. Define **Common rules** to specify the same auto-labeling conditions/rules across multiple selected non-Microsoft locations. Select **Advanced rules** to define different specific rules for each location.

9. Create a new rule and select Microsoft's prebuilt, out of the box classifiers as conditions for auto-labeling. Custom sensitive info types and advanced classifiers like named entities (All Full Names, All Physical Addresses), trainable classifiers, credentials, and Exact Data Match are **not** supported.

10. Select **Run policy in simulation mode**, and select the checkbox if you want to automatically turn on the policy after seven days.

    > ⓘ **Note**
    >
    > Microsoft Purview Data Map locations don't support simulation.

11. Review your selected configurations and select **Create policy**.

Step 4: Scan your data to apply sensitivity labels automatically

Wait 15 minutes after auto-labeling policy creation for policy to saturate. Scan your data in Data Map to automatically apply the labels you've created, based on the auto-labeling policy you've defined.

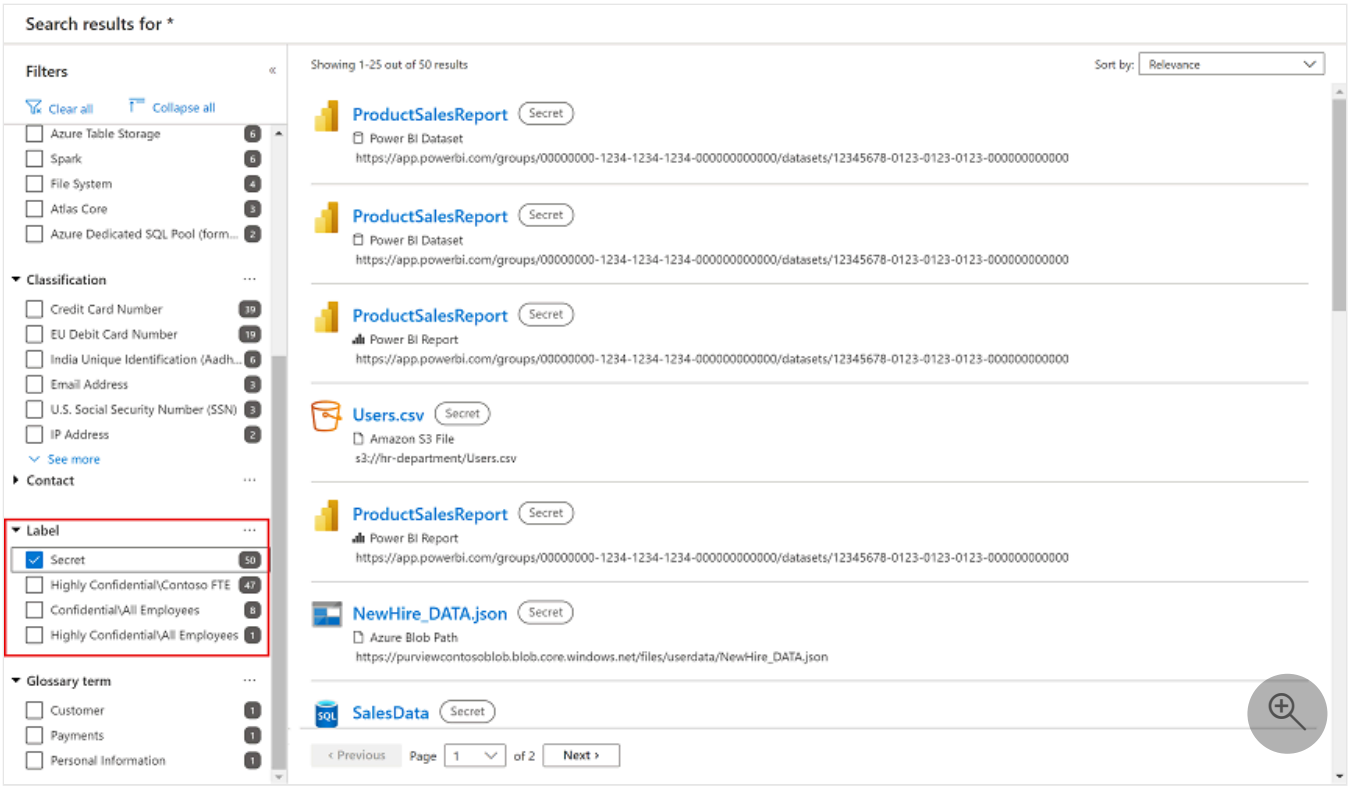
For more information on how to set up scans on various assets in Data Map, find links to each data source that connect to Data Map.

View labels on assets in the catalog

Once you've configured auto-labeling policies for the non-Microsoft 365 workloads in Microsoft Purview Information Protection and scanned your data in Data Map, labels are automatically applied to your assets in Data Map.

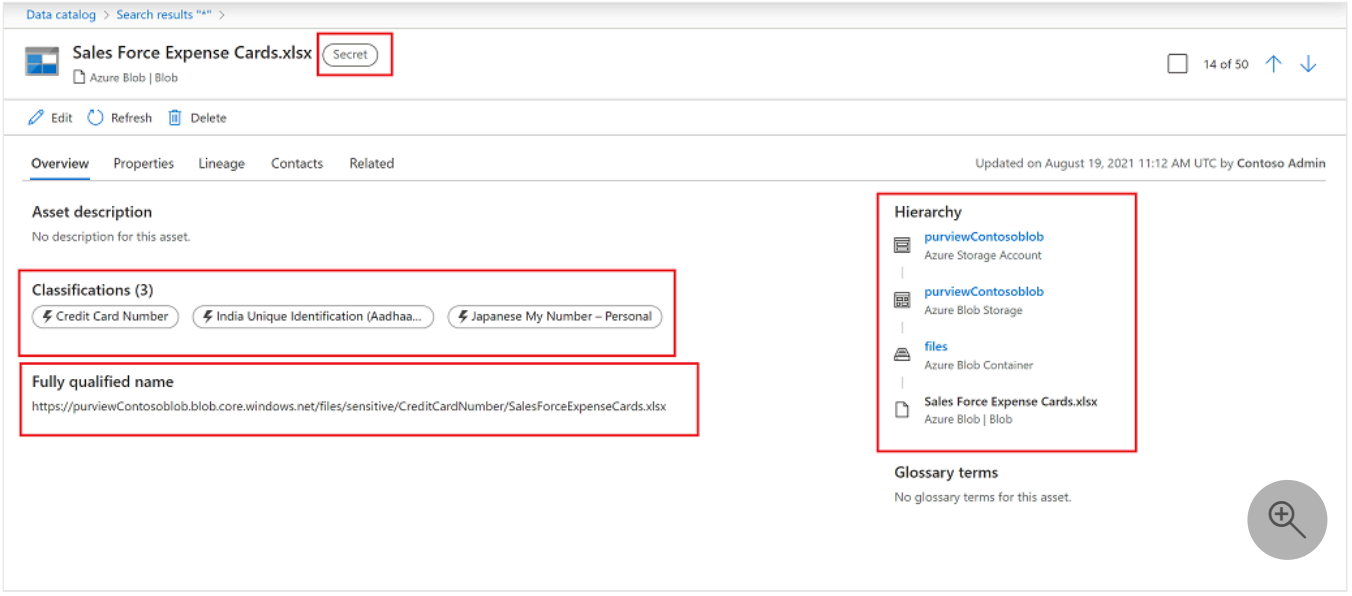
To view the labels applied to your assets in the Microsoft Purview catalog:

In Microsoft Purview Unified Catalog, use the **Label** filtering options to show assets with specific labels only. For example:



To view details of an asset including classifications found and label applied, select the asset in the results.

For example:



View Insight reports for the classifications and sensitivity labels

Find insights on your classified and labeled data in Data Map use the **Classification** and **Sensitivity labeling** reports.

Classification insights

Sensitivity label insights

Overview of Labeling in Microsoft Purview

Labeling Frequently Asked Questions

Sensitivity labels in Data Map FAQ (preview)

This article lists frequently asked questions about sensitivity labeling in the Microsoft Purview Data Map, with their answers and links to more information.

> **Important**
>
> Labeling in the Microsoft Purview Data Map is currently in preview. The Supplemental Terms of Use for Microsoft Azure Previews include other legal terms that apply to Azure features that are in beta, preview, or otherwise not yet released into general availability.

Licensing and setup

What are the licensing requirements to apply sensitivity labels in Microsoft Purview Data Map?

To use sensitivity labels in the Microsoft Purview Data Map, you'll need at least one Microsoft 365 license/account within the same Microsoft Entra tenant as your Microsoft Purview account.

The following Microsoft 365 licenses are required to automatically apply sensitivity labels to your assets in Microsoft 365 and the Microsoft Purview Data Map:

- Microsoft 365 E5/A5/G5
- Microsoft 365 E5/A5/G5 Compliance
- Microsoft 365 E5/A5/G5 Information Protection, and Governance
- Office 365 E5, Enterprise Mobility + Security E5/A5/G5, and AIP Plan 2

For more information, see the Microsoft Purview service description.

To use auto-labeling and protection policies for non-Microsoft 365 data sources, 1 Microsoft 365 E5 licenses and set up of pay as you go billing model will be required. To understand pay as you go billing by assets protected, see the pay-as-you-go billing model. For information about the specific licenses required, see this information on sensitivity labels. Microsoft 365 E5 trial licenses can be attained for your tenant by navigating here from your environment.

How are sensitivity labels applied to non-Microsoft 365 locations?

You can create and apply autolabeling policies to detect Microsoft out-of-the-box classifiers to specific or all data assets that are registered in Data Map for supported locations.

- Sensitivity labels are applied only to the asset metadata in Data Map and aren't applied to the actual files and database columns. These sensitivity labels don't modify your files and databases in any way.

Classifications vs sensitivity labels

What is the difference between classifications and sensitivity labels?

The following table lists the differences between classifications and sensitivity labels:

Expand table

Comparison	Classifications	Sensitivity labels
Definition	Classifications are regular expressions or patterns that can help identify data types that exist inside an asset.	Sensitivity labels are tags that allow organizations to categorize data based on business impact, while abstracting the type of data from the user.
Examples	Social Security Number, Drive license number, Bank account number, etc.	Highly confidential, Confidential, General, Public, etc.
Scope	The scope of classifications applied to an asset is limited to the Microsoft Purview Data Map where the classifications were applied. If the data moves to an asset managed by another Microsoft Purview Data Map, classifications applied in the original location aren't visible in the new location.	Sensitivity labels applied on an asset travel with the data no matter where the data goes. For example, this means that sensitivity labels applied to a file in Microsoft Purview Information Protection are automatically visible and remain applied to the file, even if it moves to Azure, SharePoint, or Teams.

Comparison	Classifications	Sensitivity labels
Scan Process	Scanning an asset in the Microsoft Purview Data Map looks for both system-defined and user-defined (custom) classifications in your data. If found, classifications are added in the Microsoft Purview map for the scanned asset.	If you have an auto-labeling policy for Azure Storage or Azure SQL, scanning an asset in the Microsoft Purview Data Map applies the labels to assets in the catalog based on the out-of-the-box classifications found in the scan and defined in the auto-labeling policy.
Authoring environment	Custom classifications and classification rules can be created in the Microsoft Purview Data Map. You can also create custom classifications in Microsoft Purview Information Protection. However, we don't yet support importing them to the Microsoft Purview Data Map.	Manage auto-labeling of sensitivity labels in Microsoft Purview using the Microsoft Purview Information Protection.
Assignment Limits	Assets can have no classifications, or one or more classifications assigned.	Each asset/ database column can have only one sensitivity label.
Asset application workflow	You can use Microsoft Purview Unified Catalog to manually add or modify classifications that are assigned to an asset.	In the Microsoft Purview Data Map, sensitivity labels are automatically assigned to assets based on classifications found. Applying labels manually in the Microsoft Purview Data Map isn't currently supported.
More Information	Learn more about classifications.	Learn more about sensitivity labels.

Are classifications and Sensitive Information Types (SITs) the same thing?

While classifications and SITs are fundamentally the same things, classifications are a Microsoft Purview Data Map concept and SITs are a Microsoft Purview Information Protection concept. Both classifications and SITs are used by their respective services to identify the type of data found in an asset.

Labeling capabilities in the Microsoft Purview Data Map

Which data sources can I apply sensitivity labels to in the Microsoft Purview Data Map?

You can apply sensitivity labels to all the data sources listed under Supported data sources for sensitivity labels in the Microsoft Purview Data Map.

Which file types can I apply sensitivity labels to in the Microsoft Purview Data Map?

You can apply sensitivity labels to all Microsoft Purview Data Map supported file types.

Can I use my custom Sensitive Information Type (SIT) in Microsoft Purview Information Protection for schematized data assets?

No, custom sensitive information types aren't supported in the Microsoft Purview Data Map at this time. The Microsoft Purview Data Map currently only supports Microsoft Purview Information Protection built-in sensitive information types.

Can I use the advanced classifiers from Microsoft Purview Information Protection in the Microsoft Purview Data Map?

No, advanced classifiers aren't currently supported in the Microsoft Purview Data Map and won't appear.

Can autolabeling for Purview Data Map apply to assets that might include credential content?

Data Map currently doesn't support scanning for credentials. When Data Map supports scanning for credentials, you should be able to apply labels based on credentials found.

Can I manually label an asset, or manually modify or remove a label in the Microsoft Purview Data Map?

The Microsoft Purview Data Map supports auto-labeling only. Labels are automatically applied to assets in the data map based on autolabeling policy configuration.

The Data Map doesn't currently support manually applying a label, modifying, or removing a label from an asset. However, when you apply a label on an office document and then scan it into Data Map, the label will be applied to the data asset. While the label is applied to the actual file in Microsoft Purview Information Protection, it's only added as metadata in Data Map. While there are differences in how a label is applied to an asset across various services/applications, labels travel with the data and are recognized by all the services you extend it to.

Can I apply encryption and/or content marking to files in the Microsoft Purview Data Map, as I can for Office documents and emails?

No, although the sensitivity label might be configured for these protection actions, we don't currently support encryption and content marking for files in the Microsoft Purview Data Map. We only support protection actions that are configured in protection policies.

Does the Microsoft Purview Data Map support data loss prevention?

No, Data Map doesn't currently provide data loss prevention (DLP) capabilities. Data Loss Prevention is currently supported only for Microsoft 365 apps and services, including Fabric.

Access and roles

Where can I manage my autolabeling policies for labeling beyond Microsoft 365?

Sensitivity labels and autolabeling policies are managed in Microsoft Purview Information Protection. For more information, see How to create sensitivity labels in Microsoft Purview Information Protection.

Who can manage sensitivity labels?

The following built-in admin roles include permissions to manage sensitivity labels (most to least permissive):

- Global Administrator
- Compliance Administrator
- Information Protection
- Information Protection Admins
- Information Protection Analysts
- Information Protection Investigators
- Information Protection Readers

For more information, see Permissions required to create and manage sensitivity labels. After you have compliance and global administrators configured, those administrators can give access to individual users.

> **Note**
>
> Microsoft recommends that you use roles with the fewest permissions. This helps improve security for your organization. Global Administrator is a highly privileged role that should be limited to emergency scenarios when you can't use an existing role.

Who can search and browse assets with sensitivity labels in Microsoft Purview Unified Catalog?

All users with at least data reader access to the Microsoft Purview Data Map have permissions to search and browse assets with sensitivity labels in Unified Catalog.

Who can view the sensitivity label insights report in Microsoft Purview Data Estate Insights?

All users with the insights reader role and at least data reader permissions on applicable collections will have permissions to view sensitivity label insights reports in Microsoft Purview Data Estate Insights.

Technical details

Does the Microsoft Purview Data Map scan an entire asset when applying autolabeling policies to the database columns?

The Microsoft Purview scanner samples the data. For more information, see sampling data for classification and autolabeling.

If there are multiple sensitivity labels that meet the classification criteria, which label is applied?

Sensitivity labels have a priority 'order' and the Microsoft Purview Data Map uses this order to assign labels. If there are multiple labels meeting the classification criteria, the Microsoft Purview Data Map selects the label with the highest order.

For more information, see Label priority order matters.

SQL data discovery and classification

Why does Microsoft support two classification experiences for SQL databases, Microsoft Purview and SQL data discovery and classification?

Microsoft Purview provides a classification and labeling experience for all your Azure assets including SQL databases. Microsoft Purview is intended for organizations that want to manage their entire data estate in a single place with the power of classification, labeling, alerting, and more. Microsoft Purview uses sensitivity labels, which have a global scope and travel with your data no matter where it moves to or what it transforms into.

In contrast, SQL data discovery and classification is built into SQL. SQL data discovery and classification existed before Microsoft Purview as a way to provide basic capabilities for discovering, classifying, labeling, and reporting the sensitive data in your SQL databases. SQL data discovery and classification use local labels that don't have a global scope and don't support sensitivity labels.

I applied labels in SQL data discovery and classification. Why are these labels not showing up on my asset in Microsoft Purview?

SQL classification uses local labels, while Microsoft Purview uses sensitivity labels. Labels applied in the SQL classification experience won't show up in Microsoft Purview. For more information, see Labeling for SQL databases.

For more information, see:

- How to automatically label your content
- Sensitivity label insights
- Classification insights

Last updated on 06/01/2026

Data lineage in classic Data Catalog

07/18/2025

This article provides an overview of data lineage in the classic Microsoft Purview Data Catalog. It also details how data systems can integrate with the catalog to capture lineage of data. Microsoft Purview can capture lineage for data in different parts of your organization's data estate, and at different levels of preparation including:

- Raw data staged from various platforms
- Transformed and prepared data
- Data used by visualization platforms

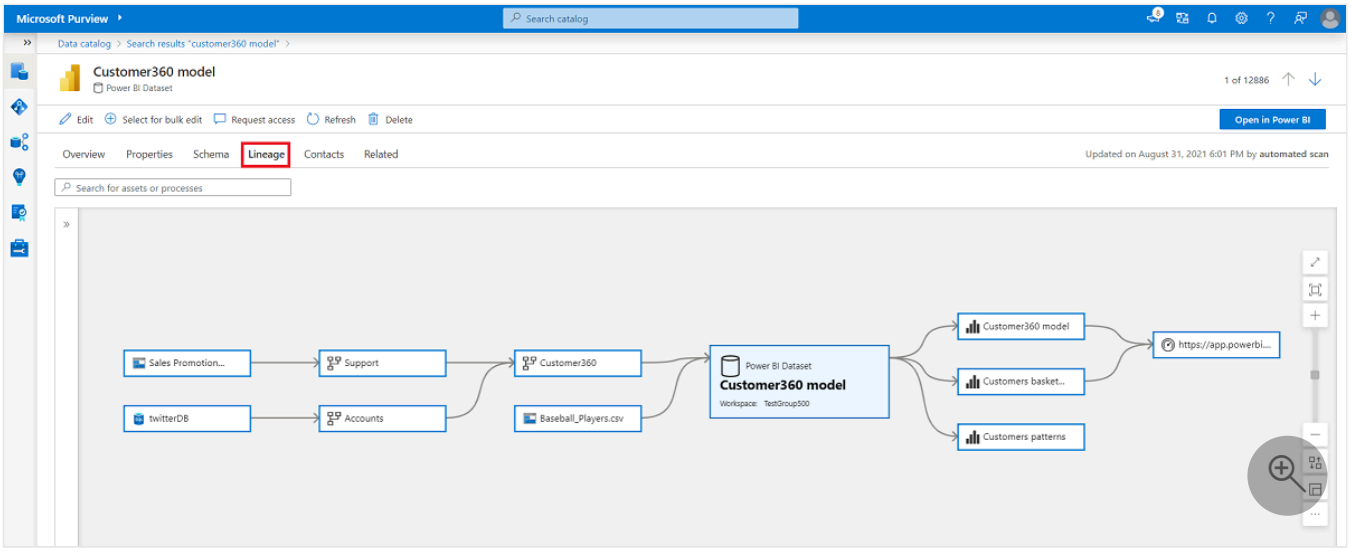
For data lineage in the Unified Catalog, see how to view data asset details and lineage information.

Use cases

Data lineage is broadly understood as the lifecycle that spans the data's origin, and where it moves over time across the data estate. It's used for different kinds of backwards-looking scenarios such as troubleshooting, tracing root cause in data pipelines and debugging. Lineage is also used for data quality analysis, compliance and "what if" scenarios often referred to as impact analysis. Lineage is represented visually to show data moving from source to destination including how the data was transformed. Given the complexity of most enterprise data environments, these views can be hard to understand without doing some consolidation or masking of peripheral data points.

Lineage experience in the classic Data Catalog

The catalog connects with other data processing, storage, and analytics systems to extract lineage information. The information is combined to represent a generic, scenario-specific lineage experience in the catalog.



Your data estate might include systems doing data extraction, transformation (ETL/ELT systems), analytics, and visualization systems. Each of the systems captures rich static and operational metadata that describes the state and quality of the data within the systems boundary. The goal of lineage in a data catalog is to extract the movement, transformation, and operational metadata from each data system at the lowest grain possible.

The following example is a typical use case of data moving across multiple systems, where a catalog would connect to each of the systems for lineage.

- Data Factory copies data from on-prem/raw zone to a landing zone in the cloud.
- Data processing systems like Synapse, Databricks would process and transform data from landing zone to Curated zone using notebooks.
- Further processing of data into analytical models for optimal query performance and aggregation.
- Data visualization systems will consume the datasets and process through their meta model to create a BI Dashboard, ML experiments and so on.

Lineage granularity

The following section covers the details about the granularity of which the lineage information is gathered by Microsoft Purview. This granularity can vary based on the data systems supported in Microsoft Purview.

Entity level lineage: Sources > Process > Targets

- Lineage is represented as a graph, typically it contains source and target entities in Data storage systems that are connected by a process invoked by a compute system.
- Data systems connect to the catalog to generate and report a unique object referencing the physical object of the underlying data system for example: SQL Stored procedure,

notebooks, and so on.
- High fidelity lineage with other metadata like ownership is captured to show the lineage in a human readable format for source & target entities. for example: lineage at a hive table level instead of partitions or file level.

Column or attribute level lineage

Identify attributes of a source entity that is used to create or derive attributes in the target entity. The name of the source attribute could be retained or renamed in a target. Systems like Azure Data Factory (ADF) can do a one-one copy from on-premises environment to the cloud. For example: `Table1/ColumnA -> Table2/ColumnA`.

Process execution status

To support root cause analysis and data quality scenarios, we capture the execution status of the jobs in data processing systems. This requirement has nothing to do with replacing the monitoring capabilities of other data processing systems, neither the goal is to replace them.

Summary

Lineage is a critical feature to support quality, trust, and audit scenarios. The goal of a data catalog is to build a robust framework where all the data systems within your environment can naturally connect and report lineage. Once the metadata is available, the catalog can bring together the metadata provided by data systems to power data governance use cases.

Next steps

- Quickstart: Create a Microsoft Purview account in the Azure portal
- Quickstart: Create a Microsoft Purview account using Azure PowerShell/Azure CLI
- Use the Microsoft Purview governance portal

Data lineage user guide for classic Data Catalog

07/18/2025

This article provides an overview of the data lineage features in the classic Microsoft Purview Data Catalog.

To view data lineage in the Unified Catalog, you need to first search for a data product, then view data asset details.

Background

One of the platform features of Microsoft Purview is the ability to show the lineage between datasets created by data processes. Systems like Data Factory, Data Share, and Power BI capture the lineage of data as it moves. Custom lineage reporting is also supported via Atlas hooks and REST API.

Lineage collection

Metadata collected in Microsoft Purview from enterprise data systems are stitched across to show an end to end data lineage. Data systems that collect lineage into Microsoft Purview are broadly categorized into following three types:

- Data processing systems
- Data storage systems
- Data analytics and reporting systems

Each system supports a different level of lineage scope. Check the sections below, or your system's individual lineage article, to confirm the scope of lineage currently available.

Known limitations

- Database Views used as source of process activity(Azure Data Factory, Synapse Pipelines, Azure SQL Database, Azure Data Share) are currently captured as Database Table objects in Microsoft Purview. If the Database is also scanned, the View assets are discovered separately in Microsoft Purview. In this scenario, two assets with same name captured in Microsoft Purview, one as a Table with data lineage and another as a View.
- If a stored procedure contains drop or create statements, they are not currently captured in lineage.

Data processing systems

Data integration and ETL tools can push lineage into Microsoft Purview at execution time. Tools such as Data Factory, Data Share, Synapse, Azure Databricks, and so on, belong to this category of data processing systems. The data processing systems reference datasets as source from different databases and storage solutions to create target datasets. The list of data processing systems currently integrated with Microsoft Purview for lineage are listed in below table.

Expand table

Data processing system	Supported scope
Airflow	Airflow lineage
Azure Data Share	Share snapshot
Azure Data Factory	Copy activity Data flow activity Execute SSIS package activity
Azure SQL Database (Preview)	Lineage extraction for stored procedure executions
Azure Synapse Analytics	Copy activity Data flow activity

Data storage systems

Databases & storage solutions such as Oracle, Teradata, and SAP have query engines to transform data using scripting language. Data lineage information from views/stored procedures/etc is collected into Microsoft Purview and stitched with lineage from other systems. Lineage is supported for the following data sources via Microsoft Purview data scan. Learn more about the supported lineage scenarios from the respective article.

Expand table

Category	Data source
Azure	Azure Databricks
Database	Cassandra
	Db2
	Google BigQuery
	Hive Metastore Database

Category	Data source
	MySQL
	Oracle
	PostgreSQL
	Snowflake
	Teradata
Services and apps	Erwin
	Looker
	SAP ECC
	SAP S/4HANA

Data analytics and reporting systems

Data analytics and reporting systems like Azure Machine Learning and Power BI report lineage into Microsoft Purview. These systems will use the datasets from storage systems and process through their meta model to create BI Dashboards, ML experiments and so on.

Expand table

Data analytics & reporting system	Supported scope
Power BI	Datasets, Dataflows, Reports & Dashboards

Get started with lineage

Lineage in Microsoft Purview includes datasets and processes. Datasets are also referred to as nodes while processes can be also called edges:

- **Dataset (Node)**: A dataset (structured or unstructured) provided as an input to a process. For example, a SQL Table, Azure blob, and files (such as .csv and .xml), are all considered datasets. In the lineage section of Microsoft Purview, datasets are represented by rectangular boxes.
- **Process (Edge)**: An activity or transformation performed on a dataset is called a process. For example, ADF Copy activity, Data Share snapshot and so on. In the lineage section of Microsoft Purview, processes are represented by round-edged boxes.

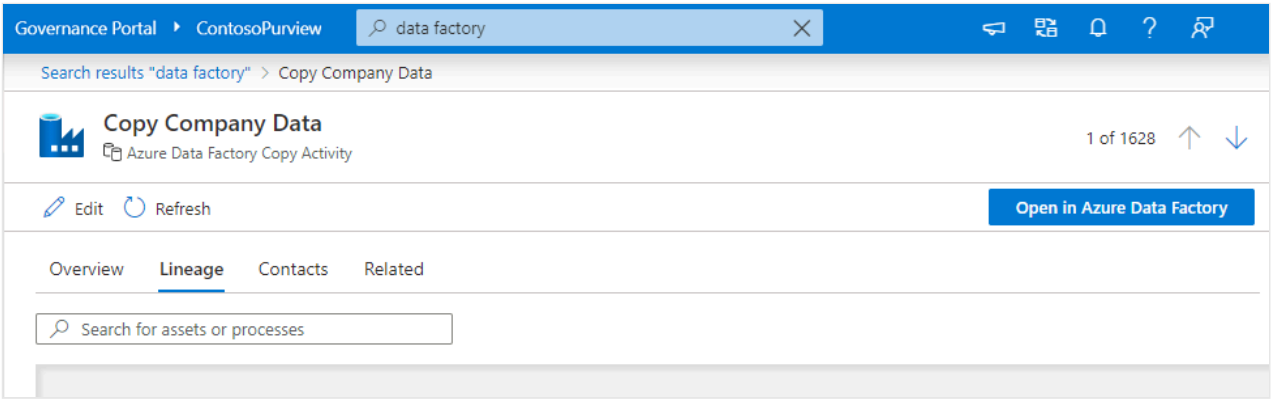
To access lineage information for an asset in Microsoft Purview, follow the steps:

1. Open the classic Microsoft Purview governance portal by:

    - Browsing directly to https://web.purview.azure.com and selecting your Microsoft Purview account.
    - Opening the Azure portal, searching for and selecting the Microsoft Purview account. Selecting the **the Microsoft Purview governance portal** button.

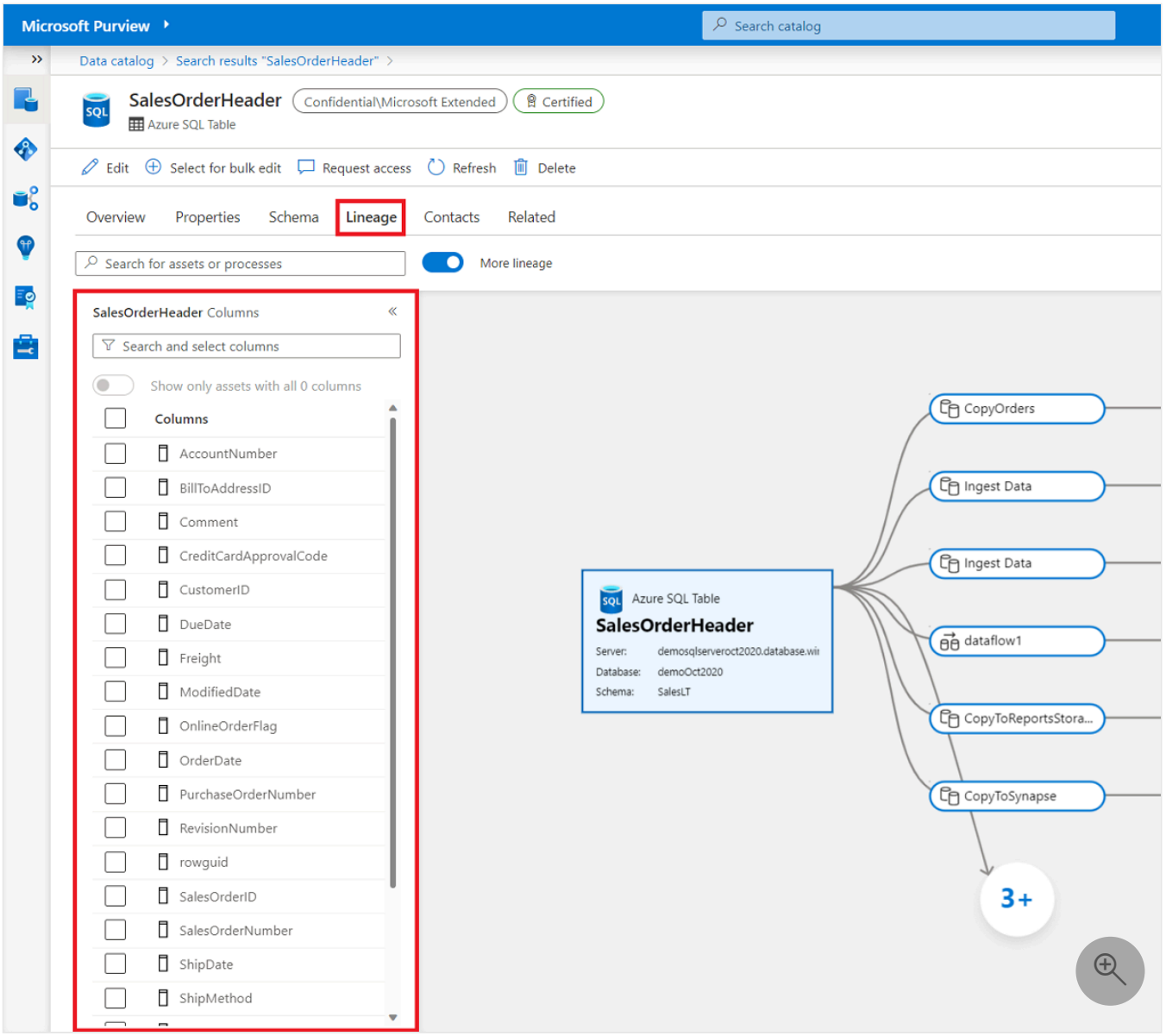
2. On the Microsoft Purview governance portal **Home** page, search for a dataset name or the process name such as ADF Copy or Data Flow activity. And then press Enter.

3. From the search results, select the asset and select its **Lineage** tab.



Asset-level lineage

Microsoft Purview supports asset level lineage for the datasets and processes. To see the asset level lineage go to the **Lineage** tab of the current asset in the catalog. Select the current dataset asset node. By default the list of columns belonging to the data appears in the left pane.



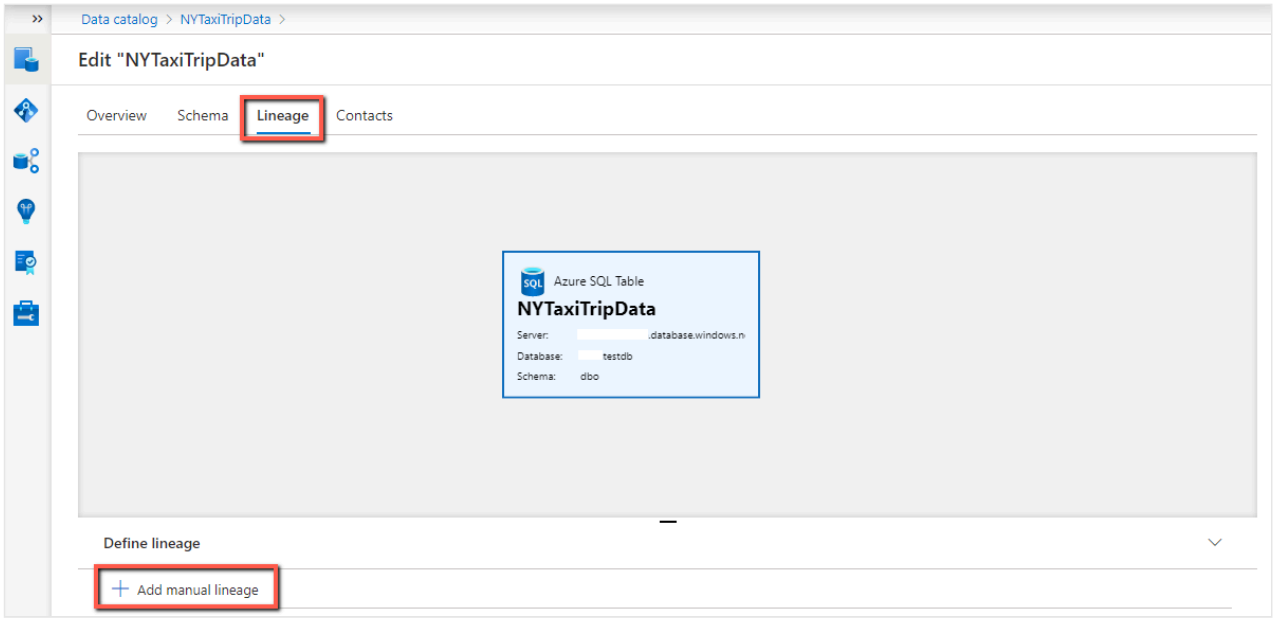
Manual lineage

Data lineage in Microsoft Purview is automated for many assets in on-premises, multicloud, and SaaS environments. While we continue to add more automated sources, manual lineage allows you to document lineage metadata for sources where automation isn't yet supported, without using any code.

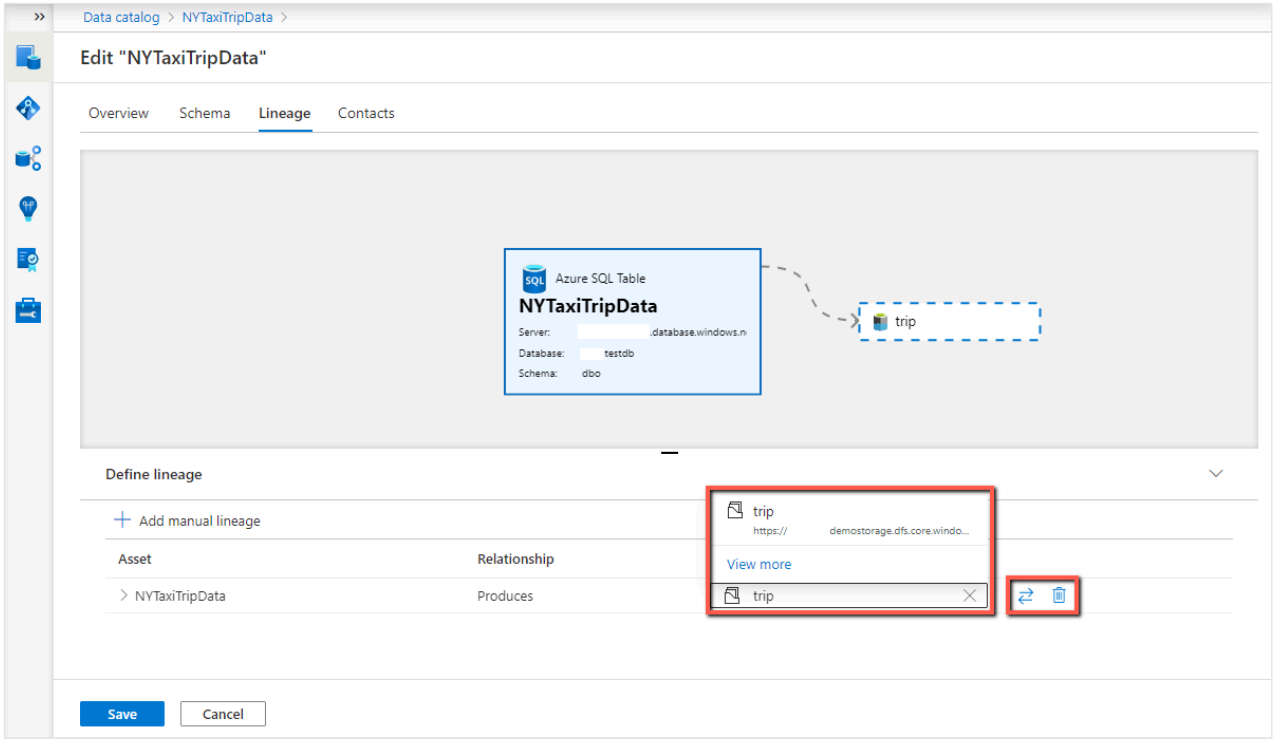
To add manual lineage for any of your assets, follow these steps:

1. Search for your asset in the classic Data Catalog and select it to view details.

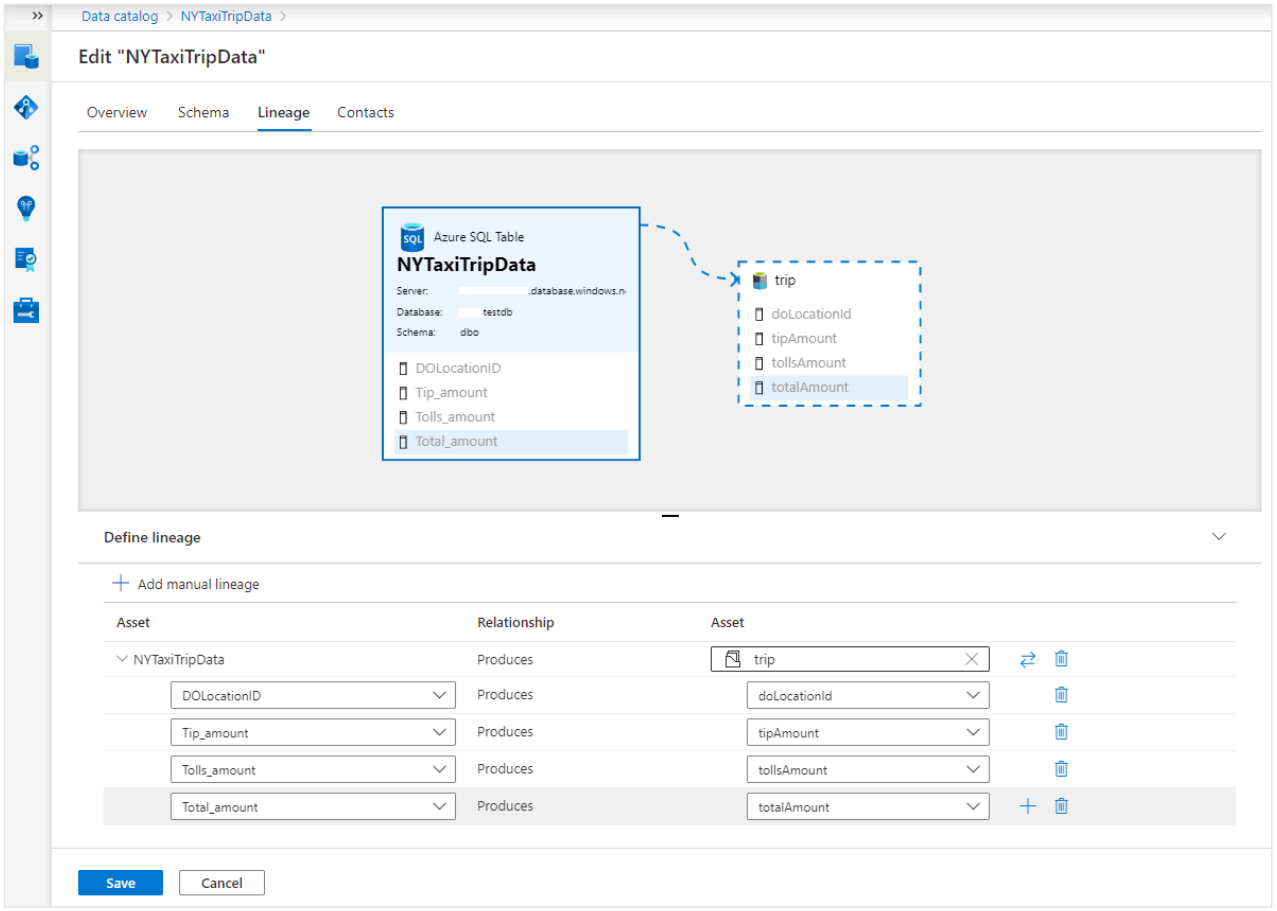
2. Select **Edit**, navigate to the **Lineage** tab, and select **Add manual lineage** in the bottom panel.



3. To configure the asset lineage:
   a. Select the asset dropdown to find the asset from the suggested list or **View more** to search the full catalog. Select the asset you'd like to link.
   b. Select the swap icon to configure the relationship direction as **Produces** (for downstream lineage) or **Consumes** (for upstream lineage).
   c. If you want to delete a lineage, select the trash can icon.



4. When you add lineage between two data assets, you can additionally configure the column level lineage. Select the expand icon at the beginning of the row, select the upstream and downstream columns from the corresponding dropdown lists to configure the column mapping. Select the plus icon to add more column lineage; select the trash bin icon to delete existing ones.



5. You can add more asset level lineage by selecting the **Add manual lineage** button again. When you're finished, select the **Save** button to save your lineage and exit edit mode.

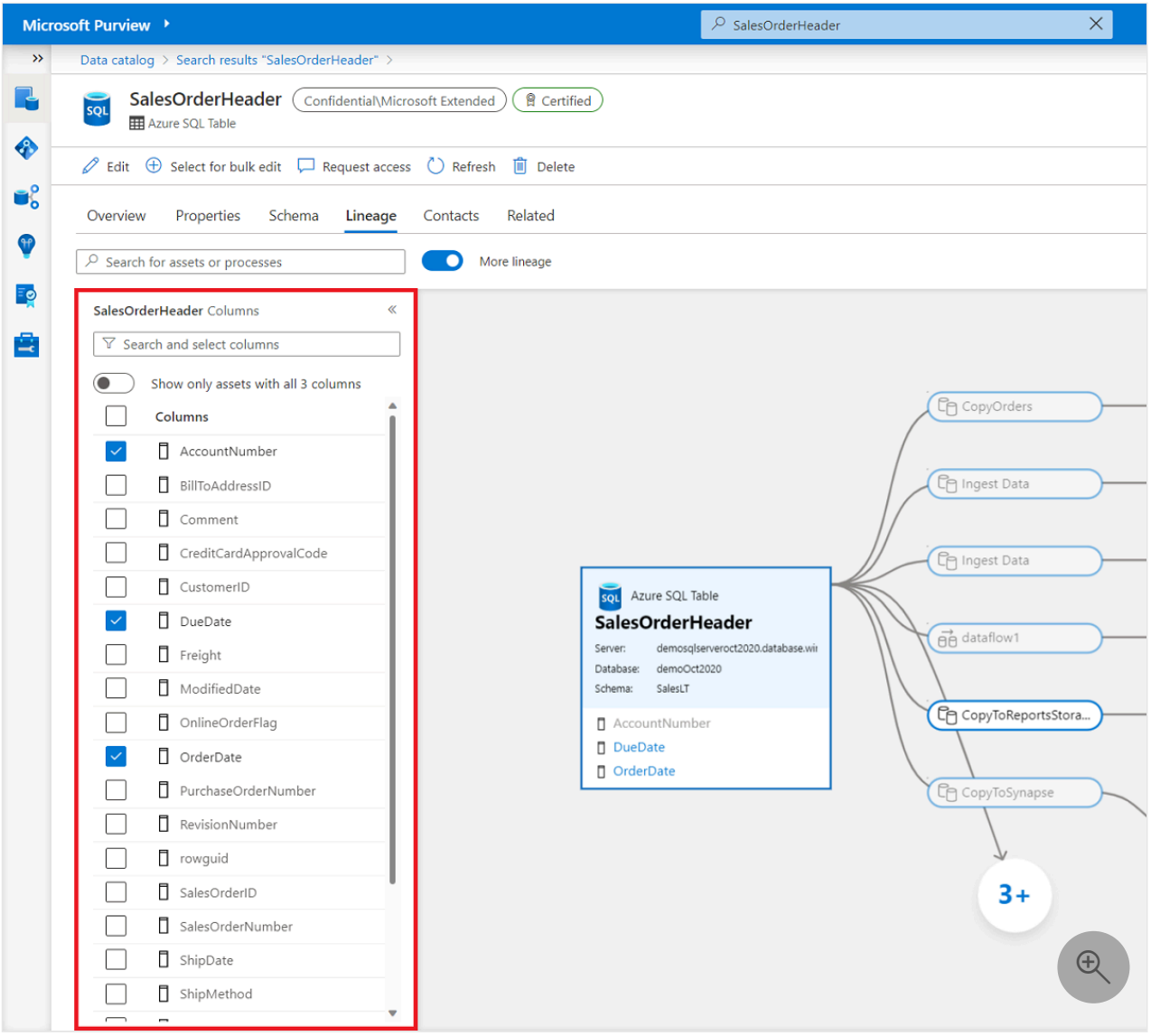
Known limitations of manual lineage

- Current asset picker experience allows selecting only one asset at a time.
- Column level manual lineage is currently supported for lineage between two data assets, while not supported when process asset is involved in-between.
- Data curation access required for both source and target assets.
- These asset types don't currently allow manual lineage because they support automated lineage:
  - Azure Data Factory
  - Synapse pipelines
  - Power BI datasets
  - Teradata stored procedure
  - Azure SQL stored procedure

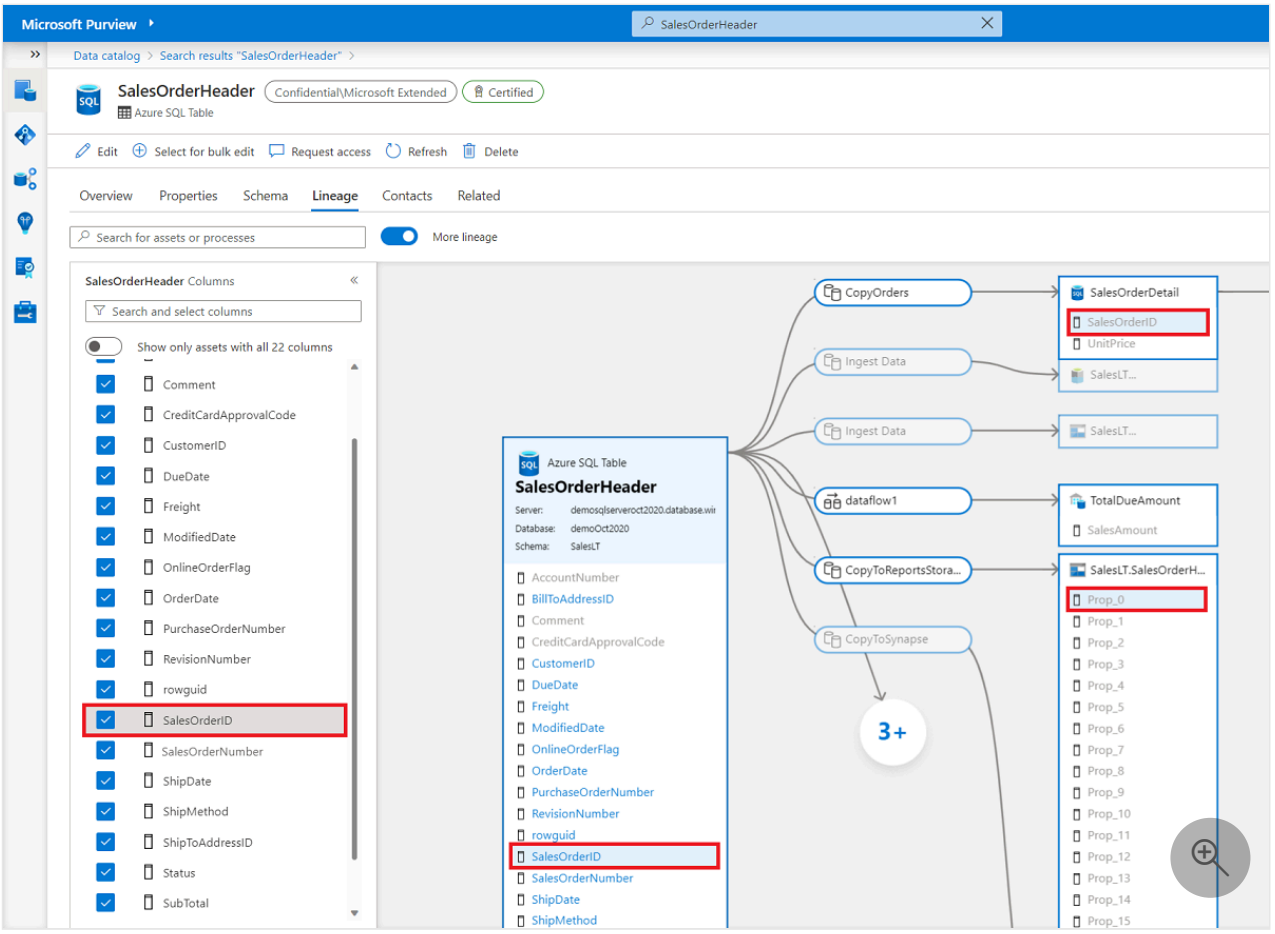
Dataset column lineage

To see column-level lineage of a dataset, go to the **Lineage** tab of the current asset in the catalog and follow below steps:

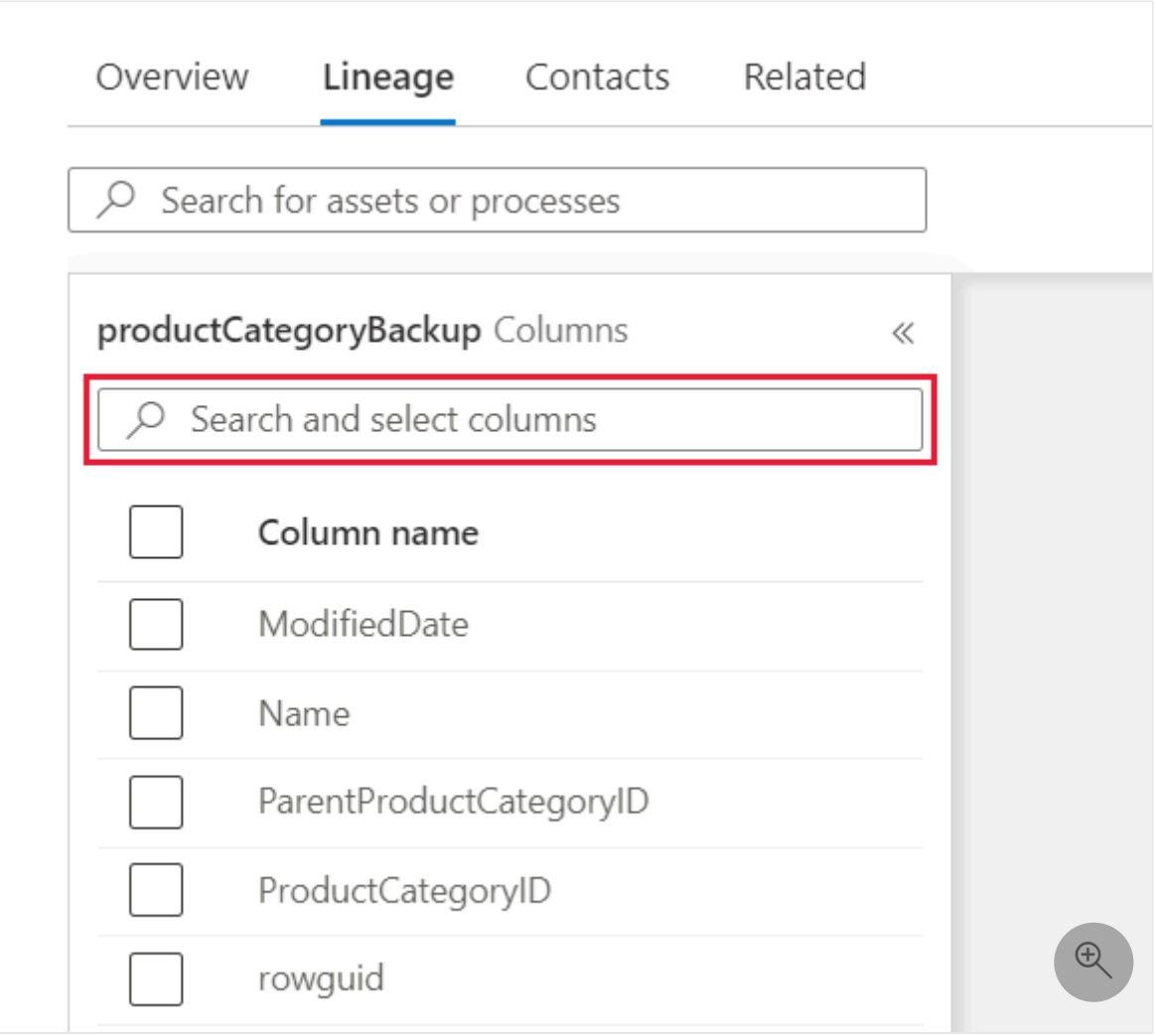
1. Once you are in the lineage tab, in the left pane, select the check box next to each column you want to display in the data lineage.



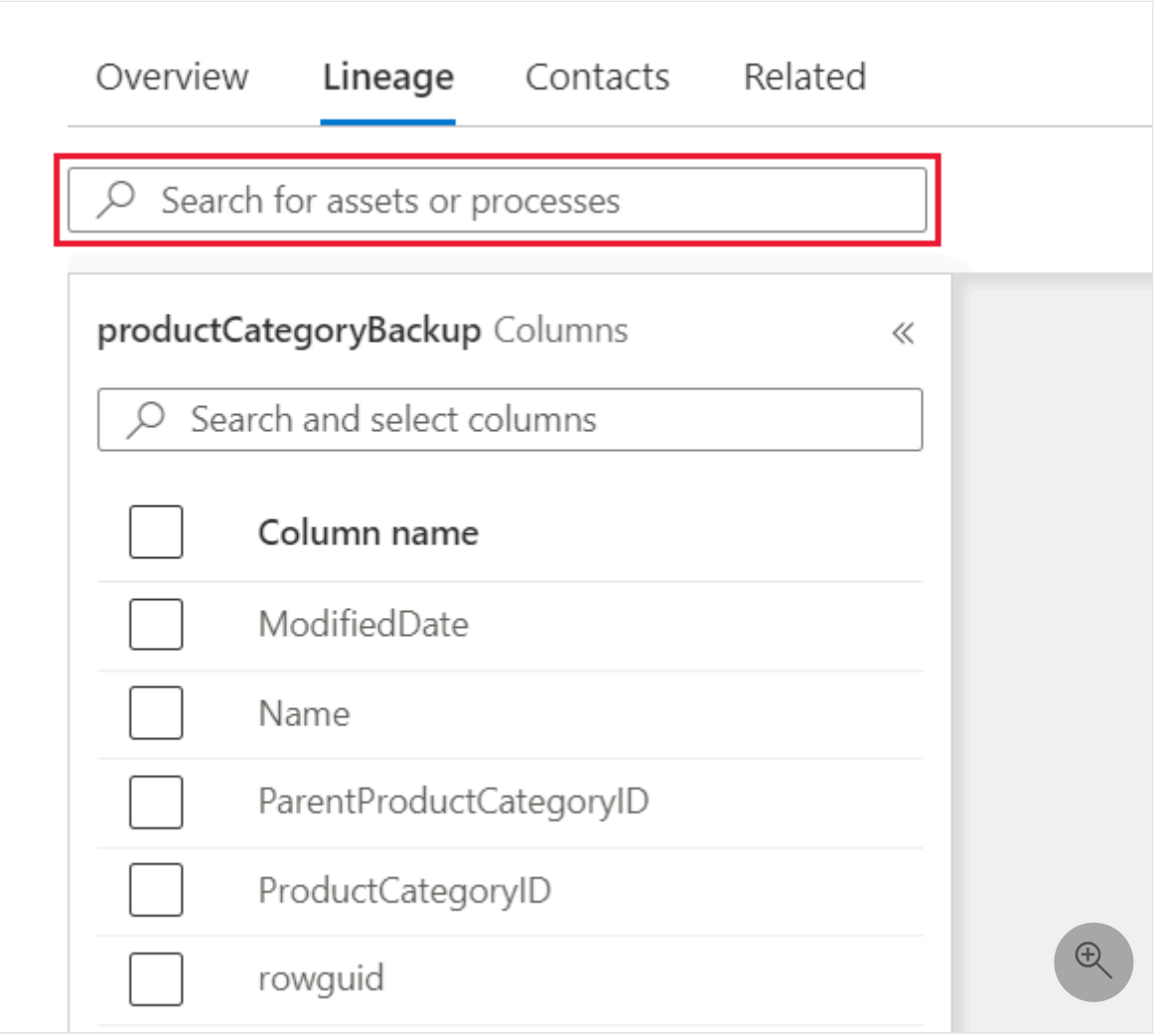
2. Hover over a selected column on the left pane or in the dataset of the lineage canvas to see the column mapping. All the column instances are highlighted.



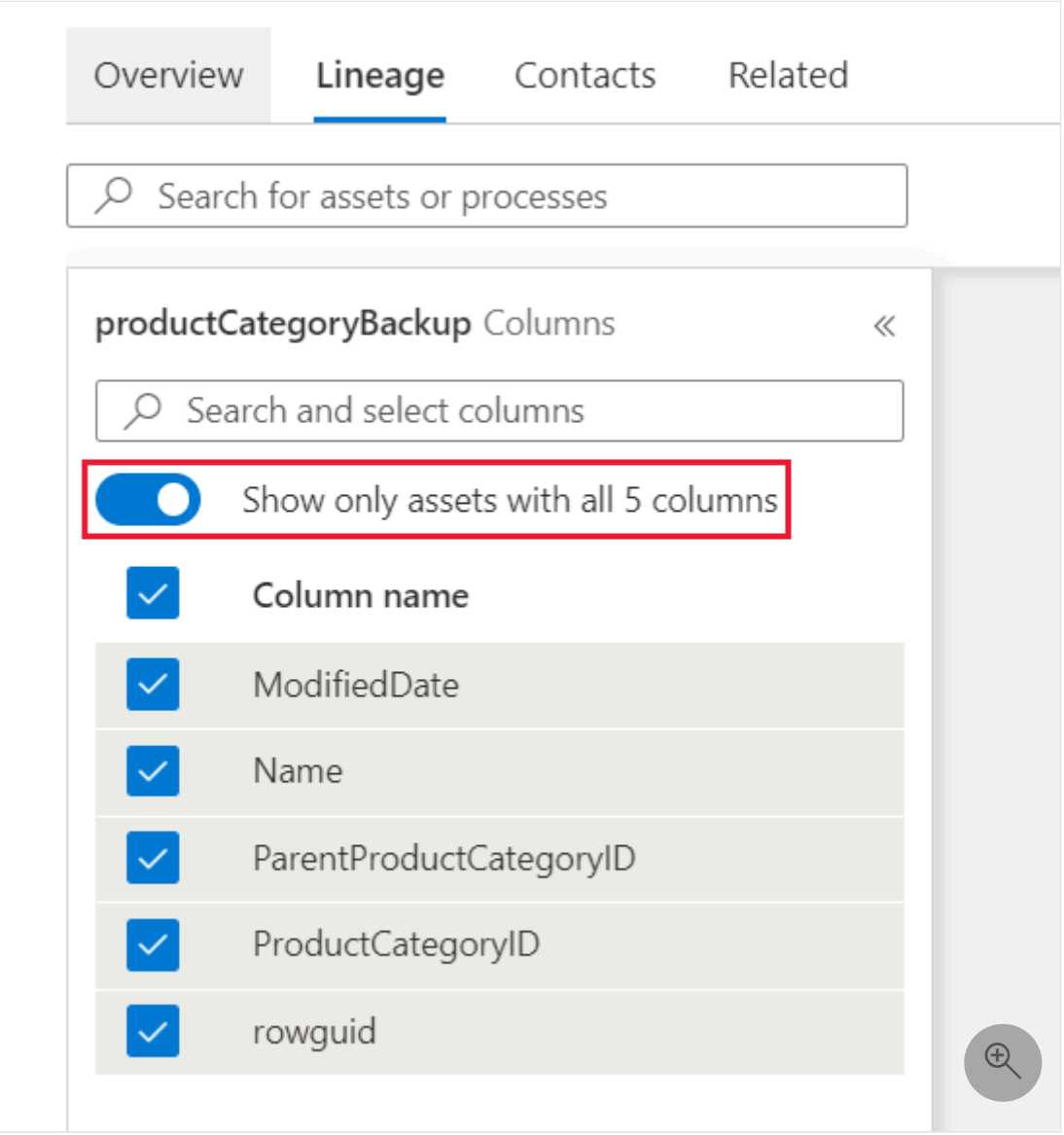
3. If the number of columns is larger than what can be displayed in the left pane, use the filter option to select a specific column by name. Alternatively, you can use your mouse to scroll through the list.



4. If the lineage canvas contains more nodes and edges, use the filter to select data asset or process nodes by name. Alternatively, you can use your mouse to pan around the lineage window.

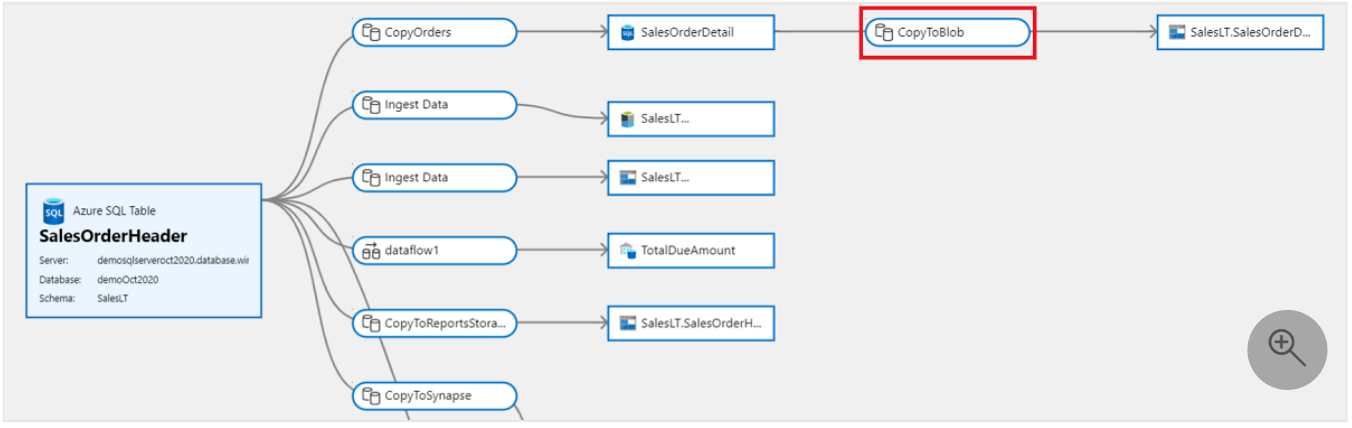


5. Use the toggle in the left pane to highlight the list of datasets in the lineage canvas. If you turn off the toggle, any asset that contains at least one of the selected columns is displayed. If you turn on the toggle, only datasets that contain all of the columns are displayed.

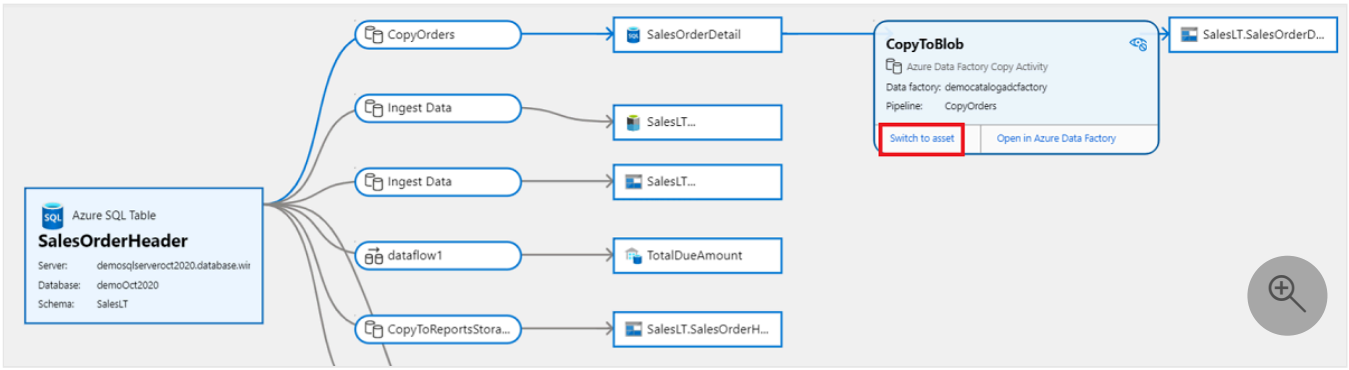


Process column lineage

You can also view data processes, like copy activities, in the catalog. For example, in this lineage flow, select the copy activity:

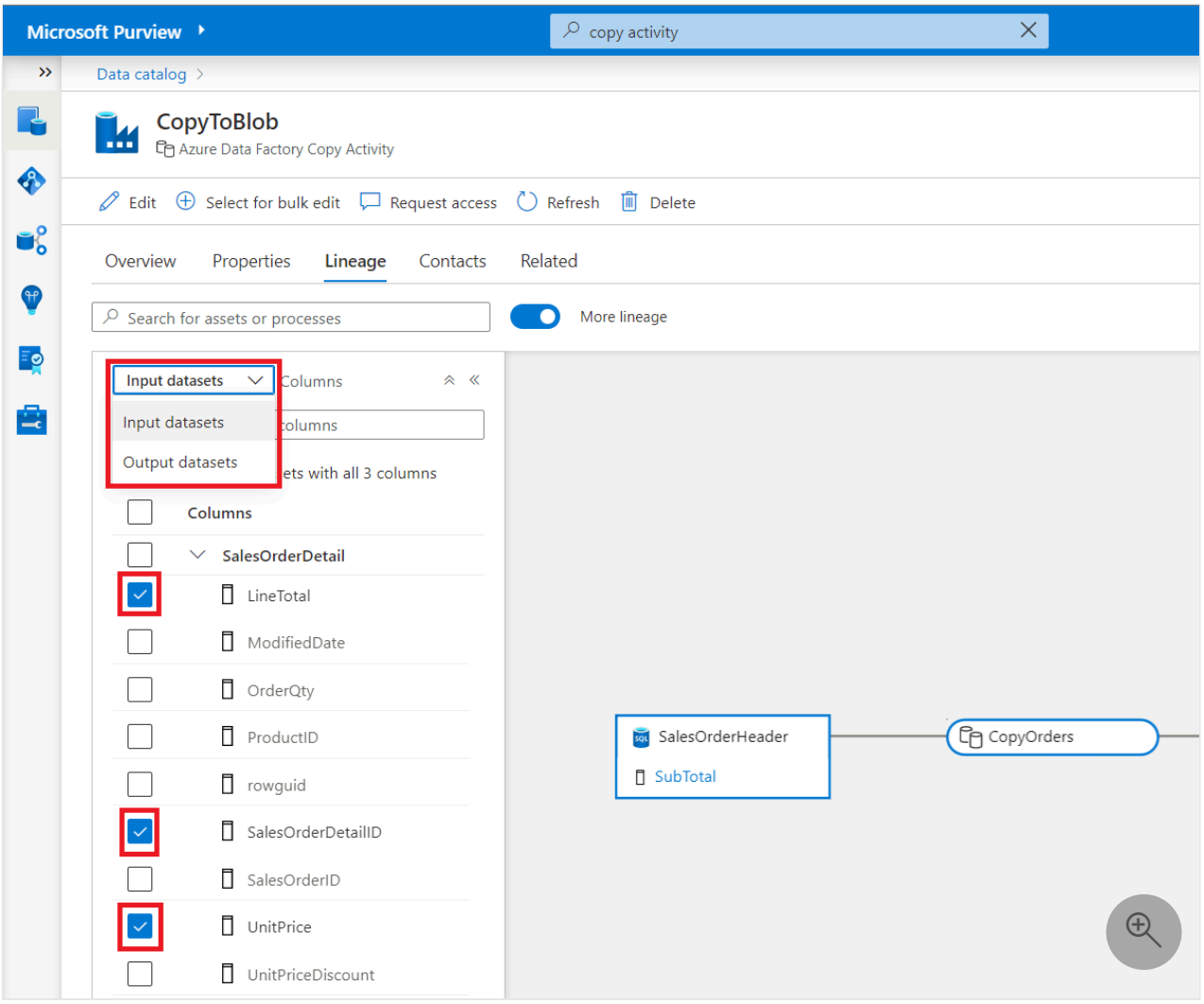


The copy activity will expand, and then you can select the **Switch to asset** button, which will give you more details about the process itself.



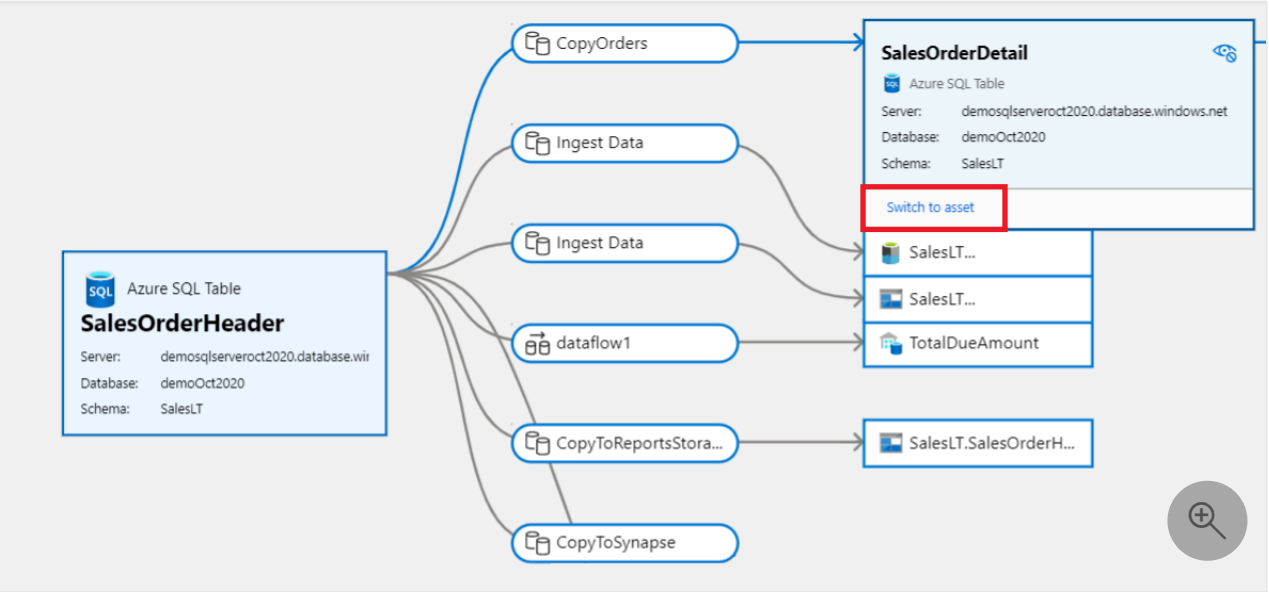
Data process can take one or more input datasets to produce one or more outputs. In Microsoft Purview, column level lineage is available for process nodes.

1. Switch between input and output datasets from a drop-down in the columns panel.
2. Select columns from one or more tables to see the lineage flowing from input dataset to corresponding output dataset.

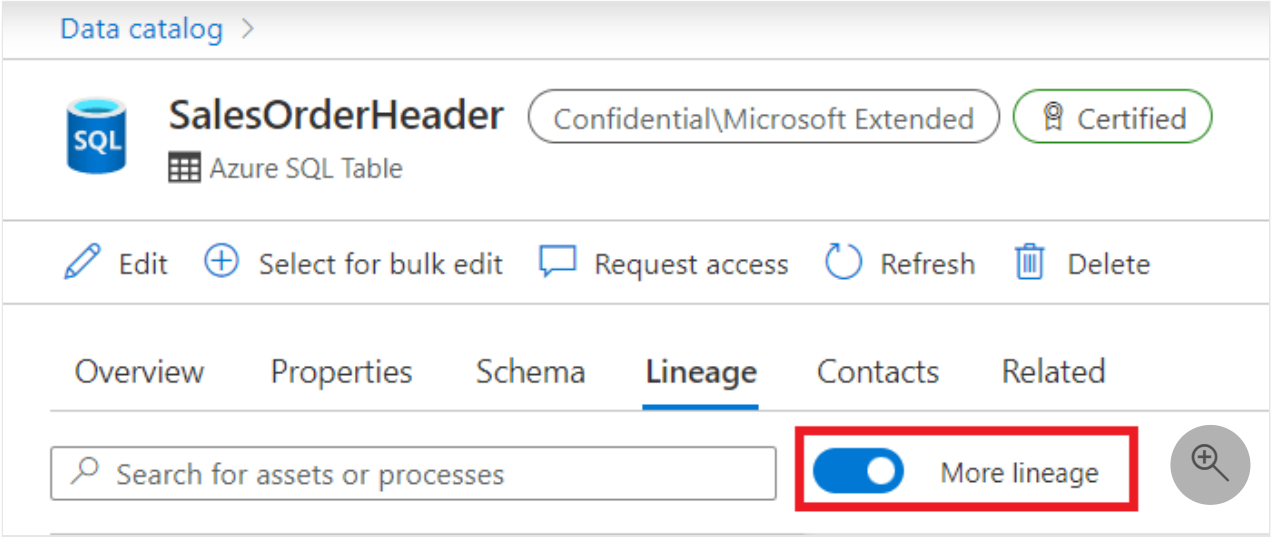


Browse assets in lineage

1. Select **Switch to asset** on any asset to view its corresponding metadata from the lineage view. Doing so is an effective way to browse to another asset in the catalog from the lineage view.

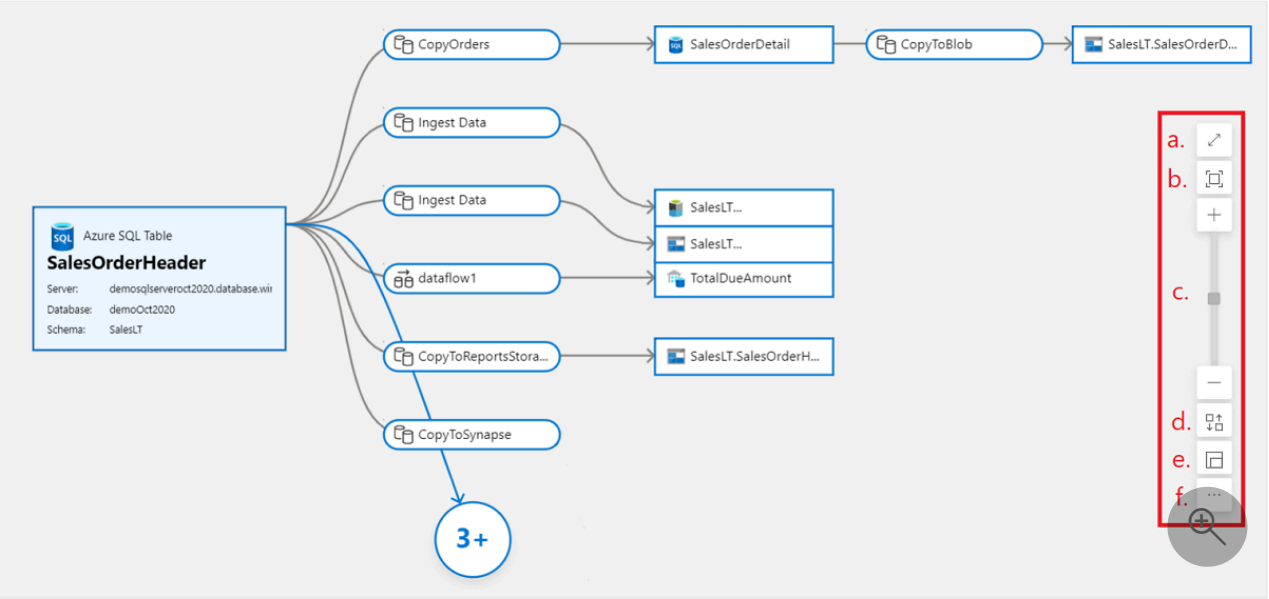


2. The lineage canvas could become complex for popular datasets. To avoid clutter, the default view will only show five levels of lineage for the asset in focus. The rest of the lineage can be expanded by selecting the bubbles in the lineage canvas. Data consumers can also hide the assets in the canvas that are of no interest. To further reduce the clutter, turn off the toggle **More Lineage** at the top of lineage canvas. This action will hide all the bubbles in lineage canvas.



3. Use the smart buttons in the lineage canvas to get an optimal view of the lineage:
   a. Full screen
   b. Zoom to fit
   c. Zoom in/out
   d. Auto align
   e. Zoom preview

f. And more options:
   i. Center the current asset
   ii. Reset to default view



Build custom lineage manually or with REST APIs

One of the important platform features of Microsoft Purview is the ability to show the lineage between datasets created by data processes. Systems like Data Factory, Data Share, and Power BI capture the lineage of data as it moves. In certain situations, automatically generated lineage by Purview is incomplete or missing for practical visualization and/or enterprise reporting purposes. In those scenarios, you can create custom lineage entries manually in the Microsoft Purview portal, or via Apache Atlas hooks and the REST API. Another major benefit of using REST APIs to report or build custom lineage is to overcome or mitigate the limitations of functionality exposed by Manual Lineage.

To build custom lineage manually, you can follow this user guide: Manual lineage entries in Microsoft Purview.

To build custom lineage in Microsoft Purview using the REST APIs, follow this user guide: Microsoft Purview - Building Custom Lineage using REST APIs.

> **Tip**
>
> In some cases, the REST APIs can provide more input and customization options than building the lineage entries manually through the portal.

Lineage best practices

Next steps

- Link to Azure Data Factory for lineage
- Link to Azure Data Share for lineage

Microsoft Purview - Build custom lineage with the REST APIs

Article • 02/24/2025

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

This article provides the steps to create data lineage entries using the REST API in the Microsoft Purview Data Catalog. In scenarios where the automatically generated lineage in Microsoft Purview is incomplete or missing, lineage can be custom built either manually in the Microsoft Purview portal or using the REST APIs. This article focuses on using REST APIs that can overcome manual lineage known limitations and provide more options.

Background

The ability to show the lineage between datasets is one of Microsoft Purview's important platform features. Systems like Data Factory, Data Share, and Power BI capture the lineage of data as it moves. In certain situations, Microsoft Purview's automatically generated lineage is incomplete or missing for practical visualization and enterprise reporting purposes. In those scenarios, custom lineage reporting is supported by Apache Atlas hooks and the REST API.

Using the REST APIs to build custom lineage allows you to overcome some of manual lineage's limitations as described in the following articles:

- Power BI lineage in Microsoft Purview
- Microsoft Purview manual lineage user guide
- Lineage - known limitations

The rest of this article explains the use of Microsoft Purview REST APIs to build and report custom lineage on Microsoft Purview.

Prerequisites

- For step-by-step guidance on how to use the Purview Python SDK Packages, refer Purview Python SDK Tutorial and Exploring Purview's REST API with Python
- For step-by-step guidance on how to use the *Purview REST APIs*, refer Purview Python SDK Tutorial

Scenarios

There are two use cases when building custom lineage becomes necessary:

A. Create new entities created and link them with lineage

B. Link existing entities or lineage to another existing entity or lineage

A. Create new entities and link them with lineage

As an example, lineage needs to be reported between entities A & B, but A & B don't exist currently.

To create the entities A & B, invoke the Microsoft Purview REST API: *Entity - Bulk Create Or Update - REST API*

HTTP

```
POST
https://{accountname}.purview.azure.com/datamap/api/atlas/v2/entity/bulk?
api-version=2023-09-01
```

Python

```
sample_entity_json = '{"entity": {"status": "ACTIVE","version": 0,"name":
ENTITY_A"}.......{"entity": ........}}'
#Send POST JSON containing entities to be created
CreateOrUpdateEntitesUrl =
'https://<purview_account_name>.purview.azure.com/datamap/api/atlas/v2/entit
y/bulk'
EntitiesResponse = requests.post(CreateOrUpdateEntitesUrl, json =
json.loads(sample_entity_json) ,headers=headers)
entitiesRes = json.loads(EntitiesResponse.text)
```

API Response "201 Created" indicates entities are successfully created and their respective GUIDs are contained in the output JSON.

Now that the entities A & B are created, move to step B to link the entities in the lineage chain using the same REST API.

B. Link existing entities or lineage to another existing entity or lineage

- If the number of entities to be linked isn't time or resource intensive (for example, less than 20-30 entities), you can connect the lineage manually in the Microsoft Purview portal. Follow the manual lineage user guide for steps to manually create lineage connections.
- If you have a high number of lineage connections to make, the process needs to be automated, or if manual lineage using the Microsoft Purview portal isn't feasible, proceed with the API process of linking and building custom lineage.

Custom Lineage JSON Payload:

Execute the *POST /entity/bulk* *Entity - Bulk Create Or Update - REST API* using the payload as illustrated:

HTTP

```
POST
https://{accountname}.purview.azure.com/datamap/api/atlas/v2/entity/bulk?
api-version=2023-09-01
```

Python

```
sample_entity_json = '{
  "entities": [
    {
      "status": "ACTIVE",
      "version": 1,
      "typeName": "Process",
      "attributes": {
        "inputs": [
          {
            "guid": "24558fd8-9cdc-47de-9310-56a58108bab0",
            “guid”: “27163581-9aca-212a-782a-213612639abc”
          }
        ],
        "outputs": [
          {
            "guid": "e33c694a-2c4f-4cae-8c27-06f6f6f60000"
          }
        ],
        "qualifiedName": "cassandra://query",
        "name": "query"
      }
    }
  ]
}'
```

```
#In this code snippet, we send the JSON as POST request containing the two GUIDs as input and "output" GUID as output. This creates lineage with 2 directional inputs and 1 directional output.
#Note: using the same API and SDK code you can create lineage with any number of inputs, any number of processes in between, any number of typedefs and any number of outputs.
#The API/SDK method is the most flexible and versatile menthod of creating lineage.

CreateLineageEntitesUrl = 'https://<purview_account_name>.purview.azure.com/datamap/api/atlas/v2/entity/bulk'
EntitiesResponse = requests.post(CreateLineageEntitesUrl, json = json.loads(sample_entity_json),headers=headers)
entitiesRes = json.loads(EntitiesResponse.text)
```

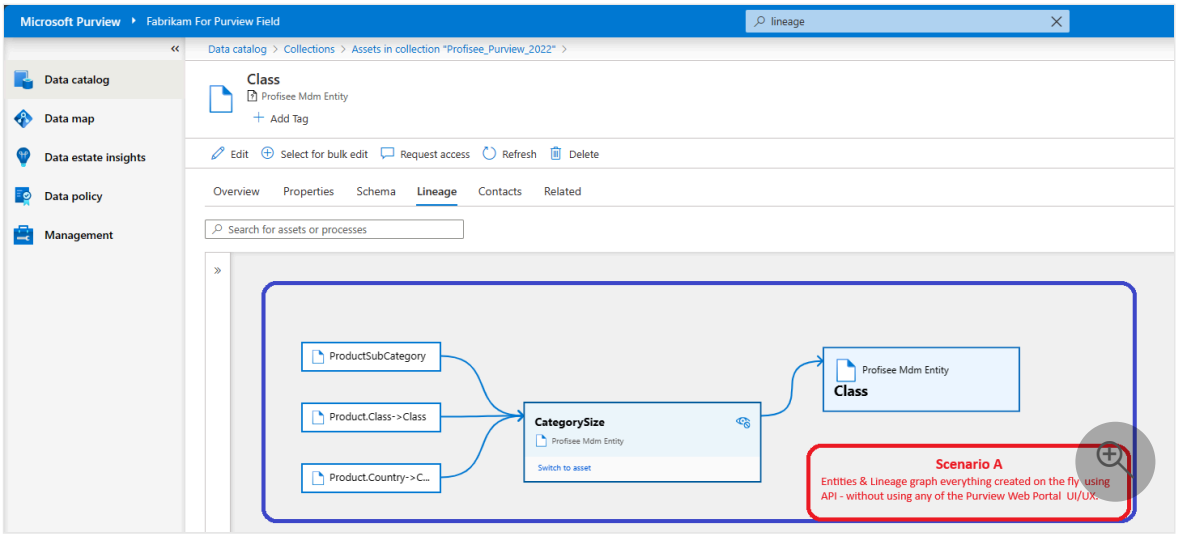
This JSON payload creates the custom lineage. It works for already existing assets whose GUIDs are supplied in the "inputs" JSON. For example "guid": "24558fd8-9cdc-47de-9310-56a58108bab0" and "27163581-9aca-212a-782a-213612639abc" refers to the directional input of the lineage and "guid": "e33c694a-2c4f-4cae-8c27-06f6f6f60000" refers to the directional output of the lineage, which was an existing asset automatically scanned by Purview. We just created the lineage between the two assets.

> ⓘ **Note**
>
> If the assets don't exist already, you need to run Bulk Entity create API before this step to create those entities before creating the lineage relationship. The bulk entity creation process using the POST /entity/bulk API and Python code snippet is described in step A.

Scenario Outcomes

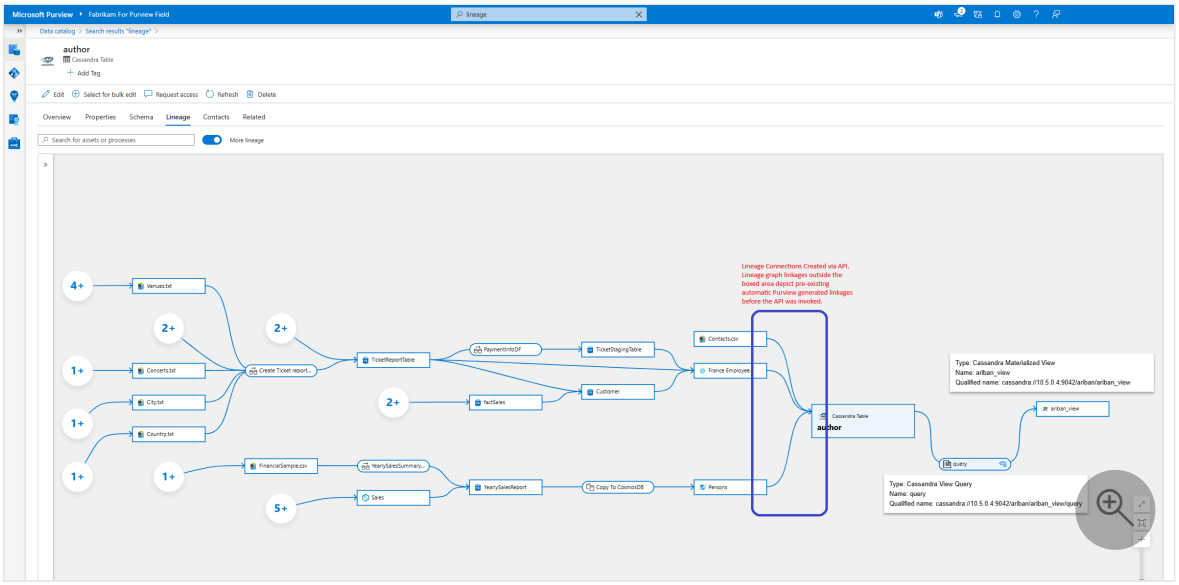
API Response "201 Created" indicates successful lineage graph linkage creation and the created GUIDs are contained in the output JSON. Lineage appears in the Microsoft Purview portal:

- Scenario A: Custom built lineage from assets created via API:



- Scenario B: Custom built lineage from pre-existing assets linked via API.

> **Note**
>
> If the lineage is created from pre-existing entities, observe that the pre-existing lineage graph remains intact and the new linkage is created and displayed additionally.


Microsoft Purview
Fabrikam For Purview Field
lineage
author
Cassandra Table
Edit
Select for bulk edit
Request access
Refresh
Delete
Overview
Properties
Schema
Lineage
Contacts
Related
Search for assets or processes
More lineage
Lineage Connections Created via API.
Lineage graph linkages outside the boxed area depict pre-existing automatic Purview generated linkages before the API was invoked.
Venues.txt
Concerts.txt
City.txt
Country.txt
Create Ticket report...
TicketReportTable
PaymentInfoDF
TicketStagingTable
Contacts.csv
France Employee
Customer
factSales
FinancialSample.csv
YearlySalesSummary...
Sales
YearlySalesReport
Copy To CosmosDB
Persons
Cassandra Table
author
Type: Cassandra Materialized View
Name: ariban_view
Qualified name: cassandra://10.5.0.4:9042/aribanaribanaribana

How to get lineage from Airflow into Microsoft Purview (preview)

07/18/2025

Airflow is an open-source workflow automation and scheduling platform that can be used to author and manage data pipelines. Microsoft Purview supports collecting Airflow lineage by integrating with OpenLineage, an open framework for data lineage collection and analysis. Learn about how Airflow works with OpenLineage from here.

Enabling OpenLineage in Airflow automatically tracks metadata and lineage about jobs and datasets as DAGs execute. The information is sent to an Azure Event Hubs that you configure. Microsoft Purview subscribes to the events, parse them and ingest into the data map.

> **Important**
>
> This feature is currently in preview. The Supplemental Terms of Use for Microsoft Azure Previews include additional legal terms that apply to Azure features that are in beta, in preview, or otherwise not yet released into general availability.

Supported capabilities

The supported Airflow versions are 1.10+ and 2.0-2.7.

Microsoft Purview supports metadata and lineage collection when the following types of data sources are used in Airflow:

- Amazon RDS for PostgreSQL
- Azure Database for PostgreSQL
- Google BigQuery
- PostgreSQL
- Snowflake

The following Airflow metadata are captured along:

- Airflow workspace
- Airflow DAG
- Airflow task

Lineage is collected into Microsoft Purview upon successful DAG runs in an event-based manner.

Known limitations

- Column level lineage is currently not supported. The schema of the data assets is captured.
- If database views are referenced in the tasks, they're currently captured as table assets.
- All the metadata are ingested into Microsoft Purview root collection. The assets already existed in the data map are preserved in the configured collection.

How to bring Airflow lineage into Microsoft Purview

As a prerequisite, you need a running Airflow instance.

To get lineage from Airflow into Microsoft Purview, you need to:

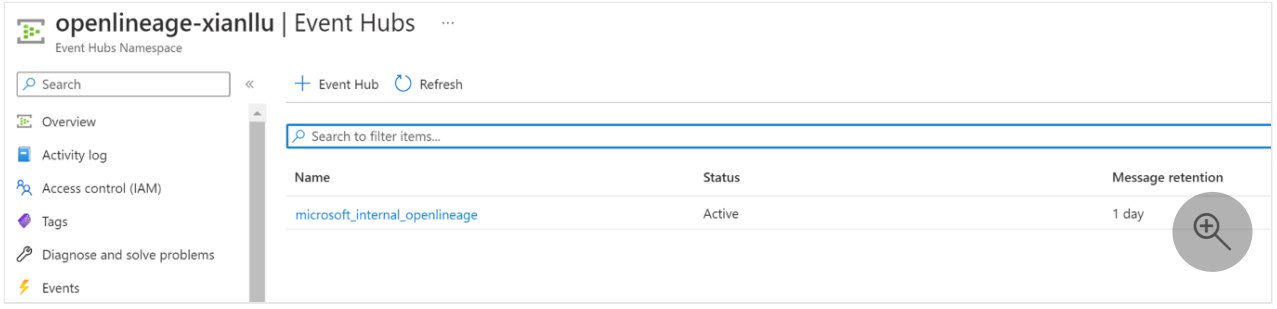
1. Set up an Azure Event Hubs
2. Configure Event Hubs to publish messages to Microsoft Purview
3. Configure your Airflow with OpenLineage
4. Run Airflow jobs and view the assets/lineage

Set up Azure Event Hubs

Set up an Azure Event Hubs as the receiver of the metadata and lineage tracked by OpenLineage in Airflow.

1. Create an Azure Event Hubs namespace.

2. Create an event hub. Name your event hub as "**microsoft_internal_openlineage**".



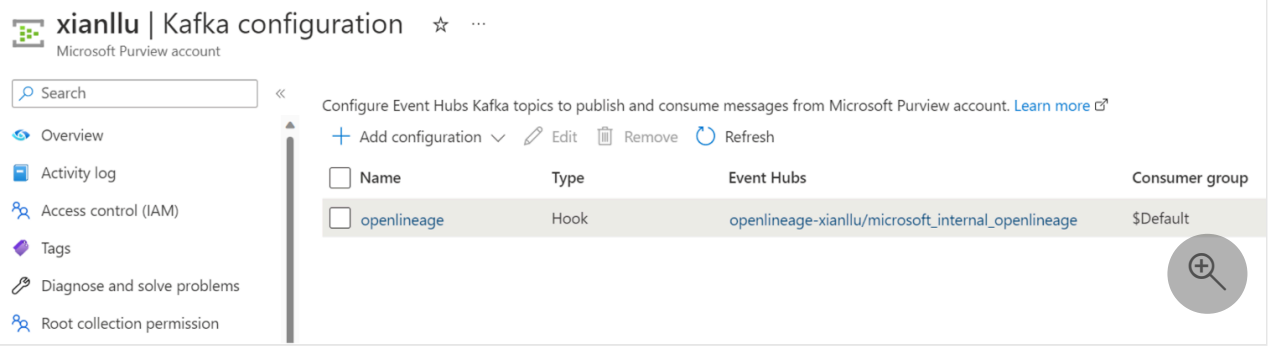
3. Go to your "microsoft_internal_openlineage" event hubs -> **Access control (IAM)** -> **Add role assignment**, assign "**Azure Event Hubs Data Receiver**" role to your Microsoft Purview account's managed identity. For detailed steps, see Assign Azure roles using the Azure portal.

Configure Event Hubs to publish messages to Microsoft Purview

Microsoft Purview supports consuming and pushing events from/to your own Event Hubs. Follow this doc to configure Event Hubs for Microsoft Purview: Configure Event Hubs with Microsoft Purview to send and receive Atlas Kafka topics messages.

In summary:

- Go to your Microsoft Purview account -> Managed resources tab, disable the managed Event Hubs namespace.
- Go to Kafka configuration tab -> **+ Add configuration** -> **Hook configuration**, input a name, and select the Event Hubs namespace and Event Hubs you created in previous step.



Configure your Airflow with OpenLineage

Installation:

To download and install the latest 'openlineage-airflow' library, update the 'requirements.txt' file of your running Airflow instance with:

```
openlineage-airflow
```

> **Note**
>
> The Airflow version and the openlineage-airflow version need to match. For example, when you use Airflow 2.7.1, you can use openlineage-airflow version 1.1.0 or 1.2.0. You can view matching versions on this website.

Configuration:

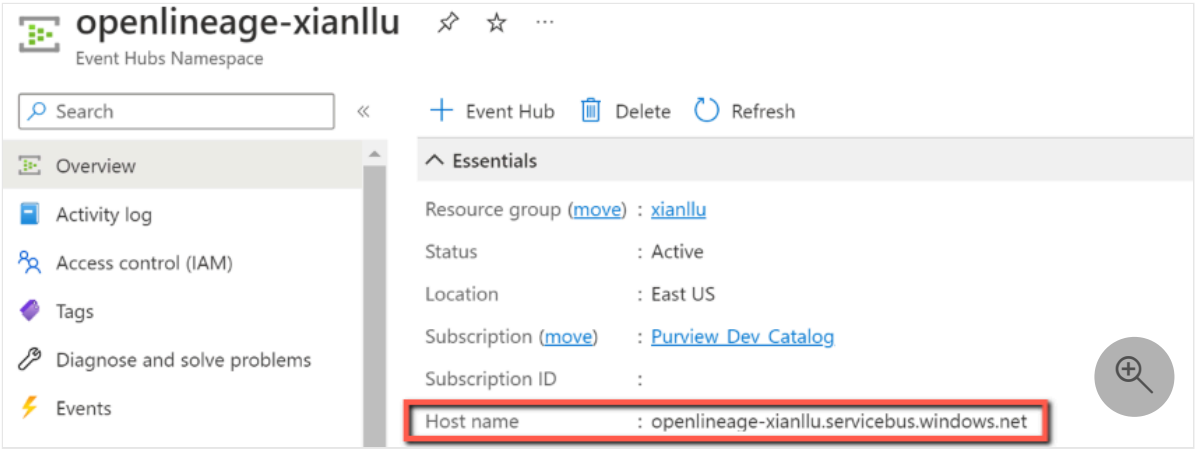
Next, configure your Azure Event Hubs instance as the target to which OpenLineage sends the events.

1. Create an 'openlineage.yml' file under your Airflow root path. The content of the file is as below:

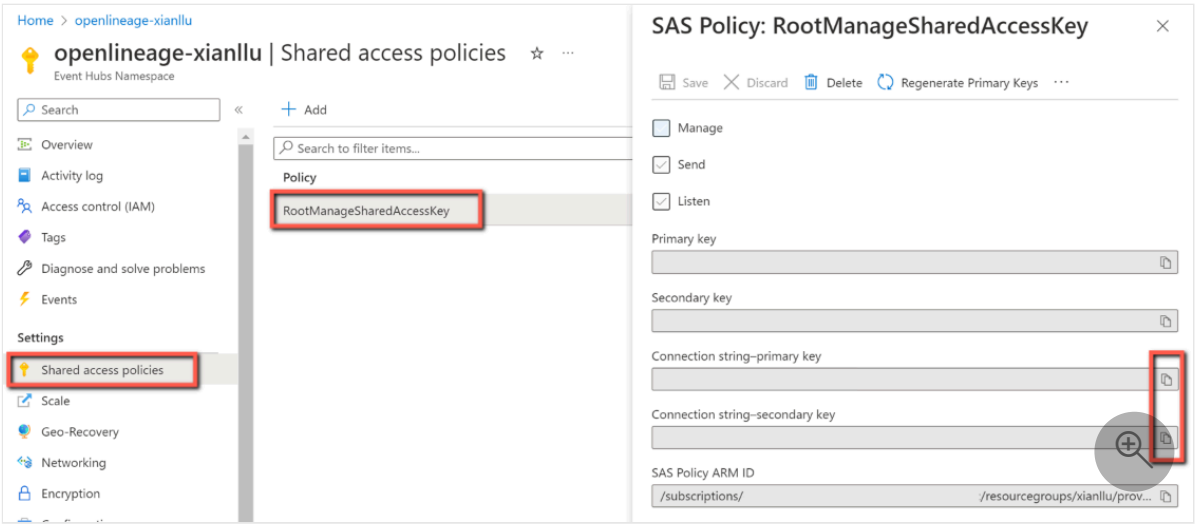
```yml
transport:
  type: "kafka"
  config:
    bootstrap.servers: "{EVENTHUB_SERVER}:9093"
    security.protocol: "SASL_SSL"
    sasl.mechanism: "PLAIN"
    sasl.username: "$ConnectionString"
    sasl.password: "{PASSWORD}"
    client.id: "airflow-client"
  topic: "microsoft_internal_openlineage"
  flash: True
```

   Replace the two place-holders with values:

   - Get the `{EVENTHUB_SERVER}` from your Event Hubs namespace -> **Host name**:



   - Get the `{PASSWORD}` from your Event Hubs namespace -> **Shared access policies** -> **Connection string - primary key** or **Connection string - secondary key**:

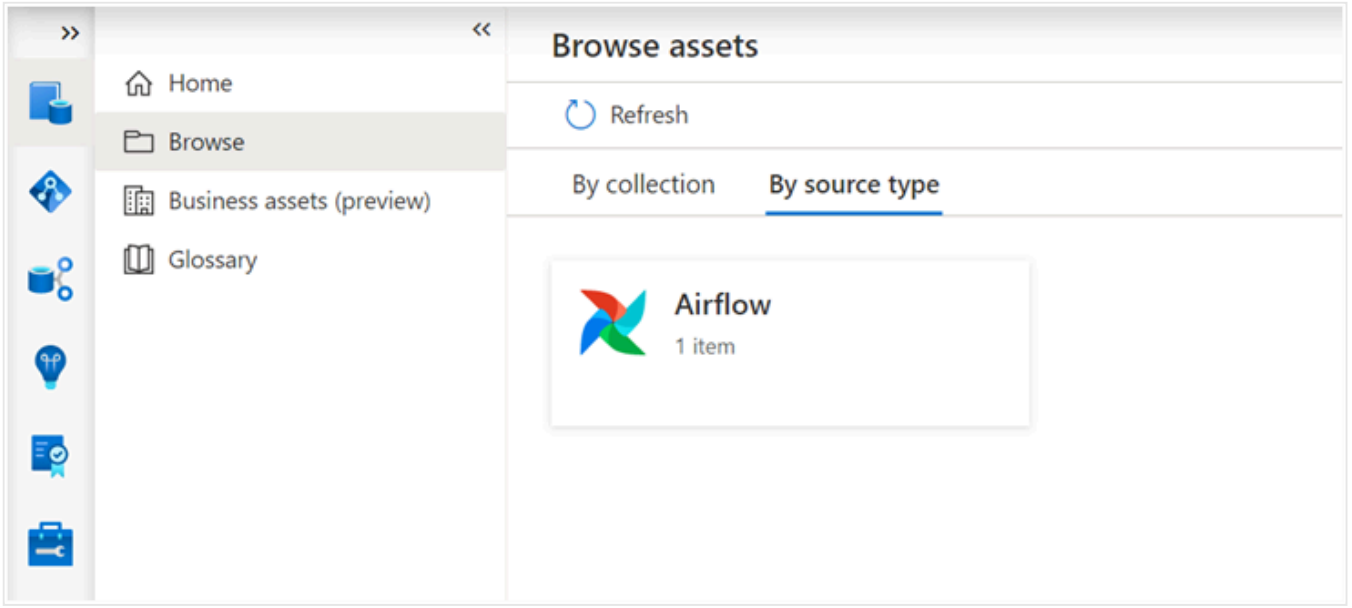


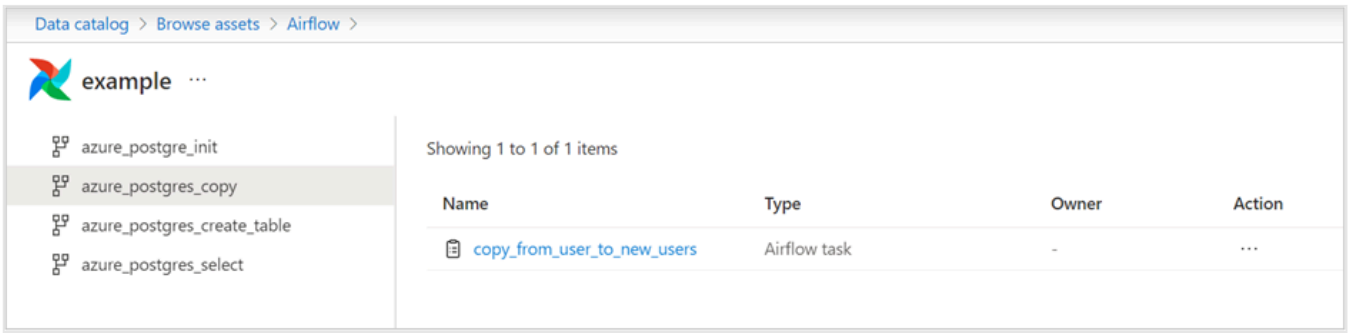
2. Restart your Airflow server.

Run Airflow jobs and view the assets/lineage

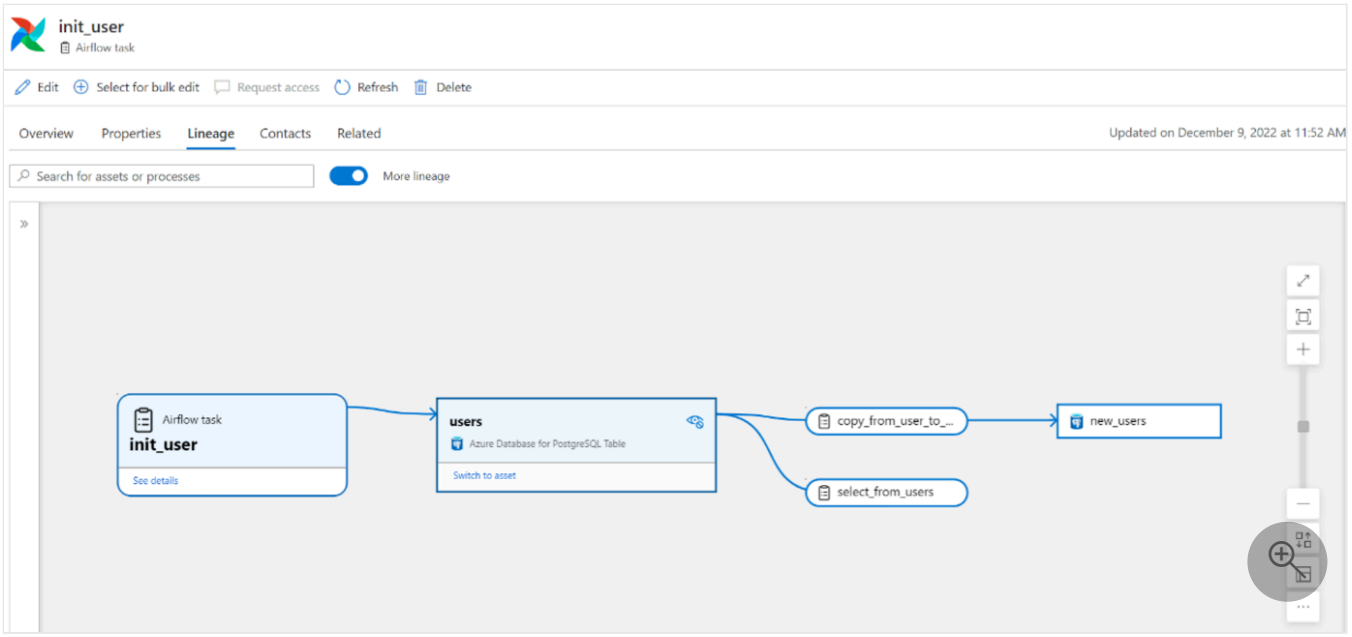
You can now run your Airflow jobs, then go to Microsoft Purview Governance portal to browse/search/view assets. The assets should show up shortly after a successful DAG run.

Browse Airflow assets:





View Airflow task asset details with lineage:



Troubleshooting tips

If you run the Airflow job but don't see the corresponding assets/lineage shown up in Microsoft Purview:

- Check if your Airflow use case is supported by Microsoft Purview. Refer to the supported capabilities section.
- Go to your Event Hubs instance to check if any incoming requests and messages. If no, double check your OpenLineage configuration in Airflow.

Next steps

- Learn about Data lineage in Microsoft Purview

How to connect Azure Data Factory and Microsoft Purview

Summarize this article for me

This article explains the steps required for connecting an Azure Data Factory account with a Microsoft Purview account to track data lineage and ingest data sources. The article also explains details of the activity coverage scope and supported lineage patterns.

When you connect an Azure Data Factory to Microsoft Purview, whenever a supported Azure Data Factory activity runs, it automatically ingests metadata about the activity's source data, output data, and the activity into Microsoft Purview Data Map.

If a data source is already scanned and exists in Data Map, the ingestion process adds the lineage information from Azure Data Factory to that existing source. If the source or output doesn't exist in Data Map and is supported by Azure Data Factory lineage, Microsoft Purview automatically adds their metadata from Azure Data Factory into Data Map under the root collection.

This connection is an excellent way to monitor your data estate as users move and transform information by using Azure Data Factory.

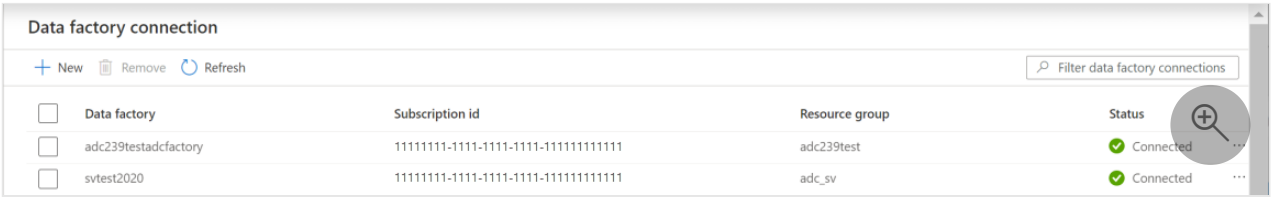
View existing Data Factory connections

Multiple Azure Data Factorys can connect to a single Purview to push lineage information. The current limit allows you to connect up to 10 Azure Data Factory accounts at a time from the Microsoft Purview management center. To show the list of Azure Data Factory accounts connected to your Microsoft Purview account, follow these steps:

1. Select **Management** on the left navigation pane.

2. Under **Lineage connections**, select **Data Factory**.

3. The Data Factory connection list appears.



4. Notice the various values for connection **Status**:

    - **Connected**: The Data Factory is connected to the Microsoft Purview account.

- **Disconnected**: The Data Factory has access to the catalog, but it's connected to another catalog. As a result, data lineage won't be reported to the catalog automatically.
- **Unknown**: The current user doesn't have access to the Data Factory, so the connection status is unknown.

> ⓘ Note
>
> To view the Data Factory connections, you need to be assigned the following role. Role inheritance from management group isn't supported. **Collection admins** role on the root collection.

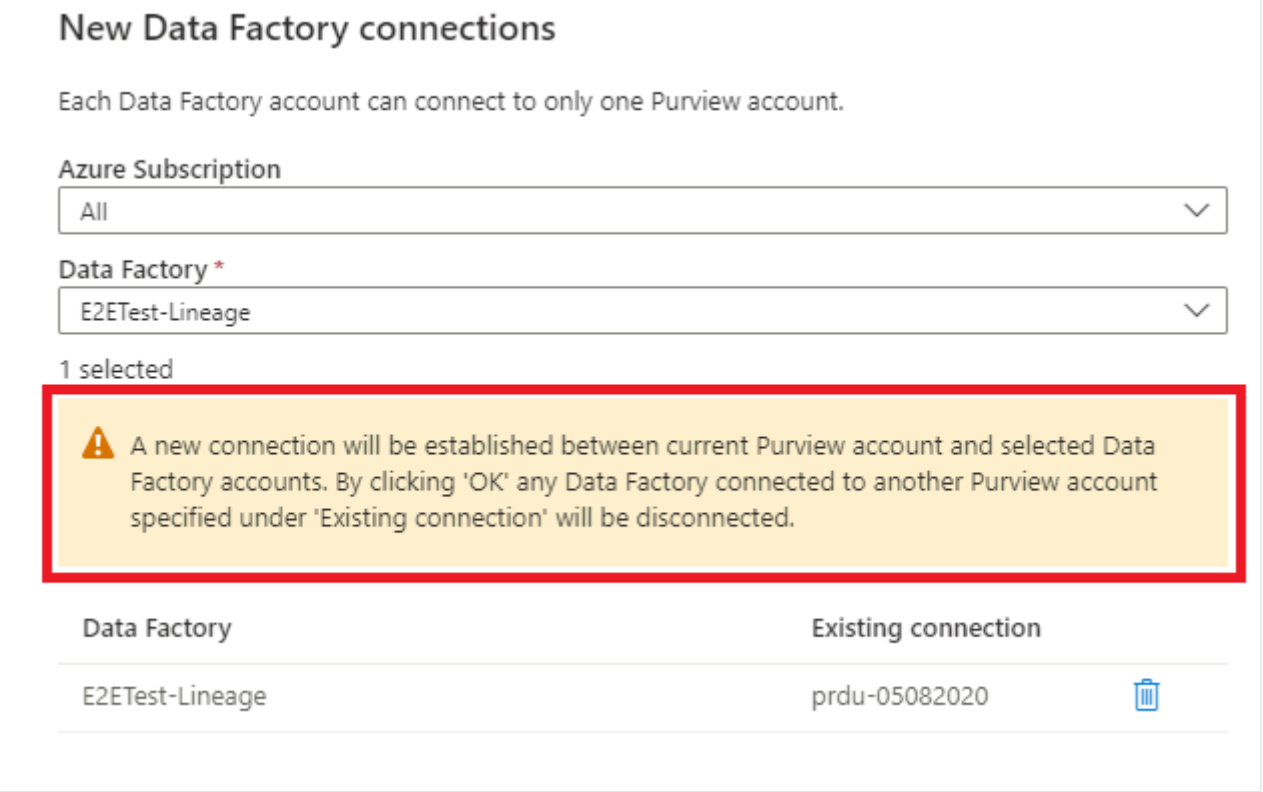
Create new Azure Data Factory connection

> ⓘ Note
>
> To add or remove the Azure Data Factory connections, you need to be assigned the following role. Role inheritance from management group is not supported. **Collection admins** role on the root collection.
>
> Also, you need to be the Azure Data Factory's "Owner" or "Contributor".
>
> Your Azure Data Factory needs to have system assigned managed identity enabled.

Follow the steps below to connect an existing Azure Data Factory to your Microsoft Purview account. You can also connect Data Factory to Microsoft Purview account from ADF.

1. Select **Management** on the left navigation pane.
2. Under **Lineage connections**, select **Data Factory**.
3. On the **Data Factory connection** page, select **New**.
4. Select your Azure Data Factory account from the list and select **OK**. You can also filter by subscription name to limit your list.

   Some Azure Data Factory instances might be disabled if the Azure Data Factory is already connected to the current Purview account, or the Azure Data Factory doesn't have a managed identity. A warning message is displayed if any of the selected Azure Data Factorys are already connected to other Microsoft Purview account. When you select **OK**, the Azure Data Factory connection with the other Microsoft Purview account is disconnected. No other confirmations are required.



> **Note**
>
> You can add up to 10 Azure Data Factory accounts at once. To add more than 10 Azure Data Factory accounts, add them in multiple batches.

How authentication works

Azure Data Factory uses its managed identity to authenticate lineage push operations from Microsoft Purview. When you connect your Azure Data Factory to Purview in the UI, it adds the role assignment automatically.

Grant the Azure Data Factory's managed identity the **Data Curator** role on Microsoft Purview **root collection**. To learn more, see Access control in Microsoft Purview and Add roles and restrict access through collections.

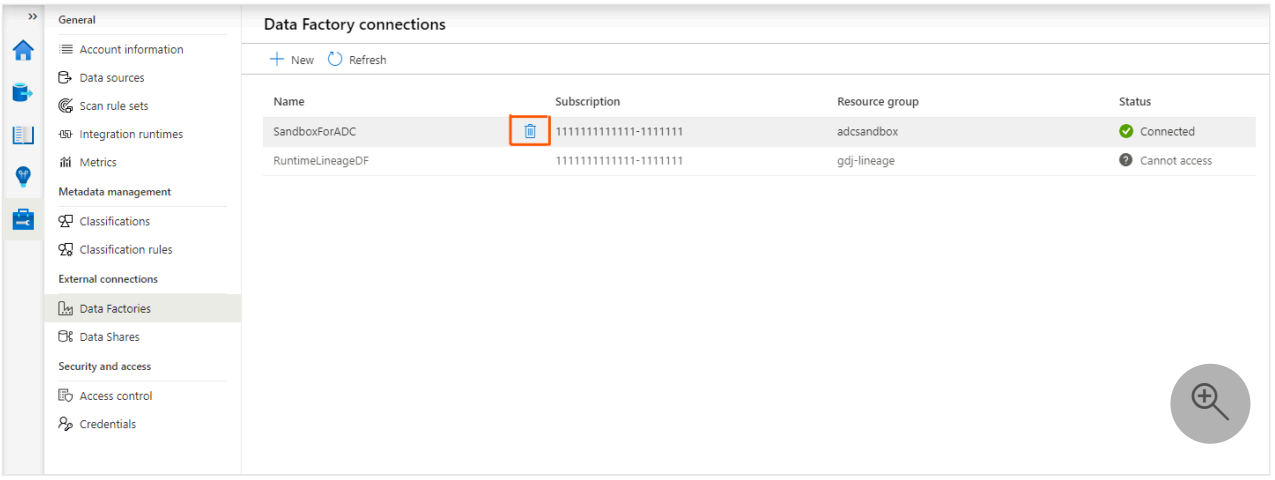
Remove Data Factory connections

> **Important**
>
> Even if the Data Factory connection is removed, any ongoing ingestion won't be canceled. Azure Data Factory lineage ingestion is an asynchronous operation.

To remove an Azure Data Factory connection, complete the following steps:

1. On the **Data Factory connection** page, select the **Remove** button next to one or more Azure Data Factory connections.

2. Select **Confirm** in the popup to delete the selected Azure Data Factory connections.



Monitor the Azure Data Factory links

In Microsoft Purview portal, you can monitor the Data Factory links.

Supported Azure Data Factory activities

Microsoft Purview captures runtime lineage from the following Azure Data Factory activities:

- Copy Data
- Data Flow
- Execute SSIS Package

> **Important**
>
> Microsoft Purview drops lineage if the source or destination uses an unsupported data storage system.

The integration between Azure Data Factory and Microsoft Purview supports only a subset of the data systems that Azure Data Factory supports. For more information, see the following sections.

Copy activity support

Expand table

Data store	Supported
Azure Blob Storage	Yes
Azure Cognitive Search	Yes
Azure Cosmos DB for NoSQL *	Yes
Azure Cosmos DB for MongoDB *	Yes
Azure Data Explorer *	Yes
Azure Data Lake Storage Gen1	Yes
Azure Data Lake Storage Gen2	Yes
Azure Database for MariaDB *	Yes
Azure Database for MySQL *	Yes
Azure Database for PostgreSQL *	Yes
Azure Files	Yes
Azure SQL Database *	Yes
Azure SQL Managed Instance *	Yes
Azure Synapse Analytics *	Yes
Azure Dedicated SQL pool (formerly SQL DW) *	Yes
Azure Table Storage	Yes
Amazon S3	Yes
Hive *	Yes
Oracle *	Yes
SAP Table *(when connecting to SAP ECC or SAP S/4HANA)*	Yes
SQL Server *	Yes
Teradata *	Yes

** Microsoft Purview currently doesn't support query or stored procedure for lineage or scanning. Lineage is limited to table and view sources only.*

If you use Self-hosted Integration Runtime, note the minimal version with lineage support for:

- Any use case: version 5.9.7885.3 or later

- Copying data from Oracle: version 5.10 or later
- Copying data into Azure Synapse Analytics via COPY command or PolyBase: version 5.10 or later

Limitations on copy activity lineage

Currently, if you use the following copy activity features, the lineage is not yet supported:

- Copy data into Azure Data Lake Storage Gen1 using Binary format.
- Compression setting for Binary, delimited text, Excel, JSON, and XML files.
- Source partition options for Azure SQL Database, Azure SQL Managed Instance, Azure Synapse Analytics, SQL Server, and SAP Table.
- Copy data to file-based sink with setting of max rows per file.
- Column level lineage is not currently supported by copy activity when source/sink is resource set.

In additional to lineage, the data asset schema (shown in Asset -> Schema tab) is reported for the following connectors:

- CSV and Parquet files on Azure Blob, Azure Files, ADLS Gen1, ADLS Gen2, and Amazon S3
- Azure Data Explorer, Azure SQL Database, Azure SQL Managed Instance, Azure Synapse Analytics, SQL Server, Teradata

Data Flow support

Expand table

Data store	Supported
Azure Blob Storage	Yes
Azure Cosmos DB for NoSQL *	Yes
Azure Data Lake Storage Gen1	Yes
Azure Data Lake Storage Gen2	Yes
Azure Database for MySQL *	Yes
Azure Database for PostgreSQL *	Yes
Azure SQL Database *	Yes
Azure SQL Managed Instance *	Yes
Azure Synapse Analytics *	Yes

Data store	Supported
Azure Dedicated SQL pool (formerly SQL DW) *	Yes

** Microsoft Purview currently doesn't support query or stored procedure for lineage or scanning. Lineage is limited to table and view sources only.*

Limitations on data flow lineage

- Data flow lineage may generate folder level resource set without visibility on the involved files.
- Column level lineage is not currently supported when source/sink is resource set.
- For the lineage of data flow activity, Microsoft Purview only supports showing the source and sink involved. The detailed lineage for data flow transformation isn't supported yet.
- Lineage is not supported when flowlets are part of the dataflow.
- Currently Purview doesn't support lineage reporting for Synapse tables (LakeHouse DB/Workspace DB)

Execute SSIS Package support

Refer to supported data stores.

Access secured Microsoft Purview account

If a firewall protects your Microsoft Purview account, learn how to let Azure Data Factory access a secured Microsoft Purview account through Microsoft Purview private endpoints.

Bring Azure Data Factory lineage into Microsoft Purview

For an end-to-end walkthrough, see the Tutorial: Push Data Factory lineage data to Microsoft Purview.

Supported lineage patterns

Microsoft Purview supports several patterns of lineage. The generated lineage data depends on the type of source and sink used in the Azure Data Factory activities. Although Azure Data Factory supports over 80 sources and sinks, Microsoft Purview supports only a subset. For the supported sources and sinks, see Supported Azure Data Factory activities.

To configure Azure Data Factory to send lineage information, see Get started with lineage.

You can find information in the lineage view by using the following methods:

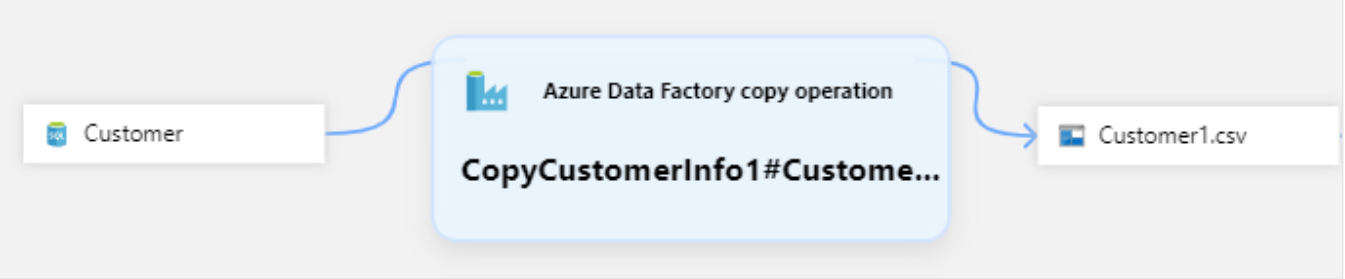
- In the **Lineage** tab, hover on shapes to preview additional information about the asset in the tooltip.
- Select the node or edge to see the asset type it belongs to or to switch assets.
- The **Lineage** tab displays columns of a dataset on the left side. For more information about column-level lineage, see Dataset column lineage.

Data lineage for 1:1 operations

The most common pattern for capturing data lineage moves data from a single input dataset to a single output dataset, with a process in between.

An example of this pattern includes the following components:

- 1 source/input: *Customer* (SQL Table)
- 1 sink/output: *Customer1.csv* (Azure Blob)
- 1 process: *CopyCustomerInfo1#Customer1.csv* (Data Factory Copy activity)

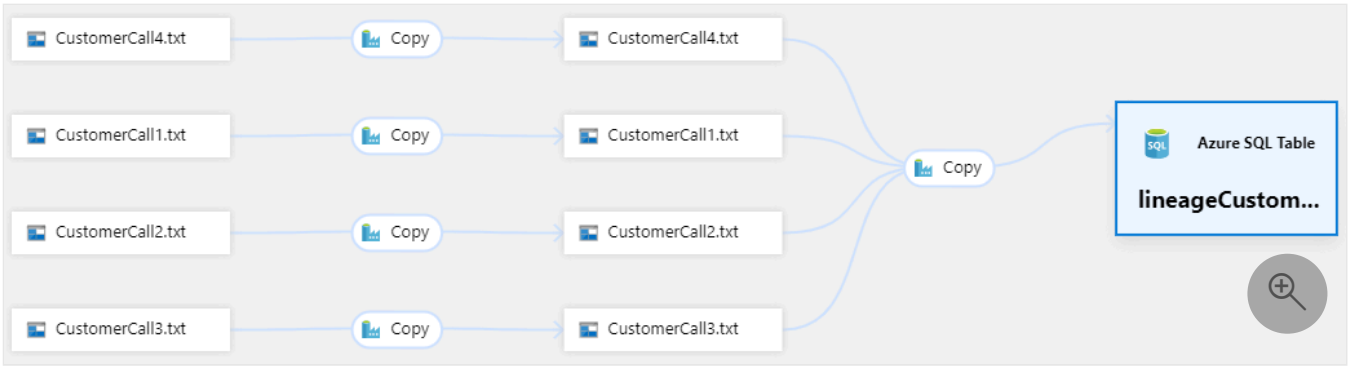


Data movement with 1:1 lineage and wildcard support

Another common scenario for capturing lineage is using a wildcard to copy files from a single input dataset to a single output dataset. The wildcard allows the copy activity to match multiple files for copying by using a common portion of the file name. Microsoft Purview captures file-level lineage for each individual file copied by the corresponding copy activity.

An example of this pattern includes the following components:

- Source/input: *CustomerCall*.csv* (ADLS Gen2 path)
- Sink/output: *CustomerCall*.csv* (Azure blob file)
- 1 process: *CopyGen2ToBlob#CustomerCall.csv* (Data Factory Copy activity)

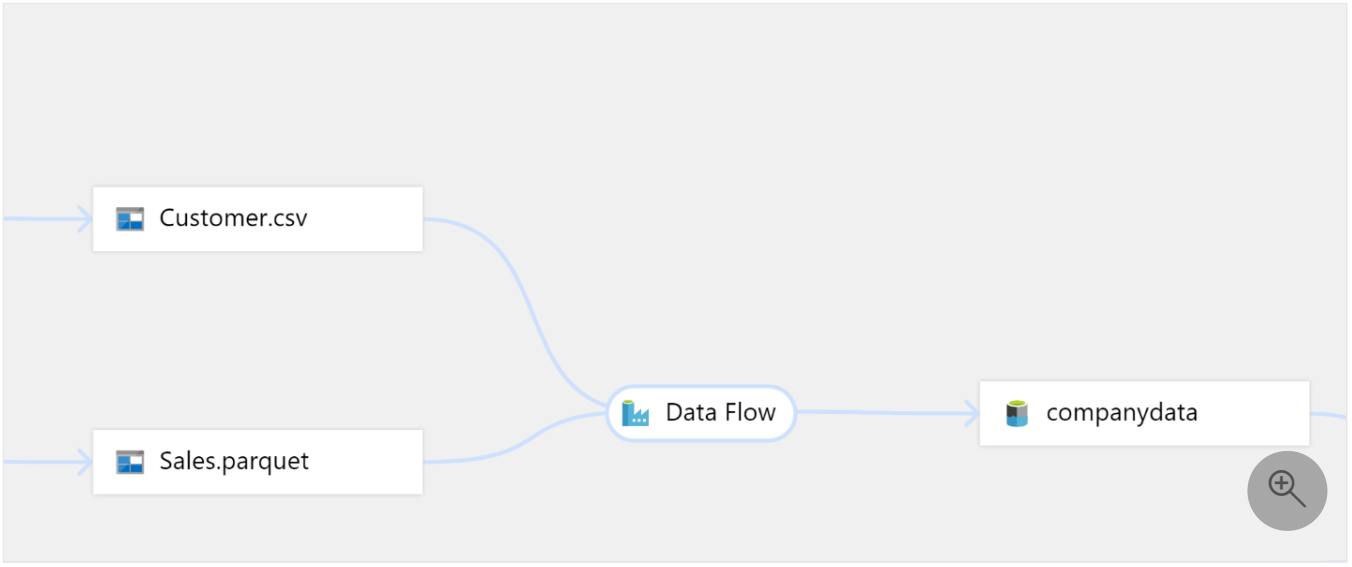


Data movement with n:1 lineage

You can use Data Flow activities to perform data operations like merge, join, and so on. You can use more than one source dataset to produce a target dataset. In this example, Microsoft Purview captures file-level lineage for individual input files to a SQL table that is part of a Data Flow activity.

An example of this pattern includes the following components:

- 2 sources/inputs: *Customer.csv*, *Sales.parquet* (ADLS Gen2 Path)
- 1 sink/output: *Company data* (Azure SQL table)
- 1 process: *DataFlowBlobsToSQL* (Data Factory Data Flow activity)

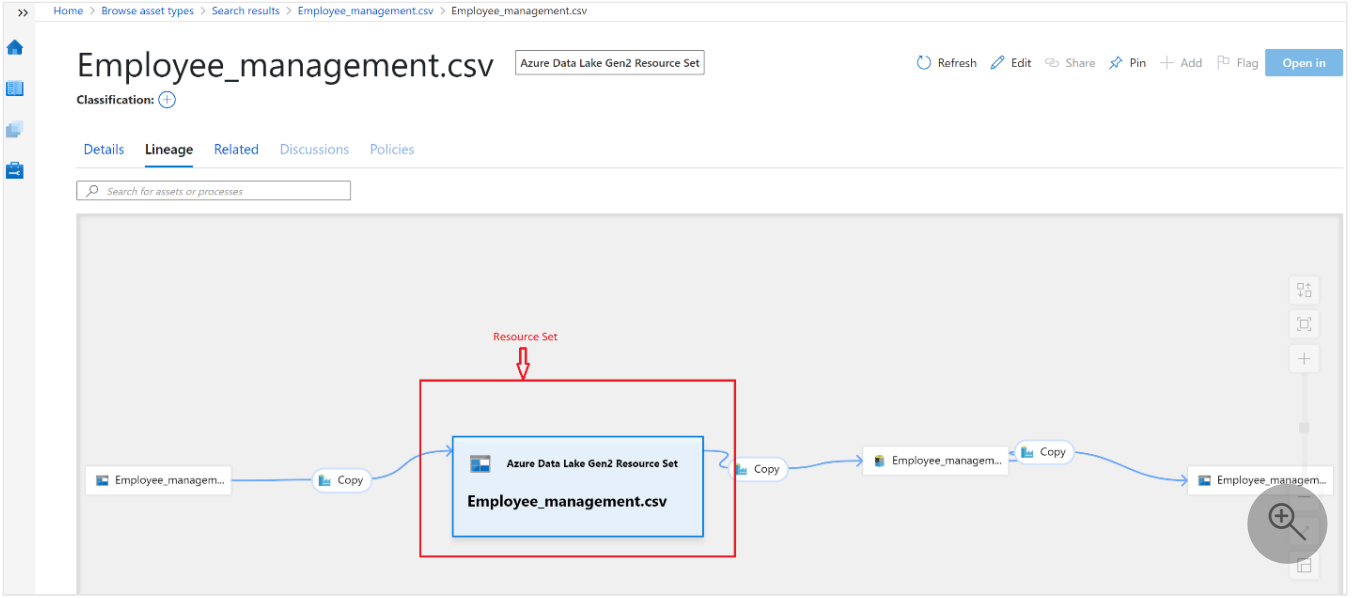


Lineage for resource sets

A resource set is a logical object in the catalog that represents many partition files in the underlying storage. For more information, see Understanding Resource sets. When Microsoft Purview captures lineage from Azure Data Factory, it applies the rules to normalize the individual partition files and create a single logical object.

In the following example, an Azure Data Lake Gen2 resource set is produced from an Azure Blob:

- 1 source/input: *Employee_management.csv* (Azure Blob)
- 1 sink/output: *Employee_management.csv* (Azure Data Lake Gen 2)
- 1 process: *CopyBlobToAdlsGen2_RS* (Azure Data Factory Copy activity)



Next steps

Tutorial: Push Data Factory lineage data to Microsoft Purview

Catalog lineage user guide

Link to Azure Data Share for lineage

Last updated on 02/24/2026

How to connect Azure Data Share and Microsoft Purview

Article • 02/26/2025

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

This article discusses how to connect your Azure Data Share account with Microsoft Purview and govern the shared datasets (both outgoing and incoming) in your data estate. Various data governance personas can discover and track lineage of data across boundaries like organizations, departments and even data centers.

> **Important**
>
> - This article is about lineage for the Azure Data Share service. If you are using data sharing in Microsoft Purview, use this article for lineage: Microsoft Purview Data Sharing lineage.
> - Data sharing is accessed through the classic Microsoft Purview governance portal: https://web.purview.azure.com

Common scenarios

Data Share Lineage is aimed to provide detailed information for root cause analysis and impact analysis.

Scenario 1: 360 view of datasets shared in/out for a partner organization or internal department

Data officers can see a list of all datasets that are bi-directionally shared with their partner organizations. They can search and discover the datasets by organization name and see a complete view of all outgoing and incoming shares.

Scenario 2: Root cause analysis - upstream dependency on datasets coming into organization (consumer view of incoming shares)

A report has incorrect information because of upstream data issues from an external Data Share account. The data engineers can understand upstream failures, be informed about the reasons, and further contact the owner of the share to fix the issues causing their data discrepancy.

Scenario 3: Impact analysis on datasets going outside organization (provider view of outgoing shares)

Data producers want to know who will be impacted upon making a change to their dataset. Using lineage, a data producer can easily understand the impact of the downstream internal or external partners who are consuming data using Azure Data Share.

Azure Data Share and Microsoft Purview connected experience

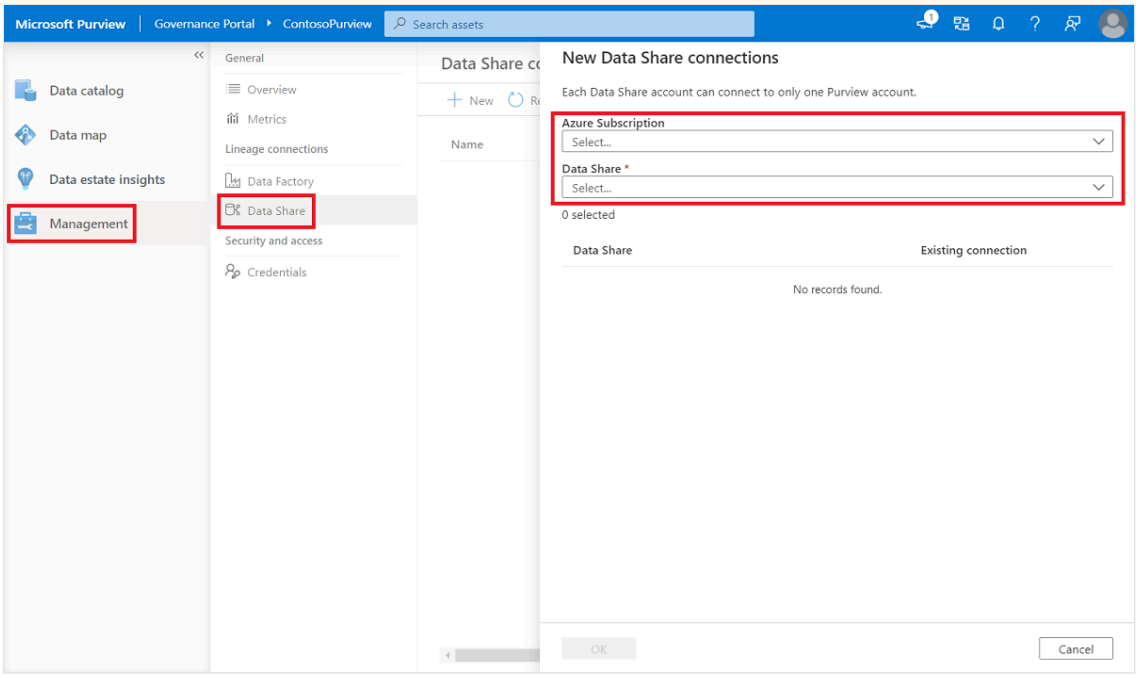
To connect your Azure Data Share and Microsoft Purview account, do the following:

1. Create a Microsoft Purview account. All the Data Share lineage information will be collected by a Microsoft Purview account. You can use an existing one or create a new Microsoft Purview account.

2. Connect your Azure Data Share to your Microsoft Purview account.

    a. In the Microsoft Purview governance portal, you can go to **Management Center** and connect your Azure Data Share under the **External connections** section.

    b. Select **+ New** on the top bar, find your Azure Data Share in the pop-up side bar and add the Data Share account. Run a snapshot job after connecting your Data Share to Microsoft Purview account, so that the Data Share assets and lineage information is visible in Microsoft Purview.

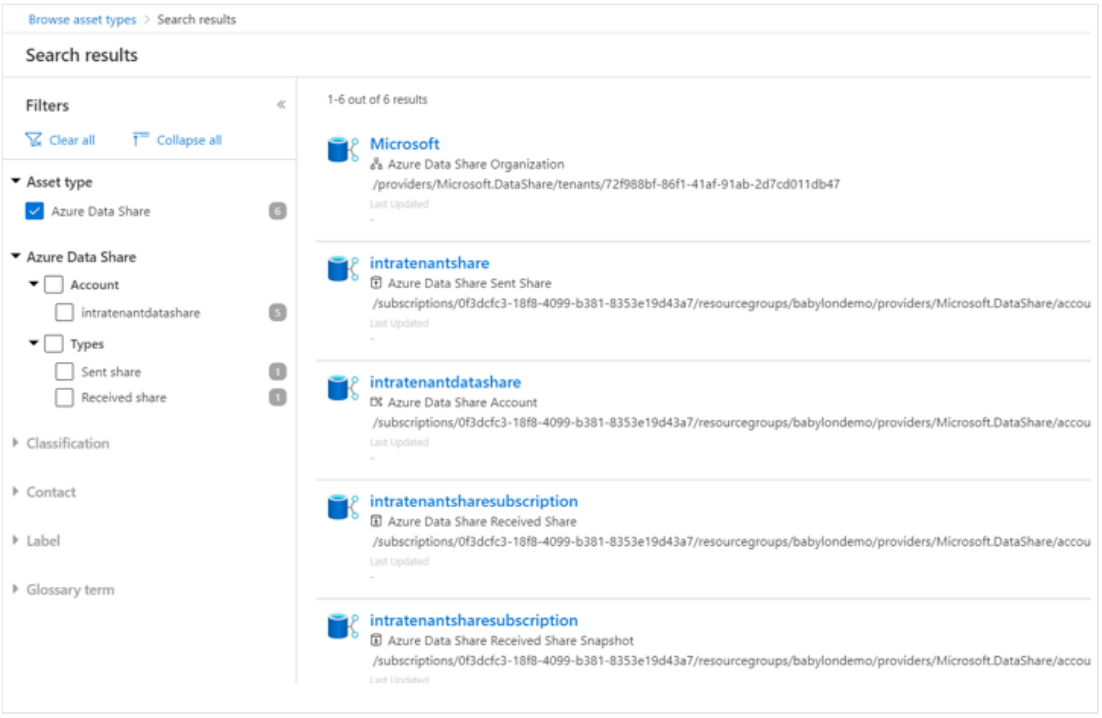


3. Execute your snapshot in Azure Data Share.

    - Once the Azure Data share connection is established with Microsoft Purview, you can execute a snapshot for your existing shares.
    - If you don't have any existing shares, go to the Azure Data Share portal to share your data and subscribe to a data share.
    - Once the share snapshot is complete, you can view associated Data Share assets and lineage in Microsoft Purview.

4. Discover Data Share accounts and share information in your Microsoft Purview account.

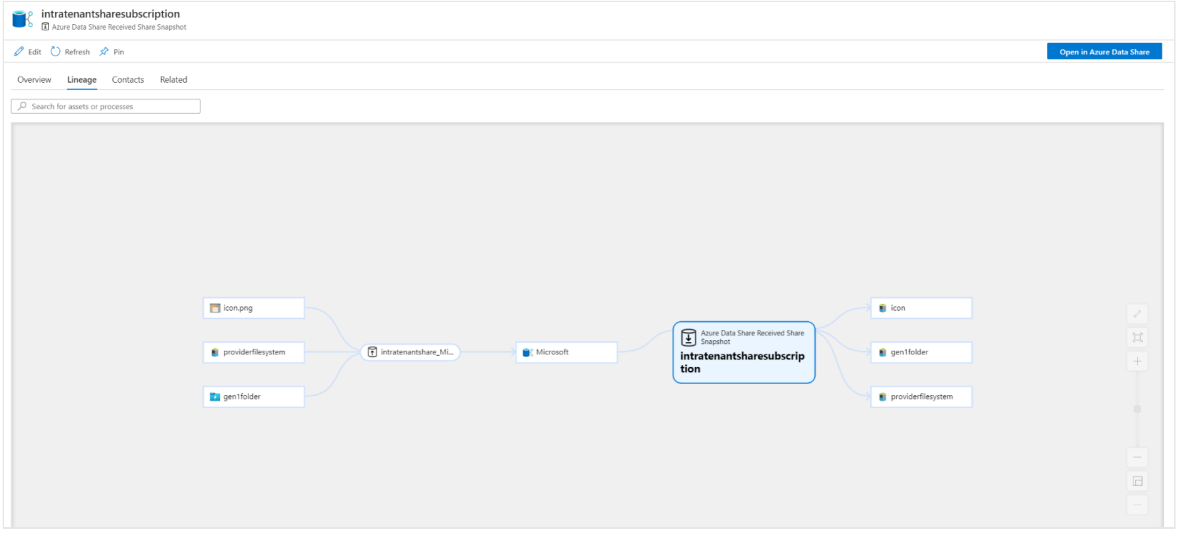
    - In the home page of Microsoft Purview account, select **Browse by asset type** and select the **Azure Data Share** tile. You can search for an account name, share name, share snapshot, or partner organization. Otherwise apply filters on the Search result page for account name, share type (sent vs received shares).



> **ⓘ Important**
>
> For Data Share assets to show in Microsoft Purview, a snapshot job must be run after you connect your Data Share to Microsoft Purview.

5. Track lineage of datasets shared with Azure Data Share.

    - From the Microsoft Purview search result page, choose the Data share snapshot (received/sent) and select the **Lineage** tab, to see a lineage graph with upstream and downstream dependencies.



Next steps

- Catalog lineage user guide

- Link to Azure Data Factory for lineage

Feedback

Was this page helpful? Yes No

How to get lineage from Azure Synapse Analytics into Microsoft Purview

07/18/2025

This article explains the steps required for connecting an Azure Synapse workspace with a Microsoft Purview account to track data lineage and ingest data sources. The article also explains details of the activity coverage scope and supported lineage capabilities.

When you connect Azure Synapse Analytics to Microsoft Purview, whenever a supported pipeline activity is run, metadata about the activity's source data, output data, and the activity will be automatically ingested into the Microsoft Purview Data Map.

If a data source has already been scanned and exists in the data map, the ingestion process will add the lineage information from Azure Synapse Analytics to that existing source. If the source or output doesn't exist in the data map and is supported by Azure Synapse Analytics lineage Microsoft Purview will automatically add their metadata from Synapse Analytics into the data map under the default domain root collection.

This can be an excellent way to monitor your data estate as users move and transform information using Azure Synapse Analytics.

Supported Azure Synapse capabilities

Currently, Microsoft Purview captures runtime lineage from the following Azure Synapse pipeline activities:

- Copy Data
- Data Flow

> **Important**
>
> Microsoft Purview drops lineage if the source or destination uses an unsupported data storage system.

Copy activity support

Expand table

Data store	Supported
Azure Blob Storage	Yes
Azure Cognitive Search	Yes
Azure Cosmos DB for NoSQL *	Yes
Azure Cosmos DB for MongoDB *	Yes
Azure Data Explorer *	Yes
Azure Data Lake Storage Gen1	Yes
Azure Data Lake Storage Gen2	Yes
Azure Database for MariaDB *	Yes
Azure Database for MySQL *	Yes
Azure Database for PostgreSQL *	Yes
Azure Files	Yes
Azure SQL Database *	Yes
Azure SQL Managed Instance *	Yes
Azure Synapse Analytics *	Yes
Azure Dedicated SQL pool (formerly SQL DW) *	Yes
Azure Table Storage	Yes
Amazon S3	Yes
Hive *	Yes
Oracle *	Yes
SAP Table *(when connecting to SAP ECC or SAP S/4HANA)*	Yes
SQL Server *	Yes
Teradata *	Yes

** Microsoft Purview currently doesn't support query or stored procedure for lineage or scanning. Lineage is limited to table and view sources only.*

If you use Self-hosted Integration Runtime, note the minimal version with lineage support for:

- Any use case: version 5.9.7885.3 or later

- Copying data from Oracle: version 5.10 or later
- Copying data into Azure Synapse Analytics via COPY command or PolyBase: version 5.10 or later

Limitations on copy activity lineage

Currently, if you use the following copy activity features, the lineage is not yet supported:

- Copy data into Azure Data Lake Storage Gen1 using Binary format.
- Compression setting for Binary, delimited text, Excel, JSON, and XML files.
- Source partition options for Azure SQL Database, Azure SQL Managed Instance, Azure Synapse Analytics, SQL Server, and SAP Table.
- Copy data to file-based sink with setting of max rows per file.
- Column level lineage is not currently supported by copy activity when source/sink is resource set.

In additional to lineage, the data asset schema (shown in Asset -> Schema tab) is reported for the following connectors:

- CSV and Parquet files on Azure Blob, Azure Files, ADLS Gen1, ADLS Gen2, and Amazon S3
- Azure Data Explorer, Azure SQL Database, Azure SQL Managed Instance, Azure Synapse Analytics, SQL Server, Teradata

Data Flow support

Expand table

Data store	Supported
Azure Blob Storage	Yes
Azure Cosmos DB for NoSQL *	Yes
Azure Data Lake Storage Gen1	Yes
Azure Data Lake Storage Gen2	Yes
Azure Database for MySQL *	Yes
Azure Database for PostgreSQL *	Yes
Azure SQL Database *	Yes
Azure SQL Managed Instance *	Yes
Azure Synapse Analytics *	Yes

Data store	Supported
Azure Dedicated SQL pool (formerly SQL DW) *	Yes

** Microsoft Purview currently doesn't support query or stored procedure for lineage or scanning. Lineage is limited to table and view sources only.*

Limitations on data flow lineage

- Data flow lineage may generate folder level resource set without visibility on the involved files.
- Column level lineage is not currently supported when source/sink is resource set.
- For the lineage of data flow activity, Microsoft Purview only supports showing the source and sink involved. The detailed lineage for data flow transformation isn't supported yet.
- Lineage is not supported when flowlets are part of the dataflow.
- Currently Purview doesn't support lineage reporting for Synapse tables (LakeHouse DB/Workspace DB)

Access secured Microsoft Purview account

If your Microsoft Purview account is protected by firewall, learn how to let Azure Synapse access a secured Microsoft Purview account through Microsoft Purview private endpoints.

Bring Azure Synapse lineage into Microsoft Purview

Step 1: Connect Azure Synapse workspace to your Microsoft Purview account

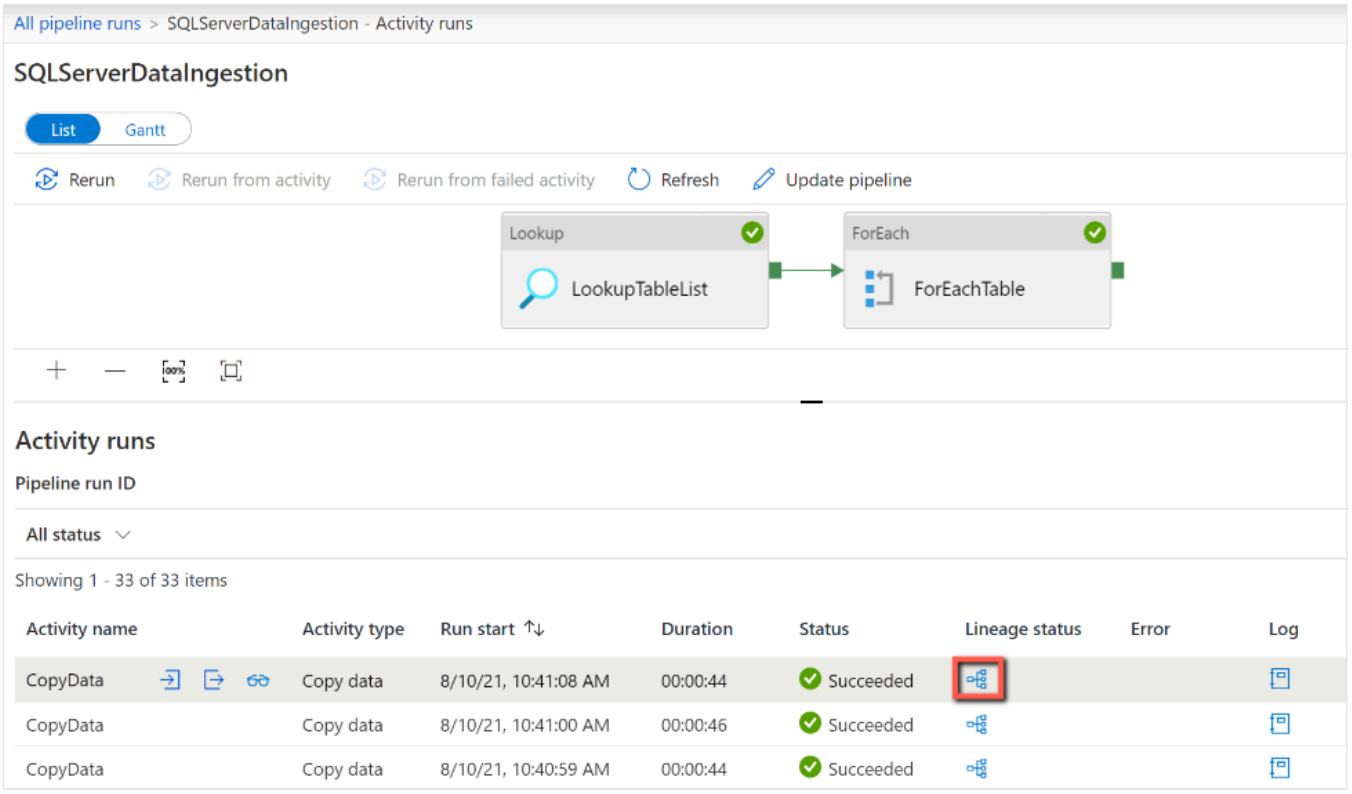
You can connect an Azure Synapse workspace to Microsoft Purview, and the connection enables Azure Synapse to push lineage information to Microsoft Purview. Follow the steps in Connect Synapse workspace to Microsoft Purview. Multiple Azure Synapse workspaces can connect to a single Microsoft Purview account for holistic lineage tracking.

Step 2: Run pipeline in Azure Synapse workspace

You can create pipelines with Copy activity in Azure Synapse workspace. You don't need any other configuration for lineage data capture. The lineage data will automatically be captured during the activities execution.

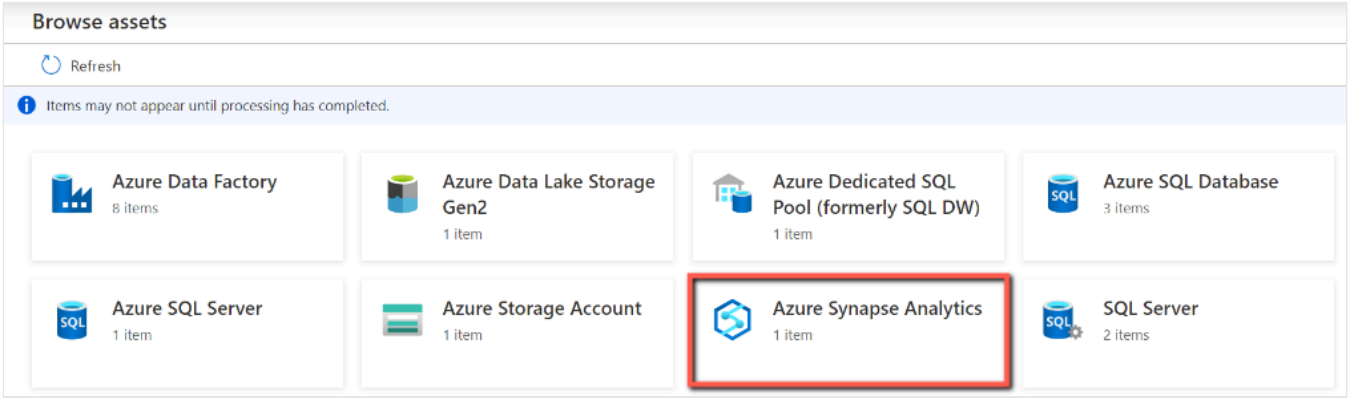
Step 3: Monitor lineage reporting status

After you run the Azure Synapse pipeline, in the Synapse pipeline monitoring view, you can check the lineage reporting status by selecting the following **Lineage status** button. The same information is also available in the activity output JSON -> `reportLineageToPurvew` section.

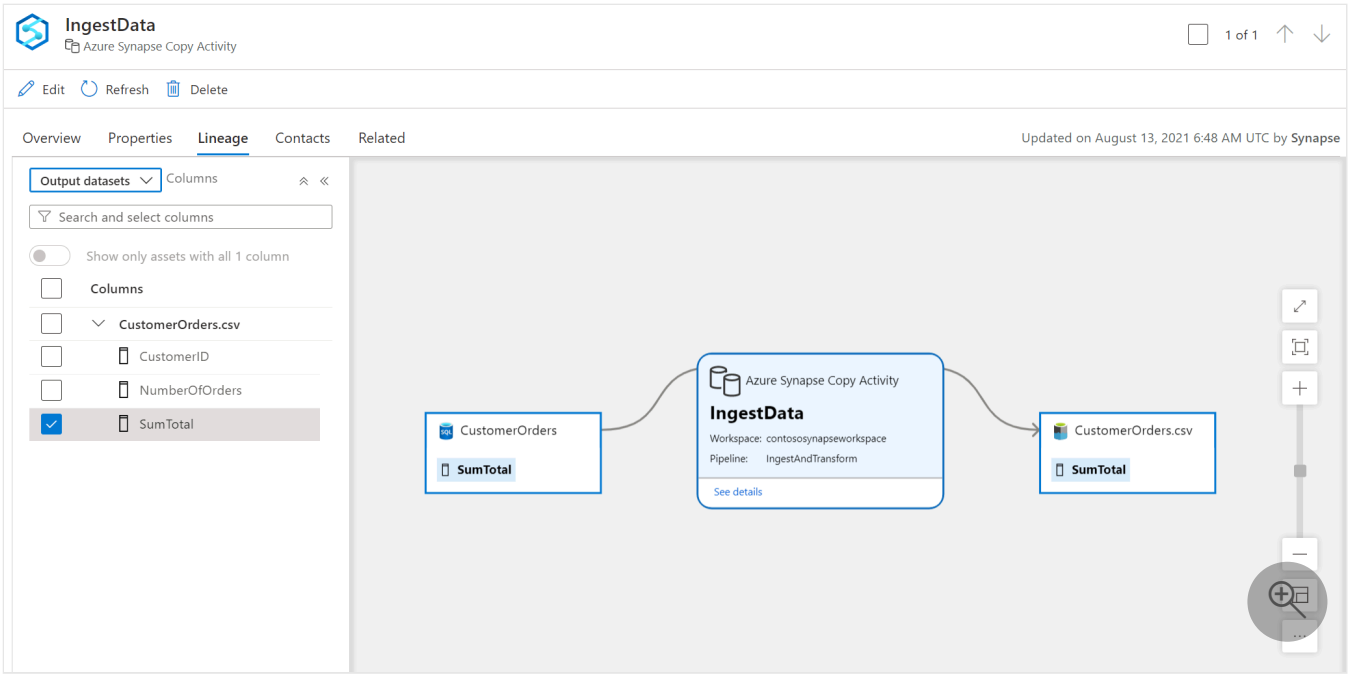


Step 4: View lineage information in your Microsoft Purview account

In your Microsoft Purview account, you can browse assets and choose type "Azure Synapse Analytics". You can also search Unified Catalog using keywords.



Select the Synapse account -> pipeline -> activity, you can view the lineage information.



Monitor the Azure Synapse Analytics links

In Microsoft Purview governance portal, you can monitor the Azure Synapse Analytics links.

Next steps

Catalog lineage user guide

Link to Azure Data Share for lineage

How to get lineage from Microsoft Fabric items into Microsoft Purview

07/18/2025

This article explains data lineage for Microsoft Fabric items in Microsoft Purview.

Prerequisites

To see data lineage in Microsoft Purview for Fabric items, you must first register and scan your Fabric tenant.

Supported capabilities

Scanning a Microsoft Fabric tenant will automatically bring in metadata and lineage from Fabric items including Power BI into Microsoft Purview Unified Catalog.

- For all Fabric items besides Power BI, only item level metadata and lineage can be scanned.
- For Lakehouse tables and files, sub-item level metadata scanning is now available in preview but sub-item level lineage is still not supported.

Lineage of Fabric items in Microsoft Purview

Once the scan of your Fabric tenant is complete, following Fabric items are inventoried in Microsoft Purview:

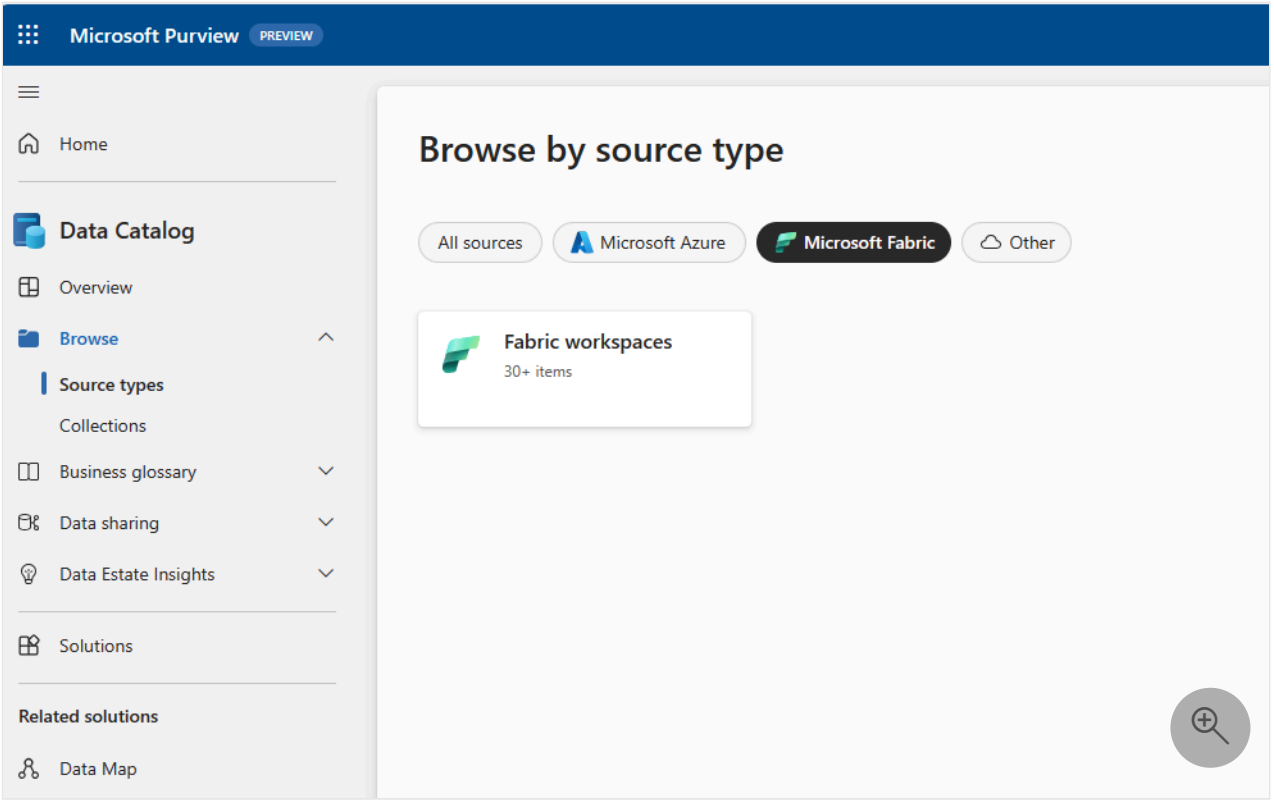
Expand table

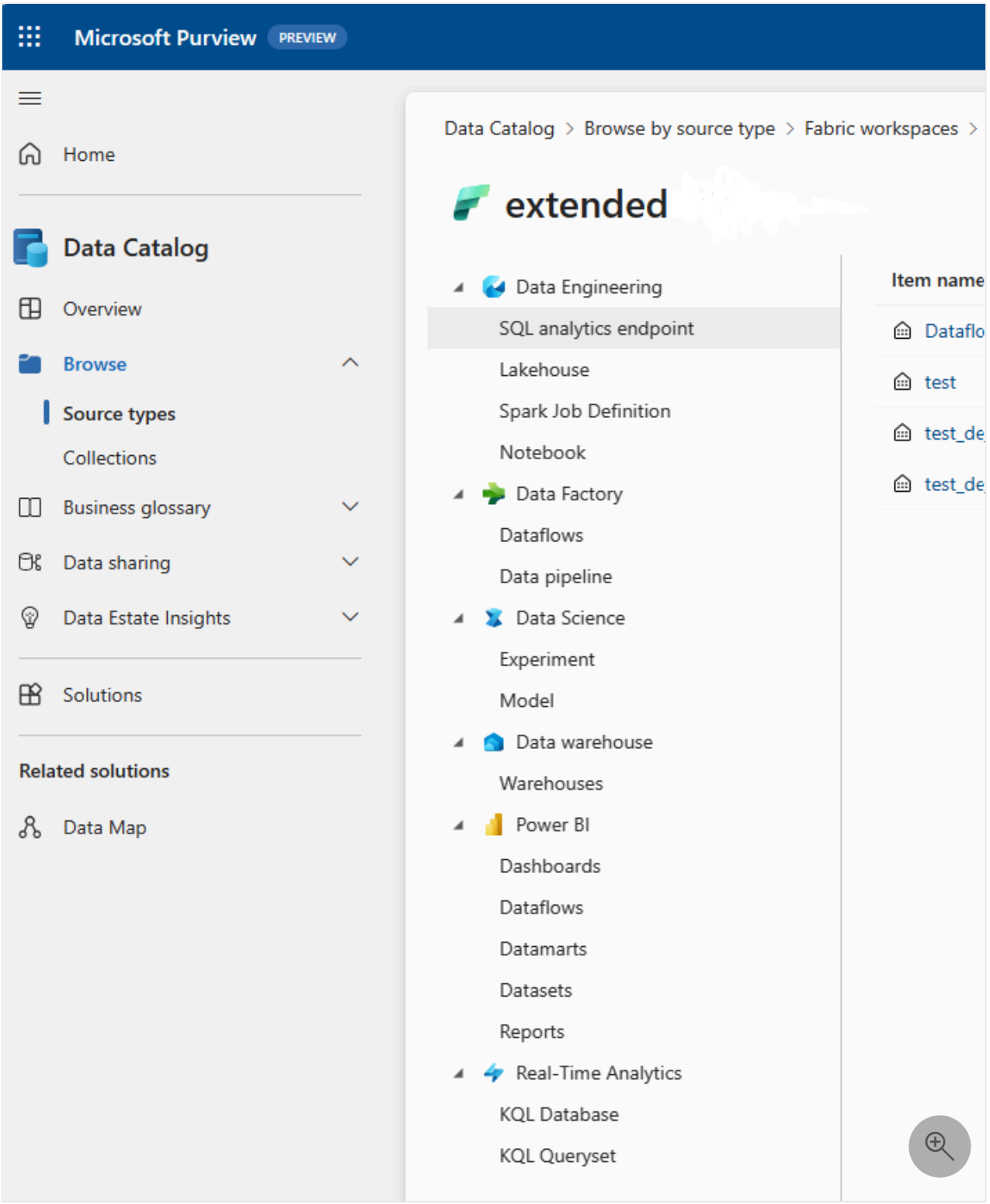
Experiences	Fabric items	Available in scan	Available in live view (same tenant only)
Real-Time Analytics	KQL Database	Yes	Yes
	KQL Queryset	Yes	Yes
Data Science	Experiment	Yes	Yes
	ML model	Yes	Yes
Data Factory	Data pipeline	Yes	Yes

Experiences	Fabric items	Available in scan	Available in live view (same tenant only)
	Dataflow Gen2	Yes	Yes
Data Engineering	Lakehouse	Yes	Yes
	Notebook	Yes	Yes
	Spark Job Definition	Yes	Yes
	SQL analytics endpoint	Yes	Yes
Data Warehouse	Warehouse	Yes	Yes
Power BI	Dashboard	Yes	Yes
	Dataflow	Yes	Yes
	Datamart	Yes	Yes
	Semantic model (Dataset)	Yes	Yes
	Report	Yes	Yes
	Paginated report	Yes	

To see lineage for Fabric items in Microsoft Purview:

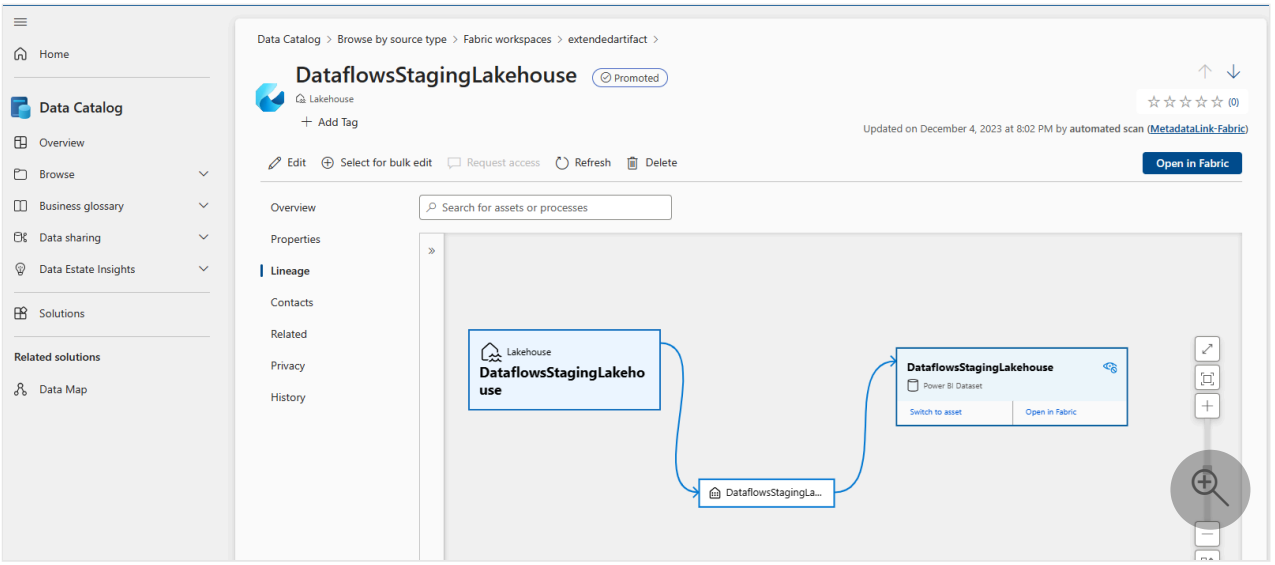
1. Select Unified Catalog -> Browse -> Source Types -> Microsoft Fabric -> Fabric Workspaces
2. Select the particular Fabric workspace, select the experience, and then the Fabric item of interest


Microsoft Purview
PREVIEW
Home
Data Catalog
Overview
Browse
Source types
Collections
Business glossary
Data sharing
Data Estate Insights
Solutions
Related solutions
Data Map
Browse by source type
All sources
Microsoft Azure
Microsoft Fabric
Other
Fabric workspaces
30+ items



Alternatively, users can search for a Fabric item by name, description, or other details to see relevant results.

3. Select the Lineage tab. Asset relationships are shown with the upstream and downstream dependencies.



Lineage of Power BI artifacts in Microsoft Purview

Refer to Power BI lineage for key details and limitations around Power BI lineage in Fabric.

Known limitations

- For non-Power BI Fabric items, external data sources as upstream sources in lineage aren't yet supported.
- Cross workspace lineage for non-Power BI items isn't supported.
- Notebook -> Pipeline lineage isn't supported.

Next steps

- Learn about data lineage in Microsoft Purview
- Learn about lineage in Microsoft Fabric

How to view Microsoft Purview Data Sharing lineage

> **Note**
>
> Microsoft Purview Data Catalog (classic), Data Health Insights (classic), and Purview Workflow (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

> **Important**
>
> This feature is currently being retired, with support provided until September 2025 to allow for migration to Microsoft Fabric external data sharing.

This article discusses how to view the lineage of datasets shared using Microsoft Purview Data Sharing. With these tools, users can discover and track lineage of data across boundaries like organizations, departments, and even data centers.

> **Important**
>
> This article is about lineage for Microsoft Purview Data Sharing. If you are using Azure Data Share, use this article for lineage: Azure Data Share lineage.

Common scenarios

Data Sharing lineage aims to provide detailed information for root cause analysis and impact analysis.

Some common scenarios include:

- Full view of datasets shared in and out of your organization
- Root cause analysis for upstream dataset dependencies
- Impact analysis for shared datasets

Full view of datasets shared in and out of your organization

Data officers can see a list of all bi-directionally shared datasets with their partner organizations. They can search and discover the datasets by organization name and see a complete view of all outgoing and incoming shares.

See the Microsoft Entra ID share lineage section to see a full view of outgoing and incoming shares.

Root cause analysis for upstream dataset dependencies

A report has incorrect information because of upstream data issues from an external Microsoft Purview Data Sharing activity. The data engineers can understand upstream failures, understand possible cause, and further contact the owner of the share to fix the issues causing their data discrepancy.

See the lineage of a share or Microsoft Entra ID share lineage sections to see the upstream dependencies for your assets.

Impact analysis for shared datasets

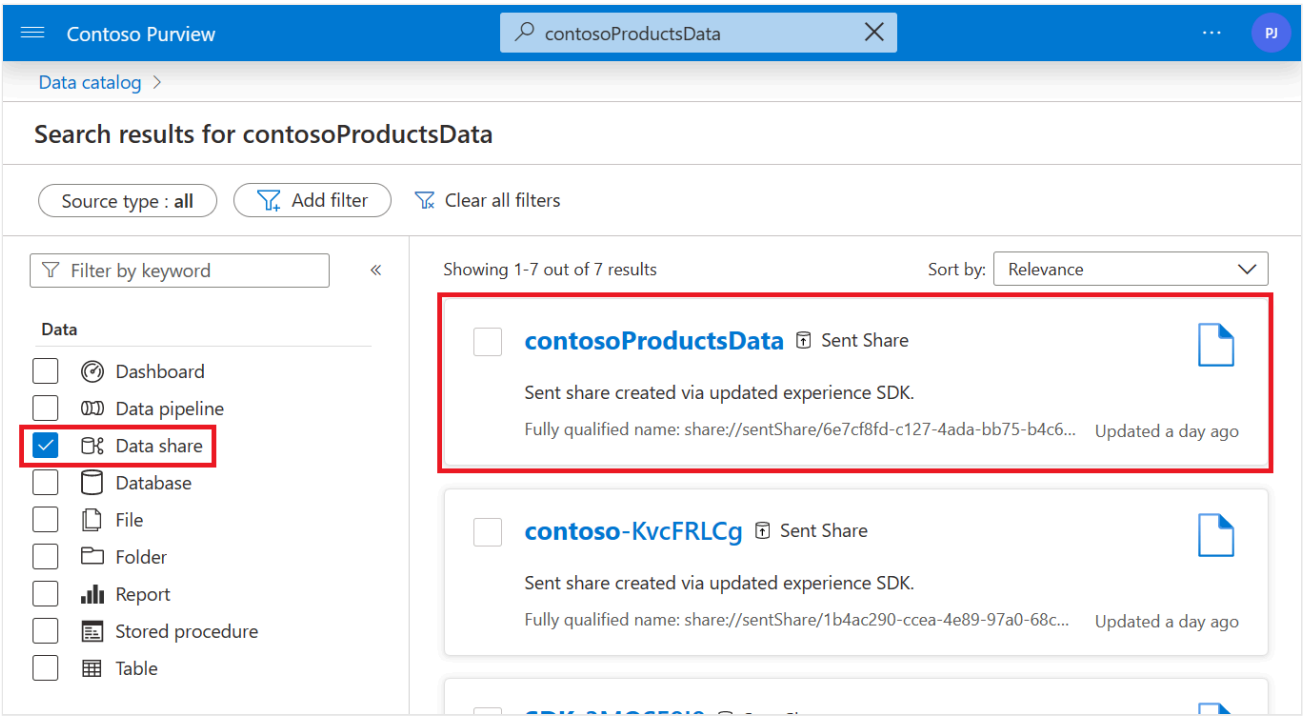
Data producers want to know who will be impacted when they make a change to their datasets. With a lineage map, a data producer can easily understand the impact of the downstream internal or external partners who are consuming data using Microsoft Purview Data Sharing.

See the lineage of a share or Microsoft Entra ID share lineage sections to see downstream dependencies of your shared assets.

Lineage of a share

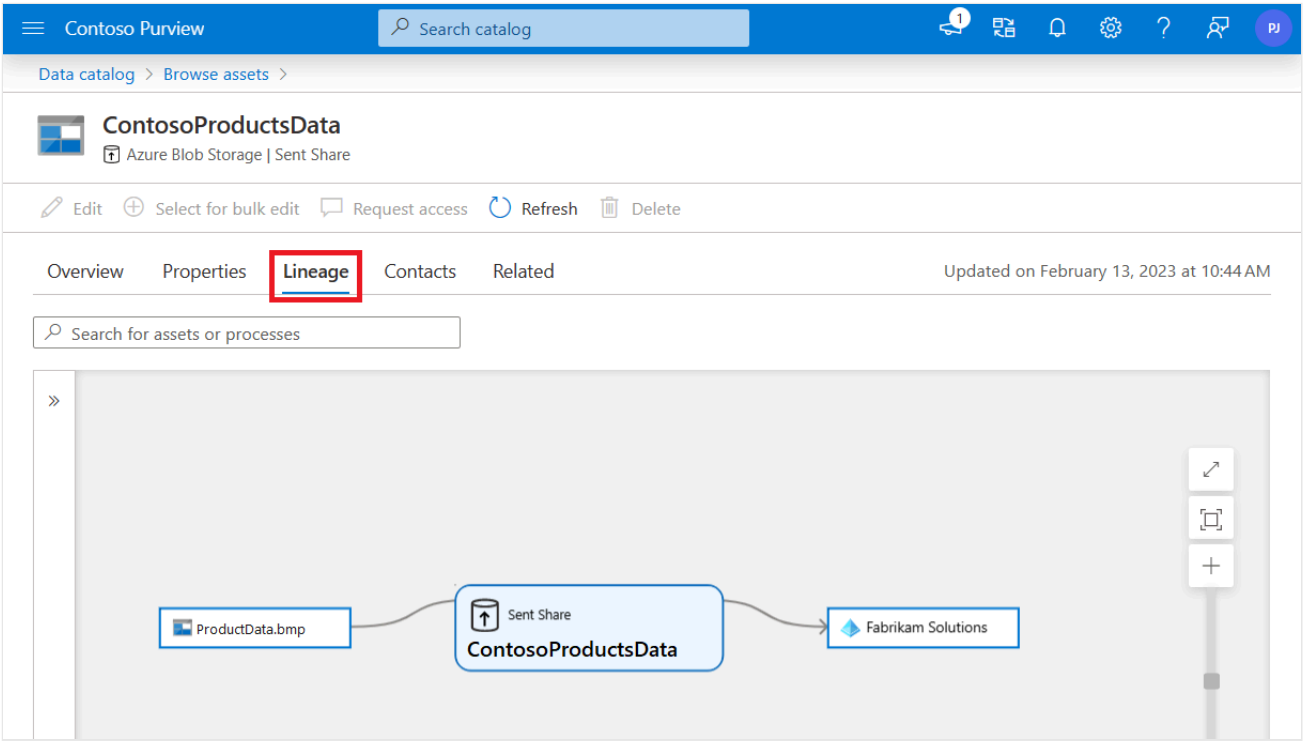
To see the data sharing lineage for your sent share or received share asset, you only need to have data reader permissions on the collection where the asset is housed.

1. Discover your sent share or received share asset in the Catalog search or browse for the share in the data catalog, and narrow results to only data shares.

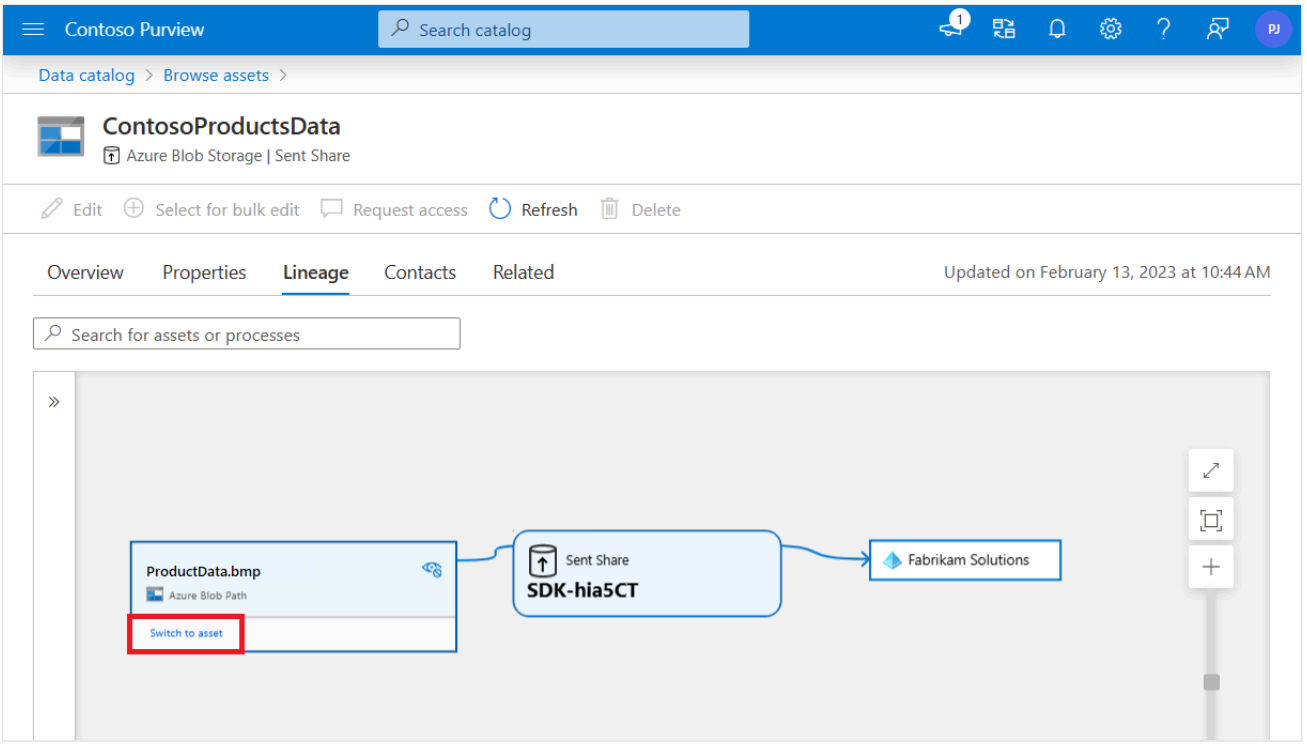


2. Select your data share asset.

3. Select the lineage tab to see a graph with upstream and downstream dependencies.



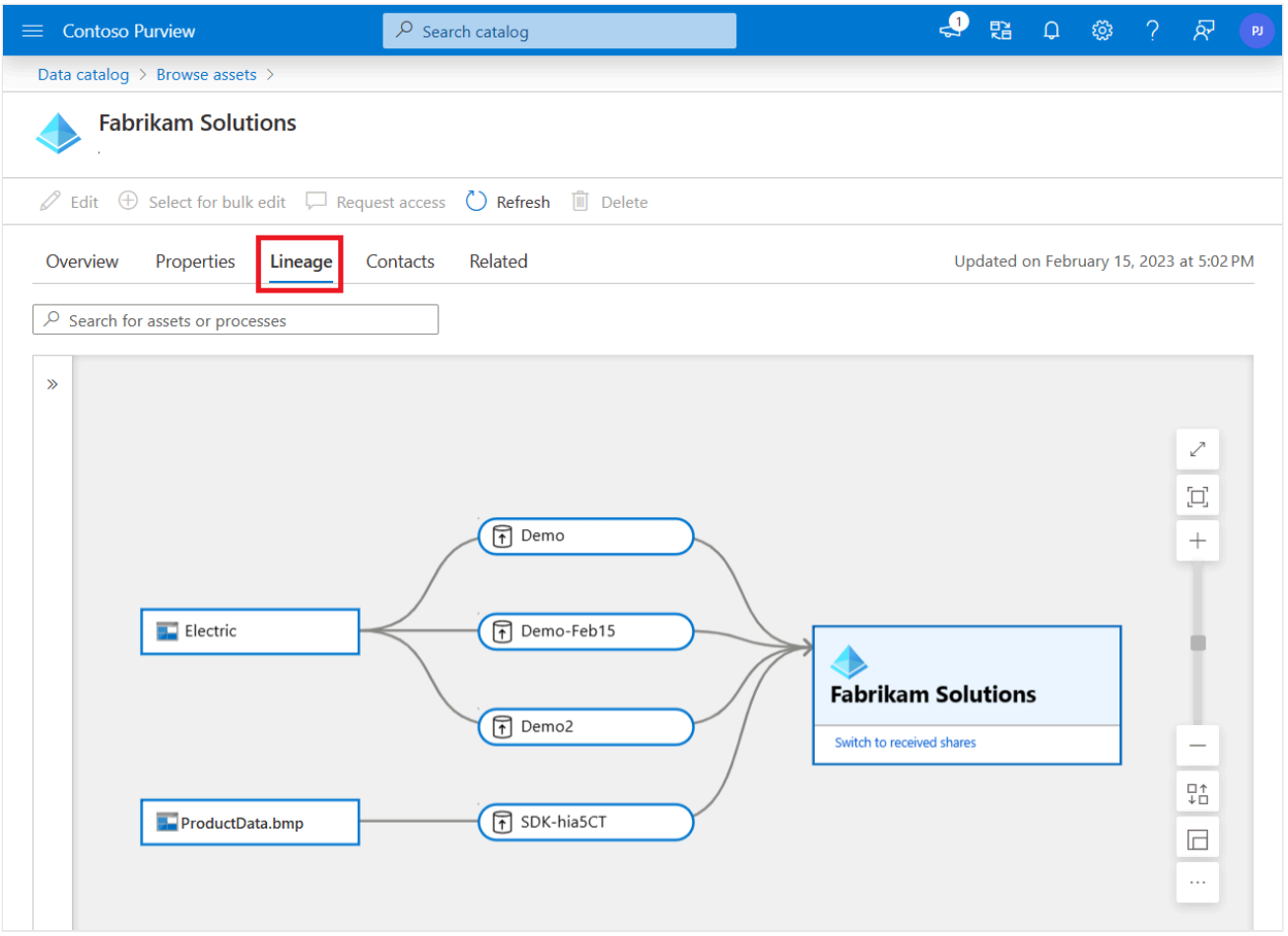
4. You can select the files or folders, or the Microsoft Entra asset in the lineage canvas and see the data sharing lineage for those assets.



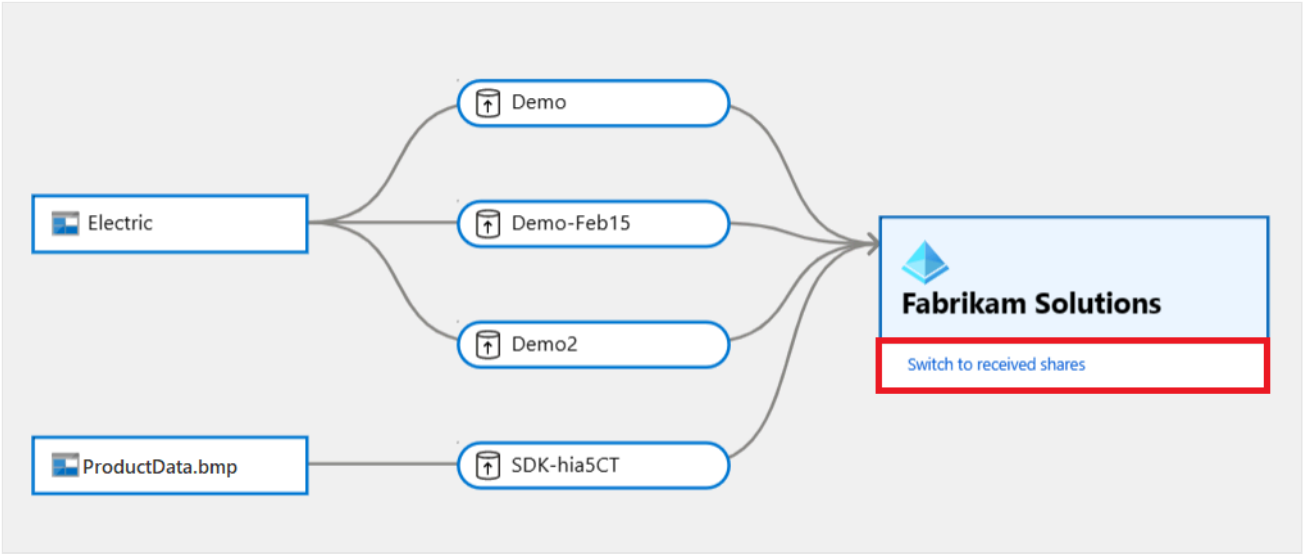
Microsoft Entra ID share lineage

1. You can search for your Microsoft Entra tenant asset in the Catalog using the search bar, or filtering by **Source type** and selecting **Microsoft Entra ID**.

2. Select your Microsoft Entra ID, then select the **Lineage** tab to see a lineage graph with either upstream or downstream dependencies.

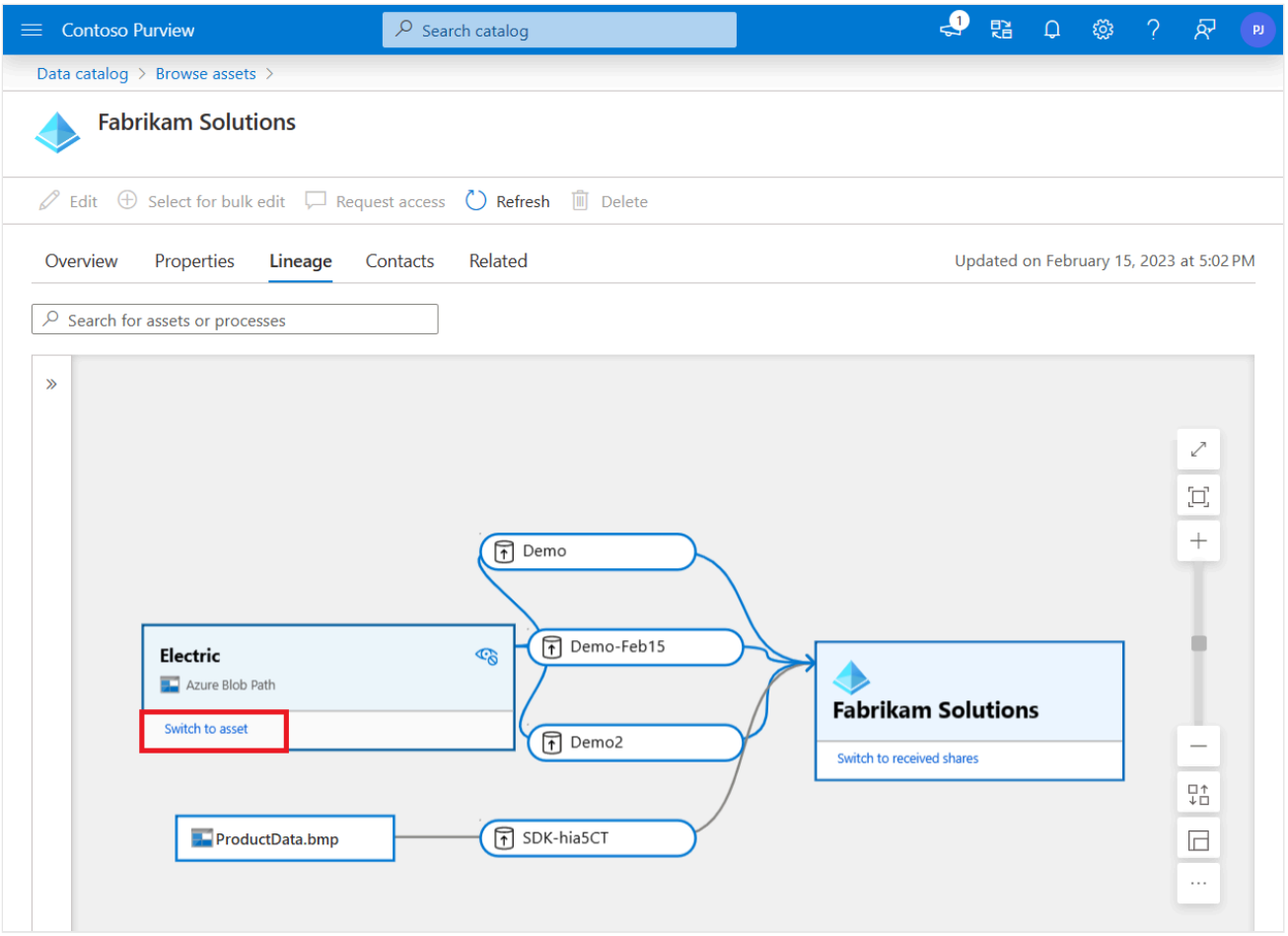


3. If you're seeing all sent shares, or all received shares, you can select the button to switch the view to see the other.



> **Note**
>
> If you do not see this button it may be because your Microsoft Entra ID does not have any sent/received shares, or that you do not have permission on the collection where the sent/received shares are housed.

4. You can also select any file or folder in the lineage canvas and see the data sharing lineage for those assets.



> **Important**
>
> For Data Share assets to show in Microsoft Purview, the ADLS Gen2 or Blob Storage account that the shares (sent share or received share) belong to should be registered with Microsoft Purview and the user needs Reader permission on the collection where they are housed.

Troubleshoot

Here are some common issues when viewing data sharing lineage and their possible resolutions.

Can't see the recipient Microsoft Entra tenant on the sent share lineage

If you're unable to see the recipient Microsoft Entra tenant on the sent share lineage, it means the recipient hasn't attached the share yet.

Can't find a sent share or received share asset in the Catalog

Sent or received share assets are housed in the same collection as their registered storage account assets.

- If the storage accounts the share assets belong to aren't registered, the share assets won't be discoverable.
- If you don't have a minimum of Data Reader permission to the collection where the share asset is housed, the share assets aren't discoverable.

Next steps

- Data sharing quickstart
- How to Share data
- How to receive a share
- REST API reference
- Data Sharing FAQ

Last updated on 10/10/2025

How to get lineage from Power BI into Microsoft Purview

07/18/2025

This article elaborates on the data lineage for Power BI sources in Microsoft Purview.

Prerequisites

To see data lineage in Microsoft Purview for Power BI, you must first register and scan your Power BI source.

Common scenarios

After a Power BI source has been scanned, lineage information for your current data assets, and data assets referenced by Power BI, will automatically be added in the Microsoft Purview Unified Catalog.

1. Data consumers can perform root cause analysis of a report or dashboard from Microsoft Purview. For any data discrepancy in a report, users can easily identify the upstream datasets and contact their owners if necessary.

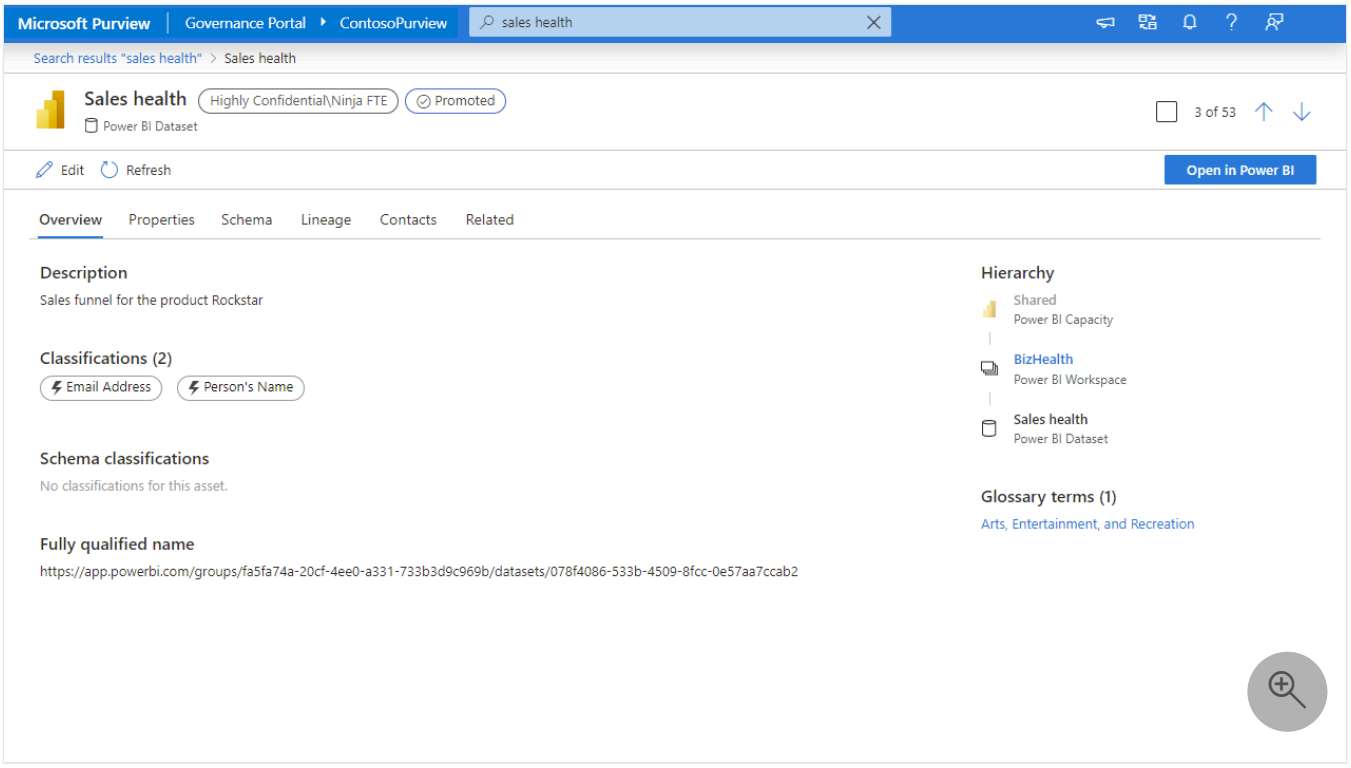
2. Data producers can see the downstream reports or dashboards consuming their dataset. Before making any changes to their datasets, the data owners can make informed decisions.

3. Users can search by name, endorsement status, sensitivity label, owner, description, and other business facets to return the relevant Power BI artifacts.

Power BI artifacts in Microsoft Purview

Once the scan of your Power BI is complete, following Power BI artifacts will be inventoried in Microsoft Purview:

- Workspaces
- Dashboards
- Reports
- Datasets
- Dataflows
- Datamarts

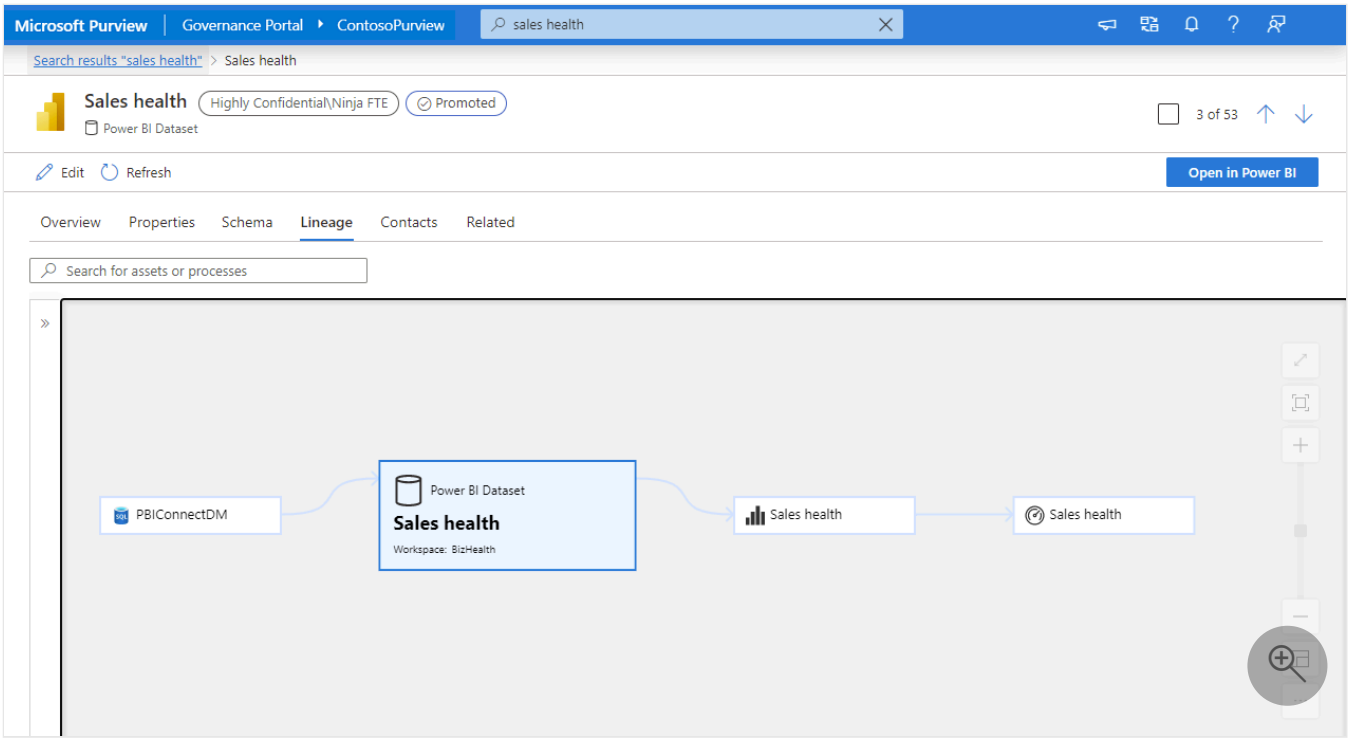


Lineage of Power BI artifacts in Microsoft Purview

Users can search for a Power BI artifact by name, description, or other details to see relevant results. Under the asset overview and properties tabs, the basic details such as description, classification are shown. Under the lineage tab, asset relationships are shown with the upstream and downstream dependencies.

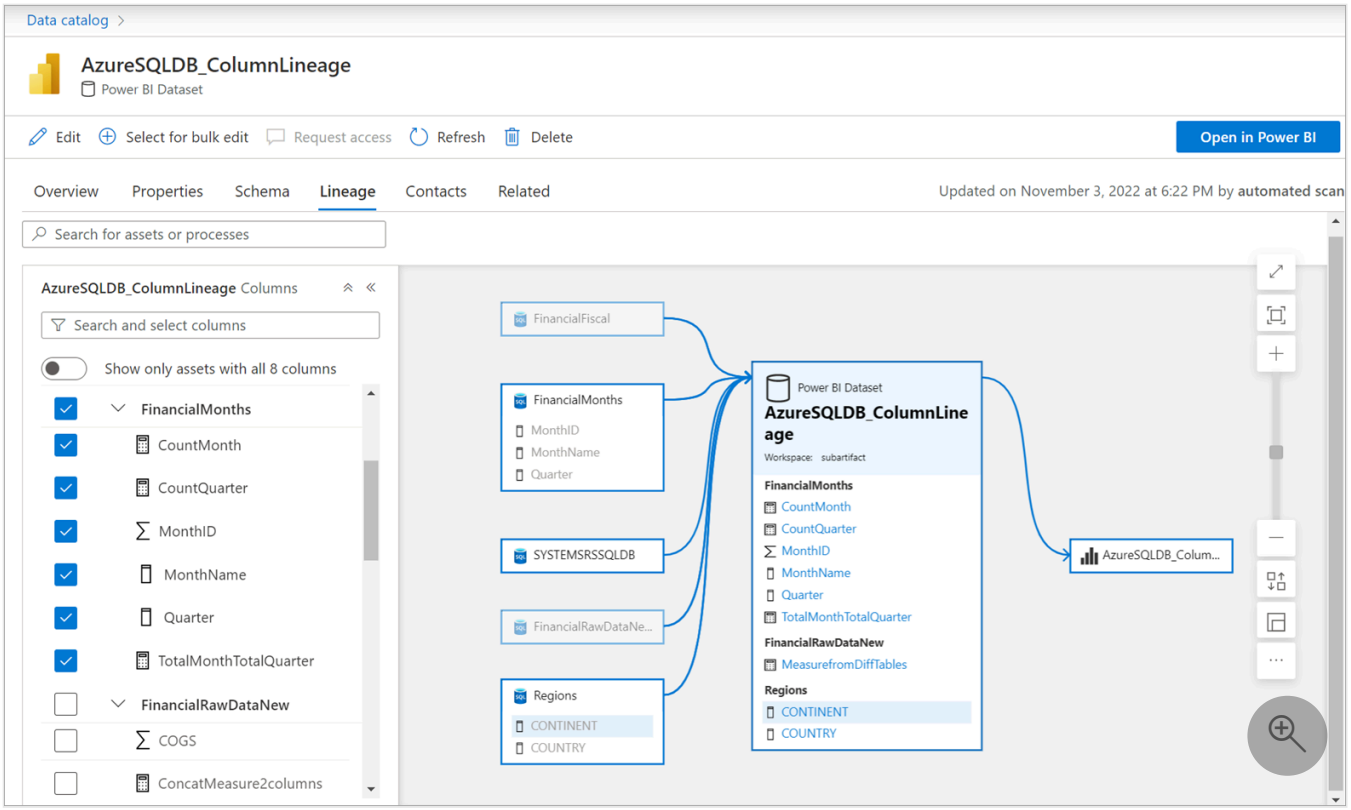
Microsoft Purview captures lineage among Power BI artifacts (for example: Dataflow -> Dataset -> Report -> Dashboard) and external data assets.

> **Note**
>
> For lineage between Power BI artifacts and external data assets, currently the supported source types are:
>
> - Azure SQL Database
> - Azure Blob Storage
> - Azure Data Lake Store Gen1
> - Azure Data Lake Store Gen2



In addition, column level lineage (Power BI subartifact lineage) and transformation inside of Power BI datasets (except for Dataflows) are captured when using Azure SQL Database as source. For measures, you can further select column -> Properties -> expression to see the transformation details.

> **Note**
>
> Column level lineage and transformations (except for Dataflows) are supported when using Azure SQL Database as source. Other sources are currently not supported.



Known limitations

- Microsoft Purview leverages the scanner API to retrieve the metadata and lineage. Learn about some API limitations from Metadata scanning - Considerations and limitations.
- In case you have the dataset table connected to another dataset table, when the middle dataset disables the "Enable load" option inside the Power BI desktop, and the lineage can't be extracted.
- For lineage between Power BI artifacts and external data assets:
  - Currently the supported source types are Azure SQL Database, Azure Blob Storage, Azure Data Lake Store Gen1 and Azure Data Lake Store Gen2.
  - Column level lineage and transformations (except for Power BI Dataflows) are only supported when using Azure SQL Database as source. Other sources are currently not supported.
  - Lineage is not captured when you use dynamic M query parameters in Power BI, e.g. pass server/database names as parameter values.
  - Limited information is currently shown for data sources where the Power BI Dataflow is created. For example, for a SQL server source of Power BI dataset, only server/database name is captured.
  - Note due to the following limitations, if you have such scenarios and scan both Power BI and the data sources that Power BI artifacts connect to, currently you may see duplicate assets in the catalog.
    - The source object names in assets and fully qualified names follow the case used in Power BI settings/queries, which may not align with the object case in original data

source.
    - When Power BI references SQL views, they are currently captured as SQL table assets.
    - When Power BI references Azure Dedicated SQL pool (formerly SQL DW) sources, it's currently listed as Azure SQL Database assets. (Lineage between Power BI artifacts and external Azure Dedicated SQL pools isn't currently supported.)
- For Power BI subartifact lineage:
    - Some measures aren't shown in the subartifact lineage, for example, `COUNTROWS`.
    - In the lineage graph, when selecting a measure that is derived by columns using the COUNT function, the underlying column isn't selected automatically. Check the measure expression in the column properties tab to identify the underlying column.
    - If you scanned your Power BI source before subartifact lineage was supported, you may see a database asset along with the new table assets in the lineage graph, which isn't removed.
- Currently for the same data asset in Azure SQL Database which has UTF-8 characters in name, the full qualified name generated from scanning Power BI may be different with the fully qualified name generated from scanning Azure SQL Database as the former will go through additional encoding, and it may introduce duplicate Azure SQL data assets.

Next steps

- Learn about Data lineage in Microsoft Purview
- Link Azure Data Factory to push automated lineage

How to get lineage from SQL Server Integration Services (SSIS) into Microsoft Purview

07/18/2025

This article elaborates on the data lineage aspects of SQL Server Integration Services (SSIS) in Microsoft Purview.

Prerequisites

- Lift and shift SQL Server Integration Services workloads to the cloud

Supported scenarios

The current scope of support includes the lineage extraction from SSIS packages executed by Azure Data Factory SSIS integration runtime.

On premises SSIS lineage extraction is not supported yet.

Only source and destination are supported for Microsoft Purview SSIS lineage running from Data Factory's SSIS Execute Package activity. Transformations under SSIS are not yet supported.

Supported data stores

Expand table

Data store	Supported
Azure Blob Storage	Yes
Azure Data Lake Storage Gen1	Yes
Azure Data Lake Storage Gen2	Yes
Azure Files	Yes
Azure SQL Database *	Yes
Azure SQL Managed Instance *	Yes
Azure Synapse Analytics *	Yes

Data store	Supported
SQL Server *	Yes

** Microsoft Purview currently doesn't support query or stored procedure for lineage or scanning. Lineage is limited to table and view sources only.*

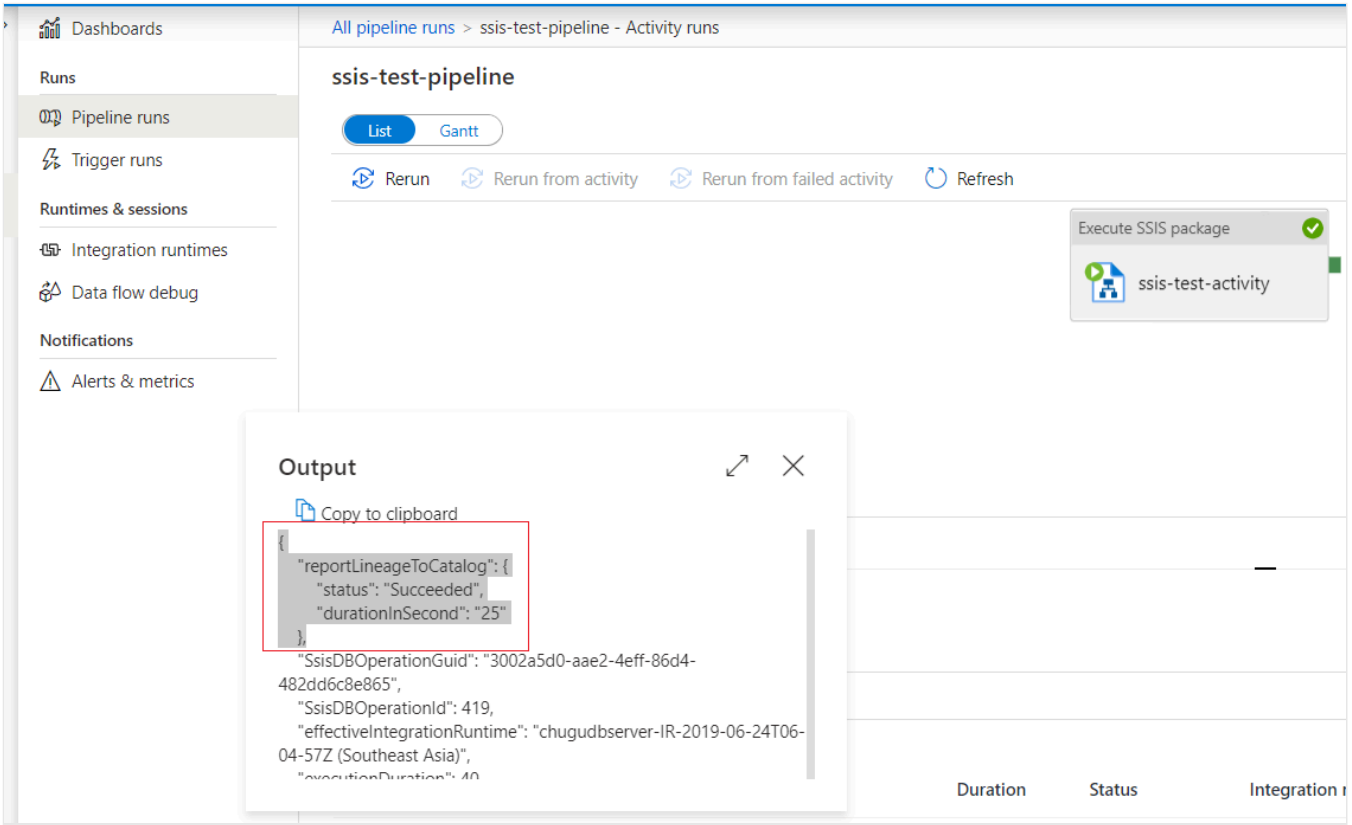
How to bring SSIS lineage into Microsoft Purview

Step 1: Connect a Data Factory to Microsoft Purview

Step 2: Trigger SSIS activity execution in Azure Data Factory

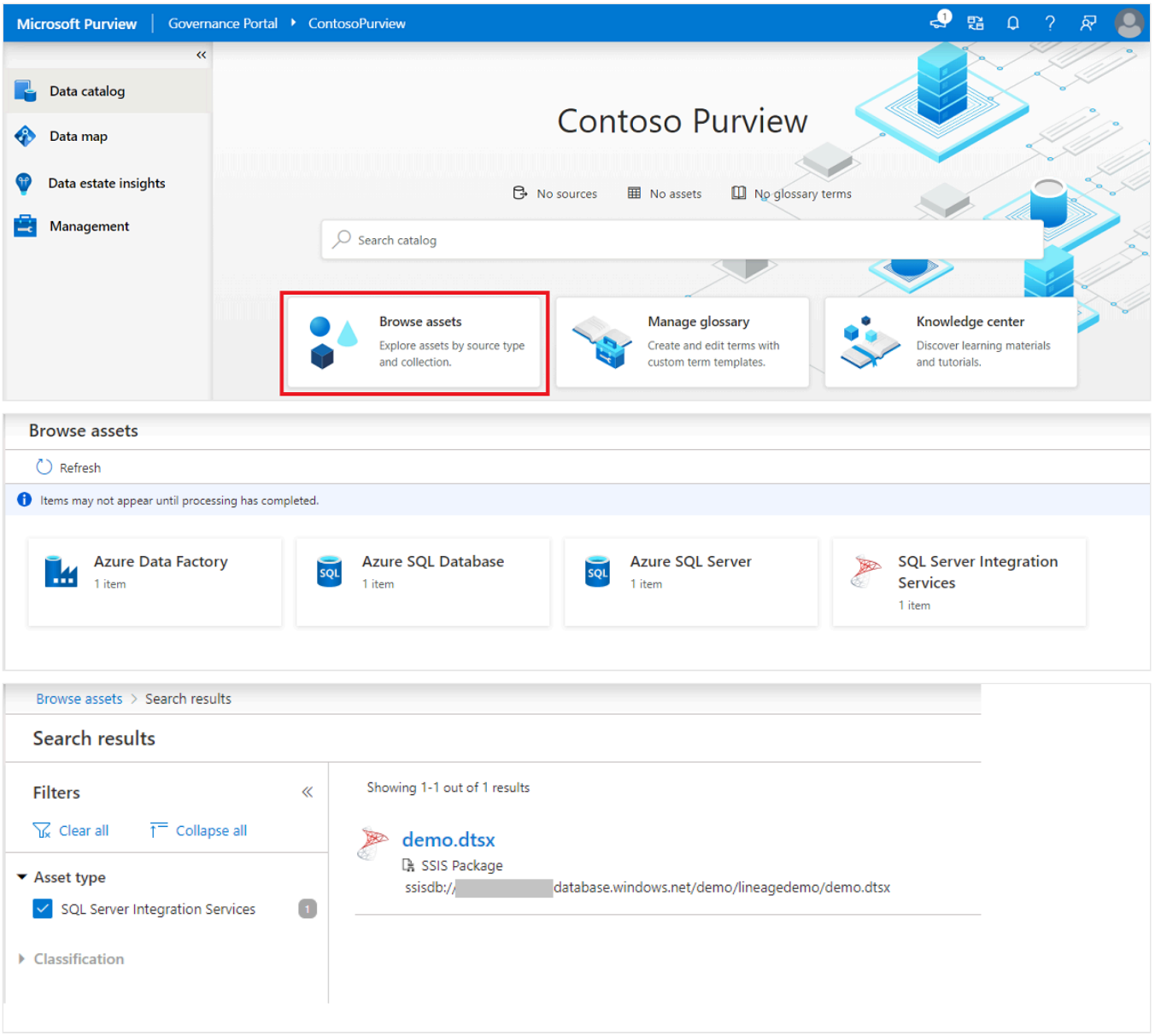
You can run SSIS package with Execute SSIS Package activity or run SSIS package with Transact-SQL in ADF SSIS Integration Runtime.

Once Execute SSIS Package activity finishes the execution, you can check lineage report status from the activity output in Data Factory activity monitor.

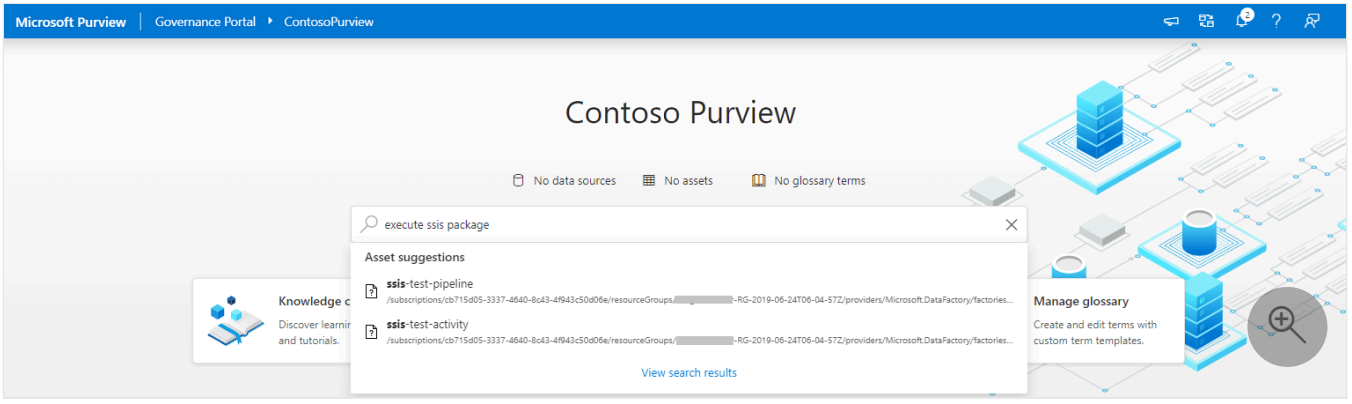


Step 3: Browse lineage Information in your Microsoft Purview account

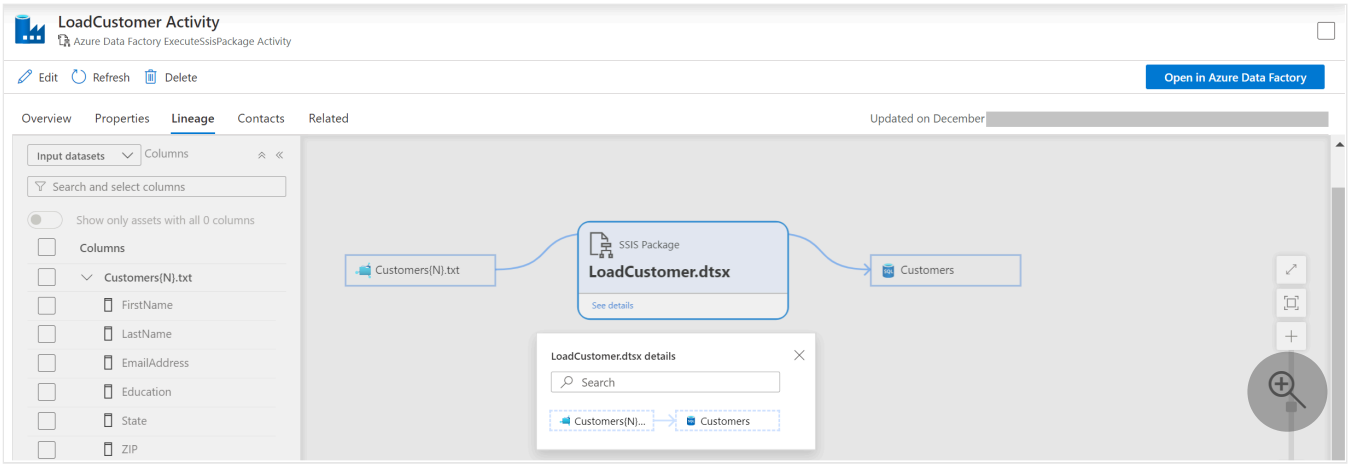
- You can browse Unified Catalog by choosing asset type “SQL Server Integration Services”.



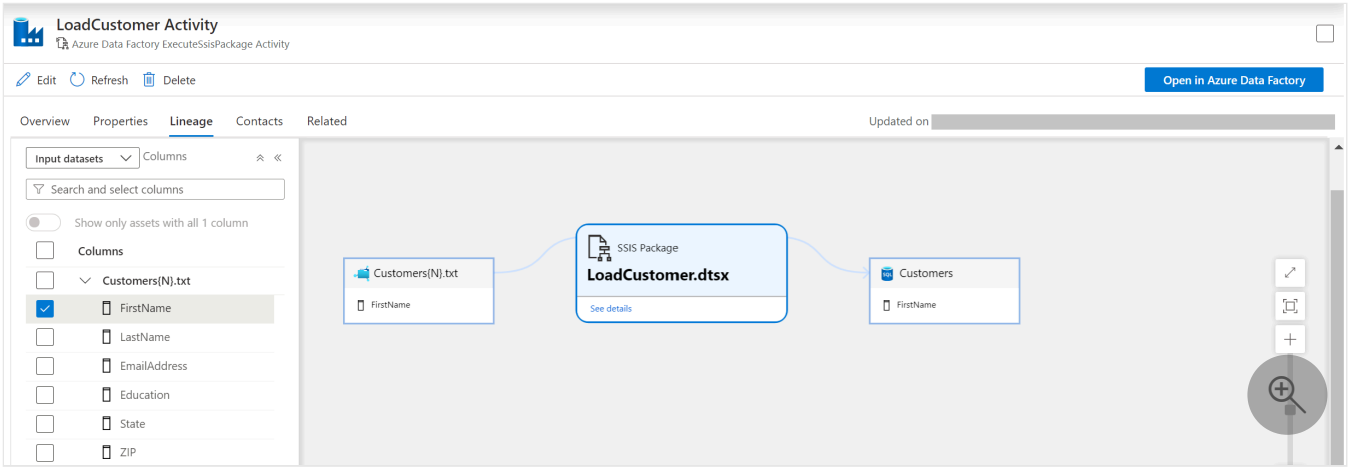
- You can also search Unified Catalog using keywords.



- You can view lineage information for an SSIS Execute Package activity and open in Data Factory to view/edit the activity settings.



- You can choose one data source to drill into how the columns in the source are mapped to the columns in the destination.



Next steps

- Lift and shift SQL Server Integration Services workloads to the cloud
- Learn about Data lineage in Microsoft Purview
- Link Azure Data Factory to push automated lineage

Connections page in the Microsoft Purview Data Map

Article • 09/10/2024

The **Connections** page in the Microsoft Purview Data Map is a page for creating unique connections to services to use more of Microsoft Purview Data Map's features.

This page is **not** for registered data sources. Instead this page is for connections in workflows and protection policies.

To view and manage connections, you can find the connections page by:

- In the Microsoft Purview portal:

    1. Select the Data Map application.
    2. Select the **Source management** dropdown.
    3. Select the **Connections** page.

- In the classic Microsoft Purview governance portal:

    1. Select the **Management** application.
    2. Under the **Security and access** list, select **Connections**.

Use

You might use the connection page when creating connections for some workflows or protection policies that require you to make special connections to partner resources.

For example:

- Connect to ServiceNow in workflows
- Create a policy enforcement connection for Amazon S3

On this page, you can edit and delete your connections by hovering over a connection and either selecting the **Edit** pencil button, or the **Delete** trashcan button.

Related content

- Workflows
- Protection policies

Feedback

Was this page helpful? Yes No

Configure Event Hubs with Microsoft Purview to send and receive Atlas Kafka topics messages

07/21/2025

This article explains how to configure Microsoft Purview to be able to send and receive *Atlas Kafka* topics events with Azure Event Hubs.

If you have already configured your environment, learn how to get started with the **Azure.Messaging.EventHubs** .NET library to send and receive messages.

Prerequisites

To configure your environment, you need certain prerequisites in place:

- **A Microsoft Azure subscription**. To use Azure services, including Event Hubs, you need an Azure subscription. If you don't have an Azure account, you can sign up for a free trial or use your MSDN subscriber benefits when you create an account.
- An active Microsoft Purview account.

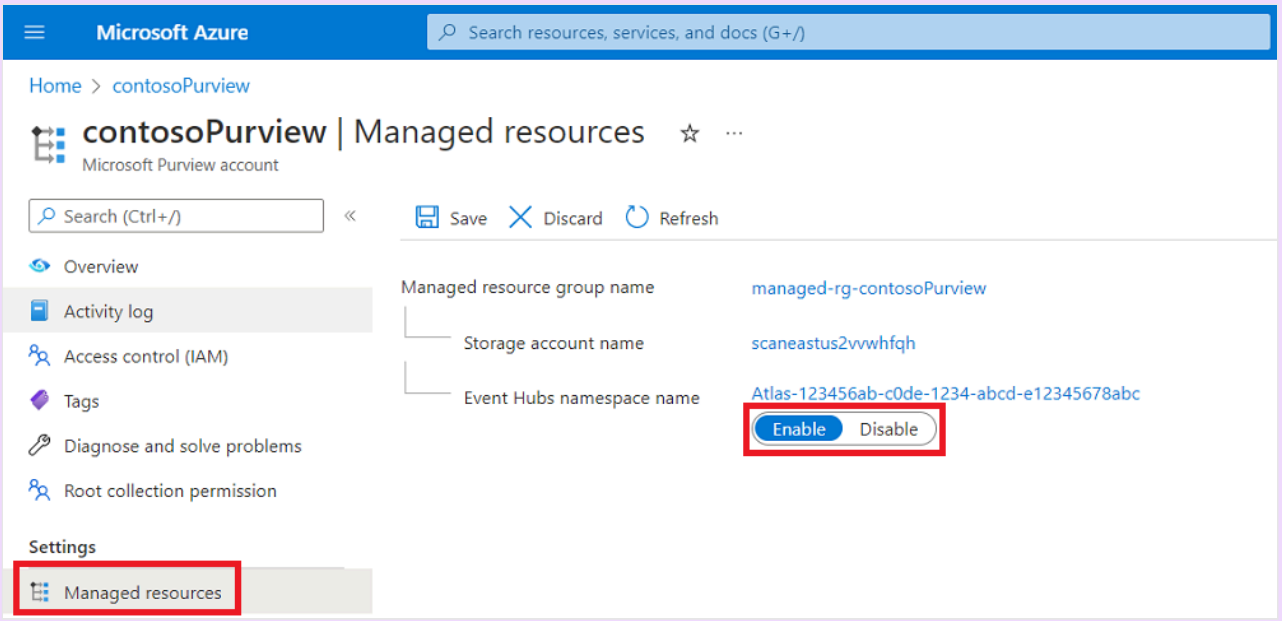
Configure Event Hubs

To send or receive Atlas Kafka topics messages, you'll need to configure at least one Event Hubs namespace. You can either:

- Create your own Azure Event Hubs namespace and configure it to work with Microsoft Purview
- Use the REST API to create a managed namespace

> **Note**
>
> If your Microsoft Purview account was created before December 15, 2022, you might have a managed Event Hubs resource already associated with your account. You can check in **Managed Resources** under settings on your Microsoft Purview account page in the Azure

portal .



- If you don't see this resource, or it's **disabled**, follow the steps below to configure your Event Hubs.
- If it's **enabled**, you can continue to use this managed Event Hubs namespace if you prefer. (There's associated cost. See see the pricing page .) If you want to manage your own Event Hubs account, you must first **disable** this feature and follow the steps below. **To keep the Event Hub's supported TLS version updated, we recommend using BYO Event Hub. If you disable the managed Event Hubs resource you won't be able to re-enable a managed Event Hub resource. You will only be able to configure your own Event Hubs.**

Bring your own Event Hubs namespace

You can configure an existing Azure Event Hubs namespace with an Event Hubs or create a new one to connect with Microsoft Purview.

Event Hubs permissions

To authenticate with your Event Hubs, you can either use:

- Microsoft Purview managed identity(Recommended)
- User assigned managed identity - only available when configuring namespaces after account creation **and** if using Event Hubs with an open network.

These identities need **at least contributor permissions on your Event Hubs** to be able to configure them to use with Microsoft Purview.

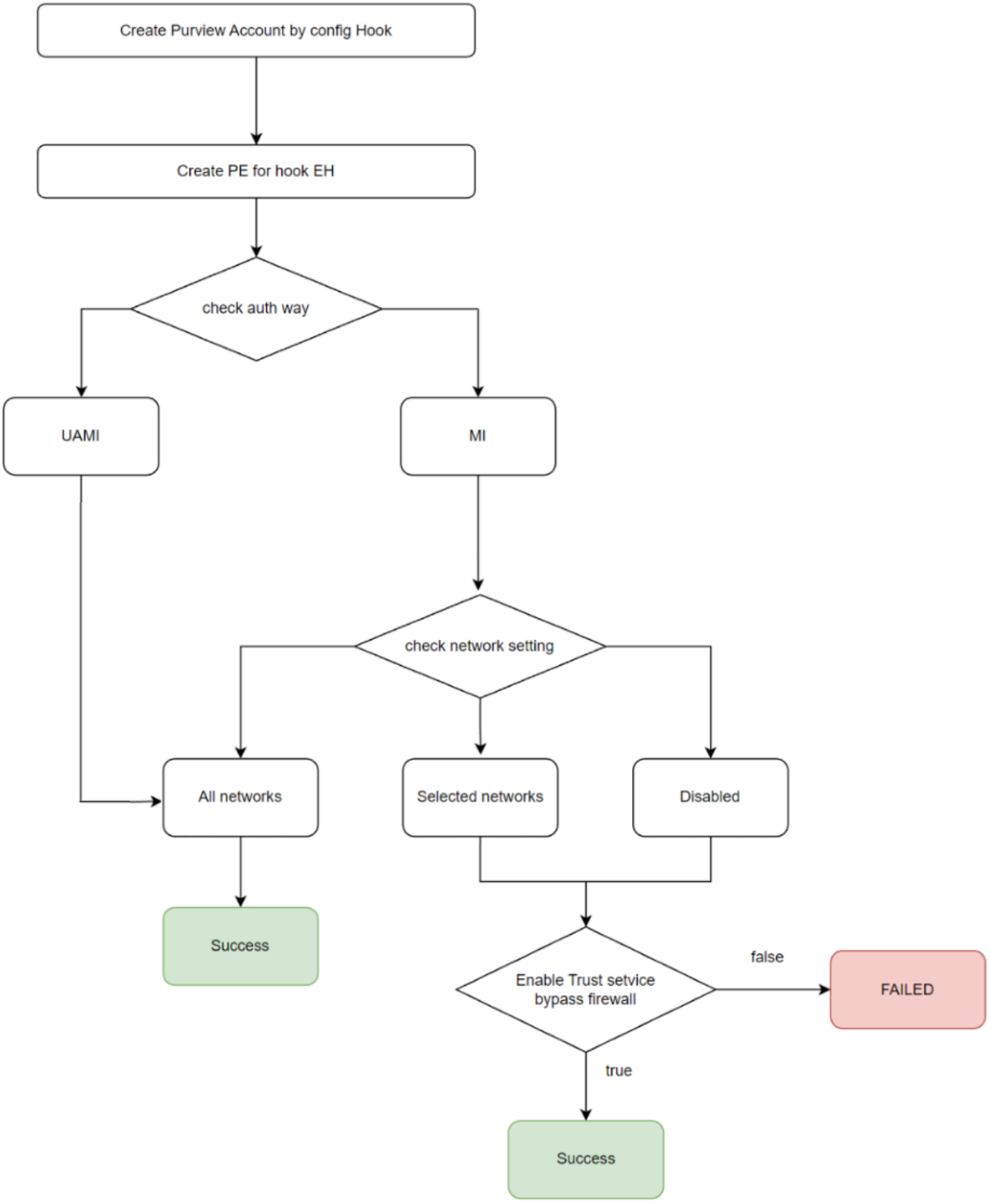
Event Hubs networking

1. If your Microsoft Purview account is using private endpoints, set up an ingestion private endpoint to connect to Event Hubs.

2. Your Event Hubs workspace network needs to be configured in one of two ways:

    - *Public network access* set to **Selected networks** or **Disabled**, *Allow trusted Microsoft services to bypass this firewall* set to **Yes**, and using the Microsoft Purview managed identity to authenticate.
    - *Public network access* set to **All networks** and using the Microsoft Purview managed identity or a user assigned managed identity to authenticate.

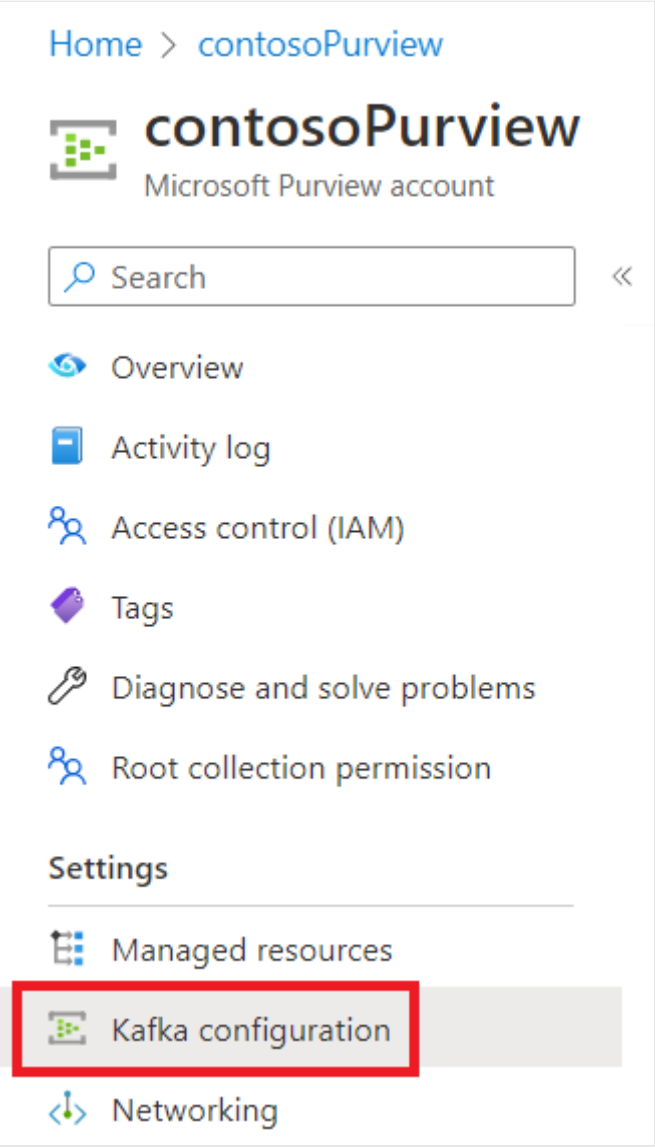
    Microsoft Purview's status as a trusted service means that, as long as *Allow trusted Microsoft services to bypass this firewall* is set to **Yes** on your Event Hubs resource, your hooks can still receive messages from Event Hubs.



For more information about networking for Event Hubs, see the Azure Event Hubs documentation.

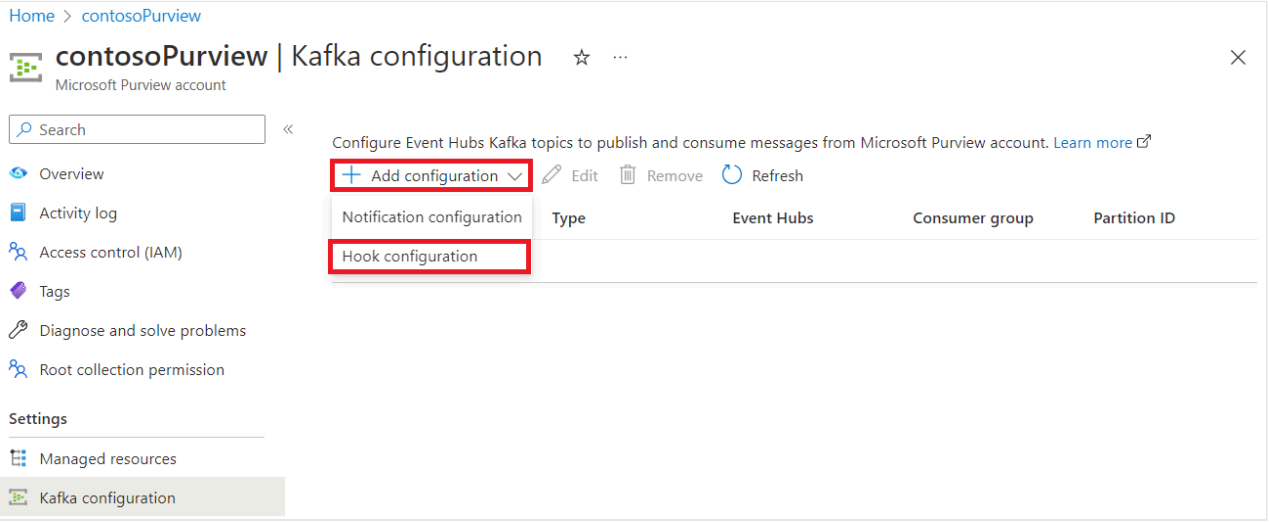
Configure Event Hubs to publish messages to Microsoft Purview

1. Navigate to **Kafka configuration** under settings on your Microsoft Purview account page in the Azure portal.



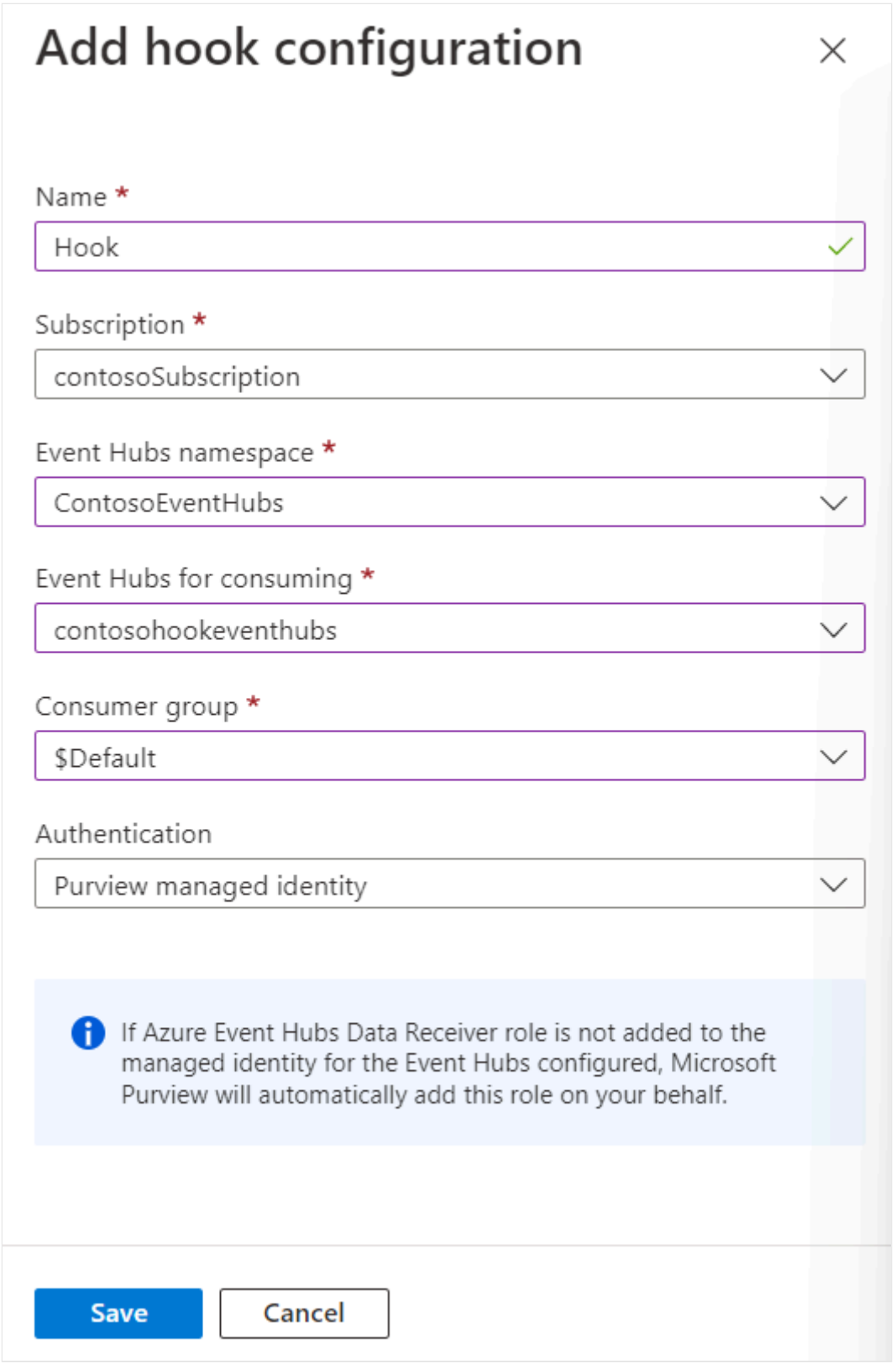
2. Select **Add configuration** and **Hook configuration**.

> **Note**
>
> You can add as many hook configurations as you need.



3. Give a name to your hook configuration, select your subscription, an existing Event Hubs namespace, an existing Event Hubs to send the notifications to, the consumer group you want to use, and the kind of authentication you would like to use.

   > **Tip**
   >
   > You can use the same Event Hubs namespace more than once, but each configuration will need its own Event Hubs.

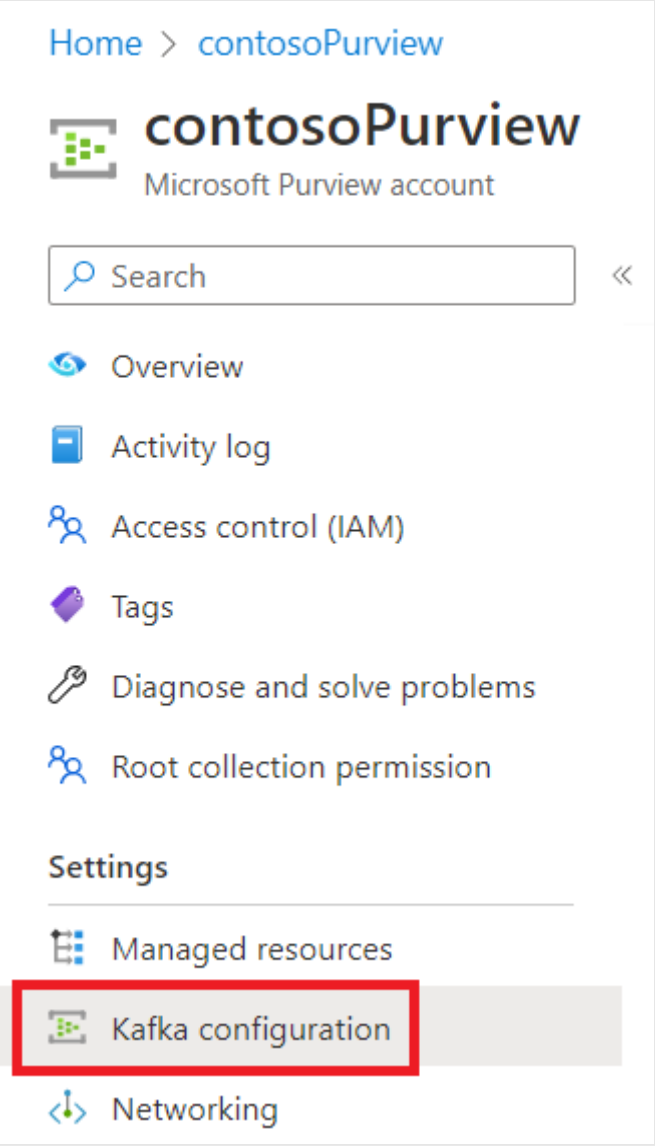


4. Select **Save**. It takes a couple minutes for your configuration to complete.

5. Once configuration is complete, you can begin the steps to publish messages to Microsoft Purview.

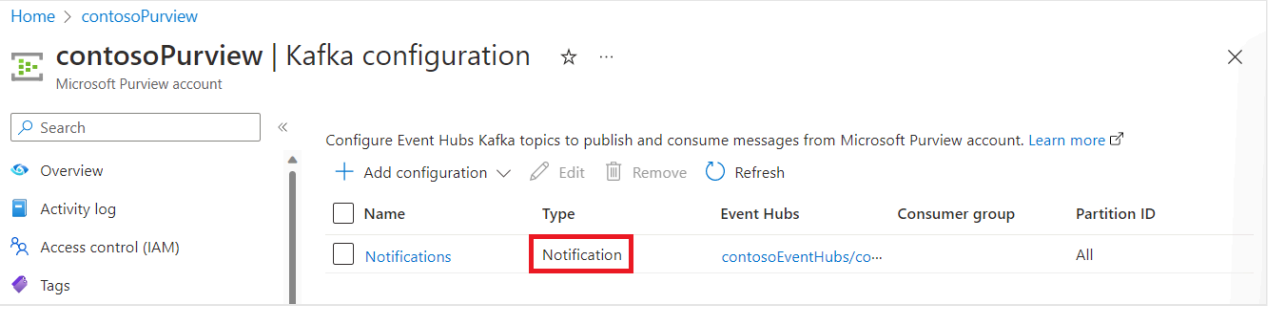
Configure Event Hubs to receive messages from Microsoft Purview

1. Navigate to **Kafka configuration** under settings on your Microsoft Purview account page in the Azure portal.

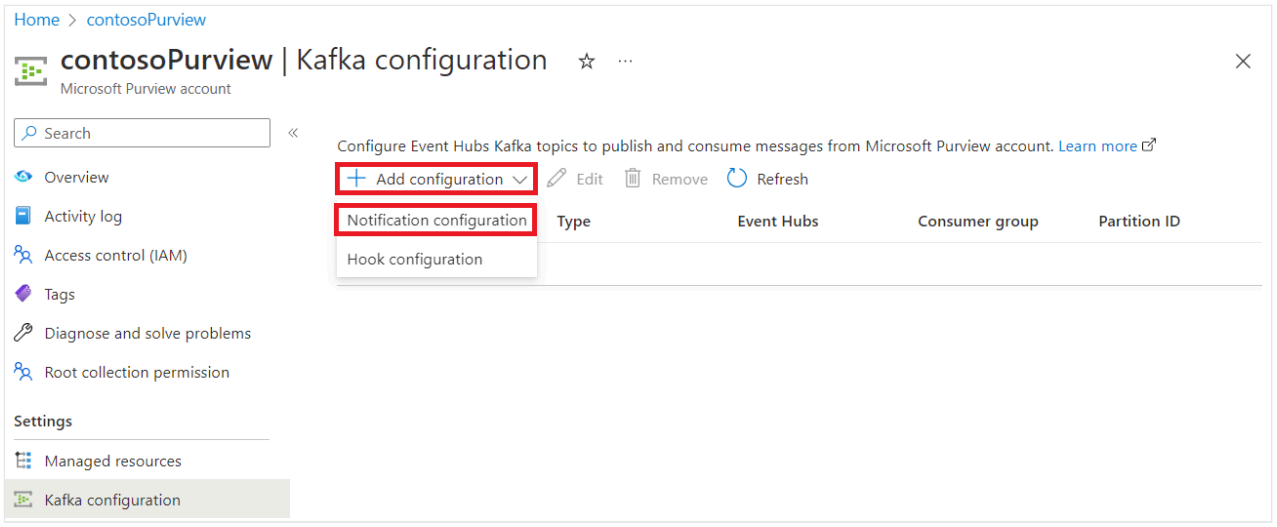


2. If there's a configuration already listed as type **Notification**, Event Hubs is already configured, and you can begin the steps to receive Microsoft Purview messages.

> **Note**
>
> Only one *Notification* Event Hubs can be configured at a time.

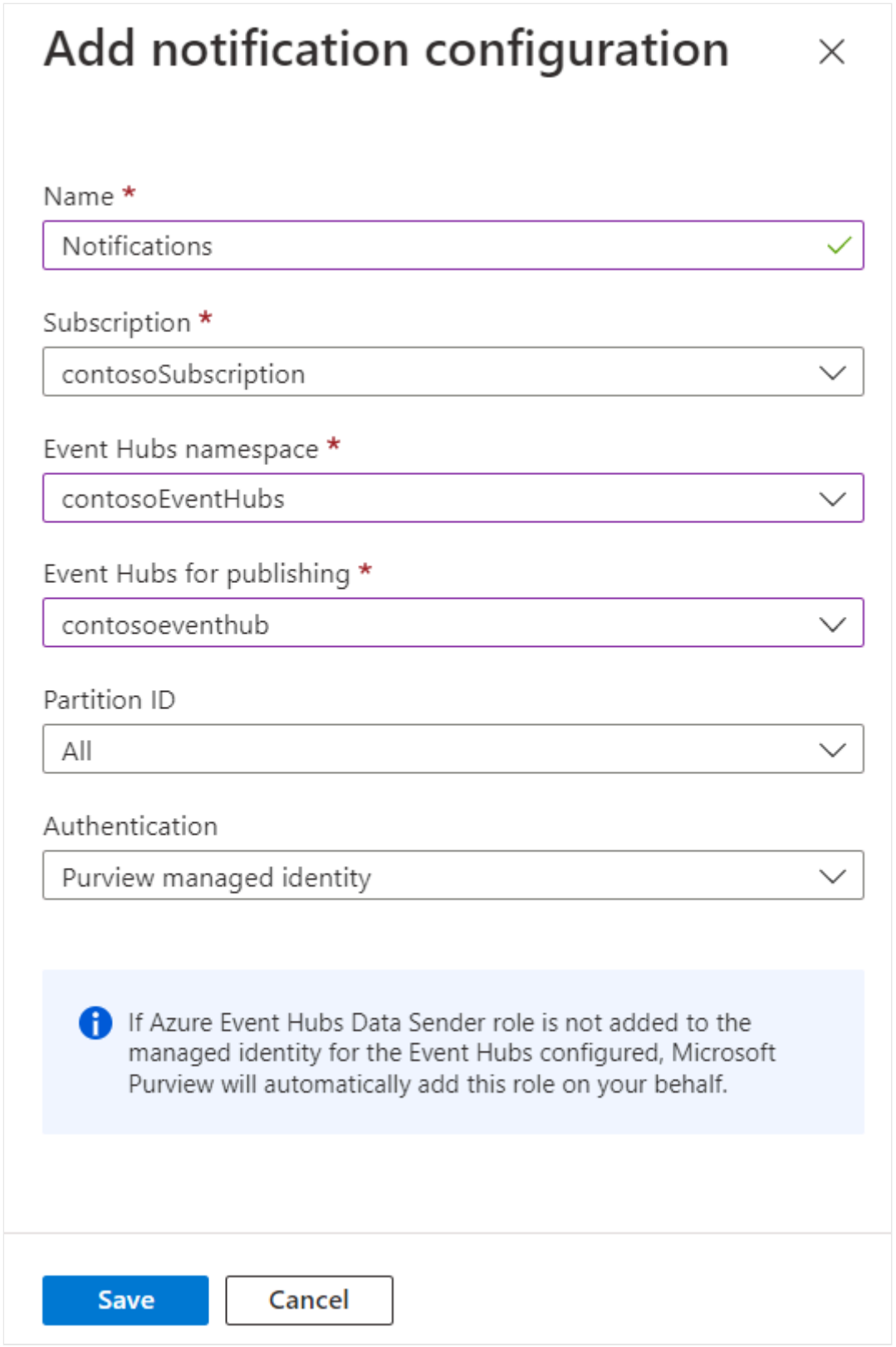


3. If there isn't a **Notification** configuration already listed, select **Add configuration** and **Notification configuration**.



4. Give a name to your notification configuration, select your subscription, an existing Event Hubs namespace, an existing Event Hubs to send the notifications to, the partitions you want to use, and the kind of authentication you would like to use.

> **Tip**
>
> You can use the same Event Hubs namespace more than once, but each configuration will need its own Event Hubs.



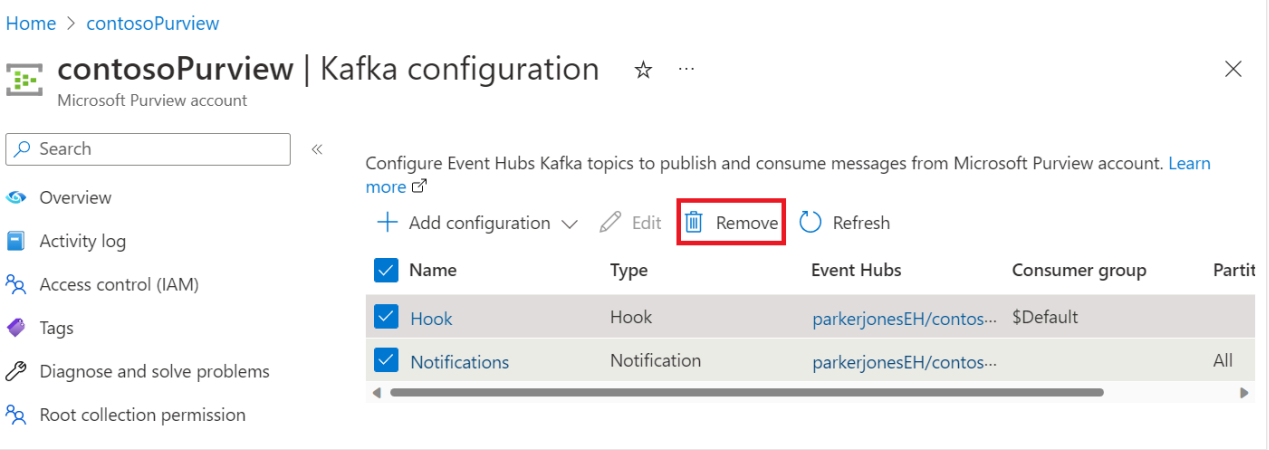
5. Select **Save**. It takes a couple minutes for your configuration to complete.

6. Once configuration is complete, you can begin the steps to receive Microsoft Purview messages.

Remove configured Event Hubs

To remove configured Event Hubs namespaces, you can follow these steps:

1. Search for and open your Microsoft Purview account in the Azure portal.
2. Select **Kafka configuration** under settings on your Microsoft Purview account page in the Azure portal.
3. Select the Event Hubs you want to disable. (Hook hubs send messages to Microsoft Purview. Notification hubs receive notifications.)
4. Select **Remove** to save the choice and begin the disablement process. This can take several minutes to complete.



Create a managed namespace

If you prefer to have a managed Event Hubs namespace, you need to use the REST API to configure your account.

You can use below script (set apiVersion = 2021-12-01) to update an account and this way will only affect the event hub configuration:

```
$body = @{"properties" = @{ "managedEventHubState" = "Enabled" }; "location" =
$($location); "identity" = @{ "type"= $($type) } } | ConvertTo-Json
$Token=Get-AzAccessToken
$Headers = @{ "Authorization" = "Bearer $($Token.Token)" }
$Uri =
"https://management.azure.com/subscriptions/$($subscription)/resourceGroups/$($rg)
/providers/Microsoft.Purview/accounts/$($accountName)?api-version=$($apiVersion)"

Invoke-WebRequest -URI $Uri -Method Put -Body $body -Headers $Headers -ContentType
"application/json"
```

Next steps

- Publish and process Atlas Kafka topics messages using your Event Hubs
- Event Hubs samples in GitHub
- Event processor samples in GitHub
- An introduction to Atlas notifications

Use Event Hubs and .NET to send and receive Atlas Kafka topics messages

07/21/2025

This quickstart teaches you how to send and receive *Atlas Kafka* topics events. We'll make use of *Azure Event Hubs* and the **Azure.Messaging.EventHubs** .NET library.

Prerequisites

If you're new to Event Hubs, see Event Hubs overview before you complete this quickstart.

To follow this quickstart, you need certain prerequisites in place:

- **A Microsoft Azure subscription**. To use Azure services, including Event Hubs, you need an Azure subscription. If you don't have an Azure account, you can sign up for a free trial or use your MSDN subscriber benefits when you create an account.
- **Microsoft Visual Studio 2022**. The Event Hubs client library makes use of new features that were introduced in C# 8.0. You can still use the library with previous C# versions, but the new syntax won't be available. To make use of the full syntax, it's recommended that you compile with the .NET Core SDK 3.0 or higher and language version set to `latest`. If you're using a Visual Studio version prior to Visual Studio 2019, it doesn't have the tools needed to build C# 8.0 projects. Visual Studio 2022, including the free Community edition, can be downloaded here.
- An active Microsoft Purview account.
- An Event Hubs configured with your Microsoft Purview account to sent and receive messages:
  - Your account may already be configured. You can check your Microsoft Purview account in the Azure portal under **Settings**, **Kafka configuration**. If it isn't already configured, follow this guide.

Publish messages to Microsoft Purview

Let's create a .NET Core console application that sends events to Microsoft Purview via Event Hubs Kafka topic, **ATLAS_HOOK**.

To publish messages to Microsoft Purview, you need either a managed Event Hubs, or at least one Event Hubs with a hook configuration..

Create a Visual Studio project

Next create a C# .NET console application in Visual Studio:

1. Launch **Visual Studio**.
2. In the Start window, select **Create a new project** > **Console App (.NET Framework)**. .NET version 4.5.2 or above is required.
3. In **Project name**, enter **PurviewKafkaProducer**.
4. Select **Create** to create the project.

Create a console application

1. Start Visual Studio 2022.
2. Select **Create a new project**.
3. On the **Create a new project** dialog box, do the following steps: If you don't see this dialog box, select **File** on the menu, select **New**, and then select **Project**.
   a. Select **C#** for the programming language.
   b. Select **Console** for the type of the application.
   c. Select **Console App (.NET Core)** from the results list.
   d. Then, select **Next**.

Add the Event Hubs NuGet package

1. Select **Tools** > **NuGet Package Manager** > **Package Manager Console** from the menu.

2. Run the following command to install the **Azure.Messaging.EventHubs** NuGet package and **Azure.Messaging.EventHubs.Producer** NuGet package:

   Windows Command Prompt

   ```
   Install-Package Azure.Messaging.EventHubs
   ```

   Windows Command Prompt

   ```
   Install-Package Azure.Messaging.EventHubs.Producer
   ```

Write code that sends messages to the event hub

1. Add the following `using` statements to the top of the **Program.cs** file:

   C#

   ```
   using System;
   using System.Text;
   ```

```
using System.Threading.Tasks;
using Azure.Messaging.EventHubs;
using Azure.Messaging.EventHubs.Producer;
```

2. Add constants to the `Program` class for the Event Hubs connection string and Event Hubs name.

```
C#

private const string connectionString = "<EVENT HUBS NAMESPACE - CONNECTION
STRING>";
private const string eventHubName = "<EVENT HUB NAME>";
```

3. Replace the `Main` method with the following `async Main` method and add an `async ProduceMessage` to push messages into Microsoft Purview. See the comments in the code for details.

```
C#

    static async Task Main()
    {
        // Read from the default consumer group: $Default
        string consumerGroup =
EventHubConsumerClient.DefaultConsumerGroupName;

        / Create an event producer client to add events in the event hub
        EventHubProducerClient producer = new
EventHubProducerClient(ehubNamespaceConnectionString, eventHubName);

        await ProduceMessage(producer);
    }

    static async Task ProduceMessage(EventHubProducerClient producer)

    {
        // Create a batch of events
        using EventDataBatch eventBatch = await
producerClient.CreateBatchAsync();

        // Add events to the batch. An event is a represented by a collection
of bytes and metadata.
        eventBatch.TryAdd(new EventData(Encoding.UTF8.GetBytes("<First
event>")));
        eventBatch.TryAdd(new EventData(Encoding.UTF8.GetBytes("<Second
event>")));
        eventBatch.TryAdd(new EventData(Encoding.UTF8.GetBytes("<Third
event>")));

        // Use the producer client to send the batch of events to the event
hub
```

```
        await producerClient.SendAsync(eventBatch);
        Console.WriteLine("A batch of 3 events has been published.");

    }
```

4. Build the project. Ensure that there are no errors.

5. Run the program and wait for the confirmation message.

> ⓘ Note
>
> For the complete source code with more informational comments, see this file in GitHub

Sample code that creates an sql table with two columns using a Create Entity JSON message

JSON

```
    {
    "msgCreatedBy":"nayenama",
    "message":{
        "type":"ENTITY_CREATE_V2",
        "user":"admin",
        "entities":{
            "entities":[
                {
                    "typeName":"azure_sql_table",
                    "attributes":{
                        "owner":"admin",
                        "temporary":false,

"qualifiedName":"mssql://nayenamakafka.eventhub.sql.net/salespool/dbo/SalesOrderTa
ble",
                        "name":"SalesOrderTable",
                        "description":"Sales Order Table added via Kafka"
                    },
                    "relationshipAttributes":{
                        "columns":[
                            {
                                "guid":"-1102395743156037",
                                "typeName":"azure_sql_column",
                                "uniqueAttributes":{

"qualifiedName":"mssql://nayenamakafka.eventhub.sql.net/salespool/dbo/SalesOrderTa
ble#OrderID"
                                }
                            },
```

```
                            {
                                "guid":"-1102395743156038",
                                "typeName":"azure_sql_column",
                                "uniqueAttributes":{

"qualifiedName":"mssql://nayenamakafka.eventhub.sql.net/salespool/dbo/SalesOrderTa
ble#OrderDate"
                                }
                            }
                        ]
                    },
                    "guid":"-1102395743156036",
                    "version":0
                }
            ],
            "referredEntities":{
                "-1102395743156037":{
                    "typeName":"azure_sql_column",
                    "attributes":{
                        "owner":null,
                        "userTypeId":61,

"qualifiedName":"mssql://nayenamakafka.eventhub.sql.net/salespool/dbo/SalesOrderTa
ble#OrderID",
                        "precision":23,
                        "length":8,
                        "description":"Sales Order ID",
                        "scale":3,
                        "name":"OrderID",
                        "data_type":"int"
                    },
                    "relationshipAttributes":{
                        "table":{
                            "guid":"-1102395743156036",
                            "typeName":"azure_sql_table",
                            "entityStatus":"ACTIVE",
                            "displayText":"SalesOrderTable",
                            "uniqueAttributes":{

"qualifiedName":"mssql://nayenamakafka.eventhub.sql.net/salespool/dbo/SalesOrderTa
ble"
                            }
                        }
                    },
                    "guid":"-1102395743156037",
                    "version":2
                },
                "-1102395743156038":{
                    "typeName":"azure_sql_column",
                    "attributes":{
                        "owner":null,
                        "userTypeId":61,

"qualifiedName":"mssql://nayenamakafka.eventhub.sql.net/salespool/dbo/SalesOrderTa
ble#OrderDate",
```

```
                        "description":"Sales Order Date",
                        "scale":3,
                        "name":"OrderDate",
                        "data_type":"datetime"
                    },
                    "relationshipAttributes":{
                        "table":{
                            "guid":"-1102395743156036",
                            "typeName":"azure_sql_table",
                            "entityStatus":"ACTIVE",
                            "displayText":"SalesOrderTable",
                            "uniqueAttributes":{

"qualifiedName":"mssql://nayenamakafka.eventhub.sql.net/salespool/dbo/SalesOrderTa
ble"
                            }
                        }
                    },
                    "guid":"-1102395743156038",
                    "status":"ACTIVE",
                    "createdBy":"ServiceAdmin",
                    "version":0
                }
            }
        }
    },
    "version":{
        "version":"1.0.0"
    },
    "msgCompressionKind":"NONE",
    "msgSplitIdx":1,
    "msgSplitCount":1
}
```


Receive Microsoft Purview messages

Next learn how to write a .NET Core console application that receives messages from event hubs using an event processor. The event processor manages persistent checkpoints and parallel receptions from event hubs. This simplifies the process of receiving events. You need to use the ATLAS_ENTITIES event hub to receive messages from Microsoft Purview.

To receive messages from Microsoft Purview, you need either a managed Event Hubs, or an Event Hubs notification configuration..

> ⚠ **Warning**

Event Hubs SDK uses the most recent version of Storage API available. That version may not necessarily be available on your Stack Hub platform. If you run this code on Azure Stack Hub, you'll experience runtime errors unless you target the specific version you're using. If you're using Azure Blob Storage as a checkpoint store, review the **supported Azure Storage API version for your Azure Stack Hub build** and in your code, target that version.

The highest available version of the Storage service is version 2019-02-02. By default, the Event Hubs SDK client library uses the highest available version on Azure (2019-07-07 at the time of the release of the SDK). If you're using Azure Stack Hub version 2005, in addition to following the steps in this section, you'll also need to add code that targets the Storage service API version 2019-02-02. To learn how to target a specific Storage API version, see **this sample in GitHub**.

Create an Azure Storage and a blob container

We'll use Azure Storage as the checkpoint store. Use the following steps to create an Azure Storage account.

1. Create an Azure Storage account
2. Create a blob container
3. Get the connection string for the storage account

   Make note of the connection string and the container name. You'll use them in the receive code.

Create a Visual Studio project for the receiver

1. In the Solution Explorer window, select and hold (or right-click) the **EventHubQuickStart** solution, point to **Add**, and select **New Project**.
2. Select **Console App (.NET Core)**, and select **Next**.
3. Enter **PurviewKafkaConsumer** for the **Project name**, and select **Create**.

Add the Event Hubs NuGet package

1. Select **Tools** > **NuGet Package Manager** > **Package Manager Console** from the menu.
2. Run the following command to install the **Azure.Messaging.EventHubs** NuGet package:

   Windows Command Prompt

```
Install-Package Azure.Messaging.EventHubs
```

3. Run the following command to install the **Azure.Messaging.EventHubs.Processor** NuGet package:

```
Windows Command Prompt
```

```
Install-Package Azure.Messaging.EventHubs.Processor
```

Update the Main method

1. Add the following `using` statements at the top of the **Program.cs** file.

```
C#
```

```
using System;
using System.Text;
using System.Threading.Tasks;
using Azure.Storage.Blobs;
using Azure.Messaging.EventHubs;
using Azure.Messaging.EventHubs.Consumer;
using Azure.Messaging.EventHubs.Processor;
```

2. Add constants to the `Program` class for the Event Hubs connection string and the event hub name. Replace placeholders in brackets with the real values that you got when you created the event hub and the storage account (access keys - primary connection string). Make sure that the `{Event Hubs namespace connection string}` is the namespace-level connection string, and not the event hub string.

```
C#
```

```
    private const string ehubNamespaceConnectionString = "<EVENT HUBS
NAMESPACE - CONNECTION STRING>";
    private const string eventHubName = "<EVENT HUB NAME>";
    private const string blobStorageConnectionString = "<AZURE STORAGE
CONNECTION STRING>";
    private const string blobContainerName = "<BLOB CONTAINER NAME>";
```

   Use **ATLAS_ENTITIES** as the event hub name when sending messages to Microsoft Purview.

3. Replace the `Main` method with the following `async Main` method. See the comments in the code for details.

C#

```
    static async Task Main()
    {
        // Read from the default consumer group: $Default
        string consumerGroup =
EventHubConsumerClient.DefaultConsumerGroupName;

        // Create a blob container client that the event processor will use
        BlobContainerClient storageClient = new
BlobContainerClient(blobStorageConnectionString, blobContainerName);

        // Create an event processor client to process events in the event
hub
        EventProcessorClient processor = new
EventProcessorClient(storageClient, consumerGroup,
ehubNamespaceConnectionString, eventHubName);

        // Register handlers for processing events and handling errors
        processor.ProcessEventAsync += ProcessEventHandler;
        processor.ProcessErrorAsync += ProcessErrorHandler;

        // Start the processing
        await processor.StartProcessingAsync();

        // Wait for 10 seconds for the events to be processed
        await Task.Delay(TimeSpan.FromSeconds(10));

        // Stop the processing
        await processor.StopProcessingAsync();
    }
```

4. Now add the following event and error handler methods to the class.

C#

```
    static async Task ProcessEventHandler(ProcessEventArgs eventArgs)
    {
        // Write the body of the event to the console window
        Console.WriteLine("\tReceived event: {0}",
Encoding.UTF8.GetString(eventArgs.Data.Body.ToArray()));

        // Update checkpoint in the blob storage so that the app receives
only new events the next time it's run
        await eventArgs.UpdateCheckpointAsync(eventArgs.CancellationToken);
    }

    static Task ProcessErrorHandler(ProcessErrorEventArgs eventArgs)
    {
        // Write details about the error to the console window
        Console.WriteLine($"\tPartition '{ eventArgs.PartitionId}': an
unhandled exception was encountered. This was not expected to happen.");
        Console.WriteLine(eventArgs.Exception.Message);
```

```
            return Task.CompletedTask;
        }
```

5. Build the project. Ensure that there are no errors.

> **Note**
>
> For the complete source code with more informational comments, see this file on the GitHub.

6. Run the receiver application.

An example of a Message received from Microsoft Purview

JSON


```
{
    "version":
        {"version":"1.0.0",
         "versionParts":[1]
        },
         "msgCompressionKind":"NONE",
         "msgSplitIdx":1,
         "msgSplitCount":1,
         "msgSourceIP":"10.244.155.5",
         "msgCreatedBy":
         "",
         "msgCreationTime":1618588940869,
         "message":{
            "type":"ENTITY_NOTIFICATION_V2",
            "entity":{
                "typeName":"azure_sql_table",
                    "attributes":{
                        "owner":"admin",
                        "createTime":0,

"qualifiedName":"mssql://nayenamakafka.eventhub.sql.net/salespool/dbo/SalesOrderTa
ble",
                        "name":"SalesOrderTable",
                        "description":"Sales Order Table"
                        },
                        "guid":"ead5abc7-00a4-4d81-8432-d5f6f6f60000",
                        "status":"ACTIVE",
                        "displayText":"SalesOrderTable"
                    },
                    "operationType":"ENTITY_UPDATE",
                    "eventTime":1618588940567
```

```
                }
}
```

Next steps

Check out more examples in GitHub.

- Event Hubs samples in GitHub
- Event processor samples in GitHub
- An introduction to Atlas notifications

Disaster recovery for Data Map (manual)

This article provides backup and recovery guidance for data governance admins who use Microsoft Purview Data Map in a production deployment. It covers manual business continuity and disaster recovery (BCDR) methods. Before you begin, review the prerequisites described first in this article. For detailed BCDR steps for Unified Catalog, see Disaster recovery for Unified Catalog (manual).

Prerequisites

Step 1. Create an Enterprise tier account

Create a Purview Data Governance account at the **Enterprise** tier. This serves as the **primary account** with full governance capabilities.

Step 2. Request Classic Purview provisioning in a secondary region

Create a support ticket through the Customer Support channel requesting enablement of Classic Purview account provisioning in a **secondary (paired) region** for BCDR purposes. This region shouldn't be the same as the primary region.

Step 3. Create a Classic Purview account

After Classic provisioning is enabled in the secondary region, create a **Classic Purview account**. This account serves as the **secondary (backup) account** with limited capabilities.

> ⓘ **Note**
>
> The first three steps are common to both Data Map and Unified Catalog BCDR.

Step 4. Make note of MSI identities

Capture the primary account's managed identity object ID (Identity A) and the secondary account's managed identity object ID (Identity B).

Step 5. Confirm Key Vault DR readiness

Make sure that either the secondary-region key vault holds the same secret names, or the key vault is geo-resilient with the paired secondary region.

Step 6. Confirm data source DR readiness

Make sure that each data source that's registered and scanned in Microsoft Purview has source-side DR (geo-replication/RA-GRS, SQL failover group, and so on) that provides a reachable copy in or near the secondary region.

Set up the secondary account and networking

Set up the secondary account to mirror the primary account, and then configure networking:

1. Recreate the collection hierarchy in the secondary account to match the primary account (same names).
2. Assign role-based access control (RBAC) roles on the secondary account collections (Collection Admin, Data Source Admin, Data Curator, Data Reader) to the same principals as the primary account.
3. Create the integration runtimes (IRs) in the secondary account (Managed virtual network (VNet) IR, Azure IR, or both) matching the primary account. Make sure the integration runtimes are present and **Running**.
4. If a self-hosted integration runtime (SHIR) is used, install a new SHIR node and register it against the secondary account by using the secondary account's authentication key.
5. Recreate the managed virtual network and managed private endpoints in the secondary region.
6. Approve each managed private endpoint on the target resource side.
7. Create private endpoints (account/portal/ingestion) and private DNS records for the secondary account.

Configure identity and access on data sources

Grant the secondary account's identity the same access on each data source that the primary account's identity has:

1. For each data source, grant Identity B the same data-plane role that Identity A has (for example, Storage Blob Data Reader, SQL `db_datareader`, and so on).

2. Open the data source firewalls and network rules for the secondary account's integration runtime or account (IP ranges, service endpoints, private endpoints, **Allow trusted Microsoft services**). This access is required for the secondary account to reach the source over the network.
3. Register the key vault connection in the secondary account and grant Identity B get/list secret permission on the key vault. Use a DR-enabled or secondary-region key vault that holds the same secret names.

Mirror scan configuration to the secondary account

Replicate the primary account's scan configuration in the secondary account:

1. Register the same data sources in the secondary account under the matching collections.
2. Create the same credentials in the secondary account (referencing the secondary account key vault connection and the same secret names).
3. Recreate custom classification rules in the secondary account (same names).
4. Recreate scan rulesets in the secondary account (system and custom) matching the primary account.
5. Create the same scans in the secondary account with matching scope selection, ruleset, and integration runtime binding.
6. Create the same triggers and schedules in the secondary account.

Validate regularly

In the steady state, check the following for your primary and secondary accounts:

1. For manually triggered or scheduled scans in the secondary account, make sure each scan reaches the **Succeeded** state (not **Failed** or **Completed with error**).
2. Compare discovered asset counts between the primary and secondary accounts after both complete a scan of the same source. Counts and asset qualified names should match within one scan interval of each other.
3. Spot-check classifications applied by rulesets on the same assets in the primary and secondary accounts. The same classifications should be present in the secondary account.
4. Verify schema parity on a representative tabular asset. Column names and types should match.

Fail over during a disaster recovery event

When a disaster recovery event occurs, create a support ticket requesting the following actions:

1. Promote the secondary (Classic) account to **Enterprise** tier.
2. Demote the primary (Enterprise) account to **Secondary**.

Share the tenant ID, account name, and region for both the primary and secondary accounts in the support ticket.

The failover completes, and the previously secondary account becomes the new primary with full Enterprise capabilities.

Resolve common errors

Expand table

Symptom	Most likely cause	Fix
Secondary account scan fails at authentication	Identity B isn't granted on the data source	Grant Identity B the data-plane role.
Secondary account scan fails with a network or timeout error	Data source firewall isn't opened for the secondary account, or the managed private endpoint isn't approved	Open the firewall or approve the managed private endpoint.
Credential **Test connection** fails in the secondary account	Key vault connection is missing, or Identity B lacks secret access	Register the key vault connection and grant secret get/list permission.
Assets present in the primary account but missing in the secondary account	Configuration drift — source or scan not mirrored	Mirror the configuration and run reconciliation.
Can't create the secondary account	Single instance per tenant, or Classic provisioning isn't enabled	Raise a support ticket.

Related content

- Disaster recovery for Unified Catalog (manual)
- Disaster recovery and migration best practices for Microsoft Purview data governance (classic)

ⓘ**Note:** The author created this article with assistance from AI. Learn more

Last updated on 07/20/2026

Private endpoints for data governance solutions

Article • 05/01/2025

> **Important**
>
> This article covers private endpoints for governance solutions in the Microsoft Purview portal. If you're using the classic governance portal, follow the documentation for private endpoints in the classic portal.

You can use private links to secure access to Microsoft Purview Unified Catalog and Data Map, and secure data traffic between Microsoft Purview and your private networks. Azure Private links and Azure Networking private endpoints are used to route traffic across Microsoft's infrastructure, instead of using the internet. To learn more about Azure Private Link in general, see: What is Azure Private Link?

A private endpoint is a single, directional technology that lets clients initiate connections to a given service, but doesn't allow the service to initiate a connection into the customer network. For multitenant services, this model provides link identifiers to prevent access to other customers' resources hosted within the same service. When you use private endpoints, only a limited set of other PaaS resources can be accessed from services using the integration.

You can deploy an *ingestion* private endpoint if you need to scan Azure IaaS and PaaS data sources inside Azure virtual networks and on-premises data sources through a private connection. This method ensures network isolation for your metadata flowing from the data sources to Microsoft Purview Data Map.

You can deploy a *platform* private endpoint to allow only client calls to the Microsoft Purview governance portal that originates from within the private network and connect to the Microsoft Purview governance portal using a private network connectivity.

> **Important**
>
> Private endpoints ensure that your organization's user traffic and resources within Azure, follow your organization's configured private link network path, and you can configure Microsoft Purview to deny all request from outside that network path.
>
> However, for external sources private endpoints do not manage all network traffic. You to configure traffic isolation from non-Azure based infrastructure, like on-premises

infrastructure, you need to configure ExpressRoute, or virtual private networks, and use integration runtimes to further secure your data sources.

Prerequisites

Before deploying private endpoints for your Microsoft Purview governance resources and the Microsoft Purview portal, ensure you meet the following prerequisites:

1. An Azure account with an active subscription. Create an account for free.
2. To configure private endpoints, you need to be a Microsoft Purview admin and have permissions in Azure to create and configure resources like virtual machines (VMs) and virtual networks (VNets).
3. Currently, Azure Data Factory, Azure Machine Learning, and Azure Synapse connections aren't supported with the platform private endpoint, and might not work after switching.

Deployment checklist

Following the instructions further in this guide, you can deploy these private endpoints for an existing Microsoft Purview account:

1. Choose an appropriate Azure virtual network and a subnet to deploy Microsoft Purview private endpoints. Select one of the following options:
    - Deploy a new virtual network in your Azure subscription.
    - Locate an existing Azure virtual network and a subnet in your Azure subscription.
2. Define an appropriate DNS name resolution method, so you can access Microsoft Purview account and scan data sources using a private network.
3. Create a platform private endpoint.
4. Enable ingestion private endpoints
5. Disable public access for Microsoft Purview.
6. Enable access to Microsoft Entra ID if your private network has network security group rules set to deny for all public internet traffic.
7. Deploy and register a Self-hosted integration runtime inside the same virtual network or a peered virtual network where Microsoft Purview account and ingestion private endpoints are deployed.
8. After completing this guide, adjust DNS configurations if needed.

Create a virtual network

> **Tip**
>
> These instructions will create a basic virtual network. For more information about virtual network options and features, see the virtual network documentation.

The next step is to create a virtual network and subnet. The number of IP address your subnet will need is made up of the number of capacities on your tenant plus three. For example, if you're creating a subnet for a tenant with seven capacities, you'll need 10 IP addresses.

Replace the sample parameters in the following table with your own to create a virtual network and subnet.

Expand table

Parameter	Value
`<resource-group-name>`	myResourceGroup
`<virtual-network-name>`	myVirtualNetwork
`<region-name>`	Central US
`<address-space>`	10.0.0.0/16
`<subnet-name>`	mySubnet
`<subnet-address-range>`	10.0.0.0/24

1. Open the Azure portal.
2. Select **Create a resource** > **Networking** > **Virtual network** or search for **Virtual network** in the search box.
3. In **Create virtual network**, enter or select the following information in the **Basics** tab:

Expand table

Settings	Value
Project details	
Subscription	Select your Azure Subscription
Resource group	Select **Create new**, enter `<resource-group-name>`, then select **OK**, or select an existing `<resource-group-name>` based on parameters.

Settings	Value
Instance details	
Name	Enter `<virtual-network-name>`

4. Select the **IP addresses** tab and enter the subnet address range.

5. Select **Review + create** > **Create**.

Define a DNS name resolution method

Follow this article to select and deploy a DNS name resolution method that bets fits your organization's needs: Configure DNS Name Resolution for private endpoints.

Create a platform private endpoint

The next step is to create the platform private endpoint for Microsoft Purview.

1. Open the Azure portal.

2. Select **Create a resource** > **Networking** > **Private Link**.

3. In **Private Link Center - Overview**, under the option to **Build a private connection to a service**, select **Create private endpoint**.

4. In the **Create a private endpoint - Basics** tab, enter or select the following information:

Expand table

Settings	Value
Project details	
Subscription	Select your Azure Subscription
Resource group	Select **myResourceGroup**. You created this in the previous section
Instance details	
Name	Enter *myPrivateEndpoint*. If this name is taken, create a unique name.
Network Interface Name	Filled automatically by the instance name.
Region	Selected automatically based on your resource group.

5. Once that information is complete, select **Next: Resource** and in the **Create a private endpoint - Resource** page, enter or select the following information:

Expand table

Settings	Value
Connection method	Select connect to an Azure resource in my directory
Subscription	Select your subscription
Resource type	Select **Microsoft.Purview/accounts**
Resource	Select your Microsoft Purview resource
Target subresource	Platform

6. Once that information is properly input, select **Next: Virtual Network** and enter or select the following information:

Expand table

Settings	Value
NETWORKING	
Virtual network	Select the virtual network you created earlier.
Subnet	Select the subnet you created earlier.
PRIVATE IP configuration	Select **Dynamically allocate IP address**.

7. Select **Next: DNS** and enter the following information:

Expand table

Settings	Value
Integrate with private DNS zone	Select **Yes**
Subscription	Select your subscription where your DNS zone is configured.
Resource group	Select the resource group where your DNS zone is configured.

8. Select **Next: Tags** and on the tags page you can optionally add any tags your organization is using in Azure.

9. Select **Next: Review + create** which displays the **Review + create** page where Azure validates your configuration. When you see the **Validation passed** message, select **Create**.

Enable *ingestion* private endpoint

1. Go to the Azure portal, search for and then select your Microsoft Purview account.
2. From your Microsoft Purview account, under **Settings** select **Networking**, and then select **Ingestion private endpoint connections**.
3. Under Ingestion private endpoint connections, select **+ New** to create a new ingestion private endpoint.
4. Fill in the basic information, selecting your existing virtual network and a subnet details. Optionally, select **Private DNS integration** to use Azure Private DNS Zones. Select correct Azure Private DNS Zones from each list.

   > **Note**
   >
   > You can also use your existing Azure Private DNS zones or create DNS records in your DNS Servers manually. For more information, see Configure DNS Name Resolution for private endpoints

5. Select **Create** to finish the setup.

Disable public access for Microsoft Purview

To cut off access to the Microsoft Purview account completely from the public internet, follow these steps. This setting applies to both private endpoint and ingestion private endpoint connections.

1. From the Azure portal, go to the Microsoft Purview account, and under **Settings**, select **Networking**.
2. Go to the **Firewall** tab, and ensure that the toggle is set to **Disable from all networks**.

Enable access to Microsoft Entra ID

> **Note**

If your VM, VPN gateway, or virtual network peering gateway has public internet access, it can access the Microsoft Purview governance portal and the Microsoft Purview account enabled with private endpoints. For this reason, you don't have to follow the rest of the instructions. If your private network has network security group rules set to deny all public internet traffic, you'll need to add some rules to enable Microsoft Entra ID access. Follow the instructions to do so.

These instructions are provided for accessing Microsoft Purview securely from an Azure VM. Similar steps must be followed if you're using VPN or other virtual network peering gateways.

1. Go to your VM in the Azure portal, and under **Settings**, select **Networking**. Then select **Outbound port rules** > **Add outbound port rule**.

2. On the **Add outbound security rule** pane:
   a. Under **Destination**, select **Service Tag**.
   b. Under **Destination service tag**, select **AzureActiveDirectory**.
   c. Under **Destination port ranges**, select *.
   d. Under **Action**, select **Allow**.
   e. Under **Priority**, the value should be higher than the rule that denied all internet traffic.

   Create the rule.

3. Follow the same steps to create another rule to allow the **AzureResourceManager** service tag. If you need to access the Azure portal, you can also add a rule for the **AzurePortal** service tag.

4. Connect to the VM and open the browser. Go to the browser console by selecting Ctrl+Shift+J, and switch to the network tab to monitor network requests. Enter web.purview.azure.com in the URL box, and try to sign in by using your Microsoft Entra credentials. Sign-in will probably fail, and on the **Network** tab on the console, you can see Microsoft Entra ID trying to access aadcdn.msauth.net but getting blocked.

5. In this case, open a command prompt on the VM, ping aadcdn.msauth.net, get its IP, and then add an outbound port rule for the IP in the VM's network security rules. Set the **Destination** to **IP Addresses** and set **Destination IP addresses** to the aadcdn IP. Because of Azure Load Balancer and Azure Traffic Manager, the Microsoft Entra Content Delivery Network IP might be dynamic. After you get its IP, it's better to add it into the VM's host file to force the browser to visit that IP to get the Microsoft Entra Content Delivery Network.

6. After the new rule is created, go back to the VM and try to sign in by using your Microsoft Entra credentials again. If sign-in succeeds, then the Microsoft Purview governance portal is ready to use. But in some cases, Microsoft Entra ID redirects to other domains to sign in

based on a customer's account type. For example, for a live.com account, Microsoft Entra ID redirects to live.com to sign in, and then those requests are blocked again. For Microsoft employee accounts, Microsoft Entra ID accesses msft.sts.microsoft.com for sign-in information.

Check the networking requests on the browser **Networking** tab to see which domain's requests are getting blocked, redo the previous step to get its IP, and add outbound port rules in the network security group to allow requests for that IP. If possible, add the URL and IP to the VM's host file to fix the DNS resolution. If you know the exact sign-in domain's IP ranges, you can also directly add them into networking rules.

7. Now your Microsoft Entra sign-in should be successful. The Microsoft Purview governance portal will load successfully, but listing all the Microsoft Purview accounts won't work because it can only access a specific Microsoft Purview account. Enter `web.purview.azure.com/resource/{PurviewAccountName}` to directly visit the Microsoft Purview account that you successfully set up a private endpoint for.

Deploy self-hosted integration runtime (IR) and scan your data sources

Once you deploy ingestion private endpoints for your Microsoft Purview, you need to setup and register at least one self-hosted integration runtime (IR):

- All on-premises source types like Microsoft SQL Server, Oracle, SAP, and others are currently supported only via self-hosted IR-based scans. The self-hosted IR must run within your private network and then be peered with your virtual network in Azure.
- For all Azure source types like Azure Blob Storage and Azure SQL Database, you must explicitly choose to run the scan by using a self-hosted integration runtime that is deployed in the same virtual network or a peered virtual network where Microsoft Purview account and ingestion private endpoints are deployed.

Follow the steps in Create and manage a self-hosted integration runtime to set up a self-hosted IR. Then set up your scan on the Azure source by choosing that self-hosted IR in the **Connect via integration runtime** dropdown list to ensure network isolation.

> **Important**
>
> Make sure you download and install the latest version of self-hosted integration runtime from Microsoft download center.

Test your private connection

To test your new private endpoints, you can create a virtual machine within your private virtual network and access your platform private endpoint to ensure it's working.

1. Create a virtual machine (VM).
2. Connect to a VM using Remote Desktop (RDP).
3. Access Microsoft Purview privately from the virtual machine.

Create a virtual machine (VM)

The next step is to create a VM.

1. On the upper-left side of the screen in your Azure portal, select **Create a resource** > **Compute** > **Virtual Machine**.

2. In the **Basics** tab, enter or select the following information:

Expand table

Settings	Value
Project details	
Subscription	Select your Azure Subscription
Resource group	Select **myResourceGroup** which you created in the previous section
Instance details	
VM name	Enter **myVM**
Region	Select **West US**
Availability options	Leave the default **No infrastructure redundancy required**
Image	Select **Windows 10 Pro**
Size	Leave the default **Standard DS1 v2**
ADMINISTRATOR ACCOUNT	
Username	Enter a username of your choosing
Password	Enter a password of your choosing. The password must be at least 12 characters long and meet the defined complexity requirements

Settings	Value
Confirm password	Reenter password
INBOUND PORT RULES	
Public inbound ports	Leave the default **None**
LICENSING	
I have an eligible Windows 10/11 license	Check the box

3. Select **Next: Disks**.
4. In the **Disks** tab, leave the defaults and select **Next: Networking**.
5. In the **Networking** tab, select the following information:

Expand table

Settings	Value
Virtual network	Leave the default **MyVirtualNetwork**
Address space	Leave the default **10.0.0.0/24**
Subnet	Leave the default **mySubnet (10.0.0.0/24)**
Public IP	Leave the default **(new) myVM-ip**
Public inbound ports	Select **Allow selected**
Select inbound ports	Select **RDP**

6. Select **Review + create**. You're taken to the **Review + create** page where Azure validates your configuration.
7. When you see the **Validation passed** message, select **Create**.

Connect to a VM using Remote Desktop (RDP)

After you create your VM, called **myVM**, connect to it from the internet using the following steps:

1. In the portal's search bar, enter *myVM*.
2. Select the **Connect** button, and choose **RDP** from the dropdown menu.

3. Enter an IP address, then select **Download RDP File**. Azure creates a Remote Desktop Protocol (.rdp) file and downloads it to your computer.
4. Open the *.rdp* file to start Remote Desktop Connection, then select **Connect**.
5. Enter the username and password you specified when creating the VM in the previous step.
6. Select **OK**.
7. You might receive a certificate warning during the sign-in process. If you receive a certificate warning, select **Yes** or **Continue**.

Access Microsoft Purview privately from the VM

The next step is to access Microsoft Purview privately, from the virtual machine you created in the previous step, using the following steps:

1. In the Remote Desktop of myVM, open PowerShell.
2. Enter `nslookup <tenant-object-id>-api.purview-service.microsoft.com`.
3. You receive a response similar to the following message:

```
Server:  UnKnown
Address:  168.63.129.16

Non-authoritative answer:
Name:    <tenantid>-api.purview-service.microsoft.com
Address:  10.5.0.4
```

4. Open the browser and go to https://purview.microsoft.com to access Microsoft Purview privately.

Completion of private endpoint configuration

Once you've followed the steps in the previous sections and the private link is successfully configured, your organization implements private links based on the following configuration selections, whether the selection is set upon initial configuration or later changed.

If Azure Private Link is properly configured and **Block public Internet access** is **enabled**:

- Microsoft Purview is only accessible for your organization from private endpoints, and isn't accessible from the public Internet.
- Traffic from the virtual network targeting endpoints and scenarios that support private links are transported through the private link.
- Traffic from the virtual network targeting endpoints and scenarios that don't* support private links will be blocked by the service, and won't work.
- There could be scenarios that don't support private links, which therefore will be blocked at the service when **Block public Internet access** is enabled.

If Azure Private Link is properly configured and **Block public Internet access** is **disabled**:

- Traffic from the public Internet will be allowed by Microsoft Purview services
- Traffic from the virtual network targeting endpoints and scenarios that support private links are transported through the private link.
- Traffic from the virtual network targeting endpoints and scenarios that don't* support private links are transported through the public Internet, and will be allowed by Microsoft Purview services.
- If the virtual network is configured to block public Internet access, scenarios that don't support private links will be blocked by the virtual network, and won't work.'

Migrate governance private endpoints from classic portal to Microsoft Purview portal

Article • 05/01/2025

> **Important**
>
> This article covers migrating private endpoints for governance solutions from the classic governance portal to the Microsoft Purview portal. If you're building your private endpoint connections for the first time, follow the full private endpoints article instead.

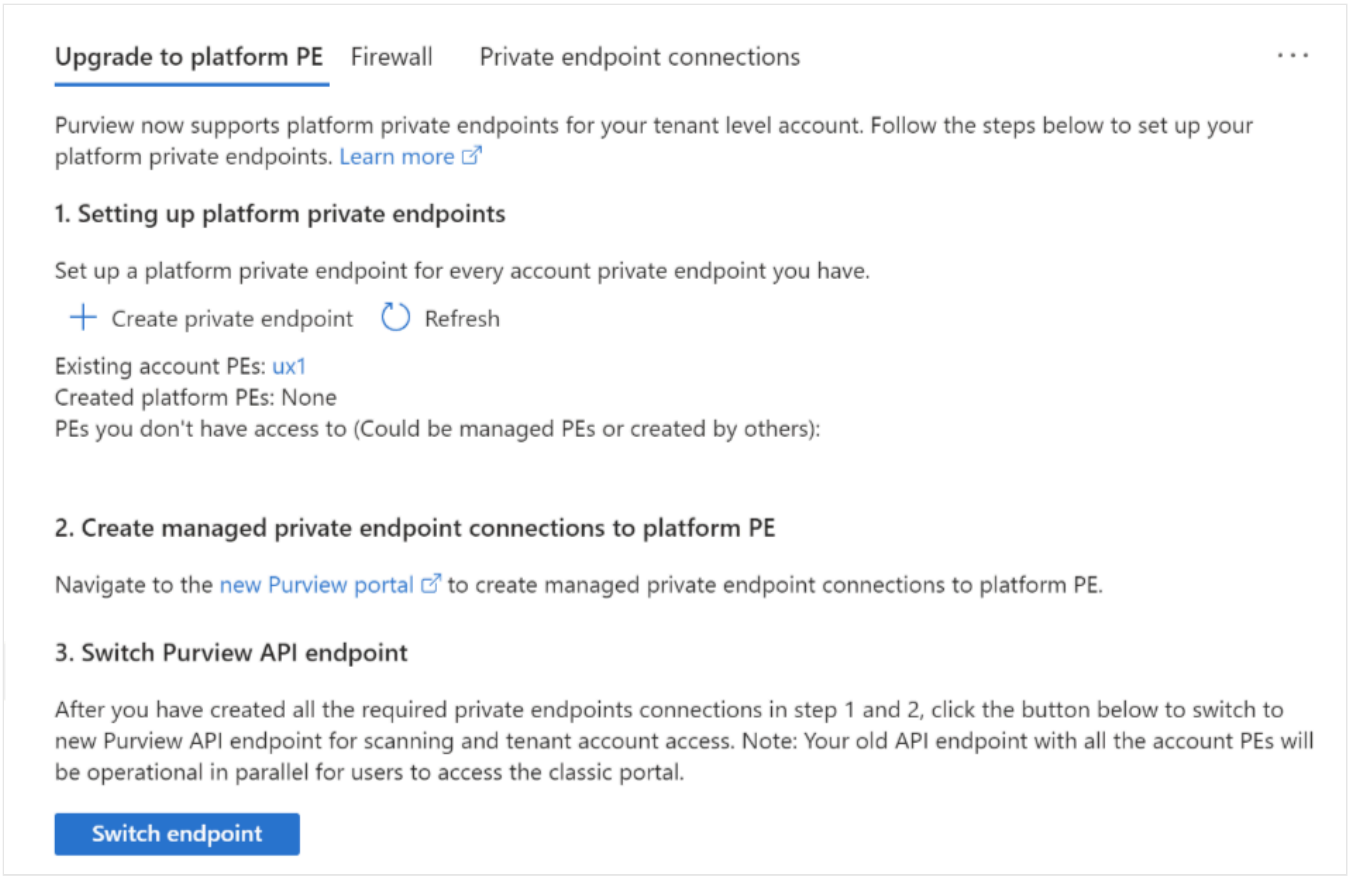
> **Note**
>
> Currently, Azure Data Factory, Azure Machine Learning, and Azure Synapse connections aren't supported with the platform private endpoint, and might not work after switching.

Microsoft Purview now supports platform private endpoints for your organization-level account. You can follow these steps to migrate your existing Microsoft Purview account private endpoints to the new platform private endpoints.

Prerequisites

If you're still using version 1 of the managed virtual network (which you can check), to migrate to the new private endpoints you need to upgrade to the latest version of the virtual network.

Migrate your existing private endpoints



1. In the Azure portal, search for and open your Microsoft Purview account.
2. Select the **Networking** page.
3. If it's not already open, select the **Upgrade to platform PE** tab.
4. Set up your platform private endpoints
5. Create managed private endpoint connections to platform private endpoints
6. Switch your Microsoft Purview API endpoint

Set up your platform private endpoints

For every account private endpoint you'll need to set up a platform private endpoint. All the account private endpoints you have are listed under the step. For each existing endpoint:

1. In the Azure portal, search for and open your Microsoft Purview account.
2. Select the **Networking** page.
3. If it's not already open, select the **Upgrade to platform PE** tab.
4. Select the **+ Create platform private endpoint** button.
5. In the **Create a private endpoint - Basics** tab, enter or select the following information:

Expand table

Settings	Value
Project details	
Subscription	Select your Azure Subscription
Resource group	Select your resource group.
Instance details	
Name	Create a unique name.
Network Interface Name	Filled automatically by the instance name.
Region	Selected automatically based on your resource group.

6. Once that information is complete, select **Next: Resource** and in the **Create a private endpoint - Resource** page, enter or select the following information:

Expand table

Settings	Value
Connection method	Select connect to an Azure resource in my directory
Subscription	Select your subscription
Resource type	Select **Microsoft.Purview/accounts**
Resource	Select your Microsoft Purview resource
Target subresource	Platform

7. Once that information is properly input, select **Next: Virtual Network** and enter or select the following information:

Expand table

Settings	Value
NETWORKING	
Virtual network	Select your existing virtual network.
Subnet	Select your existing subnet
PRIVATE IP configuration	Select **Dynamically allocate IP address**.

8. Select **Next: DNS** and enter the following information:

Expand table

Settings	Value
Integrate with private DNS zone	Select **Yes**
Subscription	Select your subscription where your DNS zone is configured.
Resource group	Select the resource group where your DNS zone is configured.

9. Select **Next: Tags** and on the tags page you can optionally add any tags your organization is using in Azure.

10. Select **Next: Review + create** which displays the **Review + create** page where Azure validates your configuration. When you see the **Validation passed** message, select **Create**.

11. If you configured firewall allowlist rules for your account endpoints (`{account-name}.purview.azure.com`), you need to update your firewall configuration for the new platform private endpoints: `{tenant id}-api.purview-service.microsoft.com`, `api.purview-service.microsoft.com`. The use of the endpoint `api.purview-service.microsoft.com` will require additional configuration and checks. This will depend on your network setup, and we recommend using `{tenant id}-api.purview-service.microsoft.com` for ease of setup.

> **Note**
>
> `api.purview-service.microsoft.com` won't be automatically configured in the DNS to point to `{tenant id}-api.purview-service.microsoft.com` when the private endpoint is created; it will continue to point to the public IP of the purview instance unless overridden by subsequent customer network or DNS configuration changes.

Create managed private endpoint connections to platform private endpoints

1. In the Microsoft Purview portal open the Microsoft Purview Data Map.

2. Navigate to **Source management**, and select **Managed private endpoints**.

3. Select the pop-up link that says, 'Click here to provision platform managed PE for all your managed VNets'.

4. Microsoft Purview will create managed platform private endpoints for all your existing managed account endpoints.

5. Select each approval link to approve the endpoints, or approve them directly by:
   a. In the Azure portal search for your Microsoft Purview account and open it.
   b. Select **Networking**
   c. Select **Private endpoints**
   d. Approve the new list of private endpoints.

Switch your Microsoft Purview API endpoint

Now that you've created the required private endpoints, connected your resources, and recreated your scans, you can update your API endpoint. On the networking page, select the **Switch endpoint** button to switch to the new Microsoft Purview API endpoint for scanning and tenant account access.

The new API endpoint you'll be using is either:

`https://api.purview-service.microsoft.com/`

`https://{your-tenant-id}-api.purview-service.microsoft.com/`

> **Note**
>
> The old endpoint will be operational in parallel for users with access to the classic portal. Scans that were already running will continue to use the existing endpoint until they finish. New scans will use the new platform endpoint.

Remove account endpoints

Once all your currently running scans have completed, remove your existing account endpoints to switch completely to your new platform private endpoints.

Use managed virtual network with your Microsoft Purview account

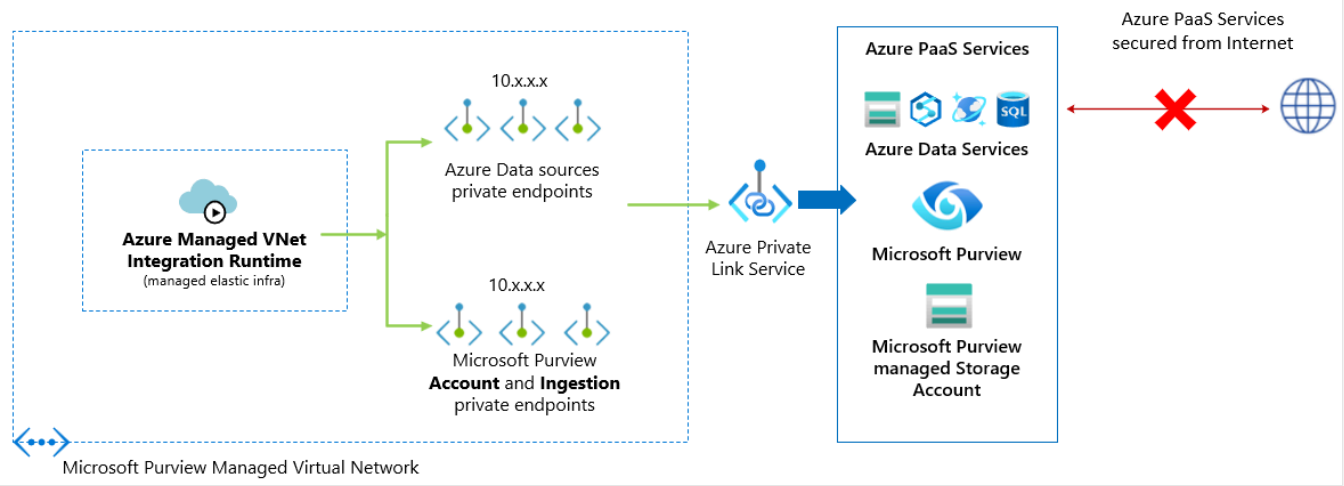
Article • 05/01/2025

This article describes managed virtual networks and managed private endpoints in Microsoft Purview.

> **Note**
>
> - Microsoft Purview added new version of Managed Virtual Network support mid November 2023. New managed private endpoints no longer support the old managed virtual network. Create new a managed virtual network, and all your newly created resources will be using the latest version. Learn about the difference here.
> - Microsoft Purview accounts that were deployed before December 15, 2023 include a managed storage account. Microsoft Purview accounts that are deployed after (or deployed using API version 2023-05-01-preview onwards) don't have manage storage accounts deployed to the Azure subscription. Instead, these accounts include an ingestion storage account that's deployed to Microsoft's internal Azure subscriptions.

Managed Virtual Network

A Managed Virtual Network (virtual network) is a virtual network that is deployed and managed by Microsoft Purview to empower scanning data sources inside a private network, without having to deploy and manage any self-hosted integration runtime virtual machines by the customer in Azure.



You can deploy Azure Managed Virtual Network Integration Runtimes (IR) within Microsoft Purview Managed Virtual Networks in any available Microsoft Purview region. From there, the managed virtual network IR can use private endpoints to securely connect to and scan the supported data sources.

Creating a Managed virtual network IR within Managed Virtual Network ensures that data integration process is isolated and secure.

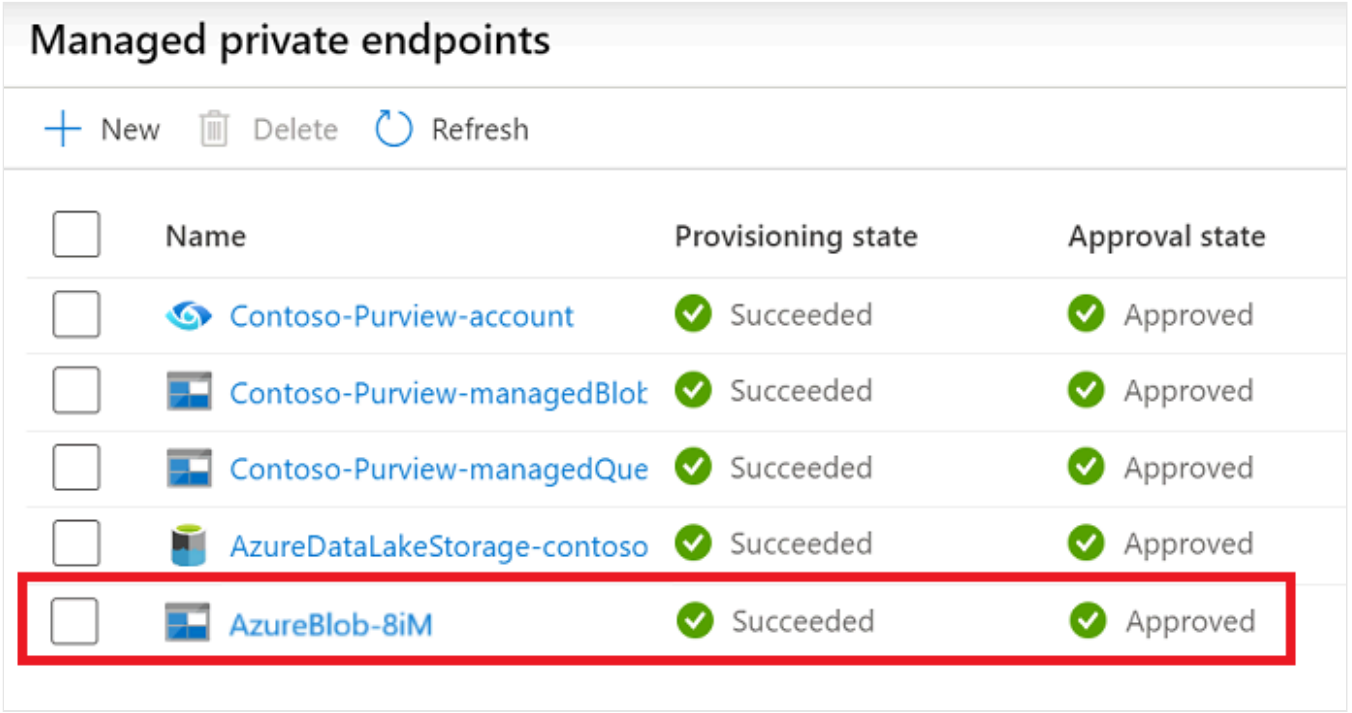
Benefits of using Managed Virtual Network:

- With a Managed Virtual Network, you can offload the burden of managing the Virtual Network to Microsoft Purview. You don't need to create and manage VNets or subnets for Azure Integration Runtime to use for scanning Azure data sources.
- It doesn't require deep Azure networking knowledge to do data integrations securely. Using a Managed Virtual Network is much simplified for data engineers.
- Managed Virtual Network along with Managed private endpoints protects against data exfiltration.

A managed virtual network is created for your Microsoft Purview account when you create a Managed Virtual Network Integration Runtime for the first time in your Microsoft Purview account. You can't view or manage the virtual networks outside of Microsoft Purview.

Managed private endpoints

Managed private endpoints are private endpoints created in the Microsoft Purview Managed Virtual Network establishing a private link to Microsoft Purview and Azure resources. Microsoft Purview manages these private endpoints on your behalf.



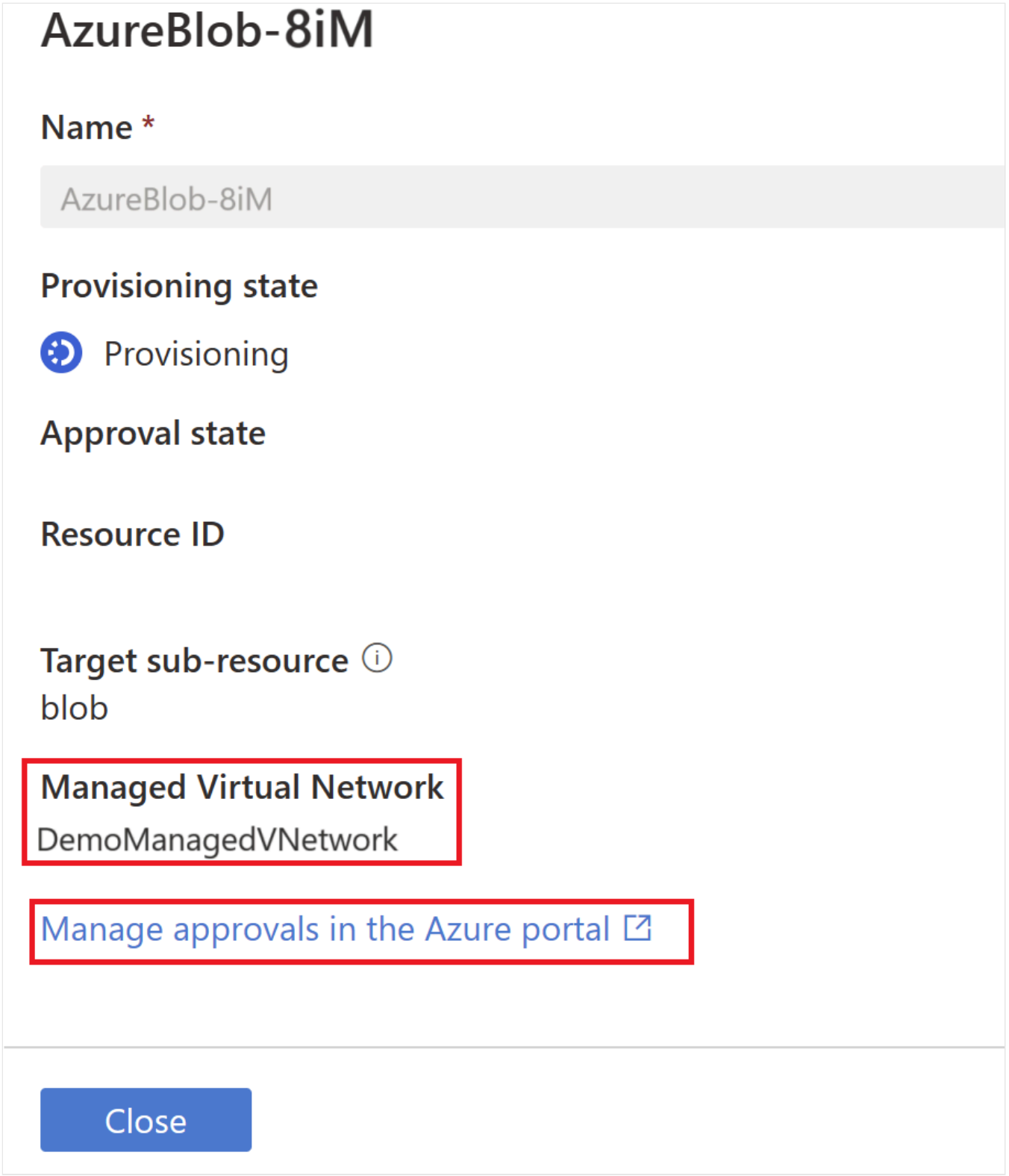
Microsoft Purview supports private links. Private link enables you to access Azure services like Azure Storage, Azure SQL Database, Azure Cosmos DB, Azure Synapse Analytics.

When you use a private link, traffic between your data sources and Managed Virtual Network traverses entirely over the Microsoft backbone network. Private Link protects against data exfiltration risks. You establish a private link to a resource by creating a private endpoint.

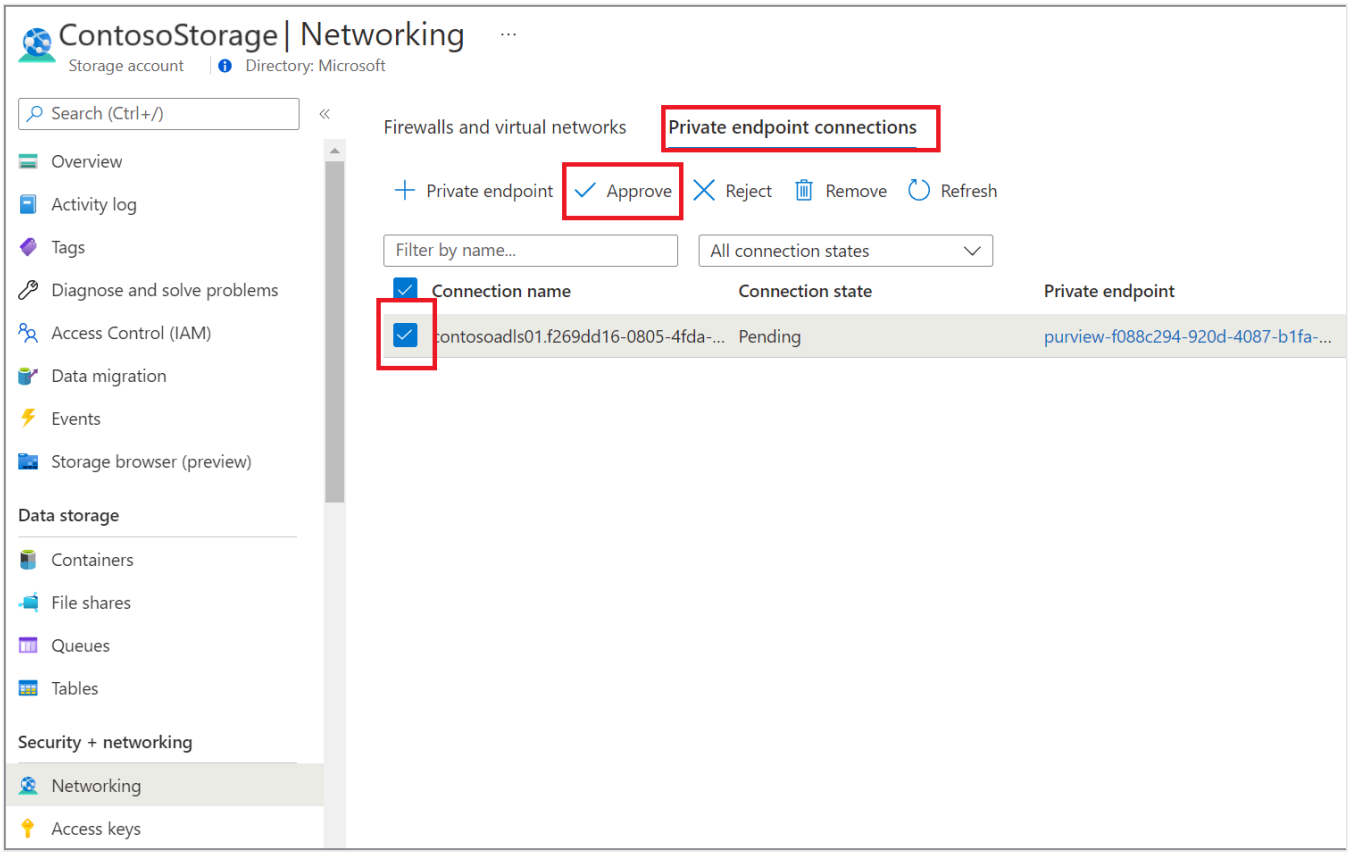
Private endpoint uses a private IP address in the Managed Virtual Network to effectively bring the service into it. Private endpoints are mapped to a specific resource in Azure and not the entire service. Customers can limit connectivity to a specific resource approved by their organization. Learn more about private links and private endpoints.

> ⚠ **Warning**
>
> If an Azure PaaS data store (Blob, Azure Data Lake Storage Gen2, Azure Synapse Analytics) has a private endpoint already created against it, and even if it allows access from all networks, Microsoft Purview would only be able to access it using a managed private endpoint. If a private endpoint does not already exist, you must create one in such scenarios.

A private endpoint connection is created in a "Pending" state when you create a managed private endpoint in Microsoft Purview. An approval workflow is initiated. The private link resource owner is responsible to approve or reject the connection.



If the owner approves the connection, the private link is established. Otherwise, the private link won't be established. In either case, the Managed private endpoint is updated with the status of the connection.



Only a Managed private endpoint in an approved state can send traffic to a given private link resource.

Supported regions and data sources

Supported regions: Managed Virtual Network is available in all the Microsoft Purview supported regions. Up to five managed virtual networks can be deployed across different regions in a single Microsoft Purview instance.

Supported data sources for scan: You can use the Managed Virtual Network Integration Runtime to scan many types of data source. Refer to the supported data sources.

Supported systems with managed private endpoint: The following services have native private endpoint support. They can be connected through private link from Microsoft Purview's Managed Virtual Network:

- Azure Blob Storage
- Azure Cosmos DB
- Azure Data Lake Storage Gen 2
- Azure Database for MySQL
- Azure Database for PostgreSQL
- Azure Databricks
- Azure Dedicated SQL pool (formerly SQL DW)
- Azure Files

- Azure Key Vault
- Azure SQL Database
- Azure SQL Managed Instance
- Azure Synapse Analytics
- Snowflake

Pricing details

When using Managed Virtual Network feature, you're charged by two parts:

- **Charge per scan run** (pay as you go): based on scan duration * used vCore hours * unit price per vCore hour. This is the common charge for scan running on any integration runtime types.
- **Charge on Managed VNet IR uptime** (always on): based on **Managed VNet IR life time * 1/8 vCore hour** * unit price per vCore hour. Managed virtual network IR life time means from the IR instance being successfully created until its deleted. The charge applies no matter you have scan run or not.

You can find details for "Automated Scanning, Ingestion & Classification" pricing from Microsoft Purview pricing page.

As an example, for a given month, you scan various data sources and the execution consumes X vCore hours of "Scanning Ingestion and Classification - Standard vCore" in total, and the managed virtual network runtime is provisioned to run for the whole month. The charge is: X vCore hours * $0.63 per vCore hour for scan runs + 1/8 * $0.63 per vCore hour * 730 hours for Managed virtual network IR uptime.

Deployment steps

Prerequisites

Before deploying a Managed virtual network and Managed Virtual Network Integration Runtime for a Microsoft Purview account, ensure you meet the following prerequisites:

1. From Microsoft Purview roles, you need Data Source Administrator permission on any collection in your Microsoft Purview account.
2. From Azure RBAC roles, you must be contributor on the Microsoft Purview account and data source to approve private links.

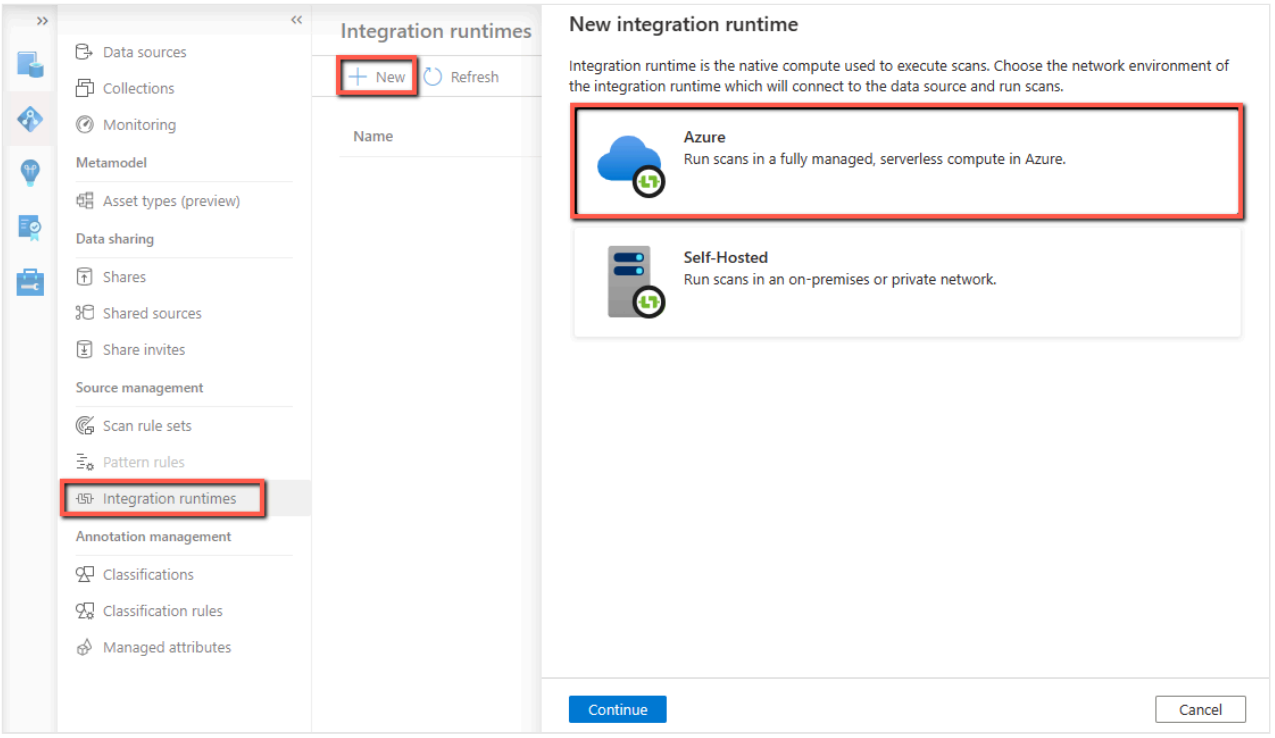
Create Managed Virtual Network Integration Runtime

1. Open the Microsoft Purview governance portal by:

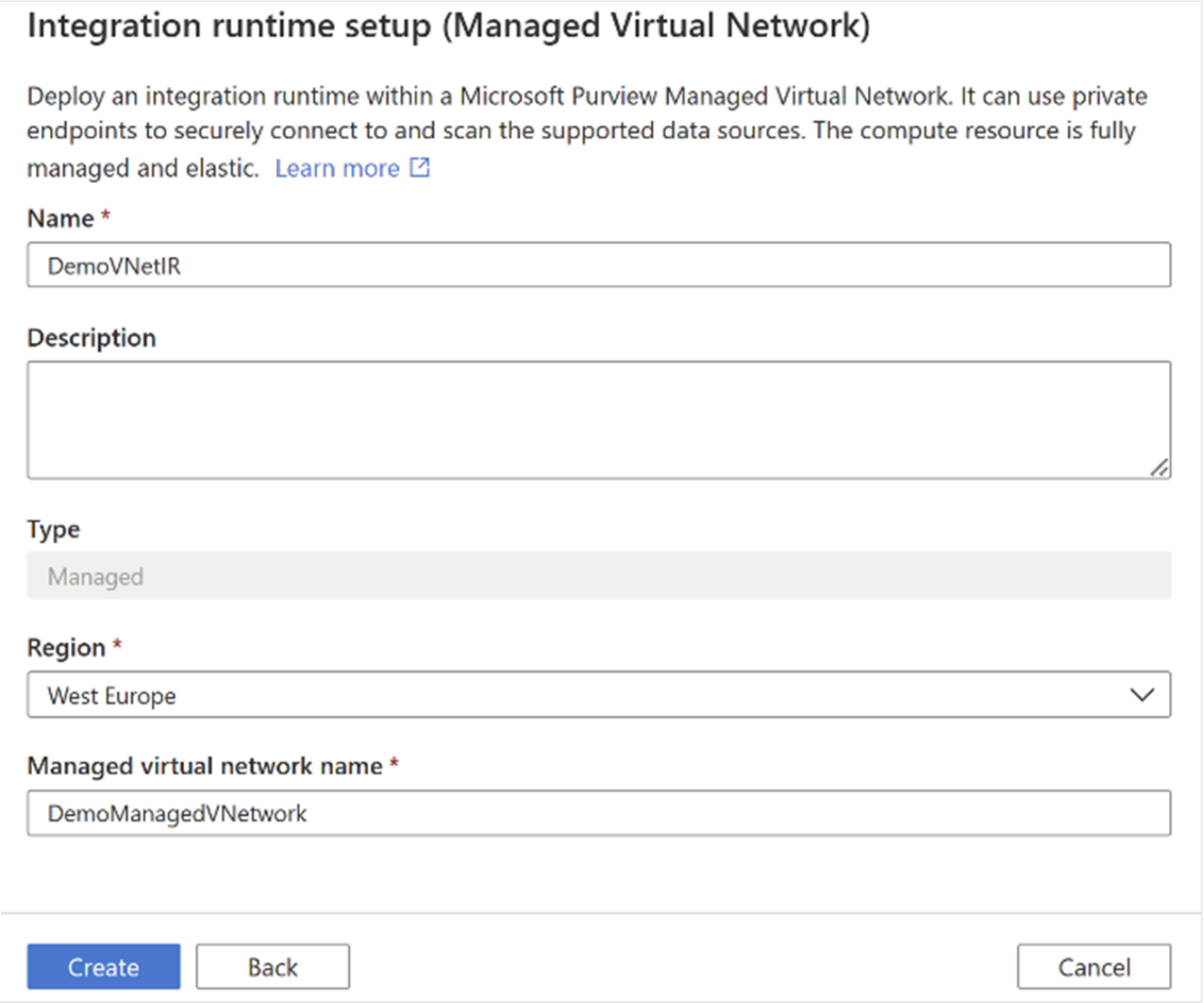
    - Browsing directly to https://web.purview.azure.com and selecting your Microsoft Purview account.
    - Opening the Azure portal, searching for and selecting the Microsoft Purview account. Selecting the **the Microsoft Purview governance portal** button.

2. Navigate to the **Data Map** -> **Integration runtimes**.

3. From **Integration runtimes** page, select + **New** icon, to create a new runtime. Select **Azure** and then select **Continue**.

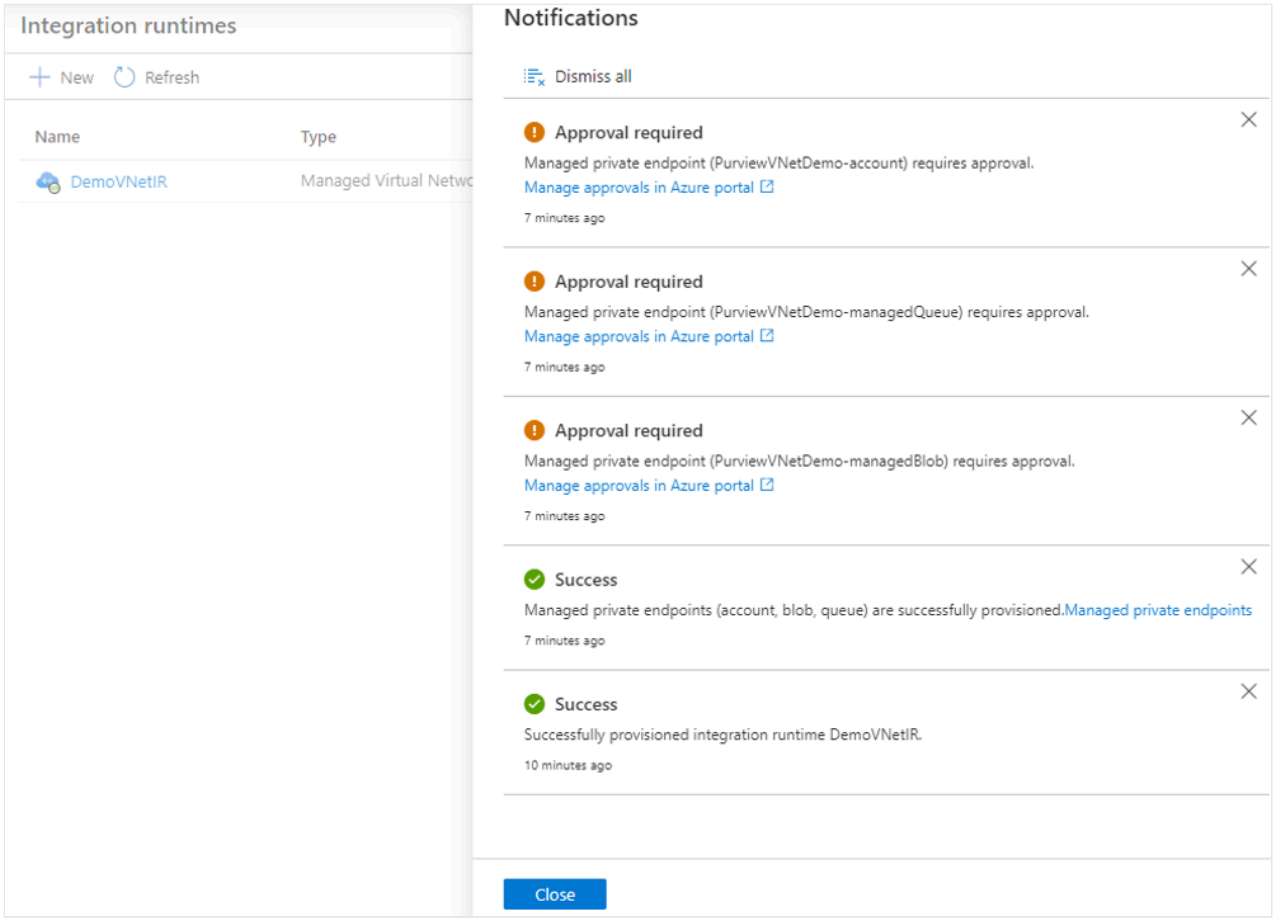


4. Provide a **name** for your Managed Virtual Network Integration Runtime, select a region, and give your managed virtual network a name.

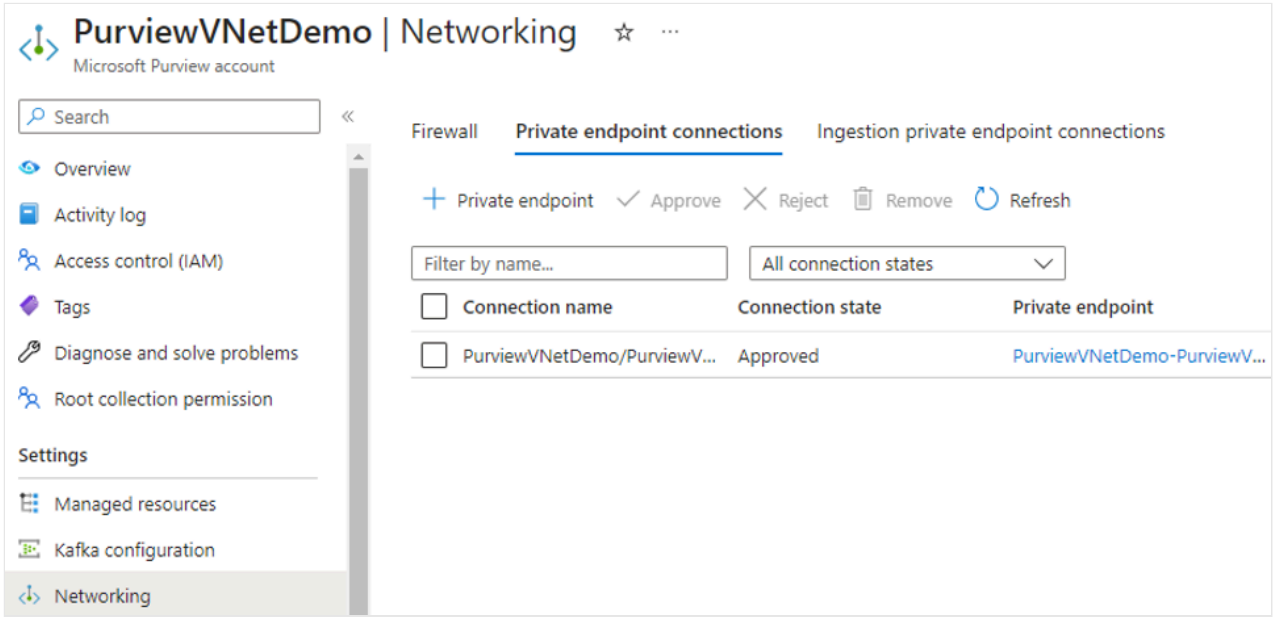


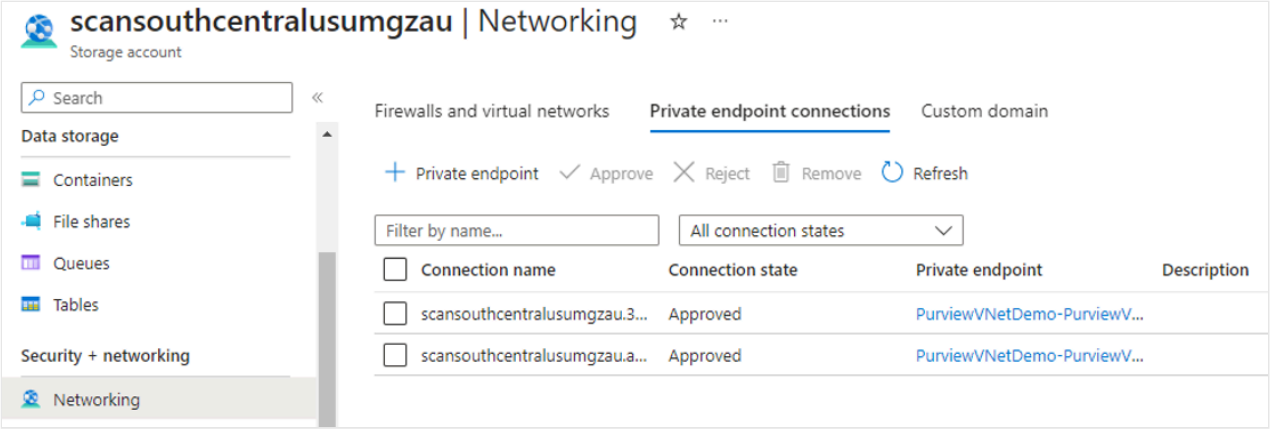
5. Select **Create**.

6. Deploying the Managed Virtual Network Integration Runtime triggers multiple workflows in the Microsoft Purview governance portal for creating managed private endpoints for Microsoft Purview and its managed Storage Account. Select each workflow to approve the private endpoint for the corresponding Azure resource.

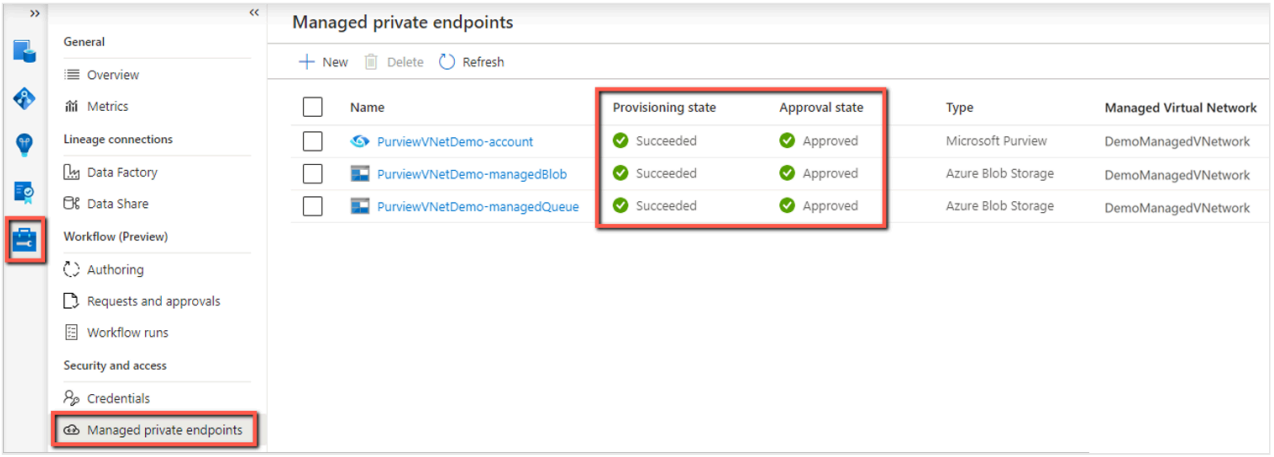


7. In Azure portal, from your Microsoft Purview account resource window, approve the managed private endpoint. From managed storage account page approve the managed private endpoints for blob and queue services:

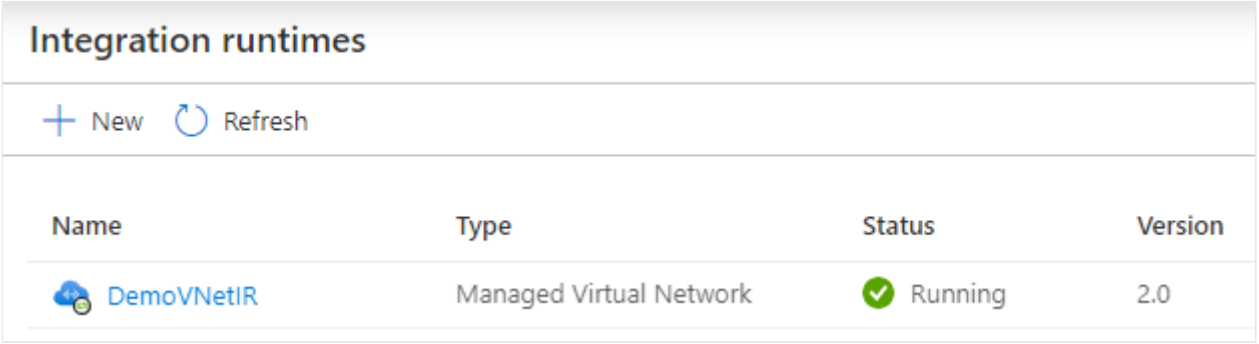




8. From Management, select Managed private endpoint to validate if all managed private endpoints are successfully deployed and approved.



9. Go to the Integration runtimes page, you'll see the IR status shown up as "Initializing" upon creation. Wait until it turns into "**Running**" state to use in scan. It usually takes several minutes.



> **Tip**
>
> You can create multiple managed virtual networks in different regions in your Microsoft Purview account to securely access resources across regions.

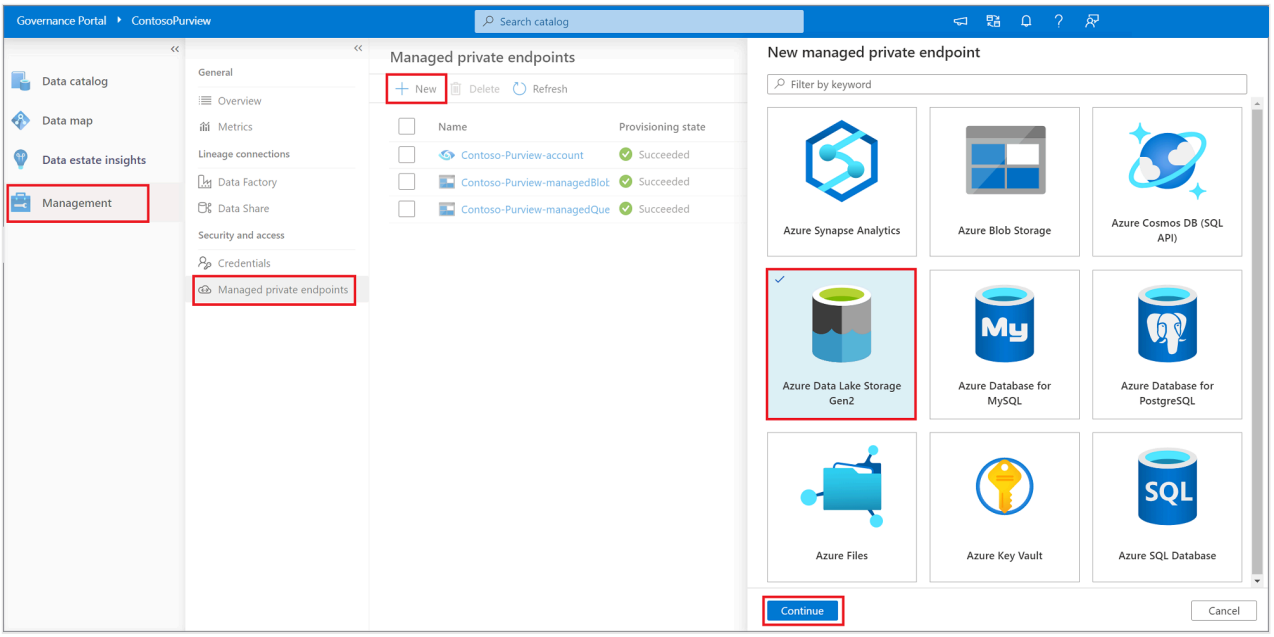
Create managed private endpoints for data sources

You can use managed private endpoints to connect your data sources to ensure data security during transmission.

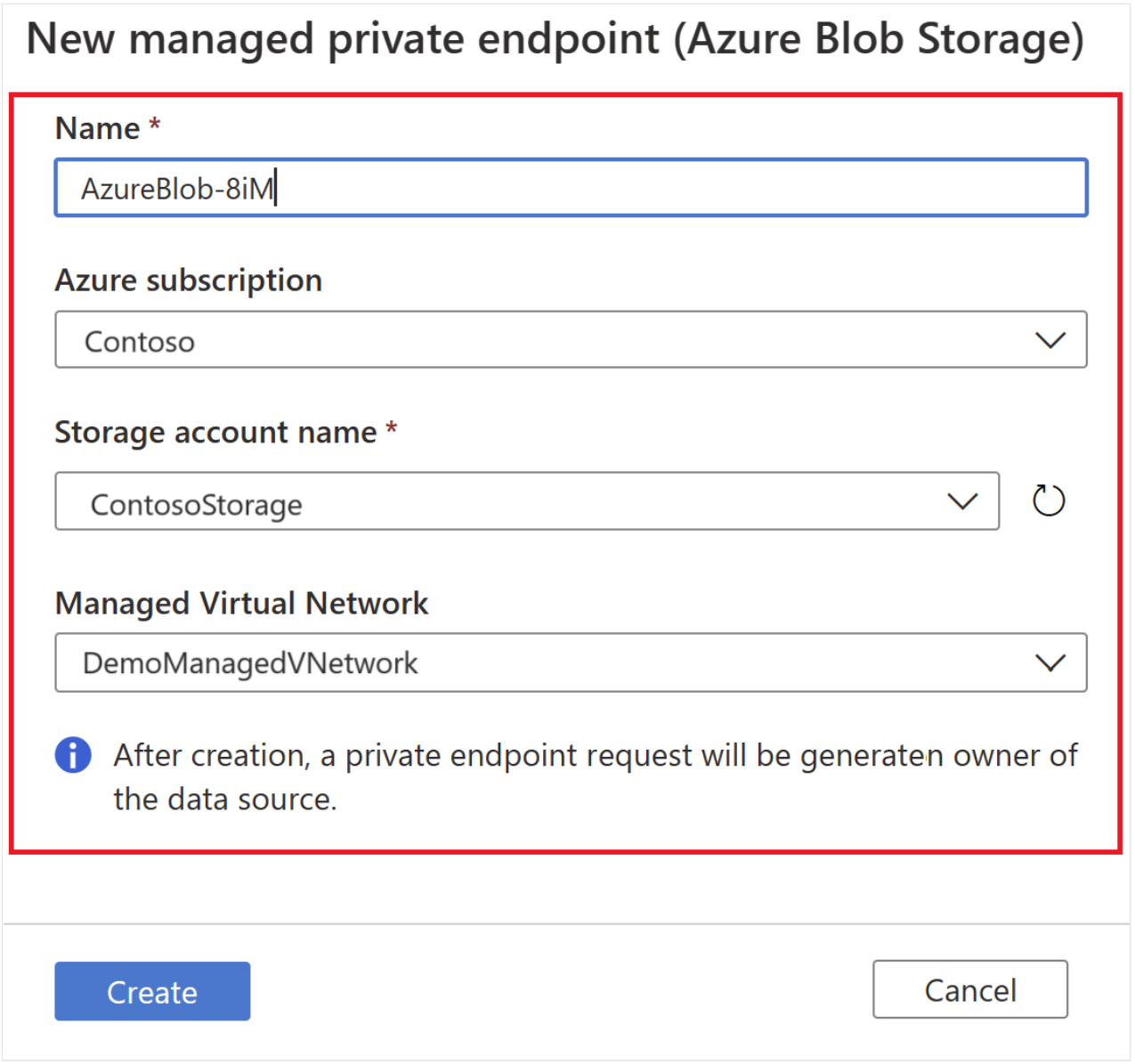
> **Tip**
>
> If your data source allows public access and you want to connect via public network, you can skip this step. Scan runs can be executed as long as the integration runtime can connect to your data source.

To deploy and approve a managed private endpoint for a data source, follow these steps selecting data source of your choice from the list:

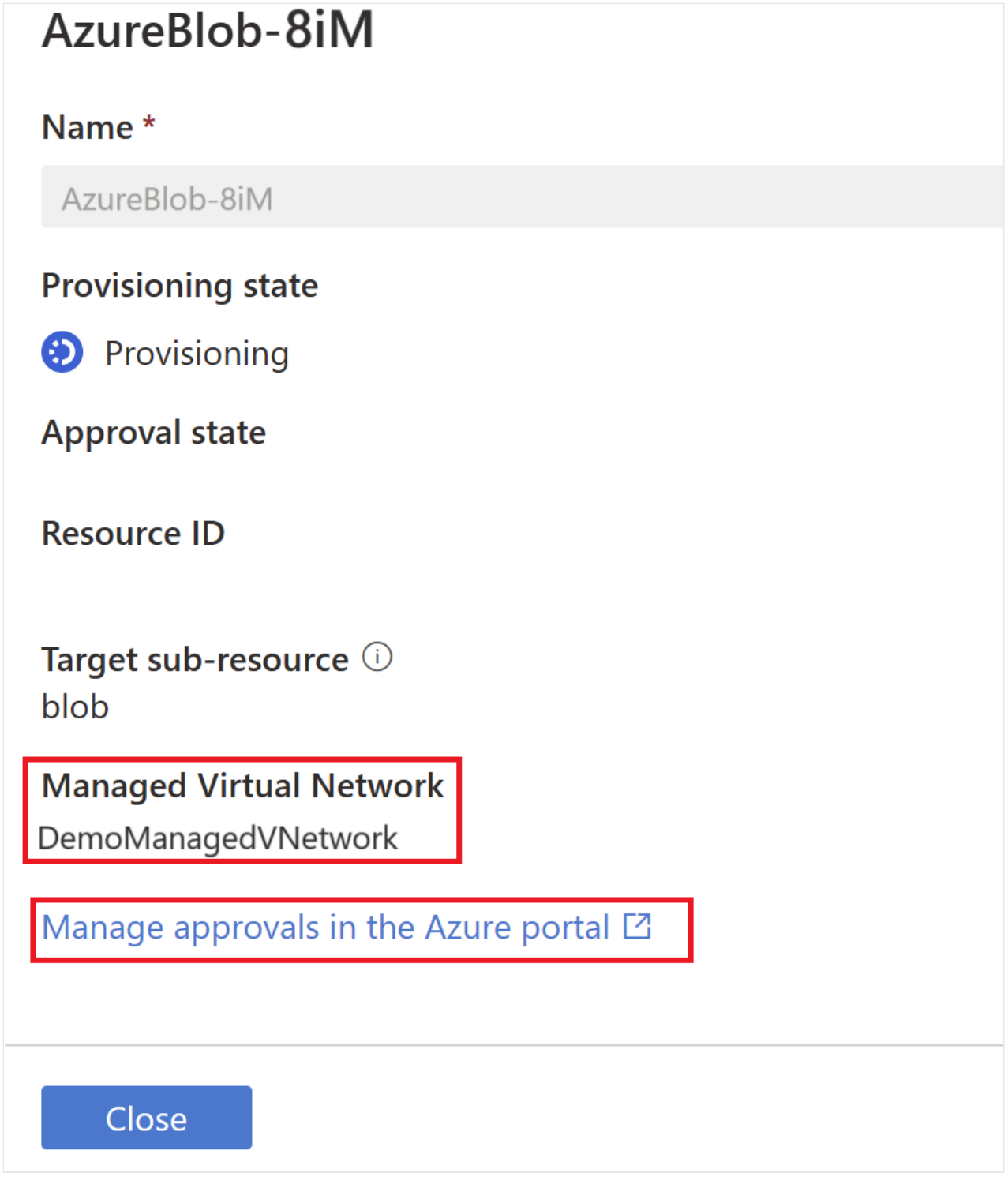
1. Navigate to **Management**, and select **Managed private endpoints**.
2. Select **+ New**.
3. From the list of supported data sources, select the type that corresponds to the data source you're planning to scan using Managed Virtual Network Integration Runtime.



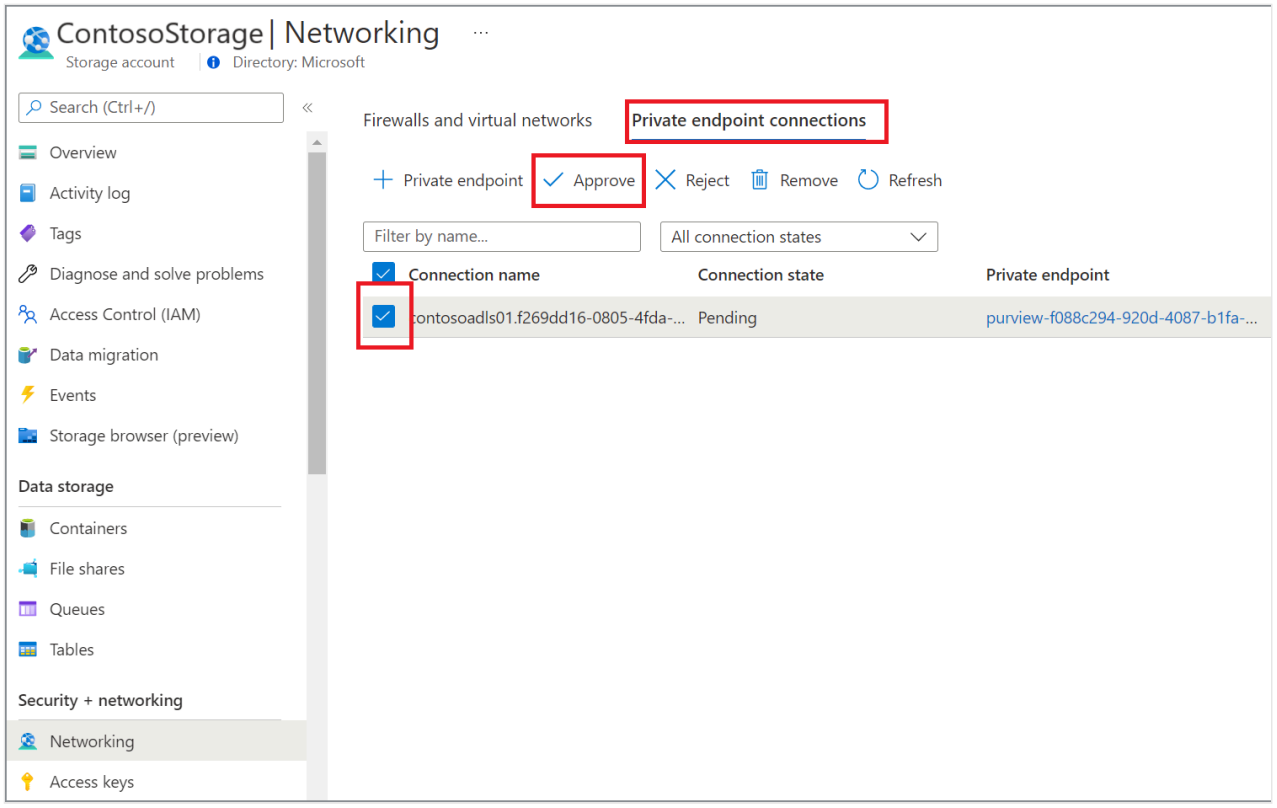
4. Provide a name for the managed private endpoint, select the Azure subscription, data source, and Managed Virtual Network from the drop-down lists. Select **Create**.

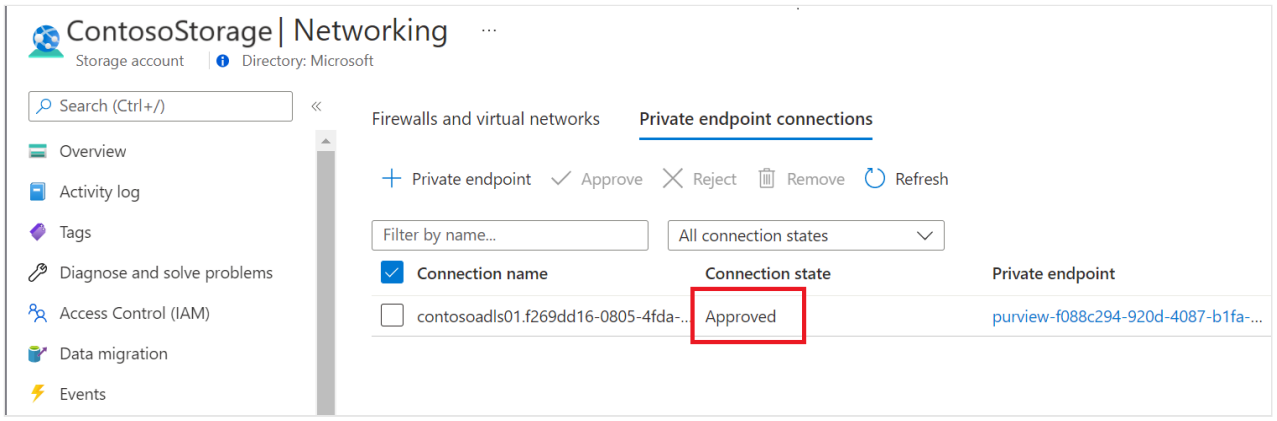


5. From the list of managed private endpoints, select the newly created managed private endpoint for your data source and then select on **Manage approvals in the Azure portal**, to approve the private endpoint in Azure portal.

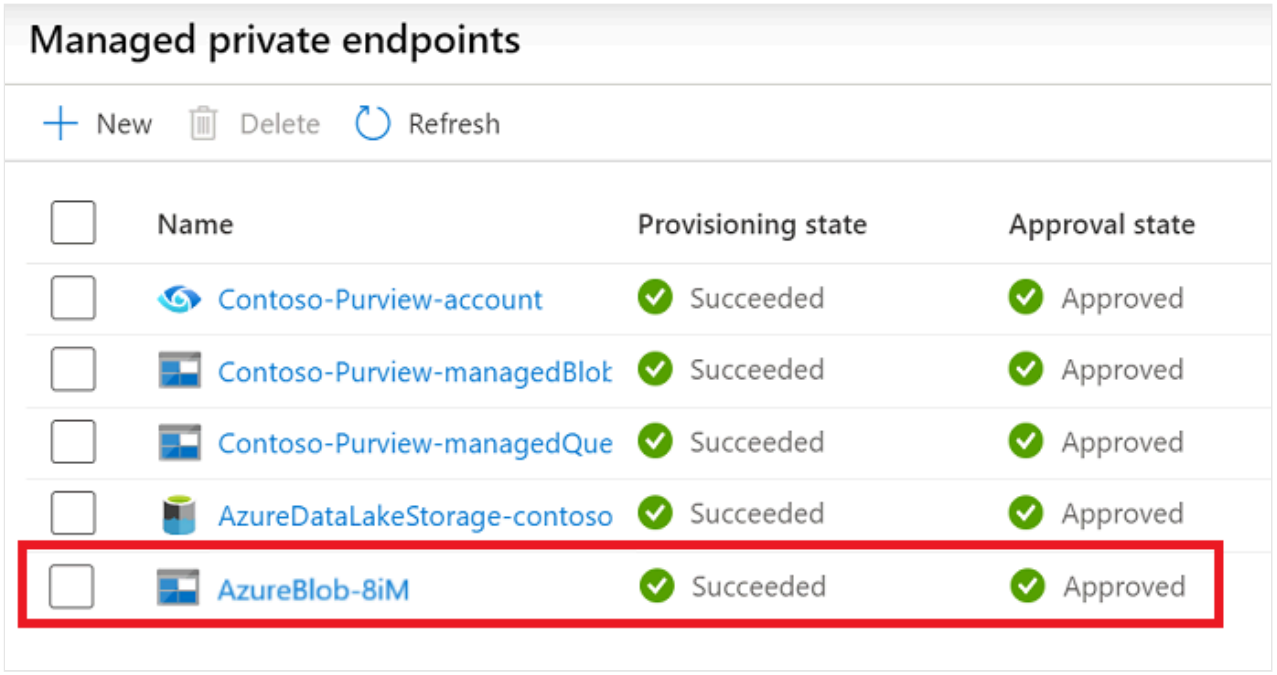


6. By selecting the link, you're redirected to Azure portal. Under the private endpoints connection, select the newly created private endpoint and select **Approve**.





7. Inside the Microsoft Purview governance portal, the managed private endpoint must be shown as approved as well.

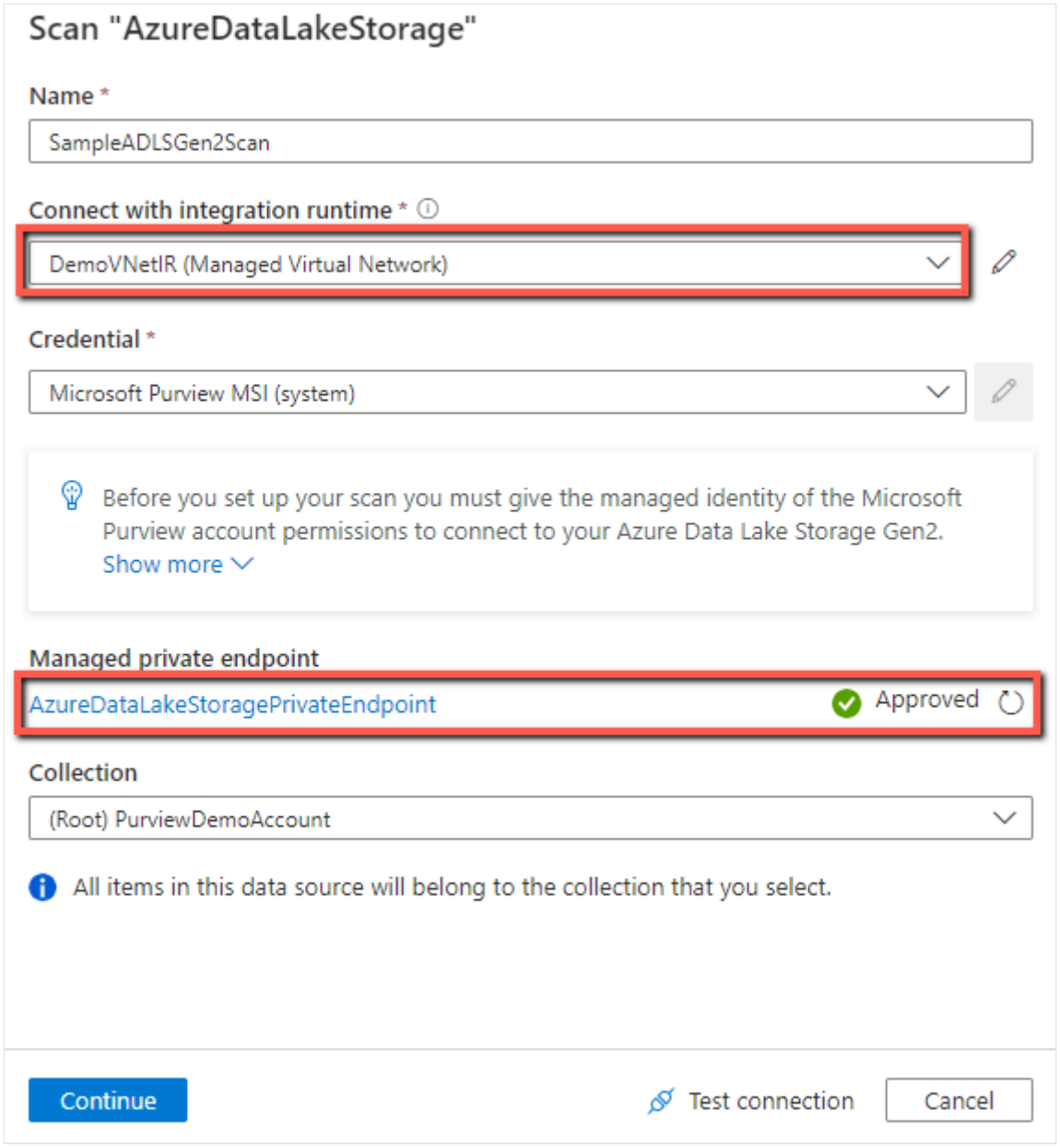


Set up scan to use managed virtual network IR

1. Go to **Data map** -> **Sources**, register the source that you want to scan if not already done.

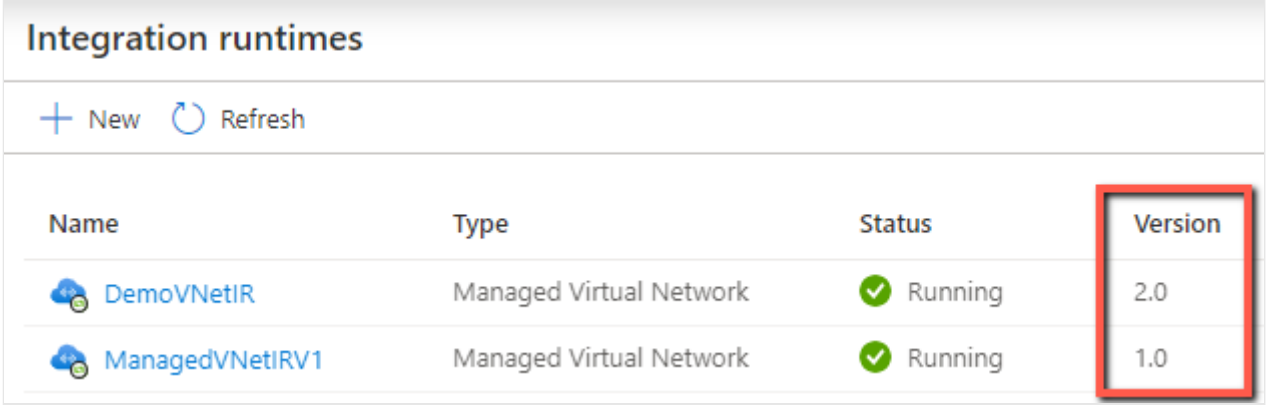
2. Go to your source -> **+ New scan**.

3. Under the **Connect via integration runtime**, select the managed virtual network IR you created. Configure the other scan settings as usual, and trigger a run.



Earlier version of Managed Virtual Network

To check the version of managed virtual network IR that you're using, go to **Data Map** -> **Integration runtimes**, and see the **Version** column.



The concept and architecture of Managed Virtual Network and managed private endpoint apply for both versions. The following table lists the differences in some aspects of the two versions. The version 2 provides better scan performance as well.

Expand table

Area	Version 1	Version 2 (current)
Supported regions	Available only in Australia East, Canada Central, East US, East US 2, North Europe, and West Europe.	Available in all Microsoft Purview supported regions.
Multi-vnet & multi-region	Only one virtual network supported at a time.	Up to five virtual networks supported across multiple regions.
Supported data source types for scan	Support scanning the following Azure data sources only. - Azure Blob Storage - Azure Cosmos DB - Azure Data Lake Storage Gen 2 - Azure Database for MySQL - Azure Database for PostgreSQL - Azure Dedicated SQL pool (formerly SQL DW) - Azure Files - Azure SQL Database - Azure SQL Managed Instance - Azure Synapse Analytics	Expanded to support scanning more data sources. See full list here.
Interactive authoring from Microsoft Purview portal (test connection, browse source during scan setup)	Need to turn on in Managed virtual network IR settings.	Always available
Name of the managed virtual network	default (autogenerated)	Configurable
Pricing	Charge on scan and ingestion jobs.	Charge on scan and ingestion jobs, and the running period of the managed virtual network IR. Learn more from Pricing details.

Update to latest version

Microsoft Purview added new version of Managed Virtual Network support mid-November 2023. All the newly created resources will be using the new offering. If you're using the earlier

version (which you can check), you can update to the latest version of the managed virtual network.

There are several advantages of using the new version of Managed Virtual Network:

- Generally available in all Microsoft Purview supported regions.
- Expanded data source support, including Databricks, Snowflake, Fabric, and all the sources that are supported by the default Azure integration runtime.
- Better scan performance.
- Interactive operations from Purview portal are always available (test connection, browse source during scan setup, etc.)

To upgrade, you can follow the instructions on this page: Upgrade to the new managed virtual network integration runtime.

Next steps

- Manage data sources in Microsoft Purview

Migrate to managed virtual network integration runtime

Article • 05/01/2025

Microsoft Purview added new version of Managed Virtual Network support mid-November 2023. All the newly created resources will be using the new offering. If you're using the earlier version (which you can check), you can update to the latest version of the managed virtual network.

There are several advantages of using the new version of Managed Virtual Network:

- Generally available in all Microsoft Purview supported regions.
- Expanded data source support, including Databricks, Snowflake, Fabric, and all the sources that are supported by the default Azure integration runtime.
- Better scan performance.
- Interactive operations from Purview portal are always available (test connection, browse source during scan setup, etc.)

> **Important**
>
> There is a pricing change to use a managed virtual network with Microsoft Purview. For more information about the capabilities of managed virtual networks and for pricing details, see the managed virtual network article.

To upgrade, you need to create a new Managed Virtual Network IR and switch your scans to this new IR. Here's an overview of what you'll need to do to migrate. All specific steps are listed after:

1. Ensure you meet prerequisites
2. Create a new Managed Virtual Network IR
3. Create managed private endpoints for data sources
4. Link your existing scans to the new managed virtual network IR
5. Delete the old managed virtual network V1

Deployment steps

Prerequisites

Before deploying a Managed virtual network and Managed Virtual Network Integration Runtime for a Microsoft Purview account, ensure you meet the following prerequisites:

1. From Microsoft Purview roles, you need Data Source Administrator permission on any collection in your Microsoft Purview account.
2. From Azure RBAC roles, you must be contributor on the Microsoft Purview account and data source to approve private links.

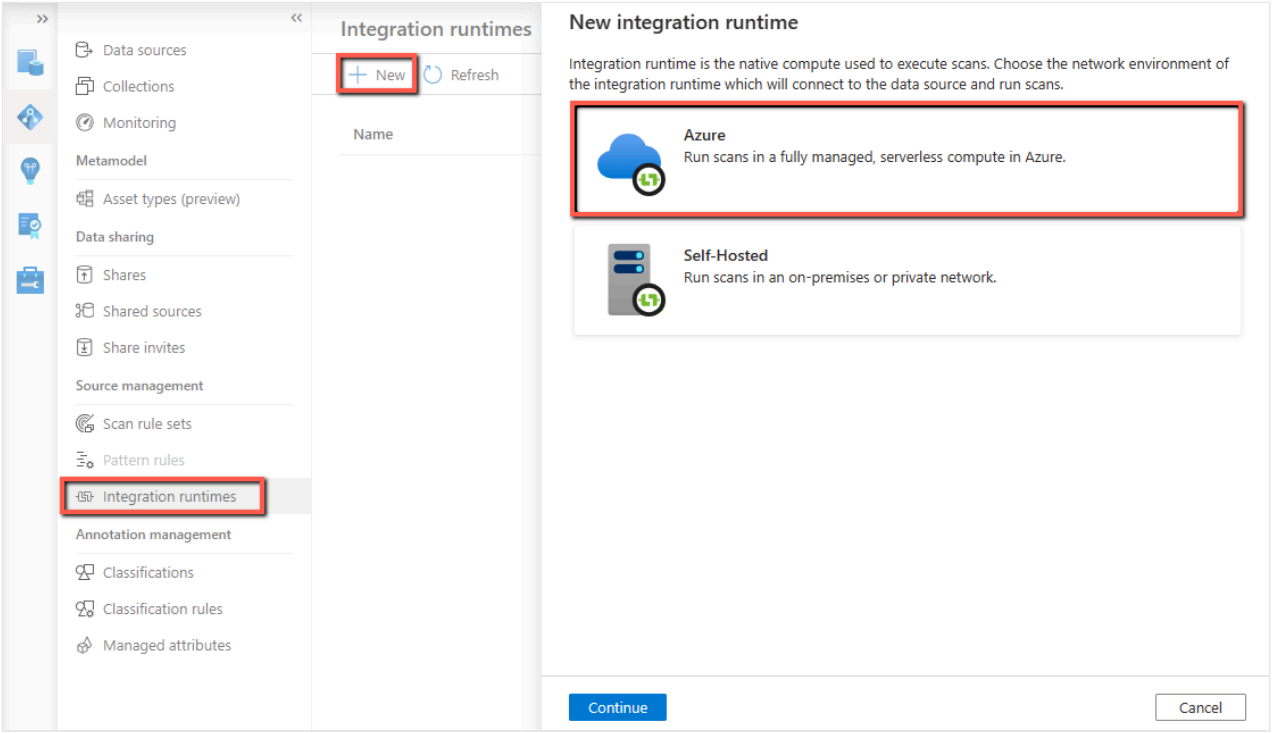
Create Managed Virtual Network Integration Runtime

1. Open the Microsoft Purview governance portal by:

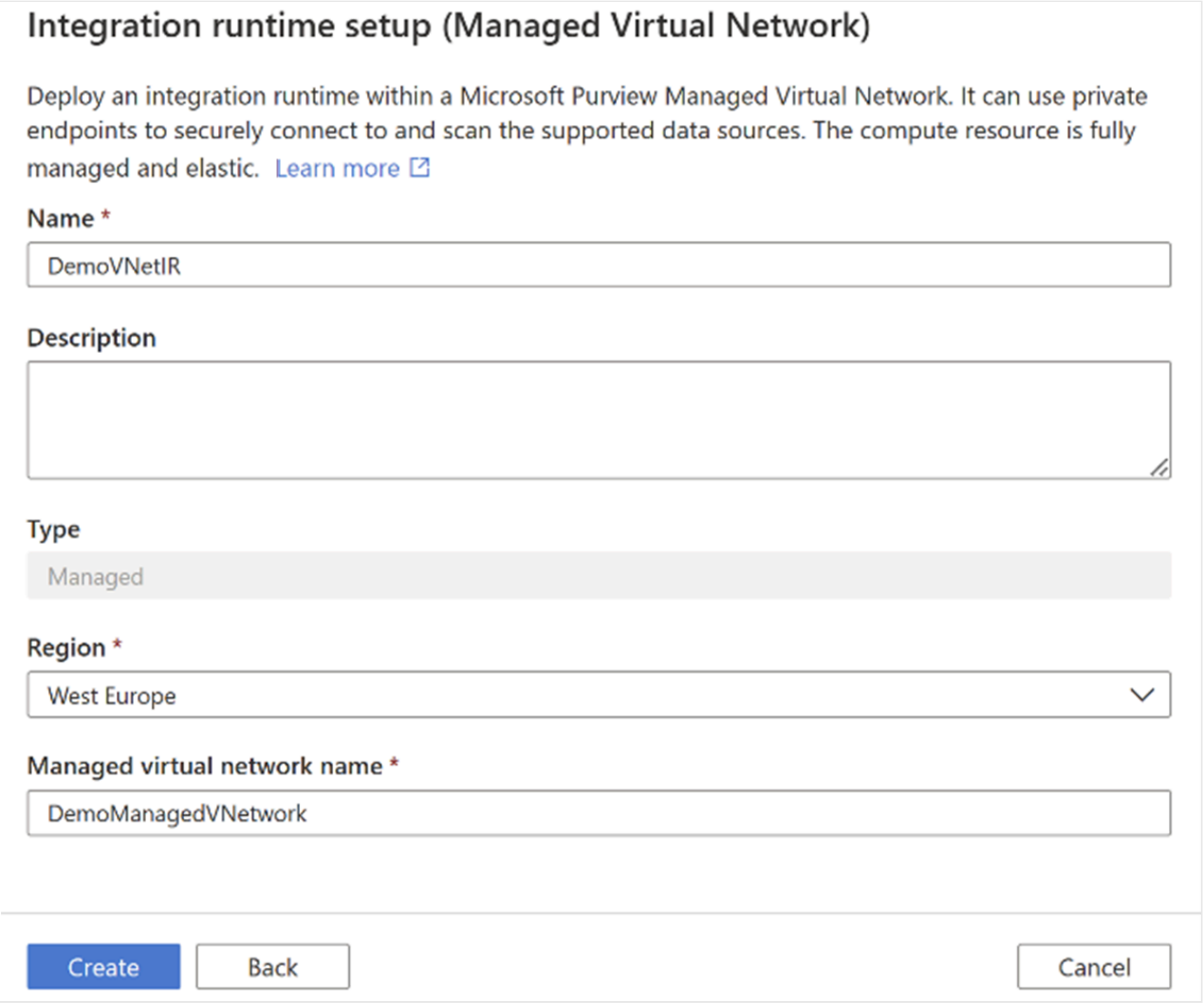
    - Browsing directly to https://web.purview.azure.com and selecting your Microsoft Purview account.
    - Opening the Azure portal, searching for and selecting the Microsoft Purview account. Selecting the **the Microsoft Purview governance portal** button.

2. Navigate to the **Data Map -> Integration runtimes**.

3. From **Integration runtimes** page, select **+ New** icon, to create a new runtime. Select **Azure** and then select **Continue**.

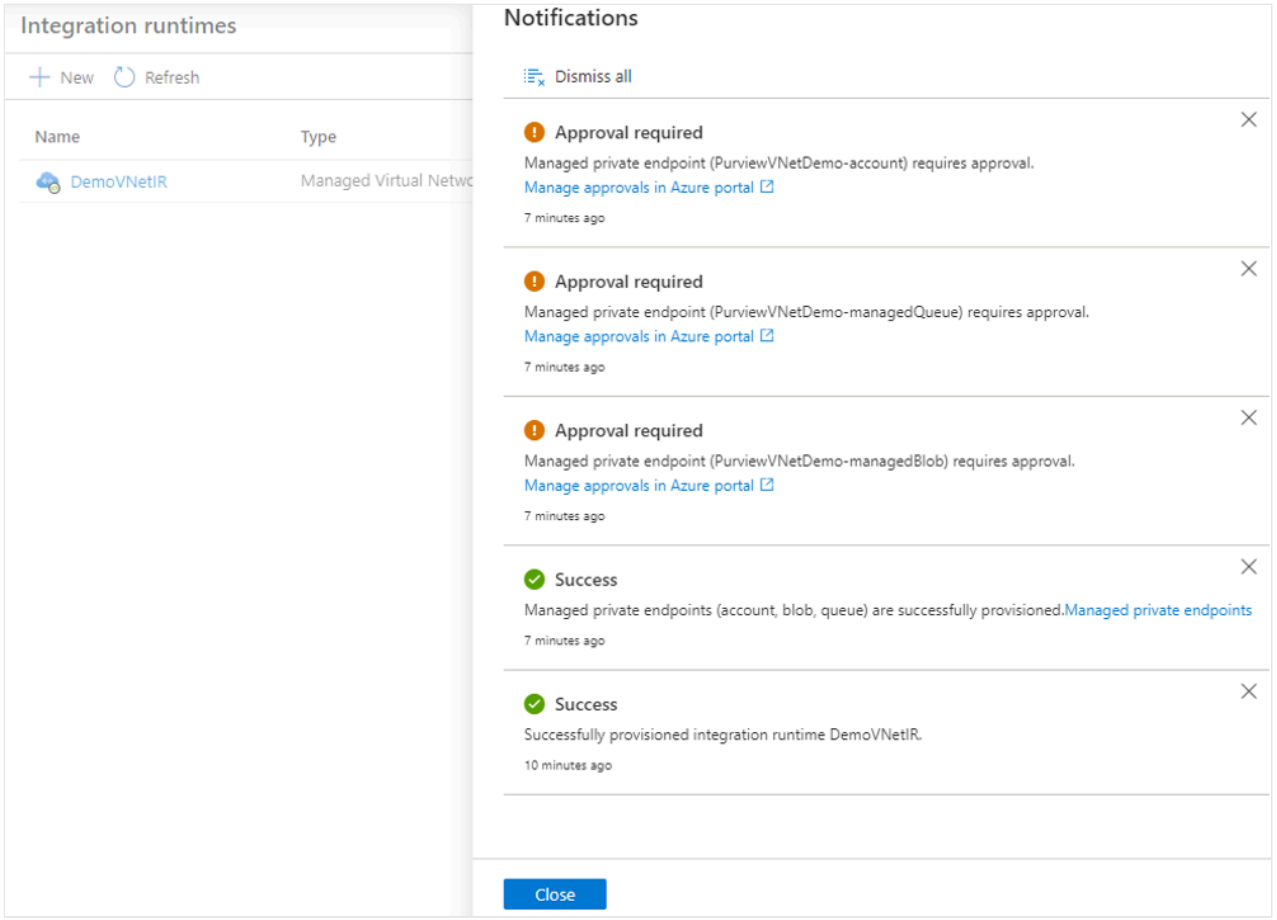


4. Provide a **name** for your Managed Virtual Network Integration Runtime, select a region, and give your managed virtual network a name.

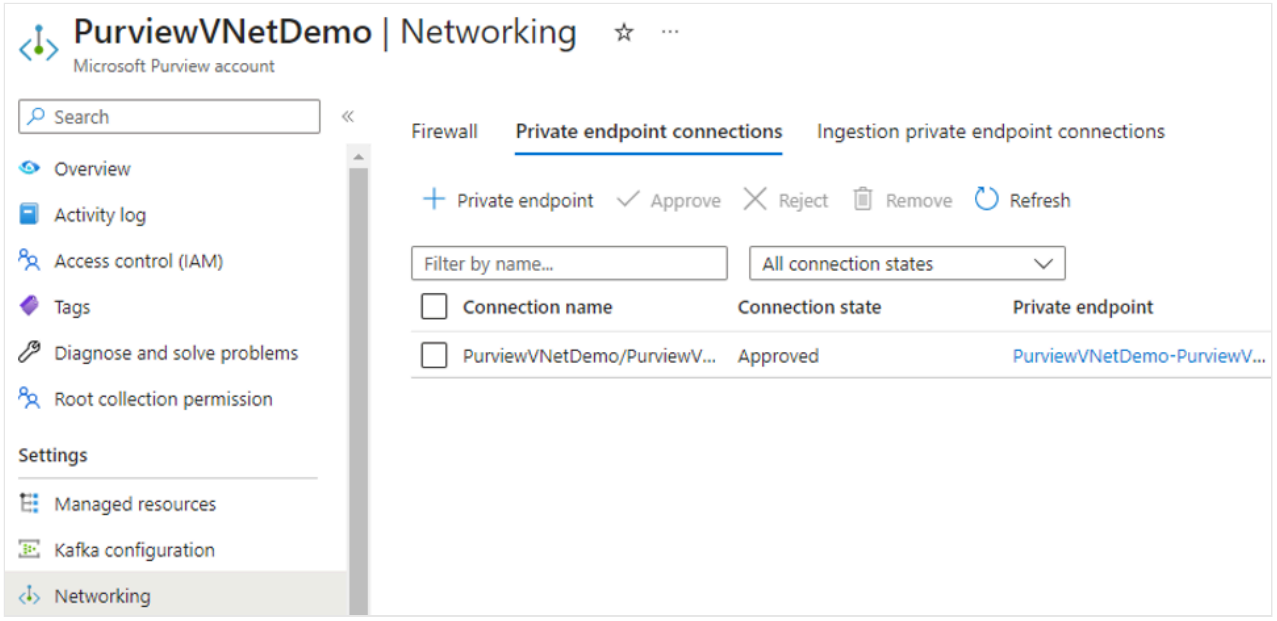


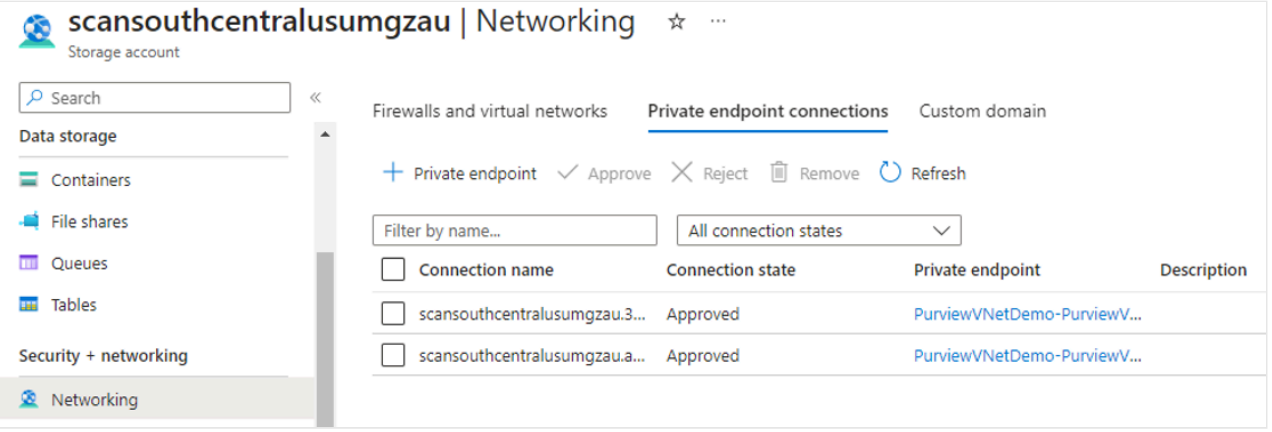
5. Select **Create**.

6. Deploying the Managed Virtual Network Integration Runtime triggers multiple workflows in the Microsoft Purview governance portal for creating managed private endpoints for Microsoft Purview and its managed Storage Account. Select each workflow to approve the private endpoint for the corresponding Azure resource.

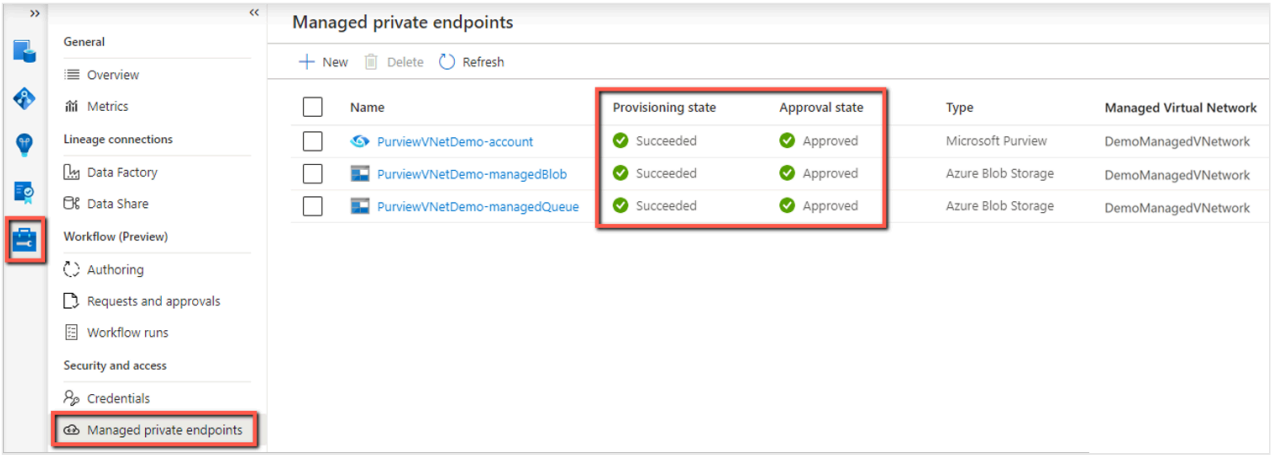


7. In Azure portal, from your Microsoft Purview account resource window, approve the managed private endpoint. From managed storage account page approve the managed private endpoints for blob and queue services:

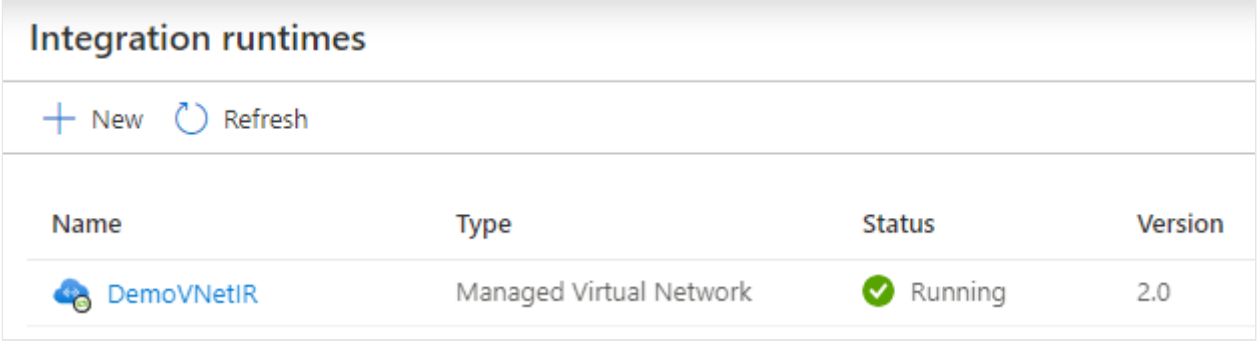




8. From Management, select Managed private endpoint to validate if all managed private endpoints are successfully deployed and approved.



9. Go to the Integration runtimes page, you'll see the IR status shown up as "Initializing" upon creation. Wait until it turns into "**Running**" state to use in scan. It usually takes several minutes.



> **Tip**
>
> You can create multiple managed virtual networks in different regions in your Microsoft Purview account to securely access resources across regions.

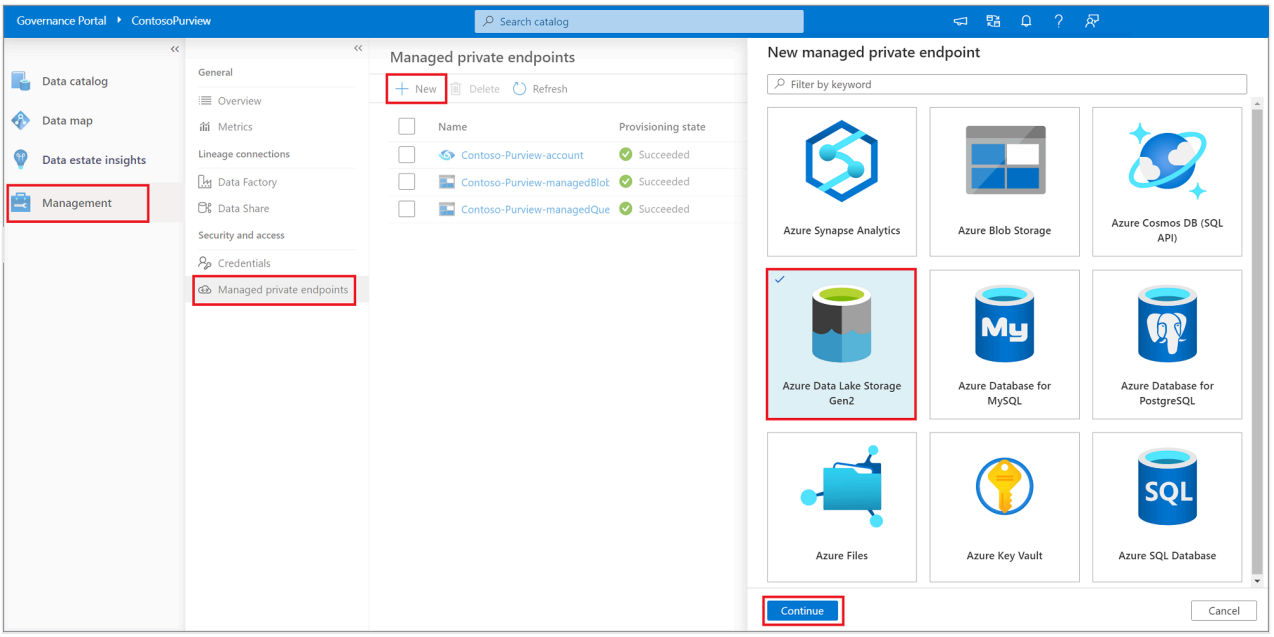
Create managed private endpoints for data sources

You can use managed private endpoints to connect your data sources to ensure data security during transmission.

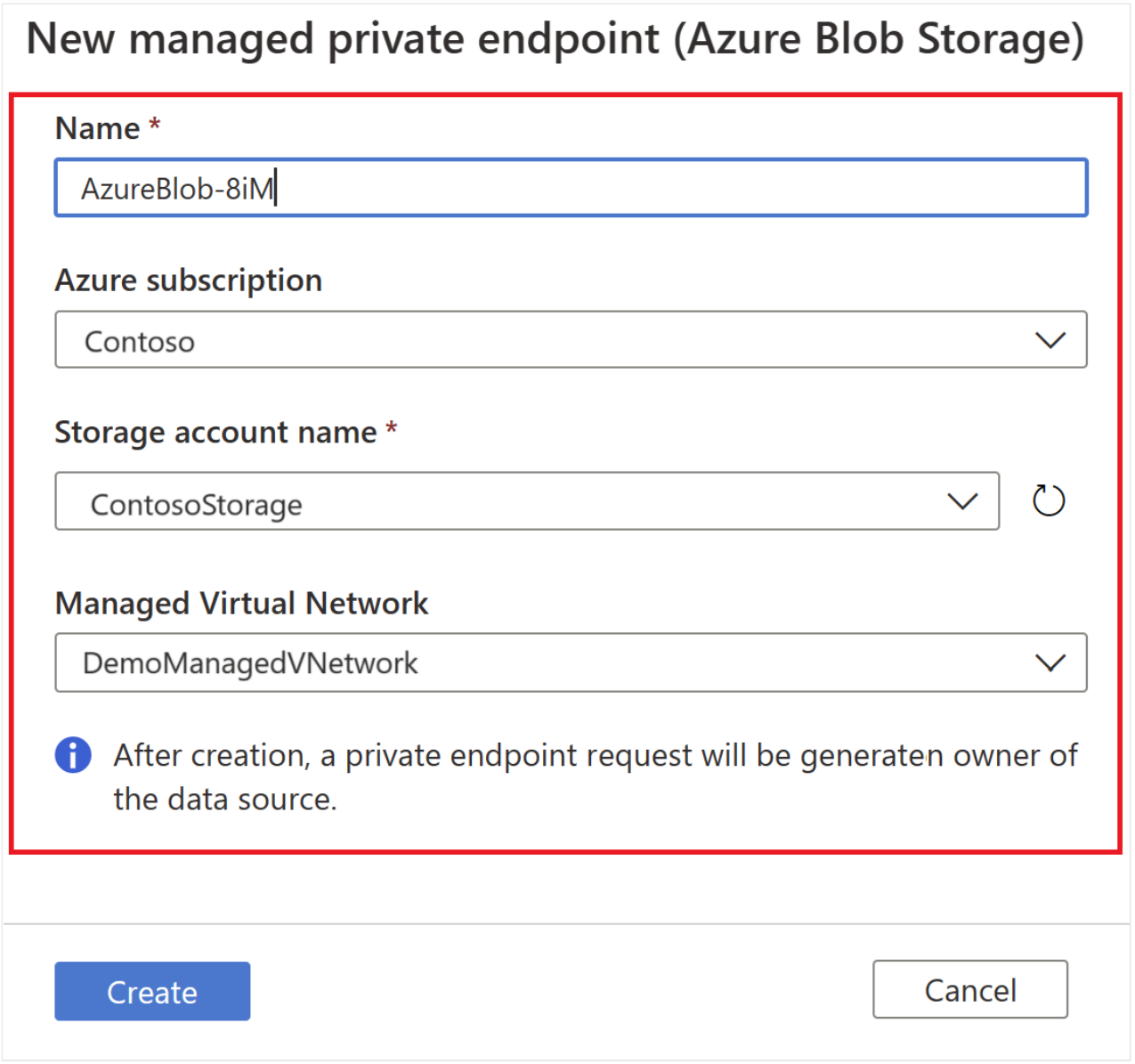
> **Tip**
>
> If your data source allows public access and you want to connect via public network, you can skip this step. Scan runs can be executed as long as the integration runtime can connect to your data source.

To deploy and approve a managed private endpoint for a data source, follow these steps selecting data source of your choice from the list:

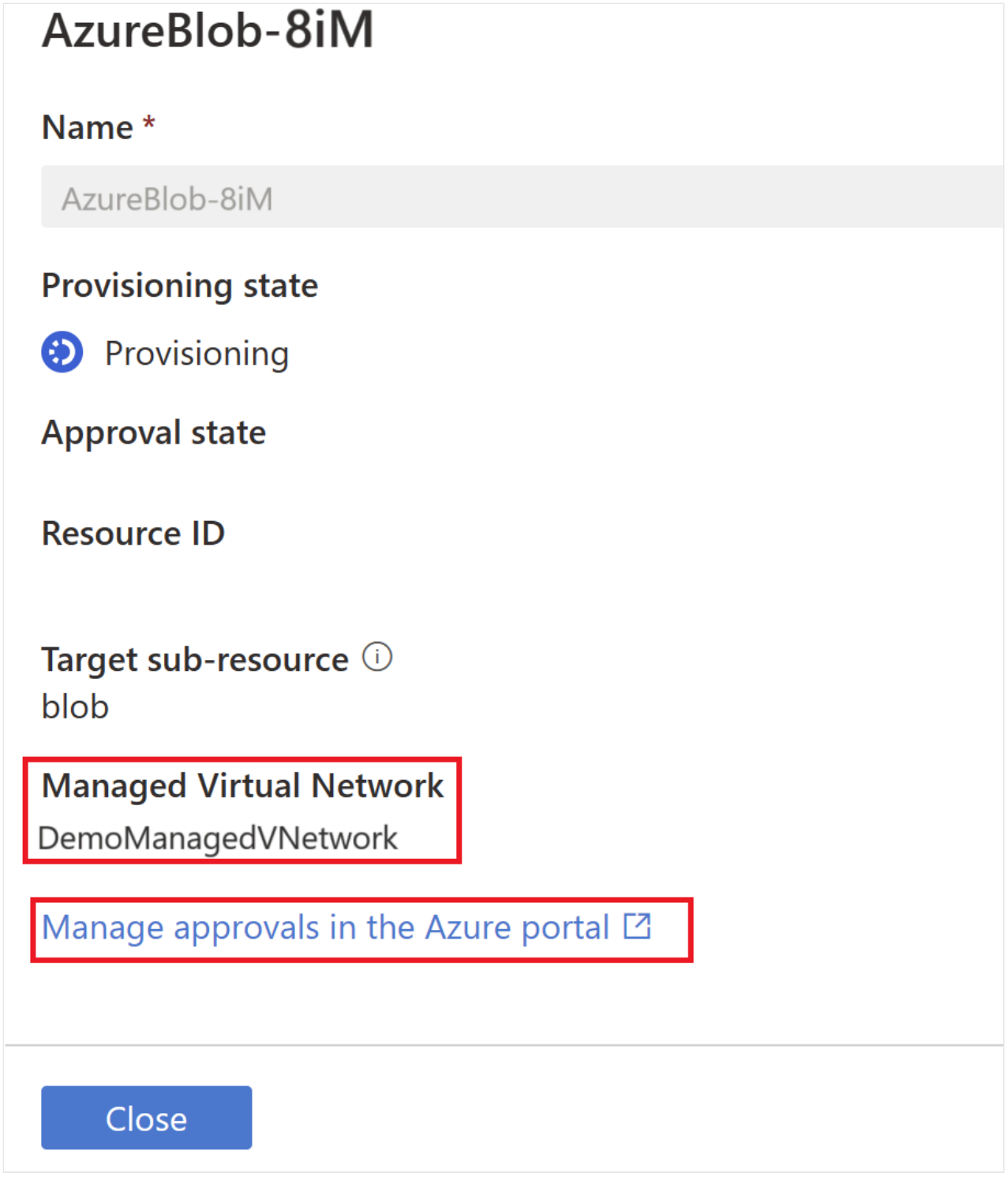
1. Navigate to **Management**, and select **Managed private endpoints**.
2. Select **+ New**.
3. From the list of supported data sources, select the type that corresponds to the data source you're planning to scan using Managed Virtual Network Integration Runtime.



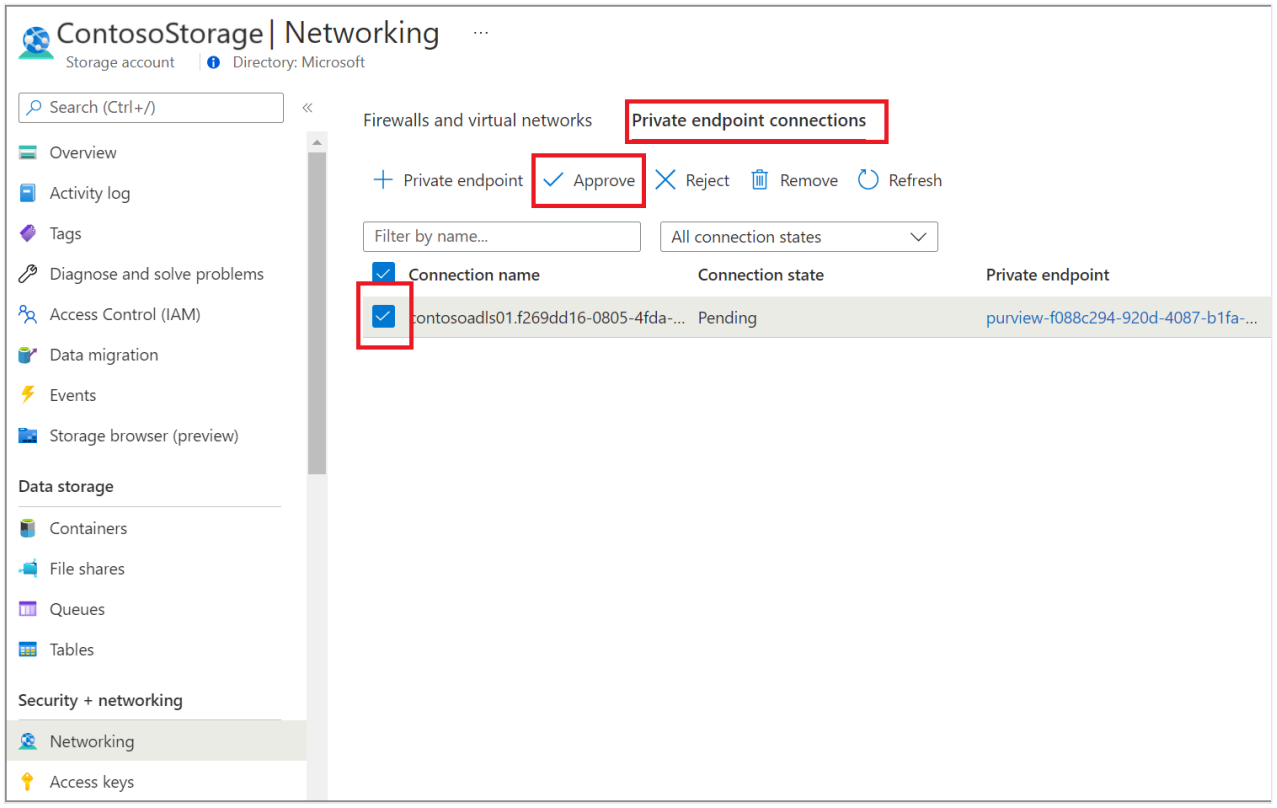
4. Provide a name for the managed private endpoint, select the Azure subscription, data source, and Managed Virtual Network from the drop-down lists. Select **Create**.

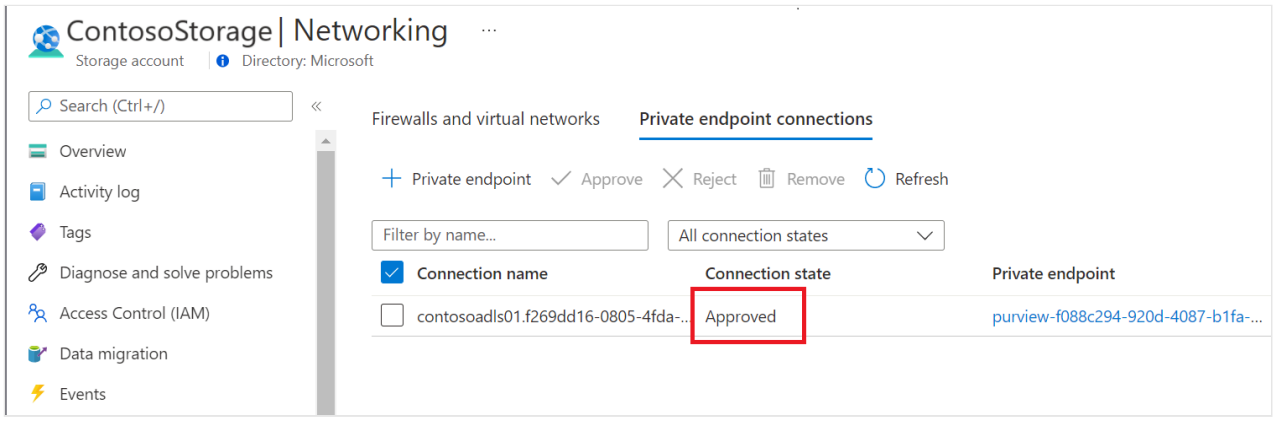


5. From the list of managed private endpoints, select the newly created managed private endpoint for your data source and then select on **Manage approvals in the Azure portal**, to approve the private endpoint in Azure portal.

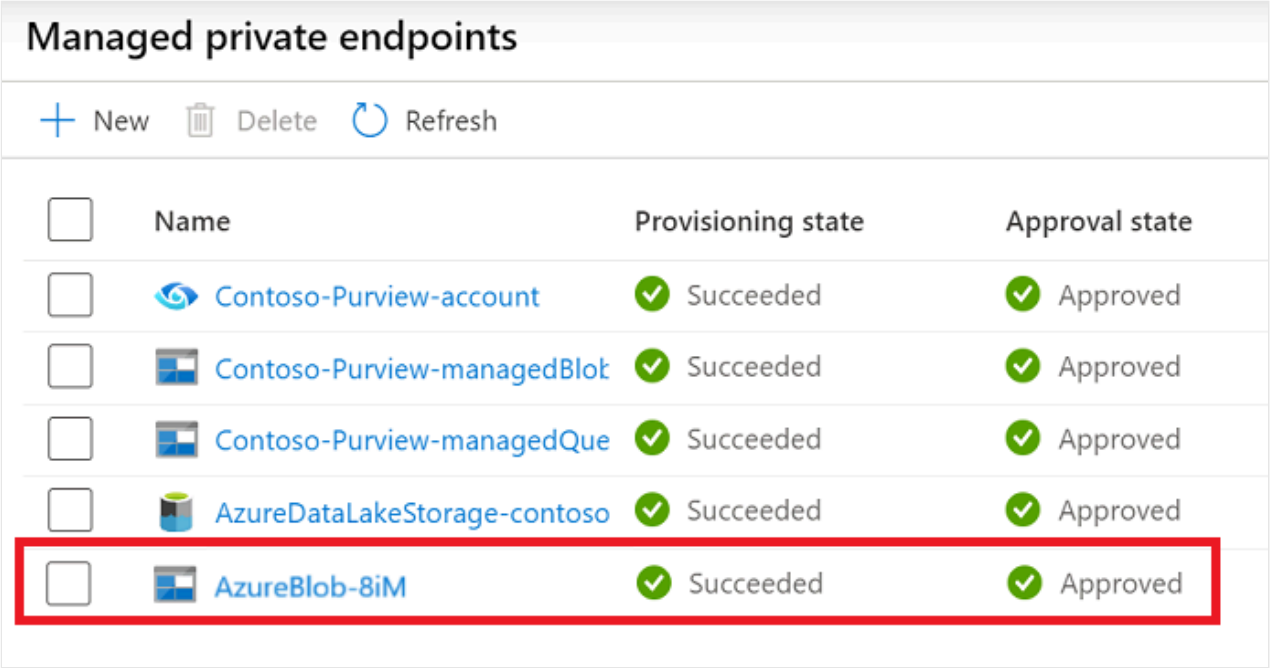


6. By selecting the link, you're redirected to Azure portal. Under the private endpoints connection, select the newly created private endpoint and select **Approve**.



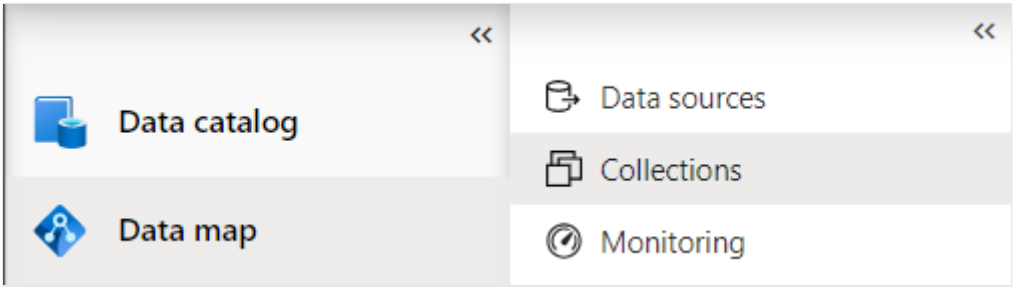


7. Inside the Microsoft Purview governance portal, the managed private endpoint must be shown as approved as well.

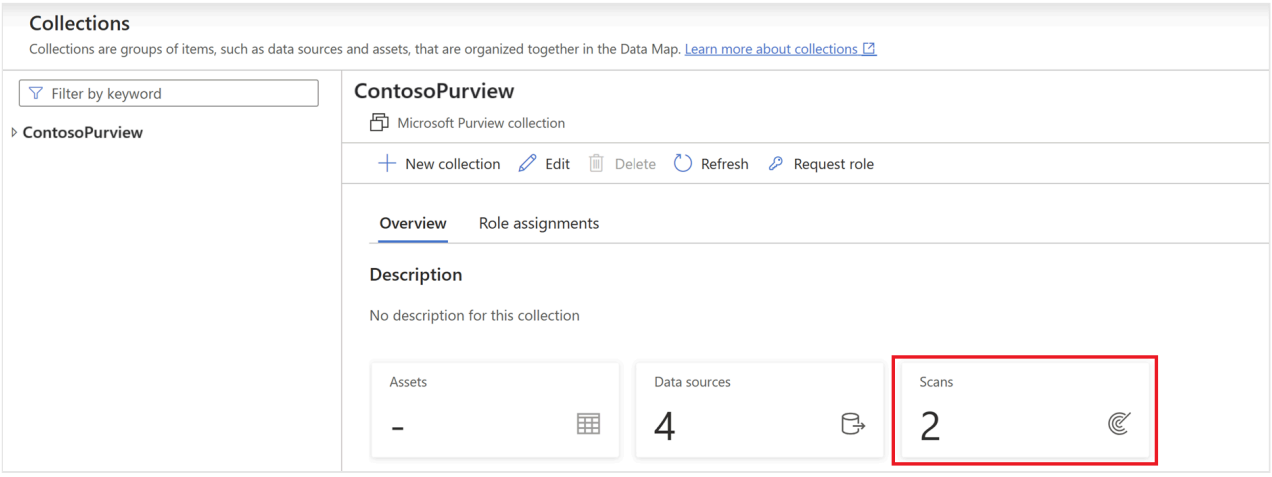


Link your existing scans to the new managed virtual network IR

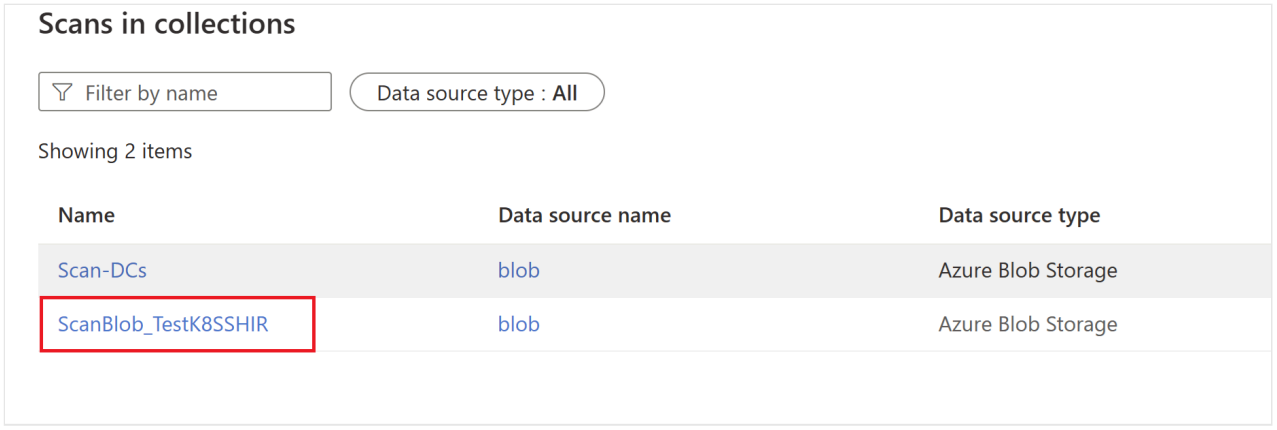
1. Inside Microsoft Purview, navigate to **Data Map**->**Collections**.



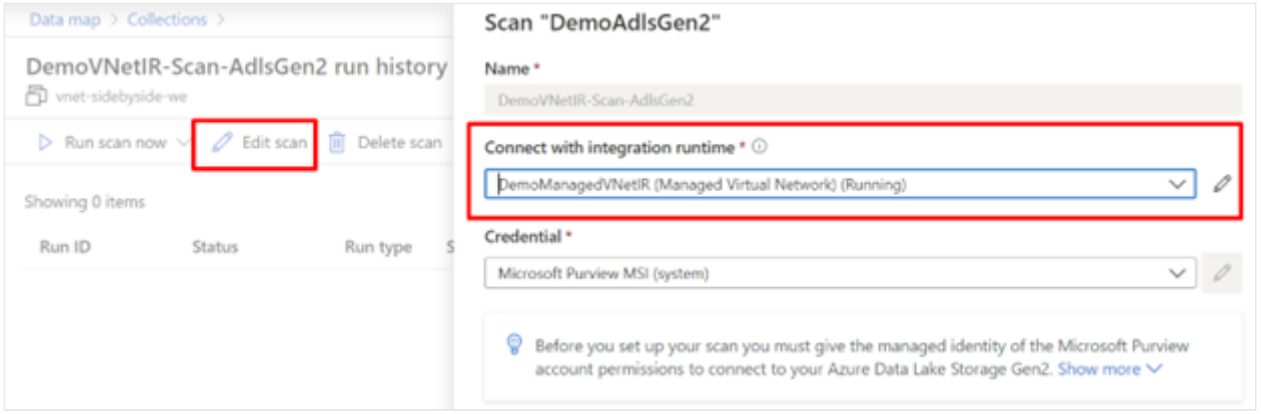
2. Select to your collections with scans and select the **Scans** button.



3. Select your scan name to see details.



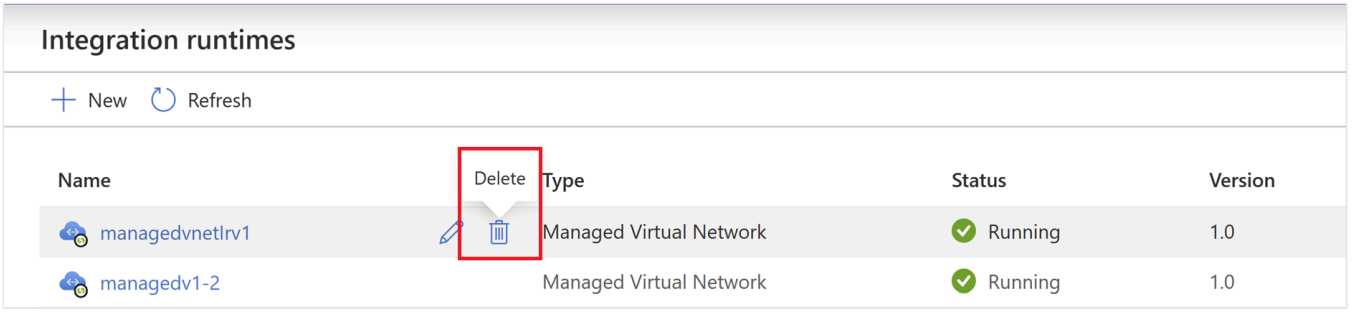
4. Select **Edit scan** and choose your new Managed Virtual Network IR from the drop-down.



5. Save your changes and edit any other scans.

Delete the old managed virtual network V1

You can look for your old Managed Virtual Network IRs under **Integration Runtimes** and delete them.



Next steps

- Manage data sources in Microsoft Purview

Master data management in Microsoft Purview

Master data management (MDM) serves as a foundational element to build comprehensive data governance framework. Master data management creates a source of truth for your data, so you can check your key assets against this source of truth to confirm your data is trustworthy. You do this by creating a deduplicated set of master data records, a golden standard by which to measure your data. The standardization process you use to create this set of master data records differs across organizations, and Microsoft Purview integrates with partner solutions to offer helpful strategies.

Partner MDM integration with Microsoft Purview

Microsoft Purview integrates with various MDM solutions to empower data product owners to create and publish golden records that represent the authoritative version of master data assets. These integrations operate bi-directionally, enabling MDM systems to access metadata from Microsoft Purview and connect with the source systems to unify, cleanse, and standardize data, thereby creating master data assets. Later, the metadata of these master data assets is published back to Microsoft Purview in Unified Catalog, where it can be curated as a data product.

Conversely, the creation of golden data assets for master data products is facilitated within the integrated MDM solution, applying its capabilities for data management and governance. This collaborative approach ensures that organizations can use the full potential of master data for analytics and insights while maintaining data integrity and compliance.

Partner solutions and integration instructions

CluedIn

- Master data management with CluedIn
- Microsoft Purview and CluedIn integration

Profisee

- Master data management with Profisee
- Microsoft Purview and Profisee integration

Reltio

- Master data management with Reltio
- Microsoft Purview and Reltio integration

Semarchy

- Master data management with Semarchy
- Microsoft Purview and Semarchy integration

Data flow and integration steps

Here's a step-by-step guide on how to accomplish the tasks mentioned using Microsoft Purview and a partner MDM application:

1. Register and scan a data source in your Microsoft Purview Data Map.
2. Create a data product to use as your master data product.
3. Create a glossary term and link it to your data product.
4. Curate the master data product:
    a. Within the master data product, review the existing data assets and determine which ones are relevant to include.
    b. Add details description, use cases, etc., to ensure that the master data product contains the most relevant and up-to-date information.
5. The master data management application receives signal and pulls metadata:
    a. Configure your master data management application to listen for signals or events indicating changes in the Microsoft Purview metadata. This process is slightly different depending on which application you're using.
    b. When a signal is received, the MDM application pulls the required metadata related to the master data product and associated data assets from Microsoft Purview.
6. The master data management application connects data sources and copies data to create master dataset:
    a. Configure the MDM application to connect to the relevant data sources where the data assets are located.
    b. Copy the data from these sources to the MDM application's storage environment.
7. Unify, clean, and standardize the data
    a. Within the MDM application, unify, clean, and standardize the copied data to ensure consistency and accuracy.
    b. Apply data governance rules and transformations as needed to prepare the data for creating the master data asset.
8. Create a master data asset of Golden Records

a. Utilize the MDM application's capabilities to create a master data asset comprised of golden records that represent the consolidated and standardized view of the data.
b. Ensure that the master data asset meets the defined criteria and quality standards.

9. Publish the metadata to Microsoft Purview and link the new data asset to your data product.

By following these steps, you can seamlessly integrate with a partner MDM solution to manage glossary terms, curate master data products, synchronize metadata, and create and publish master data assets. This integration helps maintain data consistency, quality, and governance across your organization's data landscape

Where to find master data assets in Unified Catalog

Master data assets are found on a data product's details page of data produts. On the details page, locate the **Data assets** section. Each MDM asset is listed on a card labeled with the name of the MDM solution provider, such as Profisee or Semarchy. Learn more about searching for and browsing data products.

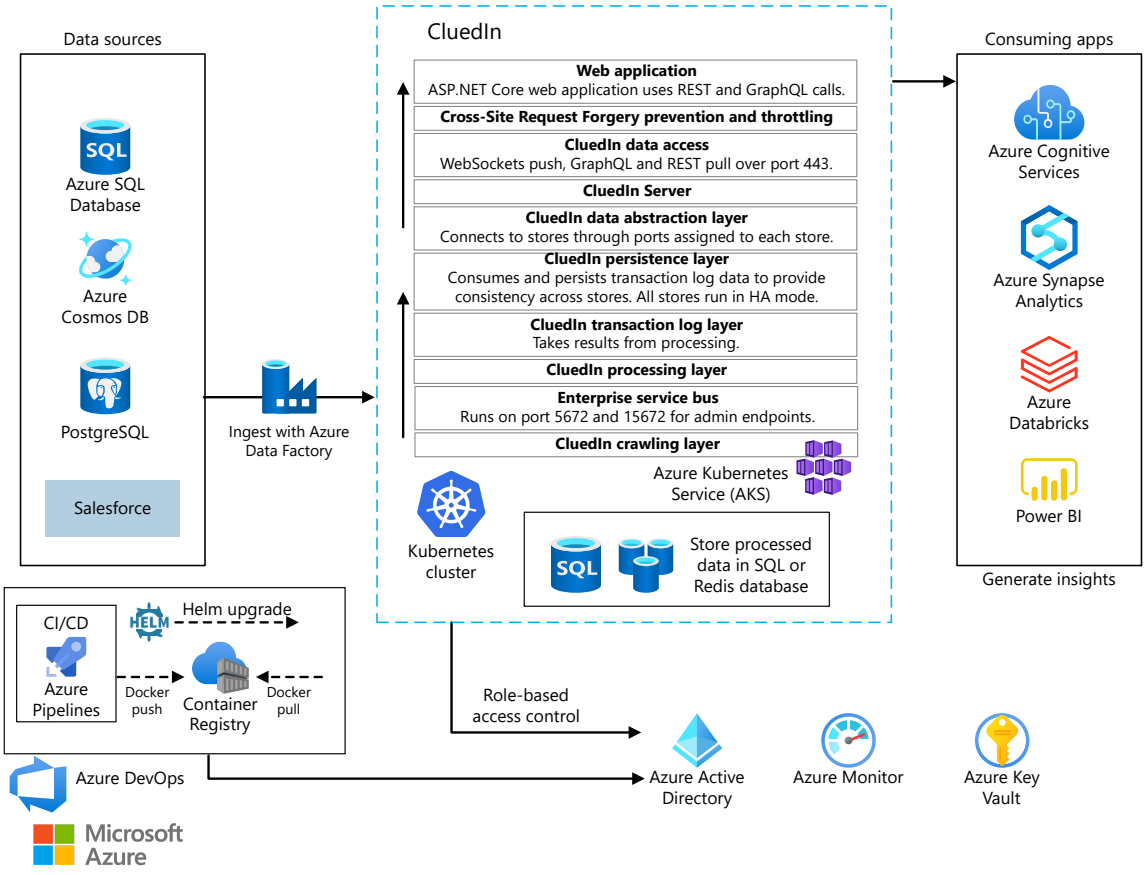
Last updated on 04/24/2026

Microsoft Purview and CluedIn integration for master data management

Article • 04/03/2025

The CluedIn architecture provides businesses with metrics about the quality of data it ingests, intelligently detecting dirty data and preparing it for cleaning by data engineers and data stewards. Proprietary fuzzy logic machine learning algorithms help business users and curators label data, and teach the system to identify, correct, and prevent data quality issues over time.

Architecture



Dataflow

The CluedIn solution consists of various functional layers that run in a Kubernetes cluster in Azure Kubernetes Service (AKS). A combination of .NET Core microservice applications

handles distinct functions like data ingestion, streaming data processing, queuing, and user interface.

1. The CluedIn crawling layer ingests data from customer cloud sources like Azure SQL DB, Azure Cosmos DB, PostgreSQL, and Salesforce databases via Azure Data Factory connectors.

   CluedIn also takes input from on-premises accessible systems like SAP, Oracle, IBM, and Hadoop, or can use on-premises agents to crawl non-public data.

2. The enterprise service bus connects through ports 5672 and 15672 for admin endpoints. Crawlers send data to the bus, and the processing layer consumes data from the bus, over port 5672.

3. The transaction log layer takes results from the processing layer.

4. In the persistence layer, databases consume data from the transaction log and persist it to provide eventual consistency across the different data stores. All the stores run in high-availability (HA) mode.

   Unlike with data virtualization, the CluedIn persistence layer ingests parts of the source data and preserves the highest fidelity version of data and its structure. This high fidelity means that the CluedIn Data Fabric can serve business requests for data in any format or model.

5. The data abstraction layer connects to the different data stores through the ports for each store.

6. Data access is through GraphQL, REST, and WebSockets calls over port 443. GraphQL and REST use a pull model, and WebSockets uses a push model.

   CluedIn protects data access through throttling and Cross-Site Request Forgery (CSRF) prevention.

7. The CluedIn ASP.NET Core web application communicates through a combination of REST and GraphQL calls over port 443.

   All communication from the browser into the application uses a set of ingress definitions, which require only a single public IP address. In a production environment, all communication is over secure socket layer (SSL).

8. The CluedIn application provides cleaned, processed data to analytics services like Power BI and Azure Synapse Analytics for generating insights. The system backs up and stores all data in SQL or Redis databases.

Components

CluedIn runs on Azure Kubernetes Service (AKS), a highly available, secure, and fully managed Kubernetes service for deploying and managing containerized applications. AKS offers serverless Kubernetes, integrated CI/CD, and enterprise-grade security and governance.

CluedIn uses and supports many database sources and services, including:

- Azure SQL Database, a managed relational cloud database service that is always up to date and can automatically scale resources on demand.
- Azure SQL Managed Instance, for broad SQL Server engine compatibility with existing SQL Server applications. SQL Managed Instance provides on-premises database infrastructures with Azure cloud benefits like elastic scale, unified management, and a cloud billing model.
- Azure Cosmos DB, a fully managed, nonrelational NoSQL serverless database for modern app development.
- Azure Data Lake, a scalable data storage and analytics service.
- Azure Data Factory, a fully managed, serverless data integration solution for ingesting, preparing, and transforming data at scale. CluedIn uses over 90 built-in Data Factory connectors to acquire data from sources like Amazon Redshift, Google BigQuery, HDFS, Oracle Exadata, Teradata, Salesforce, Marketo, ServiceNow, and all Azure data services.

CluedIn provides processed, governed data to many analytics apps and services, including:

- Azure Databricks, a fast, easy, and collaborative Apache Spark-based analytics service.
- Azure Synapse Analytics, a limitless analytics service that brings together enterprise data warehousing and Big Data analytics.
- Log Analytics, an Azure portal tool to edit, run, and analyze queries from Azure Monitor log data.
- Azure Cognitive Services, a comprehensive family of AI services and cognitive APIs for building intelligent apps.
- Power BI, a Microsoft business analytics service that combines interactive visualizations and business intelligence with an easy-to-use report creation interface.

Scenario details

Modern enterprise companies base many processes and projects on data, but the raw data has to be prepared for consumption. Data use cases from advanced analytics to machine learning all require similar data preparation processes and attention.

1. Data projects start with data *discovery*, to determine where data is and what systems it uses.
2. Data *integration* then brings multiple data sources together into a unified or connected data set.
3. The next step is to *normalize, standardize, harmonize,* and *clean* the data so machines can process it in a uniform, consistent, and high-fidelity manner.
4. Finally, the data must be made easily and readily available for business needs.

During these processes, *governance* must ensure data control and privacy protection with clear ownership, full traceability, and an audit trail of data origins, processing, and use.

The CluedIn platform encapsulates these data management processes and pillars into a coherent, consistent, end-to-end Master Data Management (MDM) solution. CluedIn uses a data integration technique called *eventual connectivity* that yields better results than classic *extract, transform, load (ETL)* or *extract, load, transform (ELT)* models. Eventual connectivity uses GraphQL queries to blend data seamlessly from across many siloed data sources.

With eventual connectivity, data isn't joined or blended upon entry or loading into other systems. Instead, CluedIn loads the data as is, and tags records using metadata. Eventually, records with the same tags merge or build a relationship in the graph.

This sophisticated data merging technique provides a foundation for data-driven solutions. The CluedIn Data Fabric integrates data into a pipeline that cleans, prepares, models, governs, enriches, deduplicates, and catalogs data to make it easily available and accessible for business uses.

CluedIn provides businesses with metrics about the quality of data it ingests, intelligently detecting dirty data and preparing it for cleaning by data engineers and data stewards. Proprietary fuzzy logic machine learning algorithms help business users and curators label data, and teach the system to identify, correct, and prevent data quality issues over time.

CluedIn includes enterprise-grade governance, for assurance that you can use your data safely and confidently. CluedIn can stream cleaned, governed data directly to analysis systems like Power BI, Azure Databricks, Azure Synapse Analytics, or Azure Cognitive Services to make it easily available to the rest of the business. Native support for

autoscaling uses the power of Azure to provide a scalable environment for the biggest data workloads.

Potential use cases

Building a single view of data

- Due to CluedIn's semantic modeling, it makes building a Single View of your Master Data an easier thing to achieve compared to traditional approaches. CluedIn's customers are using CluedIn to build a connected, historical, and high quality view of their most critical business data. CluedIn not only supports mastering of classic Master domains like People, Companies, Vendors, and Products - it supports and endless number of different domains as well as unstructured domains like files, mail, events, and more. If you require a centralized repository of master data that is clean, enriched, governed, quality-controlled, and cataloged, then CluedIn is a good fit for your use cases.

A data fabric

- CluedIn is a Gartner Cool Vendor in 2020, due to its ability to orchestrate data from across 10s, 100s and 1000s of different and complex datasources into a unified data hub. If you need to wrangle data from many different datasources with ease, then CluedIn can be used as a data fabric to achieve this. This can provide a streaming infrastructure for your data that can also proactively clean and master the data as it flows onto downstream consumers.

Sophisticated merging and linking of master data

- CluedIn's unique data modeling approach utilizes a graph database, which allows for complex data to be merged and linked with simplicity. Unlike traditional approaches, to solve this challenge, CluedIn adds more machine learning and graph analytics to merge, match, and link records with a high precision.

Considerations

These considerations implement the pillars of the Azure Well-Architected Framework, which is a set of guiding tenets that can be used to improve the quality of a workload. For more information, see Microsoft Azure Well-Architected Framework.

Reliability

Reliability ensures your application can meet the commitments you make to your customers. For more information, see Overview of the reliability pillar.

- CluedIn takes automatic daily database backups and keeps them in long-term storage for 30 days by default. The entire platform is built on redundant, fault tolerant stacks that maintain backups for all subsystems. Round the clock monitoring systems ensure that services are as untainted as possible. CluedIn follows industry standard practices for infrastructure redundancy.
- CluedIn surfaces and stores only a representation of your data, not the original version. If CluedIn detects destructive data intrusion, it can temporarily wipe the CluedIn data from your servers. Once the intrusion subsides, CluedIn regathers the data to get back to its original state.
- All data stores run in high-availability mode.

Scalability

- CluedIn runs in Docker containers and uses Kubernetes to host and orchestrate the different pieces of the application. This architecture means that CluedIn works well in elastic environments and can automatically scale to needed sizes and infrastructure.
- Native support for autoscaling applies the power of Azure to provide a scalable environment for the biggest data workloads.
- Schemaless graph modeling automatically infers a data model from the source data. New data sources automatically connect to all other data sources, rather than having to be explicitly integrated. The number of data sources can scale infinitely without increasing integration complexity.

Security

Security provides assurances against deliberate attacks and the abuse of your valuable data and systems. For more information, see Overview of the security pillar.

- CluedIn security grants permissions and controls access to different services through Azure RBAC, with Azure Key Vault security key control and Azure Monitor access tracking and logging.

- In addition to authenticated user accounts, CluedIn also supports single sign-on (SSO) and identity frameworks. Requests to the CluedIn application use encrypted access tokens that have no correlation with user identity.
- CluedIn manages stored data representations behind multiple firewall and proxy layers, and authenticates them with a set of unique keys.
- CluedIn stores all sourced data with 256-bit AES encryption, which is stronger than or equal to the encryption level of the supported data sources.
- Throttling and CSRF prevention protect data access.

DevOps

- CluedIn uses Azure Pipelines continuous integration and continuous delivery (CI/CD) pipelines to handle deployments and rolling updates to the AKS environment.
- CluedIn supports unit, integration, and functional testing to ensure that data transforms as expected. Virtualized processing pipelines can run in memory for sandbox testing. Production-grade assertions can help debug and track down data issues.
- For testing and production environments, CluedIn provides a Helm package manager chart to install CluedIn quickly in a Kubernetes cluster. Fully scripted data deployment processes support setup, test, and rollout.

Cost optimization

Cost optimization is about looking at ways to reduce unnecessary expenses and improve operational efficiencies. For more information, see Overview of the cost optimization pillar.

The pricing for CluedIn is open and transparent. You can see the pricing on their website.

Azure sizing and starting a trial

You can start a 7-day trial of CluedIn on their website, which can also help you scope your Azure hosting costs with prebuilt Azure estimates for different sized environments.

Deploy this scenario

- To deploy CluedIn for development and evaluation purposes using Docker, see CluedIn with Docker.
- To install CluedIn quickly in a Kubernetes cluster, see CluedIn with Kubernetes. The Helm chart installs the CluedIn server, website, and other required services, like storage and queues.

Next steps

- For more information about CluedIn, see the CluedIn website.
- For CluedIn documentation, see the CluedIn documentation.

Feedback

Was this page helpful? Yes No

Microsoft Purview and Profisee integration for master data management

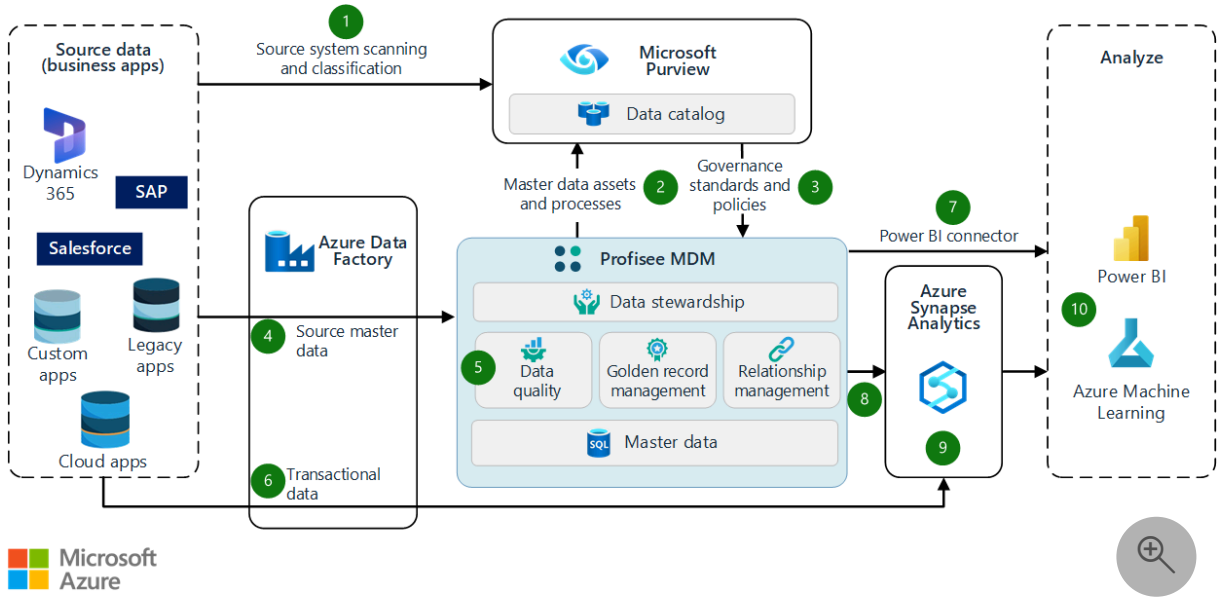
Article • 04/03/2025

Enterprise systems can have multiple sources of master data—the common data that's shared across systems. This fact can become apparent when you catalog data sources. Examples of master data include customer, product, location, asset, and vendor data. When you use Profisee to merge, validate, and correct your master data, you can make that data effective. Specifically, you can use it to build a common trusted platform for analytics and operational improvement. By using the governance definitions, insights, and expertise that are detailed in Microsoft Purview, you can build your platform effectively.

This reference architecture presents a governance and data management solution that features Microsoft Purview and the Profisee master data management (MDM) platform. These services work together to provide a foundation of high-quality, trusted data that maximizes the business value of data in Azure. For a short video about this solution, see The power of fully integrated master data management in Azure.

Architecture

The following diagram shows the steps that you take when you develop and operate your master data solution. Think of these steps as highly iterative. As your solution evolves, you might repeat these steps and phases, sometimes automatically and sometimes manually. Whether you use automatic or manual steps depends on the changes that your master data solution, metadata, and data undergo.



Dataflow

Metadata and data flow include these steps, which are shown in the preceding figure:

1. Prebuilt Microsoft Purview connectors are used to build a Unified Catalog from source business applications. The connectors scan data sources and populate Unified Catalog.

2. The master data model is published to Microsoft Purview. Master data entities that are created in Profisee MDM are seamlessly published to Microsoft Purview. This step further populates the Unified Catalog and ensures that there's a record of this critical source of data in Microsoft Purview.

3. Governance standards and policies for data stewardship are used to enrich master data entity definitions. The data is enriched in Microsoft Purview with data dictionary and glossary information, ownership data, and sensitive data classifications. Any definitions and metadata that are available in Microsoft Purview are visible in real time in Profisee as guidance for the MDM data stewards.

4. Master data from source systems is loaded into Profisee MDM. A data integration toolset like Azure Data Factory extracts data from the source systems by using any of more than 100 prebuilt connectors or a REST gateway. Multiple streams of master data are loaded into Profisee MDM.

5. The master data is standardized, matched, merged, enriched, and validated according to governance rules. Other systems, like Microsoft Purview, might define data quality and governance rules. But Profisee MDM is the system that enforces these rules. Source records are matched and merged within and across source systems to create the most complete and correct record possible. Data quality

rules check each record for compliance with business and technical requirements. Any record that fails validation or that returns a low probability score is subject to remediation. To remediate failed validations, a workflow process assigns records that require review to data stewards who are experts in their business data domain. After a record has been verified or corrected, it's ready to use as a *golden record* master.

6. Transactional data is loaded into a downstream analytics solution. A data integration toolset like Data Factory extracts transactional data from source systems by using any of more than 100 prebuilt connectors or a REST gateway. The toolset loads the data directly into an analytics data platform like Azure Synapse Analytics. Analysis on this raw information without the proper master golden data is subject to inaccuracy, because data overlaps, mismatches, and conflicts aren't yet resolved.

7. Power BI connectors provide direct access to the curated master data. Power BI users can use the master data directly in reports. A dedicated Power BI connector recognizes and enforces role-based security. It also hides various system fields to simplify use.

8. High-quality, curated master data is published to a downstream analytics solution. If master data records have been merged into a single golden record, parent-child links to the original records are preserved.

9. The analytics platform has a set of data that's certified in the sense that it's complete, consistent, and accurate. That data includes properly curated master data and associated transactional data. That combination forms a solid foundation of trusted data that's available for further analysis.

10. The high-quality master data is visualized and analyzed, and machine learning models are applied. The system delivers sound insights for driving the business.

Components

- Microsoft Purview is a data governance solution that provides broad visibility into on-premises and cloud data estates. Microsoft Purview offers a combination of data discovery and classification, lineage, metadata search and discovery, and usage insights. All these features help you manage and understand data across your enterprise data landscape.

- Profisee MDM is a fast and intuitive MDM platform that integrates seamlessly with Microsoft technologies and the Azure data management ecosystem.

- Data Factory is a hybrid data integration service. You can use Data Factory to create, schedule, and orchestrate extract, transform, and load (ETL) and extract, load, and transform (ELT) workflows. Data Factory also offers more than 100 prebuilt connectors and a REST gateway that you can use to extract data from source systems.

- Azure Synapse Analytics is a fast, flexible, and trusted cloud data warehouse that uses a massive parallel processing architecture. You can use Azure Synapse Analytics to scale, compute, and store data elastically and independently.

- Power BI is a suite of business analytics tools that delivers insights throughout your organization. You can use Power BI to connect to hundreds of data sources, simplify data preparation, and drive improvised analysis. You can also produce beautiful reports and then publish them for your organization to consume on the web and on mobile devices.

Alternatives

If you don't have a dedicated MDM application, you can find some of the technical capabilities that you need to build an MDM solution in Azure:

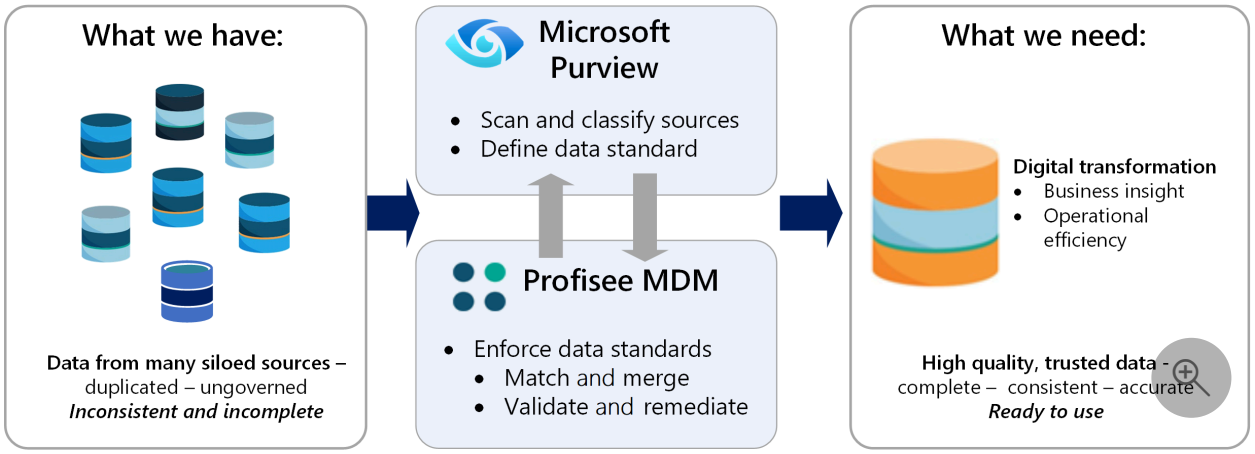
- Data quality. When you load data into an analytics platform, you can build data quality into integration processes. For example, you can use hard-coded scripts to apply data quality transformations in a Data Factory pipeline.
- Data standardization and enrichment. Azure Maps can provide data verification and standardization for address data. You can use the standardized data in Azure Functions and Data Factory. To standardize other data, you might need to develop hard-coded scripts.
- Duplicate data management. You can use Data Factory to deduplicate rows if sufficient identifiers are available for an exact match. You likely need custom hard-coded scripts to implement the logic that's needed to merge matched rows while applying appropriate data survivorship techniques.
- Data stewardship. You can use Power Apps to quickly develop basic data stewardship solutions to manage data in Azure. You can also develop appropriate user interfaces for reviews, workflows, alerts, and validations.

In Microsoft-centric environments, Azure Synapse Analytics is preferred as an analytics service. But you can use any analytics database. Snowflake and Databricks are common choices.

Scenario details

As the amount of data that you load into Azure increases, the need to properly govern and manage that data across all your data sources, and data consumers grows. Data that seems adequate in the source system is often found to be deficient when it's shared. It might have missing or incomplete information, or duplications and conflicts. Its overall quality might be poor. What's needed is data that's complete, consistent, and accurate.

Without high-quality data in your Azure data estate, the business value of Azure is undermined, perhaps critically. The solution is to build a foundation for data governance and management that can produce and deliver a source of truth for high-quality, trusted data. Microsoft Purview and Profisee MDM work together to form this enterprise platform.



Microsoft Purview catalogs all your data sources and identifies any sensitive information and lineage. It gives the data architect a place to consider the appropriate data standards to impose on all data. Microsoft Purview focuses on governance to find, classify, and define policies and standards. The tasks of enforcing policies and standards, cataloging data sources, and remediating deficient data fall to technologies like MDM systems.

Profisee MDM is designed to accept master data from any source. Profisee MDM then matches, merges, standardizes, verifies, corrects, and synchronizes the data across systems. This process ensures that data can be properly integrated and that it meets the needs of downstream systems, such as business intelligence (BI) and machine learning applications. The integrative Profisee platform enforces governance standards across multiple data silos.

Better together

Microsoft Purview and Profisee MDM work better together. When integrated, they streamline data management tasks and ensure that all systems work to enforce the same standards. Profisee MDM publishes its master data model to Microsoft Purview,

where it can participate in governance. Microsoft Purview then shares the output of governance, such as a Unified Catalog and glossary information. Profisee can review the output and enforce standards. By working jointly, Microsoft Purview and Profisee create a natural, better-together synergy that goes deeper than each independent offering.

For example, after you catalog enterprise data sources, you might determine that master data is present in multiple systems. Master data is the data that defines a domain entity. Examples of master data include customer, product, asset, location, vendor, patient, household, menu item, and ingredient data. Resolving differing definitions and matching and merging this data across systems is critical to the ability to use this data in a meaningful way. To be effective, you should merge, validate, and correct master data in Profisee MDM by using governance definitions, insights, and expertise that are detailed in Microsoft Purview. In this way, Microsoft Purview and Profisee MDM form a foundation for governance and data management, and they maximize the business value of data in Azure.

The alternative is to use whatever information you can get. But when you take this approach, you risk generating misleading results that can damage your business. When you instead use high-quality master data, you eliminate common data quality issues. Then your system delivers sound insights that you can use to drive your business, no matter which tools you use for analysis, machine learning, and visualization. Well-curated master data is a key aspect of building a solid, reliable data foundation.

When you use Profisee MDM with Microsoft Purview, you experience the following benefits:

- A common technical foundation. Profisee originated in Microsoft technologies. Profisee and Microsoft use common tools, databases, and infrastructure, which makes the Profisee solution familiar to anyone who works with Microsoft technologies. In fact, for many years, Profisee MDM was built on Microsoft Master Data Services. Now Master Data Services is nearing the end of its lifecycle, and Profisee is the premier upgrade and replacement solution.
- Developer collaboration and joint development. Profisee and Microsoft Purview developers collaborate extensively to ensure a good, complementary fit between their respective solutions. This collaboration delivers a seamless integration that meets customer needs.
- Joint sales and deployments. Profisee has more MDM deployments on Azure, and jointly with Microsoft Purview, than any other MDM vendor. You can purchase Profisee through Azure Marketplace. In fiscal year 2023, Profisee is the only MDM vendor with a top-tier Microsoft partner certification that has an infrastructure as a service (IaaS), containers as a service (CaaS), or software as a service (SaaS) offering on Azure Marketplace.

- Rapid and reliable deployment. A critical feature of all enterprise software is rapid and reliable deployment. According to the Gartner Peer Insights platform, Profisee has more implementations that take fewer than 90 days to complete than any other MDM vendor.
- Multiple domains. Profisee offers an approach to MDM that inherently uses multiple domains. There are no limitations to the number of master data domains that you can create. This design aligns well with customers who plan to modernize their data estates. A customer might start with a limited number of domains, but they ultimately benefit from maximizing their domain coverage across their whole data estate. This domain coverage is matched to their data governance coverage.
- Engineering that's designed for Azure. Profisee is engineered to be cloud native with options for SaaS and managed IaaS or CaaS deployments on Azure.

Potential use cases

For a detailed list of MDM use cases of this solution, see MDM use cases later in this article. Key MDM use cases include the following retail and manufacturing examples:

- Consolidating customer data for analytics.
- Having a 360-degree view of product data in a consistent and accessible form, such as each product's name, description, and characteristics.
- Establishing reference data to consistently augment descriptions of master data. For example, reference data includes lists of countries/regions, currencies, colors, sizes, and units of measure.

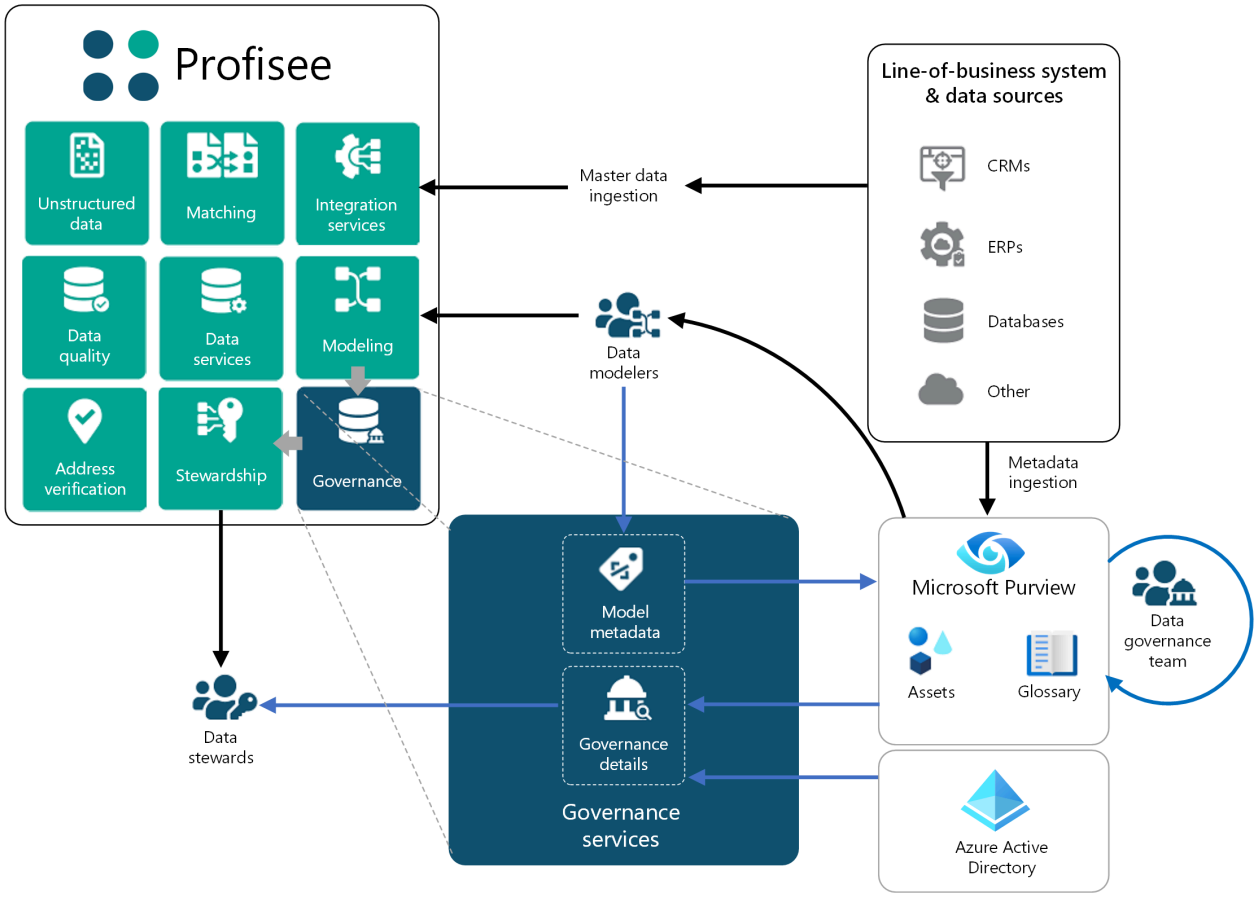
These MDM solutions also help financial organizations that rely heavily on data for critical activities, such as timely reporting.

MDM integration with Microsoft Purview

The following diagram illustrates in detail the integration of Profisee MDM in Microsoft Purview. To support this integration, the Profisee governance subsystem provides bidirectional integration with Microsoft Purview, which consists of two distinct flows:

- Solution metadata publishing occurs when your data modelers make changes to your master data model, matching strategies, and their related subartifacts. These changes are seamlessly published to Microsoft Purview as they occur. Publishing these changes syncs the metadata that's related to your master data model and solution. As a result, Unified Catalog is further populated, and Microsoft Purview has a record of this critical data source.

- Governance details are returned and provided to data stewards and business users. These details are available as the users view data, enrich data, and remediate data quality issues by using the Profisee FastApp portal.



Microsoft Purview integration capabilities

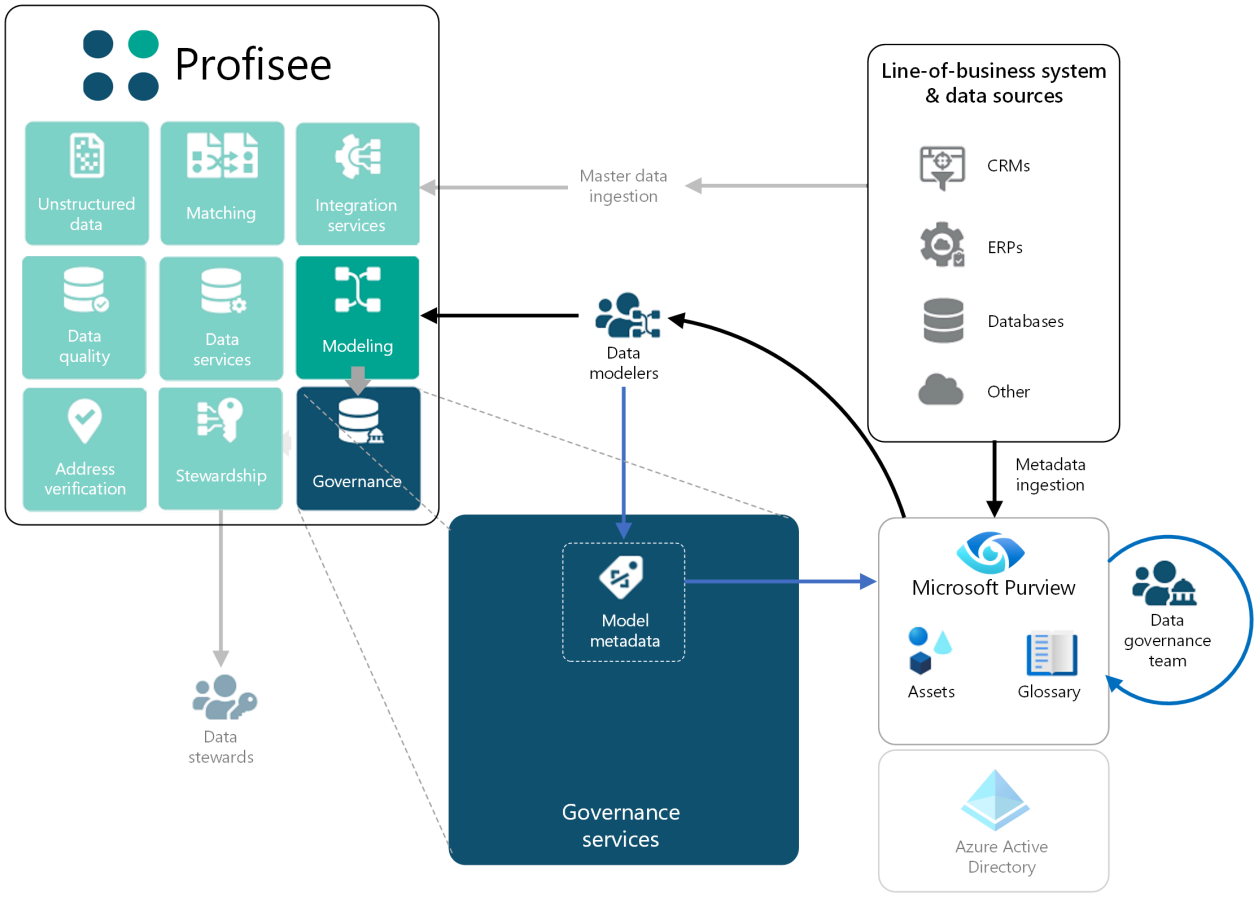
The Microsoft Purview catalog and glossary can help you maximize integration.

Master data model design

One of the challenges of preparing an MDM solution is determining what constitutes master data and which data sources to use when you populate your master data model. You can use Microsoft Purview to help with this effort. You can take advantage of the ability to scan your critical data sources, and you can engage your data stewards and subject matter experts (SMEs). This way, you can enrich Unified Catalog with information that your stewards can then access, to better align your master data model with your line-of-business systems. You can reconcile conflicting terminology. This process yields a master data model that optimally reflects the terminology and definitions that you want to standardize for your business. It also avoids outdated and misleading verbiage.

The following excerpt from the broader diagram illustrates this integration use case. First, you use Microsoft Purview system scanning functions to ingest metadata from your line-of-business systems. Next, your data stewards and SMEs prepare a solid

catalog and contacts. Then the data modelers who work with Profisee MDM modeling services can prepare and evolve your master data model. This work aligns with the standards that you define in Microsoft Purview.



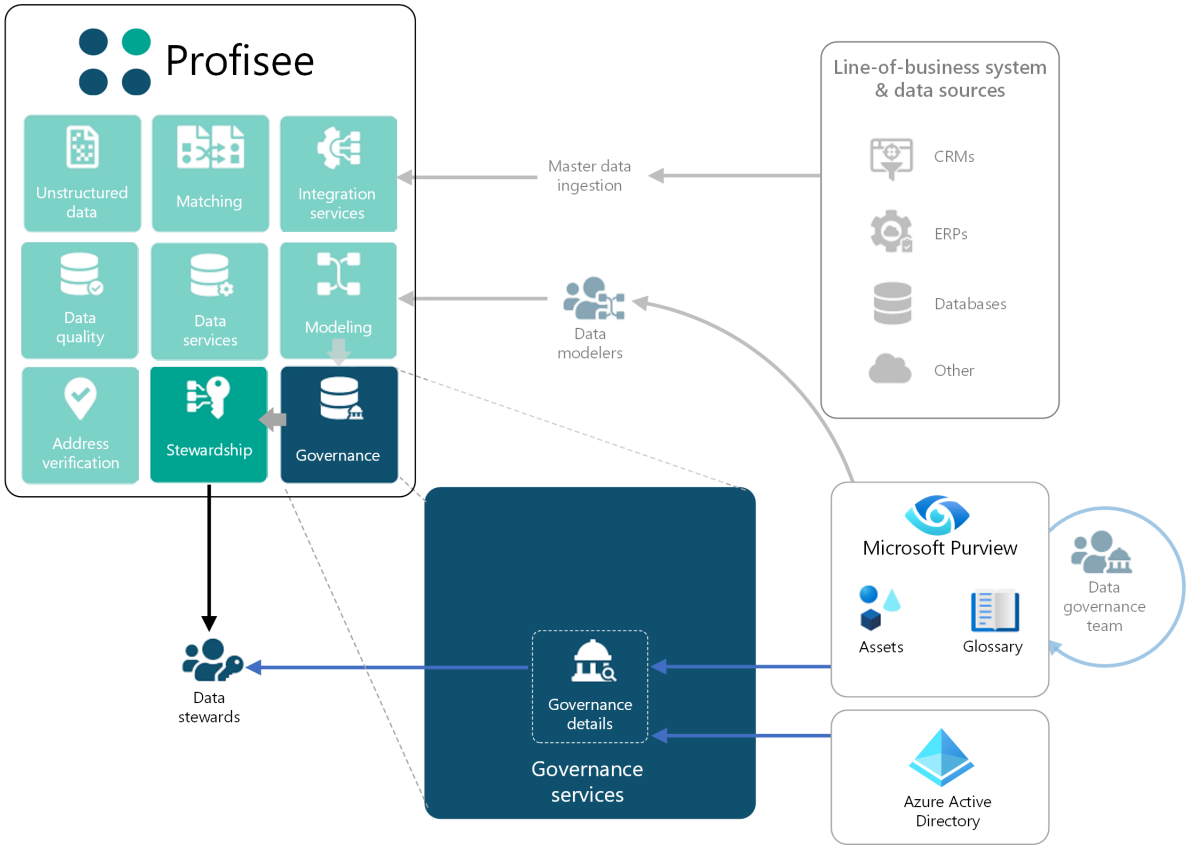
As your data stewards evolve the model, the modeling services within the Profisee MDM platform publish changes that Profisee MDM governance services receive. In turn, Profisee MDM prepares and forwards those changes to Microsoft Purview for inclusion in its updated Unified Catalog. These additions to the catalog ensure that your master data definitions are included in the broader data estate and that they can be governed and controlled in the same manner as your line-of-business system metadata. By ensuring that this information is cataloged together, you're in a better position to manage the relationships between your master data and your line-of-business system data.

Data stewardship

Large enterprises that have correspondingly complex and expansive data estates can present challenges to data stewards, who are responsible for managing and remediating issues as they arise. Key data domains can be complex, with many obscure attributes that only tenured employees who have significant institutional knowledge understand. Through the Profisee MDM integration with Microsoft Purview, you can capture this institutional knowledge within Microsoft Purview and make it available for use within

Profisee MDM. As a result, you alleviate a great need for corporate data knowledge when you manage critical and time-sensitive information.

The following figure illustrates the flow of information from Microsoft Purview to the data stewards who work in the Profisee FastApp portal. The governance data service integrates with Microsoft Purview and Microsoft Entra ID. This service provides lookup functionality. FastApp portal users can use this functionality to retrieve enriched governance data about the entities and the attributes that they work with.



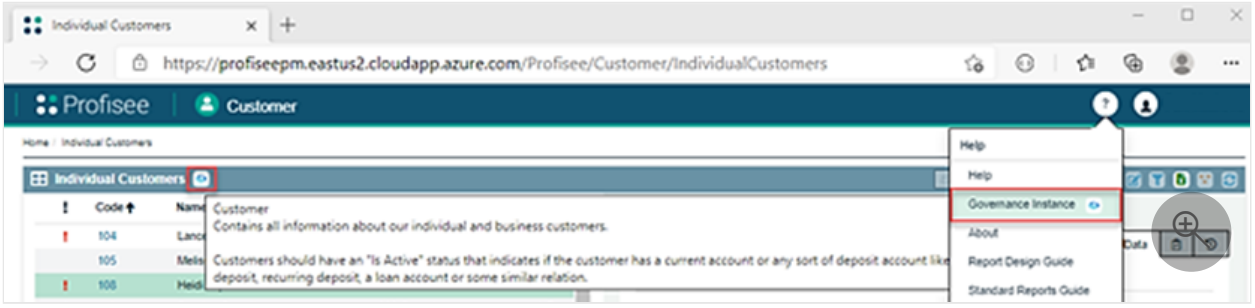
Governance services also resolve contacts that are received from Microsoft Purview to their full profile details, which are available in Microsoft Entra ID. With complete profile details, stewards can effectively collaborate with data owners and experts as they work to enhance the quality of your master data.

The Profisee MDM Governance dialog is the interface through which data stewards and users interact with governance-level details. This UI renders information that's obtained from Microsoft Purview to users. By using this information, users can review the details behind the data from which the dialog was launched. If the information that's provided in the Governance dialog is insufficient, users can go directly to the Microsoft Purview user experience.

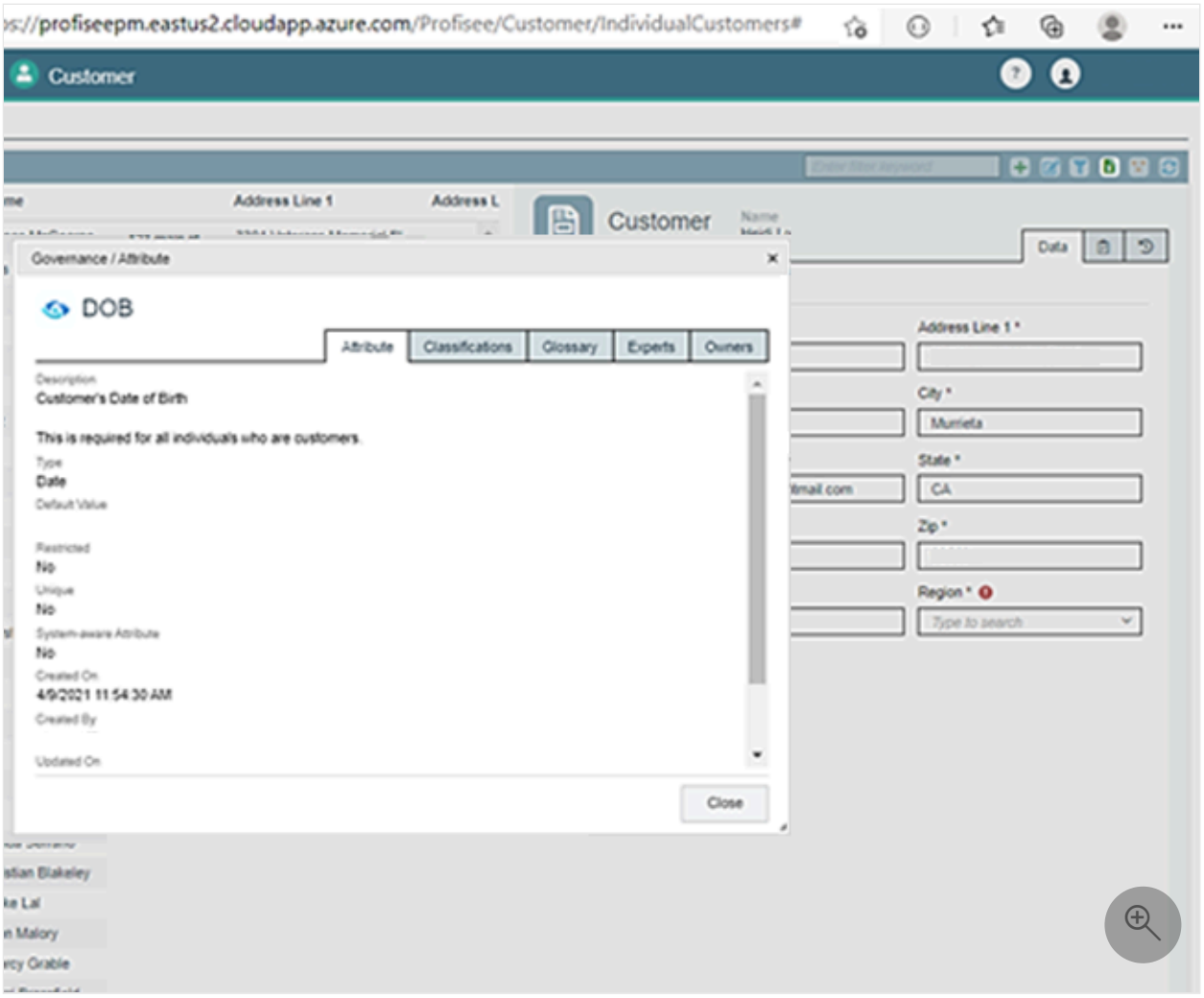
Data stewards and business users can access three Profisee MDM data asset types via the FastApp portal:

- Profisee Instance, which provides the infrastructure properties of the specific instance of the Profisee MDM platform that the user is viewing
- Profisee Entity, which provides the properties of the master data entity (the table) that the steward or user is currently viewing
- Profisee Attribute, which provides the properties of the attribute (such as the field or column) in which the user is interested

The following figure illustrates where users who are working in the FastApp portal can view governance details for each of these asset types. You can find instance-level details on the **Help** menu. You can access entity details from the page zone header, which contains an entity grid. For attribute details, go to the form that's associated with the entity grid. Access the details from the labels that are associated with the attribute.



To see summary information, hover over the governance icon, such as Microsoft Purview. Select the icon to display the full governance dialog:



To go to the full Microsoft Purview user experience, select the governance icon in the dialog header. Selecting the icon takes you to Microsoft Purview in the context of the asset that you're currently viewing. You then can easily move around in Microsoft Purview based on your discovery needs.

MDM processing

The power of an MDM solution is in the details.

Data modeling

The heart of your MDM solution is the underlying data model. It represents the definition of *master data* within your company. Developing a master data model involves the following tasks:

- Identify elements of source data from across your systems landscape that are critical to your company's operations and central to analyzing performance.
- Enrich the model with elements that you obtain from other third-party sources that render the data more useful, accurate, and trustworthy.

- Establish clear ownership and permissions related to the elements of your data model. This practice helps ensure that you factor visibility and change management into your model's design.

Data governance provides a critical foundation of support:

- Your governance Unified Catalog, dictionary, glossary, and supporting resources are invaluable sources of information to your governance data stewards. These resources help stewards determine what to include in your master data model. They also help determine ownership and sensitive data classifications in Microsoft Purview. You can reinforce terminology in your model. Through this practice, you can establish an official lexicon for your business. By integrating terminology, your master data model can also translate any esoteric terms that are in use in various source systems to the approved language of the business.
- Third-party systems are often a source of master data that's separate and apart from your line-of-business systems. It's critical to add elements to your model to capture the information that these systems add to your data, and to reflect these sources of information back into Unified Catalog.
- You can use ownership and data access, as identified in your governance catalog, to enforce access and change management permissions within your MDM solution. As a result, you align your corporate policies and needs with the tools that you use to manage and steward your master data.

Source data load

Ideally, your disparate line-of-business systems load data into your master data model with little or no change or transformation. The goal is to have a centralized version of the data as it exists in the source system. There should be as little loss of fidelity as possible between the source system and your master data repository. By limiting the complexity of your loading process, you make lineage simpler. And when you use technology like Data Factory pipelines, your governance solution can inspect the flow. Then your solution can identify the relationships between your source system and your master data model. Specifically, your solution can extract data from source systems by using any of more than 100 prebuilt connectors and a REST gateway.

Data enrichment and standardization

After you load source data into your model, you can extend it by tapping into rich sources of third-party data. You can use these systems to improve the data that you obtain from your line-of-business systems. You can also use these systems to augment

the source data with information that enhances its use for other downstream consumers. For example:

- You can use address-verification services like Bing to correct and improve source system addresses. These services can standardize and add missing information that's crucial to geo-location and mail delivery.
- Third-party information services like Dun & Bradstreet can provide general-purpose or industry-specific data. You can use this data to extend the value of your golden master record. Specifically, you might add information that was unavailable or in conflict in your disparate line-of-business systems.

Profisee's publish/subscribe infrastructure makes it easy to integrate your own third-party sources into your solution as needed.

The ability to understand the sources and meaning behind this data is as critical for third-party data as it is for your internal line-of-business systems. By integrating your master data model into your governance Unified Catalog, you can manage the relationships between internal and external sources of data while enriching your model with governance details.

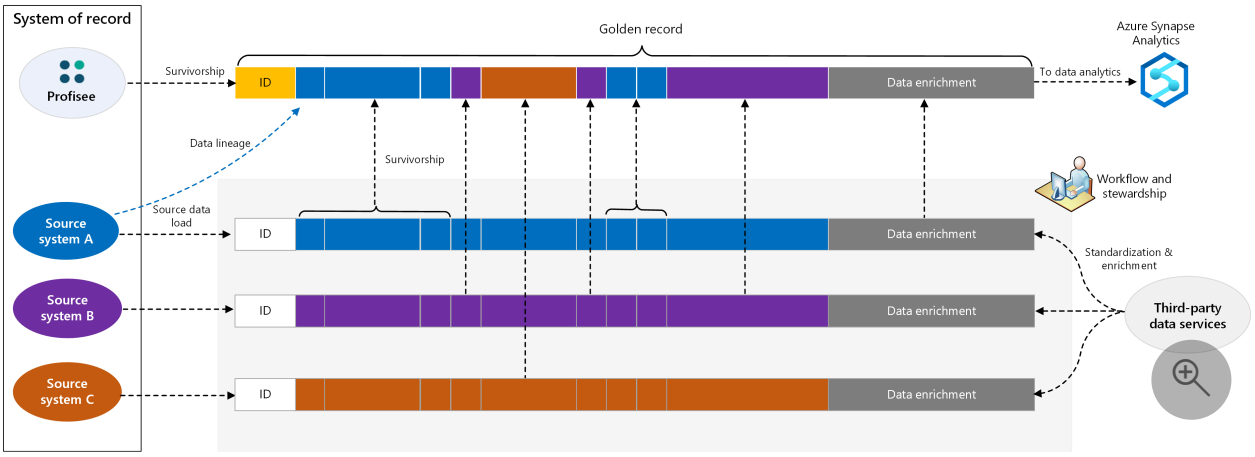
Data quality validation and stewardship

After you load and enrich your data, it's important to check it for quality and adherence to standards that you establish through your governance processes. Microsoft Purview can again be a rich source of standards information. You can use Microsoft Purview to drive the data quality rules that your MDM solution enforces. Profisee MDM can also publish data quality rules as assets to your governance catalog. The rules can be subject to review and approval, which helps you provide top-down oversight of quality standards that are associated with your master data. Your rules are tied to master data entities and attributes, and those attributes can be traced back to the source system. For these reasons, you can establish the root cause of the poor data quality that originates from your line-of-business systems.

Data stewards are experts in their governance domain. As stewards address issues that your master data solution reveals, they can use the Microsoft Purview data governance catalog. The catalog helps stewards understand and resolve quality issues as they arise. Backed by the support of data owners and experts, the stewards are prepared to address data quality issues quickly and accurately.

Matching and survivorship

With enriched, high-quality source data, you're positioned to produce a golden record master that represents the most accurate information across your disparate line-of-business systems. The following figure illustrates how all the steps culminate in high-quality data that's ready to use for business analysis. At any time, you can sync this data across your data estate.



The Profisee MDM matching engine produces a golden record master as part of the survivorship process. Survivorship rules selectively populate the golden record with information that you've chosen across all your source systems.

The Profisee MDM history and audit tracking subsystem tracks changes that users make. This subsystem also tracks changes that system processes like survivorship make. Matching and survivorship make it possible to trace the flow of information from your source records to the master. Profisee MDM has a record of the source system that's responsible for a specific source record. You also know how disparate source records populate the golden record. As a result, you can achieve data lineage from your analytics back to the source data that your reports reference.

MDM use cases

Although there are numerous use cases for MDM, a few use cases cover most real-world MDM implementations. These use cases focus on a single domain, but they're unlikely to be built from only that domain. Even these focused use cases most likely involve multiple domains. In each use case, MDM meets the goal of providing a 360 degree, or unified, view of essential data types.

Customer data

Consolidating and standardizing customer data for BI analytics is the most common MDM use case. Organizations capture customer data across an increasing number of systems and applications. Duplicate customer data records result. These duplicates are

located in and across applications, and they contain inconsistencies and discrepancies. The poor quality of the customer data limits the value of modern analytics solutions. Symptoms include the following challenges:

- It's hard to answer basic business questions like, "Who are our top customers?" and "How many new customers do we have?" Answering these questions requires significant manual effort.
- You have missing and inaccurate customer information, which makes it difficult to roll up or drill down into the data.
- You're unable to uniquely identify or verify a customer across organizational and system boundaries. As a result, you're unable to analyze your customer data across systems or business units.
- You have poor-quality insights from AI and machine learning due to the poor-quality input data.

Product data

Product data is often spread across multiple enterprise applications, such as enterprise resource planning (ERP), product lifecycle management (PLM), or e-commerce applications. As a result, it's challenging to understand the total catalog of products that have inconsistent definitions for properties, such as the product name, description, and characteristics. Different definitions of reference data complicate this situation. Symptoms include the following challenges:

- You're unable to support different alternative hierarchical roll-up and drill-down paths for product analytics.
- With finished goods or material inventory, you have difficulty evaluating product inventory and established vendors. You also have duplicate products, which leads to excess inventory.
- It's hard to rationalize products due to conflicting definitions. This situation leads to missing or inaccurate information in analytics.

Reference data

In the context of analytics, reference data exists as numerous lists of data. These lists are often used to further describe other sets of master data. For example, reference data includes lists of countries/regions, currencies, colors, sizes, and units of measure. Inconsistent reference data leads to obvious errors in downstream analytics. Symptoms are:

- Multiple representations of the same value. For example, the state of Georgia is listed as *GA* and *Georgia*, which makes it difficult to consistently aggregate and

drill down into data.
- Difficulty streamlining data across systems due to an inability to *crosswalk*, or map, reference data values between systems. For example, the color red is represented by *R* in the ERP system and *Red* in the PLM system.
- Difficulty tying numbers across organizations due to differences in established reference data values that are used for data categorization.

Financial data

Financial organizations rely heavily on data for critical activities, such as monthly, quarterly, and annual reporting. Organizations that have multiple finance and accounting systems often have financial data across multiple general ledgers that needs to be consolidated to produce financial reports. MDM can provide a centralized hub to map and manage accounts, cost centers, business entities, and other financial datasets. Through the centralized hub, MDM provides a consolidated view of these datasets. Symptoms include the following challenges:

- Difficulty aggregating financial data across multiple systems into a consolidated view
- Lack of process for adding and mapping new data elements in financial systems
- Delays in producing end-of-period financial reports

Considerations

These considerations implement the pillars of the Azure Well-Architected Framework, which is a set of guiding tenets that can be used to improve the quality of a workload. For more information, see Microsoft Azure Well-Architected Framework.

Consider these factors when you choose a data management solution for your organization.

Reliability

Reliability ensures your application can meet the commitments you make to your customers. For more information, see Overview of the reliability pillar.

Profisee runs natively on Azure Kubernetes Service (AKS) and Azure SQL Database. Both services offer out-of-the-box capabilities to support high availability.

Security

Security provides assurances against deliberate attacks and the abuse of your valuable data and systems. For more information, see Overview of the security pillar.

Profisee authenticates users by using OpenID Connect, which implements an OAuth 2.0 authentication flow. Most organizations configure Profisee MDM to authenticate users against Microsoft Entra ID, which ensures that you can apply and enforce your enterprise policies for authentication.

Cost optimization

Cost optimization is about looking at ways to reduce unnecessary expenses and improve operational efficiencies. For more information, see Overview of the cost optimization pillar.

Running costs consist of a software license and Azure consumption. For more information, contact Profisee.

Performance efficiency

Performance efficiency is the ability of your workload to scale to meet the demands placed on it by users in an efficient manner. For more information, see Performance efficiency pillar overview.

Profisee MDM runs natively on AKS and SQL Database. You can configure AKS to scale Profisee MDM up, down, and across your business functions. You can deploy SQL Database in numerous configurations to balance performance, scalability, and costs.

Dynamic scaling is inherent in the cloud-native architecture of Profisee, which uses microservices and containers. If you run Profisee in your cloud tenant via Kubernetes, you can dynamically scale up and out based on your load. With the Profisee SaaS service that runs on AKS, you can configure large node pools for your pods. These pools scale dynamically based on the load on the system across the multitenant infrastructure.

For detailed information about how to deploy Profisee and Microsoft Purview on AKS, see Microsoft Purview - Profisee MDM integration.

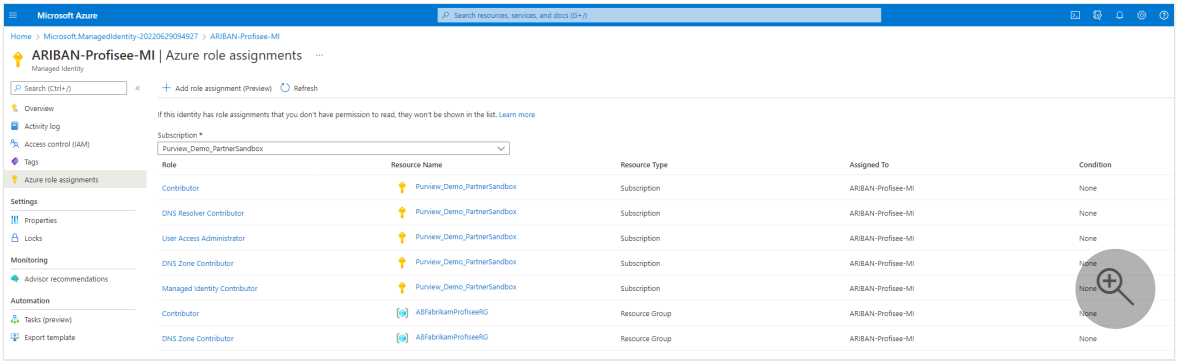
Deploy this scenario

Profisee MDM is a packaged Kubernetes service. You can deploy Profisee MDM as a PaaS in your Azure tenant, in any other cloud tenant, or on-premises. You can also deploy Profisee MDM as a SaaS that Profisee hosts and manages.

1. Get the license file from Profisee by raising a support ticket on https://support.profisee.com/. The only pre-requisite for this step is your need to pre-determine the DNS resolved URL your Profisee setup on Azure. In other words, keep the DNS HOST NAME of the load balancer used in the deployment. It will be something like "[profisee_name].[region].cloudapp.azure.com". For example, DNSHOSTNAME="purviewprofisee.southcentralus.cloudapp.azure.com". Supply this DNSHOSTNAME to Profisee support when you raise the support ticket and Profisee will revert with the license file. You'll need to supply this file during the next configuration steps below.

2. Create a user-assigned managed identity in Azure. You must have a managed identity created to run the deployment. After the deployment is done, the managed identity can be deleted. Based on your ARM template choices, you'll need some or all of the following roles and permissions assigned to your managed identity:

    - Contributor role to the resource group where AKS will be deployed. It can either be assigned directly to the resource group **OR** at the subscription level and down.
    - DNS Zone Contributor role to the particular DNS zone where the entry will be created **OR** Contributor role to the DNS Zone resource group. This DNS role is needed only if updating DNS hosted in Azure.
    - Application Administrator role in Microsoft Entra ID so the required permissions that are needed for the application registration can be assigned.
    - Managed Identity Contributor and User Access Administrator at the subscription level. Required in order for the ARM template managed identity to be able to create a Key Vault specific managed identity that will be used by Profisee to pull the values stored in the Key Vault.



3. Create an application registration that will act as the login identity once Profisee is installed. It needs to be a part of the Microsoft Entra ID that will be used to sign in to Profisee. Save the **Application (client) ID** for use later.

    - Set authentication to match the settings below:

    - Support ID tokens (used for implicit and hybrid flows)
    - Set the redirect URL to: https://<your-deployment-url>/profisee/auth/signin-microsoft
        - Your deployment URL is the URL you'll have provided Profisee in step 1

4. Create a service principal that Microsoft Purview will use to take some actions on itself during this Profisee deployment. To create a service principal, create an application like you did in the previous step, then create an application secret. Save the **Object ID** for the application, and the **Value** of the secret you created for later use.

    - Give this service principal (using the name or Object ID to locate it) **Data Curator** permissions on the root collection of your Microsoft Purview account.

5. Go to https://github.com/Profisee/kubernetes and select Microsoft Purview **Azure ARM**.

    - The ARM template will deploy Profisee on a load balanced AKS (Azure Kubernetes Service) infrastructure using an ingress controller.
    - The readme includes troubleshooting steps.
    - Read all the steps and troubleshooting wiki page carefully.

6. Select "Deploy to Azure"

    - The configurator wizard will ask for the inputs as described here - Deploying the AKS Cluster using the ARM Template
    - Make sure to give the exact same RG (Resource Group) in the deployment as you gave permissions to the managed identity in Step1.

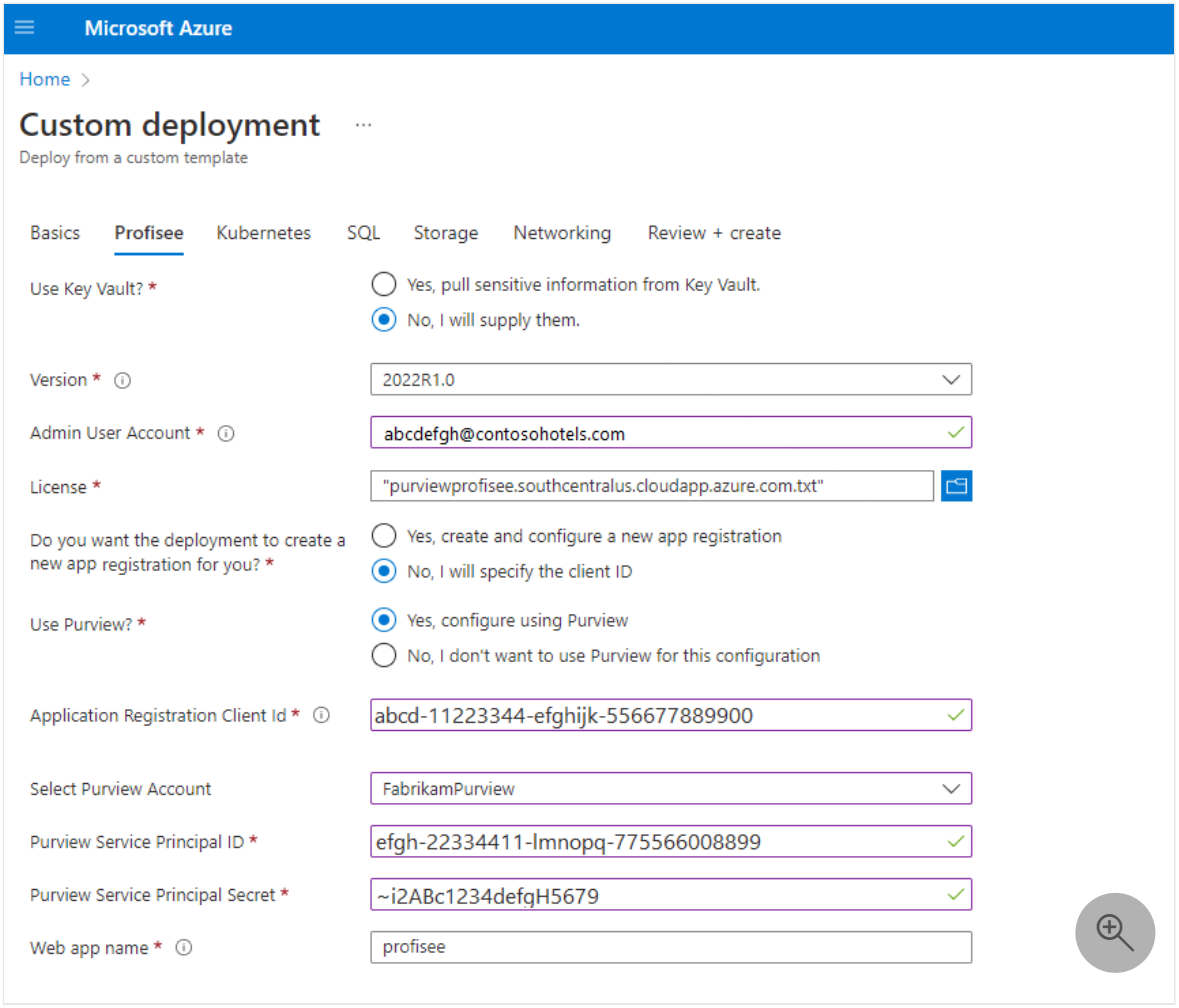
7. Once deployment completes, select Microsoft Purview "Go to Resource Group" and open the Profisee AKS Cluster.

Stages of a typical Microsoft Purview - Profisee deployment run

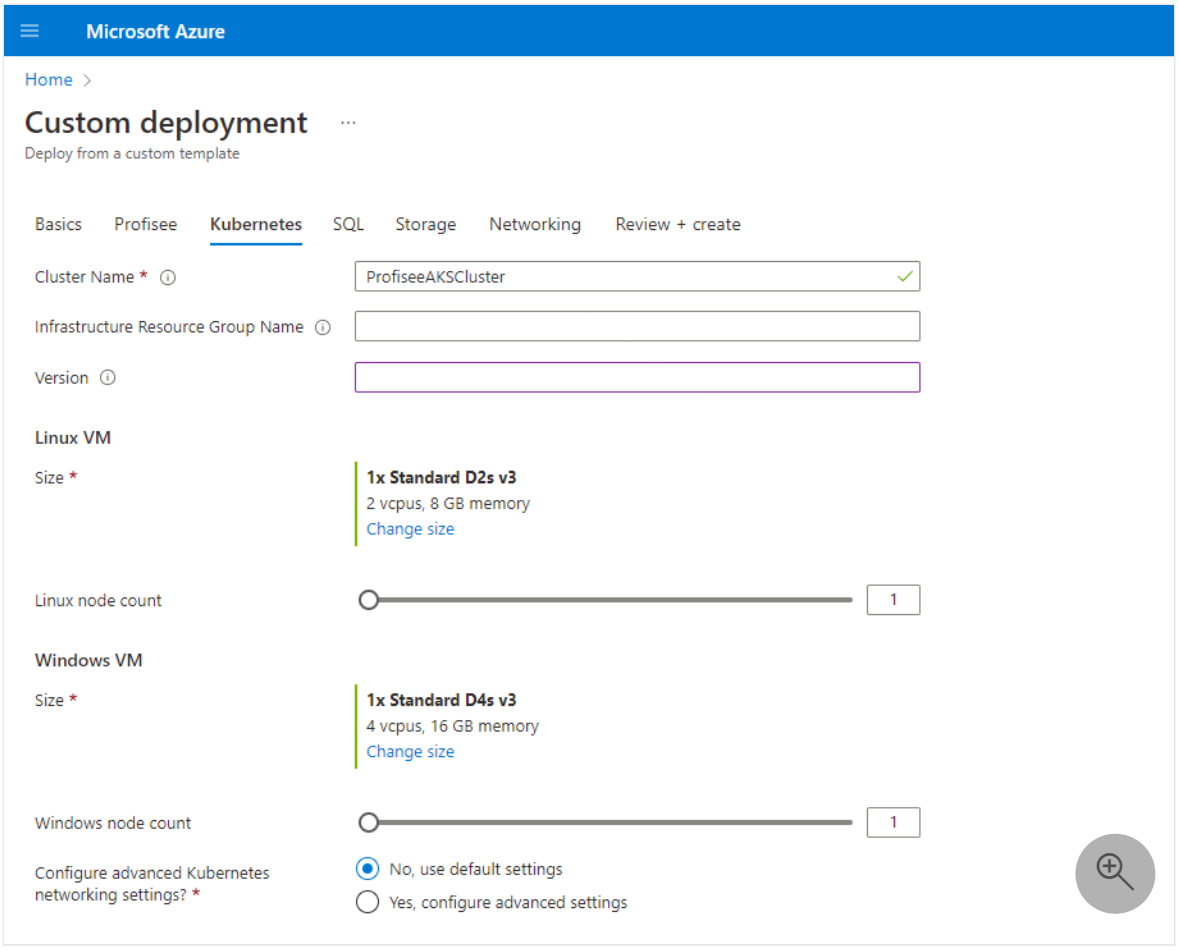
1. On the basics page, select the user-assigned managed identity you created earlier to deploy the resources.

2. For your Profisee configuration, you can have your information stored in Key Vault or supply the details during deployment.

a. Choose your Profisee version, and provide your admin user account and license.
b. Select to configure using Microsoft Purview.
c. For the Application Registration Client ID, provide the **application (client) ID** for the application registration you created earlier.
d. Select your Microsoft Purview account.
e. Add the **object ID** for the service principal you created earlier.
f. Add the value for the secret you created for that service principal.
g. Give your web application a name.

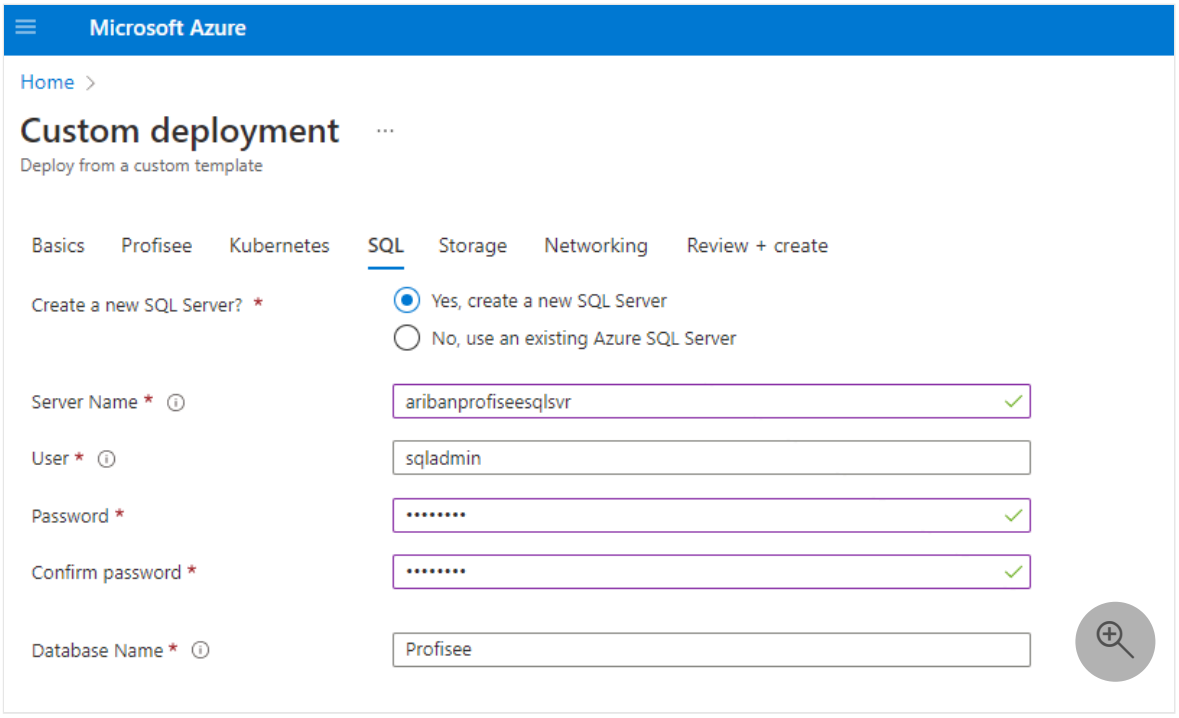


3. On the Kubernetes page, you may choose an older version of Kubernetes if needed, but leave the field **blank** to deploy the **latest** version.

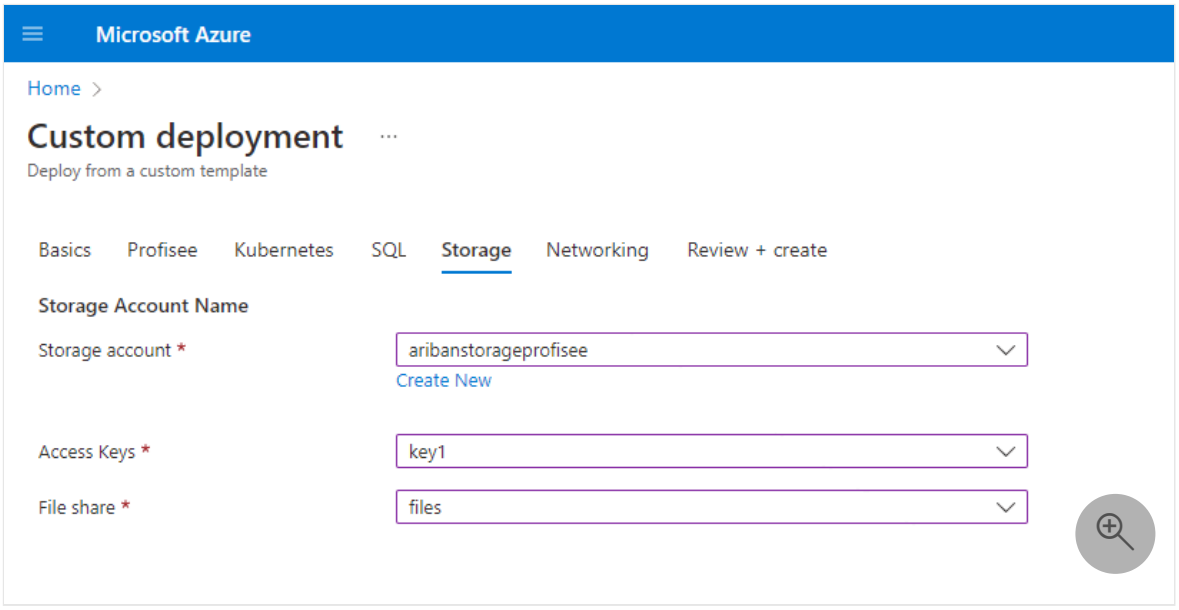


> **Tip**
>
> In most cases, leaving the version field blank is sufficient, unless there is a reason you need to deploy using an older version of Kubernetes AKS specifically.

4. On the SQL configuration page you can choose to deploy a new Azure SQL server, or use an existing Azure SQL Server. You'll provide login details and a database name to use for this deployment.

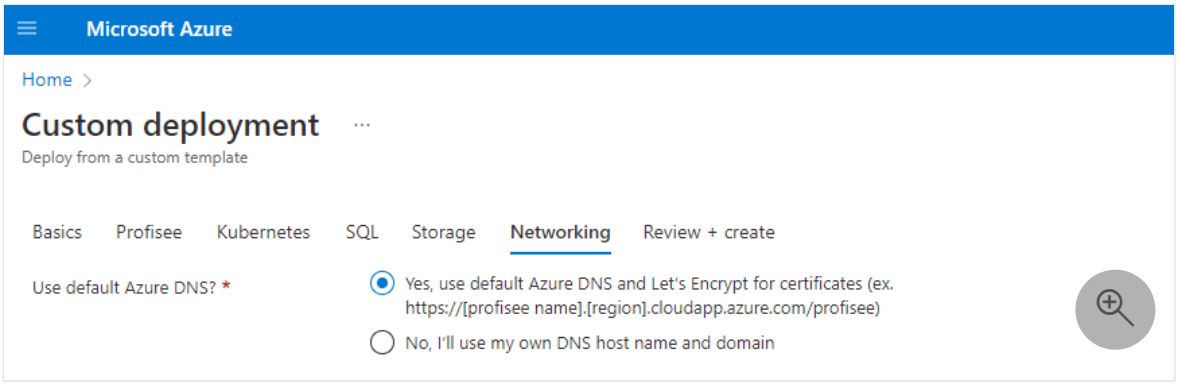


5. On the storage configuration page, you can choose to create a new storage account or use an existing one. You'll need to provide an access key and the name of an existing file share if you choose an existing account.



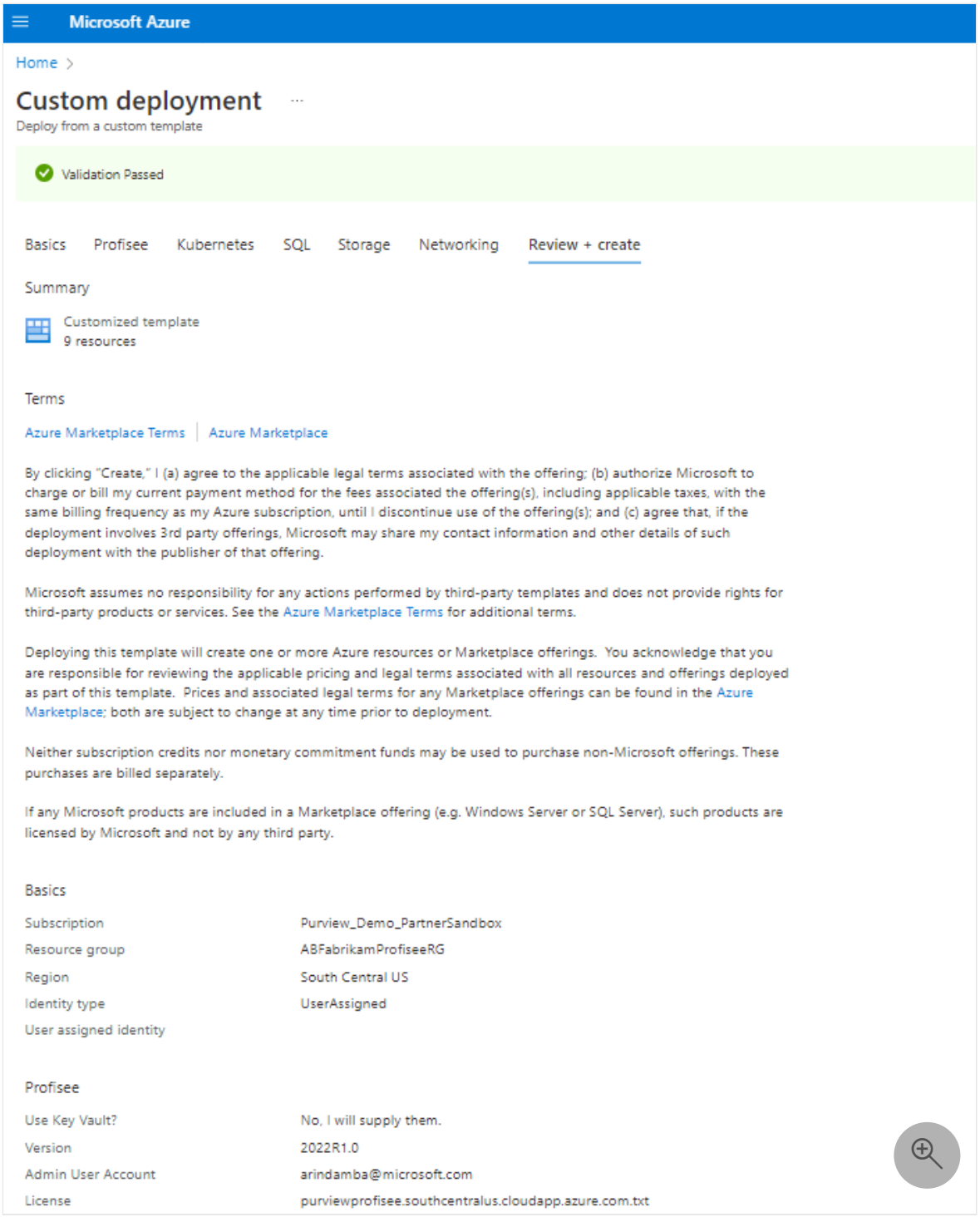
6. On the networking configuration page, you'll choose to either use the default Azure DNS or provide your own DNS host name.

> **Tip**
>
> **Yes, use default Azure DNS** is the recommended configuration. Choosing Yes, the deployer automatically creates a Let's Encrypt certificate for HTTP/TLS. If you choose **No** you'll need to supply various networking configuration parameters and your own HTTPS/TLS certificate.

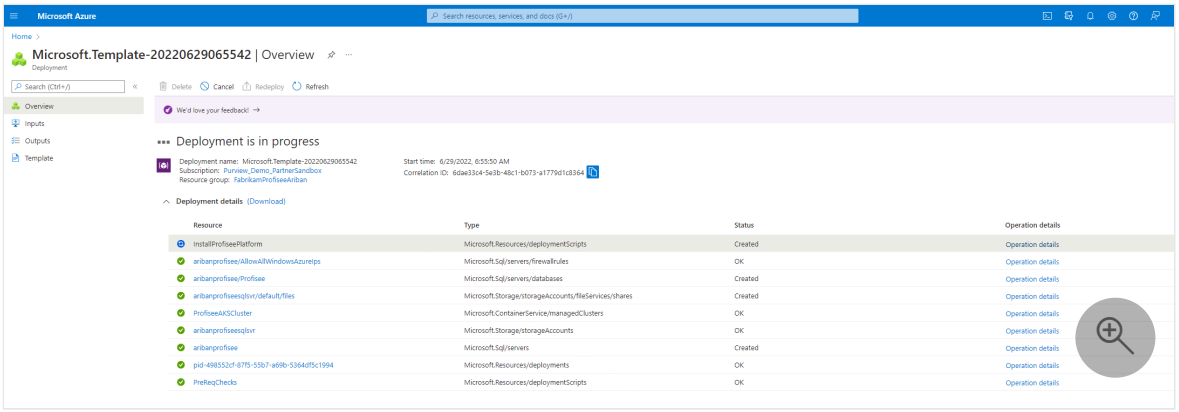


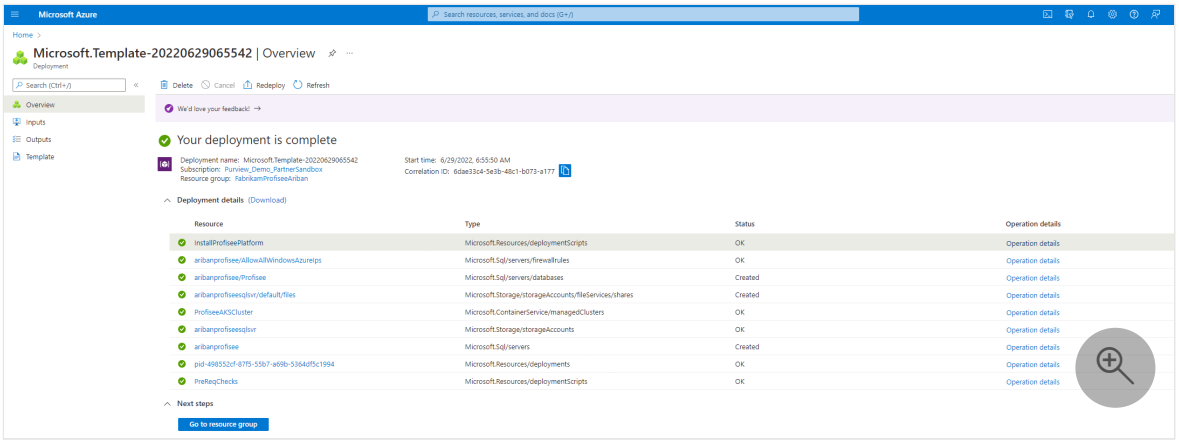
> ⚠ **Warning**
>
> The default Azure DNS URL (for example URL="https://purviewprofisee.southcentralus.cloudapp.azure.com/profisee") will be picked up by thr ARM template deployment wizard from the license file supplied to you by Profisee. If you intend to make changes and not use the default Azure DNS, make sure to communicate the full DNS and the fully qualified URL of the Profisee DNS to the Profisee support team so that they can re-generate and provide you the updated license file. Failure to do this will result in a malfunctioning installation of Profisee.

7. On the review + create page, review your details to ensure they're correct while the wizard validates your configuration. Once validation passes, select **Create**.

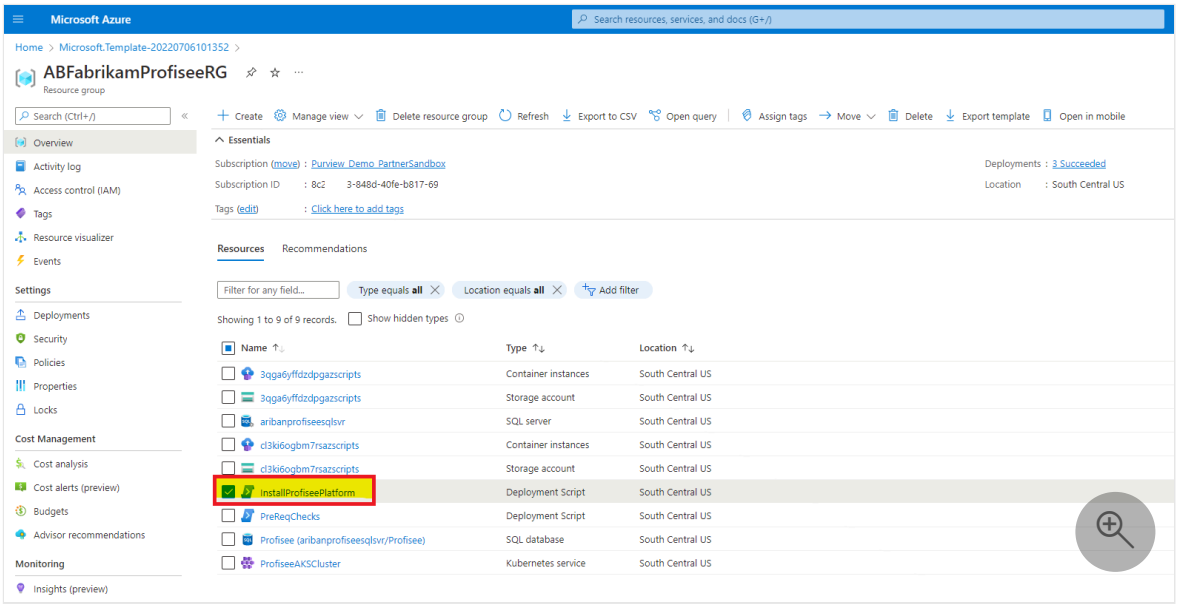


8. It will take around 45-50 Minutes for deployment to complete installing Profisee. During the deployment, you'll see the aspects that are in progress, and can refresh the page to review progress. The deployment will show as complete when all is finished. Completion of "InstallProfiseePlatform" stage also indicates deployment is complete!

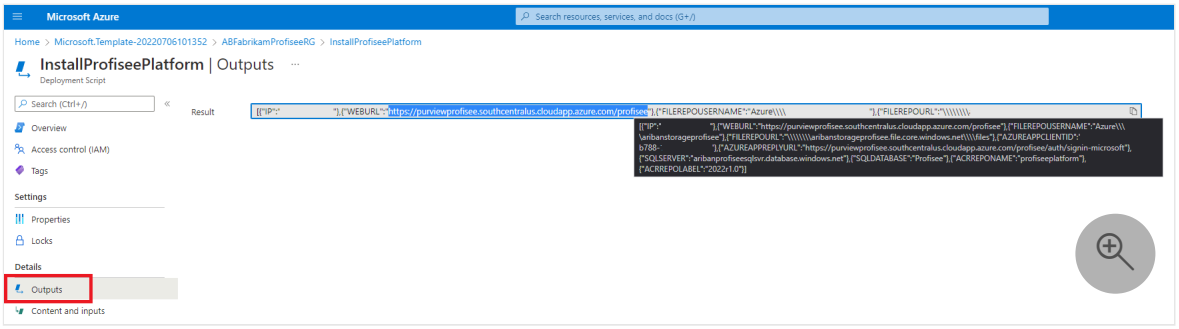




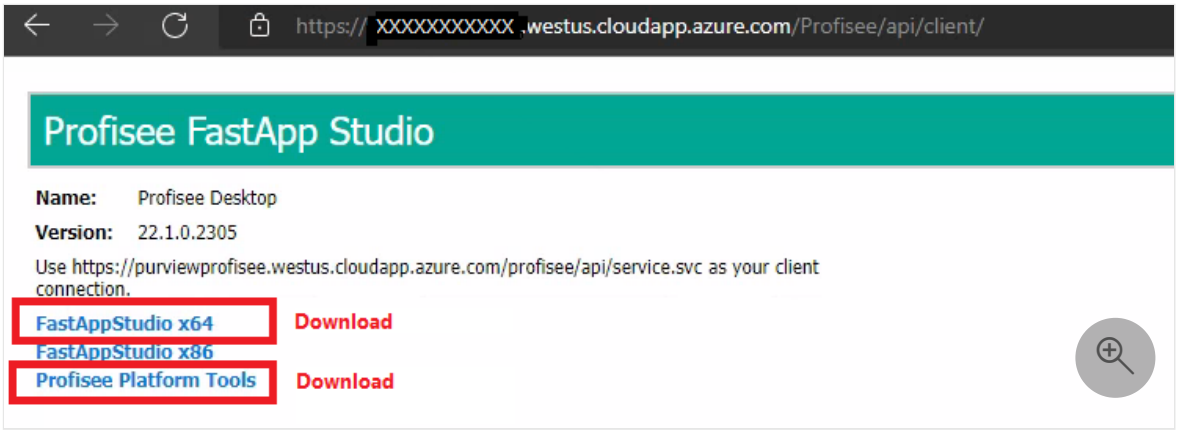
9. Once deployment is completed, open the resource group where you deployed your integration.



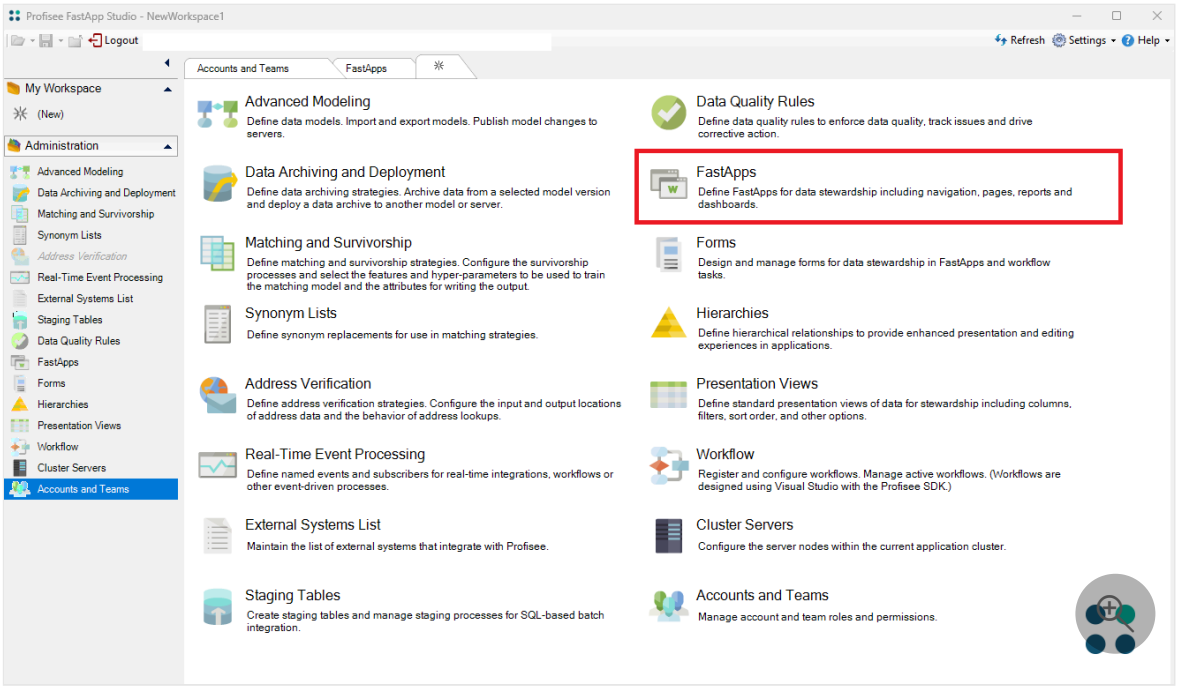
10. Under outputs, fetch the final deployment URL. The final WEBURL is what you need to paste on your browser address bar and start enjoying Profisee-Purview integration! This URL will be the same that you'd have supplied to Profisee support while obtaining the license file. Unless you chose to change the URL format, it will look something like - "https://[profisee_name].[region].cloudapp.azure.com/profisee/



11. Populate and hydrate data to the newly installed Profisee environment by installing FastApp. Go to your Profisee deployment URL and select **/Profisee/api/client**. It should look something like - "https://[profisee_name].[region].cloudapp.azure.com/profisee/api/client". Select the Downloads for "Profisee FastApp Studio" utility and the "Profisee Platform Tools". Install both these tools on your local client machine.



12. Log in to FastApp Studio and perform the rest of the MDM Administration and configuration management for Profisee. Once you log in with the administrator email address supplied during the setup; you should be able to see the administration menu on the left pane of the Profisee FastApp Studio. Navigate to these menus and perform the rest of your MDM journey using FastApp tool. Being able to see the administration menu as seen in the image below confirms successful installation of Profisee on Azure Platform.



13. As a final validation step to ensure successful installation and for checking whether Profisee has been successfully connected to your Microsoft Purview instance, go to **/Profisee/api/governance/health** It should look something like - "https://[profisee_name].[region].cloudapp.azure.com//Profisee/api/governance/health". The output response will indicate the words "**Status**": "**Healthy**" on all the Microsoft Purview subsystems.

JSON


```
{
  "OverallStatus": "Healthy",
  "TotalCheckDuration": "0:XXXXXXX",
  "DependencyHealthChecks": {
    "purview_service_health_check": {
      "Status": "Healthy",
      "Duration": "00:00:NNNN",
      "Description": "Successfully connected to Purview."
    },
    "governance_service_health_check": {
      "Status": "Healthy",
      "Duration": "00:00:NNNN",
      "Description": "Purview cache loaded successfully.
      Total assets: NNN; Instances: 1; Entities: NNN; Attributes: NNN;
Relationships: NNN; Hierarchies: NNN"
    },
    "messaging_db_health_check": {
      "Status": "Healthy",
      "Duration": "00:00:NNNN",
      "Description": null
    },
    "logging_db_health_check": {
      "Status": "Healthy",
```

```
      "Duration": "00:00:NNNN",
      "Description": null
    }
  }
}
```

An output response that looks similar as the above confirms successful installation, completes all the deployment steps; and validates whether Profisee has been successfully connected to your Microsoft Purview and indicates that the two systems are able to communicate properly.

Next steps

- Understand the capabilities of the REST copy connector in Data Factory.
- Learn more about Profisee running natively in Azure.
- Learn how to deploy Profisee to Azure by using an Azure Resource Manager template (ARM template).
- View Profisee Data Factory templates.

Feedback

Was this page helpful? Yes No

Microsoft Purview and Reltio integration for master data management

Article • 04/03/2025

Reltio Integration for Microsoft Purview is implemented through our Reltio Integration Hub (RIH). The integration consists of Reltio and Microsoft Purview connections and recipes to read the tenant business config from Reltio and create corresponding objects in Microsoft Purview.

The integration supports these objects:

- Entity types
- Entity attributes - Simple, Nested, Reference
- Relationship types
- Relationship attributes - Simple, Nested

The foundation of the integration is creating the EntityDefs in Microsoft Purview's Data Map for every Reltio's object.

Microsoft Purview's Data Map is like a central inventory which powers the Microsoft Purview Unified Catalog, the Microsoft Purview Data Estate Insights and the Microsoft Purview Data Policy as unified experiences within the Microsoft Purview governance portal.

EntityDef is like a class in object-oriented programming and provides the schema or blueprint for all the Reltio objects. It defines the structure of the entities, attributes, relationships, and their types. This schema ensures that all instances of the entity adhere to rules and constraints for the entities.

Expand table

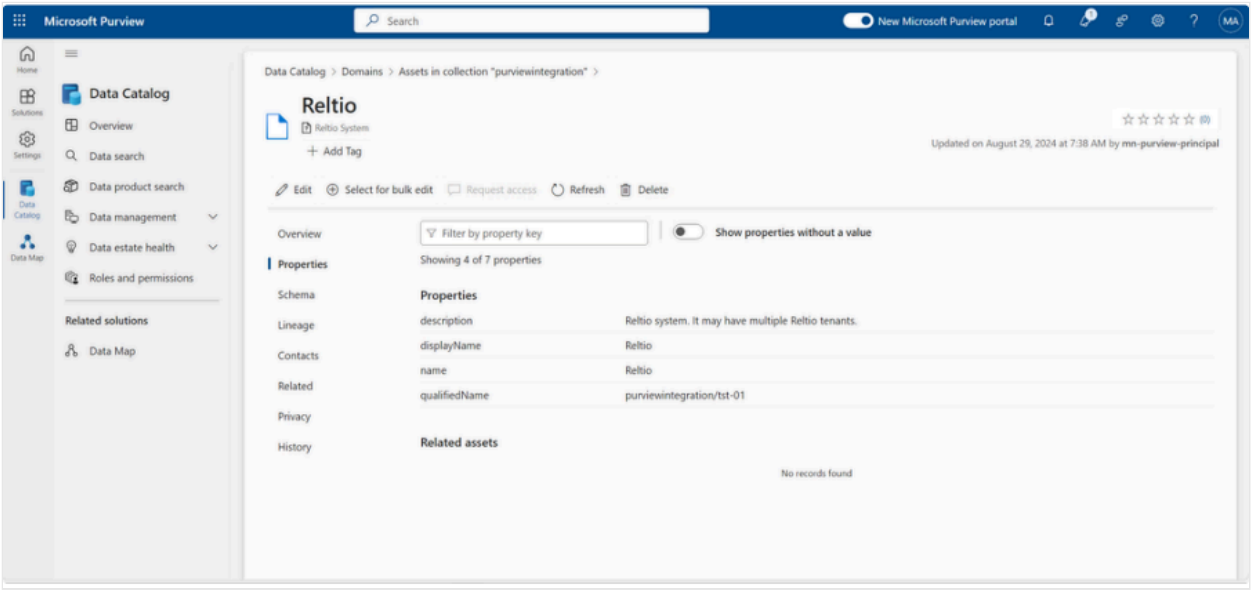
Reltio	Microsoft Microsoft Purview
System	EntityDef "reltio_system"
Tenant	EntityDef "reltio_tenant"
Entity Type	EntityDef "reltio_entity_type"
Simple attribute (Entity Type, Relationship Type)	EntityDef "reltio_attribute"
Nested attribute (Entity Type, Relationship Type)	EntityDef "reltio_attribute"

Reltio	Microsoft Microsoft Purview
Reference attribute (Entity Type)	EntityDef “reltio_attribute”
Relationship Type	EntityDef “reltio_relationship_type”

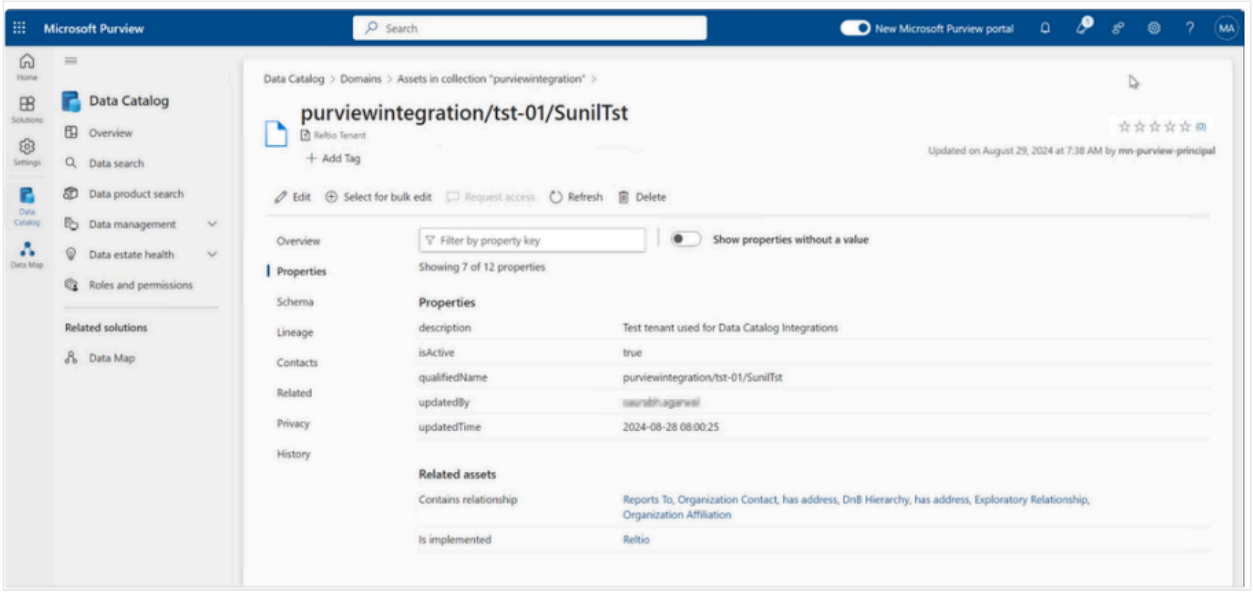
Later the integration creates a Reltio System as an Entity in Microsoft Purview's Data Map. The Reltio tenant is created also as an Entity under the Reltio System in Microsoft Purview using the implements relationship.



Reltio System contains the following properties: description, displayName (same as Reltio label), name (last part of Reltio URI), qualifiedName (unique identifier).



Reltio Tenant contains the following properties: description, displayName, isActive (if entity isn't removed on Reltio), name, qualifiedName, updatedBy (autopopulated), updatedTime (autopopulated).



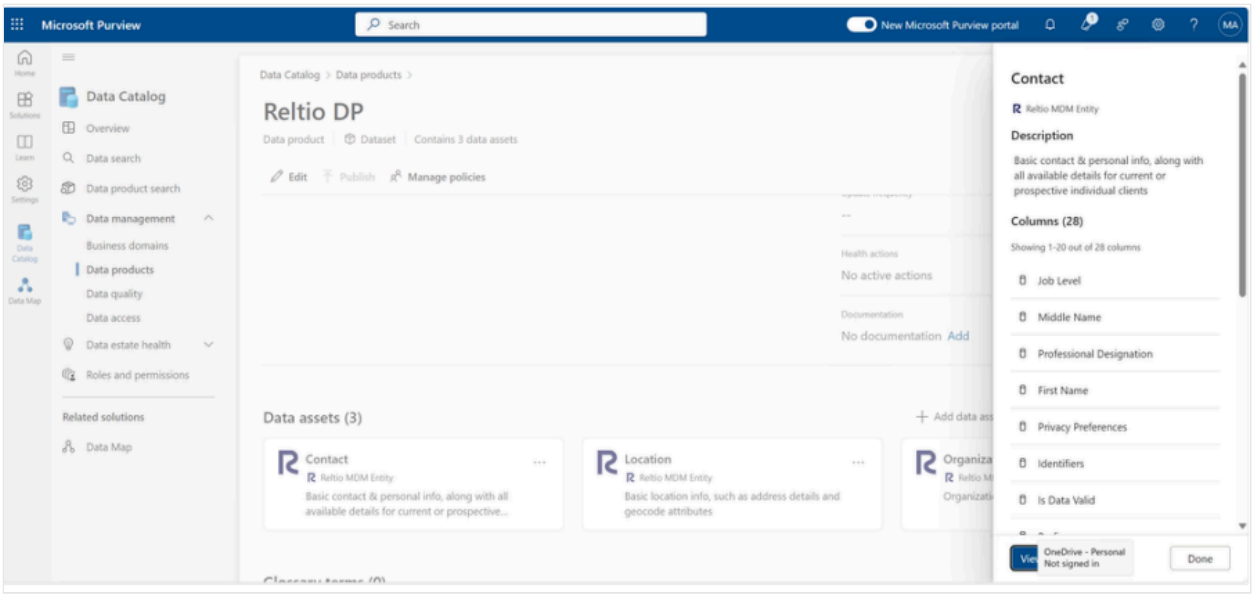
Properties are descriptive attributes or metadata that provide context about the data asset, helping with discovery, management, and governance without affecting the actual structure or format of the data.

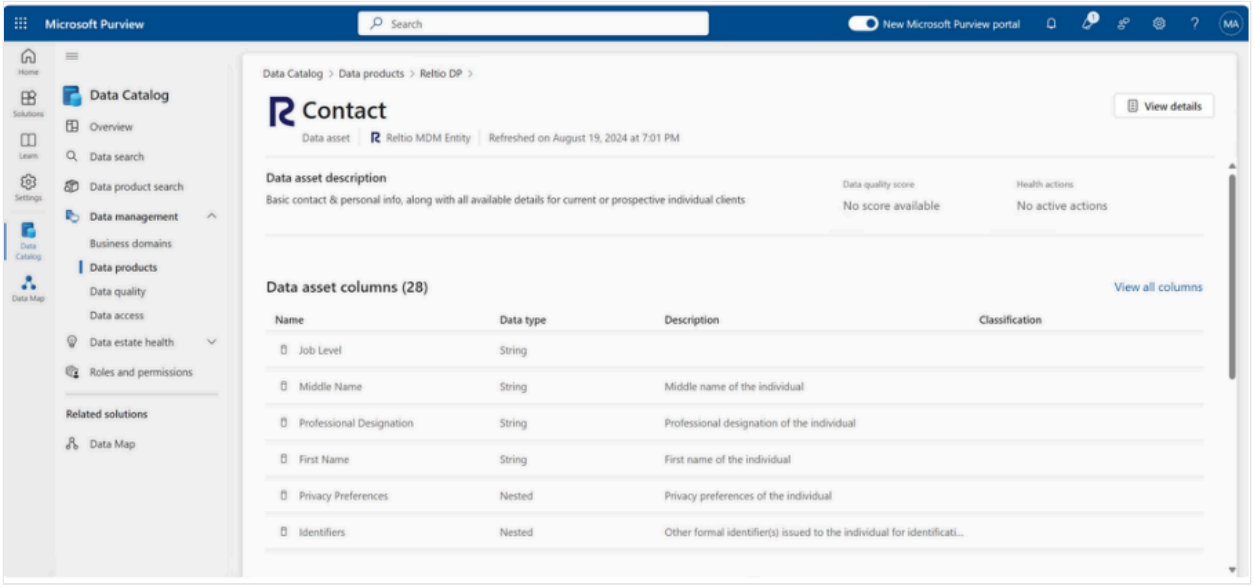
The Microsoft Purview Entity qualifiedName is derived using the

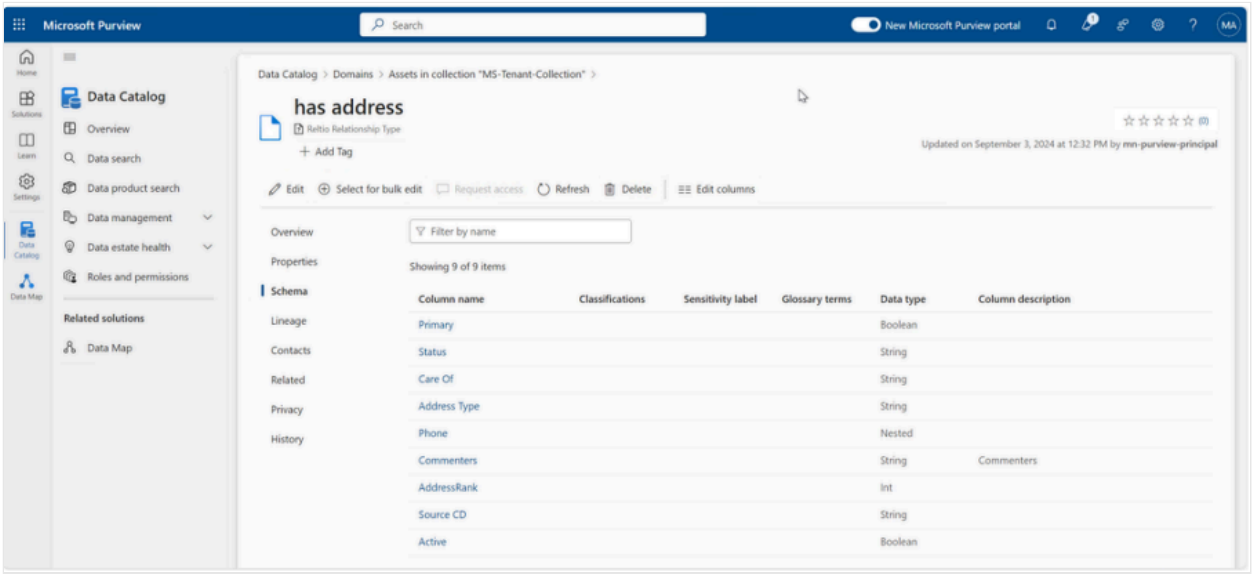
```
<collectionId>/<env>/<tenantId>/<URI>
```

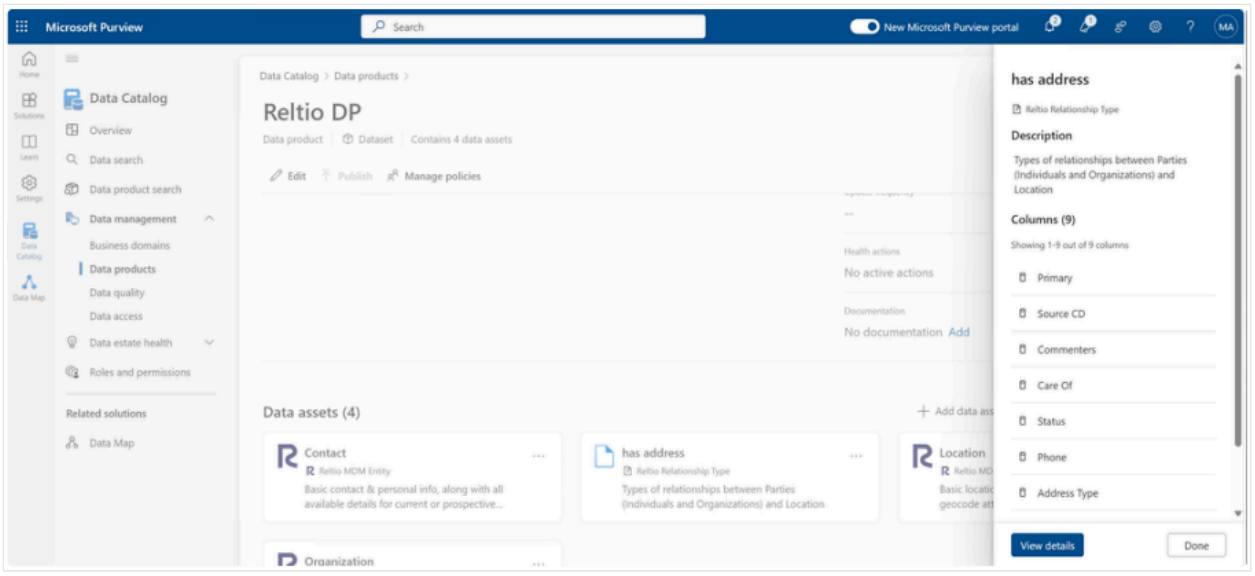
Since the Entity qualifiedName uses the tenant ID and env, it's always unique for a tenant, so the same Microsoft Purview account can be used for multiple tenants.

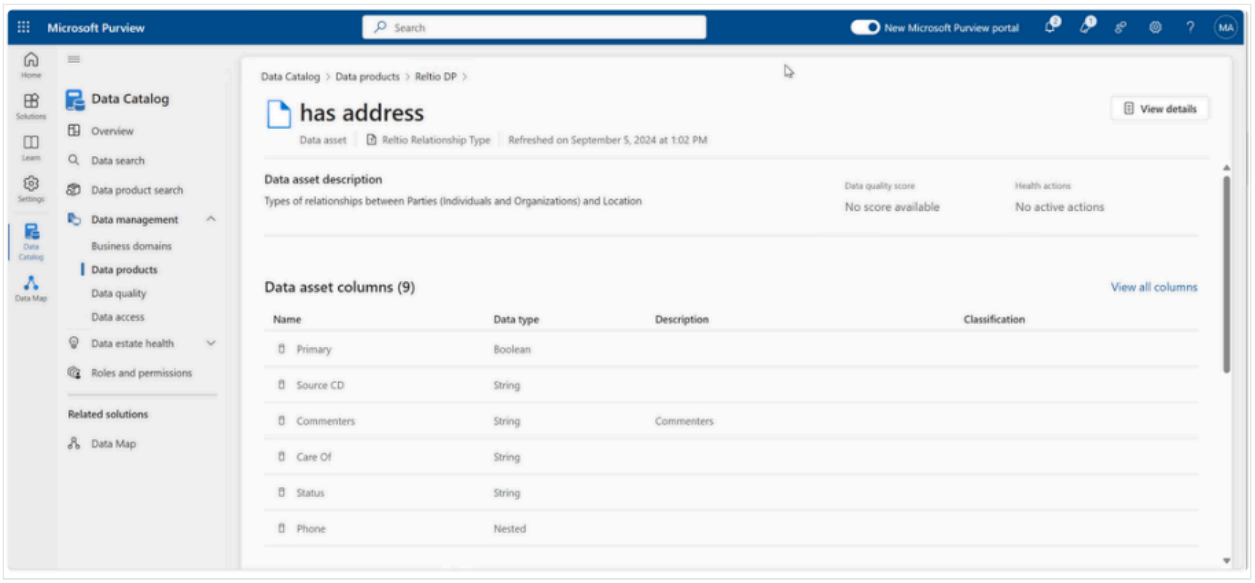
The created Entities can be verified under DataMap->Domains->Collection (by default root)-> Assets-> search the name of your Asset/Entity. To ensure Entities appear in Unified Catalog add them manually to Microsoft Purview Data Product where they need to be published. Entities appear in Unified Catalog assigned to chosen Data Products as Data Assets, with their attributes (subnested and referenced attributes not included) in the Schema section. Schema section is available only for the Entities of specific Asset types: Reltio Entity Type, Reltio Relationship Type. Schema defines the structure of the data and dictates how the data is stored, queried, and maintained. It governs how data must be formatted and what relationships exist within the data


Microsoft Purview
Search
New Microsoft Purview portal
Data Catalog
Overview
Data search
Data product search
Data management
Business domains
Data products
Data quality
Data access
Data estate health
Roles and permissions
Related solutions
Data Map
Data Catalog > Data products >
Reltio DP
Data product
Dataset
Contains 3 data assets
Edit
Publish
Manage policies
Health actions
No active actions
Documentation
No documentation Add
Data assets (3)
Contact
Reltio MDM Entity
Basic contact & personal info, along with all available details for current or prospective...
Location
Reltio MDM Entity
Basic location info, such as address details and geocode attributes
Contact
Reltio MDM Entity
Description
Basic contact & personal info, along with all available details for current or prospective individual clients
Columns (28)
Showing 1-20 out of 28 columns
Job Level
Middle Name
Professional Designation
First Name
Privacy Preferences
Identifiers
Is Data Valid
OneDrive - Personal
Not signed in
Done


Microsoft Purview
Search
New Microsoft Purview portal
Data Catalog
Overview
Data search
Data product search
Data management
Business domains
Data products
Data quality
Data access
Data estate health
Roles and permissions
Related solutions
Data Map
Data Catalog > Data products > Reltio DP >
Contact
Data asset
Reltio MDM Entity
Refreshed on August 19, 2024 at 7:01 PM
View details
Data asset description
Basic contact & personal info, along with all available details for current or prospective individual clients
Data quality score
No score available
Health actions
No active actions
Data asset columns (28)
View all columns
Name
Data type
Description
Classification
Job Level
String
Middle Name
String
Middle name of the individual
Professional Designation
String
Professional designation of the individual
First Name
String
First name of the individual
Privacy Preferences
Nested
Privacy preferences of the individual
Identifiers
Nested
Other formal identifier(s) issued to the individual for identificati...


Microsoft Purview
Search
New Microsoft Purview portal
Data Catalog
Overview
Data search
Data product search
Data management
Data estate health
Roles and permissions
Related solutions
Data Map
Data Catalog > Domains > Assets in collection "MS-Tenant-Collection" >
has address
Reltio Relationship Type
Add Tag
Updated on September 3, 2024 at 12:32 PM by mn-purview-principal
Edit
Select for bulk edit
Request access
Refresh
Delete
Edit columns
Overview
Properties
Schema
Lineage
Contacts
Related
Privacy
History
Filter by name
Showing 9 of 9 items
Column name
Classifications
Sensitivity label
Glossary terms
Data type
Column description
Primary
Boolean
Status
String
Care Of
String
Address Type
String
Phone
Nested
Commenters
String
Commenters
AddressRank
Int
Source CD
String
Active
Boolean





Schema

Schema is an important concept that reflects how data is stored and organized in the data store. It reflects the structure of the data and the data restrictions of the elements that construct the structure. Elements on the same schema can be classified differently (due to their content). Also, different transformation (lineage) can happen to only a subset of elements. Due to these aspects, Microsoft Purview can model schema and schema elements as entities, hence schema is usually a relationship attribute to the data asset entity. Examples of schema elements are: columns of a table, json properties of json schema, xml elements of xml schema etc.

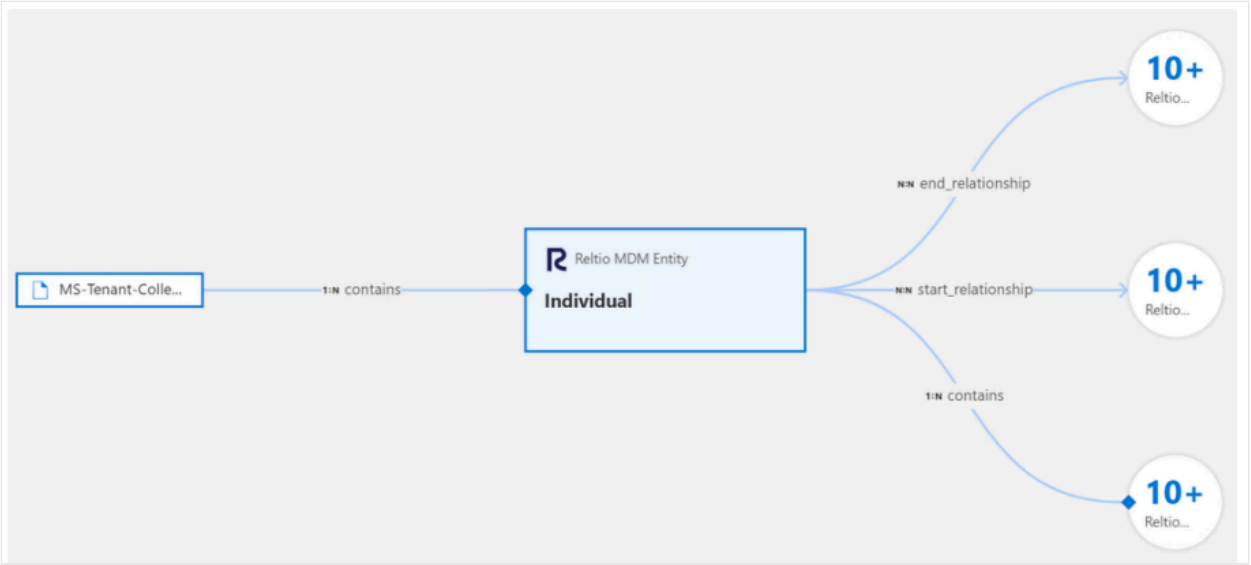
Type definitions and how to create custom types in Microsoft Purview

Reltio objects mapping to Microsoft Purview entities

Entity Types

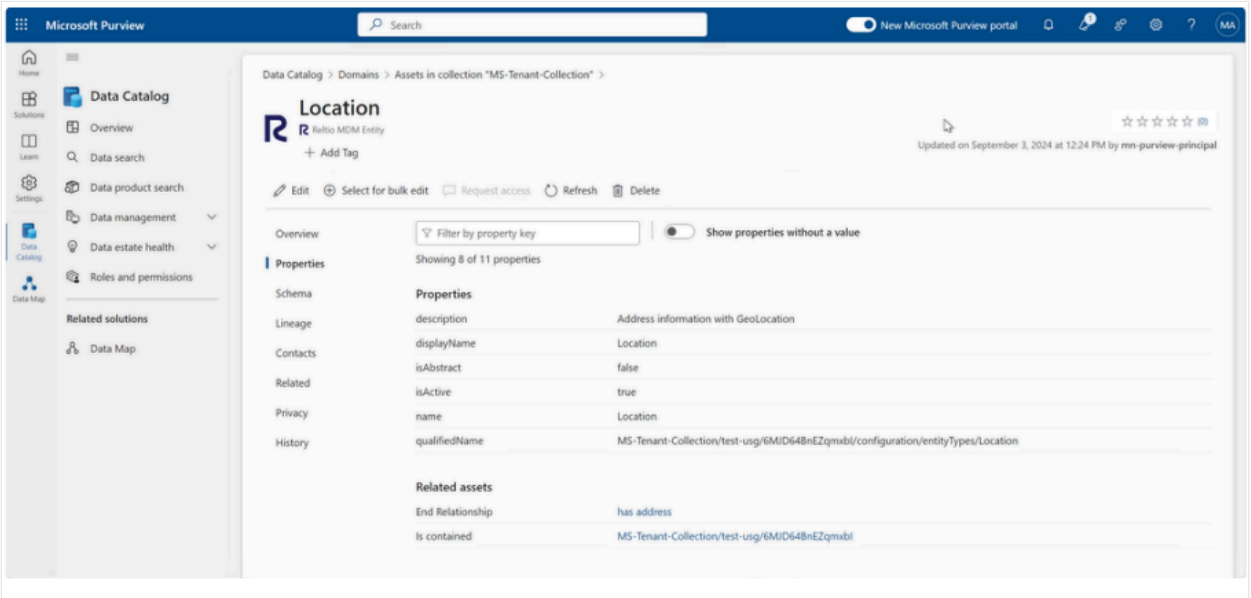
Entity Types within a tenant are created as Entities called Reltio MDM Entity in Microsoft Purview with a contains relationship with the tenant.

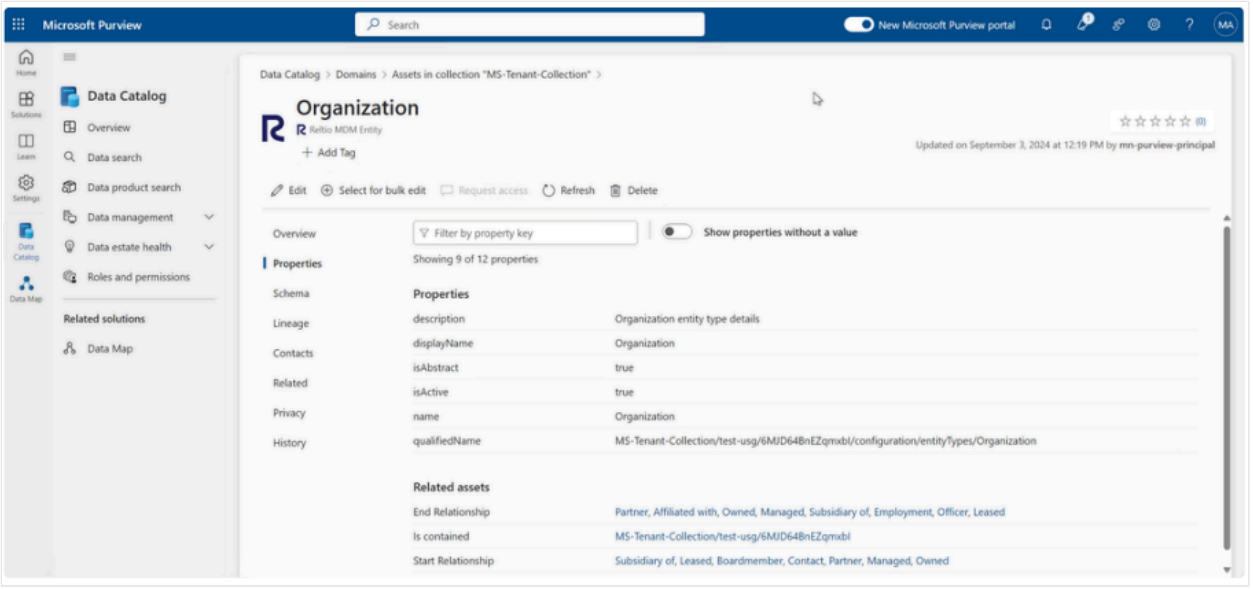




Properties of Reltio MDM Entity

Reltio MDM entity contains these properties: description, displayName, isAbstract (if the entity is abstract on Reltio), isActive, name, qualifiedName.



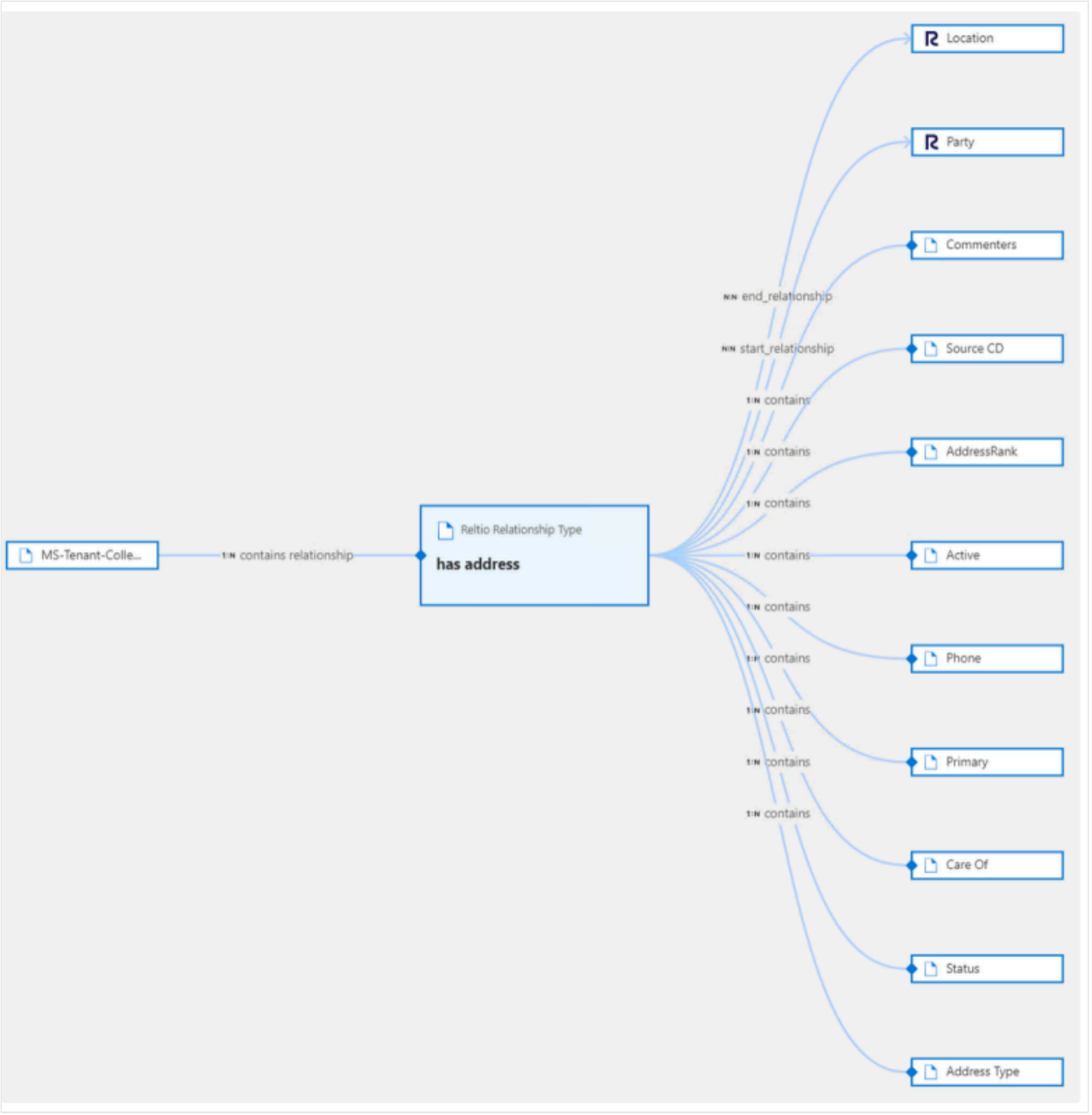


Relationship Types

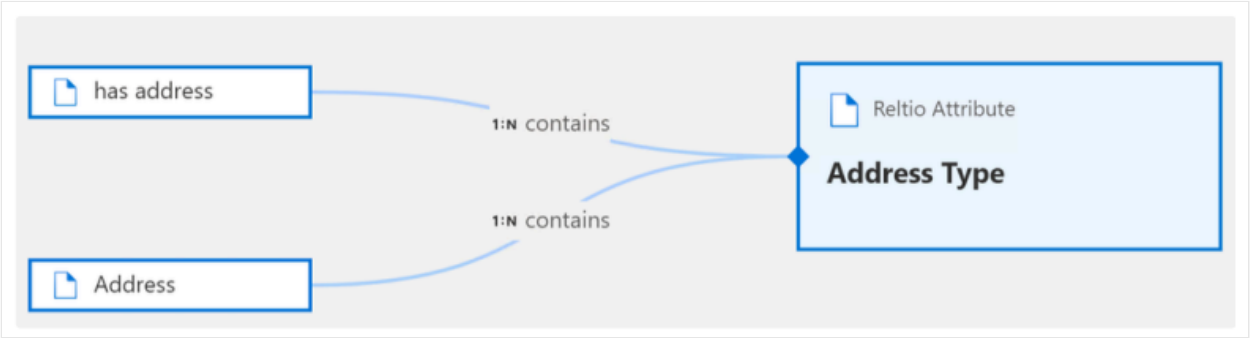
Relationship Types within a tenant are created as Entities called Reltio Relationship Type in Microsoft Purview with a contains relationship with the tenant.

Also, Reltio Relationship Type Entity has relations with two Reltio MDM Entities: start_relationship and end_relationship. Relationship attributes are added as Entities called Reltio Attributes with a contains relation with the Reltio **Relationship Type Entity.**

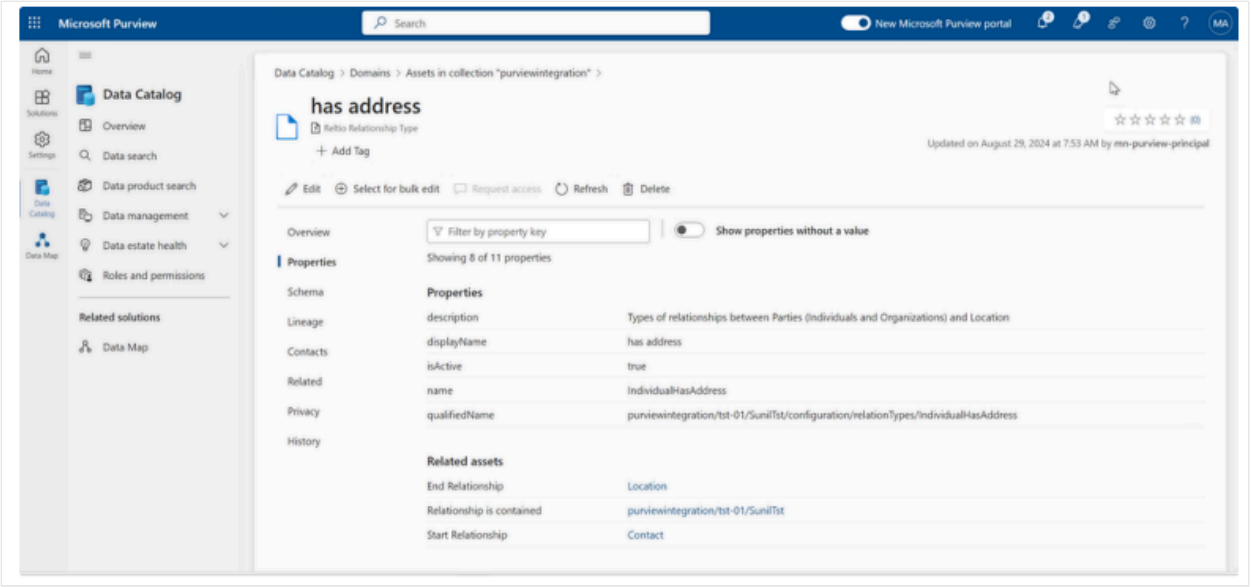
This screenshot shows a Contact Reltio MDM Entity connected to have address Reltio Relationship Type Entity with start_relationship Relationship. And later has address Reltio Relationship Type Entity is connected to Location Reltio Entity **Type Entity** with end_relationship Relationship. The **has address** Reltio Relationship Type Entity has four simple data attributes with contains relation: Address Type, Address. Rank, Active, Status.



Relationship attributes can further have a contains relation with the Reference attribute. In the example, the Address Type relationship data attributes references the Contact Reltio Entity Type Entity.



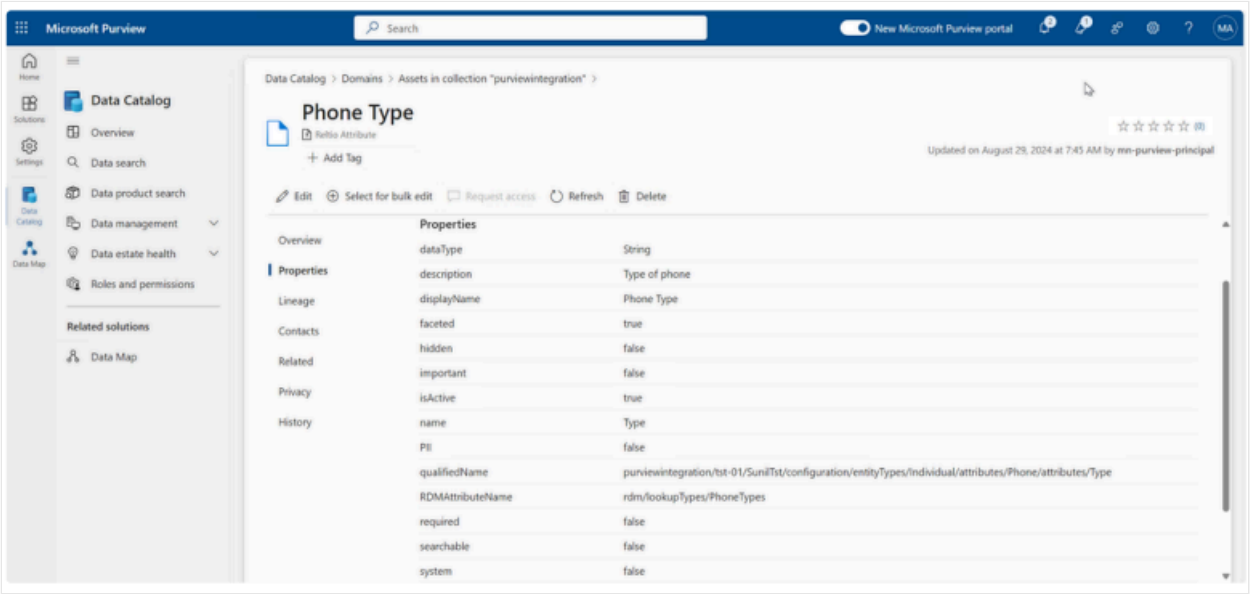
Also, Reltio Relationship Type Entities contain the same properties as Reltio Entity Type Entities: description, displayName, isActive, name, qualifiedName.



Attributes

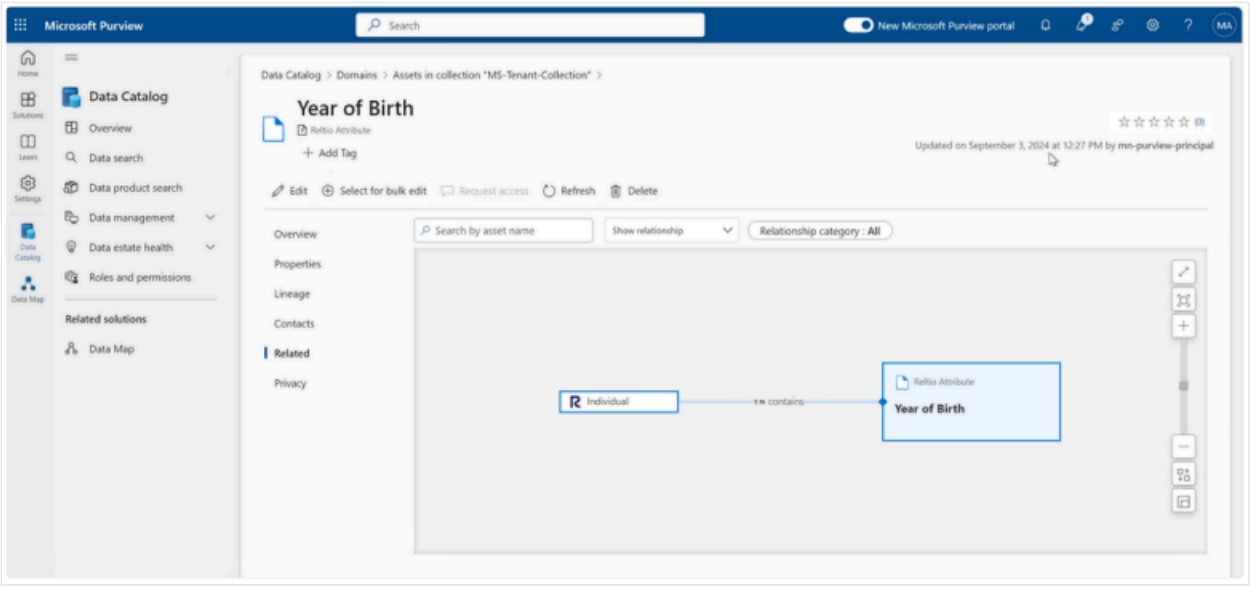
All Reltio Attribute Entities (Simple, Nested, Reference) may contain the following properties (depending on the availability on Reltio):

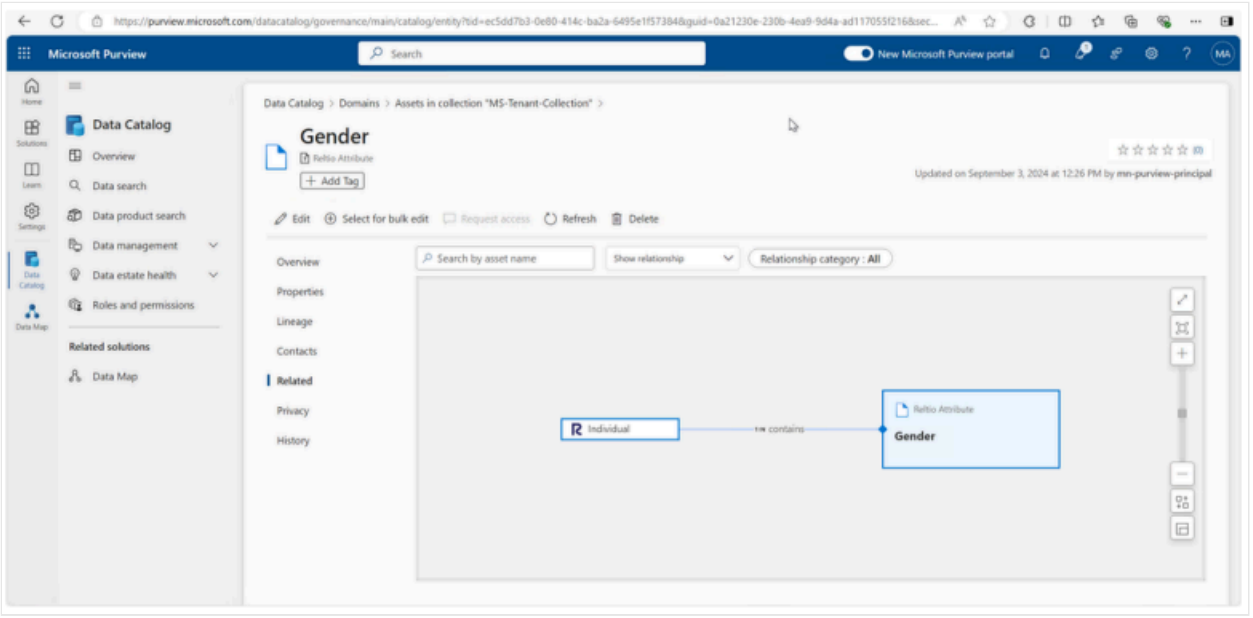
- dataType (Nested, Reference, or data type for Simple attributes),
- description, displayName, isActive, name, qualifiedName,
- faceted
- hidden
- important
- PII
- RDMAttributeName
- required
- searchable
- system

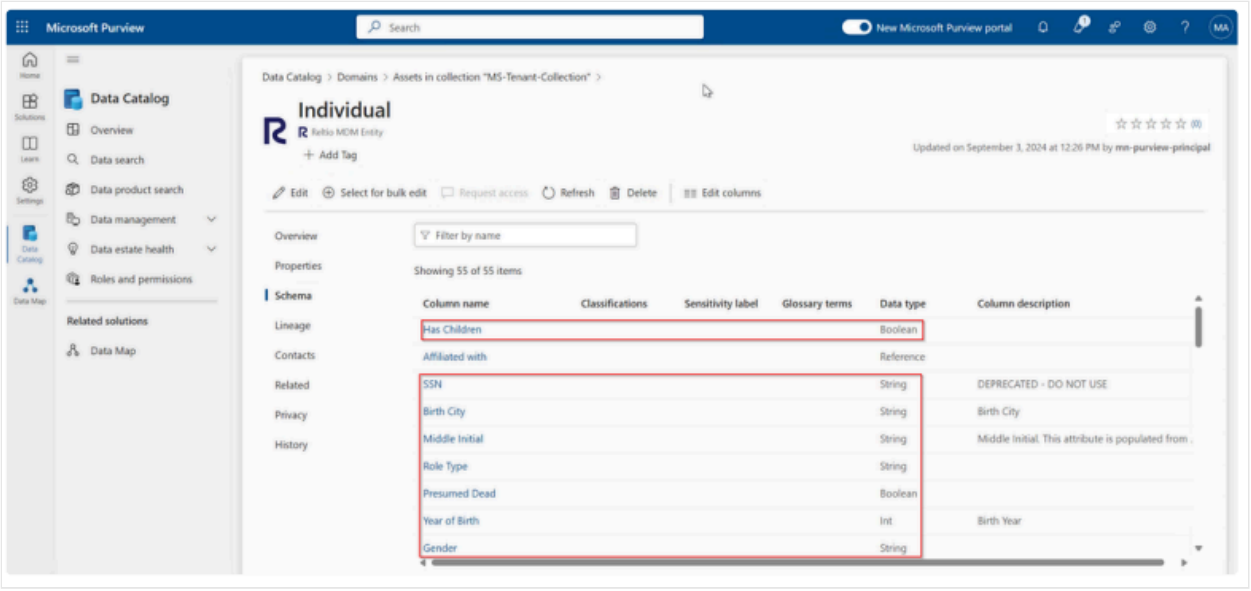


Simple Attributes

Simple attributes are created as Entities called Reltio Attribute in Microsoft Purview with a contains relationship with its entity.

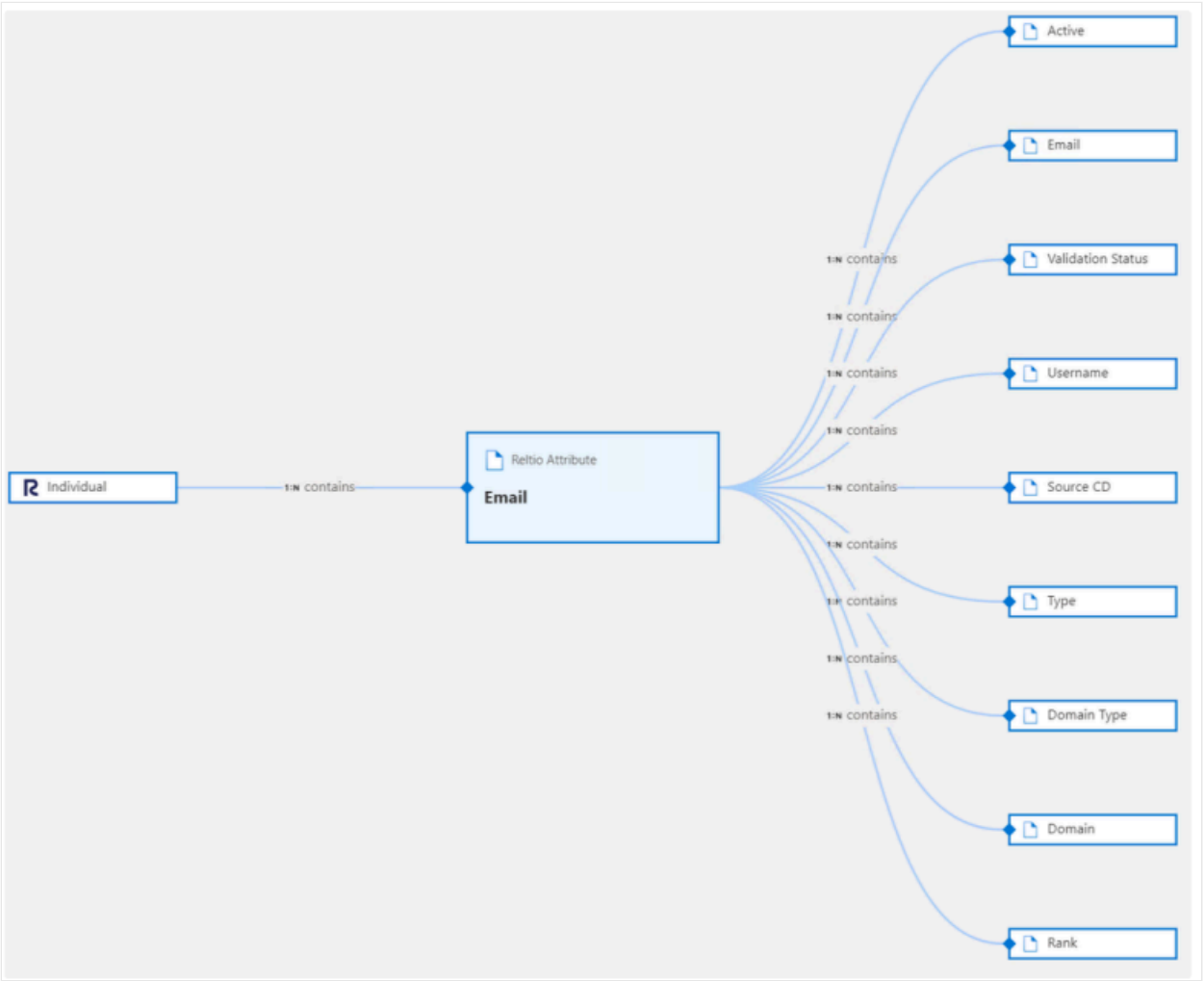


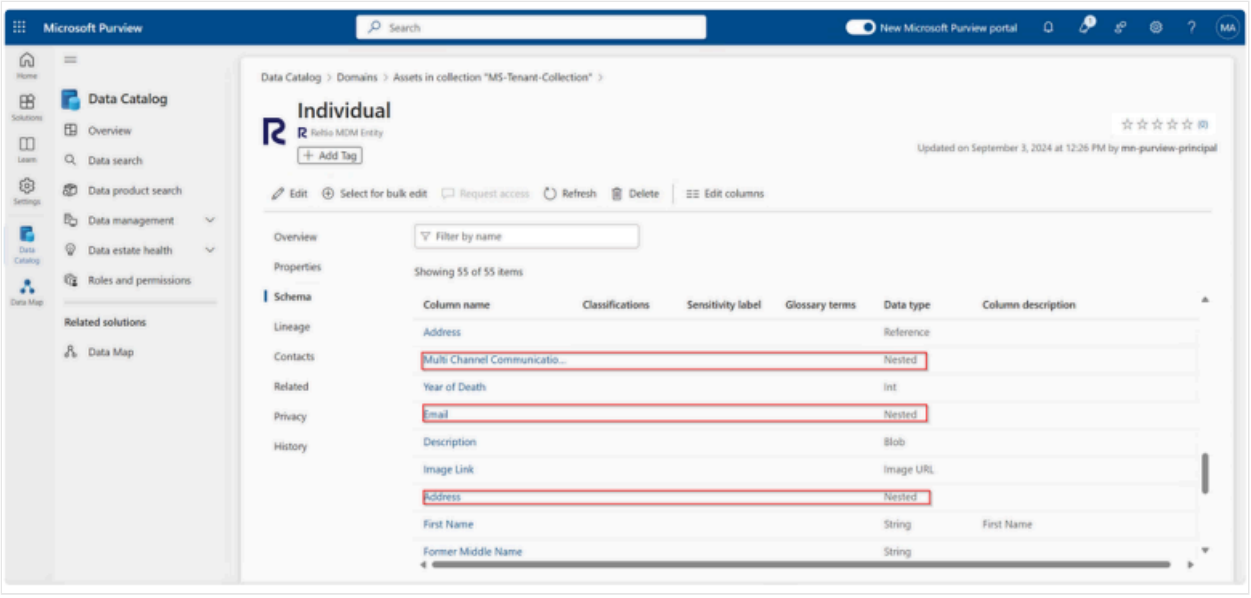




Nested Attributes

Nested attributes are created as Entities called Reltio Attribute in Microsoft Purview with a relationship with its entity. Sub nested attributes are also created as Entities called Reltio Attribute with a relationship with the parent nested attribute. This screenshot shows a Contact Reltio Entity Type that contains the Email nested attribute, which further contains multiple sub nested attributes like username, domain type, etc.

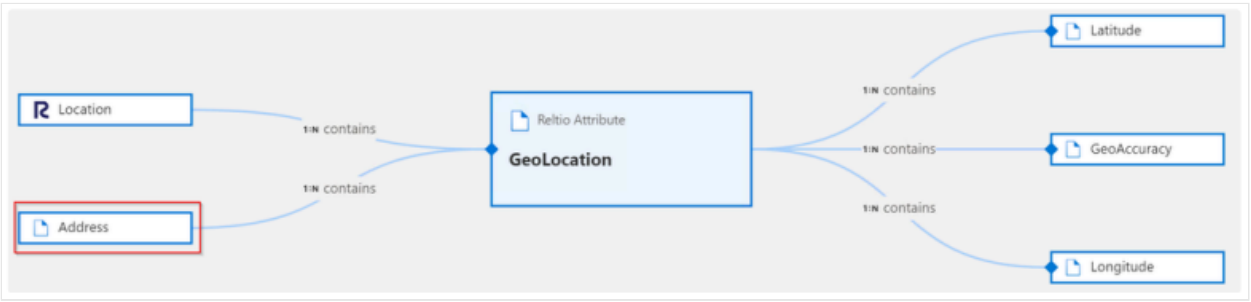


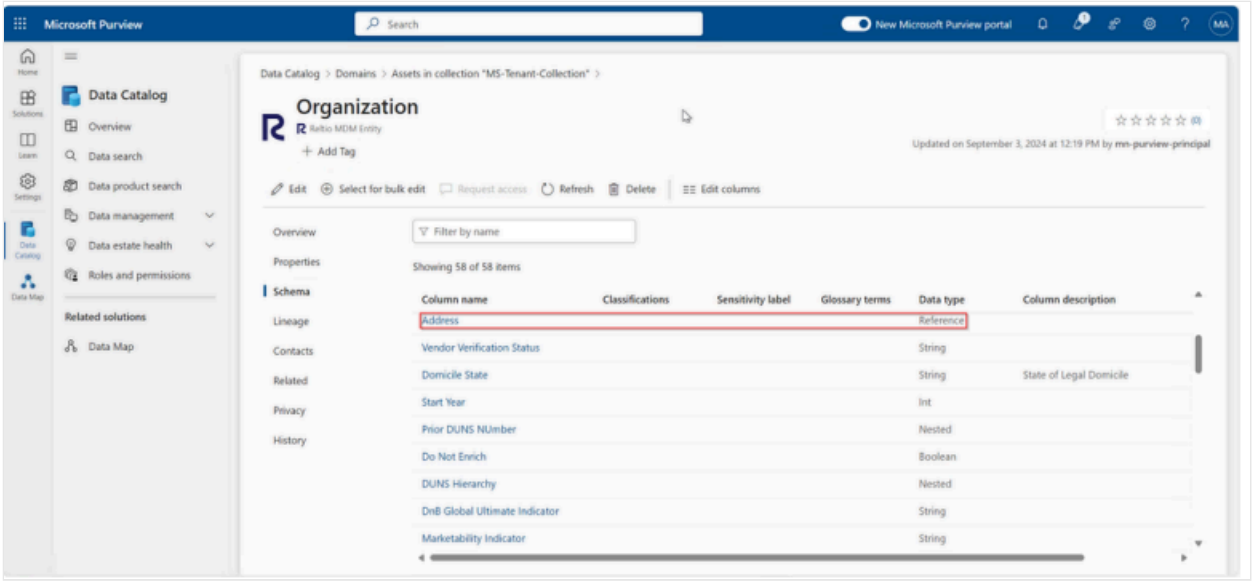


Reference Attributes

A Reference attribute in a Reltio Entity Type is also created as Entities called Reltio Attribute in Microsoft Purview with a contains relationship with its entity. Sub referenced attributes are also created as Entities called Reltio Attribute with a contains relationship with the parent reference attribute.

This screenshot shows a Contact Reltio Entity Type entity with a Reference attribute called Address, which contains multiple simple attributes like verification status, PO Box, state, country, etc., and some nested attributes like Zip and Geo Location.





Lookup tables

- Microsoft Purview Created Assets
- Stores Microsoft Purview asset GUID. Will be used for soft delete of Microsoft Purview assets.
- Microsoft Purview Reference Attributes Table
- Stores ID of reference attributes and its array of referred URIs. This is created for avoiding going through the L3 configuration twice in order to create reference attributes.

Microsoft Purview Recursion Table

Acts as a stack for storing function calls for implementing recursive calls for processing nested attributes. Stores attribute's json, asset ID of parent, asset type of parent, relationship type. Contains a dummy field called processed which is not used but is required for getting a single entry from the table.

Configure RIH recipes

You need to have the Microsoft Purview connector license in your contract. Reach out to your Reltio Account Manager for licensing needs.

Once you have the Microsoft Purview connector license in your Reltio contract, the Microsoft Purview connector recipes and connections are deployed to your tenant's Reltio Integration Hub (RIH) instance.

- Configure Connections

- Configure Environment properties.

To configure the recipes: Check the recipes

The recipes package contains two connections and seven recipe functions. The package is deployed in the Reltio Solution Templates project in your RIH account.

- Connect with your Reltio tenant
  - Enter the connection details
    - Connection Name: for example, Microsoft Purview
    - Environment URL: for example, 361, tst-01 etc
    - Tenant Id: for example, a0bcdEf1ghiJ2Kl
    - Set Client Authorization to No
    - Reltio User Name and Password
  - Select Connect
- The status changes to Connected and the button to Disconnect when the connection is established
- Connect with your Purview account

Before you're able to connect you need to generate the client ID and client secret based on this documentation - API authentication for Microsoft Purview Data Planes

1. Enter the connection details

- Connection Name - for example, Microsoft Purview Portal
- Connection Type - Cloud
- Authentication type - Azure OAuth 2
- Azure Tenant ID - for example, a0bcdEf1ghiJ2Kl
- Azure OAuth client ID & client secret
- Base URL

```
https://api.purview-service.microsoft.com
```

- OAuth scopes

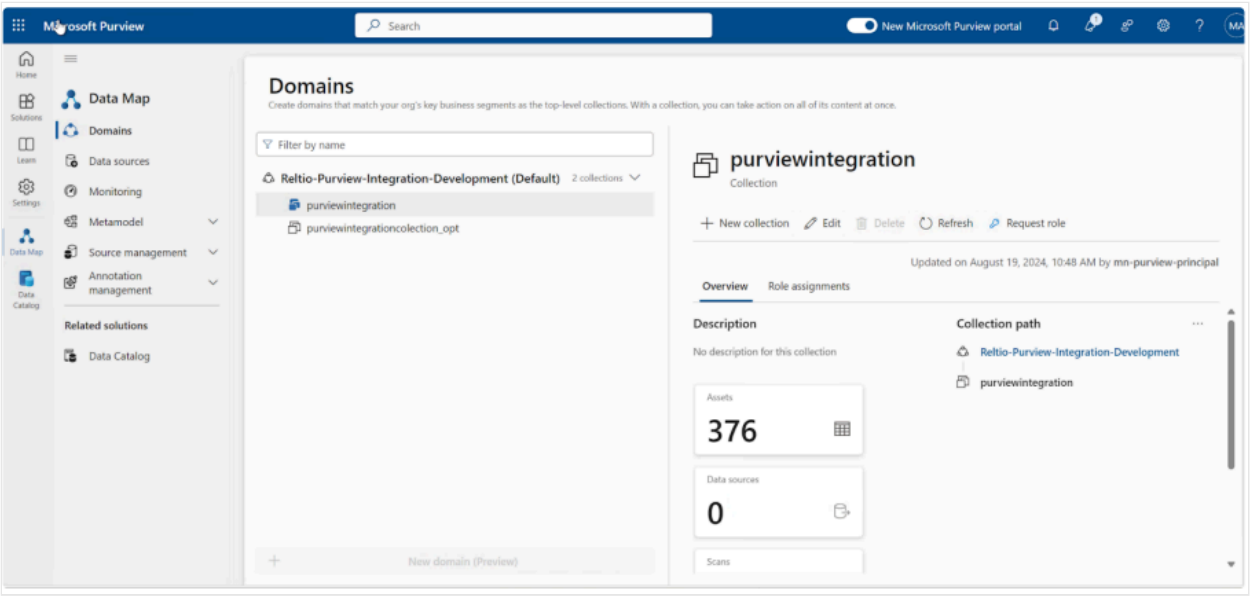
2. Select Connect

- The status changes to Connected and the button to Disconnect when the connection is established.

Manually create a Microsoft Purview Domain in the Data Map domains and add it to the recipe parameters (name of parameter) & add Collection to the recipe parameters name

to create the Reltio system in an existing Domain & Collection. Make sure the User ID of Collection creator matches the UserID of integration recipe, otherwise data won't be populated there. If you don't create the Domain then the recipe fails. If you don't provide a value for Collection then the integration uses this default value:

- Collection - purviewintegration



Use Reltio integration for Microsoft Purview

Let's cover the usage scenarios for Reltio Integration for Microsoft Purview. The following recipes and functions are deployed to the Recipes folder:

- Main
- Initial setup
- Create Entities
- Create Attributes
- Create Relationships
- Create reference attribute relationships
- Soft-delete Assets

Related content

Microsoft Purview MDM integration

Feedback

Was this page helpful? Yes No

Microsoft Purview and Semarchy integration for master data management

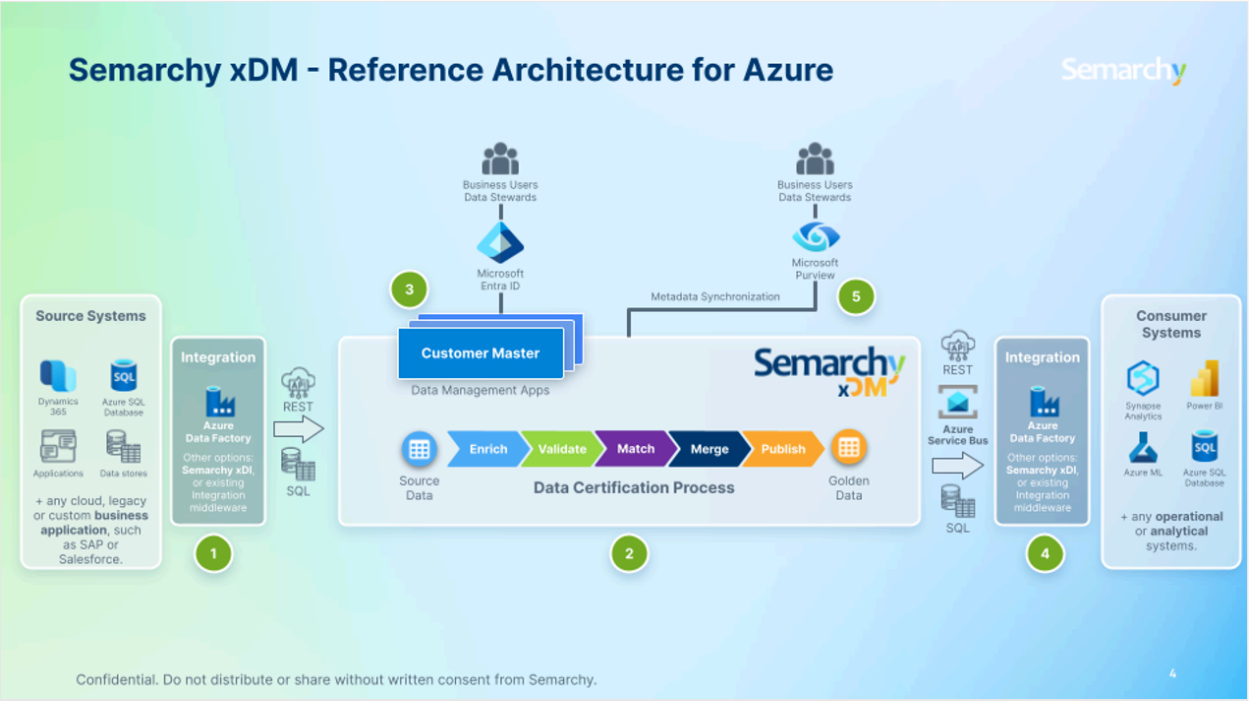
Article • 04/03/2025

Modern organizations generate large quantities of data, often from numerous, disparate sources. The Semarchy Data Platform is the intelligent data hub for data integration, master data management (MDM), reference data management (RDM), application data management (ADM), data integration, quality, and governance. Semarchy Data Management (xDM) brings extreme agility for defining and implementing data management applications and releasing them to production.

This architecture demonstrates how to include Master Data Management (MDM) into the Azure ecosystem to enable quality, validation, matching, deduplication, authoring, curation, and collaboration for your critical data assets.

Architecture

The following architecture illustrates the Semarchy xDM architecture and data flow.



Data Flow

Metadata and data flow include the following steps:

- **Source data integration** from identified source systems:
  - This integration uses Azure Data Factory, Semarchy xDI (Semarchy's data integration component), or your integration solution.
  - semarchy xDM exposes SQL or REST endpoints for batch and real-time integration.
  - Incoming data can be profiled with Semarchy xDM Discovery and reviewed to help define the master data model structure and rules.
  - At any time, the REST endpoints can also be used by applications to interact (read/write) with the master data managed in xDM, and use xDM as their master data backend.

- **Automatic Data Certification**
  - Data undergoes enrichment, standardization, and quality validation through a combination of rules, plugins, AI models (including Azure Machine Learning and Azure OpenAI), and third-party services. semarchy xDM supports these automated quality processes with data recycling and incorporates user corrections to ensure optimal data quality.
  - Matching, merging, and survivorship processes are automatically carried out by combining sophisticated automation with informed user decisions, to produce accurate, and reliable consolidated golden records.
  - Golden records produced by the certification process feature complete lineage to the source systems, user change tracking, and optional historization. This guarantees data integrity, traceability, and comprehensive historical record tracking.

- **Customized Data Management Applications** are available for users to:
  - Authenticate via Microsoft Entra ID,
  - Browse and search certified data with complete traceability through each certification stage, alongside comprehensive historization,
  - Manage and curate data:
    - Users can author and import new data,
    - review and fix errors,
    - manually match and merge records with override options,
    - and perform soft or hard deletes as necessary,
  - Collaborate with Data-Driven Workflows.

- **Golden data distribution** to consumer operational and analytical applications such as Azure Synapse Analytics, Power BI, Azure Machine Learning, and Azure OpenAI, ensures uninterrupted integration and usage across platforms:
  - This integration uses Azure Data Factory, Semarchy xDI, or your integration solution.

    - Semarchy xDM provides built-in SQL and REST endpoints, and Data Notifications for event-based propagation in Azure Service Bus.

- **Synchronize xDM metadata** with Microsoft Purview to gain comprehensive visibility and lineage of the entire master data flow.

Components

This architecture involves the following components.

Core Components

- Semarchy xDM is a no-code platform that allows data teams to quickly develop customized master data management solutions, offering a wide range of capabilities for complex data ecosystems.
- Azure Database for SQL Server and PostgreSQL are fully managed databases as a service with built-in capabilities, such as high availability and intelligence. They store both the Semarchy metadata and master data hubs managed in xDM. Data processing is performed in the database layer, guaranteeing the best performance and scalability.
- Microsoft Entra ID, for user authentication and single sign-on to the Semarchy platform.
- Azure Key Vault is a cloud service that provides secure storage for secrets. You can use it to encrypt, decrypt, and store secrets (passwords, for example) used in xDM.

AI Components

- **Azure Machine Learning** is a cloud service for accelerating and managing machine learning (ML) projects. Semarchy xDM can use customized Azure Machine Learning models in the data hub certification processes.
- **Azure OpenAI** is a suite of AI services providing access to OpenAI's powerful language models. Semarchy xDM includes built-in plug-ins using these language models to enrich and certify data, for example for content generation, summarization, or translation.

Governance Components

- **Microsoft Purview** is a data governance solution that provides broad visibility into on-premises and cloud data estates. Semarchy xDM integrates with Microsoft Purview to provide insights into Semarchy Data Hubs as data products and end-to-end master data lineage.

Source and Consumer Systems Among others, this architecture includes the following systems from which you collect master data to be managed in Semarchy xDM, or to which you send golden data produced by Semarchy xDM.

- **Azure Synapse Analytics** is a fast, flexible, and trusted cloud data warehouse that uses a massively parallel processing architecture. Semarchy Data Hubs act as providers of certified metadata for Azure Synapse.
- **Power BI** is a business analytics suite that delivers insights throughout your organization. You can use Power BI to build dashboards and reports on top of Semarchy Discovery metrics and the Semarchy Data Hubs.

Scenario Details

Data-driven initiatives, such as digital transformation, business intelligence, or AI projects require accurate and trustable data. Master Data Management is an essential step to deliver this clean, accurate data.

A common use case for an MDM solution is to consolidate master data from multiple sources while allowing collaborative authoring and stewardship of this master data to serve analytical and operational applications with golden data.

Design data applications

Semarchy xDM Data Management Applications provide all users with a customized experience to access and manage their data. Through these applications, master data records are displayed in fully customizable interfaces, supporting data management, authoring, and stewardship operations. Users with different roles and personas collaborate in Data-Driven Workflows to manage data. The power of Semarchy xDM lies in the flexibility in the design of your data applications, allowing them to adapt to your domains, organization, and business needs.

Integrate and certify master data

Data curated in external source systems, such as Customer Relationship Management (CRM), Enterprise Resource Planning (ERP), or other systems (known as the publishers) is pushed to Semarchy xDM's data hubs via an integration layer, such as Azure Data Factory or Semarchy xDI.

As data changes appear in the data hub, through data loads or authoring, it passes through the entire certification process, during which it's enriched, standardized, validated for data quality, and then matched and merged the records.

Consume data from the hub

Data can be pushed to or consumed from Semarchy xDM using REST API endpoints, or through SQL. Changes made to the data through the data hub can also be propagated in real-time to downstream systems using data notifications.

Considerations

The Semarchy Data Platform has several features that address the issues of reliability, security, cost optimization, operational excellence, and performance efficiency. Further information about architectural excellence can be found in this article on the ** pillars of well-architected Azure frameworks.**

Reliability

Reliability ensures your application fulfills the promises you make to your customers. For more information, see Overview of the reliability pillar.

Semarchy xDM runs natively on Azure Kubernetes Service and Azure SQL Database, which offers out-of-the-box capabilities to support high availability.

Security

Security shields against intentional attacks and misuse of your valuable data and systems. For more information, see Security overview in Microsoft learn page.

Semarchy xDM authenticates users via its identity management layer that supports role mapping, lookup, and profile synchronization. It provides native support for multiple IDPs, including Microsoft Entra ID. It also includes advanced security features such as fine-grained privileges to secure access and operations.

Performance Efficiency

Performance efficiency is the capability of your system to scale and effectively meet user needs. For more information, see Performance Efficiency learn page in microsoft learn.

Semarchy xDM runs natively on Azure Kubernetes Service and Azure SQL Database. You can configure Azure Kubernetes Service to scale up and out. You can deploy and configure Azure SQL Database to balance performance, scalability, and costs.

Cost Optimization

Cost optimization involves finding ways to reduce unnecessary expenses and enhance operational efficiency. For more information, see Cost Optimization learn page in Microsoft learn.

Running costs consist of the Semarchy software subscription license and Azure consumption. Contact Semarchy for more information.

Deploy this scenario

To deploy this scenario:

1. Deploy Semarchy xDM using Azure Kubernetes Services.
2. Configure Secrets Management to use Azure Key Vault.
3. Configure Authentication with Microsoft Entra ID.
4. Design and deploy your customized master data model in xDM
5. Integrate your data into xDM using Azure Data Factory.

Contributors to this document

- David Cox
- Cedric Blanc
- François-Xavier Nicolas (FX)

Reference document

- Deploy Semarchy xDM
- Copy Secrets Management
- Configure Authentication with Microsoft Entra ID
- Integrate your data into xDM using Azure Data Factory

Feedback

Was this page helpful? Yes No

Create an Azure policy exclusion for Microsoft Purview

Article • 05/14/2025

Many subscriptions have Azure Policies in place that restrict the creation or update of some resources. This is to maintain subscription security and cleanliness. However, Microsoft Purview accounts created before December 15, 2023 (or deployed using API version previous to 2023-05-01-preview) deployed an Azure Storage account when it was created. It's managed by Azure, so you don't need to maintain it, but it's necessary for Microsoft Purview to run correctly. Existing policies could block Microsoft Purview from updating this Azure Storage account, which can cause errors during scanning.

Overall, your approach to resolving this error is going to be dependent on your organizations needs and policies, but here are a few ways you could update your policies to resolve this issue.

To maintain your policies in your subscription, but still allow the creation and updates to these managed resources, you can create an exclusion. For example, adding an exclusion for the resource group where your Microsoft Purview account is being deployed.

For steps to create an exclusion for your policies, see these steps to add an exclusion to your policy.

Use resource selectors on a created policy

Depending on other policies deployed in your subscription, or depending on your region, you could need to add **Resource selectors** under the **Advanced** tab when assigning a policy. For example, you might need to add a resource selector for **resourceLocation** set to the region where you'll deploy your Microsoft Purview account. For more information about these conditions, see our documentation on location conditions.

Create an exemption

Exemptions are recommended for time-bound or specific scenarios where a resource or resource hierarchy should still be tracked and would otherwise be evaluated, but there's a specific reason it shouldn't be assessed for compliance. It would allow a temporary exception to your policy for a resource to exist for a finite amount of time.

For more information, see the documentation for policy exemptions.

Next steps

To set up Microsoft Purview by using Private Link, see Use private endpoints for your Microsoft Purview account.

GraphQL API with Microsoft Purview (preview)

Article • 05/14/2025

> **Important**
>
> This feature is currently in preview. The Supplemental Terms of Use for Microsoft Azure Previews include additional legal terms that apply to Azure features that are in beta, in preview, or otherwise not yet released into general availability.

In this tutorial, learn to programmatically interact with Microsoft Purview using the GraphQL API. For more information about GraphQL in general, see this introduction to GraphQL.

Using GraphQL is similar to using the REST APIs, in that you send a JSON payload to a service endpoint. However, GraphQL allows us to return complete information in a single fetch, eliminating the need for multiple API calls.

GraphQL also uses declarative data fetching. Declarative data fetching is useful for selective fetch on fields like classification/term/related entity linked to the original entity. By using GraphQL with these query patterns, we can optimize backend database fetching and data transmission. A good example is "Get Entity with filtered related entities, separated by aliases."

With its introspection feature, the GraphQL API becomes self-descriptive, enabling clients to retrieve schema details such as available queries, types, and query parameters. See more about introspection.

Prerequisites

- If you don't have an Azure subscription, create a free account before you begin.
- You must have an existing Microsoft Purview account. If you don't, see the quickstart for creating a Microsoft Purview account.
- To establish a bearer token and to call any APIs, see the documentation about how to authenticate APIs for Microsoft Purview.

GraphQL Endpoint

For all the requests, you can send a `POST` request to the following endpoint:

```
.NET CLI
```

```
POST https://{{endpoint}}/datamap/api/graphql
```

- If you're using the new Microsoft Purview portal, the {{endpoint}} value is: `api.purview-service.microsoft.com`.
- If you're using the classic Microsoft Purview governance portal, the {{endpoint}} value is: `{your_purview_account_name}.purview.azure.com`.

Basic query

Entity - List By Guids

JSON

```
query {
    entities(where: { guid: ["<guid1>", "<guid2>"] }) { #Values in the array are
combined as a logical-OR.
        guid
        createTime
        updateTime
        typeName
        attributes
        name
        qualifiedName
        description
    }
}
```

Sample Response:

JSON


```
{
    "data": {
        "entities": [
            {
                "guid": "9fb74c11-ac48-4650-95bc-760665c5bd92",
                "createTime": 1553072455110,
                "updateTime": 1553072455110,
                "typeName": "azure_storage_account",
                "attributes": {
                    "qualifiedName": "https://exampleaccount.core.windows.net",
                    "name": "ExampleStorageAccount",
                },
                "name": "ExampleStorageAccount",
                "qualifiedName": "https://exampleaccount.core.windows.net",
                "description": "Example Storage Account"
```

```
            }
        ]
    }
}
```

Query entities by type and qualified name

JSON

```
query {
    entities(
        where: {
            type: { typeName: ["<entityType1>", "<entityType2>"] }
            qualifiedName: { in: ["<qualifiedName1>", "<qualifiedName2>"] }
        }
    ) {
        guid
        typeName
        qualifiedName
        attributes
    }
}
```

The operator `in` is used to query multiple qualified names.

Possible Operators could be `exists|eq|ne|in|nin|gt|ge|lt|le`. You can find more server schema details with introspection.

Declarative data fetching

Get entity with selected fields

JSON

```
query {
    entities(where: { guid: "<guid>" }) {
        guid
        typeName
        attributes
        businessAttributes
        qualifiedName
    }
}
```

Get entity with all fields

JSON

```
query {
    entities(where: { guid: "<guid>" }) {
        guid
        createTime
        createdBy
        updateTime
        updatedBy
        lastModifiedTS
        typeName
        attributes
        businessAttributes
        collectionId
        customAttributes
        hierarchyInfo {
            ...hierarchyInfoFields
        }
        labels
        sensitivityLabel {
            ...sensitivityLabelFields
        }
        source
        sourceDetails
        qualifiedName
        name
        description
        displayName
        userDescription
        classifications {
            ...classificationFields
        }
        relatedEntities {
            ...relatedEntitiesFields
        }
        assignedTerms {
            ...assignedTermsFields
        }
    }
}
```

Get entity with classifications

- with all classifications

JSON

```
query {
    entities(where: { guid: "<guid>" }) {
        guid
```

```
        typeName
        attributes
        classifications {
            typeName
            attributes
        }
    }
}
```

- with filtered classifications

JSON

```
query {
    entities(where: { guid: "<guid>" }) {
        guid
        typeName
        attributes
        classifications(where: { type: { typeName: "<classificationType>" } }) {
            typeName
            attributes
        }
    }
}
```

Get entity with related entities

- with all related entities

JSON

```
query {
    entities(where: { guid: "<guid>" }) {
        guid
        typeName
        attributes
        relatedEntities {
            relationshipAttributeName
            relationship {
                guid
                typeName
                attributes
            }
            entity {
                guid
                typeName
                qualifiedName
                attributes
```

```
            }
        }
    }
}
```

- with filtered related entities

JSON

```
query {
    entities(where: { guid: "<guid>" }) {
        guid
        typeName
        attributes
        relatedEntities(where: { relationshipAttributeName: "<relAttrName1>" }) {
            entity {
                guid
                typeName
                qualifiedName
                attributes
            }
        }
    }
}
```

- with filtered related entities, separated by aliases

JSON

```
query {
    entities(where: { guid: "<guid>" }) {
        guid
        typeName
        attributes
        alias1: relatedEntities(where: { relationshipAttributeName: "
<relAttrName1>" }) {
            entity {
                guid
                typeName
                attributes
            }
        }
        alias2: relatedEntities(where: { relationshipAttributeName: "
<relAttrName2>" }) {
            entity {
                guid
                typeName
                attributes
            }
```

```
        }
    }
}
```

Get entity with glossary terms

- with all glossary terms

JSON

```
query {
    entities(where: { guid: "<guid>" }) {
        guid
        typeName
        attributes
        assignedTerms {
            confidence
            createdBy
            description
            expression
            steward
            source
            status
            term {
                qualifiedName
                name
                shortDescription
                longDescription
            }
        }
    }
}
```

- with filtered glossary terms

JSON

```
query {
    entities(where: { guid: "<guid>" }) {
        guid
        typeName
        attributes
        assignedTerms {
            confidence
            createdBy
            description
            expression
            steward
```

```
            source
            status
            term {
                qualifiedName
                name
                shortDescription
                longDescription
            }
        }
    }
}
```

Filtering (preview)

The performance of exact matching for 'GUID' and 'Qualified-Name' is guaranteed in the examples provided in the 'basic queries' section. However, there are some limitations for other filtering patterns:

- **Filtering on Non-indexed Fields**: Fields other than GUID & qualified name are currently not indexed (as examples in the "Simple filter" section). Filtering over nonindexed fields without criteria on GUID/Qualified-name will result in a table scan and might cause performance issues on large datasets.
- **Nested Filtering**: Similar to nonindexed fields, nested filtering can cause table scans, which might cause performance issues on large datasets. For example, finding an entity with a linked classification/term/related entity.

Despite these limitations, this call pattern is superior to client-side filtering and is currently used by our internal client.

Simple filter

Query entities by type and system attributes

JSON

```
query {
    entities(
        where: {
            type: { typeName: "<entityType>" }
            name: { eq: "<value>" }
            createTime: { timerange: LAST_7D }
            updateTime: { gt: "<timestamp>" }
        }
    ) {
        guid
        typeName
```

```
        qualifiedName
        attributes
    }
}
```

Query entities by type and attributes

JSON

```
query {
    entities(
        where: {
            type: { typeName: "<entityType>" }
            attribute: { field: "<attrName>",operator: eq, value: "<attrValue>" }
        }
    ) {
        guid
        typeName
        qualifiedName
        attributes
    }
}
```

Query entities by type and business attributes

JSON

```
query {
    entities(
        where: {
            type: { typeName: "<entityType>" }
            businessAttribute: { field: "<BusinessMetadataName>.
<BusinessAttributeName>", operator: eq, value: "<BizAttrValue>" }
        }
    ) {
        guid
        typeName
        qualifiedName
        attributes
    }
}
```

Query entities by collection

Currently, this query pattern isn't supported for subassets.

JSON

```
query {
    entities(
        where: {
            type: { typeName: "<entityType>" }
            collectionID: "<collectionId>"
        }
    ) {
        guid
        typeName
        qualifiedName
        attributes
    }
}
```

Filter combination

Keys in an object(map) are combined as a logical AND.

JSON

```
query {
    entities(
        where: {
            type: { typeName: "<entityType>" }
            or: [
                {
                    and: [
                        { attribute: { field: "<attrName1>", value: "<attrValue1>"
} }
                        { attribute: { field: "<attrName2>", value: "<attrValue2>"
} }
                    ]
                }
                {
                    not: {
                        businessAttribute: {
                            field: "<BusinessMetadataName>.
<BusinessAttributeName>", value: "<BizAttrValue>"
                        }
                    }
                }
            ]
        }
    ) {
        guid
        typeName
        qualifiedName
        attributes
    }
}
```

Nested filter

Currently, this query pattern isn't supported for subassets.

Query entity by classification

JSON

```
query {
    entities(
        where: {
            classification: { type: { typeName: "<classificationType>",
includeSubTypes: true } }
        }
    ) {
        guid
        typeName
        qualifiedName
        attributes
    }
}
```

Query entity by related entity

JSON

```
query {
    entities(
        where: {
            relatedEntity: {
                relationshipAttributeName: "<relAttrName>"
                entity: {
                    type: { typeName: "<entityType>" }
                }
            }
        }
    ) {
        guid
        typeName
        qualifiedName
        attributes
    }
}
```

Query entity by glossary terms

JSON

```
query {
    entities(
        where: {
            assignedTerm: {
                term: {
                    qualifiedName: { eq: "<termName>" }
                }
            }
        }
    ) {
        guid
        typeName
        qualifiedName
        attributes
    }
}
```

Other queries

Get relationship

JSON

```
query {
    relationships(where: { guid: "<guid>" }) {
        guid
        typeName
        attributes
        end1 {
            guid
            typeName
            qualifiedName
            attributes
        }
        end2 {
            guid
            typeName
            qualifiedName
            attributes
        }
    }
}
```

Sample Response:

JSON

```
{
    "data": {
        "relationships": [
            {
                "guid": "...",
                "typeName": "ExampleRelationship",
                "attributes": {},
                "end1": {
                    "guid": "...",
                    "typeName": "column",
                    "qualifiedName": "...",
                    "attributes": {}
                },
                "end2": {
                    "guid": "...",
                    "typeName": "column",
                    "qualifiedName": "...",
                    "attributes": {}
                }
            }
        ]
    }
}
```

Lineage

Get dataset with downstream datasets

This query returns downstream datasets to 2 degrees.

JSON

```
query {
    entities(where: { guid: "<guid>" }) {#dataset
        guid
        typeName
        qualifiedName
        relatedEntities(where: { relationshipAttributeName: "inputToProcesses" })
{
            entity {#process
                guid
                typeName
                relatedEntities(where: { relationshipAttributeName: "outputs" }) {
                    entity {#dataset
                        guid
                        typeName
                        qualifiedName
                        relatedEntities(where: { relationshipAttributeName:
"inputToProcesses" }) {
                            entity {#process
```

```
                                guid
                                typeName
                                qualifiedName
                                relatedEntities(where: {
relationshipAttributeName: "outputs" }) {
                                    entity {#dataset
                                        guid
                                        typeName
                                        qualifiedName
                                    }
                                }
                            }
                        }
                    }
                }
            }
        }
    }
}
```

Sample Response:

JSON

```
{
    "data": {
        "entities": [
            {
                "guid": "...",
                "typeName": "Dataset",
                "qualifiedName": "...",
                "relatedEntities": [
                    {
                        "entity": {
                            "guid": "...",
                            "typeName": "Process",
                            "relatedEntities": [
                                {
                                    "entity": {
                                        "guid": "...",
                                        "typeName": "Dataset",
                                        "qualifiedName": "...",
                                        "relatedEntities": [
                                            {
                                                "entity": {
                                                    "guid": "...",
                                                    "typeName": "Process",
                                                    "relatedEntities": [
                                                        {
                                                            "entity": {
                                                                "guid": "...",
                                                                "typeName":
"Dataset",
```

```
                                                    "qualifiedName":
"..."
                                                        }
                                                    }
                                                ]
                                            }
                                        }
                                    ]
                                }
                            }
                        ]
                    }
                ]
            }
        }
    }
```

Limitations

Query fetch-depth limitation:

Queries are limited by query depth. The maximum depth is 5. For example, the following query will fail:

JSON

```
query {
    entities { #depth 1
        relatedEntities { #depth 2
            entity {
                relatedEntities { #depth 3
                    entity {
                        assignedTerms{ #depth 4
                            term {
                                classifications { #depth 5
                                    ...
                                }
                            }
                        }
                        relatedEntities { #depth 4
                            entity {
                                assignedTerms{ #depth 5
                                    term {
                                        classifications { #depth 6
                                            ...
                                        }
                                    }
                                }
```

```
                            }
                        }
                    }
                }
            }
        }
    }
}
```

Execution-time fetch-cost limitation:

Queries are constrained by the cost of query execution. The maximum allowable cost is set at 100 units.

The execution fetch cost is computed each time we aim to retrieve related entities, assigned terms, or classifications for a given entity or term.

Example

Consider a scenario where we query 3 entities, each with two related entities. The cost calculation would be as follows:

- One unit for the root query
- Three units for each level-1 entity

Hence, the total cost for this query would be `1 (root query) + 3 (level-1 entities) = 4`.

Filtering performance

Filtering is currently preview and has some limitations. See filtering for more information.

GraphQL queries begin with a root query that retrieves the top-level nodes. It then recursively fetches the related nodes.

The performance of a nested query is primarily determined by the root query because the related nodes are fetched from a known starting point, similar to a foreign key in SQL.

To optimize performance, it's crucial to avoid wildcard root queries that could trigger a table scan on the top-level nodes.

For instance, the following query could cause performance issues on large datasets because the `name` field isn't indexed:

JSON

```
query {
    entities(where: { name: { eq: "<value>" } }) {
        guid
        typeName
        attributes
        relatedEntities {
            entity {
                guid
                typeName
                attributes
            }
        }
    }
}
```

To prevent such performance issues, ensure that you apply a filter on `guid` or `qualifiedName`.

Next steps

Manage data sources | Microsoft Purview data plane REST APIs

Azure Purview SDK for .NET - latest

Packages - latest

Expand table

Reference	Package	Source
Resource Management - Purview	Azure.ResourceManager.Purview	GitHub

Last updated on 10/12/2026

Az.Purview Module

Microsoft Azure PowerShell: Purview cmdlets

Purview

Expand table

Cmdlet	Description
Add-AzPurviewAccountRootCollectionAdmin	Add the administrator for root collection associated with this account.
Get-AzPurviewAccount	Get an account
Get-AzPurviewAccountKey	List the authorization keys associated with this account.
Get-AzPurviewClassificationRule	Get a classification rule
Get-AzPurviewClassificationRuleVersion	Lists the rule versions of a classification rule
Get-AzPurviewDataSource	Get a data source
Get-AzPurviewDefaultAccount	Get the default account for the scope.
Get-AzPurviewFilter	Get a filter
Get-AzPurviewKeyVaultConnection	Gets key vault information
Get-AzPurviewScan	Gets a scan information
Get-AzPurviewScanResultScanHistory	Lists the scan history of a scan
Get-AzPurviewScanRuleset	Get a scan ruleset
Get-AzPurviewSystemScanRuleset	Get a system scan ruleset for a data source
Get-AzPurviewSystemScanRulesetLatest	Get the latest version of a system scan ruleset
Get-AzPurviewSystemScanRulesetVersion	List system scan ruleset versions in Data catalog
Get-AzPurviewTrigger	Gets trigger information
Invoke-AzPurviewTagClassificationRuleClassificationVersion	Sets Classification Action on a specific classificat rule version.
New-AzPurviewAccount	Creates or updates an account
New-AzPurviewAdlsGen1CredentialScanObject	Create an in-memory object for AdlsGen1CredentialScan.

Cmdlet	Description
New-AzPurviewAdlsGen1DataSourceObject	Create an in-memory object for AdlsGen1DataSource.
New-AzPurviewAdlsGen1MsiScanObject	Create an in-memory object for AdlsGen1MsiSc
New-AzPurviewAdlsGen1ScanRulesetObject	Create an in-memory object for AdlsGen1ScanRuleset.
New-AzPurviewAdlsGen2CredentialScanObject	Create an in-memory object for AdlsGen2CredentialScan.
New-AzPurviewAdlsGen2DataSourceObject	Create an in-memory object for AdlsGen2DataSource.
New-AzPurviewAdlsGen2MsiScanObject	Create an in-memory object for AdlsGen2MsiSc
New-AzPurviewAdlsGen2ScanRulesetObject	Create an in-memory object for AdlsGen2ScanRuleset.
New-AzPurviewAmazonAccountCredentialScanObject	Create an in-memory object for AmazonAccountCredentialScan.
New-AzPurviewAmazonAccountDataSourceObject	Create an in-memory object for AmazonAccountDataSource.
New-AzPurviewAmazonPostgreSqlCredentialScanObject	Create an in-memory object for AmazonPostgreSqlCredentialScan.
New-AzPurviewAmazonPostgreSqlDataSourceObject	Create an in-memory object for AmazonPostgreSqlDataSource.
New-AzPurviewAmazonPostgreSqlScanRulesetObject	Create an in-memory object for AmazonPostgreSqlScanRuleset.
New-AzPurviewAmazonS3CredentialScanObject	Create an in-memory object for AmazonS3CredentialScan.
New-AzPurviewAmazonS3DataSourceObject	Create an in-memory object for AmazonS3DataSource.
New-AzPurviewAmazonS3ScanRulesetObject	Create an in-memory object for AmazonS3ScanRuleset.
New-AzPurviewAmazonSqlCredentialScanObject	Create an in-memory object for AmazonSqlCredentialScan.
New-AzPurviewAmazonSqlDataSourceObject	Create an in-memory object for AmazonSqlDataSource.
New-AzPurviewAmazonSqlScanRulesetObject	Create an in-memory object for AmazonSqlScanRuleset.
New-AzPurviewAzureCosmosDbCredentialScanObject	Create an in-memory object for AzureCosmosDbCredentialScan.

Cmdlet	Description
New-AzPurviewAzureCosmosDbDataSourceObject	Create an in-memory object for AzureCosmosDbDataSource.
New-AzPurviewAzureCosmosDbScanRulesetObject	Create an in-memory object for AzureCosmosDbScanRuleset.
New-AzPurviewAzureDataExplorerCredentialScanObject	Create an in-memory object for AzureDataExplorerCredentialScan.
New-AzPurviewAzureDataExplorerDataSourceObject	Create an in-memory object for AzureDataExplorerDataSource.
New-AzPurviewAzureDataExplorerMsiScanObject	Create an in-memory object for AzureDataExplorerMsiScan.
New-AzPurviewAzureDataExplorerScanRulesetObject	Create an in-memory object for AzureDataExplorerScanRuleset.
New-AzPurviewAzureFileServiceCredentialScanObject	Create an in-memory object for AzureFileServiceCredentialScan.
New-AzPurviewAzureFileServiceDataSourceObject	Create an in-memory object for AzureFileServiceDataSource.
New-AzPurviewAzureFileServiceScanRulesetObject	Create an in-memory object for AzureFileServiceScanRuleset.
New-AzPurviewAzureKeyVaultObject	Create an in-memory object for AzureKeyVault.
New-AzPurviewAzureMySqlCredentialScanObject	Create an in-memory object for AzureMySqlCredentialScan.
New-AzPurviewAzureMySqlDataSourceObject	Create an in-memory object for AzureMySqlDataSource.
New-AzPurviewAzureMySqlScanRulesetObject	Create an in-memory object for AzureMySqlScanRuleset.
New-AzPurviewAzurePostgreSqlCredentialScanObject	Create an in-memory object for AzurePostgreSqlCredentialScan.
New-AzPurviewAzurePostgreSqlDataSourceObject	Create an in-memory object for AzurePostgreSqlDataSource.
New-AzPurviewAzurePostgreSqlScanRulesetObject	Create an in-memory object for AzurePostgreSqlScanRuleset.
New-AzPurviewAzureResourceGroupCredentialScanObject	Create an in-memory object for AzureResourceGroupCredentialScan.
New-AzPurviewAzureResourceGroupDataSourceObject	Create an in-memory object for AzureResourceGroupDataSource.
New-AzPurviewAzureResourceGroupMsiScanObject	Create an in-memory object for AzureResourceGroupMsiScan.

Cmdlet	Description
New-AzPurviewAzureSqlDatabaseCredentialScanObject	Create an in-memory object for AzureSqlDatabaseCredentialScan.
New-AzPurviewAzureSqlDatabaseDataSourceObject	Create an in-memory object for AzureSqlDatabaseDataSource.
New-AzPurviewAzureSqlDatabaseManagedInstanceCredentialScanObject	Create an in-memory object for AzureSqlDatabaseManagedInstanceCredentialS
New-AzPurviewAzureSqlDatabaseManagedInstanceDataSourceObject	Create an in-memory object for AzureSqlDatabaseManagedInstanceDataSource
New-AzPurviewAzureSqlDatabaseManagedInstanceMsiScanObject	Create an in-memory object for AzureSqlDatabaseManagedInstanceMsiScan.
New-AzPurviewAzureSqlDatabaseManagedInstanceScanRulesetObject	Create an in-memory object for AzureSqlDatabaseManagedInstanceScanRuleset
New-AzPurviewAzureSqlDatabaseMsiScanObject	Create an in-memory object for AzureSqlDatabaseMsiScan.
New-AzPurviewAzureSqlDatabaseScanRulesetObject	Create an in-memory object for AzureSqlDatabaseScanRuleset.
New-AzPurviewAzureSqlDataWarehouseCredentialScanObject	Create an in-memory object for AzureSqlDataWarehouseCredentialScan.
New-AzPurviewAzureSqlDataWarehouseDataSourceObject	Create an in-memory object for AzureSqlDataWarehouseDataSource.
New-AzPurviewAzureSqlDataWarehouseMsiScanObject	Create an in-memory object for AzureSqlDataWarehouseMsiScan.
New-AzPurviewAzureSqlDataWarehouseScanRulesetObject	Create an in-memory object for AzureSqlDataWarehouseScanRuleset.
New-AzPurviewAzureStorageCredentialScanObject	Create an in-memory object for AzureStorageCredentialScan.
New-AzPurviewAzureStorageDataSourceObject	Create an in-memory object for AzureStorageDataSource.
New-AzPurviewAzureStorageMsiScanObject	Create an in-memory object for AzureStorageMsiScan.
New-AzPurviewAzureStorageScanRulesetObject	Create an in-memory object for AzureStorageScanRuleset.
New-AzPurviewAzureSubscriptionCredentialScanObject	Create an in-memory object for AzureSubscriptionCredentialScan.
New-AzPurviewAzureSubscriptionDataSourceObject	Create an in-memory object for AzureSubscriptionDataSource.

Cmdlet	Description
New-AzPurviewAzureSubscriptionMsiScanObject	Create an in-memory object for AzureSubscriptionMsiScan.
New-AzPurviewAzureSynapseWorkspaceCredentialScanObject	Create an in-memory object for AzureSynapseWorkspaceCredentialScan.
New-AzPurviewAzureSynapseWorkspaceDataSourceObject	Create an in-memory object for AzureSynapseWorkspaceDataSource.
New-AzPurviewAzureSynapseWorkspaceMsiScanObject	Create an in-memory object for AzureSynapseWorkspaceMsiScan.
New-AzPurviewAzureSynapseWorkspaceScanRulesetObject	Create an in-memory object for AzureSynapseWorkspaceScanRuleset.
New-AzPurviewClassificationRule	Create a classification rule
New-AzPurviewCustomClassificationRuleObject	Create an in-memory object for CustomClassificationRule.
New-AzPurviewDataSource	Create a data source
New-AzPurviewFilter	Create a filter
New-AzPurviewFilterObject	Create an in-memory object for Filter.
New-AzPurviewKeyVaultConnection	Create an instance of a key vault connection
New-AzPurviewOracleDataSourceObject	Create an in-memory object for OracleDataSour
New-AzPurviewPowerBIDataSourceObject	Create an in-memory object for PowerBIDataSource.
New-AzPurviewPowerBIDelegatedScanObject	Create an in-memory object for PowerBIDelegatedScan.
New-AzPurviewPowerBIMsiScanObject	Create an in-memory object for PowerBIMsiScar
New-AzPurviewRegexClassificationRulePatternObject	Create an in-memory object for RegexClassificationRulePattern.
New-AzPurviewSapEccDataSourceObject	Create an in-memory object for SapEccDataSou
New-AzPurviewSapS4HanaDataSourceObject	Create an in-memory object for SapS4HanaDataSource.
New-AzPurviewSapS4HanaSapS4HanaCredentialScanObject	Create an in-memory object for SapS4HanaSapS4HanaCredentialScan.
New-AzPurviewScan	Create an instance of a scan
New-AzPurviewScanRuleset	Create a scan ruleset
New-AzPurviewSqlServerDatabaseCredentialScanObject	Create an in-memory object for SqlServerDatabaseCredentialScan.

Cmdlet	Description
New-AzPurviewSqlServerDatabaseDataSourceObject	Create an in-memory object for SqlServerDatabaseDataSource.
New-AzPurviewSqlServerDatabaseScanRulesetObject	Create an in-memory object for SqlServerDatabaseScanRuleset.
New-AzPurviewTeradataDataSourceObject	Create an in-memory object for TeradataDataSource.
New-AzPurviewTrigger	Create an instance of a trigger
New-AzPurviewTriggerObject	Create an in-memory object for Trigger.
Remove-AzPurviewAccount	Deletes an account resource
Remove-AzPurviewClassificationRule	Deletes a classification rule
Remove-AzPurviewDataSource	Deletes a data source
Remove-AzPurviewDefaultAccount	Removes the default account from the scope.
Remove-AzPurviewKeyVaultConnection	Deletes the key vault connection associated with the account
Remove-AzPurviewScan	Deletes the scan associated with the data source
Remove-AzPurviewScanRuleset	Deletes a scan ruleset
Remove-AzPurviewTrigger	Deletes the trigger associated with the scan
Set-AzPurviewDefaultAccount	Sets the default account for the scope.
Start-AzPurviewScanResultScan	Runs the scan
Stop-AzPurviewScanResultScan	Cancels a scan
Test-AzPurviewAccountNameAvailability	Checks if account name is available.
Update-AzPurviewAccount	Update an account
Update-AzPurviewClassificationRule	Update a classification rule
Update-AzPurviewDataSource	Update a data source
Update-AzPurviewFilter	Update a filter
Update-AzPurviewKeyVaultConnection	Update an instance of a key vault connection
Update-AzPurviewScan	Update an instance of a scan
Update-AzPurviewScanRuleset	Update a scan ruleset
Update-AzPurviewTrigger	Update an instance of a trigger

purview Package

Packages

Expand table

aio
models
operations

Classes

Expand table

PurviewManagementClient	Creates a Microsoft.Purview management client.

Microsoft Purview Resource Provider Rest API

Article • 10/31/2023

Microsoft Purview Microsoft Purview is a unified data governance service that helps you manage and govern your on-premises, multi-cloud, and software-as-a-service (SaaS) data. Easily create a holistic, up-to-date map of your data landscape with automated data discovery, sensitive data classification, and end-to-end data lineage. Empower data consumers to find valuable, trustworthy data. Manage user access to data in your estate securely and at scale.

Authorization Token

All Azure Resource Manager and Azure REST API calls require a valid authorization token in the request header to succeed. See Create the request: Acquire an access token for details on how to obtain this authorization token.

See Also

- Microsoft Purview

Type definitions and how to create custom types

Article • 05/14/2025

This tutorial will explain what type definitions are, how to create custom types, and how to initialize assets of custom types in Microsoft Purview.

In this tutorial, you'll learn:

- ✓ How Microsoft Purview uses the *type system* from *Apache Atlas*
- ✓ How to create a new custom type
- ✓ How to create relationships between custom types
- ✓ How to initialize new entities of custom types

Prerequisites

For this tutorial you'll need:

- An Azure account with an active subscription. If you don't have one, you can create an account for free.
- An active Microsoft Purview (formerly Azure Purview) account. If you don't have one, see the guide for getting started with Microsoft Purview for free.
- A bearer token to your Microsoft Purview account. To establish a bearer token and to call any APIs, see the documentation about how to authenticate APIs for Microsoft Purview.
- Your Apache Atlas endpoint of your Microsoft Purview account. It will be one of these two endpoints, depending on your environment:
  - If you're using the classic Microsoft Purview governance portal: `https://{{ACCOUNTNAME}}.purview.azure.com/catalog`
  - If you're using the new Microsoft Purview portal: `https://api.purview-service.microsoft.com/catalog`

> **Note**
>
> Before moving to the hands-on part of the tutorial, the first four sections will explain what a System Type is and how it is used in Microsoft Purview. All the REST API calls described further will use the **bearer token** and the **endpoint** which are described in the prerequisites.
>
> To skip directly to the steps, use these links:

- Create custom type definitions
- Initialize assets of custom types

What are *asset* and *type* in Microsoft Purview?

An *asset* is a metadata element that describes a digital or physical resource. The digital or physical resources that are expected to be cataloged as assets include:

- Data sources such as databases, files, and data feed.
- Analytical models and processes.
- Business policies and terms.
- Infrastructure like the server.

Microsoft Purview provides users a flexible *type system* to expand the definition of the asset to include new kinds of resources as they become relevant. Microsoft Purview relies on the Type System from Apache Atlas. All metadata objects (assets) managed by Microsoft Purview are modeled using type definitions. Understanding the Type System is fundamental to create new custom types in Microsoft Purview.

Essentially, a *Type* can be seen as a *Class* from Object Oriented Programming (OOP):

- It defines the properties that represent that type.
- Each type is uniquely identified by its *name*.
- A *type* can inherit from a *superType*. This is an equivalent concept as inheritance from OOP. A type that extends a superType will inherit the attributes of the superType.

You can see all type definitions in your Microsoft Purview account by sending a `GET` request to the All Type Definitions endpoint:

```
HTTP

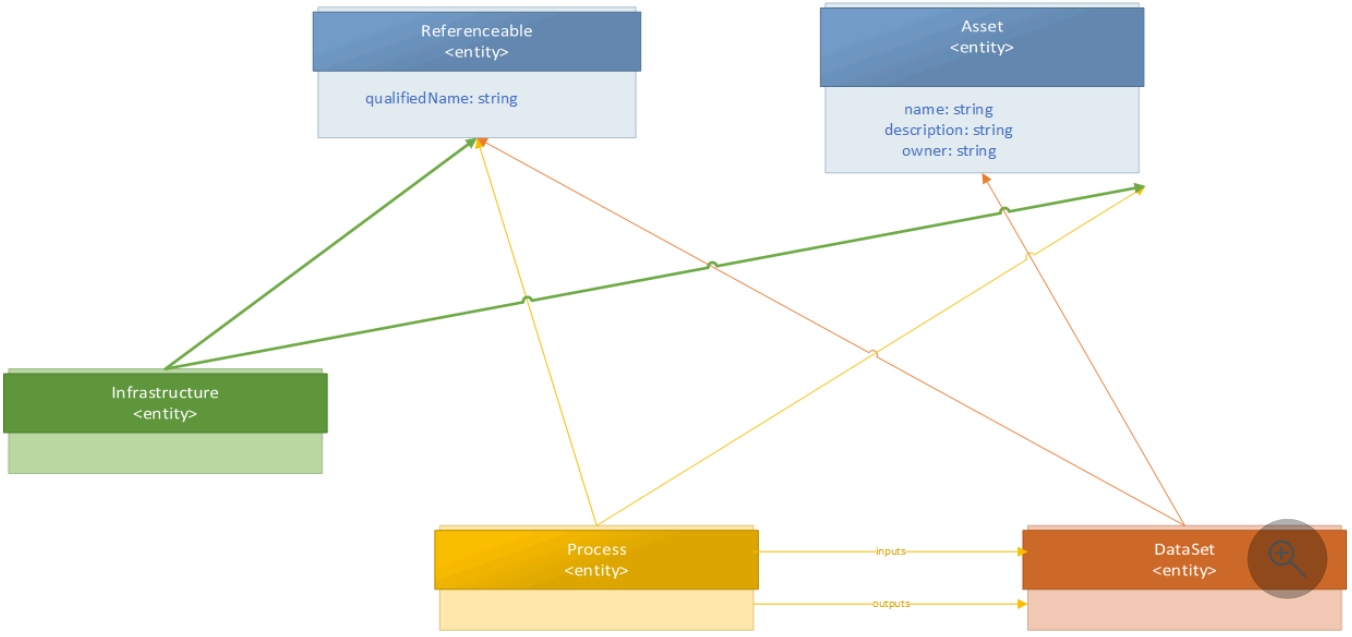
GET https://{{ENDPOINT}}/catalog/api/atlas/v2/types/typedefs
```

Apache Atlas has few predefined system types that are commonly used as supertypes.

For example:

- **Referenceable**: This type represents all entities that can be searched for using a unique attribute called *qualifiedName*.
- **Asset**: This type extends from Referenceable and has other attributes such as: *name*, *description* and *owner*.

- **DataSet**: This type extends Referenceable and Asset. Conceptually, it can be used to represent a type that stores data. Types that extend DataSet can be expected to have a Schema. For example, a SQL table.
- **Lineage**: Lineage information helps one understand the origin of data and the transformations it may have gone through before arriving in a file or table. Lineage is calculated through *DataSet* and *Process*: DataSets (input of process) impact some other DataSets (output of process) through Process.



Example of a *Type* definition

To better understand the Type system, let's look at an example and see how an **Azure SQL Table** is defined.

You can get the complete type definition by sending a `GET` request to the Type Definition endpoint:

HTTP

```
GET https://{{ENDPOINT}}/catalog/api/atlas/v2/types/typedef/name/{name}
```

> **Tip**
>
> The **{name}** property tells which definition you are interested in. In this case, you should use **azure_sql_table**.

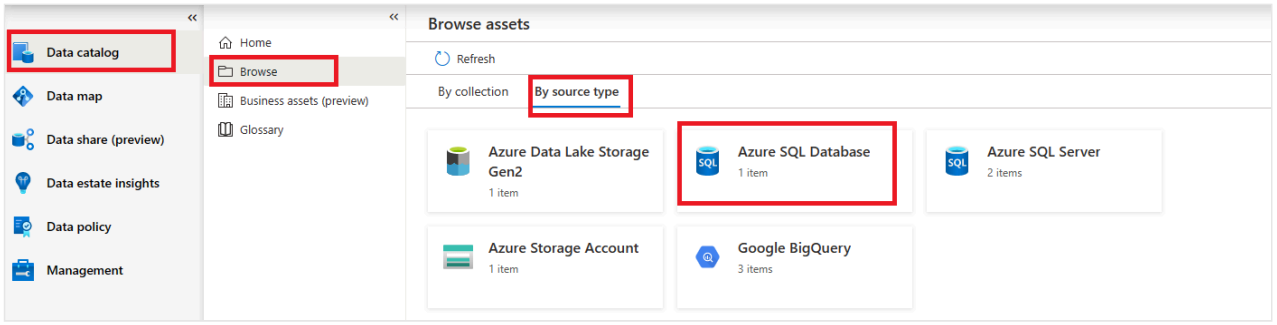
Below you can see a simplified JSON result:

JSON

```
{
  "category": "ENTITY",
  "guid": "7d92a449-f7e8-812f-5fc8-ca6127ba90bd",
  "name": "azure_sql_table",
  "description": "azure_sql_table",
  "typeVersion": "1.0",
  "serviceType": "Azure SQL Database",
  "options": {
    "schemaElementsAttribute": "columns",
  },
  "attributeDefs": [
    { "name": "principalId", ...},
    { "name": "objectType", ...},
    { "name": "createTime", ...},
    { "name": "modifiedTime", ... }
  ],
  "superTypes": [
    "DataSet",
    "Purview_Table",
    "Table"
  ],
  "subTypes": [],
  "relationshipAttributeDefs": [
    {
      "name": "dbSchema",
      "typeName": "azure_sql_schema",
      "isOptional": false,
      "cardinality": "SINGLE",
      "relationshipTypeName": "azure_sql_schema_tables",
    },
    {
      "name": "columns",
      "typeName": "array<azure_sql_column>",
      "isOptional": true,
      "cardinality": "SET",
      "relationshipTypeName": "azure_sql_table_columns",
    },
  ]
}
```

Based on the JSON type definition, let's look at some properties:

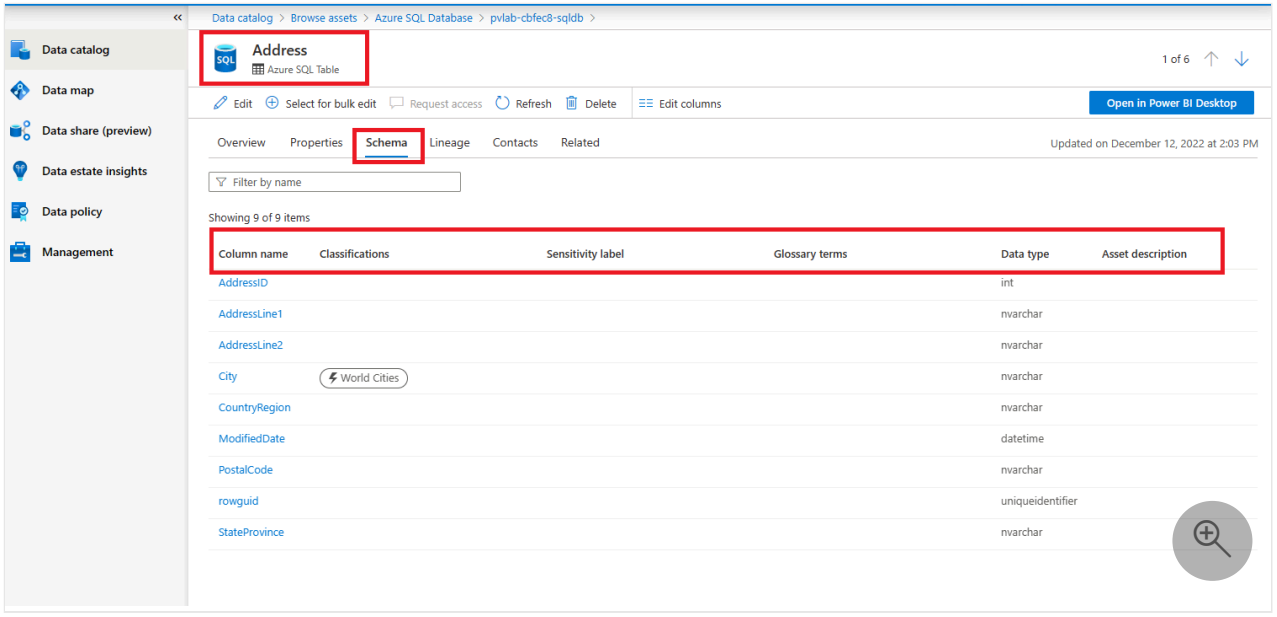
- **Category** field describes in what category your type is. The list of categories supported by Apache Atlas can be found here.
- **ServiceType** field is useful when browsing assets *by source type* in Microsoft Purview. The *service type* will be an entry point to find all assets that belong to the same *service type* - as defined on their type definition. In the below screenshot of Purview UI, the user limits the result to be the entities specified with *Azure SQL Database* in **serviceType**:



> **Note**
>
> **Azure SQL Database** is defined with the same *serviceType* as **Azure SQL Table**.

- **SuperTypes** describes the "*parent*" types you want to "*inherit*" from.
- **schemaElementsAttributes** from **options** influences what appears in the **Schema** tab of your asset in Microsoft Purview.

  Below you can see an example of how the **Schema** tab looks like for an asset of type Azure SQL Table:



- **relationshipAttributeDefs** are calculated through the relationship type definitions. In our JSON, we can see that **schemaElementsAttributes** points to the relationship attribute called **columns** - which is one of elements from **relationshipAttributeDefs** array, as shown below:

```JSON
...
"relationshipAttributeDefs": [
    ...
    {
```

```
        "name": "columns",
        "typeName": "array<azure_sql_column>",
        "isOptional": true,
        "cardinality": "SET",
        "relationshipTypeName": "azure_sql_table_columns",
      },
    ]
```

Each relationship has its own definition. The name of the definition is found in **relationshipTypeName** attribute. In this case, it's *azure_sql_table_columns*.

- The **cardinality** of this relationship attribute is set to *SET, which suggests that it holds a list of related assets.
- The related asset is of type *azure_sql_column*, as visible in the *typeName* attribute.

In other words, the *columns* relationship attribute relates the Azure SQL Table to a list of Azure SQL Columns that show up in the Schema tab.

Example of a *relationship Type definition*

Each relationship consists of two ends, called *endDef1* and *endDef2*.

In the previous example, *azure_sql_table_columns* was the name of the relationship that characterizes a table (endDef1) and its columns (endDef2).

For the full definition, you can do make a `GET` request to the following endpoint using *azure_sql_table_columns* as the name:

HTTP

```
GET
https://{{ENDPOINT}}/catalog/api/atlas/v2/types/typedef/name/azure_sql_table_colum
ns
```

Below you can see a simplified JSON result:

JSON

```
{
  "category": "RELATIONSHIP",
  "guid": "c80d0027-8f29-6855-6395-d243b37d8a93",
  "name": "azure_sql_table_columns",
  "description": "azure_sql_table_columns",
  "serviceType": "Azure SQL Database",
  "relationshipCategory": "COMPOSITION",
  "endDef1": {
    "type": "azure_sql_table",
    "name": "columns",
```

```
    "isContainer": true,
    "cardinality": "SET",
  },
  "endDef2": {
    "type": "azure_sql_column",
    "name": "table",
    "isContainer": false,
    "cardinality": "SINGLE",
  }
}
```

- **name** is the name of the relationship definition. The value, in this case *azure_sql_table_columns* is used in the *relationshipTypeName* attribute of the entity that has this relationship, as you can see it referenced in the json.
- **relationshipCategory** is the category of the relationship and it can be either COMPOSITION, AGGREGATION or ASSOCIATION as described here.
- **enDef1** is the first end of the definition and contains the attributes:
  - **type** is the type of the entity that this relationship expects as end1.
  - **name** is the attribute that will appear on this entity's relationship attribute.
  - **cardinality** is either SINGLE, SET or LIST.
  - **isContainer** is a boolean and applies to containment relationship category. When set to true in one end, it indicates that this end is the container of the other end. Therefore:
    - Only the *Composition* or *Aggregation* category relationships can and should have in one end *isContainer* set to true.
    - *Association* category relationship shouldn't have *isContainer* property set to true in any end.
- **endDef2** is the second end of the definition and describes, similarly to endDef1, the properties of the second part of the relationship.

Schema tab

What is *Schema* in Microsoft Purview?

Schema is an important concept that reflects how data is stored and organized in the data store. It reflects the structure of the data and the data restrictions of the elements that construct the structure.

Elements on the same schema can be classified differently (due to their content). Also, different transformation (lineage) can happen to only a subset of elements. Due to these aspects, Purview can model schema and schema elements **as entities**, hence schema is usually a relationship attribute to the data asset entity. Examples of schema elements are: **columns** of a table, **json properties** of json schema, **xml elements** of xml schema etc.

There are two types of schemas:

- **Intrinsic Schema** - Some systems are intrinsic to schema. For example, when you create a SQL Table, the system requires you to define the columns that construct the table; in this sense, schema of a table is reflected by its columns.

  For data store with predefined schema, Purview uses the corresponding relationship between the data asset and the schema elements to reflect the schema. This relationship attribute is specified by the keyword **schemaElementsAttribute** in **options** property of the entity type definition.

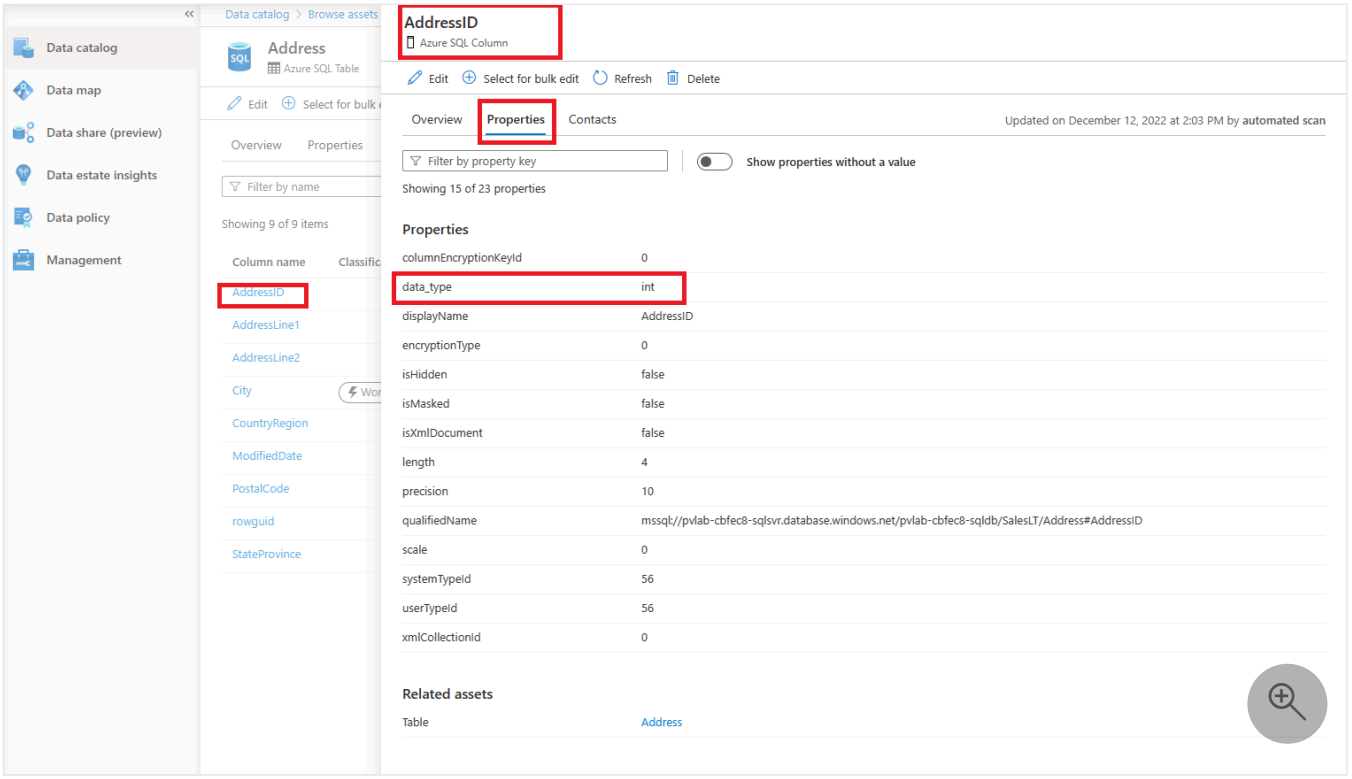
- **Non Intrinsic Schema** - Some systems don't enforce such schema restrictions, but users can use it to store structural data by applying some schema protocols to the data. For example, Azure Blobs store binary data and don't care about the data in the binary stream. Therefore, it's unaware of any schema, but the user can serialize their data with schema protocols like json before storing it in the blob. In this sense, schema is maintained by some extra protocols and corresponding validation enforced by the user.

  For data store without inherent schema, schema model is independent of this data store. For such cases, Purview defines an interface for schema and a relationship between DataSet and schema, called **dataset_attached_schemas** - this extends any entity type that inherits from DataSet to have an **attachedSchema** relationship attribute to link to their schema representation.

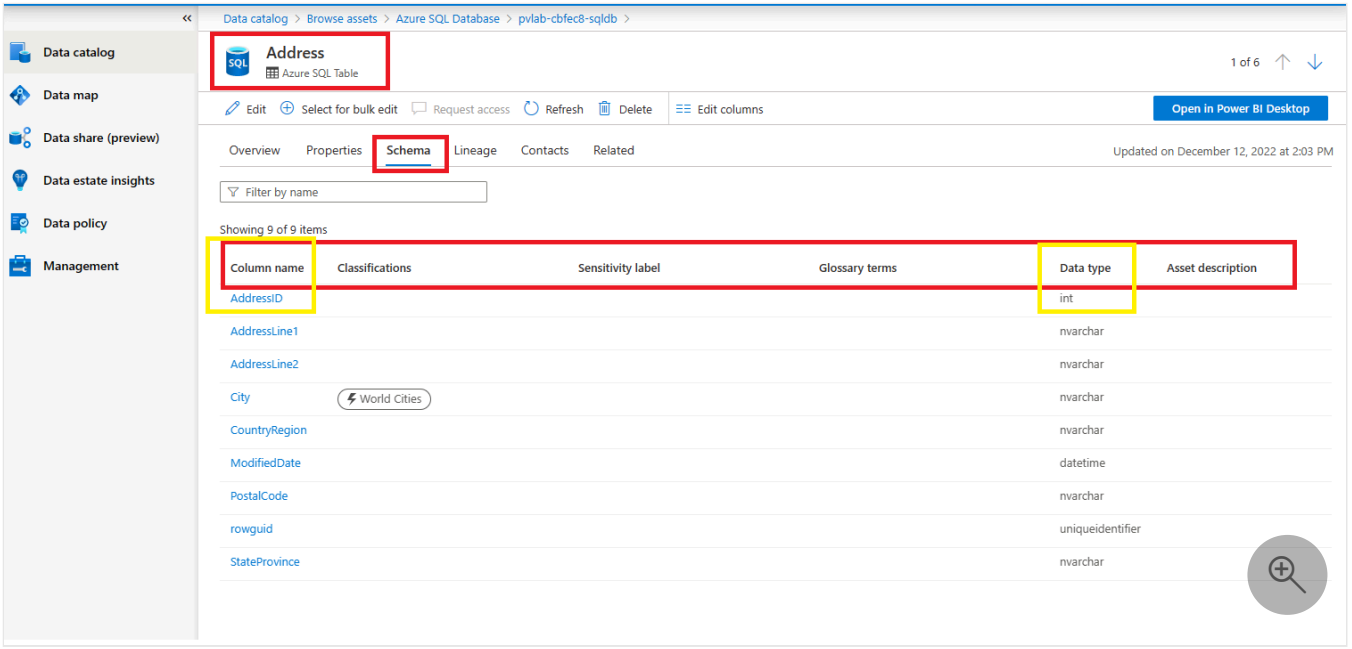
Example of *Schema tab*

The Azure SQL Table example from above has an intrinsic schema. The information that shows up in the Schema tab of the Azure SQL Table comes from the Azure SQL Column themselves.

Selecting one column item, we would see the following:



The question is, how did Microsoft Purview select the *data_tye* property from the column and showed it in the Schema tab of the table?



You can get the type definition of an Azure SQL Column by making a `GET` request to the endpoint:

HTTP

```
GET https://{{ENDPOINT}}/catalog/api/atlas/v2/types/typedef/name/{name}
```

> ⓘ **Note**

{name} in this case is: azure_sql_column

Here's a simplified JSON result:

JSON

```
{
  "category": "ENTITY",
  "guid": "58034a18-fc2c-df30-e474-75803c3a8957",
  "name": "azure_sql_column",
  "description": "azure_sql_column",
  "serviceType": "Azure SQL Database",
  "options": {
    "schemaAttributes": "[\"data_type\"]"
  },
  "attributeDefs":
  [
    {
      "name": "data_type",
      "typeName": "string",
      "isOptional": false,
      "cardinality": "SINGLE",
      "valuesMinCount": 1,
      "valuesMaxCount": 1,
      "isUnique": false,
      "isIndexable": false,
      "includeInNotification": false
    },
  ...
  ]
  ...
}
```

> **Note**
>
> serviceType is Azure SQL Database, the same as for the table

- *schemaAttributes* is set to **data_type**, which is one of the attributes of this type.

Azure SQL Table used *schemaElementAttribute* to point to a relationship consisting of a list of Azure SQL Columns. The type definition of a column has *schemaAttributes* defined.

In this way, the Schema tab in the table displays the attribute(s) listed in the *schemaAttributes* of the related assets.

Create custom type definitions

Why?

First, why would someone like to create a custom type definition?

There can be cases where there's no built-in type that corresponds to the structure of the metadata you want to import in Microsoft Purview.

In such a case, a new type definition has to be defined.

> **Note**
>
> The usage of built-in types should be favored over the creation of custom types, whenever possible.

Now that we gained an understanding of type definitions in general, let us create custom type definitions.

Scenario

In this tutorial, we would like to model a 1:n relationship between two types, called *custom_type_parent* and *custom_type_child*.

A *custom_type_child* should reference one parent, whereas a *custom_type_parent* can reference a list of children.

They should be linked together through a 1:n relationship.

> **Tip**
>
> Here you can find few tips when creating a new custom type.

Create definitions

1. Create the *custom_type_parent* type definition by making a `POST` request to one of the two following endpoints:

Classic Microsoft Purview governance portal:

```HTTP
POST https://{{ENDPOINT}}.purview.azure.com/catalog/api/atlas/v2/types/typedefs
```

New Microsoft Purview portal:

```HTTP
POST https://api.purview-service.microsoft.com/catalog/api/atlas/v2/types/typedefs
```

With the body:

```JSON
{
   "entityDefs":
   [
       {
           "category": "ENTITY",
           "version": 1,
           "name": "custom_type_parent",
           "description": "Sample custom type of a parent object",
           "typeVersion": "1.0",
           "serviceType": "Sample-Custom-Types",
           "superTypes": [
               "DataSet"
           ],
           "subTypes": [],
           "options":{
               "schemaElementsAttribute": "columns"
           }
       }
   ]
}
```

1. Create the *custom_type_child* type definition by making a `POST` request to one of the two following endpoints:

Classic Microsoft Purview governance portal:

```HTTP
POST https://{{ENDPOINT}}.purview.azure.com/catalog/api/atlas/v2/types/typedefs
```

New Microsoft Purview portal:

```HTTP
POST https://api.purview-service.microsoft.com/catalog/api/atlas/v2/types/typedefs
```

With the body:

JSON

```
{
   "entityDefs":
   [
       {
           "category": "ENTITY",
           "version": 1,
           "name": "custom_type_child",
           "description": "Sample custom type of a CHILD object",
           "typeVersion": "1.0",
           "serviceType": "Sample-Custom-Types",
           "superTypes": [
               "DataSet"
           ],
           "subTypes": [],
           "options":{
              "schemaAttributes": "data_type"
           }
       }
   ]
}
```

1. Create a custom type relationship definition by making a `POST` request to one of the two following endpoints:

Classic Microsoft Purview governance portal:

HTTP

```
POST https://{{ENDPOINT}}.purview.azure.com/catalog/api/atlas/v2/types/typedefs
```

New Microsoft Purview portal:

HTTP

```
POST https://api.purview-service.microsoft.com/catalog/api/atlas/v2/types/typedefs
```

With the body:

JSON

```
{
    "relationshipDefs": [
        {
            "category": "RELATIONSHIP",
            "endDef1" : {
                "cardinality" : "SET",
                "isContainer" : true,
```

```
                "name" : "Children",
                "type" : "custom_type_parent"
            },
            "endDef2" : {
                "cardinality" : "SINGLE",
                "isContainer" : false,
                "name" : "Parent",
                "type" : "custom_type_child"
            },
            "relationshipCategory" : "COMPOSITION",
            "serviceType": "Sample-Custom-Types",
            "name": "custom_parent_child_relationship"
        }
    ]
}
```

Initialize assets of custom types

1. Initialize a new asset of type *custom_type_parent* by making a `POST` request to one of the two following endpoints:

Classic Microsoft Purview governance portal:

HTTP

```
POST https://{{ENDPOINT}}.purview.azure.com/catalog/api/atlas/v2/entity
```

New Microsoft Purview portal:

HTTP

```
POST https://api.purview-service.microsoft.com/catalog/api/atlas/v2/entity
```

With the body:

JSON

```
{
    "entity": {
        "typeName":"custom_type_parent",
        "status": "ACTIVE",
        "version": 1,
         "attributes":{
            "name": "First_parent_object",
            "description": "This is the first asset of type custom_type_parent",
            "qualifiedName": "custom//custom_type_parent:First_parent_object"
         }
```

```
    }
}
```

Save the *guid* as you'll need it later.

1. Initialize a new asset of type *custom_type_child* by making a `POST` request to one of the two following endpoints:

Classic Microsoft Purview governance portal:

HTTP

```
POST https://{{ENDPOINT}}.purview.azure.com/catalog/api/atlas/v2/entity
```

New Microsoft Purview portal:

HTTP

```
POST https://api.purview-service.microsoft.com/catalog/api/atlas/v2/entity
```

With the body:

JSON

```
{
   "entity": {
       "typeName":"custom_type_child",
       "status": "ACTIVE",
       "version": 1,
       "attributes":{
           "name": "First_child_object",
           "description": "This is the first asset of type custom_type_child",
           "qualifiedName": "custom//custom_type_child:First_child_object"
        }
   }
}
```

Save the *guid* as you'll need it later.

1. Initialize a new relationship of type *custom_parent_child_relationship* by making a `POST` request to one of the two following endpoints:

Classic Microsoft Purview governance portal:

HTTP

```
POST https://{{ENDPOINT}}.purview.azure.com/catalog/api/atlas/v2/relationship/
```

New Microsoft Purview portal:

HTTP

```
POST https://api.purview-service.microsoft.com/catalog/api/atlas/v2/relationship/
```

With the following body:

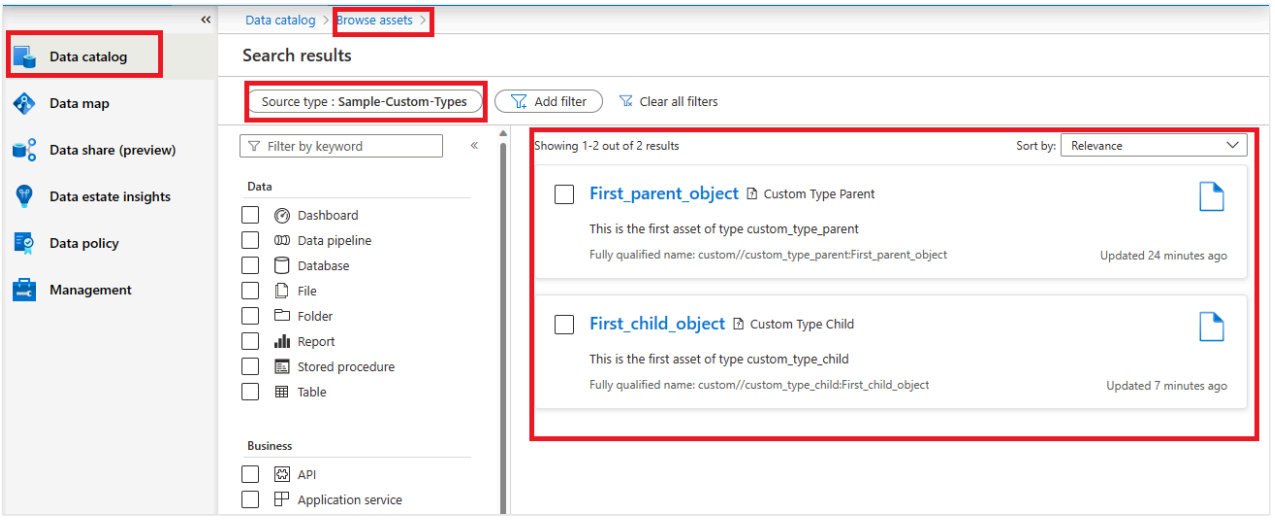
> **Note**
>
> The *guid* in end1 must be replaced with the the guid of the object created at step 6.1 The *guid* in end2 must be replaced with the guid of the object created at step 6.2

JSON

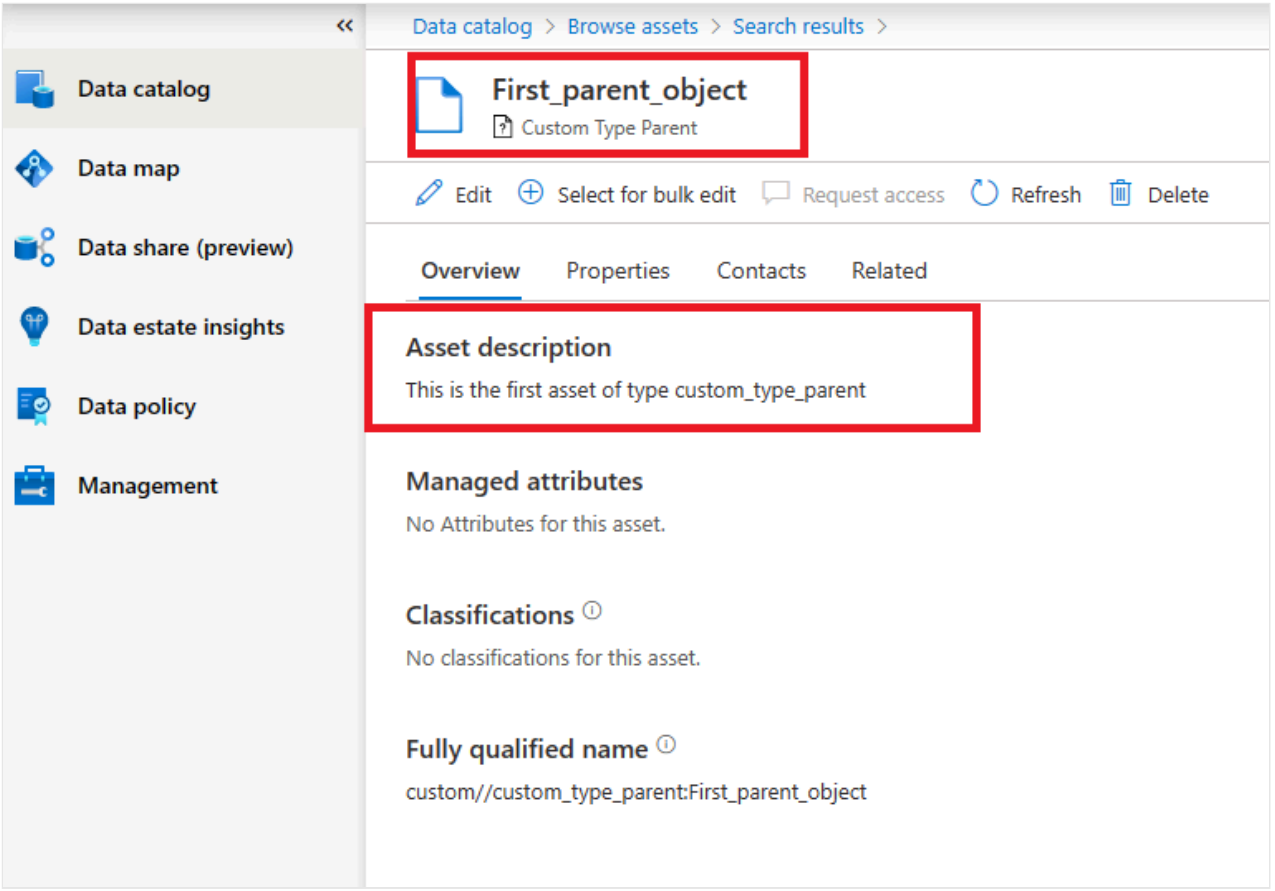
```
{
   "typeName": "custom_parent_child_relationship",
   "end1": {
         "guid": "...",
       "typeName": "custom_type_parent"
   },
   "end2": {
       "guid": "...",
       "typeName": "custom_type_child"
   }
}
```

View the assets in Microsoft Purview

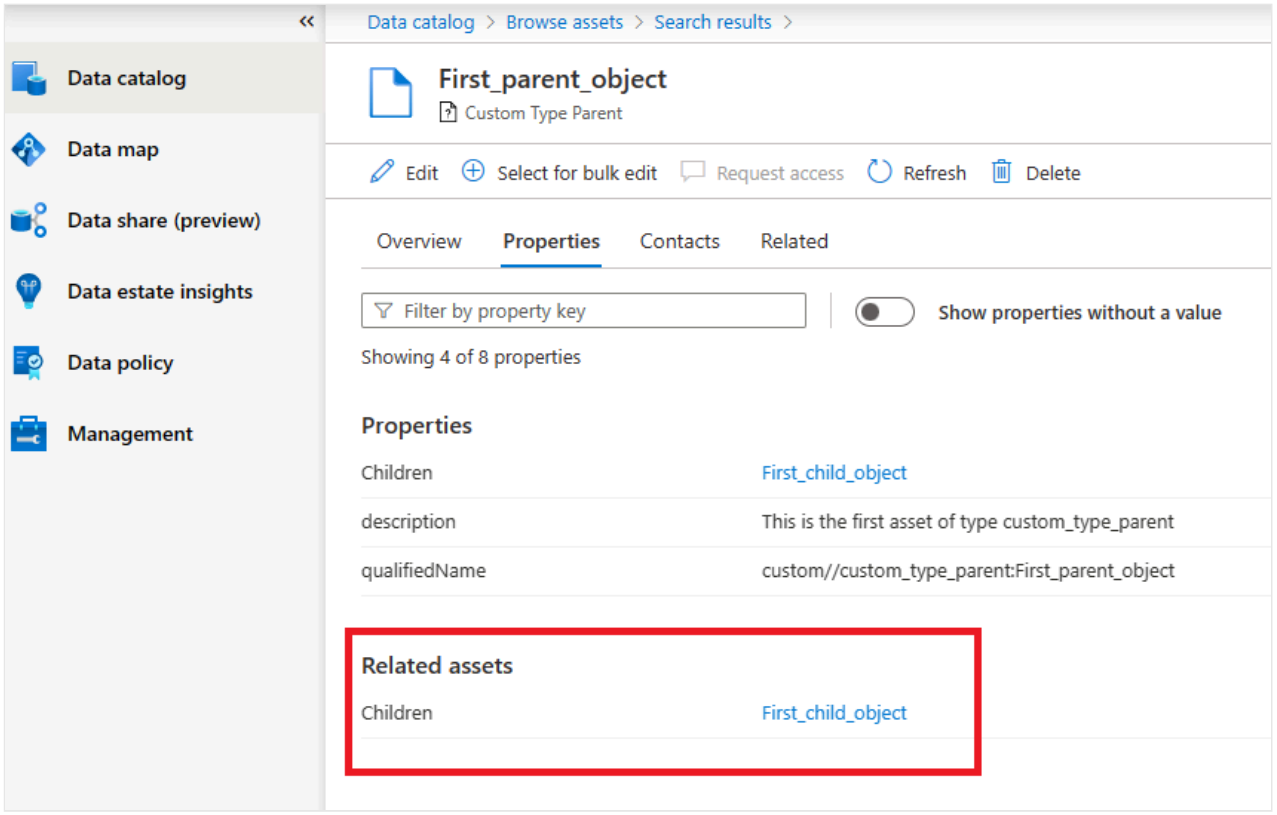
1. Go to *Unified Catalog* in Microsoft Purview.
2. Select *Browse*.
3. Select *By source type*.
4. Select *Sample-Custom-Types*.



5. Select the *First_parent_object*:

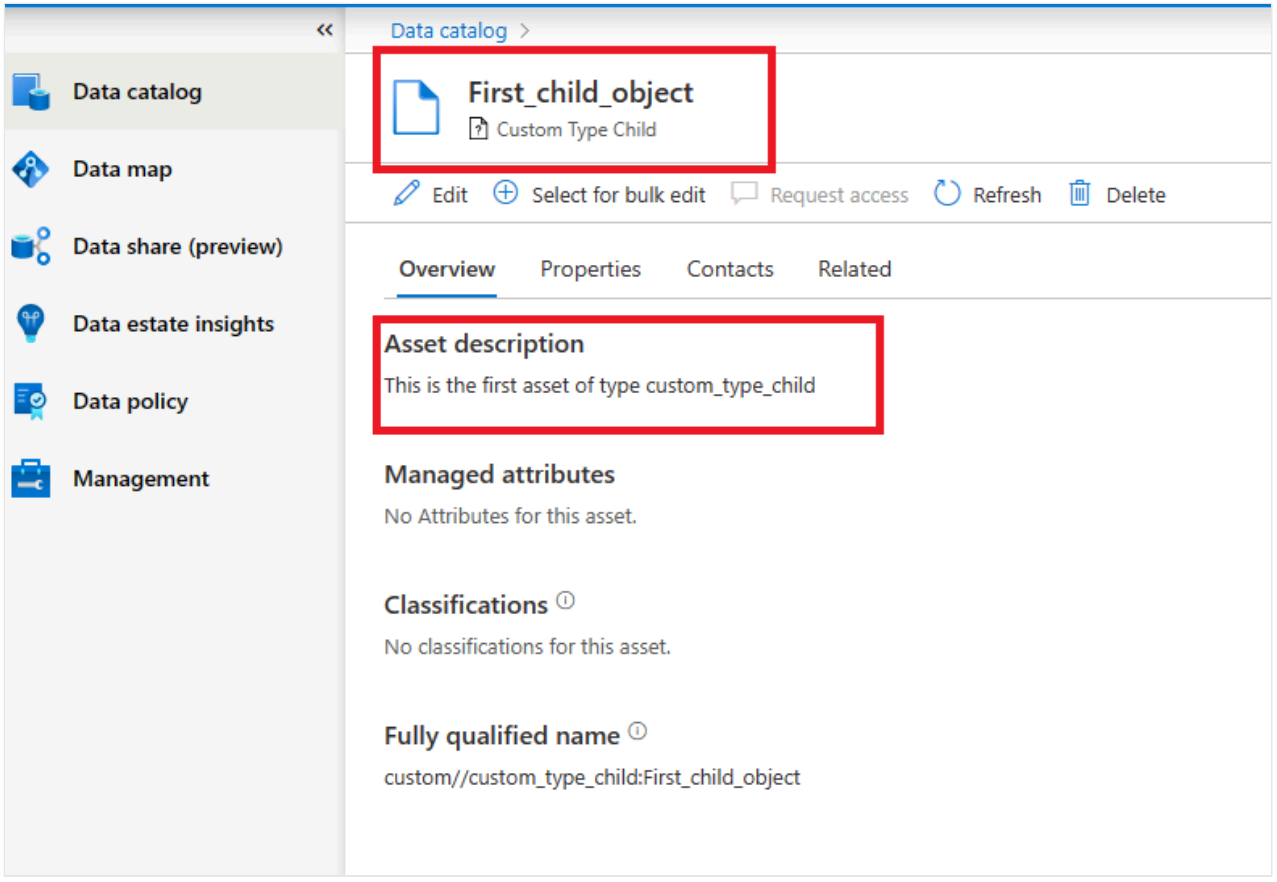


6. Select the *Properties* tab:

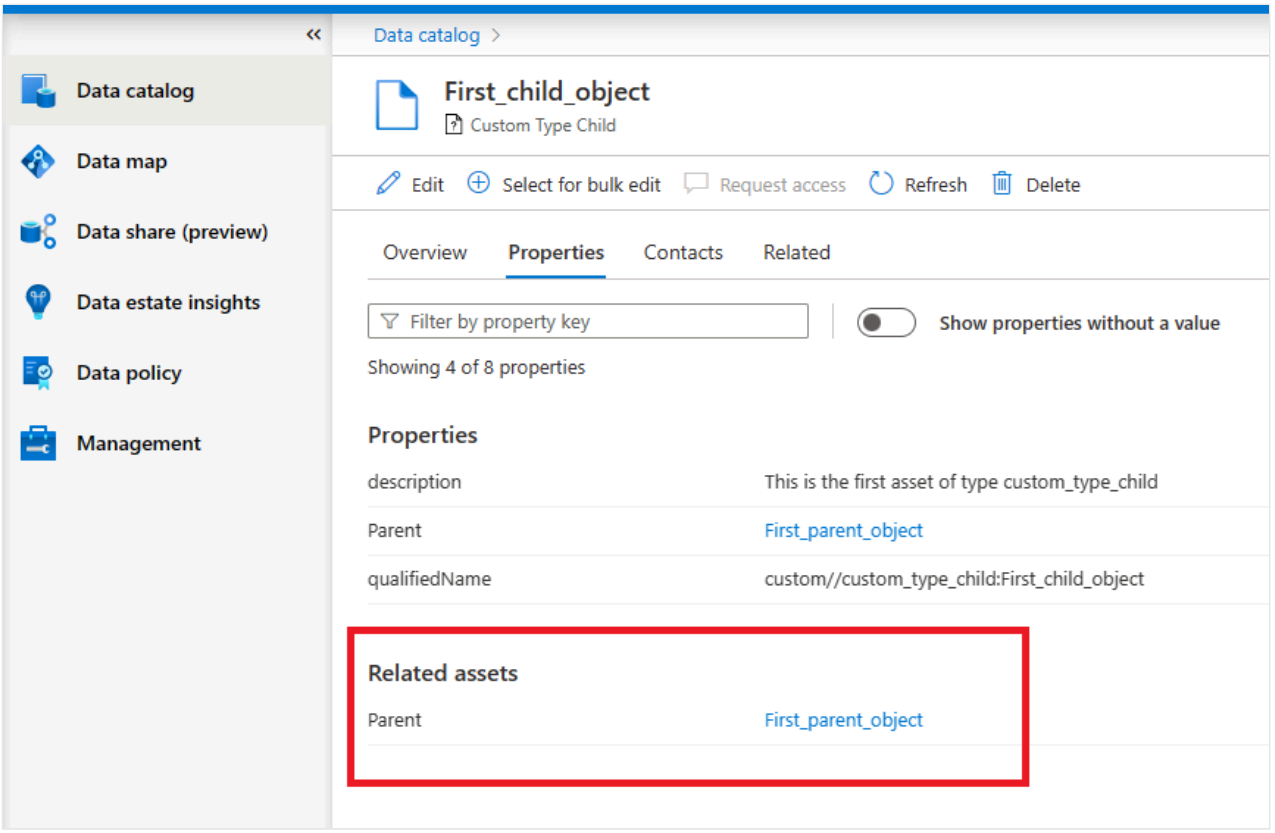


7. You can see the *First_child_object* being linked there.

8. Select the *First_child_object*:

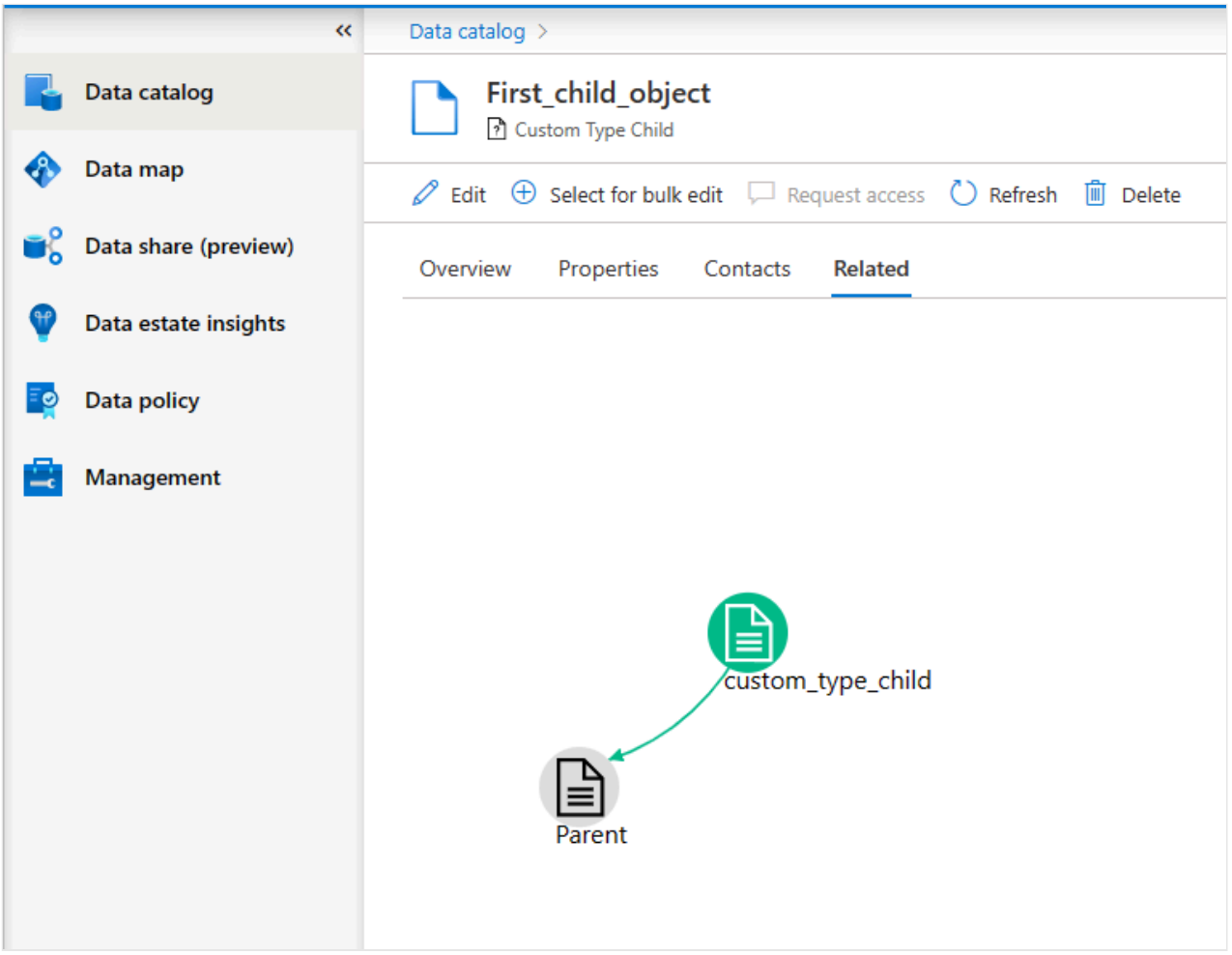


9. Select the *Properties* tab:



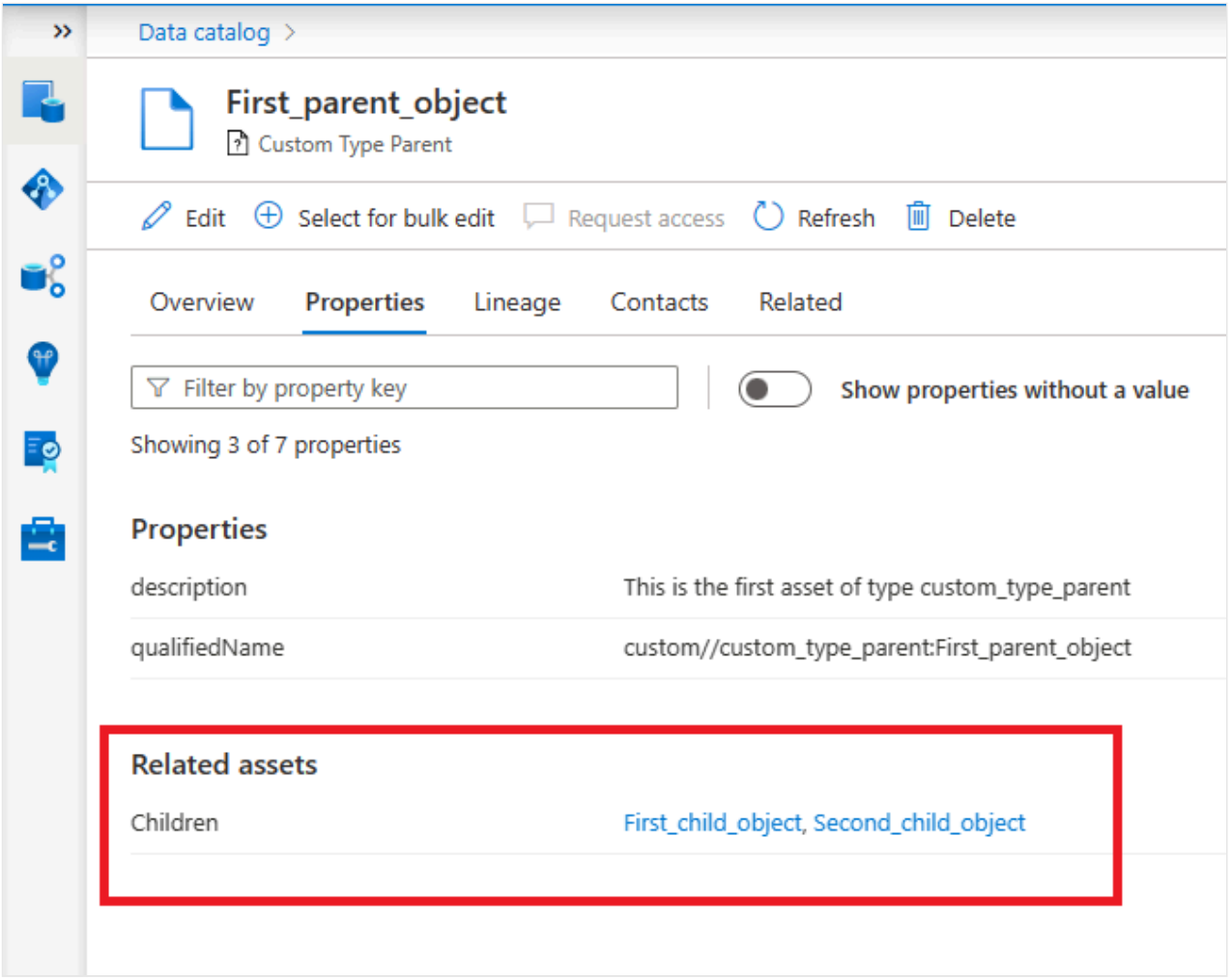
10. You can see the Parent object being linked there.

11. Similarly, you can select the *Related* tab and will see the relationship between the two objects:



12. You can create multiple children by initializing a new child asset and inititialzing a relationship

> **Note**
>
> The *qualifiedName* is unique per asset, therefore, the second child should be called differently, such as: *custom//custom_type_child:Second_child_object*



> **Tip**
>
> If you delete the *First_parent_object* you will notice that the children will also be removed, due to the *COMPOSITION* relationship that we chose in the definition.

Limitations

There are several known limitations when working with custom types that will be enhanced in future, such as:

- Relationship tab looks different compared to built-in types

- Custom types have no icons
- Hierarchy isn't supported

Next steps

Manage data sources | Microsoft Purview data plane REST APIs

Tutorial: Atlas 2.2 functionality

Article • 05/14/2025

In this tutorial, learn to programmatically interact with new Atlas 2.2 APIs with Microsoft Purview Data Map.

Prerequisites

- If you don't have an Azure subscription, create a free account before you begin.
- You must have an existing Microsoft Purview account. If you don't have a catalog, see the quickstart for creating a Microsoft Purview account.
- To establish a bearer token and to call any APIs, see the documentation about how to authenticate APIs for Microsoft Purview.

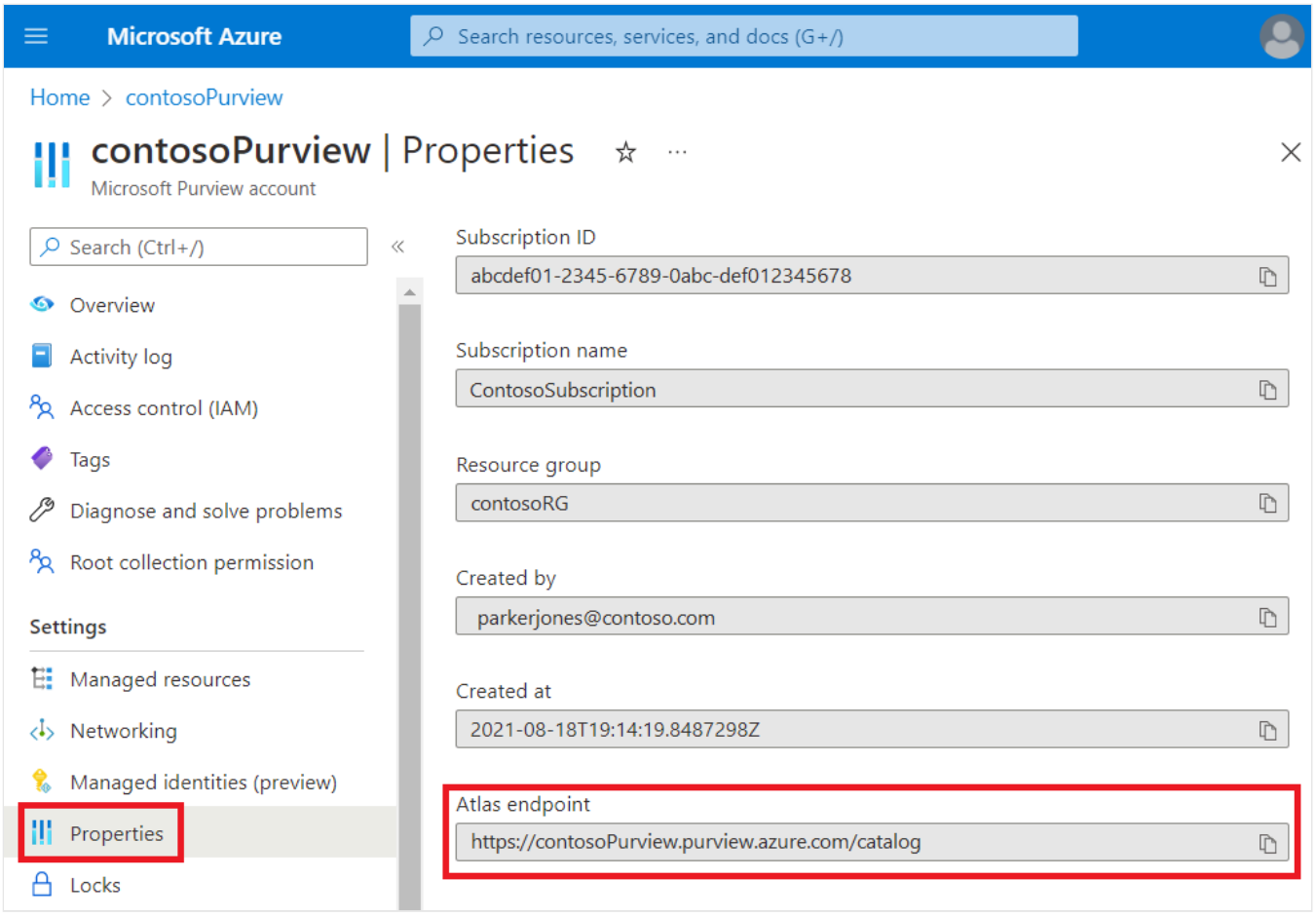
Business metadata APIs

Business metadata is a template that contains custom attributes (key values). You can create these attributes globally and then apply them across multiple typedefs.

Atlas endpoint

For all the requests, you'll need the Atlas endpoint for your Microsoft Purview account.

1. Find your Microsoft Purview account in the Azure portal
2. Select the **Properties** page on the left side menu
3. Copy the **Atlas endpoint** value



Create business metadata with attributes

You can send a `POST` request to the following endpoint:

```
POST {{endpoint}}/api/atlas/v2/types/typedefs
```

> **Tip**
>
> The **applicableEntityTypes** property tells which data types the metadata will be applied to.

Sample JSON:

JSON

```
 {
"businessMetadataDefs": [
  {
    "category": "BUSINESS_METADATA",
    "createdBy": "admin",
    "updatedBy": "admin",
    "version": 1,
```

```
      "typeVersion": "1.1",
      "name": "<Name of Business Metadata>",
      "description": "",
      "attributeDefs": [
        {
          "name": "<Attribute Name>",
          "typeName": "string",
          "isOptional": true,
          "cardinality": "SINGLE",
          "isUnique": false,
          "isIndexable": true,
          "options": {
            "maxStrLength": "50",
            "applicableEntityTypes": "[\"Referenceable\"]"
          }
        }
      ]
    }
  ]
}
```

Add or update an attribute to existing business metadata

You can send a `PUT` request to the following endpoint:

```
PUT {{endpoint}}/api/atlas/v2/types/typedefs
```

Sample JSON:

JSON

```
 {
  "businessMetadataDefs": [
   {
     "category": "BUSINESS_METADATA",
     "createdBy": "admin",
     "updatedBy": "admin",
     "version": 1,
     "typeVersion": "1.1",
     "name": "<Name of Business Metadata>",
     "description": "",
     "attributeDefs": [
       {
         "name": "<Attribute Name>",
         "typeName": "string",
         "isOptional": true,
         "cardinality": "SINGLE",
         "isUnique": false,
```

```
          "isIndexable": true,
          "options": {
            "maxStrLength": "500",
            "applicableEntityTypes": "[\"Referenceable\"]"
          }
        },
        {
          "name": "<Attribute Name 2>",
          "typeName": "int",
          "isOptional": true,
          "cardinality": "SINGLE",
          "isUnique": false,
          "isIndexable": true,
          "options": {
            "applicableEntityTypes": "[\"Referenceable\"]"
          }
        }
      ]
    }
  ]
}
```

Get a business metadata definition

You can send a `GET` request to the following endpoint:

```
GET {endpoint}}/api/atlas/v2/types/typedef/name/{{Business Metadata Name}}
```

Set a business metadata attribute to an entity

You can send a `POST` request to the following endpoint:

```
POST {{endpoint}}/api/atlas/v2/entity/guid/{{GUID}}/businessmetadata?
isOverwrite=true
```

Sample JSON:

JSON

```
{
  "myBizMetaData1": {
        "bizAttr1": "I am myBizMetaData1.bizAttr1",
        "bizAttr2": 123,
```

```
  }
 }
```

Delete a business metadata attribute from an entity

You can send a `DELETE` request to the following endpoint:

```
'DELETE' {{endpoint}}/api/atlas/v2/entity/guid/{{GUID}}/businessmetadata?
isOverwrite=true
```

Sample JSON:

JSON

```
{
  "myBizMetaData1": {
    "bizAttr1": ""
  }
}
```

Delete a business metadata type definition

> **Note**
>
> You can only delete business metadata type definition if it has no references, i.e., if it has not been assigned to any assets in the catalog.

You can send a `DELETE` request to the following endpoint:

```
DELETE {{endpoint}}/api/atlas/v2/types/typedef/name/{{Business Metadata Name}}
```

Custom attribute APIs

Custom attributes are key/value pairs that can be directly added to an Atlas entity.

Set a custom attribute to an entity

You can send a `POST` request to the following endpoint:

```
POST {{endpoint}}/api/atlas/v2/entity
```

Sample JSON:

JSON

```
{
    "entity": {
        "typeName": "azure_datalake_gen2_path",
        "attributes": {

            "qualifiedName": "<FQN of the asset>",
            "name": "data6.csv"
        },
        "guid": "3ffb28ff-138f-419e-84ba-348b0165e9e0",
        "customAttributes": {
            "custAttr1": "attr1",
            "custAttr2": "attr2"
        }
    }
}
```

Label APIs

Labels are free text tags that can be applied to any Atlas entity.

Set labels to an entity

You can send a `POST` request to the following endpoint:

```
POST {{endpoint}}/api/atlas/v2/entity/guid/{{GUID}}/labels
```

Sample JSON:

JSON

```
[
  "label1",
```

```
  "label2"
]
```

Delete labels to an entity

You can send a `DELETE` request to the following endpoint:

```
DELETE {{endpoint}}/api/atlas/v2/entity/guid/{{GUID}}/labels
```

Sample JSON:

JSON

```
[
  "label2"
]
```

Next steps

Manage data sources | Microsoft Purview data plane REST APIs

Tutorial: Authenticate for APIs

10/15/2025

In this tutorial, you learn how to authenticate for the Microsoft Purview data plane APIs. Anyone who wants to submit data to Microsoft Purview, include Microsoft Purview as part of an automated process, or build their own user experience on Microsoft Purview can use the APIs to do so.

Prerequisites

- To get started, you must have an existing Microsoft Purview account. If you don't have a catalog, see the quickstart for creating a Microsoft Purview account.

Create a service principal (application)

For an API client to access the Microsoft Purview data plane APIs, the client must have a service principal (application), and an identity that Microsoft Purview recognizes and is configured to trust. When you make API calls, that service principal's identity will be used for authorization.

Customers who have used existing service principals (application IDs) have had a high rate of failure. Therefore, we recommend creating a new service principal for calling APIs.

To create a new service principal:

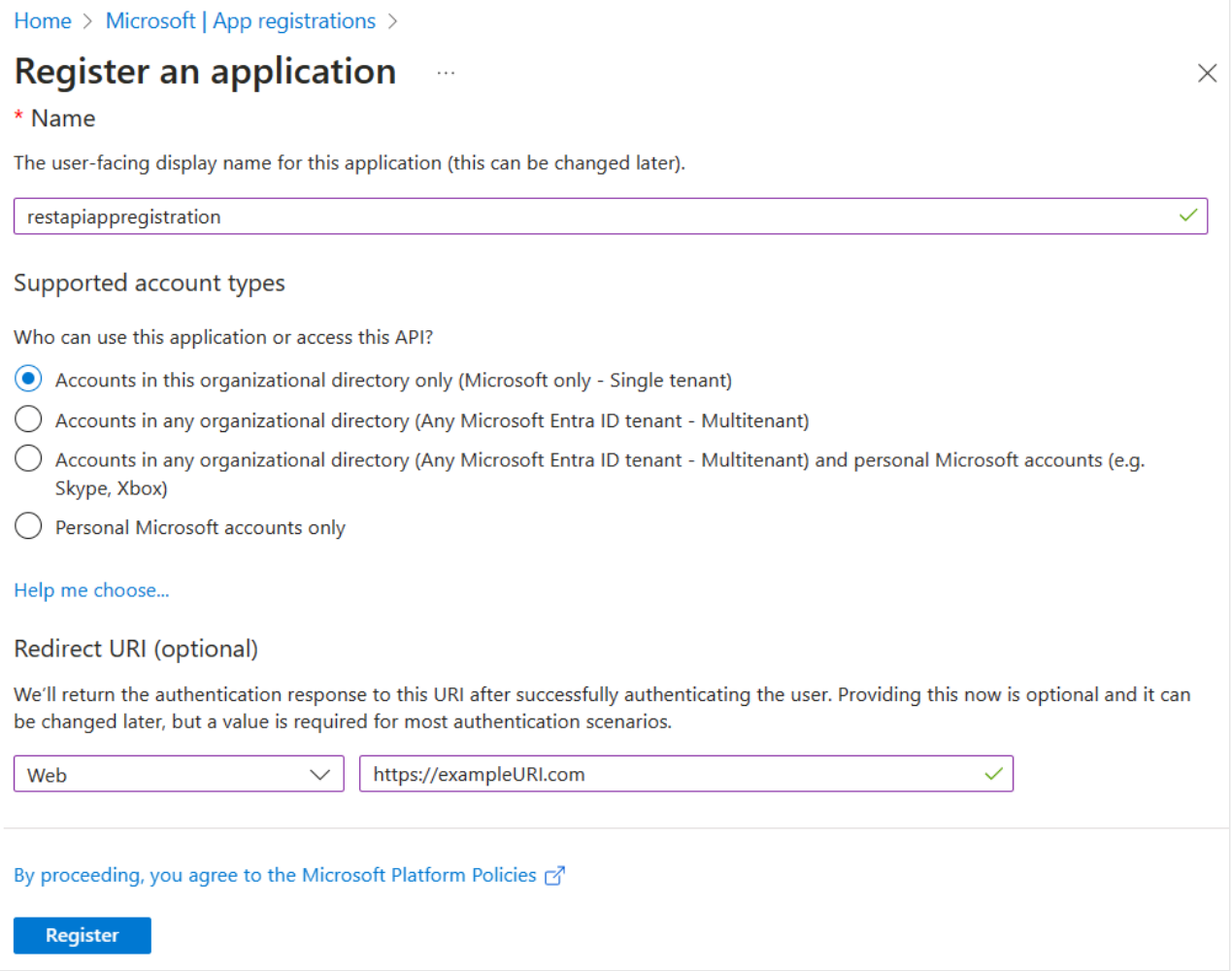
1. Sign in to the Azure portal.
2. From the portal, search for and select **Microsoft Entra ID**.
3. From the **Microsoft Entra ID** page, select **App registrations** from the left pane.
4. Select **New registration**.
5. On the **Register an application** page:
   a. Enter a **Name** for the application (the service principal name).
   b. For *Who can use this application or access this API?*, select the kinds of user accounts you expect to use this API.

> **Tip**

If you expect only users from your current Microsoft Entra ID tenant to use the REST API, select, **Accounts in this organizational directory only (*<your tenant's name>* only - Single tenant)**. Otherwise, consider the other options.

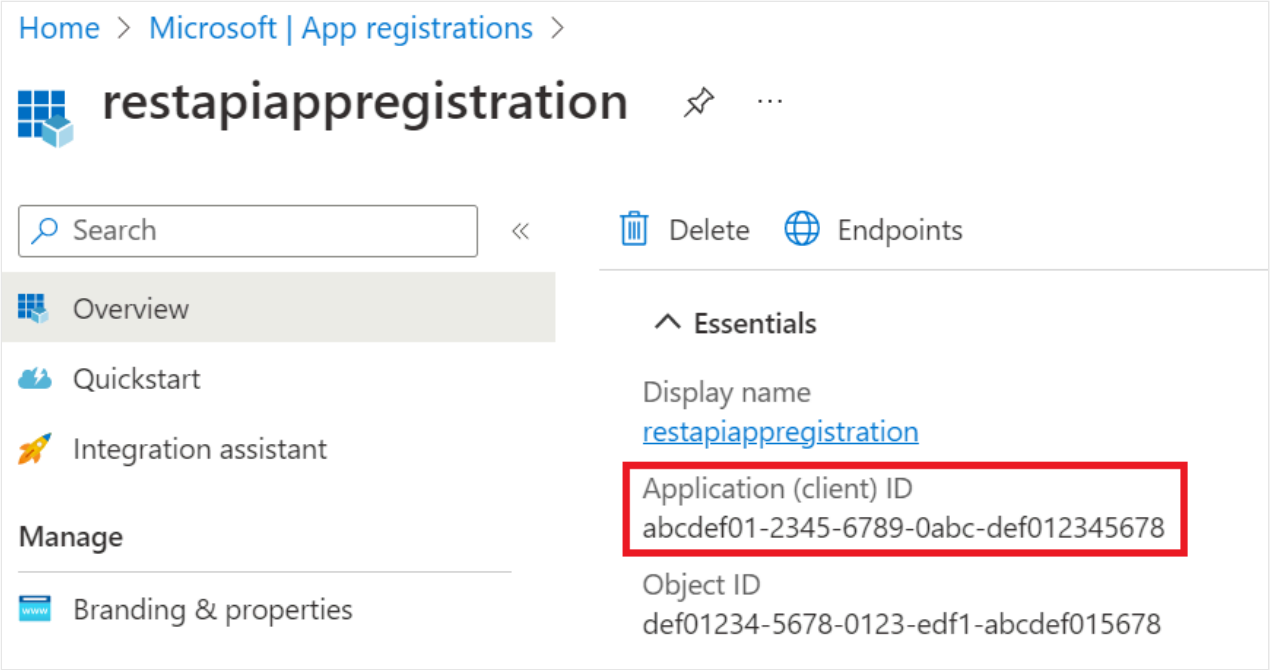
c. For **Redirect URI (optional)**, select **Web** and enter a value. This value doesn't need to be a valid endpoint. `https://exampleURI.com` will do.

d. Select **Register**.



6. On the new service principal page, copy the values of the **Display name** and the **Application (client) ID** to save for later.

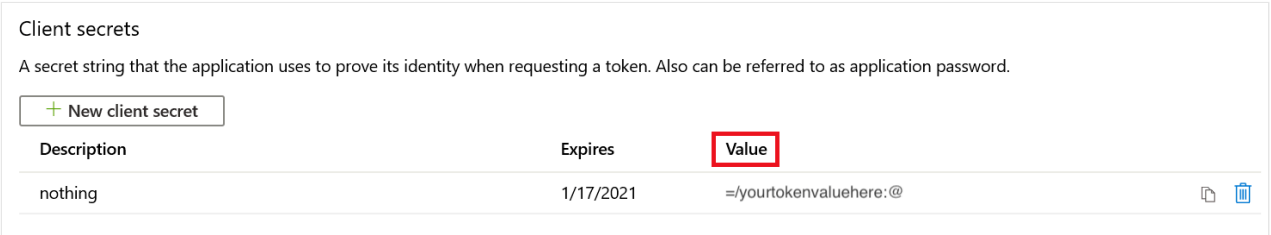
   The application ID is the `client_id` value in the sample code.



To use the service principal (application), you need to know the service principal's password that can be found by:

1. From the Azure portal, search for and select **Microsoft Entra ID**, and then select **App registrations** from the left pane.
2. Select your service principal (application) from the list.
3. Select **Certificates & secrets** from the left pane.
4. Select **New client secret**.
5. On the **Add a client secret** page, enter a **Description**, select an expiration time under **Expires**, and then select **Add**.

   On the **Client secrets** page, the string in the **Value** column of your new secret is your password. Save this value.



Set up authentication using service principal

Once the new service principal is created, you need to assign the data plane roles of your purview account to the service principal created above. Follow the steps below to assign the correct role to establish trust between the service principal and the Purview account:

Data Map

1. Navigate to your Microsoft Purview governance portal.
2. Select the Data Map in the left menu.
3. Select Collections.
4. Select the root collection in the collections menu. This will be the top collection in the list, and will have the same name as your Microsoft Purview account.

   > **Note**
   >
   > You can also assign your service principal permission to any sub-collections, instead of the root collection. However, all APIs will be scoped to that collection (and sub-collections that inherit permissions), and users trying to call the API for another collection will get errors.

5. Select the **Role assignments** tab.
6. Assign the following roles to the service principal created previously to access various data planes in Microsoft Purview. For detailed steps, see Assign Azure roles using the Microsoft Purview governance portal.

- Data Curator role to access Catalog Data plane.
- Data Source Administrator role to access Scanning Data plane.
- Collection Admin role to access Account Data Plane and Metadata policy Data Plane.
- Policy Author role to access the DevOps policies API

> **Note**
>
> Only members of the Collection Admin role can assign data plane roles in Microsoft Purview. For more information about Microsoft Purview roles, see Access Control in Microsoft Purview.

Get token

You can send a POST request to the following URL to get access token.

`https://login.microsoftonline.com/{your-tenant-id}/oauth2/token`

You can find your Tenant ID by searching for **Tenant Properties** in the Azure portal. The ID will be available on the tenant properties page.

The following parameters need to be passed to the above URL:

- **client_id**: client ID of the application registered in Microsoft Entra ID and is assigned to a data plane role for the Microsoft Purview account.
- **client_secret**: client secret created for the above application.
- **grant_type**: This should be 'client_credentials'.
- **resource**: 'https://purview.azure.net'

Here's a sample POST request in PowerShell:

Azure PowerShell

```
$tenantID = "12a345bc-67d1-ef89-abcd-efg12345abcde"

$url = "https://login.microsoftonline.com/$tenantID/oauth2/token"
$params = @{ client_id = "a1234bcd-5678-9012-abcd-abcd1234abcd"; client_secret =
"abcd~a1234bcd56789012abcdabcd1234abcd"; grant_type = "client_credentials";
resource = 'https://purview.azure.net' }

Invoke-WebRequest $url -Method Post -Body $params -UseBasicParsing | ConvertFrom-
Json
```

Sample response token:

JSON

```
{
    "token_type": "Bearer",
    "expires_in": "86399",
    "ext_expires_in": "86399",
    "expires_on": "1621038348",
    "not_before": "1620951648",
    "resource": "https://purview.azure.net",
    "access_token": "<<access token>>"
}
```

Tip

If you get an error message that reads: *Cross-origin token redemption is permitted only for the 'Single-Page Application' client-type.*

- Check your request headers and confirm that your request **doesn't** contain the 'origin' header.
- Confirm that your redirect URI is set to **web** in your service principal.
- Make sure your software is up to date for the application you're using to send your POST request.

Use the access token above to call the Data plane APIs.

Next steps

Manage data sources | Microsoft Purview Data Plane REST APIs

Tutorial: Use REST APIs to manage role-based access control on Microsoft Purview collections

Article • 02/24/2025

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

In August 2021, access control in Microsoft Purview moved from Azure Identity & Access Management (IAM) (control plane) to Microsoft Purview collections (data plane). This change gives enterprise data curators and administrators more precise, granular access control on their data sources scanned by Microsoft Purview. The change also enables organizations to audit right access and right use of their data.

This tutorial guides you through step-by-step usage of the Microsoft Purview Metadata Policy APIs to help you add users, groups, or service principals to a collection, and manage or remove their roles within that collection. REST APIs are an alternative method to using the Azure portal or Microsoft Purview governance portal to achieve the same granular role-based access control.

For more information about the built-in roles in Microsoft Purview, see the Microsoft Purview permissions guide. The guide maps the roles to the level of access permissions that are granted to users.

Metadata Policy API Reference summary

The following table gives an overview of the Microsoft Purview Metadata Policy API Reference.

> **Note**
>
> Replace {pv-acc-name} with the name of your Microsoft Purview account before running these APIs. For instance, if your Microsoft Purview account name is *FabrikamPurviewAccount*, your API endpoints will become *FabrikamPurviewAccount.purview.azure.com*. The "api-version" parameter is subject

to change. Please refer the Microsoft Purview Metadata policy REST API documentation for the latest "api-version" and the API signature.

Expand table

API function	REST method	API endpoint	Description
Read All Metadata Roles	GET	https://{pv-acc-name}.purview.azure.com /policystore/metadataroles?&api-version=2021-07-01	Reads all metadata roles from your Microsoft Purview account.
Read Metadata Policy By Collection Name	GET	https://{pv-acc-name}.purview.azure.com /policystore/collections/{collectionName}/metadataPolicy? &api-version=2021-07-01	Reads the metadata policy by using a specified collection name (the six character random name that's generated by Microsoft Purview when it creates the policy).
Read Metadata Policy By PolicyID	GET	https://{pv-acc-name}.purview.azure.com /policystore/metadataPolicies/{policyId}?&api-version=2021-07-01	Reads the metadata policy by using a specified policy ID. The policy ID is in GUID format.
Read All Metadata Policies	GET	https://{pv-acc-name}.purview.azure.com /policystore/metadataPolicies?&api-version=2021-07-01	Reads all metadata policies from your

API function	REST method	API endpoint	Description
			Microsoft Purview account. You can pick a certain policy to work with from the JSON output list that's generated by this API.
Update/PUT Metadata Policy	PUT	https://{pv-acc-name}.purview.azure.com /policystore/metadataPolicies/{policyId}?&api-version=2021-07-01	Updates the metadata policy by using a specified policy ID. The policy ID is in GUID format.

Microsoft Purview catalog collections API reference summary

The following table gives an overview of the Microsoft Purview collections APIs. For complete documentation about each API, select the API operation in the left column.

Expand table

Operation	Description
Create or update collection	Creates or updates a collection entity.
Delete collection	Deletes a collection entity.
Get collection	Gets a collection.
Get collection path	Gets the parent name and display name chains that represent the

Operation	Description
	collection path.
List child collection names	Lists the child collections names in the collection.
List collections	Lists the collections in the account.

- If you're using the API, the service principal, user, or group that executes the API should have a Collection Admin role assigned in Microsoft Purview to execute this API successfully.
- For all Microsoft Purview APIs that require {collectionName}, you'll need to use "*name*" (and not "*friendlyName*"). Replace {collectionName} with the actual six-character alphanumeric collection name string.

> **Note**
>
> This name is different from the friendly display name you supplied when you created the collection. If you don't have {collectionName} handy, use the List Collections API to select the six-character collection name from the JSON output.

Here's an example JSON file:

JSON

```
{
    "name": "74dhe7",
    "friendlyName": "Friendly Name",
    "parentCollection": {
        "type": "CollectionReference",
        "referenceName": "{your_purview_account_name}"
    },
    "systemData": {
        "createdBy": "{guid}",
        "createdByType": "Application",
        "createdAt": "2021-08-26T21:21:51.2646627Z",
        "lastModifiedBy": "7f8d47e2-330c-42f0-8744-fcfb1ecb3ea0",
        "lastModifiedByType": "Application",
        "lastModifiedAt": "2021-08-26T21:21:51.2646628Z"
    },
    "collectionProvisioningState": "Succeeded"
}
```

Policy JSON description

Here are some of the important identifiers in the JSON output that's received from the collection APIs:

Name: The name of the policy.

Id: The unique identifier for the policy.

Version: The latest version number of the policy.

> **ⓘ Important**
>
> The version number is incremented each time the Update-Metadata-Policy API is called. Be sure to fetch the latest copy of the policy by invoking the Get-Policy-by-Policy-ID API. Perform this refresh every time before you call the Update Policy (PUT) API, so that you always have the latest version of the JSON file.

DecisionRules: List the rules and effect of this policy. For metadata policies, the effect is always *"Permit"*.

Add or remove users from a collection or role

Use Microsoft Purview REST APIs to add or remove a user, group, or service principal from a collection or role. Detailed API usage is provided along with sample JSON outputs. We highly recommend that you follow the instructions in the next sections sequentially for best understanding of the Microsoft Purview metadata policy APIs.

Get all metadata roles

To list all the available metadata access permission roles, run one of the following commands, depending on the portal you are using:

Classic Microsoft Purview governance portal:

```ruby
GET
https://{your_purview_account_name}.purview.azure.com/policystore/metadataroles?api-version=2021-07-01
```

New Microsoft Purview portal:

Ruby

```
GET https://api.purview-service.microsoft.com/policystore/metadataroles?api-version=2021-07-01
```

The output JSON will describe the roles and their associated permissions in this format.

The default metadata roles are listed in the following table:

Expand table

Role ID	Permissions	Role description
purviewmetadatarole_builtin_data-source-administrator	Microsoft.Purview/accounts/scan/read Microsoft.Purview/accounts/scan/write Microsoft.Purview/accounts/collection/read	Grants access to others to read, write collection, register data sources, and trigger scans.
purviewmetadatarole_builtin_collection-administrator	Microsoft.Purview/accounts/collection/read Microsoft.Purview/accounts/collection/write	Administrator-level full access to the entire collection, including add or remove users and service principal names (SPNs) from the collection, management rights, and grant or revoke access. In some cases, the Collection Administrator might be different from the creator of the collection.

Role ID	Permissions	Role description
purviewmetadatarole_builtin_purview-reader	Microsoft.Purview/accounts/data/read Microsoft.Purview/accounts/collection/read	Grants only read access to data handling and all metadata, including classifications, sensitivity labels, insights, and read assets in a collection, except scan bindings.
purviewmetadatarole_builtin_data-curator	Microsoft.Purview/accounts/data/read Microsoft.Purview/accounts/data/write Microsoft.Purview/accounts/collection/read	Grants full access to data handling and all metadata, including classifications, sensitivity labels, insights, and read assets in a collection, except scan bindings.
purviewmetadatarole_builtin_data-share-contributor	Microsoft.Purview/accounts/share/read Microsoft.Purview/accounts/share/write	Grants access to data shares as a contributor.

JSON


```
{
  "values": [
    {
      "id": "purviewmetadatarole_builtin_data-curator",
      "name": "data-curator",
      "type": "Microsoft.Purview/role",
      "properties": {
        "provisioningState": "Provisioned",
        "roleType": "BuiltIn",
        "friendlyName": "Data Curator",
        "cnfCondition": [
```

```
          [
            {
              "attributeName": "request.azure.dataAction",
              "attributeValueIncludedIn": [
                "Microsoft.Purview/accounts/data/read",
                "Microsoft.Purview/accounts/data/write",
                "Microsoft.Purview/accounts/collection/read"
              ]
            }
          ]
        ],
        "version": 1
      }
    },
    {
      "id": "purviewmetadatarole_builtin_data-source-administrator",
      "name": "data-source-administrator",
      "type": "Microsoft.Purview/role",
      "properties": {
        "provisioningState": "Provisioned",
        "roleType": "BuiltIn",
        "friendlyName": "Data Source Administrator",
        "cnfCondition": [
          [
            {
              "attributeName": "request.azure.dataAction",
              "attributeValueIncludedIn": [
                "Microsoft.Purview/accounts/scan/read",
                "Microsoft.Purview/accounts/scan/write",
                "Microsoft.Purview/accounts/collection/read"
              ]
            }
          ]
        ],
        "version": 1
      }
    },
    {
      "id": "purviewmetadatarole_builtin_collection-administrator",
      "name": "collection-administrator",
      "type": "Microsoft.Purview/role",
      "properties": {
        "provisioningState": "Provisioned",
        "roleType": "BuiltIn",
        "friendlyName": "Collection Administrator",
        "cnfCondition": [
          [
            {
              "attributeName": "request.azure.dataAction",
              "attributeValueIncludedIn": [
                "Microsoft.Purview/accounts/collection/read",
                "Microsoft.Purview/accounts/collection/write"
              ]
            }
          ]
```

```
        ],
        "version": 1
      }
    },
    {
      "id": "purviewmetadatarole_builtin_purview-reader",
      "name": "purview-reader",
      "type": "Microsoft.Purview/role",
      "properties": {
        "provisioningState": "Provisioned",
        "roleType": "BuiltIn",
        "friendlyName": "Microsoft Purview Reader",
        "cnfCondition": [
          [
            {
              "attributeName": "request.azure.dataAction",
              "attributeValueIncludedIn": [
                "Microsoft.Purview/accounts/data/read",
                "Microsoft.Purview/accounts/collection/read"
              ]
            }
          ]
        ],
        "version": 1
      }
    },
    {
      "id": "purviewmetadatarole_builtin_data-share-contributor",
      "name": "data-share-contributor",
      "type": "Microsoft.Purview/role",
      "properties": {
        "provisioningState": "Provisioned",
        "roleType": "BuiltIn",
        "friendlyName": "Data share contributor",
        "cnfCondition": [
          [
            {
              "attributeName": "request.azure.dataAction",
              "attributeValueIncludedIn": [
                "Microsoft.Purview/accounts/share/read",
                "Microsoft.Purview/accounts/share/write"
              ]
            }
          ]
        ],
        "version": 1
      }
    }
  ]
}
```


Get all metadata policies

Classic Microsoft Purview governance portal:

```
Ruby

GET
https://{your_purview_account_name}.purview.azure.com/policystore/metadataPolicies?api-version=2021-07-01
```

New Microsoft Purview portal:

```
Ruby

GET https://api.purview-service.microsoft.com/policystore/metadataPolicies?api-version=2021-07-01
```

The preceding command lists all available metadata policies across the entire collections hierarchy in tree format, from the root collection at the top to all its child policies. Each child collection contains each of its next level children.

Example:

```
JSON

{
  "values": [
    {
      "name": "policy_FabrikamPurview",
      "id": "9b2f1cb9-584c-4a16-811e-9232884b5cac",
      "version": 30,
      "properties": {
        "description": "",
        "decisionRules": [
          {
            "kind": "decisionrule",
            "effect": "Permit",
            "dnfCondition": [
              [
                {
                  "attributeName": "resource.purview.collection",
                  "attributeValueIncludes": "fabrikampurview"
                },
                {
                  "fromRule": "permission:fabrikampurview",
                  "attributeName": "derived.purview.permission",
                  "attributeValueIncludes": "permission:fabrikampurview"
                }
              ]
            ]
          }
        ],
```

```
        "attributeRules": [
          {
            "kind": "attributerule",
            "id": "purviewmetadatarole_builtin_collection-
administrator:fabrikampurview",
            "name": "purviewmetadatarole_builtin_collection-
administrator:fabrikampurview",
            "dnfCondition": [
              [
                {
                  "attributeName": "principal.microsoft.id",
                  "attributeValueIncludedIn": [
                    "2f656762-e440-4b62-9eb6-a991d17d64b0",
                    "04314867-60a4-4e5a-ae16-8e5856f415d9",
                    "8988fe5c-5736-4179-9435-0a64c273b90b",
                    "6d563253-1d5b-48f2-baaa-5489f22ddce9",
                    "26f98046-5b02-4fa9-b709-e0519c658891",
                    "73fc02dc-becd-468b-a2a3-82238e722dae"
                  ]
                },
                {
                  "fromRule": "purviewmetadatarole_builtin_collection-
administrator",
                  "attributeName": "derived.purview.role",
                  "attributeValueIncludes":
"purviewmetadatarole_builtin_collection-administrator"
                }
              ],
              [
                {
                  "fromRule": "purviewmetadatarole_builtin_collection-
administrator",
                  "attributeName": "derived.purview.role",
                  "attributeValueIncludes":
"purviewmetadatarole_builtin_collection-administrator"
                },
                {
                  "attributeName": "principal.microsoft.groups",
                  "attributeValueIncludedIn": [
                    "ffd851fa-86ec-431b-95ea-8b84d5012383",
                    "cf84b126-4384-4952-91f1-7f705b25e569",
                    "5046aba1-5b81-411c-8fec-b84600f3f08b",
                    "b055a5c6-a04e-4d1a-8524-001ad81bfb28",
                    "cc194892-92fa-4ce3-96ae-1f98bef8211c"
                  ]
                }
              ]
            ]
          },
          {
            "kind": "attributerule",
            "id": "purviewmetadatarole_builtin_data-
curator:fabrikampurview",
            "name": "purviewmetadatarole_builtin_data-
curator:fabrikampurview",
```

```
            "dnfCondition": [
              [
                {
                  "attributeName": "principal.microsoft.id",
                  "attributeValueIncludedIn": [
                    "2f656762-e440-4b62-9eb6-a991d17d64b0",
                    "649f56ab-2dd2-40de-a731-3d3f28e7af92",
                    "c29a5809-f9ec-49fd-b762-2d4d64abb93e",
                    "04314867-60a4-4e5a-ae16-8e5856f415d9",
                    "73fc02dc-becd-468b-a2a3-82238e722dae",
                    "517a27d2-39ba-4c91-a032-dd9ecf8ad6f1",
                    "6d563253-1d5b-48f2-baaa-5489f22ddce9"
                  ]
                },
                {
                  "fromRule": "purviewmetadatarole_builtin_data-curator",
                  "attributeName": "derived.purview.role",
                  "attributeValueIncludes":
"purviewmetadatarole_builtin_data-curator"
                }
              ],
              [
                {
                  "fromRule": "purviewmetadatarole_builtin_data-curator",
                  "attributeName": "derived.purview.role",
                  "attributeValueIncludes":
"purviewmetadatarole_builtin_data-curator"
                },
                {
                  "attributeName": "principal.microsoft.groups",
                  "attributeValueIncludedIn": [
                    "b055a5c6-a04e-4d1a-8524-001ad81bfb28",
                    "cc194892-92fa-4ce3-96ae-1f98bef8211c",
                    "5046aba1-5b81-411c-8fec-b84600f3f08b"
                  ]
                }
              ]
            ]
          },
          {
            "kind": "attributerule",
            "id": "purviewmetadatarole_builtin_data-source-
administrator:fabrikampurview",
            "name": "purviewmetadatarole_builtin_data-source-
administrator:fabrikampurview",
            "dnfCondition": [
              [
                {
                  "attributeName": "principal.microsoft.id",
                  "attributeValueIncludedIn": [
                    "2f656762-e440-4b62-9eb6-a991d17d64b0",
                    "04314867-60a4-4e5a-ae16-8e5856f415d9",
                    "517a27d2-39ba-4c91-a032-dd9ecf8ad6f1",
                    "6d563253-1d5b-48f2-baaa-5489f22ddce9"
                  ]
```

```
                },
                {
                  "fromRule": "purviewmetadatarole_builtin_data-source-
administrator",
                  "attributeName": "derived.purview.role",
                  "attributeValueIncludes":
"purviewmetadatarole_builtin_data-source-administrator"
                }
              ],
              [
                {
                  "fromRule": "purviewmetadatarole_builtin_data-source-
administrator",
                  "attributeName": "derived.purview.role",
                  "attributeValueIncludes":
"purviewmetadatarole_builtin_data-source-administrator"
                },
                {
                  "attributeName": "principal.microsoft.groups",
                  "attributeValueIncludedIn": [
                    "b055a5c6-a04e-4d1a-8524-001ad81bfb28",
                    "cc194892-92fa-4ce3-96ae-1f98bef8211c",
                    "d34eb741-be5e-4098-90d7-eca8d4a5153f",
                    "664ec992-9af0-4773-88f2-dc39edc46f6f",
                    "5046aba1-5b81-411c-8fec-b84600f3f08b"
                  ]
                }
              ]
            ]
          },
          {
            "kind": "attributerule",
            "id": "permission:fabrikampurview",
            "name": "permission:fabrikampurview",
            "dnfCondition": [
              [
                {
                  "fromRule": "purviewmetadatarole_builtin_collection-
administrator:fabrikampurview",
                  "attributeName": "derived.purview.permission",
                  "attributeValueIncludes":
"purviewmetadatarole_builtin_collection-administrator:fabrikampurview"
                }
              ],
              [
                {
                  "fromRule": "purviewmetadatarole_builtin_purview-
reader:fabrikampurview",
                  "attributeName": "derived.purview.permission",
                  "attributeValueIncludes":
"purviewmetadatarole_builtin_purview-reader:fabrikampurview"
                }
              ],
              [
                {
```

```
                    "fromRule": "purviewmetadatarole_builtin_data-
curator:fabrikampurview",
                    "attributeName": "derived.purview.permission",
                    "attributeValueIncludes":
"purviewmetadatarole_builtin_data-curator:fabrikampurview"
                  }
                ],
                [
                  {
                    "fromRule": "purviewmetadatarole_builtin_data-source-
administrator:fabrikampurview",
                    "attributeName": "derived.purview.permission",
                    "attributeValueIncludes":
"purviewmetadatarole_builtin_data-source-administrator:fabrikampurview"
                  }
                ]
              ]
            }
          ],
          "collection": {
            "type": "CollectionReference",
            "referenceName": "fabrikampurview"
          }
        }
      },
      {
        "name": "policy_b2zpf1",
        "id": "12b0bb28-2acc-413e-8fe1-179ff9cc54c3",
        "version": 0,
        "properties": {
          "description": "",
          "decisionRules": [
            {
              "kind": "decisionrule",
              "effect": "Permit",
              "dnfCondition": [
                [
                  {
                    "attributeName": "resource.purview.collection",
                    "attributeValueIncludes": "b2zpf1"
                  },
                  {
                    "fromRule": "permission:b2zpf1",
                    "attributeName": "derived.purview.permission",
                    "attributeValueIncludes": "permission:b2zpf1"
                  }
                ]
              ]
            }
          ],
          "attributeRules": [
            {
              "kind": "attributerule",
              "id": "purviewmetadatarole_builtin_collection-
administrator:b2zpf1",
```

```
            "name": "purviewmetadatarole_builtin_collection-
administrator:b2zpf1",
            "dnfCondition": [
              [
                {
                  "attributeName": "principal.microsoft.id",
                  "attributeValueIncludedIn": [
                    "2f656762-e440-4b62-9eb6-a991d17d64b0"
                  ]
                },
                {
                  "fromRule": "purviewmetadatarole_builtin_collection-
administrator",
                  "attributeName": "derived.purview.role",
                  "attributeValueIncludes":
"purviewmetadatarole_builtin_collection-administrator"
                }
              ],
              [
                {
                  "fromRule": "purviewmetadatarole_builtin_collection-
administrator:ukx7pq",
                  "attributeName": "derived.purview.permission",
                  "attributeValueIncludes":
"purviewmetadatarole_builtin_collection-administrator:ukx7pq"
                }
              ]
            ]
          },
          {
            "kind": "attributerule",
            "id": "permission:b2zpf1",
            "name": "permission:b2zpf1",
            "dnfCondition": [
              [
                {
                  "fromRule": "purviewmetadatarole_builtin_collection-
administrator:b2zpf1",
                  "attributeName": "derived.purview.permission",
                  "attributeValueIncludes":
"purviewmetadatarole_builtin_collection-administrator:b2zpf1"
                }
              ],
              [
                {
                  "fromRule": "permission:ukx7pq",
                  "attributeName": "derived.purview.permission",
                  "attributeValueIncludes": "permission:ukx7pq"
                }
              ]
            ]
          }
        ],
        "collection": {
          "type": "CollectionReference",
```

```
          "referenceName": "b2zpf1"
        },
        "parentCollectionName": "ukx7pq"
      }
    },
    {
      "name": "policy_7wte2n",
      "id": "a72084e4-ccab-4aec-a364-08ab001e4999",
      "version": 0,
      "properties": {
        "description": "",
        "decisionRules": [
          {
            "kind": "decisionrule",
            "effect": "Permit",
            "dnfCondition": [
              [
                {
                  "attributeName": "resource.purview.collection",
                  "attributeValueIncludes": "7wte2n"
                },
                {
                  "fromRule": "permission:7wte2n",
                  "attributeName": "derived.purview.permission",
                  "attributeValueIncludes": "permission:7wte2n"
                }
              ]
            ]
          }
        ],
        "attributeRules": [
          {
            "kind": "attributerule",
            "id": "purviewmetadatarole_builtin_collection-
administrator:7wte2n",
            "name": "purviewmetadatarole_builtin_collection-
administrator:7wte2n",
            "dnfCondition": [
              [
                {
                  "attributeName": "principal.microsoft.id",
                  "attributeValueIncludedIn": [
                    "2f656762-e440-4b62-9eb6-a991d17d64b0"
                  ]
                },
                {
                  "fromRule": "purviewmetadatarole_builtin_collection-
administrator",
                  "attributeName": "derived.purview.role",
                  "attributeValueIncludes":
"purviewmetadatarole_builtin_collection-administrator"
                }
              ],
              [
                {
```

```
                    "fromRule": "purviewmetadatarole_builtin_collection-
administrator:ukx7pq",
                    "attributeName": "derived.purview.permission",
                    "attributeValueIncludes":
"purviewmetadatarole_builtin_collection-administrator:ukx7pq"
                  }
                ]
              ]
            },
            {
              "kind": "attributerule",
              "id": "permission:7wte2n",
              "name": "permission:7wte2n",
              "dnfCondition": [
                [
                  {
                    "fromRule": "purviewmetadatarole_builtin_collection-
administrator:7wte2n",
                    "attributeName": "derived.purview.permission",
                    "attributeValueIncludes":
"purviewmetadatarole_builtin_collection-administrator:7wte2n"
                  }
                ],
                [
                  {
                    "fromRule": "permission:ukx7pq",
                    "attributeName": "derived.purview.permission",
                    "attributeValueIncludes": "permission:ukx7pq"
                  }
                ]
              ]
            }
          ],
          "collection": {
            "type": "CollectionReference",
            "referenceName": "7wte2n"
          },
          "parentCollectionName": "ukx7pq"
        }
      }
    ]
}
```


Get the selected metadata policy

You can use either of two APIs to fetch a particular collection's metadata policy JSON structure by supplying either {collectionName} or {PolicyID}.

As described in the following two sections, both APIs serve the same purpose, and the JSON outputs of both are exactly the same.

Get the metadata policy of the collection by using the collection name

Classic Microsoft Purview governance portal:

```
GET
https://{your_purview_account_name}.purview.azure.com/policystore/collections/{collectionName}/metadataPolicy?api-version=2021-07-01
```

New Microsoft Purview portal:

```
GET https://api.purview-service.microsoft.com/policystore/collections/{collectionName}/metadataPolicy?api-version=2021-07-01
```

1. The Microsoft Purview account name is {your_purview_account_name}. Replace it with your Microsoft Purview account name.
2. In the JSON output of the previous API, "Get All Metadata Policies," look for the following section:

   { "type": "CollectionReference", "referenceName": "7xkdg2"}
3. Replace "{collectionName}" in the API URL with the value of "referenceName": "{6-char-collection-name}". For example, if your six-character collection name is "7xkdg2," the API URL will be formatted as:

   `https://{your_purview_account_name}.purview.azure.com/policystore/collections/7xkdg2/metadataPolicy?api-version=2021-07-01`
4. Execute the following API:

   ```
   {
     "name": "policy_qu45fs",
     "id": "c6639bb2-9c41-4be0-912b-775750e725de",
     "version": 0,
     "properties": {
       "description": "",
       "decisionRules": [
         {
           "kind": "decisionrule",
   ```

```
        "effect": "Permit",
        "dnfCondition": [
          [
            {
              "attributeName": "resource.purview.collection",
              "attributeValueIncludes": "qu45fs"
            },
            {
              "fromRule": "permission:qu45fs",
              "attributeName": "derived.purview.permission",
              "attributeValueIncludes": "permission:qu45fs"
            }
          ]
        ]
      }
    ],
    "attributeRules": [
      {
        "kind": "attributerule",
        "id": "purviewmetadatarole_builtin_collection-
administrator:qu45fs",
        "name": "purviewmetadatarole_builtin_collection-
administrator:qu45fs",
        "dnfCondition": [
          [
            {
              "attributeName": "principal.microsoft.id",
              "attributeValueIncludedIn": [
                "2f656762-e440-4b62-9eb6-a991d17d64b0"
              ]
            },
            {
              "fromRule": "purviewmetadatarole_builtin_collection-
administrator",
              "attributeName": "derived.purview.role",
              "attributeValueIncludes":
"purviewmetadatarole_builtin_collection-administrator"
            }
          ],
          [
            {
              "fromRule": "purviewmetadatarole_builtin_collection-
administrator:fabrikampurview",
              "attributeName": "derived.purview.permission",
              "attributeValueIncludes":
"purviewmetadatarole_builtin_collection-administrator:fabrikampurview"
            }
          ]
        ]
      },
      {
        "kind": "attributerule",
        "id": "permission:qu45fs",
        "name": "permission:qu45fs",
        "dnfCondition": [
```

```
          [
            {
              "fromRule": "purviewmetadatarole_builtin_collection-
administrator:qu45fs",
              "attributeName": "derived.purview.permission",
              "attributeValueIncludes":
"purviewmetadatarole_builtin_collection-administrator:qu45fs"
            }
          ],
          [
            {
              "fromRule": "permission:fabrikampurview",
              "attributeName": "derived.purview.permission",
              "attributeValueIncludes": "permission:fabrikampurview"
            }
          ]
        ]
      }
    ],
    "collection": {
      "type": "CollectionReference",
      "referenceName": "qu45fs"
    },
    "parentCollectionName": "fabrikampurview"
  }
}
```

Get the metadata policy of the collection by using the policy ID

Classic Microsoft Purview governance portal:

Ruby

```
GET
https://{your_purview_account_name}.purview.azure.com/policystore/metadataPo
licies/{policyId}?api-version=2021-07-01
```

New Microsoft Purview portal:

Ruby

```
GET https://api.purview-
service.microsoft.com/policystore/metadataPolicies/{policyId}?api-
version=2021-07-01
```

1. The Microsoft Purview account name is {your_purview_account_name}. Replace it with your Microsoft Purview account name.

2. In the JSON output of the previous API, "Get All Metadata Policies," look for the following section:

   {.... "name": "policy_qu45fs", "id": "{policy-guid}", "version": N ....}

3. Replace "{policyId}" in the API URL with the value of "id". For example, if your "{policy-guid}" is "c6639bb2-9c41-4be0-912b-775750e725de," the API URL will be formatted as:

   https://{your_purview_account_name}.purview.azure.com/policystore/metadataPolicies/c6639bb2-9c41-4be0-912b-775750e725de?api-version=2021-07-01

4. Now execute the following API:

> **Note**
>
> The output of this API call and the previous API call is the same. As mentioned previously, you can choose either one.

JSON


```
{
  "name": "policy_qu45fs",
  "id": "c6639bb2-9c41-4be0-912b-775750e725de",
  "version": 0,
  "properties": {
    "description": "",
    "decisionRules": [
      {
        "kind": "decisionrule",
        "effect": "Permit",
        "dnfCondition": [
          [
            {
              "attributeName": "resource.purview.collection",
              "attributeValueIncludes": "qu45fs"
            },
            {
              "fromRule": "permission:qu45fs",
              "attributeName": "derived.purview.permission",
              "attributeValueIncludes": "permission:qu45fs"
            }
          ]
        ]
      }
    ],
    "attributeRules": [
      {
        "kind": "attributerule",
```

```
        "id": "purviewmetadatarole_builtin_collection-administrator:qu45fs",
        "name": "purviewmetadatarole_builtin_collection-
administrator:qu45fs",
        "dnfCondition": [
          [
            {
              "attributeName": "principal.microsoft.id",
              "attributeValueIncludedIn": [
                "2f656762-e440-4b62-9eb6-a991d17d64b0"
              ]
            },
            {
              "fromRule": "purviewmetadatarole_builtin_collection-
administrator",
              "attributeName": "derived.purview.role",
              "attributeValueIncludes":
"purviewmetadatarole_builtin_collection-administrator"
            }
          ],
          [
            {
              "fromRule": "purviewmetadatarole_builtin_collection-
administrator:fabrikampurview",
              "attributeName": "derived.purview.permission",
              "attributeValueIncludes":
"purviewmetadatarole_builtin_collection-administrator:fabrikampurview"
            }
          ]
        ]
      },
      {
        "kind": "attributerule",
        "id": "permission:qu45fs",
        "name": "permission:qu45fs",
        "dnfCondition": [
          [
            {
              "fromRule": "purviewmetadatarole_builtin_collection-
administrator:qu45fs",
              "attributeName": "derived.purview.permission",
              "attributeValueIncludes":
"purviewmetadatarole_builtin_collection-administrator:qu45fs"
            }
          ],
          [
            {
              "fromRule": "permission:fabrikampurview",
              "attributeName": "derived.purview.permission",
              "attributeValueIncludes": "permission:fabrikampurview"
            }
          ]
        ]
      }
    ],
    "collection": {
```

```
      "type": "CollectionReference",
      "referenceName": "qu45fs"
    },
    "parentCollectionName": "fabrikampurview"
  }
}
```

Update the collection policy

Ruby

```
PUT
https://{your_purview_account_name}.purview.azure.com/policystore/metadataPo
licies/{policyId}?api-version=2021-07-01
```

In this section, you update the policy JSON that you obtained in the preceding step by adding or removing a user, group, or service principal from the collection. You then push it to the Microsoft Purview service by using a PUT REST method.

Whether you're adding or removing a user, group, or service principal, you'll follow the same API process.

1. Supply the user, group, or service principal object ID {guid} in the "attributeValueIncludedIn" array of the JSON.

2. Search the JSON output of the Get-Policy-by-ID API for the "attributeValueIncludedIn" array in the previous step, and either add or remove the user, group, or service principal object ID in the array. If you're unsure about how to fetch the user, group, or service principal object ID, see Get-MgUser.

3. There are multiple sections in the JSON mapping for each of the four roles. For the Collection Administrator permissions role, use the section bearing the ID called "purviewmetadatarole_builtin_collection-administrator". Likewise, use the corresponding section for the other roles.

4. To better understand the add/remove operation, carefully examine the difference between the JSON output from the previous API and the following output. In the JSON following output, we've added user ID "3a3a3a3a-2c2c-4b4b-1c1c-2a3b4c5d6e7f" As Collection Administrator.

JSON

```
{
  "name": "policy_qu45fs",
  "id": "c6639bb2-9c41-4be0-912b-775750e725de",
```

```
  "version": 0,
  "properties": {
    "description": "",
    "decisionRules": [
      {
        "kind": "decisionrule",
        "effect": "Permit",
        "dnfCondition": [
          [
            {
              "attributeName": "resource.purview.collection",
              "attributeValueIncludes": "qu45fs"
            },
            {
              "fromRule": "permission:qu45fs",
              "attributeName": "derived.purview.permission",
              "attributeValueIncludes": "permission:qu45fs"
            }
          ]
        ]
      }
    ],
    "attributeRules": [
      {
        "kind": "attributerule",
        "id": "purviewmetadatarole_builtin_collection-administrator:qu45fs",
        "name": "purviewmetadatarole_builtin_collection-
administrator:qu45fs",
        "dnfCondition": [
          [
            {
              "attributeName": "principal.microsoft.id",
              "attributeValueIncludedIn": [
                "2f656762-e440-4b62-9eb6-a991d17d64b0",
                "3a3a3a3a-2c2c-4b4b-1c1c-2a3b4c5d6e7f"
              ]
            },
            {
              "fromRule": "purviewmetadatarole_builtin_collection-
administrator",
              "attributeName": "derived.purview.role",
              "attributeValueIncludes":
"purviewmetadatarole_builtin_collection-administrator"
            }
          ],
          [
            {
              "fromRule": "purviewmetadatarole_builtin_collection-
administrator:fabrikampurview",
              "attributeName": "derived.purview.permission",
              "attributeValueIncludes":
"purviewmetadatarole_builtin_collection-administrator:fabrikampurview"
            }
          ]
        ]
```

```
      },
      {
        "kind": "attributerule",
        "id": "permission:qu45fs",
        "name": "permission:qu45fs",
        "dnfCondition": [
          [
            {
              "fromRule": "purviewmetadatarole_builtin_collection-
administrator:qu45fs",
              "attributeName": "derived.purview.permission",
              "attributeValueIncludes":
"purviewmetadatarole_builtin_collection-administrator:qu45fs"
            }
          ],
          [
            {
              "fromRule": "permission:fabrikampurview",
              "attributeName": "derived.purview.permission",
              "attributeValueIncludes": "permission:fabrikampurview"
            }
          ]
        ]
      }
    ],
    "collection": {
      "type": "CollectionReference",
      "referenceName": "qu45fs"
    },
    "parentCollectionName": "fabrikampurview"
  }
}
```

Add the Root Collection Administrator role

By default, the user who created the Microsoft Purview account is the Root Collection Administrator (that is, the administrator of the topmost level of the collection hierarchy). However, in some cases, an organization may want to change the Root Collection Administrator using the API.

Ruby

```
POST
https://management.azure.com/subscriptions/{subscriptionId}/resourceGroups/{
resourceGroupName}/providers/Microsoft.Purview/accounts/{accountName}/addRoo
tCollectionAdmin?api-version=2021-07-01
```

To run the preceding command, you need only to pass the new Root Collection Administrator's object ID. As we mentioned before, the object ID can be that of any

user, group, or service principal.

```JSON
{"objectId": "{guid}"}
```

> **Note**
>
> Users who call this API must have Owner or User Account and Authentication (UAA) permissions on the Microsoft Purview account to execute a write action on the account.

More resources

You may choose to execute Microsoft Purview REST APIs by using the PowerShell utility. It can be readily installed from PowerShell Gallery. With this utility, you can execute all the same commands, but from Windows PowerShell.

Next steps

Purview-API-PowerShell

Feedback

Was this page helpful? Yes No

Create assets using the REST API

Article • 05/14/2025

In this tutorial, you'll learn how to create assets in Microsoft Purview Data Map that users are able to search and browse in Microsoft Purview Unified Catalog.

> **Note**
>
> If you're using the free version of Microsoft Purview, only users in the data curator role group will be able to access these assets.

If you already have your service principal and authentication token, you can skip directly to the steps to create assets using the REST API. Otherwise, follow this guide to call the REST API.

Prerequisites

- If you don't have an Azure subscription, create a free account before you begin.
- You must have an existing Microsoft Purview account. If you don't, check if your organization has access to the free version of Microsoft Purview, or use this quickstart to create a Microsoft Purview account.

Create a service principal (application)

For a REST API client to access Microsoft Purview to create entities, the client must have a service principal (application), and an identity that Microsoft Purview recognizes and is configured to trust.

Customers who have used existing service principals (application IDs) have had a high rate of failure. Therefore, we recommend creating a new service principal for calling APIs.

To create a new service principal:

1. Sign in to the Azure portal.
2. From the portal, search for and select **Microsoft Entra ID**.
3. From the **Microsoft Entra ID** page, select **App registrations** from the left pane.
4. Select **New registration**.
5. On the **Register an application** page:
   a. Enter a **Name** for the application (the service principal name).

b. Select **Accounts in this organizational directory only (<*your tenant's name*> only - Single tenant)**.
c. For **Redirect URI (optional)**, select **Web** and enter a value. This value doesn't need to be a valid endpoint. `https://exampleURI.com` will do.
d. Select **Register**.

6. On the new service principal page, copy the values of the **Display name** and the **Application (client) ID** to save for later.

   The application ID is the `client_id` value in the sample code.

To use the service principal (application), you need to know the service principal's password:

1. Select your service principal (application) in the **App registrations** list.
2. Select **Certificates & secrets** from the left pane.
3. Select **New client secret**.
4. On the **Add a client secret** page, enter a **Description**, select an expiration time under **Expires**, and then select **Add**.
5. On the **Client secrets** page, the string in the **Value** column of your new secret is your password. Save this value.

Set up authentication using service principal

Once the new service principal is created, you need to assign permissions so your service principal can access your Microsoft Purview account. The permissions you need to assign depend on whether you're using the classic Microsoft Purview experience or the new Microsoft Purview portal.

Select the tab for the experience you're using.

Classic governance portal

1. Navigate to your Microsoft Purview governance portal.
2. Select the Data Map in the left menu.
3. Select Collections.
4. Select the root collection in the collections menu. This is the top collection in the list, and will have the same name as your Microsoft Purview account.

> ⓘ **Note**

> You can also assign your service principal permission to any sub-collections, instead of the root collection. However, all APIs will be scoped to that collection (and sub-collections that inherit permissions), and users trying to call the API for another collection will get errors.

5. Select the **Role assignments** tab.
6. Assign the Data Curator role to the service principal created previously. For detailed steps, see Assign Azure roles using the Microsoft Purview governance portal.

After you've assigned permissions, you'll need to gather your authentication token.

1. Send a POST request to the following URL to get access token:

   `https://login.microsoftonline.com/{your-tenant-id}/oauth2/token`

   You can find your Tenant ID by searching for **Tenant Properties** in the Azure portal. The ID is available on the tenant properties page.

   The following parameters need to be passed to the above URL:

   - **client_id**: client ID of the application registered in Microsoft Entra ID and is assigned to a data plane role for the Microsoft Purview account.
   - **client_secret**: client secret created for the above application.
   - **grant_type**: This should be 'client_credentials'.
   - **resource**: This should be 'https://purview.azure.net'

   Here's a sample POST request in PowerShell:

   Azure PowerShell

   ```
   $tenantID = "12a345bc-67d1-ef89-abcd-efg12345abcde"

   $url = "https://login.microsoftonline.com/$tenantID/oauth2/token"
   $params = @{ client_id = "a1234bcd-5678-9012-abcd-abcd1234abcd";
   client_secret = "abcd~a1234bcd56789012abcdabcd1234abcd"; grant_type =
   "client_credentials"; resource = 'https://purview.azure.net' }

   Invoke-WebRequest $url -Method Post -Body $params      -UseBasicParsing |
   ConvertFrom-Json
   ```

   Sample response token:

   JSON

```
{
    "token_type": "Bearer",
    "expires_in": "86399",
    "ext_expires_in": "86399",
    "expires_on": "1621038348",
    "not_before": "1620951648",
    "resource": "https://purview.azure.net",
    "access_token": "<<access token>>"
}
```

> **Tip**
>
> If you get an error message that reads: *Cross-origin token redemption is permitted only for the 'Single-Page Application' client-type.*
>
> - Check your request headers and confirm that your request **doesn't** contain the 'origin' header.
> - Confirm that your redirect URI is set to **web** in your service principal.
> - Make sure your software is up to date for the application you're using to send your POST request.

2. Use the access token above to call the Data plane APIs.

Create assets using the REST API

Now that you have a token and can authenticate, you're ready to create your assets.

> **Important**
>
> Your request URL endpoints depend on what Microsoft Purview experience you're using: the classic Microsoft Purview experience or the new Microsoft Purview portal

Create Azure assets

Here's an example you can use to create your entities for Azure resources. This example covers an Azure Storage account, but you can use it for any other Azure sources.

> **Important**
>
> To use this example for your Azure resources, replace these values in the payload:

- typeName
- Owner
- qualifiedName
- name
- Expert ID
- Expert Info
- Owner ID
- Owner Info
- Createdby
- Updatedby

Classic governance portal

Request URL: `https://{accountName}.purview.azure.com/catalog/api/atlas/v2/entity`

Method: POST

Authentication: Use the token from the previous step as your bearer token.

Payload example:

JSON

```
{
  "referredEntities": {},
  "entity": {
    "typeName": "azure_storage_account",
    "attributes": {
      "owner": "ExampleOwner",
      "modifiedTime": 0,
      "createTime": 0,
      "qualifiedName": "https://exampleaccount.core.windows.net",
      "name": "ExampleStorageAccount",
      "description": null,
      "publicAccessLevel": null
    },
    "contacts": {
      "Expert": [
        {
          "id": "30435ff9-9b96-44af-a5a9-e05c8b1ae2d2",
          "info": "Example Expert Info"
        }
      ],
      "Owner": [
        {
          "id": " 30435ff9-9b96-44af-a5a9-e05c8b1ae2d2",
```

```
            "info": "Example Owner Info"
          }
        ]
      },
      "status": "ACTIVE",
      "createdBy": "ExampleCreator",
      "updatedBy": "ExampleUpdator",
      "version": 0
    }
  }
```

Create multicloud assets

Here's an example you can use to create your entities for multicloud resources. This example creates a Snowflake resource, but you can use it for any other sources

> ⓘ **Important**
>
> To use this example for your Azure resources, replace these values in the payload:
>
> - typeName
> - Owner
> - qualifiedName
> - name
> - type
> - Expert ID
> - Expert Info
> - Owner ID
> - Owner Info
> - Createdby
> - Updatedby

Classic governance portal

Request URL: `https://{accountName}.purview.azure.com/catalog/api/atlas/v2/entity`

Method: POST

Authentication: Use the token from the previous step as your bearer token.

Payload example:

JSON


```
{
  "referredEntities": {},
  "entity": {
    "typeName": "snowflake_table",
    "attributes": {
      "owner": "ExampleOwner",
      "modifiedTime": 0,
      "createTime": 0,
      "qualifiedName": "snowflake://microsoft_partner.east-us-
2.azure.snowflakecomputing.com/databases/AZUREPURVIEW_TESTDB/schemas/COMPANY/table
s/PROJECT_INFO",
      "name": "PROJECT_INFO",
      "description": null,
      "type": "TABLE"
    },
    "contacts": {
      "Expert": [
        {
          "id": "30435ff9-9b96-44af-a5a9-e05c8b1ae2d2",
          "info": "Example Expert Info"
        }
      ],
      "Owner": [
        {
          "id": "4b27e65f-6a15-4925-a4ef-2e640445079b",
          "info": "Example Owner Info"
        }
      ]
    },
    "status": "ACTIVE",
    "createdBy": "ExampleCreator",
    "updatedBy": "ExampleUpdator",
    "version": 0
  }
}
```


Next steps

Manage data sources | Microsoft Purview Entity REST API reference

Create and get lineage relationships using the REST API

Article • 05/14/2025

In this tutorial, you learn how to use Microsoft Purview REST APIs to:

1. Create assets and lineage relationships between the data assets.
2. Query lineage relationships/paths.

APIs referenced in this article:

- Create Entity
- Create Relationship
- Get Lineage by GUID

Prerequisites

The use of these APIs requires Data Curator and Data Reader roles. Check this tutorial for details on how to obtain the access token.

Concepts

Asset

In Microsoft Purview, there are two types of base assets: DataSet and Process.

- DataSet Assets that contain data such as Azure SQL Table, Oracle Table etc. should be inherited from DataSet.
- Process Assets that process data, such as a data pipeline, query, function etc. should be inherited from Process.

Refer to assets and types to understand type definitions of DataSets and Processes.

Relationships and lineage

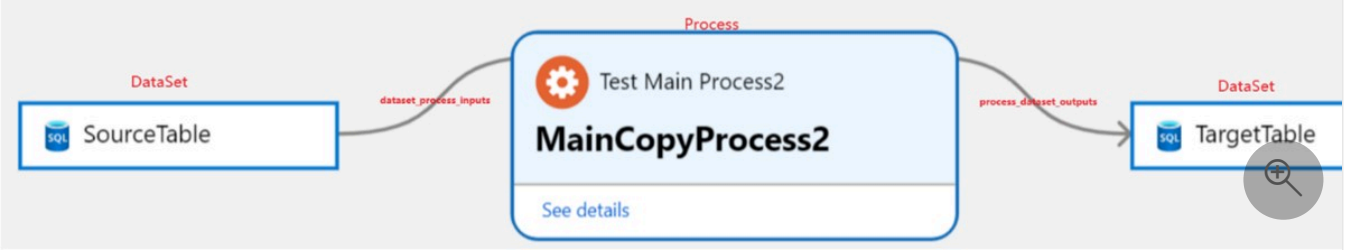
In Microsoft Purview, there are three types of relationships for lineage:

- *dataset_process_inputs*: connects DataSet to Process, which means the DataSet is the input of the Process
- *process_dataset_outputs*: connects Process to DataSet, which means the Process produces the DataSet

- *direct_lineage_dataset_dataset*: connects DataSet1 to DataSet2, which means DataSet1 is the upstream of DataSet2, although we wouldn't know exactly which Process is between them

Example 1: *Here's an example of lineage with 2 DataSets, one Process and two lineage relationships*:

Dataset ---> dataset_process_inputs ---> Process ---> process_dataset_outputs ---> DataSet



Example 2: *Another example of lineage with 2 DataSets and one lineage relationship:*

Dataset ---> direct_lineage_dataset_dataset ---> DataSet



Create assets and lineage relationships between the data assets

In the following sections, let's take hive_table as an example type for DataSet, and hive_query_process as an example type for Process. We'll create assets with those two types and create lineages between them. You can use any other type, which inherits from DataSet or Process to create lineages.

Example 1

Create assets via API

If the assets you want to create lineage for aren't created in Microsoft Purview yet, you can call the following API to create them.

API: Bulk Create Assets

Task: Create two hive_tables as datasets - table1 and table2, each with 2 hive_columns - column1 and column2.

HTTP

```
POST {endpoint}/datamap/api/atlas/v2/entity/bulk
```

With the body:

JSON

```
{
  "entities": [
    {
      "typeName": "hive_table",
      "attributes": {
        "qualifiedName": "test_lineage.table1",
        "name": "table1"
      },
      "relationshipAttributes": {
        "columns": [
          {
            "guid": "-11",
            "typeName": "hive_column"
          },{
            "guid": "-12",
            "typeName": "hive_column"
          }
        ]
      },
      "guid": "-1"
    },
    {
      "typeName": "hive_column",
      "attributes": {
        "qualifiedName": "test_lineage.table1#column1",
        "name": "column1",
        "type": "int"
      },
      "guid": "-11",
      "relationshipAttributes": {
        "table": {
          "guid": "-1",
          "typeName": "hive_table"
        }
      }
    },
    {
      "typeName": "hive_column",
      "attributes": {
```

```
      "qualifiedName": "test_lineage.table1#column2",
      "name": "column2",
      "type": "int"
    },
    "guid": "-12",
    "relationshipAttributes": {
      "table": {
        "guid": "-1",
        "typeName": "hive_table"
      }
    }
  },
  {
    "typeName": "hive_table",
    "attributes": {
      "qualifiedName": "test_lineage.table2",
      "name": "table2"
    },
    "relationshipAttributes": {
      "columns": [
        {
          "guid": "-21",
          "typeName": "hive_column"
        },{
          "guid": "-22",
          "typeName": "hive_column"
        }
      ]
    },
    "guid": "-2"
  },
  {
    "typeName": "hive_column",
    "attributes": {
      "qualifiedName": "test_lineage.table2#column1",
      "name": "column1",
      "type": "int"
    },
    "guid": "-21",
    "relationshipAttributes": {
      "table": {
        "guid": "-2",
        "typeName": "hive_table"
      }
    }
  },
  {
    "typeName": "hive_column",
    "attributes": {
      "qualifiedName": "test_lineage.table2#column2",
      "name": "column2",
      "type": "int"
    },
    "guid": "-22",
    "relationshipAttributes": {
```

```
        "table": {
          "guid": "-2",
          "typeName": "hive_table"
        }
      }
    }
  ]
}
```

Task: Create the process asset 'hive_view_query'

HTTP

```
POST {endpoint}/datamap/api/atlas/v2/entity/bulk
```

With the body:

JSON

```
{
  "entities": [
    {
      "typeName": "hive_view_query",
      "attributes": {
        "qualifiedName": "test_lineage.HiveQuery1",
        "name": "HiveQuery1",
        "columnMapping": "[{\"DatasetMapping\":
{\"Source\":\"test_lineage.table1\",\"Sink\":\"test_lineage.table2\"},\"ColumnMapp
ing\":[{\"Source\":\"column1\",\"Sink\":\"column1\"},
{\"Source\":\"column2\",\"Sink\":\"column2\"}]}]"
      },
      "guid": "-1"
    }
  ]
}
```

The above API calls result in creation of two hive_tables(Datasets) and one hive_view_query(Process).

Create lineage relationships between Datasets and the Process

API: Create Relationship

Task: Create lineage from table1 -> HiveQuery1 (that is, Dataset -> Process)

HTTP

```
POST {endpoint}/datamap/api/atlas/v2/relationship
```

With the body:

JSON

```
{
        "typeName": "dataset_process_inputs",
        "guid": "-1",
        "end1": {
            "typeName": "hive_table",
            "uniqueAttributes": {
                "qualifiedName": "test_lineage.table1"
            }
        },
        "end2": {
            "typeName": "Process",
            "uniqueAttributes": {
                "qualifiedName": "test_lineage.HiveQuery1"
            }
        }
}
```

Task: Create lineage from HiveQuery1 -> table2 (that is, Process -> Dataset)

HTTP

```
POST {endpoint}/datamap/api/atlas/v2/relationship
```

With the body:

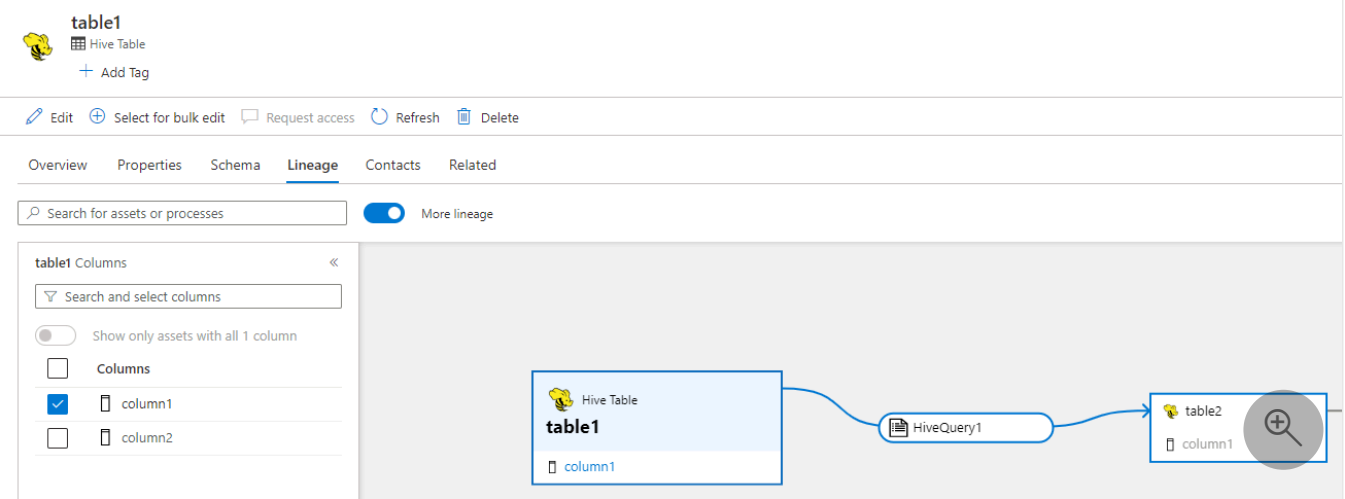
JSON

```
{
        "typeName": "process_dataset_outputs",
        "guid": "-2",
        "end1": {
            "typeName": "Process",
            "uniqueAttributes": {
                "qualifiedName": "test_lineage.HiveQuery1"
            }
        },
        "end2": {
            "typeName": "hive_table",
            "uniqueAttributes": {
                "qualifiedName": "test_lineage.table2"
            }
```

```
        }
}
```

View lineage

Once the assets and lineage relationships are created, you can check the lineage graph in Microsoft Purview:



Example 2

Create a hive table, table3, with two columns

API: Bulk Create Assets

Task: Create table3, with 2 hive_columns, column1 and column2

HTTP

```
POST {endpoint}/datamap/api/atlas/v2/entity/bulk
```

With the body:

JSON

```
{
"entities": [
    {
      "typeName": "hive_table",
      "attributes": {
        "qualifiedName": "test_lineage.table3",
        "name": "table3"
      },
      "relationshipAttributes": {
```

```
        "columns": [
          {
            "guid": "-31",
            "typeName": "hive_column"
          },{
            "guid": "-32",
            "typeName": "hive_column"
          }
        ]
      },
      "guid": "-3"
    },
    {
      "typeName": "hive_column",
      "attributes": {
        "qualifiedName": "test_lineage.table3#column1",
        "name": "column1",
        "type": "int"
      },
      "guid": "-31",
      "relationshipAttributes": {
        "table": {
          "guid": "-3",
          "typeName": "hive_table"
        }
      }
    },
    {
      "typeName": "hive_column",
      "attributes": {
        "qualifiedName": "test_lineage.table3#column2",
        "name": "column2",
        "type": "int"
      },
      "guid": "-32",
      "relationshipAttributes": {
        "table": {
          "guid": "-3",
          "typeName": "hive_table"
        }
      }
    }
  ]
}
```

Create direct lineage between table 2 and table 3, with column mapping

API: Create Relationship

Task: Create lineage from table2 -> table3 (that is, Dataset -> Dataset) with column mapping

```
HTTP
```

```
POST {endpoint}/datamap/api/atlas/v2/relationship
```

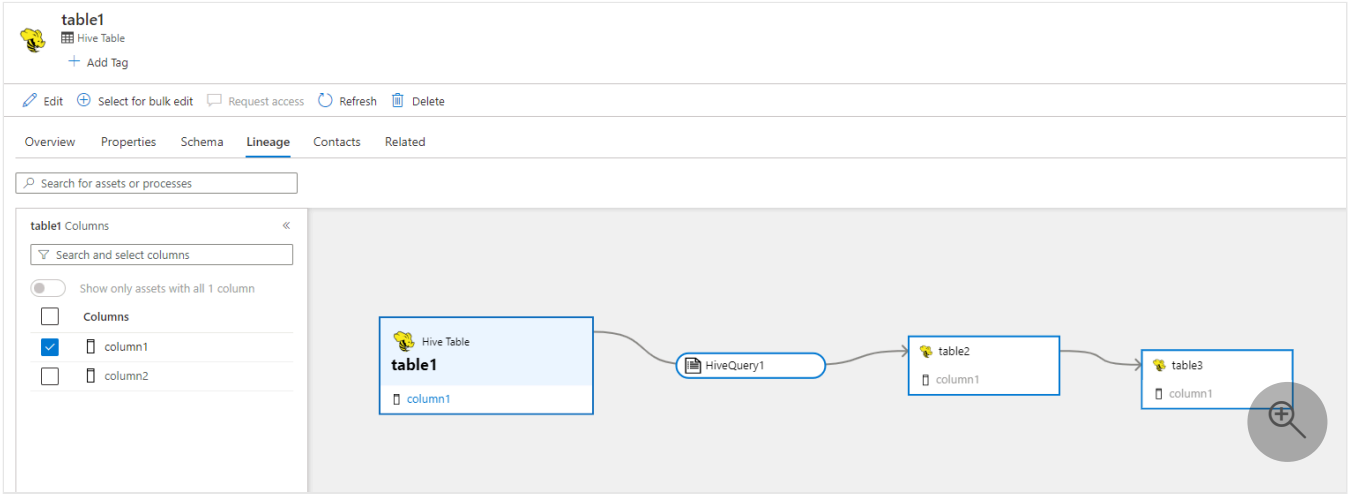
With the body:

JSON

```
{
    "typeName": "direct_lineage_dataset_dataset",
    "guid": "-1",
    "end1": {
        "typeName": "hive_table",
        "uniqueAttributes": {
            "qualifiedName": "test_lineage.table2"
        }
    },
    "end2": {
        "typeName": " hive_table ",
        "uniqueAttributes": {
            "qualifiedName": "test_lineage.table3"
        }
    },
    "attributes": {
      "columnMapping": "[{\"Source\":\"column1\",\"Sink\":\"column1\"},
{\"Source\":\"column2\",\"Sink\":\"column2\"}]"
    }
}
```

View lineage

Now the lineage graph (together from Example 1 & Example 2 above) becomes:



Notice that table2 is directly linked to table3, without a HiveQuery between them.

Query lineage relationships/paths

API: Get Lineage by GUID

Task: Get lineage of table2 via REST API

HTTP

```
GET {{endpoint}}/api/atlas/v2/lineage/{{guid_of_table2}}?direction=BOTH
```

Below you can JSON response payload:

JSON


```
{
    "baseEntityGuid": "2a12b3ff-5816-4222-833a-035bf82e06e0",
    "lineageDirection": "BOTH",
    "lineageDepth": 3,
    "lineageWidth": -1,
    "childrenCount": -1,
    "guidEntityMap": {
        "16b93b78-8683-4f88-9651-24c4a9d797b0": {
            "typeName": "hive_table",
            "attributes": {
                "temporary": false,
                "lastAccessTime": 0,
                "createTime": 0,
                "qualifiedName": "test_lineage.table3",
                "name": "table3",
                "retention": 0
            },
            "lastModifiedTS": "1",
            "guid": "16b93b78-8683-4f88-9651-24c4a9d797b0",
            "status": "ACTIVE",
            "displayText": "table3",
            "classificationNames": [],
            "meaningNames": [],
            "meanings": [],
            "isIncomplete": false,
            "labels": [],
            "isIndexed": true
        },
        "cb22ba23-47a2-4149-ade6-e3d9642fe592": {
            "typeName": "hive_table",
            "attributes": {
                "temporary": false,
                "lastAccessTime": 0,
                "createTime": 0,
                "qualifiedName": "test_lineage.table1",
                "name": "table1",
                "retention": 0
            },
            "lastModifiedTS": "1",
            "guid": "cb22ba23-47a2-4149-ade6-e3d9642fe592",
```

```
            "status": "ACTIVE",
            "displayText": "table1",
            "classificationNames": [],
            "meaningNames": [],
            "meanings": [],
            "isIncomplete": false,
            "labels": [],
            "isIndexed": true
        },
        "bbeacce6-5bde-46f7-8fe4-689cbb36ba51": {
            "typeName": "hive_view_query",
            "attributes": {
                "qualifiedName": "test_lineage.HiveQuery1",
                "name": "HiveQuery1",
                "columnMapping": "[{\"DatasetMapping\":
{\"Source\":\"test_lineage.table1\",\"Sink\":\"test_lineage.table2\"},\"ColumnMapp
ing\":[{\"Source\":\"column1\",\"Sink\":\"column1\"},
{\"Source\":\"column2\",\"Sink\":\"column2\"}]}]"
            },
            "lastModifiedTS": "1",
            "guid": "bbeacce6-5bde-46f7-8fe4-689cbb36ba51",
            "status": "ACTIVE",
            "displayText": "HiveQuery1",
            "classificationNames": [],
            "meaningNames": [],
            "meanings": [],
            "isIncomplete": false,
            "labels": [],
            "isIndexed": true
        },
        "2a12b3ff-5816-4222-833a-035bf82e06e0": {
            "typeName": "hive_table",
            "attributes": {
                "temporary": false,
                "lastAccessTime": 0,
                "createTime": 0,
                "qualifiedName": "test_lineage.table2",
                "name": "table2",
                "retention": 0
            },
            "lastModifiedTS": "1",
            "guid": "2a12b3ff-5816-4222-833a-035bf82e06e0",
            "status": "ACTIVE",
            "displayText": "table2",
            "classificationNames": [],
            "meaningNames": [],
            "meanings": [],
            "isIncomplete": false,
            "labels": [],
            "isIndexed": true
        }
    },
    "includeParent": false,
    "relations": [
        {
```

```
            "fromEntityId": "2a12b3ff-5816-4222-833a-035bf82e06e0",
            "toEntityId": "16b93b78-8683-4f88-9651-24c4a9d797b0",
            "relationshipId": "23df8e3e-b066-40b2-be29-9fd90693c51b",
            "columnMapping": "[{\"Source\":\"column1\",\"Sink\":\"column1\"},
{\"Source\":\"column2\",\"Sink\":\"column2\"}]"
        },
        {
            "fromEntityId": "bbeacce6-5bde-46f7-8fe4-689cbb36ba51",
            "toEntityId": "2a12b3ff-5816-4222-833a-035bf82e06e0",
            "relationshipId": "5fe8d378-39cd-4c6b-8ced-91b0152d3014"
        },
        {
            "fromEntityId": "cb22ba23-47a2-4149-ade6-e3d9642fe592",
            "toEntityId": "bbeacce6-5bde-46f7-8fe4-689cbb36ba51",
            "relationshipId": "73e084bf-98a3-45fb-a1e4-c56cc40661b8"
        }
    ],
    "parentRelations": [],
    "widthCounts": {
        "INPUT": {
            "cb22ba23-47a2-4149-ade6-e3d9642fe592": 0,
            "bbeacce6-5bde-46f7-8fe4-689cbb36ba51": 1,
            "2a12b3ff-5816-4222-833a-035bf82e06e0": 1
        },
        "OUTPUT": {
            "16b93b78-8683-4f88-9651-24c4a9d797b0": 0,
            "2a12b3ff-5816-4222-833a-035bf82e06e0": 1
        }
    }
}
```

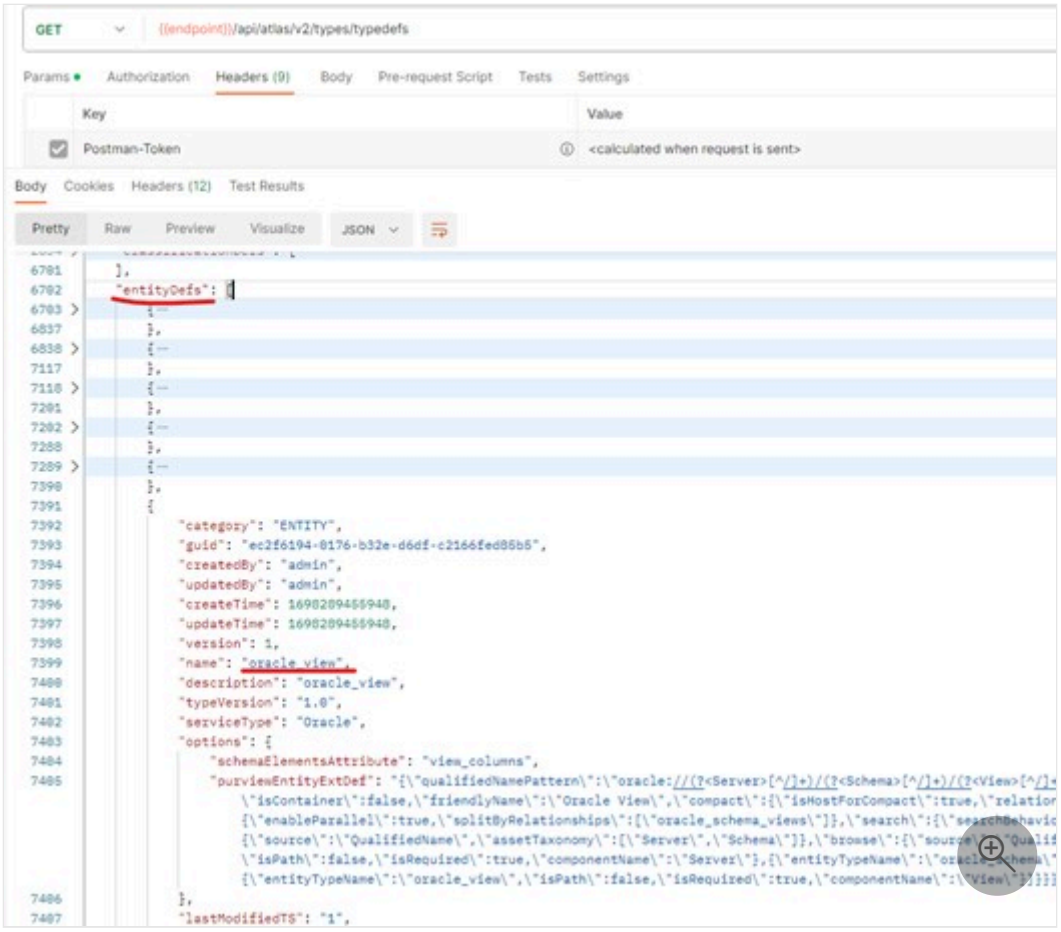

Other Resources

Type Definitions

All Assets/Entities and Relationships are defined in the type system.

Call list all type definitions API to get current type definitions in your Microsoft Purview instance. If you're not sure which typeName to use in the API calls described, you can check the type definitions API to find the appropriate Asset/Entity type.

Following is a sample response of this API:



You can see that in the entityDefs in the above response, asset types (such as oracle_table, oracle_view etc.) are defined. A further look at the definition of the asset (in this example, oracle_view) shows that the asset is inherited DataSet type.



Similarly, you can find that a process (for example, Oracle_function) is inherited from “Process” type as shown:



Create New Custom Types

If you want to create custom assets or processes, you can use the following APIs.

API: Create Typedef API

Task: Create a Custom Process Type

HTTP

```
POST {endpoint}/datamap/api/atlas/v2/types/typedefs
```

With the body:

JSON

```
{
  "enumDefs": [],
  "structDefs": [],
  "classificationDefs": [],
  "entityDefs": [
    {
      "name": "MyCustomServiceProcess",
      "superTypes": [
        "Process"
      ],
      "typeVersion": "1.0",
      "attributeDefs": []
    }
  ],
  "relationshipDefs": []
}
```

Task: Create a Custom DataSet Type

HTTP

```
POST {endpoint}/datamap/api/atlas/v2/types/typedefs
```

With the body:

JSON

```
{
  "enumDefs": [],
  "structDefs": [],
  "classificationDefs": [],
  "entityDefs": [
    {
      "name": "MyCustomDataSet",
      "superTypes": [
        "DataSet"
      ],
      "typeVersion": "1.0",
      "attributeDefs": []
    }
  ],
  "relationshipDefs": []
}
```

Microsoft Purview - Build custom lineage with the REST APIs

Article • 02/24/2025

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

This article provides the steps to create data lineage entries using the REST API in the Microsoft Purview Data Catalog. In scenarios where the automatically generated lineage in Microsoft Purview is incomplete or missing, lineage can be custom built either manually in the Microsoft Purview portal or using the REST APIs. This article focuses on using REST APIs that can overcome manual lineage known limitations and provide more options.

Background

The ability to show the lineage between datasets is one of Microsoft Purview's important platform features. Systems like Data Factory, Data Share, and Power BI capture the lineage of data as it moves. In certain situations, Microsoft Purview's automatically generated lineage is incomplete or missing for practical visualization and enterprise reporting purposes. In those scenarios, custom lineage reporting is supported by Apache Atlas hooks and the REST API.

Using the REST APIs to build custom lineage allows you to overcome some of manual lineage's limitations as described in the following articles:

- Power BI lineage in Microsoft Purview
- Microsoft Purview manual lineage user guide
- Lineage - known limitations

The rest of this article explains the use of Microsoft Purview REST APIs to build and report custom lineage on Microsoft Purview.

Prerequisites

- For step-by-step guidance on how to use the Purview Python SDK Packages, refer Purview Python SDK Tutorial and Exploring Purview's REST API with Python
- For step-by-step guidance on how to use the *Purview REST APIs*, refer Purview Python SDK Tutorial

Scenarios

There are two use cases when building custom lineage becomes necessary:

A. Create new entities created and link them with lineage

B. Link existing entities or lineage to another existing entity or lineage

A. Create new entities and link them with lineage

As an example, lineage needs to be reported between entities A & B, but A & B don't exist currently.

To create the entities A & B, invoke the Microsoft Purview REST API: *Entity - Bulk Create Or Update - REST API*

HTTP

```
POST
https://{accountname}.purview.azure.com/datamap/api/atlas/v2/entity/bulk?
api-version=2023-09-01
```

Python

```
sample_entity_json = '{"entity": {"status": "ACTIVE","version": 0,"name":
ENTITY_A"}.......{"entity": ........}}'
#Send POST JSON containing entities to be created
CreateOrUpdateEntitesUrl =
'https://<purview_account_name>.purview.azure.com/datamap/api/atlas/v2/entit
y/bulk'
EntitiesResponse = requests.post(CreateOrUpdateEntitesUrl, json =
json.loads(sample_entity_json) ,headers=headers)
entitiesRes = json.loads(EntitiesResponse.text)
```

API Response "201 Created" indicates entities are successfully created and their respective GUIDs are contained in the output JSON.

Now that the entities A & B are created, move to step B to link the entities in the lineage chain using the same REST API.

B. Link existing entities or lineage to another existing entity or lineage

- If the number of entities to be linked isn't time or resource intensive (for example, less than 20-30 entities), you can connect the lineage manually in the Microsoft Purview portal. Follow the manual lineage user guide for steps to manually create lineage connections.
- If you have a high number of lineage connections to make, the process needs to be automated, or if manual lineage using the Microsoft Purview portal isn't feasible, proceed with the API process of linking and building custom lineage.

Custom Lineage JSON Payload:

Execute the *POST /entity/bulk* *Entity - Bulk Create Or Update - REST API* using the payload as illustrated:

HTTP

```
POST
https://{accountname}.purview.azure.com/datamap/api/atlas/v2/entity/bulk?
api-version=2023-09-01
```

Python

```
sample_entity_json = '{
  "entities": [
    {
      "status": "ACTIVE",
      "version": 1,
      "typeName": "Process",
      "attributes": {
        "inputs": [
          {
            "guid": "24558fd8-9cdc-47de-9310-56a58108bab0",
            “guid”: “27163581-9aca-212a-782a-213612639abc”
          }
        ],
        "outputs": [
          {
            "guid": "e33c694a-2c4f-4cae-8c27-06f6f6f60000"
          }
        ],
        "qualifiedName": "cassandra://query",
        "name": "query"
      }
    }
  ]
}'
```

```
#In this code snippet, we send the JSON as POST request containing the two GUIDs as input and "output" GUID as output. This creates lineage with 2 directional inputs and 1 directional output.
#Note: using the same API and SDK code you can create lineage with any number of inputs, any number of processes in between, any number of typedefs and any number of outputs.
#The API/SDK method is the most flexible and versatile menthod of creating lineage.

CreateLineageEntitesUrl = 'https://<purview_account_name>.purview.azure.com/datamap/api/atlas/v2/entity/bulk'
EntitiesResponse = requests.post(CreateLineageEntitesUrl, json = json.loads(sample_entity_json),headers=headers)
entitiesRes = json.loads(EntitiesResponse.text)
```

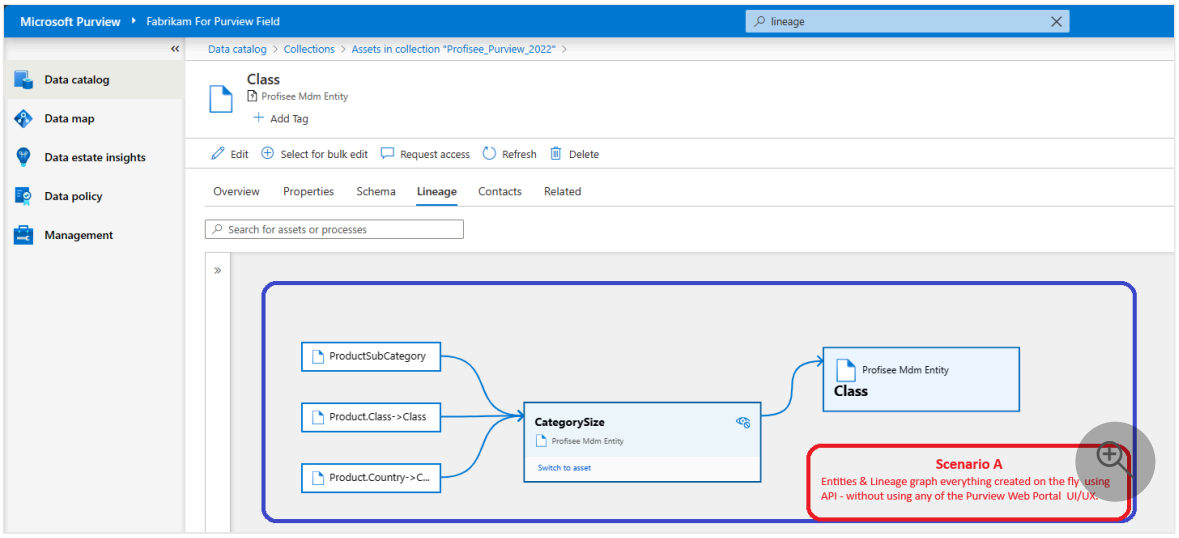
This JSON payload creates the custom lineage. It works for already existing assets whose GUIDs are supplied in the "inputs" JSON. For example "guid": "24558fd8-9cdc-47de-9310-56a58108bab0" and "27163581-9aca-212a-782a-213612639abc" refers to the directional input of the lineage and "guid": "e33c694a-2c4f-4cae-8c27-06f6f6f60000" refers to the directional output of the lineage, which was an existing asset automatically scanned by Purview. We just created the lineage between the two assets.

> ⓘ **Note**
>
> If the assets don't exist already, you need to run Bulk Entity create API before this step to create those entities before creating the lineage relationship. The bulk entity creation process using the POST /entity/bulk API and Python code snippet is described in step A.

Scenario Outcomes

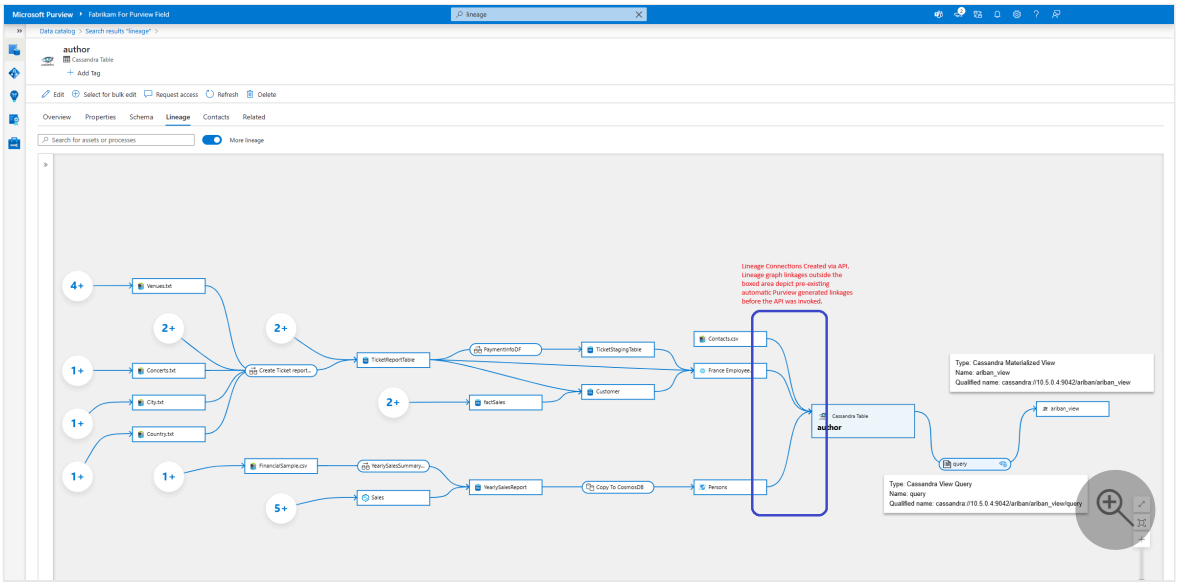
API Response "201 Created" indicates successful lineage graph linkage creation and the created GUIDs are contained in the output JSON. Lineage appears in the Microsoft Purview portal:

- Scenario A: Custom built lineage from assets created via API:



- Scenario B: Custom built lineage from pre-existing assets linked via API.

> **Note**
>
> If the lineage is created from pre-existing entities, observe that the pre-existing lineage graph remains intact and the new linkage is created and displayed additionally.


Microsoft Purview
Fabrikam For Purview Field
lineage
author
Cassandra Table
Edit
Select for bulk edit
Request access
Refresh
Delete
Overview
Properties
Schema
Lineage
Contacts
Related
Search for assets or processes
More lineage
Lineage Connections Created via API.
Lineage graph linkages outside the boxed area depict pre-existing automatic Purview generated linkages before the API was invoked.
Venues.txt
Concerts.txt
City.txt
Country.txt
Create Ticket report...
TicketReportTable
PaymentInfoDF
TicketStagingTable
Contacts.csv
France Employee
Customer
factSales
FinancialSample.csv
YearlySalesSummary...
Sales
YearlySalesReport
Copy To CosmosDB
Persons
Cassandra Table
author
Type: Cassandra Materialized View
Name: ariban_view
Qualified name: cassandra://10.5.0.4:9042/aribanaribanaribana

Tutorial: Use the Microsoft Purview Python SDK

Article • 05/14/2025

This tutorial will introduce you to using the Microsoft Purview Python SDK. You can use the SDK to do all the most common Microsoft Purview operations programmatically, rather than through the Microsoft Purview governance portal.

In this tutorial, you'll learn how us the SDK to:

- ✓ Grant the required rights to work programmatically with Microsoft Purview
- ✓ Register a Blob Storage container as a data source in Microsoft Purview
- ✓ Define and run a scan
- ✓ Search the catalog
- ✓ Delete a data source

Prerequisites

For this tutorial, you'll need:

- Python 3.6 or higher
- An active Azure Subscription. If you don't have one, you can create one for free.
- A Microsoft Entra tenant associated with your subscription.
- An Azure Storage account. If you don't already have one, you can follow our quickstart guide to create one.
- A Microsoft Purview account. If you don't already have one, you can follow our quickstart guide to create one.
- A service principal with a client secret.

> **ⓘ Important**
>
> For these scripts your endpoint value will be different depending on which Microsoft Purview portal you are using. Endpoint for the classic Microsoft Purview governance portal: purview.azure.com/ Endpoint for the New Microsoft Purview portal: purview.microsoft.com/
>
> So if you're using the new portal, your endpoint value will be something like: "https://consotopurview.scan.purview.microsoft.com"

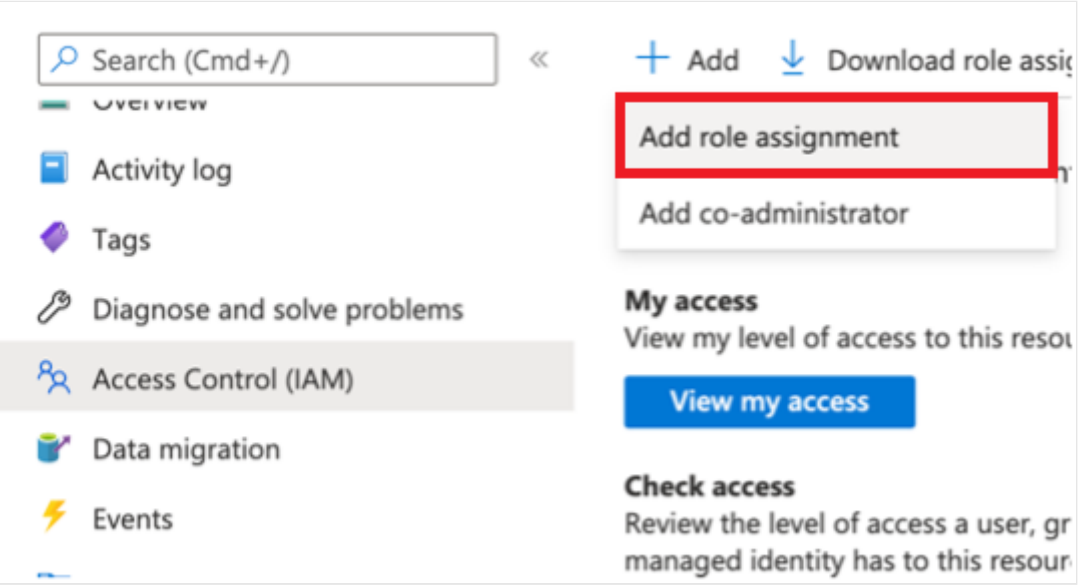
Give Microsoft Purview access to the Storage account

Before being able to scan the content of the Storage account, you need to give Microsoft Purview the right role.

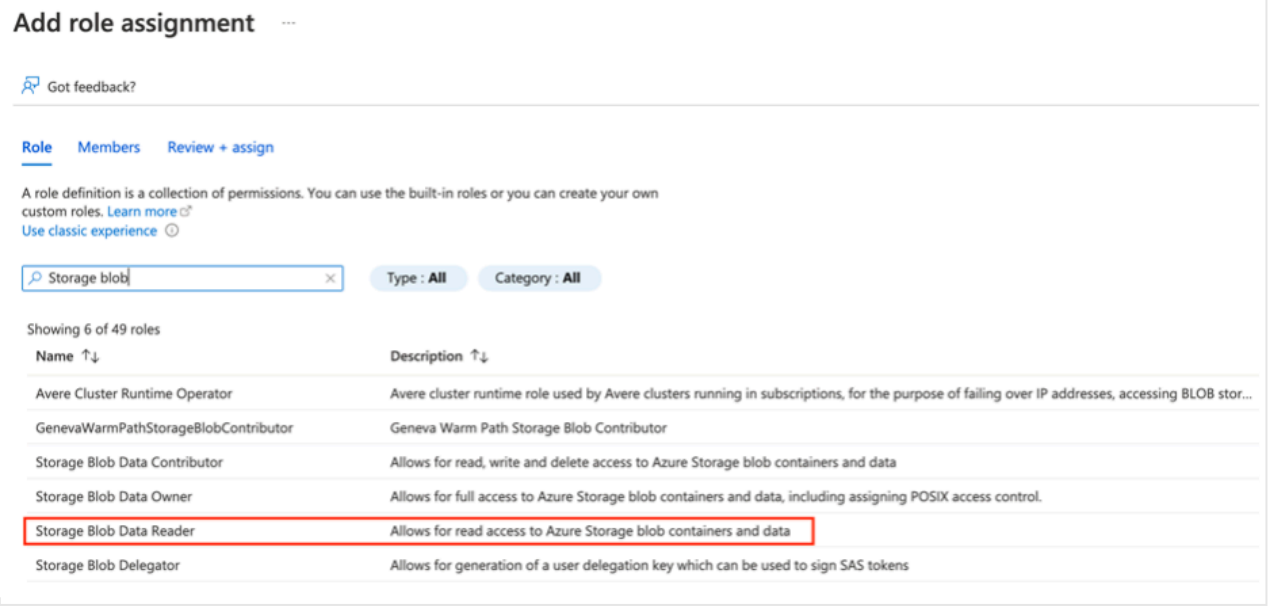
1. Go to your Storage Account through the Azure portal.

2. Select Access Control (IAM).

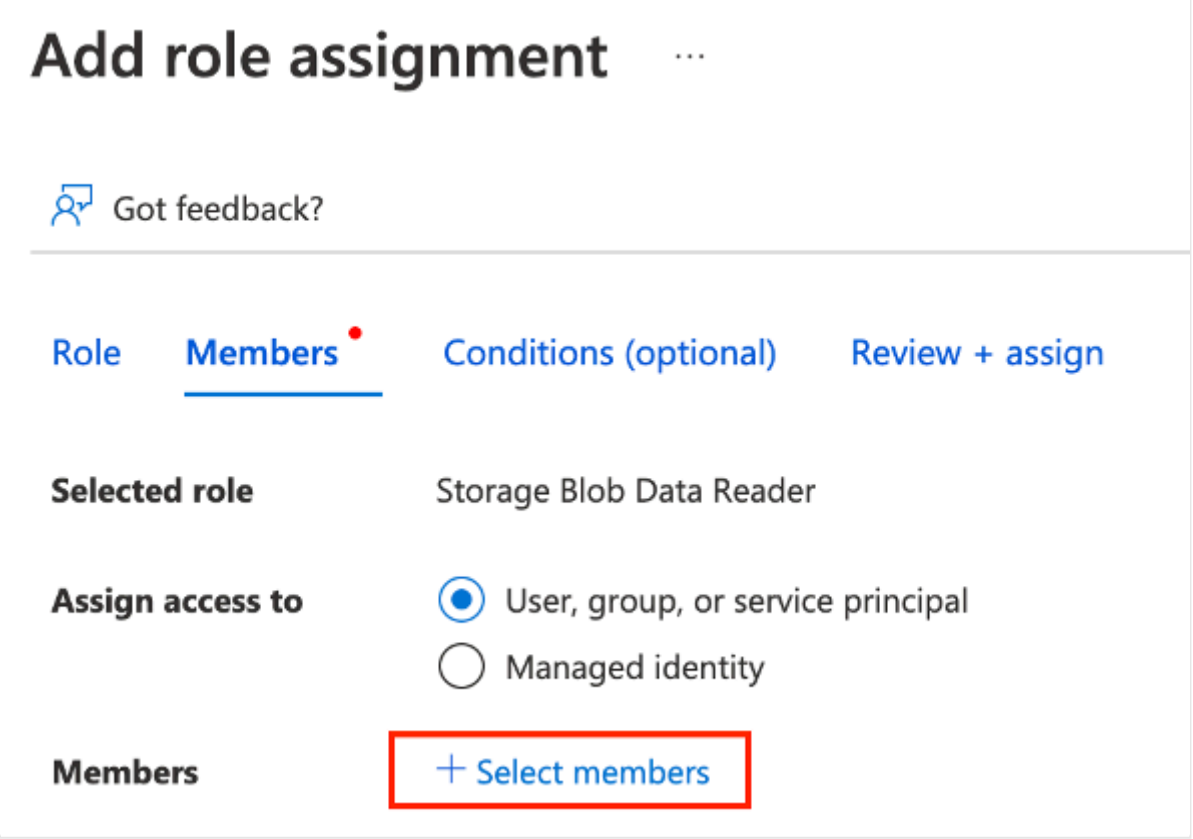
3. Select the Add button and select **Add role assignment**.



4. In the next window, search for the **Storage blob Reader** role and select it:



5. Then go on the **Members** tab and select **Select members**:



6. A new pane appears on the right. Search and select the name of your existing Microsoft Purview instance.

7. You can then select **Review + Assign**.

Microsoft Purview now has the required reading right to scan your Blob Storage.

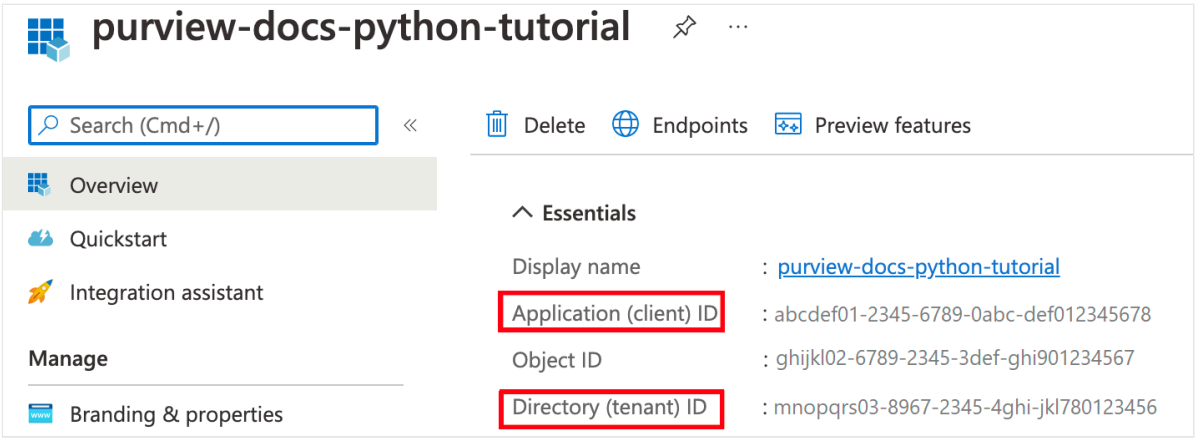
Grant your application the access to your Microsoft Purview account

1. First, you'll need the Client ID, Tenant ID, and Client secret from your service principal. To find this information, select your **Microsoft Entra ID**.

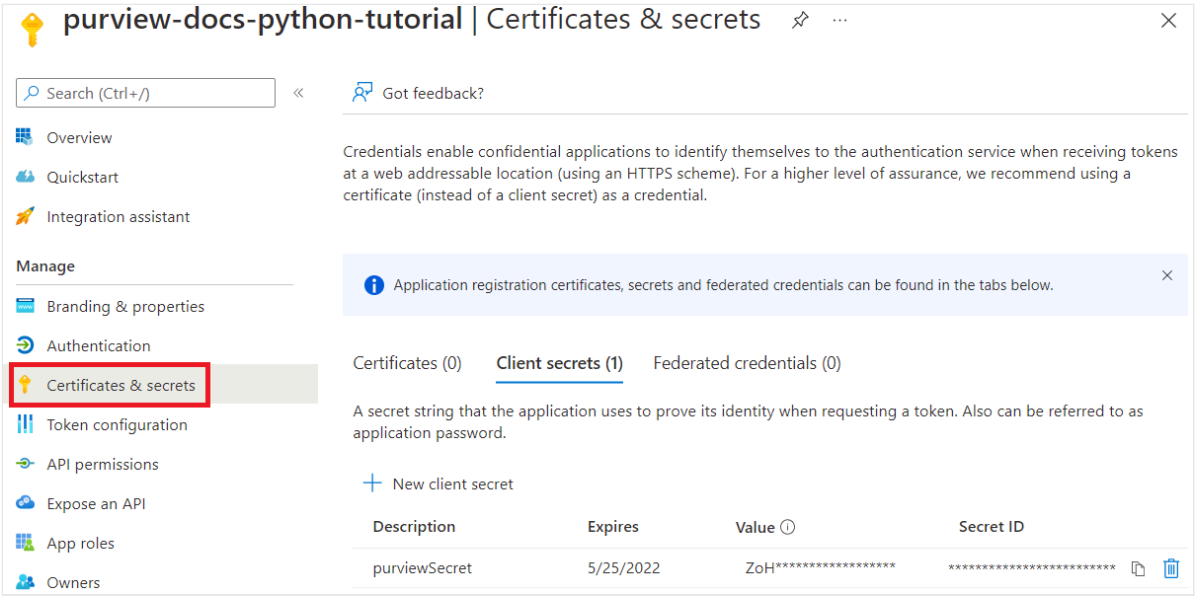
2. Then, select **App registrations**.

3. Select your application and locate the required information:

    - Name
    - Client ID (or Application ID)
    - Tenant ID (or Directory ID)

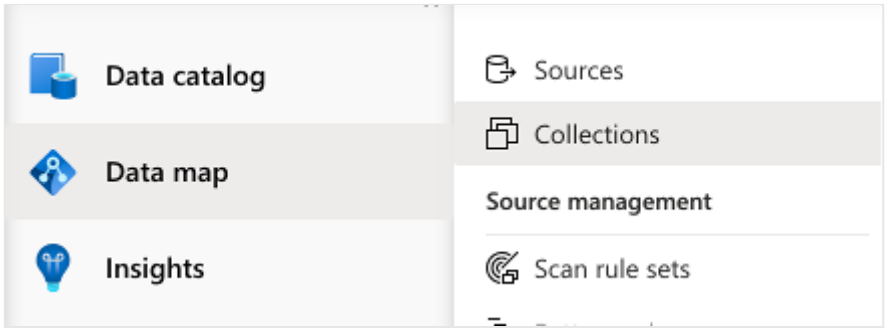


- Client secret



4. You now need to give the relevant Microsoft Purview roles to your service principal. To do so, access your Microsoft Purview instance. Select **Open Microsoft Purview governance portal** or open the Microsoft Purview's governance portal directly and choose the instance that you deployed.

5. Inside the Microsoft Purview governance portal, select **Data map**, then **Collections**:

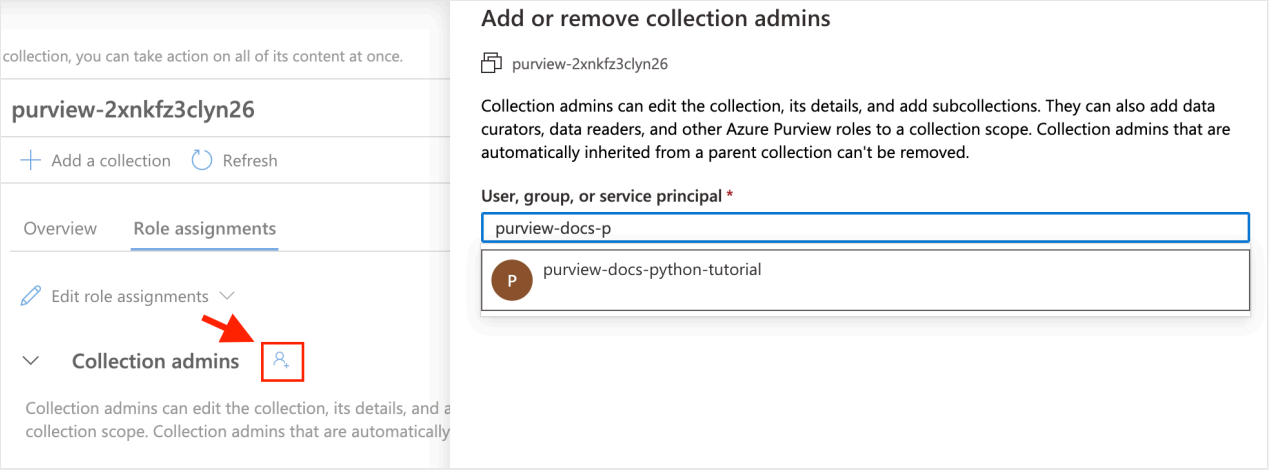


6. Select the collection you want to work with, and go on the **Role assignments** tab. Add the service principal in the following roles:

    - Collection admins

- Data source admins
- Data curators
- Data readers

7. For each role, select the **Edit role assignments** button and select the role you want to add the service principal to. Or select the **Add** button next to each role, and add the service principal by searching its name or Client ID as shown below:



Install the Python packages

1. Open a new command prompt or terminal
2. Install the Azure identity package for authentication:

```
Bash

pip install azure-identity
```

3. Install the Microsoft Purview Scanning Client package:

```
Bash

pip install azure-purview-scanning
```

4. Install the Microsoft Purview Administration Client package:

```
Bash

pip install azure-purview-administration
```

5. Install the Microsoft Purview Client package:

```
Bash
```

```
pip install azure-purview-catalog
```

6. Install the Microsoft Purview Account package:

```Bash
pip install azure-purview-account
```

7. Install the Azure Core package:

```Bash
pip install azure-core
```

Create Python script file

Create a plain text file, and save it as a Python script with the suffix .py. For example: tutorial.py.

Instantiate a Scanning, Catalog, and Administration client

In this section, you learn how to instantiate:

- A scanning client useful to registering data sources, creating and managing scan rules, triggering a scan, etc.
- A catalog client useful to interact with the catalog through searching, browsing the discovered assets, identifying the sensitivity of your data, etc.
- An administration client is useful for interacting with the Microsoft Purview Data Map itself, for operations like listing collections.

First you need to authenticate to your Microsoft Entra ID. For this, you'll use the client secret you created.

1. Start with required import statements: our three clients, the credentials statement, and an Azure exceptions statement.

```Python
from azure.purview.scanning import PurviewScanningClient
from azure.purview.catalog import PurviewCatalogClient
from azure.purview.administration.account import PurviewAccountClient
```

```
from azure.identity import ClientSecretCredential
from azure.core.exceptions import HttpResponseError
```

2. Specify the following information in the code:

   - Client ID (or Application ID)
   - Tenant ID (or Directory ID)
   - Client secret

Python

```
client_id = "<your client id>"
client_secret = "<your client secret>"
tenant_id = "<your tenant id>"
```

3. Specify your endpoint:

> **ⓘ Important**
>
> Your endpoint value will be different depending on which Microsoft Purview portal you are using. Endpoint for the classic Microsoft Purview governance portal: `https://{your_purview_account_name}.purview.azure.com/` Endpoint for the New Microsoft Purview portal: `https://api.purview-service.microsoft.com`
>
> Scan endpoint for the classic Microsoft Purview governance portal: `https://{your_purview_account_name}.scan.purview.azure.com/` Endpoint for the New Microsoft Purview portal: `https://api.scan.purview-service.microsoft.com`

Python

```
purview_endpoint = "<endpoint>"

purview_scan_endpoint = "<scan endpoint>"
```

4. You can now instantiate the three clients:

Python

```
def get_credentials():
    credentials = ClientSecretCredential(client_id=client_id,
client_secret=client_secret, tenant_id=tenant_id)
    return credentials

def get_purview_client():
    credentials = get_credentials()
```

```
    client = PurviewScanningClient(endpoint=purview_scan_endpoint,
credential=credentials, logging_enable=True)
    return client

def get_catalog_client():
    credentials = get_credentials()
    client = PurviewCatalogClient(endpoint=purview_endpoint,
credential=credentials, logging_enable=True)
    return client

def get_admin_client():
    credentials = get_credentials()
    client = PurviewAccountClient(endpoint=purview_endpoint,
credential=credentials, logging_enable=True)
    return client
```

Many of our scripts will start with these same steps, as we'll need these clients to interact with the account.

Register a data source

In this section, you'll register your Blob Storage.

1. Like we discussed in the previous section, first you'll import the clients you'll need to access your Microsoft Purview account. Also import the Azure error response package so you can troubleshoot, and the ClientSecretCredential to construct your Azure credentials.

   Python

   ```
   from azure.purview.administration.account import PurviewAccountClient
   from azure.purview.scanning import PurviewScanningClient
   from azure.core.exceptions import HttpResponseError
   from azure.identity import ClientSecretCredential
   ```

2. Gather the resource ID for your storage account by following this guide: get the resource ID for a storage account.

3. Then, in your Python file, define the following information to be able to register the Blob storage programmatically:

   Python

   ```
   storage_name = "<name of your Storage Account>"
   storage_id = "<id of your Storage Account>"
   rg_name = "<name of your resource group>"
   rg_location = "<location of your resource group>"
   reference_name_purview = "<name of your Microsoft Purview account>"
   ```

4. Provide the name of the collection where you'd like to register your blob storage. (It should be the same collection where you applied permissions earlier. If it isn't, first apply permissions to this collection.) If it's the root collection, use the same name as your Microsoft Purview instance.

```
Python

collection_name = "<name of your collection>"
```

5. Create a function to construct the credentials to access your Microsoft Purview account:

```
Python

client_id = "<your client id>"
client_secret = "<your client secret>"
tenant_id = "<your tenant id>"


def get_credentials():
     credentials = ClientSecretCredential(client_id=client_id,
client_secret=client_secret, tenant_id=tenant_id)
     return credentials
```

6. All collections in the Microsoft Purview Data Map have a **friendly name** and a **name**.

   - The **friendly name** name is the one you see on the collection. For example: Sales.
   - The **name** for all collections (except the root collection) is a six-character name assigned by the data map.

   Python needs this six-character name to reference any sub collections. To convert your **friendly name** automatically to the six-character collection name needed in your script, add this block of code:

> **ⓘ Important**
>
> Your endpoint value will be different depending on which Microsoft Purview portal you are using. Endpoint for the classic Microsoft Purview governance portal: purview.azure.com/ Endpoint for the New Microsoft Purview portal: purview.microsoft.com/
>
> So if you're using the new portal, your endpoint value will be something like: "https://consotopurview.scan.purview.microsoft.com"

```
Python
```

```
def get_admin_client():
     credentials = get_credentials()
     client = PurviewAccountClient(endpoint=purview_endpoint,
credential=credentials, logging_enable=True)
     return client

try:
  admin_client = get_admin_client()
except ValueError as e:
    print(e)

collection_list = client.collections.list_collections()
 for collection in collection_list:
  if collection["friendlyName"].lower() == collection_name.lower():
      collection_name = collection["name"]
```

7. For both clients, and depending on the operations, you also need to provide an input body. To register a source, you'll need to provide an input body for data source registration:

Python

```
ds_name = "<friendly name for your data source>"

body_input = {
        "kind": "AzureStorage",
        "properties": {
            "endpoint": f"https://{storage_name}.blob.core.windows.net/",
            "resourceGroup": rg_name,
            "location": rg_location,
            "resourceName": storage_name,
            "resourceId": storage_id,
            "collection": {
                "type": "CollectionReference",
                "referenceName": collection_name
            },
            "dataUseGovernance": "Disabled"
        }
}
```

8. Now you can call your Microsoft Purview clients and register the data source.

> **ⓘ Important**
>
> Your endpoint value will be different depending on which Microsoft Purview portal you are using. Endpoint for the classic Microsoft Purview governance portal: `https://{your_purview_account_name}.purview.azure.com/` Endpoint for the New Microsoft Purview portal: `https://api.purview-service.microsoft.com`

If you're using the classic portal, your endpoint value will be: `https://{your_purview_account_name}.scan.purview.azure.com` If you're using the new portal, your endpoint value will be: `https://scan.api.purview-service.microsoft.com`

Python

```
def get_purview_client():
     credentials = get_credentials()
     client = PurviewScanningClient(endpoint={{ENDPOINT}},
credential=credentials, logging_enable=True)
     return client

try:
    client = get_purview_client()
except ValueError as e:
    print(e)

try:
    response = client.data_sources.create_or_update(ds_name, body=body_input)
    print(response)
    print(f"Data source {ds_name} successfully created or updated")
except HttpResponseError as e:
    print(e)
```

When the registration process succeeds, you can see an enriched body response from the client.

In the following sections, you'll scan the data source you registered and search the catalog. Each of these scripts will be similarly structured to this registration script.

Full code

Python

```
from azure.purview.scanning import PurviewScanningClient
from azure.identity import ClientSecretCredential
from azure.core.exceptions import HttpResponseError
from azure.purview.administration.account import PurviewAccountClient

client_id = "<your client id>"
client_secret = "<your client secret>"
tenant_id = "<your tenant id>"
purview_endpoint = "<endpoint>"
purview_scan_endpoint = "<scan endpoint>"
storage_name = "<name of your Storage Account>"
storage_id = "<id of your Storage Account>"
rg_name = "<name of your resource group>"
```

```
rg_location = "<location of your resource group>"
collection_name = "<name of your collection>"
ds_name = "<friendly data source name>"

def get_credentials():
    credentials = ClientSecretCredential(client_id=client_id,
client_secret=client_secret, tenant_id=tenant_id)
    return credentials

def get_purview_client():
    credentials = get_credentials()
    client = PurviewScanningClient(endpoint=purview_scan_endpoint,
credential=credentials, logging_enable=True)
    return client

def get_admin_client():
    credentials = get_credentials()
    client = PurviewAccountClient(endpoint=purview_endpoint,
credential=credentials, logging_enable=True)
    return client

try:
    admin_client = get_admin_client()
except ValueError as e:
        print(e)

collection_list = admin_client.collections.list_collections()
for collection in collection_list:
    if collection["friendlyName"].lower() == collection_name.lower():
        collection_name = collection["name"]


body_input = {
    "kind": "AzureStorage",
    "properties": {
        "endpoint": f"https://{storage_name}.blob.core.windows.net/",
        "resourceGroup": rg_name,
        "location": rg_location,
        "resourceName": storage_name,
        "resourceId": storage_id,
        "collection": {
            "type": "CollectionReference",
            "referenceName": collection_name
        },
        "dataUseGovernance": "Disabled"
    }
}

try:
    client = get_purview_client()
except ValueError as e:
        print(e)

try:
    response = client.data_sources.create_or_update(ds_name, body=body_input)
```

```
    print(response)
    print(f"Data source {ds_name} successfully created or updated")
except HttpResponseError as e:
    print(e)
```

Scan the data source

Scanning a data source can be done in two steps:

1. Create a scan definition
2. Trigger a scan run

In this tutorial, you'll use the default scan rules for Blob Storage containers. However, you can also create custom scan rules programmatically with the Microsoft Purview Scanning Client.

Now let's scan the data source you registered above.

1. Add an import statement to generate unique identifier, call the Microsoft Purview scanning client, the Microsoft Purview administration client, the Azure error response package to be able to troubleshoot, and the client secret credential to gather your Azure credentials.

   Python

   ```
   import uuid
   from azure.purview.scanning import PurviewScanningClient
   from azure.purview.administration.account import PurviewAccountClient
   from azure.core.exceptions import HttpResponseError
   from azure.identity import ClientSecretCredential
   ```

2. Create a scanning client using your credentials:

   Python

   ```
   client_id = "<your client id>"
   client_secret = "<your client secret>"
   tenant_id = "<your tenant id>"

   def get_credentials():
        credentials = ClientSecretCredential(client_id=client_id,
   client_secret=client_secret, tenant_id=tenant_id)
        return credentials

   def get_purview_client():
        credentials = get_credentials()
        client =
   PurviewScanningClient(endpoint=f"https://{reference_name_purview}.scan.purvie
   w.azure.com", credential=credentials, logging_enable=True)
   ```

```
     return client

try:
     client = get_purview_client()
except ValueError as e:
     print(e)
```

3. Add the code to gather the internal name of your collection. (For more information, see the previous section):

Python

```
collection_name = "<name of the collection where you will be creating the
scan>"

def get_admin_client():
     credentials = get_credentials()
     client = PurviewAccountClient(endpoint=purview_endpoint,
credential=credentials, logging_enable=True)
     return client

try:
    admin_client = get_admin_client()
except ValueError as e:
    print(e)

collection_list = client.collections.list_collections()
 for collection in collection_list:
  if collection["friendlyName"].lower() == collection_name.lower():
      collection_name = collection["name"]
```

4. Then, create a scan definition:

Python

```
ds_name = "<name of your registered data source>"
scan_name = "<name of the scan you want to define>"
reference_name_purview = "<name of your Microsoft Purview account>"

body_input = {
        "kind":"AzureStorageMsi",
        "properties": {
            "scanRulesetName": "AzureStorage",
            "scanRulesetType": "System", #We use the default scan rule set
            "collection":
                {
                    "referenceName": collection_name,
                    "type": "CollectionReference"
                }
        }
}
```

```
try:
    response = client.scans.create_or_update(data_source_name=ds_name,
scan_name=scan_name, body=body_input)
    print(response)
    print(f"Scan {scan_name} successfully created or updated")
except HttpResponseError as e:
    print(e)
```

5. Now that the scan is defined you can trigger a scan run with a unique ID:

Python

```
run_id = uuid.uuid4() #unique id of the new scan

try:
    response = client.scan_result.run_scan(data_source_name=ds_name,
scan_name=scan_name, run_id=run_id)
    print(response)
    print(f"Scan {scan_name} successfully started")
except HttpResponseError as e:
    print(e)
```

Full code

Python

```
import uuid
from azure.purview.scanning import PurviewScanningClient
from azure.purview.administration.account import PurviewAccountClient
from azure.identity import ClientSecretCredential

ds_name = "<name of your registered data source>"
scan_name = "<name of the scan you want to define>"
reference_name_purview = "<name of your Microsoft Purview account>"
client_id = "<your client id>"
client_secret = "<your client secret>"
tenant_id = "<your tenant id>"
collection_name = "<name of the collection where you will be creating the scan>"

def get_credentials():
    credentials = ClientSecretCredential(client_id=client_id,
client_secret=client_secret, tenant_id=tenant_id)
    return credentials

def get_purview_client():
    credentials = get_credentials()
    client = PurviewScanningClient(endpoint=purview_scan_endpoint,
credential=credentials, logging_enable=True)
    return client
```

```
def get_admin_client():
    credentials = get_credentials()
    client = PurviewAccountClient(endpoint=purview_endpoint,
credential=credentials, logging_enable=True)
    return client

try:
    admin_client = get_admin_client()
except ValueError as e:
        print(e)

collection_list = admin_client.collections.list_collections()
for collection in collection_list:
    if collection["friendlyName"].lower() == collection_name.lower():
        collection_name = collection["name"]


try:
    client = get_purview_client()
except AzureError as e:
    print(e)

body_input = {
    "kind":"AzureStorageMsi",
    "properties": {
        "scanRulesetName": "AzureStorage",
        "scanRulesetType": "System",
        "collection": {
            "type": "CollectionReference",
            "referenceName": collection_name
        }
    }
}

try:
    response = client.scans.create_or_update(data_source_name=ds_name,
scan_name=scan_name, body=body_input)
    print(response)
    print(f"Scan {scan_name} successfully created or updated")
except HttpResponseError as e:
    print(e)

run_id = uuid.uuid4() #unique id of the new scan

try:
    response = client.scan_result.run_scan(data_source_name=ds_name,
scan_name=scan_name, run_id=run_id)
    print(response)
    print(f"Scan {scan_name} successfully started")
except HttpResponseError as e:
    print(e)
```

Search catalog

Once a scan is complete, it's likely that assets have been discovered and even classified. This process can take some time to complete after a scan, so you may need to wait before running this next portion of code. Wait for your scan to show **completed**, and the assets to appear in Microsoft Purview Unified Catalog.

Once the assets are ready, you can use the Microsoft Purview Catalog client to search the whole catalog.

1. This time you need to import the **catalog** client instead of the scanning one. Also include the HTTPResponse error and ClientSecretCredential.

   Python

   ```
   from azure.purview.catalog import PurviewCatalogClient
   from azure.identity import ClientSecretCredential
   from azure.core.exceptions import HttpResponseError
   ```

2. Create a function to get the credentials to access your Microsoft Purview account, and instantiate the catalog client.

   Python

   ```
   client_id = "<your client id>"
   client_secret = "<your client secret>"
   tenant_id = "<your tenant id>"
   reference_name_purview = "<name of your Microsoft Purview account>"

   def get_credentials():
        credentials = ClientSecretCredential(client_id=client_id,
   client_secret=client_secret, tenant_id=tenant_id)
        return credentials

   def get_catalog_client():
       credentials = get_credentials()
       client =
   PurviewCatalogClient(endpoint=f"https://{reference_name_purview}.scan.purview
   .azure.com", credential=credentials, logging_enable=True)
       return client

   try:
       client_catalog = get_catalog_client()
   except ValueError as e:
       print(e)
   ```

3. Configure your search criteria and keywords in the input body:

Python

```
keywords = "keywords you want to search"

body_input={
    "keywords": keywords
}
```

Here you only specify keywords, but keep in mind you can add many other fields to further specify your query.

4. Search the catalog:

Python

```
try:
    response = client_catalog.discovery.query(search_request=body_input)
    print(response)
except HttpResponseError as e:
    print(e)
```

Full code

Python

```
from azure.purview.catalog import PurviewCatalogClient
from azure.identity import ClientSecretCredential
from azure.core.exceptions import HttpResponseError

client_id = "<your client id>"
client_secret = "<your client secret>"
tenant_id = "<your tenant id>"
reference_name_purview = "<name of your Microsoft Purview account>"
keywords = "<keywords you want to search for>"

def get_credentials():
    credentials = ClientSecretCredential(client_id=client_id,
client_secret=client_secret, tenant_id=tenant_id)
    return credentials

def get_catalog_client():
    credentials = get_credentials()
    client = PurviewCatalogClient(endpoint=purview_endpoint,
credential=credentials, logging_enable=True)
    return client

body_input={
    "keywords": keywords
}
```

```
try:
    catalog_client = get_catalog_client()
except ValueError as e:
    print(e)

try:
    response = catalog_client.discovery.query(search_request=body_input)
    print(response)
except HttpResponseError as e:
    print(e)
```

Delete a data source

In this section, you'll learn how to delete the data source you registered earlier. This operation is fairly simple, and is done with the scanning client.

1. Import the **scanning** client. Also include the HTTPResponse error and ClientSecretCredential.

   Python

   ```
   from azure.purview.scanning import PurviewScanningClient
   from azure.identity import ClientSecretCredential
   from azure.core.exceptions import HttpResponseError
   ```

2. Create a function to get the credentials to access your Microsoft Purview account, and instantiate the scanning client.

   Python

   ```
   client_id = "<your client id>"
   client_secret = "<your client secret>"
   tenant_id = "<your tenant id>"
   reference_name_purview = "<name of your Microsoft Purview account>"

   def get_credentials():
        credentials = ClientSecretCredential(client_id=client_id,
   client_secret=client_secret, tenant_id=tenant_id)
        return credentials

   def get_scanning_client():
       credentials = get_credentials()

   PurviewScanningClient(endpoint=f"https://{reference_name_purview}.scan.purvie
   w.azure.com", credential=credentials, logging_enable=True)
       return client

   try:
       client_scanning = get_scanning_client()
   ```

```
except ValueError as e:
    print(e)
```

3. Delete the data source:

Python

```
    ds_name = "<name of the registered data source you want to delete>"
    try:
        response = client_scanning.data_sources.delete(ds_name)
        print(response)
        print(f"Data source {ds_name} successfully deleted")
    except HttpResponseError as e:
        print(e)
```

Full code

Python

```
from azure.purview.scanning import PurviewScanningClient
from azure.identity import ClientSecretCredential
from azure.core.exceptions import HttpResponseError


client_id = "<your client id>"
client_secret = "<your client secret>"
tenant_id = "<your tenant id>"
reference_name_purview = "<name of your Microsoft Purview account>"
ds_name = "<name of the registered data source you want to delete>"

def get_credentials():
    credentials = ClientSecretCredential(client_id=client_id,
client_secret=client_secret, tenant_id=tenant_id)
    return credentials

def get_scanning_client():
    credentials = get_credentials()
    client =
PurviewScanningClient(endpoint=f"https://{reference_name_purview}.scan.purview.azu
re.com", credential=credentials, logging_enable=True)
    return client

try:
    client_scanning = get_scanning_client()
except ValueError as e:
    print(e)

try:
    response = client_scanning.data_sources.delete(ds_name)
    print(response)
    print(f"Data source {ds_name} successfully deleted")
```

```
except HttpResponseError as e:
    print(e)
```

Microsoft Purview governance services open-source tools and utilities

Article • 02/26/2025

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

This article lists several open-source tools and utilities (command-line, python, and PowerShell interfaces) that help you get started quickly with Microsoft Purview governance services, like Microsoft Purview Data Map, Data Catalog, and Data Estate Insights in a matter of minutes! These tools have been authored & developed by collective effort of the Microsoft Purview Product Group and the open-source community. The objective of such tools is to make learning, starting up, regular usage, and long-term adoption of Microsoft Purview fast and easy.

Intended audience

- Microsoft Purview community including customers, developers, ISVs, partners, evangelists, and enthusiasts.
- The Microsoft Purview Data Catalog is based on Apache Atlas and extends full support for Apache Atlas APIs. We welcome Apache Atlas community, enthusiasts, and developers to wholeheartedly build on and evangelize Microsoft Purview.

Microsoft Purview customer journey stages

- *Microsoft Purview Learners*: Learners who are starting fresh with Microsoft Purview governance services and are keen to understand and explore how a multicloud unified data governance solution works. A section of learners includes users who want to compare and contrast Microsoft Purview with other competing solutions in the data governance market and try it before adopting for long-term usage.
- *Microsoft Purview Innovators*: Innovators who are keen to understand existing and latest features, ideate, and conceptualize features upcoming on Microsoft Purview. They're adept at building and developing solutions for customers, and have

futuristic forward-looking ideas for the next-gen cutting-edge data governance product.

- *Microsoft Purview Enthusiasts/Evangelists*: Enthusiasts who are a combination of Learners and Innovators. They have developed solid understanding and knowledge of Microsoft Purview, hence, are upbeat about adoption of Microsoft Purview. They can help evangelize Microsoft Purview as a service and educate several other Microsoft Purview users and probable customers across the globe.
- *Microsoft Purview Adopters*: Adopters who have migrated from starting up and exploring the Microsoft Purview governance portal and are smoothly using Microsoft Purview for more than a few months.
- *Microsoft Purview Long-Term Regular Users*: Long-term users who have been using the Microsoft Purview governance portal for more than one year and are now confident and comfortable using most advanced use cases on the Azure portal and Microsoft Purview governance portal; furthermore they have near perfect knowledge and awareness of the Microsoft Purview REST APIs and the other use cases supported via Microsoft Purview APIs.

Microsoft Purview open-source tools and utilities list

1. Purview-API-via-PowerShell

    - **Recommended customer journey stages**: *Learners, Innovators, Enthusiasts, Adopters, Long-Term Regular Users*
    - **Description**: This utility is based on and covers the entire set functionality described in the Microsoft Purview REST API reference. Download & Install from PowerShell Gallery. It helps you execute all the documented Microsoft Purview REST APIs through a breezy fast and easy to use PowerShell interface. Use and automate Microsoft Purview APIs for regular and long-term usage via command-line and scripted methods. This is an alternative for customers looking to do bulk tasks in automated manner, batch-mode, or scheduled cron jobs; as against the GUI method of using the Azure portal and Microsoft Purview governance portal. Detailed documentation, sample usage guide, self-help, and examples are available on GitHub:Azure-Purview-API-PowerShell.

2. Microsoft Purview Lab

    - **Recommended customer journey stages**: *Learners, Innovators, Enthusiasts*

- **Description**: A hands-on-lab introducing the myriad features of Microsoft Purview and helping you learn the concepts in a practical and hands-on approach where you execute each step on your own by hand to develop the best possible understanding of Microsoft Purview.

3. Microsoft Purview CLI

   - **Recommended customer journey stages**: *Innovators, Enthusiasts, Adopters, Long-Term Regular Users*
   - **Description**: Python-based tool to execute the Microsoft Purview APIs similar to Purview-API-via-PowerShell but has limited/lesser functionality than the PowerShell-based framework.

4. Microsoft Purview Demo

   - **Recommended customer journey stages**: *Learners, Innovators, Enthusiasts*
   - **Description**: An Azure Resource Manager (ARM) template-based tool to automatically set up and deploy fresh new Microsoft Purview account quickly and securely at the issue of just one command. It's similar to Purview-Starter-Kit, the extra feature being it deploys a few more pre-configured data sources - Azure SQL Database, Azure Data Lake Storage Gen2 Account, Azure Data Factory, Azure Synapse Analytics Workspace

5. PyApacheAtlas: Interface between Microsoft Purview and Apache Atlas using Atlas APIs

   - **Recommended customer journey stages**: *Innovators, Enthusiasts, Adopters, Long-Term Regular Users*
   - **Description**: A Python package to work with Microsoft Purview and Apache Atlas API. Supports bulk loading, custom lineage, and more from a Pythonic set of classes and Excel templates. The package supports programmatic interaction and an Excel template for low-code uploads.

6. Microsoft Purview Event Hubs Notifications Reader

   - **Recommended customer journey stages**: *Innovators, Enthusiasts, Adopters, Long-Term Regular Users*
   - **Description**: This tool demonstrates how to read Microsoft Purview's Event Hubs and catch real-time Kafka notifications from the Event Hubs in Atlas Notifications (https://atlas.apache.org/2.0.0/Notifications.html) format. Further, it generates an excel sheet CSV of the entities and assets on the fly that are discovered live during a scan, and any other notifications of interest that Microsoft Purview generates.

Feedback and disclaimer

None of the tools come with an express warranty from Microsoft verifying their efficacy or guarantees of functionality. They're certified to be free of any malicious activity or viruses, and guaranteed to not collect any private or sensitive data.

For feedback or questions about efficacy and functionality during usage, contact the respective tool owners and authors on the contact details mentioned in the respective GitHub repo.

Next steps

Purview-API-PowerShell

Feedback

Was this page helpful? Yes No

Troubleshoot Microsoft Purview Data Map API request timeout exceptions

Article • 05/08/2025

Applies to Microsoft Purview Data Map API returned HTTP 408 request timeout.

Troubleshooting steps

The following list contains known patterns and solutions for request timeout exceptions.

If it doesn't match your case, or the failure can't be resolved after applied the solution, contact Microsoft Customer Support with all request history.

Time out to list all glossaries

```HTTP
# https://learn.microsoft.com/en-us/rest/api/purview/datamapdataplane/glossary/get-term
GET {endpoint}/datamap/api/atlas/v2/glossary?ignoreTermsAndCategories={ignoreTermsAndCategories}&limit={limit}&offset={offet}&sort={sort}
```

This API by default returns all glossaries together with all terms/categories under those glossaries. If there are 1000 glossaries, each glossary has 1000 terms and 1000 categories, it returns 1000 glossary with 1000 * (1000 terms + 1000 categories).

- The size of all related terms/categories could be too large to fetch via single API within a reasonable duration. Use query parameter `ignoreTermsAndCategories=true` to reduce total count to load in server side and shrink response size for network transfer. If terms/categories are still needed, you can combine with the below APIs to fetch them.
  - List terms under specific glossary

    ```HTTP
    # https://learn.microsoft.com/en-us/rest/api/purview/datamapdataplane/glossary/list-categories
    GET {endpoint}/datamap/api/atlas/v2/glossary/{glossaryGuid}/terms?limit={limit}&offset={offset}&sort={sort}
    ```

  - List categories under specific glossary

    ```HTTP
    ```

```
# https://learn.microsoft.com/en-us/rest/api/purview/datamapdataplane/glossary/list-terms
GET {endpoint}/datamap/api/atlas/v2/glossary/{glossaryGuid}/categories?limit={limit}&offset={offset}&sort={sort}
```

- The size of the all glossary count could be too large to fetch via a single API within a reasonable duration. Use query parameter `limit` and `offset` with small pagination size to get glossaries page by page.

Time out to get glossary with detailed info

HTTP

```
GET {endpoint}/datamap/api/atlas/v2/glossary/{glossaryGuid}/detailed
```

This API by default returns the glossary together with all terms/categories under the glossary.

- The size of all related terms/categories could be too large to fetch via single API within a reasonable duration. We recommend you use the below APIs to have a better pagination experience.
  - Get specific glossary

    HTTP

    ```
    # https://learn.microsoft.com/en-us/rest/api/purview/datamapdataplane/glossary/get
    GET {endpoint}/datamap/api/atlas/v2/glossary/{glossaryGuid}
    ```

  - List terms under specific glossary

    HTTP

    ```
    # https://learn.microsoft.com/en-us/rest/api/purview/datamapdataplane/glossary/list-terms
    GET {endpoint}/datamap/api/atlas/v2/glossary/{glossaryGuid}/terms?limit={limit}&offset={offset}&sort={sort}
    ```

  - List categories under specific glossary

    HTTP

    ```
    # https://learn.microsoft.com/en-us/rest/api/purview/datamapdataplane/glossary/list-categories
    ```

```
GET {endpoint}/datamap/api/atlas/v2/glossary/{glossaryGuid}/categories?
limit={limit}&offset={offset}&sort={sort}
```

Time out to list all terms under specific glossary

HTTP

```
# https://learn.microsoft.com/en-
us/rest/api/purview/datamapdataplane/glossary/list-terms
GET {endpoint}/datamap/api/atlas/v2/glossary/{glossaryGuid}/terms
```

By default, this API returns all terms under the glossary.

- The size of all terms under the glossary could be too large to fetch via single API within a reasonable duration. Use query parameters `limit` and `offset` with small pagination size to list terms page by page.

  HTTP

  ```
  # https://learn.microsoft.com/en-
  us/rest/api/purview/datamapdataplane/glossary/list-terms
  GET {endpoint}/datamap/api/atlas/v2/glossary/{glossaryGuid}/terms?limit=
  {limit}&offset={offset}&sort={sort}
  ```

Time out to get single term

HTTP

```
# https://learn.microsoft.com/en-
us/rest/api/purview/datamapdataplane/glossary/get-term
GET {endpoint}/datamap/api/atlas/v2/glossary/term/{termGuid}
```

This API by default returns the term together with its all relationships. For example, assignedEntities, parent/child terms.

- The size of assignedEntities on the term could be too large to fetch via single API within a reasonable duration. Use query parameter `excludeRelationshipTypes` to ignore assignedEntities relationship called `AtlasGlossarySemanticAssignment` in response. Then use another dedicated API to get assignedEntities with pagination.
  - Get specific term

    HTTP

```
# https://learn.microsoft.com/en-us/rest/api/purview/datamapdataplane/glossary/get-term
GET {endpoint}/datamap/api/atlas/v2/glossary/term/{termGuid}?excludeRelationshipTypes=AtlasGlossarySemanticAssignment
```

- List assignedEntities under specific term

```HTTP
# https://learn.microsoft.com/en-us/rest/api/purview/datamapdataplane/glossary/list-entities-assigned-with-term
GET {endpoint}/datamap/api/atlas/v2/glossary/terms/{termGuid}/assignedEntities?limit={limit}&offset={offset}&sort={sort}
```

Next steps

- Microsoft Purview Data Map REST APIs
- Microsoft Customer Support

Best practices for Microsoft Purview accounts merge scenarios

Article • 05/14/2025

Background

Microsoft Purview represents the next step towards unifying your organization's governance, policy, compliance, risk, and security. If you started using Microsoft Purview data governance with the classic experience, upgrading to the new experience is necessary to obtain the full set of capabilities and enhancements. For organizations with multiple accounts, this upgrade process involves merging existing accounts.

Merging Microsoft Purview accounts involves consolidating resources, configurations, and data governance policies from multiple accounts into a single, unified account. This process requires careful planning and consideration to ensure data integrity, security, and compliance.

It's recommended to follow this guide if you have more than one Microsoft Purview account under the classic experience.

Pre-merge assessment

- Conduct a thorough assessment to obtain the total number of Microsoft Purview accounts in your tenant, and the purpose of each account to identify which account to promote as primary. This account will be converted into your default domain in the new experience.
- Evaluate the remaining accounts to detect and resolve potential conflicts, such as duplicate resources or incompatible configurations.
- Currently, you can only have up to five domains (one default domain and four custom domains) inside the new experience. If your organization has more than five accounts in the classic mode, you need to select only five accounts to upgrade and retire the remaining accounts. For example, if your organization has multiple nonproduction accounts, you can consider merging one account into a custom domain and retire remaining accounts without using the merge tool. This process may include a combination of using the merge tool and manually migrating resources from other classic accounts under a nonprod account.
- Review Microsoft Purview limitations to reduce the risk of failure during the merge and ensure the unified experience has the capacity to merge resources.
- Evaluate resources in each account to determine what needs to be merged or can be excluded during the migration.

- Run the merge tool in assess mode to identify and resolve any issues before the merge, considering that the merge is a one-time irreversible process.

Conflict resolution

- Develop a strategy for resolving conflicts, such as renaming resources or excluding certain types from the merge. Handle name conflicts and consider the option to exclude certain object types.

What to do if conflicts happen

Use the following guidelines if conflicts are raised during the Microsoft Purview account merge process:

1. Identify Conflicts: During the premerge assessment, run the merge tool in assess mode to identify potential conflicts. This generates a list of conflicts that need to be resolved before proceeding with the merge. Conflicts may raise due any of these issues:

    - Conflicts may arise due to duplicate resources, such as assets, or glossary terms. For example, if two accounts have the same custom type definition name, you need to choose one and delete the other.
    - Conflicts could also occur due to incompatible configurations between accounts. This could include differences in data source configurations. For example, Advanced Resource Set feature may be turned on in the merging account but it's off in the primary account.

2. Resolve Conflicts:

    - If there are duplicate resources, consider renaming them to avoid conflicts. For example, you can rename conflicting assets or glossary term to ensure they have unique names across accounts.
        - You may choose to exclude certain types of resources from the merge if they're not essential or if they cause conflicts. This can help streamline the merge process and reduce the risk of errors.
        - For conflicts that can't be resolved automatically, manually delete the conflicting resources before proceeding with the merge. This ensures that only the necessary and nonconflicting resources are merged.

3. Re-Run Assessments:

    - After resolving conflicts, rerun the merge tool in assess mode to ensure that all issues are addressed. This step helps validate that the merge can proceed without

further conflicts.

4. Execute the Merge:

    - Run merge tool in assess mode until all conflicts are resolved.
    - Once all conflicts are resolved, proceed with the merge by following the step-by-step process outlined in the execution phase. This includes running assessments, performing prechecks, and executing the merge.

Perform the merge

- Consider the merge as a one-time process and irreversible.
- Make sure no major changes are happening in data map in the secondary account during the merge process.
- Follow a step-by-step process for the merge, starting with a precheck, running assessments, and then executing the merge.

Post-merge review

- After merge completion, conduct a review to ensure all resources are correctly integrated and that there are no outstanding issues.
- Manually migrate resources that weren't moved.
- Remove classic accounts that are merged.
- Review domains and collections roles to ensure proper role assignments are in place. Remove unnecessary roles that were added due to the merge process or assign the roles to lower-level collections when possible.

Next steps

Merge account to domain

Best practices for Microsoft Purview private endpoints during merging accounts

Article • 05/14/2025

Background

The new Microsoft Purview experience is single, organization-wide instance of Microsoft Purview that's the next step towards unifying your organization's governance, policy, compliance, risk, and security. It's a prerequisite to a single platform used to manage and govern any data, both structured and unstructured, across your data estate. Including Azure, Microsoft 365, Azure, on-premises, or multicloud and SaaS applications in future. For organizations with multiple accounts, this upgrade process involves merging existing accounts.

When merging Microsoft Purview accounts, it's important to consider the private endpoints. Private endpoints provide a secure and private connection to the Purview governance portal and other resources, ensuring that only users within your network can access them. During the merge process, resources, configurations, and data governance policies from multiple accounts are consolidated into a single, unified account. If one or more of your classic accounts have private endpoint connections, this requires careful planning to ensure that private endpoints are correctly configured and that all necessary connections are maintained. Failure to properly manage private endpoints during the merge can result in issues and potential failure in merge process.

This article contains information to guide you resolve conflict when merging accounts with private endpoints.

Private endpoints in Microsoft Purview accounts

Private endpoints in classic mode include the following capabilities:

- **Account private endpoints**: Allowed only client calls originating from within the private network to access the Microsoft Purview account.
- **Portal private endpoints**: Enabled secure access to the Microsoft Purview governance portal using private network connectivity.
- **Ingestion private endpoints**: Used to scan Azure IaaS and PaaS data sources inside Azure virtual networks and on-premises data sources through a private connection, to provide network isolation for metadata, flowing from data sources to the Microsoft Purview Data Map.

Private endpoints in unified experience include:

- **Platform private endpoints**: The Platform private endpoint is a new private endpoint concept in the unified portal that provides you organizational-level account private endpoint connection.
- **Ingestion private endpoints**: Used to scan Azure IaaS and PaaS data sources inside Azure virtual networks and on-premises data sources through a private connection.

Private endpoints conflict resolution during account merge

After you upgrade your primary account into unified experience, you may need to plan for account merge for the remaining classic accounts in your tenant. It's recommended to follow these guidelines to avoid conflicts during the merge process:

- Identify classic accounts that have any private endpoints, including accounts with Managed Vnet IRs. Remove unnecessary private endpoints, Managed VNet IR's or self-hosted integration runtimes from secondary accounts.
- If an account private endpoint is identified as a conflict and blocking the merge process, it's recommended to delete all private endpoints or filter them out during the merge process. The filtered configurations will not be merged into the unified experience.
- If there's a conflict in an ingestion private endpoint, the ingestion private endpoint in the secondary account can be abandoned. It's recommended to use the ingestion private endpoint in the primary account instead. If ingestion private endpoint is needed for a different VNet, you can create more ingestion private endpoints in the primary account. Create a new self-hosted integration runtime for full VNet isolation if needed.
- Portal private endpoint should be filtered out during the merge and be deleted in the secondary account after the merge, as portal private endpoint will be deprecated.

Conflicts on integration runtime

- If you're using Managed VNet IR and self-hosted integration runtimes in both accounts, it's recommended to delete Managed VNet IR and self-hosted integration runtime in the secondary account before merge. Create new Managed VNet IR or self-hosted integration runtime in the primary account after the merge. Alternatively, filtering out Managed VNET IR and self-hosted integration runtime is recommended. Scans using these will be filtered out and will need to be recreated after the merge.
- If the plan is to use the same self-hosted integration runtime virtual machines, uninstall self-hosted integration runtime in the secondary account before merge and install self-hosted integration runtime in the primary account after merging the secondary account. It's possible to reuse the same VMs and machines or use new ones.

Conflicts on managed private endpoint connections:

- Ensure Managed private endpoint connections to Platform private endpoint are properly set up in the primary account.
- Delete conflicting Managed VNet private endpoints from secondary accounts. After the merge, recreate or reuse existing Managed VNet IR in the primary account for connections to the sources that got merged.

Conflicting Azure Data Factory or ADS Lineage Connections

- It's recommended to filter out Azure Data Factory and ADS lineage connections during the merge. Work with the primary account's domain source admins to reconnect it.

Next steps

Merge account to domain

Microsoft Purview automation best practices

Article • 05/14/2025

While Microsoft Purview governance solutions provide an out of the box user experience with the Microsoft Purview governance portal, not all tasks are suited to the point-and-click nature of the graphical user experience.

For example:

- Triggering a scan to run as part of an automated process.
- Monitoring for metadata changes in real time.
- Building your own custom user experience.

Microsoft Purview provides several tools in which we can use to interact with the underlying platform, in an automated, and programmatic fashion. Because of the open nature of the Microsoft Purview service, we can automate different aspects, from the control plane, made accessible via Azure Resource Manager, to Microsoft Purview's multiple data planes (catalog, scanning, administration, and more).

This article provides a summary of the options available, and guidance on what to use when.

Tools

Expand table

Tool Type	Tool	Scenario	Management	Catalog	Scanning	Logs
Resource Management	• ARM Templates • Terraform	Infrastructure as Code	✓			
Command Line	• Azure CLI	Interactive	✓			
Command Line	• Azure PowerShell	Interactive	✓	✓		
API	• REST API	On-Demand	✓	✓	✓	
Streaming (Apache Atlas)	• Event Hubs	Real-Time		✓		

Tool Type	Tool	Scenario	Management	Catalog	Scanning	Logs
Monitoring	• Azure Monitor	Monitoring				✓
SDK	• .NET • Java • JavaScript • Python	Custom Development	✓	✓	✓	

Command Line

Azure CLI and Azure PowerShell are command-line tools that enable you to manage Azure resources such as Microsoft Purview. While the list of commands will grow over time, only a subset of Microsoft Purview control plane operations is currently available. For an up-to-date list of commands currently available, check out the documentation (Azure CLI | Azure PowerShell).

- **Azure CLI** - A cross-platform tool that allows the execution of commands through a terminal using interactive command-line prompts or a script. Azure CLI has a **purview extension** that allows for the management of Microsoft Purview accounts. For example, `az purview account`.
- **Azure PowerShell** - A cross-platform task automation program, consisting of a set of cmdlets for managing Azure resources. Azure PowerShell has a module called **Az.Purview** that allows for the management of Microsoft Purview accounts. For example, `Get-AzPurviewAccount`.

When to use?

- Best suited for ad-hoc tasks and quick exploratory operations.

API

REST APIs are HTTP endpoints that surface different methods (`POST`, `GET`, `PUT`, `DELETE`), triggering actions such as create, read, update, or delete (CRUD). Microsoft Purview exposes a large portion of the Microsoft Purview platform via multiple service endpoints.

When to use?

- Required operations not available via Azure CLI, Azure PowerShell, or native client libraries.

- Custom application development or process automation.

Streaming (Apache Atlas)

Each Microsoft Purview account can configure Event Hubs that are accessible via their Atlas Kafka endpoint.

You can follow these steps to configure the Event Hubs namespaces.

> **Note**
>
> Enabling this Event Hubs namespace does incur a cost for the namespace. For specific details, see the pricing page.

Once the namespace is enabled, Microsoft Purview events can be monitored by consuming messages from the event hub. External systems can also use the event hub to publish events to Microsoft Purview as they occur.

- **Consume Events** - Microsoft Purview will send notifications about metadata changes to Kafka topic **ATLAS_ENTITIES**. Applications interested in metadata changes can monitor for these notifications. Supported operations include: `ENTITY_CREATE`, `ENTITY_UPDATE`, `ENTITY_DELETE`, `CLASSIFICATION_ADD`, `CLASSIFICATION_UPDATE`, `CLASSIFICATION_DELETE`.
- **Publish Events** - Microsoft Purview can be notified of metadata changes via notifications to Kafka topic **ATLAS_HOOK**. Supported operations include: `ENTITY_CREATE_V2`, `ENTITY_PARTIAL_UPDATE_V2`, `ENTITY_FULL_UPDATE_V2`, `ENTITY_DELETE_V2`.

When to use?

- Applications or processes that need to publish or consume Apache Atlas events in real time.

Monitoring

Microsoft Purview can send platform logs and metrics via "Diagnostic settings" to one or more destinations (Log Analytics Workspace, Storage Account, or Azure Event Hubs). Available metrics include `Data Map Capacity Units`, `Data Map Storage Size`, `Scan Canceled`, `Scan Completed`, `Scan Failed`, and `Scan Time Taken`.

Once configured, Microsoft Purview automatically sends these events to the destination as a JSON payload. From there, application subscribers that need to consume and act on these events can do so with the option of orchestrating downstream logic.

When to use?

- Applications or processes that need to consume diagnostic events.

SDK

Microsoft provides Azure SDKs to programmatically manage and interact with Azure services. Microsoft Purview client libraries are available in several languages (.NET, Java, JavaScript, and Python), designed to be consistent, approachable, and idiomatic.

- .NET
- Java
- JavaScript
- Python

When to use?

- Recommended over the REST API as the native client libraries (where available) will follow standard programming language conventions in line with the target language that will feel natural to the developer.

Next steps

- Microsoft Purview REST API

Microsoft Purview disaster recovery and migration best practices

Article • 05/14/2025

This article provides guidance on backup and recovery strategy when your organization has Microsoft Purview governance solutions in production deployment. You can also use this general guideline to implement account migration. The scope of this article is to cover manual business continuity and disaster recovery (BCDR) methods, where you could automate using APIs.

Azure data center outages are rare, but can last anywhere from a few minutes to hours. Data Center outages can cause disruption to environments that are being relied on for data governance. By following the steps detailed in this article, you can continue to govern your data in the event of a data center outage for the primary region of your Microsoft Purview account.

> **Tip**
>
> Find more information about reliability for Microsoft Purview.

Achieve business continuity for Microsoft Purview

BCDR in a Microsoft Purview instance refers to the mechanisms, policies, and procedures that enable your business to protect data loss and continue operating in the face of disruption, particularly to its scanning, classic catalog, and classic insights tiers. This page explains how to configure a disaster recovery environment for Microsoft Purview.

Today, Microsoft Purview doesn't support automated BCDR. Until that support is added, you're responsible to take care of backup and restore activities. You can manually create a secondary Microsoft Purview account as a warm standby instance in another region.

The following steps summarize how you can achieve disaster recovery manually:

1. Once the primary Microsoft Purview account is created, create one or more secondary Microsoft Purview accounts in a separate region.

   > **Important**
   >
   > Microsoft Purview currently supports a single Microsoft Purview instance per tenant. To create a second account for backup and disaster recovery, contact support

2. All activities performed on the primary Microsoft Purview account must be carried out on the secondary Microsoft Purview accounts as well. This includes:

    - Maintain Account information
    - Create and maintain custom scan rule sets, classifications, and classification rules
    - Register and scan sources
    - Create and maintain collections along with the association of sources with the collections
    - Create and maintain credentials used while scanning
    - Curate data assets
    - Create and maintain glossary terms

Specific steps to create and maintain a disaster recovery account are provided later in the article. Before you follow them, read through the limitations and considerations.

Limitations and considerations

As you create your manual BCDR plan, keep the following points in mind:

- You'll be charged for primary and secondary Microsoft Purview instances.
- The primary and secondary Microsoft Purview accounts can't be configured to the same Azure Data Factory, Azure Data Share and Synapse Analytics accounts, if applicable. As a result, the lineage from Azure Data Factory and Azure Data Share can't be seen in the secondary Microsoft Purview accounts. This limitation will be addressed when automated BCDR is supported.
- The integration runtimes are specific to a Microsoft Purview account. So, scans need to run in primary and secondary Microsoft Purview accounts in-parallel, multiple self-hosted integration runtimes must be maintained. This limitation will also be addressed when automated BCDR is supported.
- Parallel execution of scans from both primary and secondary Microsoft Purview accounts on the same source can affect the performance of the source. This can result in scan durations to vary across the Microsoft Purview accounts.
- It isn't advisable to back up "scanned" assets' details. You should only back up the curated data such as mapping of classifications and glossaries on assets. The only case when you need to back up assets' details is when you have custom assets via custom typeDef.
- The backed-up asset count should be fewer than 100,000 assets. The main driver is that you have to use the search query API to get the assets, which have limitation of 100,000

assets returned. However, if you're able to segment the search query to get smaller number of assets per API call, it's possible to back up more than 100,000 assets.

- If you want to continuously "sync" assets between two accounts, there are other steps that won't be covered in detail in this article. You have to use Microsoft Purview's Event Hubs to subscribe and create entities to another account. However, Event Hubs only has Atlas information. Microsoft Purview has added other capabilities such as **glossaries** and **contacts** that won't be available via Event Hubs.

Steps to achieve business continuity

Create the new account

- **If your organization already has multiple Microsoft Purview accounts**, you can create a new Microsoft Purview account by following this guide instruction: Quickstart: Create a Microsoft Purview account in the Azure portal
- **If your organization only has one tenant/organizational Microsoft Purview account**, to create an account for backup and recovery contact support.

Plan these configuration items that you can't change later:

- Account name
- Region
- Subscription
- Manage resource group name

Migrate configuration items

Below steps are referring to Microsoft Purview API documentation so that you can programmatically stand up the backup account quickly:

Expand table

Task	Description
Account information	Maintain Account information by granting access for the admin and/or service principal to the account at root level
Collections	Create and maintain Collections along with the association of sources with the Collections. You can call List Collections API and then get specific details of each collection via Get Collection API

Task	Description
Scan rule set	Create and maintain custom scan rule sets. You need to call List custom scan rule sets API and get details by calling Get scan rule set API
Manual classifications	Get a list of all manual classifications by calling get classifications APIs and get details of each classification
Resource set rule	Create and maintain resource set rule. You can call the Get resource set rule API to get the rule details
Data sources	Call the Get all data sources API to list data sources with details. You also have to get the triggers by calling Get trigger API. There's also Create data sources API if you need to re-create the sources in bulk in the new account.
Credentials	Create and maintain credentials used while scanning. There's no API to extract credentials, so this must be redone in the new account.
Self-hosted integration runtime (SHIR)	Get a list of SHIR and get updated keys from the new account then update the SHIRs. This must be done manually inside the SHIRs' hosts. These need to be running before you create scans.
ADF connections	Currently an ADF can be connected to one Microsoft Purview at a time. You must disconnect ADF from failed Microsoft Purview account and reconnect it to the new account later.

Run scans

> **Important**
>
> Make sure your self-hosted integration runtimes have been configured and are running and available before creating scans.

This will populate all assets with default `typedef`. There are several reasons to run the scans again vs. exporting the existing assets and importing to the new account:

- There's a limit of 100,000 assets returned from the search query to export assets.
- It's cumbersome to export assets with relationships.
- When you rerun the scans, you'll get all relationships and assets details up to date.
- Microsoft Purview comes out with new features regularly so you can benefit from other features when running new scans.

Running the scans is the most effective way to get all assets of data sources that Microsoft Purview is already supporting.

Migrate custom typedefs and custom assets

If your organization has created custom types in Microsoft Purview, you need to migrate those manually.

Custom typedefs

To identify all custom `typedef`, you can use the get all type definitions API. This will return each type. You need to identify the custom types in such format as `"serviceType": "<custom_typedef>"`

Custom assets

To export custom assets, you can search those custom assets and pass the proper custom `typedef` via the discovery API

> **Note**
>
> There is a 100,000 return limit per search result. You might have to break the search query so that it won't return more than 100,000 records.

There are several ways to scope down the search query to get a subset of assets:

- **Using `Keyword`**: Pass the parent FQN such as `Keyword: "<Parent String>/*"`
- **Using `Filter`**: Include `assetType` with the specific custom `typedef` in your search such as `"assetType": "<custom_typedef>"`

Here's an example of a search payload by customizing the `keywords` so that only assets in specific storage account (`exampleaccount`) are returned:

JSON

```
{
  "keywords": "adl://exampleaccount.azuredatalakestore.net/*",
  "filter": {
    "and": [
      {
        "not": {
          "or": [
            {
              "attributeName": "size",
              "operator": "eq",
              "attributeValue": 0
            },
```

```
            {
              "attributeName": "fileSize",
              "operator": "eq",
              "attributeValue": 0
            }
          ]
        }
      },
      {
        "not": {
          "classification": "MICROSOFT.SYSTEM.TEMP_FILE"
        }
      },
      {
        "not": {
          "or": [
            {
              "entityType": "AtlasGlossaryTerm"
            },
            {
              "entityType": "AtlasGlossary"
            }
          ]
        }
      }
    ]
  },
  "limit": 10,
  "offset": 0,
  "facets": [
    {
      "facet": "assetType",
      "count": 0,
      "sort": {
        "count": "desc"
      }
    },
    {
      "facet": "classification",
      "count": 10,
      "sort": {
        "count": "desc"
      }
    },
    {
      "facet": "contactId",
      "count": 10,
      "sort": {
        "count": "desc"
      }
    },
    {
      "facet": "label",
      "count": 10,
      "sort": {
```

```
        "count": "desc"
      }
    },
    {
      "facet": "term",
      "count": 10,
      "sort": {
        "count": "desc"
      }
    }
  ]
}
```

The returned assets will have some key/pair value that you can extract details:

JSON

```
{
    "referredEntities": {},
    "entity": {
    "typeName": "column",
    "attributes": {
        "owner": null,
        "qualifiedName":
"adl://exampleaccount.azuredatalakestore.net/123/1/DP_TFS/CBT/Extensions/DTTP.targ
ets#:xml/Project/Target/XmlPeek/@XmlInputPath",
        "name": "~XmlInputPath",
        "description": null,
        "type": "string"
    },
    "guid": "5cf8a9e5-c9fd-abe0-2e8c-d40024263dcb",
    "status": "ACTIVE",
    "createdBy": "ExampleCreator",
    "updatedBy": "ExampleUpdator",
    "createTime": 1553072455110,
    "updateTime": 1553072455110,
    "version": 0,
    "relationshipAttributes": {
        "schema": [],
        "inputToProcesses": [],
        "composeSchema": {
        "guid": "cc6652ae-dc6d-90c9-1899-252eabc0e929",
        "typeName": "tabular_schema",
        "displayText": "tabular_schema",
        "relationshipGuid": "5a4510d4-57d0-467c-888f-4b61df42702b",
        "relationshipStatus": "ACTIVE",
        "relationshipAttributes": {
            "typeName": "tabular_schema_columns"
        }
        },
        "meanings": [],
        "outputFromProcesses": [],
        "tabular_schema": null
```

```
    },
    "classifications": [
        {
        "typeName": "MICROSOFT.PERSONAL.EMAIL",
        "lastModifiedTS": "1",
        "entityGuid": "f6095442-f289-44cf-ae56-47f6f6f6000c",
        "entityStatus": "ACTIVE"
        }
    ],
    "contacts": {
        "Expert": [
        {
            "id": "30435ff9-9b96-44af-a5a9-e05c8b1ae2df",
            "info": "Example Expert Info"
        }
        ],
        "Owner": [
        {
            "id": "30435ff9-9b96-44af-a5a9-e05c8b1ae2df",
            "info": "Example Owner Info"
        }
        ]
    }
    }
}
```

> ⓘ Note
>
> You need to migrate the term templates from `typedef` output as well.

When you re-create the custom entities, you might need to prepare the payload prior to sending to the API:

> ⓘ Note
>
> The initial goal is to migrate all entities without any relationships or mappings. This will avoid potential errors.

- All `timestamp` value must be null such as `updateTime`, `updateTime`, and `lastModifiedTS`.
- The `guid` can't be regenerated exactly as before so you have to pass in a negative integer such as "-5000" to avoid error.
- The content of `relationshipAttributes` shouldn't be a part of the payload to avoid errors since it's possible that the `guids` aren't the same or haven't been created yet. You have to turn `relationshipAttributes` into an empty array prior to submitting the payload.

    - `meanings` contains all glossary mappings, which will be updated in bulk after the entities are created.

- Similarly, `classifications` needs to be an empty array as well when you submit the payload to create entities since you have to create classification mapping to bulk entities later using a different API.

Migrate relationships

To complete the asset migration, you must remap the relationships. There are three tasks:

1. Call the relationship API to get relationship information between entities by its `guid`
2. Prepare the relationship payload so that there's no hard reference to old `guids` in the old Microsoft Purview accounts. You need to update those `guids` to the new account's `guids`.
3. Finally, create a new relationship between entities

Migrate glossary terms

> **Note**
>
> Before migrating terms, you need to migrate the term templates. This step should be already covered in the custom `typedef` migration.

Using the Microsoft Purview governance portal

The quickest way to migrate glossary terms is to export terms to a .csv file. You can do this using the Microsoft Purview governance portal.

Using Microsoft Purview API

To automate glossary migration, you first need to get the glossary `guid` (`glossaryGuid`) via List Glossaries API. The `glossaryGuid` is the top/root level glossary `guid`.

The below sample response will provide the `guid` to use for subsequent API calls:

JSON

```
"guid": "c018ddaf-7c21-4b37-a838-dae5f110c3d8"
```

Once you have the `glossaryGuid`, you can start to migrate the terms via two steps:

1. Export Glossary Terms As .csv
2. Import Glossary Terms Via .csv

Assign classifications to assets

> **Note**
>
> The prerequisite for this step is to have all classifications available in the new account from Migrate configuration items step.

You must call the discovery API to get the classification assignments to assets. This is applicable to all assets. If you've migrated the custom assets, the information about classification assignments is already available in `classifications` property. Another way to get classifications is to list classification per guid in the old account.

To assign classifications to assets, you need to associate a classification to multiple entities in bulk via the API.

Assign contacts to assets

If you have extracted asset information from previous steps, the contact details are available from the discovery API.

To assign contacts to assets, you need a list of `guids` and identify all `objectid` of the contacts. You can automate this process by iterating through all assets and reassign contacts to all assets using the Create Or Update Entities API.

Related information

- Microsoft Purview reliability documentation
- Business Continuity and Disaster Recovery
- Build high availability into your BCDR strategy
- Azure status

Classification best practices in the Microsoft Purview Data Map

Article • 05/14/2025

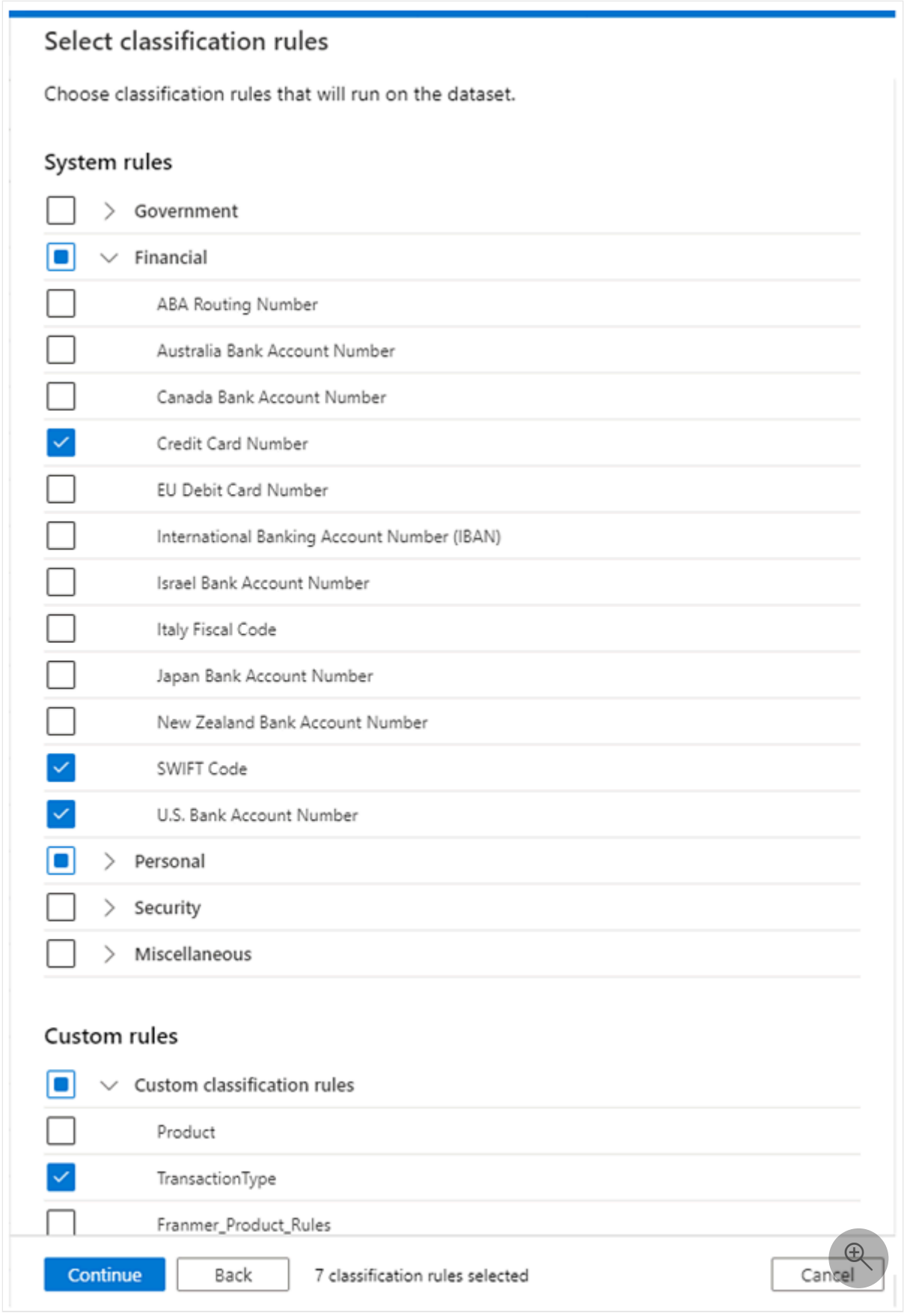
Data classification in the Microsoft Purview Data Map is a way of categorizing data assets by assigning unique logical labels or classes to the data assets. Classification is based on the business context of the data. For example, you might classify assets by *Passport Number*, *Driver's License Number*, *Credit Card Number*, *SWIFT Code*, *Person's Name*, and so on. Learn more about data classification in Data Map.

This article describes best practices to adopt when you're classifying data assets, so that your scans will be more effective and you have the most complete information possible about your entire data estate.

Scan rule set

By using a *scan rule set*, you can configure the relevant classifications that should be applied to the particular scan for the data source. Select the relevant system classifications, or select custom classifications if you've created one for the data you're scanning.

For example, in the following image, only the specific selected system and custom classifications will be applied for the data source you're scanning (for example, financial data).

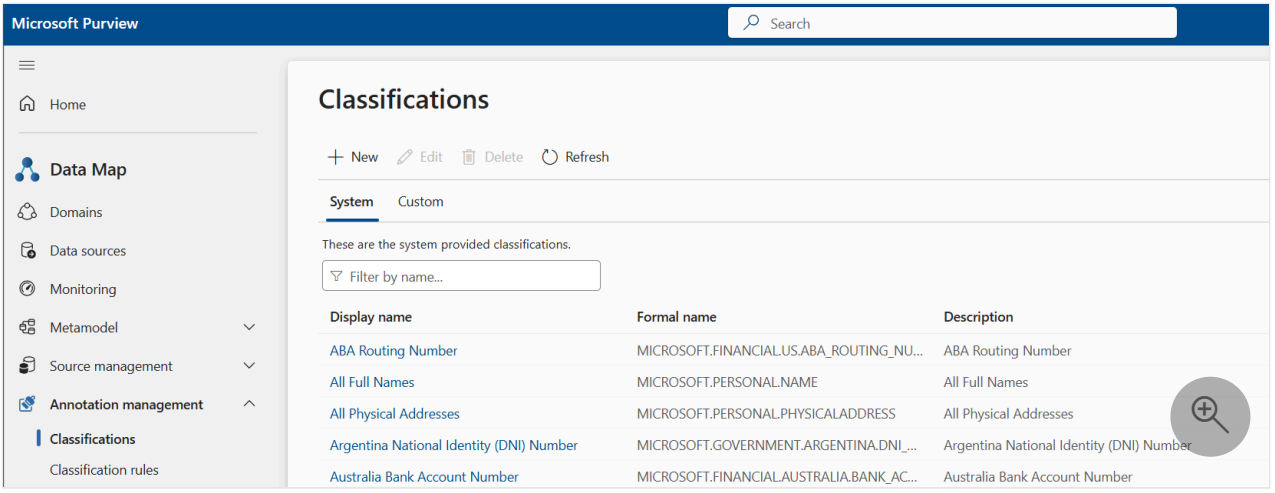


Annotation management

While you're deciding on which classifications to apply, we recommend that you:

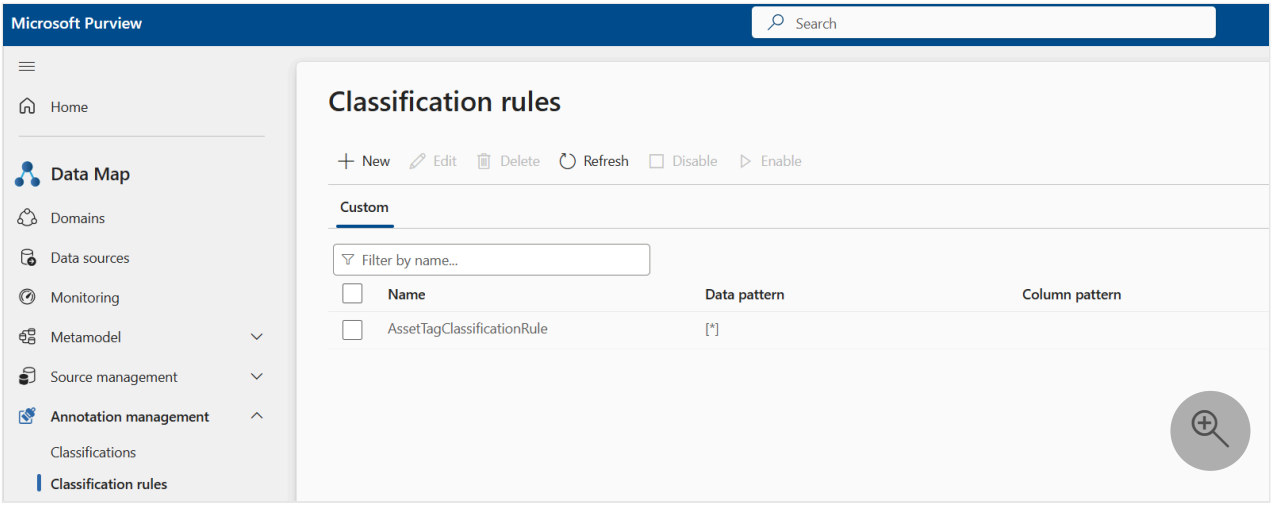
- Go to **Data Map** > **Annotation management** > **Classifications** pane.

- Review the available system classifications to be applied on the data assets you're scanning. The formal names of system classifications have a *MICROSOFT* prefix.



- Create a custom classification, if necessary. Select the **Custom** tab, and then select **+ New**. Learn how to create a custom classification.

- Create the classification rule for the custom classification that you created in the preceding step. Go to **Data Map** > **Annotation management** > **Classification rules**. Here, you can create the classification rule for the custom classification name that you created in the preceding step.



Custom classifications

Create custom classifications only if the available system classifications don't meet your needs.

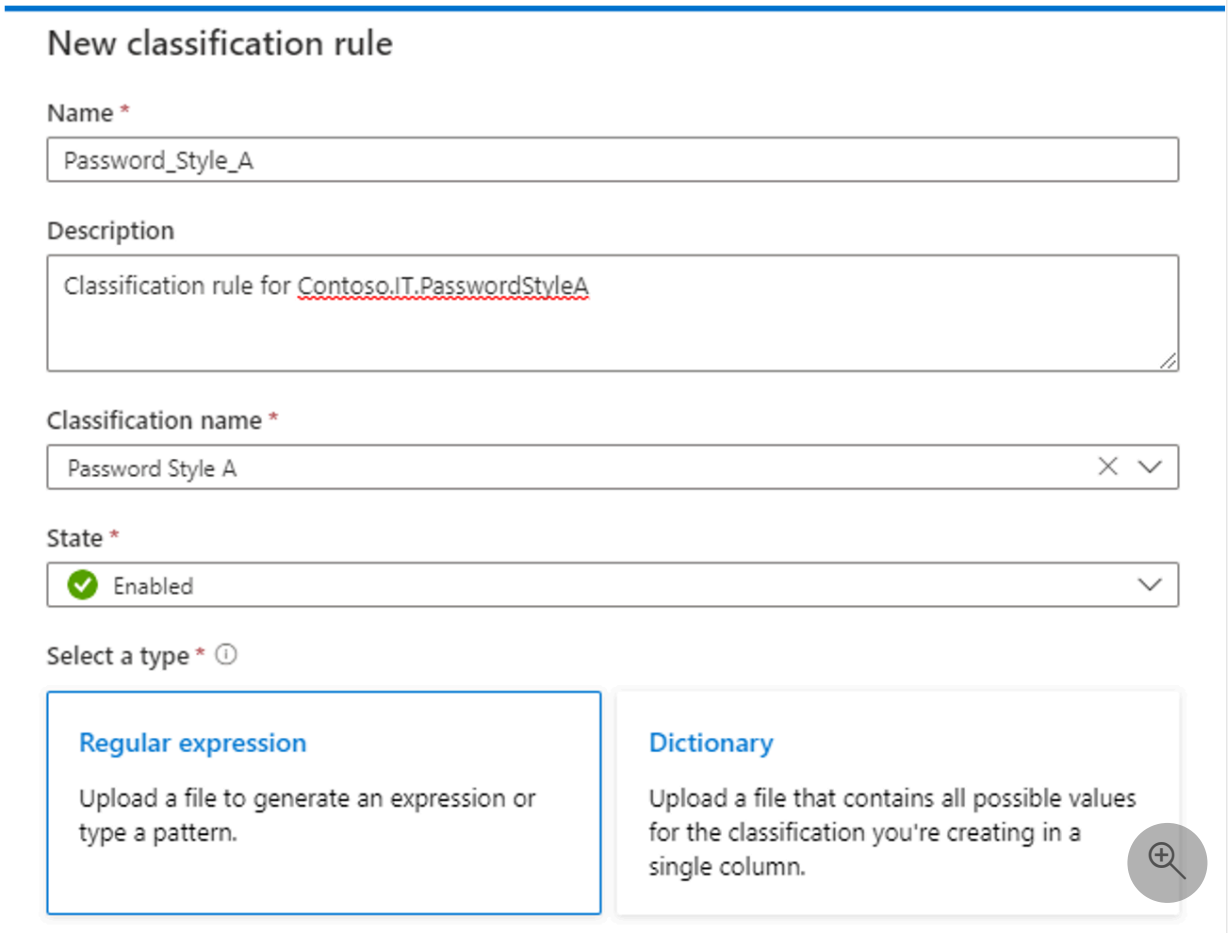
For the *name* of the custom classification, it's a good practice to use a namespace convention (for example, <*company name*>.<*business unit*>.<*custom classification name*>).

As an example, for the custom EMPLOYEE_ID classification for fictitious company Contoso, the name of your custom classification would be CONTOSO.HR.EMPLOYEE_ID, and the friendly name is stored in the system as HR.EMPLOYEE ID.



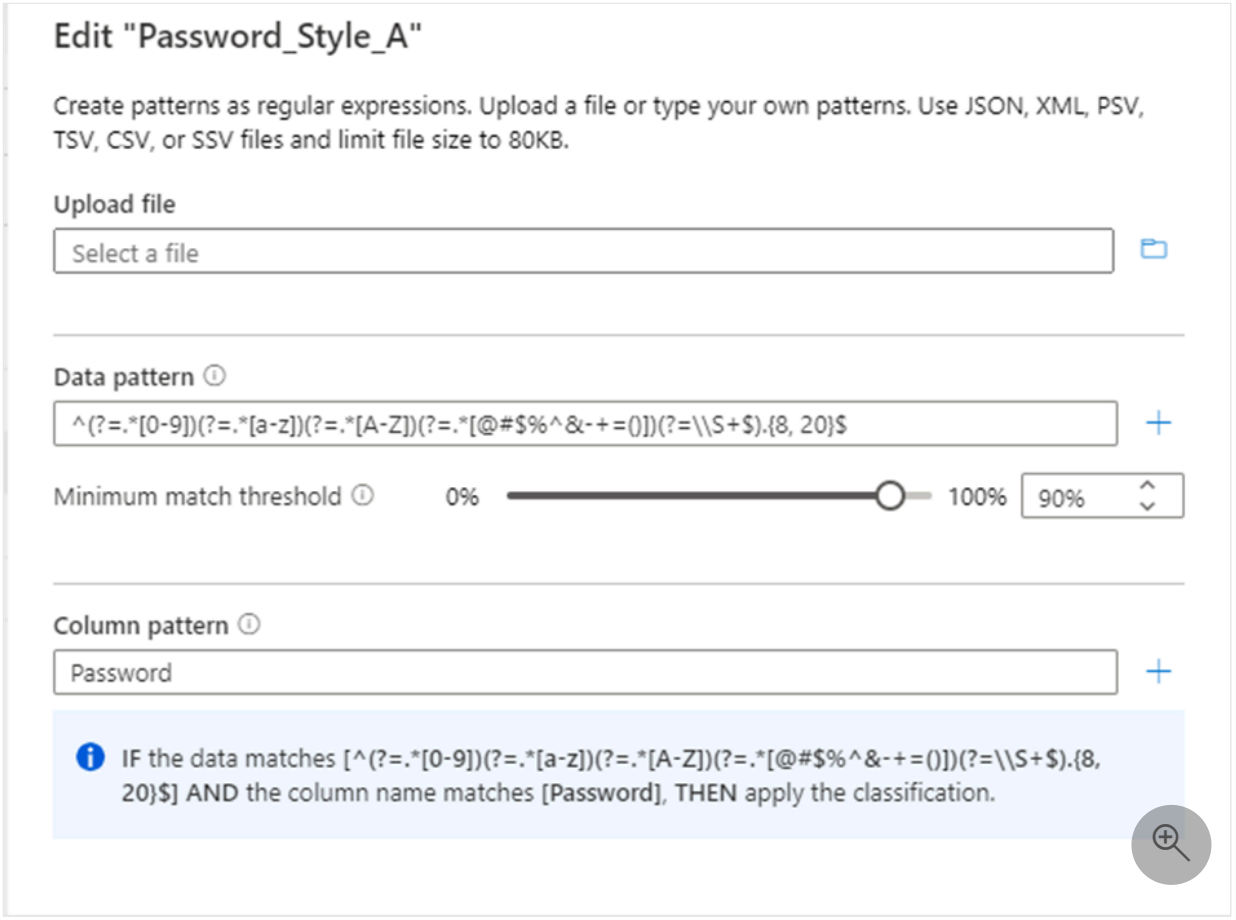
When you create and configure the classification rules for a custom classification, do the following:

- Select the appropriate classification name for which the classification rule is to be created.
- The Microsoft Purview Data Map supports the following two methods for creating custom classification rules:
    - Use the **Regular expression** (regex) method if you can consistently express the data element by using a regular expression pattern or you can generate the pattern by using a data file. Ensure that the sample data reflects the population.
    - Use the **Dictionary** method only if the list of values in the dictionary file represents all possible values of data to be classified and is expected to conform to a given set of data (considering future values as well).

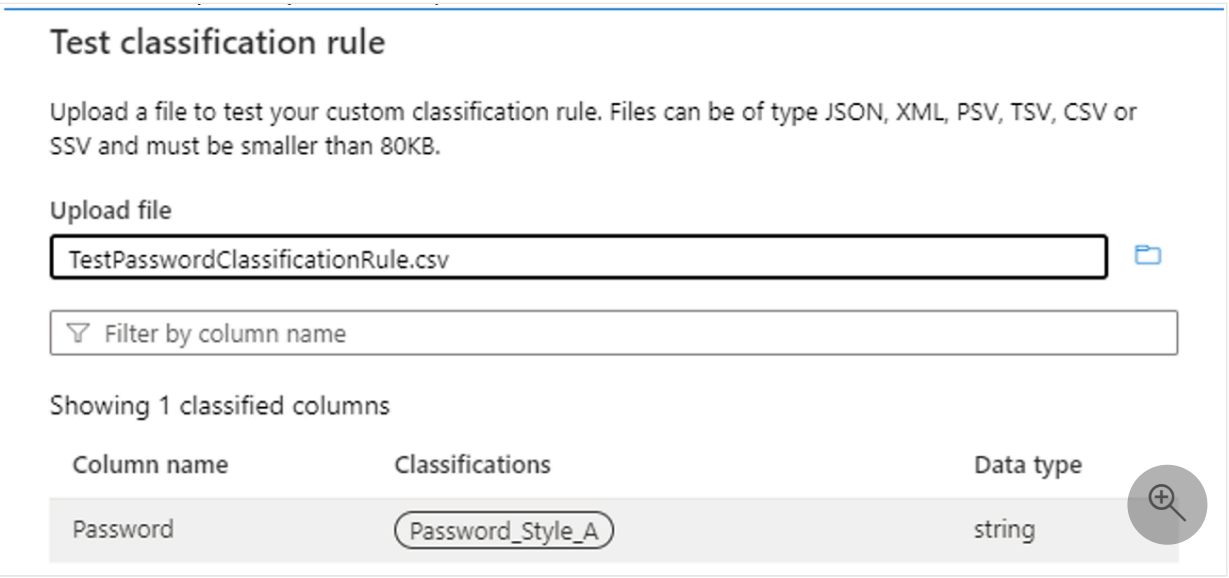


- Using the **Regular expression** method:
    - Configure the regex pattern for the data to be classified. Ensure that the regex pattern is generic enough to cater to the data being classified.

- Microsoft Purview also provides a feature to generate a suggested regex pattern. After you upload a sample data file, select one of the suggested patterns, and then select **Add to patterns** to use the suggested data and column patterns. You can modify the suggested patterns, or you can type your own patterns without having to upload a file.
- You can also configure the column name pattern, for the column to be classified to minimize false positives.
- Configure the *Minimum match threshold* parameter that's acceptable for your data that matches the data pattern to apply the classification. The threshold values can be from 1% through 100%. We suggest a value of at least 60% as the threshold to avoid false positives. However, you could configure as necessary for your specific classification scenarios. For example, your threshold might be as low as 1% if you want to detect and apply a classification for any value in the data if it matches the pattern.



- The option to set a minimum match rule is automatically disabled if more than one data pattern is added to the classification rule.
- Use the *Test classification rule* and test with sample data to verify that the classification rule is working as expected. Ensure that in the sample data (for example, in a .csv file) at least three columns are present, including the column on which the classification is to be applied. If the test is successful, you should see the classification label on the column, as shown in the following image:



- Using the **Dictionary** method:
  - You can use the Dictionary method to fit enumeration data or if the dictionary list of possible values is available.
  - This method supports .csv and .tsv files, with a file size limit of 30 megabytes (MB).

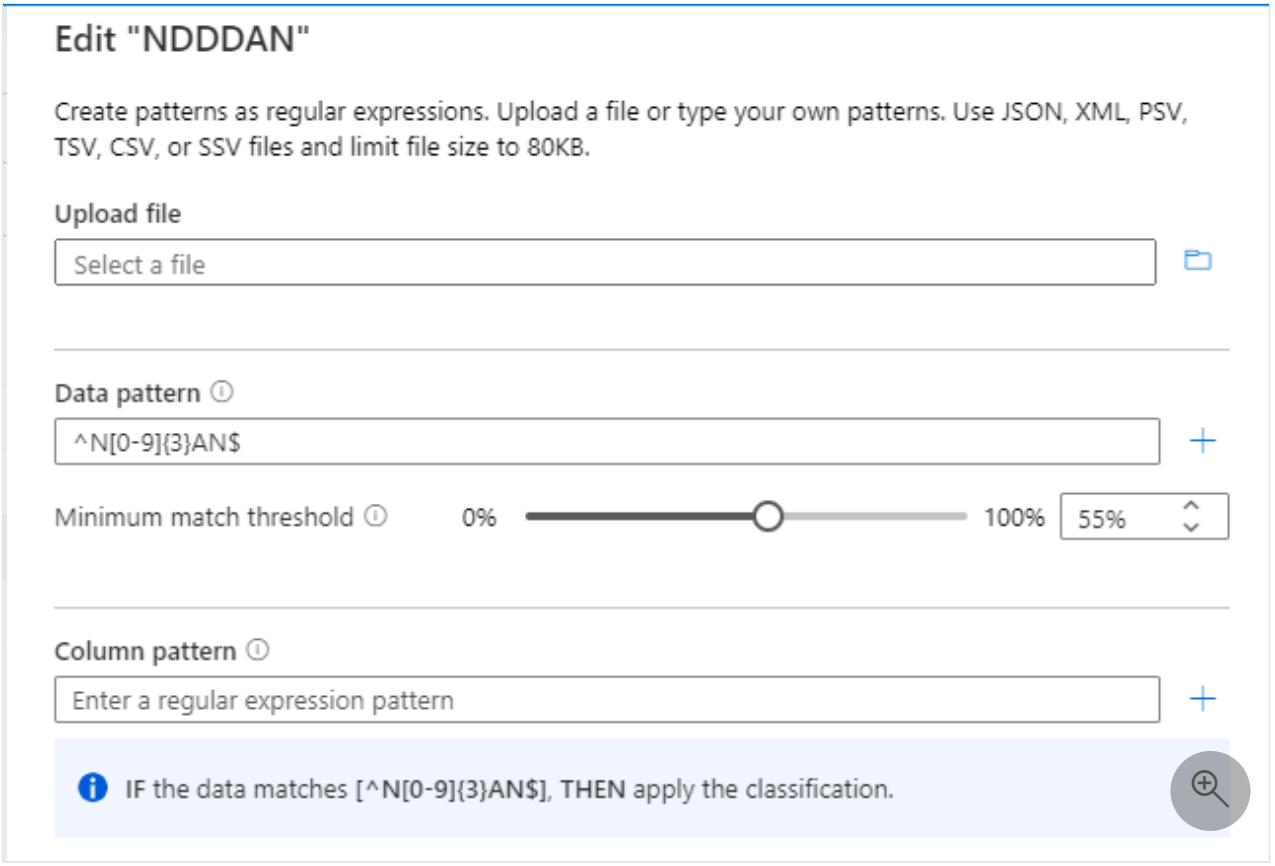
Custom classification archetypes

How the "threshold" parameter works in the regular expression

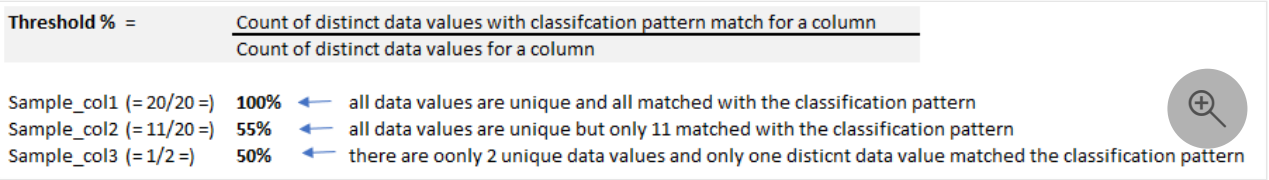
- Consider the sample source data in the following image. There are five columns, and the custom classification rule should be applied to columns **Sample_col1**, **Sample_col2**, and **Sample_col3** for the data pattern *N{Digit}{Digit}{Digit}AN*.

	A	B	C	D	E
1	ID	Sample_col1	Sample_col2	Sample_col3	Description
2	1	N280AN	N280AN	N280AN	This is test data
3	2	N857AN	N857AN	N280AN	This is test data
4	3	N471AN	N471AN	N280AN	This is test data
5	4	N670AN	N670AN	N280AN	This is test data
6	5	N240AN	N240AN	N280AN	This is test data
7	6	N809AN	N809AN	N280AN	This is test data
8	7	N638AN	N638AN	N280AN	This is test data
9	8	N519AN	N519AN	N280AN	This is test data
10	9	N499AN	N499AN	N280AN	This is test data
11	10	N853AN	N853AN	N280AN	This is test data
12	11	N565AN	N565AN	N280AN	This is test data
13	12	N535AN	XYZ123	XYZ123	This is test data
14	13	N825AN	XYZ124	XYZ123	This is test data
15	14	N891AN	XYZ125	XYZ123	This is test data
16	15	N239AN	XYZ126	XYZ123	This is test data
17	16	N889AN	XYZ127	XYZ123	This is test data
18	17	N823AN	XYZ128	XYZ123	This is test data
19	18	N249AN	XYZ129	XYZ123	This is test data
20	19	N637AN	XYZ130	XYZ123	This is test data
21	20	N999AN	XYZ131	XYZ123	This is test data

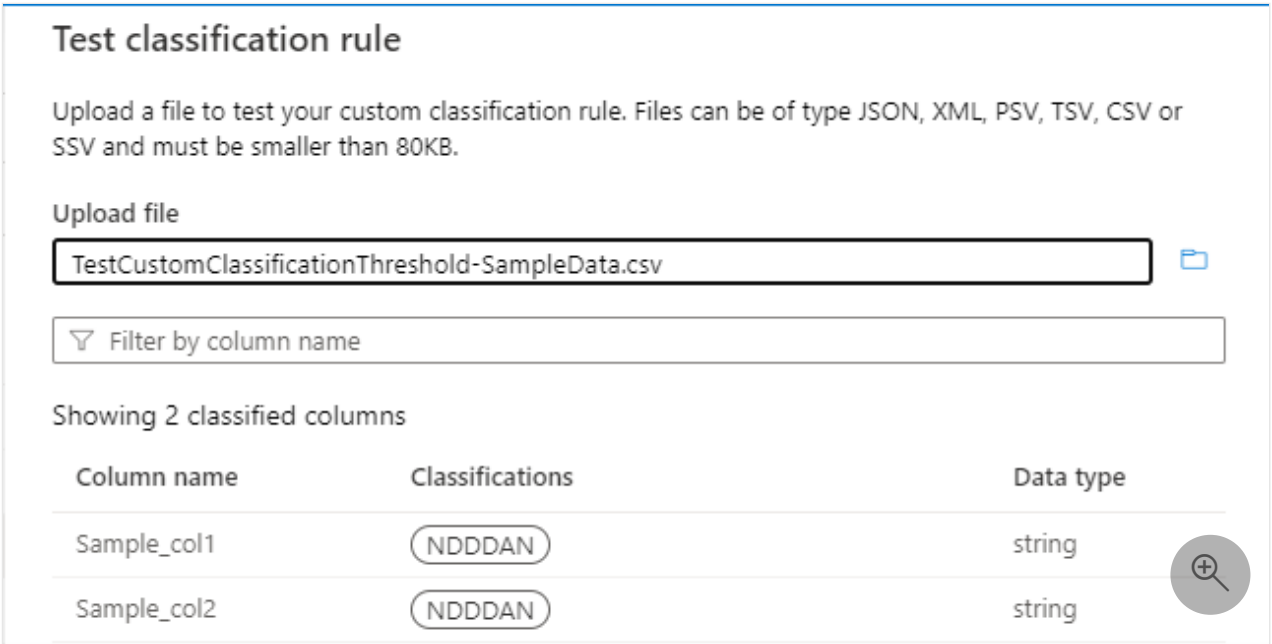
- The custom classification is named NDDDAN.
- The classification rule (regex for the data pattern) is ^N[0-9]{3}AN$.



- The threshold would be computed for the "^N[0-9]{3}AN$" pattern, as shown in the following image:

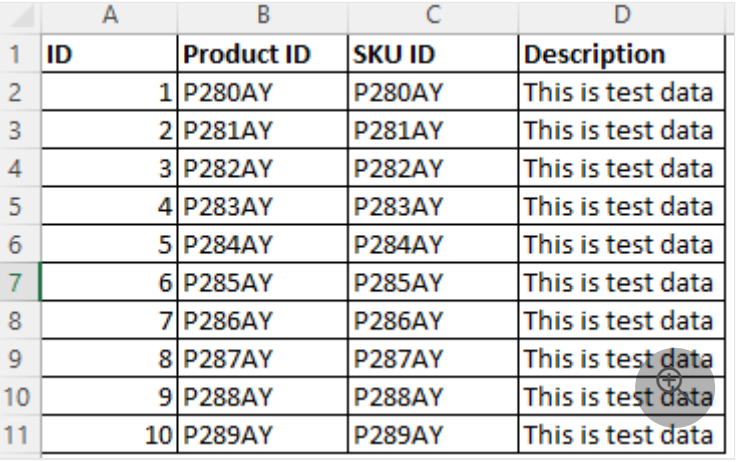


If you have a threshold of 55%, only columns **Sample_col1** and **Sample_col2** will be classified. **Sample_col3** won't be classified, because it doesn't meet the 55% threshold criterion.



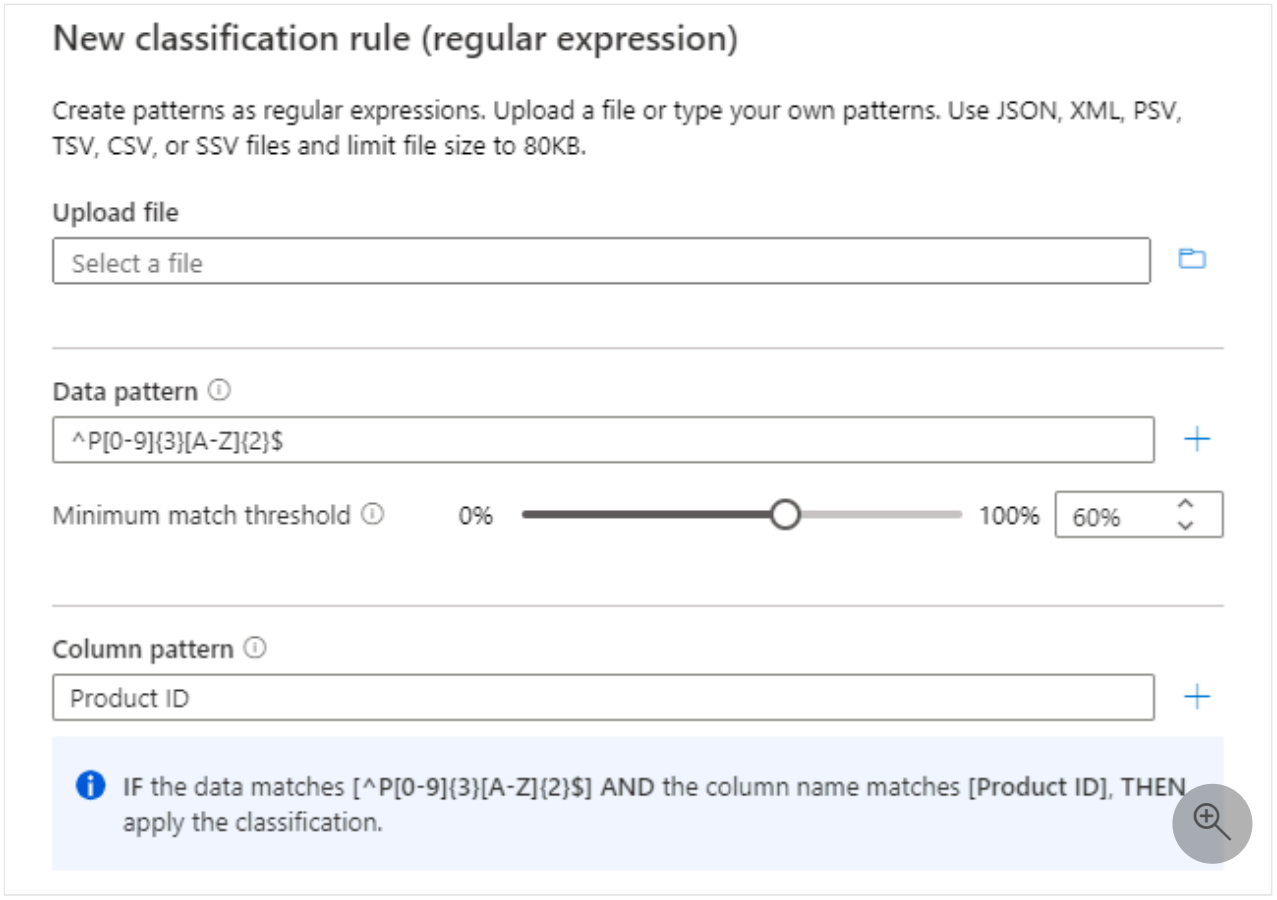
How to use both data and column patterns

- For the given sample data, where both column **B** and column **C** have similar data patterns, you can classify on column **B** based on the data pattern "^P[0-9]{3}[A-Z]{2}$".



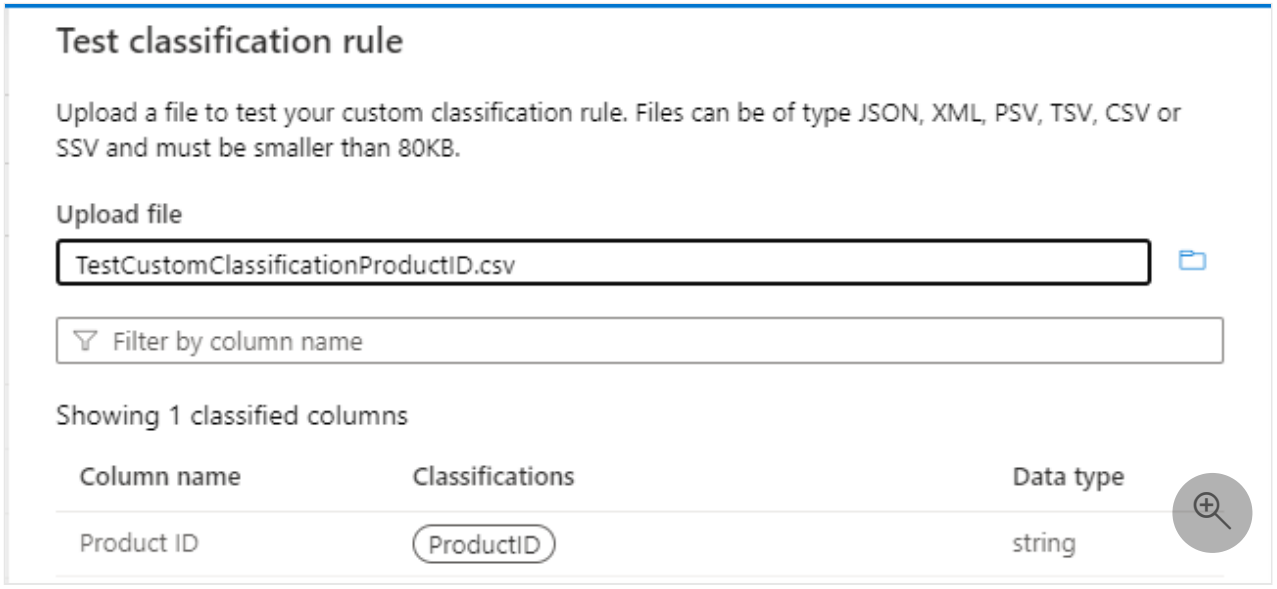
	A	B	C	D
1	**ID**	**Product ID**	**SKU ID**	**Description**
2	1	P280AY	P280AY	This is test data
3	2	P281AY	P281AY	This is test data
4	3	P282AY	P282AY	This is test data
5	4	P283AY	P283AY	This is test data
6	5	P284AY	P284AY	This is test data
7	6	P285AY	P285AY	This is test data
8	7	P286AY	P286AY	This is test data
9	8	P287AY	P287AY	This is test data
10	9	P288AY	P288AY	This is test data
11	10	P289AY	P289AY	This is test data

- Use the column pattern along with the data pattern to ensure that only **Product ID** column is classified.



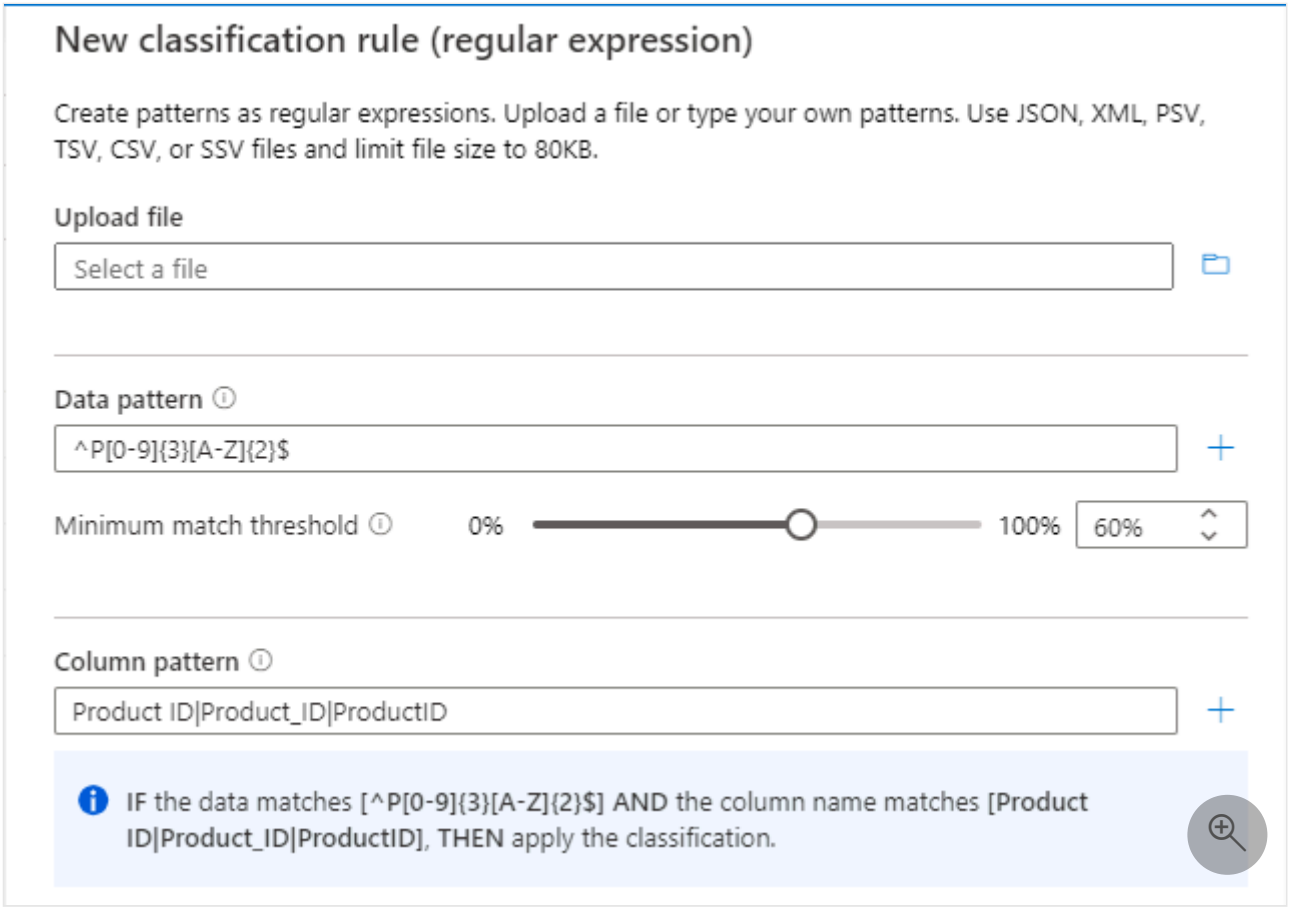
> **Note**
>
> The column pattern is verified as an AND condition with the data pattern.

- Use the *Test classification rule* and test with sample data to verify that the classification rule is working as expected.



How to use multiple column patterns

If there are multiple column patterns to be classified for the same classification rule, use pipe (|) character-separated column names. For example, for columns **Product ID**, **Product_ID**, **ProductID**, and so on, write the column pattern as shown in the following image:



For more information, see regex alternation construct.

Classification considerations

Here are some considerations to bear in mind as you're defining classifications:

- To decide what classifications are required to be applied to the assets prior to scanning, consider how your classifications are to be used. Unnecessary classification labels might look noisy and even misleading for data consumers. You can use classifications to:
  - Describe the nature of the data that exists in the data asset or schema that's being scanned. In other words, classifications should enable customers to identify the content of data asset or schema from the classification labels as they search the catalog.
  - Set priorities and develop a plan to achieve the security and compliance needs of an organization.
  - Describe the phases in the data preparation processes (raw zone, landing zone, and so on) and assign the classifications to specific assets to mark the phase in the process.

- You can assign classifications at the asset or column level automatically by including relevant classifications in the scan rule, or you can assign them manually after you ingest the metadata into the Microsoft Purview Data Map.
- For automatic assignment, see supported data stores for the Microsoft Purview Data Map.
- Before you scan your data sources in the Microsoft Purview Data Map, it's important to understand your data and configure the appropriate scan rule set for it (for example, by selecting relevant system classification, custom classifications, or a combination of both), because it could affect your scan performance. For more information, see supported classifications in Microsoft Purview Data Map.
- The Microsoft Purview scanner applies data sampling rules for deep scans (subject to classification) for both system and custom classifications. The sampling rule is based on the type of data sources. For more information, see the "Sampling within a file" section in Supported data sources and file types in Microsoft Purview.

  > ⓘ **Note**
  >
  > **Distinct data threshold**: This is the total number of distinct data values that need to be found in a column before the scanner runs the data pattern on it. Distinct data threshold has nothing to do with pattern matching but it is a pre-requisite for pattern matching. System classification rules require there to be at least 8 distinct values in each column to subject them to classification. The system requires this value to make sure that the column contains enough data for the scanner to accurately classify it. For example, a column that contains multiple rows that all contain the value 1 won't be classified. Columns that contain one row with a value and the rest of the rows have null values also won't get classified. If you specify multiple patterns, this value applies to each of them.

- The sampling rules apply to resource sets as well. For more information, see the "Resource set file sampling" section in supported data sources and file types in the Microsoft Purview Data Map.
- Custom classifications can't be applied on document type assets using custom classification rules. Classifications for such types can be applied manually only.
- Custom classifications aren't included in any default scan rules. Therefore, if automatic assignment of custom classifications is expected, you must deploy and use a custom scan rule that includes the custom classification to run the scan.
- If you apply classifications manually from the Microsoft Purview governance portal, such classifications are retained in subsequent scans.

- Subsequent scans won't remove any classifications from assets, if they were detected previously, even if the classification rules are inapplicable.
- For *encrypted source* data assets, Microsoft Purview picks only file names, fully qualified names, schema details for structured file types, and database tables. For classification to work, decrypt the encrypted data before you run scans.

Next steps

- Automatically apply classifications
- Manually apply classifications
- Create custom classification

Microsoft Purview domains and collections architectures and best practices

Article • 05/14/2025

At the core of Microsoft Purview data governance solutions, Data Map is a service that keeps an up-to-date map of assets and their metadata across your data estate. To hydrate the data map, you need to register and scan your data sources. In an organization, there might be thousands of sources of data managed and governed by either centralized or decentralized business units, teams, and environments. To manage this, you can use domains and collections in Microsoft Purview.

> **Note**
>
> The best practices recommendations in this article applies to Microsoft Purview accounts that are using a tenant-level account (https://purview.microsoft.com).

Domains

In Microsoft Purview, domains are foundational elements of the Data Map and represent a top-level hierarchy within a Microsoft Purview account. They enable the segregation of responsibilities, effective organization, and management of data governance within the organization specially when there are subsidiaries or business units that operate independently, but they share a common Entra ID tenant. By using domains, organizations can achieve several capabilities, including:

- Organization: Domains help logically group resources such as data sources, assets, scans, and security related resources that belong to a business unit or region.
- Delegation: Domains are a hierarchy above collections, allowing Microsoft Purview admins to delegate specific administrative tasks to subsets of components within Microsoft Purview data governance for specific business units or suborganization.
- Security: By isolating objects within domains, administrators can implement targeted security measures and control access more effectively. For example, resources such as connections, credentials and policies can be specific and visible for a certain domain.
- Lifecycle Management: Domains facilitate the separation of development, test, QA, preproduction, and production resources within the same tenant.
- Isolation of Resources: Domains help isolate resources due to regional, legal, or regulatory requirements.

Collections

Collections in Microsoft Purview support organizational or suborganizational mapping of metadata. By using collections, you can manage and maintain data sources, scans, and assets within a business unit in a hierarchy instead of a flat structure. Collections allow you to build a custom hierarchical model of your data landscape based on how your organization plans to use Microsoft Purview to govern your data.

A collection also provides a security boundary for your metadata in the data map. Access to collections, data sources, and metadata is set up and maintained based on the collections hierarchy in Microsoft Purview, following a least-privilege model:

- Users have the minimum amount of access they need to do their jobs.
- Users don't have access to sensitive data that they don't need.

Understanding the Relationship

- **Domains** are more strategic and policy-centric, while **collections** are more operational and access-centric. For example, in a large healthcare organization with several segments such as Hospitals, Clinics, Research, and Administration, all under the same Microsoft Entra ID tenant, domains, and collections can be defined as follows:
  - Domains: The organization creates domains for each segment. These domains are strategic and policy-centric, meaning they define high-level governance policies, compliance requirements, and data management strategies for each segment. For instance, the Hospitals domain might have policies related to patient data privacy and healthcare regulations, while the Research domain might focus on data sharing agreements and ethical guidelines for clinical trials. Each domain can have their own sets of credentials, scan rulesets, policies, and connections as well as collections, data sources, scans, and assets that aren't visible to users and admins in other domains.
  - Collections: Within the Hospitals domain, there are several operational tasks that need to be managed. The organization creates collections for different operational units such as Emergency Services, Inpatient Care, and Outpatient Services. These collections are more operational and access-centric, meaning they organize data sources, assets, and scan specific to each operational unit. Access to these collections is controlled based on the roles and responsibilities of the users within the Hospitals segment. For example, only emergency department staff might have access to the Emergency Services collection, while inpatient care managers have access to the Inpatient Care collection.
- **Collections** can exist within domains, inheriting the governance policies set at the domain level.
- In Microsoft Purview data governance, domains and collections have distinct functions. An account can have one default domain and up to four custom domains. Each domain can have its own collection hierarchy.

- A user member of **Purview Administrators** role can create and manage domains and delegate access to each business unit to manage their own domains by granting them access as **Purview Domain Manager** role.

Define a hierarchy

Design recommendations

- Start designing your domains and collections architecture based on your organization's legal, security requirements considering data management, and governance structure of your organization. Review the recommended archetypes in this article.
- Consider security and access management as part of your design decision-making process when you build domains and collections in Microsoft Purview.
- Start with default domain and build collection hierarchy inside the default domain. Use additional domains if you have any of the following requirements:
  - You need to build prod and non-prod environments under the same tenant.
  - You have multiple regions and need to logically separate resources and segregate responsibilities across these regions.
  - Your organization has multiple companies or business unit under the same tenant, and need to separate resources and segregate management and responsibilities.
- Each domain or collection has a name attribute and a friendly name attribute. If you use the Microsoft Purview governance portal to deploy a domain or collection, the system automatically assigns a random six-letter name to avoid duplication.
- Currently, a domain or collection name can contain up to 36 characters and a collection friendly name can have up to 100 characters.
- When you can, avoid duplicating your organizational structure into a deeply nested collection hierarchy. If you can't avoid doing so, be sure to use different names for every collection in the hierarchy to make the collections easy to distinguish.
- Automate deployment of domains and collections by using the API if you're planning to deploy domains and collections and role assignments in bulk.
- Use a dedicated service principal name (SPN) to run operations on Data Map for managing domains, collections and role assignments by using the API. Using an SPN

reduces the number of users who have elevated rights and follows least-privilege guidelines.

Design considerations

- Domains are available only to Microsoft Purview accounts using a tenant-level account.
- Consider that a Microsoft Purview account can have up to four domains in addition to the default domain. As part of consolidating your current Microsoft Purview accounts, the content of existing data maps including collections, data sources, assets and scans are migrated to a new domain.
- Create new domain if you're planning to onboard a new organization in your tenant that they have a different legal requirement.
- The following resources are deployed at tenant level and visible across all domains:
  - Typedefs
  - Managed attributes
  - Glossary terms
  - Classifications and classification rules
  - Metamodel
  - Integration runtimes
  - Workflows
- Domains provide separation of the following resources:
  - Credentials
  - Security connections
  - Custom scan rule sets
  - Advanced resource sets and pattern rules
  - Policies
  - ADF connections
  - Collections and all resources that can be scoped to a collection
- Collections provide separation of the following resources:
  - Data sources
  - Scans
  - Assets
- Each Microsoft Purview account is created with a default domain. The default domain name is the same as your Microsoft Purview account name. The default domain can't be removed, however, you can change the default domain's friendly name.

- A collection can have as many child collections as needed. But each collection can have only one domain and a parent collection.
- A collections hierarchy in a Microsoft Purview can support as many as 1000 collections, with a maximum of eight levels of depth. This doesn't include the root collection.
- By design, you can't register data sources multiple times in a single Microsoft Purview account. This architecture helps to avoid the risk of assigning different levels of access control to a single data source. If multiple teams consume the metadata of a single data source, you can register and manage the data source in a parent collection. You can then create corresponding scans under each subcollection so that relevant assets appear under each child collection.
- Lineage connections and artifacts are attached to the default domain even if the data sources are registered at lower-level collections.
- When you run a new scan, by default, the scan is deployed in the same collection as the data source. You can optionally select a different subcollection to run the scan. As a result, the assets belong under the subcollection.
- You can delete a domain if it's empty.
- You can delete a collection if it does not have any assets, associated scans, data sources or child collections.
- Moving data sources across collections is allowed if the user is granted the Data Source Admin role for the source and destination collections.
- Moving assets across collections is allowed if the user is granted the Data Curator role for the source and destination collections.
- To perform move and rename operations on a collection, review the following recommendations and considerations:

    1. To rename a collection, you must be member of collection admins role.
    2. To move a collection, you must be member of collection admins role on the source and destination collections.

Define an authorization model

Microsoft Purview contains roles in Microsoft Defender for Office 365 as well as roles that exist inside Microsoft Purview data plane. After you deploy a Microsoft Purview account, a default domain is automatically created and the creator of the Microsoft Purview account becomes

part of Purview Administrators role. For more information about permissions for the Microsoft Purview Data Map and Unified Catalog, see the roles and permissions documentation.

Design recommendations

- Consider implementing emergency access or a break-glass strategy for your tenant, so you can recover access to Microsoft Purview default domain when needed, to avoid Microsoft Purview account-level lockouts. Document the process for using emergency accounts.
- Minimize the number of Purview Administrators, domain admins and collection admins. Assign a maximum of three domain admin users at the default domain, including the SPN and your break-glass accounts. Assign your Collection Admin roles to the top-level collection or to subcollections instead.
- Assign roles to groups instead of individual users to reduce administrative overhead and errors in managing individual roles.
- Assign the service principal at the root collection for automation purposes.

To increase security, enable Microsoft Entra Conditional Access with multifactor authentication for purview administrators, domain admins and collection admins, data source admins, and data curators. Make sure emergency accounts are excluded from the Conditional Access policy.

Design considerations

- Microsoft Purview access management has moved into data plane and roles in Microsoft Defender for Office 365. Azure Resource Manager roles aren't used anymore, so you should use Microsoft Purview to assign roles.
- In Microsoft Purview, you can assign roles to users, security groups, and service principals (including managed identities) from Microsoft Entra ID on the same Microsoft Entra tenant where the Microsoft Purview account is deployed.
- You must first add guest accounts to your Microsoft Entra tenant as B2B users before you can assign Microsoft Purview roles to external users.
- By default, domain admins also obtain data source admins, data reader and data curator roles so they have access to read or modify assets.
- By default, Global Administrator is added as collection admins on the default domain.
- By default, all role assignments are automatically inherited by all child collections. But you can enable **Restrict inherited permissions** on any collection except for the root collection.

Restrict inherited permissions removes the inherited roles from all parent collections, except for the collection admin role.

- For Azure Data Factory connection: to connect to Azure Data Factory, you have to be a collection admin on the default domain.
- If you need to connect to Azure Data Factory for lineage, grant the Data Curator role to the data factory's managed identity at your Microsoft Purview root collection level. When you connect Data Factory to Microsoft Purview in the authoring UI, Data Factory tries to add these role assignments automatically. If you have the collection admin role on the Microsoft Purview default domain, this operation works.

Domains and collections archetypes

You can deploy your Microsoft Purview domains and collections based on centralized, decentralized, or hybrid data management and governance models. Base this decision on your business, legal and security requirements.

Example 1: Single organization with a single environment and shared legal requirements

This structure is suitable for organizations that:

- Are based in a single geographic location and operate under same legal requirements.
- Have a centralized data management and governance team where the next level of data management falls into departments, teams, or projects.

The hierarchy consists of these verticals:

Domains:

- Default domain: Contoso

Collections under default domain:

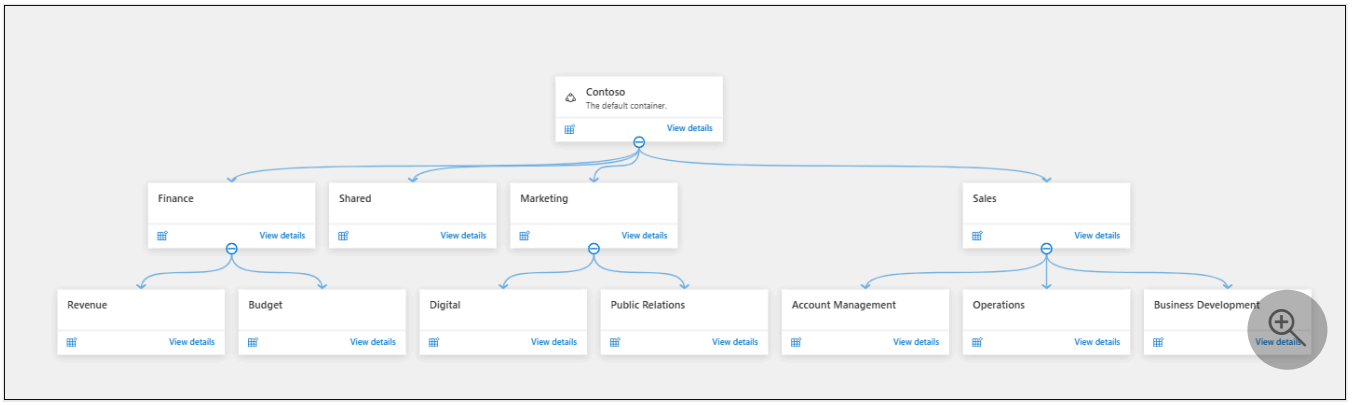
- Departments (a delegated collection for each department)
- Teams or projects (further segregation based on projects)

There's no need for more domains, since there are no specific business or legal requirements to add more.

Organization-level shared data sources are registered and scanned in the Hub collection.

The department-level shared data sources are registered and scanned in the department collections.

Each data source is registered and scanned in its corresponding collection. So assets also appear in the same collection.



Example 2: Single multi-region organization with centralized management

This scenario is useful for organizations:

- That have presence in multiple regions.
- Where the data governance team is centralized or decentralized in each region.
- Where data management teams are distributed in each geographic location, and there's also a centralized federated management.
- Teams that need to manage their own data sources and resources

The domain and collection hierarchy consists of these verticals:

Domains:

- Default domain: FourthCoffee

Collections under default domain:

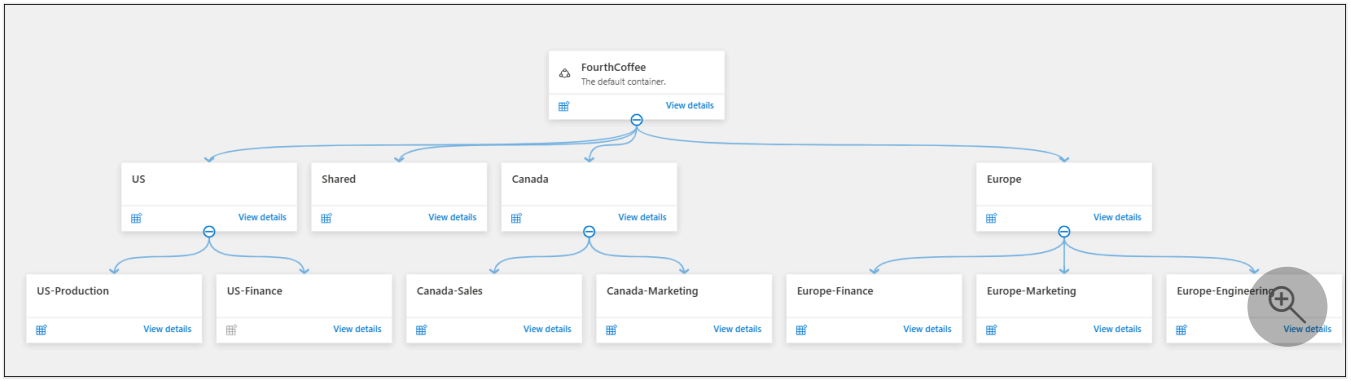
- Geographic locations (top-level collections based on geographic locations where data sources and data owners are located)
- Departments (a delegated collection for each department)
- Teams or projects (further segregation based on projects)

In this scenario, each region has a collection of its own under the default domain in the Microsoft Purview account. Data sources are registered and scanned in the corresponding collections in their own geographic locations. So assets also appear in the collection hierarchy for the region.

If you have centralized data management and governance teams, you can grant them access from the default domain. When you do, they gain oversight for the entire data estate in the data map. Optionally, the centralized team can register and scan any shared data sources. The centralized team can also manage security resources such as credentials and integration runtimes.

Region-based data management and governance teams can obtain access from their corresponding collections.

The department-level shared data sources are registered and scanned in the department collections.



Example 3: Single organization with multiple environments

This scenario can be useful if you have single tenant for all types of prod and non-prod environments and you need to isolate resources as much as possible. Data scientists and data engineers who can transform data to make it more meaningful, can manage Raw and Refine zones. They can then move the data into Produce or Curated zones in the corresponding environments.

The domain and collection hierarchy consists of these verticals:

Domains:

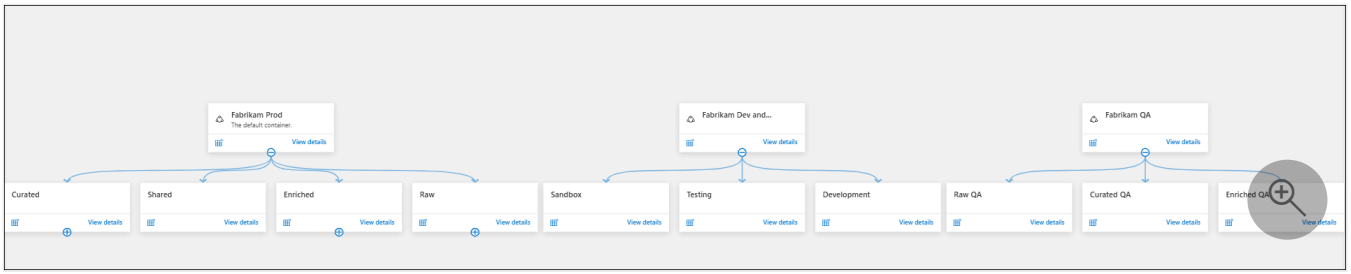
- Default domain: Fabrikam production
- Custom domain 1: Dev and Test
- Custom domain 2: QA

Collections under each domain could follow any of these verticals:

- Departments, Teams or projects (further segregation based on projects)
- Data transformation stages (Raw, Enriched, Produce/Curated, Development, etc.)

Data scientists and data engineers can have the Data Curators role on their corresponding zones so they can curate metadata. Data Reader access to the curated zone can be granted to

entire data personas and business users.



Example 4: Multiple organizations or companies, using the same Entra ID tenant with decentralized management

This option can be used for scenarios where multiple companies share the same Entra ID tenant and each organization need to organize metadata and manage their own resources

> ⓘ **Note**
>
> If you previously had multiple Microsoft Purview accounts in your tenant, the first account you select to migrate becomes default domain and you can upgrade the other accounts into separate domains.

The domain and collection hierarchy consists of these verticals:

Domains:

- Default domain: parent company or organization such as Contoso
- Custom domain 1: FourthCoffee
- Custom domain 2: Fabrikam

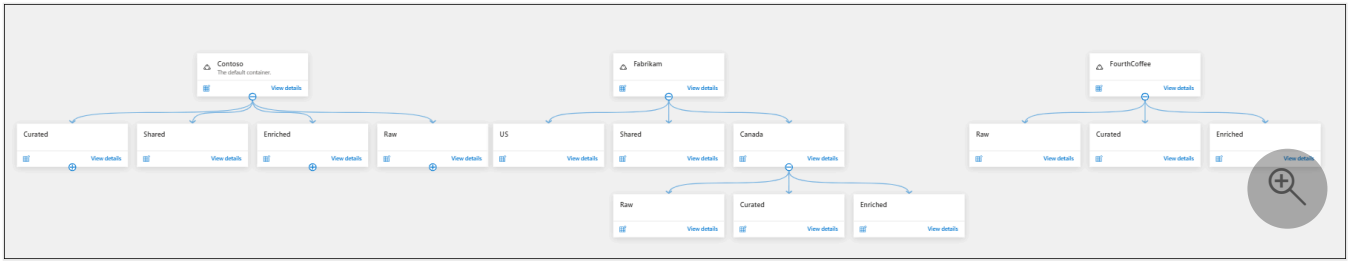
Collections under each domain could follow any of these verticals:

- Departments, teams or projects (further segregation based on projects)
- Data transformation stages (Raw, Enriched, Produce/Curated, Development, etc.)
- Regions within an organization

Each organization has a domain of its own with their own collections hierarchy in the Microsoft Purview account. Security resources are managed inside each domain and data sources are registered and scanned in the corresponding domains. Assets are added to the subcollection hierarchy for the specific domain.

If you have centralized data management and governance organization, that can be the default domain, so they can manage shared resources such as integration run-times, managed attributes, etc.

Organizational data management and governance teams can obtain access from their corresponding collections at a lower level, depending in centralized or decentralized management in each domain.



> **Note**
>
> A shared non-production domain can be created and used by multiple organizations, each org having their own top level collection in the non-prod domain.

Access management options

If you want to implement data democratization across an entire organization, use one domain and assign the Data Reader role at the default domain to data management, governance, and business users. Assign Data Source Admin and Data Curator roles at the subcollection levels to the corresponding data management and governance teams.

If you need to restrict access to metadata search and discovery in your organization, assign Data Reader and Data Curator roles at the specific collection level. For example, you could restrict US employees so they can read data only at the US collection level and not in the LATAM collection.

Create additional domains only if you need to, such as when separating prod and non-pro environments, upgrading multiple accounts into one unified account or have multiple companies inside the same tenant who have different security requirements.

You can apply a combination of these scenarios in your Microsoft Purview Data Map by using domains and collections.

Assign the domain admin role to the centralized data security and management team at the default collection. Delegate further domains or collection management of additional domains and lower-level collections to corresponding teams.

Next steps

- Manage domains and collections in the Microsoft Purview Data Map

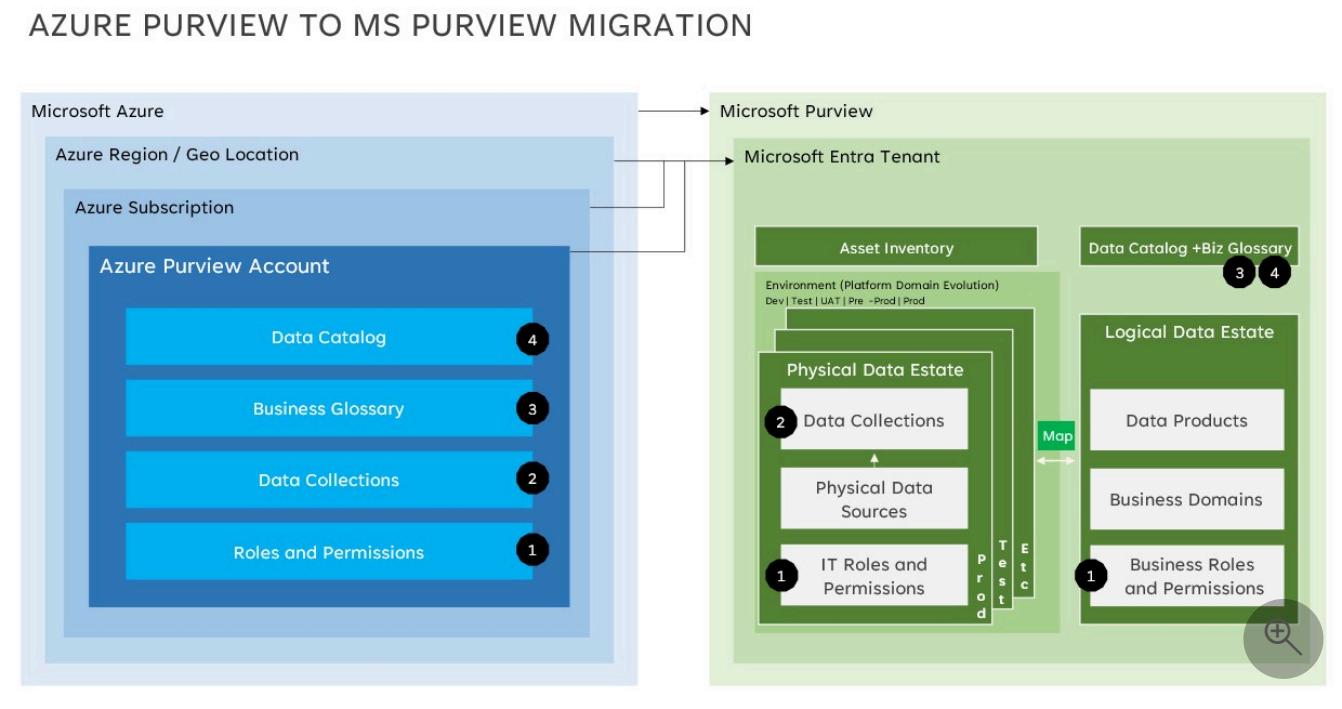
- Governance roles and permissions in free and enterprise account types

Governance domains and domain recommendations

Article • 05/14/2025

> **ⓘ Important**
>
> - Domains in Data Map are limitied to 5 total.
> - Collections are limited to 1000 across all platform domains. (One domain can have 1000 collections; 5 plaform domains would have 200 collections each if spread evenly.)
> - Data sources are registered in one collection in one domain.
> - Assets from a scanned data source must be scanned in the same domain as the data source.
> - Merging an account will create a domain.
> - Fabric domains likey don't map to domains, as the number of Fabric domains will exceed the 5 domains limit.

General recommendations



- If true separation is required separate tenants are needed to enable different implementations of Microsoft Purview. Creating separate tenants isn't a likely solution for most organizations, but in some regulatory environments could be needed.

- Many organizations already have a SaaS model for deployment they can follow from Microsoft 365 or PBI/Fabric. Following that pattern within the organization can help resolve the SaaS deployment concerns.

Collections and domains

- Using collections offer access segmentation for most needs to different users to get only the access required to Data Map.
- Use collections for technical separation along IT teams and enable onboarding, the governance domains can be used for the business team separations and logical concept boundaries (glossary terms and data products).
- For data sources that don't have production and nonproduction separate resources, data sources that have both requires the use of collections for testing because each data source is registered to one domain and one collection.

Domains and governance domains

- Use dev, test, production governance domains (Finance.dev, Finance.prod, etc.) to separate testing with concepts and consumption experience. When testing isn't needed or shouldn't be made visible to consumers of the catalog, you can expire or repurpose the governance domain for production if concepts are ready for production.
- Plan to segment billing with governance domains not collections or platform domains.
- Governance domains can be mapped to platform domains and collections to build a logical relationship between the governance domains and the parts of Data Map that house their data assets.

Domain recommendations

Nonrecommended uses of the domains in Data Map

- Isolation of geo-residency segmentation.
- Domains don't scale to create boundaries along geo-residency segmentation needs.
- Domains don't enable separate storage locations for metadata.
- Business continuity and disaster recovery scenarios aren't recommended because if one domain is down all others are down.
- Metadata would need to be downloaded to an alternate source.
- Business unit isolation is best handled with Governance domains.
- Billing separation would follow the practice of using Governance domains.

Recommended uses of the domains in Data Map

- Lifecycle Management (Development, test, Quality Assurance, Production), most customers have nonproduction and production resources, which are intended to be kept separate from production resources. Separation can be achieved with either domains or collections. When there's nonproduction and production data in the same resource, the customer has to use collections.

Microsoft Purview (formerly Azure Purview) deployment best practices

> **Note**
>
> Microsoft Purview Data Catalog (classic), Data Health Insights (classic), and Purview Workflow (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

> **Note**
>
> These best practices cover the deployment for classic Microsoft Purview governance solutions.
>
> For information about deploying the new Microsoft Purview data governance features, see our quickstart article.
>
> For more information about Microsoft Purview risk and compliance solutions, go here. For more information about Microsoft Purview in general, go here.

This article is a guide to successfully deploying Microsoft Purview (formerly Azure Purview) into production in your data estate. It's intended to help you strategize and phase your deployment from research to hardening your production environment, and is best used in tandem with our deployment checklist.

If you're looking for a strictly technical deployment guide, use the deployment checklist.

If you're creating a plan to deploy Microsoft Purview and want to consider best practices as you develop your deployment strategy, then follow the article below. This guide outlines tasks can be completed in phases over the course of a month or more to develop your deployment process for Microsoft Purview. Even organizations who have already deployed Microsoft Purview can use this guide to ensure they're getting the most out of their investment.

A well-planned deployment of your governance platform, can give the following benefits:

- Better data discovery
- Improved analytic collaboration

- Maximized return on investment

This guide provides insight on a full deployment lifecycle, from initial planning to a mature environment by following these stages:

Expand table

Stage	Description
Identify objectives and goals	Consider what your entire organization wants and needs from data governance.
Gathering questions	What questions might you and your team have as you get started, and where can you look to begin addressing them?
Create a process to move to production	Create a phased deployment strategy tailored to your organization.
Platform hardening	Continue to grow your deployment to maturity.

Many of Microsoft Purview's applications and features have their own individual best practices pages as well. They're referenced often throughout this deployment guide, but you can find all of them in the table of contents under **Concepts** and then **Best practices and guidelines**.

Identify objectives and goals

Many organizations have started their data governance journey by developing individual solutions that cater to specific requirements of isolated groups and data domains across the organization. Although experiences may vary depending on the industry, product, and culture, most organizations find it difficult to maintain consistent controls and policies for these types of solutions.

Some of the common data governance objectives that you might want to identify in the early phases to create a comprehensive data governance experience include:

- Maximizing the business value of your data
- Enabling a data culture where data consumers can easily find, interpret, and trust data
- Increasing collaboration amongst various business units to provide a consistent data experience
- Fostering innovation by accelerating data analytics to reap the benefits of the cloud
- Decreasing time to discover data through self-service options for various skill groups
- Reducing time-to-market for the delivery of analytics solutions that improve service to their customers

- Reducing the operational risks that are due to the use of domain-specific tools and unsupported technology

The general approach is to break down those overarching objectives into various categories and goals. Some examples are:

Expand table

Category	Goal
Discovery	Admin users should be able to scan Azure and non-Azure data sources (including on-premises sources) to gather information about the data assets automatically.
Classification	The platform should automatically classify data based on a sampling of the data and allow manual override using custom classifications.
Consumption	The business users should be able to find information about each asset for both business and technical metadata.
Lineage	Each asset must show a graphical view of underlying datasets so that the users understand the original sources and what changes have been made.
Collaboration	The platform must allow users to collaborate by providing additional information about each data asset.
Reporting	The users must be able to view reporting on the data estate including sensitive data and data that needs extra enrichment.
Data governance	The platform must allow the admin to define policies for access control and automatically enforce the data access based on each user.
Workflow	The platform must have the ability to create and modify workflow so that it's easy to scale out and automate various tasks within the platform.
Integration	Other third-party technologies such as ticketing or orchestration must be able to integrate into the platform via script or REST APIs.

Identify key scenarios

Microsoft Purview governance services can be used to centrally manage data governance across an organization's data estate spanning cloud and on-premises environments. To have a successful implementation, you must identify key scenarios that are critical to the business. These scenarios can cross business unit boundaries or affect multiple user personas either upstream or downstream.

These scenarios can be written up in various ways, but you should include at least these five dimensions:

1. Persona – Who are the users?
2. Source system – What are the data sources such as Azure Data Lake Storage Gen2 or Azure SQL Database?
3. Impact Area – What is the category of this scenario?
4. Detail scenarios – How the users use Microsoft Purview to solve problems?
5. Expected outcome – What is the success criteria?

The scenarios must be specific, actionable, and executable with measurable results. Some example scenarios that you can use:

Expand table

Scenario	Detail	Persona
Catalog business-critical assets	I need to have information about each data sets to have a good understanding of what it is. This scenario includes both business and technical metadata data about the data set in the catalog. The data sources include Azure Data Lake Storage Gen2, Azure Synapse DW, and/or Power BI. This scenario also includes on-premises resources such as SQL Server.	Business Analyst, Data Scientist, Data Engineer
Discover business-critical assets	I need to have a search engine that can search through all metadata in the catalog. I should be able to search using technical term, business term with either simple or complex search using wildcard.	Business Analyst, Data Scientist, Data Engineer, Data Admin
Track data to understand its origin and troubleshoot data issues	I need to have data lineage to track data in reports, predictions, or models back to its original source. I also need to understand the changes made to the data, and where the data has resided throughout the data life cycle. This scenario needs to support prioritized data pipelines Azure Data Factory and Databricks.	Data Engineer, Data Scientist
Enrich metadata on critical data assets	I need to enrich the data set in the catalog with technical metadata that is generated automatically. Classification and labeling are some examples.	Data Engineer, Domain/Business Owner
Govern data assets with friendly user experience	I need to have a Business glossary for business-specific metadata. The business users can use Microsoft Purview for self-service scenarios to annotate their data and enable the data to be discovered easily via search.	Domain/Business Owner, Business Analyst, Data Scientist, Data Engineer

Integration points with Microsoft Purview

It's likely that a mature organization already has an existing data catalog. The key question is whether to continue to use the existing technology and sync with the Microsoft Purview Data Map and Data Catalog or not. To handle syncing with existing products in an organization, Microsoft Purview provides Atlas REST APIs. Atlas APIs provide a powerful and flexible mechanism handling both push and pull scenarios. Information can be published to Microsoft Purview using Atlas APIs for bootstrapping or to push latest updates from another system into Microsoft Purview. The information available in Microsoft Purview can also be read using Atlas APIs and then synced back to existing products.

For other integration scenarios such as ticketing, custom user interface, and orchestration you can use Atlas APIs and Kafka endpoints. In general, there are four integration points with Microsoft Purview:

- **Data Asset** – This enables Microsoft Purview to scan a store's assets in order to enumerate what those assets are and collect any readily available metadata about them. So for SQL this could be a list of DBs, tables, stored procedures, views and config data about them kept in places like `sys.tables`. For something like Azure Data Factory (ADF) this could be enumerating all the pipelines and getting data on when they were created, last run, current state.
- **Lineage** – This enables Microsoft Purview to collect information from an analysis/data mutation system on how data is moving around. For something like Spark this could be gathering information from the execution of a notebook to see what data the notebook ingested, how it transformed it and where it outputted it. For something like SQL, it could be analyzing query logs to reverse engineer what mutation operations were executed and what they did. We support both push and pull based lineage depending on the needs.
- **Classification** – This enables Microsoft Purview to take physical samples from data sources and run them through our classification system. The classification system figures out the semantics of a piece of data. For example, we may know that a file is a Parquet file and has three columns and the third one is a string. But the classifiers we run on the samples will tell us that the string is a name, address, or phone number. Lighting up this integration point means that we've defined how Microsoft Purview can open up objects like notebooks, pipelines, parquet files, tables, and containers.
- **Embedded Experience** – Products that have a "studio" like experience (such as ADF, Synapse, SQL Studio, PBI, and Dynamics) usually want to enable users to discover data they want to interact with and also find places to output data. Microsoft Purview's catalog can help to accelerate these experiences by providing an embedding experience. This experience can occur at the API or the UX level at the partner's option. By embedding a call

to Microsoft Purview, the organization can take advantage of Microsoft Purview's map of the data estate to find data assets, see lineage, check schemas, look at ratings, contacts etc.

Gathering questions

Once your organization agrees on the high-level objectives and goals, there will be many questions from multiple groups. It's crucial to gather these questions in order to craft a plan to address all of the concerns. Make sure to include relevant groups as you gather these questions. You can use our documentation to start answering them.

Some example questions that you may run into during the initial phase:

- What are the main data sources and data systems in our organization?
- What data sources are supported?
- For data sources that aren't supported yet by Microsoft Purview, what are my options?
- How should we budget for Microsoft Purview?
- Who will use Microsoft Purview, and what roles will they have?
- Who can scan new data sources?
- Who can modify content inside of Microsoft Purview?
- What processes can I use to improve the data quality in Microsoft Purview?
- How to bootstrap the platform with existing critical assets, glossary terms, and contacts?
- How can we secure Microsoft Purview?
- How can we gather feedback and build a sustainable process?
- What can we do in a disaster situation?
- We're already using Azure Data Catalog, can we migrate to Microsoft Purview?

Even if you might not have the answer to most of these questions right away, gathering questions can help your organization to frame this project and ensure all "must-have" requirements can be met.

Include the right stakeholders

To ensure the success of implementing Microsoft Purview for your entire organization, it's important to involve the right stakeholders. Only a few people are involved in the initial phase. However, as the scope expands, you'll require more personas to contribute to the project and provide feedback.

Some key stakeholders that you may want to include:

Expand table

Persona	Roles
Chief Data Officer	The CDO oversees a range of functions that may include data management, data quality, master data management, data science, business intelligence, and creating data strategy. They can be the sponsor of the Microsoft Purview implementation project.
Domain/Business Owner	A business person who influences usage of tools and has budget control
Data Analyst	Able to frame a business problem and analyze data to help leaders make business decisions
Data Architect	Design databases for mission-critical line-of-business apps along with designing and implementing data security
Data Engineer	Operate and maintain the data stack, pull data from different sources, integrate and prepare data, set up data pipelines
Data Scientist	Build analytical models and set up data products to be accessed by APIs
DB Admin	Own, track, and resolve database-related incidents and requests within service-level agreements (SLAs); May set up data pipelines
DevOps	Line-of-Business application development and implementation; may include writing scripts and orchestration capabilities
Data Security Specialist	Assess overall network and data security, which involves data coming in and out of Microsoft Purview

Create a process to move to production

Below we've provided a potential four phase deployment plan that includes tasks, helpful links, and acceptance criteria for each phase:

1. Phase 1: Pilot
2. Phase 2: Minimum viable product
3. Phase 3: Pre-production
4. Phase 4: Production

Phase 1: Pilot

In this phase, Microsoft Purview must be created and configured for a small set of users. Usually, it's just a group of 2-3 people working together to run through end-to-end scenarios. They're

considered the advocates of Microsoft Purview in their organization. The main goal of this phase is to ensure key functionalities can be met and the right stakeholders are aware of the project.

Tasks to complete

Expand table

Task	Detail	Duration
Gather & agree on requirements	Discussion with all stakeholders to gather a full set of requirements. Different personas must participate to agree on a subset of requirements to complete for each phase of the project.	One Week
Navigating the Microsoft Purview governance portal	Understand how to use Microsoft Purview from the home page.	One Day
Configure ADF for lineage	Identify key pipelines and data assets. Gather all information required to connect to an internal ADF account.	One Day
Scan a data source such as Azure Data Lake Storage Gen2 or a SQL server.	Add the data source and set up a scan. Ensure the scan successfully detects all assets.	Two Day
Search and browse	Allow end users to access Microsoft Purview and perform end-to-end search and browse scenarios.	One Day

Other helpful links

- Create a Microsoft Purview account
- Create a collection
- Concept: Permissions and access
- Microsoft Purview product glossary

Acceptance criteria

- Microsoft Purview account is created successfully in organization subscription under the organization tenant.
- A small group of users with multiple roles can access Microsoft Purview.
- Microsoft Purview is configured to scan at least one data source.
- Users should be able to extract key values of Microsoft Purview such as:
  - Search and browse
  - Lineage
- Users should be able to assign asset ownership in the asset page.

- Presentation and demo to raise awareness to key stakeholders.
- Buy-in from management to approve more resources for MVP phase.

Phase 2: Minimum viable product

Once you have the agreed requirements and participated business units to onboard Microsoft Purview, the next step is to work on a Minimum Viable Product (MVP) release. In this phase, you'll expand the usage of Microsoft Purview to more users who will have more needs horizontally and vertically. There will be key scenarios that must be met horizontally for all users such as glossary terms, search, and browse. There will also be in-depth requirements vertically for each business unit or group to cover specific end-to-end scenarios such as lineage from Azure Data Lake Storage to Azure Synapse DW to Power BI.

Tasks to complete

Expand table

Task	Detail	Duration
Scan Azure Synapse Analytics	Start to onboard your database sources and scan them to populate key assets	Two Days
Create custom classifications and rules	Once your assets are scanned, your users may realize that there are other use cases for more classification beside the default classifications from Microsoft Purview.	2-4 Weeks
Scan Power BI	If your organization uses Power BI, you can scan Power BI in order to gather all data assets being used by Data Scientists or Data Analysts that have requirements to include lineage from the storage layer.	1-2 Weeks
Import glossary terms	In most cases, your organization may already develop a collection of glossary terms and term assignment to assets. This will require an import process into Microsoft Purview via .csv file.	One Week
Add contacts to assets	For top assets, you may want to establish a process to either allow other personas to assign contacts or import via REST APIs.	One Week
Add sensitive labels and scan	This might be optional for some organizations, depending on the usage of Labeling from Microsoft 365.	1-2 Weeks
Get classification and sensitive insights	For reporting and insight in Microsoft Purview, you can access this functionality to get various reports and provide presentation to management.	One Day
Onboard more users using Microsoft Purview	This step will require the Microsoft Purview Admin to work with the Microsoft Entra Admin to establish new Security Groups to	One Week

Task	Detail	Duration
managed users	grant access to Microsoft Purview.	

Other helpful links

- Collections architecture best practices
- Classification best practices
- Labeling best practices

Acceptance criteria

- Successfully onboard a larger group of users to Microsoft Purview (50+)
- Scan business critical data sources
- Import and assign all critical glossary terms
- Successfully test important labeling on key assets
- Successfully met minimum scenarios for participated business units' users

Phase 3: Pre-production

Once the MVP phase has passed, it's time to plan for pre-production milestone. You may want to include scanning on on-premises data sources such as SQL Server. If there's any gap in data sources not supported by Microsoft Purview, it's time to explore the Atlas API to understand other options.

Tasks to complete

Expand table

Task	Detail	Duration
Refine your scan using scan rule set	Your organization will have many data sources for pre-production. It's important to pre-define key criteria for scanning so that classifications and file extension can be applied consistently across the board.	1-2 Days
Assess region availability for scan for each of your sources by checking source pages	Depending on the region of the data sources and organizational requirements on compliance and security, you may want to consider what regions must be available for scanning.	One Day
Understand firewall concept when scanning	This step requires some exploration of how the organization configures its firewall and how Microsoft Purview can authenticate itself to access the data sources for scanning.	One Day

Task	Detail	Duration
Understand Private Link concept when scanning	If your organization uses Private Link, you must lay out the foundation of network security to include Private Link as a part of the requirements.	One Day
Scan on-premises SQL Server	This is optional if you have on-premises SQL Server. The scan will require setting up Self-hosted Integration Runtime and adding SQL Server as a data source.	1-2 Weeks
Use Microsoft Purview REST API for integration scenarios	If you have requirements to integrate Microsoft Purview with other third party technologies such as orchestration or ticketing system, you may want to explore REST API area.	1-4 Weeks
Understand Microsoft Purview pricing	This step will provide the organization important financial information to make decision.	1-5 Days

Other helpful links

- Labeling best practices
- Network architecture best practices

Acceptance criteria

- Successfully onboard at least one business unit with all of users
- Scan on-premises data source such as SQL Server
- POC at least one integration scenario using REST API
- Complete a plan to go to production, which should include key areas on infrastructure and security

Phase 4: Production

The above phases should be followed to create an effective data lifecycle management, which is the foundation for better governance programs. Data governance will help your organization prepare for the growing trends such as AI, Hadoop, IoT, and blockchain. It's just the start for many things data and analytics, and there's plenty more that can be discussed. The outcome of this solution would deliver:

- **Business Focused** - A solution that is aligned to business requirements and scenarios over technical requirements.
- **Future Ready** - A solution will maximize default features of the platform and use standardized industry practices for configuration or scripting activities to support the advancements/evolution of the platform.

Tasks to complete

Expand table

Task	Detail	Duration
Scan production data sources with Firewall enabled	If this is optional when firewall is in place but it's important to explore options to hardening your infrastructure.	1-5 Days
Enable Private Link	If this is optional when Private Link is used. Otherwise, you can skip this as it's a must-have criterion when Private is enabled.	1-5 Days
Create automated workflow	Workflow is important to automate process such as approval, escalation, review and issue management.	2-3 Weeks
Create operation documentation	Data governance isn't a one-time project. It's an ongoing program to fuel data-driven decision making and creating opportunities for business. It's critical to document key procedure and business standards.	One Week

Other helpful links

- Manage workflow runs
- Workflow requests and approvals
- Manage integration runtimes
- Request access to a data asset

Acceptance criteria

- Successfully onboard all business unit and their users
- Successfully meet infrastructure and security requirements for production
- Successfully meet all use cases required by the users

Platform hardening

More hardening steps can be taken:

- Increase security posture by enabling scan on firewall resources or use Private Link
- Fine-tune scope scan to improve scan performance
- Use REST APIs to export critical metadata and properties for backup and recovery
- Use workflow to automate ticketing and eventing to avoid human errors
- Use policies to manage access to data assets through the Microsoft Purview governance portal.

Lifecycle considerations

Another important aspect to include in your production process is how classifications and labels can be migrated. Microsoft Purview has over 90 system classifiers. You can apply system or custom classifications on file, table, or column assets. Classifications are like subject tags and are used to mark and identify content of a specific type found within your data estate during scanning. Sensitivity labels are used to identify the categories of classification types within your organizational data, and then group the policies you wish to apply to each category. It makes use of the same sensitive information types as Microsoft 365, allowing you to stretch your existing security policies and protection across your entire content and data estate. It can scan and automatically classify documents. For example, if you have a file named multiple.docx and it has a National ID number in its content, Microsoft Purview will add classification such as EU National Identification Number in the Asset Detail page.

In the Microsoft Purview Data Map, there are several areas where the Catalog Administrators need to ensure consistency and maintenance best practices over its life cycle:

- **Data assets** – Data sources will need to be rescanned across environments. It's not recommended to scan only in development and then regenerate them using APIs in Production. The main reason is that the Microsoft Purview scanners do a lot more "wiring" behind the scenes on the data assets, which could be complex to move them to a different Microsoft Purview instance. It's much easier to just add the same data source in production and scan the sources again. The general best practice is to have documentation of all scans, connections, and authentication mechanisms being used.
- **Scan rule sets** – This is your collection of rules assigned to specific scan such as file type and classifications to detect. If you don't have that many scan rule sets, it's possible to just re-create them manually again via Production. This will require an internal process and good documentation. However, if your rule sets change on a daily or weekly basis, this could be addressed by exploring the REST API route.
- **Custom classifications** – Your classifications may not also change regularly. During the initial phase of deployment, it may take some time to understand various requirements to come up with custom classifications. However, once settled, this will require little change. So the recommendation here's to manually migrate any custom classifications over or use the REST API.
- **Glossary** – It's possible to export and import glossary terms via the UX. For automation scenarios, you can also use the REST API.
- **Resource set pattern policies** – This functionality is advanced for any typical organizations to apply. In some cases, your Azure Data Lake Storage has folder naming conventions and

specific structure that may cause problems for Microsoft Purview to generate the resource set. Your business unit may also want to change the resource set construction with more customizations to fit the business needs. For this scenario, it's best to keep track of all changes via REST API, and document the changes through external versioning platform.

- **Role assignment** – This is where you control who has access to Microsoft Purview and which permissions they have. Microsoft Purview also has REST API to support export and import of users and roles but this isn't Atlas API-compatible. The recommendation is to assign an Azure Security Group and manage the group membership instead.

Moving tenants

Moving tenants is not currently supported for Microsoft Purview.

Moving subscriptions

It's possible to move your Microsoft Purview account between subscriptions. However, if your account was created before December 15, 2023 (or deployed using an API version previous to 2023-05-01-preview), or is using a managed Event Hubs, the managed storage account and managed Event Hubs associated with your Microsoft Purview account will not migrate with your instance. Your Microsoft Purview account will still be able to function, but you should not remove these resources.

If you need to remove the managed resources from the other subscription, you will need to create a new Microsoft Purview account and migrate your information to this new account, before removing the original and its managed resources.

Next steps

- Collections best practices
- Navigate the home page and search for an asset

Last updated on 10/28/2024

Labeling best practices for Data Map

Article • 05/14/2025

> **Note**
>
> These best practices cover applying sensitivity labels to your data in Data Map. For more information about applying sensitivity labels in Microsoft Purview risk and compliance solutions, see Automatically apply a sensitivity label to Microsoft 365 data.

Microsoft Purview Data Map supports labeling structured and unstructured data stored across various data sources. Labeling data within Data Map allows users to easily find data that matches predefined autolabeling rules that were configured in the Microsoft Purview portal. Data Map extends the use of sensitivity labels from Microsoft Purview Information Protection to assets stored in infrastructure cloud locations and structured data sources.

Protect personal data with custom sensitivity labels for Microsoft Purview

Storing and processing personal data is subject to special protection. Labeling personal data is crucial to help you identify sensitive information. You can use the detection and labeling tasks for personal data in different stages of your workflows. Because personal data is ubiquitous and fluid in your organization, you need to define identification rules for building policies that suit your individual situation.

Why do you need to use labeling within Data Map

In Data Map, you can extend your organization's investment in sensitivity labels from Microsoft Purview Information Protection to assets that are stored in files and database columns within Azure, multicloud, and on-premises locations. These locations are defined in supported data sources. When you apply sensitivity labels to your content, you can keep your data secure by stating how sensitive certain data is in your organization. The Data Map also abstracts the data itself, so you can use labels to track the type of data, without exposing sensitive data on another platform.

Microsoft Purview Data Map labeling best practices and considerations

The following sections walk you through the process of implementing labeling for your assets.

Get started

- To enable sensitivity labeling in the data map, follow the steps in automatically apply sensitivity labels to your data in the Microsoft Purview Data Map.
- To find information on required licensing and helpful answers to other questions, see Sensitivity labels in the Microsoft Purview Data Map FAQ.

Label considerations

- If you already have sensitivity labels from Microsoft Purview Information Protection in use in your environment, continue to use your existing labels. Don't make duplicate or more labels for Data Map. This approach allows you to maximize the investment you've already made in the Microsoft Purview. It also ensures consistent labeling across your data estate.
- If you haven't created sensitivity labels in Microsoft Purview Information Protection, review the documentation to get started with sensitivity labels. Creating a classification schema is a tenant-wide operation. Discuss it thoroughly before you enable it within your organization.

Label recommendations

- When you configure sensitivity labels for the Microsoft Purview Data Map, you might define autolabeling rules for files, database columns, or both within the label properties. Microsoft Purview labels files within the Microsoft Purview Data Map. When the autolabeling rule is configured, Microsoft Purview automatically applies the label or recommends that the label is applied.

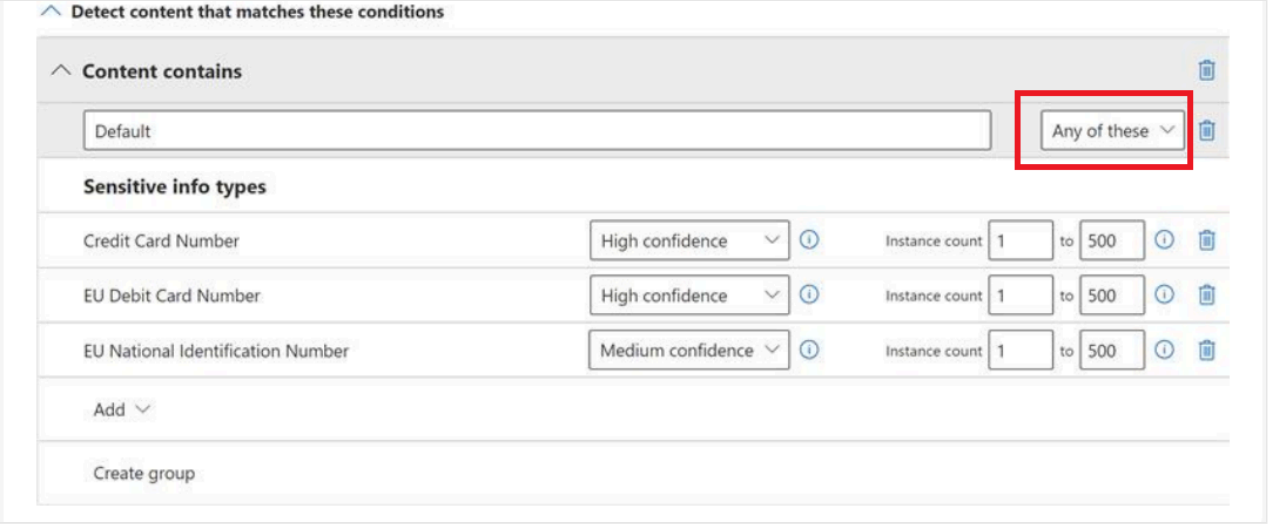
  > ⚠ **Warning**
  >
  > If you haven't configured autolabeling for items on your sensitivity labels, users might be affected within your Office and Microsoft 365 environment. You can test autolabeling on database columns without affecting users.

- If you're defining new autolabeling rules for files when you configure labels for the Microsoft Purview Data Map, make sure that you have the condition for applying the label set appropriately.

- You can set the detection criteria to **All of these** or **Any of these** in the upper right of the autolabeling for items page of the label properties.

- The default setting for detection criteria is **All of these**. This setting means that the asset must contain all the specified sensitive information types for the label to be applied.

While the default setting might be valid in some instances, many customers want to use **Any of these**. Then if at least one asset is found, the label is applied.



> **Note**
>
> Trainable classifiers from Microsoft Purview Information Protection aren't supported by Microsoft Purview Data Map.

- Maintain consistency in labeling across your data estate. If you use autolabeling rules for files, use the same sensitive information types for autolabeling database columns.
- Define your sensitivity labels via Microsoft Purview Information Protection to identify your personal data at a central place.
- Use policy templates as a starting point to build your rule sets.
- Combine data classifications to an individual rule set.
- Force labeling by using autolabel functionality.
- Build groups of sensitivity labels and store them as a dedicated sensitivity label policy. For example, store all required sensitivity labels for regulatory rules by using the same sensitivity label policy to publish.
- Capture all test cases for your labels. Test your label policies with all applications you want to secure.
- Promote sensitivity label policies to the Microsoft Purview Data Map.
- Run test scans from the Microsoft Purview Data Map on different data sources like hybrid cloud and on-premises to identify sensitivity labels.

- Gather and consider insights, for example, by using Microsoft Purview Data Estate Insights. Use alerting mechanisms to mitigate potential breaches of regulations.

By using sensitivity labels with Microsoft Purview Data Map, you can extend information protection beyond the border of your Microsoft data estate to your on-premises, hybrid cloud, multicloud, and software as a service (SaaS) scenarios.

Next steps

- Get started with sensitivity labels.
- How to automatically apply sensitivity labels to your data in the Microsoft Purview Data Map.

Microsoft Purview network architecture and best practices

Article • 02/24/2025

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

Microsoft Purview governance solutions are a platform as a service (PaaS) solutions for data governance. Microsoft Purview accounts have public endpoints that are accessible through the internet to connect to the service. However, all endpoints are secured through Microsoft Entra logins and role-based access control (RBAC).

> **Note**
>
> These best practices cover the network architecture for **Microsoft Purview unified governance solutions**. For more information about Microsoft Purview risk and compliance solutions, **go here**. For more information about Microsoft Purview in general, **go here**.

For an added layer of security, you can create private endpoints for your Microsoft Purview account. You'll get a private IP address from your virtual network in Azure to the Microsoft Purview account and its managed resources. This address will restrict all traffic between your virtual network and the Microsoft Purview account to a private link for user interaction with the APIs and Microsoft Purview governance portal, or for scanning and ingestion.

Currently, the Microsoft Purview firewall provides access control for the public endpoint of your purview account. You can use the firewall to allow all access or to block all access through the public endpoint when using private endpoints. For more information see, Microsoft Purview firewall options

Based on your network, connectivity, and security requirements, you can set up and maintain Microsoft Purview accounts to access underlying services or ingestion. Use this best practices guide to define and prepare your network environment so you can access Microsoft Purview and scan data sources from your network or cloud.

This guide covers the following network options:

- Use Azure public endpoints.
- Use private endpoints.
- Use private endpoints and allow public access on the same Microsoft Purview account.
- Use Azure public endpoints to access Microsoft Purview governance portal and private endpoints for ingestion.

This guide describes a few of the most common network architecture scenarios for Microsoft Purview. Though you're not limited to those scenarios, keep in mind the limitations of the service when you're planning networking for your Microsoft Purview accounts.

Prerequisites

To understand which network option is best for your environment, we suggest that you perform the following actions first:

- Review your network topology and security requirements before registering and scanning any data sources in Microsoft Purview. For more information, see: Define an Azure network topology.

- Define your network connectivity model for PaaS services.

Option 1: Use public endpoints

By default, you can use Microsoft Purview accounts through the public endpoints accessible over the internet. Allow public networks in your Microsoft Purview account if you have the following requirements:

- No private connectivity is required when scanning or connecting to Microsoft Purview endpoints.
- All data sources are software-as-a-service (SaaS) applications only.
- All data sources have a public endpoint that's accessible through the internet.
- Business users require access to a Microsoft Purview account and the Microsoft Purview governance portal through the internet.

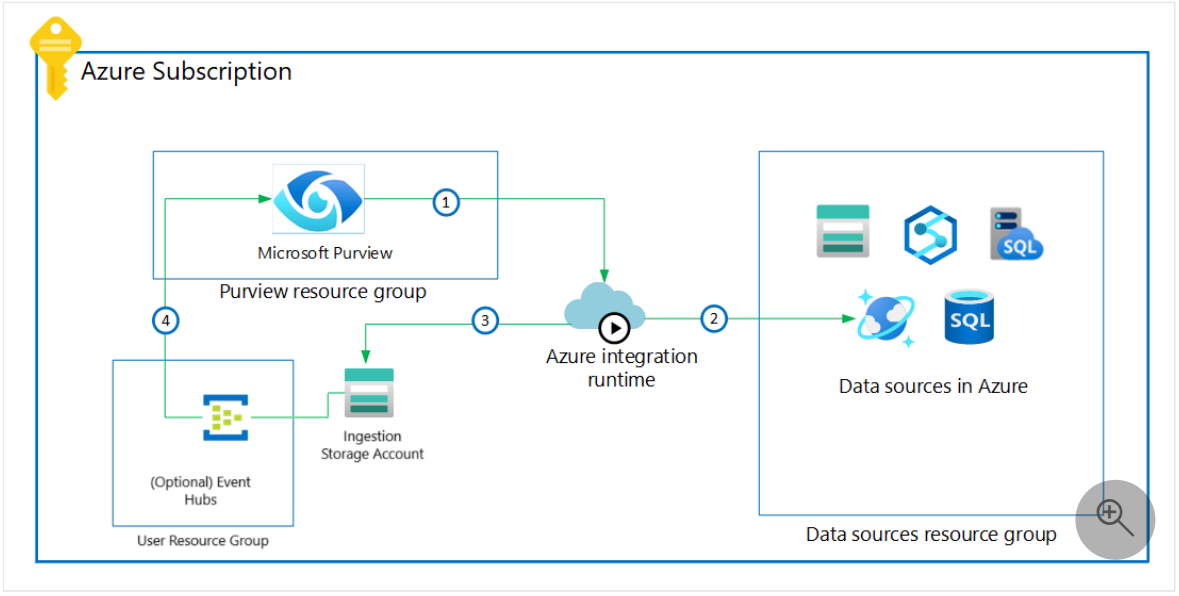
Integration runtime options

To scan data sources while the Microsoft Purview account firewall is set to allow public access, you can use all the integration runtime types - the Azure integration runtime,

Managed VNet integration runtime, and a self-hosted integration runtime. Learn more from Choose the right integration runtime configuration for your scenario.

Here are some best practices:

- Whenever applicable, we recommend that you use the Azure integration runtime or Managed VNet integration runtime to scan data sources, to reduce cost and administrative overhead.
- The following steps show the communication flow at a high level when you're using the Azure integration runtime to scan a data source:

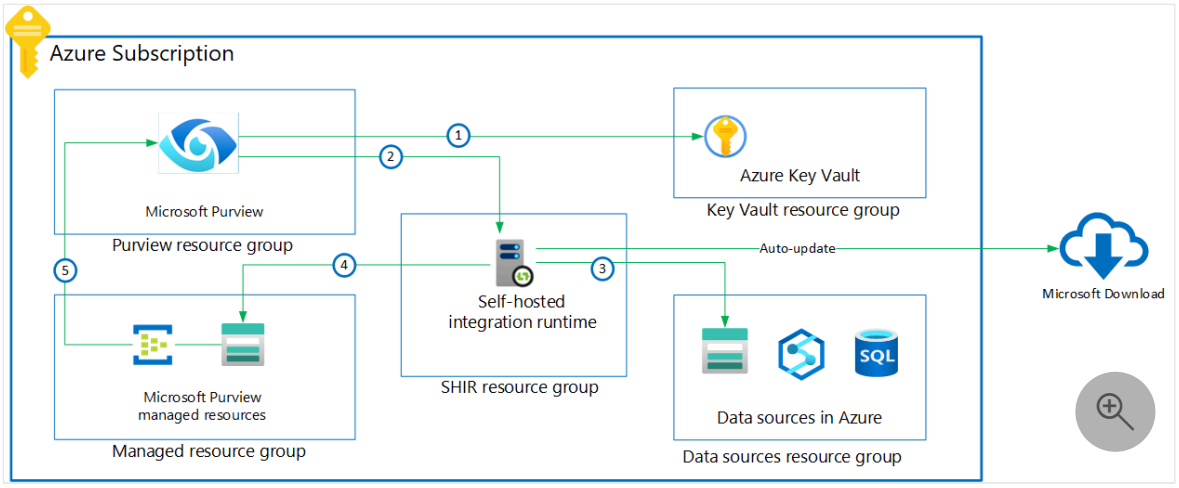


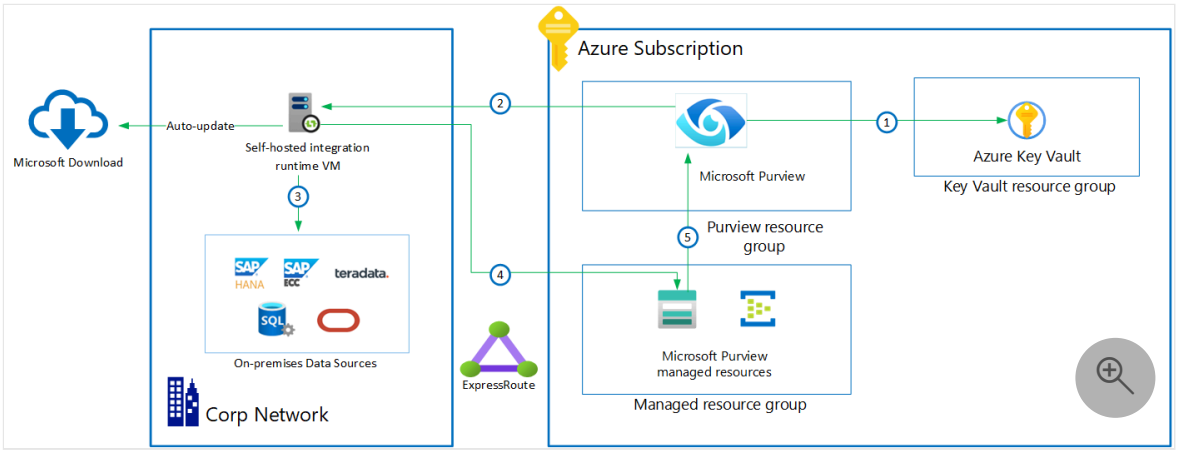
> **Note**
>
> This graphic only applies to Microsoft Purview accounts created after December 15, 2023 (or deployed using API version 2023-05-01-preview onwards).

1. A manual or automatic scan is initiated from the Microsoft Purview Data Map through the Azure integration runtime.
2. The Azure integration runtime connects to the data source to extract metadata.
3. Metadata is queued in the Microsoft Purview ingestion storage account and stored in Azure Blob Storage temporarily.
4. Metadata is sent to the Microsoft Purview Data Map.

- Scanning on-premises and VM-based data sources always requires using a self-hosted integration runtime. The Azure integration runtime isn't supported for

these data sources. The following steps show the communication flow at a high level when you're using a self-hosted integration runtime to scan a data source. The first diagram shows a scenario where resources are within Azure or on a VM in Azure. The second diagram shows a scenario with on-premises resources. The steps between the two are the same from Microsoft Purview's perspective:





1. A manual or automatic scan is triggered. Microsoft Purview connects to Azure Key Vault to retrieve the credential to access a data source.

2. The scan is initiated from the Microsoft Purview Data Map through a self-hosted integration runtime.

3. The self-hosted integration runtime service from the VM or on-premises machine connects to the data source to extract metadata.

4. Metadata is processed in the machine's memory for the self-hosted integration runtime. Metadata is queued in Microsoft Purview ingestion storage and then stored in Azure Blob Storage temporarily. Actual data never leaves the boundary of your network.

5. Metadata is sent to the Microsoft Purview Data Map.

Authentication options

When you're scanning a data source in Microsoft Purview, you need to provide a credential. Microsoft Purview can then read the metadata of the assets from the data source by using the integration runtime. Refer to each data source article for details about the supported authentication types and the needed permissions. Authentication options and requirements vary based on the following factors:

- **Data source type**. For example, if the data source is Azure SQL Database, you need to use a login with db_datareader access to each database. This can be a user-assigned managed identity or a Microsoft Purview managed identity. Or it can be a service principal in Microsoft Entra ID added to SQL Database as db_datareader.

  If the data source is Azure Blob Storage, you can use a Microsoft Purview managed identity, or a service principal in Microsoft Entra ID added as a Blob Storage Data Reader role on the Azure storage account. Or use the storage account's key.

- **Authentication type**. We recommend that you use a Microsoft Purview managed identity to scan Azure data sources when possible, to reduce administrative overhead. For any other authentication types, you need to set up credentials for source authentication inside Microsoft Purview:

  1. Generate a secret inside an Azure key vault.
  2. Register the key vault inside Microsoft Purview.
  3. Inside Microsoft Purview, create a new credential by using the secret saved in the key vault.

- **Runtime type that's used in the scan**. Currently, you can't use a Microsoft Purview managed identity with a self-hosted integration runtime.

Other considerations

- If you choose to scan data sources using public endpoints, your self-hosted integration runtime VMs must have outbound access to data sources and Azure endpoints.
- Your self-hosted integration runtime VMs must have outbound connectivity to Azure endpoints.

Option 2: Use private endpoints

Similar to other PaaS solutions, Microsoft Purview doesn't support deploying directly into a virtual network. So you can't use certain networking features with the offering's

resources, such as network security groups, route tables, or other network-dependent appliances such as Azure Firewall. Instead, you can use private endpoints that can be enabled on your virtual network. You can then disable public internet access to securely connect to Microsoft Purview.

You must use private endpoints for your Microsoft Purview account if you have any of the following requirements:

- You need to have end-to-end network isolation for Microsoft Purview accounts and data sources.
- You need to block public access to your Microsoft Purview accounts.
- Your platform-as-a-service (PaaS) data sources are deployed with private endpoints, and you've blocked all access through the public endpoint.
- Your on-premises or infrastructure-as-a-service (IaaS) data sources can't reach public endpoints.

Design considerations

- To connect to your Microsoft Purview account privately and securely, you need to deploy an account and a portal private endpoint. For example, this deployment is necessary if you intend to connect to Microsoft Purview through the API or use the Microsoft Purview governance portal.
- If you need to connect to the Microsoft Purview governance portal by using private endpoints, you have to deploy both account and portal private endpoints.
- To scan data sources through private connectivity, you need to configure at least one account and one ingestion private endpoint for Microsoft Purview.
- Review DNS requirements. If you're using a custom DNS server on your network, clients must be able to resolve the fully qualified domain name (FQDN) for the Microsoft Purview account endpoints to the private endpoint's IP address.

Integration runtime options

To scan data sources through private connectivity, you can use Managed VNet integration runtime or self-hosted integration runtime. Learn more from Choose the right integration runtime configuration for your scenario.

- Whenever applicable, we recommend that you use the Managed VNet integration runtime to scan data sources, to reduce cost and administrative overhead.

- If using self-hosted integration runtime, you need to set up and use a self-hosted integration runtime on a Windows virtual machine that's deployed inside the same or a peered virtual network where Microsoft Purview ingestion private endpoints are deployed.
- To scan on-premises data sources, you can also install a self-hosted integration runtime either on an on-premises Windows machine or on a VM inside an Azure virtual network.
- When you're using private endpoints with Microsoft Purview, you need to allow network connectivity from data sources to the self-hosted integration VM on the Azure virtual network where Microsoft Purview private endpoints are deployed.
- We recommend allowing automatic upgrade of the self-hosted integration runtime. Make sure you open required outbound rules in your Azure virtual network or on your corporate firewall to allow automatic upgrade. For more information, see Self-hosted integration runtime networking requirements.

Authentication options

- Make sure that your credentials are stored in an Azure key vault and registered inside Microsoft Purview.
- You must create a credential in Microsoft Purview based on each secret that you create in the Azure key vault. You need to assign, at minimum, *get* and *list* access for secrets for Microsoft Purview on the Key Vault resource in Azure. Otherwise, the credentials won't work in the Microsoft Purview account.

Current limitations

- Scanning multiple Azure sources by using the entire subscription or resource group through ingestion private endpoints and a self-hosted integration runtime or Managed VNet integration runtime isn't supported when you're using private endpoints for ingestion. Instead, you can register and scan data sources individually.
- For limitations related to Microsoft Purview private endpoints, see Known limitations.
- For limitations related to the Private Link service, see Azure Private Link limits.

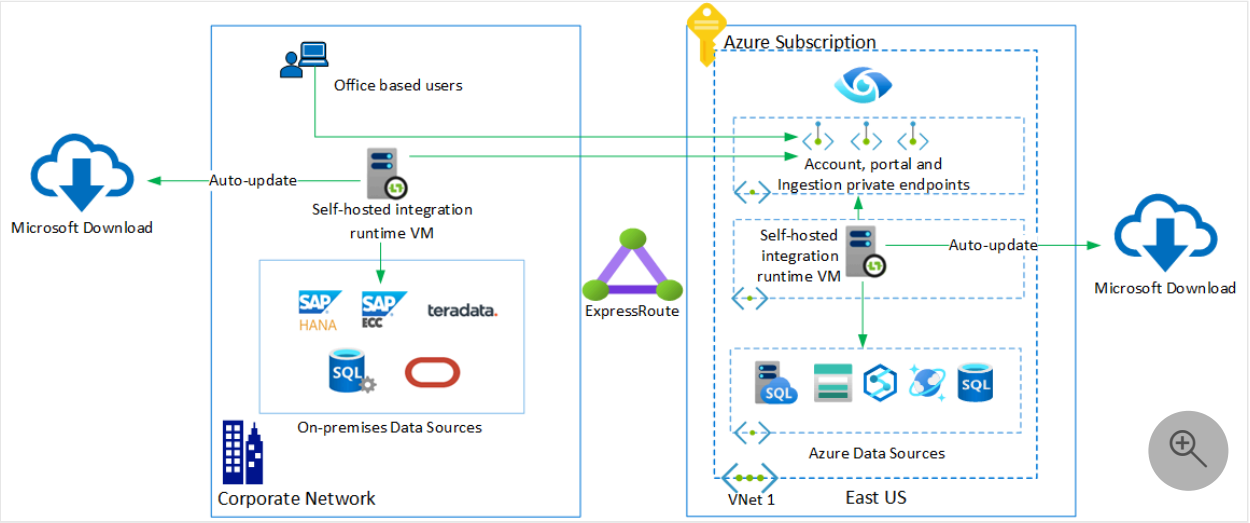
Private endpoint scenarios

Single virtual network, single region

In this scenario, all Azure data sources, self-hosted integration runtime VMs, and Microsoft Purview private endpoints are deployed in the same virtual network in an Azure subscription.

If on-premises data sources exist, connectivity is provided through a site-to-site VPN or Azure ExpressRoute connectivity to an Azure virtual network where Microsoft Purview private endpoints are deployed.

This architecture is suitable mainly for small organizations or for development, testing, and proof-of-concept scenarios.



Single region, multiple virtual networks

To connect two or more virtual networks in Azure together, you can use virtual network peering. Network traffic between peered virtual networks is private and is kept on the Azure backbone network.

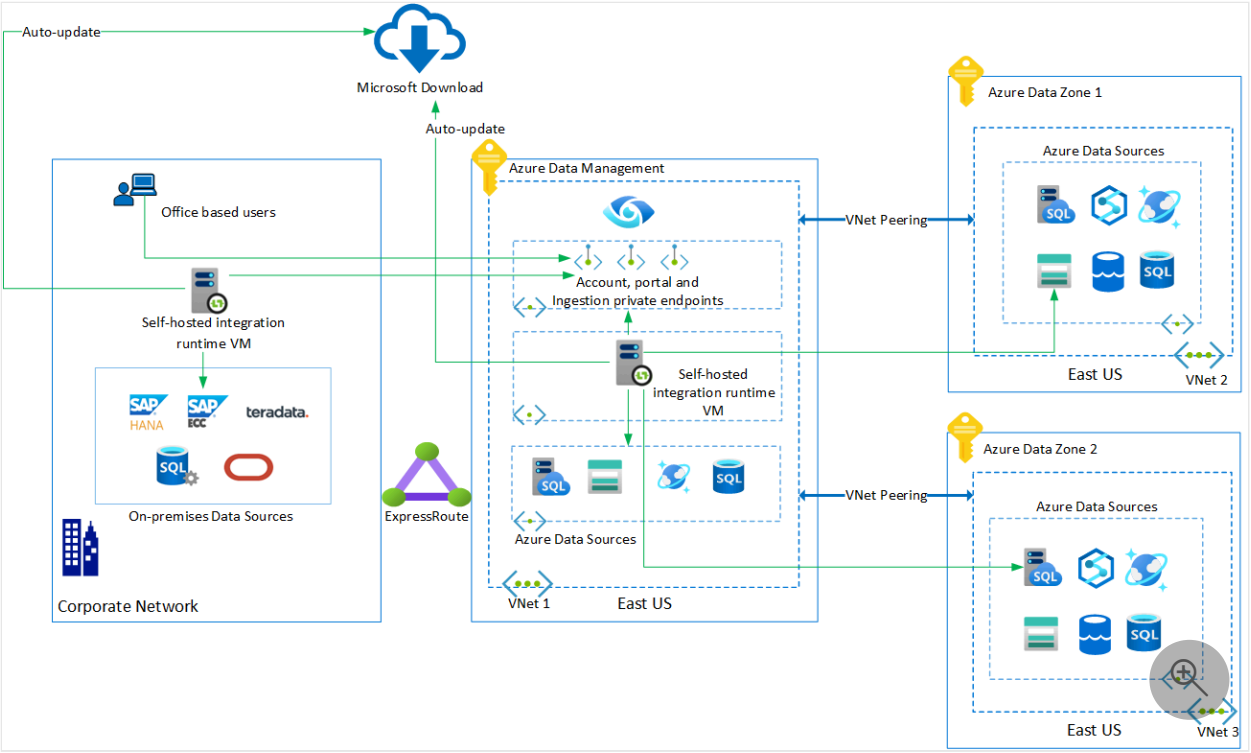
Many customers build their network infrastructure in Azure by using the hub-and-spoke network architecture, where:

- Networking shared services (such as network virtual appliances, ExpressRoute/VPN gateways, or DNS servers) are deployed in the hub virtual network.
- Spoke virtual networks consume those shared services via virtual network peering.

In hub-and-spoke network architectures, your organization's data governance team can be provided with an Azure subscription that includes a virtual network (hub). All data services can be located in a few other subscriptions connected to the hub virtual network through a virtual network peering or a site-to-site VPN connection.

In a hub-and-spoke architecture, you can deploy Microsoft Purview and one or more self-hosted integration runtime VMs in the hub subscription and virtual network. You can register and scan data sources from other virtual networks from multiple subscriptions in the same region.

The self-hosted integration runtime VMs can be deployed inside the same Azure virtual network or a peered virtual network where the account and ingestion private endpoints are deployed.

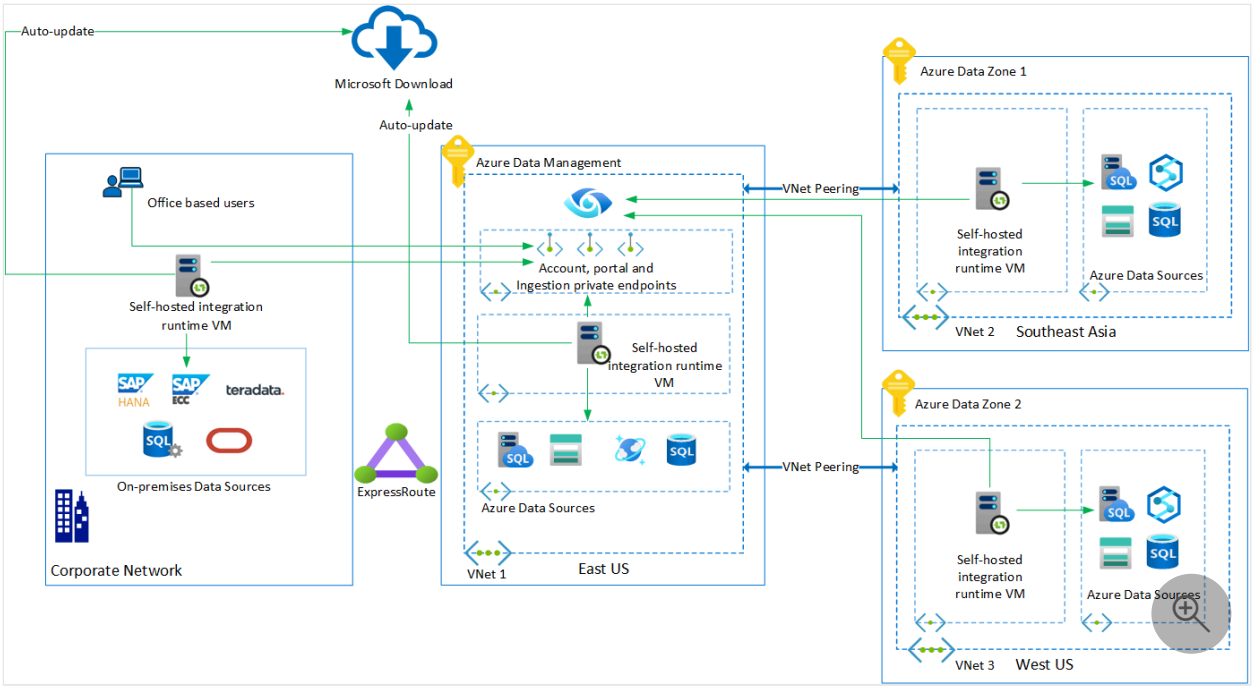


You can optionally deploy another self-hosted integration runtime in the spoke virtual networks.

Multiple regions, multiple virtual networks

If your data sources are distributed across multiple Azure regions in one or more Azure subscriptions, you can use this scenario.

For performance and cost optimization, we highly recommended deploying one or more self-hosted integration runtime VMs in each region where data sources are located.

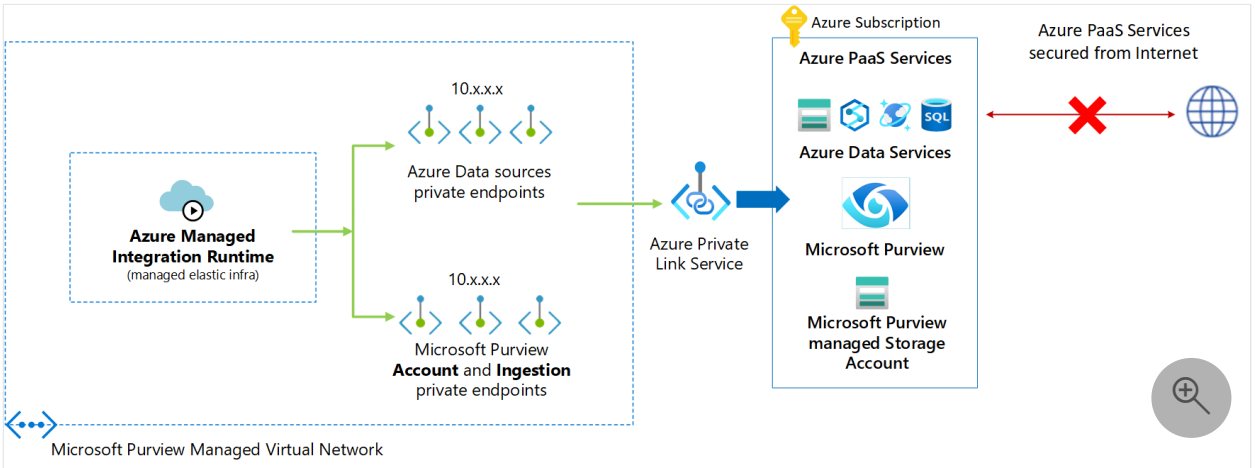


Scan using Managed Vnet Runtime

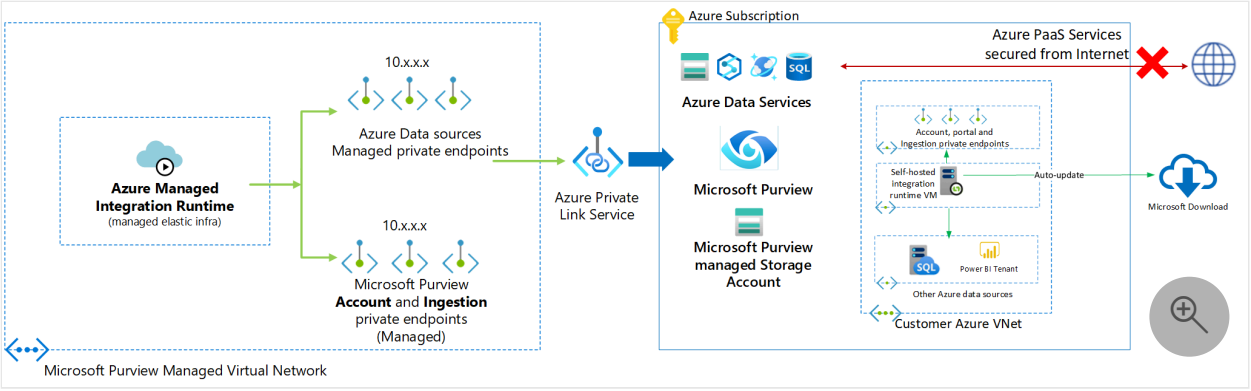
You can use Managed VNet Runtime to scan data sources in a private network. Learn more from Use a Managed VNet with your Microsoft Purview account.

Using Managed VNet Runtime helps to minimize the administrative overhead of managing the runtime and reduce overall scan duration.

To scan any Azure data sources over private network using Managed VNet Runtime, a managed private endpoint must be deployed within Microsoft Purview Managed Virtual Network, even if the data source already has a private network in your Azure subscription.



If you need to scan on-premises data sources or additional data sources in Azure that are not supported by Managed VNet Runtime, you can deploy both Managed VNet Runtime and Self-hosted integration runtime.



If Microsoft Purview isn't available in your primary region

Microsoft Purview is an Azure platform as a service solution. You can deploy a Microsoft Purview account inside your Azure subscription in any supported Azure regions.

If Microsoft Purview isn't available in your primary Azure region, consider the following factors when choosing a secondary region to deploy your Microsoft Purview account:

- Review the latency between your primary Azure region where data sources are deployed and your secondary Azure region, where Microsoft Purview account will be deployed. For more information, see Azure network round-trip latency statistics.
- Review your data residency requirements. When you scan data sources in the Microsoft Purview Data Map, information related to your metadata is ingested and stored inside your data map in the Azure region where your Microsoft Purview account is deployed. For more information, see Where is metadata stored.
- Review your network and security requirements if private network connectivity for user access or metadata ingestion is required. For more information, see If Microsoft Purview isn't available in your primary region.

Option 1: Deploy your Microsoft Purview account in a secondary region and deploy all private endpoints in the primary region, where your Azure data sources are located. For this scenario:

- This is the recommended option, if Australia Southeast is the primary region for all your data sources and you have all network resources deployed in your primary region.
- Deploy a Microsoft Purview account in your secondary region (for example, Australia East).
- Deploy all Microsoft Purview private endpoints including account, portal and ingestion in your primary region (for example, Australia Southeast).
- Deploy all [Microsoft Purview self-hosted integration runtime](../ manage-integration-runtimes.md) VMs in your primary region (for example, Australia Southeast). This helps to reduce cross region traffic as the Data Map scans will

happen in the local region where data sources are located and only metadata is ingested int your secondary region where your Microsoft Purview account is deployed.
- If you use Microsoft Purview Managed VNets for metadata ingestion, Managed VNet Runtime and all managed private endpoints will be automatically deployed in the region where your Microsoft Purview is deployed (for example, Australia East).

Option 2: Deploy your Microsoft Purview account in a secondary region and deploy private endpoints in the primary and secondary regions. For this scenario:

- This option is recommended if you have data sources in both primary and secondary regions and users are connected through the primary region.
- Deploy a Microsoft Purview account in your secondary region (for example, Australia East).
- Deploy Microsoft Purview governance portal private endpoint in the primary region (for example, Australia Southeast) for user access to Microsoft Purview governance portal.
- Deploy Microsoft Purview account and ingestion private endpoints in your primary region (for example, Australia southeast) to scan data sources locally in the primary region.
- Deploy Microsoft Purview account and ingestion private endpoints in your secondary region (for example, Australia East) to scan data sources locally in the secondary region.
- Deploy [Microsoft Purview self-hosted integration runtime](../ manage-integration-runtimes.md) VMs in both primary and secondary regions. This will help to keep data Map scan traffic in the local region and send only metadata to Microsoft Purview Data Map where is configured in your secondary region (for example, Australia East).
- If you use Microsoft Purview Managed VNets for metadata ingestion, Managed VNet Runtime and all managed private endpoints will be automatically deployed in the region where your Microsoft Purview is deployed (for example, Australia East).

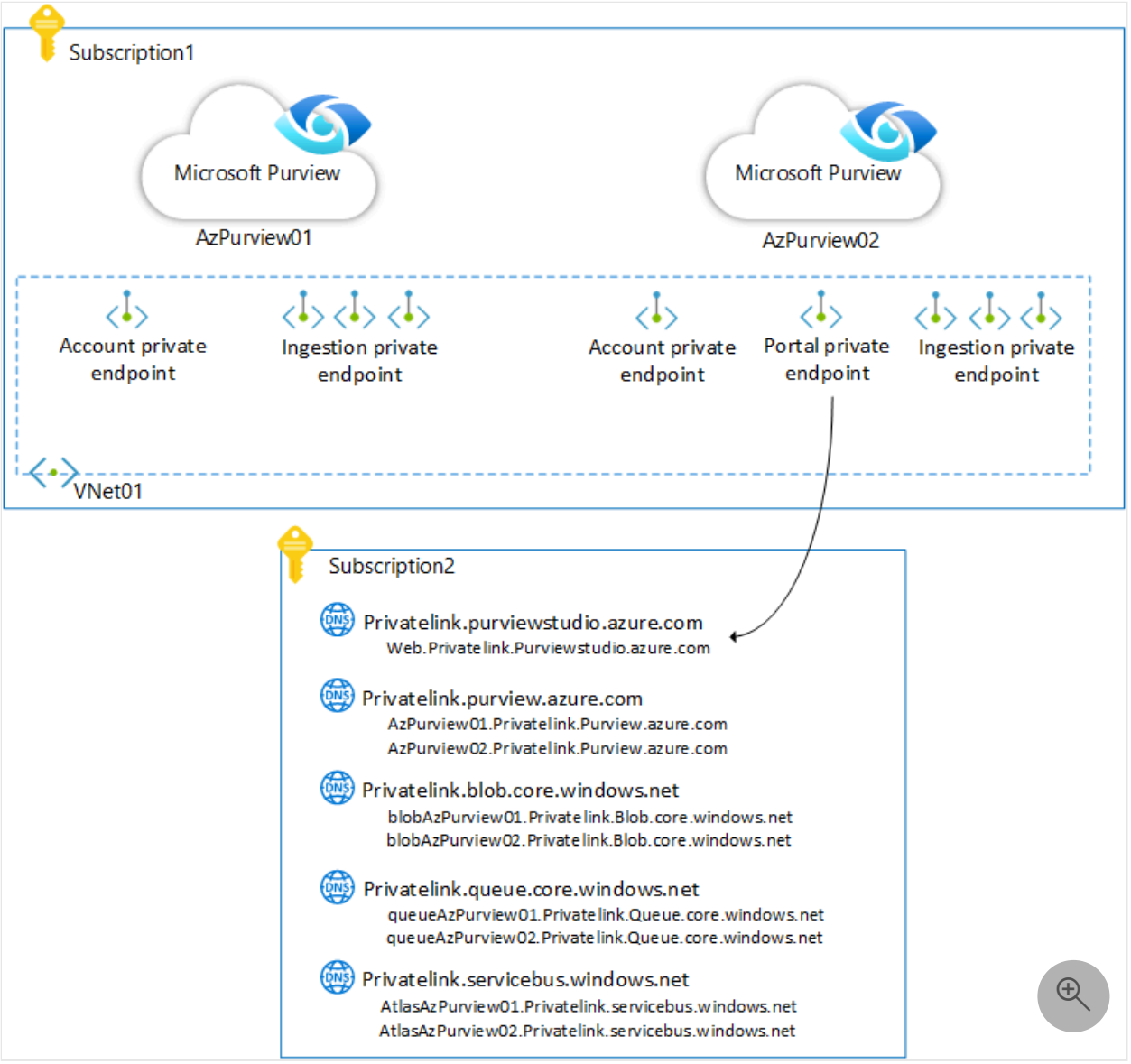
DNS configuration with private endpoints

Name resolution for multiple Microsoft Purview accounts

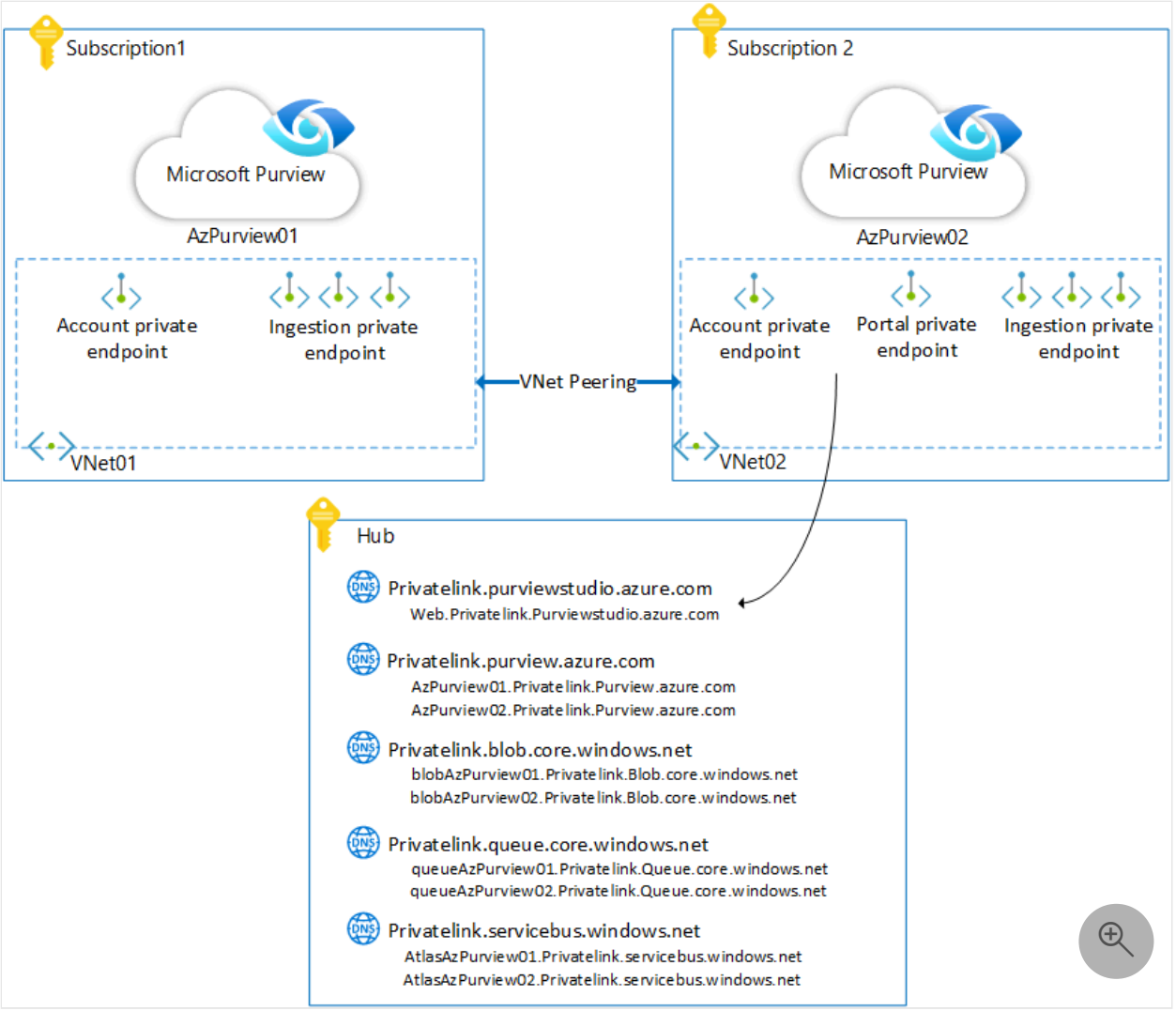
It's recommended to follow these recommendations, if your organization needs to deploy and maintain multiple Microsoft Purview accounts using private endpoints:

1. Deploy at least one *account* private endpoint for each Microsoft Purview account.

2. Deploy at least one set of *ingestion* private endpoints for each Microsoft Purview account.
3. Deploy one *portal* private endpoint for one of the Microsoft Purview accounts in your Azure environments. Create one DNS A record for *portal* private endpoint to resolve `web.purview.azure.com`. The *portal* private endpoint can be used by all purview accounts in the same Azure virtual network or virtual networks connected through VNet peering.



This scenario also applies if multiple Microsoft Purview accounts are deployed across multiple subscriptions and multiple VNets that are connected through VNet peering. *Portal* private endpoint mainly renders static assets related to the Microsoft Purview governance portal, thus, it's independent of Microsoft Purview account, therefore, only one *portal* private endpoint is needed to visit all Microsoft Purview accounts in the Azure environment if VNets are connected.



> **Note**
>
> You may need to deploy separate *portal* private endpoints for each Microsoft Purview account in the scenarios where Microsoft Purview accounts are deployed in isolated network segmentations. Microsoft Purview *portal* is static contents for all customers without any customer information. Optionally, you can use public network, (without portal private endpoint) to launch `web.purview.azure.com` if your end users are allowed to launch the Internet.

Option 3: Use both private and public endpoints

You might choose an option in which a subset of your data sources uses private endpoints, and at the same time, you need to scan either of the following:

- Other data sources that are configured with a service endpoint
- Data sources that have a public endpoint that's accessible through the internet

If you need to scan some data sources by using an ingestion private endpoint and some data sources by using public endpoints or a service endpoint, you can:

1. Use private endpoints for your Microsoft Purview account.
2. Set **Public network access** to **Enabled from all networks** on your Microsoft Purview account.

Integration runtime options

- To scan an Azure data source that's configured with a private endpoint, you need to set up and use a self-hosted integration runtime on a Windows virtual machine that's deployed inside the same or a peered virtual network where Microsoft Purview account and ingestion private endpoints are deployed.

  When you're using a private endpoint with Microsoft Purview, you need to allow network connectivity from data sources to a self-hosted integration VM on the Azure virtual network where Microsoft Purview private endpoints are deployed.

- To scan an Azure data source that's configured to allow a public endpoint, you can use the Azure integration runtime.

- To scan on-premises data sources, you can also install a self-hosted integration runtime on either an on-premises Windows machine or a VM inside an Azure virtual network.

- We recommend allowing automatic upgrade for a self-hosted integration runtime. Make sure you open required outbound rules in your Azure virtual network or on your corporate firewall to allow automatic upgrade. For more information, see Self-hosted integration runtime networking requirements.

Authentication options

- To scan an Azure data source that's configured to allow a public endpoint, you can use any authentication option, based on the data source type.

- If you use an ingestion private endpoint to scan an Azure data source that's configured with a private endpoint:

  - You can't use a Microsoft Purview managed identity. Instead, use a service principal, an account key, or SQL authentication, based on the data source type.

  - Make sure that your credentials are stored in an Azure key vault and registered inside Microsoft Purview.

- You must create a credential in Microsoft Purview based on each secret that you create in Azure Key Vault. At minimum, assign *get* and *list* access for secrets for Microsoft Purview on the Key Vault resource in Azure. Otherwise, the credentials won't work in the Microsoft Purview account.

Option 4: Use private endpoints for ingestion only

You might choose this option if you need to:

- Scan all data sources using ingestion private endpoint.
- Managed resources must be configured to disable public network.
- Enable access to Microsoft Purview governance portal through public network.

To enable this option:

1. Configure ingestion private endpoint for your Microsoft Purview account.
2. Set **Public network access** to **Disabled for ingestion only (Preview)** on your Microsoft Purview account.

Integration runtime options

Follow recommendation for option 2.

Authentication options

Follow recommendation for option 2.

Self-hosted integration runtime network and proxy recommendations

For scanning data sources across your on-premises and Azure networks, you may need to deploy and use one or multiple self-hosted integration runtime virtual machines inside an Azure VNet or an on-premises network, for any of the scenarios mentioned earlier in this document.

- To simplify management, when possible, use Azure IR and Microsoft Purview Managed VNet IR to scan data sources.

- The Self-hosted integration runtime service can communicate with Microsoft Purview through public or private network over port 443. For more information,

see, self-hosted integration runtime networking requirements.

- One self-hosted integration runtime VM can be used to scan one or multiple data sources in Microsoft Purview, however, self-hosted integration runtime must be only registered for Microsoft Purview and can't be used for Azure Data Factory or Azure Synapse at the same time.
- You can register and use one or multiple self-hosted integration runtimes in one Microsoft Purview account. It's recommended to place at least one self-hosted integration runtime VM in each region or on-premises network where your data sources reside.
- It's recommended to define a baseline for required capacity for each self-hosted integration runtime VM and scale the VM capacity based on demand.
- It's recommended to set up network connection between self-hosted integration runtime VMs and Microsoft Purview and its managed resources through private network, when possible.
- Allow outbound connectivity to download.microsoft.com, if auto-update is enabled.
- The self-hosted integration runtime service doesn't require outbound internet connectivity, if self-hosted integration runtime VMs are deployed in an Azure VNet or in the on-premises network that is connected to Azure through an ExpressRoute or Site to Site VPN connection. In this case, the scan and metadata ingestion process can be done through private network.
- Self-hosted integration runtime can communicate Microsoft Purview and its managed resources directly or through a proxy server. Avoid using proxy settings if self-hosted integration runtime VM is inside an Azure VNet or connected through ExpressRoute or Site to Site VPN connection.
- Review supported scenarios, if you need to use self-hosted integration runtime with proxy setting.

Next steps

- Use private endpoints for secure access to Microsoft Purview

---

Feedback

Was this page helpful? Yes No

Manage and increase quotas for classic data governance resources

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

This article highlights the limits that currently exist in the classic Microsoft Purview Data Governance service. These limits are also known as quotas.

> **Note**
>
> These are limits for the classic data governance experience.

Classic Microsoft Purview Data Governance limits

Expand table

Resource	Default limit	Maximum limit
Microsoft Purview accounts per region, per tenant (all subscriptions combined)	3* (dependent)	Contact Support
Data Map throughput There's no default limit on the Data Map metadata storage	10 capacity units 250 operations per second	100 capacity units 2,500 operations per second
vCores available for scanning, per account**	160	160
Concurrent scans per Purview account. The limit is based on the type of data sources scanned*	5	10
Maximum time that a scan can run for	7 days	7 days
Size of assets per account	100M physical assets	Contact support
Maximum size of an asset in a catalog	2 MB	2 MB
Maximum length of an asset name and classification name	4 KB	4 KB
Maximum length of asset property name and value	32 KB	32 KB

Resource	Default limit	Maximum limit
Maximum length of classification attribute name and value	32 KB	32 KB
Maximum number of glossary terms, per glossary	100K	100K
Maximum number of self-service policies, per account	3K	3K
Maximum number of collections	256	256
Maximum ingestion of child objects and contacts	Doesn't apply	20,000 entities

* Accounts limit is dependent on your organization. For full information, see the FAQ.

** Self-hosted integration runtime scenarios aren't included in the limits defined in the above table.

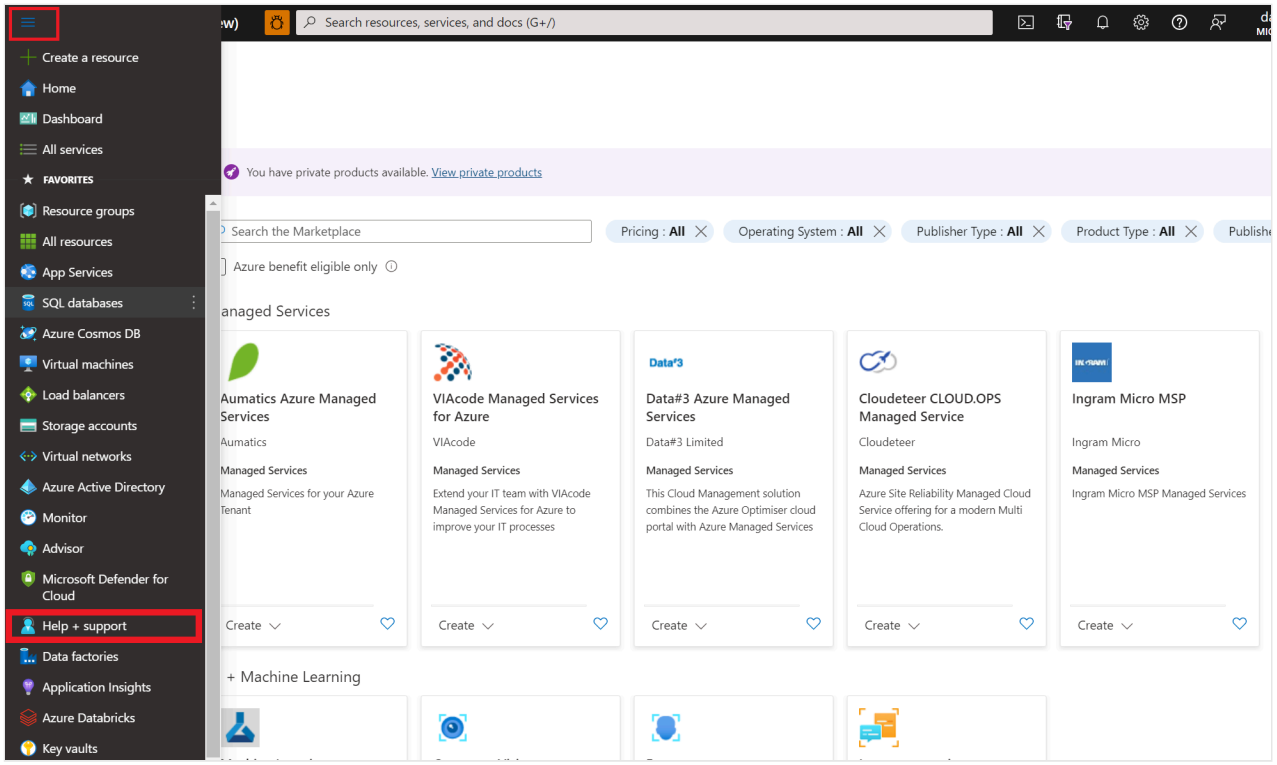
^ Increasing the Data Map throughput limit also increases the minimum number of capacity units with no usage. TO increase Data Map throughput, you can request a capacity until upgrade. See Data Map throughput for more info about Data Map in general.

Request quota increase

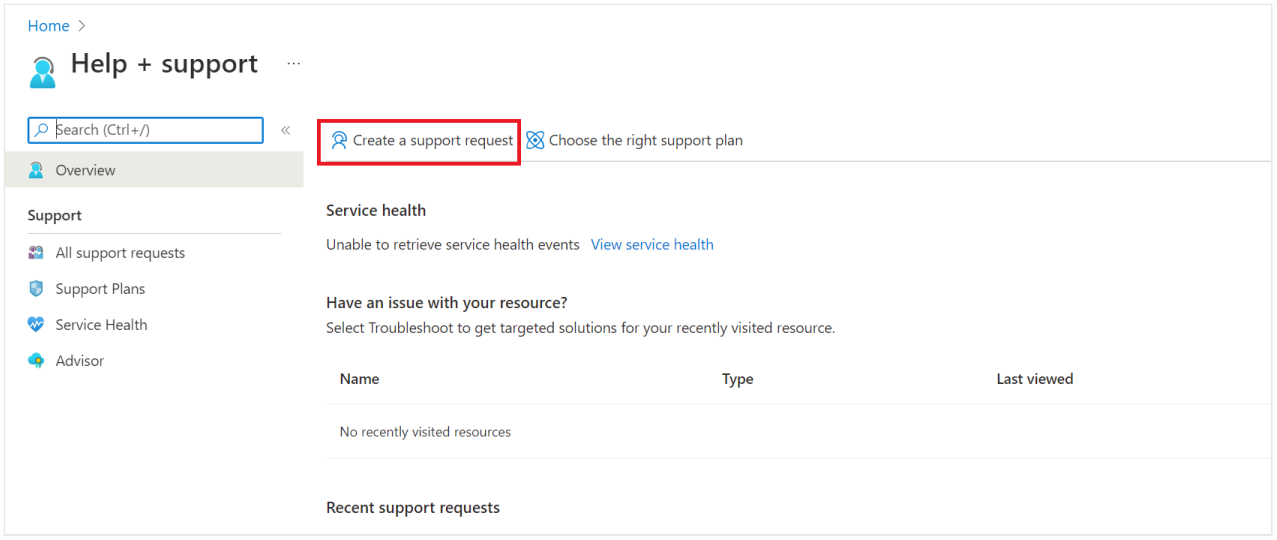
Use the following steps to create a new support request from the Azure portal to increase quota for Microsoft Purview. You can create a quota request for Microsoft Purview accounts in a subscription or tenant (dependent on your organization), and the Data Map throughput (capacity units) of a specific account.

> **⊗ Caution**
>
> Once you have upgraded your Data Map throughput, you can't downgrade.

1. On the Azure portal menu, select **Help + support**.



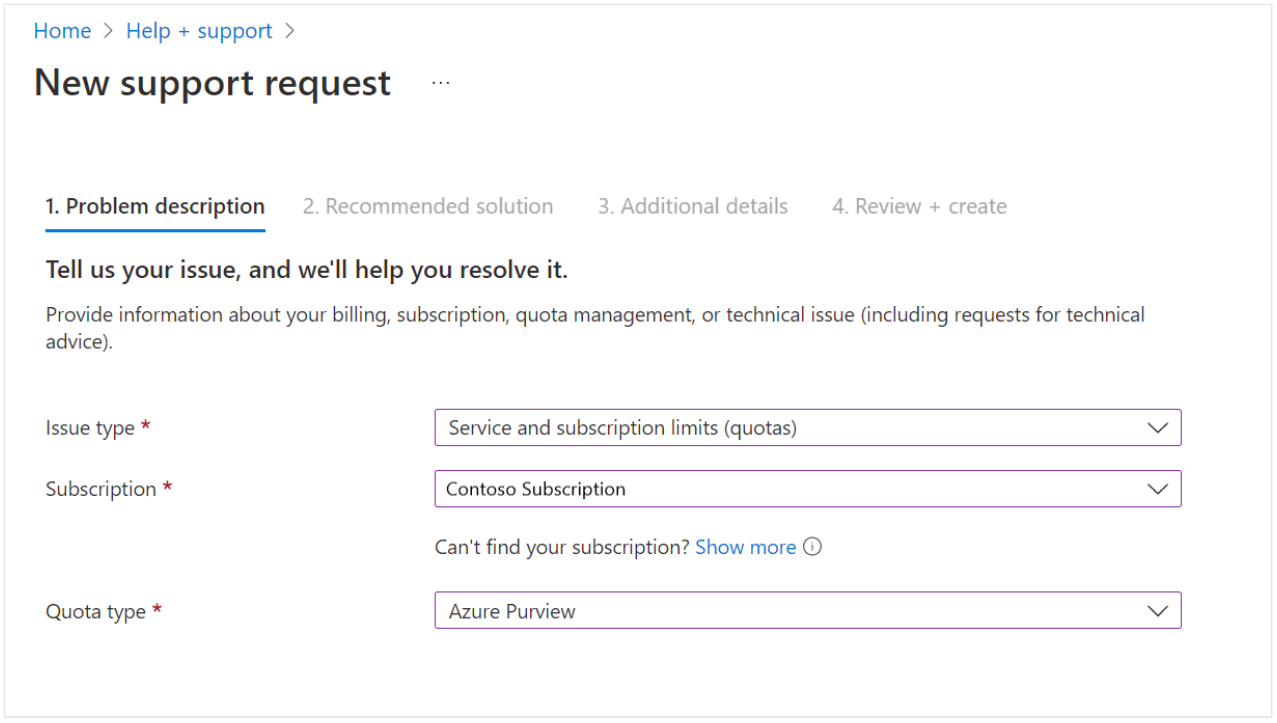
2. In **Help + support**, select **New support request**.



3. For **Issue type**, select **Service and subscription limits (quotas)**.

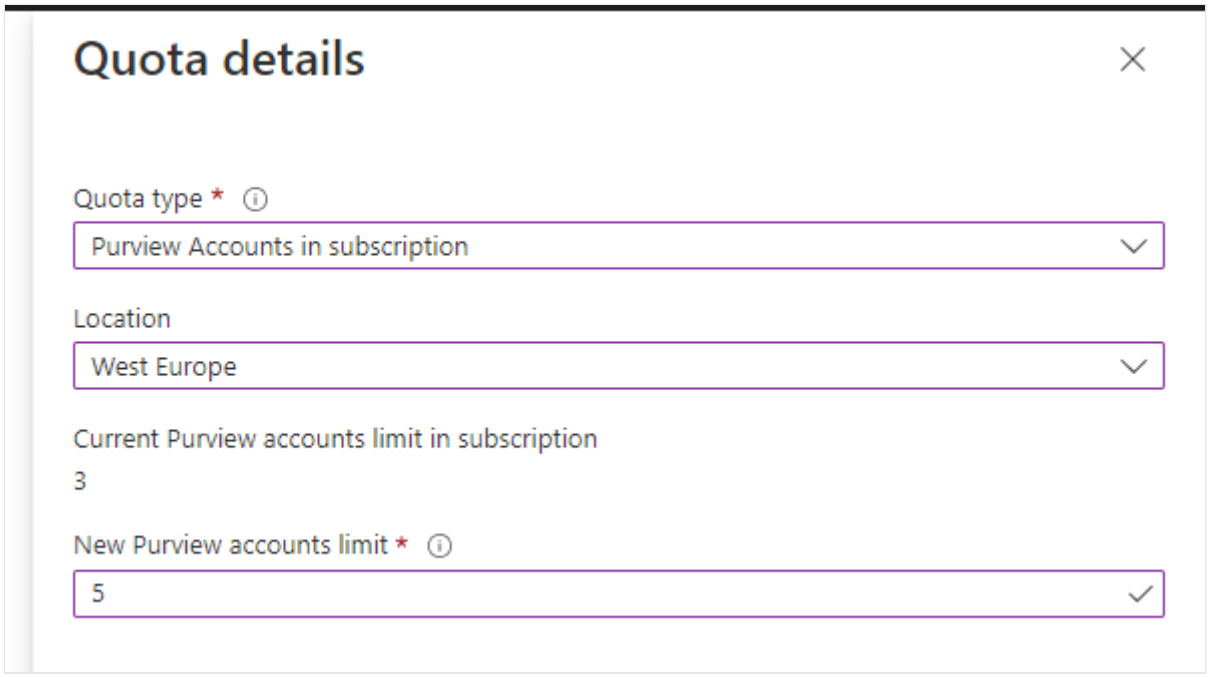
4. For **Subscription**, select the subscription whose quota you want to increase.

5. For **Quota type**, select Microsoft Purview. Then select **Next**.



6. In the **Details** window, select **Enter details** to enter additional information.

7. Choose your **Quota type**, scope (either location or account) and what you wish the new limit to be



8. Enter the rest of the required support information. Review and create the support request

Next steps

Concept: Data Map in Microsoft Purview

Tutorial: Scan data with Microsoft Purview

Tutorial: Navigate the home page and search for an asset

Tutorial: Browse assets and view their lineage

Last updated on 12/16/2025

Frequently asked questions (FAQ) about Microsoft Purview data governance solutions

Overview

Many organizations lack a holistic understanding of their data. It's challenging to understand what data exists, where data is located, and how to find and access relevant data. Data lacks context such as lineage, classification, and comprehensive metadata, making it difficult for business users to search for the right data and use that data appropriately. As a result, only a small fraction of collected data is used to inform business decisions. Finally, identifying data security issues and protecting sensitive data is inconsistent. It requires ongoing time and effort, especially while maintaining data agility.

Microsoft Purview data governance solutions help customers gain deep knowledge of all their data while maintaining control over its use. With our tools, organizations discover and curate data. They gain insights into their data estate and centrally govern access to data.

Purpose of this FAQ

This FAQ answers common questions that customers and field teams often ask about Microsoft Purview data governance solutions: Microsoft Purview Data Map and Microsoft Purview Unified Catalog.

Where can I submit feedback or feature requests?

If you have an issue with Microsoft Purview Data Map or Unified Catalog, reach out to support.

If you have feedback about the documentation, reach out using the feedback button at the bottom of every documentation page.

What are the source types available for metadata scanning and classification?

There are many sources available, and we're adding more all the time. For a full list of sources and all their available capabilities, see our supported data sources article.

Does the Microsoft Purview Data Map support scanning zipped files?

Currently, the Microsoft Purview Data Map supports scanning GZIP files that contain a single CSV. For more information, see supported file types for scanning.

Do scans in the Microsoft Purview Data Map impact operational database performance?

It depends. The amount of load on a database will be dependent on the number of assets being scanned, and complexity of the table schema (like number of columns, etc.). The Microsoft Purview Data Map only samples a subset of the data to determine classification. For sampling related details, see the supported data sources documentation. It's recommended to schedule scans outside of normal business hours. For scanning best practices, see the scanning best practices documentation.

Can I customize the number of rows scanned?

No. Number of rows scanned is determined by the data source being scanned. For more information, see our sampling documentation.

What data systems/processors can we connect and get lineage?

Our lineage offering updates often. For a full list of systems, see our lineage user guide.

How are Azure Unified Catalog (ADC) Gen 2, Azure Information Protection, and Microsoft Purview (formerly Azure Purview) related?

Microsoft Purview (formerly Azure Purview) originally began as ADC Gen 2 but has since broadened in scope. It now embraces the advanced catalog capabilities of ADC Gen 2 combined

with the data classification, labeling, and compliance policy enforcement capabilities of Azure Information Protection. Microsoft Purview now houses compliance and data governance tools to manage your full data estate. For more information, see our overview documentation.

What happens to customers using ADC Gen 1?

Microsoft Purview is the focus of all product innovation in the catalog solution space for Microsoft. ADC Gen 1 is being retired in 2024.

How do I migrate existing ADC Gen 1 data assets to Microsoft Purview?

Use the Microsoft Purview APIs to extract from ADC Gen 1 and ingest into Microsoft Purview. For the glossary, we support bulk tools based on CSV.

How do I encrypt sensitive data for SQL tables using Microsoft Purview?

Data encryption is done at the data source level. Microsoft Purview stores only the metadata. It doesn't preview data.

What's the difference between a glossary and classification?

A glossary uses a naming convention followed by non-technical/business users of the data, also known as data consumers. These types of people are business analysts or data scientists who use Microsoft Purview to search for certain types of data, based on business usage. For instance, supply chain analysts might need to search for the terms *SKU types* and *shipment details*. They search the glossary for these terms to find relevant data. Classification is a tag applied to a data asset at the table, column, or file level that identifies what data exists in the asset. Classification can be applied automatically or manually, based on the type of data found. Typically, you use classification tags to identify whether an asset contains sensitive data, and what type of sensitive data that might be.

Can the Microsoft Purview Data Map scan and classify emails, PDFs etc. in my SharePoint and OneDrive?

Scanning for on-premises SharePoint sites and libraries is provided through the Microsoft Purview Information Protection scanner. The scanner is available for use through a customer's Microsoft 365 subscription with the following SKUs: AIP P1, EMS E3, and Microsoft 365 E3. If you have any one of these SKUs, you should have the right entitlements to start using the Microsoft Purview Information Protection scanner. However, currently, the Microsoft Purview Data Map doesn't support scanning SharePoint and OneDrive into the Microsoft Purview Unified Catalog. Follow these links for supported file types for the Microsoft Purview Information Protection scanner, and supported file types for the Microsoft Purview Unified Catalog.

What is the compute used for the scan?

There's a Microsoft-managed scanning infrastructure. For most Azure/AWS resources that we support, you don't need to deploy a scanning infrastructure.

Is there a way to create a Microsoft Purview account programatically? For example, using Azure Resource Manager (ARM) template / CLI / PowerShell?

Yes, PowerShell commands, a .NET script, a Python script, an ARM template, and a Bicep template are all available.

I'm already using Atlas, can I easily move to Microsoft Purview?

Microsoft Purview is compatible with Atlas API. If you're migrating from Atlas, it's recommended to scan your data sources first using Microsoft Purview. Once the assets are available in your account, you can use similar Atlas APIs to integrate such as updating assets or adding custom lineage. Microsoft Purview modifies the Search API to use Azure Search so you should be able to use Advance Search.

Can I register multiple tenants within a single Microsoft Purview account?

No, currently in order to scan another tenant's data source, you need to create a separate Microsoft Purview account in that tenant.

Does Microsoft Purview support column level lineage?

Yes, Microsoft Purview supports column level lineage.

Does Microsoft Purview support Soft-Delete?

Yes, Microsoft Purview supports Soft Delete for Azure subscription status management perspective. Microsoft Purview can read subscription states (disabled/warned etc.) and put the account in soft-delete state until the account is restored/deleted. All the data plane API calls will be blocked when the account is in soft delete state and only GET/DELETE control plane API calls will be allowed. You can find additional information in Azure subscription states page Azure Subscription Status.

Does Microsoft Purview Data Map support data loss prevention (DLP) capabilities?

No, the data map doesn't currently provide the data loss prevention capabilities that are supported for Microsoft 365 apps and services.

Read about Data Loss Prevention in Microsoft Purview Information Protection if you're interested in data loss prevention for Microsoft 365 apps and services.

Can I use the same self-hosted integration runtime (SHIR) for both Microsoft Purview and Azure Data Factory?

No. For more information about the SHIR, see our self-hosted integration runtime documentation.

How do I use the REST API for Microsoft Purview?

You can follow our documentation for examples: using the REST API to access the Unified Catalog, and using the REST API to manage access control. We also have a tutorial for using the Python SDK, and Atlas 2.2 APIs.

What network, security, and firewall configurations are recommended for Microsoft Purview?

For network, security, and firewall configuration recommendations, see our security best practices documentation.

Last updated on 07/23/2025

Microsoft Purview (formerly Azure Purview) deployment checklist

Article • 10/28/2024

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

> **Tip**
>
> Microsoft Purview has a new experience! If you're a new Microsoft Purview customer, or if you want more information, see our article on the new portal or the article about our new data governance experience.
>
> If you're planning on deploying Microsoft Purview governance solutions in your organization, use our new portal and check out our data governance quickstart and best practices.

This article lists prerequisites that help you get started quickly on planning and deployment for your Microsoft Purview (formerly Azure Purview) account.

If you're creating a plan to deploy Microsoft Purview, and also want to consider best practices as you develop your deployment strategy, then use our deployment best practices guide to get started.

If you're looking for a strictly technical deployment guide, this deployment checklist is for you.

Expand table

No.	Prerequisite / Action	Required permission	More guidance and recommendations
1	Microsoft Entra tenant	N/A	An Microsoft Entra tenant should be associated with your subscription. • *Information Protection Administrator* role is required, if you plan to extend Microsoft 365 Sensitivity Labels to

No.	Prerequisite / Action	Required permission	More guidance and recommendations
			the Microsoft Purview Data Map for files and db columns • *Power BI Administrator* role is required, if you're planning to scan Power BI tenants.
2	An active Azure Subscription	*Subscription Owner*	An Azure subscription is needed to deploy Microsoft Purview and its managed resources. If you don't have an Azure subscription, create a free subscription before you begin.
3	Define whether you plan to deploy a Microsoft Purview with a managed event hub	N/A	You can choose to deploy configure an existing Event Hubs namespace during Microsoft Purview account creation, see Microsoft Purview account creation. With this managed namespace you can publish messages to the event hub kafka topic ATLAS_HOOK and Microsoft Purview will consume and process it. Microsoft Purview will notify entity changes to the event hub kafka topic ATLAS_ENTITIES and user can consume and process it. You can enable or disable this feature any time after account creation.
4	Register the following resource providers: • Microsoft.Storage • Microsoft.EventHub (optional) • Microsoft.Purview	*Subscription Owner* or custom role to register Azure resource providers (*/register/action*)	Register required Azure Resource Providers in the Azure Subscription that is designated for Microsoft Purview Account. Review Azure resource provider operations.
5	Update Azure Policy to allow deployment of the following resources in your Azure subscription: • Microsoft Purview • Azure Storage	*Subscription Owner*	Use this step if an existing Azure Policy prevents deploying such Azure resources. If a blocking policy exists and needs to remain in place, follow our Microsoft Purview exception tag guide and follow the steps to create an exception for Microsoft Purview accounts.
6	Define your network security requirements.	Network and Security architects.	• Review Microsoft Purview network architecture and best practices to define what scenario is more relevant to your network requirements.

No.	Prerequisite / Action	Required permission	More guidance and recommendations
			• If private network is needed, use Microsoft Purview Managed IR to scan Azure data sources when possible to reduce complexity and administrative overhead.
7	An Azure Virtual Network and Subnets for Microsoft Purview private endpoints.	*Network Contributor* to create or update Azure virtual network.	Use this step if you're planning to deploy private endpoint connectivity with Microsoft Purview: • Private endpoints for **Ingestion**. • Private endpoint for Microsoft Purview **Account**. • Private endpoint for Microsoft Purview **Portal**. Deploy Azure Virtual Network if you need one.
8	Deploy private endpoint for Azure data sources.	*Network Contributor* to set up private endpoints for each data source.	Perform this step, if you're planning to use Private Endpoint for Ingestion.
9	Define whether to deploy new or use existing Azure Private DNS Zones.	Required Azure Private DNS Zones can be created automatically during Purview Account deployment using Subscription Owner / Contributor role	Use this step if you're planning to use Private Endpoint connectivity with Microsoft Purview. Required DNS Zones for Private Endpoint: • privatelink.purview.azure.com • privatelink.purview-service.microsoft.com • privatelink.purviewstudio.azure.com • privatelink.blob.core.windows.net • privatelink.queue.core.windows.net • privatelink.servicebus.windows.net
10	A management machine in your CorpNet or inside Azure virtual network to launch the Microsoft Purview governance portal.	N/A	Use this step if you're planning to set **Allow Public Network** to **deny** on your Microsoft Purview Account.

No.	Prerequisite / Action	Required permission	More guidance and recommendations
11	Deploy a Microsoft Purview Account	Subscription Owner / Contributor	Purview account is deployed with one Capacity Unit and will scale up based on demand.
12	Deploy a Managed Integration Runtime and Managed private endpoints for Azure data sources.	*Data source admin* to set up Managed virtual network inside Microsoft Purview. *Network Contributor* to approve managed private endpoint for each Azure data source.	Perform this step if you're planning to use Managed virtual network. within your Microsoft Purview account for scanning purposes.
13	Deploy Self-hosted integration runtime VMs inside your network.	Azure: *Virtual Machine Contributor* On-premises: Application owner	Use this step if you're planning to perform any scans using Self-hosted Integration Runtime.
14	Create a Self-hosted integration runtime inside Microsoft Purview.	Data curator VM Administrator or application owner	Use this step if you're planning to use Self-hosted Integration Runtime instead of Managed Integration Runtime or Azure Integration Runtime. download
15	Register your Self-hosted integration runtime	Virtual machine administrator	Use this step if you have **on-premises** or **VM-based data sources** (for example, SQL Server). Use this step are using **Private Endpoint** to scan to **any** data sources.
16	Grant Azure RBAC **Reader** role to **Microsoft Purview MSI** at data sources' Subscriptions	*Subscription owner* or *User Access Administrator*	Use this step if you're planning to register multiple or **any** of the following data sources: • Azure Blob Storage • Azure Data Lake Storage Gen1 • Azure Data Lake Storage Gen2 • Azure SQL Database

No.	Prerequisite / Action	Required permission	More guidance and recommendations
			• Azure SQL Managed Instance • Azure Synapse Analytics
17	Grant Azure RBAC **Storage Blob Data Reader** role to **Microsoft Purview MSI** at data sources Subscriptions.	*Subscription owner* or *User Access Administrator*	**Skip** this step if you're using Private Endpoint to connect to data sources. Use this step if you have these data sources: • Azure Blob Storage • Azure Data Lake Storage Gen2
18	Enable network connectivity to allow AzureServices to access data sources: for example, Enable "**Allow trusted Microsoft services to access this storage account**".	*Owner* or *Contributor* at Data source	Use this step if **Service Endpoint** is used in your data sources. (Don't use this step if Private Endpoint is used)
19	Enable **Microsoft Entra authentication** on **Azure SQL Servers**, **Azure SQL Managed Instance** and **Azure Synapse Analytics**	Azure SQL Server Contributor	Use this step if you have **Azure SQL DB** or **Azure SQL Managed Instance** or **Azure Synapse Analytics** as data source.
20	Grant **Microsoft Purview MSI** account with **db_datareader** role to Azure SQL databases and Azure SQL Managed Instance databases	Azure SQL Administrator	Use this step if you have **Azure SQL DB** or **Azure SQL Managed Instance** as data source. **Skip** this step if you're using **Private Endpoint** to connect to data sources.
21	Grant Azure RBAC **Storage Blob Data Reader** to **Synapse SQL Server** for staging Storage Accounts	Owner or User Access Administrator at data source	Use this step if you have **Azure Synapse Analytics** as data sources. **Skip** this step if you're using Private Endpoint to connect to data sources.
22	Grant Azure RBAC **Reader** role to **Microsoft Purview MSI** at **Synapse workspace** resources	Owner or User Access Administrator at data source	Use this step if you have **Azure Synapse Analytics** as data sources. **Skip** this step if you're using Private Endpoint to connect to data sources.
23	Grant Azure **Purview MSI account** with **db_datareader** role	Azure SQL Administrator	Use this step if you have **Azure Synapse Analytics (Dedicated SQL databases)**.

No.	Prerequisite / Action	Required permission	More guidance and recommendations
			Skip this step if you're using **Private Endpoint** to connect to data sources.
24	Grant **Microsoft Purview MSI** account with **sysadmin** role	Azure SQL Administrator	Use this step if you have Azure Synapse Analytics (Serverless SQL databases). **Skip** this step if you're using **Private Endpoint** to connect to data sources.
25	Create an app registration or service principal inside your Microsoft Entra tenant	Microsoft Entra ID *Application Administrator*	Use this step if you're planning to perform a scan on a data source using Delegated Author Service Principal.
26	Create an **Azure Key Vault** and a **Secret** to save data source credentials or service principal secret.	*Contributor* or *Key Vault Administrator*	Use this step if you have **on-premises** or **VM-based data sources** (for example, SQL Server). Use this step are using **ingestion private endpoints** to scan a data source.
27	Grant Key **Vault Access Policy** to Microsoft Purview MSI: **Secret: get/list**	*Key Vault Administrator*	Use this step if you have **on-premises** / **VM-based data sources** (for example, SQL Server) Use this step if **Key Vault Permission Model** is set to Vault Access Policy.
28	Grant **Key Vault RBAC role** Key Vault Secrets User to Microsoft Purview MSI.	*Owner* or *User Access Administrator*	Use this step if you have **on-premises** or **VM-based data sources** (for example, SQL Server) Use this step if **Key Vault Permission Model** is set to Azure role-based access control.
29	Create a new connection to Azure Key Vault from the Microsoft Purview governance portal	*Data source admin*	Use this step if you're planning to use any of the following authentication options to scan a data source in Microsoft Purview: • Account key • Basic Authentication • Delegated Auth • SQL Authentication • Service Principal • Consumer Key
30	Deploy a private endpoint for Power BI tenant	*Power BI Administrator* *Network contributor*	Use this step if you're planning to register a Power BI tenant as data source and your Microsoft Purview account is set to **deny public access**.

No.	Prerequisite / Action	Required permission	More guidance and recommendations
			For more information, see How to configure private endpoints for accessing Power BI.
31	Connect Azure Data Factory to Microsoft Purview from Azure Data Factory Portal. **Manage** -> **Microsoft Purview**. Select **Connect to a Purview account**. Validate if Azure resource tag **catalogUri** exists in ADF Azure resource.	Azure Data Factory Contributor / Data curator	Use this step if you have **Azure Data Factory**.
32	Verify if you have at least one **Microsoft 365 required license** in your Microsoft Entra tenant to use sensitivity labels in Microsoft Purview.	Microsoft Entra ID *Global Reader*	Perform this step if you're planning to extend **sensitivity labels to Microsoft Purview Data Map** For more information, see licensing requirements to use sensitivity labels on files and database columns in Microsoft Purview
33	Consent "**Extend labeling to assets in Microsoft Purview Data Map**"	Compliance Administrator Azure Information Protection Administrator	Use this step if you're interested in extending sensitivity labels to your data in the data map. For more information, see Labeling in the Microsoft Purview Data Map.
34	Create new collections and assign roles in Microsoft Purview	*Collection admin*	Create a collection and assign permissions in Microsoft Purview.
36	Govern Data Sources in Microsoft Purview	*Data Source admin* *Data Reader* or *Data Curator*	For more information, see supported data sources and file types
35	Grant access to data roles in the organization	*Collection admin*	Provide access to other teams to use Microsoft Purview: • Data curator • Data reader • Collection admin • Data source admin • Policy Author • Workflow admin

No.	Prerequisite / Action	Required permission	More guidance and recommendations
			For more information, see Access control in Microsoft Purview.

Next steps

- Review Microsoft Purview deployment best practices

Feedback

Was this page helpful? Yes No

Feature availability for Microsoft Purview

07/15/2025

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

Learn what which Microsoft Purview features are partially available or unavailable in Azure Government and Azure China 21Vianet regions.

In the list of global Azure regions, there are several regions that serve specific markets in addition to the public cloud regions. Microsoft Purview is deployed into the following regions.

- Azure China 21Vianet region: China North 3.
- Azure Government regions: US GOV Virginia

The information in the rest of this document provides information on the feature availability for Microsoft Purview. To see more detailed information about a feature's regional support, select the link to that feature.

Azure China 21Vianet cloud

Feature support is represented by the following symbols:

Expand table

Symbol	Meaning	Description
❌	Unsupported	Unsupported means that the feature isn't available in that particular region.
✅	Supported	Supported means that the feature is available in that particular region and in no way differs from the public cloud offering.
⚠	Partial support	Partial support indicates that the feature offered for the public cloud is available, but with perhaps some modifications. See the guidance and recommended sections for more information about partially supported features.

Expand table

Feature	Subfeature	Public cloud status	CH-North-3
Microsoft Purview Data Map		GA	⚠
	Register and scan Azure Sources	GA	✅
	Register and scan on-premises Sources	GA	✅
	Register and scan multicloud sources	GA	⚠
	Metamodel	Public preview	✅
Microsoft Purview Data Catalog		GA	⚠
	Search and browse	GA	✅
	Bulk edit	GA	✅
	Asset Certification	GA	✅
	Managed attributes	Public preview	❌
	Business glossary	GA	✅
	Lineage	GA	✅
	Classifications	GA	✅
	Labeling	Public preview	❌
	Workflows	Public preview	✅
Data sharing		Public preview	❌
Microsoft Purview Data Estate Insights		GA	⚠
Microsoft Purview Data Policies		GA	✅
	DevOps policies	GA	✅
	Self Service policies	Public preview	✅
	Data owner policies	Public preview	✅

Azure China guidance and recommendations

Microsoft Purview Data Map

The Microsoft Purview Data Map is fully supported, except for some multicloud sources.

Sources that are **NOT** supported in Azure China are:

Expand table

Category	Source
Microsoft	Azure Data Lake Store Gen 1
	Azure Machine Learning
Database	Amazon RDS
	SQL Server on Azure-Arc
File	Amazon S3
Services and apps	Looker

For lineage, Azure Synapse Analytics pipeline lineage is not supported.

Microsoft Purview Data Estate Insights

Data Estate Insights is fully available except for sensitivity label insights, as sensitivity labels aren't supported.

Microsoft Purview Workflows

Workflows are available in preview. The Supplemental Terms of Use for Microsoft Azure Previews include other legal terms that apply to Azure features that are in beta, in preview, or otherwise not yet released into general availability.

Microsoft Purview Data policies

Only policies for Azure SQL Database are currently supported.

Azure Government

Feature support is represented by the following symbols:

Expand table

Symbol	Meaning	Description
❌	Unsupported	Unsupported means that the feature isn't available in that particular region.
✅	Supported	Supported means that the feature is available in that particular region and in no way differs from the public cloud offering.
⚠	Partial support	Partial support indicates that the feature offered for the public cloud is available, but with perhaps some modifications. See the guidance and recommended sections for more information about partially supported features.

Expand table

Feature	Subfeature	Public cloud status	US Gov Virginia
Microsoft Purview Data Map		GA	⚠
	Register and scan Azure Sources	GA	✅
	Register and scan on-premises Sources	GA	✅
	Register and scan multicloud sources	GA	⚠
	Metamodel	Public preview	✅
Microsoft Purview Data Catalog		GA	⚠
	Search and browse	GA	✅
	Bulk edit	GA	✅
	Asset Certification	GA	✅
	Managed attributes	Public preview	❌
	Business glossary	GA	✅
	Lineage	GA	✅
	Classifications	GA	✅
	Labeling	Public preview	❌
	Workflows	Public preview	✅
Data sharing		Public preview	❌

Feature	Subfeature	Public cloud status	US Gov Virginia
Microsoft Purview Data Estate Insights		GA	⚠
Microsoft Purview Data Policies		GA	✅
	DevOps policies	GA	✅
	Self Service policies	Public preview	✅
	Data owner policies	Public preview	✅

Azure Government guidance and recommendations

Microsoft Purview Data Map

The Microsoft Purview Data Map is fully supported, except for some multicloud sources.

Sources that are **NOT** supported in Azure Government are:

Expand table

Category	Source
Microsoft	Power BI (Though Power BI GCC **is** supported.)
Database	Amazon RDS
File	Amazon S3

> **Note**
>
> Any sources that are in preview are not supported. For lineage, Azure Synapse Analytics pipeline lineage is not supported.

Microsoft Purview Data Estate Insights

Data Estate Insights is fully available except for sensitivity label insights, as sensitivity labels aren't supported.

Microsoft Purview Workflows

Workflows are available in preview. The Supplemental Terms of Use for Microsoft Azure Previews include other legal terms that apply to Azure features that are in beta, in preview, or otherwise not yet released into general availability.

Microsoft Purview Data policies

Only policies for Azure SQL Database are currently supported.

Data governance known issues

This article provides information about known issues when using Microsoft Purview data governance solutions.

Browser issues

To learn about browser issues that might affect use of data governance solutions, see these resources:

- Microsoft Edge known issues
- Google Chrome Local Network Access

If you need further assistance, contact Microsoft support.

Unified Catalog management issues

Expand table

Issue ID	Description	Publish Date	Status	Fixed Date
1	Data assets from top-level hierarchies aren't selectable. While adding data assets to data products, assets that are part of a semantic model, or Power BI dataset, hierarchy are automatically expanded without getting user input on which ones to include.	9/6/2024	Active	
2	Missing duplication checks. All business concepts except governance domains are missing duplication checks during creation.	9/6/2024	Active	
3	Incomplete workflow then Privacy reviewer is selected. Privacy reviewer selection skips the last step of showing data assets table for recording asset level access provisioning status and instructions to data consumer.	9/6/2024	Active	
4	Inconsistent asset details page experience. In the asset discovery experience under **Discovery** > **Data assets**, the asset details page for assets within the	10/25/2024	Active	

Issue ID	Description	Publish Date	Status	Fixed Date
	default/merged account render in the new Microsoft Purview portal but assets within accounts that were merged to the default account render in the classic governance portal.			
5	Usage monitoring (preview) for governed assets in Unified Catalog settings has a gap for some customers in the East US region. Values for April 12, 13, and 14, 2026, won't contain the number of governed assets on those days and will appear blank.	4/14/2026	Active	

Discovery issues

Expand table

Issue ID	Title	Publish Date	Status	Fixed Date
1	Data lineage on a data asset in a data product isn't clickable to expand. To view the full lineage of the asset, it must be expanded at the asset in view details first then refreshed and the data product level.	5/9/2025	Active	
2	Enterprise glossary filters don’t interact. Filters in the enterprise glossary don't automatically update based on other filter selections.	9/6/2024	Fixed	9/26/2024
3	Data asset metadata don’t autorefresh. Data assets within data products don’t automatically reflect owner, description, or lineage updates made in the Data Map. They need to be manually refreshed in the **Catalog management** > **Data products** section by selecting **Refresh** in the asset details page.	9/6/2024	Active	
4	Linked objectives and key results (OKRs) don't show in data product search. Published OKRs that are linked to data products don't display in the data product details page within the **Discovery** > **Data products** search experience today.	9/6/2024	Active	

Health management issues

Expand table

Issue ID	Title	Publish Date	Status	Fixed Date
1	Virtual network isn't supported for Google BigQuery	9/1/2024	Active	7/2/2025: Virtual network is supported for Azure SQL, Synapse, Fabric, Azure Databricks, and Snowflake
2	No SQL query is generated for custom data quality rules failures. SQL queries are only generated for failures of out-of-the-box rules.	4/11/2024	Active	
3	Data quality profiling uses a random sample of 1 million records. Users can include a maximum of 50 columns in a single batch, but they can run multiple batches if they need to profile files or tables with more than 50 columns.	4/11/2024	Active	There's no limit for rule based data quality scan; customers can run a data quality scan for billions of records and hundreds of columns.
4	A table lookup rule can be created for the tables within the same governance domain. You need to use four characters when you search for a data asset to reference to look up.	4/11/2024	Active	
5	The freshness data quality rule isn't supported for Azure SQL, Google BigQuery, Databricks Unity Catalog, Snowflake, and Synapse.	9/1/2024	Active	
6	Azure SQL doesn't support RLIKE keywords, so any query containing RLIKE keywords won't run in Azure SQL. Custom regex rule won't work for Azure SQL.	4/11/2024	Active	
7	Fabric OneLake tables and semantic model data quality scanning isn't supported. Only Fabric Lakehouse tables data quality scanning is supported (private preview).	9/1/2024	Active	
8	Scheduled jobs have a maximum limit of 30 data assets. Customers need to set up multiple schedules if	9/1/2024	Active	

Issue ID	Title	Publish Date	Status	Fixed Date
	they have more than 30 data assets in a governance domain.			
9	Data quality threshold alerts can be set up for a maximum of 30 data assets in one batch. If customers have more than 30 data assets in the same governance domain, they need to create multiple alerts, for example: Alert 1 for first 30 data assets and Alert 2 for next 30 data asset, and so on.	9/1/2024	Active	
10	Reports show all records (metadata), no filter available for role-based access control (RBAC). This means a data steward from one org (part of collection 1) can view analytics and reports from another org (collection 2). Recommending to use Purview metadata from OneLake to create own reports and insights.	9/1/2024	Active	
11	Not support N arbitrary Hierarchies of Directories with Parquet Files. We advise users to present data in (1) A directory with Parquet Part File. For example: `./Sales/{Parquet Part Files}` or (2) A directory with Partitioned Parquet Files, partitioned by columns within the dataset like sales data partitioned by year and month. For example: `./Sales/{Year=2018}/{Month=Dec}/{Parquet Part Files}`.	9/1/2024	Active	

Last updated on 04/15/2026

Reliability in Microsoft Purview

Summarize this article for me

This article describes reliability support in Purview for governance experiences. It covers both regional resiliency with availability zones and disaster recovery and business continuity. For a more detailed overview of reliability principles in Microsoft Azure, see Azure reliability.

Availability zone support

Availability zones are physically separate groups of datacenters within an Azure region. When one zone fails, services can fail over to one of the remaining zones.

Microsoft Purview makes commercially reasonable efforts to support zone-redundant availability zones, where resources automatically replicate across zones, without any need for you to set up or configure.

Prerequisites

- Microsoft Purview governance experience currently provides partial availability zone support in a limited number of regions. This partial availability zone support covers experiences and certain functionalities within an experience.
- Zone availability might or might not be available for Microsoft Purview governance experiences or features and functionalities that are in preview.

Supported regions

The table below lists availability zone support:

Expand table

Region	Microsoft Purview Data Map	Scan	Policy	Insights
Southeast Asia		✅	✅	✅
East US		✅	✅	✅
Australia East	✅	✅	✅	✅
West US 2		✅	✅	✅
Canada Central	✅	✅	✅	✅
Central India		✅		✅

Region	Microsoft Purview Data Map	Scan	Policy	Insights
East US 2		✅	✅	✅
France Central		✅	✅	✅
Germany West Central		✅	✅	✅
Japan East		✅	✅	✅
Korea Central		✅	✅	✅
West US 3		✅	✅	✅
North Europe		✅	✅	✅
South Africa North		✅	✅	✅
Sweden Central		✅	✅	✅
Switzerland North		✅		
USGov Virginia	✅	✅	✅	✅
South Central US		✅	✅	✅
Brazil South		✅	✅	✅
UK South	✅	✅	✅	✅
Qatar Central		✅	✅	✅
China North 3	✅	✅	✅	✅
West Europe		✅	✅	✅

Disaster recovery and business continuity

Disaster recovery refers to practices that organizations use to recover from events such as natural disasters or failed deployments that result in downtime and data loss. Regardless of the cause, the best remedy for a disaster is a well-defined and tested disaster recovery plan and an application design that actively supports disaster recovery. Before you start creating your disaster recovery plan, see Recommendations for designing a disaster recovery strategy.

For disaster recovery, Microsoft uses the shared responsibility model. In this model, Microsoft ensures that the baseline infrastructure and platform services are available. However, many Azure services don't automatically replicate data or fall back from a failed region to cross-replicate to another enabled region. For those services, you're responsible for setting up a disaster recovery plan that works for your workload. Most services that run on Azure platform

as a service (PaaS) offerings provide features and guidance to support disaster recovery. Use service-specific features to support fast recovery to help develop your disaster recovery plan.

> **Important**
>
> Classic Microsoft Purview doesn't support automated disaster recovery. You're responsible for backup and restore activities. You can manually create a secondary Purview account as a warm standby instance in another region. This standby instance in another region doesn't support Microsoft Purview data governance.

To implement disaster recovery for Microsoft Purview classic data governance, see disaster recovery and migration best practices.

Last updated on 02/18/2026

Classic Microsoft Purview governance portal

Article • 01/28/2025

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

This article gives an overview of some of the features in the classic Microsoft Purview governance portal.

> **Tip**
>
> Microsoft Purview has a new experience! If you're a new Microsoft Purview customer, or if you want more information, see our **article on the new portal** or the article about our new **data governance experience**.

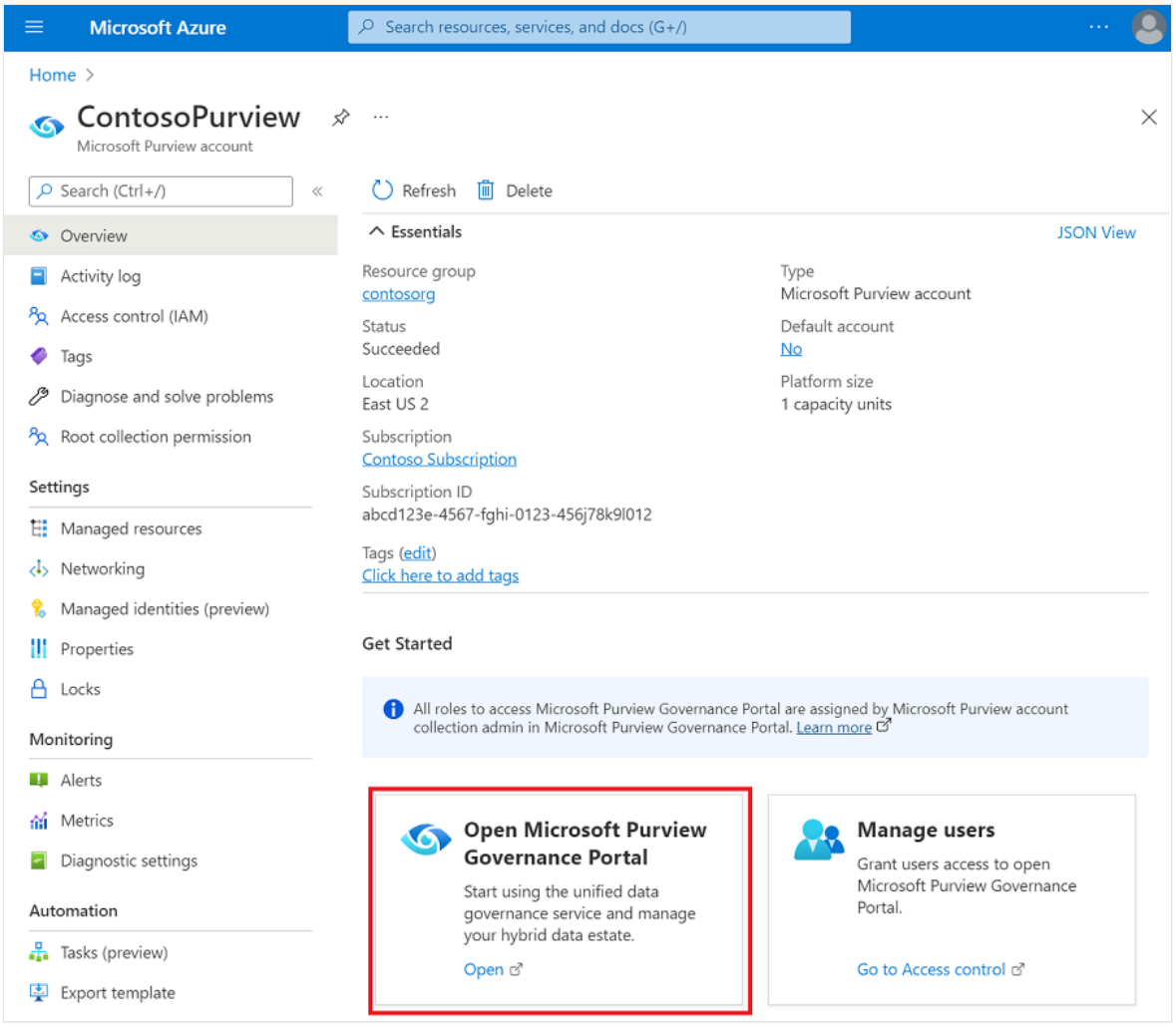
Prerequisites

- An Active Microsoft Purview account is already created in Azure portal
- The user has permissions to access the classic Microsoft Purview governance portal.

Launch Microsoft Purview account

- You can launch the Microsoft Purview account directly by going to `https://web.purview.azure.com`, selecting **Microsoft Entra ID** and the account name. Or by going to `https://web.purview.azure.com/resource/yourpurviewaccountname`

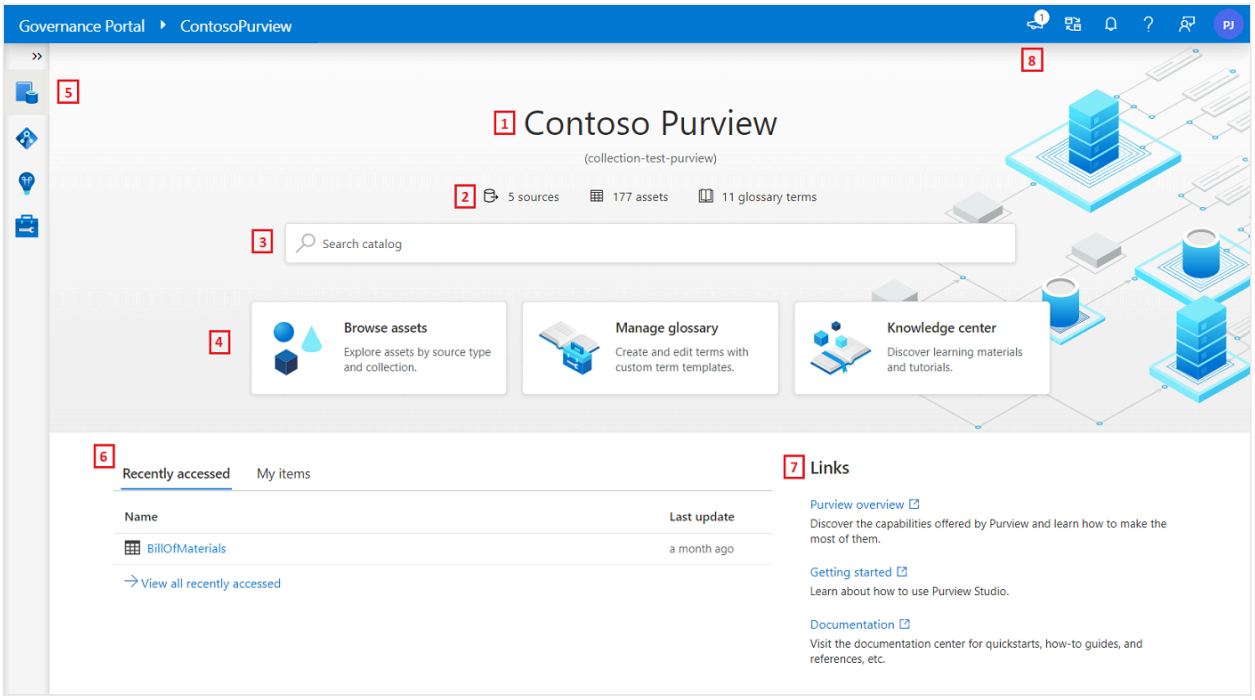
- To launch your Microsoft Purview account from the Azure portal, go to Microsoft Purview accounts in Azure portal, select the account you want to launch and launch the account.



> **Tip**
>
> If you can't access the portal, confirm you have the necessary permissions.

Home page

Home is the starting page for the Microsoft Purview governance portal.



The following list summarizes the main features of **Home page**. Each number in the list corresponds to a highlighted number in the preceding screenshot.

1. Friendly name of the catalog. You can set catalog name in **Management** > **Account information**.

2. Catalog analytics shows the number of:

    - Data sources
    - Assets
    - Glossary terms

3. The search box allows you to search for data assets across the data catalog.

4. The quick access buttons give access to frequently used functions of the application. The buttons that are presented, depend on the role assigned to your user account at the root collection.

    - For *collection admin*, the available button is **Knowledge center**.
    - For *data curator*, the buttons are **Browse assets**, **Manage glossary**, and **Knowledge center**.
    - For *data reader*, the buttons are **Browse assets**, **View glossary**, and **Knowledge center**.
    - For *data source admin* + *data curator*, the buttons are **Browse assets**, **Manage glossary**, and **Knowledge center**.
    - For *data source admin* + *data reader*, the buttons are **Browse assets**, **View glossary**, and **Knowledge center**.

> **Note**
>
> For more information about Microsoft Purview roles, see Access control in Microsoft Purview.

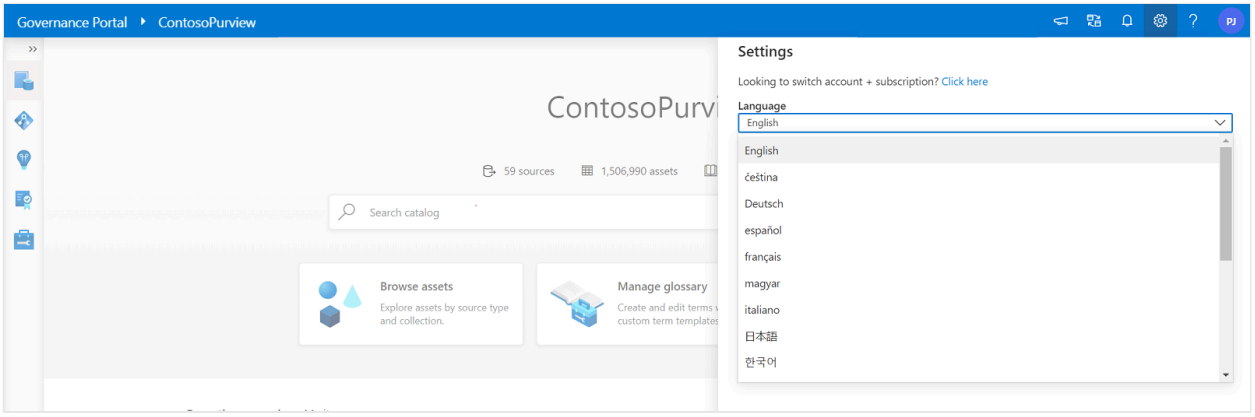
5. The left navigation bar helps you locate the main pages of the application.
6. The **Recently accessed** tab shows a list of recently accessed data assets. For information about accessing assets, see Search the Data Catalog and Browse by asset type. **My items** tab is a list of data assets owned by the logged-on user.
7. **Links** contains links to region status, documentation, pricing, overview, and Microsoft Purview status
8. The top navigation bar contains information about release notes/updates, change purview account, notifications, help, and feedback sections.

Knowledge center

The knowledge center is where you can find all the videos and tutorials related to Microsoft Purview.

Localization

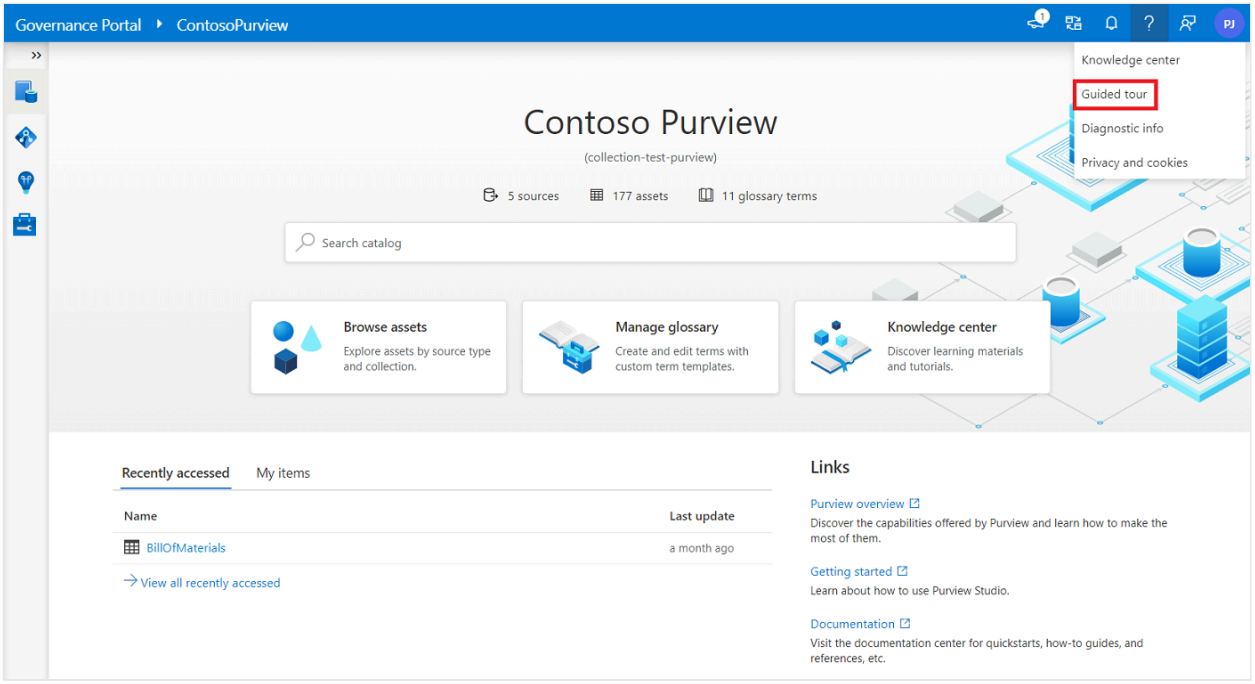
Microsoft Purview is localized in 18 languages. To change the language used, go to the **Settings** from the top bar and select the desired language from the dropdown.



> **Note**
>
> Only generally available features are localized. Features still in preview are in English regardless of which language is selected.

Guided tours

Each UX in the Microsoft Purview governance portal will have guided tours to give overview of the page. To start the guided tour, select **help** on the top bar and select **guided tours**.



Next steps

Scan data

Feedback

Was this page helpful? Yes No

Access control in the classic Microsoft Purview governance portal

Article • 05/20/2025

The Microsoft Purview governance portal uses **Collections** in the Microsoft Purview Data Map to organize and manage access across its sources, assets, and other artifacts. This article describes collections and access management for your account in the classic Microsoft Purview governance portal.

> **Important**
>
> This article covers Microsoft Purview data governance permissions in the classic Microsoft Purview governance portal.
>
> - For governance permissions the Microsoft Purview portal see governance permissions in the portal.
> - For general permissions in the new Microsoft Purview portal see permissions in the portal.

Permissions to access the classic Microsoft Purview governance portal

There are two main ways to access the classic Microsoft Purview governance portal, both of which require permissions:

- To access your Microsoft Purview governance portal directly at https://web.purview.azure.com, you'll need at least a reader role on a collection in your Microsoft Purview Data Map.
- To access your Microsoft Purview governance portal through the Azure portal by searching for your Microsoft Purview account, opening it, and selecting **Open Microsoft Purview governance portal**, you'll need at least a **Reader** role under **Access Control (IAM)**.

> **Note**
>
> If you created your account using a service principal, to be able to access the Microsoft Purview governance portal you will need to grant a user collection admin permissions on the root collection.

Collections

A collection is a tool that the Microsoft Purview Data Map uses to group assets, sources, and other artifacts into a hierarchy to manage access control. All accesses to the Microsoft Purview governance portal's resources are managed from collections in the Microsoft Purview Data Map.

Roles

The Microsoft Purview governance portal uses a set of predefined roles to control who can access what within the account. These roles are currently:

- **Domain admin** (domain level only) - Can assign permissions within a domain and manage its resources.
- **Collection administrator** - a role for users that will need to assign roles to other users in the Microsoft Purview governance portal or manage collections. Collection admins can add users to roles on collections where they're admins. They can also edit collections, their details, and add subcollections. A collection administrator on the root collection also automatically has permission to the Microsoft Purview governance portal. If your **root collection administrator** ever needs to be changed, you can follow the steps in the section below.
- **Data curators** - a role that provides access to the Microsoft Purview Unified Catalog to manage assets, configure custom classifications, create and manage glossary terms, and view data estate insights. Data curators can create, read, modify, move, and delete assets. They can also apply annotations to assets.
- **Data readers** - a role that provides read-only access to data assets, classifications, classification rules, collections and glossary terms.
- **Data source administrator** - a role that allows a user to manage data sources and scans. If a user is granted only to **Data source admin** role on a given data source, they can run new scans using an existing scan rule. To create new scan rules, the user must be also granted as either **Data reader** or **Data curator** roles.
- **Insights reader** - a role that provides read-only access to insights reports for collections where the insights reader also has at least the **Data reader** role. For more information, see insights permissions.
- **Policy author** - a role that allows a user to view, update, and delete Microsoft Purview policies through the Data policy app within Microsoft Purview.
- **Workflow administrator** - a role that allows a user to access the workflow authoring page in the Microsoft Purview governance portal, and publish workflows on collections where they have access permissions. Workflow administrator only has access to authoring, and

so will need at least Data reader permission on a collection to be able to access the Purview governance portal.

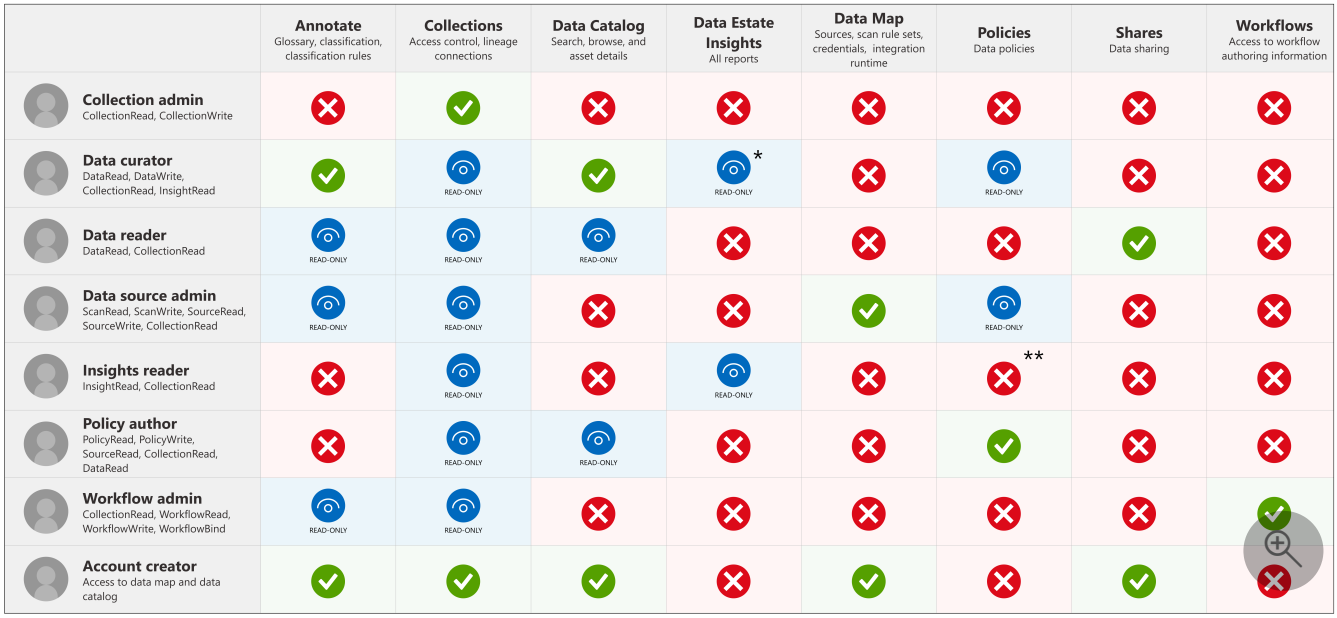
> **Note**
>
> At this time, Microsoft Purview policy author role is not sufficient to create policies. The Microsoft Purview data source admin role is also required.

Who should be assigned to what role?

Expand table

User Scenario	Appropriate Role(s)
I just need to find assets, I don't want to edit anything	Data reader
I need to edit and manage information about assets	Data curator
I want to create custom classifications	Data curator **or** data source administrator
I need to edit the business glossary	Data curator
I need to view Data Estate Insights to understand the governance posture of my data estate	Data curator
My application's Service Principal needs to push data to the Microsoft Purview Data Map	Data curator
I need to set up scans via the Microsoft Purview governance portal	Data curator on the collection **or** data curator **and** data source administrator where the source is registered.
I need to enable a Service Principal or group to set up and monitor scans in the Microsoft Purview Data Map without allowing them to access the catalog's information	Data source administrator
I need to put users into roles in the Microsoft Purview governance portal	Collection administrator
I need to create and publish access policies	Data source administrator and policy author
I need to create workflows for my Microsoft Purview account in the governance portal	Workflow administrator
I need to share data from sources registered in Microsoft Purview	Data reader

User Scenario	Appropriate Role(s)
I need to receive shared data in Microsoft Purview	Data reader
I need to view insights for collections I'm a part of	Insights reader **or** data curator
I need to create or manage our self-hosted integration runtime (SHIR)	Data source administrator
I need to create managed private endpoints	Data source administrator



> ⓘ Note
>
> ***Data curator** - Data curators can read insights only if they are assigned data curator at the root collection level.
>
> ****Data source administrator permissions on Policies** - Data source administrators are also able to publish data policies.

Understand how to use the Microsoft Purview governance portal's roles and collections

All access control is managed through collections in the Microsoft Purview Data Map. The collections can be found in the Microsoft Purview governance portal. Open your account in the Azure portal and select the Microsoft Purview governance portal tile on the Overview page. From there, navigate to the data map on the left menu, and then select the 'Collections' tab.

When a Microsoft Purview (formerly Azure Purview) account is created, it starts with a root collection that has the same name as the account itself. The creator of the account is automatically added as a Collection Admin, Data Source Admin, Data Curator, and Data Reader on this root collection, and can edit and manage this collection.

Sources, assets, and objects can be added directly to this root collection, but so can other collections. Adding collections will give you more control over who has access to data across your account.

All other users can only access information within the Microsoft Purview governance portal if they, or a group they're in, are given one of the above roles. This means, when you create an account, no one but the creator can access or use its APIs until they're added to one or more of the above roles in a collection.

Users can only be added to a collection by a collection admin, or through permissions inheritance. The permissions of a parent collection are automatically inherited by its subcollections. However, you can choose to restrict permission inheritance on any collection. If you do this, its subcollections will no longer inherit permissions from the parent and will need to be added directly, though collection admins that are automatically inherited from a parent collection can't be removed.

You can assign roles to users, security groups, and service principals from your Microsoft Entra ID that is associated with your subscription.

Assign permissions to your users

After creating a Microsoft Purview (formerly Azure Purview) account, the first thing to do is create collections and assign users to roles within those collections.

> ⓘ **Note**
>
> If you created your account using a service principal, to be able to access the Microsoft Purview governance portal and assign permissions to users, you will need to **grant a user collection admin permissions on the root collection**.

Create collections

Collections can be customized for structure of the sources in your Microsoft Purview Data Map, and can act like organized storage bins for these resources. When you're thinking about the collections you might need, consider how your users will access or discover information. Are your sources broken up by departments? Are there specialized groups within those

departments that will only need to discover some assets? Are there some sources that should be discoverable by all your users?

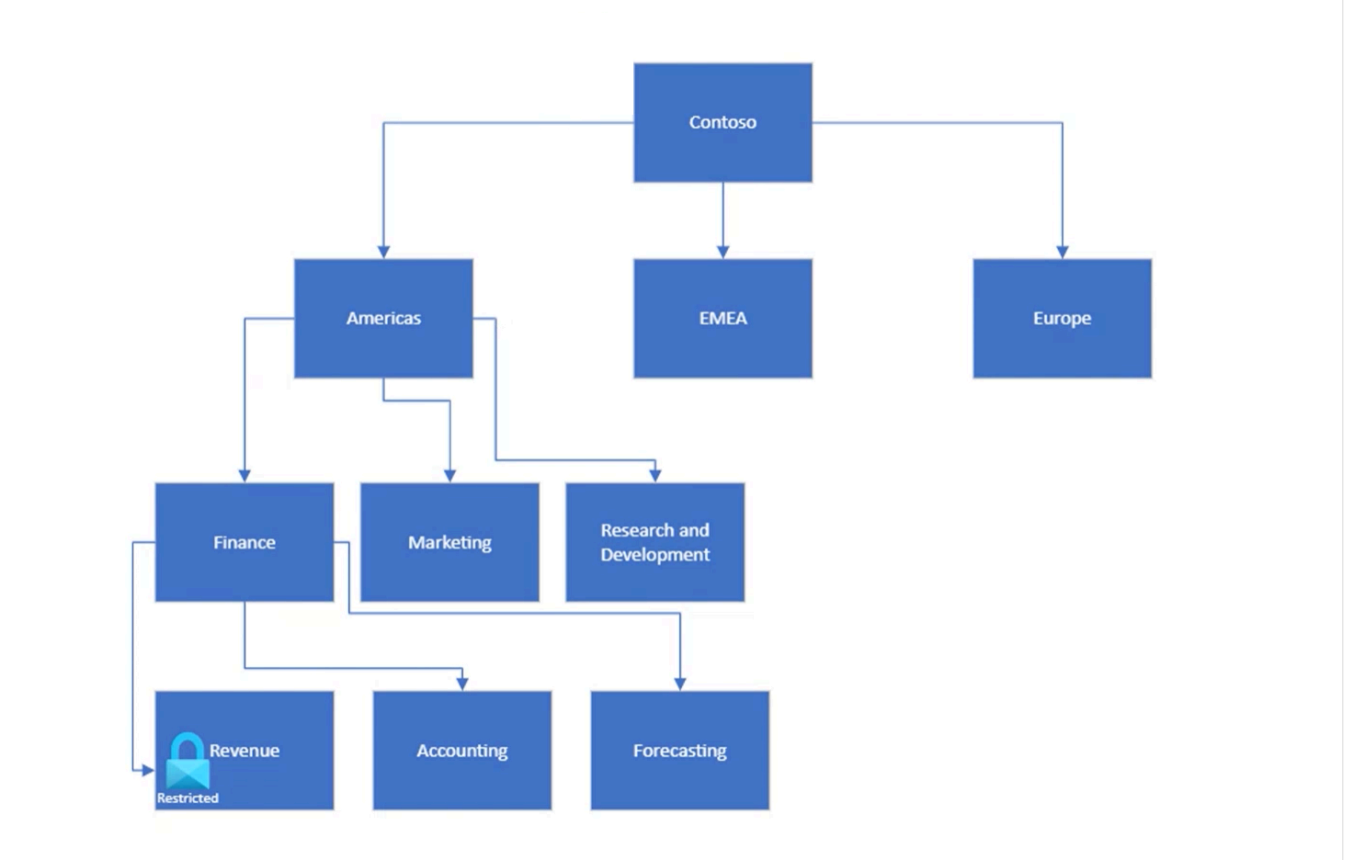
This will inform the collections and subcollections you may need to most effectively organize your data map.

New collections can be added directly to the data map where you can choose their parent collection from a drop-down, or they can be added from the parent as a sub collection. In the data map view, you can see all your sources and assets ordered by the collections, and in the list, the source's collection is listed.

For more instructions and information, you can follow our guide for creating and managing collections.

A collections example

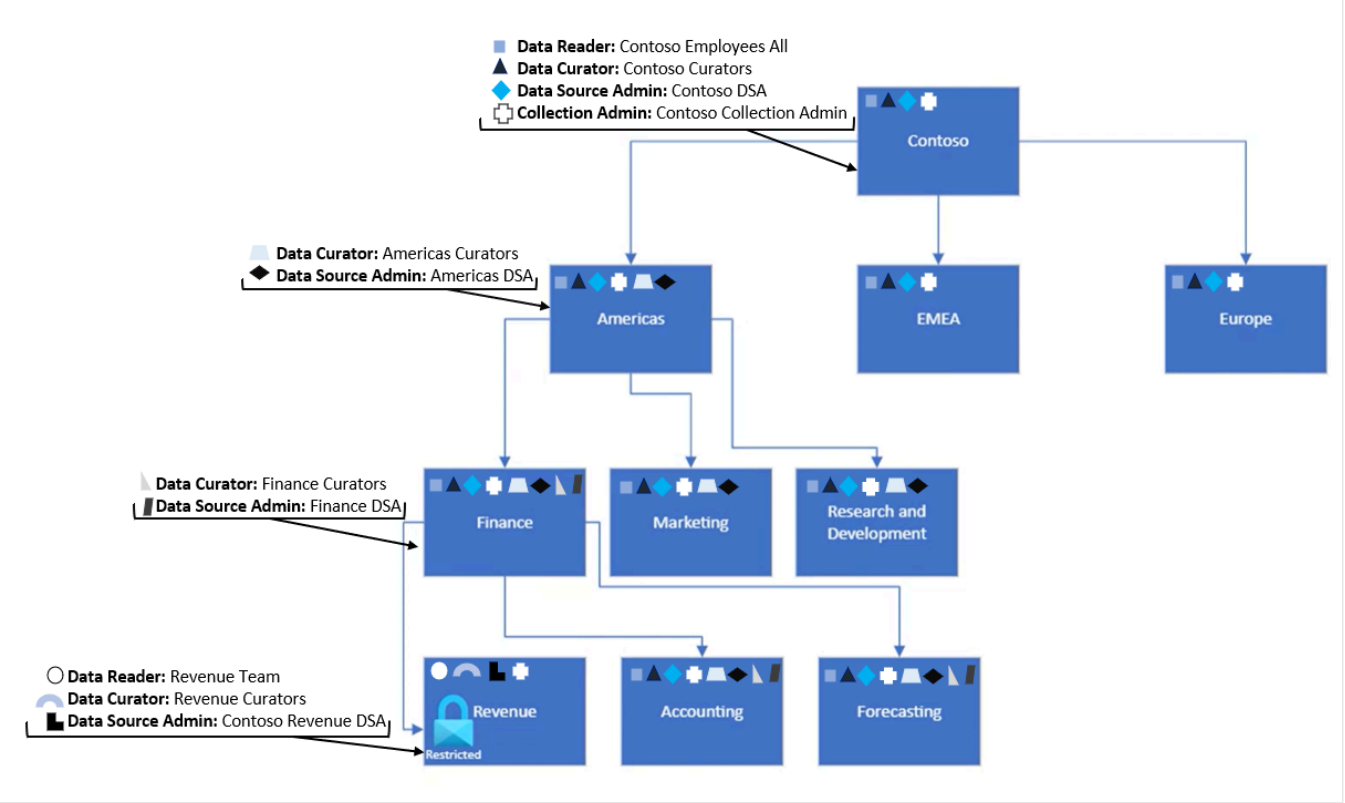
Now that we have a base understanding of collections, permissions, and how they work, let's look at an example.



This is one way an organization might structure their data: Starting with their root collection (Contoso, in this example) collections are organized into regions, and then into departments and subdepartments. Data sources and assets can be added to any one these collections to organize data resources by these regions and department, and manage access control along

those lines. There's one subdepartment, Revenue, that has strict access guidelines so permissions will need to be tightly managed.

The data reader role can access information within the catalog, but not manage or edit it. So for our example above, adding the Data Reader permission to a group on the root collection and allowing inheritance will give all users in that group reader permissions on sources and assets in the Microsoft Purview Data Map. This makes these resources discoverable, but not editable, by everyone in that group. Restricting inheritance on the Revenue group will control access to those assets. Users who need access to revenue information can be added separately to the Revenue collection. Similarly with the Data Curator and Data Source Admin roles, permissions for those groups will start at the collection where they're assigned and trickle down to subcollections that haven't restricted inheritance. Below we have assigned permissions for several groups at collections levels in the Americas sub collection.



Add users to roles

Role assignment is managed through the collections. Only a user with the collection admin role can grant permissions to other users on that collection. When new permissions need to be added, a collection admin will access the Microsoft Purview governance portal, navigate to data map, then the collections tab, and select the collection where a user needs to be added. From the Role Assignments tab, they'll be able to add and manage users who need permissions.

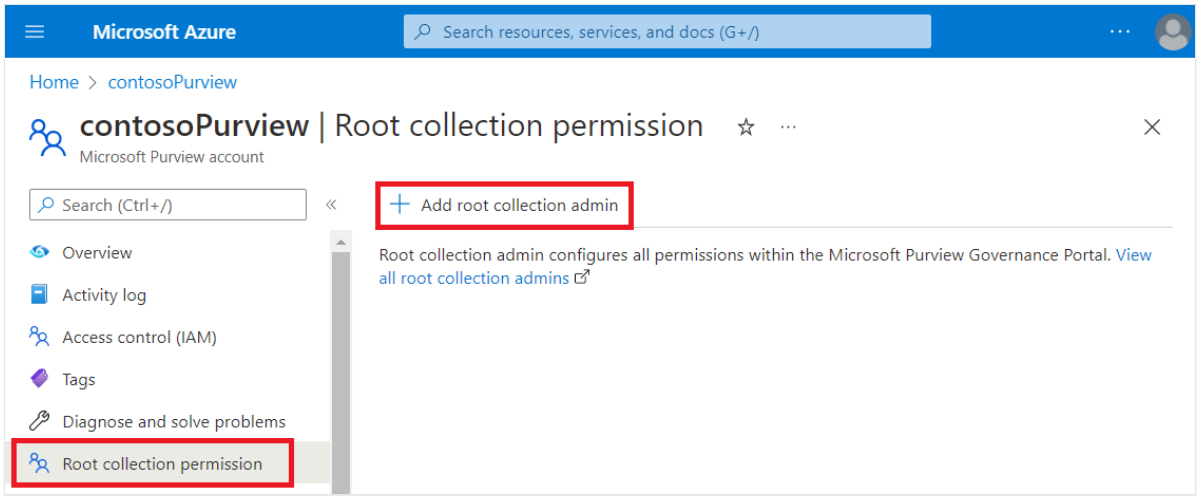
For full instructions, see our how-to guide for adding role assignments.

Administrator change

There may be a time when your root collection admin needs to change, or an admin needs to be added after an account is created by an application. By default, the user who creates the account is automatically assigned collection admin to the root collection. To update the root collection admin, there are four options:

- You can manage root collection administrators in the Azure portal:

    1. Sign in to the Azure portal and search for your Microsoft Purview account.
    2. Select **Root collection permission** from the left-side menu on your Microsoft Purview account page.
    3. Select **Add root collection admin** to add an administrator.

    4. You can also select **View all root collection admins** to be taken to the root collection in the Microsoft Purview governance portal.

- You can assign permissions through the Microsoft Purview governance portal as you have for any other role.

- You can use the REST API to add a collection administrator. Instructions to use the REST API to add a collection admin can be found in our REST API for collections documentation. For additional information, you can see our REST API reference.

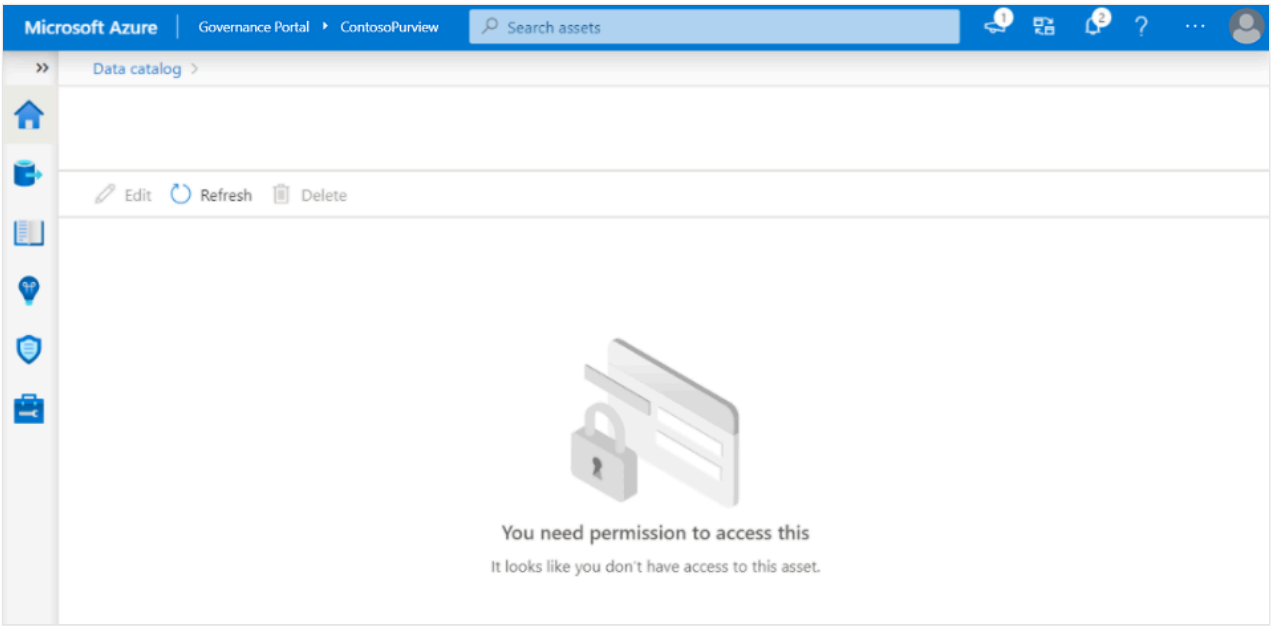
- You can also use the below Azure CLI command. The object-id is optional. For more information and an example, see the CLI command reference page.

    Azure CLI

    ```
    az purview account add-root-collection-admin --account-name [Microsoft
    Purview Account Name] --resource-group [Resource Group Name] --object-id
    [User Object Id]
    ```

Permissions on assets

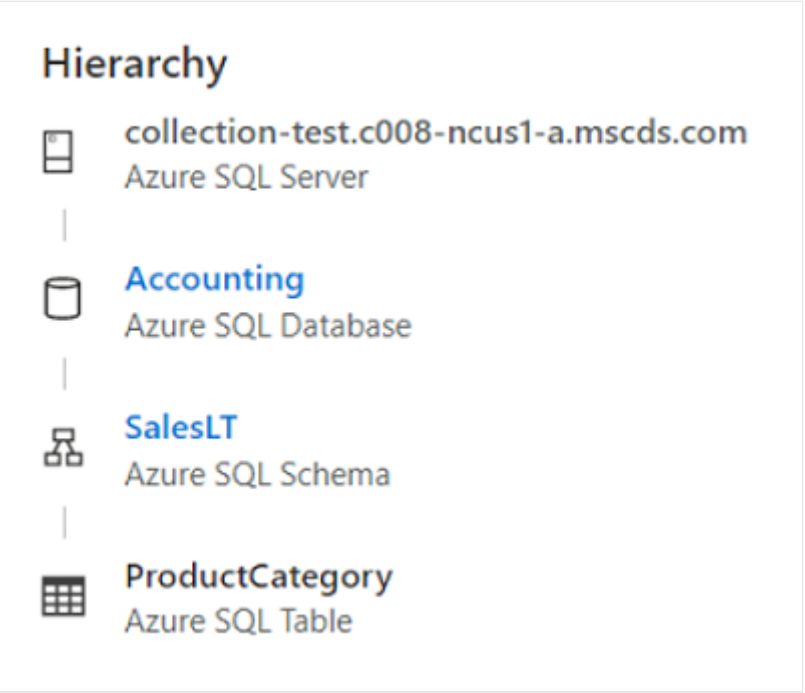
1. If you don't have read permission on a collection, the assets under that collection won't be listed in search results. If you get the direct URL of one asset and open it, you'll see the no access page. Contact your collection admin to grant you the access. You can select the **Refresh** button to check the permission again.



2. If you have the read permission to one collection but don't have the write permission, you can browse the asset details page, but the following operations are disabled:

    - Edit the asset. The **Edit** button will be disabled.
    - Delete the asset. The **Delete** button will be disabled.
    - Move asset to another collection. The ellipsis button on the right-top corner of Collection path section will be hidden.

3. The assets in **Hierarchy** section are also affected by permissions. Assets without read permission will be grayed.



Next steps

Now that you have a base understanding of collections, and access control, follow the guides below to create and manage those collections, or get started with registering sources into your Microsoft Purview Data Map.

- How to create and manage collections
- Supported data sources in the Microsoft Purview Data Map

What are Microsoft Purview's classic governance solutions?

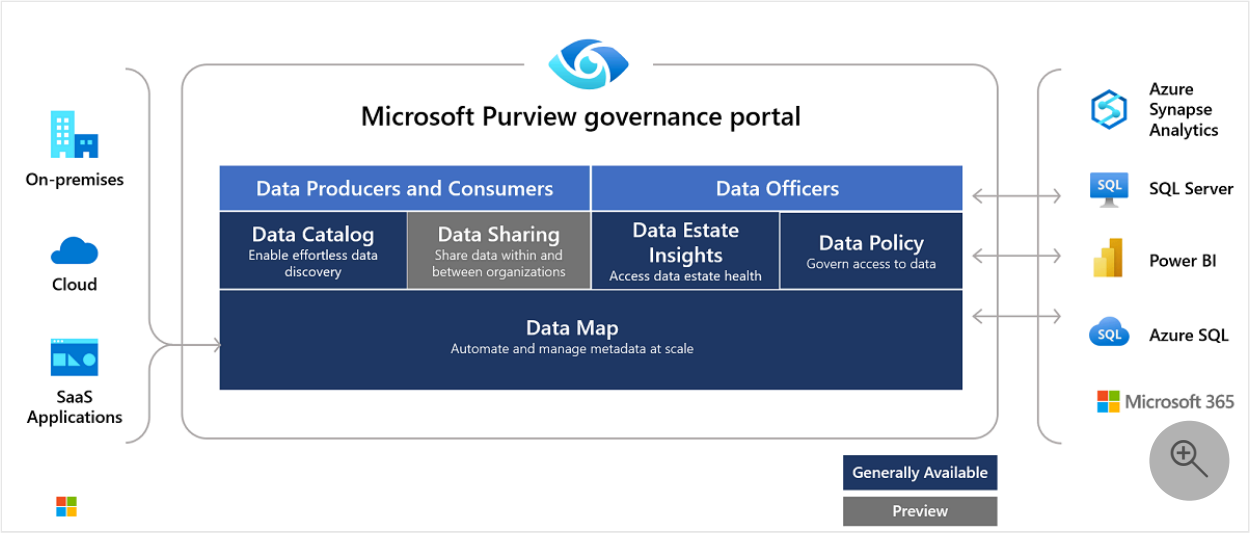
Article • 01/28/2025

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

> **Note**
>
> Data governance in Microsoft Purview has a new experience! This article covers functionality and features in the classic experience. For more information about the latest tools and features, see the overview for data governance using the data catalog.

Microsoft Purview's classic data governance solutions created one place for you to manage your on-premises, multicloud, and software-as-a-service (SaaS) data. Using the Microsoft Purview Data Catalog, Data Map, Data Estate Insights, and Policies you could:

- Create an up-to-date map of your business' entire data landscape.
- Empower your users to find useful, trustworthy data.
- Classify your sensitive information, so it's visible and you can manage it.
- Automatically map your data's lineage so you can see where it's coming from and where it's going.
- Enable your business' data curators and security administrators with the tools they need to manage and secure your data estate.



> **Tip**
>
> Looking to govern your data in Microsoft 365 by keeping what you need and deleting what you don't? Use Microsoft Purview Data Lifecycle Management.

Data Map

Microsoft Purview automates data discovery by providing data scanning and classification for assets across your data estate. Metadata and descriptions of discovered data assets are integrated into a holistic map of your data estate. Microsoft Purview Data Map provides the foundation for data discovery and data governance. Microsoft Purview Data Map is a cloud native PaaS service that captures metadata about enterprise data present in analytics and operation systems on-premises and cloud. Microsoft Purview Data Map is automatically kept up to date with built-in automated scanning and classification system. Business users can configure and use the data map through an intuitive UI and developers can programmatically interact with the Data Map using open-source Apache Atlas 2.2 APIs. Microsoft Purview Data Map powers the Microsoft Purview Data Catalog, the Microsoft Purview Data Estate Insights and the Microsoft Purview Data Policy as unified experiences within the Microsoft Purview governance portal.

For more information, see our introduction to Data Map.

Atop the Data Map, there are purpose-built apps that create environments for data discovery, access management, and insights about your data landscape.

Expand table

App	Description
Classic Data Catalog	Finds trusted data sources by browsing and searching your data assets. The data catalog aligns your assets with friendly business terms and data classification to identify data sources.
Data Estate Insights	Gives you an overview of your data estate to help you discover what kinds of data you have and where it is.
Data Policy	A set of central, cloud-based experiences that help you provision access to data securely and at scale.

Classic Data Catalog

With the classic Microsoft Purview Data Catalog, business and technical users can quickly and easily find relevant data using a search experience with filters based on lenses such as glossary terms, classifications, sensitivity labels and more. For subject matter experts, data stewards and officers, the Microsoft Purview Data Catalog provides data curation features such as business glossary management and the ability to automate tagging of data assets with glossary terms. Data consumers and producers can also visually trace the lineage of data assets: for example, starting from operational systems on-premises, through movement, transformation & enrichment with various data storage and processing systems in the cloud, to consumption in an analytics system like Power BI. For more information, see our introduction to search using Data Catalog.

Data Estate Insights

With the Microsoft Purview Data Estate Insights, the chief data officers and other governance stakeholders can get a bird's eye view of their data estate and can gain actionable insights into the governance gaps that can be resolved from the experience itself.

For more information, see our introduction to Data Estate Insights.

Data Policy

Microsoft Purview Data Policy is a set of central, cloud-based experiences that help you manage access to data sources and datasets securely and at scale.

- Manage access to data sources from a single-pane of glass, cloud-based experience
- Enables at-scale access provisioning
- Introduces a new data-plane permission model that is external to data sources
- It's seamlessly integrated with Microsoft Purview Data Map and Catalog:
  - Search for data assets and grant access only to what is required via fine-grained policies.
  - Path to support SaaS, on-premises, and multicloud data sources.
  - Path to create policies that leverage any metadata associated to the data objects.
- Based on role definitions that are simple and abstracted (for example: Read, Modify)

For more information, see our introductory guides:

- Data owner access policies (preview): Provision fine-grained to broad access to users and groups via intuitive authoring experience.
- Self-service access policies (preview): Self-Service: Workflow approval and automatic provisioning of access requests initiated by business analysts that discover data assets in Microsoft Purview's catalog.
- DevOps policies: Provision IT operations personnel access to SQL system metadata, so that they can monitor performance, health and audit security, while limiting the insider threat.

Here are the benefits of the Data Policy app:

Expand table

Principle	Benefit
Simplify	Permissions are bundled into role definitions that are abstracted and consistent across data source types, like Read and Modify.
	Reduce the need of permission expertise for each data source type.
Reduce effort	Graphical interface lets you navigate the data object hierarchy quickly.
	Supports policies on entire Azure resource groups and subscriptions.
Enhance security	Access is granted centrally and can be easily reviewed and revoked.
	Reduces the need for privileged accounts to configure access directly at the data source.

Principle	Benefit
	Supports the Principle of Least Privilege via data resource scopes and common role definitions.

Traditional challenges that Microsoft Purview seeks to address

Challenges for data consumers

Traditionally, discovering enterprise data sources has been an organic process based on communal knowledge. For companies that want the most value from their information assets, this approach presents many challenges:

- Because there's no central location to register data sources, users might be unaware of a data source unless they come into contact with it as part of another process.
- Unless users know the location of a data source, they can't connect to the data by using a client application. Data-consumption experiences require users to know the connection string or path.
- The intended use of the data is hidden to users unless they know the location of a data source's documentation. Data sources and documentation might live in several places and be consumed through different kinds of experiences.
- If users have questions about an information asset, they must locate the expert, or team responsible for that data and engage them offline. There's no explicit connection between the data and the experts that understand the data's context.
- Unless users understand the process for requesting access to the data source, discovering the data source and its documentation won't help them access the data.

Challenges for data producers

Although data consumers face the previously mentioned challenges, users who are responsible for producing and maintaining information assets face challenges of their own:

- Annotating data sources with descriptive metadata is often a lost effort. Client applications typically ignore descriptions that are stored in the data source.

- Creating documentation for data sources can be difficult and it's an ongoing responsibility to keep documentation in sync with data sources. Users might not trust documentation that's perceived as being out of date.
- Creating and maintaining documentation for data sources is complex and time-consuming. Making that documentation readily available to everyone who uses the data source can be even more so.
- Restricting access to data sources and ensuring that data consumers know how to request access is an ongoing challenge.

When such challenges are combined, they present a significant barrier for companies that want to encourage and promote the use and understanding of enterprise data.

Challenges for security administrators

Users who are responsible for ensuring the security of their organization's data may have any of the challenges listed above as data consumers and producers, and the following extra challenges:

- An organization's data is constantly growing and being stored and shared in new directions. The task of discovering, protecting, and governing your sensitive data is one that never ends. You need to ensure that your organization's content is being shared with the correct people, applications, and with the correct permissions.
- Understanding the risk levels in your organization's data requires diving deep into your content, looking for keywords, RegEx patterns, and sensitive data types. For example, sensitive data types might include Credit Card numbers, Social Security numbers or Bank Account numbers. You must constantly monitor all data sources for sensitive content, as even the smallest amount of data loss can be critical to your organization.
- Ensuring that your organization continues to comply with corporate security policies is a challenging task as your content grows and changes, and as those requirements and policies are updated for changing digital realities. Security administrators need to ensure data security in the quickest time possible.

Microsoft Purview advantages

Microsoft Purview is designed to address the issues mentioned in the previous sections and to help enterprises get the most value from their existing information assets. The catalog makes data sources easily discoverable and understandable by the users who manage the data.

Microsoft Purview provides a cloud-based service into which you can register data sources. During registration, the data remains in its existing location, but a copy of its metadata is added to Microsoft Purview, along with a reference to the data source location. The metadata is also indexed to make each data source easily discoverable via search and understandable to the users who discover it.

After you register a data source, you can then enrich its metadata. Either the user who registered the data source or another user in the enterprise can add more metadata. Any user can annotate a data source by providing descriptions, tags, or other metadata for requesting data source access. This descriptive metadata supplements the structural metadata, such as column names and data types that are registered from the data source.

Discovering and understanding data sources and their use is the primary purpose of registering the sources. Enterprise users might need data for business intelligence, application development, data science, or any other task where the correct data is required. They can use the data catalog discovery experience to quickly find data that matches their needs, understand the data to evaluate its fitness for purpose, and consume the data by opening the data source in their tool of choice.

At the same time, users can contribute to the catalog by tagging, documenting, and annotating data sources that have already been registered. They can also register new data sources, which are then discovered, understood, and consumed by the community of catalog users.

Lastly, Microsoft Purview Data Policy app provides a superior solution to keep your data secure.

In-region data residency

Microsoft Purview processes data and stores metadata information, but does not store customer data. Data is processed in its data region, and customer metadata stays within the region where Microsoft Purview is deployed. For regions with data residency requirements, customer data stays within its region, and customer metadata is always kept within the same region where Microsoft Purview is deployed.

Next steps

Tip

Check if Microsoft Purview is available in your region on the **regional availability page**.

To get started with Microsoft Purview, see Create a Microsoft Purview account. For pricing information see, Pricing for Microsoft Purview.

Feedback

Was this page helpful? Yes No

Overview of pricing for classic data governance solutions in Microsoft Purview

Article • 05/20/2025

Classic data governance solutions in Microsoft Purview in the classic governance portal provided a single pane of glass for managing data governance by enabling automated scanning and classifying data.

> **Tip**
>
> For specific price details for governance in the classic portal, see the Microsoft Purview (formerly Azure Purview) pricing page. This article will guide you through the features and factors that will affect pricing.

Why do you need to understand the components of pricing?

- While the pricing for classic data governance solutions in Microsoft Purview is on a subscription-based **Pay-As-You-Go** model, there are various dimensions that you can consider while budgeting
- This guideline is intended to help you plan the budgeting for classic data governance solutions by providing a view on the control factors that impact the budget

Factors impacting Azure Pricing

There are **direct** and **indirect** costs that need to be considered while planning budgeting and cost management.

Direct costs impacting Microsoft Purview pricing are based on these applications:

- The Microsoft Purview Data Map
- Data Estate Insights

Indirect costs

Indirect costs impacting classic data governance solutions in Microsoft Purview pricing are:

- Managed resources

    - An Event Hubs namespace can be configured at creation or enabled in the Azure portal on the Kafka configuration page of the account to enable monitoring with *Atlas Kafka* topics events. The Event Hubs will be charged separately.

- Azure private endpoint
    - Azure private end points are used for classic data governance solutions in Microsoft Purview, where it's required for users on a virtual network (virtual network) to securely access the catalog over a private link
    - The prerequisites for setting up private endpoints could result in extra costs, for example, costs if you deploy a virtual network.

- Self-hosted integration runtime related costs
    - Self-hosted integration runtime requires infrastructure, which results in extra costs
    - It's required to deploy and register Self-hosted integration runtime (SHIR) inside the same virtual network where Microsoft Purview ingestion private endpoints are deployed
    - Other memory requirements for scanning
        - Certain data sources such as SAP require more memory on the SHIR machine for scanning

- Virtual Machine Sizing
    - Plan virtual machine sizing in order to distribute the scanning workload across VMs to optimize the v-cores utilized while running scans

- Microsoft 365 license
    - Microsoft Purview Information Protection sensitivity labels can be automatically applied to your Azure assets in the Microsoft Purview Data Map.
    - Microsoft Purview Information Protection sensitivity labels are created and managed in the Microsoft Purview portal.
    - To create sensitivity labels for use in Microsoft Purview, you must have an active Microsoft 365 license, which offers the benefit of automatic labeling. For the full list of licenses, see the Sensitivity labels in Microsoft Purview FAQ.

- Azure Alerts
    - Azure Alerts can notify customers of issues found with infrastructure or applications using the monitoring data in Azure Monitor
    - The pricing for Azure Alerts is available here

- Cost Management Budgets & Alerts
    - Automatically generated cost alerts are used in Azure to monitor Azure usage and spending based on when Azure resources are consumed
    - Azure allows you to create and manage Azure budgets. Refer tutorial

- Multicloud egress charges
  - Consider the egress charges (minimal charges added as a part of the multicloud subscription) associated with scanning multicloud (for example AWS, Google) data sources running native services excepting the S3 and RDS sources

Next steps

- Classic data governance solutions in Microsoft Purview pricing page
- Pricing guideline Data Estate Insights
- Pricing guideline Data Map

Pricing for the Microsoft Purview Data Map

Article • 05/20/2025

This guide covers pricing guidelines for Microsoft Purview Data Map in the classic Microsoft Purview governance portal.

For a full pricing guideline details for Microsoft Purview (formerly Azure Purview), see the pricing guideline overview.

For specific price details, see the Microsoft Purview (formerly Azure Purview) pricing page. This article will guide you through the features and factors that will affect pricing for Data Map.

Direct costs impacting pricing for Data Map are based on the following three dimensions:

- **Elastic data map**
- **Automated scanning & classification**
- **Advanced resource sets**

Elastic Data Map

- The **Data Map** is the foundation of the Microsoft Purview governance portal architecture and so needs to be up to date with asset information in the data estate at any given point.
- The data map is charged in terms of **Capacity Unit** (CU). The data map is provisioned at one CU if the catalog is storing up to 10 GB of metadata storage and serves up to 25 data map operations/sec.
- The data map is always provisioned at one CU when an account is first created.
- However, the data map scales automatically between the minimal and maximal limits of that elasticity window, to cater to changes in the data map with respect to two key factors - **operation throughput** and **metadata storage**.

Operation throughput

- An event driven factor based on the Create, Read, Update, Delete operations performed on the data map
- Some examples of the data map operations would be:
  - Creating an asset in Data Map
  - Adding a relationship to an asset such as owner, steward, parent, lineage
  - Editing an asset to add business metadata such as description, glossary term
  - Keyword-search returning results to search result page

    - Importing or exporting information using API
- If there are multiple queries executed on the Data Map, the number of I/O operations also increases resulting in the scaling up of the data map
- The number of concurrent users also forms a factor governing the data map capacity unit
- Other factors to consider are type of search query, API interaction, workflows, approvals, and so on
- Data burst level
    - When there's a need for more operations/second throughput, the Data map can autoscale within the elasticity window to cater to the changed load
    - This constitutes the **burst characteristic** that needs to be estimated and planned for
    - The burst characteristic comprises the **burst level** and the **burst duration** for which the burst exists
        - The **burst level** is a multiplicative index of the expected consistent elasticity under steady state
        - The **burst duration** is the percentage of the month that such bursts (in elasticity) are expected because of growing metadata or higher number of operations on the data map

Metadata storage

- If the number of assets reduces in the data estate, and are then removed in the data map through subsequent incremental scans, the storage component automatically reduces and so the data map scales down

Automated scanning, classification, and ingestion

There are two major automated processes that can trigger ingestion of metadata into the Microsoft Purview Data Map:

- Automatic scans using native connectors. This process includes three main steps:
    - Metadata scan
    - Automatic classification
    - Ingestion of metadata into the Microsoft Purview Data Map

- Automated ingestion using Azure Data Factory and/or Azure Synapse pipelines. This process includes:
    - Ingestion of metadata and lineage into the Microsoft Purview Data Map if the account is connected to any Azure Data Factory or Azure Synapse pipelines.

Automatic scans using native connectors

- A **full scan** processes all assets within a selected scope of a data source whereas an **incremental scan** detects and processes assets, which have been created, modified, or deleted since the previous successful scan
- All scans (full or Incremental scans) will pick up **updated, modified, or deleted** assets
- It's important to consider and avoid the scenarios when multiple people or groups belonging to different departments set up scans for the same data source resulting in more pricing for duplicate scanning
- Schedule **frequent incremental scans** post the initial full scan aligned with the changes in the data estate. This will ensure the data map is kept up to date always and the incremental scans consume lesser v-core hours as compared to a full scan
- The **"View Details"** link for a data source will enable users to run a full scan. However, consider running incremental scans after a full scan for optimized scanning excepting when there's a change to the scan rule set (classifications/file types)
- **Register the data source at a parent collection** and **Scope scans at child collection** with different access controls to ensure there are no duplicate scanning costs being entailed
- Curtail the users who are allowed to register data sources for scanning through **fine grained access control** and **Data Source Administrator** role using Collection authorization. This will ensure only valid data sources are allowed to be registered and scanning v-core hours is controlled resulting in lower costs for scanning
- Consider that the **type of data source** and the **number of assets** being scanned affect the scan duration
- **Create custom scan rule sets** to include only the subset of **file types** available in your data estate and **classifications** that are relevant to your business requirements to ensure optimal use of the scanners
- While creating a new scan for a data source, follow the **order of preparation** recommended before actually running the scan. This includes gathering the requirements for **business specific classifications** and **file types** (for storage accounts) to enable appropriate scan rule sets to be defined to avoid multiple scans and control unnecessary costs for multiple scans through missed requirements
- Align your scan schedules with Self-Hosted Integration Runtime (SHIR) VMs (Virtual Machines) size to avoid extra costs linked to virtual machines

Automated ingestion using Azure Data Factory and/or Azure Synapse pipelines

- Metadata and lineage are ingested from Azure Data Factory or Azure Synapse pipelines every time the pipelines run in the source system.

Advanced resource sets

- The Microsoft Purview Data Map uses **resource sets** to address the challenge of mapping large numbers of data assets to a single logical resource by providing the ability to scan all the files in the data lake and find patterns (GUID, localization patterns, etc.) to group them as a single asset in the data map

- **Advanced Resource Set** is an optional feature, which allows for customers to get enriched resource set information computed such as Total Size, Partition Count, etc., and enables the customization of resource set grouping via pattern rules. If Advanced Resource Set feature isn't enabled, Unified Catalog will still contain resource set assets, but without the aggregated properties. There will be no "Resource Set" meter billed to the customer in this case.

- Use the basic resource set feature, before switching on the Advanced Resource Sets in the Microsoft Purview Data Map to verify if requirements are met

- Consider turning on Advanced Resource Sets if:
  - Your data lakes schema is constantly changing, and you're looking for more value beyond the basic Resource Set feature to enable the Microsoft Purview Data Map to compute parameters such as #partitions, size of the data estate, etc., as a service
  - There's a need to customize how resource set assets get grouped.

- It's important to note that billing for Advanced Resource Sets is based on the compute used by the offline tier to aggregate resource set information and is dependent on the size/number of resource sets in your catalog

Next steps

- Microsoft Purview, formerly Azure Purview, pricing page
- Pricing guideline overview
- Pricing guideline Data Estate Insights

Pricing for Data Estate Insights

Article • 01/28/2025

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

This guide covers pricing guidelines for the classic Data Estate Insights features in the classic Microsoft Purview governance portal.

> **Tip**
>
> There's a **new data estate health experience** in the new Microsoft Purview Data Catalog.

For a full pricing guideline details for classic data governance solutions in Microsoft Purview, see the pricing guideline overview.

For specific price details, see the classic data governance solutions in Microsoft Purview pricing page. This article guides you through the features and factors that affect pricing for Data Estate Insights.

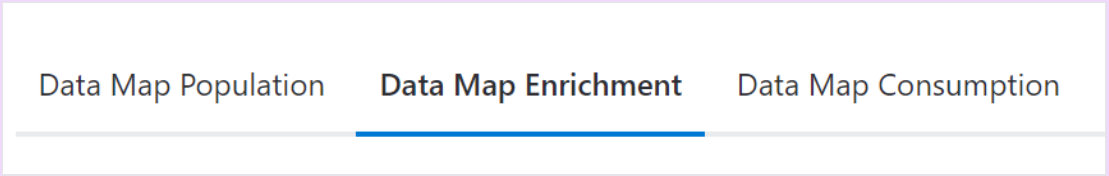
Guidelines

Data Estate Insights in the classic Microsoft Purview governance portal is billed on two dimensions:

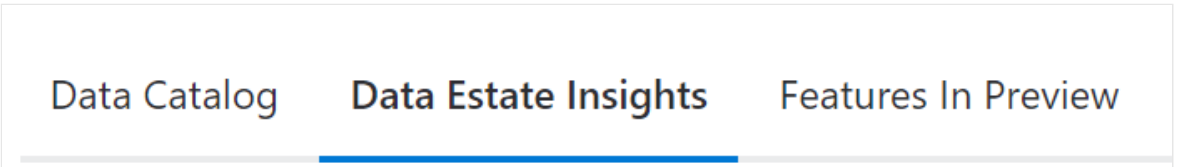
- **Report generation** - this incorporates the job that checks for any changes in your environment and the jobs that aggregate metrics about your Microsoft Purview account that will appear in specific reports.
  - If you have set your reports to refresh on a schedule, at the time of refresh there will be a job to check if any updates have been made to your environment. You'll always be billed a small amount for this check.
  - If updates have been made to your environment, you're billed for the jobs that aggregate metrics and generate your report.

    > **Note**

On the pricing page, you can find details for report generation pricing under Data Map Enrichment.



- **Report consumption** - This incorporates access of the report features (currently served through the UX). On the pricing page, you can find details for report generation pricing under Data Estate Insights.



> **Important**
>
> The Data Estate Insights application is **on** by default when you create a Microsoft Purview account. This means, "State" is on and "Refresh Frequency" is set to automatic*.
>
> * At this time automatic refresh is weekly.

If you want to reschedule or reduce how often your reports are refreshed, you can schedule your Data Estate Insights reports.

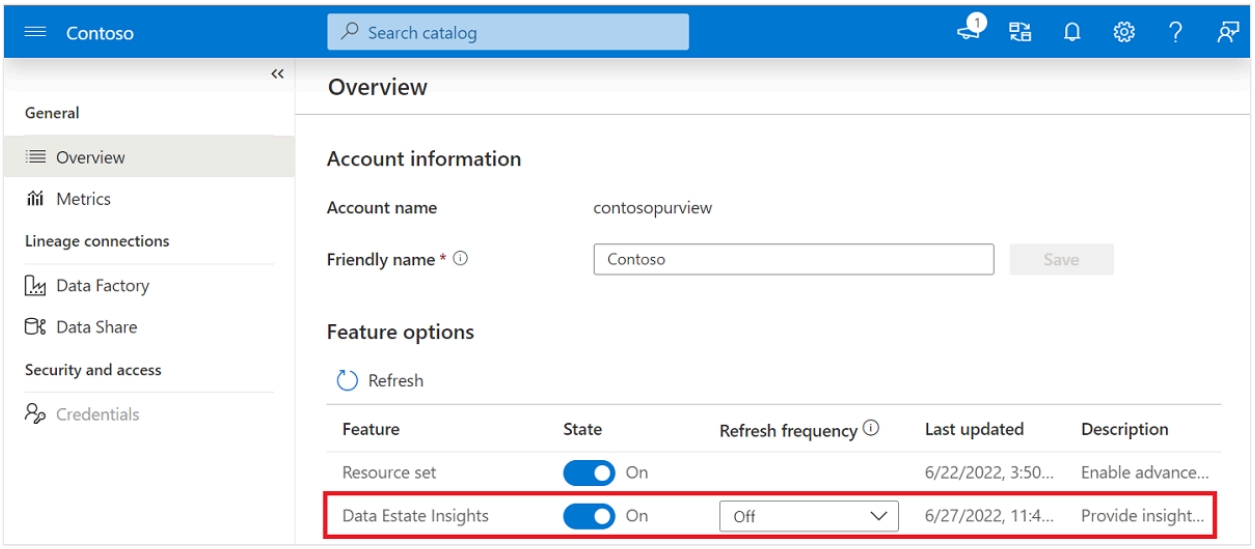
If you don't plan on using Data Estate Insights for a while, a **data curator** on the **root collection** can disable Data Estate Insights features in one of two ways:

- Disable the Data Estate Insights application - this stops billing from both report generation and report consumption.
- Disable report refresh - insights readers have access to current reports, but reports won't be refreshed. Billing occurs for report consumption but not report generation.

> **Note**
>
> The application or report refresh can be enabled again later at any time.

A **data curator** on your account's **root collection** can make these changes in the Management section of the Microsoft Purview governance portal in **Overview**, under

Feature options. For specific steps, see the disable Data Estates Insights article



Schedule Report Refresh

Your reports can be scheduled to refresh daily, weekly, or monthly depending on your needs, and this will affect billing.

Firstly, on the date/time of your scheduled refresh, insights runs a job to check if any changes have been made in your environment since the last refresh. You'll always be billed for this job.

If changes have been made since the last report refresh, billing is calculated based on the amount of compute power that is used to generate the report, and is billed at the time of report generation. For more information about specific pricing, see the pricing page **under Data Map Enrichment.**

For more information about setting your refresh schedule, see the schedule Data Estate Insights reports article.

Disable the Data Estate Insights application

Disabling Data Estate Insights disables the entire application, including these reports:

- Stewardship
- Asset
- Glossary
- Classification
- Labeling

The application icon will still show in the menu, but insights readers won't have access to reports at all, and report generation jobs are stopped. The Microsoft Purview account

won't receive any bill for Data Estate Insights.

For steps to disable the Data Estate Insights application, see the disable article.

Disable report refresh

You can choose to disable report refreshes instead of disabling the entire Data Estate Insights application.

When you disable report refreshes, insight readers will be able view reports but they'll see a banner on top of each report, warning that the report might not be current. It will also indicate the date the report was last generated.

In this case, graphs showing data from last 30 days will appear blank after 30 days. Graphs showing snapshot of the data map will continue to show graph and details. When an insights readers accesses an insight report, report consumption meter will be triggered, and the Microsoft Purview account will be billed.

For steps to disable report refresh see the disable article.

Next steps

- Schedule Data Estate Insights reports
- Disable Data Estate Insights
- Microsoft Purview, formerly Azure Purview, pricing page
- Pricing guideline overview
- Pricing guideline Data Map

Feedback

Was this page helpful? Yes No

Create a Microsoft Purview account

Article • 10/28/2024

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

> **Important**
>
> You can only create one Microsoft Purview account per tenant. If your organization already has a Microsoft Purview account, you won't be able to create a new Microsoft Purview account unless your organization already had multiple accounts and is still under the the preexisting quota. For more information, see the FAQ.

> **Tip**
>
> If your organization hasn't created a Microsoft Purview account in Azure before, start in the new Microsoft Purview portal: https://purview.microsoft.com. You can use the free version to test capabilities, and upgrade to the enterprise version for the full data governance in Microsoft Purview features.

This quickstart describes the steps to create a Microsoft Purview (formerly Azure Purview) account through the Azure portal. Then we'll get started on the process of classifying, securing, and discovering your data in the Microsoft Purview Data Map!

The Microsoft Purview governance portal surfaces tools like the Microsoft Purview Data Map and Microsoft Purview Data Catalog that help you manage and govern your data landscape. By connecting to data across your on-premises, multicloud, and software-as-a-service (SaaS) sources, the Microsoft Purview Data Map creates an up-to-date map of your data estate. It identifies and classifies sensitive data, and provides end-to-end linage. Data consumers are able to discover data across your organization, and data administrators are able to audit, secure, and ensure right use of your data.

For more information about the governance capabilities of Microsoft Purview, see our governance solutions overview page.

Prerequisites

- If you don't have an Azure subscription, create a free subscription before you begin.
- A Microsoft Entra tenant associated with your subscription.
- The user account that you use to sign in to Azure must be a member of the *contributor* or *owner* role, or an *administrator* of the Azure subscription. To view the permissions that you have in the subscription, follow these steps:

    1. Go to the Azure portal
    2. Select your username in the upper-right corner.
    3. Select the ellipsis button ("...") for more options.
    4. Then select **My permissions**.
    5. If you have access to multiple subscriptions, select the appropriate subscription.

Sign in to Azure

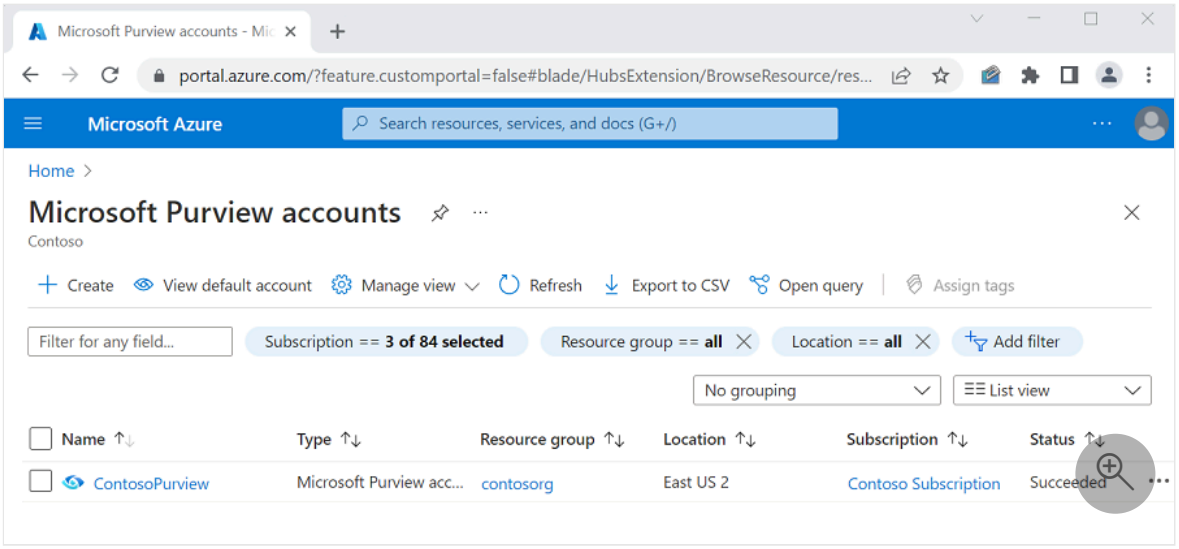
Sign in to the Azure portal with your Azure account.

Create an account

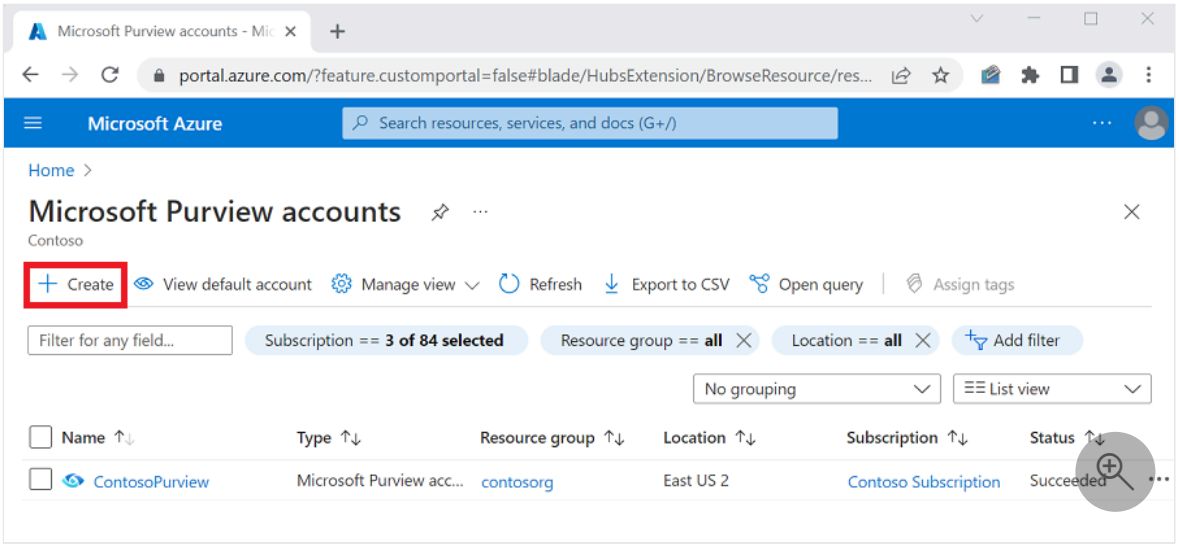
> **Tip**
>
> If you're creating the first Microsoft Purview account in your tenant, you don't need to set anything up and you should use **the free version of Microsoft Purview governance solutions** at **https://purview.microsoft.com**.

> **Important**
>
> If you have any **Azure Policies** preventing creation of **Storage accounts** or **Event Hub namespaces**, or preventing **updates to Storage accounts** first follow the **Microsoft Purview exception tag guide** to create an exception for Microsoft Purview accounts. Otherwise you will not be able to deploy Microsoft Purview.

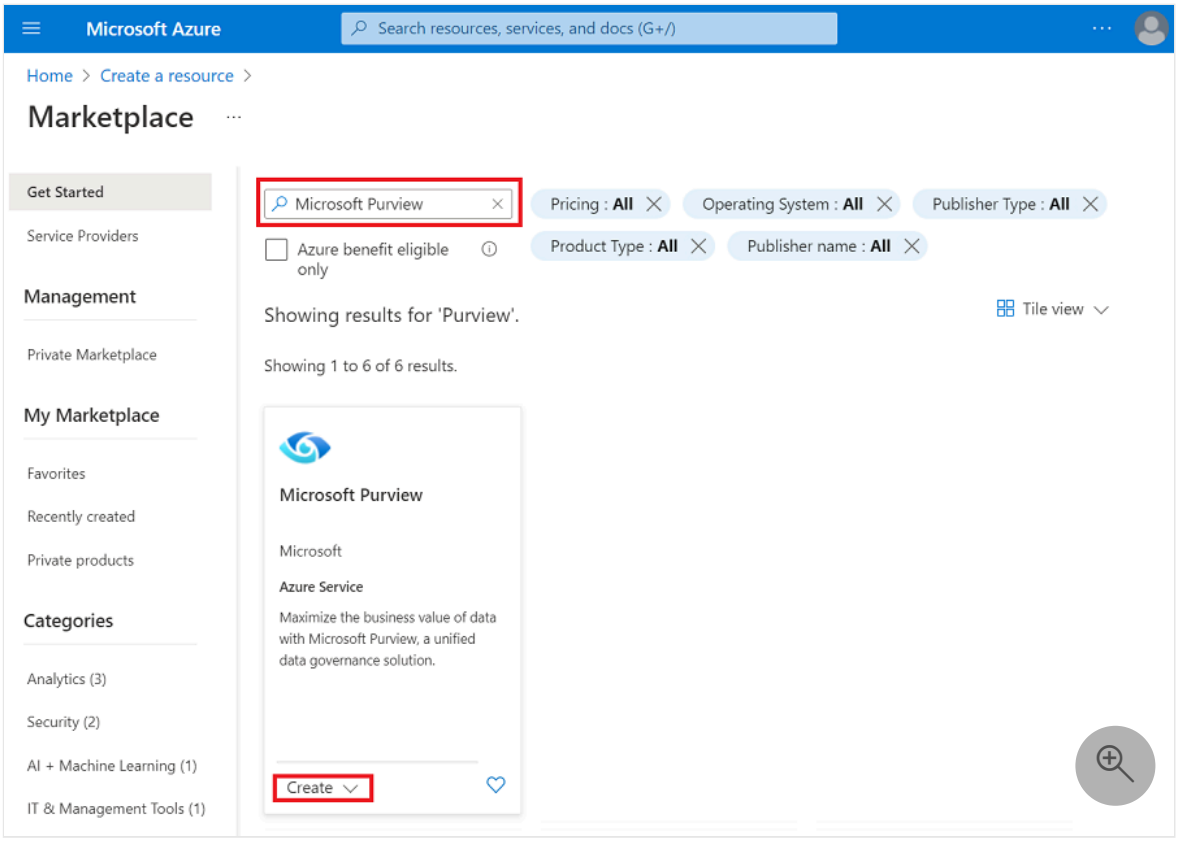
1. Search for **Microsoft Purview** in the Azure portal.



2. Select **Create** to create a new Microsoft Purview account.



Or instead, you can go to the marketplace, search for **Microsoft Purview**, and select **Create**.



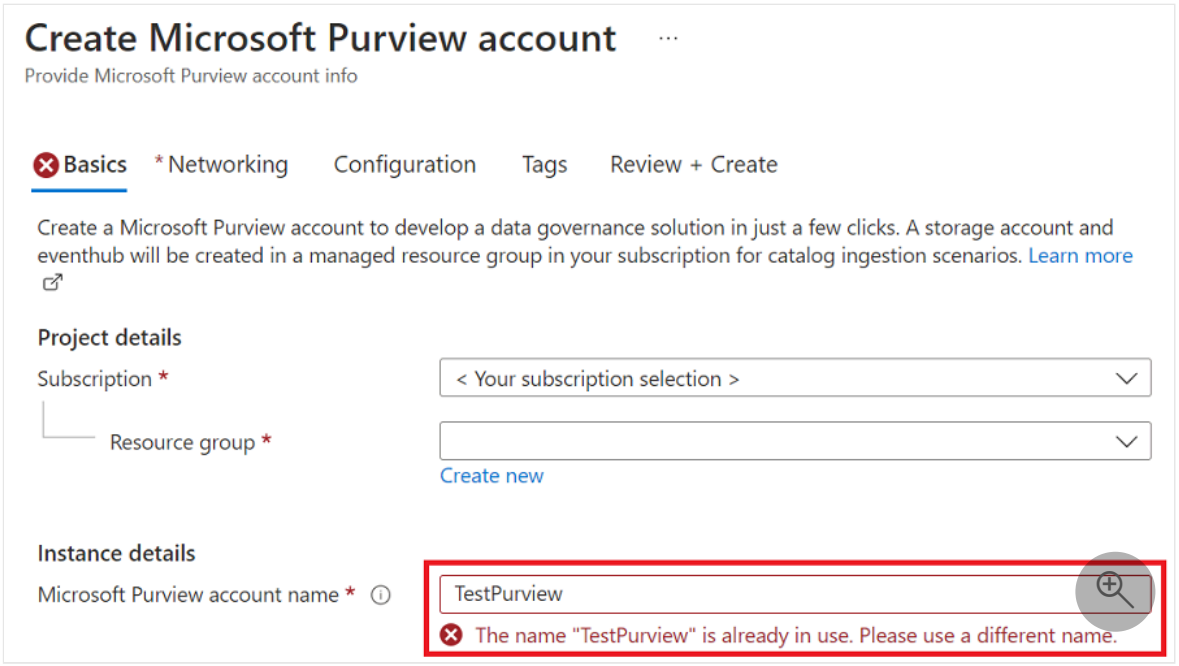
3. On the new Create Microsoft Purview account page under the **Basics** tab, select the Azure subscription where you want to create your account.

4. Select an existing **resource group** or create a new one to hold your account.

   To learn more about resource groups, see our article on using resource groups to manage your Azure resources.

5. Enter a **Microsoft Purview account name**. Spaces and symbols aren't allowed.

   The name of the Microsoft Purview account must be globally unique. If you see the following error, change the name of Microsoft Purview account and try creating again.



6. Choose a **location**.

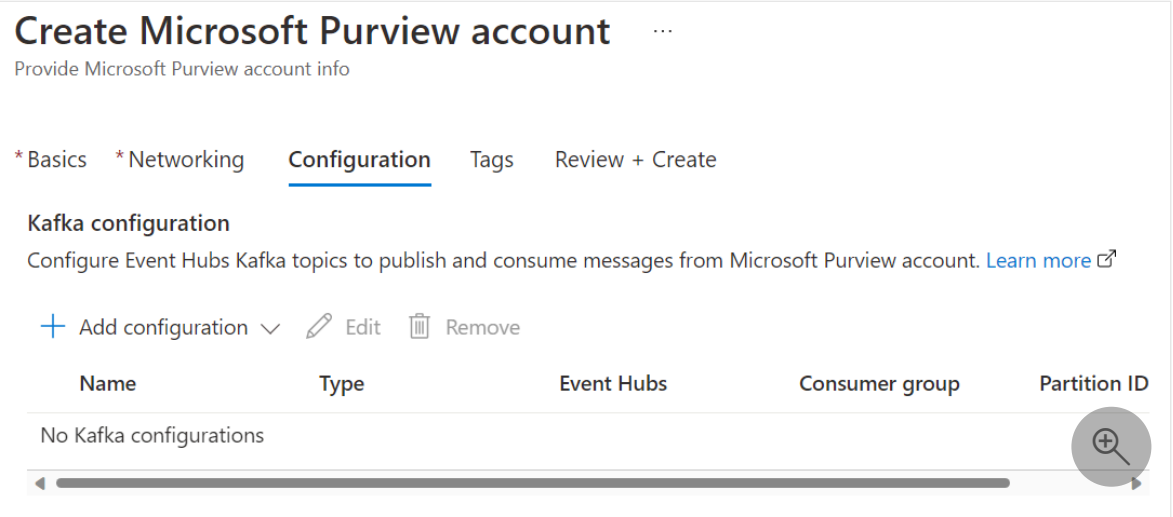
   The list shows only locations that support the Microsoft Purview governance portal, and if this is your organization's first Microsoft Purview account, the location is restricted to the one that matches your Microsoft Entra ID home region. This location is the region where your Microsoft Purview account and metadata are stored. Sources can be housed in other regions.

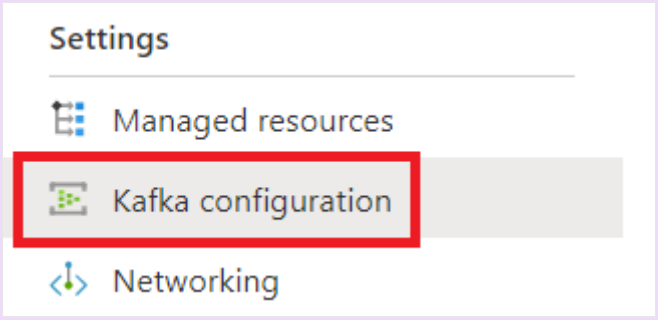
   > **Note**
   >
   > The Microsoft Purview, formerly Azure Purview, does not support moving accounts across regions, so be sure to deploy to the correction region. You can find out more information about this in move operation support for resources.

7. On the **Networking** tab you can choose to connect to all networks, or to use private endpoints. For more information and configuration options, see Configure firewall settings for your Microsoft Purview account and private endpoints for Microsoft Purview articles.

8. On **Configuration** tab you can choose to configure Event Hubs namespaces to programmatically monitor your Microsoft Purview account using Event Hubs and Atlas Kafka.

   - Steps to configure Event Hubs namespaces
   - Send and receive Atlas Kafka topics messages



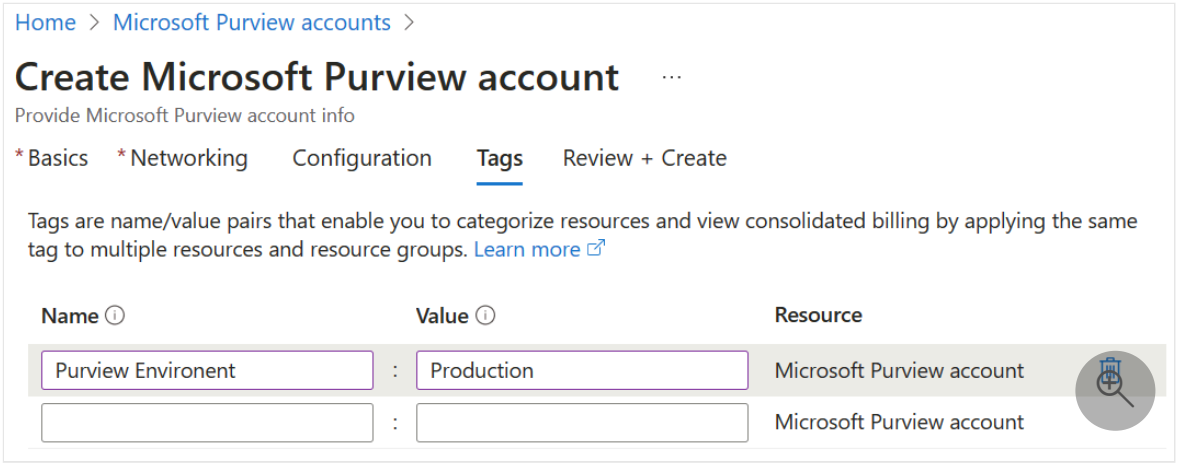
> **Note**
>
> These options can be configured after you have created your account in **Kafka configuration** under settings on your Microsoft Purview account page in the Azure Portal.
>
> 
>

9. On the **Tags** tab, add a tag called **Purview environment**, and give it one of the below values:

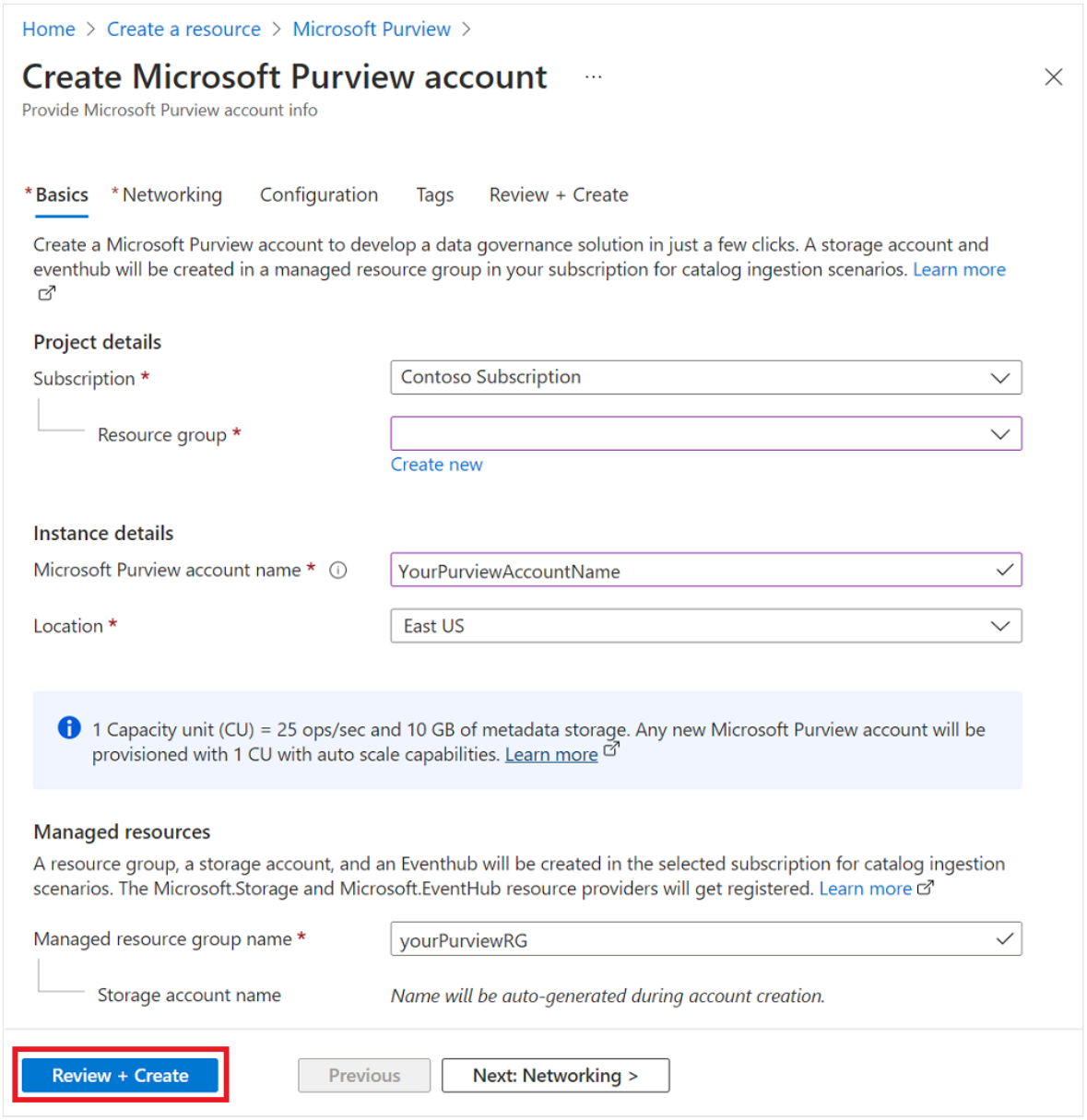
Expand table

Value	Meaning
Production	This account is being used or will be used in the future to support all my cataloging and governance requirements in production.
Pre-Production	This account is being used or will be used in the future to validate cataloging and governance requirements before making it available to my users in production.
Test	This account is being used or will be used in the future to test out capabilities in Microsoft Purview Governance.
Dev	This account is being used or will be used in the future to test out capabilities or develop custom code, scripts etc. in Microsoft Purview

Value	Meaning
	Governance.
Proof of Concept	This account is being used or will be used in the future to test out capabilities or develop custom code, scripts etc. in Microsoft Purview Governance.



10. Select **Review & Create**, and then select **Create**. It takes a few minutes to complete the creation. The newly created account appears in the list on your **Microsoft Purview accounts** page.



Protect your account

To protect your Microsoft Purview Account from accidental deletion, Microsoft recommends that you create the Azure policy ['DenyAction'](/azure/governance/policy/concepts/effect-deny-action#denyaction-example)

JSON

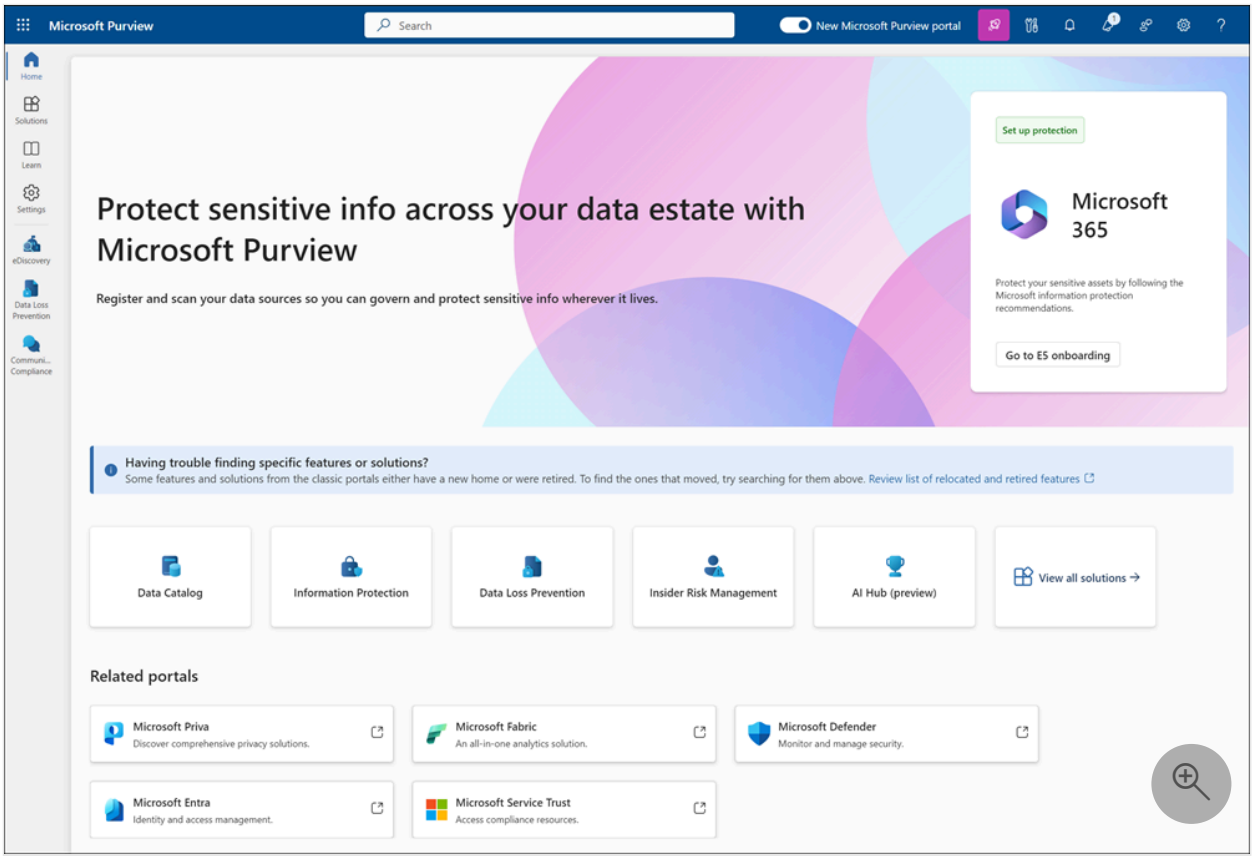
```
{
  "if": {
    "allOf": [
      {
        "field": "type",
        "equals": "Microsoft.Purview/accounts"
      },
      {
        "field": "tags.environment",
        "equals": "Production"
      }
```

```
      ]
    },
    "then": {
      "effect": "denyAction",
      "details": {
        "actionNames": [
          "delete"
        ],
        "cascadeBehaviors": {
          "resourceGroup": "deny"
        }
      }
    }
  }
```

Open the Microsoft Purview portal

After your account is created, you can use the Microsoft Purview portal to access and manage it.

In a browser, navigate directly to https://purview.microsoft.com to access the portal.



For more information about using the portal, see our documentation on the Microsoft Purview portal.

Next steps

In this quickstart, you learned how to create a Microsoft Purview (formerly Azure Purview) account, and how to access it.

Next, you can create a user-assigned managed identity (UAMI) that will enable your new Microsoft Purview account to authenticate directly with resources using Microsoft Entra authentication.

To create a UAMI, follow our guide to create a user-assigned managed identity.

Follow these next articles to learn how to navigate the Microsoft Purview governance portal, create a collection, and grant access to the Microsoft Purview Data Map:

- Using Microsoft Purview portal
- Get started with data governance in Microsoft Purview
- Add users to your Microsoft Purview account

Feedback

Was this page helpful? Yes No

Quickstart: Create a Microsoft Purview (formerly Azure Purview) account using Azure PowerShell/Azure CLI

Article • 02/26/2025

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

> **Important**
>
> You can only create one Microsoft Purview account per tenant. If your organization already has a Microsoft Purview account, you won't be able to create a new Microsoft Purview account unless your organization already had multiple accounts and is still under the the preexisting quota. For more information, see the FAQ.

In this Quickstart, you'll create a Microsoft Purview account using Azure PowerShell/Azure CLI. PowerShell reference for Microsoft Purview is available, but this article will take you through all the steps needed to create an account with PowerShell.

The Microsoft Purview governance portal surfaces tools like the Microsoft Purview Data Map and Microsoft Purview Data Catalog that help you manage and govern your data landscape. By connecting to data across your on-premises, multi-cloud, and software-as-a-service (SaaS) sources, the Microsoft Purview Data Map creates an up-to-date map of your information. It identifies and classifies sensitive data, and provides end-to-end linage. Data consumers are able to discover data across your organization, and data administrators are able to audit, secure, and ensure right use of your data.

For more information about the classic governance capabilities of Microsoft Purview, see our governance solutions overview page.

Prerequisites

- If you don't have an Azure subscription, create a free subscription before you begin.

- A Microsoft Entra tenant associated with your subscription.
- The user account that you use to sign in to Azure must be a member of the *contributor* or *owner* role, or an *administrator* of the Azure subscription. To view the permissions that you have in the subscription, follow these steps:

    1. Go to the Azure portal
    2. Select your username in the upper-right corner.
    3. Select the ellipsis button ("...") for more options.
    4. Then select **My permissions**.
    5. If you have access to multiple subscriptions, select the appropriate subscription.

Sign in to Azure

Sign in to the Azure portal with your Azure account.

Install PowerShell

Install either Azure PowerShell or Azure CLI in your client machine to deploy the template: Command-line deployment

Create an account

1. Sign in with your Azure credential

    PowerShell

    Azure PowerShell

    ```
    Connect-AzAccount
    ```

2. If you have multiple Azure subscriptions, select the subscription you want to use:

    PowerShell

    Azure PowerShell

    ```
    Set-AzContext [SubscriptionID/SubscriptionName]
    ```

3. Create a resource group for your account. You can skip this step if you already have one:

PowerShell

Azure PowerShell

```
New-AzResourceGroup -Name myResourceGroup -Location 'East US'
```

4. Create or Deploy the account:

PowerShell

Use the New-AzPurviewAccount cmdlet to create the Microsoft Purview account:

Azure PowerShell

```
New-AzPurviewAccount -Name yourPurviewAccountName -
ResourceGroupName myResourceGroup -Location eastus -IdentityType
SystemAssigned -SkuCapacity 4 -SkuName Standard -
PublicNetworkAccess Enabled
```

5. The deployment command returns results. Look for `ProvisioningState` to see whether the deployment succeeded.

6. If you deployed the account using a service principal, instead of a user account, you'll also need to run the below command in the Azure CLI:

Azure CLI

```
az purview account add-root-collection-admin --account-name [Microsoft
Purview Account Name] --resource-group [Resource Group Name] --object-
id [User Object Id]
```

This command will grant the user account collection admin permissions on the root collection in your Microsoft Purview account. This allows the user to access the Microsoft Purview governance portal and add permission for other users. For more information about permissions in Microsoft Purview, see our permissions guide. For more information about collections, see our manage collections article.

Next steps

In this quickstart, you learned how to create a Microsoft Purview (formerly Azure Purview) account.

Follow these next articles to learn how to navigate the Microsoft Purview governance portal, create a collection, and grant access to the Microsoft Purview governance portal.

- How to use the Microsoft Purview governance portal
- Grant users permissions to the governance portal
- Create a collection

Feedback

Was this page helpful? Yes No

Quickstart: Create a Microsoft Purview (formerly Azure Purview) account using .NET SDK

Article • 02/26/2025

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

> **Important**
>
> You can only create one Microsoft Purview account per tenant. If your organization already has a Microsoft Purview account, you won't be able to create a new Microsoft Purview account unless your organization already had multiple accounts and is still under the the preexisting quota. For more information, see the FAQ.

In this quickstart, you'll use the .NET SDK to create a Microsoft Purview (formerly Azure Purview) account.

The Microsoft Purview governance portal surfaces tools like the Microsoft Purview Data Map and Microsoft Purview Data Catalog that help you manage and govern your data landscape. By connecting to data across your on-premises, multicloud, and software-as-a-service (SaaS) sources, the Microsoft Purview Data Map creates an up-to-date map of your information. It identifies and classifies sensitive data, and provides end-to-end linage. Data consumers are able to discover data across your organization, and data administrators are able to audit, secure, and ensure right use of your data.

For more information about the classic governance capabilities of Microsoft Purview, see our governance solutions overview page.

Prerequisites

- If you don't have an Azure subscription, create a free subscription before you begin.
- A Microsoft Entra tenant associated with your subscription.

- The user account that you use to sign in to Azure must be a member of the *contributor* or *owner* role, or an *administrator* of the Azure subscription. To view the permissions that you have in the subscription, follow these steps:

    1. Go to the Azure portal
    2. Select your username in the upper-right corner.
    3. Select the ellipsis button ("...") for more options.
    4. Then select **My permissions**.
    5. If you have access to multiple subscriptions, select the appropriate subscription.

Sign in to Azure

Sign in to the Azure portal with your Azure account.

Visual Studio

The walkthrough in this article uses Visual Studio 2019. The procedures for Visual Studio 2013, 2015, or 2017 may differ slightly.

Azure .NET SDK

Download and install Azure .NET SDK on your machine.

Create an application in Microsoft Entra ID

1. In Create a Microsoft Entra application, create an application that represents the .NET application you're creating in this tutorial. For the sign-on URL, you can provide a dummy URL as shown in the article (`https://contoso.org/exampleapp`).
2. In Get values for signing in, get the **application ID** and **tenant ID**, and note down these values that you use later in this tutorial.
3. In Certificates and secrets, get the **authentication key**, and note down this value that you use later in this tutorial.
4. In Assign the application to a role, assign the application to the **Contributor** role at the subscription level so that the application can create data factories in the subscription.

Create a Visual Studio project

Next, create a C# .NET console application in Visual Studio:

1. Launch **Visual Studio**.
2. In the Start window, select **Create a new project** > **Console App (.NET Framework)**. .NET version 4.5.2 or above is required.
3. In **Project name**, enter **PurviewQuickStart**.
4. Select **Create** to create the project.

Install NuGet packages

1. Select **Tools** > **NuGet Package Manager** > **Package Manager Console**.

2. In the **Package Manager Console** pane, run the following commands to install packages. For more information, see the Microsoft.Azure.Management.Purview NuGet package.

```
PowerShell
```

```
Install-Package Microsoft.Azure.Management.Purview
Install-Package Microsoft.Azure.Management.ResourceManager -
IncludePrerelease
Install-Package Microsoft.IdentityModel.Clients.ActiveDirectory
```

Tip

If you are getting an error that reads: **Package <package name> is not found in the following primary sources(s):** and it is listing a local folder, you need to update your package sources in Visual Studio to include the nuget site as an online source.

1. Go to **Tools**
2. Select **NuGet Package Manager**
3. Select **Package Manage Settings**
4. Select **Package Sources**
5. Add https://nuget.org/api/v2/ as a source.

Create a Microsoft Purview client

1. Open **Program.cs**, include the following statements to add references to namespaces.

```
C#
```

```
using System;
using System.Collections.Generic;
using System.Linq;
using Microsoft.Rest;
using Microsoft.Rest.Serialization;
   using Microsoft.Azure.Management.ResourceManager;
using Microsoft.Azure.Management.Purview;
   using Microsoft.Azure.Management.Purview.Models;
   using Microsoft.IdentityModel.Clients.ActiveDirectory;
```

2. Add the following code to the **Main** method that sets the variables. Replace the placeholders with your own values. For a list of Azure regions in which Microsoft Purview is currently available, search on **Microsoft Purview** and select the regions that interest you on the following page: Products available by region.

C#

```
// Set variables
string tenantID = "<your tenant ID>";
string applicationId = "<your application ID>";
string authenticationKey = "<your authentication key for the
application>";
string subscriptionId = "<your subscription ID where the data factory
resides>";
string resourceGroup = "<your resource group where the data factory
resides>";
string region = "<the location of your resource group>";
string purviewAccountName =
    "<specify the name of purview account to create. It must be
globally unique.>";
```

3. Add the following code to the **Main** method that creates an instance of **PurviewManagementClient** class. You use this object to create a Microsoft Purview Account.

C#

```
// Authenticate and create a purview management client
var context = new AuthenticationContext("https://login.windows.net/" +
tenantID);
ClientCredential cc = new ClientCredential(applicationId,
authenticationKey);
AuthenticationResult result = context.AcquireTokenAsync(
"https://management.azure.com/", cc).Result;
ServiceClientCredentials cred = new
TokenCredentials(result.AccessToken);
var client = new PurviewManagementClient(cred)
{
```

```
    SubscriptionId = subscriptionId
};
```

Create an account

Add the following code to the **Main** method that will create the **Microsoft Purview Account**.

C#

```
// Create a purview Account
Console.WriteLine("Creating Microsoft Purview Account " + purviewAccountName
+ "...");
Account account = new Account()
{
Location = region,
Identity = new Identity(type: "SystemAssigned"),
Sku = new AccountSku(name: "Standard", capacity: 4)
};
try
{
  client.Accounts.CreateOrUpdate(resourceGroup, purviewAccountName,
account);
  Console.WriteLine(client.Accounts.Get(resourceGroup,
purviewAccountName).ProvisioningState);
}
catch (ErrorResponseModelException purviewException)
{
Console.WriteLine(purviewException.StackTrace);
  }
  Console.WriteLine(
    SafeJsonConvert.SerializeObject(account, client.SerializationSettings));
  while (client.Accounts.Get(resourceGroup,
purviewAccountName).ProvisioningState ==
         "PendingCreation")
  {
    System.Threading.Thread.Sleep(1000);
  }
Console.WriteLine("\nPress any key to exit...");
Console.ReadKey();
```

Run the code

Build and start the application, then verify the execution.

The console prints the progress of creating Microsoft Purview Account.

Sample output

JSON

```
Creating Microsoft Purview Account testpurview...
Succeeded
{
  "sku": {
    "capacity": 4,
    "name": "Standard"
  },
  "identity": {
    "type": "SystemAssigned"
  },
  "location": "southcentralus"
}

Press any key to exit...
```

Verify the output

Go to the **Microsoft Purview accounts** page in the Azure portal and verify the account created using the above code.

Delete Microsoft Purview account

To programmatically delete a Microsoft Purview account, add the following lines of code to the program:

C#

```
Console.WriteLine("Deleting the Microsoft Purview Account");
client.Accounts.Delete(resourceGroup, purviewAccountName);
```

Check if Microsoft Purview account name is available

To check availability of a purview account, use the following code:

C#

```
CheckNameAvailabilityRequest checkNameAvailabilityRequest =
newCheckNameAvailabilityRequest()
{
```

```
    Name = purviewAccountName,
    Type =  "Microsoft.Purview/accounts"
};
Console.WriteLine("Check Microsoft Purview account name");
Console.WriteLine(client.Accounts.CheckNameAvailability(checkNameAvailabilit
yRequest).NameAvailable);
```

The above code with print 'True' if the name is available and 'False' if the name isn't available.

Next steps

In this quickstart, you learned how to create a Microsoft Purview (formerly Azure Purview) account, delete the account, and check for name availability. You can now download the .NET SDK and learn about other resource provider actions you can perform for a Microsoft Purview account.

Follow these next articles to learn how to navigate the Microsoft Purview governance portal, create a collection, and grant access to the Microsoft Purview governance portal.

- How to use the Microsoft Purview governance portal
- Grant users permissions to the governance portal
- Create a collection

Feedback

Was this page helpful? Yes No

Quickstart: Create a Microsoft Purview (formerly Azure Purview) account using Python

Article • 01/28/2025

> **Note**
>
> Microsoft Purview Data Map (classic), Data Catalog (classic), Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

> **Important**
>
> You can only create one Microsoft Purview account per tenant. If your organization already has a Microsoft Purview account, you will not be able to create a new Microsoft Purview account unless your organization already had multiple accounts and is still under the the pre-existing quota. For more information, see the FAQ.

In this quickstart, you'll create a Microsoft Purview (formerly Azure Purview) account programatically using Python. The Python reference for Microsoft Purview is available, but this article will take you through all the steps needed to create an account with Python.

The Microsoft Purview governance portal surfaces tools like the Microsoft Purview Data Map and Microsoft Purview Data Catalog that help you manage and govern your data landscape. By connecting to data across your on-premises, multi-cloud, and software-as-a-service (SaaS) sources, the Microsoft Purview Data Map creates an up-to-date map of your information. It identifies and classifies sensitive data, and provides end-to-end linage. Data consumers are able to discover data across your organization, and data administrators are able to audit, secure, and ensure right use of your data.

For more information about the classic governance capabilities of Microsoft Purview, see our governance solutions overview page.

Prerequisites

- If you don't have an Azure subscription, create a free subscription before you begin.
- A Microsoft Entra tenant associated with your subscription.
- The user account that you use to sign in to Azure must be a member of the *contributor* or *owner* role, or an *administrator* of the Azure subscription. To view the permissions that you have in the subscription, follow these steps:

    1. Go to the Azure portal
    2. Select your username in the upper-right corner.
    3. Select the ellipsis button ("...") for more options.
    4. Then select **My permissions**.
    5. If you have access to multiple subscriptions, select the appropriate subscription.

Sign in to Azure

Sign in to the Azure portal with your Azure account.

Install the Python package

1. Open a terminal or command prompt with administrator privileges.
2. First, install the Python package for Azure management resources:

    ```
    Python
    ```

    ```
    pip install azure-mgmt-resource
    ```

3. To install the Python package for Microsoft Purview, run the following command:

    ```
    Python
    ```

    ```
    pip install azure-mgmt-purview
    ```

    The Python SDK for Microsoft Purview supports Python 2.7, 3.3, 3.4, 3.5, 3.6 and 3.7.

4. To install the Python package for Azure Identity authentication, run the following command:

    ```
    Python
    ```

```
pip install azure-identity
```

> **Note**
>
> The "azure-identity" package might have conflicts with "azure-cli" on some common dependencies. If you meet any authentication issue, remove "azure-cli" and its dependencies, or use a clean machine without installing "azure-cli" package.

Create a purview client

1. Create a file named **purview.py**. Add the following statements to add references to namespaces.

```Python
from azure.identity import ClientSecretCredential
 from azure.mgmt.resource import ResourceManagementClient
 from azure.mgmt.purview import PurviewManagementClient
 from azure.mgmt.purview.models import *
 from datetime import datetime, timedelta
 import time
```

2. Add the following code to the **Main** method that creates an instance of PurviewManagementClient class. You'll use this object to create a purview account, delete purview accounts, check name availability, and other resource provider operations.

```Python
def main():

# Azure subscription ID
subscription_id = '<subscription ID>'

 # This program creates this resource group. If it's an existing resource group, comment out the code that creates the resource group
rg_name = '<resource group>'

# The purview name. It must be globally unique.
purview_name = '<purview account name>'

# Location name, where Microsoft Purview account must be created.
location = '<location name>'
```

```
# Specify your Active Directory client ID, client secret, and tenant ID
credentials = ClientSecretCredential(client_id='<service principal
ID>', client_secret='<service principal key>', tenant_id='<tenant ID>')
# resource_client = ResourceManagementClient(credentials,
subscription_id)
purview_client = PurviewManagementClient(credentials, subscription_id)
```

Create a purview account

1. Add the following code to the **Main** method that creates a **purview account**. If your resource group already exists, comment out the first `create_or_update` statement.

Python

```
 # create the resource group
 # comment out if the resource group already exits
 resource_client.resource_groups.create_or_update(rg_name, rg_params)

 #Create a purview
 identity = Identity(type= "SystemAssigned")
 sku = AccountSku(name= 'Standard', capacity= 4)
 purview_resource = Account(identity=identity,sku=sku,location
=location )

 try:
     pa = (purview_client.accounts.begin_create_or_update(rg_name,
purview_name, purview_resource)).result()
     print("location:", pa.location, " Microsoft Purview Account Name:
", pa.name, " Id: " , pa.id ," tags: " , pa.tags)
 except:
     print("Error")
     print_item(pa)

 while (getattr(pa,'provisioning_state')) != "Succeeded" :
     pa = (purview_client.accounts.get(rg_name, purview_name))
     print(getattr(pa,'provisioning_state'))
     if getattr(pa,'provisioning_state') == "Failed" :
         print("Error in creating Microsoft Purview account")
         break
     time.sleep(30)
```

2. Now, add the following statement to invoke the **main** method when the program is run:

Python

```
# Start the main method
main()
```

Full script

Here's the full Python code:

Python

```
from azure.identity import ClientSecretCredential
from azure.mgmt.resource import ResourceManagementClient
from azure.mgmt.purview import PurviewManagementClient
from azure.mgmt.purview.models import *
from datetime import datetime, timedelta
import time

def main():

    # Azure subscription ID
    subscription_id = '<subscription ID>'

    # This program creates this resource group. If it's an existing resource
group, comment out the code that creates the resource group
    rg_name = '<resource group>'

    # The purview name. It must be globally unique.
    purview_name = '<purview account name>'

    # Location for your resource group and your Microsoft Purview account.
    location ="<location>"

    # Specify your Active Directory client ID, client secret, and tenant ID
    credentials = ClientSecretCredential(client_id='<service principal ID>',
client_secret='<service principal key>', tenant_id='<tenant ID>')
    resource_client = ResourceManagementClient(credentials, subscription_id)
    purview_client = PurviewManagementClient(credentials, subscription_id)

    # create the resource group
    # comment out if the resource group already exits
    resource_client.resource_groups.create_or_update(rg_name, {"location":
location})

    #Create a purview
    identity = Identity(type= "SystemAssigned")
    sku = AccountSku(name= 'Standard', capacity= 4)
    purview_resource = Account(identity=identity,sku=sku,location =location)

    try:
        pa = (purview_client.accounts.begin_create_or_update(rg_name,
purview_name, purview_resource)).result()
```

```
        print("location:", pa.location, " Microsoft Purview Account Name: ",
purview_name, " Id: " , pa.id ," tags: " , pa.tags)
    except:
        print("Error in submitting job to create account")
        print_item(pa)

    while (getattr(pa,'provisioning_state')) != "Succeeded" :
        pa = (purview_client.accounts.get(rg_name, purview_name))
        print(getattr(pa,'provisioning_state'))
        if getattr(pa,'provisioning_state') == "Failed" :
            print("Error in creating Microsoft Purview account")
            break
        time.sleep(30)

# Start the main method
main()
```

Run the code

Build and start the application. The console prints the progress of Microsoft Purview account creation. Wait until it's completed. Here's the sample output:

Console

```
location: southcentralus  Microsoft Purview Account Name:  purviewpython7
Id:  /subscriptions/8c2c7b23-848d-40fe-b817-
690d79ad9dfd/resourceGroups/Demo_Catalog/providers/Microsoft.Purview/account
s/purviewpython7  tags:  None
Creating
Creating
Succeeded
```

Verify the output

Go to the **Microsoft Purview accounts** page in the Azure portal and verify the account created using the above code.

Delete Microsoft Purview account

To delete purview account, add the following code to the program, then run:

Python

```
pa = purview_client.accounts.begin_delete(rg_name, purview_name).result()
```

Next steps

In this quickstart, you learned how to create a Microsoft Purview (formerly Azure Purview) account, delete the account, and check for name availability. You can now download the Python SDK and learn about other resource provider actions you can perform for a Microsoft Purview account.

Follow these next articles to learn how to navigate the Microsoft Purview governance portal, create a collection, and grant access to the Microsoft Purview governance portal.

- How to use the Microsoft Purview governance portal
- Grant users permissions to the governance portal
- Create a collection

Feedback

Was this page helpful? Yes No

Quickstart: Create a Microsoft Purview (formerly Azure Purview) account using a Bicep file

Article • 02/26/2025

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

> **Important**
>
> You can only create one Microsoft Purview account per tenant. If your organization already has a Microsoft Purview account, you won't be able to create a new Microsoft Purview account unless your organization already had multiple accounts and is still under the the preexisting quota. For more information, see the FAQ.

This quickstart describes the steps to deploy a Microsoft Purview (formerly Azure Purview) account using a Bicep file.

After you've created the account, you can begin registering your data sources, and using the Microsoft Purview governance portal to understand and govern your data landscape. By connecting to data across your on-premises, multicloud, and software-as-a-service (SaaS) sources, the Microsoft Purview Data Map creates an up-to-date map of your information. It identifies and classifies sensitive data, and provides end-to-end data linage. Data consumers are able to discover data across your organization and data administrators are able to audit, secure, and ensure right use of your data.

For more information about the classic governance capabilities of Microsoft Purview, see our governance solutions overview page.

To deploy a Microsoft Purview account to your subscription using a Bicep file, follow the guide below.

Prerequisites

- If you don't have an Azure subscription, create a free subscription before you begin.
- A Microsoft Entra tenant associated with your subscription.
- The user account that you use to sign in to Azure must be a member of the *contributor* or *owner* role, or an *administrator* of the Azure subscription. To view the permissions that you have in the subscription, follow these steps:

    1. Go to the Azure portal
    2. Select your username in the upper-right corner.
    3. Select the ellipsis button ("...") for more options.
    4. Then select **My permissions**.
    5. If you have access to multiple subscriptions, select the appropriate subscription.

Sign in to Azure

Sign in to the Azure portal with your Azure account.

Review the Bicep file

The Bicep file used in this quickstart is from Azure Quickstart Templates.

Bicep

```
@description('Specify a name for the Azure Purview account.')
param purviewName string = 'azurePurview${uniqueString(resourceGroup().id)}'

@description('Specify a region for resource deployment.')
param location string = resourceGroup().location

resource purview 'Microsoft.Purview/accounts@2021-12-01' = {
  name: purviewName
  location: location
  sku: {
    name: 'Standard'
    capacity: 1
  }
  identity: {
    type: 'SystemAssigned'
  }
  properties: {
    publicNetworkAccess: 'Enabled'
    managedResourceGroupName: 'managed-rg-${purviewName}'
```

```
  }
}
```

The following resources are defined in the Bicep file:

- Microsoft.Purview/accounts

The Bicep file performs the following tasks:

- Creates a Microsoft Purview account in a specified resource group.

Deploy the Bicep file

1. Save the Bicep file as **main.bicep** to your local computer.
2. Deploy the Bicep file using either Azure CLI or Azure PowerShell.

CLI

Azure CLI

```
az group create --name exampleRG --location eastus
az deployment group create --resource-group exampleRG --template-
file main.bicep
```

You'll be prompted to enter the following values:

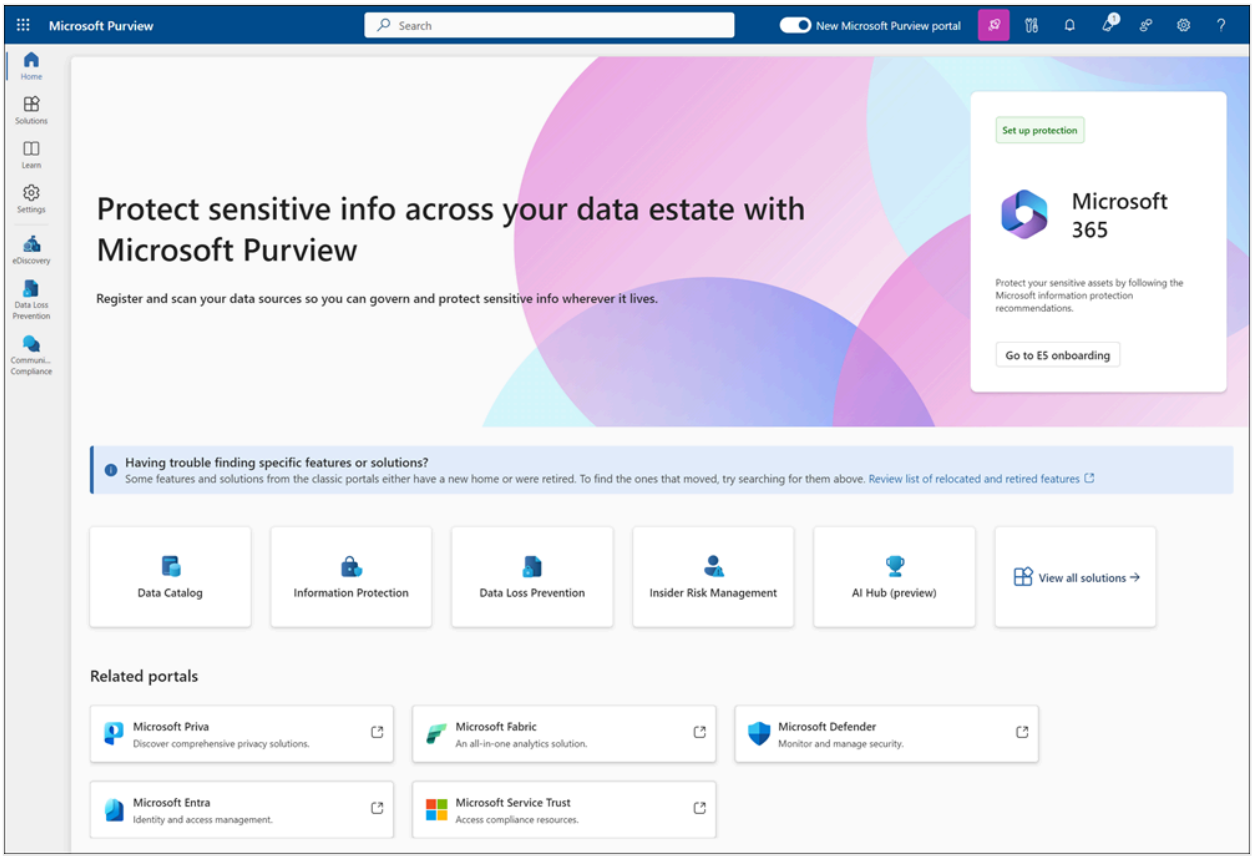
- **Purview name**: enter a name for the Microsoft Purview account.

When the deployment finishes, you should see a message indicating the deployment succeeded.

Open the Microsoft Purview governance portal

After your account is created, you can use the Microsoft Purview portal to access and manage it.

In a browser, navigate directly to https://purview.microsoft.com to access the portal.



For more information about using the portal, see our documentation on the Microsoft Purview portal.

Get started with your Purview resource

After deployment, the first activities are usually:

- Create a collection
- Register a resource
- Scan the resource

At this time, these actions aren't able to be taken through a Bicep file. Follow the guides above to get started!

Clean up resources

When you no longer need them, use the Azure portal, Azure CLI, or Azure PowerShell to remove the resource group, firewall, and all related resources.

CLI

Azure CLI

```
az group delete --name exampleRG
```

Next steps

In this quickstart, you learned how to create a Microsoft Purview (formerly Azure Purview) account and how to access the Microsoft Purview governance portal.

Next, you can create a user-assigned managed identity (UAMI) that will enable your new Microsoft Purview account to authenticate directly with resources using Microsoft Entra authentication.

To create a UAMI, follow our guide to create a user-assigned managed identity.

Follow these next articles to learn how to navigate the Microsoft Purview governance portal, create a collection, and grant access to Microsoft Purview:

Using the Microsoft Purview governance portal | Create a collection

Add users to your Microsoft Purview account

Feedback

Was this page helpful? Yes | No

Quickstart: Create a Microsoft Purview (formerly Azure Purview) account using an ARM template

Article • 05/20/2025

> **Important**
>
> You can only create one Microsoft Purview account per tenant. If your organization already has a Microsoft Purview account, you won't be able to create a new Microsoft Purview account unless your organization already had multiple accounts and is still under the the preexisting quota. For more information, see the FAQ.

This quickstart describes the steps to deploy a Microsoft Purview (formerly Azure Purview) account using an Azure Resource Manager (ARM) template.

After you've created the account, you can begin registering your data sources and using the Microsoft Purview governance portal to understand and govern your data landscape. By connecting to data across your on-premises, multicloud, and software-as-a-service (SaaS) sources, the Microsoft Purview Data Map creates an up-to-date map of your information. It identifies and classifies sensitive data, and provides end-to-end data linage. Data consumers are able to discover data across your organization and data administrators are able to audit, secure, and ensure right use of your data.

For more information about the classic governance capabilities of Microsoft Purview, see our governance solutions overview page.

To deploy a Microsoft Purview account to your subscription using an ARM template, follow the guide below.

Prerequisites

- If you don't have an Azure subscription, create a free subscription before you begin.
- A Microsoft Entra tenant associated with your subscription.
- The user account that you use to sign in to Azure must be a member of the *contributor* or *owner* role, or an *administrator* of the Azure subscription. To view the permissions that you have in the subscription, follow these steps:

    1. Go to the Azure portal

2. Select your username in the upper-right corner.
3. Select the ellipsis button ("...") for more options.
4. Then select **My permissions**.
5. If you have access to multiple subscriptions, select the appropriate subscription.

Sign in to Azure

Sign in to the Azure portal with your Azure account.

Deploy a custom template

If your environment meets the prerequisites and you're familiar with using ARM templates, select the **Deploy to Azure** button. The template will open in the Azure portal where you can customize values and deploy. The template will deploy a Microsoft Purview account into a new or existing resource group in your subscription.



Review the template

The template used in this quickstart is from Azure Quickstart Templates.

JSON

```
{
  "$schema": "https://schema.management.azure.com/schemas/2019-04-
01/deploymentTemplate.json#",
  "contentVersion": "1.0.0.0",
  "metadata": {
    "_generator": {
      "name": "bicep",
      "version": "0.11.1.770",
      "templateHash": "10388724958841997784"
    }
  },
  "parameters": {
    "purviewName": {
      "type": "string",
      "defaultValue": "[format('azurePurview{0}',
uniqueString(resourceGroup().id))]",
      "metadata": {
        "description": "Specify a name for the Azure Purview account."
      }
    },
    "location": {
```

```
      "type": "string",
      "defaultValue": "[resourceGroup().location]",
      "metadata": {
        "description": "Specify a region for resource deployment."
      }
    }
  },
  "resources": [
    {
      "type": "Microsoft.Purview/accounts",
      "apiVersion": "2021-12-01",
      "name": "[parameters('purviewName')]",
      "location": "[parameters('location')]",
      "sku": {
        "name": "Standard",
        "capacity": 1
      },
      "identity": {
        "type": "SystemAssigned"
      },
      "properties": {
        "publicNetworkAccess": "Enabled",
        "managedResourceGroupName": "[format('managed-rg-{0}',
parameters('purviewName'))]"
      }
    }
  ]
}
```

The following resources are defined in the template:

- Microsoft.Purview/accounts

The template performs the following tasks:

- Creates a Microsoft Purview account in a specified resource group.

Customize network settings for your account

When you're deploying your ARM template, you can also use the following settings in the template to manage your public network access settings:

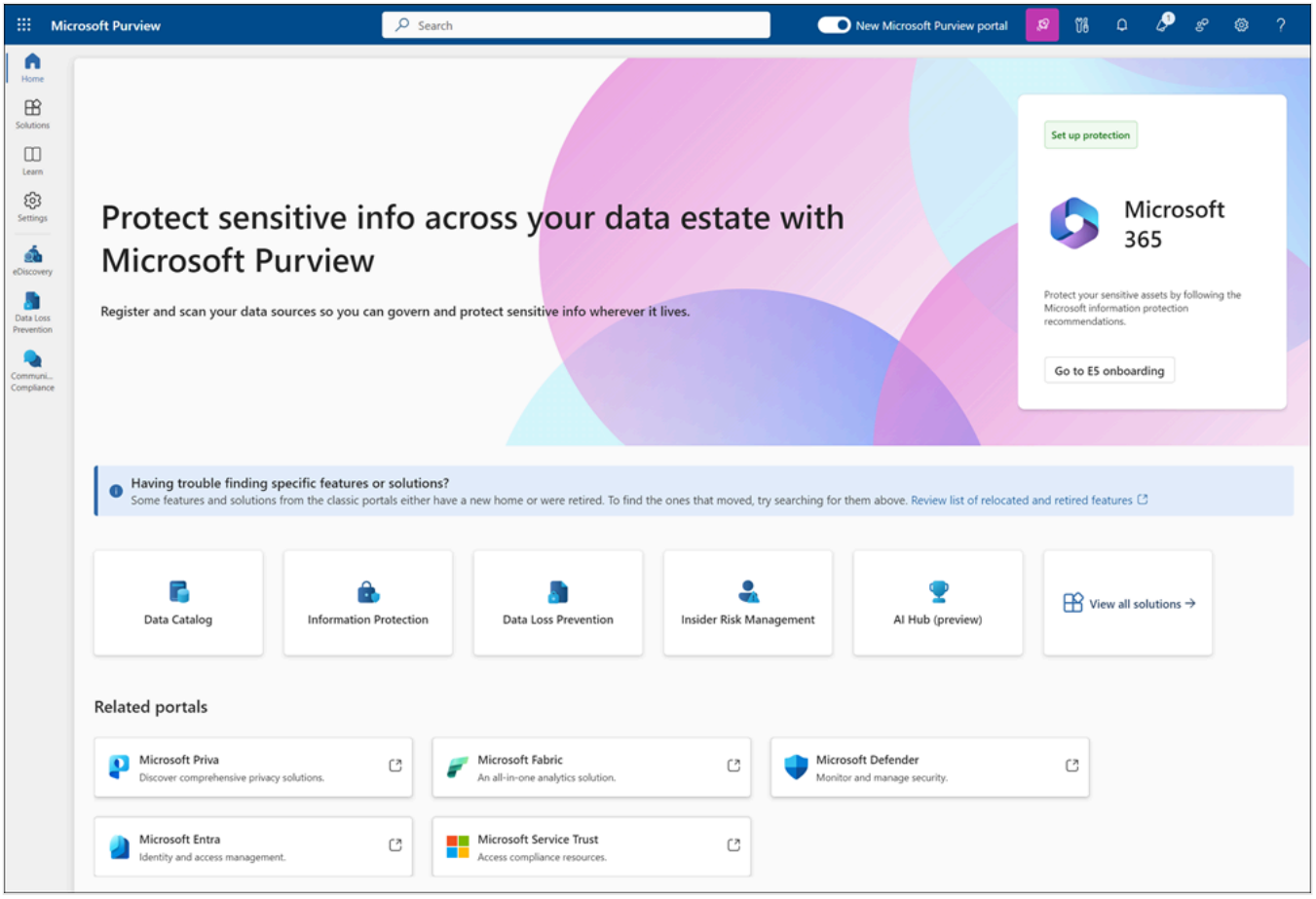
- **Enabled for all networks** `"publicNetworkAccess": "Enabled", "managedResourcesPublicNetworkAccess": "Enabled"`
- **Disabled for ingestion only (Preview)** `"publicNetworkAccess": "Enabled", "managedResourcesPublicNetworkAccess": "Disabled"`
- **Disabled from all networks** `"publicNetworkAccess": "Disables", "managedResourcesPublicNetworkAccess": "Disabled"`

For example: `"resources": [{ "type": "Microsoft.Purview/accounts", "apiVersion": "2021-12-01", "name": "[parameters('purviewName')]", "location": "[parameters('location')]", "sku": { "name": "Standard", "capacity": 1 }, "identity": { "type": "SystemAssigned" }, "properties": { "publicNetworkAccess": "Enabled", "managedResourcesPublicNetworkAccess": "Enabled" "managedResourceGroupName": "[format('managed-rg-{0}', parameters('purviewName'))]" } }]`

Open the Microsoft Purview governance portal

After your account is created, you can use the Microsoft Purview portal to access and manage it.

In a browser, navigate directly to https://purview.microsoft.com to access the portal.



For more information about using the portal, see our documentation on the Microsoft Purview portal.

Get started with your Purview resource

After deployment, the first activities are usually:

- Create a collection

- Register a resource
- Scan the resource

At this time, these actions aren't able to be taken through an Azure Resource Manager template. Follow the guides above to get started!

Clean up resources

To clean up the resources deployed in this quickstart, delete the resource group, which deletes all resources in the group. You can delete the resources either through the Azure portal, or using the PowerShell script below.

Azure PowerShell

```
$resourceGroupName = Read-Host -Prompt "Enter the resource group name"
Remove-AzResourceGroup -Name $resourceGroupName
Write-Host "Press [ENTER] to continue..."
```

Next steps

In this quickstart, you learned how to create a Microsoft Purview (formerly Azure Purview) account and how to access the Microsoft Purview governance portal.

Next, you can create a user-assigned managed identity (UAMI) that will enable your new Microsoft Purview account to authenticate directly with resources using Microsoft Entra authentication.

To create a UAMI, follow our guide to create a user-assigned managed identity.

Follow these next articles to learn how to navigate the Microsoft Purview governance portal, create a collection, and grant access to Microsoft Purview:

Use the Microsoft Purview portal

Get started with data governance in Microsoft Purview

Add users to your Microsoft Purview account

Asset certification in the Microsoft Purview Data Catalog

Article • 01/28/2025

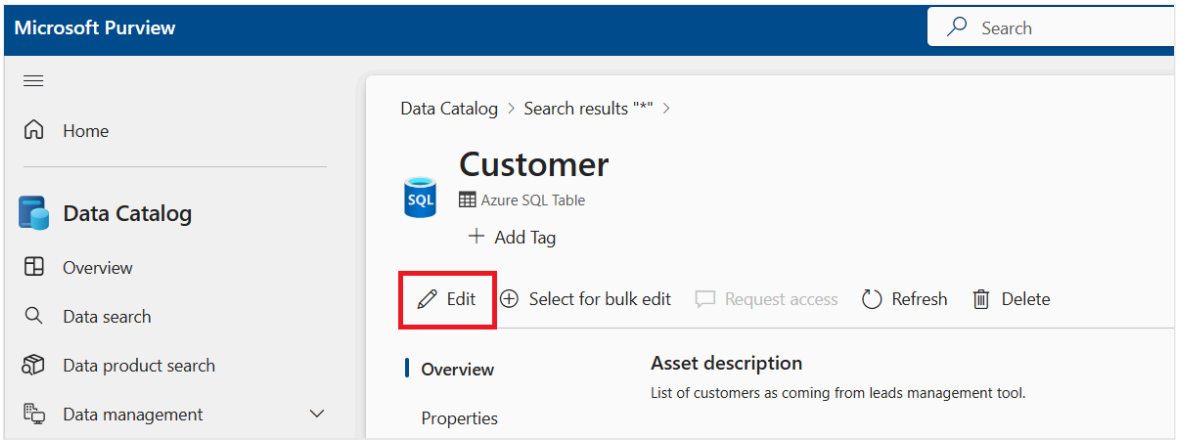
> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

As a Microsoft Purview Data Catalog grows in size, it becomes important for data consumers to understand what assets they can trust. Data consumers must know if an asset meet their organization's quality standards and can be regarded as reliable. Microsoft Purview allows data stewards to manually endorse assets to indicate that they're ready to use across an organization or business unit. This article describes how data stewards can certify assets and data consumers can view certification labels.

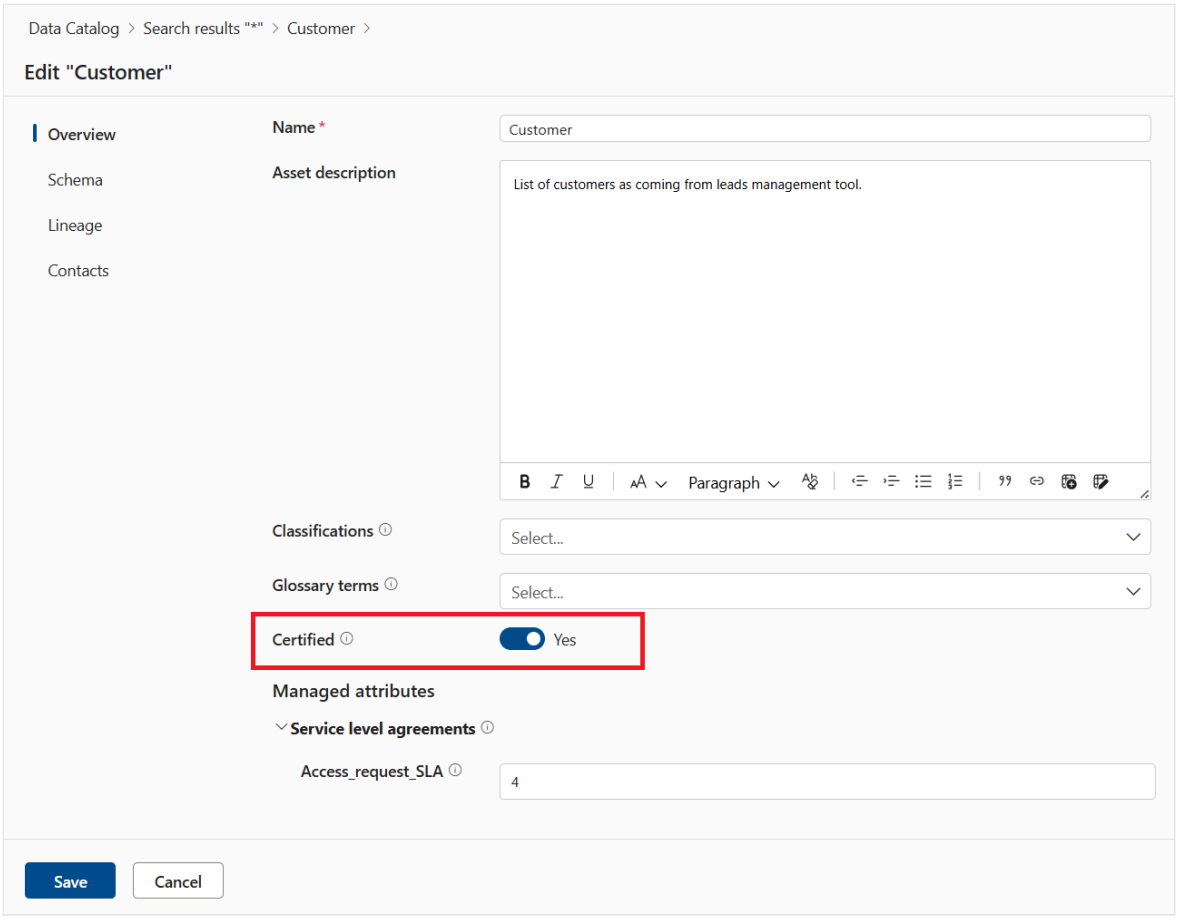
How to certify an asset

To certify an asset, you must be a **data curator** for the collection containing the asset. For steps to assign this role, see our permissions documentation.

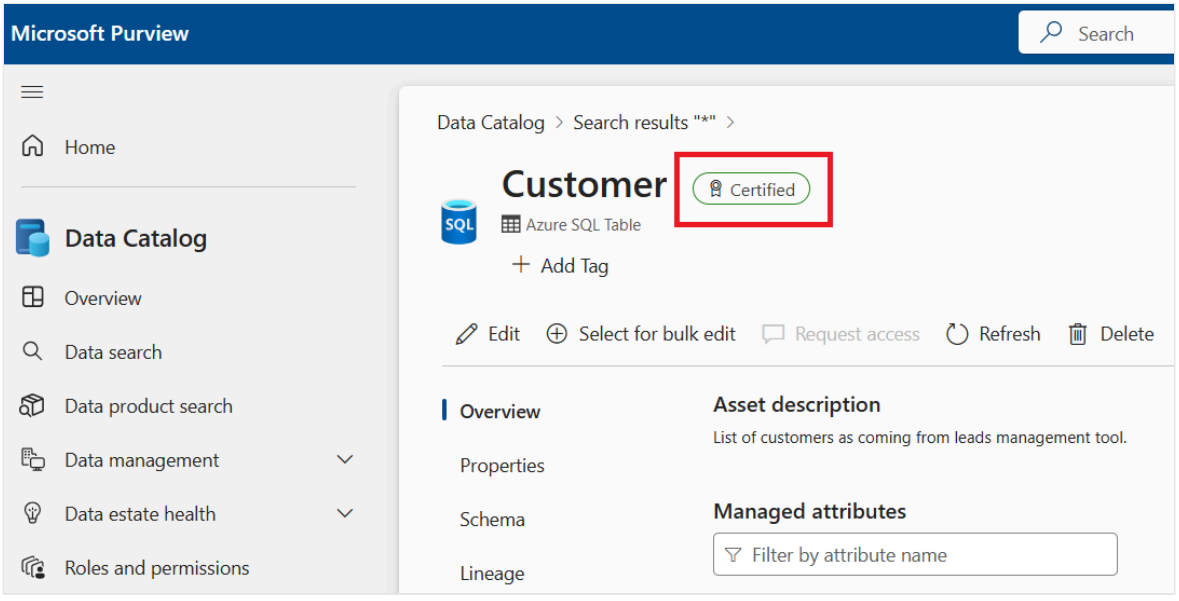
1. Navigate to the [asset details]classic-catalog-asset-details.md) of the desired asset. Select **Edit**.



2. Toggle the **Certified** field to **Yes**.



3. Save your changes. The asset has a "Certified" label next to the asset name.

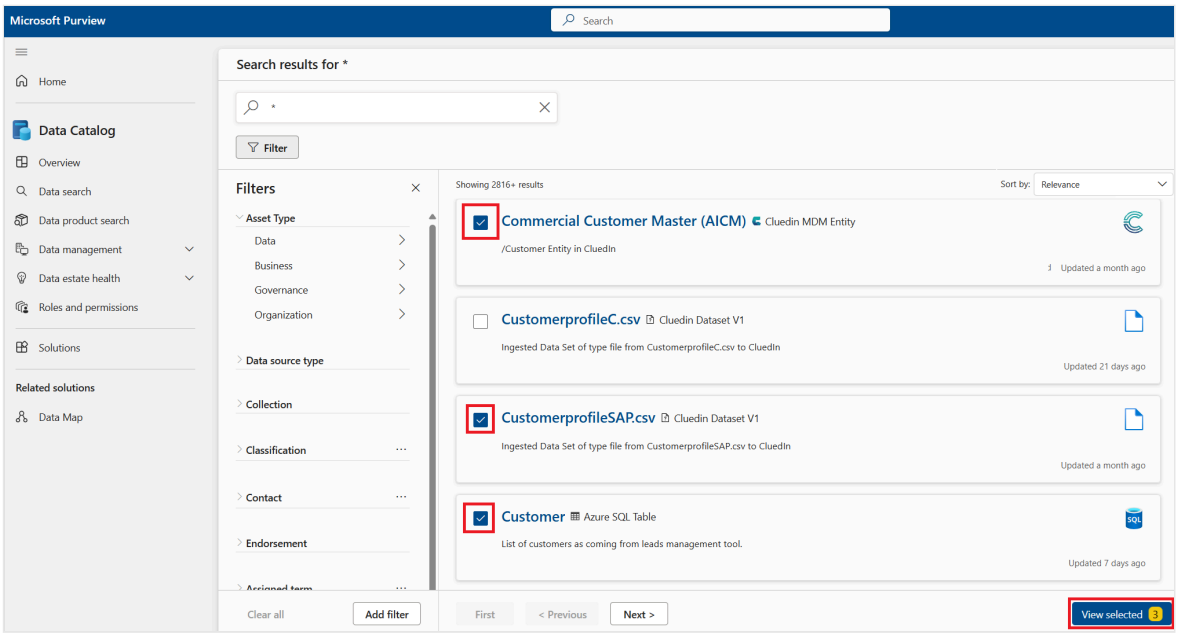


> **Note**
>
> PowerBI assets can only be certified in a PowerBI workspace. PowerBI endorsement labels are displayed in Microsoft Purview's search and browse experiences.

Certify assets in bulk

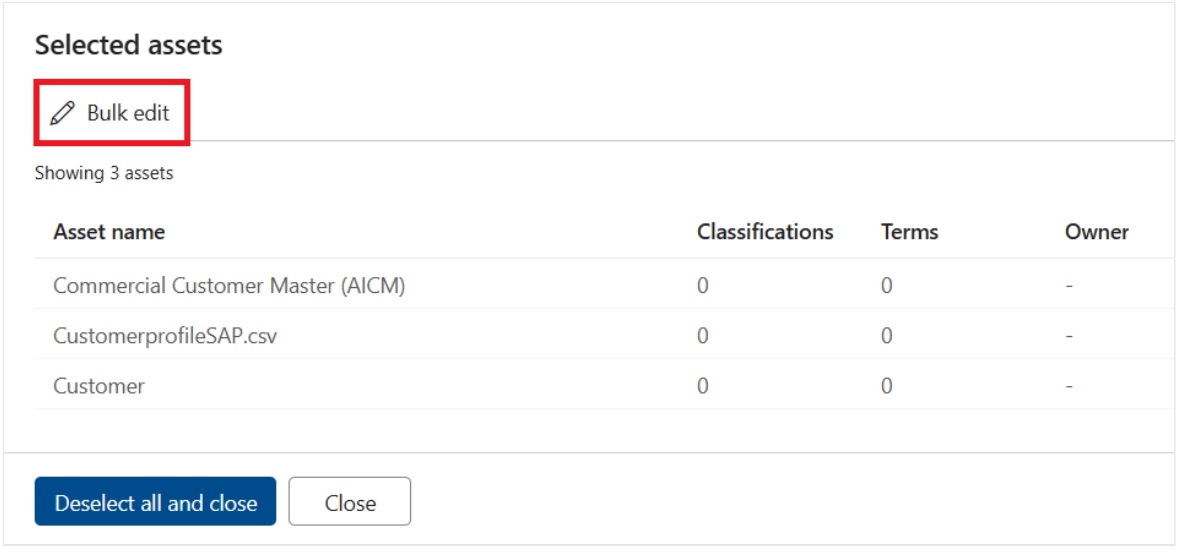
You can use the Microsoft Purview bulk edit experience to certify multiple assets at once.

1. After searching or browsing the data catalog, select checkbox next to the assets you wish to certify.

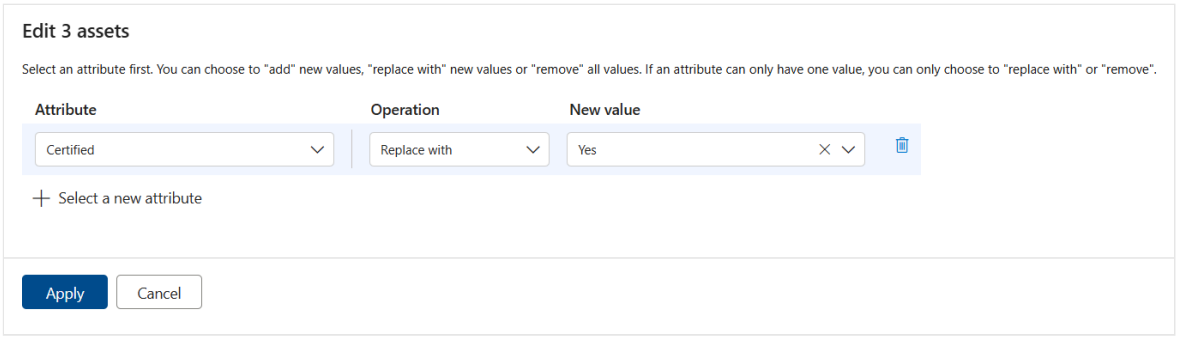


2. Select **View selected**.

3. Select **Bulk edit**



4. Choose attribute **Certified**, operation **Replace with**, and new value **Yes**.

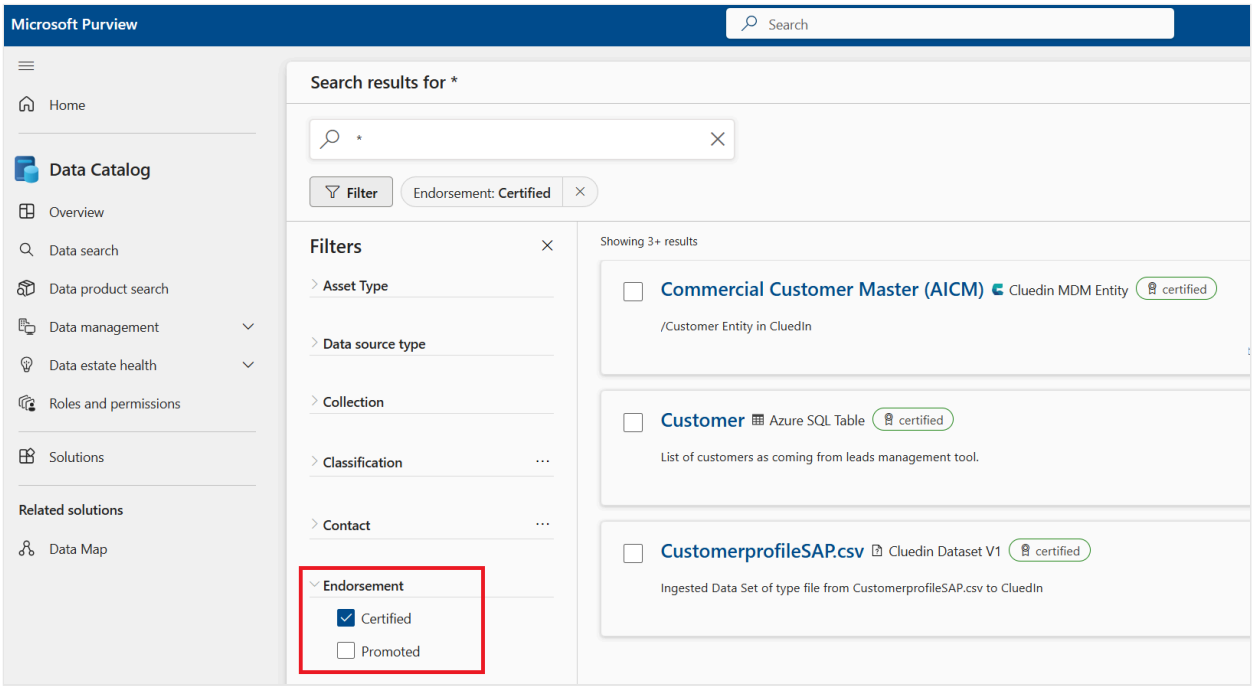


5. Select **Apply**

All assets selected have the "Certified" label.

Viewing certification labels in Search

When search or browsing the data catalog, you see a certification label on any asset that is certified. Certified assets boosted in search results to help data consumers discover them easily.



Next steps

Discover your assets in the Microsoft Purview Data Catalog by either:

- Browsing the data catalog
- Searching the data catalog

Feedback

Was this page helpful? Yes No

Asset normalization

Article • 05/20/2025

When ingesting assets into the Microsoft Purview Data Map, different sources updating the same data asset may send similar, but slightly different qualified names. While these qualified names represent the same asset, slight differences such as an extra character may cause these assets on the surface to appear different and cause duplicate entries in Microsoft Purview. To avoid storing duplicate entries and causing confusion when consuming the Unified Catalog, Microsoft Purview automatically applies normalization during ingestion to ensure all fully qualified names of the same entity type are in the same format.

For example, you scan in an Azure Blob with the qualified name `https://myaccount.file.core.windows.net/myshare/folderA/folderB/my-file.parquet`. This blob is also consumed by an Azure Data Factory pipeline that will then add lineage information to the asset. The ADF (Azure Data Factory) pipeline may be configured to read the file as `https://myAccount.file.core.windows.net//myshare/folderA/folderB/my-file.parquet`. While the qualified name is different, this ADF pipeline is consuming the same piece of data. Normalization ensures that all the metadata from both Azure Blob Storage and Azure Data Factory is visible on a single asset, `https://myaccount.file.core.windows.net/myshare/folderA/folderB/my-file.parquet`.

> **Important**
>
> The rules listed below are the only kinds of potential duplication Microsoft Purview currently recognizes. If you are experiencing accidental asset duplication, compare the assets fully qualified names to check for capitalization differences or extra characters. Update any ingestion points, for example your ADF pipelines, so that the qualified names match.

Normalization rules

These are the normalization rules that Microsoft Purview automatically applies.

Encode curly brackets

Applies to: All Assets

Before: `https://myaccount.file.core.windows.net/myshare/{folderA}/folder{B/`

After: `https://myaccount.file.core.windows.net/myshare/%7BfolderA%7D/folder%7BB/`

Trim section spaces

Applies to: Azure Blob, Azure Files, Azure Data Lake Storage Gen1, Azure Data Lake Storage Gen2, Azure Data Factory, Azure SQL Database, Azure SQL Managed Instance, Azure SQL pool, Azure Cosmos DB, Azure Cognitive Search, Azure Data Explorer, Azure Data Share, Amazon S3

Before: `https://myaccount.file.core.windows.net/myshare/ folder A/folderB /`

After: `https://myaccount.file.core.windows.net/myshare/folder A/folderB/`

Remove hostname spaces

Applies to: Azure Blob, Azure Files, Azure Data Lake Storage Gen1, Azure Data Lake Storage Gen2, Azure SQL Database, Azure SQL Managed Instance, Azure SQL pool, Azure Cosmos DB, Azure Cognitive Search, Azure Data Explorer, Azure Data Share, Amazon S3

Before: `https://myaccount .file. core.win dows. net/myshare/folderA/folderB/`

After: `https://myaccount.file.core.windows.net/myshare/folderA/folderB/`

Remove square brackets

Applies to: Azure SQL Database, Azure SQL Managed Instance, Azure SQL pool

Before: `mssql://foo.database.windows.net/[bar]/dbo/[foo bar]`

After: `mssql://foo.database.windows.net/bar/dbo/foo%20bar`

> ⓘ **Note**
>
> Spaces between two square brackets will be encoded

Lowercase scheme

Applies to: Azure Blob, Azure Files, Azure Data Lake Storage Gen1, Azure Data Lake Storage Gen2, Azure SQL Database, Azure SQL Managed Instance, Azure SQL pool, Azure Cosmos DB, Azure Cognitive Search, Azure Data Explorer, Amazon S3

Before: `HTTPS://myaccount.file.core.windows.net/myshare/folderA/folderB/`

After: `https://myaccount.file.core.windows.net/myshare/folderA/folderB/`

Lowercase hostname

Applies to: Azure Blob, Azure Files, Azure Data Lake Storage Gen1, Azure Data Lake Storage Gen2, Azure SQL Database, Azure SQL Managed Instance, Azure SQL pool, Azure Cosmos DB, Azure Cognitive Search, Azure Data Explorer, Amazon S3

Before: `https://myAccount.file.Core.Windows.net/myshare/folderA/folderB/`

After: `https://myaccount.file.core.windows.net/myshare/folderA/folderB/`

Lowercase file extension

Applies to: Azure Blob, Azure Files, Azure Data Lake Storage Gen1, Azure Data Lake Storage Gen2, Amazon S3

Before: `https://myAccount.file.core.windows.net/myshare/folderA/data.TXT`

After: `https://myaccount.file.core.windows.net/myshare/folderA/data.txt`

Remove duplicate slash

Applies to: Azure Blob, Azure Files, Azure Data Lake Storage Gen1, Azure Data Lake Storage Gen2, Azure Data Factory, Azure SQL Database, Azure SQL Managed Instance, Azure SQL pool, Azure Cosmos DB, Azure Cognitive Search, Azure Data Explorer, Azure Data Share, Amazon S3

Before: `https://myAccount.file.core.windows.net//myshare/folderA////folderB/`

After: `https://myaccount.file.core.windows.net/myshare/folderA/folderB/`

Convert to ADL scheme

Applies to: Azure Data Lake Storage Gen1

Before: `https://mystore.azuredatalakestore.net/folderA/folderB/abc.csv`

After: `adl://mystore.azuredatalakestore.net/folderA/folderB/abc.csv`

Remove Trailing Slash

Remove the trailing slash from higher level assets for Azure Blob, ADLS Gen1, and ADLS Gen2.

Applies to: Azure Blob, Azure Data Lake Storage Gen1, Azure Data Lake Storage Gen2

Asset types: "azure_blob_container", "azure_blob_service", "azure_storage_account", "azure_datalake_gen2_service", "azure_datalake_gen2_filesystem", "azure_datalake_gen1_account".

Before: `https://myaccount.core.windows.net/`

After: `https://myaccount.core.windows.net`

Troubleshooting

If your data isn't being normalized, and you're experiencing accidental asset duplication, compare the assets fully qualified names to check for capitalization differences or additional characters.

The rules listed above are the only types of duplication Microsoft Purview currently recognizes. If your data is falling outside of these rules, update any ingestion points, for example your ADF pipelines, so that the qualified names match.

If your assets meet the rules but aren't being normalized, contact support.

Next steps

Scan in an Azure Blob Storage account into the Microsoft Purview data map.

Understand business glossary features in the classic Microsoft Purview governance portal

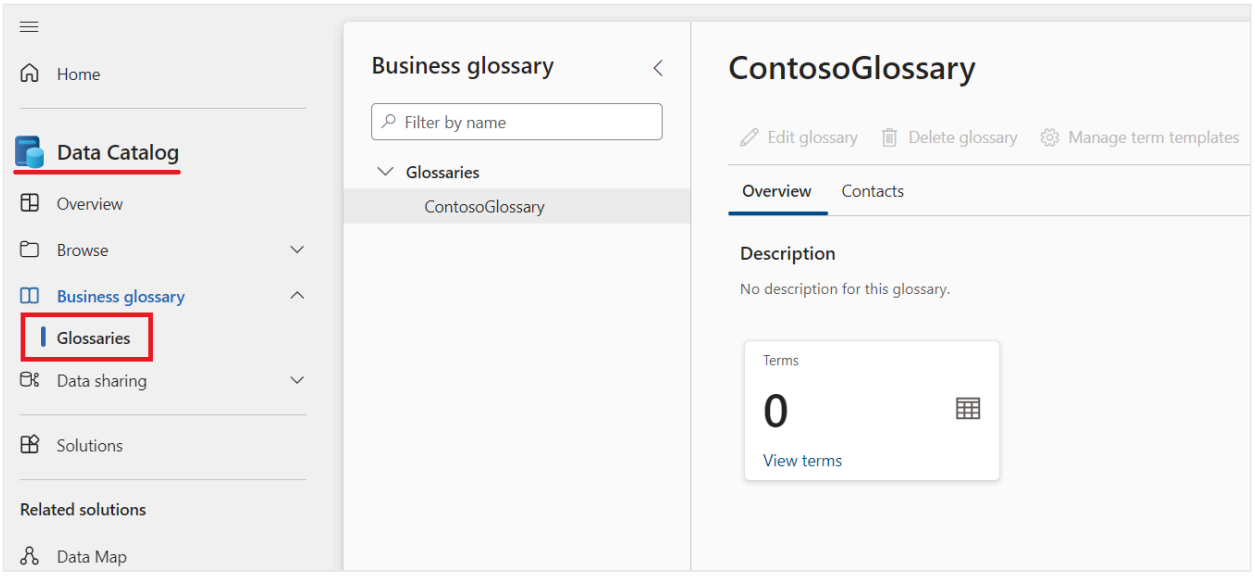
Article • 01/28/2025

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

> **Tip**
>
> Microsoft Purview has a new experience! If you're a new Microsoft Purview customer, or if you want more information, see our **article on the new portal** or the article about our new **data governance experience** and our new **glossary terms**.

A **glossary** provides vocabulary for business users. It consists of business terms that can be related to each other and allows them to be categorized so that they can be understood in different contexts. These terms can be then mapped to assets like a database, tables, columns etc. This helps in abstracting the technical jargon associated with the data repositories and allows the business user to discover and work with data in the vocabulary that is more familiar to them.

You can find Microsoft Purview's classic glossaries in the classic Microsoft Purview governance portal in the Data Catalog under **Business Glossary**.



A **business glossary** is a collection of terms. Each term represents an object in an organization and it's highly likely that there are multiple terms representing the same object. A customer could also be referred to as client, purchaser, or buyer. These multiple terms have a relationship with each other. The relationship between these terms could be:

- synonyms - different terms with the same definition
- related - different name with similar definition

The same term can also imply multiple business objects. It's important that each term is well-defined and clearly understood within the organization.

Relationships between terms

Microsoft Purview supports these out-of-the-box relationships for terms:

- Parent/child term
- Acronym
- Synonyms
- Related terms

Relationship definitions in the glossary are bi-directional: Every relationship between terms is a two-way relationship. This means that if term A is related to term B, then term B is also related to term A.

Anytime you populate a relationship in one direction, Purview automatically adds the reverse relationship for you. For example, if you add term A as a synonym for term B, Purview automatically adds term B as a synonym for term A.

Glossary vs classification vs sensitivity labels

While glossary terms, classifications and labels are annotations to a data asset, each one of them has a different meaning in the context of catalog.

Glossary

As stated above, Business glossary term defines the business vocabulary for an organization and helps in bridging the gap between various departments in your company.

Classifications

Classifications are annotations that can be assigned to entities. The flexibility of classifications enables you to use them for multiple scenarios such as:

- understanding the nature of data stored in the data assets
- defining access control policies

Microsoft Purview has more than 200 system classifiers today and you can define your own classifiers in catalog. As part of the scanning process, we automatically detect these classifications and apply them to data assets and schemas. However, you can override them at any point of time. The human overrides are never replaced by automated scans.

Sensitivity labels

Sensitivity labels are a type of annotation that allows you to classify and protect your organization's data, without hindering productivity and collaboration. Sensitivity labels are used to identify the categories of classification types within your organizational data, and group the policies that you wish to apply to each category. Microsoft Purview makes use of the same sensitive information types as Microsoft 365, which allows you to stretch your existing security policies and protection across your entire content and data estate. The same labels can be shared across Microsoft Office products and data assets in Microsoft Purview.

For more information about sensitivity labels, see: Learn about sensitivity labels

Next steps

- Create and manage glossaries
- Create and manage terms

- Manage term templates
- Browse the data catalog in Microsoft Purview

Feedback

Was this page helpful? Yes No

Create and manage classic business glossaries

Article • 01/28/2025

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

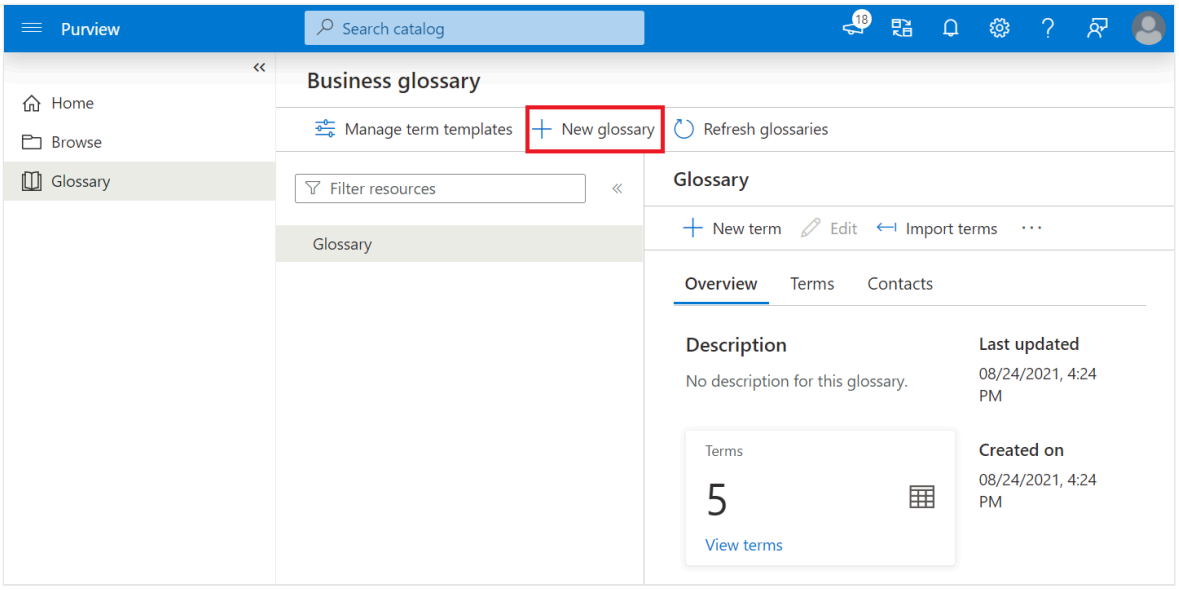
> **Tip**
>
> The new Microsoft Purview Data Catalog has a new glossary experience. For more information, see the glossary terms in the data catalog article.

In the classic Microsoft Purview data governance experience, you can create multiple business glossaries to support separate glossaries for any context in your business. This article provides the steps to create and manage business glossaries in the classic Microsoft Purview Data Catalog.

Create a new glossary

To create a classic business glossary, follow these steps:

1. Open the **Data catalog**
2. In the classic Microsoft Purview portal, select **Glossary**.
3. On the **Business glossary** page, select **+ New glossary**.

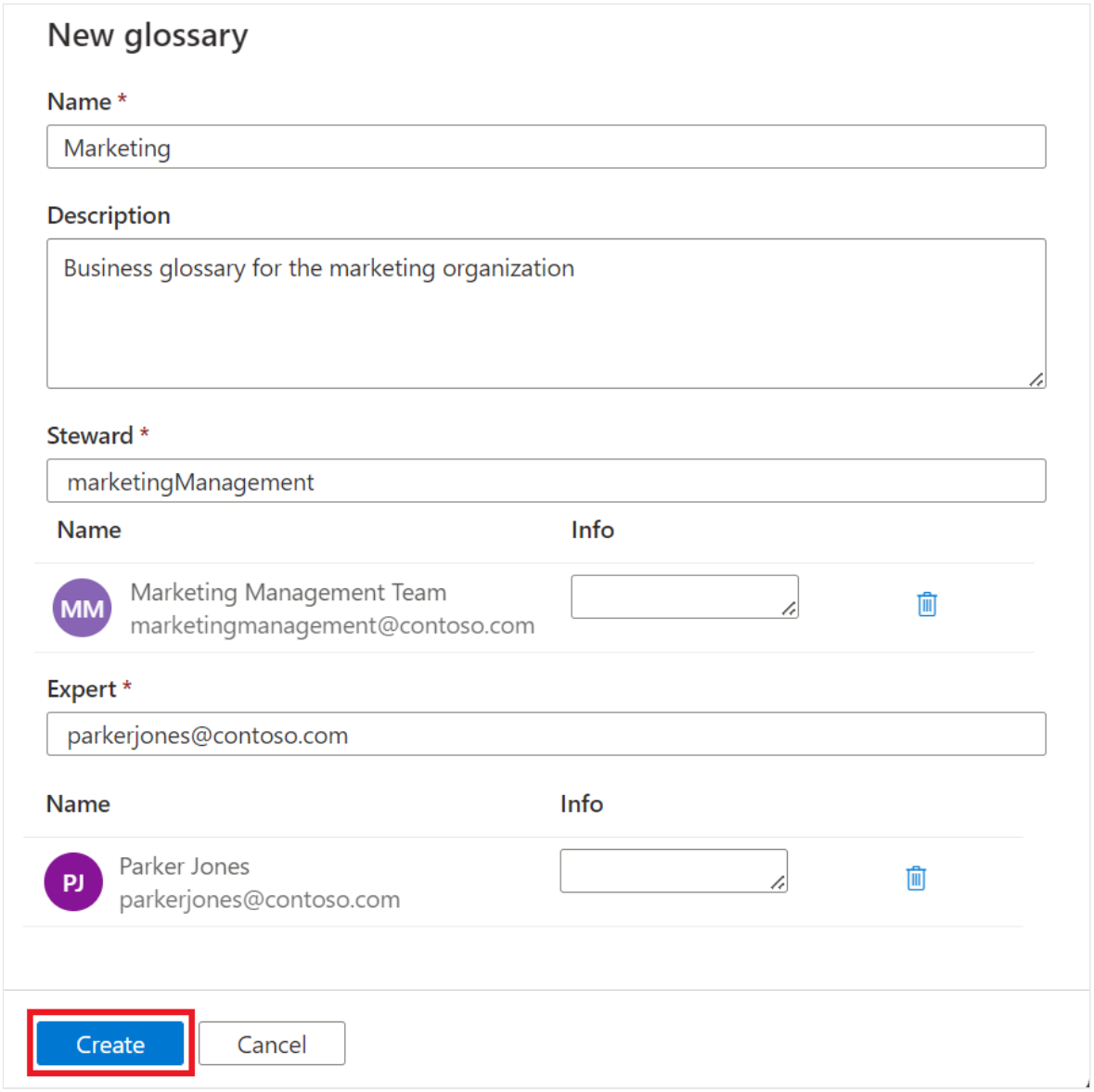


4. Give your glossary a **Name** and a **Description**.

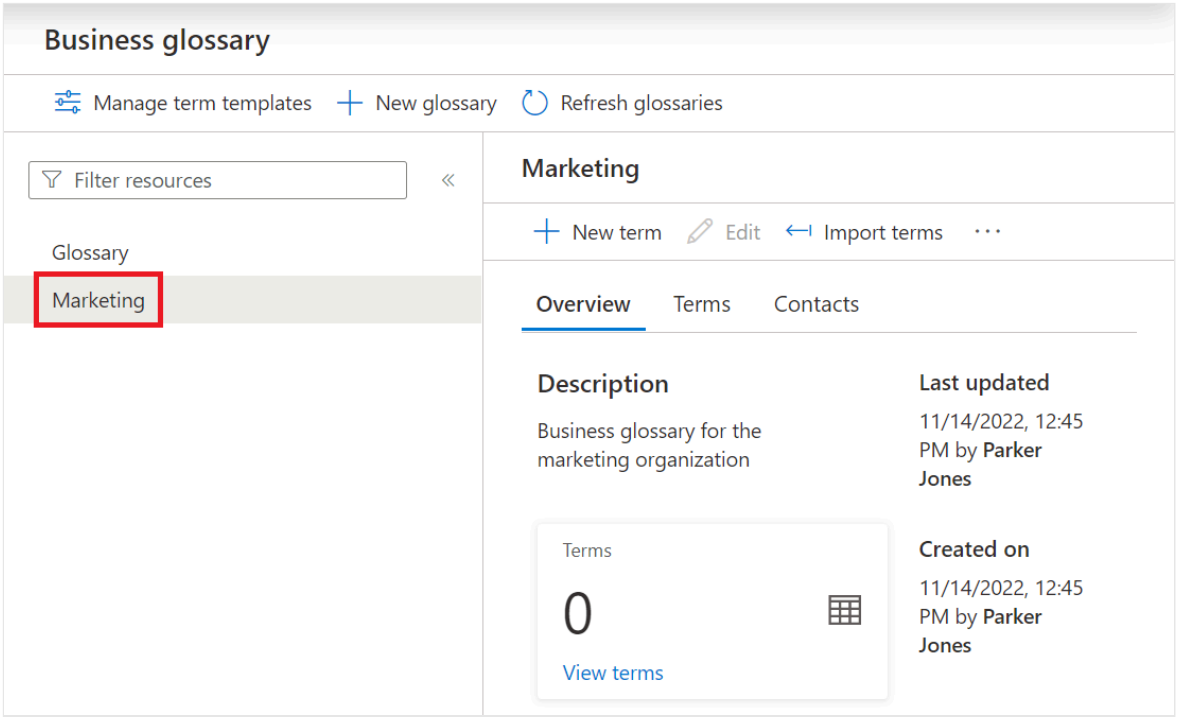
5. You need to select at least one **Steward**, a Microsoft Entra user or group who will manage the glossary.

6. You'll also need to select at least one **Expert**, a Microsoft Entra user or group who can be contacted for more information about the glossary.

7. You can give additional information about the stewards or experts here, and when you're finished, select **Create**.

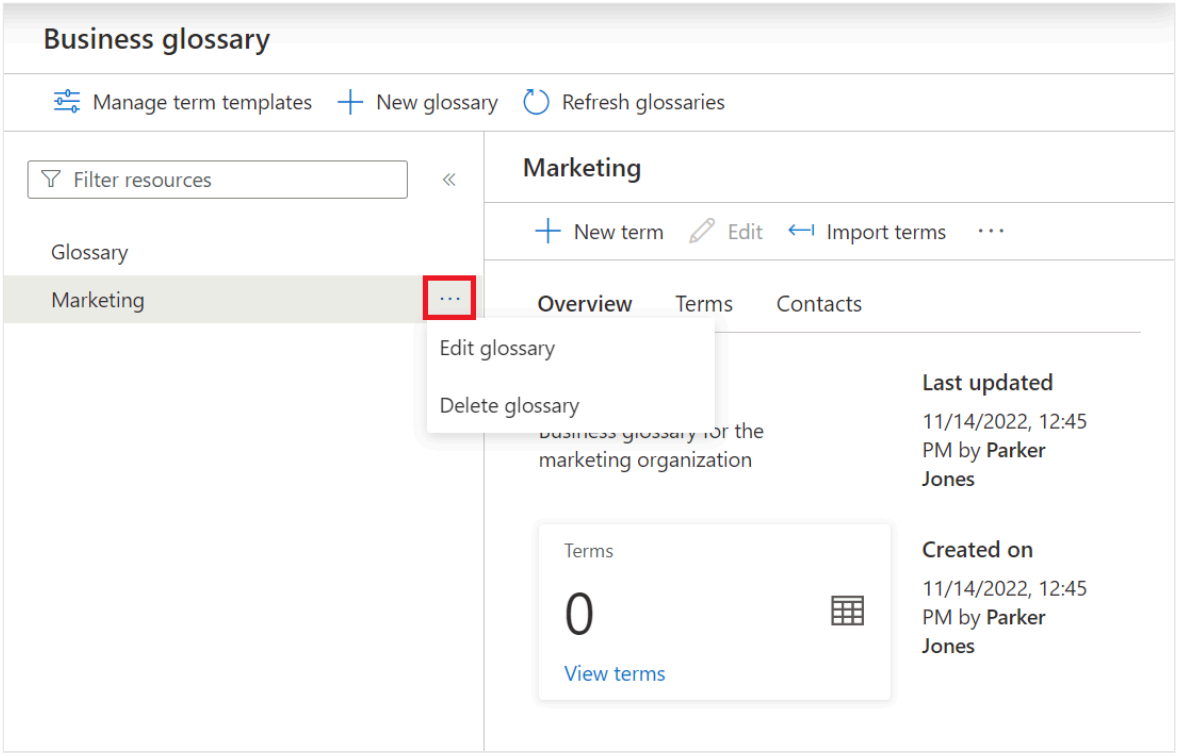


8. Once your new glossary is created, you'll be able to see it in the list of glossaries. You can switch between glossaries by selecting their names.

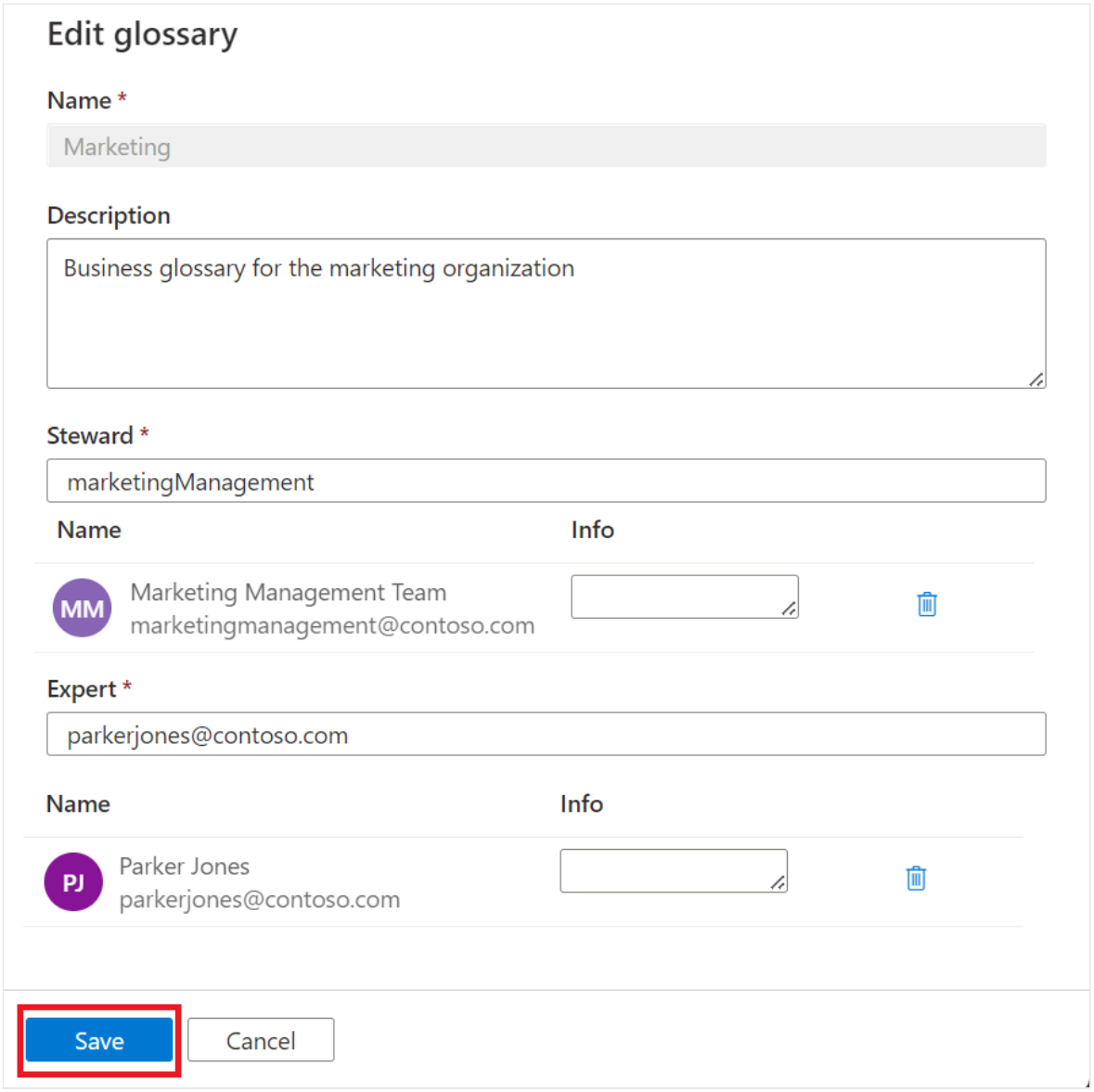


Manage or delete a glossary

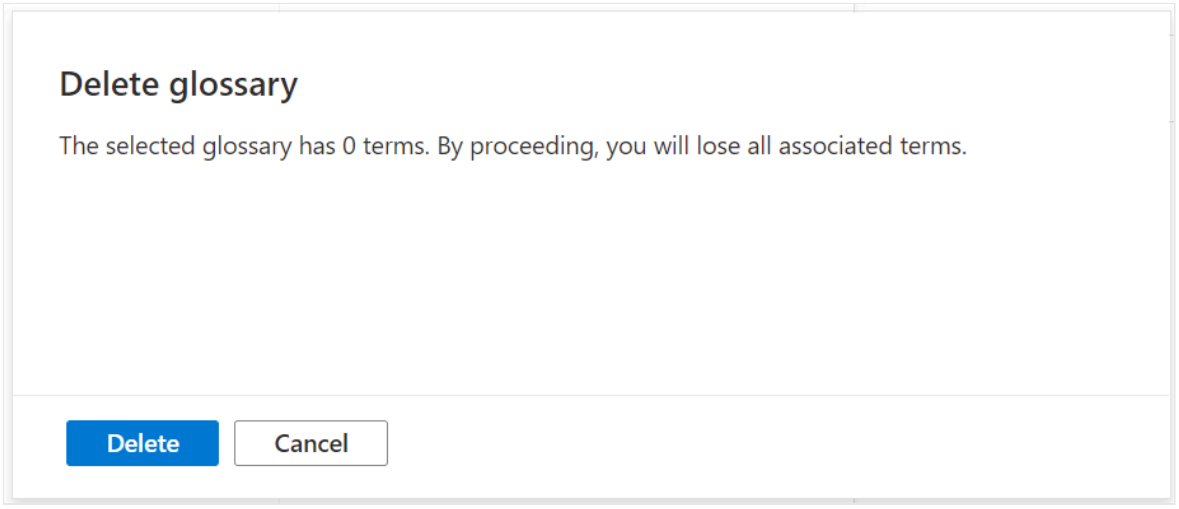
1. To edit or delete a glossary, hover over the glossary and select the ellipsis button next to the glossary's name.



2. If you select **Edit glossary**, you can edit the description and the steward or the expert, but at this time you can't change the glossary name. Select **Save** to save the changes.



3. If you select **Delete glossary**, you'll be asked to confirm the deletion. All terms associated with the glossary will be deleted if you delete the glossary. Select **Delete** again to delete the glossary.



Next steps

- For more information about creating and managing glossary terms, see the glossary terms article.
- For information about glossary terms in the new experience, see glossary terms in the new Microsoft Purview Data Catalog experience

Feedback

Was this page helpful? Yes No

Creating glossary terms in the classic experience

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

This article describes how to work with the business glossary in Microsoft Purview. It provides steps to create a business glossary term in the Microsoft Purview Data Catalog. It also shows you how to import and export glossary terms by using .CSV files, and how to delete terms that you no longer need.

> **Tip**
>
> Take advantage of the unified glossary experience. Migrate your existing glossary terms from the classic data governance experience so that you can apply any glossary term to both a business concept and a data asset. Learn how to migrate classic glossary terms (preview).

Create a term

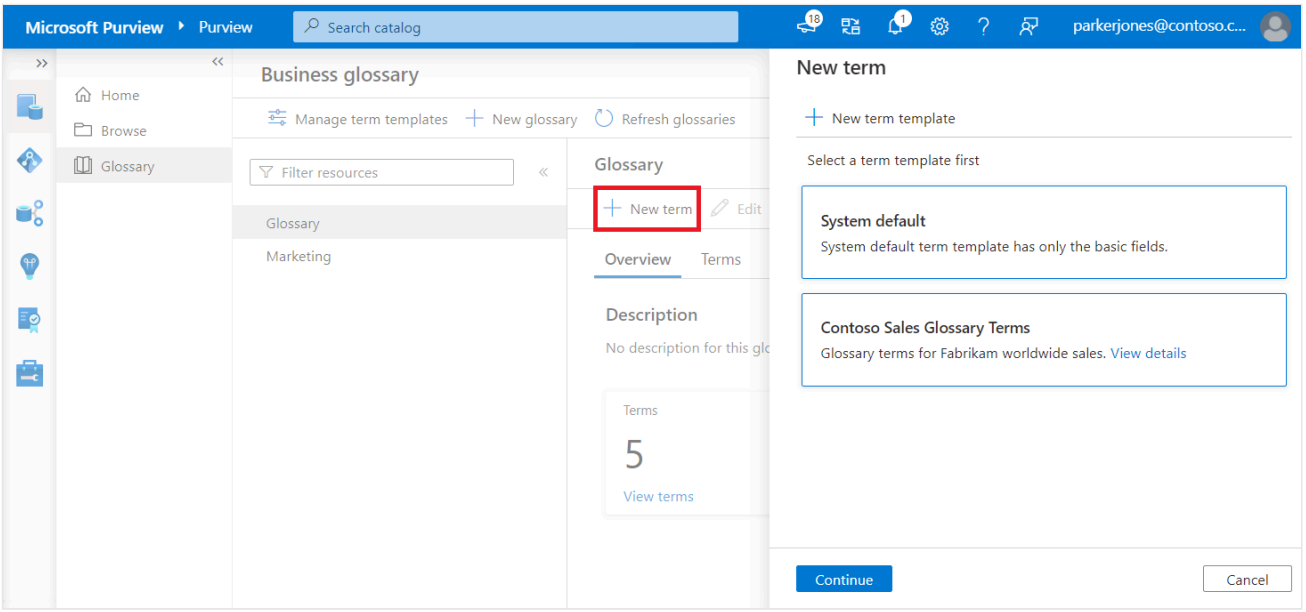
To create a glossary term, follow these steps:

1. Open the Microsoft Purview Data Catalog.
2. If you're using the classic Microsoft Purview portal, select **Glossary**. If you're using the new Microsoft Purview portal under **Business glossary**, select **Glossaries**.
3. Select the glossary you would like to create the new term for, then select **+ New term**. A term can only be added to one glossary at a time.

   > **Note**

Each glossary supports a maximum of 100,000 terms. For more information about creating and managing glossaries see the manage glossaries page.

A pane opens with the **System default** template selected. Choose the template, or templates, that you want to use to create a glossary term, and then select **Continue**.



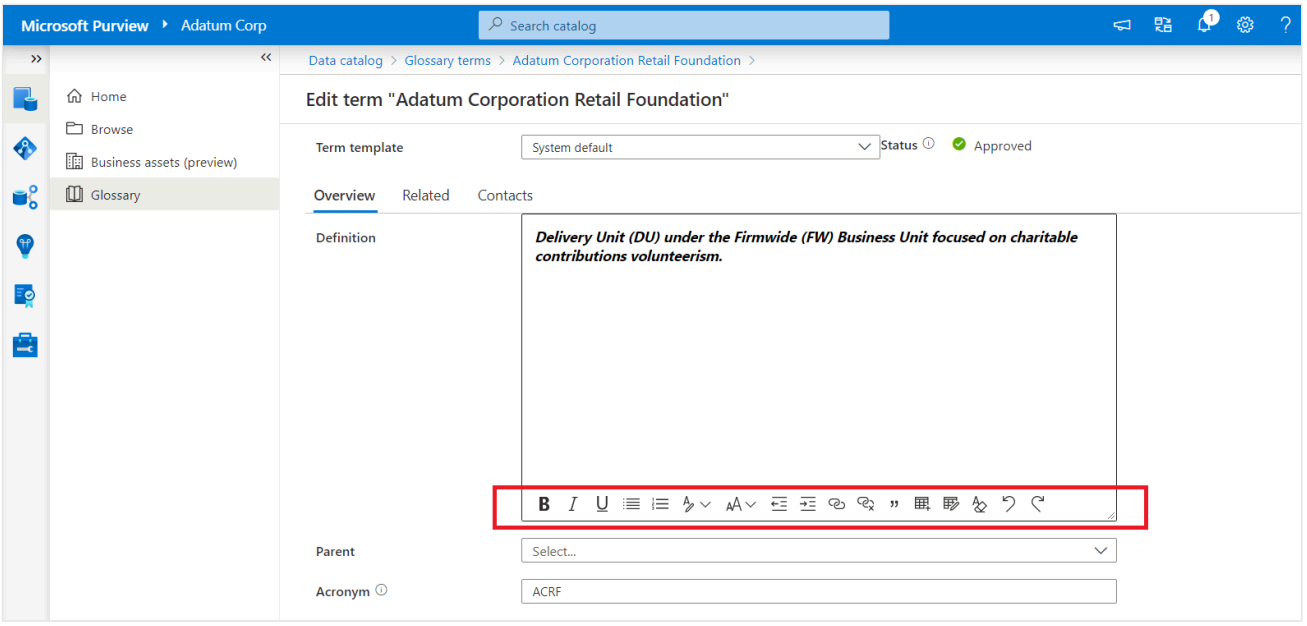
4. If you selected multiple templates, you can select and deselect templates from the **Term template** dropdown at the top of the page.

5. Give your new term a name, which must be unique in the catalog.

> **Note**
>
> Term names are case-sensitive. For example, **Sample** and **sample** could both exist in the same glossary.

6. For **Definition**, add a definition for the term.

   Microsoft Purview enables you to add rich formatting to term definitions. For example, you can add bold, underline, or italic formatting to text. You can also create tables, bulleted lists, or hyperlinks to external resources.



Here are the options for rich text formatting:

Expand table

Name	Description	Keyboard shortcut
Bold	Make your text bold. Adding the asterisk (*) character around text will also make it bold.	Ctrl+B
Italic	Make your text italic. Adding the underscore (_) character around text will also make it italic.	Ctrl+I
Underline	Underline your text.	Ctrl+U
Bullets	Create a bulleted list. Adding the hyphen (-) character before text will also create a bulleted list.	
Numbering	Create a numbered list. Adding the 1 character before text will also create a numbered list.	
Heading	Add a formatted heading.	
Font size	Change the size of your text. The default size is 12.	
Decrease indent	Move your paragraph closer to the margin.	
Increase indent	Move your paragraph farther away from the margin.	
Add hyperlink	Create a link for quick access to webpages and files.	
Remove hyperlink	Change a link to plain text.	

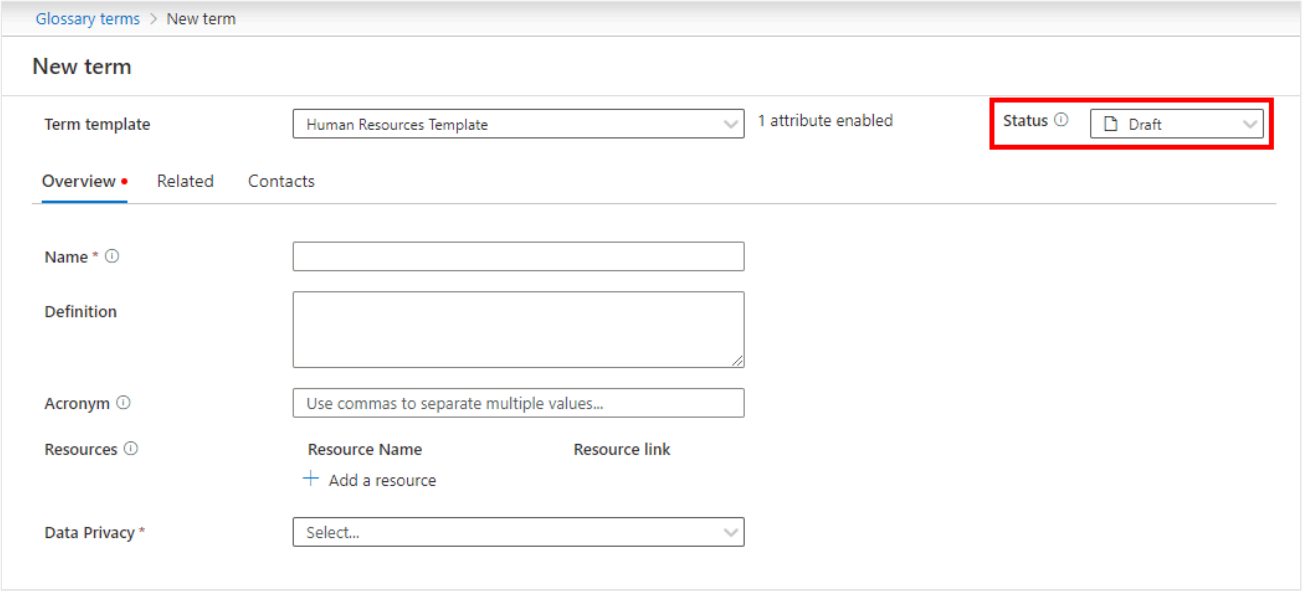
Name	Description	Keyboard shortcut
Quote	Add quote text.	
Add table	Add a table to your content.	
Edit table	Insert or delete a column or row from a table.	
Clear formatting	Remove all formatting from a selection of text.	
Undo	Undo changes that you made to the content.	Ctrl+Z
Redo	Redo changes that you made to the content.	Ctrl+Y

> **Note**
>
> Updating a definition with the rich text editor adds the attribute `"microsoft_isDescriptionRichText": "true"` in the term payload. This attribute is not visible on the user experience and is automatically populated when you take any rich text action. The right text definition is populated in the following snippet of a term's JSON message:

JSON

```
{
     "additionalAttributes": {
     "microsoft_isDescriptionRichText": "true"
     }
 }
```

7. For **Status**, select the status for the term. New terms default to **Draft**.



Status markers are metadata associated with the term. Currently, you can set the following status on each term:

- **Draft**: This term isn't yet officially implemented.
- **Approved**: This term is officially approved.
- **Expired**: This term should no longer be used.
- **Alert**: This term needs attention.

> ⓘ **Important**
>
> If an approval workflow is enabled on the term hierarchy, a new term will go through the approval process when it's created. The term is stored in the catalog only when it's approved. To learn about how to manage approval workflows for a business glossary, see Approval workflow for business terms.

8. Add **Resources** and **Acronym** information. If the term is part of a hierarchy, you can add parent terms at **Parent** on the **Overview** tab.

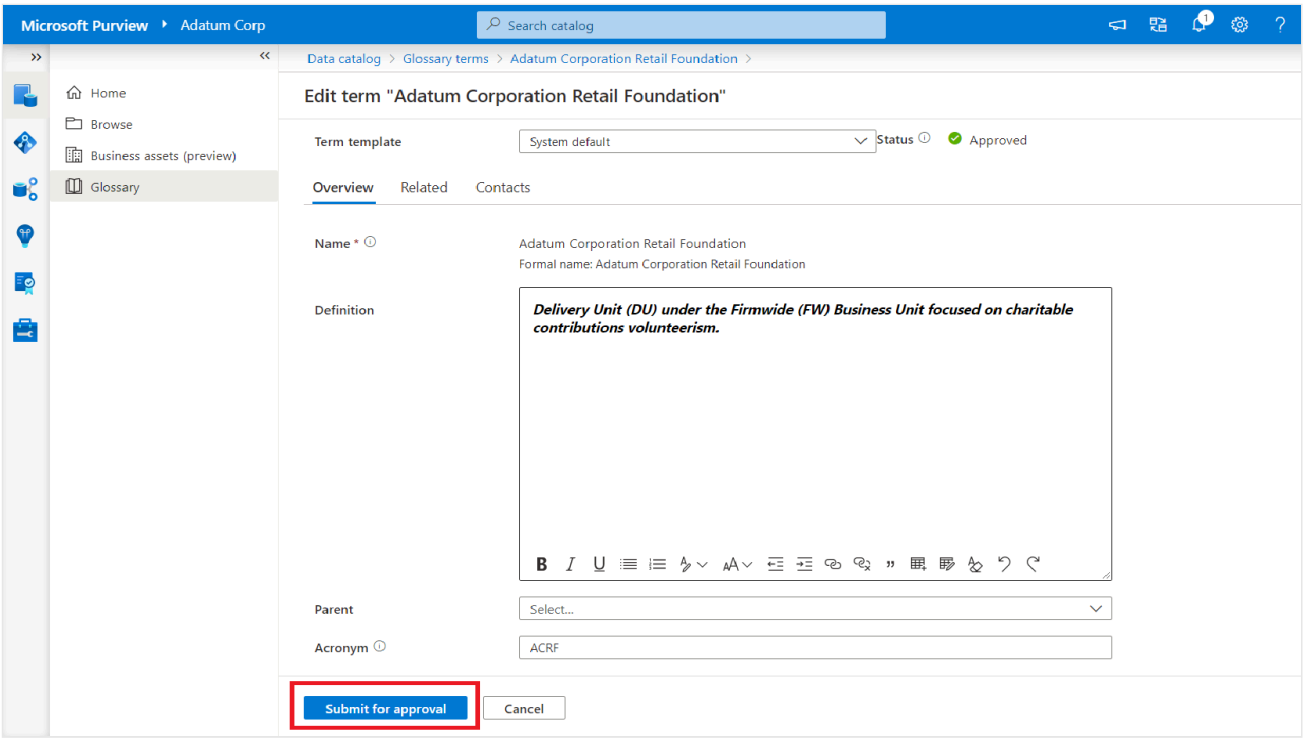
9. To establish relationships with other terms, add **Synonyms** and **Related terms** information on the **Related** tab, and then select **Apply**.

10. Optionally, select the **Contacts** tab to add experts and stewards to your term.

11. Select **Create** to create your term.

> ⓘ **Important**

If an approval workflow is enabled on the term's hierarchy path, you'll see **Submit for approval** instead of the **Create** button. Selecting **Submit for approval** will trigger the approval workflow for this term.



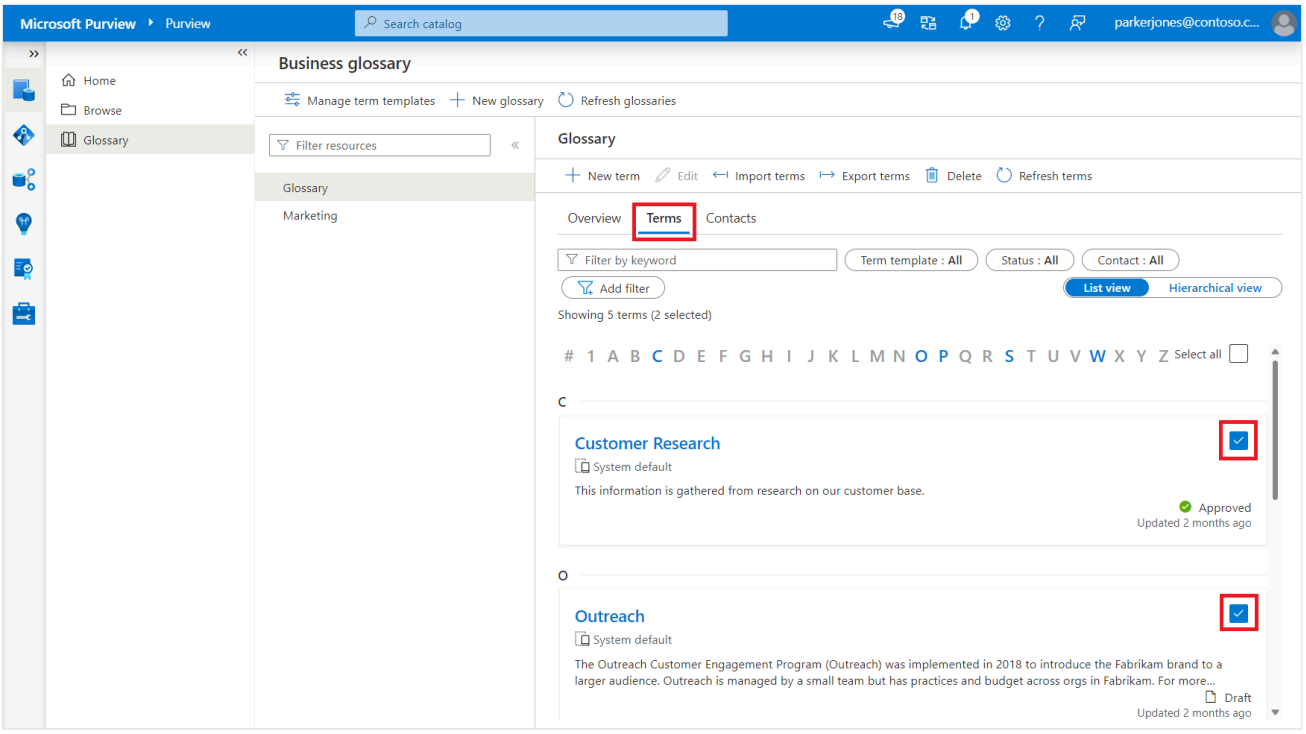
Delete terms

1. Open the **Data catalog**.

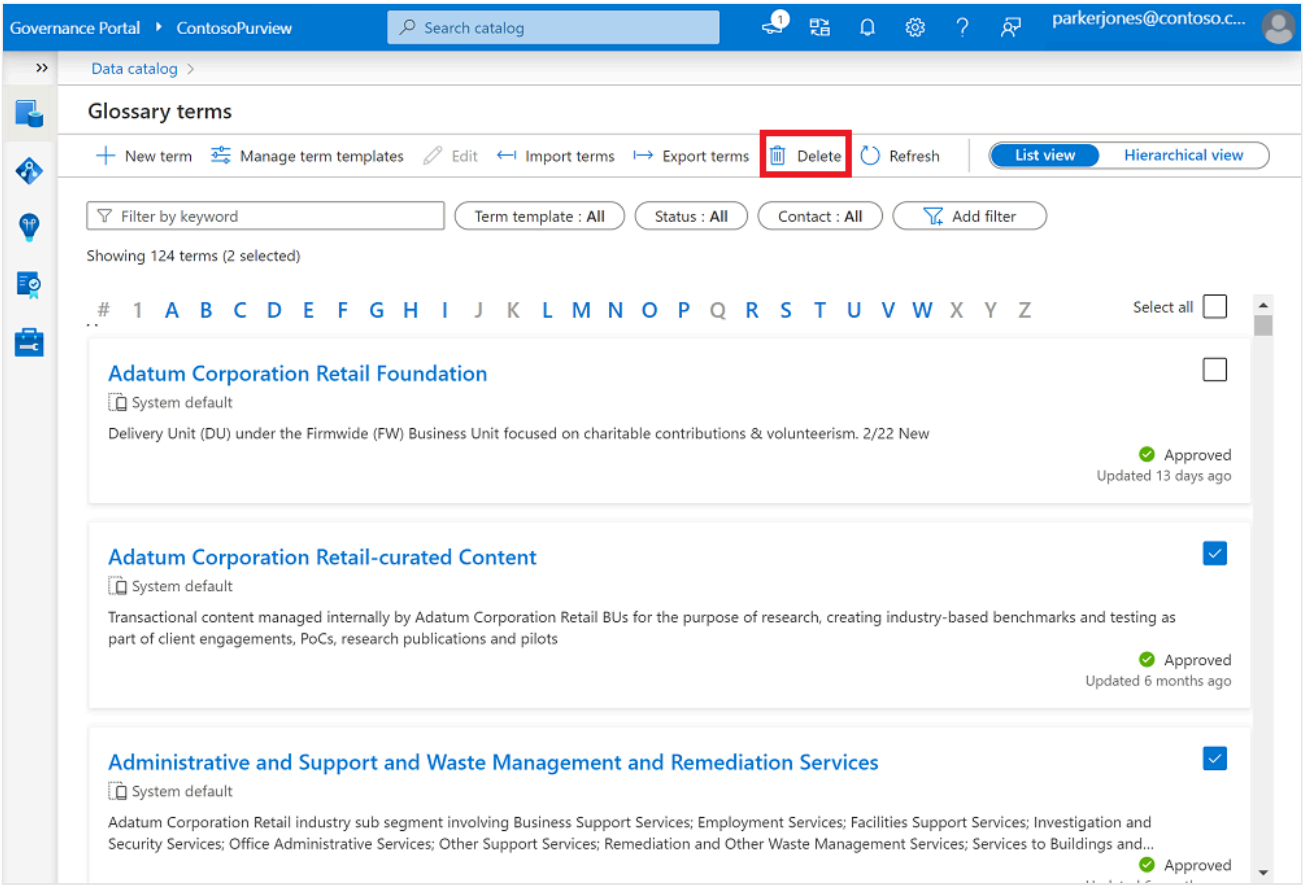
2. If you're using the classic Microsoft Purview portal, select **Glossary**. If you're using the new Microsoft Purview portal under **Business glossary**, select **Glossaries**.

3. Select the glossary that has the terms you want to delete, and select the **Terms** tab.

4. Select checkboxes for the terms that you want to delete. You can select a single term or multiple terms for deletion.



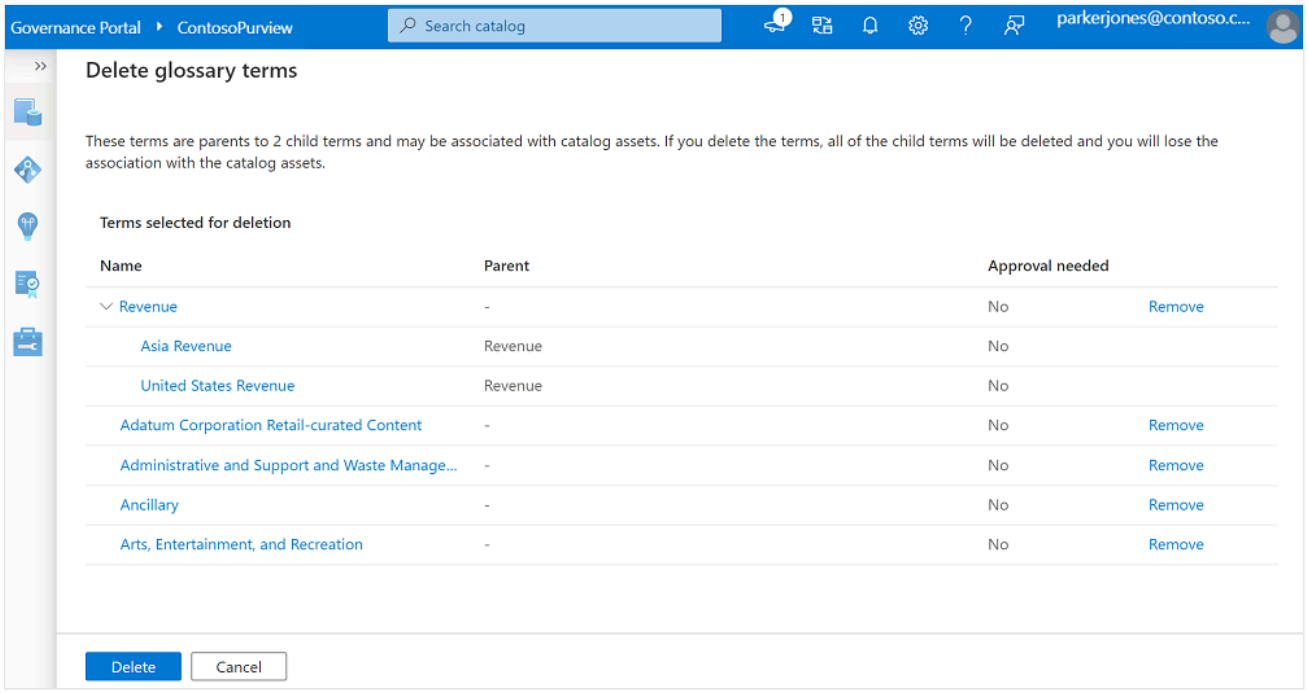
5. Select the **Delete** button on the top menu.



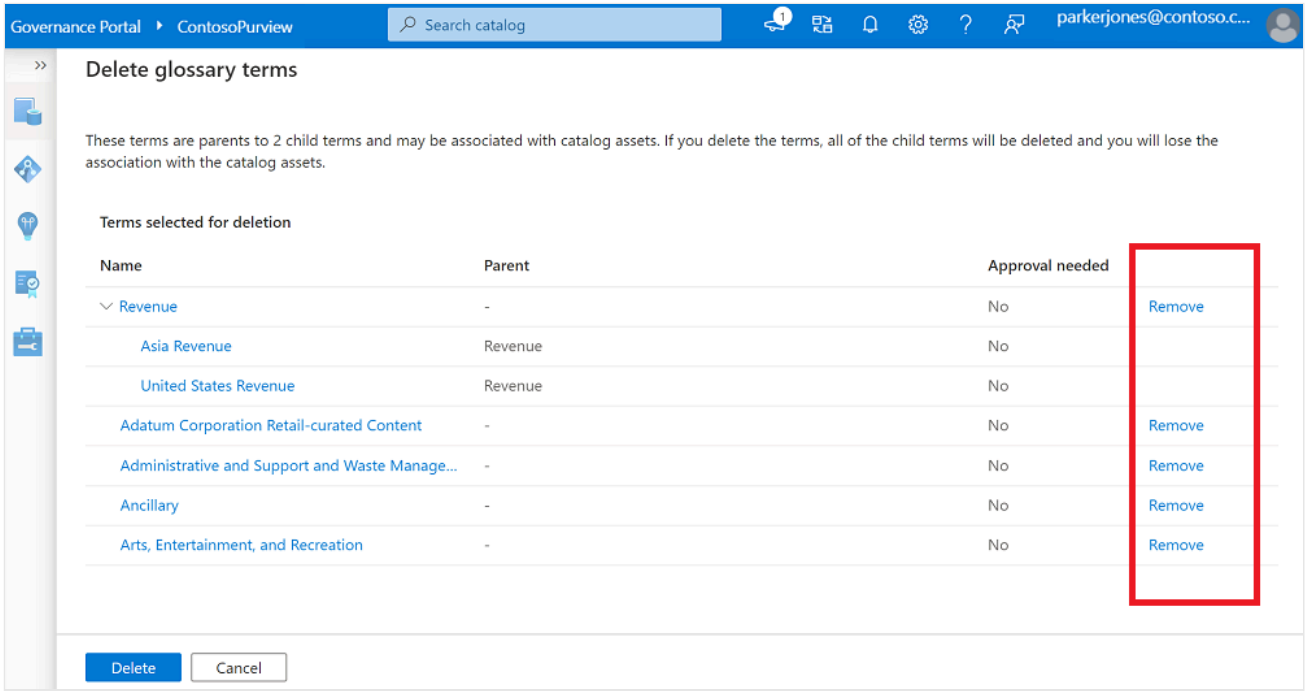
6. A new window shows all the terms selected for deletion. In the following example, the list of terms to be deleted are the parent term **Revenue** and its two child terms.

> **Note**

If a parent is selected for deletion, all the children for that parent are automatically selected for deletion.



Review the list. You can remove the terms that you don't want to delete by selecting **Remove**.

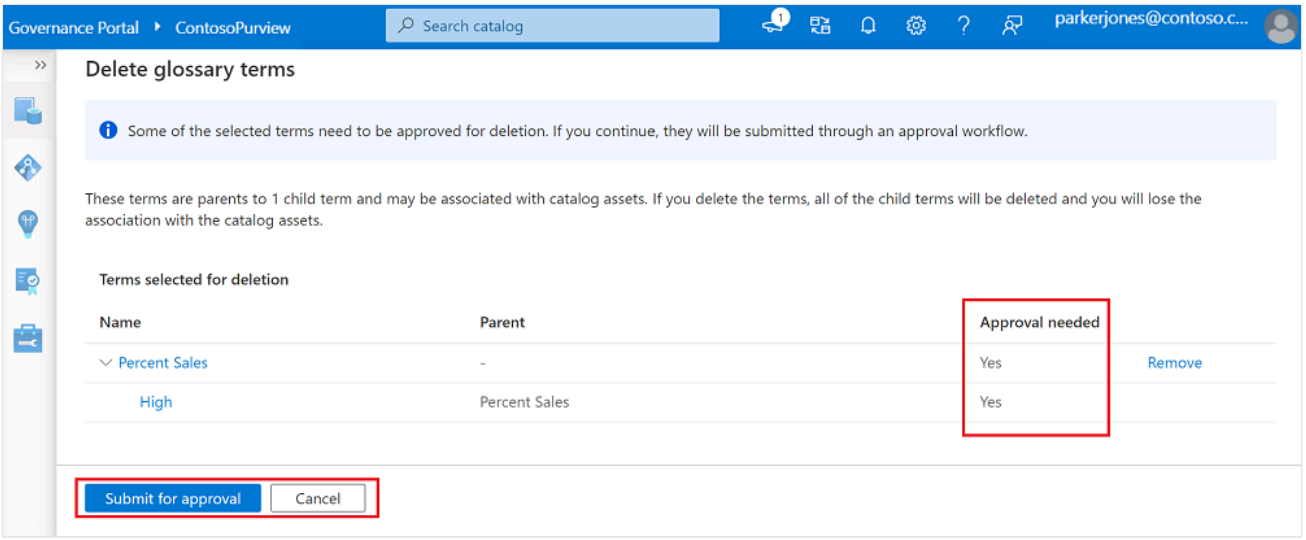


7. The **Approval needed** column shows which terms require an approval process. If the value is **Yes**, the term will go through an approval workflow before deletion. If the value is **No**, the term will be deleted without any approvals.

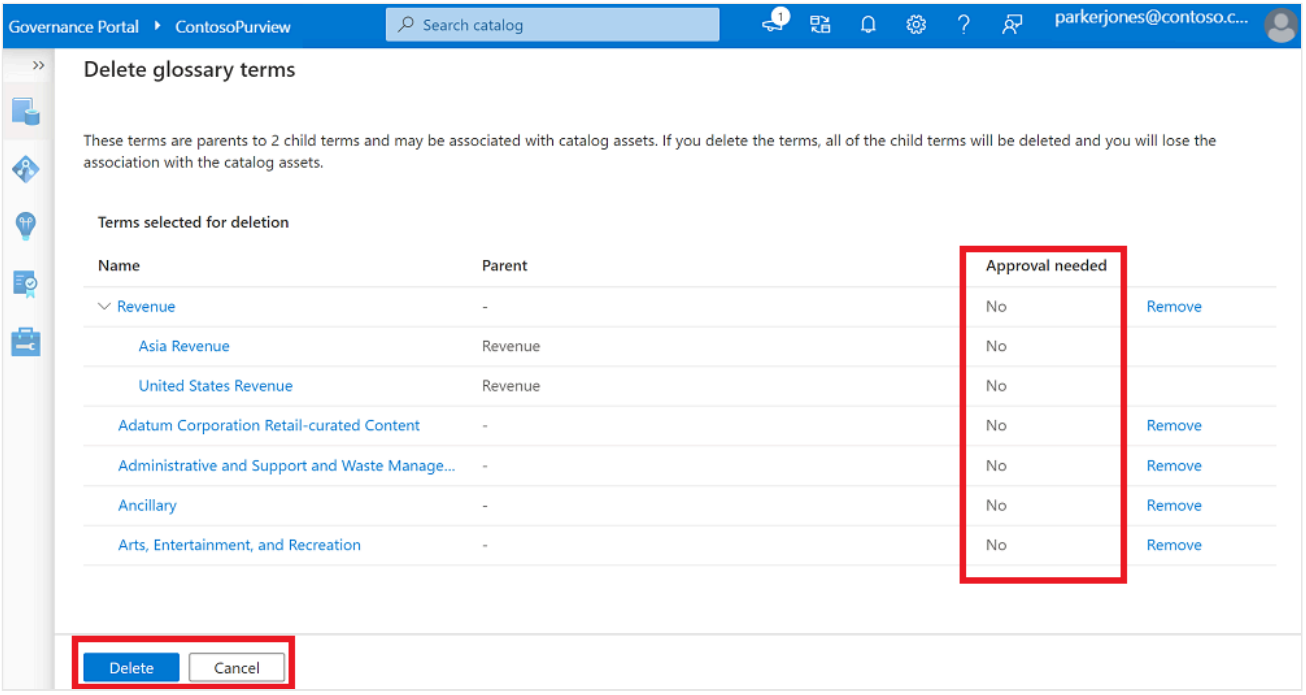
> ⓘ Note

If a parent has an associated approval process but its child doesn't, the workflow for deleting the parent term will be triggered. This is because the selection is done on the parent, and you're acknowledging the deletion of child terms along with parent.

If at least one term needs to be approved, **Submit for approval** and **Cancel** buttons appear. Selecting **Submit for approval** will delete all the terms where approval isn't needed and will trigger approval workflows for terms that require it.



If no terms need to be approved, **Delete** and **Cancel** buttons appear. Selecting **Delete** will delete all the selected terms.



Business terms with approval workflow enabled

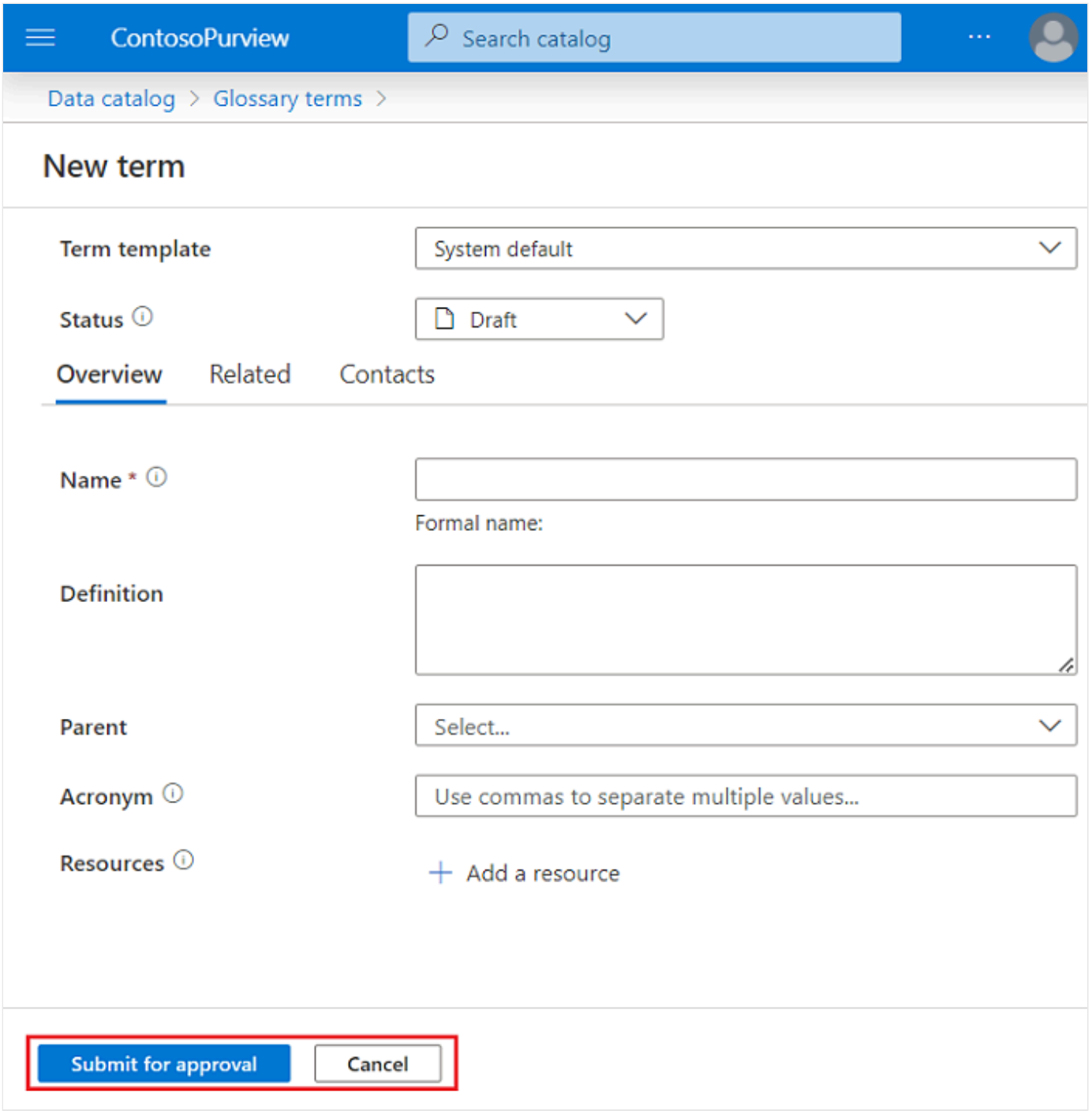
If workflows are enabled on a term, then any create, update, or delete actions for the term will go through an approval before they're saved in the data catalog.

- **New terms**: When a create approval workflow is enabled on a parent term, you see **Submit for approval** instead of **Create** after you enter all the details in the creation process. Selecting **Submit for approval** triggers the workflow. You'll get a notification when your request is approved or rejected.

- **Updates to existing terms**: When an update approval workflow is enabled on a parent term, you see **Submit for approval** instead of **Save** when you're updating the term. Selecting **Submit for approval** triggers the workflow. The changes won't be saved in catalog until all the approvals are met.

- **Deletion**: When a delete approval workflow is enabled on the parent term, you see **Submit for approval** instead of **Delete** when you're deleting the term. Selecting **Submit for approval** triggers the workflow. However, the term won't be deleted from the catalog until all the approvals are met.

- **Importing terms**: When an import approval workflow is enabled for the Microsoft Purview glossary, you see **Submit for approval** instead of **OK** in the **Import** window when you're importing terms via .CSV file. Selecting **Submit for approval** triggers the workflow. However, the terms in the file won't be updated in the catalog until all the approvals are met.



Next steps

- For more information about glossary terms, see the glossary reference.
- For more information about approval workflows of the business glossary, see Approval workflow for business terms.

Last updated on 04/24/2026

Import and export glossary terms

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

This article describes how to work with the business glossary in Microsoft Purview. It provides steps to create a business glossary term in the Microsoft Purview Data Catalog. It also shows you how to import and export glossary terms by using .CSV files.

> **Note**
>
> To edit the glossary, and import or export terms, you need to have data curator permissions.

Import terms into the glossary

The Microsoft Purview Data Catalog provides a template .CSV file for you to import terms from the catalog into your glossary. Duplicate terms include both spelling and capitalization, because term names are case-sensitive.

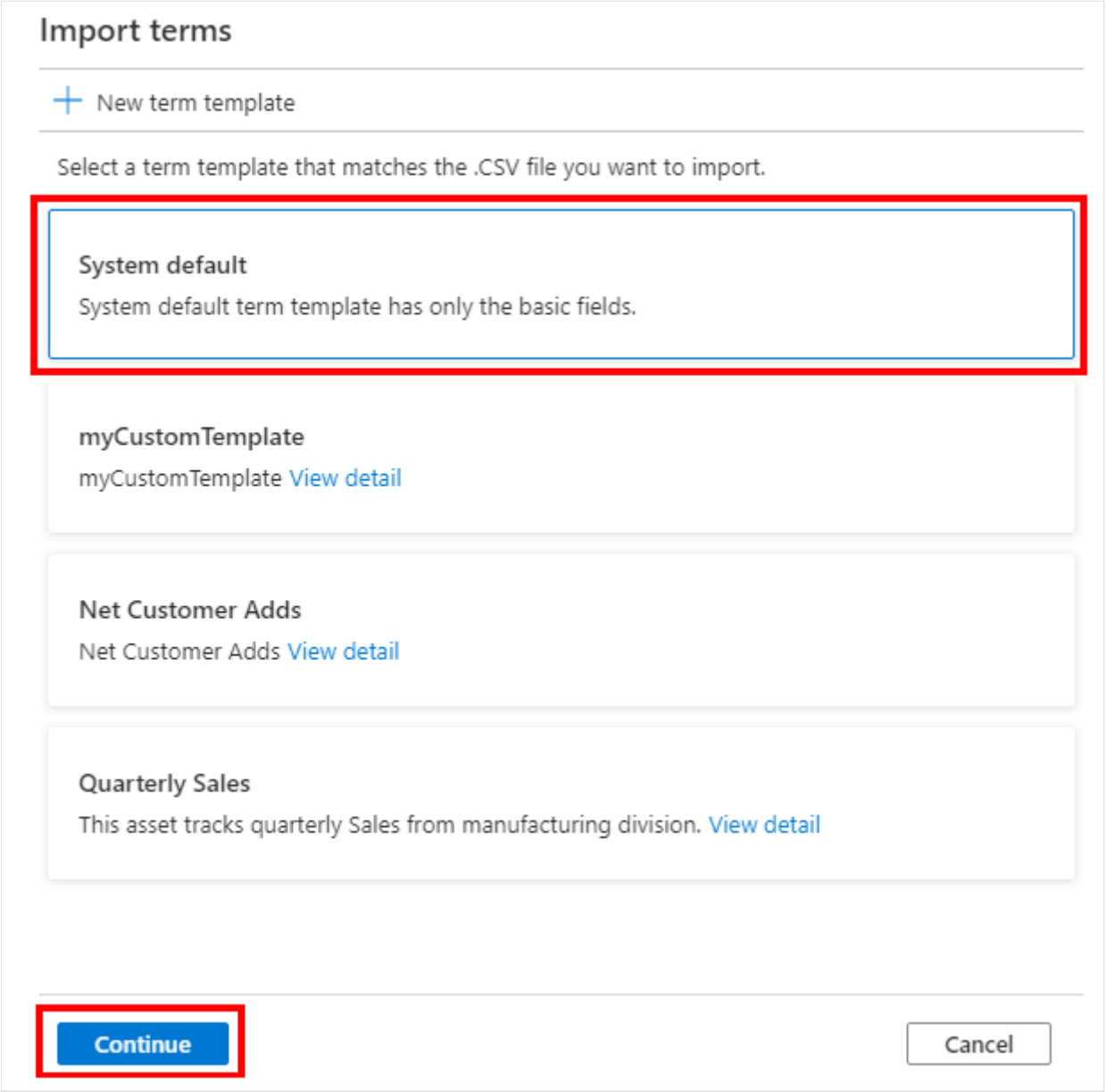
1. Open the Microsoft Purview Data Catalog.

2. If you're using the classic Microsoft Purview portal, select **Glossary**. If you're using the new Microsoft Purview portal under **Business glossary**, select **Glossaries**.

3. Select the glossary where you want to import terms.
   a. In the Microsoft Purview portal, select **View terms** and then select **Import terms**.
   b. In the classic governance portal, select **Import terms**.

4. Select the template, or templates, for the terms you want to import, and then select **Continue**.

   You can select multiple templates and import terms for different templates from a single .csv file.

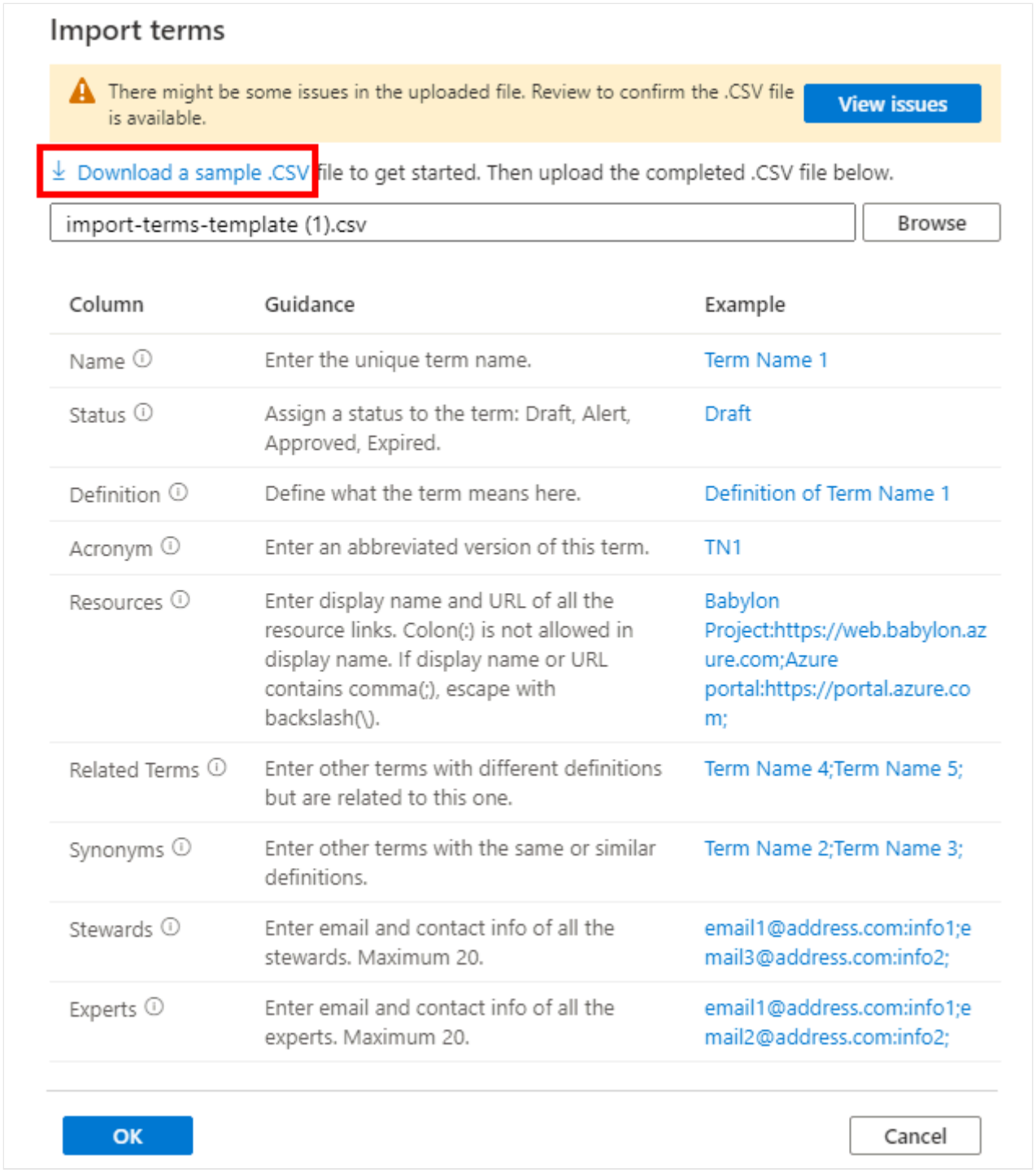


5. Download the .csv template and use it to enter the terms that you want to add.

   Give your template file a name that starts with a letter and includes only letters, numbers, spaces, an underscore (_), or other non-ASCII Unicode characters. Special characters in the file name will create an error.

> ⓘ **Important**
>
> The system supports only importing columns that are available in the template. The **System default** template will have all the default attributes.
>
> Custom term templates define out-of-the box attributes and additional custom attributes. Therefore, the .CSV file differs in the total number of columns and the column names, depending on the term template that you select. You can also review the file for problems after upload.

If you want to upload a file with a rich text definition, be sure to enter the definition with markup tags and populate the column `IsDefinitionRichText` to `true` in the .CSV file.



6. After you finish filling out your .CSV file, select your file to import, and then select **OK**.

The system will upload the file and add all the terms to your selected glossary.

> **Important**

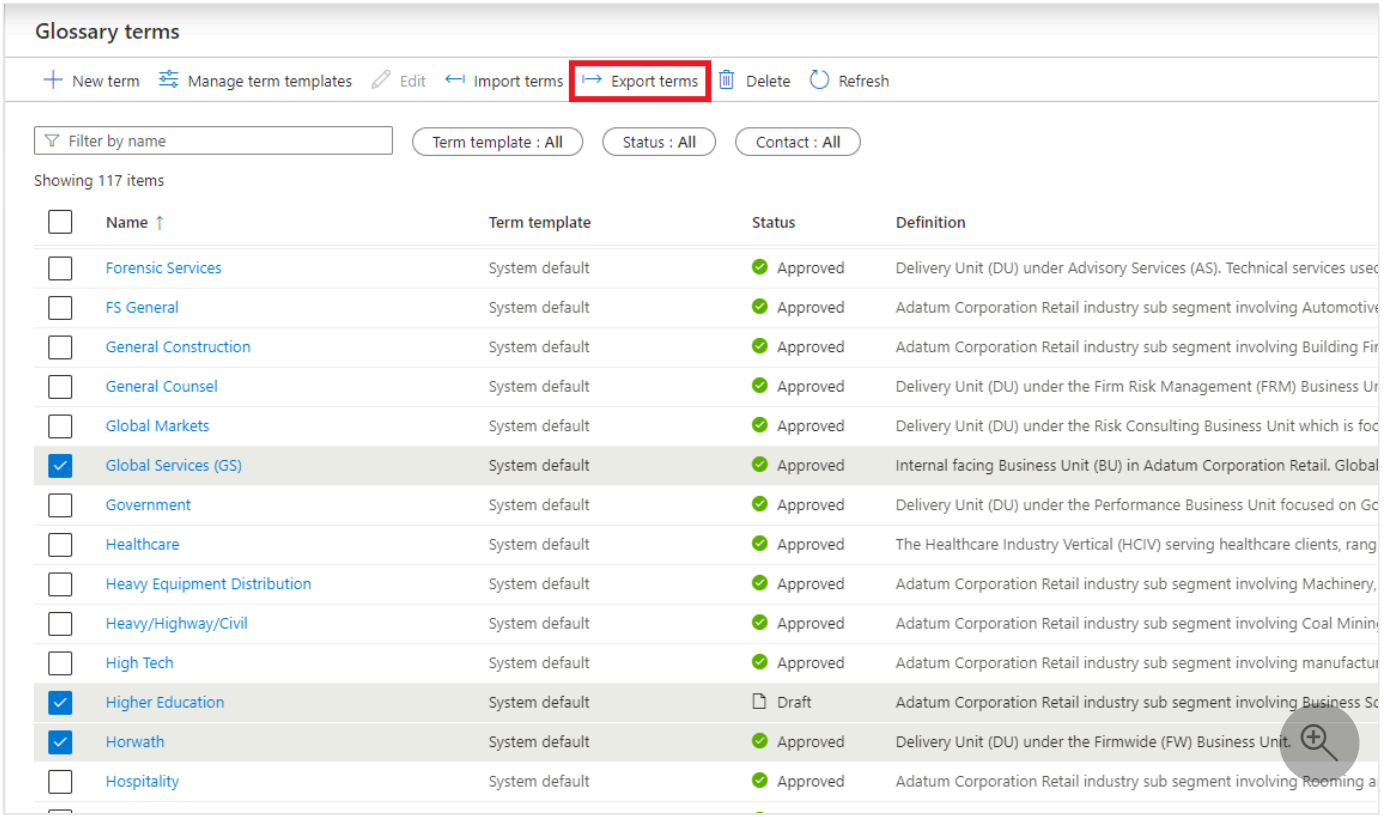
The email address for an expert or steward should be the primary address of the user from the Microsoft Entra group. Alternate emails, user principal names, and non-Microsoft Entra ID emails are not yet supported.

Export terms from the glossary

When you're in any of your glossaries, the **Export terms** button is disabled by default. After you select the terms that you want to export, the **Export terms** button is enabled.

> **Note**
>
> Selected terms **don't** need to be from the same term template to be able to export them.

Select **Export terms** to download the selected terms.



> **Important**
>
> The import process currently doesn't support updating the parent of a term.

Business terms with approval workflow enabled

If workflows are enabled on a term, then any import actions for the term will go through an approval before they're saved in the data catalog.

When an import approval workflow is enabled for the Microsoft Purview glossary, you'll see **Submit for approval** instead of **OK** in the **Import** window when you're importing terms via .CSV file. Selecting **Submit for approval** triggers the workflow. However, the terms in the file won't be updated in the catalog until all the approvals are met.

Next steps

- For more information about glossary terms, see the glossary reference.
- For more information about approval workflows of the business glossary, see Approval workflow for business terms.

Last updated on 04/22/2026

How to manage term templates for business glossary

Article • 01/28/2025

> **Note**
>
> Microsoft Purview Data Map (classic), Data Catalog (classic), Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

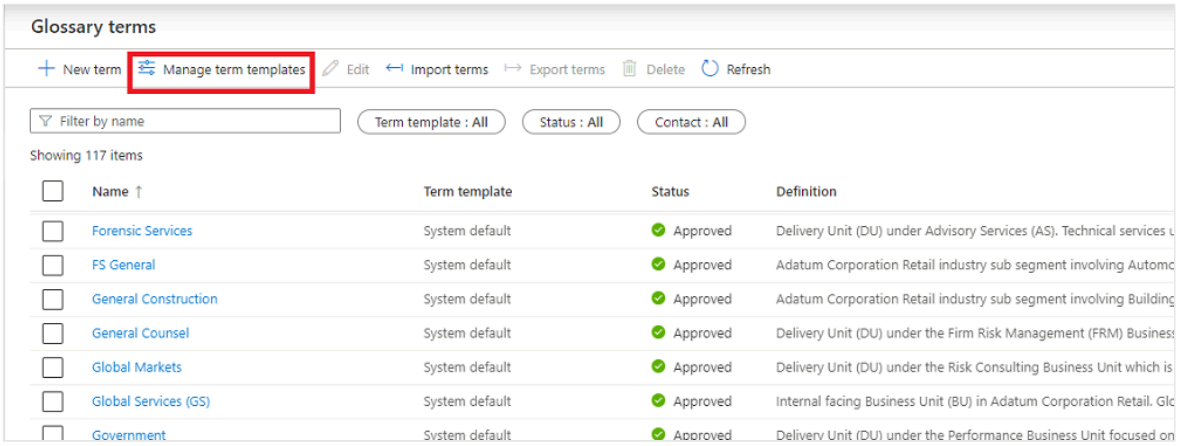
Microsoft Purview allows you to create a glossary of terms that are important for enriching your data. Each new term added to your Microsoft Purview Data Catalog Glossary is based on a term template that determines the fields for the term. This article describes how to create a term template and custom attributes that can be associated to glossary terms.

Manage term templates and custom attributes

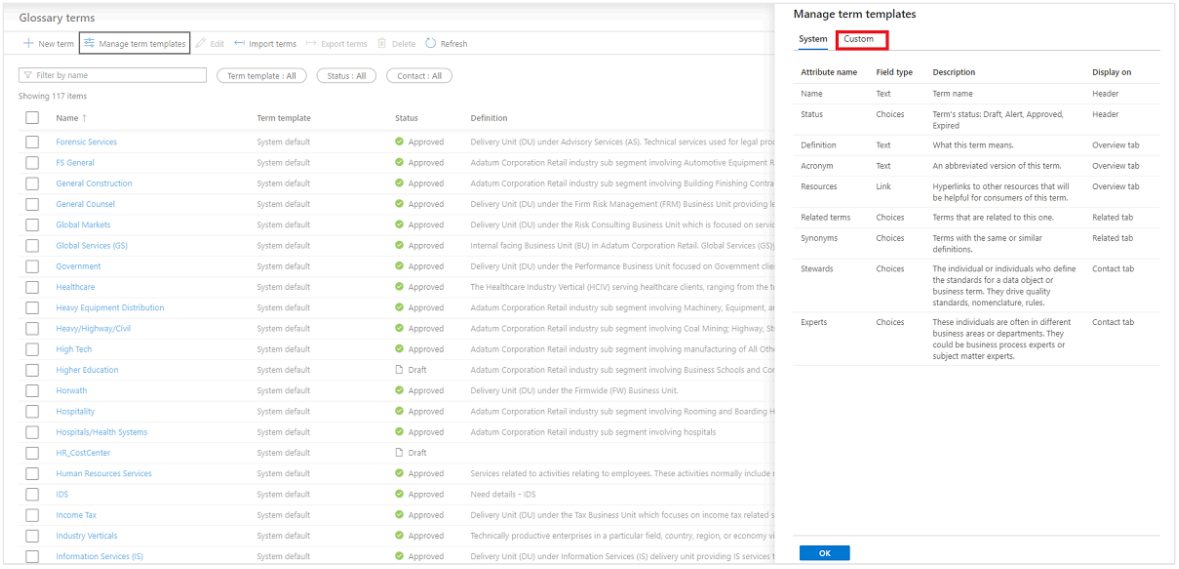
Using term templates, you can create custom attributes, group them together and apply a template while creating terms.

> **Note**
>
> Once a term is created, the template associated with the term can't be changed.

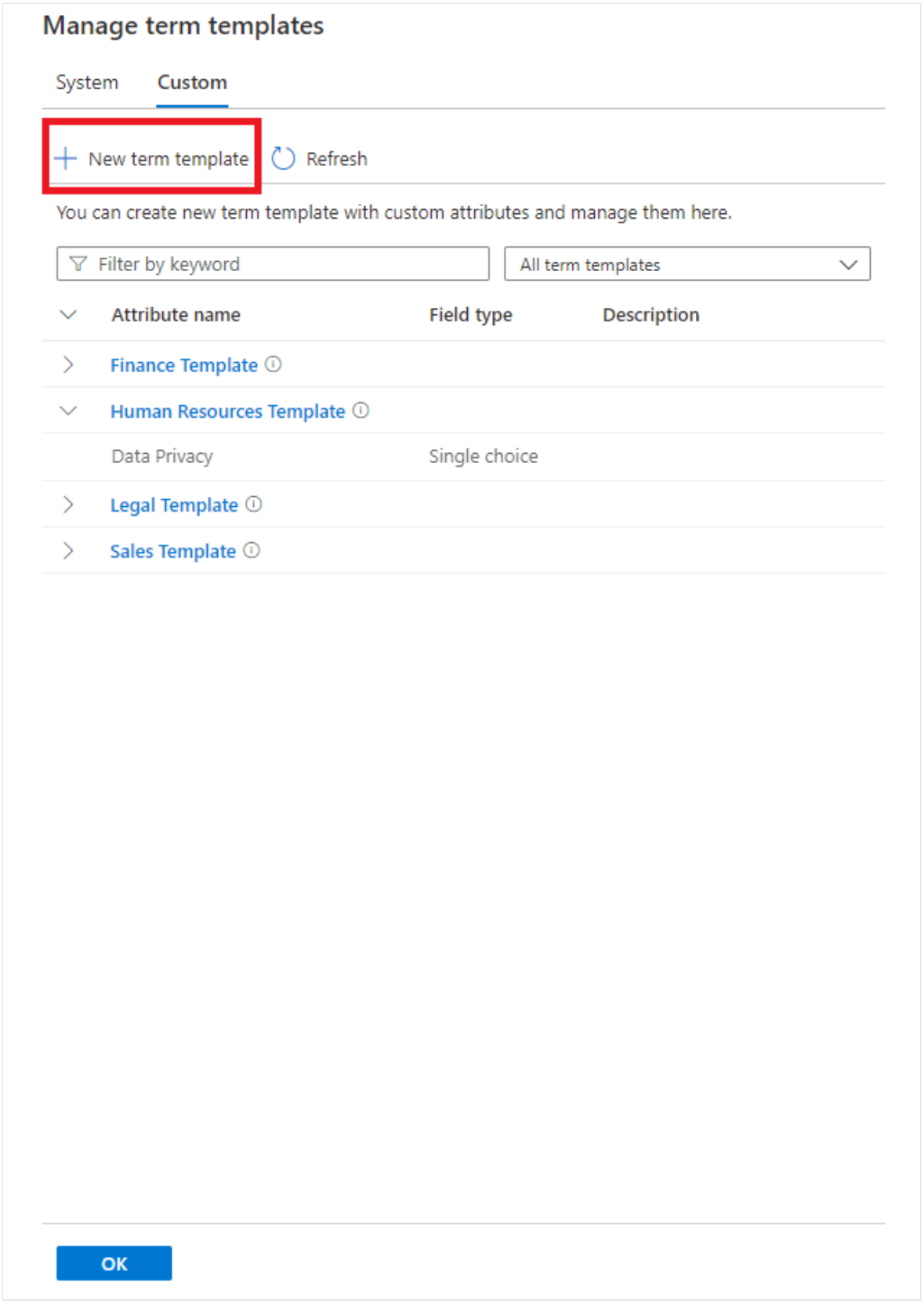
1. Open the Microsoft Purview Data Catalog.
2. If you're using the classic Microsoft Purview portal, select **Glossary**. If you're using the new Microsoft Purview portal under **Business glossary**, select **Glossaries**.
3. Select **Manage term templates**.



4. The page shows both system and custom attributes. Select the **Custom** tab to create or edit term templates.

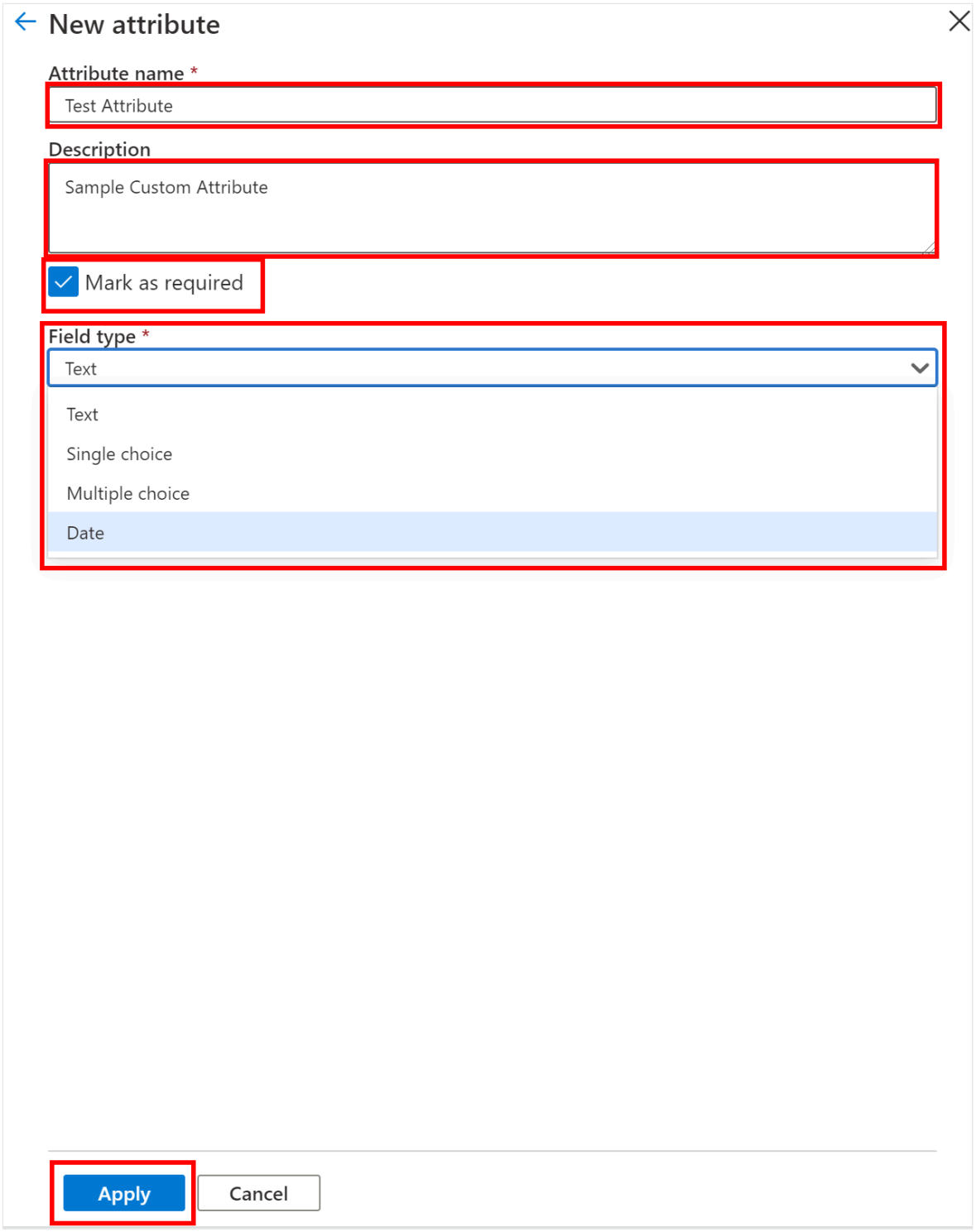


5. Select + **New term template** and enter a template name and description.

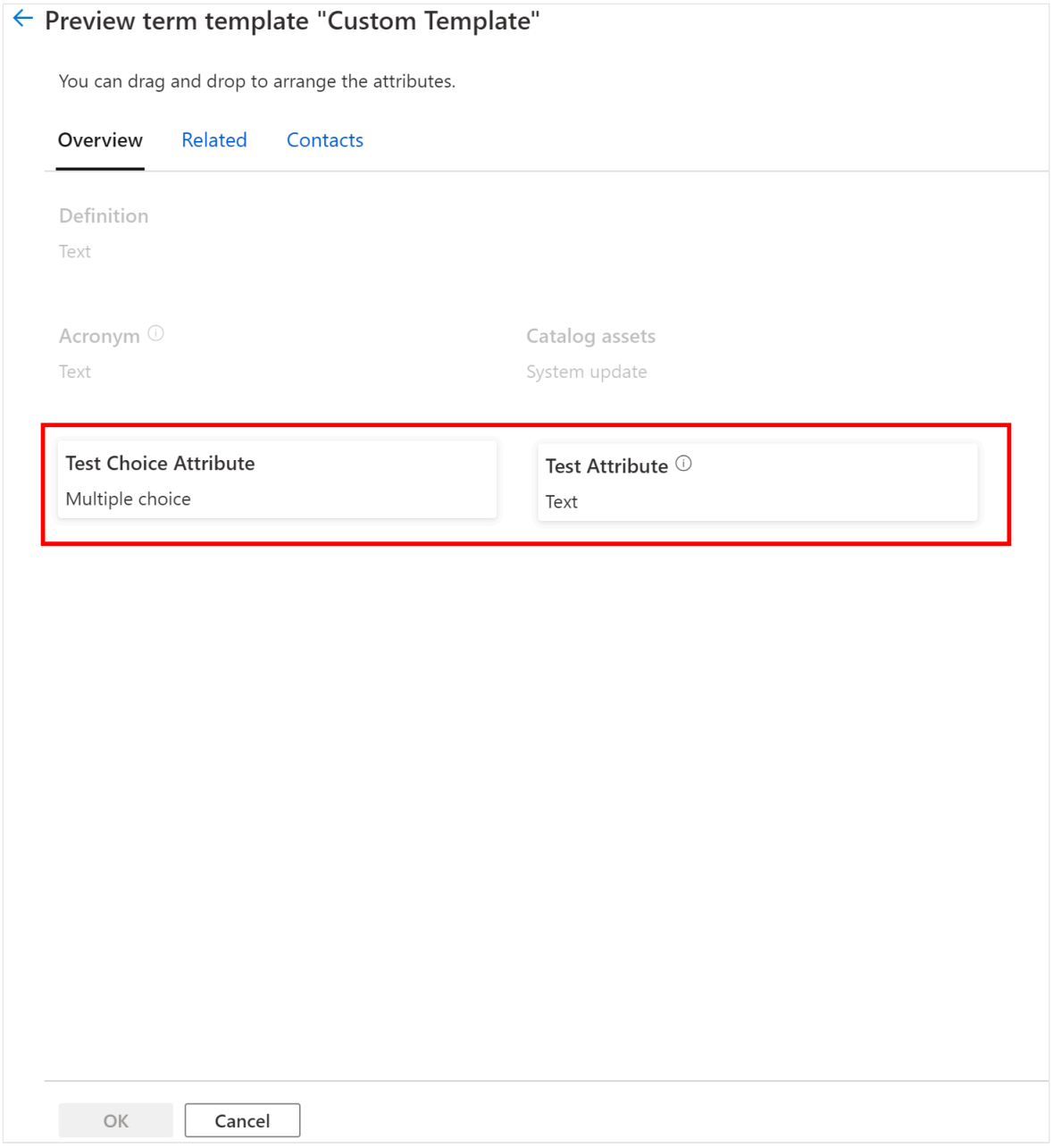


6. Select + **New attribute** to create a new custom attribute for the term template. Enter an attribute name, description. The custom attribute name must be unique within a term template but can be same name can be reused across templates.

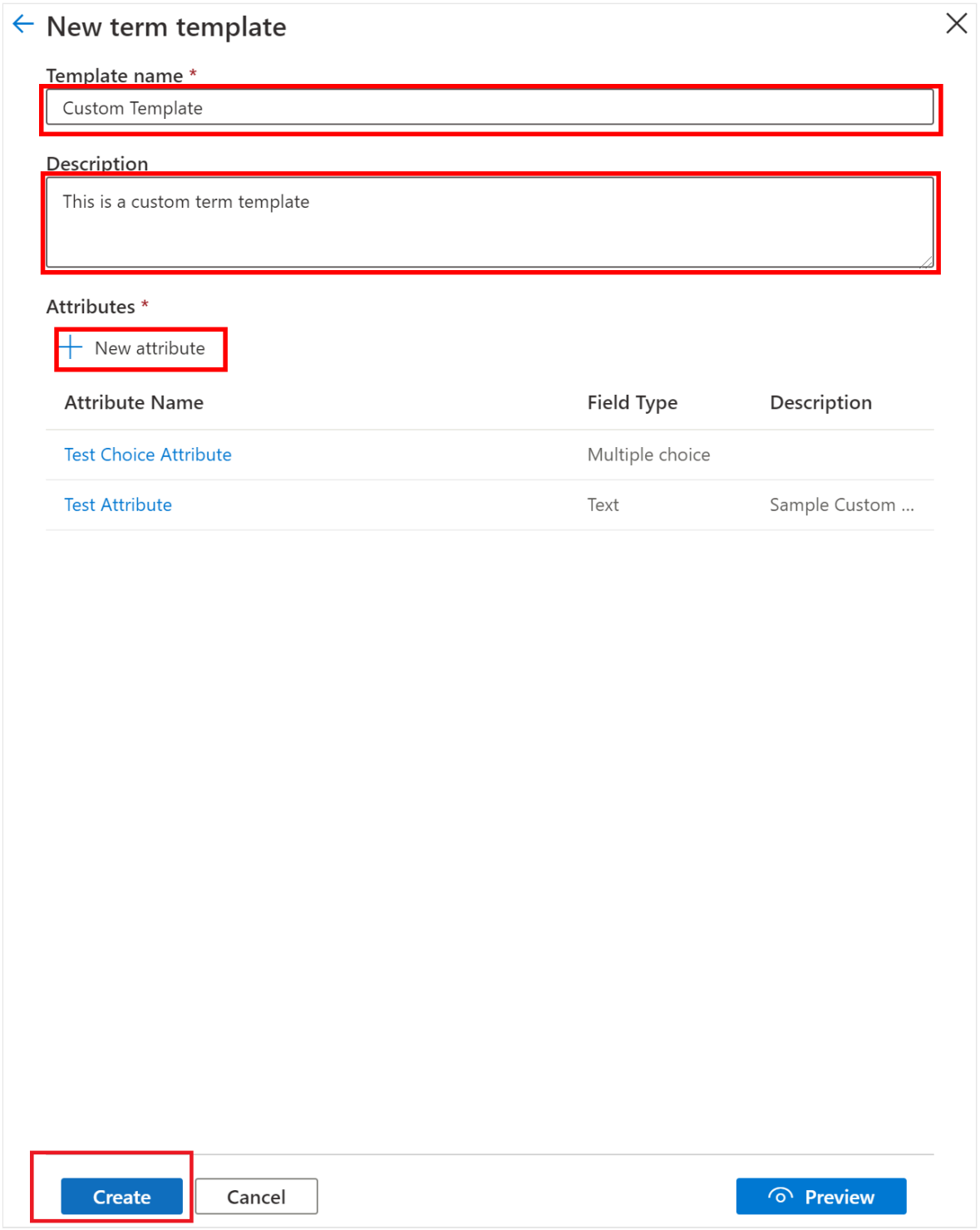
   Choose the field type from the list of options **Text**, **Single choice**, **Multi choice**, or **Date**. You can also provide a default value string for Text field types. The attribute can also be marked as **required**.



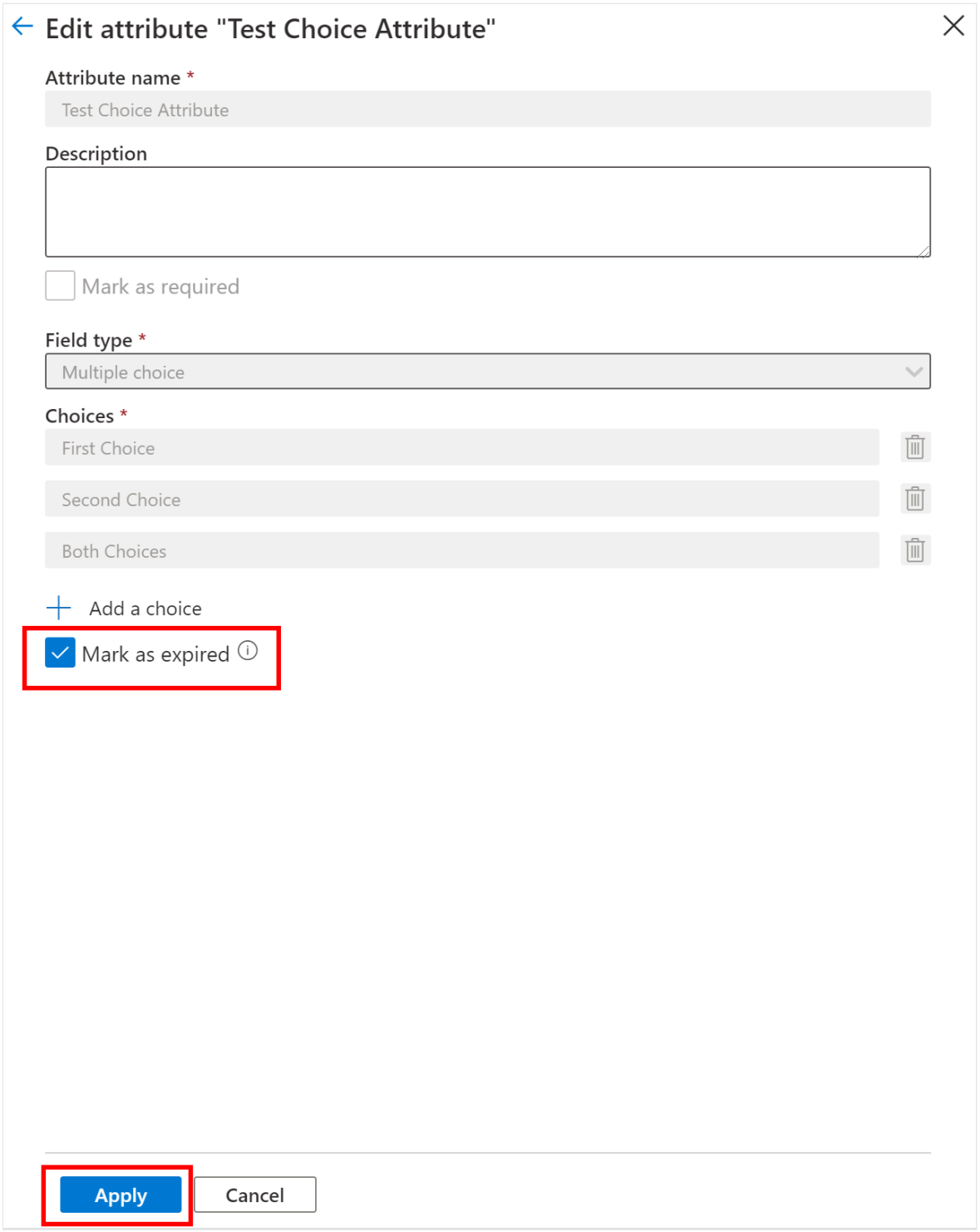
7. Once all the custom attributes are created, select **Preview** to arrange the sequence of custom attributes. You can drag and drop custom attributes in the desired sequence.



8. Once all the custom attributes are defined, select **Create** to create a term template with custom attributes.



9. Existing custom attributes can be marked as expired by checking **Mark as Expired**. Once expired, the attribute can't be reactivated. The expired attribute is greyed out for existing terms. Future new terms created with this term template will no longer show the attribute that has been marked expired.



10. To delete a term template, open the template and select **Delete**. Note, if the term template is in use, you can't delete it until you remove the glossary terms to which it's assigned.

Next steps

Follow the Tutorial: Create and import glossary terms to learn more.

Feedback

Was this page helpful? Yes No

How to bulk edit assets

Article • 02/26/2025

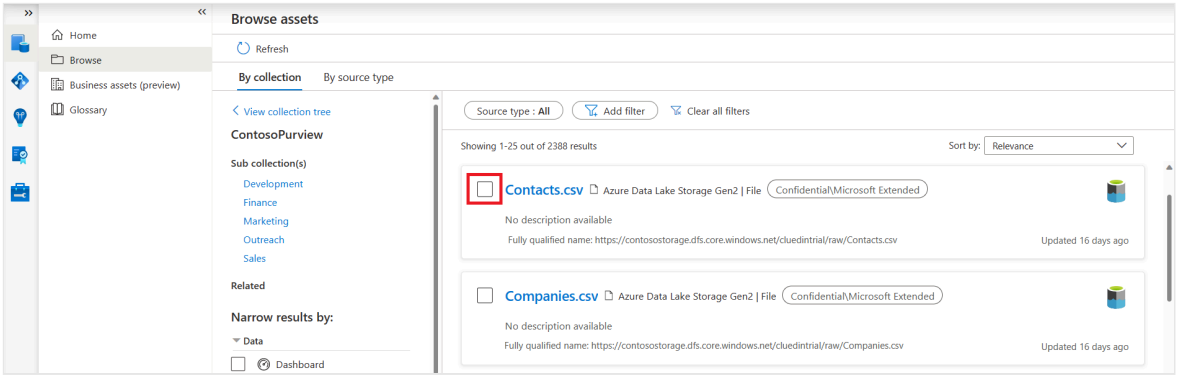
> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

This article describes how you can update assets in bulk to add glossary terms, classifications, owners, and experts on multiple assets at once.

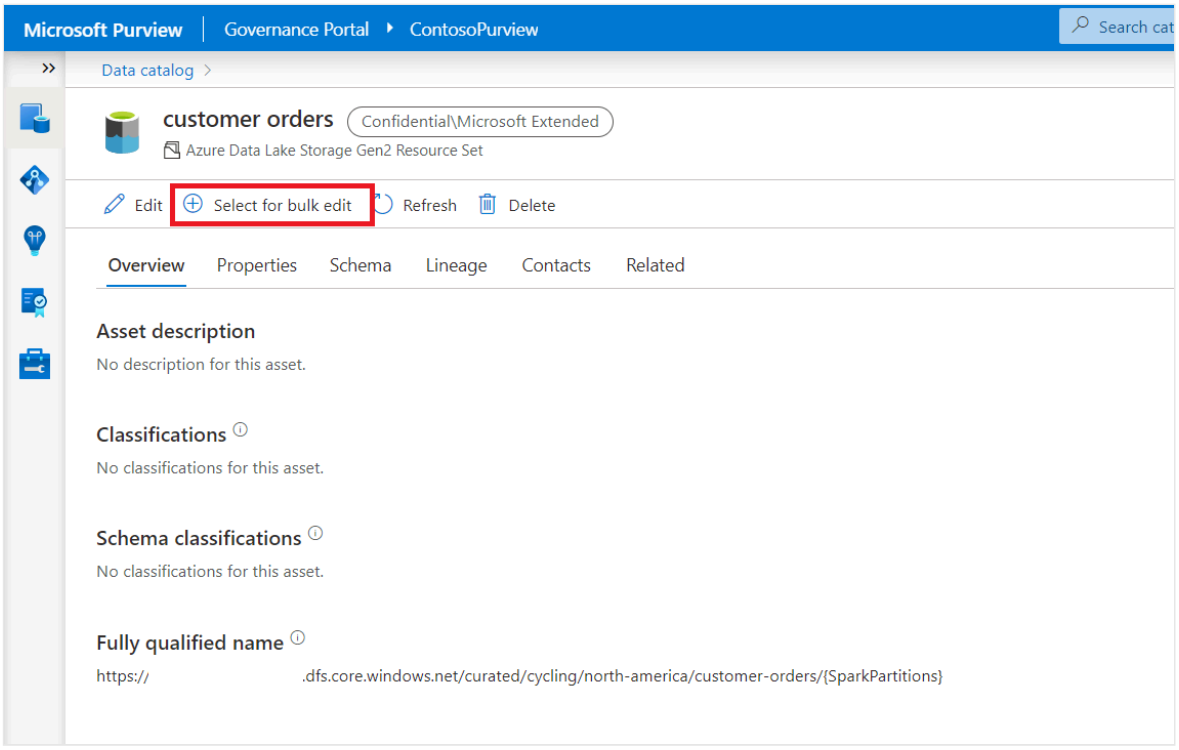
> **Note**
>
> Bulk editing as described here is for data assets. Other assets like classifications and glossary terms can't be added to the bulk edit list.

Select assets to bulk edit

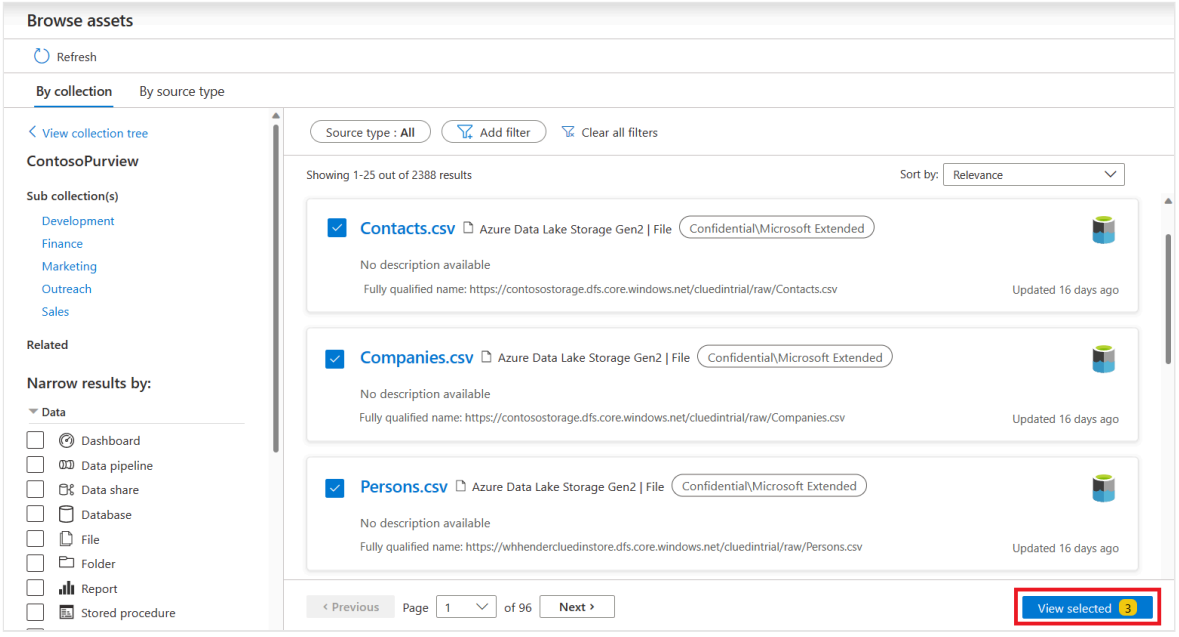
1. Use Microsoft Purview search or browse to discover assets you wish to edit.
2. In the search results, each data asset has a checkbox you can select to add the asset to the selection list.



3. You can also add an asset to the bulk edit list from the asset detail page. Select **Select for bulk edit** to add the asset to the bulk edit list.

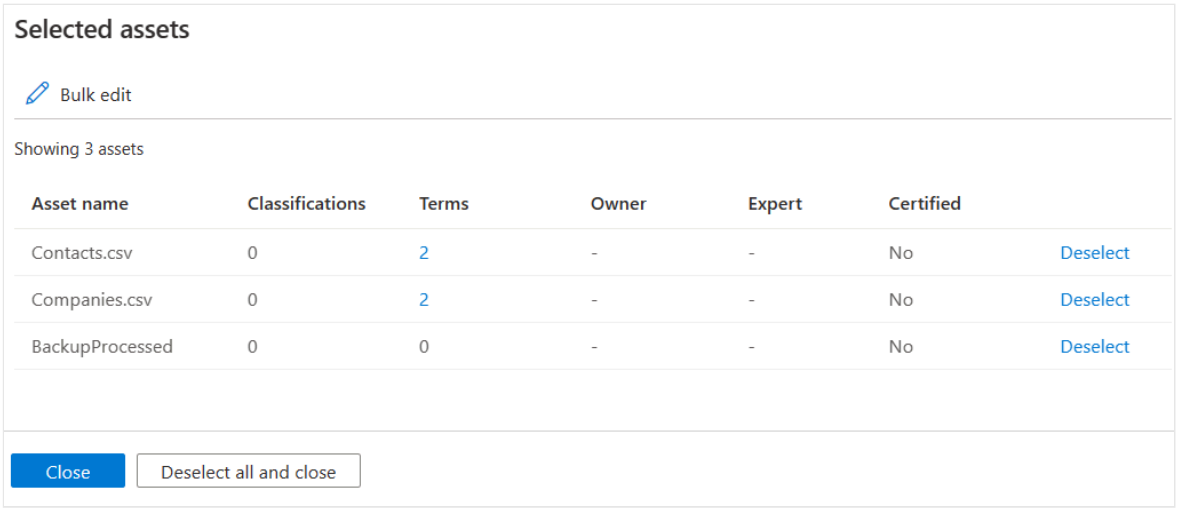


4. Select the checkbox to add it to the bulk edit list. You can see the selected assets by selecting the **View selected** button.

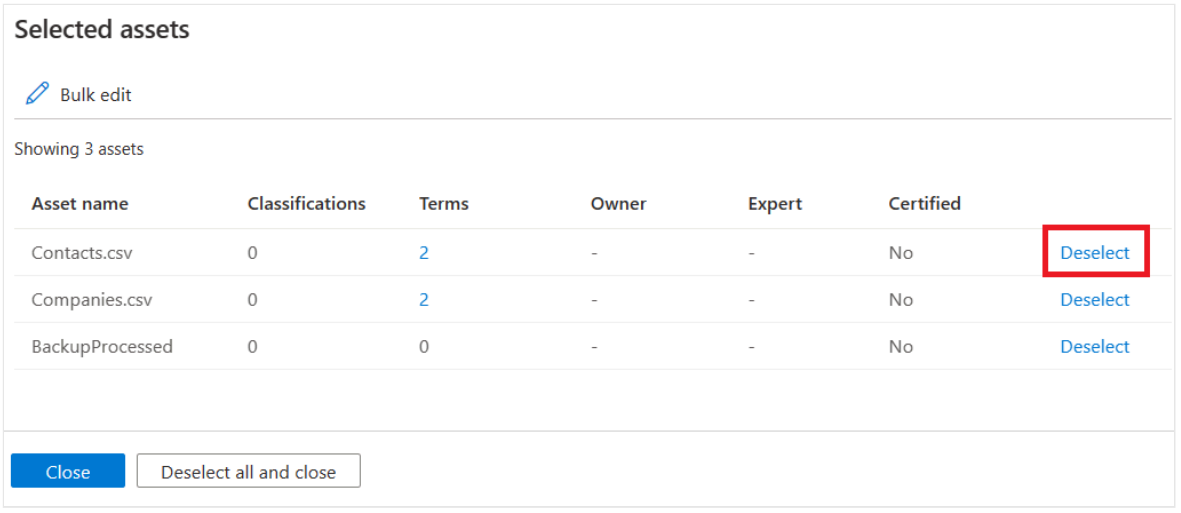


Bulk edit assets

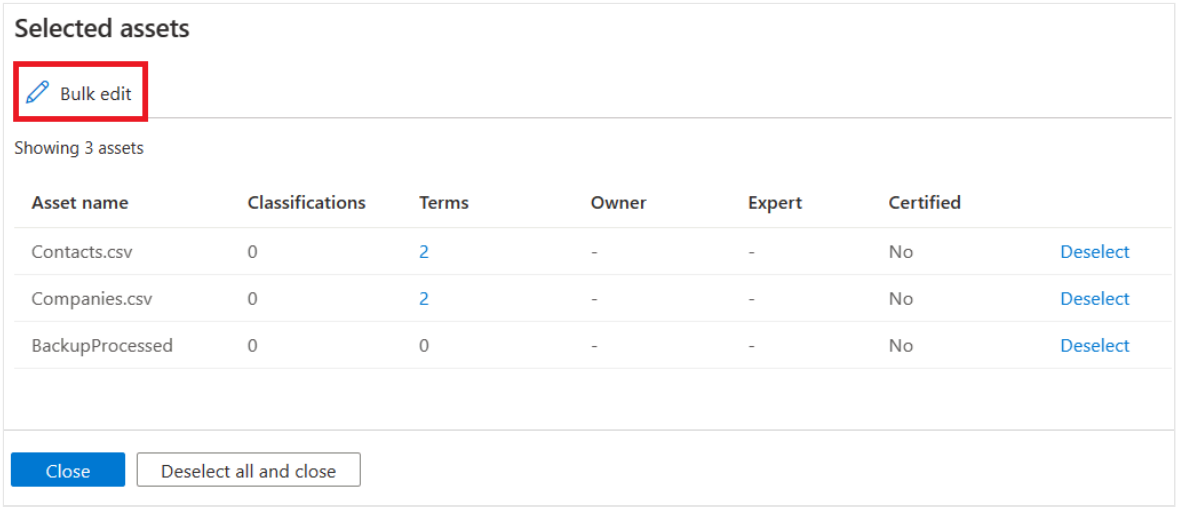
1. When all assets have been chosen, select **View selected** to pull up the selected assets.



2. Review the list and select **Deselect** if you want to remove any assets from the list.



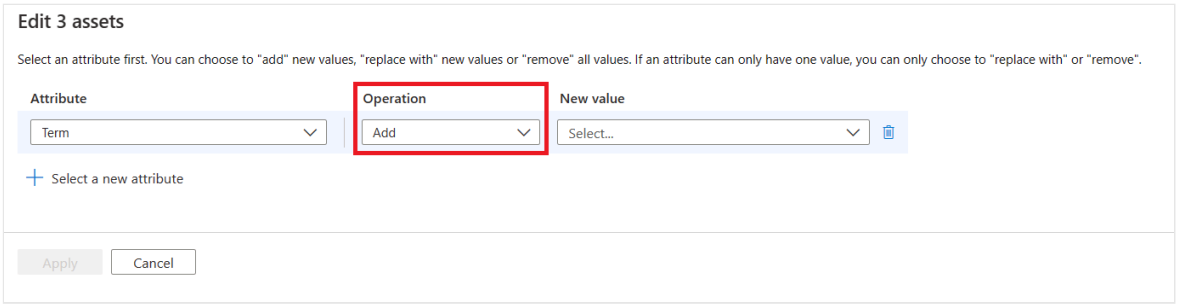
3. Select **Bulk edit** to add, remove, or replace an annotation for all the selected assets. You can edit the glossary terms, classifications, owners, or experts of an asset.



4. For each attribute selected, you can choose which edit operation to apply
   a. **Add** will append a new annotation to the selected data assets.

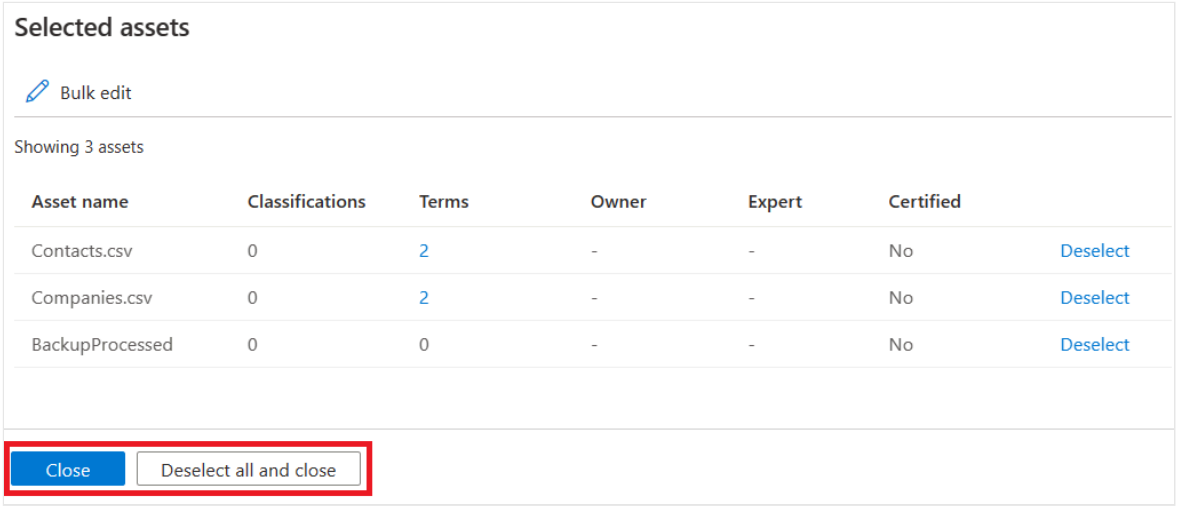
b. **Replace with** will replace all of the annotations for the selected data assets with the annotation selected.
c. **Remove** will remove all annotations for selected data assets.

You can edit multiple assets at once by selecting **Select a new attribute**.



5. When you have made all your updates, select **Apply**.

6. Once complete, close the bulk edit blade by selecting **Close** or **Remove all and close**. Close won't remove the selected assets whereas remove all and close will remove all the selected assets.



> **Important**
>
> In the UI you can currently only select up to 25 assets. The **View Selected** box will be visible only if there is at least one asset selected.

Next steps

- How to create and manage glossary terms
- How to import and export glossary terms

Feedback

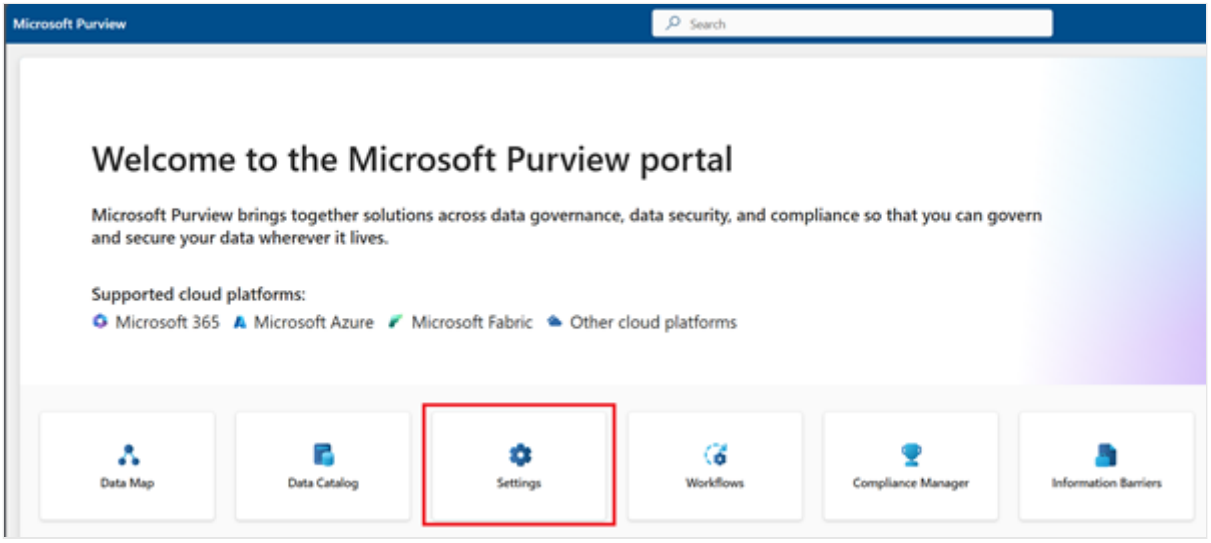
Was this page helpful? Yes No

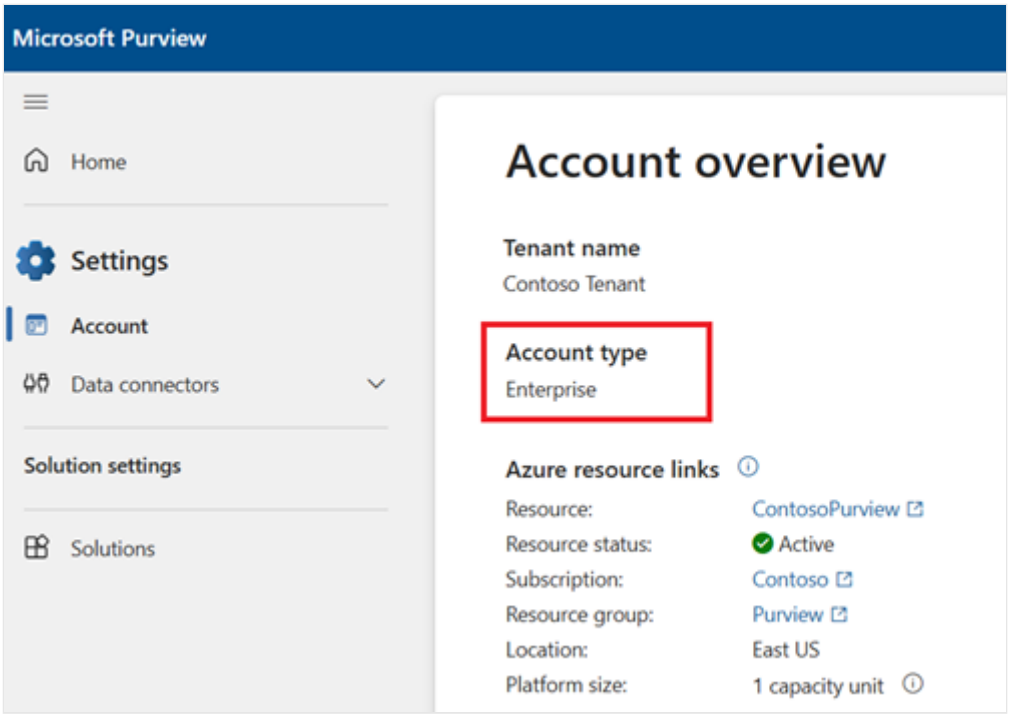
Governance roles and permissions for classic Data Catalog in free and enterprise account types

Article • 05/20/2025

> **Important**
>
> This article explains governance permissions in the free and enterprise versions of Microsoft Purview data governance solutions using the classic Microsoft Purview Data Catalog.
>
> - For permissions using the new Microsoft Purview Unified Catalog, see the data governance permissions article.
> - For general permissions in the Microsoft Purview portal see permissions in the portal.
> - For classic data governance permissions, see the permissions in the Microsoft Purview governance portal article.

Microsoft Purview data governance has two solutions in the Microsoft Purview portal, the Data Map and the catalog. These solutions use tenant/organizational-level permissions, existing data access permissions, and domain/collection permissions to provide users access to governance tools and data assets. The kind of permissions available to use depends on your Microsoft Purview account type. You can check your account type in the Microsoft Purview portal under the **Settings** card, and **Account**.





Expand table

Account Type	Tenant/Organization permissions	Data access permissions	Domain and collection permissions
Free	x	x	
Enterprise	x	x	x

For more information about each permission type, see these guides:

- Tenant/organization permissions - assigned at the organizational level, they provide general and administrative permissions.
- Domain and collection level permissions - permissions in the Microsoft Purview Data Map that grant access to data assets in Microsoft Purview.
- Data access permissions - permissions users already have on their Azure data sources.

For more information about permission based on account type, see these guides:

- Free
- Enterprise

> **Important**
>
> For users newly created in Microsoft Entra ID, it could take some time for permissions to propagate even after correct permissions have been applied.

Tenant level role groups

Assigned at the organizational level, tenant level role groups provide general, and administrative permissions for both the Microsoft Purview Data Map and classic Data Catalog. If you're managing your Microsoft Purview account or your organization's data governance strategy, you probably need one or several of these roles.

The currently available governance tenant-level role groups are:

Expand table

Role group	Description	Account type availability
Purview administrator	Create, edit, and delete domains and perform role assignments.	Free and enterprise accounts
Data Source Administrators	Manage data sources and data scans in the Microsoft Purview Data Map.	Enterprise accounts
Data Catalog Curators	Perform create, read, modify, and delete actions on catalog data objects and establish relationships between objects in the classic Data Catalog.	Enterprise accounts
Data Estate Insights Readers	Provides read-only access to all insights reports across platforms and providers in the classic Data Catalog.	Enterprise accounts
Data Estate Insights Admins	Provides admin access to all insights reports across platforms and providers in the classic Data Catalog.	Enterprise accounts

For a full list of all available roles and role groups, not just for data governance, see roles, and role groups in the Microsoft Defender XDR and Microsoft Purview portals.

How to assign and manage role groups

> **Note**
>
> Users must hold the role management role or the Global Administrator role for the organization in order to assign roles.

> **Important**

Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview **roles and permissions**.

To assign and manage roles in Microsoft Purview, follow the permissions in Microsoft Purview guide.

Permissions to search the classic Data Catalog

No specific permissions are needed in the catalog to be able to search it. However, searching the catalog will only return relevant data assets that you have permissions to view in the Data Map.

Users can find a data asset in the catalog when:

- The user has at least read permissions on an available Azure or Microsoft Fabric resource
- The user has data reader permissions on a domain or collection in the Microsoft Purview Data Map where the asset is stored

Permissions to these assets are managed at the resource level and at the Data Map level.

Domain and collection permissions

Domains and collections are tools that the Microsoft Purview Data Map uses to group assets, sources, and other artifacts into a hierarchy for discoverability and to manage access control within the Microsoft Purview Data Map.

Domain and collection roles

The Microsoft Purview Data Map uses a set of predefined roles to control who can access what within the account. These roles are currently:

- **Domain admin** (domain level only) - Can assign permissions within a domain and manage its resources.
- **Collection administrator** - a role for users that will need to assign roles to other users in the Microsoft Purview governance portal or manage collections. Collection admins can add users to roles on collections where they're admins. They can also edit collections, their details, and add subcollections. A collection administrator on the root collection also automatically has permission to the Microsoft Purview governance portal. If your **root collection administrator** ever needs to be changed, you can follow the steps in the section below.

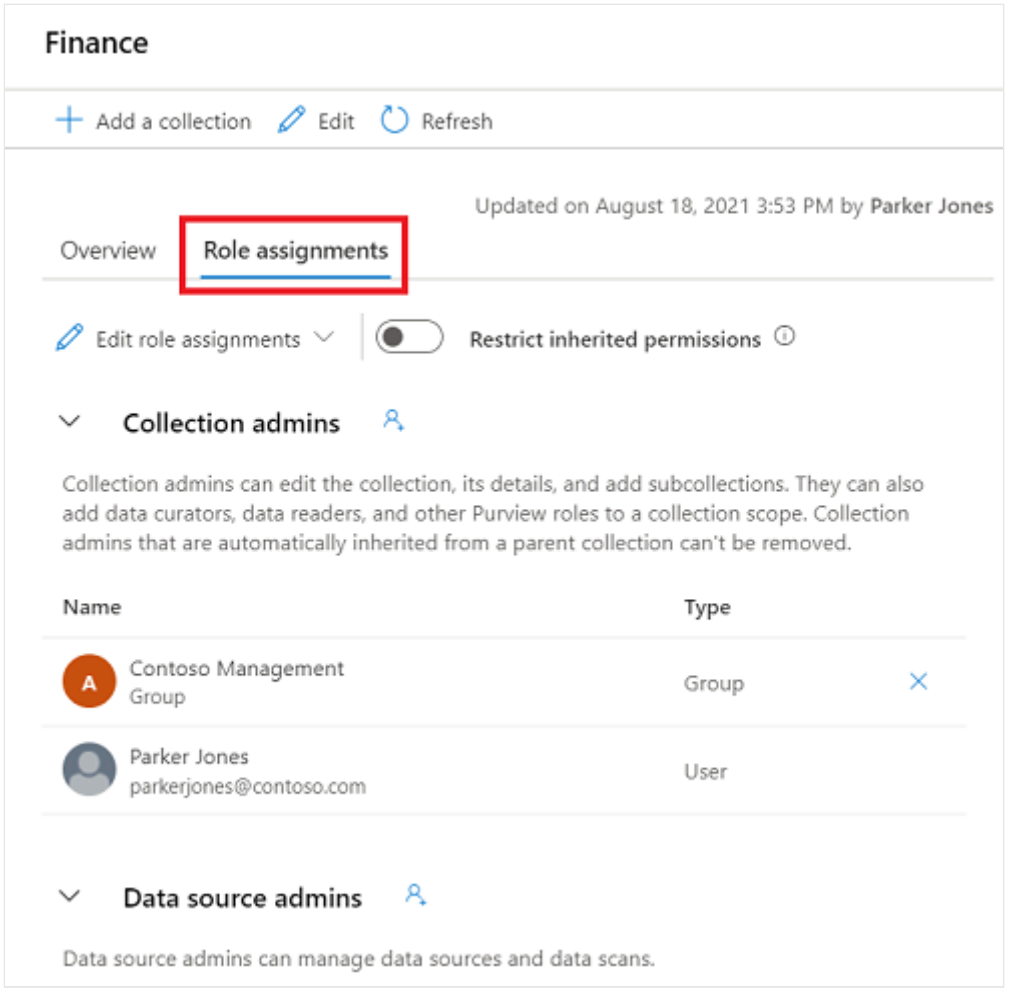
- **Data curators** - a role that provides access to the Microsoft Purview Unified Catalog to manage assets, configure custom classifications, create and manage glossary terms, and view data estate insights. Data curators can create, read, modify, move, and delete assets. They can also apply annotations to assets.
- **Data readers** - a role that provides read-only access to data assets, classifications, classification rules, collections and glossary terms.
- **Data source administrator** - a role that allows a user to manage data sources and scans. If a user is granted only to **Data source admin** role on a given data source, they can run new scans using an existing scan rule. To create new scan rules, the user must be also granted as either **Data reader** or **Data curator** roles.
- **Insights reader** - a role that provides read-only access to insights reports for collections where the insights reader also has at least the **Data reader** role. For more information, see insights permissions.
- **Policy author** - a role that allows a user to view, update, and delete Microsoft Purview policies through the Data policy app within Microsoft Purview.
- **Workflow administrator** - a role that allows a user to access the workflow authoring page in the Microsoft Purview governance portal, and publish workflows on collections where they have access permissions. Workflow administrator only has access to authoring, and so will need at least Data reader permission on a collection to be able to access the Purview governance portal.

> ⓘ **Note**
>
> At this time, Microsoft Purview policy author role is not sufficient to create policies. The Microsoft Purview data source admin role is also required.

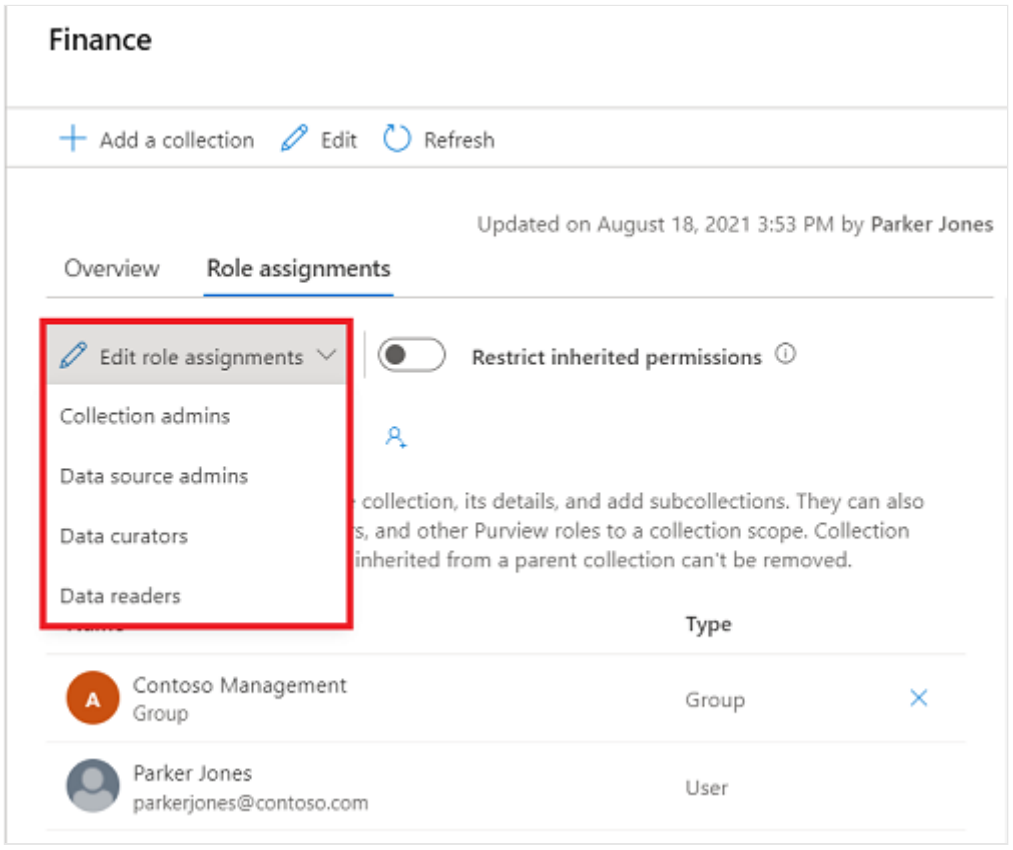
> ⓘ **Important**
>
> The user that created the account is automatically assigned domain admin on the default domain and collection admin on the root collection.

Add role assignments

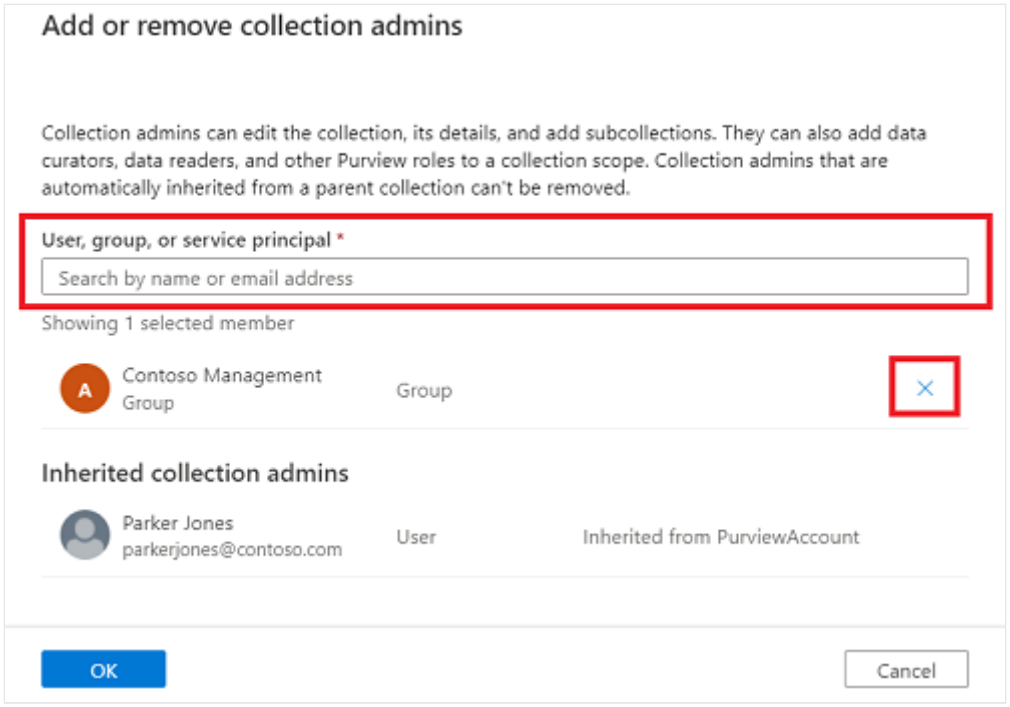
1. Open the Microsoft Purview Data Map.
2. Select the domain or collection where you want to add your role assignment.
3. Select the **Role assignments** tab to see all the roles in a collection or a domain. Only a collection admin or domain admin can manage role assignments.



4. Select **Edit role assignments** or the person icon to edit each role member.



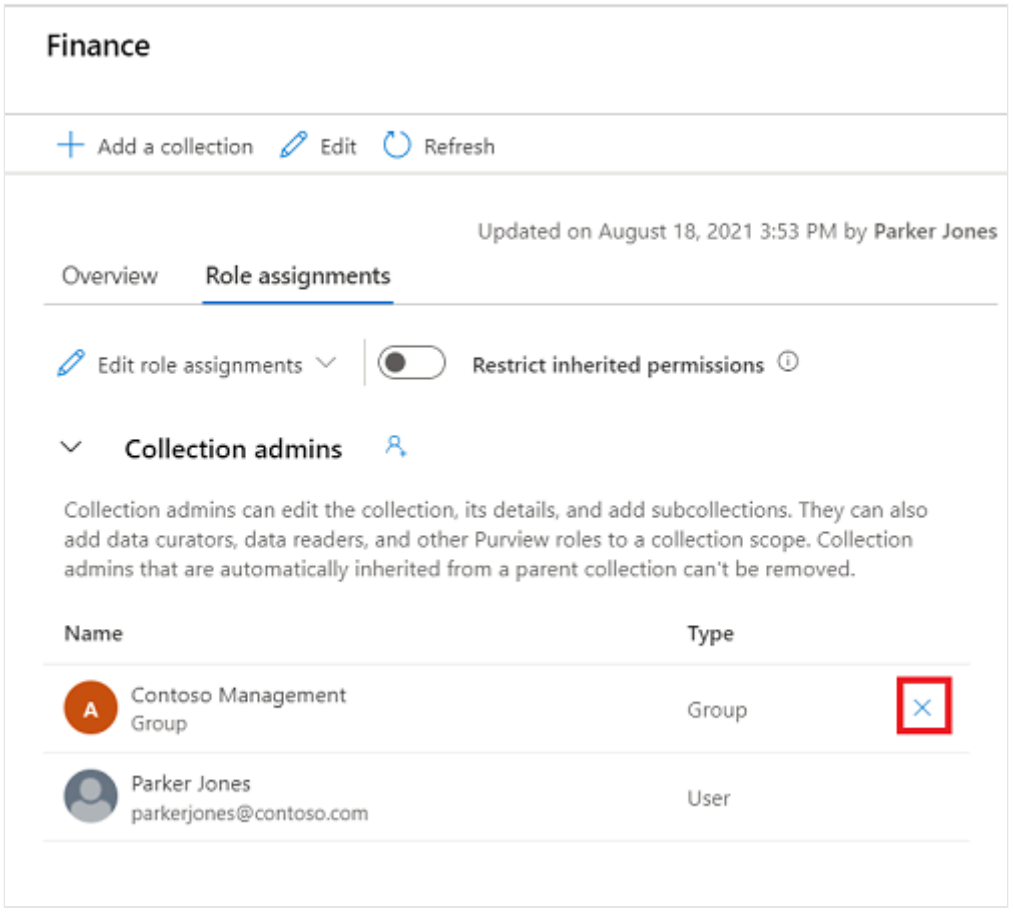
5. Type in the textbox to search for users you want to add to the role member. Select **X** to remove members you don't want to add.



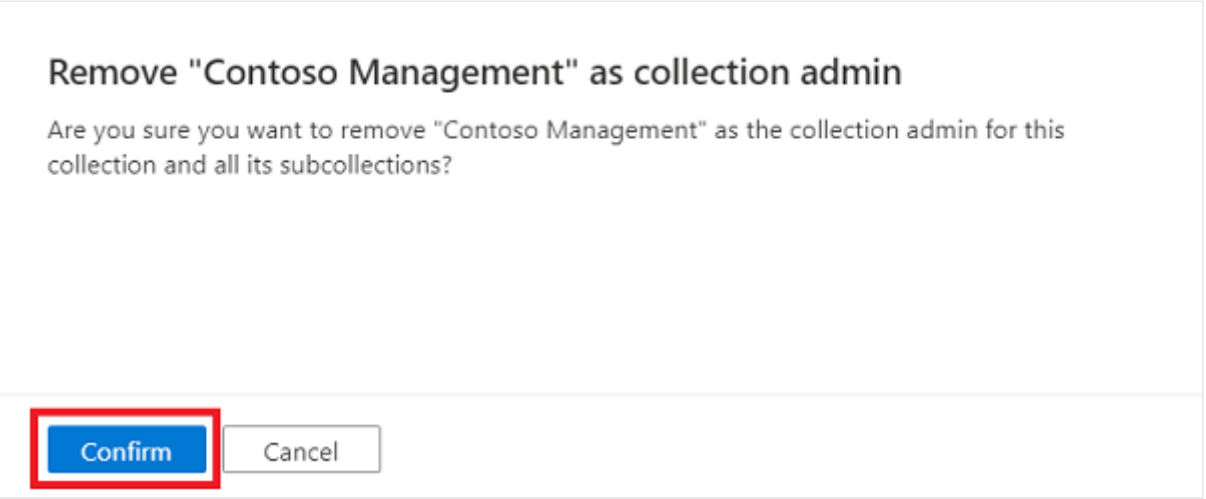
6. Select **OK** to save your changes, and you'll see the new users reflected in the role assignments list.

Remove role assignments

1. Select **X** button next to a user's name to remove a role assignment.



2. Select **Confirm** if you're sure to remove the user.



Restrict inheritance

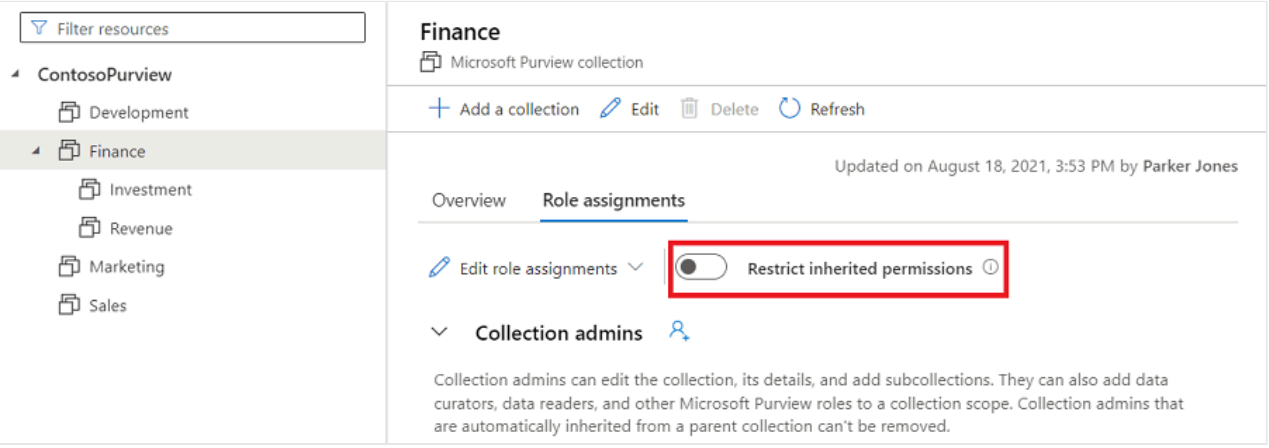
Collection permissions are inherited automatically from the parent collection. You can restrict inheritance from a parent collection at any time, using the restrict inherited permissions option.

> **Note**
>
> Currently permissions from the default domain cannot be restricted. Any permissions assigned at the default domain will be inherited by the domain's direct subcollections.

Once you restrict inheritance, you'll need to add users directly to the restricted collection to grant them access.

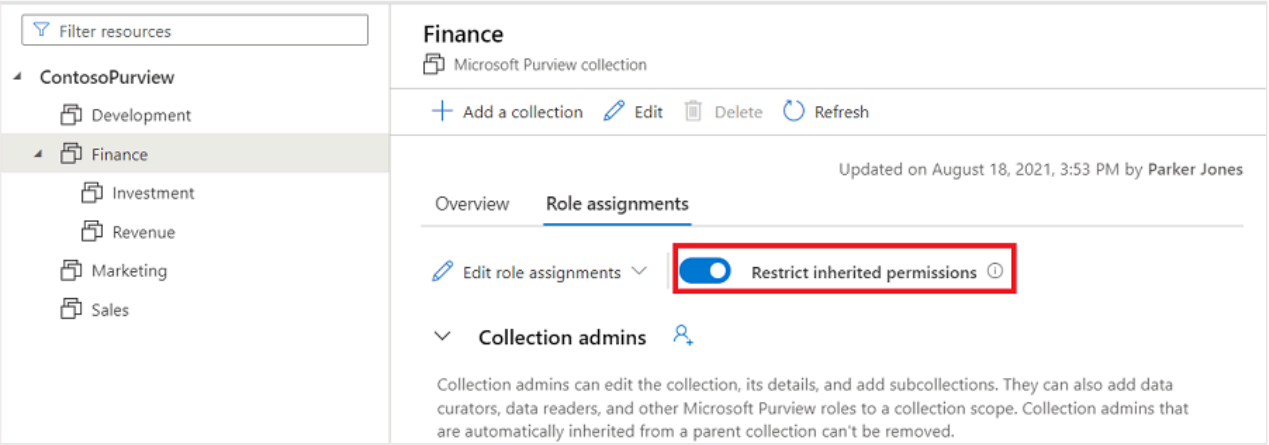
1. Navigate to the collection where you want to restrict inheritance and select the **Role assignments** tab.

2. Select **Restrict inherited permissions** and select **Restrict access** in the popup dialog to remove inherited permissions from this collection and any subcollections. Collection admin permissions won't be affected.



3. After restriction, inherited members are removed from the roles expect for collection admin.

4. Select the **Restrict inherited permissions** toggle button again to revert.



> **Tip**
>
> For more detailed information about the roles available in collections see the who should be assigned what roles table or the collections example.

Data access permissions

Data access permissions are permissions users already have on their Azure data sources. These existing permissions also grant permission to access and manage the metadata for those sources, depending on the level of permissions:

- Reader permissions
- Owner permissions

Currently, these features are only available for some Azure sources:

Expand table

Data source	Reader permission
Azure SQL Database	Reader role in Azure
Azure Blob Storage	Reader role in Azure
Azure Data Lake Storage Gen2	Reader role in Azure
Azure Subscription	Read permission on subscription

The reader role contains enough permissions, but if you're building a custom role, your users need to have these actions included:

Expand table

Data source	Reader permission
Azure SQL Database	"Microsoft.Sql/servers/read", "Microsoft.Sql/servers/databases/read", "Microsoft.Sql/servers/databases/schemas/read", "Microsoft.Sql/servers/databases/schemas/tables/read", "Microsoft.Sql/servers/databases/schemas/tables/columns/read"
Azure Blob Storage	"Microsoft.Storage/storageAccounts/read", "Microsoft.Storage/storageAccounts/blobServices/read", "Microsoft.Storage/storageAccounts/blobServices/containers/read"
Azure Data Lake Storage Gen2	"Microsoft.Storage/storageAccounts/read", "Microsoft.Storage/storageAccounts/blobServices/read", "Microsoft.Storage/storageAccounts/blobServices/containers/read"
Azure Subscription	"Microsoft.Resources/subscriptions/resourceGroups/read"

Reader permissions

Users who have at least the reader role on available Azure resources are also able to access those resources metadata in the free and enterprise account types.

Users can search and browse for assets from these sources in the classic Data Catalog and view their metadata.

These are the permissions needed on the resources for users to be considered readers:

Expand table

Data source	Reader permission
Azure SQL Database	Reader role in Azure
Azure Blob Storage	Reader role in Azure
Azure Data Lake Storage Gen2	Reader role in Azure
Azure Subscription	Read permission on subscription

Owner permissions

Users who have the owner role or write permissions on available Azure resources can access and edit metadata for those resources in free and enterprise account types.

Owning users can search and browse for assets from these sources in the catalog and view metadata. Users can also update and manage the metadata for those resources. Learn more about metadata curation.

These are the permissions needed on the resources for users to be considered owners:

Expand table

Data source	Owner permission
Azure SQL Database	"Microsoft.Sql/servers/write", "Microsoft.Sql/servers/databases/write", "Microsoft.Authorization/roleAssignments/write"
Azure Blob Storage	"Microsoft.Storage/storageAccounts/write", "Microsoft.Authorization/roleAssignments/write"
Azure Data Lake Storage Gen2	"Microsoft.Storage/storageAccounts/write", "Microsoft.Authorization/roleAssignments/write"

Permissions in the free version

All users are able to view data assets for available sources where they have at least **Read** permissions already. Owning users are able to manage the metadata for available assets where they have at least **Owner/Write** permissions already. For more information, see the data access permissions section.

Extra permissions can also be assigned using tenant level role groups.

> **Important**
>
> For users newly created in Microsoft Entra ID, it may take some time for permissions to propagate even after correct permissions have been applied.

Permissions in enterprise version of Microsoft Purview

All users are able to view data assets for available sources where they have at least **Read** permissions already. Owning users are able to manage the metadata for assets where they

have at least **Owner/Write** permissions already. For more information, see the data access permissions section.

Extra permissions can also be assigned using tenant level role groups.

Permissions can also be assigned in the Microsoft Purview Data Map so users can browse assets in the Data Map or classic Data Catalog search that they don't already have data access to.

Azure Storage in-place data sharing with Microsoft Purview (preview)

Article • 01/28/2025

> **Note**
>
> Microsoft Purview Data Map (classic), Data Catalog (classic), Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

> **Important**
>
> This feature is currently being retired, with support provided until September 2025 to allow for migration to Microsoft Fabric external data sharing.

Traditionally, organizations have shared data with internal teams or external partners by generating data feeds, requiring investment in data copy and refresh pipelines. The result is higher cost for data storage and movement, data proliferation (that is, multiple copies of data), and delay in access to time-sensitive data.

With Microsoft Purview Data Sharing, data providers can now share data **in-place** from Azure Data Lake Storage Gen2 and Azure Storage accounts, both within and across organizations. Share data directly with users and partners without data duplication and centrally manage your sharing activities from within Microsoft Purview.

With Microsoft Purview Data Sharing, data consumers can now have near real-time access to shared data. Storage data access and transactions are charged to the data consumers based on what they use, and at no more cost to the data providers.

Get started

Get started with Microsoft Purview in-place data sharing for Azure Storage by reviewing the next steps or following the Data Sharing Quickstart.

> **Important**

Data sharing is accessed through the classic Microsoft Purview governance portal: https://web.purview.azure.com

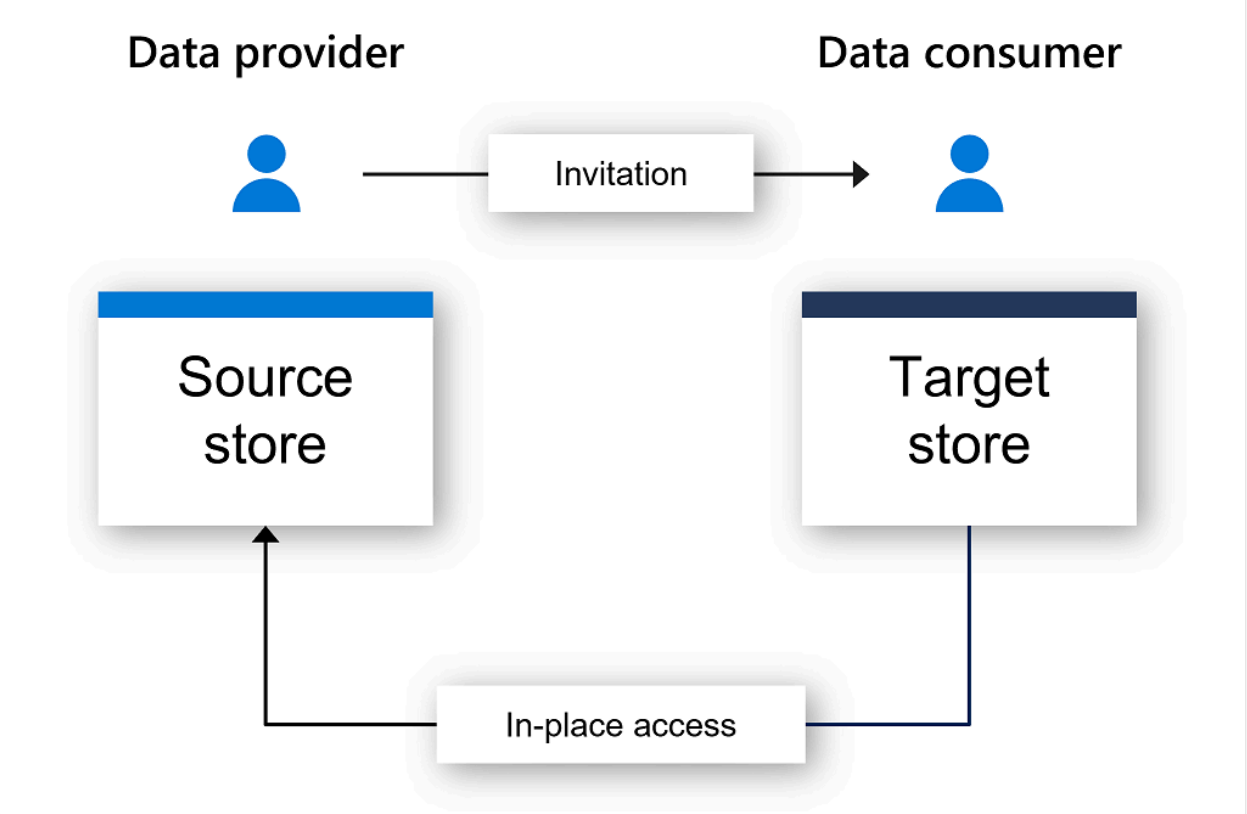
Available sources

Expand table

Supported data store	Data Sharing
Azure Blob Storage	Yes
Azure Data Lake Storage Gen2	Yes

How in-place data sharing works

Microsoft Purview enables sharing of files and folders in-place from ADLS Gen2 and Blob storage accounts.



A data provider creates a share by selecting a data source that is registered in Microsoft Purview, choosing which files and folders to share, and who to share them with. Microsoft Purview then sends an invitation to each data consumer.

When a consumer accepts the invitation, they specify a target storage account in their own Azure subscription that they'll use to access the shared data. This establishes a

sharing relationship between the provider and consumer storage accounts. This sharing relationship provides data consumer read-only access to shared data through the consumer's storage account. Any changes to the data in the provider's source storage account is reflected in near real-time in the consumer's storage account.

The data provider pays for data storage and their own data access, while the data consumer pays for their own data access transactions.

Data providers can revoke access to the shared data at any time, or set a share expiration time for time-bound access to data. Data consumers can also terminate access to the share at any time.

Where data is stored

Microsoft Purview Data Sharing only stores metadata about your share. It doesn't store a copy of the shared data itself. The data is stored in the underlying source storage account that is being shared. You can have your storage accounts in a different Azure region than your Microsoft Purview account.

Key capabilities

- Share data within the organization or with partners and customers outside of the organization (within the same Azure tenant or across different Azure tenants).
- Share data from ADLS Gen2 or Blob storage in-place without data duplication.
- Share data with multiple recipients.
- Access shared data in near real time.
- Manage sharing relationships and keep track of who the data is shared with/from, for each ADLSGen2 or Blob Storage account.
- Terminate share access at any time.
- Flexible experience through Microsoft Purview governance portal or via REST APIs.

Data sharing scenarios

Microsoft Purview Data Sharing can help with various data sharing scenarios, including:

- Collaborate with external business partners while maintaining data security in your own environment.
- Outsource data transformation and processing to third party ISVs or data aggregators by sharing raw data and receiving normalized data and analytics results back.

- Automate sharing of big data (for example: IoT data, scientific data, satellite and surveillance images or videos, financial market data) in near real time and without data duplication.
- Share data between different departments within the organization.

Support

If you have any questions, please reach out to your account manager or customer support. If you have a support plan and you need technical help, create a support request.

1. For **Issue type**, select *Technical*.
2. For **Subscription**, select your subscription.
3. For **Service**, select *My services*, then select *Service type*.
4. Select the Microsoft Purview resource that you are creating a support request for.
5. For **Summary**, type a description of your issue or question.
6. For **Problem type**, select *Data Share*.
7. For **Problem subtype**, select appropriately.

Next Steps

- Quickstart: share data
- FAQ for data sharing
- How to share data
- How to receive a data share

Feedback

Was this page helpful? Yes No

Share Azure Storage data in-place with Microsoft Purview Data Sharing (preview)

Article • 01/28/2025

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

> **Important**
>
> This feature is currently being retired, with support provided until September 2025 to allow for migration to Microsoft Fabric external data sharing.

Microsoft Purview Data Sharing supports in-place data sharing from Azure Data Lake Storage (ADLS Gen2) to ADLS Gen2, and Blob storage account to Blob storage account. This article explains how to share data using Microsoft Purview.

Prerequisites to share data

Microsoft Purview prerequisites

- A Microsoft Purview account.
- A minimum of Data Reader role is needed on a Microsoft Purview collection to use data sharing in the governance portal. Refer to Microsoft Purview permissions to learn more about the Microsoft Purview collection and roles.
- To use the SDK, no Microsoft Purview permissions are needed.
- Your data recipient's Azure sign-in email address, or the object ID and tenant ID of the recipient application, that you'll use to send the invitation to receive a share. The recipient's email alias won't work.

Azure Storage account prerequisites

- Your Azure subscription must be registered for the **AllowDataSharing** preview feature. (If you haven't already registered this preview feature, consider using Microsoft Fabric external data sharing instead, since Microsoft Purview Data Sharing will be retired in September 2025.) Or, contact support.
- A source storage account **created after the registration step is completed**. Source storage account can be in a different Azure region from your Microsoft Purview account, but needs to follow the available configurations.
- You need the **Owner** or **Storage Blob Data Owner** role on the source storage account to be able to share data. You can find more details on the ADLS Gen2 or Blob storage data source page.
- If the source storage account is in a different Azure subscription than the one for Microsoft Purview account, the Microsoft. Purview resource provider needs to be registered in the Azure subscription where the Storage account is located. It's automatically registered at the time of share provider adding an asset if the user has permission to do the `/register/action` operation and therefore, Contributor or Owner roles to the subscription where the Storage account is located. This registration is only needed the first time when sharing or receiving data into a storage account in the Azure subscription.
- A storage account needs to be registered in the collection to create a share using the Microsoft Purview portal experience. For instructions to register, see the ADLS Gen2 or Blob storage data source pages. This step isn't required to use the SDK.

Create a share

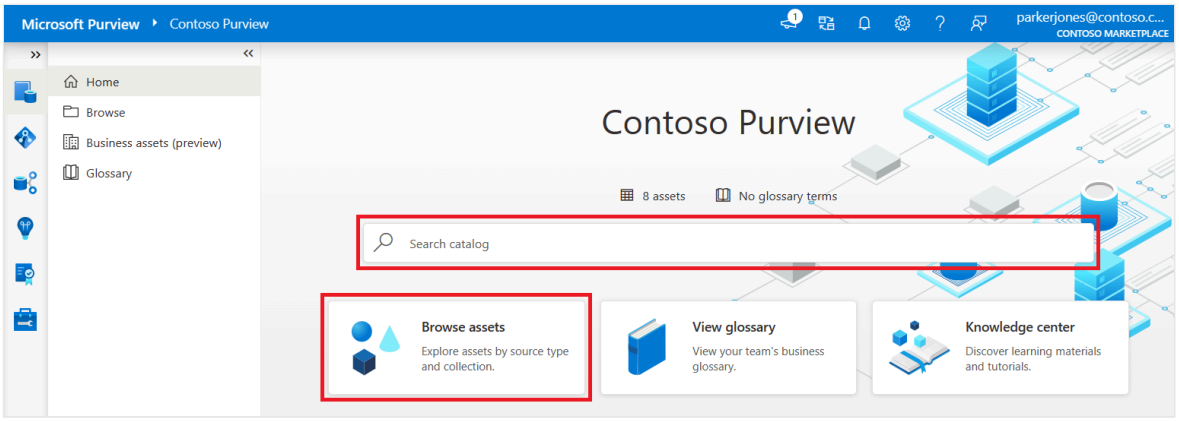
There are two ways you can create a data share:

- Directly from the data asset
- Use the data sharing application

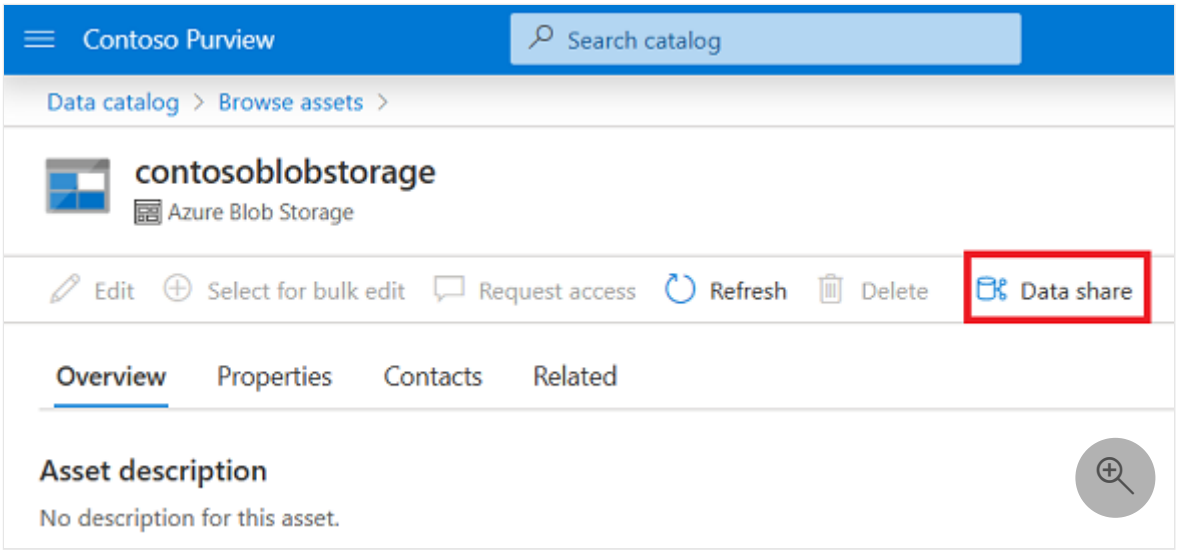
Create share from asset

1. You can create a share by starting from **Data Catalog**

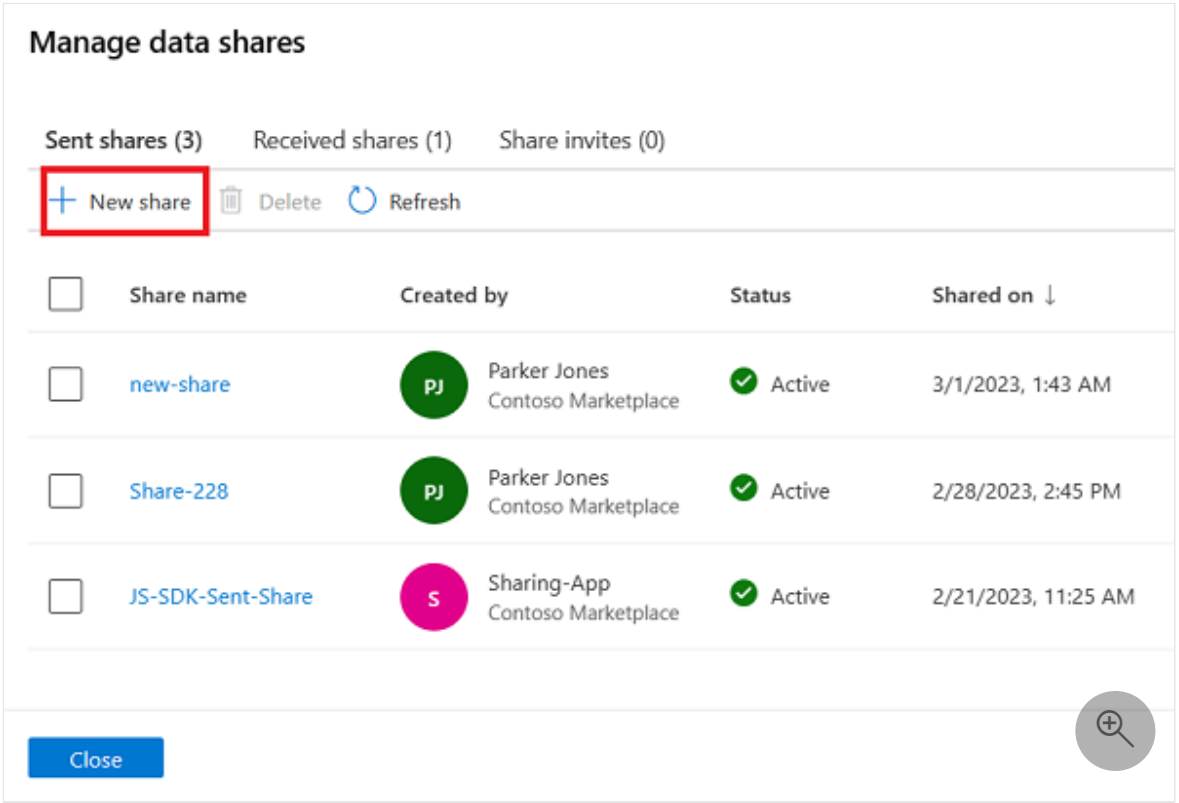
   Within the classic Microsoft Purview governance portal or the new Microsoft Purview portal, find the Azure Storage or Azure Data Lake Storage (ADLS) Gen 2 data asset you would like to share data from using either the data catalog search or browse.



2. Once you have found your data asset, select the **Data Share** button.



3. Select **+New Share**.



4. Follow the rest of the steps to create your data share.

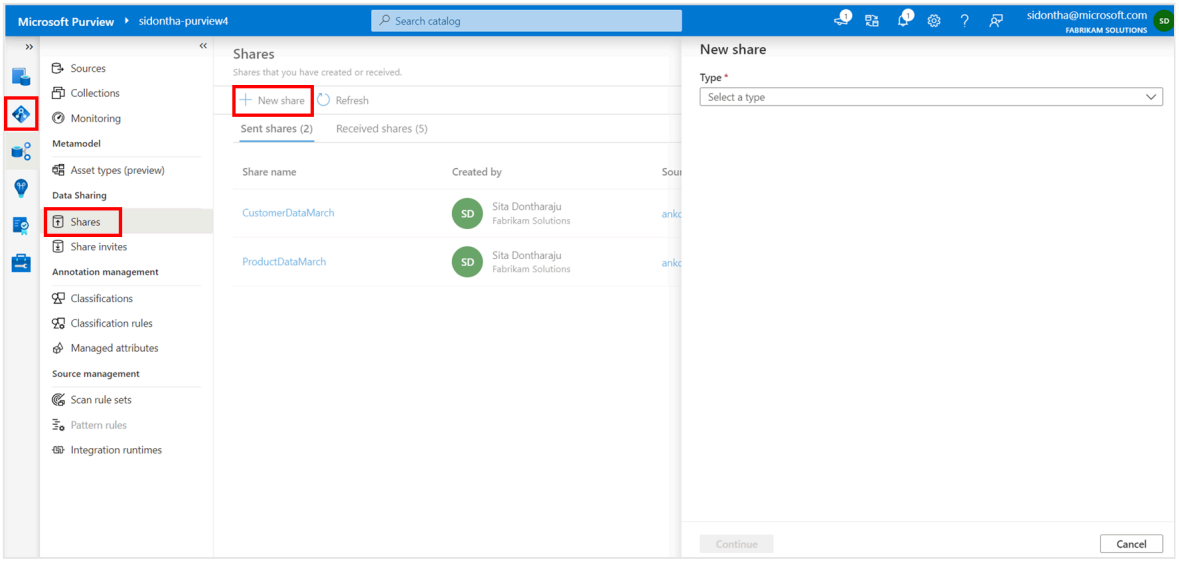
Create share from application

1. If you're using the new Microsoft Purview experience You can create a share by starting from the **Data Catalog**.

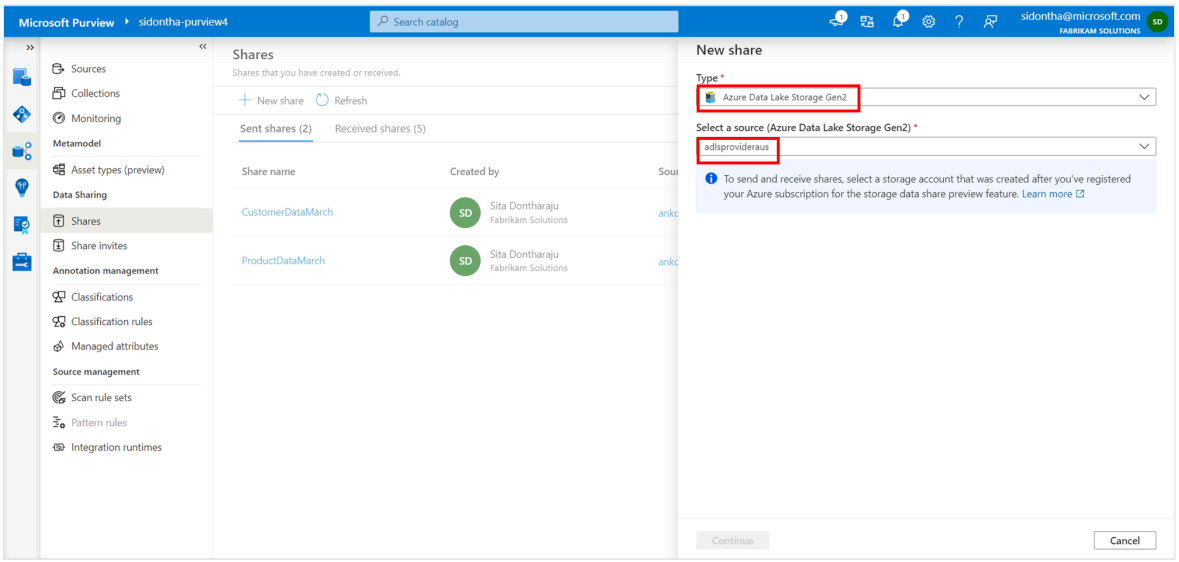
   Open the Microsoft Purview portal. Navigate to the **Data Catalog** application. Then select **Shares**. Select **+New Share**.

2. If you're using the classic Microsoft Purview experience You can create a share by starting from the **Data Map**.

   Open the Microsoft Purview governance portal. Select the **Data Map** icon from the left navigation. Then select **Shares**. Select **+New Share**.



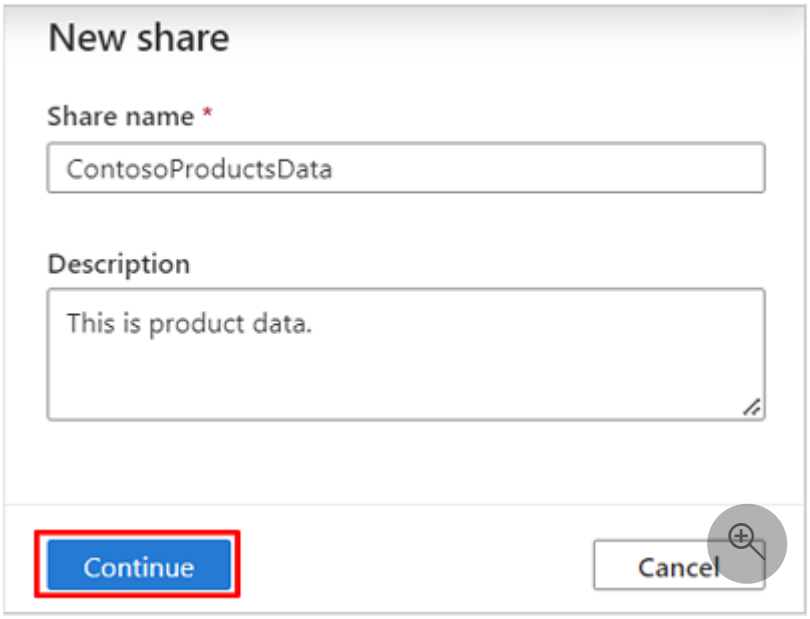
3. From either location, select the Storage account type and the Storage account you want to share data from. Then select **Continue**.



4. Follow the rest of the steps to create your data share.

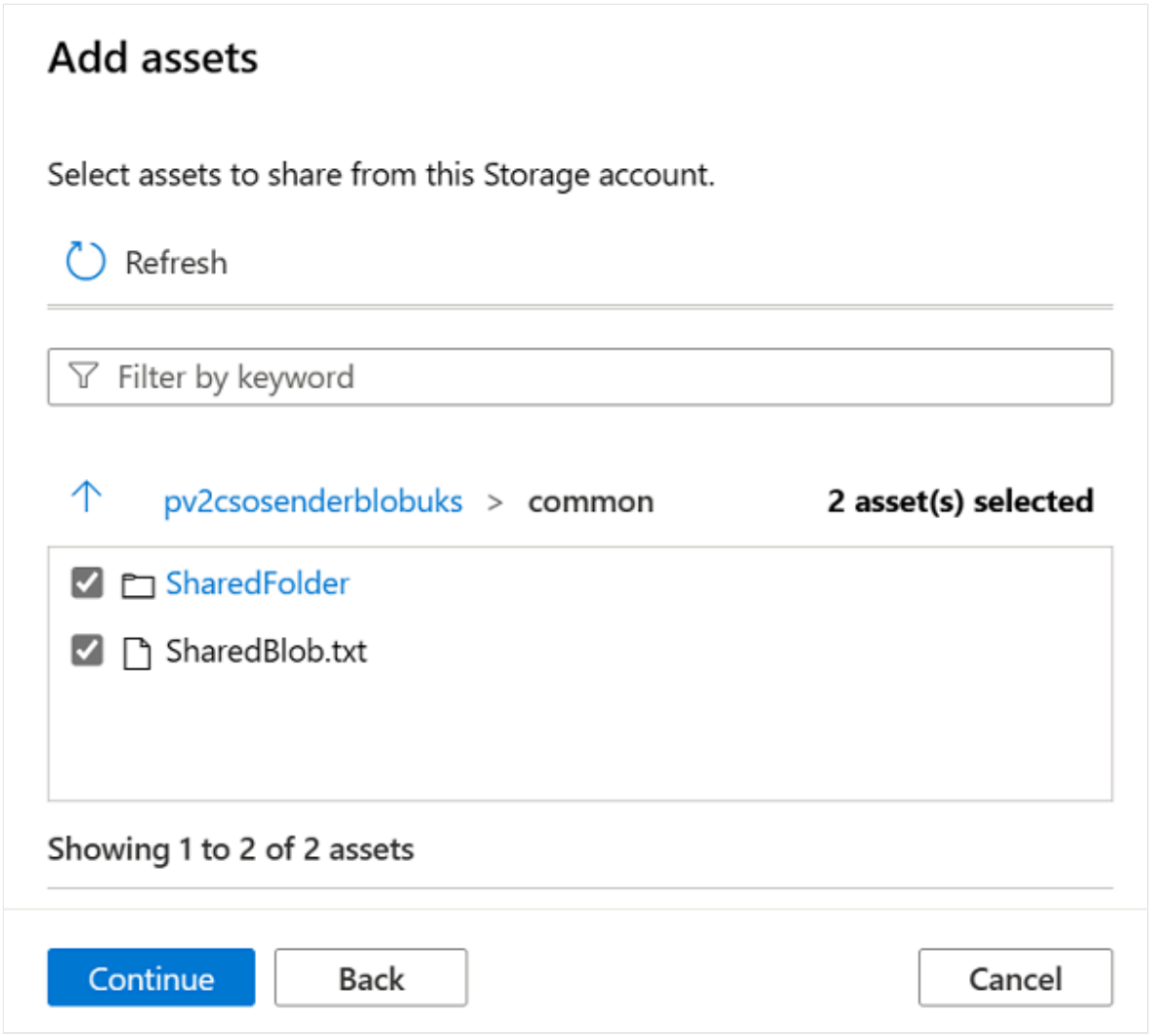
Create share

1. Specify a name and a description of share contents (optional). Then select **Continue**.

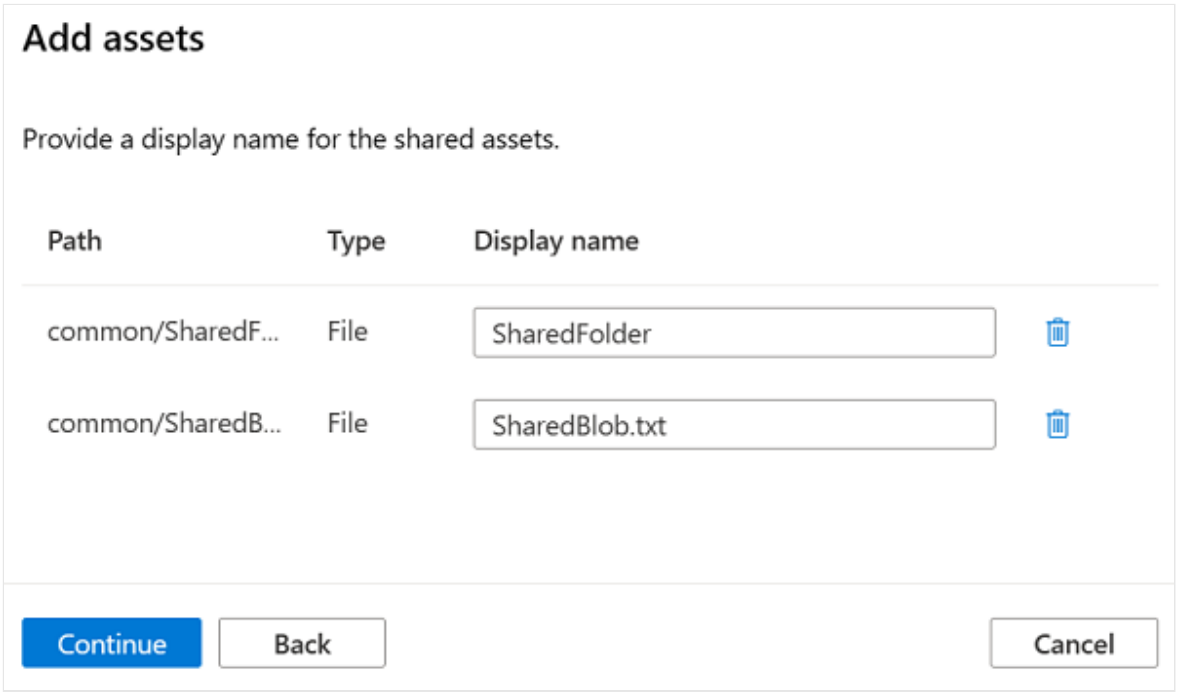


2. Search for and add all the assets you'd like to share out at the container, folder, and file level, and then select **Continue**.

> **Important**
>
> Only containers, files, and folders that belong to the current Blob or ADLSGen2 Storage account can be added to the share.

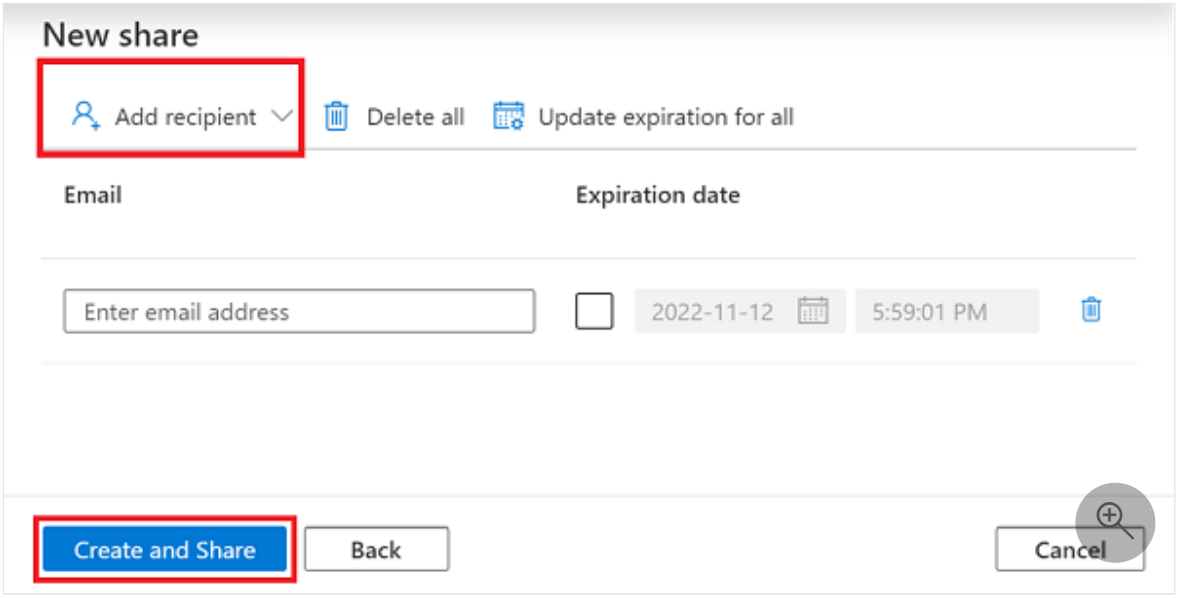


3. You can edit the display names the shared data will have, if you like. Then select **Continue**.

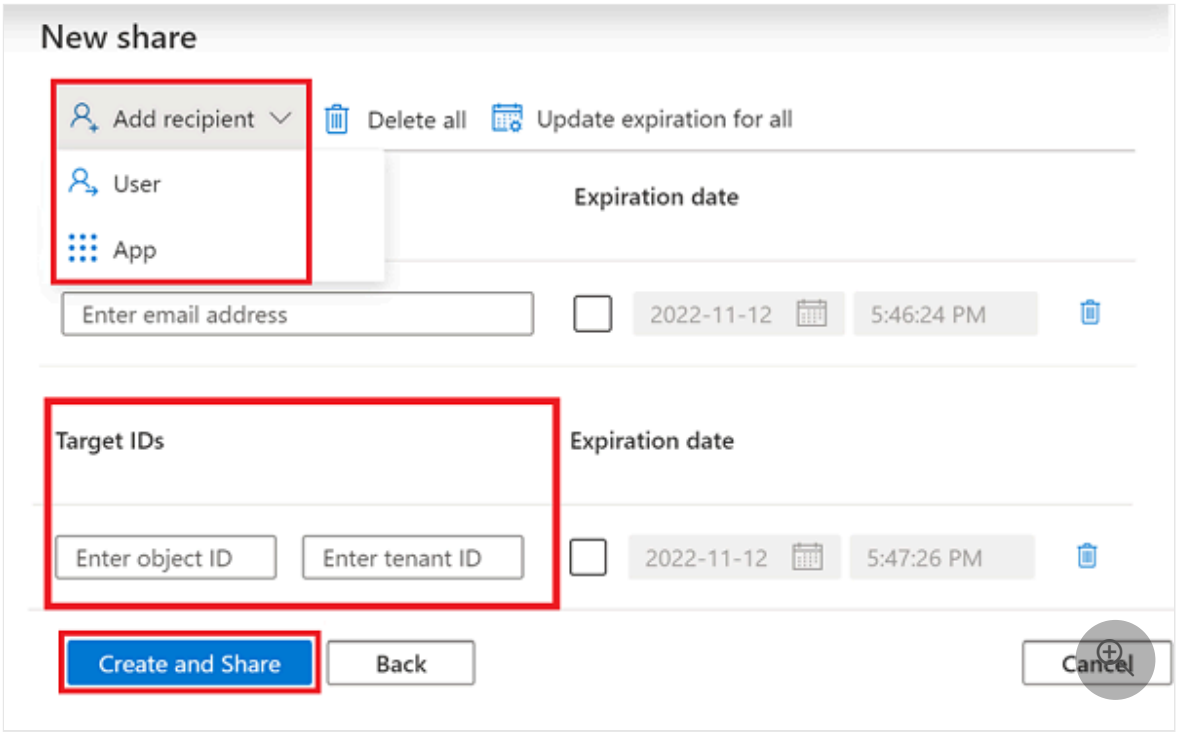


4. Select **Add Recipient** and select **User** or **App**.

To share data to a user, select **User**, then enter the Azure sign-in email address of who you want to share data with. By default, the option to enter email address of user is shown.



To to share data with a service principal, select **App**. Enter the object ID and tenant ID of the recipient you want to share data with.



5. Select **Create and Share**. Optionally, you can specify an **Expiration date** for when to terminate the share. You can share the same data with multiple recipients by selecting **Add Recipient** multiple times.

You've now created your share. The recipients of your share will receive an invitation and they can view the share invitation in their Microsoft Purview account.

When a share is created, a new asset of type sent share is ingested into the Microsoft Purview catalog, in the same collection as the storage account from which you created the share. You can search for it like any other asset in the data catalog.

You can also track lineage for data shared using Microsoft Purview. See, Microsoft Purview Data Sharing lineage to learn more about share assets and data sharing lineage.

> ⓘ **Note**
>
> Shares created using the SDK without registering the storage account with Microsoft Purview will not be ingested into the catalog. User can register their storage account if desired. If a storage account is un-registered or re-registered to a different collection, share assets of that storage account continue to be in the initial collection.

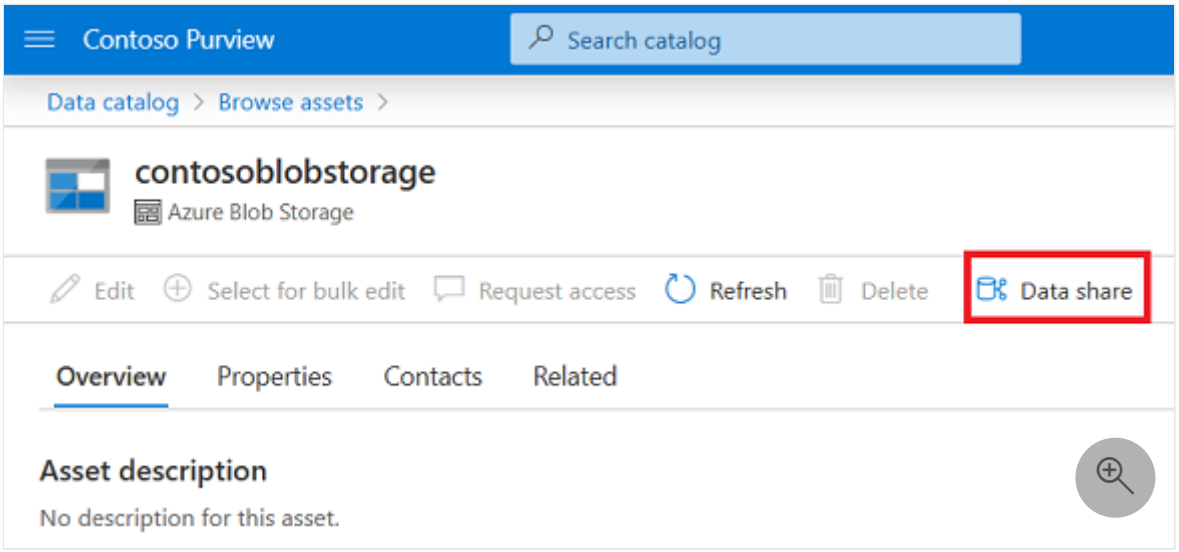
Update a sent share

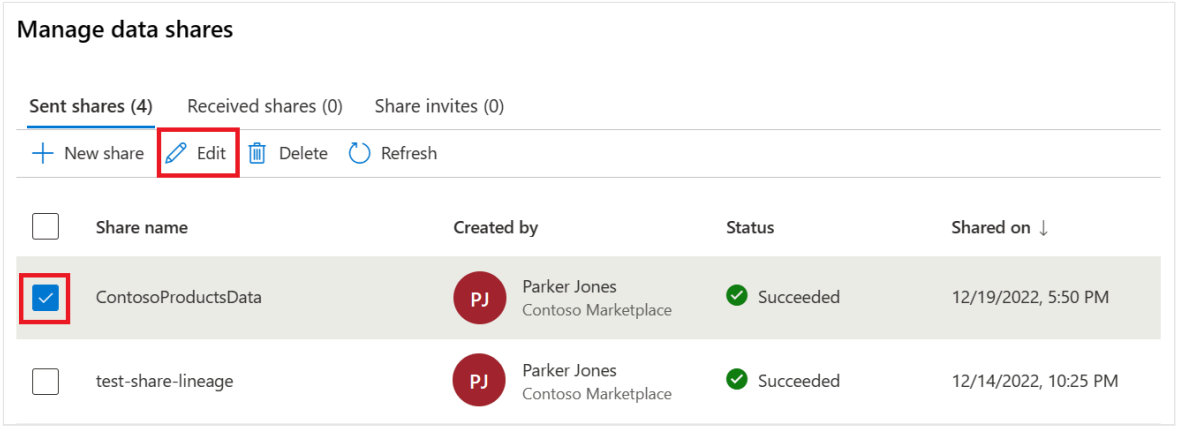
Once a share is created, you can update description, assets, and recipients.

> ⓘ **Note**
>
> If you only have the **Reader** role on the source storage account, you will be able to view list of sent shares and received shares but not edit. You can find more details on the ADLS Gen2 or Blob storage data source page.

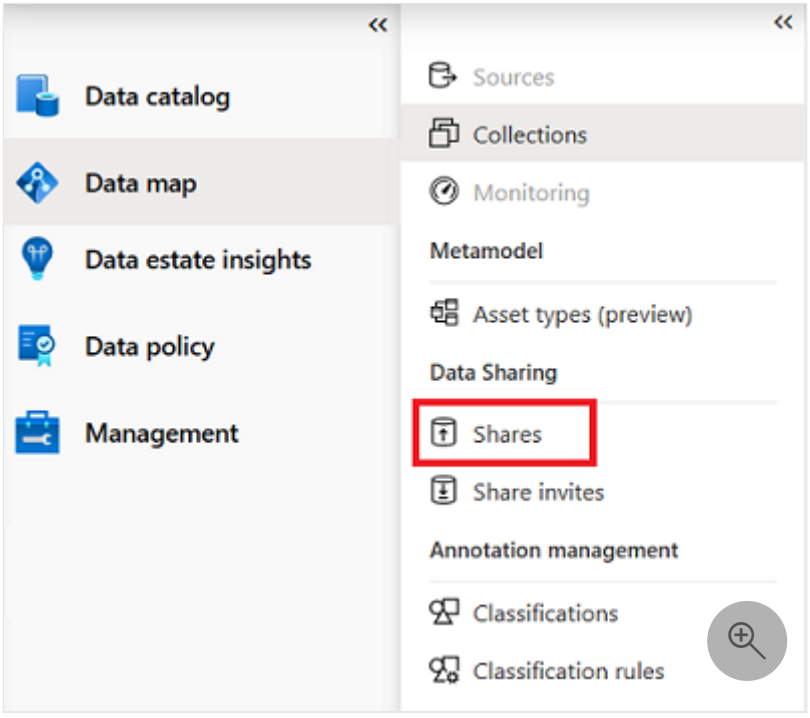
You can find your sent shares one of two ways:

- Access the blob storage or ADLS Gen2 asset where the data was shared from in the data catalog. Open it, then select **Data Share**. There you're able to see all the shares for that asset. Select a share, and then select the **Edit** option.





- For shares that you sent, in the classic Microsoft Purview governance portal you can find them in the **Shares** menu in the Microsoft Purview Data Map. In the new Microsoft Purview portal you can find them in the **Data Catalog** application in the **Shares** menu. There you're able to see all the shares you have sent. Select a share, and then select the **Edit** option.

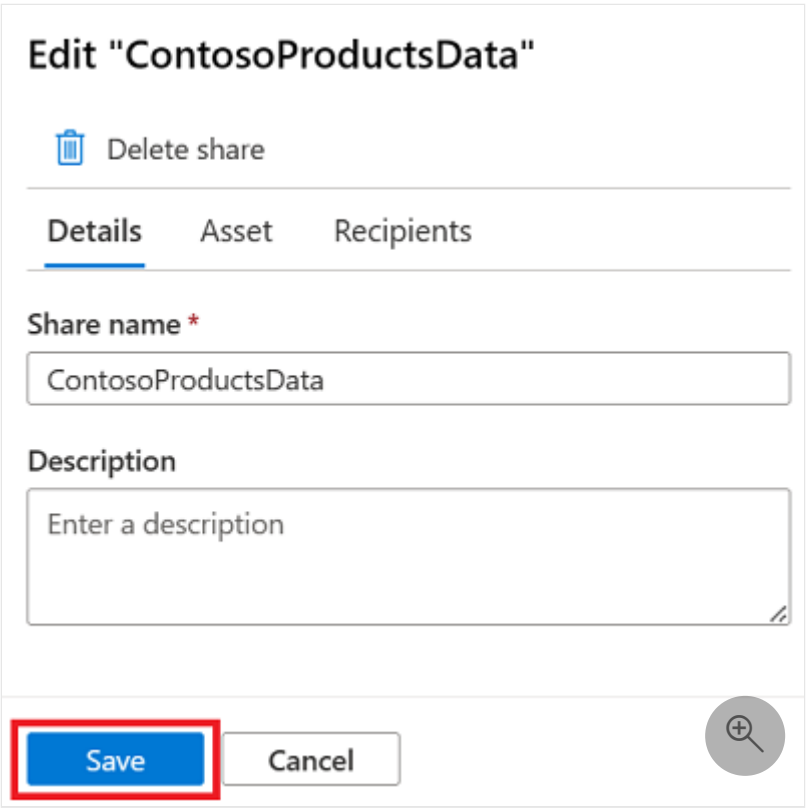


From any of these places you can:

- Edit share details
- Edit shared assets
- Edit share recipients
- Delete your share

Edit details

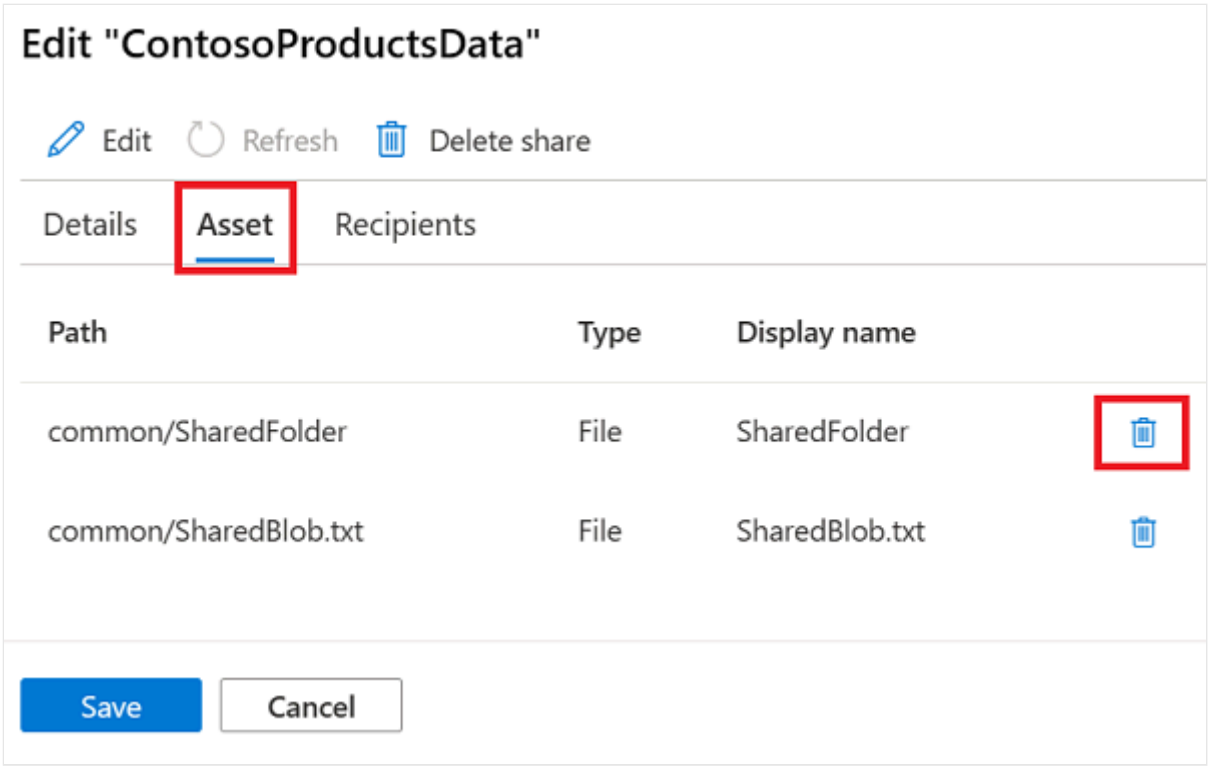
On the **Details** tab of the edit share page, you can update the share name and description. Save any changes by selecting **Save**.



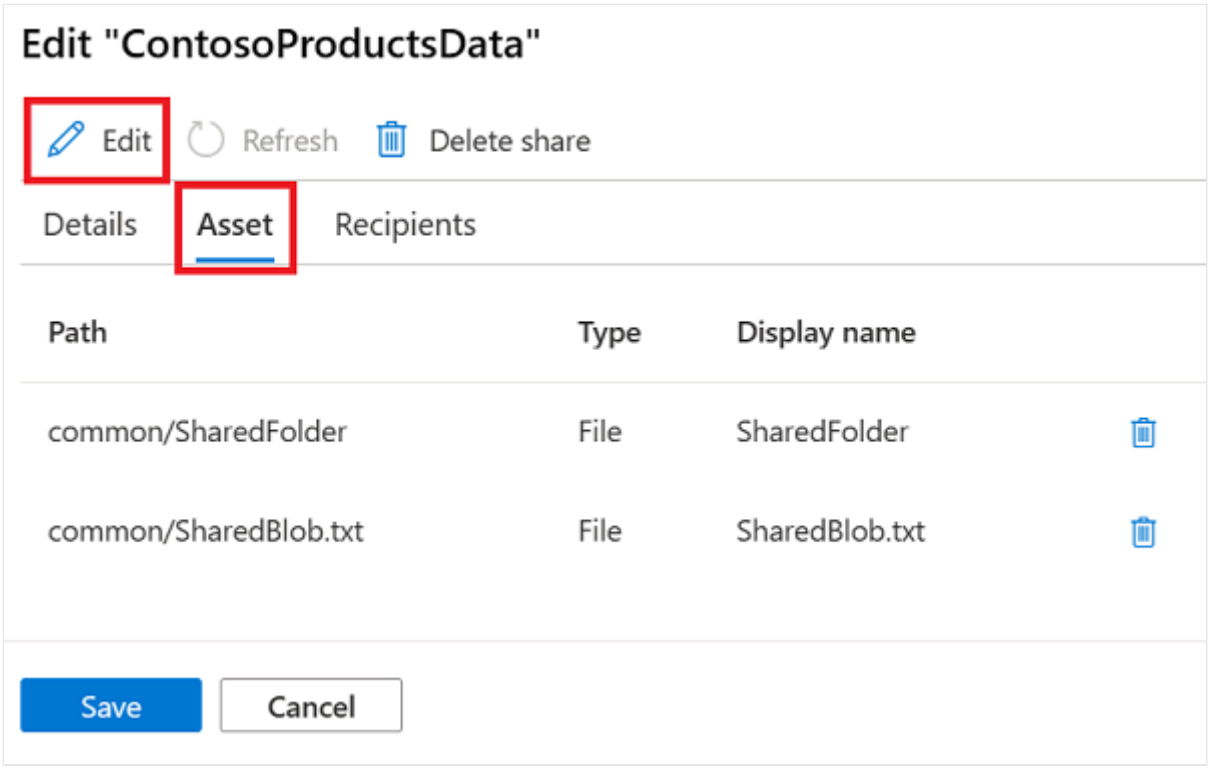
Edit assets

On the **Asset** tab of the edit share page you can see all the shared files and folders.

You can **remove** any containers, files, or folders from the share by selecting the delete button in the asset's row however you can't remove all the assets of a sent share.



You can **add new assets** by selecting the **Edit** button and then searching for and selecting any other containers, files, and folders in the asset that you would like to add.



Once you've selected your assets, select **Add**, and you'll see your new asset in the Asset tab.

Save all your changes by selecting the **Save** button.

Edit recipients

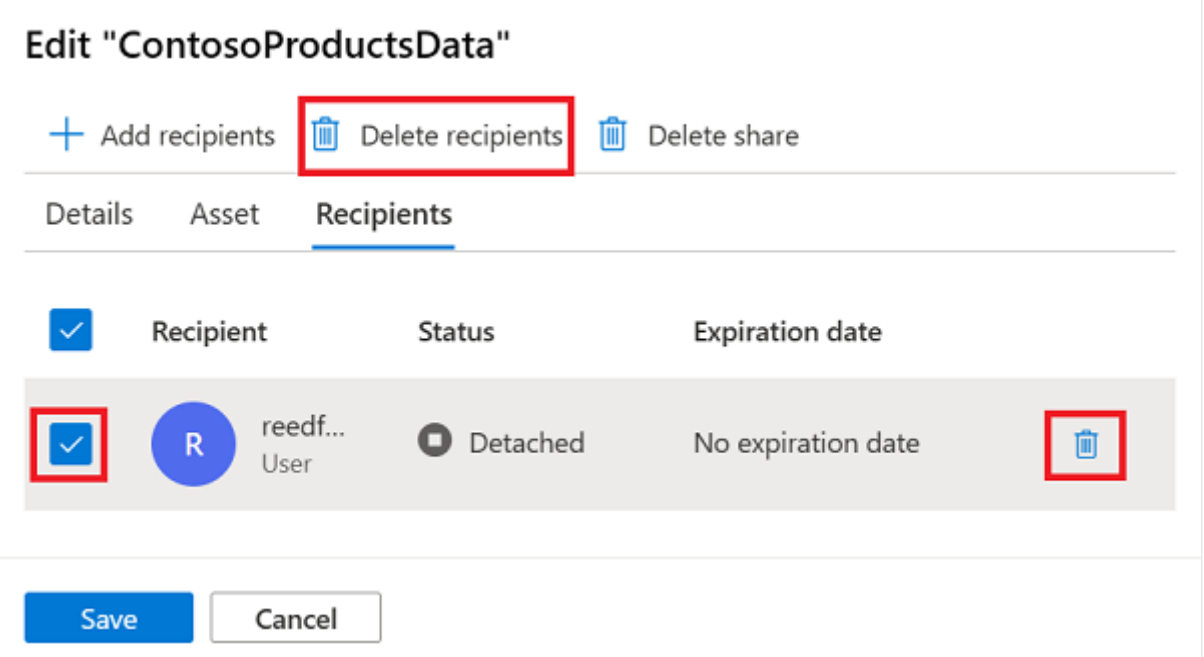
On the **Recipients** tab of the edit share page you can see all the users and groups that are receiving your shares, their status, and the expiration date for their share.

Here are what each of the recipient statuses mean:

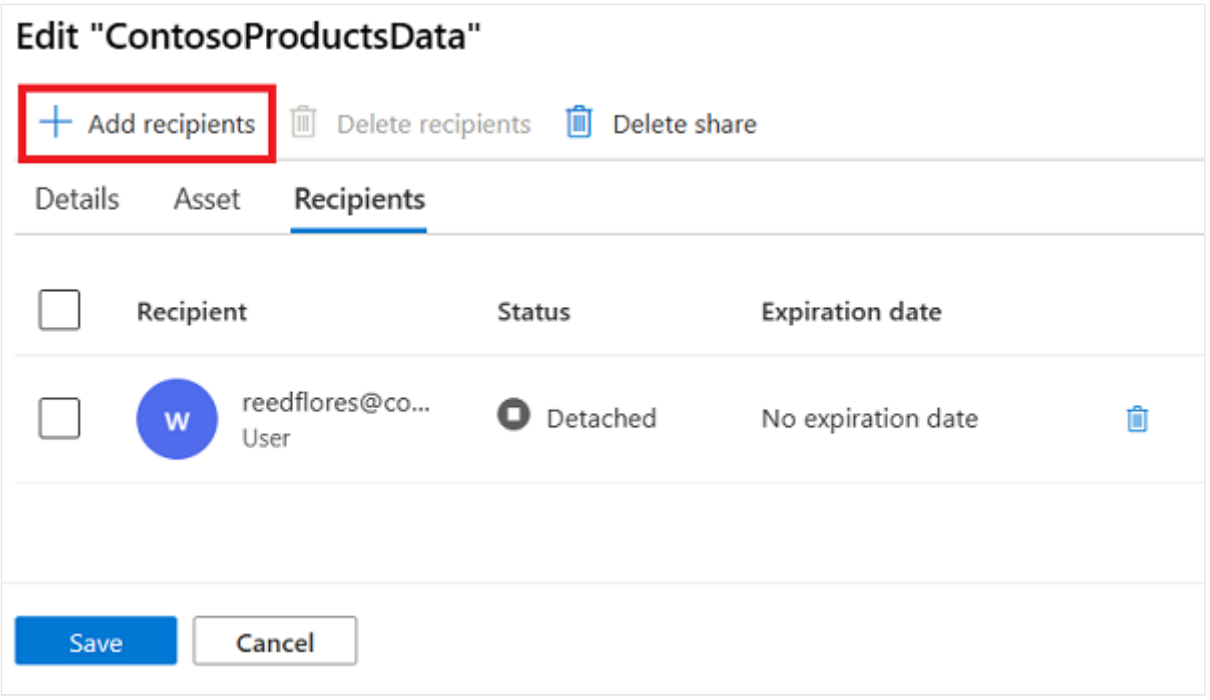
Expand table

Status	Meaning
Attached	The share has been accepted and the recipient has access to the shared data.
Detached	The recipient hasn't accepted the invitation or is no longer active. They aren't receiving the share.

You can **remove or delete recipients** by either selecting the delete button on the recipient's row, or selecting multiple recipients and then selecting the **Delete recipients** button at the top of the page.

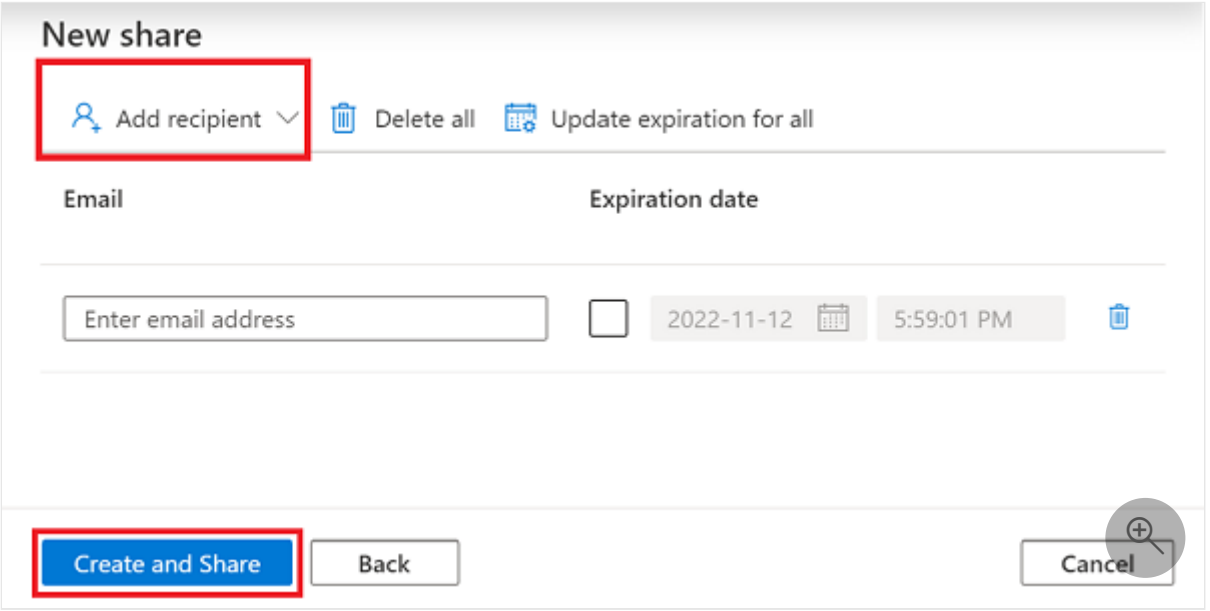


You can **add recipients** by selecting the **Add recipients** button.

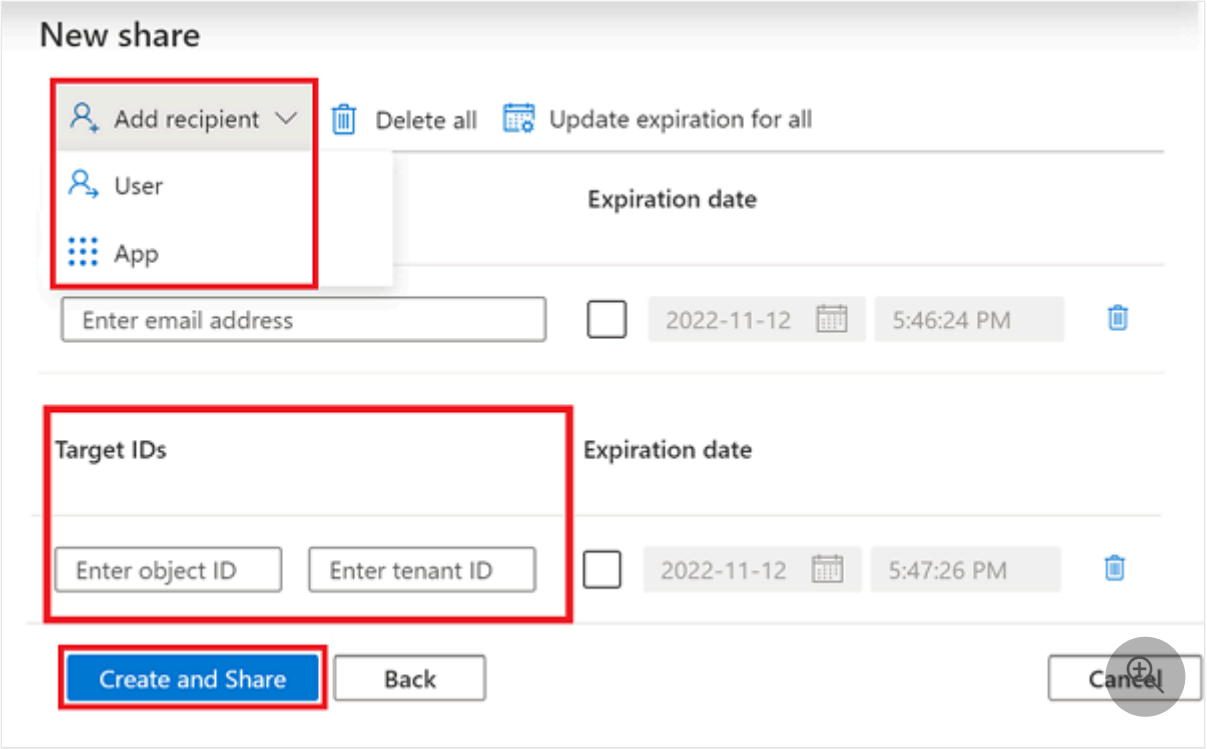


Select **Add Recipient** again and select **User** or **App**.

To share data to a user, select **User**, then enter the Azure sign-in email address of who you want to share data with. By default, the option to enter email address of user is shown.



To to share data with a service principal, select **App**. Enter the object ID and tenant ID of the recipient you want to share data with.



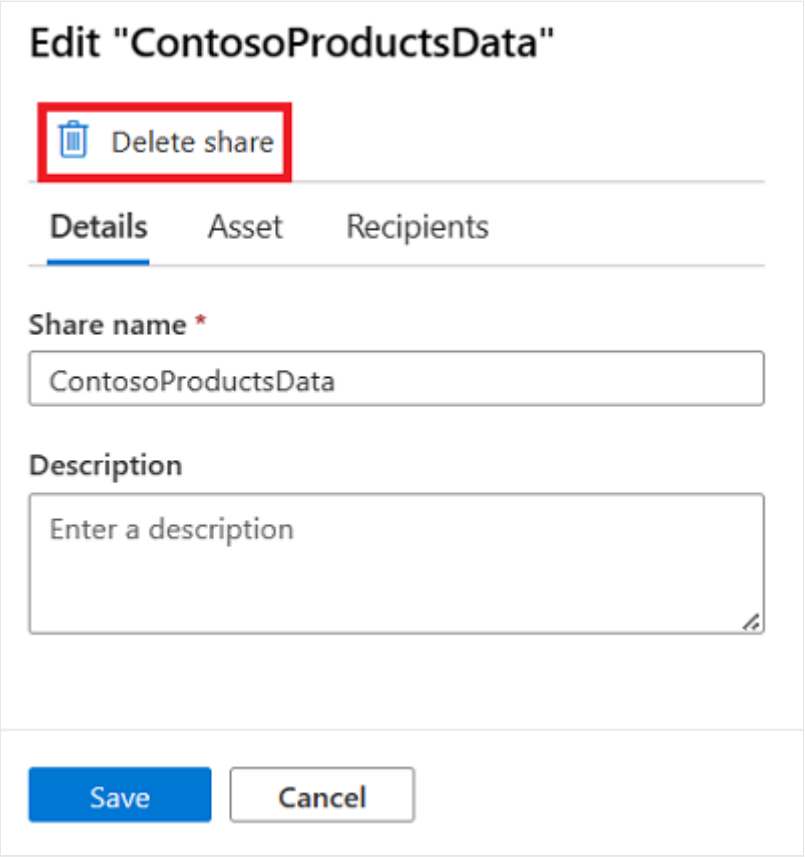
Optionally, you can specify an **Expiration date** for when to terminate the share. You can share the same data with multiple recipients by clicking on **Add Recipient** multiple times.

When you're finished, select the **Add recipients** confirmation button at the bottom of the page.

Save all your changes by selecting the **Save** button.

Delete share

To delete your share, on any tab in the edit share page, select the **Delete share** button.



Confirm that you would like to delete in the pop-up window and the share will be removed.

Troubleshoot

Here are some common issues for sharing data and how to troubleshoot.

Can't create Microsoft Purview account

If you're getting an error related to *quota* when creating a Microsoft Purview account, it means your organization has exceeded Microsoft Purview service limit. If you require an increase in limit, contact support.

Can't find my Storage account asset in the Catalog

There are a couple possible reasons:

- The data source isn't registered in Microsoft Purview. Refer to the registration steps for Blob Storage and ADLSGen2 respectively. Performing a scan isn't necessary.
- Data source is registered to a Microsoft Purview collection that you don't have a minimum of Data Reader permission to. Refer to Microsoft Purview catalog permissions and reach out to your collection admin for access.

Can't create shares or edit shares

- You don't have permission to the data store where you want to share data from. Check the prerequisites for required data store permissions.

Can't view list of shares in the storage account asset

- You don't have enough permissions the data store that you want to see shares of. You need a minimum of **Reader** role on the source storage account to see a read-only view of sent shares and received shares. You can find more details on the ADLS Gen2 or Blob storage data source page.
- Review storage account prerequisites and make sure your storage account region, performance, and redundancy options are all supported.

Next steps

- How to receive share
- FAQ for data sharing
- REST API reference

Feedback

Was this page helpful? Yes No

Receive Azure Storage in-place share with Microsoft Purview Data Sharing (preview)

Article • 10/28/2024

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

> **Important**
>
> This feature is currently being retired, with support provided until September 2025 to allow for migration to Microsoft Fabric external data sharing.

Microsoft Purview Data Sharing supports in-place data sharing from Azure Data Lake Storage (ADLS Gen2) to ADLS Gen2, and Blob storage account to Blob storage account. This article explains how to receive a share and access shared data.

Prerequisites to receive shared data

Microsoft Purview prerequisites

- A Microsoft Purview account.
- A minimum of Data Reader role on a Microsoft Purview collection to use data sharing user experience in the Microsoft Purview governance portal. Refer to Microsoft Purview permissions to learn more about the Microsoft Purview collection and roles.
- To use the SDK, no Microsoft Purview permission is needed.

Azure Storage account prerequisites

- Your Azure subscription must be registered for the **AllowDataSharing** preview feature. Follow the below steps using Azure portal or PowerShell.

Portal

1. In the Azure portal, select the Azure subscription that you use to create the source and target storage account.
2. From the left menu, select **Preview features** under *Settings*.
3. Select **AllowDataSharing** and *Register*.
4. Refresh the *Preview features* screen to verify the *State* is **Registered**. It could take 15 minutes to 1 hour for registration to complete.
5. In addition, to use data share for storage accounts in East US, East US2, North Europe, South central US, West Central US, West Europe, West US, West US2, West US3: Select AllowDataSharingInHeroRegion and Register

For more information, see the register preview feature article.

The *RegistrationState* should be **Registered**. It could take 15 minutes to 1 hour for registration to complete. For more information, see the register preview feature article.

> **Note**
>
> The following are supported storage account configurations:
> - Azure regions: Canada Central, Canada East, UK South, UK West, Australia East, Japan East, Korea South, and South Africa North
> - Additional Azure Regions: East US, East US2, North Europe, Southcentral US, West Central US, West Europe, West US, West US2, West US3
> - Performance: Standard
> - Redundancy options: LRS

- A target storage account **created after** the registration step is completed. **The target storage account must be in the same Azure region as the source storage account.** If you don't know the Azure region of the source storage account, you can find out during the share attaching step later in the process. Target storage account can be in a different Azure region from your Microsoft Purview account.

> **Important**
>
> The target storage account must be in the same Azure region as the source storage account.

- You need the **Contributor** or **Owner** or **Storage Blob Data Owner** or **Storage Blob Data Contributor** role on the target storage account. You can find more details on the ADLS Gen2 or Blob storage data source pages.

- If the target storage account is in a different Azure subscription than the one for Microsoft Purview account, the `Microsoft.Purview` resource provider needs to be registered in the Azure subscription where the Storage account is located. It's automatically registered at the time of share consumer attaching the share and if the user has permission to do the `/register/action` operation and therefore, Contributor or Owner roles to the subscription where the Storage account is located. This registration is only needed the first time when sharing or receiving data into a storage account in the Azure subscription.

- A storage account needs to be registered in the collection where you receive the share. For instructions to register, see the ADLS Gen2 or Blob storage data source pages. This step isn't required to use the SDK.

- Latest version of the storage SDK, PowerShell, CLI, and Azure Storage Explorer. Storage REST API version must be February 2020 or later.

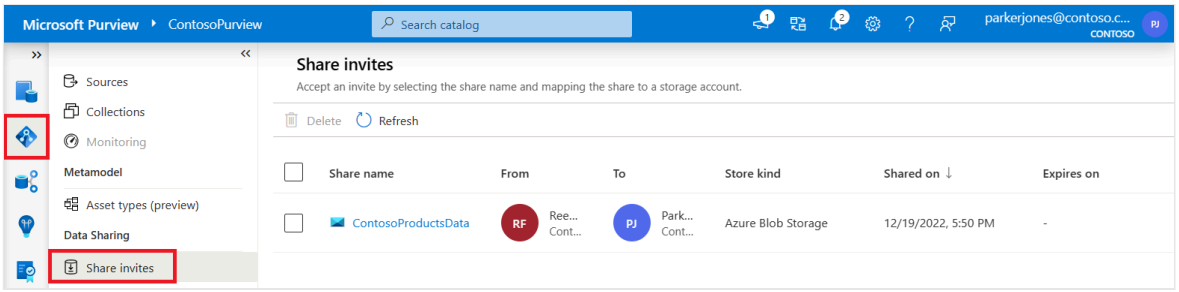
Receive share

1. You can view your share invitations in any Microsoft Purview account. Open the Microsoft Purview governance portal by browsing directly to the classic portal or the new Microsoft Purview portal and selecting your Microsoft Purview account.

2. In the classic portal, select the **Data Map** icon from the left navigation. In the new portal, navigate to the **Data Catalog** application.

3. Select **Share invites**. If you received an email invitation, you can also select the **View share invite** link in the email to select a Microsoft Purview account.

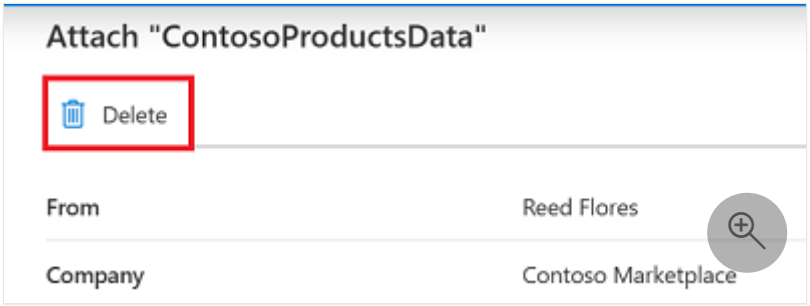
   If you're a guest user of a tenant, you are asked to verify your email address for the tenant before viewing share invitation for the first time. You can see our guide below for steps. Once verified, it's valid for 12 months.



4. Alternately, within the Microsoft Purview portal, find the Azure Storage or Azure Data Lake Storage (ADLS) Gen 2 data asset you would like to receive the share into using either the data catalog search or browse. Select the **Data Share** button. You can see all the invitations in the **Share invites** tab.

5. Select name of the share invite to view or configure.

6. If you don't want to accept the invitation, select **Delete**.



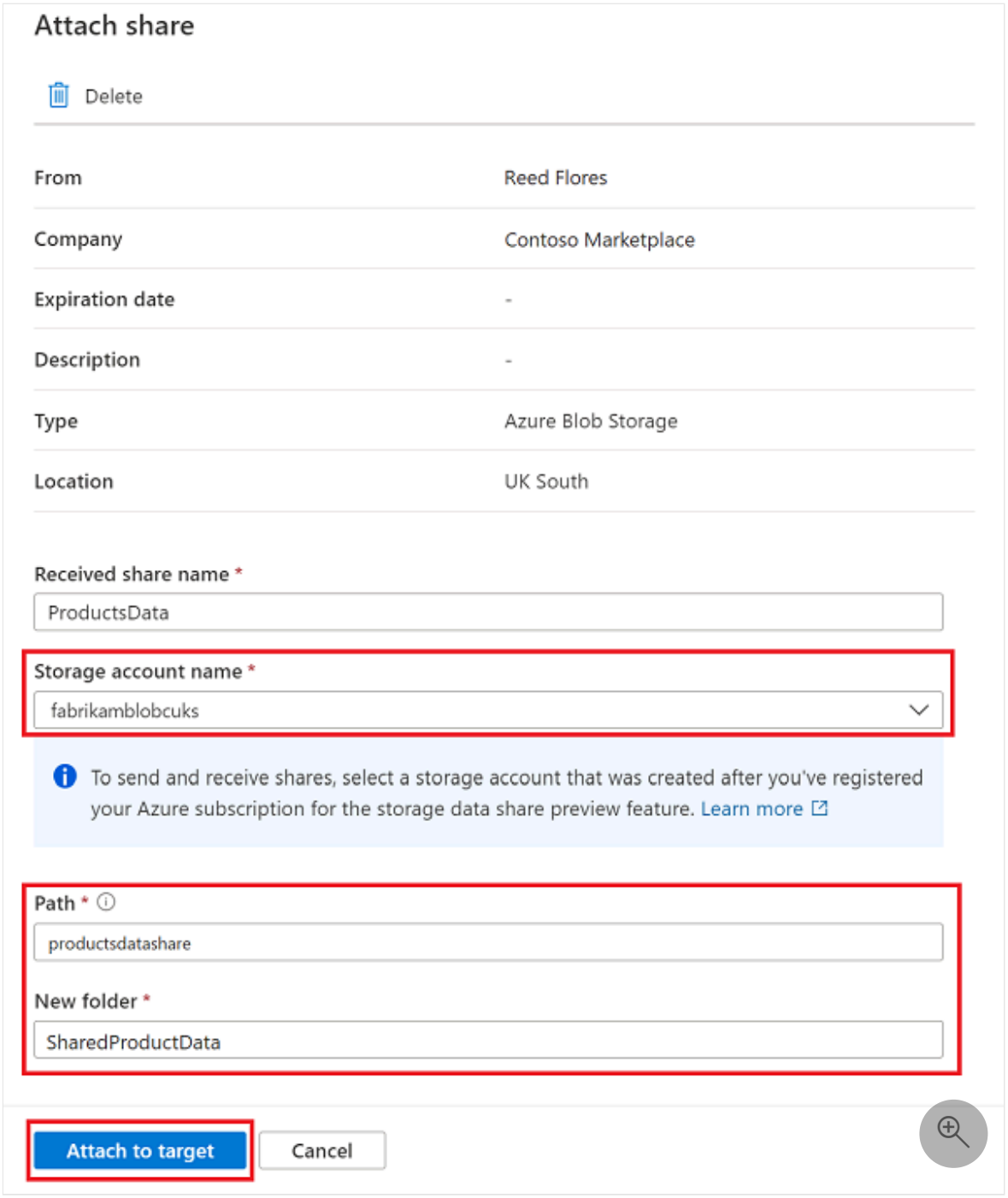
> **Note**
>
> If you delete an invitation, if you want to accept the share in future it needs to be resent. To deselect the share without deleting select the **Cancel** button instead.

7. You can edit the **Received share name** if you like. Then select a **Storage account name** for a target storage account in the same region as the source. You can choose to **Register a new storage account to attach the share** in the drop-down as well.

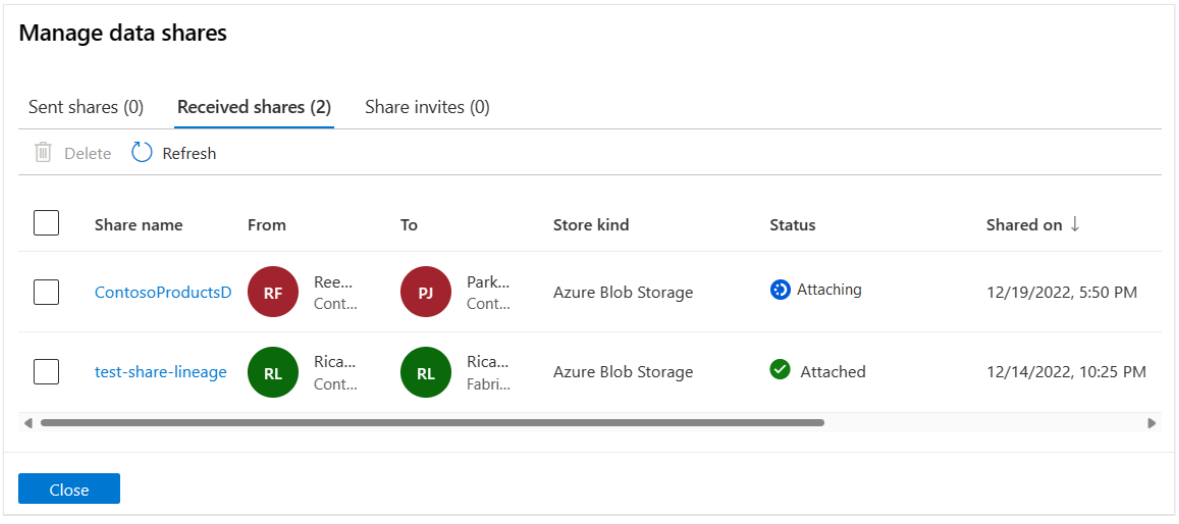
> **Important**
>
> The target storage account needs to be in the same Azure region as the source storage account.

8. Configure the **Path** (either a new container name, or the name of an existing share container) and, **New folder** (a new folder name for the share within in your container).

9. Select **Attach to target**.



10. On the Manage data shares page, you see the new share with the status of **Attaching** until it has completed and is attached.



11. You can access shared data from the target storage account through Azure portal, Azure Storage Explorer, Azure Storage SDK, PowerShell, or CLI. You can also analyze the shared data by connecting your storage account to Azure Synapse Analytics Spark or Databricks.

When a share is attached, a new asset of type received share is ingested into the Microsoft Purview catalog, in the same collection as the storage account to which you attached the share is registered. You can search or browse for it like any other asset.

Refer to Microsoft Purview Data Sharing lineage to learn more about share assets and data sharing lineage.

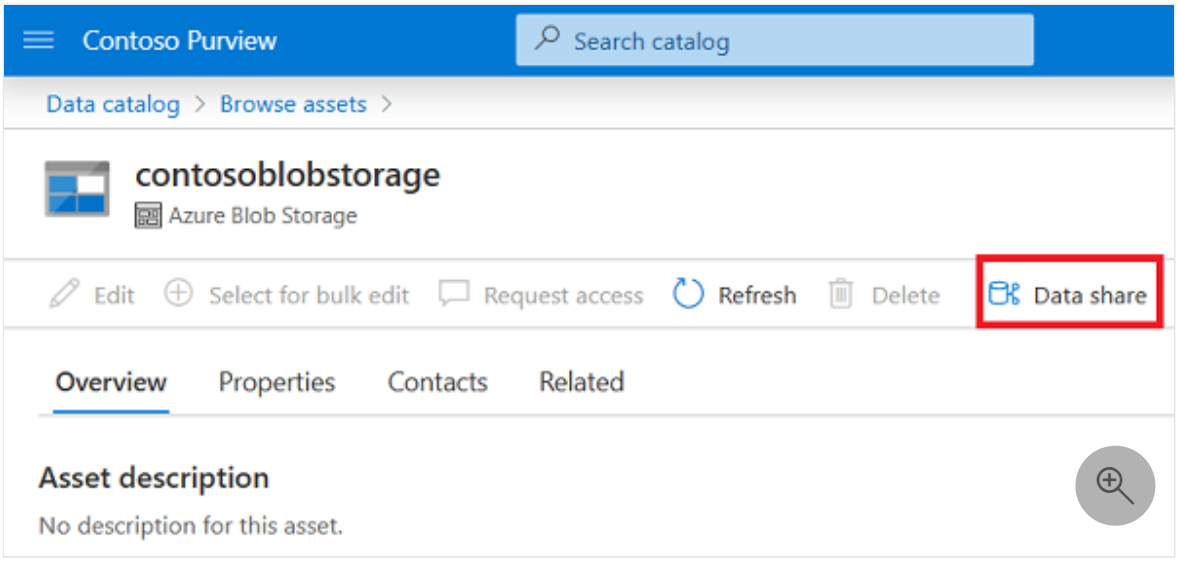
> ⓘ **Note**
>
> Shares attached using the SDK without registering the storage account with Microsoft Purview won't be ingested into the catalog. User can register their storage account if desired. If a storage account is unregistered or re-registered to a different collection, share assets of that storage account continue to be in the initial collection.

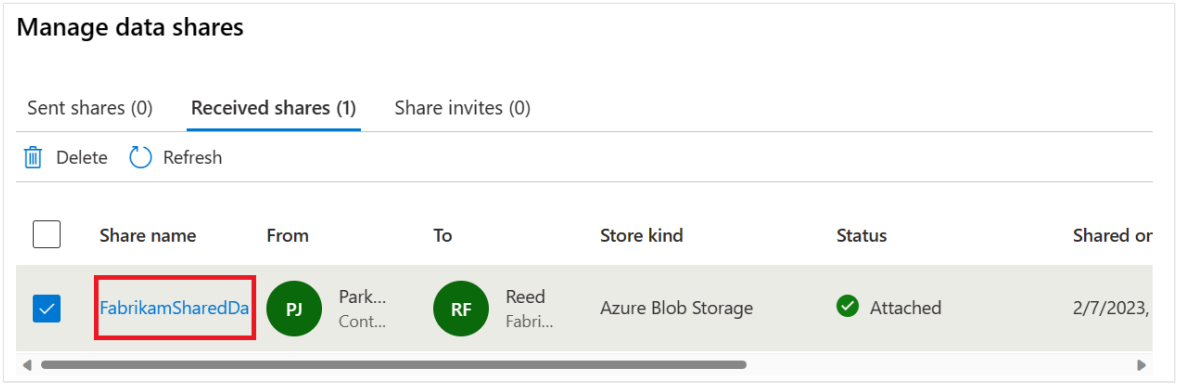
Update received share

Once you attach a share, you can edit the received share by reattaching the share to a new storage account, or stop the sharing relationship by deleting the received share.

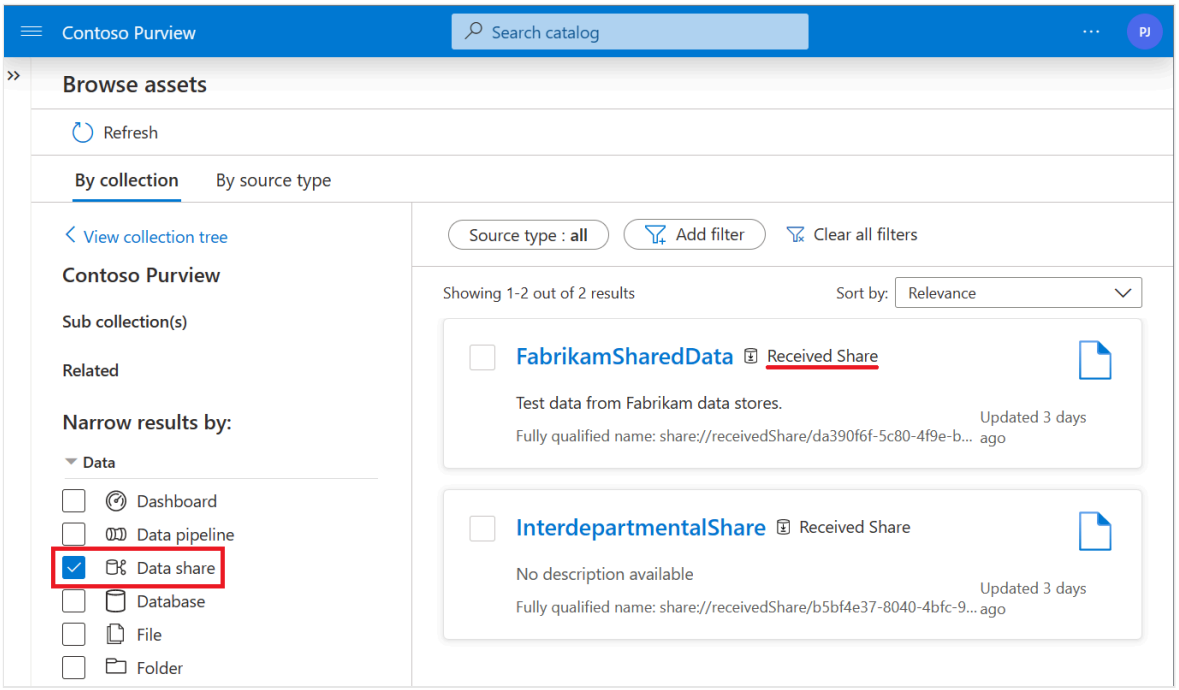
You can find and edit received share asset one of two ways:

- Access the blob storage or ADLS Gen2 asset where the data was received in the data catalog and open it, then select **Data Share** and **Manage data shares**. There you're able to see all the shares for that asset. Select the **Received shares** tab, select a share, and then select your share.



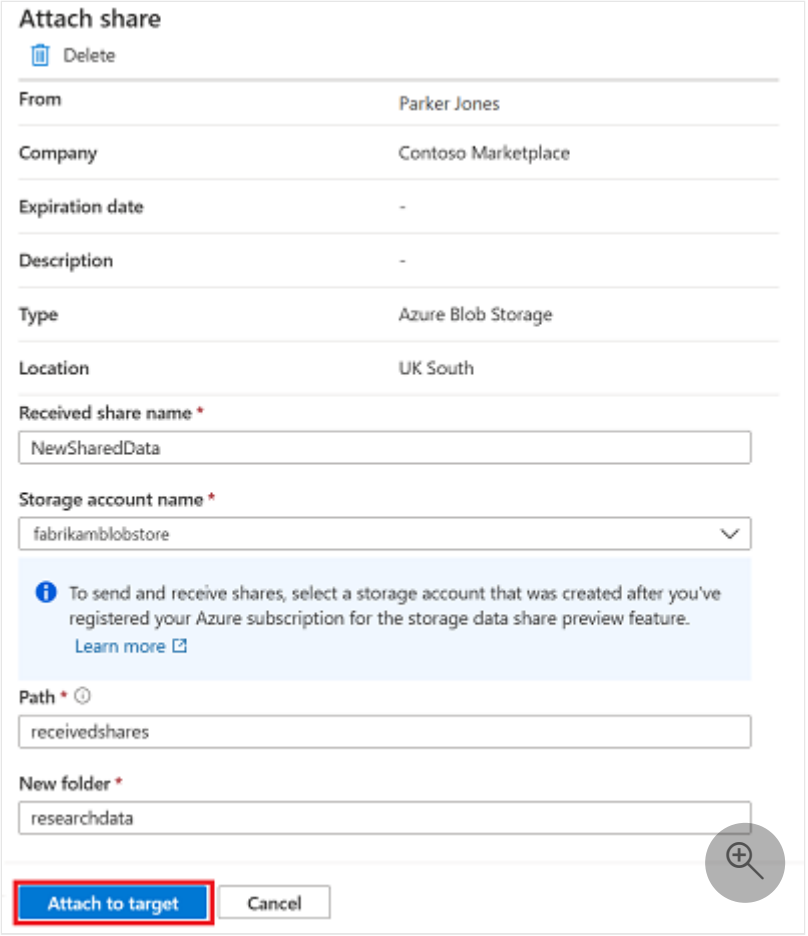


- Search or browse the data catalog for data share assets and select your received share. Then select the **Edit** button.



Reattach share

After you've selected your data share to edit, you can reattach the share to a new storage account or path in your current storage account by selecting a storage account, providing a path, and providing the folder.

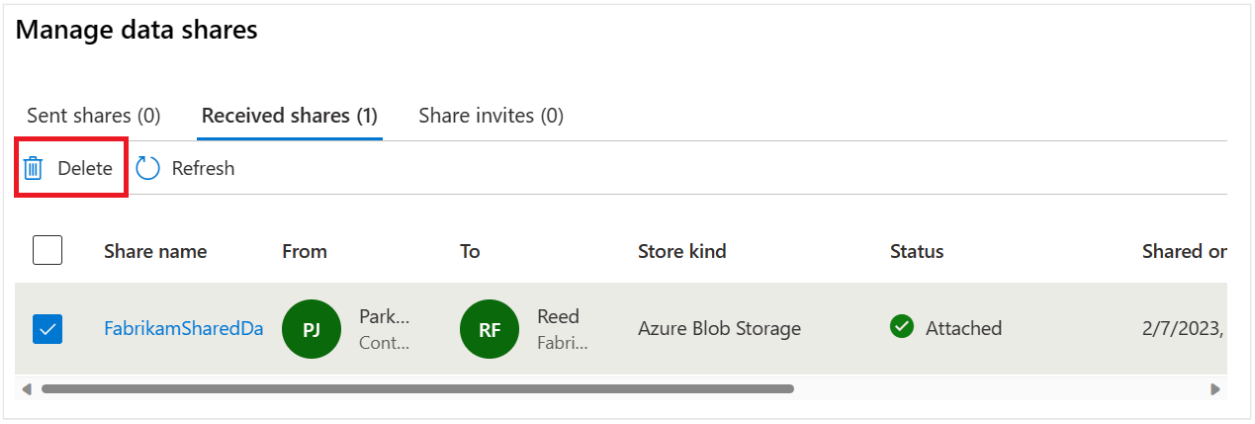


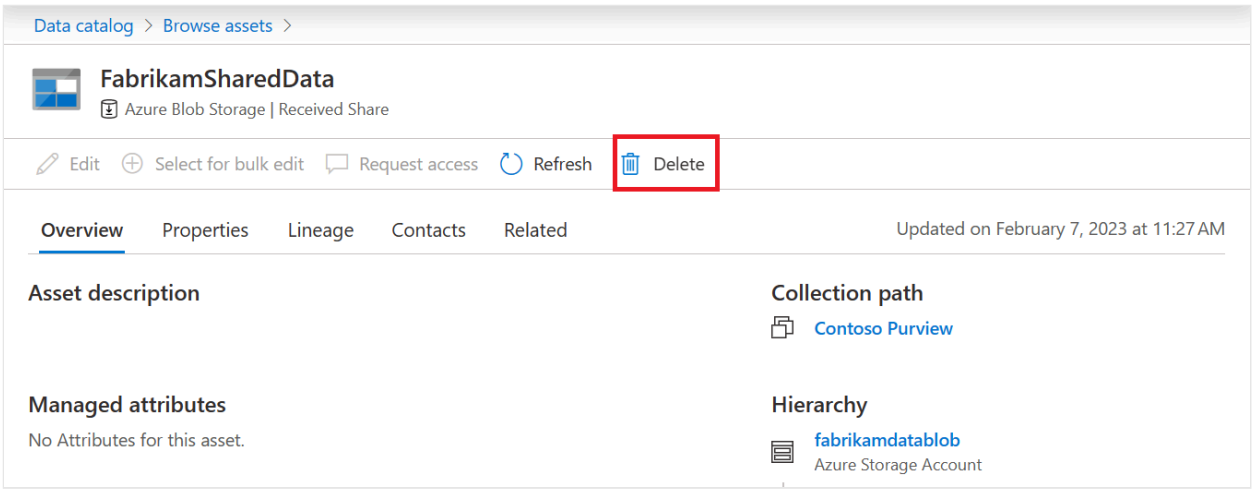
If you're updating the target, select **Attach to target** to save your changes. Attaching can take a couple minutes to complete after the process has been started.

> **Note**
>
> While reattaching, if you selected a storage account that is registered to a collection that you don't have permissions to or a storage account that isn't registered in Microsoft Purview, you'll be shown the appropriate message. You'll see the shared data in your target data store.

Delete received share

To delete a received share, select the share and then select **Delete**.





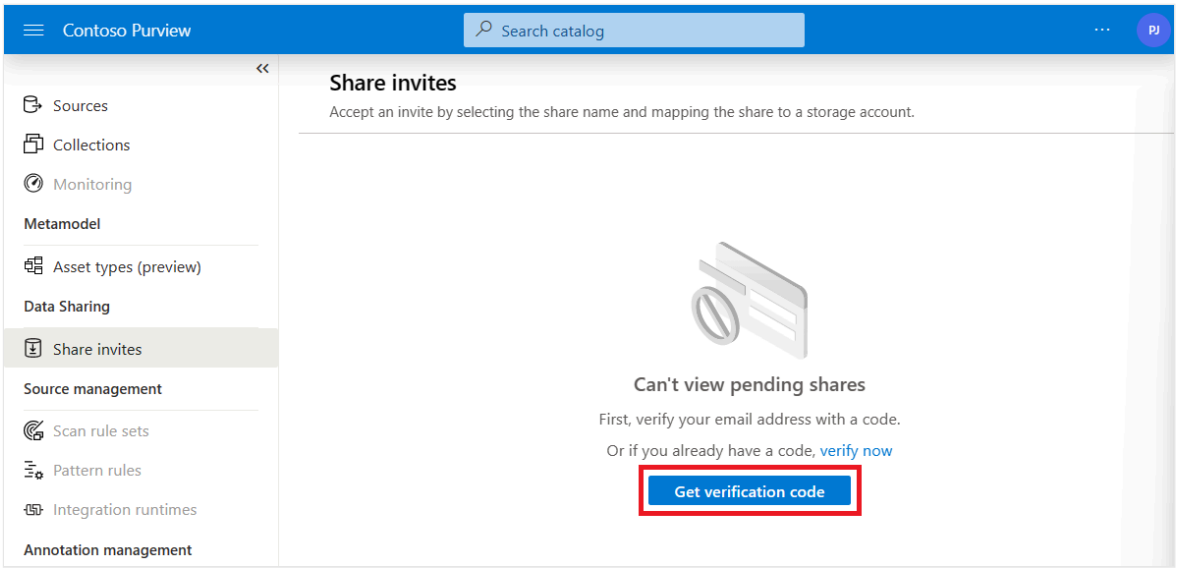
Deleting a received share stops the sharing relationship, and you won't be able to access shared data. Deleting a received share can take a few minutes.

Guest user verification

If you're a guest user in Azure, to be able to receive shares your account must first be associated with the Microsoft Entra ID. You can start this process from within Microsoft Purview.

1. In the Microsoft Purview Governance Portal, select the Data Map and then **Share invites**.

2. You see text indicating that you need to be associated with Microsoft Entra ID. Select the **Get verification code** button.



3. You'll receive an email from Microsoft Azure with the verification code. Copy it.

4. Return to the **Share invites** page in the Microsoft Purview Governance Portal, select **Verify**, and enter the code you received.

Troubleshoot

Here are some common issues for receiving share and how to troubleshoot them.

Can't create Microsoft Purview account

If you're getting an error related to *quota* when creating a Microsoft Purview account, it means your organization has exceeded Microsoft Purview service limit. If you require an increase in limit, contact support.

Can't find my Storage account asset in the Catalog

- Data source isn't registered in Microsoft Purview. Refer to the registration steps for Blob Storage and ADLSGen2 respectively. Performing a scan isn't necessary.
- Data source is registered to a Microsoft Purview collection that you don't have a minimum of Data Reader permission to. Refer to Microsoft Purview catalog permissions and reach out to your collection admin for access.

Can't view list of shares in the storage account asset

- Permission issue to the data store that you want to see shares for. You need a minimum of **Reader** role on the source storage account to see a read-only view of

sent shares and received shares. You can find more details on the ADLS Gen2 or Blob storage data source page.

Can't view received share in the storage account asset

- You might have selected a different storage account to attach the share to, that might not be registered in Microsoft Purview or maybe registered to a collection you don't have permissions to. Refer to the registration steps for Blob Storage and ADLSGen2 respectively. Refer to Microsoft Purview catalog permissions and reach out to your collection admin for access to collections.

Can't view share invite

If you've been notified that you've received a share, but can't view share invite in your Microsoft Purview account, it could be due to the following reasons:

- You don't have a minimum of **Data Reader** role to any collections in this Microsoft Purview account. Contact your *Microsoft Purview Collection Admin* to grant you access to **Data Reader** role to view, accept, and configure the received share.
- Share invitation is sent to your email alias or an email distribution group instead of your Azure sign-in email. Contact your data provider and ensure that they've sent the invitation to your Azure sign-in e-mail address.
- Share has already been accepted. If you already accepted the share, it will no longer show up in *Share invites* tab. Select *Received shares* in any Storage account that you have permissions to and see your active shares.
- You're a guest user of the tenant. If you're a guest user of a tenant, you need to verify your email address for the tenant in order to view a share invitation for the first time. Once verified, it's valid for 12 months.

Can't see target storage account in the list when attaching a share

When you attach a share to a target, if your storage account isn't listed for you to select, it's likely due to the following reasons:

- You don't have required permissions to the storage account. Check prerequisites for details on required storage permissions.

Failed to attach share

If you failed to attach a share, it's likely due to the following reasons:

- Permission issue to the target data store. Check prerequisites for required data store permissions.
- The storage account isn't supported. Microsoft Purview Data share only supports storage accounts with specific configurations.
- The *Path* you specified includes container created outside of Microsoft Purview Data Sharing. You can only receive data into containers created while attaching.
- The *New Folder* you specified to receive storage data isn't empty.
- Source and target storage account is the same. Sharing from the same source storage account to the same target isn't supported.
- Exceeding limit. Source storage account can support up to 20 targets, and target storage account can support up to 100 sources. If you require an increase in limit, contact support.

Can't access shared data in the target data store

If you can't access shared data, it's likely due to the following reasons:

- After share attaching is successful, it could take some time for the data to appear in the target data store. Try again in a few minutes. Likewise, after you delete a share, it can take a few minutes for the data to disappear in the target data store.
- You're accessing shared data using a storage API version prior to February 2020. Only storage API version February 2020 and later are supported for accessing shared data. Ensure you're using the latest version of the storage SDK, PowerShell, CLI, and Azure Storage Explorer.
- You're accessing shared data using an analytics tool that uses a storage API version February 2020. You can access shared data from Azure Synapse Analytics Spark and Databricks. You won't be able to access shared data usingprior to Azure Data Factory, Power BI, or AzCopy.

Next steps

- How to share data
- FAQ for data sharing
- REST API reference

Feedback

Was this page helpful? Yes No

Quickstart: Share and receive Azure Storage data in-place with Microsoft Purview Data Sharing (preview)

> **Note**
>
> Microsoft Purview Data Catalog (classic), Data Health Insights (classic), and Purview Workflow (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

> **Important**
>
> This feature is currently being retired, with support provided until September 2025 to allow for migration to Microsoft Fabric external data sharing.

This article provides a quick guide on how to share data and receive shares from Azure Data Lake Storage (ADLS Gen2) or Blob storage accounts.

> **Important**
>
> Data sharing is accessed through the classic Microsoft Purview governance portal: https://web.purview.azure.com

Prerequisites

Microsoft Purview prerequisites

- A Microsoft Purview account. You can also use two Microsoft Purview accounts, one for data provider and one for data consumer to test both scenarios.
- Your recipient's Azure sign-in email address that you can use to send the invitation to. The recipient's email alias won't work.

Azure Storage account prerequisites

- Your Azure subscription must be registered for the **AllowDataSharing** preview feature. (If you haven't already registered this preview feature, consider using Microsoft Fabric external data sharing instead, since Microsoft Purview Data Sharing will be retired in September 2025.) Or, contact support.
- Source and target storage accounts **created after** the registration step is completed. **Both storage accounts must be in the same Azure region as each other**. Both storage accounts need to be ADLS Gen2 or Blob Storage accounts. Your storage accounts can be in a different Azure region from your Microsoft Purview account.
- Latest version of the storage SDK, PowerShell, CLI, and Azure Storage Explorer. Storage REST API version must be February 2020 or later.
- The storage accounts need to be registered in the collections where you'll send or receive the share. If you're using one Microsoft Purview account, this can be two different collections, or the same collection. For instructions to register, see the ADLS Gen2 or Blob storage data source pages.
- If the source or target storage accounts are in a different Azure subscription than the one for Microsoft Purview account, the Microsoft. Purview resource provider is automatically registered in the Azure subscription where the data store is located at the time of share provider adding an asset or share consumer mapping an asset and **ONLY** if the user has permission to do the /register/action operation for the resource provider. The permission is included in the Contributor and Owner roles.

> **Note**
>
> This registration is only needed the first time when sharing or receiving data into a storage account in the Azure subscription.

Required roles

Here are required roles for sharing data and receiving shares.

Expand table

	Azure Storage account roles	Microsoft Purview collection roles
Data provider	One of the following roles: • Owner • Storage Blob Data Owner	Data Reader
Data consumer	One of the following roles:	Data Reader

	Azure Storage account roles	Microsoft Purview collection roles
	• Contributor • Owner • Storage Blob Data Contributor • Storage Blob Data Owner	

> **Note**
>
> If you created the Microsoft Purview account, you're automatically assigned all the roles to the root collection. Refer to Microsoft Purview permissions to learn more about the Microsoft Purview collection and roles.

Create a share

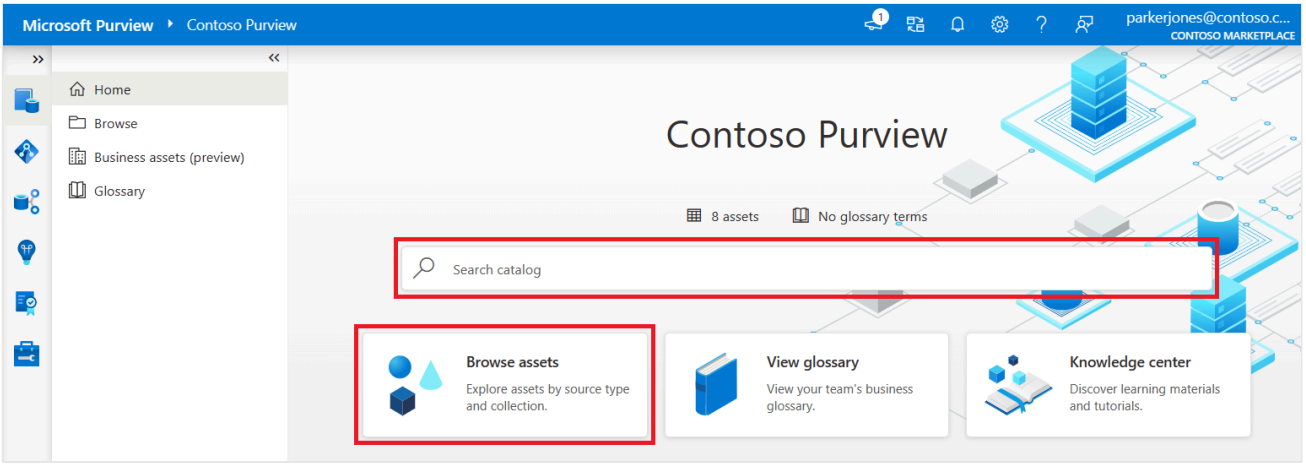
There are two ways you can create a data share:

- Directly from the data asset
- Use the data sharing application

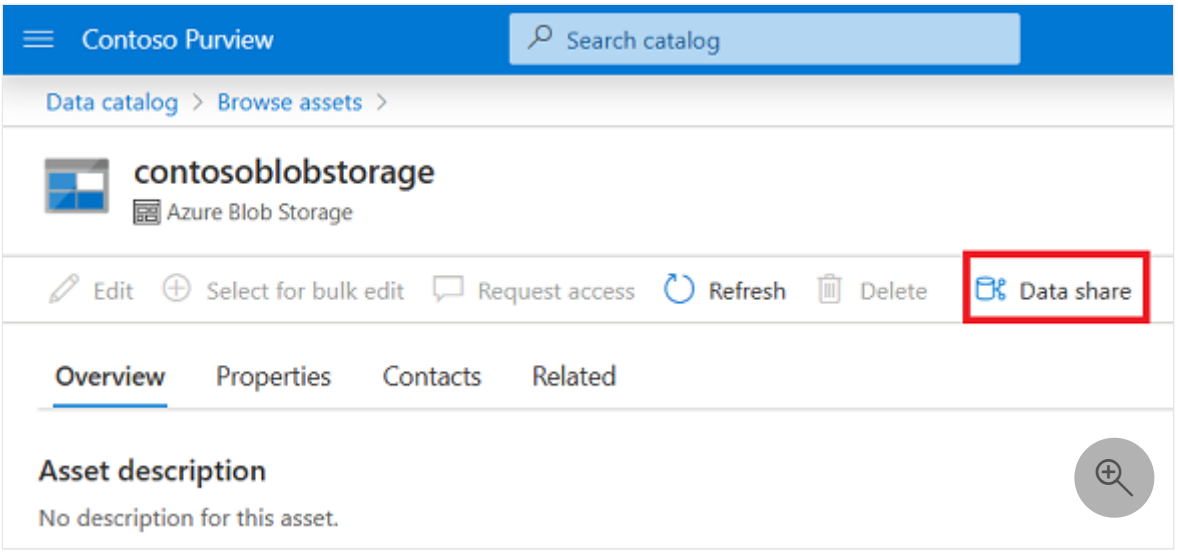
Create share from asset

1. You can create a share by starting from **Data Catalog**

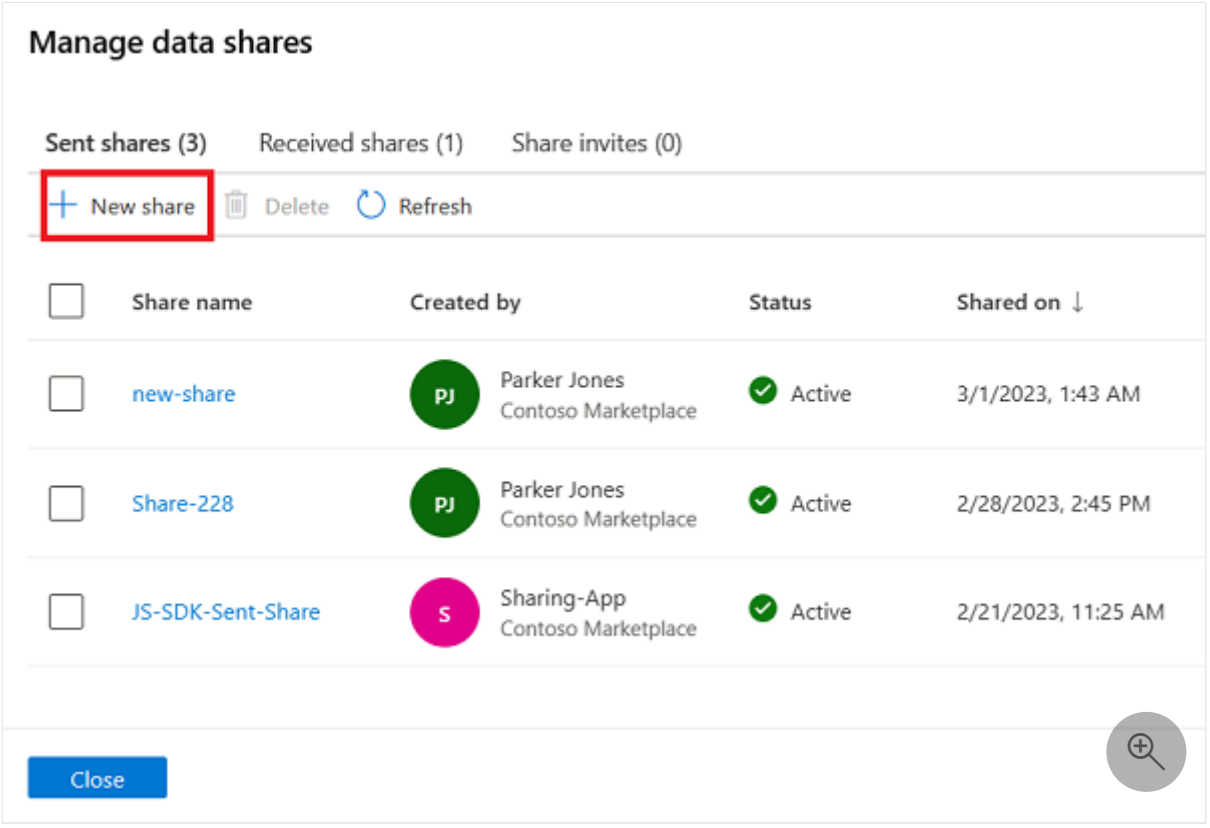
   Within the classic Microsoft Purview governance portal or the new Microsoft Purview portal, find the Azure Storage or Azure Data Lake Storage (ADLS) Gen 2 data asset you would like to share data from using either the data catalog search or browse.



2. Once you have found your data asset, select the **Data Share** button.



3. Select **+New Share**.



4. Follow the rest of the steps to create your data share.

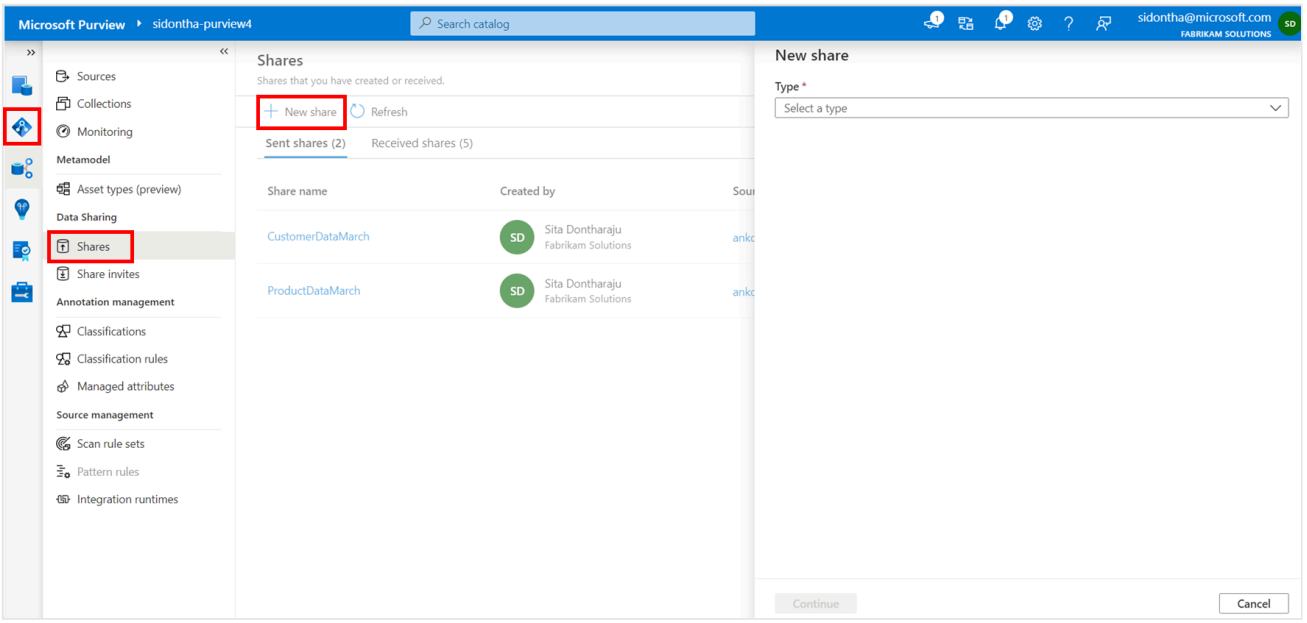
Create share from application

1. If you're using the new Microsoft Purview experience You can create a share by starting from the **Data Catalog**.

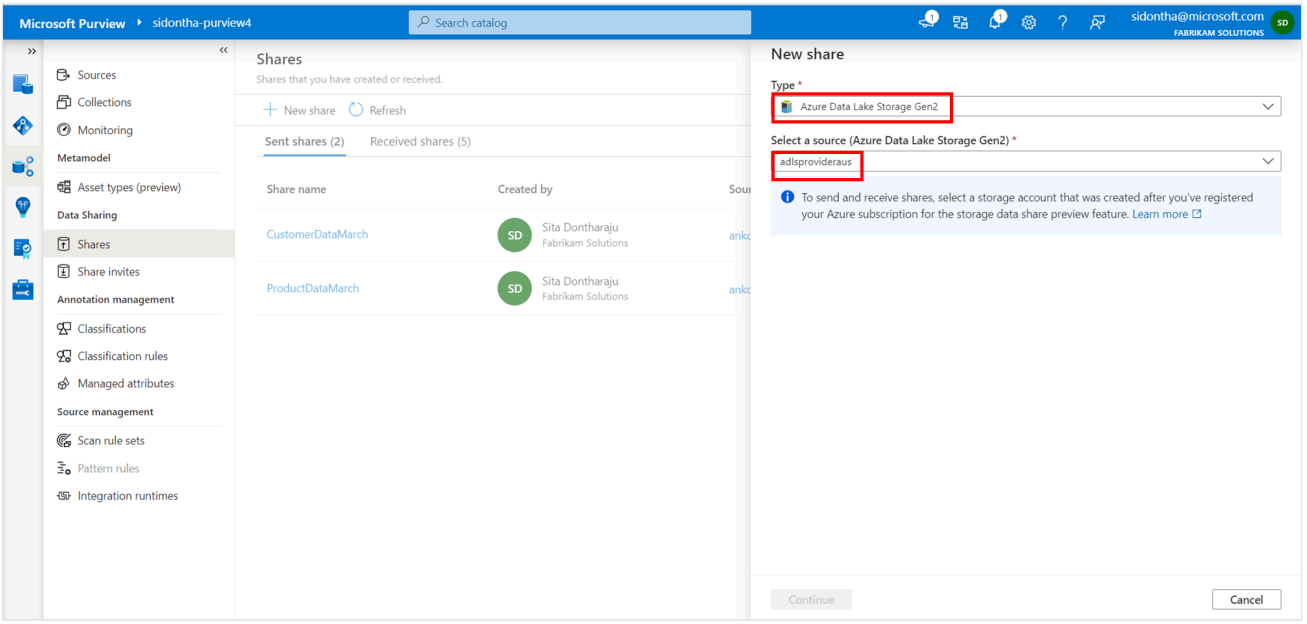
   Open the Microsoft Purview portal. Navigate to the **Data Catalog** application. Then select **Shares** from the left navigation. Select **+New Share**.

2. If you're using the classing Microsoft Purview experience You can create a share by starting from the **Data Map**.

   Open the Microsoft Purview governance portal. Select the **Data Map** icon from the left navigation. Then select **Shares**. Select **+New Share**.



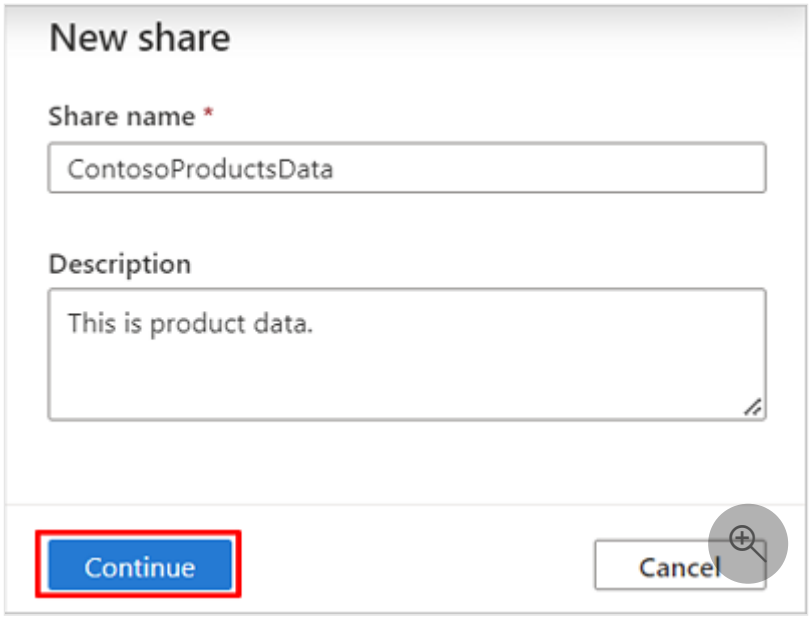
3. From either location, select the Storage account type and the Storage account you want to share data from. Then select **Continue**.



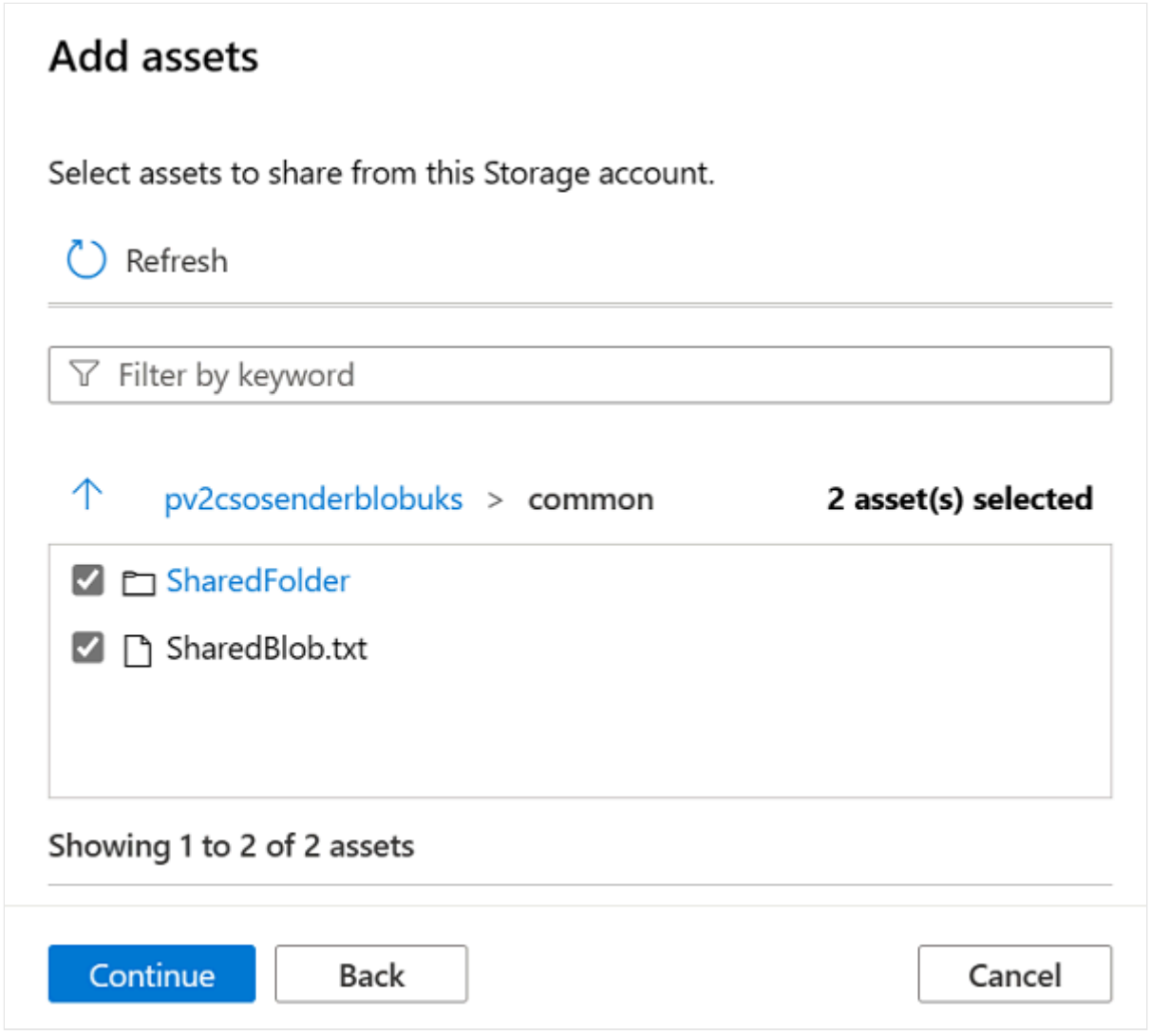
4. Follow the rest of the steps to create your data share.

Create share

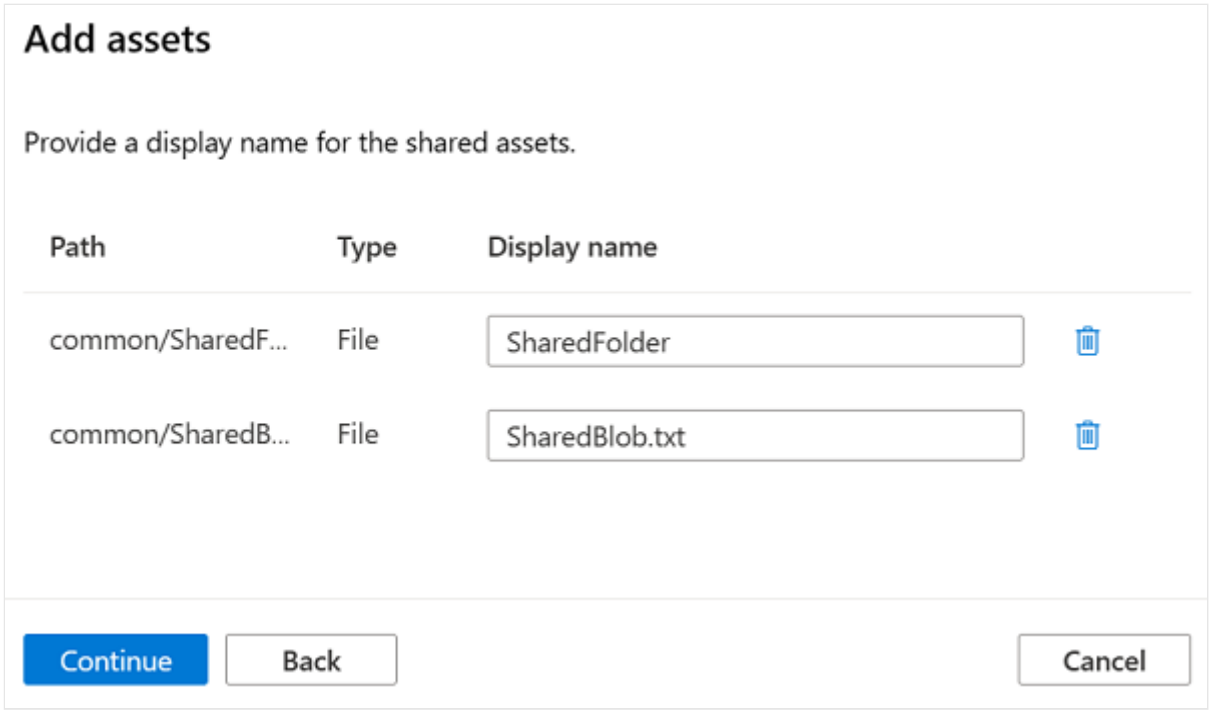
1. Specify a name and a description of share contents (optional). Then select **Continue**.



2. Search for and add all the assets you'd like to share out at the container, folder, or file level, and then select **Continue**.

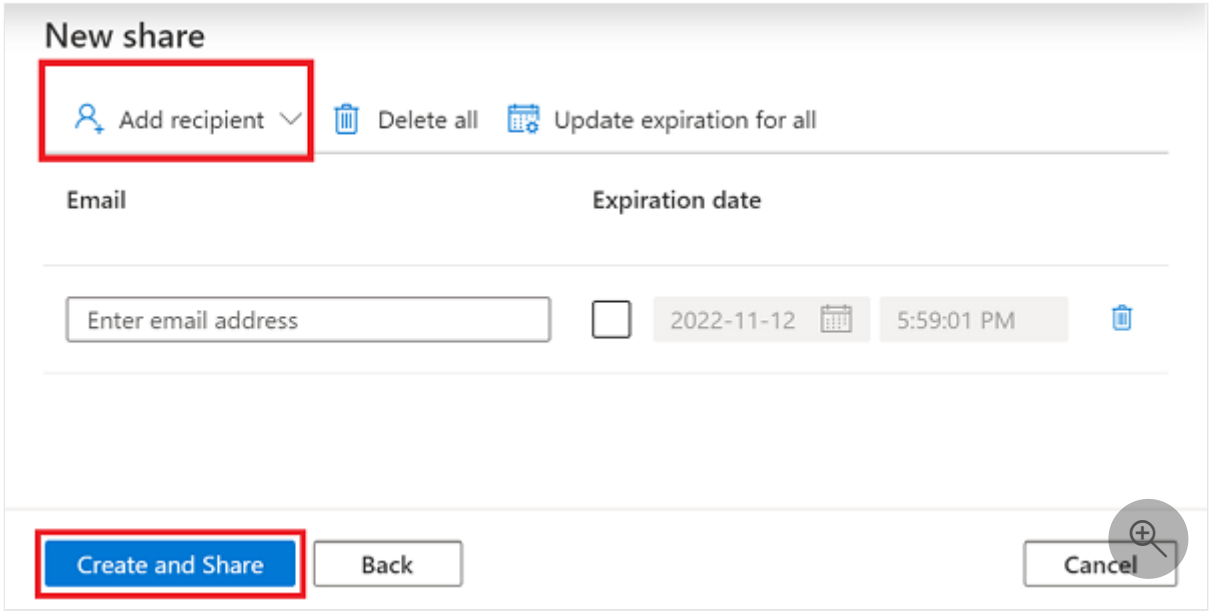


3. You can edit the display names the shared data will have, if you like. Then select **Continue**.

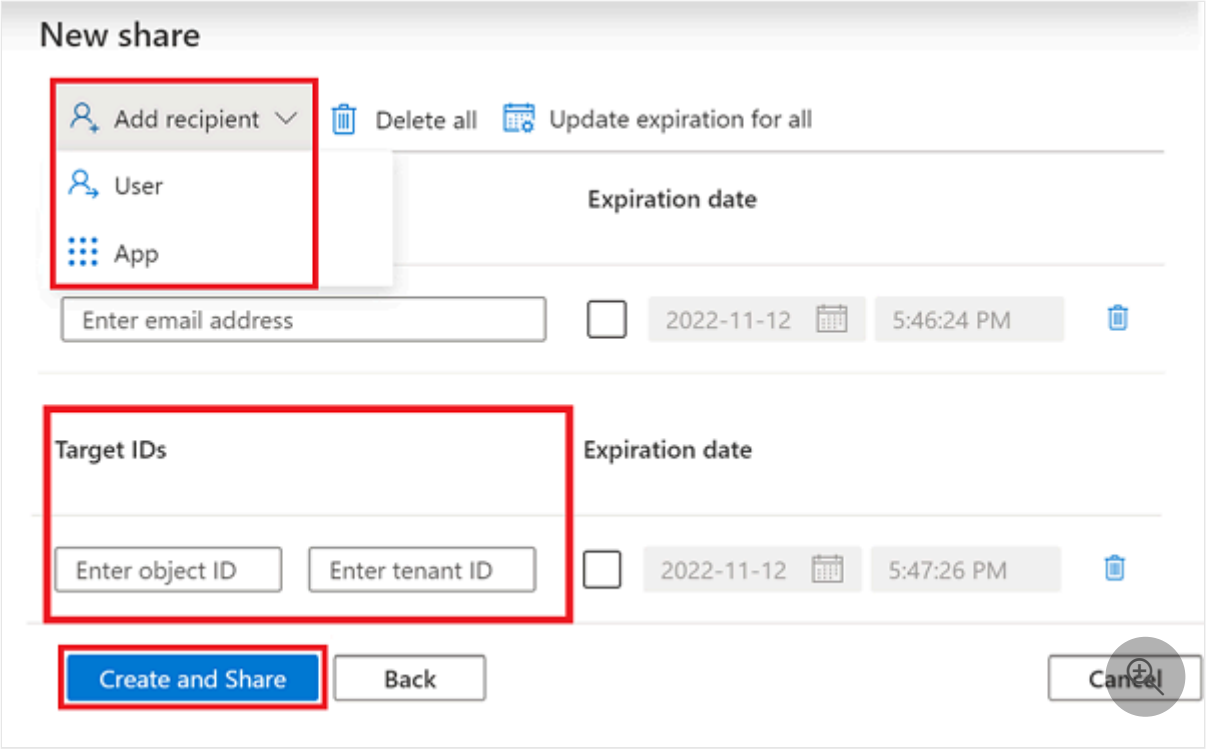


4. Select **Add Recipient** and select **User** or **App**.

   To share data to a user, select **User**, then enter the Azure sign-in email address of who you want to share data with. By default, the option to enter email address of user is shown.



   To to share data with a service principal, select **App**. Enter the object ID and tenant ID of the recipient you want to share data with.



5. Select **Create and Share**. Optionally, you can specify an **Expiration date** for when to terminate the share. You can share the same data with multiple recipients by selecting **Add Recipient** multiple times.

You've now created your share. The recipients of your share will receive an invitation and they can view the pending share in their Microsoft Purview account.

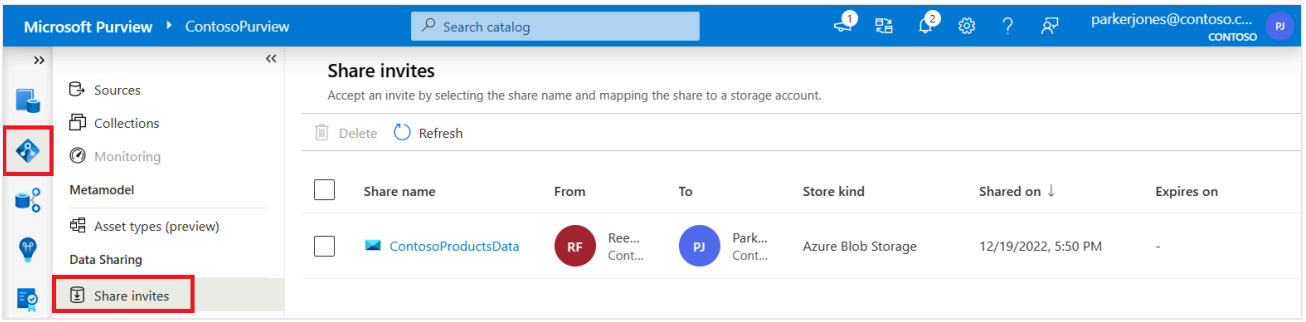
Receive share

1. You can view your share invitations in any Microsoft Purview account. Open the Microsoft Purview governance portal by browsing directly to the classic portal or the new Microsoft Purview portal and selecting your Microsoft Purview account.

2. In the classic portal, select the **Data Map** icon from the left navigation. In the new portal, navigate to the **Data Catalog** application.

3. Select **Share invites**. If you received an email invitation, you can also select the **View share invite** link in the email to select a Microsoft Purview account.

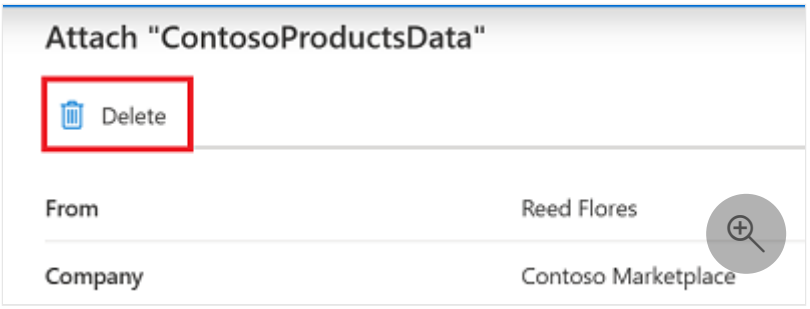
   If you're a guest user of a tenant, you'll be asked to verify your email address for the tenant before viewing share invitation for the first time. You can see our guide for steps. Once verified, it's valid for 12 months.



4. Alternately, within the Microsoft Purview portal, find the Azure Storage or Azure Data Lake Storage (ADLS) Gen 2 data asset you would like to receive the share into using either the data catalog search or browse. Select the **Data Share** button. You can see all the invitations in the **Share invites** tab.

5. Select name of the share to view or configure.

6. If you don't want to accept the invitation, select **Delete**.



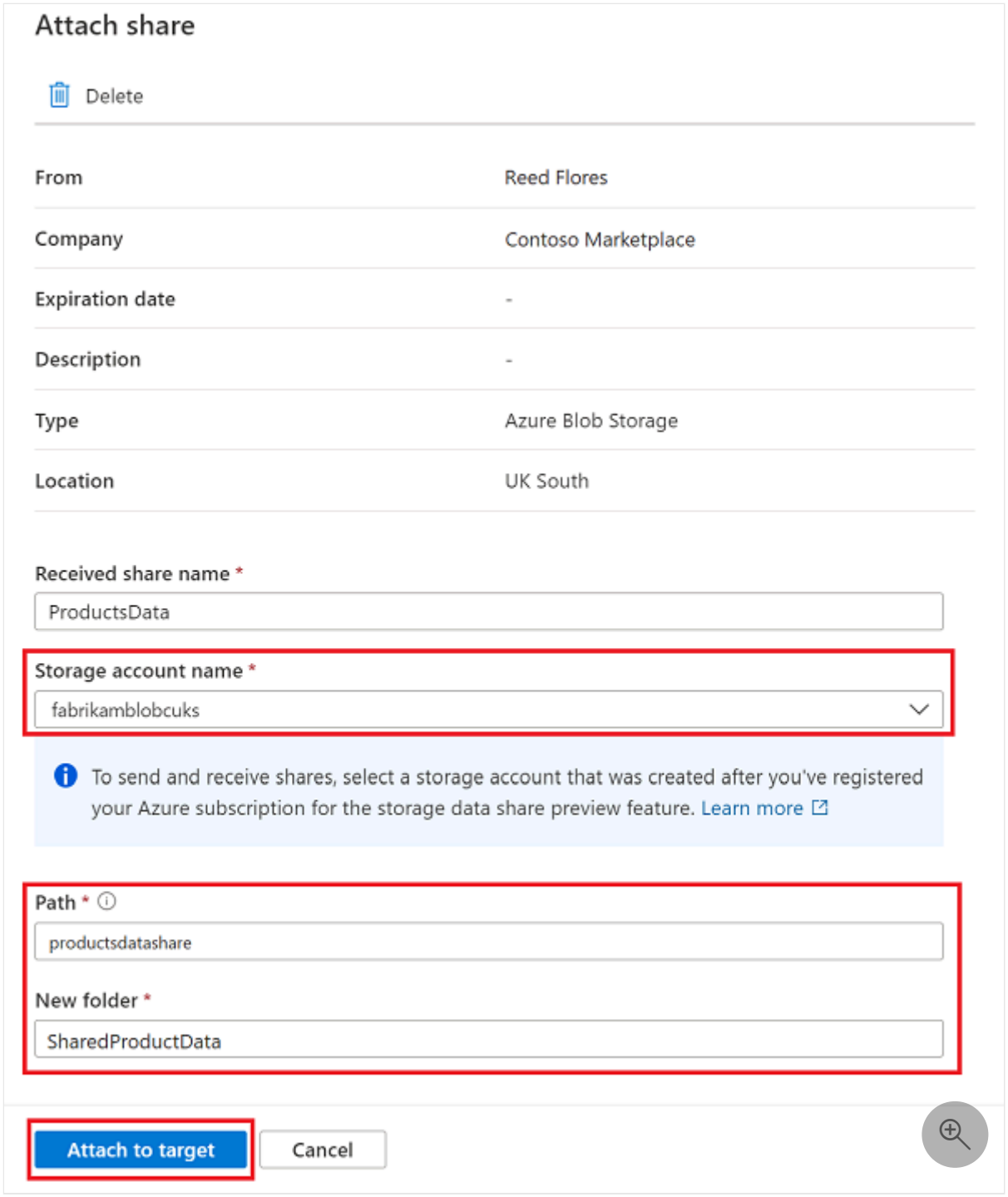
> **Note**
>
> If you delete an invitation, if you want to accept the share in future it will need to be resent. To deselect the share without deleting select the **Cancel** button instead.

7. You can edit the **Received share name** if you like. Then select a **Storage account name** for a target storage account in the same region as the source. You can choose to **Register a new storage account to attach the share** in the drop-down as well.

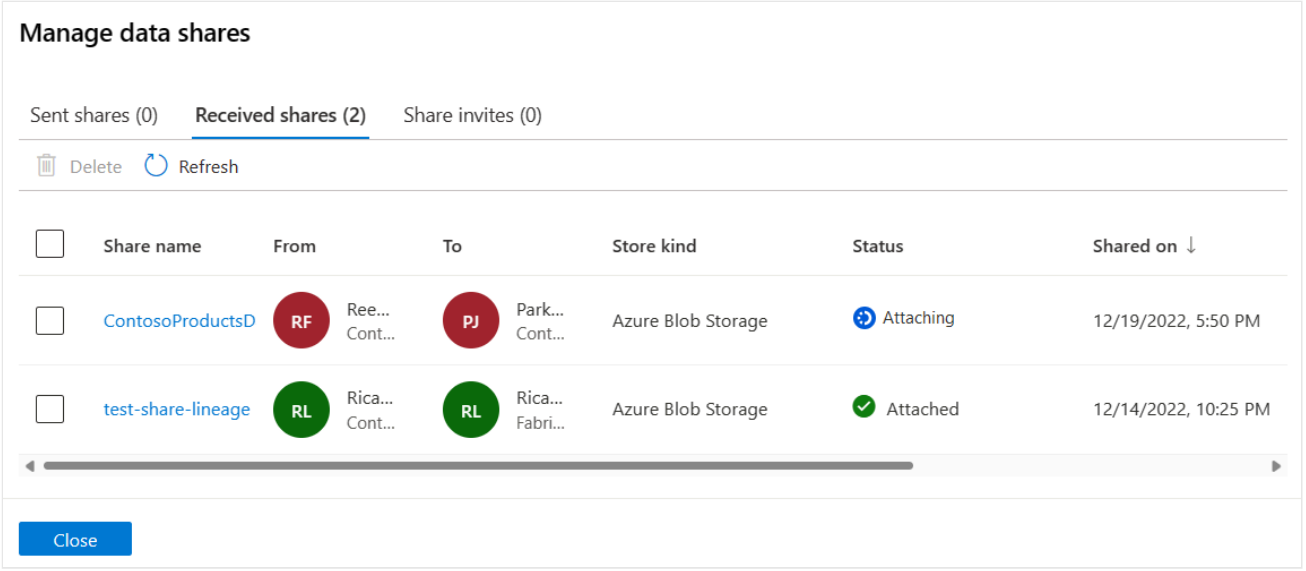
> **Important**
>
> The target storage account needs to be in the same Azure region as the source storage account.

8. Configure the **Path** (either a new container name, or the name of an existing share container) and, **New folder** (a new folder name for the share within in your container).

9. Select **Attach to target**.



10. On the Manage data shares page, you'll see the new share with the status of **Attaching** until it has completed and is attached.



11. You can access shared data from the target storage account through Azure portal, Azure Storage Explorer, Azure Storage SDK, PowerShell or CLI. You can also analyze the shared data by connecting your storage account to Azure Synapse Analytics Spark or Databricks.

Clean up resources

To clean up the resources created for the quick start, follow the steps below:

1. Within Microsoft Purview portal, delete the sent share.
2. Also delete your received share.
3. Once the shares are successfully deleted, delete the target container and folder Microsoft Purview created in your target storage account when you received shared data.

Troubleshoot

To troubleshoot issues with sharing data, refer to the troubleshooting section of the how to share data article. To troubleshoot issues with receiving share, refer to the troubleshooting section of the how to receive shared data article.

Next steps

- FAQ for data sharing
- How to share data
- How to receive share
- REST API reference

Last updated on 10/10/2025

Quickstart: Share and receive data with the Microsoft Purview Data Sharing .NET SDK

Article • 10/28/2024

> **Note**
>
> Microsoft Purview Data Map (classic), Data Catalog (classic), Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

> **Important**
>
> This feature is currently being retired, with support provided until September 2025 to allow for migration to Microsoft Fabric external data sharing.

In this quickstart, you'll use the .NET SDK to share data and receive shares from Azure Data Lake Storage (ADLS Gen2) or Blob storage accounts. The article includes code snippets that allow you to create, accept, and manage shares using Microsoft Purview Data Sharing.

For an overview of how data sharing works, watch this short demo.

> **Note**
>
> This feature has been updated in February 2023, and the SDK and permissions needed to view and manage data shares in Microsoft Purview have changed.
>
> - No permissions are now required in Microsoft Purview to use the SDK to create and manage shares. (Reader permissions are needed to use the Microsoft Purview Governance Portal for sharing data.)
> - Permissions are still required on storage accounts.
>
> See the updated NuGet package and updated code snippets to use the updated SDK.

Prerequisites

Microsoft Purview prerequisites

- A Microsoft Purview account. You can also use two Microsoft Purview accounts, one for data provider and one for data consumer to test both scenarios.
- Your recipient's Azure sign-in email address that you can use to send the invitation to. The recipient's email alias won't work.

Azure Storage account prerequisites

- Your Azure subscription must be registered for the **AllowDataSharing** preview feature. (If you haven't already registered this preview feature, consider using Microsoft Fabric external data sharing instead, since Microsoft Purview Data Sharing will be retired in September 2025.) Or, contact support.
- Source and target storage accounts **created after** the registration step is completed. **Both storage accounts must be in the same Azure region as each other**. Both storage accounts need to be ADLS Gen2 or Blob Storage accounts. Your storage accounts can be in a different Azure region from your Microsoft Purview account.
- Latest version of the storage SDK, PowerShell, CLI, and Azure Storage Explorer. Storage REST API version must be February 2020 or later.
- The storage accounts need to be registered in the collections where you'll send or receive the share. If you're using one Microsoft Purview account, this can be two different collections, or the same collection. For instructions to register, see the ADLS Gen2 or Blob storage data source pages.
- If the source or target storage accounts are in a different Azure subscription than the one for Microsoft Purview account, the Microsoft. Purview resource provider is automatically registered in the Azure subscription where the data store is located at the time of share provider adding an asset or share consumer mapping an asset and **ONLY** if the user has permission to do the /register/action operation for the resource provider. The permission is included in the Contributor and Owner roles.

  > ⓘ **Note**
  >
  > This registration is only needed the first time when sharing or receiving data into a storage account in the Azure subscription.

Required roles

Here are required roles for sharing data and receiving shares.

Expand table

	Azure Storage account roles	Microsoft Purview collection roles
Data provider	One of the following roles: • Owner • Storage Blob Data Owner	Data Reader
Data consumer	One of the following roles: • Contributor • Owner • Storage Blob Data Contributor • Storage Blob Data Owner	Data Reader

> **Note**
>
> If you created the Microsoft Purview account, you're automatically assigned all the roles to the root collection. Refer to Microsoft Purview permissions to learn more about the Microsoft Purview collection and roles.

Visual Studio

The walkthrough in this article uses Visual Studio 2022. The procedures for Visual Studio 2013, 2015, 2017, or 2019 may differ slightly.

Azure .NET SDK

Download and install Azure .NET SDK on your machine.

Use a service principal

In the code snippets in this tutorial, you can authenticate either using your own credentials or using a service principal. To set up a service principal, follow these instructions:

1. In Create a Microsoft Entra application, create an application that represents the .NET application you're creating in this tutorial. For the sign-on URL, you can provide a dummy URL as shown in the article (`https://contoso.org/exampleapp`).

2. In Get values for signing in, get the **application ID**,**tenant ID**, and **object ID**, and note down these values that you use later in this tutorial.

3. In Certificates and secrets, get the **authentication key**, and note down this value that you use later in this tutorial.

4. assign the application to these roles:

Expand table

User	Azure Storage account roles	Microsoft Purview collection roles
Data provider	One of the following roles: • Owner • Blob Storage Data Owner	Data Share Contributor
Data consumer	One of the following roles: • Contributor • Owner • Storage Blob Data Contributor • Blob Storage Data Owner	Data Share Contributor

Create a Visual Studio project

Next, create a C# .NET console application in Visual Studio:

1. Launch **Visual Studio**.
2. In the Start window, select **Create a new project** > **Console App**. .NET version 6.0 or above is required.
3. In **Project name**, enter **PurviewDataSharingQuickStart**.
4. Select **Create** to create the project.

Install NuGet packages

1. Select **Tools** > **NuGet Package Manager** > **Package Manager Console**.

2. In the **Package Manage Console** run the .NET cli add package command shown on this page to add the NuGet package: Microsoft.Azure.Analytics.Purview.Sharing NuGet package.

3. In the **Package Manager Console** pane, run the following commands to install packages.

```PowerShell
Install-Package Azure.Analytics.Purview.Sharing -IncludePrerelease
Install-Package Azure.Identity
```

> **Tip**
>
> If you get an error that reads "Could not find any project in..." when attempting these commands, you may just need to move a folder level down in your project. Try the command `dir` to list folders in your directory, then use 'cd <name of the project folder>' to move down a level to your project folder. Then try again.

Create a sent share

This script creates a data share that you can send to internal or external users. To use it, be sure to fill out these variables:

- **SenderTenantId** - the Azure Tenant ID for the sender's identity.
- **SenderPurviewAccountName** - the name of the Microsoft Purview account where the data will be sent from.
- **ShareName** - A display name for your sent share.
- **ShareDescription** - (optional) A description for your sent share.
- **SenderStorageKind** - either BlobAccount or AdlsGen2Account.
- **SenderStorageResourceId** - the resource ID for the storage account where the data will be sent from.
- **SenderStorageContainer** - the name of the container where the data to be shared is stored.
- **SenderPathToShare** - the file/folder path to the data to be shared.
- **UseServiceTokenCredentials** - (optional) If you want to use a service principal to create the shares, set this to **true**.
- **SenderClientId** - (optional) If using a service principal to create the shares, this is the Application (client) ID for the service principal.
- **SenderClientSecret** - (optional) If using a service principal to create the shares, add your client secret/authentication key.
- **SentShareID** - (optional) This option must be a GUID, and the current value generates one for you, but you can replace it with a different value if you would

like.

- **ReceiverVisiblePath** - (optional) The name for the share the receiver will see. Currently set to a GUID, but GUID is not required.
- **SenderPurviewEndPoint** - If you're using If you're using the classic Microsoft purview experience, use `$"https://{SenderPurviewAccountName}.purview.azure.com"`; If you're using the new Microsoft Purview experience, use `https://api.purview-service.microsoft.com/`

C#

```
using Azure;
using Azure.Analytics.Purview.Sharing;
using Azure.Core;
using Azure.Identity;
using System.ComponentModel;

public static class PurviewDataSharingQuickStart
{
    // [REQUIRED INPUTS] Set To Actual Values.
    private static string SenderTenantId = "<Sender Identity's Tenant ID>";
    private static string SenderPurviewAccountName = "<Sender Purview
Account Name>";

    private static string ShareName = "<Share Display Name>";
    private static string ShareDescription = "Share created using the SDK.";
    private static string SenderStorageKind = "<Sender Storage Account Kind
(BlobAccount / AdlsGen2Account)>";
    private static string SenderStorageResourceId = "<Sender Storage Account
Resource Id>";
    private static string SenderStorageContainer = "<Share Data Container
Name>";
    private static string SenderPathToShare = "<File/Folder Path To Share>";

    // Set if using Service principal to create shares
    private static bool UseServiceTokenCredentials = false;
    private static string SenderClientId = "<Sender Application (Client)
Id>";
    private static string SenderClientSecret = "<Sender Application (Client)
Secret>";

    // [OPTIONAL INPUTS] Override Value If Desired.
    private static string SentShareId = Guid.NewGuid().ToString();
    private static string ReceiverVisiblePath = Guid.NewGuid().ToString();

    // General Configs
    private static string SenderPurviewEndPoint =
$"https://{SenderPurviewAccountName}.purview.azure.com";
    private static string SenderShareEndPoint = $"
{SenderPurviewEndPoint}/share";

    private static async Task Main(string[] args)
    {
```

```
        try
        {
            /// Replace all placeholder inputs above with actual values
before running this program.
            /// This updated Share experience API will create Shares based
on callers RBAC role on the storage account.
            /// To view/manage Shares via UX in Purview Studio. Storage
accounts need to be registered (one time action) in Purview account with DSA
permissions.

            Console.WriteLine($"{DateTime.Now.ToShortTimeString()}:
CreateShare - START");
            await Sender_CreateSentShare();
            Console.WriteLine($"{DateTime.Now.ToShortTimeString()}:
CreateShare - FINISH");
        }
        catch (Exception ex)
        {
            Console.ForegroundColor = ConsoleColor.Red;
            Console.WriteLine(ex);
            Console.ForegroundColor = Console.ForegroundColor;
        }
    }

    private static async Task<BinaryData> Sender_CreateSentShare()
    {

        TokenCredential senderCredentials = UseServiceTokenCredentials
                ? new ClientSecretCredential(SenderTenantId, SenderClientId,
SenderClientSecret)
                : new DefaultAzureCredential(new
DefaultAzureCredentialOptions { AuthorityHost = new
Uri("https://login.windows.net"), TenantId = SenderTenantId });

        SentSharesClient? sentSharesClient = new
SentSharesClient(SenderShareEndPoint, senderCredentials);

        if (sentSharesClient == null)
        {
            throw new InvalidEnumArgumentException("Invalid Sent Shares
Client.");
        }

        // Create sent share
        var inPlaceSentShareDto = new
        {
            shareKind = "InPlace",
            properties = new
            {
                displayName = ShareName,
                description = ShareDescription,
                artifact = new
                {
                    storeKind = SenderStorageKind,
                    storeReference = new
```

```
                    {
                        referenceName = SenderStorageResourceId,
                        type = "ArmResourceReference"
                    },
                    properties = new
                    {
                        paths = new[]
                        {
                            new
                            {
                                receiverPath = ReceiverVisiblePath,
                                containerName = SenderStorageContainer,
                                senderPath = SenderPathToShare
                            }
                        }
                    }
                }
            },
        };

        Operation<BinaryData> sentShare = await
sentSharesClient.CreateOrReplaceSentShareAsync(WaitUntil.Completed,
SentShareId, RequestContent.Create(inPlaceSentShareDto));
        Console.ForegroundColor = ConsoleColor.Green;
        Console.WriteLine(sentShare.Value);
        Console.ForegroundColor = Console.ForegroundColor;
        return sentShare.Value;
    }
}
```

Send invitation to a user

This script sends an email invitation for a share to a user. If you want to send an invitation to a service principal, see the next code example. To use it, be sure to fill out these variables:

- **RecipientUserEmailId** - Email address for the user to send the invitation to.
- **SenderTenantId** - the Azure Tenant ID for the share sender's identity.
- **SenderPurviewAccountName** - the name of the Microsoft Purview account where the data will be sent from.
- **SenderStorageResourceId** - the resource ID for the storage account where the data will be sent from.
- **SentShareDisplayName** - the name of the sent share you're sending an invitation for.
- **UseServiceTokenCredentials** - (optional) If you want to use a service principal to create the shares, set this to **true**.

- **SenderClientId** - (optional) If using a service principal to create the shares, this is the Application (client) ID for the service principal.
- **SenderClientSecret** - (optional) If using a service principal to create the shares, add your client secret/authentication key.
- **InvitationId** - (optional) This option must be a GUID, and the current value generates one for you, but you can replace it with a different value if you would like.
- **SenderPurviewEndPoint** - If you're using If you're using the classic Microsoft purview experience, use `$"https://{SenderPurviewAccountName}.purview.azure.com"`; If you're using the new Microsoft Purview experience, use `https://api.purview-service.microsoft.com/`

C#

```
using Azure;
using Azure.Analytics.Purview.Sharing;
using Azure.Core;
using Azure.Identity;
using System.ComponentModel;
using System.Text.Json;

public static class PurviewDataSharingQuickStart
{
    // [REQUIRED INPUTS] Set To Actual Values.
    private static string RecipientUserEmailId = "<Target User's Email Id>";

    private static string SenderTenantId = "<Sender Indentity's Tenant ID>";
    private static string SenderPurviewAccountName = "<Sender Purview
Account Name>";
    private static string SenderStorageResourceId = "<Sender Storage Account
Resource Id>";

    // Set if using Service principal to send invitation
    private static bool UseServiceTokenCredentials = false;
    private static string SenderClientId = "<Sender Application (Client)
Id>";
    private static string SenderClientSecret = "<Sender Application (Client)
Secret>";

    private static string SentShareDisplayName = "<Name of share you're
sending an invite for.>";
    private static string InvitationId = Guid.NewGuid().ToString();

    // General Configs
    private static string SenderPurviewEndPoint =
$"https://{SenderPurviewAccountName}.purview.azure.com";
    private static string SenderShareEndPoint = $"
{SenderPurviewEndPoint}/share";
    private static int StepCounter = 0;
```

```
    private static async Task Main(string[] args)
    {
        try
        {

            Console.WriteLine($"{DateTime.Now.ToShortTimeString()}:
SendtoUser - START");
            await Sender_CreateUserRecipient();
            Console.WriteLine($"{DateTime.Now.ToShortTimeString()}:
SendtoUser - FINISH");
        }
        catch (Exception ex)
        {
            Console.ForegroundColor = ConsoleColor.Red;
            Console.WriteLine(ex);
            Console.ForegroundColor = Console.ForegroundColor;
        }
    }

    public static async Task<List<T>> ToResultList<T>(this AsyncPageable<T>
asyncPageable)
    {
        List<T> list = new List<T>();

        await foreach (T item in asyncPageable)
        {
            list.Add(item);
        }

        return list;
    }

    private static async Task<BinaryData> Sender_CreateUserRecipient()
    {

        TokenCredential senderCredentials = UseServiceTokenCredentials
                ? new ClientSecretCredential(SenderTenantId, SenderClientId,
SenderClientSecret)
                : new DefaultAzureCredential(new
DefaultAzureCredentialOptions { AuthorityHost = new
Uri("https://login.windows.net"), TenantId = SenderTenantId });

        SentSharesClient? sentSharesClient = new
SentSharesClient(SenderShareEndPoint, senderCredentials);

        if (string.IsNullOrEmpty(RecipientUserEmailId))
        {
            throw new InvalidEnumArgumentException("Invalid Recipient User
Email Id.");
        }

        // Create user recipient and invite
        var invitationData = new
        {
            invitationKind = "User",
```

```
            properties = new
            {
                expirationDate = DateTime.Now.AddDays(7).ToString(),
                notify = true, // Send invitation email
                targetEmail = RecipientUserEmailId
            }
        };

        var allSentShares = await
sentSharesClient.GetAllSentSharesAsync(SenderStorageResourceId).ToResultList
();

        Console.ForegroundColor = ConsoleColor.Yellow;
        Console.WriteLine("{0}. {1}...", ++StepCounter, "Get a Specific Sent
Share");
        Console.ForegroundColor = Console.ForegroundColor;

        var mySentShare = allSentShares.First(sentShareDoc =>
        {
            var doc = JsonDocument.Parse(sentShareDoc).RootElement;
            var props = doc.GetProperty("properties");
            return props.GetProperty("displayName").ToString() ==
SentShareDisplayName;
        });

        Console.ForegroundColor = ConsoleColor.Green;
        Console.WriteLine("My Sent Share Id: " +
JsonDocument.Parse(mySentShare).RootElement.GetProperty("id").ToString());
        Console.ForegroundColor = Console.ForegroundColor;

        var SentShareId =
JsonDocument.Parse(mySentShare).RootElement.GetProperty("id").ToString();

        var sentInvitation = await
sentSharesClient.CreateSentShareInvitationAsync(SentShareId, InvitationId,
RequestContent.Create(invitationData));

        Console.ForegroundColor = ConsoleColor.Green;
        Console.WriteLine(sentInvitation.Content);
        Console.ForegroundColor = Console.ForegroundColor;

        return sentInvitation.Content;
    }
}
```

Send invitation to a service

This script sends an email invitation for a share to a service principal. If you want to send an invitation to a user, see the previous sample. To use it, be sure to fill out these variables:

- **RecipientApplicationTenantId** - the Azure Tenant ID for the receiving service principal.
- **RecipientApplicationObjectId** - the object ID for the receiving service principal.
- **SenderTenantId** - the Azure Tenant ID for the share sender's identity.
- **SenderPurviewAccountName** - the name of the Microsoft Purview account where the data will be sent from.
- **SenderStorageResourceId** - the resource ID for the storage account where the data will be sent from.
- **SentShareDisplayName** - the name of the sent share you're sending an invitation for.
- **UseServiceTokenCredentials** - (optional) If you want to use a service principal to create the shares, set this to **true**.
- **SenderClientId** - (optional) If using a service principal to create the shares, this is the Application (client) ID for the service principal.
- **SenderClientSecret** - (optional) If using a service principal to create the shares, add your client secret/authentication key.
- **InvitationId** - (optional) This option must be a GUID, and the current value generates one for you, but you can replace it with a different value if you would like.
- **SenderPurviewEndPoint** - If you're using If you're using the classic Microsoft purview experience, use `$"https://{SenderPurviewAccountName}.purview.azure.com"`; If you're using the new Microsoft Purview experience, use `https://api.purview-service.microsoft.com/`

C#

```
using Azure;
using Azure.Analytics.Purview.Sharing;
using Azure.Core;
using Azure.Identity;
using System.ComponentModel;
using System.Text.Json;

public static class PurviewDataSharingQuickStart
{
    // [REQUIRED INPUTS] Set To Actual Values.
    private static string RecipientApplicationTenantId = "<Target
Application's Tenant Id>";
    private static string RecipientApplicationObjectId = "<Target
Application's Object Id>";

    private static string SentShareDisplayName = "<Name of share you're
sending an invite for.>";
    private static string InvitationId = Guid.NewGuid().ToString();

    private static string SenderTenantId = "<Sender Indentity's Tenant ID>";
```

```
    private static string SenderPurviewAccountName = "<Sender Purview
Account Name>";

    private static string SenderStorageResourceId = "<Resource ID for
storage account that has been shared>";

    // Set if using Service principal to send invitation
    private static bool UseServiceTokenCredentials = false;
    private static string SenderClientId = "<Sender Application (Client)
Id>";
    private static string SenderClientSecret = "<Sender Application (Client)
Secret>";

    // General Configs
    private static string SenderPurviewEndPoint =
$"https://{SenderPurviewAccountName}.purview.azure.com";
    private static string SenderShareEndPoint = $"
{SenderPurviewEndPoint}/share";

    private static async Task Main(string[] args)
    {
        try
        {

            Console.WriteLine($"{DateTime.Now.ToShortTimeString()}:
SendtoService - START");
            await Sender_CreateServiceRecipient();
            Console.WriteLine($"{DateTime.Now.ToShortTimeString()}:
SendtoService - FINISH");
        }
        catch (Exception ex)
        {
            Console.ForegroundColor = ConsoleColor.Red;
            Console.WriteLine(ex);
            Console.ForegroundColor = Console.ForegroundColor;
        }
    }

    public static async Task<List<T>> ToResultList<T>(this AsyncPageable<T>
asyncPageable)
    {
        List<T> list = new List<T>();

        await foreach (T item in asyncPageable)
        {
            list.Add(item);
        }

        return list;
    }

    private static async Task<BinaryData> Sender_CreateServiceRecipient()
    {

        TokenCredential senderCredentials = UseServiceTokenCredentials
```

```
                ? new ClientSecretCredential(SenderTenantId, SenderClientId,
SenderClientSecret)
                : new DefaultAzureCredential(new
DefaultAzureCredentialOptions { AuthorityHost = new
Uri("https://login.windows.net"), TenantId = SenderTenantId });

        SentSharesClient? sentSharesClient = new
SentSharesClient(SenderShareEndPoint, senderCredentials);

        if (!Guid.TryParse(RecipientApplicationTenantId, out Guid _))
        {
            throw new InvalidEnumArgumentException("Invalid Recipient
Service Tenant Id.");
        }

        if (!Guid.TryParse(RecipientApplicationObjectId, out Guid _))
        {
            throw new InvalidEnumArgumentException("Invalid Recipient
Service Object Id.");
        }

        // Create service recipient
        var invitationData = new
        {
            invitationKind = "Service",
            properties = new
            {
                expirationDate = DateTime.Now.AddDays(5).ToString(),
                targetActiveDirectoryId = RecipientApplicationTenantId,
                targetObjectId = RecipientApplicationObjectId
            }
        };


        var allSentShares = await
sentSharesClient.GetAllSentSharesAsync(SenderStorageResourceId).ToResultList
();
        var mySentShare = allSentShares.First(sentShareDoc =>
        {
            var doc = JsonDocument.Parse(sentShareDoc).RootElement;
            var props = doc.GetProperty("properties");
            return props.GetProperty("displayName").ToString() ==
SentShareDisplayName;
        });

        var SentShareId =
JsonDocument.Parse(mySentShare).RootElement.GetProperty("id").ToString();

        var sentInvitation = await
sentSharesClient.CreateSentShareInvitationAsync(SentShareId, InvitationId,
RequestContent.Create(invitationData));
        Console.ForegroundColor = ConsoleColor.Green;
        Console.WriteLine(sentInvitation.Content);
        Console.ForegroundColor = Console.ForegroundColor;
        return sentInvitation.Content;
```

```
    }

}
```

List sent shares

This script lists all the sent shares for a specific storage resource. To use it, be sure to fill out these variables:

- **SenderTenantId** - the Azure Tenant ID for the share sender's identity.
- **SenderPurviewAccountName** - the name of the Microsoft Purview account where the data was sent from.
- **SenderStorageResourceId** - the resource ID for the storage account where shares have been sent from.
- **UseServiceTokenCredentials** - (optional) If you want to use a service principal to create the shares, set this to **true**.
- **SenderClientId** - (optional) If using a service principal to create the shares, this is the Application (client) ID for the service principal.
- **SenderClientSecret** - (optional) If using a service principal to create the shares, add your client secret/authentication key.
- **SenderPurviewEndPoint** - If you're using If you're using the classic Microsoft purview experience, use `$"https://{SenderPurviewAccountName}.purview.azure.com"`; If you're using the new Microsoft Purview experience, use `https://api.purview-service.microsoft.com/`

C#

```
using Azure;
using Azure.Analytics.Purview.Sharing;
using Azure.Core;
using Azure.Identity;

public static class PurviewDataSharingQuickStart
{
    // [REQUIRED INPUTS] Set To Actual Values.
    private static string SenderTenantId = "<Sender Tenant ID>";
    private static string SenderPurviewAccountName = "<Name of the Microsoft
Purview account>";
    private static string SenderStorageResourceId = "<Sender Storage Account
Resource Id>";

    // Set if using Service principal to list shares
    private static bool UseServiceTokenCredentials = false;
    private static string SenderClientId = "<Sender Application (Client)
Id>";
    private static string SenderClientSecret = "<Sender Application (Client)
```

```
Secret>";

    // General Configs
    private static string SenderPurviewEndPoint =
$"https://{SenderPurviewAccountName}.purview.azure.com";
    private static string SenderShareEndPoint = $"
{SenderPurviewEndPoint}/share";

    private static async Task Main(string[] args)
    {


        try
        {
            TokenCredential senderCredentials = UseServiceTokenCredentials
                ? new ClientSecretCredential(SenderTenantId, SenderClientId,
SenderClientSecret)
                : new DefaultAzureCredential(new
DefaultAzureCredentialOptions { AuthorityHost = new
Uri("https://login.windows.net"), TenantId = SenderTenantId });

            SentSharesClient? sentSharesClient = new
SentSharesClient(SenderShareEndPoint, senderCredentials);

            var allSentShares = await
sentSharesClient.GetAllSentSharesAsync(SenderStorageResourceId).ToResultList
();
            Console.WriteLine(allSentShares);
        }
        catch (Exception ex)
        {
            Console.ForegroundColor = ConsoleColor.Red;
            Console.WriteLine(ex);
            Console.ForegroundColor = Console.ForegroundColor;
        }
    }

    public static async Task<List<T>> ToResultList<T>(this AsyncPageable<T>
asyncPageable)
    {
        List<T> list = new List<T>();

        await foreach (T item in asyncPageable)
        {
            list.Add(item);
        }

        return list;
    }
}
```

List all share recipients

This script lists all recipients for a specific share. To use it, be sure to fill out these variables:

- **SenderTenantId** - the Azure Tenant ID for the share sender's identity.
- **SenderPurviewAccountName** - the name of the Microsoft Purview account where the data was sent from.
- **SentShareDisplayName** - the name of the sent share you're listing recipients for.
- **SenderStorageResourceId** - the resource ID for the storage account where the data will be sent from.
- **UseServiceTokenCredentials** - (optional) If you want to use a service principal to create the shares, set this to **true**.
- **SenderClientId** - (optional) If using a service principal to create the shares, this is the Application (client) ID for the service principal.
- **SenderClientSecret** - (optional) If using a service principal to create the shares, add your client secret/authentication key.
- **SenderPurviewEndPoint** - If you're using If you're using the classic Microsoft purview experience, use `$"https://{SenderPurviewAccountName}.purview.azure.com"`; If you're using the new Microsoft Purview experience, use `https://api.purview-service.microsoft.com/`

C#

```
using Azure;
using Azure.Analytics.Purview.Sharing;
using Azure.Core;
using Azure.Identity;
using System.Text.Json;

public static class PurviewDataSharingQuickStart
{
    // [REQUIRED INPUTS] Set To Actual Values.
    private static string SentShareDisplayName = "<Name of share you're
listing recipients for.>";
    private static string SenderTenantId = "<Sender Tenant ID>";
    private static string SenderPurviewAccountName = "<Name of the Microsoft
Purview account>";

    private static string SenderStorageResourceId = "<Sender Storage Account
Resource Id>";

    // Set if using Service principal to list recipients
    private static bool UseServiceTokenCredentials = false;
    private static string SenderClientId = "<Sender Application (Client)
Id>";
    private static string SenderClientSecret = "<Sender Application (Client)
Secret>";

    // General Configs
```

```
    private static string SenderPurviewEndPoint =
$"https://{SenderPurviewAccountName}.purview.azure.com";
    private static string SenderShareEndPoint = $"
{SenderPurviewEndPoint}/share";

    private static async Task Main(string[] args)
    {
        try
        {
            TokenCredential senderCredentials = UseServiceTokenCredentials
                ? new ClientSecretCredential(SenderTenantId, SenderClientId,
SenderClientSecret)
                : new DefaultAzureCredential(new
DefaultAzureCredentialOptions { AuthorityHost = new
Uri("https://login.windows.net"), TenantId = SenderTenantId });

            SentSharesClient? sentSharesClient = new
SentSharesClient(SenderShareEndPoint, senderCredentials);

            var allSentShares = await
sentSharesClient.GetAllSentSharesAsync(SenderStorageResourceId).ToResultList
();
            var mySentShare = allSentShares.First(sentShareDoc =>
            {
                var doc = JsonDocument.Parse(sentShareDoc).RootElement;
                var props = doc.GetProperty("properties");
                return props.GetProperty("displayName").ToString() ==
SentShareDisplayName;
            });

            Console.ForegroundColor = ConsoleColor.Green;
            Console.WriteLine("My Sent Share Id: " +
JsonDocument.Parse(mySentShare).RootElement.GetProperty("id").ToString());
            Console.ForegroundColor = Console.ForegroundColor;

            var SentShareId =
JsonDocument.Parse(mySentShare).RootElement.GetProperty("id").ToString();

            var allRecipients = await
sentSharesClient.GetAllSentShareInvitationsAsync(SentShareId).ToResultList()
;
            Console.WriteLine(allRecipients);

        }
        catch (Exception ex)
        {
            Console.ForegroundColor = ConsoleColor.Red;
            Console.WriteLine(ex);
            Console.ForegroundColor = Console.ForegroundColor;
        }
    }

    public static async Task<List<T>> ToResultList<T>(this AsyncPageable<T>
asyncPageable)
    {
```

```
        List<T> list = new List<T>();

        await foreach (T item in asyncPageable)
        {
            list.Add(item);
        }

        return list;
    }
}
```

Delete recipient

This script removes a share invitation, and therefore the share, for a recipient. To use it, be sure to fill out these variables:

- **SenderTenantId** - the Azure Tenant ID for the share sender's identity.
- **SenderPurviewAccountName** - the name of the Microsoft Purview account where the data was sent from.
- **SentShareDisplayName** - the name of the sent share you're removing a recipient for.
- **SenderStorageResourceId** - the resource ID for the storage account where the data will be sent from.
- **RecipientUserEmailId** - Email address for the user you want to delete.
- **UseServiceTokenCredentials** - (optional) If you want to use a service principal to create the shares, set this to **true**.
- **SenderClientId** - (optional) If using a service principal to create the shares, this is the Application (client) ID for the service principal.
- **SenderClientSecret** - (optional) If using a service principal to create the shares, add your client secret/authentication key.
- **SenderPurviewEndPoint** - If you're using If you're using the classic Microsoft purview experience, use `$"https://{SenderPurviewAccountName}.purview.azure.com"`; If you're using the new Microsoft Purview experience, use `https://api.purview-service.microsoft.com/`

C#

```
using Azure;
using Azure.Analytics.Purview.Sharing;
using Azure.Core;
using Azure.Identity;
using System.Text.Json;

public static class PurviewDataSharingQuickStart
{
```

```
    // [REQUIRED INPUTS] Set To Actual Values.
    private static string SentShareDisplayName = "<Name of share you're
removing a recipient for.>";
    private static string SenderTenantId = "<Sender Tenant ID>";
    private static string SenderPurviewAccountName = "<Name of the Microsoft
Purview account>";
    private static string RecipientUserEmailId = "<Target User's Email Id>";

    private static string SenderStorageResourceId = "<Sender Storage Account
Resource Id>";

    // Set if using Service principal to delete recipients
    private static bool UseServiceTokenCredentials = false;
    private static string SenderClientId = "<Sender Application (Client)
Id>";
    private static string SenderClientSecret = "<Sender Application (Client)
Secret>";

    // General Configs
    private static string SenderPurviewEndPoint =
$"https://{SenderPurviewAccountName}.purview.azure.com";
    private static string SenderShareEndPoint = $"
{SenderPurviewEndPoint}/share";

    private static async Task Main(string[] args)
    {
        try
        {
            TokenCredential senderCredentials = UseServiceTokenCredentials
                ? new ClientSecretCredential(SenderTenantId, SenderClientId,
SenderClientSecret)
                : new DefaultAzureCredential(new
DefaultAzureCredentialOptions { AuthorityHost = new
Uri("https://login.windows.net"), TenantId = SenderTenantId });

            SentSharesClient? sentSharesClient = new
SentSharesClient(SenderShareEndPoint, senderCredentials);

            var allSentShares = await
sentSharesClient.GetAllSentSharesAsync(SenderStorageResourceId).ToResultList
();

            var mySentShare = allSentShares.First(sentShareDoc =>
            {
                var doc = JsonDocument.Parse(sentShareDoc).RootElement;
                var props = doc.GetProperty("properties");
                return props.GetProperty("displayName").ToString() ==
SentShareDisplayName;
            });

            var SentShareId =
JsonDocument.Parse(mySentShare).RootElement.GetProperty("id").ToString();

            var allRecipients = await
sentSharesClient.GetAllSentShareInvitationsAsync(SentShareId).ToResultList()
```

```
;

            var recipient = allRecipients.First(recipient =>
            {
                var doc = JsonDocument.Parse(recipient).RootElement;
                var props = doc.GetProperty("properties");
                return props.TryGetProperty("targetEmail", out JsonElement
rcpt) && rcpt.ToString() == RecipientUserEmailId;
            });

            var recipientId =
JsonDocument.Parse(recipient).RootElement.GetProperty("id").ToString();

            await
sentSharesClient.DeleteSentShareInvitationAsync(WaitUntil.Completed,
SentShareId, recipientId);

            Console.ForegroundColor = ConsoleColor.Green;
            Console.WriteLine("Remove Id: " +
JsonDocument.Parse(recipient).RootElement.GetProperty("id").ToString());
            Console.WriteLine("Complete");
            Console.ForegroundColor = Console.ForegroundColor;

        }

        catch (Exception ex)
        {
            Console.ForegroundColor = ConsoleColor.Red;
            Console.WriteLine(ex);
            Console.ForegroundColor = Console.ForegroundColor;
        }
    }

    public static async Task<List<T>> ToResultList<T>(this AsyncPageable<T>
asyncPageable)
    {
        List<T> list = new List<T>();

        await foreach (T item in asyncPageable)
        {
            list.Add(item);
        }

        return list;
    }
}
```

Delete sent share

This script deletes a sent share. To use it, be sure to fill out these variables:

- **SenderTenantId** - the Azure Tenant ID for the share sender's identity.

- **SenderPurviewAccountName** - the name of the Microsoft Purview account where the data was sent from.
- **SentShareDisplayName** - the name of the sent share you're listing recipients for.
- **SenderStorageResourceId** - the resource ID for the storage account where the data will be sent from.
- **UseServiceTokenCredentials** - (optional) If you want to use a service principal to create the shares, set this to **true**.
- **SenderClientId** - (optional) If using a service principal to create the shares, this is the Application (client) ID for the service principal.
- **SenderClientSecret** - (optional) If using a service principal to create the shares, add your client secret/authentication key.
- **SenderPurviewEndPoint** - If you're using If you're using the classic Microsoft purview experience, use `$"https://{SenderPurviewAccountName}.purview.azure.com"`; If you're using the new Microsoft Purview experience, use `https://api.purview-service.microsoft.com/`

C#

```
using Azure;
using Azure.Analytics.Purview.Sharing;
using Azure.Core;
using Azure.Identity;
using System.Text.Json;

public static class PurviewDataSharingQuickStart
{
    // [REQUIRED INPUTS] Set To Actual Values.
    private static string SenderTenantId = "<Sender Tenant ID>";
    private static string SenderPurviewAccountName = "<Name of the Microsoft
Purview account>";
    private static string SentShareDisplayName = "<Name of share you're
removing.>";

    private static string SenderStorageResourceId = "<Sender Storage Account
Resource Id>";

    // Set if using Service principal to delete share
    private static bool UseServiceTokenCredentials = false;
    private static string SenderClientId = "<Sender Application (Client)
Id>";
    private static string SenderClientSecret = "<Sender Application (Client)
Secret>";

    // General Configs
    private static string SenderPurviewEndPoint =
$"https://{SenderPurviewAccountName}.purview.azure.com";
    private static string SenderShareEndPoint = $"
{SenderPurviewEndPoint}/share";
```

```
    private static async Task Main(string[] args)
    {
        try
        {
            TokenCredential senderCredentials = UseServiceTokenCredentials
                ? new ClientSecretCredential(SenderTenantId, SenderClientId,
SenderClientSecret)
                : new DefaultAzureCredential(new
DefaultAzureCredentialOptions { AuthorityHost = new
Uri("https://login.windows.net"), TenantId = SenderTenantId });

            SentSharesClient? sentSharesClient = new
SentSharesClient(SenderShareEndPoint, senderCredentials);

            var allSentShares = await
sentSharesClient.GetAllSentSharesAsync(SenderStorageResourceId).ToResultList
();

            var mySentShare = allSentShares.First(sentShareDoc =>
            {
                var doc = JsonDocument.Parse(sentShareDoc).RootElement;
                var props = doc.GetProperty("properties");
                return props.GetProperty("displayName").ToString() ==
SentShareDisplayName;
            });

            var SentShareId =
JsonDocument.Parse(mySentShare).RootElement.GetProperty("id").ToString();

            await sentSharesClient.DeleteSentShareAsync(WaitUntil.Completed,
SentShareId);

            Console.ForegroundColor = ConsoleColor.Green;
            Console.WriteLine("Remove Id: " + SentShareId);
            Console.WriteLine("Complete");
            Console.ForegroundColor = Console.ForegroundColor;

        }
        catch (Exception ex)
        {
            Console.ForegroundColor = ConsoleColor.Red;
            Console.WriteLine(ex);
            Console.ForegroundColor = Console.ForegroundColor;
        }
    }

    public static async Task<List<T>> ToResultList<T>(this AsyncPageable<T>
asyncPageable)
    {
        List<T> list = new List<T>();

        await foreach (T item in asyncPageable)
        {
            list.Add(item);
        }
```

```
        return list;
    }
}
```

Create a received share

This script allows you to receive a data share. To use it, be sure to fill out these variables:

- **ReceiverTenantId** - the Azure Tenant ID for the user/service that is receiving the shared data.
- **ReceiverPurviewAccountName** - the name of the Microsoft Purview account where the data will be received.
- **ReceiverStorageKind** - either BlobAccount or AdlsGen2Account.
- **ReceiverStorageResourceId** - the resource ID for the storage account where the data will be received.
- **ReceiverStorageContainer** - the name of the container where the shared data will be stored.
- **ReceiverTargetFolderName** - the folder path to where the shared data will be stored.
- **ReceiverTargetMountPath** - the mount path you'd like to use to store your data in the folder.
- **UseServiceTokenCredentials** - (optional) If you want to use a service principal to receive the shares, set this to **true**.
- **ReceiverClientId** - (optional) If using a service principal to receive the shares, this is the Application (client) ID for the service principal.
- **ReceiverClientSecret** - (optional) If using a service principal to receive the shares, add your client secret/authentication key.
- **ReceivedShareId** - (optional) This option must be a GUID, and the current value will generate one for you, but you can replace it with a different value if you would like.
- **ReceiverVisiblePath** - (optional) Name you want to use for the path for your received share.
- **ReceivedShareDisplayName** - (optional) A display name for your received share.
- **ReceiverPurviewEndPoint** - If you're using If you're using the classic Microsoft purview experience, use `$"https://{ReceiverPurviewAccountName}.purview.azure.com"`; If you're using the new Microsoft Purview experience, use `https://api.purview-service.microsoft.com/`

C#

```
using Azure;
using Azure.Analytics.Purview.Sharing;
using Azure.Core;
using Azure.Identity;
using System.ComponentModel;
using System.Text.Json;

public static class PurviewDataSharingQuickStart
{
    // [REQUIRED INPUTS] Set To Actual Values.
    private static string ReceiverTenantId = "<Receiver Indentity's Tenant
ID>";
    private static string ReceiverPurviewAccountName = "<Receiver Purview
Account Name>";

    private static string ReceiverStorageKind = "<Receiver Storage Account
Kind (BlobAccount / AdlsGen2Account)>";
    private static string ReceiverStorageResourceId = "<Receiver Storage
Account Resource Id>";
    private static string ReceiverStorageContainer = "<Container Name To
Receive Data Under>";
    private static string ReceiverTargetFolderName = "<Folder Name to
Received Data Under>";
    private static string ReceiverTargetMountPath = "<Mount Path to store
Received Data Under>";

    //Use if using a service principal to receive a share
    private static bool UseServiceTokenCredentials = false;
    private static string ReceiverClientId = "<Receiver Caller Application
(Client) Id>";
    private static string ReceiverClientSecret = "<Receiver Caller
Application (Client) Secret>";

    // [OPTIONAL INPUTS] Override Values If Desired.
    private static string ReceivedShareId = Guid.NewGuid().ToString();
    private static string ReceiverVisiblePath = "ReceivedSharePath";

    private static string ReceivedShareDisplayName = "ReceivedShare";

    // General Configs
    private static string ReceiverPurviewEndPoint =
$"https://{ReceiverPurviewAccountName}.purview.azure.com";
    private static string ReceiverShareEndPoint = $"
{ReceiverPurviewEndPoint}/share";

    private static async Task Main(string[] args)
    {
        try
        {


            Console.WriteLine($"{DateTime.Now.ToShortTimeString()}:
CreateReceivedShare - START");
            await Receiver_CreateReceivedShare();
```

```
            Console.WriteLine($"{DateTime.Now.ToShortTimeString()}:
CreateReceivedShare - FINISH");
        }
        catch (Exception ex)
        {
            Console.ForegroundColor = ConsoleColor.Red;
            Console.WriteLine(ex);
            Console.ForegroundColor = Console.ForegroundColor;
        }
    }

    private static async Task<BinaryData> Receiver_CreateReceivedShare()
    {

        TokenCredential receiverCredentials = UseServiceTokenCredentials
            ? new ClientSecretCredential(ReceiverTenantId, ReceiverClientId,
ReceiverClientSecret)
            : new DefaultAzureCredential(new DefaultAzureCredentialOptions {
AuthorityHost = new Uri("https://login.windows.net"), TenantId =
ReceiverTenantId });

        ReceivedSharesClient? receivedSharesClient = new
ReceivedSharesClient(ReceiverShareEndPoint, receiverCredentials);

        if (receivedSharesClient == null)
        {
            throw new InvalidEnumArgumentException("Invalid Received Shares
Client.");
        }

        var results = await
receivedSharesClient.GetAllDetachedReceivedSharesAsync().ToResultList();
        var detachedReceivedShare = results;

        if (detachedReceivedShare == null)
        {
            throw new InvalidOperationException("No received shares
found.");
        }



        var myReceivedShare = detachedReceivedShare.First(recShareDoc =>
        {
            var doc = JsonDocument.Parse(recShareDoc).RootElement;
            var props = doc.GetProperty("properties");
            return props.GetProperty("displayName").ToString() ==
ReceivedShareDisplayName;
        });

        var ReceivedShareId =
JsonDocument.Parse(myReceivedShare).RootElement.GetProperty("id").ToString()
;
```

```
        var attachedReceivedShareData = new
        {
            shareKind = "InPlace",
            properties = new
            {
                displayName = ReceivedShareDisplayName,
                sink = new
                {
                    storeKind = ReceiverStorageKind,
                    properties = new
                    {
                        containerName = ReceiverStorageContainer,
                        folder = ReceiverTargetFolderName,
                        mountPath = ReceiverTargetMountPath
                    },
                    storeReference = new
                    {
                        referenceName = ReceiverStorageResourceId,
                        type = "ArmResourceReference"
                    }
                }
            }
        };

        var receivedShare = await
receivedSharesClient.CreateOrReplaceReceivedShareAsync(WaitUntil.Completed,
ReceivedShareId, RequestContent.Create(attachedReceivedShareData));

        Console.ForegroundColor = ConsoleColor.Green;
        Console.WriteLine(receivedShare.Value);
        Console.ForegroundColor = Console.ForegroundColor;

        return receivedShare.Value;
    }

    public static async Task<List<T>> ToResultList<T>(this AsyncPageable<T>
asyncPageable)
    {
        List<T> list = new List<T>();

        await foreach (T item in asyncPageable)
        {
            list.Add(item);
        }

        return list;
    }
}
```

List all received shares

This script lists all received shares on a storage account. To use it, be sure to fill out these variables:

- **ReceiverTenantId** - the Azure Tenant ID for the user/service that is receiving the shared data.
- **ReceiverPurviewAccountName** - the name of the Microsoft Purview account where the data was received.
- **ReceiverStorageResourceId** - the resource ID for the storage account where the data has been shared.
- **UseServiceTokenCredentials** - (optional) If you want to use a service principal to receive the shares, set this to **true**.
- **ReceiverClientId** - (optional) If using a service principal to receive the shares, this is the Application (client) ID for the service principal.
- **ReceiverClientSecret** - (optional) If using a service principal to receive the shares, add your client secret/authentication key.
- **ReceiverPurviewEndPoint** - If you're using If you're using the classic Microsoft purview experience, use `$"https://{ReceiverPurviewAccountName}.purview.azure.com"`; If you're using the new Microsoft Purview experience, use `https://api.purview-service.microsoft.com/`

C#

```
using Azure;
using Azure.Analytics.Purview.Sharing;
using Azure.Core;
using Azure.Identity;

public static class PurviewDataSharingQuickStart
{
    // [REQUIRED INPUTS] Set To Actual Values.
    private static string ReceiverTenantId = "<Receiver Indentity's Tenant
ID>";
    private static string ReceiverPurviewAccountName = "<Receiver Purview
Account Name>";

    private static string ReceiverStorageResourceId = "<Storage Account
Resource Id that is housing shares>";

    //Use if using a service principal to list shares
    private static bool UseServiceTokenCredentials = false;
    private static string ReceiverClientId = "<Receiver Caller Application
(Client) Id>";
    private static string ReceiverClientSecret = "<Receiver Caller
Application (Client) Secret>";

    // General Configs
    private static string ReceiverPurviewEndPoint =
```

```
$"https://{ReceiverPurviewAccountName}.purview.azure.com";
    private static string ReceiverShareEndPoint = $"
{ReceiverPurviewEndPoint}/share";

    private static async Task Main(string[] args)
    {
        try
        {
            TokenCredential receiverCredentials = UseServiceTokenCredentials
            ? new ClientSecretCredential(ReceiverTenantId, ReceiverClientId,
ReceiverClientSecret)
            : new DefaultAzureCredential(new DefaultAzureCredentialOptions {
AuthorityHost = new Uri("https://login.windows.net"), TenantId =
ReceiverTenantId });

            ReceivedSharesClient? receivedSharesClient = new
ReceivedSharesClient(ReceiverShareEndPoint, receiverCredentials);

            var allReceivedShares = await
receivedSharesClient.GetAllAttachedReceivedSharesAsync(ReceiverStorageResour
ceId).ToResultList();
            Console.WriteLine(allReceivedShares);
        }
        catch (Exception ex)
        {
            Console.ForegroundColor = ConsoleColor.Red;
            Console.WriteLine(ex);
            Console.ForegroundColor = Console.ForegroundColor;
        }
    }

    public static async Task<List<T>> ToResultList<T>(this AsyncPageable<T>
asyncPageable)
    {
        List<T> list = new List<T>();

        await foreach (T item in asyncPageable)
        {
            list.Add(item);
        }

        return list;
    }
}
```

Update received share

This script allows you to update the storage location for a received share. Just like creating a received share, you add the information for the storage account where you want the data to be housed. To use it, be sure to fill out these variables:

- **ReceiverTenantId** - the Azure Tenant ID for the user/service that is receiving the shared data.
- **ReceiverPurviewAccountName** - the name of the Microsoft Purview account where the data will be received.
- **ReceiverStorageKind** - either BlobAccount or AdlsGen2Account.
- **ReceiverStorageResourceId** - the resource ID for the storage account where the data has been shared.
- **ReAttachStorageResourceId** - the resource ID for the storage account where the data will be received.
- **ReceiverStorageContainer** - the name of the container where the shared data will be stored.
- **ReceiverTargetFolderName** - the folder path to where the shared data will be stored.
- **ReceiverTargetMountPath** - the mount path you'd like to use to store your data in the folder.
- **ReceivedShareDisplayName** - The display name for your received share.
- **UseServiceTokenCredentials** - (optional) If you want to use a service principal to receive the shares, set this to **true**.
- **ReceiverClientId** - (optional) If using a service principal to receive the shares, this is the Application (client) ID for the service principal.
- **ReceiverClientSecret** - (optional) If using a service principal to receive the shares, add your client secret/authentication key.
- **ReceiverPurviewEndPoint** - If you're using If you're using the classic Microsoft purview experience, use `$"https://{ReceiverPurviewAccountName}.purview.azure.com"`; If you're using the new Microsoft Purview experience, use `https://api.purview-service.microsoft.com/`

C#

```
using Azure;
using Azure.Analytics.Purview.Sharing;
using Azure.Core;
using Azure.Identity;
using System.ComponentModel;
using System.Text.Json;

public static class PurviewDataSharingQuickStart
{
    // [REQUIRED INPUTS] Set To Actual Values.
    private static string ReceiverTenantId = "<Receiver Indentity's Tenant
ID>";
    private static string ReceiverPurviewAccountName = "<Receiver Purview
Account Name>";
```

```
    private static string ReceiverStorageKind = "<Receiver Storage Account
Kind (BlobAccount / AdlsGen2Account)>";
    private static string ReceiverStorageResourceId = "<Storage Account
Resource Id for the account where the share is currently attached.>";
    private static string ReAttachStorageResourceId = "<Storage Account
Resource Id For Reattaching Received Share>";
    private static string ReceiverStorageContainer = "<Container Name To
Receive Data Under>";
    private static string ReceiverTargetFolderName = "<Folder Name to
Received Data Under>";
    private static string ReceiverTargetMountPath = "<Mount Path to Received
Data Under>";

    private static string ReceivedShareDisplayName = "<Display name of your
received share>";

    //Use if using a service principal to update the share
    private static bool UseServiceTokenCredentials = false;
    private static string ReceiverClientId = "<Receiver Caller Application
(Client) Id>";
    private static string ReceiverClientSecret = "<Receiver Caller
Application (Client) Secret>";

    // General Configs
    private static string ReceiverPurviewEndPoint =
$"https://{ReceiverPurviewAccountName}.purview.azure.com";
    private static string ReceiverShareEndPoint = $"
{ReceiverPurviewEndPoint}/share";

    private static async Task Main(string[] args)
    {
        try
        {

            Console.WriteLine($"{DateTime.Now.ToShortTimeString()}:
UpdateReceivedShare - START");
            await Receiver_UpdateReceivedShare();
            Console.WriteLine($"{DateTime.Now.ToShortTimeString()}:
UpdateReceivedShare - FINISH");
        }
        catch (Exception ex)
        {
            Console.ForegroundColor = ConsoleColor.Red;
            Console.WriteLine(ex);
            Console.ForegroundColor = Console.ForegroundColor;
        }
    }

    private static async Task<BinaryData> Receiver_UpdateReceivedShare()
    {

        TokenCredential receiverCredentials = UseServiceTokenCredentials
            ? new ClientSecretCredential(ReceiverTenantId, ReceiverClientId,
ReceiverClientSecret)
            : new DefaultAzureCredential(new DefaultAzureCredentialOptions {
```

```
AuthorityHost = new Uri("https://login.windows.net"), TenantId =
ReceiverTenantId });

        ReceivedSharesClient? receivedSharesClient = new
ReceivedSharesClient(ReceiverShareEndPoint, receiverCredentials);

        if (receivedSharesClient == null)
        {
            throw new InvalidEnumArgumentException("Invalid Received Shares
Client.");
        }

        var attachedReceivedShareData = new
        {
            shareKind = "InPlace",
            properties = new
            {
                displayName = ReceivedShareDisplayName,
                sink = new
                {
                    storeKind = ReceiverStorageKind,
                    properties = new
                    {
                        containerName = ReceiverStorageContainer,
                        folder = ReceiverTargetFolderName,
                        mountPath = ReceiverTargetMountPath
                    },
                    storeReference = new
                    {
                        referenceName = ReAttachStorageResourceId,
                        type = "ArmResourceReference"
                    }
                }
            }
        };

        var allReceivedShares = await
receivedSharesClient.GetAllAttachedReceivedSharesAsync(ReceiverStorageResour
ceId).ToResultList();

        var myReceivedShare = allReceivedShares.First(recShareDoc =>
        {
            var doc = JsonDocument.Parse(recShareDoc).RootElement;
            var props = doc.GetProperty("properties");
            return props.GetProperty("displayName").ToString() ==
ReceivedShareDisplayName;
        });

        Console.ForegroundColor = ConsoleColor.Green;
        Console.WriteLine("My Received Share Id: " +
JsonDocument.Parse(myReceivedShare).RootElement.GetProperty("id").ToString()
);
        Console.ForegroundColor = Console.ForegroundColor;
```

```
        var ReceivedShareId =
JsonDocument.Parse(myReceivedShare).RootElement.GetProperty("id").ToString()
;

        var receivedShare = await
receivedSharesClient.CreateOrReplaceReceivedShareAsync(WaitUntil.Completed,
ReceivedShareId, RequestContent.Create(attachedReceivedShareData));

        Console.ForegroundColor = ConsoleColor.Green;
        Console.WriteLine(receivedShare.Value);
        Console.ForegroundColor = Console.ForegroundColor;

        return receivedShare.Value;
    }

    public static async Task<List<T>> ToResultList<T>(this AsyncPageable<T>
asyncPageable)
    {
        List<T> list = new List<T>();

        await foreach (T item in asyncPageable)
        {
            list.Add(item);
        }

        return list;
    }
}
```

Delete received share

This script deletes a received share. To use it, be sure to fill out these variables:

- **ReceiverTenantId** - the Azure Tenant ID for the user/service that is receiving the shared data.
- **ReceiverPurviewAccountName** - the name of the Microsoft Purview account where the data will be received.
- **ReceivedShareDisplayName** - The display name for your received share.
- **ReceiverStorageResourceId** - the resource ID for the storage account where the data has been shared.
- **UseServiceTokenCredentials** - (optional) If you want to use a service principal to receive the shares, set this to **true**.
- **ReceiverClientId** - (optional) If using a service principal to receive the shares, this is the Application (client) ID for the service principal.
- **ReceiverClientSecret** - (optional) If using a service principal to receive the shares, add your client secret/authentication key.

- **ReceiverPurviewEndPoint** - If you're using If you're using the classic Microsoft purview experience, use `$"https://{ReceiverPurviewAccountName}.purview.azure.com"`; If you're using the new Microsoft Purview experience, use `https://api.purview-service.microsoft.com/`

C#

```
using Azure;
using Azure.Analytics.Purview.Sharing;
using Azure.Core;
using Azure.Identity;
using System.Text.Json;

public static class PurviewDataSharingQuickStart
{
    // [REQUIRED INPUTS] Set To Actual Values.
    private static string ReceiverTenantId = "<Receiver Indentity's Tenant ID>";
    private static string ReceiverPurviewAccountName = "<Receiver Purview Account Name>";

    private static string ReceivedShareDisplayName = "<Display name of your received share>";

    private static string ReceiverStorageResourceId = "<Storage Account Resource Id for the account where the share is currently attached.>";

    //Use if using a service principal to delete share.
    private static bool UseServiceTokenCredentials = false;
    private static string ReceiverClientId = "<Receiver Caller Application (Client) Id>";
    private static string ReceiverClientSecret = "<Receiver Caller Application (Client) Secret>";

    // General Configs
    private static string ReceiverPurviewEndPoint = $"https://{ReceiverPurviewAccountName}.purview.azure.com";
    private static string ReceiverShareEndPoint = $"{ReceiverPurviewEndPoint}/share";

    private static async Task Main(string[] args)
    {
        try
        {
            TokenCredential receiverCredentials = UseServiceTokenCredentials
            ? new ClientSecretCredential(ReceiverTenantId, ReceiverClientId, ReceiverClientSecret)
            : new DefaultAzureCredential(new DefaultAzureCredentialOptions { AuthorityHost = new Uri("https://login.windows.net"), TenantId = ReceiverTenantId });
```

```
            ReceivedSharesClient? receivedSharesClient = new
ReceivedSharesClient(ReceiverShareEndPoint, receiverCredentials);

            var allReceivedShares = await
receivedSharesClient.GetAllAttachedReceivedSharesAsync(ReceiverStorageResour
ceId).ToResultList();

            var myReceivedShare = allReceivedShares.First(recShareDoc =>
            {
                var doc = JsonDocument.Parse(recShareDoc).RootElement;
                var props = doc.GetProperty("properties");
                return props.GetProperty("displayName").ToString() ==
ReceivedShareDisplayName;
            });

            Console.ForegroundColor = ConsoleColor.Green;
            Console.WriteLine("My Received Share Id: " +
JsonDocument.Parse(myReceivedShare).RootElement.GetProperty("id").ToString()
);
            Console.ForegroundColor = Console.ForegroundColor;

            var ReceivedShareId =
JsonDocument.Parse(myReceivedShare).RootElement.GetProperty("id").ToString()
;

            await
receivedSharesClient.DeleteReceivedShareAsync(WaitUntil.Completed,
ReceivedShareId);

            Console.ForegroundColor = ConsoleColor.Green;
            Console.WriteLine("Delete Complete");
            Console.ForegroundColor = Console.ForegroundColor;
        }
        catch (Exception ex)
        {
            Console.ForegroundColor = ConsoleColor.Red;
            Console.WriteLine(ex);
            Console.ForegroundColor = Console.ForegroundColor;
        }
    }
    public static async Task<List<T>> ToResultList<T>(this AsyncPageable<T>
asyncPageable)
    {
        List<T> list = new List<T>();

        await foreach (T item in asyncPageable)
        {
            list.Add(item);
        }

        return list;
    }
}
```

Clean up resources

To clean up the resources created for the quick start, use the following guidelines:

1. Within Microsoft Purview portal, delete the sent share.
2. Also delete your received share.
3. Once the shares are successfully deleted, delete the target container and folder Microsoft Purview created in your target storage account when you received shared data.

Next steps

- FAQ for data sharing
- How to share data
- How to receive share
- REST API reference

Feedback

Was this page helpful? Yes No

FAQ: Azure Storage in-place data share with Microsoft Purview Data Sharing (preview)

Article • 10/28/2024

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

Here are some frequently asked questions for Microsoft Purview Data Sharing.

Is data sharing going to be retired?

Yes, Microsoft Purview Data Sharing will be retired in September 2025. To share data, use Microsoft Fabric external data sharing.

Is data sharing available in the new Microsoft Purview Data Catalog?

No, data sharing is accessed through the classic Microsoft Purview governance portal: https://web.purview.azure.com

What are the key terms related to data sharing?

- **Data Provider** - Organization that shares data.
- **Data Consumer** - Organization that receives shared data from a data provider.
- **Asset** - For storage in-place sharing, an asset is a storage account, and the list of files and folders you want to share from the storage account.
- **Share** - A share is a set of data that can be shared from provider to consumer. It's a set of assets. You can have one asset with files/folders from one storage account, and another asset with files/folders from a different storage account.
- **Collection** - A collection is a tool Microsoft Purview uses to group assets, sources, shares, and other artifacts into a hierarchy for discoverability and to manage access control. A root collection is created automatically when you create your Microsoft

Purview account and you're granted all the roles to the root collection. You can use the root collection (default) or create child collections for data sharing.

- **Recipient** - A recipient is a user or service principal to which the share is sent.

Can I use the API or SDK for storage in-place sharing?

Yes, you can use REST API or .NET SDK for programmatic experience to share data.

We have a guide for getting started with the .NET SDK.

What are the roles and permissions required to share data or receive shares?

Expand table

Operations	Roles and Permissions
Data provider: create share, add asset and recipients, revoke access	**Microsoft Purview collection role**: minimum of Data Reader to use the Microsoft Purview portal experience, none to use API or SDK
	Storage account role checked when adding and updating asset: Owner or Storage Blob Data Owner
	Storage account permissions checked when adding and updating asset: Microsoft.Authorization/roleAssignments/write OR Microsoft.Storage/storageAccounts/blobServices/containers/blobs/modifyPermissions/
Data consumer: Receive share, attach share, delete share	**Microsoft Purview collection role**: minimum of Data Reader to use the Microsoft Purview portal experience, none to use API or SDK
	Storage account role checked when attaching share: Contributor OR Owner OR Storage Blob Data Contributor OR Storage Blob Data Owner

Operations	Roles and Permissions
	Storage account permissions checked when attaching share: Microsoft.Storage/storageAccounts/write OR Microsoft.Storage/storageAccounts/blobServices/containers/write
Data consumer: Access shared data	No share-specific role required. You can access shared data with regular storage account permission just like any other data. Data consumer's ability to apply ACLs for shared data is currently not supported.

How can I share data from containers?

When adding assets, you can select the container(s) that you would like to share.

Can I share data in-place with storage account in a different Azure region?

Cross-region in-place data sharing isn't currently supported for storage account. Data provider and data consumer's storage accounts need to be in the same Azure region.

Is there support for read-write shares?

Storage in-place sharing supports read-only shares. Data consumer can't write to the shared data.

To share data back to the data provider, the data consumer can create a share and share with the data provider.

Can I access shared data from analytics tools like Azure Synapse?

You can access shared data from storage clients like Azure Synapse Analytics Spark and Databricks. You won't be able to access shared data using Azure Data Factory, Power BI, or AzCopy.

Does the recipient of the share need to be a user's email address or can I share data with an

application?

Through the UI, you can share data with recipient's Azure sign-in email or using service principal's object ID and tenant ID.

Through API and SDK, you also send invitation to object ID of a user principal or service principal. Also, you can optionally specify a tenant ID that you want the share to be received into.

Is the recipient accepting the share only for themselves?

When the recipient attaches the share to a target storage account, any user or application that has access to the target storage account will be able to access shared data.

If the recipient leaves the organization, what happens to the received share?

Once the received share is accepted and attached to a target storage account, any users with appropriate permissions to the target storage account can continue to access the shared data even after the recipient has left the organization.

Once the received share is accepted, any user with data reader permission to the Microsoft Purview collection that the share is received into can view and update the received share.

How do I request an increase in limits for the number of shares?

Data provider's source storage account can support up to 20 targets, and data consumer's target storage account can support up to 100 sources. To request a limit increase, contact support.

How do I troubleshoot data sharing issues?

To troubleshoot issues with sharing data, refer to the troubleshooting section of the how to share data article. To troubleshoot issues with receiving share, refer to the

troubleshooting section of the how to receive share article.

Is there support for Private endpoints, VNET and IP restrictions?

Private endpoints, VNET, and IP restrictions are supported for data share for storage. Blob should be chosen as the target subresource when creating a private endpoint for storage accounts.

Related content

- Data sharing quickstart
- How to Share data
- How to receive a share
- REST API reference

Feedback

Was this page helpful? Yes No

Understand the Microsoft Purview Data Estate Insights application

Article • 02/24/2025

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

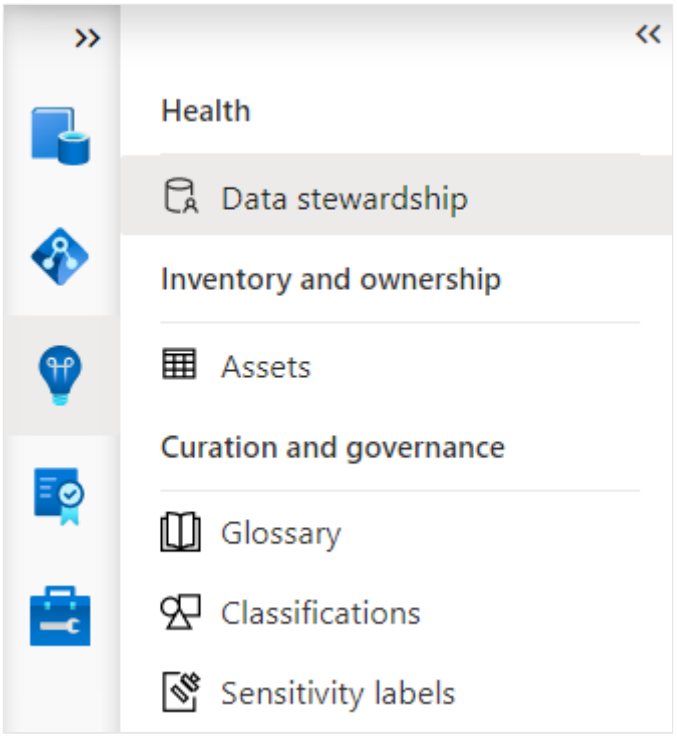
This article provides an overview of the Data Estate Insights application in Microsoft Purview.

The Data Estate Insights application is purpose-built for governance stakeholders, primarily for roles focused on data management, compliance, and data use: like a Chief Data Officer. The application provides actionable insights into the organization's data estate, catalog usage, adoption, and processes. As organizations scan and populate their Microsoft Purview Data Map, the Data Estate Insights application automatically extracts valuable governance gaps and highlights them in its top metrics. Then it also provides drill-down experience that enables all stakeholders, such as data owners and data stewards, to take appropriate action to close the gaps.

All the reports within the Data Estate Insights application are automatically generated and populated, so governance stakeholders can focus on the information itself, rather than building the reports.

The dashboards and reports available within Microsoft Purview Data Estate Insights are categorized in three sections:

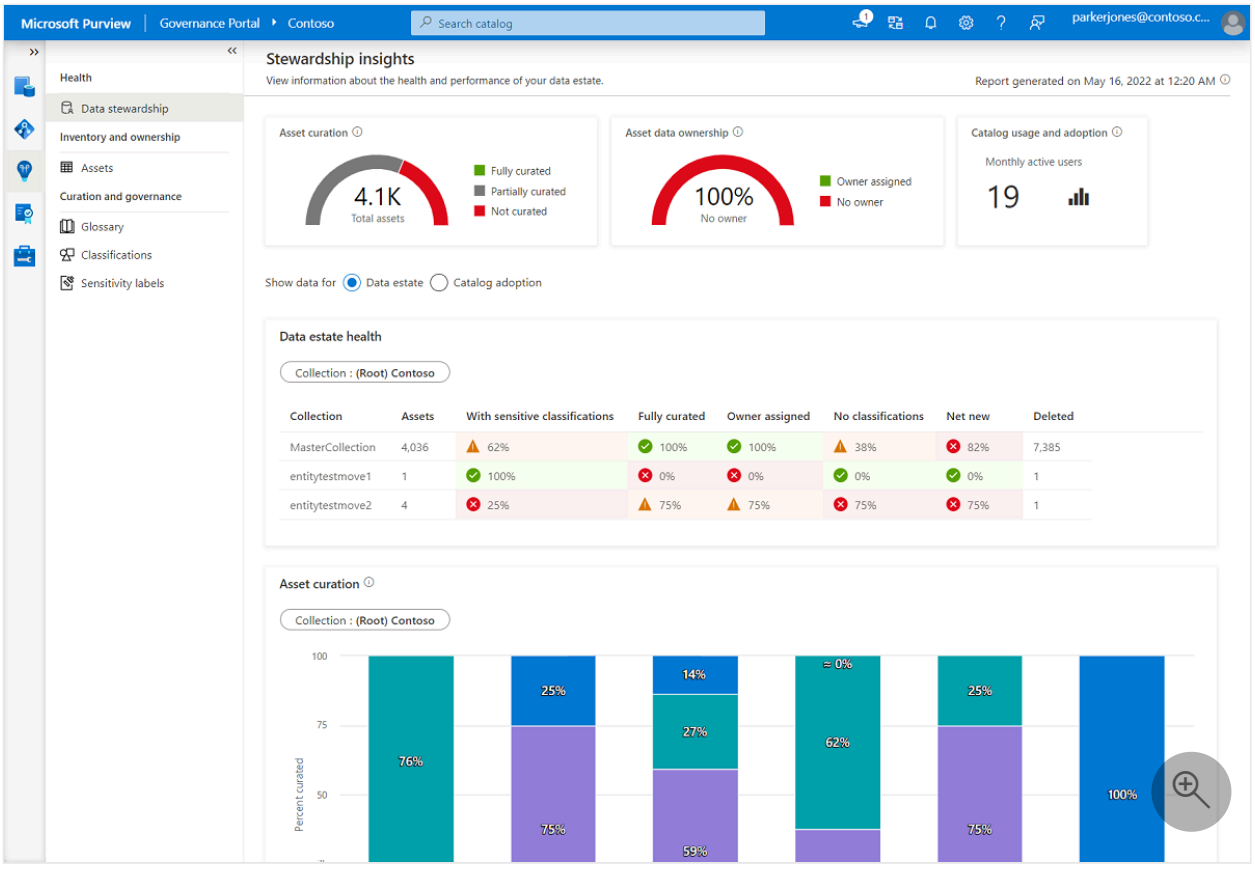
- Health
- Inventory and Ownership
- Curation and governance



> **Note**
>
> If you're using the new Microsoft Purview portal, these same reports are still available, but are accessed inside the **Data Catalog**.

Health

Data, governance, and quality focused users like chief data officers and data stewards can start at the health dashboards to understand the current health status of their data estate, current return on investment on their catalog, and begin to address any outstanding issues.



Data stewardship

The data stewardship dashboard highlights key performing indicators that the governance stakeholders need to focus on, to attain a clean and governance-ready data estate. Information like asset curation rates, data ownership rates, and classification rates are calculated out of the box and trended over time.

Management-focused users, like a Chief Data Officer, can also get a high-level understanding of weekly and monthly active users of their catalog, and information about how the catalog is being used. Is the catalog being adopted across their organization, as better adoption leads to better overall governance penetration in the organization?

For more information about these dashboards, see the data Stewardship documentation.

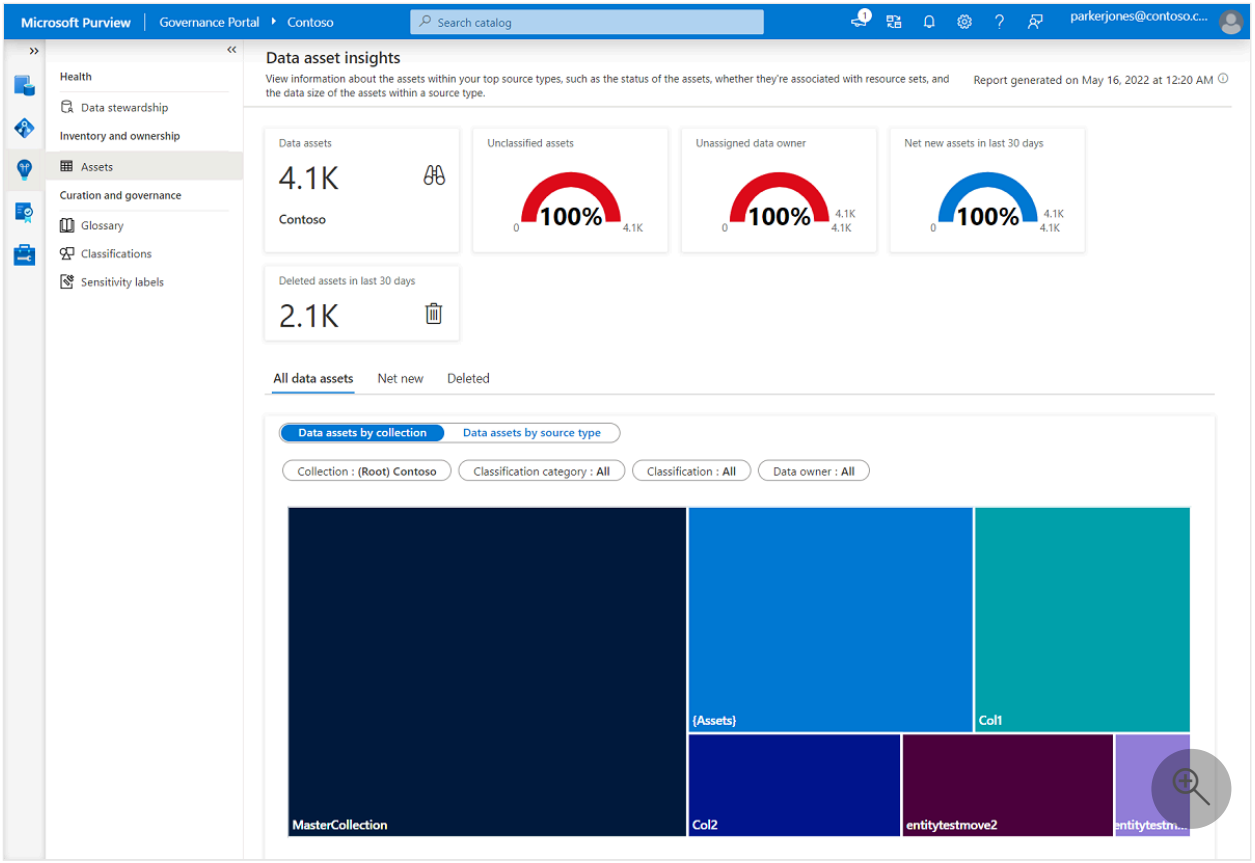
Catalog adoption

The catalog adoption dashboard highlights active users, searches, viewed assets, and top searched keywords. This report helps you understand how and if your data catalog is being used, so you can see the impact it's having on data usage and discoverability.

For more information about these dashboards, see the catalog adoption documentation.

Inventory and ownership

This area focuses on summarizing data estate inventory for data quality and management focused users, like data stewards and data curators. These dashboards provide key metrics and overviews to give users the ability to find and resolve gaps in their assets, all from within the data estate insights application.



Assets

This report provides a summary of your data estate and its distribution by collection and source type. You can also view new assets, deleted assets, updated assets, and stale assets from the last 30 days.

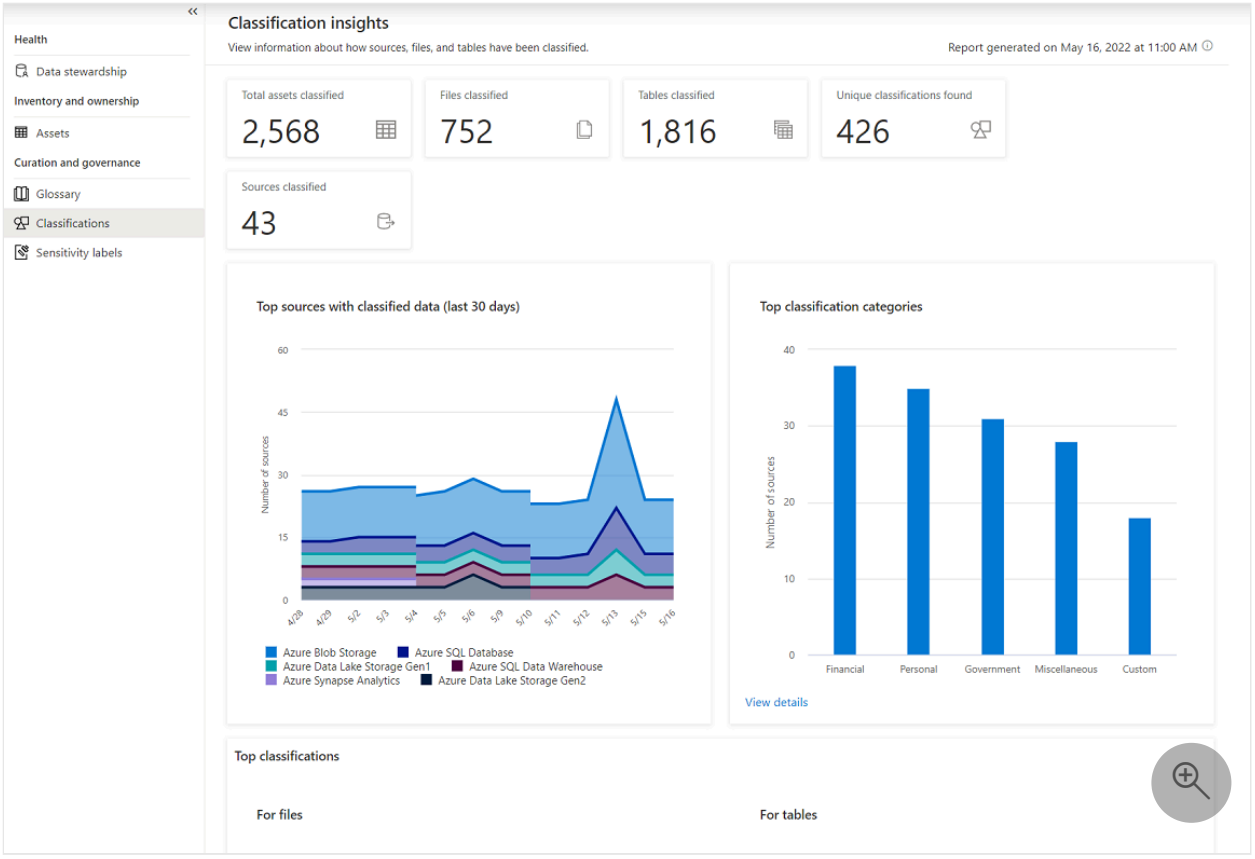
Explore your data by classification, investigate why assets didn't get classified, and see how many assets exist without a data owner assigned. To take action, the report provides a “View Detail” button to view and edit the specific assets that need treatment.

You can also view data asset trends by asset count and data size, as we record this metadata during the data map scanning process.

For more information, see the asset insights documentation.

Curation and governance

This area focuses on giving a summary of how curated your assets are by several curation contexts. Currently we focus on showcasing assets with glossary, classification, and sensitivity labels.



Glossary

Data, governance, and quality focused users like chief data officers and data stewards a status check on their business glossary. Data maintenance and collection focused users like Data Stewards can view this report to understand distribution of glossary terms by status, learn how many glossary terms are attached to assets, and how many aren't yet attached to any asset. Business users can also learn about completeness of their glossary terms.

This report summarizes top items that use needs to focus on to create a complete and usable glossary for their organization. Users can also navigate into the "Glossary" experience from "Glossary Insights" experience, to make changes on a specific glossary term.

For more information, see the glossary insights in Microsoft Purview documentation.

Classifications

This report provides details about where classified data is located, the classifications found during a scan, and a drill-down to the classified files themselves. It enables data quality and data security focused users like data stewards, data curators, and security administrators to understand the types of information found in their organization's data estate.

In Microsoft Purview, classifications are similar to subject tags, and are used to mark and identify content of a specific type in your data estate.

Use the classification insights report to identify content with specific classifications and understand required actions, such as adding extra security to the repositories, or moving content to a more secure location.

For more information, see the classification insights about your data from Microsoft Purview documentation.

Sensitivity Labels

This report provides details about the sensitivity labels found during a scan and a drill-down to the labeled files themselves. It enables security administrators to ensure the security of the data found in their organization's data estate by identifying where sensitive data is stored.

In Microsoft Purview, sensitivity labels are used to identify classification type categories, and the group security policies that you want to apply to each category.

Use the labeling insights report to identify the sensitivity labels found in your content and understand required actions, such as managing access to specific repositories or files.

For more information, see the sensitivity label insights about your data in Microsoft Purview documentation.

Next steps

Learn how to use Data Estate Insights with resources below:

- Learn how to use Asset insights
- Learn how to use Classification insights
- Learn how to use Glossary insights
- Learn how to use Label insights

Feedback

Was this page helpful? Yes No

Access control in Data Estate Insights within Microsoft Purview

> **Note**
>
> Microsoft Purview Data Catalog (classic), Data Health Insights (classic), and Purview Workflow (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

Like all other permissions in Microsoft Purview, Data Estate Insights access is given through collections. This article describes what permissions are needed to access Data Estate Insights in Microsoft Purview.

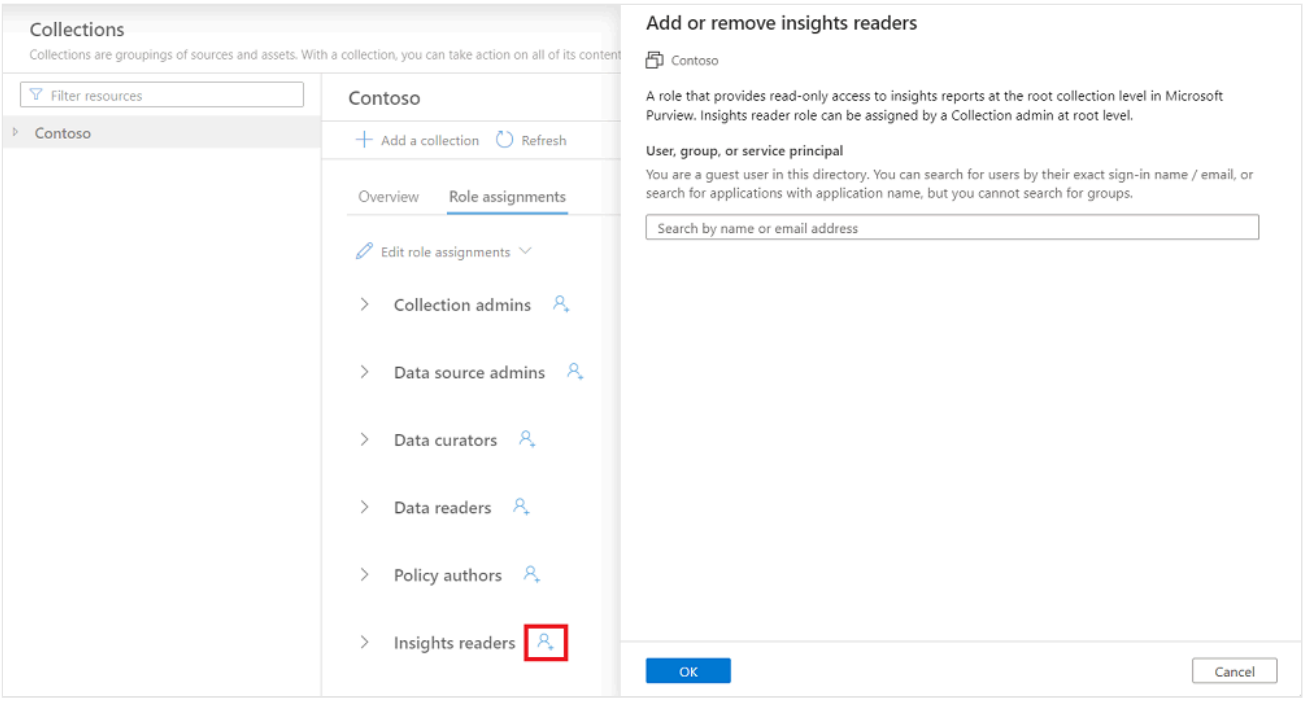
Insights reader role

The insights reader role gives users read permission to the Data Estate Insights application in Microsoft Purview. However, a user with this role will only have access to information for collections that they also have at least data reader access to.

As Data Estate Insights application gives a bird's eye view of your data estate and catalog usage from a governance and risk perspective, it's intended for users who need to manage and report on this high-level information, like a Chief Data Officer. You may not want, or need, all your data readers to have access to the Data Estate Insights dashboards.

Role assignment

- **Insights Reader** role can be assigned to any Data Map user, by the **Data Curator of the root collection**. Users assigned the Data Curator role on subcollections don't have the privilege of assigning insights reader.



- A **Data Curator** of any collection also has read permission to Data Estate Insights application. Their scope of insights will be limited to the metadata assigned to collections. In other words, a Data Curator at a subcollection will only view KPIs and aggregations on collections they have access to. A Data Curator can still view and edit assets from Data Estate Insights app, without any extra permissions.

- A **Data Reader** at any collection node, can see Data Estate Insights app on the left navigation bar, however, when they hover on the icon, they'll receive a message saying they need to contact Data Curator at root collection for access. Once a Data Reader has been assigned the Insights Reader role, they can view KPIs and aggregations based on the collections they have Data Reader permission on. A Data Reader can't edit assets or select "***Export to CSV***" from the app.

> ⓘ **Note**
>
> All roles other than Data Curator, will need explicit role assignment as **Insights Reader** to be able to click into the Data Estate Insights app.

Next steps

Learn how to use Data Estate Insights with sources below:

- Learn how to use Asset insights
- Learn how to use Data Stewardship
- Learn how to use Classification insights

- Learn how to use Glossary insights
- Learn how to use Label insights

Last updated on 10/10/2025

Disable Data Estate Insights or report refresh

> **Note**
>
> Microsoft Purview Data Catalog (classic), Data Health Insights (classic), and Purview Workflow (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

Microsoft Purview Data Estate Insights automatically aggregates metrics and creates reports about your Microsoft Purview account and your data estate. When you scan registered sources and populate your Microsoft Purview Data Map, the Data Estate Insights application automatically identifies governance gaps and highlights them in its top metrics. It also provides drill-down experience that enables all stakeholders, such as data owners and data stewards, to take appropriate action to close the gaps.

These features are optional and can be enabled or disabled at any time. This article provides the specific steps required to enable or disable Microsoft Purview Data Estate Insights features.

To reschedule without disabling, see the article to schedule Data Estate Insights reports.

> **Important**
>
> The Data Estate Insights application is **on** by default when you create a Microsoft Purview account.
>
> **State** is set to On.
>
> **Refresh Schedule** is set to a weekly refresh that begins 7 days after the account is created.
>
> As the Data Map is populated and curated, Insights App shows data in the reports. The reports are ready for consumption to anyone with Insights Reader role.

If you don't plan on using Data Estate Insights for a time, a **data curator** **on the** **root collection** can disable the Data Estate Insights in one of two ways:

- Disable the Data Estate Insights application - this will stop billing from both report generation and report consumption.
- Disable report refresh - Insights readers have access to current reports, but reports won't be refreshed. Billing will occur for report consumption but not report generation.

Steps for both methods, and for re-enablement, are below.

For more information about billing for Data Estates Insights, see our pricing guidelines.

Disable the Data Estate Insights application

> **Note**
>
> To be able to disable this application, you will need to have the **data curator role** on your account's **root collection.**

Disabling Data Estate Insights will disable the entire application, including these reports:

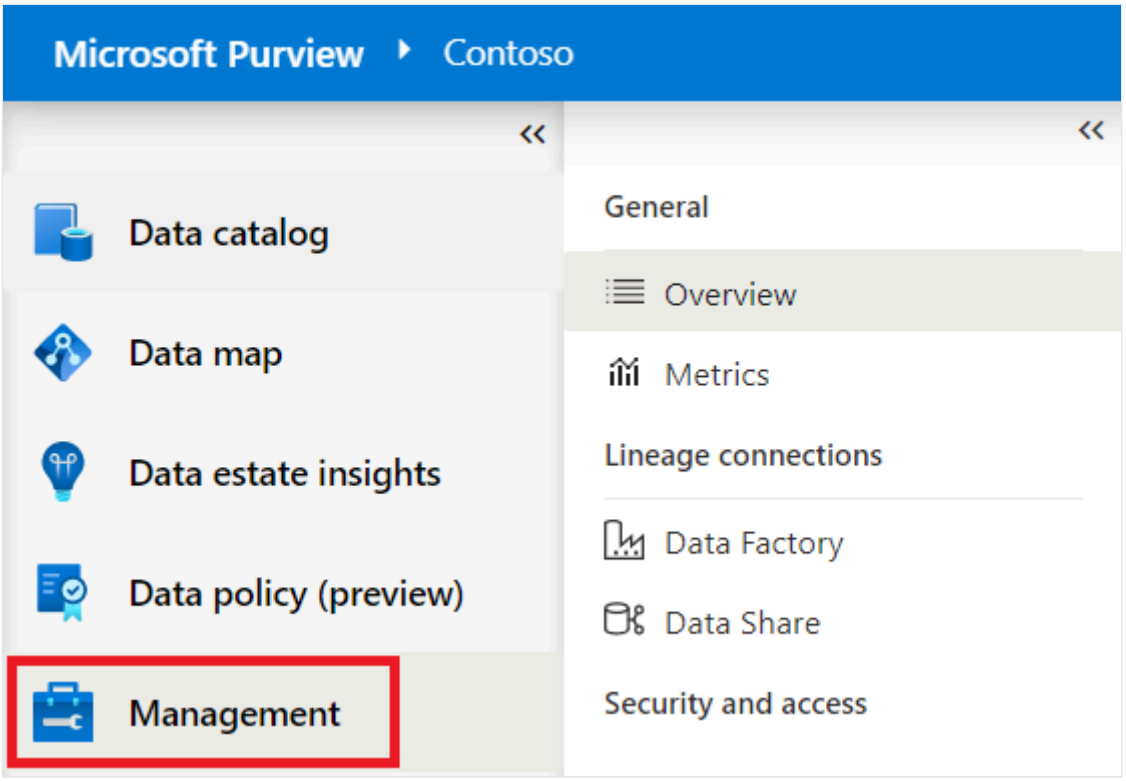
- Stewardship
- Asset
- Glossary
- Classification
- Labeling

The application icon will still show in the menu, but insights readers won't have access to reports at all, and report generation jobs will be stopped. The Microsoft Purview account won't receive any bill for Data Estate Insights.

To disable the Data Estate Insights application, a user with the data curator role at the root collection can follow these steps:

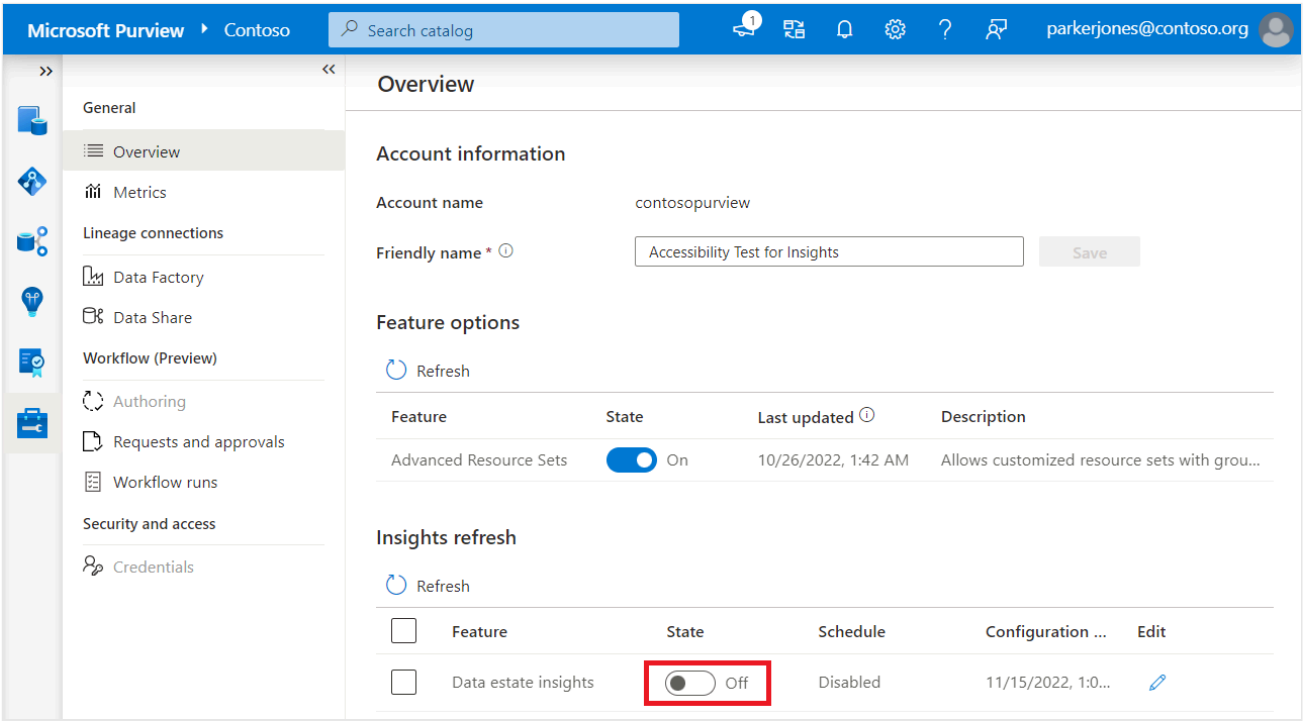
1. If you're using the Microsoft Purview portal, the schedule can't be changed; all classic reports are set to refresh weekly by default. Classic reports in the new Microsoft Purview don't impact your billing.

2. In the Microsoft Purview governance portal, go to the **Management** section.

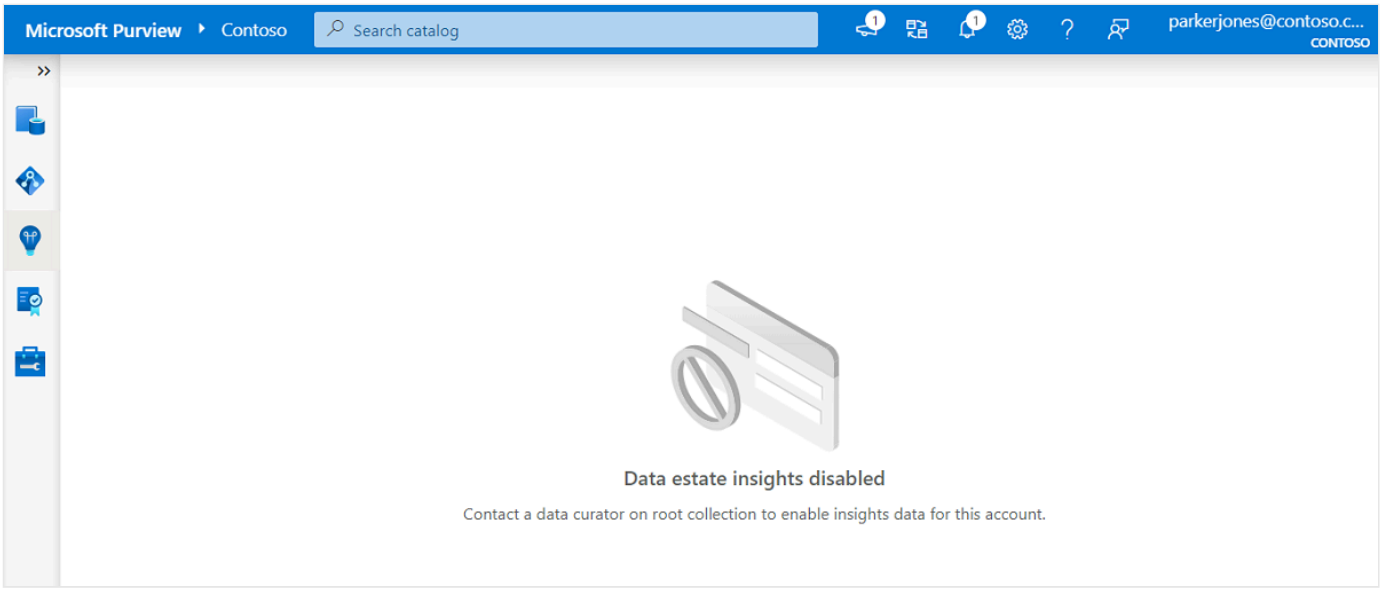


3. Then select **Overview**.

4. In the **Feature options** menu, locate Data Estate Insights, and select the **State** toggle to change it to **Off**.



Once you have disabled Data Estate Insights, the icon will still appear in the left hand menu, but users will receive a warning stating that the application has been disabled when attempting to access it.

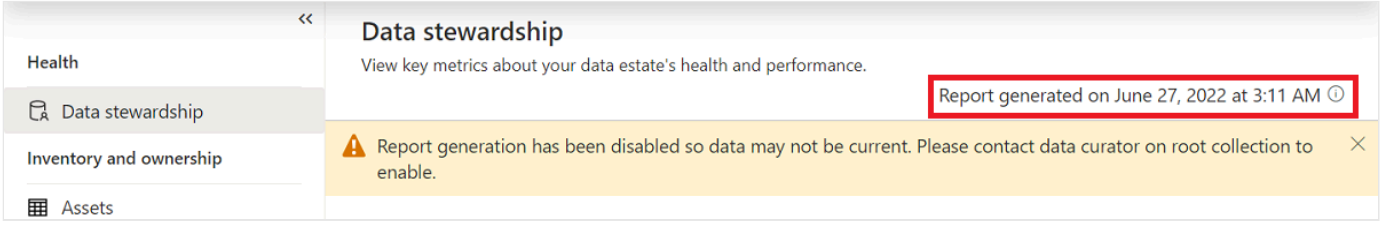


Disable report refresh

> **Note**
>
> To be able to disable or edit report refresh, you will need to have the data curator role on your account's root collection.

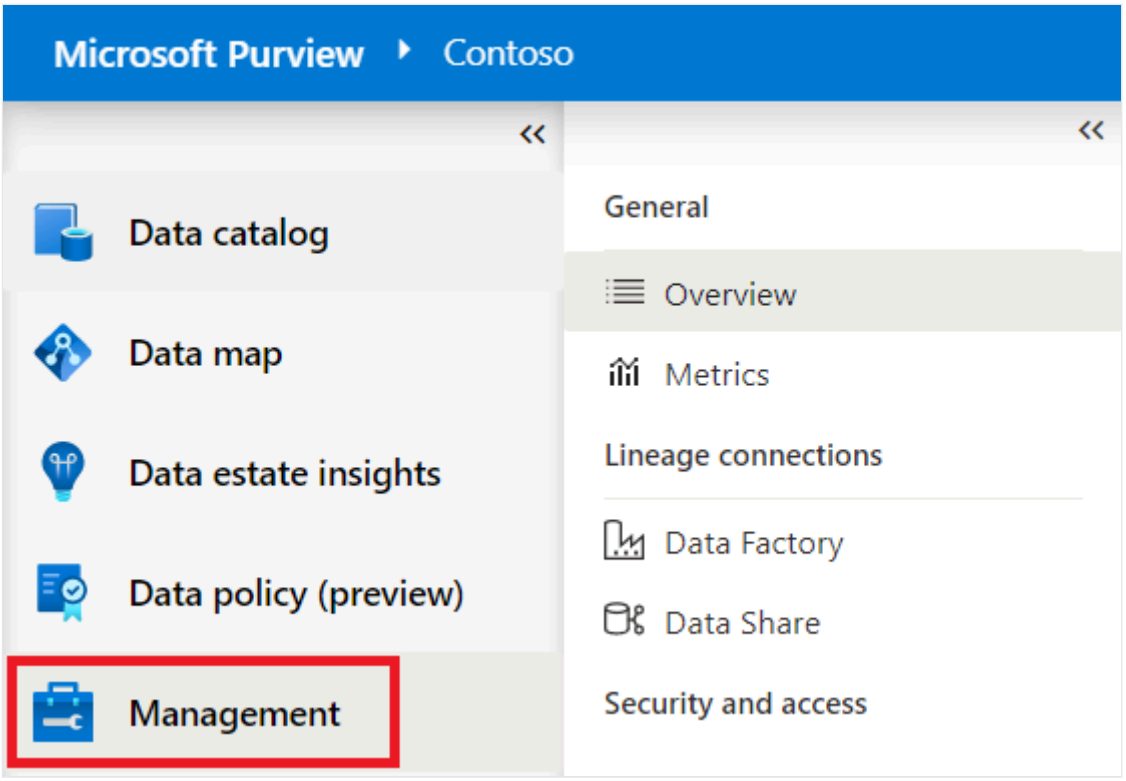
You can choose to disable report refreshes instead of disabling the entire Data Estate Insights application. When you disable report refreshes, users with the insights reader role will still be able view reports, but they'll see warning at the top of each report indicating that the data may not be current and the date of the last refresh.

Graphs that show data from the last 30 days will appear blank after 30 days while graphs showing snapshot of the data map will continue to show graphs and details.



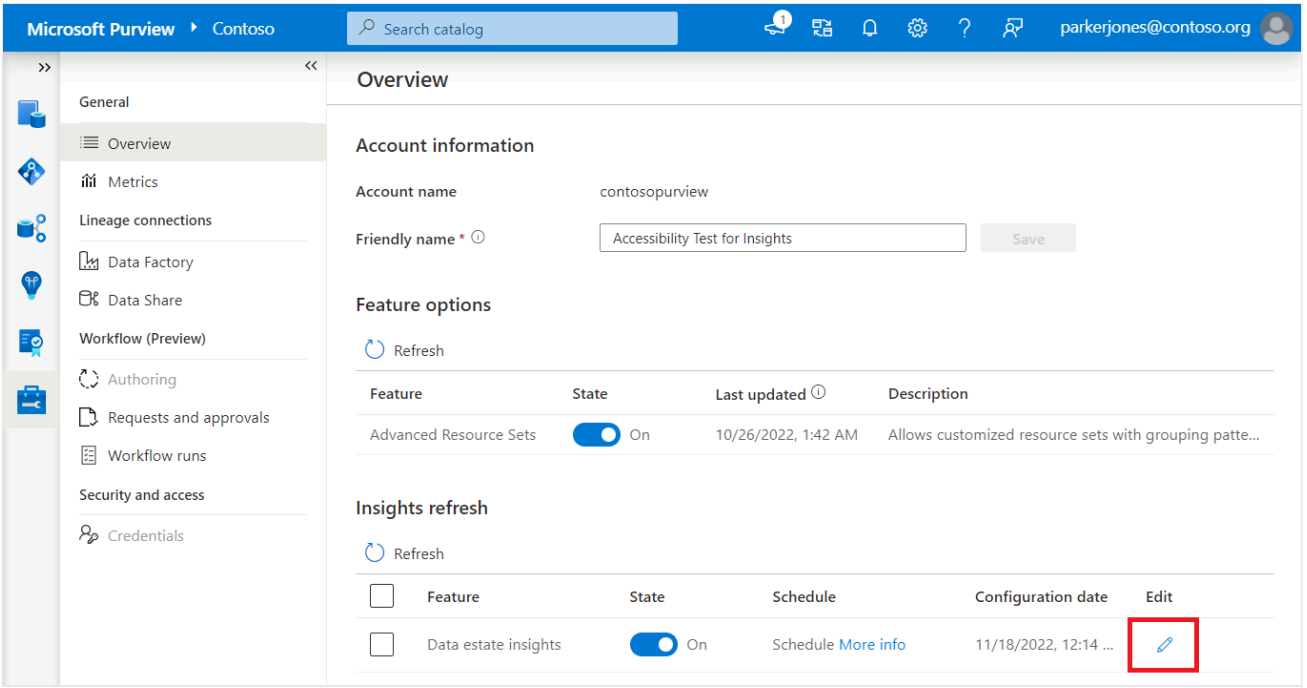
To disable the Data Estate Insights report refresh, a user with the data curator role at the root collection can follow these steps:

1. In the Microsoft Purview governance portal, go to the **Management** section.

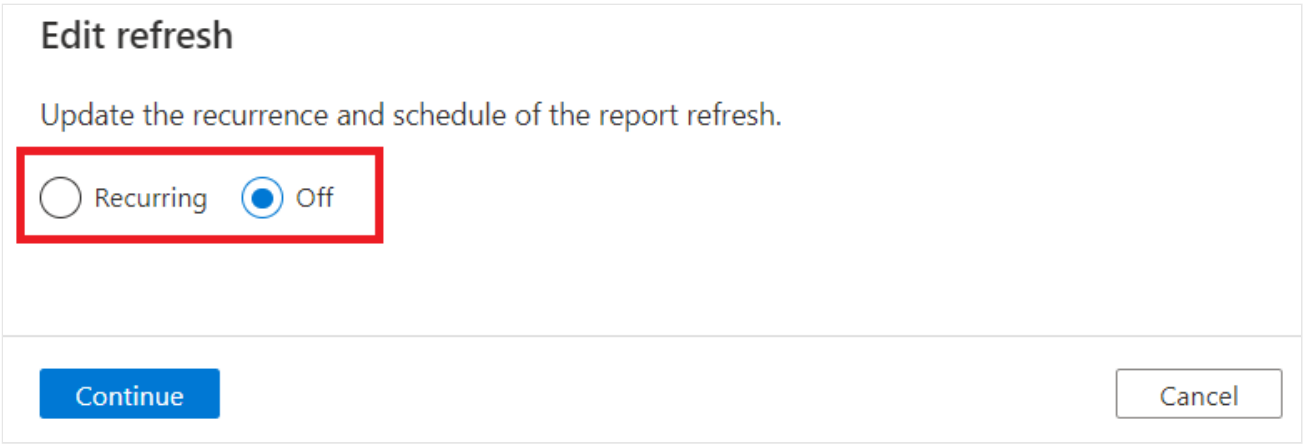


2. Then select **Overview**.

3. In the **Insights refresh** menu, locate Data Estate Insights and select the **Edit pencil**



4. Select the **Off** radio button and select **Continue**. If prompted, review your edits and select **Save.**



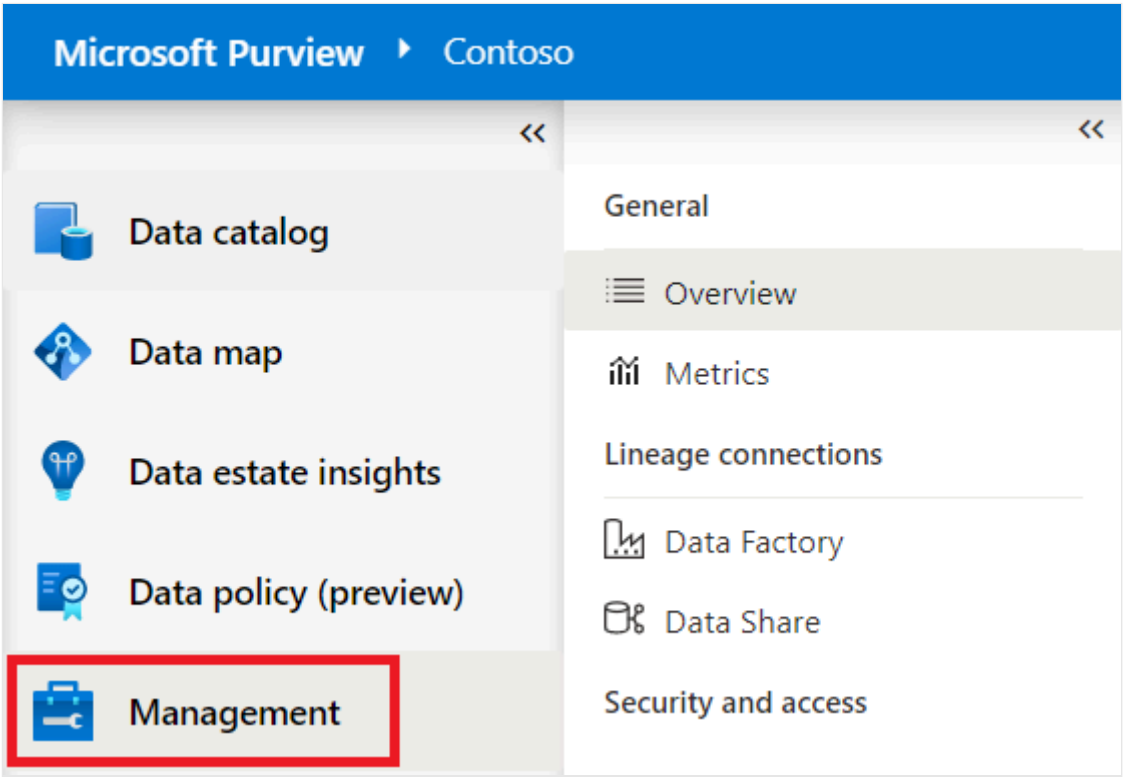
5. You can now see under the **Schedule** column, the refresh schedule reads as **Disabled**.

Re-enable Data Estate Insights and report refresh

> **Note**
>
> To enable Data Estate Insights, enable report refresh, or edit report refresh, you will need to have the data curator role on your account's root collection.

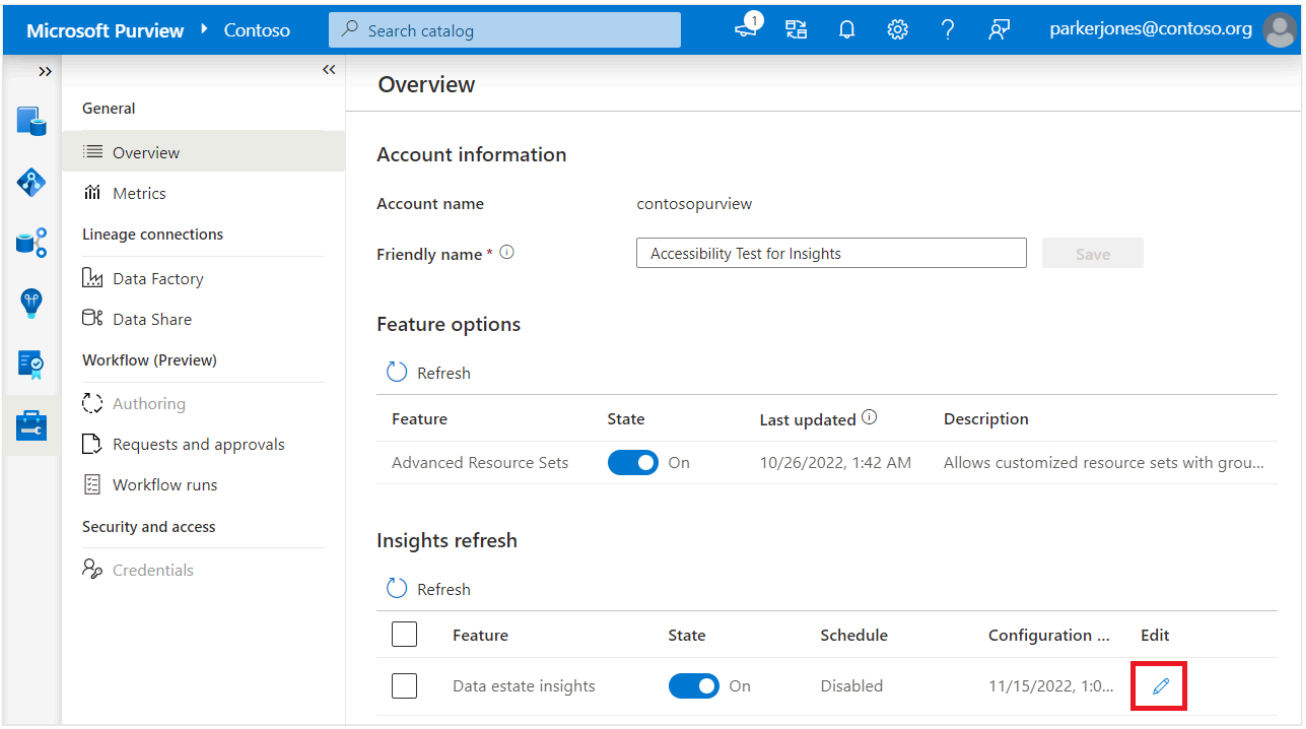
If Data Estate Insights or report refresh has been disabled in your Microsoft Purview governance portal environment, a user with the data curator role at the root collection can re-enable either at any time by following these steps:

1. In the Microsoft Purview governance portal, go to the **Management** section.



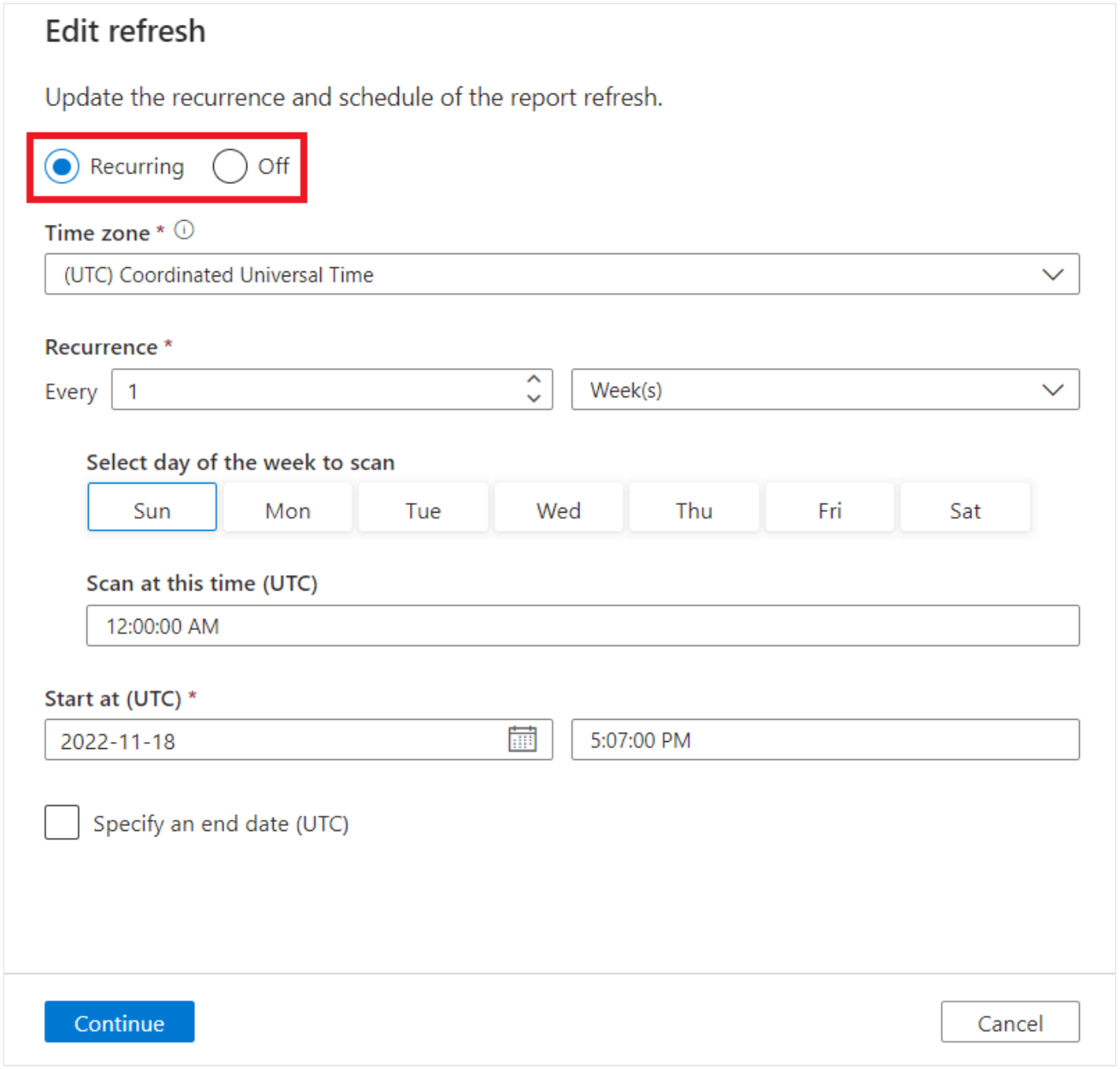
2. Then select **Overview**.

3. In the **Insights refresh** menu, locate Data Estate Insights, and select the **Edit** pencil.



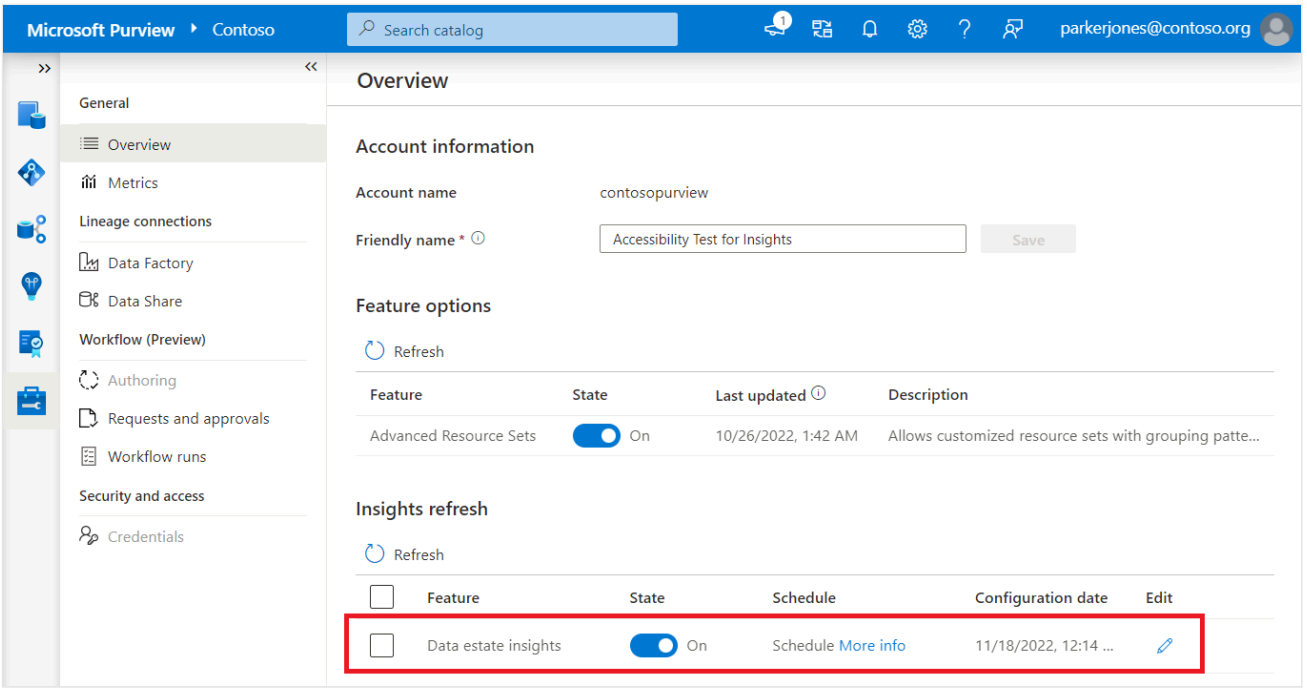
4. In the **Edit** menu, select **Recurring**.

5. Then select your time zone, set your recurrence to **Month(s)** or **Weeks(s)**, select your day and time to run, specify a start time, and optionally specify an end time. For more information about the available options, see the article to schedule Data Estate Insights reports.



6. Select **Continue**. If prompted, review the updates and select **Save**.

7. Now you can see your schedule is set. Selecting **More info** in the schedule columns will give you the recurrence details.



Next steps

- Learn how to use Asset insights
- Learn how to use Classification insights
- Learn how to use Glossary insights
- Learn how to use Label insights
- Schedule Data Estate Insights reports

Last updated on 11/07/2025

Schedule Data Estate Insights reports refresh

Article • 02/24/2025

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

Microsoft Purview Data Estate Insights automatically creates reports that help you to identify governance gaps in your data estate. However, you can set the schedule for this automatic refresh to fit your reporting patterns and needs. In this article, we'll describe how you can schedule your refresh and what options are available.

If you would like to disable Data Estate Insights for a time, you can follow our article to disable Data Estate Insights.

> **Important**
>
> The Data Estate Insights application is **on** by default when you create a Microsoft Purview account.
>
> **Refresh Schedule** is set to a weekly refresh that begins 7 days after the account is created.
>
> As the Data Map is populated and curated, Insights App shows data in the reports. The reports are ready for consumption to anyone with Insights Reader role.

Permissions needed

To update the report refresh schedule for Data Estate Insights, you need **data curator permissions** on the **root collection**.

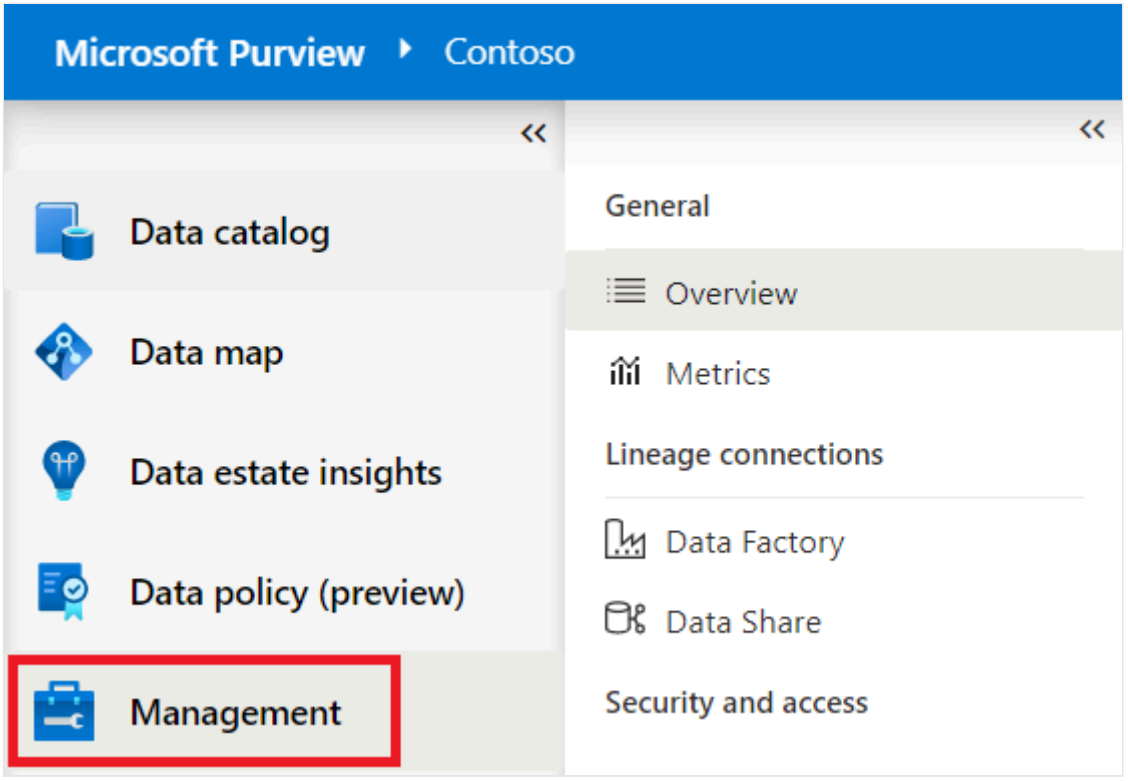
How to update the refresh schedule

> **Important**

Report refresh rate does impact your billing.

1. If you're using the Microsoft Purview portal, the schedule can't be changed; all classic reports are set to refresh weekly by default. Classic reports in the new Microsoft Purview don't impact your billing.

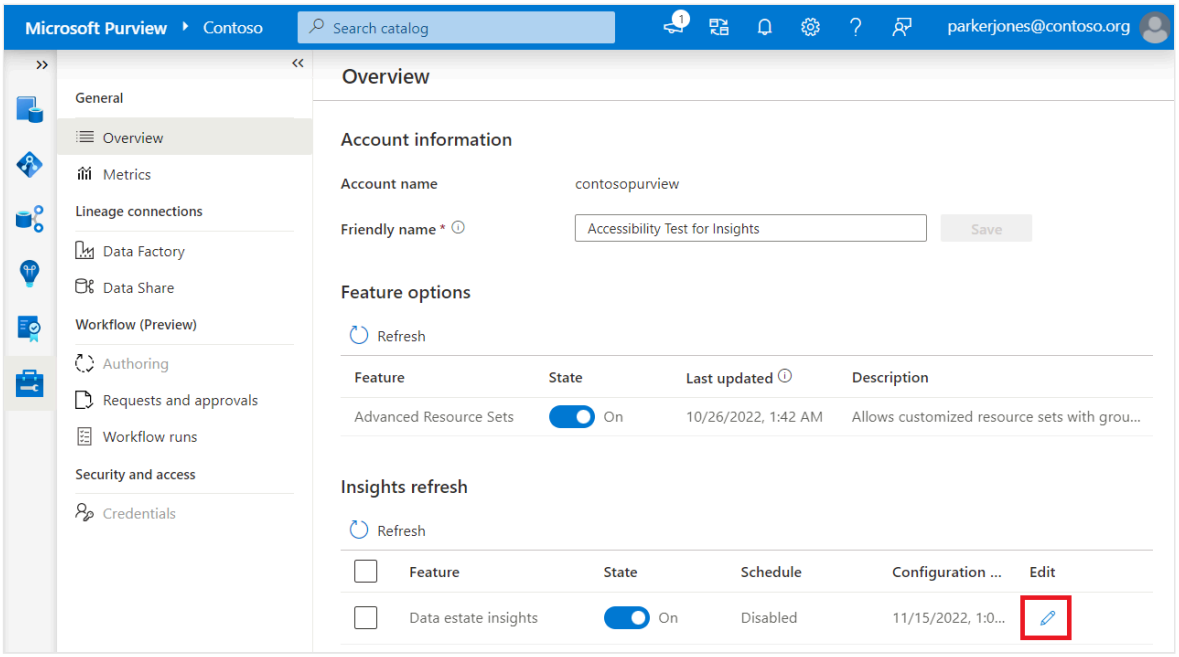
2. In the Microsoft Purview governance portal, go to the **Management** section.



3. Then select **Overview**.

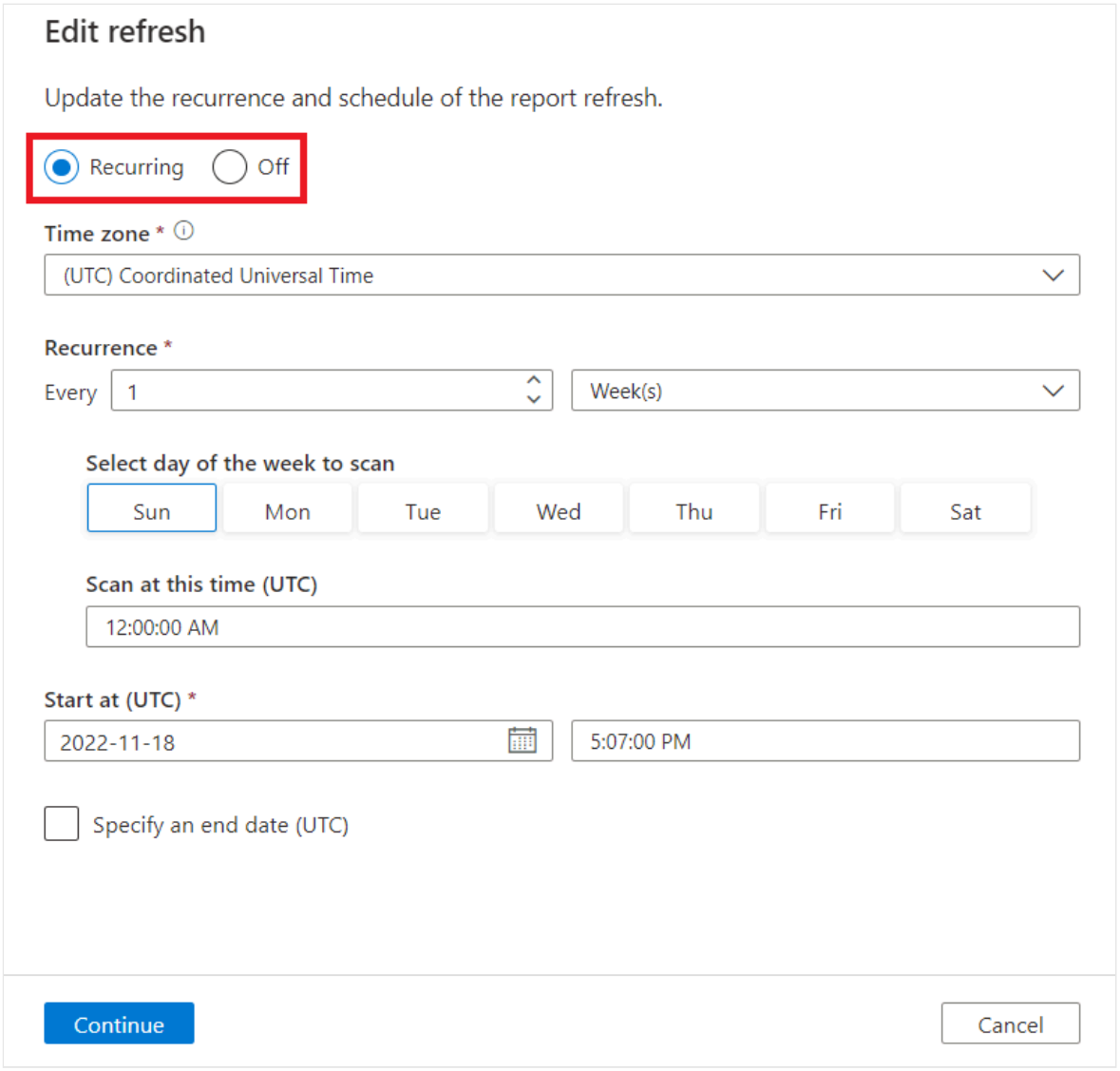
4. In the **Insights refresh** menu, locate Data Estate Insights, and select the **Edit** pencil. For more information about the insights refresh menu, see the insights refresh menu section below.

> **Note**
>
> If the State is currently set to **Off** you need to first turn it **On** before you can update the refresh schedule.



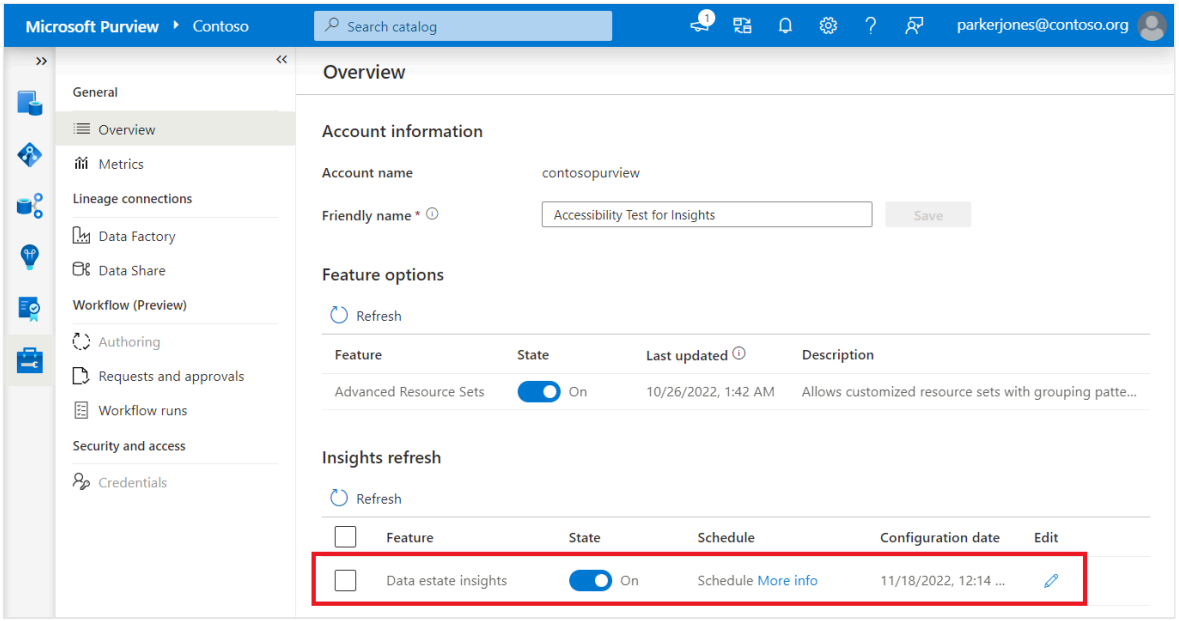
5. In the **Edit** menu, select **Recurring**.

6. Then select your time zone, set your recurrence to **Month(s)**, **Weeks(s)**, or **Days(s)**, select your day and time to run, specify a start time, and optionally specify an end time. For more information about the schedule options, see the schedule options section below.



7. Select **Continue**. If prompted, review the updates and select **Save**.

8. Now you can see your schedule is set. Selecting **More info** in the schedule columns give you the recurrence details.



Does the refresh schedule affect billing?

Yes. You're billed when Microsoft Purview Data Estate Insights checks for updates in your environment and when a report is generated.

1. A check for updates to your Microsoft Purview instance is made at your scheduled report refresh time. You'll always be billed for this check at your report refresh time.

2. If updates have been made, you'll be billed to generate your insights report. If no updates were made since the last report, you'll not be billed to generate a report.

   > ⓘ **Note**
   >
   > An asset is considered to be **updated** if it's updated in Microsoft Purview. An update in Microsoft Purview means a change in the following properties: asset name, description, schema, classifications, experts, owners, etc. New rows in a data asset and a re-scan won't result in an update to your insights reports if the metadata properties haven't changed.

For more information about billing for Data Estates Insights, see our pricing guidelines.

Schedule options

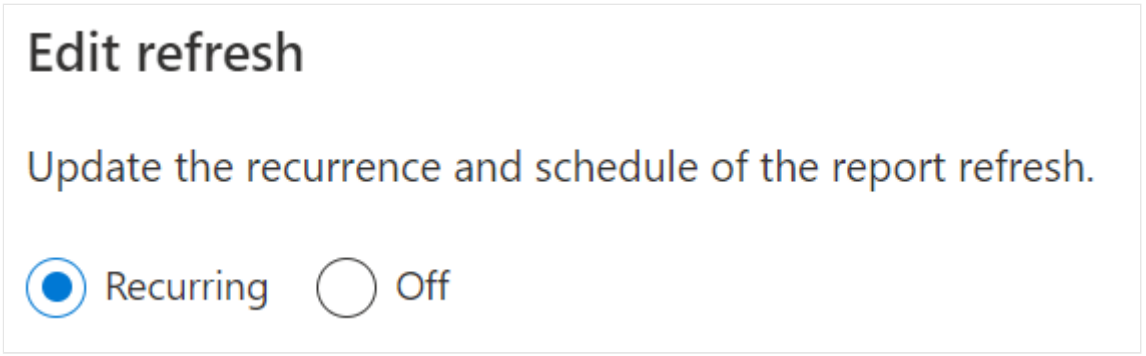
Below are listed all the schedule options currently available when you modify your refresh schedule:

- Recurring or off
- Time zone
- Recurrence
- Start at
- End date (optional)

Recurring or off

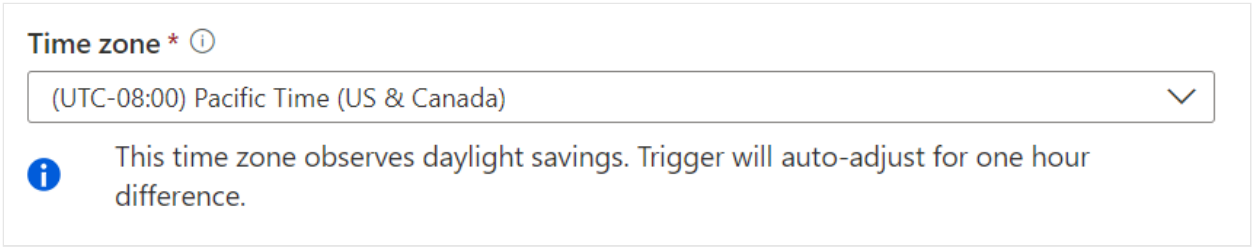
- **Recurring** - your Data Estate Insights report refreshes on the recurrence schedule you specify below, based on the time zone you indicate.

- **Off** - your Data Estate Insights reports will no longer refresh information, but the current reports remain available. For more information, see our disable Data Estate Insights article.



Time zone

Select the time zone you'd like to align your refresh schedule with. If the time zone you select observes daylight savings, your trigger auto-adjusts for the difference.

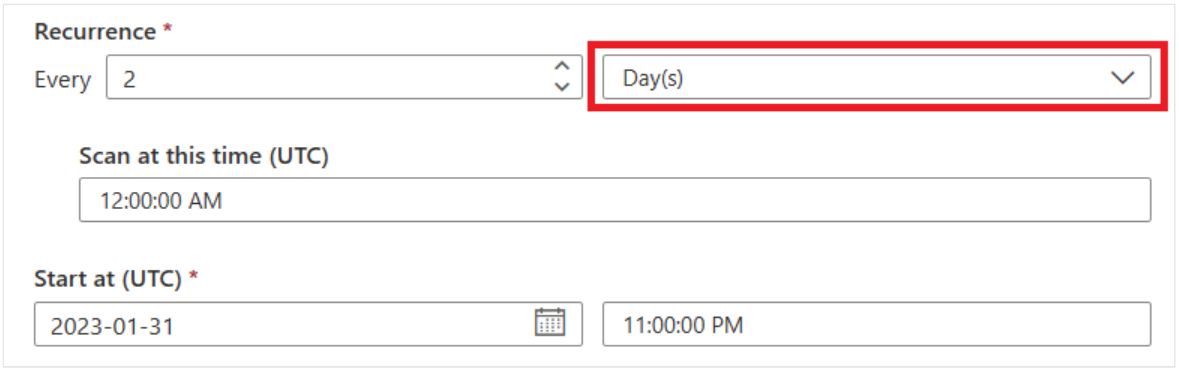


Recurrence

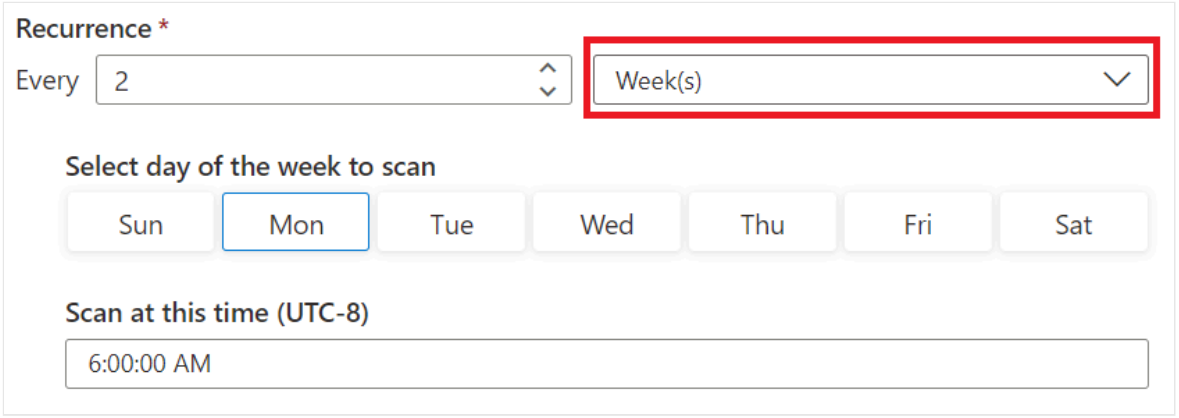
You can select a daily, weekly, or monthly refresh recurrence.

> **Important**
>
> No matter what you set your refresh recurreance to, Data Estate Insights run a check before refreshing to see if any updates have been made to the assets. If no update has been made, the reports won't refresh.
>
> An asset is considered to be **updated** only if it is updated in Microsoft Purview. An update in Microsoft Purview means a change in the following properties: asset name, description, schema, classifications, experts, owners, etc... New rows in a data asset and a re-scan won't result in an update to your insights reports if the metadata properties of the asset have not changed.

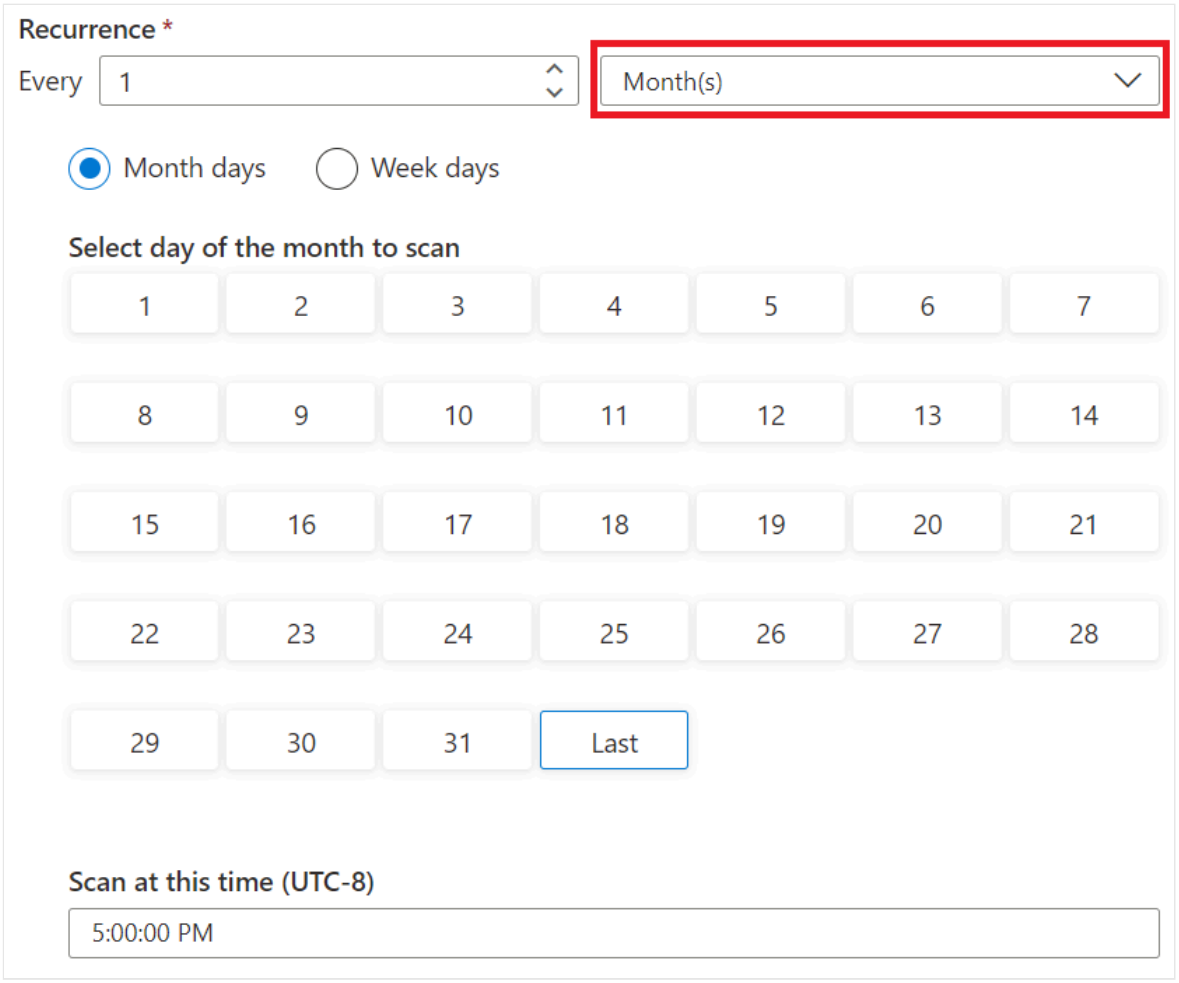
- **Daily recurrence** - schedule daily recurrence by setting recurrence to every (1-5) days, and choosing the time of day you would like to refresh the report. For example, the below report refreshes every other day at 12 AM UTC.



- **Weekly recurrence** - schedule weekly recurrence by setting recurrence to every (1-5) weeks, and choosing the day of the week and time of day you would like to refresh the report. For example, the below report refreshes on Monday every two weeks, at six AM.



- **Monthly recurrence** - schedule monthly recurrence by setting recurrence to every (1-5) months, and choosing the day of the month and time of day you would like to refresh the report. Choose **Last** to always run on the last day of the month, and **1** to always run on the first day of the month. For example, the below report refreshes on the last day of every month at five PM.

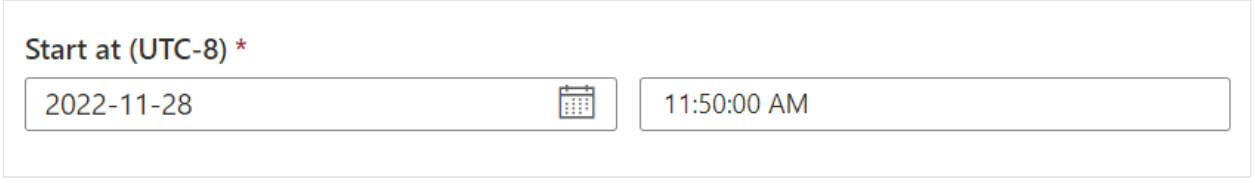


Set **Scan at this time** to the time of day you'd like to begin your refresh.

> **Tip**
>
> The time you select will be the time that the report begins its refresh. Microsoft Purview will adjust compute for the amount of data being aggregated, but **schedule your recurrence a little before you will need the report** so the report has time to fully refresh and will be ready.

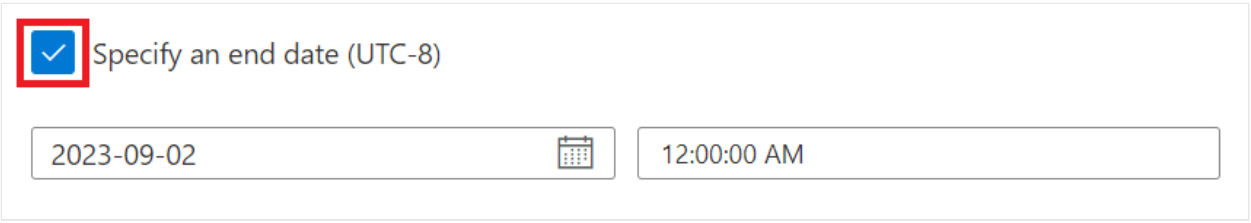
Start at

The **Start at** option allows you to set when your reports will begin their refresh schedule, and is relative to the time zone you selected. Set it for the current date and time to set the schedule immediately.



Specify an end date

If you want your reports to stop refreshing after a certain amount of time, you can enable this option by selecting the check box and providing an end date. For example: a monthly recurrence schedule refreshing on the first of the month with the below end date would refresh on September 1 2023, and wouldn't refresh again in October.

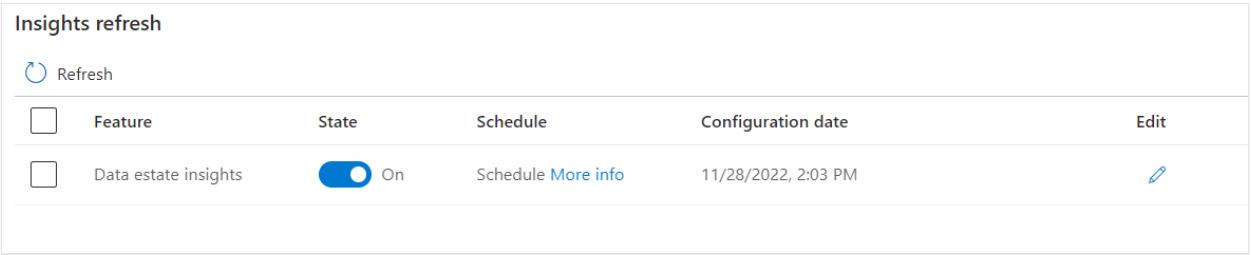


Insights refresh menu

In the management menu of the Microsoft Purview governance portal, you'll find the Insights refresh menu that lets you set your Data Estate Insights report availability. Here's all the information you can find there:

Expand table

Feature	State	Schedule	Configuration date	Edit
Lists the feature your settings are for	Either **On** or **Off**. For more information about turning off reports, see our disable Data Estate Insights article.	Select **More info** to see your recurrence schedule, start time, and end time.	The day your recurrence schedule was last set. By default this will be the creation date of your Microsoft Purview account.	Allows you to update your recurrence schedule when reports are turned on.



Next steps

- Learn how to use Asset insights
- Learn how to use Classification insights
- Learn how to use Glossary insights
- Learn how to use Label insights

- Disable Data Estate Insights

Feedback

Was this page helpful? Yes No

Understand the classic data stewardship report in Unified Catalog

As described in the insights concepts, the data stewardship report is part of the "Health" section of Data Estate Insights. This report offers a one-stop shop experience for data, governance, and quality focused users like chief data officers and data stewards to get actionable insights into key areas of gap in their data estate.

In this guide, you learn how to:

- ✓ Navigate and view data stewardship report from your Microsoft Purview account.
- ✓ Drill down for more asset count details.

Prerequisites

Before getting started with Data Estate Insights, make sure that you complete the following steps:

- Set up a storage resource and populate the account with data.
- Set up and complete a scan your storage source. For more information on how to create and complete a scan, see Manage data sources in Microsoft Purview.
- Permissions depending on your current experience:
  - If using the classic Microsoft Purview Catalog, you need the data curator role or insight reader role.
  - If using the new Microsoft Purview Unified Catalog, you need data health reader permissions.

Understand your data estate and catalog health in Data Estate Insights

In Data Estate Insights, you can get an overview of all assets inventoried in Data Map, and any key gaps that governance stakeholders can close for better governance of the data estate.

1. Open the Microsoft Purview portal.

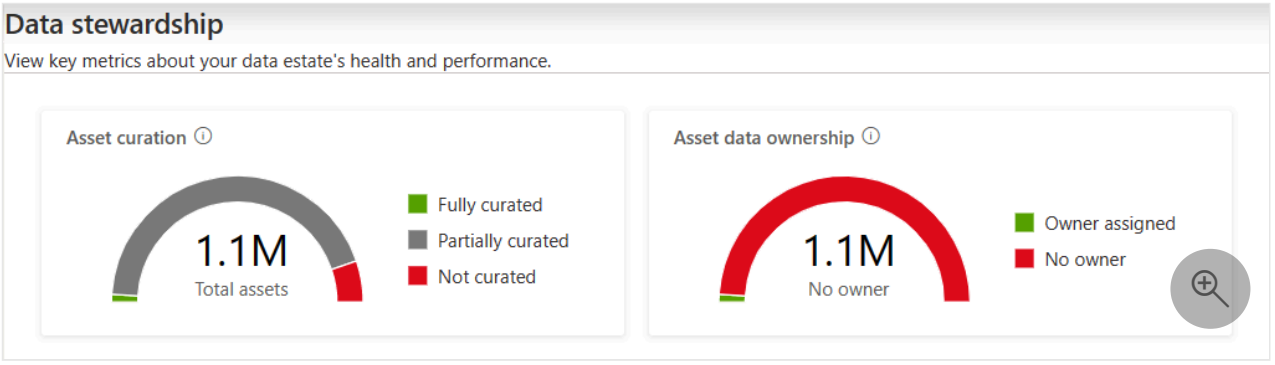
2. If you're using the classic Microsoft Purview portal, select **Data estate insights**. If you're using the new Microsoft Purview portal open **Unified Catalog**, select **Health Management**, then select **Reports**.

3. Select **Classic data stewardship**.

View data stewardship dashboard

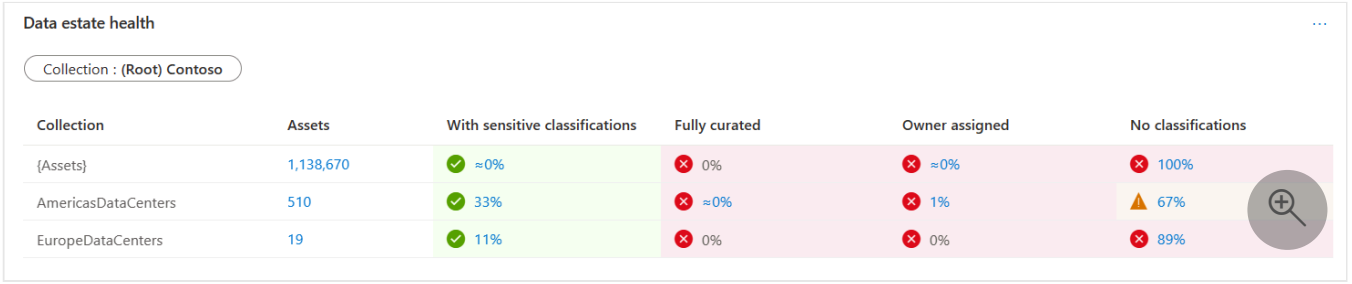
The dashboard is purpose-built for governance and quality-focused users, like data stewards and chief data officers, to understand the health management of their organization. The dashboard shows high-level KPIs that help reduce governance risks:

- **Asset curation**: The dashboard categorizes all data assets into three areas: "Fully curated," "Partially curated," and "Not curated," based on certain attributes of assets being present. An asset is "Fully curated" if it has at least one classification tag, an assigned data owner, and a description. If any of these attributes is missing, but not all, then the asset is categorized as "Partially curated" and if all of them are missing, then it's "Not curated."
- **Asset data ownership**: The dashboard categorizes assets that have the owner attribute within the "Contacts" tab as blank as "No owner". Otherwise, it categorizes them as "Owner assigned".



Health management

Health management is a scorecard view that helps management and governance focused users, like chief data officers, understand critical governance metrics that can be looked at by collection hierarchy.

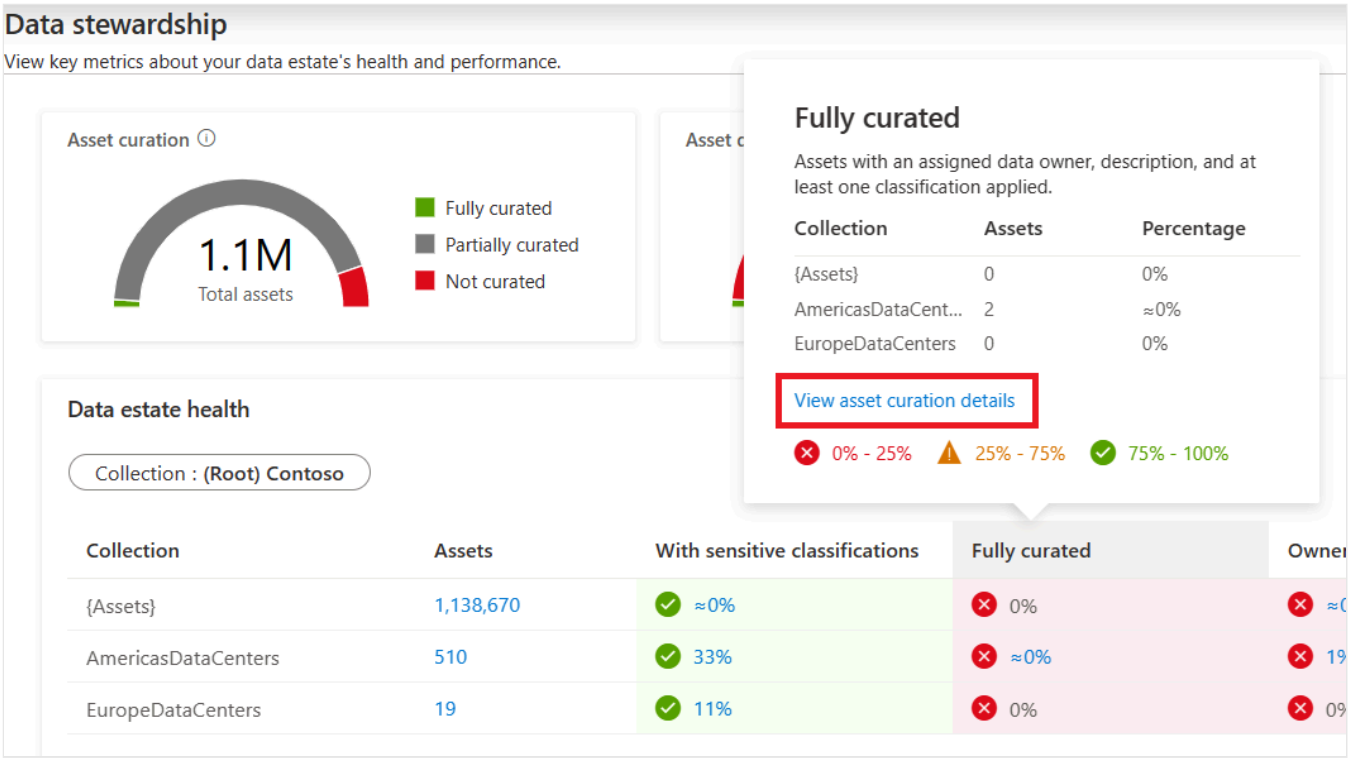


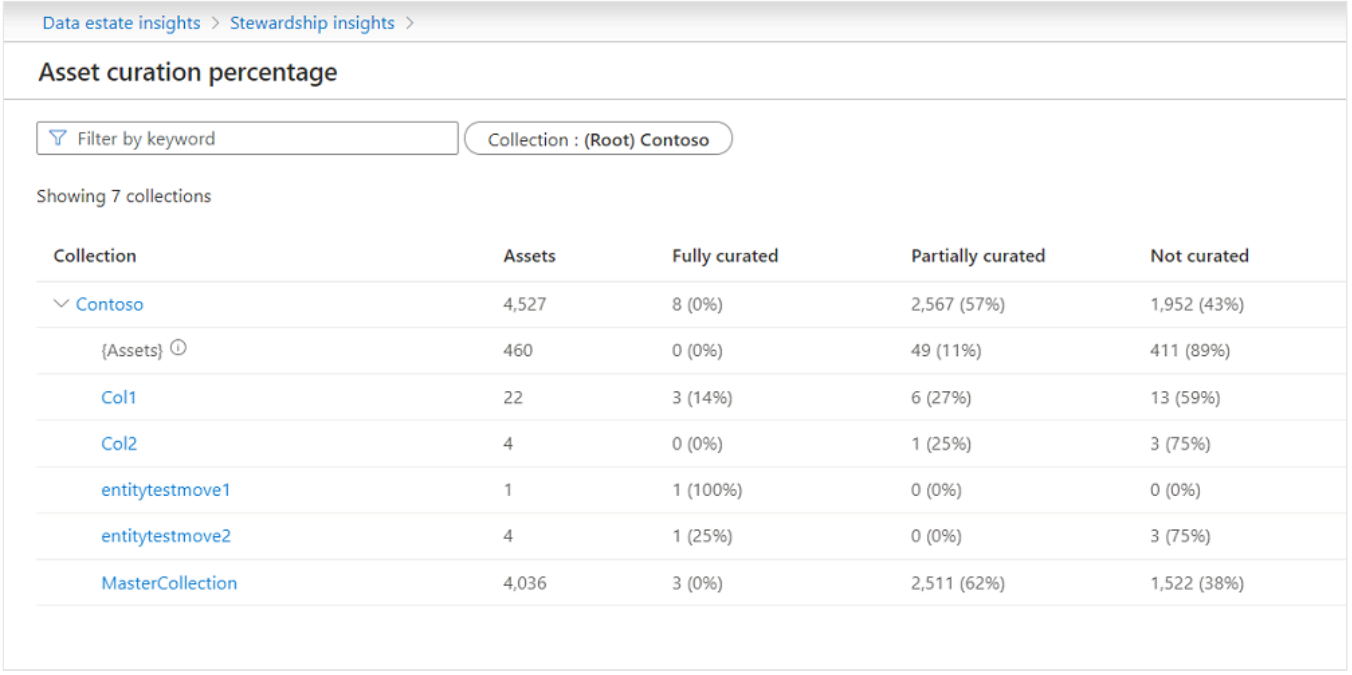
You can view the following metrics:

- **Assets**: Count of assets by collection drill-down.
- **With sensitive classifications**: Count of assets with any system classification applied.
- **Fully curated assets**: Count of assets that have a data owner, at least one classification, and a description.

- **Owner assigned**: Count of assets with data owner assigned on them.
- **No classifications**: Count of assets with no classification tag.
- **Out of date**: Percentage of assets that haven't been updated in over 365 days.
- **New**: Count of new assets pushed in the Data Map in the last 30 days.
- **Updated**: Count of assets updated in the Data Map in the last 30 days.
- **Deleted**: Count of deleted assets from the Data Map in the last 30 days.

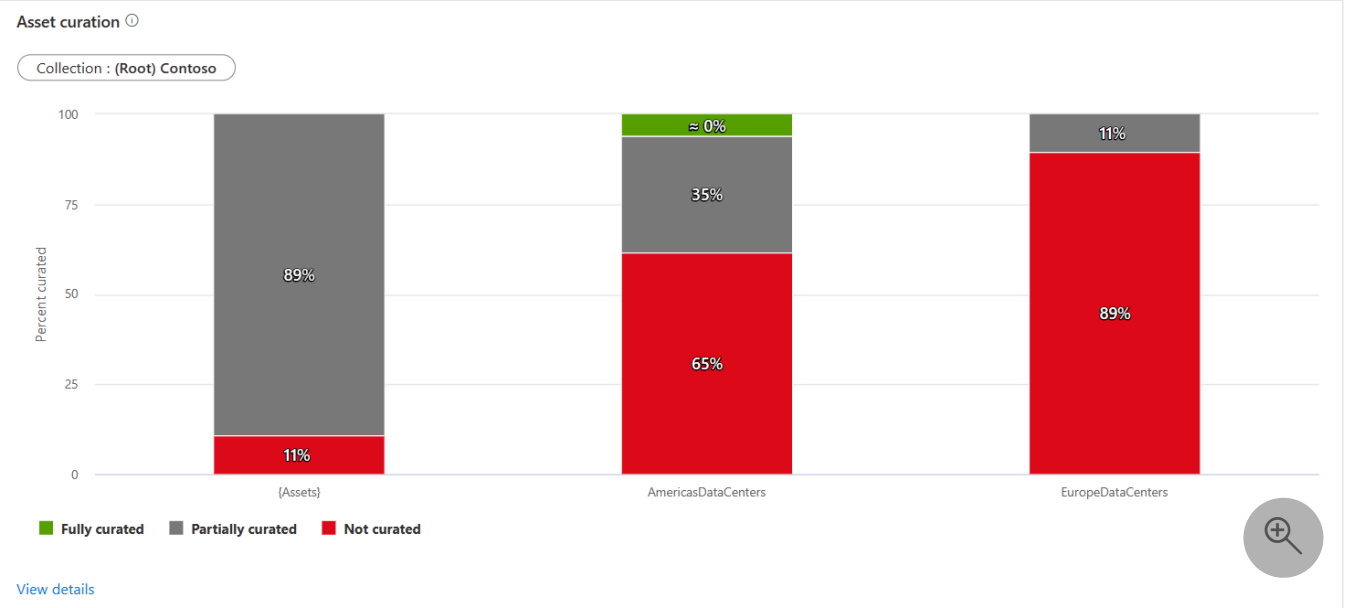
You can also drill down by collection paths. As you hover on each column name, it provides description of the column, provides recommended percentage ranges, and takes you to the detailed graph for further drill-down.





Asset curation

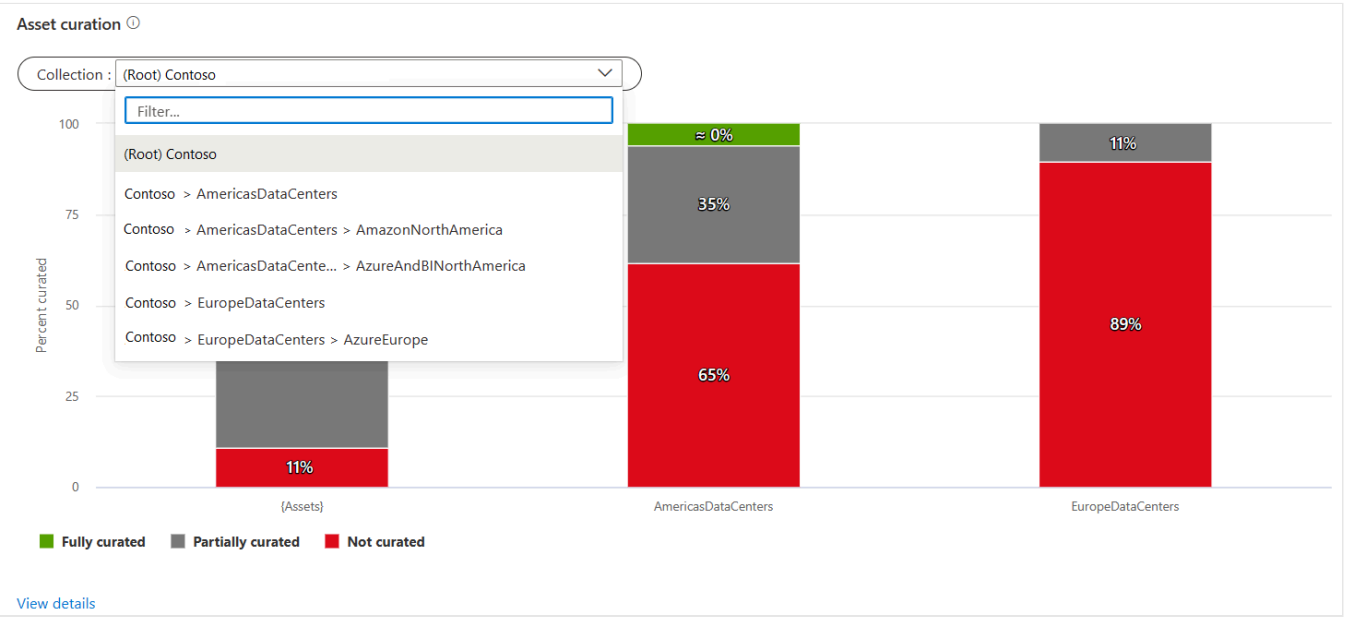
The system categorizes all data assets into three areas: **Fully curated**, **Partially curated**, and **Not curated**, based on whether assets have certain attributes.



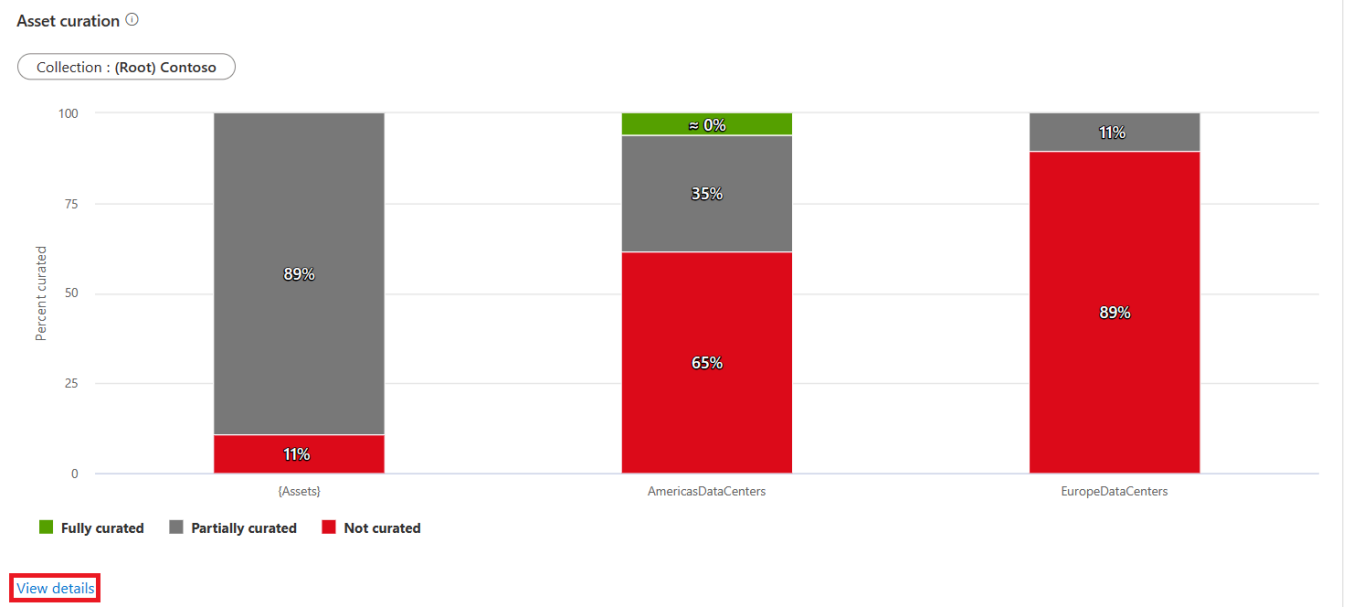
An asset is ***Fully curated*** if it has at least one classification tag, an assigned data owner, and a description.

If an asset is missing any of these attributes, but not all, the system categorizes it as **Partially curated**. If an asset is missing all of these attributes, the system lists it as **Not curated**.

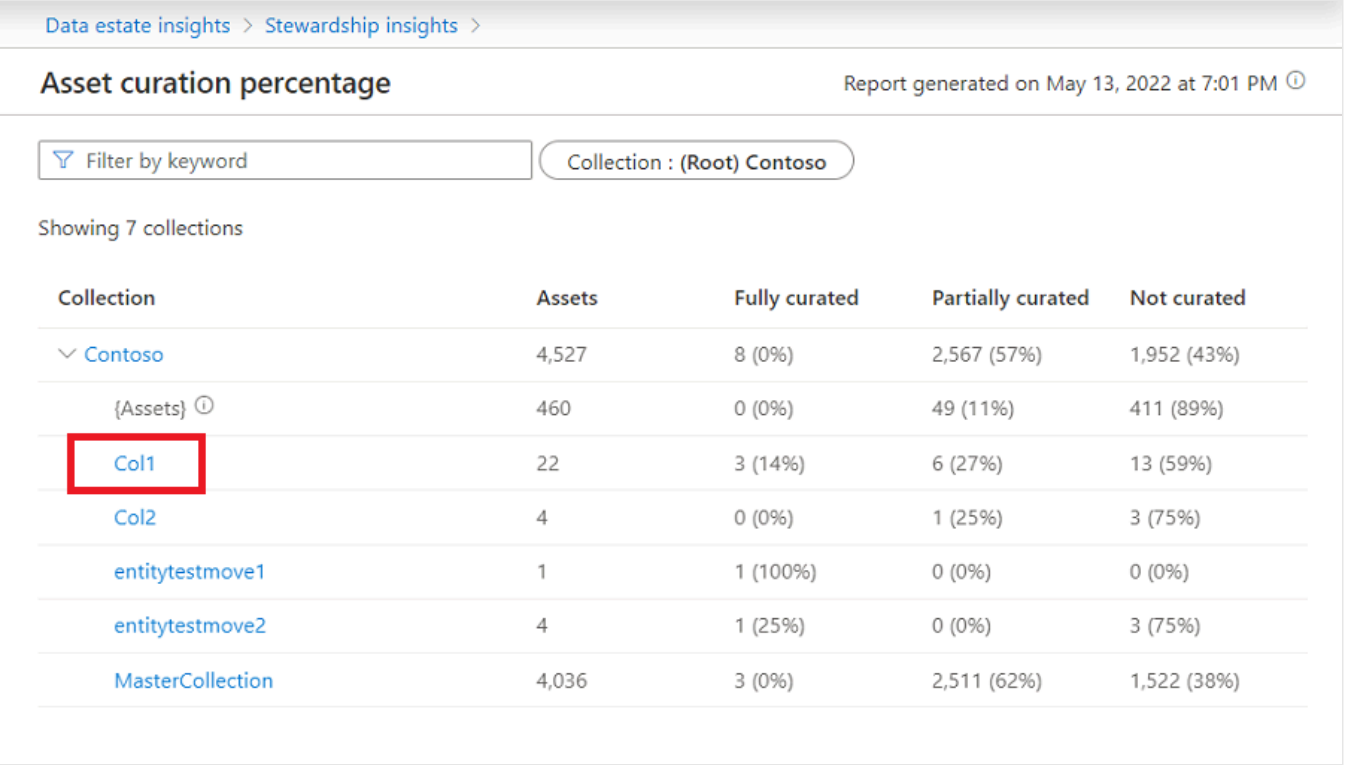
You can drill down by collection hierarchy.



For further information about which assets aren't fully curated, you can select **View details** link that takes you into the deeper view.



In the **View details** page, if you select a specific collection, it lists all assets with attribute values or blanks that make up the ***fully curated*** assets.



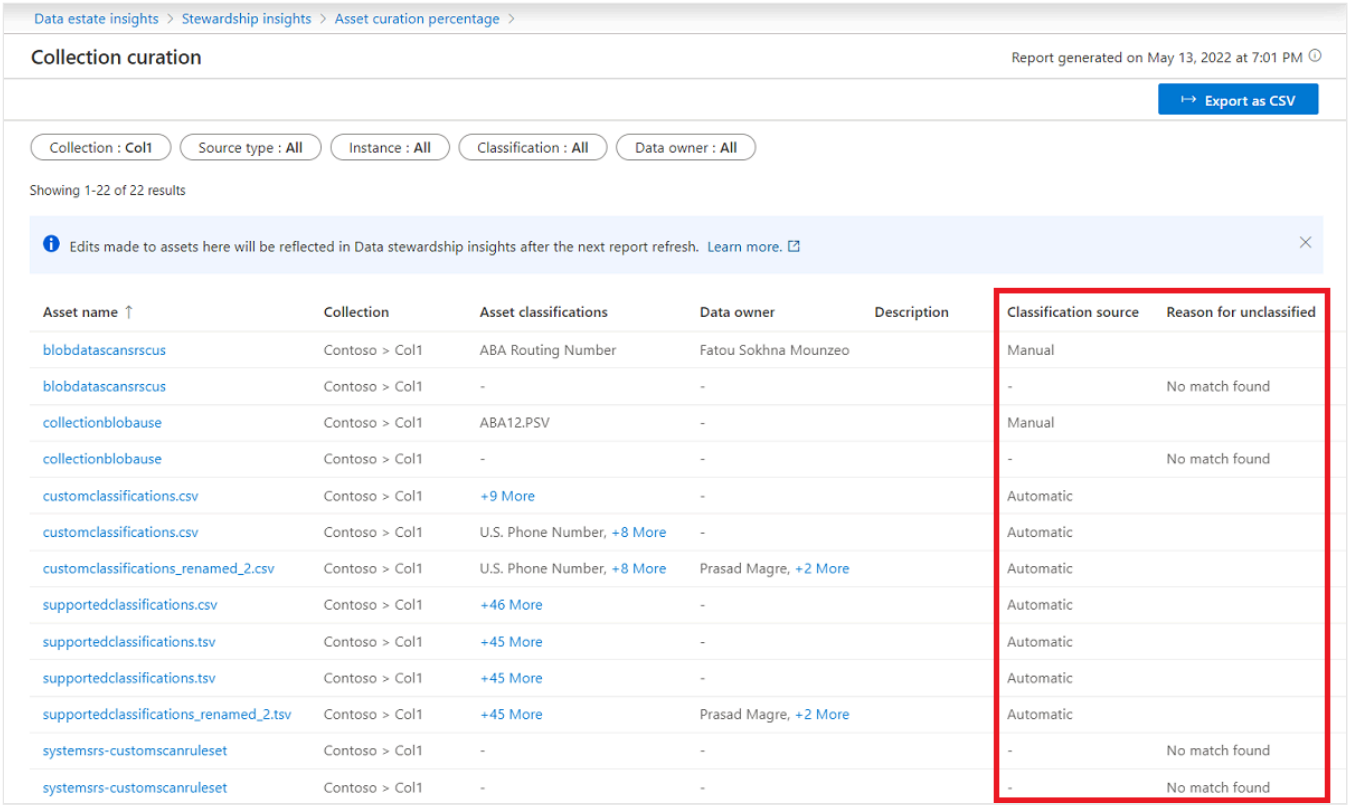
The detail page shows two key pieces of information:

First, it tells you what the ***classification source*** is, if the asset is classified. It's **Manual** if a curator or data owner applies the classification manually. It's **Automatic** if the system classifies the asset during scanning. This page only provides the last applied classification state.

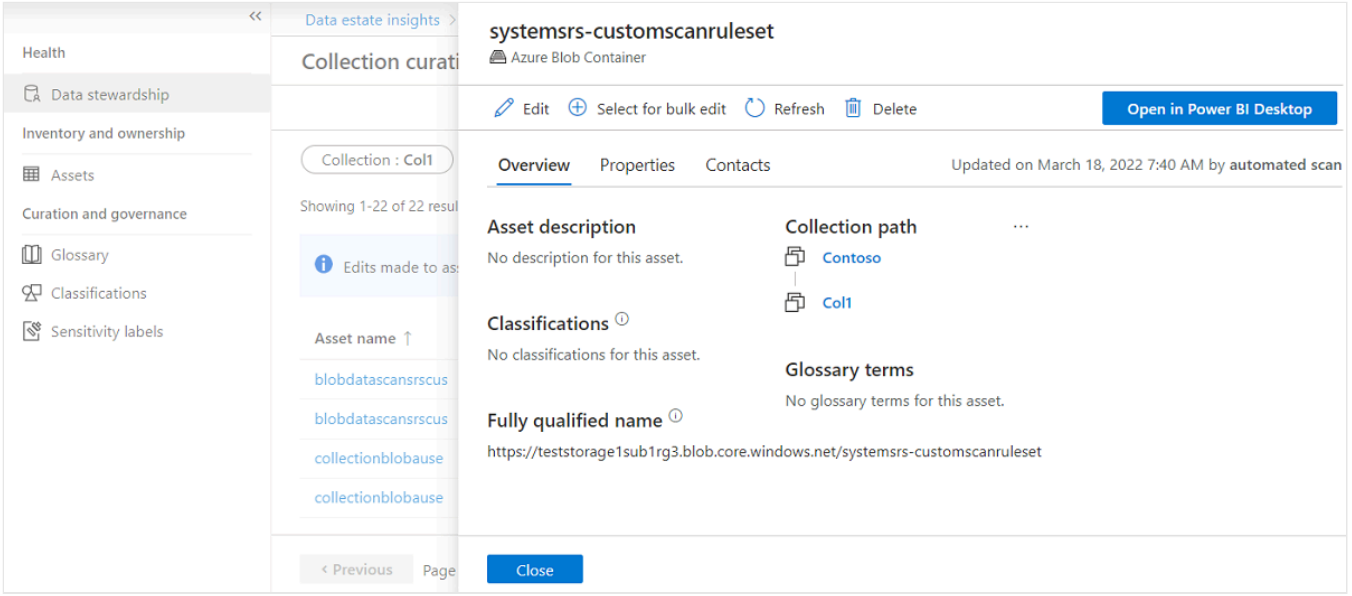
Second, if an asset is unclassified, it tells us why it's not classified, in the column ***Reasons for unclassified***. Currently, Data Estate Insights can provide one of the following reasons:

- No match found

- Low confidence score
- Not applicable



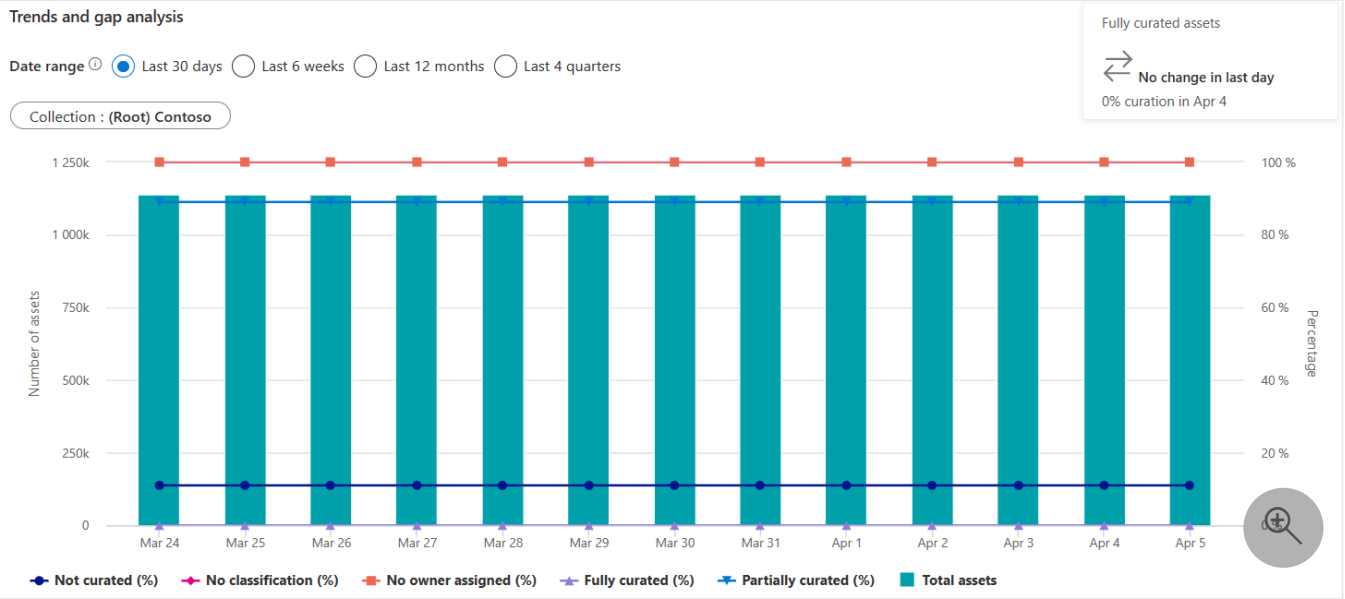
You can select any asset and add missing attributes, without leaving the **Data estate insights** app.



Trends and gap analysis

This graph shows how the assets and key metrics have been trending over different time periods:

- Last 30 days: The graph uses the last run of the day or the recording of the last run across days as a data point.
- Last six weeks: The graph uses the last run of the week, with the week ending on Sunday. If there's no run on Sunday, the graph uses the last recorded run.
- Last 12 months: The graph uses the last run of the month.
- Last four quarters: The graph uses the last run of the calendar quarter.



Next steps

Learn more about Classic catalog adoption insights.

Last updated on 11/14/2025

Understand the classic catalog adoption report in Unified Catalog

As described in the insights concepts, the catalog adoption report is part of the "Health" section of classic Microsoft Purview Data Estate Insights. This report offers a one-stop shop experience for administrators to determine if and how the classic Microsoft Purview Data Catalog is being used. It helps answer questions like:

- What are my users searching for?
- How many people used the catalog last month?
- What are the most used data assets?

Prerequisites

Before getting started with Data Estate Insights, make sure that you complete the following steps:

- Set up a storage resource and populate the account with data.
- Set up and complete a scan of your storage source. For more information about creating and completing a scan, see manage data sources in Microsoft Purview.
- Have the right permissions depending on your current experience:
  - If using the classic Microsoft Purview Data Catalog, you need the data curator role or insight reader role.
  - If using the new Microsoft Purview Unified Catalog, you need data health reader permissions.

Understand your data estate and catalog health in Data Estate Insights

In Data Estate Insights, you can get an overview of all assets inventoried in the Data Map, and any key gaps that governance stakeholders can close for better governance of the data estate.

1. Open the Microsoft Purview portal.

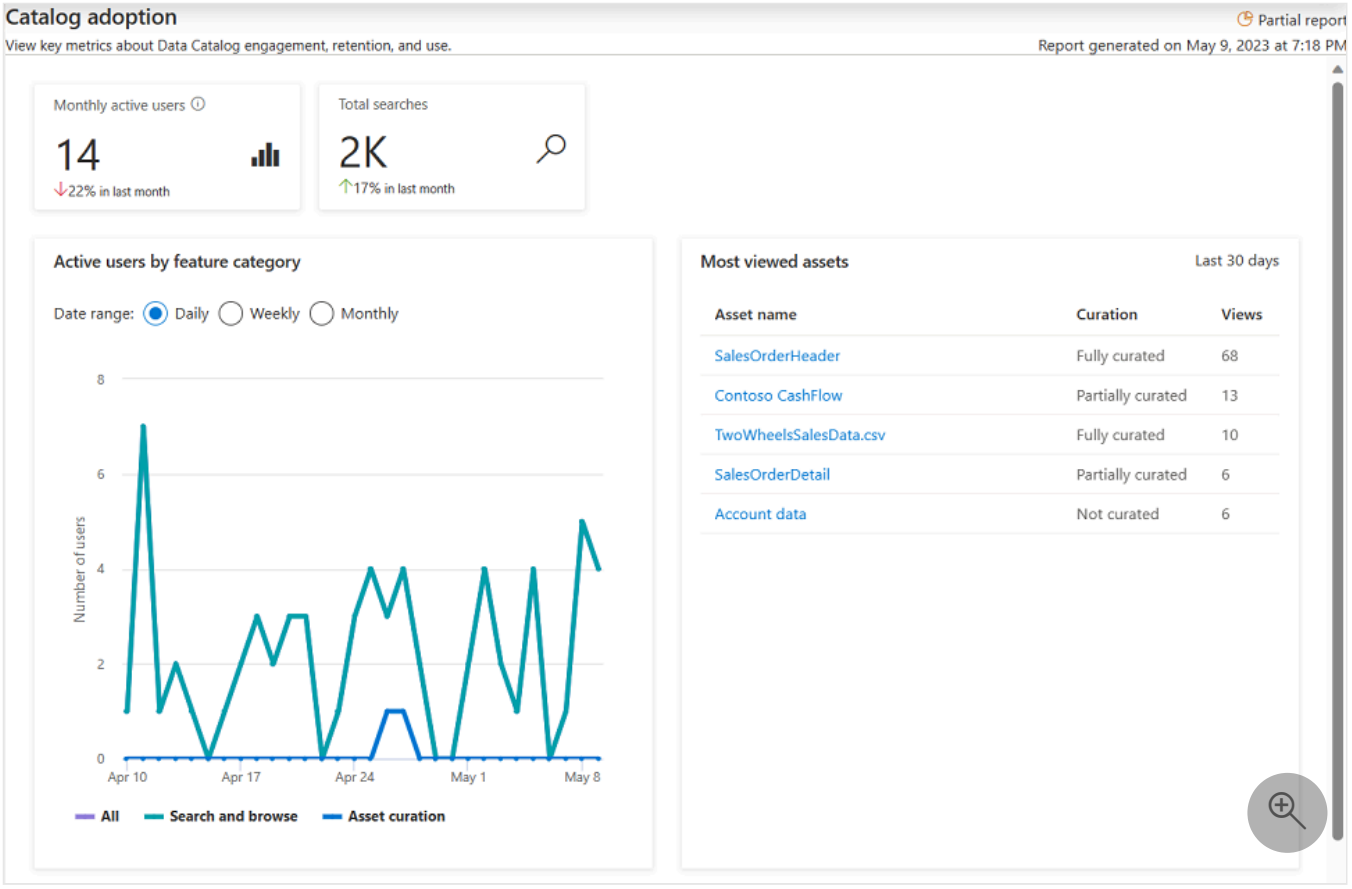
2. If you're using the classic Microsoft Purview portal, select **Data estate insights**. If you're using the new Microsoft Purview portal open **Unified Catalog**, select **Health Management**, then select **Reports**.

3. Select **Classic catalog adoption**.

View catalog adoption dashboard

The catalog adoption dashboard has several curated tiles and charts to identify:

- How many active users your catalog had in the last month.
- How many total searches were performed in the last month.
- Which catalog features are being used.
- Most viewed assets.
- Top searched keywords in catalog.



Monthly active users

The **Monthly active users** count is the number of unique users who perform any action, such as search, browse, or curation.

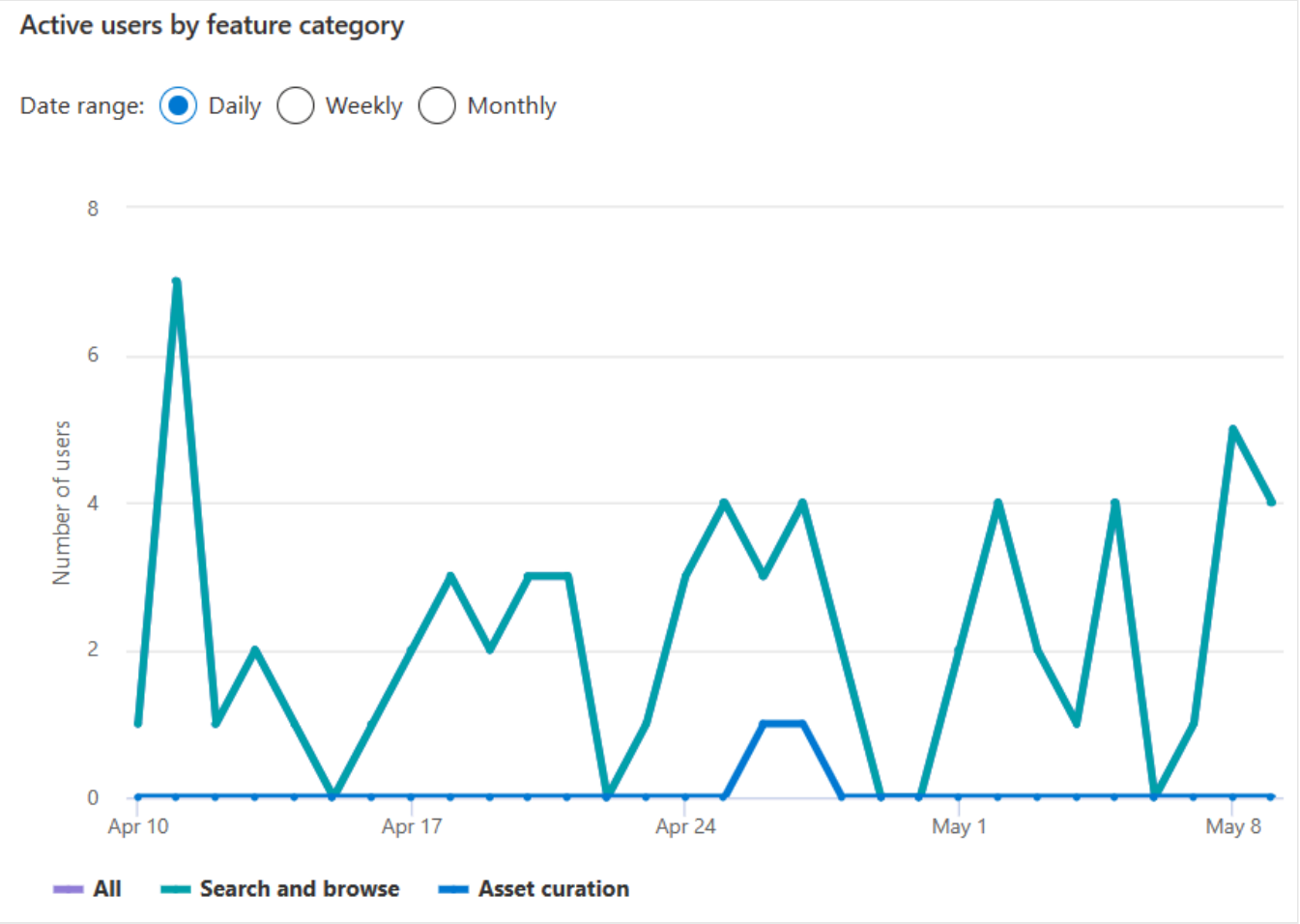
The **Monthly active users** tile shows how many users took at least one action in the catalog in the last 30 days. An action in the catalog includes searching for a term, browsing in the catalog, and updating an asset in the catalog. This tile also includes an indicator of a percentage increase or decrease in users from the previous month.

Total searches

The **Total searches** tile shows a count of all searches performed in the catalog over the last 30 days. The tile also shows the percentage increase or decrease in the number of searches from the previous month.

Active users by feature category

The **Active users by feature category** chart helps you monitor user activity trends in the catalog. Here's an example of the chart:



At the top of the chart, select your date range to view user activity on a daily, weekly, or monthly basis.

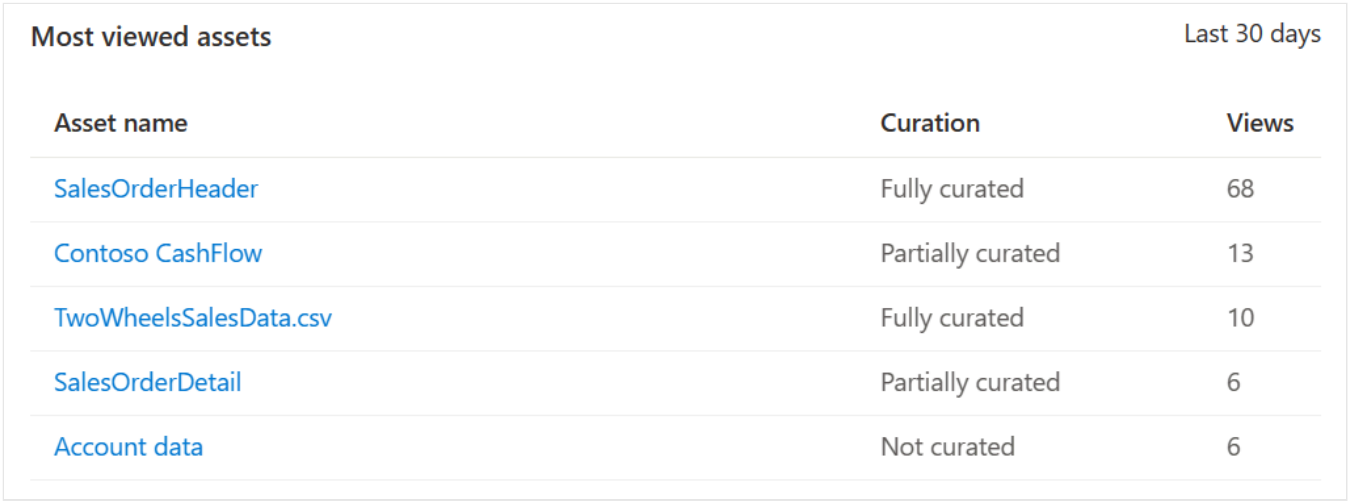
The chart shows the number of users in the date range who perform actions in one of three categories:

- **Search and browse**: The count of users per action category (search or browse). A user may appear in both counts if they perform both actions. For example, if a user performs both search and browse, the same user is counted twice in this number (once under search, once under browse). One user can contribute two counts in Search/Browse, but still counts as one in Monthly Active Users. This is why **Search and browse** can be higher than **Monthly active users**.

- **Asset curation**: The count of users who performed at least any one of these actions on an asset:
  - Added, removed, or updated a rating
  - Added or removed a tag
  - Added or removed a glossary term
  - Added or removed a classification
  - Edited an asset's name or description
  - Added or removed a certification
  - Added or removed a column level classification, glossary term, description, name, or data type
  - Added or removed manual lineage
  - Added or removed a contact
- **All**: The count of users who performed both a search or browse action, *and* at least one asset curation action within the date range.

Most viewed assets

The most viewed assets table shows the top assets in the catalog, by sum of views in the last 30 days.

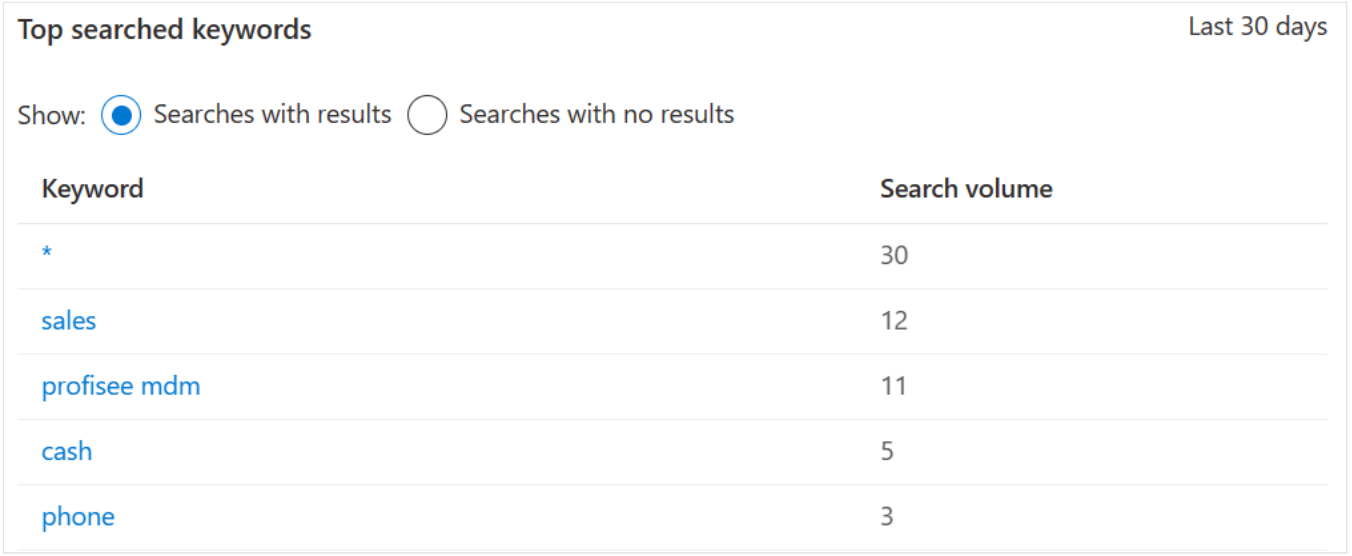


This table allows you to select through to your most viewed assets for more information, but the table also provides these other details:

- Curation status - There are three possible statuses: **Fully curated**, **Partially curated**, and **Not curated**, based on certain attributes of assets being present. An asset is **Fully curated** if it has at least one classification tag, an assigned Data Owner, and a description. If any of these attributes is missing, but not all, then the asset is categorized as **Partially curated** and if all of them are missing, then it's **Not curated**. For more information about the curation status of your data assets, see the data stewardship dashboard.
- Views - count of the number of views the asset received in the last 30 days.

Top searched keywords

The top searched keywords table shows your top keywords both for searches that produced results, and searches that didn't. That way you can know what users are finding with the catalog, and what they're still looking for.



At the top of the table, you can select one of the two radio buttons to select whether to show keywords for searches with results, or searches without results.

You can select any of the keywords to run the search in the catalog and see the results for yourself as well.

The table also provides search volume, which is the number of times that keyword was searched in the last 30 days.

Report generated timestamp

The timestamp displayed, such as, "Report generated on December 4, 2025 at 11:37 AM", indicates when the page was loaded or refreshed. It doesn't represent the time when the backend actually computed the KPI data. The actual calculation time comes from the backend timePoint field (for example, 2025-11-05T06:59:48) from the API response.

Monthly aggregation behavior for usage metrics

The **Monthly active users** and **Total searches** metrics display data from the previous fully completed month. For example, December views show November data. The KPI is updated during the job run after the month changes, even though the job runs every week.

Next steps

Learn more about Data Estate Insights:

- Data Estate Insights Concepts
- Data stewardship insights

Last updated on 01/15/2026

Understand the classic assets report in Unified Catalog

This article describes how to access, view, and filter data asset insight reports for your data.

In this article, you learn how to:

- ✓ View data estate insights from your Microsoft Purview account.
- ✓ Get a bird's eye view of your data.
- ✓ Drill down for more asset count details.

Prerequisites

Before getting started with Microsoft Purview Data Estate Insights, make sure that you've completed the following steps:

- Set up a storage resource and populated the account with data.
- Set up and completed a scan your storage source. For more information to create and complete a scan, see Manage data sources in Microsoft Purview.
- Permissions depending on your current experience:
  - If using the classic Microsoft Purview Catalog, you need the data curator role or insight reader role.
  - If using the new Microsoft Purview Unified Catalog, you need data health reader permissions.

Understand your asset inventory in Data Estate Insights

In Microsoft Purview Data Estate Insights, you can get an overview of the assets that you scanned into the Data Map. You can also view key gaps that governance stakeholders can close for better governance of the data estate.

1. Open the Microsoft Purview portal.

2. If you're using the classic Microsoft Purview portal, select **Data estate insights**. If you're using the new Microsoft Purview portal open **Unified Catalog**, select **Health Management**, then select **Reports**.

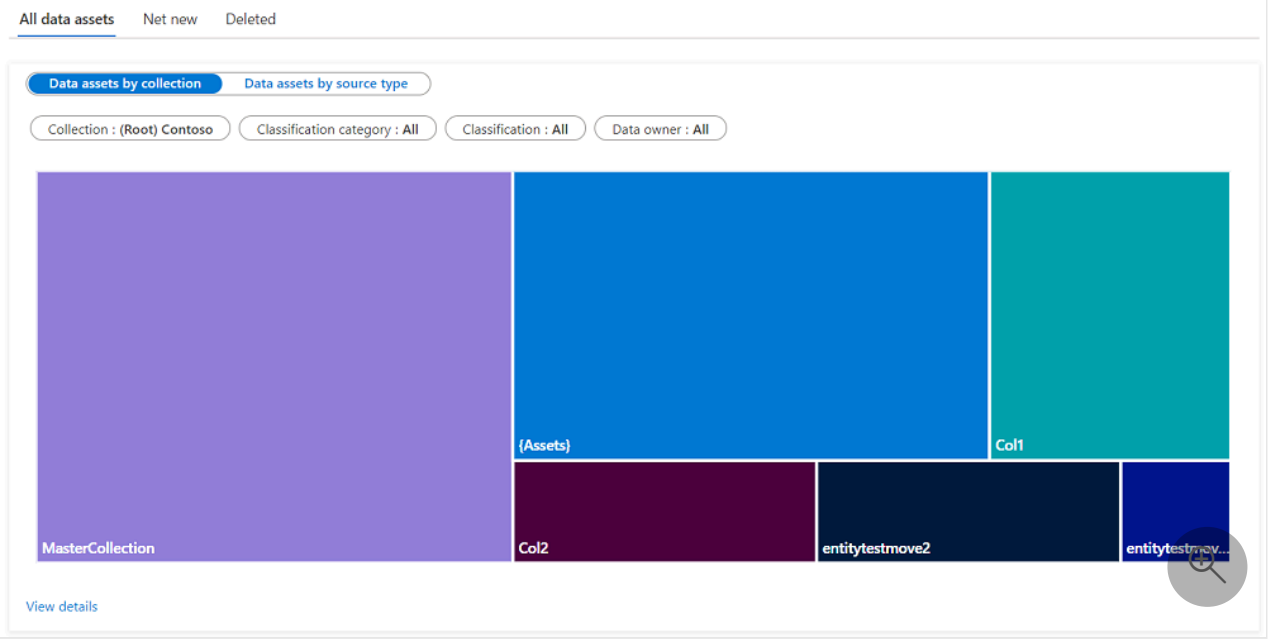
3. Select the **Classic assets** report.

4. The **Assets Summary** report provides several high-level KPIs, with these graphs:

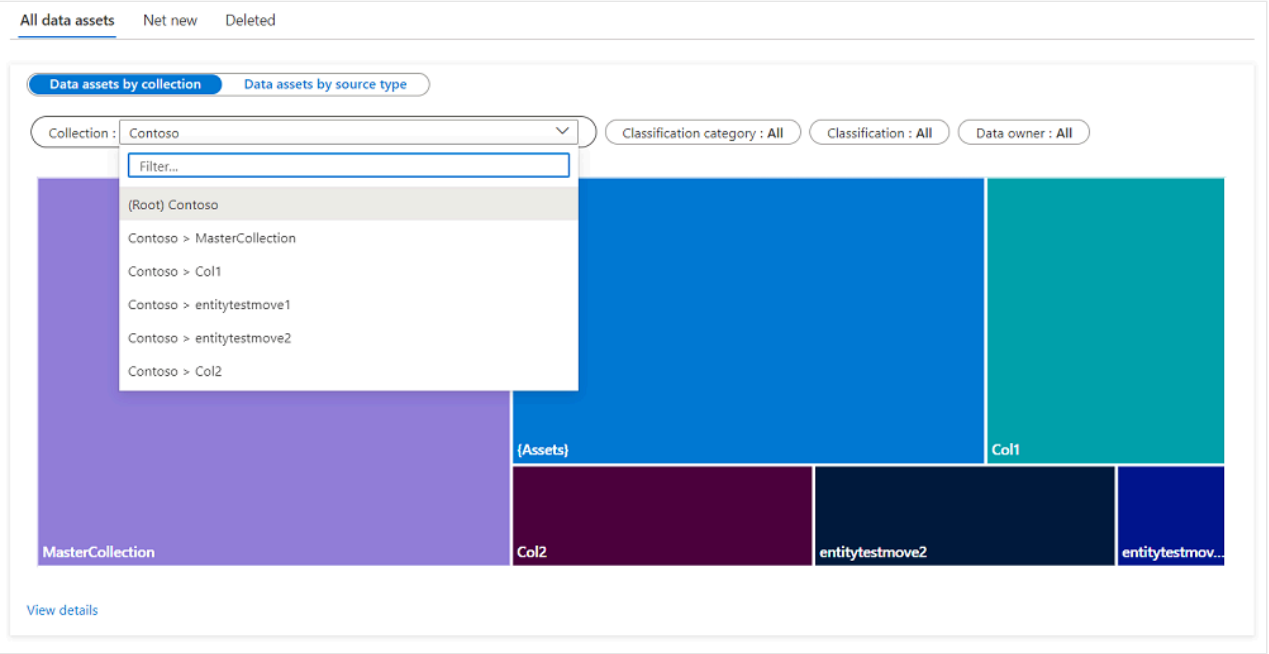
- **Unclassified assets**: Assets with no system or custom classification on the entity or its columns.
- **Unassigned data owner**: Assets that have the owner attribute within "Contacts" tab as blank.
- **Net new assets in last 30 days**: Assets that you added to the Purview account, via data scan or Atlas API pushes.
- **Deleted assets in last 30 days**: Assets that you deleted from the Purview account, as a result of deletion from data sources.



5. Below these KPIs, you can also view your data asset distribution by collection.

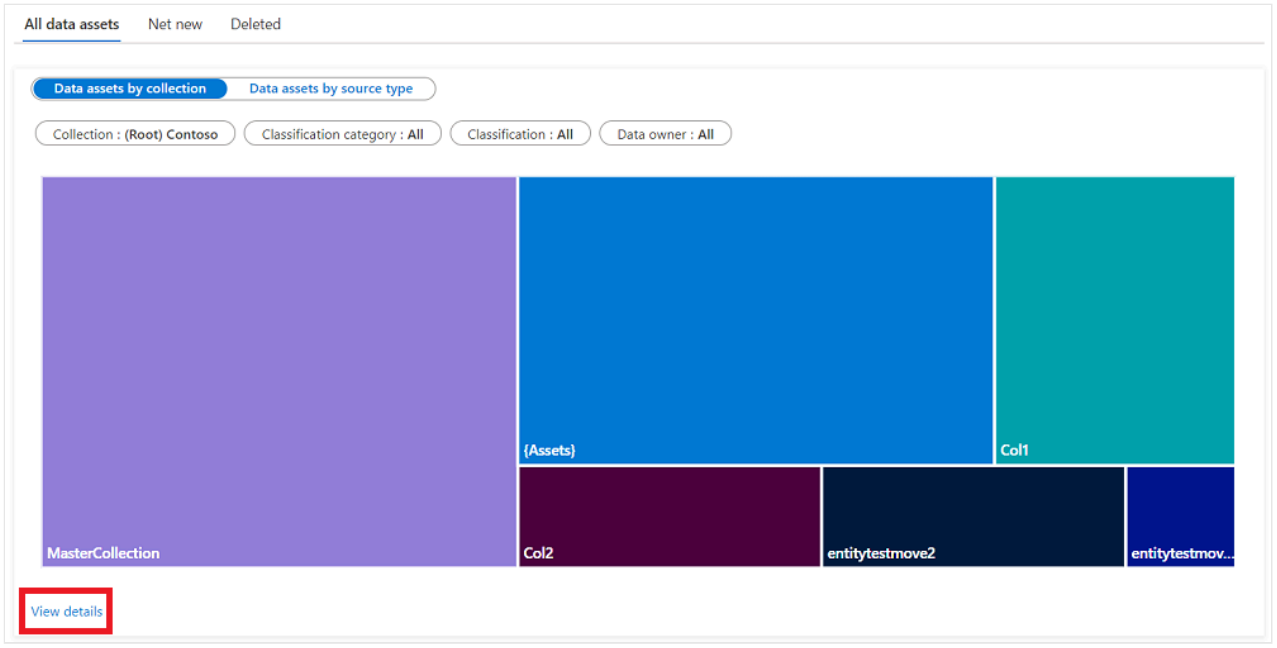


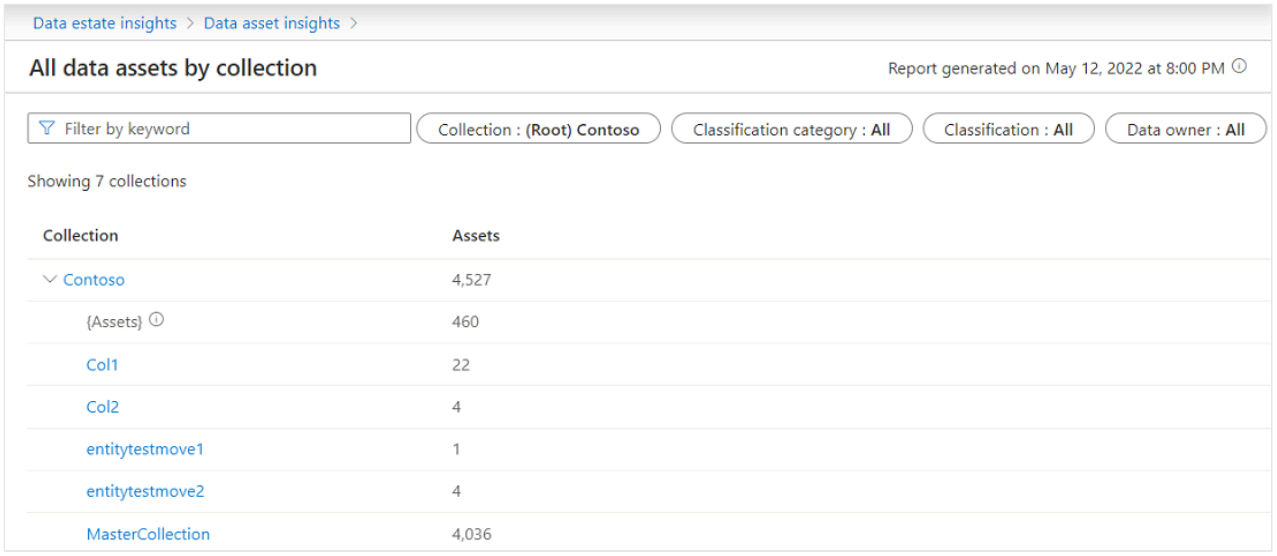
6. Using filters you can drill down to assets within a specific collection or classification category.



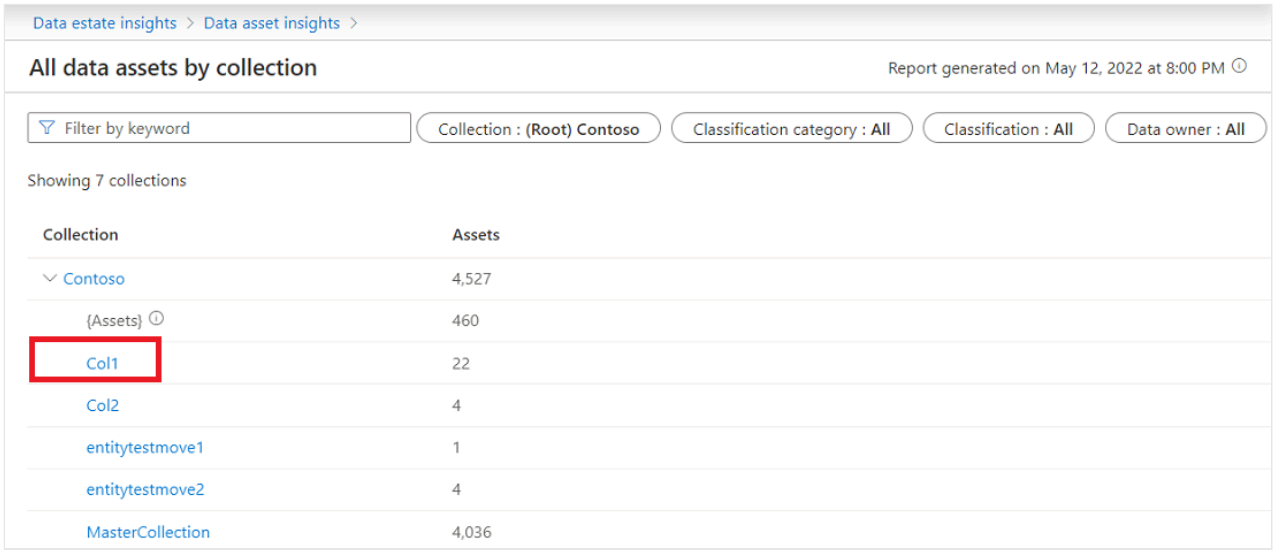
> ⓘ Note
>
> ***Each classification filter has some common values:***
>
> - **Applied**: Any filter value is applied
> - **Not Applied**: No filter value is applied. For example, if you pick a classification filter with value as "Not Applied", the graph shows all assets with no classification.
> - **All**: Filter values are cleared. Meaning the graph shows all assets, with or without classification.
> - **Specific**: You pick a specific classification from the filter, and only that classification is shown.

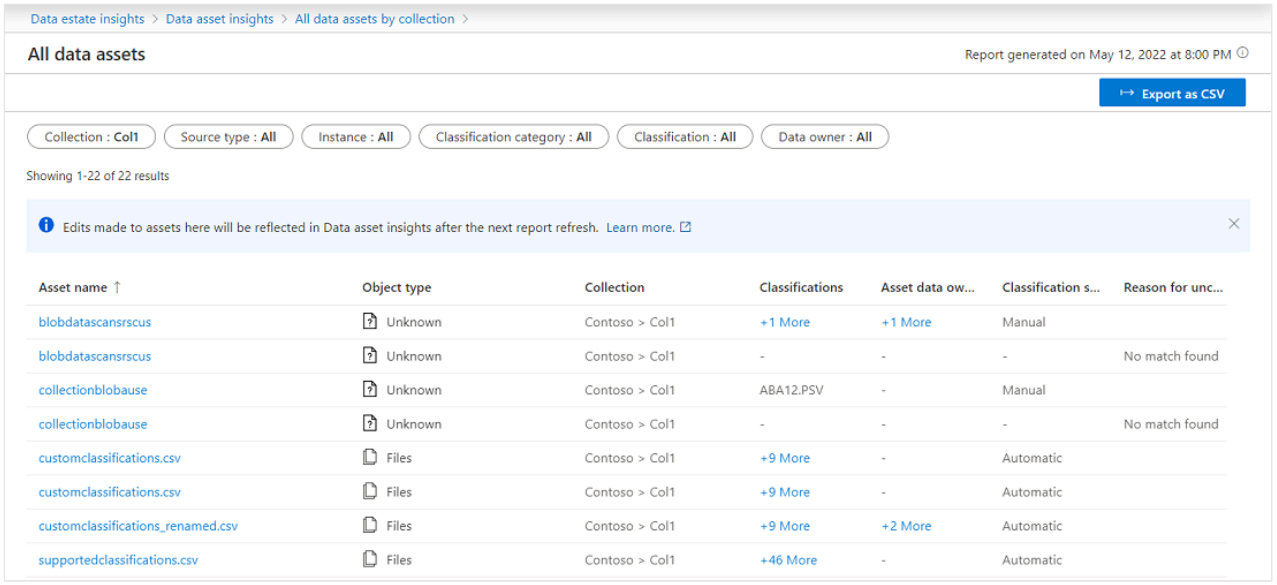
7. To learn more about which specific assets are shown in the graph, select **View details**.



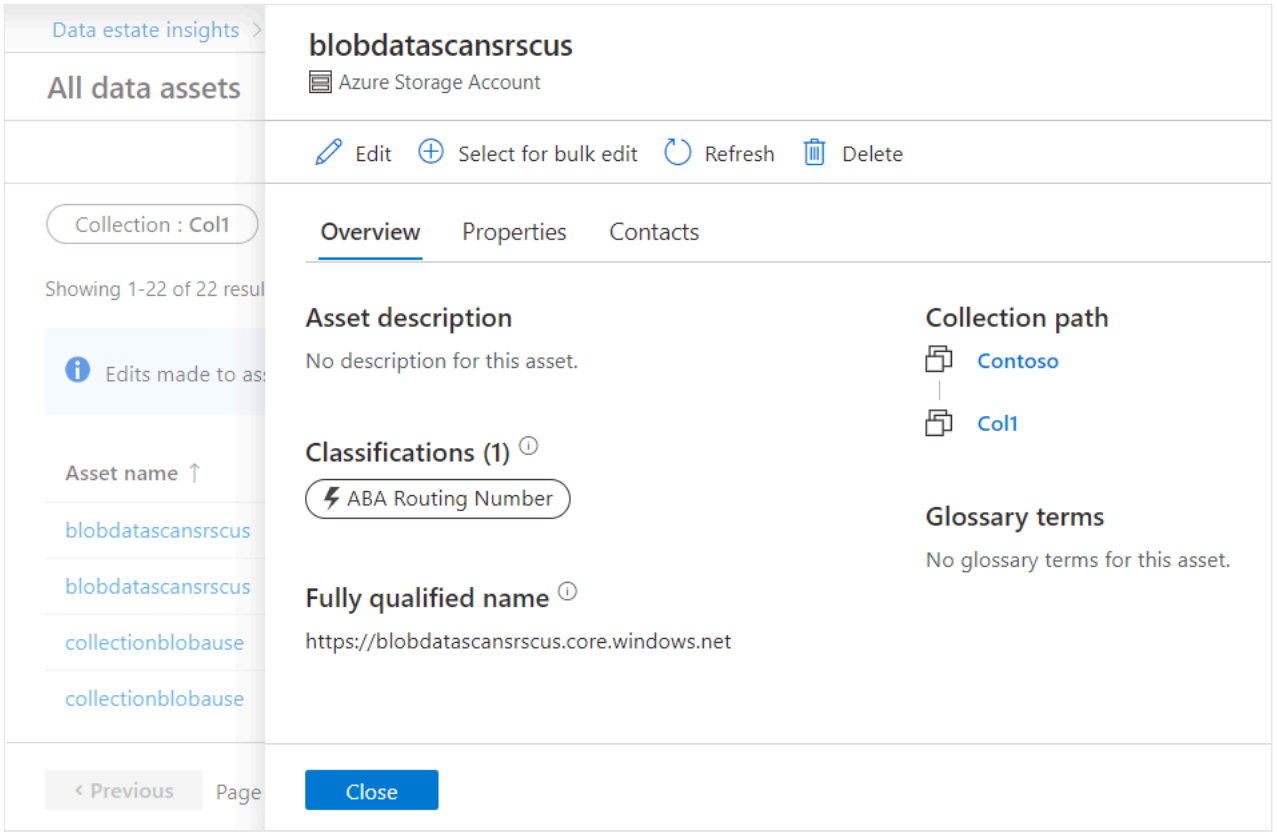


8. You can select any collection to view the collection's asset list.





9. You can also select an asset to edit without leaving the Data Estate Insights App.



File-based source types

The next graphs in asset insights show a distribution of file-based source types. The first graph, **Size trend (GB) of file type within source types**, shows top file type size trends over the last 30 days.

1. Pick your source type to view the file type within the source.

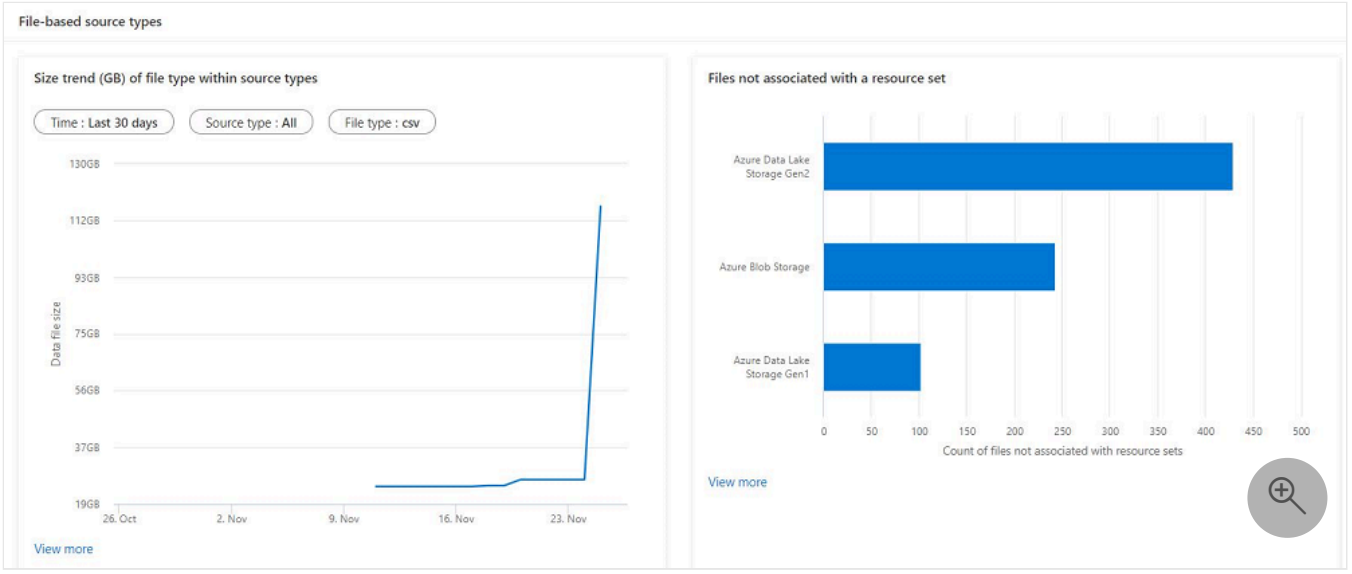
2. Select **View details** to see the current data size, change in size, current asset count, and change in asset count.

> **Note**
>
> If the scan runs only once in the last 30 days, or any catalog change such as classification addition or removal happens only once in 30 days, then the change information appears blank.

3. See the top folders with change top asset counts when you select source type.

4. Select the path to see the asset list.

The second graph in file-based source types is **Files not associated with a resource set**. If you expect that all files should roll up into a resource set, this graph can help you understand which assets aren't rolled up. Missing assets can be an indication of the wrong file-pattern in the folder. You can select **View details** below the graph for more information.



Next steps

Learn how to use Data Estate Insights with the following resources:

- Learn how to use data stewardship insights.
- Learn how to use classification insights.
- Learn how to use glossary insights.
- Learn how to use label insights.

Last updated on 11/14/2025

Understand the classic glossary report in Unified Catalog

This article describes how to access, view, and filter Microsoft Purview glossary insight reports for your data. In this how-to guide, you learn how to:

- ✓ Find Data Estate Insights.
- ✓ Get a bird's eye view of your data.

Prerequisites

Before getting started with Data Estate Insights glossary insights, make sure that you complete the following steps:

- Set up a storage resource and populate the account with data.
- Set up and complete a scan your storage source.
- Set up at least one business glossary term and attach it to an asset.
- Permissions depending on your current experience:
  - If using the classic Microsoft Purview Catalog, you need the data curator role or insight reader role.
  - If using the new Microsoft Purview Unified Catalog, you need data health reader permissions.

Use Microsoft Purview glossary insights

In Microsoft Purview, you can create glossary terms and attach them to assets. As you make use of these terms in your data map, you can view the glossary distribution in glossary insights. These insights give you the state of your glossary based on:

- Number of terms attached to assets.
- Status of terms.
- Distribution of roles by users.

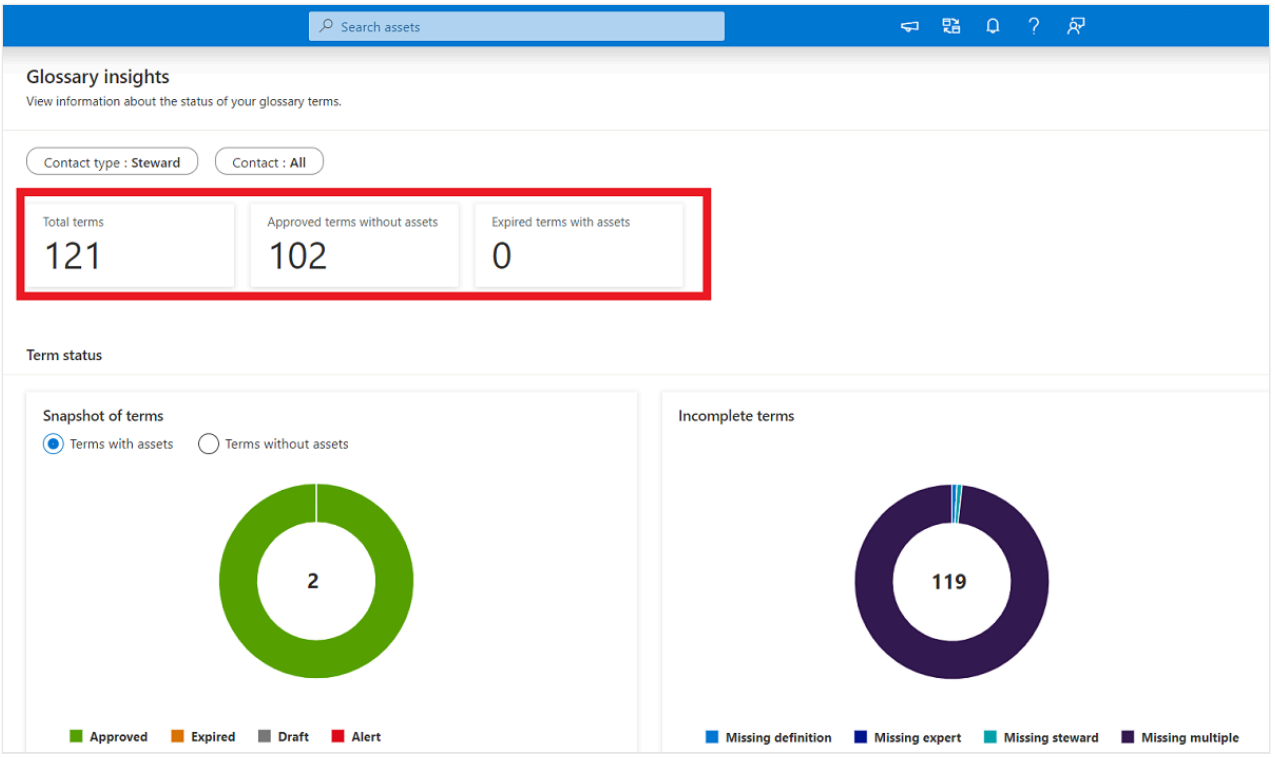
View glossary insights

1. Open the Microsoft Purview portal.

2. If you're using the classic Microsoft Purview portal, select **Data estate insights**. If you're using the new Microsoft Purview portal open **Unified Catalog**, select **Health Management**, then select **Reports**.

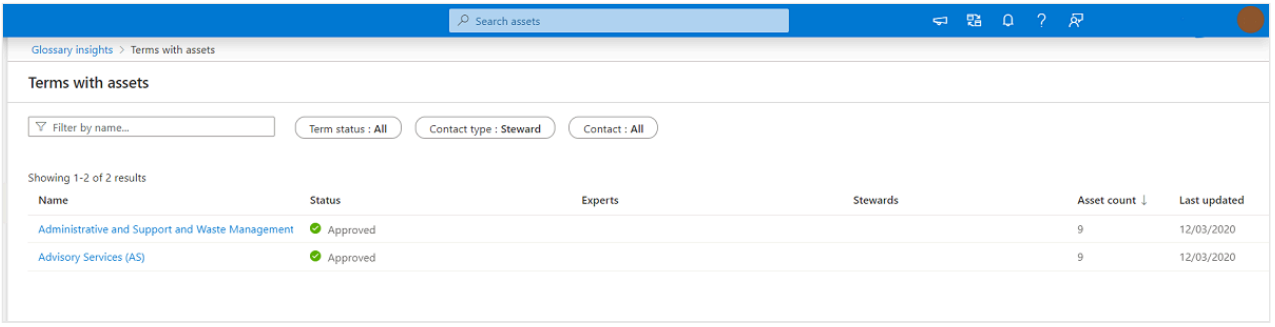
3. Select the **Classic glossary** report.

4. The report starts with **High-level KPIs** that shows ***Total terms*** in your Microsoft Purview account, ***Approved terms without assets***, and ***Expired terms with assets***. Each of these values helps you understand the current health of your Glossary.

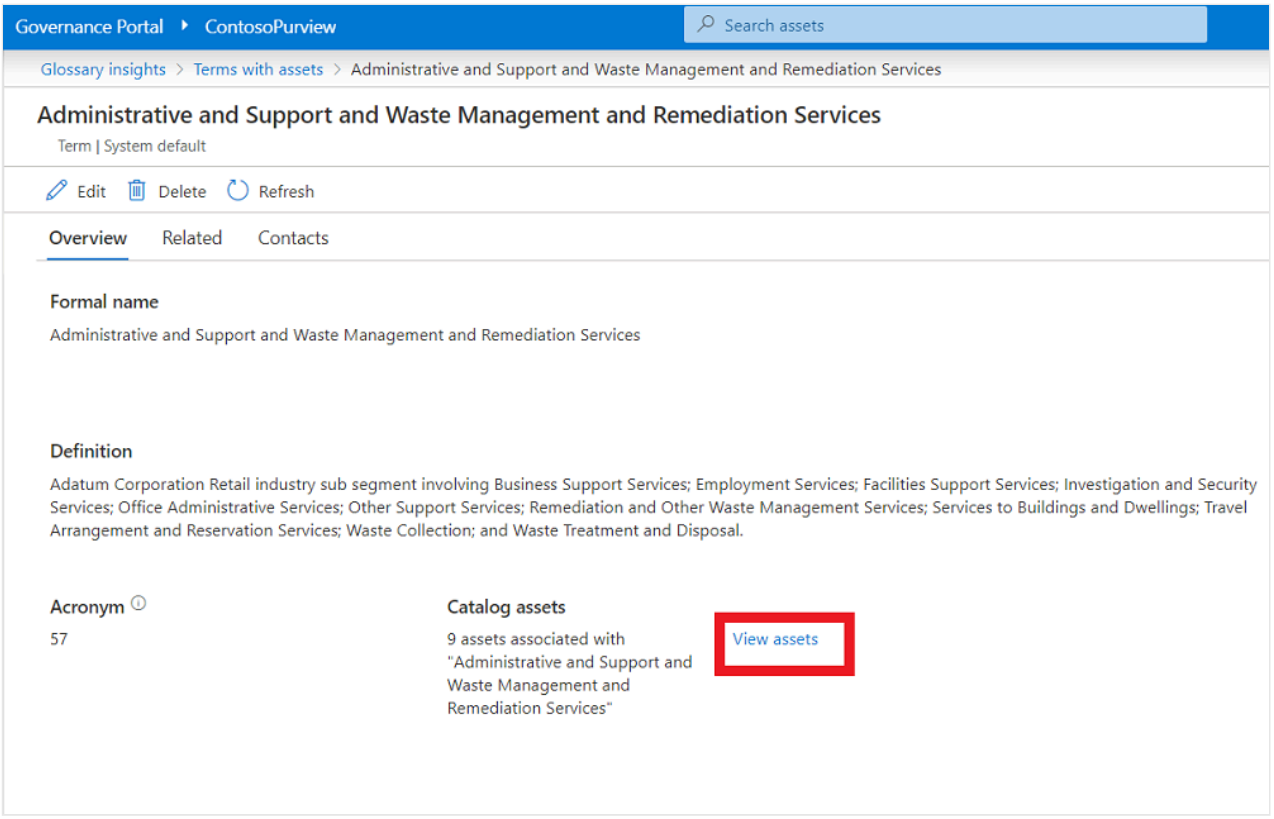


5. The **Snapshot of terms** section (displayed above) shows the term status as ***Draft***, ***Approved***, ***Alert***, and ***Expired*** for terms with assets and terms without assets.

6. Select **View details** to see the term names with various status and more details about ***Stewards*** and ***Experts***.



7. When you select "View more" for ***Approved terms with assets***, Data Estate Insights allows you to navigate to the **Glossary** term detail page, from where you can further navigate to the list of assets with the attached terms.



8. In Glossary insights page, view a distribution of **Incomplete terms** by type of information missing. The graph shows count of terms with ***Missing definition***, ***Missing expert***, ***Missing steward***, and ***Missing multiple*** fields.

9. Select ***View more*** from **Incomplete terms**, to view the terms that have missing information. You can navigate to Glossary term detail page to input the missing information and ensure the glossary term is complete.

Last updated on 11/14/2025

Understand the classic classifications report in Unified Catalog

This article describes how to access, view, and filter the classic classifications report for your data.

In this article, you learn how to:

- ✓ Launch your Microsoft Purview account from Microsoft Azure.
- ✓ View classification insights on your data.
- ✓ Drill down for more classification details on your data.

Prerequisites

Before getting started with Data Estate Insights, make sure you complete the following steps:

- Set up a storage resource and populate the account with data.
- Set up and complete a scan on the data in each data source. For more information, see Manage data sources in Microsoft Purview and Create a scan rule set.
- Permissions depending on your current experience:
  - If using the classic Microsoft Purview Catalog, you need the data curator role or insight reader role.
  - If using the new Microsoft Purview Unified Catalog, you need data health reader permissions.

Use Data Estate Insights for classifications

In Microsoft Purview, classifications are similar to subject tags, and are used to mark and identify data of a specific type that's found within your data estate during scanning.

Microsoft Purview uses the same sensitive information types as Microsoft 365, allowing you to stretch your existing security policies and protection across your entire data estate.

> ⓘ **Note**
>
> After you scan your source types, wait a couple of hours for the classic classifications report to reflect the new assets.

View classification insights

1. Open the Microsoft Purview portal.
2. If you're using the classic Microsoft Purview portal, select **Data estate insights**. If you're using the new Microsoft Purview portal open **Unified Catalog**, select **Health management**, then select **Reports**.
3. Select the **Classic classifications** report. The **Classification insights** page displays the following areas:

Expand table

Area	Description
Overview of sources with classifications	Displays tiles that provide: - The number of subscriptions found in your data - The number of unique classifications found in your data - The number of classified sources found - The number of classified files found - The number of classified tables found
Top sources with classified data (last 30 days)	Shows the trend, over the past 30 days, of the number of sources found with classified data.
Top classification categories by sources	Shows the number of sources found by classification category, such as **Financial** or **Government**.
Top classifications for files	Shows the top classifications applied to files in your data, such as credit card numbers or national/regional identification numbers.
Top classifications for tables	Shows the top classifications applied to tables in your data, such as personal identifying information.
Classification activity (files and tables)	Displays separate graphs for files and tables, each showing the number of files or tables classified over the selected timeframe. **Default**: 30 days Select the **Time** filter above the graphs to select a different time frame to display.

Classification insights drilldown

In any of the following **Classification insights** graphs, select the **View details** link to drill down for more details:

- **Top classification categories by sources**
- **Top classifications for files**
- **Top classifications for tables**

- **Classification activity > Classification data**

To learn more, use the following options:

Expand table

Option	Description
Filter your data	Use the filters above the grid to filter the data shown, including the classification name, subscription name, or source type. If you're not sure of the exact classification name, enter part or all of the name in the **Filter by keyword** box.
Sort the grid	Select a column header to sort the grid by that column.
Edit columns	To display more or fewer columns in your grid, select **Edit Columns**, then select the columns you want to view or change the order.
Drill down further	To drill down to a specific classification, select a name in the **Classification** column to view the **Classification by source** report. This report displays data for the selected classification, including the source name, source type, subscription ID, and the numbers of classified files and tables.
Browse assets	To browse through the assets found with a specific classification or source, select a classification or source, depending on the report you're viewing, then select **Browse assets** above the filters. The search results display all of the classified assets found for the selected filter. For more information, see Search the Microsoft Purview Unified Catalog.

Next steps

Learn how to use Data Estate Insights with the following resources:

- Learn how to use Asset insights.
- Learn how to use Data Stewardship.
- Learn how to use Classification insights.
- Learn how to use Glossary insights.
- Learn how to use Label insights.

Last updated on 11/14/2025

Understand the classic sensitivity labels report in Unified Catalog

This article describes how to access, view, and filter security insights provided by sensitivity labels applied to your data.

Supported data sources include: Azure Blob Storage, Azure Data Lake Storage (ADLS) GEN 1, Azure Data Lake Storage (ADLS) GEN 2, SQL Server, Azure SQL Database, Azure SQL Managed Instance, Amazon S3 buckets, Amazon RDS databases (public preview), Power BI.

In this article, you learn how to:

- ✓ Launch your Microsoft Purview account from Microsoft Azure.
- ✓ View sensitivity labeling insights on your data.
- ✓ Drill down for more sensitivity labeling details on your data.

Prerequisites

Before getting started with Data Estate Insights, make sure that you complete the following steps:

- Set up your Azure resources and populate the relevant accounts with test data.
- Extended sensitivity labels to assets in Microsoft Purview Data Map, and create or select the labels you want to apply to your data.
- Set up and complete a scan on the test data in each data source. For more information, see Manage data sources in Microsoft Purview and Create a scan rule set.
- Permissions depending on your current experience:
  - If using the classic Microsoft Purview Catalog, you need the data curator role or insight reader role.
  - If using the new Microsoft Purview Unified Catalog, you need data health reader permissions.

For more information, see Manage data sources in Microsoft Purview and Automatically label your data in Microsoft Purview.

Use Data Estate Insights for sensitivity labels

Classifications are similar to subject tags. Use them to mark and identify data of a specific type that you find within your data estate during scanning.

Sensitivity labels enable you to state how sensitive certain data is in your organization. For example, a specific project name might be highly confidential within your organization, while that same term isn't confidential to other organizations.

Classifications match directly, such as a social security number, which has a classification of **Social Security Number**.

In contrast, sensitivity labels are applied when one or more classifications and conditions are found together. In this context, conditions refer to all the parameters that you can define for unstructured data, such as **proximity to another classification**, and **% confidence**.

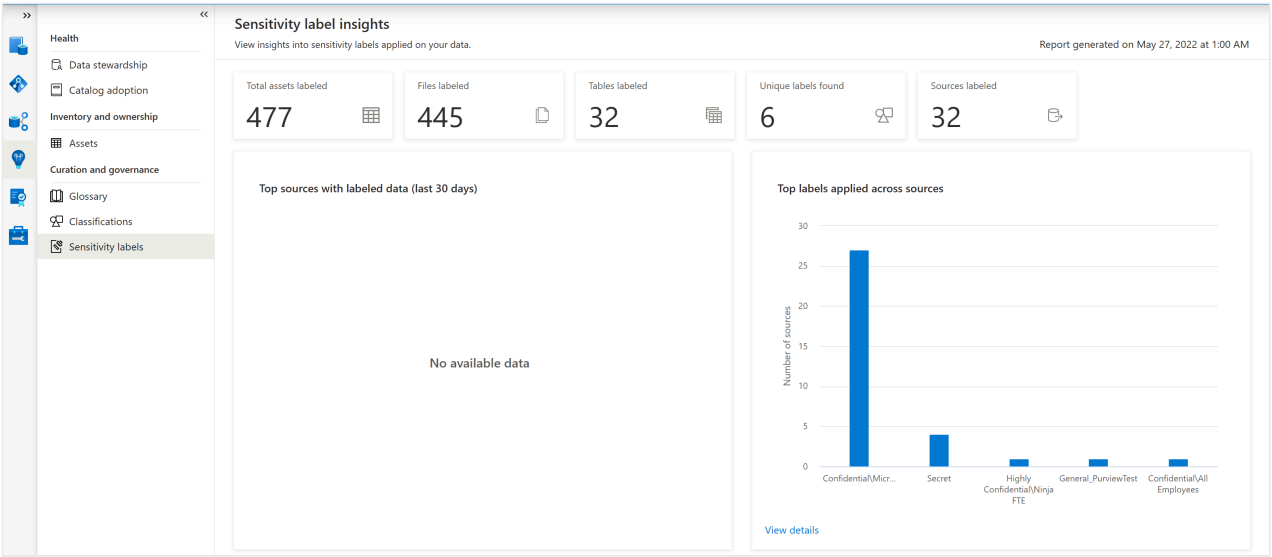
Data Estate Insights uses the same classifications, also known as sensitive information types, as those used with Microsoft 365 apps and services. This enables you to extend your existing sensitivity labels to assets in Data Map.

> ⓘ **Note**
>
> After you scan your source types, wait a couple of hours for the **Classic sensitivity labels** report to reflect the new assets.

View sensitivity labeling insights

1. Open the Microsoft Purview portal.
2. If you're using the classic Microsoft Purview portal, select **Data estate insights**. If you're using the new Microsoft Purview portal open **Unified Catalog**, select **Health Management**, then select **Reports**.
3. Select the **Classic sensitivity labels** report.

   > ⓘ **Note**
   >
   > If this report is empty, you might not have extended your sensitivity labels to Data Map. For more information, see Learn about sensitivity labels in Data Map.



The main **Sensitivity labeling insights** page displays the following areas:

Expand table

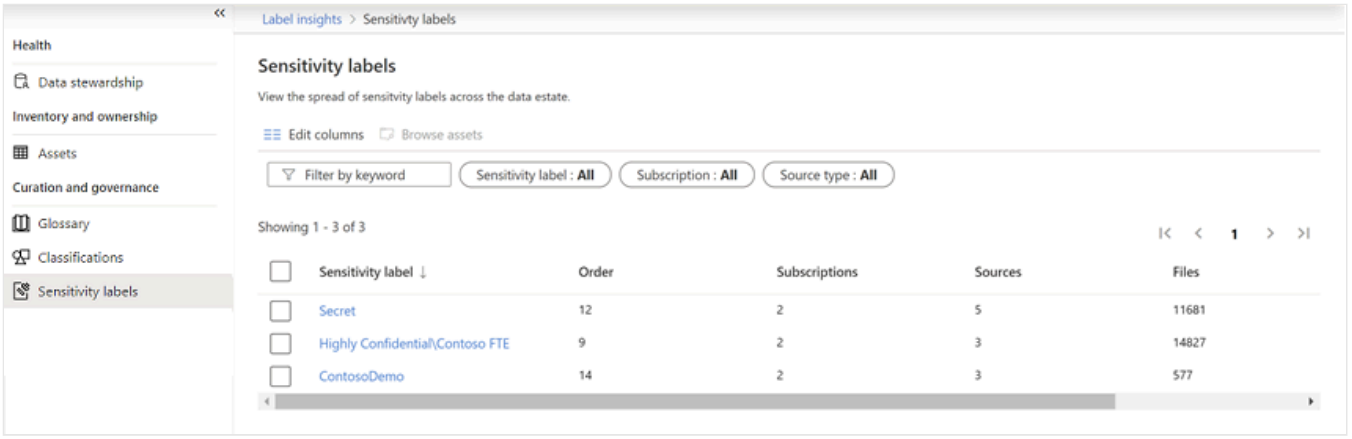
Area	Description
Overview of sources with sensitivity labels	Displays tiles that provide: - The number of subscriptions found in your data. - The number of unique sensitivity labels applied on your data - The number of sources with sensitivity labels applied - The number of files and tables found with sensitivity labels applied
Top sources with labeled data (last 30 days)	Shows the trend, over the past 30 days, of the number of sources with sensitivity labels applied.
Top labels applied across sources	Shows the top labels applied across all of your Microsoft Purview data resources.
Top labels applied on files	Shows the top sensitivity labels applied to files in your data.
Top labels applied on tables	Shows the top sensitivity labels applied to database tables in your data.
Labeling activity	Displays separate graphs for files and tables, each showing the number of files or tables labeled over the selected time frame. **Default**: 30 days Select the **Time** filter above the graphs to select a different time frame to display.

Sensitivity labeling insights drilldown

In any of the following **Sensitivity labeling insights** graphs, select the **View more** link to drill down for more details:

- **Top labels applied across sources**
- **Top labels applied on files**
- **Top labels applied on tables**
- **Labeling activity > Labeled data**

For example:



To learn more, use the following options:

Expand table

Option	Description
Filter your data	Use the filters above the grid to filter the data shown, including the label name, subscription name, or source type. If you're not sure of the exact label name, enter part or all of the name in the **Filter by keyword** box.
Sort the grid	Select a column header to sort the grid by that column.
Edit columns	To display more or fewer columns in your grid, select **Edit Columns**, then select the columns you want to view or change the order. Select a column header to sort the grid by that column.
Drill down further	To drill down to a specific label, select a name in the **Sensitivity label** column to view the **Label by source** report. This report displays data for the selected label, including the source name, source type, subscription ID, and the numbers of classified files and tables.
Browse assets	To browse through the assets found with a specific label or source, select one or more labels or sources, depending on the report you're viewing, then select **Browse assets** above the filters.

Option	Description
	The search results display all of the labeled assets found for the selected filter. For more information, see Search the Microsoft Purview Unified Catalog.

Sensitivity label integration with Microsoft Purview Information Protection

Close integration with Microsoft Purview Information Protection means that you have direct ways to extend visibility into your data estate, and classify and label your data. For more information about sensitivity labels, see: Learn about sensitivity labels.

To extend sensitivity labels to your assets in Data Map, you must turn on this capability in the Microsoft Purview portal.

For more information, see Apply sensitivity labels to your data in Data Map.

Next steps

Learn how to use Data Estate Insights with sources below:

- Learn how to use Asset insights.
- Learn how to use Data Stewardship.
- Learn how to use Classification insights.
- Learn how to use Glossary insights.

Last updated on 11/14/2025

Managed attributes in Microsoft Purview Data Map

07/16/2025

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

Managed attributes are user-defined attributes that provide a business or organization level context to an asset. When applied to assets, managed attributes enable data consumers using the data catalog to gain context on the role an asset plays in a business.

Terminology

> **Important**
>
> Attributes for data assets are referred to by different names, depending on which portal you're using.
>
> - In the Microsoft Purview portal, https://purview.microsoft.com, they're called **Data asset attributes**.
> - In the classic Microsoft Purview governance portal, https://web.purview.azure.com, they're called **Managed attributes**.
>
> This article explains working with **Managed attributes** in the classic governance portal (although the process for working with data asset attributes in the Microsoft Purview portal is the same).

- **Managed attribute:** A set of user-defined attributes that provide a business or organization level context to an asset. A managed attribute has a name and a value. For example, 'Department' is an attribute name and 'Finance' is its value.
- **Attribute group:** A grouping of managed attributes that allow for easier organization and consumption.

Create managed attributes

1. Access managed attributes by following the steps corresponding to the portal you're using:

    - In the Microsoft Purview portal (https://purview.microsoft.com):
        a. Under **Solutions**, select **Unified Catalog**.
        b. Under **Catalog management**, select **Custom metadata**, then select the **Data asset attributes** tab.
    - In the classic Microsoft Purview governance portal (https://web.purview.azure.com):
        a. Open **Data Map**.
        b. Under the **Annotation management**, select **Managed attributes**.

2. Select **New**, then select either **New attribute** or **New attribute group**.

3. To create an attribute group, enter a name and a description.

4. Managed attributes have a name, attribute group, data type, and associated asset types. They also have a required flag that can only be enabled when created as part of creating a new attribute group. Associated asset types are the data asset types you can apply the attribute to. For example, if you select "Azure SQL Table" for an attribute, it will be applied to Azure SQL Table assets, but not Azure Synapse Dedicated Table assets.

5. Select **Create** to save your attribute.

These attributes will be automatically applied to the specified resources without a value. You'll need to apply a value for each asset.

Required managed attributes

When you create a managed attribute as part of a managed attribute group, you can add the **required** flag. The required flag means that a value must be provided for this managed attribute. When you edit a data asset, the required attribute must be filled out before you can close the editor.

> **Important**
>
> - You can only add the **required** flag when you create an attribute during the process of creating an attribute group. You can't add the **required** flag while creating a new attribute outside of an attribute group.
> - If an existing attribute doesn't have the **required** flag, you can't add it during the editing process.

- Once you add the **required** flag, you can't remove it from the attribute.

1. Open Data Map and navigate to **Managed attributes** in the **Annotation management** section.
2. Select **New** and select **Attribute group**.
3. Select **New attribute**.
4. Fill out your attribute details, and select **Mark as required**.
5. Select **Apply** and finish adding other attributes to complete your attribute group.

Add value for managed attribute

Once a managed attribute has been created, you'll need to add a value for each of your assets. You can add values to your assets by:

1. Search for your data asset.
2. On the overview for your asset, see the **Managed attributes** section with all attributes that have values. (You can see attributes without values by using the **Show attributes without a value** toggle.)
3. Select the **Edit** button.
4. Under **Managed attributes**, add values for each of your attributes.
5. If any attributes are **Required**, you won't be able to save until you add a value for that attribute.

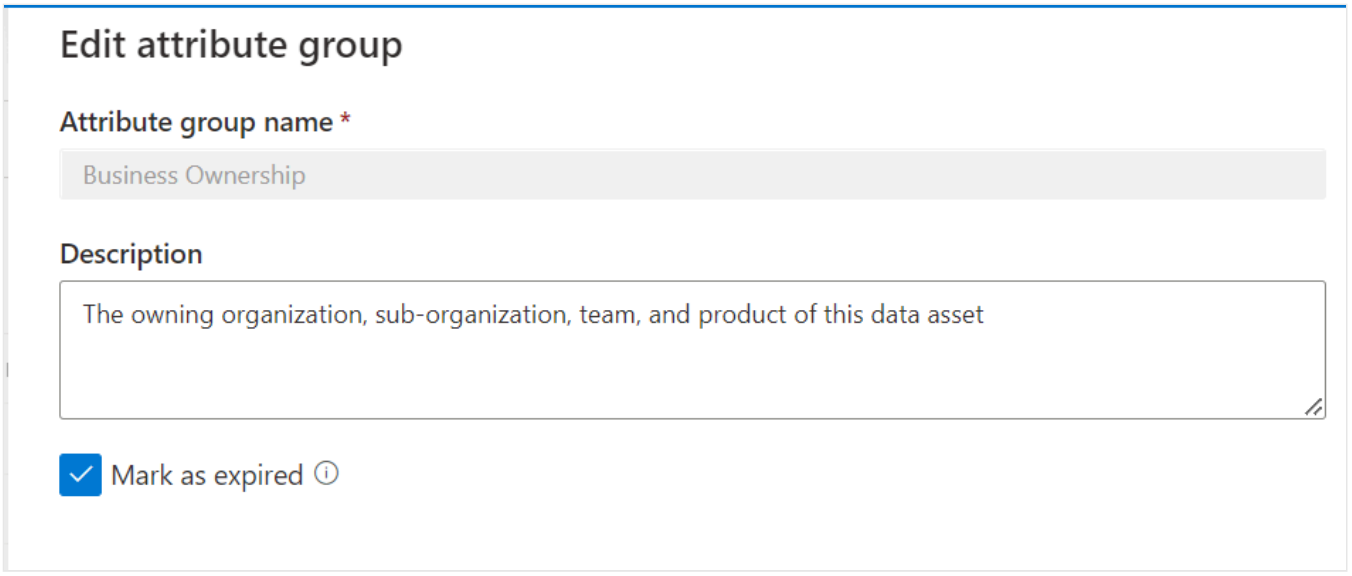
Expiring managed attributes

In the managed attribute management experience, managed attributes can't be deleted, only expired. Expired attributes can't be applied to any assets and are, by default, hidden in the user experience. By default, expired managed attributes aren't removed from an asset. If an asset has an expired managed attribute applied, it can only be removed, not edited.

Both attribute groups and individual managed attributes can be expired. To mark an attribute group or managed attribute as expired, select the **Edit** icon.

Name	Description	Type	Associated asset types
Business Ownership	The owning organization, sub-organization, team, a...	-	-
Organization	-	Single choice	Azure SQL Table

Select **Mark as expired** and confirm your change. Once expired, attribute groups and managed attributes can't be reactivated.



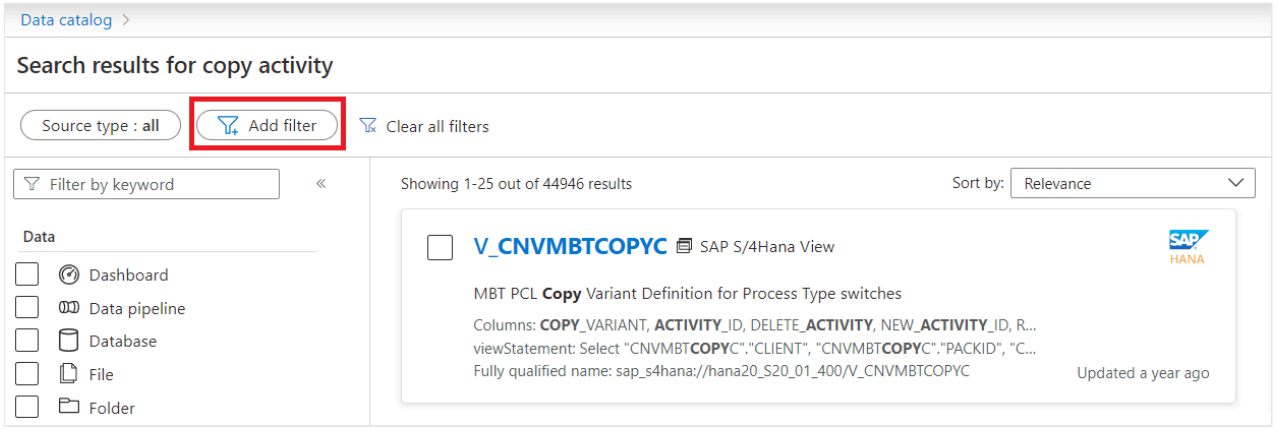
Create managed attributes using APIs

Managed attributes can be programmatically created and applied using the business metadata APIs in Apache Atlas 2.2. For more information, see the Use Atlas 2.2 APIs tutorial.

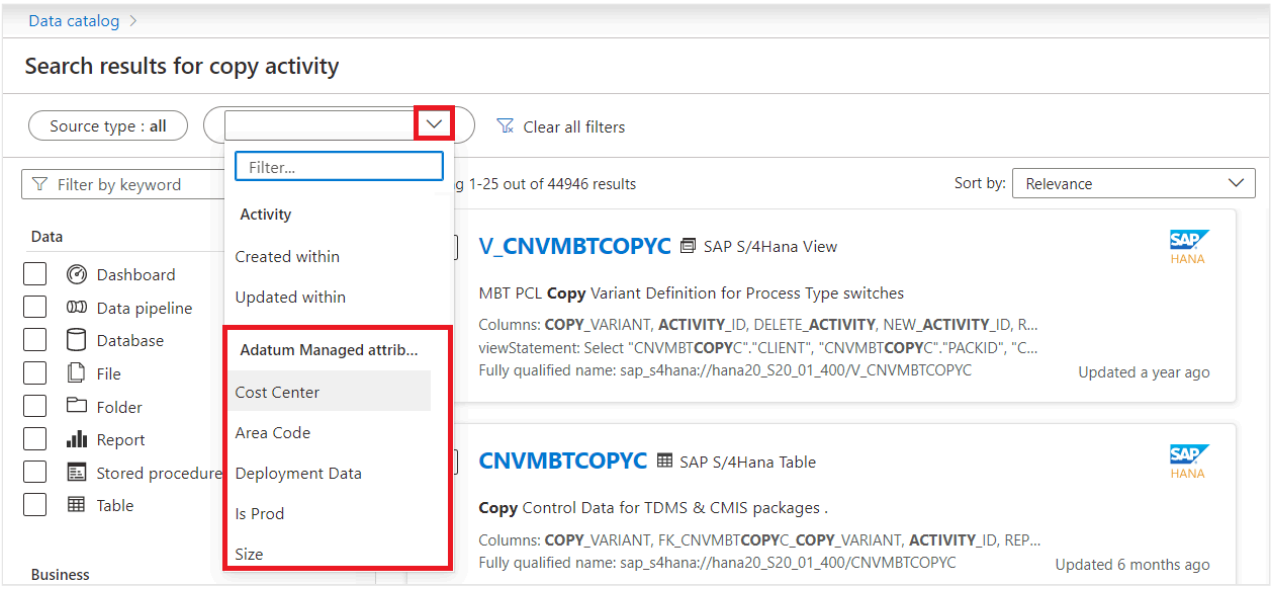
Searching by managed attributes

Once you have created managed attributes, you can refine your data catalog searches using these attributes.

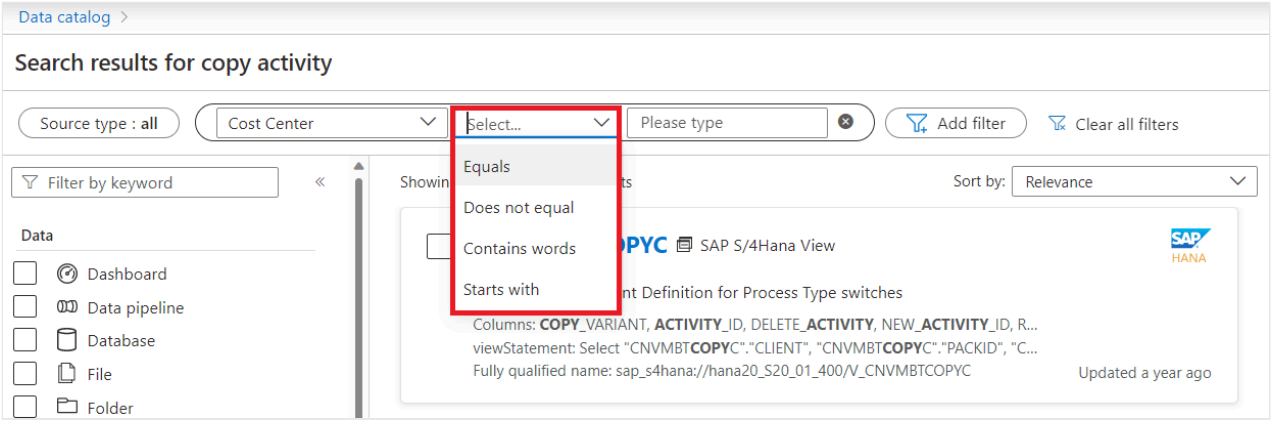
1. In a data catalog search, to refine by a managed attribute, first select **Add filter** at the top of the search.



2. Select the drop-down, scroll to your list of managed attributes, and select one.



3. Select your operator, which will be different based on the kinds of values allowed by the attribute. In this example, we've selected Cost Center, which is a text value, so we can compare Cost Center with the text we'll enter.

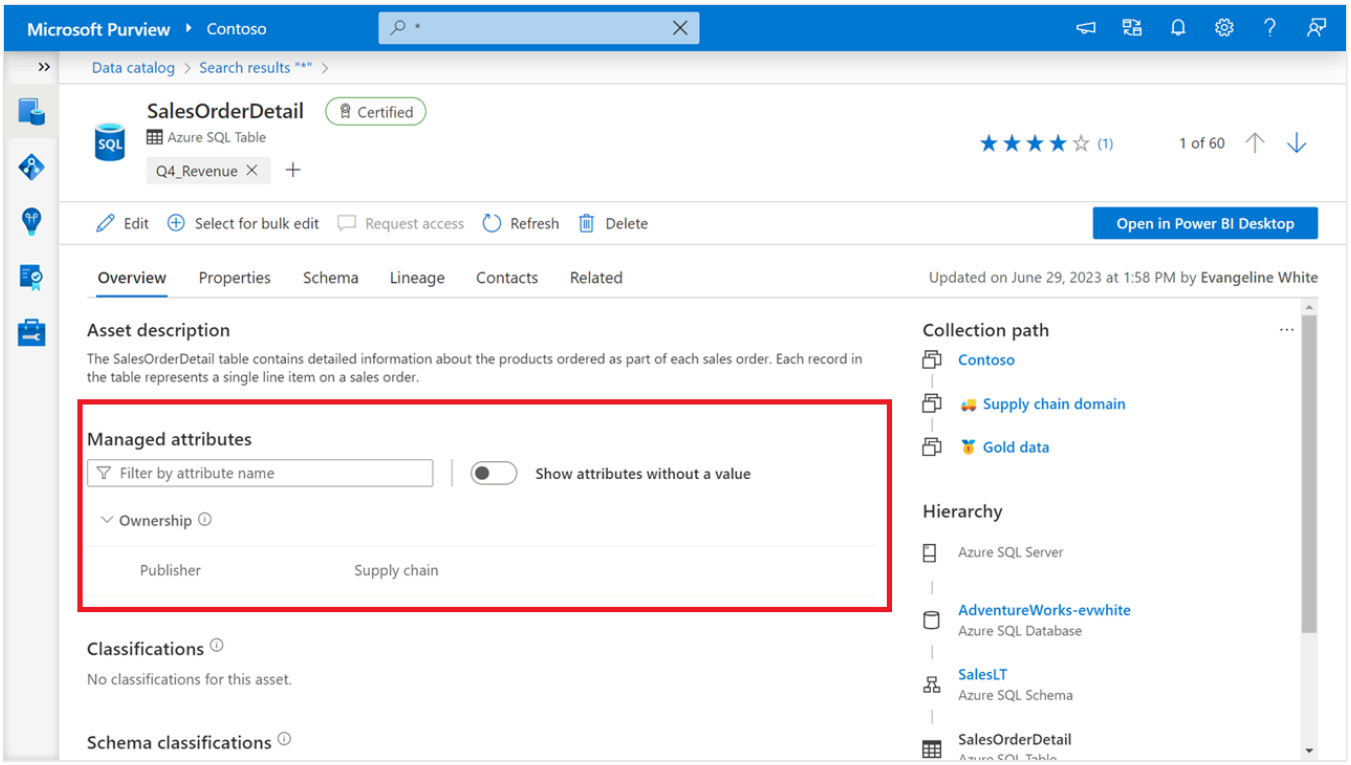


4. Enter your values, and the search runs with your new filter.

Best practices for using managed attributes

Use managed attributes to extend the fields available for an asset in the Microsoft Purview Data Catalog. Managed attributes are key-value pairs that add structured metadata to your data catalog. When Microsoft Purview scans data, it adds technical information about the data like data type, classification, etc. If you want to add more fields, you need to define managed attributes.

In the following example, a managed attribute lets me tag tables with the department that publishes them. I use a managed attribute because I want to make sure assets are always tagged in exactly the same way with this information. I also want to filter by the publisher field when I search for data.



The managed attribute in this example helps people quickly find all data published by the supply chain team, but doesn't help someone understand the definition of a publisher or what it means if supply chain is the publisher of the data. For any information that needs a business explanation, use terms.

Known limitations

Below are the known limitations of the managed attribute feature as it currently exists in Microsoft Purview.

- Managed attributes can only be expired.
- Managed attributes can't be applied via the bulk edit experience.
- After creating an attribute group, you can't edit the name of the attribute group.
- After creating a managed attribute, you can't update the attribute name, attribute group, or the field type.
- A managed attribute can only be marked as required during the creation of an attribute group.

Asset management in the Microsoft Purview Data Catalog

Article • 02/24/2025

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

This article discusses how assets are displayed in the Microsoft Purview Data Catalog, and all the features and details available to them. It describes how you can view relevant information or take action on assets in your catalog.

Prerequisites

- Set up your data sources and scan the assets into your catalog.
- *Or* Use the Microsoft Purview Atlas APIs to ingest assets into the catalog.

Discover assets

You can discover your assets in the Microsoft Purview Data Catalog by either:

- Browsing the data catalog
- Searching the data catalog

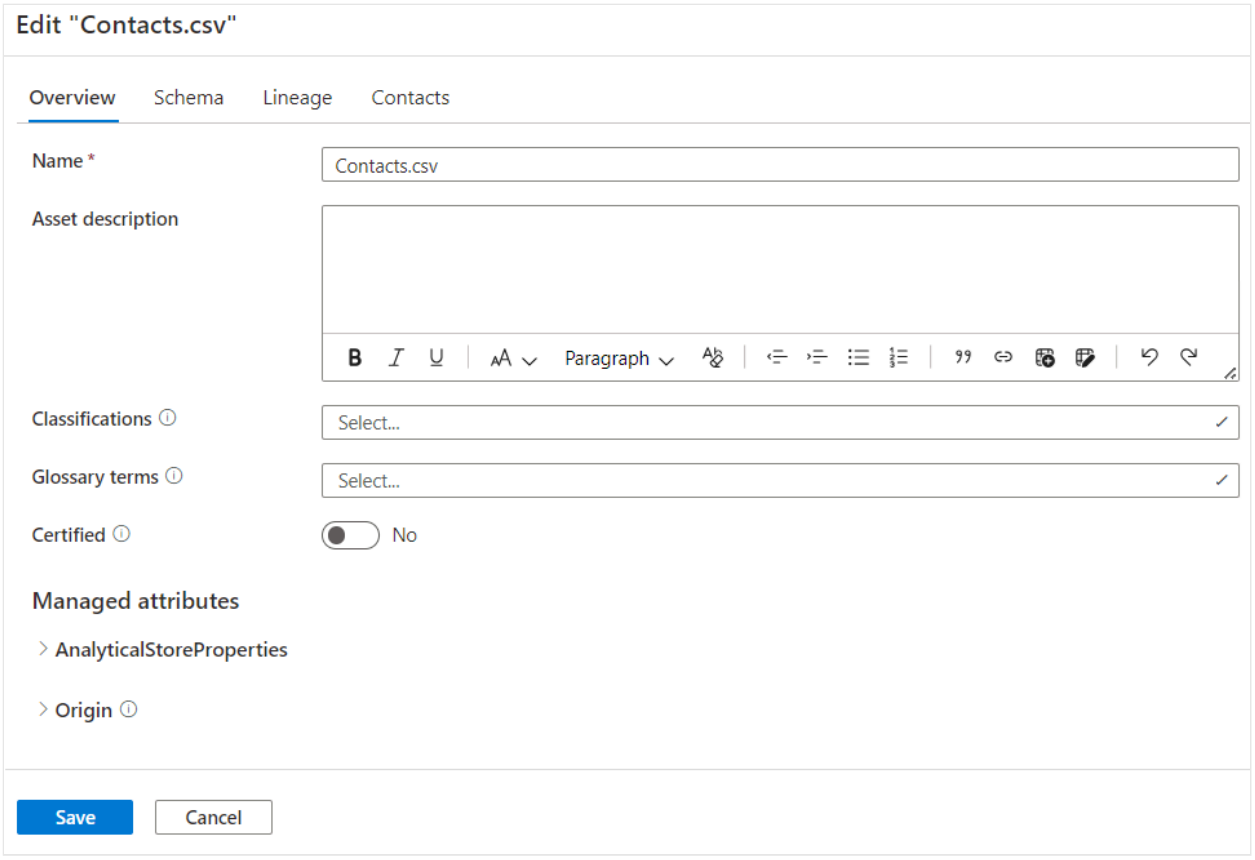
Once you've discovered an asset, select it to view the asset details page, which contains all the asset details, and more actions you can take.

Edit assets

To edit an asset, other than adding a rating, you will need the data curator role on the collection where the asset is housed.

To edit assets you can either:

- Select multiple assets in the catalog to bulk edit assets
- Select a single asset and select the **Edit** button at the top of the asset details page



When you're editing an asset you can manage nearly all its information including:

- Overview: Manually update name, description, classification, and glossary terms.
- Schema: Manually update column name, classification, glossary terms, data types, and description.
- Lineage: Manually add lineage connections to other resources.
- Contacts: Add experts and owners users can contact for more information on this resource.

Scan behavior after editing assets

Microsoft Purview works to reflect the truth of the source system whenever possible. For example, if you edit a column and later it's deleted from the source table. A scan will remove the column metadata from the asset in Microsoft Purview.

Both column-level and asset-level updates such as adding a description, glossary term or classification don't impact scan updates. Scans will update new columns and classifications regardless if these changes are made.

Delete asset

If you're a data curator on the collection containing an asset, you can delete an asset by selecting the delete icon under the name of the asset.

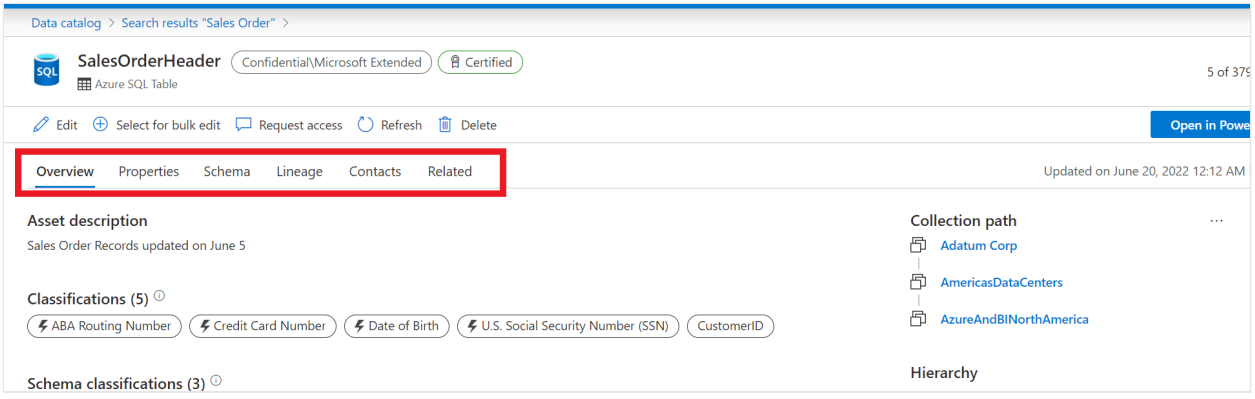
> **Important**
>
> Deleting the asset using the delete button will also delete all the child assets under the hierarchy. For example, if you attempt to delete a storage account asset in your catalog, all the assets under it - containers, folders and files - will also be deleted together.

Any asset you delete using the delete button is permanently deleted in Microsoft Purview. However, if you run a **full scan** on the source from which the asset was ingested into the catalog, then the asset is reingested and you can discover it using the Microsoft Purview Data Catalog.

When you run an incremental scan on the source, the **deleted asset won't get re-ingested** into the catalog unless the asset is modified by an end user since the previous run of the scan. For example, say you manually delete a SQL table from the Microsoft Purview Data Map. Later, a data engineer adds a new column to the source table. When Microsoft Purview scans the database, the table will be reingested into the data map and be discoverable in the data catalog.

Asset details page

At the top of the asset details page there are several tabs:



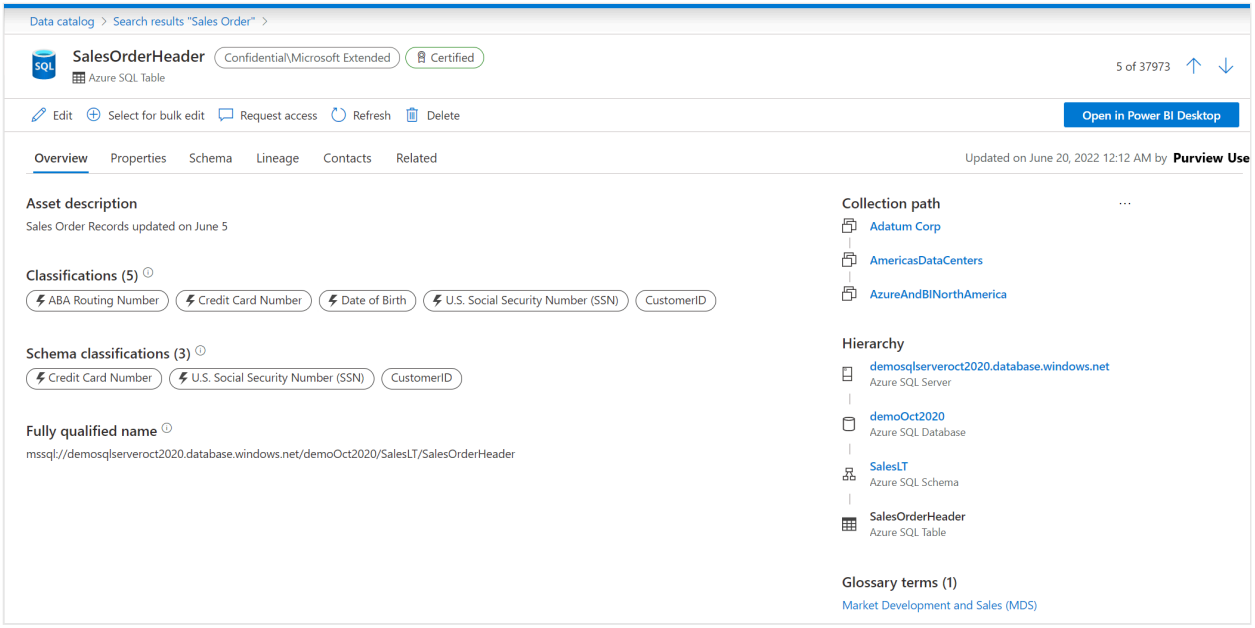
- **Overview** - An asset's basic details like description, classification, hierarchy, and glossary terms.
- **Properties** - The technical metadata and relationships discovered in the data source.
- **Schema** - The schema of the asset including column names, data types, column level classifications, terms, and descriptions are represented in the schema tab.
- **Lineage** - This tab contains lineage graph details for assets where it's available.
- **Contacts** - Every asset can have an assigned owner and expert that can be viewed and managed from the contacts tab.

- **Related** - This tab lets you navigate through the technical hierarchy of assets that are related to the current asset you're viewing.

> **Tip**
>
> You can information on all these tabs by editing the asset.

Asset overview

The overview section of the asset details gives a summarized view of an asset. The sections that follow explains the different parts of the overview page.

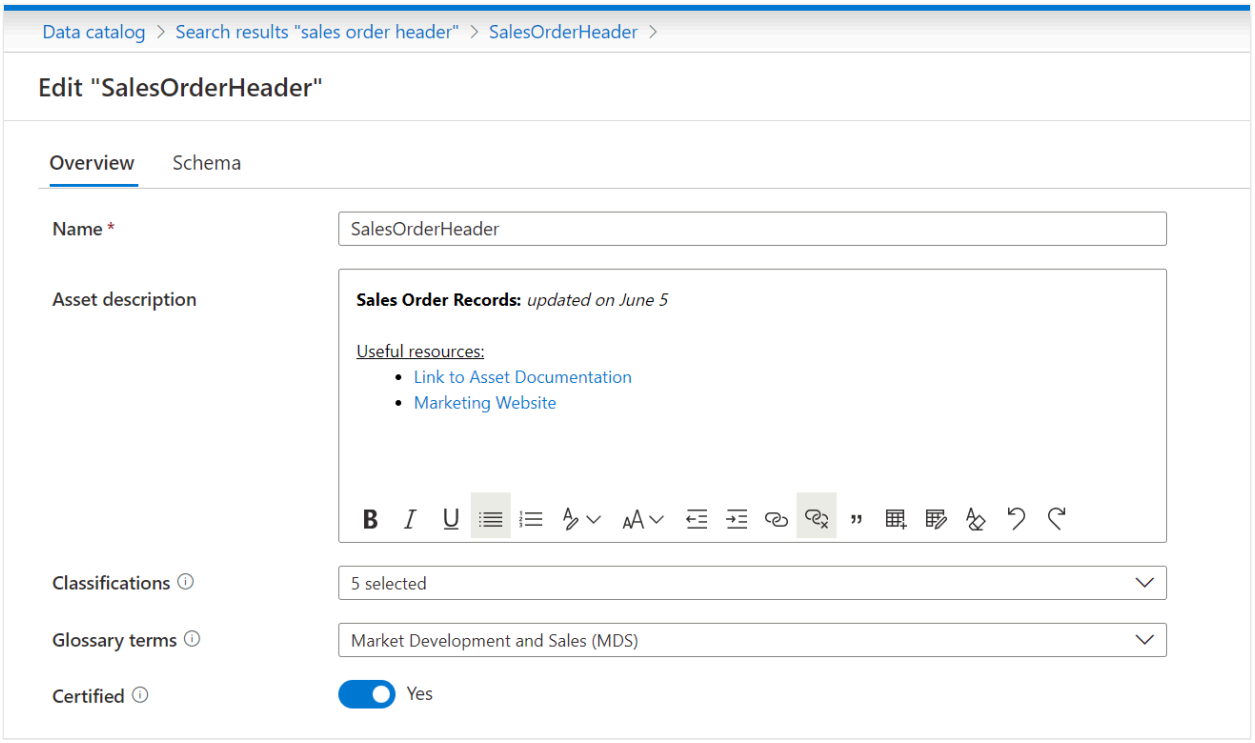


Asset description

An asset description gives a synopsis of what the asset represents. You can add or update an asset description by editing the asset.

Adding rich text to a description

Microsoft Purview enables users to add rich formatting to asset descriptions such as adding bolding, underlining, or italicizing text. Users can also create tables, bulleted lists, or hyperlinks to external resources.



Below are the rich text formatting options:

Expand table

Name	Description	Shortcut key
Bold	Make your text bold. Adding the '*' character around text will also bold it.	Ctrl+B
Italic	Italicize your text. Adding the '_' character around text will also italicize it.	Ctrl+I
Underline	Underline your text.	Ctrl+U
Bullets	Create a bulleted list. Adding the '-' character before text will also create a bulleted list.	
Numbering	Create a numbered list Adding the '1' character before text will also create a bulleted list.	
Heading	Add a formatted heading	
Font size	Change the size of your text. The default size is 12.	
Decrease indent	Move your paragraph closer to the margin.	
Increase indent	Move your paragraph farther away from the margin.	
Add hyperlink	Create a link in your document for quick access to web pages and files.	

Name	Description	Shortcut key
Remove hyperlink	Change a link to plain text.	
Quote	Add quote text	
Add table	Add a table to your content.	
Edit table	Insert or delete a column or row from a table	
Clear formatting	Remove all formatting from a selection of text, leaving only the normal, unformatted text.	
Undo	Undo changes you made to the content.	Ctrl+Z
Redo	Redo changes you made to the content.	Ctrl+Y

> **Note**
>
> Updating a description with the rich text editor updates the `userDescription` field of an entity. If you have already added an asset description before the release of this feature, that description is stored in the `description` field. When overwriting a plain text description with rich text, the entity model will persist both `userDescription` and `description`. The asset details overview page will only show `userDescription`. The `description` field can't be edited in the Microsoft Purview studio user experience.

Classifications

Classifications identify the kind of data being represented by an asset or column such as "ABA routing number", "Email Address", or "U.S. Passport number". These attributes can be assigned during scans or added manually. For a full list of classifications, see the supported classifications in Data Map. You can see classifications assigned both to the asset and columns in the schema from the overview page.which you can also view as part of the schema.

Glossary terms

Glossary terms are a managed vocabulary for business terms that can be used to categorize and relate assets across your organization. For more information, see the business glossary page. You can view the assigned glossary terms for an asset in the

overview section. If you're a data curator on the asset, you can add or remove a glossary term on an asset by editing the asset.

Collection hierarchy

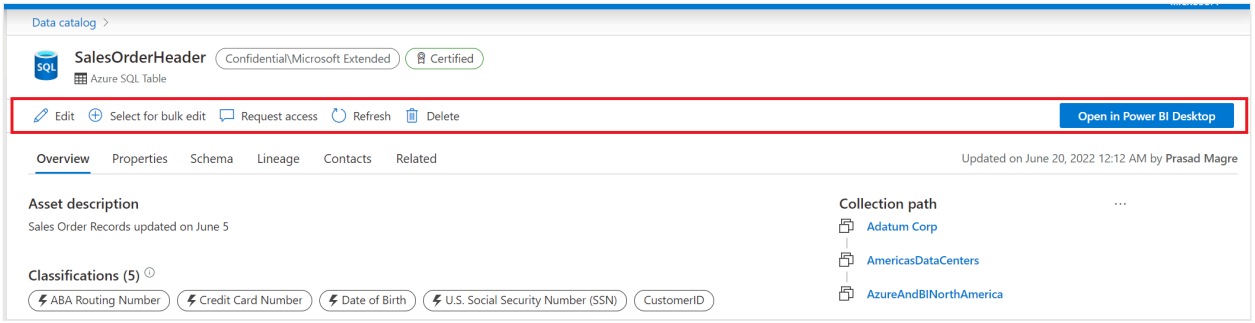
In Microsoft Purview, collections organize assets and data sources. They also manage access across the Microsoft Purview governance portal. You can view an assets containing collection under the **Collection path** section.

Asset hierarchy

You can view the full asset hierarchy within the overview tab. As an example: if you navigate to a SQL table, then you can see the schema, database, and the server the table belongs to.

Asset actions

Below are a list of actions you can take from an asset details page. Actions available to you vary depending on your permissions and the type of asset you're looking at. Available actions are generally available on the global actions bar.



Request access to data

If a self-service data access workflow has been created, you can request access to a desired asset directly from the asset details page! To learn more about Microsoft Purview's data policy applications, see how to enable data policy enforcement.

Open in Power BI

Microsoft Purview makes it easy to work with useful data you find the data catalog. You can open certain assets in Power BI Desktop from the asset details page. Power BI Desktop integration is supported for the following sources.

- Azure Blob Storage
- Azure Cosmos DB
- Azure Data Lake Storage Gen2
- Azure Dedicated SQL pool (formerly SQL DW)
- Azure SQL Database
- Azure SQL Managed Instance
- Azure Synapse Analytics
- Azure Database for MySQL
- Azure Database for PostgreSQL
- Oracle DB
- SQL Server
- Teradata

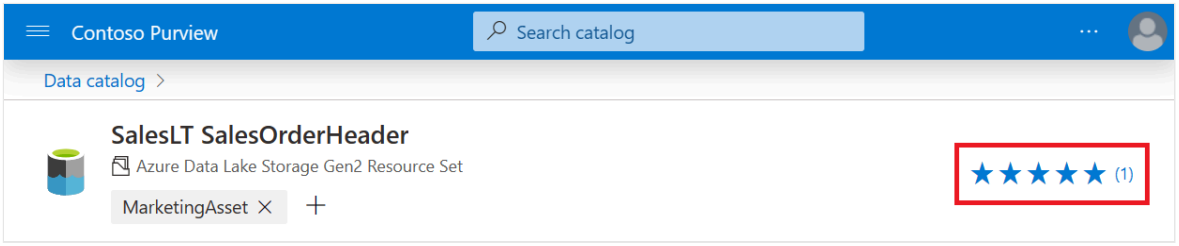
Ratings

Assets can be rated by all users with read access, or better, to that asset in Microsoft Purview. Ratings allow users to give an asset a rating from 1 to 5 stars, and leave a comment about the asset.

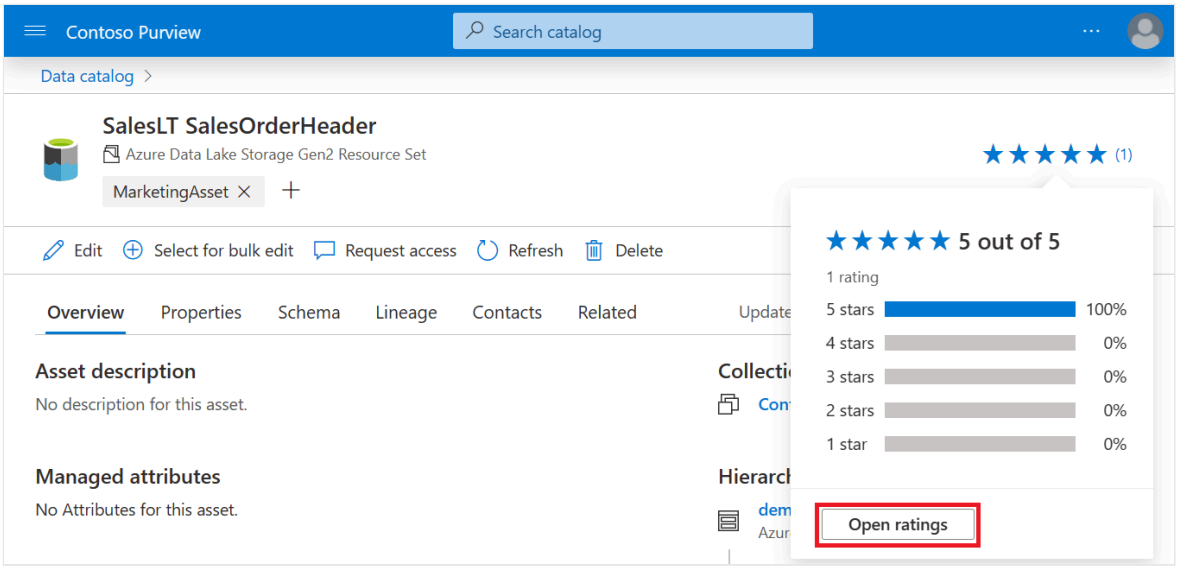
These ratings can be seen by all users with read access, and rating can be added as a filter when searching the data catalog so users can find assets with a certain rating.

View ratings

1. Search or browse for your asset in Microsoft Purview and select it.
2. In the header of the asset you can see a rating, which will show an aggregate star rating of the asset, and the number of reviews.

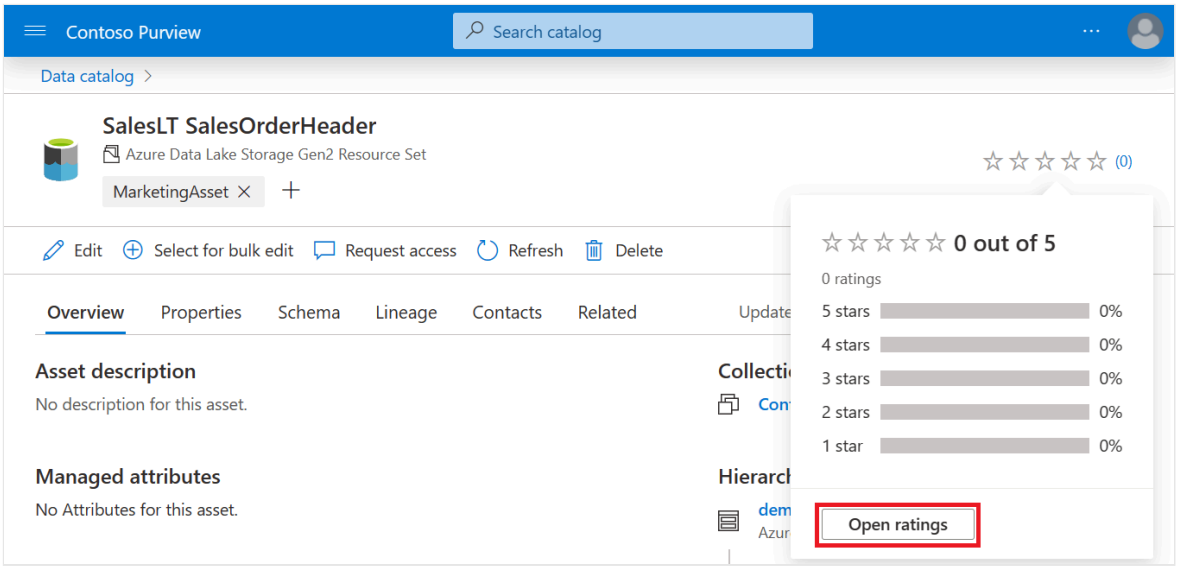


3. To see a percentage breakdown of the ratings, select the rating.
4. To see specific ratings and their comments, or to add your own rating, select **Open ratings**.

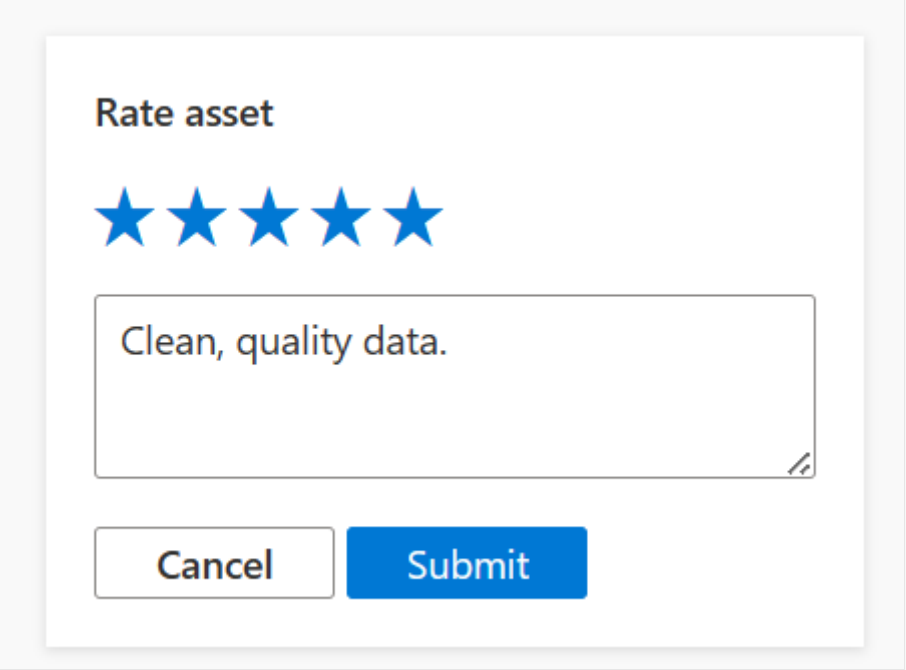


Add a rating to an asset

1. Search or browse for your asset in Microsoft Purview and select it.
2. Select the ratings button in the asset's header.
3. Select the **Open ratings** button.

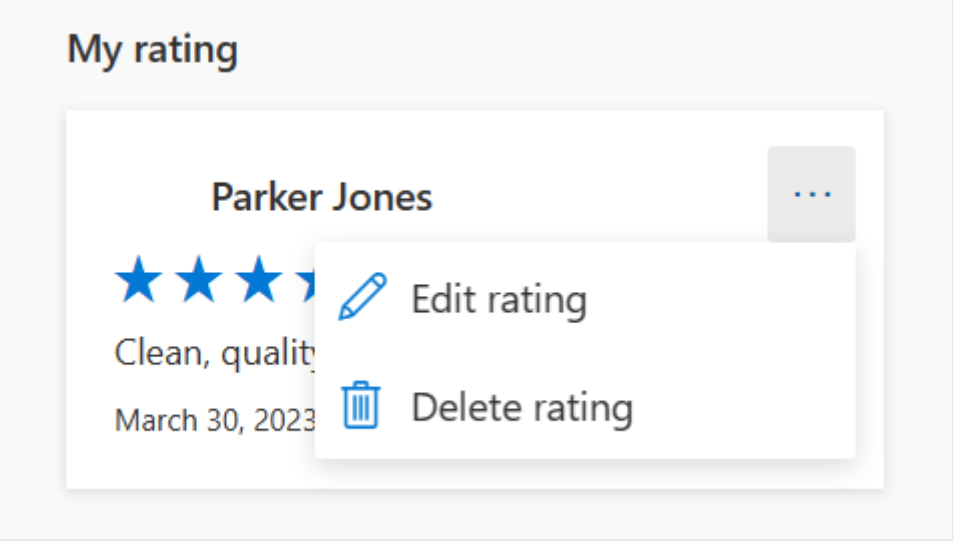


4. Choose a star rating, add a comment, and select **Submit**.



Edit or delete your rating

1. Select the ratings button in the asset's header.
2. Select the **Open ratings** button.
3. Under **My rating** select the ellipsis button in your rating.



4. To delete your rating, select **Delete rating**.
5. To edit your rating, select **Edit rating**, then update your score and comment and select **Submit**.

Tags

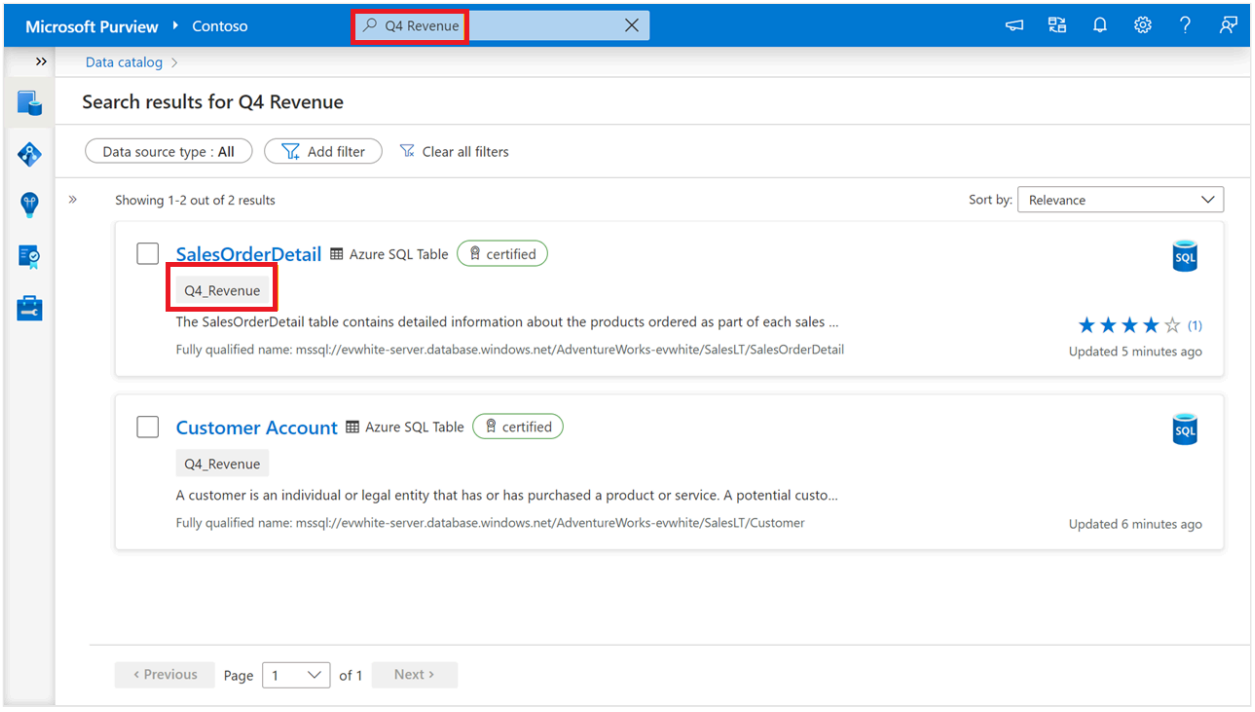
Asset can be tagged by users with data curator permissions or better, and any users with reader permissions on these assets in Microsoft Purview can see these tags. Users can

add tags as a filter when searching the data catalog, so users can see all assets with certain tags.

> **Note**
>
> Tag limitations:
>
> - An asset can only have up to 50 tags
> - Tags can only be 50 characters
> - Allowed characters: numbers, letters, -, and _

Use tags when you want to quickly label your data assets without the need for consistency or control. Tags are simple keywords or phrases that can be applied to data assets to provide informal metadata. They're useful for categorizing data assets, making them easier to discover and understand. They're also a great way to see how your data consumers describe your data so you can incorporate this language into your business glossary over time.

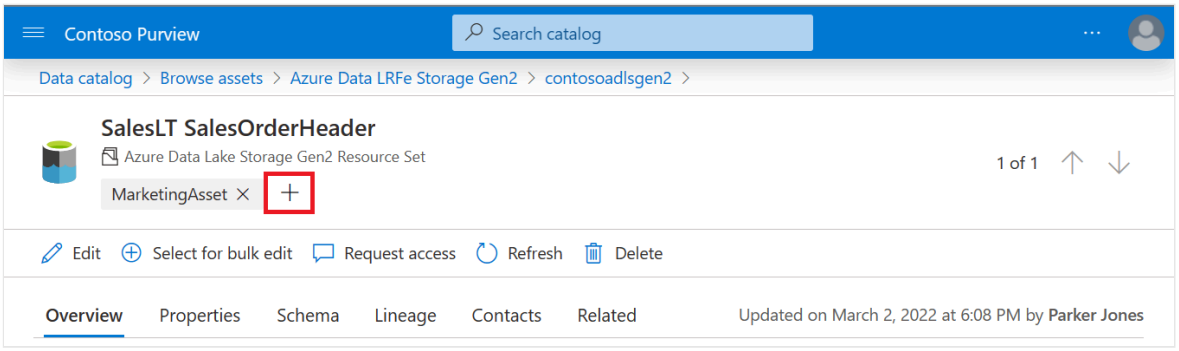
In the following example, a few assets with Q4 Revenue so users can easily find the data assets to use for a new report with this information. Searching the keyword returns all data with that tag applied:



Add a tag to an asset

If you have data curator permissions Microsoft Purview, you can add a tag to an asset by:

1. Search or browse for your asset in Microsoft Purview and select it.
2. Select the + **Add Tag** button under the asset's name.

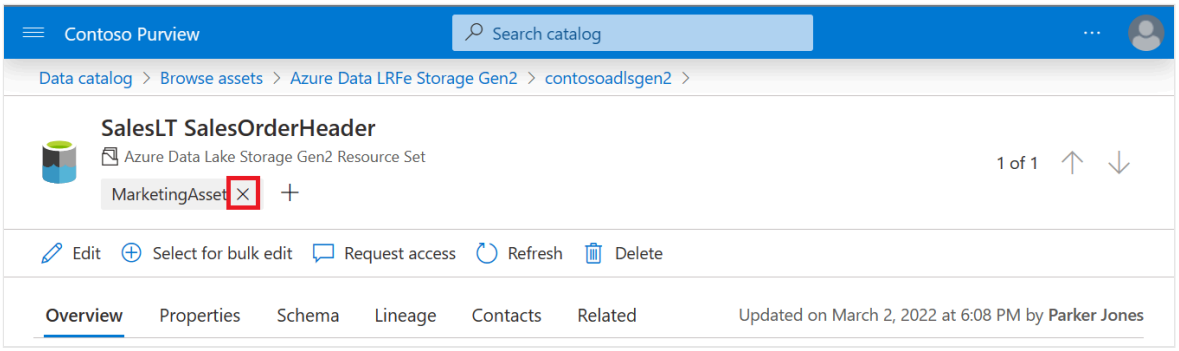


3. Select an existing available tag, or input a new tag.

Remove a tag from an asset

If you have data curator permissions Microsoft Purview, you can remove a tag from an asset by:

1. Search or browse for your asset in Microsoft Purview and select it.
2. Select the **X** button next to an existing tag under the asset's name.



3. Confirm the removal of the tag.

Duplicate assets

If you notice duplicate assets in your Microsoft Purview Data Catalog, review the asset normalization documentation. Microsoft Purview normalizes assets to prevent duplication, but not all possible scenarios are covered.

Compare the fully qualified asset names for duplicate assets, and update ingestion points to resolve capitalization or character differences. Then, delete the duplicated asset in your catalog.

Next steps

- Browse the Microsoft Purview Data Catalog
- Search the Microsoft Purview Data Catalog

- Asset normalization

Feedback

Was this page helpful? Yes No

Metadata curation in Microsoft Purview Data Map

Article • 05/20/2025

This article describes how users can update and curate the metadata of an asset in the Microsoft Purview Data Map. Data assets in the map automatically have technical metadata like name and schema, but users can also add other information like descriptions and contacts. This extra information allows other users to better understand data assets so they can discover, govern, and use them.

Who can update metadata?

- Users who have the owner role or write permissions on available, live Azure resources can access and edit those resources in free and enterprise account types. For more information about data owner metadata curation, see the data owner curation section on this page.
- Users who have data curator permissions on a collection in a Microsoft Purview account can edit the metadata details for all data assets in that collection. This is only available in the enterprise version of Microsoft Purview.
- Users who are using the classic Microsoft Purview Unified Catalog and have been assigned to the the data curator role group can edit the metadata details for all data assets in a Microsoft Purview account. For more information, see the role group curation section on this page.

Data owner curation

If a user is a data owner on an available, live Azure resource, they're able to access and manage that asset's metadata in Microsoft Purview. In this case, a 'data owner' isn't limited to users who have the Azure RBAC owner role on a resource. It also applies to users who have certain write permissions on the available Azure resources. The permissions needed are listed in the supported live sources section.

With those permissions, data owners able to update these aspects of a data asset:

- Description
- Contacts
- Column descriptions

> **Tip**
>
> Any of these updates are tracked by Microsoft Purview and can be viewed in **Data Map history**.

Role group curation

If a user has been added to the data curator role group in the classic Microsoft Purview Unified Catalog, they're able to access and manage asset metadata in Microsoft Purview depending on the version of Microsoft Purview they're using.

With the free version of Microsoft Purview governance solutions you can update these aspects:

- Display name
- Description
- Contacts
- Column descriptions
- Column names
- Column data types
- Source classifications
- Column classifications

With the enterprise version of Microsoft Purview, after extending live view, or registering and scanning an asset, you can edit all the other aspects of that asset, including adding glossary terms.

> **Tip**
>
> Any of these updates are tracked by Microsoft Purview and can be viewed in **Data Map history**.

Collection-level role curation

If a user has been given the data curator role on a collection in Data Map, they're able to access and manage metadata for any assets in that collection.

> **Note**

A user can only be assigned the data curator role on a collection in the enterprise version of Microsoft Purview.

With the data curator role, a user can update aspects of their data assets depending on the kind of asset it is.

If an asset is only available as a live view asset you can edit:

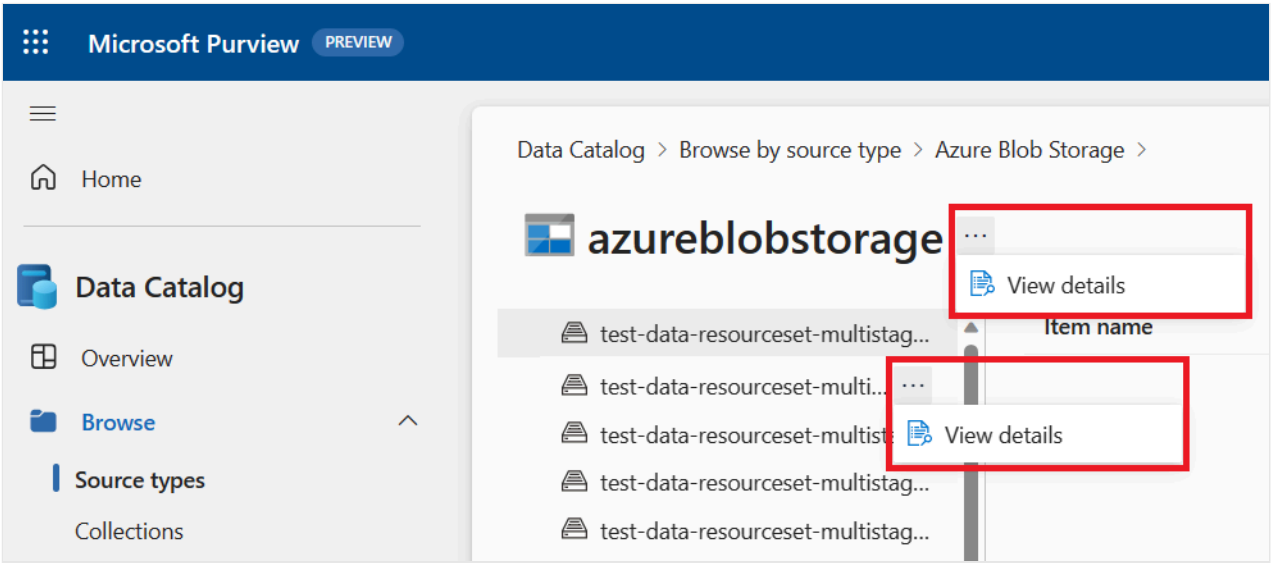
- Display name
- Description
- Contacts
- Column descriptions
- Column names
- Column data types
- Source classifications
- Column classifications

Once an asset has had live view extended for all users, or if an asset has been registered and scanned, you can edit all the other aspects of an asset, including adding glossary terms.

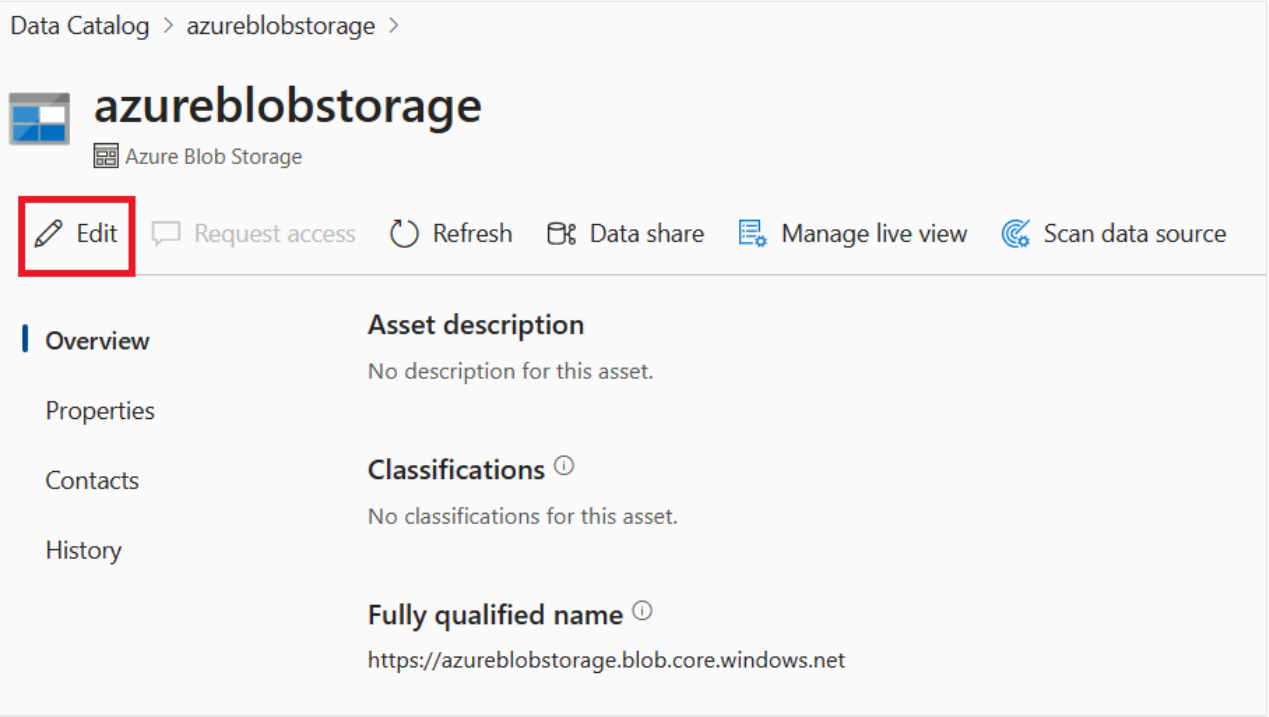
How to edit assets

1. In the Microsoft Purview portal (or the classic Microsoft Purview governance portal) search or browse for a data source or asset.

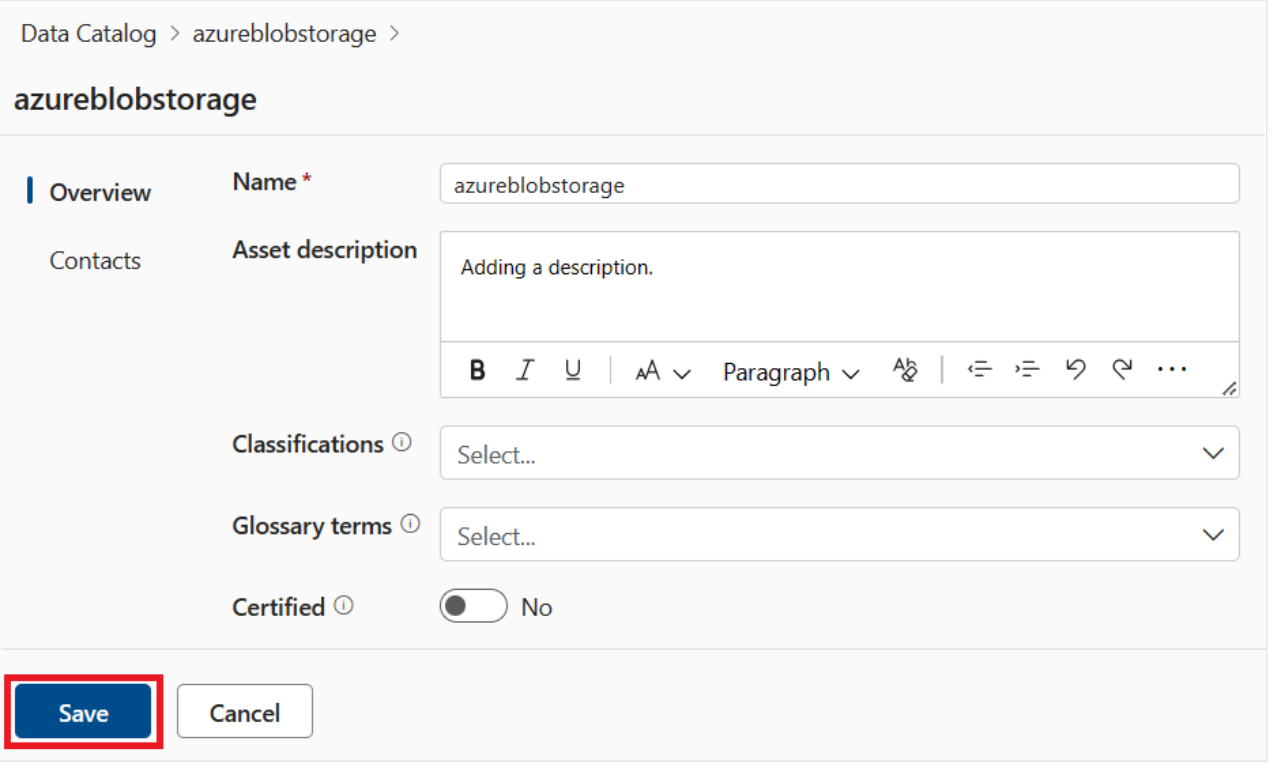
2. Select the asset to view details.



3. Select **Edit**.



4. Make any changes to the metadata, and select **Save** when you're finished.



For more information about editing assets, including rich text editing for descriptions, see the [manage assets in the Unified Catalog article]classic-catalog-asset-details.md).

Supported live sources

Expand table

Data source	Owner permission
Azure SQL Database	"Microsoft.Sql/servers/write", "Microsoft.Sql/servers/databases/write", "Microsoft.Authorization/roleAssignments/write"
Azure Blob Storage	"Microsoft.Storage/storageAccounts/write", "Microsoft.Authorization/roleAssignments/write"
Azure Data Lake Storage Gen2	"Microsoft.Storage/storageAccounts/write", "Microsoft.Authorization/roleAssignments/write"

Limitations

These are the current limitations for metadata curation in Microsoft Purview:

- When a user in the data curator role group edits a live mode asset for the first time, it will not work. If you edit a second time with a different edit than the first attempt, then it works.

Next steps

- What is live view?
- The new Microsoft Purview portal
- Roles and permissions

How to request access for a data asset

Article • 02/26/2025

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

> **Important**
>
> This feature is currently in preview. The Supplemental Terms of Use for Microsoft Azure Previews include additional legal terms that apply to Azure features that are in beta, in preview, or otherwise not yet released into general availability.

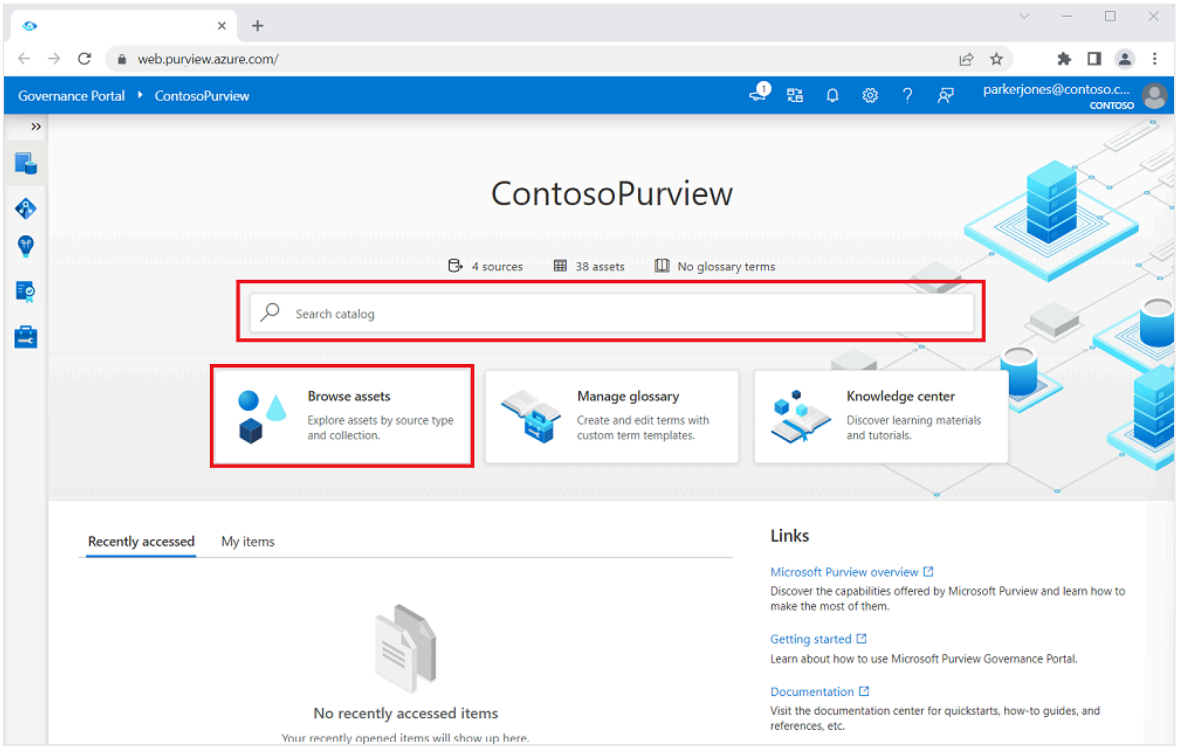
If you discover a data asset in the catalog that you would like to access, you can request access directly through Microsoft Purview. The request will trigger a workflow that will request that the owners of the data resource grant you access to that data source.

This article outlines how to make an access request.

> **Note**
>
> For this option to be available for a resource, a self-service access workflow needs to be created and assigned to the collection where the resource is registered. Contact the collection administrator, data source administrator, or workflow administrator of your collection for more information. Or, for information on how to create a self-service access workflow, see our self-service access workflow documentation.

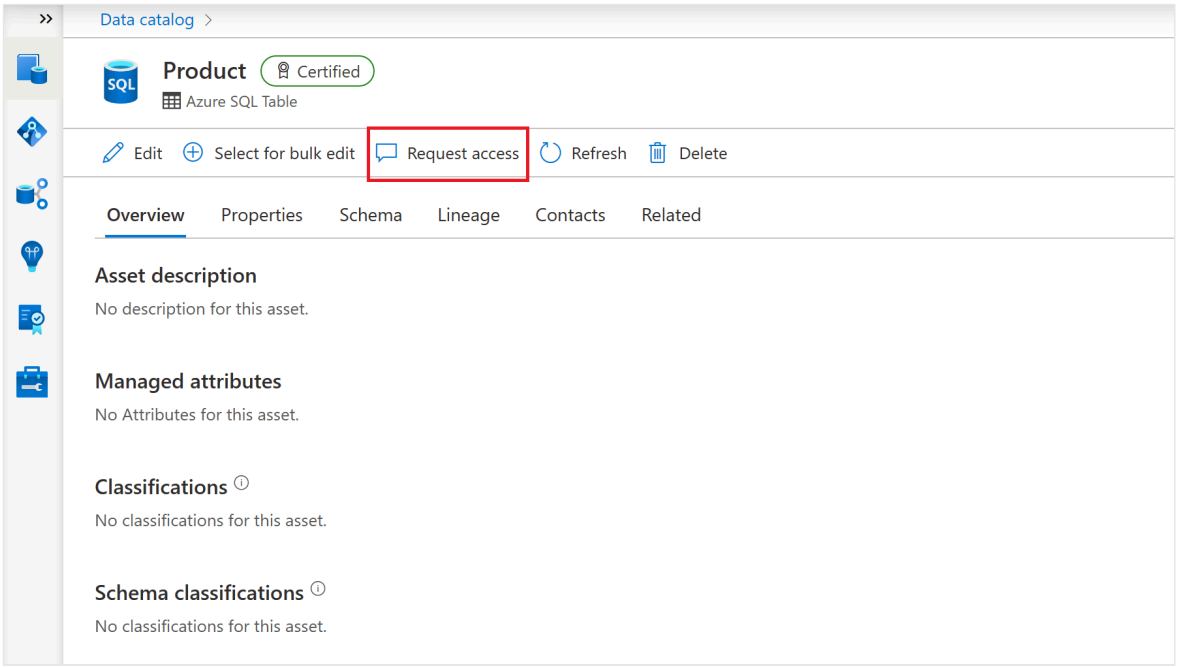
Request access

1. To find a data asset, use Microsoft Purview's search or browse functionality.



2. Select the asset to go to asset details.

3. Select **Request access**.



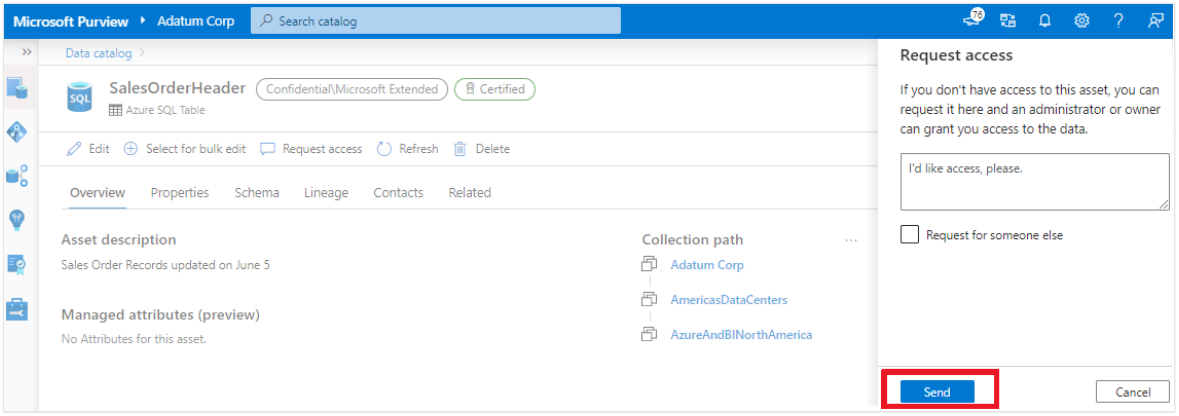
> **Note**
>
> If this option isn't available, a self-service access workflow either hasn't been created, or hasn't been assigned to the collection where the resource is registered. Contact the collection administrator, data source administrator, or workflow administrator of your collection for more information. Or, for

information on how to create a self-service access workflow, see our self-service access workflow documentation.

4. The **Request access** window will open. You can provide comments on why data access is requested.

5. Select **Send** to trigger the self-service data access workflow.

> ⓘ Note
>
> If you want to request access on behalf of another user, select the checkbox **Request for someone else** and populate the email id of that user.



> ⓘ Note
>
> A request access to resource set will actually submit the data access request for the folder one level up which contains all these resource set files.

6. Data owners will be notified of your request and will either approve or reject the request.

Next steps

- What are Microsoft Purview workflows
- Approval workflow for business terms
- Self-service data access workflow for hybrid data estates
- Self-service policies

Feedback

Was this page helpful? Yes No

Search for data assets

After data is scanned and ingested into the Microsoft Purview Data Map, data consumers need to easily find the data they need for their analytics or governance workloads. Data discovery can be time consuming because you might not know where to find the data that you want. Even after finding the data, you might have doubts about whether you can trust the data and take a dependency on it.

This article outlines how to search Microsoft Purview Unified Catalog to quickly find the data you're looking for. You can search across all data assets, or focus on governed assets that have been curated with governance metadata such as glossary terms, data products, critical data elements, or data quality scores.

Permissions to search the Unified Catalog

Searching Unified Catalog only returns relevant data assets that you have permissions to view.

You can find a data asset in Unified Catalog when:

- You have at least read permissions on an available Azure or Microsoft Fabric resource.
- You have at least data reader permissions on a domain or collection in the Microsoft Purview Data Map where the asset is stored.

Manage permissions to these assets at the resource level and at the Microsoft Purview Data Map level, respectively. For more information on providing this access, see the links above.

To search for curated, ready-to-use data assets, search for data products your organization has developed.

> **Tip**
>
> If your Unified Catalog is well-curated, day-to-day business users shouldn't need to search the full catalog. They should be able to find data they need in data products, or by using the governed assets search to find curated data. For more information about setting up Unified Catalog, see Get started with data governance, and plan for Unified Catalog.

Permissions to search governed assets

Governed assets are visible to users with the Global Catalog Reader role. Local Catalog Readers can view governed assets in the governance domain they have access to. Published governance concepts, such as glossary terms, data products, and critical data elements, are also visible to catalog readers. Draft concepts are visible only to Data Product Owners and Data Stewards.

Existing Data Map reader and curator roles aren't changed by the governed assets search experience.

Searching the catalog

To search for data assets in Unified Catalog:

1. In the Microsoft Purview portal, open **Unified Catalog**.
2. Under **Discovery**, select **Data assets**.
3. In the search bar, enter keywords to narrow your search, such as name, data type, classifications, and glossary terms. Search suggestions appear as you enter keywords. Your search history is also presented so you can resume a previous data exploration.
4. To complete your search, select **View search results** or select **Enter**. For full search syntax information, see keyword search syntax.

The search results page shows a list of data assets and glossary terms that match the search terms. The keyword is highlighted in the return results, so you can see where the term was found in the asset. The **Governed assets** tab displays only governed assets search results, and the **All assets** tab displays the search results in all assets in your data estate, whether they're governed or not.

> **Note**
>
> Search only returns items in collections where you're a data reader or curator. For more information, see roles and permissions in collections.

The Microsoft Purview relevance engine sorts through all the matches and ranks them based on what it believes is useful to a user. For example, a data consumer is likely more interested in a table curated by a data steward that matches on multiple keywords than an unannotated folder. Many factors determine an asset's relevance score. The Microsoft Purview search team constantly tunes the relevance engine to ensure the top search results have value to you.

Keywords and keyword search syntax

Search queries consist of keywords and operators such as `AND`, `OR`, `NOT`, exact phrase quotes, field-scoped search, and parentheses for grouping. You can combine these operators to build precise queries that return exactly the assets you need.

For the full list of supported operators, field names, and detailed syntax guidance, see Keyword search syntax.

Filter results

If the top results don't include the assets you're looking for, you can filter your results by using the filter pane to the left of the search results.

On the search results page, select **Filter** to open and close the filter pane. With the filter pane open, select a filter category, then select the checkbox next to the values you want to filter on. For some filters, you can select the ellipsis to choose **Or** or **And**.

Available filters

You can add filters by selecting **Add filter** on the filter pane and selecting one of the following options:

- **Activity**: Refine your search to attributes created or updated within a certain timeframe.
- **Asset type**: Refine your search to specified asset types. For example: dashboards, files, glossary terms, or metamodel assets.
- **Assigned term**: Refine your search to assets with the selected terms applied.
- **Classification**: Refine your search to assets with certain classifications.
- **Collection**: Refine your search by assets in a specific collection.
- **Contact**: Refine your search to assets that have selected users listed as a contact.
- **Data source type**: Refine your search to assets from specified source types. For example: Azure Blob Storage or Power BI.
- **Endorsement**: Refine your search to assets with specified endorsements, like **Certified** or **Promoted**.
- **Label**: Refine your search to assets with specific security labels.
- **Data asset attributes**: Refine your search to assets with specified data asset attributes. Attributes are listed under their attribute group, and use operators to help search for specific values. For example: Equals or Doesn't equal.
- **Rating**: Refine your search to only data assets with a specified rating.
- **Tags**: Refine your search to assets with selected tags.

Search for governed assets (preview)

The governed assets search provides a focused view of data assets that are actively governed in Unified Catalog. Instead of searching across all assets, you can search only for assets that have governance metadata, such as glossary terms, data products, critical data elements, or data quality scores. Learn more about governed assets.

The governed assets search is helpful because organizations have large data estates but only a small percentage of assets are curated and ready for consumption. Data consumers can use this search to find meaningful, governed data without navigating the entire data estate.

To search for governed assets:

1. In the Microsoft Purview portal, open **Unified Catalog**.
2. Under **Discovery**, select **Data assets**.
3. Enter keywords in the search bar and select **Enter** or **View search results**.
4. On the search results page, select the **Governed assets** tab.

The search results show only assets that are governed in Unified Catalog. Each asset has a **Governed** label next to its name. Governed assets search supports semantic search, so you can use natural language queries to find assets.

Search capabilities

Governed assets search includes the following capabilities:

- **Partial keyword match**: Enter a partial keyword, and the search returns results that match the partial term. For example, entering *cust* returns results that match *customer*.
- **Exact match with double quotes**: Use double quotes to search for an exact phrase. For example, "*customer profile*" searches for the phrase as a single unit, rather than matching *customer* and *profile* separately.
- **Search highlights**: Hover over a search result to see why it was returned. The highlight shows which attributes matched your search query, such as name, description, or term.

> **Note**
>
> Search isn't case sensitive, but spaces within search terms matter. Underscores and other non-space characters are treated as part of the keyword.

Filter governed asset results

You can filter governed asset search results by using the following filters (see filter definitions):

- Asset type
- Data source type
- Classification
- Contact
- Glossary term
- Data product
- Critical data element

> **Note**
>
> The **Terms** filter in governed assets search uses Unified Catalog terms. After you migrate glossary terms and enable data asset curation, the **Assigned term** filter is no longer available in the all data assets search, and classic terms no longer appear on asset details pages.

View governed asset details

Select a governed asset from the search results to view its details. On the asset details page, select the **Governance** tab to see how the asset is governed. The **Governance** tab shows:

- Data products the asset is part of.
- Glossary terms linked to the asset.
- Critical data elements associated with columns in the asset.
- The latest data quality score, if a data quality run was performed on the asset.

The **Governance** tab follows existing permission models. If you have a Catalog Reader role, you see only published governance concepts. Draft concepts, such as draft glossary terms, are visible only to Data Product Owners and Data Stewards.

If an asset is deleted from Data Map, a warning icon appears next to the asset in search results. The asset continues to appear until the Data Product Owner removes it.

> **Note**

If an asset is ungoverned (all links to data products, glossary terms, and critical data elements are removed), the asset currently still appears in governed assets search results.

Bulk edit search results

If you want to make changes to multiple assets returned by search, you can modify glossary terms, classifications, and contacts in bulk. For more information, see How to bulk edit assets.

View assets

From the search results page, select an asset to view details, such as description, attributes, schema, lineage, and classifications.

Browse the Unified Catalog

Unified Catalog offers a browse experience that enables users to explore what data is available to them either by collection or by exploring the hierarchy of each data source in the catalog.

In Unified Catalog, go to **Discovery** > **Data assets**. Under **Explore your data**, select a card for how you want to explore your data: in Azure, in Microsoft Fabric, by source type, or by collection.

You can only see the data sources you have permission to view.

Searching Microsoft Purview in connected services

After you register your Microsoft Purview instance to an Azure Data Factory or an Azure Synapse Analytics workspace, you can search Unified Catalog directly from those services. For more information, see Discover data in ADF using Microsoft Purview and Discover data in Synapse using Microsoft Purview.

Link glossary terms to assets and columns (preview)

You can add business context to your data by linking glossary terms to data assets and columns.

Users need two roles in order to attach terms to assets and columns:

1. **Data Curator**: All users need to hold this role in the collection where the asset belongs. Only Data Curators can edit data assets. Learn more about this role and collections permissions in Data Map.

2. In addition, users need either one of these roles:
    a. **Data Steward**: Users can be a Data Steward in any governance domian. These users can see published and unpublished glossary terms, and can attach both kinds of terms to assets and columns.
    b. **Global Asset Curator**: This role allows users to attach published glossary terms to assets and columns. Users with this role can only see published terms and attach published terms to assets and columns.

See Data governance roles and permissions for role descriptions and how to assign roles.

Link terms to a data asset:

1. Search for an asset or column in Unified Catalog.
2. Select the data asset to open its details page.
3. Select **Edit**.
4. On the **Overview** page, at **Glossary terms**, select **Add term**.
5. Search for a term by entering keywords, or select the checkbox next to a term listed on the page. You can select multiple terms.
6. Select **Add**.
7. Select **Save**.

The terms are attached to the asset.

Link terms to a column within an asset:

1. Search for an asset or column in Unified Catalog.
2. Select the data asset to open its details page.
3. Select **Edit**.
4. On the **Schema** page, find the row for the column you want to add a term to, and select **Add term**.
5. Search for a term by entering keywords, or select the checkbox next to a term listed on the page. You can select multiple terms.
6. Select **Add**.
7. Select **Save**.

The terms are attached to the column.

ⓘ **Note:** The author created this article with assistance from AI. Learn more

Last updated on 05/21/2026

Workflows in Microsoft Purview

Article • 02/24/2025

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

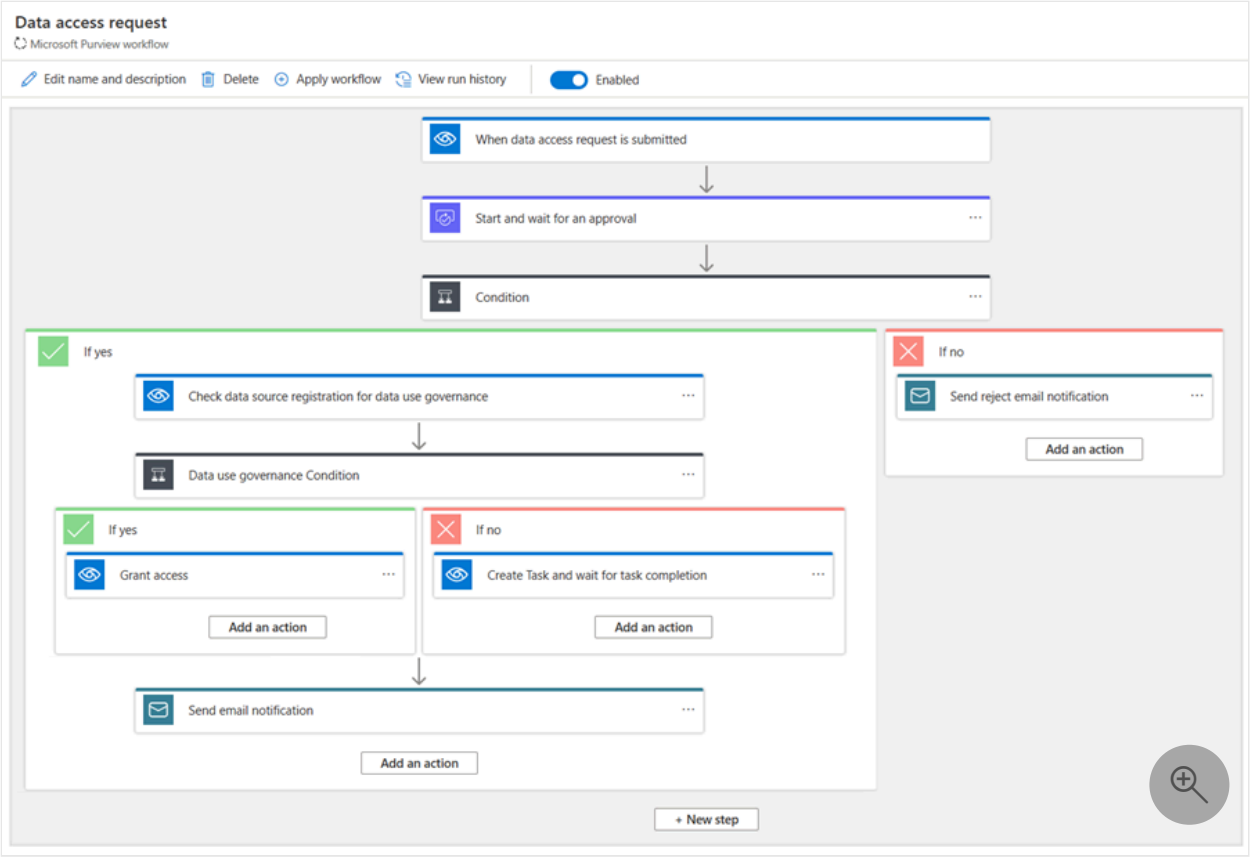
> **Important**
>
> This feature is currently in preview. The Supplemental Terms of Use for Microsoft Azure Previews include additional legal terms that apply to Azure features that are in beta, in preview, or otherwise not yet released into general availability.

Workflows are automated, repeatable business processes that users can create within Microsoft Purview to validate and orchestrate CUD (create, update, delete) operations on their data entities. Enabling these processes allows organizations to track changes, enforce policy compliance, and ensure quality data across their data landscape.

Workflows automate the approval process for updates your data assets, so you can ensure quality updates.

What are workflows?

Workflows are automated processes that are made up of connectors that contain a common set of pre-established actions and are run when specified operations occur in your data catalog.



For example: A user attempts to delete a business glossary term that is bound to a workflow. When the user submits this operation, the workflow runs through its actions instead of, or before, the original delete operation.

Workflow actions include things like generating approval requests or sending a notification, that allow users to automate validation and notification systems across their organization.

Currently, there are two kinds of workflows:

- **Data governance** - for data policy, access governance, and loss prevention. Scoped at the collection level.
- **Data catalog** - to manage approvals for CUD (create, update, delete) operations for glossary terms. Scoped at the glossary level.

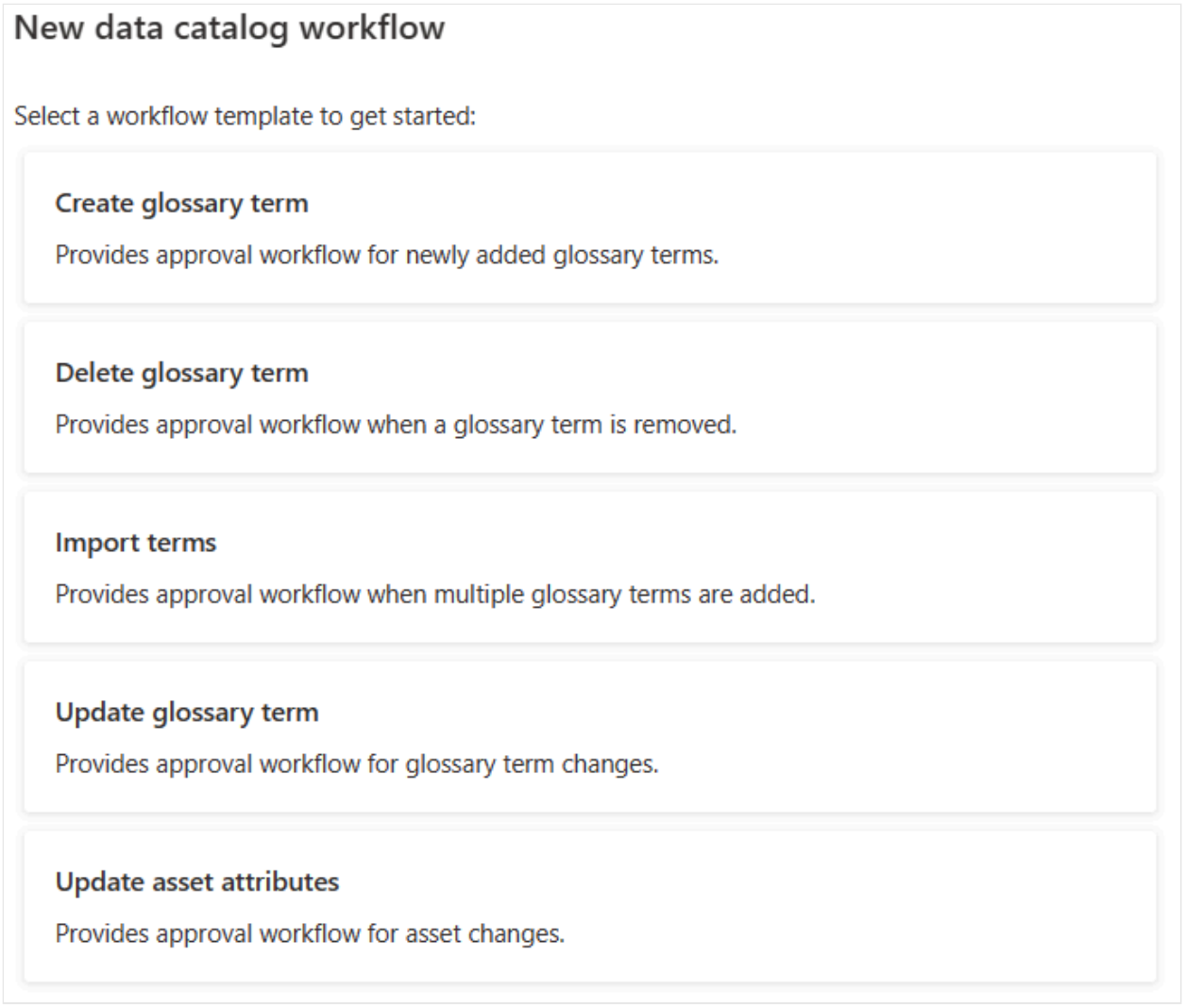
These workflows can be built from pre-established workflow templates provided in the Microsoft Purview governance portal, but are fully customizable using the available workflow connectors.

Workflow templates

For all the different types of user defined workflows enabled and available for your use, Microsoft Purview provides templates to help workflow administrators create workflows without needing to build them from scratch. The templates are built into the authoring

experience and automatically populate based on the workflow being created, so there's no need to search for them.

Templates are available to launch the workflow authoring experience. However, a workflow admin can customize the template to meet the requirements in their organization.

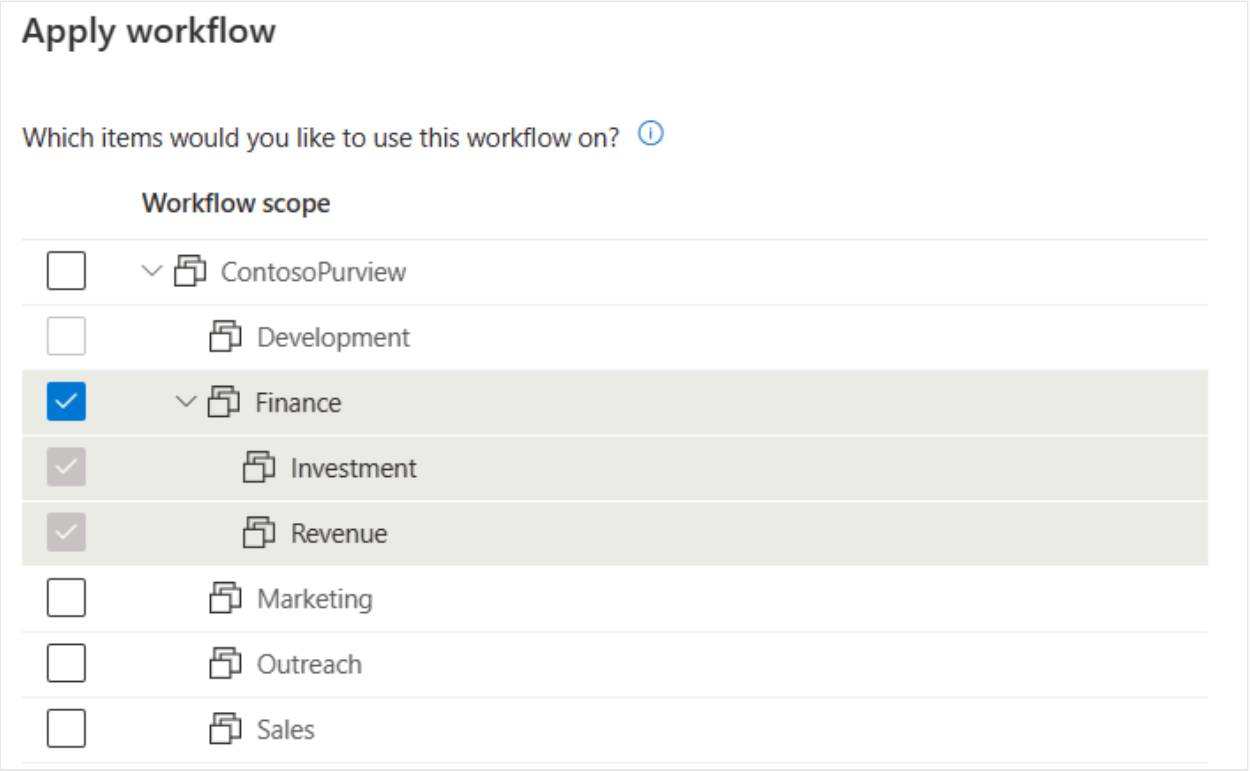


Workflow connectors

Workflow connectors are a common set of actions applicable across some workflows. They can be used in any workflow in Microsoft Purview to create processes customized to your organization. To view the list of existing workflow connectors in Microsoft Purview, see the workflow connectors documentation..

Workflow scope

Once a workflow is created and enabled, it can be bound to a particular scope. This gives you the flexibility to run different workflows for different areas/departments in your organization.



Data governance workflows are scoped to collections, and can be bound to the root collection to govern the whole Microsoft Purview catalog, or any subcollection.

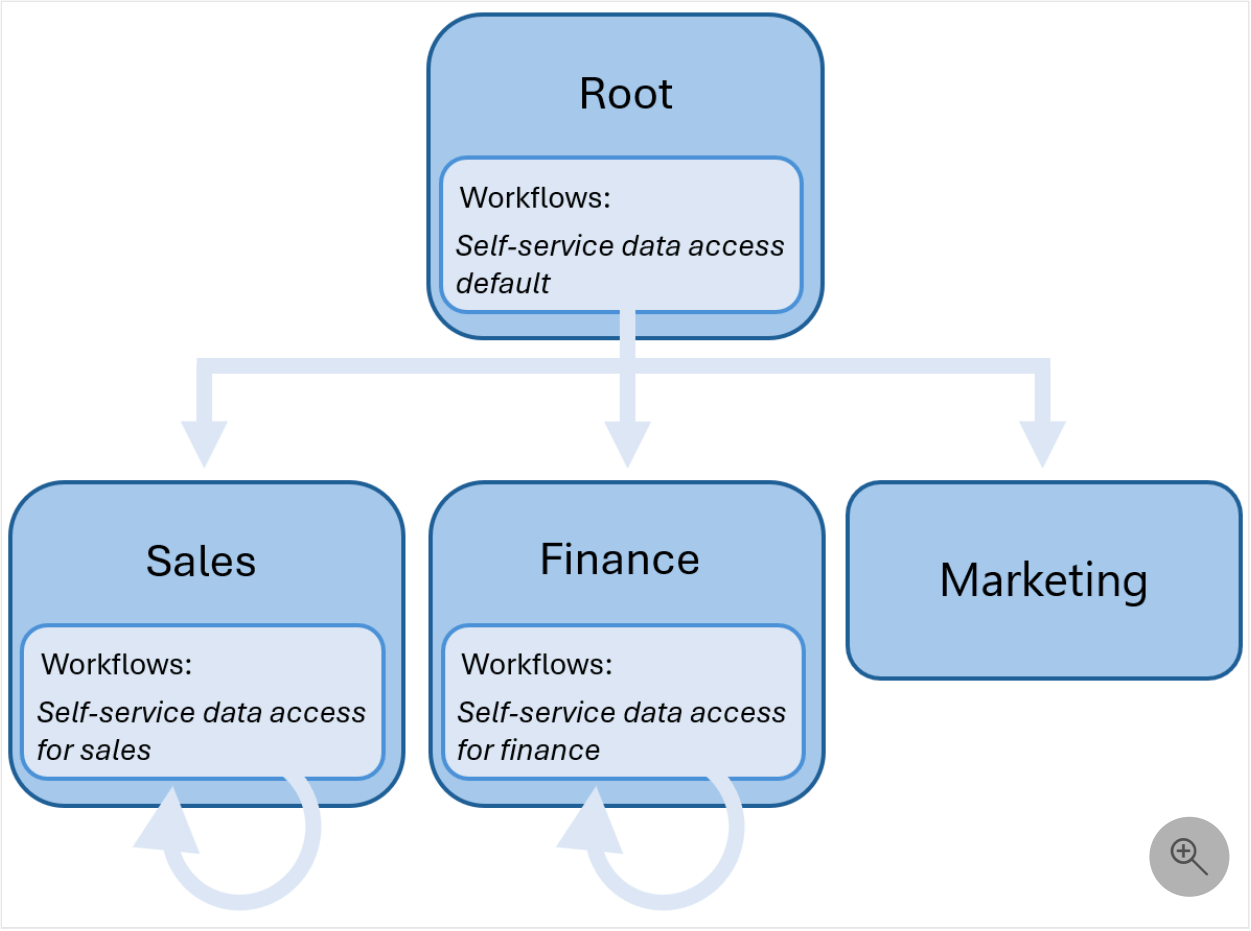
Data catalog workflows are scoped to the glossary and can be bound to the entire glossary, any single term, or any parent term to manage child-terms.

If there's no workflow directly associated with a scope, the workflow engine will traverse upward in the scope hierarchy to determine closest workflow, and run that workflow for the operation.

For example, the AdatumCorp Purview account has the following collection hierarchy:

Root Collection > Sales | Finance | Marketing

- **Root collection** has the workflow *Self-Service data access default* defined and bound.
- **Sales** has *Self-Service data access for sales* defined and bound.
- **Finance** has *Self-Service data access for finance* defined and bound.
- **Marketing** has no workflows directly bound.



When an access request is made for a data asset in Finance collection, the *Self-Service data access for finance collection* workflow is run.

However, when a request access is made for a data asset in Marketing collection, *Self-Service data access default workflow* is triggered. Because there are no workflows bound at the Marketing scope, the workflow engine traverses to the next level in the scope hierarchy, which is Marketing's parent: the root collection. The workflow at the parent scope, the root collection scope, is run.

Who can manage workflows?

A new role, **Workflow Admin** is being introduced with workflow functionality.

A Workflow admin defined for a collection can create self-service workflows and bind these workflows to the collections they have access to.

A Workflow admin defined for any collection can create approval workflows for the business glossary. In order to bind the glossary workflows to a term you need to have at least Data reader permissions.

Next steps

Now that you understand what workflows are, you can follow these guides to use them in your Microsoft Purview account:

- Self-service data access workflow for hybrid data estates
- Approval workflow for business terms
- Manage workflow requests and approvals

Feedback

Was this page helpful? Yes No

Self-service access workflows for hybrid data estates

Article • 02/26/2025

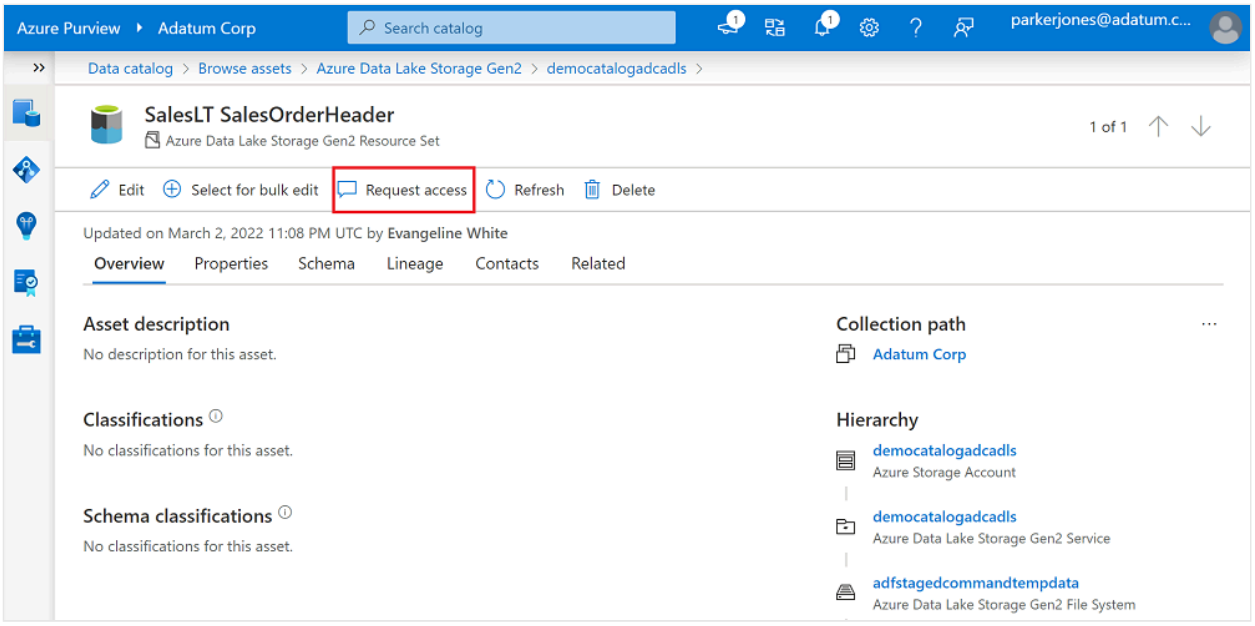
> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

> **Important**
>
> This feature is currently in preview. The Supplemental Terms of Use for Microsoft Azure Previews include additional legal terms that apply to Azure features that are in beta, in preview, or otherwise not yet released into general availability.

You can use workflows to automate some business processes through Microsoft Purview. Self-service access workflows allow you to create a process for your users to request access to datasets they've discovered in Microsoft Purview.

Let's say your team has a new data analyst who will do some business reporting. You add that data analyst to your department's collection in Microsoft Purview. From there, they can browse through the data assets and read descriptions about the data that your department has available.

The data analyst notices that one of the Azure Data Lake Storage Gen2 accounts seems to have the exact data that they need to get started. Because a self-service access workflow has been set up for that resource, they can request access to that Azure Data Lake Storage account from within Microsoft Purview.



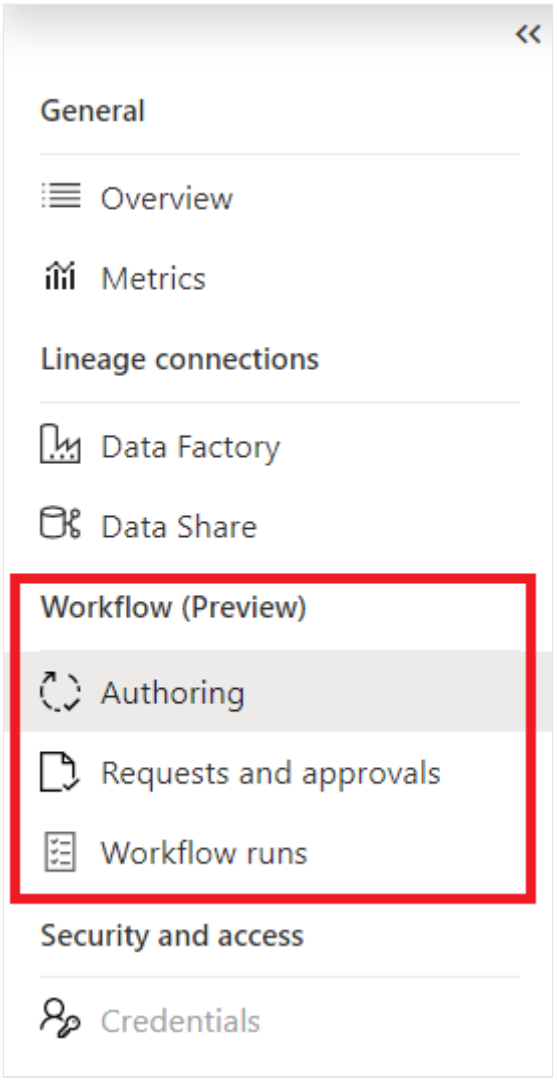
You can create these workflows for any of your resources across your data estate to automate the access request process. Workflows are assigned at the collection level, so they automate business processes along the same organizational lines as your permissions.

This guide shows you how to create and manage self-service access workflows in Microsoft Purview.

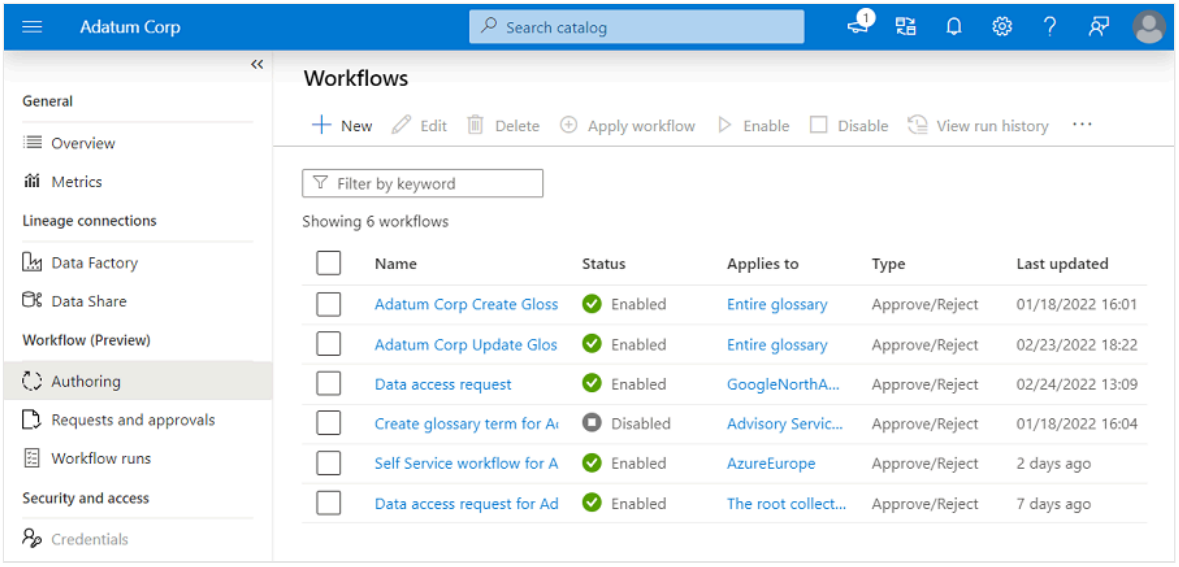
> **Note**
>
> To create or edit a workflow, you need the workflow admin role in Microsoft Purview. You can also contact the workflow admin in your collection, or reach out to your collection administrator, for permissions.

Create and enable the self-service access workflow

1. Sign in to the Microsoft Purview governance portal and select the management center. Three new icons appear in the table of contents.



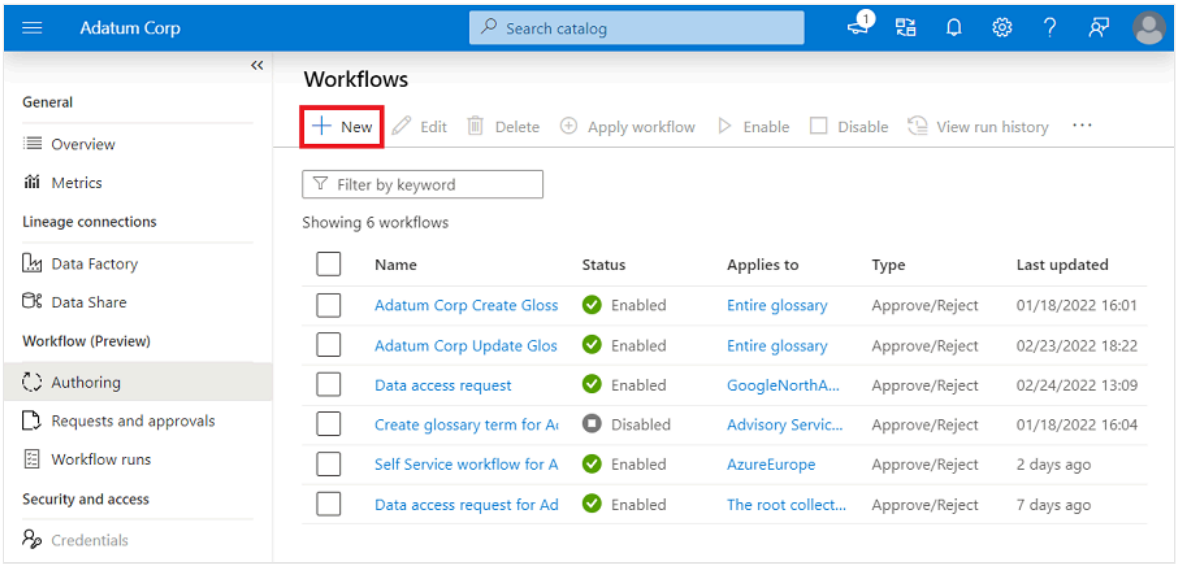
2. To create new workflows, select **Authoring**. This step takes you to the workflow authoring experience.



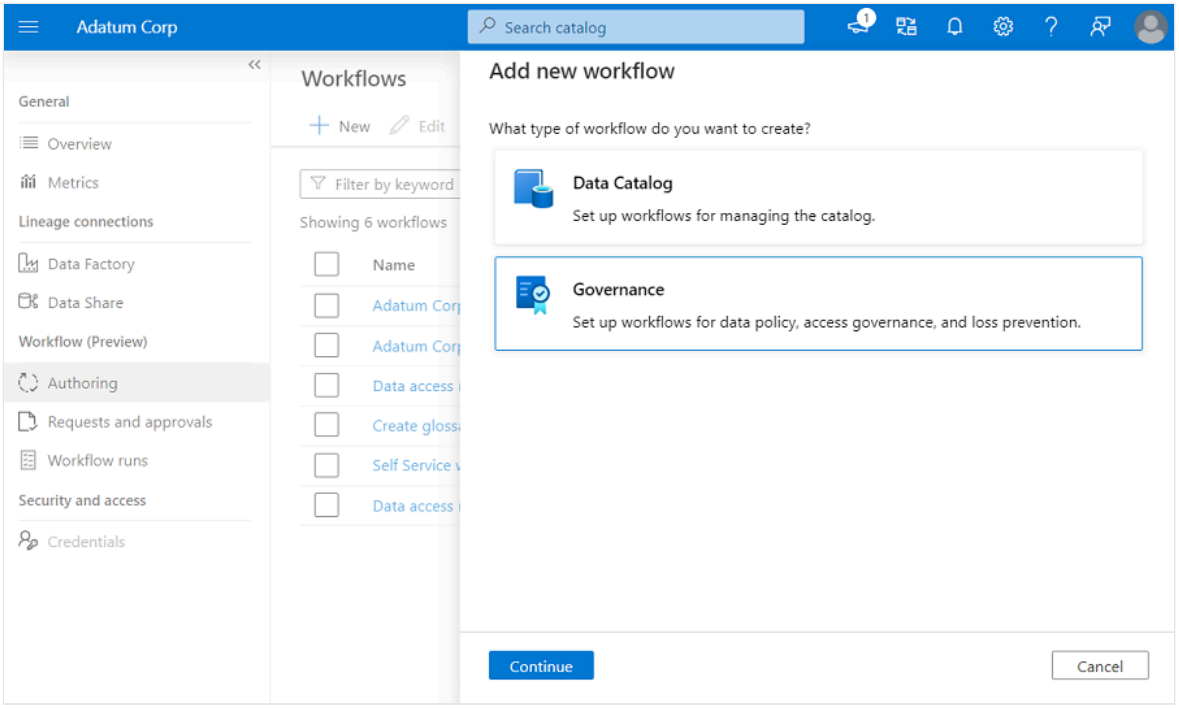
> **Note**

If the **Authoring** tab is unavailable, you don't have the permissions to author workflows. You need the workflow admin role.

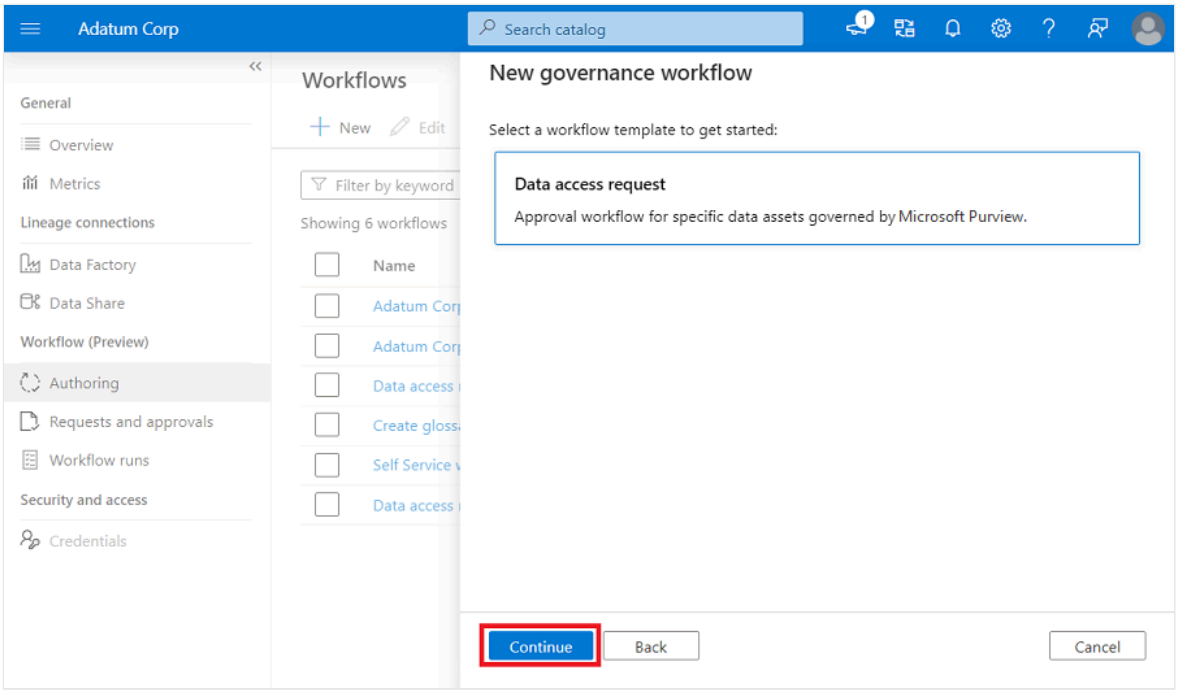
3. To create a new self-service workflow, select the **+New** button.



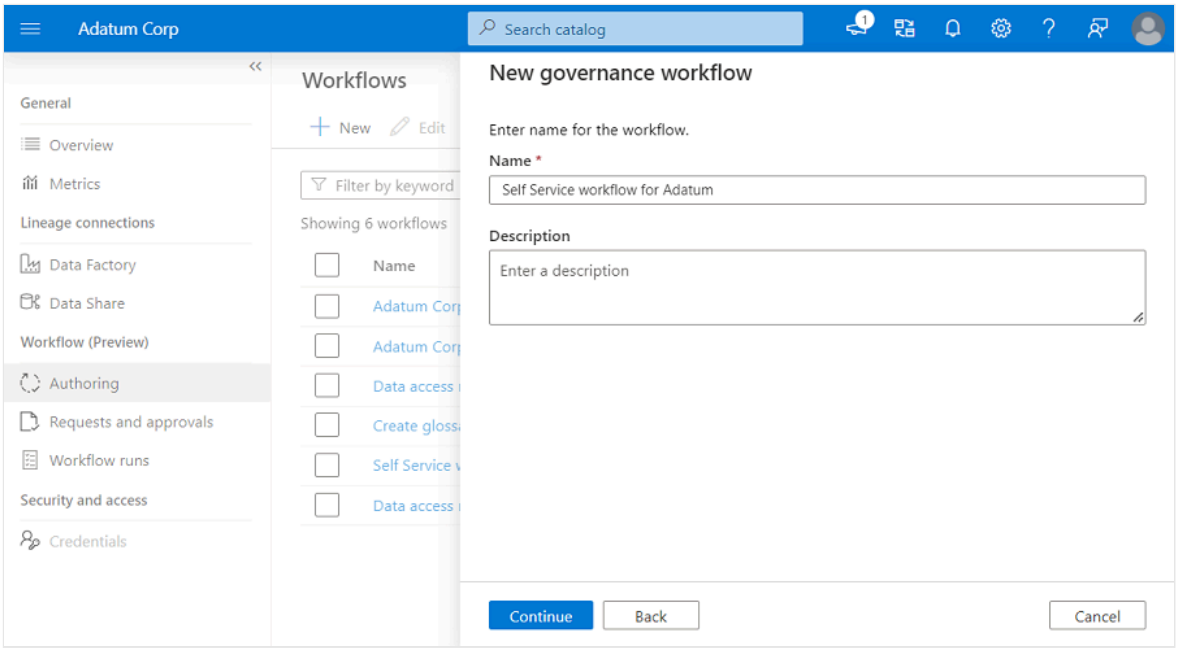
4. You're presented with categories of workflows that you can create in Microsoft Purview. To create an access request workflow, select **Governance**, and then select **Continue**.



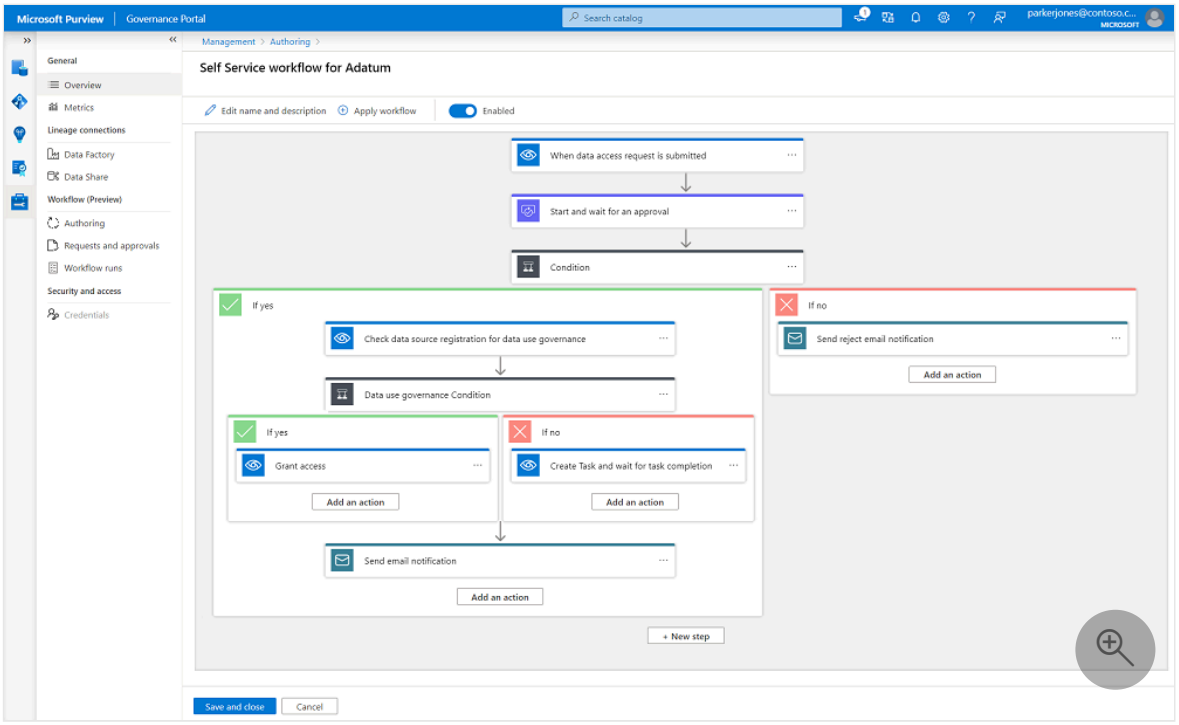
5. The next screen shows all the templates that Microsoft Purview provides to create a self-service data access workflow. Select the **Data access request** template, and then select **Continue**.



6. Enter a workflow name, optionally add a description, and then select **Continue**.



7. You're presented with a canvas where the selected template is loaded by default.



The template has the following steps:

a. Trigger when a data access request is made.

b. Get an approval connector that specifies a user or group that will be contacted to approve the request.

Assign data owners as approvers. Using the dynamic variable **Asset.Owner** as approvers in the approval connector will send approval requests to the data owners on the entity.

> **Note**
>
> Using the **Asset.Owner** variable might result in errors if an entity doesn't have a data owner.

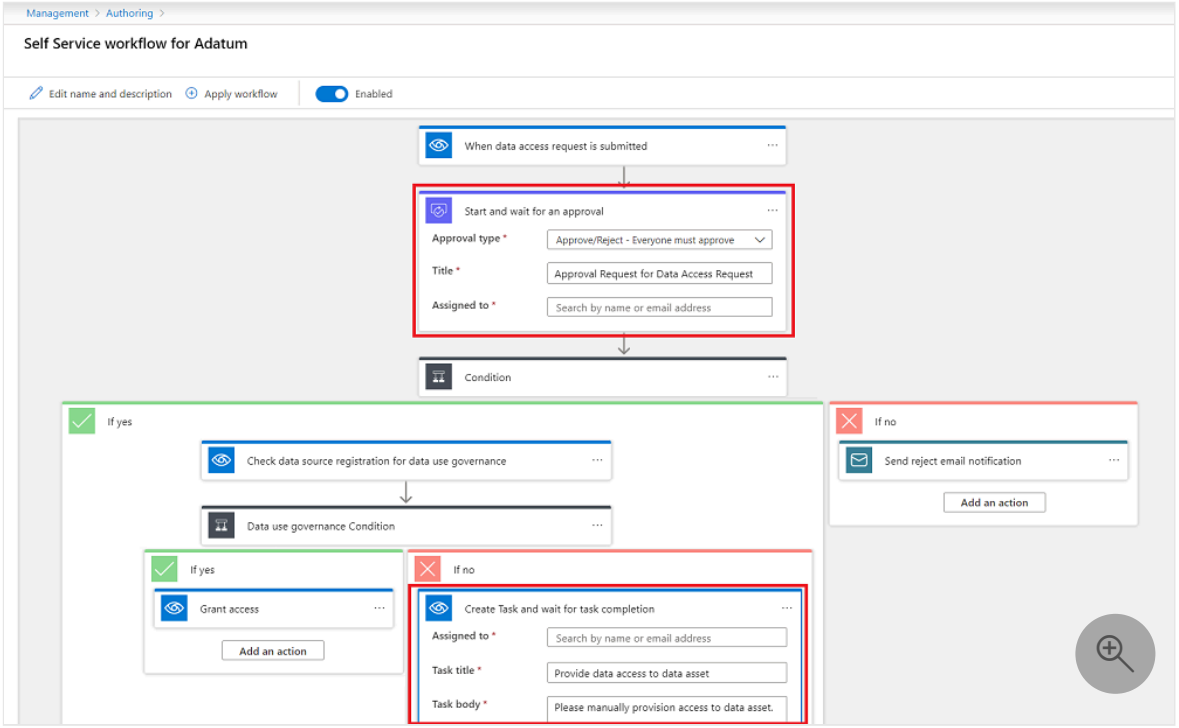
8. If the condition to check approval status is approved, take the following steps:

   - If a data source is registered for data policy enforcement with the policy:
     a. Create a self-service policy.
     b. Send email to the requestor that confirms access.
   - If a data source isn't registered with the policy:
     a. Use a connector to assign a task to a user or a Microsoft Entra group to manually provide access to the requestor.
     b. Send an email to requestor to explain that access is provided after the task is marked as complete.

   If the condition to check approval status is rejected, send an email to the requestor to say that the data access request is denied.

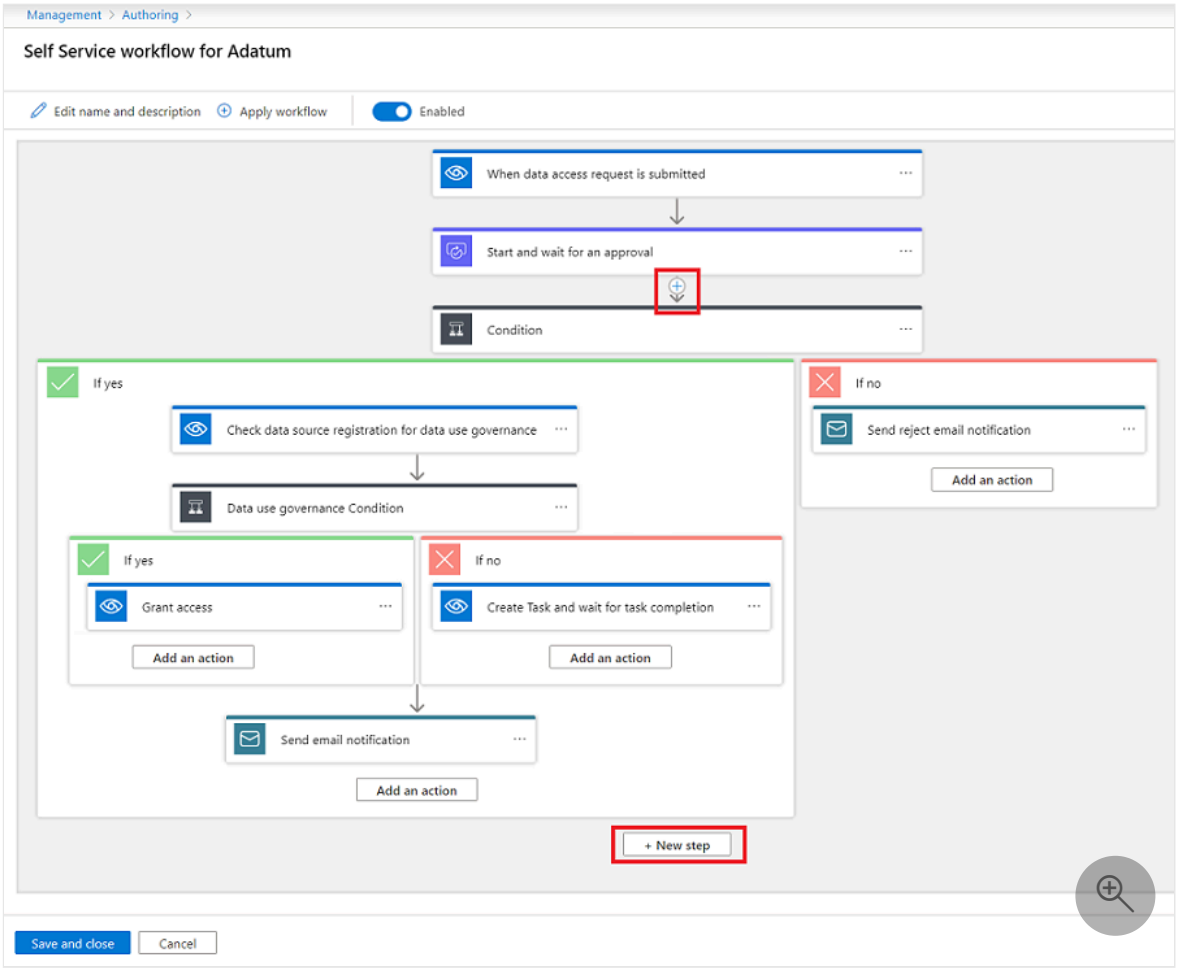
9. You can use the default template as it is by populating two fields:

   - Add an approver's email address or Microsoft Entra group in the **Start and wait for an approval** connector.
   - Add a user's email address or Microsoft Entra group in the **Create task and wait for task completion** connector to denote who is responsible for manually providing access if the source isn't registered with the policy.



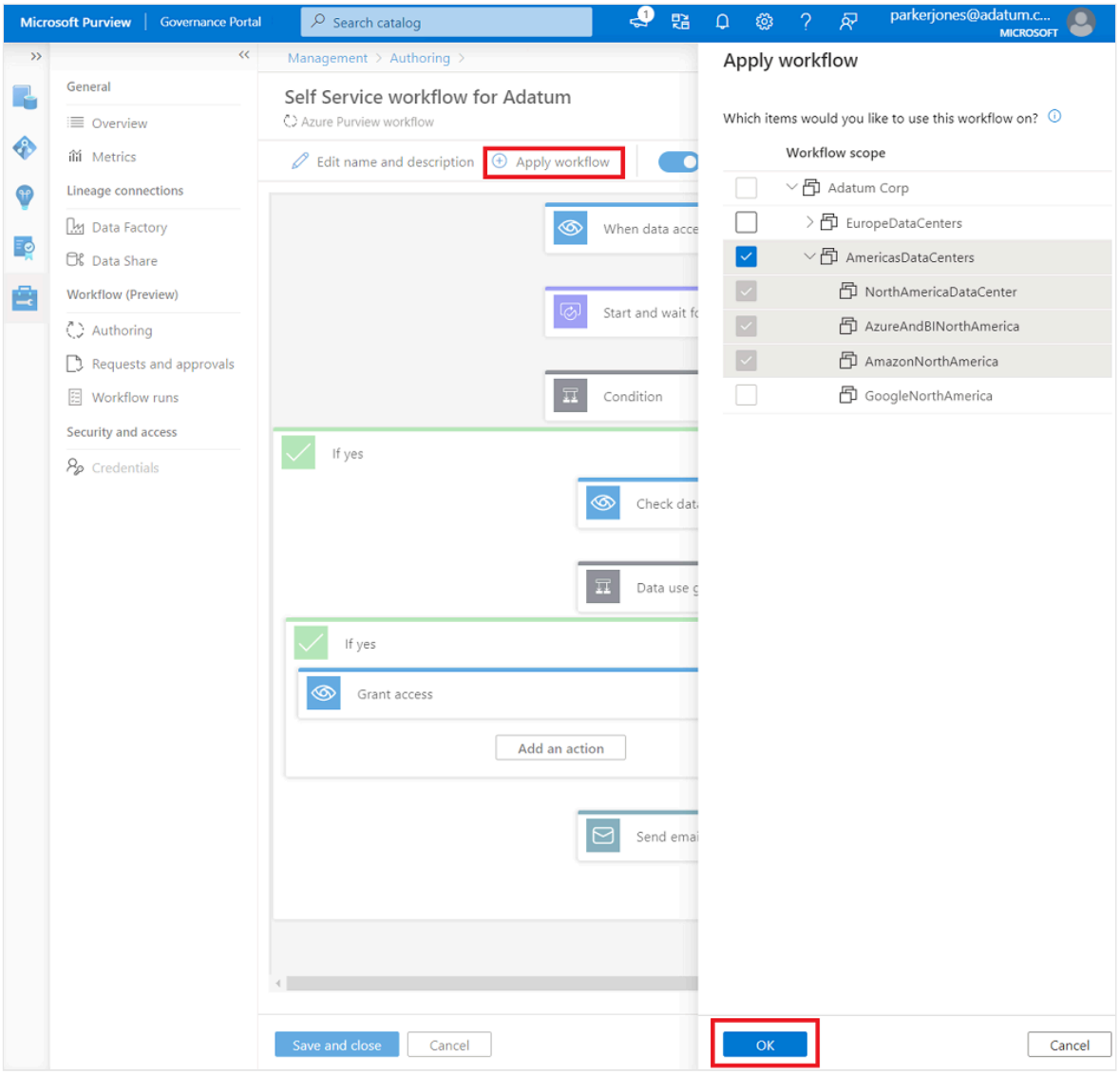
> **Note**
>
> Configure the workflow to create self-service policies only for sources that the Microsoft Purview policy supports. To see what the policy supports, check the documentation about data owner policies.
>
> If the Microsoft Purview policy doesn't support your source, use the **Create task and wait for task completion** connector to assign tasks to users or groups that can provide access.

You can also modify the template by adding more connectors to suit your organizational needs.



10. After you define a workflow, you need to bind the workflow to a collection hierarchy path. The binding (or scoping) implies that this workflow is triggered only for data access requests in that collection.

    To bind a workflow or to apply a scope to a workflow, select **Apply workflow**. Select the scope that you want to associate with this workflow, and then select **OK**.

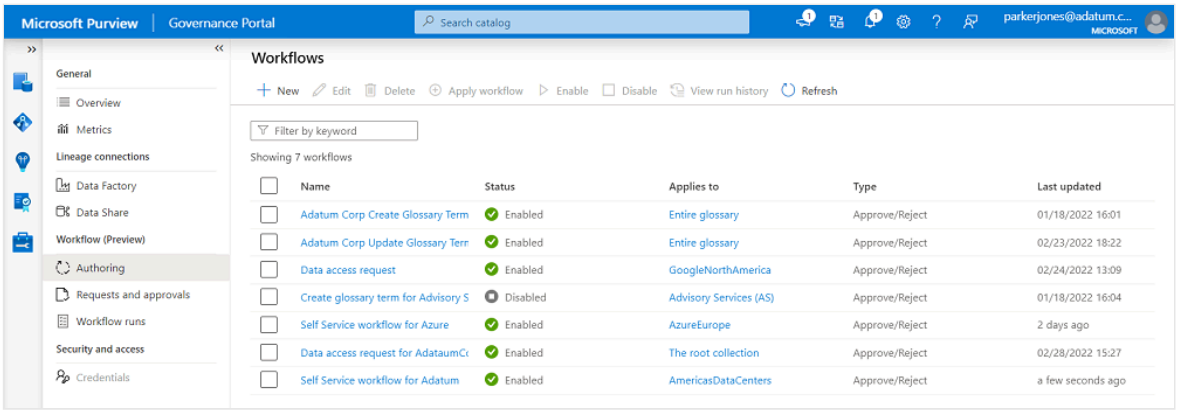


> **Note**
>
> The Microsoft Purview workflow engine will always resolve to the closest workflow that the collection hierarchy path is associated with. If the workflow engine doesn't find a direct binding, it will look for the workflow that's associated with the closest parent in the collection tree.

11. Make sure that the **Enable** toggle is on. The workflow should be enabled by default.

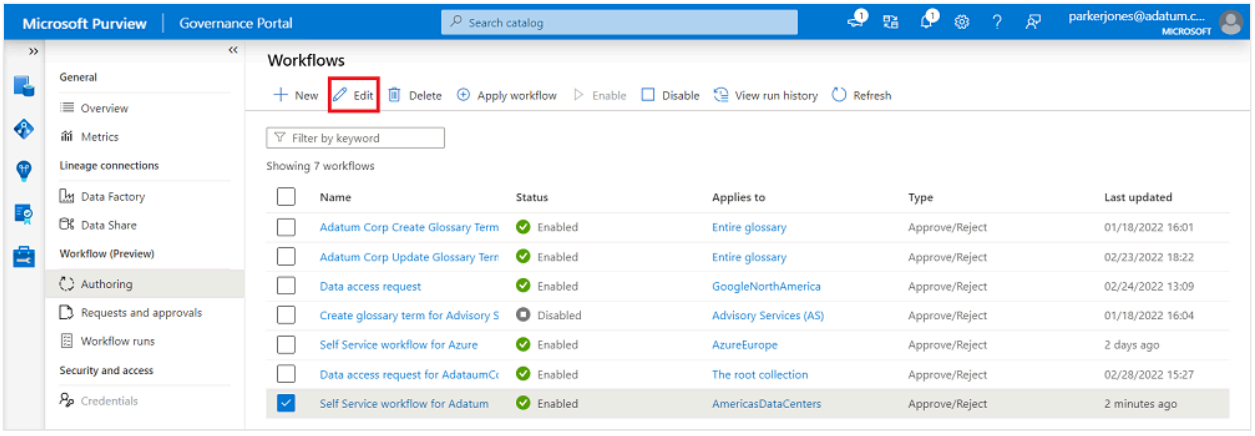
12. Select **Save and close** to create and enable the workflow.

    Your new workflow now appears in the list of workflows.



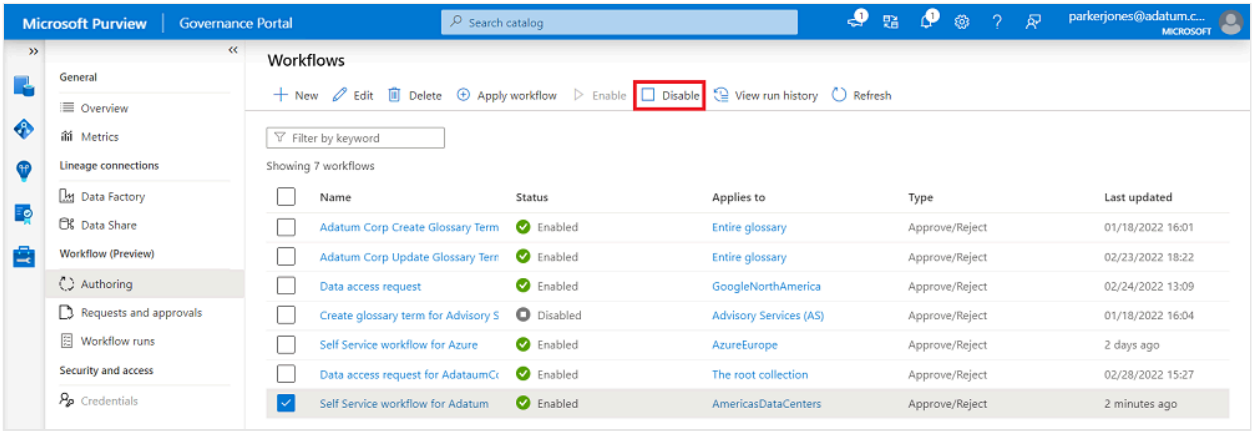
Edit an existing workflow

To modify an existing workflow, select the workflow, and then select the **Edit** button. You're presented with the canvas that contains the workflow definition. Modify the workflow, and then select **Save** to commit the changes.



Disable a workflow

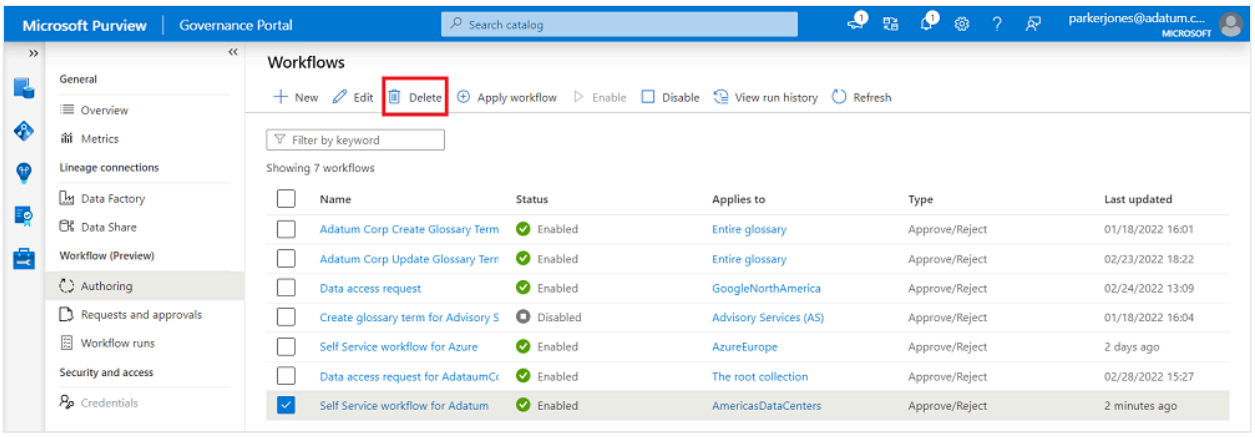
To disable a workflow, select the workflow, and then select **Disable**.



Another way is to select the workflow, select **Edit**, turn off the **Enable** toggle in the workflow canvas, and then select **Save and close**.

Delete a workflow

To delete a workflow, select the workflow, and then select **Delete**.



Next steps

For more information about workflows, see these articles:

- Workflows in Microsoft Purview
- Approval workflow for business terms
- Manage workflow requests and approvals
- Self-service access policies

Feedback

Was this page helpful? Yes No

Approval workflow for business terms

Article • 02/26/2025

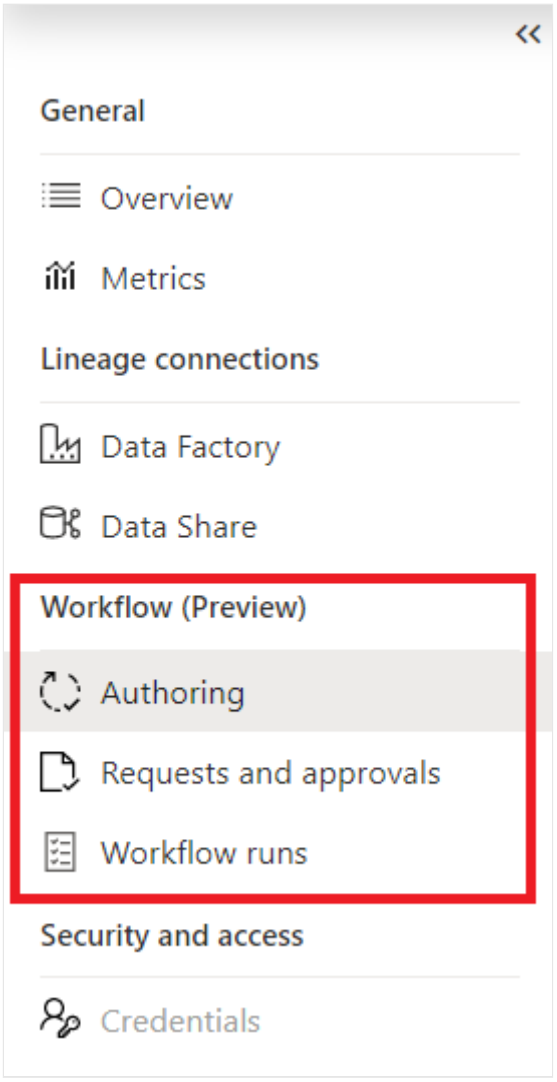
> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

> **Important**
>
> This feature is currently in preview. The Supplemental Terms of Use for Microsoft Azure Previews include additional legal terms that apply to Azure features that are in beta, in preview, or otherwise not yet released into general availability.

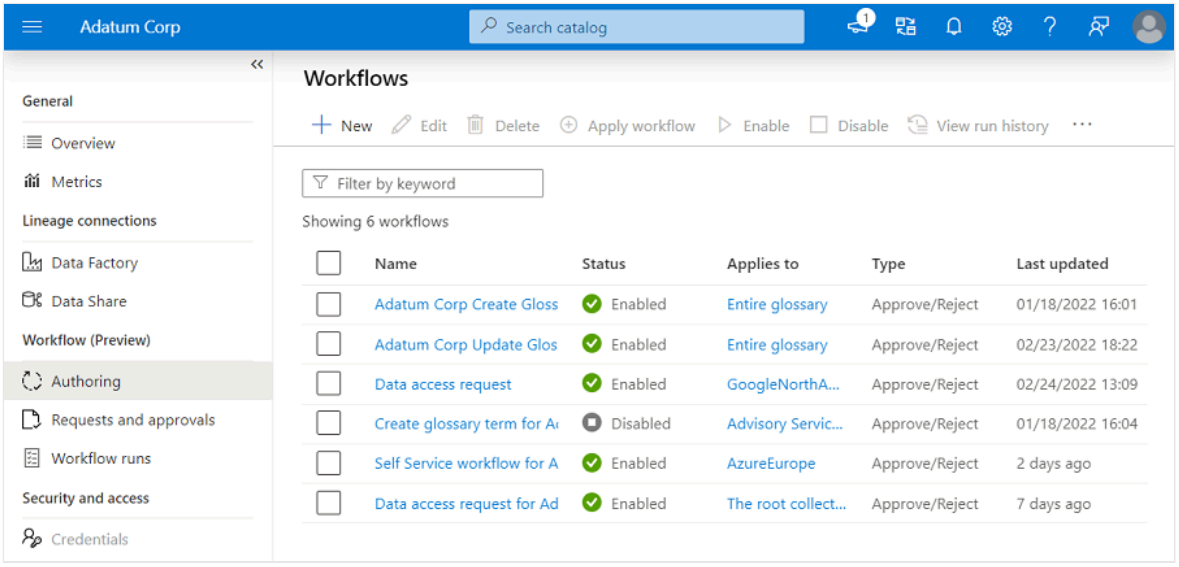
This guide will take you through the creation and management of approval workflows for business terms.

Create and enable a new approval workflow for business terms

1. Sign in to the Microsoft Purview governance portal and select the Management center. You'll see three new icons in the table of contents.



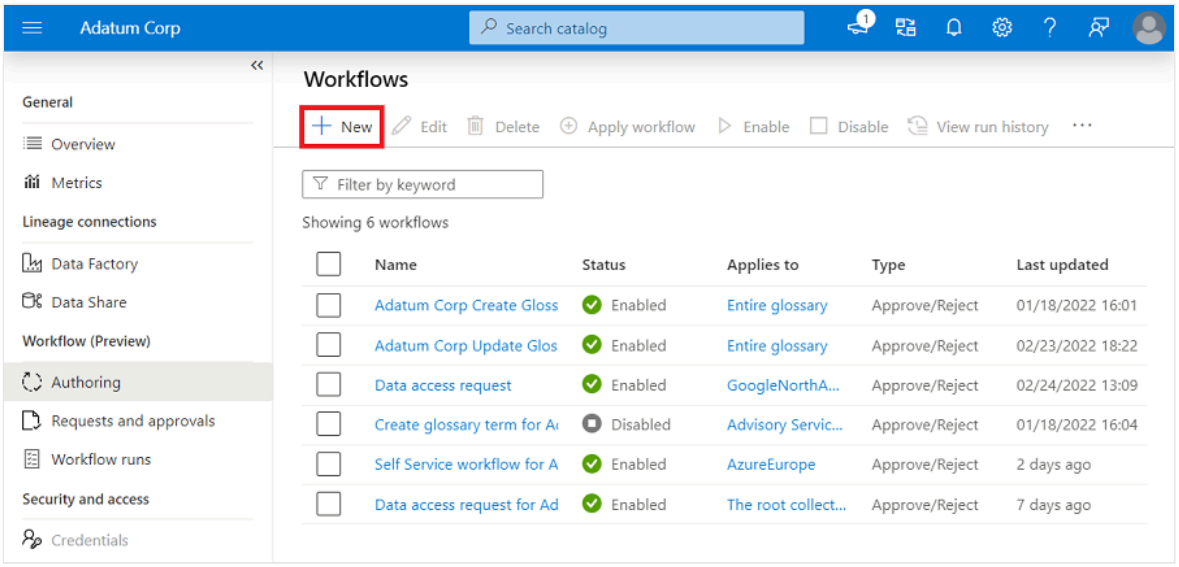
2. To create new workflows, select **Authoring** in the workflow section. This will take you to the workflow authoring experiences.



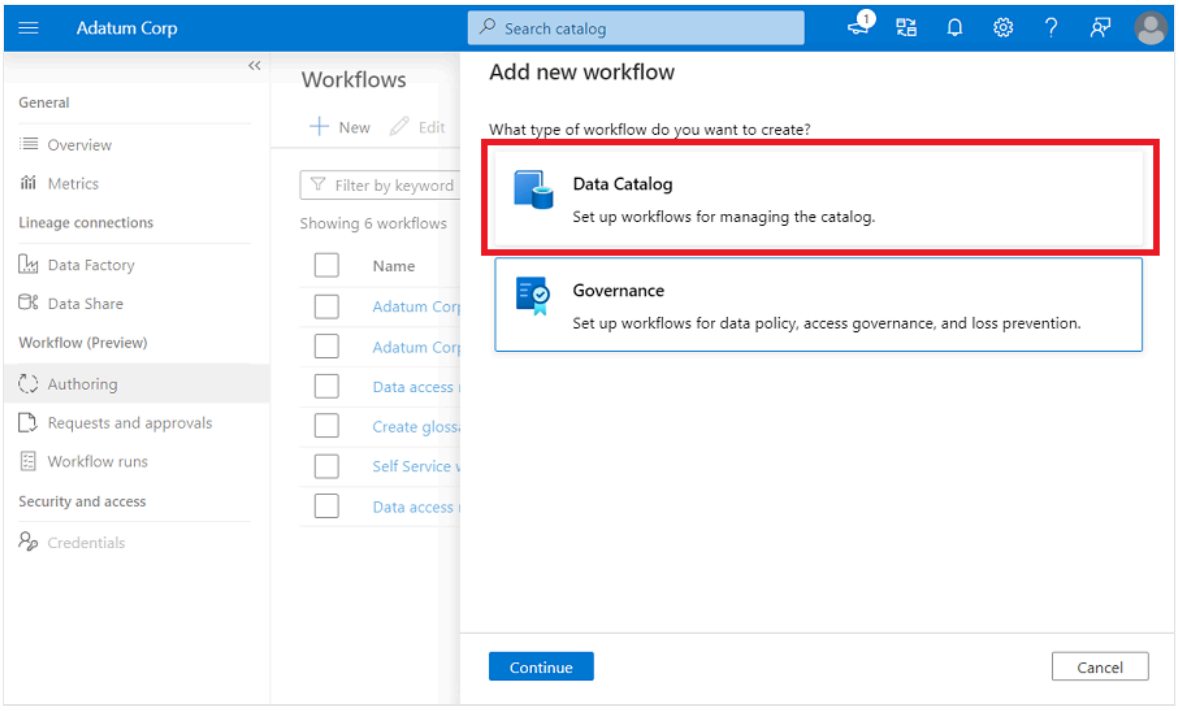
> **Note**

If the authoring tab is greyed out, you don't have the permissions to be able to author workflows. You'll need the workflow admin role.

3. To create a new workflow, select **+New** button.

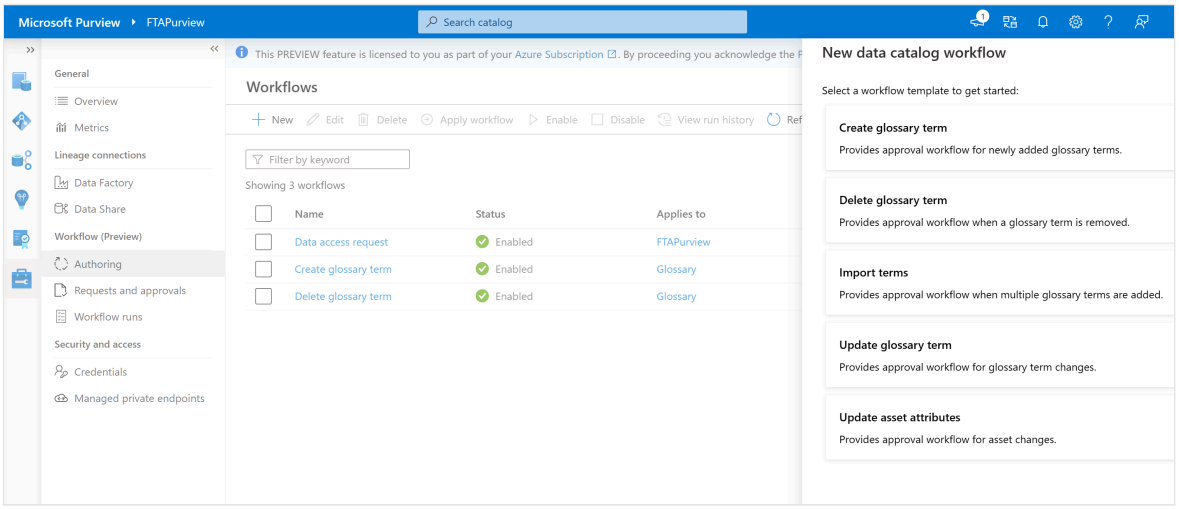


4. To create **Approval workflows for business terms** Select **Data Catalog** and select **Continue**

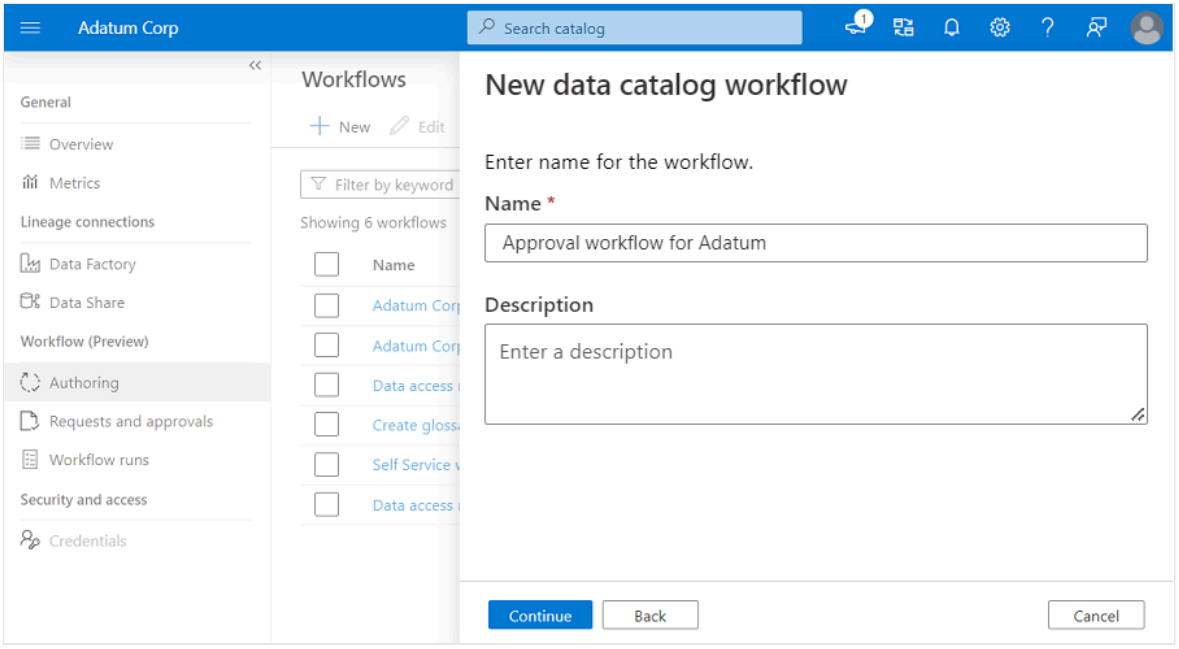


5. In the next screen, you'll see all the templates provided by Microsoft Purview to create a workflow. Select the template using which you want to start your authoring experiences and select **Continue**. Each of these templates specifies the kind of action that will trigger the workflow. In the screenshot below we've selected **Create glossary term**. The four different templates available for business glossary are:

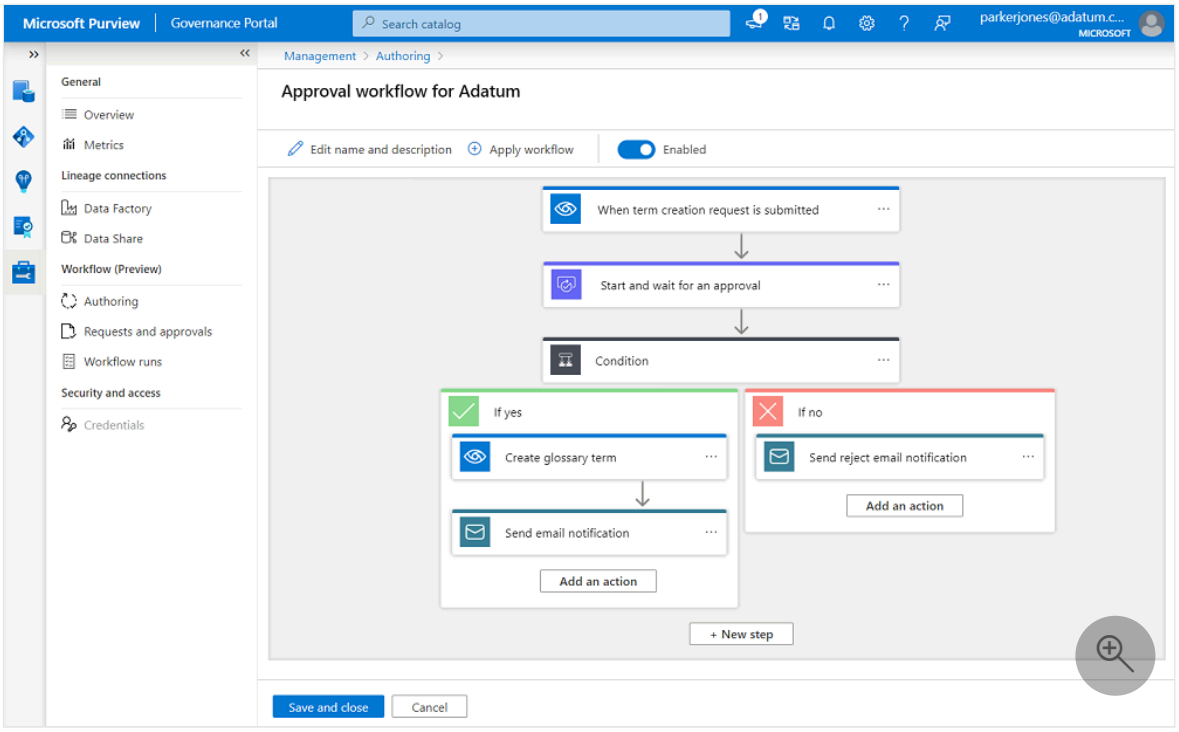
- Create glossary term - when a term is created, approval will be requested.
- Update glossary term - when a term is updated, approval will be requested.
- Delete glossary term - when a term is deleted, approval will be requested.
- Import terms - when terms are imported, approval will be requested.



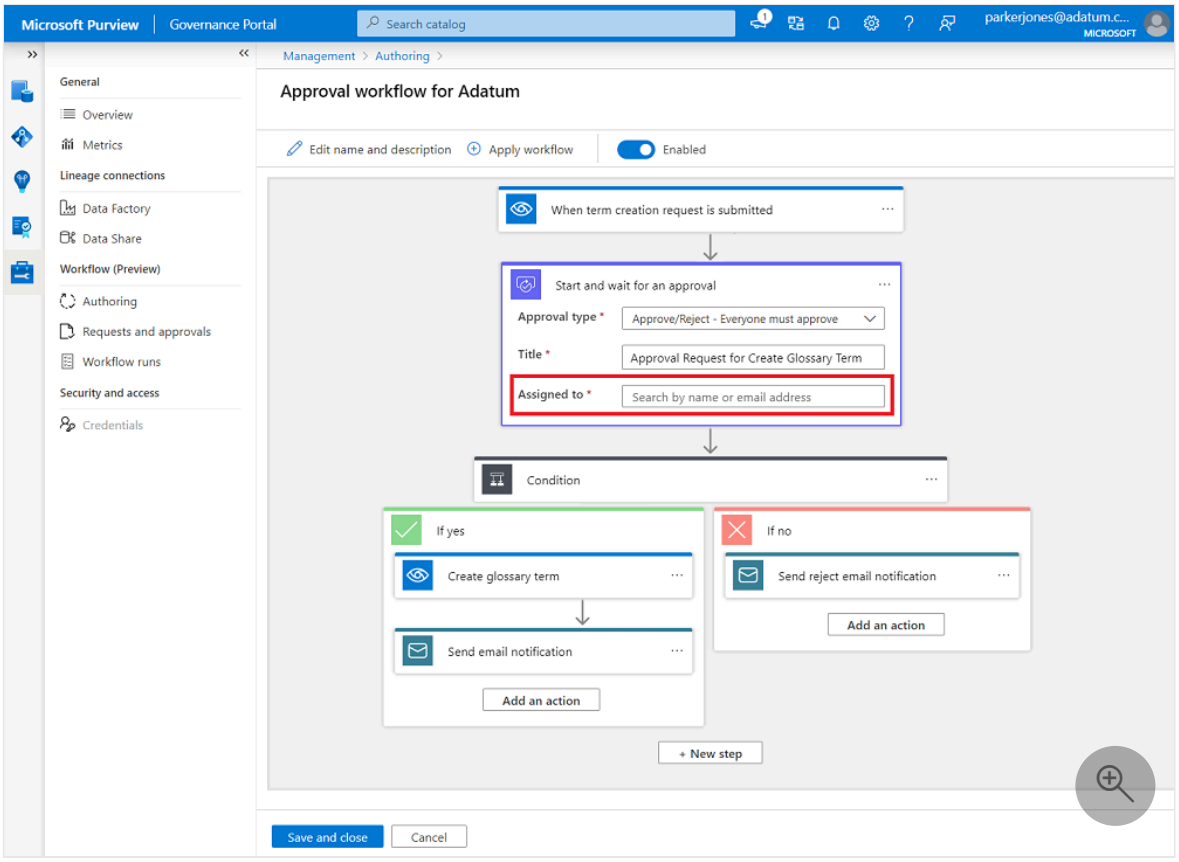
6. Next, enter a workflow name and optionally add a description. Then select **Continue**.



7. You'll now be presented with a canvas where the selected template is loaded by default.



8. The default template can be used as it is by populating the approver's email address in **Start and Wait for approval** Connector.



The default template has the following steps:

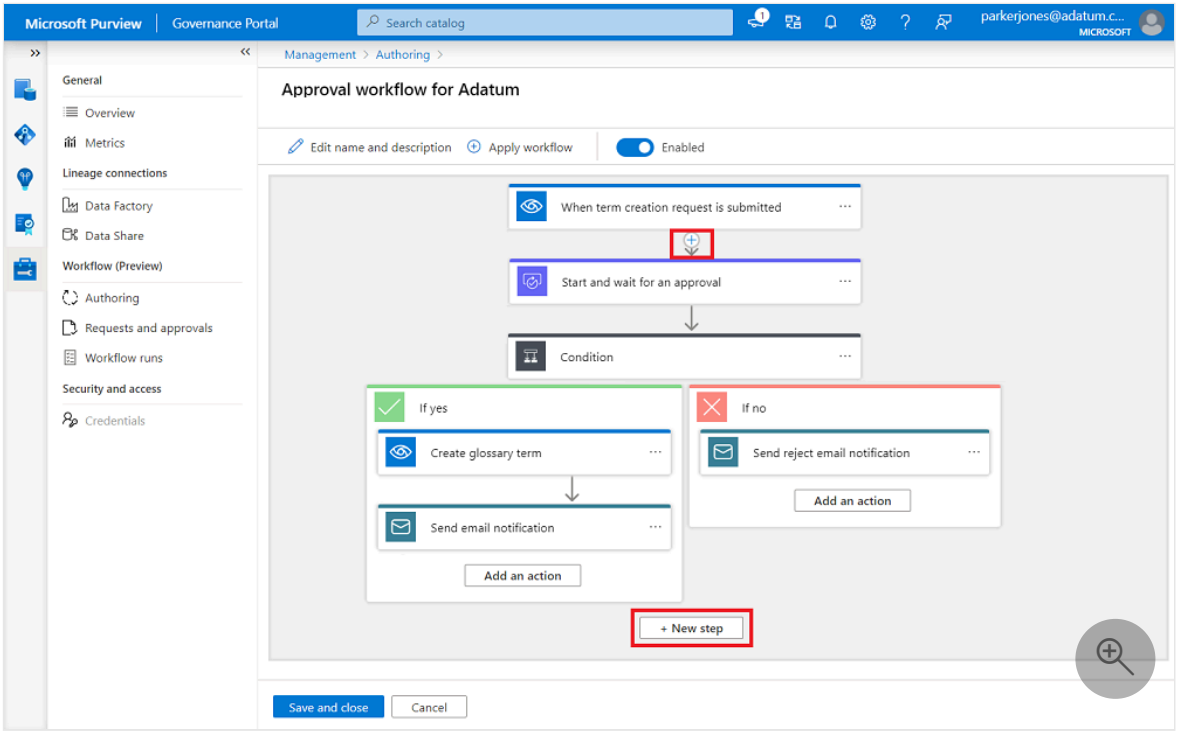
a. Trigger when a glossary term is created/updated/deleted/imported depending on the template selected.

b. Approval connector that specifies a user or group that will be contacted to approve the request.

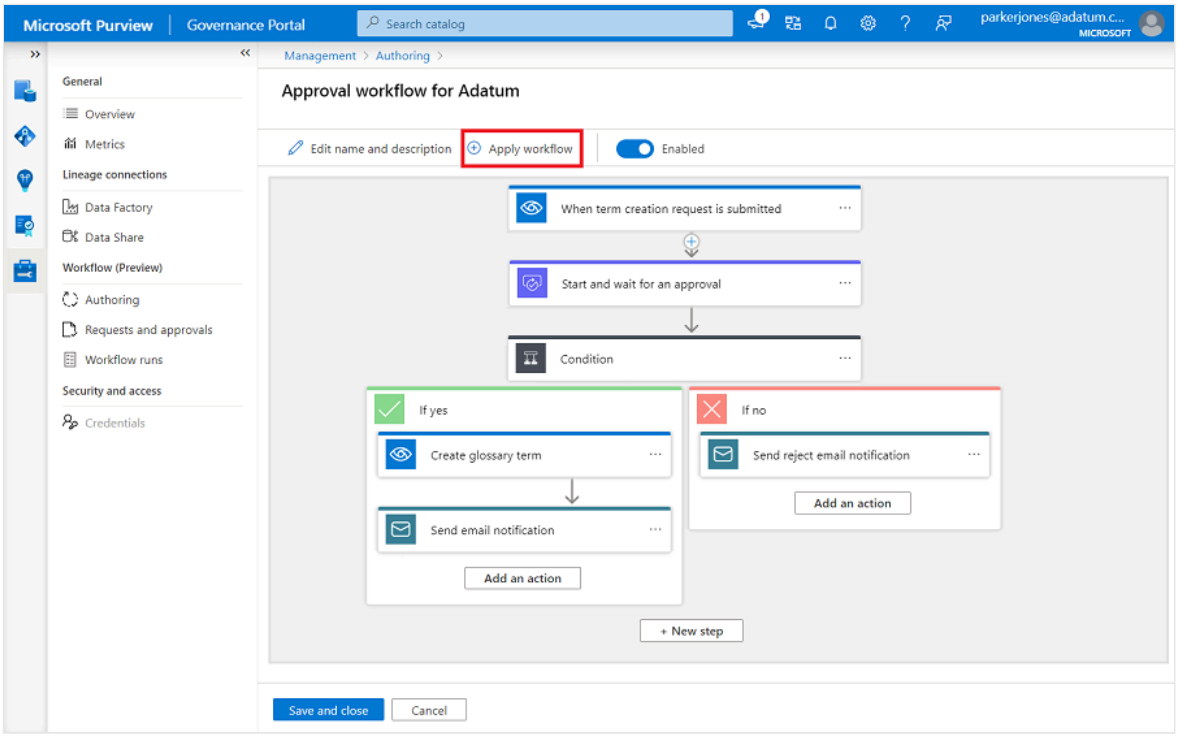
c. Condition to check approval status

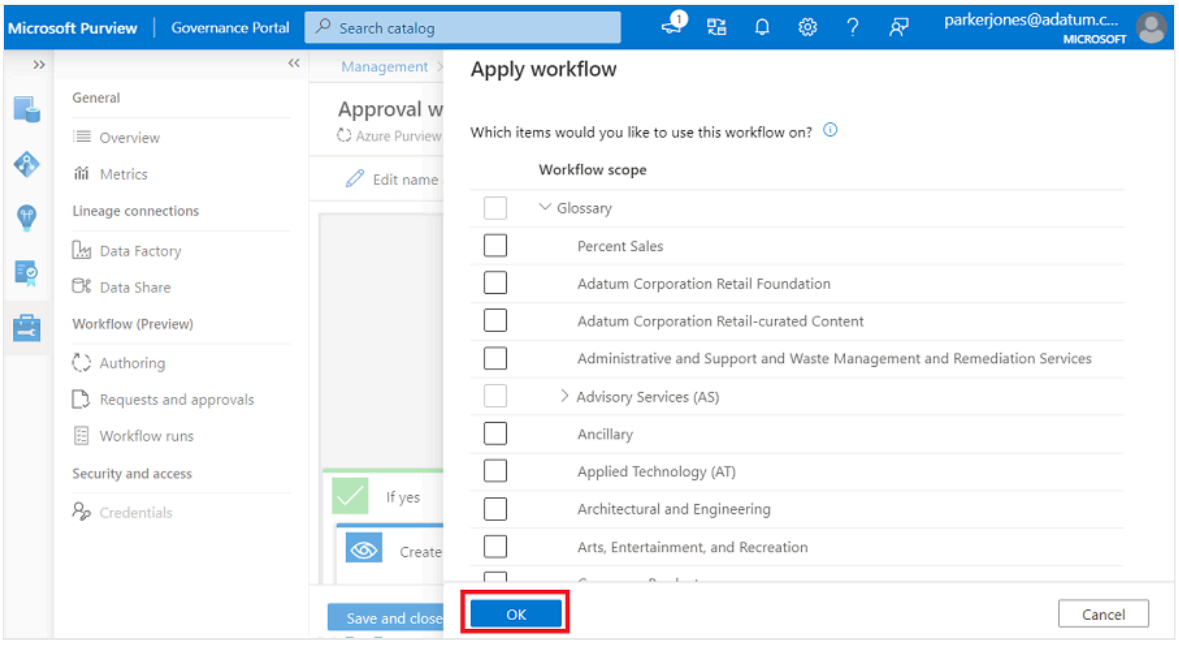
- If approved:
  i. Create/update/delete/import the glossary term
  ii. Send an email to requestor that their request is approved, and term CUD (create, update, delete) operation is successful.
- If rejected:
  i. Send email to requestor that their request is denied.

9. You can also modify the template by adding more connectors to suit your organizational needs. Add a new step to the end of the template by selecting the **New step** button. Add steps between any already existing steps by selecting the arrow icon between any steps.



10. Once you're done defining a workflow, you need to bind the workflow to a glossary hierarchy path. The binding implies that this workflow is triggered only for CUD operations within the specified glossary hierarchy path. A workflow can be bound to only one hierarchy path. To bind a workflow or to apply a scope to a workflow, you need to select 'Apply workflow'. Select the scopes you want this workflow to be associated with and select **OK**.





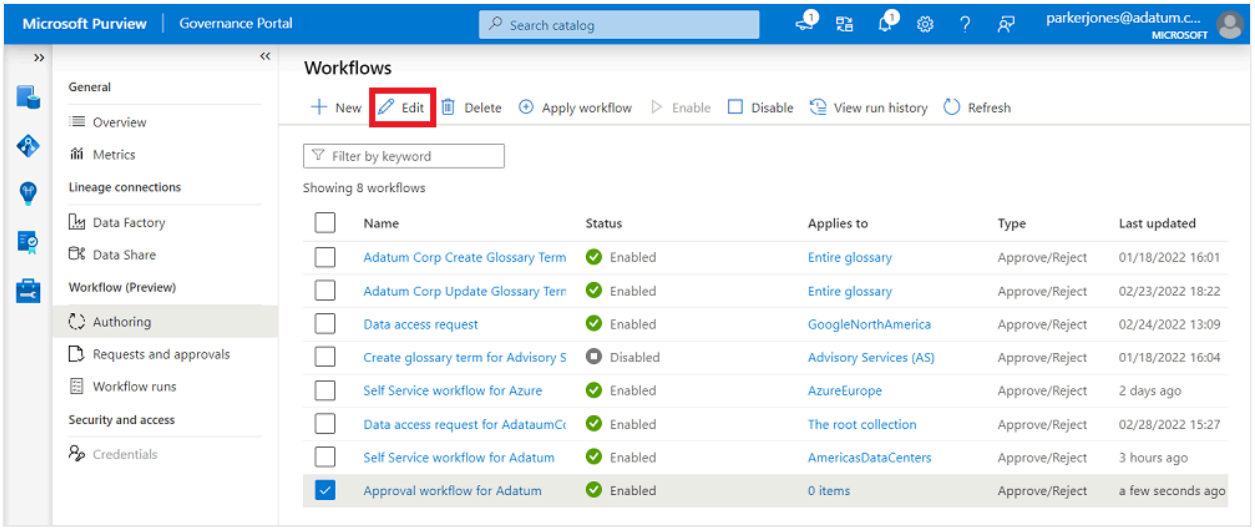
> **Note**
>
> - The Microsoft Purview workflow engine will always resolve to the closest workflow that the term hierarchy path is associated with. In case a direct binding is not found, it will traverse up in the tree to find the workflow associated with the closest parent in the glossary tree.
> - Import terms can only be bound to root glossary path as the .CSV can contain terms from different hierarchy paths.

11. By default, the workflow will be enabled. To disable, toggle the Enable button in the top menu.

12. Finally select **Save and close** to create the workflow.

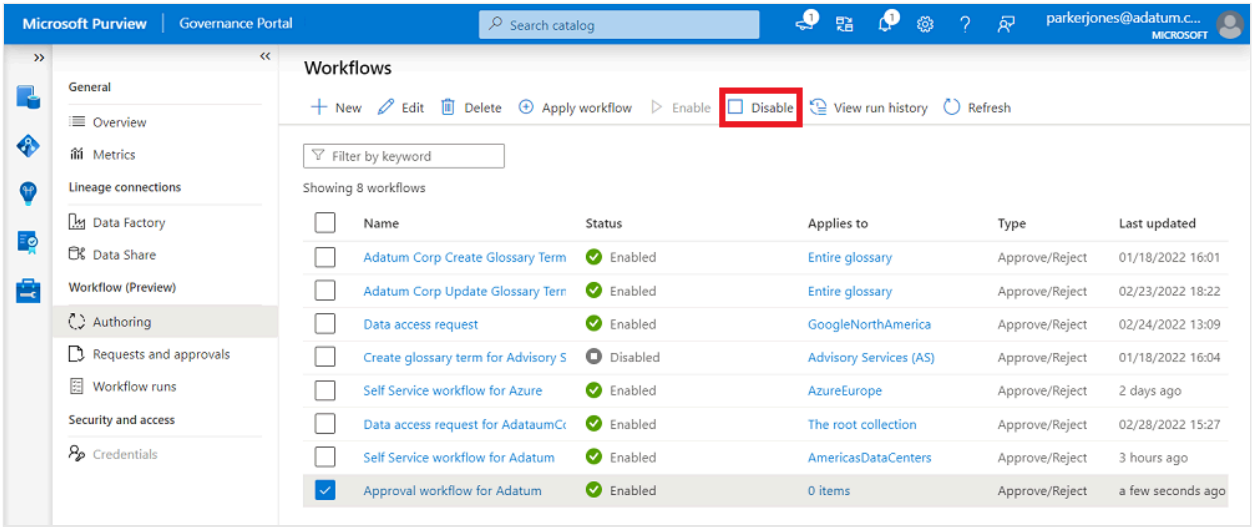
Edit an existing workflow

To modify an existing workflow, select the workflow and then select **Edit** in the top menu. You'll then be presented with the canvas containing workflow definition. Modify the workflow and select **Save** to commit changes.



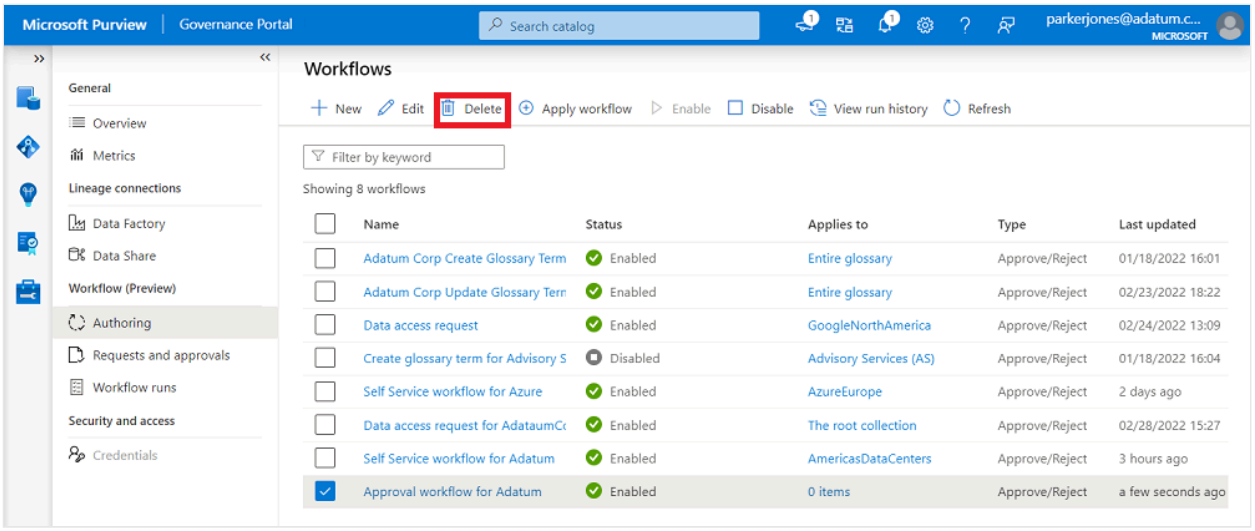
Disable a workflow

To disable a workflow, select the workflow and then select **Disable** in the top menu. You can also disable the workflow by selecting **Edit** and changing the enable toggle in workflow canvas.



Delete a workflow

To delete a workflow, select the workflow and then select **Delete** in the top menu.



Limitations for business terms with approval workflow enabled

- Non-approved glossary terms aren't saved in Purview catalog.
- The behavior of tagging terms to assets/schemas is same as today. That is, previously created draft terms can be tagged to assets/schemas.

Next steps

For more information about workflows, see these articles:

- What are Microsoft Purview workflows
- Self-service data access workflow for hybrid data estates
- Manage workflow requests and approvals

Feedback

Was this page helpful? Yes No

Approval workflow for asset curation

Article • 02/26/2025

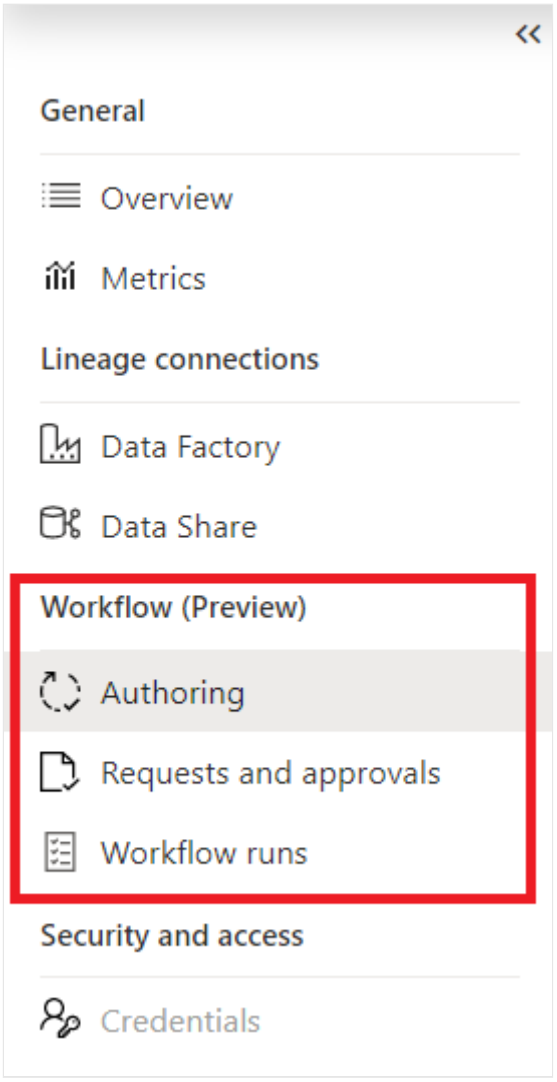
> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

> **Important**
>
> This feature is currently in preview. The Supplemental Terms of Use for Microsoft Azure Previews include additional legal terms that apply to Azure features that are in beta, in preview, or otherwise not yet released into general availability.

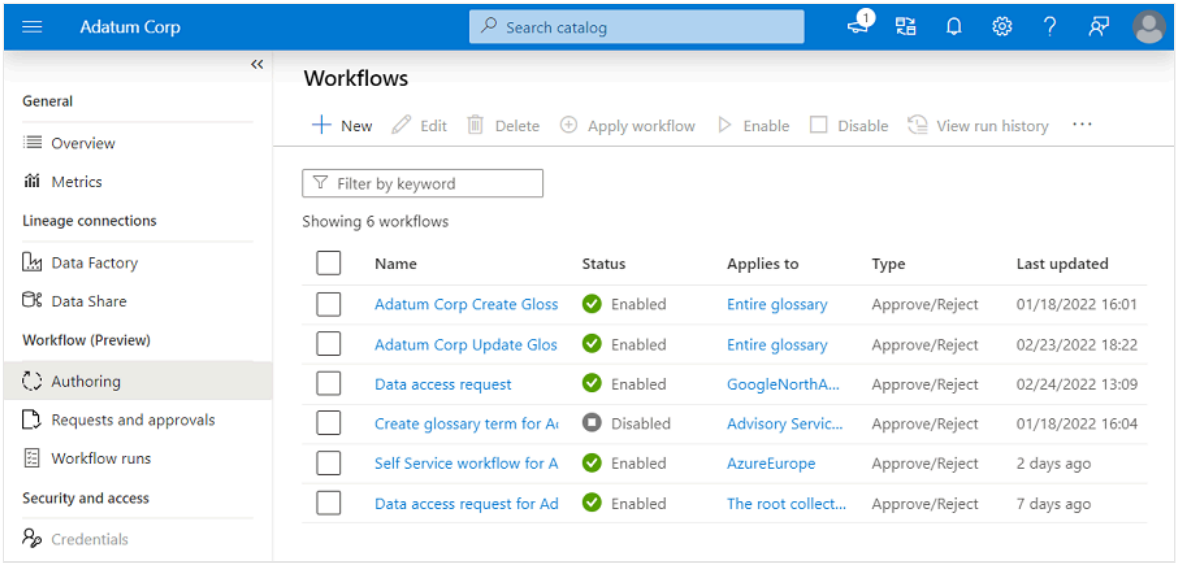
This guide will take you through the creation and management of approval workflows for asset curation.

Create and enable a new approval workflow for asset curation

1. Sign in to the Microsoft Purview governance portal and select the Management center. You'll see three new icons in the table of contents.



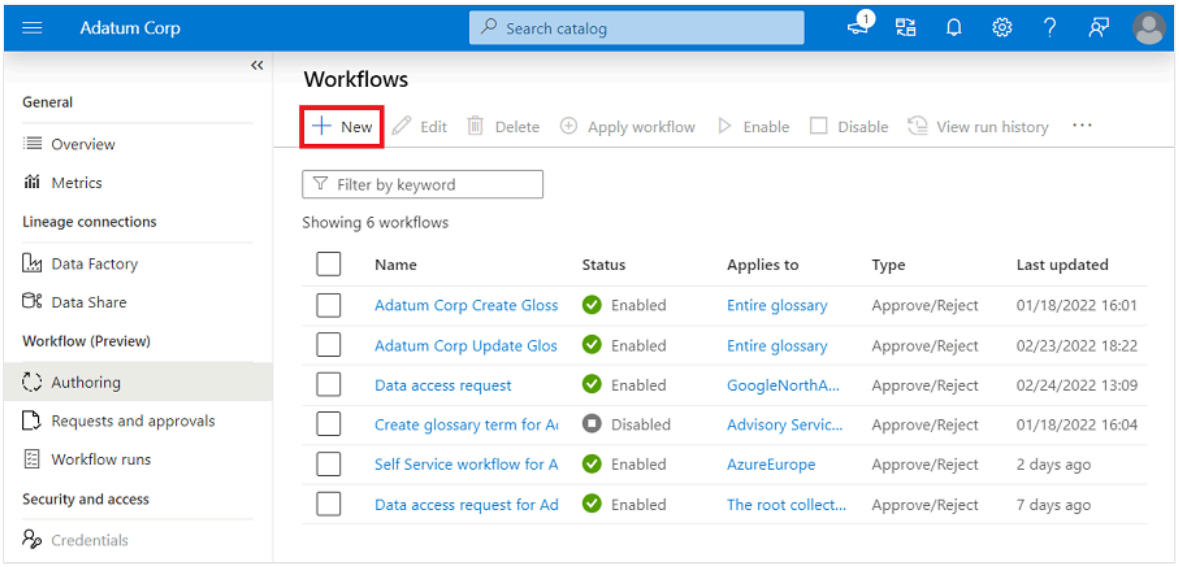
2. To create new workflows, select **Authoring** in the workflow section. This will take you to the workflow authoring experiences.



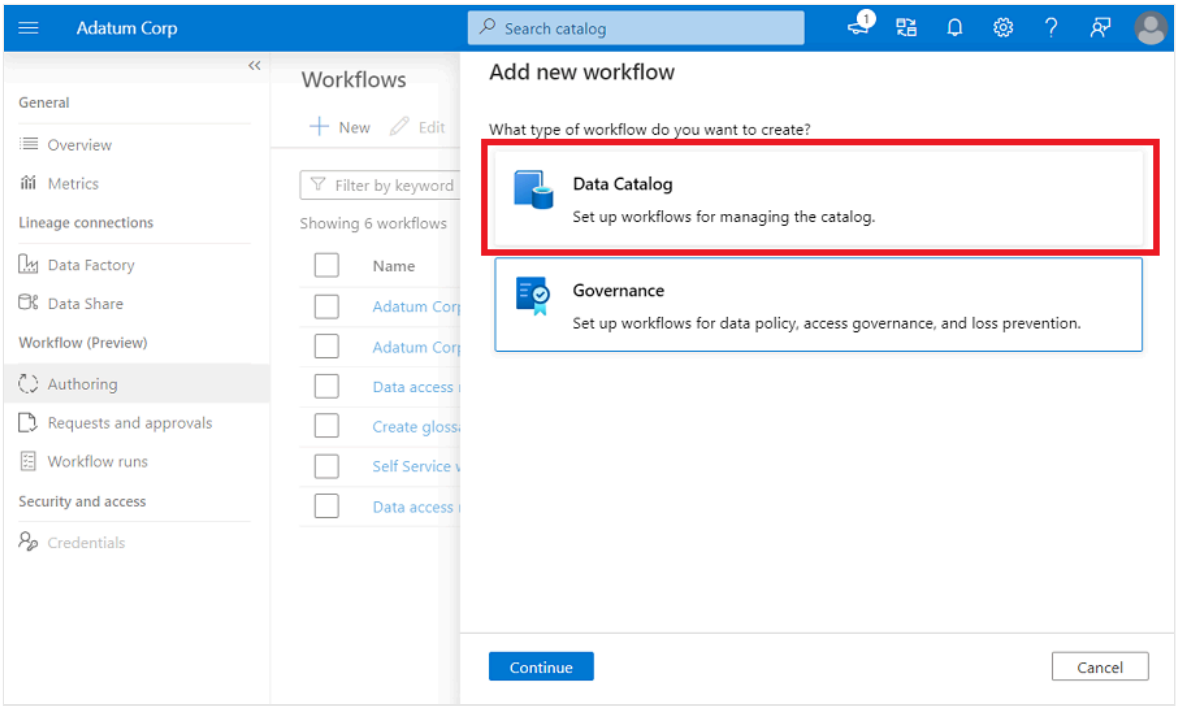
> **Note**

If the authoring tab is greyed out, you don't have the permissions to be able to author workflows. You'll need the workflow admin role.

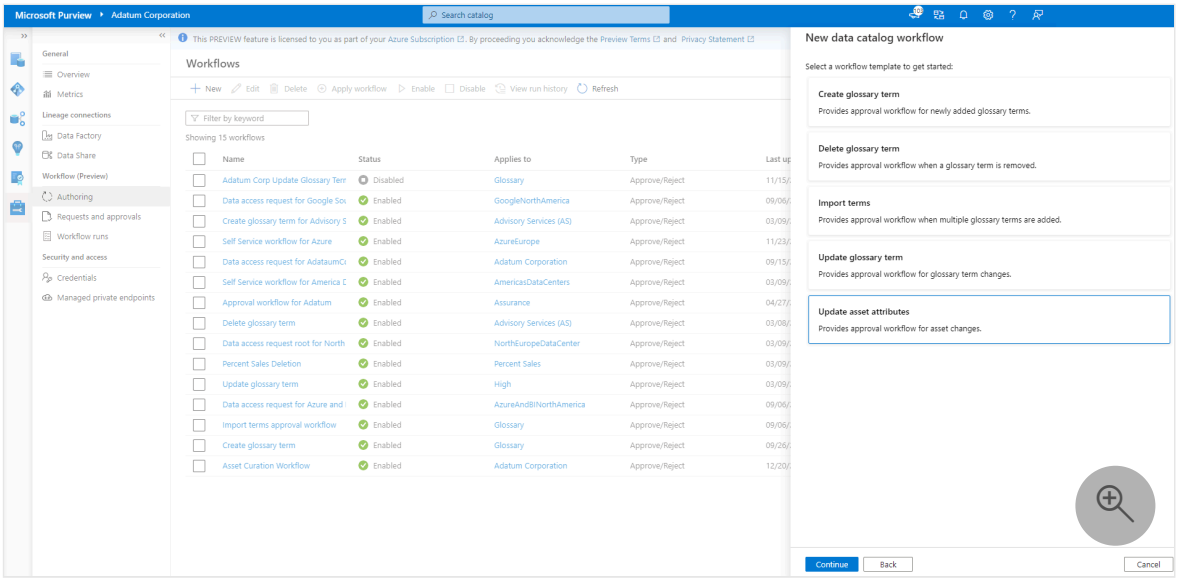
3. To create a new workflow, select the **+New** button.



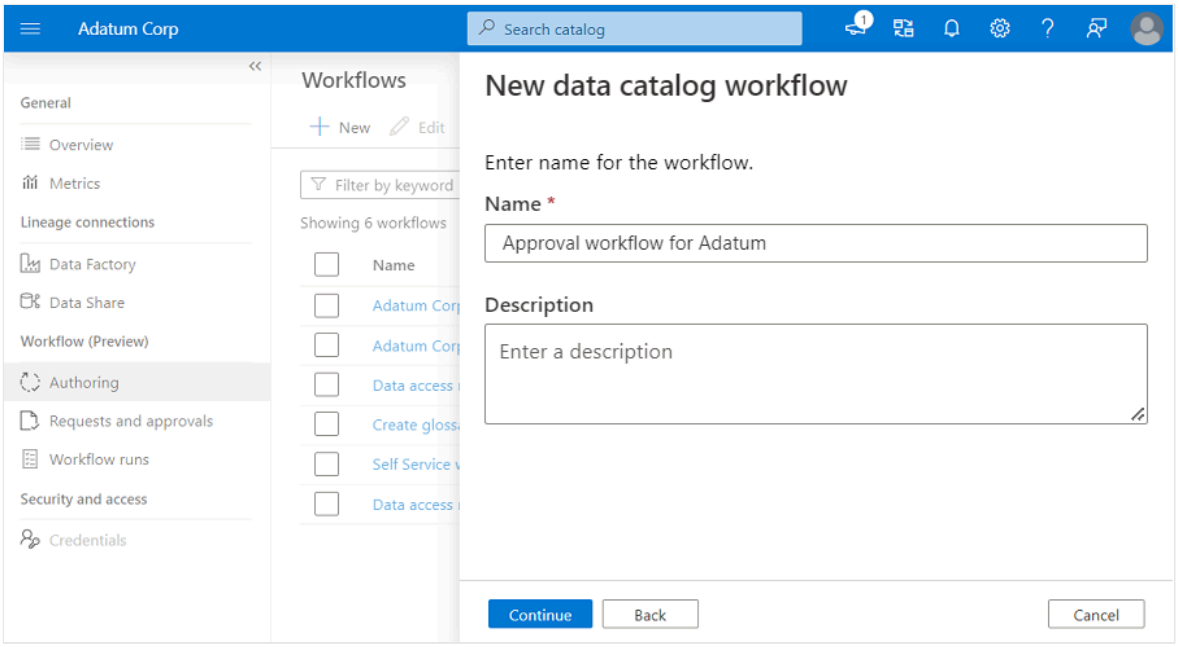
4. To create **Approval workflows for asset curation** Select **Data Catalog** and select **Continue**



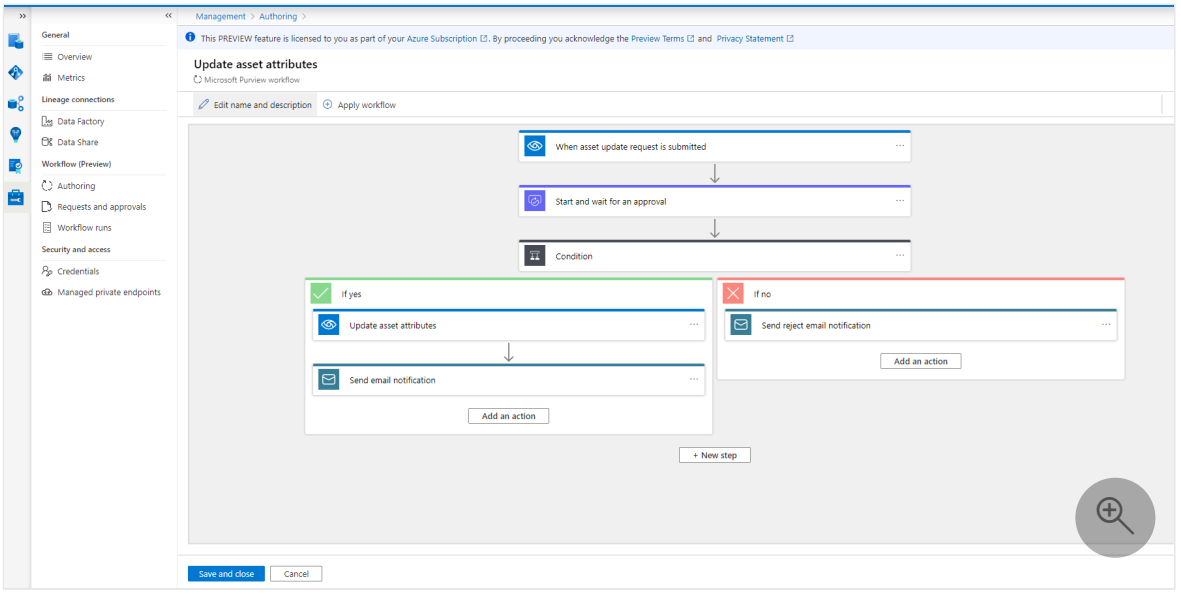
5. In the next screen, you'll see all the templates provided by Microsoft Purview to create a workflow. Select the template using which you want to start your authoring experiences and select **Continue**. Each of these templates specifies the kind of action that will trigger the workflow. In the screenshot below we've selected **Update asset attributes** to create approval workflow for asset updates.



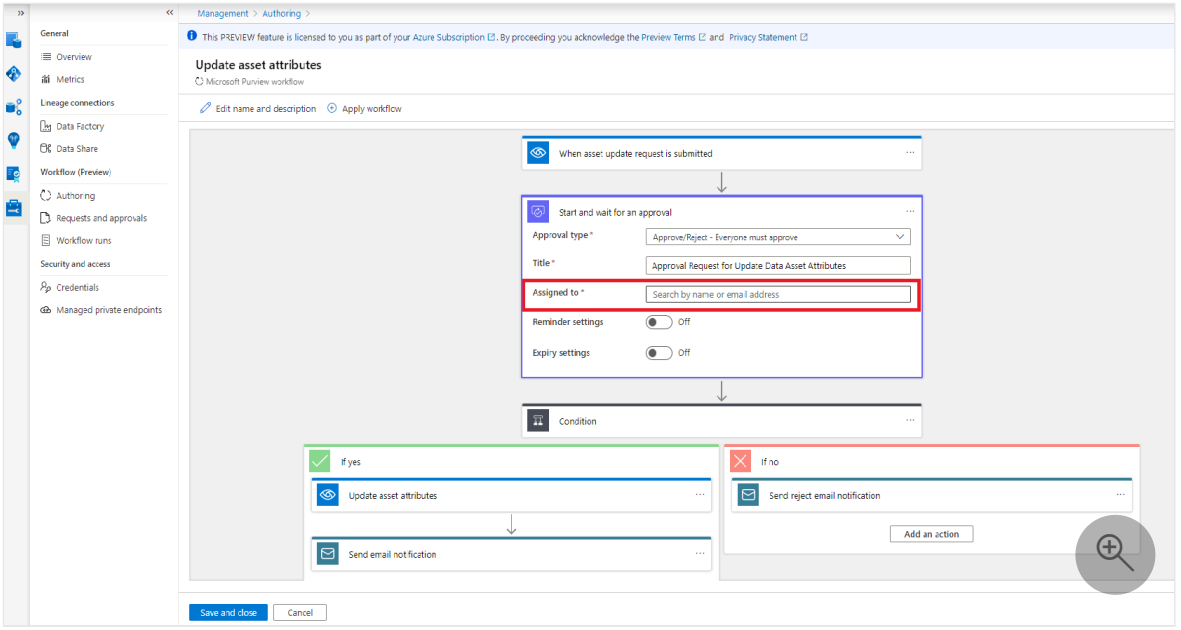
6. Next, enter a workflow name and optionally add a description. Then select **Continue**.



7. You'll now be presented with a canvas where the selected template is loaded by default.



8. The default template can be used as it is by populating the approver's email address in **Start and Wait for approval** Connector.



The default template has the following steps:

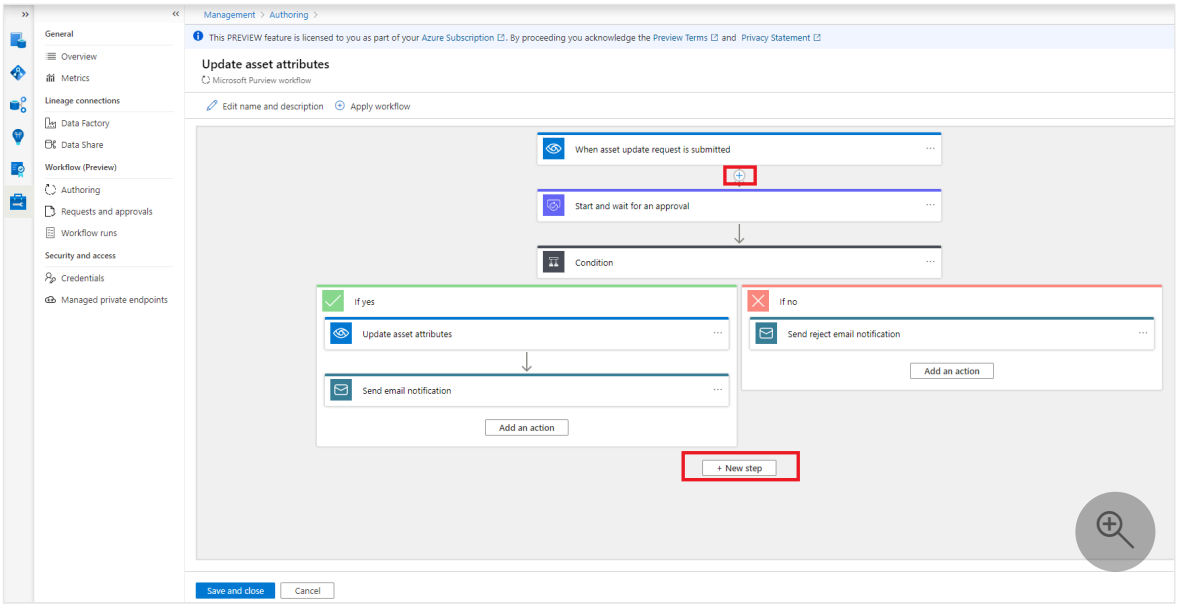
a. Trigger when an asset is updated. The update can be done in overview, schema or contacts tab.

b. Approval connector that specifies a user or group that will be contacted to approve the request.

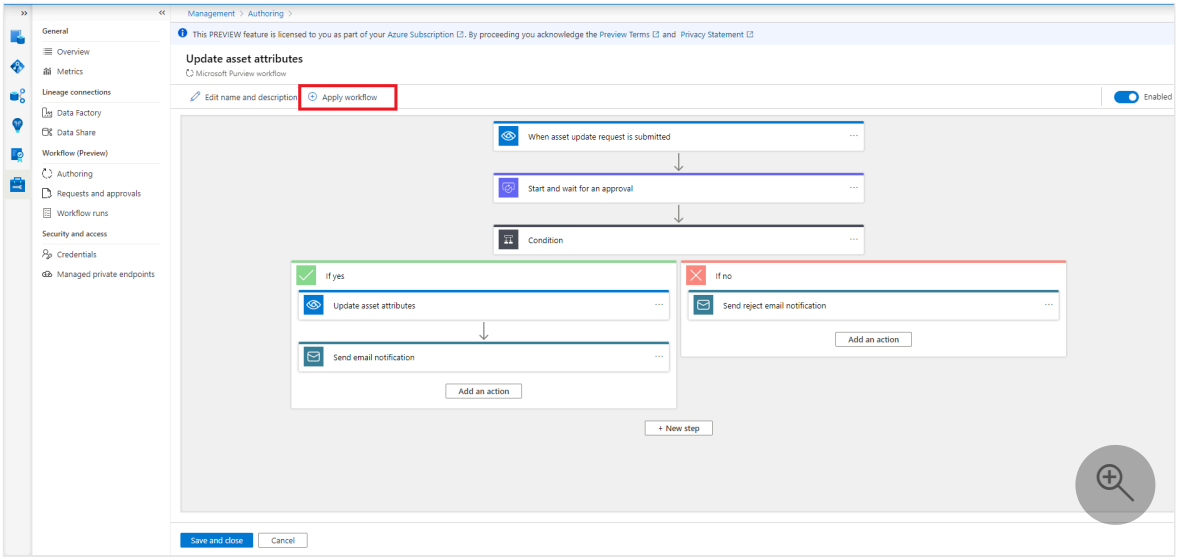
c. Condition to check approval status

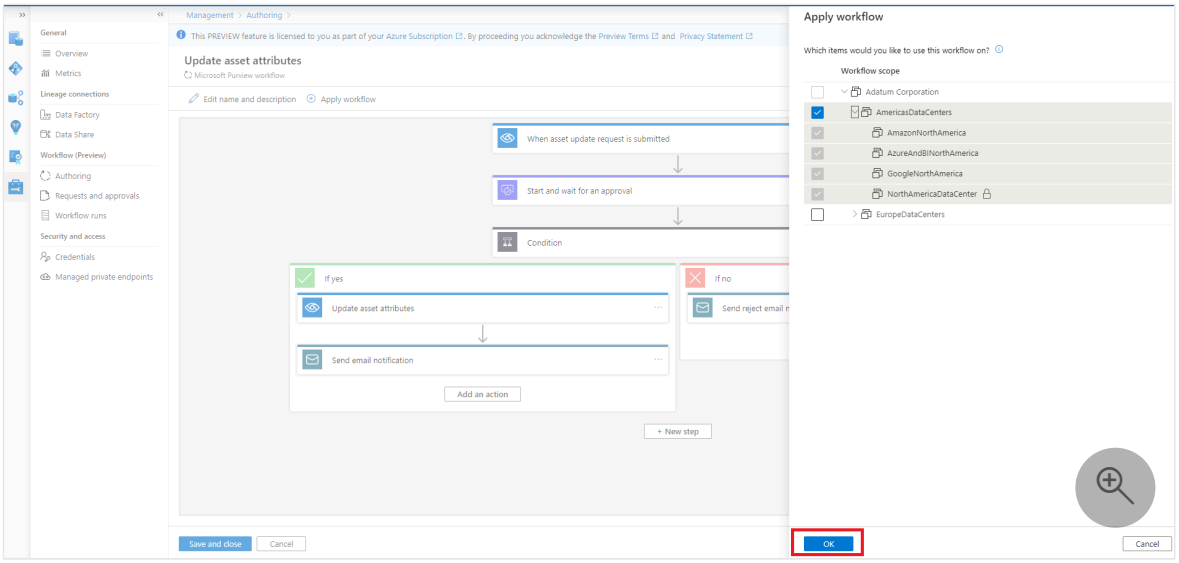
- If approved:
  i. Update the asset in Purview data catalog.
  ii. Send an email to requestor that their request is approved, and asset update operation is successful.
- If rejected:
  i. Send email to requestor that their asset update request is denied.

9. You can also modify the template by adding more connectors to suit your organizational needs. Add a new step to the end of the template by selecting the **New step** button. Add steps between any already existing steps by selecting the arrow icon between any steps.



10. Once you're done defining a workflow, you need to bind the workflow to a collection hierarchy path. The binding implies that this workflow is triggered only for update operation on data assets in that collection. A workflow can be bound to only one hierarchy path. To bind a workflow or to apply a scope to a workflow, you need to select 'Apply workflow'. Select the scopes you want this workflow to be associated with and select **OK**.

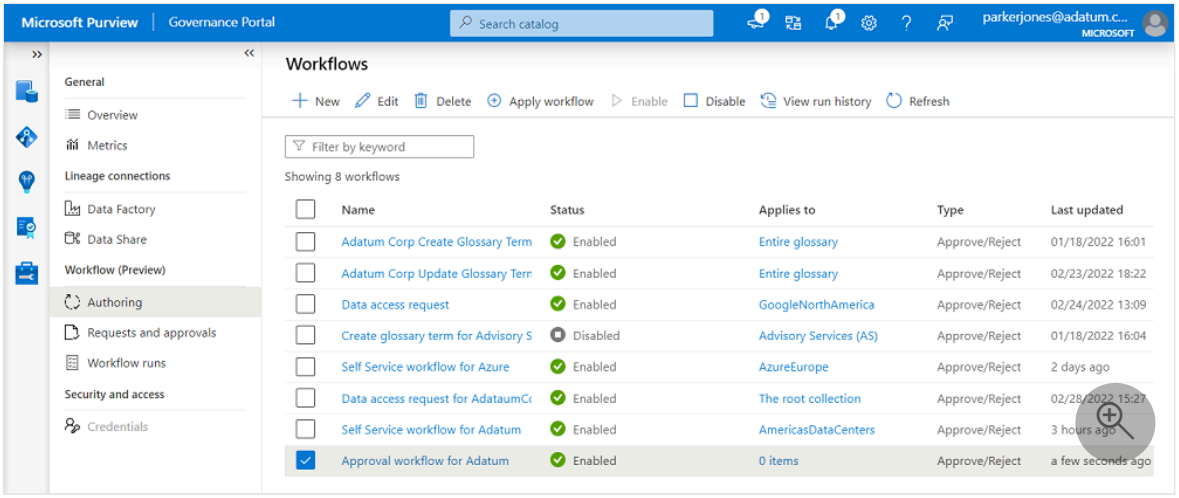




> **Note**
>
> The Microsoft Purview workflow engine will always resolve to the closest workflow that the collection hierarchy path is associated with. In case a direct binding is not found, it will traverse up in the tree to find the workflow associated with the closest parent in the collection tree.

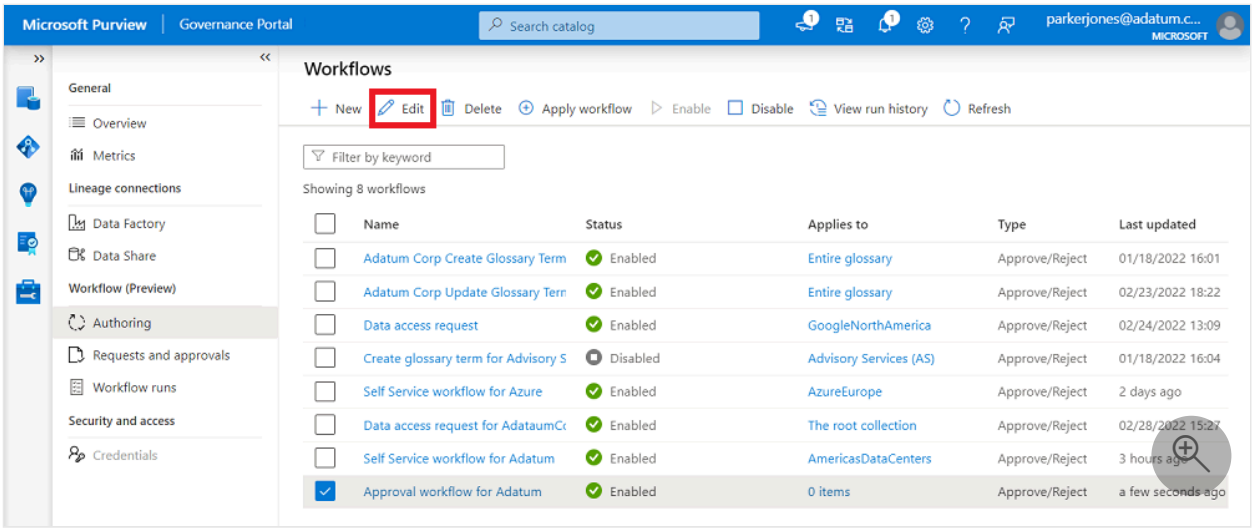
11. By default, the workflow will be enabled. To disable, toggle the Enable button in the top menu.

12. Finally select **Save and close** to create and the workflow.



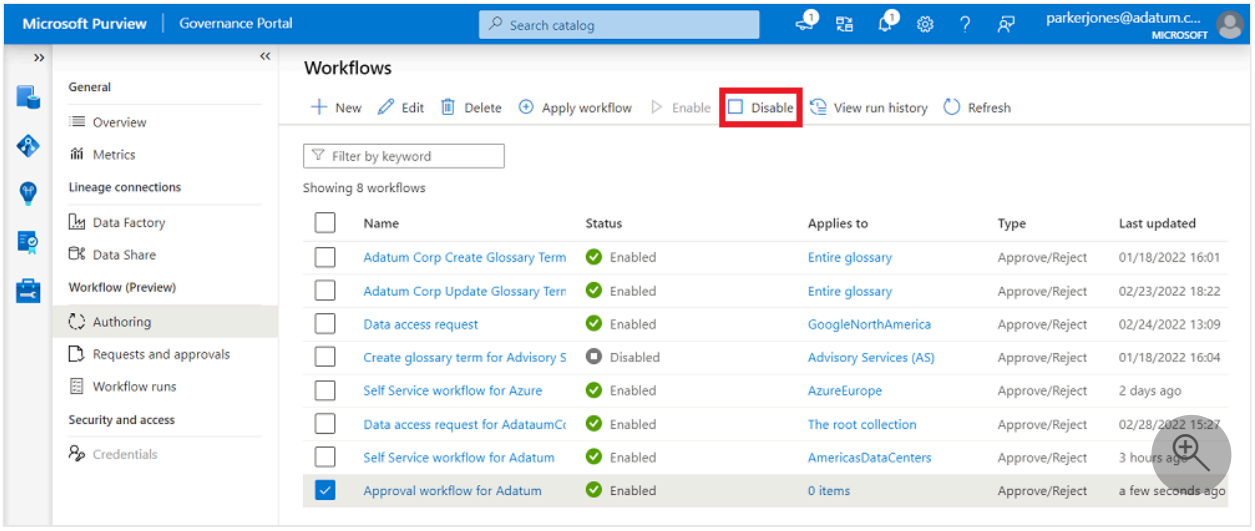
Edit an existing workflow

To modify an existing workflow, select the workflow and then select **Edit** in the top menu. You'll then be presented with the canvas containing workflow definition. Modify the workflow and select **Save** to commit changes.



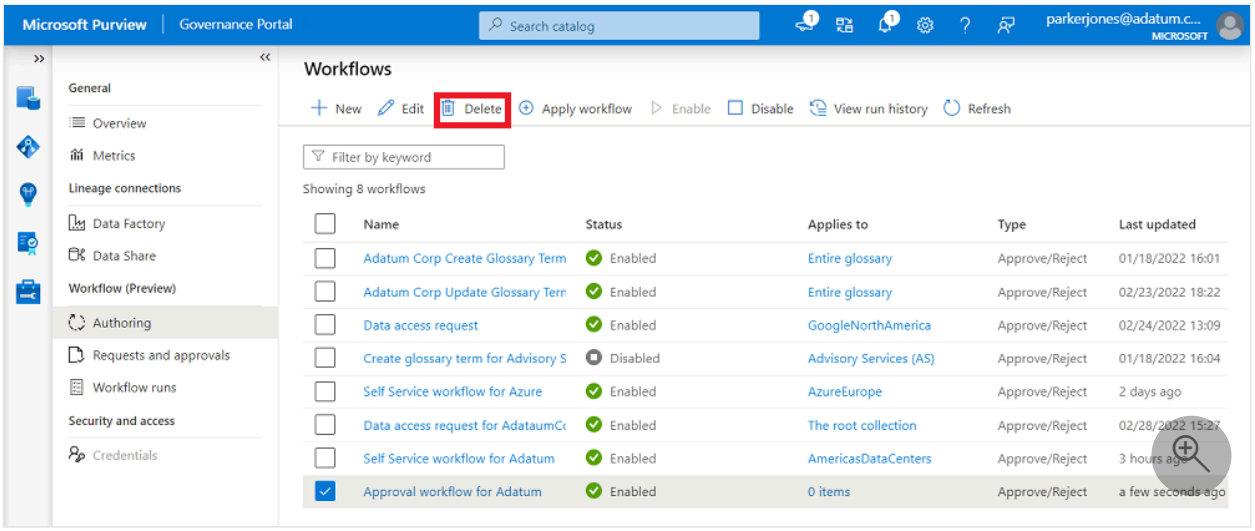
Disable a workflow

To disable a workflow, select the workflow and then select **Disable** in the top menu. You can also disable the workflow by selecting **Edit** and changing the enable toggle in workflow canvas.



Delete a workflow

To delete a workflow, select the workflow and then select **Delete** in the top menu.



Limitations for asset curation with approval workflow enabled

- Lineage updates are directly stored in Purview data catalog without any approvals.

Next steps

For more information about workflows, see these articles:

- What are Microsoft Purview workflows
- Self-service data access workflow for hybrid data estates
- Manage workflow requests and approvals

Feedback

Was this page helpful? Yes No

Manage workflow requests and approvals

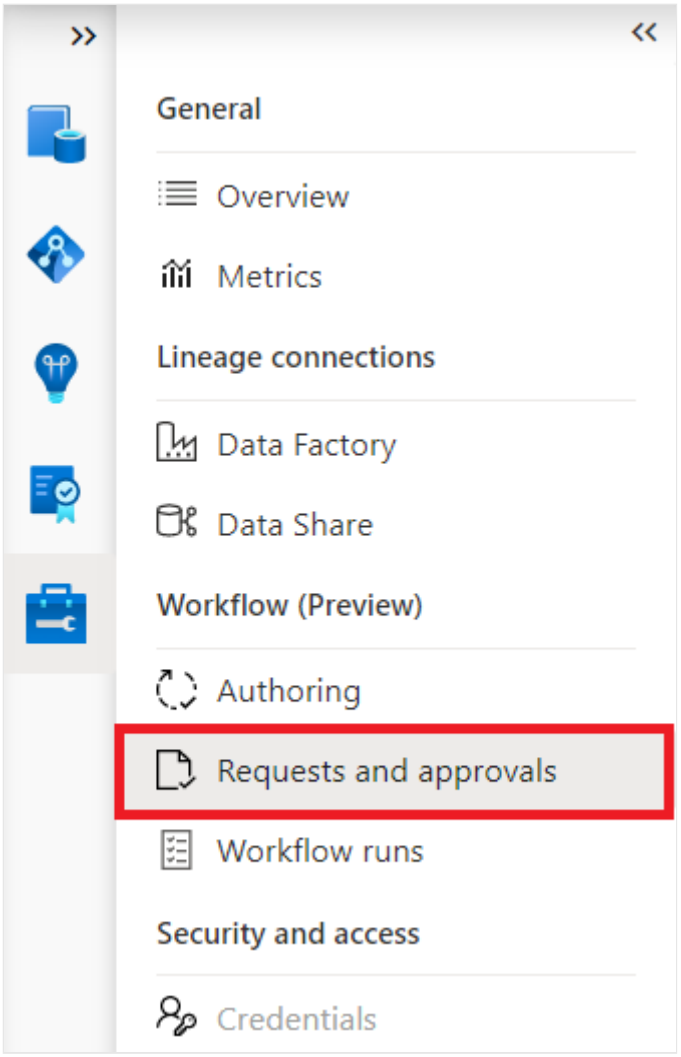
Article • 02/26/2025

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

> **Important**
>
> This feature is currently in preview. The Supplemental Terms of Use for Microsoft Azure Previews include additional legal terms that apply to Azure features that are in beta, in preview, or otherwise not yet released into general availability.

This article outlines how to manage requests and approvals that generated by a workflow in Microsoft Purview.

To view requests you've made or request for approvals that have been sent to you by a workflow instance, navigate to management center in the Microsoft Purview governance portal, and select **Requests and Approvals**.

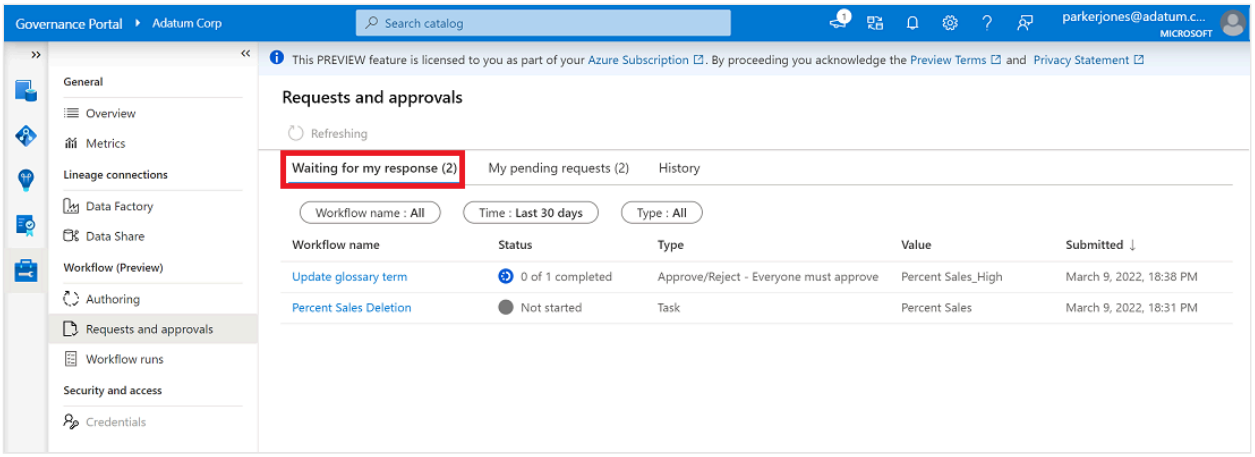


You'll be presented with three tabs:

- Waiting for a response - This tab shows the requests (tasks) and approvals that are waiting for you to act on.
- Pending requests - You can view all the approval requests and tasks you've submitted in this tab.
- History - All the completed approvals and tasks are moved to this tab.

Waiting for a response

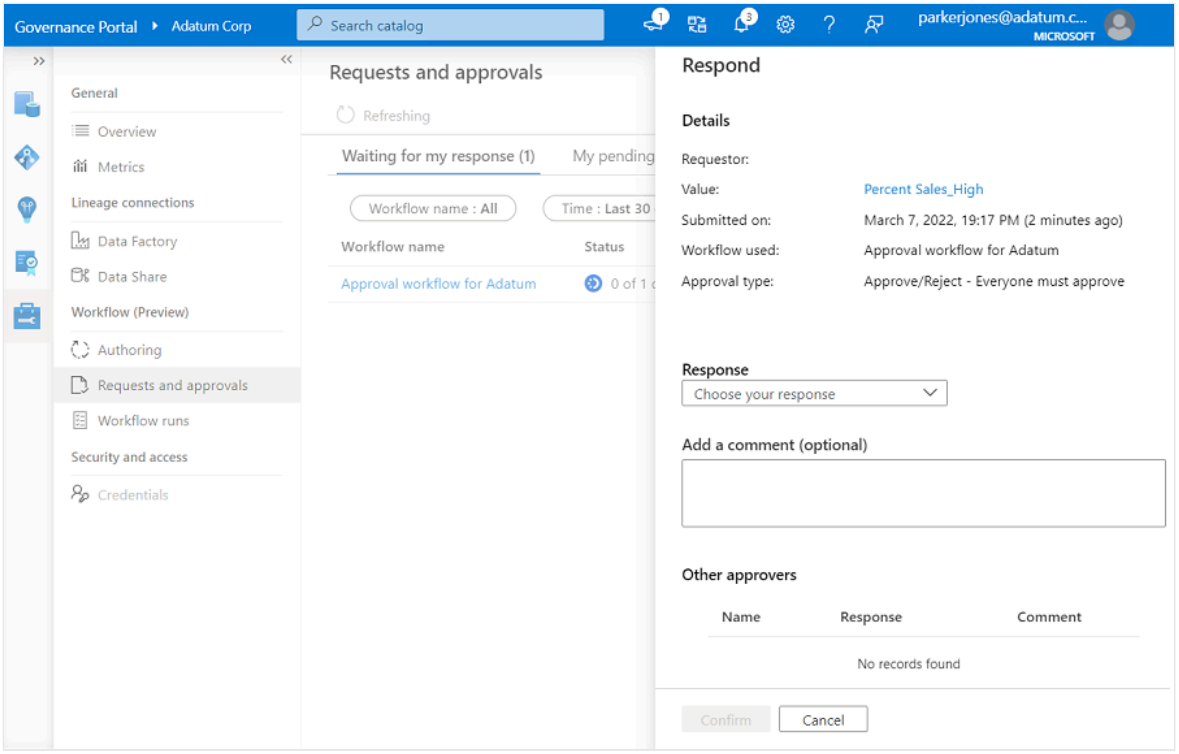
This tab shows the requests (tasks) and approvals that are waiting for your action.



Select the request to take action.

Approvals

1. To approve/reject a request, select the request and you'll be presented with the following window:

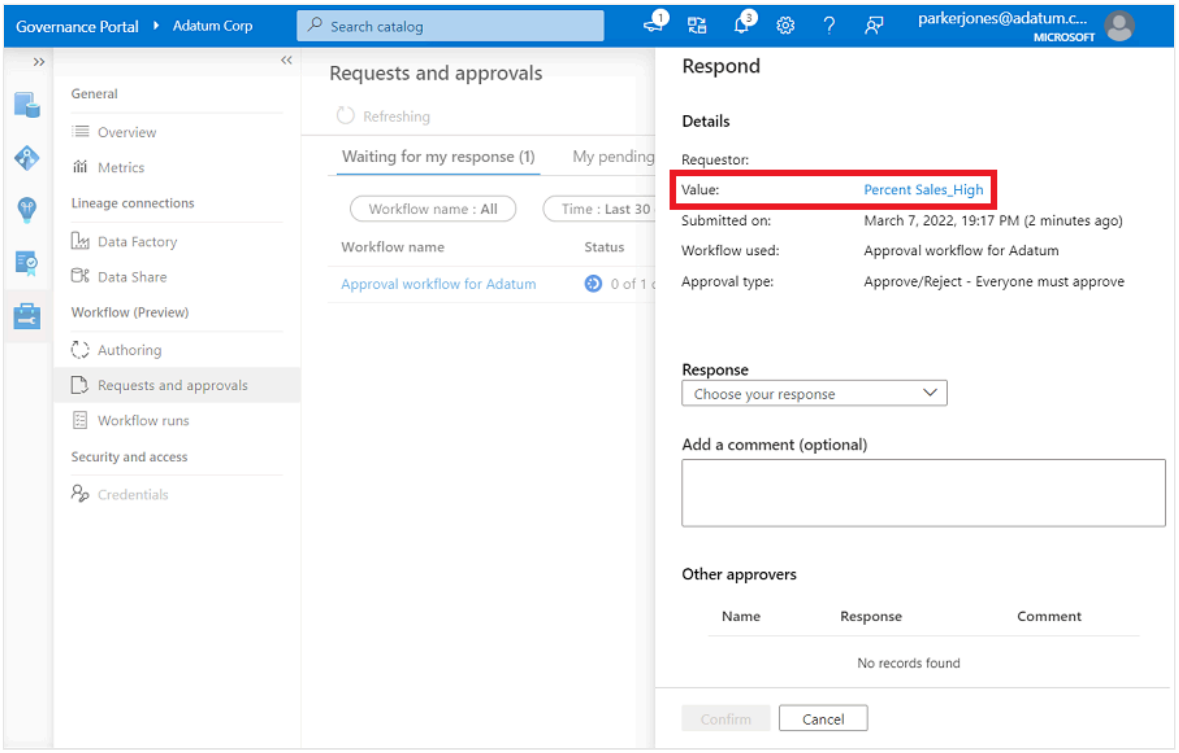


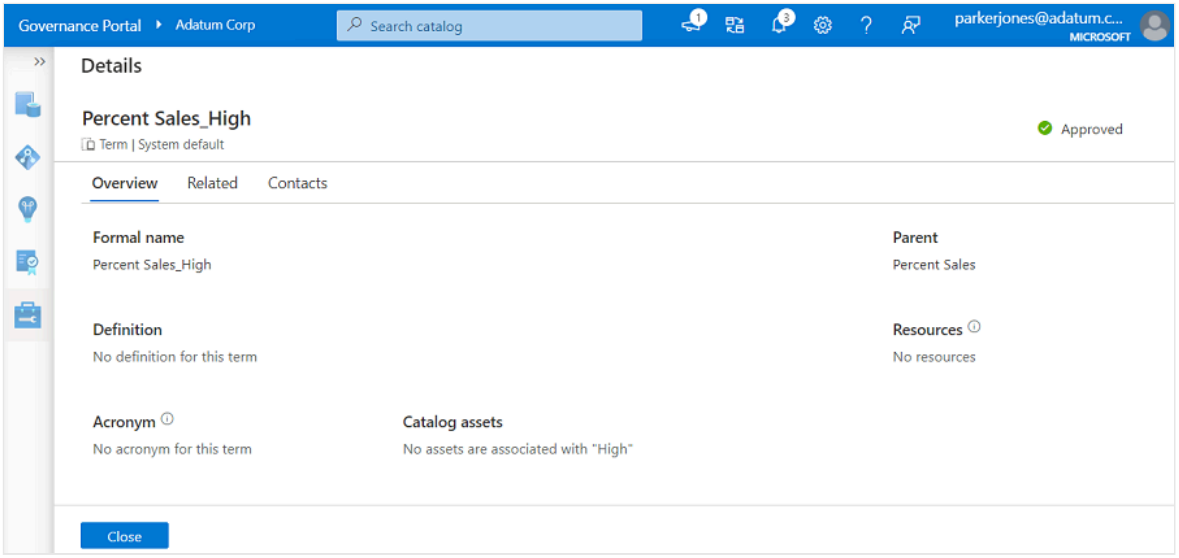
2. An approval activity has the following available responses:

   - **Approved** – An approver can mark the response as **Approved** indicating their approval for the changes proposed by the requestor.
   - **Rejected** - An approver can mark the response as **Rejected** indicating that they don't approve the changes proposed by the requestor.

3. Optionally, select the value to see details of the request. In the screenshot below, it shows the term details. If the approval is for a data asset, you'll be able to view the

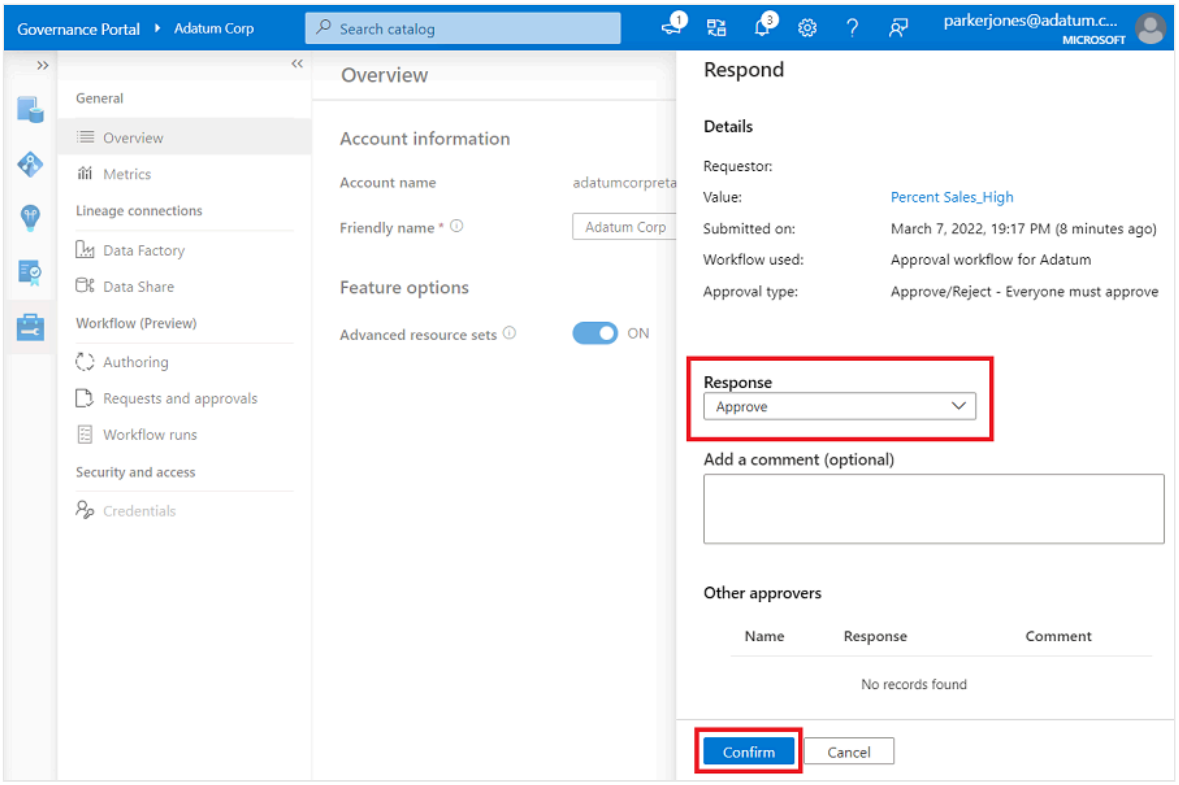
details, as shown below.





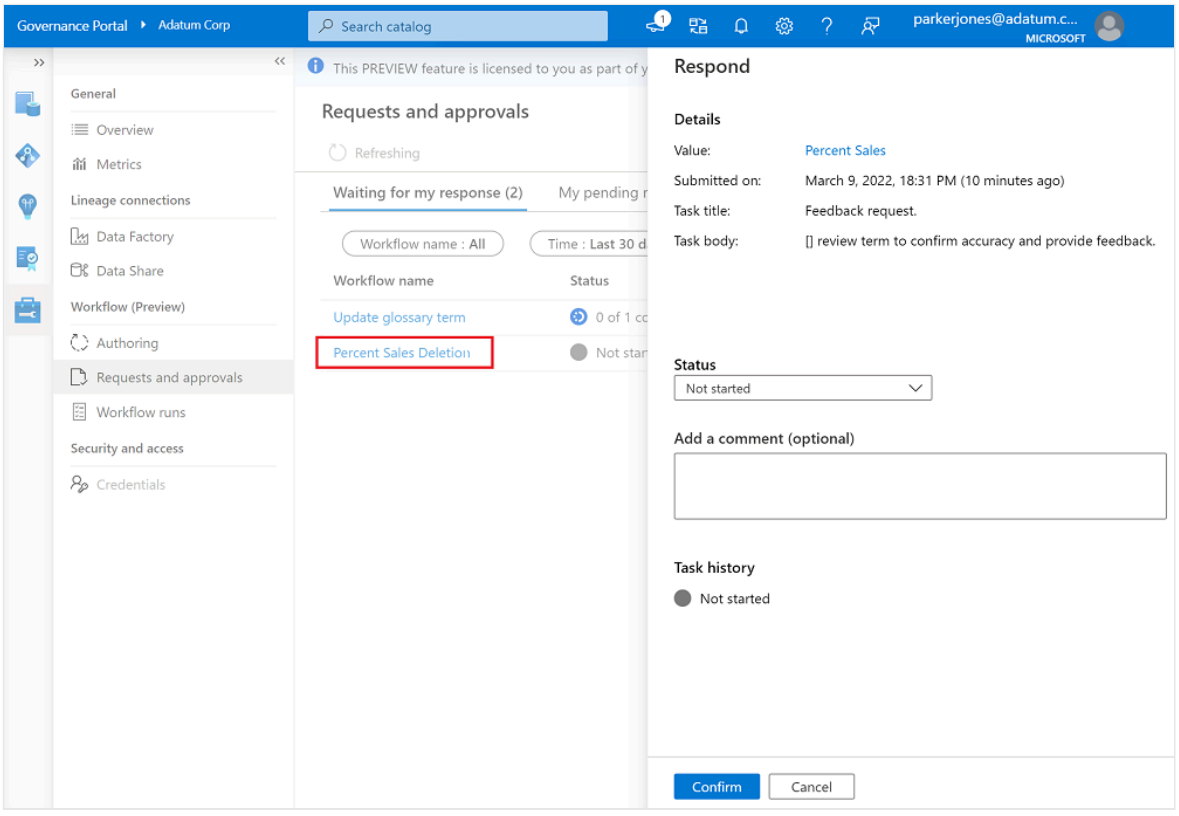
4. If there are updates, you'll also to be able to see the current value and proposed value.

5. Choose your response, optionally add comments, and select **Confirm**.



Tasks

1. To complete a task, select the task request and you'll be presented with the following window:



2. A task has the following statues:

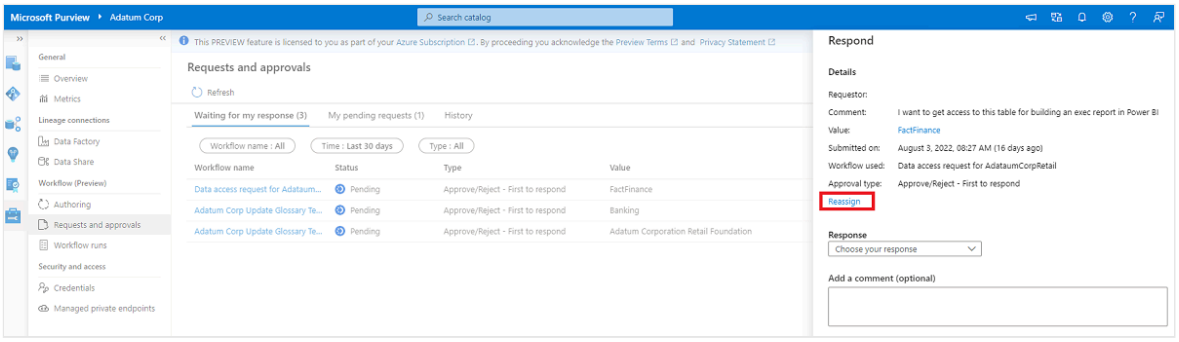
- Not started – This is the status of the task when it's initially created by a workflow.
- In Progress – A task owner can mark the task as **In progress** to indicate that they're currently working on it.
- Complete – Once the task is complete, a task owner can change the status as **Complete**. This marks the completion of task activity, and the workflow will now move to the next step.

3. Select the correct status, add any comments, and select **Confirm**.

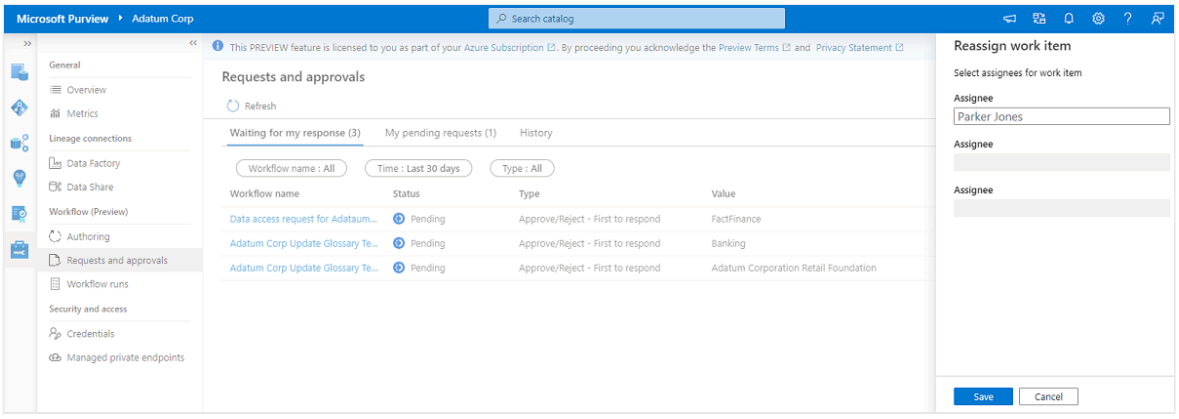
Reassign requests

You can reassign requests both approvals and tasks that are assigned to you to a different user.

1. To reassign, select the request or task you're assigned and select **Reassign** in the following window.



2. You'll be presented with a list of all the users who are assigned to the request. Select **Assignee** where your user name or the group you're part of is displayed and change it from your user name to the new user name. Select **Save** to complete the reassignment.



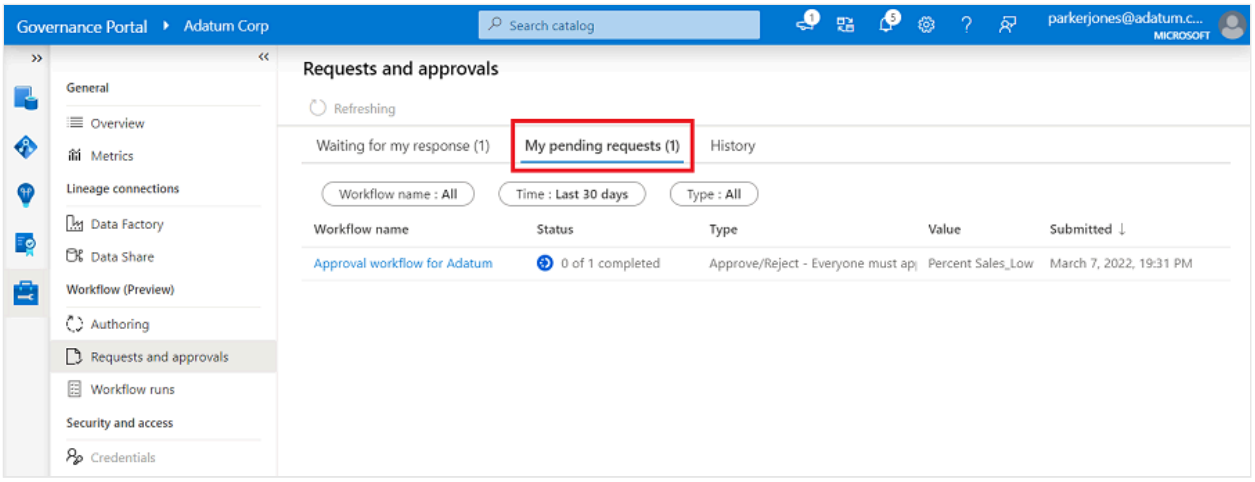
> **Note**

You can only re-assign your user ID or group you are part of to another user or group. The other assignees will be greyed out and will not available for re-assignment.

Pending requests

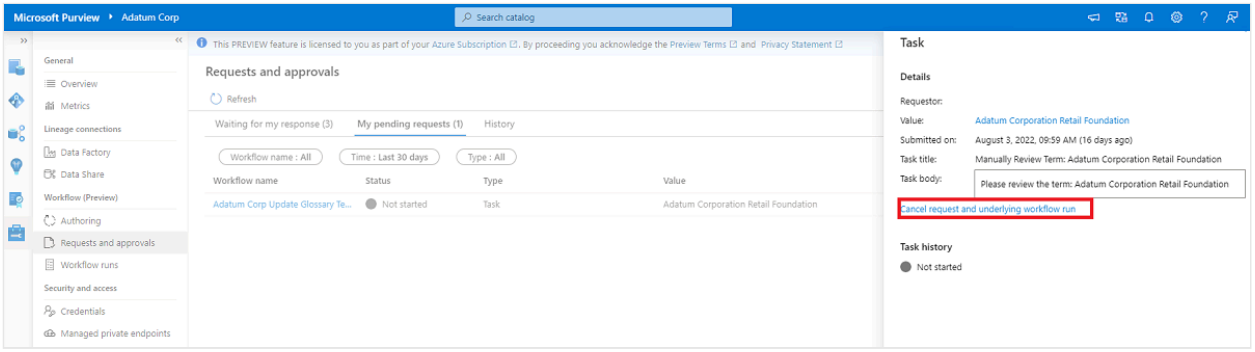
In this tab you can view all the approval requests and tasks that you've submitted.

Select the request to see the status and the outcomes for each approver/task owner.



Cancel workflows

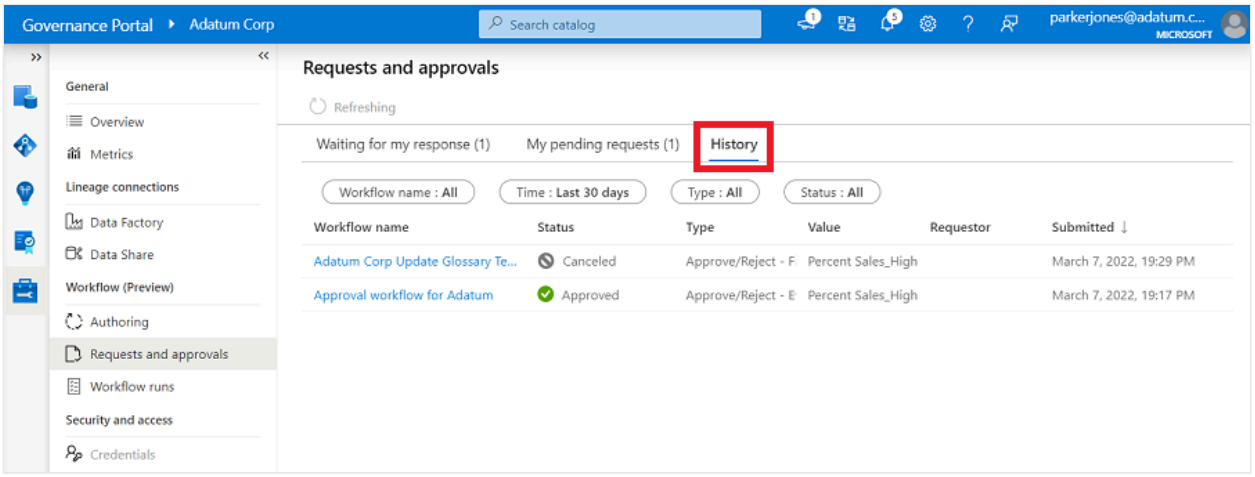
You can cancel a submitted request and its underlying workflow by selecting **Cancel request and its underlying workflow run**.



> **Note**
>
> You can only cancel workflows that are in progress. When you cancel a request from the **requests and approvals** section, it will cancel the underlying workflow run.

History

All the completed approvals and tasks are moved to this tab.



Select an approval or task to see details and responses from all approvals or task owners.

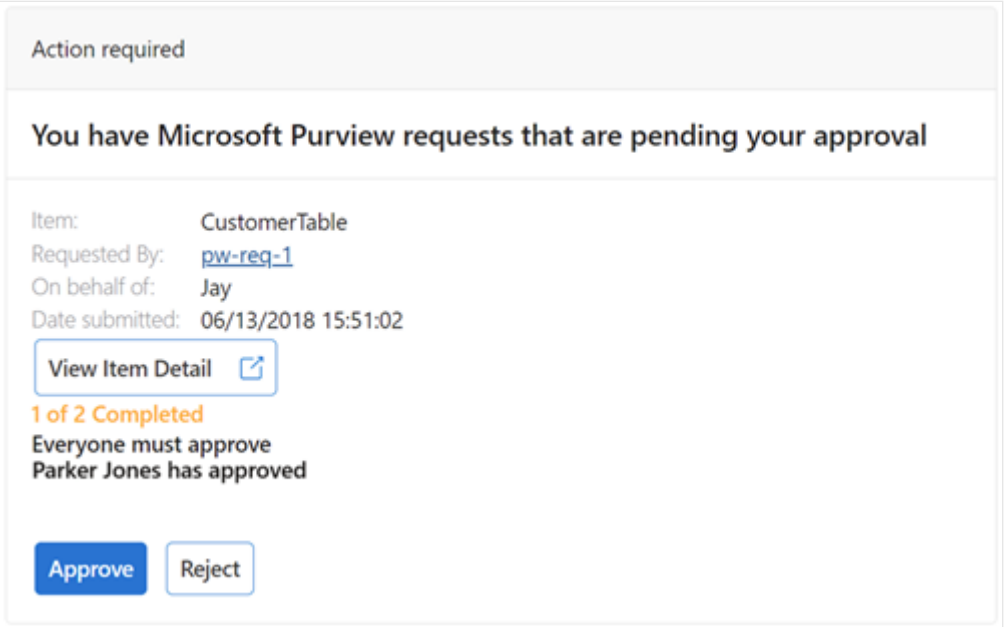
Email notifications

Purview approvals and task connectors have in-built email capabilities. Every time an approval/task action is triggered in workflow; it sends email to all the users who need to act on it.

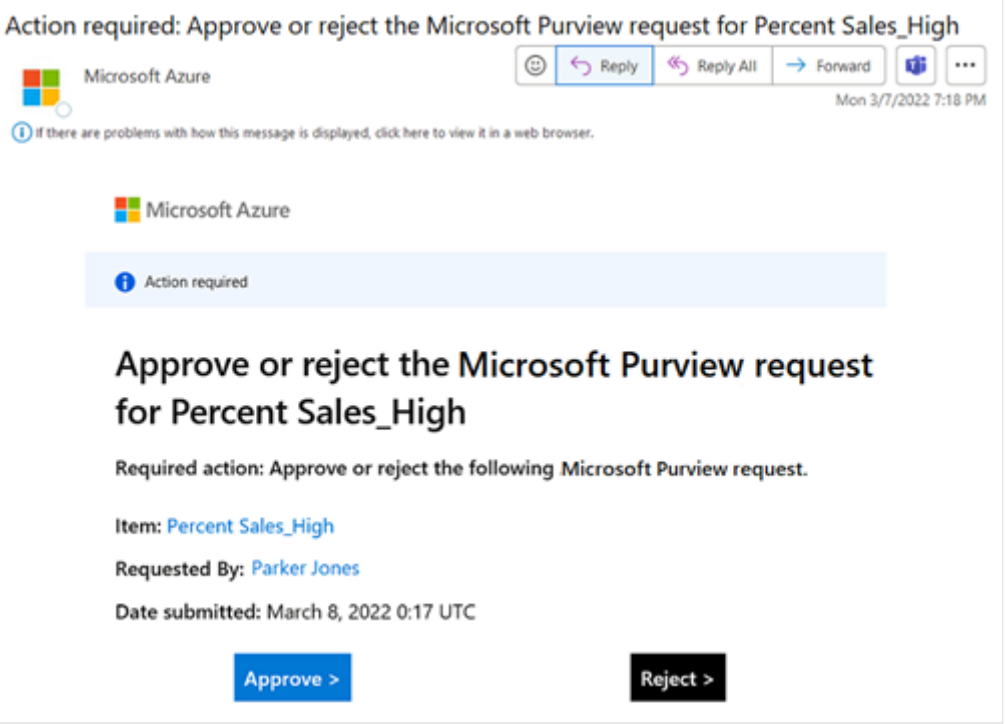
There are two kinds of emails that can be sent:

- **Actionable messages** - if your workflow and email client meet the requirements, your approvers can receive single and aggregated actionable approval emails where they can leave comments and reject or approve requests within outlook. For more information, the approver can also select the approval item in the message to be taken to request details in the Microsoft Purview portal.

  > **ⓘ Important**
  >
  > If your workflow or email client don't meet the **requirements**, your users might receive 400/401/403 errors when using actionable messages.



- **Standard messages** - if your workflow and email client don't meet the requirements, your approvers will receive single and aggregated approval emails that will direct them to the approval experience in the Microsoft Purview portal.



Users can respond by selecting the links in the email, for **actionable messages** they can submit responses directly in the email, or by navigating to the Microsoft Purview governance portal and viewing their pending tasks.

Next steps

- What are Microsoft Purview workflows
- Approval workflow for business terms

- Self-service data access workflow for hybrid data estates
- Manage workflow runs

Feedback

Was this page helpful? Yes No

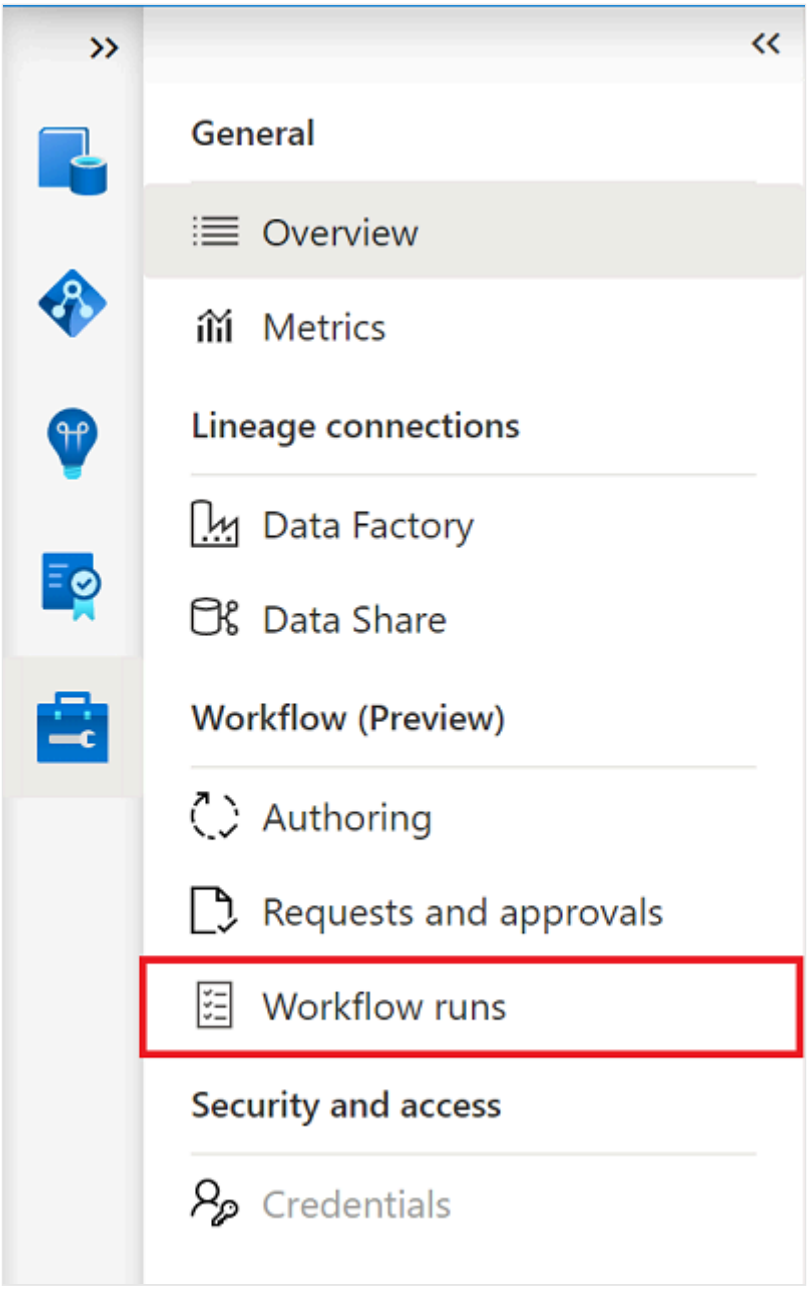
Manage workflow runs

> **Note**
>
> Microsoft Purview Data Catalog (classic), Data Health Insights (classic), and Purview Workflow (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

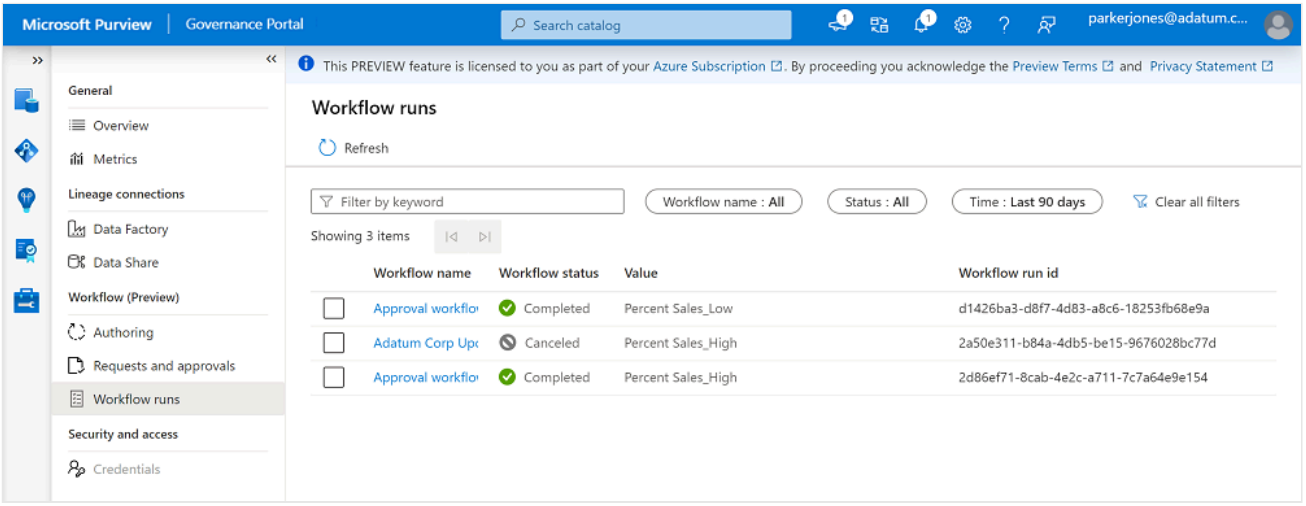
> **Important**
>
> This feature is currently in preview. The Supplemental Terms of Use for Microsoft Azure Previews include additional legal terms that apply to Azure features that are in beta, in preview, or otherwise not yet released into general availability.

This article outlines how to manage workflows that are already running.

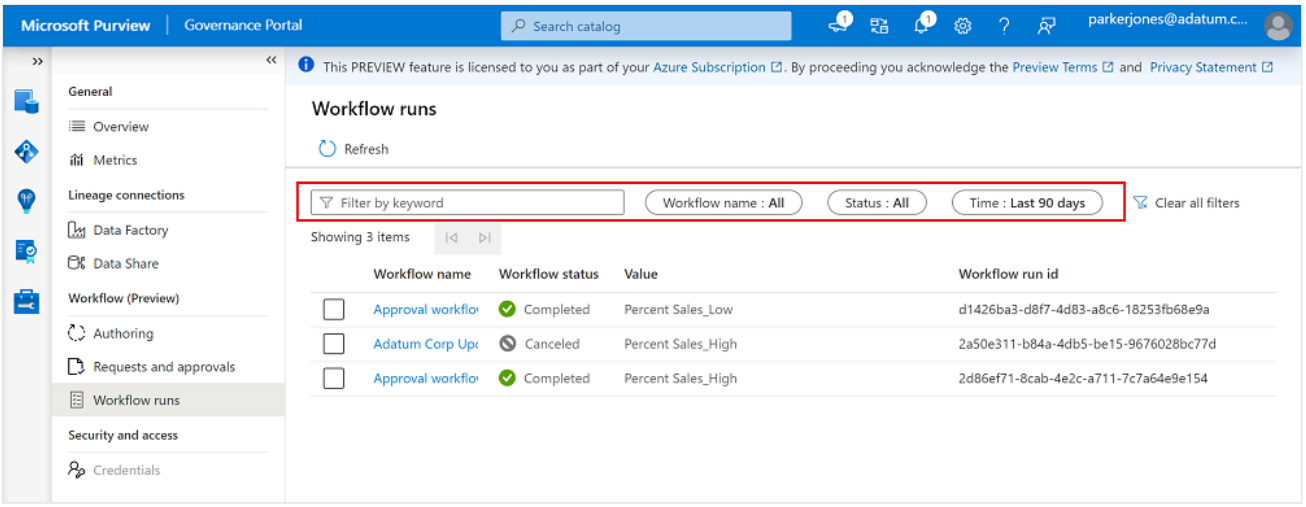
1. To view workflow runs you triggered, sign in to the Microsoft Purview governance portal, select the Management center, and select **Workflow runs**.



2. You'll be presented with the list of workflow runs and their statuses.

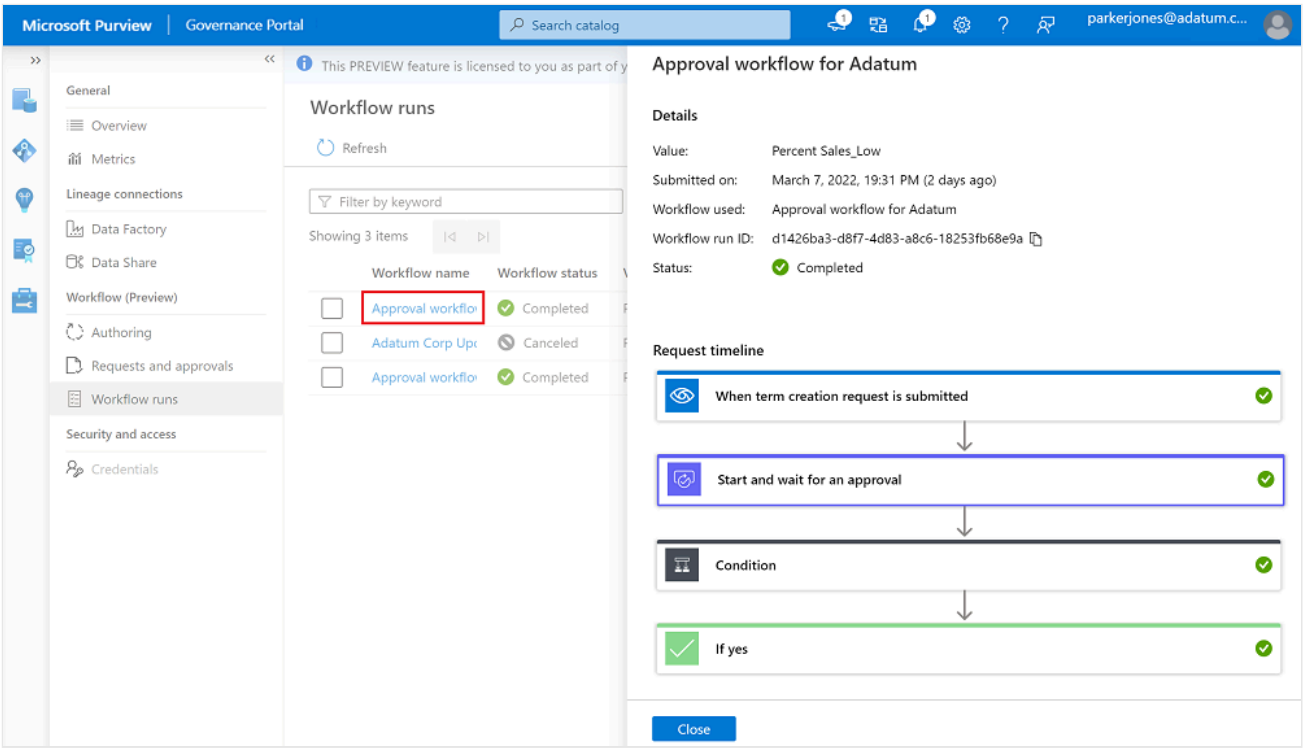


3. You can filter the results by using workflow name, status, or time.

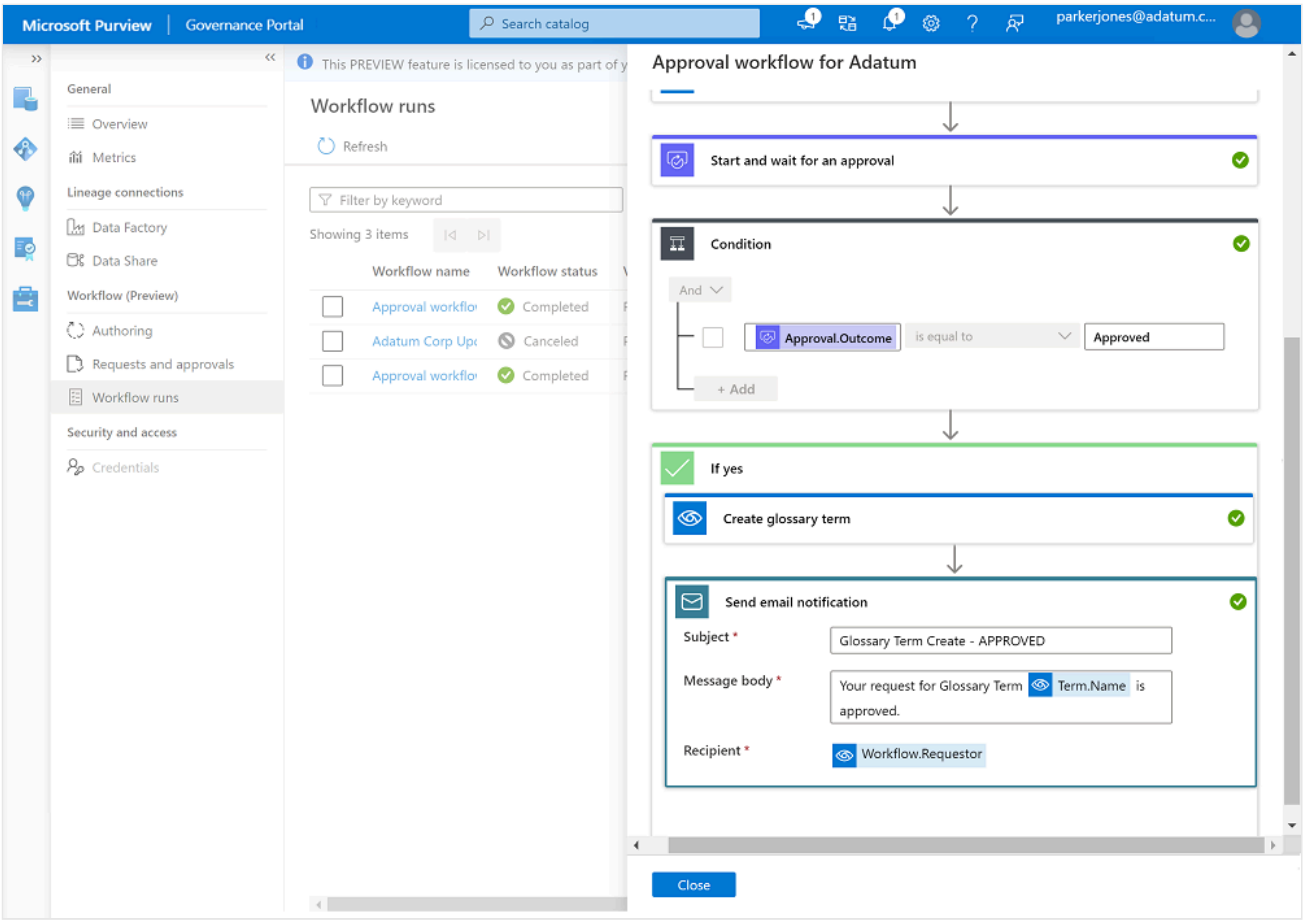


4. Select a workflow name to see the details of the workflow run.

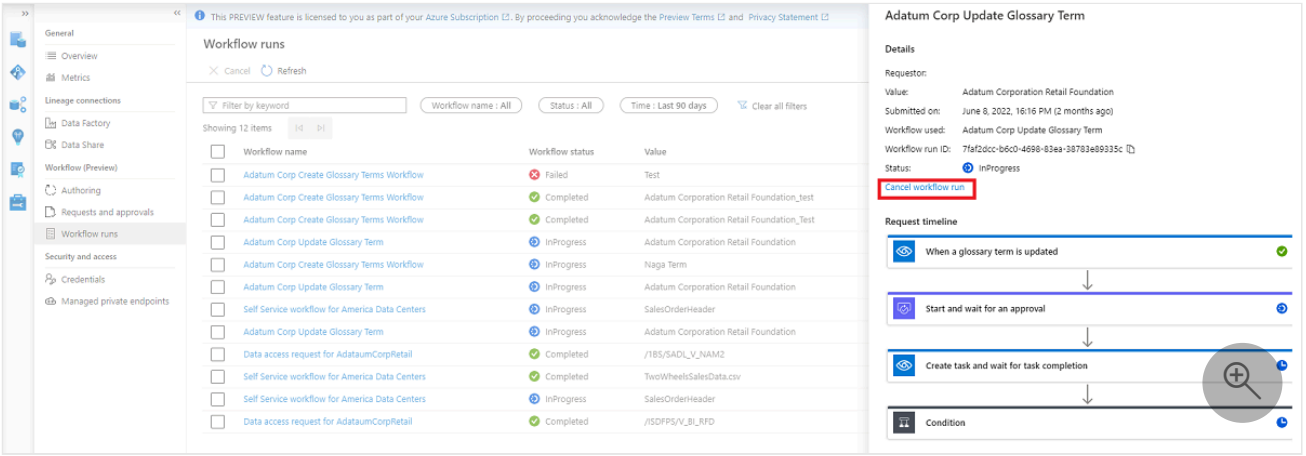
5. This will present a window that shows all the actions that are completed, actions that are in-progress, and the next action for that workflow run.



6. You can select any of the actions in the request timeline to see the specific status and sub steps details.



7. You can cancel a running workflow by selecting **Cancel workflow run**.



> **Note**
>
> You can only cancel workflows that are in progress.

Next steps

- What are Microsoft Purview workflows
- Approval workflow for business terms
- Self-service data access workflow for hybrid data estates

- Manage workflow requests and approvals

Last updated on 10/10/2025

ServiceNow actions in Microsoft Purview workflows

Article • 02/24/2025

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

> **Important**
>
> This feature is currently in preview. The Supplemental Terms of Use for Microsoft Azure Previews include additional legal terms that apply to Azure features that are in beta, in preview, or otherwise not yet released into general availability.

You can use ServiceNow actions in Microsoft Purview workflows to manage and update your ServiceNow records during your data catalog and data governance workflows.

For a list of all available workflow actions, see our workflow actions page. To connect to your ServiceNow instance, follow the instructions below.

Set up authentication

Microsoft Purview uses Azure Key Vault to securely store the credentials it uses to connect to other services. Before we can connect to ServiceNow we need to create our ServiceNow credential in the Azure Key vault. Follow these steps:

1. If you don't have an Azure Key Vault, you can follow this documentation to create one.
2. Create a secret in your key vault. The value will be your ServiceNow basic authentication password.
3. Grant Microsoft Purview access to your Azure Key Vault.
4. Create a connection in Microsoft Purview to your Azure Key Vault.
5. Create a new basic authentication credential in Microsoft Purview, providing your ServiceNow username and your key vault secret.

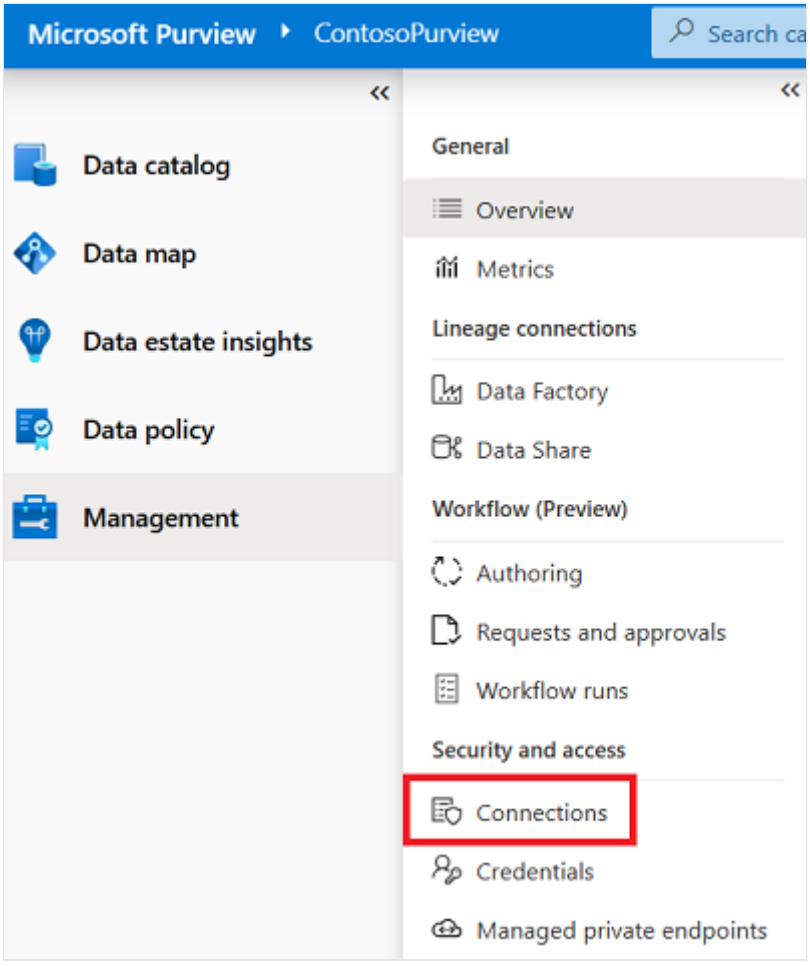
Now we have all the authentication pieces in place to connect to your ServiceNow instance.

Connect to ServiceNow

To access to your ServiceNow instance from Microsoft Purview workflows, first we'll create a connection. This allows Microsoft Purview to authenticate with ServiceNow.

Microsoft Purview uses basic auth credentials to connect to ServiceNow.

1. In the Microsoft Purview governance portal, select the **Management** menu, and then select **Connections**.

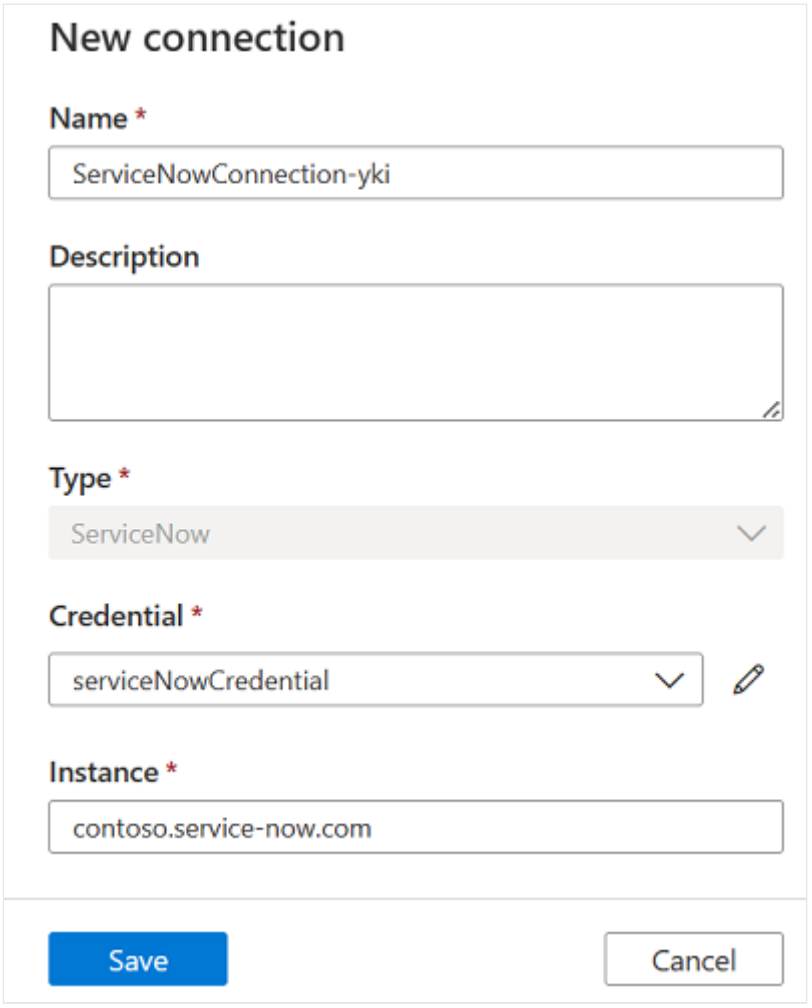


2. Select the + **New** button to create a new connection.

3. Select the **ServiceNow** data source.

4. Name your connection, select the credential you created when you set up authentication, and provide your ServiceNow instance in this format:

   `<instance>.service-now.com`

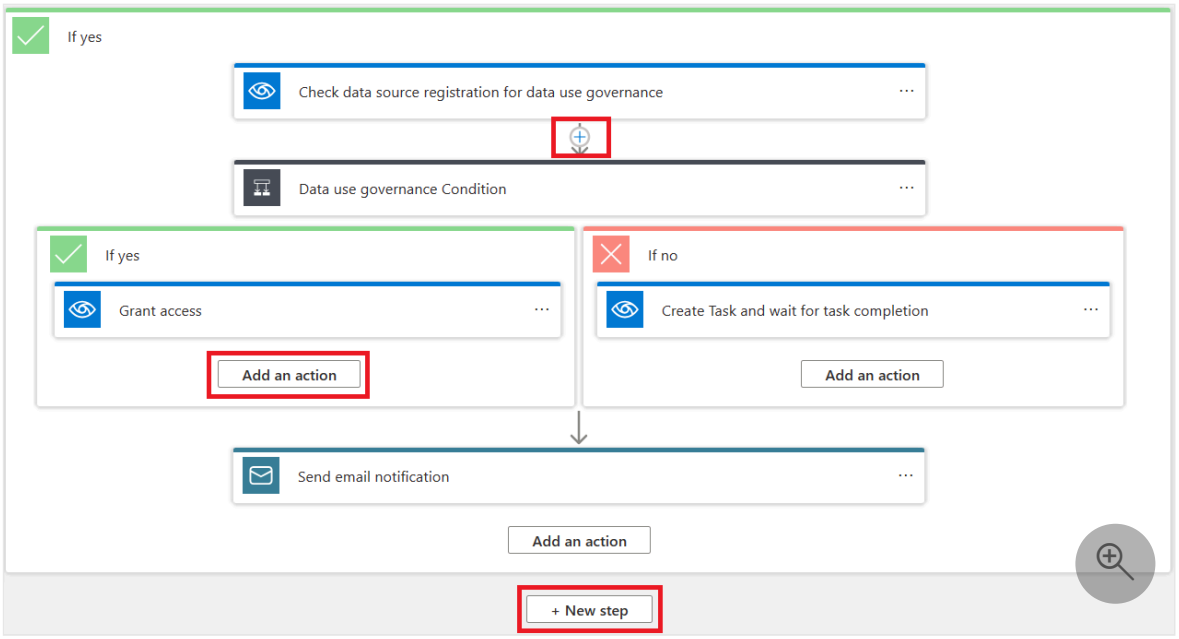


5. Select **Save**

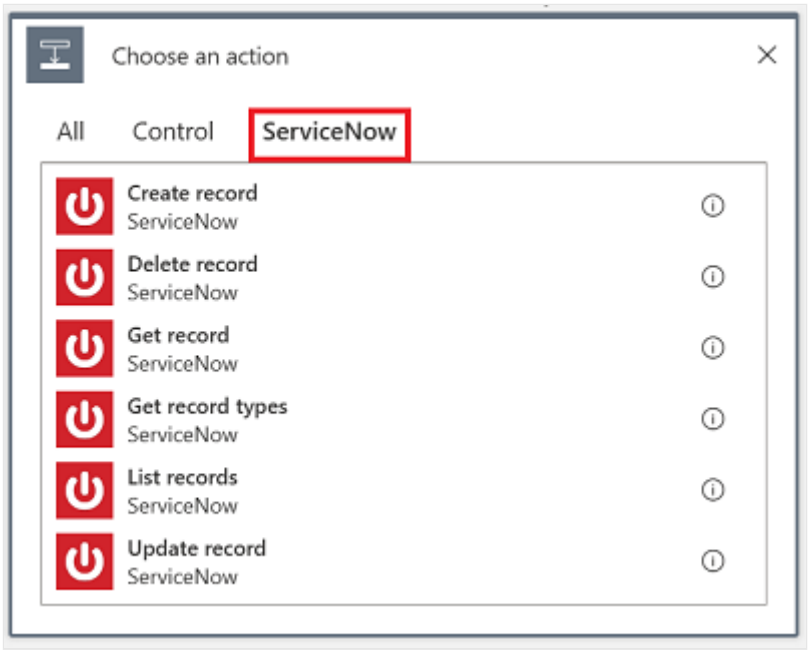
Create ServiceNow action

When you're authoring any workflow (for example an approval workflow for business terms), you can add a ServiceNow action by following these steps.

1. Add an action to your workflow by selecting the + button on the arrow between actions, select the **Add an action** button, or the + **New Step** button at the end of the flow. Where you add the action is where it will be executed in the workflow.



2. In the **Choose an action** menu, select **ServiceNow**.



3. Select the ServiceNow action you'd like to use, and fill out the parameters. For more information about each action, see the available ServiceNow actions list.

4. Select **Save** to save your additions to the workflow.

Available actions

Expand table

Connector Type	Functionality	Parameters	Customizable	Workflow templates
Create record	Create a record for a ServiceNow object type.	-Connection (your ServiceNow connection) -Record type -Record body -Display system references toggle -Exclude reference links toggle	- Renamable: Yes - Deletable: Yes - Multiple per workflow	All workflows templates
Delete record	Delete a record for a ServiceNow object type.	-Connection (your ServiceNow connection) -Record type -System ID	- Renamable: Yes - Deletable: Yes - Multiple per workflow	All workflows templates
Get record	Get a record for a ServiceNow object type.	-Connection (your ServiceNow connection) -Record type -System ID -Display system references toggle -Exclude reference links toggle	- Renamable: Yes - Deletable: Yes - Multiple per workflow	All workflows templates
Get record types	List available record types.	-Connection (your ServiceNow connection)	- Renamable: Yes - Deletable: Yes - Multiple per workflow	All workflows templates
List records	Get records of a ServiceNow object type.	-Connection (your ServiceNow connection) -Record type -Display system references toggle -Exclude reference links toggle -Query -Limit (default is 10000) -Offset	- Renamable: Yes - Deletable: Yes - Multiple per workflow	All workflows templates
Update record	Updates a record for a ServiceNow object	-Connection (your ServiceNow	- Renamable: Yes	All workflows templates

Connector Type	Functionality	Parameters	Customizable	Workflow templates
	type.	connection) -Record type -System ID -Record body -Display system references toggle -Exclude reference links toggle	- Deletable: Yes - Multiple per workflow	

Next steps

For more information about workflows, see these articles:

- Workflows in Microsoft Purview
- Workflow actions and connections
- Approval workflow for business terms
- Manage workflow requests and approvals

Feedback

Was this page helpful? Yes No

Workflow connectors and actions

Article • 02/24/2025

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

> **Important**
>
> This feature is currently in preview. The Supplemental Terms of Use for Microsoft Azure Previews include additional legal terms that apply to Azure features that are in beta, in preview, or otherwise not yet released into general availability.

You can use workflows to automate some business processes through Microsoft Purview. A Connector in a workflow provides a way to connect to different systems and apply a set of prebuilt actions and triggers.

Current workflow connectors and actions

Currently the following connectors and actions are available for a workflow in Microsoft Purview:

Expand table

Connector Type	Functionality	Parameters	Customizable	Workflow templates
Apply to each	Apply an action or set of actions to all returned values in an output.	-Output to process -Actions	- Renamable: Yes - Deletable: Yes - Multiple per workflow	All workflows templates
Check data source registration for data use governance	Validate if data source has been registered with Data policy enforcement enabled.	None	- Renamable: Yes - Deletable: Yes - Multiple per workflow	Data access request

Connector Type	Functionality	Parameters	Customizable	Workflow templates
Condition	Evaluate a value to true or false. Based on the evaluation the workflow will be redirected to different branches	- Add row - Title - Add group	- Renamable: Yes - Deletable: Yes - Multiple per workflow	All workflows templates
Create Glossary Term	Create a new glossary term	None	- Renamable: Yes - Deletable: Yes - Multiple per workflow	Create glossary term template
Create task and wait for task completion	Creates, assigns, and tracks a task to a user or Microsoft Entra group as part of a workflow. - Reminder settings - You can set reminders to periodically remind the task owner until they complete the task. - Expiry settings - You can set an expiration or deadline for the task activity. Also, you can set who needs to be notified (user/AAD group) after the expiry. - Email aggregation available	- Assigned to - Task title - Task body	- Renamable: Yes - Deletable: Yes - Multiple per workflow	All workflows templates
Delay	Delay the next action until a specified number of units has passed.	-Count: Number of units that need to pass until the delay resolves -Unit: Month, Week, Day, Hour, Minute, Second	- Renamable: Yes - Deletable: Yes - Multiple per workflow	All workflows templates
Delete glossary term	Delete an existing glossary term	None	- Renamable: Yes - Deletable: Yes - Multiple per workflow	Delete glossary term

Connector Type	Functionality	Parameters	Customizable	Workflow templates
Grant access	Create an access policy to grant access to the requested user.	None	- Renamable: Yes - Deletable: Yes - Multiple per workflow	Data access request
Http	Integrate with external applications through http or https call. For more information, see Workflows HTTP connector	- Host - Method - Path - Headers - Queries - Body - Authentication	- Renamable: Yes - Deletable: Yes - Settings: Secured Input and Secure outputs (Enabled by default) - Multiple per workflow	All workflows templates
Import glossary terms	Import one or more glossary terms	None	- Renamable: Yes - Deletable: No - Multiple per workflow	Import terms
Parse JSON	Parse an incoming JSON to extract parameters	- Content - Schema	- Renamable: Yes - Deletable: No - Multiple per workflow	All workflows templates
Send email notification	Send email notification to one or more recipients	- Subject - Message body - Recipient	- Renamable: Yes - Deletable: Yes - Settings: Secured Input and Secure outputs (Enabled by default), Single or aggregated emails - Multiple per workflow	All workflows templates
ServiceNow actions	Connect to an instance of	Depending. See ServiceNow in workflows	- Renamable: Yes	All workflows

Connector Type	Functionality	Parameters	Customizable	Workflow templates
	ServiceNow and complete one of these actions: -Create record -Delete record -Get record -Get record types -List records -Update record	for more information.	- Deletable: Yes - Multiple per workflow	templates
Start and wait for an approval	Generates approval requests and assigns the requests to individual users or Microsoft Entra groups. There are two approval types: - First to Respond – First approver's outcome (Approve/Reject) is considered final. - Everyone must approve – This implies everyone identified as an approver must approve. If one approver rejects the request, the request is rejected. - Email aggregation available	- Approval Type - Title - Assigned To	- Renamable: Yes - Deletable: Yes - Multiple per workflow	All workflows templates
Switch	Evaluates a dynamic content expression and chooses the case that matches the result, running specific actions for only that case. **Currently up to 15 cases are supported.**	- On: Dynamic content to evaluate. (string, integer) - Equals: Value to evaluate 'on' against. When matched, the workflow diverts to this case. (string, integer)	- Renamable: Yes - Deletable: Yes - Multiple per workflow	All workflows templates
Until	Repeat actions within the until control until the and/or condition	-Actions: Actions to repeat -Count: Default 60,	- Renamable: Yes - Deletable: Yes	All workflows templates

Connector Type	Functionality	Parameters	Customizable	Workflow templates
	is satisfied, or the count or time out limits have been reached.	Minimum 1, Maximum 1000 The number of cycles that an Until loop can have during a workflow run. -Time-out: Default PT1H(1 hour), Maximum PT1H(1 hour) The amount of time that the Until loop can run before exiting and is specified in ISO 8601 format. If cumulative time counted for all actions and iterations within Until action takes longer than the time-out limit, the current cycle doesn't stop. However, the next cycle doesn't start because the limit condition isn't met.	- Multiple per workflow	
Update glossary term	Update an existing glossary term	None	- Renamable: Yes - Deletable: Yes - Multiple per workflow	Update glossary term
Variable actions	Define and update custom variables with these actions: -Initialize variable -Set variable -Increment variable -Decrement variable -Append array to variable -Append string to variable	Depending. See the variable actions section for more information.	- Renamable: Yes - Deletable: Yes - Multiple per workflow	All workflows templates
When term creation request is submitted	Triggers a workflow with all term details when a new term request is submitted	None	- Renamable: Yes - Deletable: No	Create glossary term template

Connector Type	Functionality	Parameters	Customizable	Workflow templates
			- Only one per workflow	
When term deletion request is submitted	Triggers a workflow with all term details when a request to delete an existing term is submitted	None	- Renamable: Yes - Deletable: No - Only one per workflow	Delete glossary term
When term Import request is submitted	Triggers a workflow with all term details in a csv file, when a request to import terms is submitted	None	- Renamable: Yes - Deletable: No - Only one per workflow	Import terms
When term update request is submitted	Triggers a workflow with all term details when a request to update an existing term is submitted	None	- Renamable: Yes - Deletable: No - Only one per workflow	Update glossary term

Start and wait for an approval

The start and wait for an approval connector generates approval requests and assigns the requests to individual users or Microsoft Entra groups.

Currently, the connector supports two types of approval types:

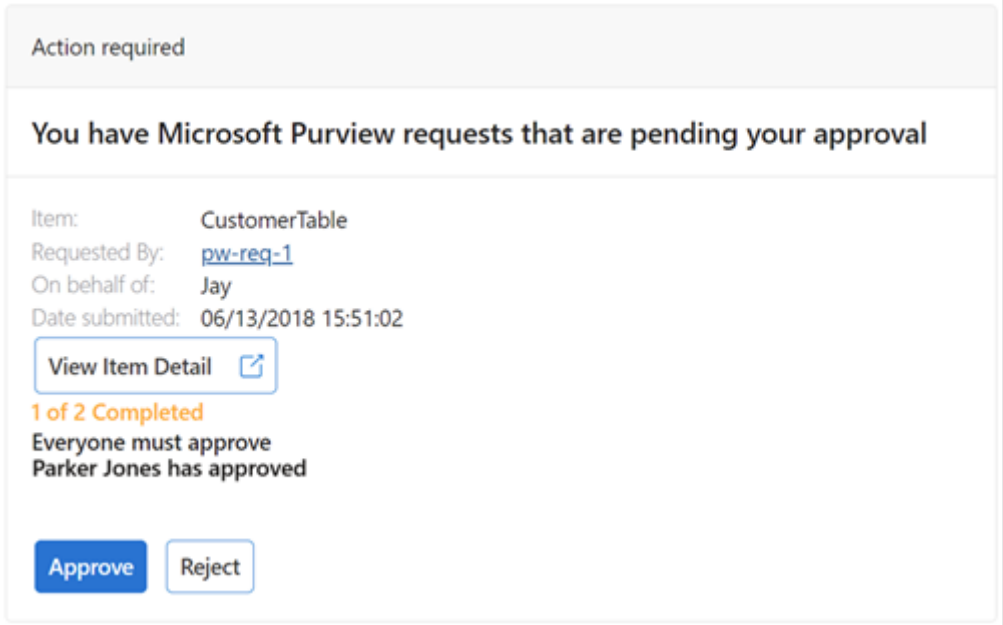
- **First to Respond** – This implies that the first approver's outcome (Approve/Reject) is considered final.
- **Everyone must approve** – This implies everyone identified as an approver must approve the request for the request to be considered approved. If one approver rejects the request, regardless of other approvers, the request is rejected.

You can also set reminders to periodically remind the approver until they approve or reject, or set an expiration or deadline for the approval activity. When the request expires, you can choose who will receive a notification about the expiration.

Approval requests

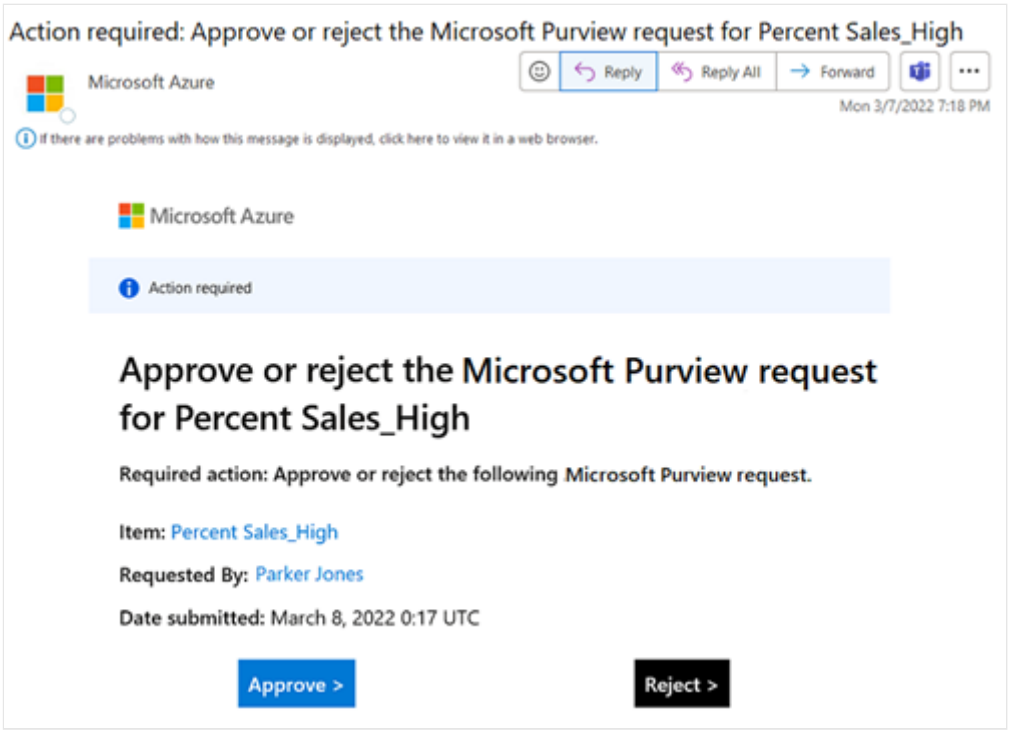
Approval requests are sent as emails to the groups or users specified in the approval connector. There are two types of possible approval emails depending on whether your workflow and your email client meet some requirements.

- **Actionable messages** - if your workflow and email client meet the requirements, your approvers can receive single and aggregated actionable approval emails where they can leave comments and reject or approve requests within outlook. For more information, the approver can also select the approval item in the message to be taken to request details in the Microsoft Purview portal.



> **ⓘ Important**
>
> If your workflow or email client don't meet the **requirements**, your users might receive 400/401/403 errors when using actionable messages.

- **Standard messages** - if your workflow and email client don't meet the requirements, your approvers receive single and aggregated approval emails that direct them to the approval experience in the Microsoft Purview portal.



Actionable approval requirements

For your approvers to use actionable messages in Outlook, your Microsoft Purview, workflow, and email client must meet these requirements:

- Your Microsoft Purview account must have public network access enabled.
- When the user submits an approval from the actionable email, they need to be logged into the same tenant where the Microsoft Purview account resides.
- The **Assigned to** configured in the Start and Wait for approval Connector must be an individual not a group.
- The **Assigned to** configured in the Start and Wait for approval Connector must have a mailbox on Outlook.com or Exchange Online in Office 365.
- The email client used to open the approval request email must be Outlook with these versions: Outlook version requirements for actionable messages.
- The Microsoft Purview account isn't in an Azure Government and Azure China 21Vianet region.

> **Tip**
>
> If users receive a 403 message when attempting to approve an item from an actionable email, your Microsoft Purview account likely has network settings that are rejecting the request.

Send email notification

When you send an email notification in a workflow, either using the send email activity or another activity, by default individual emails are sent to each user specified. However, for some activities there's also an option to aggregate emails to be sent to batches of users. This can greatly reduce the number of emails sent from activities.

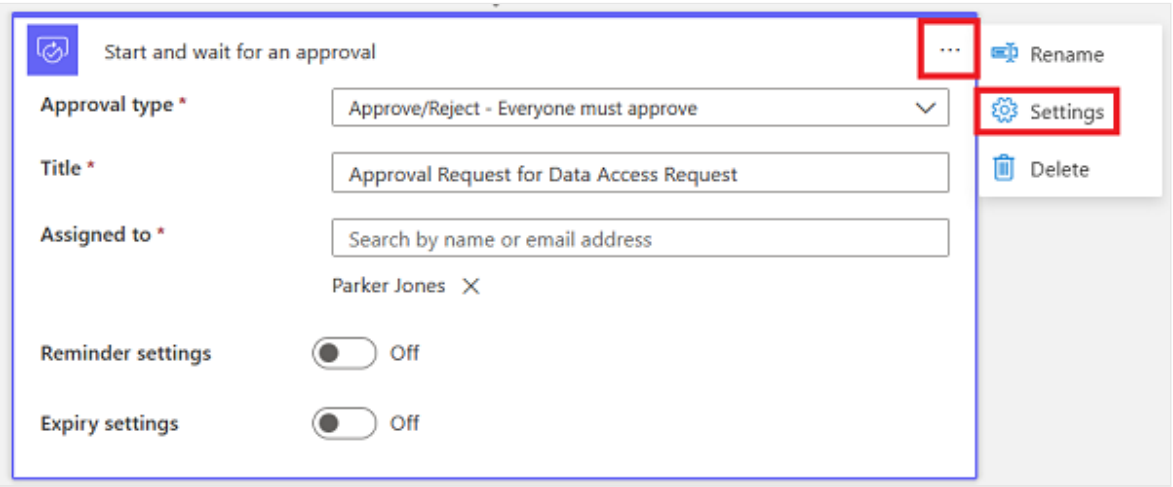
These activities are available to aggregate emails:

- Start and wait for an approval
- Send email notification
- Create task and wait for task completion

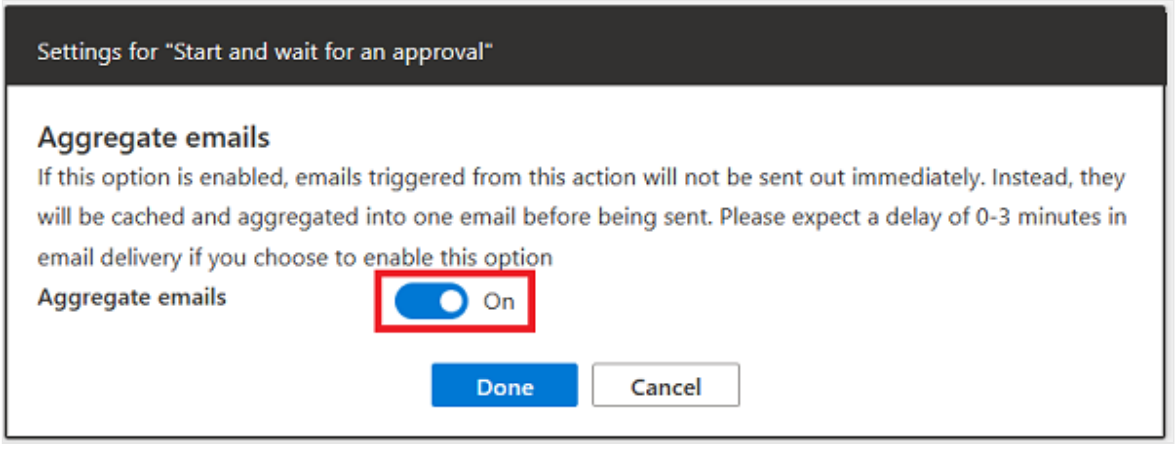
How to aggregate emails

When you're using one of the available activities and want to aggregate your emails, follow these steps.

1. Open your activity settings by selecting the ellipsis button in the activity, and then the **Settings** option.



2. Select the **Aggregate emails** toggle to switch it to **On**.



3. Select **Done**

4. Finish making any other edits needed to your workflow, and then select **Save** to save the changes to your workflow.

Aggregation information

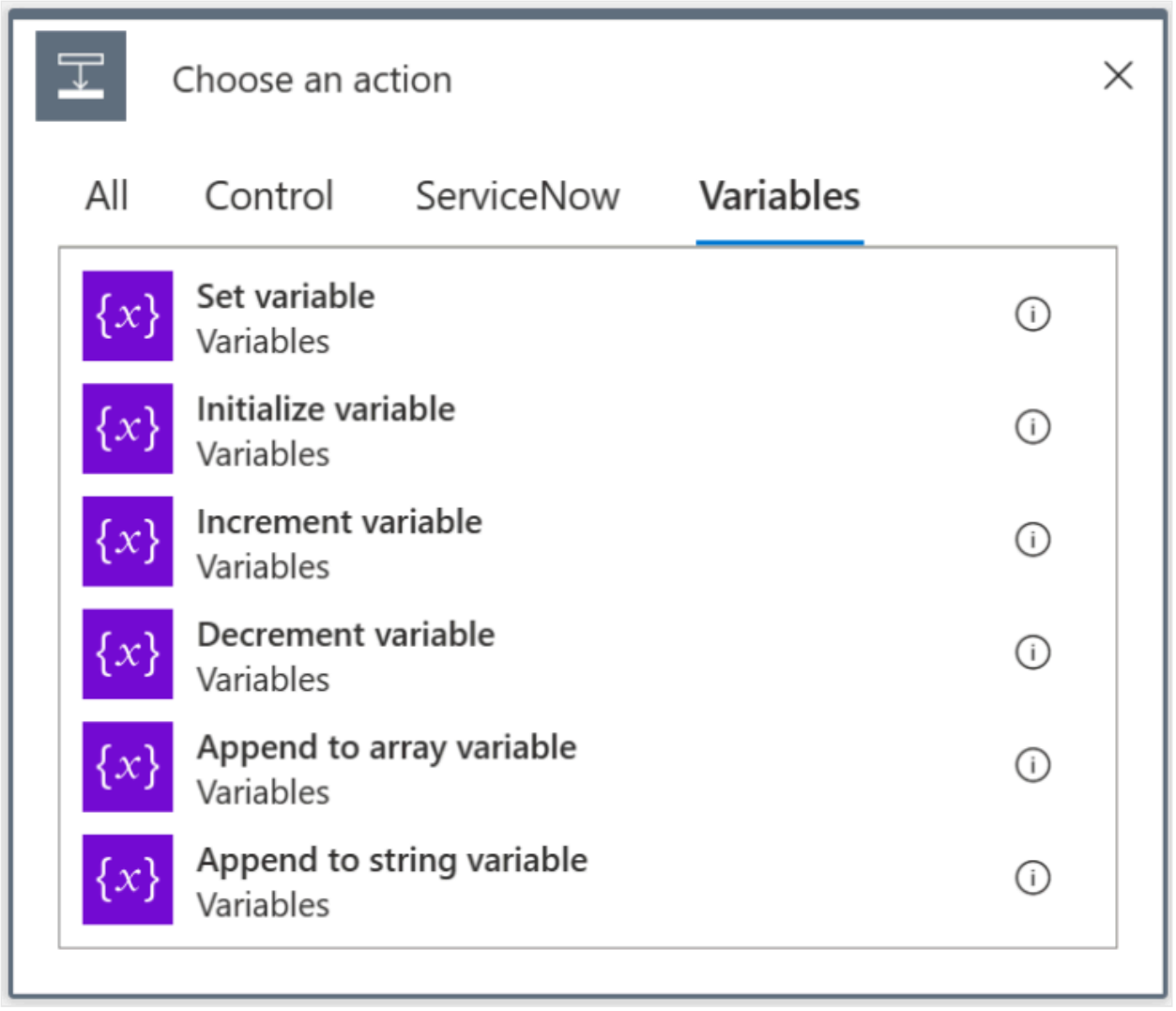
When you enable aggregation, the emails sent during this action are sent immediately. Instead the individual emails would be aggregated into one email and sent to receivers. The aggregation causes a slight delay in email delivery, between 0-3 minutes.

Aggregated emails are batched in groups of at most, 20. That is, if there are 110 emails that need to be aggregated, in the end 6 emails would be sent.

Variable actions

Create, update, and reference custom variables that you can reference in workflows in any places that support dynamic content. You can create variables of these data types:

- Integer
- Float
- Boolean
- Array
- String
- Object



Here are all the actions available and their parameters and descriptions:

Expand table

Action	Functionality	Parameters
Initialize variable	Create a new variable of one of the available types	-Name -Type -Value: Appropriate value based on variable type
Set variable	Update variable to specified value	-Name: Variable name selected from drop-down -Value: New value for variable
Increment variable	Increase integer and float variables by a specified value	-Name: Variable name selected from drop-down -Value: Amount to increment variable
Decrement variable	Reduce integer and float variables by a specified value	-Name: Variable name selected from drop-down

Action	Functionality	Parameters
		-Value: Amount to decrement variable
Append to array variable	Append specified value to array variable	-Name: Variable name selected from drop-down -Value: New value to append
Append to string variable	Append specified value to array variable	-Name: Variable name selected from drop-down -Value: New value to append

Variable limitations

- 50 variables per workflow
- Numeric variable range is $[-(2^{53}-1), 2^{53}-1]$
- names are limited to 46 characters
- Self-references aren't allowed in actions that update variable values: set, increment, decrement, append to array, append to string
- Initialize variable can't be used in local scope actions like "apply to each," "condition," or "until."
- Set variable can't be used in "apply to each" actions.

Next steps

For more information about workflows, see these articles:

- Workflows in Microsoft Purview
- Approval workflow for business terms
- Manage workflow requests and approvals

Feedback

Was this page helpful? Yes No

Workflow dynamic content

Article • 02/26/2025

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

> **Important**
>
> This feature is currently in preview. The Supplemental Terms of Use for Microsoft Azure Previews include additional legal terms that apply to Azure features that are in beta, in preview, or otherwise not yet released into general availability.

You can use dynamic content inside Microsoft Purview workflows to associate certain variables in the workflow or add other expressions to process these values.

When you add dynamic content to your workflows, you're building expressions from provided building blocks that reference and process information in your workflow so you can get the values you need in real-time.

In the dynamic content menu, the currently available options are:

- Built-in variables - variables that represent values coming to the workflow from the items that triggered it
- Expressions - formulas built from functions and variables that can process values in-workflow.

Built-in variables

Currently, the following variables are available for a workflow connector in Microsoft Purview:

Expand table

Prerequisite connector	Built-in variable	Functionality	Type	Possible values
When data access request is submitted When asset update request is submitted	Workflow.Requestor	The ID of the requestor of the workflow	*string*	
	Asset.Name	The name of the asset	*string*	
	Asset.Description	The description of the asset	*string* or *null*	
	Asset.Type	The type of the asset	*string*	
	Asset.Fully Qualified Name	The fully qualified name of the asset	*string*	
	Asset.Owner	The IDs of the data owners of the asset	*array* of *string*s	
	Asset.Classification	The display names of classifications of the asset	*array* of *string*s	
	Asset.Schema.Classification	The display names of classifications applied to a column in the asset schema	*array* of *string*s	
	Asset.Is Certified	The indicator of whether the asset meets your organization's quality standards and can be regarded as reliable	*boolean*	`true` or `false`
When data access request is submitted	Workflow.Request Recipient	The ID of the request recipient of the workflow	*string*	
Start and wait for an approval	Approval.Outcome	The outcome of the approval	*string*	`'Approved'` or `'Rejected'`
	Approval.Assigned To	The IDs of the approvers	*array* of *string*s	

Prerequisite connector	Built-in variable	Functionality	Type	Possible values
	Approval.Comments	The names of the approvers and their comments	*string*	
Check data source registration for data use governance	Data Use Governance	The result of the data use governance check	*boolean*	`true` or `false`
When term creation request is submitted	Workflow.Requestor	The ID of the requestor of the workflow	*string*	
	Term.Name	The name of the term	*string*	
	Term.Formal Name	The formal name of the term	*string*	
	Term.Definition	The definition of the term	*string*	
	Term.Experts	The IDs of the experts of the term	*array* of *string*s	
	Term.Stewards	The IDs of the stewards of the term	*array* of *string*s	
	Term.Parent.Name	The name of parent term if exists	*string*	
	Term.Parent.Formal Name	The formal name of parent term if exists	*string*	
When term update request is submitted When term deletion request is submitted	Workflow.Requestor	The requestor of the workflow	*string*	
	Term.Name	The name of the term	*string*	
	Term.Formal Name	The formal name of the term	*string*	
	Term.Definition	The definition of the term	*string* or *null*	

Prerequisite connector	Built-in variable	Functionality	Type	Possible values
	Term.Experts	The IDs of the experts of the term	*array* of *string*s	
	Term.Stewards	The IDs of the stewards of the term	*array* of *string*s	
	Term.Parent.Name	The name of parent term if exists	*string*	
	Term.Parent.Formal Name	The formal name of parent term if exists	*string*	
	Term.Created By	The ID of the creator of the term	*string*	
	Term.Last Updated By	The ID of the last updater of the term	*string*	
When term import request is submitted	Workflow.Requestor	The ID of the requestor of the workflow	*string*	
	Import File.Name	The name of the file to import	*string*	
Http	Http.Body	HTTP response body	*object* or *string*	
	Http.Headers	HTTP response headers	*object*	
	Http.Status Code	HTTP status code	*string*	
Apply to each	Item	The current item of Apply to Each action	Any	
Until	Index	The current iteration index of Until action	*integer*	
Create record Delete record Get record Update record	Result	The result of ServiceNow action	*object*	
Get record types List records	Result	The result of ServiceNow action	*array* of *object*s	

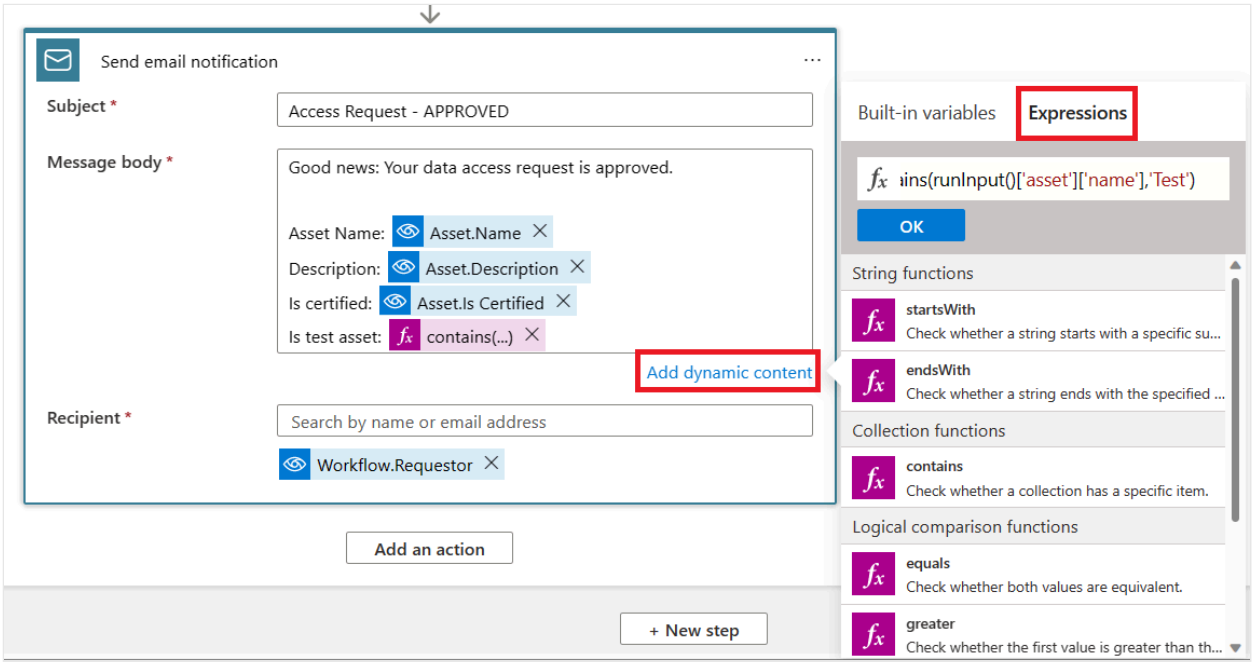
Expressions

Workflow definitions in Microsoft Purview allow you to use functions in your expressions to process values in your workflows.

To find functions based on their general purpose, review the following tables. Or, for detailed information about each function, see the alphabetical list.

When you're building a workflow and want to add a function to your expressions, follow these steps:

1. Select the value you're going to edit.
2. Select the **Add dynamic content** button that appears underneath the textbox.
3. Select the **Expressions** tab in the dynamic content window and scroll to select your value.
4. Update your expression and select **OK** to add it.



Considerations

- Function parameters are evaluated from left to right.
- Functions that appear inline with plain text require enclosing curly braces ({}) to use the expression's interpolated format instead. This format helps avoid parsing problems. If your function expression doesn't appear inline with plain text, no curly braces are necessary.

  The following example shows the correct and incorrect syntax:

  Correct: `"<text>/@{<function-name>('<parameter-name>')}/<text>"`

Incorrect: `"<text>/@<function-name>('<parameter-name>')/<text>"`

- Functions may have overloads, which are multiple versions of a function with the same name but different ways of doing a task, depending on the input you give them.

The following sections organize functions based on their general purpose, or you can browse these functions in alphabetical order.

String functions

To work with strings, you can use these string functions and also some collection functions. String functions work only on strings.

Expand table

String function	Task
endsWith	Check whether a string ends with the specified substring.
startsWith	Check whether a string starts with a specific substring.

Collection functions

To work with collections, generally arrays, strings, and sometimes, dictionaries, you can use these collection functions.

Expand table

Collection function	Task
contains	Check whether a collection has a specific item.
length	Return the number of items in a collection, or the length of a string.

Logical comparison functions

To work with conditions, compare values and expression results, or evaluate various kinds of logic, you can use these logical comparison functions. For the full reference about each function, see the alphabetical list.

Expand table

Logical comparison function	Task
not	Check whether an expression is false. Return true when the expression is false, or return false when true.
and	Check whether all expressions are true. Return true when all expressions are true, or return false when at least one expression is false.
or	Check whether at least one expression is true. Return true when at least one expression is true, or return false when all are false.
equals	Check whether both values are equivalent.
greater	Check whether the first value is greater than the second value.
greaterOrEquals	Check whether the first value is greater than or equal to the second value.
less	Check whether the first value is less than the second value.
lessOrEquals	Check whether the first value is less than or equal to the second value.

User functions

To work with users, their relationships with other users and groups, and the resources they access for example their mails, calendars, files, administrative roles, group memberships. For the full reference about each function, see the alphabetical list.

Expand table

User function	Task
getUserName	Get the name of the user.
getManager	Get the ID of the user's manager.
memberOfGroup	Check whether the user is a member of the group.

Conversion functions

To change a value's type or format, you can use these conversion functions. For the full reference about each function, see the alphabetical list.

Expand table

Conversion function	Task
createArray	Return an array from any number of inputs.
createObject	Return an empty object.
string	Return the string version for a value.
int	Try to convert the input value into an integer number.
float	Try to convert the input value into a float number..
bool	Try to convert the input value into a boolean value.
object	Try to convert the input value into an object.
array	Try to convert the input value into an array.
nullOrElse	Give a default value to a nullable value.

All functions - alphabetical list

and

Check whether all expressions are true. Return true when all expressions are true, or return false when at least one expression is false.

Expand table

Parameter	Type	Description
value	*boolean*	The first expression to check
values...	*boolean*	The rest expressions to check

Expand table

Return value type	Description
boolean	Return true when all expressions are true. Return false when at least one expression is false.

Examples

These examples check whether the specified expressions are false:

```
and(equals(1, 1), equals(2, 2))
and(equals(1, 1), equals(1, 2))
and(equals(1, 2))
```

And returns these results:

- First example: Both expressions are true, so returns `true`.
- Second example: One expression is false, so returns `false`.
- Third example: The expression is false, so returns `false`.

array

This function tries to convert the input value into an array. It is used to cast the value from a broader type to *array*. To create an array from items, see createArray.

Expand table

Parameter	Type	Description
value	Any	The value to convert

Expand table

Return value type	Description
array	The converted array value

Example

This example asserts **Item** of an "Apply to each" action to be of *array* type:

```
array(items('Appy to each'))
```

bool

This function tries to convert the input value into a boolean value. It is used to cast the value from a broader type to *boolean*.

Expand table

Parameter	Type	Description
value	Any	The value to convert

Expand table

Return value type	Description
boolean	The converted boolean value

Example

This example asserts **Item** of an "Apply to each" action to be of *boolean* type:

```
bool(items('Appy to each'))
```

contains

Check whether a collection has a specific item. Return true when the item is found, or return false when not found.

Overload 1: A *string* to find a substring

This overload is case-sensitive.

Expand table

Parameter	Type	Description
collection	*string*	The string to check
value	*string*	The substring to find

Expand table

Return value type	Description
boolean	Return true when the substring is found. Return false when not found.

Overload 2: A *array* to find a value

Expand table

Parameter	Type	Description
collection	*array*	The collection to check
value	Any	The item to find

Expand table

Return value type	Description
boolean	Return true when the item is found. Return false when not found.

Example 1

This example checks the string "hello world" for the substring "world" and returns `true`:

```
contains('hello world', 'world')
```

Example 2

This example checks the variable **Asset.Owner** for the item equal to the variable **Workflow.Requestor**:

```
contains(runInput()['asset']['dataOwner'], runInput()['requestor'])
```

createArray

Return an array from any number of inputs. This function doesn't require the input values to be of a same type.

Expand table

Parameter	Type	Description
values...	Any	Values to create the array.

Expand table

Return value type	Description
array	The array created from all the input items

Examples

These examples create an array from inputs:

```
createArray('h', 'e', 'l', 'l', 'o')
createArray(true, null)
createArray()
```

createObject

Return an empty object.

Expand table

Return value type	Description
object	An empty object

Example

This example creates an array from inputs, and returns {}.

```
createObject()
```

endsWith

Check whether a string ends with a specific substring. Return true when the substring is found, or return false when not found. This function is case-sensitive.

Expand table

Parameter	Type	Description
text	*string*	The string to check
searchText	*string*	The ending substring to find

Expand table

Return value type	Description
boolean	Return true when the ending substring is found. Return false when not found.

Example 1

This example checks whether the "hello world" string ends with the "world" string:

```
endsWith('hello world', 'world')
```

And returns this result: `true`

Example 2

This example checks whether the "hello world" string ends with the "universe" string:

```
endsWith('hello world', 'universe')
```

And returns this result: `false`

equals

Check whether both values, expressions, or objects are equivalent. Return true when both are equivalent, or return false when they're not equivalent.

Expand table

Parameter	Type	Description
value	Any	The values, expressions, or objects to compare
compareTo	Any	The comparison values, expressions, or objects

Expand table

Return value type	Description
boolean	Return true when both are equivalent. Return false when not equivalent.

Examples

These examples check whether the specified inputs are equivalent.

```
equals(true, 1)
equals('abc', 'abcd')
```

And returns these results:

- First example: Both values are equivalent, so the function returns `true`.
- Second example: Both values aren't equivalent, so the function returns `false`.

float

This function tries to convert the input value into a float number. It is used to cast the value from a broader type to *float*.

Expand table

Parameter	Type	Description
value	Any	The value to convert

Expand table

Return value type	Description
float	The converted float value

Example

This example asserts **Item** of an "Apply to each" action to be of *float* type:

```
float(items('Appy to each'))
```

getManager

Get the ID of the user's manager in Microsoft Entra ID.

Overload 1: Get for a single user

Expand table

Parameter	Type	Description
userId	*string*	The ID of the input user

Expand table

Return value type	Description
string or *null*	The ID of the input user's manager or null if not exist

Overload 2: Get for multiple users

Expand table

Parameter	Type	Description
userId	*array*	The IDs of the input users

Expand table

Return value type	Description
array	The union of IDs of the input users' managers

Examples

These examples get the manager(s) of the input user(s):

```
getManager(runInput()['requestor'])
getManager(runInput()['term']['stewards'])
```

- First example: The input parameter is a string, so returns the ID of the manager of **Workflow.Requestor**

- Second example: the input parameter is an array, so returns the union set of IDs of managers of **Term.Stewards**.

getUserName

Get the name of the user in Microsoft Entra ID.

Overload 1: Get for a single user

Expand table

Parameter	Type	Description
userId	*string*	The ID of the input user

Expand table

Return value type	Description
string	The name of the input user

Overload 2: Get for multiple users

Expand table

Parameter	Type	Description
userId	*array*	The IDs of the input users

Expand table

Return value type	Description
array	The names of the input users

Examples

These examples get the name(s) of the input user(s):

```
getUserName(runInput()['requestor'])
getUserName(runInput()['term']['stewards'])
```

- First example: The input parameter is a string, so returns the name of **Workflow.Requestor**
- Second example: The input parameter is an array, so returns an array of names of **Term.Stewards**.

greater

Check whether the first value is greater than the second value. Return true when the first value is more, or return false when less.

Overload 1: Compare numbers

Expand table

Parameter	Type	Description
value	*integer* or *float*	The first value to check whether greater than the second value
compareTo	*integer* or *float*	The comparison value

Expand table

Return value type	Description
boolean	Return true when the first value is greater than the second value. Return false when the first value is equal to or less than the second value.

Overload 2: Compare strings

Expand table

Parameter	Type	Description
value	*string*	The first value to check whether greater than the second value
compareTo	*string*	The comparison value

Expand table

Return value type	Description
boolean	Return true when the first value is greater than the second value. Return false when the first value is equal to or less than the second value.

Examples

These examples check whether the first value is greater than the second value:

```
greater(10, 5)
greater('apple', 'banana')
```

And return these results:

- First example: `true`
- Second example: `false`

greaterOrEquals

Check whether the first value is greater than or equal to the second value. Return true when the first value is greater or equal, or return false when the first value is less.

Overload 1: Compare numbers

Expand table

Parameter	Type	Description
value	*integer* or *float*	The first value to check whether greater than or equal to the second value
compareTo	*integer* or *float*	The comparison value

Expand table

Return value type	Description
boolean	Return true when the first value is greater than or equal to the second value. Return false when the first value is less than the second value.

Overload 2: Compare strings

Expand table

Parameter	Type	Description
value	*string*	The first value to check whether greater than or equal to the second value
compareTo	*string*	The comparison value

Expand table

Return value type	Description
boolean	Return true when the first value is greater than or equal to the second value. Return false when the first value is less than the second value.

Examples

These examples check whether the first value is greater or equal than the second value:

```
greaterOrEquals(5, 5)
greaterOrEquals('apple', 'banana')
```

And return these results:

- First example: `true`
- Second example: `false`

int

This function tries to convert the input value into an integer number. It is used to cast the value from a broader type to *integer*.

Expand table

Parameter	Type	Description
value	Any	The value to convert

Expand table

Return value type	Description
integer	The converted integer value

Example

This example asserts **Item** of an "Apply to each" action to be of *integer* type:

```
int(items('Appy to each'))
```

length

Return the number of items in a collection, or the length of a string.

Expand table

Parameter	Type	Description
value	*string* or *array*	The string or the collection with the items to count

Expand table

Return value type	Description
boolean	The number of items in the collection or the length of the string

Examples

These examples count the number of items in these collections:

```
length('abcd')
length(createArray(0, 1, 2, 3))
```

And both return this result: `4`

less

Check whether the first value is less than the second value. Return true when the first value is less, or return false when the first value is more.

Overload 1: Compare numbers

Expand table

Parameter	Type	Description
value	*integer* or *float*	The first value to check whether less than the second value
compareTo	*integer* or *float*	The comparison value

Expand table

Return value type	Description
boolean	Return true when the first value is less than the second value. Return false when the first value is equal to or greater than the second value.

Overload 2: Compare strings

Expand table

Parameter	Type	Description
value	*string*	The first value to check whether greater than the second value
compareTo	*string*	The comparison value

Expand table

Return value type	Description
boolean	Return true when the first value is less than the second value. Return false when the first value is equal to or greater than the second value.

Examples

These examples check whether the first value is less than the second value.

```
less(5, 10)
less('banana', 'apple')
```

And return these results:

- First example: `true`
- Second example: `false`

lessOrEquals

Check whether the first value is less than or equal to the second value. Return true when the first value is less than or equal, or return false when the first value is more.

Overload 1: Compare numbers

Expand table

Parameter	Type	Description
value	*integer* or *float*	The first value to check whether less than or equal to the second value
compareTo	*integer* or *float*	The comparison value

Expand table

Return value type	Description
boolean	Return true when the first value is less than or equal to the second value. Return false when the first value is greater than the second value.

Overload 2: Compare strings

Expand table

Parameter	Type	Description
value	*string*	The first value to check whether less than or equal to the second value
compareTo	*string*	The comparison value

Expand table

Return value type	Description
boolean	Return true when the first value is less than or equal to the second value. Return false when the first value is greater than the second value.

Examples

These examples check whether the first value is less or equal than the second value.

```
lessOrEquals(10, 10)
lessOrEquals('apply', 'apple')
```

And return these results:

- First example: `true`
- Second example: `false`

memberOfGroup

Check whether the user is a member of the group in Microsoft Entra ID. Group ID can be found on Microsoft Entra group management page.

Expand table

Parameter	Type	Description
userId	*string*	The ID of the input user
groupId	*string*	The ID of the group to check

Expand table

Return value type	Description
boolean	Return true when the user is a member of the group. Return false when the user is not a member of the group or the group doesn't exist.

Example

This example check whether **Workflow.Requestor** is a member of the specified group:

```
memberOfGroup(runInput()['requestor'], '00000000-0000-0000-0000-
111111111111')
```

not

Check whether an expression is false. Return true when the expression is false, or return false when true.

Expand table

Parameter	Type	Description
value	*boolean*	The expression to check

Expand table

Return value type	Description
boolean	Return true when the expression is false. Return false when the expression is true.

Examples

These examples check whether the specified expressions are false:

```
not(equals(1, 2))
not(equals(1, 1))
```

And return these results:

- First example: The expression is false, so the function returns `true`.
- Second example: The expression is true, so the function returns `false`.

nullOrElse

Give a default value to a nullable value. Return a non-nullable value.

Overload 1: String version

Expand table

Parameter	Type	Description
value	*string* or *null*	The nullable value
defaultValue	*string*	The fallback value

Expand table

Return value type	Description
string	Return the input value if it is not null. Return the default value if the input value is null.

Overload 2: Integer version

Expand table

Parameter	Type	Description
value	*integer* or *null*	The nullable value
defaultValue	*integer*	The fallback value

Expand table

Return value type	Description
integer	Return the input value if it is not null. Return the default value if the input value is null.

Overload 3: Float version

Expand table

Parameter	Type	Description
value	*float* or *null*	The nullable value
defaultValue	*float*	The fallback value

Expand table

Return value type	Description
float	Return the input value if it is not null. Return the default value if the input value is null.

Overload 4: Boolean version

Expand table

Parameter	Type	Description
value	*boolean* or *null*	The nullable value
defaultValue	*boolean*	The fallback value

Expand table

Return value type	Description
boolean	Return the input value if it is not null. Return the default value if the input value is null.

Overload 5: Array version

Expand table

Parameter	Type	Description
value	*array* or *null*	The nullable value
defaultValue	*array*	The fallback value

Expand table

Return value type	Description
array	Return the input value if it is not null. Return the default value if the input value is null.

Overload 6: Object version

Expand table

Parameter	Type	Description
value	*object* or *null*	The nullable value
defaultValue	*object*	The fallback value

Expand table

Return value type	Description
object	Return the input value if it is not null. Return the default value if the input value is null.

Example

This example gives the nullable variable **Term.Definition** a default value `'Empty'`:

```
nullOrElse(runInput()['term']['definition'], 'Empty')
```

object

This function tries to convert the input value into an object value. It is used to cast the value from a broader type to *object*.

Expand table

Parameter	Type	Description
value	Any	The value to convert

Expand table

Return value type	Description
boolean	The converted object value

Example

This example asserts **Item** of an "Apply to each" action to be of *object* type:

```
object(items('Appy to each'))
```

or

Check whether at least one expression is true. Return true when at least one expression is true, or return false when all are false.

Expand table

Parameter	Type	Description
value	*boolean*	The first expression to check
values...	*boolean*	The rest expressions to check

Expand table

Return value type	Description
boolean	Return true when at least one expression is true. Return false when all expressions are false.

Examples

These examples check whether at least one expression is true:

```
or(equals(1, 1), equals(1, 2))
or(equals(1, 2), equals(1, 3))
or(equals(1, 2))
```

And return these results:

- First example: At least one expression is true, so the function returns `true`.
- Second example: Both expressions are false, so the function returns `false`.
- Third example: The expression is false, so the function returns `false`.

startsWith

Check whether a string starts with a specific substring. Return true when the substring is found, or return false when not found. This function is case-sensitive.

Expand table

Parameter	Type	Description
text	*string*	The string to check
searchText	*string*	The starting substring to find

Expand table

Return value type	Description
boolean	Return true when the starting substring is found. Return false when not found.

Example 1

This example checks whether the "hello world" string starts with the "hello" substring:

```
startsWith('hello world', 'hello')
```

And returns this result: `true`

Example 2

This example checks whether the "hello world" string starts with the "greetings" substring:

```
startsWith('hello world', 'greetings')
```

And returns this result: `false`

string

This function returns the string version of the value.

Expand table

Parameter	Type	Description
value	Any	The value to convert

Expand table

Return value type	Description
string	The string version for the specified value. If the value parameter is null, this value is returned as an empty string ("") value.

Examples

These examples convert the input value into a string:

```
string('abc')
string(true)
string(createArray())
```

And return these results:

- First example: The input value is a string, so returns an identity value `'abc'`.
- Second example: String version of boolean values are capitalized, so the function returns `'True'`.
- Third example: Arrays or objects are converted to JSON form, so the function returns `'[]'`.

Next steps

For more information about workflows, see these articles:

- Workflows in Microsoft Purview
- Approval workflow for business terms
- Manage workflow requests and approvals

Feedback

Was this page helpful? Yes No

Use external APIs in Workflows

Article • 02/26/2025

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

> **Important**
>
> This feature is currently in preview. The Supplemental Terms of Use for Microsoft Azure Previews include additional legal terms that apply to Azure features that are in beta, in preview, or otherwise not yet released into general availability.

You can use workflows to automate some business processes through Microsoft Purview, and the HTTP connector and parse JSON action allow you to integrate your workflows with external applications.

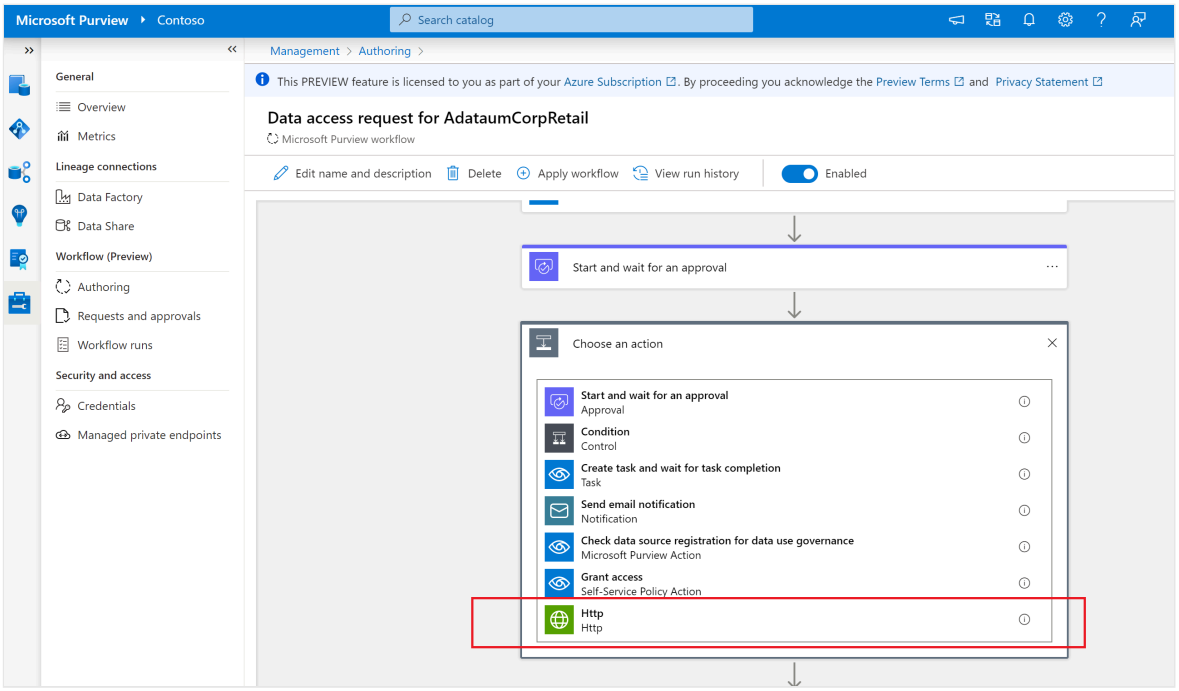
HTTP connector

HTTP connectors use Representational State Transfer (REST) architecture, which allows Microsoft Purview workflows to interact directly with non-Microsoft applications using web requests.

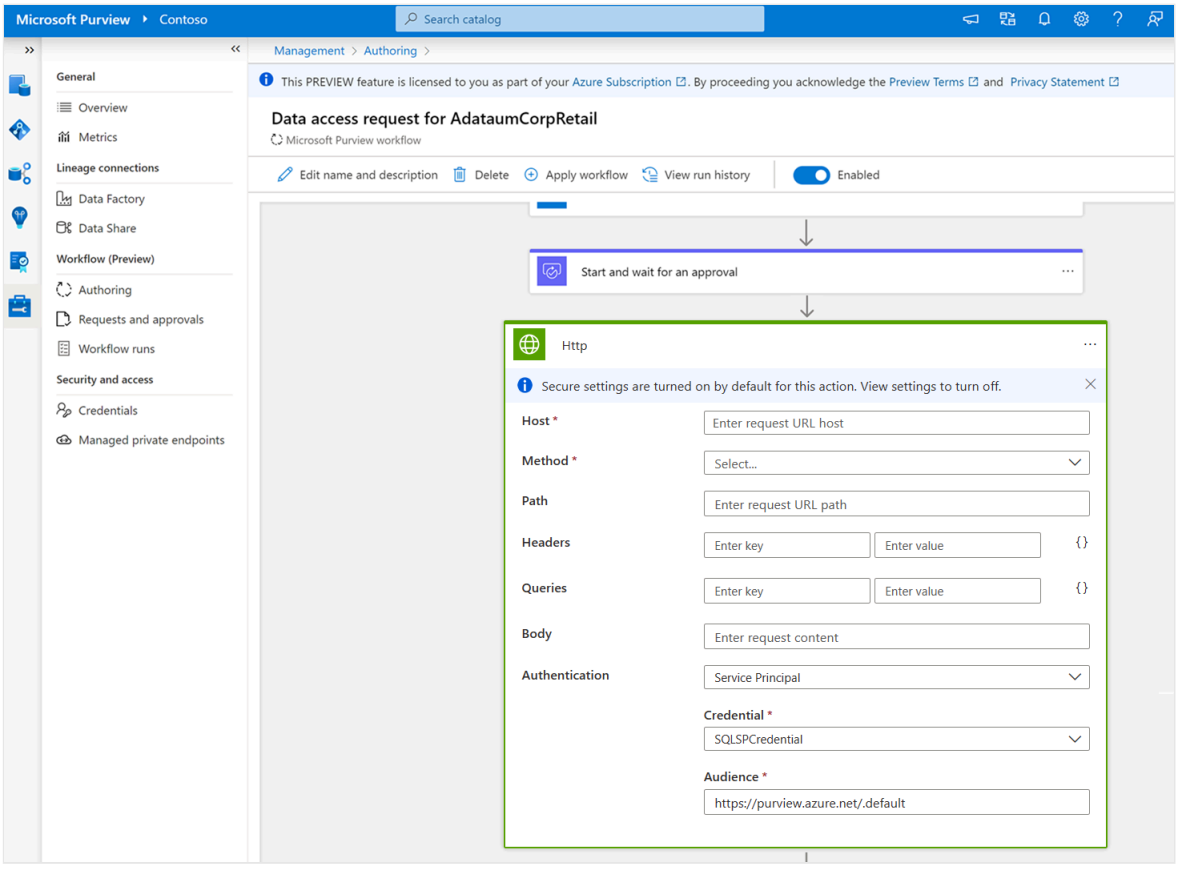
HTTP connector is available in all workflow templates.

> **Note**
>
> To create or edit a workflow, you need the workflow admin role in Microsoft Purview. You can also contact the workflow admin in your collection, or reach out to your collection administrator, for permissions.

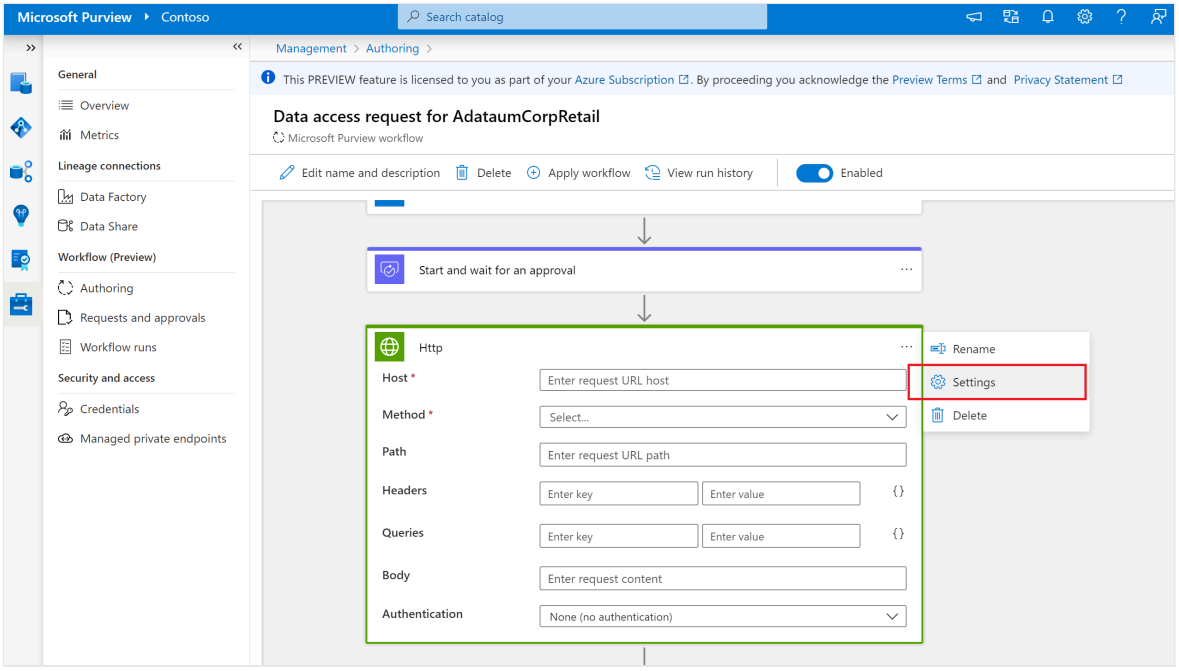
1. To add an HTTP connector, select the + icon in the template where you want to add and select HTTP connector.



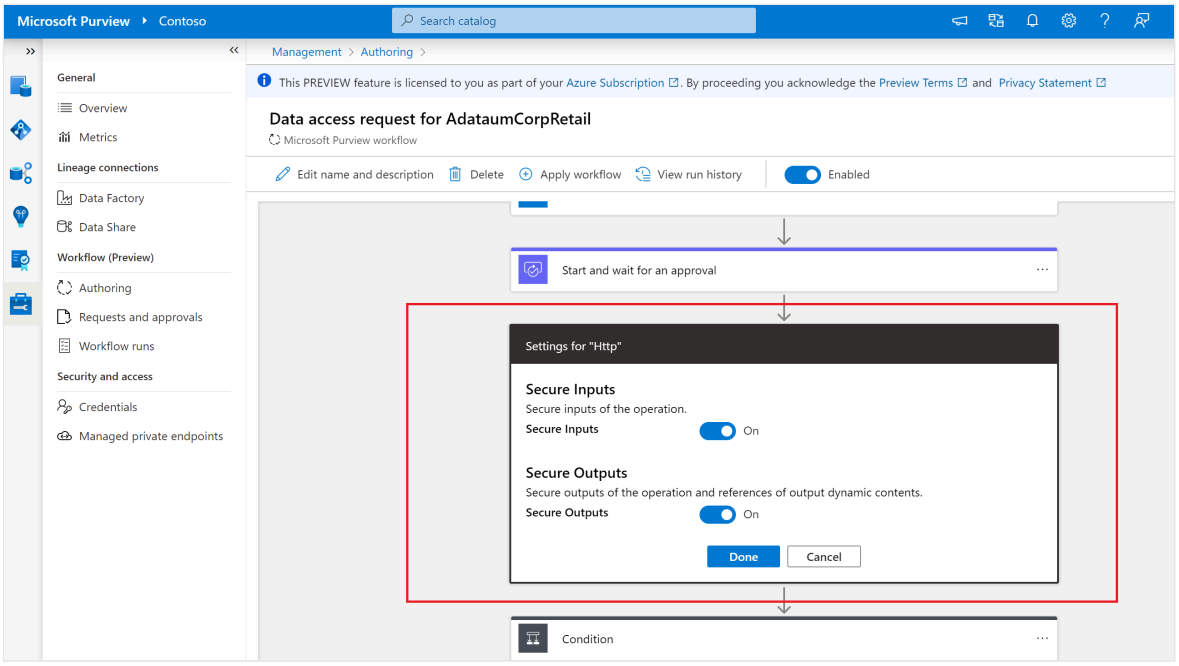
2. Once you select HTTP connector, fill the following parameters:
    a. **Host** - Request URL you want to call when this connector is executed.
    b. **Method** - Select one of the following methods. GET, PUT, PATCH, POST and DELETE. These methods correspond to create, read, update and delete operations.
    c. **Path** - Optionally you can enter request URL Path. You can use dynamic content for this parameter.
    d. **Headers** - Optionally, you can enter HTTP headers. HTTP headers let the client and the server pass additional information with an HTTP request or response
    e. **Queries** - Optionally, you can pass queries.
    f. **Body** - Optionally, you can pass HTTP body while invoking the URL
    g. **Authentication** - HTTP connector is integrated with your Microsoft Purview credentials. Depending on the URL, you can invoke one of these three types:
        i. **None** (no authentication)
        ii. **Basic authentication** - using a basic authentication credential you've already created.
        iii. **Service principal** - using a service principal credential you've already created, and an audience. The audience will be the intended consumer of the access token. The Microsoft Purview audience for example: `https://purview.azure.net/.default` Make sure your service principal is authorized to call corresponding services, like ARM.



3. By default, secure settings are turned on for HTTP connectors. To turn OFF secure inputs and outputs select the ellipsis icon (...) to go to settings.



4. You're now presented with the settings for HTTP connector and you can turn secure inputs and outputs OFF.



Parse JSON action

The parse JSON action in workflows allows you to take an incoming JSON from HTTP (or any other action/connector), and parse the JSON to extract values for use in your workflow.

The parse JSON action is available in all workflows.

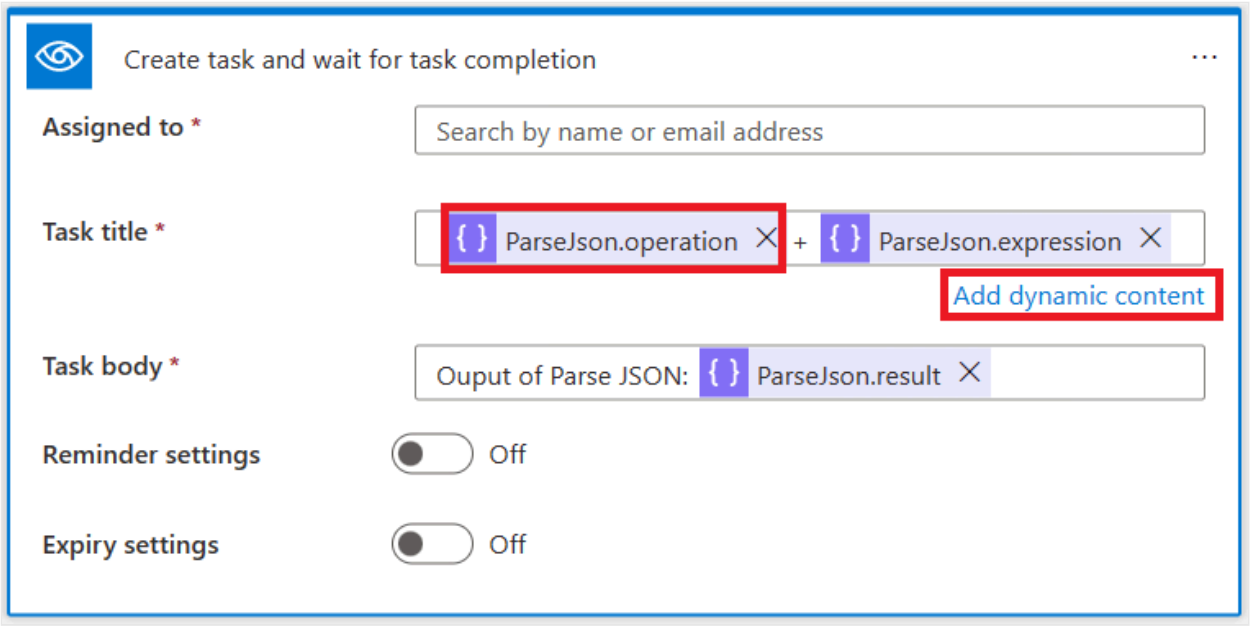


The parse JSON action has two parameters:

- Content - a variable that should contain the JSON you want to parse.
- Schema - the schema of the incoming JSON, which allows the workflow to parse the incoming information. You can supply your own, or use the **Generate from**

sample button. If you generate from a sample, you'll enter a sample JSON payload and a schema will be automatically generated for you.

Actions and connectors in the workflow after the parse JSON action will be able to use the values extracted from the JSON by selecting **Add dynamic content** for any parameters.



Next steps

For more information about workflows, see these articles:

- Workflows in Microsoft Purview
- Approval workflow for business terms
- Manage workflow requests and approvals

Feedback

Was this page helpful? Yes No

Concepts for Microsoft Purview Data owner policies (preview)

Article • 02/24/2025

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

> **Important**
>
> This feature is currently in preview. The Supplemental Terms of Use for Microsoft Azure Previews include additional legal terms that apply to Azure features that are in beta, in preview, or otherwise not yet released into general availability.

This article discusses concepts related to managing read or modify access to assets in your data estate from within the Microsoft Purview governance portal.

> **Note**
>
> This capability does not offer access control for Microsoft Purview itself. Granting access to Microsoft Purview internal roles is described in Access control in Microsoft Purview. This capability is used to grant data plane access, i.e., access to the data itself in data systems like Azure Storage. It does not allow you to grant control plane access. Control plane access provides visibility and ability manage resources in your subscription. You can manage control plane access via Identity and Access Management (IAM)

Overview

Access policies in Microsoft Purview enable you to manage access to different data systems across your entire data estate. For example:

A user needs read access to an Azure Storage account that has been registered in Microsoft Purview. You can grant this access directly in Microsoft Purview by creating a

data access policy through the **Data policy** app in the Microsoft Purview governance portal.

Data owner policies can be enforced only on data systems that have been enabled for policy enforcement in Microsoft Purview i.e., with the **Data policy enforcement** option turned on in data source registration.

Concepts

Data owner policy

A **Data owner policy** is a named set of policy statements. When a policy is published to one or more data systems under Microsoft Purview's governance, it's then enforced by them. A policy definition includes a policy name, description, and a list of one or more policy statements.

> ⓘ **Note**
>
> Currently only a single policy statement is supported per policy.

Policy statement

A **policy statement** is a human readable instruction that dictates how the data source should handle a specific data access operation. The policy statement comprises **Effect**, **Action, Data Resource** and **Subject**.

Action

An **action** is the operation being permitted or denied as part of this policy. For example: Read or Modify. These high-level logical actions map to one (or more) data actions in the data system where they are enforced.

Effect

The **effect** indicates what should be outcome if there is a match on the Data resource and Subject of the policy statement. Currently, the only supported value is **Allow**.

Data resource

The **data resource** is the fully qualified data asset path to the object the policy statement applies. It conforms to the following format:

/subscription/<subscription-id>/resourcegroups/<resource-group-name>/providers/<provider-name>/<data-asset-path>

Azure Storage data-asset-path format:

Microsoft.Storage/storageAccounts/<account-name>/blobservice/default/containers/<container-name>

Azure SQL DB data-asset-path format:

Microsoft.Sql/servers/<server-name>

Subject

It is a list of the end-user identities from Microsoft Entra ID for whom this policy statement is applicable. Each identity can be a service principal, an individual user, a group, or a managed service identity (MSI).

Example

Allow Read on Data Asset:
/subscription/finance/resourcegroups/prod/providers/Microsoft.Storage/storageAccounts/finDataLake/blobservice/default/containers/FinData to group Finance-analyst

In the above policy statement, the effect is *Allow*, the action is *Read*, the data resource is Azure Storage container *FinData*, and the subject is Microsoft Entra group *Finance-analyst*. If any user that belongs to this group attempts to read data from the storage container *FinData*, the request will be allowed.

Hierarchical enforcement of policies

The data resource specified in a policy statement is hierarchical by default. This means that the policy statement applies to the data object itself and to **all** the children objects contained by the data object. For example, a policy statement on Azure Storage container applies to all the blobs contained within it.

Policy combining algorithm

A data source combines all the applicable local policies with all policies from Microsoft Purview and provides a consolidated decision when a user tries to access an asset. The combining strategy picks the most restrictive policy.

For example, let's assume two different policies on an Azure Storage container *FinData* as follows,

Policy 1 - *Allow Read on Data Asset /subscription/..../containers/FinData to group Finance-analyst*

Policy 2 - *Deny Read on Data Asset /subscription/..../containers/FinData to group Finance-contractors*

Then let's assume that user 'user1', who is part of two groups: *Finance-analyst* and *Finance-contractors*, executes a call to the blob read API. Since both policies will be applicable, Azure Storage will choose the most restrictive one, which is *Deny* of *Read*. Thus, the access request will be denied.

Policy publishing

A newly created policy exists in the draft mode state, only visible in Microsoft Purview. The act of publishing initiates enforcement of a policy in the specified data systems. It's an asynchronous action that can take between 5 minutes and 2 hours to be effective, depending on the data source type. For more information, consult the Data owner policies how-to guides related to each data source type.

A policy published to a data source could contain policy statements that refer to a different data source. Such references will be ignored since the asset in question does not exist in the data source where the policy is applied.

Next steps

Check the guides on how to create policies in Microsoft Purview that get enforced in specific data systems. Beyond the UI, you can now try the Data owner policies API.

- Doc: How to enable policy enforcement for a data source
- Doc: Provision access to Azure Storage datasets
- Doc: Provision access to all data sources in a subscription or a resource group
- Doc: Provision access to Azure SQL Database assets
- Doc: Provision access to SQL Server 2022 (Arc-enabled) assets
- Blog: Grant users access to data assets in your enterprise via API

Feedback

Was this page helpful? Yes No

Provision access to system metadata in Azure SQL Database using Microsoft Purview DevOps policies

Article • 10/28/2024

> **Note**
>
> Microsoft Purview Data Map (classic), Data Catalog (classic), Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

DevOps policies are a type of Microsoft Purview access policies. They allow you to manage access to system metadata on data sources that have been registered for *Data policy enforcement* in Microsoft Purview. These policies are configured directly from the Microsoft Purview governance portal, and after they're saved, they get automatically published and then enforced by the data source. Microsoft Purview policies only manage access for Microsoft Entra principals.

This how-to guide covers how to configure Azure SQL Database to enforce policies created in Microsoft Purview. It covers the configuration steps for Azure SQL Database and the ones in Microsoft Purview to provision access to Azure SQL Database system metadata (DMVs and DMFs) using the DevOps policies actions *SQL Performance Monitoring* or *SQL Security Auditing*.

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- A new or existing Microsoft Purview account. Follow this quickstart guide to create one.
- Create a new Azure SQL Database instance, or use an existing one, in one of the currently available regions for this feature. You can follow this guide to create an Azure SQL Database instance.

Region support

All Microsoft Purview regions are supported.

Enforcement of Microsoft Purview policies is available only in the following regions for Azure SQL Database:

Public cloud:

- East US
- East US2
- South Central US
- West Central US
- West US3
- Canada Central
- Brazil South
- West Europe
- North Europe
- France Central
- UK South
- South Africa North
- Central India
- South East Asia
- East Asia
- Australia East

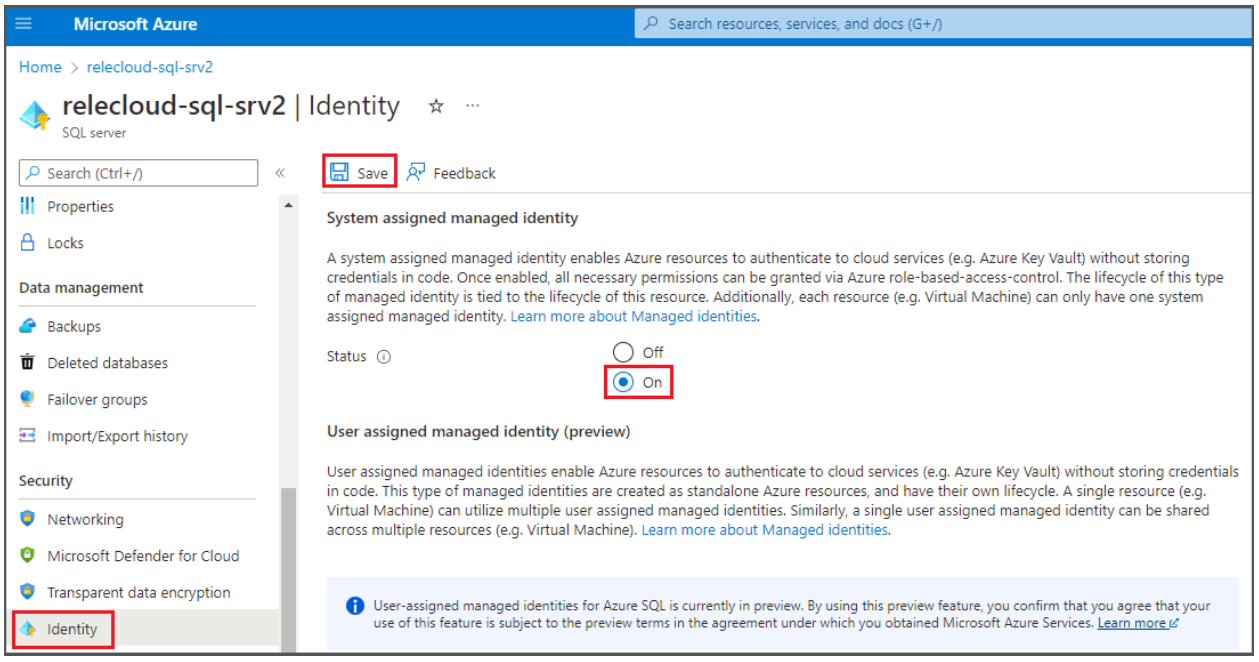
Sovereign clouds:

- USGov Virginia
- China North 3

Configure the Azure SQL Database instance for policies from Microsoft Purview

For the logical server associated with Azure SQL Database to honor policies from Microsoft Purview, you need to configure a Microsoft Entra admin. In the Azure portal, go to the logical server that hosts the Azure SQL Database instance. On the side menu, select **Microsoft Entra ID**. Set an admin name to any Microsoft Entra user or group that you prefer, and then select **Save**.

Then, on the side menu, select **Identity**. Under **System assigned managed identity**, turn the status to **On**, and then select **Save**.



Microsoft Purview Configuration

Register the data source in Microsoft Purview

Before a policy can be created in Microsoft Purview for a data resource, you must register that data resource in Microsoft Purview Studio. You will find the instructions related to registering the data resource later in this guide.

> **Note**
>
> Microsoft Purview policies rely on the data resource ARM path. If a data resource is moved to a new resource group or subscription it will need to be de-registered and then registered again in Microsoft Purview.

Configure permissions to enable Data policy enforcement on the data source

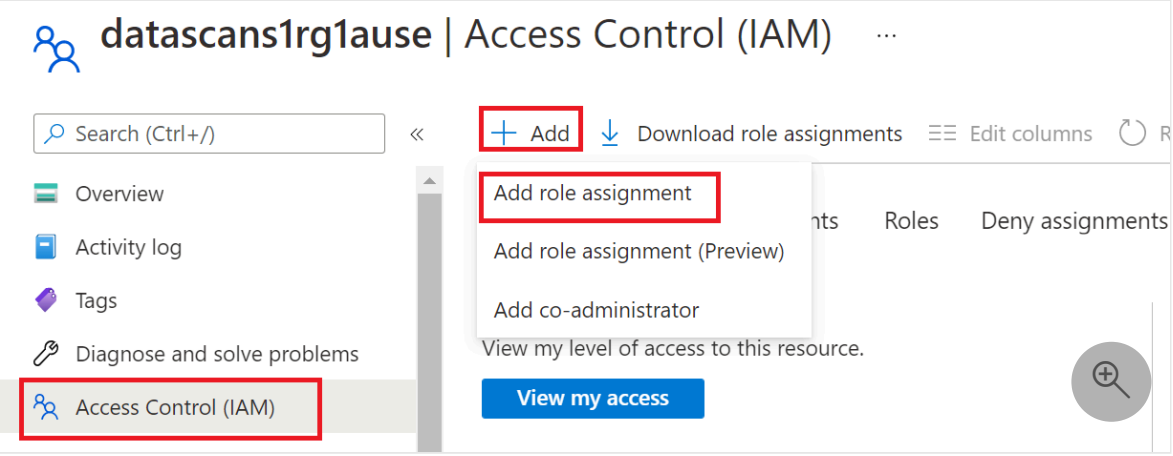
Once a resource is registered, but before a policy can be created in Microsoft Purview for that resource, you must configure permissions. A set of permissions are needed to enable the **Data policy enforcement**. This applies to data sources, resource groups, or subscriptions. To enable *Data policy enforcement*, you must have *both* specific Identity and Access Management (IAM) privileges on the resource as well as specific Microsoft Purview privileges:

- You must have *either one* of the following IAM role combinations on the resource's Azure Resource Manager path or any parent of it (that is, using IAM permission

inheritance):

- IAM Owner
- Both IAM Contributor and IAM User Access Administrator

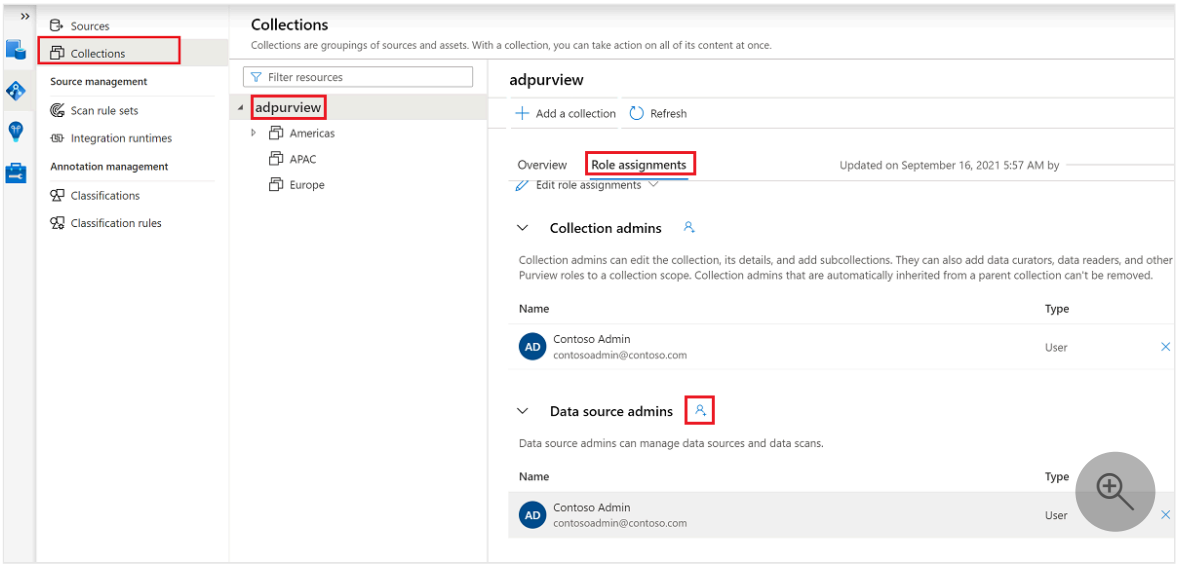
To configure Azure role-based access control (RBAC) permissions, follow this guide. The following screenshot shows how to access the **Access Control** section in the Azure portal for the data resource to add a role assignment.



> **Note**
>
> The *IAM Owner* role for a data resource can be inherited from a parent resource group, a subscription, or a subscription management group. Check which Microsoft Entra users, groups, and service principals hold or are inheriting the *IAM Owner* role for the resource.

- You also need to have the Microsoft Purview *Data source admin* role for the collection or a parent collection (if inheritance is enabled). For more information, see the guide on managing Microsoft Purview role assignments.

  The following screenshot shows how to assign the *Data source admin* role at the root collection level.



Configure Microsoft Purview permissions to create, update, or delete access policies

To create, update or delete policies, you need to get the *Policy author* role in Microsoft Purview at *root collection level*:

- The *Policy author* role can create, update, and delete DevOps and Data Owner policies.
- The *Policy author* role can delete self-service access policies.

For more information about managing Microsoft Purview role assignments, see Create and manage collections in the Microsoft Purview Data Map.

> **Note**
>
> Policy author role must be configured at the root collection level.

In addition, to easily search Microsoft Entra users or groups when creating or updating the subject of a policy, you can greatly benefit from getting the Directory Readers permission in Microsoft Entra ID. This is a common permission for users in an Azure tenant. Without the Directory Reader permission, the Policy Author will have to type the complete username or email for all the principals included in the subject of a data policy.

Configure Microsoft Purview permissions for publishing Data Owner policies

Data Owner policies allow for checks and balances if you assign the Microsoft Purview *Policy author* and *Data source admin* roles to different people in the organization. Before

a Data owner policy takes effect, a second person (*Data source admin*) must review it and explicitly approve it by publishing it. This does not apply to DevOps or Self-service access policies as publishing is automatic for them when those policies are created or updated.

To publish a Data owner policy you need to get the *Data source admin* role in Microsoft Purview at *root collection level*.

For more information about managing Microsoft Purview role assignments, see Create and manage collections in the Microsoft Purview Data Map.

> ⓘ **Note**
>
> To publish Data owner policies, the Data source admin role must be configured at the root collection level.

Delegate access provisioning responsibility to roles in Microsoft Purview

After a resource has been enabled for **Data policy enforcement**, any Microsoft Purview user with the *Policy author* role at the root collection level can provision access to that data source from Microsoft Purview.

> ⓘ **Note**
>
> Any Microsoft Purview root *Collection admin* can assign new users to root *Policy author* roles. Any *Collection admin* can assign new users to a *Data source admin* role under the collection. Minimize and carefully vet the users who hold Microsoft Purview *Collection admin*, *Data source admin*, or *Policy author* roles.

If a Microsoft Purview account with published policies is deleted, such policies will stop being enforced within an amount of time that depends on the specific data source. This change can have implications on both security and data access availability. The Contributor and Owner roles in IAM can delete Microsoft Purview accounts. You can check these permissions by going to the **Access control (IAM)** section for your Microsoft Purview account and selecting **Role Assignments**. You can also use a lock to prevent the Microsoft Purview account from being deleted through Resource Manager locks.

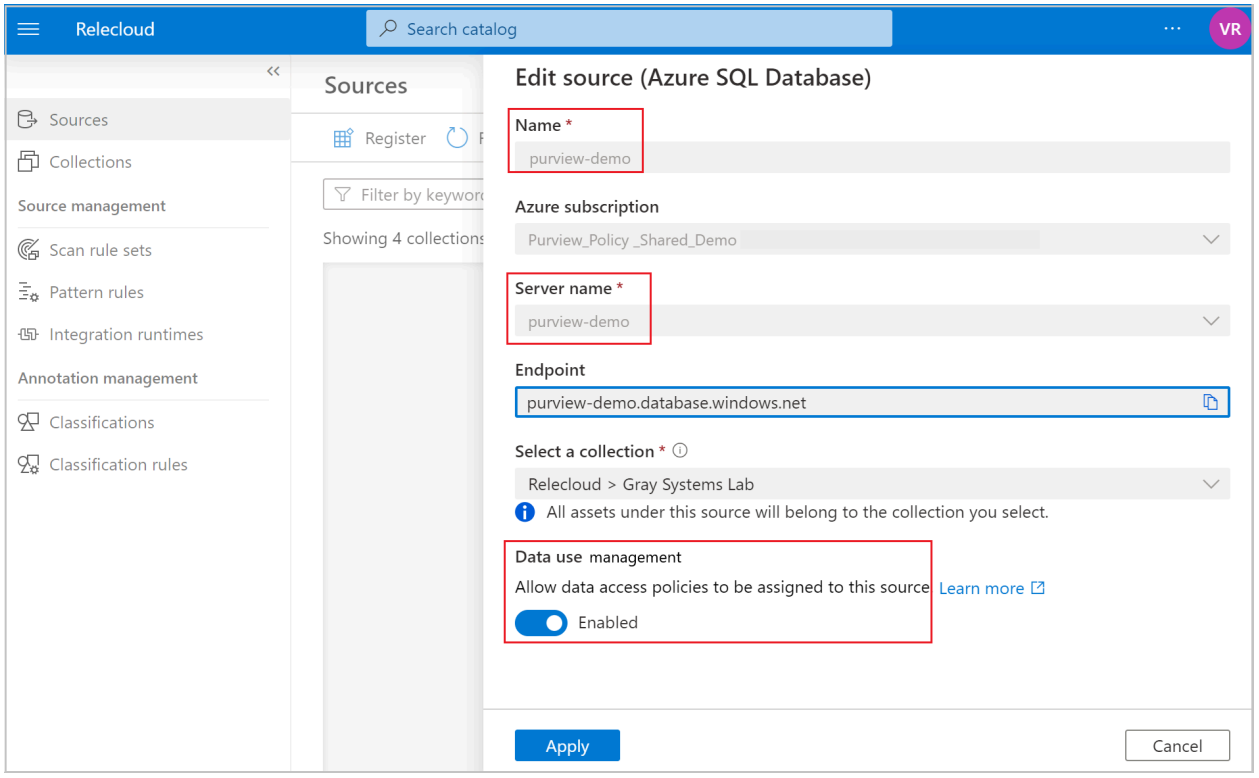
Register the data sources in Microsoft Purview

The Azure SQL Database data source needs to be registered first with Microsoft Purview, before access policies can be created. You can follow these guides:

Register and scan Azure SQL Database.

After you've registered your resources, you'll need to enable Data policy enforcement. Data policy enforcement needs certain permissions and can affect the security of your data, as it delegates to certain Microsoft Purview roles the ability to manage access to data sources. **Go through the secure practices related to Data policy enforcement in this guide**: How to enable Data policy enforcement

Once your data source has the **Data policy enforcement** toggle *Enabled*, it looks like this screenshot. This enables the access policies to be used with the given data source.



Return to the Azure portal for Azure SQL Database to verify it's now governed by Microsoft Purview:

1. Sign in to the Azure portal through this link
2. Select the Azure SQL Server that you want to configure.
3. Go to **Microsoft Entra ID** on the left pane.
4. Scroll down to **Microsoft Purview access policies**.
5. Select the button to **Check for Microsoft Purview Governance**. Wait while the request is processed. It could take a few minutes.



6. Confirm that the Microsoft Purview Governance Status shows `Governed`. Note that **it could take a few minutes** after you enable *Data policy enforcement* in Microsoft Purview for the correct status to be reflected.

> **Note**
>
> If you disable Data policy enforcement for this Azure SQL Database data source, it might take up to 24 hours for the Microsoft Purview Governance Status to be updated automatically to `Not Governed`. This can be accelerated by selecting **Check for Microsoft Purview Governance**. Before you enable *Data policy enforcement* for the data source in another Microsoft Purview account, ensure that the Purview Governance Status shows as `Not Governed`. Then repeat the steps above with the new Microsoft Purview account.

Create a new DevOps policy

Follow this link for the steps to create a new DevOps policy in Microsoft Purview.

List DevOps policies

Follow this link for the steps to list DevOps policies in Microsoft Purview.

Update a DevOps policy

Follow this link for the steps to update a DevOps policies in Microsoft Purview.

Delete a DevOps policy

Follow this link for the steps to delete a DevOps policies in Microsoft Purview.

> **Important**

DevOps policies are auto-published and changes can take up to **5 minutes** to be enforced by the data source.

Test the DevOps policy

See how to test the policy you created.

Role definition detail

See the mapping of DevOps role to data source actions.

Next steps

See related videos, blogs, and documents.

Feedback

Was this page helpful? Yes No

What can I accomplish with Microsoft Purview DevOps policies?

Article • 02/24/2025

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

This article describes how to manage access to data sources in your data estate by using the Microsoft Purview governance portal. It focuses on basic concepts of DevOps policies. That is, it provides background information about DevOps policies that you should know before you follow other articles to get configuration steps.

> **Note**
>
> This capability is different from the internal access control in the Microsoft Purview governance portal.

Access to system metadata is crucial for IT and DevOps personnel to ensure that critical database systems are healthy, are performing to expectations, and are secure. You can grant and revoke that access efficiently and at scale through Microsoft Purview DevOps policies.

Any user who holds the Policy Author role at the root collection level in Microsoft Purview can create, update, and delete DevOps policies. After DevOps policies are saved, they're published automatically.

Access policies vs. DevOps policies

Microsoft Purview access policies enable customers to manage access to data systems across their entire data estate, all from a central location in the cloud. You can think of these policies as access grants that can be created through Microsoft Purview Studio, avoiding the need for code. They dictate whether a list of Microsoft Entra principals, such as users and groups, should be allowed or denied a specific type of access to a data source or an asset within it. Microsoft Purview communicates these policies to the data sources, where they're natively enforced.

DevOps policies are a special type of Microsoft Purview access policies. They grant access to database system metadata instead of user data. They simplify access provisioning for IT operations and security auditing personnel. DevOps policies only grant access. They don't deny access.

Elements of a DevOps policy

Three elements define a DevOps policy:

- **Subject**

  This is a list of Microsoft Entra users, groups, or service principals that are granted access.

- **Data resource**

  This is the scope where the policy is enforced. The data resource path is the composition of subscription > resource group > data source.

  Microsoft Purview DevOps policies currently support SQL-type data sources. You can configure them on individual data sources and on entire resource groups and subscriptions. You can create DevOps policies only after you register the data resource in Microsoft Purview with the **Data policy enforcement** option turned on.

- **Role**

  A role maps to a set of actions that the policy permits on the data resource. DevOps policies support the SQL Performance Monitor and SQL Security Auditor roles. Both of these roles provide access to SQL system metadata, and more specifically to dynamic management views (DMVs) and dynamic management functions (DMFs). But the set of DMVs and DMFs that these roles grant is different. We provide some popular examples later in this article.

  The Create, list, update, and delete Microsoft Purview DevOps policies article details the role definition for each data source type. That is, it provides a mapping of roles in Microsoft Purview to the actions that are permitted in that type of data source. For example, the role definition for SQL Performance Monitor and SQL Security Auditor includes *Connect* actions at the server and database level on the data source side.

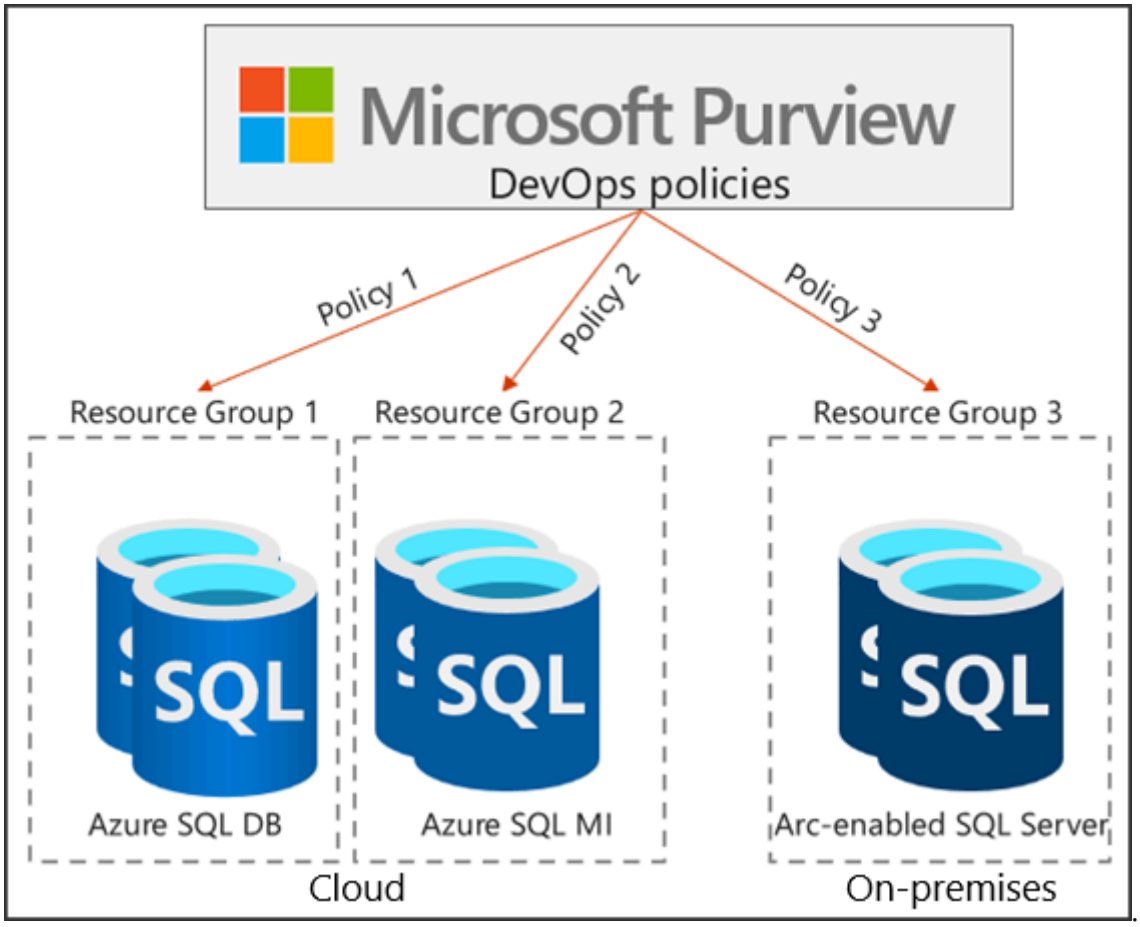
In essence, the DevOps policy assigns the role's related permissions to the subject and is enforced in the scope of the data resource's path.

Hierarchical enforcement of policies

A DevOps policy on a data resource is enforced on the data resource itself and all child resources that it contains. For example, a DevOps policy on an Azure subscription applies to all resource groups, to all policy-enabled data sources within each resource group, and to all databases within each data source.

Example scenario to demonstrate the concept and the benefits

Bob and Alice are involved with the DevOps process at their company. They need to log in to dozens of SQL Server instances on-premises and Azure SQL logical servers to monitor their performance so that critical DevOps processes don't break. Their manager, Mateo, puts all these SQL data sources into Resource Group 1. He then creates a Microsoft Entra group and includes Alice and Bob. Next, he uses Microsoft Purview DevOps policies (Policy 1 in the following diagram) to grant this Microsoft Entra group access to Resource Group 1, which hosts the logical servers.



These are the benefits:

- Mateo doesn't have to create local logins in each server.

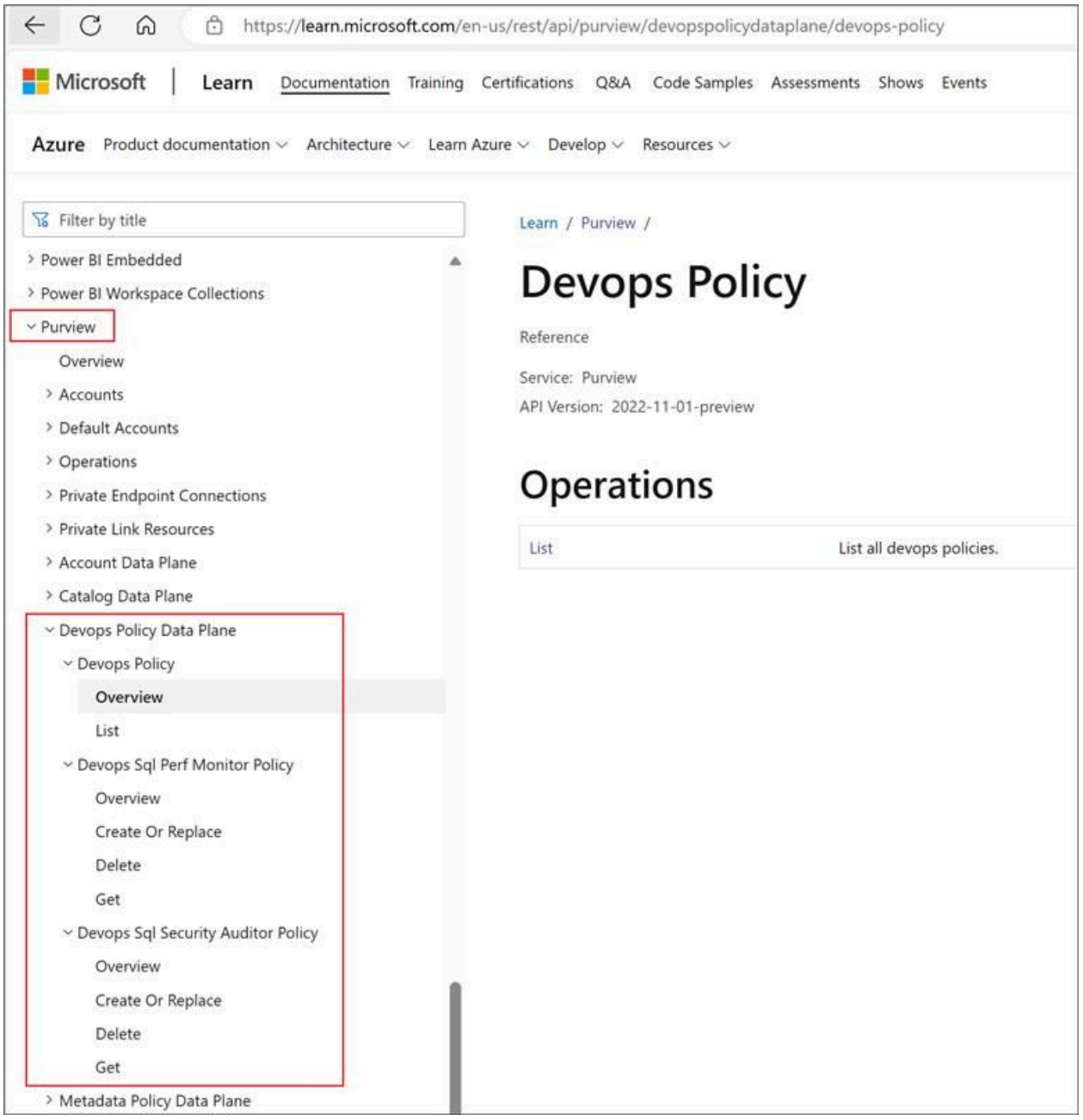
- The policies from Microsoft Purview improve security by limiting local privileged access. They support the principle of least privilege. In the scenario, Mateo grants only the minimum access that Bob and Alice need to perform the task of monitoring system health and performance.
- When new servers are added to the resource group, Mateo doesn't need to update the policy in Microsoft Purview for it to be enforced on the new servers.
- If Alice or Bob leaves the organization and the job is backfilled, Mateo just updates the Microsoft Entra group. He doesn't have to make any changes to the servers or to the policies that he created in Microsoft Purview.
- At any point in time, Mateo or the company's auditor can see all the permissions that were granted directly in Microsoft Purview Studio.

Expand table

Principle	Benefit
Simplify	The role definitions SQL Performance Monitor and SQL Security Auditor capture the permissions that typical IT and DevOps personas need to execute their job.
	There's less need for permission expertise on each data source type.
Reduce effort	A graphical interface lets you move through the data object hierarchy quickly.
	Microsoft Purview supports policies on entire Azure resource groups and subscriptions.
Enhance security	Access is granted centrally and can be easily reviewed and revoked.
	There's less need for privileged accounts to configure access directly at the data source.
	DevOps policies support the principle of least privilege via data resource scopes and role definitions.

DevOps policies API

Many sophisticated customers prefer to interact with Microsoft Purview via scripts rather than via the UI. Microsoft Purview DevOps policies now support a REST API that offers full create, read, update, and delete (CRUD) capability. This capability includes listing, policies for SQL Performance Monitor, and policies for SQL Security Auditor. For more information, see the API specification.



.

Mapping of popular DMVs and DMFs

SQL dynamic metadata includes a list of more than 700 DMVs and DMFs. The following table illustrates some of the most popular ones. The table maps the DMVs and DMFs to their role definitions in Microsoft Purview DevOps policies. It also provides links to reference content.

Expand table

DevOps role	Category	Example DMV or DMF
SQL Performance Monitor	Query system parameters to understand your system	sys.configurations

DevOps role	Category	Example DMV or DMF
		sys.dm_os_sys_info
	Identify performance bottlenecks	sys.dm_os_wait_stats
	Analyze currently running queries	sys.dm_exec_query_stats
	Analyze blocking issues	sys.dm_tran_locks
		sys.dm_exec_requests
		sys.dm_os_waiting_tasks
	Analyze memory usage	sys.dm_os_memory_clerks
	Analyze file usage and performance	sys.master_files
		sys.dm_io_virtual_file_stats
	Analyze index usage and fragmentation	sys.indexes
		sys.dm_db_index_usage_stats
		sys.dm_db_index_physical_stats
	Manage active user connections and internal tasks	sys.dm_exec_sessions
	Get procedure execution statistics	sys.dm_exec_procedure_stats
	Use the Query Store	sys.query_store_plan
		sys.query_store_query
		sys.query_store_query_text
	Get Error Log (not yet supported)	sys.sp_readerrorlog
SQL Security Auditor	Get audit details	sys.dm_server_audit_status
Both SQL Performance Monitor and SQL Security Auditor		sys.dm_audit_actions
		sys.dm_audit_class_type_map

DevOps role	Category	Example DMV or DMF

For more information on what IT support personnel can do when you grant them access via the Microsoft Purview roles, see the following resources:

- SQL Performance Monitor: Use Microsoft Purview to provide at-scale access to performance data in Azure SQL and SQL Server
- SQL Security Auditor: Security-related dynamic management views and functions

Next steps

To get started with DevOps policies, consult the following resources:

- Try DevOps policies for Azure SQL Database: Quickstart guide.
- See other videos, blogs, and articles.

Feedback

Was this page helpful? Yes No

Microsoft Purview Self-service data discovery and access (Preview)

Article • 02/26/2025

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

This article helps you understand Microsoft Purview Self-service data access policy.

> **Important**
>
> Microsoft Purview Self-service data access policy is currently in PREVIEW. The Supplemental Terms of Use for Microsoft Azure Previews include additional legal terms that apply to Azure features that are in beta, preview, or otherwise not yet released into general availability.

Important limitations

The self-service data access policy is only supported when the prerequisites mentioned in Data Policy Enforcement are satisfied.

Overview

Microsoft Purview Self-service data access workflow allows data consumer to request access to data when browsing or searching for data. Once the data access request is approved, a policy gets auto-generated to grant access to the requestor provided the data source is enabled for Data Policy Enforcement. Currently, self-service data access policy is supported for storage accounts, containers, folders, and files.

A **workflow admin** needs to map a self-service data access workflow to a collection. Collection is logical grouping of data sources that are registered within Microsoft Purview. **Only data source(s) that are registered** for Data Policy Enforcement has self-service policies auto-generated.

Terminology

- **Data consumer** is anyone who uses the data. Example, a data analyst accessing marketing data for customer segmentation. Data consumer and data requestor will be used interchangeably within this document.
- **Collection** is logical grouping of data sources that are registered within Microsoft Purview.
- **Self-service data access workflow** is the workflow that is initiated when a data consumer requests access to data.
- **Approver** is either a security group, a Microsoft Entra user, or a Microsoft Entra group that can approve self-service access requests.

How to use Microsoft Purview self-service data access policy

Microsoft Purview allows organizations to catalog metadata about all registered data assets. It allows data consumers to search for or browse to the required data asset.

With self-service data access workflow, data consumers can not only find data assets but also request access to the data assets. When the data consumer requests access to a data asset, the associated self-service data access workflow is triggered.

A default self-service data access workflow template is provided with every Microsoft Purview account. The default template can be amended to add more approvers and/or set the approver's email address. For more details, refer Create and enable self-service data access workflow.

Whenever a data consumer requests access to a dataset, the notification is sent to the workflow approver(s). The approver(s) can view the request and approve it either from Microsoft Purview governance portal or from within the email notification. When the request is approved, a policy is auto-generated and applied against the respective data source. Self-service data access policy gets auto-generated only if the data source is registered for **Data Policy Enforcement**. The prerequisites mentioned within the Data Policy Enforcement have to be satisfied.

Data consumer can access the requested dataset using tools such as Power BI or Azure Synapse Analytics workspace.

> ⓘ **Note**

Users won't be able to browse to the asset using the Azure portal or Storage explorer if the only permission granted is read/modify access at the file or folder level of the storage account.

⊗ Caution

Folder level permission is required to access data in ADLS Gen 2 using Power BI. Additionally, resource sets aren't supported by self-service policies. Hence, folder level permission needs to be granted to access resource set files such as CSV or parquet.

Next steps

If you would like to preview these features in your environment, follow the link below.

- Enable Data Policy Enforcement
- create self-service data access workflow
- working with policies at file level
- working with policies at folder level
- self-service policies for Azure SQL Database tables and views

Feedback

Was this page helpful? Yes No

Impact of deleting Microsoft Purview account on access policies

> **Note**
>
> Microsoft Purview Data Catalog (classic), Data Health Insights (classic), and Purview Workflow (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

Important considerations

Deleting a Microsoft Purview account that has active (that is, published) policies removes those policies. Any access to a data source or a dataset that was previously provisioned from Microsoft Purview gets revoked. This can lead to outages, that is, users or groups in your organization not able to access critical data. Review the decision to delete the Microsoft Purview account with the people in Policy Author role at root collection level before proceeding. To find out who holds that role in the Microsoft Purview account, review the section on managing role assignments in this guide.

Before deleting the Microsoft Purview account, it's advisable that you provision access to the users in your organization that need access to datasets using an alternate mechanism or a different Purview account. Then orderly delete or unpublish any active policies

- Deleting DevOps policies - You need to delete DevOps policies for them to be unpublished.
- Unpublishing Data Owner policies.
- Deleting Self-service access policies - You need to delete Self-service access policies for them to be unpublished.

Next steps

Check these concept guides

- DevOps policies
- Data owner access policies
- Self-service access policies

Last updated on 10/10/2025

Role definitions SQL Performance Monitor and SQL Security Auditor supported only from DevOps policies experience

> **Note**
>
> Microsoft Purview Data Catalog (classic), Data Health Insights (classic), and Purview Workflow (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

This guide discusses actions that have been retired from Data owner policies experience and are now supported only from the DevOps policies experience.

Important considerations

The following two actions that were previously available in the Microsoft Purview Data owner policies experience are now supported only from the Microsoft Purview DevOps policies, which is a more focused experience.

- SQL Performance Monitor
- SQL Security Auditor

If you currently have Data owner policies with these actions we encourage you to give the DevOps policies experience a try. Creating new policies or editing existing policies that involve these two actions is now unsupported from the Data owner experience.

Next steps

Check these concept guides

- DevOps policies

Last updated on 10/10/2025

Enable data policy enforcement on your Microsoft Purview sources

Article • 02/24/2025

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

Data policy enforcement is an option within the data source registration in Microsoft Purview. This option lets Microsoft Purview manage data access for your resources. The high level concept is that the data owner allows its data resource to be available for access policies by enabling *data policy enforcement*.

Currently, a data owner can enable data policy enforcement on a data resource, which enables it for these types of access policies:

- DevOps policies
- Data owner access policies
- Self-service access policies - access policies automatically generated by Microsoft Purview after a self-service access request is approved.
- Protection-policies

To be able to create any data policy on a resource, data policy enforcement must first be enabled on that resource. This article will explain how to enable data policy enforcement on your resources in Microsoft Purview.

> **Important**
>
> Because data policy enforcement directly affects access to your data, it directly affects your data security. Review additional considerations and best practices below before enabling data policy enforcement in your environment.

Prerequisites

Register the data source in Microsoft Purview

Before a policy can be created in Microsoft Purview for a data resource, you must register that data resource in Microsoft Purview Studio. You will find the instructions related to registering the data resource later in this guide.

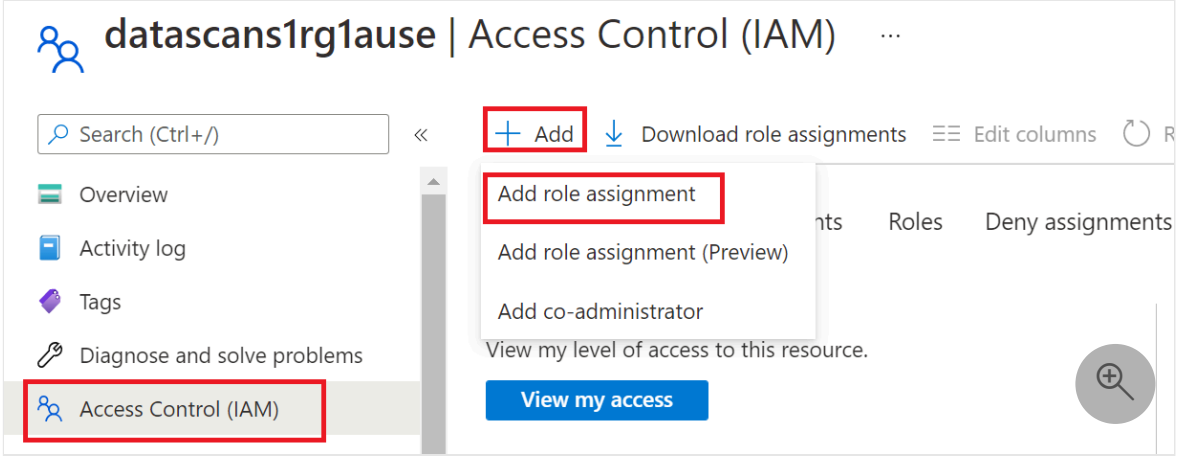
> **Note**
>
> Microsoft Purview policies rely on the data resource ARM path. If a data resource is moved to a new resource group or subscription it will need to be de-registered and then registered again in Microsoft Purview.

Configure permissions to enable Data policy enforcement on the data source

Once a resource is registered, but before a policy can be created in Microsoft Purview for that resource, you must configure permissions. A set of permissions are needed to enable the **Data policy enforcement**. This applies to data sources, resource groups, or subscriptions. To enable *Data policy enforcement*, you must have *both* specific Identity and Access Management (IAM) privileges on the resource as well as specific Microsoft Purview privileges:

- You must have *either one* of the following IAM role combinations on the resource's Azure Resource Manager path or any parent of it (that is, using IAM permission inheritance):
  - IAM Owner
  - Both IAM Contributor and IAM User Access Administrator

  To configure Azure role-based access control (RBAC) permissions, follow this guide. The following screenshot shows how to access the **Access Control** section in the Azure portal for the data resource to add a role assignment.

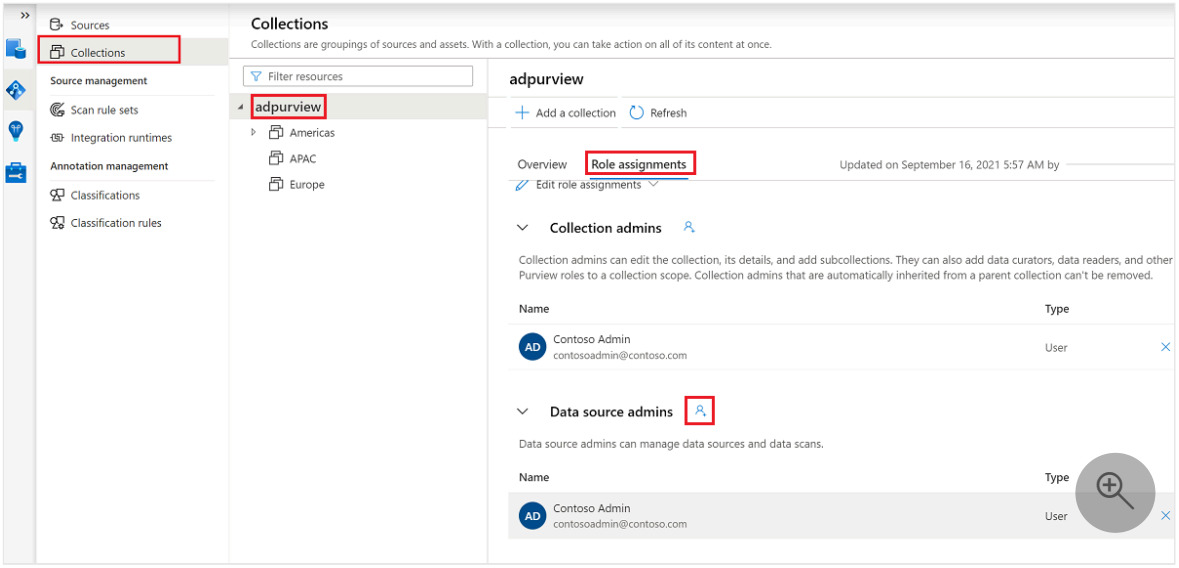


> **Note**

> The *IAM Owner* role for a data resource can be inherited from a parent resource group, a subscription, or a subscription management group. Check which Microsoft Entra users, groups, and service principals hold or are inheriting the *IAM Owner* role for the resource.

- You also need to have the Microsoft Purview *Data source admin* role for the collection or a parent collection (if inheritance is enabled). For more information, see the guide on managing Microsoft Purview role assignments.

  The following screenshot shows how to assign the *Data source admin* role at the root collection level.



Configure Microsoft Purview permissions to create, update, or delete access policies

To create, update or delete policies, you need to get the *Policy author* role in Microsoft Purview at *root collection level*:

- The *Policy author* role can create, update, and delete DevOps and Data Owner policies.
- The *Policy author* role can delete self-service access policies.

For more information about managing Microsoft Purview role assignments, see Create and manage collections in the Microsoft Purview Data Map.

> **Note**
>
> Policy author role must be configured at the root collection level.

In addition, to easily search Microsoft Entra users or groups when creating or updating the subject of a policy, you can greatly benefit from getting the Directory Readers permission in Microsoft Entra ID. This is a common permission for users in an Azure tenant. Without the Directory Reader permission, the Policy Author will have to type the complete username or email for all the principals included in the subject of a data policy.

Configure Microsoft Purview permissions for publishing Data Owner policies

Data Owner policies allow for checks and balances if you assign the Microsoft Purview *Policy author* and *Data source admin* roles to different people in the organization. Before a Data owner policy takes effect, a second person (*Data source admin*) must review it and explicitly approve it by publishing it. This does not apply to DevOps or Self-service access policies as publishing is automatic for them when those policies are created or updated.

To publish a Data owner policy you need to get the *Data source admin* role in Microsoft Purview at *root collection level*.

For more information about managing Microsoft Purview role assignments, see Create and manage collections in the Microsoft Purview Data Map.

> ⓘ **Note**
>
> To publish Data owner policies, the Data source admin role must be configured at the root collection level.

Delegate access provisioning responsibility to roles in Microsoft Purview

After a resource has been enabled for **Data policy enforcement**, any Microsoft Purview user with the *Policy author* role at the root collection level can provision access to that data source from Microsoft Purview.

> ⓘ **Note**
>
> Any Microsoft Purview root *Collection admin* can assign new users to root *Policy author* roles. Any *Collection admin* can assign new users to a *Data source admin*

role under the collection. Minimize and carefully vet the users who hold Microsoft Purview *Collection admin*, *Data source admin*, or *Policy author* roles.

If a Microsoft Purview account with published policies is deleted, such policies will stop being enforced within an amount of time that depends on the specific data source. This change can have implications on both security and data access availability. The Contributor and Owner roles in IAM can delete Microsoft Purview accounts. You can check these permissions by going to the **Access control (IAM)** section for your Microsoft Purview account and selecting **Role Assignments**. You can also use a lock to prevent the Microsoft Purview account from being deleted through Resource Manager locks.

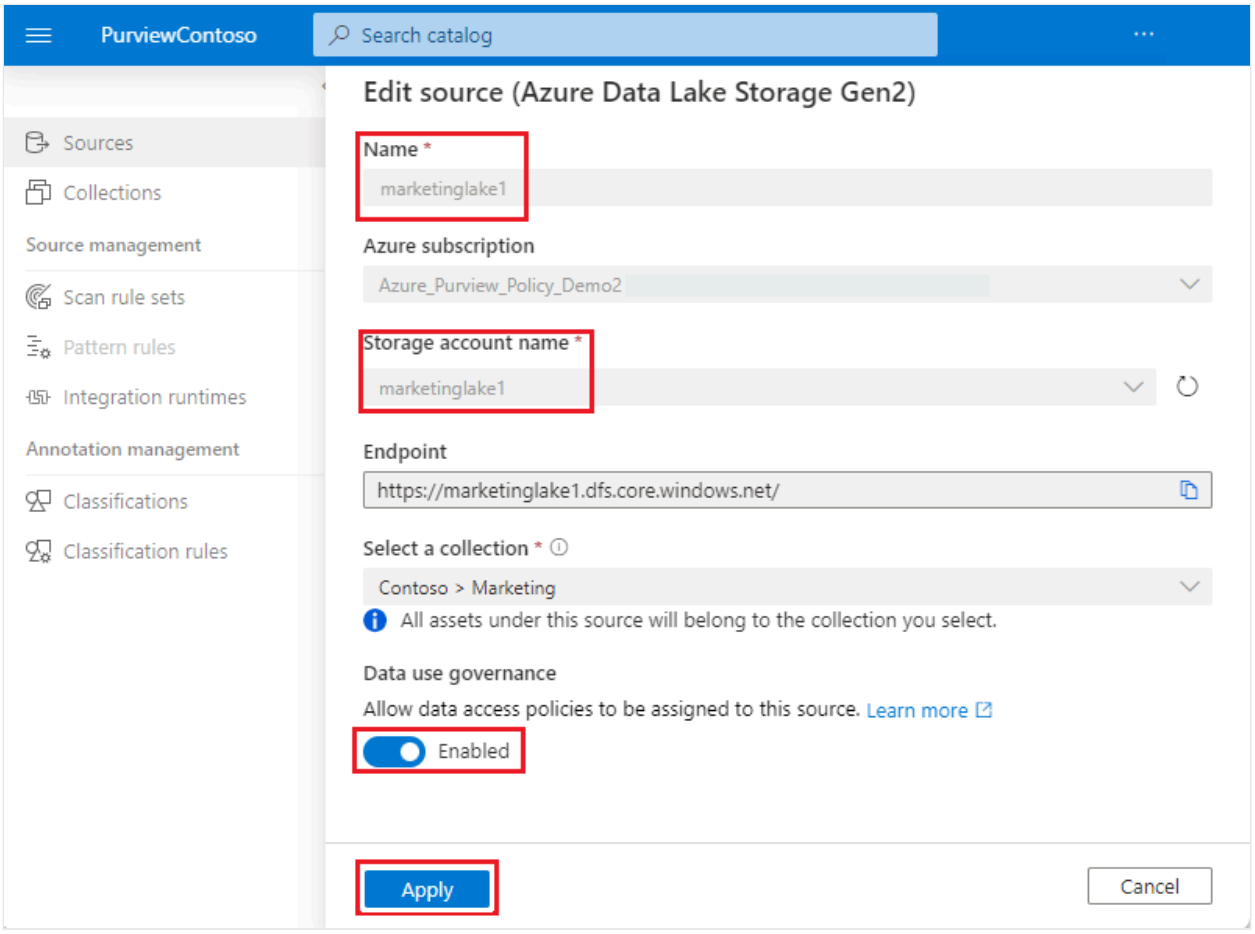
Enable data policy enforcement

To enable *data policy enforcement* for a resource, the resource will first need to be registered in Microsoft Purview. To register a resource, follow the **Prerequisites** and **Register** sections of the source pages for your resources.

Once you have your resource registered, follow the rest of the steps to enable an individual resource for *data policy enforcement*.

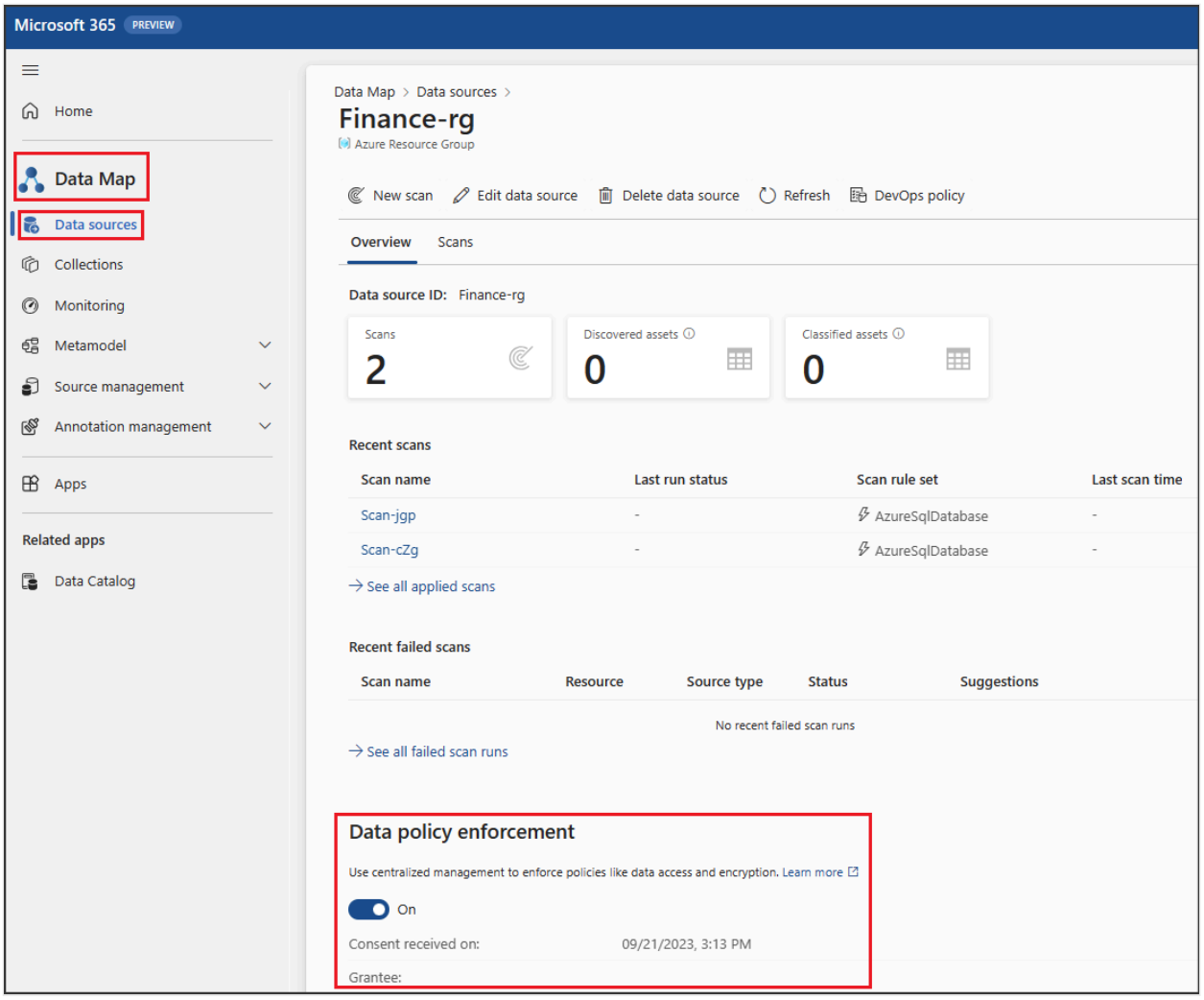
From the classic Microsoft Purview governance portal

1. Go to the classic Microsoft Purview governance portal.
2. Select the **Data map** tab in the left menu.
3. Select the **Sources** tab in the left menu.
4. Select the source where you want to enable *data policy enforcement*.
5. At the top of the source page, select **Edit source**.
6. Set the **Data policy enforcement** toggle to **Enabled**, as shown in the image below.



From the new Microsoft Purview portal

1. Go to the new Microsoft Purview portal.
2. Select the **Data Map** tab in the left menu.
3. Select the **Data sources** tab in the left menu.
4. Select the source where you want to enable *data policy enforcement*.
5. Set the **Data policy enforcement** toggle to **On**, as shown in the image below.



Disable data policy enforcement

To disable data policy enforcement for a source, resource group, or subscription, a user needs to either be a resource IAM **Owner** or a Microsoft Purview **Data source admin**. Once you have those permissions follow these steps:

From the classic Microsoft Purview governance portal

1. Go to the Microsoft Purview governance portal.
2. Select the **Data map** tab in the left menu.
3. Select the **Sources** tab in the left menu.
4. Select the source you want to disable data policy enforcement for.
5. At the top of the source page, select **Edit source**.
6. Set the **Data policy enforcement** toggle to **Disabled**.

From the new Microsoft Purview portal

1. Go to the new Microsoft Purview portal.
2. Select the **Data map** tab in the left menu.
3. Select the **Sources** tab in the left menu.
4. Select the source you want to disable data policy enforcement for.
5. Set the **Data policy enforcement** toggle to **Off**.

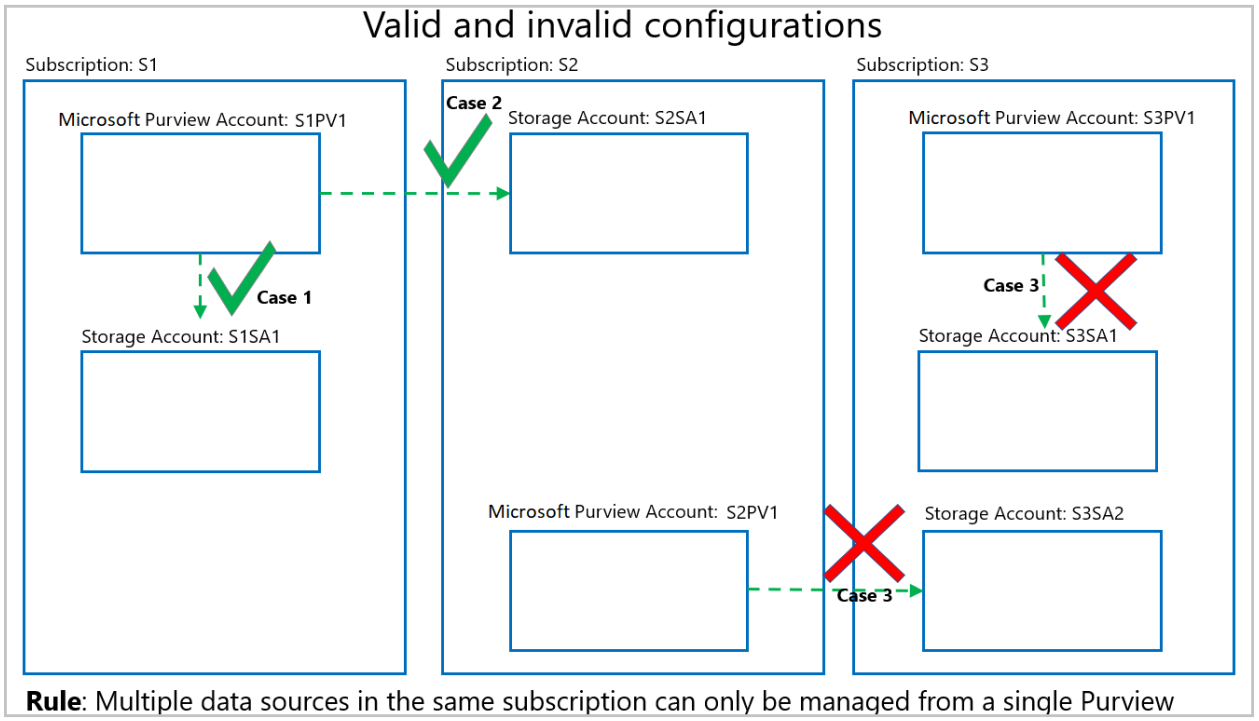
Additional considerations related to data policy enforcement

- Make sure you write down the **Name** you use when registering in Microsoft Purview. You will need it when you publish a policy. The recommended practice is to make the registered name exactly the same as the endpoint name.
- To disable a source for *data policy enforcement*, you first have to remove any published policies on that data source.
- While user needs to have both data source *Owner* and Microsoft Purview *Data source admin* to enable a source for *data policy enforcement*, **any** Data Source admin for the collection can disable it.
- Disabling *data policy enforcement* for a subscription will disable it also for all assets registered in that subscription.

> ⚠ **Warning**
>
> **Known issues** related to source registration
>
> - Moving data sources to a different resource group or subscription is not supported. If want to do that, de-register the data source in Microsoft Purview before moving it and then register it again after that happens. Note that policies are bound to the data source ARM path. Changing the data source subscription or resource group makes policies ineffective.
> - Once a subscription gets disabled for *data policy enforcement* any underlying assets that are enabled for *data policy enforcement* will be disabled, which is the right behavior. However, policy statements based on those assets will still be allowed after that.

Data policy enforcement best practices

- We highly encourage registering data sources for *data policy enforcement* and managing all associated access policies in a single Microsoft Purview account.
- Should you have multiple Microsoft Purview accounts, be aware that **all** data sources belonging to a subscription must be registered for *data policy enforcement* in a single Microsoft Purview account. That Microsoft Purview account can be in any subscription in the tenant. The *data policy enforcement* toggle will become greyed out when there are invalid configurations. Some examples of valid and invalid configurations follow in the diagram below:
  - **Case 1** shows a valid configuration where a Storage account is registered in a Microsoft Purview account in the same subscription.
  - **Case 2** shows a valid configuration where a Storage account is registered in a Microsoft Purview account in a different subscription.
  - **Case 3** shows an invalid configuration arising because Storage accounts S3SA1 and S3SA2 both belong to Subscription 3, but are registered to different Microsoft Purview accounts. In that case, the *data policy enforcement* toggle will only enable in the Microsoft Purview account that wins and registers a data source in that subscription first. The toggle will then be greyed out for the other data source.
- If the *data policy enforcement* toggle is greyed out and cannot be enabled, hover over it to know the name of the Microsoft Purview account that has registered the data resource first.



Next steps

- Create data owner policies for your resources
- Enable Microsoft Purview data owner policies on all data sources in a subscription or a resource group
- Enable Microsoft Purview data owner policies on an Azure Storage account

Feedback

Was this page helpful? Yes No

Tutorial: Troubleshoot Microsoft Purview policies for SQL data sources

Article • 02/26/2025

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

In this tutorial, you learn how issue SQL commands to inspect the Microsoft Purview policies that have been communicated to the SQL instance, where they'll be enforced. You'll also learn how to force a download of the policies to the SQL instance. These commands are only used for troubleshooting and aren't required during the normal operation of Microsoft Purview policies. These commands require a higher level of privileges in the SQL instance.

For more information about Microsoft Purview policies, see the concept guides listed in the Next steps section.

Prerequisites

- An Azure subscription. If you don't already have one, create a free subscription.
- A Microsoft Purview account. If you don't have one, see the quickstart for creating a Microsoft Purview account.
- Register a data source, enable *Data policy enforcement*, and create a policy. To do so, use one of the Microsoft Purview policy guides. To follow along with the examples in this tutorial, you can create a DevOps policy for Azure SQL Database.

Test the policy

Once you create a policy, the Microsoft Entra principals referenced in the Subject of the policy should be able to connect to any database in the server to which the policies are published.

Force policy download

It's possible to force an immediate download of the latest published policies to the current SQL database by running the following command. The minimal permission required to run it's membership in ##MS_ServerStateManager##-server role.

SQL

```
-- Force immediate download of latest published policies
exec sp_external_policy_refresh reload
```

Analyze downloaded policy state from SQL

The following DMVs can be used to analyze which policies have been downloaded and are currently assigned to Microsoft Entra principals. The minimal permission required to run them is VIEW DATABASE SECURITY STATE - or assigned Action Group *SQL Security Auditor*.

SQL

```
-- Lists generally supported actions
SELECT * FROM sys.dm_server_external_policy_actions

-- Lists the roles that are part of a policy published to this server
SELECT * FROM sys.dm_server_external_policy_roles

-- Lists the links between the roles and actions, could be used to join the
two
SELECT * FROM sys.dm_server_external_policy_role_actions

-- Lists all Azure AD principals that were given connect permissions
SELECT * FROM sys.dm_server_external_policy_principals

-- Lists Azure AD principals assigned to a given role on a given resource
scope
SELECT * FROM sys.dm_server_external_policy_role_members

-- Lists Azure AD principals, joined with roles, joined with their data
actions
SELECT * FROM sys.dm_server_external_policy_principal_assigned_actions
```

Next steps

Concept guides for Microsoft Purview access policies:

- DevOps policies
- Self-service access policies

- Data owner policies

Feedback

Was this page helpful? Yes No

Create, list, update, and delete Microsoft Purview DevOps policies

Article • 02/26/2025

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

DevOps policies are a type of Microsoft Purview access policies. You can use them to manage access to system metadata on data sources that have been registered for **Data policy enforcement** in Microsoft Purview.

You can configure DevOps policies directly from the Microsoft Purview governance portal. After they're saved, they're automatically published and then enforced by the data source. Microsoft Purview policies manage only access for Microsoft Entra principals.

This guide covers the configuration steps in Microsoft Purview to provision access to database system metadata by using the DevOps policy actions for the SQL Performance Monitor and SQL Security Auditor roles. It shows how to create, list, update, and delete DevOps policies.

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- A new or existing Microsoft Purview account. Follow this quickstart guide to create one.

Configuration

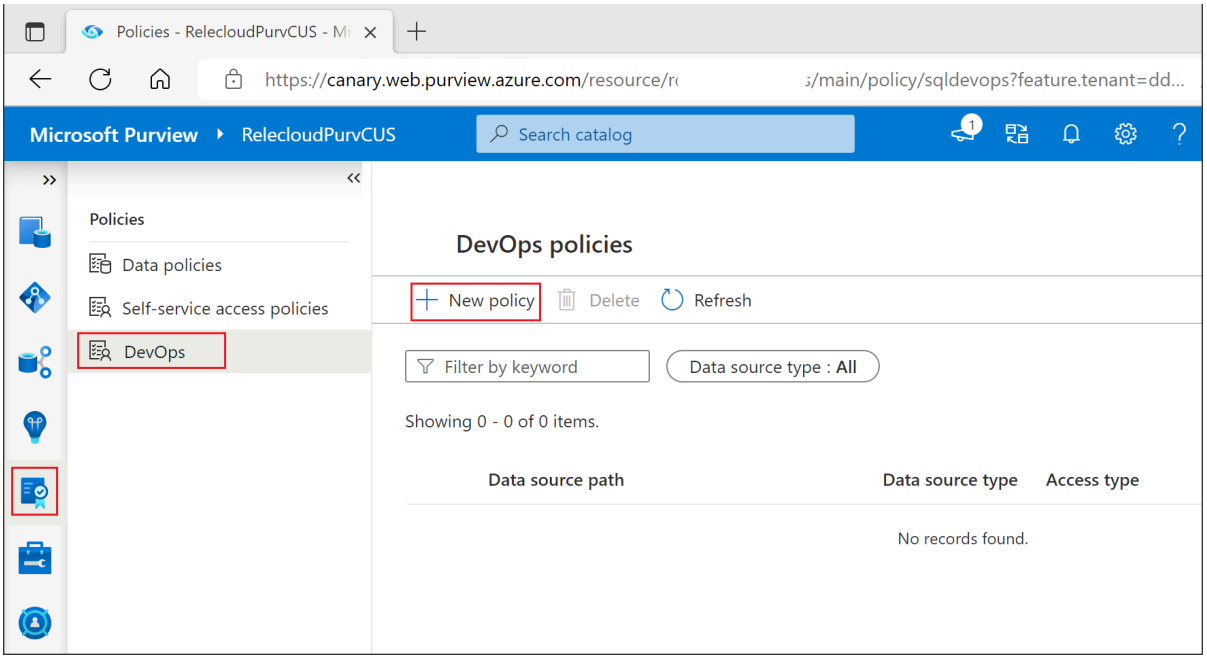
Before you author policies in the Microsoft Purview policy portal, you need to configure the data sources so that they can enforce those policies:

1. Follow any policy-specific prerequisites for your source. Check the table of Microsoft Purview supported data sources and select the link in the **Access policy** column for sources where access policies are available. Follow any steps listed in the "Access policy" and "Prerequisites" sections.
2. Register the data source in Microsoft Purview. Follow the "Prerequisites" and "Register" sections of the source pages for your resources.
3. Turn on the **Data policy enforcement** toggle in the data source registration. For more information, including additional permissions that you need for this step, see Enable Data policy enforcement on your Microsoft Purview sources.

Create a new DevOps policy

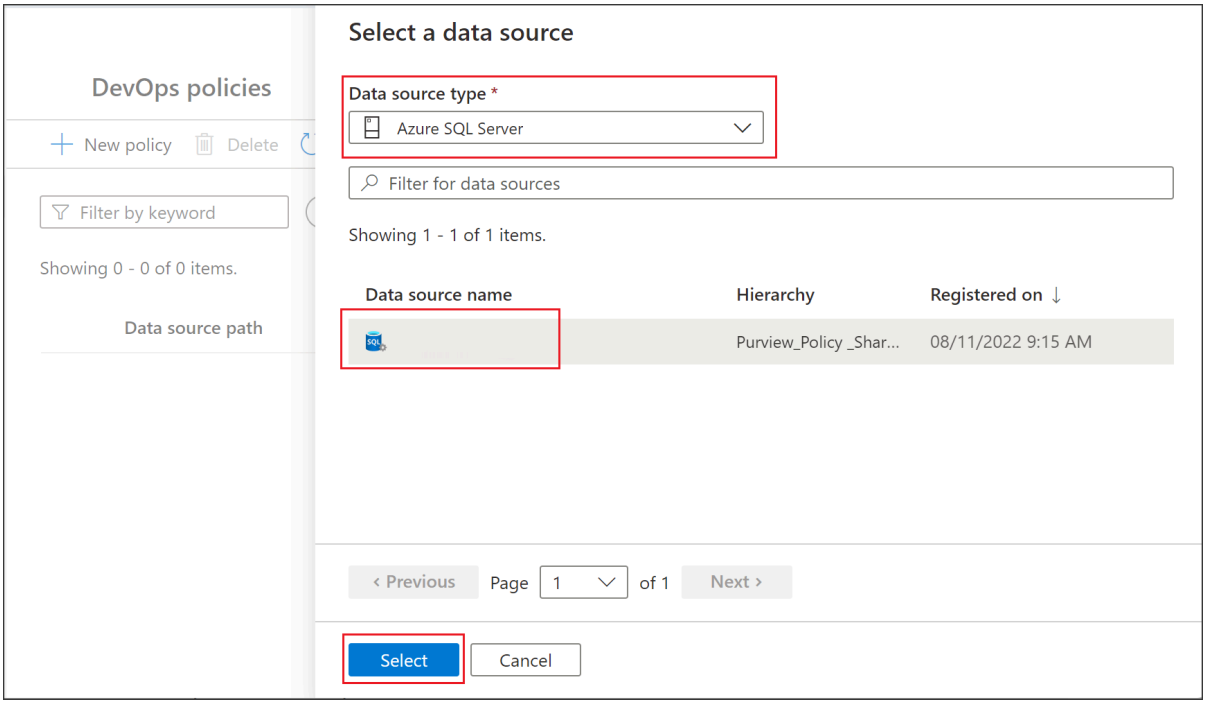
To create a DevOps policy, first ensure that you have the Microsoft Purview Policy Author role at the root collection level. Check the section on managing Microsoft Purview role assignments in this guide. Then, follow these steps:

1. Sign in to the Microsoft Purview governance portal.
2. On the left pane, select **Data policy**. Then select **DevOps policies**.
3. Select the **New policy** button.



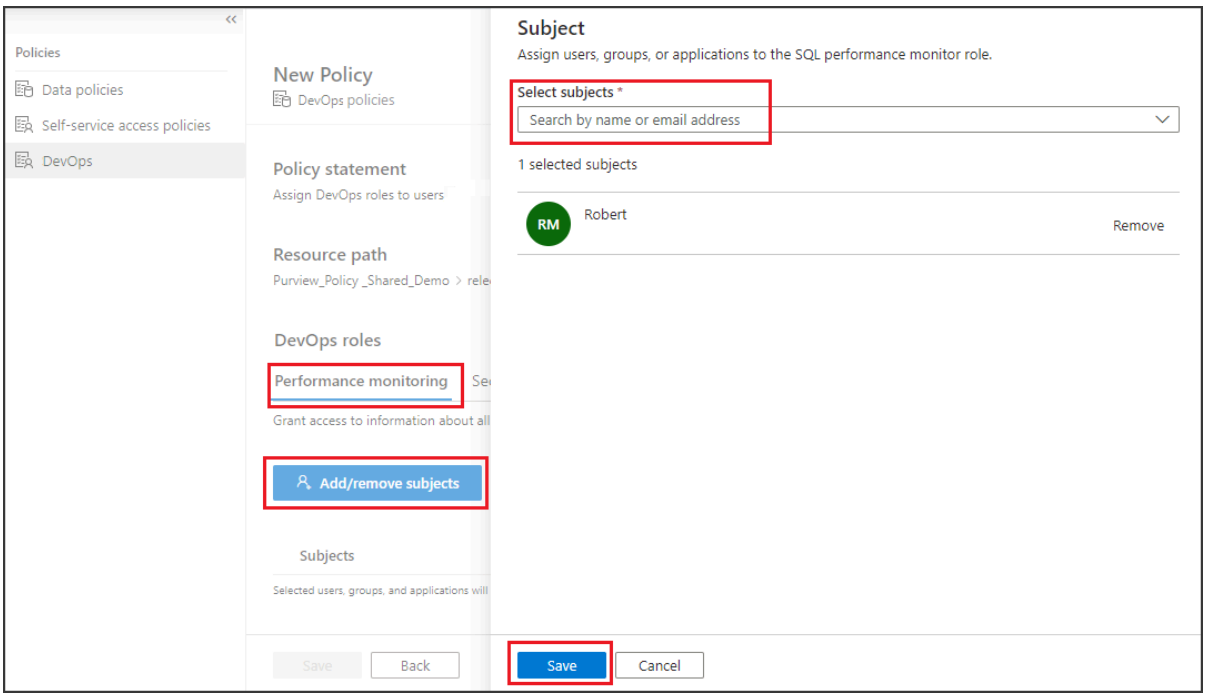
The policy detail panel opens.

4. For **Data source type**, select a data source. Under **Data source name**, select one of the listed data sources. Then click **Select** to return to the **New Policy** pane.



5. Select one of two roles, **Performance monitoring** or **Security auditing**. Then select **Add/remove subjects** to open the **Subject** panel.

   In the **Select subjects** box, enter the name of a Microsoft Entra principal (user, group, or service principal). Microsoft 365 groups are supported, but updates to group membership take up to one hour to be reflected in Microsoft Entra ID. Keep adding or removing subjects until you're satisfied, and then select **Save**.



6. Select **Save** to save the policy. The policy is published automatically. Enforcement starts at the data source within five minutes.

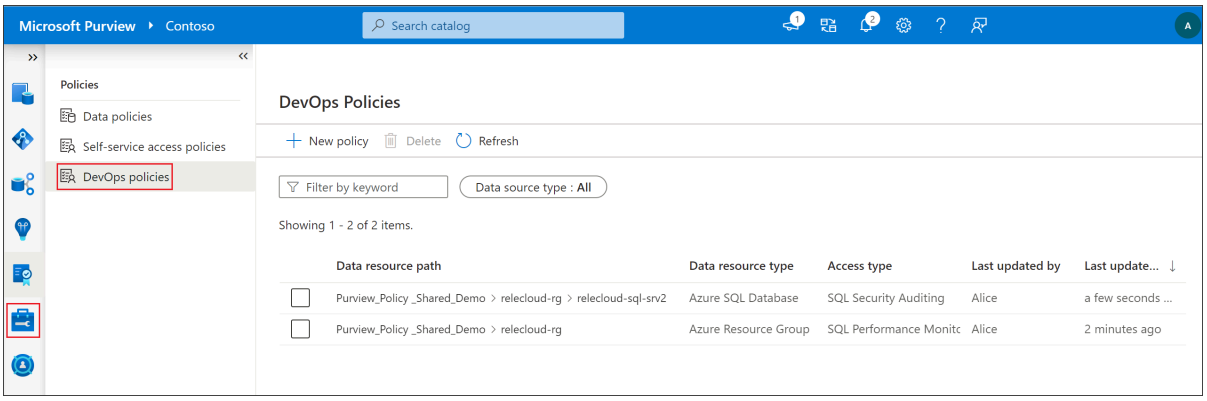
List DevOps policies

To list DevOps policies, first ensure that you have one of the following Microsoft Purview roles at the root collection level: Policy Author, Data Source Admin, Data Curator, or Data Reader. Check the section on managing Microsoft Purview role assignments in this guide. Then, follow these steps:

1. Sign in to the Microsoft Purview governance portal.

2. On the left pane, select **Data policy**. Then select **DevOps policies**.

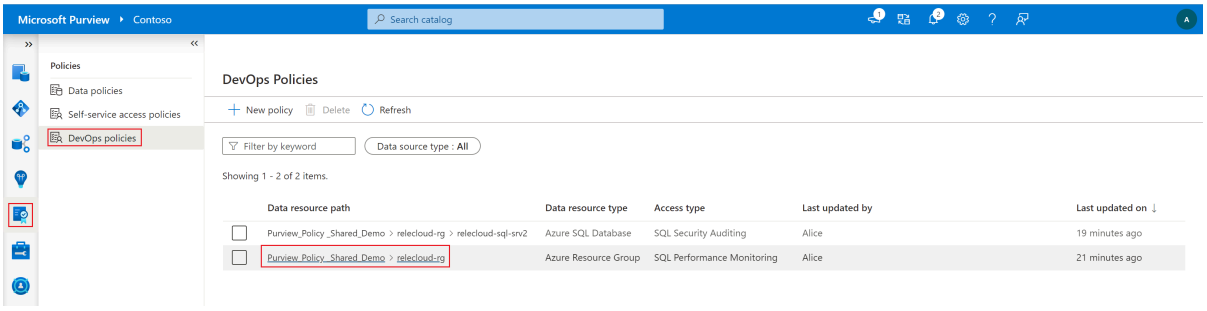
   The **DevOps Policies** pane lists any policies that have been created.



Update a DevOps policy

To update a DevOps policy, first ensure that you have the Microsoft Purview Policy Author role at the root collection level. Check the section on managing Microsoft Purview role assignments in this guide. Then, follow these steps:

1. Sign in to the Microsoft Purview governance portal.
2. On the left pane, select **Data policy**. Then select **DevOps policies**.
3. On the **DevOps Policies** pane, open the policy details for one of the policies by selecting it from its data resource path.

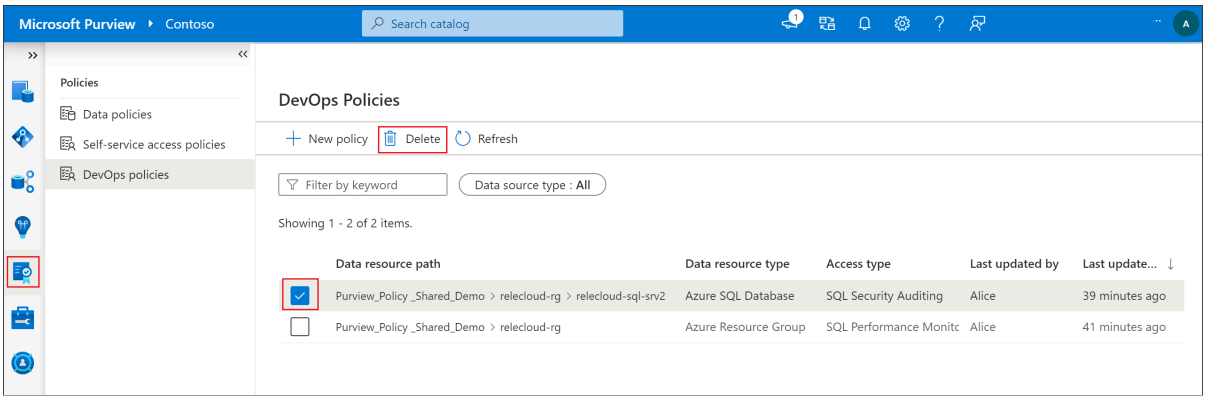


4. On the pane for policy details, select **Edit**.
5. Make your changes, and then select **Save**.

Delete a DevOps policy

To delete a DevOps policy, first ensure that you have the Microsoft Purview Policy Author role at the root collection level. Check the section on managing Microsoft Purview role assignments in this guide. Then, follow these steps:

1. Sign in to the Microsoft Purview governance portal.
2. On the left pane, select **Data policy**. Then select **DevOps policies**.
3. Select the checkbox for one of the policies, and then select **Delete**.



Test the DevOps policy

After you create a policy, any of the Microsoft Entra users that you selected as subjects can now connect to the data sources in the scope of the policy. To test, use SQL Server Management Studio (SSMS) or any SQL client and try to query some dynamic management views (DMVs) and dynamic management functions (DMFs). The following sections list a few examples. For more examples, consult the mapping of popular DMVs and DMFs in What can I accomplish with Microsoft Purview DevOps policies?.

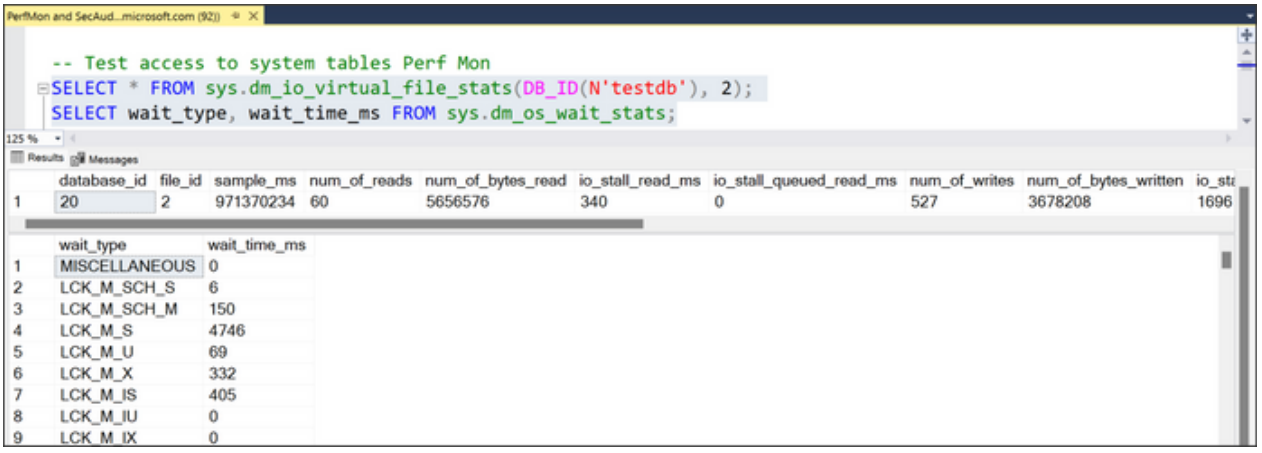
If you require more troubleshooting, see the Next steps section in this guide.

Test SQL Performance Monitor access

If you provided the subjects of the policy for the SQL Performance Monitor role, you can issue the following commands:

SQL

```
-- Returns I/O statistics for data and log files
SELECT * FROM sys.dm_io_virtual_file_stats(DB_ID(N'testdb'), 2)
-- Waits encountered by threads that executed. Used to diagnose performance issues
SELECT wait_type, wait_time_ms FROM sys.dm_os_wait_stats
```



Test SQL Security Auditor access

If you provided the subjects of the policy for the SQL Security Auditor role, you can issue the following commands from SSMS or any SQL client:

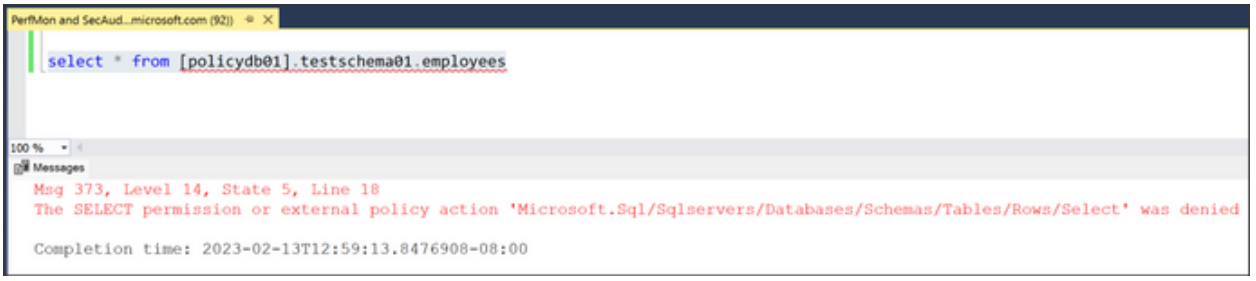
```sql
-- Returns the current state of the audit
SELECT * FROM sys.dm_server_audit_status
-- Returns information about the encryption state of a database and its associated database encryption keys
SELECT * FROM sys.dm_database_encryption_keys
```

Ensure no access to user data

Try to access a table in one of the databases by using the following command:

```sql
-- Test access to user data
SELECT * FROM [databaseName].schemaName.tableName
```

The Microsoft Entra principal that you're testing with should be denied, which means the data is protected from insider threats.



Role definition detail

The following table maps Microsoft Purview data policy roles to specific actions in SQL data sources.

Expand table

Microsoft Purview policy role	Actions in data sources
SQL Performance Monitor	Microsoft.Sql/Sqlservers/Connect
	Microsoft.Sql/Sqlservers/Databases/Connect
	Microsoft.Sql/Sqlservers/Databases/SystemViewsAndFunctions/DatabasePerformanceState/Rows/Select

Microsoft Purview policy role	Actions in data sources
	Microsoft.Sql/Sqlservers/SystemViewsAndFunctions/ServerPerformanceState/Rows/Select
	Microsoft.Sql/Sqlservers/Databases/SystemViewsAndFunctions/DatabaseGeneralMetadata/Rows/Select
	Microsoft.Sql/Sqlservers/SystemViewsAndFunctions/ServerGeneralMetadata/Rows/Select
	Microsoft.Sql/Sqlservers/Databases/DBCCs/ViewDatabasePerformanceState/Execute
	Microsoft.Sql/Sqlservers/DBCCs/ViewServerPerformanceState/Execute
	Microsoft.Sql/Sqlservers/Databases/ExtendedEventSessions/Create
	Microsoft.Sql/Sqlservers/Databases/ExtendedEventSessions/Options/Alter
	Microsoft.Sql/Sqlservers/Databases/ExtendedEventSessions/Events/Add
	Microsoft.Sql/Sqlservers/Databases/ExtendedEventSessions/Events/Drop
	Microsoft.Sql/Sqlservers/Databases/ExtendedEventSessions/State/Enable
	Microsoft.Sql/Sqlservers/Databases/ExtendedEventSessions/State/Disable
	Microsoft.Sql/Sqlservers/Databases/ExtendedEventSessions/Drop
	Microsoft.Sql/Sqlservers/Databases/ExtendedEventSessions/Target/Add
	Microsoft.Sql/Sqlservers/Databases/ExtendedEventSessions/Target/Drop
	Microsoft.Sql/Sqlservers/ExtendedEventSessions/Create
	Microsoft.Sql/Sqlservers/ExtendedEventSessions/Options/Alter
	Microsoft.Sql/Sqlservers/ExtendedEventSessions/Events/Add
	Microsoft.Sql/Sqlservers/ExtendedEventSessions/Events/Drop
	Microsoft.Sql/Sqlservers/ExtendedEventSessions/State/Enable
	Microsoft.Sql/Sqlservers/ExtendedEventSessions/State/Disable
	Microsoft.Sql/Sqlservers/ExtendedEventSessions/Drop
	Microsoft.Sql/Sqlservers/ExtendedEventSessions/Target/Add
	Microsoft.Sql/Sqlservers/ExtendedEventSessions/Target/Drop
SQL Security Auditor	Microsoft.Sql/Sqlservers/Connect
	Microsoft.Sql/Sqlservers/Databases/Connect
	Microsoft.Sql/sqlservers/SystemViewsAndFunctions/ServerSecurityState/rows/select
	Microsoft.Sql/Sqlservers/Databases/SystemViewsAndFunctions/DatabaseSecurityState/rows/select

Microsoft Purview policy role	Actions in data sources
	Microsoft.Sql/sqlservers/SystemViewsAndFunctions/ServerSecurityMetadata/rows/select
	Microsoft.Sql/Sqlservers/Databases/SystemViewsAndFunctions/DatabaseSecurityMetadata/rows/select

Next steps

Check the following blogs, videos, and related articles:

- Blog: Microsoft Purview DevOps policies for Azure SQL Database is now generally available
- Blog: Inexpensive solution for managing access to SQL health, performance, and security information
- Blog: Microsoft Purview DevOps policies enable at-scale access provisioning for IT operations
- Blog: Microsoft Purview DevOps policies API is now public
- Video: Prerequisite for policies: The "Data policy enforcement" option
- Video: Quick overview of DevOps policies
- Video: Deep dive for DevOps policies
- Article: Microsoft Purview DevOps policies concept guide
- Article: Microsoft Purview DevOps policies on Azure Arc-enabled SQL Server
- Article: Microsoft Purview DevOps policies on Azure SQL Database
- Article: Microsoft Purview DevOps policies on entire resource groups or subscriptions
- Article: Troubleshoot Microsoft Purview policies for SQL data sources

Feedback

Was this page helpful? Yes No

Provision access to system metadata in Azure SQL Database using Microsoft Purview DevOps policies

Article • 10/28/2024

> **Note**
>
> Microsoft Purview Data Map (classic), Data Catalog (classic), Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

DevOps policies are a type of Microsoft Purview access policies. They allow you to manage access to system metadata on data sources that have been registered for *Data policy enforcement* in Microsoft Purview. These policies are configured directly from the Microsoft Purview governance portal, and after they're saved, they get automatically published and then enforced by the data source. Microsoft Purview policies only manage access for Microsoft Entra principals.

This how-to guide covers how to configure Azure SQL Database to enforce policies created in Microsoft Purview. It covers the configuration steps for Azure SQL Database and the ones in Microsoft Purview to provision access to Azure SQL Database system metadata (DMVs and DMFs) using the DevOps policies actions *SQL Performance Monitoring* or *SQL Security Auditing*.

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- A new or existing Microsoft Purview account. Follow this quickstart guide to create one.
- Create a new Azure SQL Database instance, or use an existing one, in one of the currently available regions for this feature. You can follow this guide to create an Azure SQL Database instance.

Region support

All Microsoft Purview regions are supported.

Enforcement of Microsoft Purview policies is available only in the following regions for Azure SQL Database:

Public cloud:

- East US
- East US2
- South Central US
- West Central US
- West US3
- Canada Central
- Brazil South
- West Europe
- North Europe
- France Central
- UK South
- South Africa North
- Central India
- South East Asia
- East Asia
- Australia East

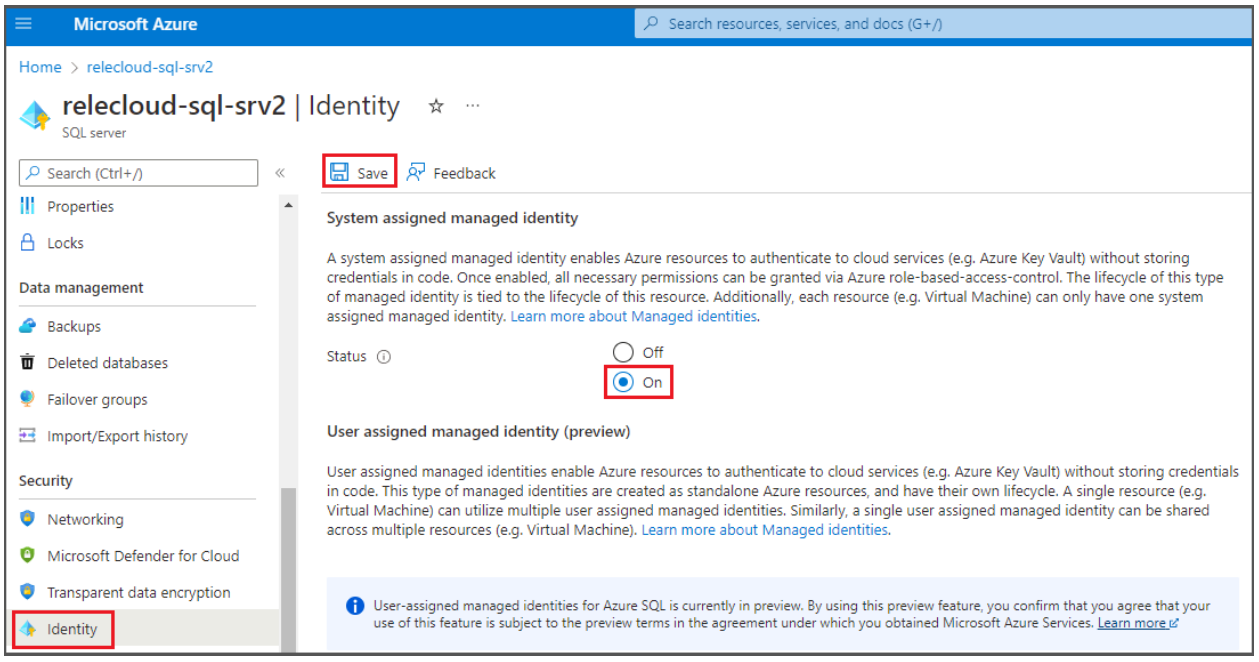
Sovereign clouds:

- USGov Virginia
- China North 3

Configure the Azure SQL Database instance for policies from Microsoft Purview

For the logical server associated with Azure SQL Database to honor policies from Microsoft Purview, you need to configure a Microsoft Entra admin. In the Azure portal, go to the logical server that hosts the Azure SQL Database instance. On the side menu, select **Microsoft Entra ID**. Set an admin name to any Microsoft Entra user or group that you prefer, and then select **Save**.

Then, on the side menu, select **Identity**. Under **System assigned managed identity**, turn the status to **On**, and then select **Save**.



Microsoft Purview Configuration

Register the data source in Microsoft Purview

Before a policy can be created in Microsoft Purview for a data resource, you must register that data resource in Microsoft Purview Studio. You will find the instructions related to registering the data resource later in this guide.

> **Note**
>
> Microsoft Purview policies rely on the data resource ARM path. If a data resource is moved to a new resource group or subscription it will need to be de-registered and then registered again in Microsoft Purview.

Configure permissions to enable Data policy enforcement on the data source

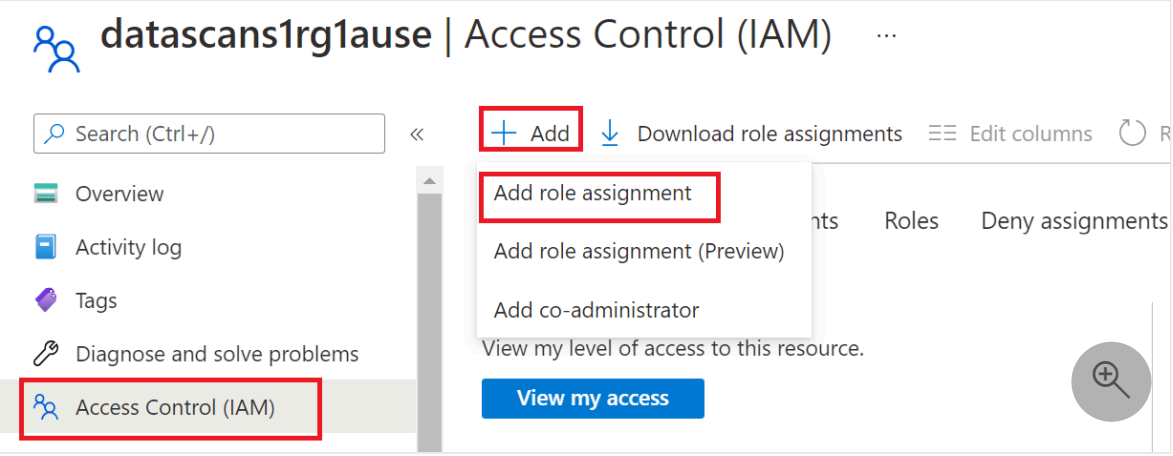
Once a resource is registered, but before a policy can be created in Microsoft Purview for that resource, you must configure permissions. A set of permissions are needed to enable the **Data policy enforcement**. This applies to data sources, resource groups, or subscriptions. To enable *Data policy enforcement*, you must have *both* specific Identity and Access Management (IAM) privileges on the resource as well as specific Microsoft Purview privileges:

- You must have *either one* of the following IAM role combinations on the resource's Azure Resource Manager path or any parent of it (that is, using IAM permission

inheritance):

- IAM Owner
- Both IAM Contributor and IAM User Access Administrator

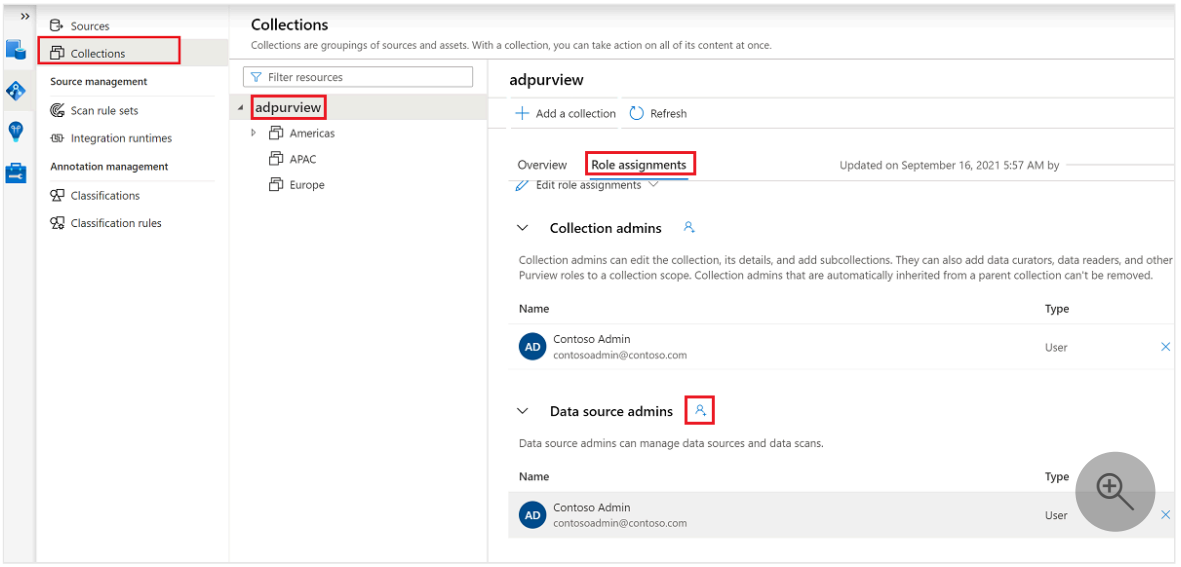
To configure Azure role-based access control (RBAC) permissions, follow this guide. The following screenshot shows how to access the **Access Control** section in the Azure portal for the data resource to add a role assignment.



> **Note**
>
> The *IAM Owner* role for a data resource can be inherited from a parent resource group, a subscription, or a subscription management group. Check which Microsoft Entra users, groups, and service principals hold or are inheriting the *IAM Owner* role for the resource.

- You also need to have the Microsoft Purview *Data source admin* role for the collection or a parent collection (if inheritance is enabled). For more information, see the guide on managing Microsoft Purview role assignments.

  The following screenshot shows how to assign the *Data source admin* role at the root collection level.



Configure Microsoft Purview permissions to create, update, or delete access policies

To create, update or delete policies, you need to get the *Policy author* role in Microsoft Purview at *root collection level*:

- The *Policy author* role can create, update, and delete DevOps and Data Owner policies.
- The *Policy author* role can delete self-service access policies.

For more information about managing Microsoft Purview role assignments, see Create and manage collections in the Microsoft Purview Data Map.

> **Note**
>
> Policy author role must be configured at the root collection level.

In addition, to easily search Microsoft Entra users or groups when creating or updating the subject of a policy, you can greatly benefit from getting the Directory Readers permission in Microsoft Entra ID. This is a common permission for users in an Azure tenant. Without the Directory Reader permission, the Policy Author will have to type the complete username or email for all the principals included in the subject of a data policy.

Configure Microsoft Purview permissions for publishing Data Owner policies

Data Owner policies allow for checks and balances if you assign the Microsoft Purview *Policy author* and *Data source admin* roles to different people in the organization. Before

a Data owner policy takes effect, a second person (*Data source admin*) must review it and explicitly approve it by publishing it. This does not apply to DevOps or Self-service access policies as publishing is automatic for them when those policies are created or updated.

To publish a Data owner policy you need to get the *Data source admin* role in Microsoft Purview at *root collection level*.

For more information about managing Microsoft Purview role assignments, see Create and manage collections in the Microsoft Purview Data Map.

> ⓘ **Note**
>
> To publish Data owner policies, the Data source admin role must be configured at the root collection level.

Delegate access provisioning responsibility to roles in Microsoft Purview

After a resource has been enabled for **Data policy enforcement**, any Microsoft Purview user with the *Policy author* role at the root collection level can provision access to that data source from Microsoft Purview.

> ⓘ **Note**
>
> Any Microsoft Purview root *Collection admin* can assign new users to root *Policy author* roles. Any *Collection admin* can assign new users to a *Data source admin* role under the collection. Minimize and carefully vet the users who hold Microsoft Purview *Collection admin*, *Data source admin*, or *Policy author* roles.

If a Microsoft Purview account with published policies is deleted, such policies will stop being enforced within an amount of time that depends on the specific data source. This change can have implications on both security and data access availability. The Contributor and Owner roles in IAM can delete Microsoft Purview accounts. You can check these permissions by going to the **Access control (IAM)** section for your Microsoft Purview account and selecting **Role Assignments**. You can also use a lock to prevent the Microsoft Purview account from being deleted through Resource Manager locks.

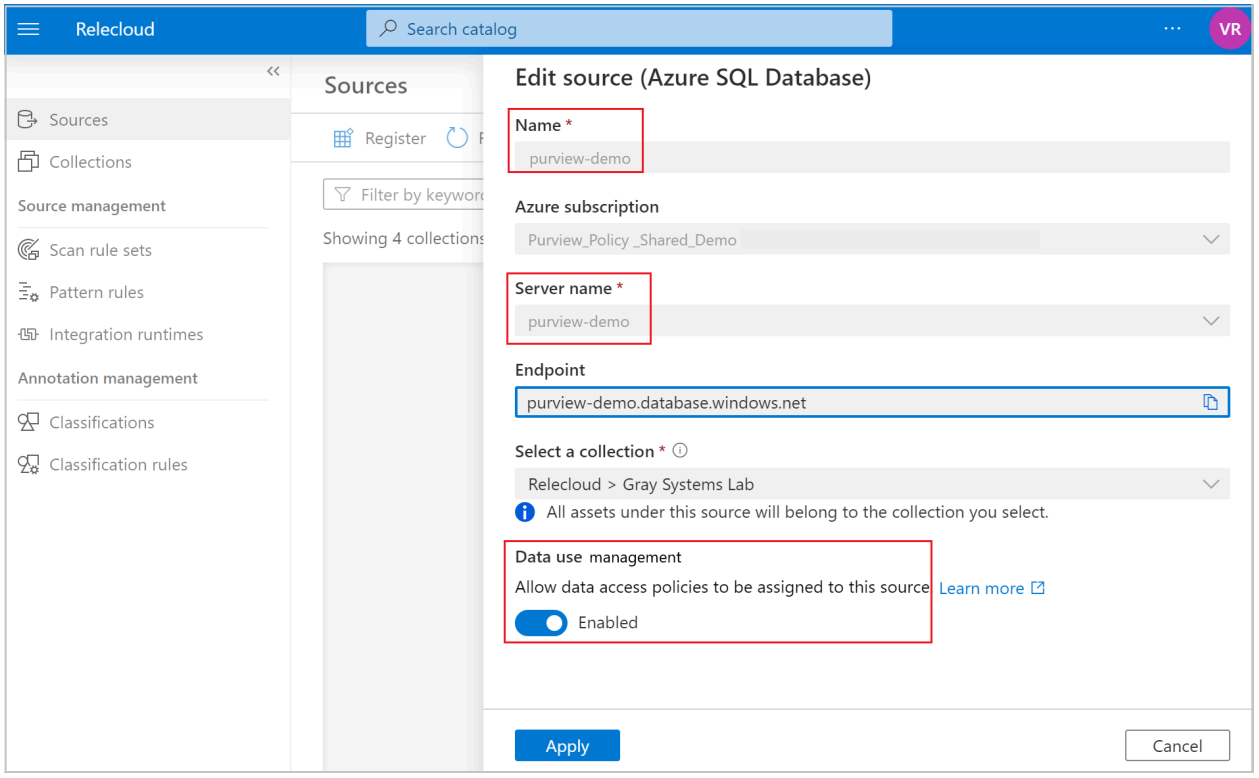
Register the data sources in Microsoft Purview

The Azure SQL Database data source needs to be registered first with Microsoft Purview, before access policies can be created. You can follow these guides:

Register and scan Azure SQL Database.

After you've registered your resources, you'll need to enable Data policy enforcement. Data policy enforcement needs certain permissions and can affect the security of your data, as it delegates to certain Microsoft Purview roles the ability to manage access to data sources. **Go through the secure practices related to Data policy enforcement in this guide**: How to enable Data policy enforcement

Once your data source has the **Data policy enforcement** toggle *Enabled*, it looks like this screenshot. This enables the access policies to be used with the given data source.



Return to the Azure portal for Azure SQL Database to verify it's now governed by Microsoft Purview:

1. Sign in to the Azure portal through this link
2. Select the Azure SQL Server that you want to configure.
3. Go to **Microsoft Entra ID** on the left pane.
4. Scroll down to **Microsoft Purview access policies**.
5. Select the button to **Check for Microsoft Purview Governance**. Wait while the request is processed. It could take a few minutes.



6. Confirm that the Microsoft Purview Governance Status shows `Governed`. Note that **it could take a few minutes** after you enable *Data policy enforcement* in Microsoft Purview for the correct status to be reflected.

> **Note**
>
> If you disable Data policy enforcement for this Azure SQL Database data source, it might take up to 24 hours for the Microsoft Purview Governance Status to be updated automatically to `Not Governed`. This can be accelerated by selecting **Check for Microsoft Purview Governance**. Before you enable *Data policy enforcement* for the data source in another Microsoft Purview account, ensure that the Purview Governance Status shows as `Not Governed`. Then repeat the steps above with the new Microsoft Purview account.

Create a new DevOps policy

Follow this link for the steps to create a new DevOps policy in Microsoft Purview.

List DevOps policies

Follow this link for the steps to list DevOps policies in Microsoft Purview.

Update a DevOps policy

Follow this link for the steps to update a DevOps policies in Microsoft Purview.

Delete a DevOps policy

Follow this link for the steps to delete a DevOps policies in Microsoft Purview.

> **Important**

DevOps policies are auto-published and changes can take up to **5 minutes** to be enforced by the data source.

Test the DevOps policy

See how to test the policy you created.

Role definition detail

See the mapping of DevOps role to data source actions.

Next steps

See related videos, blogs, and documents.

Feedback

Was this page helpful? Yes No

Provision access to system metadata in Azure SQL Managed Instance using Microsoft Purview DevOps policies

Article • 10/28/2024

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

DevOps policies are a type of Microsoft Purview access policies. They allow you to manage access to system metadata on data sources that have been registered for *data policy enforcement* in Microsoft Purview. These policies are configured directly from the Microsoft Purview governance portal, and after they're saved, they get automatically published and then enforced by the data source. Microsoft Purview policies only manage access for Microsoft Entra principals.

This how-to guide covers how to configure Azure SQL Managed Instance to enforce policies created in Microsoft Purview. It covers the configuration steps for Azure SQL MI and the ones in Microsoft Purview to provision access to Azure SQL MI system metadata (DMVs and DMFs) using the DevOps policies actions *SQL Performance Monitoring* or *SQL Security Auditing*.

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- A new or existing Microsoft Purview account. Follow this quickstart guide to create one.
- Create a new Azure SQL MI or use an existing one in one of the currently available regions for this feature. You can follow this guide to create a new Azure SQL MI.

Region support

All Microsoft Purview regions are supported.

Enforcement of Microsoft Purview policies is available only in the following regions for Azure SQL MI:

Public cloud:

- East US
- East US2
- South Central US
- West Central US
- West US3
- Canada Central
- Brazil South
- West Europe
- North Europe
- France Central
- UK South
- South Africa North
- Central India
- South East Asia
- East Asia
- Australia East

Azure SQL MI configuration

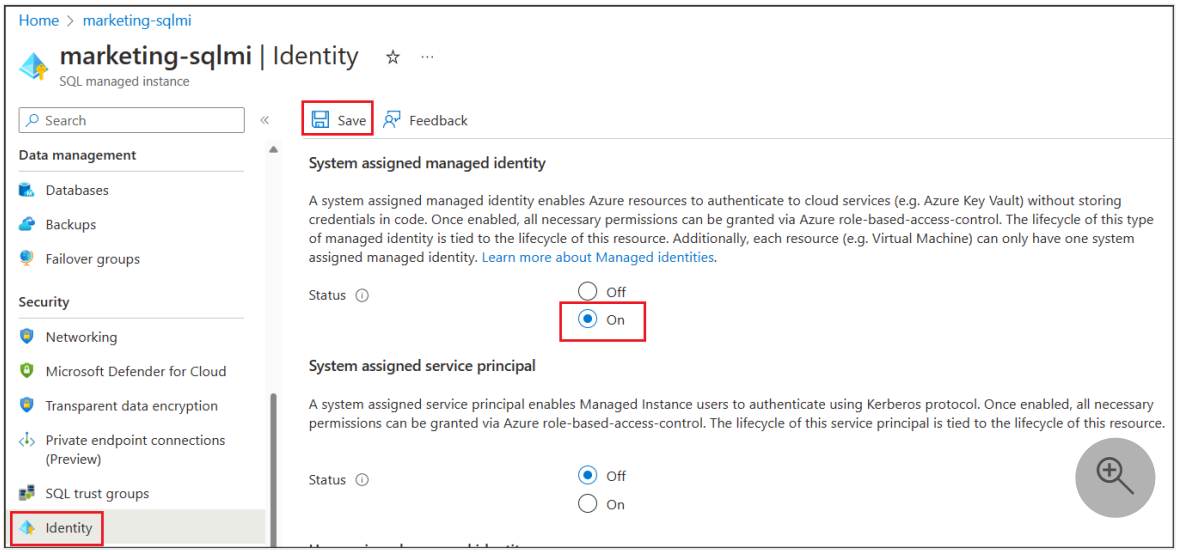
This section explains how you can set up Azure SQL MI to honor policies from Microsoft Purview. Check first whether Azure SQL MI is configured for public or private endpoint. This guide explains how to do so.

Configuration for SQL MI public endpoint

If Azure SQL MI is configured for public endpoint, follow these steps:

- Configure a Microsoft Entra Admin. In Azure portal navigate to the Azure SQL MI then navigate to Microsoft Entra ID on the side menu (formerly called Active Directory admin). Set an Admin name and then select **Save**.

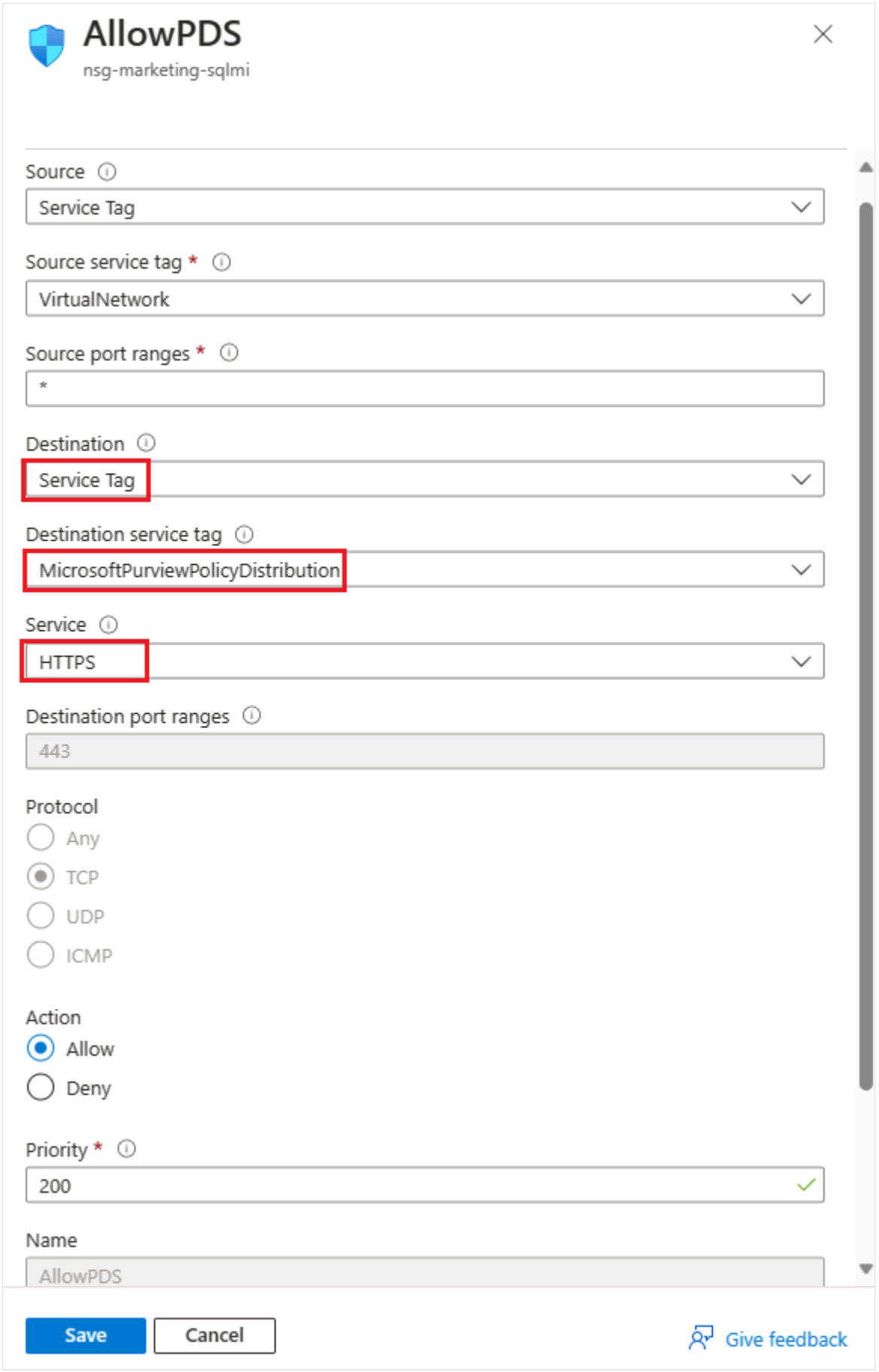
- Then navigate to Identity on the side menu. Under System assigned managed identity check status to *On* and then select **Save**. See screenshot:



Configuration for SQL MI private endpoint

If Azure SQL MI is configured to use private endpoint, execute the same steps outlined in the configuration for public endpoint, and in addition, do the following:

- Navigate to the Network Security Group (NSG) associated to your Azure SQL MI.
- Add an outbound security rule similar to the one in the following screenshot. *Destination = Service Tag, Destination service tag = MicrosoftPurviewPolicyDistribution, Service = HTTPS, Action = Allow.* Ensure also that this rule's priority is lower than the one for the *deny_all_outbound* rule.



Microsoft Purview Configuration

Register the data source in Microsoft Purview

Before a policy can be created in Microsoft Purview for a data resource, you must register that data resource in Microsoft Purview Studio. You will find the instructions related to registering the data resource later in this guide.

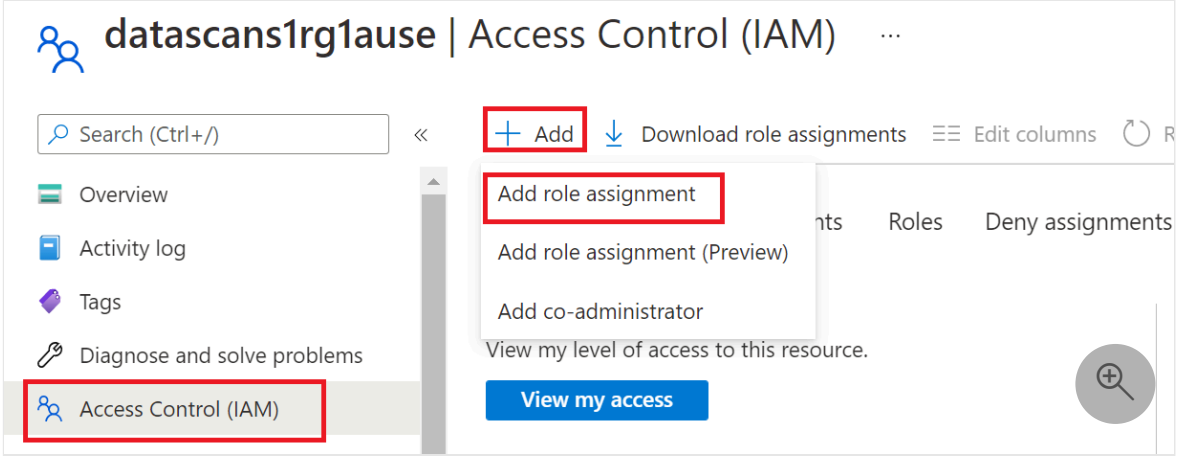
> **Note**
>
> Microsoft Purview policies rely on the data resource ARM path. If a data resource is moved to a new resource group or subscription it will need to be de-registered and then registered again in Microsoft Purview.

Configure permissions to enable Data policy enforcement on the data source

Once a resource is registered, but before a policy can be created in Microsoft Purview for that resource, you must configure permissions. A set of permissions are needed to enable the **Data policy enforcement**. This applies to data sources, resource groups, or subscriptions. To enable *Data policy enforcement*, you must have *both* specific Identity and Access Management (IAM) privileges on the resource as well as specific Microsoft Purview privileges:

- You must have *either one* of the following IAM role combinations on the resource's Azure Resource Manager path or any parent of it (that is, using IAM permission inheritance):
  - IAM Owner
  - Both IAM Contributor and IAM User Access Administrator

  To configure Azure role-based access control (RBAC) permissions, follow this guide. The following screenshot shows how to access the **Access Control** section in the Azure portal for the data resource to add a role assignment.

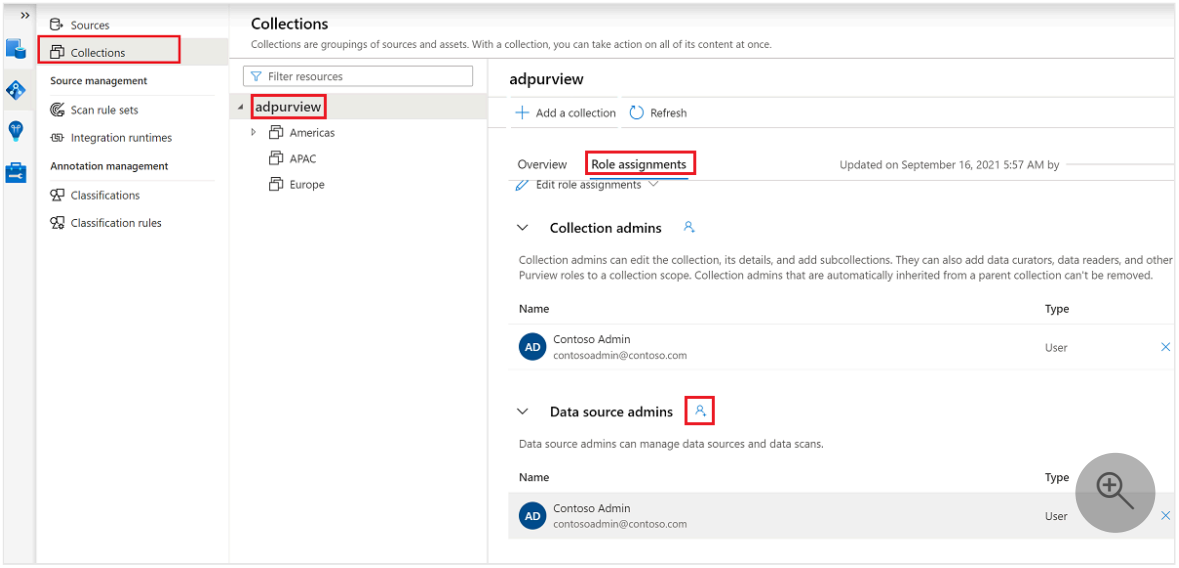


> **Note**

> The *IAM Owner* role for a data resource can be inherited from a parent resource group, a subscription, or a subscription management group. Check which Microsoft Entra users, groups, and service principals hold or are inheriting the *IAM Owner* role for the resource.

- You also need to have the Microsoft Purview *Data source admin* role for the collection or a parent collection (if inheritance is enabled). For more information, see the guide on managing Microsoft Purview role assignments.

  The following screenshot shows how to assign the *Data source admin* role at the root collection level.



Configure Microsoft Purview permissions to create, update, or delete access policies

To create, update or delete policies, you need to get the *Policy author* role in Microsoft Purview at *root collection level*:

- The *Policy author* role can create, update, and delete DevOps and Data Owner policies.
- The *Policy author* role can delete self-service access policies.

For more information about managing Microsoft Purview role assignments, see Create and manage collections in the Microsoft Purview Data Map.

> **Note**
>
> Policy author role must be configured at the root collection level.

In addition, to easily search Microsoft Entra users or groups when creating or updating the subject of a policy, you can greatly benefit from getting the Directory Readers permission in Microsoft Entra ID. This is a common permission for users in an Azure tenant. Without the Directory Reader permission, the Policy Author will have to type the complete username or email for all the principals included in the subject of a data policy.

Configure Microsoft Purview permissions for publishing Data Owner policies

Data Owner policies allow for checks and balances if you assign the Microsoft Purview *Policy author* and *Data source admin* roles to different people in the organization. Before a Data owner policy takes effect, a second person (*Data source admin*) must review it and explicitly approve it by publishing it. This does not apply to DevOps or Self-service access policies as publishing is automatic for them when those policies are created or updated.

To publish a Data owner policy you need to get the *Data source admin* role in Microsoft Purview at *root collection level*.

For more information about managing Microsoft Purview role assignments, see Create and manage collections in the Microsoft Purview Data Map.

> ⓘ **Note**
>
> To publish Data owner policies, the Data source admin role must be configured at the root collection level.

Delegate access provisioning responsibility to roles in Microsoft Purview

After a resource has been enabled for **Data policy enforcement**, any Microsoft Purview user with the *Policy author* role at the root collection level can provision access to that data source from Microsoft Purview.

> ⓘ **Note**
>
> Any Microsoft Purview root *Collection admin* can assign new users to root *Policy author* roles. Any *Collection admin* can assign new users to a *Data source admin*

role under the collection. Minimize and carefully vet the users who hold Microsoft Purview *Collection admin*, *Data source admin*, or *Policy author* roles.

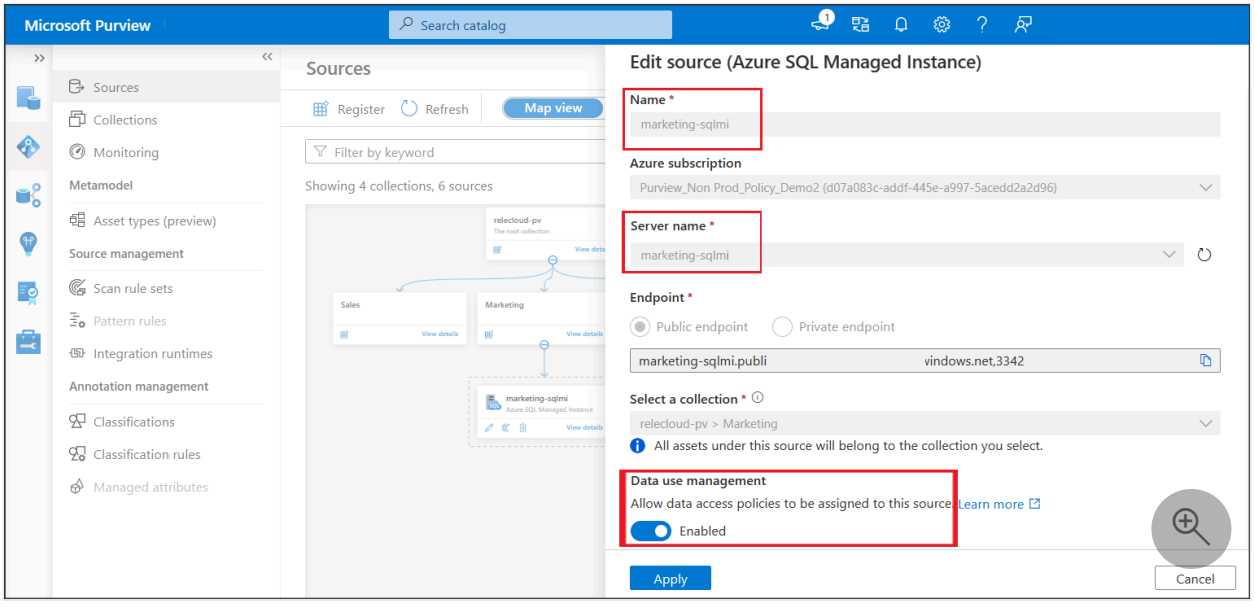
If a Microsoft Purview account with published policies is deleted, such policies will stop being enforced within an amount of time that depends on the specific data source. This change can have implications on both security and data access availability. The Contributor and Owner roles in IAM can delete Microsoft Purview accounts. You can check these permissions by going to the **Access control (IAM)** section for your Microsoft Purview account and selecting **Role Assignments**. You can also use a lock to prevent the Microsoft Purview account from being deleted through Resource Manager locks.

Register the data sources in Microsoft Purview

The Azure SQL Managed Instance data source needs to be registered first with Microsoft Purview, before you can create access policies. You can follow the "Prerequisites" and "Register the data source" sections in this guide: Register and scan Azure SQL MI.

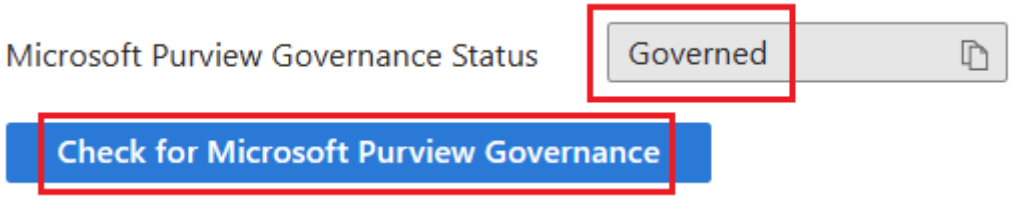
After you've registered your resources, you'll need to enable data policy enforcement (formerly Data Use Management). Data policy enforcement needs certain permissions and can affect the security of your data, as it delegates to certain Microsoft Purview roles the ability to manage access to data sources. **Go through the secure practices related to data policy enforcement in this guide**: How to enable policy enforcement

Once your data source has the **Data policy enforcement** toggle *Enabled*, it will look like this screenshot. This will enable the access policies to be used with the given data source.



Return to the Azure portal for Azure SQL Database to verify it's now governed by Microsoft Purview

1. Sign in to the Azure portal through this link
2. Select the Azure SQL Managed Instance that you want to configure.
3. Go to **Microsoft Entra ID admin** on the left pane.
4. Scroll down to **Microsoft Purview access policies**.
5. Select the button to **Check for Microsoft Purview Governance**. Wait while the request is processed. It could take a few minutes.



6. Confirm that the Microsoft Purview Governance Status shows `Governed`. Note that **it could take a few minutes** after you enable *data policy enforcement* in Microsoft Purview for the correct status to be reflected.

> **Note**
>
> If you disable data policy enforcement for this Azure SQL MI in Microsoft Purview, select **Check for Microsoft Purview Governance** for the Microsoft Purview Governance Status to be updated to `Not Governed`. This step must be done for the impacted Azure SQL MI each time data policy enforcement status is changed for an existing or new access policy in Microsoft Purview. Before you enable *data policy enforcement* for the data source in another Microsoft Purview account, ensure that the Purview Governance Status shows as `Not Governed`. Then repeat the steps above with the new Microsoft Purview account.

Create a new DevOps policy

Follow this link for the steps to create a new DevOps policy in Microsoft Purview.

List DevOps policies

Follow this link for the steps to list DevOps policies in Microsoft Purview.

Update a DevOps policy

Follow this link for the steps to update a DevOps policies in Microsoft Purview.

Delete a DevOps policy

Follow this link for the steps to delete a DevOps policies in Microsoft Purview.

> **Important**
>
> DevOps policies are auto-published and changes can take up to **5 minutes** to be enforced by the data source.

Test the DevOps policy

See how to test the policy you created.

Role definition detail

See the mapping of DevOps role to data source actions.

Next steps

See related videos, blogs, and documents.

Feedback

Was this page helpful? Yes No

Provision access to system metadata in SQL Server 2022 (Azure Arc-enabled) using Microsoft Purview DevOps policies

Article • 10/28/2024

[Dev> [!NOTE] > Microsoft Purview Data Map (classic), Data Catalog (classic), Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

Ops policies](../concept-policies-devops.md) are a type of Microsoft Purview access policies. They allow you to manage access to system metadata on data sources that have been registered for *Data policy enforcement* in Microsoft Purview. These policies are configured directly from the Microsoft Purview governance portal, and after they're saved, they get automatically published and then enforced by the data source. Microsoft Purview policies only manage access for Microsoft Entra principals.

This how-to guide covers how to configure SQL Server 2022 to enforce policies created in Microsoft Purview. It covers onboarding with Azure Arc, enabling Microsoft Entra ID on the SQL Server, and provisioning access to its system metadata (DMVs and DMFs) using the DevOps policies actions *SQL Performance Monitoring* or *SQL Security Auditing*.

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- A new or existing Microsoft Purview account. Follow this quickstart guide to create one.
- Get SQL Server on-premises version 2022 and install it. Versions 2022 or more recent are supported on both Windows and Linux. You can try the free Developer edition.
- Configure your permissions and then register a list of resource providers in the subscription you'll use to onboard the SQL Server instance to Azure Arc.
- Complete prerequisites and onboard the SQL Server instance with Azure Arc. A simple configuration for Windows SQL Server is here. An alternate configuration for Linux SQL Server is here.
- Enable Microsoft Entra authentication in SQL Server. For a simpler setup, complete the prerequisites and process outlined in this article.

- Remember to grant application permissions and grant admin consent
- You need to configure a Microsoft Entra admin for the SQL Server instance, but it isn't necessary to set up other Microsoft Entra logins or users. You'll grant access to those users by using Microsoft Purview policies.

Region support

Data policy enforcement is available in all Microsoft Purview regions except:

- West US2
- East Asia
- US Gov Virginia
- China North 3

Security considerations for Azure Arc-enabled SQL Server

- The server admin can turn off the Microsoft Purview policy enforcement.
- Azure Arc admin and server admin permissions provide the ability to change the Azure Resource Manager path of the server. Because mappings in Microsoft Purview use Resource Manager paths, it can lead to wrong policy enforcements.
- A SQL Server admin (database admin) can gain the power of a server admin and can tamper with the cached policies from Microsoft Purview.
- The recommended configuration is to create a separate app registration for each SQL server instance. This configuration prevents the second SQL Server instance from reading the policies meant for the first SQL Server instance, in case a rogue admin in the second SQL Server instance tampers with the Resource Manager path.

Verify the prerequisites

1. Sign in to the Azure portal through this link
2. Navigate to **SQL servers** on the left pane. You'll see a list of SQL Server instances on Azure Arc.
3. Select the SQL Server instance that you want to configure.
4. Go to **Microsoft Entra ID** on the left pane.
5. Ensure that Microsoft Entra authentication is configured with an admin login. If not, refer to the access policy prerequisites section in this guide.

6. Ensure that a certificate has been provided to for SQL Server to authenticate to Azure. If not, refer to the access policy prerequisites section in this guide.

7. Ensure that an app registration has been entered to create a trust relationship between SQL Server and Microsoft Entra ID. If not, refer to the access policy prerequisites section in this guide.

8. If you made any changes, select the **Save** button to save the configuration and wait until the operation completes successfully. This could take a few minutes. The message "*Saved successfully*" will be displayed at the top of the page in green background. You might need to scroll up to see it.

Microsoft Purview configuration

Register the data source in Microsoft Purview

Before a policy can be created in Microsoft Purview for a data resource, you must register that data resource in Microsoft Purview Studio. You will find the instructions related to registering the data resource later in this guide.

> ⓘ **Note**
>
> Microsoft Purview policies rely on the data resource ARM path. If a data resource is moved to a new resource group or subscription it will need to be de-registered and then registered again in Microsoft Purview.

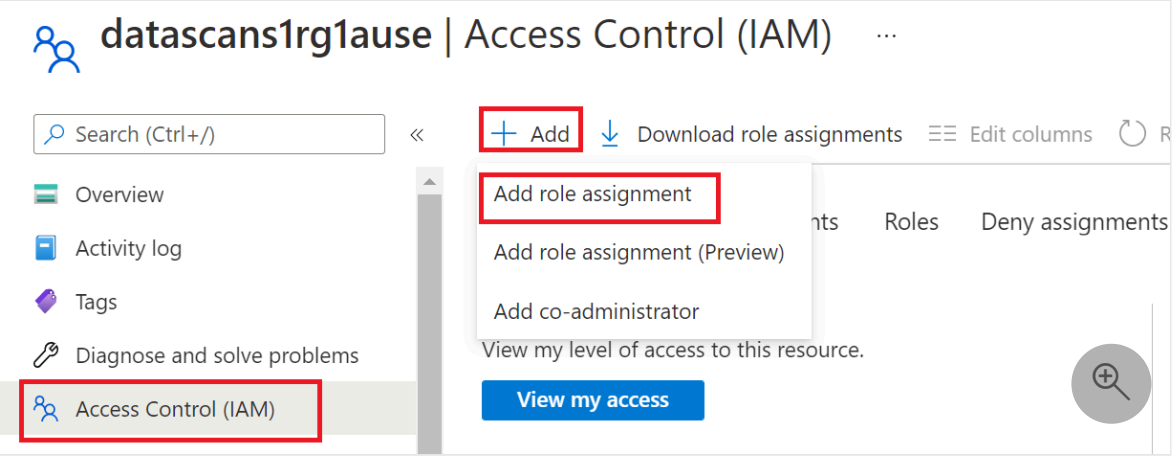
Configure permissions to enable Data policy enforcement on the data source

Once a resource is registered, but before a policy can be created in Microsoft Purview for that resource, you must configure permissions. A set of permissions are needed to enable the **Data policy enforcement**. This applies to data sources, resource groups, or subscriptions. To enable *Data policy enforcement*, you must have *both* specific Identity and Access Management (IAM) privileges on the resource as well as specific Microsoft Purview privileges:

- You must have *either one* of the following IAM role combinations on the resource's Azure Resource Manager path or any parent of it (that is, using IAM permission inheritance):
  - IAM Owner

- Both IAM Contributor and IAM User Access Administrator

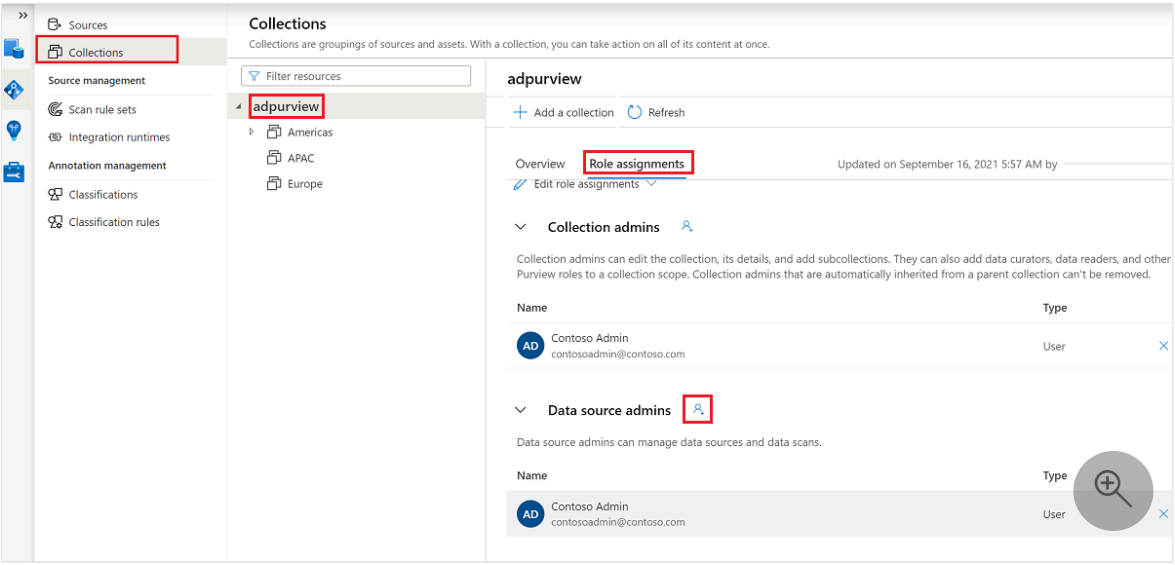
To configure Azure role-based access control (RBAC) permissions, follow this guide. The following screenshot shows how to access the **Access Control** section in the Azure portal for the data resource to add a role assignment.



> **Note**
>
> The *IAM Owner* role for a data resource can be inherited from a parent resource group, a subscription, or a subscription management group. Check which Microsoft Entra users, groups, and service principals hold or are inheriting the *IAM Owner* role for the resource.

- You also need to have the Microsoft Purview *Data source admin* role for the collection or a parent collection (if inheritance is enabled). For more information, see the guide on managing Microsoft Purview role assignments.

  The following screenshot shows how to assign the *Data source admin* role at the root collection level.



Configure Microsoft Purview permissions to create, update, or delete access policies

To create, update or delete policies, you need to get the *Policy author* role in Microsoft Purview at *root collection level*:

- The *Policy author* role can create, update, and delete DevOps and Data Owner policies.
- The *Policy author* role can delete self-service access policies.

For more information about managing Microsoft Purview role assignments, see Create and manage collections in the Microsoft Purview Data Map.

> **Note**
>
> Policy author role must be configured at the root collection level.

In addition, to easily search Microsoft Entra users or groups when creating or updating the subject of a policy, you can greatly benefit from getting the Directory Readers permission in Microsoft Entra ID. This is a common permission for users in an Azure tenant. Without the Directory Reader permission, the Policy Author will have to type the complete username or email for all the principals included in the subject of a data policy.

Configure Microsoft Purview permissions for publishing Data Owner policies

Data Owner policies allow for checks and balances if you assign the Microsoft Purview *Policy author* and *Data source admin* roles to different people in the organization. Before a Data owner policy takes effect, a second person (*Data source admin*) must review it and explicitly approve it by publishing it. This does not apply to DevOps or Self-service access policies as publishing is automatic for them when those policies are created or updated.

To publish a Data owner policy you need to get the *Data source admin* role in Microsoft Purview at *root collection level*.

For more information about managing Microsoft Purview role assignments, see Create and manage collections in the Microsoft Purview Data Map.

> **Note**

To publish Data owner policies, the Data source admin role must be configured at the root collection level.

Delegate access provisioning responsibility to roles in Microsoft Purview

After a resource has been enabled for **Data policy enforcement**, any Microsoft Purview user with the *Policy author* role at the root collection level can provision access to that data source from Microsoft Purview.

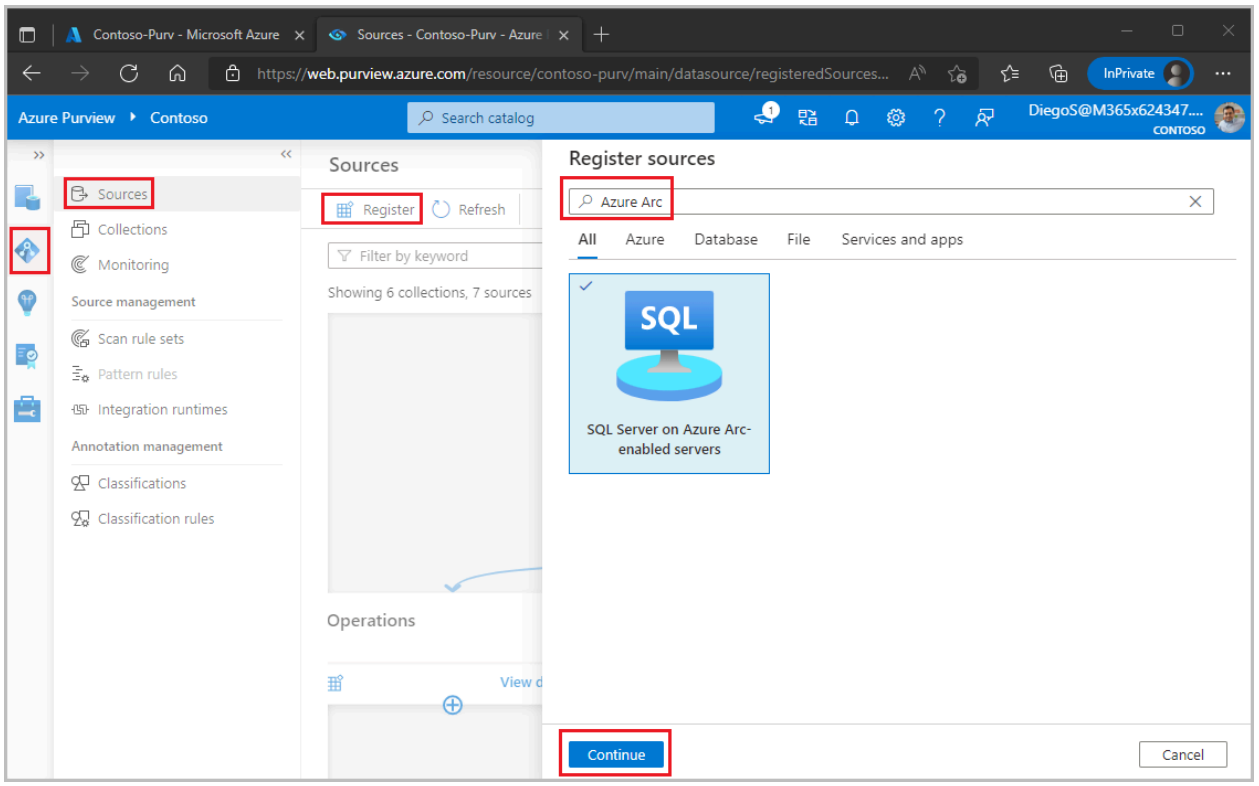
> **Note**
>
> Any Microsoft Purview root *Collection admin* can assign new users to root *Policy author* roles. Any *Collection admin* can assign new users to a *Data source admin* role under the collection. Minimize and carefully vet the users who hold Microsoft Purview *Collection admin*, *Data source admin*, or *Policy author* roles.

If a Microsoft Purview account with published policies is deleted, such policies will stop being enforced within an amount of time that depends on the specific data source. This change can have implications on both security and data access availability. The Contributor and Owner roles in IAM can delete Microsoft Purview accounts. You can check these permissions by going to the **Access control (IAM)** section for your Microsoft Purview account and selecting **Role Assignments**. You can also use a lock to prevent the Microsoft Purview account from being deleted through Resource Manager locks.

Register data sources in Microsoft Purview

The Azure Arc-enabled SQL Server data source needs to be registered first with Microsoft Purview, before policies can be created.

1. Sign in to Microsoft Purview Studio.
2. Navigate to the **Data map** feature on the left pane, select **Sources**, then select **Register**. Type "Azure Arc" in the search box and select **SQL Server on Azure Arc**. Then select **Continue**.



1. Enter a **Name** for this registration. It's best practice to make the name of the registration the same as the server name in the next step.

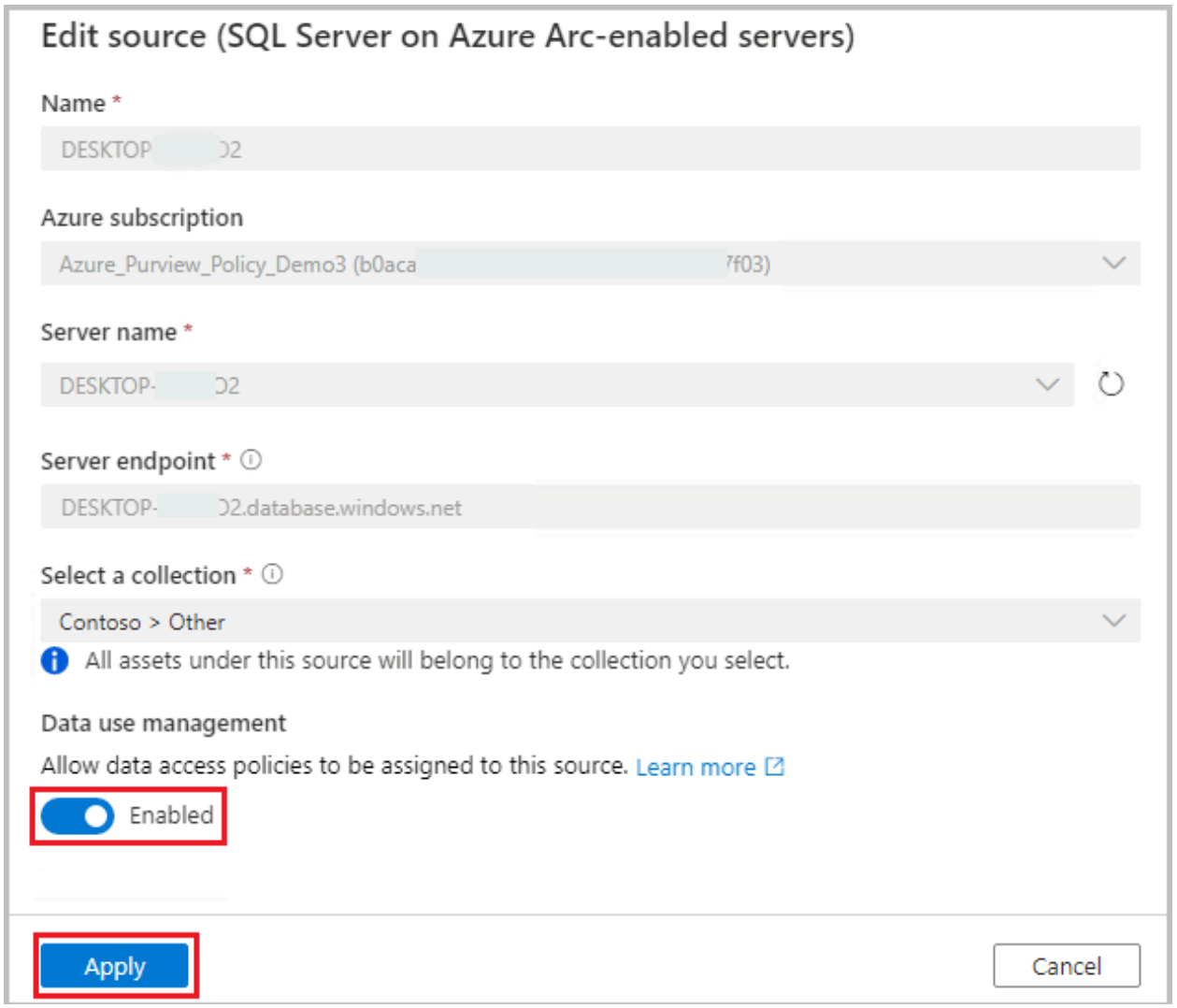
2. select an **Azure subscription**, **Server name**, and **Server endpoint**.

3. **Select a collection** to put this registration in.

4. Enable Data Policy Enforcement. Data Policy Enforcement needs certain permissions and can affect the security of your data, as it delegates to certain Microsoft Purview roles to manage access to the data sources. **Go through the secure practices related to Data Policy Enforcement in this guide**: How to enable Data Policy Enforcement

5. Select **Register** or **Apply** at the bottom.

Once your data source has the **Data policy enforcement** toggle *Enabled*, it will look like this picture.

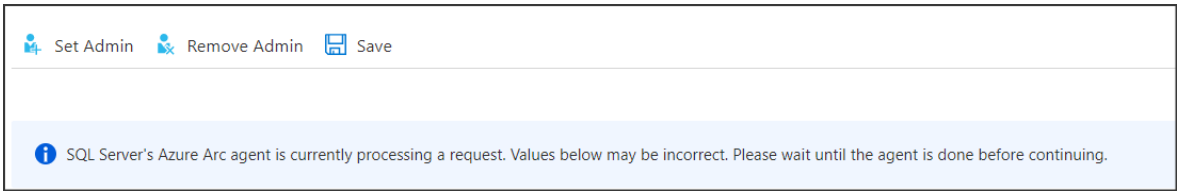


Enable policies in Azure Arc-enabled SQL Server

This section describes the steps to configure SQL Server on Azure Arc to use Microsoft Purview. Execute these steps after you enable the *Data policy enforcement* option for this data source in the Microsoft Purview account.

1. Sign in to the Azure portal through this link
2. Navigate to **SQL servers** on the left pane. You'll see a list of SQL Server instances on Azure Arc.
3. Select the SQL Server instance that you want to configure.
4. Go to **Microsoft Entra ID** on the left pane.
5. Scroll down to **Microsoft Purview access policies**.
6. Select the button to **Check for Microsoft Purview Governance**. Wait while the request is processed. While that happens, this message will be displayed at the top

of the page. You might need to scroll up to see it.



7. At the bottom of the page, confirm that the Microsoft Purview Governance Status shows `Governed`. Note that **it may take up to 30 minutes** for the correct status to be reflected. Continue doing a browser refresh until that happens.

8. Confirm that the Microsoft Purview Endpoint points to the Microsoft Purview account where you registered this data source and enabled the *data policy enforcement*

Create a new DevOps policy

Follow this link for the steps to create a new DevOps policy in Microsoft Purview.

List DevOps policies

Follow this link for the steps to list DevOps policies in Microsoft Purview.

Update a DevOps policy

Follow this link for the steps to update a DevOps policies in Microsoft Purview.

Delete a DevOps policy

Follow this link for the steps to delete a DevOps policies in Microsoft Purview.

> **Important**
>
> DevOps policies are auto-published and changes can take up to **5 minutes** to be enforced by the data source.

Test the DevOps policy

See how to test the policy you created.

Role definition detail

See the mapping of DevOps role to data source actions.

Next steps

See related videos, blogs, and documents.

Feedback

Was this page helpful? Yes No

Provision access to system metadata on entire resource groups or subscriptions using Microsoft Purview DevOps policies

Article • 10/28/2024

> **Note**
>
> Microsoft Purview Data Map (classic), Data Catalog (classic), Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

DevOps policies are a type of Microsoft Purview access policies. They allow you to manage access to system metadata on data sources that have been registered for *Data policy enforcement* in Microsoft Purview. These policies are configured directly in the Microsoft Purview governance portal, and after they're saved, they get automatically published and then enforced by the data source. Microsoft Purview policies only manage access for Microsoft Entra principals.

This how-to guide covers how to register an entire resource group or subscription and then create a single policy that will provision access to **all** data sources in that resource group or subscription. That single policy will cover all existing data sources and any data sources that are created afterwards. And provisioning access to its system metadata (DMVs and DMFs) using the DevOps policies actions *SQL Performance Monitoring* or *SQL Security Auditing*.

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- A new or existing Microsoft Purview account. Follow this quickstart guide to create one.

Only these data sources are enabled for access policies on resource group or subscription. Follow the **Prerequisites** section that is specific to the data sources in these guides:

- DevOps policies on an Azure SQL Database

- DevOps policies on an Azure Arc-enabled SQL Server

Microsoft Purview Configuration

Register the data source in Microsoft Purview

Before a policy can be created in Microsoft Purview for a data resource, you must register that data resource in Microsoft Purview Studio. You will find the instructions related to registering the data resource later in this guide.

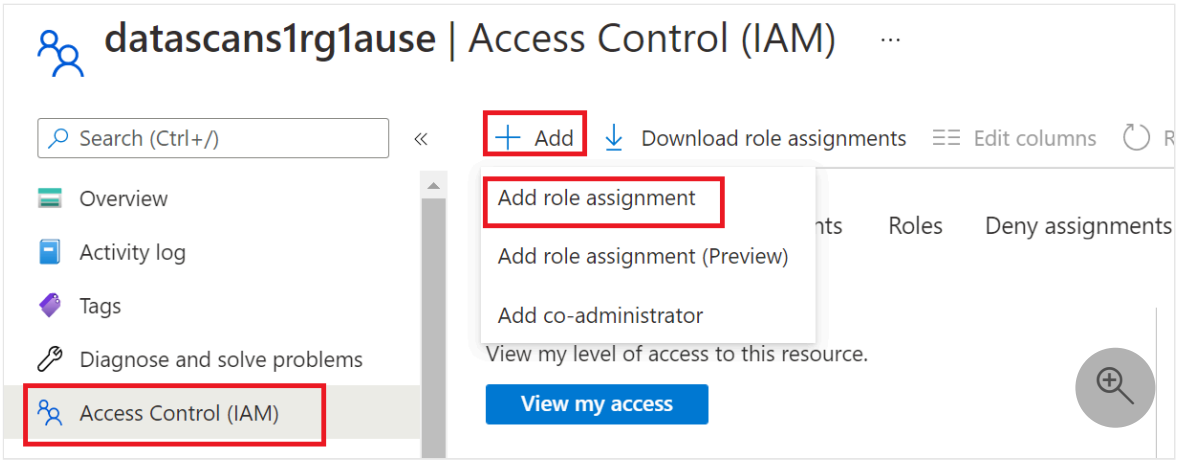
> ⓘ **Note**
>
> Microsoft Purview policies rely on the data resource ARM path. If a data resource is moved to a new resource group or subscription it will need to be de-registered and then registered again in Microsoft Purview.

Configure permissions to enable Data policy enforcement on the data source

Once a resource is registered, but before a policy can be created in Microsoft Purview for that resource, you must configure permissions. A set of permissions are needed to enable the **Data policy enforcement**. This applies to data sources, resource groups, or subscriptions. To enable *Data policy enforcement*, you must have *both* specific Identity and Access Management (IAM) privileges on the resource as well as specific Microsoft Purview privileges:

- You must have *either one* of the following IAM role combinations on the resource's Azure Resource Manager path or any parent of it (that is, using IAM permission inheritance):
  - IAM Owner
  - Both IAM Contributor and IAM User Access Administrator

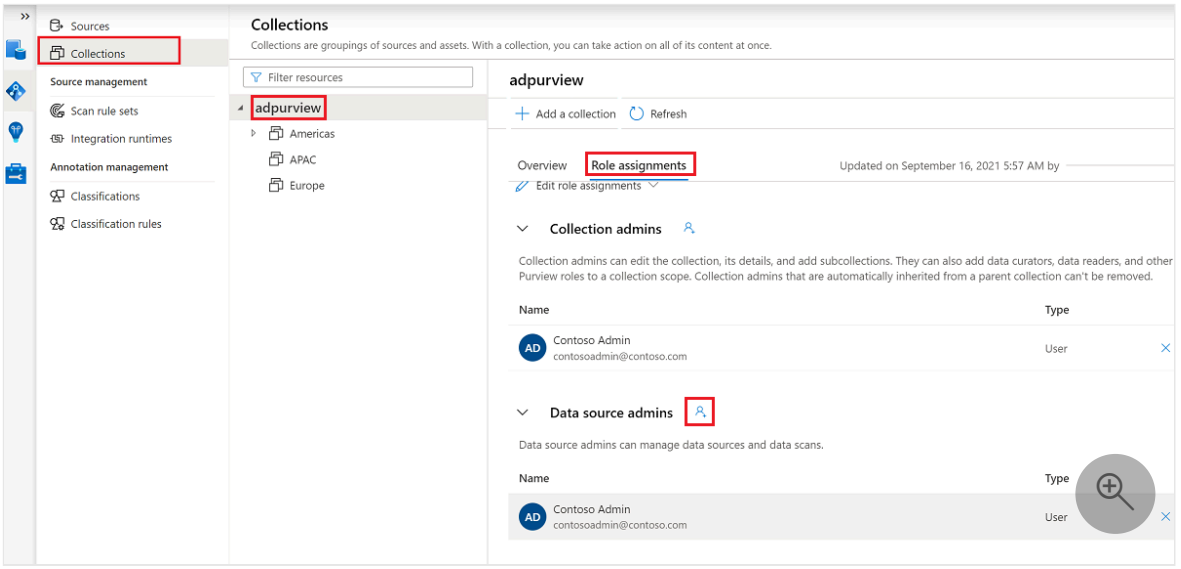
  To configure Azure role-based access control (RBAC) permissions, follow this guide. The following screenshot shows how to access the **Access Control** section in the Azure portal for the data resource to add a role assignment.



> **Note**
>
> The *IAM Owner* role for a data resource can be inherited from a parent resource group, a subscription, or a subscription management group. Check which Microsoft Entra users, groups, and service principals hold or are inheriting the *IAM Owner* role for the resource.

- You also need to have the Microsoft Purview *Data source admin* role for the collection or a parent collection (if inheritance is enabled). For more information, see the guide on managing Microsoft Purview role assignments.

  The following screenshot shows how to assign the *Data source admin* role at the root collection level.



Configure Microsoft Purview permissions to create, update, or delete access policies

To create, update or delete policies, you need to get the *Policy author* role in Microsoft Purview at *root collection level*:

- The *Policy author* role can create, update, and delete DevOps and Data Owner policies.
- The *Policy author* role can delete self-service access policies.

For more information about managing Microsoft Purview role assignments, see Create and manage collections in the Microsoft Purview Data Map.

> **Note**
>
> Policy author role must be configured at the root collection level.

In addition, to easily search Microsoft Entra users or groups when creating or updating the subject of a policy, you can greatly benefit from getting the Directory Readers permission in Microsoft Entra ID. This is a common permission for users in an Azure tenant. Without the Directory Reader permission, the Policy Author will have to type the complete username or email for all the principals included in the subject of a data policy.

Configure Microsoft Purview permissions for publishing Data Owner policies

Data Owner policies allow for checks and balances if you assign the Microsoft Purview *Policy author* and *Data source admin* roles to different people in the organization. Before a Data owner policy takes effect, a second person (*Data source admin*) must review it and explicitly approve it by publishing it. This does not apply to DevOps or Self-service access policies as publishing is automatic for them when those policies are created or updated.

To publish a Data owner policy you need to get the *Data source admin* role in Microsoft Purview at *root collection level*.

For more information about managing Microsoft Purview role assignments, see Create and manage collections in the Microsoft Purview Data Map.

> **Note**
>
> To publish Data owner policies, the Data source admin role must be configured at the root collection level.

Delegate access provisioning responsibility to roles in Microsoft Purview

After a resource has been enabled for **Data policy enforcement**, any Microsoft Purview user with the *Policy author* role at the root collection level can provision access to that data source from Microsoft Purview.

> **Note**
>
> Any Microsoft Purview root *Collection admin* can assign new users to root *Policy author* roles. Any *Collection admin* can assign new users to a *Data source admin* role under the collection. Minimize and carefully vet the users who hold Microsoft Purview *Collection admin*, *Data source admin*, or *Policy author* roles.

If a Microsoft Purview account with published policies is deleted, such policies will stop being enforced within an amount of time that depends on the specific data source. This change can have implications on both security and data access availability. The Contributor and Owner roles in IAM can delete Microsoft Purview accounts. You can check these permissions by going to the **Access control (IAM)** section for your Microsoft Purview account and selecting **Role Assignments**. You can also use a lock to prevent the Microsoft Purview account from being deleted through Resource Manager locks.

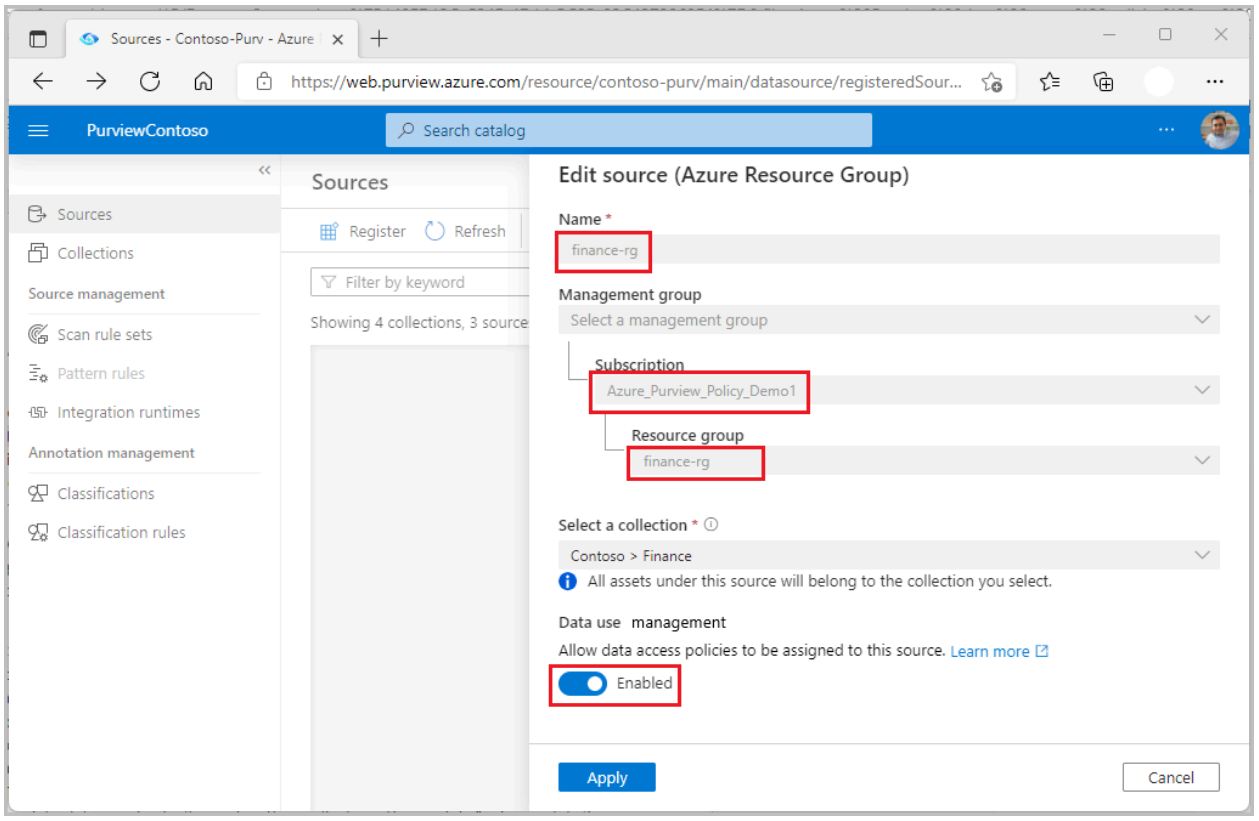
Register the subscription or resource group for Data Policy Enforcement

The subscription or resource group needs to be registered with Microsoft Purview before you can create access policies. To register your subscription or resource group, follow the **Prerequisites** and **Register** sections of this guide:

- Register multiple sources in Microsoft Purview

After you've registered your resources, you'll need to enable the Data Policy Enforcement option. Data Policy Enforcement needs certain permissions and can affect the security of your data, as it delegates to certain Microsoft Purview roles to manage access to the data sources. **Go through the secure practices related to Data Policy Enforcement in this guide**: How to enable Data Policy Enforcement

In the end, your resource will have the **Data Policy Enforcement** toggle **Enabled**, as shown in the screenshot:



Create a new DevOps policy

Follow this link for the steps to create a new DevOps policy in Microsoft Purview.

List DevOps policies

Follow this link for the steps to list DevOps policies in Microsoft Purview.

Update a DevOps policy

Follow this link for the steps to update a DevOps policies in Microsoft Purview.

Delete a DevOps policy

Follow this link for the steps to delete a DevOps policies in Microsoft Purview.

Test the DevOps policy

See how to test the policy you created.

Role definition detail

See the mapping of DevOps role to data source actions.

Next steps

See related videos, blogs, and documents.

Feedback

Was this page helpful? Yes No

Tutorial: Access provisioning by data owner to Azure Storage datasets (preview)

Article • 02/26/2025

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

> **Important**
>
> This feature is currently in preview. The Supplemental Terms of Use for Microsoft Azure Previews include additional legal terms that apply to Azure features that are in beta, in preview, or otherwise not yet released into general availability.

Policies in Microsoft Purview allow you to enable access to data sources that have been registered to a collection. This tutorial describes how a data owner can use Microsoft Purview to enable access to datasets in Azure Storage through Microsoft Purview.

In this tutorial, you learn how to:

- ✓ Prepare your Azure environment
- ✓ Configure permissions to allow Microsoft Purview to connect to your resources
- ✓ Register your Azure Storage resource for Data Policy Enforcement
- ✓ Create and publish a policy for your resource group or subscription

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- A new or existing Microsoft Purview account. Follow this quickstart guide to create one.

Region support

- All Microsoft Purview regions are supported.

- Storage accounts in the following regions are supported without the need for additional configuration. However, zone-redundant storage (ZRS) accounts are not supported.
  - Australia Central
  - Australia East
  - Australia Southeast
  - Brazil South
  - Canada Central
  - Canada East
  - Central India
  - Central US
  - East Asia
  - East US 2
  - East US
  - France Central
  - Germany West Central
  - Japan East
  - Japan West
  - Korea Central
  - North Central US
  - North Europe
  - Norway East
  - Poland Central
  - Qatar Central
  - South Central US
  - South Africa North
  - Southeast Asia
  - South India
  - Sweden Central
  - Switzerland North
  - West Central US
  - West Europe
  - West US
  - West US 2
  - West US 3
  - UAE North
  - UK South
  - UK West
- Storage accounts in other regions in Public Cloud are supported after setting feature flag *AllowPurviewPolicyEnforcement*, as outlined in the next section. Newly

created ZRS Storage accounts are supported, if created after setting the feature flag *AllowPurviewPolicyEnforcement*.

If needed, you can create a new Storage account by following this guide.

Configure the subscription where the Azure Storage account resides for policies from Microsoft Purview

This step is only necessary in certain regions (see prior section). To enable Microsoft Purview to manage policies for one or more Azure Storage accounts, execute the following PowerShell commands in the subscription where you'll deploy your Azure Storage account. These PowerShell commands will enable Microsoft Purview to manage policies on all Azure Storage accounts in that subscription.

If you're executing these commands locally, be sure to run PowerShell as an administrator. Alternatively, you can use the Azure Cloud Shell in the Azure portal: https://shell.azure.com.

PowerShell

```
# Install the Az module
Install-Module -Name Az -Scope CurrentUser -Repository PSGallery -Force
# Login into the subscription
Connect-AzAccount -Subscription <SubscriptionID>
# Register the feature
Register-AzProviderFeature -FeatureName AllowPurviewPolicyEnforcement -
ProviderNamespace Microsoft.Storage
```

If the output of the last command shows *RegistrationState* as *Registered*, then your subscription is enabled for access policies. If the output is *Registering*, wait at least 10 minutes, and then retry the command. **Do not continue unless the RegistrationState shows as *Registered***.

Configuration

Register the data source in Microsoft Purview

Before a policy can be created in Microsoft Purview for a data resource, you must register that data resource in Microsoft Purview Studio. You will find the instructions related to registering the data resource later in this guide.

Note

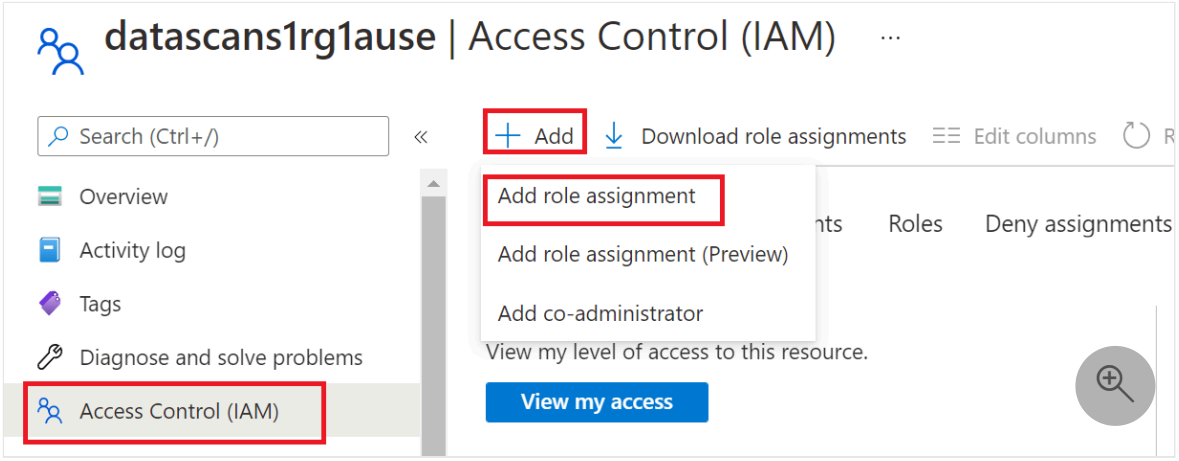
Microsoft Purview policies rely on the data resource ARM path. If a data resource is moved to a new resource group or subscription it will need to be de-registered and then registered again in Microsoft Purview.

Configure permissions to enable Data policy enforcement on the data source

Once a resource is registered, but before a policy can be created in Microsoft Purview for that resource, you must configure permissions. A set of permissions are needed to enable the **Data policy enforcement**. This applies to data sources, resource groups, or subscriptions. To enable *Data policy enforcement*, you must have *both* specific Identity and Access Management (IAM) privileges on the resource as well as specific Microsoft Purview privileges:

- You must have *either one* of the following IAM role combinations on the resource's Azure Resource Manager path or any parent of it (that is, using IAM permission inheritance):
  - IAM Owner
  - Both IAM Contributor and IAM User Access Administrator

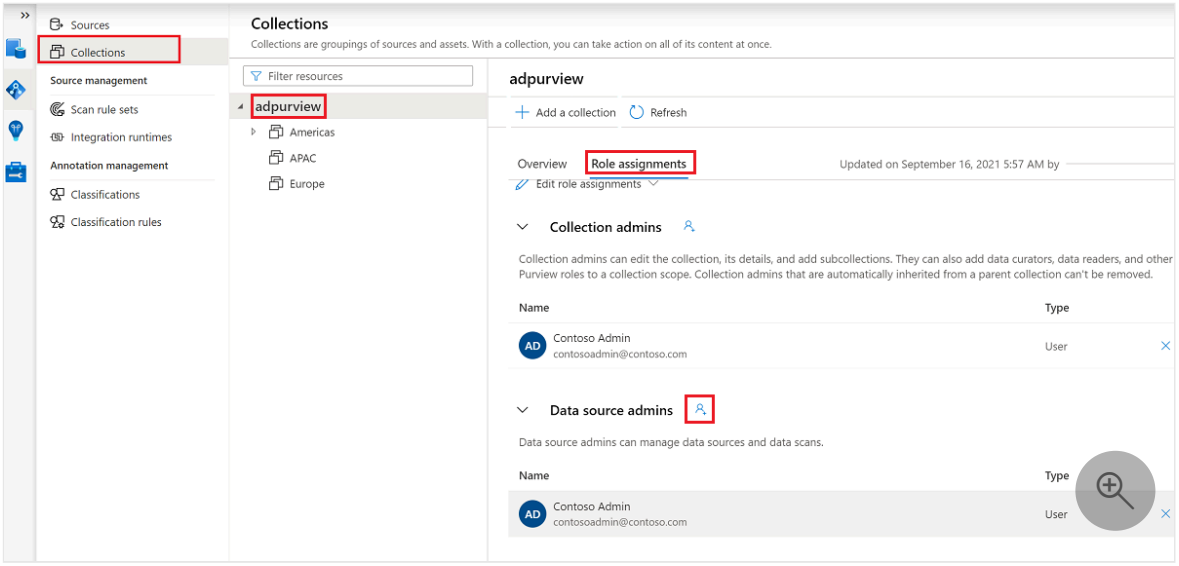
  To configure Azure role-based access control (RBAC) permissions, follow this guide. The following screenshot shows how to access the **Access Control** section in the Azure portal for the data resource to add a role assignment.



> **Note**
>
> The *IAM Owner* role for a data resource can be inherited from a parent resource group, a subscription, or a subscription management group. Check which Microsoft Entra users, groups, and service principals hold or are inheriting the *IAM Owner* role for the resource.

- You also need to have the Microsoft Purview *Data source admin* role for the collection or a parent collection (if inheritance is enabled). For more information, see the guide on managing Microsoft Purview role assignments.

  The following screenshot shows how to assign the *Data source admin* role at the root collection level.



Configure Microsoft Purview permissions to create, update, or delete access policies

To create, update or delete policies, you need to get the *Policy author* role in Microsoft Purview at *root collection level*:

- The *Policy author* role can create, update, and delete DevOps and Data Owner policies.
- The *Policy author* role can delete self-service access policies.

For more information about managing Microsoft Purview role assignments, see Create and manage collections in the Microsoft Purview Data Map.

> **Note**
>
> Policy author role must be configured at the root collection level.

In addition, to easily search Microsoft Entra users or groups when creating or updating the subject of a policy, you can greatly benefit from getting the Directory Readers permission in Microsoft Entra ID. This is a common permission for users in an Azure tenant. Without the Directory Reader permission, the Policy Author will have to type the complete username or email for all the principals included in the subject of a data policy.

Configure Microsoft Purview permissions for publishing Data Owner policies

Data Owner policies allow for checks and balances if you assign the Microsoft Purview *Policy author* and *Data source admin* roles to different people in the organization. Before a Data owner policy takes effect, a second person (*Data source admin*) must review it and explicitly approve it by publishing it. This does not apply to DevOps or Self-service access policies as publishing is automatic for them when those policies are created or updated.

To publish a Data owner policy you need to get the *Data source admin* role in Microsoft Purview at *root collection level*.

For more information about managing Microsoft Purview role assignments, see Create and manage collections in the Microsoft Purview Data Map.

> **Note**
>
> To publish Data owner policies, the Data source admin role must be configured at the root collection level.

Delegate access provisioning responsibility to roles in Microsoft Purview

After a resource has been enabled for **Data policy enforcement**, any Microsoft Purview user with the *Policy author* role at the root collection level can provision access to that data source from Microsoft Purview.

> **Note**
>
> Any Microsoft Purview root *Collection admin* can assign new users to root *Policy author* roles. Any *Collection admin* can assign new users to a *Data source admin* role under the collection. Minimize and carefully vet the users who hold Microsoft Purview *Collection admin*, *Data source admin*, or *Policy author* roles.

If a Microsoft Purview account with published policies is deleted, such policies will stop being enforced within an amount of time that depends on the specific data source. This change can have implications on both security and data access availability. The Contributor and Owner roles in IAM can delete Microsoft Purview accounts. You can check these permissions by going to the **Access control (IAM)** section for your Microsoft Purview account and selecting **Role Assignments**. You can also use a lock to

prevent the Microsoft Purview account from being deleted through Resource Manager locks.

Register the data sources in Microsoft Purview for Data Policy Enforcement

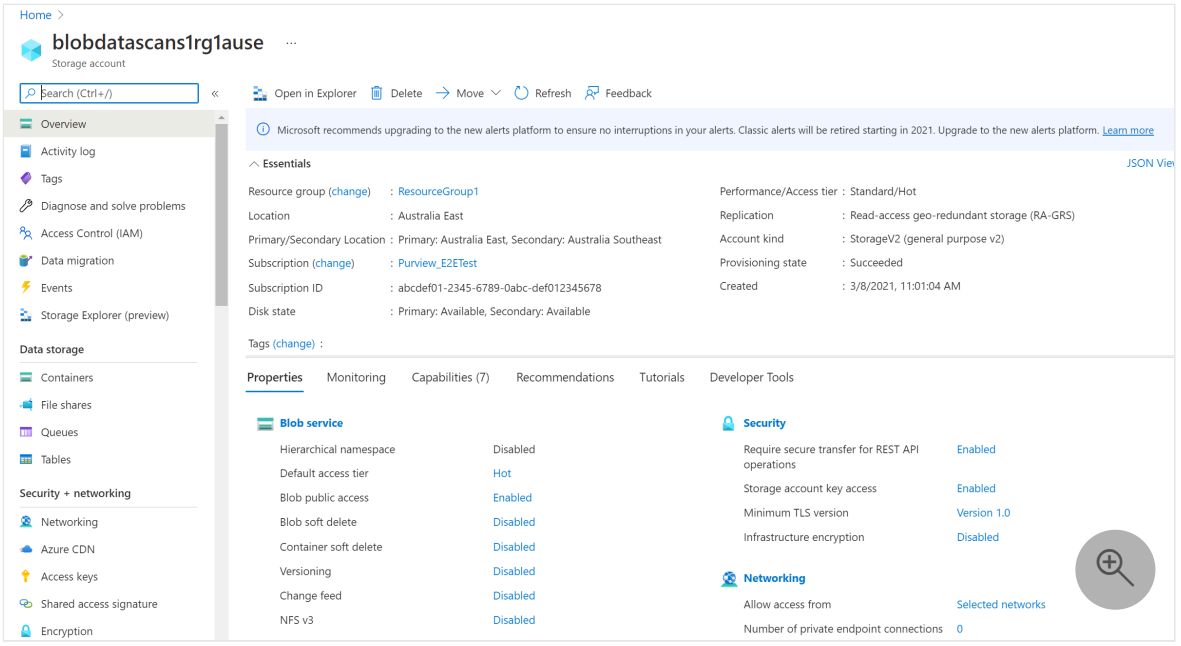
Your Azure Storage account needs to be registered in Microsoft Purview to later define access policies, and during registration we'll enable Data Policy Enforcement. **Data Policy Enforcement** is an available feature in Microsoft Purview that allows users to manage access to a resource from within Microsoft Purview. This allows you to centralize data discovery and access management, however it's a feature that directly impacts your data security.

> ⚠ **Warning**
>
> Before enabling Data Policy Enforcement for any of your resources, read through our **Data Policy Enforcement article**.
>
> This article includes Data Policy Enforcement best practices to help you ensure that your information is secure.

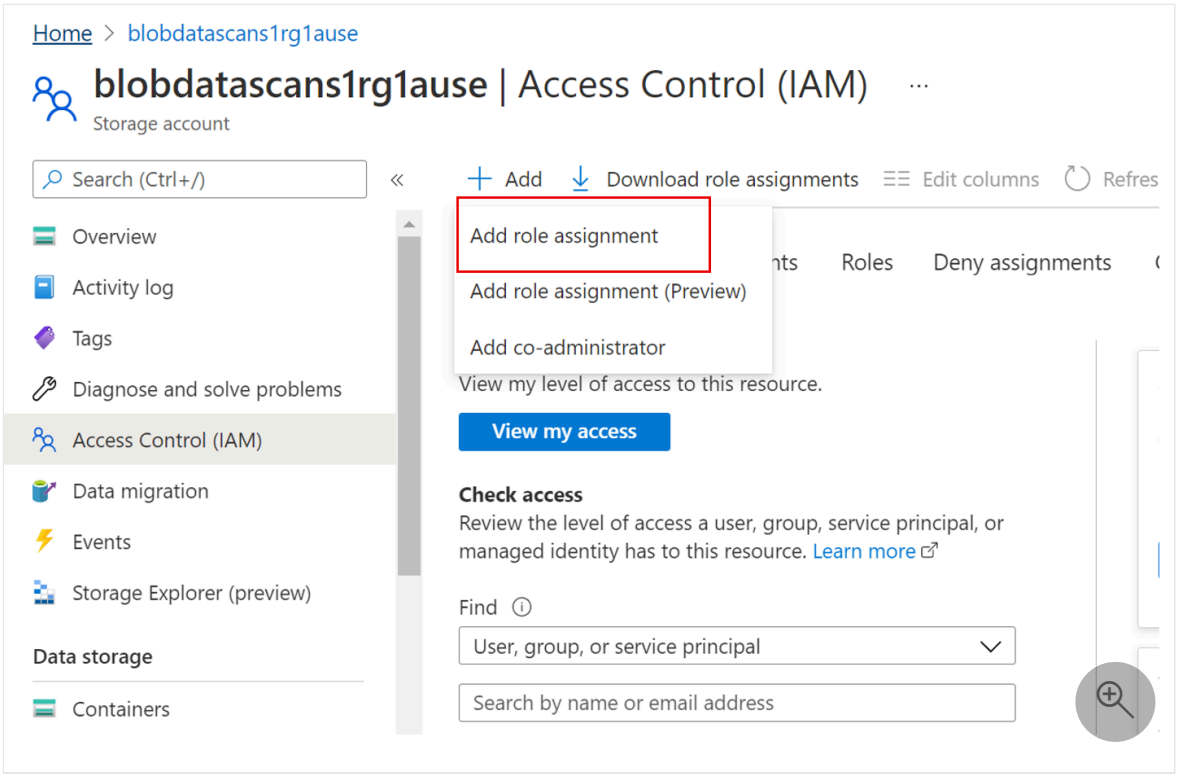
To register your resource and enable Data Policy Enforcement, follow these steps:

> ⓘ **Note**
>
> You need to be an owner of the subscription or resource group to be able to add a managed identity on an Azure resource.

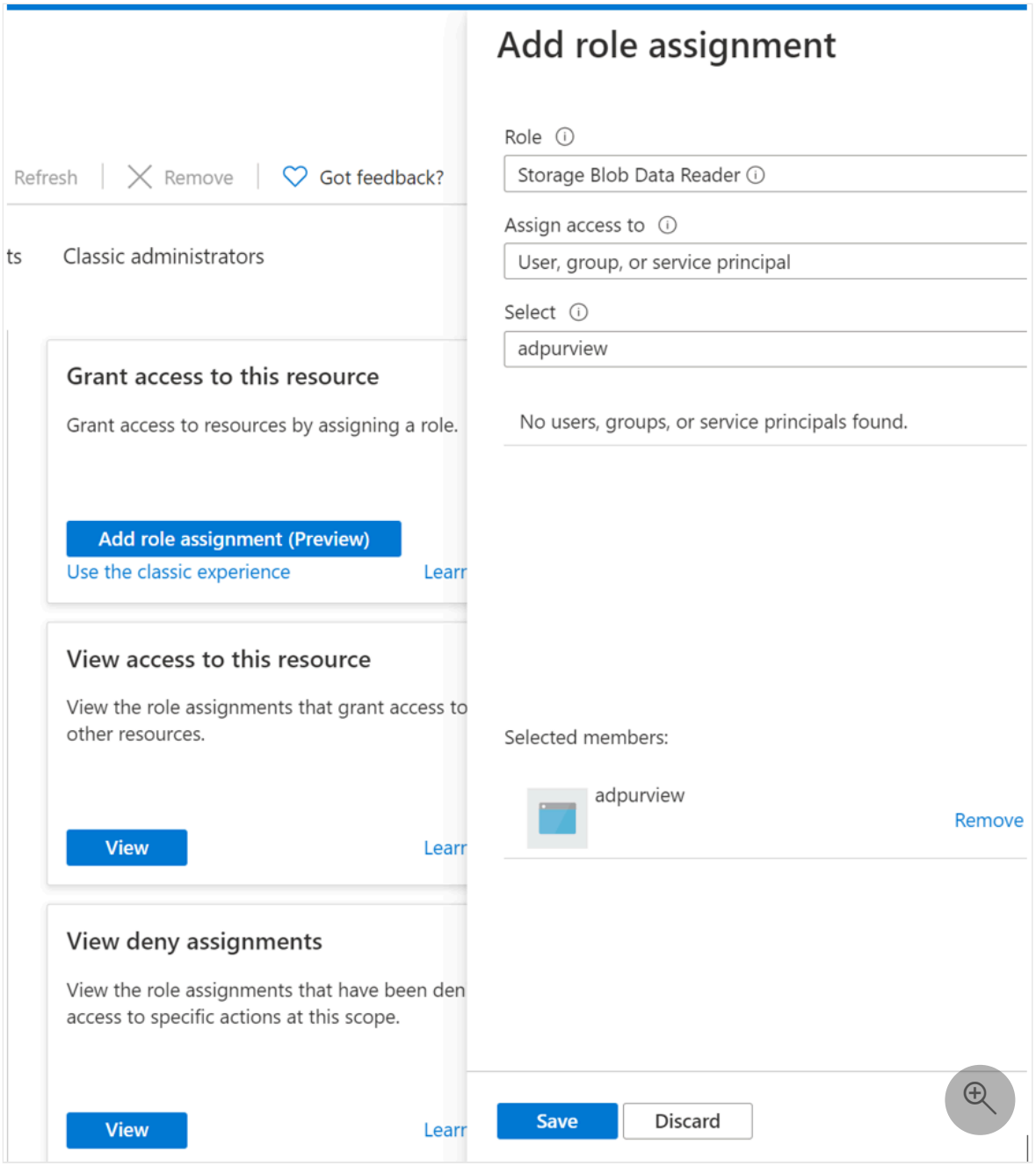
1. From the Azure portal, find the Azure Blob storage account that you would like to register.



2. Select **Access Control (IAM)** in the left navigation and then select **+ Add** --> **Add role assignment**.



3. Set the **Role** to **Storage Blob Data Reader** and enter your *Microsoft Purview account name* under the **Select** input box. Then, select **Save** to give this role assignment to your Microsoft Purview account.



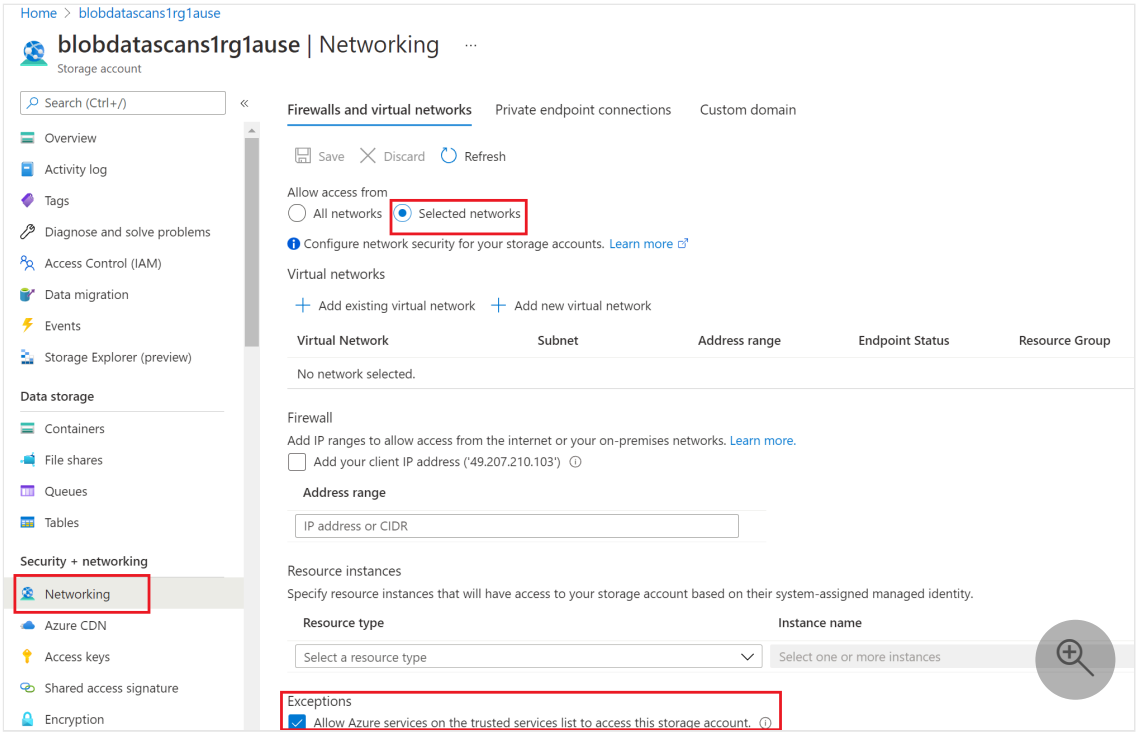
4. If you have a firewall enabled on your Storage account, follow these steps as well:

   a. Go into your Azure Storage account in Azure portal.

   b. Navigate to **Security + networking** > **Networking**.

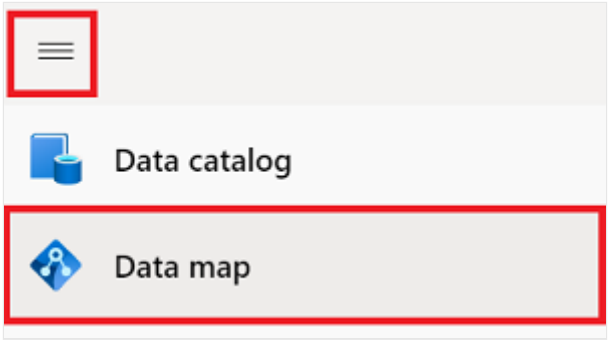
   c. Choose **Selected Networks** under **Allow access from**.

   d. In the **Exceptions** section, select **Allow trusted Microsoft services to access this storage account** and select **Save**.

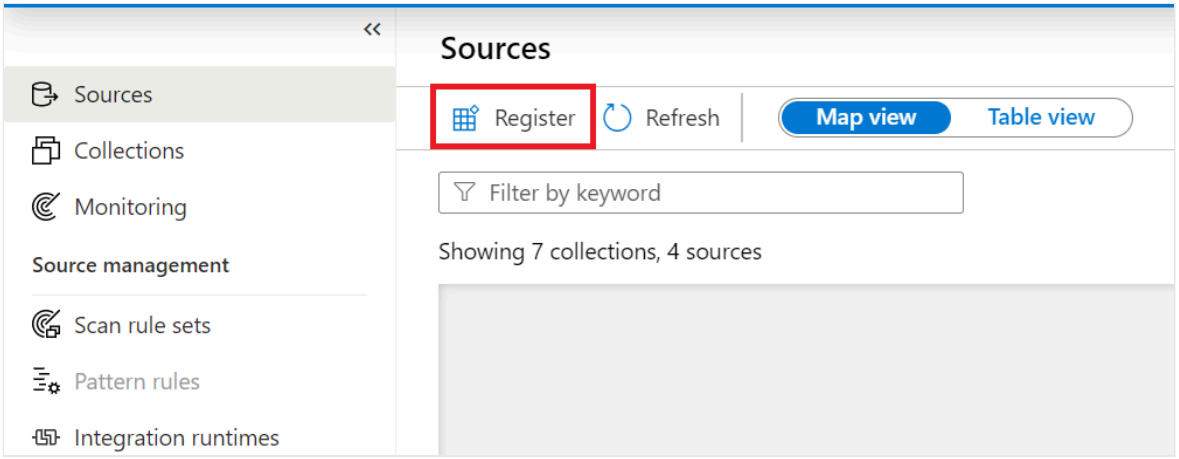


5. Once you have set up authentication for your storage account, go to the Microsoft Purview governance portal.

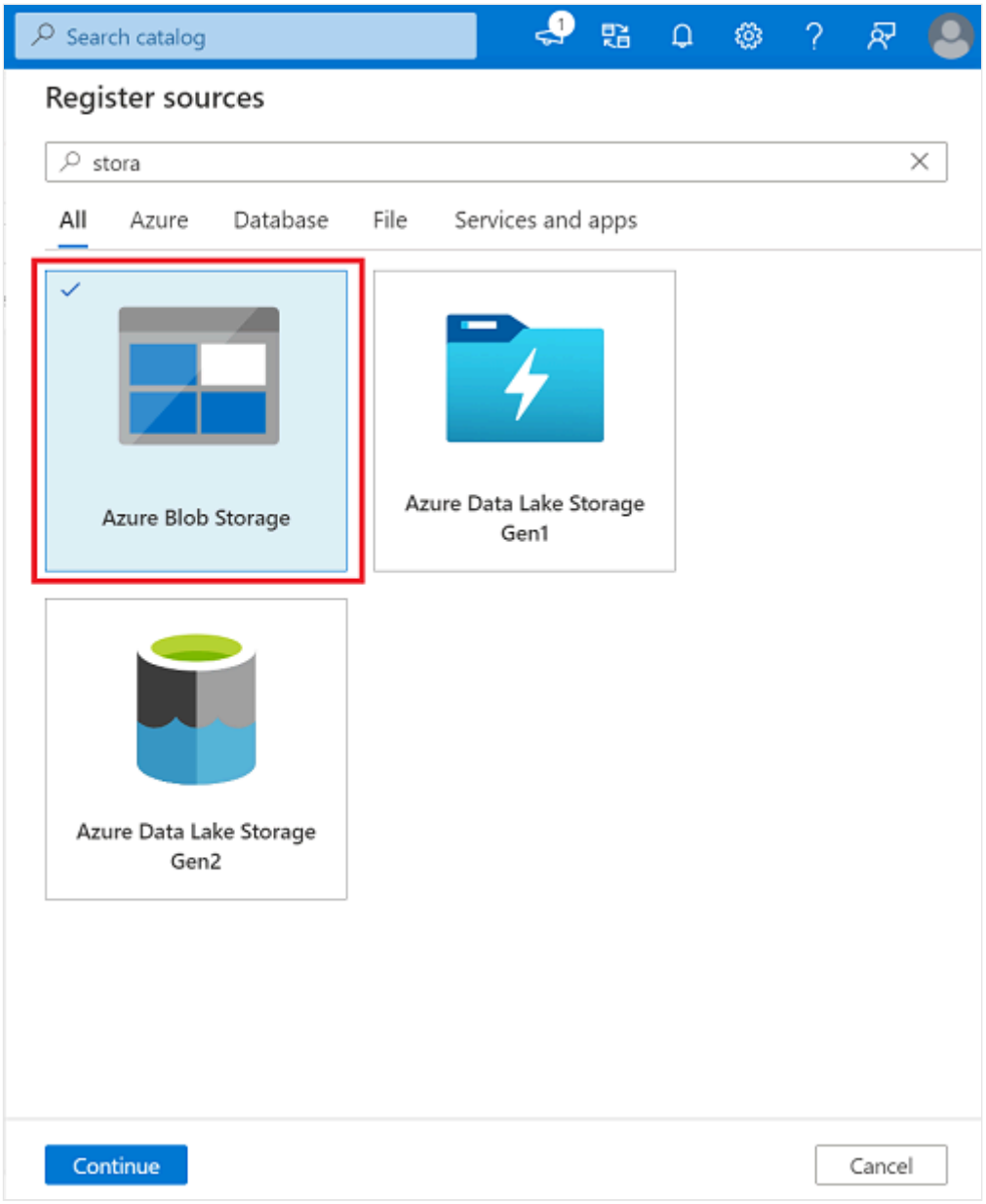
6. Select **Data Map** on the left menu.



7. Select **Register**.



8. On **Register sources**, select **Azure Blob Storage**.

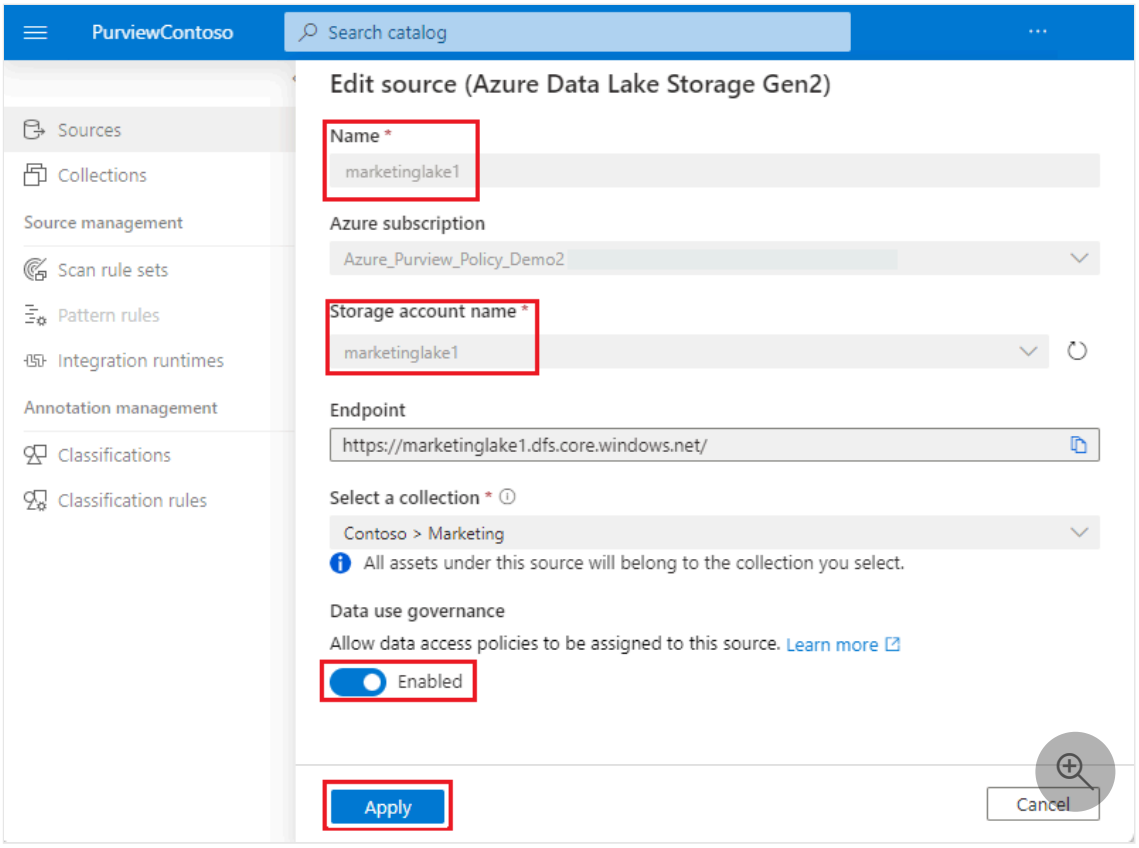


9. Select **Continue**.

10. On the **Register sources (Azure)** screen, do the following:

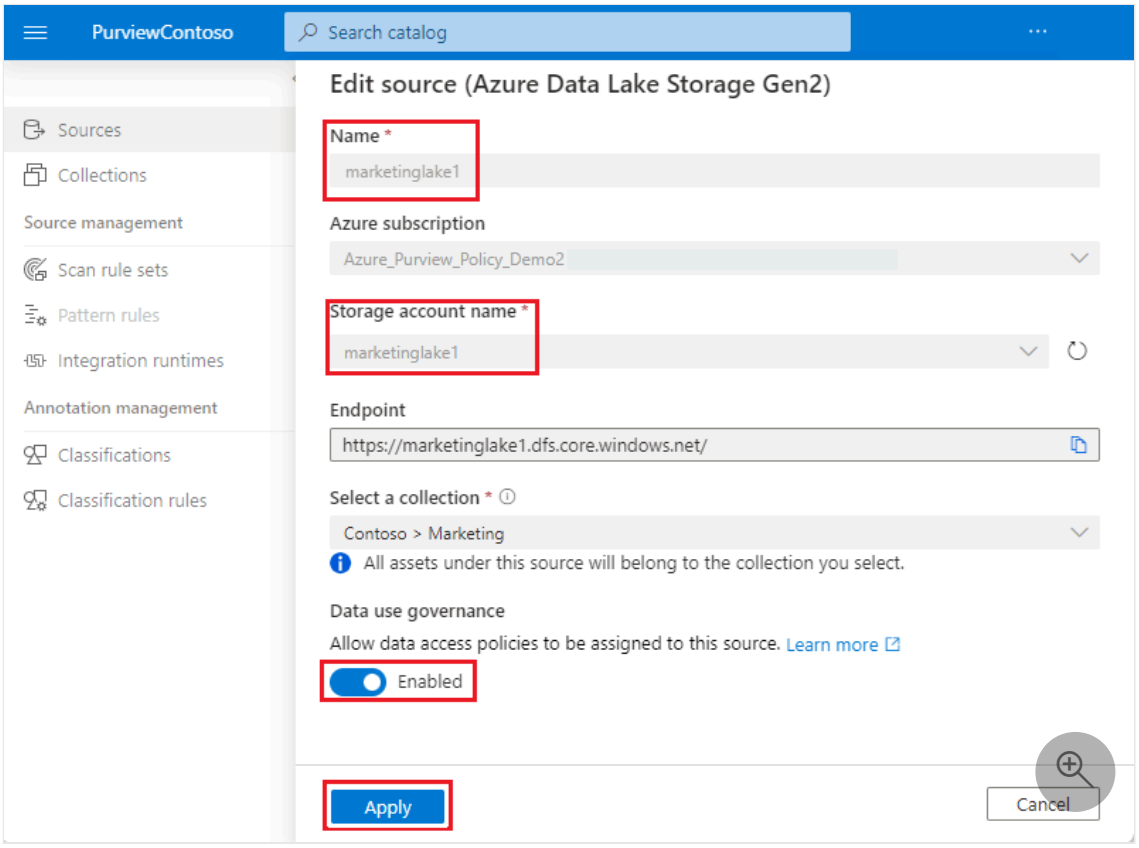
    a. In the **Name** box, enter a friendly name that the data source will be listed with in the catalog.

    b. In the **Subscription** dropdown list boxes, select the subscription where your storage account is housed. Then select your storage account under **Storage account name**. In **Select a collection** select the collection where you'd like to register your Azure Storage account.



c. In the **Select a collection** box, select a collection or create a new one (optional).

d. Set the *Data Policy Enforcement* toggle to **Enabled**, as shown in the image below.



> **Tip**

If the Data Policy Enforcement toggle is greyed out and unable to be selected:

i. Confirm you have followed all prerequisites to enable Data Policy Enforcement across your resources.

ii. Confirm that you have selected a storage account to be registered.

iii. It may be that this resource is already registered in another Microsoft Purview account. Hover over it to know the name of the Microsoft Purview account that has registered the data resource.first. Only one Microsoft Purview account can register a resource for Data Policy Enforcement at at time.

e. Select **Register** to register the resource group or subscription with Microsoft Purview with Data Policy Enforcement enabled.

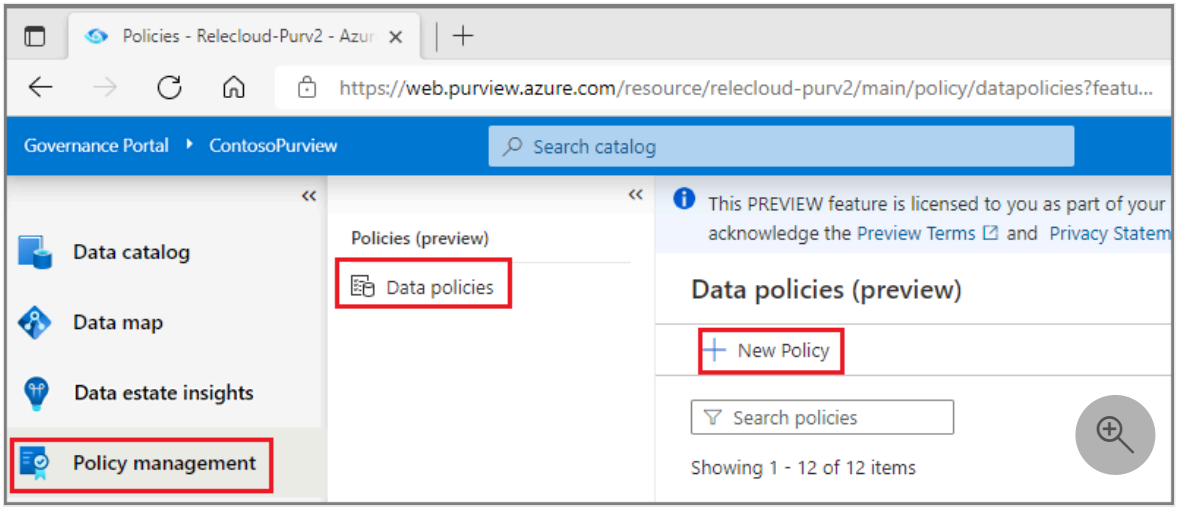
> **Tip**
>
> For more information about Data Policy Enforcement, including best practices or known issues, see our Data Policy Enforcement article.

Create a data owner policy

1. Sign in to the Microsoft Purview governance portal.

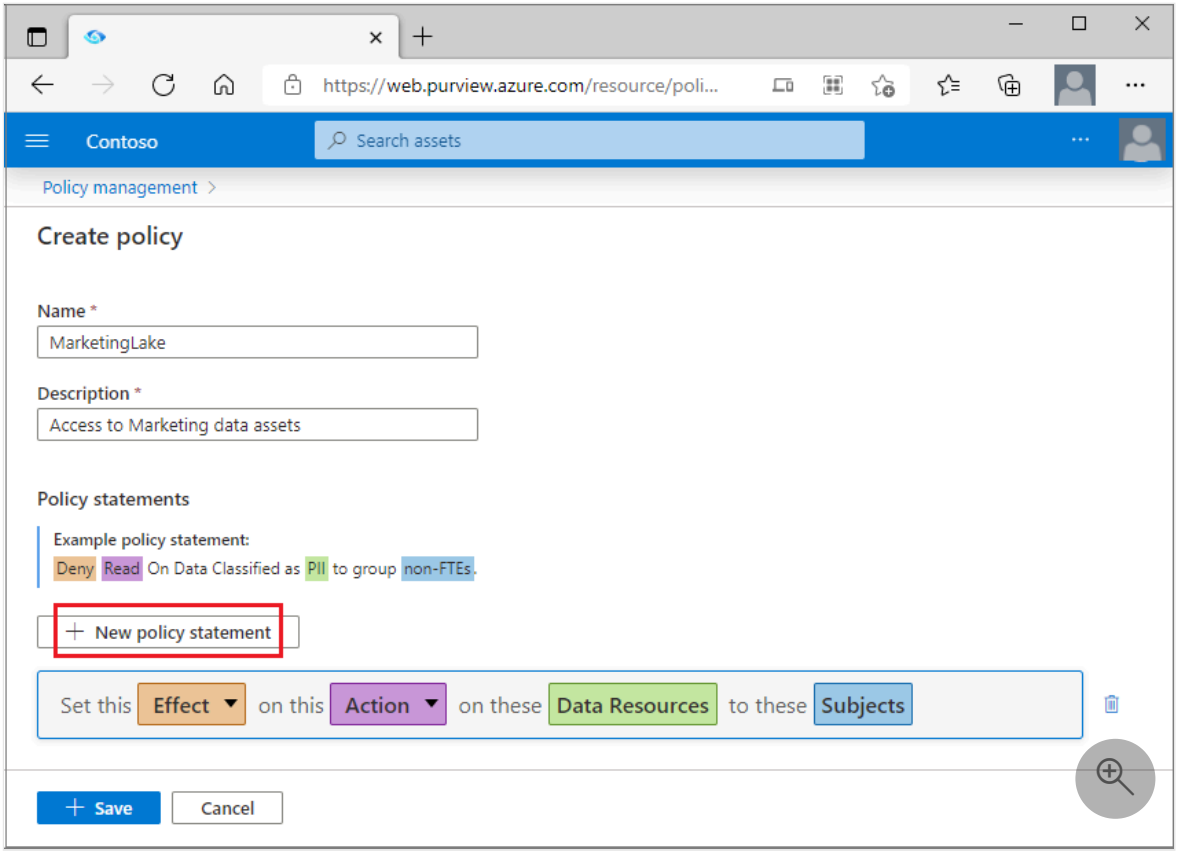
2. Navigate to the **Data policy** feature using the left side panel. Then select **Data policies**.

3. Select the **New Policy** button in the policy page.



4. The new policy page will appear. Enter the policy **Name** and **Description**.

5. To add a policy statement to the new policy, select the **New policy statement** button. This will bring up the policy statement builder.



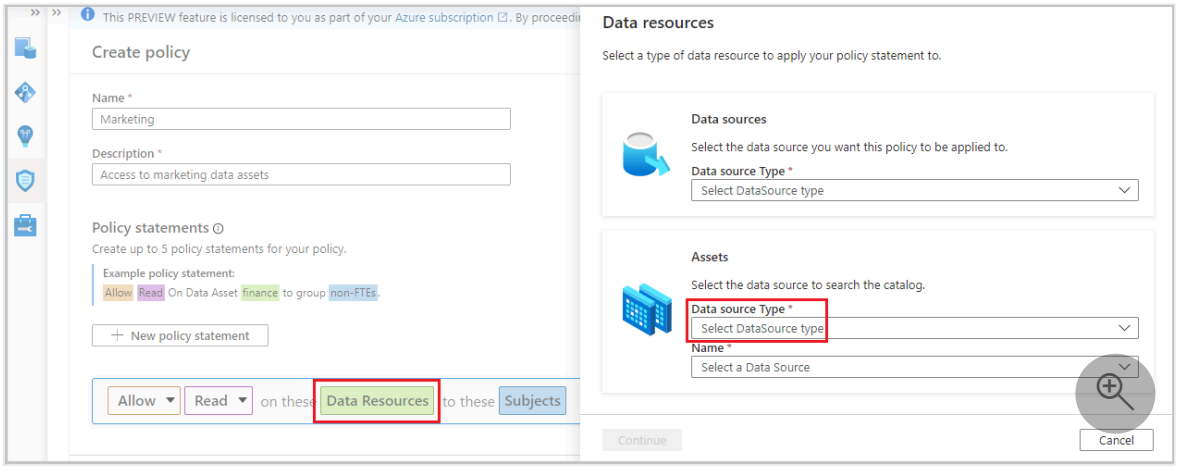
6. Select the **Effect** button and choose *Allow* from the drop-down list.

7. Select the **Action** button and choose *Read* or *Modify* from the drop-down list.

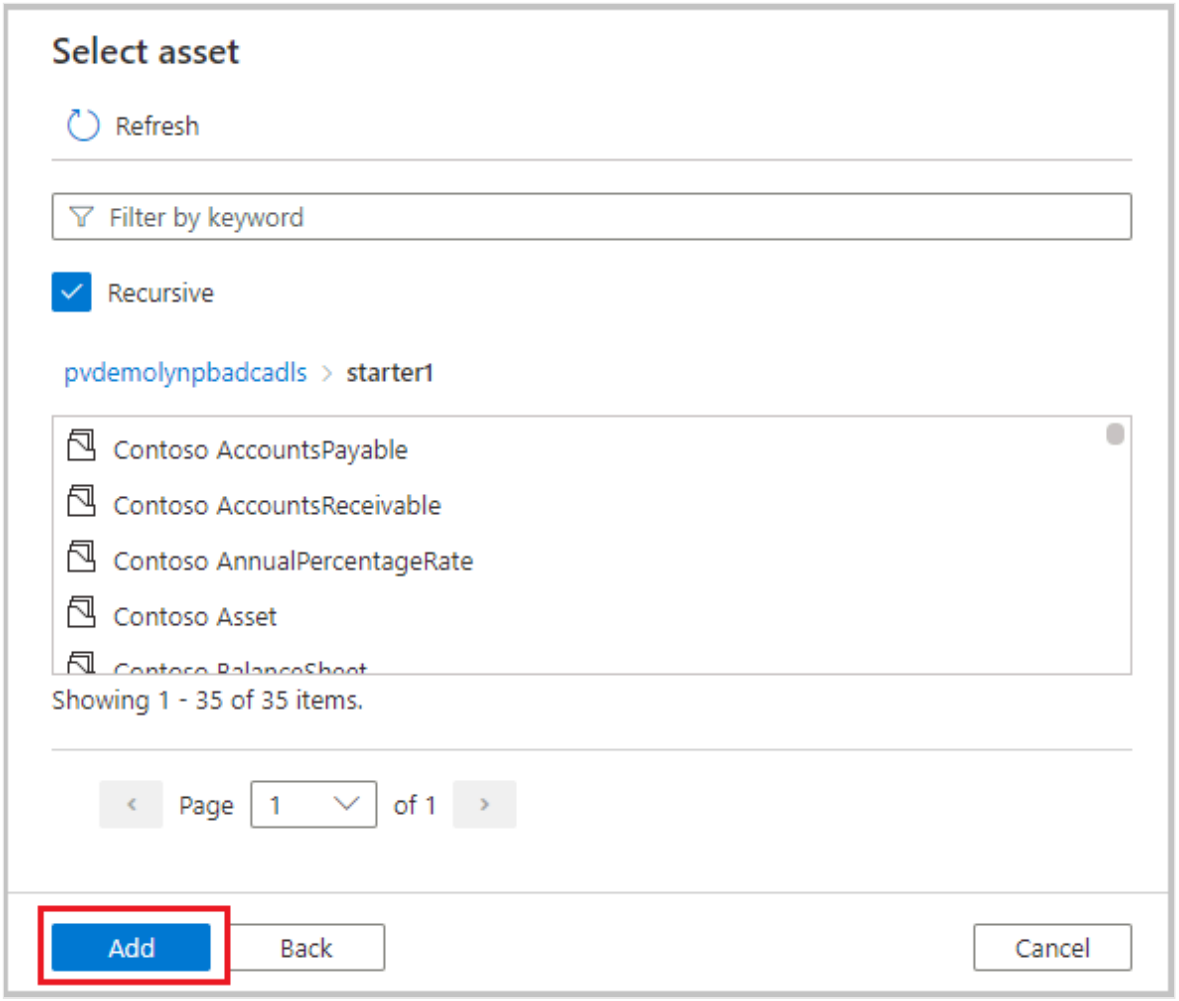
8. Select the **Data Resources** button to bring up the window to enter Data resource information, which will open to the right.

9. Under the **Data Resources** Panel do one of two things depending on the granularity of the policy:

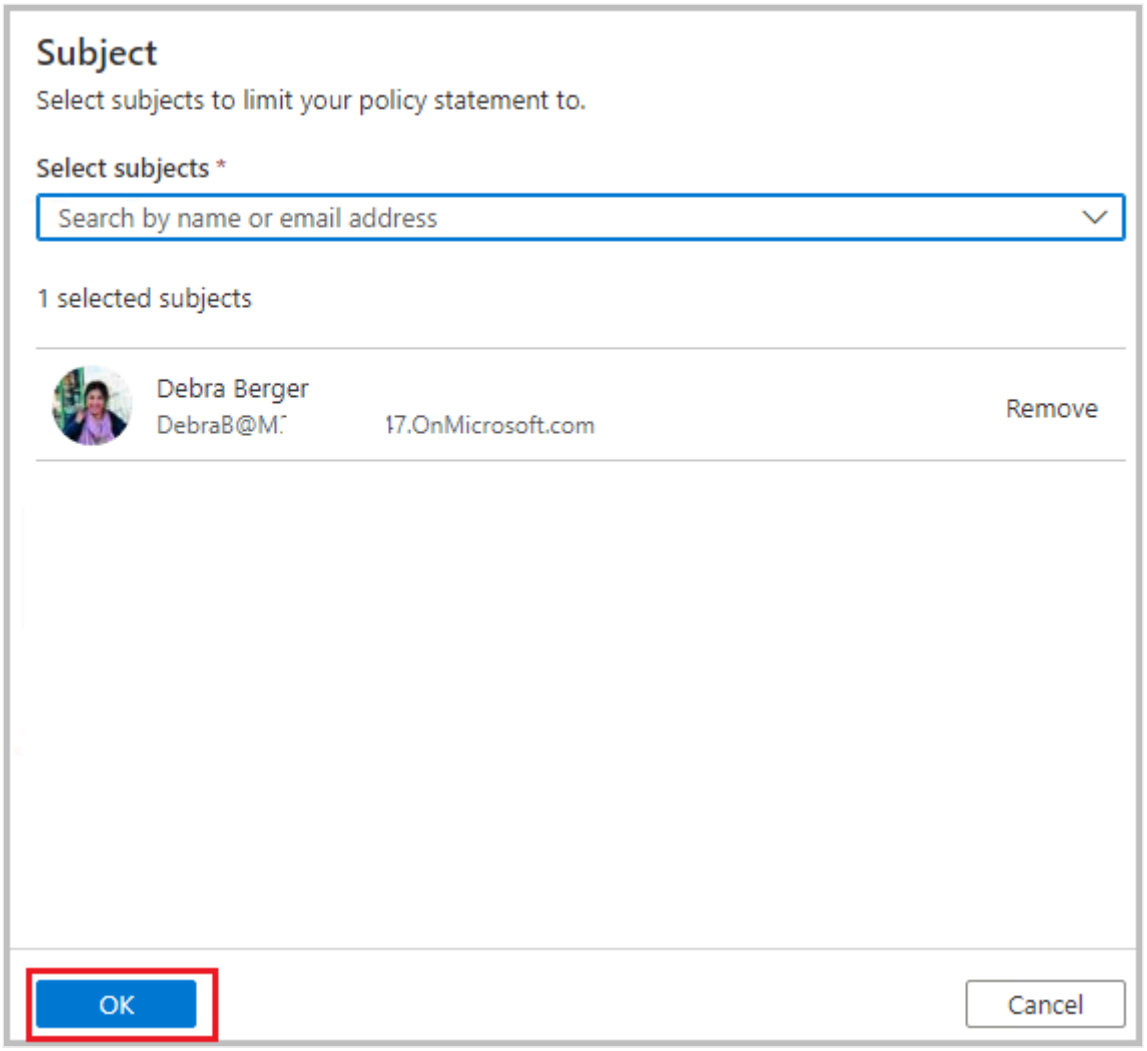
    - To create a broad policy statement that covers an entire data source, resource group, or subscription that was previously registered, use the **Data sources** box and select its **Type**.
    - To create a fine-grained policy, use the **Assets** box instead. Enter the **Data Source Type** and the **Name** of a previously registered and scanned data source. See example in the image.



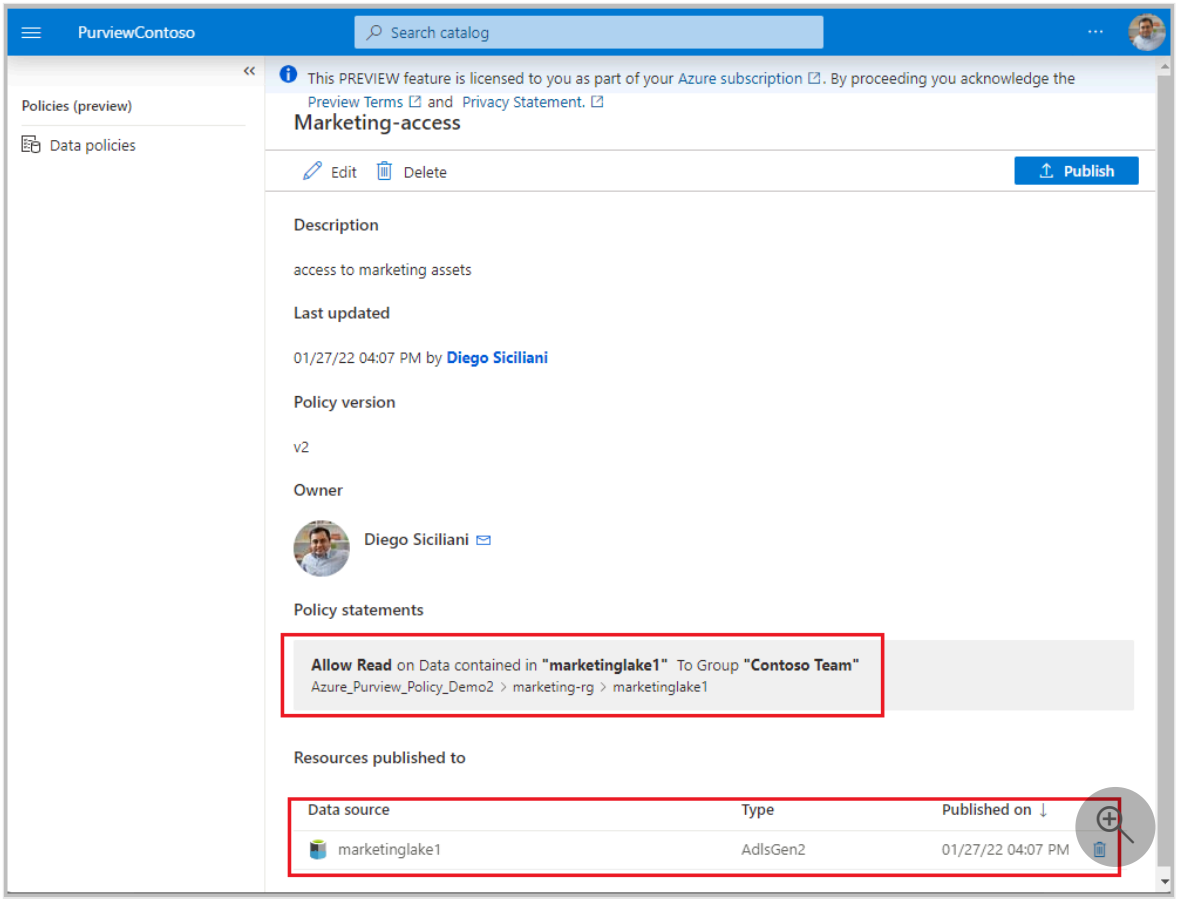
10. Select the **Continue** button and traverse the hierarchy to select and underlying data-object (for example: folder, file, etc.). Select **Recursive** to apply the policy from that point in the hierarchy down to any child data-objects. Then select the **Add** button. This will take you back to the policy editor.



11. Select the **Subjects** button and enter the subject identity as a principal, group, or MSI. Then select the **OK** button. This will take you back to the policy editor



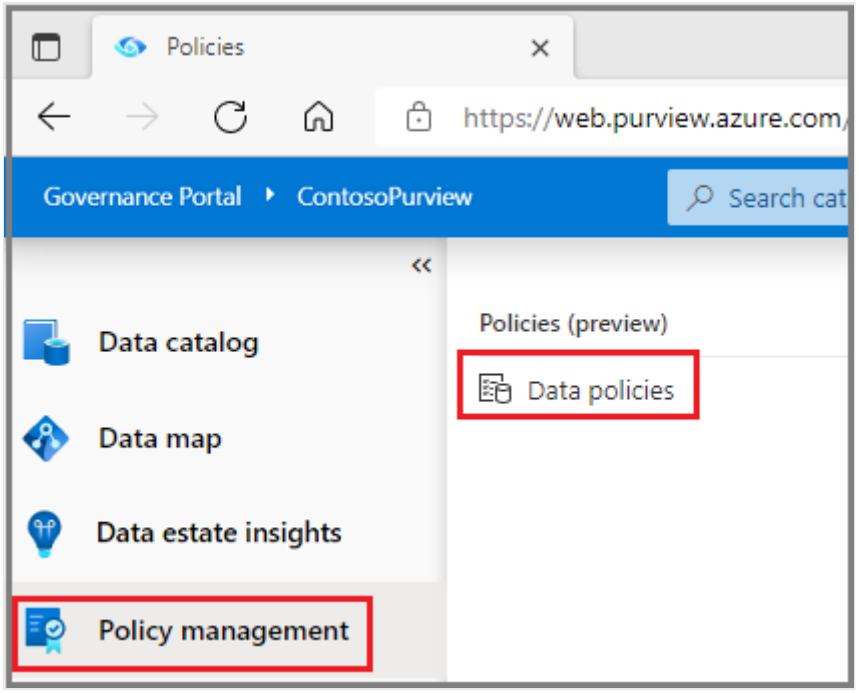
12. Select the **Save** button to save the policy.



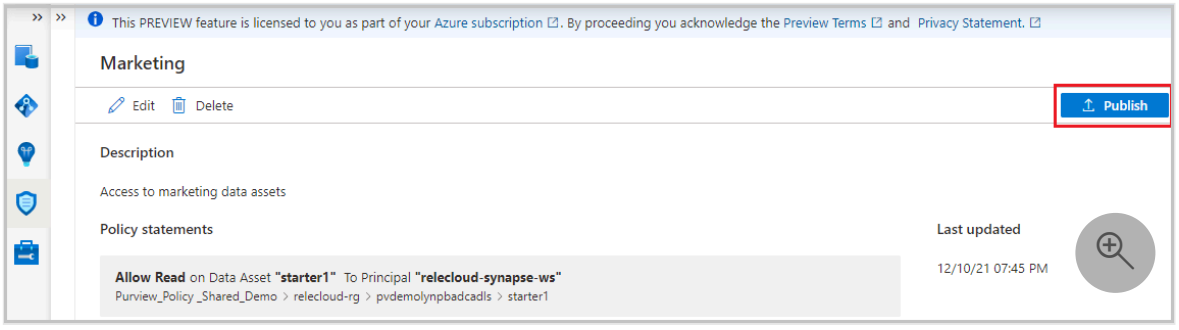
Publish a data owner policy

1. Sign in to the Microsoft Purview governance portal.

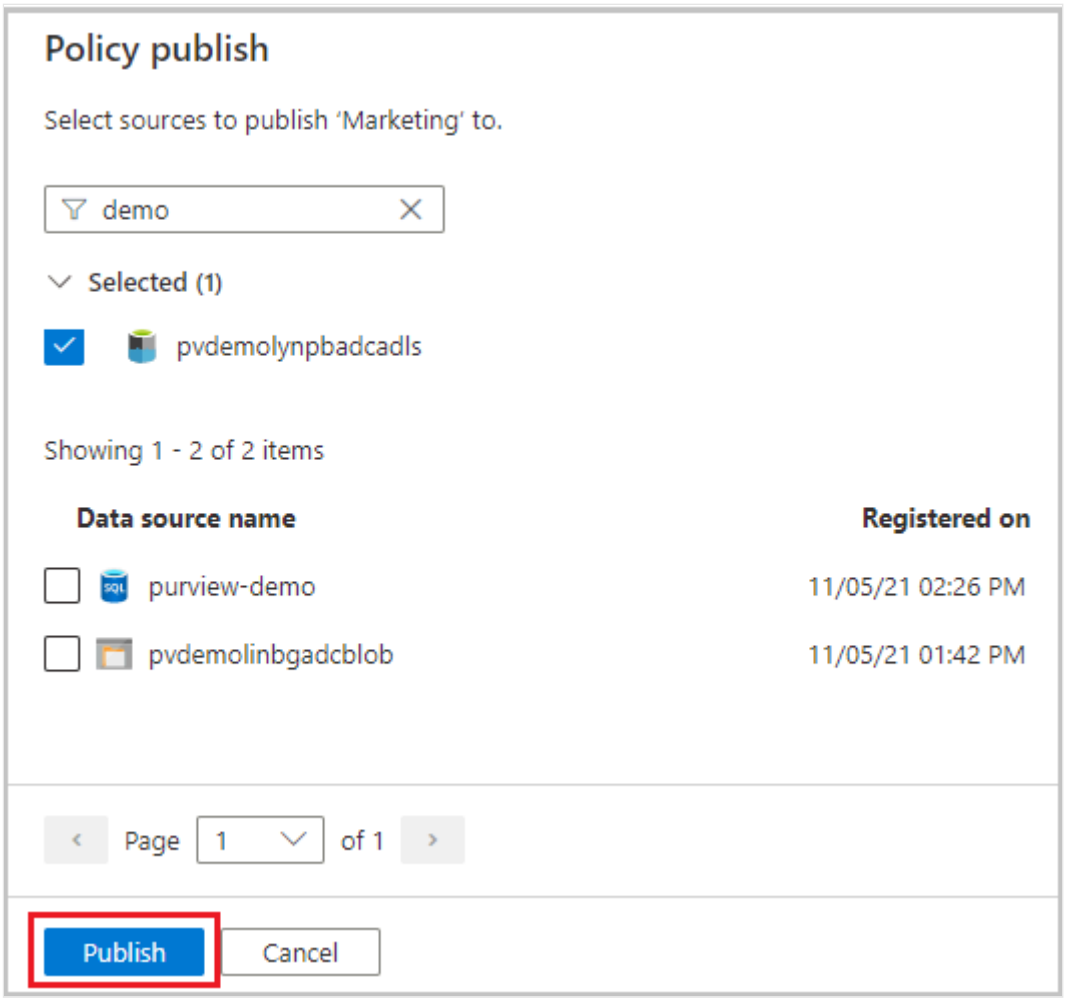
2. Navigate to the **Data policy** feature using the left side panel. Then select **Data policies**.



3. The Policy portal will present the list of existing policies in Microsoft Purview. Locate the policy that needs to be published. Select the **Publish** button on the right top corner of the page.



4. A list of data sources is displayed. You can enter a name to filter the list. Then, select each data source where this policy is to be published and then select the **Publish** button.

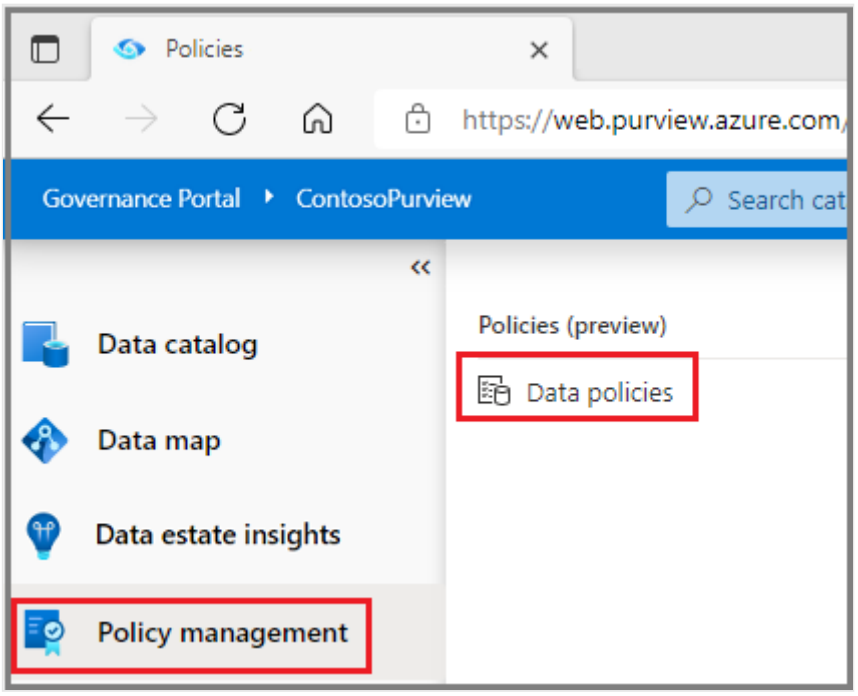


> **ⓘ Important**
>
> - Publish is a background operation. It can take up to **2 hours** for the changes to be reflected in Storage account(s).

Clean up resources

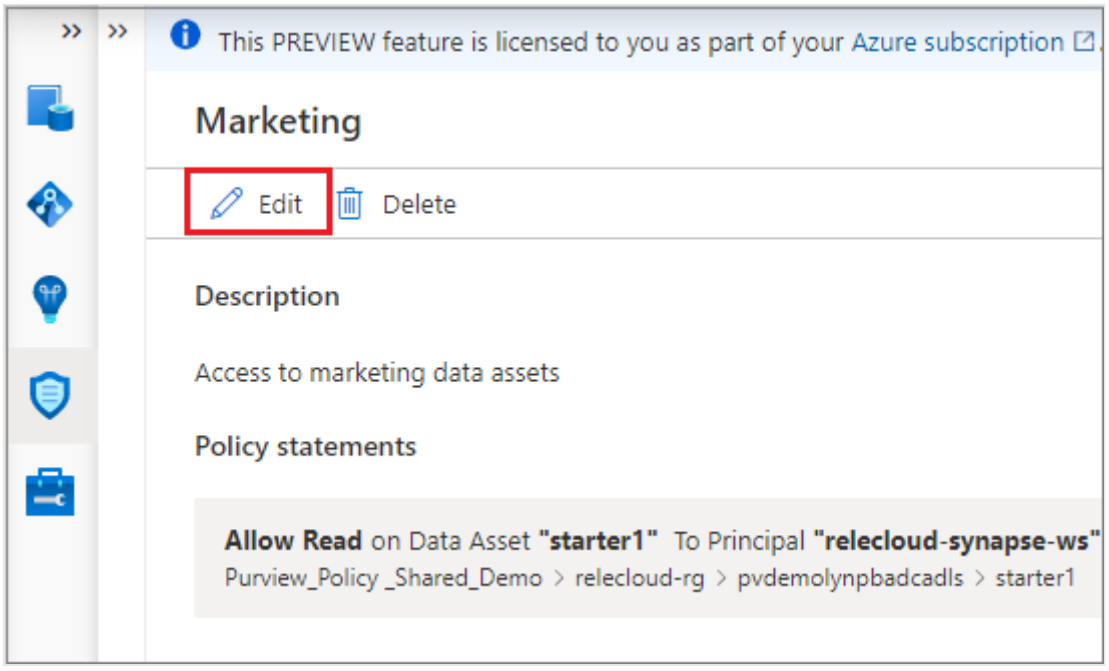
To delete a policy in Microsoft Purview, follow these steps:

1. Sign in to the Microsoft Purview governance portal.
2. Navigate to the **Data policy** feature using the left side panel. Then select **Data policies**.



3. The Policy portal will present the list of existing policies in Microsoft Purview. Select the policy that needs to be updated.

4. The policy details page will appear, including Edit and Delete options. Select the **Edit** button, which brings up the policy statement builder. Now, any parts of the statements in this policy can be updated. To delete the policy, use the **Delete** button.



Next steps

Check our demo and related tutorials:

Demo of access policy for Azure Storage

Concepts for Microsoft Purview data owner policies

Enable Microsoft Purview data owner policies on all data sources in a subscription or a resource group

Feedback

Was this page helpful? Yes No

Authoring and publishing Microsoft Purview Data owner policies (preview)

Article • 02/24/2025

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

> **Important**
>
> This feature is currently in preview. The Supplemental Terms of Use for Microsoft Azure Previews include additional legal terms that apply to Azure features that are in beta, in preview, or otherwise not yet released into general availability.

Data owner policies are a type of Microsoft Purview access policies. They allow you to manage access to user data in sources that have been registered for *Data Policy Enforcement* in Microsoft Purview. These policies can be authored directly in the Microsoft Purview governance portal, and after publishing, they get enforced by the data source.

This guide describes how to create, update, and publish data owner policies in the Microsoft Purview governance portal.

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- A new or existing Microsoft Purview account. Follow this quickstart guide to create one.

Configuration

Before authoring policies in the Microsoft Purview policy portal, you'll need to configure Microsoft Purview and the data sources so that they can enforce those policies.

1. Follow any policy-specific prerequisites for your source. Check the Microsoft Purview supported data sources table and select the link in the **Access Policy** column for sources where access policies are available. Follow any steps listed in the Access policy or Prerequisites sections.
2. Register the data source in Microsoft Purview. Follow the **Prerequisites** and **Register** sections of the source pages for your resources.
3. Enable the Data policy enforcement option on the data source registration. Data Policy Enforcement needs certain permissions and can affect the security of your data, as it delegates to certain Microsoft Purview roles to manage access to the data sources. **Go through the secure practices related to Data Policy Enforcement in this guide**: How to enable Data Policy Enforcement

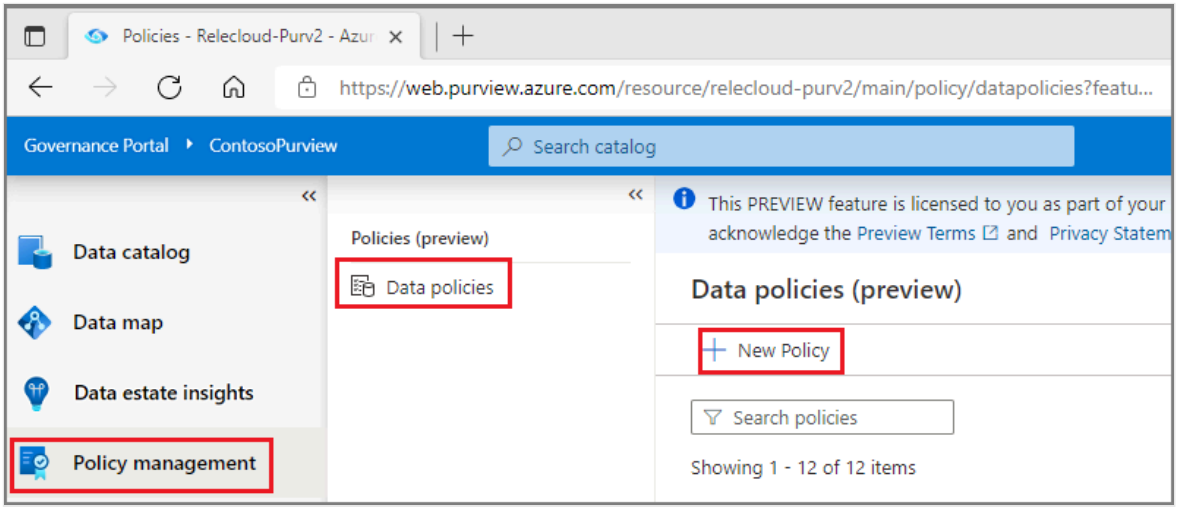
Create a new policy

This section describes the steps to create a new policy in Microsoft Purview. Ensure you have the *Policy Author* permission as described here.

1. Sign in to the Microsoft Purview governance portal.

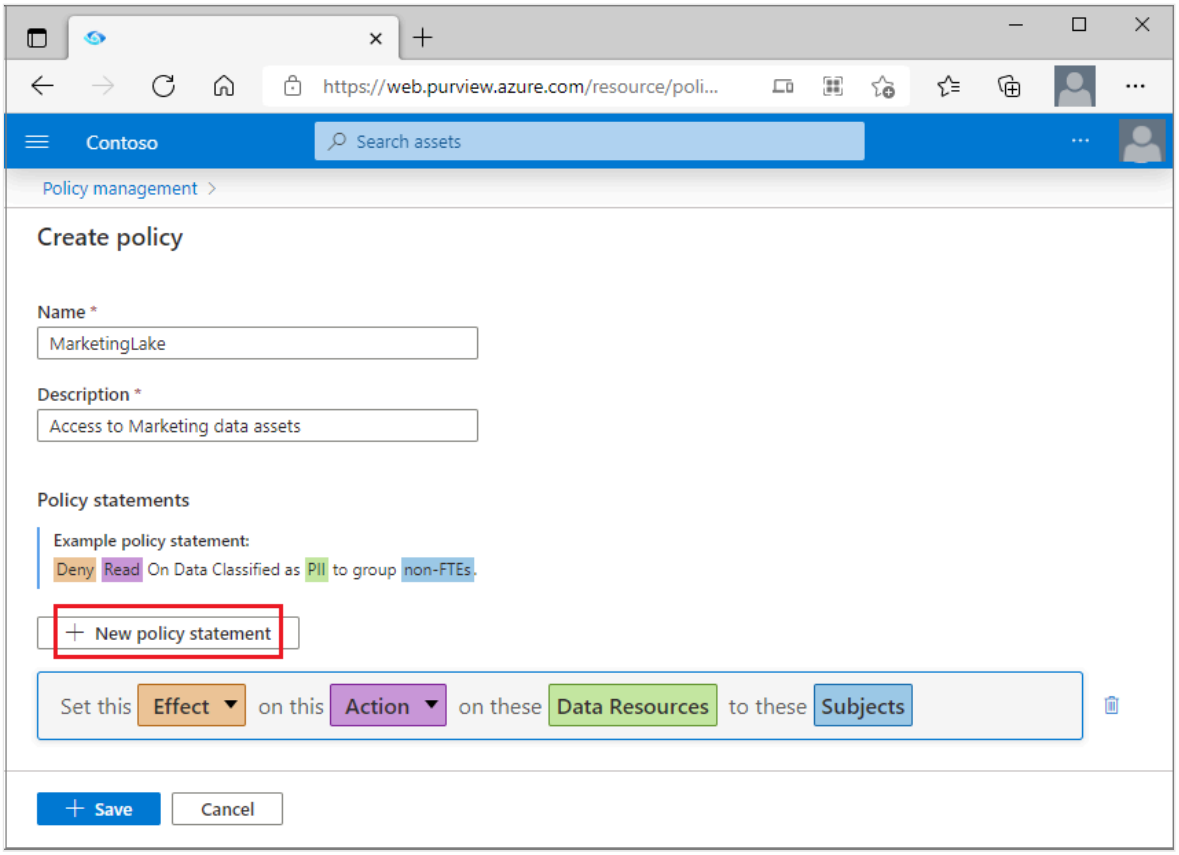
2. Navigate to the **Data policy** feature using the left side panel. Then select **Data policies**.

3. Select the **New Policy** button in the policy page.



4. The new policy page will appear. Enter the policy **Name** and **Description**.

5. To add a policy statement to the new policy, select the **New policy statement** button. This will bring up the policy statement builder.



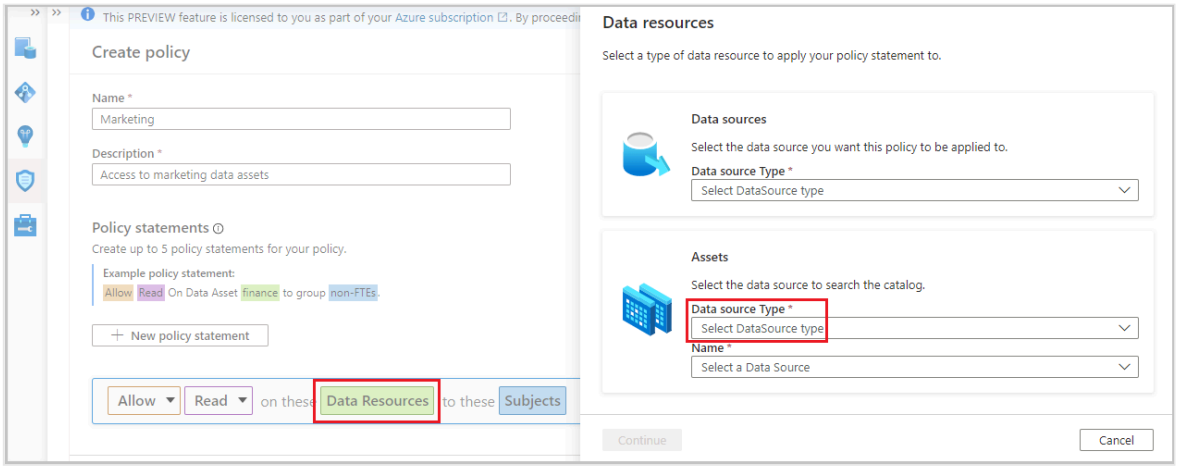
6. Select the **Effect** button and choose *Allow* from the drop-down list.

7. Select the **Action** button and choose *Read* or *Modify* from the drop-down list.

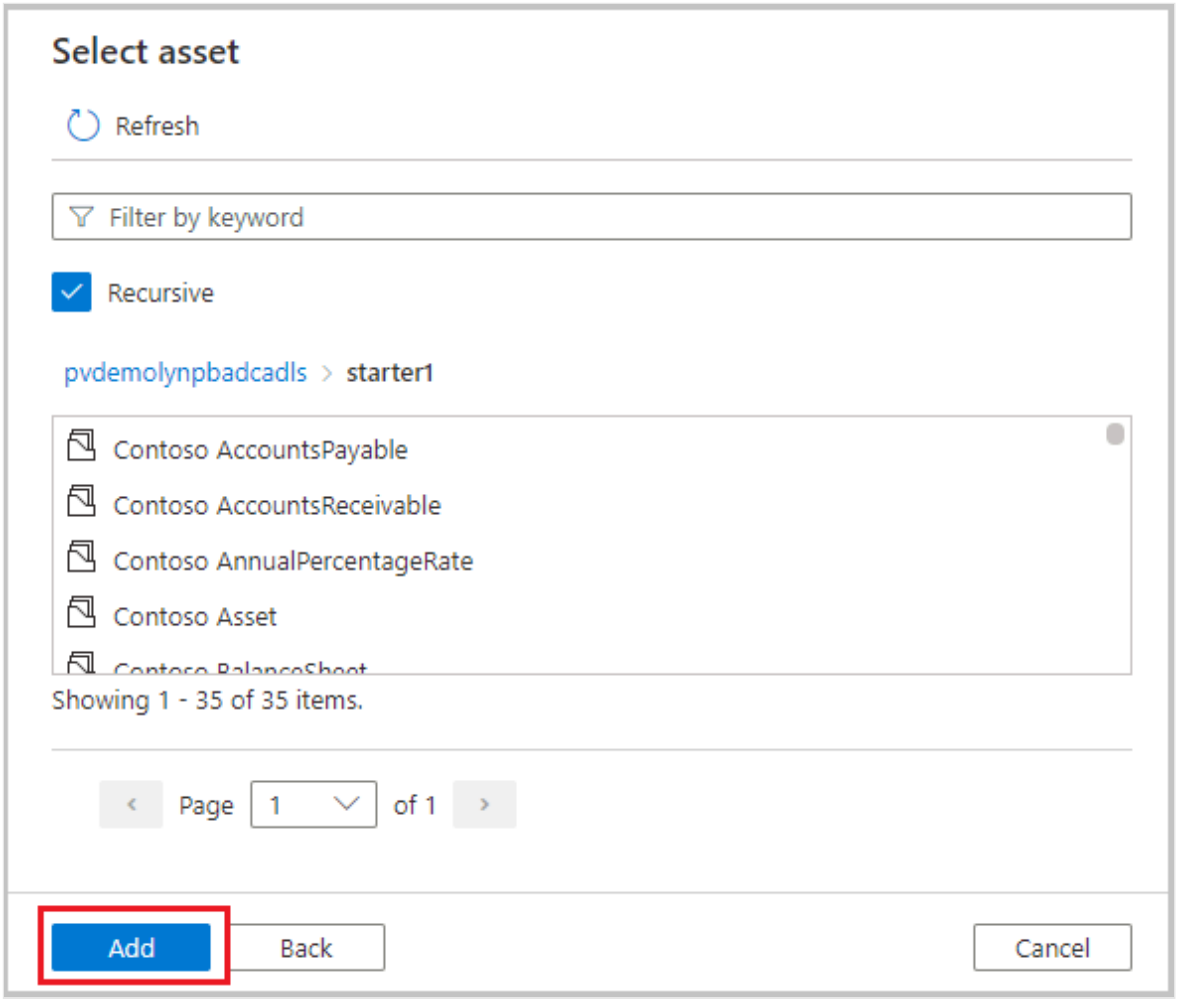
8. Select the **Data Resources** button to bring up the window to enter Data resource information, which will open to the right.

9. Under the **Data Resources** Panel do **one of two things** depending on the granularity of the policy:

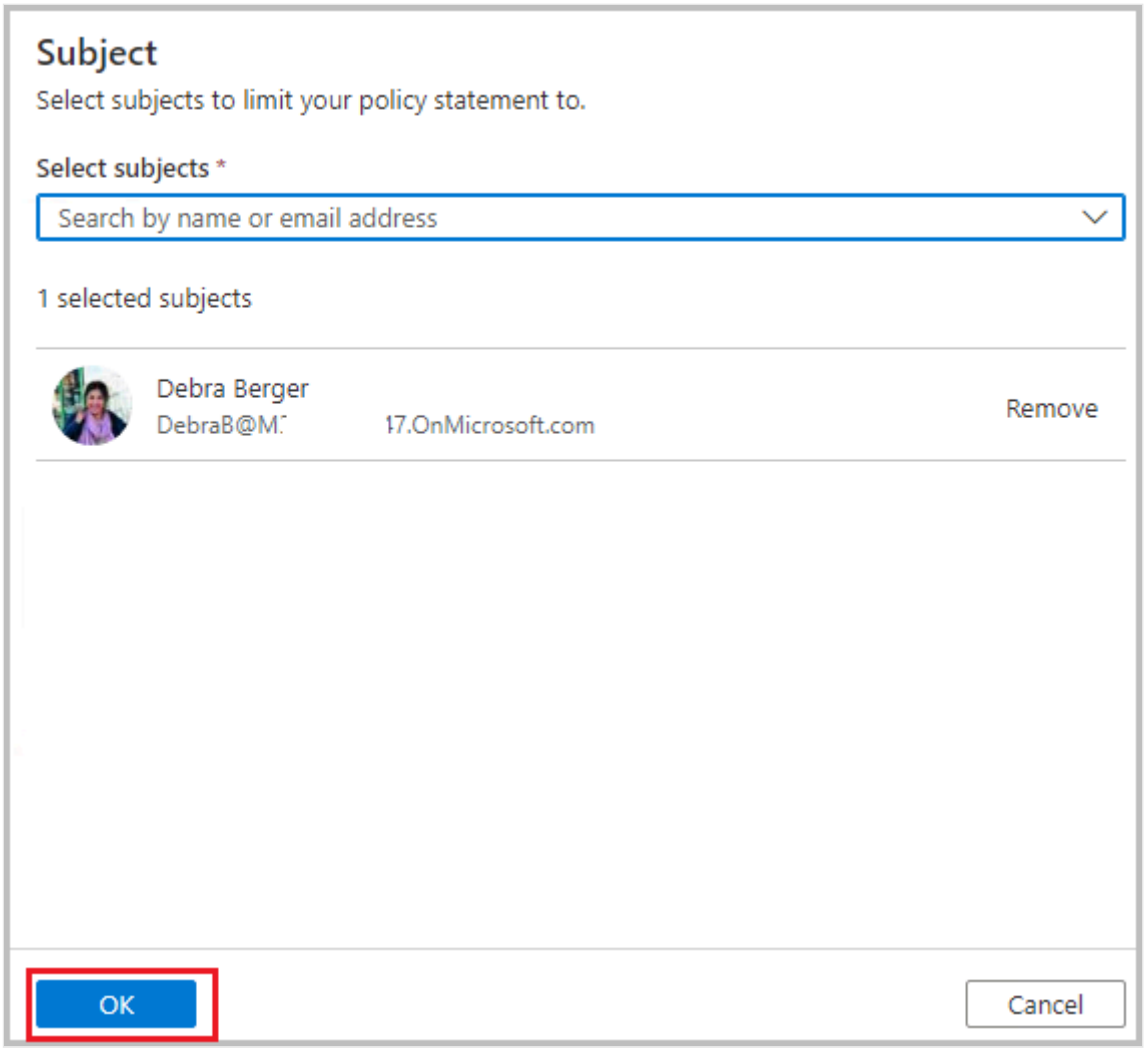
    - To create a broad policy statement that covers an entire data source, resource group, or subscription that was previously registered, use the **Data sources** box and select its **Type**.
    - To create a fine-grained policy, use the **Assets** box instead. Enter the **Data Source Type** and the **Name** of a previously registered and scanned data source. See example in the image.



10. Select the **Continue** button and transverse the hierarchy to select and underlying data-object (for example: folder, file, etc.). Select **Recursive** to apply the policy from that point in the hierarchy down to any child data-objects. Then select the **Add** button. This will take you back to the policy editor.



11. Select the **Subjects** button and enter the subject identity as a principal, group, or MSI. Note that Microsoft 365 groups are supported but updates to group membership take up to 1 hour to get reflected by Microsoft Entra ID. Then select the **OK** button. This will take you back to the policy editor.



12. Select the **Save** button to save the policy.

Now that you have created your policy, you will need to publish it for it to become active.

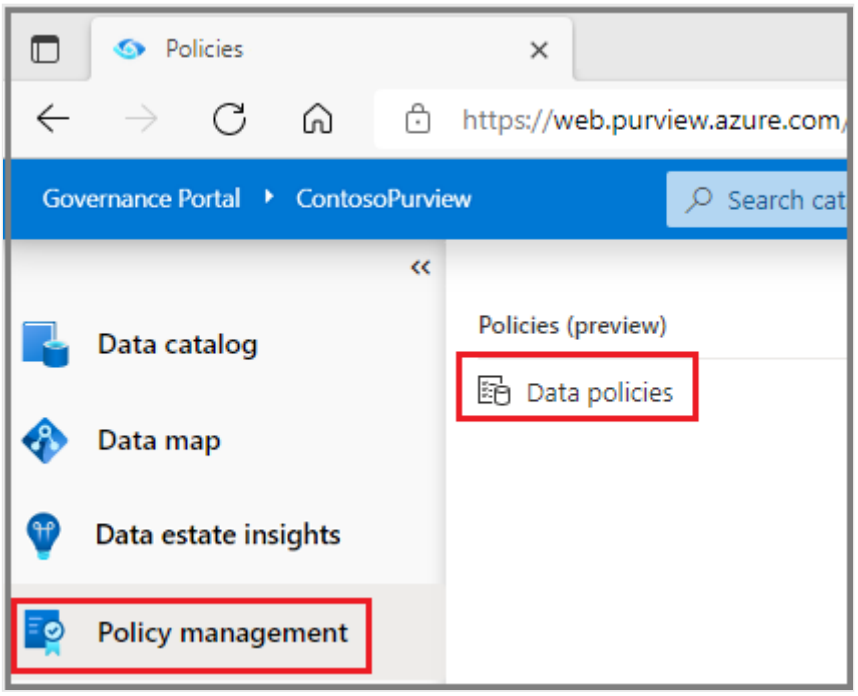
Publish a policy

A newly created policy is in the **draft** state. The process of publishing associates the new policy with one or more data sources under governance. This is called "binding" a policy to a data source.

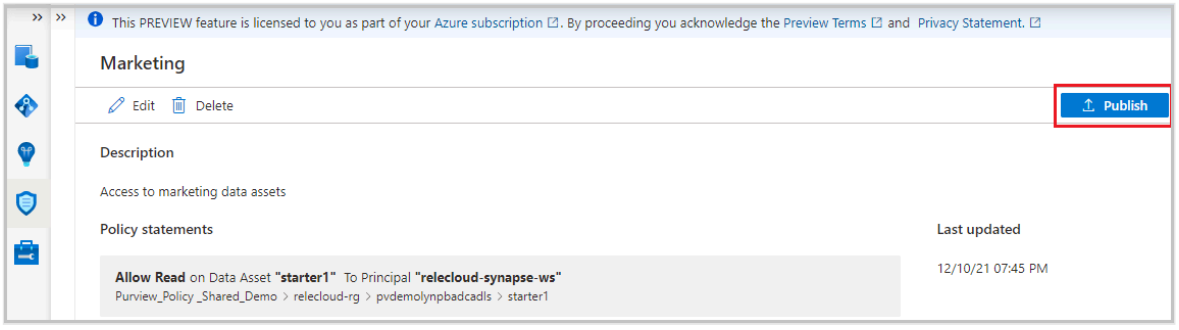
Ensure you have the *Data Source Admin* permission as described here

The steps to publish a policy are as follows:

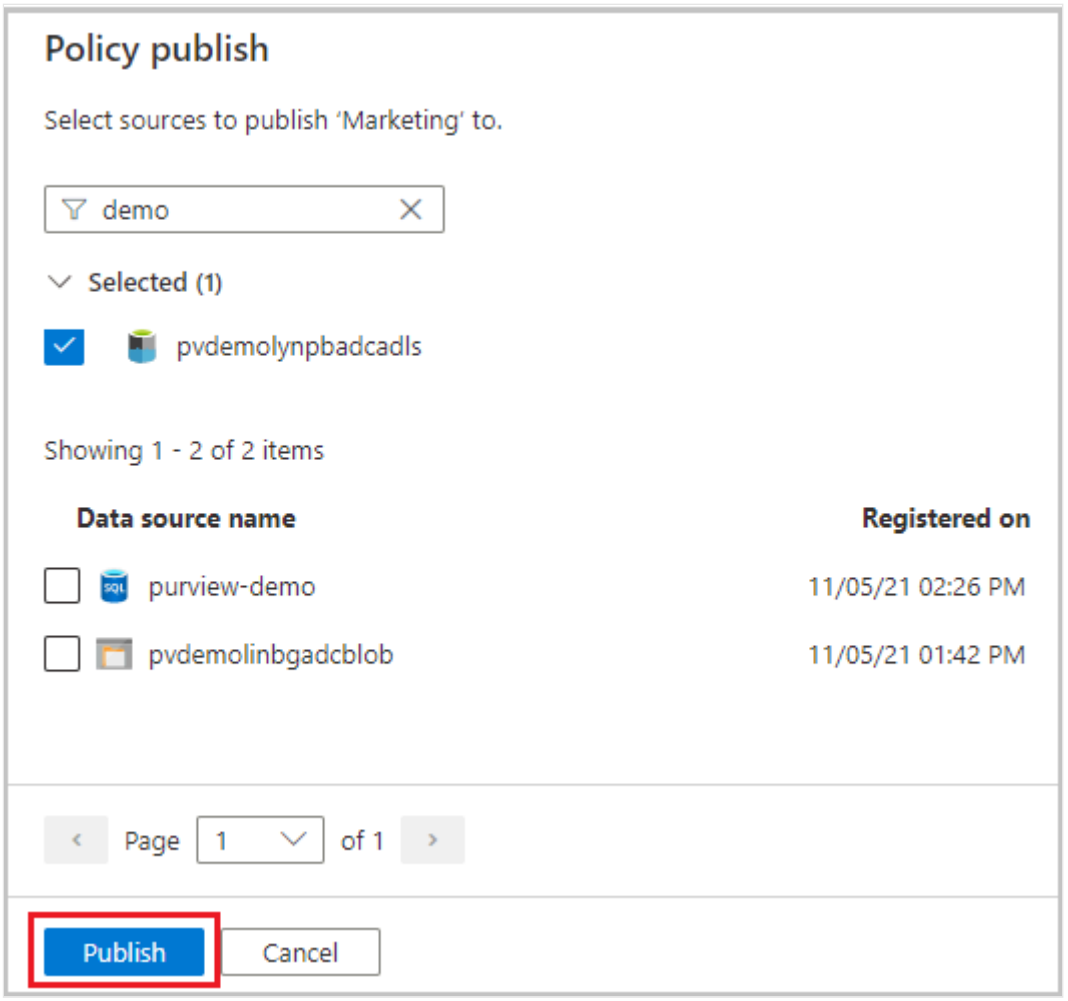
1. Sign in to the Microsoft Purview governance portal.
2. Navigate to the **Data policy** feature using the left side panel. Then select **Data policies**.



3. The Policy portal will present the list of existing policies in Microsoft Purview. Locate the policy that needs to be published. Select the **Publish** button on the right top corner of the page.



4. A list of data sources is displayed. You can enter a name to filter the list. Then, select each data source where this policy is to be published and then select the **Publish** button.



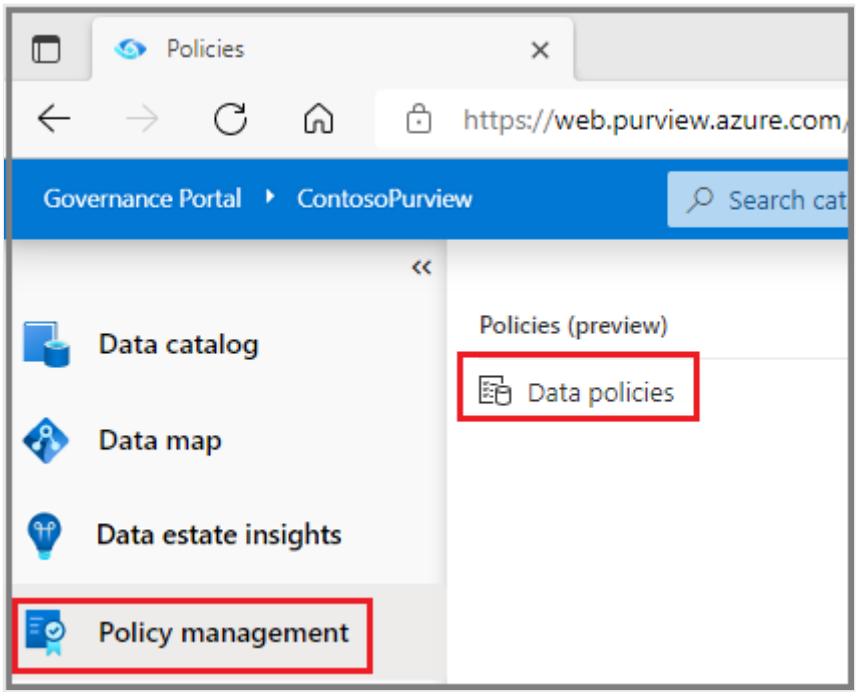
> ⓘ Note
>
> After making changes to a policy, there is no need to publish it again for it to take effect if the data source(s) continues to be the same.

Unpublish a policy

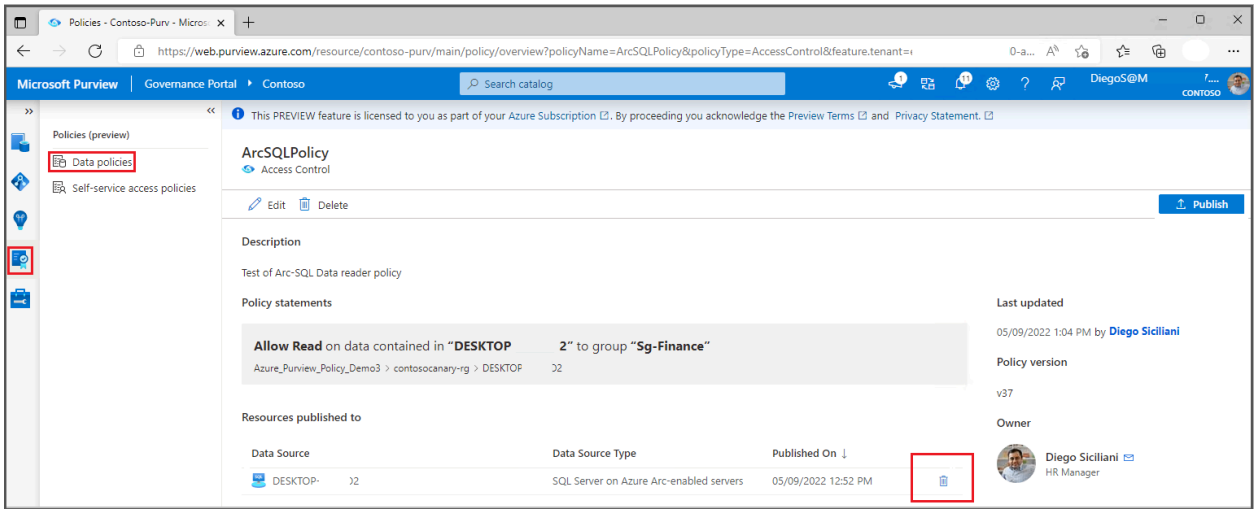
Ensure you have the *Data Source Admin* permission as described here

The steps to publish a policy are as follows:

1. Sign in to the Microsoft Purview governance portal.
2. Navigate to the **Data policy** feature using the left side panel. Then select **Data policies**.



3. The Policy portal will present the list of existing policies in Microsoft Purview. Locate the policy that needs to be unpublished. Select the trash can icon.

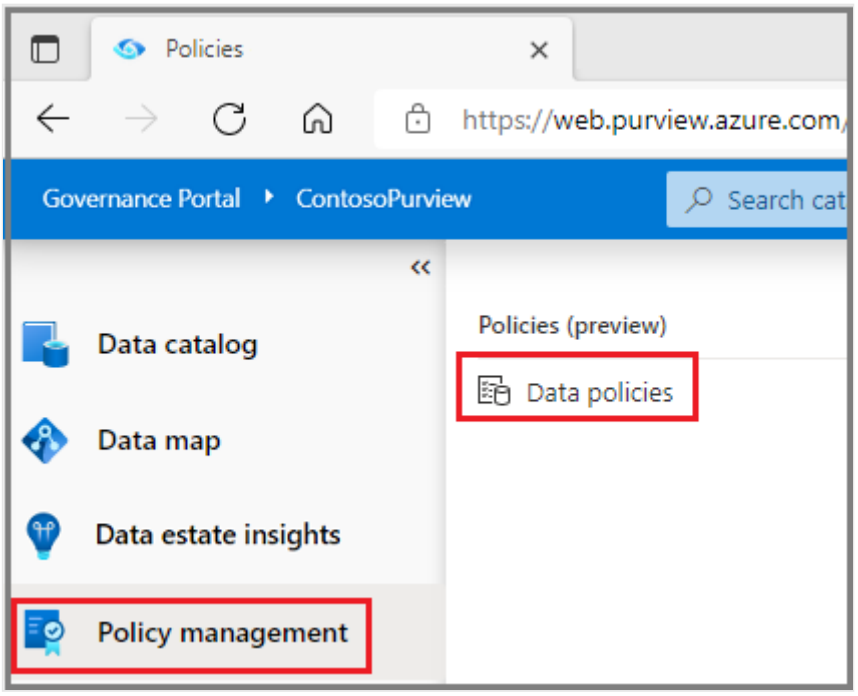


Update or delete a policy

Steps to update or delete a policy in Microsoft Purview are as follows. Ensure you have the *Policy Author* permission as described here

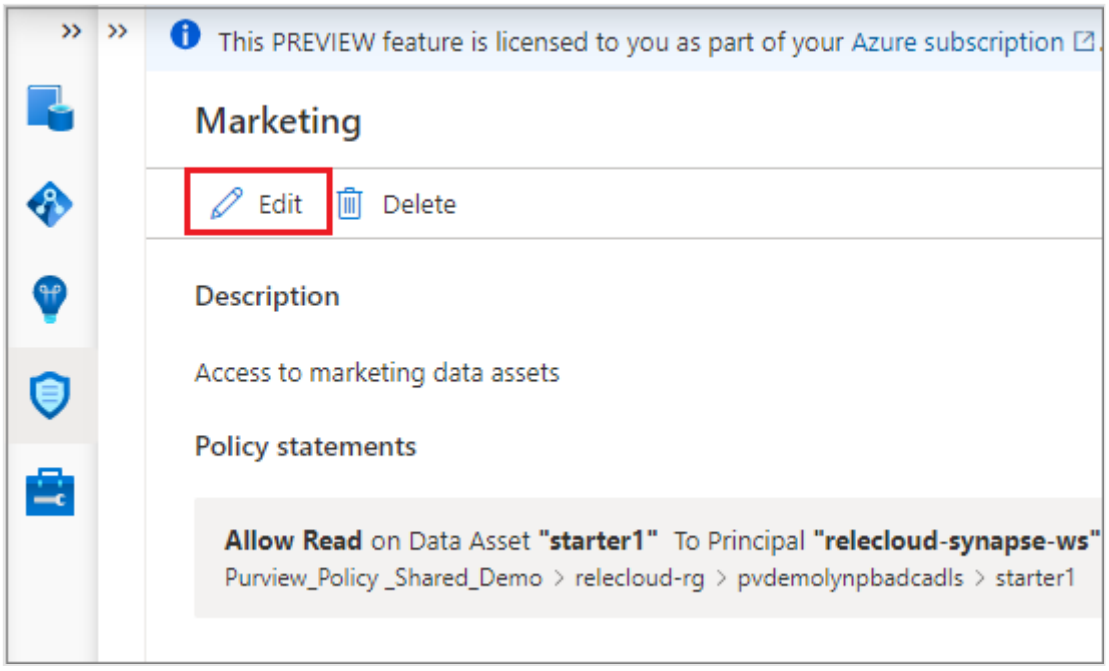
1. Sign in to the Microsoft Purview governance portal.

2. Navigate to the **Data policy** feature using the left side panel. Then select **Data policies**.



3. The Policy portal will present the list of existing policies in Microsoft Purview. Select the policy that needs to be updated.

4. The policy details page will appear, including Edit and Delete options. Select the **Edit** button, which brings up the policy statement builder. Now, any parts of the statement in this policy can be updated. To delete the policy, use the **Delete** button.



Next steps

For specific guides on creating policies, you can follow these tutorials:

- Doc: Microsoft Purview data owner policies on all data sources in a subscription or a resource group
- Doc: Microsoft Purview data owner policies on an Azure Storage account
- Blog: Grant users access to data assets in your enterprise via API

Feedback

Was this page helpful? Yes No

Provision read access to Azure Storage using Microsoft Purview Data owner policies (preview)

Article • 10/28/2024

> **Note**
>
> Microsoft Purview Data Map (classic), Data Catalog (classic), Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

> **Important**
>
> This feature is currently in preview. The Supplemental Terms of Use for Microsoft Azure Previews include additional legal terms that apply to Azure features that are in beta, in preview, or otherwise not yet released into general availability.

Data owner policies are a type of Microsoft Purview access policies. They allow you to manage access to user data in sources that have been registered for *Data Policy Enforcement* in Microsoft Purview. These policies can be authored directly in the Microsoft Purview governance portal, and after publishing, they get enforced by the data source.

This guide covers how a data owner can delegate in Microsoft Purview management of access to Azure Storage datasets. Currently, these two Azure Storage sources are supported:

- Blob storage
- Azure Data Lake Storage (ADLS) Gen2

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- A new or existing Microsoft Purview account. Follow this quickstart guide to create one.

Region support

- All Microsoft Purview regions are supported.
- Storage accounts in the following regions are supported without the need for additional configuration. However, zone-redundant storage (ZRS) accounts are not supported.
  - Australia Central
  - Australia East
  - Australia Southeast
  - Brazil South
  - Canada Central
  - Canada East
  - Central India
  - Central US
  - East Asia
  - East US 2
  - East US
  - France Central
  - Germany West Central
  - Japan East
  - Japan West
  - Korea Central
  - North Central US
  - North Europe
  - Norway East
  - Poland Central
  - Qatar Central
  - South Central US
  - South Africa North
  - Southeast Asia
  - South India
  - Sweden Central
  - Switzerland North
  - West Central US
  - West Europe
  - West US
  - West US 2
  - West US 3
  - UAE North
  - UK South
  - UK West

- Storage accounts in other regions in Public Cloud are supported after setting feature flag *AllowPurviewPolicyEnforcement*, as outlined in the next section. Newly created ZRS Storage accounts are supported, if created after setting the feature flag *AllowPurviewPolicyEnforcement*.

If needed, you can create a new Storage account by following this guide.

Configure the subscription where the Azure Storage account resides for policies from Microsoft Purview

This step is only necessary in certain regions (see prior section). To enable Microsoft Purview to manage policies for one or more Azure Storage accounts, execute the following PowerShell commands in the subscription where you'll deploy your Azure Storage account. These PowerShell commands will enable Microsoft Purview to manage policies on all Azure Storage accounts in that subscription.

If you're executing these commands locally, be sure to run PowerShell as an administrator. Alternatively, you can use the Azure Cloud Shell in the Azure portal: https://shell.azure.com.

PowerShell

```
# Install the Az module
Install-Module -Name Az -Scope CurrentUser -Repository PSGallery -Force
# Login into the subscription
Connect-AzAccount -Subscription <SubscriptionID>
# Register the feature
Register-AzProviderFeature -FeatureName AllowPurviewPolicyEnforcement -
ProviderNamespace Microsoft.Storage
```

If the output of the last command shows *RegistrationState* as *Registered*, then your subscription is enabled for access policies. If the output is *Registering*, wait at least 10 minutes, and then retry the command. **Do not continue unless the RegistrationState shows as *Registered***.

Microsoft Purview configuration

Register the data source in Microsoft Purview

Before a policy can be created in Microsoft Purview for a data resource, you must register that data resource in Microsoft Purview Studio. You will find the instructions related to registering the data resource later in this guide.

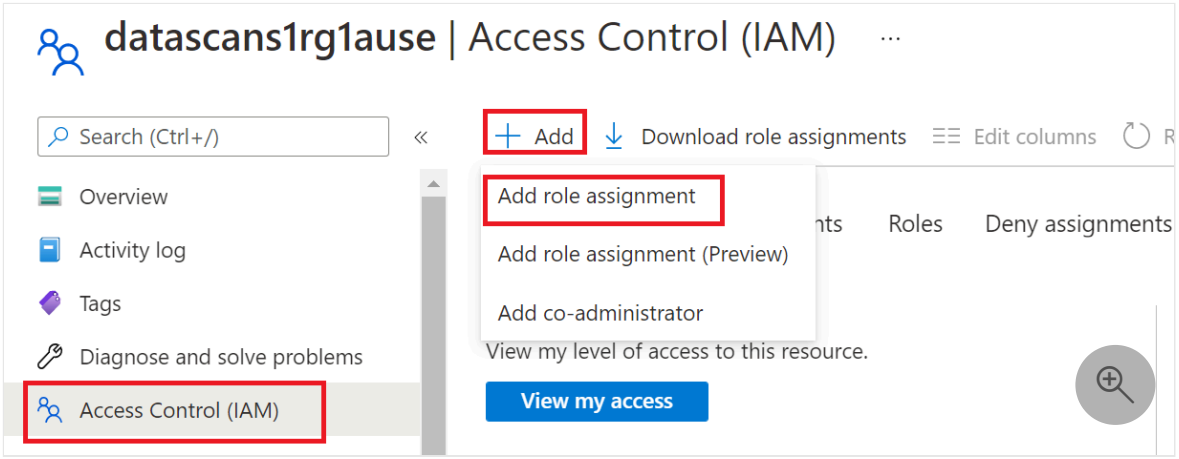
> **Note**
>
> Microsoft Purview policies rely on the data resource ARM path. If a data resource is moved to a new resource group or subscription it will need to be de-registered and then registered again in Microsoft Purview.

Configure permissions to enable Data policy enforcement on the data source

Once a resource is registered, but before a policy can be created in Microsoft Purview for that resource, you must configure permissions. A set of permissions are needed to enable the **Data policy enforcement**. This applies to data sources, resource groups, or subscriptions. To enable *Data policy enforcement*, you must have *both* specific Identity and Access Management (IAM) privileges on the resource as well as specific Microsoft Purview privileges:

- You must have *either one* of the following IAM role combinations on the resource's Azure Resource Manager path or any parent of it (that is, using IAM permission inheritance):
  - IAM Owner
  - Both IAM Contributor and IAM User Access Administrator

  To configure Azure role-based access control (RBAC) permissions, follow this guide. The following screenshot shows how to access the **Access Control** section in the Azure portal for the data resource to add a role assignment.

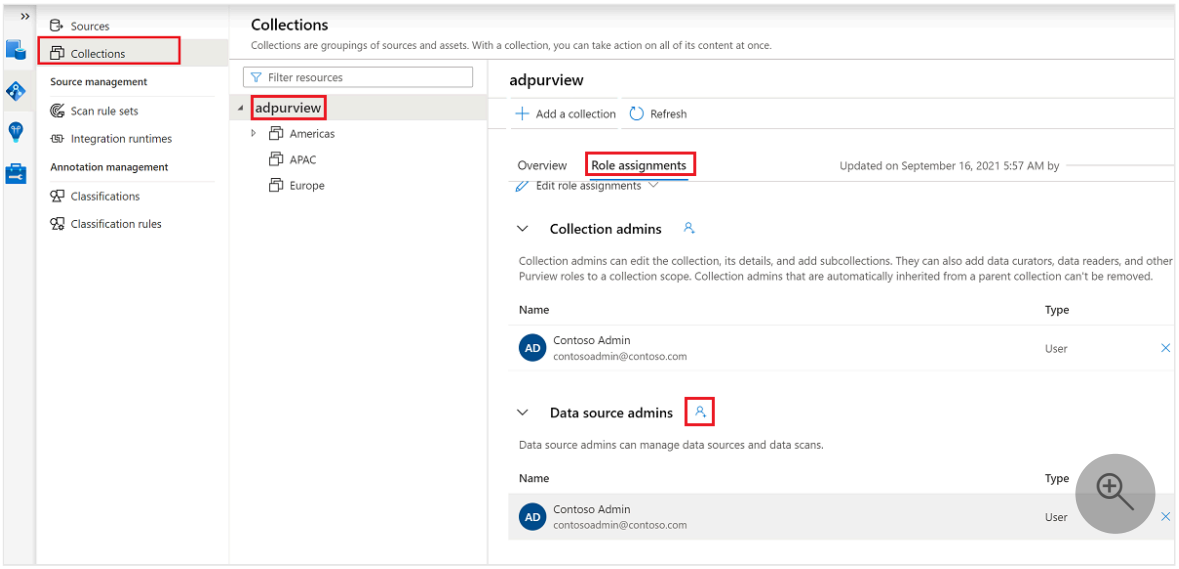


> **Note**
>
> The *IAM Owner* role for a data resource can be inherited from a parent resource group, a subscription, or a subscription management group. Check

which Microsoft Entra users, groups, and service principals hold or are inheriting the *IAM Owner* role for the resource.

- You also need to have the Microsoft Purview *Data source admin* role for the collection or a parent collection (if inheritance is enabled). For more information, see the guide on managing Microsoft Purview role assignments.

  The following screenshot shows how to assign the *Data source admin* role at the root collection level.



Configure Microsoft Purview permissions to create, update, or delete access policies

To create, update or delete policies, you need to get the *Policy author* role in Microsoft Purview at *root collection level*:

- The *Policy author* role can create, update, and delete DevOps and Data Owner policies.
- The *Policy author* role can delete self-service access policies.

For more information about managing Microsoft Purview role assignments, see Create and manage collections in the Microsoft Purview Data Map.

> **Note**
>
> Policy author role must be configured at the root collection level.

In addition, to easily search Microsoft Entra users or groups when creating or updating the subject of a policy, you can greatly benefit from getting the Directory Readers

permission in Microsoft Entra ID. This is a common permission for users in an Azure tenant. Without the Directory Reader permission, the Policy Author will have to type the complete username or email for all the principals included in the subject of a data policy.

Configure Microsoft Purview permissions for publishing Data Owner policies

Data Owner policies allow for checks and balances if you assign the Microsoft Purview *Policy author* and *Data source admin* roles to different people in the organization. Before a Data owner policy takes effect, a second person (*Data source admin*) must review it and explicitly approve it by publishing it. This does not apply to DevOps or Self-service access policies as publishing is automatic for them when those policies are created or updated.

To publish a Data owner policy you need to get the *Data source admin* role in Microsoft Purview at *root collection level*.

For more information about managing Microsoft Purview role assignments, see Create and manage collections in the Microsoft Purview Data Map.

> **Note**
>
> To publish Data owner policies, the Data source admin role must be configured at the root collection level.

Delegate access provisioning responsibility to roles in Microsoft Purview

After a resource has been enabled for **Data policy enforcement**, any Microsoft Purview user with the *Policy author* role at the root collection level can provision access to that data source from Microsoft Purview.

> **Note**
>
> Any Microsoft Purview root *Collection admin* can assign new users to root *Policy author* roles. Any *Collection admin* can assign new users to a *Data source admin* role under the collection. Minimize and carefully vet the users who hold Microsoft Purview *Collection admin*, *Data source admin*, or *Policy author* roles.

If a Microsoft Purview account with published policies is deleted, such policies will stop being enforced within an amount of time that depends on the specific data source. This change can have implications on both security and data access availability. The Contributor and Owner roles in IAM can delete Microsoft Purview accounts. You can check these permissions by going to the **Access control (IAM)** section for your Microsoft Purview account and selecting **Role Assignments**. You can also use a lock to prevent the Microsoft Purview account from being deleted through Resource Manager locks.

Register the data sources in Microsoft Purview for Data Policy Enforcement

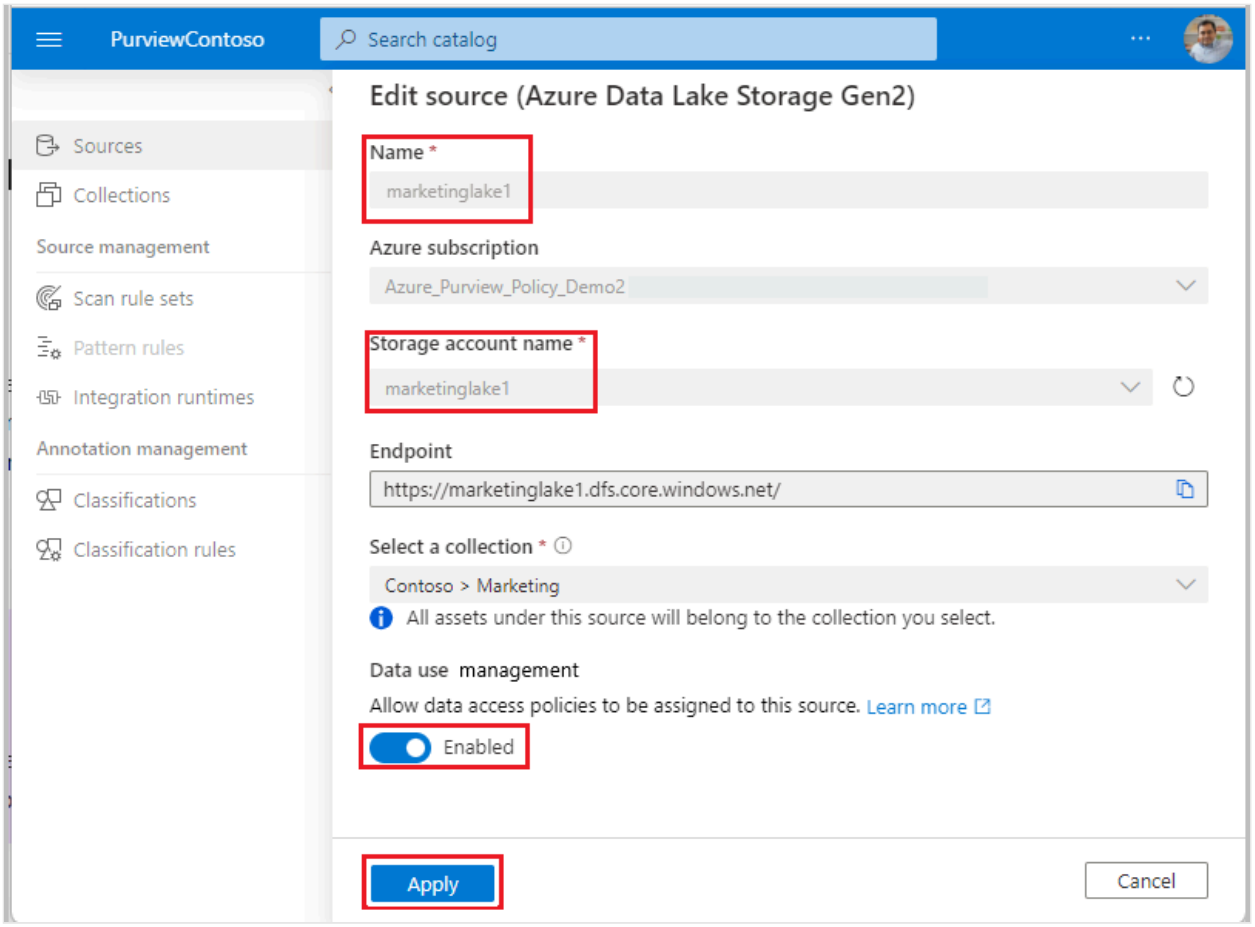
The Azure Storage resources need to be registered first with Microsoft Purview to later define access policies.

To register your resources, follow the **Prerequisites** and **Register** sections of these guides:

- Register and scan Azure Storage Blob - Microsoft Purview
- Register and scan Azure Data Lake Storage (ADLS) Gen2 - Microsoft Purview

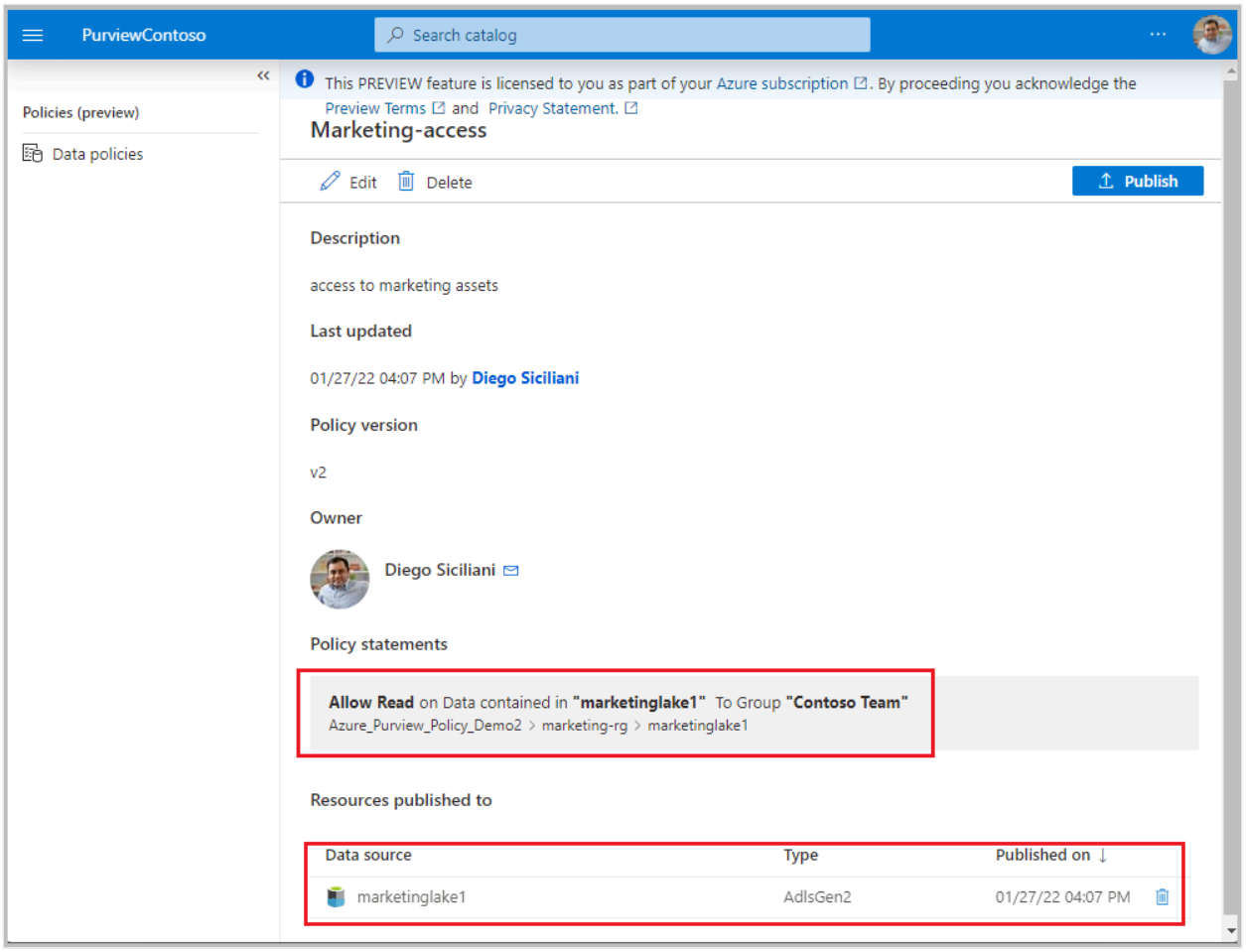
After you've registered your resources, you'll need to enable Data Policy Enforcement. Data Policy Enforcement needs certain permissions and can affect the security of your data, as it delegates to certain Microsoft Purview roles to manage access to the data sources. **Go through the secure practices related to Data Policy Enforcement in this guide**: How to enable Data Policy Enforcement

Once your data source has the **Data Policy Enforcement** toggle **Enabled**, it will look like this screenshot:



Create and publish a data owner policy

Execute the steps in the **Create a new policy** and **Publish a policy** sections of the data-owner policy authoring tutorial. The result will be a data owner policy similar to the example shown in the image: a policy that provides group *Contoso Team read* access to Storage account *marketinglake1*:



> **Important**
>
> - Publish is a background operation. Azure Storage accounts can take up to **2 hours** to reflect the changes.

Unpublish a data owner policy

Follow this link for the steps to unpublish a data owner policy in Microsoft Purview.

Update or delete a data owner policy

Follow this link for the steps to update or delete a data owner policy in Microsoft Purview.

Data Consumption

- Data consumer can access the requested dataset using tools such as Power BI or Azure Synapse Analytics workspace.

- The Copy and Clone commands in Azure Storage Explorer require other IAM permissions to work in addition to the Allow Modify policy from Purview. Provide Microsoft.Storage/storageAccounts/blobServices/generateUserDelegationKey/action permission in IAM to the Microsoft Entra principal.
- Subcontainer access: Policy statements set below container level on a Storage account are supported. However, users won't be able to browse to the data asset using Azure portal's Storage Browser or Microsoft Azure Storage Explorer tool if access is granted only at file or folder level of the Azure Storage account. This is because these apps attempt to crawl down the hierarchy starting at container level, and the request fails because no access has been granted at that level. Instead, the App that requests the data must execute a direct access by providing a fully qualified name to the data object. The following documents show examples of how to perform a direct access. See also the blogs in the *Next steps* section of this how-to-guide.
  - *abfs* for ADLS Gen2
  - *az storage blob download* for Blob Storage

Additional information

- Creating a policy at Storage account level enables the Subjects to access system containers, for example *$logs*. If this is undesired, first scan the data sources and then create finer-grained policies for each (that is, at container or subcontainer level).
- The root blob in a container will be accessible to the Microsoft Entra principals in a Microsoft Purview *allow*-type RBAC policy if the scope of such policy is either subscription, resource group, Storage account or container in Storage account.
- The root container in a Storage account will be accessible to the Microsoft Entra principals in a Microsoft Purview *allow*-type RBAC policy if the scope of such policy is either subscription, resource group, or Storage account.

Limits

- The limit for Microsoft Purview policies that can be enforced by Storage accounts is 100 MB per subscription, which roughly equates to 5,000 policies.

Known issues

Known issues related to Policy creation

- Don't create policy statements based on Microsoft Purview resource sets. Even if displayed in Microsoft Purview policy authoring UI, they aren't yet enforced. Learn more about resource sets.

Policy action mapping

This section contains a reference of how actions in Microsoft Purview data policies map to specific actions in Azure Storage.

Expand table

Microsoft Purview policy action	Data source specific actions
Read	Microsoft.Storage/storageAccounts/blobServices/containers/read
	Microsoft.Storage/storageAccounts/blobServices/containers/blobs/read
Modify	Microsoft.Storage/storageAccounts/blobServices/containers/blobs/read
	Microsoft.Storage/storageAccounts/blobServices/containers/blobs/write
	Microsoft.Storage/storageAccounts/blobServices/containers/blobs/add/action
	Microsoft.Storage/storageAccounts/blobServices/containers/blobs/move/action
	Microsoft.Storage/storageAccounts/blobServices/containers/blobs/delete
	Microsoft.Storage/storageAccounts/blobServices/containers/read
	Microsoft.Storage/storageAccounts/blobServices/containers/write
	Microsoft.Storage/storageAccounts/blobServices/containers/delete

Next steps

Check blog, demo, and related tutorials:

- Demo of access policy for Azure Storage
- Doc: Concepts for Microsoft Purview data owner policies
- Doc: Provision access to all data sources in a subscription or a resource group
- Blog: What's New in Microsoft Purview at Microsoft Ignite 2021
- Blog: Accessing data when folder level permission is granted

- Blog: Accessing data when file level permission is granted
- Blog: Grant users access to data assets in your enterprise via API

Feedback

Was this page helpful? Yes No

Provision read access to Azure SQL Database using Microsoft Purview Data owner policies (preview)

Article • 10/28/2024

> **Note**
>
> Microsoft Purview Data Map (classic), Data Catalog (classic), Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

> **Important**
>
> This feature is currently in preview. The Supplemental Terms of Use for Microsoft Azure Previews include additional legal terms that apply to Azure features that are in beta, in preview, or otherwise not yet released into general availability.

Data owner policies are a type of Microsoft Purview access policies. They allow you to manage access to user data in sources that have been registered for *Data Policy Enforcement* in Microsoft Purview. These policies can be authored directly in the Microsoft Purview governance portal, and after publishing, they get enforced by the data source.

This guide covers how a data owner can delegate authoring policies in Microsoft Purview to enable access to Azure SQL Database. The following actions are currently enabled: *Read*. *Modify* isn't supported at this point.

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- A new or existing Microsoft Purview account. Follow this quickstart guide to create one.
- Create a new Azure SQL Database instance, or use an existing one, in one of the currently available regions for this feature. You can follow this guide to create an Azure SQL Database instance.

Region support

All Microsoft Purview regions are supported.

Enforcement of Microsoft Purview policies is available only in the following regions for Azure SQL Database:

Public cloud:

- East US
- East US2
- South Central US
- West Central US
- West US3
- Canada Central
- Brazil South
- West Europe
- North Europe
- France Central
- UK South
- South Africa North
- Central India
- South East Asia
- East Asia
- Australia East

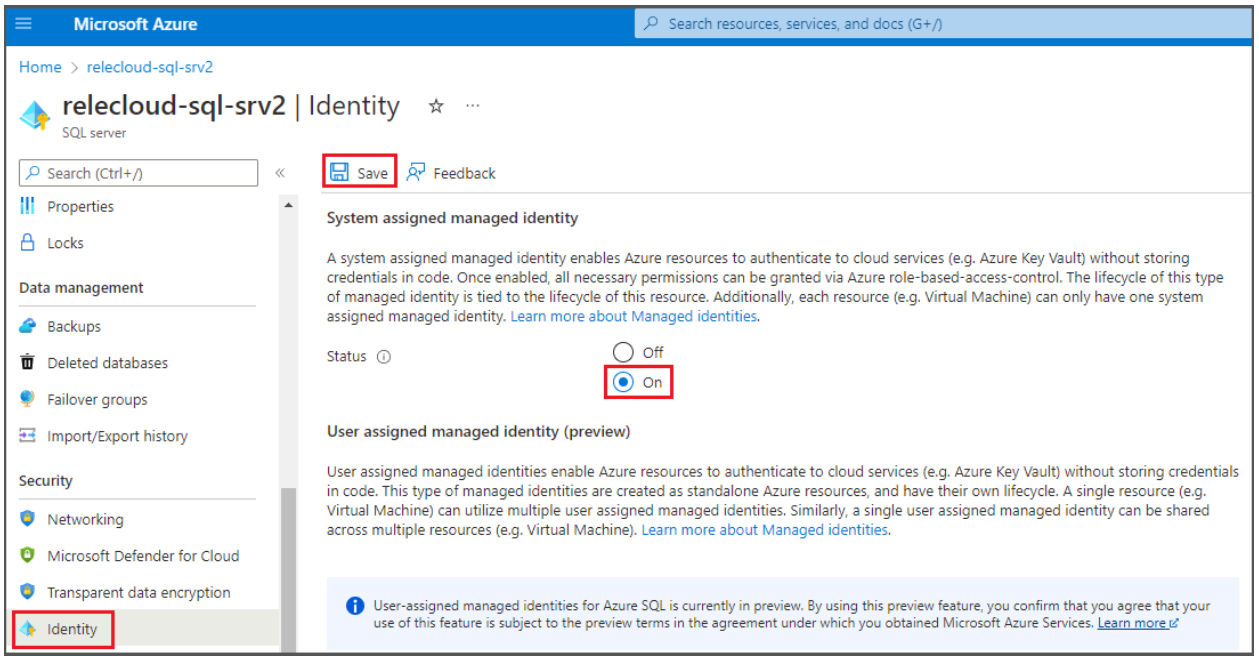
Sovereign clouds:

- USGov Virginia
- China North 3

Configure the Azure SQL Database instance for policies from Microsoft Purview

For the logical server associated with Azure SQL Database to honor policies from Microsoft Purview, you need to configure a Microsoft Entra admin. In the Azure portal, go to the logical server that hosts the Azure SQL Database instance. On the side menu, select **Microsoft Entra ID**. Set an admin name to any Microsoft Entra user or group that you prefer, and then select **Save**.

Then, on the side menu, select **Identity**. Under **System assigned managed identity**, turn the status to **On**, and then select **Save**.



Microsoft Purview configuration

Register the data source in Microsoft Purview

Before a policy can be created in Microsoft Purview for a data resource, you must register that data resource in Microsoft Purview Studio. You will find the instructions related to registering the data resource later in this guide.

> **Note**
>
> Microsoft Purview policies rely on the data resource ARM path. If a data resource is moved to a new resource group or subscription it will need to be de-registered and then registered again in Microsoft Purview.

Configure permissions to enable Data policy enforcement on the data source

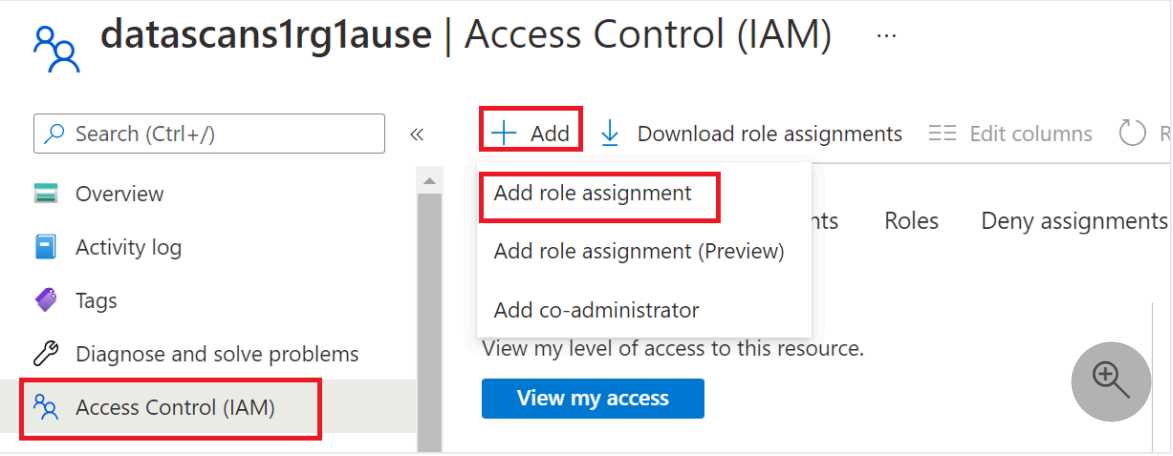
Once a resource is registered, but before a policy can be created in Microsoft Purview for that resource, you must configure permissions. A set of permissions are needed to enable the **Data policy enforcement**. This applies to data sources, resource groups, or subscriptions. To enable *Data policy enforcement*, you must have *both* specific Identity and Access Management (IAM) privileges on the resource as well as specific Microsoft Purview privileges:

- You must have *either one* of the following IAM role combinations on the resource's Azure Resource Manager path or any parent of it (that is, using IAM permission

inheritance):

- IAM Owner
- Both IAM Contributor and IAM User Access Administrator

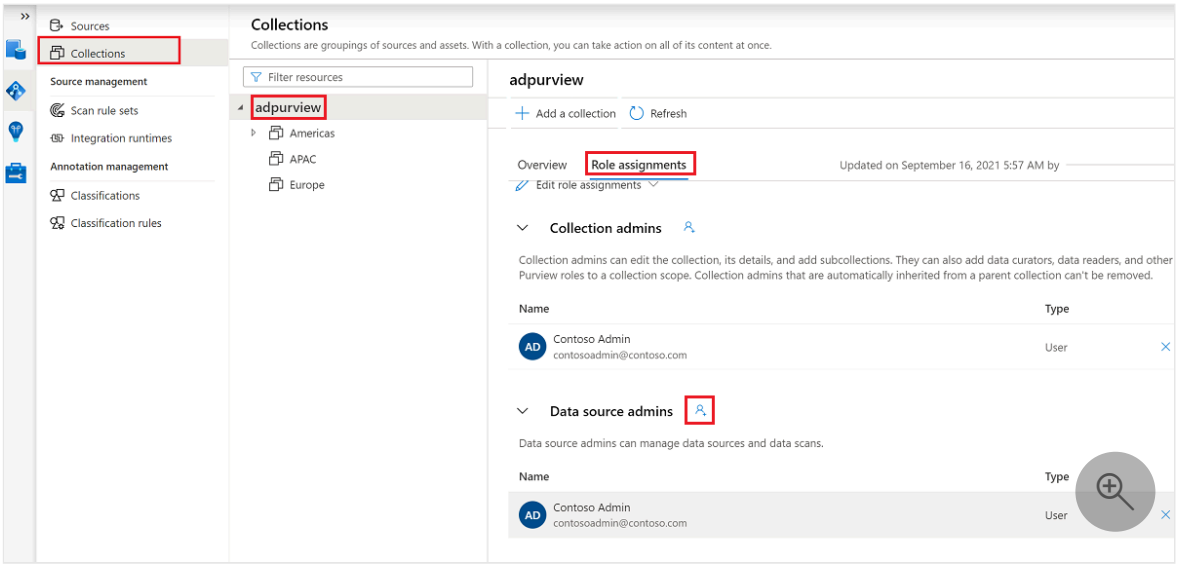
To configure Azure role-based access control (RBAC) permissions, follow this guide. The following screenshot shows how to access the **Access Control** section in the Azure portal for the data resource to add a role assignment.



> **Note**
>
> The *IAM Owner* role for a data resource can be inherited from a parent resource group, a subscription, or a subscription management group. Check which Microsoft Entra users, groups, and service principals hold or are inheriting the *IAM Owner* role for the resource.

- You also need to have the Microsoft Purview *Data source admin* role for the collection or a parent collection (if inheritance is enabled). For more information, see the guide on managing Microsoft Purview role assignments.

  The following screenshot shows how to assign the *Data source admin* role at the root collection level.



Configure Microsoft Purview permissions to create, update, or delete access policies

To create, update or delete policies, you need to get the *Policy author* role in Microsoft Purview at *root collection level*:

- The *Policy author* role can create, update, and delete DevOps and Data Owner policies.
- The *Policy author* role can delete self-service access policies.

For more information about managing Microsoft Purview role assignments, see Create and manage collections in the Microsoft Purview Data Map.

> **Note**
>
> Policy author role must be configured at the root collection level.

In addition, to easily search Microsoft Entra users or groups when creating or updating the subject of a policy, you can greatly benefit from getting the Directory Readers permission in Microsoft Entra ID. This is a common permission for users in an Azure tenant. Without the Directory Reader permission, the Policy Author will have to type the complete username or email for all the principals included in the subject of a data policy.

Configure Microsoft Purview permissions for publishing Data Owner policies

Data Owner policies allow for checks and balances if you assign the Microsoft Purview *Policy author* and *Data source admin* roles to different people in the organization. Before

a Data owner policy takes effect, a second person (*Data source admin*) must review it and explicitly approve it by publishing it. This does not apply to DevOps or Self-service access policies as publishing is automatic for them when those policies are created or updated.

To publish a Data owner policy you need to get the *Data source admin* role in Microsoft Purview at *root collection level*.

For more information about managing Microsoft Purview role assignments, see Create and manage collections in the Microsoft Purview Data Map.

> ⓘ **Note**
>
> To publish Data owner policies, the Data source admin role must be configured at the root collection level.

Delegate access provisioning responsibility to roles in Microsoft Purview

After a resource has been enabled for **Data policy enforcement**, any Microsoft Purview user with the *Policy author* role at the root collection level can provision access to that data source from Microsoft Purview.

> ⓘ **Note**
>
> Any Microsoft Purview root *Collection admin* can assign new users to root *Policy author* roles. Any *Collection admin* can assign new users to a *Data source admin* role under the collection. Minimize and carefully vet the users who hold Microsoft Purview *Collection admin*, *Data source admin*, or *Policy author* roles.

If a Microsoft Purview account with published policies is deleted, such policies will stop being enforced within an amount of time that depends on the specific data source. This change can have implications on both security and data access availability. The Contributor and Owner roles in IAM can delete Microsoft Purview accounts. You can check these permissions by going to the **Access control (IAM)** section for your Microsoft Purview account and selecting **Role Assignments**. You can also use a lock to prevent the Microsoft Purview account from being deleted through Resource Manager locks.

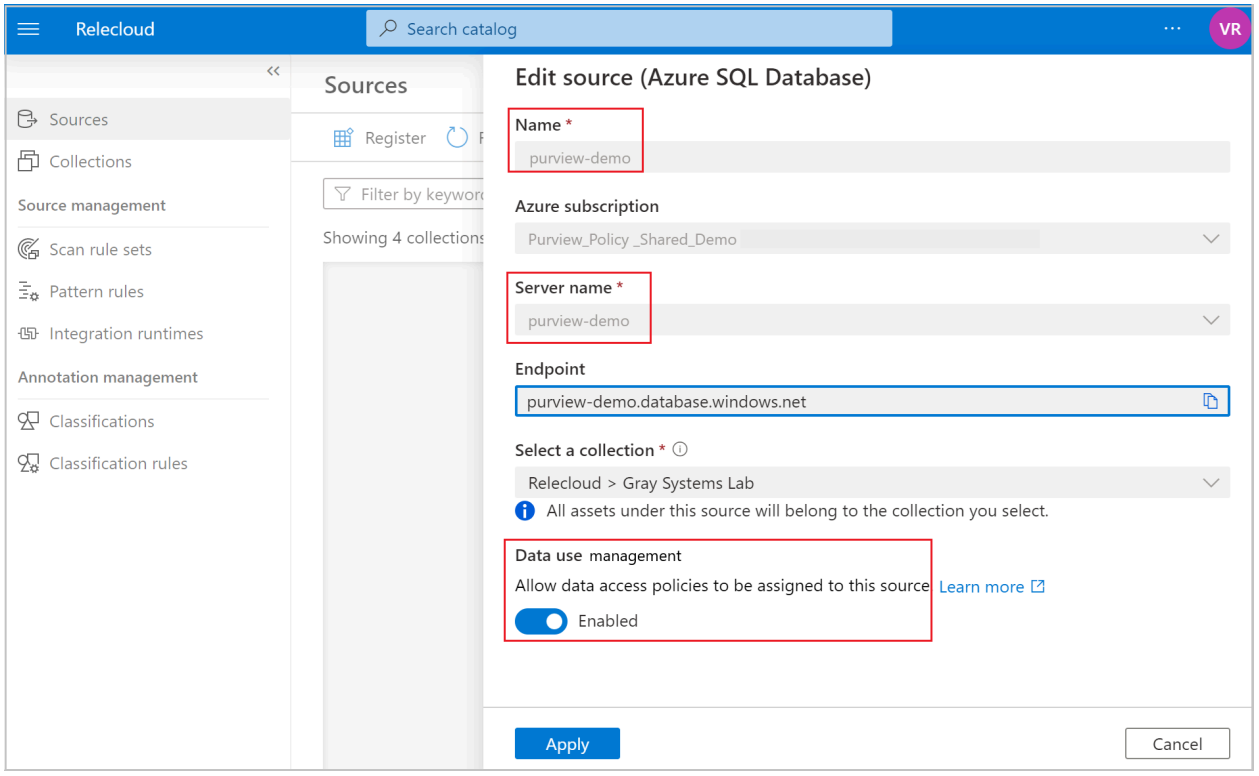
Register the data sources in Microsoft Purview

The Azure SQL Database data source needs to be registered first with Microsoft Purview before creating access policies. You can follow these guides:

Register and scan Azure SQL Database

After you've registered your resources, you'll need to enable Data Policy Enforcement. Data Policy Enforcement can affect the security of your data, as it delegates to certain Microsoft Purview roles to manage access to the data sources. **Go through the secure practices related to Data Policy Enforcement in this guide**: How to enable Data Policy Enforcement

Once your data source has the **Data policy enforcement** toggle *Enabled*, it will look like this screenshot. This will enable the access policies to be used with the given Azure SQL server and all its contained databases.



Return to the Azure portal for Azure SQL Database to verify it's now governed by Microsoft Purview:

1. Sign in to the Azure portal through this link
2. Select the Azure SQL Server that you want to configure.
3. Go to **Microsoft Entra ID** on the left pane.
4. Scroll down to **Microsoft Purview access policies**.
5. Select the button to **Check for Microsoft Purview Governance**. Wait while the request is processed. It could take a few minutes.



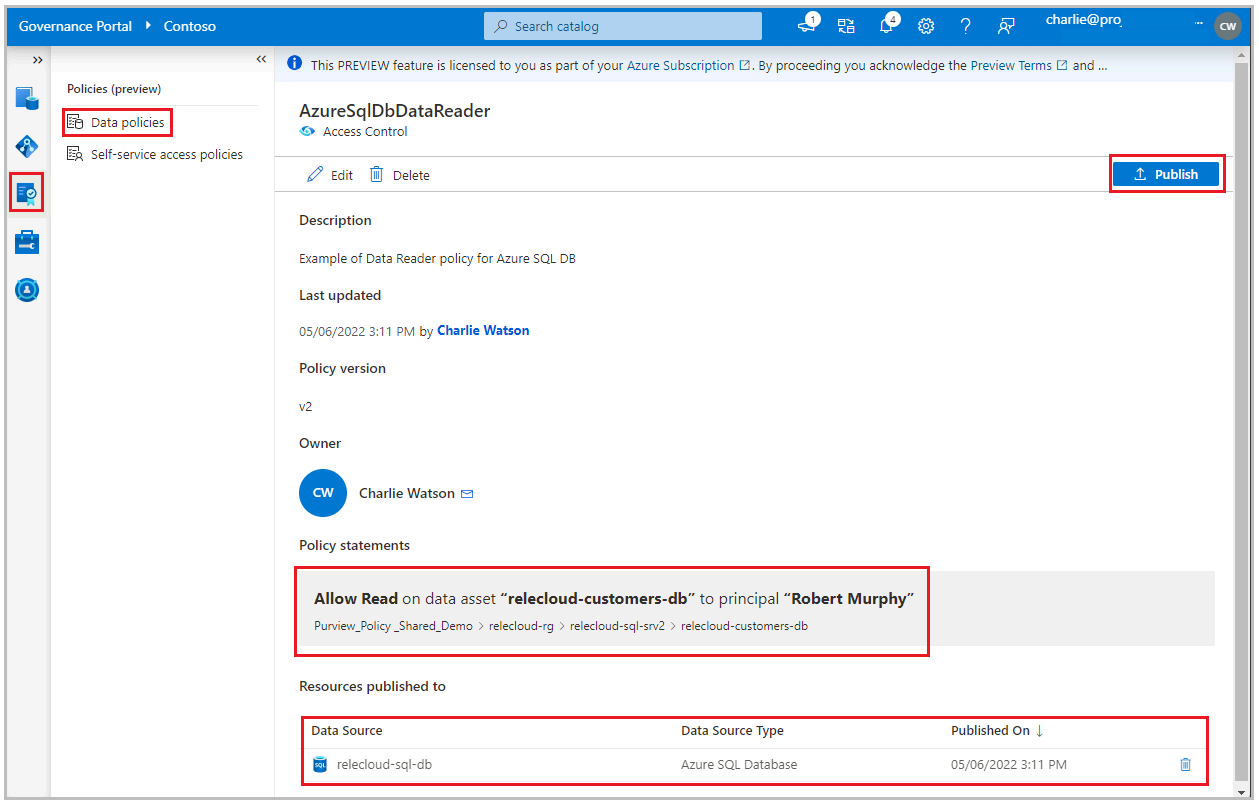
6. Confirm that the Microsoft Purview Governance Status shows `Governed`. Note that **it could take a few minutes** after you enable *Data policy enforcement* in Microsoft Purview for the correct status to be reflected.

> **Note**
>
> If you disable Data policy enforcement for this Azure SQL Database data source, it might take up to 24 hours for the Microsoft Purview Governance Status to be updated automatically to `Not Governed`. This can be accelerated by selecting **Check for Microsoft Purview Governance**. Before you enable *Data policy enforcement* for the data source in another Microsoft Purview account, ensure that the Purview Governance Status shows as `Not Governed`. Then repeat the steps above with the new Microsoft Purview account.

Create and publish a data owner policy

Execute the steps in the **Create a new policy** and **Publish a policy** sections of the data-owner policy authoring tutorial. The result will be a data owner policy similar to the example shown.

Example: Read policy. This policy assigns the Microsoft Entra principal 'Robert Murphy' to the *SQL Data reader* action, in the scope of SQL server *relecloud-sql-srv2*. This policy has also been published to that server. Note that policies related to this action are supported below server level (for example, database, table).



> **Important**
>
> - Publish is a background operation. It can take up to **5 minutes** for the changes to be reflected in this data source.
> - Changing a policy does not require a new publish operation. The changes will be picked up with the next pull.

Unpublish a data owner policy

Follow this link for the steps to unpublish a data owner policy in Microsoft Purview.

Update or delete a data owner policy

Follow this link for the steps to update or delete a data owner policy in Microsoft Purview.

Test the policy

After you create the policy, any of the Microsoft Entra users in the Subject should now be able to connect to the data sources in the scope of the policy. To test, use SSMS or any SQL client and try to query. Attempt access to a SQL table you have provided read access to.

If you require more troubleshooting, see the Next steps section in this guide.

Role definition detail

This section contains a reference of how relevant Microsoft Purview data policy roles map to specific actions in SQL data sources.

Expand table

Microsoft Purview policy role definition	Data source specific actions
Read	Microsoft.Sql/sqlservers/Connect
	Microsoft.Sql/sqlservers/databases/Connect
	Microsoft.Sql/Sqlservers/Databases/Schemas/Tables/Rows
	Microsoft.Sql/Sqlservers/Databases/Schemas/Views/Rows

Next steps

Check blog, demo, and related how-to guides.

- Doc: Concepts for Microsoft Purview data owner policies
- Doc: Microsoft Purview data owner policies on all data sources in a subscription or a resource group
- Doc: Microsoft Purview data owner policies on an Azure Arc-enabled SQL Server
- Doc: Troubleshoot Microsoft Purview policies for SQL data sources
- Blog: Grant users access to data assets in your enterprise via API

Feedback

Was this page helpful? Yes No

Provision read access to Azure Arc-enabled SQL Server 2022 using Microsoft Purview Data owner policies (preview)

Article • 02/24/2025

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

> **Important**
>
> This feature is currently in preview. The Supplemental Terms of Use for Microsoft Azure Previews include additional legal terms that apply to Azure features that are in beta, in preview, or otherwise not yet released into general availability.

Data owner policies are a type of Microsoft Purview access policies. They allow you to manage access to user data in sources that have been registered for *Data Policy Enforcement* in Microsoft Purview. These policies can be authored directly in the Microsoft Purview governance portal, and after publishing, they get enforced by the data source.

This guide covers how a data owner can delegate authoring policies in Microsoft Purview to enable access to Azure Arc-enabled SQL Server. The following actions are currently enabled: *Read*. This action is only supported for policies at server level. *Modify* is not supported at this point.

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- A new or existing Microsoft Purview account. Follow this quickstart guide to create one.

- Get SQL Server on-premises version 2022 and install it. Versions 2022 or more recent are supported on both Windows and Linux. You can try the free Developer edition.
- Configure your permissions and then register a list of resource providers in the subscription you'll use to onboard the SQL Server instance to Azure Arc.
- Complete prerequisites and onboard the SQL Server instance with Azure Arc. A simple configuration for Windows SQL Server is here. An alternate configuration for Linux SQL Server is here.
- Enable Microsoft Entra authentication in SQL Server. For a simpler setup, complete the prerequisites and process outlined in this article.
- Remember to grant application permissions and grant admin consent
- You need to configure a Microsoft Entra admin for the SQL Server instance, but it isn't necessary to set up other Microsoft Entra logins or users. You'll grant access to those users by using Microsoft Purview policies.

Region support

Data policy enforcement is available in all Microsoft Purview regions except:

- West US2
- East Asia
- US Gov Virginia
- China North 3

Security considerations for Azure Arc-enabled SQL Server

- The server admin can turn off the Microsoft Purview policy enforcement.
- Azure Arc admin and server admin permissions provide the ability to change the Azure Resource Manager path of the server. Because mappings in Microsoft Purview use Resource Manager paths, it can lead to wrong policy enforcements.
- A SQL Server admin (database admin) can gain the power of a server admin and can tamper with the cached policies from Microsoft Purview.
- The recommended configuration is to create a separate app registration for each SQL server instance. This configuration prevents the second SQL Server instance from reading the policies meant for the first SQL Server instance, in case a rogue admin in the second SQL Server instance tampers with the Resource Manager path.

Verify the prerequisites

1. Sign in to the Azure portal through this link

2. Navigate to **SQL servers** on the left pane. You'll see a list of SQL Server instances on Azure Arc.

3. Select the SQL Server instance that you want to configure.

4. Go to **Microsoft Entra ID** on the left pane.

5. Ensure that Microsoft Entra authentication is configured with an admin login. If not, refer to the access policy prerequisites section in this guide.

6. Ensure that a certificate has been provided to for SQL Server to authenticate to Azure. If not, refer to the access policy prerequisites section in this guide.

7. Ensure that an app registration has been entered to create a trust relationship between SQL Server and Microsoft Entra ID. If not, refer to the access policy prerequisites section in this guide.

8. If you made any changes, select the **Save** button to save the configuration and wait until the operation completes successfully. This could take a few minutes. The message "*Saved successfully*" will be displayed at the top of the page in green background. You might need to scroll up to see it.

Microsoft Purview configuration

Register the data source in Microsoft Purview

Before a policy can be created in Microsoft Purview for a data resource, you must register that data resource in Microsoft Purview Studio. You will find the instructions related to registering the data resource later in this guide.

> ⓘ **Note**
>
> Microsoft Purview policies rely on the data resource ARM path. If a data resource is moved to a new resource group or subscription it will need to be de-registered and then registered again in Microsoft Purview.

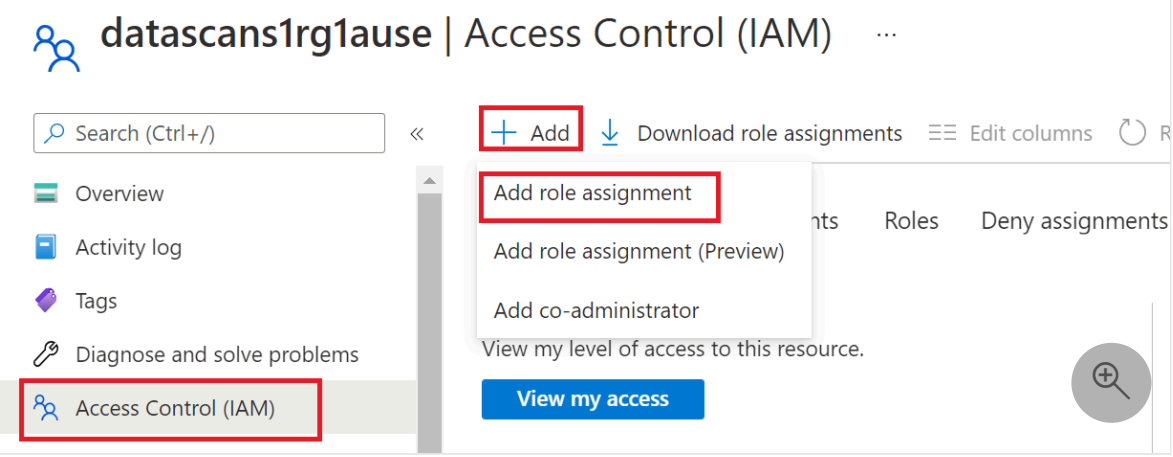
Configure permissions to enable Data policy enforcement on the data source

Once a resource is registered, but before a policy can be created in Microsoft Purview for that resource, you must configure permissions. A set of permissions are needed to

enable the **Data policy enforcement**. This applies to data sources, resource groups, or subscriptions. To enable *Data policy enforcement*, you must have *both* specific Identity and Access Management (IAM) privileges on the resource as well as specific Microsoft Purview privileges:

- You must have *either one* of the following IAM role combinations on the resource's Azure Resource Manager path or any parent of it (that is, using IAM permission inheritance):
  - IAM Owner
  - Both IAM Contributor and IAM User Access Administrator

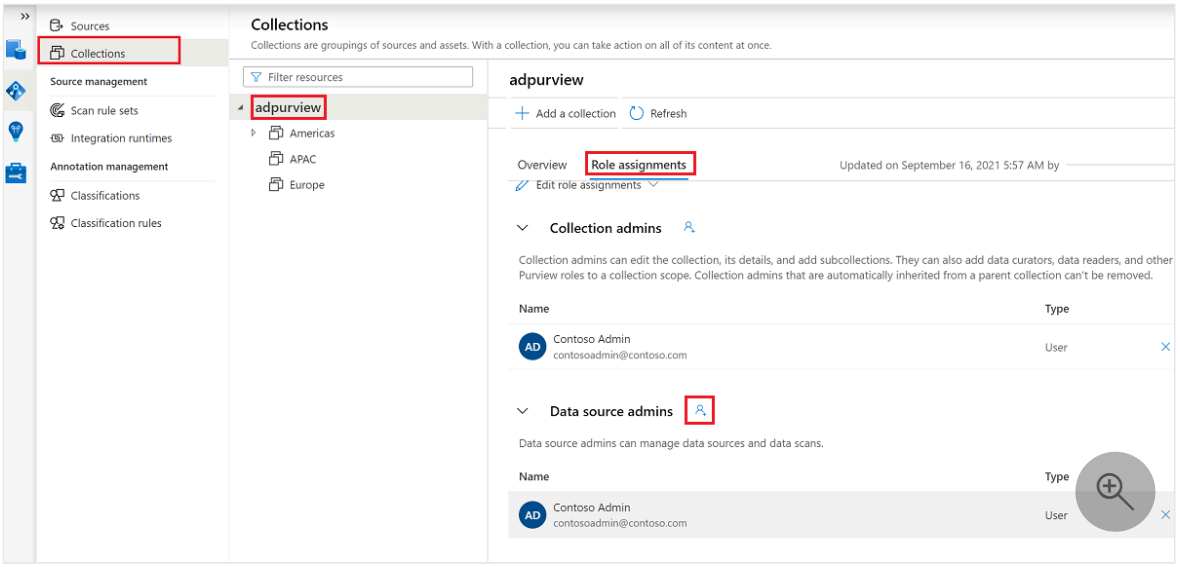
  To configure Azure role-based access control (RBAC) permissions, follow this guide. The following screenshot shows how to access the **Access Control** section in the Azure portal for the data resource to add a role assignment.



> **Note**
>
> The *IAM Owner* role for a data resource can be inherited from a parent resource group, a subscription, or a subscription management group. Check which Microsoft Entra users, groups, and service principals hold or are inheriting the *IAM Owner* role for the resource.

- You also need to have the Microsoft Purview *Data source admin* role for the collection or a parent collection (if inheritance is enabled). For more information, see the guide on managing Microsoft Purview role assignments.

  The following screenshot shows how to assign the *Data source admin* role at the root collection level.



Configure Microsoft Purview permissions to create, update, or delete access policies

To create, update or delete policies, you need to get the *Policy author* role in Microsoft Purview at *root collection level*:

- The *Policy author* role can create, update, and delete DevOps and Data Owner policies.
- The *Policy author* role can delete self-service access policies.

For more information about managing Microsoft Purview role assignments, see Create and manage collections in the Microsoft Purview Data Map.

> **Note**
>
> Policy author role must be configured at the root collection level.

In addition, to easily search Microsoft Entra users or groups when creating or updating the subject of a policy, you can greatly benefit from getting the Directory Readers permission in Microsoft Entra ID. This is a common permission for users in an Azure tenant. Without the Directory Reader permission, the Policy Author will have to type the complete username or email for all the principals included in the subject of a data policy.

Configure Microsoft Purview permissions for publishing Data Owner policies

Data Owner policies allow for checks and balances if you assign the Microsoft Purview *Policy author* and *Data source admin* roles to different people in the organization. Before

a Data owner policy takes effect, a second person (*Data source admin*) must review it and explicitly approve it by publishing it. This does not apply to DevOps or Self-service access policies as publishing is automatic for them when those policies are created or updated.

To publish a Data owner policy you need to get the *Data source admin* role in Microsoft Purview at *root collection level*.

For more information about managing Microsoft Purview role assignments, see Create and manage collections in the Microsoft Purview Data Map.

> ⓘ **Note**
>
> To publish Data owner policies, the Data source admin role must be configured at the root collection level.

Delegate access provisioning responsibility to roles in Microsoft Purview

After a resource has been enabled for **Data policy enforcement**, any Microsoft Purview user with the *Policy author* role at the root collection level can provision access to that data source from Microsoft Purview.

> ⓘ **Note**
>
> Any Microsoft Purview root *Collection admin* can assign new users to root *Policy author* roles. Any *Collection admin* can assign new users to a *Data source admin* role under the collection. Minimize and carefully vet the users who hold Microsoft Purview *Collection admin*, *Data source admin*, or *Policy author* roles.

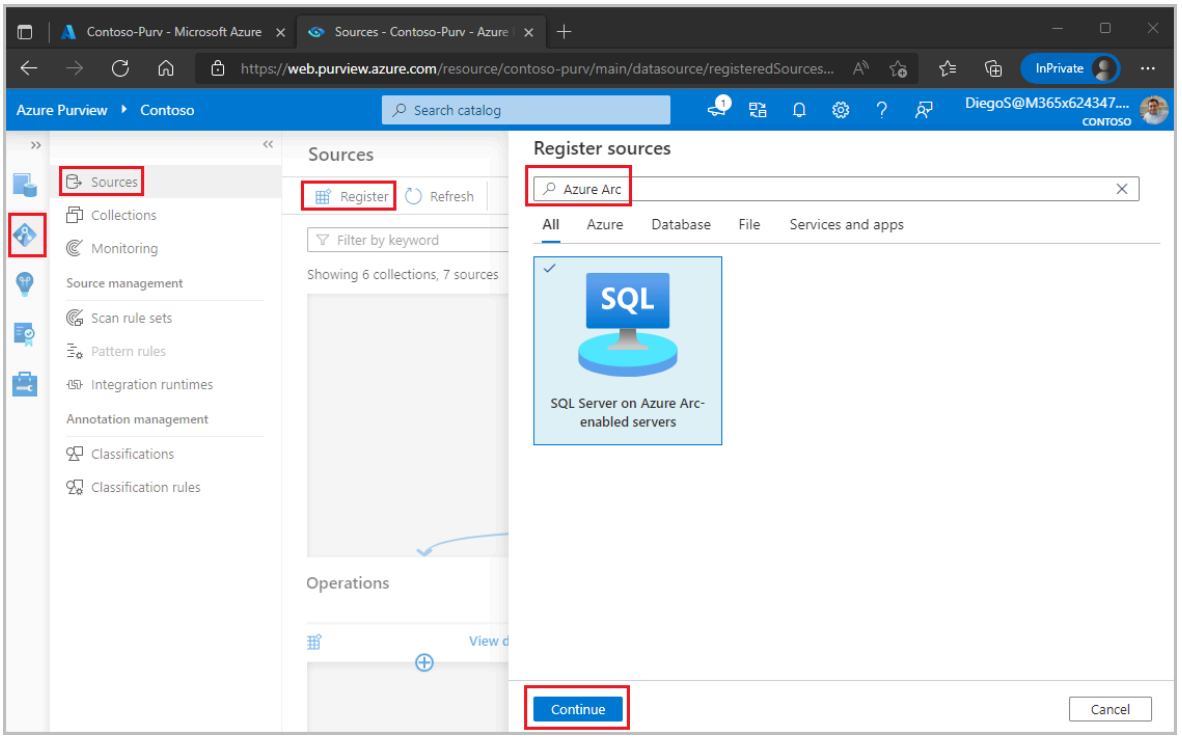
If a Microsoft Purview account with published policies is deleted, such policies will stop being enforced within an amount of time that depends on the specific data source. This change can have implications on both security and data access availability. The Contributor and Owner roles in IAM can delete Microsoft Purview accounts. You can check these permissions by going to the **Access control (IAM)** section for your Microsoft Purview account and selecting **Role Assignments**. You can also use a lock to prevent the Microsoft Purview account from being deleted through Resource Manager locks.

Register data sources in Microsoft Purview

Register each data source with Microsoft Purview to later define access policies.

1. Sign in to Microsoft Purview Studio.

2. Navigate to the **Data map** feature on the left pane, select **Sources**, then select **Register**. Type "Azure Arc" in the search box and select **SQL Server on Azure Arc**. Then select **Continue**.



3. Enter a **Name** for this registration. It is best practice to make the name of the registration the same as the server name in the next step.

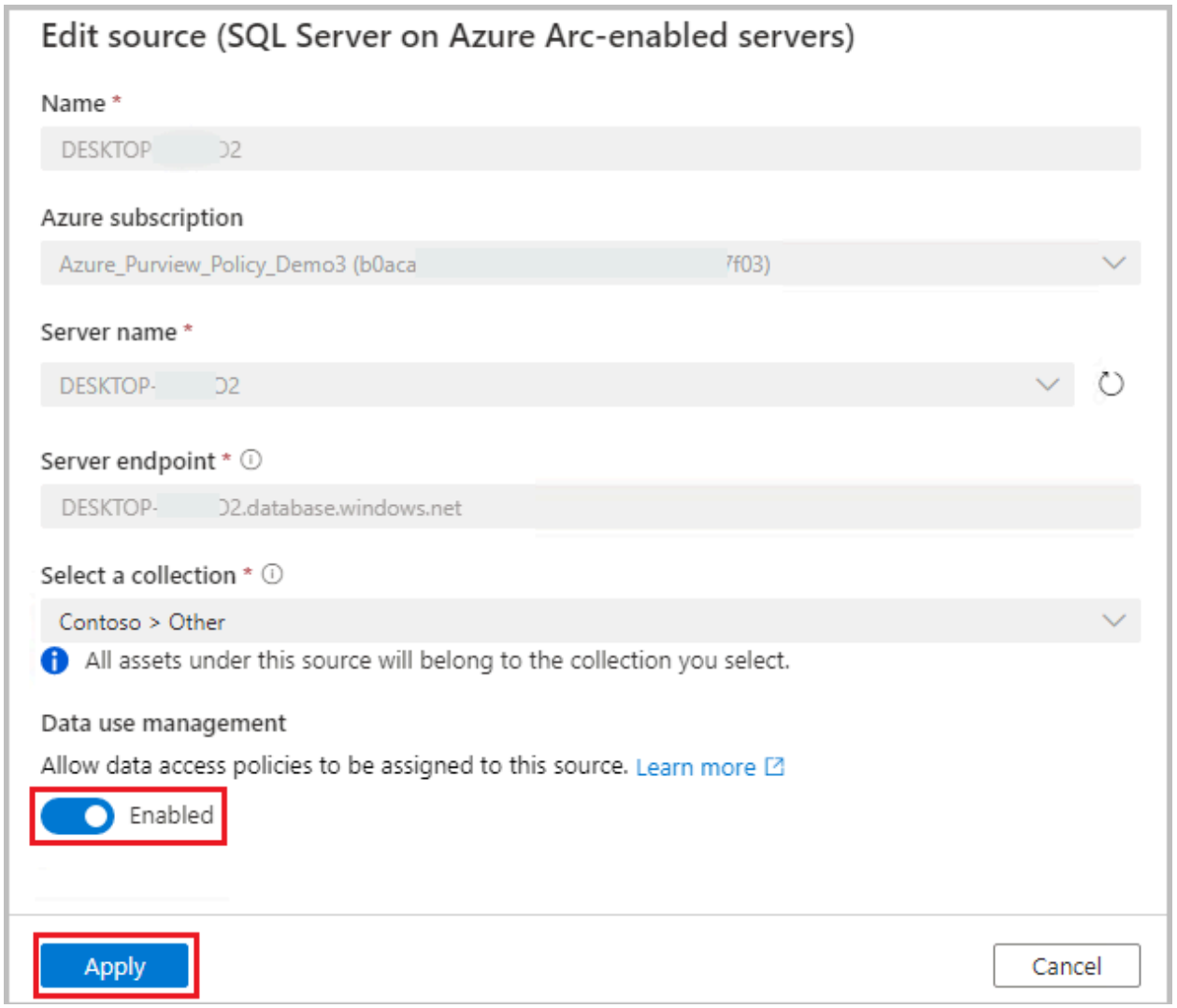
4. Select an **Azure subscription**, **Server name** and **Server endpoint**.

5. **Select a collection** to put this registration in.

6. Enable Data Policy Enforcement. Data Policy Enforcement needs certain permissions and can affect the security of your data, as it delegates to certain Microsoft Purview roles to manage access to the data sources. **Go through the secure practices related to Data Policy Enforcement in this guide**: How to enable Data Policy Enforcement

7. Select **Register** or **Apply** at the bottom.

Once your data source has the **Data policy enforcement** toggle *Enabled*, it will look like this picture.

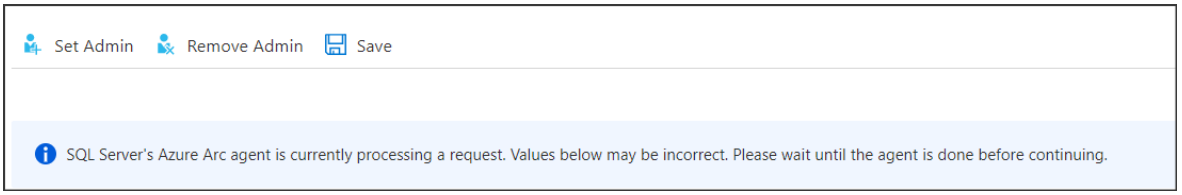


Enable policies in Azure Arc-enabled SQL Server

This section describes the steps to configure SQL Server on Azure Arc to use Microsoft Purview. Execute these steps after you enable the *Data policy enforcement* option for this data source in the Microsoft Purview account.

1. Sign in to the Azure portal through this link
2. Navigate to **SQL servers** on the left pane. You'll see a list of SQL Server instances on Azure Arc.
3. Select the SQL Server instance that you want to configure.
4. Go to **Microsoft Entra ID** on the left pane.
5. Scroll down to **Microsoft Purview access policies**.
6. Select the button to **Check for Microsoft Purview Governance**. Wait while the request is processed. While that happens, this message will be displayed at the top

of the page. You might need to scroll up to see it.



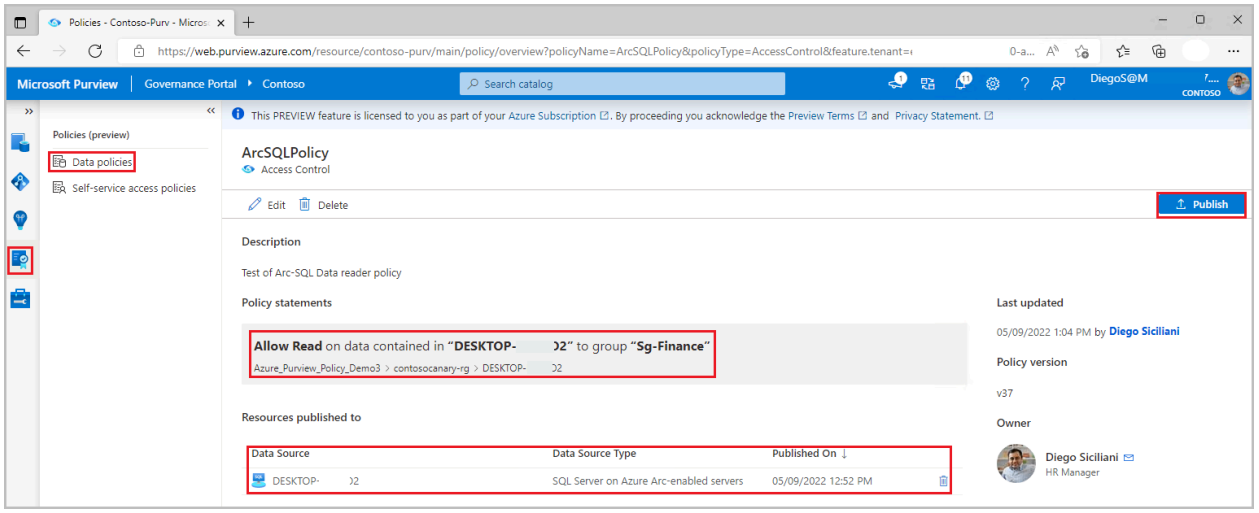
7. At the bottom of the page, confirm that the Microsoft Purview Governance Status shows `Governed`. Note that **it may take up to 30 minutes** for the correct status to be reflected. Continue doing a browser refresh until that happens.

8. Confirm that the Microsoft Purview Endpoint points to the Microsoft Purview account where you registered this data source and enabled the *data policy enforcement*

Create and publish a Data owner policy

Execute the steps in the **Create a new policy** and **Publish a policy** sections of the data-owner policy authoring tutorial. The result will be a data owner policy similar to the example:

Example: Read policy. This policy assigns the Microsoft Entra principal 'sg-Finance' to the *SQL Data reader* action, in the scope of SQL server *DESKTOP-xxx*. This policy has also been published to that server. Note that policies related to this action are not supported below server level.



> **Note**
>
> - Given that scan is not currently available for this data source, data reader policies can only be created at server level. Use the **Data sources** box instead of the Asset box when authoring the **data resources** part of the policy.

- There is a know issue with SQL Server Management Studio that prevents right-clicking on a table and choosing option “Select Top 1000 rows”.

> **Important**
>
> - Publish is a background operation. It can take up to **5 minutes** for the changes to be reflected in this data source.
> - Changing a policy does not require a new publish operation. The changes will be picked up with the next pull.

Unpublish a data owner policy

Follow this link for the steps to unpublish a data owner policy in Microsoft Purview.

Update or delete a data owner policy

Follow this link for the steps to update or delete a data owner policy in Microsoft Purview.

Test the policy

After the policy is published and has been communicated to the data source, any of the Microsoft Entra users in the Subject should be able to connect and execute a query that targets the assets to which the policy granted access. To test, use SSMS or any SQL client and try to query. For example, attempt access to a SQL table you have provided read access to.

If you require additional troubleshooting, see the Next steps section in this guide.

Role definition detail

This section contains a reference of how relevant Microsoft Purview data policy roles map to specific actions in SQL data sources.

Expand table

Microsoft Purview policy role definition	Data source specific actions
Read	Microsoft.Sql/sqlservers/Connect
	Microsoft.Sql/sqlservers/databases/Connect
	Microsoft.Sql/Sqlservers/Databases/Schemas/Tables/Rows
	Microsoft.Sql/Sqlservers/Databases/Schemas/Views/Rows

Next steps

Check blog, demo and related how-to guides

- Doc: Concepts for Microsoft Purview data owner policies
- Doc: Microsoft Purview data owner policies on all data sources in a subscription or a resource group
- Doc: Microsoft Purview data owner policies on an Azure SQL Database
- Doc: Troubleshoot Microsoft Purview policies for SQL data sources
- Blog: Grant users access to data assets in your enterprise via API

Feedback

Was this page helpful? Yes No

Provision access to entire resource groups or subscriptions using Microsoft Purview Data owner policies (preview)

Article • 02/24/2025

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

> **Important**
>
> This feature is currently in preview. The Supplemental Terms of Use for Microsoft Azure Previews include additional legal terms that apply to Azure features that are in beta, in preview, or otherwise not yet released into general availability.

Data owner policies are a type of Microsoft Purview access policies. They allow you to manage access to user data in sources that have been registered for *Data Policy Enforcement* in Microsoft Purview. These policies can be authored directly in the Microsoft Purview governance portal, and after publishing, they get enforced by the data source.

In this guide we cover how to register an entire resource group or subscription and then create a single policy that will manage access to **all** data sources in that resource group or subscription. That single policy will cover all existing data sources and any data sources that are created afterwards.

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- A new or existing Microsoft Purview account. Follow this quickstart guide to create one.

Only these data sources are enabled for access policies on resource group or subscription. Follow the **Prerequisites** section that is specific to the data source(s) in these guides:

- Data owner policies on an Azure Storage account
- Data owner policies on an Azure SQL Database(*)
- Data owner policies on an Azure Arc-enabled SQL Server(*)

(*) The *Modify* action is not currently supported for SQL-type data sources.

Microsoft Purview configuration

Register the data source in Microsoft Purview

Before a policy can be created in Microsoft Purview for a data resource, you must register that data resource in Microsoft Purview Studio. You will find the instructions related to registering the data resource later in this guide.

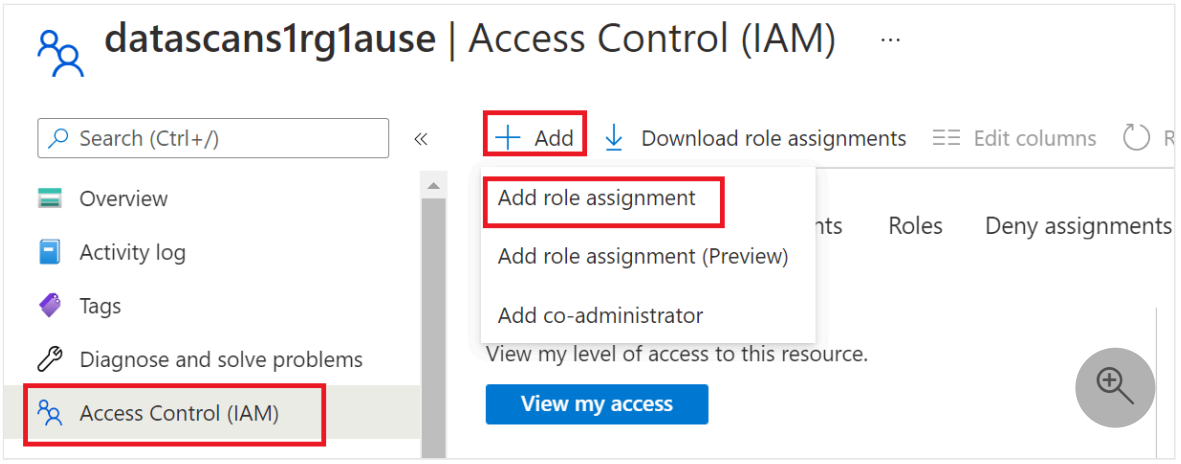
> ⓘ **Note**
>
> Microsoft Purview policies rely on the data resource ARM path. If a data resource is moved to a new resource group or subscription it will need to be de-registered and then registered again in Microsoft Purview.

Configure permissions to enable Data policy enforcement on the data source

Once a resource is registered, but before a policy can be created in Microsoft Purview for that resource, you must configure permissions. A set of permissions are needed to enable the **Data policy enforcement**. This applies to data sources, resource groups, or subscriptions. To enable *Data policy enforcement*, you must have *both* specific Identity and Access Management (IAM) privileges on the resource as well as specific Microsoft Purview privileges:

- You must have *either one* of the following IAM role combinations on the resource's Azure Resource Manager path or any parent of it (that is, using IAM permission inheritance):
  - IAM Owner
  - Both IAM Contributor and IAM User Access Administrator

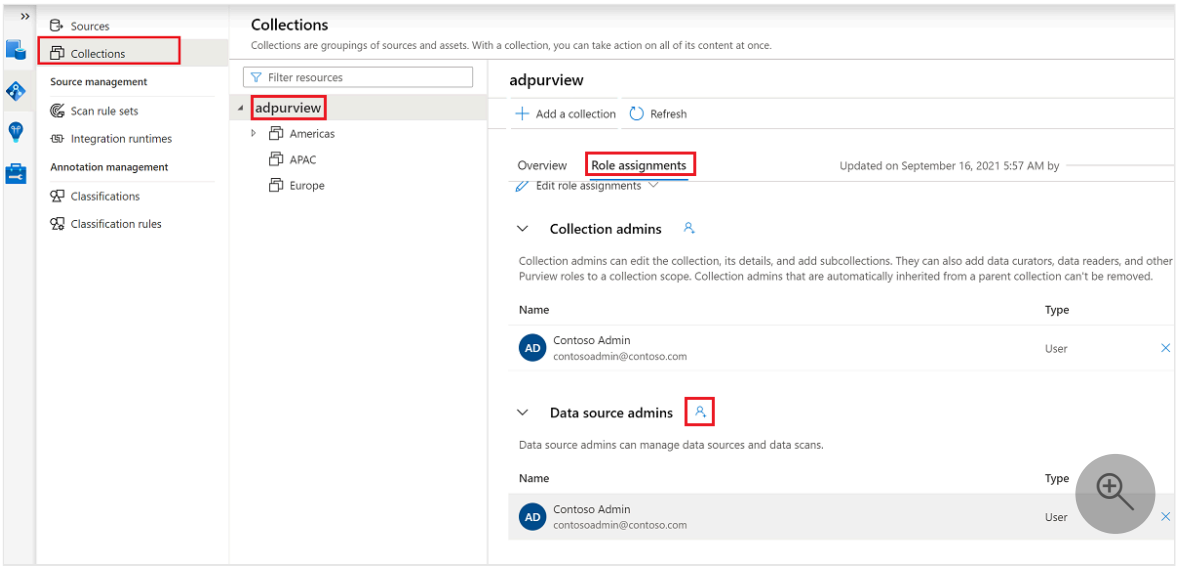
  To configure Azure role-based access control (RBAC) permissions, follow this guide. The following screenshot shows how to access the **Access Control** section in the Azure portal for the data resource to add a role assignment.



> **Note**
>
> The *IAM Owner* role for a data resource can be inherited from a parent resource group, a subscription, or a subscription management group. Check which Microsoft Entra users, groups, and service principals hold or are inheriting the *IAM Owner* role for the resource.

- You also need to have the Microsoft Purview *Data source admin* role for the collection or a parent collection (if inheritance is enabled). For more information, see the guide on managing Microsoft Purview role assignments.

  The following screenshot shows how to assign the *Data source admin* role at the root collection level.



Configure Microsoft Purview permissions to create, update, or delete access policies

To create, update or delete policies, you need to get the *Policy author* role in Microsoft Purview at *root collection level*:

- The *Policy author* role can create, update, and delete DevOps and Data Owner policies.
- The *Policy author* role can delete self-service access policies.

For more information about managing Microsoft Purview role assignments, see Create and manage collections in the Microsoft Purview Data Map.

> **Note**
>
> Policy author role must be configured at the root collection level.

In addition, to easily search Microsoft Entra users or groups when creating or updating the subject of a policy, you can greatly benefit from getting the Directory Readers permission in Microsoft Entra ID. This is a common permission for users in an Azure tenant. Without the Directory Reader permission, the Policy Author will have to type the complete username or email for all the principals included in the subject of a data policy.

Configure Microsoft Purview permissions for publishing Data Owner policies

Data Owner policies allow for checks and balances if you assign the Microsoft Purview *Policy author* and *Data source admin* roles to different people in the organization. Before a Data owner policy takes effect, a second person (*Data source admin*) must review it and explicitly approve it by publishing it. This does not apply to DevOps or Self-service access policies as publishing is automatic for them when those policies are created or updated.

To publish a Data owner policy you need to get the *Data source admin* role in Microsoft Purview at *root collection level*.

For more information about managing Microsoft Purview role assignments, see Create and manage collections in the Microsoft Purview Data Map.

> **Note**
>
> To publish Data owner policies, the Data source admin role must be configured at the root collection level.

Delegate access provisioning responsibility to roles in Microsoft Purview

After a resource has been enabled for **Data policy enforcement**, any Microsoft Purview user with the *Policy author* role at the root collection level can provision access to that data source from Microsoft Purview.

> **Note**
>
> Any Microsoft Purview root *Collection admin* can assign new users to root *Policy author* roles. Any *Collection admin* can assign new users to a *Data source admin* role under the collection. Minimize and carefully vet the users who hold Microsoft Purview *Collection admin*, *Data source admin*, or *Policy author* roles.

If a Microsoft Purview account with published policies is deleted, such policies will stop being enforced within an amount of time that depends on the specific data source. This change can have implications on both security and data access availability. The Contributor and Owner roles in IAM can delete Microsoft Purview accounts. You can check these permissions by going to the **Access control (IAM)** section for your Microsoft Purview account and selecting **Role Assignments**. You can also use a lock to prevent the Microsoft Purview account from being deleted through Resource Manager locks.

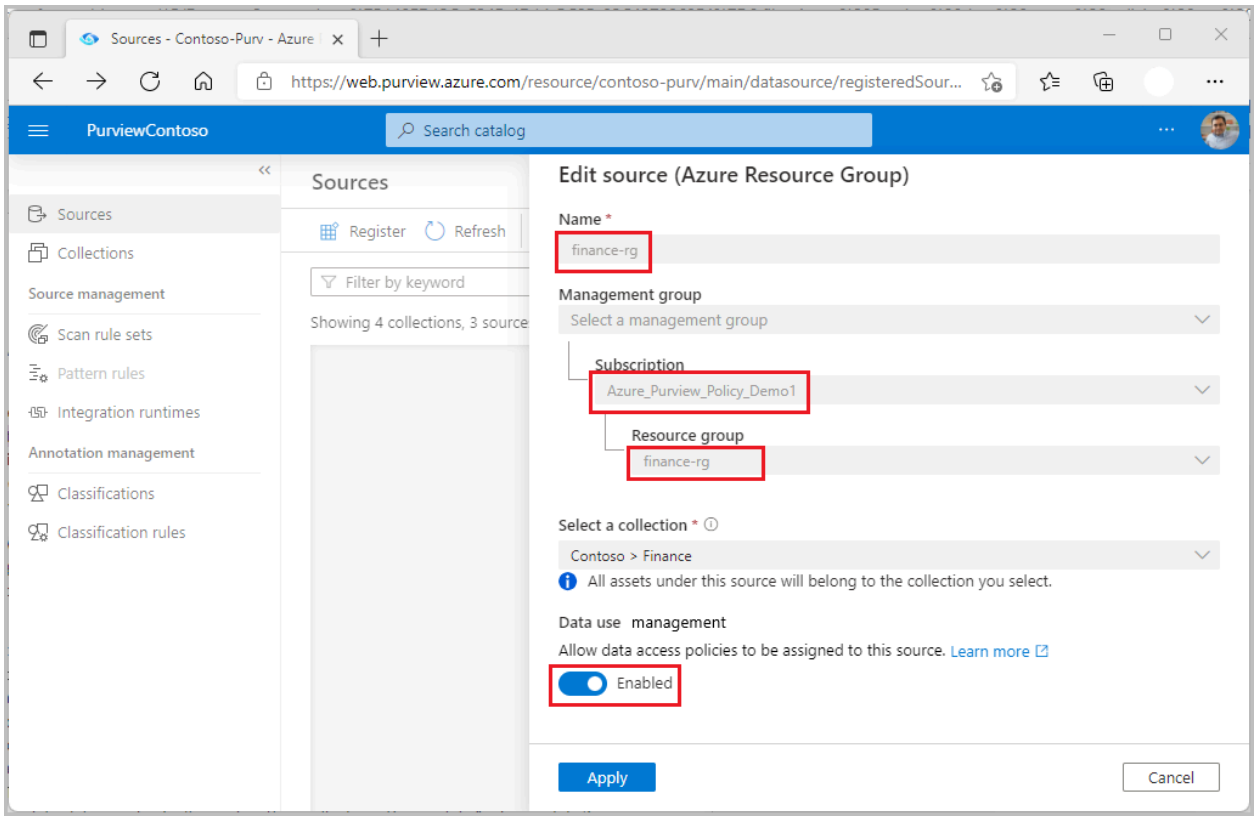
Register the subscription or resource group for Data Policy Enforcement

The subscription or resource group needs to be registered with Microsoft Purview before you can create access policies. To register your subscription or resource group, follow the **Prerequisites** and **Register** sections of this guide:

- Register multiple sources in Microsoft Purview

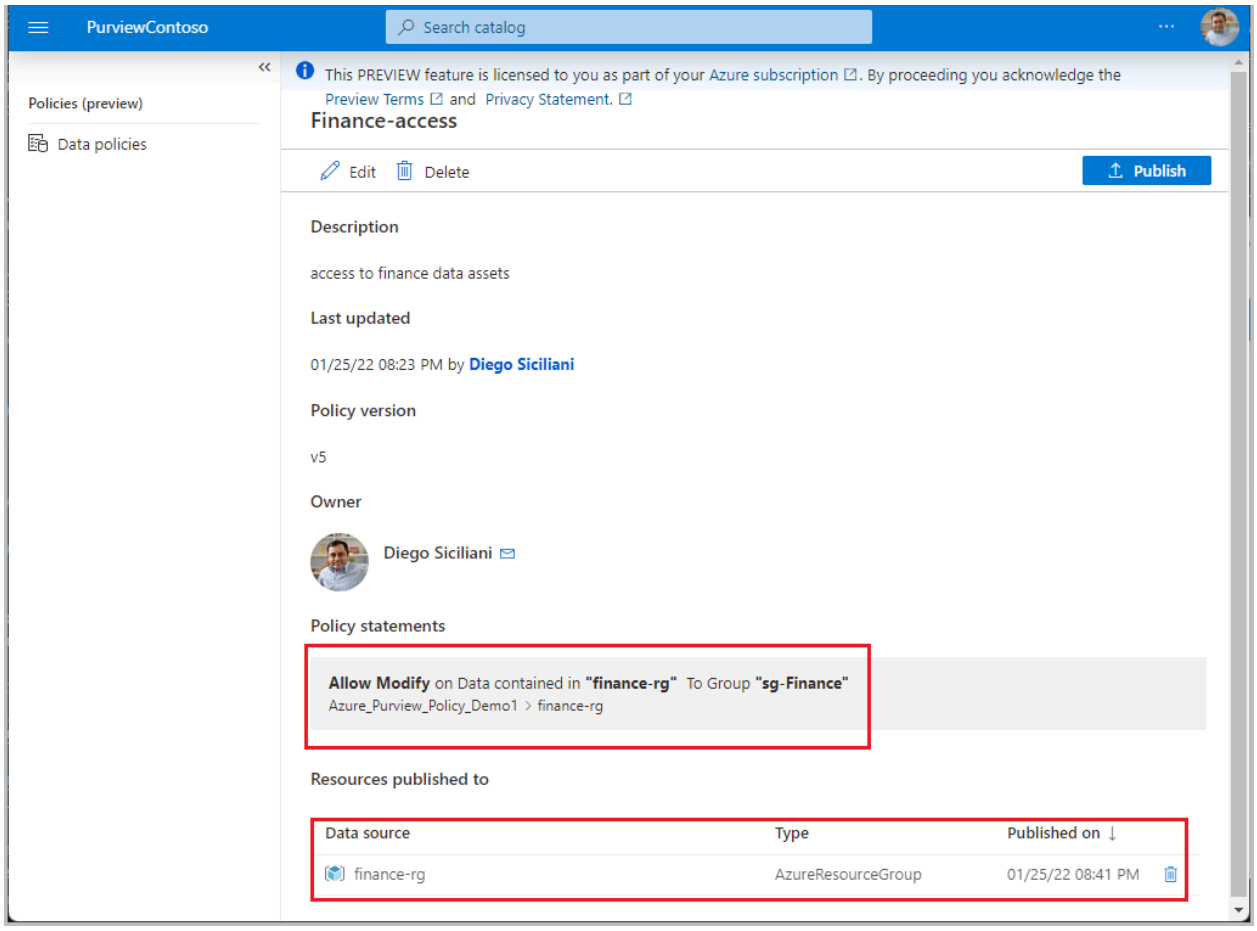
After you've registered your resources, you'll need to enable Data Policy Enforcement. Data Policy Enforcement needs certain permissions and can affect the security of your data, as it delegates to certain Microsoft Purview roles to manage access to the data sources. **Go through the secure practices related to Data Policy Enforcement in this guide**: How to enable Data Policy Enforcement

In the end, your resource will have the **Data Policy Enforcement** toggle **Enabled**, as shown in the screenshot:



Create and publish a data owner policy

Execute the steps in the **Create a new policy** and **Publish a policy** sections of the data-owner policy authoring tutorial. The result will be a data owner policy similar to the example shown in the image: a policy that provides security group *sg-Finance modify* access to resource group *finance-rg*. Use the Data source box in the Policy user experience.



> **Important**
>
> - Publish is a background operation. For example, Azure Storage accounts can take up to **2 hours** to reflect the changes.
> - Changing a policy does not require a new publish operation. The changes will be picked up with the next pull.

Unpublish a data owner policy

Follow this link for the steps to unpublish a data owner policy in Microsoft Purview.

Update or delete a data owner policy

Follow this link for the steps to update or delete a data owner policy in Microsoft Purview.

Additional information

- Creating a policy at subscription or resource group level will enable the Subjects to access Azure Storage system containers, for example, *$logs*. If this is undesired, first scan the data source and then create finer-grained policies for each (that is, at container or sub-container level).

Limits

The limit for Microsoft Purview policies that can be enforced by Storage accounts is 100 MB per subscription, which roughly equates to 5000 policies.

Next steps

Check blog, demo and related tutorials:

- Doc: Concepts for Microsoft Purview data owner policies
- Blog: Resource group-level governance can significantly reduce effort
- Video: Demo of data owner access policies for Azure Storage
- Blog: Grant users access to data assets in your enterprise via API

Feedback

Was this page helpful? Yes No

How to view self-service data access policies

Article • 01/28/2025

> **Note**
>
> Microsoft Purview Data Map (classic), Data Catalog (classic), Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

In a Microsoft Purview catalog, you can now request access to data assets. If policies are currently available for the data source type and the data source has Data Policy Enforcement enabled, a self-service policy is generated when a data access request is approved.

This article describes how to view self-service data access policies that have been autogenerated by approved access requests.

Prerequisites

> **Important**
>
> To view self-service policies, make sure that the below prerequisites are completed.

Self-service policies must exist for them to be viewed. To enable and create self-service policies, follow these articles:

1. Enable Data Policy Enforcement - this will allow Microsoft Purview to create policies for your sources.
2. Create a self-service data access workflow - this will enable users to request access to data sources from within Microsoft Purview.
3. Approve a self-service data access request - after approving a request, if your workflow from the previous step includes the ability to create a self-service data policy, your policy will be created and will be viewable.

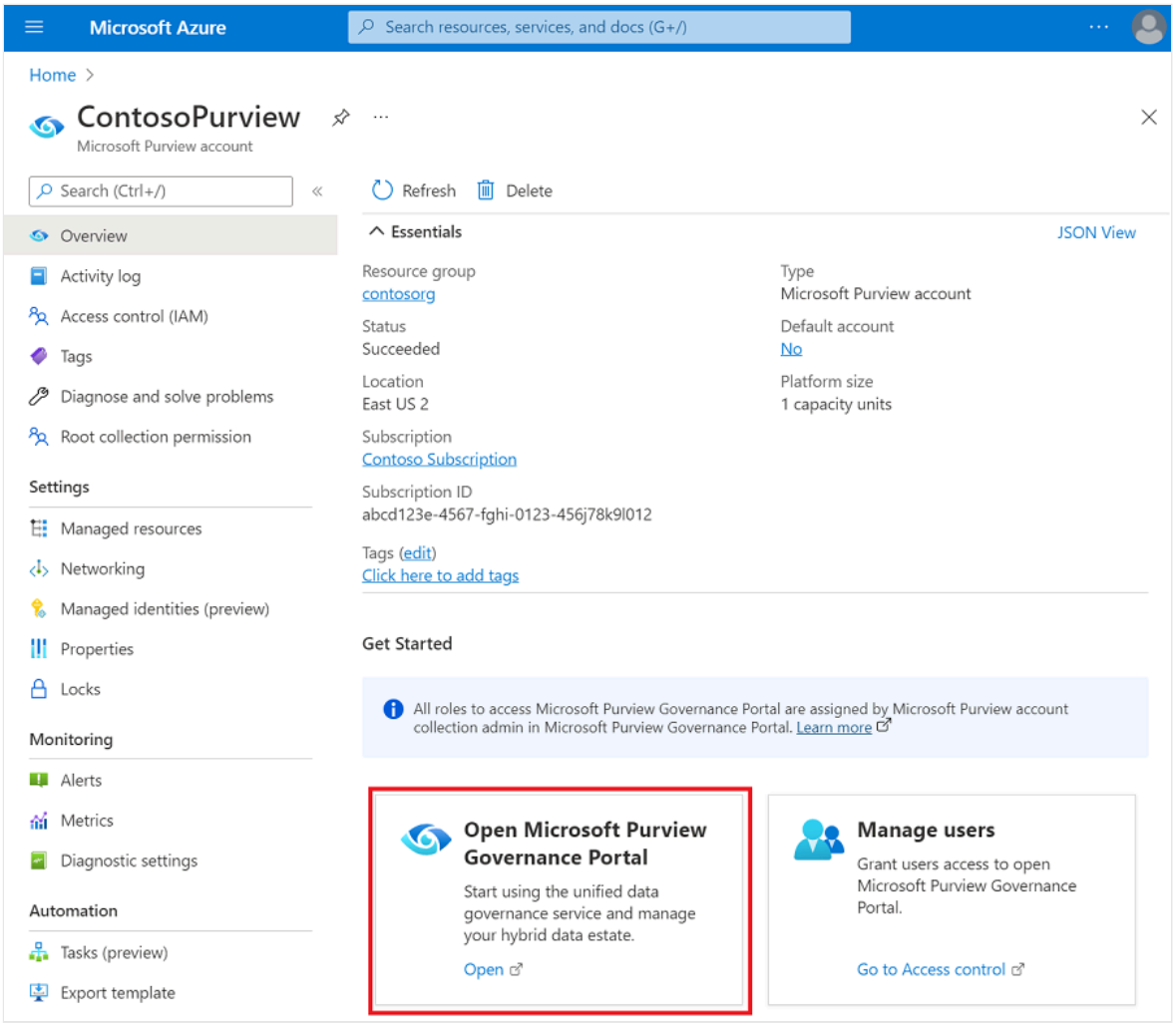
Permission

Only the creator of your Microsoft Purview account, or users with **Policy Admin** permissions can view self-service data access policies.

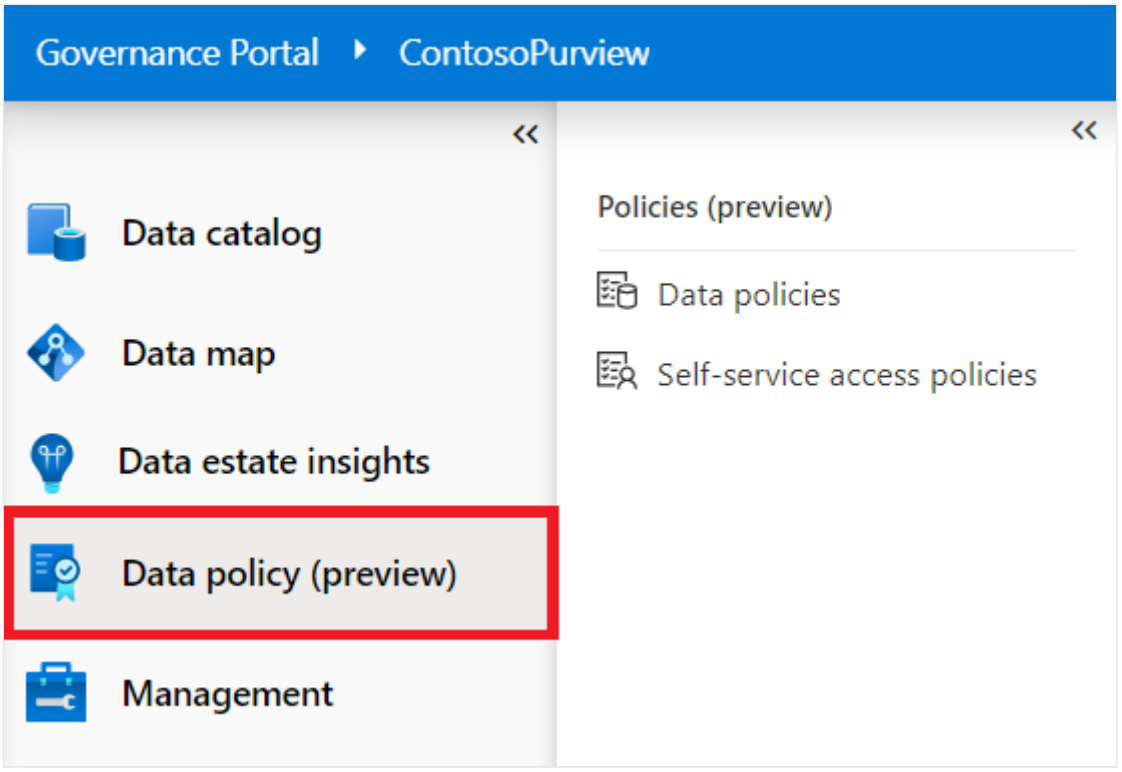
If you need to add or request permissions, follow the Microsoft Purview permissions documentation.

Steps to view self-service data access policies

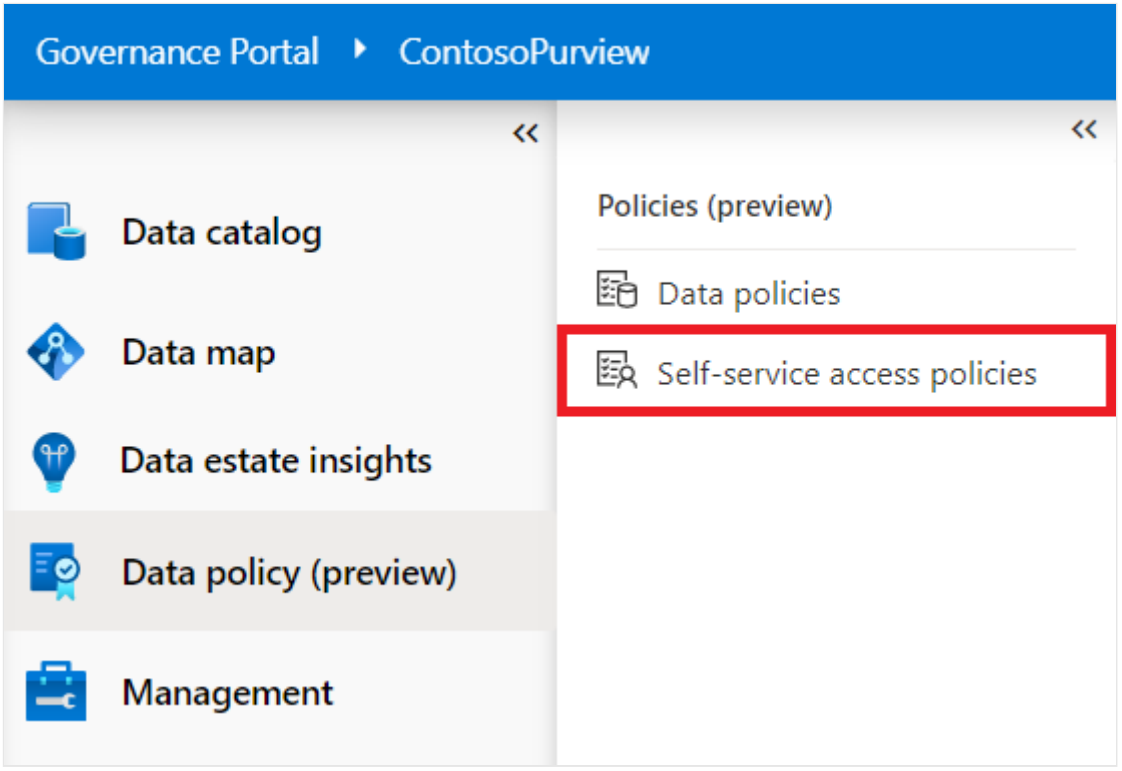
1. Launch the Microsoft Purview governance portal. The Microsoft Purview governance portal can be launched as shown below or by using the url directly.



2. Select the policy management tab to launch the self-service access policies.



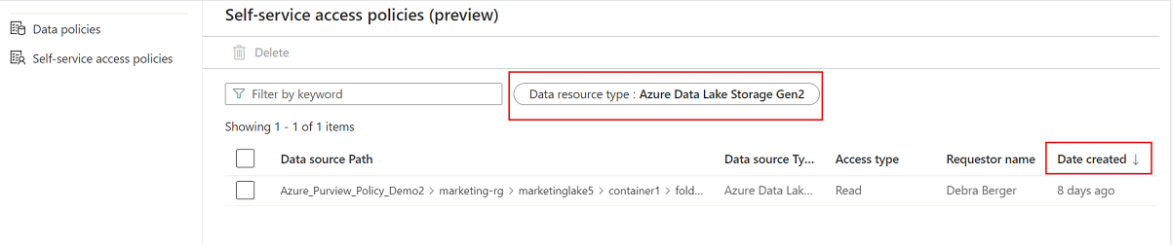
3. Open the self-service access policies tab.



4. Here you'll see all your policies. The policies can be sorted and filtered by any of the displayed columns to improve your search.

> ⓘ Note
>
> The only policies listed here will be policies from access requests that have already been approved. Access policies are not created until a self-service

access request has been submitted and approved.



5. Each policy has these details to tell you more about it. These details make up the simple rules of the policy: A user has been granted specific access to a specific data source.

    - **Data source path** - the full path of the data source in your Azure subscription that a user now has access to.
    - **Access type** - what kind of access the user has.
    - **Requestor name** and **On behalf of** - Who requested the access and whom the access was for.
    - **Date created** - when the policy was created (when the access request was approved.)

You can also delete existing self-service access policies, if you're ready to remove access for a user.

Next steps

- Self-service data access policies
- How to delete self-service access policies

Feedback

Was this page helpful? Yes No

How to delete self-service data access policies

Article • 02/26/2025

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

In a Microsoft Purview catalog, you can now request access to data assets. If policies are currently available for the data source type and the data source has Data Policy Enforcement enabled, a self-service policy is generated when a data access request is approved.

This article describes how to delete self-service data access policies that have been auto-generated by approved access requests.

Prerequisites

> **Important**
>
> To delete self-service policies, make sure that the below prerequisites are completed.

Self-service policies must exist to be deleted. To enable and create self-service policies, follow these articles:

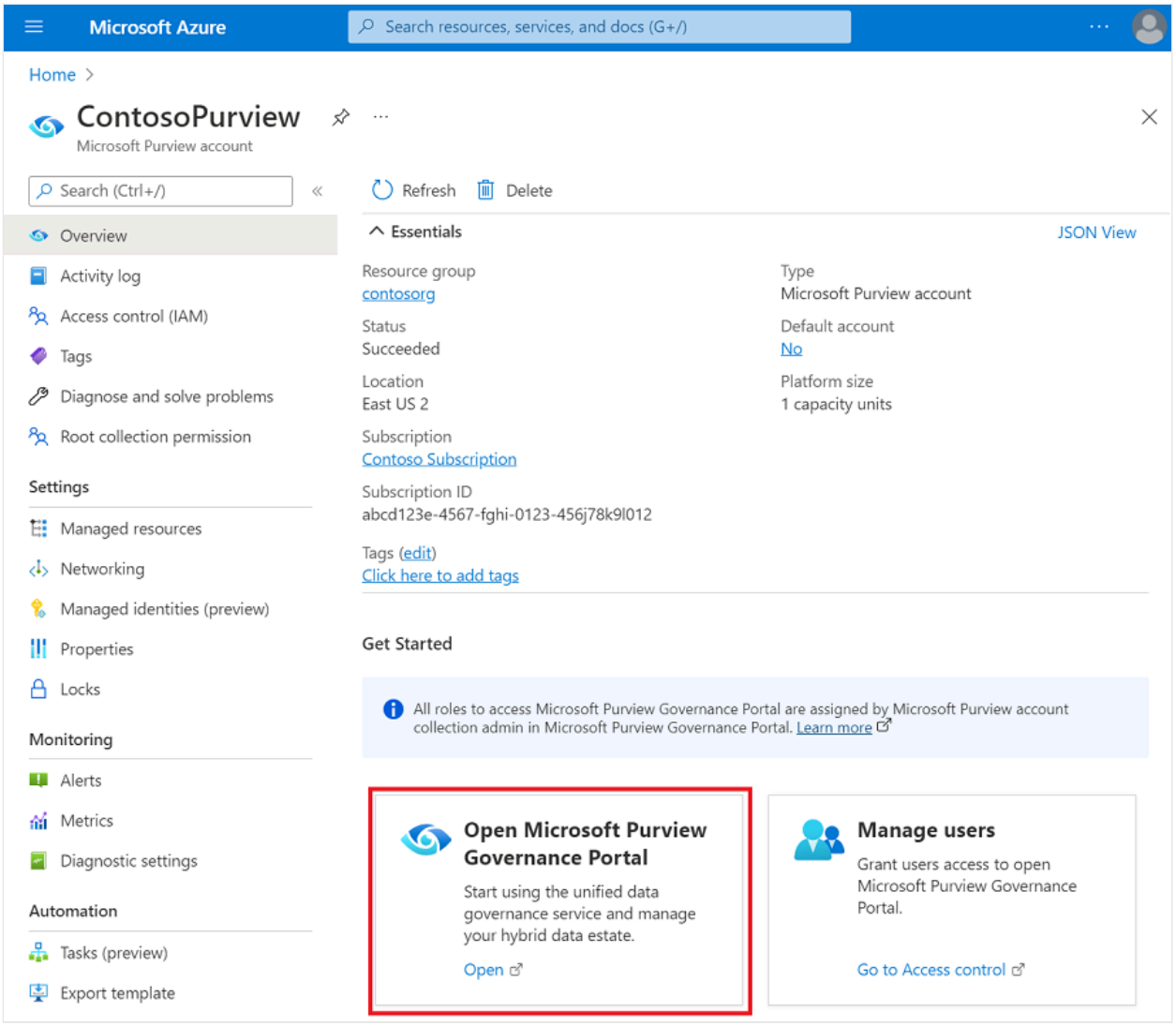
1. Enable Data Policy Enforcement - this will allow Microsoft Purview to create policies for your sources.
2. Create a self-service data access workflow - this will enable users to request access to data sources from within Microsoft Purview.
3. Approve a self-service data access request - after approving a request, if your workflow from the previous step includes the ability to create a self-service data policy, your policy will be created and will be viewable.

Permission

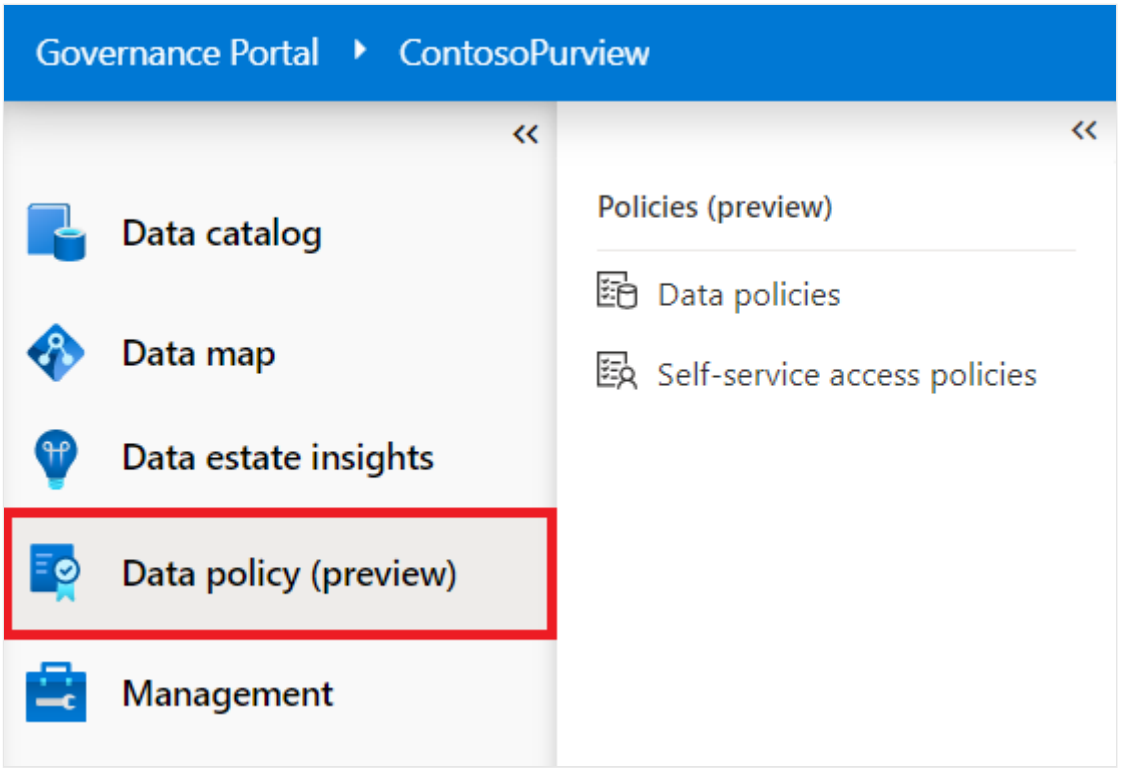
Only users with **Policy Admin** privilege can delete self-service data access policies.

Steps to delete self-service data access policies

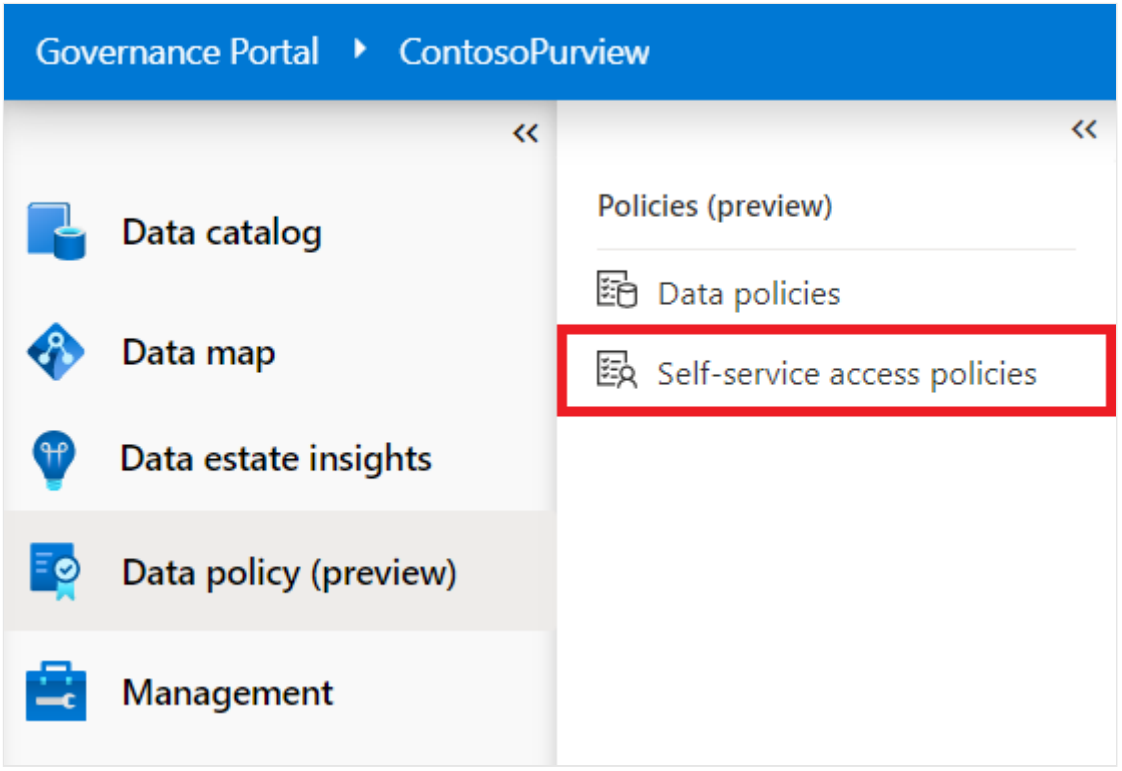
The Microsoft Purview governance portal can be launched as shown below from the Azure portal or by using the url directly.



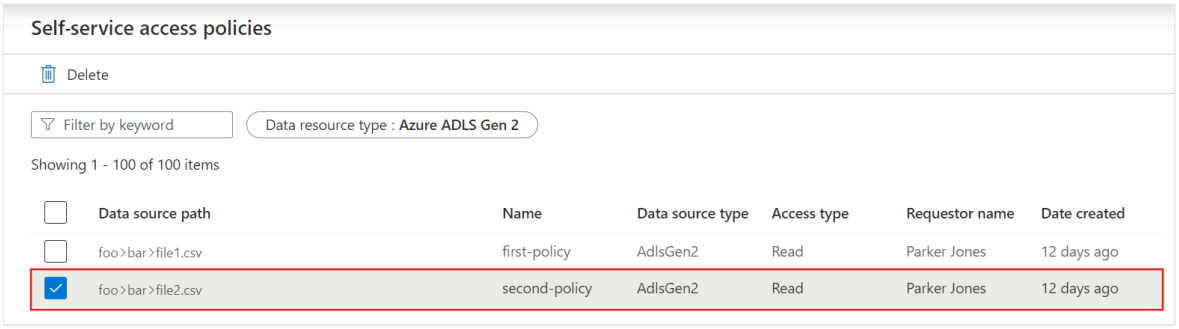
1. Select the policy management tab to launch the self-service access policies.



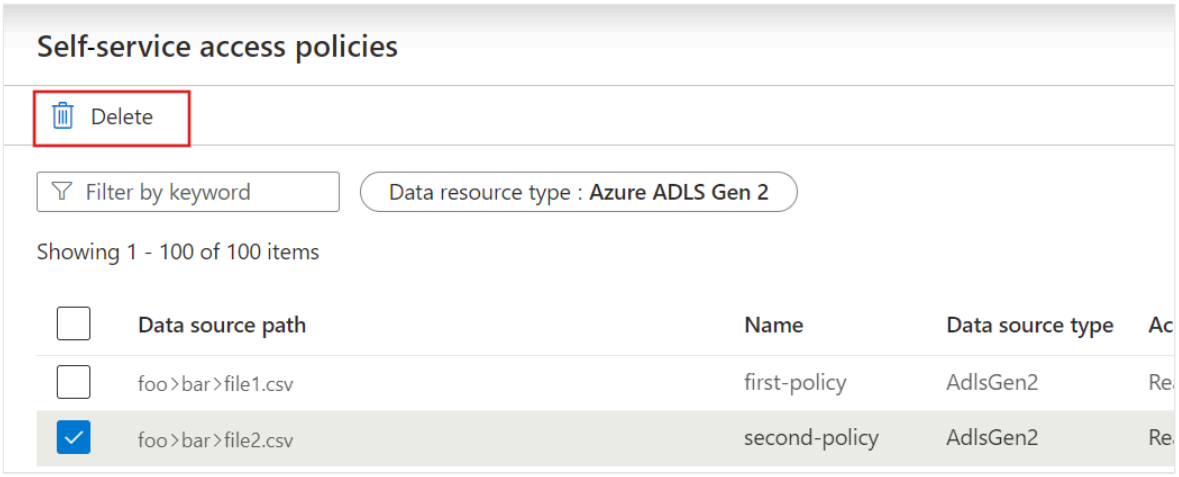
2. Open the self-service access policies tab.



3. Here you'll see all your policies. Select the policies that need to be deleted. The policies can be sorted and filtered by any of the displayed columns to improve your search.



4. Select the delete button to delete all selected policies.



5. Select **OK** on the confirmation dialog box to delete the policy. Refresh the screen to confirm that the policies have been deleted.

Next steps

- Self-service data access policy

Feedback

Was this page helpful? Yes No

Self-service policies for Azure SQL Database (preview)

Article • 02/26/2025

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

> **Important**
>
> This feature is currently in preview. The Supplemental Terms of Use for Microsoft Azure Previews include additional legal terms that apply to Azure features that are in beta, in preview, or otherwise not yet released into general availability.

Self-service policies allow you to manage access from Microsoft Purview to data sources that have been registered for **Data Policy Enforcement**.

This how-to guide describes how self-service policies get created in Microsoft Purview to enable access to Azure SQL Database. The following actions are currently enabled: *Read Tables*, and *Read Views*.

> **Caution**
>
> *Ownership chaining* must exist for *select* to work on Azure SQL Database *views*.

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- A new or existing Microsoft Purview account. Follow this quickstart guide to create one.
- Create a new Azure SQL Database instance, or use an existing one, in one of the currently available regions for this feature. You can follow this guide to create an Azure SQL Database instance.

Region support

All Microsoft Purview regions are supported.

Enforcement of Microsoft Purview policies is available only in the following regions for Azure SQL Database:

Public cloud:

- East US
- East US2
- South Central US
- West Central US
- West US3
- Canada Central
- Brazil South
- West Europe
- North Europe
- France Central
- UK South
- South Africa North
- Central India
- South East Asia
- East Asia
- Australia East

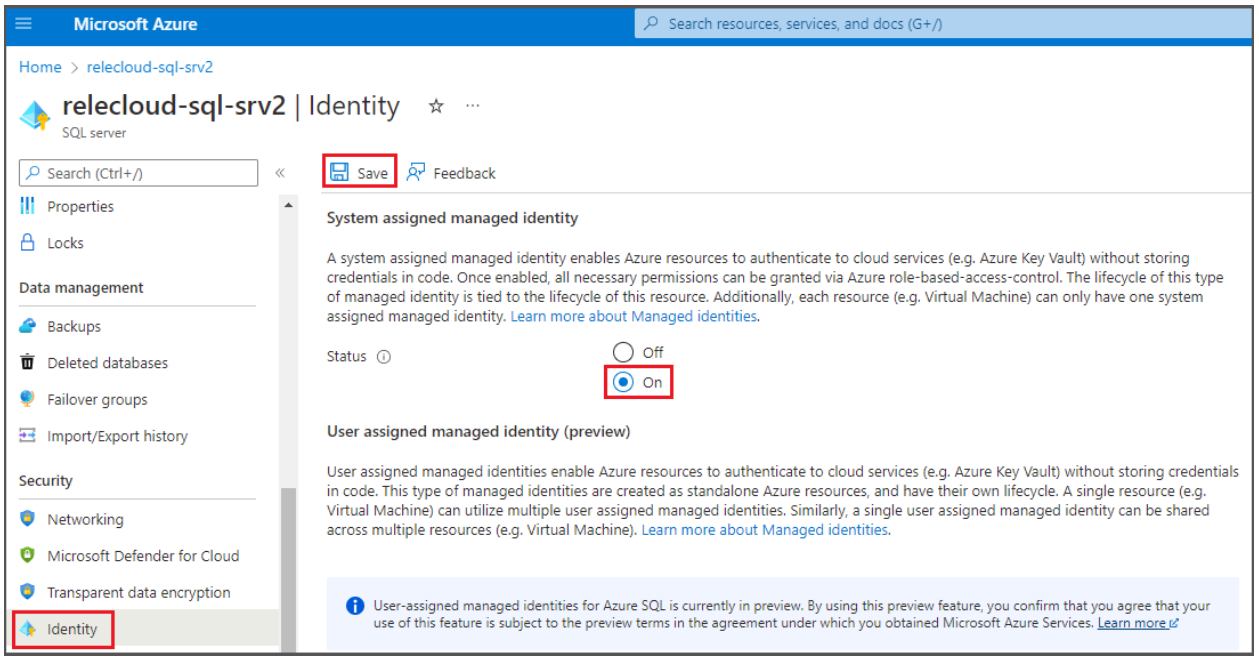
Sovereign clouds:

- USGov Virginia
- China North 3

Configure the Azure SQL Database instance for policies from Microsoft Purview

For the logical server associated with Azure SQL Database to honor policies from Microsoft Purview, you need to configure a Microsoft Entra admin. In the Azure portal, go to the logical server that hosts the Azure SQL Database instance. On the side menu, select **Microsoft Entra ID**. Set an admin name to any Microsoft Entra user or group that you prefer, and then select **Save**.

Then, on the side menu, select **Identity**. Under **System assigned managed identity**, turn the status to **On**, and then select **Save**.



Microsoft Purview Configuration

Register the data source in Microsoft Purview

Before a policy can be created in Microsoft Purview for a data resource, you must register that data resource in Microsoft Purview Studio. You will find the instructions related to registering the data resource later in this guide.

> **Note**
>
> Microsoft Purview policies rely on the data resource ARM path. If a data resource is moved to a new resource group or subscription it will need to be de-registered and then registered again in Microsoft Purview.

Configure permissions to enable Data policy enforcement on the data source

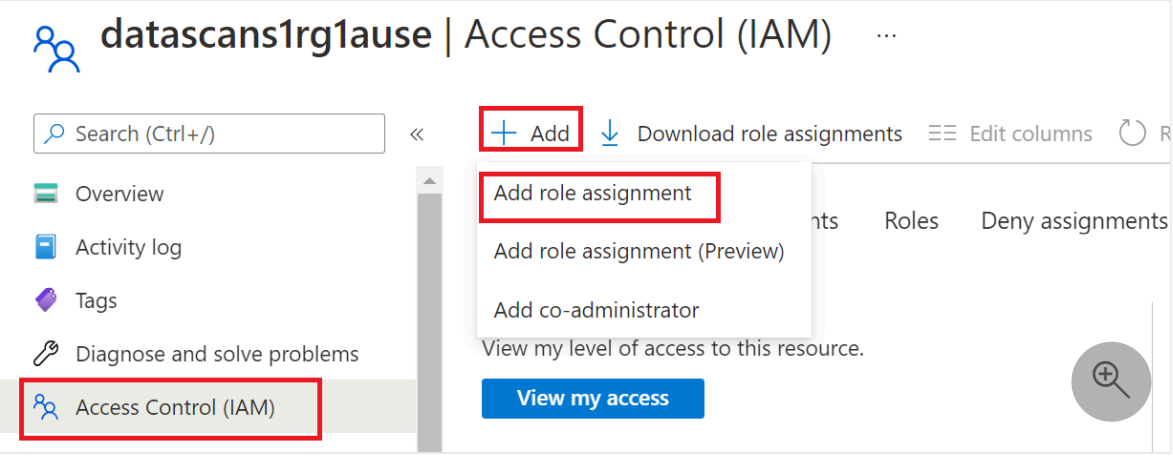
Once a resource is registered, but before a policy can be created in Microsoft Purview for that resource, you must configure permissions. A set of permissions are needed to enable the **Data policy enforcement**. This applies to data sources, resource groups, or subscriptions. To enable *Data policy enforcement*, you must have *both* specific Identity and Access Management (IAM) privileges on the resource as well as specific Microsoft Purview privileges:

- You must have *either one* of the following IAM role combinations on the resource's Azure Resource Manager path or any parent of it (that is, using IAM permission

inheritance):

- IAM Owner
- Both IAM Contributor and IAM User Access Administrator

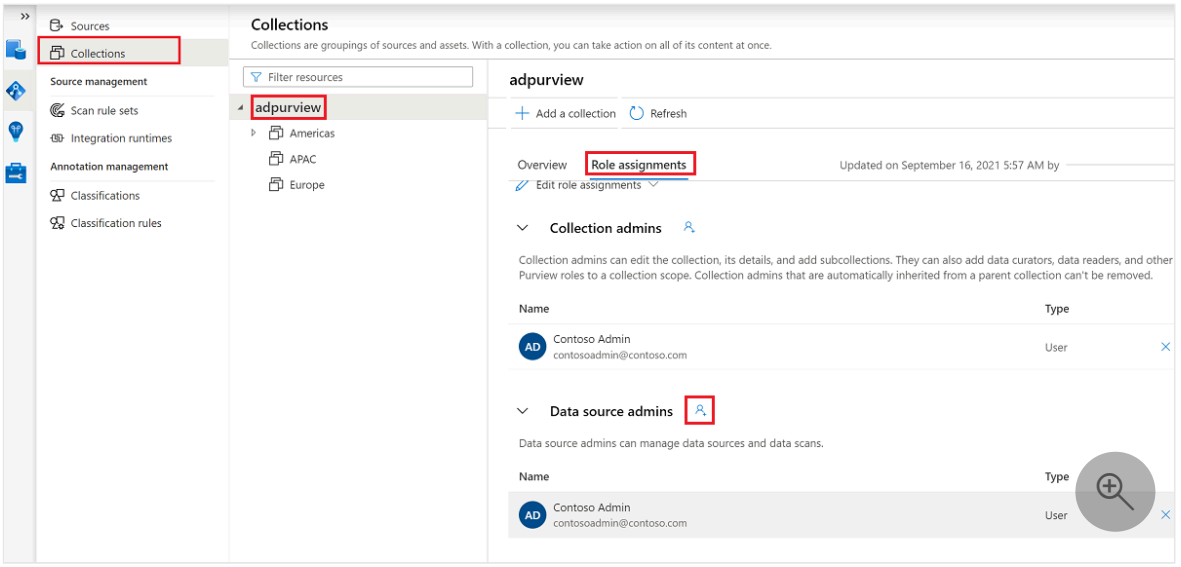
To configure Azure role-based access control (RBAC) permissions, follow this guide. The following screenshot shows how to access the **Access Control** section in the Azure portal for the data resource to add a role assignment.



> **Note**
>
> The *IAM Owner* role for a data resource can be inherited from a parent resource group, a subscription, or a subscription management group. Check which Microsoft Entra users, groups, and service principals hold or are inheriting the *IAM Owner* role for the resource.

- You also need to have the Microsoft Purview *Data source admin* role for the collection or a parent collection (if inheritance is enabled). For more information, see the guide on managing Microsoft Purview role assignments.

  The following screenshot shows how to assign the *Data source admin* role at the root collection level.



Configure Microsoft Purview permissions to create, update, or delete access policies

To create, update or delete policies, you need to get the *Policy author* role in Microsoft Purview at *root collection level*:

- The *Policy author* role can create, update, and delete DevOps and Data Owner policies.
- The *Policy author* role can delete self-service access policies.

For more information about managing Microsoft Purview role assignments, see Create and manage collections in the Microsoft Purview Data Map.

> **Note**
>
> Policy author role must be configured at the root collection level.

In addition, to easily search Microsoft Entra users or groups when creating or updating the subject of a policy, you can greatly benefit from getting the Directory Readers permission in Microsoft Entra ID. This is a common permission for users in an Azure tenant. Without the Directory Reader permission, the Policy Author will have to type the complete username or email for all the principals included in the subject of a data policy.

Configure Microsoft Purview permissions for publishing Data Owner policies

Data Owner policies allow for checks and balances if you assign the Microsoft Purview *Policy author* and *Data source admin* roles to different people in the organization. Before

a Data owner policy takes effect, a second person (*Data source admin*) must review it and explicitly approve it by publishing it. This does not apply to DevOps or Self-service access policies as publishing is automatic for them when those policies are created or updated.

To publish a Data owner policy you need to get the *Data source admin* role in Microsoft Purview at *root collection level*.

For more information about managing Microsoft Purview role assignments, see Create and manage collections in the Microsoft Purview Data Map.

> ⓘ **Note**
>
> To publish Data owner policies, the Data source admin role must be configured at the root collection level.

Delegate access provisioning responsibility to roles in Microsoft Purview

After a resource has been enabled for **Data policy enforcement**, any Microsoft Purview user with the *Policy author* role at the root collection level can provision access to that data source from Microsoft Purview.

> ⓘ **Note**
>
> Any Microsoft Purview root *Collection admin* can assign new users to root *Policy author* roles. Any *Collection admin* can assign new users to a *Data source admin* role under the collection. Minimize and carefully vet the users who hold Microsoft Purview *Collection admin*, *Data source admin*, or *Policy author* roles.

If a Microsoft Purview account with published policies is deleted, such policies will stop being enforced within an amount of time that depends on the specific data source. This change can have implications on both security and data access availability. The Contributor and Owner roles in IAM can delete Microsoft Purview accounts. You can check these permissions by going to the **Access control (IAM)** section for your Microsoft Purview account and selecting **Role Assignments**. You can also use a lock to prevent the Microsoft Purview account from being deleted through Resource Manager locks.

Register the data sources in Microsoft Purview

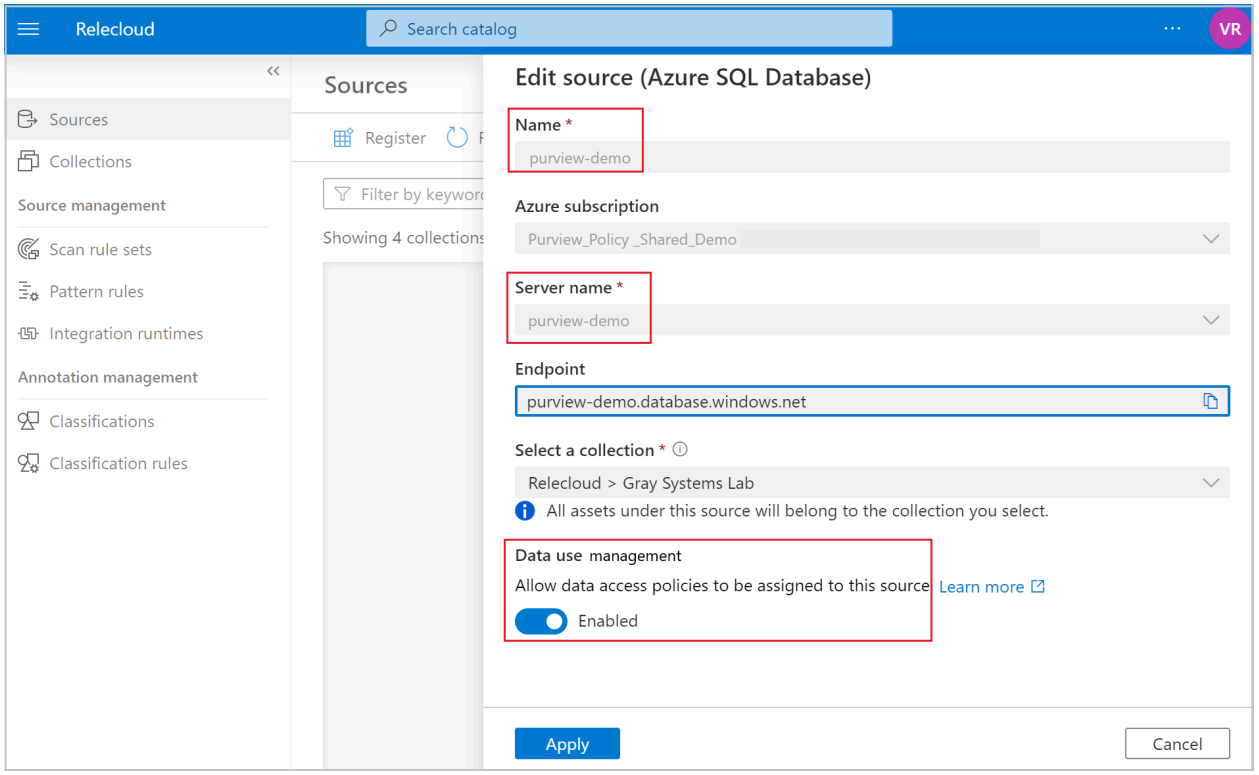
The Azure SQL Database resources need to be registered first with Microsoft Purview to later define access policies. You can follow these guides:

Register and scan Azure SQL DB

After you've registered your resources, you'll need to enable data policy enforcement. Data policy enforcement can affect the security of your data, as it delegates to certain Microsoft Purview roles to manage access to the data sources. **Go through the secure practices related to data policy enforcement in this guide**:

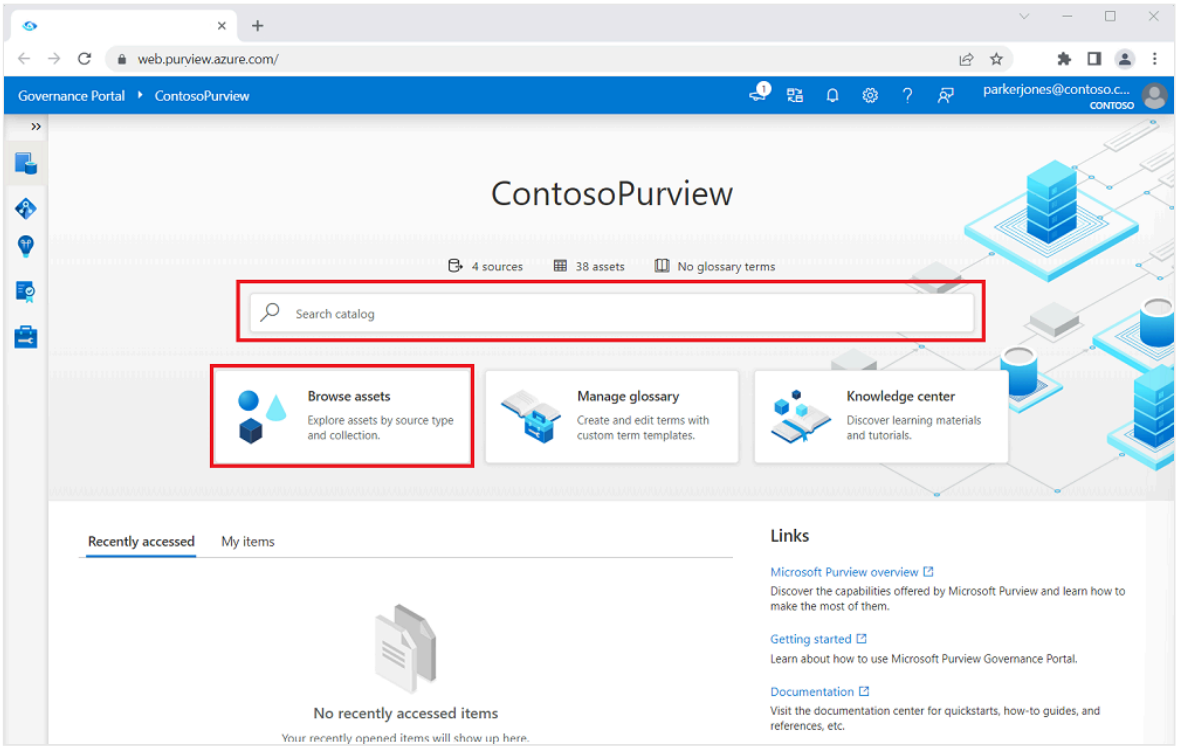
How to enable data policy enforcement

Once your data source has the **Data policy enforcement** toggle *Enabled*, it will look like this picture. This will enable the access policies to be used with the given SQL server and all its contained databases.



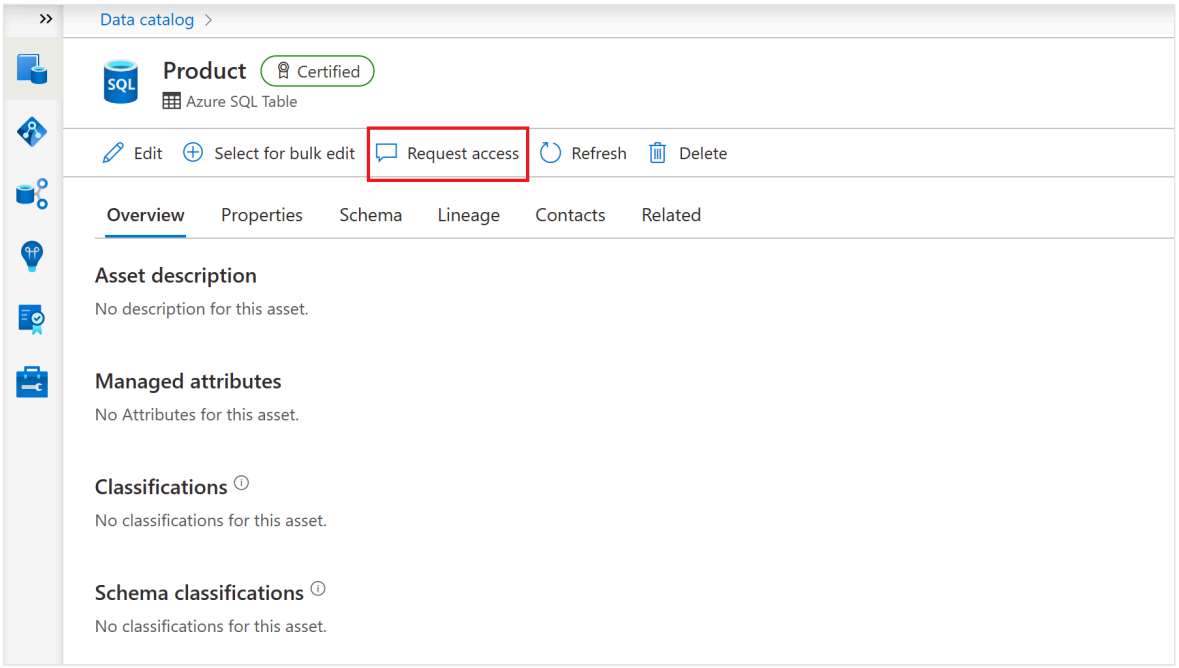
Create a self-service data access request

1. To find a data asset, use Microsoft Purview's search or browse functionality.



2. Select the asset to go to asset details.

3. Select **Request access**.



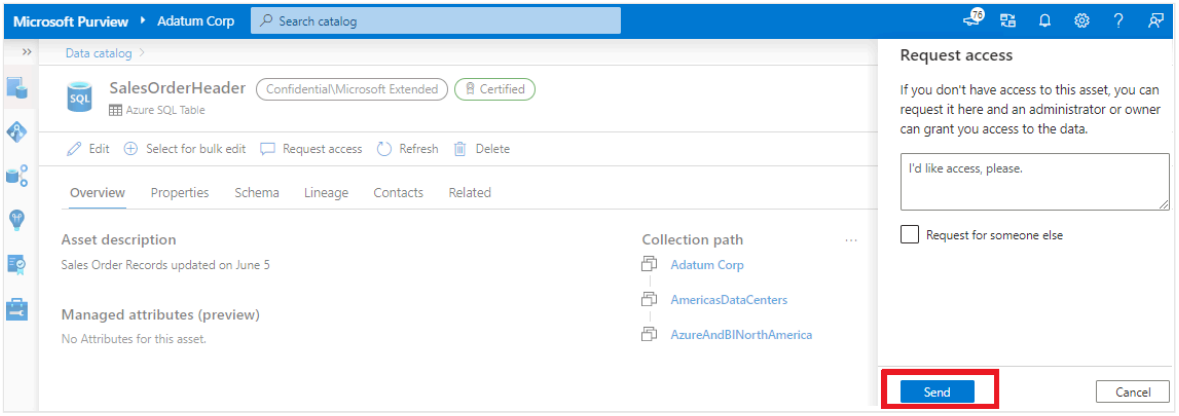
> **Note**
>
> If this option isn't available, a self-service access workflow either hasn't been created, or hasn't been assigned to the collection where the resource is registered. Contact the collection administrator, data source administrator, or workflow administrator of your collection for more information. Or, for

information on how to create a self-service access workflow, see our [self-service access workflow documentation](#).

4. The **Request access** window will open. You can provide comments on why data access is requested.

5. Select **Send** to trigger the self-service data access workflow.

> **Note**
>
> If you want to request access on behalf of another user, select the checkbox **Request for someone else** and populate the email id of that user.



> **Note**
>
> A request access to resource set will actually submit the data access request for the folder one level up which contains all these resource set files.

6. Data owners will be notified of your request and will either approve or reject the request.

> **Important**
>
> - Publish is a background operation. It can take up to **5 minutes** for the changes to be reflected in this data source.
> - Changing a policy does not require a new publish operation. The changes will be picked up with the next pull.

View a self-service Policy

To view the policies you've created, follow the article to view the self-service policies.

Test the policy

The Microsoft Entra account, group, MSI, or SPN for which the self-service policies were created, should now be able to connect to the database on the server and execute a select query against the requested table or view.

Force policy download

It's possible to force an immediate download of the latest published policies to the current SQL database by running the following command. The minimal permission required to run the command is membership in the ##MS_ServerStateManager##-server role.

SQL

```
-- Force immediate download of latest published policies
exec sp_external_policy_refresh reload
```

Analyze downloaded policy state from SQL

The following DMVs can be used to analyze which policies have been downloaded and are currently assigned to Microsoft Entra accounts. The minimal permission required to run them is VIEW DATABASE SECURITY STATE - or assigned Action Group *SQL Security Auditor*.

SQL

```
-- Lists generally supported actions
SELECT * FROM sys.dm_server_external_policy_actions

-- Lists the roles that are part of a policy published to this server
SELECT * FROM sys.dm_server_external_policy_roles

-- Lists the links between the roles and actions, could be used to join the
two
SELECT * FROM sys.dm_server_external_policy_role_actions

-- Lists all Azure AD principals that were given connect permissions
SELECT * FROM sys.dm_server_external_policy_principals

-- Lists Azure AD principals assigned to a given role on a given resource
scope
```

```
SELECT * FROM sys.dm_server_external_policy_role_members

-- Lists Azure AD principals, joined with roles, joined with their data
actions
SELECT * FROM sys.dm_server_external_policy_principal_assigned_actions
```

Additional information

Policy action mapping

This section contains a reference of how actions in Microsoft Purview data policies map to specific actions in Azure SQL Database.

Expand table

Microsoft Purview policy action	Data source specific actions
Read	Microsoft.Sql/sqlservers/Connect
	Microsoft.Sql/sqlservers/databases/Connect
	Microsoft.Sql/Sqlservers/Databases/Schemas/Tables/Rows
	Microsoft.Sql/Sqlservers/Databases/Schemas/Views/Rows

Next steps

Check blog, demo and related how-to guides

- self-service policies
- What are Microsoft Purview workflows
- Self-service data access workflow for hybrid data estates

Feedback

Was this page helpful? Yes No

Self-service access provisioning for Azure Storage datasets (Preview)

Article • 02/26/2025

> **Note**
>
> Microsoft Purview Data Catalog (classic) and Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

> **Important**
>
> This feature is currently in preview. The Supplemental Terms of Use for Microsoft Azure Previews include additional legal terms that apply to Azure features that are in beta, in preview, or otherwise not yet released into general availability.

Access policies allow you to manage access from Microsoft Purview to data sources that have been registered for *Data Policy Enforcement*.

This how-to guide describes how self-service policies get created in Microsoft Purview to enable access to Azure storage datasets. Currently, these two Azure Storage sources are supported:

- Blob storage
- Azure Data Lake Storage (ADLS) Gen2

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- A new or existing Microsoft Purview account. Follow this quickstart guide to create one.

Region support

- All Microsoft Purview regions are supported.
- Storage accounts in the following regions are supported without the need for additional configuration. However, zone-redundant storage (ZRS) accounts are not

supported.
    - Australia Central
    - Australia East
    - Australia Southeast
    - Brazil South
    - Canada Central
    - Canada East
    - Central India
    - Central US
    - East Asia
    - East US 2
    - East US
    - France Central
    - Germany West Central
    - Japan East
    - Japan West
    - Korea Central
    - North Central US
    - North Europe
    - Norway East
    - Poland Central
    - Qatar Central
    - South Central US
    - South Africa North
    - Southeast Asia
    - South India
    - Sweden Central
    - Switzerland North
    - West Central US
    - West Europe
    - West US
    - West US 2
    - West US 3
    - UAE North
    - UK South
    - UK West
- Storage accounts in other regions in Public Cloud are supported after setting feature flag *AllowPurviewPolicyEnforcement*, as outlined in the next section. Newly created ZRS Storage accounts are supported, if created after setting the feature flag *AllowPurviewPolicyEnforcement*.

If needed, you can create a new Storage account by following this guide.

Configure the subscription where the Azure Storage account resides for policies from Microsoft Purview

This step is only necessary in certain regions (see prior section). To enable Microsoft Purview to manage policies for one or more Azure Storage accounts, execute the following PowerShell commands in the subscription where you'll deploy your Azure Storage account. These PowerShell commands will enable Microsoft Purview to manage policies on all Azure Storage accounts in that subscription.

If you're executing these commands locally, be sure to run PowerShell as an administrator. Alternatively, you can use the Azure Cloud Shell in the Azure portal: https://shell.azure.com.

PowerShell

```
# Install the Az module
Install-Module -Name Az -Scope CurrentUser -Repository PSGallery -Force
# Login into the subscription
Connect-AzAccount -Subscription <SubscriptionID>
# Register the feature
Register-AzProviderFeature -FeatureName AllowPurviewPolicyEnforcement -
ProviderNamespace Microsoft.Storage
```

If the output of the last command shows *RegistrationState* as *Registered*, then your subscription is enabled for access policies. If the output is *Registering*, wait at least 10 minutes, and then retry the command. **Do not continue unless the RegistrationState shows as *Registered***.

Configuration

Register the data source in Microsoft Purview

Before a policy can be created in Microsoft Purview for a data resource, you must register that data resource in Microsoft Purview Studio. You will find the instructions related to registering the data resource later in this guide.

> **Note**
>
> Microsoft Purview policies rely on the data resource ARM path. If a data resource is moved to a new resource group or subscription it will need to be de-registered and

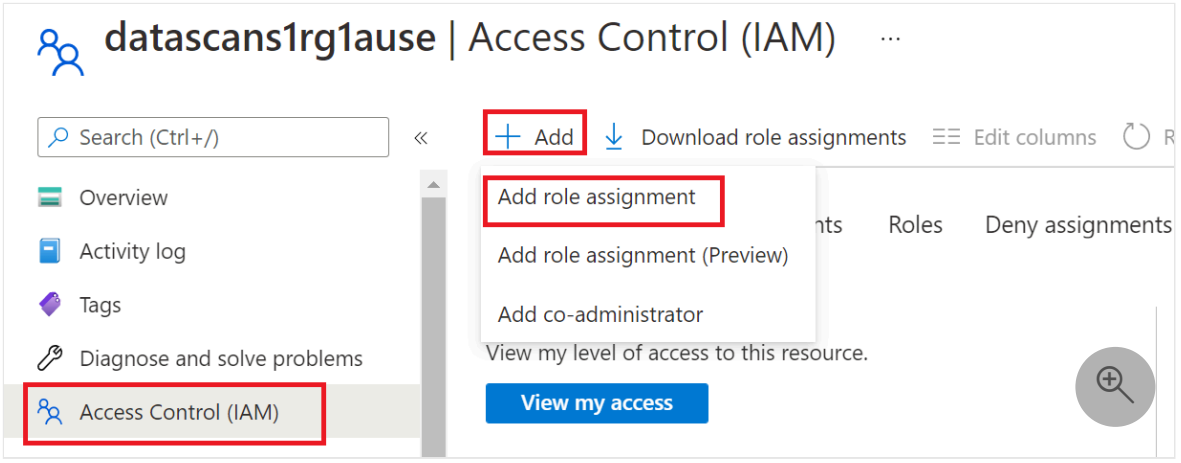
then registered again in Microsoft Purview.

Configure permissions to enable Data policy enforcement on the data source

Once a resource is registered, but before a policy can be created in Microsoft Purview for that resource, you must configure permissions. A set of permissions are needed to enable the **Data policy enforcement**. This applies to data sources, resource groups, or subscriptions. To enable *Data policy enforcement*, you must have *both* specific Identity and Access Management (IAM) privileges on the resource as well as specific Microsoft Purview privileges:

- You must have *either one* of the following IAM role combinations on the resource's Azure Resource Manager path or any parent of it (that is, using IAM permission inheritance):
  - IAM Owner
  - Both IAM Contributor and IAM User Access Administrator

  To configure Azure role-based access control (RBAC) permissions, follow this guide. The following screenshot shows how to access the **Access Control** section in the Azure portal for the data resource to add a role assignment.

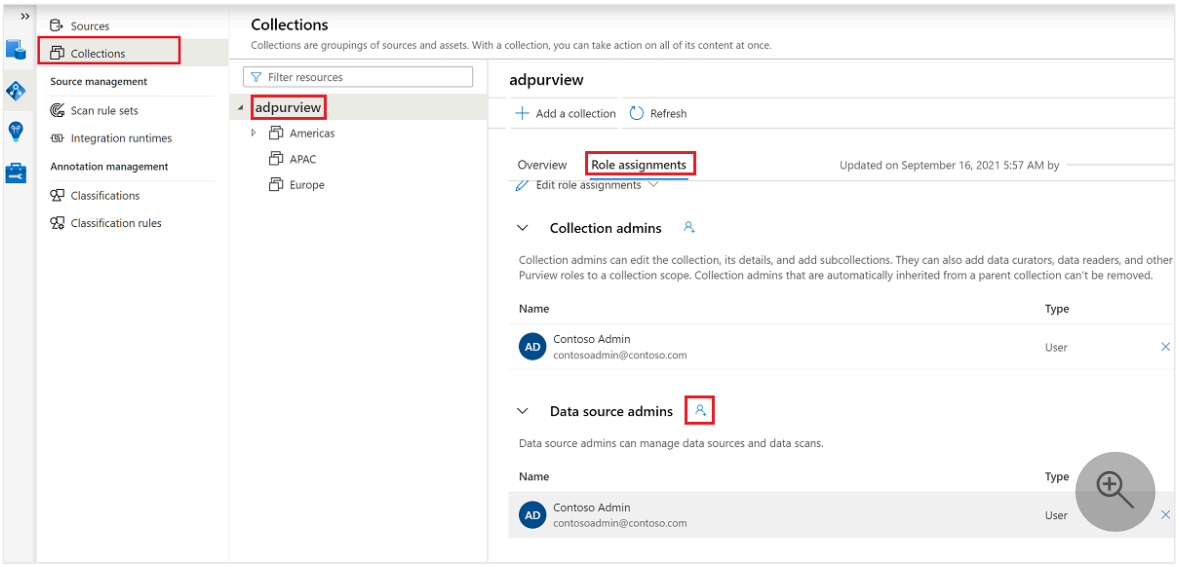


> **Note**
>
> The *IAM Owner* role for a data resource can be inherited from a parent resource group, a subscription, or a subscription management group. Check which Microsoft Entra users, groups, and service principals hold or are inheriting the *IAM Owner* role for the resource.

- You also need to have the Microsoft Purview *Data source admin* role for the collection or a parent collection (if inheritance is enabled). For more information,

see the guide on managing Microsoft Purview role assignments.

The following screenshot shows how to assign the *Data source admin* role at the root collection level.



Configure Microsoft Purview permissions to create, update, or delete access policies

To create, update or delete policies, you need to get the *Policy author* role in Microsoft Purview at *root collection level*:

- The *Policy author* role can create, update, and delete DevOps and Data Owner policies.
- The *Policy author* role can delete self-service access policies.

For more information about managing Microsoft Purview role assignments, see Create and manage collections in the Microsoft Purview Data Map.

> **Note**
>
> Policy author role must be configured at the root collection level.

In addition, to easily search Microsoft Entra users or groups when creating or updating the subject of a policy, you can greatly benefit from getting the Directory Readers permission in Microsoft Entra ID. This is a common permission for users in an Azure tenant. Without the Directory Reader permission, the Policy Author will have to type the complete username or email for all the principals included in the subject of a data policy.

Configure Microsoft Purview permissions for publishing Data Owner policies

Data Owner policies allow for checks and balances if you assign the Microsoft Purview *Policy author* and *Data source admin* roles to different people in the organization. Before a Data owner policy takes effect, a second person (*Data source admin*) must review it and explicitly approve it by publishing it. This does not apply to DevOps or Self-service access policies as publishing is automatic for them when those policies are created or updated.

To publish a Data owner policy you need to get the *Data source admin* role in Microsoft Purview at *root collection level*.

For more information about managing Microsoft Purview role assignments, see Create and manage collections in the Microsoft Purview Data Map.

> **Note**
>
> To publish Data owner policies, the Data source admin role must be configured at the root collection level.

Delegate access provisioning responsibility to roles in Microsoft Purview

After a resource has been enabled for **Data policy enforcement**, any Microsoft Purview user with the *Policy author* role at the root collection level can provision access to that data source from Microsoft Purview.

> **Note**
>
> Any Microsoft Purview root *Collection admin* can assign new users to root *Policy author* roles. Any *Collection admin* can assign new users to a *Data source admin* role under the collection. Minimize and carefully vet the users who hold Microsoft Purview *Collection admin*, *Data source admin*, or *Policy author* roles.

If a Microsoft Purview account with published policies is deleted, such policies will stop being enforced within an amount of time that depends on the specific data source. This change can have implications on both security and data access availability. The Contributor and Owner roles in IAM can delete Microsoft Purview accounts. You can check these permissions by going to the **Access control (IAM)** section for your Microsoft Purview account and selecting **Role Assignments**. You can also use a lock to

prevent the Microsoft Purview account from being deleted through Resource Manager locks.

Register the data sources in Microsoft Purview for Data Policy Enforcement

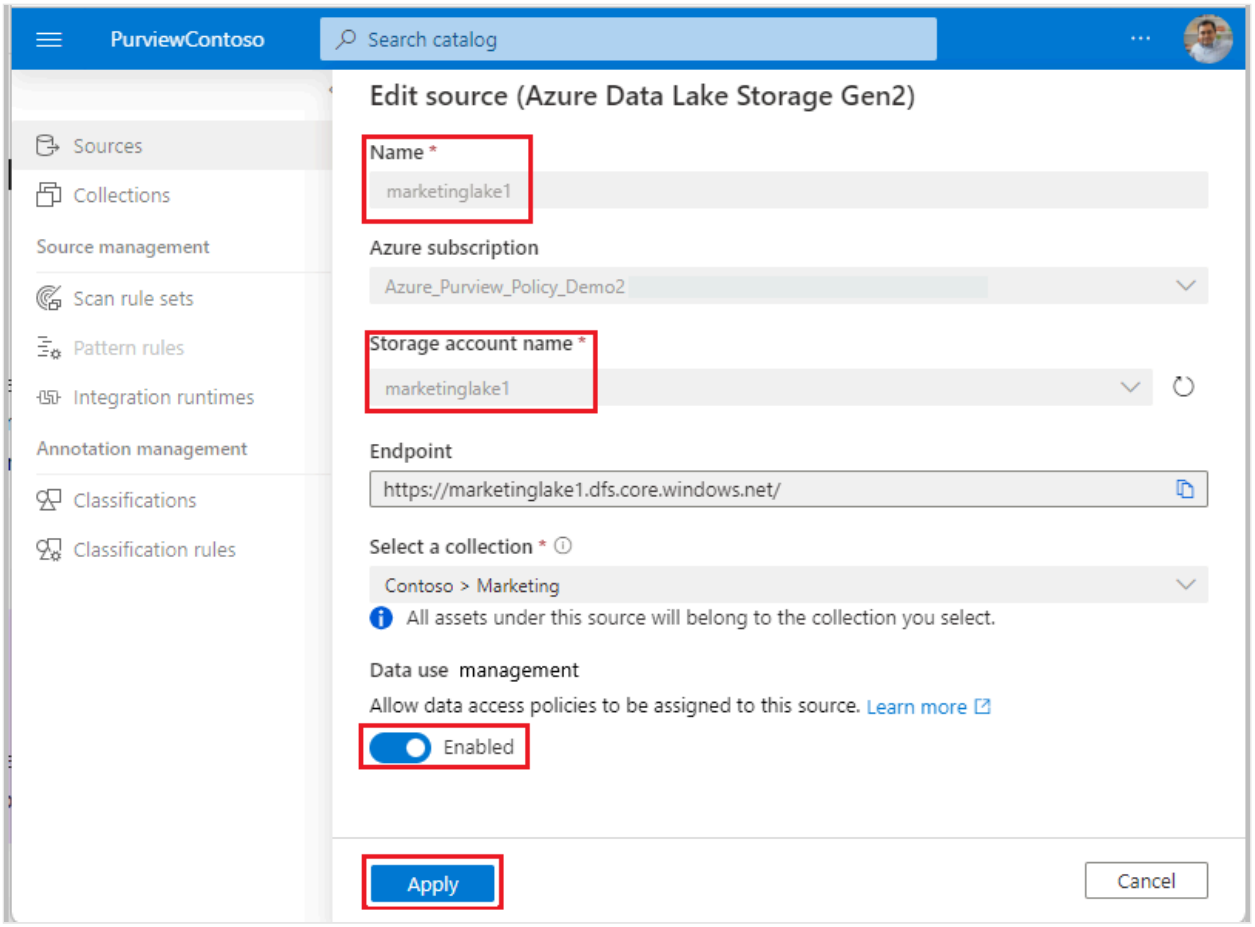
The Azure Storage resources need to be registered first with Microsoft Purview to later define access policies.

To register your resources, follow the **Prerequisites** and **Register** sections of these guides:

- Register and scan Azure Storage Blob - Microsoft Purview
- Register and scan Azure Data Lake Storage (ADLS) Gen2 - Microsoft Purview

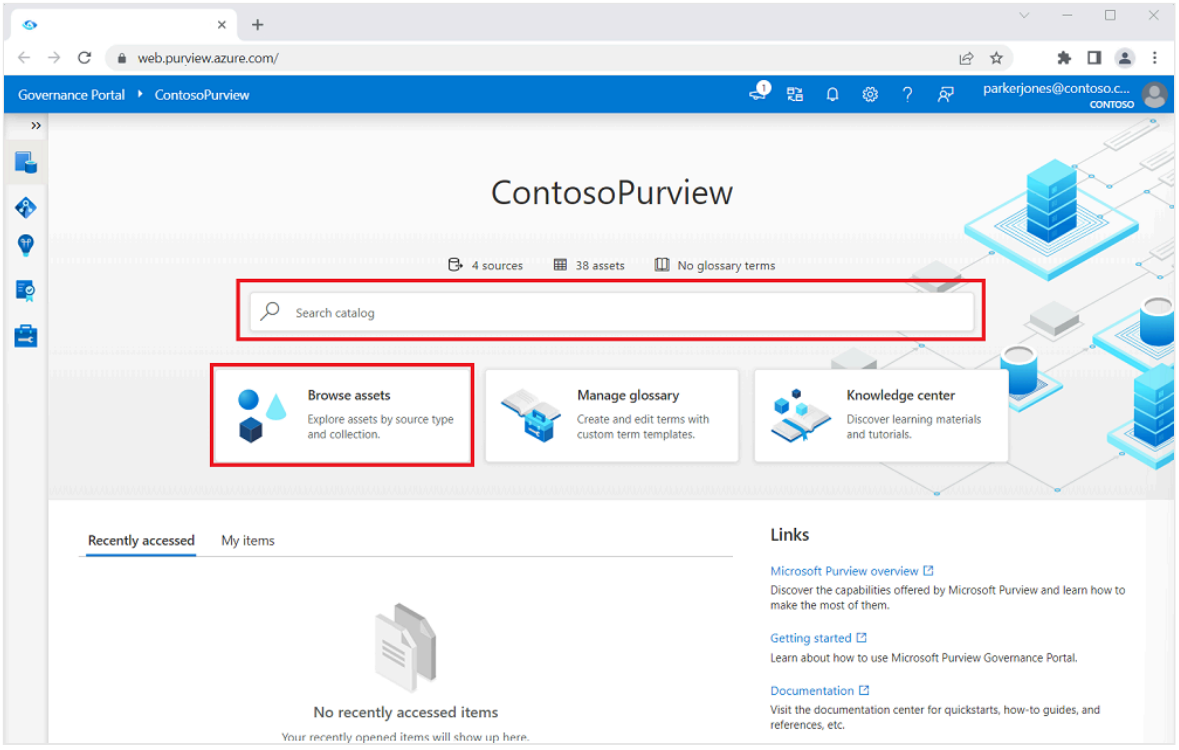
After you've registered your resources, you'll need to enable data policy enforcement. Data policy enforcement needs certain permissions and can affect the security of your data, as it delegates to certain Microsoft Purview roles to manage access to the data sources. **Go through the secure practices related to Data Policy Enforcement in this guide**: How to enable data policy enforcement

Once your data source has the **Data Policy Enforcement** toggle **Enabled**, it will look like this picture:



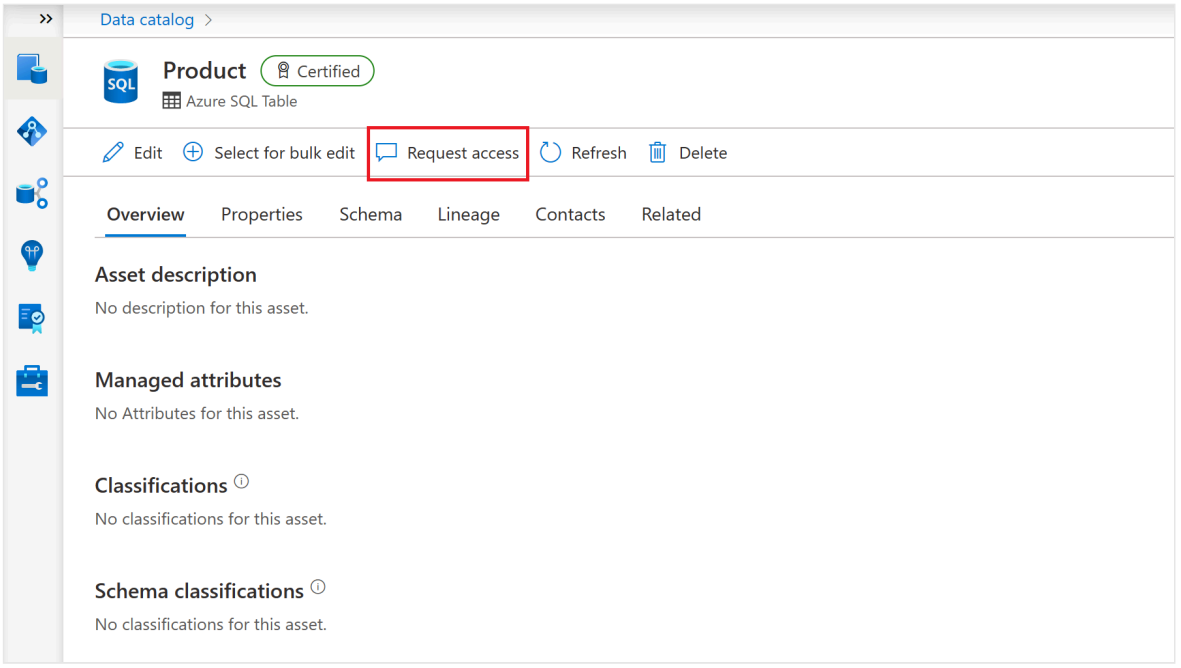
Create a self-service data access request

1. To find a data asset, use Microsoft Purview's search or browse functionality.



2. Select the asset to go to asset details.

3. Select **Request access**.

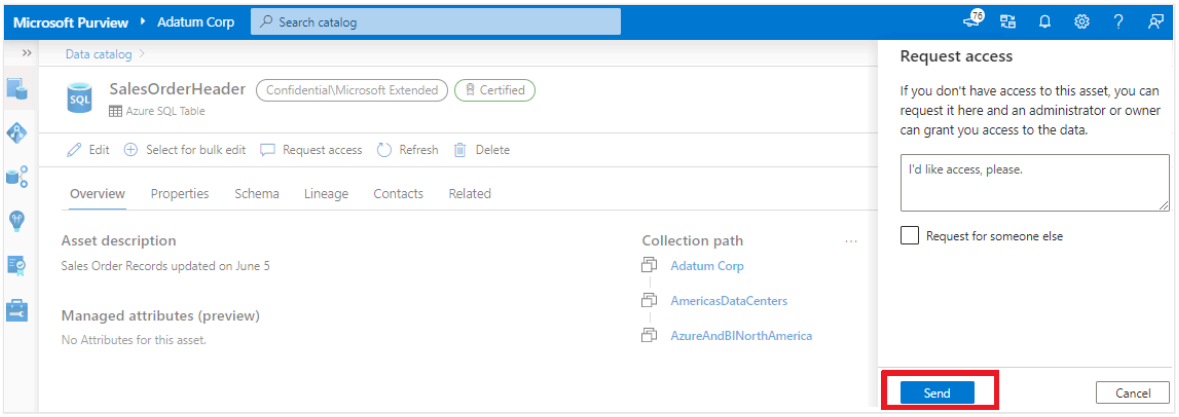


> **Note**
>
> If this option isn't available, a self-service access workflow either hasn't been created, or hasn't been assigned to the collection where the resource is registered. Contact the collection administrator, data source administrator, or workflow administrator of your collection for more information. Or, for information on how to create a self-service access workflow, see our self-service access workflow documentation.

4. The **Request access** window will open. You can provide comments on why data access is requested.

5. Select **Send** to trigger the self-service data access workflow.

> **Note**
>
> If you want to request access on behalf of another user, select the checkbox **Request for someone else** and populate the email id of that user.



> **Note**
>
> A request access to resource set will actually submit the data access request for the folder one level up which contains all these resource set files.

6. Data owners will be notified of your request and will either approve or reject the request.

> **Important**
>
> - Publish is a background operation. Azure Storage accounts can take up to **2 hours** to reflect the changes.

View a self-service policy

To view the policies you've created, follow the article to view the self-service policies.

Data consumption

- Data consumer can access the requested dataset using tools such as Power BI or Azure Synapse Analytics workspace.

> **Note**
>
> Users won't be able to browse to the asset using the Azure portal or Storage explorer if the only permission granted is read/modify access at the file or folder level of the storage account.

> **Caution**

Folder level permission is required to access data in ADLS Gen 2 using Power BI. Additionally, resource sets aren't supported by self-service policies. Hence, folder level permission needs to be granted to access resource set files such as CSV or parquet.

Known issues

Known issues related to Policy creation

- self-service policies aren't supported for Microsoft Purview resource sets. Even if displayed in Microsoft Purview, it isn't yet enforced. Learn more about resource sets.

Next steps

Check blog, demo and related tutorials:

- self-service policies concept
- Demo of self-service policies for storage
- Blog: Accessing data when folder level permission is granted
- Blog: Accessing data when file level permission is granted

Feedback

Was this page helpful? Yes No

Get started with protection policies (preview)

Protection policies help organizations automatically protect sensitive data across data sources. Microsoft Purview already scans data assets and identifies sensitive data elements. Protection policies help you automatically restrict access to that data by using sensitivity labels from Microsoft Purview Information Protection.

Protection policies ensure that enterprise admins must authorize data access for a sensitivity label. After you enable these policies, access control is automatically imposed whenever Information Protection detects sensitive information.

Actions supported

- Restrict access on labeled data assets so that only users and groups you select can access them.
- Configure action on sensitivity labels in the Information Protection solution.

Data sources supported

- Azure Sources:
  - Azure Blob storage
  - Azure Data Lake Storage Gen2
- Microsoft Fabric

> **Tip**
>
> This article covers general steps for all protection policies. The available data source articles cover specifics like available regions, limitations, and specific features for those data sources.

Prerequisites

- Microsoft 365 E5 license and setup of the pay-as-you-go billing model. To understand pay-as-you-go billing by assets protected, see the pay-as-you-go billing model. For information about the specific licenses required, see sensitivity label licensing information. To get a Microsoft 365 E5 trial license for your tenant, sign up from your environment.

Azure sources

- Configure users and permissions.
- Create or extend sensitivity labels from Microsoft Purview Information Protection to Data Map assets.
- Register sources - register any of these sources you like:
  - Azure Blob Storage
  - Azure Data Lake Storage Gen2

  > **Note**
  >
  > To continue, you need to be a data source administrator in the collection where your Azure storage source is registered.

- Enable data policy enforcement:
  - Go to the new Microsoft Purview portal.
  - Select the **Data Map** tab in the left menu.
  - Select the **Data sources** tab in the left menu.
  - Select the source where you want to enable *data policy enforcement*.
  - Set the **Data policy enforcement** toggle to **On**.
- Scan sources - register any of the sources you've registered:
  - Azure Blob Storage
  - Azure Data Lake Storage Gen2

  > **Note**
  >
  > Wait at least 24 hours after scanning.

Users and permissions for Azure sources

There are several types of users you need, and you need to set up the corresponding roles and permissions for these users:

1. Microsoft Purview Information Protection Admin - Broad rights to manage Information Protection solution: reviewing / creating / updating / deleting protection policies, sensitivity labels and label/auto-labeling policies, all classifier types. They should also have full access to data explorer, activity explorer, Microsoft Purview Information Protection insights, and reports.

    - User needs the roles from within the built-in role group “Information Protection”, along with new roles for data map reader, insights reader, scan reader, source reader. Full permissions would be:
        - Information Protection reader
        - Data Map Reader
        - Insights Reader
        - Source Reader
        - Scan Reader
        - Information Protection admin
        - Information Protection analyst
        - Information protection investigator
        - Data classification list viewer
        - Data classification content viewer
        - Microsoft Purview evaluation administrator
    - **Option 1** - Recommended:
        a. Within the Microsoft Purview role groups panel, search for **Information Protection**.
        b. Select the Information Protection role group, select **Copy**.
        c. Name it: "Preview - Information Protection", and select **Create copy**.
        d. Select **Preview - Information Protection** and select **Edit**.
        e. On the **Roles** page, **+ Choose roles** and search for “reader”.
        f. Select these four roles: Data map reader, insights reader, scan reader, source reader.
        g. Add the Microsoft Purview Information Protection admin test user account to this new copied group and complete the wizard.
    - **Option 2**- uses built-in groups (will provide more permissions than needed)
        a. Place a new Microsoft Purview Information Protection admin test user account within the built-in groups for Information Protection, Data Estate Insights Readers, Data Source Administrators.

2. Data Owner/Admin - This user will enable your source for data policy enforcement in Microsoft Purview for Azure and Amazon S3 sources.

Create protection policy

After you check the prerequisites and prepare your Microsoft Purview instance and source for protection policies and wait at least 24 hours after your most recent scan, complete the following steps to create your protection policies:

1. Sign in to the Microsoft Purview portal > **Information Protection** card > **Policies**. If the **Information Protection** solution card isn't displayed, select **View all solutions** and then select **Information Protection** from the **Data Security** section.

2. Select **Protection policies**.

3. Select **+ New protection policy**.

4. Enter a name and description, and select **Next**.

5. Select **+ Add sensitivity label** to add sensitivity labels to detect for the policy, and select all the labels you want the policy to apply to.

6. Select **Add** and then select **Next**.

7. Select the sources you want to apply the policy to.
   a. For Microsoft Fabric sources, select only Fabric and select **Next**. (For more information, see the Fabric documentation.)
   b. For Azure sources, you can select multiple sources. Select the **Edit** button to manage the scope for each source you select.

8. Depending on your source, select the **+ Include** button at the top to add up to 10 resources to your scope list. The policy applies to all the resources you select.

   > **Note**
   >
   > Currently, a maximum of 10 resources are supported, and you must select them under **Edit** for them to be enabled. If the Azure storage account enables a customer-managed key, the storage account doesn't work for protection policies.

9. Select **Add** and then select **Done** when your list of sources is complete.

10. Depending on your source, select the type of protection policy you want to create.
    a. For Fabric sources, follow the protection policies for Fabric documentation.
    b. For other sources, select the users who are **NOT** denied access based on the label. Everyone in your org is denied access to labeled items except for the users and groups you add here.

> **Important**
>
> Ensure that you include any service principal accounts running Microsoft Purview Data Governance Scans for the scoped sources in a security group. Add this security group to the allowed list of your protection policies. This action prevents future scans from being blocked by access restrictions on labeled files.

1. Select **Next**.
2. Select whether to turn on the policy right away or not, and select **Next**.
3. Select **Submit**.
4. Select **Done**.
5. You should now see your new policy in the list of protection policies. Select it to confirm that all the details are correct.

Manage protection policy

To edit or delete an existing protection policy, follow these steps:

1. Open the Microsoft Purview portal.
2. Open the **Information Protection** solution.
3. Select **Policies**, then select **Protection policies**.
4. Select the policy you want to manage.
5. To change any of the details, select the **Edit policy** button.
6. To delete the policy, select **Delete policy**.

Next steps

- Protection policies for Azure sources: Learn about region availability, limitations, and specific features for Azure data sources.
- Protection policies for Fabric: Learn about protection policies for Fabric items.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 07/01/2026

Authoring and publishing protection policies for Azure sources (preview)

Protection access control policies (protection policies) enable organizations to automatically protect sensitive data across data sources. Microsoft Purview already scans data assets and identifies sensitive data elements. This new feature allows you to automatically restrict access to that data by using sensitivity labels from Microsoft Purview Information Protection.

Protection policies ensure that Enterprise Admins must authorize data access for a sensitivity type. After you enable these policies, access control is automatically imposed whenever Information Protection detects sensitive information.

Actions supported

- Restrict access on labeled data assets so that only users and groups you select can access them.
- Configure action on Sensitivity Labels in the Information Protection solution.

Sources supported

- Azure Blob Storage
- Azure Data Lake Storage Gen2

Prerequisites

- Microsoft 365 E5 license and setup of the pay-as-you-go billing model. To understand pay-as-you-go billing by assets protected, see the pay-as-you-go billing model. For information about the specific licenses required, see sensitivity label licensing information. To get a Microsoft 365 E5 trial license for your tenant, sign up from your environment.

- Existing Purview accounts upgraded to Microsoft Purview single tenant model and new Portal experience, using the enterprise version of Microsoft Purview.

- Configure users and permissions.

- Create or extend sensitivity labels from Microsoft Purview Information Protection to Data Map assets.

- Register sources - register any of these sources you like:
  - Azure Blob Storage
  - Azure Data Lake Storage Gen2

  > ⓘ **Note**
  >
  > To continue, you need to be a **data source administrator** in the collection where your Azure storage source is registered.

- Enable data policy enforcement:
  - Go to the new Microsoft Purview portal.
  - Select the **Data Map** tab in the left menu.
  - Select the **Data sources** tab in the left menu.
  - Select the source where you want to enable *data policy enforcement*.
  - Set the **Data policy enforcement** toggle to **On**.

- Scan sources - register any of the sources you've registered:
  - Azure Blob Storage
  - Azure Data Lake Storage Gen2

  > ⓘ **Note**
  >
  > Wait at least 24 hours after scanning.

Users and permissions

There are several types of users you need, and you need to set up the corresponding roles and permissions for these users:

1. Microsoft Purview Information Protection Admin - Broad rights to manage Information Protection solution: reviewing / creating / updating / deleting protection policies, sensitivity labels and label/auto-labeling policies, all classifier types. They should also have full access to data explorer, activity explorer, Microsoft Purview Information Protection insights, and reports.

   - User needs the roles from within the built-in role group “Information Protection”, along with new roles for data map reader, insights reader, scan reader, source reader. Full permissions would be:

    - Information Protection reader
    - Data Map Reader
    - Insights Reader
    - Source Reader
    - Scan Reader
    - Information Protection admin
    - Information Protection analyst
    - Information protection investigator
    - Data classification list viewer
    - Data classification content viewer
    - Microsoft Purview evaluation administrator
- **Option 1** - Recommended:
    a. Within the Microsoft Purview role groups panel, search for **Information Protection**.
    b. Select the Information Protection role group, select **Copy**.
    c. Name it: "Preview - Information Protection", and select **Create copy**.
    d. Select **Preview - Information Protection** and select **Edit**.
    e. On the **Roles** page, **+ Choose roles** and search for “reader”.
    f. Select these four roles: Data map reader, insights reader, scan reader, source reader.
    g. Add the Microsoft Purview Information Protection admin test user account to this new copied group and complete the wizard.
- **Option 2**- uses built-in groups (will provide more permissions than needed)
    a. Place a new Microsoft Purview Information Protection admin test user account within the built-in groups for Information Protection, Data Estate Insights Readers, Data Source Administrators.

2. Data Owner/Admin - This user will enable your source for data policy enforcement in Microsoft Purview for Azure and Amazon S3 sources.

Create protection policy

After you check the prerequisites and prepare your Microsoft Purview instance and source for protection policies and wait at least 24 hours after your most recent scan, complete the following steps to create your protection policies:

1. Sign in to the Microsoft Purview portal and select the **Information Protection** card. If the **Information Protection** solution card isn't displayed, select **View all solutions** and then select **Information Protection** from the **Data Security** section.

2. On the left navigation, select select **Policies**, then select **Protection policies**.

3. Select **New protection policy**.
4. Provide a name and description and select **Next**.
5. Select + **Add sensitivity label** to add sensitivity label(s) to detect for the policy, and select all the labels you want the policy to apply to.
6. Select **Add** then select **Next**.
7. Select the sources you want to apply the policy to. You can select multiple sources. Select **Edit** to manage the scope for each you select.
8. Depending on your source, select the + **Include** button at the top to add up to 10 resources your scope list. The policy will be applied to all the resources you select.

   > ⓘ **Note**
   >
   > Currently a maximum of 10 resources is supported, and they must be selected under **Edit** for them to be enabled.

9. Select **Add** and then select **Done** when your list of sources is complete.
10. Depending on your source, select the type of protection policy you want to create.
11. Select the users who will **NOT** be denied access based on the label. Everyone in your org will be denied access to labeled items except for the users and groups you add here.

    > ⓘ **Important**
    >
    > Ensure that any service principal accounts running data governance scans for the scoped sources are included in a Security Group. Add this Security Group to the allowed list of your protection policies. This will prevent future scans from being blocked by access restrictions on labeled files.

12. Select **Next**.
13. Choose whether turn on the policy right away or not, and select **Next**.
14. Select **Submit**.
15. Select **Done**.

You should now see your new policy in the list of protection policies. Select it to confirm that all the details are correct.

Manage protection policy

To edit or delete an existing protection policy, follow these steps:

1. Open the Microsoft Purview portal.
2. Open the **Information Protection** solution.
3. Select **Policies**, then select **Protection policies**.
4. Select the policy you want to manage.
5. To change any of the details, select the **Edit policy** button.
6. To delete the policy, select **Delete policy**.

Next steps

- Get started with protection policies: Learn about protection policies across all supported data sources.
- Register and scan Azure Data Lake Storage Gen2: Register your ADLS Gen2 source and apply protection policies.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 07/01/2026

Protection policies in Microsoft Fabric

Microsoft Purview protection access control policies (protection policies) enable organizations to control access to items in Fabric using sensitivity labels.

The target audience for this article is security and compliance admins, Fabric admins and users, and anyone else who wants to learn about how protection policies control access to items in Fabric. If you want to see how to create a protection policy for Fabric, see Create and manage protection policies for Fabric.

How do protection policies for Fabric work?

Each protection policy for Fabric is associated with a sensitivity label. The policy controls access to an item that has the associated label by allowing users and groups specified in the policy to retain permissions they have on the item, while blocking access for everyone else.

The policy can allow specified users and groups to retain full control on the labeled item if they currently have it. All other users and groups are blocked from accessing the item.

> **Note**
>
> A protection policy doesn't apply to a label issuer. That is, the user that last applied a label associated with a protection policy to an item won't be denied access to that item, even if they aren't specified in the policy. For example, if a protection policy is associated with label A, and a user applies label A to an item, that user will be able to access the item even if they're not specified in the policy.

View an item's restricted users

In the Fabric admin portal, you can use the OneLake catalog to view an item's permissions and determine which users are restricted from accessing it. The Permissions tab in the item's details is visible to you if you have a role of Admin or Member in the workspace containing the item.

To view an item's permissions in the Fabric admin portal, open the **OneLake catalog**, locate the item (optionally filter by workspace), and select its name to open the item details. Then, select the **Permissions** tab to see the list of users and groups that have access to the item, including those restricted by a protection policy.

Use cases

Protection policies are useful in scenarios where organizations need to restrict access to sensitive items. For example, an organization might want to ensure that only specific internal users can access items labeled as "Confidential," while blocking access for everyone else.

Who creates protection policies for Fabric?

Protection policies for Fabric are generally configured by an organization's Purview security and compliance teams. The protection policy creator must have role of Information Protection Admin or higher. For more information, see Create and manage protection policies for Fabric.

Requirements

- A Microsoft 365 E3/E5 license is required for sensitivity labels from Microsoft Purview Information Protection. For more information, see Microsoft Purview Information Protection: Sensitivity labeling.
- At least one "appropriately configured" sensitivity label from Microsoft Purview Information Protection must exist in the tenant. "Appropriately configured" in the context of protection policies for Fabric means that when the label was configured, it was scoped to **Files & other data assets**, and its protection settings were set to include **Control access** (for information about sensitivity label configuration, see Create and configure sensitivity labels and their policies). Only such "appropriately configured" sensitivity labels can be used to create the protection policies for Fabric.
- For protection policies to be enforced in Fabric, the Fabric tenant setting **Allow users to apply sensitivity labels for content** must be enabled. This setting is required for all sensitivity label related policy enforcement in Fabric, so if sensitivity labels are already being used in Fabric, this setting will already be on. For more information about enabling sensitivity labels in Fabric, see Enable sensitivity labels.

Supported item types

Protection policies are supported for all native Fabric item types, including lakehouses, notebooks, pipelines, and other core Fabric assets.

Additionally, protection policies are supported for Power BI semantic models. Other Power BI item types, such as reports and dashboards, aren't currently supported.

Considerations and limitations

- Up to 50 protection policies can be created.
- Up to 100 users and groups can be added to a protection policy.
- Protection policies for Fabric don't support guest/external users.
- Protection policies don't integrate with Fabric CI/CD solutions. Deployment pipelines and Git integration use only workspace permissions, not item-level permissions from protection policy labels.
- After a policy has been created, it can take up to 24 hours for it to start detecting and protecting items labeled with the sensitivity label that was associated with the policy.
- For native Fabric item types, if the label is applied automatically by the system—such as in the case of downstream inheritance—and there is no designated label issuer, the user who created the artifact will not be restricted by protection policies.

Related content

- Create and manage protection policies for Fabric
- Authoring and publishing protection policies

Last updated on 12/03/2025

Microsoft Purview metamodel

Article • 01/28/2025

> **Note**
>
> Microsoft Purview Data Map (classic), Data Catalog (classic), Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

> **Important**
>
> This feature is currently in preview. The Supplemental Terms of Use for Microsoft Azure Previews include additional legal terms that apply to Azure features that are in beta, in preview, or otherwise not yet released into general availability.

> **Note**
>
> On April 22, 2024 asset types moved to be under the Microsoft Purview Data Catalog feature in the Microsoft Purview portal. Access instructions have been updated.

Metamodel is a feature in the Microsoft Purview Data Catalog that helps add rich business context to your data catalog. It tells a story about how your data is grouped in data domains, how it's used in business processes, what projects are impacted by the data, and ultimately how the data fits in the day to day of your business.

The context metamodel provides is important because business users, the people who are consuming the data, often have nontechnical questions about the data. Questions like: What department produces this dataset? Are there any projects that are using this dataset? Where does this report come from?

When you scan data into Microsoft Purview, the technical metadata can tell you what the data looks like, if the data has been classified, or if it has glossary terms assigned, but it can't tell you where and how that data is used. The metamodel gives your users that information. It can tell you what data is mission critical for a product launch or used by a high performing sales team to convert leads to prospects.

So the metamodel not only helps data consumers find the data they're looking for, but it also tells your data managers what data is critical and how healthy that critical data is,

so you can better manage and govern it. For example: you could have different privacy obligations if you use personal information for marketing activities vs. analytics that improve a product or service, and metamodel can help you determine where data is being used.

So what is the metamodel?

What does it look like? The metamodel is built from **assets** and the **relationships** between them.

For example, you might have a sales reporting team (asset) that consumes data (relationship) from some SQL tables (assets).

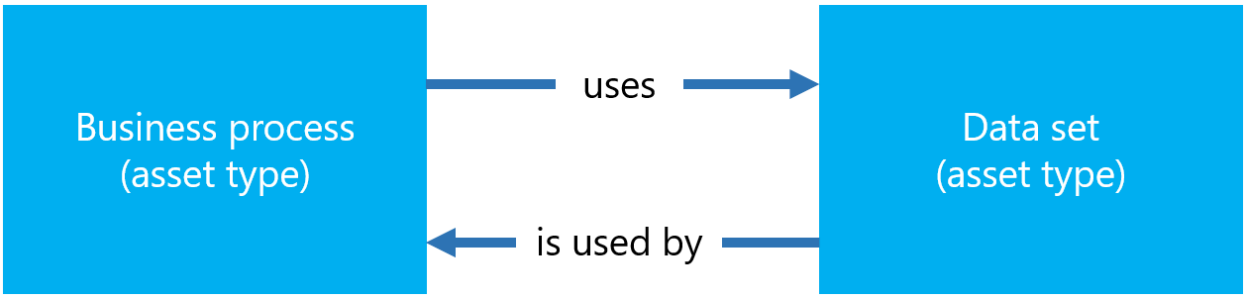
When you scan data sources into the data map, you already have the technical data assets like SQL tables available for your metamodel. But what about assets like a sales reporting team, or a marketing strategy that represent processes or people instead of a data source? Metamodel provides **asset types** that allow you to describe other important parts of your business.

An **asset type** is a template for important concepts like business processes, departments, lines of business, or even products. They're the building blocks you'll use to describe how data is used in your business. The **asset type** creates a template you can use over and over to describe specific assets. For example, you can define an asset type "department" and then create new department assets for each of your business departments. These new assets are stored in Microsoft Purview like any other data assets that were scanned in the data map, so you can search and browse for them in the data catalog.

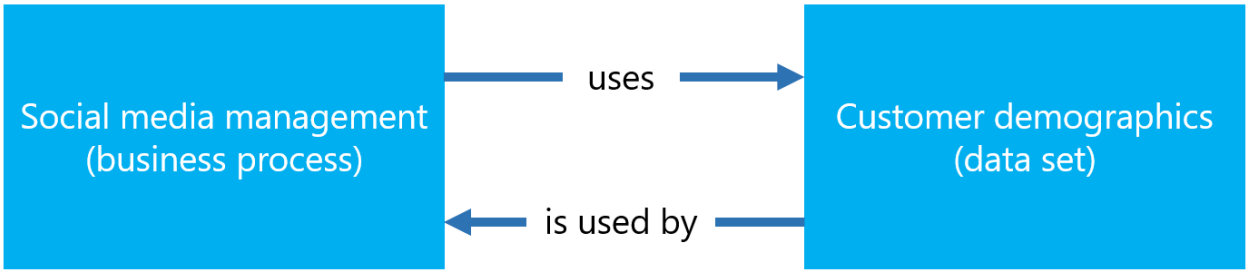
Metamodel includes several predefined asset types to help you get started, but you can also create your own.

Similarly a **relationship definition** is the template for the kinds of interactions between assets that you want to represent. For example, a department *manages* a business process. An organization *has* departments. The business process, organization, and departments are the assets, "manages" and "has" are the relationships between them.

For example, if we want to use Microsoft Purview to show how key data sets are used in our business processes, we can represent that information as a template:



Which we can then use to describe how a specific business process uses a specific data set:

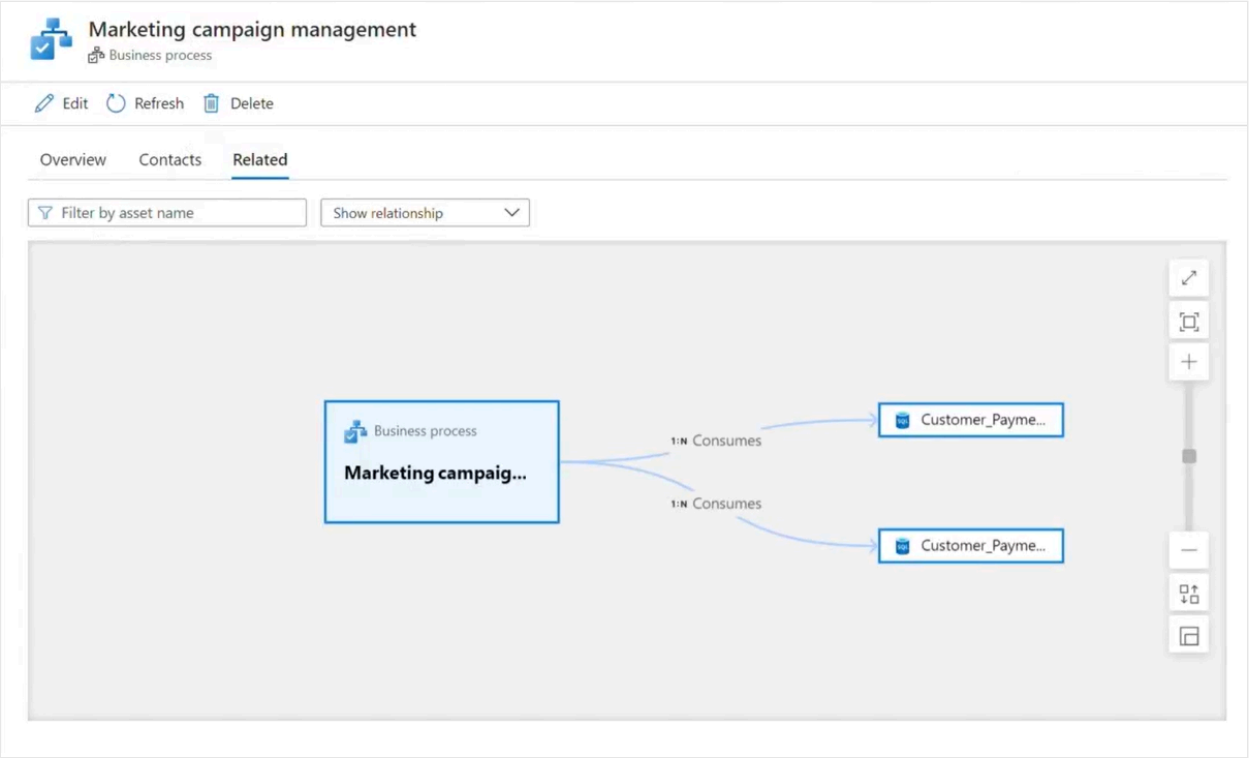


Metamodel example

For a simple example of a metamodel, let's consider marketing campaign management in a business. This process is an asset that we'll add to our metamodel. It's not a data source we can scan, since it's a set of business processes, but during this process real data will be used and referenced. It's important to show what data is being used and how, so we can properly manage and govern it. We'll create an asset for the marketing campaign management from the asset type "Business Process".

We know there are two tables in SQL that the marketing campaign management team uses. These are assets too, but they were data source assets created when the SQL database was scanned into the Microsoft Purview Data Map.

The relationship between the marketing campaign management asset and the SQL table assets is that campaign management **consumes**, or uses, the SQL tables when developing campaigns. We can record that in our metamodel, so now anyone that looks at those SQL tables can see that they're being used to develop marketing campaigns. We'll also be able to see if there are any other teams that use this data, or maybe which department develops this data as well. So now, with metamodel, we know not only what the data is, but we have a story about how it's being used that helps us understand and manage it.



And now that marketing campaign management is an asset in your data catalog like any other asset, so you can find it in a search. A user could search for that process and quickly know what data it uses and produces, without needing to know anything about the data beforehand.

Next steps

If you're ready to get started, follow the how to create a metamodel article.

Feedback

Was this page helpful? Yes No

Manage assets with metamodel

Article • 10/28/2024

> **Note**
>
> Microsoft Purview Data Map (classic), Data Catalog (classic), Data Health Insights (classic) are no longer taking on new customers and these services, previously Azure Purview, are now in customer support mode.

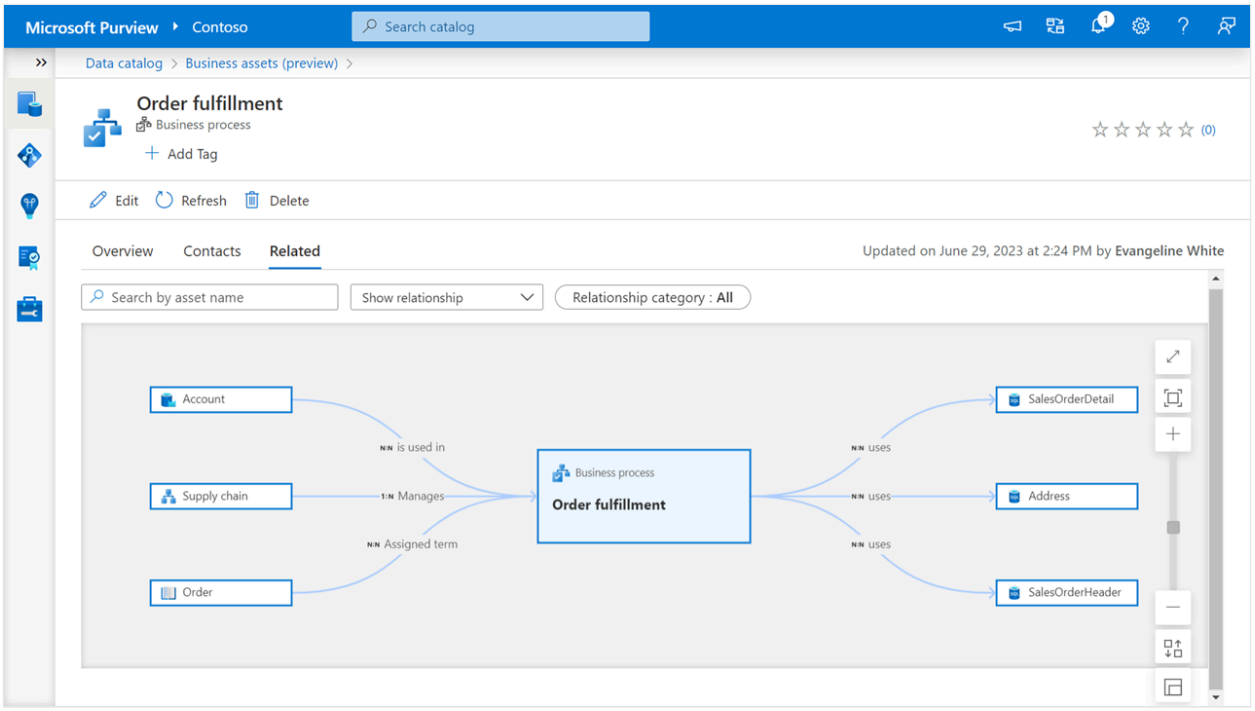
> **Important**
>
> This feature is currently in preview. The Supplemental Terms of Use for Microsoft Azure Previews include additional legal terms that apply to Azure features that are in beta, in preview, or otherwise not yet released into general availability.

> **Note**
>
> On April 22, 2024 asset types moved to be under the Microsoft Purview Data Catalog feature in the Microsoft Purview portal. Access instructions have been updated.

Metamodel is a feature in the Microsoft Purview Data Map that gives the technical data in your data map relationships and reference points to make it easier to navigate and understand in the day to day. Like adding streets and cities to a map, the metamodel orients users so they know where they're and can discover the information they need.

Create additional asset types for describing real-world things in your organization such as departments, projects, products, and lines of business. When you look at your data estate, it's often helpful to understand how your data fits into your business. Use business assets whenever you want to associate data assets to specific organizational structures, business processes, or other anything else that could be convincingly modeled as an entity.

In the example below, I describe more business context for the SalesOrderDetail table by showing that the Supply chain department (a business asset) manages the order fulfillment business process (a business asset) which uses the SalesOrderDetail table. Visualizing business context in this way can help others identify the “official” dataset that's used for a particular business purpose and understand whether data is being used compliantly.



This article will get you started in building a metamodel for your Microsoft Purview Data Map.

Prerequisites

- An Azure account with an active subscription. Create an account for free.
- Create a new, or use an existing Microsoft Purview account. You can follow our quick-start guide to create one.
- Data Curator role on the collection where the data asset is housed and/or the root collection, depending on what you need. See the guide on managing Microsoft Purview role assignments.
  - Create and modify asset types, modify assets - Data Curator on the collection where the data asset is housed. An asset will need to be moved to your collection after creation for you to be able to modify it.
  - Create and modify assets - Data curator on the root collection.

> **Note**
>
> Even if you are using the new data catalog experience these permissions are required.

Current limitations

> **Note**

Since this feature is in preview, available abilities are regularly updated.

- When a new asset created, you have to refresh the asset to see relationships
- You can't set relationships between two data assets in the Microsoft Purview governance portal
- The related tab only shows a "business lineage" view for business assets, not data assets

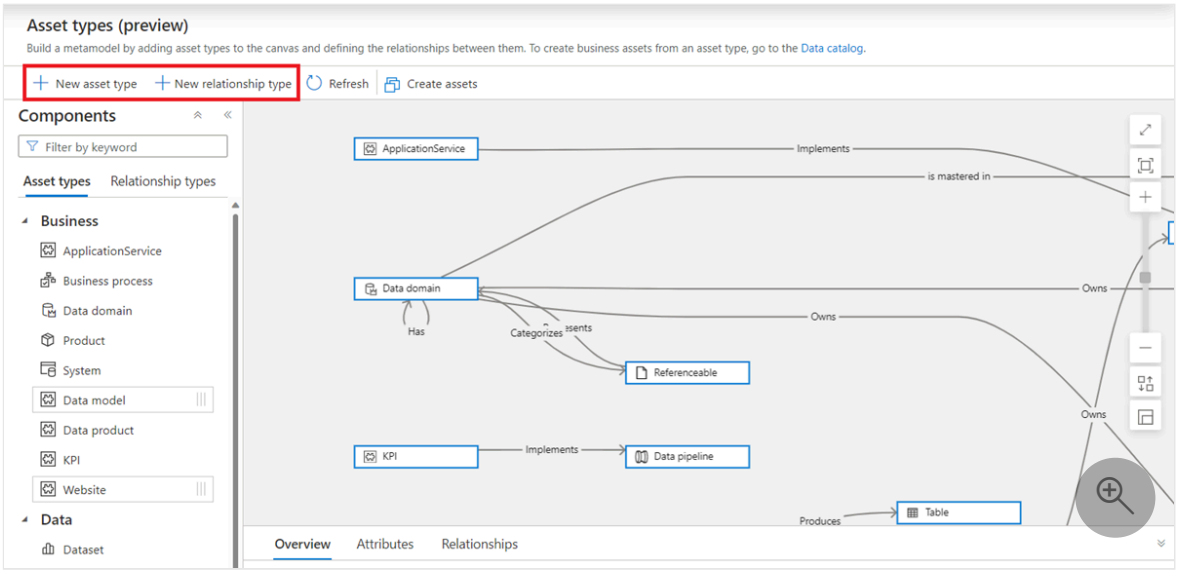
Create and modify asset types

1. To open asset types in the Microsoft Purview Data Catalog:

    - If you're using the new data catalog experience in the Microsoft Purview portal (https://purview.microsoft.com):
        a. Open the data catalog application
        b. Select the **Data management** dropdown
        c. Select **Governance domains**
        d. Select a domain where you want to add your managed attribute.
        e. Select the **Governance domains** header
        f. Select **Asset types**
    - If you're using the classic data catalog experience in the Microsoft Purview portal (https://purview.microsoft.com):
        a. Open the data catalog application
        b. Select the **Metamodel** dropdown
        c. Select **Asset types**
    - If you're using the classic Microsoft Purview governance portal (https://web.purview.azure.com):
        a. Open the Microsoft Purview Data Map
        b. Select **Asset types** under the **Metamodel** section.

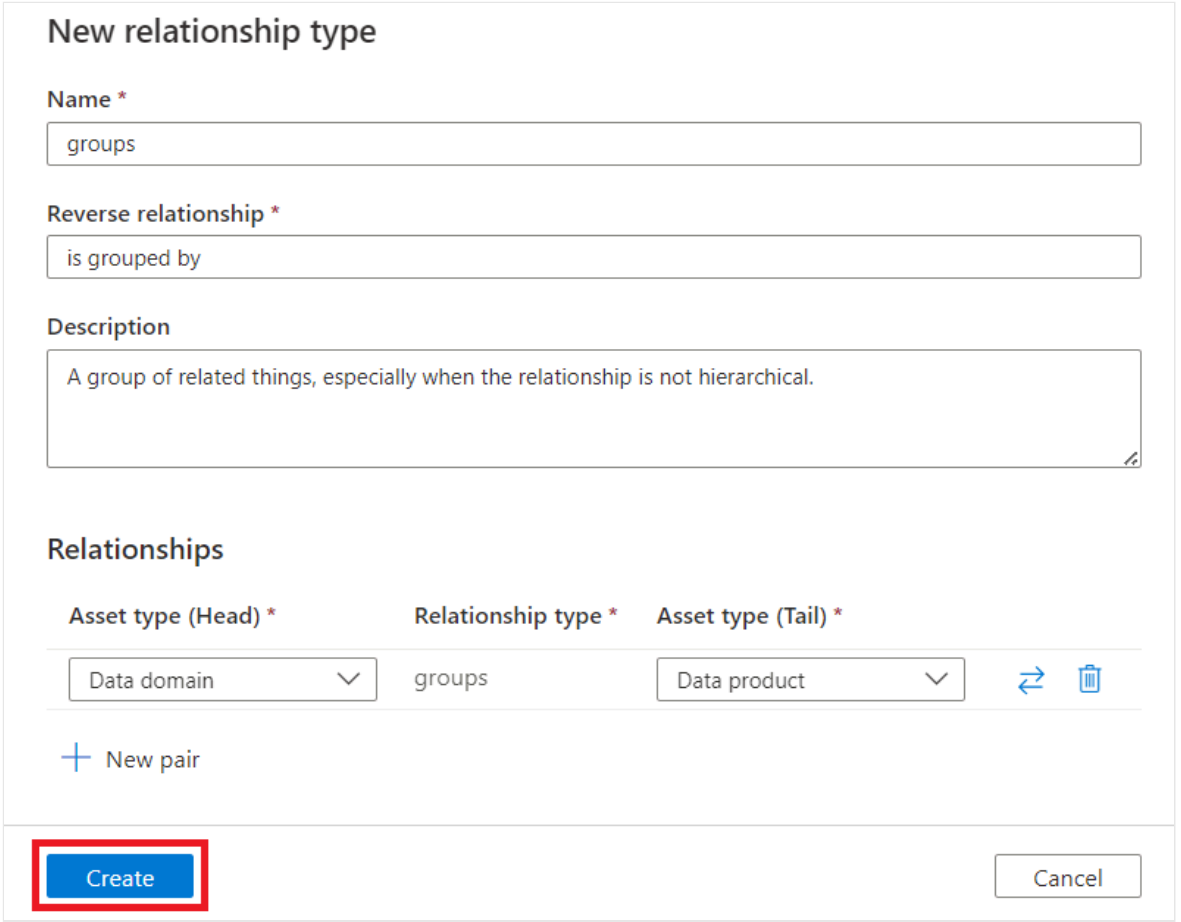
2. You'll see a list of available asset types. Predefined asset types will have unique icons. All custom assets are designated with a puzzle piece icon.

3. To create a new asset type, select **New asset type** and add a name, description, and attributes.

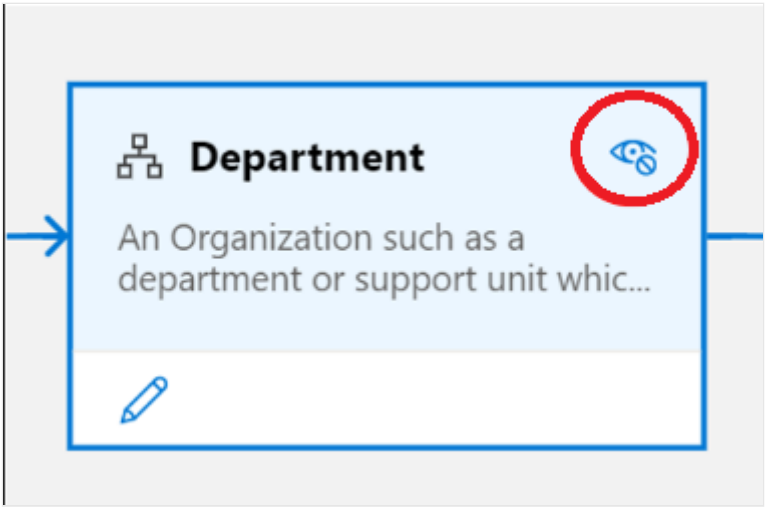


4. To define a relationship between two asset types, select **New relationship type**.

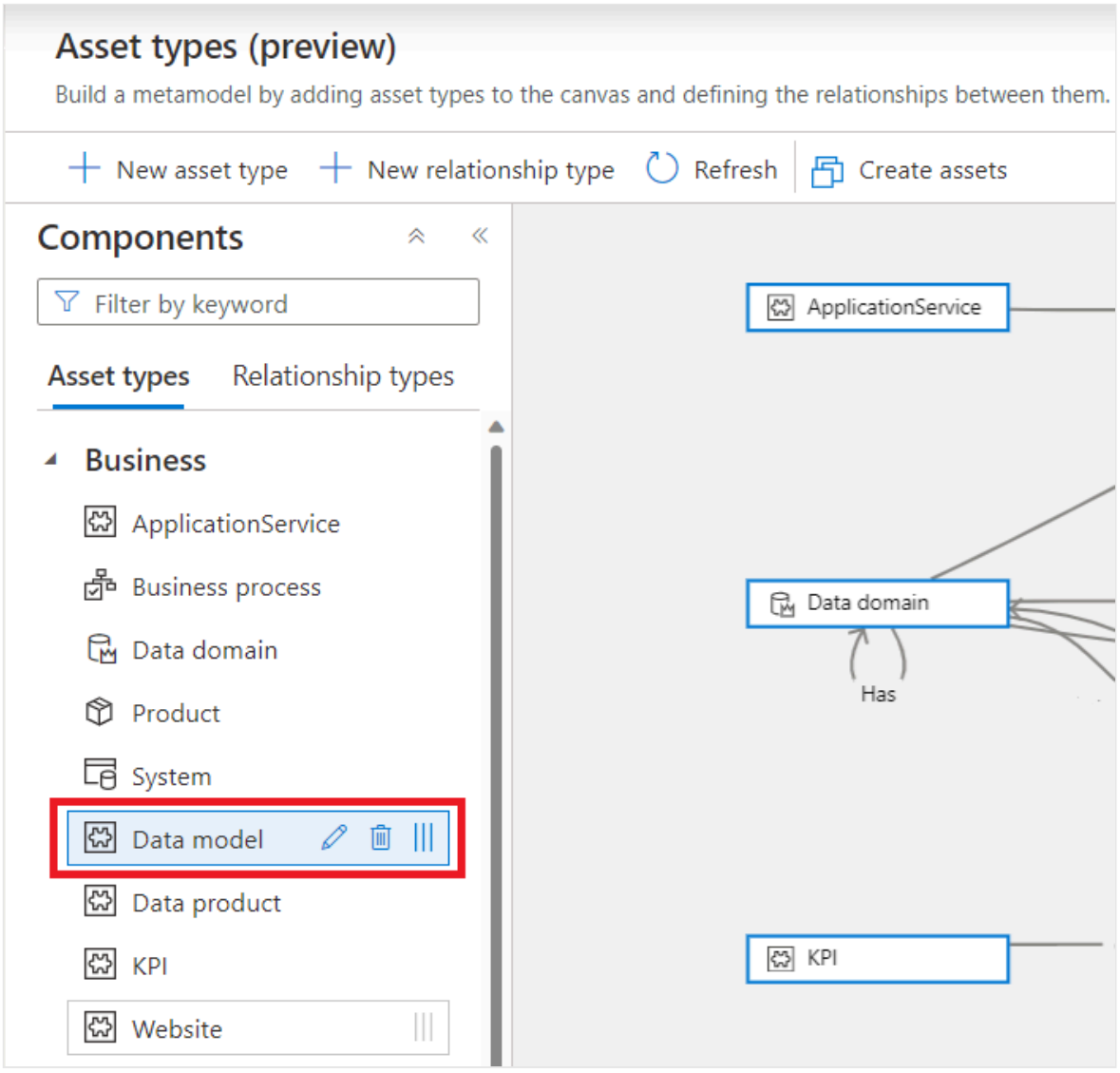
5. Give the relationship a name and define its reverse direction. Assign it to one or more pairs of assets. Select **Create** to save your new relationship type.



6. As you create more asset types, your canvas might get crowded with asset types. To hide an asset from the canvas, select the eye icon on the asset card.



7. To add an asset type back to the canvas, drag it from the left panel.



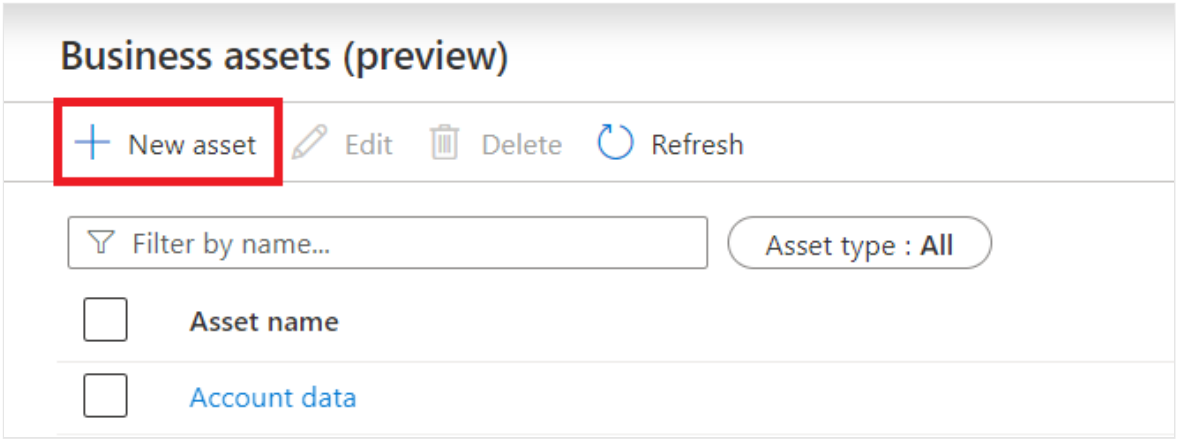
Create and modify assets

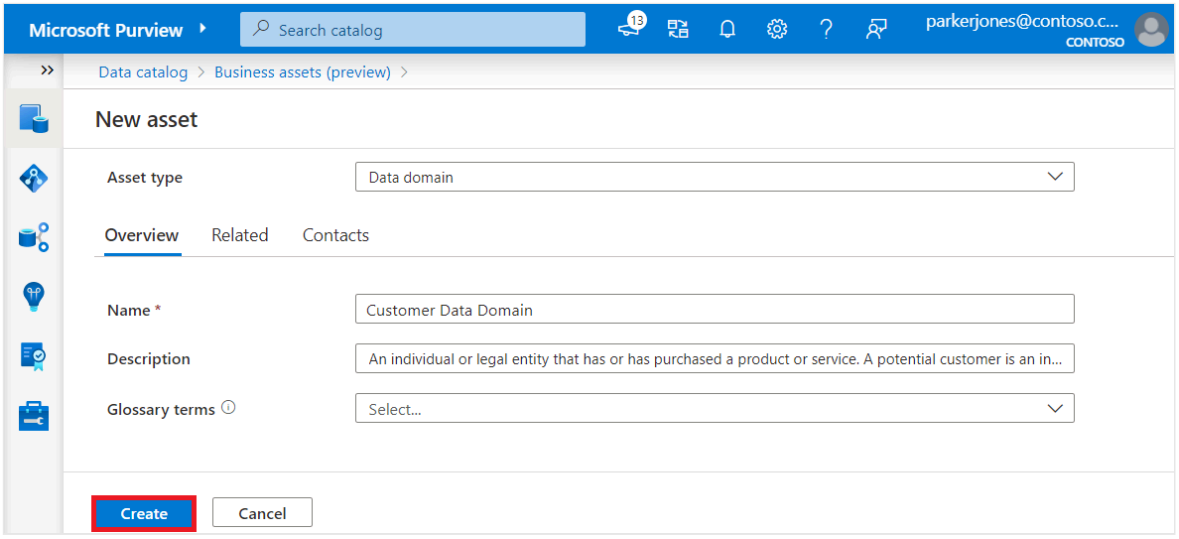
1. When you're ready to begin working with assets, navigate to your data catalog. If you're using the classic Microsoft Purview portal, select **Business assets**. If you're

using the new Microsoft Purview portal under **Business glossary** select **Business assets**.

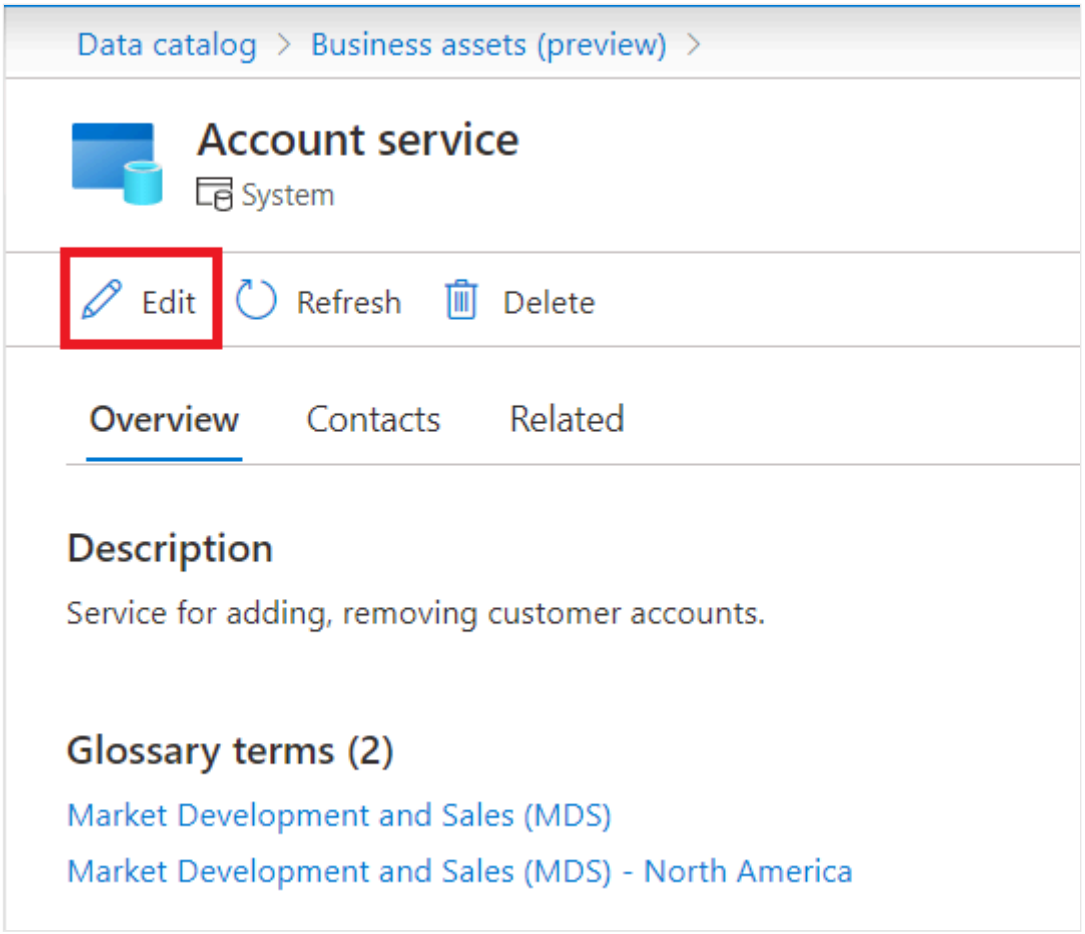
> **Note**
>
> If you're using the new data catalog experience, you'll need to switch back to the original experience to update your business assets.

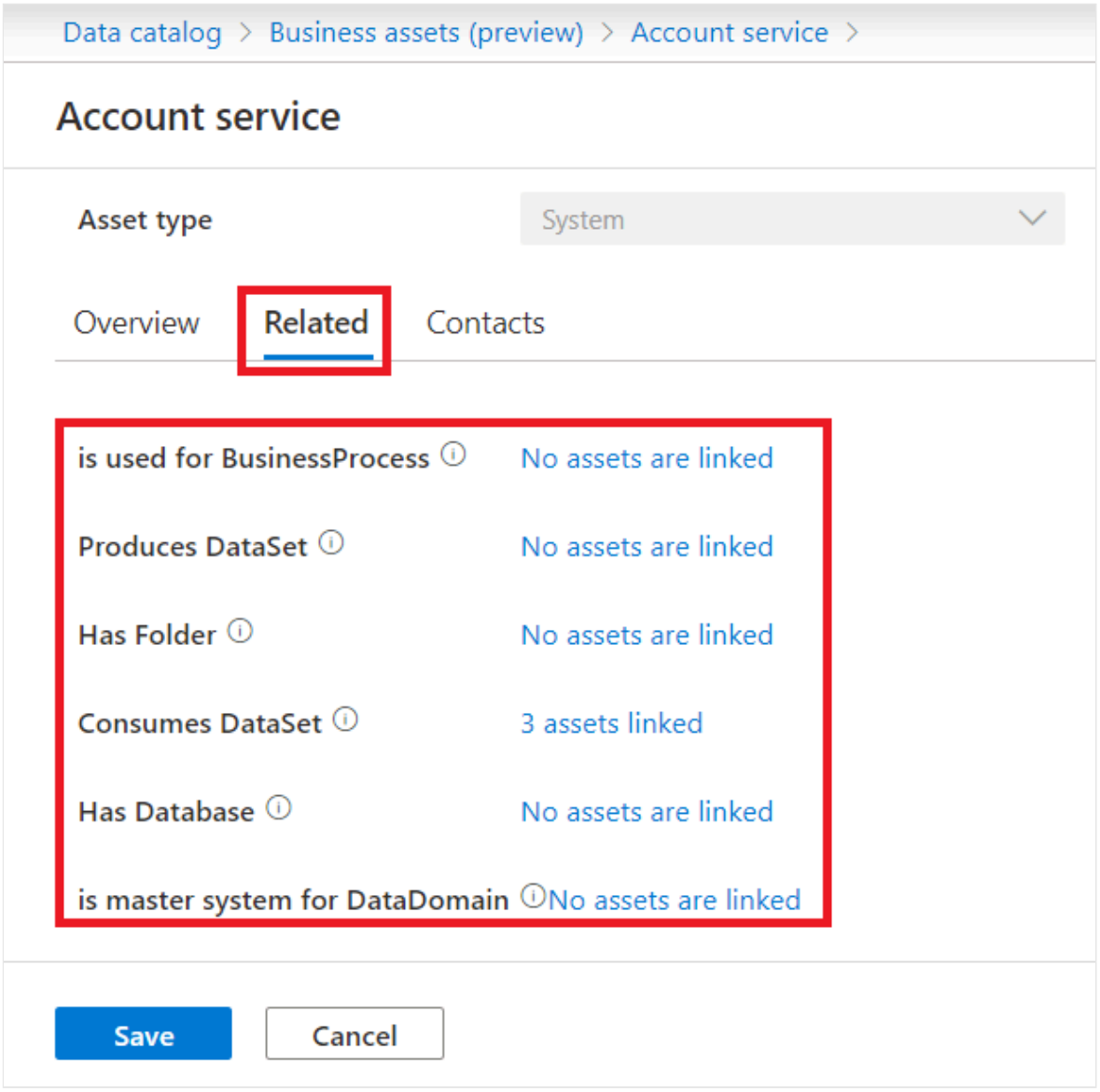
2. To create a new asset, select **New asset**, select the asset type from the drop-down menu, give it a name, description, and complete any required attributes. Select **Create** to save your new asset.



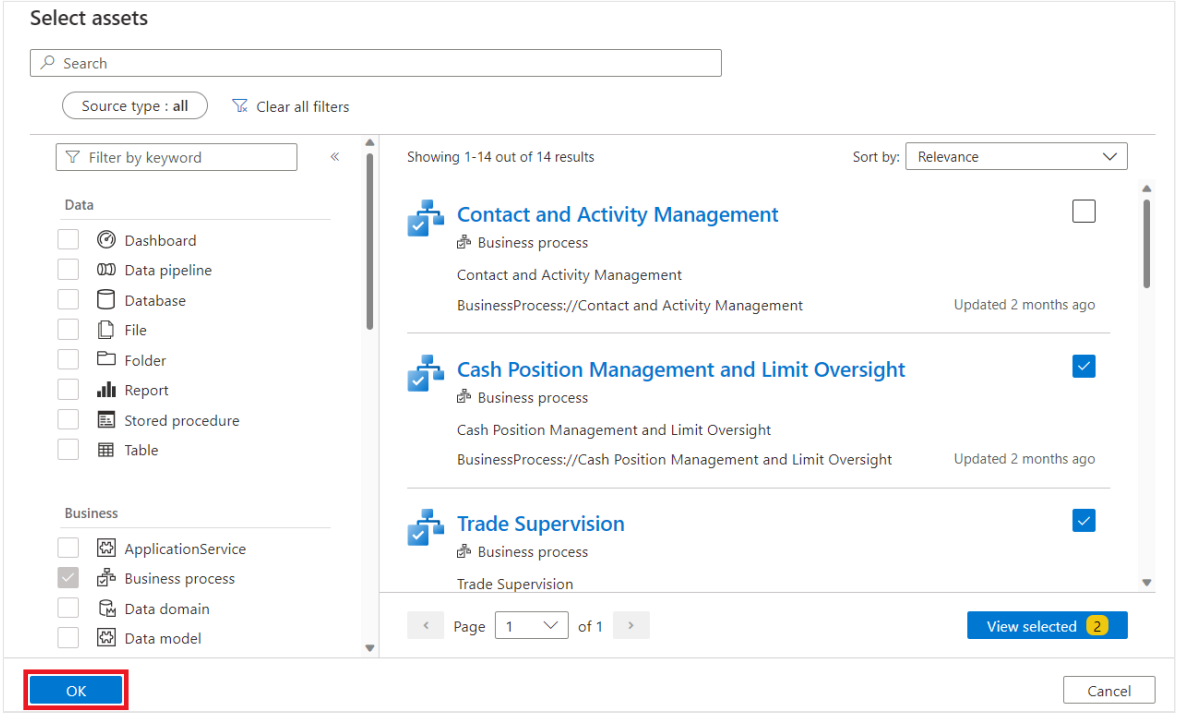


3. To establish a relationship between two assets, go to the asset detail page and select **Edit** > **Related**, and the relationship you'd like to populate.


Data catalog > Business assets (preview) >
Account service
System
Edit
Refresh
Delete
Overview
Contacts
Related
Description
Service for adding, removing customer accounts.
Glossary terms (2)
Market Development and Sales (MDS)
Market Development and Sales (MDS) - North America

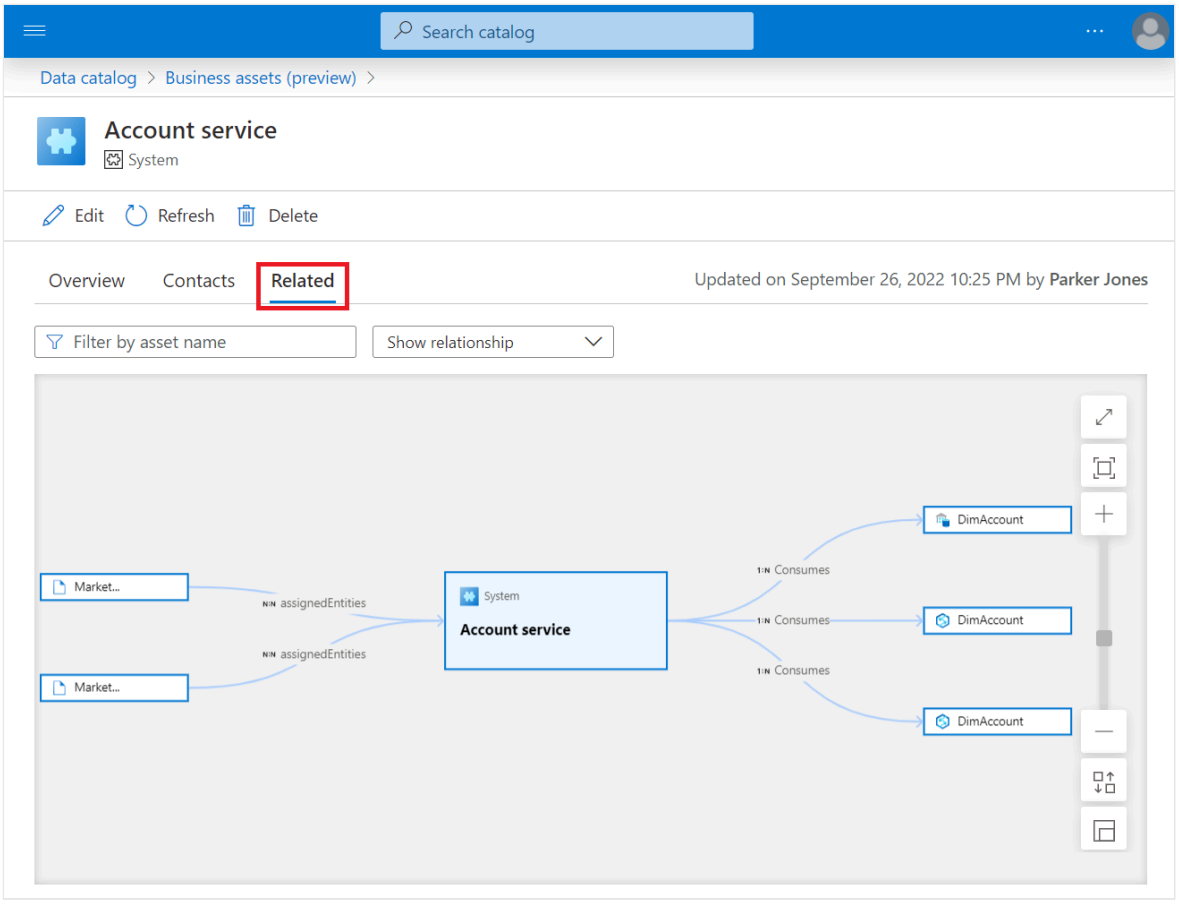


4. Select the assets or assets you'd like to link from the data catalog and select **OK**.



5. Save your changes. You can see the relationships you established in the asset overview.

6. In the **Related** tab of the asset you can also explore a visual representation of related assets.



> **Note**
>
> This is the experience provided by default from Atlas.

Predefined asset types

An asset type is a template for storing a concept that's important to your organization—anything you might want to represent in your data map alongside your physical metadata. You can create your own, but Purview also comes with a prepackaged set of business asset types you can modify to meet your needs.

Expand table

Asset Type	Description
Application service	A well-defined software component, especially one that implements a specific business function such as on-boarding a new customer, taking an order, or sending an invoice.
Business process	A set of activities that are performed in coordination in an organizational or technical environment that jointly realizes a business goal.
Data Domain	A category of data that is governed or explicitly managed for master data management.
Department	An organizational subunit that only has full recognition within the context of that organization. A department wouldn't be regarded as a legal entity in its own right.
Line of business	An organization subdivision focused on a single product or family of products.
Organization	A collection of people organized together into a community or other social, commercial, or political structure. The group has some common purpose or reason for existence that goes beyond the set of people belonging to it and can act as a unit. Organizations are often decomposable into hierarchical structures.
Product	Any offered product or service.
Project	A specific activity used to control the use of resources and associated costs so they're used appropriately in order to successfully achieve the project's goals, such as building a new capability or improving an existing capability.
System	An IT system including hardware and software.

Next steps

For more information about the metamodel, see the metamodel concept page.

Feedback

Was this page helpful? Yes No

Audit logs, diagnostics, and activity history

Article • 05/20/2025

This tutorial lists the step-by-step configuration required to enable and capture audit and diagnostics logs for the Microsoft Purview Data Map via Azure Event Hubs.

A Microsoft Purview administrator or Microsoft Purview data-source admin needs the ability to monitor audit and diagnostics logs captured from Data Map. Audit and diagnostics information consists of the timestamped history of actions taken and changes made to a Microsoft Purview account by every user. Captured activity history includes actions made using the Microsoft Purview portal, or using the REST API.

This tutorial takes you through the steps to enable audit logging. It also shows you how to configure and capture streaming audit events from the Microsoft Purview portal via Azure Diagnostics event hubs.

Enable audit and diagnostics

The following sections walk you through the process of enabling audit and diagnostics.

Configure Event Hubs

Create an Azure Event Hubs namespace by using an Azure Resource Manager (ARM) template (GitHub). This automated Azure ARM template will deploy and finish creating your Event Hubs instance with the required configuration.

For step-by-step explanations and manual setup:

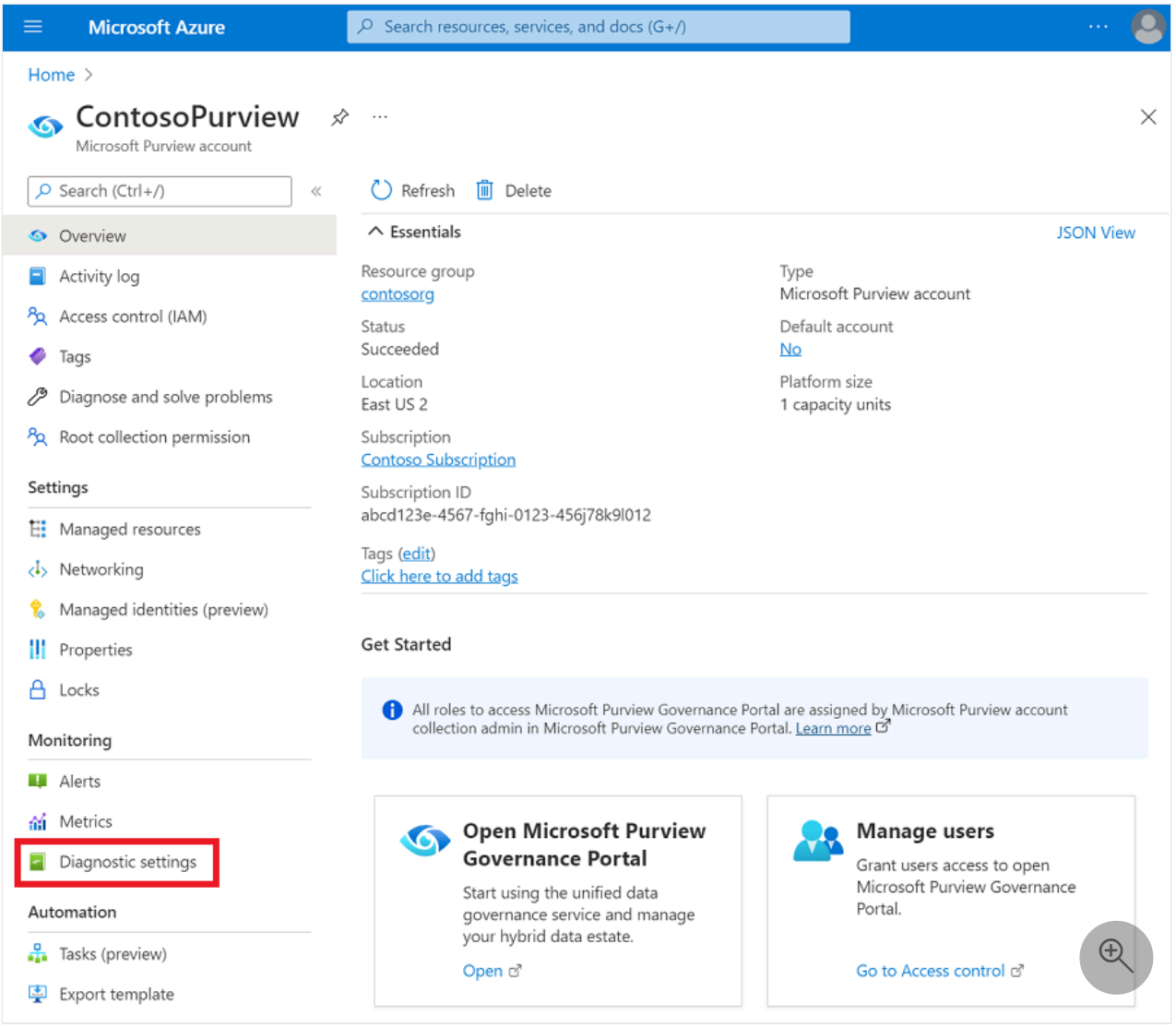
- Event Hubs: Use an ARM template to enable Event Hubs capture
- Event Hubs: Enable capturing of events streaming manually by using the Azure portal

Connect a Microsoft Purview account to Diagnostics event hubs

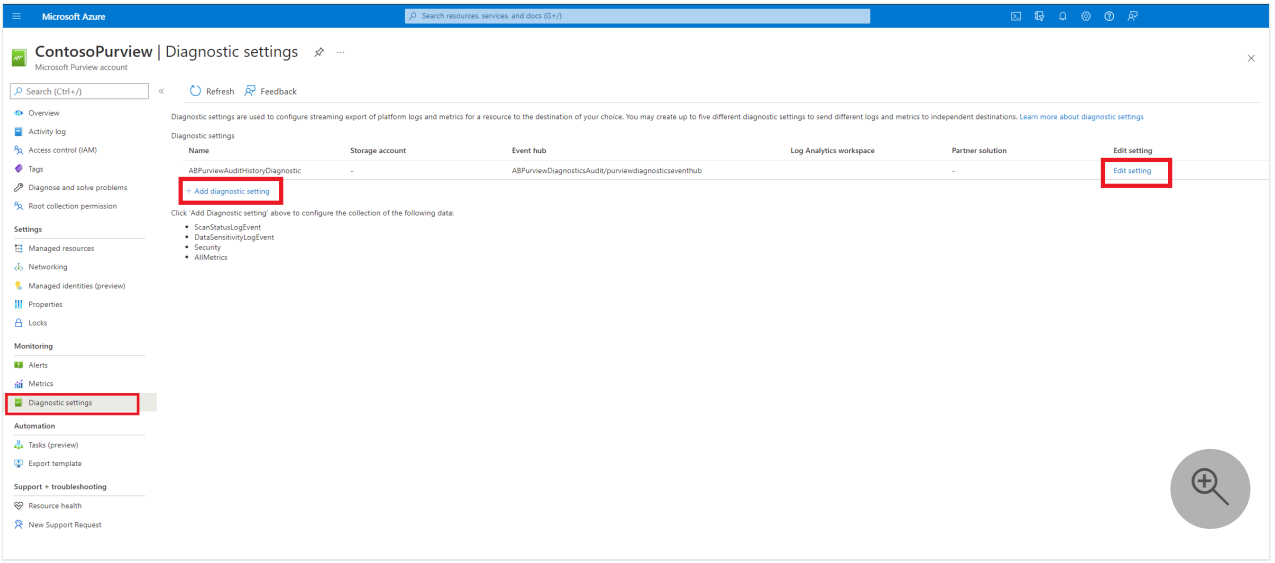
Now that Event Hubs is deployed and created, connect your Microsoft Purview account diagnostics audit logging to Event Hubs.

1. Go to your Microsoft Purview account home page. This page is where the overview information is displayed in the Azure portal. It's not the Microsoft Purview governance portal home page.

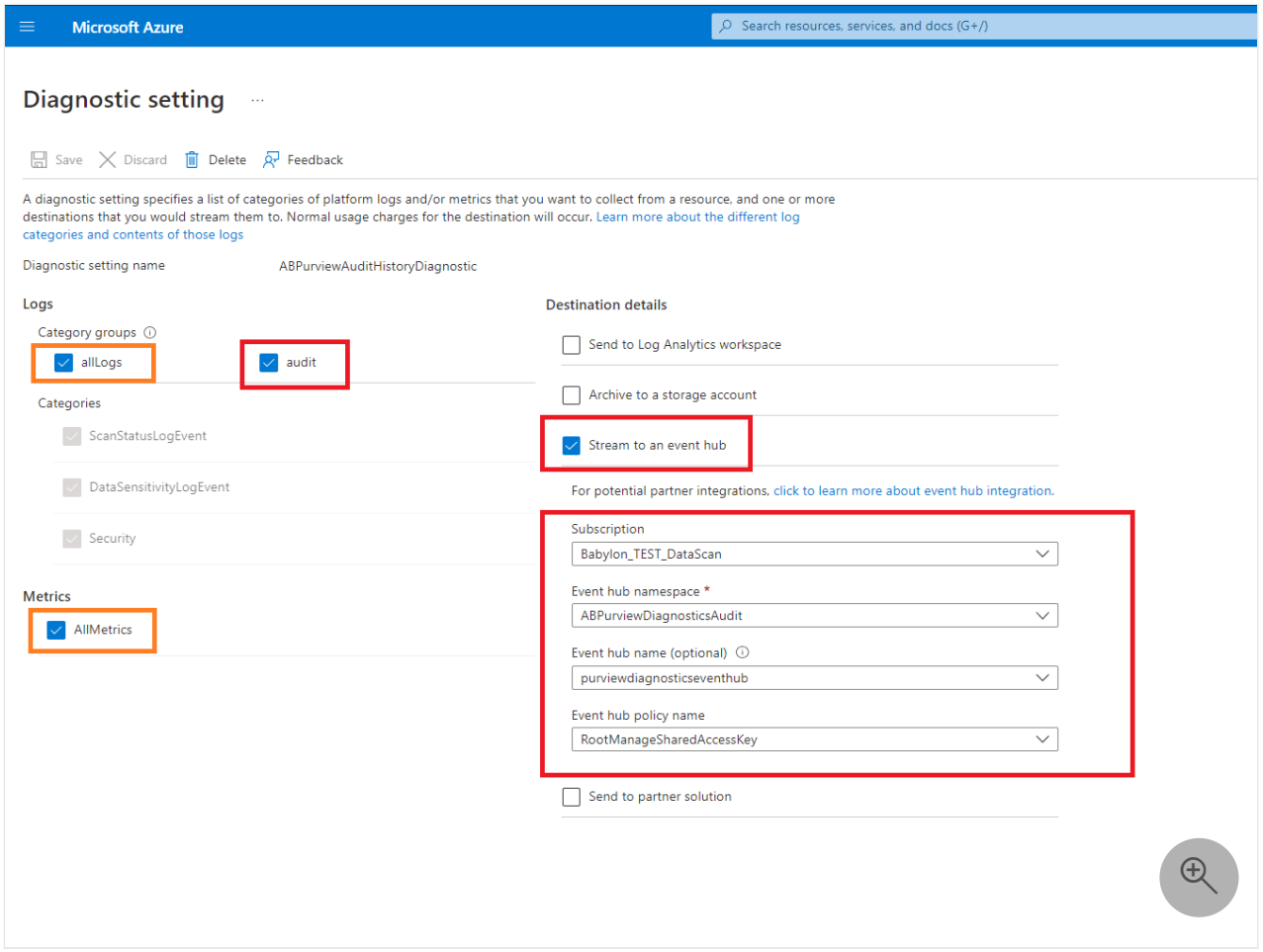
2. On the left menu, select **Monitoring** > **Diagnostic settings**.



3. Select **Add diagnostic setting** or **Edit setting**. Adding more than one diagnostic setting row in the context of Microsoft Purview isn't recommended. In other words, if you already have a diagnostic setting row, don't select **Add diagnostic**. Select **Edit** instead.



4. Select the **audit** and **allLogs** checkboxes to enable collection of audit logs. Optionally, select **AllMetrics** if you also want to capture Data Map capacity units and Data Map size metrics of the account.



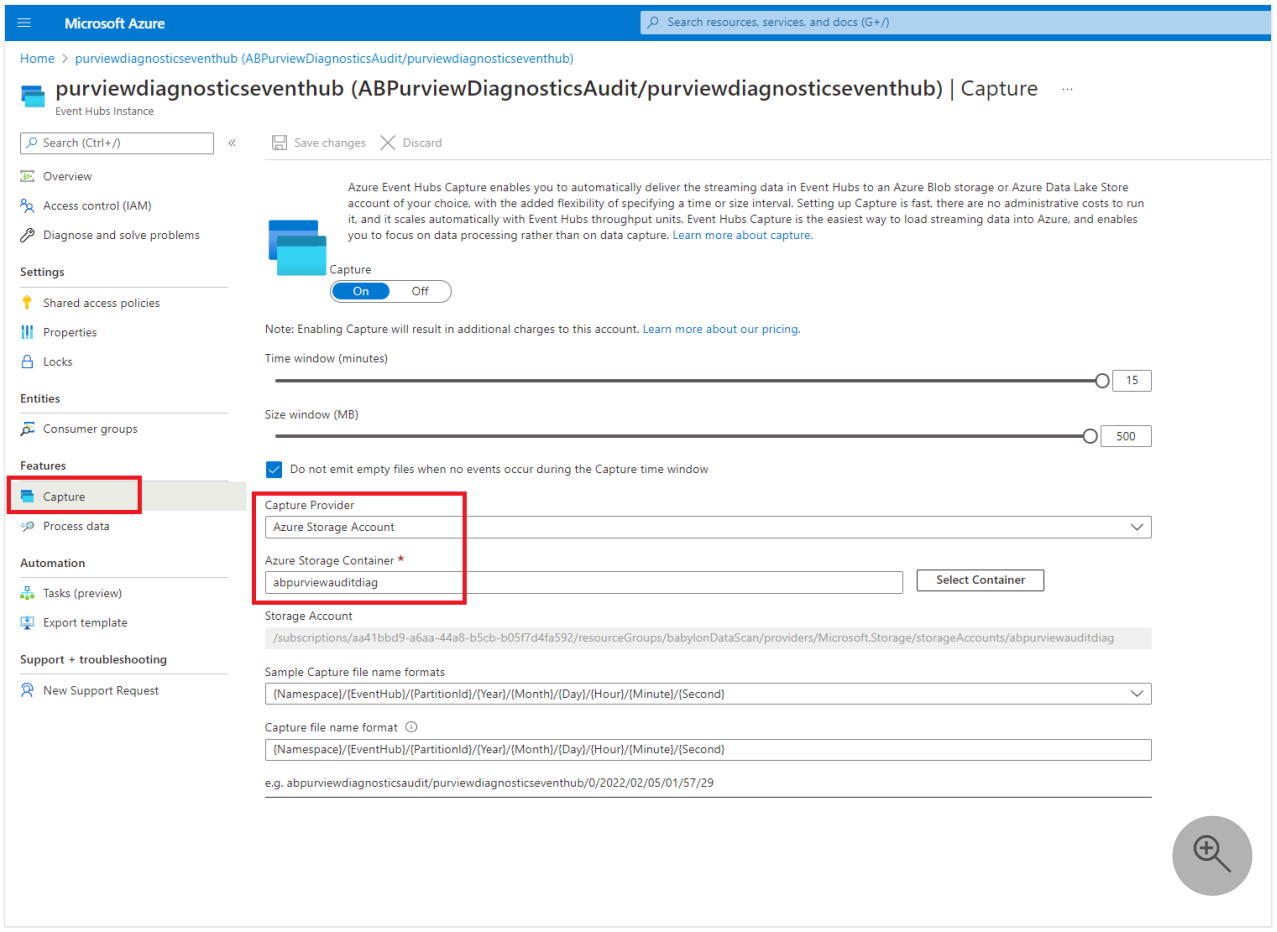
Diagnostics configuration on the Microsoft Purview account is complete.

Now that diagnostics audit logging configuration is complete, configure the data capture and data retention settings for Event Hubs.

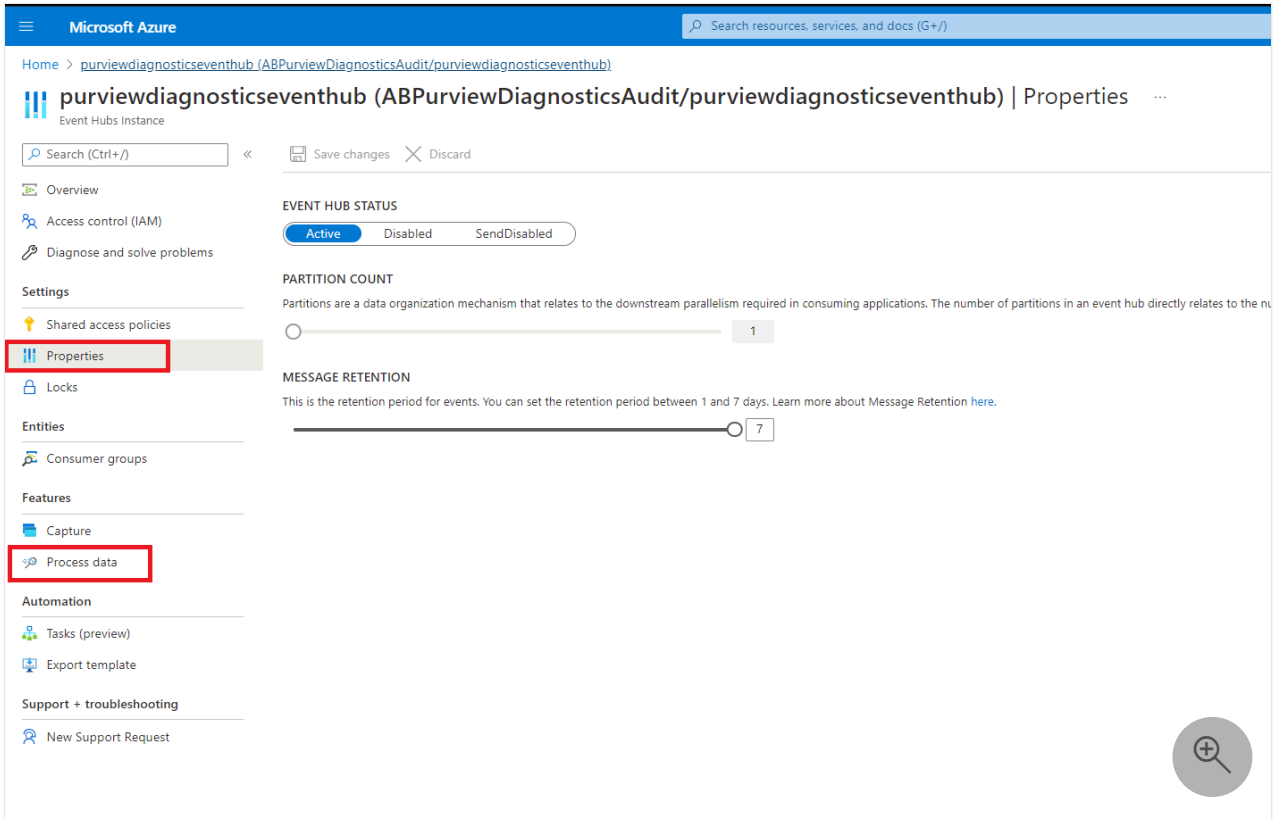
1. Go to the Azure portal home page, and search for the name of the Event Hubs namespace you created earlier.

2. Go to the Event Hubs namespace. Select **Event Hubs** > **Capture Data**.

3. Supply the name of the Event Hubs namespace and the event hub where you want the audit and diagnostics to be captured and streamed. Modify the **Time Window** and **Size Window** values for the retention period of the streaming events. Select **Save**.



4. Optionally, on the left menu, go to **Properties** and change **Message Retention** to any value between one and seven days. The retention period value depends on the frequency of scheduled jobs or the scripts you've created to continuously listen and capture the streaming events. If you schedule a capture once every week, move the slider to seven days.

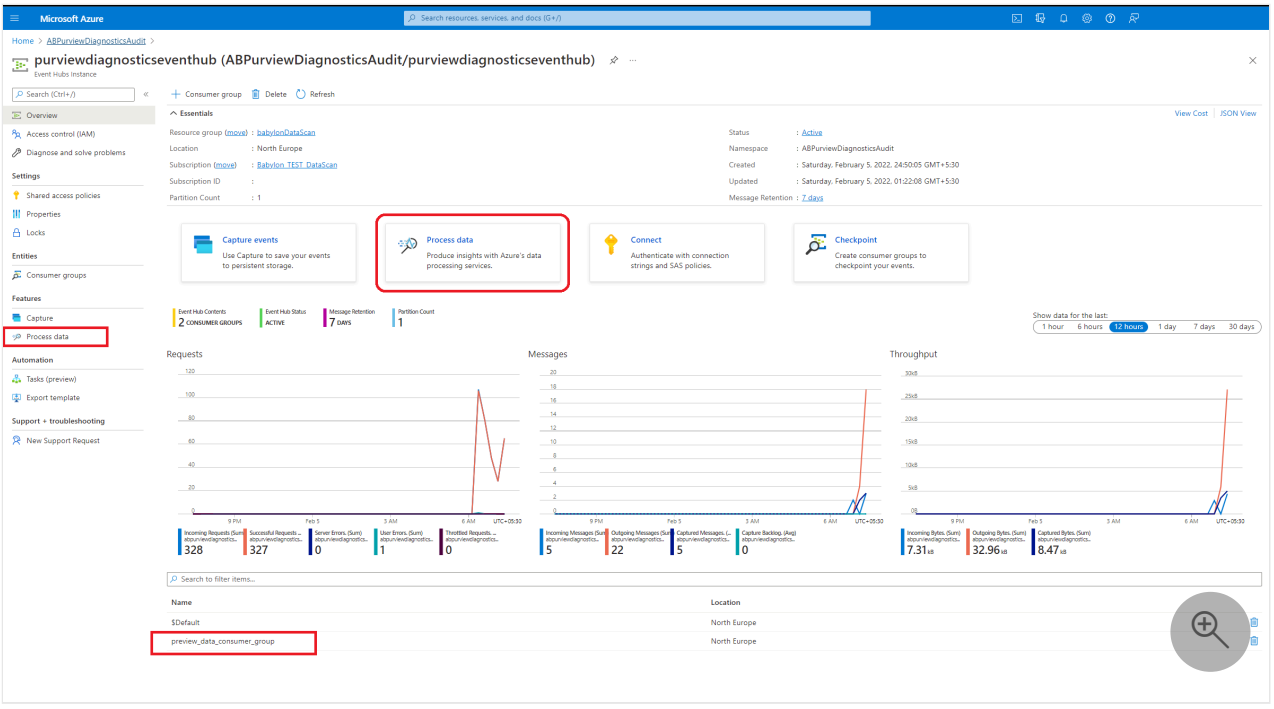


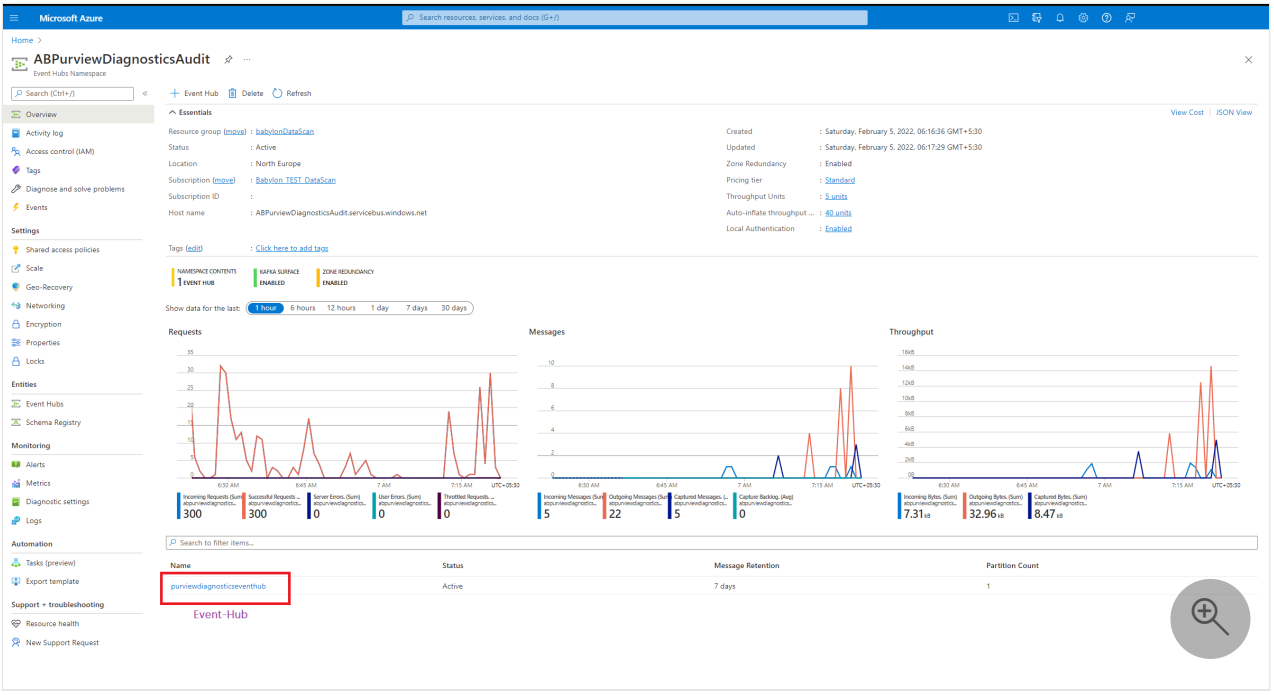
5. At this stage, the Event Hubs configuration is complete. The Microsoft Purview governance portal will start streaming all its audit history and diagnostics data to this event hub. You can now proceed to read, extract, and perform further analytics and operations on the captured diagnostics and audit events.

Read captured audit events

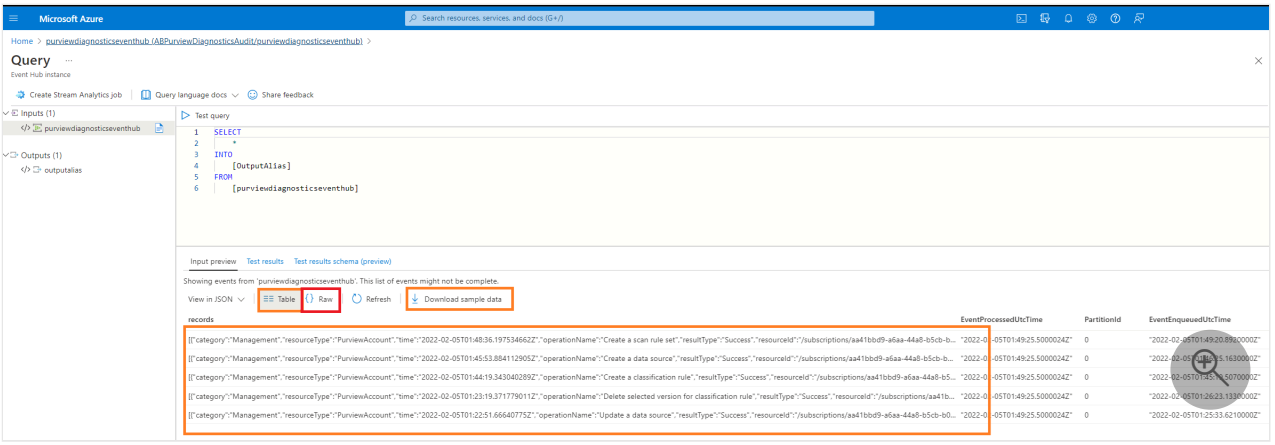
To analyze the captured audit and diagnostics log data:

1. Go to **Process data** on the Event Hubs page to see a preview of the captured audit logs and diagnostics.

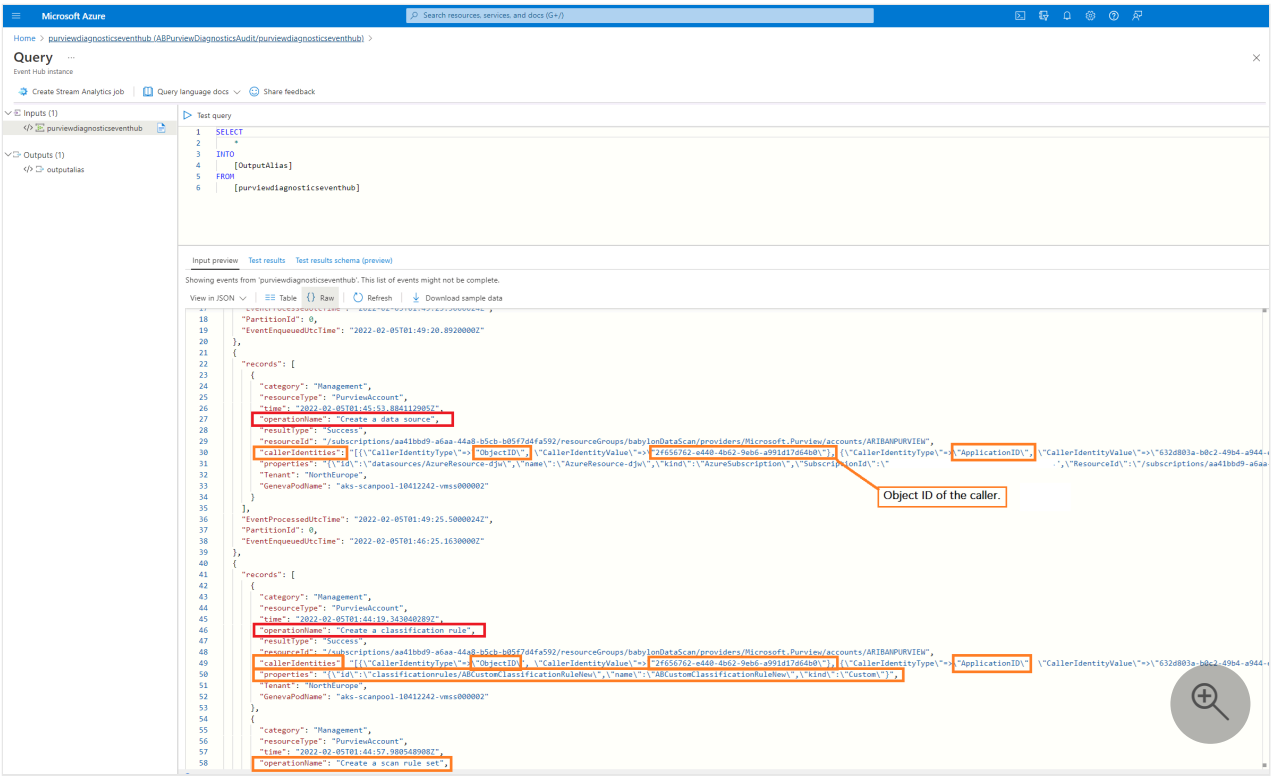




2. Switch between the **Table** and **Raw** views of the JSON output.



3. Select **Download sample data** and analyze the results carefully.



Now that you know how to gather this information, you can use automatic, scheduled scripts to extract, read, and perform further analytics on the Event Hubs audit and diagnostics data. You can even build your own utilities and custom code to extract business value from captured audit events.

These audit logs can also be transformed to Excel, any database, Dataverse, or Synapse Analytics database for analytics and reporting by using Power BI.

While you're free to use any programming or scripting language of your choice to read the event hubs, here's a ready-made Python-based script. See this Python tutorial on how to capture Event Hubs data in Azure Storage and read it by using Python (azure-eventhub).

Audit event categories

Some of the important categories of Microsoft Purview governance portal audit events that are currently available for capture and analysis are listed in the table.

More types and categories of activity audit events will be added.

Expand table

Category	Activity	Operation
Management	Collections	Create
Management	Collections	Update
Management	Collections	Delete
Management	Role assignments	Create
Management	Role assignments	Update
Management	Role assignments	Delete
Management	Scan rule set	Create
Management	Scan rule set	Update
Management	Scan rule set	Delete
Management	Classification rule	Create
Management	Classification rule	Update
Management	Classification rule	Delete
Management	Scan	Create
Management	Scan	Update
Management	Scan	Delete
Management	Scan	Run
Management	Scan	Cancel
Management	Scan	Create
Management	Scan	Schedule
Management	Data source	Register
Management	Data source	Update

Category	Activity	Operation
Management	Data source	Delete

Create an Azure policy exclusion for Microsoft Purview

Article • 05/14/2025

Many subscriptions have Azure Policies in place that restrict the creation or update of some resources. This is to maintain subscription security and cleanliness. However, Microsoft Purview accounts created before December 15, 2023 (or deployed using API version previous to 2023-05-01-preview) deployed an Azure Storage account when it was created. It's managed by Azure, so you don't need to maintain it, but it's necessary for Microsoft Purview to run correctly. Existing policies could block Microsoft Purview from updating this Azure Storage account, which can cause errors during scanning.

Overall, your approach to resolving this error is going to be dependent on your organizations needs and policies, but here are a few ways you could update your policies to resolve this issue.

To maintain your policies in your subscription, but still allow the creation and updates to these managed resources, you can create an exclusion. For example, adding an exclusion for the resource group where your Microsoft Purview account is being deployed.

For steps to create an exclusion for your policies, see these steps to add an exclusion to your policy.

Use resource selectors on a created policy

Depending on other policies deployed in your subscription, or depending on your region, you could need to add **Resource selectors** under the **Advanced** tab when assigning a policy. For example, you might need to add a resource selector for **resourceLocation** set to the region where you'll deploy your Microsoft Purview account. For more information about these conditions, see our documentation on location conditions.

Create an exemption

Exemptions are recommended for time-bound or specific scenarios where a resource or resource hierarchy should still be tracked and would otherwise be evaluated, but there's a specific reason it shouldn't be assessed for compliance. It would allow a temporary exception to your policy for a resource to exist for a finite amount of time.

For more information, see the documentation for policy exemptions.

Next steps

To set up Microsoft Purview by using Private Link, see Use private endpoints for your Microsoft Purview account.

Conditional Access with Microsoft Purview

Article • 05/20/2025

Microsoft Purview governance solutions support Microsoft Conditional Access.

The following steps show how to configure Microsoft Purview to enforce a Conditional Access policy.

Prerequisites

- When multi-factor authentication is enabled, to sign in to the Microsoft Purview governance portal, you must perform multi-factor authentication.

Configure conditional access

1. Sign in to the Azure portal, select **Microsoft Entra ID**.
2. Browse to **Security** > **Conditional Access**. For more information, see Microsoft Entra Conditional Access technical reference.
3. Select **Create new policy**, provide a name, and then select **Configure rules**.
4. Under **Assignments**, select **Users and groups**, check **Select users and groups**, and then select the user or group for Conditional Access. Select **Select**, and then select **Done** to accept your selection.
5. Select **Cloud apps**, select **Select apps**. You see all apps available for Conditional Access. Select **Azure Purview**, at the bottom select **Select**, and then select **Done**.
6. Select **Access controls**, select **Grant**, and then check the policy you want to apply.
7. Set **Enable policy** to **On** and select **Create**.

Next steps

- Use the Microsoft Purview governance portal
- More information about conditional access

Microsoft Purview metrics in Azure Monitor

Article • 05/20/2025

This article describes how to configure metrics, alerts, and diagnostic settings for Microsoft Purview using Azure Monitor.

Monitor Microsoft Purview

Microsoft Purview admins can use Azure Monitor to track the operational state of Microsoft Purview account. Metrics are collected to provide data points for you to track potential problems, troubleshoot, and improve the reliability of the Microsoft Purview account. The metrics are sent to Azure monitor for events occurring in Microsoft Purview.

Aggregated metrics

The metrics can be accessed from the Azure portal of a Microsoft Purview account. Access to the metrics is controlled by the role assignment of Microsoft Purview account. Users need to be part of the "Monitoring Reader" role in Microsoft Purview to see the metrics. Learn more about Monitoring role permissions.

The person who created the Microsoft Purview account automatically gets permissions to view metrics. If anyone else wants to see metrics, add them to the **Monitoring Reader** role, by following these steps:

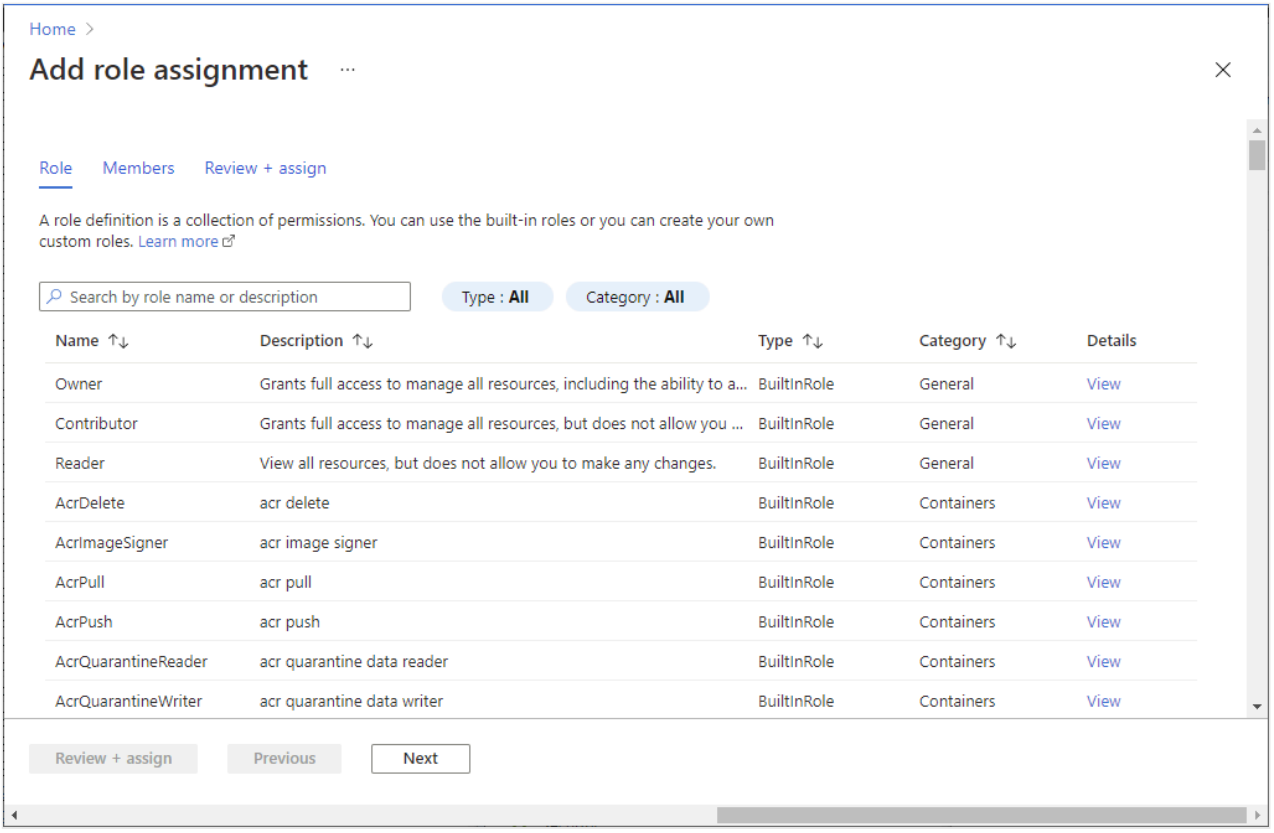
Add a user to the Monitoring Reader role

To add a user to the **Monitoring Reader** role, the owner of Microsoft Purview account or the Subscription owner can follow these steps:

1. Go to the Azure portal and search for the Microsoft Purview account name.
2. Select **Access control (IAM)**.
3. Select **Add** > **Add role assignment** to open the **Add role assignment** page.
4. Assign the role listed below. Get details on assigning Azure roles.

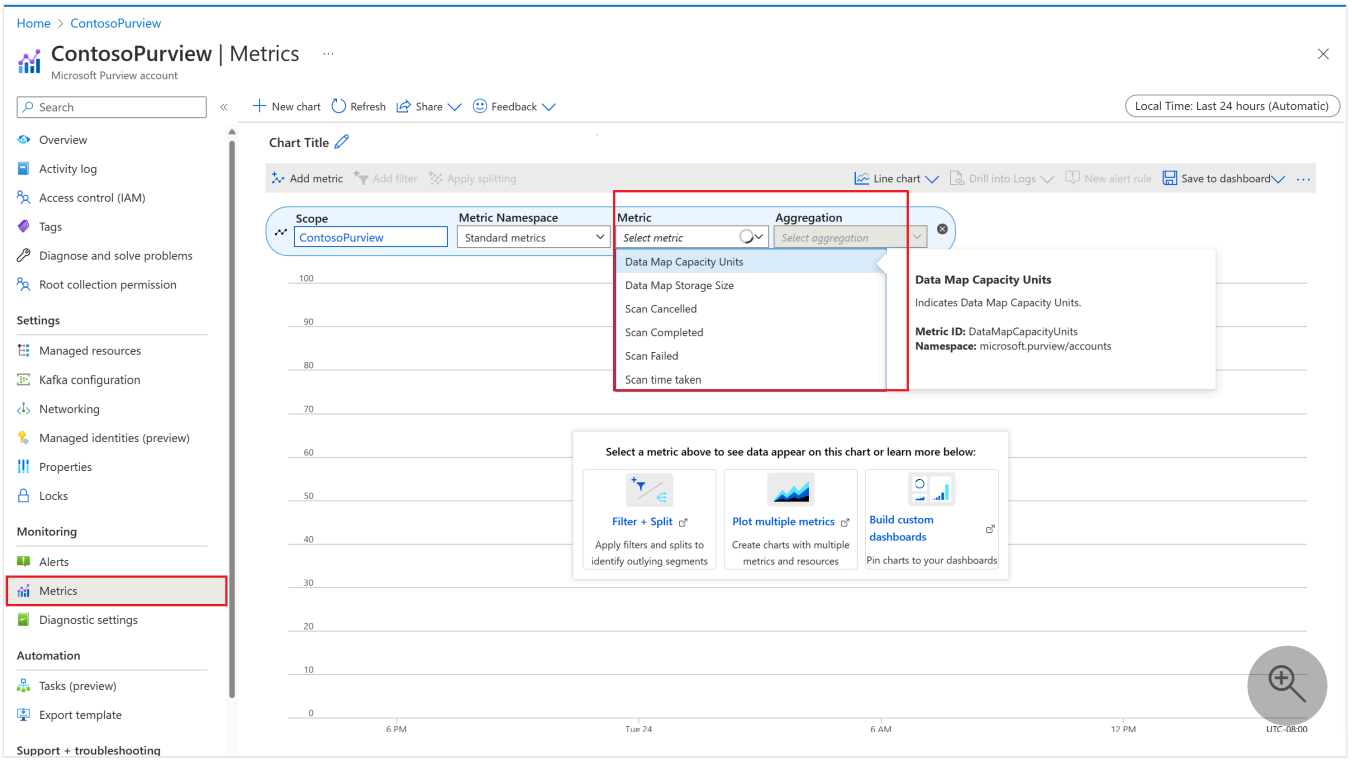
Expand table

Setting	Value
Role	Monitoring Reader
Assign access to	User, group, or service principal
Members	<Microsoft Entra account user>

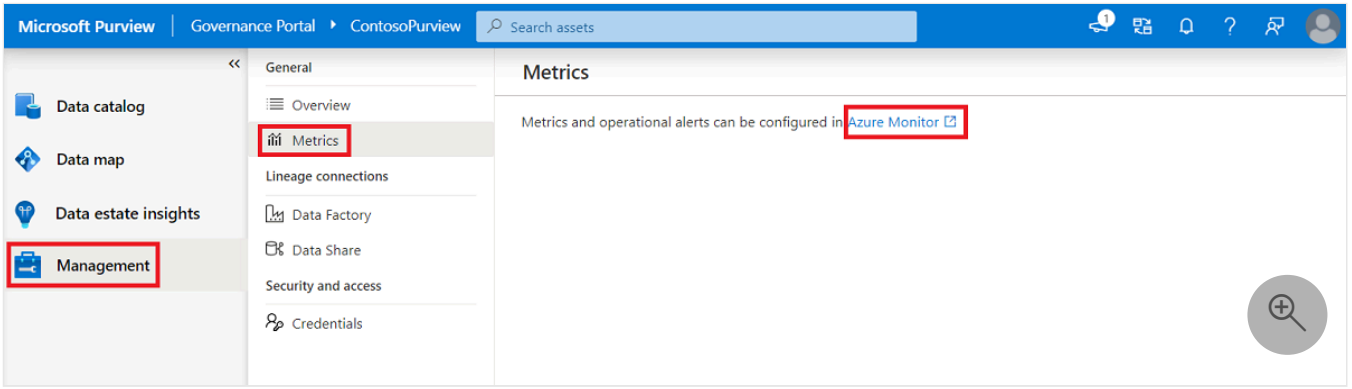


Metrics visualization

Users in the **Monitoring Reader** role can see the aggregated metrics and diagnostic logs sent to Azure Monitor. The metrics are listed in the Azure portal for the corresponding Microsoft Purview account. In the Azure portal, select the Metrics section to see the list of all available metrics.



Microsoft Purview users can also access the metrics page directly from the management center of the Microsoft Purview account. Select Azure Monitor in the main page of Microsoft Purview management center to launch Azure portal.



Available metrics

To get familiarized with how to use the metric section in the Azure portal pre read the following two documents. Getting started with Metric Explorer and Advanced features of Metric Explorer.

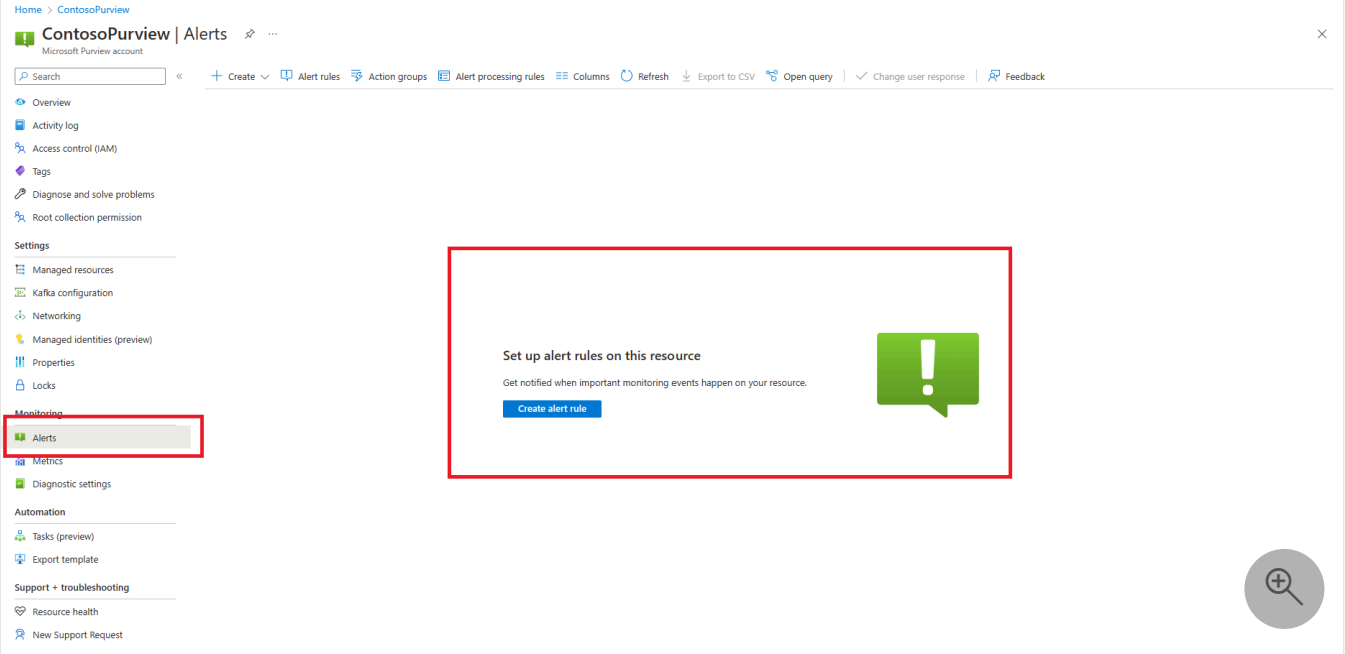
The following table contains the list of metrics available to explore in the Azure portal:

Expand table

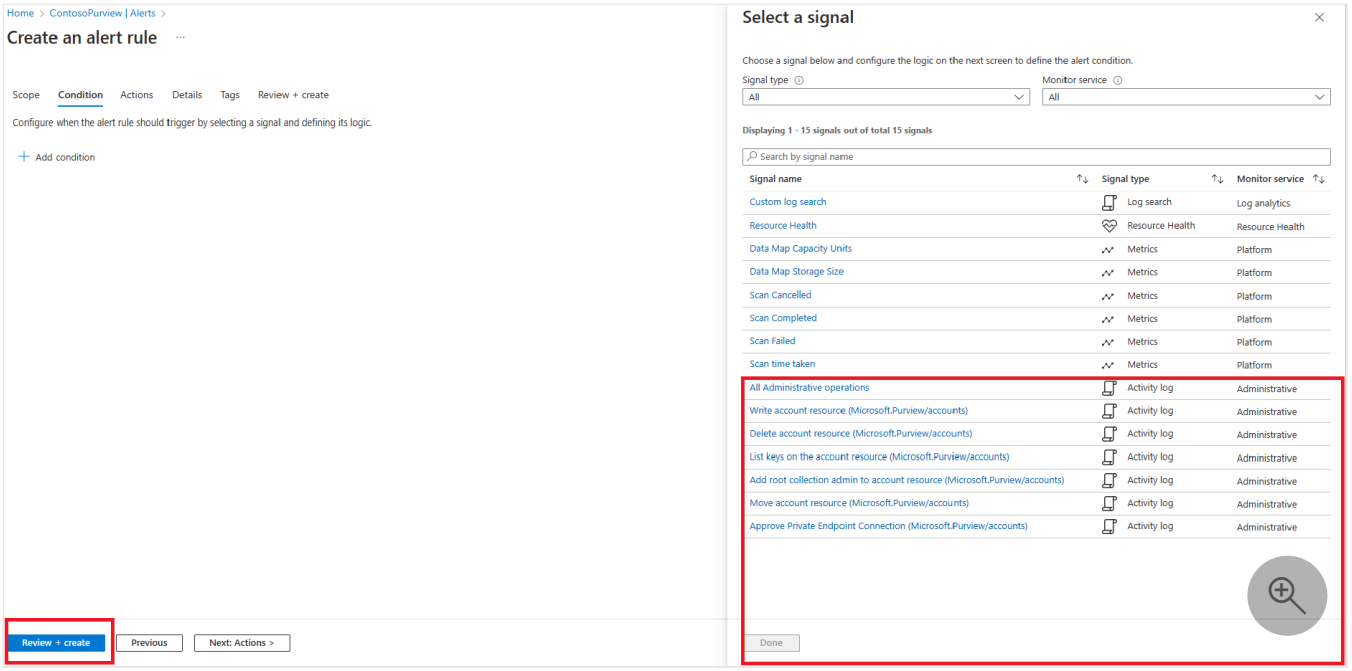
Metric Name	Metric Namespace	Aggregation type	Description
Data Map Capacity Units	Elastic data map	Sum Count	Aggregate the elastic data map capacity units over time period
Data Map Storage Size	Elastic data map	Sum Avg	Aggregate the elastic data map storage size over time period
Scan Canceled	Automated scan	Sum Count	Aggregate the canceled data source scans over time period
Scan Completed	Automated scan	Sum Count	Aggregate the completed data source scans over time period
Scan Failed	Automated scan	Sum Count	Aggregate the failed data source scans over time period
Scan time taken	Automated scan	Min Max Sum Avg	Aggregate the total time taken by scans over time period

Monitoring alerts

Alerts can be accessed from the Azure portal of a Microsoft Purview account. Access to the alerts is controlled by the role assignment of Microsoft Purview account just like metrics. A user can setup alert rules in their purview account to get notified when important monitoring events happen.



The user can also create specific alert rules and conditions for signals within purview.

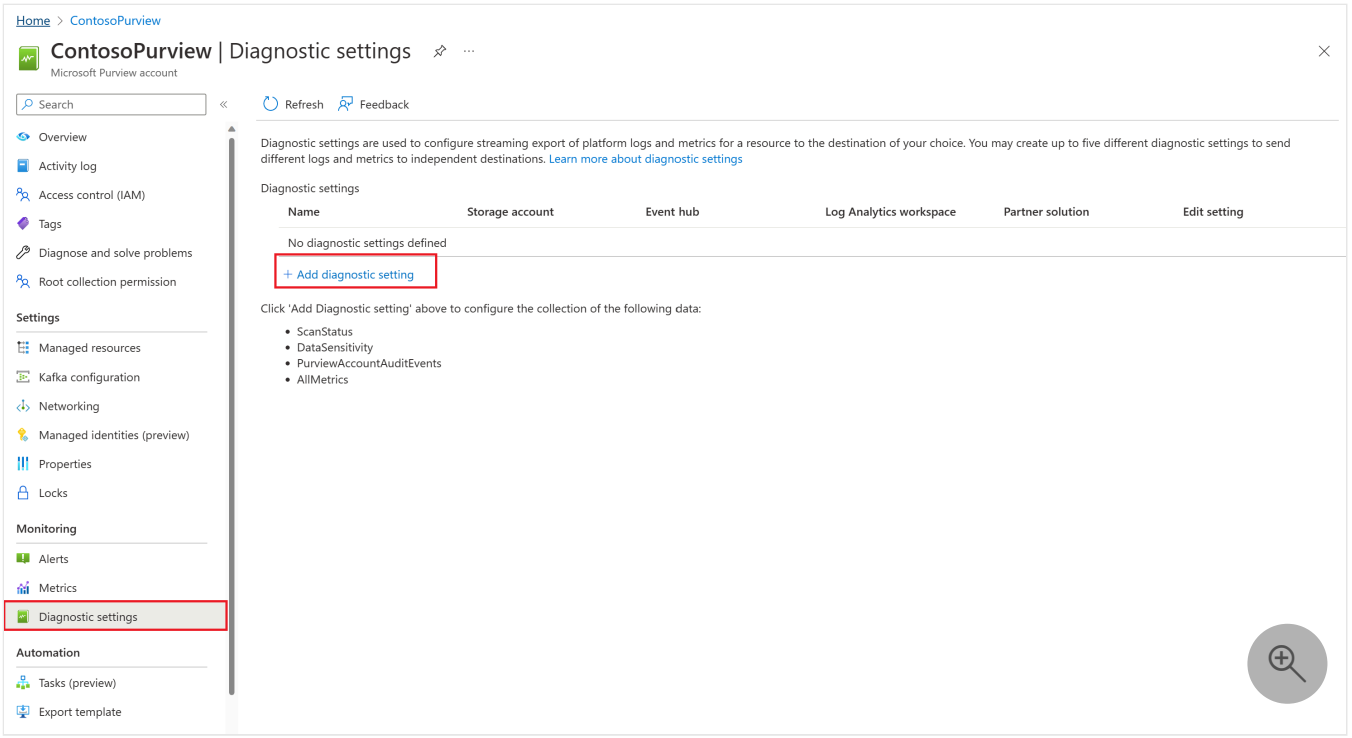


Sending diagnostic logs

Raw telemetry events are sent to Azure Monitor. Events can be sent to a Log Analytics Workspace, archived to a customer storage account of choice, streamed to an event hub, or sent to a partner solution for further analysis. Exporting of logs is done via the Diagnostic settings for the Microsoft Purview account on the Azure portal.

Follow these steps to create a diagnostic setting for your Microsoft Purview account and send to your preferred destination:

1. Locate your Microsoft Purview account in the Azure portal.
2. In the menu under **Monitoring** select **Diagnostic settings**.
3. Select **Add diagnostic setting** to create a new diagnostic setting to collect platform logs and metrics. For more information about these settings and logs, see the Azure Monitor documentation..

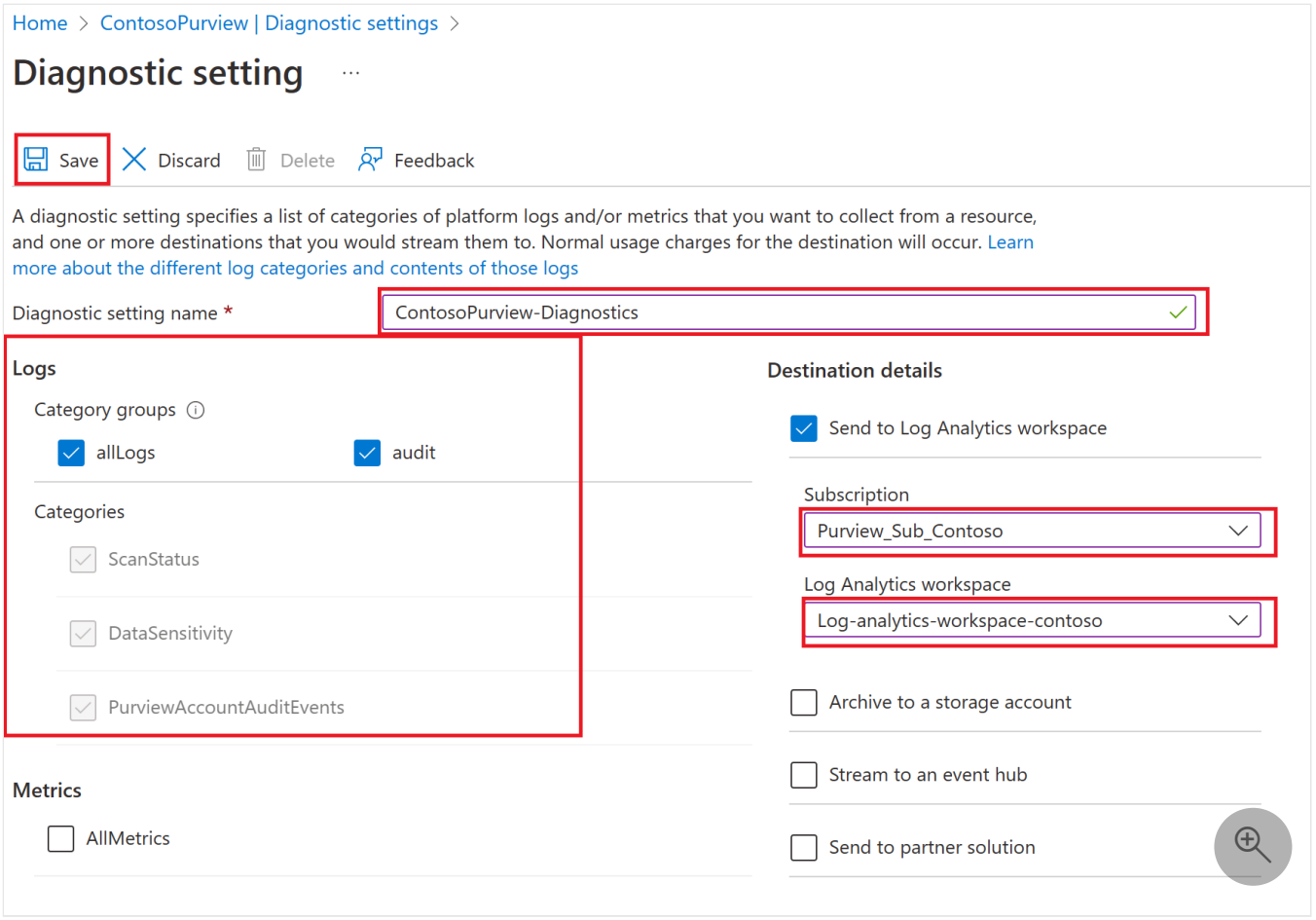


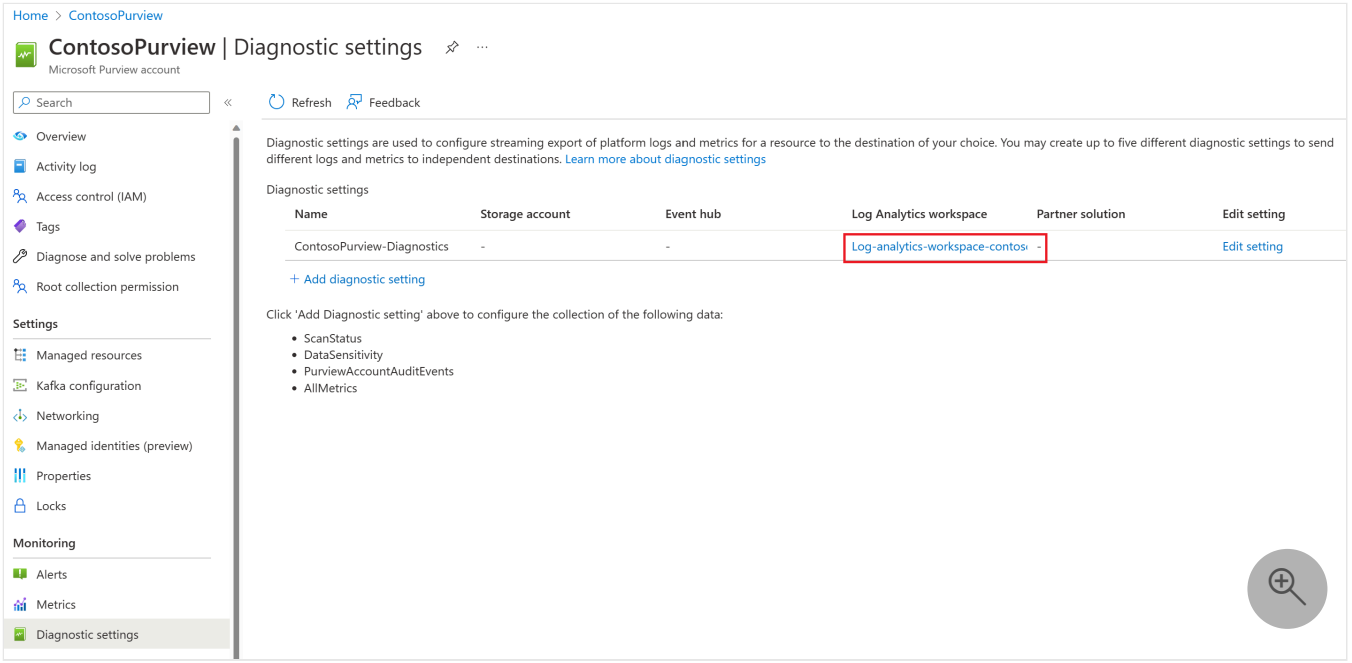
4. You can send your logs to:

- A log analytics workspace
- A storage account
- An event hub

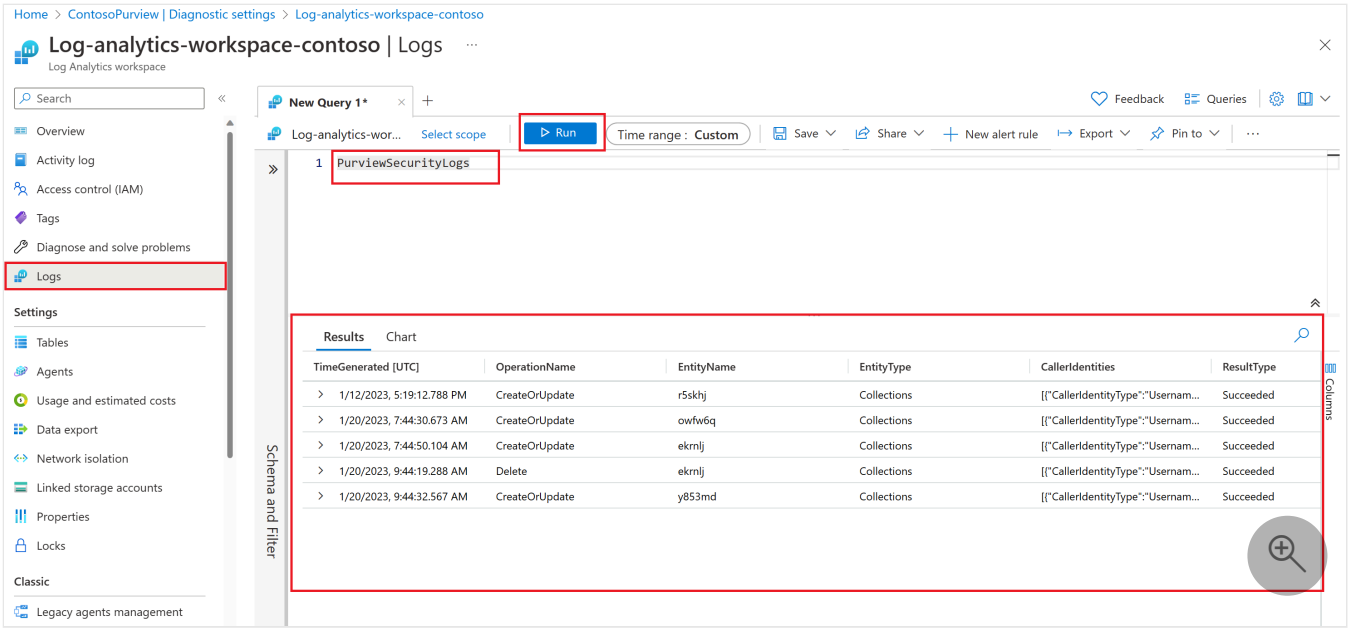
Destination - Log Analytics Workspace

1. In the **Destination details**, select **Send to Log Analytics workspace**.
2. Create a name for the diagnostic setting, select the applicable log category group and select the right subscription and workspace, then select save. The workspace doesn't have to be in the same region as the resource being monitored. You to create a new workspace, you can follow this article: Create a New Log Analytics Workspace.



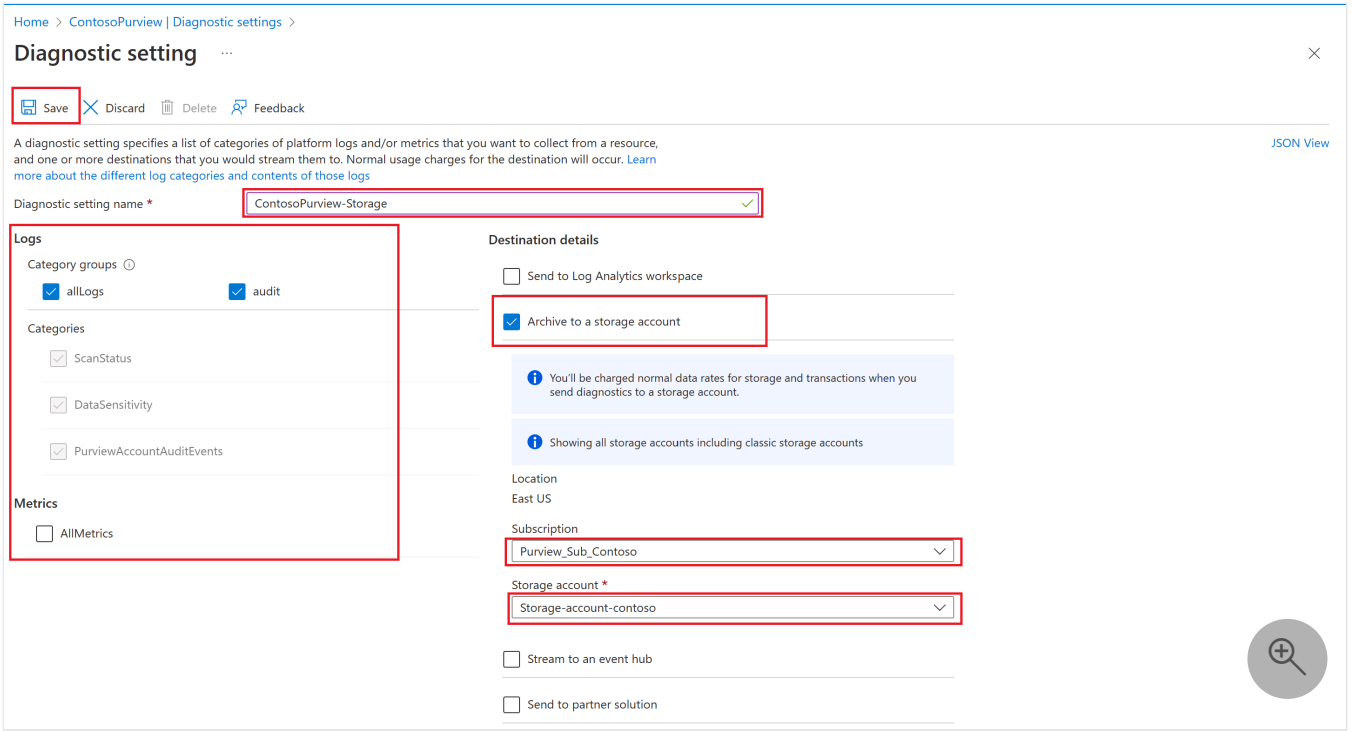


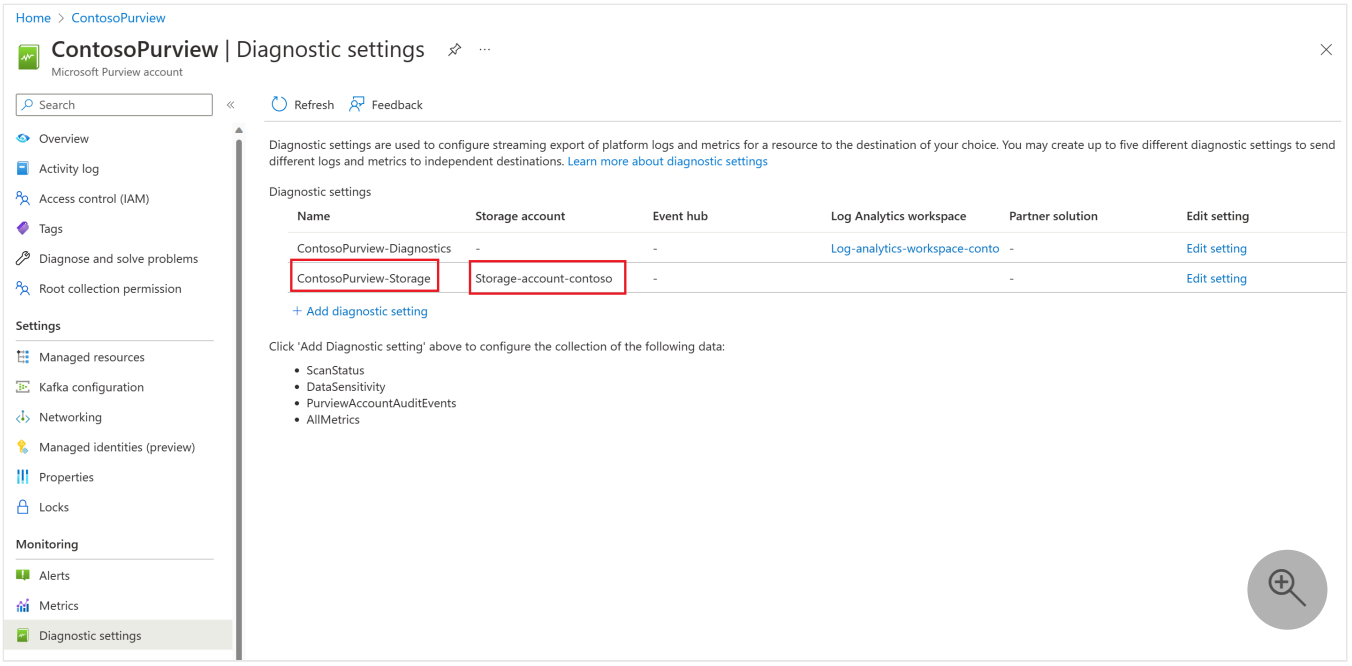
3. Verify the changes in your Log Analytics Workspace by performing some operations to populate data. For example, creating/updating/deleting a policy. After which you can open the **Log Analytics Workspace**, navigate to **Logs**, enter query filter as "**purviewsecuritylogs**", then select "**Run**" to execute the query.



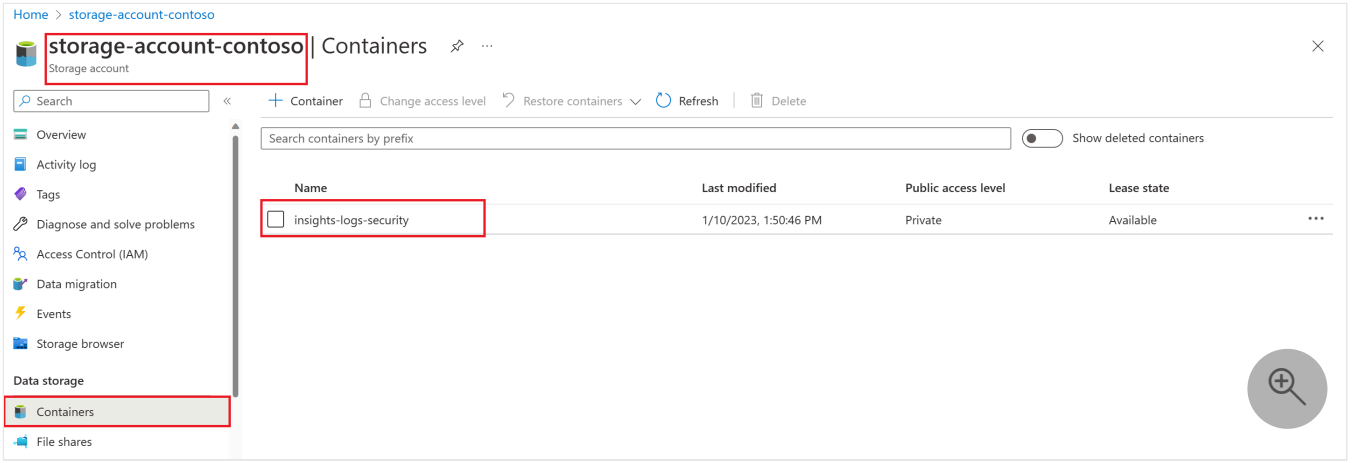
Destination - Storage account

1. In the **Destination details**, select **Archive to a storage account**.
2. Create a diagnostic setting name, select the log category, select the destination as archive to a storage account, select the right subscription and storage account then select save. A dedicated storage account is recommended for archiving the diagnostic logs. If you need a storage account, you can follow this article: Create a storage account.

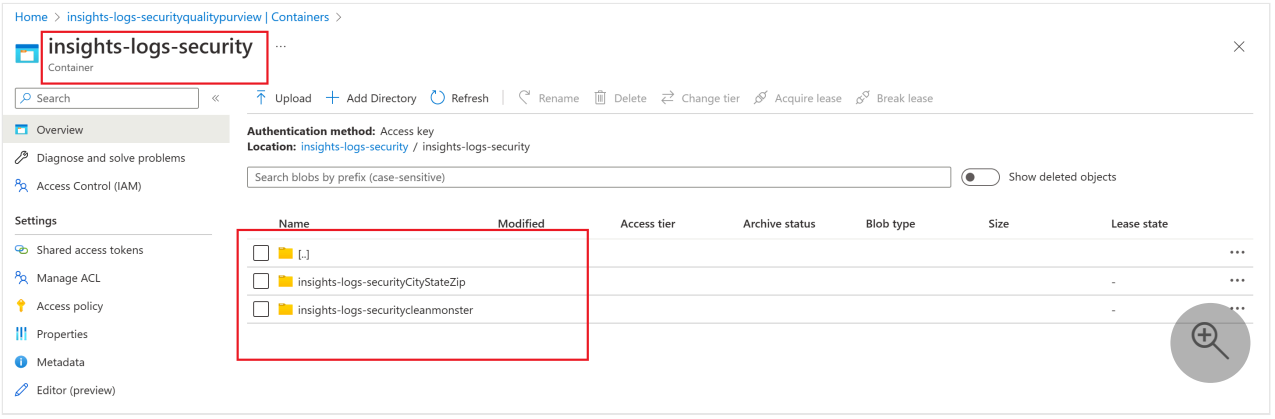


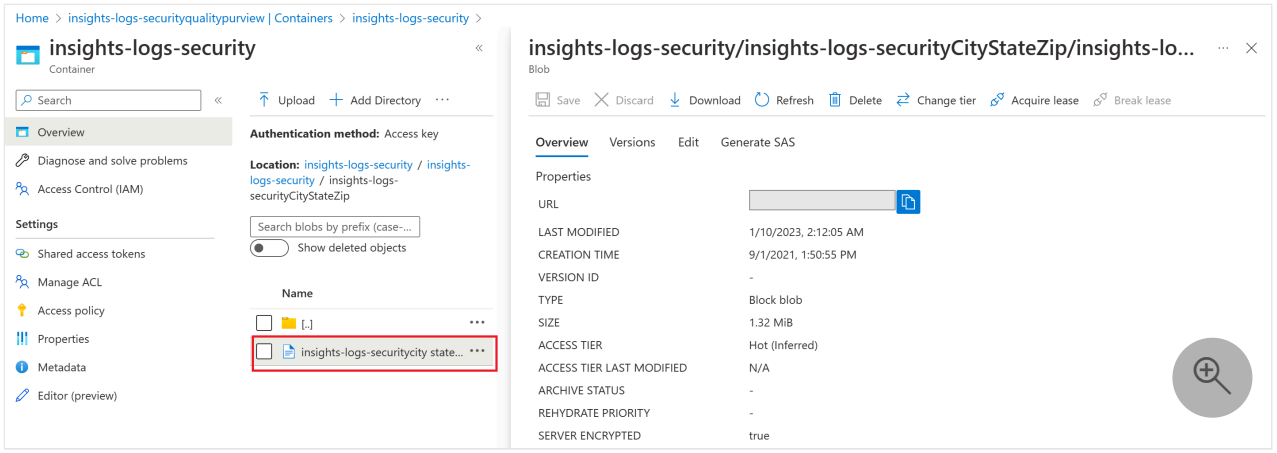


3. To see logs in the **Storage Account**, perform a sample action (for example: create/update/delete a policy), then open the **Storage Account**, navigate to **Containers**, and select the container name.



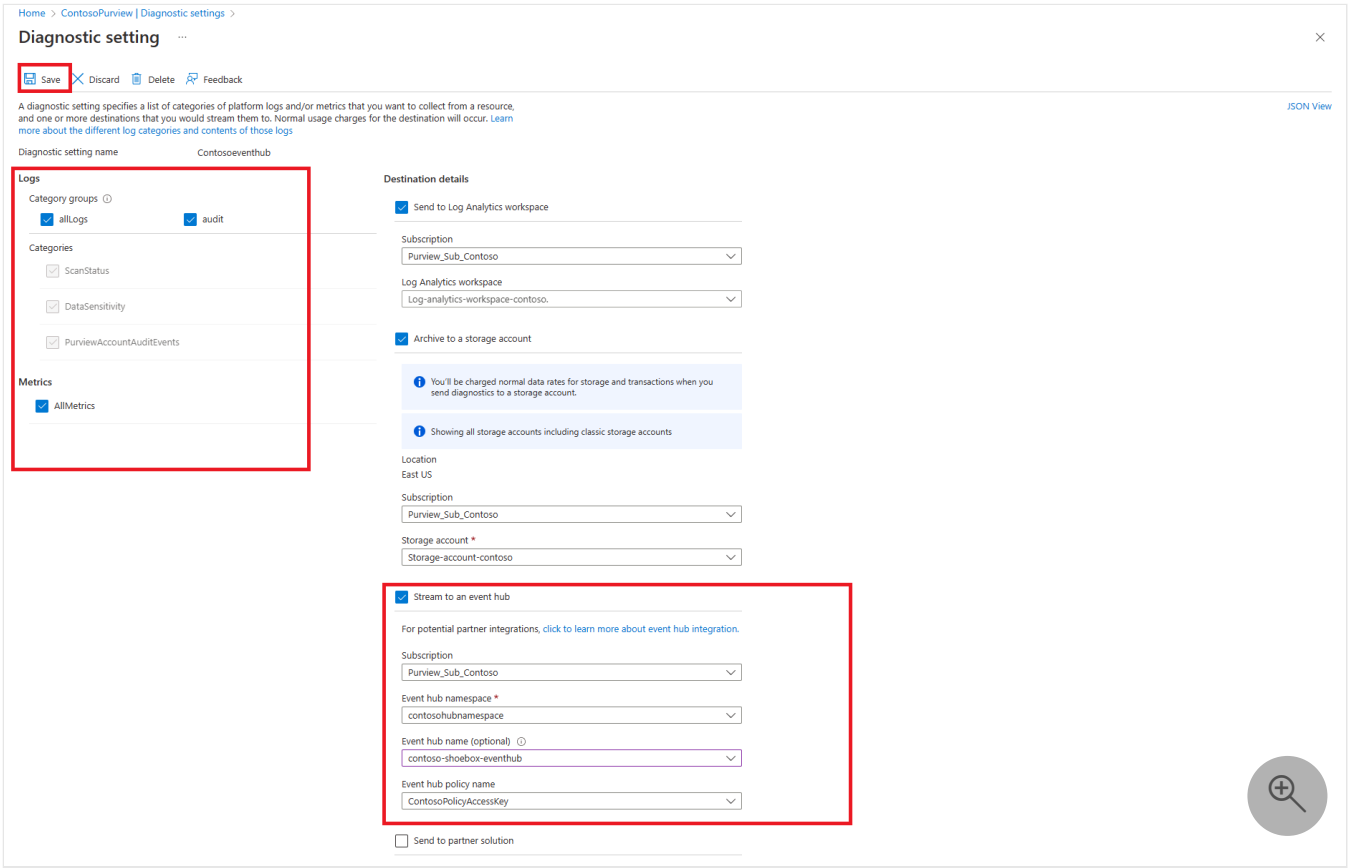
4. Navigate to the file and download it to see the logs.

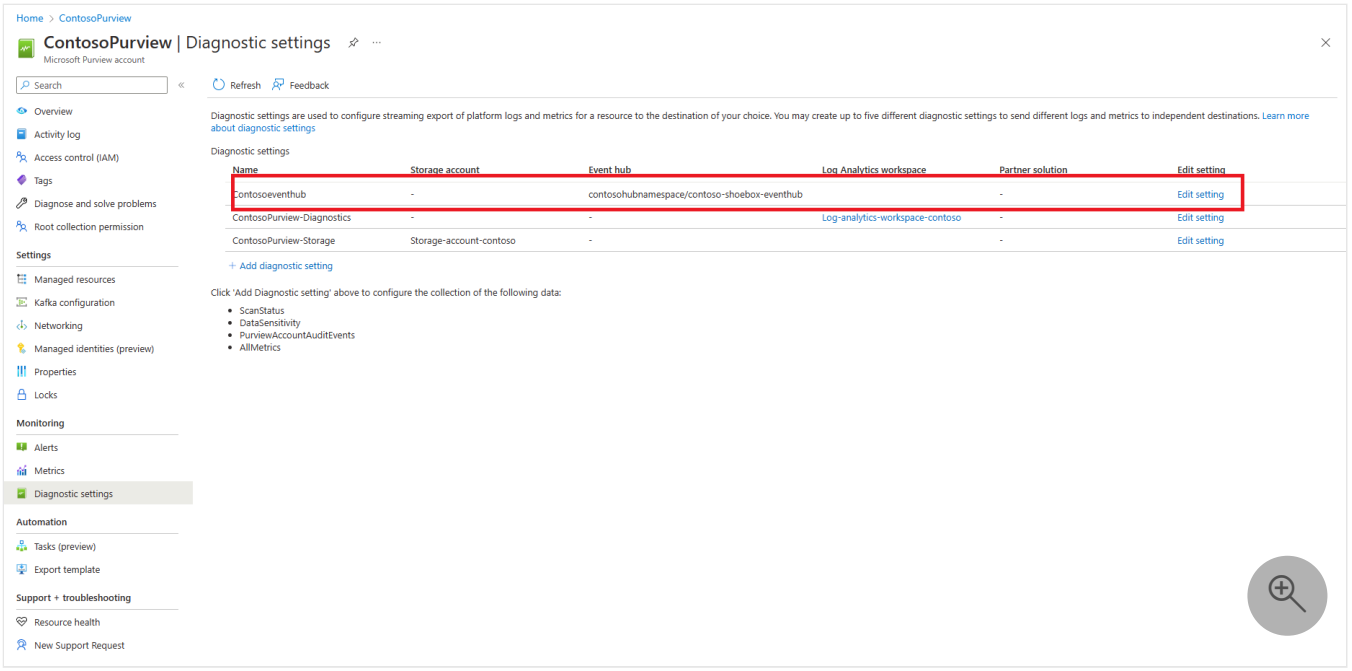




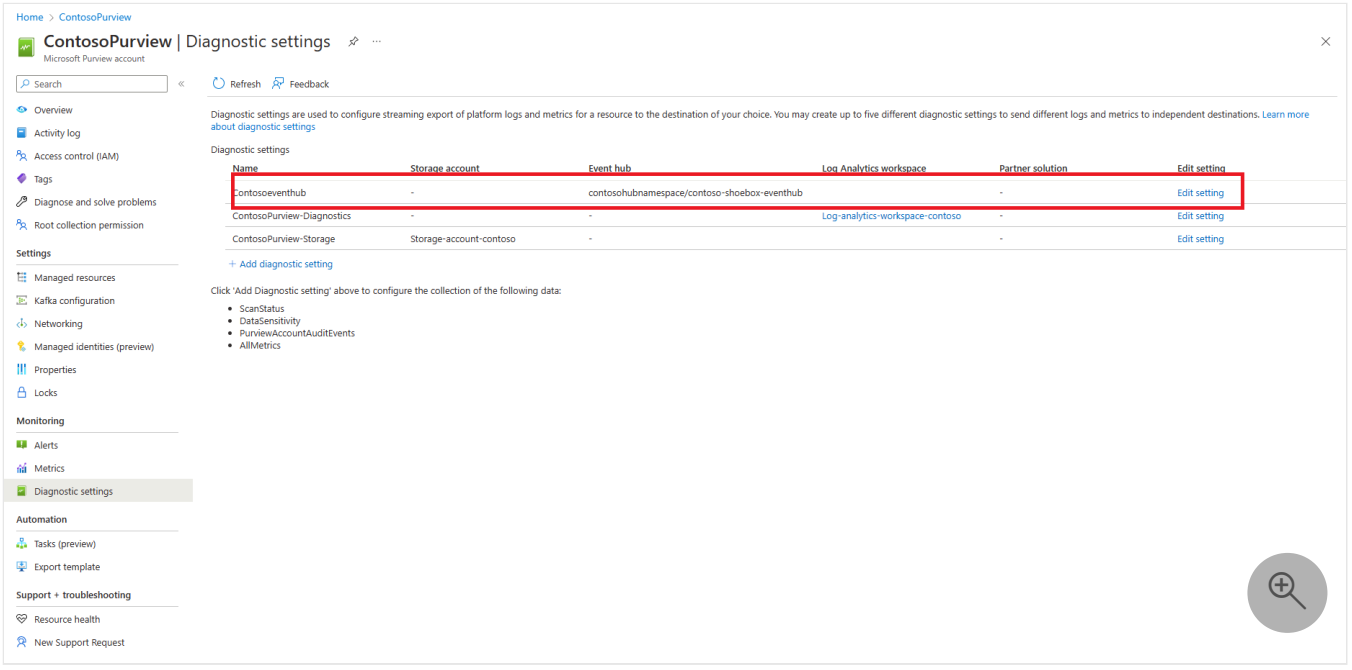
Destination - Event hub

1. In the **Destination details**, select **Stream to an event hub**.
2. Create a diagnostic setting name, select the log category, select the destination as stream to event hub, select the right subscription, event hubs namespace, event hub name and event hub policy name then select save. An event hub name space is required before you can stream to an event hub. If you need to create an event hub namespace, you can follow this article: Create an event hub & event hubs namespace storage account





3. To see logs in the **Event Hubs Namespace**, Go to the Azure portal, and search for the name of the event hubs namespace you created earlier, go the Event Hubs Namespace and click on overview. To find our more about capturing and reading captured audit events in the event hubs namespace, you can follow this article: Audit Logs & diagnostics



Sample Log

Here's a sample log you'd receive from a diagnostic setting.

The event tracks the scan life cycle. A scan operation follows progress through a sequence of states, from Queued, Running and finally a terminal state of Succeeded | Failed | Canceled. An event is logged for each state transition and the schema of the event will have the following properties.

JSON

```
{
  "time": "<The UTC time when the event occurred>",
  "properties": {
    "dataSourceName": "<Registered data source friendly name>",
    "dataSourceType": "<Registered data source type>",
    "scanName": "<Scan instance friendly name>",
    "assetsDiscovered": "<If the resultType is succeeded, count of assets
discovered in scan run>",
    "assetsClassified": "<If the resultType is succeeded, count of assets
classified in scan run>",
    "scanQueueTimeInSeconds": "<If the resultType is succeeded, total seconds the
scan instance in queue>",
    "scanTotalRunTimeInSeconds": "<If the resultType is succeeded, total seconds
the scan took to run>",
    "runType": "<How the scan is triggered>",
    "errorDetails": "<Scan failure error>",
    "scanResultId": "<Unique GUID for the scan instance>"
  },
  "resourceId": "<The azure resource identifier>",
  "category": "<The diagnostic log category>",
  "operationName": "<The operation that cause the event Possible values for
ScanStatusLogEvent category are:
                    |AdhocScanRun
                    |TriggeredScanRun
                    |StatusChangeNotification>",
  "resultType": "Queued - indicates a scan is queued.
                 Running - indicates a scan entered a running state.
                 Succeeded - indicates a scan completed successfully.
                 Failed - indicates a scan failure event.
                 Cancelled - indicates a scan was cancelled. ",
  "resultSignature": "<Not used for ScanStatusLogEvent category. >",
  "resultDescription": "<This will have an error message if the resultType is
Failed. >",
  "durationMs": "<Not used for ScanStatusLogEvent category. >",
  "level": "<The log severity level. Possible values are:
            |Informational
            |Error >",
  "location": "<The location of the Microsoft Purview account>",
}
```

The Sample log for an event instance is shown in the below section.

JSON

```
{
  "time": "2020-11-24T20:25:13.022860553Z",
  "properties": {
    "dataSourceName": "AzureDataExplorer-swD",
    "dataSourceType": "AzureDataExplorer",
    "scanName": "Scan-Kzw-shoebox-test",
    "assetsDiscovered": "0",
```

```
    "assetsClassified": "0",
    "scanQueueTimeInSeconds": "0",
    "scanTotalRunTimeInSeconds": "0",
    "runType": "Manual",
    "errorDetails": "empty_value",
    "scanResultId": "0dc51a72-4156-40e3-8539-b5728394561f"
  },
  "resourceId": "/SUBSCRIPTIONS/111111111111-111-4EB2/RESOURCEGROUPS/FOOBAR-TEST-
RG/PROVIDERS/MICROSOFT.PURVIEW/ACCOUNTS/FOOBAR-HEY-TEST-NEW-MANIFEST-EUS",
  "category": "ScanStatusLogEvent",
  "operationName": "TriggeredScanRun",
  "resultType": "Delayed",
  "resultSignature": "empty_value",
  "resultDescription": "empty_value",
  "durationMs": 0,
  "level": "Informational",
  "location": "eastus",
}
```


Resources

Microsoft Purview Data Map

Use private endpoints for your Microsoft Purview account

Article • 05/09/2025

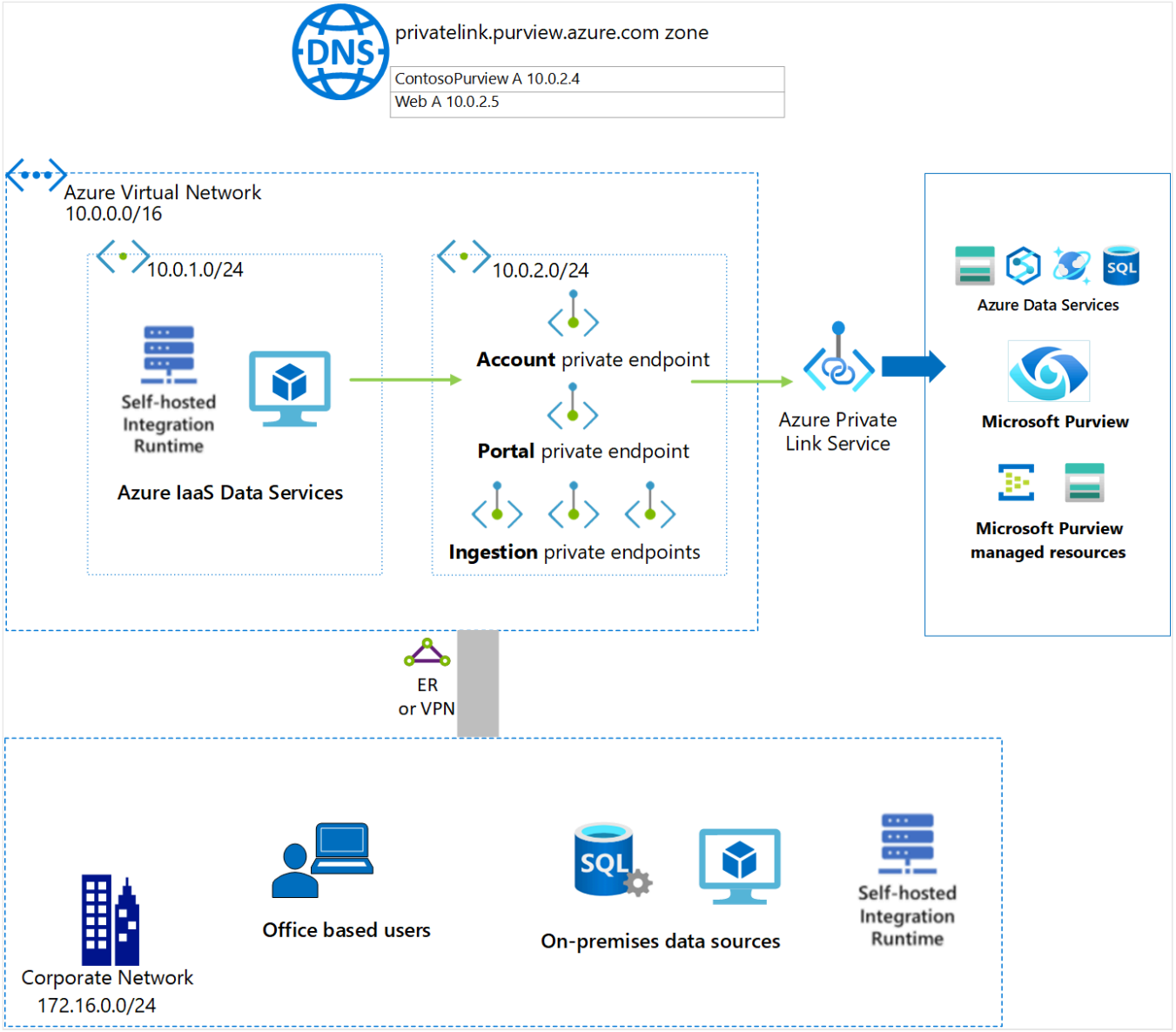
> **Important**
>
> This article covers private endpoints for the classic Microsoft Purview governance portal (https://web.purview.azure.com). If you're using the new Microsoft Purview portal (https://purview.microsoft.com/), follow the documentation for private endpoints in the Microsoft Purview portal.

Conceptual Overview

You can use Azure private endpoints for your Microsoft Purview accounts to allow users on a virtual network (VNet) to securely access the catalog over a Private Link. A private endpoint uses an IP address from the VNet address space for your Microsoft Purview account. Network traffic between the clients on the VNet and the Microsoft Purview account traverses over the VNet and a private link on the Microsoft backbone network.

If you're still using the classic portal experience, you can deploy Microsoft Purview *account* private endpoint, to allow only client calls to Microsoft Purview that originate from within the private network. To connect to the classic governance portal using a private network connectivity, you can deploy *portal* private endpoint.

For both the new and classic experience, you can deploy *ingestion* private endpoints if you need to scan Azure IaaS and PaaS data sources inside Azure virtual networks and on-premises data sources through a private connection. This method ensures network isolation for your metadata flowing from the data sources to Microsoft Purview Data Map.



Prerequisites

Before deploying private endpoints for Microsoft Purview account, ensure you meet the following prerequisites:

1. An Azure account with an active subscription. Create an account for free.
2. An existing Azure Virtual network. Deploy a new Azure virtual network if you don't have one.

Microsoft Purview private endpoint deployment scenarios

Use the following recommended checklist to perform deployment of Microsoft Purview account with private endpoints:

Expand table

Scenario	Objectives
Scenario 1 - Connect to your Microsoft Purview and scan data sources privately and securely	You need to restrict access to your Microsoft Purview account only via a private endpoint, including access to the classic governance portal, Atlas APIs and scan data sources in on-premises and Azure (but inside a virtual network) using self-hosted integration runtime ensuring end to end network isolation. (Deploy *account*, _portal, and *ingestion* private endpoints.)
Scenario 2 - Connect privately and securely to your Microsoft Purview account	You need to enable access to your Microsoft Purview account, including access to *the Microsoft Purview governance portal* and Atlas API through private endpoints. (Deploy *account* and *portal* private endpoints).
Scenario 3 - Scan data source securely using Managed Virtual Network	You need to scan Azure data sources securely, without having to manage a virtual network or a self-hosted integration runtime VM. (Deploy managed private endpoints for Microsoft Purview Azure data sources).
Scenario 4 - Using the new Microsoft Purview portal	If you're using the new portal experience, you can set up ingestion and platform private endpoints.

Frequently Asked Questions

For FAQs related to private endpoint deployments in Microsoft Purview, see FAQ about Microsoft Purview private endpoints.

Troubleshooting guide

For troubleshooting private endpoint configuration for Microsoft Purview accounts, see Troubleshooting private endpoint configuration for Microsoft Purview accounts.

Known limitations

To view list of current limitations related to Microsoft Purview private endpoints, see Microsoft Purview private endpoints known limitations.

Configure firewall settings for your Microsoft Purview account

Article • 05/09/2025

> **Important**
>
> This article covers private endpoints for the classic Microsoft Purview governance portal (https://web.purview.azure.com). If you're using the new Microsoft Purview portal (https://purview.microsoft.com/), follow the documentation for private endpoints in the Microsoft Purview portal.

Prerequisites

To configure Microsoft Purview account firewall settings, ensure you meet the following prerequisites:

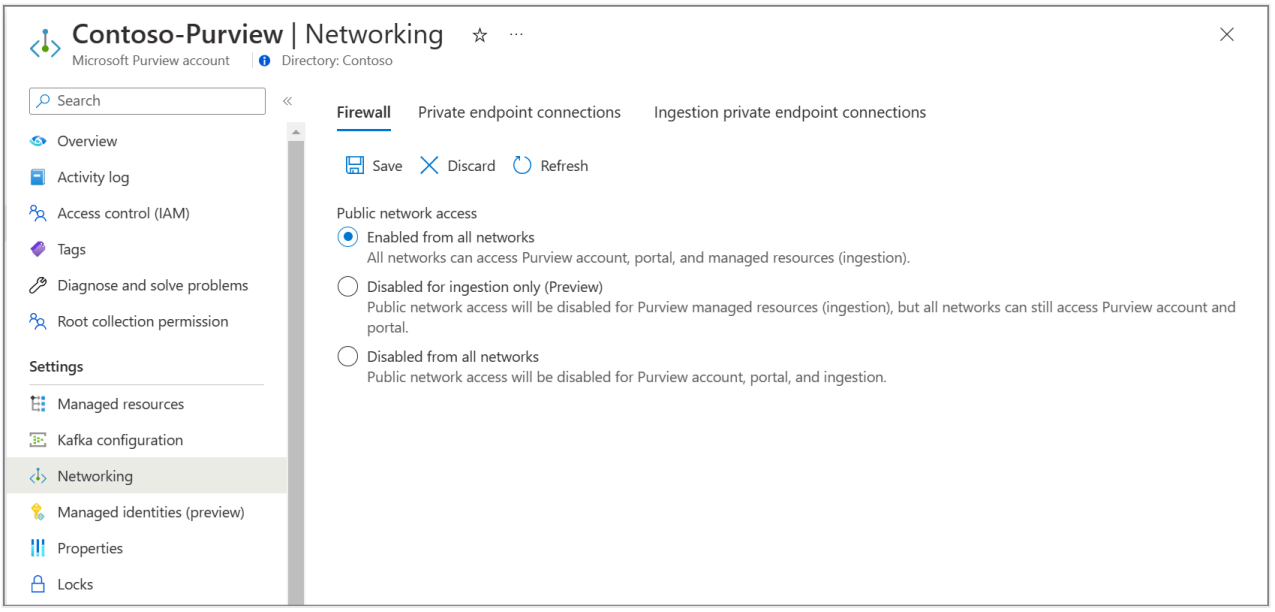
1. An Azure account with an active subscription. Create an account for free.
2. An existing Microsoft Purview account.
3. Must have IAM Contributor role on the Microsoft Purview account.
4. Must have IAM Reader role on the Azure subscription that hosts the Microsoft Purview account.

Microsoft Purview firewall deployment scenarios

To configure Microsoft Purview firewall, follow these steps:

1. Sign in to the Azure portal.
2. Navigate to your Microsoft Purview account in the portal.
3. Under *Settings*, choose **Networking**.
4. In the **Firewall** tab, under **Public network access**, change the firewall settings to the option that suits your scenario:

- **Enabled from all networks**

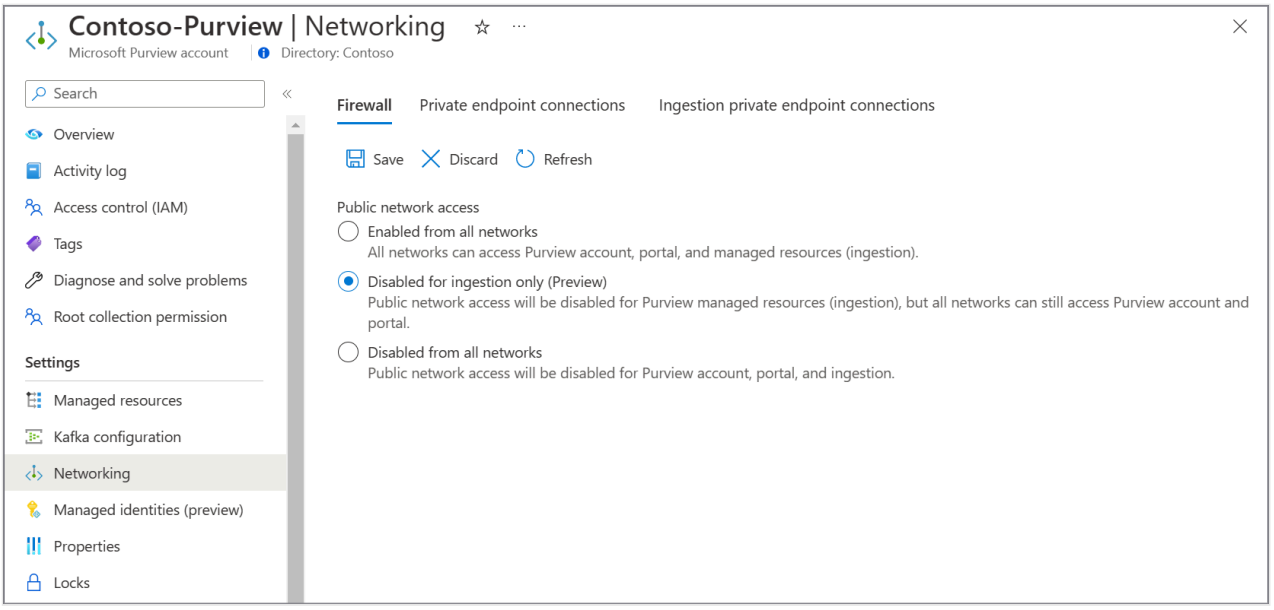


By choosing this option:

- All public network access into your Microsoft Purview account is allowed.
- If you're using a managed Event Hubs for kafka notifications, public network access is set to *All networks* on your Microsoft Purview account's Event Hubs.
- If you have an account created before December 15, 2023, public network access is set to *Enabled from all networks* on your Microsoft Purview account's managed storage account. If you have an account created after that date (or deployed using API version 2023-05-01-preview onwards), the public network access on your Microsoft Purview account's ingestion storage account *Enabled from all networks*.

> **Note**
>
> Even though the network access is enabled through public internet, to gain access to the classic governance portal, users must be first authenticated and authorized.

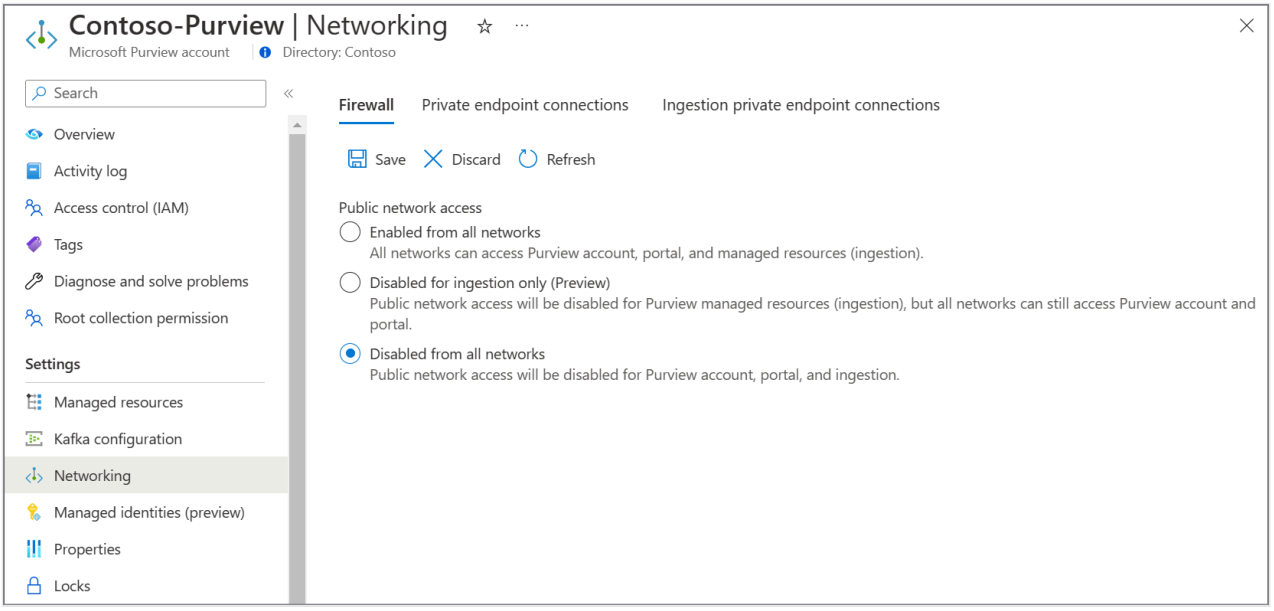
- **Disabled for ingestion only (preview)**



By choosing this option:

- Public network access to your Microsoft Purview account through API and the classic governance portal is allowed.
- All public network traffic for ingestion is disabled. In this case, you must configure a private endpoint for ingestion before setting up any scans. For more information, see Use private endpoints for your Microsoft Purview account.
- If you're using a managed Event Hubs for kafka notifications, public network access is set to *Disabled* on your Microsoft Purview account's Event Hubs.
- If you have an account created before December 15, 2023, public network access is set to *Disabled* on your Microsoft Purview account's managed storage account. If you have an account created after that date (or deployed using API version 2023-05-01-preview onwards), the public network access on your Microsoft Purview account's ingestion storage account *Disabled*.

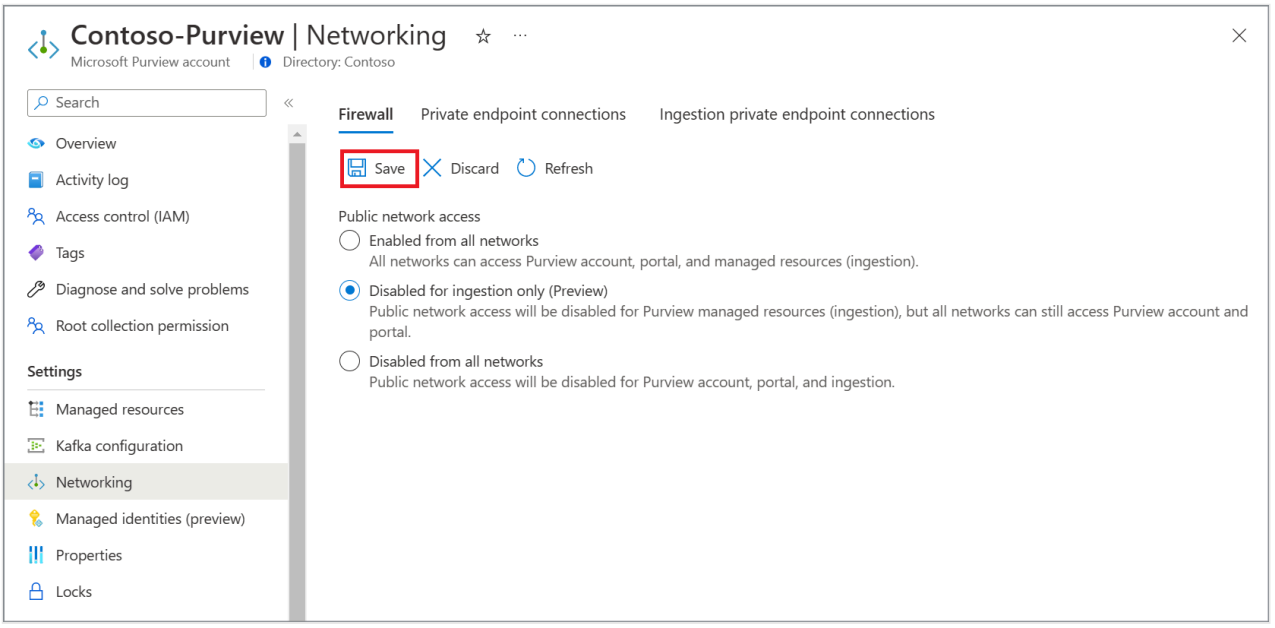
- **Disabled from all networks**



By choosing this option:

- All public network access into your Microsoft Purview account is disabled.
- All network access to your Microsoft Purview account through APIs or the classic governance portal including traffic to run scans is allowed only through private network using private endpoints. For more information, see Connect to your Microsoft Purview and scan data sources privately and securely.
- If you're using a managed Event Hubs for kafka notifications, public network access is set to *Disabled* on your Microsoft Purview account's Event Hubs.
- If you have an account created before December 15, 2023, public network access is set to *Disabled* on your Microsoft Purview account's managed storage account. If you have an account created after that date (or deployed using API version 2023-05-01-preview onwards), the public network access on your Microsoft Purview account's ingestion storage account is set to *Disabled*.

5. Select **Save**.



Connect to your Microsoft Purview and scan data sources privately and securely

Article • 05/09/2025

> **Important**
>
> This article covers private endpoints for the classic Microsoft Purview governance portal (https://web.purview.azure.com). If you're using the new Microsoft Purview portal (https://purview.microsoft.com/), follow the documentation for private endpoints in the Microsoft Purview portal.

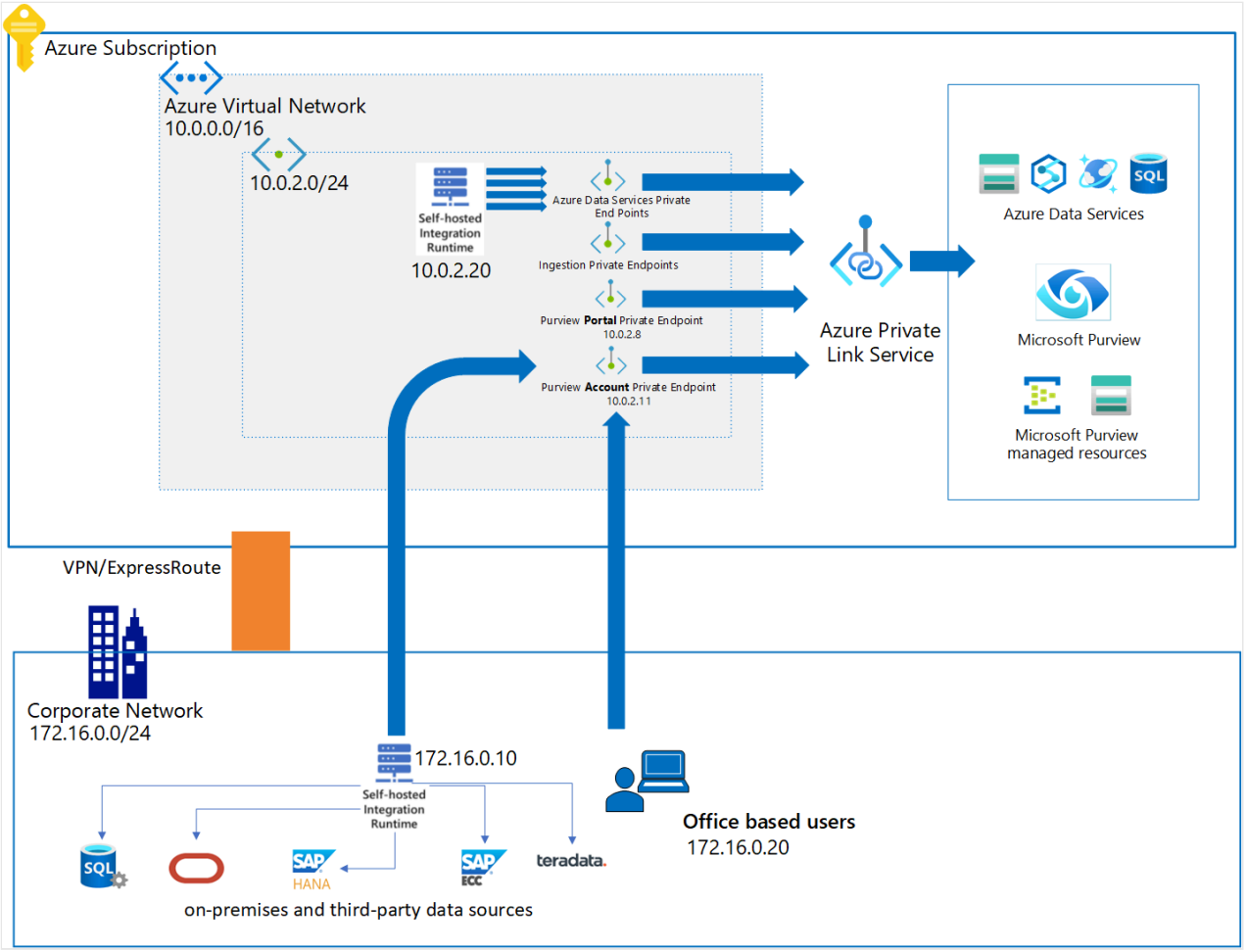
In this guide, you'll learn how to deploy *account*, _portal, and *ingestion* private endpoints for your Microsoft Purview account to access purview account and scan data sources using a self-hosted integration runtime securely and privately, enabling end-to-end network isolation.=

The Microsoft Purview *account* private endpoint is used to add another layer of security by enabling scenarios where only client calls that originate from within the virtual network are allowed to access the Microsoft Purview account. This private endpoint is also a prerequisite for the portal private endpoint.

The Microsoft Purview *portal* private endpoint is required to enable connectivity to the classic Microsoft Purview governance portal using a private network.

Microsoft Purview can scan data sources in Azure or an on-premises environment by using *ingestion* private endpoints. Depending on your accounts deployment date and configuration, up to three private endpoint resources are required to be deployed and linked to Microsoft Purview configured resources when ingestion private endpoint is deployed:

- If you're using a managed Event Hubs for kafka notifications, a namespace private endpoint is linked to a Microsoft Purview configured Event Hubs namespace.
- If your account was created before December 15, 2023:
  - Blob private endpoint is linked to a Microsoft Purview managed storage account.
  - Queue private endpoint is linked to a Microsoft Purview managed storage account.
- If your account was created after December 15, 2023 (or deployed using API version 2023-05-01-preview onwards):
  - Blob private endpoint is linked to a Microsoft Purview ingestion storage.
  - Queue private endpoint is linked to a Microsoft Purview ingestion storage.



Deployment checklist

Following the instructions further in this guide, you can deploy these private endpoints for an existing Microsoft Purview account:

1. Choose an appropriate Azure virtual network and a subnet to deploy Microsoft Purview private endpoints. Select one of the following options:

    - Deploy a new virtual network in your Azure subscription.
    - Locate an existing Azure virtual network and a subnet in your Azure subscription.

2. Define an appropriate DNS name resolution method, so you can access Microsoft Purview account and scan data sources using private network. You can use any of the following options:

    - Deploy new Azure DNS zones using the steps explained further in this guide.
    - Add required DNS records to existing Azure DNS zones using the steps explained further in this guide.
    - After completing the steps in this guide, add required DNS A records in your existing DNS servers manually.

3. Deploy private endpoints for an existing Microsoft Purview account.

4. Enable access to Microsoft Entra ID if your private network has network security group rules set to deny for all public internet traffic.

5. Deploy and register Self-hosted integration runtime inside the same virtual network or a peered virtual network where Microsoft Purview account and ingestion private endpoints are deployed.

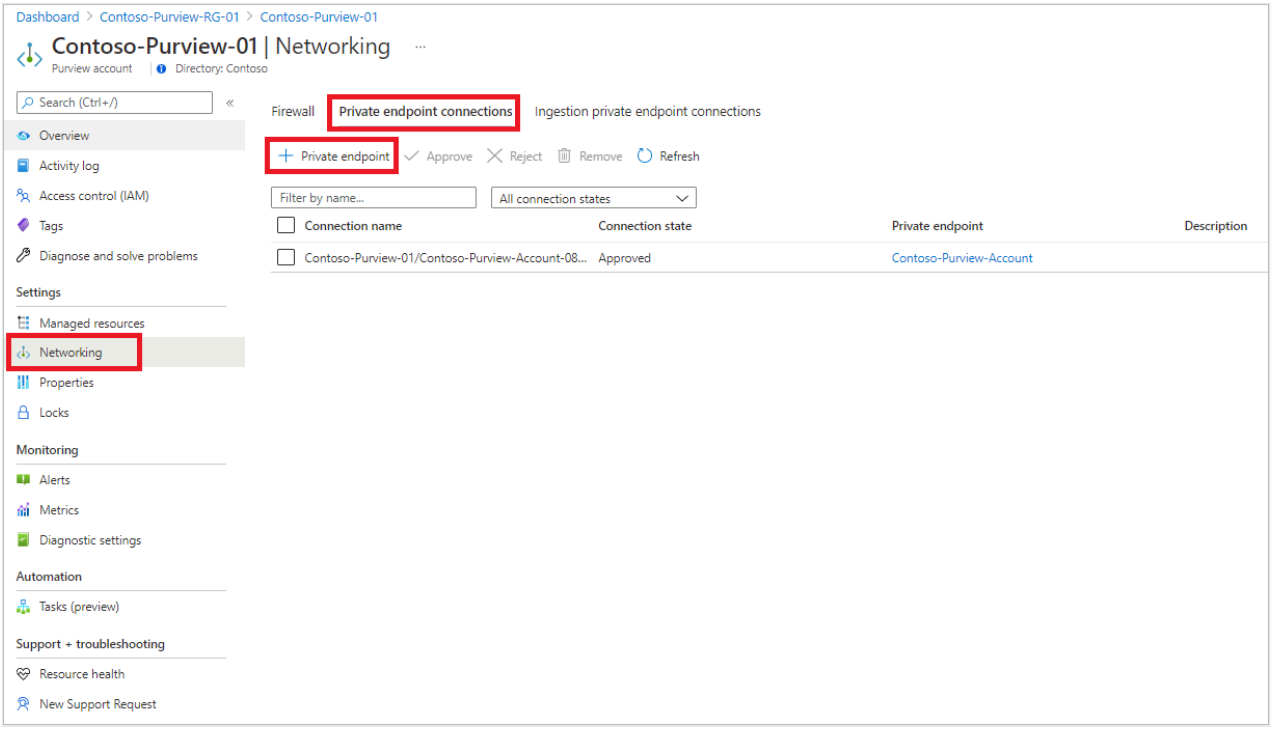
6. After completing this guide, adjust DNS configurations if needed.

7. Validate your network and name resolution between management machine, self-hosted IR VM and data sources to Microsoft Purview.

> **Note**
>
> If you configure your own Event Hubs namespace or enable a managed Event Hubs namespace after deploying your ingestion private endpoint, you'll need to redeploy the ingestion private endpoint.

Enable *account* and *portal* private endpoint

1. Go to the Azure portal, and then select your Microsoft Purview account, and under **Settings** select **Networking**, and then select **Private endpoint connections**.



2. Select + **Private endpoint** to create a new private endpoint.

3. Fill in the basic information.

4. On the **Resource** tab, for **Resource type**, select **Microsoft.Purview/accounts**.

5. For **Resource**, select the Microsoft Purview account, and for **Target sub-resource**, select **account**.

6. On the **Configuration** tab, select the virtual network and optionally, select Azure Private DNS zone to create a new Azure DNS Zone.

   > **Note**
   >
   > For DNS configuration, you can also use your existing Azure Private DNS Zones from the dropdown list or add the required DNS records to your DNS Servers manually later. For more information, see Configure DNS Name Resolution for private endpoints

7. Go to the summary page, and select **Create** to create the account private endpoint.

8. Repeat steps 2 through 7 to create the portal private endpoint. Make sure you select **portal** for **Target sub-resource**.

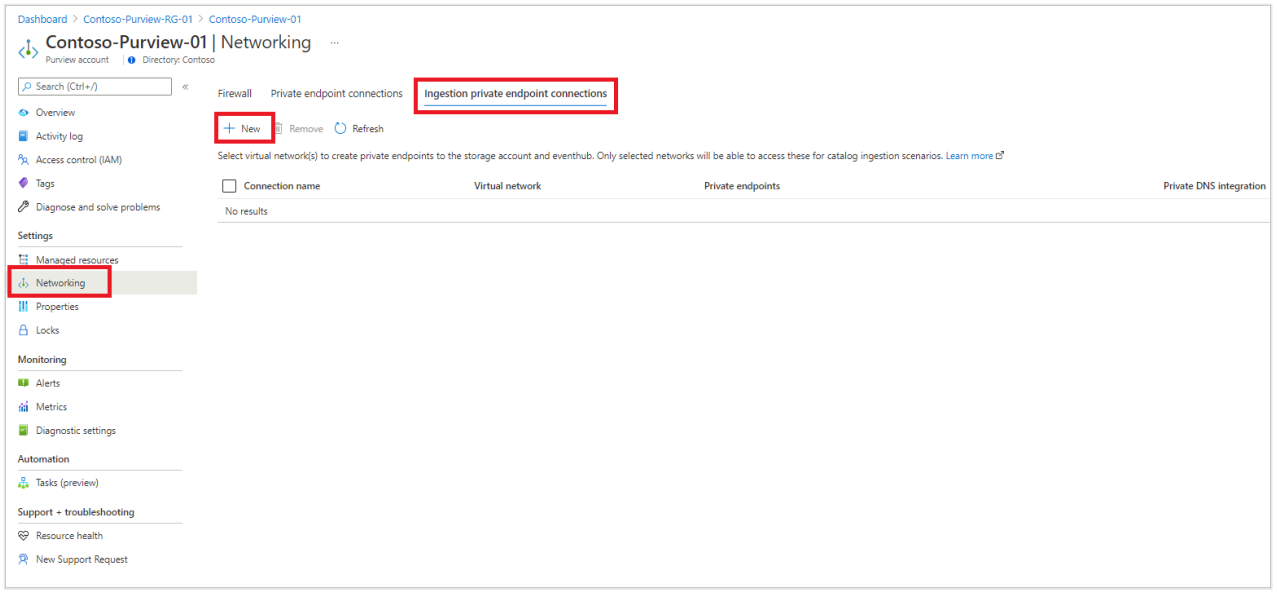
Deploy *ingestion* private endpoint

Ingestion private endpoints are deployed from the Azure portal for all Microsoft Purview instances, regardless of whether you're using the Microsoft Purview portal or the classic governance portal.

1. Go to the Azure portal, and then search for and select your Microsoft Purview account.

2. From your Microsoft Purview account, under **Settings** select **Networking**, and then select **Ingestion private endpoint connections**.

3. Under Ingestion private endpoint connections, select **+ New** to create a new ingestion private endpoint.



4. Fill in the basic information, selecting your existing virtual network and a subnet details. Optionally, select **Private DNS integration** to use Azure Private DNS Zones. Select correct Azure Private DNS Zones from each list.

   > **Note**
   >
   > You can also use your existing Azure Private DNS zones or create DNS records in your DNS Servers manually later. For more information, see Configure DNS Name Resolution for private endpoints

5. Select **Create** to create the private endpoint.

6. If you don't have one already, you need to deploy a self-hosted integration runtime to connect to the sources with private endpoints.

Configure networking for Event Hubs namespace

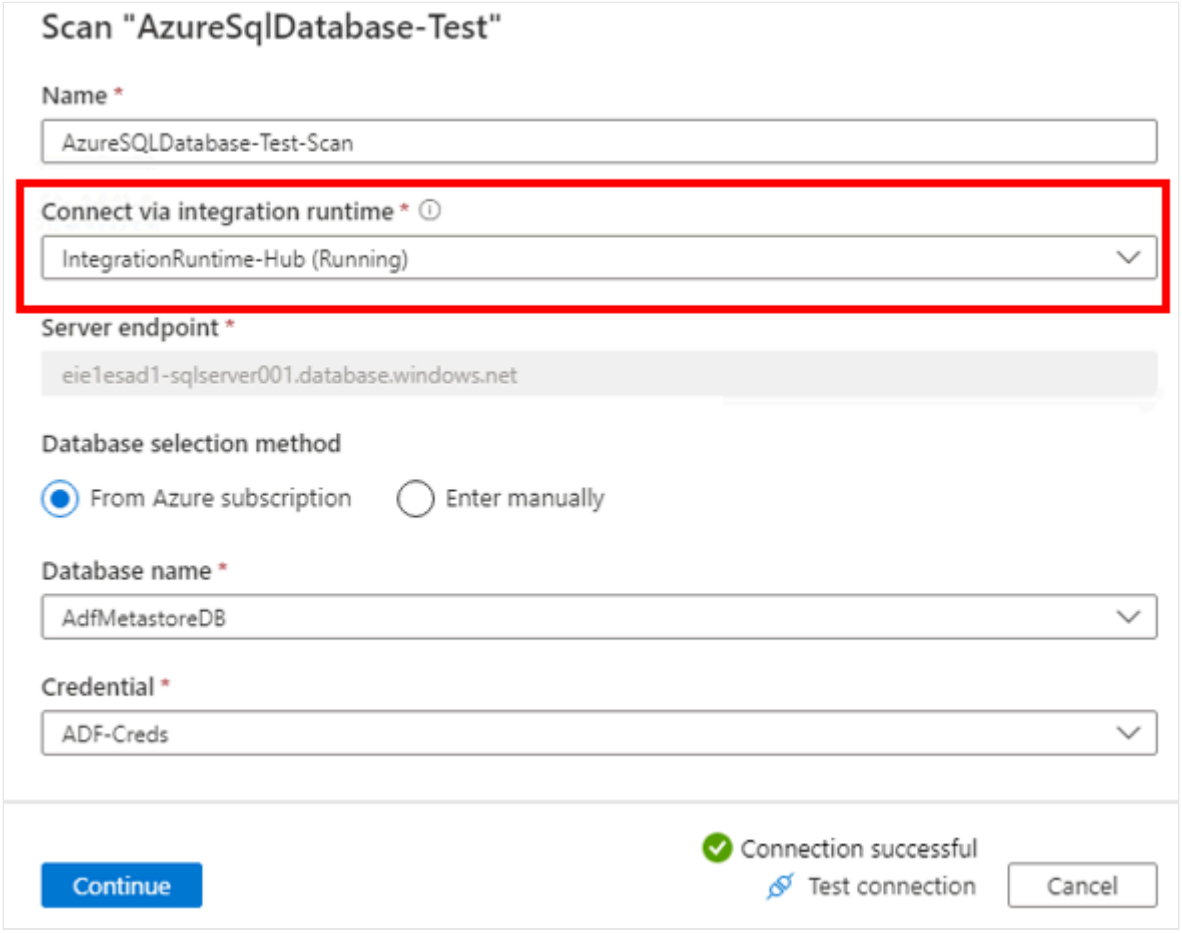
If you've configured an Event Hubs to send/receive events to/from Microsoft Purview's Apache Atlas Kafka topics, you need to confirm your Event Hubs network settings allow communication.

Deploy self-hosted integration runtime (IR) and scan your data sources

Once you deploy ingestion private endpoints for your Microsoft Purview, you need to setup and register at least one self-hosted integration runtime (IR):

- All on-premises source types like Microsoft SQL Server, Oracle, SAP, and others are currently supported only via self-hosted IR-based scans. The self-hosted IR must run within your private network and then be peered with your virtual network in Azure.
- For all Azure source types like Azure Blob Storage and Azure SQL Database, you must explicitly choose to run the scan by using a self-hosted integration runtime that is deployed in the same virtual network or a peered virtual network where Microsoft Purview account and ingestion private endpoints are deployed.

Follow the steps in Create and manage a self-hosted integration runtime to set up a self-hosted IR. Then set up your scan on the Azure source by choosing that self-hosted IR in the **Connect via integration runtime** dropdown list to ensure network isolation.



> **Important**
>
> Make sure you download and install the latest version of self-hosted integration runtime from Microsoft download center.

Enable access to Microsoft Entra ID

> **Note**

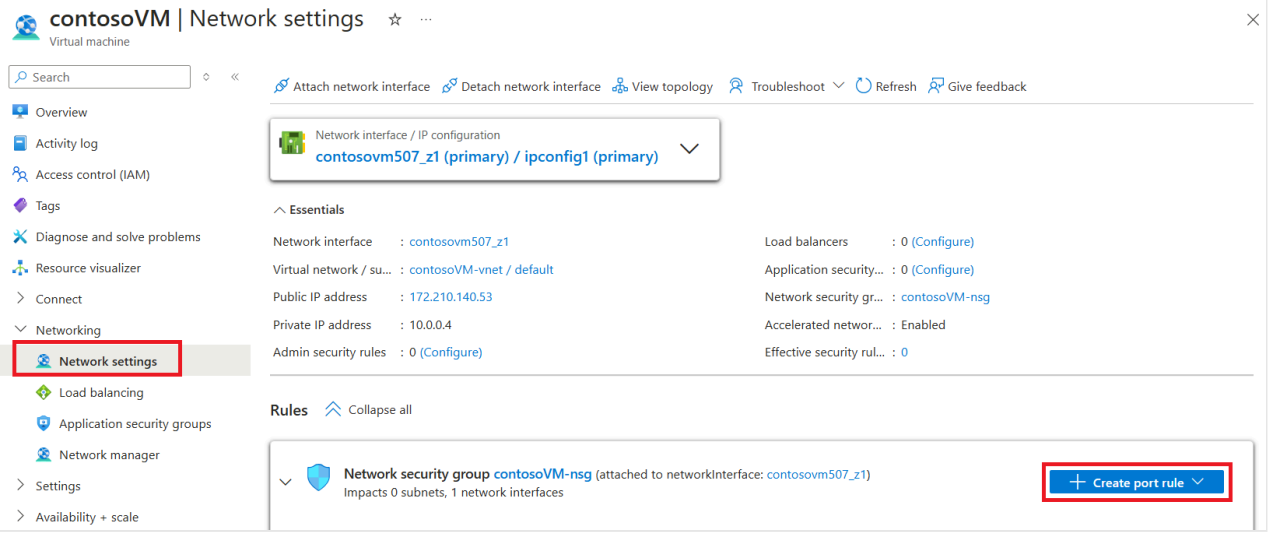
If your VM, VPN gateway, or VNet Peering gateway has public internet access, it can access the classic Microsoft Purview governance portal and the Microsoft Purview account enabled with private endpoints. For this reason, you don't have to follow the rest of the instructions. If your private network has network security group rules set to deny all public internet traffic, you'll need to add some rules to enable Microsoft Entra ID access. Follow the instructions to do so.

These instructions are provided for accessing Microsoft Purview securely from an Azure VM. Similar steps must be followed if you're using VPN or other virtual network Peering gateways.

1. Go to your virtual machine in the Azure portal, and under **Networking**, select **Network settings**.

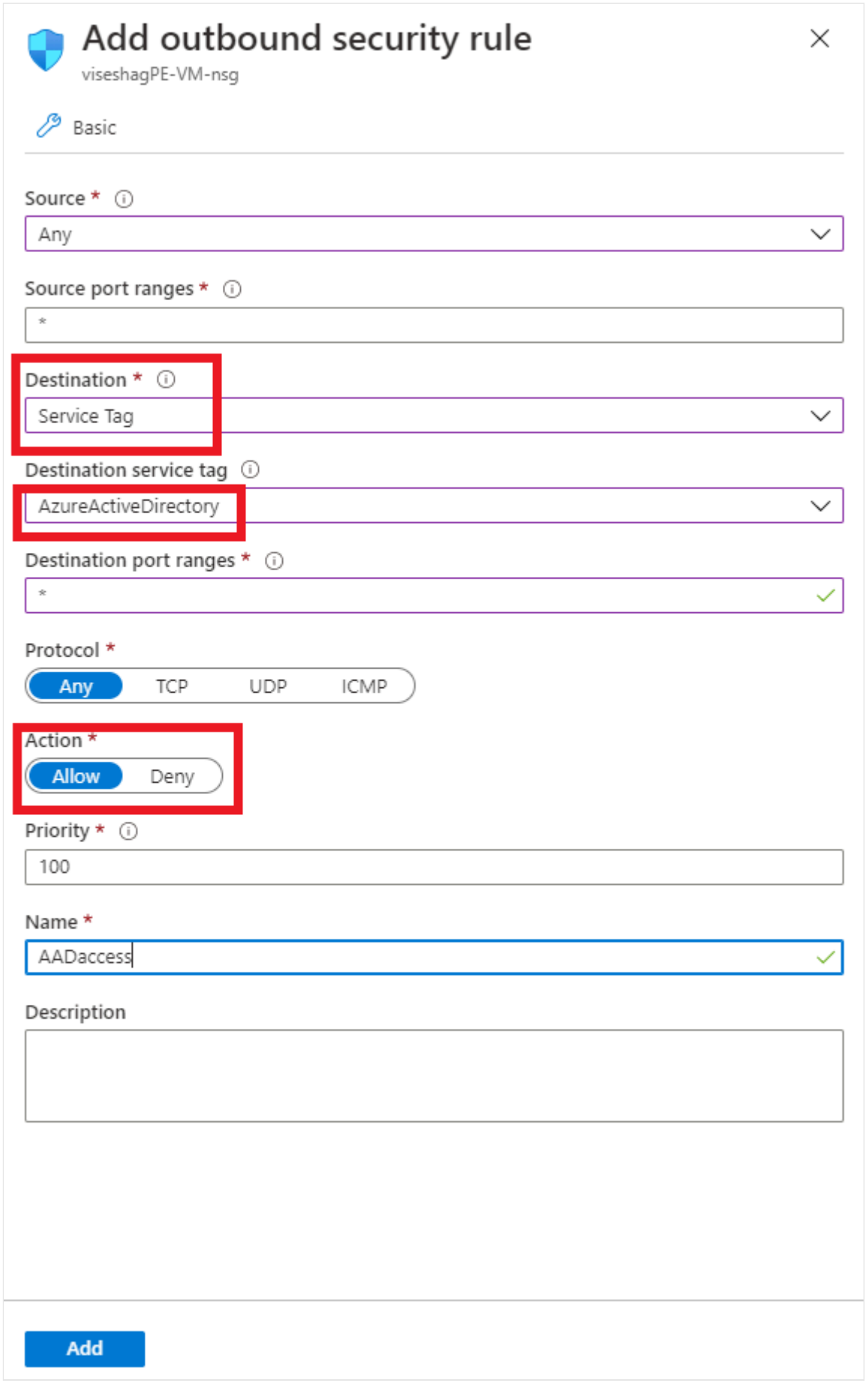
2. Select **+ Create port rule**

3. Select **Outbound port rule**



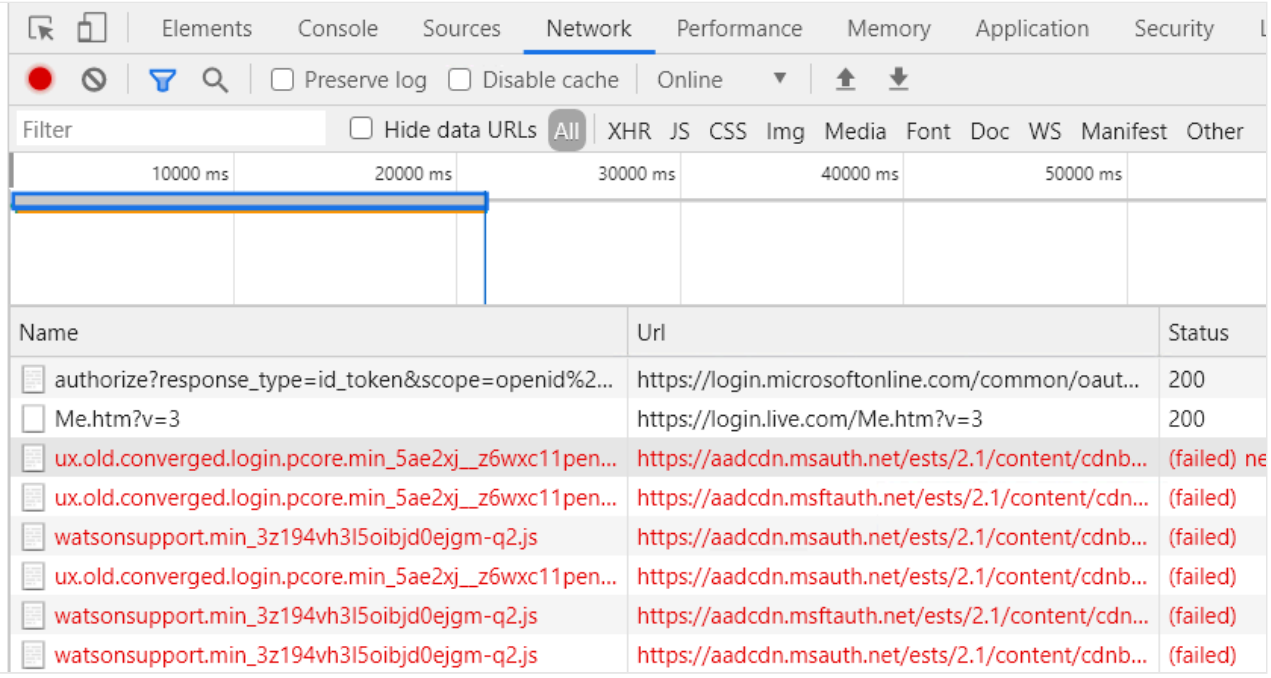
4. On the **Add outbound security rule** pane:
   a. Under **Destination**, select **Service Tag**.
   b. Under **Destination service tag**, select **AzureActiveDirectory**.
   c. Under **Destination port ranges**, select *.
   d. Under **Action**, select **Allow**.
   e. Under **Priority**, the value should be higher than the rule that denied all internet traffic.

   Create the rule.

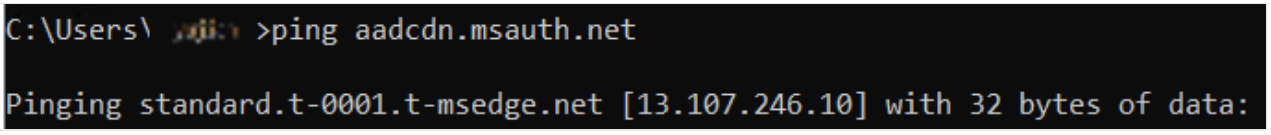


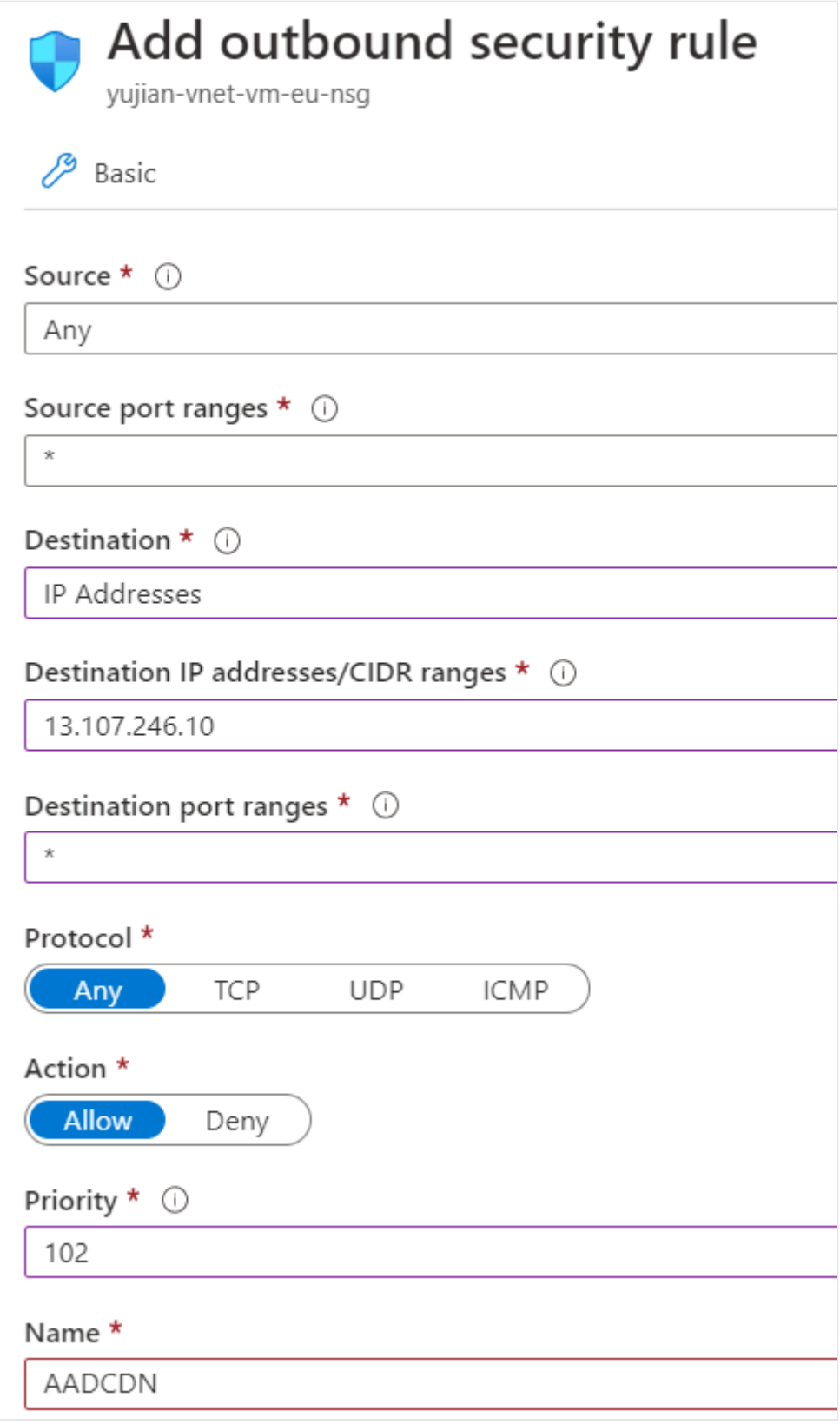
5. Follow the same steps to create another rule to allow the **AzureResourceManager** service tag. If you need to access the Azure portal, you can also add a rule for the **AzurePortal** service tag.

6. Connect to the VM and open the browser. Go to the browser console by selecting Ctrl+Shift+J, and switch to the network tab to monitor network requests. Enter web.purview.azure.com in the URL box, and try to sign in by using your Microsoft Entra credentials. Sign-in will probably fail, and on the **Network** tab on the console, you can see Microsoft Entra ID trying to access aadcdn.msauth.net but getting blocked.



7. In this case, open a command prompt on the VM, ping aadcdn.msauth.net, get its IP, and then add an outbound port rule for the IP in the VM's network security rules. Set the **Destination** to **IP Addresses** and set **Destination IP addresses** to the aadcdn IP. Because of Azure Load Balancer and Azure Traffic Manager, the Microsoft Entra Content Delivery Network IP might be dynamic. After you get its IP, it's better to add it into the VM's host file to force the browser to visit that IP to get the Microsoft Entra Content Delivery Network.





8. After the new rule is created, go back to the VM and try to sign in by using your Microsoft Entra credentials again. If signin succeeds, then the classic Microsoft Purview governance portal is ready to use. But in some cases, Microsoft Entra ID redirects to other domains to sign in based on a customer's account type. For example, for a live.com account, Microsoft Entra ID redirects to live.com to sign in, and then those requests are blocked again. For Microsoft employee accounts, Microsoft Entra ID accesses msft.sts.microsoft.com for sign-in information.

Check the networking requests on the browser **Networking** tab to see which domain's requests are getting blocked, redo the previous step to get its IP, and add outbound port rules in the network security group to allow requests for that IP. If possible, add the URL and IP to the VM's host file to fix the DNS resolution. If you know the exact sign-in domain's IP ranges, you can also directly add them into networking rules.

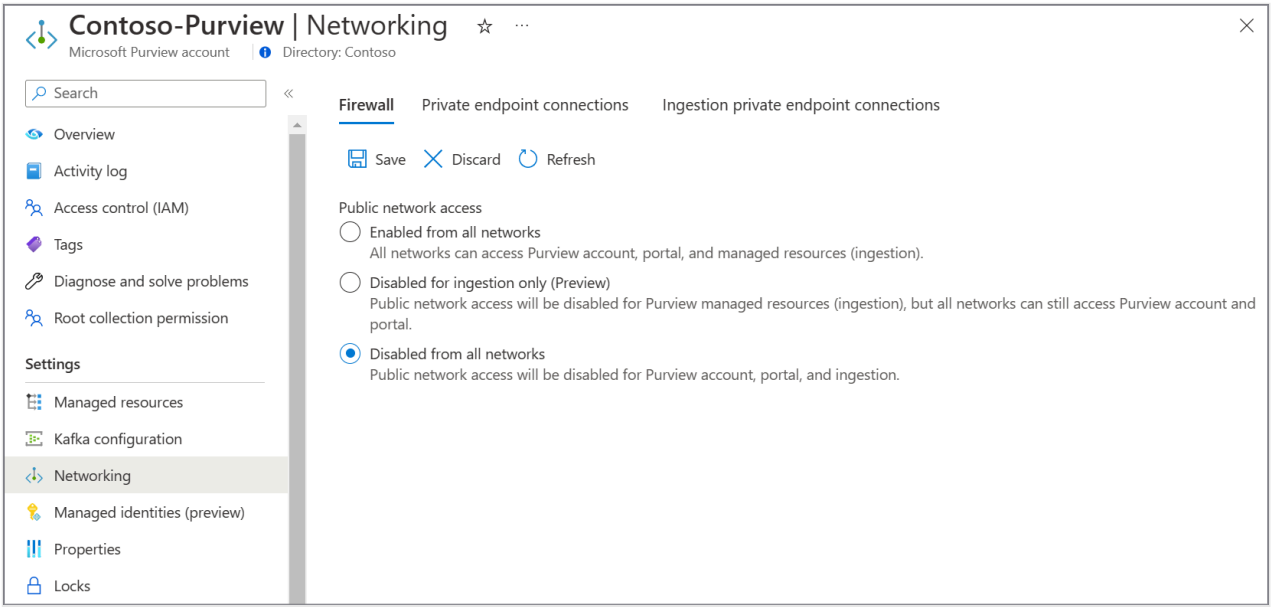
9. Now your Microsoft Entra sign-in should be successful. The classic Microsoft Purview governance portal will load successfully, but listing all the Microsoft Purview accounts won't work because it can only access a specific Microsoft Purview account. Enter `web.purview.azure.com/resource/{PurviewAccountName}` to directly visit the Microsoft Purview account that you successfully set up a private endpoint for.

Firewalls to restrict public access

To cut off access to the Microsoft Purview account completely from the public internet, follow these steps. This setting applies to both private endpoint and ingestion private endpoint connections.

1. From the Azure portal, go to the Microsoft Purview account, and under **Settings**, select **Networking**.

2. Go to the **Firewall** tab, and ensure that the toggle is set to **Disable from all networks**.



Next steps

- Verify resolution for private endpoints
- Manage data sources in Microsoft Purview
- Troubleshooting private endpoint configuration for your Microsoft Purview account

Connect privately and securely to your Microsoft Purview account

Article • 05/09/2025

> **Important**
>
> This article covers private endpoints for the classic Microsoft Purview governance portal (https://web.purview.azure.com). If you're using the new Microsoft Purview portal (https://purview.microsoft.com/), follow the documentation for private endpoints in the Microsoft Purview portal.

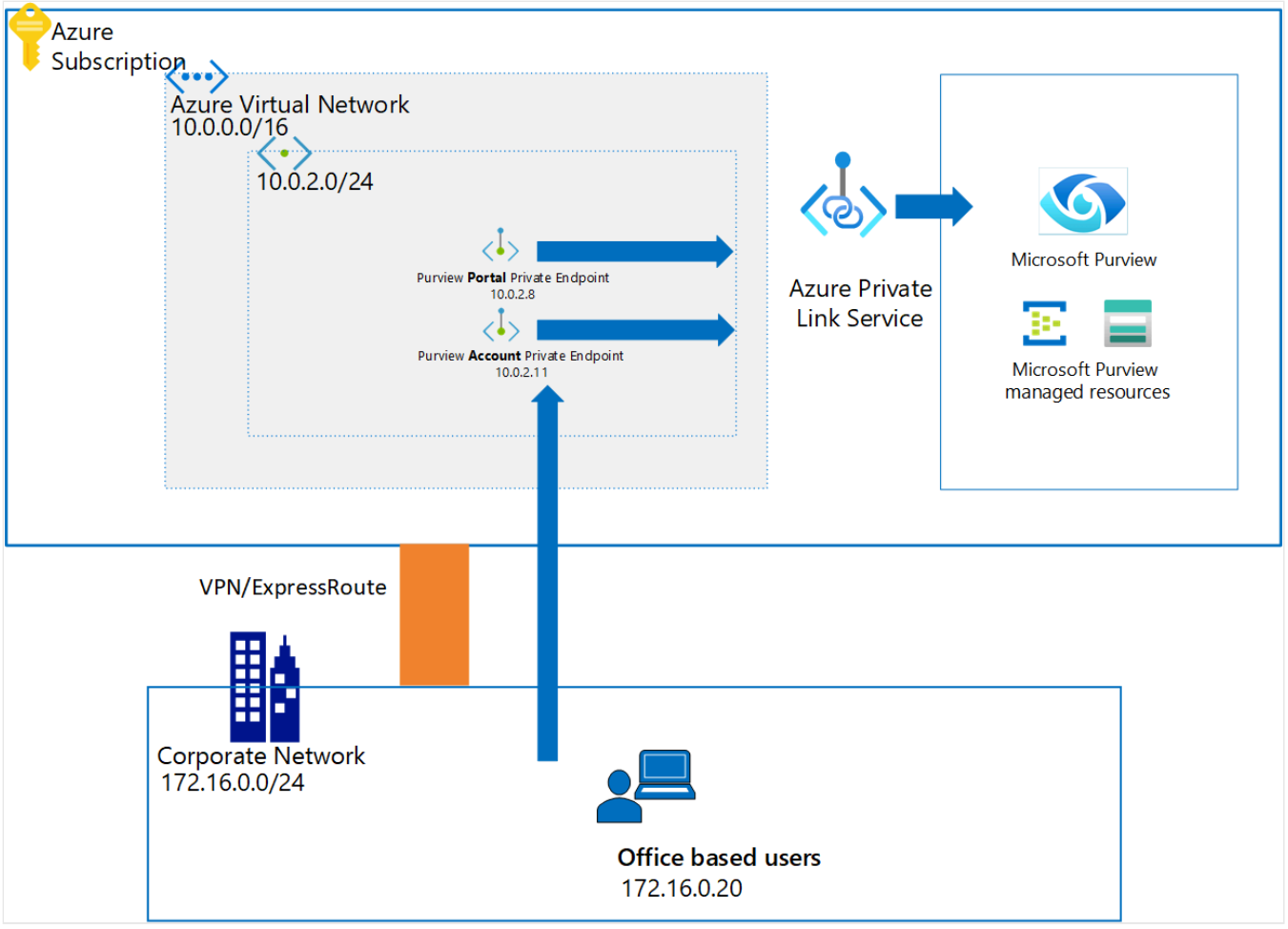
In this guide, you'll learn how to deploy private endpoints for your Microsoft Purview account to allow you to connect to your Microsoft Purview account only from VNets and private networks. To achieve this goal, you need to deploy *account* and *portal* private endpoints for your Microsoft Purview account.

> **Important**
>
> Microsoft Purview instances using the new Microsoft Purview portal can only use ingestion private endpoints.

The Microsoft Purview *account* private endpoint is used to add another layer of security by enabling scenarios where only client calls that originate from within the virtual network are allowed to access the Microsoft Purview account. This private endpoint is also a prerequisite for the portal private endpoint.

The Microsoft Purview *portal* private endpoint is required to enable connectivity to the classic Microsoft Purview governance portal using a private network.

> **Note**
>
> If you only create *account* and *portal* private endpoints, you won't be able to run any scans. To enable scanning on a private network, you will also need to create an ingestion private endpoint.



For more information about Azure Private Link service, see private links and private endpoints to learn more.

Deployment checklist

Using this guide, you can deploy these private endpoints for an existing Microsoft Purview account:

1. Choose an appropriate Azure virtual network and a subnet to deploy Microsoft Purview private endpoints. Select one of the following options:

    - Deploy a new virtual network in your Azure subscription.
    - Locate an existing Azure virtual network and a subnet in your Azure subscription.

2. Define an appropriate DNS name resolution method, so Microsoft Purview account and web portal can be accessible through private IP addresses. You can use any of the following options:

    - Deploy new Azure DNS zones using the steps explained further in this guide.
    - Add required DNS records to existing Azure DNS zones using the steps explained further in this guide.

- After completing the steps in this guide, add required DNS A records in your existing DNS servers manually.

3. Deploy account and portal private endpoints for an existing Microsoft Purview account.

4. Enable access to Microsoft Entra ID if your private network has network security group rules set to deny for all public internet traffic.

5. After completing this guide, adjust DNS configurations if needed.

6. Validate your network and name resolution from management machine to Microsoft Purview.

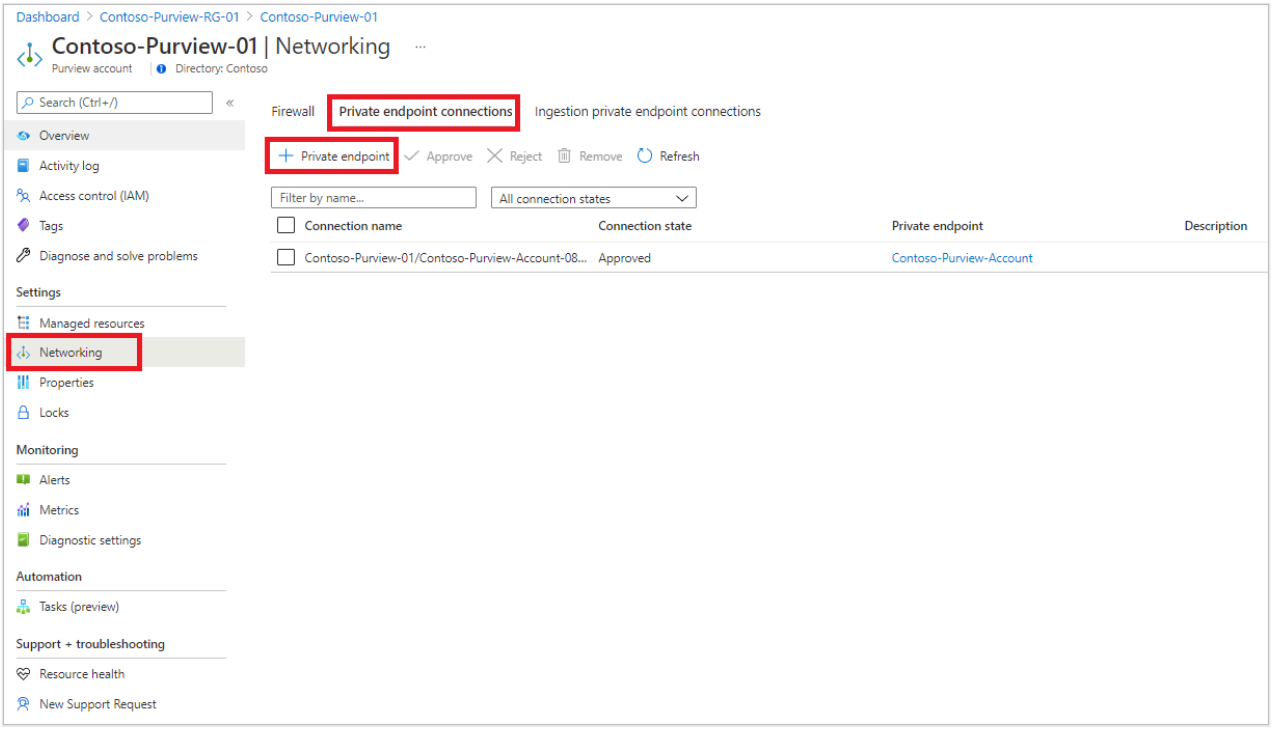
Enable *account* and *portal* private endpoint

There are two ways you can add Microsoft Purview *account* and *portal* private endpoints for an existing Microsoft Purview account:

- Use the Azure portal (Microsoft Purview account).
- Use the Private Link Center.

Use the Azure portal (Microsoft Purview account)

1. Go to the Azure portal, and then select your Microsoft Purview account, and under **Settings** select **Networking**, and then select **Private endpoint connections**.



2. Select + **Private endpoint** to create a new private endpoint.

3. Fill in the basic information.

4. On the **Resource** tab, for **Resource type**, select **Microsoft.Purview/accounts**.

5. For **Resource**, select the Microsoft Purview account, and for **Target sub-resource**, select **account**.

6. On the **Configuration** tab, select the virtual network and optionally, select Azure Private DNS zone to create a new Azure DNS Zone.

   > **Note**
   >
   > For DNS configuration, you can also use your existing Azure Private DNS Zones from the dropdown list or add the required DNS records to your DNS Servers manually later. For more information, see Configure DNS Name Resolution for private endpoints

7. Go to the summary page, and select **Create** to create the portal private endpoint.

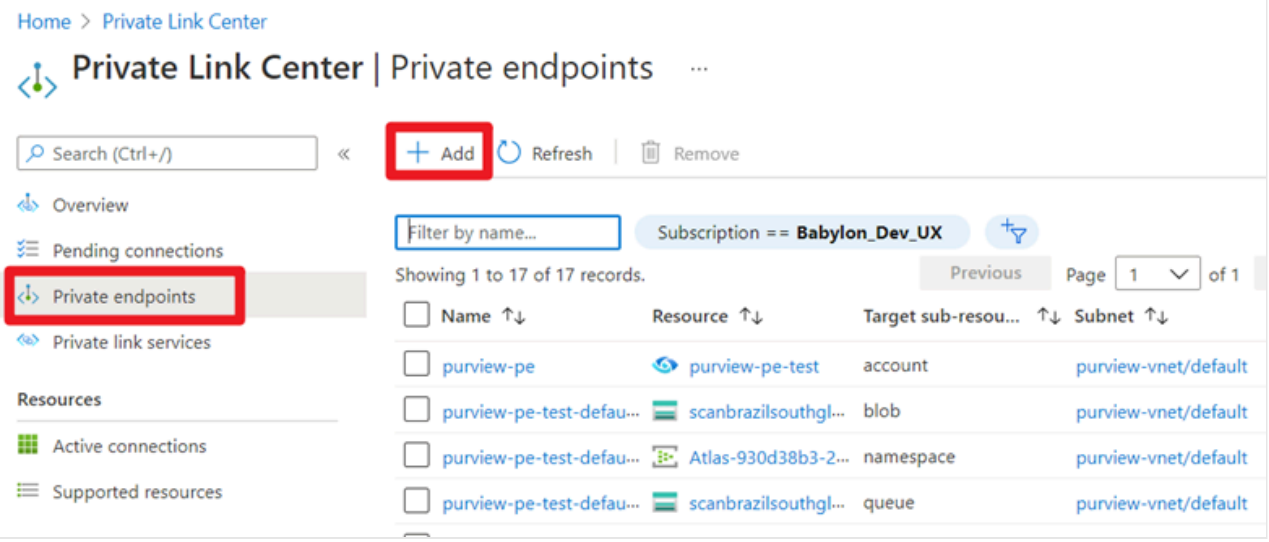
8. Follow the same steps when you select **portal** for **Target sub-resource**.

Use the Private Link Center

1. Go to the Azure portal.

2. In the search bar at the top of the page, search for **private link** and go to the **Private Link** pane by selecting the first option.

3. Select + **Add**, and fill in the basic details.



4. For **Resource**, select the already created Microsoft Purview account. For **Target sub-resource**, select **account**.

5. On the **Configuration** tab, select the virtual network and private DNS zone. Go to the summary page, and select **Create** to create the account private endpoint.

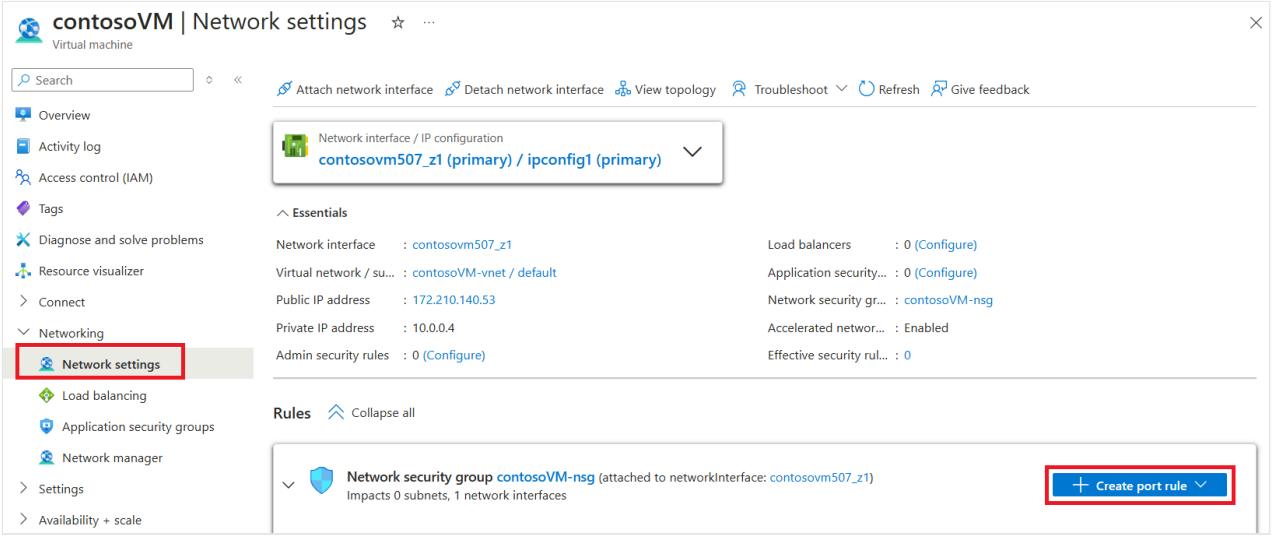
> **Note**
>
> Follow the same steps when you select **portal** for **Target sub-resource**.

Enable access to Microsoft Entra ID

> **Note**
>
> If your VM, VPN gateway, or VNet Peering gateway has public internet access, it can access the Microsoft Purview portal and the Microsoft Purview account enabled with private endpoints. For this reason, you don't have to follow the rest of the instructions. If your private network has network security group rules set to deny all public internet traffic, you'll need to add some rules to enable Microsoft Entra ID access. Follow the instructions to do so.

These instructions are provided for accessing Microsoft Purview securely from an Azure VM. Similar steps must be followed if you're using VPN or other virtual network Peering gateways.

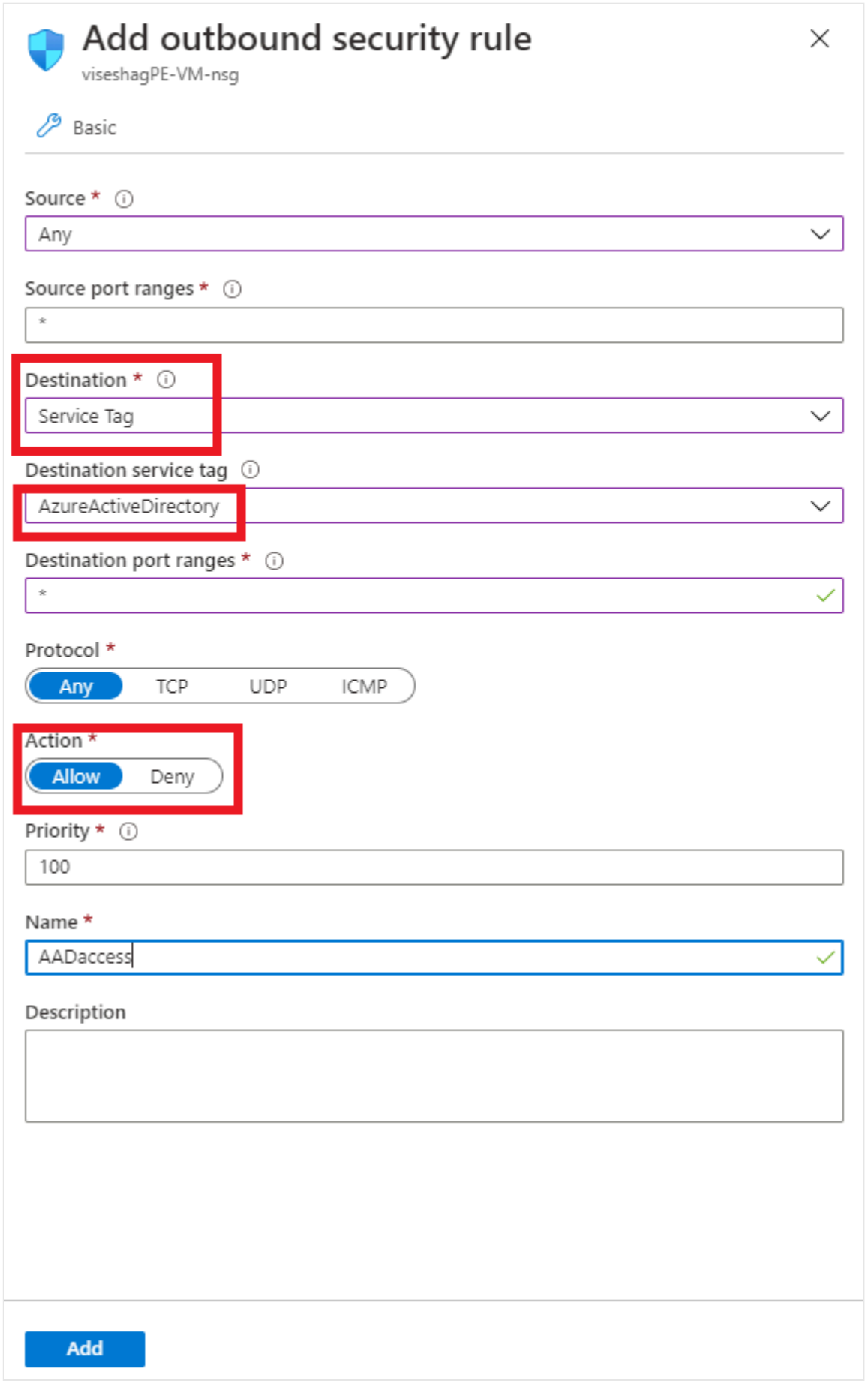
1. Go to your VM in the Azure portal, and under **Settings**, select **Networking**. Then select **Outbound port rules**, **Add outbound port rule**.



2. On the **Add outbound security rule** pane:
   a. Under **Destination**, select **Service Tag**.

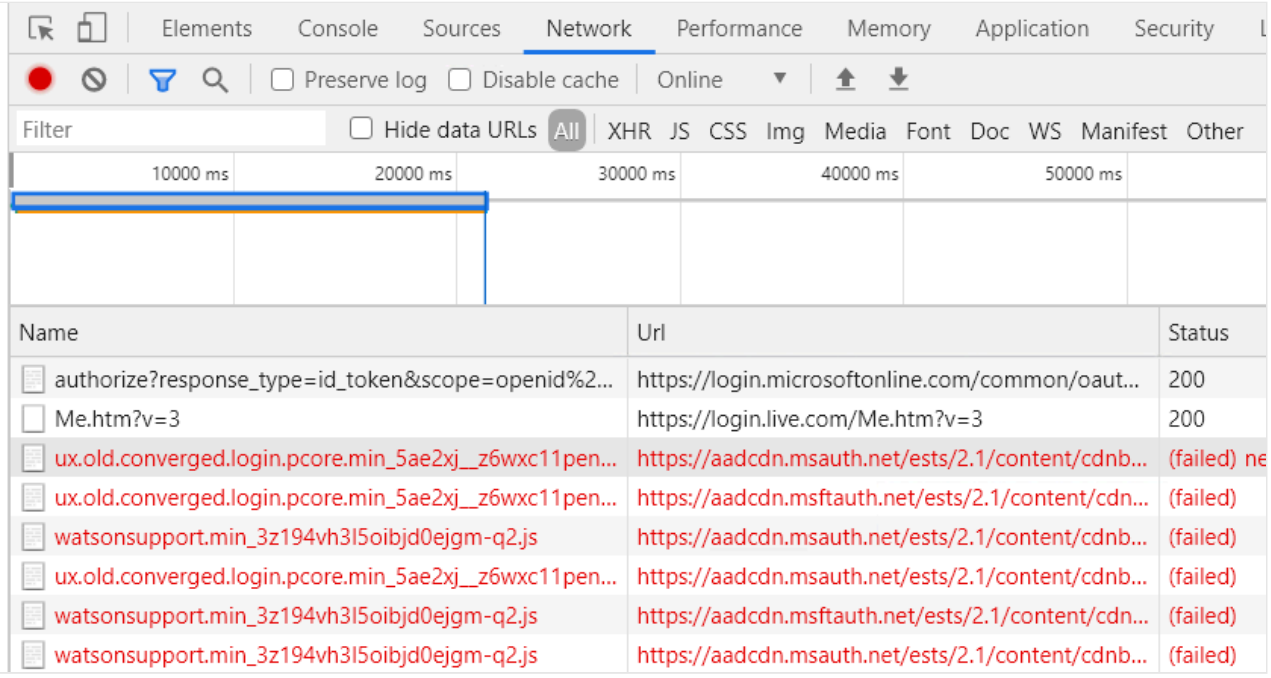
b. Under **Destination service tag**, select **AzureActiveDirectory**.
c. Under **Destination port ranges**, select *.
d. Under **Action**, select **Allow**.
e. Under **Priority**, the value should be higher than the rule that denied all internet traffic.

Create the rule.

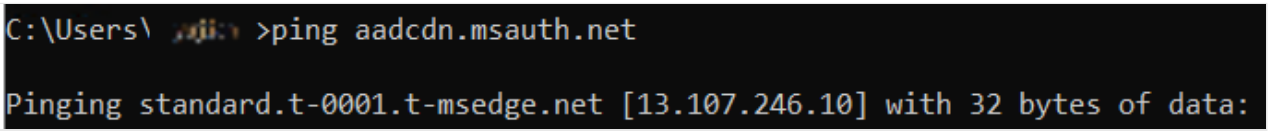


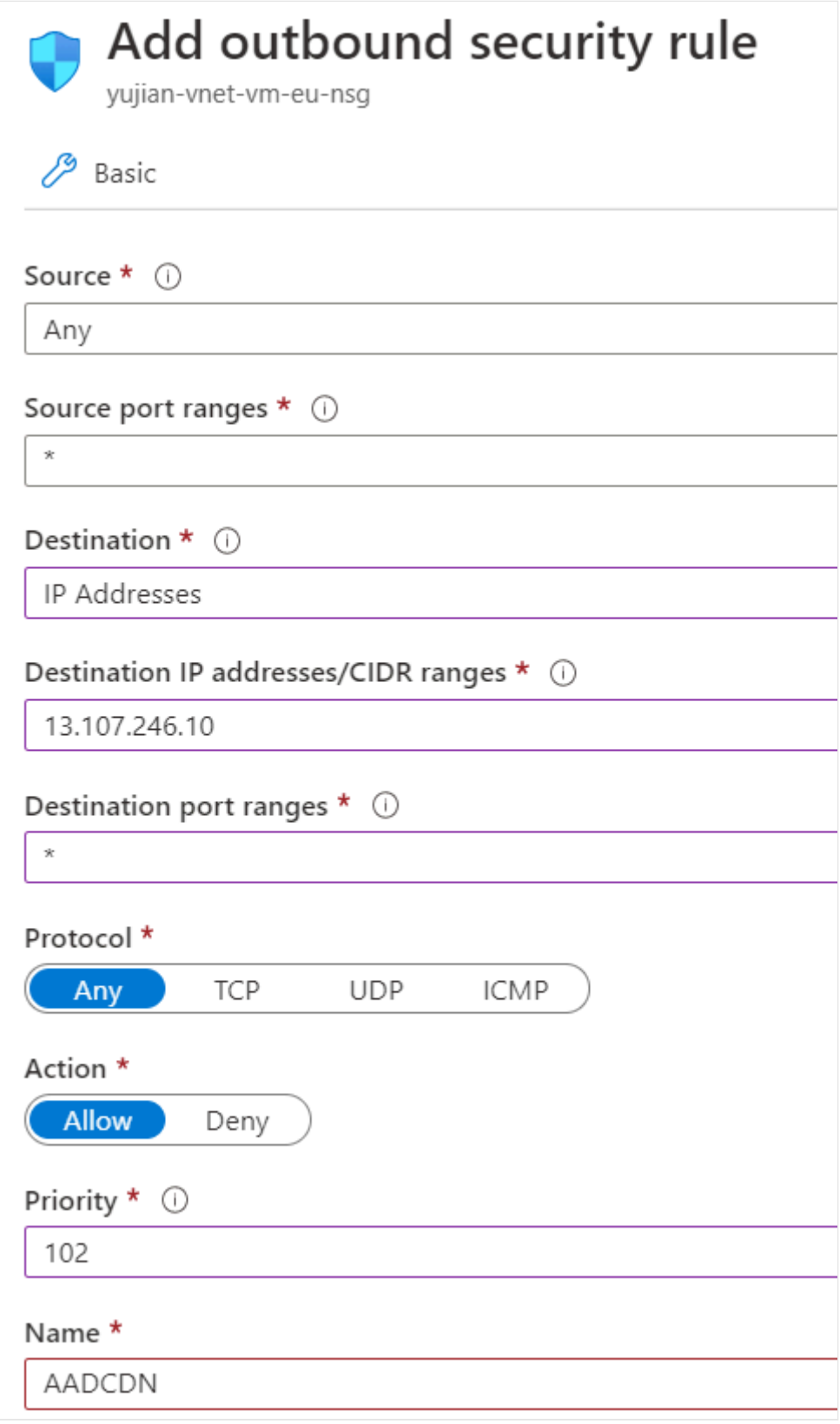
3. Follow the same steps to create another rule to allow the **AzureResourceManager** service tag. If you need to access the Azure portal, you can also add a rule for the **AzurePortal** service tag.

4. Connect to the VM and open the browser. Go to the browser console by selecting Ctrl+Shift+J, and switch to the network tab to monitor network requests. Enter web.purview.azure.com in the URL box, and try to sign in by using your Microsoft Entra credentials. Sign-in will probably fail, and on the **Network** tab on the console, you can see Microsoft Entra ID trying to access aadcdn.msauth.net but getting blocked.



5. In this case, open a command prompt on the VM, ping aadcdn.msauth.net, get its IP, and then add an outbound port rule for the IP in the VM's network security rules. Set the **Destination** to **IP Addresses** and set **Destination IP addresses** to the aadcdn IP. Because of Azure Load Balancer and Azure Traffic Manager, the Microsoft Entra Content Delivery Network IP might be dynamic. After you get its IP, it's better to add it into the VM's host file to force the browser to visit that IP to get the Microsoft Entra Content Delivery Network.





6. After the new rule is created, go back to the VM and try to sign in by using your Microsoft Entra credentials again. If sign-in succeeds, then the Microsoft Purview portal is ready to use. But in some cases, Microsoft Entra ID redirects to other domains to sign in based on a customer's account type. For example, for a live.com account, Microsoft Entra ID redirects to live.com to sign in, and then those requests are blocked again. For Microsoft employee accounts, Microsoft Entra ID accesses msft.sts.microsoft.com for sign-in information.

Check the networking requests on the browser **Networking** tab to see which domain's requests are getting blocked, redo the previous step to get its IP, and add outbound port rules in the network security group to allow requests for that IP. If possible, add the URL and IP to the VM's host file to fix the DNS resolution. If you know the exact sign-in domain's IP ranges, you can also directly add them into networking rules.

7. Now your Microsoft Entra sign-in should be successful. The Microsoft Purview portal will load successfully, but listing all the Microsoft Purview accounts won't work because it can only access a specific Microsoft Purview account. Enter `web.purview.azure.com/resource/{PurviewAccountName}` to directly visit the Microsoft Purview account that you successfully set up a private endpoint for.

Next steps

- Verify resolution for private endpoints
- Manage data sources in Microsoft Purview
- Troubleshooting private endpoint configuration for your Microsoft Purview account

Configure and verify DNS Name Resolution for Microsoft Purview private endpoints

Article • 05/09/2025

> **Important**
>
> This article covers private endpoints for the classic Microsoft Purview governance portal (https://web.purview.azure.com). If you're using the new Microsoft Purview portal (https://purview.microsoft.com/), follow the documentation for private endpoints in the Microsoft Purview portal.

Conceptual overview

Accurate name resolution is a critical requirement when setting up private endpoints for your Microsoft Purview accounts.

You might need to enable internal name resolution in your DNS settings to resolve the private endpoint IP addresses to the fully qualified domain name (FQDN) from data sources and your management machine to Microsoft Purview account and self-hosted integration runtime, depending on scenarios that you're deploying.

> **Tip**
>
> When you deploy your DNS Zones, don't use specific IP addresses. IP addresses for Azure resources are not static, and building your DNS Zones with static IP addresses will eventually cause errors.

Deployment options

Use any of the following options to set up internal name resolution when using private endpoints for your Microsoft Purview account:

- Deploy new Azure Private DNS Zones in your Azure environment part of private endpoint deployment. (Default option)
- Use existing Azure Private DNS Zones. Use this option if you're using a private endpoint in a hub-and-spoke model from a different subscription or even within the same subscription.
- Use your own DNS Servers if you don't use DNS forwarders and instead you manage A records directly in your on-premises DNS servers.

> **Tip**
>
> - Classic Microsoft Purview governance portal (https://web.purview.azure.com): *account*, *portal*, and *ingestion* private endpoints are supported.
> - New Microsoft Purview portal (https://purview.microsoft.com/): *platform* and *ingestion* private endpoints are supported.

Option 1 - Deploy new Azure Private DNS Zones

Deploy new Azure Private DNS Zones

To enable internal name resolution, you can deploy the required Azure DNS Zones inside your Azure subscription where Microsoft Purview account is deployed.



Name	Type	Location
privatelink.blob.core.windows.net	Private DNS zone	Global
privatelink.purview.azure.com	Private DNS zone	Global
privatelink.purviewstudio.azure.com	Private DNS zone	Global
privatelink.queue.core.windows.net	Private DNS zone	Global
privatelink.servicebus.windows.net	Private DNS zone	Global

When you create ingestion, portal and account private endpoints, the DNS CNAME resource records for Microsoft Purview is automatically updated to an alias in few subdomains with the prefix `privatelink`:

- By default, during the deployment of *account* or *platform* private endpoints for your Microsoft Purview account, we also create a private DNS zone that corresponds to the `privatelink` subdomain for Microsoft Purview as `privatelink.purview.azure.com` for the classic Microsoft Purview governance portal and `privatelink.purview-service.microsoft.com` for the Microsoft Purview portal, including DNS A resource records for the private endpoints.
- During the deployment of *portal* private endpoint for your Microsoft Purview account, we also create a new private DNS zone that corresponds to the `privatelink` subdomain for Microsoft Purview as `privatelink.purviewstudio.azure.com` including DNS A resource records for *Web*.
- If you enable ingestion private endpoints, other DNS zones are required for managed or configured resources.

The following table shows an example of Azure Private DNS zones and DNS A Records that are deployed as part of configuration of private endpoint for a Microsoft Purview account if you enable *Private DNS integration* during the deployment:

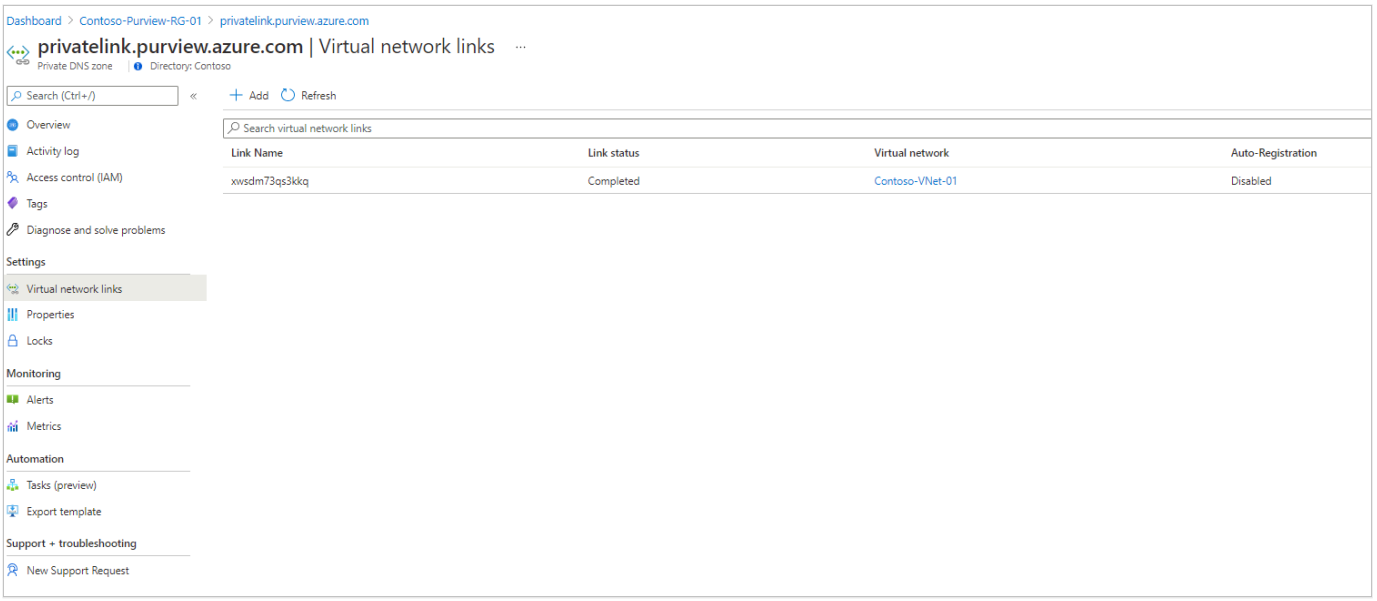
Expand table

Private endpoint	Private endpoint associated to	Portal availability	DNS Zone (new)	A Record (example)
Account	Microsoft Purview	`privatelink.purview.azure.com`	Contoso-Purview	
Platform	Microsoft Purview	`privatelink.purview-service.microsoft.com`	Contoso-Purview	
Portal	Microsoft Purview	`privatelink.purviewstudio.azure.com`	Web	
Platform	Microsoft Purview	`privatelink.purview-service.microsoft.com`	Web	
Ingestion	Microsoft Purview ingestion - Blob*	`privatelink.blob.core.windows.net`	ingestioneus2eastusksqky	
Ingestion	Microsoft Purview ingestion - Blob*	`privatelink.blob.storage.azure.net`	ingestioneus2eastusksqky	
Ingestion	Microsoft Purview ingestion - Queue*	`privatelink.queue.core.windows.net`	ingestioneus2eastusksqky	
Ingestion	Microsoft Purview ingestion - Queue*	`privatelink.queue.storage.azure.net`	ingestioneus2eastusksqky	
Ingestion	Microsoft Purview configured Event Hubs - Event Hubs**	`privatelink.servicebus.windows.net`	atlas-12345678-1234-1234-abcd-123456789abc	

> **Note**
>
> *If your account was created before December 15, 2023, the endpoint is deployed to your managed storage account. If it was created after November 10 (or deployed using API version 2023-05-01-preview onwards), it points to the ingestion storage.
>
> **Your account only has an associated Event Hubs namespace if it is configured for kafka notifications or created before December 15, 2022.

Validate virtual network links on Azure Private DNS Zones

Once the private endpoint deployment is completed, make sure there's a Virtual network link on all corresponding Azure Private DNS zones to Azure virtual network where private endpoint was deployed.



For more information, see Azure private endpoint DNS configuration.

Verify internal name resolution

When you resolve the Microsoft Purview endpoint URL from outside the virtual network with the private endpoint, it resolves to the public endpoint of Microsoft Purview. When resolved from the virtual network hosting the private endpoint, the Microsoft Purview endpoint URL resolves to the private endpoint's IP address.

As an example, if a Microsoft Purview account name is 'Contoso-Purview', when it's resolved from outside the virtual network that hosts the private endpoint, it will be:

Expand table

Name	Type	Value
`Contoso-Purview.purview.azure.com`	CNAME	`Contoso-Purview.privatelink.purview.azure.com`
`Contoso-Purview.privatelink.purview.azure.com`	CNAME	<Microsoft Purview public endpoint>
<Microsoft Purview public endpoint>	A	<Microsoft Purview public IP address>
`Web.purview.azure.com`	CNAME	<Microsoft Purview governance portal public endpoint>

The DNS resource records for Contoso-Purview, when resolved in the virtual network hosting the private endpoint, will be:

Expand table

Name	Type	Value
`Contoso-Purview.purview.azure.com`	CNAME	`Contoso-Purview.privatelink.purview.azure.com`
`Contoso-Purview.privatelink.purview.azure.com`	A	<Microsoft Purview account private endpoint IP address>
`Web.purview.azure.com`	CNAME	<Microsoft Purview portal private endpoint IP address>

Option 2 - Use existing Azure Private DNS Zones

Use existing Azure Private DNS Zones

During the deployment of Microsoft Purview private endpoints, you can choose *Private DNS integration* using existing Azure Private DNS zones. This is common case for organizations where private endpoint is used for other services in Azure. In this case, during the deployment of private endpoints, make sure you select the existing DNS zones instead of creating new ones.

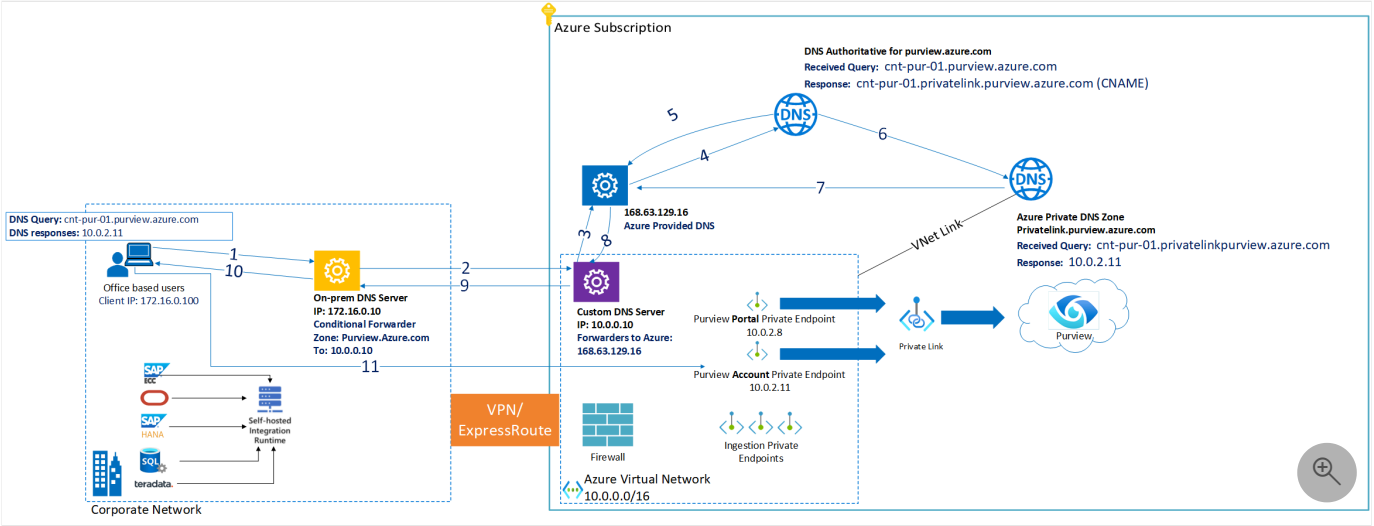
This scenario also applies if your organization uses a central or hub subscription for all Azure Private DNS Zones.

The following list shows the required Azure DNS zones and A records for Microsoft Purview private endpoints:

> **Important**
>
> Update all names with `Contoso-Purview`,`ingestioneus2eastusksqky` and `atlas-12345678-1234-1234-abcd-123456789abc` with corresponding Azure resources name in your environment.
>
> - `Contoso-Purview` is the name of your Microsoft Purview account.
> - You only need the endpoints associated with the portal you're using:
>   - Classic Microsoft Purview governance portal (https://web.purview.azure.com): *account*, *portal*, and *ingestion* private endpoints are supported.
>   - New Microsoft Purview portal (https://purview.microsoft.com/): *platform* and *ingestion* private endpoints are supported.
> - If your account is configured for kafka notifications or was created before December 15, 2022, `atlas-12345678-1234-1234-abcd-123456789abc` is your Event Hubs namespace.
> - If your account was created **before** December 15, 2023, for `ingestioneus2eastusksqky` use the name of your Microsoft Purview managed storage account.
> - If your account was created **after** December 15, 2023 (or deployed using API version 2023-05-01-preview onwards), leave `ingestioneus2eastusksqky` as is.

Expand table

Private endpoint	Private endpoint associated to	Portal availability	DNS Zone (existing)	A Record (example)
Account	Microsoft Purview	`privatelink.purview.azure.com`	Contoso-Purview	
Platform	Microsoft Purview	`privatelink.purview-service.microsoft.com`	Contoso-Purview	
Portal	Microsoft Purview	`privatelink.purviewstudio.azure.com`	Web	
Ingestion	Microsoft Purview ingestion - Blob*	`privatelink.blob.core.windows.net`	ingestioneus2eastusksqky	
Ingestion	Microsoft Purview ingestion - Blob*	`privatelink.blob.storage.azure.net`	ingestioneus2eastusksqky	
Ingestion	Microsoft Purview ingestion - Queue*	`privatelink.queue.core.windows.net`	ingestioneus2eastusksqky	
Ingestion	Microsoft Purview ingestion - Queue*	`privatelink.queue.storage.azure.net`	ingestioneus2eastusksqky	
Ingestion	Microsoft Purview configured Event Hubs**	`privatelink.servicebus.windows.net`	atlas-12345678-1234-1234-abcd-123456789abc	

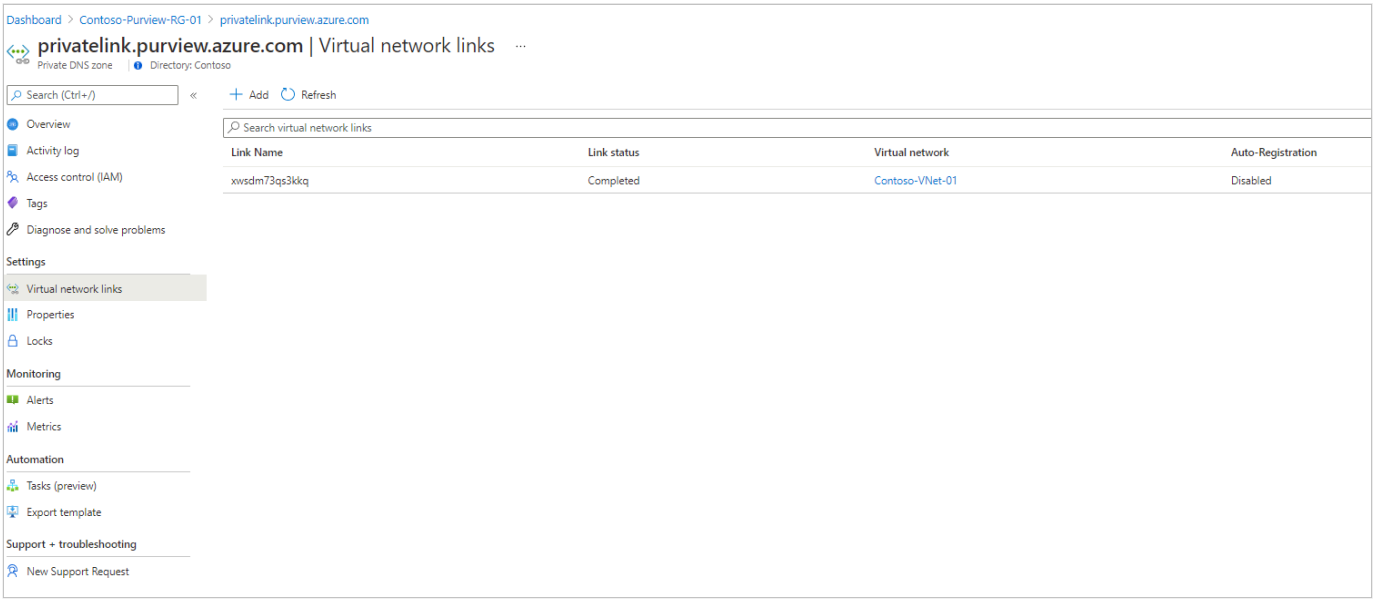


> **Note**
>
> *If your account was created before December 15, 2023, the endpoint is deployed to your managed storage account. If it was created after November 10 (or deployed using API version 2023-05-01-preview onwards), it points to the ingestion storage. **Your account only has an associated Event Hubs namespace if it is configured for kafka notifications or created before December 15, 2022.

For more information, see Virtual network workloads without custom DNS server and On-premises workloads using a DNS forwarder scenarios in Azure Private Endpoint DNS configuration.

Verify virtual network links on Azure Private DNS Zones

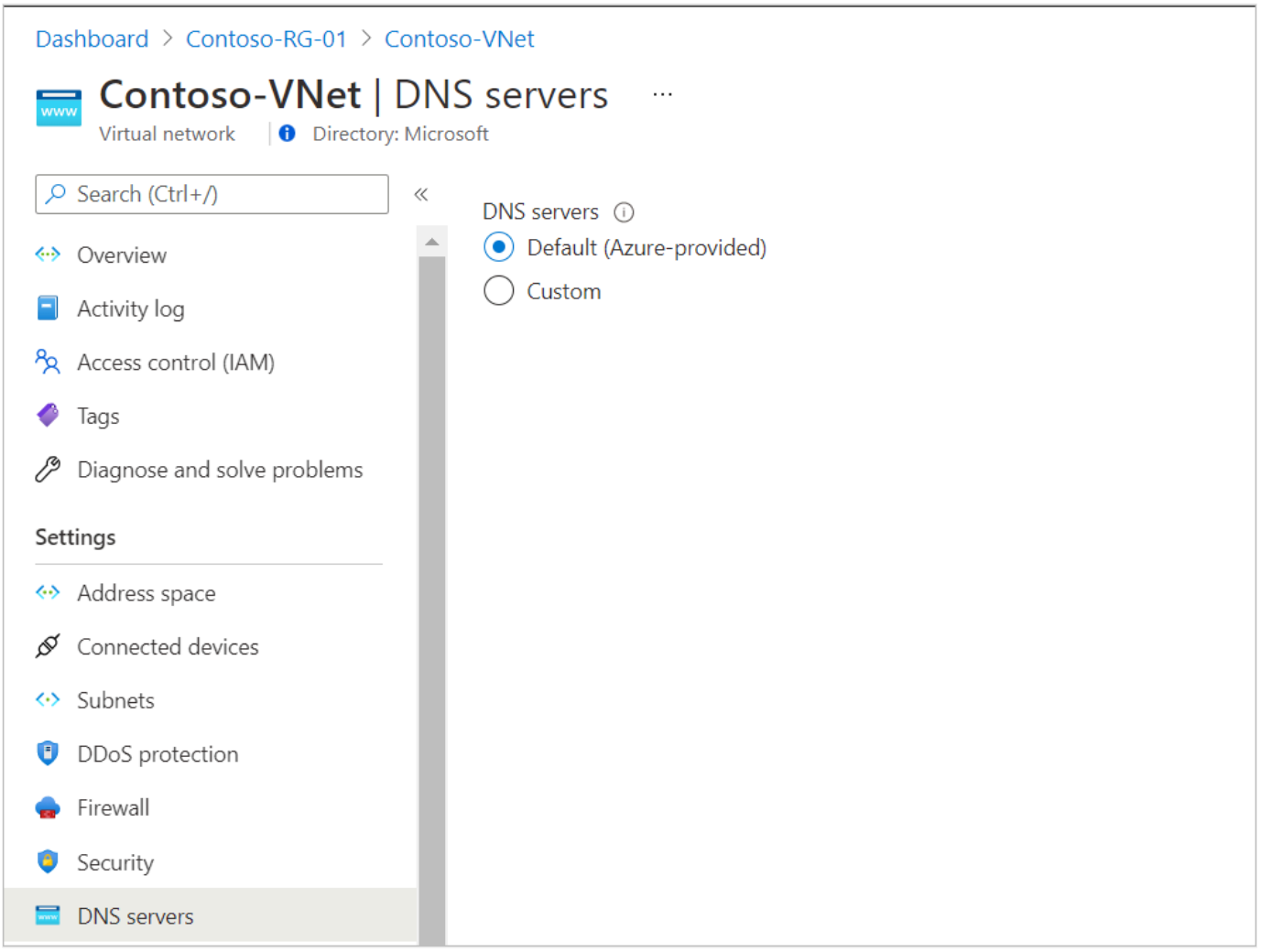
Once the private endpoint deployment is completed, make sure there's a Virtual network link on all corresponding Azure Private DNS zones to Azure virtual network where private endpoint was deployed.



For more information, see Azure private endpoint DNS configuration.

Configure DNS Forwarders if custom DNS is used

Additionally it's required to validate your DNS configurations on Azure virtual network where self-hosted integration runtime VM or management PC is located.



- If it's configured to *Default*, no further action is required in this step.
- If custom DNS server is used, you should add corresponding DNS forwarders inside your DNS servers for the following zones:
  - Purview.azure.com
  - purviewstudio.azure.com
  - Blob.core.windows.net
  - Queue.core.windows.net
  - Blob.storage.azure.net
  - Queue.storage.azure.net
  - Servicebus.windows.net

Verify internal name resolution

When you resolve the Microsoft Purview endpoint URL from outside the virtual network with the private endpoint, it resolves to the public endpoint of Microsoft Purview. When resolved from the virtual network hosting the private endpoint, the Microsoft Purview endpoint URL resolves to the private endpoint's IP address.

As an example, if a Microsoft Purview account name is 'Contoso-Purview', when it's resolved from outside the virtual network that hosts the private endpoint, it will be:

Expand table

Name	Type	Value
`Contoso-Purview.purview.azure.com`	CNAME	`Contoso-Purview.privatelink.purview.azure.com`
`Contoso-Purview.privatelink.purview.azure.com`	CNAME	<Microsoft Purview public endpoint>

Name	Type	Value
<Microsoft Purview public endpoint>	A	<Microsoft Purview public IP address>
`Web.purview.azure.com`	CNAME	<Microsoft Purview governance portal public endpoint>

The DNS resource records for Contoso-Purview, when resolved in the virtual network hosting the private endpoint, will be:

Expand table

Name	Type	Value
`Contoso-Purview.purview.azure.com`	CNAME	`Contoso-Purview.privatelink.purview.azure.com`
`Contoso-Purview.privatelink.purview.azure.com`	A	<Microsoft Purview account private endpoint IP address>
`Web.purview.azure.com`	CNAME	<Microsoft Purview portal private endpoint IP address>

Option 3 - Use your own DNS Servers

If you don't use DNS forwarders and instead you manage A records directly in your on-premises DNS servers to resolve the endpoints through their private IP addresses, you might need to create the following A records in your DNS servers.

> **Important**
>
> Update all names with `Contoso-Purview`, `ingestioneus2eastusksqky` and `atlas-12345678-1234-1234-abcd-123456789abc` with corresponding Azure resources name in your environment.
>
> - `Contoso-Purview` is the name of your Microsoft Purview account.
> - You only need the endpoints associated with the portal you're using:
>   - Classic Microsoft Purview governance portal (https://web.purview.azure.com): *account*, *portal*, and *ingestion* private endpoints are supported.
>   - New Microsoft Purview portal (https://purview.microsoft.com/): *platform* and *ingestion* private endpoints are supported.
> - If your account is configured for kafka notifications or was created before December 15, 2022, `atlas-12345678-1234-1234-abcd-123456789abc` is your Event Hubs namespace.
> - If your account was created **before** December 15, 2023, for `ingestioneus2eastusksqky` use the name of your Microsoft Purview managed storage account.
> - If your account was created **after** December 15, 2023 (or deployed using API version 2023-05-01-preview onwards), leave `ingestioneus2eastusksqky` as is.

Expand table

Name	Type	Value	Portal availability
`web.purview.azure.com`	A	<portal private endpoint IP address of Microsoft Purview>	Classic portal
`ingestioneus2eastusksqky.blob.core.windows.net`, `ingestioneus2eastusksqky.blob.storage.azure.net`	A	<blob-ingestion private endpoint IP address of Microsoft Purview>	Classic and new portals
`ingestioneus2eastusksqky.queue.core.windows.net`, `ingestioneus2eastusksqky.queue.storage.azure.net`	A	<queue-ingestion private endpoint IP address of Microsoft Purview>	Classic and new portals
`atlas-12345678-1234-1234-abcd-123456789abc.servicebus.windows.net`	A	<namespace-ingestion private endpoint IP address of Microsoft Purview>	Classic and new portals

Name	Type	Value	Portal availability
`Contoso-Purview.Purview.azure.com`	A	<account/platform private endpoint IP address of Microsoft Purview>	Classic and new portals
`Contoso-Purview.scan.Purview.azure.com`	A	<account/platform private endpoint IP address of Microsoft Purview>	Classic and new portals
`Contoso-Purview.catalog.Purview.azure.com`	A	<account/platform private endpoint IP address of Microsoft Purview>	Classic and new portals
`Contoso-Purview.proxy.purview.azure.com`	A	<account/platform private endpoint IP address of Microsoft Purview>	Classic and new portals
`Contoso-Purview.guardian.purview.azure.com`	A	<account/platform private endpoint IP address of Microsoft Purview>	Classic and new portals
`gateway.purview.azure.com`	A	<account/platform private endpoint IP address of Microsoft Purview>	Classic and new portals
`insight.prod.ext.web.purview.azure.com`	A	<portal private endpoint IP address of Microsoft Purview>	Classic portal
`manifest.prod.ext.web.purview.azure.com`	A	<portal private endpoint IP address of Microsoft Purview>	Classic portal
`cdn.prod.ext.web.purview.azure.com`	A	<portal private endpoint IP address of Microsoft Purview>	Classic portal
`hub.prod.ext.web.purview.azure.com`	A	<portal private endpoint IP address of Microsoft Purview>	Classic portal
`catalog.prod.ext.web.purview.azure.com`	A	<portal private endpoint IP address of Microsoft Purview>	Classic portal
`cseo.prod.ext.web.purview.azure.com`	A	<portal private endpoint IP address of Microsoft Purview>	Classic portal
`datascan.prod.ext.web.purview.azure.com`	A	<portal private endpoint IP address of Microsoft Purview>	Classic portal
`datashare.prod.ext.web.purview.azure.com`	A	<portal private endpoint IP address of Microsoft Purview>	Classic portal
`datasource.prod.ext.web.purview.azure.com`	A	<portal private endpoint IP address of Microsoft Purview>	Classic portal
`policy.prod.ext.web.purview.azure.com`	A	<portal private endpoint IP address of Microsoft Purview>	Classic portal
`sensitivity.prod.ext.web.purview.azure.com`	A	<portal private endpoint IP address of Microsoft Purview>	Classic portal
`web.privatelink.purviewstudio.azure.com`	A	<portal private endpoint IP address of Microsoft Purview>	Classic portal
`workflow.prod.ext.web.purview.azure.com`	A	<portal private endpoint IP address of Microsoft Purview>	Classic portal

Verify and DNS test name resolution and connectivity

1. If you're using Azure Private DNS Zones, make sure the following DNS Zones and the corresponding A records are created in your Azure Subscription:

Expand table

Private endpoint	Private endpoint associated to	Portal availability	DNS Zone (existing)	A Record (example)
Account	Microsoft Purview	Classic portal	`privatelink.purview.azure.com`	Contoso-Purview
Platform	Microsoft Purview	New portal	`privatelink.purview-service.microsoft.com`	Contoso-Purview
Portal	Microsoft Purview	Classic portal	`privatelink.purviewstudio.azure.com`	Web
Ingestion	Microsoft Purview ingestion - Blob*	Classic and new portals	`privatelink.blob.core.windows.net`, `privatelink.blob.storage.azure.net`	ingestioneus2eastusksqky
Ingestion	Microsoft Purview ingestion - Queue*	Classic and new portals	`privatelink.queue.core.windows.net`, `privatelink.queue.storage.azure.net`	ingestioneus2eastusksqky
Ingestion	Event Hubs**	Classic and new portals	`privatelink.servicebus.windows.net`	atlas-12345678-1234-1234-abcd-123456789abc

> **Note**
>
> *If your account was created before December 15, 2023, the endpoint is deployed to your managed storage account. If it was created after November 10 (or deployed using API version 2023-05-01-preview onwards), it points to the ingestion storage. **Your account only has an associated Event Hubs namespace if it is configured for kafka notifications or created before December 15, 2022.

2. Create Virtual network links in your Azure Private DNS Zones for your Azure Virtual Networks to allow internal name resolution.

3. From your management PC and self-hosted integration runtime VM, test name resolution and network connectivity to your Microsoft Purview account using tools such as Nslookup.exe and PowerShell

To test name resolution you need to resolve the following FQDNs through their private IP addresses: (Instead of Contoso-Purview, ingestioneus2eastusksqky or atlas-12345678-1234-1234-abcd-123456789abc, use the hostname associated with your purview account name and managed or configured resources names)

- `Contoso-Purview.purview.azure.com`
- `web.purview.azure.com`
- `ingestioneus2eastusksqky.blob.core.windows.net`
- `ingestioneus2eastusksqky.queue.core.windows.net`
- `atlas-12345678-1234-1234-abcd-123456789abc.servicebus.windows.net`

To test network connectivity, from self-hosted integration runtime VM you can launch PowerShell console and test connectivity using `Test-NetConnection`. You must resolve each endpoint by their private endpoint and obtain TcpTestSucceeded as True. (Instead of Contoso-Purview, ingestioneus2eastusksqky or atlas-12345678-1234-1234-abcd-123456789abc, use the hostname associated with your purview account name and managed or configured resources names)

- `Test-NetConnection -ComputerName Contoso-Purview.purview.azure.com -port 443`
- `Test-NetConnection -ComputerName web.purview.azure.com -port 443`
- `Test-NetConnection -ComputerName ingestioneus2eastusksqky.blob.core.windows.net -port 443`
- `Test-NetConnection -ComputerName ingestioneus2eastusksqky.queue.core.windows.net -port 443`

- `Test-NetConnection -ComputerName atlas-12345678-1234-1234-abcd-123456789abc.servicebus.windows.net -port 443`

The following example shows Microsoft Purview DNS name resolution from outside the virtual network or when an Azure private endpoint isn't configured.

```
C:\>nslookup.exe
Default Server:  UnKnown
Address:  168.63.129.16

> Contoso-Purview-01.purview.azure.com
Server:  UnKnown
Address:  168.63.129.16

Non-authoritative answer:
Name:    purview-gateway-prod.eastus.cloudapp.azure.com
Address:  52.226.150.205
Aliases:  Contoso-Purview-01.purview.azure.com
          contoso-purview-01.privatelink.purview.azure.com
          gateway-prod-eastus.babylon-service.azure.com

> web.purview.azure.com
Server:  UnKnown
Address:  168.63.129.16

Non-authoritative answer:
Name:    waws-prod-blu-655b4fde.cloudapp.net
Address:  40.117.171.112
Aliases:  web.purview.azure.com
          web.privatelink.purviewstudio.azure.com
          web.privatelink.purview.azure.com
          purviewstudio-prod-tm.trafficmanager.net
          purviewstudio-prod-wa-prod-shell-eastus.purviewstudio-prod-ase-eastus.p.azurewebsites.net
          waws-prod-blu-655b4fde.sip.p.azurewebsites.windows.net
```

The following example shows Microsoft Purview DNS name resolution from inside the virtual network.

```
C:\>nslookup.exe
Default Server:  UnKnown
Address:  168.63.129.16

> Contoso-Purview-01.purview.azure.com
Server:  UnKnown
Address:  168.63.129.16

Non-authoritative answer:
Name:    contoso-purview-01.privatelink.purview.azure.com
Address:  10.40.0.4
Aliases:  Contoso-Purview-01.purview.azure.com

> web.purview.azure.com
Server:  UnKnown
Address:  168.63.129.16

Non-authoritative answer:
Name:    web.privatelink.purviewstudio.azure.com
Address:  10.40.0.5
Aliases:  web.purview.azure.com
```

> **Note**
>
> The values in these images are examples. Use the information in the article to correctly configure your DNS Zones.

Next steps

- Troubleshooting private endpoint configuration for your Microsoft Purview account
- Manage data sources in Microsoft Purview

Troubleshooting private endpoint configuration for Microsoft Purview accounts

Article • 05/09/2025

> **Important**
>
> This article covers private endpoints for the classic Microsoft Purview governance portal (https://web.purview.azure.com). If you're using the new Microsoft Purview portal (https://purview.microsoft.com/), follow the documentation for private endpoints in the Microsoft Purview portal.

This article summarizes known limitations related to using private endpoints for Microsoft Purview and provides a list of steps and solutions for troubleshooting some of the most common relevant issues.

Known limitations

- We currently don't support ingestion private endpoints that work with your AWS sources.
- Scanning Azure Multiple Sources using self-hosted integration runtime isn't supported.
- Using Azure integration runtime to scan data sources behind private endpoint isn't supported.
- The ingestion private endpoints can be created via the classic Microsoft Purview governance portal experience described in the steps here. They can't be created from the Private Link Center.
- Creating a DNS record for ingestion private endpoints inside existing Azure DNS Zones, while the Azure Private DNS Zones are located in a different subscription than the private endpoints isn't supported via the classic Microsoft Purview governance portal experience. A record can be added manually in the destination DNS Zones in the other subscription.
- If you configure your own Event Hubs namespace or enable a managed Event Hubs namespace after deploying an ingestion private endpoint, you'll need to redeploy the ingestion private endpoint.
- Self-hosted integration runtime machine must be deployed in the same virtual network or a peered virtual network where Microsoft Purview account and ingestion private endpoints are deployed.
- For limitation related to Private Link service, see Azure Private Link limits.
- Currently, Microsoft Purview instances using the new Microsoft Purview portal can only use ingestion private endpoints.

Recommended troubleshooting steps

1. Once you deploy private endpoints for your Microsoft Purview account, review your Azure environment to make sure private endpoint resources are deployed successfully. Depending on your scenario, one or more of the following Azure private endpoints must be deployed in your Azure subscription:

Expand table

Private endpoint	Private endpoint assigned to	Example
Account	Microsoft Purview Account	mypurview-private-account
Portal	Microsoft Purview Account	mypurview-private-portal
Ingestion	Storage (Blob)*	mypurview-ingestion-blob
Ingestion	Storage (Queue)*	mypurview-ingestion-queue
Ingestion	Event Hubs Namespace**	mypurview-ingestion-namespace

> **Note**
>
> *If your account was created before December 15, 2023, the endpoint is deployed to your managed storage account. If it was created after November 10 (or deployed using API version 2023-05-01-preview onwards), it points to the ingestion storage.
>
> **Your account only has an associated Event Hubs namespace if it is configured for kafka notifications or created before December 15, 2022.

2. If portal private endpoint is deployed, make sure you also deploy account private endpoint.

3. If portal private endpoint is deployed, and public network access is set to deny in your Microsoft Purview account, make sure you launch the classic Microsoft Purview governance portal from internal network.

    - To verify the correct name resolution, you can use a **NSlookup.exe** command line tool to query `web.purview.azure.com`. The result must return a private IP address that belongs to portal private endpoint.
    - To verify network connectivity, you can use any network test tools to test outbound connectivity to `web.purview.azure.com` endpoint to port **443**. The connection must be successful.

4. If Azure Private DNS Zones are used, make sure the required Azure DNS Zones are deployed and there's DNS (A) record for each private endpoint.

5. Test network connectivity and name resolution from management machine to Microsoft Purview endpoint and purview web url. If account and portal private endpoints are deployed, the endpoints must be resolved through private IP addresses.

   PowerShell

   ```
   Test-NetConnection -ComputerName web.purview.azure.com -Port 443
   ```

   Example of successful outbound connection through private IP address:

   ```
   ComputerName     : web.purview.azure.com
   RemoteAddress    : 10.9.1.7
   RemotePort       : 443
   InterfaceAlias   : Ethernet 2
   SourceAddress    : 10.9.0.10
   TcpTestSucceeded : True
   ```

   PowerShell

   ```
   Test-NetConnection -ComputerName purview-test01.purview.azure.com -Port 443
   ```

   Example of successful outbound connection through private IP address:

   ```
   ComputerName     : purview-test01.purview.azure.com
   RemoteAddress    : 10.9.1.8
   RemotePort       : 443
   InterfaceAlias   : Ethernet 2
   SourceAddress    : 10.9.0.10
   TcpTestSucceeded : True
   ```

6. If you've created your Microsoft Purview account after 18 August 2021, make sure you download and install the latest version of self-hosted integration runtime from Microsoft download center.

7. From self-hosted integration runtime VM, test network connectivity and name resolution to Microsoft Purview endpoint.

8. From self-hosted integration runtime, test network connectivity and name resolution to Microsoft Purview managed resources such as blob queue, and secondary resources like Event Hubs through port 443 and private IP addresses. (If either are needed, replace the managed storage account and Event Hubs namespace with corresponding resource names).

```
PowerShell

Test-NetConnection -ComputerName
`scansoutdeastasiaocvseab`.blob.core.windows.net -Port 443
```

Example of successful outbound connection to managed blob storage through private IP address:

```

ComputerName     : scansoutdeastasiaocvseab.blob.core.windows.net
RemoteAddress    : 10.15.1.6
RemotePort       : 443
InterfaceAlias   : Ethernet 2
SourceAddress    : 10.15.0.4
TcpTestSucceeded : True
```

```
PowerShell

Test-NetConnection -ComputerName
`scansoutdeastasiaocvseab`.queue.core.windows.net -Port 443
```

Example of successful outbound connection to managed queue storage through private IP address:

```

ComputerName     : scansoutdeastasiaocvseab.blob.core.windows.net
RemoteAddress    : 10.15.1.5
RemotePort       : 443
InterfaceAlias   : Ethernet 2
SourceAddress    : 10.15.0.4
TcpTestSucceeded : True
```

```
PowerShell

Test-NetConnection -ComputerName `Atlas-1225cae9-d651-4039-86a0-
b43231a17a4b`.servicebus.windows.net -Port 443
```

Example of successful outbound connection to Event Hubs namespace through private IP address:

```
ComputerName     : Atlas-1225cae9-d651-4039-86a0-
b43231a17a4b.servicebus.windows.net
RemoteAddress    : 10.15.1.4
RemotePort       : 443
InterfaceAlias   : Ethernet 2
SourceAddress    : 10.15.0.4
TcpTestSucceeded : True
```

9. From the network where data source is located, test network connectivity and name resolution to Microsoft Purview endpoint and managed or configured resources endpoints.

10. If data sources are located in on-premises network, review your DNS forwarder configuration. Test name resolution from within the same network where data sources are located to self-hosted integration runtime, Microsoft Purview endpoints and managed or configured resources. It's expected to obtain a valid private IP address from DNS query for each endpoint.

    For more information, see Virtual network workloads without custom DNS server and On-premises workloads using a DNS forwarder scenarios in Azure Private Endpoint DNS configuration.

11. If management machine and self-hosted integration runtime VMs are deployed in on-premises network and you have set up DNS forwarder in your environment, verify DNS and network settings in your environment.

12. If ingestion private endpoint is used, make sure self-hosted integration runtime is registered successfully inside Microsoft Purview account and shows as running both inside the self-hosted integration runtime VM and in the classic Microsoft Purview governance portal .

Common errors and messages

Issue

You could receive the following error message when running a scan:

`Internal system error. Please contact support with correlationId:xxxxxxxx-xxxx-xxxx-xxxx-xxxxxxxxxxxx System Error, contact support.`

Cause

This can be an indication of issues related to connectivity or name resolution between the VM running self-hosted integration runtime and Microsoft Purview's managed storage account or configured Event Hubs.

Resolution

Validate if name resolution is successful between the VM running the Self-Hosted Integration Runtime and the Microsoft Purview managed blob queue or configured Event Hubs through port 443 and private IP addresses (step 8 above.)

Issue

You could receive the following error message when running a new scan:

`message: Unable to setup config overrides for this scan. Exception:'Type=Microsoft.WindowsAzure.Storage.StorageException,Message=The remote server returned an error: (404) Not Found.,Source=Microsoft.WindowsAzure.Storage,StackTrace= at Microsoft.WindowsAzure.Storage.Core.Executor.Executor.EndExecuteAsync[T](IAsyncResult result)`

Cause

This can be an indication of running an older version of self-hosted integration runtime. You'll need to use the self-hosted integration runtime version 5.9.7885.3 or greater.

Resolution

Upgrade self-hosted integration runtime to 5.9.7885.3.

Issue

Microsoft Purview account with private endpoint deployment failed with Azure Policy validation error during the deployment.

Cause

This error suggests that there might be an existing Azure Policy Assignment on your Azure subscription that is preventing the deployment of any of the required Azure resources.

Resolution

Review your existing Azure Policy Assignments and make sure deployment of the following Azure resources are allowed in your Azure subscription.

> ⓘ **Note**
>
> Depending on your scenario, you may need to deploy one or more of the following Azure resource types:
>
> - Microsoft Purview (Microsoft.Purview/Accounts)
> - Private Endpoint (Microsoft.Network/privateEndpoints)
> - Private DNS Zones (Microsoft.Network/privateDnsZones)
> - Event Hub Name Space (Microsoft.EventHub/namespaces)
> - Storage Account (Microsoft.Storage/storageAccounts)

Issue

Not authorized to access this Microsoft Purview account. This Microsoft Purview account is behind a private endpoint. Access the account from a client in the same virtual network (virtual network) that has been configured for the Microsoft Purview account's private endpoint.

Cause

User is trying to connect to Microsoft Purview from a public endpoint or using Microsoft Purview public endpoints where **Public network access** is set to **Deny**.

Resolution

In this case, to open the classic Microsoft Purview governance portal, either use a machine that is deployed in the same virtual network as the classic Microsoft Purview governance portal private endpoint or use a VM that is connected to your CorpNet in which hybrid connectivity is allowed.

Issue

You could receive the following error message when scanning a SQL server, using a self-hosted integration runtime:

```
Message=This implementation is not part of the Windows Platform FIPS validated cryptographic algorithms
```

Cause

Self-hosted integration runtime machine has enabled the FIPS mode. Federal Information Processing Standards (FIPS) defines a certain set of cryptographic algorithms that are allowed to be used. When FIPS mode is enabled on the machine, some cryptographic classes that the invoked processes depend on are blocked in some scenarios.

Resolution

Disable FIPS mode on self-hosted integration server.

Next steps

If your problem isn't listed in this article or you can't resolve it, get support by visiting one of the following channels:

- Get answers from experts through Microsoft Q&A.
- Connect with @AzureSupport. This official Microsoft Azure resource on Twitter helps improve the customer experience by connecting the Azure community to the right answers, support, and experts.
- If you still need help, go to the Azure support site and select **Submit a support request**.

FAQ about Microsoft Purview private endpoints and managed virtual networks

Article • 05/09/2025

> **Important**
>
> This article covers private endpoints for the classic Microsoft Purview governance portal (https://web.purview.azure.com). If you're using the new Microsoft Purview portal (https://purview.microsoft.com/), follow the documentation for private endpoints in the Microsoft Purview portal.

This article answers common questions that customers and field teams often ask about Microsoft Purview network configurations by using Azure Private Link or Microsoft Purview Managed VNets. It's intended to clarify questions about Microsoft Purview firewall settings, private endpoints, DNS configuration, and related configurations.

To set up Microsoft Purview by using Private Link, see Use private endpoints for your Microsoft Purview account. To configure Managed VNets for a Microsoft Purview account, see Use a Managed virtual network with your Microsoft Purview account.

When should I use a self-hosted integration runtime, Managed virtual network IR, or Azure IR?

Learn more from Choose the right integration runtime configuration for your scenario.

Can I use both self-hosted integration runtime and Managed virtual network IR inside a Microsoft Purview account?

Yes. You can use one or all of the runtime options in a single Microsoft Purview account: Azure IR, Managed virtual network IR, and self-hosted integration runtime. You can use only one runtime option in a single scan.

What's the purpose of deploying the Microsoft Purview account private endpoint?

The Microsoft Purview account private endpoint is used to add another layer of security by enabling scenarios where only client calls that originate from within the virtual network are allowed to access the account. This private endpoint is also a prerequisite for the portal private endpoint.

What's the purpose of deploying the Microsoft Purview portal private endpoint?

The Microsoft Purview portal private endpoint provides private connectivity to the classic Microsoft Purview governance portal.

What's the purpose of deploying the Microsoft Purview ingestion private endpoints?

Microsoft Purview can scan data sources in Azure or an on-premises environment by using ingestion private endpoints. Three other private endpoint resources are deployed and linked to Microsoft Purview managed or configured resources when ingestion private endpoints are created:

- If you're using a managed Event Hubs for kafka notifications, **namespace** is linked to a Microsoft Purview configured Event Hubs namespace.
- If your account was created before December 15, 2023:
  - **Blob** is linked to a Microsoft Purview managed storage account.
  - **Queue** is linked to a Microsoft Purview managed storage account.
- If your account was created after December 15, 2023 (or deployed using API version 2023-05-01-preview onwards):
  - **Blob** is linked to a Microsoft Purview ingestion storage.
  - **Queue** is linked to a Microsoft Purview ingestion storage.

Can I scan a data source through a public endpoint if a private endpoint is enabled on my Microsoft Purview account?

Yes. Data sources that aren't connected through a private endpoint can be scanned by using a public endpoint while Microsoft Purview is configured to use a private endpoint.

Can I scan a data source through a service endpoint if a private endpoint is enabled?

Yes. Data sources that aren't connected through a private endpoint can be scanned by using a service endpoint while Microsoft Purview is configured to use a private endpoint. Learn more from Choose the right integration runtime configuration for your scenario.

Can I access the classic Microsoft Purview governance portal from a public network if Public network access is set to Deny in Microsoft Purview account networking?

No. Connecting to Microsoft Purview from a public endpoint where **Public network access** is set to **Deny** results in the following error message:

"Not authorized to access this Microsoft Purview account. This Microsoft Purview account is behind a private endpoint. Access the account from a client in the same virtual network (virtual network) that has been configured for the Microsoft Purview account's private endpoint."

In this case, to open the classic Microsoft Purview governance portal, either use a machine that's deployed in the same virtual network as the Microsoft Purview portal private endpoint or use a VM that's connected to your CorpNet in which hybrid connectivity is allowed.

Is it possible to restrict access to the Microsoft Purview managed storage account or ingestion storage account and Event Hubs namespace (for private endpoint ingestion only) but keep portal access enabled for users across the web?

> **Note**
>
> Your account only has a managed storage account if it was created before December 15, 2023 (or deployed using API version previous to 2023-05-01-preview). Your account only has an associated Event Hubs namespace if it is configured for kafka notifications or was created before December 15, 2022.

Yes. You can configure Microsoft Purview firewall setting to Disabled for ingestion only (Preview). By choosing this option, public network access to your Microsoft Purview account through API and classic Microsoft Purview governance portal is allowed, however public network access is set to disabled on your Microsoft Purview account's Managed storage account. You also need to confirm your Event Hubs network settings allow communication.

If public network access is set to Allow, does it mean the managed storage account or ingestion storage account and Event Hubs namespace are accessible by anyone?

> **Note**
>
> Your account only has a managed storage account if it was created before December 15, 2023 (or deployed using API version previous to 2023-05-01-preview). Your account only has an associated Event Hubs namespace if it is **configured for kafka notifications** or was created before December 15, 2022.

No. As protected resources, access to the Microsoft Purview managed storage account and any Event Hubs namespace is restricted to Microsoft Purview only using RBAC authentication schemes. These resources are deployed with a deny assignment to all principals, which prevents any applications, users, or groups from gaining access to them.

To read more about Azure deny assignment, see Understand Azure deny assignments.

What private DNS zones are required for Microsoft Purview for a private endpoint?

For Microsoft Purview *account*, *portal*, and *platform* private endpoints:

- `privatelink.purview.azure.com` - for the classic Microsoft Purview governance portal.
- `privatelink.purview-service.microsoft.com` - for the Microsoft Purview portal.

For Microsoft Purview *ingestion* private endpoints:

- `privatelink.blob.core.windows.net`
- `privatelink.queue.core.windows.net`
- `privatelink.servicebus.windows.net`

Do I have to use a dedicated virtual network and dedicated subnet when I deploy Microsoft Purview private endpoints?

No. However, `PrivateEndpointNetworkPolicies` must be disabled in the destination subnet before you deploy the private endpoints. Consider deploying Microsoft Purview into a virtual network that has network connectivity to data source virtual networks through virtual network Peering and access to an on-premises network if you plan to scan data sources cross-premises.

Read more about Disable network policies for private endpoints.

Can I deploy Microsoft Purview private endpoints and use existing private DNS zones in my subscription to register the A records?

Yes. Your private endpoint DNS zones can be centralized in a hub or data management subscription for all internal DNS zones required for Microsoft Purview and all data source records. We recommend this method to allow Microsoft Purview to resolve data sources by using their private endpoint internal IP addresses.

You're also required to set up a virtual network link for virtual networks for the existing private DNS zone.

What are the outbound ports and firewall requirements for virtual machines with self-hosted integration runtime for Microsoft Purview when you use a private endpoint?

The VMs in which self-hosted integration runtime is deployed must have outbound access to Azure endpoints and a Microsoft Purview private IP address through port 443.

Do I need to enable outbound internet access from the virtual machine running self-hosted integration runtime if a private endpoint is enabled?

No. However, it's expected that the virtual machine running self-hosted integration runtime can connect to your instance of Microsoft Purview through an internal IP address by using port 443. Use common troubleshooting tools for name resolution and connectivity testing, such as nslookup.exe and Test-NetConnection.

Do I still need to deploy private endpoints for my Microsoft Purview account if I'm using Managed virtual network?

At least one account and portal private endpoints are required, if public access in Microsoft Purview account is set to **deny**. At least one account, portal and ingestion private endpoint are required, if public access in Microsoft Purview account is set to **deny** and you're planning to scan more data sources using a self-hosted integration runtime.

What inbound and outbound communications are allowed through public endpoint for Microsoft Purview Managed VNets?

No inbound communication is allowed into a Managed virtual network from public network. All ports are opened for outbound communications. In Microsoft Purview, a Managed virtual network can be used to privately connect to Azure data sources to extract metadata during scan.

Why do I receive the following error message when I try to launch the classic Microsoft Purview governance portal from my machine?

"This Microsoft Purview account is behind a private endpoint. Access the account from a client in the same virtual network (virtual network) that has been configured for the Microsoft Purview account's private endpoint."

It's likely your Microsoft Purview account is deployed by using Private Link and public access is disabled on your Microsoft Purview account. As a result, you have to browse the clssic Microsoft Purview governance portal from a virtual machine that has internal network connectivity to Microsoft Purview.

If you're connecting from a VM behind a hybrid network or using a jump machine connected to your virtual network, use common troubleshooting tools for name resolution and connectivity testing, such as nslookup.exe and Test-NetConnection.

1. Validate if you can resolve the following addresses through your Microsoft Purview account's private IP addresses.

   - `Web.Purview.Azure.com`
   - `<YourPurviewAccountName>.Purview.Azure.com`

2. Verify network connectivity to your Microsoft Purview account by using the following PowerShell command:

   PowerShell

   ```
   Test-NetConnection -ComputerName <YourPurviewAccountName>.Purview.Azure.com -Port 443
   ```

3. Verify your cross-premises DNS configuration if you use your own DNS resolution infrastructure.

For more information about DNS settings for private endpoints, see Azure private endpoint DNS configuration.

Can I move private endpoints associated with Microsoft Purview account or its managed resources to another Azure subscription or resource group?

No. Move operations for Account, Portal, or Ingestion private endpoints aren't supported. For more information, see Move networking resources to new resource group or subscription.

Can I create multiple managed virtual networks in different regions?

Yes. You can create multiple managed virtual networks across different regions in a single Microsoft Purview instance so you can access data sources available in different regions. This feature provides the ability to:

- Create multiple managed virtual networks (five maximum) across different regions within a single Microsoft Purview instance.
- Network isolation within your own organization to address potential data residency or scan performance concerns.

Next steps

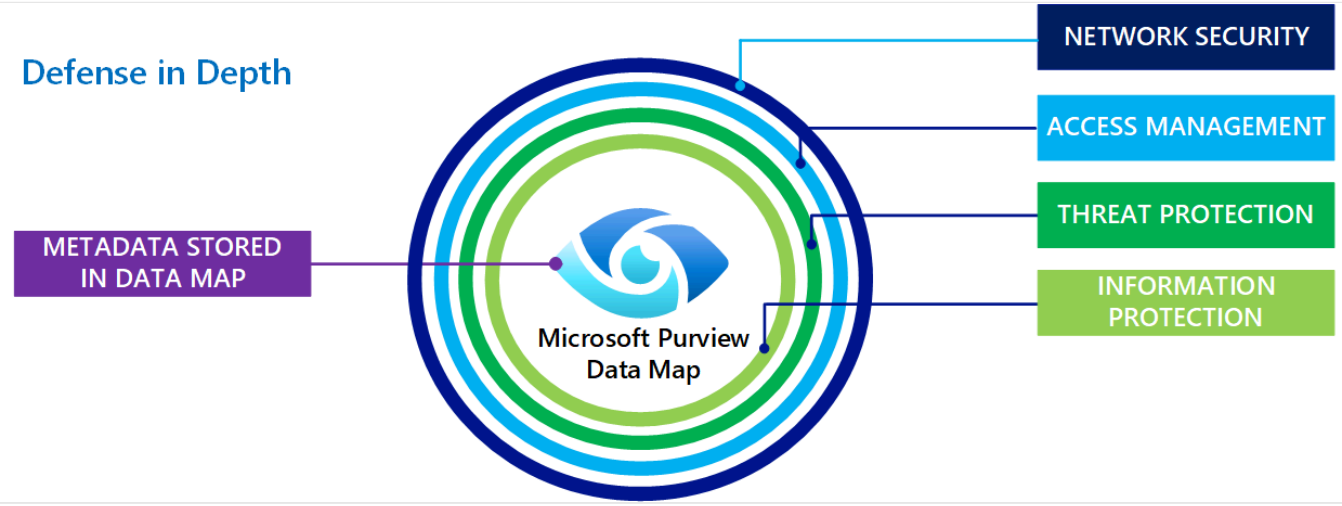
To set up Microsoft Purview by using Private Link, see Use private endpoints for your Microsoft Purview account.

Data governance best practices for security

Article • 05/20/2025

This article provides best practices for common security requirements for classic Microsoft Purview governance solutions. The security strategy described follows the layered defense-in-depth approach.

> **Note**
>
> These best practices cover security for Microsoft Purview data governance solutions. Learn more about Microsoft Purview in general, or about risk and compliance solutions specifically.

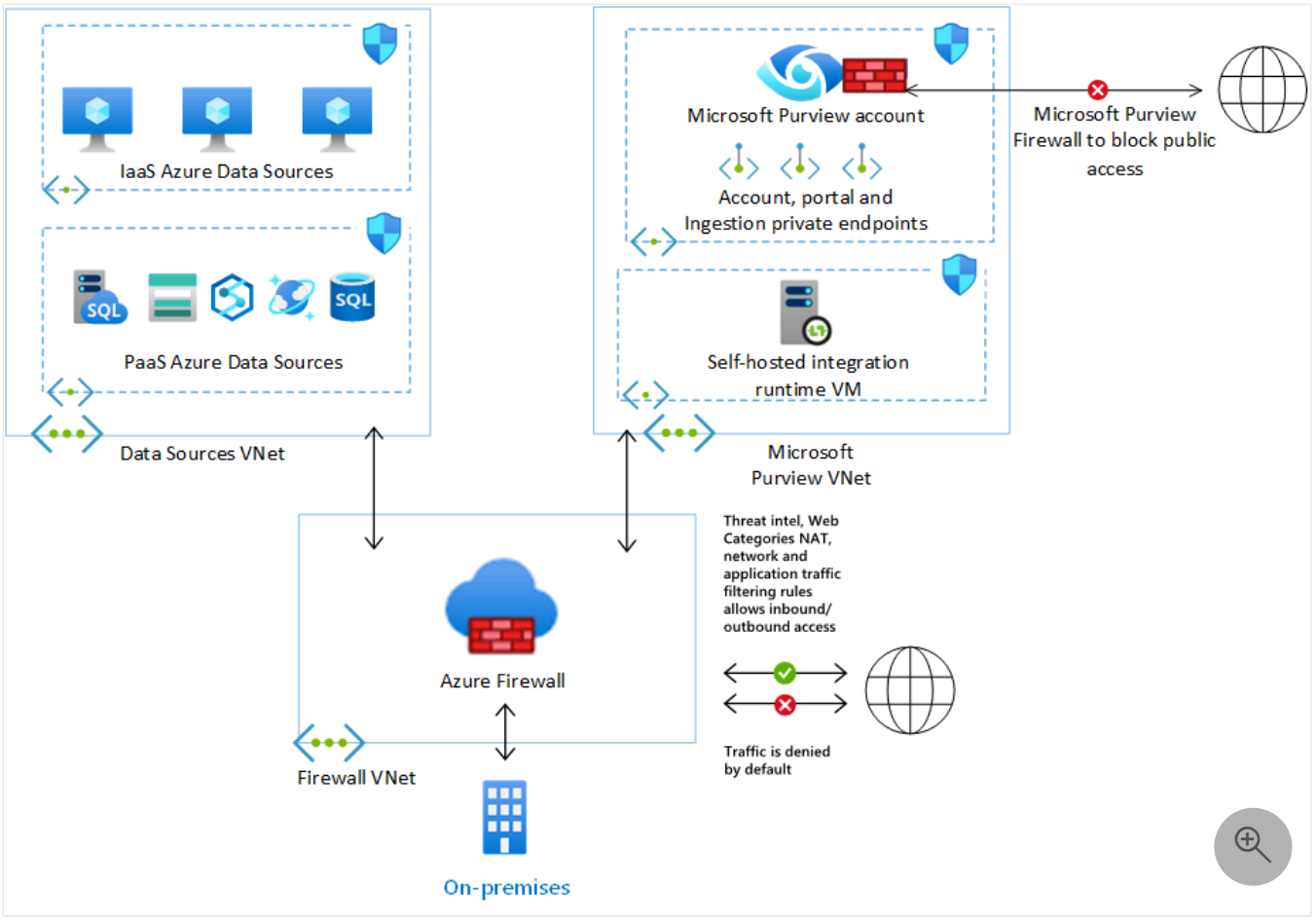


Before applying these recommendations to your environment, you should consult your security team as some may not be applicable to your security requirements.

Network security

You can enable the following network security capabilities for the Microsoft Purview data map:

- Enable end-to-end network isolation using Private Link Service.
- Use Microsoft Purview Firewall to disable Public access.
- Deploy Network Security Group (NSG) rules for subnets where Azure data sources private endpoints, Microsoft Purview private endpoints and self-hosted runtime VMs are deployed.
- Implement Microsoft Purview with private endpoints managed by a Network Virtual Appliance, such as Azure Firewall for network inspection and network filtering.



For more information, see Best practices related to connectivity to Azure PaaS Services.

Deploy private endpoints for Microsoft Purview accounts

If you need to use Microsoft Purview from inside your private network, it's recommended to use Azure Private Link Service with your Microsoft Purview accounts for partial or end-to-end isolation to connect to Microsoft Purview governance portal, access Microsoft Purview endpoints and to scan data sources.

The Microsoft Purview *account* private endpoint is used to add another layer of security, so only client calls that are originated from within the virtual network are allowed to access the Microsoft Purview account. This private endpoint is also a prerequisite for the portal private endpoint.

The Microsoft Purview *portal* private endpoint is required to enable connectivity to Microsoft Purview governance portal using a private network.

Microsoft Purview can scan data sources in Azure or an on-premises environment by using ingestion private endpoints.

For more information, see Microsoft Purview network architecture and best practices.

Block public access using Microsoft Purview firewall

You can disable Microsoft Purview Public access to cut off access to the Microsoft Purview account completely from the public internet. In this case, you should consider the following requirements:

- Microsoft Purview must be deployed based on end-to-end network isolation scenario.
- To access Microsoft Purview governance portal and Microsoft Purview endpoints, you need to use a management machine that is connected to private network to access Microsoft Purview through private network.
- Review known limitations.

For more information, see Firewalls to restrict public access.

Use Network Security Groups

You can use an Azure network security group to filter network traffic to and from Azure resources in an Azure virtual network. A network security group contains security rules that allow or deny inbound network traffic to, or outbound network traffic from, several types of Azure resources. For each rule, you can specify source and destination, port, and protocol.

Network Security Groups can be applied to network interface or Azure virtual networks subnets, where Microsoft Purview private endpoints, self-hosted integration runtime VMs and Azure data sources are deployed.

For more information, see apply NSG rules for private endpoints.

The following NSG rules are required on **data sources** for Microsoft Purview scanning:

Expand table

Direction	Source	Source port range	Destination	Destination port	Protocol	Action
Inbound	Self-hosted integration runtime VMs' private IP addresses or subnets	*	Data Sources private IP addresses or Subnets	443	Any	Allow

The following NSG rules are required on from the **management machines** to access Microsoft Purview governance portal:

Expand table

Direction	Source	Source port range	Destination	Destination port	Protocol	Action
Outbound	Management machines' private IP addresses or subnets	*	Microsoft Purview account and portal private endpoint IP addresses or subnets	443	Any	Allow
Outbound	Management machines' private IP addresses or subnets	*	Service tag: `AzureCloud`	443	Any	Allow

The following NSG rules are required on **self-hosted integration runtime VMs** for Microsoft Purview scanning and metadata ingestion:

> **Important**
>
> Consider adding additional rules with relevant Service Tags, based on your data source types.

Expand table

Direction	Source	Source port range	Destination	Destination port	Protocol	Action
Outbound	Self-hosted integration runtime VMs' private IP addresses or subnets	*	Data Sources private IP addresses or subnets	443	Any	Allow
Outbound	Self-hosted integration runtime VMs' private IP addresses or subnets	*	Microsoft Purview account and ingestion private endpoint IP addresses or Subnets	443	Any	Allow
Outbound	Self-hosted integration runtime VMs' private IP	*	Service tag: `Servicebus`	443	Any	Allow

Direction	Source	Source port range	Destination	Destination port	Protocol	Action
	addresses or subnets					
Outbound	Self-hosted integration runtime VMs' private IP addresses or subnets	*	Service tag: `Storage`	443	Any	Allow
Outbound	Self-hosted integration runtime VMs' private IP addresses or subnets	*	Service tag: `AzureActiveDirectory`	443	Any	Allow
Outbound	Self-hosted integration runtime VMs' private IP addresses or subnets	*	Service tag: `DataFactory`	443	Any	Allow
Outbound	Self-hosted integration runtime VMs' private IP addresses or subnets	*	Service tag: `KeyVault`	443	Any	Allow

The following NSG rules are required on for **Microsoft Purview account, portal and ingestion private endpoints**:

Expand table

Direction	Source	Source port range	Destination	Destination port	Protocol	Action
Inbound	Self-hosted integration runtime VMs' private IP addresses or subnets	*	Microsoft Purview account and ingestion private endpoint IP addresses or subnets	443	Any	Allow

Direction	Source	Source port range	Destination	Destination port	Protocol	Action
Inbound	Management machines' private IP addresses or subnets	*	Microsoft Purview account and ingestion private endpoint IP addresses or subnets	443	Any	Allow

For more information, see Self-hosted integration runtime networking requirements.

Access management

Identity and Access Management provides the basis of a large percentage of security assurance. It enables access based on identity authentication and authorization controls in cloud services. These controls protect data and resources and decide which requests should be permitted.

Related to roles and access management in Microsoft Purview, you can apply the following security best practices:

- Define roles and responsibilities to manage Microsoft Purview in control plane and data plane:
  - Define roles and tasks required to deploy and manage Microsoft Purview inside an Azure subscription.
  - Define roles and task needed to perform data management and governance using Microsoft Purview.
- Assign roles to Microsoft Entra groups instead of assigning roles to individual users.
- Use Azure Active Directory Entitlement Management to map user access to Microsoft Entra groups using Access Packages.
- Enforce multifactor authentication for Microsoft Purview users, especially, for users with privileged roles such as collection admins, data source admins or data curators.

Manage a Microsoft Purview account in control plane and data plane

Control plane refers to all operations related to Azure deployment and management of Microsoft Purview inside Azure Resource Manager.

Data plane refers to all operations, related to interacting with Microsoft Purview inside Data Map and Unified Catalog.

You can assign control plane and data plane roles to users, security groups and service principals from your Microsoft Entra tenant that is associated to Microsoft Purview instance's Azure subscription.

Examples of control plane operations and data plane operations:

Expand table

Task	Scope	Recommended role	What roles to use?
Deploy a Microsoft Purview account	Control plane	Azure subscription owner or contributor	Azure RBAC roles
Set up a Private Endpoint for Microsoft Purview	Control plane	Contributor	Azure RBAC roles
Delete a Microsoft Purview account	Control plane	Contributor	Azure RBAC roles
Add or manage a self-hosted integration runtime (SHIR)	Control plane	Data source administrator	Microsoft Purview roles
View Microsoft Purview metrics to get current capacity units	Control plane	Reader	Azure RBAC roles
Create a collection	Data plane	Collection Admin	Microsoft Purview roles
Register a data source	Data plane	Collection Admin	Microsoft Purview roles
Scan a SQL Server	Data plane	Data source admin and data reader or data curator	Microsoft Purview roles
Search inside Microsoft Purview Unified Catalog	Data plane	Data source admin and data reader or data curator	Microsoft Purview roles

Microsoft Purview plane roles are defined and managed inside Microsoft Purview instance in Microsoft Purview collections. For more information, see Access control in Microsoft Purview.

Follow Microsoft Purview Data Governance roles for guidance on users' roles and permissions to manage in Unified Catalog.

Follow Azure role-based access recommendations for Azure control plane tasks.

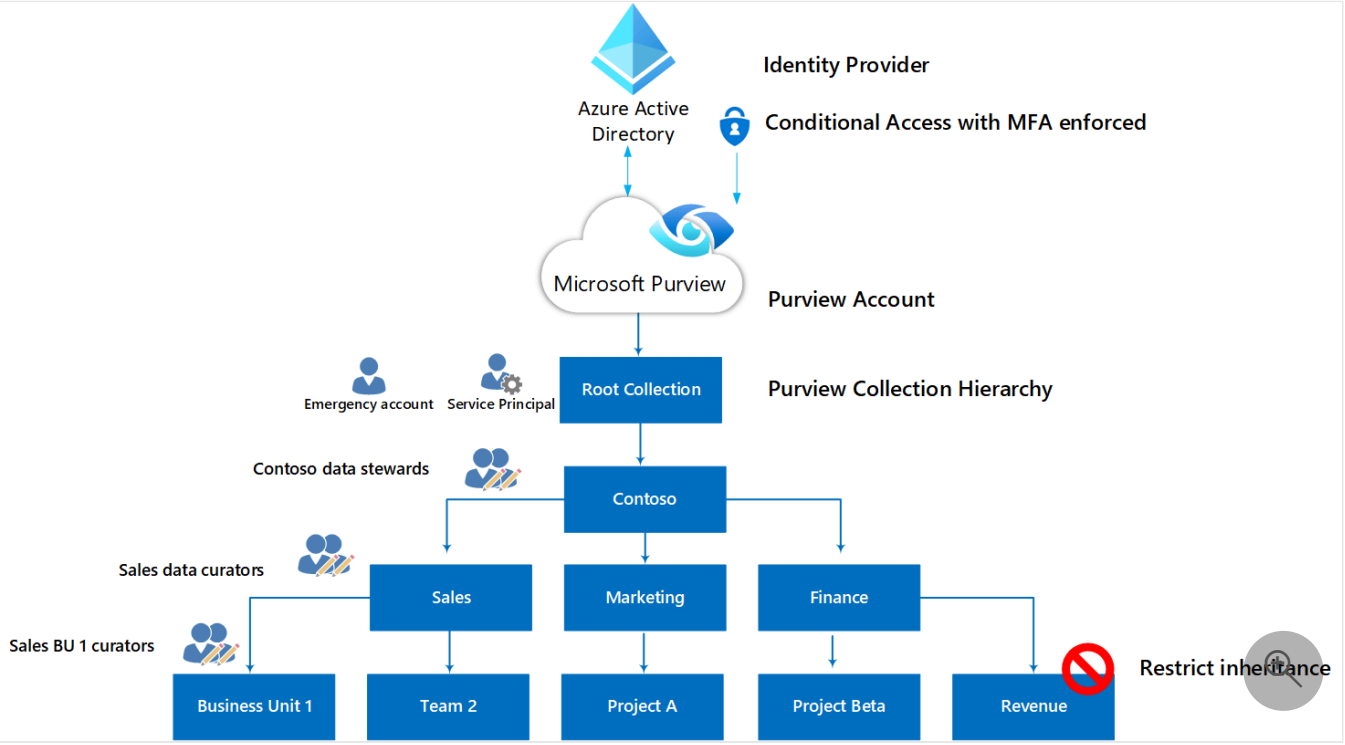
Authentication and authorization

To gain access to Microsoft Purview, users must be authenticated and authorized. Authentication is the process of proving the user is who they claim to be. Authorization refers to controlling access inside Microsoft Purview assigned on collections.

We use Microsoft Entra ID to provide authentication and authorization mechanisms for Microsoft Purview inside Collections. You can assign Microsoft Purview roles to the following security principals from your Microsoft Entra tenant that is associated with Azure subscription where your Microsoft Purview instance is hosted:

- Users and guest users (if they're already added into your Microsoft Entra tenant)
- Security groups
- Managed Identities
- Service Principals

Microsoft Purview fine-grained roles can be assigned to a flexible Collections hierarchy inside the Microsoft Purview instance.



Define Least Privilege model

As a general rule, restrict access based on the need to know and least privilege security principles is imperative for organizations that want to enforce security policies for data access.

In Microsoft Purview, data sources, assets and scans can be organized using Microsoft Purview Collections. Collections are hierarchical grouping of metadata in Microsoft Purview, but at the same time they provide a mechanism to manage access across Microsoft Purview. Roles in Microsoft Purview can be assigned to a collection based on your collection's hierarchy.

Use Microsoft Purview collections to implement your organization's metadata hierarchy for centralized or delegated management and governance hierarchy based on least privileged model.

Follow least privilege access model when assigning roles inside Microsoft Purview collections by segregating duties within your team and grant only the amount of access to users that they need to perform their jobs.

For more information how to assign least privilege access model in Microsoft Purview, based on Microsoft Purview collection hierarchy, see Access control in Microsoft Purview.

Lower exposure of privileged accounts

Securing privileged access is a critical first step to protecting business assets. Minimizing the number of people who have access to secure information or resources, reduces the chance of a malicious user getting access, or an authorized user inadvertently affecting a sensitive resource.

Reduce the number of users with write access inside your Microsoft Purview instance. Keep the number of collection admins and data curator roles minimum at root collection.

Use multifactor authentication and conditional access

Microsoft Entra multifactor authentication provides another layer of security and authentication. For more security, we recommend enforcing conditional access policies for all privileged accounts.

By using Microsoft Entra Conditional Access policies, apply Microsoft Entra multifactor authentication at sign-in for all individual users who are assigned to Microsoft Purview roles with modify access inside your Microsoft Purview instances: Collection Admin, Data Source Admin, Data Curator.

Enable multifactor authentication for your admin accounts and ensure that admin account users have registered for MFA.

You can define your Conditional Access policies by selecting Microsoft Purview as a Cloud App.

Prevent accidental deletion of Microsoft Purview accounts

In Azure, you can apply resource locks to an Azure subscription, a resource group, or a resource to prevent accidental deletion or modification for critical resources.

Enable Azure resource lock for your Microsoft Purview accounts to prevent accidental deletion of Microsoft Purview instances in your Azure subscriptions.

Adding a `CanNotDelete` or `ReadOnly` lock to Microsoft Purview account doesn't prevent deletion or modification operations inside Microsoft Purview data plane, however, it prevents any operations in control plane, such as deleting the Microsoft Purview account, deploying a private endpoint or configuration of diagnostic settings.

For more information, see Understand scope of locks.

Resource locks can be assigned to Microsoft Purview resource groups or resources, however, you can't assign an Azure resource lock to Microsoft Purview Managed resources or managed Resource Group.

Implement a break glass strategy

Plan for a break glass strategy for your Microsoft Entra tenant, Azure subscription and Microsoft Purview accounts to prevent tenant-wide account lockout.

For more information about Microsoft Entra ID and Azure emergency access planning, see Manage emergency access accounts in Microsoft Entra ID.

For more information about Microsoft Purview break glass strategy, see Microsoft Purview collections best practices and design recommendations.

Threat protection and preventing data exfiltration

Microsoft Purview provides rich insights into the sensitivity of your data, which makes it valuable to security teams using Microsoft Defender for Cloud to manage the organization's security posture and protect against threats to their workloads. Data resources remain a popular target for malicious actors, making it crucial for security teams to identify, prioritize, and secure sensitive data resources across their cloud environments. To address this challenge, we're announcing the integration between Microsoft Defender for Cloud and Microsoft Purview in public preview.

Integrate with Microsoft 365 and Microsoft Defender for Cloud

Often, one of the biggest challenges for security organization in a company is to identify and protect assets based on their criticality and sensitivity. Microsoft recently announced integration between Microsoft Purview and Microsoft Defender for Cloud in Public Preview to help overcome these challenges.

If you've extended your Microsoft 365 sensitivity labels for assets and database columns in Microsoft Purview, you can keep track of highly valuable assets using Microsoft Defender for

Cloud from inventory, alerts and recommendations based on assets detected sensitivity labels.

- For recommendations, we've provided **security controls** to help you understand how important each recommendation is to your overall security posture. Microsoft Defender for Cloud includes a **secure score** value for each control to help you prioritize your security work. Learn more in Security controls and their recommendations.
- For alerts, we've assigned **severity labels** to each alert to help you prioritize the order in which you attend to each alert. Learn more in How are alerts classified?.

For more information, see Integrate Microsoft Purview with Azure security products.

Information Protection

Secure metadata extraction and storage

Microsoft Purview is a data governance solution in cloud. You can register and scan different data sources from various data systems from your on-premises, Azure, or multicloud environments into Microsoft Purview. While data source is registered and scanned in Microsoft Purview, the actual data and data sources stay in their original locations, only metadata is extracted from data sources and stored in Microsoft Purview Data Map, which means you don't need to move data out of the region or their original location to extract the metadata into Microsoft Purview.

If your account was created **before** December 15, 2023, then your account was deployed with a managed resource group and an Azure Storage account was deployed in that group. These resources are consumed by Microsoft Purview and can't be accessed by any other users or principals. An Azure role-based access control (RBAC) deny assignment is added automatically to this resource group when the Microsoft Purview account was deployed, and that prevents access by other users or any CRUD operations that aren't initiated by Microsoft Purview.

Accounts deployed **after** December 15, 2023 (or deployed using API version 2023-05-01-preview onwards) use an ingestion storage account which is deployed in internal Microsoft Azure subscriptions. Since these resources are not on the Microsoft Purview's Azure subscription, they can't be accessed by any other users or principals, except the Microsoft Purview account.

(If you've deployed your accounts using the API, accounts deployed using the API version 2023-05-01-preview onwards will use an ingestion storage account deployed on internal Microsoft Azure subscriptions, instead of a managed storage account deployed on your Azure subscription.)

Where is metadata stored?

Microsoft Purview extracts only the metadata from different data source systems into Microsoft Purview Data Map during the scanning process.

You can deploy a Microsoft Purview account inside your Azure subscription in any supported Azure regions.

All metadata is stored inside Data Map inside your Microsoft Purview instance. This means the metadata is stored in the same region as your Microsoft Purview instance.

How metadata is extracted from data sources?

Microsoft Purview allows you to use any of the following options to extract metadata from data sources:

- **Azure runtime**. Metadata data is extracted and processed inside the same region as your data sources.

    1. A manual or automatic scan is initiated from the Microsoft Purview Data Map through the Azure integration runtime.
    2. The Azure integration runtime connects to the data source to extract metadata.
    3. Metadata is queued in Microsoft Purview managed or ingestion storage and stored in Azure Blob Storage.
    4. Metadata is sent to the Microsoft Purview Data Map.

- **Self-hosted integration runtime**. Metadata is extracted and processed by self-hosted integration runtime inside self-hosted integration runtime VMs' memory before they're sent to Microsoft Purview Data Map. In this case, customers have to deploy and manage one or more self-hosted integration runtime Windows-based virtual machines inside their Azure subscriptions or on-premises environments. Scanning on-premises and VM-based data sources always requires using a self-hosted integration runtime. The Azure integration runtime isn't supported for these data sources. The following steps show the communication flow at a high level when you're using a self-hosted integration runtime to scan a data source.

    1. A manual or automatic scan is triggered. Microsoft Purview connects to Azure Key Vault to retrieve the credential to access a data source.
    2. The scan is initiated from the Microsoft Purview Data Map through a self-hosted integration runtime.

3. The self-hosted integration runtime service from the VM connects to the data source to extract metadata.

4. Metadata is processed in VM memory for the self-hosted integration runtime. Metadata is queued in Microsoft Purview ingestion storage and then stored in Azure Blob Storage.

5. Metadata is sent to the Microsoft Purview Data Map.

If you need to extract metadata from data sources with sensitive data that can't leave the boundary of your on-premises network, it's highly recommended to deploy the self-hosted integration runtime VM inside your corporate network, where data sources are located, to extract and process metadata in on-premises, and send only metadata to Microsoft Purview.

1. A manual or automatic scan is triggered. Microsoft Purview connects to Azure Key Vault to retrieve the credential to access a data source.

2. The scan is initiated through the on-premises self-hosted integration runtime.

3. The self-hosted integration runtime service from the VM connects to the data source to extract metadata.

4. Metadata is processed in VM memory for the self-hosted integration runtime. Metadata is queued in Microsoft Purview ingestion storage and then stored in Azure Blob Storage. Actual data never leaves the boundary of your network.

5. Metadata is sent to the Microsoft Purview Data Map.

Information protection and encryption

Azure offers many mechanisms for keeping data private in rest and as it moves from one location to another. For Microsoft Purview, data is encrypted at rest using Microsoft-managed keys and when data is in transit, using Transport Layer Security (TLS) v1.2 or greater.

Transport Layer Security (Encryption-in-transit)

Data in transit (also known as data in motion) is encrypted in Microsoft Purview.

To add another layer of security in addition to access controls, Microsoft Purview secures customer data by encrypting data in motion with Transport Layer Security (TLS) and protect data in transit against 'out of band' attacks (such as traffic capture). It uses encryption to make sure attackers can't easily read or modify the data.

Microsoft Purview supports data encryption in transit with Transport Layer Security (TLS) v1.2 or greater.

For more information, see, Encrypt sensitive information in transit.

Transparent data encryption (Encryption-at-rest)

Data at rest includes information that resides in persistent storage on physical media, in any digital format. The media can include files on magnetic or optical media, archived data, and data backups inside Azure regions.

To add another layer of security in addition to access controls, Microsoft Purview encrypts data at rest to protect against 'out of band' attacks (such as accessing underlying storage). It uses encryption with Microsoft-managed keys. This practice helps make sure attackers can't easily read or modify the data.

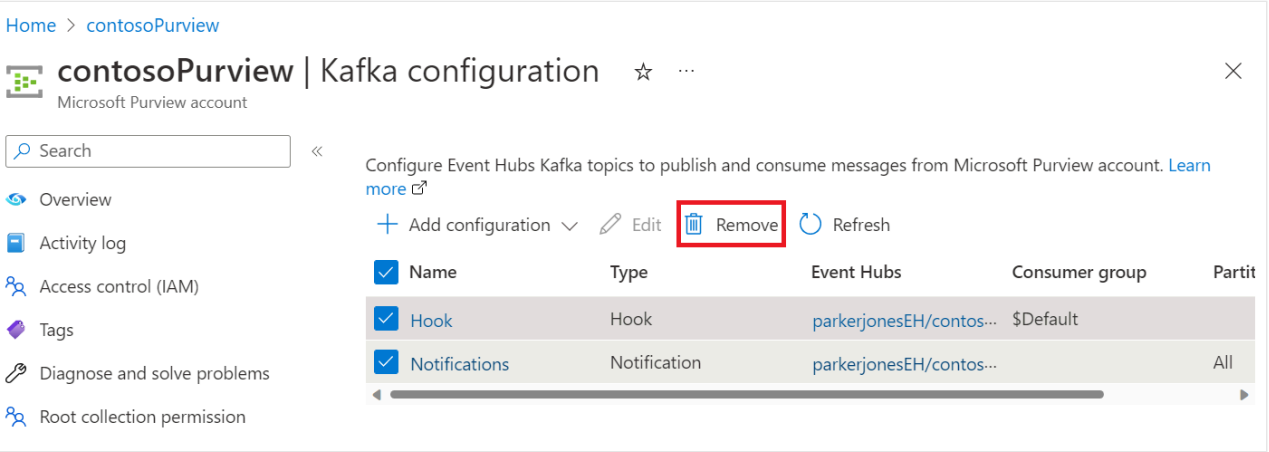
For more information, see Encrypt sensitive data at rest.

Optional Event Hubs namespace configuration

Each Microsoft Purview account can configure Event Hubs that are accessible via their Atlas Kafka endpoint. This can be enabled at creation under *Configuration*, or from the Azure portal under *Kafka configuration*. It's recommended to only enable optional managed event hub if it's used to distribute events into or outside of Microsoft Purview account Data Map. To remove this information distribution point, either don't configure these endpoints, or remove them.

To remove configured Event Hubs namespaces, you can follow these steps:

1. Search for and open your Microsoft Purview account in the Azure portal.
2. Select **Kafka configuration** under settings on your Microsoft Purview account page in the Azure portal.
3. Select the Event Hubs you want to disable. (Hook hubs send messages to Microsoft Purview. Notification hubs receive notifications.)
4. Select **Remove** to save the choice and begin the disablement process. This can take several minutes to complete.



> **Note**
>
> If you have an ingestion private endpoint when you disable this Event Hubs namespace, after disabling the ingestion private endpoint may show as disconnected.

For more information about configuring these Event Hubs namespaces, see: Configure Event Hubs for Atlas Kafka topics

Credential management

To extract metadata from a data source system into Microsoft Purview Data Map, it's required to register and scan the data source systems in Microsoft Purview Data Map. To automate this process, we have made available connectors for different data source systems in Microsoft Purview to simplify the registration and scanning process.

To connect to a data source Microsoft Purview requires a credential with read-only access to the data source system.

It's recommended prioritizing the use of the following credential options for scanning, when possible:

1. Microsoft Purview Managed Identity
2. User Assigned Managed Identity
3. Service Principals
4. Other options such as Account key, SQL Authentication, etc.

If you use any options rather than managed identities, all credentials must be stored and protected inside an Azure key vault. Microsoft Purview requires get/list access to secret on the Azure Key Vault resource.

As a general rule, you can use the following options to set up integration runtime, and credentials to scan data source systems:

Expand table

Scenario	Runtime option	Supported Credentials
Data source is an Azure Platform as a Service, such as Azure Data Lake Storage Gen 2 or Azure SQL inside public network	Option 1: Azure Runtime	Microsoft Purview Managed Identity, Service Principal or Access Key / SQL Authentication (depending on Azure data source type)
Data source is an Azure Platform as a Service, such as Azure Data Lake Storage Gen 2 or Azure SQL inside public network	Option 2: Self-hosted integration runtime	Service Principal or Access Key / SQL Authentication (depending on Azure data source type)
Data source is an Azure Platform as a Service, such as Azure Data Lake Storage Gen 2 or Azure SQL inside private network using Azure Private Link Service	Self-hosted integration runtime	Service Principal or Access Key / SQL Authentication (depending on Azure data source type)
Data source is inside an Azure IaaS VM such as SQL Server	Self-hosted integration runtime deployed in Azure	SQL Authentication or Basic Authentication (depending on Azure data source type)
Data source is inside an on-premises system such as SQL Server or Oracle	Self-hosted integration runtime deployed in Azure or in the on-premises network	SQL Authentication or Basic Authentication (depending on Azure data source type)
Multicloud	Azure runtime or self-hosted integration runtime based on data source types	Supported credential options vary based on data sources types
Power BI tenant	Azure Runtime	Microsoft Purview Managed Identity

Use this guide to read more about each source and their supported authentication options.

Other recommendations

Apply security best practices for Self-hosted runtime VMs

Consider securing the deployment and management of self-hosted integration runtime VMs in Azure or your on-premises environment, if self-hosted integration runtime is used to scan data sources in Microsoft Purview.

For self-hosted integration runtime VMs deployed as virtual machines in Azure, follow security best practices recommendations for Windows virtual machines.

- Lock down inbound traffic to your VMs using Network Security Groups and Azure Defender access Just-in-Time.
- Install antivirus or antimalware.
- Deploy Azure Defender to get insights around any potential anomaly on the VMs.
- Limit the amount of software in the self-hosted integration runtime VMs. Although it isn't a mandatory requirement to have a dedicated VM for a self-hosted runtime for Microsoft Purview, we highly suggest using dedicated VMs especially for production environments.
- Monitor the VMs using Azure Monitor for VMs. By using Log analytics agent, you can capture content such as performance metrics to adjust required capacity for your VMs.
- By integrating virtual machines with Microsoft Defender for Cloud, you can you prevent, detect, and respond to threats.
- Keep your machines current. You can enable Automatic Windows Update or use Update Management in Azure Automation to manage operating system level updates for the OS.
- Use multiple machines for greater resilience and availability. You can deploy and register multiple self-hosted integration runtimes to distribute the scans across multiple self-hosted integration runtime machines or deploy the self-hosted integration runtime on a Virtual Machine Scale Set for higher redundancy and scalability.
- Optionally, you can plan to enable Azure backup from your self-hosted integration runtime VMs to increase the recovery time of a self-hosted integration runtime VM if there's a VM level disaster.

Next steps

- Microsoft Purview network architecture and best practices
- Credentials for source authentication in Microsoft Purview

Microsoft Purview data security solutions

> *Microsoft Purview service description*

Microsoft Purview data security solutions help you manage and monitor your data and protect information. This article helps you learn about Microsoft Purview data security solutions and quickly get started with deploying these solutions to meet specific security needs for your organization.

Protect sensitive data across clouds, apps, and devices

Your information protection strategy should be driven by your business needs, but every organization has a requirement to protect some or all of its data. Use the capabilities from Microsoft Purview Information Protection (formerly Microsoft Information Protection) to help you discover, classify, protect, and govern sensitive information wherever it lives or travels.

Know your data

Your organization likely has information residing across the Microsoft 365 services, other cloud services, and on-premises. Identifying which items are sensitive and gaining visibility into how they're being used is central to your information protection practice. Microsoft Purview includes:

- Sensitive information types to identify sensitive items by using built-in or custom regular expressions, or a function.
- Trainable classifiers to identify sensitive items by using examples of the data you're interested in rather than identifying elements in the item.
- Data classification provides a graphical identification of items in your organization that have a sensitivity label, a retention label, or have been classified and the actions your users are taking on them

Protect your data

There are many capabilities that you can use from the Microsoft Purview Information Protection solution to help protect your data, wherever it's stored and however it's accessed. However, sensitivity labels are the foundational capability that both provide protection actions and interact with other Purview solutions and capabilities.

Sensitivity labels provide users and admins with visibility into the sensitivity of the data that they're using, and the labels themselves can apply protection actions that include encryption, access restrictions, and visual markings. For more information about the range of labeling scenarios supported, see the Common scenarios for sensitivity labels section from the getting started documentation. For more information about sensitivity labels, see Learn about sensitivity labels.

Encrypt your data and control your encryption keys

Encryption is an important part of your information protection strategy. The encryption process encodes your data (referred to as plaintext) into ciphertext. Unlike plaintext, ciphertext can't be used by people or computers unless and until the ciphertext is decrypted. Decryption requires an encryption key that only authorized users have. Encryption helps ensure that only authorized recipients can decrypt your content.

Microsoft Purview Double Key Encryption helps secure your most sensitive data that is subject to the strictest protection requirements. Microsoft Purview Customer Key helps you meet regulatory or compliance obligations for controlling root keys. You explicitly authorize Microsoft 365 services to use your encryption keys to provide value added cloud services, such as eDiscovery, anti-malware, anti-spam, search indexing, and so on.

Prevent data loss

Unintentional sharing of sensitive items can cause financial harm to your organization and may result in a violation of laws and regulations. Microsoft Purview Data Loss Prevention can help protect your organization against unintentional or accidental sharing of sensitive information both inside and outside of your organization. In a data loss prevention policy, you:

- Define the sensitive information you want to monitor for, like financial, health, medical, and privacy data.
- Where to monitor, like Microsoft 365 services or Windows and macOS devices.
- The conditions that must be matched for a policy to be applied to an item, like items containing credit card, driver's license, or social security numbers.
- The actions to take when a match is found, like audit, block the activity, and block the activity with override.

Address your data security objectives with Data Security Posture Management

Microsoft Data Security Posture Management (DSPM) helps you discover, protect, and investigate sensitive data risks across your digital estate. Unifying other Microsoft Purview solutions by focusing on data security objectives with remediation plans that have guided workflows and reports, DSPM provides visibility and control for both traditional applications and AI apps and agents, regardless of where sensitive data resides.

Analyze and respond to data security incidents with Data Security Investigations

Microsoft Purview Data Security Investigations helps cybersecurity teams in your organization harness generative artificial intelligence (AI) to analyze and respond to data security incidents, risky insiders, and data breaches. Investigations help you quickly identify risks from sensitive data exposure and more effectively collaborate with your partner teams to remediate the issues and simplify tasks that traditionally are time consuming and complex.

Detect and act on risk activities with Insider Risk Management

Microsoft Purview Insider Risk Management uses the full breadth of service and 3rd-party indicators to help you quickly identify, triage, and act on risky user activity in your organization. By using logs from Microsoft 365 and Microsoft Graph, insider risk management allows you to define specific policies to identify risk indicators. After identifying risky activities, you can take action to mitigate these risks.

Restrict communication and collaboration between users with Information Barriers

Microsoft Purview Information Barriers (IB) is a compliance solution that allows you to restrict two-way communication and collaboration between groups and users in Microsoft Teams, SharePoint, and OneDrive. Often used in highly regulated industries, IB can help to avoid conflicts of interest and safeguard internal information between users and organizational areas.

Limit access to sensitive data with Privileged Access Management

Having standing access by some users to sensitive information or critical network configuration settings in Microsoft Exchange is a potential pathway for compromised accounts or internal

threat activities. Microsoft Purview Privileged Access Management helps protect your organization from breaches and helps to meet compliance best practices by limiting standing access to sensitive data or access to critical configuration settings. Instead of administrators having constant access, just-in-time access rules are implemented for tasks that need elevated permissions.

Last updated on 05/01/2026

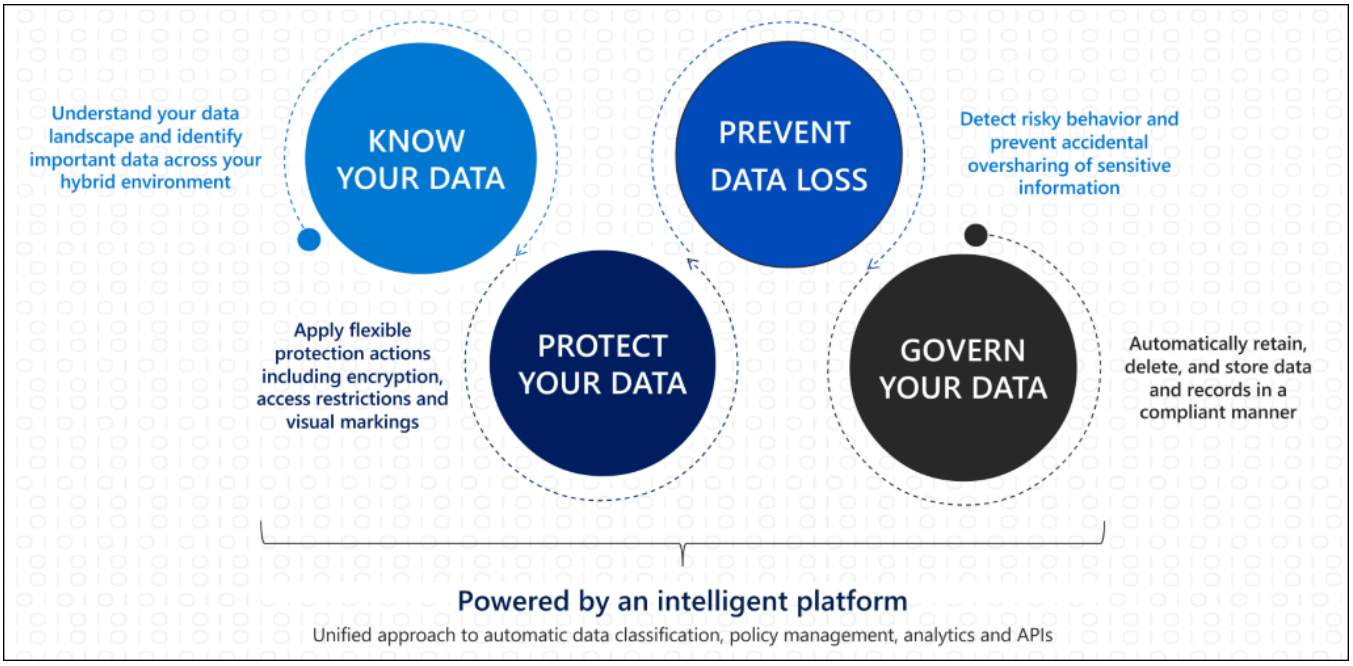
Protect your sensitive data with Microsoft Purview

Summarize this article for me

> *Microsoft Purview service description*

Implement capabilities from **Microsoft Purview Information Protection** (formerly Microsoft Information Protection) to help you discover, classify, and protect sensitive information wherever it lives or travels.

These information protection capabilities give you the tools to know your data, protect your data, and prevent data loss.



Use the following sections to learn more about the available capabilities and how to get started with each one. However, if you're looking for a guided deployment, see Deploy an information protection solution with Microsoft Purview.

For information about governing your data for compliance or regulatory requirements, see Govern your data with Microsoft Purview.

Know your data

To understand your data landscape and identify sensitive data across your hybrid environment, use the following capabilities:

Expand table

Capability	What problems does it solve?	Get started
Sensitive information types	Identifies sensitive data by using built-in or custom regular expressions or a function. Corroborative evidence includes keywords, confidence levels, and proximity.	Customize a built-in sensitive information type
Trainable classifiers	Identifies sensitive data by using examples of the data you're interested in rather than identifying elements in the item (pattern matching). You can use built-in classifiers or train a classifier with your own content.	Get started with trainable classifiers
Data classification	A graphical identification of items in your organization that have a sensitivity label, a retention label, or have been classified. You can also use this information to gain insights into the actions that your users are taking on these items.	Get started with content explorer Get started with activity explorer

Protect your data

To apply flexible protection actions that include encryption, access restrictions, and visual markings, use the following capabilities:

Expand table

Capability	What problems does it solve?	Get started
Sensitivity labels	A single labeling solution across apps, services, and devices to protect your data as it travels inside and outside your organization. Example scenarios: - Manage sensitivity labels for Office apps - Encrypt documents and emails - Protect calendar items, Teams meetings, and chat For a comprehensive list of supported scenarios for sensitivity labels, see the Get started documentation.	Get started with sensitivity labels
Microsoft Purview Information Protection client	For Windows computers, extends labeling to File Explorer and PowerShell	Extend sensitivity labeling on Windows
Double Key Encryption	Under all circumstances, only your organization can ever decrypt protected content or for regulatory requirements, you must hold encryption keys within a geographical boundary.	Deploy Double Key Encryption

Capability	What problems does it solve?	Get started
Message Encryption	Encrypts email messages and attached documents that are sent to any user on any device, so only authorized recipients can read emailed information. Example scenario: Revoke email encrypted by Advanced Message Encryption	Set up Message Encryption
Service encryption with Customer Key	Protects against viewing of data by unauthorized systems or personnel, and complements BitLocker disk encryption in Microsoft datacenters.	Set up Customer Key
SharePoint Information Rights Management (IRM)	Protects SharePoint lists and libraries so that when a user checks out a document, the downloaded file is protected so that only authorized people can view and use the file according to policies that you specify.	Set up Information Rights Management (IRM) in SharePoint admin center
Rights Management connector	Protection-only for existing on-premises deployments that use Exchange or SharePoint Server, or file servers that run Windows Server and File Classification Infrastructure (FCI).	Steps to deploy the RMS connector
Information protection scanner	Discovers, labels, and protects sensitive information that resides in data stores that are on premises.	Configuring and installing the information protection scanner
Microsoft Defender for Cloud Apps	Discovers, labels, and protects sensitive information that resides in data stores that are in the cloud.	Discover, classify, label, and protect regulated and sensitive data stored in the cloud
Microsoft Purview Data Map	Identifies sensitive data and applies automatic labeling to content in Microsoft Purview Data Map assets. These include files in storage such as Azure Data Lake and Azure Files, and schematized data such as columns in Azure SQL DB and Azure Cosmos DB.	Labeling in Microsoft Purview Data Map
Microsoft Information Protection SDK	Extends sensitivity labels to third-party apps and services. Example scenario: Set and get a sensitivity label (C++)	Microsoft Information Protection (MIP) SDK setup and configuration

Prevent data loss

To help prevent accidental oversharing of sensitive information, use the following capabilities:

Expand table

Capability	What problems does it solve?	Get started
Microsoft Purview Data Loss Prevention	Helps prevent unintentional sharing of sensitive items.	Learn about data loss prevention
Endpoint data loss prevention	Extends DLP capabilities to items that are used and shared on Windows 10 computers.	Get started with Endpoint data loss prevention
Microsoft Purview extension for Chrome	Extends DLP capabilities to the Chrome browser	Get started with the Microsoft Purview extension for Chrome
Microsoft Purview data loss prevention on-premises repositories	Extends DLP monitoring of file activities and protective actions for those files to on-premises file shares and SharePoint folders and document libraries.	Get started with Microsoft Purview data loss prevention on-premises repositories
Protect sensitive information in Microsoft Teams chat and channel messages	Extends some DLP functionality to Teams chat and channel messages	Learn about the default data loss prevention policy in Microsoft Teams

Licensing requirements

License requirements to protect your sensitive data depend on the scenarios and features you use. Rather than set licensing requirements for each capability listed on this page, for more information, see the Microsoft Purview service description and the related PDF download for feature-level licensing requirements.

Last updated on 02/27/2026

Deploy an information protection solution with Microsoft Purview

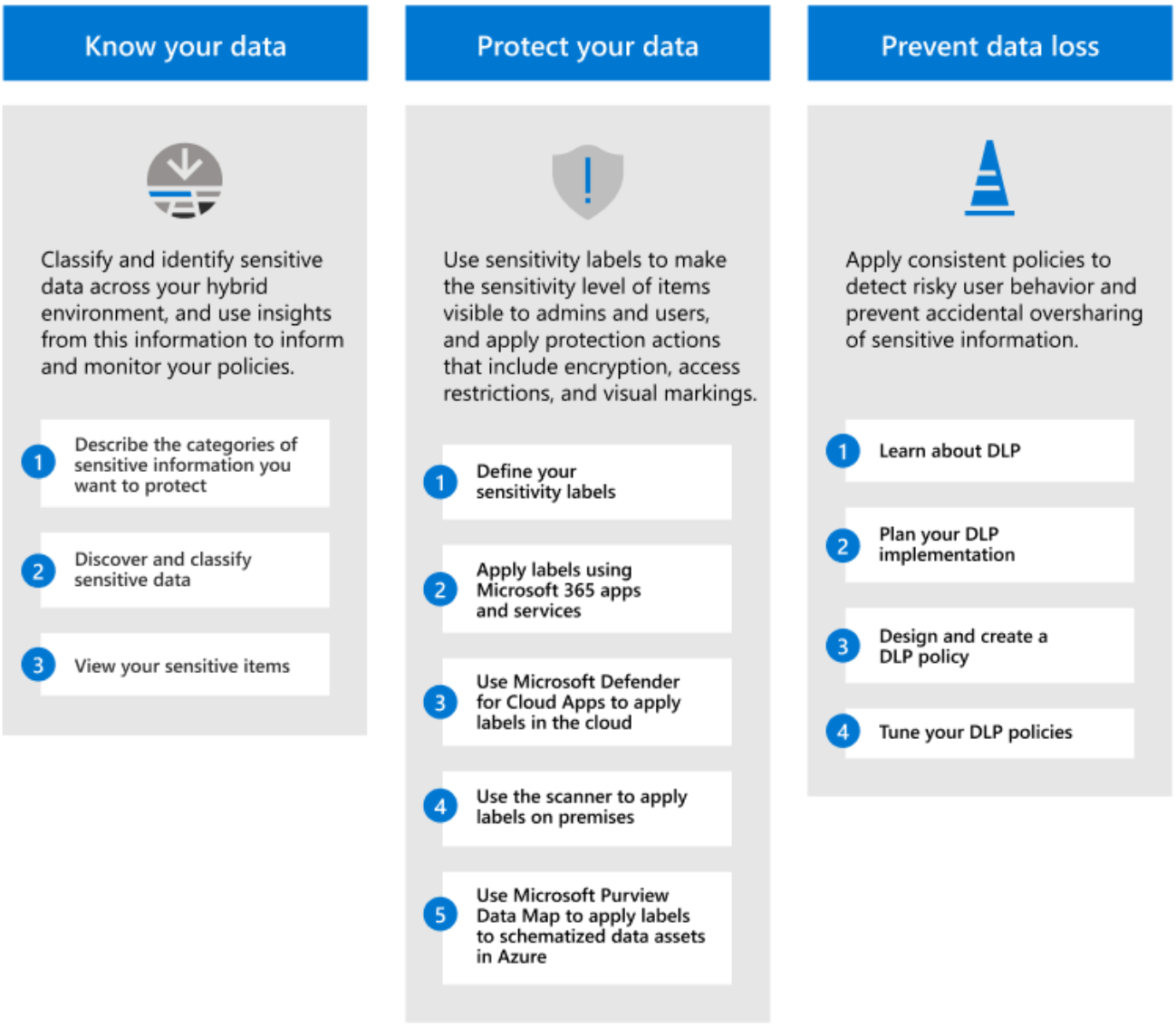
> *Microsoft Purview service description*

Your information protection strategy is driven by your business needs. Many organizations must comply with regulations, laws, and business practices. Additionally, organizations need to protect proprietary information, such as data for specific projects.

Microsoft Purview Information Protection (formerly Microsoft Information Protection) provides a framework, process, and capabilities you can use to protect sensitive data across clouds, apps, and devices.

Microsoft Purview Information Protection framework

Use Microsoft Purview Information Protection to help you discover, classify, protect, and govern sensitive information wherever it lives or travels.



For data governance, see Deploy a data governance solution with Microsoft Purview.

> **Tip**
>
> Consider using the new version of Data Security Posture Management (DSPM) to supplement the information on this page and jumpstart securing your data with Microsoft Purview Information Protection.
>
> DSPM helps you discover, protect, and investigate sensitive data risks across your digital estate by unifying Microsoft Purview Information Protection with other Microsoft Purview solutions and capabilities. With a focus on data security objectives with guided workflows that offer default sensitivity labels, default policies, and more, DSPM helps you monitor, assess, and remediate data risks regardless of where sensitive data resides.

Licensing

Microsoft Purview Information Protection capabilities are included with Microsoft Purview. The licensing requirements can vary even within capabilities, depending on configuration options. To identify licensing requirements and options, see the Microsoft Purview service description.

Know your data

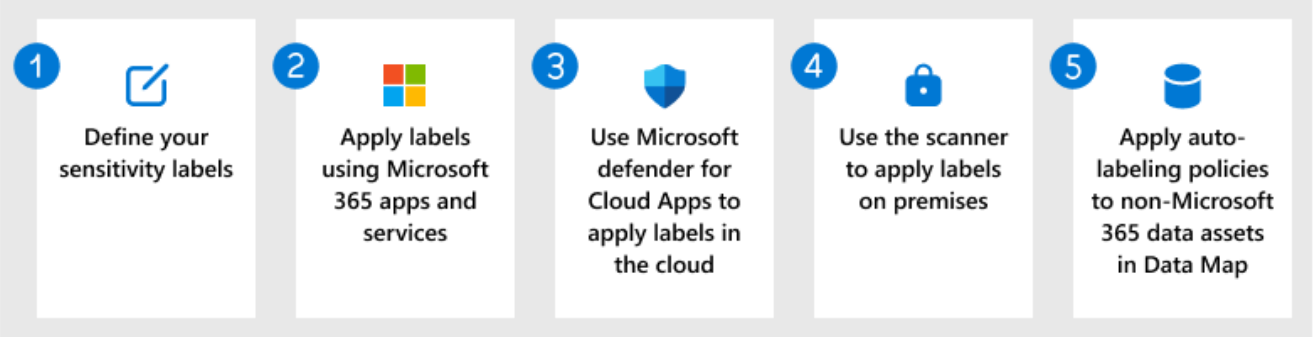


Knowing where your sensitive data resides is often the biggest challenge for many organizations. Microsoft Purview Information Protection data classification helps you to discover and accurately classify ever-increasing amounts of data that your organization creates. Graphical representations help you gain insights into this data so you can set up and monitor policies to protect and govern it.

Expand table

Step	Description	More information
1	Describe the categories of sensitive information you want to protect. You already have an idea of what types of information are most valuable to your org and what types aren't. Work with stakeholders to describe these categories that are your starting point.	Learn about sensitive information types Learn about trainable classifiers
2	Discover and classify sensitive data. Sensitive data in items can be found by using many different methods that include default DLP policies, manual labeling by users, and automated pattern recognition using sensitive information types or machine learning.	Classifiers overview
3	View your sensitive items. Use content explorer and activity explorer for a deeper analysis of sensitive items and the actions that users are taking on these items.	Get started with content explorer Get started with activity explorer

Protect your data



Use the information from knowing where your sensitive data resides to help you more efficiently protect it. However, there's no need to wait—you can start to protect your data immediately with a combination of manual, default, and automatic labeling. Then, use content explorer and activity explorer from the previous section to confirm what items are labeled and how your labels are being used.

Expand table

Step	Description	More information
1	Define your sensitivity labels and policies that will protect your organization's data. In addition to identifying the sensitivity of content, these labels can apply protection actions such as content markings (headers, footers, watermarks), encryption, and other access controls. Example sensitivity labels: **Personal** **Public** **General** - **Anyone (unrestricted)** - **All Employees (unrestricted)** **Confidential** - **Anyone (unrestricted)** - **All Employees** - **Trusted People** **Highly Confidential** - **All Employees** - **Specific People** Example sensitivity label policy: 1. Publish all labels to all users in the tenant 2. Default label of **General \ All Employees (unrestricted)** for items	Get started with sensitivity labels Create and configure sensitivity labels and their policies Restrict access to content by using sensitivity labels to apply encryption

Step	Description	More information
	3. Users must provide a justification to remove a label or lower its classification	
2	Label and protect data for Microsoft 365 apps and services. Sensitivity labels are supported for Microsoft 365 Word, Excel, PowerPoint, Outlook, Teams meetings, and also containers that include SharePoint and OneDrive sites, and Microsoft 365 groups. Use a combination of labeling methods such as manual labeling, automatic labeling, a default label, and mandatory labeling. Example configuration for client-side auto-labeling: 1. Recommend **Confidential \ Anyone (unrestricted)** if 1-9 credit card numbers 2. Recommend **Confidential \ All Employees** if 10+ credit card numbers -- typical end user experience, and the user selects the button to show sensitive content (Word only) Example configuration for service-side auto-labeling: Apply to all locations (Exchange, SharePoint, OneDrive) 1. Apply **Confidential \ Anyone (unrestricted)** if 1-9 credit card numbers 2. Apply **Confidential \ All Employees** if 10+ credit card numbers 3. Apply **Confidential \ Anyone (unrestricted)** if 1-9 US personal data and full names 4. Apply **Confidential \ All Employees** if 10+ US personal data and full names	Manage sensitivity labels in Office apps Enable sensitivity labels for files in SharePoint and OneDrive Enable co-authoring for files encrypted with sensitivity labels Configure a default sensitivity label for a SharePoint document library Automatically apply a sensitivity label to Microsoft 365 data Use sensitivity labels to protect collaborative workspaces (groups and sites) Use sensitivity labels to protect calendar items, Teams meetings, and chat Use sensitivity labels with Microsoft Loop Use sensitivity labels to set the default sharing link for sites and documents in SharePoint and OneDrive Apply a sensitivity label to a model in Microsoft Syntex Sensitivity labels in Power BI
3	Discover, label, and protect sensitive items that reside in data stores in the cloud (Box, GSuite, SharePoint, and OneDrive) by using Microsoft Defender for Cloud Apps with your sensitivity labels.	Discover, classify, label, and protect regulated and sensitive data stored in the cloud

Step	Description	More information
	Example configuration for a file policy: Looks for credit card numbers in files stored in a Box account, and then applies a sensitivity label to identify the highly confidential info and encrypt it.	
4	Discover, label, and protect sensitive items that reside in data stores on premises by deploying the information protection scanner with your sensitivity labels.	Configuring and installing the information protection scanner
5	Auto-apply sensitivity labels in Microsoft Purview Data Map to discover and label items for Azure Blob Storage, Azure files, and Azure Data Lake Storage Gen2.	Learn about sensitivity labels in Microsoft Purview Data Map

If you're a developer who wants to extend sensitivity labels to line-of-business apps or third-party SaaS apps, see Microsoft Information Protection (MIP) SDK setup and configuration.

Additional protection capabilities

Microsoft Purview includes additional capabilities to help protect data. Not every customer needs these capabilities, and some might be superseded by more recent releases.

Refer to the Protect your data with Microsoft Purview page for the full list of protection capabilities.

Prevent data loss



Deploy Microsoft Purview Data Loss Prevention (DLP) policies to govern and prevent the inappropriate sharing, transfer, or use of sensitive data across apps and services. These policies help users make the right decisions and take the right actions when they're using sensitive data.

Expand table

Step	Description	More information
1	Learn about DLP. Organizations have sensitive information under their control, such as financial data, proprietary data, credit card numbers, health records, and social security numbers. To help protect this sensitive data and reduce risk, they need a way to prevent their users from inappropriately sharing it with people who shouldn't have it. This practice is called data loss prevention (DLP).	Learn about data loss prevention
2	Plan your DLP implementation. Every organization will plan for and implement data loss prevention (DLP) differently, because every organization's business needs, goals, resources, and situation are unique to them. However, there are elements that are common to all successful DLP implementations.	Plan for data loss prevention
3	Design and create a DLP policy. Creating a data loss prevention (DLP) policy is quick and easy, but getting a policy to yield the intended results can be time consuming if you have to do a lot of tuning. Taking the time to design a policy before you implement it gets you to the desired results faster, and with fewer unintended issues, than tuning by trial and error alone. Help prevent sharing credit card numbers through email: Prevents emails being sent if they contain credit card numbers or the email has a specific sensitivity label that identities highly confidential info.	Design a DLP policy DLP policy reference Create and deploy data loss prevention policies
4	Tune your DLP policies. After you deploy a DLP policy, you'll see how well it meets the intended purpose. Use that information to adjust your policy settings for better performance.	Create and deploy data loss prevention policies

Deployment strategies

The credit card number examples are often helpful for initial testing and end user education. Even if your organization doesn't typically need to protect credit card numbers, the concept of these being sensitive items that need protection is easily understood by users. Many websites provide credit card numbers that are suitable for testing purposes only. You can also search for sites that provide credit card number generators so that you can paste the numbers into documents and emails.

When you're ready to move your automatic labeling and DLP policies into production, change to classifiers and configurations that are suitable for the type of data used by your organization. For example, you might need to use trainable classifiers for intellectual property and specific types of documents, or exact data match (EDM) sensitive information types for privacy data that's related to customers or employees.

Or, you might want to start by discovering and protecting IT-related information that is frequently the target of security attacks. Then, supplement this by checking for and preventing the sharing of passwords with DLP policies for email and Teams chat:

- Use the trainable classifiers **IT** and **IT Infra and Network Security Documents**
- Use the built-in sensitive info type **General Password** and create a custom sensitive info type for "password is" for the different languages used by your users

Deploying an information protection solution isn't a linear deployment but iterative, and often circular. The more you know your data, the more accurately you can label it, and prevent data leakage. The results of those applied labels and policies flow into the Reports and tools, which in turn makes more sensitive data visible for you to protect. Or, if you're already protecting that sensitive data, consider whether it requires additional protective actions.

You can start to manually label data as soon as you've defined sensitivity labels. The same classifiers that you use for DLP can be used to automatically find and label more data. You can even use sensitivity labels as a classifier, for example, block sharing items that are labeled highly confidential.

Most customers already have some solutions in place to protect their data. Your deployment strategy might be to build on what you already have, or focus on gaps that offer the most business value or addresses high risk areas.

To help you plan your unique deployment strategy, reference Get started with sensitivity labels and Plan for data loss prevention.

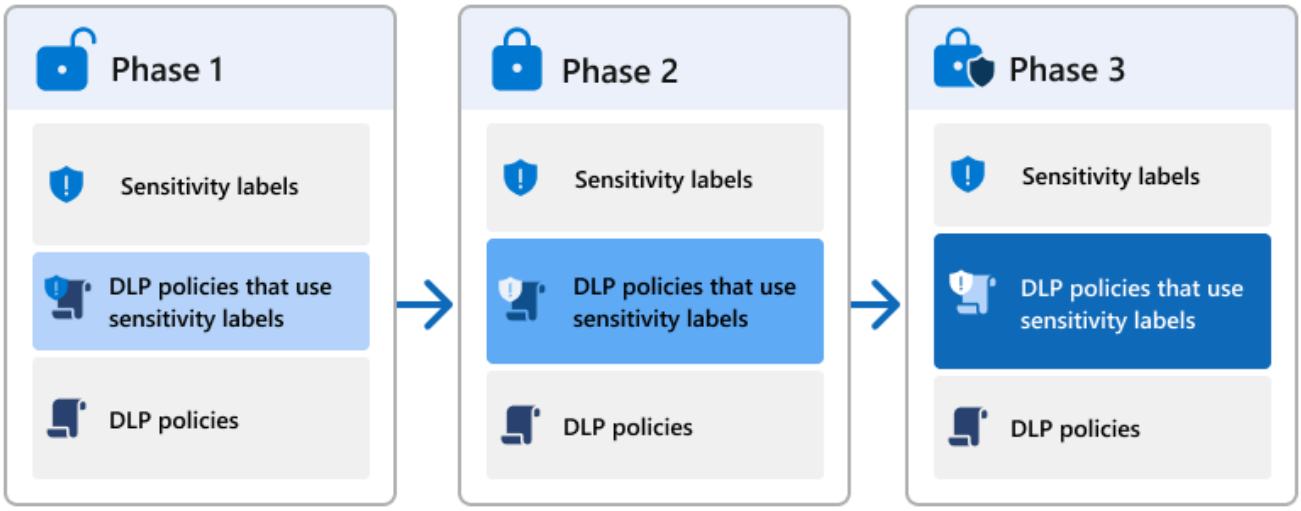
Consider a phased deployment

You might prefer to deploy information protection by using a phased deployment that implements progressively restrictive controls. This approach gradually introduces new protection measures for users as you gain familiarity and confidence with the technology. For example:

- From default labels and no encryption, to recommending labels that apply encryption when sensitive data is found, and then automatically applying labels when sensitive data is found.

- DLP policies that progress from auditing oversharing actions, to more restrictive blocking with warning to educate users, and then blocking all sharing.

Details of such a phased deployment might look something like the following plan, where sensitivity labels and DLP policies become more integrated with each other to provide greater data protection than if they were used independently:



Phase 1

Sensitivity label configurations:

- **General\All Employees**: Default label for email. No encryption. If applied on emails, block users from over-sharing.
- **Confidential\All Employees**: Default label for documents. No encryption. If applied on emails, block users from over-sharing.
- **Highly Confidential\All Employees**: No encryption. If applied on emails, block users from over-sharing.

DLP policy A:

- If 1-2 instances of credit cards are found, block external sharing except if the item is labeled as **Personal** or **Confidential\Anyone (unrestricted)**. Use logging and reporting for analysis.

DLP policy B:

- If 3-9 instances of credit cards are found, block external sharing, cloud egress, and copy to removable drives except if the item is labeled as **Confidential\Anyone (unrestricted)**. Use logging and reporting for analysis.

DLP policy C:

- If 10+ instances of credit cards are found, block external sharing, cloud egress, and copy to removable drives with no exceptions. Use logging and reporting for analysis.

Configuration details for this example phased deployment:

- Default sublabel for a parent label or label group
- Data loss prevention Exchange conditions and actions reference
- Confidence levels and other elements of a sensitivity type
- Use sensitivity labels as conditions in DLP policies
- Encryption that lets users assign permissions
- Encryption for specific usage rights
- Configure and view alerts for data loss prevention policies

For other deployment guidance that includes a phased deployment strategy, see Secure by default with Microsoft Purview and protect against oversharing from our Product Engineering team.

Training resources

Interactive guides: Microsoft Purview Information Protection

Learning modules for consultants and admins:

- Introduction to information protection and data lifecycle management in Microsoft Purview
- Classify data for protection and governance
- Protect information in Microsoft Purview
- Prevent data loss in Microsoft Purview

To help train your users to apply and use the sensitivity labels that you configure for them, see End-user documentation for sensitivity labels.

When you deploy data loss prevention policies for Teams, you might find the following end-user guidance useful as an introduction to this technology. It includes some potential messages that users might see: Teams messages about data loss prevention (DLP) and communication compliance policies.

Last updated on 03/31/2026

Microsoft Purview Information Protection for Office 365 operated by 21Vianet

09/02/2025

This article covers the differences between Microsoft Purview Information Protection support for Office 365 operated by 21Vianet and commercial offerings that are limited to offering previously known as Azure Information Protection (AIP), as well as specific configuration instructions for customers in China— including how to install the information protection scanner and manage content scan jobs.

Differences between 21Vianet and commercial offerings

While our goal is to deliver all commercial features and functionality to customers in China with our Microsoft Purview Information Protection support for Office 365 operated by 21Vianet offer, there's some missing functionality:

- Active Directory Rights Management Services (AD RMS) encryption is supported only in Microsoft 365 Apps for enterprise (build 11731.10000 or later). Office Professional Plus doesn't support AD RMS.
- Migration from AD RMS to the Azure Rights Management service isn't currently available.
- Sharing of encrypted emails with users in the commercial cloud is supported.
- Sharing of documents and email attachments with users in the commercial cloud isn't currently available. This includes Office 365 operated by 21Vianet users in the commercial cloud, non-Office 365 operated by 21Vianet users in the commercial cloud, and users with an RMS for Individuals license.
- IRM with SharePoint (IRM-protected sites and libraries) isn't currently available.
- The Mobile Device Extension for AD RMS isn't currently available.
- The Microsoft Purview Information Protection viewer for iOS and Android isn't supported by Azure China 21Vianet.
- The scanner area of the Microsoft Purview portal is unavailable to customers in China. Use PowerShell commands instead of performing actions in the portal, such as managing and running your content scan jobs.

- Network endpoints for the Microsoft Purview Information Protection client within the 21Vianet environment are different than the endpoints required for other cloud services. Network connectivity from clients to the following endpoints is required:
  - Download label and label policies: `*.protection.partner.outlook.cn`
  - Azure Rights Management service: `*.aadrm.cn`

- Document Tracking and Revocation *by users* isn't currently available.

Configuration for customers in 21Vianet

To configure Microsoft Purview Information Protection support for Office 365 operated by 21Vianet:

1. Enable Rights Management for the tenant.
2. Add the Microsoft Information Protection Sync Service service principal.
3. Configure DNS encryption.
4. Install and configure the Microsoft Purview Information Protection client.
5. Configure Windows settings.
6. Install the information protection scanner and manage content scan jobs.

Step 1: Enable Rights Management for the tenant

For the encryption to work correctly, the Azure Rights Management service must be enabled for the tenant.

1. Check if this service is enabled:
   a. Launch PowerShell as an administrator.
   b. If the AIPService module isn't installed, run `Install-Module AipService`.
   c. Import the module using `Import-Module AipService`.
   d. Connect to the service using `Connect-AipService -environmentname azurechinacloud`.
   e. Run `(Get-AipServiceConfiguration).FunctionalState` and check if the state is `Enabled`.
2. If the functional state is `Disabled`, run `Enable-AipService`.

Step 2: Add the Microsoft Information Protection Sync Service service principal

The **Microsoft Information Protection Sync Service** service principal isn't available in Azure China tenants by default, and is required for Microsoft Purview Information Protection. Create this service principal manually via the Azure Az PowerShell module.

1. If you don't have the Azure Az module installed, install it or use a resource where the Azure Az module comes preinstalled, such as Azure Cloud Shell. For more information, see Install the Azure Az PowerShell module.

2. Connect to the service using the Connect-AzAccount cmdlet and the `azurechinacloud` environment name:

   PowerShell

   ```
   Connect-azaccount -environmentname azurechinacloud
   ```

3. Create the **Microsoft Information Protection Sync Service** service principal manually using the New-AzADServicePrincipal cmdlet and the `870c4f2e-85b6-4d43-bdda-6ed9a579b725` application ID for the Microsoft Purview Information Protection Sync Service:

   PowerShell

   ```
   New-AzADServicePrincipal -ApplicationId 870c4f2e-85b6-4d43-bdda-6ed9a579b725
   ```

4. After adding the service principal, add the relevant permissions required to the service.

Step 3: Configure DNS encryption

For encryption to work correctly, Office client applications must connect to the China instance of the service and bootstrap from there. To redirect client applications to the right service instance, the tenant admin must configure a DNS SRV record with information about the URL for the Azure Rights Management service. Without the DNS SRV record, the client application attempts to connect to the public cloud instance by default and will fail.

Also, the assumption is that users will log in with a username based off the tenant-owned domain (for example, `joe@contoso.cn`), and not the `onmschina` username (for example, `joe@contoso.onmschina.cn`). The domain name from the username is used for DNS redirection to the correct service instance.

Configure DNS encryption - Windows

1. Get the Azure Rights Management service ID:

a. Launch PowerShell as an administrator.
b. If the AIPService module isn't installed, run `Install-Module AipService`.
c. Connect to the service using `Connect-AipService -environmentname azurechinacloud`.
d. Run `(Get-AipServiceConfiguration).RightsManagementServiceId` to get the service ID.

2. Log in to your DNS provider, navigate to the DNS settings for the domain, and then add a new SRV record.

   - Service = `_rmsredir`
   - Protocol = `_http`
   - Name = `_tcp`
   - Target = `[GUID].rms.aadrm.cn` (where GUID is the Azure Rights Management service ID)
   - Priority, Weight, Seconds, TTL = default values

3. Associate the custom domain with the tenant in the Azure portal. This adds an entry in DNS, which might take several minutes to get verified after you add the value to the DNS settings.

4. Log in to the Microsoft 365 admin center with the corresponding global admin credentials and add the domain (for example, `contoso.cn`) for user creation. In the verification process, additional DNS changes might be required. Once verification is done, users can be created.

Configure DNS encryption - Mac, iOS, Android

Log in to your DNS provider, navigate to the DNS settings for the domain, and then add a new SRV record.

- Service = `_rmsdisco`
- Protocol = `_http`
- Name = `_tcp`
- Target = `api.aadrm.cn`
- Port = `80`
- Priority, Weight, Seconds, TTL = default values

Step 4: Install and configure the labeling client

Download and install the Microsoft Purview Information Protection client from the Microsoft Download Center.

For more information, see:

- Extend sensitivity labeling on Windows
- Microsoft Purview Information Protection client - Release management and supportability

Step 5: Configure Windows settings

Windows needs the following registry key for authentication to point to the correct sovereign cloud for Azure China:

- Registry node = `HKEY_LOCAL_MACHINE\SOFTWARE\WOW6432Node\Microsoft\MSIP`
- Name = `CloudEnvType`
- Value = `6` (default = 0)
- Type = `REG_DWORD`

> **Important**
>
> Make sure you don't delete the registry key after an uninstall. If the key is empty, incorrect, or non-existent, the functionality will behave as the default value (default value = 0 for the commercial cloud). If the key is empty or incorrect, a print error is also added to the log.

Step 6: Install the information protection scanner and manage content scan jobs

Install the Microsoft Purview Information Protection scanner to scan your network and content shares for sensitive data, and apply classification and protection labels as configured in your organization's policy.

When configuring and managing your content scan jobs, use the following procedure instead of the Microsoft Purview portal that's used by the commercial offerings.

For more information, see Learn about the information protection scanner and Manage your content scan jobs using PowerShell only.

To install and configure your scanner:

1. Sign in to the Windows Server computer that will run the scanner. Use an account that has local administrator rights and that has permissions to write to the SQL Server master database.

2. Start with PowerShell closed. If you've previously installed the information protection scanner, make sure the Microsoft Purview Information Protection Scanner service is stopped.

3. Open a Windows PowerShell session with the **Run as an administrator** option.

4. Run the Install-Scanner cmdlet, specifying your SQL Server instance on which to create a database for the Microsoft Purview Information Protection scanner, and a meaningful name for your scanner cluster.

   ```PowerShell
   Install-Scanner -SqlServerInstance <name> -Cluster <cluster name>
   ```

   > **Tip**
   >
   > You can use the same cluster name in the **Install-Scanner** command to associate multiple scanner nodes to the same cluster. Using the same cluster for multiple scanner nodes enables multiple scanners to work together to perform your scans.

5. Verify that the service is now installed by using **Administrative Tools** > **Services**.

   The installed service is named **Microsoft Purview Information Protection Scanner** and is configured to run by using the scanner service account that you created.

6. Get an Azure token to use with your scanner. A Microsoft Entra token allows the scanner to authenticate to the Azure Information Protection service, enabling the scanner to run non-interactively.

   a. Open the Azure portal and create a Microsoft Entra application to specify an access token for authentication. For more information, see Run information protection labeling cmdlets unattended.

      > **Tip**
      >
      > When creating and configuring Microsoft Entra applications for the **Set-Authentication** command, the **Request API permissions** pane shows the **APIs my organization uses** tab instead of the **Microsoft APIs** tab. Select the **APIs my organization uses** to then select **Azure Rights Management Services**.

   b. From the Windows Server computer, if your scanner service account has been granted the **Log on locally** right for the installation, sign in with this account and start a PowerShell session.

      If your scanner service account can't be granted the **Log on locally** right for the installation, use the *OnBehalfOf* parameter with Set-Authentication, as described in

Run information protection labeling cmdlets unattended.

c. Run Set-Authentication, specifying values copied from your Microsoft Entra application:

```
PowerShell

Set-Authentication -AppId <ID of the registered app> -AppSecret <client
secret sting> -TenantId <your tenant ID> -DelegatedUser <Azure AD account>
```

For example:

```
PowerShell

$pscreds = Get-Credential CONTOSO\scanner
Set-Authentication -AppId "77c3c1c3-abf9-404e-8b2b-4652836c8c66" -AppSecret
"OAkk+rnuYc/u+]ah2kNxVbtrDGbS47L4" -DelegatedUser scanner@contoso.com -
TenantId "9c11c87a-ac8b-46a3-8d5c-f4d0b72ee29a" -OnBehalfOf $pscreds
Acquired application access token on behalf of CONTOSO\scanner.
```

The scanner now has a token to authenticate to Microsoft Entra ID. This token is valid for one year, two years, or never, according to your configuration of the **Web app /API** client secret in Microsoft Entra ID. When the token expires, you must repeat this procedure.

7. Run the Set-ScannerConfiguration cmdlet to set the scanner to function in offline mode. Run:

```
PowerShell

Set-ScannerConfiguration -OnlineConfiguration Off
```

8. Run the Set-ScannerContentScanJob cmdlet to create a default content scan job.

The only required parameter in the **Set-ScannerContentScanJob** cmdlet is **Enforce**. However, you may want to define other settings for your content scan job at this time. For example:

```
PowerShell

Set-ScannerContentScanJob -Schedule Manual -DiscoverInformationTypes
PolicyOnly -Enforce Off -DefaultLabelType PolicyDefault -RelabelFiles Off -
PreserveFileDetails On -IncludeFileTypes '' -ExcludeFileTypes '.msg,.tmp' -
DefaultOwner <account running the scanner>
```

The syntax above configures the following settings while you continue the configuration:

- Keeps the scanner run scheduling to *manual*
- Sets the information types to be discovered based on the sensitivity labeling policy
- Does *not* enforce a sensitivity labeling policy
- Automatically labels files based on content, using the default label defined for the sensitivity labeling policy
- Does *not* allow for relabeling files
- Preserves file details while scanning and auto-labeling, including *date modified*, *last modified*, and *modified by* values
- Sets the scanner to exclude .msg and .tmp files when running
- Sets the default owner to the account you want to use when running the scanner

9. Use the Add-ScannerRepository cmdlet to define the repositories you want to scan in your content scan job. For example, run:

```
PowerShell

Add-ScannerRepository -OverrideContentScanJob Off -Path 'c:\repoToScan'
```

Use one of the following syntaxes, depending on the type of repository you're adding:

- For a network share, use `\\Server\Folder`.
- For a SharePoint library, use `http://sharepoint.contoso.com/Shared%20Documents/Folder`.
- For a local path: `C:\Folder`
- For a UNC path: `\\Server\Folder`

> **Note**
>
> Wildcards aren't supported and WebDav locations aren't supported.
>
> To modify the repository later on, use the Set-ScannerRepository cmdlet instead.

Continue with the following steps as needed:

- Run a discovery cycle and view reports for the scanner
- Use PowerShell to configure the scanner to apply classification and protection
- Use PowerShell to configure a DLP policy with the scanner

The following table lists PowerShell cmdlets that are relevant for installing the scanner and managing your content scan jobs:

Expand table

Cmdlet	Description
Add-ScannerRepository	Adds a new repository to your content scan job.
Get-ScannerConfiguration	Returns details about your cluster.
Get-ScannerContentScan	Gets details about your content scan job.
Get-ScannerRepository	Gets details about repositories defined for your content scan job.
Remove-ScannerContentScan	Deletes your content scan job.
Remove-ScannerRepository	Removes a repository from your content scan job.
Set-ScannerContentScan	Defines settings for your content scan job.
Set-ScannerRepository	Defines settings for an existing repository in your content scan job.

For more information, see:

- Learn about the information protection scanner
- Configuring and installing the information protection scanner
- Manage your content scan jobs using PowerShell only

Default sensitivity labels and policies to protect your data

> *Microsoft Purview service description*

When you use Data Security Posture Management or Data Security Posture Management for AI (classic), you can activate some default sensitivity labels and their policies for Microsoft Purview Information Protection:

- Sensitivity labels and a sensitivity label publishing policy
- Client-side auto-labeling
- Service-side auto-labeling

> **Note**
>
> The default labels and policies are now automatically created for new customers who have an eligible license for Microsoft Purview.

These default configurations get you up and running quickly with Microsoft Purview Information Protection to help you protect your data in Microsoft 365. You can use them as-is, make just a few changes, or fully customize them to better suit your business requirements. Then, extend the labels beyond Microsoft 365, to protect data in other sources, such as Microsoft Azure and Microsoft Foundry.

If you're an existing customer who wasn't eligible for the automatic creation of the default labels and policies, you can activate the default configurations that are supported by your licenses if you haven't yet configured an equivalent:

Expand table

Default configuration	Equivalent
Sensitivity labels and a sensitivity label policy	Published sensitivity labels
Client-side auto-labeling	One or more sensitivity labels configured to automatically apply (or recommend to users) in Office apps

Default configuration	Equivalent
Service-side auto-labeling	At least one auto-labeling policy that's turned on

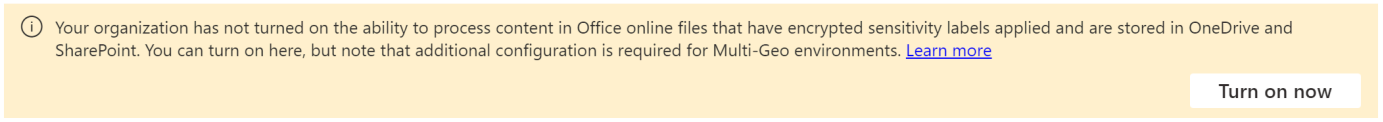
Activate the default labels and policies

If the default labels and polices aren't automatically created for you, you can activate them:

- From the current version of Data Security Posture Management > **Actions** > **Setup tasks**, select **Protect your data with sensitivity labels**
- From Data Security Posture Management for AI (classic) > **Recommendations**, select **Protect your data with sensitivity labels**

After the labels and policies are created, you'll need to enable sensitivity labels for SharePoint and OneDrive. This one-time, manual step is a prerequisite to use sensitivity labels in Office for the web, and auto-labeling policies for SharePoint and OneDrive.

To do this step, use the following banner at the top of the **Information Protection** > **Overview** page, and select **Turn on now**. If you don't see this banner, sensitivity labels for SharePoint and OneDrive have already been enabled for your tenant.



For more information about this capability, see Enable sensitivity labels for files in SharePoint and OneDrive.

Default sensitivity labels

When you don't have sensitivity labels that are published, we'll create the following labels for you:

Expand table

Label name	Label description for users	Settings
Personal	Non-business data, for personal use only.	**Scope**: Files & other data assets, Emails, Meetings* **Content marking**: No **Auto-labeling**: No

Label name	Label description for users	Settings
		Group settings: No **Site settings**: No **Auto-labeling for database columns**: None
Public	Business data that is specifically prepared and approved for public consumption.	**Scope**: Files & other data assets, Emails, Meetings* **Content marking**: No **Auto-labeling**: No **Group settings**: No **Site settings**: No **Auto-labeling for database columns**: None
General	Business data that is not intended for public consumption. However, this can be shared with external partners, as required. Examples include a company internal telephone directory, organizational charts, internal standards, and most internal communication.	**Scope**: Files & other data assets, Emails **Content marking**: No **Auto-labeling**: No **Group settings**: No **Site settings**: No **Auto-labeling for database columns**: None
General \ Anyone (unrestricted)	Organization data that isn't intended for public consumption but can be shared with external partners if appropriate. Examples include customer conversations that don't include sensitive info or released marketing materials.	**Scope**: Files & other data assets, Emails, Meetings* **Content marking**: No **Auto-labeling**: No

Label name	Label description for users	Settings
		Group settings: No **Site settings**: No **Auto-labeling for database columns**: None
General \ All Employees (unrestricted)	Organization data that isn't intended for public consumption. If you need to share this content with external partners, confirm with other data owners that it's OK to share and then change the label to General \ Anyone (unrestricted) . Examples include a company internal telephone directory, organizational charts, internal standards, and most internal communication.	**Scope**: Files & other data assets, Emails, Meetings* **Content marking**: No **Auto-labeling**: No **Group settings**: No **Site settings**: No **Auto-labeling for database columns**: None
Confidential	Sensitive business data that could cause damage to the business if shared with unauthorized people. Examples include contracts, security reports, forecast summaries, and sales account data.	**Scope**: Files & other data assets, Emails **Content marking**: No **Auto-labeling**: No **Group settings**: No **Site settings**: No **Auto-labeling for database columns**: None
Confidential \ Anyone (unrestricted)	Confidential data that doesn't need to be encrypted. Use this option with care and appropriate business justification.	This label is selected for client-side auto-labeling and service-side auto-labeling. **Scope**: Files & other data assets, Emails, Meetings*

Label name	Label description for users	Settings
		Content marking: Footer: Classified as Confidential **Auto-labeling**: Recommend that users apply the label **Group settings**: No **Site settings**: No **Auto-labeling for database columns**: None
Confidential \ All Employees	Confidential data that requires protection, which allows all employees full permissions. Data owners can track and revoke content.	This label is selected for client-side auto-labeling and service-side auto-labeling. **Scope**: Files & other data assets, Emails, Meetings* **Encryption**: All users and groups in the org: Co-Author **Content marking**: Footer: Classified as Confidential **Auto-labeling**: Recommend that users apply the label **Group settings**: No **Site settings**: No **Auto-labeling for database columns**: None
Confidential \ Trusted People	Confidential data that can be shared with trusted people inside and outside your organization. These people can also reshare the data as needed.	**Scope**: Files & other data assets, Emails, Meetings*

Label name	Label description for users	Settings
		Encryption: Let users assign permissions: - Encrypt-Only for Outlook - Prompt users in Word, PowerPoint, and Excel **Content marking**: Footer: Classified as Confidential **Auto-labeling**: No **Group settings**: No **Site settings**: No **Auto-labeling for database columns**: None
Highly Confidential	Very sensitive business data that would cause damage to the business if it was shared with unauthorized people. Examples include employee and customer information, passwords, source code, and pre-announced financial reports.	**Scope**: Files & other data assets, Emails **Content marking**: Watermark: HIGHLY CONFIDENTIAL **Auto-labeling**: No **Group settings**: No **Site settings**: No **Auto-labeling for database columns**: None
Highly Confidential \ All Employees	Highly confidential data that allows all employees view, edit, and reply permissions to this content. Data owners can track and revoke content.	**Scope**: Files & other data assets, Emails, Meetings* **Encryption**: All users and groups in the org: Co-Author **Content marking**:

Label name	Label description for users	Settings
		Footer: Classified as Highly Confidential **Auto-labeling**: No **Group settings**: No **Site settings**: No **Auto-labeling for database columns**: None
Highly Confidential \ Specific People	Highly confidential data that requires protection and can be viewed only by people you specify and with the permission level you choose.	**Scope**: Files & other data assets, Emails, Meetings* **Encryption**: Let users assign permissions: - Do Not Forward for Outlook - Prompt users in Word, PowerPoint, and Excel **Content marking**: Footer: Classified as Highly Confidential **Auto-labeling**: No **Group settings**: No **Site settings**: No **Auto-labeling for database columns**: None

* The **Meetings** scope is included for customers starting October 2024 if the tenant has licenses to manually apply a label for scheduled meetings. Additionally:

- The sensitivity label policy includes a default Teams meeting label.
- If the tenant has licenses to manually apply the label to Teams meetings, some of the sensitivity labels also have settings configured to protect these meetings.

> **Note**
>
> The label names and descriptions are automatically available for the following locales: US English, Chinese Simplified and Traditional, French, German, Italian, Japanese, Korean, Portuguese Brazilian, Russian, and Spanish.
>
> If you need additional languages, you can specify your translations by using PowerShell.

For more information about these configuration settings and what sensitivity labels can do, see What sensitivity labels can do.

If you need to edit these default sensitivity labels, see Create and configure sensitivity labels.

Default sensitivity label policy

The default sensitivity label policy makes the labels available for users to start labeling their documents and emails with sensitivity labels. It has the following configuration:

- Publish the default labels to all users in your tenant
- Default label of **General \ All Employees (unrestricted)** for unlabeled documents, email, and meetings
- Users must provide a justification to remove a label or lower its classification

For more information about these policy settings, and other policy settings that are available, see What label policies can do.

If you need to edit these default policy settings, see Publish sensitivity labels by creating a label policy.

When you use these labels in Office apps on Windows, macOS, iOS, and Android, users see new labels within four hours, and within one hour for Word, Excel, and PowerPoint on the web when you refresh the browser. However, you might need to allow up to 24 hours for changes to replicate to all apps and services.

Client-side auto-labeling

The default client-side auto-labeling configuration automatically recommends users apply a sensitivity label when we detect credit card numbers in documents or emails they're working with. As a recommendation rather than automatically applied, this configuration serves as a good

first step for highlighting concerning content and introduces users to the practice of labeling their documents and emails.

Client-side auto-labeling only works for documents and emails in use by the Office apps Word, Excel, PowerPoint, and Outlook.

The default client-side auto-labeling has the following configuration:

- If there are 1-9 instances of credit card numbers found in a document or email, recommend the user applies the sensitivity label **Confidential \ Anyone (unrestricted)**
- If there are 10 or more instances of credit card numbers found in a document or email, recommend the user applies the sensitivity label **Confidential \ All Employees**

> **Note**
>
> If we detected you have your own sensitivity labels published, we'll prompt you to select one of your own labels for auto-labeling and configure it for you.

If you want to edit the client-side auto-labeling configuration, see How to configure auto-labeling for Office apps.

Service-side auto-labeling

Service-side auto-labeling helps label sensitive documents at rest, and emails in transit. The default service-side auto-labeling policy creates policies that run in simulation mode for documents stored in all SharePoint or OneDrive sites, and all emails that are sent via Exchange Online.

In simulation mode, items aren't actually labeled until the policy is turned on. You can manually turn on the policy. Alternatively, if you don't change the default setting, the policy will be automatically turned on for you if there aren't any changes to the policy within a set number of days from when the simulation completes.

In most cases, the number of days before an unedited policy is automatically turned on is 7. However, specific to new customers from June 23, 2022, the initial number of days is 25, and then 7 after the policy is edited.

Simulation mode allows you to preview what items would get labeled when the policy is turned on, so you can have confidence in the labeling feature before you deploy the policy to your

tenant for actual labeling.

The default service-side auto-labeling policies have the following configuration: For all customers:

- If there are 1-9 instances of credit card numbers found in a document or email, apply the sensitivity label **Confidential \ Anyone (unrestricted)**
- If there are 10 or more instances of credit card numbers found in a document or email, apply the sensitivity label **Confidential \ All Employees**

> **Note**
>
> If we detected you have your own sensitivity labels published, we'll prompt you to select one of your own labels for your auto-labeling policy.

There are two auto-labeling policies for each setting. One policy is for the Exchange location, and the other for the SharePoint and OneDrive locations. Although the policies are created at the same time, simulation isn't immediately turned on for SharePoint and OneDrive:

- Exchange location: The auto-labeling policy is created and immediately starts simulation.
- SharePoint and OneDrive locations: The auto-labeling policy is created but waits 25 days before it automatically starts simulation. This delay ensures that there is time for files to be created and saved to these locations.

When the simulation is complete, review the results. If you are happy with them, turn on the policies.

For more information about simulation mode, see Learn about simulation mode.

If you want to edit the service-side auto-labeling policy, see How to configure auto-labeling policies for SharePoint, OneDrive, and Exchange.

Additional resources

To learn more about sensitivity labels and all the capabilities available with Microsoft Purview Information Protection, see the following resources:

- Learn about sensitivity labels
- Protect your data with Microsoft Purview

Last updated on 05/01/2026

Microsoft Purview posture reports overview

> **Note**
>
> Information Protection posture reports, Data Loss Prevention posture reports, and Custom posture reports are in preview. The capabilities, terminology, and UX may change before general availability.

Prerequisites

Permissions

To access **Reports**, your account must belong to one of the following Microsoft 365 role groups, Microsoft Purview roles, or Microsoft Purview role groups.

Microsoft 365 role groups

- Compliance administrator
- Security administrator
- Compliance data administrator
- Global administrator

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

Microsoft Purview roles

To fine-tune access controls, use Microsoft Purview roles. To learn more, see Permissions in the Microsoft Purview portal.

- Information Protection Admin
- Information Protection Analyst
- Information Protection Investigator

- Information Protection Reader

Microsoft Purview role groups

To grant access through role groups, use the following Microsoft Purview role groups. To learn more, see Permissions in the Microsoft Purview portal.

- Information Protection
- Information Protection Admins
- Information Protection Analysts
- Information Protection Investigators
- Information Protection Readers

Posture reports

Microsoft Purview collects and analyzes large volumes of data about how information in your organization is used and protected. This data surfaces through audit logs, Activity explorer, Content explorer, solution-level dashboards, and reports. Posture reports give SOC admins and solution-level admins prebuilt views that turn that data into key insights about your information protection and DLP posture.

Posture reports help you answer questions for executives and operations teams, such as:

- Is Purview doing what we intend it to do?
- Are our policies effective?
- Where are the gaps in our protection strategy?

You find **Posture Reports** in the Microsoft Purview portal:

- **Information Protection** > **Reports**
- **Data Loss Prevention** > **Reports**

> **Note**
>
> The data shown in posture reports is gathered on a 30-day rolling-window basis. Save any reports (as .PDF) that you want to keep for future reference. You can also configure and save your preferred filter settings for future use.

Information Protection posture reports

These reports are in preview.

- Label distribution and adoption
- Auto-labeling policy coverage
- Sensitivity label activity

Data Loss Prevention posture reports

These reports are in preview.

- Most triggered DLP rules and activities
- DLP policies with highest trigger volume
- Top DLP policy violators

> **Note**
>
> DLP policies in simulation mode don't show up in posture reports.

How the Reports tool works

As a Microsoft 365 administrator or compliance administrator, you can evaluate and then tag content in your organization to control where it goes, protect it no matter where it is, and ensure that it's preserved and deleted according to your organization's needs. You do this through the application of sensitivity labels, retention labels, sensitive information types (SIT) classification, and classification by trainable classifiers. There are various ways to do the discovery, evaluation, and tagging, but the end result is that you might have large numbers of documents and emails that are tagged and classified with one or more of these labels. You need to see:

- How labels are used across your tenant and what's done with those items.
- The protections placed on items.
- The most common activities.

The **Reports** tool, available in the **Information Protection** and **Data Loss Prevention (DLP)** solutions, provides visibility into that information and more. The reports include information from both Microsoft 365 source services (Exchange, SharePoint, OneDrive) and non-Microsoft 365 sources.

For example, the **Reports** tool shows you:

- Information on protection policies (preview).

- The trainable classifiers detected most often across your cloud platforms.
- The number of items classified as a sensitive information type and what those classifications are.
- The SITs detected most often across your cloud platforms.
- The top applied sensitivity labels and activities.
- A summary of activities users take on your sensitive content.
- The locations of your sensitive and retained data.
- Summary information on email encryption.

The sections in the **Reports** page are interactive, so you can drill down for more detail. The **Reports** page also links to:

- Activity explorer
- Data explorer
- Content explorer (classic)

Custom posture reports

> **Note**
>
> Custom posture reports are in preview. The capabilities, terminology, and UX may change before general availability.

Custom posture reports work alongside the built-in posture reports to give your team flexibility. Use the built-in reports for ready-made insights that cover common information-protection and DLP scenarios, and create custom reports to assemble the metrics, charts, and filters that answer questions specific to your organization.

Use custom posture reports to answer questions such as:

- Are sensitivity labels being adopted consistently across workloads?
- Where is sensitive content concentrated, and is that concentration growing?
- Are DLP policies reducing risk for specific data types?

To create or open a custom posture report, go to one of the following locations in the Microsoft Purview portal:

- **Information Protection** > **Reports** > **Add report**
- **Data Loss Prevention** > **Reports** > **Add report**

Key concepts

A custom report is a container that holds the cards you use to visualize labeling, classification, and protection activity. Every custom report includes:

- A **report name** and description that state the report's purpose.
- One or more **sections** that group cards by purpose, audience, or theme.
- One or more **cards** within each section that each answer a specific question.

Sections

Sections give a report structure by grouping related cards. They make multi-card reports easier to read and navigate. A report can contain one or more sections, and each section can hold multiple cards.

Add sections before you add cards—sections are the layout framework for the report.

Card types

Cards are the building blocks of a custom report. You can mix and match built-in cards within the same report, or add a custom card.

Expand table

Card type	Description	Use for
Metric card	Displays a single, high-level value or KPI.	Pairing a primary metric with its historical trend. Common uses include adoption, growth, and compliance-health indicators.
Chart card	Provides richer visualizations that reveal patterns and trends.	Distributions, breakdowns, and trend analysis across locations, labels, or workloads.

Custom card configuration

Custom cards let you tailor views to the specific questions your organization needs to answer. They complement built-in cards by adding organization-specific perspectives to the foundation that built-in cards already provide.

Metric-based custom cards support these configuration options:

- **Time range**: Choose any window up to 30 days, extending the 7-day view that built-in metric cards provide.
- **Display format**: **Number** (a raw count), **Percentage** (a proportional view), or **Compound** (a value combined with a trend direction).
- **Filters**: Limit data to specific criteria, such as a sensitivity label, location, or workload.

Chart-based custom cards support these chart types:

- **Vertical bar**: Compares values across categories using vertical bars.
- **Horizontal bar**: Compares values across categories using horizontal bars. Useful when category labels are long.
- **Pie**: Shows the proportional distribution of values across categories.
- **Donut**: Similar to a pie chart, with a central area that improves readability.
- **Line chart**: Shows trends or changes over time.

Apply filters and pivot by grouping dimensions—such as **Activity** or **Sensitive info type**—to tailor chart-based cards to specific scenarios.

Example: configure a custom card

This example tracks adoption of the *Confidential* sensitivity label over the last 30 days.

Expand table

Setting	Value
Card type	Metric card
Metric	Number of items labeled *Confidential*
Time range	Last 30 days (custom)
Display format	**Compound** (total count with trend direction)
Sensitivity label filter	*Confidential*
Workload filter	SharePoint

What the card shows:

- The current total number of items labeled *Confidential*.
- Whether labeling activity is increasing or decreasing over the last 30 days.
- A focused view of adoption for one label and one workload.

Best practices

- Start with clear questions, then choose the cards that answer them.
- Avoid overcrowding a report. A few well-chosen cards are more effective than many.
- Use metric cards for status at a glance and chart cards for deeper analysis.
- Treat custom reports as living assets and iterate as needs evolve.
- For larger reports, save after every two sections.

See also

- View label activity
- View labeled content
- Learn about sensitivity labels
- Learn about retention policies and retention labels
- Learn about sensitive information types
- Sensitive information type entity definitions
- Learn about trainable classifiers (preview)
- Monitor Endpoint DLP device health

To learn how to use data classification to comply with data privacy regulations, see Deploy information protection for data privacy regulations with Microsoft 365.

ⓘ **Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 05/28/2026

Microsoft Purview classifiers overview

Microsoft Purview classifiers help administrators identify sensitive content so they can protect and manage it. The Classifiers page brings classification types together in one experience where you can discover, create, inspect, and manage classifiers.

Search and filtering help you find classifiers. When you create a new classifier, you can view how each classifier type works. This helps you understand the available options and choose a classifier for your scenario.

The **Classified items** count shows how many items were found with that specific classification. You can also see how many policies use a classifier and inspect those dependencies before changing it.

In the Microsoft Purview portal, you can access **Classifiers** under **Information Protection**, **Data Loss Prevention**, or **DSPM** (Data Security Posture Management).

Classification types and capabilities

> **Important**
>
> The new Classifiers experience in Microsoft Purview is in preview and is available alongside the existing classification experiences. Classifiers you create, edit, or delete in either experience are reflected in the other. Features and functionality in the new experience might change before general availability.

Explore the following classification types and related capabilities.

Sensitive information types

Sensitive information types (SITs) identify content that matches patterns of sensitive information, such as social security, credit card, or bank account numbers. You can use preexisting sensitive information types or create your own.

For more information, see Learn about sensitive information types. For information about predefined types, see Sensitive information type entity definitions.

Named entities

Named entities are sensitive information types that use dictionaries and patterns to detect person names, physical addresses, and medical terms and conditions.

For more information, see Learn about named entities.

Exact data match

Exact data match identifies sensitive content by matching exact values.

For more information, see Learn about exact data match based sensitive information types.

Trainable classifiers

Microsoft Purview provides pretrained trainable classifiers that are ready to use. You can also create custom trainable classifiers by providing samples of the content you want to recognize.

For more information, see Learn about trainable classifiers.

Credentials

Credential sensitive information types detect credentials in content. The **All credentials** sensitive information type groups individual credential types to detect credentials from supported services and environments.

For more information, see All credential types.

Fingerprinting

Document fingerprinting recognizes content that is a variation of a template.

For more information, see Document fingerprinting.

Optical character recognition

Optical character recognition (OCR) lets Microsoft Purview scan images for sensitive information. You can enable OCR to extend classification to images without creating separate classifiers.

For setup requirements and supported locations, see Learn about optical character recognition.

On-demand classification

On-demand classification identifies and classifies sensitive content in files at rest, including files that haven't been classified or need updated classification.

For more information, see Create and analyze an on-demand classification scan.

Explore classified content

The Classifiers page helps you manage classifiers. Use Data Explorer to review items with classification matches and see where those items are located. Access to item lists and item contents requires the appropriate Data Explorer permissions.

You can find Data Explorer on the Microsoft Purview portal > **Information Protection** > **Explorers** page.

For supported content, access requirements, and instructions, see Get started with Data Explorer.

Permissions

Access to the Classifiers page can come from a Microsoft Entra administrator role or a Microsoft Purview role group.

> **Note**
>
> As a best practice, always use the role with least privilege to grant access to the Classifiers page.

Microsoft Entra administrator roles

The following Microsoft Entra roles provide access:

- Global administrator
- Compliance administrator
- Security administrator
- Compliance data administrator

Microsoft Purview roles and role groups

In Microsoft Purview, a **role** defines permissions, and a **role group** contains a set of roles. Adding a user to a role group grants the permissions of the roles in that group. For details, see Permissions in the Microsoft Purview portal.

Applicable Microsoft Purview roles

- Information Protection Admin
- Information Protection Analyst
- Information Protection Investigator
- Information Protection Reader

Applicable Microsoft Purview role groups

Use the following role groups to assign access. Their permissions differ; choose the group that matches the user's responsibilities. For assignment guidance, see Permissions in the Microsoft Purview portal.

- Information Protection
- Information Protection Admins
- Information Protection Analysts
- Information Protection Investigators
- Information Protection Readers

Next steps

- Get started with trainable classifiers
- Learn about sensitive information types
- Learn about trainable classifiers
- Explore classified content
- Learn about sensitivity labels
- Learn about retention policies and retention labels

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 09/30/2026

Learn about sensitive information types

Identifying and classifying sensitive items that are under your organization's control is the first step in the Information Protection discipline. Microsoft Purview provides three ways of identifying items so that they can be classified:

- manually, by users
- via automated pattern recognition, as with sensitive information types
- via machine learning

Sensitive information types (SITs) are pattern-based classifiers. They detect sensitive information like social security, credit card, or bank account numbers to identify sensitive items, see Sensitive information type entity definitions for a complete list of all SITs.

Microsoft provides a large number of preconfigured SITs or you can create your own.

Licensing

E5 license is required to make use of credential scanning SITs. For a list of all the credential scanning SITs, see All credentials sensitive information types. This SIT contains all the credential scanning SITs that are available in the portal. Each member of this SIT is a credential scanning SIT and can be used as a standalone. For a list of many Microsoft created SITs, see Sensitive information type entity definitions.

Sensitive information types are used in

- Microsoft Purview Data Loss Prevention policies
- Sensitivity labels
- Retention labels
- Insider risk management
- Communication compliance
- Auto-labeling policies
- Microsoft Priva

Categories of sensitive information types

Built in sensitive information types

Microsoft created these SITs and they show up in the Purview portal by default. These SITs can't be edited, but you can use them as templates by copying them to create custom sensitive information types. See, Sensitive information type entity definitions for a full list of all SITs.

Named entity sensitive information types

Named entity SITs also show up in the Purview portal by default. They detect person names, physical addresses, and medical terms and conditions. They can't be edited or copied. For more information, see Learn about named entities.

Named entity SITs come in two types:

un-bundled

These named entity SITs have a narrower focus, such as a single country or region, or a single class of terms. Use them when you need a data loss prevention (DLP) policy with a narrower detection scope. See, Examples of named entity SITs.

bundled

Bundled named entity SITs detect all possible matches in a class, such as *All physical addresses*. Use them as broad criteria in your DLP policies for detecting sensitive items. See, Examples of named entity SITs.

> **Tip**
>
> You must enable **Advanced classification scanning and protection**if you want to use a bundled SIT in an endpoint DLP policy. This requirement is specific to the combination of bundled SITS andendpoint DLP policies.

Custom sensitive information types

If the preconfigured sensitive information types don't meet your needs, you can create your own custom sensitive information types that you fully define or you can copy one of the built-in ones and modify it. For more information, see

Create a custom sensitive information type in the Microsoft Purview portal.

Exact data match sensitive information types

All exact data match (EDM)-based SITs are created from scratch. You use them to detect items that have exact values, which you define in a database of sensitive information. For more information, see Learn about exact data match based sensitive information types.

Fundamental parts of a sensitive information type

Every sensitive information type (SIT) entity consists of the following fields:

- ***Name:*** Indicates how the sensitive information type is referred to.
- ***Description:*** Explanation of what the sensitive information type is looking for.
- ***Pattern:*** Defines what a SIT detects. It consists of the following components: primary element, supporting elements, confidence level, and proximity.

The following table describes each component of the patterns used in defining sensitive information types.

Expand table

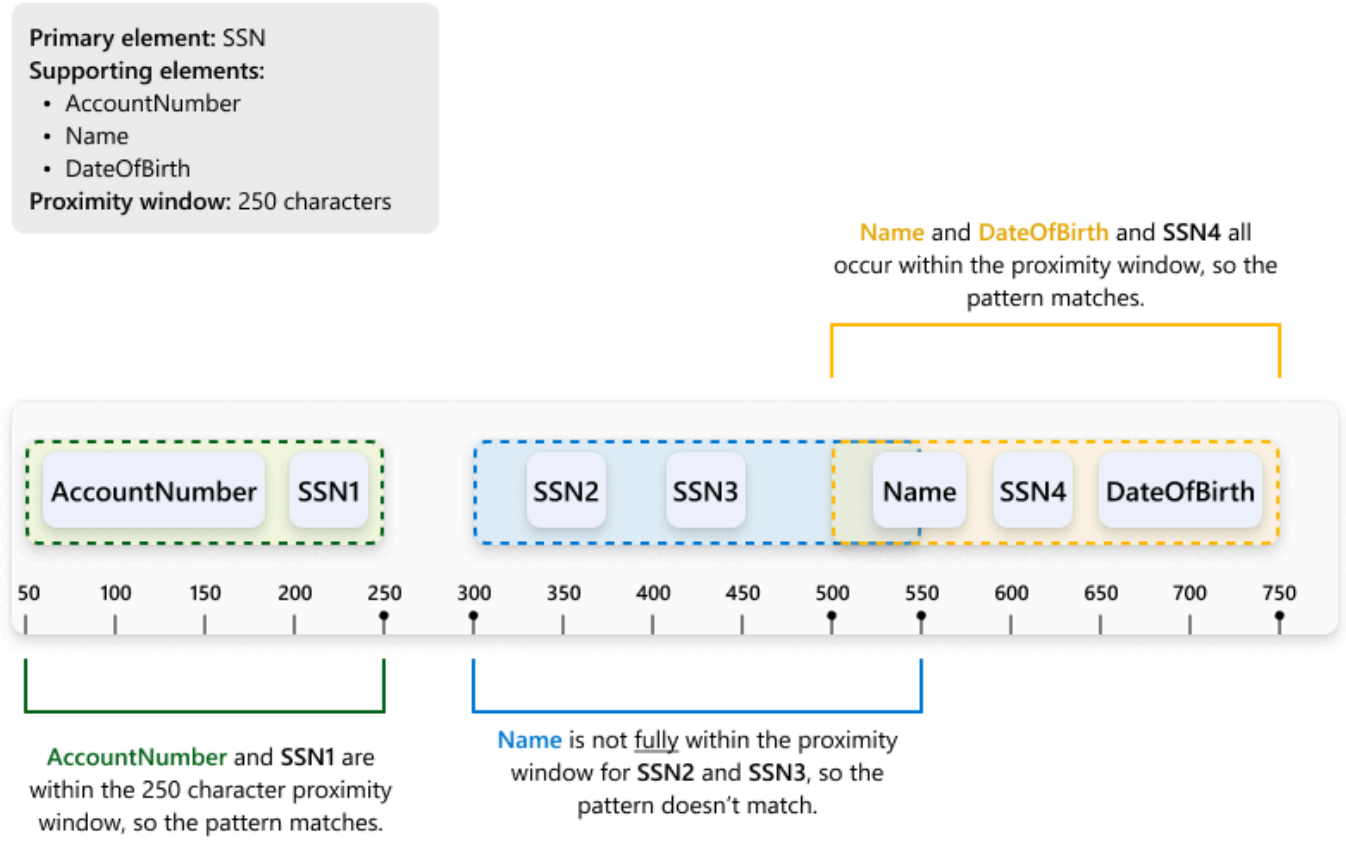
Pattern component	Description
Primary element	The main element that the sensitive information type is looking for. It can be a **regular expression** with or without a checksum validation, a **keyword list**, a **keyword dictionary**, or a **function**.Each of these types of elements can either be selected from the list of existing SITs or can be custom-defined by a user with admin permissions. Once an element is defined, it appears in the list of existing elements, along with those that come built-in.
Supporting element	An element that acts as corroborative evidence. When included, supporting elements help increase the confidence level with respect to the accuracy of detected matches. For example, if the primary element is defined as `SSN` (composed of nine digits), and the keyword *Social Security Number (SSN)* is used as a supporting element when found in proximity to `SSN`, the confidence that the `SSN` detected is truly a Social Security number is higher than if the *Social Security Number (SSN)* keyword is not present. A supporting element can be a regular expression (with or without a checksum validation), a keyword list, or a keyword dictionary.
Confidence Level	There are three confidence levels with respect to detected matches: high, medium, and low. The confidence level reflects how much supporting evidence is detected along with the primary element. The more supporting evidence a detected item contains, the higher the confidence that a matched item contains the sensitive info you're looking for. For more information about confidence levels, see the video included later in this article.
Proximity	Specifies how close a supporting element is to a primary element, in terms of the number of characters between them.

Understanding proximity

The following diagram shows how match detection works with respect to proximity. In this example, the primary element is the `SSN` field and the SIT definition requires that each instance of an `SSN` value must be within a specified proximity to at least one of the following elements:

- `AccountNumber`
- `Name`
- `DateOfBirth`

In the diagram, we see that the data being checked includes three different instances of the `SSN` field: `SSN1`, `SSN2`, `SSN3`, and `SSN4`.



.

To understand how proximity works, let's start by taking a look at some sample detection criteria. Here, were want to detect nine-digit social security numbers. The detection criteria require that a nine-digit regular expression (primary element) is found in conjunction with supporting evidence (among the `AccountNumber`, `Name`, and `DateOfBirth` fields) that appears within 250 characters (the *proximity*).

As illustrated in the diagram, only the primary elements `SSN1` and `SSN4` meet the detection criteria described. Let's take a closer look.

- In the case of `SSN1`, the `AccountNumber` value is within the specified proximity window of 250 characters, so a match is detected.
- In both the cases of `SSN2` and `SSN3`, none of the supporting elements occur within 250 characters of the primary element, so those values aren't detected as a match. However, as you look at the proximity window for `SSN2` in the diagram, you might ask: *Why isn't there a match for `SSN2`? Doesn't the `SSN2` proximity window extend to the `Name` element?* This is a good question. The answer is: *Not quite*. While the proximity window ***extends into*** the `Name` value, it doesn't include the ***entire*** value, so the pattern doesn't match.
- Finally, in the case of `SSN4`, there are two supporting elements within the proximity window, both `Name` and `DateOfBirth`, so this pattern matches as well.

Learn more about confidence levels in this short video.

https://learn-video.azurefd.net/vod/player?id=ace82ff5-4e4c-4542-b0f4-09d8d3335486&locale=en-us&embedUrl=%2Fpurview%2Fsit-sensitive-information-type-learn-about

Example sensitive information type

Argentina national identity (DNI) number

Format

Eight digits separated by periods

Pattern

Eight digits:

- two digits
- a period
- three digits
- a period
- three digits

Checksum

No

Definition

A DLP policy has medium confidence that it has detected this type of sensitive information if, within a proximity of 250 characters:

- The regular expression Regex_argentina_national_id finds content that matches the pattern.
- A keyword from Keyword_argentina_national_id is found.

XML

```
<!-- Argentina National Identity (DNI) Number -->
<Entity id="00aa00aa-bb11-cc22-dd33-44ee44ee44ee" recommendedConfidence="75"
patternsProximity="250">
   <Pattern confidenceLevel="75">
      <IdMatch idRef="Regex_argentina_national_id"/>
      <Match idRef="Keyword_argentina_national_id"/>
  </Pattern>
</Entity>
```

Keywords

Keyword_argentina_national_id

- Argentina National Identity number
- Identity
- Identification National Identity Card
- DNI
- National Registry of Persons (NIC)
- Documento Nacional de Identidad
- Registro Nacional de las Personas
- Identidad
- Identificación

More on confidence levels

In a sensitive information type entity definition, **confidence level** reflects how much supporting evidence is detected in addition to the primary element. The more supporting evidence an item contains, the higher the confidence that a matched item contains the sensitive info you're looking for. For example, matches with a high confidence level contain more supporting evidence in close proximity to the primary element, whereas matches with a low confidence level would contain little to no supporting evidence in close proximity.

A high confidence level returns the fewest false positives but might result in more false negatives. Low or medium confidence levels return more false positives but few to zero false negatives.

- **low confidence**: Matched items contain the fewest false negatives but the most false positives. Low confidence returns all low, medium, and high confidence matches. The low confidence level has a value of 65.
- **medium confidence**: Matched items contain an average number of false positives and false negatives. Medium confidence returns all medium, and high confidence matches. The medium confidence level has a value of 75.
- **high confidence**: Matched items contain the fewest false positives but the most false negatives. High confidence only returns high confidence matches and has a value of 85.

You should use high confidence level patterns with low counts, say five to 10, and low confidence patterns with higher counts, say 20 or more.

> ⓘ **Note**
>
> If you have existing policies or custom sensitive information types (SITs) defined using number-based confidence levels (also known as *accuracy*), they'll automatically be mapped to the three discrete confidence levels; low confidence, medium confidence, and high confidence, across the Security @ Compliance Center UI.
>
> - All policies with minimum accuracy or custom SIT patterns with confidence levels of between 76 and 100 will be mapped to high confidence.
> - All policies with minimum accuracy or custom SIT patterns with confidence levels of between 66 and 75 will be mapped to medium confidence.
> - All policies with minimum accuracy or custom SIT patterns with confidence levels less than or equal to 65 will be mapped to low confidence.

Creating custom sensitive information types

You can choose from several options to create custom sensitive information types.

- **Use the UI** - You can set up a custom sensitive information type using the Purview portal UI. With this method, you can use regular expressions, keywords, and keyword dictionaries. To learn more, see Create a custom sensitive information type.

- **Use EDM** - You can set up custom sensitive information types using Exact Data Match (EDM)-based classification. This method enables you to create a dynamic sensitive information type using a secure database that you can refresh periodically. See Learn about exact data match based sensitive information types.

- **Use PowerShell** - You can set up custom sensitive information types using PowerShell. Although this method is more complex than using the UI, you have more configuration options. See Create a custom sensitive information type in Security & Compliance PowerShell.

Tuning trainable classifiers

Endpoint DLP classifies files based on all sensitive information types available in the tenant, including custom sensitive information types, regardless of their utilization in any DLP policies. This can cause excessive classification traffic if the sensitive information types aren't well tuned and end up matching many files. You should optimize all custom sensitive information types. Do this by removing unused sensitive information types and redesigning SITs if they match most of the files in your organization. For guidance on utilizing SIT Regex validators to tune SITs, see: Sensitive information type REGEX validators and additional check

Double byte character set support

Improved confidence levels are available for immediate use within Microsoft Purview Data Loss Prevention services, information protection, Communication Compliance, data lifecycle management, and records management.

- Information Protection now supports double byte character set languages for:
- Chinese (simplified)
- Chinese (traditional)
- Korean
- Japanese

This support is available for sensitive information types. For more information, see Information protection support for double byte character sets release notes.

Single byte character set support

To detect patterns containing Chinese/Japanese characters and single byte characters or to detect patterns containing Chinese/Japanese and English, define two variants of the keyword or regex.

For example, to detect a keyword like "机密的document", use two variants of the keyword; one with a space between the Japanese and English text and another without a space between the Japanese and English text. So, the keywords to be added in the SIT should be "机密的 document"

and "机密的document". Similarly, to detect a phrase "東京オリンピック2020", two variants should be used; "東京オリンピック 2020" and "東京オリンピック2020".

Along with Chinese/Japanese/double byte characters, if the list of keywords/phrases also contains non-Chinese/Japanese words also (for instance, English only), you should create two dictionaries/keyword lists. One for keywords containing Chinese/Japanese/double byte characters and another one for English-only keywords. For example, if you want to create a keyword dictionary/list with three phrases "Highly confidential", "機密性が高い" and "机密的document", the you should create two keyword lists.

- Highly confidential
- 機密性が高い, 机密的document and 机密的 document While creating a regex using a double byte hyphen or a double byte period, make sure to escape both the characters like you would escape a hyphen or period in a regex. Here's a sample regex for reference: `(?<!\d)([4][0-9]{3}[\-?\-\t]*[0-9]{4}`

> We recommend using string match instead of word match in a keyword list.

Detection differences across workloads

The same sensitive information type can produce different results depending on where content is evaluated—for example, email in transit (Exchange), files at rest (SharePoint and OneDrive), and files on managed devices (Endpoint DLP). Endpoint DLP classifies files against *all* sensitive information types available in the tenant, and using a bundled SIT in an endpoint DLP policy requires Advanced classification scanning and protection.

Before you rely on a custom SIT in production, test it and validate detection in each workload where you intend to use it—for example, a mail flow rule, a SharePoint or OneDrive policy, and an endpoint policy. Detection can also be affected by content-extraction limits such as maximum file size and the number of characters scanned; see DLP platform considerations.

Test sensitive information type

You can test the SIT by uploading a sample file. The test results show the number of matches for each confidence level. You can test built-in SITs, custom SITs, trainable classifiers, and exact data match.

Test Built-in and Custom sensitive information type

Test exact data match sensitive information type.

To test any custom or default SIT tenant, there must be at least one Exchange Online license added to the tenant. Otherwise, the Test SIT option will be greyed out when any SIT is selected.

Provide match/not a match accuracy feedback in sensitive info types

You can view the number of matches a SIT has in **Sensitive info types** and **Content explorer**. You can also provide feedback on whether an item is actually a match or not using the **Match**, **Not a Match** feedback mechanism and use that feedback to tune your SITs. For more information, see Increase classifier accuracy.

For further information

- Sensitive information type entity definitions
- Create a custom sensitive information type
- Create a custom sensitive information type in PowerShell

To learn how to use sensitive information types to comply with data privacy regulations, see Deploy information protection for data privacy regulations with Microsoft 365 (aka.ms/m365dataprivacy).

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 07/24/2026

Create custom sensitive information types

If the preconfigured sensitive information types (SITs) don't meet your needs, you can create and define customized SITs that meet your needs. You can also copy and then edit a built-in SIT.

The custom SITs are added to the `Microsoft.SCCManaged.CustomRulePack` rule package.

There are two methods for creating a new SIT:

- Create a new SIT from scratch
- Copy and modify an existing SIT

Before you begin

- You should be familiar with sensitive information types and what they're composed of. To get this understanding, see, Learn about sensitive information types. It's critical to understand the roles of:
  - regular expressions - Microsoft 365 sensitive information types uses the Boost.RegEx 5.1.3 engine
  - keyword lists - you can create your own as you define your sensitive information type or choose from existing keyword lists
  - keyword dictionary
  - Sensitive information type functions
  - confidence levels
- Familiarize yourself with Sensitive information type limits.

For information on licensing, see

- Microsoft 365 Enterprise Plans
- Microsoft 365 Service Descriptions

> **Important**
>
> Microsoft Customer Service & Support can't assist with creating custom classifications or regular expression patterns. Support engineers can provide limited support for the feature, such as, providing sample regular expression patterns for simulation purposes, or helping to troubleshoot an existing regular expression pattern that's not triggering as expected.

However, they can't provide assurances that any custom content-matching development fulfills your requirements or obligations.

Create a custom SIT from scratch

> **Note**
>
> Microsoft Purview supports creating custom SITs that use double-byte character languages, such as Chinese, Japanese, and Korean. Because these languages don't use delimiters the way that single-byte languages do, Purview adds a space between each word in languages that use double-byte characters. It also removes special characters, such as punctuation.

Use the following procedure to fully define a brand new sensitive information type.

1. Sign in to the Microsoft Purview portal.
2. In the Microsoft Purview portal, navigate to **Information Protection** > **Classifiers** > **Sensitive info types** and choose **Create sensitive info type**.
3. Fill in values for **Name** and **Description** and choose **Next**.
4. Choose **Create pattern**. You can create multiple patterns, each with different elements and confidence levels, as you define your new sensitive information type.
5. Choose the default confidence level for the pattern. The values are **Low confidence**, **Medium confidence**, and **High confidence**.
6. Choose and define the **Primary element**. The primary element can be a **Regular expression** with an optional validator, a **Keyword list**, a **Keyword dictionary**, or one of the preconfigured **Functions**. For more information on the SIT functions used for data loss prevention, see Sensitive information type functions. For more information on the date and the checksum validators, see Sensitive Information Type regular expression validators.

   > **Important**
   >
   > Custom SIT regex patterns must follow these rules:
   >
   > - **Don't use positional anchors** like `^` and `$`. When content is scanned, there are no guarantees about where in the content the starting and ending anchors correspond to.

- Use **one primary capturing group** (the only capturing group), for example: `(?:prefix)(primary capturing group)(?:suffix)` (use noncapturing groups for any additional grouping).
- Handle all match variants **inside** that single group using `|` (alternation). Multiple capturing groups separated by `|` at the top level are **not supported** and are blocked during validation.

Invalid pattern (three top-level capturing groups):

```
(?:[\s,;])([A-Z]{2}[0-9]{3})(?:[\s,;])|(?:[\s,;])([A-Z]{2}[A-Z]{4}[0-9])(?:[\s,;])|(?:[\s,;])([A-Z]{2}[0-9]{5})(?:[\s,;])
```

Valid pattern (single capturing group with alternation):

```
(?:[\s,;])([A-Z]{2}[0-9]{3}|[A-Z]{2}[A-Z]{4}[0-9]|[A-Z]{2}[0-9]{5})(?:[\s,;])
```

For more information on limits, see Sensitive information type limits.

7. Fill in a value for **Character proximity**.
8. (Optional) Add supporting elements if you have any. Supporting elements can be a regular expression with an optional validator, a keyword list, a keyword dictionary or one of the predefined functions. Supporting elements can have their own **Character proximity** configuration.
9. (Optional) Add any **sensitive information type additional checks** from the list of available checks.
10. Choose **Create**.
11. Choose **Next**.
12. Choose the **recommended confidence level** for this sensitive information type.
13. Check your settings and choose **Save**.

> ⓘ **Important**
>
> Microsoft 365 uses the search crawler to identify and classify sensitive information in SharePoint and OneDrive sites. To identify your new custom sensitive information type in existing content, the content must be recrawled. Content is crawled based on a schedule, but you can manually recrawl content for a site collection, list, or library. For

more information, see Manually request crawling and reindexing of a site, a library, or a list.

14. The **Sensitive info types** tab of the **Classifiers** page, lists all of the sensitive information types. Choose **Refresh** and then or use the search tool or browse the list to find your new SIT.

Copy and modify an existing SIT

This procedure explains how to copy and modify an existing SIT using the Purview Portal.

Alternatively, you can copy and modify custom SITs using PowerShell and using Purview's Exact Data Match (EDM) capabilities. To learn more about those methods, see:

- Create a custom sensitive information type in Microsoft Purview PowerShell
- Learn about exact data match based sensitive information types

> **Note**
>
> These SITs can't be copied:
>
> - Canada driver's license number
> - EU driver's license number
> - EU national identification number
> - EU passport number
> - EU social security number or equivalent identification
> - EU tax identification number
> - International classification of diseases (ICD-10-CM)
> - International classification of diseases (ICD-9-CM)
> - U.S. driver's license number

When you have 10 text processors or restricted text processors, such as keyword lists, regex, keyword dictionaries, or functions, you won't be able to copy the SIT from the UI directly. Instead, you need to use PowerShell for this task. To learn more about how to clone or copy your SIT using PowerShell, see Modify a custom sensitive information type using PowerShell

Copy and modify an existing SIT using the Microsoft Purview portal

1. Sign in to the Microsoft Purview portal.

2. **Information Protection** > **Classifiers** > **Sensitive info types** and select the sensitive information type that you want to copy.

3. The overview page for the sensitive information type opens. Choose **Copy**. When the copy is ready, a message stating that the copy was created appears with an option to edit it. Choose **Yes**.

4. Give your new sensitive information type a new **Name** and **Description**.

5. You can choose to create a new pattern, or edit or remove some or all of the existing patterns.
   a. To create a new pattern, choose **Create**.
   b. To edit an existing pattern, choose the **Edit** (pencil) icon next to the pattern you want to change.
   c. To remove a pattern, choose the **Delete** icon next to the pattern you want to remove.

6. When creating or editing a pattern, choose the default confidence level for the pattern. The values are **Low confidence**, **Medium confidence**, and **High confidence**.

7. Choose and define **Primary element**. The primary element can be a **Regular expression**, a **Keyword list**, a **Keyword dictionary**, or one of the preconfigured **Functions**. See, Sensitive information type functions.

8. Fill in a value for **Character proximity**.

9. (Optional) If you have **Supporting elements** or any sensitive information type additional checks you want to run, add them. If needed, you can organize your **Supporting elements** into groups.

10. If you're creating a new pattern, choose **Create**. If you're editing an existing pattern, choose **Update**.

11. Choose **Next**.

12. Confirm the confidence level selection for this sensitive information type and then choose **Next**.

13. Review your settings and then choose **Save**.

14. Your new sensitive information type is created. At the confirmation message, choose **Done*

> ⓘ **Note**

Microsoft Purview Information Protection supports double byte character set languages for:

- Chinese (simplified)
- Chinese (traditional)
- Korean
- Japanese

This support is available for sensitive information types. For more information, see Information protection support for double byte character sets release notes (preview).

Tip

To detect patterns containing Chinese/Japanese characters and single byte characters, or to detect patterns containing Chinese/Japanese and English, define two variants of the keyword or regex.

- For example, to detect a keyword like "机密的document", use two variants of the keyword; one with a space between the Japanese and English text and another without a space between the Japanese and English text. So, the keywords to be added in the SIT should be "机密的 document" and "机密的document". Similarly, to detect a phrase "東京オリンピック2020", two variants should be used; "東京オリンピック 2020" and "東京オリンピック2020".

Along with Chinese/Japanese/double byte characters, if the list of keywords/phrases also contain non Chinese/Japanese words also (for instance, English only), creating two dictionaries/keyword lists is recommended. Create one for keywords containing Chinese/Japanese/double byte characters and another for English-only.

- For example, if you want to create a keyword dictionary/list with three phrases "Highly confidential", "機密性が高い", and "机密的document", you should create two keyword lists.

    1. Highly confidential
    2. 機密性が高い, 机密的document and 机密的 document

While creating a regex using a double byte hyphen or a double byte period, make sure to escape both the characters in the same way that you would escape a hyphen or period in a regex. Here's a sample regex for reference:

```
(?<!\d)([4][0-9]{3}[\-?\-\t]*[0-9]{4})
```

Double-byte special characters shouldn't be used in the keyword.

We recommend using a string match instead of a word match in a keyword list.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Test sensitive information types

At any time during the life of a sensitive information type (SIT), you can test how it performs. This article explains how to do that in the Microsoft Purview portal.

Before you begin

Before you test a sensitive information type, make sure you have the following background knowledge and preparation.

- You should be familiar with sensitive information types and what they're composed of. For more information, see, Learn about sensitive information types.
- It's critical to understand the roles of:
  - regular expressions - Microsoft 365 sensitive information types uses the Boost.RegEx 5.1.3 engine
  - keyword lists - you can create your own as you define your sensitive information type or choose from existing keyword lists
  - keyword dictionary
  - Sensitive information type functions
  - confidence levels
- Familiarize yourself with Sensitive information type limits.

> **Important**
>
> Microsoft Customer Service & Support can't assist with creating custom classifications or regular expression patterns. Support engineers can provide limited support for the feature, such as, providing sample regular expression patterns for simulation purposes, or helping to troubleshoot an existing regular expression pattern that's not triggering as expected. However, they can't provide assurances that any custom content-matching development fulfills your requirements or obligations.

Permissions and role groups

The account you use to perform tests of SITs must be assigned one of these roles in the Microsoft Purview *and* one of these roles or role groups in Exchange.

Role groups assignments in the Microsoft Purview portal

The account you use must be assigned one of these role groups in the Microsoft Purview portal:

- Compliance Administrator
- Compliance Data Administrator
- Security Administrator
- Communication Compliance Admins
- Information Protection Admins
- Information Protection Investigators
- Organization Management

To assign a user to a role group in the Microsoft Purview portal, follow the steps in Add users or groups to a Microsoft Purview built-in role group.

Microsoft Entra Roles in Exchange Online

The following roles can be assigned either in the Exchange Admin center or in Microsoft Entra.

- Communication Compliance Admin
- Communication Compliance Investigation
- Compliance Admin
- Data Loss Prevention
- Information Protection Admin
- Information Protection Investigator
- Information Protection Reader
- Insider Risk Management Admin
- Insider Risk Management Investigation
- Privacy Management Admin
- Privacy Management Investigation
- Security Admin

Microsoft Entra Role Groups in Exchange Online

The following Exchange Online role groups satisfy this requirement:

- Communication Compliance
- Communication Compliance Administrators
- Compliance Administrator
- Organization Management
- Compliance Management
- Information Protection

- Information Protection Admins
- Information Protection Investigators
- Information Protection Readers
- Insider Risk Management
- Insider Risk Management Admins
- Insider Risk Management Investigators
- Privacy Management
- Privacy Management Administrators
- Privacy Management Investigators
- Security Administrator

Enable access to simulation mode

To enable access to simulation mode, your account must be assigned one of the following role groups:

Applies to these role groups:

- Communication Compliance Admins
- Information Protection Admins
- Information Protection Investigators
- Organization Management

Test the effects of a sensitive information type

You can test the effects of any sensitive information type in the **Sensitive info types** list. Test each sensitive information type that you create before using it in a policy.

This uploaded-file test is different from auto-labeling policy simulation. It tests the SIT against the file you upload; it doesn't refresh classification for existing SharePoint or OneDrive files. A policy simulation uses available classification results for those files. See Auto-labeling classification and simulation concepts for how to interpret differences and when classification needs updating.

> ⓘ **Note**
>
> This procedure is only supported for unencrypted files.

1. Prepare two files, for example, two Word documents. One should have content that matches the elements you specified in your sensitive information type. The other should

have content that doesn't match.

2. Sign in to the Microsoft Purview portal > **Information Protection** > **Classifiers** > **Sensitive info types** and choose the sensitive information type from the list to open the details pane. Choose **Test**.
3. Upload a file and choose **Test**. (You can only upload and test one file at a time.)
4. On the **Match results** page, review the results and choose **Finish**.
5. Perform the same steps for the second file.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 10/14/2026

Customize a built-in sensitive information type

When looking for sensitive information in content, you need to describe that information in what's called a *rule*. Microsoft Purview Data Loss Prevention (DLP) includes rules for the most common sensitive information types (SITs). You can use these rules right away. To use them, you must include them in a policy. You can adjust these built-in rules to meet your organization's specific needs by creating a custom sensitive information type. This topic shows you how to customize the XML file that contains the existing rule collection so you can detect a wider range of potential credit card information.

You can take this example, which modifies the Credit Card Number SIT, and apply this approach to other built-in sensitive information types. For a list of default SITs and XML definitions, see Sensitive information type entity definitions.

Export the XML file of the current rules

To export the XML, you need to connect to Security & Compliance PowerShell.

1. In the PowerShell, type the following to display your organization's rules on screen. If you haven't created your own, you'll only see the default, built-in rules, labeled "Microsoft Rule Package."

   PowerShell

   ```
   Get-DlpSensitiveInformationTypeRulePackage
   ```

2. Store your organization's rules in a variable by typing the following. Storing something in a variable makes it easily available later in a format that works for PowerShell commands.

   PowerShell

   ```
   $ruleCollections = Get-DlpSensitiveInformationTypeRulePackage
   ```

3. Make a formatted XML file with all that data by typing the following.

   PowerShell

```
[System.IO.File]::WriteAllBytes('C:\custompath\exportedRules.xml',
$ruleCollections.SerializedClassificationRuleCollection)
```

> **Important**
>
> Make sure that you use the file location where your rule pack is actually stored. `C:\custompath\` is a placeholder.

Find the rule that you want to modify in the XML

The cmdlets above exported the entire *rule collection*, which includes the default rules that Microsoft provides. Next, you'll need to look specifically for the Credit Card Number rule that you want to modify.

1. Use a text editor to open the exported XML file (for example, `exportedRules.xml`) that you created in the preceding procedure.

2. Scroll down to the `<Rules>` tag, which marks the start of the rules section. Because this XML file contains the information for the entire rule collection, it contains other information at the top that you need to scroll past to get to the rules.

3. Look for *Func_credit_card* to find the Credit Card Number rule definition. In the XML, rule names can't contain spaces, so the spaces are usually replaced with underscores, and rule names are sometimes abbreviated. An example of this is the U.S. Social Security number rule, which is abbreviated *SSN*. The XML for the Credit Card Number rule should look like the following code sample:

XML

```
<Entity id="50842eb7-edc8-4019-85dd-5a5c1f2bb085"
       patternsProximity="300" recommendedConfidence="85">
      <Pattern confidenceLevel="85">
       <IdMatch idRef="Func_credit_card" />
        <Any minMatches="1">
          <Match idRef="Keyword_cc_verification" />
          <Match idRef="Keyword_cc_name" />
          <Match idRef="Func_expiration_date" />
        </Any>
      </Pattern>
    </Entity>
```

Now that you have located the Credit Card Number rule definition in the XML, you can customize the rule's XML to meet your needs. For a refresher on the XML definitions, see the Term glossary later in this article.

Modify the XML and create a new sensitive information type

First, you need to create a new sensitive information type because you can't directly modify the default rules. You can do a wide variety of things with custom sensitive information types, which are outlined in Create a custom sensitive information type in Security & Compliance PowerShell. For this example, we'll keep it simple and only remove corroborative evidence and add keywords to the Credit Card Number rule.

All XML rule definitions are built on the following general template. You need to copy and paste the Credit Card Number definition XML in the template, modify some values (notice the ". . ." placeholders in the following example), and then upload the modified XML as a new rule that can be used in policies.

XML

```
<?xml version="1.0" encoding="utf-16"?>
<RulePackage xmlns="https://schemas.microsoft.com/office/2011/mce">
  <RulePack id=". . .">
    <Version major="1" minor="0" build="0" revision="0" />
    <Publisher id=". . ." />
    <Details defaultLangCode=". . .">
      <LocalizedDetails langcode=" . . . ">
         <PublisherName>. . .</PublisherName>
         <Name>. . .</Name>
         <Description>. . .</Description>
      </LocalizedDetails>
    </Details>
  </RulePack>

 <Rules>
   <!-- Paste the Credit Card Number rule definition here.-->
      <LocalizedStrings>
         <Resource idRef=". . .">
           <Name default="true" langcode=" . . . ">. . .</Name>
           <Description default="true" langcode=". . .">  . . .</Description>
         </Resource>
      </LocalizedStrings>
   </Rules>
</RulePackage>
```

Now, you have something that looks similar to the following XML. Because rule packages and rules are identified by their unique GUIDs, you need to generate two GUIDs: one for the rule package and one to replace the GUID for the Credit Card Number rule. The GUID for the entity ID in the following code sample is the GUID for the built-in rule definition, which you need to replace with a new GUID. There are several ways to generate GUIDs, but you can do it easily in PowerShell by typing **[guid]::NewGuid()**.

XML

```
<?xml version="1.0" encoding="utf-16"?>
<RulePackage xmlns="https://schemas.microsoft.com/office/2011/mce">
  <RulePack id="8aac8390-e99f-4487-8d16-7f0cdee8defc">
    <Version major="1" minor="0" build="0" revision="0" />
    <Publisher id="..." />
    <Details defaultLangCode="en">
      <LocalizedDetails langcode="en">
        <PublisherName>Contoso Ltd.</PublisherName>
        <Name>Financial Information</Name>
        <Description>Modified versions of the Microsoft rule package</Description>
      </LocalizedDetails>
    </Details>
  </RulePack>

 <Rules>
    <Entity id="00aa00aa-bb11-cc22-dd33-44ee44ee44ee"
       patternsProximity="300" recommendedConfidence="85">
      <Pattern confidenceLevel="85">
        <IdMatch idRef="Func_credit_card" />
        <Any minMatches="1">
          <Match idRef="Keyword_cc_verification" />
          <Match idRef="Keyword_cc_name" />
          <Match idRef="Func_expiration_date" />
        </Any>
      </Pattern>
    </Entity>
      <LocalizedStrings>
         <Resource idRef="00aa00aa-bb11-cc22-dd33-44ee44ee44ee">
<!-- This is the GUID for the preceding Credit Card Number entity because the
following text is for that Entity. -->
           <Name default="true" langcode="en-us">Modified Credit Card Number</Name>
           <Description default="true" langcode="en-us">Credit Card Number that looks
for additional keywords, and another version of Credit Card Number that doesn't
require keywords (but has a lower confidence level)</Description>
         </Resource>
      </LocalizedStrings>
   </Rules>
</RulePackage>
```

Remove the corroborative evidence requirement from a sensitive information type

Now you have a new sensitive information type that you're able to upload to the Microsoft Purview portal. The next step is to make the rule more specific. Modify the rule so that it only looks for a 16-digit number that passes the checksum but that doesn't require additional (corroborative) evidence, such as keywords. To require only a 16-digit number that passes the checksum, remove the part of the XML that looks for corroborative evidence. Corroborative evidence is very helpful in reducing false positives. When detecting credit card numbers, there are usually certain keywords or an expiration date near the credit card number. If you remove that evidence, you should also adjust how confident you are that you found a credit card number by lowering the `confidenceLevel`, which is 85 in the example.

XML

```
<Entity id="00aa00aa-bb11-cc22-dd33-44ee44ee44ee" patternsProximity="300"
      <Pattern confidenceLevel="85">
        <IdMatch idRef="Func_credit_card" />
      </Pattern>
    </Entity>
```

Look for keywords that are specific to your organization

You might want to require corroborative evidence but want different or additional keywords, and perhaps you want to change where to look for that evidence. You can adjust the `patternsProximity` to expand or shrink the window for corroborative evidence around the 16-digit number. To add your own keywords, you must define a keyword list and reference it within your rule. The following XML adds the keywords "company card" and "Contoso card", so that any message that contains those phrases within 150 characters of a credit card number will be identified as a credit card number.

XML

```
<Rules>
<! -- Modify the patternsProximity to be "150" rather than "300." -->
    <Entity id="00aa00aa-bb11-cc22-dd33-44ee44ee44ee" patternsProximity="150"
recommendedConfidence="85">
      <Pattern confidenceLevel="85">
        <IdMatch idRef="Func_credit_card" />
        <Any minMatches="1">
          <Match idRef="Keyword_cc_verification" />
```

```
          <Match idRef="Keyword_cc_name" />
<!-- Add the following XML, which references the keywords at the end of the XML
sample. -->
          <Match idRef="My_Additional_Keywords" />
          <Match idRef="Func_expiration_date" />
        </Any>
      </Pattern>
    </Entity>
<!-- Add the following XML, and update the information inside the <Term> tags with the
keywords that you want to detect. -->
    <Keyword id="My_Additional_Keywords">
      <Group matchStyle="word">
        <Term caseSensitive="false">company card</Term>
        <Term caseSensitive="false">Contoso card</Term>
      </Group>
    </Keyword>
```

Upload your rule

Before you begin, make sure you can connect to Security & Compliance PowerShell.

To upload your rule, you need to do the following.

1. Save it as an .xml file with Unicode encoding. Saving the file with Unicode encoding is important because the rule won't work if the file is saved with a different encoding.

2. Connect to Security & Compliance PowerShell.

3. In the PowerShell, type the following.

   PowerShell

   ```
   New-DlpSensitiveInformationTypeRulePackage -FileData
   ([System.IO.File]::ReadAllBytes('C:\custompath\MyNewRulePack.xml'))
   ```

   > **ⓘ Important**
   >
   > Make sure that you use the file location where your rule pack is actually stored. `C:\custompath\` is a placeholder.

4. To confirm, type Y, and then press **Enter**.

5. Verify the display name of your new rule and that it was uploaded, by entering:

   PowerShell

```
Get-DlpSensitiveInformationType
```

To start using the new rule to detect sensitive information, you need to add the rule to a DLP policy. To learn how to add the rule to a policy, see Create and Deploy data loss prevention policies.

Term glossary

These are the definitions for the terms you encountered during this procedure.

Expand table

Term	Definition
Entity	*Entities* are what we call sensitive information types, such as credit card numbers. Each entity has a unique GUID as its ID. If you copy a GUID and search for it in the XML, you'll find the XML rule definition and all the localized translations of that XML rule. You can also find this definition by locating the GUID for the translation and then searching for that GUID.
Functions	The XML file references `Func_credit_card`, which is a function in compiled code. Functions are used to run complex regexes and verify that checksums match for our built-in rules. Because this happens in the code, some of the variables don't appear in the XML file.
IdMatch	This is the identifier that the pattern is to trying to match—for example, a credit card number.
Keyword lists	The XML file also references `keyword_cc_verification` and `keyword_cc_name`, which are lists of keywords that we are looking to match within the `patternsProximity` for the entity. These aren't currently displayed in the XML.
Pattern	The *pattern* contains the list of what the sensitive type is looking for. This includes keywords, regexes, and internal functions, which perform tasks like verifying checksums. Sensitive information types can have multiple patterns with unique confidence levels. This is useful when creating a sensitive information type that returns a high confidence if corroborative evidence is found and a lower confidence if little or no corroborative evidence is found.
Pattern confidenceLevel	This is the level of confidence that the DLP engine found a match. This level of confidence is associated with a match for the pattern if the pattern's requirements are met. This is the confidence measure you should consider when using Exchange mail flow rules (also known as transport rules).

Term	Definition
patternsProximity	When we find what looks like a credit card number pattern, `patternsProximity` is the distance around that number where we'll look for corroborative evidence.
recommendedConfidence	This is the confidence level we recommend for this rule. The recommended confidence level applies to entities and affinities. For entities, this number is never evaluated against the `confidenceLevel` for the pattern. It's merely a suggestion to help you choose a confidence level if you want to apply one. For affinities, the `confidenceLevel` of the pattern must be higher than the `recommendedConfidence` number for a mail flow rule action to be invoked. The `recommendedConfidence` is the default confidence level used in mail flow rules that invokes an action. If you want, you can manually change the mail flow rule to be invoked based off the pattern's confidence level, instead.

Related content

- Sensitive information type entity definitions
- Create a custom sensitive information type
- Learn about data loss prevention

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Manage custom sensitive information types

This article walks you through the steps to modify and remove an existing custom sensitive information type.

Modify custom sensitive information types

To modify a custom sensitive information type, perform the following steps:

1. Sign in to the Microsoft Purview portal > **Information Protection** > **Classifiers** > **Sensitive info types** and choose the sensitive information type from the list that you want to modify.
2. Choose **Edit**.
3. You can add other patterns, with unique primary and supporting elements, confidence levels, character proximity, and **additional checks** or edit/remove the existing ones.

Delete custom sensitive information types

> **Note**
>
> You can only remove custom sensitive information types; you can't remove built-in sensitive information types.

> **Important**
>
> Before your remove a custom sensitive information type, verify that no DLP policies or Exchange mail flow rules (also known as *transport rules*) still reference the custom sensitive information type that you plan to remove.

1. Sign in to the Microsoft Purview portal > **Information Protection** > **Classifiers** > **Sensitive info types** and choose the sensitive information type that you want to delete.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Sensitive information type regex validators and additional checks

> **Important**
>
> Microsoft Customer Service & Support can't assist with creating custom classifications or regular expression patterns. Support engineers can provide limited support for the feature, such as providing sample regular expression patterns for testing the feature, or assisting with troubleshooting an existing regular expression pattern that's not triggering as expected. However, support engineers can't assure you that any custom content-matching development fulfills your requirements or obligations.

Sensitive Information Type regular expression validators

Checksum validator

To run a checksum on a digit in a regular expression, you can use the *checksum validator*. For example, if you need to create a SIT for an eight-digit license number where the last digit is a checksum digit validated using a mod 9 calculation, set up the checksum algorithm as shown in the following example:

Console

```
Sum = digit 1 * Weight 1 + digit 2 * weight 2 + digit 3 * weight 3 + digit 4 * weight
4 + digit 5 * weight 5 + digit 6 * weight 6 + digit 7 * weight 7 + digit 8 * weight 8
Mod value = Sum % 9
If Mod value == digit 8
    Account number is valid
If Mod value != digit 8
    Account number is invalid
```

1. Define the primary element with a regular expression that matches any eight-digit number:

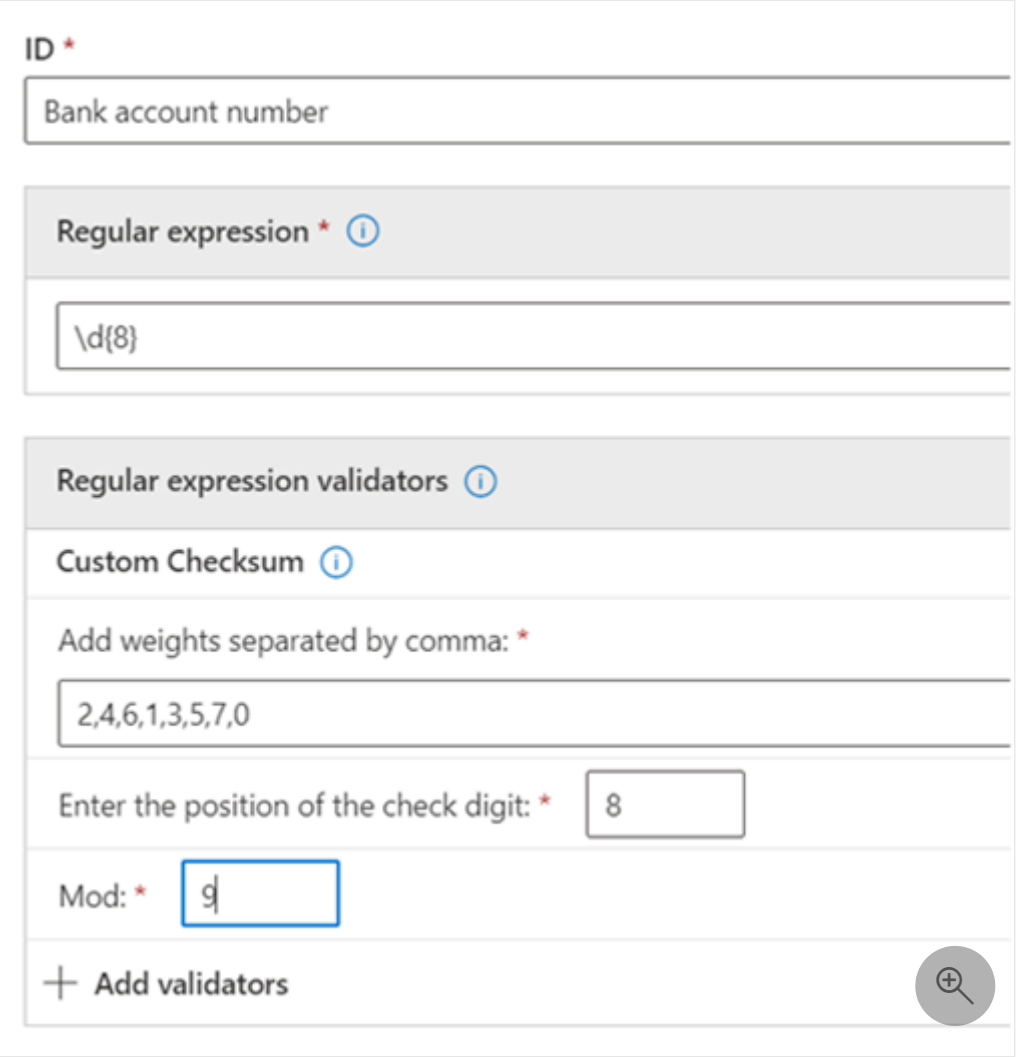
   Console

   ```
   \d{8}
   ```

2. Add the checksum validator.

3. Add the weight values separated by commas, the position of the check digit, and the mod value. For more information on the Modulo operation, see Modulo operation.

> **Note**
>
> If the check digit isn't part of the checksum calculation, use 0 as the weight for the check digit. For example, in the eight-digit license number example, weight 8 is equal to 0 if the check digit won't be used for calculating the check digit.



Parameters

The following parameters define how the checksum validator performs its calculation:

- **Weights:** To define the series of numbers with which each digit starting from position 1 to last position of the regex needs to be multiplied. This calculates the sum product. Weight

positions refer to the order of the digits only, it doesn't consider any nondigit characters like dashes.
- **Mod:** Perform Modulo operation on the result from previous operation.
- **ModCoefficient:** Perform addition or subtraction on the modulo result.
- **CheckDigit:** Define the position of the check digit with which the calculated number will be compared against.

Advanced Checksum Validator

Advanced Checksum can be used without the need of scripting, by using parameters such as PositionBasedUpdate, UseAscii, MultiDigitResult, CheckDigitValue.

- **Digit Replacement Before Computation:** Define rules to replace digits based on position or value before checksum calculation.
- **Letter-to-ASCII Conversion:** Nondigit characters can now be converted to their ASCII values instead of being ignored, enabling checksum support for alphanumeric inputs.
- **Single-Digit Reduction of Multi-Digit Results:** Intermediate results can now be reduced to a single digit by summing their digits (12 → 1+2 = 3), allowing for more compact and consistent outputs.
- **Post-Processing of Two-Digit Results:** Apply mathematical operations like division or modulo to two-digit results to derive final values.
- **Exclusion of Specific Check Digit Values:** Define a list of disallowed check digit values. If the computed result matches one, the system modifies the input and rerun the checksum logic.
- **Final Check Digit Substitution:** Post-computation, specific check digit values can be substituted with alternatives.

Advanced Checksum Validator Parameters

The following parameters control advanced checksum validator behavior:

- **UseAscii:** Replace alphabets with their ascii value
- **PositionBasedUpdate:** Pre checksum computation. We update the digits based on the attributes, match-position-replacewith.
- **CheckDigitValue:** Post checksum computation. If checksum calculated is part of repeat list, perform defined operation on it.

- ltiDigitResult:** Post/Intermediate checksum computation. If Post/intermediate calculation result is of multiple digits, perform the defined operation on it until it's a single digit.

For example, in the XML below, we have passed the following parameters: weights, mod, checkdigit, and ascii.

Console

```
<Validators id="Validator_test_id_card_number">
<Validator type="Checksum">
<Param name="Weights">1,1,2,1,1,1,1,1,1,1,1,1,1,1,1,1</Param>
<Param name="Mod">9</Param>
<Param name="CheckDigit">7</Param>
<Param name="UseAscii">1</Param>
</Validator>
</Validators>
```

Limitations

Be aware of the following limitations when using advanced checksum validation:

- No UI Support Creating or editing SITs with advanced checksum logic isn't available via the UX.
- In alphanumeric SITs, letters are skipped if no weights are defined for them.

Date validator

If a date value that's embedded in a regular expression is part of a new pattern you're creating, you can use the *date validator* to test whether that date value meets your criteria. For example, you want to create a SIT for a nine-digit employee identification number. The first six digits are the date of hire in DDMMYY format and the last three are randomly generated numbers. Use these steps to validate that the first six digits are in the correct format:

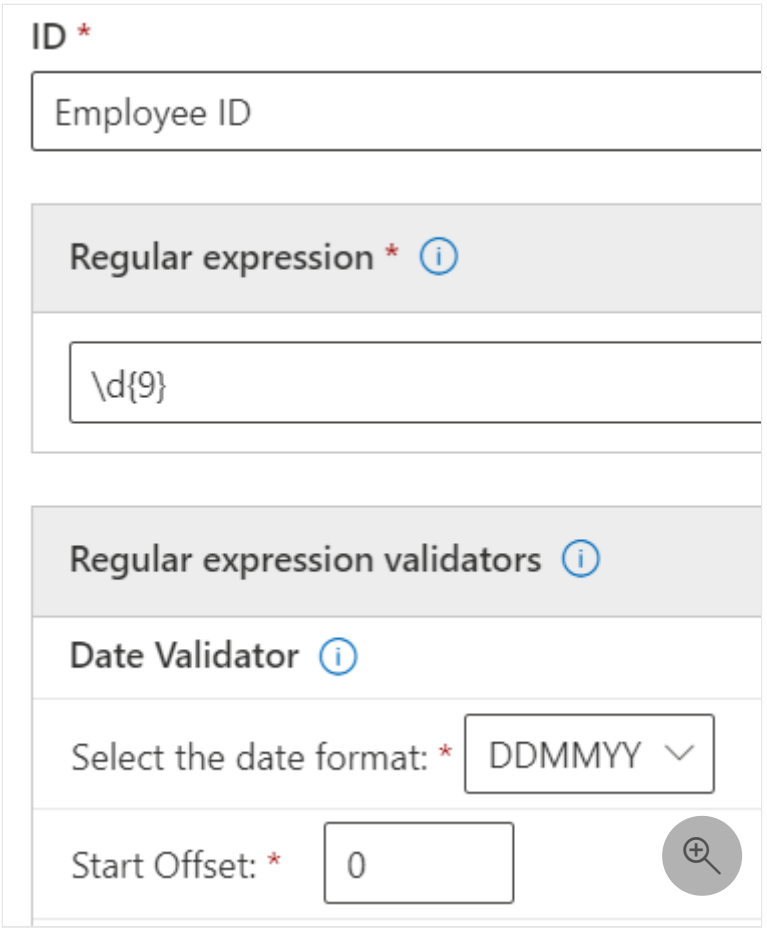
1. Define the primary element with this regular expression:

   Console

   ```
   \d{9}
   ```

2. Add the date validator.

3. Select the date format and the start offset. Since the date string is the first six digits, the offset is `0`.



Functional processors as validators

You can use function processors for some of the most commonly used SITs as validators. Using function processors allows you to define your own regular expressions while ensuring that they pass the additional checks required by the SIT. For example, Func_India_Aadhar ensures that the custom regular expression you defined passes the validation logic required for the Indian Aadhar card. For more information on the DLP functions that you can use as validators, see Sensitive information type functions.

Luhn check validator

You can use the Luhn check validator if you have a custom sensitive information type that includes a regular expression, which should pass the Luhn algorithm.

Sensitive information type additional checks

Here are the definitions and some examples for the available additional checks.

Exclude specific matches: This check lets you define keywords to exclude when detecting matches for the pattern you're editing. For example, you might exclude test credit card numbers like '4111111111111111' so that they're not matched as a valid number.

Starts or doesn't start with characters: This check lets you define the characters that the matched items must or must not start with. For example, if you want the pattern to detect only credit card numbers that start with 41, 42, or 43, select **Starts with** and add 41, 42, and 43 to the list, separated by commas.

Ends or doesn't end with characters: This check lets you define the characters that the matched items must or must not end with. For example, if your Employee ID number can't end with 0 or 1, select **Doesn't end with** and add 0 and 1 to the list, separated by commas.

Exclude duplicate characters: This check lets you ignore matches in which all the digits are the same. For example, if the six digit employee ID number can't have all the digits be the same, you can select **Exclude duplicate characters** to exclude 111111, 222222, 333333, 444444, 555555, 666666, 777777, 888888, 999999, and 000000 from the list of valid matches for the employee ID.

Include or exclude prefixes: This check lets you define the keywords that must or must not be found immediately before the matching entity. Depending on your selection, entities are matched or not matched if they're preceded by the prefixes you include here. For example, if you **Exclude** the prefix `GUID`, any entity that's preceded by **GUID:** won't match.

Include or exclude suffixes: This check lets you define the keywords that must or must not be found immediately after the matching entity. Depending on your selection, entities match or not match if they're followed by the suffixes you include here. For example, if you **Exclude** the suffix `GUID`, any text that's followed by `:GUID` won't match.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Create a custom sensitive information type using PowerShell

This article shows you how to create an XML *rule package* file that defines custom sensitive information types. This article describes a custom sensitive information type that detects employee ID numbers. You can use the sample XML in this article as a starting point for your own XML file.

For more information about sensitive information types, see Learn about sensitive information types.

After you've created a well-formed XML file, you can upload it to Microsoft 365 using PowerShell. Then, you're ready to use your custom sensitive information type in policies. You can test its effectiveness in detecting the sensitive information as you intended.

> **Note**
>
> If you don't need the fine-grained control that PowerShell provides, you can create custom sensitive information types in the Microsoft Purview portal. For more information, see **Create a custom sensitive information type**.

Important disclaimer

Microsoft Support can't help you create content-matching definitions.

For custom content-matching development, testing, and debugging, you need to use your own internal IT resources, or use consulting services, such as Microsoft Consulting Services (MCS). Microsoft Support engineers can provide limited support for this feature, but they can't guarantee that custom content-matching suggestions will fully meet your needs.

MCS can provide regular expressions for testing purposes. They can also provide assistance in troubleshooting an existing RegEx pattern that's not working as expected with a single specific content example.

For more information, see the Potential validation issues to be aware of section.

For more information about the Boost.RegEx (formerly known as RegEx++) engine that's used for processing the text, see Boost.Regex 5.1.3 .

> **Note**
>
> If you use an ampersand character (&) as part of a keyword in your custom sensitive information type, you need to add an additional term with spaces around the character. For example, use `L & P` *not* `L&P`.

Sample XML of a rule package

Here's the sample XML of the rule package that we create in this article. Elements and attributes are explained in the sections on Rule, Entity, and Pattern elements; Match elements; evidence combinations; proximity; confidence levels; LocalizedStrings; and RulePack markup.

XML

```
<?xml version="1.0" encoding="UTF-16"?>
<RulePackage xmlns="http://schemas.microsoft.com/office/2011/mce">
<RulePack id="DAD86A92-AB18-43BB-AB35-96F7C594ADAA">
  <Version build="0" major="1" minor="0" revision="0"/>
  <Publisher id="619DD8C3-7B80-4998-A312-4DF0402BAC04"/>
  <Details defaultLangCode="en-us">
    <LocalizedDetails langcode="en-us">
      <PublisherName>Contoso</PublisherName>
      <Name>Employee ID Custom Rule Pack</Name>
      <Description>
      This rule package contains the custom Employee ID entity.
      </Description>
    </LocalizedDetails>
  </Details>
</RulePack>
<Rules>
<!-- Employee ID -->
  <Entity id="00aa00aa-bb11-cc22-dd33-44ee44ee44ee" patternsProximity="300"
recommendedConfidence="75">
    <Pattern confidenceLevel="65">
      <IdMatch idRef="Regex_employee_id"/>
    </Pattern>
    <Pattern confidenceLevel="75">
      <IdMatch idRef="Regex_employee_id"/>
      <Match idRef="Func_us_date"/>
    </Pattern>
    <Pattern confidenceLevel="85">
      <IdMatch idRef="Regex_employee_id"/>
      <Match idRef="Func_us_date"/>
      <Any minMatches="1">
        <Match idRef="Keyword_badge" minCount="2"/>
```

```
        <Match idRef="Keyword_employee"/>
      </Any>
      <Any minMatches="0" maxMatches="0">
        <Match idRef="Keyword_false_positives_local"/>
        <Match idRef="Keyword_false_positives_intl"/>
      </Any>
    </Pattern>
  </Entity>
  <Regex id="Regex_employee_id">(\s)(\d{9})(\s)</Regex>
  <Keyword id="Keyword_employee">
    <Group matchStyle="word">
      <Term>Identification</Term>
      <Term>Contoso Employee</Term>
    </Group>
  </Keyword>
  <Keyword id="Keyword_badge">
    <Group matchStyle="string">
      <Term>card</Term>
      <Term>badge</Term>
      <Term caseSensitive="true">ID</Term>
    </Group>
  </Keyword>
  <Keyword id="Keyword_false_positives_local">
    <Group matchStyle="word">
      <Term>credit card</Term>
      <Term>national ID</Term>
    </Group>
  </Keyword>
  <Keyword id="Keyword_false_positives_intl">
    <Group matchStyle="word">
      <Term>identity card</Term>
      <Term>national ID</Term>
      <Term>EU debit card</Term>
    </Group>
  </Keyword>
  <LocalizedStrings>
    <Resource idRef="00aa00aa-bb11-cc22-dd33-44ee44ee44ee">
      <Name default="true" langcode="en-us">Employee ID</Name>
      <Description default="true" langcode="en-us">
      A custom classification for detecting Employee IDs.
      </Description>
      <Description default="false" langcode="de-de">
      Description for German locale.
      </Description>
    </Resource>
  </LocalizedStrings>
</Rules>
</RulePackage>
```

What are your key requirements? [Rule, Entity, Pattern elements]

It's important that you understand the basic structure of the XML schema for a rule. Your understanding of the structure helps your custom sensitive information type to identify the right content.

A rule defines one or more entities (also known as sensitive information types). Each entity defines one or more patterns. A pattern is what a policy looks for when it evaluates content (for example, email and documents).

In XML markup, "rules" mean the patterns that define the sensitive information type. Don't associate references to rules in this article with "conditions" or "actions" that are common in other Microsoft features.

Simplest scenario: entity with one pattern

Here's a simple scenario: You want your policy to identify content that contains nine-digit employee IDs that are used in your organization. A pattern refers to the regular expression in the rule that identifies nine-digit numbers. Any content that contains a nine-digit number satisfies the pattern.



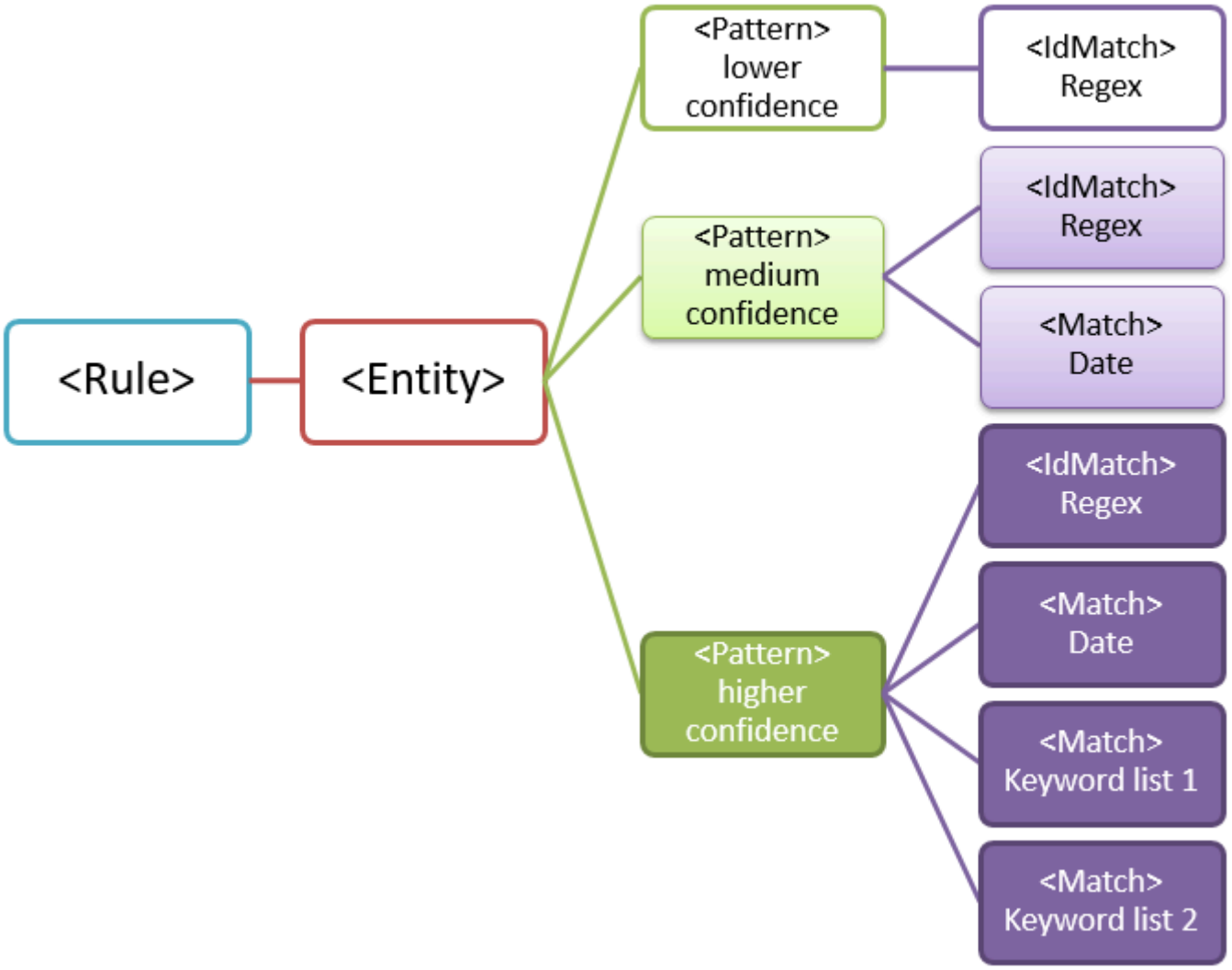
But, this pattern might identify **any** nine-digit number, including longer numbers or other types of nine-digit numbers that aren't employee IDs. This type of unwanted match is known as a *false positive*.

More common scenario: entity with multiple patterns

Because of the potential for false positives, you typically use more than one pattern to define an entity. Multiple patterns provide supporting evidence for the target entity. For example, additional keywords, dates, or other text can help identify the original entity (for example, the nine-digit employee number).

For example, to increase the likelihood of identifying content that contains an employee ID, you can define other patterns to look for:

- A pattern that identifies a hire date.
- A pattern that identifies both a hire date and the "employee ID" keyword.



There are important points to consider for multiple-pattern matches:

- Patterns that require more evidence have a higher confidence level. Based on the confidence level, you can take the following actions:
  - Use more restrictive actions (such as block content) with higher-confidence matches.
  - Use less restrictive actions (such as send notifications) with lower-confidence matches.

- The supporting `IdMatch` and `Match` elements reference RegExes and keywords that are actually children of the `Rule` element, not the `Pattern`. The `Pattern` references supporting elements, but they're included in the `Rule`. This behavior means that a single definition of a supporting element, such as a regular expression or a keyword list, can be referenced by multiple entities and patterns.

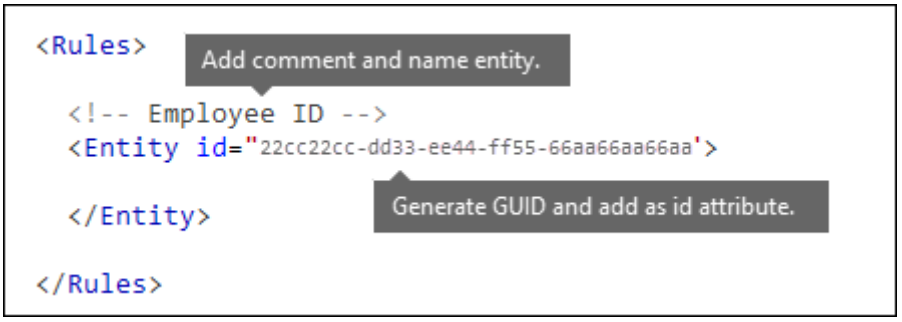
What entity do you need to identify? [Entity element, ID attribute]

An entity is a sensitive information type, such as a credit card number, that has a well-defined pattern. Each entity has a unique GUID as its ID.

Name the entity and generate its GUID

Use the following steps to create the entity markup and generate its unique identifier:

1. In your XML editor of choice, add the `Rules` and `Entity` elements.
2. Add a comment that contains the name of your custom entity, such as Employee ID. Later, you add the entity name to the localized strings section, and that name appears in the admin center when you create a policy.
3. Generate a unique GUID for your entity. For example, in Windows PowerShell, you can run the command `[guid]::NewGuid()`. Later, you'll also add the GUID to the localized strings section of the entity.




```
<Rules>
  <!-- Employee ID -->
  <Entity id="22cc22cc-dd33-ee44-ff55-66aa66aa66aa">
  </Entity>
</Rules>
```

What pattern do you want to match? [Pattern element, IdMatch element, Regex element]

The pattern contains the list of what the sensitive information type is looking for. The pattern can include RegExes, keywords, and built-in functions. Functions do task like running RegExes to find dates or addresses. Sensitive information types can have multiple patterns with unique confidences.

In the following diagram, all of the patterns reference the same regular expression. This RegEx looks for a nine-digit number `(\d{9})` surrounded by white space `(\s) ... (\s)`. The `IdMatch` element references this regular expression, and it's the common requirement for all patterns that look for the Employee ID entity. `IdMatch` is the identifier that the pattern is to trying to match. A `Pattern` element must have exactly one `IdMatch` element.

```
<Rules>

  <!-- Employee ID -->
  <Entity id="                                    ">

    <Pattern>
      <IdMatch idRef="Regex_employee_id" />
    </Pattern>

    <Pattern>
      <IdMatch idRef="Regex_employee_id" />
    </Pattern>

    <Pattern>
      <IdMatch idRef="Regex_employee_id" />
    </Pattern>

  </Entity>

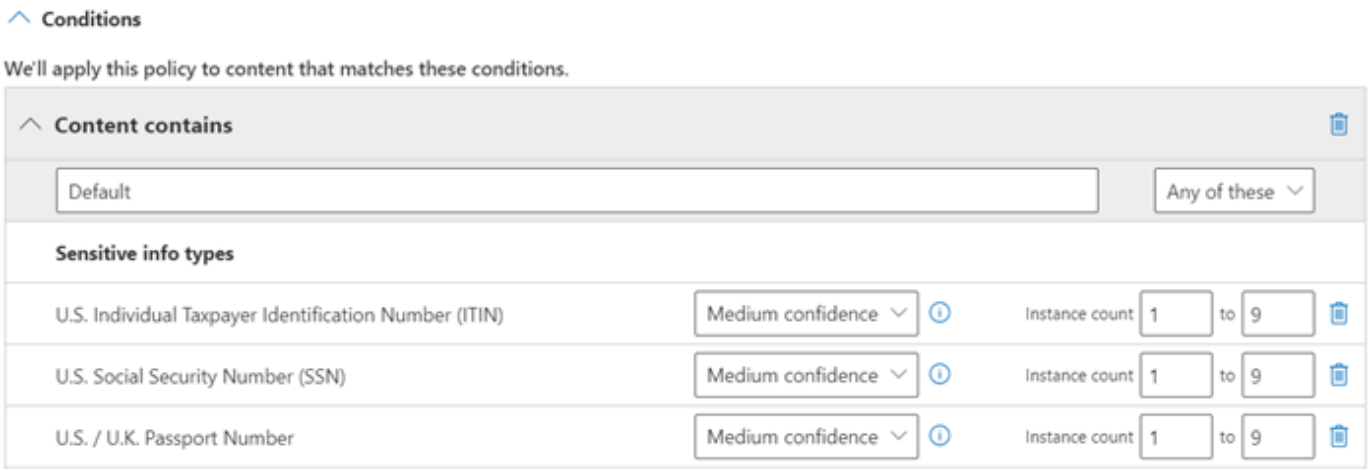
  <Regex id="Regex_employee_id">(\s)(\d{9})(\s)</Regex>

</Rules>
```

Each Pattern has exactly one required IdMatch element, which references the same regex.

Regex defines the identifier the pattern is trying to match.

A satisfied pattern match returns a count and confidence level, which you can use in the conditions in your policy. When you add a condition for detecting a sensitive information type to a policy, you can edit the count and confidence level as shown in the following diagram. Confidence level (also called match accuracy) is explained in the confidenceLevel and recommendedConfidence attributes section.



Regular expressions are powerful, so there are issues that you need to know about. For example, a RegEx that identifies too much content can affect performance. To learn more about regular expression validation and performance issues, see the Potential validation issues to be aware of section.

Do you want to require additional evidence? [Match element, minCount attribute]

In addition to `IdMatch`, a pattern can use the `Match` element to require additional supporting evidence, such as a keyword, RegEx, date, or address.

A `Pattern` might include multiple `Match` elements:

- Directly in the `Pattern` element.
- Combined by using the `Any` element.

`Match` elements are joined with an implicit AND operator. In other words, all `Match` elements must be satisfied for the pattern to be matched.

You can use the `Any` element to introduce AND or OR operators. For details, see the Different combinations of evidence section.

You can use the optional `minCount` attribute to specify how many instances of a match need to be found for each `Match` element. For example, you can specify that a pattern is satisfied only when at least two keywords from a keyword list are found.

```
<Rules>

  <!-- Employee ID -->
  <Entity id="                                    ">

    <Pattern>
      <IdMatch idRef="Regex_employee_id" />
    </Pattern>

    <Pattern>
      <IdMatch idRef="Regex_employee_id" />
    </Pattern>

    <Pattern>
      <IdMatch idRef="Regex_employee_id" />
      <Match idRef="Keyword_employee" />
      <Match idRef="Keyword_badge" minCount="2" />
    </Pattern>

  </Entity>

  <Regex id="Regex_employee_id">(\s)(\d{9})(\s)</Regex>
```

Use minCount to make the Match require more evidence.

Use the Match element to require corroborative evidence.

Keywords [Keyword, Group, and Term elements, matchStyle, and caseSensitive attributes]

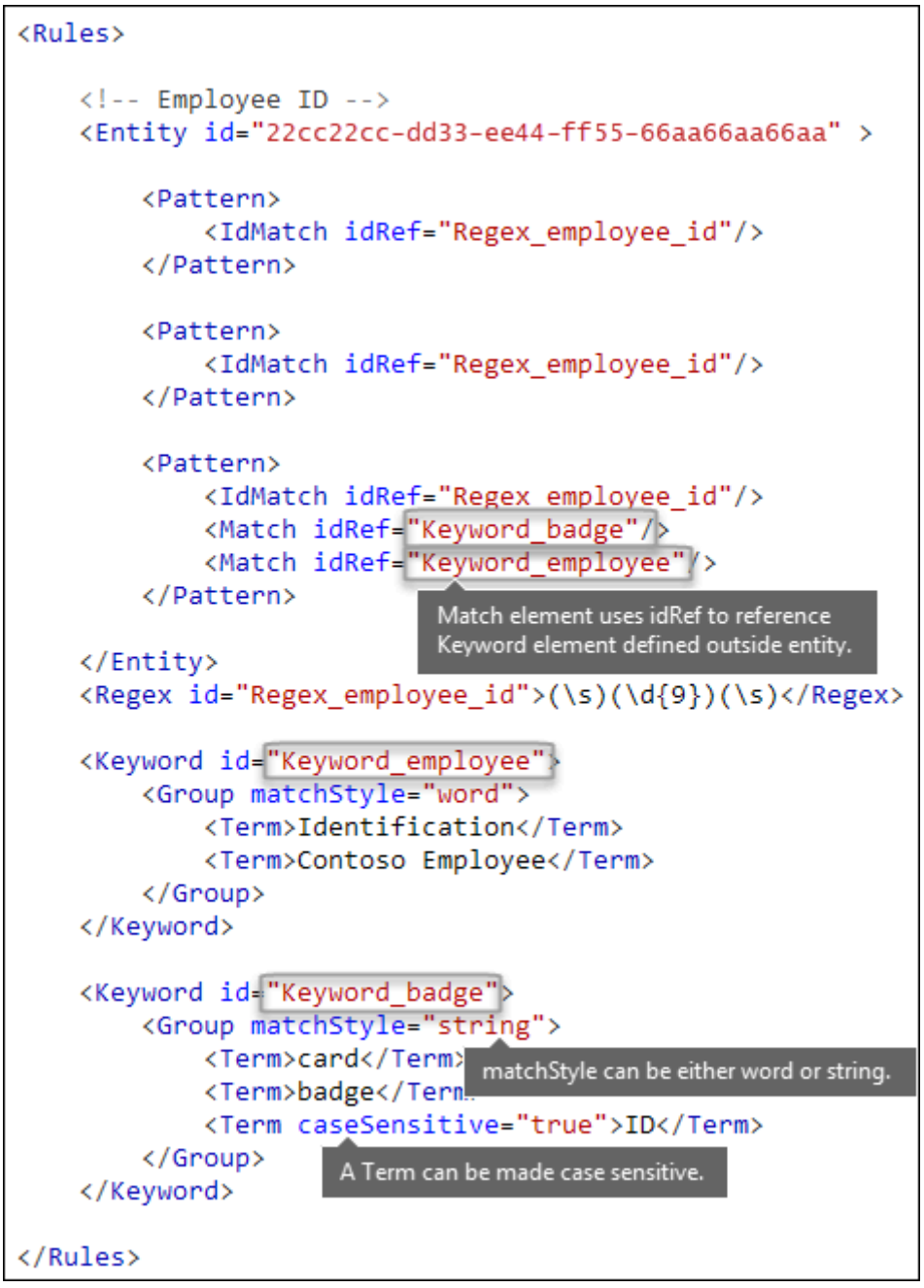
Identifying sensitive information often requires additional keywords as corroborative evidence. For example, in addition to matching a nine-digit number, you can look for words like "card", "badge", or "ID" using the Keyword element. The `Keyword` element has an `ID` attribute that can be referenced by multiple `Match` elements in multiple patterns or entities.

Keywords are included as a list of `Term` elements in a `Group` element. The `Group` element has a `matchStyle` attribute with two possible values:

- **matchStyle="word"**: A word match identifies whole words surrounded by white space or other delimiters. You should always use **word** unless you need to match parts of words or words in Asian languages.

- **matchStyle="string"**: A string match identifies strings no matter what they're surrounded by. For example, "ID" matches both "b**id**" and "**id**ea". Use `string` only when you need to match Asian words or if your keyword might be included in other strings.

Finally, you can use the `caseSensitive` attribute of the `Term` element to specify that the content must match the keyword exactly, including lower-case and upper-case letters.




```
<Rules>

    <!-- Employee ID -->
    <Entity id="22cc22cc-dd33-ee44-ff55-66aa66aa66aa" >

        <Pattern>
            <IdMatch idRef="Regex_employee_id"/>
        </Pattern>

        <Pattern>
            <IdMatch idRef="Regex_employee_id"/>
        </Pattern>

        <Pattern>
            <IdMatch idRef="Regex_employee_id"/>
            <Match idRef="Keyword_badge"/>
            <Match idRef="Keyword_employee"/>
        </Pattern>

    </Entity>
    <Regex id="Regex_employee_id">(\s)(\d{9})(\s)</Regex>

    <Keyword id="Keyword_employee">
        <Group matchStyle="word">
            <Term>Identification</Term>
            <Term>Contoso Employee</Term>
        </Group>
    </Keyword>

    <Keyword id="Keyword_badge">
        <Group matchStyle="string">
            <Term>card</Term>
            <Term>badge</Term>
            <Term caseSensitive="true">ID</Term>
        </Group>
    </Keyword>

</Rules>
```

Regular expressions [Regex element]

In this example, the employee `ID` entity already uses the `IdMatch` element to reference a regular expression for the pattern: a nine-digit number surrounded by whitespace. In addition, a pattern can use a `Match` element to reference an additional `Regex` element to identify corroborative evidence, such as a five-digit or nine-digit number in the format of a US postal code.

Additional patterns such as dates or addresses [built-in functions]

Sensitive information types can also use built-in functions to identify corroborating evidence. For example, a US date, EU date, expiration date, or US address. Microsoft 365 doesn't support uploading your own custom functions. But, when you create a custom sensitive information type, your entity can reference built-in functions.

For example, an employee ID badge has a hire date on it, so this custom entity can use the built-in `Func_us_date` function to identify a date in the format that's commonly used in the US.

For more information, see Sensitive information type functions.



Different combinations of evidence [Any element, minMatches, and maxMatches attributes]

In a `Pattern` element, all `IdMatch` and `Match` elements are joined with an implicit AND operator. In other words, all of the matches must be satisfied before the pattern can be satisfied.

You can create more flexible matching logic by using the `Any` element to group `Match` elements. For example, you can use the `Any` element to match all, none, or an exact subset of its child `Match` elements.

The `Any` element has optional `minMatches` and `maxMatches` attributes that you can use to define how many of the child `Match` elements must be satisfied before the pattern is matched. The

`minMatches` and `maxMatches` attributes define the *number* of `Match` elements, not the number of instances of evidence found for the matches. To define a minimum number of instances for a specific match, such as two keywords from a list, use the `minCount` attribute for a `Match` element, as described in the Match element, minCount attribute section.

Match at least one child Match element

To require only a minimum number of `Match` elements, you can use the `minMatches` attribute. In effect, these `Match` elements are joined with an implicit OR operator. This `Any` element is satisfied if a US-formatted date or a keyword from either list is found.

XML

```
<Any minMatches="1" >
     <Match idRef="Func_us_date" />
     <Match idRef="Keyword_employee" />
     <Match idRef="Keyword_badge" />
</Any>
```

Match an exact subset of any children Match elements

To require an exact number of `Match` elements, set `minMatches` and `maxMatches` to the same value. This `Any` element is satisfied only if exactly one date or keyword is found. If there are any more matches, the pattern isn't matched.

XML

```
<Any minMatches="1" maxMatches="1" >
     <Match idRef="Func_us_date" />
     <Match idRef="Keyword_employee" />
     <Match idRef="Keyword_badge" />
</Any>
```

Match none of children "Match" elements

If you want to require the absence of specific evidence for a pattern to be satisfied, you can set both minMatches and maxMatches to 0. This can be useful if you have a keyword list or other evidence that is likely to indicate a false positive.

For example, the employee ID entity looks for the keyword "card" because it might refer to an "ID card". However, if card appears only in the phrase "credit card", "card" in this content is unlikely

to mean "ID card". So you can add "credit card" as a keyword to a list of terms that you want to exclude from satisfying the pattern.

XML

```
<Any minMatches="0" maxMatches="0" >
    <Match idRef="Keyword_false_positives_local" />
    <Match idRef="Keyword_false_positives_intl" />
</Any>
```

Match several unique terms

If you want to match several unique terms, use the *uniqueResults* parameter, set to *true*, as shown in the following example:

XML

```
<Pattern confidenceLevel="75">
    <IdMatch idRef="Salary_Revision_terms" />
    <Match idRef=" Salary_Revision_ID " minCount="3" uniqueResults="true" />
</Pattern>
```

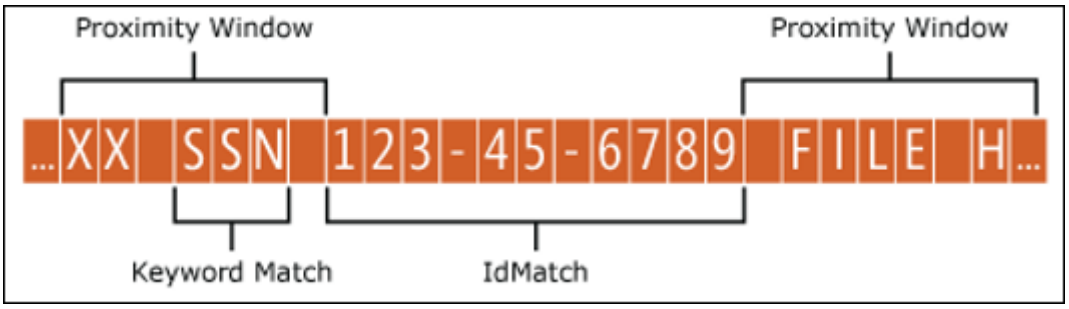
In this example, a pattern is defined for salary revision using at least three unique matches.

How close to the entity must the other evidence be? [patternsProximity attribute]

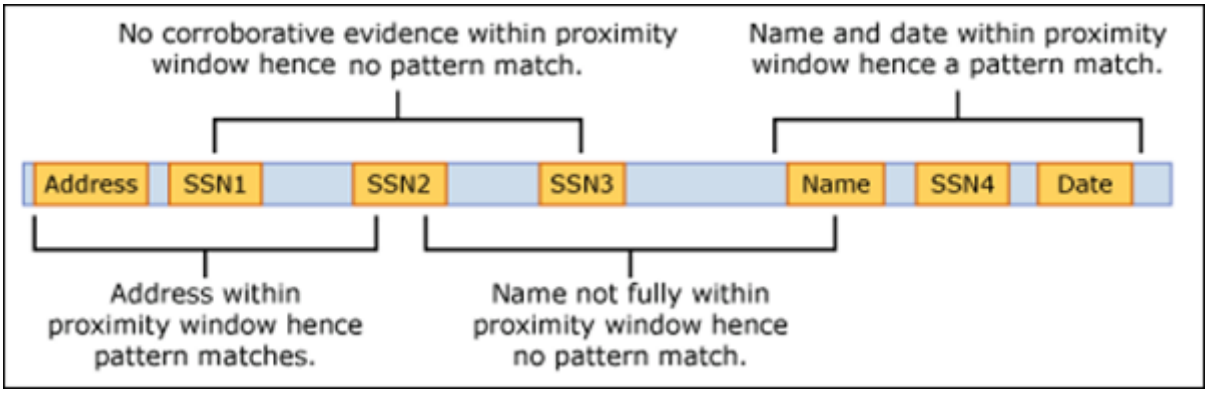
Your sensitive information type is looking for a pattern that represents an employee ID, and as part of that pattern it's also looking for corroborative evidence like a keyword such as "ID". It makes sense that the closer together this evidence is, the more likely the pattern is to be an actual employee ID. You can determine how close other evidence in the pattern must be to the entity by using the required patternsProximity attribute of the Entity element.



For each pattern in the entity, the patternsProximity attribute value defines the distance (in Unicode characters) from the IdMatch location for all other Matches specified for that Pattern. The proximity window is anchored by the IdMatch location, with the window extending to the left and right of the IdMatch.



The example below illustrates how the proximity window affects the pattern matching where IdMatch element for the employee ID custom entity requires at least one corroborating match of keyword or date. Only ID1 matches because for ID2 and ID3, either no or only partial corroborating evidence is found within the proximity window.



For email, the message body and each attachment are treated as separate items. This means that the proximity window doesn't extend beyond the end of each of these items. For each item (attachment or body), both the idMatch and corroborative evidence needs to reside in that item.

What are the right confidence levels for different patterns? [confidenceLevel attribute, recommendedConfidence attribute]

The more evidence that a pattern requires, the more confidence you have that an actual entity (such as employee ID) has been identified when the pattern is matched. For example, you have more confidence in a pattern that requires a nine-digit ID number, hire date, and keyword in close proximity, than you do in a pattern that requires only a nine-digit ID number.

The Pattern element has a required confidenceLevel attribute. You can think of the value of confidenceLevel (a value among 65/75/85 indicating Low/Medium/High confidence levels) as a unique ID for each pattern in an entity. After you upload your custom sensitive information type and then create a policy, you can reference these confidence levels in the conditions of the rules that you create.

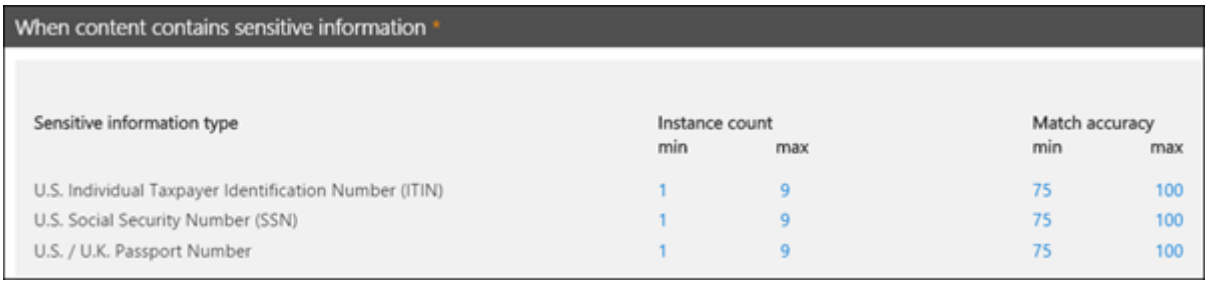
```
<!--  Employee ID  -->
<Entity id="33dd33dd-ee44-ff55-aa66-77bb77bb77bb" patternsProximity="300">
  <Pattern confidenceLevel="65">
    <IdMatch idRef="Regex_employee_id"/>
  </Pattern>
  <Pattern confidenceLevel="75">
    <IdMatch idRef="Regex_employee_id"/>
    <Match idRef="Func_us_date"/>
  </Pattern>
  <Pattern confidenceLevel="85">
    <IdMatch idRef="Regex_employee_id"/>
    <Match idRef="Func_us_date"/>
    <Any minMatches="1">
      <Match idRef="Keyword_badge" minCount="2"/>
      <Match idRef="Keyword_employee"/>
    </Any>
    <Any minMatches="0" maxMatches="0">
      <Match idRef="Keyword_false_positives_local"/>
      <Match idRef="Keyword_false_positives_intl"/>
    </Any>
  </Pattern>
</Entity>
```

The more evidence the Pattern requires, the higher the confidence level.

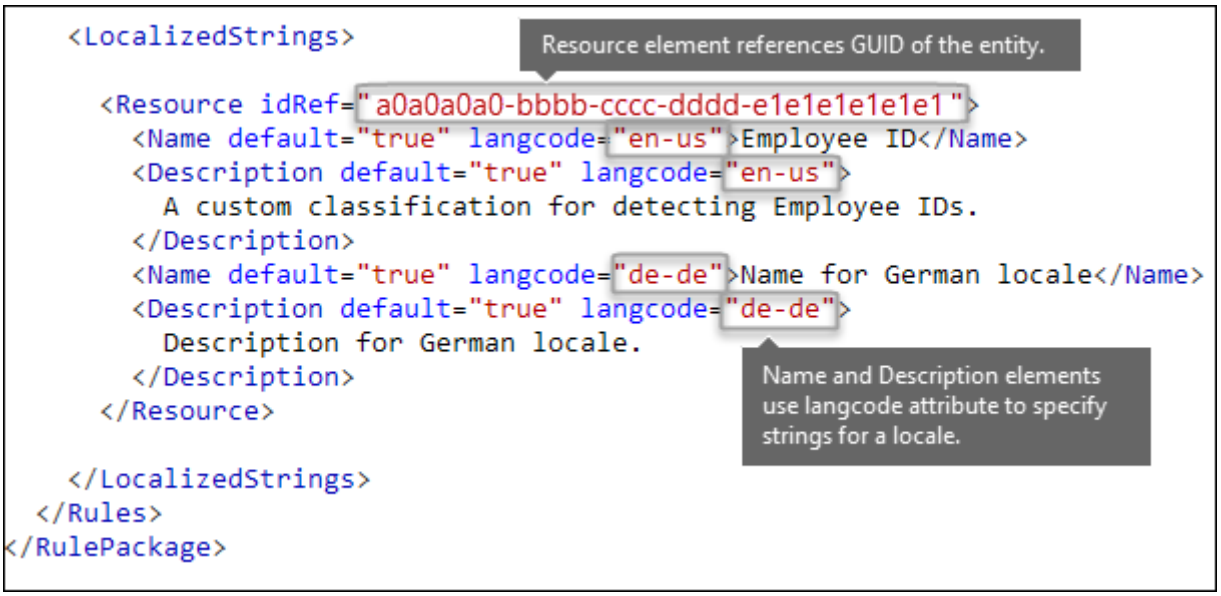
In addition to confidenceLevel for each Pattern, the Entity has a recommendedConfidence attribute. The recommended confidence attribute can be thought of as the default confidence level for the rule. When you create a rule in a policy, if you don't specify a confidence level for the rule to use, that rule matches based on the recommended confidence level for the entity. The recommendedConfidence attribute is mandatory for each Entity ID in the Rule Package, if missing you won't be able to save policies that use the Sensitive Information Type.

Do you want to support other languages in the UI of the Purview portal? [LocalizedStrings element]

If your compliance team uses the Microsoft Purview portal to create policies in different locales and in different languages, you can provide localized versions of the name and description of your custom sensitive information type. When your compliance team uses Microsoft 365 in a language that you support, they see the localized name in the UI.



The Rules element must contain a LocalizedStrings element, which contains a Resource element that references the GUID of your custom entity. In turn, each Resource element contains one or more Name and Description elements that each use the `langcode` attribute to provide a localized string for a specific language.




```
    <LocalizedStrings>
      <Resource idRef="a0a0a0a0-bbbb-cccc-dddd-e1e1e1e1e1e1">
        <Name default="true" langcode="en-us">Employee ID</Name>
        <Description default="true" langcode="en-us">
          A custom classification for detecting Employee IDs.
        </Description>
        <Name default="true" langcode="de-de">Name for German locale</Name>
        <Description default="true" langcode="de-de">
          Description for German locale.
        </Description>
      </Resource>
    </LocalizedStrings>
  </Rules>
</RulePackage>
```

You use localized strings only for how your custom sensitive information type appears in the UI of the Compliance center. You can't use localized strings to provide different localized versions of a keyword list or regular expression.

Other rule package markup [RulePack GUID]

Finally, the beginning of each RulePackage contains some general information that you need to fill in. You can use the following markup as a template and replace the ". . ." placeholders with your own info.

Most importantly, you need to generate a GUID for the RulePack. Earlier, you generated a GUID for the entity; this is a second GUID for the RulePack. There are several ways to generate GUIDs, but you can do it easily in PowerShell by typing [guid]::NewGuid().

The Version element is also important. When you upload your rule package for the first time, Microsoft 365 notes the version number. Later, if you update the rule package and upload a new

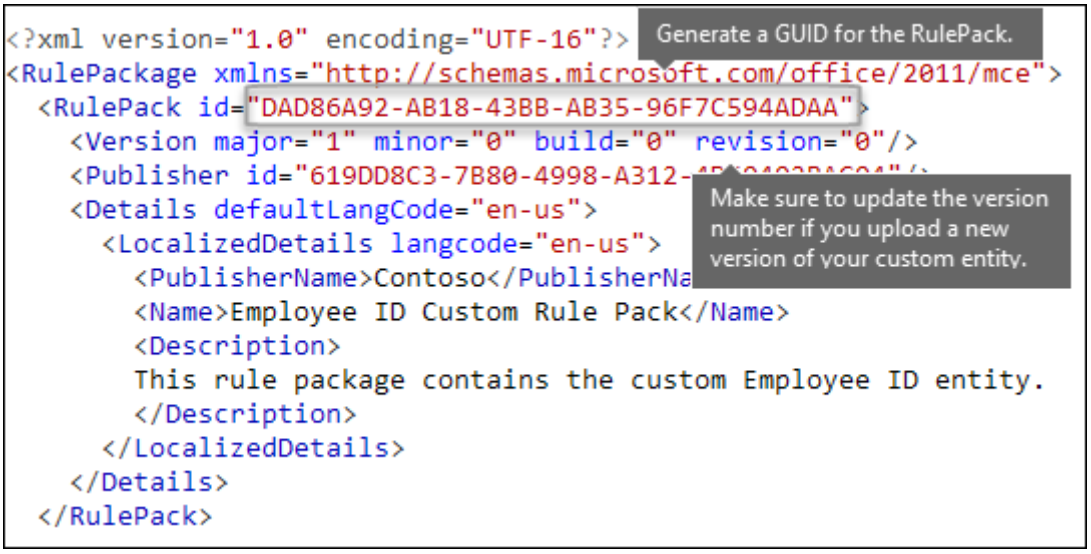
version, make sure to update the version number or Microsoft 365 won't deploy the rule package.

XML

```
<?xml version="1.0" encoding="utf-16"?>
<RulePackage xmlns="http://schemas.microsoft.com/office/2011/mce">
  <RulePack id=". . .">
    <Version major="1" minor="0" build="0" revision="0" />
    <Publisher id=". . ." />
    <Details defaultLangCode=". . .">
      <LocalizedDetails langcode=" . . . ">
         <PublisherName>. . .</PublisherName>
         <Name>. . .</Name>
         <Description>. . .</Description>
      </LocalizedDetails>
    </Details>
  </RulePack>

 <Rules>
  . . .
 </Rules>
</RulePackage>
```

When complete, your RulePack element should look like this.



Validators

Microsoft 365 exposes function processors for commonly used SITs as validators. Here's a list of them.

List of currently available validators

The following validator functions are currently available for use in custom sensitive information type definitions:

- `Func_credit_card`
- `Func_ssn`
- `Func_unformatted_ssn`
- `Func_randomized_formatted_ssn`
- `Func_randomized_unformatted_ssn`
- `Func_aba_routing`
- `Func_south_africa_identification_number`
- `Func_brazil_cpf`
- `Func_iban`
- `Func_brazil_cnpj`
- `Func_swedish_national_identifier`
- `Func_india_aadhaar`
- `Func_uk_nhs_number`
- `Func_Turkish_National_Id`
- `Func_australian_tax_file_number`
- `Func_usa_uk_passport`
- `Func_canadian_sin`
- `Func_formatted_itin`
- `Func_unformatted_itin`
- `Func_dea_number_v2`
- `Func_dea_number`
- `Func_japanese_my_number_personal`
- `Func_japanese_my_number_corporate`

This gives you the ability to define your own RegEx and validate them. To use validators, define your own RegEx and use the `Validator` property to add the function processor of your choice. Once defined, you can use this RegEx in an SIT.

In the example below, a regular expression - Regex_credit_card_AdditionalDelimiters is defined for Credit card, which is then validated using the checksum function for credit card by using Func_credit_card as a validator.

XML

```
<Regex id="Regex_credit_card_AdditionalDelimiters" validators="Func_credit_card">
(?:^|[\s,;\:\(\)\[\]"'])([0-9]{4}[ -_][0-9]{4}[ -_][0-9]{4}[ -_][0-9]{4})(?:$|[\s,;\:\
(\)\[\]"'])</Regex>
```

```
<Entity id="675634eb7-edc8-4019-85dd-5a5c1f2bb085" patternsProximity="300"
recommendedConfidence="85">
<Pattern confidenceLevel="85">
<IdMatch idRef="Regex_credit_card_AdditionalDelimiters" />
<Any minMatches="1">
<Match idRef="Keyword_cc_verification" />
<Match idRef="Keyword_cc_name" />
<Match idRef="Func_expiration_date" />
</Any>
</Pattern>
</Entity>
```

Microsoft 365 provides two generic validators

Checksum validator

In this example, a checksum validator for employee ID is defined to validate the RegEx for EmployeeID.

XML

```
<Validators id="EmployeeIDChecksumValidator">
<Validator type="Checksum">
<Param name="Weights">2, 2, 2, 2, 2, 1</Param>
<Param name="Mod">28</Param>
<Param name="CheckDigit">2</Param> <!-- Check 2nd digit -->
<Param name="AllowAlphabets">1</Param> <!– 0 if no Alphabets -->
</Validator>
</Validators>
<Regex id="Regex_EmployeeID" validators="ChecksumValidator">(\d{5}[A-Z])</Regex>
<Entity id="675634eb7-edc8-4019-85dd-5a5c1f2bb085" patternsProximity="300"
recommendedConfidence="85">
<Pattern confidenceLevel="85">
<IdMatch idRef="Regex_EmployeeID"/>
</Pattern>
</Entity>
```

Date Validator

In this example, a date validator is defined for a RegEx part of which is date.

XML

```
<Validators id="date_validator_1"> <Validator type="DateSimple"> <Param
name="Pattern">DDMMYYYY</Param> <!–supported patterns DDMMYYYY, MMDDYYYY, YYYYDDMM,
YYYYMMDD, DDMMYYYY, DDMMYY, MMDDYY, YYDDMM, YYMMDD --> </Validator> </Validators>
<Regex id="date_regex_1" validators="date_validator_1">\d{8}</Regex>
```

Changes for Exchange Online

Previously, you might have used Exchange Online PowerShell to import your custom sensitive information types for DLP. Now your custom sensitive information types can be used in both the Exchange admin center and the Microsoft Purview portal. As part of this improvement, you should use Security & Compliance PowerShell to import your custom sensitive information types—you can't import them from Exchange Online PowerShell anymore. Your custom sensitive information types will continue to work just like before; however, it may take up to one hour for changes made to custom sensitive information types in the Compliance center to appear in the Exchange admin center.

Note that in the Compliance center, you use the **New-DlpSensitiveInformationTypeRulePackage** cmdlet to upload a rule package. (Previously, in the Exchange admin center, you used the **ClassificationRuleCollection**` cmdlet.)

Upload your rule package

To upload your rule package, do the following steps:

1. Save it as an .xml file with Unicode encoding.
2. Connect to Security & Compliance PowerShell
3. Use the following syntax:

   PowerShell

   ```
   New-DlpSensitiveInformationTypeRulePackage -FileData
   ([System.IO.File]::ReadAllBytes('PathToUnicodeXMLFile'))
   ```

   This example uploads the Unicode XML file named MyNewRulePack.xml from C:\My Documents.

   PowerShell

   ```
   New-DlpSensitiveInformationTypeRulePackage -FileData
   ([System.IO.File]::ReadAllBytes('C:\My Documents\MyNewRulePack.xml'))
   ```

   For detailed syntax and parameter information, see New-DlpSensitiveInformationTypeRulePackage.

> **Note**
>
> The maximum number of rule packages supported is 10, but each package can contain the definition of multiple sensitive information types.

4. To verify that you've successfully created a new sensitive information type, do any of the following steps:

   - Run the Get-DlpSensitiveInformationTypeRulePackage cmdlet to verify the new rule package is listed:

     ```
     PowerShell
     Get-DlpSensitiveInformationTypeRulePackage
     ```

   - Run the Get-DlpSensitiveInformationType cmdlet to verify the sensitive information type is listed:

     ```
     PowerShell
     Get-DlpSensitiveInformationType
     ```

     For custom sensitive information types, the Publisher property value will be something other than Microsoft Corporation.

   - Replace <Name> with the Name value of the sensitive information type (example: Employee ID) and run the Get-DlpSensitiveInformationType cmdlet:

     ```
     PowerShell
     Get-DlpSensitiveInformationType -Identity "<Name>"
     ```

Potential validation issues to be aware of

When you upload your rule package XML file, the system validates the XML and checks for known bad patterns and obvious performance issues. Here are some known issues that the validation checks for a regular expression:

- Lookbehind assertions in the regular expression should be of fixed length only. Variable length assertions results in errors.

For example, `"(?<=^|\s|_)"` won't pass validation. The first pattern (`^`) is zero length, while the next two patterns (`\s` and `_`) have a length of one. An alternate way to write this regular expression is `"(?:^|(?<=\s|_))"`.

- Can't begin or end with alternator `|`, which matches everything because it's considered an empty match.

  For example, `|a` or `b|` won't pass validation.

- Can't begin or end with a `.{0,m}` pattern, which has no functional purpose and only impairs performance.

  For example, `.{0,50}ASDF` or `ASDF.{0,50}` won't pass validation.

- Can't have `.{0,m}` or `.{1,m}` in groups, and can't have `.\*` or `.+` in groups.

  For example, `(.{0,50000})` won't pass validation.

- Can't have any character with `{0,m}` or `{1,m}` repeaters in groups.

  For example, `(a\*)` won't pass validation.

- Can't begin or end with `.{1,m}`; instead, use `.`.

  For example, `.{1,m}asdf` won't pass validation. Instead, use `.asdf`.

- Can't have an unbounded repeater (such as `*` or `+`) on a group.

  For example, `(xx)\*` and `(xx)+` won't pass validation.

- The regex pattern must contain exactly one primary capturing group (the only capturing group). Use noncapturing groups for any prefix/suffix (for example: `(?:prefix)(primary capturing group)(?:suffix)`). Consolidate match variants inside the single capturing group using `|`; patterns like `(group1)|(group2)` are rejected.

  For example, `(?:[\s,;])([A-Z]{2}[0-9]{3})(?:[\s,;])|(?:[\s,;])([A-Z]{2}[A-Z]{4}[0-9])(?:[\s,;])` won't pass validation because it contains multiple capturing groups. Instead, use `(?:[\s,;])([A-Z]{2}[0-9]{3}|[A-Z]{2}[A-Z]{4}[0-9])(?:[\s,;])`.

- Keywords have a maximum of 50 characters in Length. If you have a keyword within a Group exceeding this, a suggested solution is to create the Group of terms as a Keyword Dictionary and reference the GUID of the Keyword Dictionary within the XML structure as part of the Entity for Match or idMatch in the file.

- Each Custom Sensitive Information Type can have a maximum of 2,048 keywords total.
- The maximum size of Keyword Dictionaries in a single tenant is 480 KB compressed to comply with AD Schema limits. Reference the same dictionary as many times as necessary when creating custom sensitive information types. Start with creating custom keyword lists in the sensitive information type and use keyword dictionaries if you have more than 2048 keywords in a keyword list or a keyword is larger than 50 characters in length.
- A maximum of 50 keyword dictionary based sensitive information types are allowed in a tenant.
- Ensure each Entity element contains a recommendedConfidence attribute.
- When using the PowerShell Cmdlet, there's a maximum return size of the Deserialized Data of approximately 1 megabyte. This affects the size of your rule pack XML file. Keep the uploaded file limited to a 770 kilobyte maximum as a suggested limit for consistent results without error when processing.
- The XML structure doesn't require formatting characters such as spaces, tabs, or carriage return/linefeed entries. Take note of this when optimizing for space on uploads. Tools such as Microsoft Visual Code provide join line features to compact the XML file.

If a custom sensitive information type contains an issue that may affect performance, it won't be uploaded and you may see one of these error messages:

- `Generic quantifiers which match more content than expected (e.g., '+', '*')`
- `Lookaround assertions`
- `Complex grouping in conjunction with general quantifiers`

Recrawl your content to identify the sensitive information

Microsoft 365 uses the search crawler to identify and classify sensitive information in site content. Content in SharePoint Online and OneDrive for Business sites is recrawled automatically whenever it's updated. But to identify your new custom type of sensitive information in all existing content, that content must be recrawled.

To re-evaluate content across multiple locations at once, you can create a scan using On-demand classification. Learn more about On-demand classification\

Reference: Rule package XML schema definition

You can copy this markup, save it as an XSD file, and use it to validate your rule package XML file.

XML

```xml
<?xml version="1.0" encoding="utf-8"?>
<xs:schema xmlns:mce="http://schemas.microsoft.com/office/2011/mce"
           targetNamespace="http://schemas.microsoft.com/office/2011/mce"
           xmlns:xs="https://www.w3.org/2001/XMLSchema"
           elementFormDefault="qualified"
           attributeFormDefault="unqualified"
           id="RulePackageSchema">
  <!-- Use include if this schema has the same target namespace as the schema being
referenced, otherwise use import -->
  <xs:element name="RulePackage" type="mce:RulePackageType"/>
  <xs:simpleType name="LangType">
    <xs:union memberTypes="xs:language">
      <xs:simpleType>
        <xs:restriction base="xs:string">
          <xs:enumeration value=""/>
        </xs:restriction>
      </xs:simpleType>
    </xs:union>
  </xs:simpleType>
  <xs:simpleType name="GuidType" final="#all">
    <xs:restriction base="xs:token">
      <xs:pattern value="[0-9a-fA-F]{8}\-([0-9a-fA-F]{4}\-){3}[0-9a-fA-F]{12}"/>
    </xs:restriction>
  </xs:simpleType>
  <xs:complexType name="RulePackageType">
    <xs:sequence>
      <xs:element name="RulePack" type="mce:RulePackType"/>
      <xs:element name="Rules" type="mce:RulesType">
        <xs:key name="UniqueRuleId">
          <xs:selector
xpath="mce:Entity|mce:Affinity|mce:Version/mce:Entity|mce:Version/mce:Affinity"/>
          <xs:field xpath="@id"/>
        </xs:key>
        <xs:key name="UniqueProcessorId">
          <xs:selector xpath="mce:Regex|mce:Keyword|mce:Fingerprint"></xs:selector>
          <xs:field xpath="@id"/>
        </xs:key>
        <xs:key name="UniqueResourceIdRef">
          <xs:selector xpath="mce:LocalizedStrings/mce:Resource"/>
          <xs:field xpath="@idRef"/>
        </xs:key>
        <xs:keyref name="ReferencedRuleMustExist" refer="mce:UniqueRuleId">
          <xs:selector xpath="mce:LocalizedStrings/mce:Resource"/>
          <xs:field xpath="@idRef"/>
        </xs:keyref>
        <xs:keyref name="RuleMustHaveResource" refer="mce:UniqueResourceIdRef">
          <xs:selector
```

```
xpath="mce:Entity|mce:Affinity|mce:Version/mce:Entity|mce:Version/mce:Affinity"/>
          <xs:field xpath="@id"/>
        </xs:keyref>
      </xs:element>
    </xs:sequence>
  </xs:complexType>
  <xs:complexType name="RulePackType">
    <xs:sequence>
      <xs:element name="Version" type="mce:VersionType"/>
      <xs:element name="Publisher" type="mce:PublisherType"/>
      <xs:element name="Details" type="mce:DetailsType">
        <xs:key name="UniqueLangCodeInLocalizedDetails">
          <xs:selector xpath="mce:LocalizedDetails"/>
          <xs:field xpath="@langcode"/>
        </xs:key>
        <xs:keyref name="DefaultLangCodeMustExist"
refer="mce:UniqueLangCodeInLocalizedDetails">
          <xs:selector xpath="."/>
          <xs:field xpath="@defaultLangCode"/>
        </xs:keyref>
      </xs:element>
      <xs:element name="Encryption" type="mce:EncryptionType" minOccurs="0"
maxOccurs="1"/>
    </xs:sequence>
    <xs:attribute name="id" type="mce:GuidType" use="required"/>
  </xs:complexType>
  <xs:complexType name="VersionType">
    <xs:attribute name="major" type="xs:unsignedShort" use="required"/>
    <xs:attribute name="minor" type="xs:unsignedShort" use="required"/>
    <xs:attribute name="build" type="xs:unsignedShort" use="required"/>
    <xs:attribute name="revision" type="xs:unsignedShort" use="required"/>
  </xs:complexType>
  <xs:complexType name="PublisherType">
    <xs:attribute name="id" type="mce:GuidType" use="required"/>
  </xs:complexType>
  <xs:complexType name="LocalizedDetailsType">
    <xs:sequence>
      <xs:element name="PublisherName" type="mce:NameType"/>
      <xs:element name="Name" type="mce:RulePackNameType"/>
      <xs:element name="Description" type="mce:OptionalNameType"/>
    </xs:sequence>
    <xs:attribute name="langcode" type="mce:LangType" use="required"/>
  </xs:complexType>
  <xs:complexType name="DetailsType">
    <xs:sequence>
      <xs:element name="LocalizedDetails" type="mce:LocalizedDetailsType"
maxOccurs="unbounded"/>
    </xs:sequence>
    <xs:attribute name="defaultLangCode" type="mce:LangType" use="required"/>
  </xs:complexType>
  <xs:complexType name="EncryptionType">
    <xs:sequence>
      <xs:element name="Key" type="xs:normalizedString"/>
      <xs:element name="IV" type="xs:normalizedString"/>
    </xs:sequence>
```

```
  </xs:complexType>
  <xs:simpleType name="RulePackNameType">
    <xs:restriction base="xs:token">
      <xs:minLength value="1"/>
      <xs:maxLength value="64"/>
    </xs:restriction>
  </xs:simpleType>
  <xs:simpleType name="NameType">
    <xs:restriction base="xs:normalizedString">
      <xs:minLength value="1"/>
      <xs:maxLength value="256"/>
    </xs:restriction>
  </xs:simpleType>
  <xs:simpleType name="OptionalNameType">
    <xs:restriction base="xs:normalizedString">
      <xs:minLength value="0"/>
      <xs:maxLength value="256"/>
    </xs:restriction>
  </xs:simpleType>
  <xs:simpleType name="RestrictedTermType">
    <xs:restriction base="xs:string">
      <xs:minLength value="1"/>
      <xs:maxLength value="100"/>
    </xs:restriction>
  </xs:simpleType>
  <xs:complexType name="RulesType">
    <xs:sequence>
      <xs:choice maxOccurs="unbounded">
        <xs:element name="Entity" type="mce:EntityType"/>
        <xs:element name="Affinity" type="mce:AffinityType"/>
        <xs:element name="Version" type="mce:VersionedRuleType"/>
      </xs:choice>
      <xs:choice minOccurs="0" maxOccurs="unbounded">
        <xs:element name="Regex" type="mce:RegexType"/>
        <xs:element name="Keyword" type="mce:KeywordType"/>
        <xs:element name="Fingerprint" type="mce:FingerprintType"/>
        <xs:element name="ExtendedKeyword" type="mce:ExtendedKeywordType"/>
      </xs:choice>
      <xs:element name="LocalizedStrings" type="mce:LocalizedStringsType"/>
    </xs:sequence>
  </xs:complexType>
  <xs:complexType name="EntityType">
    <xs:sequence>
      <xs:element name="Pattern" type="mce:PatternType" maxOccurs="unbounded"/>
      <xs:element name="Version" type="mce:VersionedPatternType" minOccurs="0"
maxOccurs="unbounded" />
    </xs:sequence>
    <xs:attribute name="id" type="mce:GuidType" use="required"/>
    <xs:attribute name="patternsProximity" type="mce:ProximityType" use="required"/>
    <xs:attribute name="recommendedConfidence" type="mce:ProbabilityType"/>
    <xs:attribute name="workload" type="mce:WorkloadType"/>
  </xs:complexType>
  <xs:complexType name="PatternType">
    <xs:sequence>
      <xs:element name="IdMatch" type="mce:IdMatchType"/>
```

```
      <xs:choice minOccurs="0" maxOccurs="unbounded">
        <xs:element name="Match" type="mce:MatchType"/>
        <xs:element name="Any" type="mce:AnyType"/>
      </xs:choice>
    </xs:sequence>
    <xs:attribute name="confidenceLevel" type="mce:ProbabilityType" use="required"/>
  </xs:complexType>
  <xs:complexType name="AffinityType">
    <xs:sequence>
      <xs:element name="Evidence" type="mce:EvidenceType" maxOccurs="unbounded"/>
      <xs:element name="Version" type="mce:VersionedEvidenceType" minOccurs="0"
maxOccurs="unbounded" />
    </xs:sequence>
    <xs:attribute name="id" type="mce:GuidType" use="required"/>
    <xs:attribute name="evidencesProximity" type="mce:ProximityType" use="required"/>
    <xs:attribute name="thresholdConfidenceLevel" type="mce:ProbabilityType"
use="required"/>
    <xs:attribute name="workload" type="mce:WorkloadType"/>
  </xs:complexType>
  <xs:complexType name="EvidenceType">
    <xs:sequence>
      <xs:choice maxOccurs="unbounded">
        <xs:element name="Match" type="mce:MatchType"/>
        <xs:element name="Any" type="mce:AnyType"/>
      </xs:choice>
    </xs:sequence>
    <xs:attribute name="confidenceLevel" type="mce:ProbabilityType" use="required"/>
  </xs:complexType>
  <xs:complexType name="IdMatchType">
    <xs:attribute name="idRef" type="xs:string" use="required"/>
  </xs:complexType>
  <xs:complexType name="MatchType">
    <xs:attribute name="idRef" type="xs:string" use="required"/>
    <xs:attribute name="minCount" type="xs:positiveInteger" use="optional"/>
    <xs:attribute name="uniqueResults" type="xs:boolean" use="optional"/>
  </xs:complexType>
  <xs:complexType name="AnyType">
    <xs:sequence>
      <xs:choice maxOccurs="unbounded">
        <xs:element name="Match" type="mce:MatchType"/>
        <xs:element name="Any" type="mce:AnyType"/>
      </xs:choice>
    </xs:sequence>
    <xs:attribute name="minMatches" type="xs:nonNegativeInteger" default="1"/>
    <xs:attribute name="maxMatches" type="xs:nonNegativeInteger" use="optional"/>
  </xs:complexType>
  <xs:simpleType name="ProximityType">
    <xs:union>
      <xs:simpleType>
        <xs:restriction base='xs:string'>
          <xs:enumeration value="unlimited"/>
        </xs:restriction>
      </xs:simpleType>
      <xs:simpleType>
        <xs:restriction base="xs:positiveInteger">
```

```
          <xs:minInclusive value="1"/>
        </xs:restriction>
      </xs:simpleType>
    </xs:union>
  </xs:simpleType>
  <xs:simpleType name="ProbabilityType">
    <xs:restriction base="xs:integer">
      <xs:minInclusive value="1"/>
      <xs:maxInclusive value="100"/>
    </xs:restriction>
  </xs:simpleType>
  <xs:simpleType name="WorkloadType">
    <xs:restriction base="xs:string">
      <xs:enumeration value="Exchange"/>
      <xs:enumeration value="Outlook"/>
    </xs:restriction>
  </xs:simpleType>
  <xs:simpleType name="EngineVersionType">
    <xs:restriction base="xs:token">
      <xs:pattern value="^\d{2}\.01?\.\d{3,4}\.\d{1,3}$"/>
    </xs:restriction>
  </xs:simpleType>
  <xs:complexType name="VersionedRuleType">
    <xs:choice maxOccurs="unbounded">
      <xs:element name="Entity" type="mce:EntityType"/>
      <xs:element name="Affinity" type="mce:AffinityType"/>
    </xs:choice>
    <xs:attribute name="minEngineVersion" type="mce:EngineVersionType" use="required"
/>
  </xs:complexType>
  <xs:complexType name="VersionedPatternType">
    <xs:sequence>
      <xs:element name="Pattern" type="mce:PatternType" maxOccurs="unbounded"/>
    </xs:sequence>
    <xs:attribute name="minEngineVersion" type="mce:EngineVersionType" use="required"
/>
  </xs:complexType>
  <xs:complexType name="VersionedEvidenceType">
    <xs:sequence>
      <xs:element name="Evidence" type="mce:EvidenceType" maxOccurs="unbounded"/>
    </xs:sequence>
    <xs:attribute name="minEngineVersion" type="mce:EngineVersionType" use="required"
/>
  </xs:complexType>
  <xs:simpleType name="FingerprintValueType">
    <xs:restriction base="xs:string">
      <xs:minLength value="2732"/>
      <xs:maxLength value="2732"/>
    </xs:restriction>
  </xs:simpleType>
  <xs:complexType name="FingerprintType">
    <xs:simpleContent>
      <xs:extension base="mce:FingerprintValueType">
        <xs:attribute name="id" type="xs:token" use="required"/>
        <xs:attribute name="threshold" type="mce:ProbabilityType" use="required"/>
```

```
      <xs:attribute name="shingleCount" type="xs:positiveInteger" use="required"/>
      <xs:attribute name="description" type="xs:string" use="optional"/>
    </xs:extension>
  </xs:simpleContent>
</xs:complexType>
<xs:complexType name="RegexType">
  <xs:simpleContent>
    <xs:extension base="xs:string">
      <xs:attribute name="id" type="xs:token" use="required"/>
    </xs:extension>
  </xs:simpleContent>
</xs:complexType>
<xs:complexType name="KeywordType">
  <xs:sequence>
    <xs:element name="Group" type="mce:GroupType" maxOccurs="unbounded"/>
  </xs:sequence>
  <xs:attribute name="id" type="xs:token" use="required"/>
</xs:complexType>
<xs:complexType name="GroupType">
  <xs:sequence>
    <xs:choice>
      <xs:element name="Term" type="mce:TermType" maxOccurs="unbounded"/>
    </xs:choice>
  </xs:sequence>
  <xs:attribute name="matchStyle" default="word">
    <xs:simpleType>
      <xs:restriction base="xs:NMTOKEN">
        <xs:enumeration value="word"/>
        <xs:enumeration value="string"/>
      </xs:restriction>
    </xs:simpleType>
  </xs:attribute>
</xs:complexType>
<xs:complexType name="TermType">
  <xs:simpleContent>
    <xs:extension base="mce:RestrictedTermType">
      <xs:attribute name="caseSensitive" type="xs:boolean" default="false"/>
    </xs:extension>
  </xs:simpleContent>
</xs:complexType>
<xs:complexType name="ExtendedKeywordType">
  <xs:simpleContent>
    <xs:extension base="xs:string">
      <xs:attribute name="id" type="xs:token" use="required"/>
    </xs:extension>
  </xs:simpleContent>
</xs:complexType>
<xs:complexType name="LocalizedStringsType">
  <xs:sequence>
    <xs:element name="Resource" type="mce:ResourceType" maxOccurs="unbounded">
    <xs:key name="UniqueLangCodeUsedInNamePerResource">
      <xs:selector xpath="mce:Name"/>
      <xs:field xpath="@langcode"/>
    </xs:key>
    <xs:key name="UniqueLangCodeUsedInDescriptionPerResource">
```

```
        <xs:selector xpath="mce:Description"/>
        <xs:field xpath="@langcode"/>
      </xs:key>
    </xs:element>
    </xs:sequence>
  </xs:complexType>
  <xs:complexType name="ResourceType">
    <xs:sequence>
      <xs:element name="Name" type="mce:ResourceNameType" maxOccurs="unbounded"/>
      <xs:element name="Description" type="mce:DescriptionType" minOccurs="0"
maxOccurs="unbounded"/>
    </xs:sequence>
    <xs:attribute name="idRef" type="mce:GuidType" use="required"/>
  </xs:complexType>
  <xs:complexType name="ResourceNameType">
    <xs:simpleContent>
      <xs:extension base="xs:string">
        <xs:attribute name="default" type="xs:boolean" default="false"/>
        <xs:attribute name="langcode" type="mce:LangType" use="required"/>
      </xs:extension>
    </xs:simpleContent>
  </xs:complexType>
  <xs:complexType name="DescriptionType">
    <xs:simpleContent>
      <xs:extension base="xs:string">
        <xs:attribute name="default" type="xs:boolean" default="false"/>
        <xs:attribute name="langcode" type="mce:LangType" use="required"/>
      </xs:extension>
    </xs:simpleContent>
  </xs:complexType>
</xs:schema>
```

Related resources for custom sensitive information types

- Learn about Microsoft Purview Data Loss Prevention
- Sensitive information type entity definitions
- Sensitive information type functions

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Remove a custom sensitive information type using PowerShell

In Security & Compliance PowerShell, there are two methods to remove custom sensitive information types:

- **Remove individual custom sensitive information types**: Use the method documented in Modify a custom sensitive information type using PowerShell. Export the custom rule package containing the custom sensitive information type. Remove the sensitive information type from the XML file, and then import the updated XML file back into the existing custom rule package.
- **Remove a custom rule package and all custom sensitive information types that it contains**: Removing an entire custom rule package is documented in this section.

> **Note**
>
> Before your remove a custom sensitive information type, verify that no DLP policies or Exchange mail flow rules (also known as transport rules) still reference the sensitive information type.

1. Security & Compliance PowerShell
2. To remove a custom rule package, use the Remove-DlpSensitiveInformationTypeRulePackage cmdlet. Run the following command to remove the specified rule package when you no longer need it or want to replace it:

   ```
   PowerShell
   Remove-DlpSensitiveInformationTypeRulePackage -Identity "RulePackageIdentity"
   ```

   You can use the Name value (for any language) or the `RulePack id` (GUID) value to identify the rule package.

   The following example removes a custom rule package named "Employee ID Custom Rule Pack":

   ```
   PowerShell
   ```

```
Remove-DlpSensitiveInformationTypeRulePackage -Identity "Employee ID Custom Rule Pack"
```

For detailed syntax and parameter information, see the Remove-DlpSensitiveInformationTypeRulePackage cmdlet reference.

3. To verify that you've successfully removed a custom sensitive information type, do any of the following steps:

   - Run the Get-DlpSensitiveInformationTypeRulePackage cmdlet and verify the rule package is no longer listed:

     PowerShell

     ```
     Get-DlpSensitiveInformationTypeRulePackage
     ```

   - Run the Get-DlpSensitiveInformationType cmdlet to verify the sensitive information types in the removed rule package are no longer listed:

     PowerShell

     ```
     Get-DlpSensitiveInformationType
     ```

     For custom sensitive information types, the Publisher property value will be something other than "Microsoft Corporation".

   - Replace <Name> with the Name value of the sensitive information type (for example, Employee ID) and run the Get-DlpSensitiveInformationType cmdlet to verify the sensitive information type is no longer listed:

     PowerShell

     ```
     Get-DlpSensitiveInformationType -Identity "<Name>"
     ```

Related resources

- Learn about Microsoft Purview Data Loss Prevention
- Sensitive information type entity definitions
- Sensitive information type functions

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Modify a custom sensitive information type using PowerShell

Before you begin, connect to Security & Compliance PowerShell. For instructions, see Security & Compliance PowerShell.

In Security & Compliance PowerShell, modifying a custom sensitive information type requires you to:

1. Export the existing rule package that contains the custom sensitive information type to an XML file (or use the existing XML file if you have it).
2. Modify the custom sensitive information type in the exported XML file.
3. Import the updated XML file back into the existing rule package.

To connect to Security & Compliance PowerShell, see Security & Compliance PowerShell.

Step 1: Export the existing rule package to an XML file

Use the following steps to export the existing rule package to an XML file.

> **Note**
>
> If you have a copy of the XML file (for example, if you just created and imported it), you can skip to Step: 2 Modify the sensitive information type in the exported XML file.

1. If you don't already know the name of the custom rule package, run the Get-DlpSensitiveInformationTypeRulePackage cmdlet to find it:

   PowerShell

   ```
   Get-DlpSensitiveInformationTypeRulePackage
   ```

   > **Note**
   >
   > The built-in rule package that contains the built-in sensitive information types is named *Microsoft Rule Package*. The rule package that contains the custom sensitive

information types that you created in the Compliance center UI is named *Microsoft.SCCManaged.CustomRulePack*.

2. Use the Get-DlpSensitiveInformationTypeRulePackage cmdlet to store the custom rule package to a variable:

```
PowerShell
$rulepak = Get-DlpSensitiveInformationTypeRulePackage -Identity "RulePackageName"
```

For example, if the name of the rule package is "Employee ID Custom Rule Pack", run the following cmdlet:

```
PowerShell
$rulepak = Get-DlpSensitiveInformationTypeRulePackage -Identity "Employee ID Custom Rule Pack"
```

3. Use the following syntax to export the custom rule package to an XML file:

```
PowerShell
[System.IO.File]::WriteAllBytes('XMLFileAndPath', $rulepak.SerializedClassificationRuleCollection)
```

This example exports the rule package to the file named ExportedRulePackage.xml in the C:\My Documents folder.

```
PowerShell
[System.IO.File]::WriteAllBytes('C:\My Documents\ExportedRulePackage.xml', $rulepak.SerializedClassificationRuleCollection)
```

Step 2: Modify the sensitive information type in the exported XML file

For XML schema details and editing guidance, see Customize a built-in sensitive information type.

Step 3: Import the updated XML file back into the existing rule package

To import the updated XML back into the existing rule package, use the Set-DlpSensitiveInformationTypeRulePackage cmdlet:

PowerShell

```
Set-DlpSensitiveInformationTypeRulePackage -FileData
([System.IO.File]::ReadAllBytes('C:\My Documents\External Sensitive Info Type Rule
Collection.xml'))
```

For detailed syntax and parameter information, see Set-DlpSensitiveInformationTypeRulePackage.

Related content

- Learn about Microsoft Purview Data Loss Prevention
- Sensitive information type entity definitions
- Sensitive information type functions

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Create a keyword dictionary

Microsoft Purview can identify, monitor, and protect your sensitive items. Identifying sensitive items sometimes requires looking for keywords, particularly when identifying generic content (such as healthcare-related communication), or inappropriate or explicit language. Although you can create keyword lists when you create custom sensitive information types, keyword lists are limited in size and if you're creating custom sensitive information types in PowerShell, require modifying XML to create or edit them.

In contrast, keyword dictionaries provide simpler management of keywords and at a larger scale, supporting up to 1 MB of terms (post-compression) in the dictionary. Additionally, keyword dictionaries can support any language. The tenant limit is also 1 MB after compression. A post-compression limit of 1 MB means that all dictionaries combined across a tenant can have close to one million characters.

Keyword dictionary limits

You can create keyword dictionary, subject to a combined size limit of 1MB (post compression) per tenant. To find out how many keyword dictionaries you have in your tenant, connect to Security & Compliance PowerShell and then run this PowerShell script:

PowerShell

```
$rawFile = $env:TEMP + "\rule.xml"

$kd = Get-DlpKeywordDictionary
$ruleCollections = Get-DlpSensitiveInformationTypeRulePackage
[System.IO.File]::WriteAllBytes((Resolve-Path $rawFile),
$ruleCollections.SerializedClassificationRuleCollection)
$UnicodeEncoding = New-Object System.Text.UnicodeEncoding
$FileContent = [System.IO.File]::ReadAllText((Resolve-Path $rawFile),
$unicodeEncoding)

if($kd.Count -gt 0)
{
$count = 0
$entities = $FileContent -split "Entity id"
for($j=1;$j -lt $entities.Count;$j++)
{
for($i=0;$i -lt $kd.Count;$i++)
{
$Matches = Select-String -InputObject $entities[$j] -Pattern $kd[$i].Identity -
AllMatches
```

```
$count = $Matches.Matches.Count + $count
if($Matches.Matches.Count -gt 0) {break}
}
}

Write-Output "Total Keyword Dictionary SIT:"
$count
}
else
{
$Matches = Select-String -InputObject $FileContent -Pattern $kd.Identity -AllMatches
Write-Output "Total Keyword Dictionary SIT:"
$Matches.Matches.Count
}

Remove-Item $rawFile
```

Basic steps to creating a keyword dictionary

Most commonly you compile your keywords for your dictionary in a file, such as a .csv or .txt list. You upload the dictionary file into a SIT during creation or editing or import them via a PowerShell cmdlet. Alternatively, you can start from an existing keyword dictionary. Lastly, you can enter keywords manually in the **Add keyword dictionary** dialog. When you create a keyword dictionary, you follow the same core steps:

Create a keyword dictionary using the Microsoft Purview portal

Use these steps to create or import keywords for a custom dictionary:

1. Sign in to the Microsoft Purview portal **Information Protection** > **Classifiers** > **Sensitive info types**.
2. Select + **Create sensitive info type** and then enter a **Name** and **Description** for your sensitive info type. Choose **Next**.
3. On the **Define patterns for this sensitive info type** page, choose + **Create pattern**.
4. In the **New pattern** window, select a **Confidence level**.
5. Choose **Add a Primary element** and select **Keyword dictionary**.
6. On the **Add a keyword dictionary** flyout, you can:
   a. **Upload a dictionary** file in **TXT** or **CSV** format.
   b. **Choose from existing dictionaries**.

c. or create a new dictionary by entering keywords manually and giving it a name.

7. Still in the **New Pattern** window, for **Character proximity**, specify how far away (in number of characters) that any supporting elements must be to be detected. The closer the primary and supporting elements are to each other, the more likely the detected content is going to be what you're looking for.

8. Add the **Supporting elements** you wish to use to increase the accuracy of detecting what you're looking for.

9. Add any **Additional checks** and then choose **Create**.

10. Choose **Next** to continue creating your sensitive information type. When you're finished, choose **Done**.

Create a keyword dictionary from a file using PowerShell

Often when you need to create a large dictionary, it's so you can use keywords from a file or a list exported from some other source. This procedure creates a keyword dictionary containing a list of diseases to screen in external email. To begin, you need to connect to Security & Compliance PowerShell.

1. Copy your keywords into a text file and make sure that each keyword is on a separate line.

2. Save the text file with Unicode encoding. In Notepad, navigate to > **Save As** > **Encoding** > **Unicode**.

3. Read the file into a variable by running this cmdlet:

```
PowerShell
$fileData = [System.IO.File]::ReadAllBytes('<filename>')
```

4. Create the dictionary by running this cmdlet:

```
PowerShell
New-DlpKeywordDictionary -Name <name> -Description <description> -FileData
$fileData
```

Using keyword dictionaries in custom sensitive information types and DLP policies

Keyword dictionaries can be used as part of the match requirements for a custom sensitive information type, or as a sensitive information type themselves. Both require you to create a custom sensitive information type using PowerShell. Follow the instructions in Create a custom sensitive information type using PowerShell to create a sensitive information type. Once you have the XML, you need the GUID identifier from the XML in order to use the keyword dictionary.

XML

```
<Entity id="00aa00aa-bb11-cc22-dd33-44ee44ee44ee" patternsProximity="300"
recommendedConfidence="75">
    <Pattern confidenceLevel="75">
        <IdMatch idRef=". . ."/>
    </Pattern>
</Entity>
```

To get the identity of your dictionary, run this command and copy the **Identity** property value:

PowerShell

```
Get-DlpKeywordDictionary -Name "Diseases"
```

The output of the command looks like this:

```
RunspaceId : 138e55e7-ea1e-4f7a-b824-79f2c4252255
Identity : 8d2d44b0-91f4-41f2-94e0-21c1c5b5fc9f
Name : Diseases
Description : Names of diseases and injuries from ICD-10-CM lexicon
KeywordDictionary : aarskog's syndrome, abandonment, abasia, abderhalden-kaufmann-lignac,
abdominalgia, abduction contracture, abetalipo proteinemia, abiotrophy, ablatio, ablation,
ablepharia,abocclusion, abolition, aborter, abortion, abortus, aboulomania, abrami's
disease, abramo
IsValid : True
ObjectState : Unchanged
```

Paste the **identity** value into the XML for your custom sensitive information type as the **idRef**. Next, upload the XML file. Your dictionary now appears in your list of sensitive information types and you can use it right in your policy, specifying how many keywords are required to match.

XML

```
<Entity id="11bb11bb-cc22-dd33-ee44-55ff55ff55ff" patternsProximity="300"
recommendedConfidence="85">
      <Pattern confidenceLevel="85">
        <IdMatch idRef="8d2d44b0-91f4-41f2-94e0-21c1c5b5fc9f" />
      </Pattern>
    </Entity>
    <LocalizedStrings>
      <Resource idRef="11bb11bb-cc22-dd33-ee44-55ff55ff55ff">
        <Name default="true" langcode="en-us">Diseases</Name>
        <Description default="true" langcode="en-us">Detects various
diseases</Description>
      </Resource>
    </LocalizedStrings>
```

> **Note**
>
> Microsoft 365 Information Protection supports double-byte character set languages for:
>
> - Chinese (simplified)
> - Chinese (traditional)
> - Korean
> - Japanese
>
> This support is available for sensitive information types. See, Information protection support for double byte character sets release notes (preview) for more information.

> **Tip**
>
> To detect patterns containing Chinese/Japanese characters and single byte characters or to detect patterns containing Chinese/Japanese and English, define two variants of the keyword or regex.
>
> - For example, to detect a keyword like "机密的document", use two variants of the keyword; one with a space between the Japanese and English text and another without a space between the Japanese and English text. So, the keywords to be added in the SIT should be "机密的 document" and "机密的document". Similarly, to detect a phrase "東京オリンピック2020", two variants should be used; "東京オリンピック 2020" and "東京オリンピック2020".
>
> Along with Chinese/Japanese/double byte characters, if the list of keywords/phrases also contains non-Chinese/Japanese words also (for instance, stand-alone English words), you

should create two dictionaries/keyword lists. One for keywords containing Chinese/Japanese/double byte characters and another one for English words.

- For example, if you want to create a keyword dictionary/list with three phrases "Highly confidential", "機密性が高い" and "机密的document", you should create two keyword lists.

    1. Highly confidential
    2. 機密性が高い, 机密的document and 机密的 document

While creating a regex using a double byte hyphen or a double byte period, make sure to escape both the characters like one would escape a hyphen or period in a regex. Here's a sample regex for reference:

- `(?<!\d)([4][0-9]{3}[\-?\-\t]*[0-9]{4}`

We recommend using a string match instead of a word match in a keyword list.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/22/2026

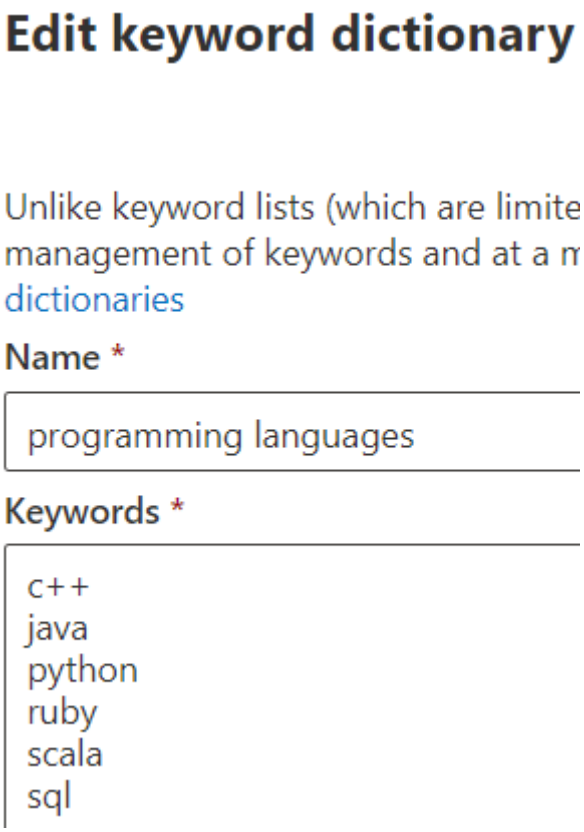
Modify a keyword dictionary

This article explains how to modify keywords in custom or built-in keyword dictionaries used by sensitive information types (SITs) in Microsoft Purview. You can modify a keyword dictionary by using the Microsoft Purview portal or PowerShell.

Modify a keyword dictionary in the portals

Keyword dictionaries can be used as **Primary elements** or **Supporting elements** in sensitive information type (SIT) patterns. You can edit a keyword dictionary while creating a SIT or in an existing custom SIT. For example to edit an existing keyword dictionary:

1. Sign in to the Microsoft Purview portal > **Information Protection** > **Classifiers** > **Sensitive info types**.
2. Choose the SIT that uses the keyword dictionary you want to update.
3. Select **Edit**.
4. Choose **Next**.
5. **Edit** the pattern that has the keyword dictionary you want to update.
6. Edit the keyword dictionary in the **Primary element** or the **Supporting element** sections. Make your edits, using one keyword per line.



7. Choose **Done**.

Modify a keyword dictionary using PowerShell

For example, we'll modify some terms in PowerShell, save the terms locally where you can modify them in an editor, and then update the previous terms in place.

First, retrieve the dictionary object:

PowerShell

```
$dict = Get-DlpKeywordDictionary -Name "Diseases"
```

Printing `$dict` will show the various properties. The keywords themselves are stored in an object on the backend, but `$dict.KeywordDictionary` contains a string representation of them, which you'll use to modify the dictionary.

Before you modify the dictionary, you need to turn the string of terms back into an array using the `.split(',')` method. Then you'll clean up the unwanted spaces between the keywords with the `.trim()` method, leaving just the keywords to work with.

PowerShell

```
$terms = $dict.KeywordDictionary.split(',').trim()
```

Now you'll remove some terms from the dictionary. Because the example dictionary has only a few keywords, you could as easily skip to exporting the dictionary and editing it in Notepad, but dictionaries generally contain a large amount of text, so you'll first learn this way to edit them easily in PowerShell.

In the previous command, you saved the keywords to the `$terms` array. There are several ways to remove items from an array, but as a straightforward approach, you'll create an array of the terms you want to remove from the dictionary, and then copy only the dictionary terms to it that aren't in the list of terms to remove.

Run the command `$terms` to show the current list of terms. The output of the command looks like this:

PowerShell

```
aarskog's syndrome
abandonment
abasia
abderhalden-kaufmann-lignac
abdominalgia
abduction contracture
abetalipoproteinemia
abiotrophy
ablatio
ablation
ablepharia
abocclusion
abolition
aborter
abortion
abortus
aboulomania
abrami's disease
```

Run this command to specify the terms that you want to remove:

PowerShell

```
$termsToRemove = @('abandonment','ablatio')
```

Run this command to actually remove the terms from the list:

PowerShell

```
$updatedTerms = $terms | Where-Object {$_ -notin $termsToRemove}
```

Run the command `$updatedTerms` to show the updated list of terms. The output of the command looks like this (the specified terms have been removed):

PowerShell

```
aarskog's syndrome
abasia
abderhalden-kaufmann-lignac
abdominalgia
abduction contracture
abetalipoproteinemia
abiotrophy
ablation
ablepharia
abocclusion
abolition
aborter
abortion
abortus
aboulomania
abrami's disease
```

Now save the dictionary locally and add a few more terms. You could add the terms right here in PowerShell, but you'll still need to export the file locally to ensure it's saved with Unicode encoding and contains the BOM.

Save the dictionary locally by running the following:

PowerShell

```
Set-Content $updatedTerms -Path "C:\myPath\terms.txt"
```

Now open the file, add your other terms, and save with Unicode encoding (UTF-16). Now you'll upload the updated terms and update the dictionary in place.

PowerShell

```
Set-DlpKeywordDictionary -Identity "Diseases" -FileData
([System.IO.File]::ReadAllBytes('C:myPath\terms.txt'))
```

Now the dictionary has been updated in place. The `Identity` field takes the name of the dictionary. If you wanted to also change the name of your dictionary using the `Set-` cmdlet, you would just need to add the `-Name` parameter to the `Set-DlpKeywordDictionary` command with your new dictionary name.

See also

- Create a keyword dictionary
- Create a custom sensitive information type

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Common usage scenarios for sensitive information types

This article describes how to implement some common sensitive information type (SIT) use case scenarios. You can use these procedures as examples and adapt them to your specific needs.

Protect credit card numbers

Contoso Bank needs to classify the credit card numbers that they issue as sensitive. Their credit cards start with a set of six-digit patterns. They would like to customize the out-of-the-box credit card definition to detect only those credit card numbers starting with their six-digit patterns.

Suggested solution

1. Create a copy of the credit card SIT. Use the steps to copy and modify an existing SIT to copy the credit card SIT.
2. Edit the high confidence pattern. Follow the steps in Edit or delete a SIT pattern.
3. Add the **starts with** check and add the list of pattern digits (formatted & unformatted). For example to ensure that the SIT only considers credit cards starting with 411111 & 433512 should be considered valid, add the following to the list 4111 11, 4111-11, 411111, 4335 12, 4335-12, 433512.
4. Repeat step 2 & 3 for the low confidence pattern.

Test numbers similar to Social Security numbers

Contoso has identified a few nine-digit test numbers that trigger false positive matches in the Social Security Number (SSN) Microsoft Purview Data Loss Prevention (DLP) policy. They would like to exclude these numbers from the list of valid matches SSN matches.

Suggested solution

1. Create a copy of the SSN SIT. Use the steps to copy and modify an existing SIT to copy the SSN SIT.
2. Edit the high confidence pattern. Follow the steps in Edit or delete a SIT pattern.
3. Add the numbers you want to exclude in the **exclude specific values** check. For example, to exclude 239-23-532 & 23923532, just adding 23923532 is sufficient.
4. Repeat step 2 & 3 for other confidence patterns as well

Prevent phone numbers in signatures from triggering matches

Australia-based Contoso finds that phone numbers in email signatures are triggering a match for their Australia company number DLP policy.

Suggested solution

Add a 'NOT' group in supporting elements using a keyword list containing commonly used keywords in email signatures, such as *Phone*, *Mobile*, *email*, *Thanks and regards*, etc. Keep the proximity of this keyword list to a smaller value (for instance, 50 characters) for better accuracy. For more information, see Get started with custom sensitive information types.

Troubleshoot why an ABA routing policy doesn't trigger

DLP policy is unable to trigger ABA routing number policy in large excel files because the required keyword isn't within 300 characters.

Suggested solution

Create a copy of the built-in SIT and edit it to change the proximity of the keyword list from *300 characters* to *Anywhere in the document*.

> **Tip**
>
> You may edit the keyword list to include/exclude keywords that are relevant to your organization.

Unable to detect credit card numbers with unusual delimiters

Contoso Bank has noticed some of their employees share credit card numbers with '/' as a delimiter, for example, 4111/1111/1111/1111, which the out-of-the-box credit card definition doesn't detect. Contoso would like to define their own regex and validate it using LuhnCheck.

Suggested solution

1. Create a copy of the credit card SIT using the steps in Customize a built-in sensitive information type

2. Add a new pattern
3. In the primary element, select **regular expression**
4. Define the regular expression that includes '/' as part of the regular expression; choose **validator** and then select **luhncheck** or **func_credit_card** to ensure the regex also passes the LuhnCheck.

Ignore disclaimer notices when configuring sensitive information type detection

Many organizations add legal disclaimers, disclosure statements, signatures, or other information to the top or bottom of email messages that enter or leave their organizations. In some cases, emails sent within an organization itself can contain such text. For example, employees might add signatures with motivational quotes, social messages, and so on. A disclaimer or signature can contain the terms that are present in the lexicon of a CC and might generate many false positives.

For example, a typical disclaimer might contain words like *sensitive*, or *confidential*. Policies looking for sensitive info will detect such disclaimers as incidents, leading to lot of false positives. Thus, providing customers with an option to ignore disclaimers can reduce the number of false positives and increase the efficiency of the compliance team.

Review an example disclaimer notice

Consider the following disclaimer:

"IMPORTANT NOTICE: This e-mail message is intended to be received only by persons entitled to receive the confidential information it may contain. E-mail messages to clients of Contoso may contain information that is confidential and legally privileged. Please do not read, copy, forward, or store this message unless you are an intended recipient. If you have received this message in error, please forward it to the sender and delete it completely from your computer system."

If the SIT is configured to detect *confidential* as a keyword, the SIT pattern invokes a match every time an email includes the disclaimer, leading to considerable number of false positives.

Ignore disclaimer using prefix and suffix in SIT

One way to ignore keyword instances in disclaimer text is to exclude keywords that are preceded by a specific prefix and followed by a specific suffix.

Consider this disclaimer:

"IMPORTANT NOTICE: This e-mail message is intended to be received only by persons *entitled to receive the* confidential **information it may contain**. E-mail messages to clients of Contoso may contain information that is confidential and legally privileged. Please do not read, copy, forward, or store this message unless you are an intended recipient. If you have received this message in error, please forward it to the sender and delete it completely from your computer system."

Say we have two instances of the keyword *confidential*. If we configure the SIT to ignore instances of the keyword *confidential* that are preceded by prefixes (italicized in the example) and followed by suffixes (bolded in the example), then we can successfully ignore disclaimers in most cases.

To ignore the disclaimer using prefix and suffix:

1. Add additional checks in the current SIT to exclude prefix and suffix text to the keyword that we want to ignore in the disclaimer.
2. Choose to exclude the prefix and in the **Prefixes** text box enter **contain information that is**.
3. Choose to exclude the suffix and in the **Suffixes** text box enter **and legally privileged**.
4. Repeat this process for other instances of the keywords in the disclaimer, as shown in the following graphic.

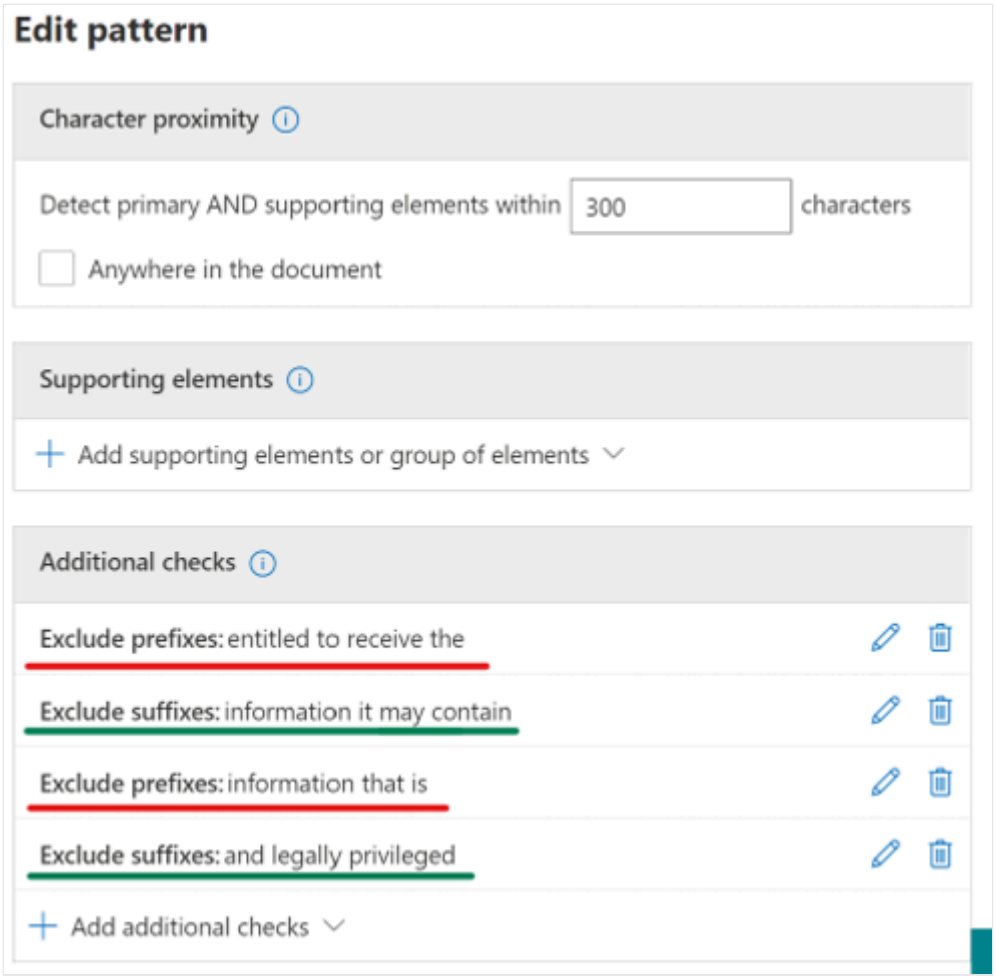
Ignore disclaimer by excluding secondary elements

Another way to add a list of supporting elements (keyword instances found in disclaimer text) that need to be excluded is to exclude secondary elements.

Consider this disclaimer:

"IMPORTANT NOTICE: This e-mail message is intended to be received only by persons entitled to receive the confidential information it may contain. E-mail messages to clients of Contoso may contain information that is confidential and legally privileged. Please do not read, copy, forward, or store this message unless you are an intended recipient. If you have received this message in error, please forward it to the sender and delete it completely from your computer system."

We have two instances of the keyword “confidential” in this example. If we configure the SIT to ignore instances of this keyword in the disclaimer (underlined as red in the following image), we can achieve ignoring disclaimers in most of the cases.



To ignore the disclaimer using secondary elements:

1. Select **Not any of these** group in the supporting elements.
2. Add the instances of disclaimer that we want to ignore as a keyword list/dictionary.
3. Add the keywords as a new line that we want to ignore. Remember that the length of each text can't be more than 50 characters.
4. Set the proximity of this element to be within 50-60 characters of the primary element.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Sensitive information type entity definitions

This article is a list of all sensitive information type (SIT) entity definitions. Each link takes you to the definition of that specific SIT and shows what a DLP policy looks for to detect each type. To learn more about sensitive information types, see Sensitive information types.

> **Note**
>
> Mapping of confidence level (high/medium/low) with accuracy number (numeric value of 1 to 100)
>
> - Low confidence: 65 or below
> - Medium confidence: 75
> - High confidence: 85

- ABA routing number
- All credentials
- All full names
- All medical terms and conditions
- All Physical Addresses
- Amazon S3 Client Secret Access Key
- Argentina national identity (DNI) number
- Argentina Unique Tax Identification Key (CUIT/CUIL)
- ASP.NET machine Key
- Australia bank account number
- Australia business number
- Australia company number
- Australia drivers license number
- Australia medical account number
- Australia passport number
- Australia physical addresses
- Australia tax file number
- Austria drivers license number
- Austria identity card
- Austria passport number
- Austria physical addresses

- Austria social security number
- Austria tax identification number
- Austria value added tax
- Azure App Service deployment password
- Azure Batch shared access key
- Azure Bot Framework secret key
- Azure Bot service app secret
- Azure Cognitive Search API key
- Azure Cognitive Service key
- Azure Container Registry access key
- Azure Cosmos DB account access key
- Azure Databricks personal access token
- Azure DevOps app secret
- Azure DevOps personal access token
- Azure DocumentDB auth key
- Azure EventGrid access key
- Azure Function Master / API key
- Azure IAAS database connection string and Azure SQL connection string
- Azure IoT connection string
- Azure IoT shared access key
- Azure Logic app shared access signature
- Azure Machine Learning web service API key
- Azure Maps subscription key
- Azure publish setting password
- Azure Redis cache connection string
- Azure Redis cache connection string password
- Azure SAS
- Azure service bus connection string
- Azure service bus shared access signature
- Azure Shared Access key / Web Hook token
- Azure SignalR access key
- Azure SQL connection string
- Azure storage account access key
- Azure storage account key
- Azure Storage account key (generic)
- Azure Storage account shared access signature
- Azure Storage account shared access signature for high risk resources

- Azure subscription management certificate
- Belgium driver's license number
- Belgium national number
- Belgium passport number
- Belgium physical addresses
- Belgium value added tax number
- Blood test terms
- Brand medication names
- Brazil CPF number
- Brazil legal entity number (CNPJ)
- Brazil national identification card (RG)
- Brazil physical addresses
- Bulgaria driver's license number
- Bulgaria passport number
- Bulgaria physical addresses
- Bulgaria uniform civil number
- Canada bank account number
- Canada driver's license number
- Canada health service number
- Canada passport number
- Canada personal health identification number (PHIN)
- Canada physical addresses
- Canada social insurance number
- Chile identity card number
- China resident identity card (PRC) number
- China physical addresses
- Client secret / API key
- Credit card number
- Colombia national ID
- Colombia tax identification number
- Croatia driver's license number
- Croatia identity card number
- Croatia passport number
- Croatia personal identification (OIB) number
- Croatia physical addresses
- Cyprus drivers license number
- Cyprus identity card

- Cyprus passport number
- Cyprus physical addresses
- Cyprus tax identification number
- Czech driver's license number
- Czech passport number
- Czech personal identity number
- Czech Republic physical addresses
- Denmark driver's license number
- Denmark passport number
- Denmark personal identification number
- Denmark physical addresses
- Diseases
- Drug Enforcement Agency (DEA) number
- Ecuador Unique Identification Number
- Estonia driver's license number
- Estonia passport number
- Estonia Personal Identification Code
- Estonia physical addresses
- EU debit card number
- EU driver's license number
- EU national identification number
- EU passport number
- EU social security number or equivalent identification
- EU Tax identification number
- Finland driver's license number
- Finland European health insurance number
- Finland national ID
- Finland passport number
- Finland physical addresses
- France driver's license number
- France health insurance number
- France national id card (CNI)
- France passport number
- France physical addresses
- France social security number (INSEE)
- France tax identification number
- France value added tax number

- General password
- General Symmetric key
- Generic medication names
- German Driver's License Number
- Germany identity card number
- German Passport Number
- Germany physical addresses
- Germany tax identification number
- Germany value added tax number
- GitHub Personal Access Token
- Google API key
- Greenland physical addresses
- Greece driver's license number
- Greece national ID card
- Greece passport number
- Greece physical addresses
- Greece Social Security Number (AMKA)
- Greek Tax Identification Number
- Hong Kong identity card (HKID) number
- Http authorization header
- Hungary driver's license number
- Hungary passport number
- Hungary personal identification number
- Hungary physical addresses
- Hungary social security number (TAJ)
- Hungary tax identification number
- Hungary value added tax number
- Iceland physical addresses
- Impairments Listed In The U.S. Disability Evaluation Under Social Security
- India driver's License Number
- India GST Number
- India permanent account number (PAN)
- India unique identification (Aadhaar) number
- India Voter Id Card
- Indonesia driver's license number
- Indonesia identity card (KTP) number
- Indonesia passport number

- International banking account number (IBAN)
- International classification of diseases (ICD-10-CM)
- International classification of diseases (ICD-9-CM)
- IP address
- IP Address v4
- IP Address v6
- Ireland driver's license number
- Ireland passport number
- Ireland personal public service (PPS) number
- Ireland physical addresses
- Israel bank account number
- Israel national identification number
- Italy driver's license number
- Italy fiscal code
- Italy passport number
- Italy physical addresses
- Italy value added tax number
- Japan bank account number
- Japan driver's license number
- Japanese My Number Corporate
- Japanese My Number Personal
- Japan passport number
- Japan physical addresses
- Japanese Residence Card Number
- Japan resident registration number
- Japan social insurance number (SIN)
- Lab test terms
- Latvia driver's license number
- Latvia passport number
- Latvia personal code
- Latvia physical addresses
- Liechtenstein physical addresses
- Lifestyles that relate to medical conditions
- Lithuania driver's license number
- Lithuania passport number
- Lithuania personal code
- Lithuania physical addresses

- Luxemburg driver's license number
- Luxemburg national identification number (natural persons)
- Luxemburg national identification number (non-natural persons)
- Luxemburg passport number
- Luxemburg physical addresses
- Malaysia identification card number
- Malaysia passport number
- Malta driver's license number
- Malta identity card number
- Malta passport number
- Malta physical addresses
- Malta tax identification number
- Medical specialities
- Medicare Beneficiary Identifier (MBI) card
- Mexico Unique Population Registry Code (CURP)
- Microsoft Bing maps key
- Microsoft Entra client access token
- Microsoft Entra client secret
- Microsoft Entra user Credentials
- Netherlands citizen's service (BSN) number
- Netherlands driver's license number
- Netherlands passport number
- Netherlands physical addresses
- Netherlands tax identification number
- Netherlands value added tax number
- New Zealand bank account number
- New Zealand driver's license number
- New Zealand inland revenue number
- New Zealand ministry of health number
- New Zealand physical addresses
- New Zealand social welfare number
- Norway identification number
- Norway physical addresses
- Philippines national identification number
- Philippines passport number
- Philippines unified multi-purpose identification number
- Poland driver's license number

- Poland identity card
- Poland national ID (PESEL)
- Poland Passport Number
- Poland physical addresses
- Poland REGON number
- Poland tax identification number
- Portugal citizen card number
- Portugal driver's license number
- Portugal passport number
- Portugal physical addresses
- Portugal tax identification number
- Qatari ID Card Number
- Romania driver's license number
- Romania passport number
- Romania personal numeric code (CNP)
- Romania physical addresses
- Russia passport number domestic
- Russia passport number international
- Russia physical addresses
- Russia taxpayer identification number
- Saudi Arabia National ID
- Singapore driver's license number
- Singapore passport number
- Singapore national registration identity card (NRIC) number
- Singapore physical addresses
- Slack access token
- Slovakia driver's license number
- Slovakia passport number
- Slovakia personal number
- Slovakia physical addresses
- Slovenia driver's license number
- Slovenia passport number
- Slovenia physical addresses
- Slovenia tax identification number
- Slovenia Unique Master Citizen Number
- South Africa identification number
- South Africa physical addresses

- South Korea driver's license number
- South Korea passport number
- South Korea resident registration number
- Spain DNI
- Spain driver's license number
- Spain passport number
- Spain physical addresses
- Spain social security number (SSN)
- Spain tax identification number
- SQL Server connection string
- Surgical procedures
- Sweden driver's license number
- Sweden national ID
- Sweden passport number
- Sweden physical addresses
- Sweden tax identification number
- SWIFT code
- Switzerland physical addresses
- Swiss Social Security Number AHV
- Taiwan National ID
- Taiwan passport number
- Taiwan-resident certificate (ARC/TARC) number
- Thai population identification code
- Turkey national identification number
- Turkey physical addresses
- Types of medication
- U.A.E. identity card number
- U.A.E. passport number
- U.K. driver's license number
- U.K. electoral roll number
- U.K. national health service number
- U.K. national insurance number (NINO)
- U.K. physical addresses
- U.K. Unique Taxpayer Reference Number
- U.S. bank account number
- U.S. driver's license number
- U.S. individual taxpayer identification number (ITIN)

- U.S. physical addresses
- U.S. social security number (SSN)
- U.S./U.K. passport number
- Ukraine passport domestic
- Ukraine passport international
- Ukraine physical addresses
- User login credentials
- X.509 certificate private key

Last updated on 06/15/2026

ABA routing number

Article • 08/19/2024

Format

nine digits that may be in a formatted or unformatted pattern

Pattern

- two digits in the ranges 00-12, 21-32, 61-72, or 80
- two digits
- an optional hyphen
- four digits
- an optional hyphen
- a digit

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function Func_aba_routing finds content that matches the pattern.
- A keyword from Keyword_ABA_Routing is found.

A DLP policy has low confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function Func_aba_routing finds content that matches the pattern.

XML

```
    <!-- ABA Routing Number -->
    <Entity id="cb353f78-2b72-4c3c-8827-92ebe4f69fdf"
patternsProximity="300" recommendedConfidence="75">
      <Pattern confidenceLevel="75">
        <IdMatch idRef="Func_aba_routing" />
        <Match idRef="Keyword_ABA_Routing" />
      </Pattern>
      <Pattern confidenceLevel="65">
        <IdMatch idRef="Func_aba_routing" />
      </Pattern>
    </Entity>
```

Keywords

Keyword_aba_routing

- aba number
- aba#
- aba
- abarouting#
- abaroutingnumber
- americanbankassociationrouting#
- americanbankassociationroutingnumber
- bankrouting#
- bankroutingnumber
- routing #
- routing no
- routing number
- routing transit number
- routing#
- RTN

All credentials sensitive information types

"All credentials" is a bundled-entity sensitive information type (SIT). It detects credentials from all supported services and environments, which include Amazon, Azure, GitHub, Google, Microsoft general, Slack and more.

> **Important**
>
> Because the **All Credetials** SIT is a bundled SIT, it ***does not*** contain any keywords. The keywords that it detects are based on the individual SITs it contains. Refer to the individual SITs for keyword information.

For information about implementing bundled SITs, see Learn about named entities.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a bundled SIT in an endpoint DLP policy. This requirement is specific to the combination of bundled SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a credential scanning SIT in an endpoint DLP policy. This requirement is specific to the combination of credential scanning SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

Format

Various

Pattern

Various

Checksum

No

Keyword Highlighting

Not supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Description

The All credentials SIT is a bundled SIT that scans for individual credential types. This includes checksum evaluated values and keyword/pattern combinations. The All credential entity scans for any of the underlying credential information type checksums and/or patterns are detected in a document.

Contains

This bundled entity SIT contains these individual SITs:

- Amazon S3 Client Secret Access Key
- ASP.NET Machine Key
- Microsoft Entra Client Access Token
- Microsoft Entra Client Secret
- Microsoft Entra user Credentials
- Azure App Service Deployment Password
- Azure Batch Shared Access Key
- Azure Bot Framework Secret Key
- Azure Bot Service App Secret
- Azure Cognitive Search API Key
- Azure Cognitive Service Key
- Azure Container Registry Access Key
- Azure COSMOS DB Account Access Key
- Azure Databricks Personal Access Token

- Azure DevOps App Secret
- Azure DevOps Personal Access Token
- Azure EventGrid Access Key
- Azure Function Master / API Key
- Azure IoT Shared Access Key
- Azure Logic App Shared Access Signature
- Azure Machine Learning Web Service API Key
- Azure Maps Subscription Key
- Azure Redis Cache Connection String Password
- Azure Service Bus Shared Access Signature
- Azure Shared Access Key / Web Hook Token
- Azure SignalR Access Key
- Azure SQL Connection String
- Azure Storage Account Access Key
- Azure Storage Account Shared Access Signature
- Azure Storage Account Shared Access Signature for High Risk Resources
- Azure Subscription Management Certificate
- Client Secret / API Key
- General Password

> **Tip**
>
> General Password detects password related keywords in English only.

- General Symmetric Key
- GitHub Personal Access Token
- Google API key
- Http Authorization Header
- Microsoft Bing Maps Key
- Slack Access Token
- User Login Credentials
- X.509 Certificate Private Key

Supported languages

- English
- Bulgarian

- Chinese
- Croatian
- Czech
- Danish
- Estonian
- Finnish
- French
- German
- Hungarian
- Icelandic
- Irish
- Italian
- Japanese
- Latvian
- Lithuanian
- Maltese
- Dutch
- Norwegian
- Polish
- Portuguese
- Romanian
- Slovak
- Slovenian
- Spanish
- Swedish
- Turkish

Last updated on 06/15/2026

All full names

All full names is a bundled named entity. It detects full names for people from all supported countries/regions, which include Australia, China, Japan, U.S., and countries/regions in the EU. Use this SIT to detect all possible matches of full names.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a bundled SIT in an endpoint DLP policy. This requirement is specific to the combination of bundled SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

Format

Various.

Pattern

Various.

Checksum

No.

Keyword Highlighting

Not supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Description

This named entity SIT matches personal names that a human would identify as a name with high confidence. For example, if a string is found consisting of a given name and is followed by a family name then a match is made with high confidence. It uses three primary resources:

- A dictionary of given names.
- A dictionary of family names.
- Patterns of how names are formed.

The three resources are different for each country/region. The strings *Olivia Wilson* would trigger a match. Common given/family names are given a higher confidence than rarer names. However, the pattern also allows partial matches. If a given name from the dictionary is found and it's followed by a family name that isn't in the dictionary, then a partial match is triggered. For example, *Tomas Richard* would trigger a partial match. Partial matches are given lower confidence.

In addition, patterns that a human would see as indicative of names are also matched with appropriate confidence. Like *O. Wilson*, *O.P. Wilson*, *Dr. O. P. Wilson*, *Wilson, O.P.* or *T. Richard, Jr.* would be matches.

Supported languages

- English
- Bulgarian
- Croatian
- Czech
- Danish
- Estonian
- Finnish
- French
- German
- Hungarian
- Icelandic

- Irish
- Italian
- Japanese
- Latvian
- Lithuanian
- Maltese
- Dutch
- Norwegian
- Polish
- Portuguese
- Romanian
- Slovak
- Slovenian
- Spanish
- Swedish
- Turkish

Last updated on 11/18/2025

All medical terms and conditions

All medical terms and conditions is a bundled named entity that detects medical terms and medical conditions. It detects English terms only. Use this SIT to detect all possible matches of medical terms and conditions.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a bundled SIT in an endpoint DLP policy. This requirement is specific to the combination of bundled SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

Format

Dictionary

Pattern

Dictionary

Checksum

No

Keyword Highlighting

Not supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Description

This bundled named entity matches text that mentions medical conditions that are present in curated dictionaries. There's one curated dictionary per supported language. The dictionaries are from many international medical resources. The dictionaries include as many medical conditions as possible without risking a large number of false positives. Each entry contains the different forms that a single condition is commonly written in to ensure coverage, for example:

- *TB*
- *tuberculosis*
- *phthisis pulmonalis*

Contains

This bundled named entity SIT contains these individual SITs.

- Blood test terms
- Types of medication
- Diseases
- Generic medication names
- Impairments listed in the U.S. Disability Evaluation Under Social Security
- Lab test terms
- Lifestyles that relate to medical conditions
- Medical specialties
- Surgical procedures
- Brand medication names

Last updated on 11/18/2025

All Physical Addresses

All physical addresses are a bundled entity SIT, which detects patterns related to physical addresses from all supported countries/regions.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a bundled SIT in an endpoint DLP policy. This requirement is specific to the combination of bundled SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

Format

Various

Pattern

Various

Checksum

No

Keyword Highlighting

Not supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following services:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Description

The matching of street addresses is designed to match strings that a human would identify as a street address. To achieve this matching, it uses several primary resources:

- A dictionary of settlements, counties, and regions.
- A dictionary of street suffixes, like Road, Street, or Avenue.
- Patterns of postal codes.
- Patterns of address formats.

The resources are different for each country/region. The primary resources are the patterns of address formats that are used in a given country/region. Different formats are chosen to make sure that as many addresses as possible are matched. These formats allow flexibility, for example, an address may omit the postal code or omit a town name or have a street with no street suffix. In all cases, such matches are used to increase the confidence of the match.

The patterns are designed to match individual single addresses, not generic locations. So strings such as *Redmond, WA 98052*, or *Main Street, Albuquerque* aren't matched.

Contains

This bundled named entity SIT contains these individual SITs:

- Australia physical addresses
- Austria physical addresses
- Belgium physical addresses
- Brazil physical addresses
- Bulgaria physical addresses
- Canada physical addresses
- China physical addresses
- Croatia physical addresses
- Cyprus physical addresses
- Czech Republic physical addresses
- Denmark physical addresses

- Estonia physical addresses
- Finland physical addresses
- France physical addresses
- Germany physical addresses
- Greece physical addresses
- Greenland physical addresses
- Hungary physical addresses
- Iceland physical addresses
- Ireland physical addresses
- Italy physical addresses
- Japan physical addresses
- Latvia physical addresses
- Liechtenstein physical addresses
- Lithuania physical addresses
- Luxembourg physical addresses
- Malta physical addresses
- Netherlands physical addresses
- New Zealand physical addresses
- Norway physical addresses
- Poland physical addresses
- Portugal physical addresses
- Romania physical addresses
- Russia physical addresses
- Singapore physical addresses
- Slovakia physical addresses
- Slovenia physical addresses
- South Africa physical addresses
- Spain physical addresses
- Sweden physical addresses
- Switzerland physical addresses
- Turkey physical addresses
- Ukraine physical addresses
- United Kingdom physical addresses
- United States physical addresses

Supported languages

- English
- Bulgarian
- Chinese
- Croatian
- Czech
- Danish
- Estonian
- Finnish
- French
- German
- Hungarian
- Icelandic
- Irish
- Italian
- Japanese
- Latvian
- Lithuanian
- Maltese
- Dutch
- Norwegian
- Polish
- Portuguese
- Romanian
- Slovak
- Slovenian
- Spanish
- Swedish
- Turkish

Last updated on 06/19/2026

Amazon S3 client secret access key

This SIT is also included in the All credentials bundled SIT.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a credential scanning SIT in an endpoint DLP policy. This requirement is specific to the combination of credential scanning SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

Format

A combination of 40 characters consisting of letters, digits, and special characters.

Pattern

13-digit number:

A combination of 40 characters consisting of:

- a-z (case insensitive)
- 0-9
- forward slash (/) or plus sign (+)

for example:

`Aa1Bb~2Cc3.-Dd4Ee5Ff6Gg7Hh8Ii9_Jj0Kk1Ll2`

Credential example

Expand table

Confidence Band	Example
High	`Amazon AWS Access Key: Aa1Bb~2Cc3.-Dd4Ee5Ff6Gg7Hh8Ii9_Jj0Kk1Ll2`
Medium	`Client secret (AWS): Aa1Bb~2Cc3.-Dd4Ee5Ff6Gg7Hh8Ii9_Jj0Kk1Ll2`

Confidence Band	Example
Low	N/A

Checksum

No

SITs that have checksums use a unique calculation to check if the information is valid. This means when the **Checksum** value is **Yes**, the service can make a positive detection based on the sensitive data alone. When the **Checksum** value is **No** additional (secondary) elements must also be detected for the service to make a positive detection.

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

This SIT is designed to match the security information that's used to access Amazon Web Services.

It uses several primary resources:

- Patterns of Base64 encoded 240-bits symmetric key.
- Patterns of CredentialName, CredentialFeatures, AccountIdentityName, AccountIdentityValue, ResourceType, ResourceName.
- Patterns of mockup values, redactions, and placeholders.
- A dictionary of vocabulary words.

The patterns are designed to match actual credentials with reasonable confidence. The patterns don't match credentials formatted as examples. Mockup values, redacted values, and placeholders, like credential type or usage descriptions, in the position where an actual secret value should present won't be matched.

Keywords

Keyword_SymmetricKey240:

- secret
- key

Last updated on 11/18/2025

Argentina national identity (DNI) number

Article • 08/19/2024

Format

Eight digits with or without periods

Pattern

Eight digits:

- two digits
- an optional period
- three digits
- an optional period
- three digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression Regex_argentina_national_id finds content that matches the pattern.
- A keyword from Keyword_argentina_national_id is found.

XML

```xml
<!-- Argentina National Identity (DNI) Number -->
<Entity id="eefbb00e-8282-433c-8620-8f1da3bffdb2" recommendedConfidence="75"
patternsProximity="300">
   <Pattern confidenceLevel="75">
      <IdMatch idRef="Regex_argentina_national_id"/>
      <Match idRef="Keyword_argentina_national_id"/>
  </Pattern>
</Entity>
```

Keywords

Keyword_argentina_national_id

- Argentina National Identity number
- cedula
- cédula
- dni
- documento nacional de identidad
- documento número
- documento numero
- registro nacional de las personas
- rnp

Argentina Unique Tax Identification Key (CUIT/CUIL)

Article • 08/19/2024

Format

11 digits with dash

Pattern

11 digits with a dash:

- two digits in 20, 23, 24, 27, 30, 33 or 34
- a hyphen (-)
- eight digits
- a hyphen (-)
- one check digit

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_Argentina_Unique_Tax_Key` finds content that matches the pattern.
- A keyword from `Keyword_Argentina_Unique_Tax_Key` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_Argentina_Unique_Tax_Key` finds content that matches the pattern.

XML

```
    <!-- Argentina Unique Tax Identification Key (CUIT/CUIL) -->
      <Entity id="98da3da1-9199-4571-b7c4-b6522980b507"
patternsProximity="300" recommendedConfidence="75" relaxProximity="true">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_Argentina_Unique_Tax_Key" />
          <Match idRef="Keyword_Argentina_Unique_Tax_Key" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_Argentina_Unique_Tax_Key" />
        </Pattern>
      </Entity>
```

Keywords

Keyword_Argentina_Unique_Tax_Key

- Clave Unica de Identificacion Tributaria
- CUIT
- unique code of labour identification
- Clave Única de Identificación Tributaria
- unique labour identification code
- CUIL
- Unique Tax Identification Key
- Unique Labour Identification Key
- Unique Key of Labour Identification
- Unique Work Identification Code
- Unique Code of Work Identification
- Unique Work Identification Key
- Unique Key of Work Identification
- Unique Code of Tax Identification
- Unique Key of Tax Identification
- Unique Labor Identification Code
- Unique Code of Labor Identification
- Unique Labor Identification Key
- Unique Key of Labor Identification

- tax ID
- taxID#
- taxId
- taxidnumber
- tax number
- tax no
- tax #
- tax#
- taxpayer ID
- taxpayer number
- taxpayer no
- taxpayer #
- taxpayer#
- tax identity
- tax identification
- Número de Identificación Fiscal
- número de contribuyente

ASP.NET machine key

This sensitive information type (SIT) is also included in the All credentials bundled SIT.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a credential scanning SIT in an endpoint DLP policy. This requirement is specific to the combination of credential scanning SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

Format

Symmetric keys in XML configuration.

Pattern

Various symmetric key formats in XML, for example:

XML

```
<machineKey decryptionKey="******** </br>
<machineKey validationKey="********
```

Credential example

XML

```
<configuration>
    <system.web>
        <machineKey
            validationKey="Validation Key here"
            decryptionKey="Decryption key here"
            validation="SHA1"
            decryption="AES" />
    </system.web>
</configuration>
```

See **How to generate a `<machineKey> element`** in the Resolving view state message authentication code (MAC) errors support article for information on generating keys.

Checksum

No

SITs that have checksums use a unique calculation to check if the information is valid. This means when the **Checksum** value is **Yes**, the service can make a positive detection based on the sensitive data alone. When the **Checksum** value is **No** additional (secondary) elements must also be detected for the service to make a positive detection.

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

This SIT is designed to match the security information that's used to encrypt or hash data in ASP.NET forms authentication and view state.

It uses several primary resources:

- Patterns of Symmetric key context in xml files.
- Patterns of CredentialName, CredentialFeatures, AccountIdentityName, AccountIdentityValue, ResourceType, ResourceName.

The patterns are designed to match actual credentials with reasonable confidence. The patterns don't match credentials formatted as examples. Mockup values, redacted values, and placeholders, like credential type or usage descriptions, in the position where an actual secret value should present won't be matched.

Keywords

Keyword_SymmetricKeyContextInXml:

- password

- key
- connectionstring

Last updated on 06/01/2026

Australia bank account number

Article • 08/19/2024

Format

six to 10 digits with or without a bank state branch number

Pattern

Account number is 6 to 10 digits.

Australia bank state branch number:

- three digits
- a hyphen
- three digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression Regex_australia_bank_account_number finds content that matches the pattern.
- A keyword from Keyword_australia_bank_account_number is found.
- The regular expression Regex_australia_bank_account_number_bsb finds content that matches the pattern.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression Regex_australia_bank_account_number finds content that matches the pattern.
- A keyword from Keyword_australia_bank_account_number is found.

XML

```
<!-- Australia Bank Account Number -->
<Entity id="74a54de9-2a30-4aa0-a8aa-3d9327fc07c7" patternsProximity="300"
recommendedConfidence="75">
  <Pattern confidenceLevel="85">
        <IdMatch idRef="Regex_australia_bank_account_number" />
        <Match idRef="Keyword_australia_bank_account_number" />
        <Match idRef="Regex_australia_bank_account_number_bsb" />
  </Pattern>
  <Pattern confidenceLevel="75">
        <IdMatch idRef="Regex_australia_bank_account_number" />
        <Match idRef="Keyword_australia_bank_account_number" />
  </Pattern>
 </Entity>
```

Keywords

Keyword_australia_bank_account_number

- swift bank code
- correspondent bank
- base currency
- usa account
- holder address
- bank address
- information account
- fund transfers
- bank charges
- bank details
- banking information
- full names
- iaea

Australia business number

Article • 08/19/2024

Format

11 digits with optional delimiters

Pattern

11 digits with optional delimiters:

- two digits
- an optional hyphen or space
- three digits
- an optional hyphen or space
- three digits
- an optional hyphen or space
- three digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function Func_australian_business_number finds content that matches the pattern.
- A keyword from Keywords_australian_business_number is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function Func_australian_business_number finds content that matches the pattern.

XML

```
      <!-- Australia Business Number -->
      <Entity id="76e83b3b-01ee-4530-aced-e667a6609f49"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_australian_business_number" />
          <Match idRef="Keywords_australian_business_number" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_australian_business_number" />
        </Pattern>
      </Entity>
```

Keywords

Keyword_australia_business_number

- australia business no
- business number
- abn#
- businessid#
- business ID
- abn
- businessno#

Australia company number

Article • 08/19/2024

Format

nine digits with delimiters

Pattern

nine digits with delimiters:

- three digits
- a space
- three digits
- a space
- three digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function Func_Australian_Company_Number finds content that matches the pattern.
- A keyword from Keyword_Australian_Company_Number is found.

A DLP policy has low confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function Func_Australian_Company_Number finds content that matches the pattern.

XML

```
      <!-- Australia Company Number -->
      <Entity id="b1fba4f7-7b3e-4bb9-8f9a-9366df604dbb"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_Australian_Company_Number" />
          <Match idRef="Keyword_Australian_Company_Number" />
        </Pattern>
        <Pattern confidenceLevel="65">
          <IdMatch idRef="Func_Australian_Company_Number" />
        </Pattern>
      </Entity>
```

Keywords

Keyword_australia_company_number

- acn
- australia company no
- australia company no#
- australia company number
- australian company no
- australian company no#
- australian company number

Australia drivers license number

Article • 02/20/2025

Format

Supports format of all states and territories with either six to 10 digits or six alphanumeric characters

Pattern

six to 10 digits and six alphanumeric characters

- two digits
- optional space or hyphen
- two digits
- optional space or hyphen
- four digits

OR

- three digits
- optional space or hyphen
- three digits
- optional space or hyphen
- three or four digits

OR

- one digit
- optional space or hyphen
- three digits
- optional space or hyphen
- three digits
- optional space or hyphen
- three digits

OR

- seven digits

OR

- one letter
- five digits

OR

- two letters
- four digits

OR

- four digits
- two letters

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A Data Loss Prevention policy has medium confidence that detects this type of sensitive information if, within a proximity of 300 characters:

- The regular expression Regex_australia_drivers_license_number finds content that matches the pattern.
- A keyword from Keyword_australia_drivers_license_number is found.

XML

```
<!-- Australia Drivers License Number -->
<Entity id="1cbbc8f5-9216-4392-9eb5-5ac2298d1356" patternsProximity="300"
recommendedConfidence="75">
   <Pattern confidenceLevel="75">
        <IdMatch idRef="Regex_australia_drivers_license_number" />
        <Match idRef="Keyword_australia_drivers_license_number" />
    </Pattern>
 </Entity>
```

Keywords

Keyword_australia_drivers_license_number

- international driving permits
- australian automobile association
- international driving permit
- DriverLicence
- DriverLicences
- Driver Lic
- Driver Licence
- Driver Licences
- DriversLic
- DriversLicence
- DriversLicences
- Drivers Lic
- Drivers Lics
- Drivers Licence
- Drivers Licences
- Driver'Lic
- Driver'Lics
- Driver'Licence
- Driver'Licences
- Driver' Lic
- Driver' Lics
- Driver' Licence
- Driver' Licences
- Driver'sLic
- Driver'sLics
- Driver'sLicence
- Driver'sLicences
- Driver's Lic
- Driver's Lics
- Driver's Licence
- Driver's Licences
- DriverLic#
- DriverLics#
- DriverLicence#
- DriverLicences#
- Driver Lic#

- Driver Lics#
- Driver Licence#
- Driver Licences#
- DriversLic#
- DriversLics#
- DriversLicence#
- DriversLicences#
- Drivers Lic#
- Drivers Lics#
- Drivers Licence#
- Drivers Licences#
- Driver'Lic#
- Driver'Lics#
- Driver'Licence#
- Driver'Licences#
- Driver' Lic#
- Driver' Lics#
- Driver' Licence#
- Driver' Licences#
- Driver'sLic#
- Driver'sLics#
- Driver'sLicence#
- Driver'sLicences#
- Driver's Lic#
- Driver's Lics#
- Driver's Licence#
- Driver's Licences#

Australia medical account number

Article • 08/19/2024

Format

10-11 digits

Pattern

10-11 digits:

- First digit is in the range 2-6
- Ninth digit is a check digit
- Tenth digit is the issue digit
- 11th digit (optional) is the individual number

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function Func_australian_medical_account_number finds content that matches the pattern.
- A keyword from Keyword_Australia_Medical_Account_Number is found.
- The checksum passes.

```
XML
```

```
  <!-- Australia Medical Account Number -->
<Entity id="104a99a0-3d3b-4542-a40d-ab0b9e1efe63" recommendedConfidence="85"
patternsProximity="300">
    <Pattern confidenceLevel="85">
     <IdMatch idRef="Func_australian_medical_account_number"/>
     <Match idRef="Keyword_Australia_Medical_Account_Number"/>
    </Pattern>
</Entity>
```

Keywords

Keyword_Australia_Medical_Account_Number

- bank account details
- medicare payments
- mortgage account
- bank payments
- information branch
- credit card loan
- department of human services
- local service
- medicare

Australia passport number

Article • 08/19/2024

Format

eight or nine alphanumeric characters

Pattern

- one letter (N, E, D, F, A, C, U, X) followed by seven digits or
- Two letters (PA, PB, PC, PD, PE, PF, PU, PW, PX, PZ) followed by seven digits.

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_australia_passport_number` finds content that matches the pattern.
- A keyword from `Keyword_australia_passport_number` is found.

A DLP policy has low confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_australia_passport_number` finds content that matches the pattern.

XML

```xml
    <!-- Australia Passport Number -->
    <Entity id="29869db6-602d-4853-ab93-3484f905df50"
patternsProximity="300" recommendedConfidence="75" relaxProximity="true">
      <Pattern confidenceLevel="75">
        <IdMatch idRef="Regex_australia_passport_number" />
        <Match idRef="Keyword_australia_passport_number" />
      </Pattern>
      <Pattern confidenceLevel="65">
        <IdMatch idRef="Regex_australia_passport_number" />
      </Pattern>
    </Entity>
```

Keywords

Keyword_australia_passport_number

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers
- passport details
- immigration and citizenship
- commonwealth of australia
- department of immigration
- national identity card
- travel document
- issuing authority

Australia physical addresses

Article • 08/19/2024

Unbundled named entity, detects patterns related to physical address from Australia. It's also included in the All Physical Addresses bundled named entity SIT.

Confidence level

medium

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Australia tax file number

Format

Eight to nine digits

Pattern

Eight to nine digits typically presented with spaces as follows:

- three digits
- an optional space
- three digits
- an optional space
- two to three digits where the last digit is a check digit

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it detects this type of sensitive information if, within a proximity of 300 characters:

- The function Func_australian_tax_file_number finds content that matches the pattern.
- A keyword from Keyword_Australia_Tax_File_Number is found.
- The checksum passes.

```
XML
```

```
   <!-- Australia Tax File Number -->
    <Entity id="e29bc95f-ff70-4a37-aa01-04d17360a4c5" patternsProximity="300"
recommendedConfidence="85">
      <Pattern confidenceLevel="85">
        <IdMatch idRef="Func_australian_tax_file_number" />
        <Match idRef="Keyword_Australia_Tax_File_Number" />
      </Pattern>
    </Entity>
```

Keywords

Keyword_australia_tax_file_number

- australian business number
- marginal tax rate
- medicare levy
- portfolio number
- service veterans
- withholding tax
- individual tax return
- tax file number
- tfn

Last updated on 06/22/2026

Austria drivers license number

Article • 08/19/2024

Format

eight digits without spaces and delimiters

Pattern

eight digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_austria_eu_driver's_license_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_driver's_license_number` or `Keywords_austria_eu_driver's_license_number` is found.

XML

```
      <!-- Austria Driver's License Number -->
      <Entity id="682f18ce-44eb-482b-8198-2bcb96a0761e"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_austria_eu_driver's_license_number" />
```

```
          <Any minMatches="1">
            <Match idRef="Keywords_eu_driver's_license_number" />
            <Match idRef="Keywords_austria_eu_driver's_license_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences

- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#

- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses

- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_austria_eu_driver's_license_number

- fuhrerschein
- führerschein
- Führerscheine
- Führerscheinnummer
- Führerscheinnummern

Austria identity card

Article • 08/19/2024

Format

A 24-character combination of letters, digits, and special characters

Pattern

24 characters:

- 22 letters (not case-sensitive), digits, backslashes, forward slashes, or plus signs
- two letters (not case-sensitive), digits, backslashes, forward slashes, plus signs, or equal signs

Checksum

Not applicable

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_austria_eu_national_id_card` finds content that matches the pattern.
- A keyword from `Keywords_austria_eu_national_id_card` is found.

XML

```
      <!-- Austria Identity Card -->
      <Entity id="5ec06c3b-007e-4820-8343-7ff73b889735"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_austria_eu_national_id_card" />
          <Match idRef="Keywords_austria_eu_national_id_card" />
        </Pattern>
      </Entity>
```

Keywords

Keywords_austria_eu_national_id_card

- identity number
- national id
- personalausweis republik österreich

Austria passport number

Article • 08/19/2024

Format

One letter followed by an optional space and seven digits

Pattern

A combination of one letter, seven digits, and one space:

- one letter (not case-sensitive)
- one space (optional)
- seven digits

Checksum

not applicable

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_austria_eu_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_austria_eu_passport_number` is found.
- The regular expression `Regex_eu_passport_date1` finds date in the format DD.MM.YYYY or a keyword from `Keywords_eu_passport_date` is found

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_austria_eu_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_austria_eu_passport_number` is found.

XML

```
      <!-- Austria Passport Number -->
      <Entity id="1c96ae4e-303b-447d-86c7-77113ac266bf"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Regex_austria_eu_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_austria_eu_passport_number" />
          </Any>
          <Any minMatches="1">
            <Match idRef="Regex_eu_passport_date1" />
            <Match idRef="Keywords_eu_passport_date" />
          </Any>
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_austria_eu_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_austria_eu_passport_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_passport_number

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number

- passportnumbers
- passport numbers

Keywords_austria_eu_passport_number

- reisepassnummer
- reisepasse
- No-Reisepass
- Nr-Reisepass
- Reisepass-Nr
- Passnummer
- reisepässe

Keywords_eu_passport_date

- date of issue
- date of expiry

Austria physical addresses

Article • 08/19/2024

This unbundled named entity detects patterns related to physical address from Austria. It's also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Austria social security number

Article • 01/27/2025

Format

10 digits in the specified format

Pattern

10 digits:

- three digits that correspond to a serial number
- one check digit
- six digits that correspond to the birth date (DDMMYY with valid MM value of 01 to 12)

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_austria_eu_ssn_or_equivalent` finds content that matches the pattern.
- a keyword from `Keywords_austria_eu_ssn_or_equivalent` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_austria_eu_ssn_or_equivalent` finds content that matches the pattern.

XML

```
      <!-- Austria Social Security Number -->
      <Entity id="6896a906-86c9-4d19-a2da-6e43ccd19b7b"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_austria_eu_ssn_or_equivalent" />
          <Match idRef="Keywords_austria_eu_ssn_or_equivalent" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_austria_eu_ssn_or_equivalent" />
          <Any minMatches="0" maxMatches="0">
            <Match idRef="Keywords_austria_eu_telephone_number" />
            <Match idRef="Keywords_austria_eu_mobile_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_austria_eu_ssn_or_equivalent

- austrian ssn
- ehic number
- ehic no
- insurance code
- insurancecode#
- insurance number
- insurance no
- krankenkassennummer
- krankenversicherung
- socialsecurityno
- socialsecurityno#
- social security no
- social security number
- social security code
- sozialversicherungsnummer
- sozialversicherungsnummer#
- soziale sicherheit kein
- sozialesicherheitkein#

- ssn#
- ssn
- versicherungscode
- versicherungsnummer
- zdravstveno zavarovanje

Austria tax identification number

Article • 08/19/2024

Format

nine digits with optional hyphen and forward slash

Pattern

nine digits with optional hyphen and forward slash:

- two digits
- a hyphen (optional)
- three digits
- a forward slash (optional)
- four digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_austria_eu_tax_file_number` finds content that matches the pattern.
- A keyword from `Keywords_austria_eu_tax_file_number` is found.

A DLP policy has low confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_austria_eu_tax_file_number` finds content that matches the pattern.

XML

```
      <!-- Austria Tax Identification Number -->
      <Entity id="4fd58d22-af28-4451-b18a-6f722430a56d"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_austria_eu_tax_file_number" />
          <Match idRef="Keywords_austria_eu_tax_file_number" />
        </Pattern>
        <Pattern confidenceLevel="65">
          <IdMatch idRef="Func_austria_eu_tax_file_number" />
        </Pattern>
      </Entity>
```

Keywords

Keywords_austria_eu_tax_file_number

- österreich
- st.nr.
- steuernummer
- tax ID
- tax identification no
- tax identification number
- tax no#
- tax no
- tax number
- tax registration number
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin ID
- tin no
- tin#

- tax number

Austria value added tax

Article • 08/19/2024

Format

11-character alphanumeric pattern

Pattern

11-character alphanumeric pattern:

- A or a
- T or t
- Optional space
- U or u
- optional space
- two or three digits
- optional space
- four digits
- optional space
- one or two digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function Func_Austria_Value_Added_Tax finds content that matches the pattern.
- A keyword from Keyword_Austria_Value_Added_Tax is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function Func_Austria_Value_Added_Tax finds content that matches the pattern.

XML

```
      <!-- Austria Value Added Tax -->
      <Entity id="b6a3eda2-c56c-4b69-a5f7-dca34db00f48"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_Austria_Value_Added_Tax" />
          <Match idRef="Keyword_Austria_Value_Added_Tax" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_Austria_Value_Added_Tax" />
        </Pattern>
      </Entity>
```

Keywords

Keyword_austria_value_added_tax

- vat number
- vat#
- austrian vat number
- vat no.
- vatno#
- value added tax number
- austrian vat
- mwst
- umsatzsteuernummer
- mwstnummer
- ust.-identifikationsnummer
- umsatzsteuer-identifikationsnummer
- vat identification number
- atu number
- uid number

Azure App Service deployment password

This SIT is also included in the All credentials bundled SIT.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a credential scanning SIT in an endpoint DLP policy. This requirement is specific to the combination of credential scanning SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

Format

A combination of 60 characters consisting of letters, digits, and special characters.

Pattern

Any combination of 60 characters consisting of:

- a-z (not case-sensitive)
- 0-9
- forward slashes (/)
- or plus signs (+)

format example:

```
Ee5Ff~6Gg7.-Hh8Ii9Jj0Kk1Ll2Mm3_Nn4Oo5Pp6
```

Credential format example

```
userPWD=Ff6Gg~7Hh8.-Ii9Jj0Kk1Ll2Mm3Nn4_Oo5Pp6Qq7
```

Checksum

No

SITs that have checksums use a unique calculation to check if the information is valid. This means when the **Checksum** value is **Yes**, the service can make a positive detection based on the sensitive data alone. When the **Checksum** value is **No** additional (secondary) elements must also be detected for the service to make a positive detection.

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

This SIT is designed to match the security information that's used to secure Azure App Service deployment from a local computer.

It uses several primary resources:

- Patterns of Base64 encoded 360-bits symmetric key.
- Patterns of CredentialName, CredentialFeatures, AccountIdentityName, AccountIdentityValue, ResourceType, ResourceName.
- Patterns of mockup values, redactions, and placeholders.
- A dictionary of vocabulary.

The patterns are designed to match actual credentials with reasonable confidence. The patterns don't match credentials formatted as examples. Mockup values, redacted values, and placeholders, like credential type or usage descriptions, in the position where an actual secret value should present won't be matched.

Keywords

Keyword_SymmetricKey360:

- password
- pw

Last updated on 06/01/2026

Azure Batch Shared Access Key

This SIT is also included in the All credentials bundled SIT.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a credential scanning SIT in an endpoint DLP policy. This requirement is specific to the combination of credential scanning SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

Format

A combination of 43 characters consisting of letters, digits, and special characters ending in an equals (=) sign that isn't part of the pattern.

Pattern

A combination of 43 characters consisting of:

- a-z (not case-sensitive)
- 0-9
- forward slashes (/)
- or plus signs (+)

format example:

```
A1bC2dE3fH4iJ5kL6mN7oP8qR9sT0u/+=
```

Credential format example

```
Account=account.batch.azure.net;AccountKey=A1bC2dE3fH4iJ5kL6mN7oP8qR9sT0u/+
```

Checksum

No

SITs that have checksums use a unique calculation to check if the information is valid. This means when the **Checksum** value is **Yes**, the service can make a positive detection based on the sensitive data alone. When the **Checksum** value is **No** additional (secondary) elements must also be detected for the service to make a positive detection.

Definition

This SIT is designed to match the security information that's used to access Azure Batch accounts.

It uses several primary resources:

- Patterns of Base64 encoded 256-bits symmetric key.
- Patterns of CredentialName, CredentialFeatures, AccountIdentityName, AccountIdentityValue, ResourceType, ResourceName, ID.
- Patterns of mockup values, redactions, and placeholders.
- A dictionary of vocabulary.

The patterns are designed to match actual credentials with reasonable confidence. The patterns don't match credentials formatted as examples. Mockup values, redacted values, and placeholders, like credential type or usage descriptions, in the position where an actual secret value should present won't be matched.

Keywords

Keyword_SymmetricKey256:

- SharedAccessKey
- AccountKey

Last updated on 11/18/2025

Azure Bot Framework secret key

This SIT is also included in the All credentials bundled SIT.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a credential scanning SIT in an endpoint DLP policy. This requirement is specific to the combination of credential scanning SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

Format

A combination of 55 characters consisting of letters, digits, and special characters.

or

A combination of 63 characters consisting of letters, digits, and special characters.

Pattern

A combination of 55 characters:

- a-z (not case-sensitive)
- 0-9
- underlines (_)
- or dots (.)

for the 63 characters

A combination of 11 characters:

- a-z (not case-sensitive)
- 0-9
- dashes (-)
- or underlines (_)
- a dot

A combination of three characters:

- a-z (not case-sensitive)
- 0-9
- dashes (-)
- or underlines (_)
- a dot

A combination of three characters:

- a-z (not case-sensitive)
- 0-9
- dashes (-)
- or underlines (_)
- a dot

A combination of 43 characters

- a-z (not case-sensitive)
- 0-9
- dashes (-)
- or underlines (_)

format example:

`Gg7Hh~8Ii9.-Jj0Kk1Ll2Mm3Nn4Oo5_Pp6Qq7Rr8`

Credential example

`host: webchat.botframework.com/?s=Gg7Hh~8Ii9.-Jj0Kk1Ll2Mm3Nn4Oo5_Pp6Qq7Rr`

Checksum

No

SITs that have checksums use a unique calculation to check if the information is valid. This means when the **Checksum** value is **Yes**, the service can make a positive detection based on the sensitive data alone. When the **Checksum** value is **No** additional (secondary) elements must also be detected for the service to make a positive detection.

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

This SIT is designed to match the security information that's used to connect to WebChat channels from Azure Bot services.

It uses several primary resources:

- Patterns of Base64 URL encoded 328-bits symmetric key.
- Patterns of Base64 URL encoded 360-bits symmetric key.
- Patterns of CredentialName, CredentialFeatures, AccountIdentityName, AccountIdentityValue, ResourceType, ResourceName.
- Patterns of mockup values, redactions, and placeholders.
- A dictionary of vocabulary.

The patterns are designed to match actual credentials with reasonable confidence. The patterns don't match credentials formatted as examples. Mockup values, redacted values, and placeholders, like credential type or usage descriptions, in the position where an actual secret value should present won't be matched.

Keywords

Keyword_SymmetricKey328Url:

- botframework
- key

Keyword_SymmetricKey360Url:

- botframework
- key

Last updated on 11/18/2025

Azure Bot service app secret

This SIT is also included in the All credentials bundled SIT.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a credential scanning SIT in an endpoint DLP policy. This requirement is specific to the combination of credential scanning SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

Format

A combination of up to 40 characters consisting of letters, digits, and special characters.

Pattern

A combination of up to 40 characters consisting of:

- a-z (not case-sensitive)
- 0-9
- dashes (-)
- underlines (_)
- dots (.)
- or tilde accents (~)

for example:

`Dd4Ee~5Ff6.-Gg7Hh8Ii9Jj0Kk1Ll2_Mm3Nn4Oo5`

Credential example

`"account.azurewebsites.net/api/messages;AppId=00001111-aaaa-2222-bbbb-3333cccc4444;AppSecret="Dd4Ee~5Ff6.-Gg7Hh8Ii9Jj0Kk1Ll2_Mm3Nn4Oo5"`

Checksum

Yes

SITs that have checksums use a unique calculation to check if the information is valid. This means when the **Checksum** value is **Yes**, the service can make a positive detection based on the sensitive data alone. When the **Checksum** value is **No** additional (secondary) elements must also be detected for the service to make a positive detection.

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

This SIT is designed to match the security information that's used to establish secure communication between an Azure Bot, WebChat channels and client applications.

It uses several primary resources:

- Patterns of Client Secret with specific format.
- Patterns of CredentialName, CredentialFeatures, AccountIdentityName, AccountIdentityValue, ResourceType, ResourceName.
- Patterns of mockup values, redactions, and placeholders.
- A dictionary of vocabulary.

The patterns are designed to match actual credentials with reasonable confidence. The patterns don't match credentials formatted as examples. Mockup values, redacted values, and placeholders, like credential type or usage descriptions, in the position where an actual secret value should present won't be matched.

Keywords

Keyword_AppSecret:

- secret
- password
- key

Last updated on 06/19/2026

Azure Cognitive Search API key

This SIT is also included in the All credentials bundled SIT.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a credential scanning SIT in an endpoint DLP policy. This requirement is specific to the combination of credential scanning SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

Format

A combination of 32 characters consisting of letters and digits.

Pattern

A combination of 32 characters consisting of:

- a-f or A-F (case-sensitive)
- or 0-9

for example:

```
A1bC2dE3fH4iJ5kL6mN7oP8qR9sT0u
```

Credential example

```
host: account.search.windows.net; apikey: A1bC2dE3fH4iJ5kL6mN7oP8qR9sT0u
```

Checksum

No

SITs that have checksums use a unique calculation to check if the information is valid. This means when the **Checksum** value is **Yes**, the service can make a positive detection based on the sensitive

data alone. When the **Checksum** value is **No** additional (secondary) elements must also be detected for the service to make a positive detection.

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

This SIT is designed to match the security information that's used to authenticate inbound requests to Azure Cognitive Search APIs.

It uses several primary resources:

Patterns of Hex encoded 128-bits symmetric key. Patterns of CredentialName, CredentialFeatures, AccountIdentityName, AccountIdentityValue, ResourceType, ResourceName. Patterns of mockup values, redactions, and placeholders. A dictionary of vocabulary.

The patterns are designed to match actual credentials with reasonable confidence. The patterns don't match credentials formatted as examples. Mockup values, redacted values, and placeholders, like credential type or usage descriptions, in the position where an actual secret value should present won't be matched.

Keywords

Keyword_SymmetricKey128Hex:

- dapi
- key
- secret
- token
- password
- pw

Last updated on 11/18/2025

Azure Cognitive Service key

This SIT is also included in the All credentials bundled SIT.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a credential scanning SIT in an endpoint DLP policy. This requirement is specific to the combination of credential scanning SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

Format

A combination of 32 characters consisting of letters and digits.

Pattern

Any combination of 32 characters consisting of:

- a-f or A-F (case-sensitive)
- or 0-9

for example:

`A1bC2dE3fH4iJ5kL6mN7oP8qR9sT0u`

Credential example

`cognitiveservices.azure.com...apikey= A1bC2dE3fH4iJ5kL6mN7oP8qR9sT0u`

Checksum

No

SITs that have checksums use a unique calculation to check if the information is valid. This means when the **Checksum** value is **Yes**, the service can make a positive detection based on the sensitive

data alone. When the **Checksum** value is **No** additional (secondary) elements must also be detected for the service to make a positive detection.

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

This SIT is designed to match the security information that's used to authenticate requests to Azure Cognitive Services.

It uses several primary resources:

- Patterns of Hex encoded 128-bits symmetric key.
- Patterns of CredentialName, CredentialFeatures, AccountIdentityName, AccountIdentityValue, ResourceType, ResourceName.
- Patterns of mockup values, redactions, and placeholders.
- A dictionary of vocabulary.

The patterns are designed to match actual credentials with reasonable confidence. The patterns don't match credentials formatted as examples. Mockup values, redacted values, and placeholders, like credential type or usage descriptions, in the position where an actual secret value should present won't be matched. Keywords

Keywords

Keyword_SymmetricKey128Hex:

- dapi
- key
- secret
- token
- password
- pw

Last updated on 11/18/2025

Azure Container Registry access key

This SIT is also included in the All credentials bundled SIT.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a credential scanning SIT in an endpoint DLP policy. This requirement is specific to the combination of credential scanning SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

Format

A combination of 32 characters consisting of letters, digits, and special characters.

Pattern

A combination of 32 characters consisting of:

- a-z (not case-sensitive)
- 0-9
- forward slashes (/)
- or plus signs (+)

Credential example

```
account.azurecr.io/ #docker password: A1bC2dE3fH4iJ5kL6mN7oP8qR9sT0u/+
```

Checksum

No

SITs that have checksums use a unique calculation to check if the information is valid. This means when the **Checksum** value is **Yes**, the service can make a positive detection based on the sensitive data alone. When the **Checksum** value is **No** additional (secondary) elements must also be detected for the service to make a positive detection.

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

This SIT is designed to match the security information that's used to access Azure Container Registry services as an admin account.

It uses several primary resources:

- Patterns of Base64 encoded 192-bits symmetric key.
- Patterns of CredentialName, CredentialFeatures, AccountIdentityName, AccountIdentityValue, ResourceType, ResourceName.
- Patterns of mockup values, redactions, and placeholders.
- A dictionary of vocabulary.

The patterns are designed to match actual credentials with reasonable confidence. The patterns don't match credentials formatted as examples. Mockup values, redacted values, and placeholders, like credential type or usage descriptions, in the position where an actual secret value should present won't be matched.

Keywords

Keyword_SymmetricKey192:

- password
- -p
- azurecr

Last updated on 11/18/2025

Azure Cosmos DB account access key

This SIT is also included in the All credentials bundled SIT.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a credential scanning SIT in an endpoint DLP policy. This requirement is specific to the combination of credential scanning SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

Format

A combination of 88 characters consisting of letters, digits, and special characters.

Pattern

Any combination of 86 characters consisting of:

- a-z (not case-sensitive)
- 0-9
- forward slashes (/)
- or plus signs (+)
- ends with two equal signs (=)

format example:

```
C2dE3fH4iJ5kL6mN7oP8qR9sT0uV1w/+C2dE3fH4iJ5kL6mN7oP8qR9sT0uV1w==
```

Credential example

```
AccountEndpoint=https://account.documents.azure.com;AccountKey=C2dE3fH4iJ5kL6mN7oP8qR9sT0uV1w/+C2dE3fH4iJ5kL6mN7oP8qR9sT0uV1w==
```

Checksum

No

SITs that have checksums use a unique calculation to check if the information is valid. This means when the **Checksum** value is **Yes**, the service can make a positive detection based on the sensitive data alone. When the **Checksum** value is **No** additional (secondary) elements must also be detected for the service to make a positive detection.

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

This SIT is designed to match the security information that's used to provide access to administrative resources for Azure Cosmos DB accounts.

It uses several primary resources:

- Patterns of Base64 encoded 512-bits symmetric key.
- Patterns of CredentialName, CredentialFeatures, AccountIdentityName, AccountIdentityValue, ResourceType, ResourceName, ID, AccountName.
- Patterns of mockup values, redactions, and placeholders.
- A dictionary of vocabulary.

The patterns are designed to match actual credentials with reasonable confidence. The patterns don't match credentials formatted as examples. Mockup values, redacted values, and placeholders, like credential type or usage descriptions, in the position where an actual secret value should present won't be matched.

Keywords

Keyword_SymmetricKey512:

- SharedAccessKey
- AccountKey

Last updated on 11/18/2025

Azure Databricks personal access token

This SIT is also included in the All credentials bundled SIT.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a credential scanning SIT in an endpoint DLP policy. This requirement is specific to the combination of credential scanning SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

Format

A combination of 32 characters consisting of letters and digits.

Pattern

A combination of 32 characters consisting of:

- a-f or A-F (case-sensitive)
- or 0-9

format example:

```
A1bC2dE3fH4iJ5kL6mN7oP8qR9sT0u
```

Credential format example

```
account.azuredatabricks.net;PAT=A1bC2dE3fH4iJ5kL6mN7oP8qR9sT0u
```

Checksum

No

SITs that have checksums use a unique calculation to check if the information is valid. This means when the **Checksum** value is **Yes**, the service can make a positive detection based on the sensitive

data alone. When the **Checksum** value is **No** additional (secondary) elements must also be detected for the service to make a positive detection.

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

This SIT is designed to match the security information that's used to authenticate to the Azure Databricks REST API.

It uses several primary resources:

- Patterns of Hex encoded 128-bits symmetric key.
- Patterns of CredentialName, CredentialFeatures, AccountIdentityName, AccountIdentityValue, ResourceType, ResourceName.
- Patterns of mockup values, redactions, and placeholders.
- A dictionary of vocabulary.

The patterns are designed to match actual credentials with reasonable confidence. The patterns don't match credentials formatted as examples. Mockup values, redacted values, and placeholders, like credential type or usage descriptions, in the position where an actual secret value should present won't be matched.

Keywords

Keyword_SymmetricKey128Hex:

dapi key secret token password pw

Last updated on 11/18/2025

Azure DevOps app secret

This SIT is also included in the All credentials bundled SIT.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a credential scanning SIT in an endpoint DLP policy. This requirement is specific to the combination of credential scanning SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

Format

A combination of 52 characters consisting of letters, digits, and special characters.

Pattern

Any combination of 52 characters consisting of:

- a-z or A-Z (case-sensitive)
- or 2-7

format example:

```
A2bC2dE3fH4iJ5kL6mN7oP7qR7sT2u
```

Credential example

```
AdoAppId=...;AdoAppSecret=2bC2dE3fH4iJ5kL6mN7oP7qR7sT2u;
```

> **Important**
>
> This example has been invalidated. It's not a detectable example of this SIT.

Checksum

Yes

SITs that have checksums use a unique calculation to check if the information is valid. This means when the **Checksum** value is **Yes**, the service can make a positive detection based on the sensitive data alone. When the **Checksum** value is **No**, other (secondary) elements must also be detected for the service to make a positive detection.

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

This SIT is designed to match the security information that's used to authenticate web app users for Azure DevOps REST API access.

It uses several primary resources:

- Patterns of Base32 encoded 256-bits symmetric key.
- Patterns of CredentialName, CredentialFeatures, AccountIdentityName, AccountIdentityValue, ResourceType, ResourceName.
- Patterns of mockup values, redactions, and placeholders.
- A dictionary of vocabulary

The patterns are designed to match actual credentials with reasonable confidence. The patterns don't match credentials formatted as examples. Mockup values, redacted values, and placeholders, like credential type or usage descriptions, in the position where an actual secret value should present won't be matched.

Keywords

Keyword_SymmetricKey256B32:

- pat
- token
- ado

- vsts
- azuredevops
- visualstudio.com
- dev.azure.com

Last updated on 11/18/2025

Azure DevOps personal access token

This SIT is also included in the All credentials bundled SIT.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a credential scanning SIT in an endpoint DLP policy. This requirement is specific to the combination of credential scanning SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

Format

A combination of 84 characters consisting of letters, digits, and special characters.

Pattern

Any combination of 84 characters consisting of:

- a-z or A-Z (case-sensitive)
- or 0-9
- with a fixed signature `AZDO` at position 76-80

Credential example

```
URL="org.visualstudio.com/proj"; PAT =
"abcdefghijklmnopqrstuvwxyz012345679ABCDEFGHIJKLMNOPQRSTUVWXYZabcdefghijklmnoAZDOabcd"
```

> **Important**
>
> This example has been invalidated. It's not a detectable example of this SIT.

Checksum

Yes

SITs that have checksums use a unique calculation to check if the information is valid. This means when the **Checksum** value is **Yes**, the service can make a positive detection based on the sensitive data alone. When the **Checksum** value is **No**, other (secondary) elements must also be detected for the service to make a positive detection.

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

This SIT is designed to match the security information that's used as an alternate password to authenticate into Azure DevOps.

It uses several primary resources:

- Patterns of Base32 encoded 256-bits symmetric key.
- Patterns of CredentialName, CredentialFeatures, AccountIdentityName, AccountIdentityValue, ResourceType, ResourceName.
- Patterns of mockup values, redactions, and placeholders.
- A dictionary of vocabulary

The patterns are designed to match actual credentials with reasonable confidence. The patterns don't match credentials formatted as examples. Mockup values, redacted values, and placeholders, like credential type or usage descriptions, in the position where an actual secret value should present won't be matched.

Keywords

Keyword_SymmetricKey256B32:

- pat
- token
- ado
- vsts
- azuredevops

- visualstudio.com
- dev.azure.com

Last updated on 11/18/2025

Azure DocumentDB auth key

Article • 02/07/2025

Format

The string `DocumentDb` followed by the characters and strings outlined in the pattern below.

Pattern

- The string `DocumentDb`
- Any combination of between 3-200 lower- or uppercase letters, digits, symbols, special characters, or spaces
- A greater than symbol (>), an equal sign (=), a quotation mark ("), or an apostrophe (')
- Any combination of 86 lower- or uppercase letters, digits, forward slash (/), or plus sign (+)
- Two equal signs (=)

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `CEP_Regex_AzureDocumentDBAuthKey` finds content that matches the pattern.

- The regular expression `CEP_CommonExampleKeywords` doesn't find content that matches the pattern.

XML

```
<!-- Azure Document DB Auth Key -->
<Entity id="0f587d92-eb28-44a9-bd1c-90f2892b47aa" patternsProximity="300"
recommendedConfidence="85">
  <Pattern confidenceLevel="85">
        <IdMatch idRef="CEP_Regex_AzureDocumentDBAuthKey" />
        <Any minMatches="0" maxMatches="0">
            <Match idRef="CEP_CommonExampleKeywords" />
          </Any>
  </Pattern>
</Entity>
```

Keywords

CEP_CommonExampleKeywords

This sensitive information type identifies these keywords by using a regular expression, not a keyword list.

- contoso
- fabrikam
- northwind
- sandbox
- onebox
- localhost
- 127.0.0.1
- testacs.com
- s-int.net

Azure EventGrid access key

This SIT is also included in the All credentials bundled SIT.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a credential scanning SIT in an endpoint DLP policy. This requirement is specific to the combination of credential scanning SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

Format

A combination of 43 characters consisting of letters, digits, and special characters ending in an equals sign (=) that isn't part of the pattern.

Pattern

A combination of 43 characters consisting of:

- a-z (not case-sensitive)
- 0-9
- forward slashes (/)
- or plus signs (+)

for example:

```
A1bC2dE3fH4iJ5kL6mN7oP8qR9sT0u+
```

Credential example

```
host: account.eventgrid.azure.net; accesskey: A1bC2dE3fH4iJ5kL6mN7oP8qR9sT0u+
```

Checksum

No

SITs that have checksums use a unique calculation to check if the information is valid. This means when the **Checksum** value is **Yes**, the service can make a positive detection based on the sensitive data alone. When the **Checksum** value is **No** additional (secondary) elements must also be detected for the service to make a positive detection.

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

This SIT is designed to match the security information that's used to authenticate an application publishing events to Azure Event Grid resources (topics and domains).

It uses several primary resources:

- Patterns of Base64 encoded 256-bits symmetric key.
- Patterns of CredentialName, CredentialFeatures, AccountIdentityName, AccountIdentityValue, ResourceType, ResourceName, ID.
- Patterns of mockup values, redactions, and placeholders.
- A dictionary of vocabulary.

The patterns are designed to match actual credentials with reasonable confidence. The patterns don't match credentials formatted as examples. Mockup values, redacted values, and placeholders, like credential type or usage descriptions, in the position where an actual secret value should present won't be matched.

Keywords

Keyword_SymmetricKey256:

- SharedAccessKey
- AccountKey

Last updated on 11/18/2025

Azure Function Master / API key

This SIT is also included in the All credentials bundled SIT.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a credential scanning SIT in an endpoint DLP policy. This requirement is specific to the combination of credential scanning SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

Format

A combination of 56 characters consisting of letters, digits, and special characters.

or

A combination of up to 90 characters consisting of letters, digits, and special characters.

Pattern

Any combination of 54 characters consisting of:

- a-z (not case-sensitive)
- 0-9
- forward slashes (/)
- or plus signs (+)
- ends with two equal signs (=)

format example:

```
A1bC2dE3fH4iJ5kL6mN7oP8qR9sT0u/==
```

or

A combination of 54 to 84 characters consisting of:

- a-z (not case-sensitive)

- 0-9
- or percent signs (%)
- ends with a suffix '%3d%3d' (not case-sensitive)

for example:

abcdefghijklmnopqrstuvwxyz0123456789%2F%2BABCDEF0123456789%3D%3D

Credential example

`https://account.azurewebsites.net/api/function?code=A1bC2dE3fH4iJ5kL6mN7oP8qR9sT0u/==`

Checksum

No

SITs that have checksums use a unique calculation to check if the information is valid. This means when the **Checksum** value is **Yes**, the service can make a positive detection based on the sensitive data alone. When the **Checksum** value is **No** additional (secondary) elements must also be detected for the service to make a positive detection.

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

This SIT is designed to match the security information that's used to request Azure Function API when its authorization level is set a value other than anonymous.

It uses several primary resources:

- Patterns of Base64 encoded 320-bits symmetric key.
- Patterns of URL Encoded 320-bits symmetric key.
- Patterns of CredentialName, CredentialFeatures, AccountIdentityName, AccountIdentityValue, ResourceType, ResourceName.
- Patterns of mockup values, redactions, and placeholders.

- A dictionary of vocabulary.

The patterns are designed to match actual credentials with reasonable confidence. The patterns don't match credentials formatted as examples. Mockup values, redacted values, and placeholders, like credential type or usage descriptions, in the position where an actual secret value should present won't be matched.

Keywords

Keyword_SymmetricKey320:

- code=
- key

Keyword_SymmetricKey320UrlEncoded:

- code=
- key

Last updated on 11/18/2025

Azure IAAS database connection string and Azure SQL connection string

Article • 08/19/2024

Format

The string `Server`, `server`, or `data source` followed by the characters and strings outlined in the pattern below, including the string `cloudapp.azure.com` or `cloudapp.azure.net` or `database.windows.net`, and the string `Password` or `password` or `pwd`.

Pattern

- the string `Server`, `server`, or `data source`
- zero to two whitespace characters
- an equal sign (=)
- zero to two whitespace characters
- any combination of between 1-200 lower- or uppercase letters, digits, symbols, special characters, or spaces
- The string "cloudapp.azure.com", "cloudapp.azure.net", or "database.windows.net"
- any combination of between 1-300 lower- or uppercase letters, digits, symbols, special characters, or spaces
- the string `Password`, `password`, or `pwd`
- zero to two whitespace characters
- an equal sign (=)
- zero to two whitespace characters
- one or more characters that aren't a semicolon (;), quotation mark ("), or apostrophe (')
- a semicolon (;), quotation mark ("), or apostrophe (')

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `CEP_Regex_AzureConnectionString` finds content that matches the pattern.
- The regular expression `CEP_CommonExampleKeywords` doesn't find content that matches the pattern.

XML

```
<!--Azure IAAS Database Connection String and Azure SQL Connection String-->
<Entity id="ce1a126d-186f-4700-8c0c-486157b953fd" patternsProximity="300"
recommendedConfidence="85">
  <Pattern confidenceLevel="85">
        <IdMatch idRef="CEP_Regex_AzureConnectionString" />
        <Any minMatches="0" maxMatches="0">
            <Match idRef="CEP_CommonExampleKeywords" />
        </Any>
    </Pattern>
</Entity>
```

Keywords

CEP_common_example_keywords

This sensitive information type identifies these keywords by using a regular expression, not a keyword list.

- contoso
- fabrikam
- northwind
- sandbox
- onebox
- localhost
- 127.0.0.1
- testacs.com

- s-int.net

Azure IoT connection string

This SIT is also included in the All credentials bundled SIT.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a credential scanning SIT in an endpoint DLP policy. This requirement is specific to the combination of credential scanning SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

Format

The string `HostName` followed by the characters and strings outlined in the pattern below, including the strings `azure-devices.net` and `SharedAccessKey`.

Pattern

- the string `HostName`
- zero to two whitespace characters
- an equal sign (=)
- zero to two whitespace characters
- any combination of between 1-200 lower- or uppercase letters, digits, symbols, special characters, or spaces
- the string "azure-devices.net"
- any combination of between 1-200 lower- or uppercase letters, digits, symbols, special characters, or spaces
- the string `SharedAccessKey`
- zero to two whitespace characters
- an equal sign (=)
- zero to two whitespace characters
- any combination of 43 lower- or uppercase letters, digits, forward slash (/), or plus sign (+)
- an equal sign (=)

Credential example

`HostName=account.azure-devices.net;SharedAccessKeyName=key;SharedAccessKey=abcdefghijklmnopqrstuvwxyz0123456789/+ABCDE=`

Checksum

No

SITs that have checksums use a unique calculation to check if the information is valid. This means when the **Checksum** value is **Yes**, the service can make a positive detection based on the sensitive data alone. When the **Checksum** value is **No** additional (secondary) elements must also be detected for the service to make a positive detection.

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `CEP_Regex_AzureIoTConnectionString` finds content that matches the pattern.
- The regular expression `CEP_CommonExampleKeywords` doesn't find content that matches the pattern.

XML

```xml
<!--Azure IoT Connection String-->
<Entity id="0b34bec3-d5d6-4974-b7b0-dcdb5c90c29d" patternsProximity="300"
recommendedConfidence="85">
  <Pattern confidenceLevel="85">
        <IdMatch idRef="CEP_Regex_AzureIoTConnectionString" />
        <Any minMatches="0" maxMatches="0">
            <Match idRef="CEP_CommonExampleKeywords" />
        </Any>
  </Pattern>
</Entity>
```

Keywords

CEP_common_example_keywords

This sensitive information type identifies these keywords by using a regular expression, not a keyword list.

- contoso
- fabrikam
- northwind
- sandbox
- onebox
- localhost
- 127.0.0.1
- testacs.com
- s-int.net

Last updated on 06/01/2026

Azure IoT shared access key

Article • 02/18/2025

Format

A combination of 44-characters consisting of letters, digits, and special characters ending with and equals sign that isn't part of the pattern.

Pattern

Any combination of 43 characters consisting of:

- a-z (not case-sensitive)
- 0-9
- forward slashes (/)
- or plus signs (+)
- ends with an equal sign (=) that isn't part of the pattern.

Format example:

```
A1bC2dE3fH4iJ5kL6mN7oP8qR9sT0u+=
```

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

This SIT is designed to match the security information that's used to authenticate Azure IoT devices and services.

It uses several primary resources:

- Patterns of Base64 encoded 256-bits symmetric key.
- Patterns of CredentialName, CredentialFeatures, AccountIdentityName, AccountIdentityValue, ResourceType, ResourceName, ID.
- Patterns of mockup values, redactions, and placeholders.
- A dictionary of vocabulary.

The patterns are designed to match actual credentials with reasonable confidence. The patterns don't match credentials formatted as examples. Mockup values, redacted values, and placeholders, like credential type or usage descriptions, in the position where an actual secret value should present won't be matched.

Keywords

Keyword_SymmetricKey256:

- SharedAccessKey
- AccountKey

Azure Logic App shared access signature

This SIT is also included in the All credentials bundled SIT.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a credential scanning SIT in an endpoint DLP policy. This requirement is specific to the combination of credential scanning SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

Format

A combination of up to 76 characters consisting of letters, digits, and special characters.

Pattern

Any combination of 43 to 73 characters consisting of:

- a-z (not case-sensitive)
- 0-9
- or percent signs (%)
- ends with a suffix '%3d' (not case-sensitive)

for example:

```
abcdefghijklmnopqrstuvwxyz0123456789%2F%2BABCDE%3D
```

Credential example

```
https://account.logic.azure.com/?...&sig=abcdefghijklmnopqrstuvwxyz0123456789%2F%2BABCDE%3D
```

Checksum

No

SITs that have checksums use a unique calculation to check if the information is valid. This means when the **Checksum** value is **Yes**, the service can make a positive detection based on the sensitive data alone. When the **Checksum** value is **No** additional (secondary) elements must also be detected for the service to make a positive detection. sit-defn-azure-iot-shared-access-key.md

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

This SIT is designed to match the security information that's used to grant access to a request endpoint on Azure Logic Apps.

It uses several primary resources:

Patterns of URL Encoded 256-bits symmetric key. Patterns of CredentialName, CredentialFeatures, AccountIdentityName, AccountIdentityValue, ResourceType, ResourceName. Patterns of mockup values, redactions, and placeholders. A dictionary of vocabulary

The patterns are designed to match actual credentials with reasonable confidence. The patterns don't match credentials formatted as examples. Mockup values, redacted values, and placeholders, like credential type or usage descriptions, in the position where an actual secret value should present won't be matched.

Keywords

Keyword_SymmetricKey256UrlEncoded:

- sig=
- key
- token
- secret
- password

Last updated on 11/18/2025

Azure Machine Learning web service API key

This SIT is also included in the All credentials bundled SIT.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a credential scanning SIT in an endpoint DLP policy. This requirement is specific to the combination of credential scanning SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

Format

A combination of 88 characters consisting of letters, digits, and special characters ending in two equals signs (==).

Pattern

Any combination of 86 characters:

- a-z (not case-sensitive)
- 0-9
- forward slashes (/)
- or plus signs (+)

format example:

```
A1bC2dE3fH4iJ5kL6mN7oP8qR9sT0u/
```

Credential example

```
host: account.azureml.net/services/33334444-dddd-5555-eeee-6666ffff7777/workspaces/33334444-dddd-5555-eeee-6666ffff7777/;
apikey:bC2dE3fH4iJ5kL6mN7oP8qR9sT0u/;
```

Checksum

No

SITs that have checksums use a unique calculation to check if the information is valid. This means when the **Checksum** value is **Yes**, the service can make a positive detection based on the sensitive data alone. When the **Checksum** value is **No** additional (secondary) elements must also be detected for the service to make a positive detection.

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

This SIT is designed to match the security information that's used to connect to Azure Machine Learning Web services.

It uses several primary resources:

- Patterns of Base64 encoded 512-bits symmetric key.
- Patterns of CredentialName, CredentialFeatures, AccountIdentityName, AccountIdentityValue, ResourceType, ResourceName, ID, AccountName.
- Patterns of mockup values, redactions, and placeholders.
- A dictionary of vocabulary.

The patterns are designed to match actual credentials with reasonable confidence. The patterns don't match credentials formatted as examples. Mockup values, redacted values, and placeholders, like credential type or usage descriptions, in the position where an actual secret value should present won't be matched.

Keywords

Keyword_SymmetricKey512:

- SharedAccessKey
- AccountKey

Last updated on 11/18/2025

Azure Maps subscription key

This SIT is also included in the All credentials bundled SIT.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a credential scanning SIT in an endpoint DLP policy. This requirement is specific to the combination of credential scanning SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

Format

A combination of 43 characters consisting of letters, digits, and special characters.

Pattern

Any combination of 43 characters consisting of:

- a-z (not case-sensitive)
- 0-9
- dashes (-)
- or underlines (_)

format example:

```
C2dE3fH4iJ5kL6mN7oP8qR9sT0uV1w-
```

Credential example

```
host: atlas.microsoft.com; key: C2dE3fH4iJ5kL6mN7oP8qR9sT0uV1w-;
```

Checksum

No

SITs that have checksums use a unique calculation to check if the information is valid. This means when the **Checksum** value is **Yes**, the service can make a positive detection based on the sensitive data alone. When the **Checksum** value is **No** additional (secondary) elements must also be detected for the service to make a positive detection.

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

This SIT is designed to match the security information that's used to access resources in Azure Maps accounts.

It uses several primary resources:

- Patterns of Base64 URL encoded 256-bits symmetric key.
- Patterns of CredentialName, CredentialFeatures, AccountIdentityName, AccountIdentityValue, ResourceType, ResourceName.
- Patterns of mockup values, redactions, and placeholders.
- A dictionary of vocabulary.

The patterns are designed to match actual credentials with reasonable confidence. The patterns don't match credentials formatted as examples. Mockup values, redacted values, and placeholders, like credential type or usage descriptions, in the position where an actual secret value should present won't be matched.

Keywords

Keyword_SymmetricKey256Url:

- key
- microsoft.maps

Last updated on 11/18/2025

Azure publish setting password

Article • 02/10/2025

Format

The string `userpwd=` followed by an alphanumeric string.

Pattern

- the string `userpwd=`
- any combination of 60 lowercase letters or digits
- a quotation mark (")

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `CEP_Regex_AzurePublishSettingPasswords` finds content that matches the pattern.
- The regular expression `CEP_CommonExampleKeywords` doesn't find content that matches the pattern.

XML

```
<!--Azure Publish Setting Password-->
<Entity id="75f4cc8a-a68e-49e5-89ce-fa8f03d286a5" patternsProximity="300"
recommendedConfidence="85">
```

```
  <Pattern confidenceLevel="85">
       <IdMatch idRef="CEP_Regex_AzurePublishSettingPasswords" />
       <Any minMatches="0" maxMatches="0">
           <Match idRef="CEP_CommonExampleKeywords" />
       </Any>
  </Pattern>
</Entity>
```

Keywords

CEP_common_example_keywords

This sensitive information type identifies these keywords by using a regular expression, not a keyword list.

- contoso
- fabrikam
- northwind
- sandbox
- onebox
- localhost
- 127.0.0.1
- testacs.com
- s-int.net

Azure Redis cache connection string

Article • 02/10/2025

Format

The string `redis.cache.windows.net` followed by the characters and strings outlined in the pattern below, including the string `password` or `pwd`.

Pattern

- the string `redis.cache.windows.net`
- any combination of between 1-200 lower- or uppercase letters, digits, symbols, special characters, or spaces
- the string `password` or `pwd`
- zero to two whitespace characters
- an equal sign (=)
- zero to two whitespace characters
- any combination of 43 characters that are lower- or uppercase letters, digits, forward slash (/), or plus sign (+)
- an equal sign (=)

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `CEP_Regex_AzureRedisCacheConnectionString` finds content that matches the pattern.
- The regular expression `CEP_CommonExampleKeywords` doesn't find content that matches the pattern.

XML

```
<!--Azure Redis Cache Connection String-->
<Entity id="095a7e6c-efd8-46d5-af7b-5298d53a49fc" patternsProximity="300"
recommendedConfidence="85">
  <Pattern confidenceLevel="85">
        <IdMatch idRef="CEP_Regex_AzureRedisCacheConnectionString" />
        <Any minMatches="0" maxMatches="0">
            <Match idRef="CEP_CommonExampleKeywords" />
        </Any>
  </Pattern>
</Entity>
```

Keywords

CEP_common_example_keywords

This sensitive information type identifies these keywords by using a regular expression, not a keyword list.

- contoso
- fabrikam
- northwind
- sandbox
- onebox
- localhost
- 127.0.0.1
- testacs.com
- s-int.net

Azure Redis cache connection string password

This SIT is also included in the All credentials bundled SIT.

> **Important**
>
> Because credential scanning SITs are a type of advanced classifier, you must enable advanced classification before implementing them. For information about enabling advanced classification, and about configuring endpoints and the supported file types, see Advanced classification scanning and protection.

Format

A combination of up to 20,000-characters consisting of letters, digits, and special characters.

or

A combination of 44-characters consisting of letters, digits, and special characters.

Pattern

Any combination of up to 20,000 characters consisting of:

- a-z (not case-sensitive)
- 0-9
- forward slashes (/)
- or plus signs (+)
- Up to 2 equal signs (=)

format example:

```
aB1cD2eF/3gH4iJ5kL6-mN7oP8qR=
```

or

Any combination of 43 characters consisting of:

- a-z (not case-sensitive)
- 0-9
- forward slashes (/)
- or plus signs (+)
- ends with an equal sign (=)

format example:

`iJ5kL6mN7oP8qR9sT0+uV1wX2yZ==`

Credential example

`HostName=account.redis.cache.windows.net;Password=****************************

Checksum

Yes

SITs that have checksums use a unique calculation to check if the information is valid. This means when the **Checksum** value is **Yes**, the service can make a positive detection based on the sensitive data alone. When the **Checksum** value is **No** additional (secondary) elements must also be detected for the service to make a positive detection.

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

This SIT is designed to match the security information that's used to connect to Azure Cache for Redis servers.

It uses several primary resources:

- Patterns of Base64 encoded string literal.
- Patterns of Base64 encoded 256-bits symmetric key.

- Patterns of CredentialName, CredentialFeatures, AccountIdentityName, AccountIdentityValue, ResourceType, ResourceName, ID.
- Patterns of mockup values, redactions, and placeholders.
- A dictionary of vocabulary.

The patterns are designed to match actual credentials with reasonable confidence. The patterns don't match credentials formatted as examples. Mockup values, redacted values, and placeholders, like credential type or usage descriptions, in the position where an actual secret value should present won't be matched.

Keywords

Keyword_Base64EncodedStringLiteral:

- MII

Keyword_SymmetricKey256:

- SharedAccessKey
- AccountKey

Last updated on 04/09/2026

Azure SAS

Article • 02/10/2025

Format

The string `sig` followed by the characters and strings outlined in the pattern below.

Pattern

- the string `sig`
- zero to two whitespace characters
- an equal sign (=)
- zero to two whitespace characters
- any combination of between 43-53 characters that are lower- or uppercase letters, digits, or the percent sign (%)
- the string `%3d`
- any character that isn't a lower- or uppercase letter, digit, or percent sign (%)

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `CEP_Regex_AzureSAS` finds content that matches the pattern.

```
XML
```

```
<!--Azure SAS-->
<Entity id="4d235014-e564-47f4-a6fb-6ebb4a826834" patternsProximity="300"
recommendedConfidence="85">
  <Pattern confidenceLevel="85">
        <IdMatch idRef="CEP_Regex_AzureSAS" />
  </Pattern>
</Entity>
```

Azure service bus connection string

Article • 02/10/2025

Format

The string `EndPoint` followed by the characters and strings outlined in the pattern below, including the strings `servicebus.windows.net` and `SharedAccesKey`.

Pattern

- the string `EndPoint`
- zero to two whitespace characters
- an equal sign (=)
- zero to two whitespace characters
- any combination of between 1-200 lower- or uppercase letters, digits, symbols, special characters, or spaces
- the string `servicebus.windows.net`
- any combination of between 1-200 lower- or uppercase letters, digits, symbols, special characters, or spaces
- the string `SharedAccessKey`
- zero to two whitespace characters
- an equal sign (=)
- zero to two whitespace characters
- any combination of 43 characters that are lower- or uppercase letters, digits, forward slash (/), or plus sign (+)
- an equal sign (=)

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `CEP_Regex_AzureServiceBusConnectionString` finds content that matches the pattern.
- The regular expression `CEP_CommonExampleKeywords` doesn't find content that matches the pattern.

XML

```
<!--Azure Service Bus Connection String-->
<Entity id="b9a6578f-a83f-4fcd-bf44-2130bae49a6f" patternsProximity="300"
recommendedConfidence="85">
  <Pattern confidenceLevel="85">
        <IdMatch idRef="CEP_Regex_AzureServiceBusConnectionString" />
        <Any minMatches="0" maxMatches="0">
            <Match idRef="CEP_CommonExampleKeywords" />
        </Any>
  </Pattern>
</Entity>
```

Keywords

CEP_common_example_keywords

This sensitive information type identifies these keywords by using a regular expression, not a keyword list.

- contoso
- fabrikam
- northwind
- sandbox
- onebox
- localhost
- 127.0.0.1
- testacs.com
- s-int.net

Azure service bus shared access signature

This SIT is also included in the All credentials bundled SIT.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a credential scanning SIT in an endpoint DLP policy. This requirement is specific to the combination of credential scanning SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

Format

A combination of 44-characters consisting of letters, digits, and special characters ending with an equals sign (=) that isn't part of the pattern.

or

A combination of up to 76-characters consisting of letters, digits, and special characters ending with an equals sign (=) that isn't part of the pattern.

Pattern

Any combination of 43 characters consisting of:

- a-z (not case-sensitive)
- 0-9
- forward slashes (/)
- or plus signs (+)
- ends with an equal sign (=) that isn't part of the pattern

format example:

```
kL6mN7oP8qR9sT0uV1+wX2yZ3aB=
```

or

Any combination of 43 to 73 characters consisting of:

- a-z (not case-sensitive)
- 0-9
- or percent signs (%)
- ends with a suffix '%3d' (not case-sensitive)

format example:

`mN7oP8qR9sT0uV1wX2%yZ3aB4dE%3D`

Credential example

`Endpoint=sb://account.servicebus.windows.net;SharedAccessKey=mN7oP8qR9sT0uV1wX2%yZ3aB4dE%3D`

Checksum

No

SITs that have checksums use a unique calculation to check if the information is valid. This means when the **Checksum** value is **Yes**, the service can make a positive detection based on the sensitive data alone. When the **Checksum** value is **No** additional (secondary) elements must also be detected for the service to make a positive detection.

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

This SIT is designed to match the security information that's used to grant a user access to Azure Service Bus resources with specific rights.

It uses several primary resources:

- Patterns of Base64 encoded 256-bits symmetric key.
- Patterns of URL Encoded 256-bits symmetric key.
- Patterns of CredentialName, CredentialFeatures, AccountIdentityName, AccountIdentityValue, ResourceType, ResourceName, ID.

- Patterns of mockup values, redactions, and placeholders.
- A dictionary of vocabulary.

The patterns are designed to match actual credentials with reasonable confidence. The patterns don't match credentials formatted as examples. Mockup values, redacted values, and placeholders, like credential type or usage descriptions, in the position where an actual secret value should present won't be matched.

Keywords

Keyword_SymmetricKey256:

- SharedAccessKey
- AccountKey

Keyword_SymmetricKey256UrlEncoded:

- sig=
- key
- token
- secret
- password

Last updated on 11/18/2025

Azure Shared Access key / Web Hook token

This SIT is also included in the All credentials bundled SIT.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a credential scanning SIT in an endpoint DLP policy. This requirement is specific to the combination of credential scanning SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

Format

A combination of 44 characters consisting of letters, digits, and special characters.

or

A combination of up to 76 characters consisting of letters, digits, and special characters.

Pattern

Any combination of 43 characters consisting of:

- a-z (not case-sensitive)
- 0-9
- forward slashes (/)
- or plus signs (+)
- ends with an equal sign (=)

format example:

```
gH4iJ5kL6mN7oP8qR9/sT0uV1wX=
```

or

A combination of 43 to 73 characters consisting of:

- a-z (case insensitive)

- 0-9
- percent signs (%)
- ends with a suffix '%3d' (case insensitive)

for example:

```
abcdefghijklmnopqrstuvwxyz0123456789%2F%2BABCDE%3D
```

Credential format example

```
PrimaryKey=mN7oP8qR9sT0uV1wX2%yZ3aB4dE%3d;
```

Checksum

No

SITs that have checksums use a unique calculation to check if the information is valid. This means when the **Checksum** value is **Yes**, the service can make a positive detection based on the sensitive data alone. When the **Checksum** value is **No** additional (secondary) elements must also be detected for the service to make a positive detection.

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

This SIT is designed to match the security information that's used to access general Azure resources with restricted permission.

It uses several primary resources:

- Patterns of Base64 encoded 256-bits symmetric key.
- Patterns of URL Encoded 256-bits symmetric key.
- Patterns of CredentialName, CredentialFeatures, AccountIdentityName, AccountIdentityValue, ResourceType, ResourceName, ID.
- Patterns of mockup values, redactions, and placeholders.

- A dictionary of vocabulary

The patterns are designed to match actual credentials with reasonable confidence. The patterns don't match credentials formatted as examples. Mockup values, redacted values, and placeholders, like credential type or usage descriptions, in the position where an actual secret value should present won't be matched.

Keywords

Keyword_SymmetricKey256:

- SharedAccessKey
- AccountKey

Keyword_SymmetricKey256UrlEncoded:

- sig=
- key
- token
- secret
- password

Last updated on 11/18/2025

Azure SignalR access key

This SIT is also included in the All credentials bundled SIT.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a credential scanning SIT in an endpoint DLP policy. This requirement is specific to the combination of credential scanning SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

Format

A combination of 43 characters consisting of letters, digits, and special characters ending in an equals sign (=) that isn't part of the pattern.

Pattern

A combination of 43 characters consisting of:

- a-z (not case-sensitive)
- 0-9
- forward slashes (/)
- or plus signs (+)

format example:

```
C2dE3fH4iJ5kL6mN7oP8qR9sT0+V1w
```

Credential example

```
host: account.service.signalr.net; accesskey: C2dE3fH4iJ5kL6mN7oP8qR9sT0+V1w;
```

Checksum

No

SITs that have checksums use a unique calculation to check if the information is valid. This means when the **Checksum** value is **Yes**, the service can make a positive detection based on the sensitive data alone. When the **Checksum** value is **No** additional (secondary) elements must also be detected for the service to make a positive detection.

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

This SIT is designed to match the security information that's used to authenticate Azure SignalR clients when requests are made to the service.

It uses several primary resources:

- Patterns of Base64 encoded 256-bits symmetric key.
- Patterns of CredentialName, CredentialFeatures, AccountIdentityName, AccountIdentityValue, ResourceType, ResourceName, Id.
- Patterns of mockup values, redactions, and placeholders.
- A dictionary of vocabulary.

The patterns are designed to match actual credentials with reasonable confidence. The patterns don't match credentials formatted as examples. Mockup values, redacted values, and placeholders, like credential type or usage descriptions, in the position where an actual secret value should present won't be matched.

Keywords

Keyword_SymmetricKey256:

- SharedAccessKey
- AccountKey

Last updated on 11/18/2025

Azure SQL connection string

This SIT is also included in the All credentials bundled SIT.

> **Important**
>
> Because credential scanning SITs are a type of advanced classifier, you must enable advanced classification before implementing them. For information about enabling advanced classification, and about configuring endpoints and the supported file types, see Advanced classification scanning and protection.

Format

Up to 20,000-character combination of letters, digits, and special characters.

or

A pair of username and password used in general authentication process.

Pattern

Any combination of up to 20,000 characters consisting of:

- a-z (not case-sensitive)
- 0-9
- forward slashes (/)
- or plus signs (+)
- Up to 2 equal signs (=)

format example:

```
H4iJ5kL6mN7o==P8qR9sT0uV1wX2y/Z3a+
```

or

Variant username and password formats, for example:

```
username=...;password=********;
```

```
user id=...;password=********;
```

`uid=...;pwd=********;`

`DB_USER=...;DB_PASS=********;`

`Service Account=...;Password=********;`

Credential example

`server=server.database.windows.net;database=database;user=user;pwd=**********

Checksum

Yes

SITs that have checksums use a unique calculation to check if the information is valid. This means when the **Checksum** value is **Yes**, the service can make a positive detection based on the sensitive data alone. When the **Checksum** value is **No** additional (secondary) elements must also be detected for the service to make a positive detection.

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

This SIT is designed to match the security information that's used to connect to Azure SQL Databases.

It uses several primary resources:

- Patterns of Base64 encoded string literal.
- Patterns of Plain-text username and password.
- Patterns of CredentialName, CredentialFeatures, AccountIdentityName, AccountIdentityValue, ResourceType, ResourceName.
- Patterns of mockup values, redactions, and placeholders.

The patterns are designed to match actual credentials with reasonable confidence. The patterns don't match credentials formatted as examples. Mockup values, redacted values, and

placeholders, like credential type or usage descriptions, in the position where an actual secret value should present won't be matched.

Keywords

Keyword_Base64EncodedStringLiteral:

- MII

Keyword_LoginCredentials:

- password
- pw
- DB_

Last updated on 04/09/2026

Azure storage account access key

This SIT is also included in the All credentials bundled SIT.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a credential scanning SIT in an endpoint DLP policy. This requirement is specific to the combination of credential scanning SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

Format

A combination of up to 20,000 characters consisting of letters, digits, and special characters.

or

A combination of 88 characters consisting of letters, digits, and special characters.

Pattern

Any combination of up to 20,000 characters consisting of:

- a-z (not case-sensitive)
- 0-9
- forward slashes (/)
- or plus signs (+)
- Up to 2 equal signs (=)

format example:

`H4iJ5kL6mN7o==P8qR9sT0uV1wX2y/Z3a` or

Any combination of 86 characters consisting of:

a-z (not case-sensitive) 0-9 forward slashes (/) or plus signs (+) ends with two equal signs (=)

format example:

`N7oP8qR9sT0uV1wX2yZ3aB4/cD5eF6g==`

Credential example

`Endpoint=account.table.core.windows.net;AccountName=account;AccountKey=N7oP8qR9sT0uV1wX2yZ3aB4/cD5eF6g==`

Checksum

Yes

SITs that have checksums use a unique calculation to check if the information is valid. This means when the **Checksum** value is **Yes**, the service can make a positive detection based on the sensitive data alone. When the **Checksum** value is **No** additional (secondary) elements must also be detected for the service to make a positive detection.

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

This SIT is designed to match the security information that's used to make request against Azure Storage services, like Blob, Queue, Table and File services.

It uses several primary resources:

- Patterns of Base64 encoded string literal.
- Patterns of Base64 encoded 512-bits symmetric key.
- Patterns of CredentialName, CredentialFeatures, AccountIdentityName, AccountIdentityValue, ResourceType, ResourceName, Id, AccountName.
- Patterns of mockup values, redactions, and placeholders.
- A dictionary of vocabulary.

The patterns are designed to match actual credentials with reasonable confidence. The patterns don't match credentials formatted as examples. Mockup values, redacted values, and

placeholders, like credential type or usage descriptions, in the position where an actual secret value should present won't be matched.

Keywords

Keyword_Base64EncodedStringLiteral:

- MII

Keyword_SymmetricKey512:

- SharedAccessKey
- AccountKey

Last updated on 11/18/2025

Azure storage account key

Article • 02/10/2025

Format

The string `DefaultEndpointsProtocol` followed by the characters and strings outlined in the pattern below, including the string `AccountKey`.

Pattern

- the string `DefaultEndpointsProtocol`
- zero to two whitespace characters
- an equal sign (=)
- zero to two whitespace characters
- any combination of between 1-200 lower- or uppercase letters, digits, symbols, special characters, or spaces
- the string `AccountKey`
- zero to two whitespace characters
- an equal sign (=)
- zero to two whitespace characters
- any combination of 86 characters that are lower- or uppercase letters, digits, forward slash (/), or plus sign (+)
- two equal signs (=)

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `CEP_Regex_AzureStorageAccountKey` finds content that matches the pattern.
- The regular expression `CEP_AzureEmulatorStorageAccountFilter` doesn't find content that matches the pattern.

- The regular expression `CEP_CommonExampleKeywords` doesn't find content that matches the pattern.

XML

```
<!--Azure Storage Account Key-->
<Entity id="c7bc98e8-551a-4c35-a92d-d2c8cda714a7" patternsProximity="300"
recommendedConfidence="85">
  <Pattern confidenceLevel="85">
        <IdMatch idRef="CEP_Regex_AzureStorageAccountKey" />
        <Any minMatches="0" maxMatches="0">
            <Match idRef="CEP_AzureEmulatorStorageAccountFilter" />
            <Match idRef="CEP_CommonExampleKeywords" />
        </Any>
  </Pattern>
</Entity>
```

Keywords

CEP_azure_emulator_storage_account_filter

This sensitive information type identifies these keywords by using a regular expression, not a keyword list.

- Eby8vdM02xNOcqFlqUwJPLlmEtlCDXJ1OUzFT50uSRZ6IFsuFq2UVErCz4I6tq/K1SZFPTOtr/KBHBeksoGMGw==

CEP_common_example_keywords

This sensitive information type identifies these keywords by using a regular expression, not a keyword list.

- contoso
- fabrikam
- northwind
- sandbox
- onebox
- localhost
- 127.0.0.1
- testacs.com
- s-int.net

Azure Storage account key (generic)

Article • 02/10/2025

Format

Any combination of 86 lower- or uppercase letters, digits, the forward slash (/), or plus sign (+), preceded or followed by the characters outlined in the pattern below.

Pattern

- zero to one of the greater than symbol (>), apostrophe ('), equal sign (=), quotation mark ("), or number sign (#)
- any combination of 86 characters that are lower- or uppercase letters, digits, the forward slash (/), or plus sign (+)
- two equal signs (=)

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `CEP_Regex_AzureStorageAccountKeyGeneric` finds content that matches the pattern.

XML

```
<!--Azure Storage Account Key (Generic)-->
<Entity id="7ff41bd0-5419-4523-91d6-383b3a37f084" patternsProximity="300"
recommendedConfidence="85">
  <Pattern confidenceLevel="85">
        <IdMatch idRef="CEP_Regex_AzureStorageAccountKeyGeneric" />
  </Pattern>
</Entity>
```

Azure Storage account shared access signature

This SIT is also included in the All credentials bundled SIT.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a credential scanning SIT in an endpoint DLP policy. This requirement is specific to the combination of credential scanning SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

Format

A combination of 44 characters consisting of letters, digits, and special characters.

or

A combination of up to 76 characters consisting of letters, digits, and special characters.

Pattern

Any combination of 43 characters consisting of:

- a-z (not case-sensitive)
- 0-9
- forward slashes (/)
- or plus signs (+)
- ends with an equal sign (=)

format example:

```
aB1cD2eF+3gH4iJ5kL6/mN7oP8qR=
```

or

A combination of 43 to 73 characters consisting of:

- a-z (not case-sensitive)
- 0-9
- or percent signs (%)
- ends with a suffix '%3d' (not case-sensitive)

for example:

`aB1cD2eF3gH4iJ5kL6mN7oP8q%R%3D`

Credential example

`https://account.blob.core.windows.net/?`

`sr=...&sv=...&st=...&se=...&sp=...&sig=aB1cD2eF3gH4iJ5kL6mN7oP8q%R%3D`

Checksum

No

SITs that have checksums use a unique calculation to check if the information is valid. This means when the **Checksum** value is **Yes**, the service can make a positive detection based on the sensitive data alone. When the **Checksum** value is **No** additional (secondary) elements must also be detected for the service to make a positive detection.

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

This SIT is designed to match the security information that's used to grant restricted access rights to Azure Storage resources.

It uses several primary resources:

- Patterns of Base64 encoded 256-bits symmetric key.
- Patterns of URL Encoded 256-bits symmetric key.

- Patterns of CredentialName, CredentialFeatures, AccountIdentityName, AccountIdentityValue, ResourceType, ResourceName, ID.
- Patterns of mockup values, redactions, and placeholders.
- A dictionary of vocabulary.

The patterns are designed to match actual credentials with reasonable confidence. The patterns don't match credentials formatted as examples. Mockup values, redacted values, and placeholders, like credential type or usage descriptions, in the position where an actual secret value should present won't be matched.

Keywords

Keyword_SymmetricKey256:

- SharedAccessKey
- AccountKey

Keyword_SymmetricKey256UrlEncoded:

- sig=
- key
- token
- secret
- password

Last updated on 11/18/2025

Azure Storage account shared access signature for high risk resources (preview)

This SIT is also included in the All credentials bundled SIT.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a credential scanning SIT in an endpoint DLP policy. This requirement is specific to the combination of credential scanning SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

Format

A combination of 44 characters consisting of letters, digits, and special characters.

or

A combination of up to 76 characters consisting of letters, digits, and special characters.

Pattern

Any combination of 43 characters consisting of:

- a-z (not case-sensitive)
- 0-9
- forward slashes (/)
- or plus signs (+)
- ends with an equal sign (=)

format example:

```
aB1cD2eF3gH4iJ5kL6mN7oP8qR/+=
```

or

Any combination of 43 to 73 characters consisting of:

- a-z (not case-sensitive)
- 0-9
- or percent signs (%)
- ends with a suffix '%3d' (not case-sensitive)

format example:

`A1bC2dE3fH4iJ5kL6mN7oP8qR9sT0u%3D`

Credential example

`https://account.blob.core.windows.net/file.cspkg?...&sig=eF3gH4iJ5kL6mN7oP8%qR9sT0uV=3D`

Checksum

No

SITs that have checksums use a unique calculation to check if the information is valid. This means when the **Checksum** value is **Yes**, the service can make a positive detection based on the sensitive data alone. When the **Checksum** value is **No** additional (secondary) elements must also be detected for the service to make a positive detection.

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

This SIT is designed to match the security information that's used to grant restricted access rights to high risk Azure Storage resources such as certificates, configurations, or deployment packages.

It uses several primary resources:

- Patterns of Base64 encoded 256-bits symmetric key.
- Patterns of URL Encoded 256-bits symmetric key.
- Patterns of CredentialName, CredentialFeatures, AccountIdentityName, AccountIdentityValue, ResourceType, ResourceName, Id.

- Patterns of mockup values, redactions, and placeholders.
- A dictionary of vocabulary

The patterns are designed to match actual credentials with reasonable confidence. The patterns don't match credentials formatted as examples. Mockup values, redacted values, and placeholders, like credential type or usage descriptions, in the position where an actual secret value should present won't be matched.

Keywords

Keyword_SymmetricKey256:

- SharedAccessKey
- AccountKey

Keyword_SymmetricKey256UrlEncoded:

- sig=
- key
- token
- secret
- password

Last updated on 11/18/2025

Azure subscription management certificate

This SIT is also included in the All credentials bundled SIT.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a credential scanning SIT in an endpoint DLP policy. This requirement is specific to the combination of credential scanning SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

Format

A combination of up to 20,000 characters consisting of letters, digits, and special characters.

Pattern

A combination of up to 20,000 characters:

- a-z (not case-sensitive)
- 0-9
- forward slash (/) or plus sign (+)
- Up to two equals signs (==)

format example:

```
cD2eF3gH4iJ5kL6mN7/oP8qR9sT==
```

Credential example

```
<Subscription id="f70163aa-03a8-4f4a-8a30-d38e3f38fde4"
ManagementCertificate="MIIPuQIBAzCCD38GCSqGS...
```

> **Important**
>
> This example has been truncated. It's not a detectable example of this SIT.

Checksum

Yes

SITs that have checksums use a unique calculation to check if the information is valid. This means when the **Checksum** value is **Yes**, the service can make a positive detection based on the sensitive data alone. When the **Checksum** value is **No** additional (secondary) elements must also be detected for the service to make a positive detection.

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

This SIT is designed to match the security information that's used to authenticate with the classic deployment model provided by Azure. Many programs and tools, like Visual Studio or the Azure SDK, use these certificates to automate configuration and deployment of various Azure services.

It uses several primary resources:

- Patterns of Base64 encoded string literal.
- Patterns of CredentialName, CredentialFeatures, AccountIdentityName, AccountIdentityValue, ResourceType, ResourceName.

The patterns are designed to match actual credentials with reasonable confidence. The patterns don't match credentials formatted as examples. Mockup values, redacted values, and placeholders, like credential type or usage descriptions, in the position where an actual secret value should present won't be matched.

Keywords

Keyword_Base64EncodedStringLiteral:

- MII

Last updated on 11/18/2025

Belgium drivers license number

Article • 08/19/2024

Format

10 digits without spaces and delimiters

Pattern

10 digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_belgium_eu_driver's_license_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_driver's_license_number` or `Keywords_belgium_eu_driver's_license_number` is found.

XML

```
      <!-- Belgium Driver's License Number -->
      <Entity id="d89fd329-9324-433c-b687-2c37bd5166f3"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_belgium_eu_driver's_license_number" />
```

```
          <Any minMatches="1">
            <Match idRef="Keywords_eu_driver's_license_number" />
            <Match idRef="Keywords_belgium_eu_driver's_license_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences

- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#

- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses

- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_belgium_eu_driver's_license_number

- rijbewijs
- rijbewijsnummer
- führerschein
- führerscheinnummer
- füehrerscheinnummer
- fuhrerschein
- fuehrerschein
- fuhrerscheinnummer
- fuehrerscheinnummer
- permis de conduire
- numéro permis conduire

Belgium national number

Article • 08/19/2024

Format

11 digits plus optional delimiters

Pattern

11 digits plus delimiters:

- six digits and two optional periods in the format YY.MM.DD for date of birth
- An optional delimiter from dot, dash, space
- three sequential digits (odd for males, even for females)
- An optional delimiter from dot, dash, space
- two check digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_belgium_national_number` finds content that matches the pattern.
- A keyword from `Keyword_belgium_national_number` is found.
- The checksum passes.

A DLP policy has low confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_belgium_national_number` finds content that matches the pattern.
- The checksum passes.

XML

```
<!-- Belgium National Number -->
       <Entity id="fb969c9e-0fd1-4b18-8091-a2123c5e6a54"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_belgium_national_number" />
          <Match idRef="Keyword_belgium_national_number" />
        </Pattern>
        <Pattern confidenceLevel="65">
          <IdMatch idRef="Func_belgium_national_number" />
        </Pattern>
      </Entity>
```

Keywords

Keyword_belgium_national_number

- belasting aantal
- bnn#
- bnn
- carte d'identité
- identifiant national
- identifiantnational#
- identificatie
- identification
- identifikation
- identifikationsnummer
- identifizierung
- identité
- identiteit
- identiteitskaart
- identity
- inscription
- national number
- national register

- nationalnumber#
- nationalnumber
- nif#
- nif
- numéro d'assuré
- numéro de registre national
- numéro de sécurité
- numéro d'identification
- numéro d'immatriculation
- numéro national
- numéronational#
- personal id number
- personalausweis
- personalidnumber#
- registratie
- registration
- registrationsnumme
- registrierung
- social security number
- ssn#
- ssn
- steuernummer
- tax id
- tax identification no
- tax identification number
- tax no#
- tax no
- tax number
- tax registration number
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id
- tin no
- tin#

Belgium passport number

Article • 08/19/2024

Format

two letters followed by six digits with no spaces or delimiters

Pattern

two letters and followed by six digits

Checksum

not applicable

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_belgium_eu_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_belgium_eu_passport_number` is found.
- The regular expression `Regex_eu_passport_date2` finds date in the format DD MM YY or a keyword from `Keywords_eu_passport_date` or `Keywords_belgium_eu_passport_number` is found

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_belgium_eu_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_belgium_eu_passport_number` is found.

XML

```
      <!-- Belgium Passport Number -->
      <Entity id="d7b1315b-21ca-4774-a32a-596010ff78fd"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Regex_belgium_eu_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_belgium_eu_passport_number" />
          </Any>
          <Any minMatches="1">
            <Match idRef="Regex_eu_passport_date2" />
            <Match idRef="Keywords_eu_passport_date" />
            <Match idRef="Keywords_belgium_eu_passport_date" />
          </Any>
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_belgium_eu_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_belgium_eu_passport_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_passport_number

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers

- passport numbers

Keywords_belgium_eu_passport_number

- numéro passeport
- paspoort nr
- paspoort-nr
- paspoortnummer
- paspoortnummers
- Passeport carte
- Passeport livre
- Pass-Nr
- Passnummer
- reisepass kein

Keywords_eu_passport_date

- date of issue
- date of expiry

Belgium physical addresses

Article • 08/19/2024

This unbundled named entity detects patterns related to physical addresses from Belgium. It's also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Belgium value added tax number

Article • 08/19/2024

Format

12-character alphanumeric pattern

Pattern

12-character alphanumeric pattern:

- a letter B or b
- a letter E or e
- a digit 0
- a digit from 1 to 9
- an optional dot or Hyphen or space
- four digits
- an optional dot or Hyphen or space
- four digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_belgium_value_added_tax_number` finds content that matches the pattern.

- A keyword from `Keywords_belgium_value_added_tax_number` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_belgium_value_added_tax_number` finds content that matches the pattern.

XML

```
      <!-- Belgium Value Added Tax Number -->
      <Entity id="85b5b3c3-f2de-4ae8-ac46-fd3cb38bf9ed"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_belgium_value_added_tax_number" />
          <Match idRef="Keywords_belgium_value_added_tax_number" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_belgium_value_added_tax_number" />
        </Pattern>
      </Entity>
    </Version>
```

Keywords

Keyword_belgium_value_added_tax_number

- nº tva
- vat number
- vat no
- numéro t.v.a
- umsatzsteuer-identifikationsnummer
- umsatzsteuernummer
- btw
- btw#
- vat#

Blood test terms

Article • 08/19/2024

This unbundled named entity detects terms related to blood tests, such as *hCG*. It supports English terms only. It's also included in the All medical terms and conditions bundled named entity SIT.

Confidence level

High

Keyword Highlighting

***UNKNOWN**

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Brand medication names

Article • 08/19/2024

This unbundled named entity detects names of brand medication, such as *Tylenol*. It supports English terms only. It's also included in the All medical terms and conditions bundled named entity SIT.

Confidence level

High

Keyword Highlighting

***UNKNOWN**

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Brazil CPF number

Article • 08/19/2024

Format

11 digits that include a check digit and can be formatted or unformatted

Pattern

Formatted:

- three digits
- a period
- three digits
- a period
- three digits
- a hyphen
- two digits that are check digits

Unformatted:

- 11 digits where the last two digits are check digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_brazil_cpf` finds content that matches the pattern.
- A keyword from `Keyword_brazil_cpf` is found.
- The checksum passes.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_brazil_cpf` finds content that matches the pattern.
- The checksum passes.

XML

```
<!-- Brazil CPF Number -->
<Entity id="78e09124-f2c3-4656-b32a-c1a132cd2711" recommendedConfidence="85"
patternsProximity="300">
  <Pattern confidenceLevel="85">
     <IdMatch idRef="Func_brazil_cpf"/>
     <Match idRef="Keyword_brazil_cpf"/>
  </Pattern>
  <Pattern confidenceLevel="75">
     <IdMatch idRef="Func_brazil_cpf"/>
  </Pattern>
</Entity>
```

Keywords

Keyword_brazil_cpf

- CPF
- Identification
- Registration
- Revenue
- Cadastro de Pessoas Físicas
- Imposto
- Identificação
- Inscrição
- Receita

Brazil legal entity number (CNPJ)

Article • 08/19/2024

Format

14 digits that include a registration number, branch number, and check digits, plus delimiters

Pattern

14 digits, plus delimiters:

- two digits
- a period
- three digits
- a period
- three digits (these first eight digits are the registration number)
- a forward slash
- four-digit branch number
- a hyphen
- two digits that are check digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_brazil_cnpj` finds content that matches the pattern.
- A keyword from `Keyword_brazil_cnpj` is found.
- The checksum passes.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_brazil_cnpj` finds content that matches the pattern.
- The checksum passes.

XML

```
<!-- Brazil Legal Entity Number (CNPJ) -->
<Entity id="9b58b5cd-5e90-4df6-b34f-1ebcc88ceae4" recommendedConfidence="85"
patternsProximity="300">
   <Pattern confidenceLevel="85">
     <IdMatch idRef="Func_brazil_cnpj"/>
     <Match idRef="Keyword_brazil_cnpj"/>
  </Pattern>
  <Pattern confidenceLevel="75">
     <IdMatch idRef="Func_brazil_cnpj"/>
  </Pattern>
</Entity>
```

Keywords

Keyword_brazil_cnpj

- CNPJ
- CNPJ/MF
- CNPJ-MF
- National Registry of Legal Entities
- Taxpayers Registry
- Legal entity
- Legal entities
- Registration Status
- Business
- Company
- CNPJ
- Cadastro Nacional da Pessoa Jurídica
- Cadastro Geral de Contribuintes
- CGC
- Pessoa jurídica

- Pessoas jurídicas
- Situação cadastral
- Inscrição
- Empresa

Brazil national identification card (RG)

Article • 08/19/2024

Format

Registro Geral (old format): Nine digits

Registro de Identidade (RIC) (new format): 11 digits

Pattern

Registro Geral (old format):

- two digits
- a period
- three digits
- a period
- three digits
- a hyphen
- one digit that is a check digit

Registro de Identidade (RIC) (new format):

- 10 digits
- a hyphen
- one digit that is a check digit

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_brazil_rg` finds content that matches the pattern.
- A keyword from `Keyword_brazil_rg` is found.
- The checksum passes.

XML

```
      <!-- Brazil National ID Card (RG) -->
      <Entity id="486de900-db70-41b3-a886-abdf25af119c"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_brazil_rg" />
          <Match idRef="Keyword_brazil_rg" />
        </Pattern>
      </Entity>
```

Keywords

Keyword_brazil_rg

- Cédula de identidade
- identity card
- national id
- número de rregistro
- registro de lidentidade
- registro geral
- RG (this keyword is case-sensitive)
- RIC (this keyword is case-sensitive)

Brazil physical addresses

Article • 08/19/2024

This unbundled named entity detects patterns related to physical address from Brazil. It's also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Bulgaria drivers license number

Article • 08/19/2024

Format

nine digits without spaces and delimiters

Pattern

nine digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_bulgaria_eu_driver's_license_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_driver's_license_number` or `Keywords_bulgaria_eu_driver's_license_number` is found.

XML

```
      <!-- Bulgaria Driver's License Number -->
      <Entity id="66d39258-94c2-43b2-804b-aa312258e54b"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_bulgaria_eu_driver's_license_number" />
```

```
          <Any minMatches="1">
            <Match idRef="Keywords_eu_driver's_license_number" />
            <Match idRef="Keywords_bulgaria_eu_driver's_license_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences

- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#

- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses

- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_bulgaria_eu_driver's_license_number

- свидетелство за управление на мпс
- свидетелство за управление на моторно превозно средство
- сумпс
- шофьорска книжка
- шофьорски книжки

Bulgaria passport number

Article • 08/19/2024

Format

nine digits without spaces and delimiters

Pattern

nine digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_bulgaria_eu_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_bulgaria_eu_passport_number` is found.
- The regular expression `Regex_eu_passport_date1` finds date in the format DD.MM.YYYY or a keyword from `Keywords_eu_passport_date` is found

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_bulgaria_eu_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_bulgaria_eu_passport_number` is found.

XML

```
      <!-- Bulgaria Passport Number -->
      <Entity id="f7172b82-c588-4216-845e-4e54e397f29a"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Regex_bulgaria_eu_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_bulgaria_eu_passport_number" />
          </Any>
          <Any minMatches="1">
            <Match idRef="Regex_eu_passport_date1" />
            <Match idRef="Keywords_eu_passport_date" />
          </Any>
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_bulgaria_eu_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_bulgaria_eu_passport_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_passport_number

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers

Keywords_bulgaria_eu_passport_number

- номер на паспорта
- номер на паспорт
- паспорт №

Keywords_eu_passport_date

- date of issue
- date of expiry

Bulgaria physical addresses

Article • 08/19/2024

This unbundled named entity detects patterns related to physical address from Bulgaria. It's also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Bulgaria uniform civil number

Article • 08/19/2024

Format

10 digits without spaces and delimiters

Pattern

10 digits without spaces and delimiters

- six digits that correspond to the birth date (YYMMDD)
- two digits that correspond to the birth order
- one digit that corresponds to gender: An even digit for male and an odd digit for female
- one check digit

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_bulgaria_eu_national_id_card` finds content that matches the pattern.
- A keyword from `Keywords_bulgaria_eu_national_id_card` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_bulgaria_eu_national_id_card` finds content that matches the pattern.

XML

```
      <!-- Bulgaria Uniform Civil Number -->
      <Entity id="100d58b1-0a35-4fb1-aa89-e4a86fb53fcc"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_bulgaria_eu_national_id_card" />
          <Match idRef="Keywords_bulgaria_eu_national_id_card" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_bulgaria_eu_national_id_card" />
          <Any minMatches="0" maxMatches="0">
            <Match idRef="Keywords_bulgaria_eu_telephone_number" />
            <Match idRef="Keywords_bulgaria_eu_mobile_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_bulgaria_eu_national_id_card

- bnn#
- bnn
- bucn#
- bucn
- edinen grazhdanski nomer
- egn#
- egn
- identification number
- national id
- national number
- nationalnumber#
- nationalnumber
- personal id
- personal no
- personal number
- personalidnumber#

- social security number
- ssn#
- ssn
- uniform civil id
- uniform civil no
- uniform civil number
- uniformcivilno#
- uniformcivilno
- uniformcivilnumber#
- uniformcivilnumber
- unique citizenship number
- егн#
- егн
- единен граждански номер
- идентификационен номер
- личен номер
- лична идентификация
- лично не
- национален номер
- номер на гражданството
- униформ id
- униформ граждански id
- униформ граждански не
- униформ граждански номер
- униформгражданскиid#
- униформгражданскине.#

Canada bank account number

Article • 08/19/2024

Format

7 or 12 digits

Pattern

A Canada Bank Account Number is 7 or 12 digits.

A Canada bank account transit number is:

- five digits
- a hyphen
- three digits OR
- a zero "0"
- eight digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_canada_bank_account_number` finds content that matches the pattern.
- A keyword from `Keyword_canada_bank_account_number` is found.

- The regular expression `Regex_canada_bank_account_transit_number` finds content that matches the pattern.

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_canada_bank_account_number` finds content that matches the pattern.
- A keyword from `Keyword_canada_bank_account_number` is found.

XML

```
<!-- Canada Bank Account Number -->
<Entity id="552e814c-cb50-4d94-bbaa-bb1d1ffb34de" patternsProximity="300"
recommendedConfidence="75">
  <Pattern confidenceLevel="85">
        <IdMatch idRef="Regex_canada_bank_account_number" />
        <Match idRef="Keyword_canada_bank_account_number" />
        <Match idRef="Regex_canada_bank_account_transit_number" />
   </Pattern>
   <Pattern confidenceLevel="75">
        <IdMatch idRef="Regex_canada_bank_account_number" />
        <Match idRef="Keyword_canada_bank_account_number" />
   </Pattern>
</Entity>
```

Keywords

Keyword_canada_bank_account_number

- canada savings bonds
- canada revenue agency
- canadian financial institution
- direct deposit form
- canadian citizen
- legal representative
- notary public
- commissioner for oaths
- child care benefit
- universal child care
- canada child tax benefit
- income tax benefit
- harmonized sales tax

- social insurance number
- income tax refund
- child tax benefit
- territorial payments
- institution number
- deposit request
- banking information
- direct deposit

Canada drivers license number

Article • 08/19/2024

Format

Varies by province

Pattern

Various patterns covering:

Alberta

- six digits
- a hyphen
- three digits

Or

- five to nine digits

British Columbia

- seven digits

Manitoba

- two letters
- optional hyphen
- two letters
- optional hyphen
- two letters
- optional hyphen
- one letter
- three digits
- two letters

New Brunswick

- five to seven digits

Newfoundland/Labrador

- one letter
- nine digits

Nova Scotia

- five letters
- optional hyphen
- one digit; any of 0, 1, 2 or 3
- one digit
- one digit zero or one
- six digits

Ontario

- one letter
- four digits
- optional hyphen
- five digits
- one digit
- one digit; any of 0, 1, 5, or 6
- one digit
- one digit; any of 0, 1, 2 or 3
- one digit

Prince Edward Island

- five to six digits

Quebec

- one letter
- twelve digits

Saskatchewan

- eight digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_[province_name]_drivers_license_number` finds content that matches the pattern.
- A keyword from `Keyword_[province_name]_drivers_license_name` is found.
- A keyword from `Keyword_canada_drivers_license` is found.

XML

```
<!-- Canada Driver's License Number -->
    <Entity id="37186abb-8e48-4800-ad3c-e3d1610b3db0"
patternsProximity="300" recommendedConfidence="75">
      <Pattern confidenceLevel="75">
        <IdMatch idRef="Func_alberta_drivers_license_number" />
        <Match idRef="Keyword_alberta_drivers_license_name" />
        <Match idRef="Keyword_canada_drivers_license" />
      </Pattern>
      <Pattern confidenceLevel="75">
        <IdMatch idRef="Func_british_columbia_drivers_license_number" />
        <Match idRef="Keyword_british_columbia_drivers_license_name" />
        <Match idRef="Keyword_canada_drivers_license" />
      </Pattern>
      <Pattern confidenceLevel="75">
        <IdMatch idRef="Func_manitoba_drivers_license_number" />
        <Match idRef="Keyword_manitoba_drivers_license_name" />
        <Match idRef="Keyword_canada_drivers_license" />
      </Pattern>
      <Pattern confidenceLevel="75">
        <IdMatch idRef="Func_new_brunswick_drivers_license_number" />
        <Match idRef="Keyword_new_brunswick_drivers_license_name" />
        <Match idRef="Keyword_canada_drivers_license" />
      </Pattern>
      <Pattern confidenceLevel="75">
        <IdMatch idRef="Func_newfoundland_labrador_drivers_license_number"
/>
        <Match idRef="Keyword_newfoundland_labrador_drivers_license_name" />
        <Match idRef="Keyword_canada_drivers_license" />
      </Pattern>
      <Pattern confidenceLevel="75">
        <IdMatch idRef="Func_nova_scotia_drivers_license_number" />
        <Match idRef="Keyword_nova_scotia_drivers_license_name" />
        <Match idRef="Keyword_canada_drivers_license" />
      </Pattern>
```

```
      <Pattern confidenceLevel="75">
        <IdMatch idRef="Func_ontario_drivers_license_number" />
        <Match idRef="Keyword_ontario_drivers_license_name" />
        <Match idRef="Keyword_canada_drivers_license" />
      </Pattern>
      <Pattern confidenceLevel="75">
        <IdMatch idRef="Func_prince_edward_island_drivers_license_number" />
        <Match idRef="Keyword_prince_edward_island_drivers_license_name" />
        <Match idRef="Keyword_canada_drivers_license" />
      </Pattern>
      <Pattern confidenceLevel="75">
        <IdMatch idRef="Func_quebec_drivers_license_number" />
        <Match idRef="Keyword_quebec_drivers_license_name" />
        <Match idRef="Keyword_canada_drivers_license" />
      </Pattern>
      <Pattern confidenceLevel="75">
        <IdMatch idRef="Func_saskatchewan_drivers_license_number" />
        <Match idRef="Keyword_saskatchewan_drivers_license_name" />
        <Match idRef="Keyword_canada_drivers_license" />
      </Pattern>
    </Entity>
```

Keywords

Keyword_[province_name]_drivers_license_name

- The province abbreviation, for example AB
- The province name, for example Alberta

Keyword_canada_drivers_license

- DL
- DLS
- CDL
- CDLS
- DriverLic
- DriverLics
- DriverLicense
- DriverLicenses
- DriverLicence
- DriverLicences
- Driver Lic
- Driver Lics
- Driver License

- Driver Licenses
- Driver Licence
- Driver Licences
- DriversLic
- DriversLics
- DriversLicence
- DriversLicences
- DriversLicense
- DriversLicenses
- Drivers Lic
- Drivers Lics
- Drivers License
- Drivers Licenses
- Drivers Licence
- Drivers Licences
- Driver'Lic
- Driver'Lics
- Driver'License
- Driver'Licenses
- Driver'Licence
- Driver'Licences
- Driver' Lic
- Driver' Lics
- Driver' License
- Driver' Licenses
- Driver' Licence
- Driver' Licences
- Driver'sLic
- Driver'sLics
- Driver'sLicense
- Driver'sLicenses
- Driver'sLicence
- Driver'sLicences
- Driver's Lic
- Driver's Lics
- Driver's License
- Driver's Licenses
- Driver's Licence
- Driver's Licences
- Permis de Conduire
- id

- ids
- idcard number
- idcard numbers
- idcard #
- idcard #s
- idcard card
- idcard cards
- idcard
- identification number
- identification numbers
- identification #
- identification #s
- identification card
- identification cards
- identification
- DL#
- DLS#
- CDL#
- CDLS#
- DriverLic#
- DriverLics#
- DriverLicense#
- DriverLicenses#
- DriverLicence#
- DriverLicences#
- Driver Lic#
- Driver Lics#
- Driver License#
- Driver Licenses#
- Driver License#
- Driver Licences#
- DriversLic#
- DriversLics#
- DriversLicense#
- DriversLicenses#
- DriversLicence#
- DriversLicences#
- Drivers Lic#
- Drivers Lics#
- Drivers License#
- Drivers Licenses#

- Drivers Licence#
- Drivers Licences#
- Driver'Lic#
- Driver'Lics#
- Driver'License#
- Driver'Licenses#
- Driver'Licence#
- Driver'Licences#
- Driver' Lic#
- Driver' Lics#
- Driver' License#
- Driver' Licenses#
- Driver' Licence#
- Driver' Licences#
- Driver'sLic#
- Driver'sLics#
- Driver'sLicense#
- Driver'sLicenses#
- Driver'sLicence#
- Driver'sLicences#
- Driver's Lic#
- Driver's Lics#
- Driver's License#
- Driver's Licenses#
- Driver's Licence#
- Driver's Licences#
- Permis de Conduire#
- id#
- ids#
- idcard card#
- idcard cards#
- idcard#
- identification card#
- identification cards#
- identification#

Canada health service number

Article • 08/19/2024

Format

10 digits

Pattern

10 digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_canada_health_service_number` finds content that matches the pattern.
- A keyword from `Keyword_canada_health_service_number` is found.

XML

```
<!-- Canada Health Service Number -->
<Entity id="59c0bf39-7fab-482c-af25-00faa4384c94" patternsProximity="300"
recommendedConfidence="75">
  <Pattern confidenceLevel="75">
        <IdMatch idRef="Regex_canada_health_service_number" />
        <Any minMatches="1">
```

```
          <Match idRef="Keyword_canada_health_service_number" />
        </Any>
  </Pattern>
</Entity>
```

Keywords

Keyword_canada_health_service_number

- personal health number
- patient information
- health services
- speciality services
- automobile accident
- patient hospital
- psychiatrist
- workers compensation
- disability

Canada passport number

Article • 08/19/2024

Format

two uppercase letters followed by six digits

Pattern

two uppercase letters followed by six digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_canada_passport_number` finds content that matches the pattern.
- A keyword from `Keyword_canada_passport_number` or `Keyword_passport` is found.

XML

```
<!-- Canada Passport Number -->
<Entity id="14d0db8b-498a-43ed-9fca-f6097ae687eb" patternsProximity="300"
recommendedConfidence="75">
  <Pattern confidenceLevel="75">
        <IdMatch idRef="Regex_canada_passport_number" />
        <Any minMatches="1">
```

```
          <Match idRef="Keyword_canada_passport_number" />
          <Match idRef="Keyword_passport" />
        </Any>
  </Pattern>
</Entity>
```

Keywords

Keyword_canada_passport_number

- canadian citizenship
- canadian passport
- passport application
- passport photos
- certified translator
- canadian citizens
- processing times
- renewal application

Keyword_passport

- Passport Number
- Passport No
- Passport #
- Passport#
- PassportID
- Passportno
- passportnumber
- パスポート
- パスポート番号
- パスポートのNum
- パスポート＃
- Numéro de passeport
- Passeport n °
- Passeport Non
- Passeport #
- Passeport#
- PasseportNon
- Passeportn °

Canada personal health identification number (PHIN)

Article • 08/19/2024

Format

nine digits

Pattern

nine digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_canada_phin` finds content that matches the pattern.
- At least two keywords from `Keyword_canada_phin` or `Keyword_canada_provinces` are found.

XML

```
<!-- Canada PHIN -->
<Entity id="722e12ac-c89a-4ec8-a1b7-fea3469f89db" patternsProximity="300"
recommendedConfidence="75">
  <Pattern confidenceLevel="75">
```

```
        <IdMatch idRef="Regex_canada_phin" />
        <Any minMatches="2">
          <Match idRef="Keyword_canada_phin" />
          <Match idRef="Keyword_canada_provinces" />
        </Any>
  </Pattern>
</Entity>
```

Keywords

Keyword_canada_phin

- social insurance number
- health information act
- income tax information
- manitoba health
- health registration
- prescription purchases
- benefit eligibility
- personal health
- power of attorney
- registration number
- personal health number
- practitioner referral
- wellness professional
- patient referral
- health and wellness

Keyword_canada_provinces

- Nunavut
- Quebec
- Northwest Territories
- Ontario
- British Columbia
- Alberta
- Saskatchewan
- Manitoba
- Yukon
- Newfoundland and Labrador
- New Brunswick

- Nova Scotia
- Prince Edward Island
- Canada

Canada physical addresses

Article • 08/19/2024

This unbundled named entity detects patterns related to physical address from Canada. It's also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Canada social insurance number

Article • 08/19/2024

Format

nine digits with optional hyphens or spaces

Pattern

Formatted:

- three digits
- a hyphen or space
- three digits
- a hyphen or space
- three digits

Unformatted: nine digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function Func_canadian_sin finds content that matches the pattern.
- At least one of the following patterns:
  - A keyword from Keyword_sin is found.

    - A keyword from Keyword_sin_collaborative is found.
- The checksum passes.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function Func_unformatted_canadian_sin finds content that matches the pattern.
- A keyword from Keyword_sin is found.
- The checksum passes.

XML

```
    <!-- Canada Social Insurance Number -->
    <Entity id="a2f29c85-ecb8-4514-a610-364790c0773e"
patternsProximity="300" recommendedConfidence="75">
      <Pattern confidenceLevel="85">
        <IdMatch idRef="Func_canadian_sin" />
        <Any minMatches="1">
          <Match idRef="Keyword_sin" />
          <Match idRef="Keyword_sin_collaborative" />
        </Any>
      </Pattern>
      <Pattern confidenceLevel="75">
        <IdMatch idRef="Func_unformatted_canadian_sin" />
        <Match idRef="Keyword_sin" />
      </Pattern>
    </Entity>
```

Keywords

Keyword_sin

- sin
- social insurance
- numero d'assurance sociale
- sins
- ssn
- ssns
- social security
- numero d'assurance social
- national identification number
- national id
- sin#

- soc ins
- social ins
- NAS (case sensitive)
- numéro d’assurance social
- numéro d’assurance sociale
- assurance social
- carte d’assurance sociale

Keyword_sin_collaborative

- driver's license
- drivers license
- driver's licence
- drivers licence
- DOB
- Birthdate
- Birthday
- Date of Birth

Chile identity card number

Article • 08/19/2024

Format

seven to eight digits plus delimiters a check digit or letter

Pattern

seven to eight digits plus delimiters:

- one to two digits
- an optional period
- three digits
- an optional period
- three digits
- a dash
- one digit or letter (not case-sensitive) which is a check digit

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_chile_id_card` finds content that matches the pattern.
- A keyword from `Keyword_chile_id_card` is found.

- The checksum passes.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_chile_id_card` finds content that matches the pattern.
- The checksum passes.

XML

```
<!-- Chile Identity Card Number -->
<Entity id="4e979794-49a0-407e-a0b9-2c536937b925" recommendedConfidence="85"
patternsProximity="300">
  <Pattern confidenceLevel="85">
     <IdMatch idRef="Func_chile_id_card"/>
     <Match idRef="Keyword_chile_id_card"/>
  </Pattern>
  <Pattern confidenceLevel="75">
     <IdMatch idRef="Func_chile_id_card"/>
  </Pattern>
</Entity>
```

Keywords

Keyword_chile_id_card

- cédula de identidad
- identificación
- national identification
- national identification number
- national id
- número de identificación nacional
- rol único nacional
- rol único tributario
- RUN
- RUT
- tarjeta de identificación
- Rol Unico Nacional
- Rol Unico Tributario
- RUN#
- RUT#
- nationaluniqueroleID#
- nacional identidad

- número identificación
- identidad número
- numero identificacion
- identidad numero
- Chilean identity no.
- Chilean identity number
- Chilean identity #
- Unique Tax Registry
- Unique Tributary Role
- Unique Tax Role
- Unique Tributary Number
- Unique National Number
- Unique National Role
- National unique role
- Chile identity no.
- Chile identity number
- Chile identity #
- R.U.T
- R.U.N

China physical addresses

This unbundled named entity detects patterns related to physical address from China. It is also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following services:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Last updated on 06/19/2026

China resident identity card (PRC) number

Article • 08/19/2024

Format

18 digits

Pattern

18 digits:

- six digits that are an address code
- eight digits in the form YYYYMMDD, which are the date of birth
- three digits that are an order code
- one digit that is a check digit

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_china_resident_id` finds content that matches the pattern.
- A keyword from `Keyword_china_resident_id` is found.
- The checksum passes.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_china_resident_id` finds content that matches the pattern.
- The checksum passes.

XML

```
<!-- China Resident Identity Card (PRC) Number -->
<Entity id="c92daa86-2d16-4871-901f-816b3f554fc1" recommendedConfidence="85"
patternsProximity="300">
  <Pattern confidenceLevel="85">
     <IdMatch idRef="Func_china_resident_id"/>
     <Match idRef="Keyword_china_resident_id"/>
  </Pattern>
  <Pattern confidenceLevel="75">
     <IdMatch idRef="Func_china_resident_id"/>
  </Pattern>
</Entity>
```

Keywords

Keyword_china_resident_id

- Resident Identity Card
- PRC
- National Identification Card
- 身份证
- 居民 身份证
- 居民身份证
- 鉴定
- 身分證
- 居民 身份證
- 鑑定

Client secret / API key

This SIT is also included in the All credentials bundled SIT.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a credential scanning SIT in an endpoint DLP policy. This requirement is specific to the combination of credential scanning SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

Format

A client secret or refresh token used in OAuth 2.0 protocol.

or

A combination of 24 characters consisting of letters, digits, and special characters.

or

A combination of 32 characters consisting of letters and digits.

or

A combination of 40 characters consisting of letters and digits.

or

A combination of 44 characters consisting of letters, digits, and special characters.

or

A combination of 56 characters consisting of letters, digits, and special characters

or

A combination of 88 characters consisting of letters, digits, and special characters.

Pattern

Various client secret or refresh token formats for example:

`ClientSecret:********`

`AppSecret=********`

`ConsumerKey:=********`

`Refresh_Token:********`

or

A combination of 22 characters:

- a-z (not case-sensitive)
- digits, forward slashes, or plus signs
- ends with two equal signs (=)

for example:

`abcdefgh0123456789/+AB==`

or

A combination of 32 characters:

- a-f or A-F (case-sensitive)
- or 0-9

for example:

`abcdef0123456789abcdef0123456789`

or

A combination of 40 characters:

- a-f or A-F (case-sensitive)

or

- 0-9

for example:

`Aa1Bb~2Cc3.-Dd4Ee5Ff6Gg7Hh8Ii9_Jj0Kk1Ll2`

or

A combination of 43 characters:

- a-z (not case-sensitive)
- 0-9
- forward slashes (/)
- or plus signs (+)
- ends with an equal sign (=)

for example:

`abcdefghijklmnopqrstuvwxyz0123456789/+ABCDE=`

or

A combination of 54 characters:

- a-z (not case-sensitive)
- 0-9
- forward slashes (/) or plus signs (+)
- ends with two equal signs (==)

for example:

`abcdefghijklmnopqrstuvwxyz0123456789/+ABCDEFGHIJKLMNOP==`

or

A combination of 86 characters:

- a-z (not case-sensitive)
- 0-9
- forward slashes (/) or plus signs (+)
- ends with two equal signs (=)

for example:

`abcdefghijklmnopqrstuvwxyz0123456789/+ABCDEabcdefghijklmnopqrstuvwxyz0123456789/+ABCDE==`

Credential example

`client_secret=abcdefghijklmnopqrstuvwxyz0123456789/+ABCDE=`

Checksum

No

SITs that have checksums use a unique calculation to check if the information is valid. This means when the **Checksum** value is **Yes**, the service can make a positive detection based on the sensitive data alone. When the **Checksum** value is **No** additional (secondary) elements must also be detected for the service to make a positive detection.

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

This SIT is designed to match the security information that's known only to the OAuth application and the authorization server to exchange for an access token at runtime.

It uses several primary resources:

- Patterns of Client secret context.
- Patterns of Base64 encoded 128 bits symmetric key.
- Patterns of Hex encoded 128 bits symmetric key.
- Patterns of Hex encoded 160 bits Symmetric Key.
- Patterns of Base64 encoded 256 bits symmetric key.
- Patterns of Base64 encoded 320 bits symmetric key.
- Patterns of Base64 encoded 512 bits symmetric key.
- Patterns of CredentialName, CredentialFeatures, AccountIdentityName, AccountIdentityValue, ResourceType, ResourceName, Id, AccountName.
- Patterns of mockup values, redactions, and placeholders.
- A dictionary of vocabulary.

The patterns are designed to match actual credentials with reasonable confidence. The patterns don't match credentials formatted as examples. Mockup values, redacted values, and placeholders, like credential type or usage descriptions, in the position where an actual secret value should present won't be matched.

Keywords

Keyword_ClientSecretContext:

- secret
- token
- auth
- securestring
- key

Keyword_SymmetricKey128:

- secret
- key
- password
- pw

Keyword_SymmetricKey128Hex:

- dapi
- key
- secret
- token
- password
- pw

Keyword_SymmetricKey160Hex:

- token

Keyword_SymmetricKey256:

- SharedAccessKey
- AccountKey

Keyword_SymmetricKey320:

- code=
- key

Keyword_SymmetricKey512:

- SharedAccessKey
- AccountKey

Last updated on 11/18/2025

Colombia national ID

Format

Eight to 10 digits with optional separators (dots or commas).

Pattern

Eight to 10 digits:

- Two digits followed by six digits (with optional dot or comma separators between groups of three), or
- The digit 1 followed by nine digits (with optional dot or comma separators between groups of three)

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Checksum

No

Definition

A data loss prevention (DLP) policy has medium confidence that it detects this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_colombia_national_id_number` finds content that matches the pattern.
- A keyword from `Keywords_colombia_national_id_number` is found.

XML

```
    <!-- Colombia National ID -->
    <Entity id="060c591a-0fd7-4598-a715-8a94b283dfcf" patternsProximity="300"
recommendedConfidence="75">
      <Pattern confidenceLevel="75">
        <IdMatch idRef="Regex_colombia_national_id_number" />
        <Match idRef="Keywords_colombia_national_id_number" />
      </Pattern>
    </Entity>
```

Keywords

Keywords_colombia_national_id_number

- Cedula
- Cédula
- documento de identificacion
- documento de identificación
- NUIP
- Número Único de Identificación

Last updated on 06/19/2026

Colombia tax identification number

Format

Ten digits with a verification digit, validated by a checksum function.

Pattern

Ten digits:

- Nine digits followed by
- One check digit (verification digit)

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Checksum

Yes

Definition

A data loss prevention (DLP) policy has high confidence that it detects this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_Colombian_Tax_Identification_Number` finds content that matches the pattern.
- A keyword from `Keyword_Colombian_Tax_Identification_Number` is found.

A DLP policy has medium confidence that it detects this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_Colombian_Tax_Identification_Number` finds content that matches the pattern.

XML

```xml
    <!-- Colombian Tax Identification Number-->
    <Entity id="df275e7b-40d0-4519-b668-59cee178df65" patternsProximity="300"
recommendedConfidence="85" filters="colombian_tax_id_filter">
      <Pattern confidenceLevel="85">
        <IdMatch idRef="Func_Colombian_Tax_Identification_Number" />
        <Match idRef="Keyword_Colombian_Tax_Identification_Number" />
      </Pattern>
      <Pattern confidenceLevel="75">
        <IdMatch idRef="Func_Colombian_Tax_Identification_Number" />
      </Pattern>
    </Entity>
```

Keywords

Keyword_Colombian_Tax_Identification_Number

- NIT.
- NIT
- nit.
- nit
- Nit.
- tax ID
- taxID#
- taxId
- taxidnumber
- tax number
- tax no
- tax #
- tax#
- taxpayer ID
- taxpayer number
- taxpayer no
- taxpayer #
- taxpayer#
- tax identity
- tax identification
- tin id
- tin no
- tin#

Last updated on 06/19/2026

Credit card number

Article • 04/23/2025

Format

14 to 19 digits that can be formatted or unformatted (ddddddddddddddd).

- Must pass the Luhn test.
- Must match number formats used by major credit card companies.

Pattern

Detects cards from all major brands. This includes Visa, MasterCard, Discover Card, JCB, American Express, gift cards, diner's cards, Rupay, and China UnionPay.

This sensitive information type doesn't flag reserved credit card numbers, such as the ones financial institutions use for testing

Checksum

Yes - the Luhn test, and verify match with number formats of major credit card companies.

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it detects this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_credit_card` finds content that matches the pattern.
- One of the following conditions is true:
  - A keyword from `Keyword_cc_verification` is found.
  - A keyword from `Keyword_cc_name` is found.
  - The function `Func_expiration_date` finds a date in the right date format.

- The checksum passes.

A DLP policy has low confidence that it detects this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_credit_card` finds content that matches the pattern.
- The checksum passes.

XML

```
<!-- Credit Card Number -->
<Entity id="50842eb7-edc8-4019-85dd-5a5c1f2bb085" patternsProximity="300"
recommendedConfidence="85">
  <Pattern confidenceLevel="85">
        <IdMatch idRef="Func_credit_card" />
        <Any minMatches="1">
          <Match idRef="Keyword_cc_verification" />
          <Match idRef="Keyword_cc_name" />
          <Match idRef="Func_expiration_date" />
        </Any>
  </Pattern>
  <Pattern confidenceLevel="65">
        <IdMatch idRef="Func_credit_card" />
  </Pattern>
</Entity>
```

Keywords

Keyword_cc_verification

- card verification
- card identification number
- cvn
- cid
- cvc2
- cvv2
- pin block
- security code
- security number
- security no
- issue number
- issue no
- cryptogramme
- numéro de sécurité

- numero de securite
- kreditkartenprüfnummer
- kreditkartenprufnummer
- prüfziffer
- prufziffer
- sicherheits Kode
- sicherheitscode
- sicherheitsnummer
- verfalldatum
- codice di verifica
- cod. sicurezza
- cod sicurezza
- n autorizzazione
- código
- codigo
- cod. seg
- cod seg
- código de segurança
- codigo de seguranca
- codigo de segurança
- código de seguranca
- cód. segurança
- cod. seguranca
- cod. segurança
- cód. seguranca
- cód segurança
- cod seguranca
- cod segurança
- cód seguranca
- número de verificação
- numero de verificacao
- ablauf
- gültig bis
- gültigkeitsdatum
- gultig bis
- gultigkeitsdatum
- scadenza
- data scad
- fecha de expiracion
- fecha de venc
- vencimiento

- válido hasta
- valido hasta
- vto
- data de expiração
- data de expiracao
- data em que expira
- validade
- valor
- vencimento
- transaction
- transaction number
- reference number
- セキュリティコード
- セキュリティ コード
- セキュリティナンバー
- セキュリティ ナンバー
- セキュリティ番号

Keyword_cc_name

- amex
- american express
- americanexpress
- americano espresso
- Visa
- mastercard
- master card
- mc
- mastercards
- master cards
- diner's Club
- diners club
- dinersclub
- discover
- discover card
- discovercard
- discover cards
- JCB
- BrandSmart
- japanese card bureau

- carte blanche
- carteblanche
- credit card
- cc#
- cc#:
- expiration date
- exp date
- expiry date
- date d'expiration
- date d'exp
- date expiration
- bank card
- bankcard
- card number
- card num
- cardnumber
- cardnumbers
- card numbers
- creditcard
- credit cards
- creditcards
- ccn
- card holder
- cardholder
- card holders
- cardholders
- check card
- checkcard
- check cards
- checkcards
- debit card
- debitcard
- debit cards
- debitcards
- atm card
- atmcard
- atm cards
- atmcards
- enroute
- en route
- card type

- Cardmember Acct
- cardmember account
- Cardno
- Corporate Card
- Corporate cards
- Type of card
- card account number
- card member account
- Cardmember Acct.
- card no.
- card no
- card number
- carte bancaire
- carte de crédit
- carte de credit
- numéro de carte
- numero de carte
- nº de la carte
- nº de carte
- kreditkarte
- karte
- karteninhaber
- karteninhabers
- kreditkarteninhaber
- kreditkarteninstitut
- kreditkartentyp
- eigentümername
- kartennr
- kartennummer
- kreditkartennummer
- kreditkarten-nummer
- carta di credito
- carta credito
- n. carta
- n carta
- nr. carta
- nr carta
- numero carta
- numero della carta
- numero di carta
- tarjeta credito

- tarjeta de credito
- tarjeta crédito
- tarjeta de crédito
- tarjeta de atm
- tarjeta atm
- tarjeta debito
- tarjeta de debito
- tarjeta débito
- tarjeta de débito
- nº de tarjeta
- no. de tarjeta
- no de tarjeta
- numero de tarjeta
- número de tarjeta
- tarjeta no
- tarjetahabiente
- cartão de crédito
- cartão de credito
- cartao de crédito
- cartao de credito
- cartão de débito
- cartao de débito
- cartão de debito
- cartao de debito
- débito automático
- debito automatico
- número do cartão
- numero do cartão
- número do cartao
- numero do cartao
- número de cartão
- numero de cartão
- número de cartao
- numero de cartao
- nº do cartão
- nº do cartao
- nº. do cartão
- no do cartão
- no do cartao
- no. do cartão
- no. do cartao

- rupay
- union pay
- unionpay
- diner's
- diners
- クレジットカード番号
- クレジットカードナンバー
- クレジットカード＃
- クレジットカード
- クレジット
- クレカ
- カード番号
- カードナンバー
- カード＃
- アメックス
- アメリカンエクスプレス
- アメリカン エクスプレス
- Visaカード
- Visa カード
- マスターカード
- マスター カード
- マスター
- ダイナースクラブ
- ダイナース クラブ
- ダイナース
- 有効期限
- 期限
- キャッシュカード
- キャッシュ カード
- カード名義人
- カードの名義人
- カードの名義
- デビット カード
- デビットカード
- 中国银联
- 银联

Croatia drivers license number

Article • 08/19/2024

Format

eight digits without spaces and delimiters

Pattern

eight digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_croatia_eu_driver's_license_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_driver's_license_number` or `Keywords_croatia_eu_driver's_license_number` is found.

XML

```
      <!-- Croatia Driver's License Number -->
      <Entity id="005b3ef1-47dd-4e68-bb02-c6db484d00f2"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_croatia_eu_driver's_license_number" />
```

```
          <Any minMatches="1">
            <Match idRef="Keywords_eu_driver's_license_number" />
            <Match idRef="Keywords_croatia_eu_driver's_license_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences

- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#

- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses

- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_croatia_eu_driver's_license_number

- vozačka dozvola
- vozačke dozvole

Croatia identity card number

Article • 08/19/2024

This entity is included in the EU National Identification Number sensitive information type. It's available as a stand-alone sensitive information type entity.

Format

nine digits

Pattern

nine consecutive digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_croatia_id_card` finds content that matches the pattern.
- A keyword from `Keyword_croatia_id_card` is found.

XML

```
<!--Croatia Identity Card Number-->
<Entity id="ff12f884-c20a-4189-b185-34c8e7258d47" recommendedConfidence="75"
patternsProximity="300">
  <Pattern confidenceLevel="75">
```

```
      <IdMatch idRef="Func_croatia_id_card"/>
      <Match idRef="Keyword_croatia_id_card"/>
    </Pattern>
  </Entity>
```

Keywords

Keyword_croatia_id_card

- majstorski broj građana
- master citizen number
- nacionalni identifikacijski broj
- national identification number
- oib#
- oib
- osobna iskaznica
- osobni id
- osobni identifikacijski broj
- personal identification number
- porezni broj
- porezni identifikacijski broj
- tax id
- tax identification no
- tax identification number
- tax no#
- tax no
- tax number
- tax registration number
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id
- tin no
- tin#

Croatia passport number

Article • 08/19/2024

Format

nine digits without spaces and delimiters

Pattern

nine digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_croatia_eu_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_croatia_eu_passport_number` is found.
- The regular expression `Regex_eu_passport_date1` finds date in the format DD.MM.YYYY or a keyword from `Keywords_eu_passport_date` is found

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_croatia_eu_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_croatia_eu_passport_number` is found.

XML

```
      <!-- Croatia Passport Number -->
      <Entity id="7d7a729d-32d8-4204-8d01-d5e6a6c25581"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Regex_croatia_eu_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_croatia_eu_passport_number" />
          </Any>
          <Any minMatches="1">
            <Match idRef="Regex_eu_passport_date1" />
            <Match idRef="Keywords_eu_passport_date" />
          </Any>
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_croatia_eu_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_croatia_eu_passport_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_passport_number_common

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers

Keywords_croatia_eu_passport_number

- broj putovnice
- br. Putovnice
- br putovnice

Croatia personal identification (OIB) number

Article • 08/19/2024

Format

11 digits

Pattern

11 digits:

- 10 digits
- final digit is a check digit

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_croatia_oib_number` finds content that matches the pattern.
- A keyword from `Keywords_croatia_eu_tax_file_number` is found.
- The checksum passes.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_croatia_oib_number` finds content that matches the pattern.
- The checksum passes.

XML

```
      <!-- Croatia Personal Identification (OIB) Number -->
      <Entity id="31983b6d-db95-4eb2-a630-b44bd091968d"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_croatia_oib_number" />
          <Match idRef="Keywords_croatia_eu_tax_file_number" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_croatia_oib_number" />
        </Pattern>
      </Entity>
```

Keywords

Keyword_croatia_oib_number

- majstorski broj građana
- master citizen number
- nacionalni identifikacijski broj
- national identification number
- oib#
- oib
- osobna iskaznica
- osobni id
- osobni identifikacijski broj
- personal identification number
- porezni broj
- porezni identifikacijski broj
- tax id
- tax identification no
- tax identification number
- tax no#
- tax no
- tax number
- tax registration number
- taxid#
- taxidno#

- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id
- tin no
- tin#

Croatia physical addresses

Article • 08/19/2024

This unbundled named entity detects patterns related to physical address from Croatia. It's also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Cyprus drivers license number

Article • 08/19/2024

Format

12 digits without spaces and delimiters

Pattern

12 digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_cyprus_eu_driver's_license_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_driver's_license_number` or `Keywords_cyprus_eu_driver's_license_number` is found.

XML

```
      <!-- Cyprus Driver's License Number -->
      <Entity id="356fa104-f9ac-4aff-a0e4-2e6e65ea06c4"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_cyprus_eu_driver's_license_number" />
```

```
          <Any minMatches="1">
            <Match idRef="Keywords_eu_driver's_license_number" />
            <Match idRef="Keywords_cyprus_eu_driver's_license_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences

- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#

- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses

- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_cyprus_eu_driver's_license_number

- άδεια οδήγησης
- αριθμό άδειας οδήγησης
- άδειες οδήγησης

Cyprus identity card

Article • 08/19/2024

Format

10 digits without spaces and delimiters

Pattern

10 digits

Checksum

Not applicable

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_cyprus_eu_national_id_card` finds content that matches the pattern.
- A keyword from `Keywords_cyprus_eu_national_id_card` is found.

XML

```
      <!-- Cyprus Identity Card -->
      <Entity id="3ba8afe5-7a6c-4929-8247-0001b6878438"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_cyprus_eu_national_id_card" />
          <Match idRef="Keywords_cyprus_eu_national_id_card" />
```

```
        </Pattern>
      </Entity>
```

Keywords

Keywords_cyprus_eu_national_id_card

- id card number
- identity card number
- kimlik karti
- national identification number
- personal id number
- ταυτοτηταϲ

Cyprus passport number

Article • 08/19/2024

Format

one letter followed by 6-8 digits with no spaces or delimiters

Pattern

one letter followed by six to eight digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_cyprus_eu_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_cyprus_eu_passport_number` is found.
- The regular expression `Regex_cyprus_eu_passport_date` finds date in the format DD/MM/YYYY or a keyword from `Keywords_cyprus_eu_passport_date` is found

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_cyprus_eu_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_cyprus_eu_passport_number` is found.

XML

```
      <!-- Cyprus Passport Number -->
      <Entity id="9193e2e8-7f8c-43c1-a274-ac40d651936f"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Regex_cyprus_eu_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_cyprus_eu_passport_number" />
          </Any>
          <Any minMatches="1">
            <Match idRef="Regex_cyprus_eu_passport_date" />
            <Match idRef="Keywords_cyprus_eu_passport_date" />
          </Any>
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_cyprus_eu_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_cyprus_eu_passport_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_passport_number_common

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers

Keywords_cyprus_eu_passport_number

- αριθμό διαβατηρίου
- pasaportu
- Αριθμός Διαβατηρίου
- κυπριακό διαβατήριο
- διαβατήριο#
- διαβατήριο
- αριθμός διαβατηρίου
- Pasaport Kimliği
- pasaport numarası
- Pasaport no.
- Αρ. Διαβατηρίου

Keywords_cyprus_eu_passport_date

- expires on
- issued on

Cyprus physical addresses

Article • 08/19/2024

This unbundled named entity detects patterns related to physical address from Cyprus. It's also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Cyprus tax identification number

Article • 08/19/2024

Format

eight digits and one letter in the specified pattern

Pattern

eight digits and one letter:

- a "0" or "9"
- seven digits
- one letter (not case-sensitive)

Checksum

Not applicable

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_cyprus_eu_tax_file_number` finds content that matches the pattern.
- A keyword from `Keywords_cyprus_eu_tax_file_number` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_cyprus_eu_tax_file_number` finds content that matches the pattern.

XML

```
      <!-- Cyprus Tax Identification Number -->
      <Entity id="40e64bd9-55f3-4a09-9bd6-1db18dced9dd"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_cyprus_eu_tax_file_number" />
          <Match idRef="Keywords_cyprus_eu_tax_file_number" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_cyprus_eu_tax_file_number" />
        </Pattern>
      </Entity>
```

Keywords

Keywords_cyprus_eu_tax_file_number

- tax id
- tax identification code
- tax identification no
- tax identification number
- tax no#
- tax no
- tax number
- tax registration number
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tic#
- tic
- tin id
- tin no
- tin#
- vergi kimlik kodu
- vergi kimlik numarası

- αριθμός φορολογικού μητρώου
- κωδικός φορολογικού μητρώου
- φορολογική ταυτότητα
- φορολογικού κωδικού

Czech drivers license number

Article • 08/19/2024

Format

two letters followed by six digits

Pattern

eight letters and digits:

- letter 'E' (not case-sensitive)
- a letter
- a space (optional)
- six digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_czech_republic_eu_driver's_license_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_driver's_license_number` or `Keywords_czech_republic_eu_driver's_license_number` is found.

```
XML
```

```
      <Entity id="86b40d3b-d8ea-4c36-aab0-ef9416a6769c"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_czech_republic_eu_driver's_license_number"
/>
          <Any minMatches="1">
            <Match idRef="Keywords_eu_driver's_license_number" />
            <Match
idRef="Keywords_czech_republic_eu_driver's_license_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences

- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#

- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences

- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses
- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_czech_republic_eu_driver's_license_number

- řidičský prúkaz
- řidičské průkazy
- číslo řidičského průkazu
- čísla řidičských průkazů

Czech passport number

Article • 08/19/2024

Format

eight digits without spaces or delimiters

Pattern

eight digits without spaces or delimiters

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_czech_republic_eu_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_czech_republic_eu_passport_number` is found.
- The regular expression `Regex_eu_passport_date1` finds date in the format DD.MM.YYYY or a keyword from `Keywords_eu_passport_date` is found

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_czech_republic_eu_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_czech_republic_eu_passport_number` is found.

XML

```
      <!-- Czech Republic Passport Number -->
      <Entity id="7bcd8ce8-5e92-4bbe-bc92-fa669f0369fa"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Regex_czech_republic_eu_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_czech_republic_eu_passport_number" />
          </Any>
          <Any minMatches="1">
            <Match idRef="Regex_eu_passport_date1" />
            <Match idRef="Keywords_eu_passport_date" />
          </Any>
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_czech_republic_eu_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_czech_republic_eu_passport_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_passport_number_common

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers

Keywords_czech_republic_eu_passport_number

- cestovní pas
- číslo pasu
- cestovní pasu
- passeport no
- čísla pasu

Keywords_eu_passport_date

- date of issue
- date of expiry

Czech personal identity number

Article • 08/19/2024

Format

nine digits with optional forward slash (old format)

10 digits with optional forward slash (new format)

Pattern

nine digits (old format):

- six digits that represent date of birth
- an optional forward slash
- three digits

10 digits (new format):

- six digits that represent date of birth
- an optional forward slash
- four digits where last digit is a check digit

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_czech_id_card` finds content that matches the pattern.
- A keyword from `Keyword_czech_id_card` is found.
- The checksum passes.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_czech_id_card_new_format` finds content that matches the pattern.
- The checksum passes.

XML

```
<!-- Czech Personal Identity Number -->
      <!-- Czech Personal Identity Number -->
      <Entity id="60c0725a-4eb6-455b-9dda-05d8a7396497"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_czech_id_card" />
          <Match idRef="Keyword_czech_id_card" />
        </Pattern>
        <Version minEngineVersion="15.20.3000.000">
          <Pattern confidenceLevel="75">
            <IdMatch idRef="Func_czech_id_card_new_format" />
          </Pattern>
        </Version>
      </Entity>
```

Keywords

Keyword_czech_id_card

- birth number
- czech republic id
- czechidno#
- daňové číslo
- identifikační číslo
- identity no
- identity number
- identityno#
- identityno
- insurance number
- national identification number
- nationalnumber#

- national number
- osobní číslo
- personalidnumber#
- personal id number
- personal identification number
- personal number
- pid#
- pid
- pojištění číslo
- rč
- rodne cislo
- rodné číslo
- ssn
- ssn#
- social security number
- tax id
- tax identification no
- tax identification number
- tax no#
- tax no
- tax number
- tax registration number
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id
- tin no
- tin#
- unique identification number

Czech Republic physical addresses

Article • 08/19/2024

This unbundled named entity detects patterns related to physical address from the Czech Republic. It's also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Denmark drivers license number

Article • 08/19/2024

Format

eight digits without spaces and delimiters

Pattern

eight digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_denmark_eu_driver's_license_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_driver's_license_number` or `Keywords_denmark_eu_driver's_license_number` is found.

XML

```
      <!-- Denmark Driver's License Number -->
      <Entity id="98a95812-6203-451a-a220-d39870ebef0e"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_denmark_eu_driver's_license_number" />
```

```
          <Any minMatches="1">
            <Match idRef="Keywords_eu_driver's_license_number" />
            <Match idRef="Keywords_denmark_eu_driver's_license_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences

- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#

- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses

- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_denmark_eu_driver's_license_number

- kørekort
- kørekortnummer

Denmark passport number

Article • 08/19/2024

Format

nine digits without spaces and delimiters

Pattern

nine digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_denmark_eu_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_denmark_eu_passport_number` is found.
- The regular expression `Regex_eu_passport_date2` finds date in the format DD MM YY or a keyword from `Keywords_eu_passport_date` is found

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_denmark_eu_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_denmark_eu_passport_number` is found.

XML

```
      <!-- Denmark Passport Number -->
      <Entity id="25e8c47e-e6fe-4884-a211-74898f8c0196"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Regex_denmark_eu_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_denmark_eu_passport_number" />
          </Any>
          <Any minMatches="1">
            <Match idRef="Regex_eu_passport_date2" />
            <Match idRef="Keywords_eu_passport_date" />
          </Any>
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_denmark_eu_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_denmark_eu_passport_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_passport_number_common

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers

Keywords_denmark_eu_passport_number

- pasnummer
- Passeport n°
- pasnumre

Keywords_eu_passport_date

- date of issue
- date of expiry

Denmark personal identification number

Article • 08/19/2024

Format

10 digits containing a hyphen

Pattern

10 digits:

- six digits in the format DDMMYY, which are the date of birth
- an optional space or hyphen
- four digits where the final digit is a check digit

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Func_denmark_eu_tax_file_number` finds content that matches the pattern.
- A keyword from `Keyword_denmark_id` is found.
- The checksum passes.

A DLP policy has low confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Func_denmark_eu_tax_file_number` finds content that matches the pattern.
- The checksum passes.

XML

```
<!-- Denmark Personal Identification Number -->
    <!-- Denmark Personal Identification Number -->
      <Entity id="6c4f2fef-56e1-4c00-8093-88d7a01cf460"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_denmark_eu_tax_file_number" />
          <Match idRef="Keyword_denmark_id" />
        </Pattern>
        <Pattern confidenceLevel="65">
          <IdMatch idRef="Func_denmark_eu_tax_file_number" />
        </Pattern>
      </Entity>
```

Keywords

Keyword_denmark_id

- centrale personregister
- civilt registreringssystem
- cpr
- cpr#
- gesundheitskarte nummer
- gesundheitsversicherungkarte nummer
- health card
- health insurance card number
- health insurance number
- identification number
- identifikationsnummer
- identifikationsnummer#
- identity number
- krankenkassennummer
- nationalid#
- nationalnumber#
- national number

- personalidnumber#
- personalidentityno#
- personal id number
- personnummer
- personnummer#
- reisekrankenversicherungskartenummer
- rejsesygesikringskort
- ssn
- ssn#
- skat id
- skat kode
- skat nummer
- skattenummer
- social security number
- sundhedsforsikringskort
- sundhedsforsikringsnummer
- sundhedskort
- sundhedskortnummer
- sygesikring
- sygesikringkortnummer
- tax code
- travel health insurance card
- uniqueidentityno#
- tax number
- tax registration number
- tax id
- tax identification number
- taxid#
- taxnumber#
- tax no
- taxno#
- taxnumber
- tax identification no
- tin#
- taxidno#
- taxidnumber#
- tax no#
- tin id
- tin no
- cpr.nr
- cprnr

- cprnummer
- personnr
- personregister
- sygesikringsbevis
- sygesikringsbevisnr
- sygesikringsbevisnummer
- sygesikringskort
- sygesikringskortnr
- sygesikringskortnummer
- sygesikringsnr
- sygesikringsnummer

Denmark physical addresses

Article • 08/19/2024

This unbundled named entity detects patterns related to physical address from Denmark. It's also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Diseases

Article • 08/19/2024

This unbundled named entity detects text that matches disease names, such as *diabetes*. It supports English terms only. It's also included in the All medical terms and conditions bundled named entity SIT.

Confidence level

High

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Drug Enforcement Agency (DEA) number

Article • 08/19/2024

Format

two letters followed by seven digits

Pattern

Pattern must include all of the following:

- one letter (not case-sensitive) from this set of possible letters: A/B/F/G/M/P/R, which is a registrant code
- one letter (not case-sensitive), which is the first letter of the registrant's last name or digit '9'
- seven digits, the last of which is the check digit

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_dea_number` finds content that matches the pattern.
- A keyword from `Keyword_dea_number` is found
- The checksum passes.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_dea_number` finds content that matches the pattern.
- The checksum passes.

XML

```
    <!-- DEA Number -->
    <Entity id="9a5445ad-406e-43eb-8bd7-cac17ab6d0e4"
patternsProximity="300" recommendedConfidence="85">
      <Pattern confidenceLevel="75">
        <IdMatch idRef="Func_dea_number" />
      </Pattern>
      <Version minEngineVersion="15.20.1207.000"
maxEngineVersion="15.20.3134.000">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_dea_number" />
        </Pattern>
      </Version>
      <Version minEngineVersion="15.20.3135.000">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_dea_number" />
          <Match idRef="Keyword_dea_number" />
        </Pattern>
      </Version>
    </Entity>
```

Keywords

Keyword_dea_number

- dea
- dea#
- drug enforcement administration
- drug enforcement agency

Ecuador unique identification number

09/26/2025

Format

Ten digits

Pattern

- two digits in the range of 01 to 30 or can be either 90 or 99
- followed by eight digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regex `Regex_Ecuador_Unique_Identification_Number` finds content that matches the pattern.
- A keyword from `Keyword_Ecuador_Unique_Identification_Number` is found.
- The checksum passes.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regex `Regex_Ecuador_Unique_Identification_Number` finds content that matches the pattern.

- The checksum passes.

XML

```
    <!-- Ecuador Unique Identification Number -->
    <Entity id="bc99c8a0-95ea-410e-9b7c-05e51b02e469" patternsProximity="300"
recommendedConfidence="85" relaxProximity="true">
      <Pattern confidenceLevel="85">
        <IdMatch idRef="Regex_Ecuador_Unique_Identification_Number" />
        <Match idRef="Keyword_Ecuador_Unique_Identification_Number" />
      </Pattern>
      <Pattern confidenceLevel="75">
        <IdMatch idRef="Regex_Ecuador_Unique_Identification_Number" />
      </Pattern>
    </Entity>
```

Keywords

Keyword_Ecuador_Unique_Identification_Number

- Identity card
- Id card
- idnumber#
- id number
- identity no
- identitynumber#
- identity number
- National insurance
- national identification number
- national id
- NI Number
- NI No.
- NI #
- NI#
- insurance#
- insurancenumber
- nationalinsurance#
- nationalinsurancenumber
- national insurance number
- Seguro
- número de seguro Nacional
- numero de seguro

- numero de identidad
- Identificación nacional

Estonia drivers license number

Article • 08/19/2024

Format

two letters followed by six digits

Pattern

two letters and six digits:

- the letters "ET" (not case-sensitive)
- six digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_estonia_eu_driver's_license_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_driver's_license_number` or `Keywords_estonia_eu_driver's_license_number` is found.

```
XML
```

```
      <!-- Estonia Driver's License Number -->
      <Entity id="51da8171-da70-4cc1-9d65-055a59ca4f83"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_estonia_eu_driver's_license_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_driver's_license_number" />
            <Match idRef="Keywords_estonia_eu_driver's_license_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic

- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#

- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen

- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses
- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_estonia_eu_driver's_license_number

- permis de conduire
- juhilubade numbrid
- juhiloa number
- juhiluba

Estonia passport number

Article • 08/19/2024

Format

one letter followed by seven digits with no spaces or delimiters

Pattern

one letter followed by seven digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_estonia_eu_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_estonia_eu_passport_number` is found.
- The regular expression `Regex_eu_passport_date1` finds date in the format DD.MM.YYYY or a keyword from `Keywords_eu_passport_date` is found

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_estonia_eu_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_estonia_eu_passport_number` is found.

XML

```
      <!-- Estonia Passport Number -->
      <Entity id="61f7073a-509e-425b-a754-bc01bb5d5b8c"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Regex_estonia_eu_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_estonia_eu_passport_number" />
          </Any>
          <Any minMatches="1">
            <Match idRef="Regex_eu_passport_date1" />
            <Match idRef="Keywords_eu_passport_date" />
          </Any>
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_estonia_eu_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_estonia_eu_passport_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_passport_number_common

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers

Keywords_estonia_eu_passport_number

eesti kodaniku pass passi number passinumbrid document number document no dokumendi nr

Keywords_eu_passport_date

- date of issue
- date of expiry

Estonia personal identification code

Article • 08/19/2024

Format

11 digits without spaces and delimiters

Pattern

11 digits:

- one digit that corresponds to sex and century of birth (odd number male, even number female; 1-2: 19th century; 3-4: 20th century; 5-6: 21st century)
- six digits that correspond to date of birth (YYMMDD)
- three digits that correspond to a serial number separating persons born on the same date
- one check digit

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_estonia_eu_national_id_card` finds content that matches the pattern.
- A keyword from `Keywords_estonia_eu_national_id_card` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_estonia_eu_national_id_card` finds content that matches the pattern.

XML

```
      <!-- Estonia Personal Identification Code -->
      <Entity id="bfb26de6-dad5-4d48-ab72-4789cdd0654c"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_estonia_eu_national_id_card" />
          <Match idRef="Keywords_estonia_eu_national_id_card" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_estonia_eu_national_id_card" />
          <Any minMatches="0" maxMatches="0">
            <Match idRef="Keywords_estonia_eu_telephone_number" />
            <Match idRef="Keywords_estonia_eu_mobile_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_estonia_eu_national_id_card

- id-kaart
- ik
- isikukood#
- isikukood
- maksu id
- maksukohustuslase identifitseerimisnumber
- maksunumber
- national identification number
- national number
- personal code
- personal id number
- personal identification code
- personal identification number
- personalidnumber#
- tax id
- tax identification no

- tax identification number
- tax no#
- tax no
- tax number
- tax registration number
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id
- tin no
- tin#

Estonia physical addresses

Article • 08/19/2024

This unbundled named entity detects patterns related to physical address from Estonia. It's also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

EU debit card number

Article • 08/19/2024

Format

16 to 19 digits

Pattern

16 to 19 formatted or unformatted digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_eu_debit_card` finds content that matches the pattern.
- At least one of the following cases is true:
  - A keyword from `Keyword_eu_debit_card` is found.
  - A keyword from `Keyword_card_terms_dict` is found.
  - A keyword from `Keyword_card_security_terms_dict` is found.
  - A keyword from `Keyword_card_expiration_terms_dict` is found.
  - The function `Func_expiration_date` finds a date in the right date format.
- The checksum passes.

A DLP policy has low confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function Func_eu_debit_card finds content that matches the pattern.
- The checksum passes.

XML

```
    <!-- EU Debit Card Number -->
    <Entity id="0e9b3178-9678-47dd-a509-37222ca96b42"
patternsProximity="300" recommendedConfidence="85" relaxProximity="true">
      <Pattern confidenceLevel="85">
        <IdMatch idRef="Func_eu_debit_card" />
        <Any minMatches="1">
          <Match idRef="Keyword_eu_debit_card" />
          <Match idRef="Keyword_card_terms_dict" />
          <Match idRef="Keyword_card_security_terms_dict" />
          <Match idRef="Keyword_card_expiration_terms_dict" />
          <Match idRef="Func_expiration_date" />
        </Any>
      </Pattern>

      <Pattern confidenceLevel="65">
        <IdMatch idRef="Func_eu_debit_card" />
      </Pattern>
    </Entity>
```

Keywords

Keyword_eu_debit_card

- account number
- card number
- card no.
- security number
- cc#

Keyword_card_terms_dict

- acct nbr
- acct num
- acct no
- american express
- americanexpress

- americano espresso
- amex
- atm card
- atm cards
- atm kaart
- atmcard
- atmcards
- atmkaart
- atmkaarten
- bancontact
- bank card
- bankkaart
- card holder
- card holders
- card num
- card number
- card numbers
- card type
- cardano numerico
- cardholder
- cardholders
- cardnumber
- cardnumbers
- carta bianca
- carta credito
- carta di credito
- cartao de credito
- cartao de crédito
- cartao de debito
- cartao de débito
- carte bancaire
- carte blanche
- carte bleue
- carte de credit
- carte de crédit
- carte di credito
- carteblanche
- cartão de credito
- cartão de crédito
- cartão de debito
- cartão de débito

- cb
- ccn
- check card
- check cards
- checkcard
- checkcards
- chequekaart
- cirrus
- cirrus-edc-maestro
- controlekaart
- controlekaarten
- credit card
- credit cards
- creditcard
- creditcards
- debetkaart
- debetkaarten
- debit card
- debit cards
- debitcard
- debitcards
- debito automatico
- diners club
- dinersclub
- discover
- discover card
- discover cards
- discovercard
- discovercards
- débito automático
- edc
- eigentümername
- european debit card
- hoofdkaart
- hoofdkaarten
- in viaggio
- japanese card bureau
- japanse kaartdienst
- jcb
- kaart
- kaart num

- kaartaantal
- kaartaantallen
- kaarthouder
- kaarthouders
- karte
- karteninhaber
- karteninhabers
- kartennr
- kartennummer
- kreditkarte
- kreditkarten-nummer
- kreditkarteninhaber
- kreditkarteninstitut
- kreditkartennummer
- kreditkartentyp
- maestro
- master card
- master cards
- mastercard
- mastercards
- mc
- mister cash
- n carta
- carta
- no de tarjeta
- no do cartao
- no do cartão
- no. de tarjeta
- no. do cartao
- no. do cartão
- nr carta
- nr. carta
- numeri di scheda
- numero carta
- numero de cartao
- numero de carte
- numero de cartão
- numero de tarjeta
- numero della carta
- numero di carta
- numero di scheda

- numero do cartao
- numero do cartão
- numéro de carte
- nº carta
- nº de carte
- nº de la carte
- nº de tarjeta
- nº do cartao
- nº do cartão
- nº. do cartão
- número de cartao
- número de cartão
- número de tarjeta
- número do cartao
- scheda dell'assegno
- scheda dell'atmosfera
- scheda dell'atmosfera
- scheda della banca
- scheda di controllo
- scheda di debito
- scheda matrice
- schede dell'atmosfera
- schede di controllo
- schede di debito
- schede matrici
- scoprono la scheda
- scoprono le schede
- solo
- supporti di scheda
- supporto di scheda
- switch
- tarjeta atm
- tarjeta credito
- tarjeta de atm
- tarjeta de credito
- tarjeta de debito
- tarjeta debito
- tarjeta no
- tarjetahabiente
- tipo della scheda
- ufficio giapponese della

- scheda
- v pay
- v-pay
- visa
- visa plus
- visa electron
- visto
- visum
- vpay

Keyword_card_security_terms_dict

- card identification number
- card verification
- cardi la verifica
- cid
- cod seg
- cod seguranca
- cod segurança
- cod sicurezza
- cod. seg
- cod. seguranca
- cod. segurança
- cod. sicurezza
- codice di sicurezza
- codice di verifica
- codigo
- codigo de seguranca
- codigo de segurança
- crittogramma
- cryptogram
- cryptogramme
- cv2
- cvc
- cvc2
- cvn
- cvv
- cvv2
- cód seguranca
- cód segurança

- cód. seguranca
- cód. segurança
- código
- código de seguranca
- código de segurança
- de kaart controle
- geeft nr uit
- issue no
- issue number
- kaartidentificatienummer
- kreditkartenprufnummer
- kreditkartenprüfnummer
- kwestieaantal
- no. dell'edizione
- no. di sicurezza
- numero de securite
- numero de verificacao
- numero dell'edizione
- numero di identificazione della
- scheda
- numero di sicurezza
- numero van veiligheid
- numéro de sécurité
- nº autorizzazione
- número de verificação
- perno il blocco
- pin block
- prufziffer
- prüfziffer
- security code
- security no
- security number
- sicherheits kode
- sicherheitscode
- sicherheitsnummer
- speldblok
- veiligheid nr
- veiligheidsaantal
- veiligheidscode
- veiligheidsnummer
- verfalldatum

Keyword_card_expiration_terms_dict

- ablauf
- data de expiracao
- data de expiração
- data del exp
- data di exp
- data di scadenza
- data em que expira
- data scad
- data scadenza
- date de validité
- datum afloop
- datum van exp
- de afloop
- espira
- espira
- exp date
- exp datum
- expiration
- expire
- expires
- expiry
- fecha de expiracion
- fecha de venc
- gultig bis
- gultigkeitsdatum
- gültig bis
- gültigkeitsdatum
- la scadenza
- scadenza
- valable
- validade
- valido hasta
- valor
- venc
- vencimento
- vencimiento
- verloopt
- vervaldag
- vervaldatum

- vto
- válido hasta

EU drivers license number

Article • 09/13/2023

These entities are in the EU Driver's License Number and are sensitive information types.

- Austria
- Belgium
- Bulgaria
- Croatia
- Cyprus
- Czech
- Denmark
- Estonia
- Finland
- France
- Germany
- Greece
- Hungary
- Ireland
- Italy
- Latvia
- Lithuania
- Luxemburg
- Malta
- Netherlands
- Poland
- Portugal
- Romania
- Slovakia
- Slovenia
- Spain
- Sweden
- U.K.

EU national identification number

Article • 09/13/2023

These entities are in the EU National Identification Number and are sensitive information types.

- Austria
- Belgium
- Bulgaria
- Croatia
- Cyprus
- Czech
- Denmark
- Estonia
- Finland
- France
- Germany
- Greece
- Hungary
- Ireland
- Italy
- Latvia
- Lithuania
- Luxemburg
- Malta
- Netherlands
- Portugal
- Romania
- Slovakia
- Slovenia
- Spain
- U.K.

EU passport number

Article • 10/30/2023

These entities are in the EU passport number and are sensitive information types. These entities are in the EU passport number bundle.

- Austria
- Belgium
- Bulgaria
- Croatia
- Cyprus
- Czech
- Denmark
- Estonia
- Finland
- France
- Germany
- Greece
- Hungary
- Ireland
- Italy
- Latvia
- Lithuania
- Malta
- Poland
- Portugal
- Romania
- Slovakia
- Slovenia
- Spain
- Sweden
- U.S./U.K. passport number

EU social security number or equivalent identification

Article • 03/18/2025

These entities are in the EU Social Security Number or equivalent identification and are sensitive information types.

- Austria
- Belgium
- Croatia
- Czech
- Denmark
- Finland
- France
- Germany
- Greece
- Hungary
- Spain
- Sweden

EU Tax identification number

Article • 09/13/2023

These entities are in the EU Tax identification number sensitive information type.

- Austria
- Belgium
- Bulgaria
- Croatia
- Cyprus
- Czech
- Denmark
- Estonia
- Finland
- France
- Germany
- Greece
- Hungary
- Ireland
- Italy
- Latvia
- Lithuania
- Luxemburg
- Malta
- Netherlands
- Poland
- Portugal
- Romania
- Slovakia
- Slovenia
- Spain
- Sweden
- U.K.

Finland drivers license number

Article • 08/19/2024

Format

10 digits containing a hyphen

Pattern

10 digits containing a hyphen:

- six digits
- a hyphen
- three digits
- a digit or letter

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_finland_eu_driver's_license_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_driver's_license_number` or `Keywords_finland_eu_driver's_license_number` is found.

XML

```
      <!-- Finland Driver's License Number -->
      <Entity id="bb3b27a3-79bd-4ac4-81a7-f9fca3c7d1a7"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_finland_eu_driver's_license_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_driver's_license_number" />
            <Match idRef="Keywords_finland_eu_driver's_license_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic

- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#

- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen

- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses
- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_finland_eu_driver's_license_number

- ajokortti
- permis de conduire
- ajokortin numero
- kuljettaja lic.
- körkort
- körkortnummer
- förare lic.
- ajokortit
- ajokortin numerot

Finland european health insurance number

Article • 08/19/2024

Format

20-digit number

Pattern

20-digit number:

- 10 digits - 8024680246
- an optional space or hyphen
- 10 digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regex `Regex_Finland_European_Health_Insurance_Number` finds content that matches the pattern.
- A keyword from `Keyword_Finland_European_Health_Insurance_Number` is found.

```
XML
```

```
      <!-- Finland European Health Insurance Number -->
      <Entity id="60f75aed-81bf-4625-89b0-0846b9248ee7"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_Finland_European_Health_Insurance_Number"/>
          <Match idRef="Keyword_Finland_European_Health_Insurance_Number"/>
        </Pattern>
      </Entity>
```

Keywords

Keyword_finland_european_health_insurance_number

- ehic#
- ehic
- finlandehicnumber#
- finska sjukförsäkringskort
- health card
- health insurance card
- health insurance number
- hälsokort
- sairaanhoitokortin
- sairausvakuutuskortti
- sairausvakuutusnumero
- sjukförsäkring nummer
- sjukförsäkringskort
- suomen sairausvakuutuskortti
- terveyskortti

Finland national ID

Article • 08/19/2024

Format

six digits plus a character indicating a century plus three digits plus a check digit

Pattern

Pattern must include all of these patterns:

- six digits in the format DDMMYY, which are a date of birth
- century marker (either '-', '+' or 'a')
- three-digit personal identification number
- a digit or letter (case insensitive) which is a check digit

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- the function `Func_finnish_national_id` finds content that matches the pattern
- a keyword from `Keyword_finnish_national_id` is found
- the checksum passes

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- the function `Func_finnish_national_id` finds content that matches the pattern
- the checksum passes

XML

```
      <!-- Finnish National ID-->
      <Entity id="338FD995-4CB5-4F87-AD35-79BD1DD926C1"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_finnish_national_id" />
          <Match idRef="Keyword_finnish_national_id" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_finnish_national_id" />
        </Pattern>
      </Entity>
```

Keywords

- ainutlaatuinen henkilökohtainen tunnus
- henkilökohtainen tunnus
- henkilötunnus
- henkilötunnusnumero#
- henkilötunnusnumero
- hetu
- id no
- id number
- identification number
- identiteetti numero
- identity number
- idnumber
- kansallinen henkilötunnus
- kansallisen henkilökortin
- national id card
- national id no.
- personal id
- personal identity code
- personalidnumber#
- personbeteckning
- personnummer
- social security number
- sosiaaliturvatunnus
- tax id

- tax identification no
- tax identification number
- tax no#
- tax no
- tax number
- tax registration number
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id
- tin no
- tin#
- tunnistenumero
- tunnus numero
- tunnusluku
- tunnusnumero
- verokortti
- veronumero
- verotunniste
- verotunnus

Finland passport number

Article • 08/19/2024

This entity is available in the EU Passport Number sensitive information type and is available as a stand-alone sensitive information type entity.

Format

combination of nine letters and digits

Pattern

combination of nine letters and digits:

- two letters (not case-sensitive)
- seven digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_finland_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keyword_finland_passport_number` is found.

- The regular expression `Regex_eu_passport_date1` finds date in the format DD.MM.YYYY or a keyword from `Keywords_eu_passport_date` is found

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_finland_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keyword_finland_passport_number` is found.

XML

```
      <!-- Finland Passport Number -->
      <Entity id="d1685ac3-1d3a-40f8-8198-32ef5669c7a5"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Regex_finland_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keyword_finland_passport_number" />
          </Any>
          <Any minMatches="1">
            <Match idRef="Regex_eu_passport_date1" />
            <Match idRef="Keywords_eu_passport_date" />
          </Any>
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_finland_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keyword_finland_passport_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_passport_number

- passport#
- passport #
- passportid
- passports
- passportno
- passport no

- passportnumber
- passport number
- passportnumbers
- passport numbers

Keyword_finland_passport_number

- suomalainen passi
- passin numero
- passin numero.#
- passin numero#
- passin numero.
- passi#
- passi number

Keywords_eu_passport_date

- date of issue
- date of expiry

Finland physical addresses

Article • 08/19/2024

This unbundled named entity detects patterns related to physical address from Finland. It's also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

France drivers license number

Article • 08/19/2024

This entity is available in the EU Driver's License Number sensitive information type and is available as a stand-alone sensitive information type entity.

Format

12 digits

Pattern

12 digits with validation to discount similar patterns such as French telephone numbers

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- the function `Func_french_drivers_license` finds content that matches the pattern.
- a keyword from `Keyword_french_drivers_license` is found.

XML

```
    <!-- France Driver's License Number -->
    <Entity id="18e55a36-a01b-4b0f-943d-dc10282a1824"
patternsProximity="300" recommendedConfidence="75">
      <Pattern confidenceLevel="75">
```

```
        <IdMatch idRef="Func_french_drivers_license" />
        <Match idRef="Keyword_french_drivers_license" />
      </Pattern>
    </Entity>
```

Keywords

Keyword_french_drivers_license

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences
- driver' lic

- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#
- drivers licence#

- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses
- driving licence

- driving licences
- driving permit
- dl no
- dlno
- dl number
- permis de conduire
- licence number
- license number
- licence numbers
- license numbers
- numéros de licence

France health insurance number

Article • 08/19/2024

Format

21-digit number

Pattern

21-digit number:

- 10 digits
- an optional space
- 10 digits
- an optional space
- a digit

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- the regex `Regex_France_Health_Insurance_Number` finds content that matches the pattern.
- a keyword from `Keyword_France_Health_Insurance_Number` is found.

```
XML
```

```
      <!-- France Health Insurance Number -->
      <Entity id="9bc2069e-76df-4ff9-ac02-2f519469e236"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_France_Health_Insurance_Number"/>
          <Match idRef="Keyword_France_Health_Insurance_Number"/>
        </Pattern>
      </Entity>
```

Keywords

Keyword_France_health_insurance_number

- insurance card
- carte vitale
- carte d'assuré social

France national ID card (CNI)

Article • 08/19/2024

Format

Twelve digits

or

Twelve letters and digits

or

Nine letters and digits

Pattern

Twelve digits

or

Combination of twelve digits and letters

or

Combination of nine digits and letters

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has low confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_france_cni` finds content that matches the pattern.
- A keyword from `Keywords_france_eu_national_id_card` is found.

XML

```
    <!-- France CNI -->
    <Entity id="f741ac74-1bc0-4665-b69b-f0c7f927c0c4"
patternsProximity="300" recommendedConfidence="65">
      <Pattern confidenceLevel="65">
        <IdMatch idRef="Regex_france_cni" />
        <Match idRef="Keywords_france_eu_national_id_card" />
      </Pattern>
    </Entity>
```

Keywords

Keywords_france_eu_national_id_card

- card number
- carte nationale d'identité
- carte nationale d'idenite no
- cni#
- cni
- compte bancaire
- national identification number
- national identity
- nationalidno#
- numéro d'assurance maladie
- numéro de carte vitale

France passport number

Article • 08/19/2024

This entity is available in the EU Passport Number sensitive information type. It's also available as a stand-alone sensitive information type entity.

Format

nine digits and letters

Pattern

nine digits and letters:

- two digits
- two letters (not case-sensitive)
- five digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_fr_passport` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_france_eu_passport_number` is found.

- The regular expression `Regex_eu_passport_date3` finds date in the format DD MM YYYY or a keyword from `Keywords_eu_passport_date` is found

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_fr_passport` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_france_eu_passport_number` is found.

XML

```
    <!-- France Passport Number -->
    <Entity id="3008b884-8c8c-4cd8-a289-99f34fc7ff5d"
patternsProximity="300" recommendedConfidence="75">
      <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_fr_passport" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_france_eu_passport_number" />
          </Any>
          <Any minMatches="1">
            <Match idRef="Regex_eu_passport_date3" />
            <Match idRef="Keywords_eu_passport_date" />
          </Any>
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_fr_passport" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_france_eu_passport_number" />
          </Any>
        </Pattern>
    </Entity>
```

Keywords

Keywords_eu_passport_number

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber

- passport number
- passportnumbers
- passport numbers

Keywords_france_eu_passport_number

- numéro de passeport
- passeport n °
- passeport non
- passeport #
- passeport#
- passeportnon
- passeportn °
- passeport français
- passeport livre
- passeport carte
- numéro passeport
- passeport n°
- n° du passeport
- n° passeport

Keywords_eu_passport_date

- date of issue
- date of expiry

France physical addresses

Article • 08/19/2024

This unbundled named entity detects patterns related to physical address from France. It's also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

France social security number (INSEE)

Article • 08/19/2024

Format

15 digits

Pattern

Must match one of two patterns:

- 13 digits followed by a space followed by two digits

  or

- 15 consecutive digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_french_insee` finds content that matches the pattern.
- A keyword from `Keyword_fr_insee` is found.
- The checksum passes.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_french_insee` or `Func_fr_insee` finds content that matches the pattern.
- The checksum passes.

XML

```
    <!-- France INSEE -->
    <Entity id="71f62b97-efe0-4aa1-aa49-e14de253619d"
patternsProximity="300" recommendedConfidence="75">
      <Pattern confidenceLevel="75">
        <IdMatch idRef="Func_french_insee" />
        <Any minMatches="0" maxMatches="0">
          <Match idRef="Keyword_fr_insee" />
        </Any>
      </Pattern>
      <Pattern confidenceLevel="85">
        <IdMatch idRef="Func_french_insee" />
        <Match idRef="Keyword_fr_insee" />
      </Pattern>
    </Entity>
```

Keywords

Keyword_fr_insee

- code sécu
- d'identité nationale
- insee
- fssn#
- le numéro d'identification nationale
- le code de la sécurité sociale
- national id
- national identification
- no d'identité
- no. d'identité
- numéro d'assurance
- numéro d'identité
- numero d'identite
- numéro de sécu
- numéro de sécurité sociale
- no d'identite
- no. d'identite
- ssn

- ssn#
- sécurité sociale
- securité sociale
- securite sociale
- socialsecuritynumber
- social security number
- social security code
- social insurance number

France tax identification number

Article • 08/19/2024

Format

13 digits

Pattern

13 digits

- One digit that must be 0, 1, 2, or 3
- One digit
- A space (optional)
- Two digits
- A space (optional)
- Three digits
- A space (optional)
- Three digits
- A space (optional)
- Three check digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_france_eu_tax_file_number` finds content that matches the pattern.
- A keyword from `Keywords_france_eu_tax_file_number` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_france_eu_tax_file_number` finds content that matches the pattern.

XML

```
      <!-- France Tax Identification Number (numéro SPI.) -->
      <Entity id="ed59e77e-171d-442c-9ec1-88e2ebcb5b0a"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_france_eu_tax_file_number" />
          <Match idRef="Keywords_france_eu_tax_file_number" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_france_eu_tax_file_number" />
          <Any minMatches="0" maxMatches="0">
            <Match idRef="Keywords_france_eu_telephone_number" />
            <Match idRef="Keywords_france_eu_mobile_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_france_eu_tax_file_number

- numéro d'identification fiscale
- tax id
- tax identification no
- tax identification number
- tax no#
- tax no
- tax number
- tax registration number
- taxid#
- taxidno#
- taxidnumber#

- taxno#
- taxnumber#
- taxnumber
- tin id
- tin no
- tin#

France value added tax number

Article • 08/19/2024

Format

13 character alphanumeric pattern

Pattern

13 character alphanumeric pattern:

- two letters - FR (case insensitive)
- an optional space or hyphen
- two letters or digits
- an optional space, dot, hyphen, or comma
- three digits
- an optional space, dot, hyphen, or comma
- three digits
- an optional space, dot, hyphen, or comma
- three digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_france_value_added_tax_number` finds content that matches the pattern.
- A keyword from `Keywords_france_value_added_tax_number` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_france_value_added_tax_number` finds content that matches the pattern.

XML

```
      <!-- France Value Added Tax Number -->
      <Entity id="949121e6-ad9f-4379-8731-710342baea78"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_france_value_added_tax_number" />
          <Match idRef="Keywords_france_value_added_tax_number" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_france_value_added_tax_number" />
        </Pattern>
      </Entity>
```

Keywords

Keyword_France_value_added_tax_number

- vat number
- vat no
- vat#
- value added tax
- siren identification no numéro d'identification taxe sur valeur ajoutée
- taxe valeur ajoutée
- taxe sur la valeur ajoutée
- n° tva
- numéro de tva
- numéro d'identification siren

General Password

This SIT is also included in the All credentials bundled SIT.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a credential scanning SIT in an endpoint DLP policy. This requirement is specific to the combination of credential scanning SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

Format

Up to 20,000-character combination of letters, digits, and special characters.

or

Sign in credentials used in command lines

or

Plain-text password used in code snippets

or

Plain-text password used in script

or

Plain-text password used in XML configuration

or

A combination of 24 characters consisting of letters, digits, and special characters.

or

A combination of 32 characters consisting of letters and digits.

or

A combination of 32 characters consisting of letters, digits, and special characters.

or

A combination of 44 characters consisting of letters, digits, and special characters.

or

An 88-character combination of letters, digits, and special characters.

Patterns

1. Any combination of up to 20,000 characters consisting of:

    - a-z (case insensitive)
    - 0-9
    - forward slashes (/) or plus signs (+)
    - Up to two equal signs (=)

2. Various command line sign-in credentials formats

3. Various password formats in code snippets

4. Various password formats in script

5. Various password formats in XML

6. Any combination of 22 characters consisting of:

    - a-z (case insensitive)
    - digits, forward slashes, or plus signs
    - ends with two equal signs (=)

7. Any combination of 32 characters consisting of:

    - a-f or A-F (case-sensitive) or 0-9

8. Any combination of 32 characters consisting of:

    - a-z (case insensitive)
    - 0-9
    - forward slashes (/) or plus signs (+)

9. Any combination of 43 characters consisting of:

- a-z (case insensitive)
- 0-9
- forward slashes (/) or plus signs (+)
- ends with an equal sign (=)

10. Any combination of 86 characters consisting of:

    - a-z (case insensitive)
    - 0-9
    - forward slashes (/) or plus signs (+)
    - ends with two equal signs (=)

Credential example

Expand table

Confidence Band	Example
High	`password = D3m0P@sswd!`
Medium	`secret : DemoPasswd!`
Low	`password = demopasswd2`

Checksum

Yes

SITs that have checksums use a unique calculation to check if the information is valid. This means when the **Checksum** value is **Yes**, the service can make a positive detection based on the sensitive data alone. When the **Checksum** value is **No** additional (secondary) elements must also be detected for the service to make a positive detection.

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Description

This SIT is designed to match security information, such as usernames and passwords, that are used in the sign-in process user login process. It uses several primary resources:

- Patterns of Base64 encoded string literal.
- Patterns of Password context in command line.
- Patterns of Password context in code.
- Patterns of Password context in script.
- Patterns of Password context in XML.
- Patterns of Base64 encoded 128-bits symmetric key.
- Patterns of Hex encoded 128-bits symmetric key.
- Patterns of Base64 encoded 192-bits symmetric key.
- Patterns of Base64 encoded 256-bits symmetric key.
- Patterns of Base64 encoded 512-bits symmetric key.
- Patterns of CredentialName, CredentialFeatures, AccountIdentityName, AccountIdentityValue, ResourceType, ResourceName, ID, AccountName.
- Patterns of mockup values, redactions, and placeholders.
- A dictionary of vocabulary words.

The patterns are designed to match actual credentials with reasonable confidence. The patterns don't match credentials formatted as examples. For instance, mockup values, redacted values, and placeholders, like credential type or usage descriptions, that are in the position where an actual secret value should present aren't matched.

Keywords

> **Important**
>
> This SIT detects passwords related keywords in English only.

Keyword_Base64EncodedStringLiteral

- MII

Keyword_PasswordContextInCmdLine

- certutil
- zdbg
- secret

- VSTS_TOKEN
- curl
- PowerShell
- ps1
- -u
- Smc
- AutoLogon
- ldifde
- Rclone
- --env
- SignTool
- winexe
- net

Keyword_PasswordContextInCode

- key
- x509c
- credential
- password
- pw
- securestring

Keyword_PasswordContextInScript

- secret
- password
- pw

Keyword_PasswordContextInXml

- userpass
- password
- pw
- connectionstring
- key
- credential
- token

- sas
- secret

Keyword_SymmetricKey128

- secret
- key
- password
- pw

Keyword_SymmetricKey128Hex

- dapi
- key
- secret
- token
- password
- pw

Keyword_SymmetricKey192

- password
- -p
- azurecr

Keyword_SymmetricKey256

- SharedAccessKey
- AccountKey

Keyword_SymmetricKey512

- SharedAccessKey
- AccountKey

Last updated on 06/09/2026

General Symmetric key

This SIT is also included in the All credentials bundled SIT.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a credential scanning SIT in an endpoint DLP policy. This requirement is specific to the combination of credential scanning SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

Format

A combination of 44 characters consisting of letters, digits, and special characters.

or

A combination of 88 characters consisting of letters, digits, and special characters.

Pattern

A combination of 43 characters:

- a-z (not case-sensitive)
- 0-9
- forward slashes (/) or plus signs (+)
- ends with an equal sign (=)

format example:

```
Cc3Dd4Ee5Ff6Gg7Hh8Ii9Jj0Kk1Ll2Mm34/+Nn=
```

or

A combination of 86 characters:

- a-z (not case-sensitive)
- 0-9

- forward slashes (/) or plus signs (+)
- ends with two equal signs (=)

format example:

`Cc3Dd4Ee5Ff6Gg7Hh8Ii9Jj0Kk1Ll2Mm3Nn4/+==`

Credential example

Expand table

Confidence Band	Example
High	`N/A`
Medium	`N/A`
Low	`key=Cc3Dd4Ee5Ff6Gg7Hh8Ii9Jj0Kk1Ll2Mm34/+Nn=;`

Checksum

No

SITs that have checksums use a unique calculation to check if the information is valid. This means when the **Checksum** value is **Yes**, the service can make a positive detection based on the sensitive data alone. When the **Checksum** value is **No** additional (secondary) elements must also be detected for the service to make a positive detection.

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

This SIT is designed to match the security information that's used in general authentication process.

It uses several primary resources:

- Patterns of Base64 encoded 256-bits symmetric key.
- Patterns of Base64 encoded 512-bits symmetric key.
- Patterns of CredentialName, CredentialFeatures, AccountIdentityName, AccountIdentityValue, ResourceType, ResourceName, Id, AccountName.
- Patterns of mockup values, redactions, and placeholders.
- A dictionary of vocabulary.

The patterns are designed to match actual credentials with reasonable confidence. The patterns don't match credentials formatted as examples. Mockup values, redacted values, and placeholders, like credential type or usage descriptions, in the position where an actual secret value should present won't be matched.

Keywords

Keyword_SymmetricKey256:

- SharedAccessKey
- AccountKey

Keyword_SymmetricKey512:

- SharedAccessKey
- AccountKey

Last updated on 11/18/2025

Generic medication names

Article • 08/19/2024

This unbundled named entity detects names of generic medications, such as *acetaminophen*. It supports English terms only. It's also included in the All medical terms and conditions bundled named entity SIT.

Confidence level

High

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

German Driver's License Number

This sensitive information type entity is included in the EU Driver's License Number sensitive information type. It's also available as a stand-alone sensitive information type entity.

Format

Combination of 11 digits and letters

Pattern

Eleven digits and letters (not case-sensitive):

- a digit or letter
- two digits
- six digits or letters
- a digit
- a digit or letter

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A data loss prevention (DLP) policy has medium confidence that it detects this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_german_drivers_license` finds content that matches the pattern.
- A keyword from `Keyword_german_drivers_license_number` is found.

- The checksum passes.

XML

```
    <!-- German Driver's License Number -->
    <Entity id="91da9335-1edb-45b7-a95f-5fe41a16c63c" patternsProximity="300"
recommendedConfidence="75">
      <Pattern confidenceLevel="75">
        <IdMatch idRef="Func_german_drivers_license" />
        <Match idRef="Keyword_german_drivers_license" />
      </Pattern>
    </Entity>
```

Keywords

Keyword_german_drivers_license_number

- ausstellungsdatum
- ausstellungsort
- ausstellende behöde
- ausstellende behorde
- ausstellende behoerde
- führerschein
- fuhrerschein
- fuehrerschein
- führerscheinnummer
- fuhrerscheinnummer
- fuehrerscheinnummer
- führerschein-
- fuhrerschein-
- fuehrerschein-
- führerscheinnummernr
- fuhrerscheinnummernr
- fuehrerscheinnummernr
- führerscheinnummerklasse
- fuhrerscheinnummerklasse
- fuehrerscheinnummerklasse
- nr-führerschein
- nr-fuhrerschein
- nr-fuehrerschein

- no-führerschein
- no-fuhrerschein
- no-fuehrerschein
- n-führerschein
- n-fuhrerschein
- n-fuehrerschein
- permis de conduire
- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences

- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#

- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses

- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses
- driving licence
- driving licences
- driving permit
- dlno

Last updated on 06/19/2026

Germany identity card number

Article • 01/24/2025

Format

since 1 November 2010: Nine to 11 letters and digits

from 1 April 1987 until 31 October 2010: 10 digits

Pattern

since 1 November 2010: 9 to 11 characters alphanumeric pattern

- one L, M, N, P, R, T, V, W, X, Y (case insensitive)
- eight digits or letters in C, F, G, H, J, K, L, M, N, P, R, T, V, W, X, Y and Z (case insensitive)
- optional check digit
- Optional d/D

from 1 April 1987 until 31 October 2010:

- 10 digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_german_id_card_with_check` finds content that matches the pattern (10 to 11 characters alphanumeric pattern (post 2010)).
- A keyword from `Keyword_germany_id_card` is found.
- The checksum passes.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_germany_id_card` finds content that matches the pattern (9 characters without check digit issued post-2010 or 10 digits pattern issued pre-2010).
- A keyword from `Keyword_germany_id_card` is found.

A DLP policy has low confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_german_id_card_with_check` finds content that matches the pattern (10 to 11 characters alphanumeric pattern (post 2010)).
- The checksum passes.

XML

```
      <!-- Germany Identity Card Number -->
      <Entity id="e577372f-c42e-47a0-9d85-bebed1c237d4"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="75">
         <IdMatch idRef="Regex_germany_id_card" />
         <Match idRef="Keyword_germany_id_card" />
        </Pattern>
        <Version minEngineVersion="15.20.4545.000">
          <Pattern confidenceLevel="85">
           <IdMatch idRef="Func_german_id_card_with_check" />
            <Match idRef="Keyword_germany_id_card" />
          </Pattern>
          <Pattern confidenceLevel="65">
           <IdMatch idRef="Func_german_id_card_with_check" />
          </Pattern>
        </Version>
      </Entity>
```

Keywords

Keyword_germany_id_card

- ausweis
- gpid
- identification
- identifikation
- identifizierungsnummer
- identity card
- identity number
- id-nummer
- personal id
- personalausweis
- persönliche id nummer
- persönliche identifikationsnummer
- persönliche-id-nummer

German Passport Number

Format

9 to 11 characters

Pattern

- one letter in C, F, G, H, J, K (case insensitive)
- eight digits or letters in C, F, G, H, J, K, L, M, N, P, R, T, V, W, X, Y, and Z (case insensitive)
- optional check digit
- Optional d/D

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A data loss prevention (DLP) policy has high confidence that it detects this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_german_passport_checksum` finds content that matches the pattern.
- A keyword from `Keyword_german_passport` or `Keywords_eu_passport_number_common` is found.
- The checksum passes.

A DLP policy has medium confidence that it detects this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_german_passport` finds content that matches the nine characters pattern (without check digit and optional d/D).
- A keyword from `Keyword_german_passport` or `Keywords_eu_passport_number_common` is found.

A DLP policy has low confidence that it detects this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_german_passport_checksum` finds content that matches the pattern.
- The checksum passes.

XML

```
    <!-- German Passport Number -->
    <Entity id="2e3da144-d42b-47ed-b123-fbf78604e52c" patternsProximity="300"
recommendedConfidence="75">
      <Pattern confidenceLevel="75">
        <IdMatch idRef="Func_german_passport" />
        <Any minMatches="1">
          <Match idRef="Keyword_german_passport" />
          <Match idRef="Keywords_eu_passport_number_common" />
        </Any>
      </Pattern>
      <Version minEngineVersion="15.20.4570.0">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_german_passport_checksum" />
          <Any minMatches="1">
            <Match idRef="Keyword_german_passport" />
            <Match idRef="Keywords_eu_passport_number_common" />
          </Any>
        </Pattern>
        <Pattern confidenceLevel="65">
          <IdMatch idRef="Func_german_passport_checksum" />
        </Pattern>
      </Version>
    </Entity>
```

Keywords

Keyword_german_passport

- reisepasse
- reisepassnummer
- No-Reisepass
- Nr-Reisepass
- Reisepass-Nr
- Passnummer

- reisepässe
- passeport no.
- passeport no

Keywords_eu_passport_number_common

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers

Last updated on 06/19/2026

Germany physical addresses

Article • 08/19/2024

This unbundled named entity detects patterns related to physical address from Germany. It's also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Germany tax identification number

Article • 08/19/2024

Format

11 digits without spaces and delimiters

Pattern

11 digits

- Two digits
- An optional space
- Three digits
- An optional space
- Three digits
- An optional space
- Two digits
- one check digit

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_germany_eu_tax_file_number` finds content that matches the pattern.

- A keyword from `Keywords_germany_eu_tax_file_number` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_germany_eu_tax_file_number` finds content that matches the pattern.

XML

```
      <!-- Germany Tax Identification Number -->
      <Entity id="43316a89-9880-40cf-b980-04bc7eefcec5"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_germany_eu_tax_file_number" />
          <Match idRef="Keywords_germany_eu_tax_file_number" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_germany_eu_tax_file_number" />
        </Pattern>
      </Entity>
```

Keywords

Keywords_germany_eu_tax_file_number

- identifikationsnummer
- steuer id
- steueridentifikationsnummer
- steuernummer
- tax id
- tax identification no
- tax identification number
- tax no#
- tax no
- tax number
- tax registration number
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber

- tin id
- tin no
- tin#
- zinn#
- zinn
- zinnnummer

Germany value added tax number

Article • 08/19/2024

Format

11 character alphanumeric pattern

Pattern

11-character alphanumeric pattern:

- a letter D or d
- a letter E or e
- an optional space
- three digits
- an optional space or comma
- three digits
- an optional space or comma
- three digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_germany_value_added_tax_number` finds content that matches the pattern.

- A keyword from `Keywords_germany_value_added_tax_number` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_germany_value_added_tax_number` finds content that matches the pattern.

XML

```
      <!-- Germany Value Added Tax Number -->
      <Entity id="db177eb2-8811-4842-bffc-128c14aa219f"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_germany_value_added_tax_number" />
          <Match idRef="Keywords_germany_value_added_tax_number" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_germany_value_added_tax_number" />
        </Pattern>
      </Entity>
```

Keywords

Keyword_germany_value_added_tax_number

- vat number
- vat no
- vat#
- vat# mehrwertsteuer
- mwst
- mehrwertsteuer identifikationsnummer
- mehrwertsteuer nummer

GitHub personal access token

This SIT is also included in the All credentials bundled SIT.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a credential scanning SIT in an endpoint DLP policy. This requirement is specific to the combination of credential scanning SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

Format

A combination of 40 characters consisting of letters, digits, and special characters.

or

Paired username and password used in URL.

or

A combination of 40 characters consisting of letters and digits.

Pattern

- A token prefix (case-sensitive) 'ghp_', 'gho_', 'ghu_', 'ghs_', or 'ghr_'
- Any combination of 36 a-z (not case-sensitive) or 0-9

format example:

`ghp_A1bC2dE3fH4iJ5kL6mN7oP8qR9sT0u`

or

Various URL username and password formats for example:

`https://username:********@contoso.com/`

`ftp://username:********@contoso.com:20/`

or

A combination of 40 characters:

- a-f or A-F (case-sensitive) or 0-9

format example:

`A1bC2dE3fH4iJ5kL6mN7oP8qR9sT0u`

Credential format example

`pat=ghp_A1bC2dE3fH4iJ5kL6mN7oP8qR9sT0u`

Checksum

Yes

SITs that have checksums use a unique calculation to check if the information is valid. This means when the **Checksum** value is **Yes**, the service can make a positive detection based on the sensitive data alone. When the **Checksum** value is **No** additional (secondary) elements must also be detected for the service to make a positive detection.

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

This SIT is designed to match the security information that's used as an alternate password for authentication to GitHub when using GitHub API or the command line.

It uses several primary resources:

- Patterns of Identifiable GitHub PAT.
- Patterns of User Login Credentials in URL.
- Patterns of Hex encoded 160 bits Symmetric Key.

- Patterns of CredentialName, CredentialFeatures, AccountIdentityName, AccountIdentityValue, ResourceType, ResourceName.
- Patterns of mockup values, redactions, and placeholders.

The patterns are designed to match actual credentials with reasonable confidence. The patterns don't match credentials formatted as examples. Mockup values, redacted values, and placeholders, like credential type or usage descriptions, in the position where an actual secret value should present won't be matched.

Keywords

Keyword_GitHubPatIdentifiableSecret:

- gh_

Keyword_LoginCredentialsInUrl:

- ://

Keyword_SymmetricKey160Hex:

- token

Last updated on 11/18/2025

Google API key

This SIT is also included in the All credentials bundled SIT.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a credential scanning SIT in an endpoint DLP policy. This requirement is specific to the combination of credential scanning SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

Format

A combination of 39 characters consisting of letters, digits, and special characters.

Pattern

A token prefix (case-sensitive) 'AIza'

A combination of 35 characters:

- a-z (not case-sensitive)
- 0-9
- hyphen (-)
- underline (_)
- or backward slashes (\)

format example:

```
A1bC2dE3f-4iJ5_L6mN\oP8qR9sT0u
```

Credential format example

```
apiKey=A1bC2dE3f-4iJ5_L6mN\oP8qR9sT0u;
```

Checksum

No

SITs that have checksums use a unique calculation to check if the information is valid. This means when the **Checksum** value is **Yes**, the service can make a positive detection based on the sensitive data alone. When the **Checksum** value is **No** additional (secondary) elements must also be detected for the service to make a positive detection.

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

This SIT is designed to match the security information that's used as a simple encrypted string that identifies a Google REST API client without any principal that is used to associate API requests with your project for quota and billing.

It uses several primary resources:

- Patterns of Base64 encoded 210 bits symmetric key.
- Patterns of CredentialName, CredentialFeatures, AccountIdentityName, AccountIdentityValue, ResourceType, ResourceName.
- Patterns of mockup values, redactions, and placeholders.
- A dictionary of vocabulary.

The patterns are designed to match actual credentials with reasonable confidence. The patterns don't match credentials formatted as examples. Mockup values, redacted values, and placeholders, like credential type or usage descriptions, in the position where an actual secret value should present won't be matched.

Keywords

Keyword_SymmetricKey210:

- AIza

Last updated on 11/18/2025

Greece drivers license number

Article • 08/19/2024

This entity is included in the EU Driver's License Number sensitive information type. It's also available as a stand-alone sensitive information type entity.

Format

nine digits without spaces and delimiters

Pattern

nine digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_greece_eu_driver's_license_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_driver's_license_number` or `Keywords_greece_eu_driver's_license_number` is found.

```
XML
```

```
      <!-- Greece Driver's License Number -->
      <Entity id="7a2200b5-aacf-4e3c-ab36-136d3e68b7da"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_greece_eu_driver's_license_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_driver's_license_number" />
            <Match idRef="Keywords_greece_eu_driver's_license_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic

- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#

- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen

- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses
- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_greece_eu_driver's_license_number

- δεια οδήγησης
- Adeia odigisis
- Άδεια οδήγησης
- Δίπλωμα οδήγησης

Greece national ID card

Article • 08/19/2024

Format

Combination of 7-8 letters and numbers plus a dash

Pattern

Seven letters and numbers (old format):

- One letter (any letter of the Greek alphabet)
- A dash
- Six digits

Eight letters and numbers (new format):

- Two letters whose uppercase character occurs in both the Greek and Latin alphabets (ABEZHIKMNOPTYX)
- A dash
- Six digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_greece_id_card` finds content that matches the pattern.
- A keyword from `Keyword_greece_id_card` is found.

A DLP policy has low confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_greece_id_card` finds content that matches the pattern.

XML

```
      <!-- Greece National ID Card -->
      <Entity id="82568215-1da1-46d3-874a-d2294d81b5ac"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Regex_greece_id_card" />
          <Match idRef="Keyword_greece_id_card" />
        </Pattern>
        <Pattern confidenceLevel="65">
          <IdMatch idRef="Regex_greece_id_card" />
        </Pattern>
      </Entity>
```

Keywords

Keyword_greece_id_card

- greek id
- greek national id
- greek personal id card
- greek police id
- identity card
- tautotita
- ταυτότητα
- ταυτότητας

Greece passport number

Article • 08/19/2024

Format

Two letters followed by seven digits with no spaces or delimiters

Pattern

Two letters followed by seven digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_greece_eu_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_greece_eu_passport_number` is found.
- The regular expression `Regex_greece_eu_passport_date` finds date in the format DD MMM YY (Example - 28 Aug 19) or a keyword from `Keywords_greece_eu_passport_date` is found

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_greece_eu_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_greece_eu_passport_number` is found.

XML

```
      <!-- Greece Passport Number -->
      <Entity id="7e65eb47-cdf9-4f52-8f90-2a27d5ee67e3"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Regex_greece_eu_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_greece_eu_passport_number" />
          </Any>
          <Any minMatches="1">
            <Match idRef="Regex_greece_eu_passport_date" />
            <Match idRef="Keywords_greece_eu_passport_date" />
          </Any>
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_greece_eu_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_greece_eu_passport_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_passport_number

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers

Keywords_greece_eu_passport_number

- αριθμός διαβατηρίου
- αριθμούς διαβατηρίου
- αριθμός διαβατηριο

Greece physical addresses

Article • 08/19/2024

This unbundled named entity detects patterns related to physical address from Greece. It's also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Greece Social Security Number (AMKA)

Article • 08/19/2024

Format

11 digits without spaces and delimiters

Pattern

- Six digits as date of birth YYMMDD
- Four digits
- a check digit

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_greece_eu_ssn` finds content that matches the pattern.
- A keyword from `Keywords_greece_eu_ssn_or_equivalent` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_greece_eu_ssn` finds content that matches the pattern.

```
XML
```

```
      <!-- Greece Social Security Number (AMKA) -->
      <Entity id="e39b03f4-50ea-41ae-af7a-a4b9539596ad"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_greece_eu_ssn" />
          <Match idRef="Keywords_greece_eu_ssn_or_equivalent" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_greece_eu_ssn" />
        </Pattern>
      </Entity>
```

Keywords

Keywords_greece_eu_ssn_or_equivalent

- ssn
- ssn#
- social security no
- socialsecurityno#
- social security number
- amka
- a.m.k.a.
- Αριθμού Μητρώου Κοινωνικής Ασφάλισης

Greek Tax Identification Number

Format

Nine digits without spaces and delimiters

Pattern

Nine digits

Checksum

Not applicable

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A data loss prevention (DLP) policy has medium confidence that it detects this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_greece_eu_tax_file_number` finds content that matches the pattern.
- A keyword from `Keywords_greece_eu_tax_file_number` is found.

XML

```
      <!-- Greek Tax Identification Number -->
      <Entity id="15a54a5a-53d4-4080-ad43-a2a4fe1d3bf7" patternsProximity="300"
recommendedConfidence="75">
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_greece_eu_tax_file_number" />
          <Match idRef="Keywords_greece_eu_tax_file_number" />
```

```
        </Pattern>
      </Entity>
```

Keywords

The following keywords are used for detection. These include the Greek tax file number (AFM) and related terms.

Keywords_greece_eu_tax_file_number

- `afm#`
- `afm`
- `αφμ|αφμ αριθμός`
- `αφμ`
- `tax id`
- `tax identification no`
- `tax identification number`
- `tax no#`
- `tax no`
- `tax number`
- `tax registration number`
- `tax registry no`
- `tax registry number`
- `taxid#`
- `taxidno#`
- `taxidnumber#`
- `taxno#`
- `taxnumber#`
- `taxnumber`
- `taxregistryno#`
- `tin id`
- `tin no`
- `tin#`
- `αριθμός φορολογικού μητρώου`
- `τον αριθμό φορολογικού μητρώου`
- `φορολογικού μητρώου νο`

Last updated on 06/19/2026

Greenland physical addresses

This unbundled named entity detects patterns related to physical address from Greenland. It is also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following services:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Last updated on 06/19/2026

Hong Kong identity card (HKID) number

Article • 08/19/2024

Format

Combination of 8-9 letters and numbers plus optional parentheses around the final character

Pattern

Combination of 8-9 letters:

- 1-2 letters (not case-sensitive)
- Six digits
- optional space
- a check character (any digit or the letter A) which is optionally enclosed in parentheses

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_hong_kong_id_card` finds content that matches the pattern.
- A keyword from `Keyword_hong_kong_id_card` is found.
- The checksum passes.

A DLP policy has low confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_hong_kong_id_card` finds content that matches the pattern.
- The checksum passes.

XML

```
<!-- Hong Kong Identity Card (HKID) number -->
<Entity id="e63c28a7-ad29-4c17-a41a-3d2a0b70fd9c" recommendedConfidence="75"
patternsProximity="300">
  <Pattern confidenceLevel="75">
     <IdMatch idRef="Func_hong_kong_id_card"/>
     <Match idRef="Keyword_hong_kong_id_card"/>
  </Pattern>
  <Pattern confidenceLevel="65">
     <IdMatch idRef="Func_hong_kong_id_card"/>
  </Pattern>
</Entity>
```

Keywords

Keyword_hong_kong_id_card

- hkid
- hong kong identity card
- HKIDC
- id card
- identity card
- hk identity card
- hong kong id
- 香港身份證
- 香港永久性居民身份證
- 身份證
- 身份証
- 身分證
- 身分証
- 香港身份証
- 香港身分證
- 香港身分証
- 香港身份證
- 香港居民身份證
- 香港居民身份証

- 香港居民身分證
- 香港居民身分証
- 香港永久性居民身份証
- 香港永久性居民身分證
- 香港永久性居民身分証
- 香港永久性居民身份證
- 香港非永久性居民身份證
- 香港非永久性居民身份証
- 香港非永久性居民身分證
- 香港非永久性居民身分証
- 香港特別行政區永久性居民身份證
- 香港特別行政區永久性居民身份証
- 香港特別行政區永久性居民身分證
- 香港特別行政區永久性居民身分証
- 香港特別行政區非永久性居民身份證
- 香港特別行政區非永久性居民身份証
- 香港特別行政區非永久性居民身分證
- 香港特別行政區非永久性居民身分証

Http authorization header

This SIT is also included in the All credentials bundled SIT.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a credential scanning SIT in an endpoint DLP policy. This requirement is specific to the combination of credential scanning SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

Format

An authorization header used in HTTP request.

Pattern

Various authentication header formats for example:

```
authorization: basic ********
authorization: bearer ********
authorization: digest ********
authorization: negotiate ********
```

Credential example

```
Authorization: Basic ABCDEFGHIJKLMNOPQRS0123456789;
```

Checksum

No

SITs that have checksums use a unique calculation to check if the information is valid. This means when the **Checksum** value is **Yes**, the service can make a positive detection based on the sensitive data alone. When the **Checksum** value is **No** additional (secondary) elements must also be detected for the service to make a positive detection.

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

This SIT is designed to match the security information that's used in the header of an HTTP request for authentication and authorization.

It uses several primary resources:

- Patterns of Http authorization header.
- Patterns of CredentialName, CredentialFeatures, ResourceType.
- Patterns of mockup values, redactions, and placeholders.

The patterns are designed to match actual credentials with reasonable confidence. The patterns don't match credentials formatted as examples. Mockup values, redacted values, and placeholders, like credential type or usage descriptions, in the position where an actual secret value should present won't be matched.

Keywords

Keyword_HttpAuthorizationHeader:

- authorization

Last updated on 11/18/2025

Hungary drivers license number

Article • 08/19/2024

Format

Two letters followed by six digits

Pattern

Two letters and six digits:

- Two letters (not case-sensitive)
- Six digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_hungary_eu_driver's_license_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_driver's_license_number` or `Keywords_hungary_eu_driver's_license_number` is found.

XML

```
      <Entity id="9d31c46b-6e6b-444c-aeb1-6dd7e604bb24"
patternsProximity="300" recommendedConfidence="75">
```

```
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_hungary_eu_driver's_license_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_driver's_license_number" />
            <Match idRef="Keywords_hungary_eu_driver's_license_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses

- driver'licence
- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#

- drivers license#
- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic

- driving licen
- driving licenses
- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_hungary_eu_driver's_license_number

- vezetoi engedely
- vezetői engedély
- vezetői engedélyek

Hungary passport number

Article • 08/19/2024

Format

Two letters followed by six or seven digits with no spaces or delimiters

Pattern

Two letters followed by six or seven digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_hungary_eu_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_hungary_eu_passport_number` is found.
- The regular expression `Regex_hungary_eu_passport_date` finds date in the format DD MMM/MMM YY (Example - 01 MÁR/MAR 12) or a keyword from `Keywords_eu_passport_date` is found

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_hungary_eu_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_hungary_eu_passport_number` is found.

XML

```
      <!-- Hungary Passport Number -->
      <Entity id="5b483910-9aa7-4c99-9917-f4001464bda7"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Regex_hungary_eu_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_hungary_eu_passport_number" />
          </Any>
          <Any minMatches="1">
            <Match idRef="Regex_hungary_eu_passport_date" />
            <Match idRef="Keywords_eu_passport_date" />
          </Any>
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_hungary_eu_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_hungary_eu_passport_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_passport_number

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers

Keywords_hungary_eu_passport_number

- útlevél száma
- Útlevelek száma
- útlevél szám

Keywords_eu_passport_date

- date of issue
- date of expiry

Hungary personal identification number

Article • 08/19/2024

Format

11 digits

Pattern

11 digits:

- One digit that corresponds to gender, 1 for male, 2 for female. Other numbers are also possible for citizens born before 1900 or citizens with double citizenship.
- Six digits that correspond to birth date (YYMMDD)
- Three digits that correspond to a serial number
- One check digit

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_hungary_eu_national_id_card` finds content that matches the pattern.
- A keyword from `Keywords_hungary_eu_national_id_card` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_hungary_eu_national_id_card` finds content that matches the pattern.

XML

```
      <!-- Hungary Personal Identification Number -->
      <Entity id="7b5cc218-7046-47d9-80c9-f325b50896ca"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_hungary_eu_national_id_card" />
          <Match idRef="Keywords_hungary_eu_national_id_card" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_hungary_eu_national_id_card" />
          <Any minMatches="0" maxMatches="0">
            <Match idRef="Keywords_hungary_eu_telephone_number" />
            <Match idRef="Keywords_hungary_eu_mobile_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_hungary_eu_national_id_card

- id number
- identification number
- sz ig
- sz. ig.
- sz.ig.
- személyazonosító igazolvány
- személyi igazolvány

Hungary physical addresses

Article • 08/19/2024

This unbundled named entity detects patterns related to physical address from Hungary. It's also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Hungary social security number (TAJ)

Article • 08/19/2024

Format

Nine digits without spaces and delimiters

Pattern

Nine digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_hungary_eu_ssn_or_equivalent` finds content that matches the pattern.
- A keyword from `Keywords_hungary_eu_ssn_or_equivalent` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_hungary_eu_ssn_or_equivalent` finds content that matches the pattern.

```
XML
```

```
      <!-- Hungarian Social Security Number (TAJ) -->
      <Entity id="0de78315-9537-47f5-95ab-b3e77eba3993"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_hungary_eu_ssn_or_equivalent" />
          <Match idRef="Keywords_hungary_eu_ssn_or_equivalent" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_hungary_eu_ssn_or_equivalent" />
        </Pattern>
      </Entity>
```

Keywords

Keywords_hungary_eu_ssn_or_equivalent

- hungarian social security number
- social security number
- socialsecuritynumber#
- hssn#
- socialsecuritynno
- hssn
- taj
- taj#
- ssn
- ssn#
- social security no
- áfa
- közösségi adószám
- általános forgalmi adó szám
- hozzáadottérték adó
- áfa szám
- magyar áfa szám

Hungary tax identification number

Article • 08/19/2024

Format

10 digits with no spaces or delimiters

Pattern

10 digits:

- One digit that must be "8"
- Eight digits
- One check digit

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_hungary_eu_tax_file_number` finds content that matches the pattern.
- A keyword from `Keywords_hungary_eu_tax_file_number` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_hungary_eu_tax_file_number` finds content that matches the pattern.

XML

```
      <!-- Hungary Tax Identification Number -->
      <Entity id="ede42eb4-59d9-49eb-9603-d7853fbda91d"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_hungary_eu_tax_file_number" />
          <Match idRef="Keywords_hungary_eu_tax_file_number" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_hungary_eu_tax_file_number" />
          <Any minMatches="0" maxMatches="0">
            <Match idRef="Keywords_hungary_eu_telephone_number" />
            <Match idRef="Keywords_hungary_eu_mobile_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_hungary_eu_tax_file_number

- adóazonosító szám
- adóhatóság szám
- adószám
- hungarian tin
- hungatiantin#
- tax authority no
- tax id
- tax identification no
- tax identification number
- tax no#
- tax no
- tax number
- tax registration number
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#

- taxnumber
- tin id
- tin no
- tin#
- vat number

Hungary value added tax number

Article • 08/19/2024

Format

10 character alphanumeric pattern

Pattern

10 character alphanumeric pattern:

- two letters - HU or hu
- optional space
- eight digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_hungarian_value_added_tax_number` finds content that matches the pattern.
- A keyword from `Keywords_hungarian_value_added_tax_number` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_hungarian_value_added_tax_number` finds content that matches the pattern.

XML

```
      <!-- Hungarian Value Added Tax Number -->
      <Entity id="976349a0-683b-477a-90f8-ff0a220d5592"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_hungarian_value_added_tax_number" />
          <Match idRef="Keywords_hungarian_value_added_tax_number" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_hungarian_value_added_tax_number" />
        </Pattern>
      </Entity>
```

Keywords

Keyword_Hungary_value_added_tax_number

- vat
- value added tax number
- vat#
- vatno#
- hungarianvatno#
- tax no.
- value added tax áfa
- közösségi adószám
- általános forgalmi adó szám
- hozzáadottérték adó
- áfa szám

Iceland physical addresses

Article • 08/19/2024

This unbundled named entity detects patterns related to physical address from Iceland. It's also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Impairments Listed In The U.S. Disability Evaluation Under Social Security

Article • 08/19/2024

This unbundled named entity detects names of impairments listed in the U.S. Disability Evaluation Under Social Security, such as *muscular dystrophy*. It supports English terms only. It's also included in the All medical terms and conditions bundled named entity SIT.

Confidence level

High

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

India Drivers License Number

Article • 08/19/2024

Format

15 character alphanumeric pattern

Pattern

15 letters or digits:

- two letters indicating state code
- optional space or dash
- two digits indicating city code
- optional space or dash
- four digits indicating year of issue
- optional space or dash
- seven digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_india_driving_license` finds content that matches the pattern.

- A keyword from `Keywords_eu_driver's_license_number_common` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_india_driving_license` finds content that matches the pattern.

XML

```
      <!-- India Driver's License Number -->
        <Entity id="680788a3-53b6-455a-b891-c38cd76dc917"
patternsProximity="300" recommendedConfidence="85" relaxProximity="true">
          <Pattern confidenceLevel="85">
            <IdMatch idRef="Regex_india_driving_license" />
            <Match idRef="Keywords_eu_driver's_license_number_common" />
          </Pattern>
          <Pattern confidenceLevel="75">
            <IdMatch idRef="Regex_india_driving_license" />
            </Pattern>
        </Entity>
```

Keywords

Keywords_eu_driver's_license_number_common

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense

- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#

- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license

- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses
- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

India GST Number

Article • 08/19/2024

Format

15 character alphanumeric pattern

Pattern

15 letters or digits:

- two digits representing valid state code
- an optional space or dash
- ten characters representing Permanent Account Number (PAN)
- one letter or digit
- an optional space or dash
- one letter 'z' or 'Z'
- an optional space or dash
- one check digit

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_india_gst_number` finds content that matches the pattern.

- A keyword from `Keyword_india_gst_number` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_india_gst_number` finds content that matches the pattern.

XML

```
    <!-- India GST number  -->
      <Entity id="9f5a721c-2fd2-446a-a27e-0c02fbe4630c"
patternsProximity="300" recommendedConfidence="85" relaxProximity="true">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_india_gst_number" />
          <Match idRef="Keyword_india_gst_number" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_india_gst_number" />
        </Pattern>
      </Entity>
```

Keywords

Keyword_india_gst_number

- gst
- gstin
- goods and services tax
- goods and service tax

India permanent account number (PAN)

Article • 08/19/2024

Format

10 letters or digits

Pattern

10 letters or digits:

- Three letters (not case-sensitive)
- A letter in C, P, H, F, A, T, B, L, J, G (not case-sensitive)
- A letter
- Four digits
- A letter that is an alphabetic check digit

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_india_permanent_account_number` finds content that matches the pattern.
- A keyword from `Keyword_india_permanent_account_number` is found.

A DLP policy has low confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_india_permanent_account_number` finds content that matches the pattern.

XML

```
      <!-- India Permanent Account Number -->
      <Entity id="2602bfee-9bb0-47a5-a7a6-2bf3053e2804"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Regex_india_permanent_account_number" />
          <Match idRef="Keyword_india_permanent_account_number" />
        </Pattern>
        <Version minEngineVersion="15.20.3520.000">
          <Pattern confidenceLevel="65">
            <IdMatch idRef="Regex_india_permanent_account_number" />
          </Pattern>
        </Version>
      </Entity>
```

Keywords

Keyword_india_permanent_account_number

- Permanent Account Number
- PAN

India unique identification (Aadhaar) number

Article • 08/19/2024

Format

12 digits containing optional spaces or dashes

Pattern

12 digits:

- A digit that isn't 0 or 1
- Three digits
- An optional space or dash
- Four digits
- An optional space or dash
- Three digits
- The final digit, which is the check digit

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_india_aadhaar` finds content that matches the pattern.

- A keyword from `Keyword_india_aadhar` is found.
- The checksum passes.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_india_aadhaar` finds content that matches the pattern.
- The checksum passes.

XML

```
<!-- India Unique Identification (Aadhaar) number -->
<Entity id="1ca46b29-76f5-4f46-9383-cfa15e91048f" recommendedConfidence="85"
patternsProximity="300">
  <Pattern confidenceLevel="85">
     <IdMatch idRef="Func_india_aadhaar"/>
     <Match idRef="Keyword_india_aadhar"/>
  </Pattern>
  <Pattern confidenceLevel="75">
     <IdMatch idRef="Func_india_aadhaar"/>
  </Pattern>
</Entity>
```

Keywords

Keyword_india_aadhar

- aadhaar
- aadhar
- aadhar#
- uid
- आधार
- uidai

India Voter ID Card

Article • 08/19/2024

Format

10 character alphanumeric pattern

Pattern

10 letters or digits:

- three letters
- seven digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_india_voter_id_card` finds content that matches the pattern.
- A keyword from `Keyword_india_voter_id_card` is found.

A DLP policy has low confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_india_voter_id_card` finds content that matches the pattern.

XML

```
      <!-- India Voter Id Card  -->
        <Entity id="646d643f-5228-4408-acc8-f2e81a6df897"
patternsProximity="300" recommendedConfidence="75" relaxProximity="true">
           <Pattern confidenceLevel="75">
             <IdMatch idRef="Regex_india_voter_id_card" />
             <Match idRef="Keyword_india_voter_id_card" />
            </Pattern>
           <Pattern confidenceLevel="65">
              <IdMatch idRef="Regex_india_voter_id_card" />
            </Pattern>
        </Entity>
```

Keywords

Keyword_india_voter_id_card

- voter
- voterid
- votercard
- voteridcard
- electoral photo identity card
- EPIC
- ECI
- election commmision

Indonesia driver's license number

Article • 08/19/2024

Format

Twelve or Fourteen digits

Pattern

Twelve digits:

- any digit between 4-9
- followed by any digit between 0-9
- followed either 0 or 1
- followed by any nine digits between 0-9

Fourteen digits

- four digits
- followed by hyphen -
- four digits
- followed by hyphen -
- six digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regex `Regex_indonesia_drivers_license_number` finds content that matches the pattern.
- A keyword from `Keyword_indonesia_drivers_license_number` is found.

A DLP policy has low confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regex `Regex_indonesia_drivers_license_number` finds content that matches the pattern.
- A keyword from `Keyword_indonesia_drivers_license_number` is found.

XML

```
<!-- Indonesia Drivers License Number -->
    <Entity id="67d902ee-c3ce-44a2-81ac-e9746e7d60a1"
patternsProximity="300" recommendedConfidence="75" relaxProximity="true">
      <Pattern confidenceLevel="75">
        <IdMatch idRef="Regex_indonesia_drivers_license_number" />
        <Match idRef="Keyword_indonesia_drivers_license_number" />
      </Pattern>
      <Pattern confidenceLevel="65">
        <IdMatch idRef="Regex_indonesia_drivers_license_number" />
      </Pattern>
    </Entity>
```

Keywords

Keyword_indonesia_drivers_license_number

- SIM
- Surat Ijin Mengemudi
- SMART SIM
- Non-SMART SIM
- nomor SIM
- nomor lisensi
- DriverLic
- DriverLics
- DriverLicense
- DriverLicenses
- Driver Lic

- Driver Lics
- Driver License
- Driver Licenses
- DriversLic
- DriversLics
- DriversLicense
- DriversLicenses
- Drivers Lic
- Drivers Lics
- Drivers License
- Drivers Licenses
- Driver'Lic
- Driver'Lics
- Driver'License
- Driver'Licenses
- Driver' Lic
- Driver' Lics
- Driver' License
- Driver' Licenses
- Driver'sLic
- Driver'sLics
- Driver'sLicense
- Driver'sLicenses
- Driver's Lic
- Driver's Lics
- Driver's License
- Driver's Licenses
- DriverLic#
- DriverLics#
- DriverLicense#
- DriverLicenses#
- Driver Lic#
- Driver Lics#
- Driver License#
- Driver Licenses#
- DriversLic#
- DriversLics#
- DriversLicense#
- DriversLicenses#
- Drivers Lic#
- Drivers Lics#

- Drivers License#
- Drivers Licenses#
- Driver'Lic#
- Driver'Lics#
- Driver'License#
- Driver'Licenses#
- Driver' Lic#
- Driver' Lics#
- Driver' License#
- Driver' Licenses#
- Driver'sLic#
- Driver'sLics#
- Driver'sLicense#
- Driver'sLicenses#
- Driver's Lic#
- Driver's Lics#
- Driver's License#
- Driver's Licenses#

Indonesia identity card (KTP) number

Article • 08/19/2024

Format

Sixteen digits

Pattern

Sixteen digits:

- Two-digit province code
- Four digits
- Six digits in the format DDMMYY, which are the date of birth
- Four digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_indonesia_id_card` finds content that matches the pattern.
- A keyword from `Keyword_indonesia_id_card` is found.

```
XML
```

```
<!-- Indonesia Identity Card (KTP) Number -->
<Entity id="da68fdb0-f383-4981-8c86-82689d3b7d55" recommendedConfidence="85"
patternsProximity="300">
  <Pattern confidenceLevel="85">
     <IdMatch idRef="Regex_indonesia_id_card"/>
     <Match idRef="Keyword_indonesia_id_card"/>
</Entity>
```

Keywords

Keyword_indonesia_id_card

- KTP
- Kartu Tanda Penduduk
- Nomor Induk Kependudukan

Indonesia passport number

09/26/2025

Format

Seven to Nine alphanumeric characters

Pattern

- one or two letters
- followed by six or seven digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regex `Regex_indonesia_passport_number` finds content that matches the pattern.
- A keyword from `Keyword_indonesia_passport_numbe` is found.

A DLP policy has low confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regex `Regex_indonesia_passport_number` finds content that matches the pattern.

```
XML
```

```
    <!-- Indonesia Passport Number -->
    <Entity id="d0001c83-e72f-4360-98d3-f5a41dc5a380" patternsProximity="300"
recommendedConfidence="75" relaxProximity="true">
      <Pattern confidenceLevel="75">
        <IdMatch idRef="Regex_indonesia_passport_number" />
        <Match idRef="Keyword_indonesia_passport_number" />
      </Pattern>
      <Pattern confidenceLevel="65">
        <IdMatch idRef="Regex_indonesia_passport_number" />
      </Pattern>
    </Entity>
```

Keywords

Keyword_indonesia_passport_number

- passport#
- passport #
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers
- passport
- passport no.
- passportno#
- passport book
- passportbook#
- paspor
- nomor paspor
- paspornomor
- paspor no
- no. paspor.

International banking account number (IBAN)

Article • 08/19/2024

Format

Country code (two letters) plus check digits (two digits) plus bban number (up to 30 characters)

Pattern

Pattern must include all of these patterns:

- Two-letter country code
- Two check digits (followed by an optional space)
- 1-7 groups of four letters or digits (can be separated by spaces)
- 1-3 letters or digits

The format for each country/region is slightly different. The IBAN sensitive information type covers these 68 countries:

- ad
- ae
- al
- at
- az
- ba
- be
- bg
- bh
- br
- ch
- cr
- cy
- cz
- de
- dk
- do
- ee

- es
- fi
- fo
- fr
- gb
- ge
- gi
- gl
- gr
- gt
- hr
- hu
- ie
- il
- is
- it
- jo
- kw
- kz
- lb
- li
- lt
- lu
- lv
- mc
- md
- me
- mk
- mr
- mt
- mu
- nl
- no
- pk
- pl
- ps
- pt
- qa
- ro
- rs
- sa

- se
- si
- sk
- sm
- tl
- tn
- tr
- vg
- xk

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_iban` finds content that matches the pattern.
- The checksum passes.

XML

```
<Entity id="e7dc4711-11b7-4cb0-b88b-2c394a771f0e" patternsProximity="300"
recommendedConfidence="85">
  <Pattern confidenceLevel="85">
        <IdMatch idRef="Func_iban" />
  </Pattern>
</Entity>
```

Keywords

None

International classification of diseases (ICD-10-CM)

Article • 08/19/2024

Format

Dictionary

Pattern

Keyword

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- A keyword from `Dictionary_icd_10_updated` is found.
- A keyword from `Dictionary_icd_10_codes` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- A keyword from `Dictionary_icd_10_ updated` is found.

```
XML
```

```
      <!-- ICD-10 CM -->
      <Entity id="3356946c-6bb7-449b-b253-6ffa419c0ce7"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Dictionary_icd_10_updated" />
          <Match idRef="Dictionary_icd_10_codes" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Dictionary_icd_10_updated" />
        </Pattern>
```

Keywords

Any term from the `Dictionary_icd_10_updated` keyword dictionary, which is based on the International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM). This type looks only for the term, not the insurance codes.

Any term from the `Dictionary_icd_10_codes` keyword dictionary, which is based on the International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM). This type looks only for insurance codes, not the description.

International classification of diseases (ICD-9-CM)

Article • 08/19/2024

Format

Dictionary

Pattern

Keyword

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- A keyword from `Dictionary_icd_9_updated` is found.
- A keyword from `Dictionary_icd_9_codes` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- A keyword from `Dictionary_icd_9_updated` is found.

```
XML
```

```
    <Entity id="fa3f9c74-ee07-4c52-b5f2-085d6b2c0ec4"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Dictionary_icd_9_updated" />
          <Match idRef="Dictionary_icd_9_codes" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Dictionary_icd_9_updated" />
        </Pattern>
      </Entity>
```

Keywords

IP address

Article • 08/19/2024

Format

IPv4:

Complex pattern that accounts for formatted (periods) and unformatted (no periods) versions of the IPv4 addresses

IPv6:

Complex pattern that accounts for formatted IPv6 numbers (which include colons)

Pattern

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

For **IPv4**, a DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_ipv4_address` finds content that matches the pattern.
- A keyword from `Keyword_ipaddress` is found.

For **IPv6**, a DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_ipv6_address` finds content that matches the pattern.
- No keyword from `Keyword_ipaddress` is found.

XML

```
    <!-- IP Address -->
    <Entity id="1daa4ad5-e2dd-4ca4-a788-54722c09efb2"
patternsProximity="300" recommendedConfidence="85">
      <Pattern confidenceLevel="85">
        <IdMatch idRef="Regex_ipv6_address" />
        <Any minMatches="0" maxMatches="0">
          <Match idRef="Keyword_ipaddress" />
        </Any>
      </Pattern>
      <Pattern confidenceLevel="95">
        <IdMatch idRef="Regex_ipv4_address" />
        <Any minMatches="1">
          <Match idRef="Keyword_ipaddress" />
        </Any>
      </Pattern>
      <Pattern confidenceLevel="95">
        <IdMatch idRef="Regex_ipv6_address" />
        <Any minMatches="1">
          <Match idRef="Keyword_ipaddress" />
        </Any>
      </Pattern>
    </Entity>
```

Keywords

Keyword_ipaddress

- IP (this keyword is case-sensitive)
- ip address
- ip addresses
- internet protocol
- כתובת ה-IP

IP Address v4

Article • 08/19/2024

Format

Complex pattern that accounts for formatted (periods) and unformatted (no periods) versions of the IPv4 addresses.

Pattern

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_ipv4_address` finds content that matches the pattern.
- A keyword from `Keyword_ipaddress` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_ipv4_address` finds content that matches the pattern.

XML

```
      <!-- IP Address v4-->
      <Entity id="a7dd5e5f-e7f9-4626-a2c6-86a8cb6830d2"
patternsProximity="300" recommendedConfidence="75" relaxProximity="true">
```

```
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Regex_ipv4_address" />
          <Match idRef="Keyword_ipaddress" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_ipv4_address" />
        </Pattern>
      </Entity>
```

Keywords

Keyword_ipaddress

- IP (case sensitive)
- ip address
- ip addresses
- internet protocol
- כתובת ה-IP

IP Address v6

Article • 08/19/2024

Format

Complex pattern that accounts for formatted IPv6 numbers (which include colons)

Pattern

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_ipv6_address` finds content that matches the pattern.
- A keyword from `Keyword_ipaddress` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_ipv6_address` finds content that matches the pattern.

XML

```
      <!-- IP Address v6-->
      <Entity id="3f691089-7413-4926-ab3b-3c5ea8a1c17e"
patternsProximity="300" recommendedConfidence="75" relaxProximity="true">
        <Pattern confidenceLevel="85">
```

```
          <IdMatch idRef="Regex_ipv6_address" />
          <Match idRef="Keyword_ipaddress" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_ipv6_address" />
        </Pattern>
      </Entity>
```

Keywords

Keyword_ipaddress

- IP (case sensitive)
- ip address
- ip addresses
- internet protocol
- כתובת ה-IP

Ireland drivers license number

Article • 08/19/2024

Format

Six digits followed by four letters

Pattern

Six digits and four letters:

- Six digits
- Four letters (not case-sensitive)

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_ireland_eu_driver's_license_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_driver's_license_number` or `Keywords_ireland_eu_driver's_license_number` is found.

```
XML
```

```
      <!-- Ireland Driver's License Number -->
      <Entity id="e01bccd9-eb4d-414f-ace1-e9b6a4c4a2ca"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_ireland_eu_driver's_license_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_driver's_license_number" />
            <Match idRef="Keywords_ireland_eu_driver's_license_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic

- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#

- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen

- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses
- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_ireland_eu_driver's_license_number

- ceadúnas tiomána
- ceadúnais tiomána
- The regular expression Regex_ipv4_address finds content that matches the pattern.
- A keyword from `Keyword_ipaddress` is found.

For IPv6, a DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression Regex_ipv6_address finds content that matches the pattern.
- No keyword from `Keyword_ipaddress` is found.

XML

```
    <!-- IP Address -->
    <Entity id="1daa4ad5-e2dd-4ca4-a788-54722c09efb2"
patternsProximity="300" recommendedConfidence="85">
      <Pattern confidenceLevel="85">
        <IdMatch idRef="Regex_ipv6_address" />
        <Any minMatches="0" maxMatches="0">
          <Match idRef="Keyword_ipaddress" />
        </Any>
      </Pattern>
      <Pattern confidenceLevel="95">
        <IdMatch idRef="Regex_ipv4_address" />
        <Any minMatches="1">
          <Match idRef="Keyword_ipaddress" />
        </Any>
      </Pattern>
      <Pattern confidenceLevel="95">
        <IdMatch idRef="Regex_ipv6_address" />
```

```
        <Any minMatches="1">
          <Match idRef="Keyword_ipaddress" />
        </Any>
      </Pattern>
    </Entity>
```

Keyword_ipaddress

- IP (this keyword is case-sensitive)
- ip address
- ip addresses
- internet protocol
- כתובת ה-IP

Ireland passport number

Article • 08/19/2024

Format

Two letters or digits followed by seven digits with no spaces or delimiters

Pattern

Two letters or digits followed by seven digits:

- Two digits or letters (not case-sensitive)
- Seven digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_ireland_eu_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_ireland_eu_passport_number` is found.
- The regular expression `Regex_ireland_eu_passport_date` finds date in the format DD MMM/MMM YYYY (Example - 01 BEA/MAY 1988) or a keyword from `Keywords_eu_passport_date` is found

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_ireland_eu_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_ireland_eu_passport_number` is found.

XML

```
      <!-- Ireland Passport Number -->
      <Entity id="a2130f27-9ee2-4103-84f9-a6b1ee7d0cbf"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Regex_ireland_eu_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_ireland_eu_passport_number" />
          </Any>
          <Any minMatches="1">
            <Match idRef="Regex_ireland_eu_passport_date" />
            <Match idRef="Keywords_eu_passport_date" />
          </Any>
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_ireland_eu_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_ireland_eu_passport_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_passport_number_common

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number

- passportnumbers
- passport numbers

Keywords_ireland_eu_passport_number

- passeport numero
- uimhreacha pasanna
- uimhir pas
- uimhir phas
- uimhreacha pas
- uimhir cárta
- uimhir chárta

Keywords_eu_passport_date

- date of issue
- date of expiry

Ireland personal public service (PPS) number

Article • 08/19/2024

Format

Old format (until 31 December 2012):

- seven digits followed by 1-2 letters

New format (1 January 2013 and after):

- seven digits followed by two letters

Pattern

Old format (until 31 December 2012):

- seven digits
- one to two letters (not case-sensitive)

New format (1 January 2013 and after):

- seven digits
- a letter (not case-sensitive) which is an alphabetic check digit
- An optional letter in the range A-I, or "W"

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_ireland_pps finds` content that matches the pattern.
- A keyword from `Keywords_ireland_eu_national_id_card` is found.
- The checksum passes.

A DLP policy has low confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_ireland_pps` finds content that matches the pattern.
- The checksum passes.

XML

```
      <!-- Ireland Personal Public Service (PPS) Number -->
      <Entity id="1cdb674d-c19a-4fcf-9f4b-7f56cc87345a"
patternsProximity="300" recommendedConfidence="85" relaxProximity="true">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_ireland_pps" />
          <Match idRef="Keywords_ireland_eu_national_id_card" />
        </Pattern>
        <Pattern confidenceLevel="65">
          <IdMatch idRef="Func_ireland_pps" />
        </Pattern>
      </Entity>
```

Keywords

Keywords_ireland_eu_national_id_card

- client identity service
- identification number
- personal id number
- personal public service number
- personal service no
- phearsanta seirbhíse poiblí
- pps no
- pps number
- pps num
- pps service no

- ppsn
- ppsno#
- ppsno
- psp
- public service no
- publicserviceno#
- publicserviceno
- revenue and social insurance number
- rsi no
- rsi number
- rsin
- seirbhís aitheantais cliant
- uimh
- uimhir aitheantais chánach
- uimhir aitheantais phearsanta
- uimhir phearsanta seirbhíse poiblí
- tax id
- tax identification no
- tax identification number
- tax no#
- tax no
- tax number
- tax registration number
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id
- tin no
- tin#

Ireland physical addresses

Article • 08/19/2024

This unbundled named entity detects patterns related to physical address from Ireland. It's also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Israel bank account number

Article • 08/19/2024

Format

13 digits

Pattern

Formatted:

- two digits
- a dash
- three digits
- a dash
- eight digits

Unformatted:

- 13 consecutive digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_israel_bank_account_number` finds content that matches the pattern.
- A keyword from `Keyword_israel_bank_account_number` is found.

XML

```xml
<!-- Israel Bank Account Number -->
<Entity id="7d08b2ff-a0b9-437f-957c-aeddbf9b2b25" patternsProximity="300"
recommendedConfidence="75">
    <Pattern confidenceLevel="75">
        <IdMatch idRef="Regex_israel_bank_account_number" />
        <Any minMatches="1">
          <Match idRef="Keyword_israel_bank_account_number" />
        </Any>
    </Pattern>
</Entity>
```

Keywords

Keyword_israel_bank_account_number

- Bank Account Number
- Bank Account
- Account Number
- מספר חשבון בנק

Israel national identification number

Article • 08/19/2024

Format

nine digits

Pattern

nine consecutive digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_israeli_national_id_number` finds content that matches the pattern.
- A keyword from `Keyword_Israel_National_ID` is found.
- The checksum passes.

XML

```
<!-- Israel National ID Number -->
<Entity id="e05881f5-1db1-418c-89aa-a3ac5c5277ee" patternsProximity="300"
recommendedConfidence="75">
    <Pattern confidenceLevel="75">
        <IdMatch idRef="Func_israeli_national_id_number" />
```

```
        <Any minMatches="1">
          <Match idRef="Keyword_Israel_National_ID" />
        </Any>
    </Pattern>
</Entity>
```

Keywords

Keyword_Israel_National_ID

- תעודת זהות
- ת.ז.
- מספר זיהוי
- מספר זהות
- זהותישר אלית
- מספר הזהות
- هوية إسرائ يلية
- رقم الهوية
- بطاقة تعريف
- idnumber#
- id number
- identity no
- identitynumber#
- identity number
- israeliidentitynumber
- personal id
- unique id

Italy drivers license number

This entity is included in the EU Driver's License Number sensitive information type. It's also available as a stand-alone sensitive information type entity.

Format

a combination of 10 letters and digits

Pattern

a combination of 10 letters and digits:

- one letter (not case-sensitive)
- the letter "A" or "V" (not case-sensitive)
- space | hyphen | spacehyphenspace (Optional)
- seven digits
- space | hyphen | spacehyphenspace (Optional)
- one letter (not case-sensitive)

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_italy_drivers_license_number` finds content that matches the pattern.

- A keyword from `Keywords_eu_driver's_license_number` or `Keyword_italy_drivers_license_number` is found.

XML

```
    <!-- Italy Driver's license Number -->
    <Entity id="97d6244f-9157-41bd-8e0c-9d669a5c4d71" patternsProximity="300"
recommendedConfidence="75">
      <Pattern confidenceLevel="75">
        <IdMatch idRef="Regex_italy_drivers_license_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_driver's_license_number" />
            <Match idRef="Keyword_italy_drivers_license_number" />
          </Any>
        </Pattern>
    </Entity>
```

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license

- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#

- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license

- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses
- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keyword_italy_drivers_license_number

- numero di patente
- patente di guida
- patente guida
- patenti di guida
- patenti guida

Last updated on 01/08/2026

Italy fiscal code

Article • 08/19/2024

Format

a 16-character combination of letters and digits in the specified pattern

Pattern

A 16-character combination of letters and digits:

- three letters that correspond to the first three consonants in the family name
- three letters that correspond to the first, third, and fourth consonants in the first name
- two digits that correspond to the last digits of the birth year
- one letter that corresponds to the letter for the month of birth—letters are used in alphabetical order, but only the letters A to E, H, L, M, P, R to T are used (so, January is A and October is R)
- two digits that correspond to the day of the month of birth in order to differentiate between genders, 40 is added to the day of birth for women
- four digits that correspond to the area code specific to the municipality where the person was born (country/region-wide codes are used for foreign countries/regions)
- one parity digit

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_italy_eu_national_id_card` finds content that matches the pattern.
- A keyword from `Keywords_italy_eu_national_id_card` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_italy_eu_national_id_card` finds content that matches the pattern.

XML

```
      <!-- Italy Fiscal Code -->
      <Entity id="4cd79172-8da9-4ff5-9188-98b1e7e2eca6"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_italy_eu_national_id_card" />
          <Match idRef="Keywords_italy_eu_national_id_card" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_italy_eu_national_id_card" />
        </Pattern>
      </Entity>
```

Keywords

Keywords_italy_eu_national_id_card

- codice fiscal
- codice fiscale
- codice id personale
- codice personale
- fiscal code
- numero certificato personale
- numero di identificazione fiscale
- numero id personale
- numero personale
- personal certificate number
- personal code
- personal id code

- personal id number
- personalcodeno#
- tax code
- tax id
- tax identification no
- tax identification number
- tax identity number
- tax no#
- tax no
- tax number
- tax registration number
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id
- tin no
- tin#

Italy passport number

Article • 08/19/2024

Format

two letters or digits followed by seven digits with no spaces or delimiters

Pattern

two letters or digits followed by seven digits:

- two digits or letters (not case-sensitive)
- seven digits

Checksum

Not applicable

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_italy_eu_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_italy_eu_passport_number` is found.
- The regular expression `Regex_italy_eu_passport_date` finds date in the format DD MMM/MMM YYYY (Example - 01 GEN/JAN 1988) or a keyword from `Keywords_eu_passport_date` is found

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_italy_eu_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_italy_eu_passport_number` is found.

XML

```
      <!-- Italy Passport Number -->
      <Entity id="39811019-4750-445f-b26d-4c0e6c431544"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Regex_italy_eu_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_italy_eu_passport_number" />
          </Any>
          <Any minMatches="1">
            <Match idRef="Regex_italy_eu_passport_date" />
            <Match idRef="Keywords_eu_passport_date" />
          </Any>
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_italy_eu_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_italy_eu_passport_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_passport_number_common

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number

- passportnumbers
- passport numbers

Keywords_italy_eu_passport_number

- italiana passaporto
- passaporto italiana
- passaporto numero
- numéro passeport
- numero di passaporto
- numeri del passaporto
- passeport italien

Keywords_eu_passport_date

- date of issue
- date of expiry

Italy physical addresses

Article • 08/19/2024

This unbundled named entity detects patterns related to physical address from Italy. It's also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Italy value added tax number

Article • 08/19/2024

Format

13 character alphanumeric pattern with optional delimiters

Pattern

13 character alphanumeric pattern with optional delimiters:

- I or i
- T or t
- optional space, dot, hyphen, or comma
- 11 digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_italy_value_added_tax_number` finds content that matches the pattern.
- A keyword from `Keywords_italy_value_added_tax_number` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_italy_value_added_tax_number` finds content that matches the pattern.

XML

```
      <!-- Italy Value Added Tax -->
      <Entity id="26a8cc07-2283-4a2a-ab1d-4ab643c4c67f"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_italy_value_added_tax_number" />
          <Match idRef="Keywords_italy_value_added_tax_number" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_italy_value_added_tax_number" />
        </Pattern>
      </Entity>
```

Keywords

Keyword_italy_value_added_tax_number

- vat number
- vat no
- vat#
- iva
- iva#

Japan bank account number

Article • 08/19/2024

Format

seven or eight digits

Pattern

bank account number:

- seven or eight digits
- bank account branch code:
- four digits
- a space or dash (optional)
- three digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_jp_bank_account` finds content that matches the pattern.
- A keyword from `Keyword_jp_bank_account` is found.

- One of the following is true:
- The function `Func_jp_bank_account_branch_code` finds content that matches the pattern.
- A keyword from `Keyword_jp_bank_branch_code` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_jp_bank_account` finds content that matches the pattern.
- A keyword from `Keyword_jp_bank_account` is found.

XML

```
<!-- Japan Bank Account Number -->
<Entity id="d354f95b-96ee-4b80-80bc-4377312b55bc" patternsProximity="300"
recommendedConfidence="75">
  <Version minEngineVersion="15.01.0131.000">
    <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_jp_bank_account" />
          <Match idRef="Keyword_jp_bank_account" />
          <Any minMatches="1">
            <Match idRef="Func_jp_bank_account_branch_code" />
            <Match idRef="Keyword_jp_bank_branch_code" />
          </Any>
      </Pattern>
  </Version>
     <Pattern confidenceLevel="75">
        <IdMatch idRef="Func_jp_bank_account" />
        <Match idRef="Keyword_jp_bank_account" />
    </Pattern>
</Entity>
```

Keywords

Keyword_jp_bank_account

- Checking Account Number
- Checking Account
- Checking Account #
- Checking Acct Number
- Checking Acct #
- Checking Acct No.
- Checking Account No.

- Bank Account Number
- Bank Account
- Bank Account #
- Bank Acct Number
- Bank Acct #
- Bank Acct No.
- Bank Account No.
- Savings Account Number
- Savings Account
- Savings Account #
- Savings Acct Number
- Savings Acct #
- Savings Acct No.
- Savings Account No.
- Debit Account Number
- Debit Account
- Debit Account #
- Debit Acct Number
- Debit Acct #
- Debit Acct No.
- Debit Account No.
- 口座番号
- 銀行口座
- 銀行口座番号
- 総合口座
- 普通預金口座
- 普通口座
- 当座預金口座
- 当座口座
- 預金口座
- 振替口座
- 銀行
- バンク

Keyword_jp_bank_branch_code

- 支店番号
- 支店コード
- 店番号

Japan drivers license number

Article • 08/19/2024

Format

12 digits

Pattern

12 consecutive digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_jp_drivers_license_number` finds content that matches the pattern.
- A keyword from `Keyword_jp_drivers_license_number` is found.

XML

```
<!-- Japan Driver's License Number -->
<Entity id="c6011143-d087-451c-8313-7f6d4aed2270" patternsProximity="300"
recommendedConfidence="75">
    <Pattern confidenceLevel="75">
        <IdMatch idRef="Func_jp_drivers_license_number" />
        <Match idRef ="Keyword_jp_drivers_license_number" />
```

```
    </Pattern>
</Entity>
```

Keywords

Keyword_jp_drivers_license_number

- driverlicense
- driverslicense
- driver'slicense
- driverslicenses
- driver'slicenses
- driverlicenses
- dl#
- dls#
- lic#
- lics#
- 運転免許証
- 運転免許
- 免許証
- 免許
- 運転免許証番号
- 運転免許番号
- 免許証番号
- 免許番号
- 運転免許証ナンバー
- 運転免許ナンバー
- 免許証ナンバー
- 運転免許証no
- 運転免許no
- 免許証no
- 免許no
- 運転経歴証明書番号
- 運転経歴証明書
- 運転免許証No.
- 運転免許No.
- 免許証No.
- 免許No.
- 運転免許証#
- 運転免許#

- 免許証#
- 免許#

Japanese My Number Corporate

Format

13-digit number

Pattern

13-digit number:

- one digit from one to nine
- 12 digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A data loss prevention (DLP) policy has high confidence that it detects this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_japanese_my_number_corporate` finds content that matches the pattern.
- A keyword from `Keywords_japanese_my_number_corporate` is found.

A DLP policy has medium confidence that it detects this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_japanese_my_number_corporate` finds content that matches the pattern.

XML

```
      <!-- Japanese My Number – Corporate -->
      <Entity id="9e0eaf79-ff20-4ffb-b3e4-e7368d5db6ff" patternsProximity="300"
recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_japanese_my_number_corporate" />
          <Match idRef="Keywords_japanese_my_number_corporate" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_japanese_my_number_corporate" />
        </Pattern>
      </Entity>
```

Keywords

Keyword_japan_my_number_corporate

- corporate number
- マイナンバー
- 共通番号
- マイナンバーカード
- マイナンバーカード番号
- 個人番号カード
- 個人識別番号
- 個人識別ナンバー
- 法人番号
- 指定通知書

Last updated on 06/19/2026

Japanese My Number Personal

Format

12-digit number

Pattern

12-digit number:

- four digits
- an optional space, dot, or hyphen
- four digits
- an optional space, dot, or hyphen
- four digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A data loss prevention (DLP) policy has high confidence that it detects this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_japanese_my_number_personal` finds content that matches the pattern.
- A keyword from `Keywords_japanese_my_number_personal` is found.

A DLP policy has low confidence that it detects this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_japanese_my_number_personal` finds content that matches the pattern.

XML

```
      <!-- Japanese My Number - Personal -->
      <Entity id="98da8e66-7299-4ebd-9f82-c871ab37d3ef" patternsProximity="300"
recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_japanese_my_number_personal" />
          <Match idRef="Keywords_japanese_my_number_personal" />
        </Pattern>
        <Pattern confidenceLevel="65">
          <IdMatch idRef="Func_japanese_my_number_personal" />
        </Pattern>
      </Entity>
```

Keywords

Keyword_japan_my_number_personal

- my number
- マイナンバー
- 個人番号
- 共通番号
- マイナンバーカード
- マイナンバーカード番号
- 個人番号カード
- 個人識別番号
- 個人識別ナンバー
- 通知カード

Last updated on 06/19/2026

Japan passport number

Article • 08/19/2024

Format

two letters followed by seven digits

Pattern

two letters (not case-sensitive) followed by seven digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_jp_passport` finds content that matches the pattern.
- A keyword from `Keyword_jp_passport` is found.

XML

```
<!-- Japan Passport Number -->
<Entity id="75177310-1a09-4613-bf6d-833aae3743f8" patternsProximity="300"
recommendedConfidence="75">
    <Pattern confidenceLevel="75">
        <IdMatch idRef="Func_jp_passport" />
        <Match idRef="Keyword_jp_passport" />
```

```
    </Pattern>
</Entity>
```

Keywords

Keyword_jp_passport

- Passport
- Passport Number
- Passport No.
- Passport #
- パスポート
- パスポート番号
- パスポートナンバー
- パスポート＃
- パスポート#
- パスポートNo.
- 旅券番号
- 旅券番号＃
- 旅券番号♯
- 旅券ナンバー

Japan physical addresses

Article • 08/19/2024

This unbundled named entity detects patterns related to physical address from Japan. It's also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Japanese Residence Card Number

Format

Twelve letters and digits

Pattern

Twelve letters and digits:

- two letters (not case-sensitive)
- eight digits
- two letters (not case-sensitive)

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A data loss prevention (DLP) policy has medium confidence that it detects this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_jp_residence_card_number` finds content that matches the pattern.
- A keyword from `Keyword_jp_residence_card_number` is found.

XML

```
<!--Japan Residence Card Number-->
-<Entity id="ac36fef2-a289-4e2c-bb48-b02366e89fc0" recommendedConfidence="75"
```

```
patternsProximity="300">
   -<Pattern confidenceLevel="75">
      <IdMatch idRef="Regex_jp_residence_card_number"/>
      <Match idRef="Keyword_jp_residence_card_number"/>
   </Pattern>
</Entity>
```

Keywords

Keyword_jp_residence_card_number

- Residence card number
- Residence card no
- Residence card #
- 在留カード番号
- 在留カード
- 在留番号

Last updated on 06/19/2026

Japan resident registration number

Article • 08/19/2024

Format

11 digits

Pattern

11 consecutive digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_jp_resident_registration_number` finds content that matches the pattern.
- A keyword from `Keyword_jp_resident_registration_number` is found.

XML

```
<!-- Japan Resident Registration Number -->
<Entity id="01c1209b-6389-4faf-a5f8-3f7e13899652" patternsProximity="300"
recommendedConfidence="75">
    <Pattern confidenceLevel="75">
        <IdMatch idRef="Func_jp_resident_registration_number" />
        <Match idRef ="Keyword_jp_resident_registration_number" />
```

```
    </Pattern>
</Entity>
```

Keywords

Keyword_jp_resident_registration_number

- Resident Registration Number
- Residents Basic Registry Number
- Resident Registration No.
- Resident Register No.
- Residents Basic Registry No.
- Basic Resident Register No.
- 外国人登録証明書番号
- 証明書番号
- 登録番号
- 外国人登録証

Japan social insurance number (SIN)

Article • 08/19/2024

Format

7-12 digits

Pattern

7-12 digits:

- four digits
- a hyphen (optional)
- six digits OR
- 7-12 consecutive digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_jp_sin finds` content that matches the pattern.
- A keyword from `Keyword_jp_sin` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_jp_sin_pre_1997` finds content that matches the pattern.
- A keyword from `Keyword_jp_sin` is found.

XML

```
<!-- Japan Social Insurance Number -->
<Entity id="c840e719-0896-45bb-84fd-1ed5c95e45ff" patternsProximity="300"
recommendedConfidence="75">
    <Pattern confidenceLevel="85">
        <IdMatch idRef="Func_jp_sin" />
        <Match idRef="Keyword_jp_sin" />
    </Pattern>
    <Pattern confidenceLevel="75">
        <IdMatch idRef="Func_jp_sin_pre_1997" />
        <Match idRef="Keyword_jp_sin" />
    </Pattern>
</Entity>
```

Keywords

Keyword_jp_sin

- Social Insurance No.
- Social Insurance Num
- Social Insurance Number
- 健康保険被保険者番号
- 健保番号
- 基礎年金番号
- 雇用保険被保険者番号
- 雇用保険番号
- 保険証番号
- 社会保険番号
- 社会保険No.
- 社会保険
- 介護保険
- 介護保険被保険者番号
- 健康保険被保険者整理番号
- 雇用保険被保険者整理番号
- 厚生年金
- 厚生年金被保険者整理番号

Lab test terms

Article • 08/19/2024

This unbundled named entity detects terms related to lab tests, such as *Insulin C-peptide*. It supports English terms only. It's also included in the All medical terms and conditions bundled named entity SIT.

Confidence level

High

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Latvia drivers license number

Article • 08/19/2024

Format

three letters followed by six digits

Pattern

three letters and six digits:

- three letters (not case-sensitive)
- six digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_latvia_eu_driver's_license_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_driver's_license_number` or `Keywords_latvia_eu_driver's_license_number` is found.

```
XML
```

```
      <!-- Latvia Driver's License Number -->
      <Entity id="ec996de0-30f2-46b1-b192-4d2ff8805fa7"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_latvia_eu_driver's_license_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_driver's_license_number" />
            <Match idRef="Keywords_latvia_eu_driver's_license_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic

- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#

- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen

- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses
- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_latvia_eu_driver's_license_number

- autovadītāja apliecība
- autovadītāja apliecības
- vadītāja apliecība

Latvia passport number

Article • 08/19/2024

Format

two letters or digits followed by seven digits with no spaces or delimiters

Pattern

two letters or digits followed by seven digits:

- two digits or letters (not case-sensitive)
- seven digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_latvia_eu_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_latvia_eu_passport_number` is found.
- The regular expression `Regex_eu_passport_date1` finds date in the format DD.MM.YYYY or a keyword from `Keywords_eu_passport_date` is found

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_latvia_eu_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_latvia_eu_passport_number` is found.

XML

```
      <!-- Latvia Passport Number -->
      <Entity id="23ae25ec-cc28-421b-b77a-3054eadf1ede"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Regex_latvia_eu_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_latvia_eu_passport_number" />
          </Any>
          <Any minMatches="1">
            <Match idRef="Regex_eu_passport_date1" />
            <Match idRef="Keywords_eu_passport_date" />
          </Any>
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_latvia_eu_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_latvia_eu_passport_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_passport_number_common

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number

- passportnumbers
- passport numbers

Keywords_latvia_eu_passport_number

- pase numurs
- pase numur
- pases numuri
- pases nr
- passeport no
- n° du Passeport

Keywords_eu_passport_date

- date of issue
- date of expiry

Latvia personal code

Article • 08/19/2024

Format

11 digits and an optional hyphen

Pattern

Old format

11 digits and a hyphen:

- six digits that correspond to the birth date (DDMMYY)
- a hyphen
- one digit that corresponds to the century of birth ("0" for 19th century, "1" for 20th century, and "2" for 21st century)
- four digits, randomly generated

New format

11 digits

- Two digits "32"
- Nine digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_latvia_eu_national_id_card` or the regex `Regex_latvia_eu_national_id_card_new_format` finds content that matches the pattern.
- A keyword from `Keywords_latvia_eu_national_id_card` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_latvia_eu_national_id_card` or the regex `Regex_latvia_eu_national_id_card_new_format` finds content that matches the pattern.

XML

```
      <!-- Latvia Personal Code -->
      <Entity id="03fcf763-27c2-49ed-9422-2641c6c895c9"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_latvia_eu_national_id_card" />
          <Match idRef="Keywords_latvia_eu_national_id_card" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_latvia_eu_national_id_card" />
          <Any minMatches="0" maxMatches="0">
            <Match idRef="Keywords_latvia_eu_telephone_number" />
            <Match idRef="Keywords_latvia_eu_mobile_number" />
          </Any>
        </Pattern>
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Regex_latvia_eu_national_id_card_new_format" />
          <Match idRef="Keywords_latvia_eu_national_id_card" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_latvia_eu_national_id_card_new_format" />
          <Any minMatches="0" maxMatches="0">
            <Match idRef="Keywords_latvia_eu_telephone_number" />
            <Match idRef="Keywords_latvia_eu_mobile_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_latvia_eu_national_id_card

- administrative number
- alvas nē
- birth number
- citizen number
- civil number
- electronic census number
- electronic number
- fiscal code
- healthcare user number
- id#
- id-code
- identification number
- identifikācijas numurs
- id-number
- individual number
- latvija alva
- nacionālais id
- national id
- national identifying number
- national identity number
- national insurance number
- national register number
- nodokļa numurs
- nodokļu id
- nodokļu identifikācija numurs
- personal certificate number
- personal code
- personal id code
- personal id number
- personal identification code
- personal identifier
- personal identity number
- personal number
- personal numeric code
- personalcodeno#
- personas kods
- population identification code
- public service number
- registration number

- revenue number
- social insurance number
- social security number
- state tax code
- tax file number
- tax id
- tax identification no
- tax identification number
- tax no#
- tax no
- tax number
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id
- tin no
- tin#
- voter's number

Latvia physical addresses

Article • 08/19/2024

This unbundled named entity detects patterns related to physical address from Latvia. It's also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Liechtenstein physical addresses

Article • 08/19/2024

This unbundled named entity detects patterns related to physical address from Liechtenstein. It's also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Lifestyles that relate to medical conditions

Article • 08/19/2024

This unbundled named entity detects terms related to lifestyles that might result in a medical condition, such as *smoking*. It supports English terms only. It's also included in the All medical terms and conditions bundled named entity SIT.

Confidence level

High

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Lithuania drivers license number

Article • 08/19/2024

Format

eight digits without spaces and delimiters

Pattern

eight digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_lithuania_eu_driver's_license_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_driver's_license_number` or `Keywords_lithuania_eu_driver's_license_number` is found.

XML

```
      <!-- Lithuania Driver's License Number -->
      <Entity id="86f7628b-e0f4-4dc3-9fbc-e4300e4c7d78"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_lithuania_eu_driver's_license_number" />
```

```
          <Any minMatches="1">
            <Match idRef="Keywords_eu_driver's_license_number" />
            <Match idRef="Keywords_lithuania_eu_driver's_license_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences

- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#

- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses

- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_lithuania_eu_driver's_license_number

- vairuotojo pažymėjimas
- vairuotojo pažymėjimo numeris
- vairuotojo pažymėjimo numeriai

Lithuania passport number

Article • 08/19/2024

Format

eight digits or letters with no spaces or delimiters

Pattern

eight digits or letters (not case-sensitive)

Checksum

Not applicable

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_lithuania_eu_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_lithuania_eu_passport_number` is found.
- The regular expression `Regex_eu_passport_date3` finds date in the format DD MM YYYY or a keyword from `Keywords_eu_passport_date` is found

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_lithuania_eu_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_lithuania_eu_passport_number` is found.

XML

```
      <!-- Lithuania Passport Number -->
      <Entity id="1b79900f-047b-4c3f-846f-7d73b5534bce"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Regex_lithuania_eu_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_lithuania_eu_passport_number" />
          </Any>
          <Any minMatches="1">
            <Match idRef="Regex_eu_passport_date3" />
            <Match idRef="Keywords_eu_passport_date" />
          </Any>
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_lithuania_eu_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_lithuania_eu_passport_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_passport_number

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers

Keywords_lithuania_eu_passport_number

- paso numeris
- paso numeriai
- paso nr

Keywords_eu_passport_date

- date of issue
- date of expiry

Lithuania personal code

Article • 08/19/2024

Format

11 digits without spaces and delimiters

Pattern

11 digits without spaces and delimiters:

- one digit (1-6) that corresponds to the person's gender and century of birth
- six digits that correspond to birth date (YYMMDD)
- three digits that correspond to the serial number of the date of birth
- one check digit

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_lithuania_eu_tax_file_number` finds content that matches the pattern.
- A keyword from `Keywords_lithuania_eu_tax_file_number` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_lithuania_eu_tax_file_number` finds content that matches the pattern.

XML

```
      <!-- Lithuania Personal Code -->
      <Entity id="cd6d3786-8ec3-4524-a2cf-1e0095379171"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_lithuania_eu_tax_file_number" />
          <Match idRef="Keywords_lithuania_eu_tax_file_number" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_lithuania_eu_tax_file_number" />
          <Any minMatches="0" maxMatches="0">
            <Match idRef="Keywords_lithuania_eu_telephone_number" />
            <Match idRef="Keywords_lithuania_eu_mobile_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_lithuania_eu_national_id_card

- asmeninis skaitmeninis kodas
- asmens kodas
- citizen service number
- mokesčių id
- mokesčių identifikavimas numeris
- mokesčių identifikavimo numeris
- mokesčių numeris
- national identification number
- personal code
- personal numeric code
- piliečio paslaugos numeris
- tax id
- tax identification no
- tax identification number
- tax no#
- tax no
- tax number
- tax registration number

- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id
- tin no
- tin#
- unikalus identifikavimo kodas
- unikalus identifikavimo numeris
- unique identification number
- unique identity number
- uniqueidentityno#

Lithuania physical addresses

Article • 08/19/2024

This unbundled named entity detects patterns related to physical address from Lithuania. It's also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Luxemburg drivers license number

Article • 08/19/2024

Format

six digits without spaces and delimiters

Pattern

six digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_luxemburg_eu_driver's_license_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_driver's_license_number` or `Keywords_luxemburg_eu_driver's_license_number` is found.

XML

```
      <!-- Luxemburg Driver's License Number -->
      <Entity id="89daf717-1544-4860-9a2e-fc9166dd8852"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_luxemburg_eu_driver's_license_number" />
```

```
          <Any minMatches="1">
            <Match idRef="Keywords_eu_driver's_license_number" />
            <Match idRef="Keywords_luxemburg_eu_driver's_license_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences

- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#

- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses

- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_luxemburg_eu_driver's_license_number

- fahrerlaubnis
- Führerschäin

Luxemburg national identification number (natural persons)

Article • 08/19/2024

Format

13 digits with no spaces or delimiters

Pattern

13 digits:

- 11 digits
- two check digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_luxemburg_eu_tax_file_number` finds content that matches the pattern.
- A keyword from `Keywords_luxemburg_eu_national_id_card` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_luxemburg_eu_tax_file_number` finds content that matches the pattern.

XML

```
      <!-- Luxemburg National Identification Number (Natural persons) -->
      <Entity id="aaf661ed-29ec-426d-8bf9-880cad298ebb"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_luxemburg_eu_tax_file_number" />
          <Match idRef="Keywords_luxemburg_eu_national_id_card" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_luxemburg_eu_tax_file_number" />
          <Any minMatches="0" maxMatches="0">
            <Match idRef="Keywords_luxemburg_eu_telephone_number" />
            <Match idRef="Keywords_luxemburg_eu_mobile_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_luxemburg_eu_national_id_card

- eindeutige id
- eindeutige id-nummer
- eindeutigeid#
- id personnelle
- idpersonnelle#
- idpersonnelle
- individual code
- individual id
- individual identification
- individual identity
- numéro d'identification personnel
- personal id
- personal identification
- personal identity
- personalidno#
- personalidnumber#
- persönliche identifikationsnummer
- unique id

- unique identity
- uniqueidkey#

Luxemburg national identification number (non-natural persons)

Article • 08/19/2024

Format

11 digits

Pattern

11 digits

- two digits
- an optional space
- three digits
- an optional space
- three digits
- an optional space
- two digits
- one check digit

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_luxemburg_eu_tax_file_number_non_natural` finds content that matches the pattern.
- A keyword from `Keywords_luxemburg_eu_tax_file_number` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_luxemburg_eu_tax_file_number_non_natural` finds content that matches the pattern.

XML

```
      <!-- Luxemburg National Identification Number (Non-natural persons) --
>
      <Entity id="84bffa3a-d805-4788-a613-b1e4df3804cf"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_luxemburg_eu_tax_file_number_non_natural" />
          <Match idRef="Keywords_luxemburg_eu_tax_file_number" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_luxemburg_eu_tax_file_number_non_natural" />
          <Any minMatches="0" maxMatches="0">
            <Match idRef="Keywords_luxemburg_eu_telephone_number" />
            <Match idRef="Keywords_luxemburg_eu_mobile_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_luxemburg_eu_tax_file_number

- carte de sécurité sociale
- étain non
- étain#
- identifiant d'impôt
- luxembourg tax identifikatiounsnummer
- numéro d'étain
- numéro d'identification fiscal luxembourgeois
- numéro d'identification fiscale
- social security
- sozialunterstützung
- sozialversécherung

- sozialversicherungsausweis
- steier id
- steier identifikatiounsnummer
- steier nummer
- steuer id
- steueridentifikationsnummer
- steuernummer
- tax id
- tax identification no
- tax identification number
- tax no#
- tax no
- tax number
- tax registration number
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id
- tin no
- tin#
- zinn#
- zinn
- zinnzahl

Luxemburg passport number

Article • 08/19/2024

Format

seven or eight alphanumeric characters with no spaces or delimiters

Pattern

seven or eight digits or letters (not case-sensitive)

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_luxemburg_eu_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_luxemburg_eu_passport_number` is found.
- The regular expression `Regex_eu_passport_date3` finds date in the format DD MM YYYY or a keyword from `Keywords_eu_passport_date` is found

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_luxemburg_eu_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_luxemburg_eu_passport_number` is found.

XML

```
      <!-- Luxemburg Passport Number -->
      <Entity id="81d5c027-bed9-4421-91a0-3b2e55b3eb85"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Regex_luxemburg_eu_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_luxemburg_eu_passport_number" />
          </Any>
          <Any minMatches="1">
            <Match idRef="Regex_eu_passport_date3" />
            <Match idRef="Keywords_eu_passport_date" />
          </Any>
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_luxemburg_eu_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_luxemburg_eu_passport_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_passport_number

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers

Keywords_luxemburg_eu_passport_number

- ausweisnummer
- luxembourg pass
- luxembourg passeport
- luxembourg passport
- no de passeport
- no-reisepass
- nr-reisepass
- numéro de passeport
- pass net
- pass nr
- passnummer
- passeport nombre
- reisepässe
- reisepass-nr
- reisepassnummer

Keywords_eu_passport_date

- date of issue
- date of expiry

Luxemburg physical addresses

Article • 08/19/2024

This unbundled named entity detects patterns related to physical address from Luxemburg. It's also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Malaysia identification card number

Article • 08/19/2024

Format

12 digits containing optional hyphens

Pattern

12 digits:

- six digits in the format YYMMDD, which are the date of birth
- a dash (optional)
- two-letter place-of-birth code
- a dash (optional)
- three random digits
- one-digit gender code

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_malaysia_id_card_number` finds content that matches the pattern.
- A keyword from `Keyword_malaysia_id_card_number` is found.

XML

```xml
<!-- Malaysia ID Card Number -->
</Entity>
      <Entity id="7f0e921c-9677-435b-aba2-bb8f1013c749"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
            <IdMatch idRef="Regex_malaysia_id_card_number" />
            <Match idRef="Keyword_malaysia_id_card_number" />
        </Pattern>
</Entity>
```

Keywords

Keyword_malaysia_id_card_number

- digital application card
- i/c
- i/c no
- ic
- ic no
- id card
- identification Card
- identity card
- k/p
- k/p no
- kad akuan diri
- kad aplikasi digital
- kad pengenalan malaysia
- kp
- kp no
- mykad
- mykas
- mykid
- mypr
- mytentera
- malaysia identity card
- malaysian identity card
- nric
- personal identification card

Malaysia passport number

Article • 08/19/2024

Format

Nine alphanumeric characters with only specific values for letters

Pattern

- a letter (A/H/K/a/h/k) followed by
- eight digits

Checksum

Not applicable

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A Data Loss Protection (DLP) policy has high confidence of detection for this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_malaysia_passport_number` finds content that matches the pattern.
- A keyword from `Keyword_malaysia_passport_number` is found.

A DLP policy has low confidence of detection for this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_malaysia_passport_number` finds content that matches the pattern.

XML

```
        <!-- Malaysia Passport Number -->
      <Entity id="62ea2a27-798e-4de3-a31f-a54c5d02ff37"
patternsProximity="300" recommendedConfidence="85" relaxProximity="true">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Regex_malaysia_passport_number" />
          <Match idRef="Keyword_malaysia_passport_number" />
        </Pattern>
        <Pattern confidenceLevel="65">
          <IdMatch idRef="Regex_malaysia_passport_number" />
        </Pattern>
      </Entity>
```

Keywords

Keyword_malaysia_passport_number

- passport#
- passport #
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers
- pasport#
- pasport #
- pasportno
- pasport no
- pasportnumber
- pasport number
- pasportnumbers
- pasport numbers
- no. pasport
- nombor pasport
- nomborpasport

Malta drivers license number

Article • 08/19/2024

Format

Combination of two characters and six digits in the specified pattern

Pattern

combination of two characters and six digits:

- two characters (digits or letters, not case-sensitive)
- a space (optional)
- three digits
- a space (optional)
- three digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_malta_eu_driver's_license_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_driver's_license_number` or `Keywords_malta_eu_driver's_license_number` is found.

XML

```xml
      <!-- Malta Driver's License Number -->
      <Entity id="a3bdaa4a-8371-4735-8fa5-56ee0fb4afc4"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_malta_eu_driver's_license_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_driver's_license_number" />
            <Match idRef="Keywords_malta_eu_driver's_license_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences

- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#

- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences

- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses
- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_malta_eu_driver's_license_number

- liċenzja tas-sewqan
- liċenzji tas-sewwieq

Malta identity card number

Article • 08/19/2024

Format

seven digits followed by one letter

Pattern

seven digits followed by one letter:

- seven digits
- one letter in "M, G, A, P, L, H, B, Z" (case insensitive)

Checksum

Not applicable

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_malta_eu_national_id_card` finds content that matches the pattern.
- A keyword from `Keywords_malta_eu_national_id_card` is found.

A DLP policy has low confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_malta_eu_national_id_card` finds content that matches the pattern.

XML

```
      <!-- Malta Identity Card Number -->
      <Entity id="854b36b3-a388-4ac8-a4ec-677c2b5e4356"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_malta_eu_national_id_card" />
          <Match idRef="Keywords_malta_eu_national_id_card" />
        </Pattern>
        <Pattern confidenceLevel="65">
          <IdMatch idRef="Regex_malta_eu_national_id_card" />
        </Pattern>
      </Entity>
```

Keywords

Keywords_malta_eu_national_id_card

- citizen service number
- id tat-taxxa
- identifika numru tal-biljett
- kodiċi numerali personali
- numru ta 'identifikazzjoni personali
- numru ta 'identifikazzjoni tat-taxxa
- numru ta 'identifikazzjoni uniku
- numru ta' identità uniku
- numru tas-servizz taċ-ċittadin
- numru tat-taxxa
- personal numeric code
- unique identification number
- unique identity number
- uniqueidentityno#

Malta passport number

Article • 08/19/2024

Format

seven digits without spaces or delimiters

Pattern

seven digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_malta_eu_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_malta_eu_passport_number` is found.
- A keyword from `Keywords_eu_passport_date` is found

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_malta_eu_passport_number` finds content that matches the pattern.

- A keyword from `Keywords_eu_passport_number` or `Keywords_malta_eu_passport_number` is found.

XML

```
      <!-- Malta Passport Number -->
      <Entity id="b2b21198-48f9-4d13-b2a5-03969bff0fb8"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Regex_malta_eu_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_malta_eu_passport_number" />
          </Any>
          <Match idRef="Keywords_eu_passport_date" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_malta_eu_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_malta_eu_passport_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_passport_number

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers

Keywords_malta_eu_passport_number

- numru tal-passaport
- numri tal-passaport

- Nru tal-passaport

Keywords_eu_passport_date

- date of issue
- date of expiry

Malta physical addresses

Article • 08/19/2024

This unbundled named entity detects patterns related to physical address from Malta. It's also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Malta tax identification number

Article • 08/19/2024

Format

For Maltese nationals:

- seven digits and one letter in the specified pattern

Non-Maltese nationals and Maltese entities:

- nine digits

Pattern

Maltese nationals: seven digits and one letter

- seven digits
- one letter (not case-sensitive)

Non-Maltese nationals and Maltese entities: nine digits

- nine digits

Checksum

Not applicable

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regex `Regex_malta_eu_tax_file_number` or `Regex_malta_eu_tax_file_number_non_maltese_national` finds content that matches the pattern.
- A keyword from `Keywords_malta_eu_tax_file_number` is found.

A DLP policy has low confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regex `Regex_malta_eu_tax_file_number` or `Regex_malta_eu_tax_file_number_non_maltese_national` finds content that matches the pattern.

XML

```
      <!-- Malta Tax ID Number -->
      <Entity id="ec830c63-65f4-45d0-9d8c-910dc8334b20"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_malta_eu_tax_file_number" />
          <Match idRef="Keywords_malta_eu_tax_file_number" />
        </Pattern>
        <Pattern confidenceLevel="65">
          <IdMatch idRef="Regex_malta_eu_tax_file_number" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch
idRef="Regex_malta_eu_tax_file_number_non_maltese_national" />
          <Match idRef="Keywords_malta_eu_tax_file_number" />
        </Pattern>
        <Pattern confidenceLevel="65">
          <IdMatch
idRef="Regex_malta_eu_tax_file_number_non_maltese_national" />
        </Pattern>
      </Entity>
```

Keywords

Keywords_malta_eu_tax_file_number

- citizen service number
- id tat-taxxa
- identifika numru tal-biljett
- kodiċi numerali personali
- numru ta 'identifikazzjoni personali
- numru ta 'identifikazzjoni tat-taxxa

- numru ta 'identifikazzjoni uniku
- numru ta' identità uniku
- numru tas-servizz taċ-ċittadin
- numru tat-taxxa
- personal numeric code
- tax id
- tax identification no
- tax identification number
- tax no#
- tax no
- tax number
- tax registration number
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id
- tin no
- tin#
- unique identification number
- unique identity number
- uniqueidentityno#

Medical specialties

Article • 08/19/2024

This unbundled named entity detects terms related to medical specialties, such as *dermatology*. It supports English terms only. It's also included in the All medical terms and conditions bundled named entity SIT.

Confidence level

High

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Medicare Beneficiary Identifier (MBI) card

Article • 08/19/2024

Format

11 character alphanumeric pattern

Pattern

- one digit between 1 to 9
- one letter excluding S, L, O, I, B, Z
- one digit or letter excluding S, L, O, I, B, Z
- one digit
- an optional Hyphen
- one letter excluding S, L, O, I, B, Z
- one digit or letter excluding S, L, O, I, B, Z
- one digit
- an optional Hyphen
- two letters excluding S, L, O, I, B, Z
- two digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_mbi_card` finds content that matches the pattern.
- A keyword from `Keyword_mbi_card` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_mbi_card` finds content that matches the pattern.

XML

```
    <!-- Medicare Beneficiary Identifier (MBI) card -->
      <Entity id="f753a286-f5cc-47e6-a592-4be25fd02591"
patternsProximity="300" recommendedConfidence="75" relaxProximity="true">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Regex_mbi_card" />
          <Match idRef="Keyword_mbi_card" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_mbi_card" />
        </Pattern>
      </Entity>
```

Keywords

Keyword_mbi_card

- mbi
- mbi#
- medicare beneficiary #
- medicare beneficiary identifier
- medicare beneficiary no
- medicare beneficiary number
- medicare beneficiary#

Mexico Unique Population Registry Code (CURP)

Article • 08/19/2024

Format

18 character alphanumeric pattern

Pattern

- four letters (case insensitive)
- six digits indicating a valid date
- a letter - H/h or M/m
- two letters indicating a valid Mexican state code
- three letters
- one letter or digit
- one digit

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_mexico_population_registry_code` finds content that matches the pattern.

- A keyword from `Keyword_mexico_population_registry_code` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_mexico_population_registry_code` finds content that matches the pattern.

XML

```
    <!-- Mexico Unique Population Registry Code (CURP) -->
      <Entity id="e905ad4d-5a74-406d-bf36-b1efca798af4"
patternsProximity="300" recommendedConfidence="75" relaxProximity="true">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_mexico_population_registry_code" />
          <Match idRef="Keyword_mexico_population_registry_code" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_mexico_population_registry_code" />
        </Pattern>
      </Entity>
```

Keywords

Keyword_mexico_population_registry_code

- Clave Única de Registro de Población
- Clave Unica de Registro de Poblacion
- Unique Population Registry Code
- unique population code
- CURP
- Personal ID
- Unique ID
- personalid
- personalidnumber
- uniqueidkey
- uniqueidnumber
- clave única
- clave unica
- clave personal Identidad
- personal Identidad Clave
- ClaveÚnica
- claveunica

- clavepersonalIdentidad

Microsoft Bing maps key

This SIT is also included in the All credentials bundled SIT.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a credential scanning SIT in an endpoint DLP policy. This requirement is specific to the combination of credential scanning SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

Format

A combination of 64 characters consisting of letters, digits, and special characters.

Pattern

A combination of 64 characters:

- a-z (not case-sensitive)
- 0-9
- underlines (_)
- or hyphens (-)

format example:

```
aB1cD2eF-3gH4iJ5kL6_mN7oP8qR
```

Credential format example

```
bingMapsKey=aB1cD2eF-3gH4iJ5kL6-mN7oP8qR
```

Checksum

No

SITs that have checksums use a unique calculation to check if the information is valid. This means when the **Checksum** value is **Yes**, the service can make a positive detection based on the sensitive data alone. When the **Checksum** value is **No** additional (secondary) elements must also be detected for the service to make a positive detection.

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

This SIT is designed to match the security information that's used to call Bing Maps API.

It uses several primary resources:

- Patterns of Base64 URL encoded 384-bits symmetric key.
- Patterns of CredentialName, CredentialFeatures, AccountIdentityName, AccountIdentityValue, ResourceType, ResourceName.
- Patterns of mockup values, redactions, and placeholders.
- A dictionary of vocabulary.

The patterns are designed to match actual credentials with reasonable confidence. The patterns don't match credentials formatted as examples. Mockup values, redacted values, and placeholders, like credential type or usage descriptions, in the position where an actual secret value should present won't be matched.

Keywords

Keyword_SymmetricKey384Url:

- virtualearth
- api/maps
- key

Last updated on 11/18/2025

Microsoft Entra client access token

This SIT is also included in the All credentials bundled SIT.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a credential scanning SIT in an endpoint DLP policy. This requirement is specific to the combination of credential scanning SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

Format

A combination of up to 10,000 characters consisting of letters, digits, and special characters.

or

A client secret or refresh token used in OAuth2.0 protocol.

or

A combination of up to 1,000 characters consisting of letters, digits, and special characters.

Pattern

Any combination of:

- up to 10,000
- a-z (not case-sensitive)
- 0-9
- forward slashes (/)
- or plus signs (+)
- Up to 2
- equal signs (=)

for example:

`"VersionProfile": null, "TokenCache": { "CacheData": A1bC2dE3fH4iJ5kL6mN7oP8qR9sT0u

or

Variant client secret or refresh token formats for example.
`ClientSecret:********`
`AppSecret=********`
`ConsumerKey:=********`
`Refresh_Token:********`

or

3 letters: eyJ (case-sensitive)

And

A combination of up to 1,000 characters consisting

- a-z (not case-sensitive)
- 0-9
- dashes (-)
- underlines (_)

- or dots (.)

format example:

`A1bC2dE3fH4iJ5kL6mN7oP8qR9sT0u`

Credential example

Expand table

Confidence Band	Example
High	`Authorization: Bearer eyJ0eXAiOiJKV1QiLCJhbGciOiJSUzI1NiIsIm5vbmNlIjoibm9uY2UifQ.eyJhdWQiOiJodHRwczovL2hvc3QiLCJleHAiOjk5OTk5OTk5OTksImFwcGlkI`
Medium	`Authorization: Bearer eyJ0eXAiOiJKV1QiLCJhbGciOiJSUzI1NiJ9.eyJhdWQiOiJodHRwczovL2hvc3QiLCJleHAiOjk5OTk5OTk5OTksImFwcGlkIjoiaWQiLCJ2ZXIiOiIxLjA`
Low	`Authorization: Bearer eyJ0eXAiOiJKV1QiLCJhbGciOiJSUzI1NiJ9.eyJhdWQiOiJodHRwczovL2hvc3QiLCJleHAiOjk5OTk5OTk5OTl9.UVQYT1FE`

Checksum

Yes

SITs that have checksums use a unique calculation to check if the information is valid. This means when the **Checksum** value is **Yes**, the service can make a positive detection based on the sensitive data alone. When the **Checksum** value is **No** additional (secondary) elements must also be detected for the service to make a positive detection.

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

This SIT is designed to match the security information that's contains claims that one can use in Azure Active Directory B2C (Azure AD B2C) to identify the granted permissions to Azure resources.

It uses several primary resources:

- Patterns of Azure PowerShell Token Cache
- Patterns of Client secret context
- Patterns of Json Web Token
- Patterns of CredentialName, CredentialFeatures, AccountIdentityName, AccountIdentityValue, ResourceType, ResourceName
- Patterns of mockup values, redactions, and placeholders
- A dictionary of vocabulary

The patterns are designed to match actual credentials with reasonable confidence. The patterns don't match credentials formatted as examples. Mockup values, redacted values, and placeholders, like credential type or usage descriptions, in the position where an actual secret value should present won't be matched.

Keywords

Keyword_SymmetricKeyContextInXml:

- tokencache

Keyword_ClientSecretContext:

- secret
- token
- auth
- securestring
- key

Keyword_JsonWebToken:

- eyJ

Last updated on 11/18/2025

Microsoft Entra client secret

This SIT is also included in the All credentials bundled SIT.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a credential scanning SIT in an endpoint DLP policy. This requirement is specific to the combination of credential scanning SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

Format

A combination of up to 40 characters consisting of letters, digits, and special characters.

Pattern

A combination of up to 40 characters consisting of:

- a-z (not case-sensitive)
- 0-9
- dashes (-)
- underlines (_)
- dots (.)
- or tilde accents (~)

for example:

`Bb2Cc~3Dd4.-Ee5Ff6Gg7Hh8Ii9Jj0_Kk1Ll2Mm3`

Credential example

Expand table

Confidence Band	Example
High	`"TenentId=aaaa0a0a-bb1b-cc2c-dd3d-eeeeee4e4e4e;AppId=22223333-cccc-4444-dddd-5555eeee6666;AppSecret=Cc3Dd~4Ee5.-Ff6Gg7Hh8Ii9Jj0Kk1_Ll2Mm3Nn4"`
Medium	`AppId=22223333-cccc-4444-dddd-5555eeee6666;AppSecret=Dd4Ee~5Ff6.-Gg7Hh8Ii9Jj0Kk1Ll2_Mm3Nn4Oo5`
Low	N/A

Checksum

Yes

SITs that have checksums use a unique calculation to check if the information is valid. This means when the **Checksum** value is **Yes**, the service can make a positive detection based on the sensitive data alone. When the **Checksum** value is **No** additional (secondary) elements must also be detected for the service to make a positive detection.

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

This SIT is designed to match the security information that's used to secure Microsoft Entra service principals.

It uses several primary resources:

- Patterns of Client Secret with specific format.
- Patterns of CredentialName, CredentialFeatures, AccountIdentityName, AccountIdentityValue, ResourceType, ResourceName.
- Patterns of mockup values, redactions, and placeholders.
- A dictionary of vocabulary.

The patterns are designed to match actual credentials with reasonable confidence. The patterns don't match credentials formatted as examples. Mockup values, redacted values, and placeholders, like credential type or usage descriptions, in the position where an actual secret value should present won't be matched.

Keywords

Keyword_AppSecret:

- secret
- password
- key

Last updated on 11/18/2025

Microsoft Entra user credentials

This SIT is also included in the All credentials bundled SIT.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a credential scanning SIT in an endpoint DLP policy. This requirement is specific to the combination of credential scanning SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

Format

A paired username and password related to *.onmicrosoft.com domain.

or

Plain-text password used in code snippets.

or

Plain-text password used in XML configuration.

Pattern

Various username and password formats, for example:

`username=...password=********`

`/user:.../pass:********`

`SharePointOnlineAuthenticatedContext`

`sign_in`

or

Various password formats in code snippets, for example:

`new X509Certificates2( ...`

`ConvertTo-SecureString -String ********...`

`password = "********"...`

`"password" : "********"...`

`UserPasswordCredential( ...`

or

Various password formats in XML, for example:

```xml
... <secret>********</secret> ...
... <password>********</password> ...
... <setting name="password" value="********" > ...
... <setting name="password">********</setting> ...
... <setting name="password"><value>********</value></setting> ...
```

Credential example

`username=user@contoso.onmicrosoft.com;password=********`

Checksum

No

SITs that have checksums use a unique calculation to check if the information is valid. This means when the **Checksum** value is **Yes**, the service can make a positive detection based on the sensitive data alone. When the **Checksum** value is **No** additional (secondary) elements must also be detected for the service to make a positive detection.

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

This SIT is designed to match the security information that's used as individual user passwords to authenticate against Microsoft Entra ID.

It uses several primary resources:

- Patterns of Plain-text username and password for Microsoft Entra tenants.
- Patterns of Password context in code.
- Patterns of Password context in XML.
- Patterns of CredentialName, CredentialFeatures, AccountIdentityName, AccountIdentityValue, ResourceType, ResourceName.
- Patterns of mockup values, redactions, and placeholders.
- A dictionary of vocabulary.

The patterns are designed to match actual credentials with reasonable confidence. The patterns don't match credentials formatted as examples. Mockup values, redacted values, and placeholders, like credential type or usage descriptions, in the position where an actual secret value should present won't be matched.

Keywords

Keyword_AzureActiveDirectoryLoginCredentials:

- password
- pw
- userpass
- credentials
- cmdkey
- Authenti
- sign_in

Keyword_PasswordContextInCode:

- key
- x509c
- credential
- password
- pw
- securestring

Keyword_PasswordContextInXml:

- userpass

- password
- pw
- connectionstring
- key
- credential
- token
- sas
- secret

Last updated on 06/01/2026

Netherlands citizens service (BSN) number

Article • 08/19/2024

Format

eight or nine digits containing optional spaces

Pattern

eight-nine digits:

- three digits
- a space (optional)
- three digits
- a space (optional)
- two-three digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_netherlands_bsn finds` content that matches the pattern.
- A keyword from `Keyword_netherlands_bsn` is found.
- The checksum passes.

XML

```xml
      <!-- Netherlands Citizen's Service (BSN) Number -->
      <Entity id="c5f54253-ef7e-44f6-a578-440ed67e946d"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_netherlands_bsn" />
          <Match idRef="Keywords_netherlands_eu_national_id_card" />
        </Pattern>
      </Entity>
```

Keywords

Keywords_netherlands_eu_national_id_card

- bsn#
- bsn
- burgerservicenummer
- citizen service number
- person number
- personal number
- personal numeric code
- person-related number
- persoonlijk nummer
- persoonlijke numerieke code
- persoonsgebonden
- persoonsnummer
- sociaal-fiscaal nummer
- social-fiscal number
- sofi
- sofinummer
- uniek identificatienummer
- uniek identiteitsnummer
- unique identification number
- unique identity number
- uniqueidentityno#

Netherlands drivers license number

Article • 08/19/2024

Format

10 digits without spaces and delimiters

Pattern

10 digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_netherlands_eu_driver's_license_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_driver's_license_number` or `Keywords_netherlands_eu_driver's_license_number` is found.

XML

```
      <!-- Netherlands Driver's License Number -->
      <Entity id="6247fbea-ab80-4be5-8233-308b7c031401"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_netherlands_eu_driver's_license_number" />
```

```
          <Any minMatches="1">
            <Match idRef="Keywords_eu_driver's_license_number" />
            <Match idRef="Keywords_netherlands_eu_driver's_license_number"
/>
            </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence

- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#

- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen

- driving licenses
- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_netherlands_eu_driver's_license_number

- permis de conduire
- rijbewijs
- rijbewijsnummer
- rijbewijzen
- rijbewijs nummer
- rijbewijsnummers

Netherlands passport number

Article • 08/19/2024

Format

nine letters or digits with no spaces or delimiters

Pattern

nine letters or digits

Checksum

Not applicable

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_netherlands_eu_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_netherlands_eu_passport_number` is found.
- The regular expression `Regex_netherlands_eu_passport_date` finds date in the format DD MMM/MMM YYYY (Example - 26 MAA/MAR 2012)

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_netherlands_eu_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_netherlands_eu_passport_number` is found.

XML

```
      <!-- Netherlands Passport Number -->
      <Entity id="61786727-bafd-45f6-94d9-888d815e228e"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Regex_netherlands_eu_passport_number" />
          <Match idRef="Regex_netherlands_eu_passport_date" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_netherlands_eu_passport_number" />
          </Any>
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_netherlands_eu_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_netherlands_eu_passport_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_passport_number

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers

Keywords_netherlands_eu_passport_number

- paspoort nummer
- paspoortnummers
- paspoortnummer
- paspoort nr

Netherlands physical addresses

Article • 08/19/2024

This unbundled named entity detects patterns related to physical address from the Netherlands. It's also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Netherlands tax identification number

Article • 08/19/2024

Format

nine digits without spaces or delimiters

Pattern

nine digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_netherlands_eu_tax_file_number` finds content that matches the pattern.
- A keyword from `Keywords_netherlands_eu_tax_file_number` is found.

A DLP policy has low confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_netherlands_eu_tax_file_number` finds content that matches the pattern.

```
XML
```

```
      <!-- Netherlands Tax Identification Number -->
      <Entity id="01f42a64-eba7-4892-a67b-398237e4ade2"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_netherlands_eu_tax_file_number" />
          <Match idRef="Keywords_netherlands_eu_tax_file_number" />
        </Pattern>
        <Pattern confidenceLevel="65">
          <IdMatch idRef="Func_netherlands_eu_tax_file_number" />
        </Pattern>
      </Entity>
```

Keywords

Keywords_netherlands_eu_tax_file_number

- btw nummer
- hollânske tax identification
- hulandes impuesto id number
- hulandes impuesto identification
- identificatienummer belasting
- identificatienummer van belasting
- impuesto identification number
- impuesto number
- nederlands belasting id nummer
- nederlands belasting identificatie
- nederlands belasting identificatienummer
- nederlands belastingnummer
- nederlandse belasting identificatie
- netherlands tax identification
- netherland's tax identification
- netherlands tin
- netherland's tin
- tax id
- tax identification no
- tax identification number
- tax identification tal
- tax no#
- tax no
- tax number
- tax registration number

- tax tal
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id
- tin no
- tin#

Netherlands value added tax number

Article • 08/19/2024

Format

14 character alphanumeric pattern

Pattern

14-character alphanumeric pattern:

- N or n
- L or l
- optional space, dot, or hyphen
- nine digits
- optional space, dot, or hyphen
- B or b
- two digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_netherlands_value_added_tax_number` finds content that matches the pattern.

- A keyword from `Keywords_netherlands_value_added_tax_number` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_netherlands_value_added_tax_number` finds content that matches the pattern.

XML

```
      <!-- Netherlands Value Added Tax Number -->
      <Entity id="4f320d9b-4972-41ae-b337-88d499bb1ade"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_netherlands_value_added_tax_number" />
          <Match idRef="Keywords_netherlands_value_added_tax_number" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_netherlands_value_added_tax_number" />
        </Pattern>
      </Entity>
```

Keywords

Keyword_netherlands_value_added_tax_number

- vat number
- vat no
- vat#
- wearde tafoege tax getal
- btw nûmer
- btw-nummer

New Zealand bank account number

Article • 08/19/2024

Format

14-digit to 16-digit pattern with optional delimiter

Pattern

14-digit to 16-digit pattern with optional delimiter:

- two digits
- an optional hyphen or space
- three to four digits
- an optional hyphen or space
- seven digits
- an optional hyphen or space
- two to three digits
- an options hyphen or space

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_new_zealand_bank_account_number` finds content that matches the pattern.

- A keyword from `Keywords_new_zealand_bank_account_number` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_new_zealand_bank_account_number` finds content that matches the pattern.

XML

```
      <!-- New Zealand Bank Account Number -->
      <Entity id="1a97fc2b-dd2f-48f1-bc4e-2ddf25813956"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_new_zealand_bank_account_number" />
          <Match idRef="Keywords_new_zFealand_bank_account_number" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_new_zealand_bank_account_number" />
        </Pattern>
      </Entity>
```

Keywords

Keyword_new_zealand_bank_account_number

- account number
- bank account
- bank_acct_id
- bank_acct_branch
- bank_acct_nbr

New Zealand drivers license number

Article • 08/19/2024

Format

eight character alphanumeric pattern

Pattern

eight character alphanumeric pattern

- two letters
- six digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_newzealand_driver_license_number` finds content that matches the pattern.
- A keyword from `Keywords_newzealand_driver_license_number` is found.

A DLP policy has low confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_newzealand_driver_license_number` finds content that matches the pattern.

XML

```
      <!-- New Zealand Driver License Number -->
      <Entity id="1924b377-d287-49c9-a737-cfe7a8a2615a"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_newzealand_driver_license_number" />
          <Match idRef="Keywords_newzealand_driver_license_number" />
        </Pattern>
        <Pattern confidenceLevel="65">
          <IdMatch idRef="Func_newzealand_driver_license_number" />
        </Pattern>
      </Entity>
```

Keywords

Keyword_new_zealand_drivers_license_number

- driverlicence
- driverlicences
- driver lic
- driver licence
- driver licences
- driverslic
- driverslicence
- driverslicences
- drivers lic
- drivers lics
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'licence
- driver'licences
- driver' lic
- driver' lics
- driver' licence
- driver' licences
- driver'slic

- driver'slics
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's licence
- driver's licences
- driverlic#
- driverlics#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver licence#
- driver licences#
- driverslic#
- driverslics#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's licence#
- driver's licences#
- international driving permit
- international driving permits

- nz automobile association
- new zealand automobile association

New Zealand inland revenue number

Article • 08/19/2024

Format

eight or nine digits with optional delimiters

Pattern

eight or nine digits with optional delimiters

- two or three digits
- an optional space or hyphen
- three digits
- an optional space or hyphen
- three digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_new_zealand_inland_revenue_number` finds content that matches the pattern.
- A keyword from `Keywords_new_zealand_inland_revenue_number` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_new_zealand_inland_revenue_number` finds content that matches the pattern.

XML

```
      <!-- New Zealand Inland Revenue Number -->
      <Entity id="dd0fe2bc-7bcf-455f-bac1-83b1e3eb25fd"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_new_zealand_inland_revenue_number" />
          <Match idRef="Keywords_new_zealand_inland_revenue_number" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_new_zealand_inland_revenue_number" />
        </Pattern>
      </Entity>
```

Keywords

Keyword_new_zealand_inland_revenue_number

- ird no.
- ird no#
- nz ird
- new zealand ird
- ird number
- inland revenue number

New Zealand ministry of health number

Article • 08/19/2024

Format

three letters and four digits

Pattern

- three letters (not case-sensitive) except 'I' and 'O'
- four digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_new_zealand_ministry_of_health_number` finds content that matches the pattern.
- A keyword from `Keyword_nz_terms` is found.
- The checksum passes.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_new_zealand_ministry_of_health_number` finds content that matches the pattern.

- The checksum passes.

XML

```
    <!-- New Zealand Health Number -->
    <Entity id="2b71c1c8-d14e-4430-82dc-fd1ed6bf05c7"
patternsProximity="300" recommendedConfidence="85">
      <Pattern confidenceLevel="85">
        <IdMatch idRef="Func_new_zealand_ministry_of_health_number" />
          <Match idRef="Keyword_nz_terms" />
      </Pattern>
      <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_new_zealand_ministry_of_health_number" />
       </Pattern>
    </Entity>
```

Keywords

Keyword_nz_terms

- NHI
- New Zealand
- National Health Index
- NHI#
- National Health Index#

New Zealand physical addresses

Article • 08/19/2024

This unbundled named entity detects patterns related to physical address from New Zealand. It's also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

New Zealand social welfare number

Article • 08/19/2024

Format

nine digits

Pattern

nine digits

- three digits
- an optional hyphen
- three digits
- an optional hyphen
- three digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_newzealand_social_welfare_number` finds content that matches the pattern.
- A keyword from `Keywords_newzealand_social_welfare_number` is found.

A DLP policy has low confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_newzealand_social_welfare_number` finds content that matches the pattern.

XML

```
      <!-- Newzealand Social Welfare Number -->
      <Entity id="20f3c48d-4ac1-4cd2-86bd-34ecc1826e9d"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_newzealand_social_welfare_number" />
          <Match idRef="Keywords_newzealand_social_welfare_number" />
        </Pattern>
        <Pattern confidenceLevel="65">
          <IdMatch idRef="Func_newzealand_social_welfare_number" />
        </Pattern>
      </Entity>
    </Version>
```

Keywords

Keyword_new_zealand_social_welfare_number

- social welfare #
- social welfare#
- social welfare No.
- social welfare number
- swn#

Norway identification number

Article • 08/19/2024

Format

11 digits

Pattern

11 digits:

- six digits in the format DDMMYY, which are the date of birth
- three-digit individual number
- two check digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_norway_id_number` finds content that matches the pattern.
- A keyword from `Keyword_norway_id_number` is found.
- The checksum passes.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_norway_id_number` finds content that matches the pattern.
- The checksum passes.

XML

```xml
<!-- Norway Identification Number -->
<Entity id="d4c8a798-e9f2-4bd3-9652-500d24080fc3" recommendedConfidence="85"
patternsProximity="300">
  <Pattern confidenceLevel="85">
     <IdMatch idRef="Func_norway_id_number"/>
     <Match idRef="Keyword_norway_id_number"/>
  </Pattern>
  <Pattern confidenceLevel="75">
     <IdMatch idRef="Func_norway_id_number"/>
  </Pattern>
</Entity>
```

Keywords

Keyword_norway_id_number

- Personal identification number
- Norwegian ID Number
- ID Number
- Identification
- Personnummer
- Fødselsnummer

Norway physical addresses

Article • 08/19/2024

This unbundled named entity detects patterns related to physical address from Norway. It's also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Philippines national identification number

Article • 08/19/2024

Format

Twelve digits plus optional delimiters

Pattern

- four digits
- followed by hyphen (optional)
- four digits
- followed by hyphen (optional)
- four digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regx `Regex_phillipines_national_id` finds content that matches the pattern.
- A keyword from `Keyword_phillipines_national_id` is found.

A DLP policy has low confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regx `Regex_phillipines_national_id` finds content that matches the pattern.

XML

```
<!-- Philippines National ID -->
    <Entity id="3257da62-3d98-44ca-a923-6f8b0cd94b7f"
patternsProximity="300" recommendedConfidence="75" relaxProximity="true">
      <Pattern confidenceLevel="75">
        <IdMatch idRef="Regex_phillipines_national_id" />
        <Match idRef="Keyword_phillipines_national_id" />
      </Pattern>
      <Pattern confidenceLevel="65">
        <IdMatch idRef="Regex_phillipines_national_id" />
      </Pattern>
    </Entity>
```

Keywords

Keyword_phillipines_national_id

- PhilSys
- Phil ID
- National ID
- PSN
- PhilSys Number
- ID numner
- National identification number

Philippines passport number

Article • 08/19/2024

Format

seven or eight or nine alphanumeric characters

Pattern

Seven or eight or nine alphanumeric characters:

- one letter followed by six digits or
- two letters followed by six digits or
- two letters followed by seven digits or
- one letter followed by seven digits followed by one letter.

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression Regex_philippines_passport_number finds content that matches the pattern.
- A keyword from Keyword_philippines_passport_number is found.

A DLP policy has low confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression Regex_philippines_passport_number finds content that matches the pattern.

XML

```
     <!-- Philippines Passport Number -->
     <Entity id="6fa57f91-314a-4561-8248-7ab921957448"
patternsProximity="300" recommendedConfidence="85" relaxProximity="true"
filters="philippines_passport_filter">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Regex_philippines_passport_number" />
          <Match idRef="Keyword_philippines_passport_number" />
        </Pattern>
        <Pattern confidenceLevel="65">
          <IdMatch idRef="Regex_philippines_passport_number" />
        </Pattern>
     </Entity>
```

Keywords

Keyword_philippines_passport_number

- passport
- passport no
- passport number
- passport#
- passportno
- passport no.
- passportno#
- passport book
- passportbook#
- pasaporte
- numero ng pasaporte
- libro ng pasaporte

Philippines unified multi-purpose identification number

Article • 08/19/2024

Format

12 digits separated by hyphens

Pattern

12 digits:

- four digits
- a hyphen
- seven digits
- a hyphen
- one digit

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_philippines_unified_id` finds content that matches the pattern.
- A keyword from `Keyword_philippines_id` is found.

XML

```xml
<!-- Philippines Unified Multi-Purpose ID number -->
<Entity id="019b39dd-8c25-4765-91a3-d9c6baf3c3b3" recommendedConfidence="75"
patternsProximity="300">
  <Pattern confidenceLevel="75">
     <IdMatch idRef="Regex_philippines_unified_id"/>
     <Match idRef="Keyword_philippines_id"/>
  </Pattern>
</Entity>
```

Keywords

Keyword_philippines_id

- Unified Multi-Purpose ID
- UMID
- Identity Card
- Pinag-isang Multi-Layunin ID

Poland drivers license number

Article • 08/19/2024

Format

11 or 14 digits containing two forward slashes

Pattern

11 or 14 digits containing two forward slashes

- five digits
- a forward slash
- two digits
- a forward slash
- four or seven digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_poland_eu_driver's_license_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_driver's_license_number` or `Keywords_poland_eu_driver's_license_number` is found.

XML

```xml
      <!-- Poland Driver's License Number -->
      <Entity id="24d51f99-ee9e-4060-a077-cae58cab1ee4"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_poland_eu_driver's_license_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_driver's_license_number" />
            <Match idRef="Keywords_poland_eu_driver's_license_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences

- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#

- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences

- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses
- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_poland_eu_driver's_license_number

- prawo jazdy
- prawa jazdy

Poland identity card

Article • 08/19/2024

Format

three letters and six digits

Pattern

three letters (not case-sensitive) followed by six digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_polish_national_id` finds content that matches the pattern.
- A keyword from `Keyword_polish_national_id_passport_number` is found.
- The checksum passes.

XML

```
<!-- Poland Identity Card-->
<Entity id="25E64989-ED5D-40CA-A939-6C14183BB7BF" patternsProximity="300"
recommendedConfidence="85">
      <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_polish_national_id" />
          <Match idRef="Keyword_polish_national_id_passport_number" />
```

```
        </Pattern>
    </Entity>
```

Keywords

Keyword_poland_national_id_passport_number

- Dowód osobisty
- Numer dowodu osobistego
- Nazwa i numer dowodu osobistego
- Nazwa i nr dowodu osobistego
- Nazwa i nr dowodu tożsamości
- Dowód Tożsamości
- dow. os.

Poland national ID (PESEL)

Article • 08/19/2024

Format

11 digits

Pattern

- six digits representing date of birth in the format YYMMDD
- four digits
- one check digit

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_pesel_identification_number` finds content that matches the pattern.
- A keyword from `Keyword_pesel_identification_number` is found.
- The checksum passes.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_pesel_identification_number` finds content that matches the pattern.
- The checksum passes.

XML

```
      <!-- Poland National ID (PESEL) -->
      <Entity id="E3AAF206-4297-412F-9E06-BA8487E22456"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_pesel_identification_number" />
          <Match idRef="Keyword_pesel_identification_number" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_pesel_identification_number" />
        </Pattern>
      </Entity>
```

Keywords

Keyword_pesel_identification_number

- dowód osobisty
- dowódosobisty
- niepowtarzalny numer
- niepowtarzalnynumer
- nr.-pesel
- nr-pesel
- numer identyfikacyjny
- pesel
- tożsamości narodowej

Poland Passport Number

This sensitive information type entity is included in the EU Passport Number sensitive information type. It's also available as a stand-alone sensitive information type entity.

Format

Two letters and seven digits

Pattern

Two letters (not case-sensitive) followed by seven digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A data loss prevention (DLP) policy has high confidence that it detects this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_polish_passport_number_v2` finds content that matches the pattern.
- The checksum passes.
- A keyword from `Keywords_eu_passport_number` or `Keyword_polish_national_passport_number` is found.
- A keyword from `Keywords_eu_passport_date` is found.

A DLP policy has medium confidence that it detects this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_polish_passport_number_v2` finds content that matches the pattern.
- The checksum passes.
- A keyword from `Keywords_eu_passport_number` or `Keyword_polish_national_passport_number` is found.

A DLP policy has low confidence that it detects this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_polish_passport_number_v2` finds content that matches the pattern.
- The checksum passes.

XML

```
      <!-- Poland Passport Number -->
      <Entity id="03937FB5-D2B6-4487-B61F-0F8BFF7C3517" patternsProximity="300"
recommendedConfidence="75">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_polish_passport_number_v2" />
          <Match idRef="Keywords_eu_passport_date" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keyword_polish_national_passport_number" />
          </Any>
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_polish_passport_number_v2" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keyword_polish_national_passport_number" />
          </Any>
        </Pattern>
        <Pattern confidenceLevel="65">
          <IdMatch idRef="Func_polish_passport_number_v2" />
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_passport_number

- passport#
- passport #
- passportid
- passports
- passportno
- passport no

- passportnumber
- passport number
- passportnumbers
- passport numbers

Keyword_polish_national_passport_number

- numer paszportu
- numery paszportów
- numery paszportowe
- nr paszportu
- nr. paszportu
- nr paszportów
- n° passeport
- passeport n°

Keywords_eu_passport_date

- date of issue
- date of expiry

Last updated on 06/19/2026

Poland physical addresses

Article • 08/19/2024

This unbundled named entity detects patterns related to physical address from Poland. It's also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Poland REGON number

Article • 08/19/2024

Format

9-digit or 14-digit number

Pattern

nine digit or 14-digit number:

- nine digits or
- nine digits
- hyphen
- five digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_polish_regon_number` finds content that matches the pattern.
- A keyword from `Keywords_polish_regon_number` is found.

A DLP policy has low confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_polish_regon_number` finds content that matches the pattern.

XML

```
      <!-- Polish REGON Number  -->
      <Entity id="fc87b421-f437-4f8b-b739-29a735ead0d9"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_polish_regon_number" />
          <Match idRef="Keywords_polish_regon_number" />
        </Pattern>
        <Pattern confidenceLevel="65">
          <IdMatch idRef="Func_polish_regon_number" />
        </Pattern>
      </Entity>
```

Keywords

Keywords_poland_regon_number

- regon id
- statistical number
- statistical id
- statistical no
- regon number
- regonid#
- regonno#
- company id
- companyid#
- companyidno#
- numer statystyczny
- numeru regon
- numerstatystyczny#
- numeruregon#

Poland tax identification number

Article • 08/19/2024

Format

11 digits with no spaces or delimiters

Pattern

11 digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_poland_eu_tax_file_number` finds content that matches the pattern.
- A keyword from `Keywords_poland_eu_tax_file_number` is found.

XML

```
      <!-- Poland Tax Identification Number -->
      <Entity id="1ff28b4d-40f2-49e9-b677-9606a88e2bca"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_poland_eu_tax_file_number" />
          <Match idRef="Keywords_poland_eu_tax_file_number" />
```

```
        </Pattern>
      </Entity>
```

Keywords

Keywords_poland_eu_tax_file_number

- nip#
- nip
- numer identyfikacji podatkowej
- numeridentyfikacjipodatkowej#
- tax id
- tax identification no
- tax identification number
- tax no#
- tax no
- tax number
- tax registration number
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id
- tin no
- tin#
- vat id#
- vat id
- vat no
- vat number
- vatid#
- vatid
- vatno#

Portugal citizen card number

Article • 08/19/2024

Format

Twelve alphanumeric characters

Pattern

Twelve alphanumeric characters:

- Eight digits
- An optional space or dash "-"
- One digit
- An optional space
- Two characters
- One check digit

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_portugal_citizen_card` finds content that matches the pattern.
- A keyword from `Keyword_portugal_citizen_card` is found.

XML

```xml
<!-- Portugal Citizen Card Number -->
<Entity id="91a7ece2-add4-4986-9a15-c84544d81ecd" recommendedConfidence="85"
patternsProximity="300">
  <Pattern confidenceLevel="85">
     <IdMatch idRef="Regex_portugal_citizen_card"/>
     <Match idRef="Keyword_portugal_citizen_card"/>
  </Pattern>
</Entity>
```

Keywords

Keyword_portugal_citizen_card

- bilhete de identidade
- cartão de cidadão
- citizen card
- document number
- documento de identificação
- id number
- identification no
- identification number
- identity card no
- identity card number
- national id card
- nic
- número bi de portugal
- número de identificação civil
- número de identificação fiscal
- número do documento
- portugal bi number

Portugal drivers license number

Article • 08/19/2024

Format

two patterns - two letters followed by 5-8 digits with special characters

Pattern

Pattern 1: Two letters followed by 5/6 with special characters:

- Two letters (not case-sensitive)
- A hyphen
- Five or Six digits
- A space
- One digit

Pattern 2: One letter followed by 6/8 digits with special characters:

- One letter (not case-sensitive)
- A hyphen
- Six or eight digits
- A space
- One digit

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_portugal_eu_driver's_license_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_driver's_license_number` or `Keywords_portugal_eu_driver's_license_number` is found.

XML

```
      <!-- Portugal Driver's License Number -->
      <Entity id="977f1e5a-2c33-4bcc-b516-95bb275cff23"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_portugal_eu_driver's_license_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_driver's_license_number" />
            <Match idRef="Keywords_portugal_eu_driver's_license_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences

- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#

- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence

- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses
- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_portugal_eu_driver's_license_number

- carteira de motorista
- carteira motorista
- carteira de habilitação
- carteira habilitação
- número de licença
- número licença
- permissão de condução
- permissão condução
- Licença condução Portugal
- carta de condução

Portugal passport number

Article • 08/19/2024

Format

one letter followed by six digits with no spaces or delimiters

Pattern

one letter followed by six digits:

- one letter (not case-sensitive)
- six digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_portugal_eu_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_portugal_eu_passport_number` is found.
- The regular expression `Regex_eu_passport_date1` finds date in the format DD.MM.YYYY or a keyword from `Keywords_eu_passport_date` is found

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_portugal_eu_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_portugal_eu_passport_number` is found.

XML

```
      <!-- Portugal Passport Number -->
      <Entity id="080a52fd-a7bc-431e-b54d-51f08f59db11"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Regex_portugal_eu_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_portugal_eu_passport_number" />
          </Any>
          <Any minMatches="1">
            <Match idRef="Regex_eu_passport_date1" />
            <Match idRef="Keywords_eu_passport_date" />
          </Any>
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_portugal_eu_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_portugal_eu_passport_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_passport_number

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number

- passportnumbers
- passport numbers

Keywords_portugal_eu_passport_number

- número do passaporte
- portuguese passport
- portuguese passeport
- portuguese passaporte
- passaporte nº
- passeport nº
- números de passaporte
- portuguese passports
- número passaporte
- números passaporte

Keywords_eu_passport_date

- date of issue
- date of expiry

Portuguese physical addresses

Article • 02/18/2025

This unbundled named entity detects patterns related to physical address from Portugal. It's also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Portugal tax identification number

Article • 08/19/2024

Format

nine digits with optional spaces

Pattern

- three digits
- an optional space
- three digits
- an optional space
- three digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_portugal_eu_tax_file_number` finds content that matches the pattern.
- A keyword from `Keywords_portugal_eu_tax_file_number` is found.

A DLP policy has low confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_portugal_eu_tax_file_number` finds content that matches the pattern.

XML

```
      <!-- Portugal Tax Identification Number -->
      <Entity id="65372402-3131-4f1e-9983-4439841d1f15"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_portugal_eu_tax_file_number" />
          <Match idRef="Keywords_portugal_eu_tax_file_number" />
        </Pattern>
        <Pattern confidenceLevel="65">
          <IdMatch idRef="Func_portugal_eu_tax_file_number" />
        </Pattern>
      </Entity>
```

Keywords

Keywords_portugal_eu_tax_file_number

- cpf#
- cpf
- nif#
- nif
- número de identificação fisca
- numero fiscal
- tax id
- tax identification no
- tax identification number
- tax no#
- tax no
- tax number
- tax registration number
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id
- tin no

- tin#

Qatari ID Card Number

Format

11 digits

Pattern

11 digits:

- a digit, either 2 or 3
- two digits representing the last two numbers of year of birth
- three digits representing ISO country code
- five digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A data loss prevention (DLP) policy has high confidence that it detects this type of sensitive information if, within a proximity of 300 characters:

- The regular expression Regex_qatari_id_card finds content that matches the pattern.
- A keyword from Keyword_qatari_id_card is found.

A DLP policy has medium confidence that it detects this type of sensitive information if, within a proximity of 300 characters:

- The regular expression Regex_qatari_id_card finds content that matches the pattern.

XML

```xml
     <!-- Qatari ID Card Number-->
        <Entity id="52b1b60e-a4be-4b5a-a67b-6f9bbb7811da" patternsProximity="300"
recommendedConfidence="85" relaxProximity="true">
          <Pattern confidenceLevel="85">
            <IdMatch idRef="Regex_qatari_id_card" />
            <Match idRef="Keyword_qatari_id_card" />
          </Pattern>
          <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_qatari_id_card" />
          </Pattern>
        </Entity>
```

Keywords

Keyword_qatari_id_card

- GP
- DOB
- D.O.B
- Date of Birth
- Birth Date
- date of issue
- date of expiry

Last updated on 06/19/2026

Romania drivers license number

Article • 08/19/2024

Format

one character followed by eight digits

Pattern

one character followed by eight digits:

- one letter (not case-sensitive) or digit
- eight digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_romania_eu_driver's_license_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_driver's_license_number` or `Keywords_romania_eu_driver's_license_number` is found.

```
XML
```

```
      <!-- Romania Driver's License Number -->
      <Entity id="b5511ace-2fd8-4ae4-b6fc-c7c6e4689e3c"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_romania_eu_driver's_license_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_driver's_license_number" />
            <Match idRef="Keywords_romania_eu_driver's_license_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic

- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#

- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen

- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses
- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_romania_eu_driver's_license_number

- permis de conducere
- permisului de conducere
- permisului conducere
- permisele de conducere
- permisele conducere
- permis conducere

Romania passport number

Article • 08/19/2024

Format

eight or nine digits without spaces and delimiters

Pattern

eight or nine digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_romania_eu_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_romania_eu_passport_number` is found.
- The regular expression `Regex_romania_eu_passport_date` finds date in the format DD MMM/MMM YY (Example- 01 FEB/FEB 10) or a keyword from `Keywords_eu_passport_date` is found

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_romania_eu_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_romania_eu_passport_number` is found.

XML

```
      <!-- Romania Passport Number -->
      <Entity id="5d31b90c-7fe2-4a76-a14b-767b8fd19d6c"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Regex_romania_eu_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_romania_eu_passport_number" />
          </Any>
          <Any minMatches="1">
            <Match idRef="Regex_romania_eu_passport_date" />
            <Match idRef="Keywords_eu_passport_date" />
          </Any>
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_romania_eu_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_romania_eu_passport_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_passport_number

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers

Keywords_romania_eu_passport_number

numărul pașaportului numarul pasaportului numerele pașaportului Pașaport nr

Keywords_eu_passport_date

- date of issue
- date of expiry

Romania personal numeric code (CNP)

Article • 08/19/2024

Format

13 digits without spaces and delimiters

Pattern

- one digit from 1-9
- six digits representing date of birth (YYMMDD)
- two digits, which can be 01-52 or 99
- four digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_romania_eu_national_id_card` finds content that matches the pattern.
- A keyword from `Keywords_romania_eu_national_id_card` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_romania_eu_national_id_card` finds content that matches the pattern.

XML

```
      <!-- Romania Personal Numerical Code (CNP) -->
      <Entity id="eb5fa399-fe28-4c67-8188-d63a616ed89c"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_romania_eu_national_id_card" />
          <Match idRef="Keywords_romania_eu_national_id_card" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_romania_eu_national_id_card" />
        </Pattern>
      </Entity>
```

Keywords

Keywords_romania_eu_national_id_card

- cnp#
- cnp
- cod identificare personal
- cod numeric personal
- cod unic identificare
- codnumericpersonal#
- codul fiscal nr.
- identificarea fiscală nr#
- id-ul taxei
- insurance number
- insurancenumber#
- national id#
- national id
- national identification number
- număr identificare personal
- număr identitate
- număr personal unic
- număridentitate#
- număridentitate
- numărpersonalunic#
- numărpersonalunic

- număru de identificare fiscală
- numărul de identificare fiscală
- personal numeric code
- pin#
- pin
- tax file no
- tax file number
- tax id
- tax identification no
- tax identification number
- tax no#
- tax no
- tax number
- tax registration number
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id
- tin no
- tin#
- unique identification number
- unique identity number
- uniqueidentityno#
- uniqueidentityno

Romania physical addresses

Article • 08/19/2024

This unbundled named entity detects patterns related to physical address from Romania. It's also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Russia passport number domestic

Article • 08/19/2024

Format

10-digit number

Pattern

10-digit number:

- two digits
- an optional space or hyphen
- two digits
- an optional space
- six digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regex `Regex_Russian_Passport_Number_Domestic` finds content that matches the pattern.
- A keyword from `Keyword_Russian_Passport_Number` is found.

```
XML
```

```
      <!-- Russian Passport Number Domestic -->
      <Entity id="76ec2f5d-cedb-48e1-8070-1998794af445"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_Russian_Passport_Number_Domestic" />
          <Match idRef="Keyword_Russian_Passport_Number" />
        </Pattern>
      </Entity>
```

Keywords

Keyword_russia_passport_number_domestic

- passport number
- passport no
- passport #
- passport id
- passportno#
- passportnumber#
- паспорт нет
- паспорт id
- российской паспорт
- русский номер паспорта
- паспорт#
- паспортid#
- номер паспорта
- номерпаспорта#

Russia passport number international

Article • 08/19/2024

Format

nine-digit number

Pattern

nine-digit number:

- two digits
- an optional space or hyphen
- seven digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regex `Regex_Russian_Passport_Number_International` finds content that matches the pattern.
- A keyword from `Keyword_Russian_Passport_Number` is found.

```
XML
```

```
      <!-- Russian Passport Number International -->
      <Entity id="ac5f4878-75e4-4b82-af2d-02e13ea9f411"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_Russian_Passport_Number_International" />
          <Match idRef="Keyword_Russian_Passport_Number" />
        </Pattern>
      </Entity>
```

Keywords

Keywords_russia_passport_number_international

- passport number
- passport no
- passport #
- passport id
- passportno#
- passportnumber#
- паспорт нет
- паспорт id
- российской паспорт
- русский номер паспорта
- паспорт#
- паспортid#
- номер паспорта
- номерпаспорта#

Russia physical addresses

This unbundled named entity detects patterns related to physical address from Russia. It is also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following services:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Last updated on 06/19/2026

Russia taxpayer identification number

Format

10 or 12 digits, validated by a checksum function.

Pattern

10 digits (organizations) or 12 digits (individuals):

- Four digits (tax authority code)
- Five or six digits (serial number)
- One or two check digits

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Checksum

Yes

Definition

A data loss prevention (DLP) policy has high confidence that it detects this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_Russian_Taxpayer_Identification_Number` finds content that matches the pattern.
- A keyword from `Keyword_Russian_Taxpayer_Identification_Number` is found.

A DLP policy has medium confidence that it detects this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_Russian_Taxpayer_Identification_Number` finds content that matches the pattern.

XML

```
    <!-- Russian Taxpayer Identification Number -->
    <Entity id="81145f47-868a-464c-956a-eed7696fdcdf" patternsProximity="300"
recommendedConfidence="85" filters="russian_taxpayer_id_filter">
      <Pattern confidenceLevel="85">
        <IdMatch idRef="Func_Russian_Taxpayer_Identification_Number" />
        <Match idRef="Keyword_Russian_Taxpayer_Identification_Number" />
      </Pattern>
      <Pattern confidenceLevel="75">
        <IdMatch idRef="Func_Russian_Taxpayer_Identification_Number" />
      </Pattern>
    </Entity>
```

Keywords

The following keywords are used for detection. These include Taxpayer Identification Number (TIN) variations and Russian-language equivalents.

Keyword_Russian_Taxpayer_Identification_Number

- `TIN`
- `taxpayer number`
- `taxpayer ID`
- `taxpayer no`
- `tax ID`
- `Tin`
- `tinno#`
- `Inn`
- `inn#`
- `taxpayerno#`
- `taxid#`
- `taxpayeridno#`
- `taxpayerid#`
- `taxpayer`
- `tax identification`
- `tax payer`
- `taxid`

- `НДС`
- `номер налогоплательщика`
- `Налогоплательщика ИД`
- `налог число`
- `налогчисло#`
- `ИНН`
- `НДС#`

Last updated on 06/19/2026

Saudi Arabia National ID

Article • 09/13/2024

Format

10 digits

Pattern

10 consecutive digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_saudi_arabia_national_id` finds content that matches the pattern.
- A keyword from `Keyword_saudi_arabia_national_id` is found.

XML

```
<!-- Saudi Arabia National ID -->
<Entity id="8c5a0ba8-404a-41a3-8871-746aa21ee6c0" patternsProximity="300"
recommendedConfidence="75">
    <Pattern confidenceLevel="75">
        <IdMatch idRef="Regex_saudi_arabia_national_id" />
        <Any minMatches="1">
```

```
            <Match idRef="Keyword_saudi_arabia_national_id" />
          </Any>
      </Pattern>
  </Entity>
```

Keywords

Keyword_saudi_arabia_national_id

- Identification Card
- I card number
- ID number
- الوطنية الهوية بطاقة رقم

Singapore driver's license number

Article • 08/19/2024

Format

Nine alphanumeric characters.

Pattern

- One letter followed by
- seven digits followed by
- One letter

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Checksum

NA

Definition

A DLP policy has high confidence of detection for this type of sensitive information if, within a proximity of 300 characters:

- The regular expression Regex_singapore_driving_license finds content that matches the pattern.
- A keyword from Keyword_singapore_driving_license or Keywords_eu_driver's_license_number_common is found.

A DLP policy has low confidence of detection for this type of sensitive information if, within a proximity of 300 characters:

- The regular expression Regex_singapore_driving_license finds content that matches the pattern.

XML

```
      <!-- Singapore Driving License Number -->
      <Entity id="beaa54d7-ec6a-4160-a37b-f3916b68934b"
patternsProximity="300" recommendedConfidence="85" relaxProximity="true">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Regex_singapore_driving_license" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_driver's_license_number_common" />
            <Match idRef="Keyword_singapore_driving_license" />
          </Any>
        </Pattern>
        <Pattern confidenceLevel="65">
          <IdMatch idRef="Regex_singapore_driving_license" />
        </Pattern>
      </Entity>
```

Keywords

Keyword_singapore_driving-license

- Licence number
- licence no
- licence#
- lesen memandu
- lesen pemandu
- nombor lesen
- 驾驶执照
- 驾驶 执照
- 驾驶执照号码
- 驾驶 执照 号码

Keywords_eu_driver's_license_number_common

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics

- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#

- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#

- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses
- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Singapore national registration identity card (NRIC) number

Article • 12/13/2024

Format

Nine letters and digits

Pattern

- Nine letters and digits:
  - the letter "F," "G," "M," "S," or "T" (not case-sensitive)
  - seven digits
  - an alphabetic check digit

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_singapore_nric` finds content that matches the pattern.
- A keyword from `Keyword_singapore_nric` is found.
- The checksum passes.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_singapore_nric` finds content that matches the pattern.
- The checksum passes.

XML

```
<!-- Singapore National Registration Identity Card (NRIC) Number -->
<Entity id="cead390a-dd83-4856-9751-fb6dc98c34da" recommendedConfidence="75"
patternsProximity="300">
  <Pattern confidenceLevel="85">
     <IdMatch idRef="Regex_singapore_nric"/>
     <Match idRef="Keyword_singapore_nric"/>
  </Pattern>
  <Pattern confidenceLevel="75">
     <IdMatch idRef="Regex_singapore_nric"/>
  </Pattern>
</Entity>
```

Keywords

Keyword_singapore_nric

- National Registration Identity Card
- Identity Card Number
- NRIC
- IC
- Foreign Identification Number
- FIN
- 身份证
- 身份證

Singapore passport number

Format

Nine alphanumeric characters with only specific values for letters.

Pattern

- a letter e/E or k/K followed by
- seven digits followed by
- a letter

Checksum

Not applicable

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A Data Loss Protection (DLP) policy has high confidence of detection for this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_singapore_passport_number` finds content that matches the pattern.
- A keyword from `Keyword_singapore_passport_number` is found.

A DLP policy has low confidence of detection for this type of sensitive information if, within a proximity of 300 characters:

- The regular expression Regex_singapore_passport_number finds content that matches the pattern.

XML

```xml
      <!-- Singapore Passport Number -->
      <Entity id="4081ae73-c4ec-405d-932c-8b36ae35b20c" patternsProximity="300"
recommendedConfidence="85" relaxProximity="true">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Regex_singapore_passport_number" />
          <Match idRef="Keyword_singapore_passport_number" />
        </Pattern>
        <Pattern confidenceLevel="65">
          <IdMatch idRef="Regex_singapore_passport_number" />
        </Pattern>
      </Entity>
```

Keywords

Keyword_singapore_passport_number

- passport#
- passport #
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers
- pasport#
- pasport #
- pasportno
- pasport no
- pasportnumber
- pasport number
- pasportnumbers
- pasport numbers
- no. pasport
- nombor pasport
- nomborpasport

Last updated on 12/09/2025

Singapore physical addresses

This unbundled named entity detects patterns related to physical address from Singapore. It is also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following services:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Last updated on 06/19/2026

Slack access token

This SIT is also included in the All credentials bundled SIT.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a credential scanning SIT in an endpoint DLP policy. This requirement is specific to the combination of credential scanning SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

Format

A combination of up to 34 characters consisting of letters, digits, and special characters.

Pattern

A token prefix (case-sensitive) `xoxp-`, `xoxb-`, `xoxa-`, `xoxr-`, `xoxo-`, `xoxs-`, or `xoxe-`

A combination of up to 29 characters:

- 29 a-z (not case-sensitive)
- 0-9 or hyphens (-)

for example:

`xoxp-abcdef-abcdef-abcdef-abcdef`

Credential example

`slack_token= xoxp-abcdef-abcdef-abcdef-abcdef;`

Checksum

No

SITs that have checksums use a unique calculation to check if the information is valid. This means when the **Checksum** value is **Yes**, the service can make a positive detection based on the sensitive data alone. When the **Checksum** value is **No**, additional (secondary) elements must also be detected for the service to make a positive detection.

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

This SIT is designed to match the security information that's used to access Slack platform functionalities (for example Bot tokens, User tokens and App-level tokens).

It uses several primary resources:

- Patterns of Slack user/bot/workspace token.
- Patterns of CredentialName, CredentialFeatures, AccountIdentityName, AccountIdentityValue, ResourceType, ResourceName.
- Patterns of mockup values, redactions, and placeholders.

The patterns are designed to match actual credentials with reasonable confidence. The patterns don't match credentials formatted as examples. Mockup values, redacted values, and placeholders such as credential type or usage descriptions that are in the position where an actual secret value should be present won't be matched.

Keywords

Keyword_SlackTokens:

- `xox`

Last updated on 11/18/2025

Slovakia drivers license number

Article • 08/19/2024

Format

one character followed by seven digits

Pattern

one character followed by seven digits

- one letter (not case-sensitive) or digit
- seven digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_slovakia_eu_driver's_license_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_driver's_license_number` or `Keywords_slovakia_eu_driver's_license_number` is found.

```
XML
```

```
      <!-- Slovakia Driver's License Number -->
      <Entity id="14240c22-b6de-4ce5-a90b-137f74252513"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_slovakia_eu_driver's_license_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_driver's_license_number" />
            <Match idRef="Keywords_slovakia_eu_driver's_license_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic

- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#

- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen

- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses
- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_slovakia_eu_driver's_license_number

- vodičský preukaz
- vodičské preukazy
- vodičského preukazu
- vodičských preukazov

Slovakia passport number

Article • 08/19/2024

Format

eight or nine character alphanumeric pattern

Pattern

one letter (not case-sensitive) followed by seven digits or two letters (not case-sensitive) followed by six or seven digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_slovakia_eu_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_slovakia_eu_passport_number` is found.
- The regular expression `Regex_eu_passport_date1` finds date in the format DD.MM.YYYY or a keyword from `Keywords_eu_passport_date` is found

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_slovakia_eu_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_slovakia_eu_passport_number` is found.

XML

```
      <!-- Slovakia Passport Number -->
      <Entity id="238e1f08-d80e-4793-af33-9b57918335b7"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Regex_slovakia_eu_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_slovakia_eu_passport_number" />
          </Any>
          <Any minMatches="1">
            <Match idRef="Regex_eu_passport_date1" />
            <Match idRef="Keywords_eu_passport_date" />
          </Any>
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_slovakia_eu_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_slovakia_eu_passport_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_passport_number

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers

Keywords_slovakia_eu_passport_number

- číslo pasu
- čísla pasov
- pas č.
- Passeport n°
- n° Passeport

Keywords_eu_passport_date

- date of issue
- date of expiry

Slovakia personal number

Article • 08/19/2024

Format

Nine or 10 digits containing optional backslash

Pattern

- six digits representing date of birth
- optional slash (/)
- three digits
- one optional check digit

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_slovakia_eu_national_id_card` finds content that matches the pattern.
- A keyword from `Keywords_slovakia_eu_national_id_card` is found.

A DLP policy has low confidence that it detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_slovakia_eu_national_id_card` finds content that matches the pattern.

XML

```
 <!-- Slovakia Personal Number -->
      <Entity id="951c26b7-3b35-4f73-924b-15dd599cb9ab"
patternsProximity="300" recommendedConfidence="85" relaxProximity="true">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_slovakia_eu_national_id_card" />
          <Match idRef="Keywords_slovakia_eu_national_id_card" />
        </Pattern>
        <Pattern confidenceLevel="65">
          <IdMatch idRef="Func_slovakia_eu_national_id_card" />
        </Pattern>
      </Entity>
    </Version>
```

Keywords

Keywords_slovakia_eu_national_id_card

- azonosító szám
- birth number
- číslo národnej identifikačnej karty
- číslo občianského preukazu
- daňové číslo
- id number
- identification no
- identification number
- identifikačná karta č
- identifikačné číslo
- identity card no
- identity card number
- národná identifikačná značka č
- national number
- nationalnumber#
- nemzeti személyazonosító igazolvány
- personalidnumber#
- rč
- rodne cislo
- rodné číslo
- social security number

- ssn#
- ssn
- személyi igazolvány szám
- személyi igazolvány száma
- személyigazolvány szám
- tax file no
- tax file number
- tax id
- tax identification no
- tax identification number
- tax no#
- tax no
- tax number
- tax registration number
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id
- tin no
- tin#

Slovakia physical addresses

Article • 08/19/2024

This unbundled named entity detects patterns related to physical address from Slovakia. It's also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Slovenia drivers license number

Article • 08/19/2024

Format

nine digits without spaces and delimiters

Pattern

nine digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_slovenia_eu_driver's_license_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_driver's_license_number` or `Keywords_slovenia_eu_driver's_license_number` is found.

XML

```
      <!-- Slovenia Driver's License Number -->
      <Entity id="d5bc089a-f2ee-433d-a6b1-5c253051d6f2"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_slovenia_eu_driver's_license_number" />
```

```
          <Any minMatches="1">
            <Match idRef="Keywords_eu_driver's_license_number" />
            <Match idRef="Keywords_slovenia_eu_driver's_license_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences

- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#

- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses

- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_slovenia_eu_driver's_license_number

- vozniško dovoljenje
- vozniška številka licence
- vozniških dovoljenj
- številka vozniškega dovoljenja
- številke vozniških dovoljenj

Slovenia passport number

Article • 08/19/2024

Format

two letters followed by seven digits with no spaces or delimiters

Pattern

two letters followed by seven digits:

- the letter "P"
- one uppercase letter
- seven digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_slovenia_eu_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_slovenia_eu_passport_number` is found.
- The regular expression `Regex_eu_passport_date1` finds date in the format DD.MM.YYYY or a keyword from `Keywords_eu_passport_date` is found

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_slovenia_eu_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_slovenia_eu_passport_number` is found.

XML

```
      <!-- Slovenia Passport Number -->
      <Entity id="235b7976-7bbe-4df5-bb40-08678e749d1a"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Regex_slovenia_eu_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_slovenia_eu_passport_number" />
          </Any>
          <Any minMatches="1">
            <Match idRef="Regex_eu_passport_date1" />
            <Match idRef="Keywords_eu_passport_date" />
          </Any>
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_slovenia_eu_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_slovenia_eu_passport_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_passport_number

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number

- passportnumbers
- passport numbers

Keywords_slovenia_eu_passport_number

- številka potnega lista
- potek veljavnosti
- potni list#
- datum rojstva
- potni list
- številke potnih listov

Keywords_eu_passport_date

- date of issue
- date of expiry

Slovenia physical addresses

Article • 08/19/2024

This unbundled named entity detects patterns related to physical address from Slovenia. It's also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Slovenia tax identification number

Article • 08/19/2024

Format

eight digits with no spaces or delimiters

Pattern

- one digit from 1-9
- six digits
- one check digit

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_slovenia_eu_tax_file_number` finds content that matches the pattern.
- A keyword from `Keywords_slovenia_eu_tax_file_number` is found.

A DLP policy has low confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_slovenia_eu_tax_file_number` finds content that matches the pattern.

XML

```xml
      <!-- Slovenia Tax Identification Number -->
      <Entity id="e47b071e-c352-4d70-8241-8c215ad65505"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_slovenia_eu_tax_file_number" />
          <Match idRef="Keywords_slovenia_eu_tax_file_number" />
        </Pattern>
        <Pattern confidenceLevel="65">
          <IdMatch idRef="Func_slovenia_eu_tax_file_number" />
        </Pattern>
      </Entity>
```

Keywords

Keywords_slovenia_eu_tax_file_number

- davčna številka
- identifikacijska številka davka
- številka davčne datoteke
- tax file no
- tax file number
- tax id
- tax identification no
- tax identification number
- tax no#
- tax no
- tax number
- tax registration number
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id
- tin no
- tin#

Slovenia Unique Master Citizen Number

Article • 08/19/2024

Format

13 digits without spaces or delimiters

Pattern

13 digits in the specified pattern:

- seven digits that correspond to the birth date (DDMMLLL) where "LLL" corresponds to the last three digits of the birth year
- two digits that correspond to the area of birth "50"
- three digits that correspond to a combination of gender and serial number for persons born on the same day. 000-499 for male and 500-999 for female.
- one check digit

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_slovenia_eu_national_id_card` finds content that matches the pattern.
- A keyword from `Keywords_slovenia_eu_national_id_card` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_slovenia_eu_national_id_card` finds content that matches the pattern.

XML

```
      <!-- Slovenia Unique Master Citizen Number -->
      <Entity id="68948b27-803d-41e4-adf1-13e05eb541bb"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_slovenia_eu_national_id_card" />
          <Match idRef="Keywords_slovenia_eu_national_id_card" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_slovenia_eu_national_id_card" />
        </Pattern>
      </Entity>
```

Keywords

Keywords_slovenia_eu_national_id_card

- edinstvena številka glavnega državljana
- emšo
- enotna maticna številka obcana
- id card
- identification number
- identifikacijska številka
- identity card
- nacionalna id
- nacionalni potni list
- national id
- osebna izkaznica
- osebni koda
- osebni ne
- osebni številka
- personal code
- personal number
- personal numeric code
- številka državljana
- unique citizen number

- unique id number
- unique identity number
- unique master citizen number
- unique registration number
- uniqueidentityno #
- uniqueidentityno#

South Africa identification number

Article • 08/19/2024

Format

13 digits that may contain spaces

Pattern

13 digits:

- six digits in the format YYMMDD, which are the date of birth
- four digits
- a single-digit citizenship indicator
- the digit "8" or "9"
- one digit, which is a checksum digit

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_south_africa_identification_number` finds content that matches the pattern.
- A keyword from `Keyword_south_africa_identification_number` is found.
- The checksum passes.

XML

```xml
<!-- South Africa Identification Number -->
<Entity id="e2adf7cb-8ea6-4048-a2ed-d89eb65f2780" recommendedConfidence="85"
patternsProximity="300">
  <Pattern confidenceLevel="85">
     <IdMatch idRef="Func_south_africa_identification_number"/>
     <Match idRef="Keyword_south_africa_identification_number"/>
  </Pattern>
</Entity>
```

Keywords

Keyword_south_africa_identification_number

- Identity card
- ID
- Identification

South Africa physical addresses

This unbundled named entity detects patterns related to physical address from South Africa. It is also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following services:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Last updated on 06/19/2026

South Korea driver's license number

Article • 08/19/2024

Format

10 or 12 digits with optional hyphens.

Pattern

Pattern 1:

- two digits
- optional hyphen
- six digits
- optional hyphen
- two digits

Pattern 2:

- optional 2 digits (allowed digits are 11-26 and 28)
- optional hyphen
- two digits
- optional hyphen
- six digits
- optional hyphen
- two digits

Checksum

Not applicable

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A Data Loss Passport (DLP) policy has high confidence of detection for this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_south_korea_driver's_license_10digits` or `Regex_south_korea_driver's_license_12digits` finds content that matches the pattern.
- A keyword from `Keyword_south_korea_driver's_license` or `Keywords_eu_driver's_license_number_common` is found.

A DLP policy has low confidence of detection for this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_south_korea_driver's_license_12digits` finds content that matches the pattern.

XML

```
      <!-- South Korea Driver's License Number -->
      <Entity id="3da0b024-4c08-406b-9f9e-de96accd0209"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Regex_south_korea_driver's_license_10digits" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_driver's_license_number_common" />
            <Match idRef="Keyword_south_korea_driver's_license" />
          </Any>
        </Pattern>
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Regex_south_korea_driver's_license_12digits" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_driver's_license_number_common" />
            <Match idRef="Keyword_south_korea_driver's_license" />
          </Any>
        </Pattern>
        <Pattern confidenceLevel="65">
          <IdMatch idRef="Regex_south_korea_driver's_license_12digits" />
        </Pattern>
      </Entity>
```

Keywords

Keyword_south_korea_driver's_license

- 운전면허번호

- 운전면허 번호
- 운전 면허 번호
- 운전면허증 번호
- 운전면허증

Keywords_eu_driver's_license_number_common

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses

- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#

- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses
- driving licence
- driving licences
- driving permit
- dl no

- dlno
- dl number

South Korea passport number

Article • 08/19/2024

Format

Nine alphanumeric characters; only specific values are allowed.

Pattern

- a letter (M|S|R|O|D) followed by
- eight digits

Or

- a letter (M|S|R|O|D) followed by
- three digits followed by
- one letter followed by
- four digits

Checksum

Not applicable

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_south_korea_passport_number` finds content that matches the pattern.
- A keyword from `Keyword_south_korea_passport_number` is found.

A DLP policy has low confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression Regex_south_korea_passport_number finds content that matches the pattern.

XML

```
      <!-- South Korea Passport Number -->
      <Entity id="cb725748-0f69-4282-8a03-e9b222a10a18"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Regex_south_korea_passport_number" />
          <Match idRef="Keyword_south_korea_passport_number" />
        </Pattern>
        <Pattern confidenceLevel="65">
          <IdMatch idRef="Regex_south_korea_passport_number" />
        </Pattern>
      </Entity>
```

Keywords

Keyword_south_korea_passport_number

- passport#
- passport #
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers
- 여권 번호
- 한국어 여권
- 여권 번호
- 여권번호
- 한국 여권
- 한국여권
- 한국 여권 번호

South Korea resident registration number

Article • 08/19/2024

Format

Thirteen digits containing a hyphen

Pattern

13 digits:

- six digits in the format YYMMDD, which are the date of birth
- a hyphen
- seven digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_south_korea_resident_number` finds content that matches the pattern.
- A keyword from `Keyword_south_korea_resident_number` is found.
- The checksum passes.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_south_korea_resident_number` finds content that matches the pattern.
- The checksum passes.

XML

```
<!-- South Korea Resident Registration Number -->
<Entity id="5b802e18-ba80-44c4-bc83-bf2ad36ae36a" recommendedConfidence="85"
patternsProximity="300">
  <Pattern confidenceLevel="85">
     <IdMatch idRef="Func_south_korea_resident_number"/>
     <Match idRef="Keyword_south_korea_resident_number"/>
  </Pattern>
  <Pattern confidenceLevel="75">
     <IdMatch idRef="Func_south_korea_resident_number"/>
  </Pattern>
</Entity>
```

Keywords

Keyword_south_korea_resident_number

- National ID card
- Citizen's Registration Number
- Jumin deungnok beonho
- RRN
- 주민등록번호

Spain DNI

Summarize this article for me

Format

Eight digits followed by one character.

Pattern

Eight digits followed by one character.

- eight digits
- An optional space or hyphen
- one check letter (not case-sensitive)

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_spain_eu_DL_and_NI_number_citizen` or `Func_spain_eu_DL_and_NI_number_foreigner` finds content that matches the pattern.
- A keyword from `Keywords_spain_eu_national_id_card"` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_spain_eu_DL_and_NI_number_citizen` or `Func_spain_eu_DL_and_NI_number_foreigner` finds content that matches the pattern.

XML

```
      <!-- Spain DNI -->
      <Entity id="8e6251b9-47b4-40e8-a42b-0f80876be192" patternsProximity="300"
recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_spain_eu_DL_and_NI_number_citizen" />
          <Match idRef="Keywords_spain_eu_national_id_card" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_spain_eu_DL_and_NI_number_citizen" />
        </Pattern>
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_spain_eu_DL_and_NI_number_foreigner" />
          <Match idRef="Keywords_spain_eu_national_id_card" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_spain_eu_DL_and_NI_number_foreigner" />
        </Pattern>
      </Entity>
```

Keywords

Keywords_spain_eu_national_id_card

- carné de identidad
- dni#
- dni
- dninúmero#
- documento nacional de identidad
- identidad único
- identidadúnico#
- insurance number
- national identification number
- national identity
- nationalid#
- nationalidno#
- nie#
- nie
- nienúmero#
- número de identificación

- número nacional identidad
- personal identification number
- personal identity no
- unique identity number
- uniqueid#

Last updated on 02/19/2026

Spain drivers license number

Article • 08/19/2024

Format

eight digits followed by one character

Pattern

eight digits followed by one character:

- eight digits
- one digit or letter (not case-sensitive)

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_spain_eu_DL_and_NI_number_citizen` or `Func_spain_eu_DL_and_NI_number_foreigner` finds content that matches the pattern.
- A keyword from `Keywords_eu_driver's_license_number` or `Keywords_spain_eu_driver's_license_number` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_spain_eu_DL_and_NI_number_citizen` or `Func_spain_eu_DL_and_NI_number_foreigner` finds content that matches the pattern.

XML

```
      <!-- Spain Driver's License Number -->
      <Entity id="d5a82922-b501-4f40-8868-341321146aa2"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_spain_eu_DL_and_NI_number_citizen" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_driver's_license_number" />
            <Match idRef="Keywords_spain_eu_driver's_license_number" />
          </Any>
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_spain_eu_DL_and_NI_number_citizen" />
        </Pattern>
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_spain_eu_DL_and_NI_number_foreigner" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_driver's_license_number" />
            <Match idRef="Keywords_spain_eu_driver's_license_number" />
          </Any>
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_spain_eu_DL_and_NI_number_foreigner" />
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences

- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#

- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#

- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses
- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_spain_eu_driver's_license_number

- permiso de conducción
- permiso conducción
- licencia de conducir
- licencia conducir
- permiso conducir
- permiso de conducir
- permisos de conducir
- permisos conducir
- carnet conducir
- carnet de conducir
- licencia de manejo
- licencia manejo

Spain passport number

Article • 08/19/2024

Format

an eight- or nine-character combination of letters and numbers with no spaces or delimiters

Pattern

an eight- or nine-character combination of letters and numbers:

- two digits or letters
- one digit or letter (optional)
- six digits

Checksum

Not applicable

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_spain_eu_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_spain_eu_passport_number` is found.

- The regular expression `Regex_spain_eu_passport_date` finds date in the format DD-MM-YYYY or a keyword from `Keywords_eu_passport_date` is found

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_spain_eu_passport_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_spain_eu_passport_number` is found.

XML

```
      <!-- Spain Passport Number -->
      <Entity id="d17a57de-9fa5-4e9f-85d3-85c26d89686e"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Regex_spain_eu_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_spain_eu_passport_number" />
          </Any>
          <Any minMatches="1">
            <Match idRef="Regex_spain_eu_passport_date" />
            <Match idRef="Keywords_eu_passport_date" />
          </Any>
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_spain_eu_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_spain_eu_passport_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_passport_number

- passport#
- passport #
- passportid
- passports
- passportno
- passport no

- passportnumber
- passport number
- passportnumbers
- passport numbers

Keywords_spain_eu_passport_number

- libreta pasaporte
- número pasaporte
- españa pasaporte
- números de pasaporte
- número de pasaporte
- números pasaporte
- pasaporte no
- Passeport n°
- n° Passeport
- pasaporte no.
- pasaporte n°
- spain passport

Keywords_eu_passport_date

- date of issue
- date of expiry

Spain physical addresses

Article • 08/19/2024

This unbundled named entity detects patterns related to physical address from Spain. It's also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Spain social security number (SSN)

Article • 08/19/2024

Format

11-12 digits

Pattern

11-12 digits:

- two digits
- a forward slash (optional)
- seven to eight digits
- a forward slash (optional)
- two digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_spanish_social_security_number` finds content that matches the pattern.
- The checksum passes.
  - A keyword from `Keywords_spain_eu_ssn_or_equivalent` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_spanish_social_security_number` finds content that matches the pattern.
- The checksum passes.

XML

```
    <!-- Spain SSN -->
    <Entity id="5df987c0-8eae-4bce-ace7-b316347f3070"
patternsProximity="300" recommendedConfidence="85" relaxProximity="true" >
      <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_spanish_social_security_number" />
          <Match idRef="Keywords_spain_eu_ssn_or_equivalent" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_spanish_social_security_number" />
        </Pattern>
    </Entity>
```

Keywords

Keywords_spain_eu_passport_number

- ssn
- ssn#
- socialsecurityno
- social security no
- social security number
- número de la seguridad social

Spain tax identification number

Article • 08/19/2024

Format

seven or eight digits and one or two letters in the specified pattern

Pattern

Spanish Natural Persons with a Spain National Identity Card:

- eight digits
- one uppercase letter (case-sensitive)

Non-resident Spaniards without a Spain National Identity Card

- one uppercase letter "L" (case-sensitive)
- seven digits
- one uppercase letter (case-sensitive)

Resident Spaniards under the age of 14 years without a Spain National Identity Card:

- one uppercase letter "K" (case-sensitive)
- seven digits
- one uppercase letter (case-sensitive)

Foreigners with a Foreigner's Identification Number

- one uppercase letter that is "X", "Y", or "Z" (case-sensitive)
- seven digits
- one uppercase letter (case-sensitive)

Foreigners without a Foreigner's Identification Number

- one uppercase letter that is "M" (case-sensitive)
- seven digits
- one uppercase letter (case-sensitive)

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_spain_eu_tax_file_number` or `Func_spain_eu_DL_and_NI_number_citizen` finds content that matches the pattern.
- A keyword from `Keywords_spain_eu_tax_file_number` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_spain_eu_tax_file_number` or `Func_spain_eu_DL_and_NI_number_citizen` finds content that matches the pattern.

XML

```
      <!-- Spain Tax Identification Number -->
      <Entity id="10f0d113-b0e1-47dc-872a-a4f45b9376a3"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_spain_eu_tax_file_number" />
          <Match idRef="Keywords_spain_eu_tax_file_number" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_spain_eu_tax_file_number" />
        </Pattern>
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_spain_eu_DL_and_NI_number_citizen" />
          <Match idRef="Keywords_spain_eu_tax_file_number" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_spain_eu_DL_and_NI_number_citizen" />
        </Pattern>
      </Entity>
```

Keywords

Keywords_spain_eu_tax_file_number

- cif
- cifid#
- cifnúmero#
- número de contribuyente
- número de identificación fiscal
- número de impuesto corporativo
- spanishcifid#
- spanishcifid
- spanishcifno#
- spanishcifno
- tax file no
- tax file number
- tax id
- tax identification no
- tax identification number
- tax no#
- tax no
- tax number
- tax registration number
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id
- tin no
- tin#

SQL Server connection string

Article • 08/19/2024

Format

The string `User Id`, `User ID`, `uid`, or `UserId` followed by the characters and strings outlined in the pattern below.

Pattern

- the string `User Id`, `User ID`, `uid`, or `UserId`
- any combination of between 1-200 lower- or uppercase letters, digits, symbols, special characters, or spaces
- the string `Password` or `pwd` where `pwd` isn't preceded by a lowercase letter
- an equal sign (=)
- any character that isn't a dollar sign ($), percent symbol (%), greater than symbol (>), at symbol (@), quotation mark ("), semicolon (;), left brace([), or left bracket ({)
- any combination of 7-128 characters that aren't a semicolon (;), forward slash (/), or quotation mark (")
- a semicolon (;) or quotation mark (")

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `CEP_Regex_SQLServerConnectionString` finds content that matches the pattern.
- A keyword from `CEP_GlobalFilter` isn't found.
- The regular expression `CEP_PasswordPlaceHolder` doesn't find content that matches the pattern.
- The regular expression `CEP_CommonExampleKeywords` doesn't find content that matches the pattern.

SQL

```
<!---SQL Server Connection String>
<Entity id="e76b6205-d3cb-46f2-bd63-c90153f2f97d" patternsProximity="300"
recommendedConfidence="85">
  <Pattern confidenceLevel="85">
        <IdMatch idRef="CEP_Regex_SQLServerConnectionString" />
        <Any minMatches="0" maxMatches="0">
            <Match idRef="CEP_GlobalFilter" />
            <Match idRef="CEP_PasswordPlaceHolder" />
            <Match idRef="CEP_CommonExampleKeywords" />
        </Any>
    </Pattern>
</Entity>
```

Keywords

CEP_GlobalFilter

- some-password
- somepassword
- secretPassword
- sample

CEP_PasswordPlaceHolder

This sensitive information type identifies these keywords by using a regular expression, not a keyword list.

- `Password` or `pwd` followed by 0-2 spaces, an equal sign (=), 0-2 spaces, and an asterisk (*) -OR-
- `Password` or `pwd` followed by:
  - Equal sign (=)
  - Less than symbol (<)

    - Any combination of 1-200 characters that are upper- or lowercase letters, digits, an asterisk (*), hyphen (-), underline (_), or whitespace character
    - Greater than symbol (>)

CEP_CommonExampleKeywords

This sensitive information type identifies these keywords by using a regular expression, not a keyword list.

- contoso
- fabrikam
- northwind
- sandbox
- onebox
- localhost
- 127.0.0.1
- testacs.com
- s-int.net

Surgical procedures

Article • 08/19/2024

This unbundled named entity detects terms related to surgical procedures, such as *appendectomy*. It supports English terms only. It's also included in the All medical terms and conditions bundled named entity SIT.

Confidence level

High

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Sweden drivers license number

Article • 08/19/2024

Format

10 digits containing a hyphen

Pattern

10 digits containing a hyphen:

- six digits
- a hyphen
- four digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_sweden_eu_driver's_license_number` finds content that matches the pattern.
- A keyword from `Keywords_eu_driver's_license_number` or `Keywords_sweden_eu_driver's_license_number` is found.

```
XML
```

```
      <!-- Sweden Driver's License Number -->
      <Entity id="70088720-90dd-47f5-805e-5525f3567391"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_sweden_eu_driver's_license_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_driver's_license_number" />
            <Match idRef="Keywords_sweden_eu_driver's_license_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic
- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic

- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#
- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#

- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic
- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen

- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses
- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

Keywords_sweden_eu_driver's_license_number

- ajokortti
- permis de conducere
- ajokortin numero
- kuljettajat lic.
- drivere lic.
- körkort
- numărul permisului de conducere
- שאָפער דערלויבעניש נומער
- förare lic.
- דריווערס דערלויבעניש
- körkortsnummer

Sweden national ID

Article • 08/19/2024

Format

10 or 12 digits and an optional delimiter

Pattern

10 or 12 digits and an optional delimiter:

- two digits (optional)
- Six digits in date format YYMMDD
- delimiter of "-" or "+" (optional)
- four digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_swedish_national_identifier` finds content that matches the pattern.
- A keyword from `Keywords_swedish_national_identifier` is found
- The checksum passes.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_swedish_national_identifier` finds content that matches the pattern.
- The checksum passes.

XML

```
    <!-- Sweden National ID -->
    <Entity id="f69aaf40-79be-4fac-8f05-fd1910d272c8"
patternsProximity="300" recommendedConfidence="85">
      <Pattern confidenceLevel="85">
        <IdMatch idRef="Func_swedish_national_identifier" />
        <Match idRef="Keywords_swedish_national_identifier" />
      </Pattern>
      <Pattern confidenceLevel="75">
        <IdMatch idRef="Func_swedish_national_identifier" />
      </Pattern>
    </Entity>
```

Keywords

Keywords_swedish_national_identifier

- id no
- id number
- id#
- identification no
- identification number
- identifikationsnumret#
- identifikationsnumret
- identitetshandling
- identity document
- identity no
- identity number
- id-nummer
- personal id
- personnummer#
- personnummer
- skatteidentifikationsnummer

Sweden passport number

Article • 08/19/2024

Format

eight digits

Pattern

eight consecutive digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- the regular expression Regex_sweden_passport_number finds content that matches the pattern.
- a keyword from `Keywords_eu_passport_number` or `Keyword_sweden_passport` is found.
- the regular expression `Regex_sweden_eu_passport_date` finds a date in the format DD MMM/MMM YY (01 JAN/JAN 12) or a keyword from `Keywords_eu_passport_date` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- the regular expression Regex_sweden_passport_number finds content that matches the pattern.
- a keyword from `Keywords_eu_passport_number` or `Keyword_sweden_passport` is found.

XML

```
    <!-- Sweden Passport Number -->
    <Entity id="ba4e7456-55a9-4d89-9140-c33673553526"
patternsProximity="300" recommendedConfidence="75">
      <Pattern confidenceLevel="85">
          <IdMatch idRef="Regex_sweden_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keyword_sweden_passport" />
          </Any>
          <Any minMatches="1">
            <Match idRef="Regex_sweden_eu_passport_date" />
            <Match idRef="Keywords_eu_passport_date" />
          </Any>
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_sweden_passport_number" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keyword_sweden_passport" />
          </Any>
      </Pattern>
    </Entity>
```

Keywords

Keywords_eu_passport_number

- passport#
- passport #
- passportid
- passports
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers

Keyword_sweden_passport

- alien registration card
- g3 processing fees
- multiple entry
- Numéro de passeport
- passeport n °
- passeport non
- passeport #
- passeport#
- passeportnon
- passeportn °
- passnummer
- pass nr
- schengen visa
- schengen visas
- single entry
- sverige pass
- visa requirements
- visa processing
- visa type

Keywords_eu_passport_date

- date of issue
- date of expiry

Sweden physical addresses

Article • 08/19/2024

This unbundled named entity detects patterns related to physical address from Sweden. It's also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Sweden tax identification number

Article • 08/19/2024

Format

10 digits and a symbol in the specified pattern

Pattern

10 digits and a symbol:

- six digits that correspond to the birth date (YYMMDD)
- a plus sign or minus sign
- three digits that make the identification number unique where:
  - for numbers issued before 1990, the seventh and eighth digit identify the county of birth or foreign-born people
  - the digit in the ninth position indicates gender by either odd for male or even for female
- one check digit

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_sweden_eu_tax_file_number` finds content that matches the pattern.

- A keyword from `Keywords_sweden_eu_tax_file_number` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_sweden_eu_tax_file_number` finds content that matches the pattern.

XML

```
      <!-- Sweden Tax Identification Number -->
      <Entity id="139acba0-a5bc-4fbb-876d-f7a493ae8a40"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_sweden_eu_tax_file_number" />
          <Match idRef="Keywords_sweden_eu_tax_file_number" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_sweden_eu_tax_file_number" />
          <Any minMatches="0" maxMatches="0">
            <Match idRef="Keywords_sweden_eu_telephone_number" />
            <Match idRef="Keywords_sweden_eu_mobile_number" />
          </Any>
        </Pattern>
      </Entity>
```

Keywords

Keywords_sweden_eu_tax_file_number

- personal id number
- personnummer
- skatt id nummer
- skatt identifikation
- skattebetalarens identifikationsnummer
- sverige tin
- tax file
- tax id
- tax identification no
- tax identification number
- tax no#
- tax no
- tax number
- tax registration number

- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id
- tin no
- tin#

SWIFT code

Article • 08/19/2024

Format

four letters followed by 5-31 letters or digits

Pattern

four letters followed by 5-31 letters or digits:

- four-letter bank code (not case-sensitive)
- an optional space
- 4-28 letters or digits (the Basic Bank Account Number (BBAN))
- an optional space
- one to three letters or digits (remainder of the BBAN)

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_swift` finds content that matches the pattern.
- A keyword from `Keyword_swift` is found.

```
XML
```

```
<Entity id="cb2ab58c-9cb8-4c81-baf8-a4e106791df4" patternsProximity="300"
recommendedConfidence="75">
<Pattern confidenceLevel="75">
        <IdMatch idRef="Regex_swift" />
        <Match idRef="Keyword_swift" />
    </Pattern>
</Entity>
```

Keywords

Keyword_swift

- international organization for standardization 9362
- iso 9362
- iso9362
- swift#
- swiftcode
- swiftnumber
- swiftroutingnumber
- swift code
- swift number #
- swift routing number
- bic number
- bic code
- bic #
- bic#
- bank identifier code
- Organisation internationale de normalisation 9362
- rapide #
- code SWIFT
- le numéro de swift
- swift numéro d'acheminement
- le numéro BIC
- # BIC
- code identificateur de banque
- SWIFTコード
- SWIFT番号
- BIC番号
- BICコード
- SWIFT コード

- SWIFT 番号
- BIC 番号
- BIC コード
- 金融機関識別コード
- 金融機関コード
- 銀行コード

Switzerland physical addresses

Article • 08/19/2024

This unbundled named entity detects patterns related to physical address from Switzerland. It's also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Swiss Social Security Number AHV

The Swiss Social Security Number AHV (Alters- und Hinterlassenenversicherung) is a unique personal identification number used in Switzerland.

Format

13-digit number

Pattern

13-digit number:

- three digits - 756
- an optional dot
- four digits
- an optional dot
- four digits
- an optional dot
- two digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A data loss prevention (DLP) policy has high confidence that it detects this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_swiss_social_security_number_ahv` finds content that matches the pattern.
- A keyword from `Keywords_swiss_social_security_number_ahv` is found.

A DLP policy has medium confidence that it detects this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_swiss_social_security_number_ahv` finds content that matches the pattern.

XML

```
      <!-- Swiss SSN AHV Number -->
      <Entity id="277cfa4b-6eaa-4a1b-9492-099dec849971" patternsProximity="300"
recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_swiss_social_security_number_ahv" />
          <Match idRef="Keywords_swiss_social_security_number_ahv" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_swiss_social_security_number_ahv" />
        </Pattern>
      </Entity>
```

Keywords

The following keywords are used for detection. These include Social Security Number (SSN) and AHV variations in multiple languages.

Keyword_swiss_ssn_AHV_number

- `ahv`
- `ssn`
- `pid`
- `insurance number`
- `personalidno#`
- `social security number`
- `personal id number`
- `personal identification no.`
- `insuranceno#`
- `uniqueidno#`
- `unique identification no.`

- `avs number`
- `personal identity no versicherungsnummer`
- `identifikationsnummer`
- `einzigartige identität nicht`
- `sozialversicherungsnummer`
- `identification personnelle id`
- `numéro de sécurité sociale`

Last updated on 06/19/2026

Taiwan National ID

Format

One letter (in English) followed by nine digits

Pattern

One letter (in English) followed by nine digits:

- One letter (in English, not case-sensitive)
- the digit "1" or "2"
- eight digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A data loss prevention (DLP) policy has high confidence that it detects this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_taiwanese_national_id` finds content that matches the pattern.
- A keyword from `Keyword_taiwanese_national_id` is found.
- The checksum passes.

A DLP policy has medium confidence that it detects this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_taiwanese_national_id` finds content that matches the pattern.

- The checksum passes.

XML

```xml
<!-- Taiwanese National ID -->
<Entity id="4C7BFC34-8DD1-421D-8FB7-6C6182C2AF03" patternsProximity="300"
recommendedConfidence="85">
      <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_taiwanese_national_id" />
          <Match idRef="Keyword_taiwanese_national_id" />
      </Pattern>
       <Pattern confidenceLevel="75">
         <IdMatch idRef="Func_taiwanese_national_id" />
       </Pattern>
</Entity>
```

Keywords

Keyword_taiwan_national_id

- 身份證字號
- 身份證
- 身份證號碼
- 身份證號
- 身分證字號
- 身分證
- 身分證號碼
- 身份證號
- 身分證統一編號
- 國民身分證統一編號
- 簽名
- 蓋章
- 簽名或蓋章
- 簽章

Last updated on 06/19/2026

Taiwan passport number

Article • 08/19/2024

Format

- biometric passport number: nine digits
- non-biometric passport number: nine digits

Pattern

biometric passport number:

- the character "3"
- eight digits

non-biometric passport number:

- nine digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_taiwan_passport` finds content that matches the pattern.
- A keyword from `Keyword_taiwan_passport` is found.

XML

```xml
<!-- Taiwan Passport Number -->
<Entity id="e7251cb4-4c2c-41df-963e-924eb3dae04a" recommendedConfidence="75"
patternsProximity="300">
  <Pattern confidenceLevel="75">
     <IdMatch idRef="Regex_taiwan_passport"/>
     <Match idRef="Keyword_taiwan_passport"/>
  </Pattern>
</Entity>
```

Keywords

Keyword_taiwan_passport

- ROC passport number
- Passport number
- Passport no
- Passport Num
- Passport #
- 护照
- 中華民國護照
- Zhōnghuá Mínguó hùzhào

Taiwan-resident certificate (ARC/TARC) number

Article • 08/19/2024

Format

10 letters and digits

Pattern

10 letters and digits:

- two letters (not case-sensitive)
- eight digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_taiwan_resident_certificate` finds content that matches the pattern.
- A keyword from `Keyword_taiwan_resident_certificate` is found.

```
XML
```

```
<!-- Taiwan Resident Certificate (ARC/TARC) -->
<Entity id="48269fec-05ea-46ea-b326-f5623a58c6e9" recommendedConfidence="75"
patternsProximity="300">
  <Pattern confidenceLevel="75">
     <IdMatch idRef="Regex_taiwan_resident_certificate"/>
     <Match idRef="Keyword_taiwan_resident_certificate"/>
  </Pattern>
</Entity>
```

Keywords

Keyword_taiwan_resident_certificate

- Resident Certificate
- Resident Cert
- Resident Cert.
- Identification card
- Alien Resident Certificate
- ARC
- Taiwan Area Resident Certificate
- TARC
- 居留證
- 外僑居留證
- 台灣地區居留證

Thai population identification code

Article • 08/19/2024

Format

13 digits

Pattern

13 digits:

- first digit isn't zero or nine
- 12 digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_Thai_Citizen_Id` finds content that matches the pattern.
- A keyword from `Keyword_Thai_Citizen_Id` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_Thai_Citizen_Id` finds content that matches the pattern.

XML

```xml
<!-- Thai Citizen ID -->
-<Entity id="44ca9e86-ead7-4c5d-884a-e2eaa401515e"
recommendedConfidence="75" patternsProximity="300">
   -<Pattern confidenceLevel="85">
      <IdMatch idRef="Func_Thai_Citizen_Id"/>
      <Match idRef="Keyword_Thai_Citizen_Id"/>
   </Pattern>
   -<Pattern confidenceLevel="75">
      <IdMatch idRef="Func_Thai_Citizen_Id"/>
   </Pattern>
</Entity>
```

Keywords

Keyword_thai_citizen_Id

- ID Number
- Identification Number
- **บัตรประชาชน**
- **รหัสบัตรประชาชน**
- **บัตรประชาชน**
- **รหัสบัตรประชาชน**

Türkiye national identification number

Format

11 digits

Pattern

11 digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_Turkish_National_Id` finds content that matches the pattern.
- A keyword from `Keyword_Turkish_National_Id` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_Turkish_National_Id` finds content that matches the pattern.

XML

```
<!-- Turkish National Identity -->
-<Entity id="fb621f20-3876-4cfc-acec-8c8e73ca32c7" recommendedConfidence="75"
patternsProximity="300">
```

```
    -<Pattern confidenceLevel="85">
       <IdMatch idRef="Func_Turkish_National_Id"/>
       <Match idRef="Keyword_Turkish_National_Id"/>
    </Pattern>
    -<Pattern confidenceLevel="75">
       <IdMatch idRef="Func_Turkish_National_Id"/>
    </Pattern>
</Entity>
```

Keywords

Keyword_turkish_national_id

- TC Kimlik No
- TC Kimlik numarası
- Vatandaşlık numarası
- Vatandaşlık no

Last updated on 09/30/2026

Türkiye physical addresses

This unbundled named entity detects patterns related to physical address from Türkiye. It's also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Last updated on 09/30/2026

Types of medication

Article • 08/19/2024

This unbundled named entity detects medication names, such as *insulin*. It supports English terms only. It's also included in the All medical terms and conditions bundled named entity SIT.

Confidence level

High

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

U.A.E. identity card number

Article • 08/19/2024

Format

Fifteen digits with optional hyphens and must pass the Luhn check

Pattern

Fifteen digits pattern

- three digits
- optional hyphen
- four digits
- optional hyphen
- seven digits
- optional hyphen
- one digit

Checksum

Luhn check

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_uae_id_card_number` finds content that matches the pattern.
- A keyword from Keyword_uae_id_card_number is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_uae_id_card_number` finds content that matches the pattern.

XML

```
     <!-- UAE Identity Card Number -->
      <Entity id="853a8051-ad90-417c-9345-c73ac0adc1c3"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Regex_uae_id_card_number" />
          <Match idRef="Keyword_uae_id_card_number" />
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_uae_id_card_number" />
        </Pattern>
      </Entity>
```

Keywords

Keyword_uae_id_card_number

- Emirates id
- ID number
- id no
- id #
- id#
- identification number
- identification no
- identification #
- identification#
- emirates identity card
- Insurance Number
- insurance no
- insurance#
- insurance #
- التأمين رقم
- #التأمينرقم
- رقم الهوية
- #رقم الهوية

U.A.E. passport number

Article • 08/19/2024

Format

Nine alphanumeric characters that must include letters of specific values

Pattern

Nine alphanumeric characters from numerals 0–9 and letters C, F, G, H, J, K, L, M, N, P, R, T, V, W, X, Y, Z

Checksum

All nine characters shouldn't be digits.

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_uae_passport_number` finds content that matches the pattern.
- A keyword from `Keyword_uae_passport_number` is found.

A DLP policy has low confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

The regular expression `Regex_uae_passport_number` finds content that matches the pattern.

XML

```xml
  <!-- UAE Passport Number -->
      <Entity id="111b5e63-6f4e-47c6-99bc-dd95625a5acf"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Regex_uae_passport_number" />
          <Match idRef="Keyword_uae_passport_number" />
        </Pattern>
        <Pattern confidenceLevel="65">
          <IdMatch idRef="Regex_uae_passport_number" />
        </Pattern>
      </Entity>
```

Keywords

Keyword_uae_passport_number

- جواز سفر
- #جواز سفر
- passport#
- passport #
- passportno
- passport no
- passportnumber
- passport number
- passportnumbers
- passport numbers

U.K. drivers license number

Article • 08/19/2024

Format

Combination of 18 letters and digits in the specified format

Pattern

18 letters and digits:

- Five letters (not case-sensitive) or the digit "9" in place of a letter.
- One digit.
- Five digits in the date format MMDDY for date of birth. The seventh character is incremented by 50 if driver is female; for example, 51 to 62 instead of 01 to 12.
- Two letters (not case-sensitive) or the digit "9" in place of a letter.
- Five digits.

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_uk_drivers_license` finds content that matches the pattern.
- A keyword from `Keywords_eu_driver's_license_number` is found.
- The checksum passes.

A DLP policy has low confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_uk_drivers_license` finds content that matches the pattern.
- The checksum passes.

XML

```
    <!-- U.K. Driver's License Number -->
    <Entity id="f93de4be-d94c-40df-a8be-461738047551"
patternsProximity="300" recommendedConfidence="75" relaxProximity="true" >
      <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_uk_drivers_license" />
          <Match idRef="Keywords_eu_driver's_license_number" />
        </Pattern>
        <Pattern confidenceLevel="65">
          <IdMatch idRef="Func_uk_drivers_license" />
        </Pattern>
    </Entity>
```

Keywords

Keywords_eu_driver's_license_number

- driverlic
- driverlics
- driverlicense
- driverlicenses
- driverlicence
- driverlicences
- driver lic
- driver lics
- driver license
- driver licenses
- driver licence
- driver licences
- driverslic
- driverslics
- driverslicence
- driverslicences
- driverslicense
- driverslicenses
- drivers lic

- drivers lics
- drivers license
- drivers licenses
- drivers licence
- drivers licences
- driver'lic
- driver'lics
- driver'license
- driver'licenses
- driver'licence
- driver'licences
- driver' lic
- driver' lics
- driver' license
- driver' licenses
- driver' licence
- driver' licences
- driver'slic
- driver'slics
- driver'slicense
- driver'slicenses
- driver'slicence
- driver'slicences
- driver's lic
- driver's lics
- driver's license
- driver's licenses
- driver's licence
- driver's licences
- dl#
- dls#
- driverlic#
- driverlics#
- driverlicense#
- driverlicenses#
- driverlicence#
- driverlicences#
- driver lic#
- driver lics#
- driver license#
- driver licenses#

- driver licences#
- driverslic#
- driverslics#
- driverslicense#
- driverslicenses#
- driverslicence#
- driverslicences#
- drivers lic#
- drivers lics#
- drivers license#
- drivers licenses#
- drivers licence#
- drivers licences#
- driver'lic#
- driver'lics#
- driver'license#
- driver'licenses#
- driver'licence#
- driver'licences#
- driver' lic#
- driver' lics#
- driver' license#
- driver' licenses#
- driver' licence#
- driver' licences#
- driver'slic#
- driver'slics#
- driver'slicense#
- driver'slicenses#
- driver'slicence#
- driver'slicences#
- driver's lic#
- driver's lics#
- driver's license#
- driver's licenses#
- driver's licence#
- driver's licences#
- driving licence
- driving license
- dlno#
- driv lic

- driv licen
- driv license
- driv licenses
- driv licence
- driv licences
- driver licen
- drivers licen
- driver's licen
- driving lic
- driving licen
- driving licenses
- driving licence
- driving licences
- driving permit
- dl no
- dlno
- dl number

U.K. electoral roll number

Article • 08/19/2024

Format

two letters followed by 1-4 digits

Pattern

two letters (not case-sensitive) followed by 1-4 numbers

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_uk_electoral` finds content that matches the pattern.
- A keyword from `Keyword_uk_electoral` is found.

XML

```
<!-- U.K. Electoral Number -->
<Entity id="a3eea206-dc0c-4f06-9e22-aa1be3059963" patternsProximity="300"
recommendedConfidence="75">
    <Pattern confidenceLevel="75">
        <IdMatch idRef="Regex_uk_electoral" />
        <Any minMatches="1">
          <Match idRef="Keyword_uk_electoral" />
        </Any>
```

```
    </Pattern>
</Entity>
```

Keywords

Keyword_uk_electoral

- council nomination
- nomination form
- electoral register
- electoral roll

U.K. national health service number

Format

Ten digits separated by optional spaces

Pattern

Ten digits:

- three digits
- a space (optional)
- three digits
- a space (optional)
- four digits

Checksum

Yes

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_uk_nhs_number` finds content that matches the pattern.
- One of the following cases is true:
  - A keyword from `Keyword_uk_nhs_number` is found.
  - A keyword from `Keyword_uk_nhs_number1` is found.

    - A keyword from `Keyword_uk_nhs_number_dob` is found.
- The checksum passes.

XML

```xml
<!-- U.K. NHS Number -->
<Entity id="3192014e-2a16-44e9-aa69-4b20375c9a78" patternsProximity="300"
recommendedConfidence="85">
    <Pattern confidenceLevel="85">
        <IdMatch idRef="Func_uk_nhs_number" />
        <Any minMatches="1">
          <Match idRef="Keyword_uk_nhs_number" />
          <Match idRef="Keyword_uk_nhs_number1" />
          <Match idRef="Keyword_uk_nhs_number_dob" />
        </Any>
    </Pattern>
</Entity>
```

Keywords

Keyword_uk_nhs_number

- national health service
- nhs
- health services authority
- health authority

Keyword_uk_nhs_number1

- patient id
- patient identification
- patient no
- patient number

Keyword_uk_nhs_number_dob

- GP
- DOB
- D.O.B
- Date of Birth
- Birth Date

Last updated on 12/10/2025

U.K. national insurance number (NINO)

Article • 08/19/2024

This sensitive information type entity is included in the EU National Identification Number sensitive information type. It's also available as a stand-alone sensitive information type entity.

Format

seven characters or nine characters separated by spaces or dashes

Pattern

two possible patterns:

- two letters (valid NINOs use only certain characters in this prefix, which this pattern validates; not case-sensitive)
- six digits
- either 'A', 'B', 'C', or 'D' (like the prefix, only certain characters are allowed in the suffix; not case-sensitive)

OR

- two letters
- a space or dash
- two digits
- a space or dash
- two digits
- a space or dash
- two digits
- a space or dash
- either 'A', 'B', 'C', or 'D'

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_uk_nino` finds content that matches the pattern.
- A keyword from `Keyword_uk_nino` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_uk_nino` finds content that matches the pattern.

XML

```
    <!-- U.K. NINO -->
    <Entity id="16c07343-c26f-49d2-a987-3daf717e94cc"
patternsProximity="300" recommendedConfidence="75" relaxProximity="true">
      <Pattern confidenceLevel="85">
        <IdMatch idRef="Func_uk_nino" />
        <Match idRef="Keyword_uk_nino" />
      </Pattern>
      <Pattern confidenceLevel="75">
        <IdMatch idRef="Func_uk_nino" />
      </Pattern>
    </Entity>
```

Keywords

Keyword_uk_nino

- national insurance number
- national insurance contributions
- protection act
- insurance
- social security number
- insurance application

- medical application
- social insurance
- medical attention
- social security
- great britain
- NI Number
- NI No.
- NI #
- NI#
- insurance#
- insurancenumber
- nationalinsurance#
- nationalinsurancenumber

U.K. physical addresses

Article • 08/19/2024

This unbundled named entity detects patterns related to physical address from the U.K.. It's also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

***UNKNOWN**

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

U.K. Unique Taxpayer Reference Number

Article • 08/19/2024

Format

10 digits without spaces and delimiters

Pattern

10 digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_uk_eu_tax_file_number` finds content that matches the pattern.
- A keyword from `Keywords_uk_eu_tax_file_number` is found.

XML

```
      <!-- U.K. Unique Taxpayer Reference Number -->
      <Entity id="ad4a8116-0db8-439a-b545-6d967642f0ec"
patternsProximity="300" recommendedConfidence="85">
        <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_uk_eu_tax_file_number" />
          <Match idRef="Keywords_uk_eu_tax_file_number" />
```

```
          </Pattern>
        </Entity>
```

Keywords

Keywords_uk_eu_tax_file_number

- tax number
- tax file
- tax id
- tax identification no
- tax identification number
- tax no#
- tax no
- tax registration number
- taxid#
- taxidno#
- taxidnumber#
- taxno#
- taxnumber#
- taxnumber
- tin id
- tin no
- tin#

U.S. bank account number

Article • 08/19/2024

Format

6-17 digits

Pattern

6-17 consecutive digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regular expression `Regex_usa_bank_account_number` finds content that matches the pattern.
- A keyword from `Keyword_usa_Bank_Account` is found.

XML

```
<!-- U.S. Bank Account Number -->
<Entity id="a2ce32a8-f935-4bb6-8e96-2a5157672e2c" patternsProximity="300"
recommendedConfidence="75">
    <Pattern confidenceLevel="75">
        <IdMatch idRef="Regex_usa_bank_account_number" />
        <Match idRef="Keyword_usa_Bank_Account" />
```

```
    </Pattern>
</Entity>
```

Keywords

Keyword_usa_Bank_Account

- Checking Account Number
- Checking Account
- Checking Account #
- Checking Acct Number
- Checking Acct #
- Checking Acct No.
- Checking Account No.
- Bank Account Number
- Bank Account #
- Bank Acct Number
- Bank Acct #
- Bank Acct No.
- Bank Account No.
- Savings Account Number
- Savings Account.
- Savings Account #
- Savings Acct Number
- Savings Acct #
- Savings Acct No.
- Savings Account No.
- Debit Account Number
- Debit Account
- Debit Account #
- Debit Acct Number
- Debit Acct #
- Debit Acct No.
- Debit Account No.

U.S. drivers license number

Article • 08/19/2024

Format

Depends on the state

Pattern

depends on the state - for example, New York:

- nine digits formatted like ddd ddd ddd will match.
- nine digits like ddddddddd won't match.

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_new_york_drivers_license_number` finds content that matches the pattern.
- A keyword from `Keyword_[state_name]_drivers_license_name` is found.
- A keyword from `Keyword_us_drivers_license` is found.

A DLP policy has low confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_new_york_drivers_license_number` finds content that matches the pattern.
- A keyword from `Keyword_[state_name]_drivers_license_name` is found.
- A keyword from `Keyword_us_drivers_license_abbreviations` is found.
- No keyword from `Keyword_us_drivers_license` is found.

XML

```xml
<Entity id="dfeb356f-61cd-459e-bf0f-7c6d28b458c6 patternsProximity="300">
    <Pattern confidenceLevel="75">
        <IdMatch idRef="Func_new_york_drivers_license_number" />
        <Match idRef="Keyword_new_york_drivers_license_name" />
        <Match idRef="Keyword_us_drivers_license" />
    </Pattern>
    <Pattern confidenceLevel="65">
        <IdMatch idRef="Func_new_york_drivers_license_number" />
        <Match idRef="Keyword_new_york_drivers_license_name" />
        <Match idRef="Keyword_us_drivers_license_abbreviations" />
        <Any minMatches="0" maxMatches="0">
          <Match idRef="Keyword_us_drivers_license" />
        </Any>
    </Pattern>
</Entity>
```

Keywords

Keyword_us_drivers_license_abbreviations

- DL
- DLS
- CDL
- CDLS
- ID
- IDs
- DL#
- DLS#
- CDL#
- CDLS#
- ID#
- IDs#
- ID number
- ID numbers
- LIC

- LIC#
- DLN

Keyword_us_drivers_license

- DriverLic
- DriverLics
- DriverLicense
- DriverLicenses
- Driver Lic
- Driver Lics
- Driver License
- Driver Licenses
- DriversLic
- DriversLics
- DriversLicense
- DriversLicenses
- Drivers Lic
- Drivers Lics
- Drivers License
- Drivers Licenses
- Driver'Lic
- Driver'Lics
- Driver'License
- Driver'Licenses
- Driver' Lic
- Driver' Lics
- Driver' License
- Driver' Licenses
- Driver'sLic
- Driver'sLics
- Driver'sLicense
- Driver'sLicenses
- Driver's Lic
- Driver's Lics
- Driver's License
- Driver's Licenses
- identification number
- identification numbers
- identification #

- id card
- id cards
- identification card
- identification cards
- DriverLic#
- DriverLics#
- DriverLicense#
- DriverLicenses#
- Driver Lic#
- Driver Lics#
- Driver License#
- Driver Licenses#
- DriversLic#
- DriversLics#
- DriversLicense#
- DriversLicenses#
- Drivers Lic#
- Drivers Lics#
- Drivers License#
- Drivers Licenses#
- Driver'Lic#
- Driver'Lics#
- Driver'License#
- Driver'Licenses#
- Driver' Lic#
- Driver' Lics#
- Driver' License#
- Driver' Licenses#
- Driver'sLic#
- Driver'sLics#
- Driver'sLicense#
- Driver'sLicenses#
- Driver's Lic#
- Driver's Lics#
- Driver's License#
- Driver's Licenses#
- id card#
- id cards#
- identification card#
- identification cards#

Keyword_[state_name]_drivers_license_name

- state abbreviation (for example, "NY")
- state name (for example, "New York")

U.S. individual taxpayer identification number (ITIN)

Article • 08/19/2024

Format

nine digits, optionally formatted with spaces or dashes

Pattern

formatted:

- the digit "9"
- two digits
- a space or dash (optional)
- two digits "50" to "65", "70" to "88", “90” to “92” and “94” to “99” for the fourth and fifth digit
- a space, or dash (optional)
- four digits

unformatted:

- the digit "9"
- two digits
- two digits "50" to "65", "70" to "88", “90” to “92”, and “94” to “99” for the fourth and fifth digits
- four digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_formatted_itin` finds content that matches the pattern.
- A keyword from `Keyword_itin` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_unformatted_itin` finds content that matches the pattern.
- A keyword from `Keyword_itin` is found.

A DLP policy has low confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_formatted_itin` or `Func_unformatted_itin` finds content that matches the pattern.

XML

```
    <!-- U.S. Individual Taxpayer Identification Number (ITIN) -->
    <Entity id="e55e2a32-f92d-4985-a35d-a0b269eb687b"
patternsProximity="300" recommendedConfidence="75">
      <Pattern confidenceLevel="85">
        <IdMatch idRef="Func_formatted_itin" />
        <Match idRef="Keyword_itin" />
      </Pattern>
      <Pattern confidenceLevel="75">
        <IdMatch idRef="Func_unformatted_itin" />
        <Match idRef="Keyword_itin" />
      </Pattern>
      <Pattern confidenceLevel="65">
        <IdMatch idRef="Func_formatted_itin" />
      </Pattern>
      <Pattern confidenceLevel="65">
        <IdMatch idRef="Func_unformatted_itin" />
      </Pattern>
    </Entity>
```

Keywords

Keyword_itin

- taxpayer

- tax id
- tax identification
- itin
- i.t.i.n.
- ssn
- tin
- social security
- tax payer
- itins
- taxid
- individual taxpayer

U.S. physical addresses

Article • 08/19/2024

This unbundled named entity detects patterns related to physical address from the U.S.. It's also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

***UNKNOWN**

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

U.S. social security number (SSN)

Article • 08/19/2024

Format

nine digits, which may be in a formatted or unformatted pattern

> **Note**
>
> If issued before mid-2011, an SSN has strong formatting where certain parts of the number must fall within certain ranges to be valid (but there's no checksum).

Pattern

four functions look for SSNs in four different patterns:

- `Func_ssn` finds SSNs with pre-2011 strong formatting that are formatted with dashes or spaces (ddd-dd-dddd OR ddd dd dddd)
- `Func_unformatted_ssn` finds SSNs with pre-2011 strong formatting that are unformatted as nine consecutive digits (ddddddddd)
- `Func_randomized_formatted_ssn` finds post-2011 SSNs that are formatted with dashes or spaces (ddd-dd-dddd OR ddd dd dddd)
- `Func_randomized_unformatted_ssn` finds post-2011 SSNs that are unformatted as nine consecutive digits (ddddddddd)

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has high confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_ssn` finds content that matches the pattern.
- A keyword from `Keyword_ssn` is found.

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_unformatted_ssn` finds content that matches the pattern.
- A keyword from `Keyword_ssn` is found.

A DLP policy has low confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_randomized_formatted_ssn` or `Func_randomized_unformatted_ssn` finds content that matches the pattern.
- A keyword from `Keyword_ssn` is found.

XML

```
<!-- U.S. Social Security Number (SSN) -->
  <Entity id="a44669fe-0d48-453d-a9b1-2cc83f2cba77" patternsProximity="300"
recommendedConfidence="75">
      <Pattern confidenceLevel="85">
        <IdMatch idRef="Func_ssn" />
        <Match idRef="Keyword_ssn" />
      </Pattern>
      <Pattern confidenceLevel="75">
        <IdMatch idRef="Func_unformatted_ssn" />
        <Match idRef="Keyword_ssn" />
      </Pattern>
      <Pattern confidenceLevel="65">
        <IdMatch idRef="Func_randomized_formatted_ssn" />
        <Match idRef="Keyword_ssn" />
      </Pattern>
      <Pattern confidenceLevel="55">
        <IdMatch idRef="Func_randomized_unformatted_ssn" />
        <Match idRef="Keyword_ssn" />
      </Pattern>
  </Entity>
```

Keywords

Keyword_ssn

- SSA Number
- social security number
- social security #
- social security#
- social security no
- Social Security#
- Soc Sec
- SSN
- SSNS
- SSN#
- SS#
- SSID

U.S./U.K. passport number

Article • 11/22/2024

Format

Nine digits

Pattern

- one letter or digit
- eight digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A Data Loss Prevention (DLP) policy has high confidence that it detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_usa_uk_passport` finds content that matches the pattern.
- A keyword from `Keywords_eu_passport_number` or `Keywords_uk_eu_passport_number` is found.
- A keyword from `Keywords_eu_passport_date` is found

A DLP policy has medium confidence that it detected this type of sensitive information if, within a proximity of 300 characters:

- The function `Func_usa_uk_passport` finds content that matches the pattern.

- A keyword from `Keywords_eu_passport_number` or `Keywords_uk_eu_passport_number` is found.

XML

```
    <!-- U.S. / U.K. Passport Number -->
    <Entity id="178ec42a-18b4-47cc-85c7-d62c92fd67f8"
patternsProximity="300" recommendedConfidence="75">
       <Pattern confidenceLevel="85">
          <IdMatch idRef="Func_usa_uk_passport" />
          <Match idRef="Keywords_eu_passport_date" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_uk_eu_passport_number" />
          </Any>
        </Pattern>
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Func_usa_uk_passport" />
          <Any minMatches="1">
            <Match idRef="Keywords_eu_passport_number" />
            <Match idRef="Keywords_uk_eu_passport_number" />
          </Any>
        </Pattern>
    </Entity>
```

Keywords

Keywords_eu_passport_number

- passport#
- passport #
- passportid
- passports
- passport no
- passport no
- passportnumber
- passport number
- passport numbers
- passport numbers

Keywords_uk_eu_passport_number

- british passport
- uk passport

Keywords_eu_passport_date

- date of issue
- date of expiry

Ukraine passport domestic

Article • 08/19/2024

Format

nine digits

Pattern

nine digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regex `Regex_Ukraine_Passport_Domestic` finds content that matches the pattern.
- A keyword from `Keyword_Ukraine_Passport_Domestic` is found.

XML

```
      <!-- Ukraine Passport Domestic -->
      <Entity id="1817a540-221f-4459-9202-3bd78b81d803"
patternsProximity="300" recommendedConfidence="75">
        <Pattern confidenceLevel="75">
        <IdMatch idRef="Regex_Ukraine_Passport_Domestic"/>
          <Match idRef="Keyword_Ukraine_Passport_Domestic"/>
```

```
          </Pattern>
        </Entity>
```

Keywords

Keyword_ukraine_passport_domestic

- ukraine passport
- passport number
- passport no
- паспорт України
- номер паспорта
- персональний

Ukraine passport international

Article • 08/19/2024

Format

eight-character alphanumeric pattern

Pattern

eight-character alphanumeric pattern:

- two letters or digits
- six digits

Checksum

No

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

A DLP policy has medium confidence that it's detected this type of sensitive information if, within a proximity of 300 characters:

- The regex `Regex_Ukraine_Passport_International` finds content that matches the pattern.
- A keyword from `Keyword_Ukraine_Passport_International` is found.

XML

```
      <!-- Ukraine Passport International -->
      <Entity id="cfbe032d-22e0-4f28-ab68-d66e9641f1e2"
patternsProximity="300" recommendedConfidence="75">
```

```
        <Pattern confidenceLevel="75">
          <IdMatch idRef="Regex_Ukraine_Passport_International"/>
          <Match idRef="Keyword_Ukraine_Passport_International"/>
        </Pattern>
      </Entity>
```

Keywords

Keyword_ukraine_passport_international

- ukraine passport
- passport number
- passport no
- паспорт України
- номер паспорта

Ukraine physical addresses

This unbundled named entity detects patterns related to physical address from Ukraine. It is also included in the All Physical Addresses bundled named entity SIT.

Confidence level

Medium

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Cloud Support

Cloud support is available for the following services:

- Microsoft Commercial Cloud
- Microsoft Government Community Cloud
- Microsoft 365 Government Community Cloud High
- Microsoft 365 Department of Defense

Last updated on 06/19/2026

User login credentials

This SIT is also included in the All credentials bundled SIT.

> **ⓘ Important**
>
> Because credential scanning SITs are a type of advanced classifier, you must enable advanced classification before implementing them. For information about enabling advanced classification, and about configuring endpoints and the supported file types, see Advanced classification scanning and protection.

Format

A paired username and password used in general authentication process.

or

A paired username and password used in PuTTY connection manager.

or

Plain text password used in code snippets.

or

A combination of 88 characters consisting of letters, digits, and special characters.

Pattern

Various username and password formats, for example:

`username=...;password=********;`
`user id=...;password=********;`
`uid=...;pwd=********;`
`DB_USER=...;DB_PASS=********;`
`Service Account=...;Password=********;`

or

XML

```
An XML element <login>
An embeded XML element <login>
Inner XML content
An embeded XML element </login>
An embeded XML element <password>
Inner XML content
An embeded XML element </password>
An XML element </login>
```

for example

`<login> <login>ZYXWVU_1</login> <password>ZY…`

or

Various password formats in code snippets, for example:

`new X509Certificates2(`

`ConvertTo-SecureString -String ********`

`password = "********"`

`"password" : "********"`

`UserPasswordCredential(`

or

A combination of 86 characters:

- a-z (not case-sensitive)
- 0-9
- forward slashes (/)
- or plus signs (+)
- ends with two equal signs (=)

format example:

`aB1cD2eF/3gH4iJ5kL6+mN7oP8qR==`

Credential example

|Confidence Band|Example| |----|----| |High| `initial catalog=host_name;database=db_name;uid=user_name;password=********** | |Medium|` user=user_name;password=********** | |Low| N/A |

Checksum

No

SITs that have checksums use a unique calculation to check if the information is valid. This means when the **Checksum** value is **Yes**, the service can make a positive detection based on the sensitive data alone. When the **Checksum** value is **No** additional (secondary) elements must also be detected for the service to make a positive detection.

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

This SIT is designed to match the security information that's used in general user login process.

It uses several primary resources:

- Patterns of Plain-text username and password.
- Patterns of Plain-text username and password in PuTTYcm database file.
- Patterns of Password context in code.
- Patterns of Base64 encoded 512 bits symmetric key.
- Patterns of CredentialName, CredentialFeatures, AccountIdentityName, AccountIdentityValue, ResourceType, ResourceName, Id, AccountName.
- Patterns of mockup values, redactions, and placeholders.
- A dictionary of vocabulary.

The patterns are designed to match actual credentials with reasonable confidence. The patterns don't match credentials formatted as examples. Mockup values, redacted values, and placeholders, like credential type or usage descriptions, in the position where an actual secret value should present won't be matched.

Keywords

Keyword_LoginCredentials:

- password
- pw
- DB_

Keyword_LoginCredentialsPutty:

- login

Keyword_PasswordContextInCode:

- key
- x509c
- credential
- password
- pw
- securestring

Keyword_SymmetricKey512:

- SharedAccessKey
- AccountKey

Last updated on 04/09/2026

X.509 certificate private key (preview)

This SIT is also included in the All credentials bundled SIT.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a credential scanning SIT in an endpoint DLP policy. This requirement is specific to the combination of credential scanning SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

Format

A combination of up to 20,000 characters consisting of letters, digits, and special characters.

or

A combination of up to 40 characters consisting of upper-case letters, space, and dashes.

Pattern

A combination of up to 20,000 characters:

- a-z (not case-sensitive)
- 0-9
- forward slashes (/)
- or plus signs (+)
- Up to two equal signs (==)

format example:

```
cD2eF3gH4iJ5kL6mN7/oP+qR9sT=
```

or

five dashes (-)

And a combination of up to 30 characters:

- A-Z (case-sensitive)
- spaces
- 5 dashes (-)

for example:

`-----BEGIN PRIVATE KEY-----`

Credential format example

`-----BEGIN PRIVATE KEY----- cD2eF3gH4iJ5kL6mN7/oP+qR9sT=...`

> **Important**
>
> This example has been truncated. It's not a detectable example of this SIT.

Checksum

Yes

SITs that have checksums use a unique calculation to check if the information is valid. This means when the **Checksum** value is **Yes**, the service can make a positive detection based on the sensitive data alone. When the **Checksum** value is **No** additional (secondary) elements must also be detected for the service to make a positive detection.

Keyword Highlighting

Supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Definition

This SIT is designed to match the security information that's used as a private component in SSL certificates.

It uses several primary resources:

- Patterns of Base64 encoded string literal.
- Patterns of Certificate private key header.
- Patterns of CredentialName, CredentialFeatures, AccountIdentityName, AccountIdentityValue, ResourceType, ResourceName.
- Patterns of mockup values, redactions, and placeholders.
- A dictionary of vocabulary.

The patterns are designed to match actual credentials with reasonable confidence. The patterns don't match credentials formatted as examples. Mockup values, redacted values, and placeholders, like credential type or usage descriptions, in the position where an actual secret value should present won't be matched.

Keywords

Keyword_Base64EncodedStringLiteral:

- MII

Keyword_CertificatePrivateKeyHeader:

- key

Last updated on 11/18/2025

Sensitive information type functions

Article • 07/21/2023

Sensitive information types (SIT) can use functions as primary elements for identifying sensitive items. For example, the Credit Card Number SIT uses the Func_credit_card function to detect credit card number.

This article explains what these functions look for, so you can better understand how the predefined sensitive information types work. For more information, see Sensitive information type entity definitions

> **Tip**
>
> If you're not an E5 customer, use the 90-day Microsoft Purview solutions trial to explore how additional Purview capabilities can help your organization manage data security and compliance needs. Start now at the **Microsoft Purview trials hub**. Learn details about **signing up and trial terms**.

Table of functions

Expand table

Function name	Function action	Is a validator
Func_aba_routing	detects ABA routing number	yes
Func_alabama_drivers_license_number	detects Alabama driver's license number	no
Func_alaska_delaware_oregon_drivers_license_number	detects Alaska, Delaware, Oregon driver's license number	no
Func_alaska_drivers_license_number	detects Alaska driver's license number	no

Function name	Function action	Is a validator
Func_alberta_drivers_license_number	detects Alberta driver's license number	no
Func_argentina_Unique_Tax_Key	detects and validates Argentina Unique tax key	no
Func_Argentina_Unique_Tax_Key	detects Argentina Unique tax key	no
Func_arizona_drivers_license_number	detects Arizona driver's license number	no
Func_arkansas_drivers_license_number	detects Arkansas driver's license number	no
Func_australian_business_number	detects Australia business number	no
Func_Australian_Company_Number	detects Australia company number	no
Func_australian_medical_account_number	detects Australia medical account number	no
Func_australian_tax_file_number	detects Australia tax file number	yes
Func_austria_eu_ssn_or_equivalent	detects Austria social security number	no
Func_austria_eu_tax_file_number	detects Austria tax file number	no

Function name	Function action	Is a validator
Func_Austria_Value_Added_Tax	detects Austria Value Added Tax	no
Func_belgium_national_number	detects Belgium national number	no
Func_belgium_value_added_tax_number	detects Belgium value added tax number	no
Func_brazil_cnpj	detects Brazil legal entity number (CNPJ)	yes
Func_brazil_cpf	detects Brazil CPF	yes
Func_brazil_rg	detects Brazil RG	no
Func_british_columbia_drivers_license_number	detects British Columbia driver's license number	no
Func_bulgaria_eu_national_id_card	detects Bulgaria uniform civil number	no
Func_california_drivers_license_number	detects California driver's license number	no
Func_canadian_sin	detects Canada sin	yes
Func_chile_id_card	detects Chile ID card	no
Func_china_resident_id	detects China-resident ID	no

Function name	Function action	Is a validator
Func_colorado_drivers_license_number	detects Colorado driver's license number	no
Func_connecticut_drivers_license_number	detects Connecticut driver's license number	no
Func_credit_card	detects credit card	yes
Func_croatia_id_card	detects Croatia ID card	no
Func_croatia_oib_number	detects Croatia OIB number	no
Func_cyprus_eu_tax_file_number	detects Cyprus tax file number	no
Func_czech_id_card_new_format	detects Czech ID card in new format	no
Func_czech_id_card	detects Czech ID card	no
Func_dea_number	detects DEA number	yes
Func_denmark_eu_tax_file_number	detects Denmark personal identification number	no
Func_district_of_columbia_drivers_license_number	detects District of Columbia driver's license number	no
Func_estonia_eu_national_id_card	detects Estonia Personal Identification Code	no

Function name	Function action	Is a validator
Func_eu_debit_card	detects EU debit card	no
Func_finnish_national_id	detects Finnish national ID	no
Func_florida_drivers_license_number	detects Florida driver's license number	no
Func_florida_maryland_michigan_minnesota_drivers_license_number	detects Florida, Maryland, Michigan, Minnesota driver's license number	no
Func_formatted_itin	detects formatted US ITIN	yes
Func_fr_insee	detects France INSEE	no
Func_fr_passport	detects France passport	no
Func_france_eu_tax_file_number	detects France tax file number	no
Func_france_value_added_tax_number	detects France value added tax number	no
Func_french_drivers_license	detects French driver's license	no
Func_french_insee	detects French INSEE	no
Func_georgia_drivers_license_number	detects Georgia driver's license number	no
Func_german_drivers_license	detects Germany driver's license	no

Function name	Function action	Is a validator
Func_german_passport_data	detects Germany passport	no
Func_german_passport	detects Germany passport	no
Func_germany_eu_tax_file_number	detects Germany tax file number	no
Func_germany_value_added_tax_number	detects Germany value added tax number	no
Func_greece_eu_ssn	detects Greece sin (AMKA)	no
Func_hawaii_drivers_license_number	detects Hawaii driver's license number	no
Func_hong_kong_id_card	detects Hong Kong ID card	no
Func_hungarian_value_added_tax_number	detects Hungary value added tax number	no
Func_hungary_eu_national_id_card	detects Hungary personal identification number	no
Func_hungary_eu_ssn_or_equivalent	detects Hungary social security number	no
Func_hungary_eu_tax_file_number	detects Hungary tax file number	no
Func_iban	detects IBAN	yes

Function name	Function action	Is a validator
Func_idaho_drivers_license_number	detects Idaho driver's license number	no
Func_illinois_drivers_license_number	detects Illinois driver's license number	no
Func_india_aadhaar	detects India aadhaar	yes
Func_indiana_drivers_license_number	detects Indiana driver's license number	no
Func_iowa_drivers_license_number	detects Iowa driver's license number	no
Func_ireland_pps	detects Ireland PPS	no
Func_israeli_national_id_number	detects Israel national ID number	no
Func_italy_eu_national_id_card	detects Italy fiscal code	no
Func_italy_value_added_tax_number	detects Italy value added tax number	no
Func_japanese_my_number_corporate	detects Japan my number corporate	yes
Func_japanese_my_number_personal	detects Japan my number personal	yes
Func_jp_bank_account_branch_code	detects Japan bank account branch code	no
Func_jp_bank_account	detects Japan bank account	no

Function name	Function action	Is a validator
Func_jp_drivers_license_number	detects Japan driver's license number	no
Func_jp_passport	detects Japan passport	no
Func_jp_resident_registration_number	detects Japan-resident registration number	no
Func_jp_sin_pre_1997	detects Japan sin pre 1997	no
Func_jp_sin	detects Japan SIN	no
Func_kansas_drivers_license_number	detects Kansas driver's license number	no
Func_kentucky_drivers_license_number	detects Kentucky driver's license number	no
Func_kentucky_massachusetts_virginia_drivers_license_number	detects Kentucky, Massachusetts, Virginia driver's license number	no
Func_latvia_eu_national_id_card	detects Latvia personal code	no
Func_lithuania_eu_tax_file_number	detects Lithuania personal code	no
Func_louisiana_drivers_license_number	detects Louisiana driver's license number	no
Func_luxemburg_eu_tax_file_number_non_natural	detects Luxemburg national	no

Function name	Function action	Is a validator
	identification number (non-natural persons)	
Func_luxemburg_eu_tax_file_number	detects Luxemburg national identification number (natural persons)	no
Func_maine_drivers_license_number	detects Maine driver's license number	no
Func_manitoba_drivers_license_number	detects Manitoba driver's license number	no
Func_maryland_drivers_license_number	detects Maryland driver's license number	no
Func_massachusetts_drivers_license_number	detects Massachusetts driver's license number	no
Func_mexico_population_registry_code	detects Mexico population registry code	no
Func_michigan_minnesota_drivers_license_number	detects Michigan, Minnesota driver's license number	no
Func_minnesota_drivers_license_number	detects Minnesota driver's license number	no

Function name	Function action	Is a validator
Func_mississippi_oklahoma_drivers_license_number	detects Mississippi, Oklahoma driver's license number	no
Func_missouri_drivers_license_number	detects Missouri driver's license number	no
Func_montana_drivers_license_number	detects Montana driver's license number	no
Func_nebraska_drivers_license_number	detects Nebraska driver's license number	no
Func_netherlands_bsn	detects Netherlands BSN	no
Func_netherlands_eu_tax_file_number	detects Netherlands tax file number	no
Func_netherlands_value_added_tax_number	detects Netherlands value added tax number	no
Func_nevada_drivers_license_number	detects Nevada driver's license number	no
Func_new_brunswick_drivers_license_number	detects New Brunswick driver's license number	no
Func_new_hampshire_drivers_license_number	detects New Hampshire driver's license number	no

Function name	Function action	Is a validator
Func_new_jersey_drivers_license_number	detects New Jersey driver's license number	no
Func_new_mexico_drivers_license_number	detects New Mexico driver's license number	no
Func_new_york_drivers_license_number	detects New York driver's license number	no
Func_new_zealand_bank_account_number	detects New Zealand bank account number	no
Func_new_zealand_inland_revenue_number	detects New Zealand inland revenue number	no
Func_new_zealand_ministry_of_health_number	detects New Zealand ministry of health number	no
Func_newfoundland_labrador_drivers_license_number	detects Newfoundland Labrador driver's license number	no
Func_newzealand_driver_license_number	detects New Zealand driver license number	no
Func_newzealand_social_welfare_number	detects New Zealand social welfare number	no
Func_north_carolina_drivers_license_number	detects North Carolina driver's license number	no

Function name	Function action	Is a validator
Func_north_dakota_drivers_license_number	detects North Dakota driver's license number	no
Func_norway_id_number	detects Norway ID number	no
Func_nova_scotia_drivers_license_number	detects Nova Scotia driver's license number	no
Func_ohio_drivers_license_number	detects Ohio driver's license number	no
Func_ontario_drivers_license_number	detects Ontario driver's license number	no
Func_pennsylvania_drivers_license_number	detects Pennsylvania driver's license number	no
Func_pesel_identification_number	detects Poland National ID (PESEL)	no
Func_poland_eu_tax_file_number	detects Poland tax file number	no
Func_polish_national_id	detects Poland identity card	no
Func_polish_passport_number	detects Polish passport number	no
Func_polish_regon_number	detects Polish REGON number	no
Func_portugal_eu_tax_file_number	detects Portugal Tax Identification Number	no

Function name	Function action	Is a validator
Func_prince_edward_island_drivers_license_number	detects Prince Edward Island driver's license number	no
Func_quebec_drivers_license_number	detects Quebec driver's license number	no
Func_randomized_formatted_ssn	detects randomized formatted US SSN	yes
Func_randomized_unformatted_ssn	detects randomized unformatted US SSN	yes
Func_rhode_island_drivers_license_number	detects Rhode Island driver's license number	no
Func_romania_eu_national_id_card	detects Romania personal numeric code (CNP)	no
Func_saskatchewan_drivers_license_number	detects Saskatchewan driver's license number	no
Func_slovakia_eu_national_id_card	detects Slovakia personal number	no
Func_slovenia_eu_national_id_card	detects Slovenia Unique Master Citizen Number	no
Func_slovenia_eu_tax_file_number	detects Slovenia tax	no

Function name	Function action	Is a validator
	file number	
Func_south_africa_identification_number	detects South Africa identification number	yes
Func_south_carolina_drivers_license_number	detects South Carolina driver's license number	no
Func_south_dakota_drivers_license_number	detects South Dakota driver's license number	no
Func_south_korea_resident_number	detects South Korea resident number	no
Func_spain_eu_DL_and_NI_number_citizen	detects Spain DL and NI number citizen	no
Func_spain_eu_DL_and_NI_number_foreigner	detects Spain DL and NI number foreigner	no
Func_spain_eu_driver's_license_number	detects Spain driver's license number	no
Func_spain_eu_tax_file_number	detects Spain tax file number	no
Func_spanish_social_security_number	detects Spanish social security number	no
Func_ssn	Function to detect non-randomized formatted US SSN	yes
Func_sweden_eu_tax_file_number	detects Sweden tax file	no

Function name	Function action	Is a validator
	number	
Func_swedish_national_identifier	detects Swedish national identifier	yes
Func_swiss_social_security_number_ahv	detects Swiss social security number AHV	no
Func_taiwanese_national_id	detects Taiwanese national ID	no
Func_tennessee_drivers_license_number	detects Tennessee driver's license number	no
Func_texas_drivers_license_number	detects Texas driver's license number	no
Func_Thai_Citizen_Id	detects Thai Citizen ID	no
Func_Turkish_National_Id	detects Turkish National ID	yes
Func_uk_drivers_license	detects UK driver's license	no
Func_uk_eu_tax_file_number	detects UK unique taxpayer number	no
Func_uk_nhs_number	detects UK NHS number	yes
Func_uk_nino	detects UK NINO	no
Func_unformatted_canadian_sin	detects unformatted Canadian SIN	no

Function name	Function action	Is a validator
Func_unformatted_itin	detects unformatted US ITIN	yes
Func_unformatted_ssn	detects non-randomized unformatted US SSN	yes
Func_usa_uk_passport	detects USA and UK passport	yes
Func_utah_drivers_license_number	detects Utah driver's license number	no
Func_vermont_drivers_license_number	detects Vermont driver's license number	no
Func_virginia_drivers_license_number	detects Virginia driver's license number	no
Func_washington_drivers_license_number	detects Washington driver's license number	no
Func_west_virginia_drivers_license_number	detects West Virginia driver's license number	no
Func_wisconsin_drivers_license_number	detects Wisconsin driver's license number	no
Func_wyoming_drivers_license_number	detects Wyoming driver's license number	no

Func_us_date

Func_us_date looks for dates in common U.S. formats. The common formats are "month/day/year", "month-day-year", and "month day year ". The names or abbreviations of months aren't case-sensitive.

Examples:

- December 2, 2016
- Dec 2, 2016
- dec 02 2016
- 12/2/2016
- 12/02/16
- Dec-2-2016
- 12-2-16

Accepted month names:

- English
  - January, February, march, April, may, June, July, August, September, October, November, December
  - Jan. Feb. Mar. Apr. May June July Aug. Sept. Oct. Nov. Dec.

Func_eu_date

Fund_eu_dates looks for dates in common E.U. formats (and most places outside the U.S.), such as "day/month/year", "day-month-year", and "day month year". The names or abbreviations of months aren't case-sensitive.

Examples:

- 2 Dec 2016
- 02 dec 2016
- 2 Dec 16
- 2/12/2016
- 02/12/16
- 2-Dec-2016
- 2-12-16

Accepted month names:

- English
  - January, February, march, April, may, June, July, August, September, October, November, December
  - Jan. Feb. Mar. Apr. May June July Aug. Sept. Oct. Nov. Dec.

- Dutch
  - januari, februari, maart, April, mei, juni, juli, augustus, September, ocktober, October, November, December
  - jan feb maart apr mei jun jul aug sep sept oct okt nov dec
- French
  - janvier, février, mars, avril, mai, juin juillet, août, septembre, octobre, novembre, décembre
  - janv. févr. mars avril mai juin juil. août sept. oct. nov. déc.
- German
  - jänuar, februar, märz, April, mai, juni juli, August, September, oktober, November, dezember
  - Jan./Jän. Feb. März Apr. Mai Juni Juli Aug. Sept. Okt. Nov. Dez.
- Italian
  - gennaio, febbraio, marzo, aprile, maggio, giugno, luglio, agosto, settembre, ottobre, novembre, dicembre
  - genn. febbr. mar. apr. magg. giugno luglio ag. sett. ott. nov. dic.
- Portuguese
  - janeiro, fevereiro, março, marco, abril, maio, junho, julho, agosto, setembro, outubro, novembro, dezembro
  - jan fev mar abr mai jun jul ago set out nov dez
- Spanish
  - enero, febrero, marzo, abril, mayo, junio, julio, agosto, septiembre, octubre, noviembre, diciembre
  - enero feb. marzo abr. mayo jun. jul. agosto sept./set. oct. nov. dic.

Func_eu_date1 (deprecated)

> ⓘ **Note**
>
> This function is deprecated because it supports only Portuguese month names, which are now included in the `Func_eu_date` function above.

This function looks for a date in the format commonly used in Portuguese. The format for this function is the same as `Func_eu_date`, differing only in the language used.

Examples:

- 2 Dez 2016
- 02 dez 2016
- 2 Dez 16

- 2/12/2016
- 02/12/16
- 2-Dez-2016
- 2-12-16

Accepted month names:

- Portuguese
  - janeiro, fevereiro, março, marco, abril, maio, junho, julho, agosto, setembro, outubro, novembro, dezembro
  - jan fev mar abr mai jun jul ago set out nov dez

Func_eu_date2 (deprecated)

> ⓘ **Note**
>
> This function is deprecated because it supports only Dutch month names, which are now included in the `Func_eu_date` function above.

This function looks for a date in the format commonly used in Dutch. The format for this function is the same as `Func_eu_date`, differing only in the language used.

Examples:

- 2 Mei 2016
- 02 mei 2016
- 2 Mei 16
- 2/12/2016
- 02/12/16
- 2-Mei-2016
- 2-12-16

Accepted month names:

- Dutch
  - januari, februari, maart, April, mei, juni, juli, augustus, September, ocktober, October, November, December
  - jan feb maart apr mei jun jul aug sep sept out okt nov dec

Func_expiration_date

Func_expiration_date looks for dates that are in formats commonly used by credit and debit cards. This function will match dates in format of "month/year", "month-year", "[month name] year", and "[month abbreviation] year". The names or abbreviations of months aren't case-sensitive.

Examples:

- MM/YY -- for example, 01/11 or 1/11
- MM/YYYY -- for example, 01/2011 or 1/2011
- MM-YY -- for example, 01-22 or 1-11
- MM-YYYY -- for example, 01-2000 or 1-2000

The following formats support YY or YYYY:

- Month-YYYY -- for example Jan-2010 or january-2010 or Jan-10 or january-10
- Month YYYY -- for example, 'january 2010' or 'Jan 2010' or 'january 10' or 'Jan 10'
- MonthYYYY -- for example, 'january2010' or 'Jan2010' or 'january10' or 'Jan10'
- Month/YYYY -- for example, 'january/2010' or 'Jan/2010' or 'january/10' or 'Jan/10'

Accepted month names:

- English
  - January, February, march, April, may, June, July, August, September, October, November, December
  - Jan Feb Mar Apr May June July Aug Sept Oct Nov Dec

Func_us_address

Func_us_address looks for a U.S. state name or postal abbreviation followed by a valid zip code. The zip code must be one of the correct zip codes associated with the U.S. state name or abbreviation. The U.S. state name and zip code can't be separated by punctuation or letters.

Examples:

- Washington 98052
- Washington 98052-9998
- WA 98052
- WA 98052-9998

Feedback

Was this page helpful? Yes No

Custom sensitive information type filters reference

09/26/2025

In Microsoft Purview, you can define filters or other checks while creating a custom sensitive information type (SIT).

List of supported filters and use cases

AllDigitsSame Exclude

Description: Allows you to exclude matches that have all digits as duplicate digits, such as 111111111 or 111-111-111

Defining filters:

XML

```xml
<Filters id="ssn_filters">
    <Filter type="AllDigitsSameFilter"></Filter>
</Filters>
```

Using it in rule package at the entity level:

XML

```xml
<Entity id="50842eb7-edc8-4019-85dd-5a5c1f2bb085" patternsProximity="300"
recommendedConfidence="85"  filters="ssn_filters">
      <Pattern confidenceLevel="85">
        <IdMatch idRef="Func_ssn" />
      </Pattern>
</Entity>
```

Using it in rule package at the pattern level:

XML

```xml
<Entity id="50842eb7-edc8-4019-85dd-5a5c1f2bb085" patternsProximity="300"
recommendedConfidence="85">
      <Pattern confidenceLevel="85"  filters="ssn_filters">
        <IdMatch idRef="Func_ssn" />
      </Pattern>
</Entity>
```

TextMatchFilter StartsWith

Description: Allows you to define the starting characters for the entity. It has two variants, *exclude* and *include*.

For example, to *exclude* the numbers starting with 0500, 91, 091, 010 in a list like this:

- 0500-4500-027
- 91564721450
- 91-8523697410
- 700-8956-7844
- 1000-3265-9874
- 0100-7892-3012

you can use the following XML:

XML

```
<Filters id="phone_number_filters_exc">
    <Filter type="TextMatchFilter" direction="StartsWith" logic="Exclude"
textProcessorId="Keyword_false_positives_sw">
</Filter>
</Filters>

  <Keyword id="Keyword_false_positives_sw">
    <Group matchStyle="string">
      <Term>0500</Term>
      <Term>91</Term>
      <Term>091</Term>
      <Term>0100</Term>
    </Group>
  </Keyword>
```

Similarly, to *include* the numbers starting with 0500, 91, 091, 0100 in a list like this:

- 0500-4500-027
- 91564721450
- 91-8523697410
- 700-8956-7844
- 1000-3265-9874
- 0100-7892-3012

you can use the following XML:

XML

```
<Filters id="phone_filters_inc">
    <Filter type="TextMatchFilter" direction="StartsWith" logic="Include"
textProcessorId="Keyword_false_positives_sw">
</Filter>
```

TextMatchFilter EndsWith

Description: Allows you to define the ending characters for the entity.

For example, to *exclude* the numbers ending with 0500,91,091, 0100 in a list like this:

- 1234567891
- 1234-5678-0091
- 1234.4567.7091
- 1234-8091-4564

you can use the following XML:

XML

```
<Filters id="phone_number_filters_exc">
    <Filter type="TextMatchFilter" direction="EndsWith" logic="Exclude"
textProcessorId="Keyword_false_positives_sw">
</Filter>

  <Keyword id="Keyword_false_positives_sw">
    <Group matchStyle="string">
      <Term>0500</Term>
      <Term>91</Term>
      <Term>091</Term>
      <Term>0100</Term>
    </Group>
  </Keyword>
```

For example, to *include* the numbers ending with 0500, 91, 091, 0100, in a list like this:

- 1234567891
- 1234-5678-0091
- 1234.4567.7091
- 1234-8091-4564

You can use the following XML:

XML

```
<Filters id="phone_filters_inc">
    <Filter type="TextMatchFilter" direction=" EndsWith" logic="Include"
textProcessorId="Keyword_false_positives_sw">
</Filter>
```

TextMatchFilter Full

Description: Allows you to prohibit certain matches to prevent them from triggering the rule, such as excluding 4111111111111111 from the list of valid credit card matches.

For example, to *exclude* credit card numbers like 4111111111111111 and 3241891031113111 in a list like this:

- 4485 3647 3952 7352
- 4111111111111111
- 3241891031113111

you can use the following XML:

XML

```
<Filters id="cc_number_filters_exc">
    <Filter type="TextMatchFilter" direction="Full" logic="Exclude"
textProcessorId="Keyword_false_positives_full">
</Filter>

  <Keyword id="Keyword_false_positives_full">
    <Group matchStyle="string">
      <Term>4111111111111111</Term>
      <Term>3241891031113111</Term>
    </Group>
  </Keyword>
```

Likewise, to *include* credit card numbers like 4111111111111111 and 3241891031113111 in a list like this:

- 4485 3647 3952 7352
- 4111111111111111
- 3241891031113111

you can use the following XML:

XML

```
<Filters id="cc_filters_inc">
    <Filter type="TextMatchFilter" direction="Full" logic="Include"
```

```
textProcessorId="Keyword_false_positives_full">
</Filter>
```

TextMatchFilter Prefix

Description: Allows you to define the preceding characters that should be always excluded or included. For example, if **Credit card number** is preceded by 'Order ID:', then remove the match from the valid matches.

For example, to exclude occurrences of phone numbers that have **Phone number** and **call me at** strings before the phone number, in a list like this:

- Phone number 091-8974-653278
- Phone 45-124576532-123
- 45-124576532-123

you can use the following XML:

XML

```
<Filters id="cc_number_filters_exc">
    <Filter type="TextMatchFilter" direction="Prefix" logic="Exclude"
textProcessorId="Keyword_false_positives_prefix">
</Filter>
  <Keyword id="Keyword_false_positives_prefix">
    <Group matchStyle="string">
      <Term>phone number</Term>
      <Term>call me at</Term>
    </Group>
  </Keyword>
```

Similarly, to include occurrences that have **credit card** and **card #** strings before the credit card number, in a list like this:

- Credit card 45-124576532-123
- 45-124576532-123 (which could be phone number)

you can use the following XML:

XML

```
<Filters id="cc_filters_inc">
    <Filter type="TextMatchFilter" direction="Full" logic="Include"
textProcessorId="Keyword_true_positives_prefix">
</Filter>

  <Keyword id="Keyword_true_positives_prefix">
```

```
    <Group matchStyle="string">
      <Term>credit card</Term>
      <Term>card #</Term>
    </Group>
  </Keyword
```

TextMatchFilter Suffix

Description: Allows you to define the following characters that should be always excluded or included. For example, if Credit card number is followed by '/xuid' then remove the match from the valid matches.

For example, top exclude occurrences if there are five more instances of four digits as suffix in a list like this:

- 1234-5678-9321 4500 9870 6321 48925566
- 1234-5678-9321

you can use the following XML:

XML

```
<Filters id="cc_number_filters_exc">
    <Filter type="TextMatchFilter" direction="Prefix" logic="Exclude"
textProcessorId="Regex_false_positives_suffix">
</Filter>

  <Regexid="Regex_false_positives_suffix">(\d{4}){5,}</Regex>
```

You can also exclude occurrences if they are followed by **/xuidsuffix**, like the one in this list:

- 1234-5678-9321 /xuid
- 1234-5678-9321

you can use this XML:

XML

```
<Filters id="cc_number_filters_exc">
    <Filter type="TextMatchFilter" direction="Prefix" logic="Exclude"
textProcessorId="Keyword_false_positives_suffix">
</Filter>

  <Keyword id="Keyword_false_positives_suffix">
    <Group matchStyle="string">
      <Term>/xuid</Term>
```

```
    </Group>
  </Keyword>
```

Similarly, to include an occurrence only if it is followed by **cvv** or **expires**, such as the two in this list:

- 45-124576532-123
- 45-124576532-123 cvv 966
- 45-124576532-123 expires 03/23

you can use this XML:

XML

```
<Filters id="cc_filters_inc">
    <Filter type="TextMatchFilter" direction="Full" logic="Include"
textProcessorId="Keyword_true_positives_suffix">
</Filter>

  <Keyword id="Keyword_true_positives_suffix">
    <Group matchStyle="string">
      <Term>cvv</Term>
      <Term>expires</Term>
    </Group>
  </Keyword>
```

Using filters in rule packages

Filters can be defined on the entire SIT or on a pattern. Here are some examples.

At the sensitive information type level

Filters at Entity - will cover all child patterns

The filters will be applied to **all** the instances classified by any of the patterns in that entity / sensitive information type.

XML

```
<Entity id="6443b88f-2808-482a-8e1a-3ae5026645e1" patternsProximity="300"
recommendedConfidence="85" filters="CompositeFiltersAtEntityLevel">
      <Pattern confidenceLevel="85">
        <IdMatch idRef="Regex_denmark_id" />
      </Pattern>
</Entity>
```

At the individual pattern of the sensitive information type level

Filters only at the pattern level.

The filter will be applied to only the instances matched by the pattern.

XML

```
<Entity id="50842eb7-edc8-4019-85dd-5a5c1f2bb085" patternsProximity="300"
recommendedConfidence="85">
      <Pattern confidenceLevel="85"  filters="CompositeFiltersAtPattern">
        <IdMatch idRef="Keyword_cc_verification" />
      </Pattern>
</Entity>
```

At the sensitive information type level with an additional filter on some of the patterns of that entity

Filters at Entity + pattern

The filters will be applied to **all** the instances classified by any of the patterns in that entity / sensitive information type. The pattern level filter will filter the instances matched by that pattern.

XML

```
<Entity id="6443b88f-2808-482a-8e1a-3ae5026645e1" patternsProximity="300"
recommendedConfidence="85" filters="CompositeFiltersAtEntityLevel">
      <Pattern confidenceLevel="85" filters="CompositeFiltersAtPattern">
        <IdMatch idRef="Regex_denmark_id" />
      </Pattern>
</Entity>
```

More information

- Learn about Microsoft Purview Data Loss Prevention
- Sensitive information type entity definitions
- Sensitive information type functions

Sensitive information type limits

These limits apply to all Microsoft Purview policies that use SITs.

To ensure high performance and lower latency, there are limitations in custom SIT configurations, as described in the following table.

Expand table

Limit	Value
Maximum number of custom SITs created through the Microsoft Purview portal	500
Maximum length of regular expression	1024 characters
Maximum length for a given term in a keyword list	50 characters
Maximum number of terms in keyword list	2048
Maximum number of distinct regexes per SIT	20
Maximum combined size for all keyword dictionaries (post compression)	1MB (~1,000,000 characters)
Maximum number of rule packages	10
Maximum number of SITs per rule package	50
Maximum size of a rule package	150KB
Capturing group in a regular expression	1

> **Note**
>
> The regex pattern must use exactly one primary capturing group. Use noncapturing groups for any prefix/suffix (for example: `(?:prefix)(primary capturing group)(?:suffix)`). Multiple top-level capturing groups separated by `|` outside the primary group aren't supported; put alternation variants inside the single capturing group. For more information, see Create a custom sensitive information type.

> **Note**

If you have a business need to create more than 500 custom SITs, please raise a support ticket.

Instance count supported values for SIT

The SIT instance count limit applies when using SITs in these solutions:

- Microsoft Purview Data Loss Prevention policies
- Microsoft Purview Information Protection policies
- Microsoft Purview Data Lifecycle Management
- Communication Compliance
- Microsoft Purview Records Management
- Microsoft Defender for Cloud Apps
- Microsoft Priva

For a scanned item to satisfy the rule criteria, the number of unique instances of a SIT in any single item must fall between the min and max values. This is called the **Instance count**.

- **Min** field: the lower limit (minimum number) of unique instances of a SIT that must be found in an item to trigger a match. The min field supports values of 1 to 500.
- **Max** field: the upper limit on the number of unique instances of a SIT that can be found in an item and still trigger a match. The max field supports the following values:
  - 1 to 500 - Use a value between `1` and `500` when you want to set a specific upper limit of 500 or less on the number of instances of a SIT in an item.
  - Any - Use `Any` when you want to ensure that the unique instance count criterion is satisfied when an undefined number of unique instances of a SIT are found in a scanned item and that the specified number of unique instances meets or exceeds the minimum number of unique instances value. In other words, the unique instance count criteria are met as long as the min value is met.

For example, if you want the rule to trigger a match when at least 500 unique instances of a SIT are found in a single item, set the **min** value to `500` and the **max** value to `Any`.

Last updated on 06/26/2026

Learn about named entities

09/26/2025

Named entities are sensitive information types (SIT). They're complex dictionary and pattern-based classifiers that you can use to detect person names, physical addresses, and medical terms and conditions. You can see them in the **Microsoft Purview portal** > **Data classification** > **Sensitive info types**. Here's a partial list of where you can use SITs:

- Microsoft Purview Data Loss Prevention policies (DLP)
- Sensitivity labels
- Insider risk management
- Microsoft Defender for Cloud Apps
- Microsoft Purview Information Protection
- Data Lifecycle Management
- Records management
- Microsoft Purview eDiscovery
- Microsoft Priva
- Exact data match sensitive information types

DLP makes special use of named entities in *enhanced policy templates*, which are preconfigured DLP policies that you can customize for your organizations needs. You can also create your own DLP policies from a blank template and use a named entity SIT as a condition. Entity Match and other advanced classifiers are offered only to users with E5 licenses.

Examples of named entity SITs

Named entity SITs come in two flavors, *bundled* and *unbundled*

Bundled named entity SITs detect all possible matches. Use them as broad criteria in your DLP policies for detecting sensitive items.

Unbundled named entity SITs have a narrower focus, like a single country. Use them when you need a DLP policy with a narrower detection scope.

> **Note**
>
> To use named entity SITs, you must activate Advanced classification scanning and protection for the relevant data loss prevention settings before they will be discoverable.

Here are some examples of named entity SITs. You can find all of them in Sensitive information type entity definitions.

Expand table

Named Entity	Description	Bundled/Unbundled
All full names	Detects all possible matches of full names	bundled
All physical addresses	Detects all possible matches of physical addresses	bundled
All medical terms and conditions	Detects all possible matches of medical terms and conditions	bundled
Australia Physical Addresses	Detects patterns related to physical addresses from Australia. Included in All physical addresses SIT.	unbundled
Blood Test Terms	Detects terms related to blood tests, such as 'hCG'. English terms only. Included in All medical terms and conditions SIT	unbundled
Brand Medication Names	Detects names of brand medication, such as 'Tylenol'. English terms only. Included in All medical terms and conditions.	unbundled

Examples of enhanced DLP policies

Here are some examples of enhanced DLP policies that use named entity SITs. You can find all 10 of them in the **Microsoft Purview portal** Navigate to **Data loss prevention** > **Create policy**. Enhanced templates can be used in DLP and auto-labeling.

Expand table

Policy category	Template	Description
Financial	U.S. Gramm-Leach-Bliley Act (GLBA) Enhanced	Helps detect the presence of information subject to Gramm-Leach-Bliley Act (GLBA), including information like social security numbers or credit card numbers. This enhanced template extends the original by also detecting people's full names, U.S./U.K. passport number, U.S. driver's license number and U.S. physical addresses.
Medical and health	Australia Health Records Act (HRIP Act) Enhanced	Helps detect the presence of information commonly considered to be subject to the Health Records and Information Privacy (HRIP) act in Australia, like medical account number and tax file number. This enhanced template extends the original by also detecting people's full names, medical terms and conditions, and Australia physical addresses.
Privacy	General Data Protection	Helps detect the presence of personal information for individuals inside the European Union (EU) to help meet GDPR privacy obligations.

Policy category	Template	Description
	Regulation (GDPR) Enhanced	This enhanced template detects people's full names and physical addresses for countries in the EU.

Next steps

- Use named entities in your data loss prevention policies

For further information

- Sensitive information type entity definitions
- Learn about Sensitive information types
- Create a custom sensitive information type
- Create a custom sensitive information type in PowerShell
- Sensitivity labels
- Retention labels
- Communication compliance
- Autolabeling policies
- Create and Deploy data loss prevention policies

Use named entities in your data loss prevention policies

Read through Named entities in sensitive information types before you start to use named entity SITs.

Before you begin

Licensing

For information on licensing, see

- Microsoft 365 Enterprise Plans
- Microsoft 365 Service Descriptions

Permissions

The account you use to create and edit data loss prevention (DLP) policies, must have the **DLP Compliance Management** role permissions. For more information, see Add users or groups to a Microsoft Purview built-in role group.

Supported locations

You can use named entity SITs and enhanced policies to detect and protect sensitive items in these locations:

- SharePoint sites
- OneDrive accounts
- Teams chat and channel messages
- Devices (Windows 10/11 endpoint devices)
- Exchange mailboxes
- Microsoft Defender for Cloud Apps

Named entity SITs and enhanced policies aren't supported for:

- On-premises repositories
- Power BI

Create and edit enhanced policies

To create or edit a DLP policy, use the procedures in Create and Deploy data loss prevention policies.

Workloads and services that support named entities

The following workloads and services support named entities:

- **Microsoft 365 eDiscovery** supports the use of named entities in Substrate services.
- **Microsoft Defender for Cloud Apps** supports the use of named entities in Defender for Cloud Apps policies in the Defender for Cloud Apps portal.
- **Insider Risk Management** supports the use of named entities in Substrate services.
- **Records Management** supports the use of named entities.
- **Exact Data Match Sensitive Information Types** supports the use of named entities.

Unified DLP support for named entities

The following table shows named entity support in Unified DLP workloads and services.

Expand table

Workload/Services	Support for Named Entities
Office Win32 clients policy tip	Not supported
Office WAC clients policy tip	Supported
OWA policy tip	Not supported
Outlook for Microsoft 365 policy tip	Supported
Endpoints (Windows 10, and 11 devices)	Supported
Exchange Transport rules	Supported
OneDrive for Business data-at-rest	Supported
SharePoint Online data-at-rest	Supported
Teams data-at-rest	Supported
Email messages data-at-rest	Supported for tenants with Privacy Service Plan
Microsoft Defender for Cloud Apps	Supported

Autolabeling support for named entities

The following table lists named entity support for autolabeling workloads and services.

Expand table

Workload/Services	Support for Named Entities
Office Win32 clients offline	Supported, user must select label and manually apply
Online Office Win32 clients online	Supported with old confidence scheme
Outlook online	Supported with old confidence scheme
Office WAC client	Supported
OWA	Supported
Exchange transport	Supported
OneDrive for Business data-at-rest	Supported
SharePoint Online data-at-rest	Supported
Microsoft Purview Information Protection scanner	Not supported

Known issues

The following table summarizes current known issues when using named entities and their impact.

Expand table

Issue	Impact
DLP Policy tips (OWA, Outlook, Office Win32 clients)	Policy tips with entity condition results in "no match"
Asian language support for person name (Chinese, Japanese, Korean)	Named entities supported for Latin-based character set only (that is, kanji isn't supported) for person name
On-premises repositories	Not supported as a workload
Power BI (preview)	Not supported

Best practices for using named entity SITs

Here are some practices you can use when you create or edit a policy that uses a named entity SIT.

- Use low instance counts (three to five) when you're looking for data that's in a spreadsheet and the keyword that's required by the SIT for that data is only in the column header. For example, let's say you're looking for US Social Security numbers, and the keyword `Social Security Number` only occurs in the column header. Since the values (the corroborative evidence) are in the cells below, it's likely that only the first few instances would be in close enough proximity to the keyword to be detected.

- If you're using a named entity SIT, like All Full Names, to help find US Social Security numbers, use larger instance counts such as 10 or 50. Then, when both the person names and the SSNs are detected together, you're more likely to get true positives.

- You can use Autolabeling simulations to test the accuracy of named entity SITs. Run a simulation using a named entity SIT to see what items match the policy. With this information, you can fine tune accuracy by adjusting the instance counts and confidence levels in your custom policies or the enhanced template conditions. You can iterate simulations until the accuracy is where you want it before deploying a DLP or autolabeling policy containing named entities in production. Here's an overview of the autolabeling simulation workflow:

1. Identify the SIT or combination of SITs you want to test in simulation mode, either custom or cloned and edited
2. Identify or create a sensitivity label to be applied when the autolabeling policy finds a match in Exchange, SharePoint sites, or OneDrive accounts
3. Create a sensitivity autolabeling policy that uses the SIT from step 1 and with same Conditions and Exceptions that are used in your DLP policy
4. Run the policy simulation
5. View the results
6. Tune the SIT or policy and the instance count and confidence levels to reduce false positives.
7. Repeat until you get the accuracy results you want

Related content

- Sensitive information type entity definitions
- Named entities overview.
- Create and Deploy data loss prevention policies

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Learn about exact data match based sensitive information types

Article • 01/21/2025

Sensitive information types (SITs) are used to help identify sensitive data so that you can prevent it from being inadvertently or inappropriately shared. They're also used to help in locating relevant data in eDiscovery, and to apply governance actions to certain types of information. You define a custom SIT based on:

- patterns
- keyword evidence such as *employee*, *social security number*, or *ID*
- character proximity to evidence in a particular pattern
- confidence levels

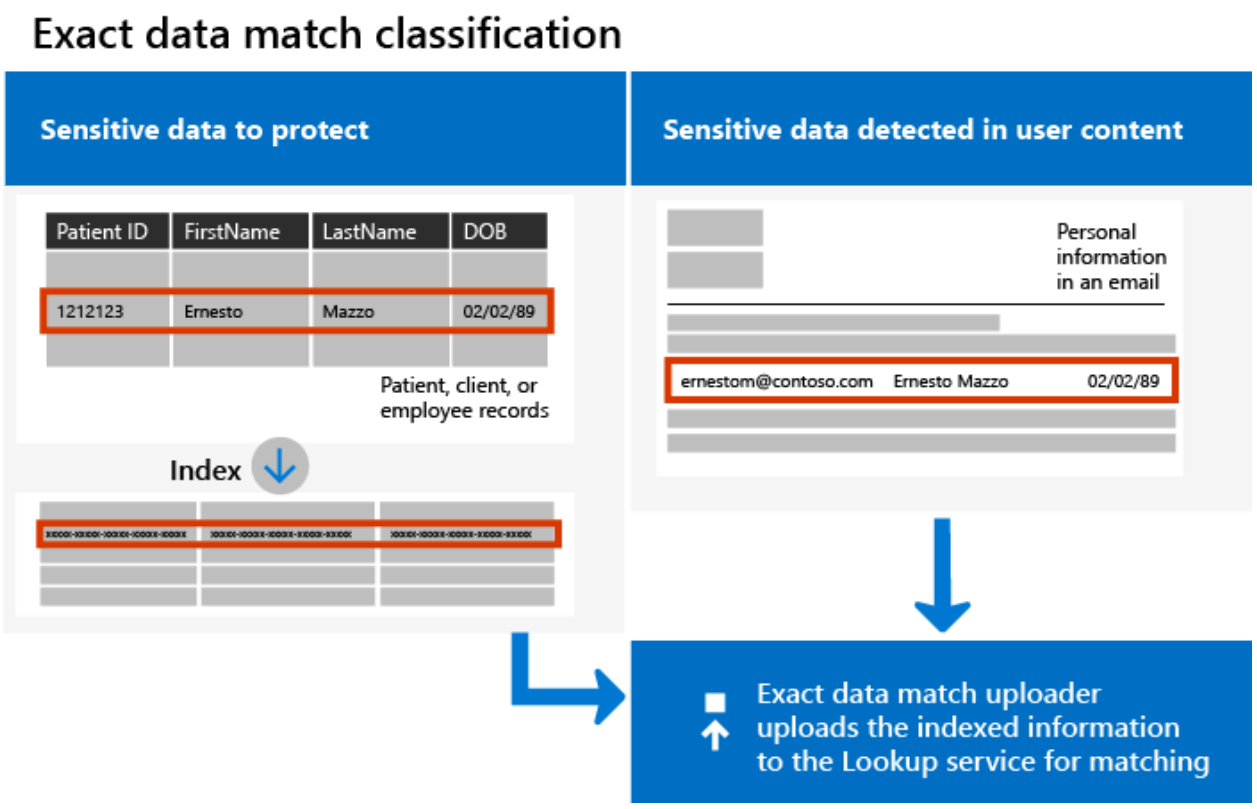
But what if you want a custom SIT that uses exact or nearly exact data values, instead of one that finds matches based on generic patterns? With Exact Data Match (EDM) based classification, you can create a custom sensitive information type that is designed to:

- be dynamic and easily refreshed
- result in fewer false positives
- work with structured sensitive data
- handle sensitive information more securely, not sharing it with anyone, including Microsoft
- be used with several Microsoft cloud services

> **Tip**
>
> If you're not an E5 customer, use the 90-day Microsoft Purview solutions trial to explore how additional Purview capabilities can help your organization manage data security and compliance needs. Start now at the **Microsoft Purview trials hub**. Learn details about **signing up and trial terms**.

EDM-based classification enables you to create custom SITs that refer to exact values in a database of sensitive information. The database can be refreshed daily and can contain up to 100 million rows of data. So, as employees, patients, and clients come and go, and as records change, your custom sensitive information types remain current and applicable. And, you can use EDM-based classification with policies, such as Microsoft Purview Data Loss Prevention policies or Microsoft Cloud App Security file policies.

The following diagram shows the fundamental workings of EDM classification:



> ⓘ Note
>
> Microsoft Purview Information Protection supports the following languages that use double-byte character sets:
>
> - Chinese (simplified)
> - Chinese (traditional)
> - Korean
> - Japanese
>
> This support is available for sensitive information types. For more information, see Information protection support for double byte character sets: Release Notes (preview).

What's different in an EDM SIT

When you work with EDM SITs, it's helpful to understand a few concepts that are unique to them.

Schema

A *schema* is an XML file. Microsoft Purview uses the schema to determine whether or not your data contains strings that match those that your sensitive information types are

designed to detect.

The schema XML file defines:

- The name of the schema, later referred to as the *DataStore*.
- The field names that your sensitive information source table contains. There's a 1:1 mapping of schema field names to the column names in the sensitive information source table.
- Which corroborative evidence fields require multi-token match mode.
- Which data fields are searchable.
- Whether or not configurable matches are supported for each field. A *configurable match* is one with parameters that modify a search, such as ignoring delimiters and case in searched values.

Sensitive information source table

The sensitive information source table contains the values that the EDM SIT looks for. The table is made up of columns and rows. The column headers are the field names, the rows are instances of items and each cell in a row contains the values for that item instance for that field.

Here's a simple example of a sensitive information source table.

Expand table

First Name	Last Name	Date of Birth
Isaiah	Langer	05-05-1960
Ana	Bowman	11-24-1971
Oscar	Ward	02-12-1998

Rule package

Every sensitive information type has a rule package. You use the rule package in an EDM SIT to define the various components of your EDM SIT. The following table provides a description of each component.

Expand table

Component	Description
Match	Specifies the primary element (data field) to be used in exact lookup. It can be a regular expression with or without a checksum validation, a keyword list, a keyword dictionary, or a function.
Classification	Specifies the sensitive information type match that triggers an EDM lookup.
Supporting elements	Elements that, when found, provide evidence that helps increase the confidence of the match. For example, the occurrence of a last name in close proximity to an actual social security number. A supporting element can be a regular expression with or without a checksum validation, a keyword list, a keyword dictionary, or a single- or multi-token string match.
Confidence level (High, Medium, Low)	Indication of how much supporting evidence is detected in addition to the primary element. The more supporting evidence an item contains, the higher the confidence that a matched item contains the sensitive info you're looking for. For more information about confidence levels, see Fundamental parts of a sensitive information type.
Proximity	The number of characters between primary and supporting element.

You supply your own schema and data

Microsoft Purview comes with many built-in SITs that are predefined. These SITs come with schemas, REGEX patterns, keywords, and confidence levels. However, with EDM SITs, you're responsible for defining the schema, as well as the primary and secondary fields that identify sensitive items. Because the schema and primary and secondary data values are all highly sensitive, you encrypt them via a hash function that includes a randomly generated or self-supplied salt value. Only the hashed values are uploaded to the service, so your sensitive data is never in the open.

Primary and secondary support elements

When you create an EDM SIT, you define a *primary element* field in the rule package. EDM then searches all of your content for the primary element. So that EDM can detect them, primary elements must be discoverable through an existing SIT.

> **Note**
>
> For a complete list of the available SITs., see Sensitive information type entity definitions

You need to find a built-in SIT that detects the sensitive information that you want your EDM SIT to detect. For example, if your EDM SIT schema has *U.S. social security number* as the primary element, when you create your EDM schema, you'd associated it with the U.S. Social Security Number (SSN) SIT. Primary elements must follow a defined pattern in order to be detected.

When the primary element is found in a scanned item, EDM then looks for *secondary* elements (also called *supporting* elements). Unlike primary elements, secondary elements have the option of following a pattern. If secondary elements contain multiple tokens, those elements need to either be associated with a SIT that can detect that content or that can be configured for multi-token matching. In all cases, secondary elements must be within a certain proximity to the primary element in order for a match to be detected.

How matching works

EDM works by comparing strings in your documents and emails against values in the sensitive information source table. It uses this comparison to determine whether the values in the scanned content are present in the table. The determination is done by comparing one-way cryptographic hashes.

> **Tip**
>
> You can use both EDM SITs and the predefined SITs that they are based on, together in DLP rules to improve the detection of sensitive data. Use the EDM SIT with higher confidence levels, and the predefined SIT with lower confidence levels. For example, use an EDM SIT that looks for social security number and other supporting data with strict requirements with high confidence. If configured for high-confidence matches, EDM generates a DLP match when only a few instances are detected. To trigger a DLP match when greater numbers of occurrences are detected, use a built-in SIT, such as the *U.S. Social Security Number*.

How supporting elements work with EDM

As discussed in What's different in an EDM SIT, supporting elements are *elements that, when found, provide evidence that helps increase the confidence of the match*.

With support for EDM SITs, you can look for and detect supporting elements that are composed of multiple fields. Supporting element matches can consist of keyword lists, keyword dictionaries, single alphanumeric strings, or multi-token strings.

Let's look at an example. Presume that you want to detect U.S. Social Security numbers. To increase the match confidence, your supporting elements include `first name`, `last name`, and `date of birth` (DoB). So, your source table looks something like this:

Expand table

SSN	FirstName	LastName	DoB
987-65-4320	Isaiah	Langer	05-05-1960
078-05-1120	Ana	Bowman	11-24-1971
219-09-9999	Oscar	Ward	02-12-1998

When looking for matching supporting elements in a protected file, your EDM SIT checks for each supporting element (both individually and in combination) once the primary element is detected.

For instance, say that the first social security number is detected. The exact data match functionality next looks for combinations of supporting elements across all the columns in your source table:

- Isaiah
- Langer
- 05-05-1960
- Isaiah Langer
- Isaiah 05-05-1960
- Langer 05-05-1960
- Isaiah Langer 05-05-1960

Multi-token matching

Multi-token matching is designed to be used when your corroborative evidence field contains multi-token values but matching such values to a SIT isn't easily accomplished. For instance, when you have an `Address` field containing values like `1 Microsoft Way, Redmond, WA` or `123 Main Street, New York, NY`.

This feature allows EDM to compare the hashes of consecutive words in the content with the hashes of the multi-token fields in your data source. If they're identical, EDM produces a match. This way, EDM can detect multi-token fields such as names, addresses, medical conditions, or any other corroborative evidence fields that might contain more than one word, as long as they're marked as multi-token in your EDM schema.

For example, if you select multi-token matching as the match option, you get two additional benefits:

1. Your policies will detect content that matches multiple fields across the columns in your source table.
2. Your source table can include fields with string values that consist of a preconfigured number of words. The following table shows a sample source table:

Expand table

SSN	Name	Street Address
987-65-4320	*Isaiah Langer*	*1432 Lincoln Road*
078-05-1120	*Ana Bowman*	*8250 First Street*
219-09-9999	*Oscar Ward*	*424 205th Avenue*

With multi-token matching, the **Name** and **Street Address** fields are matched both as independent supporting element strings and in combination as individual fields. So, when matched as multi-token strings as supporting elements for Social Security number 987-65-4320, the matches are:

- Isaiah Langer
- 1432 Lincoln Road

When matched in combination, the match is like this:

- Isaiah Langer + 1432 Lincoln Road

Multi-token matching is also supported for double-byte character sets, which generally don't use spaces to separate words.

Services that EDM supports

Expand table

Service	Locations
Microsoft Purview Data Loss Prevention	- SharePoint - OneDrive - Teams Chat - Exchange Online - Devices

Service	Locations
Microsoft Defender for Cloud Apps	- SharePoint - OneDrive
Auto-labeling (service side)	- SharePoint - OneDrive - Exchange Online
Auto-labeling (client side)	- Word - Excel - PowerPoint - Exchange desktop clients
Customer Managed Key	- SharePoint - OneDrive - Teams Chat - Exchange Online - Word - Excel - PowerPoint - Exchange desktop clients - Devices
eDiscovery	- SharePoint - OneDrive - Teams Chat - Exchange Online - Word - Excel - PowerPoint - Exchange desktop clients
Insider Risk Management	- SharePoint - OneDrive - Teams Chat - Exchange Online - Word - Excel - PowerPoint - Exchange desktop clients

See also

- Get started with exact data match based sensitive information types

Feedback

Was this page helpful? Yes No

Get started with exact data match sensitive information types in Microsoft Purview

Creating an exact data match (EDM) sensitive information type (SIT) and making it available is a multiphase process. You can use the *new experience*, *classic experience*, or PowerShell. This article helps Microsoft Purview administrators compare the new and classic experiences and choose the right one for their needs. Before you begin, review the supported regions and required licenses and permissions.

Applies to:

- Create EDM SIT using the new experience
- Create EDM SIT using the classic experience

EDM SITs can be used in:

- Microsoft Purview Data Loss Prevention
- Auto-labeling (service and client side)
- Microsoft Purview Insider Risk Management
- Microsoft Purview eDiscovery
- Microsoft Defender for Cloud Apps

Before you begin

Familiarize yourself with the concepts and terminology in these articles:

- Learn about sensitive information types
- Learn about exact data match based sensitive information types

Supported regions

The exact data match feature is available in these regions:

- Asia Pacific
- Australia
- Brazil
- Canada

- Europe
- France
- Germany
- India
- Japan
- Korea
- Norway
- Qatar
- South Africa
- Switzerland
- United Arab Emirates
- United Kingdom
- United States
- US Department of Defense (DoD)
- US Government Community Cloud (GCC)
- US Government Community Cloud High (GCC High)

To find out where your tenant is hosting data at rest, see Where your Microsoft 365 customer data is stored and refer to the data center city locations listed there.

Required licenses and permissions

You must be a Global Admin, Compliance Admin, or Exchange Admin to perform the tasks described in this article. To learn more about Purview permissions, see Permissions in the Microsoft Purview portal.

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. This helps improve security for your organization. Global Administrator is a highly privileged role that should only be used in scenarios where a lesser privileged role can't be used.

See the Microsoft 365 security and compliance licensing guidance for complete licensing information.

Portal links for your subscription

The following table lists the correct portal URLs for each subscription environment.

Expand table

Portal	Worldwide / GCC	GCC High	DoD
Office SCC	purview.microsoft.com	scc.office365.us	scc.protection.apps.mil
Microsoft Defender portal	security.microsoft.com	security.microsoft.us	security.apps.mil
Microsoft Purview portal	purview.microsoft.com	purview.microsoft.us	purview.apps.mil

New EDM experience

The new EDM experience combines the functionality of the EDM schema and EDM sensitive information type tools into a single user experience. The new experience offers the following benefits:

- A simplified workflow
- Automated schema and SIT creation
- Additional guardrails to ensure better performance

Simplified workflow

With the new experience, the schema and SIT are created in a single workflow. Creating the schema and SIT together means there are fewer steps, better guidance on mapping primary elements to default SITs, and clearer descriptions of default confidence levels for the rules.

When you need to see the status of an EDM SIT in the creation process, the new experience reports the current creation status in the UI.

- *Data not yet uploaded*
- *Data upload percent*
- *Data upload complete*
- *Indexing complete*
- *Data upload failed*
- *Data indexing failed*

Automated schema and SIT creation

In the new experience you can provide a sample data file that has the same header values and enough rows (10-20) of representative data to the system. The system validates the format and creates the schema based on the headers. You then identify the primary fields in the schema and

the system recommends the SITs that best match the primary fields to associate them with. If you don't want to upload the file, you can enter the same values manually in the UI.

> **Important**
>
> Be sure to use sample data values that aren't sensitive; however, also make sure that the sample values are in the same format as your actual sensitive data. Using nonsensitive data is essential because the sample data file doesn't get encrypted and hashed when you upload it in the same way that your actual sensitive information table does. The data from the sample data file isn't retained nor accessible once the EDM SIT is created.

The system generates the EDM SIT detection rules, one for each primary field. Based on detection of the primary fields, the system creates high and medium confidence rules using all the other fields as corroborative evidence. You can manually add low confidence rules if needed.

Additional guardrails to ensure better performance

The system warns you if it finds a primary field mapped to a SIT that detects a broad range of values, called a *loosely defined SIT*. Mapping a primary field to a loosely defined SIT can cause the system to perform lookups on large numbers of strings that aren't related to the kind of content that you're looking for. Mapping between these types of SITs and primary fields can result in false negatives and decrease performance.

> **Note**
>
> A *loosely defined SIT*, such as a custom SIT that looks for all personal identification numbers, has detection rules that allow for greater variability in the items detected. A *strongly defined SIT*, such as a U.S. Social Security Number, has detection rules that only allow a narrow, well-defined set of items to be detected.

The system will also warn you if the values in the primary field you select occur multiple times in a large number of rows. Repeated primary-field values across many rows can cause large numbers of result sets to be returned and processed, which could cause a time-out. Timeouts can result in missed detections and poor performance.

Choosing the right EDM SIT creation experience for you

You can toggle back and forth between the new and classic experiences, but we recommend using the new experience unless your needs fall into one or more of these four use cases, as described below.

Mapping multiple EDM SITs to the same schema

In EDM, you can create a maximum of 10 schemas. Each time you create an EDM SIT using the new experience, a new schema is created. This results in a 1:1 mapping between EDM schema and EDM SIT. The new experience doesn't support mapping multiple SITs to the same schema.

Creating or managing more than 10 EDM SITs

Because the new experience doesn't support mapping multiple SITs to the same schema, you're limited to creating and managing 10 EDM SITs. In the classic experience, you can map multiple EDM SITs to the same schema and so have more than 10 EDM SITs. Using the new flow, you receive an error if you try to create an eleventh EDM schema and you won't be able to view more than 10 EDM SITs.

Specifying the name of your EDM schema

If you need to specify a name for your EDM SIT schemas, you have to use the classic experience to create and manage them. Because the new experience automatically creates the schema, you don't get the opportunity to give your schema a custom name. The autogenerated name is a concatenation of the EDM SIT name and the word *schema*. For example, if the EDM SIT name is *PatientNumber*, the schema name would be *PatientNumberschema*.

Editing EDM schemas created in the classic experience

All schemas that are created using the classic experience or uploaded as an XML file using PowerShell aren't viewable or manageable in the new experience.

Once you choose your method of creating EDM SITs, select the link for your chosen experience in the Next steps section: new experience or classic experience.

Next steps

Continue with the EDM SIT creation experience you selected:

- Create exact data match sensitive information type new experience

or

- Create exact data match sensitive information type classic experience

See also

For related guidance, see the following articles:

- Learn about exact data match based sensitive information types
- Export source data for exact data match based sensitive information type
- Create exact data match sensitive information type workflow new experience
- Create exact data match sensitive information type workflow classic experience

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 09/18/2026

Create exact data match sensitive information type in the New Experience workflow

Creating and making an exact data match (EDM) based sensitive information type (SIT) available is a multi-phase process. EDM-based sensitive information types can be used in Microsoft Purview data loss prevention policies, eDiscovery, and certain content governance tasks. This article outlines the workflow and links to the procedures for each of the phases

Applies to these EDM scenarios

This article covers the new experience. If you want to create an EDM SIT using the classic experience, see Create EDM SIT classic experience.

Before you begin

Make sure you review:

- Learn about EDM based SITs
- Get started with exact data match based sensitive information types (Overview)

Overview of the EDM workflow (New experience)

The following table summarizes the phases of the EDM workflow and the requirements for each phase.

Expand table

Phase	Requirements
Phase 1: Export source data for exact data match based sensitive information type	*Read* access to the sensitive data
Phase 2:Create the sample file	Knowledge of the data fields (the column headers in your source table) and the format of the data you will be looking for in each column.
Phase 3: Create the EDM SIT	Access to the **Microsoft Purview portal** and permissions to create and edit EDM SITs

Phase	Requirements
Phase 4: Hash and upload the sensitive information source table for exact data match sensitive information types	- Custom security group and user account - To **Hash and upload from one computer**: Local admin access to a computer with direct internet access and to host the EDM Upload Agent - To **Hash and upload from separate computers**: Local admin access to a computer with direct internet access to host the EDM Upload Agent for uploading your data and *Local Admin* access to a secure computer for hosting the EDM Upload Agent and for hashing the sensitive information source table. - *Read* access to the sensitive information source table file.
Phase 5: Test an exact data match sensitive information type	Access to the Microsoft Purview portal.

Related content

- Learn about exact data match based sensitive information types

Next steps

- **For the new experience**: Export source data for exact data match based sensitive information types

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Export source data for exact data match based sensitive information types

This article explains how to prepare and export your sensitive data table for use with exact data match (EDM) based sensitive information types in Microsoft Purview. You learn how to choose primary and secondary fields, structure your data, and save it in a supported format so you can create an EDM schema.

Products and services this export procedure applies to

- New experience
- Classic experience

The sensitive data table is a text file containing rows of values against which you compare the content in your documents to identify sensitive data. These values might be personally identifiable information, product records, or other sensitive data in text form that you want to detect in your content and protect.

Once you export the data in your table (in one of the supported formats), you can create an EDM schema.

Define your EDM sensitive information type

When you define your EDM sensitive type, one of the most critical decisions is to define which fields are your primary fields. Primary fields need to follow a detectable pattern and be defined as searchable fields (columns) in your EDM schema. Secondary fields don't need to follow any pattern since they'll be compared against all the text surrounding matches to the primary fields.

Use these rules to help you decide which columns you should use as primary fields:

- If you must detect sensitive data based on the presence of a single value matching a field in your sensitive data table, regardless of the presence of any other sensitive data surrounding it, that column must be defined as a primary element for an EDM SIT.
- If multiple combinations of different fields in your sensitive data table must be detected in content, identify the columns that are common to most such combinations and designate them as primary elements. Designate combinations of the other fields as secondary elements.

- If a column you want to use as a primary element doesn't follow a detectable pattern, like `any text string` or follows detectable patterns that would be present somewhere in a large percentage of documents or emails, choose other, better structured, columns as primary elements.

For example, if you have the columns `full name`, `date of birth`, `account number`, and `Social Security Number`, even if the first and last names are the columns that are common to the different combinations of data you want to detect, such strings don't follow patterns that are easily identifiable and might be difficult to define as a sensitive information type. There are a number of reasons for this:

- some names might not start with an uppercase character
- some might be formed by two, three,or more words/strings
- some might contain numbers or other non-alphabetical characters. Dates of birth can be identified more easily but, since every email and most documents will contain at least one date, a `DateOfBirth` field is also not a good candidate. Instead, use fields such as Social Security numbers and account numbers, which are good candidates for primary fields.

Sample file templates

To make selecting your primary fields easier, we've put together some sample file templates for:

- U.S. Healthcare data
- U.S. Financial data
- U.S. Insurance data

These are comma separated value (.csv) files that have the most commonly used values across those industry verticals as column headers, along with Microsoft-generated synthetic values in the rows. Use the column headers to help you decide on your primary fields. Best practice is to export only the source data that is required. The column headers suggest the most relevant fields.

To learn how to use the sample file templates, see How to use the sample file templates.

Save sensitive data in *.csv*, *.tsv*, or *pipe-separated* format

Save your sensitive data as a supported text format by following these steps:

1. Identify the sensitive information you want to use. Export the data to an app such as Microsoft Excel and save the file as a text file. The file can be saved in any of the following formats: *.csv* (comma-separated values), *.tsv* (tab-separated values), or (|)(pipe-separated) format. The .tsv format is recommended in cases where your data values might include commas, such as street addresses. The data file can include a maximum of:

    - Up to 100 million rows of sensitive data
    - Up to 32 columns (fields) per data source
    - Up to 10 columns (fields) marked as searchable

2. Structure the sensitive data in the .csv or .tsv file such that the first row includes the names of the fields used for EDM-based classification. In your file you might have field names such as "ssn", "birthdate", "firstname", "lastname". The column header names can't include spaces or underscores. For example, a sample .csv file named *PatientRecords.csv*, and its columns include *PatientID*, *MRN*, *LastName*, *FirstName*, *SSN*, and more.

3. Pay attention to the format of the sensitive data fields; in particular, fields that might contain commas in their content. For example, a street address that contains the value "Seattle, WA" would be parsed as two separate fields if the .csv format is selected. To avoid this, use the .tsv format or surrounded the comma containing values by double quotes in the sensitive data table. If comma containing values also contain spaces, you need to create a custom SIT that matches the corresponding format. For example, a SIT that detects multi-word string with commas and spaces in it.

Next steps

Choose the next article based on the experience you're using:

- **For new experience**: Create EDM SIT sample file for the new experience

or

- **For classic experience**: Create the schema for exact data match based sensitive information types

See also

- Get started with exact data match based sensitive information types
- Learn about exact data match based sensitive information types

ⓘ**Note:** The author created this article with assistance from AI. Learn more

Last updated on 06/26/2026

Create an EDM SIT sample file (New experience)

Creating and making an exact data match (EDM) based sensitive information type (SIT) available is a multi-phase process. EDM-based SITs can be used in Microsoft Purview data loss prevention policies, eDiscovery, and certain content governance tasks.

Applies to: EDM SIT new experience

- New experience

If you want to create an EDM SIT using the classic experience see, Create EDM SIT classic experience.

Before you begin

Before you create the sample file, complete the following prerequisite:

- Make sure you've complete the steps in Export source data for exact data match based sensitive information type.

Formatting the sample file

The system will extract the column names from the sample file to create the schema, and will recommend base SITs to map the sample field data to. The sample file must be formatted identically to your source sensitive information table file and should contain synthetic values that are representative of your actual data. The file can be saved in **.csv** (comma-separated values), **.tsv** (tab-separated values), or pipe-separated (|) format, but should be the same as your actual source sensitive information table file. The **.tsv** format is recommended in cases where your data values include commas, such as street addresses.

- Use about 10-20 rows of data to ensure that the system has enough samples to work with.
- Field values that contain commas must be enclosed in quotes ".
- The first row must be the header row and contain column names.
- The file must contain at least one row of data.
- Each row of data must contain the correct number of fields, corresponding to the headers.
- The sample file can contain up to 32 columns.

- The sample file can't exceed 2.5 MB in size.
- Column (field) names must start with a letter, be at least three characters long, and consist of only alphanumeric characters (A-Z, a-z, 0-9) and can't include spaces, underscores, or other special characters.

For example, if your actual data uses tab delimited (**.tsv**) format and looks like this:

```
FirstName       LastName        PatientNumber   CreditCardNumber
Nathan  Rigby   111-22-3333     4111111111111111
Remy    Morris  444-55-6666     5111111111111111
Markus  Long    77-88-9999      6111111111111111
...
```

Then your sample file must have the same column headers, but use synthetic values for the rows, like this

Expand table

FirstName	LastName	PatientNumber	CreditCardNumber
Eric	Solomon	987-65-4321	9000000000000000
Lisa	Taylor	123-45-6789	500000000000000
Andre	Lawson	234-56-7890	200000000000000

How to use the sample file templates

If you're in the U.S. Healthcare, U.S. Financial Services, or U.S. Insurance industry verticals, you can start with the following sample file templates to speed up the sample file creation process. These files contain the most commonly used column headers across the respective industries as a well as synthetic values in the fields.

- U.S. Healthcare data
- U.S. Financial data
- U.S. Insurance data

To use these templates:

1. Download the sample file template for your industry.
2. Compare the column headers in the template to your actual source data and pick the ones you want to use as primary fields in your customized sample file.
3. Compare the formatting of your actual source data with the formatting of the synthetic values. Change the formatting of the synthetic values to match the formatting of your

source data values.

4. Save your customized sample file to use when you create EDM SIT schema and rule package.

> **Tip**
>
> When working in the new experience, you have the option to upload a sample file or enter the sample file values manually. We recommend creating the sample file.

Next steps

- **For new experience**: Create EDM SIT schema and rule package

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Create an EDM SIT schema and rule package (New experience)

You can create both the exact data match (EDM) schema and EDM sensitive information type (SIT) in the new experience by using a single workflow in the Microsoft Purview Portal.

Supported experiences

- New experience

If you want to create an EDM SIT using the classic experience, see Create an EDM SIT (Classic experience). If you need help deciding whether to use the new experience or the classic experience, see Choosing the right EDM SIT creation experience for you.

Before you begin

Make sure to complete the steps in the following articles before beginning the procedures described in this article.

1. Export source data for exact data match based sensitive information types
2. Create the EDM SIT sample file (New experience)

If you aren't familiar with EDM-based SITs or their implementation, ***it's essential*** to familiarize yourself with the concepts in the following articles:

- Learn about sensitive information types
- Learn about exact data match based sensitive information types
- Learn about multi-token matching
- Get started with exact data match based sensitive information types
- Create an exact data match SIT (New experience)
- Sensitive information type entity definitions

Permissions

To create, test, and deploy a custom sensitive information type (which an EDM SIT is) in the Microsoft Purview portal, use an account that is in *one* of these roles:

- Compliance administrator

- Global administrator

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

For more information about roles and permissions, see About admin roles in Office 365.

Important considerations

Following are several important considerations you should take into account when creating your EDM schema and EDM-based SITs.

- **Before selecting your primary elements,** review the built-in SITs to get an idea of which ones best meet your needs.
- **For each EDM SIT, you must select at least one primary element** and no more than 10 primary elements. If you have a multi-token corroborative data field, choose one of the following options:
    1. Select **Multi-token matching**.
    2. Map the elements in your multi-token fields to a SIT that can detect that content. (The more fields with unique values that you map, the more accurate your EDM SIT will be. Mapping multiple fields also helps improve performance, reduce the processing load, and avoid system timeouts.)
- When selecting the primary elements for your SIT, **select fields that ensure each row in your data table is unique**. For example, **don't** use fields like `FirstName` or `DateOfBirth`. Why? Because first names and dates of birth are likely to be duplicated throughout your sensitive data table. Instead, use fields with unique values, such as `SocialSecurityNumber` or `BankAccountNumber`.
- **Recommendation:** Build your EDM schema from a sample data file. If you build your EDM schema from a sample data file, make sure that the file adheres to the following requirements:
    - Data must be organized as a table, with columns and rows. Use your field names for the column headers. (The rows in your table correspond with your individual data items.)
    - Field names can include only alphanumeric characters.

    - Field names must start with a letter and must consist of at least three alphanumeric characters.
    - If these naming requirements are not met, errors may occur when uploading your sample data file.

- If you use a sample file of sensitive information to configure your EDM SIT, the system suggests the existing SIT for each field that best detects the uploaded data, if such a SIT is available. Microsoft Purview defaults to single-token matching for detecting sensitive content. So, if no existing SITs are able to detect the data for a field in your EDM schema, it applies the single-token matching mode. ***It is critical*** that you make sure that the SIT suggested for each element will detect the exact string you want to monitor:

    1. Make sure that the suggested SIT **doesn't contain any surrounding characters that differ from the content you want to detect**.
    2. Make sure that the suggested SIT **doesn't exclude any valid portion of the string as stored in your sensitive information table**.
    3. Make sure that the SIT you use **closely matches the *format*** of the data you want to detect. For example, look for something like `Nine digits with optional hyphens or spaces` rather than simply `digits`, or check for `A combination of 32 characters consisting of letters and digits` rather than simply `text strings`.

       Using SITs that closely match the format of the data you're trying to detect is another way to improve the accuracy of your results and shorten the time it takes for the matching to complete.

> ⓘ **Note**
>
> All of your data is saved as you navigate forward (**Next**) and backward (**Back**) through the tool while making your selections. Backward navigation only supports moving from top-level page to top-level page and from sub-page to sub-page. You can't navigate backward from a top-level page to the preceding sub-page or from a sub-page to a preceding top-level page.

Create your EDM schema and SIT

The following procedure provides step-by step guidance for creating your EDM schema and SITs using the new experience. For a conceptual overview and diagram of the EDM schema and SIT creation workflow, see Overview of the EDM workflow (New experience).

Create the EDM schema and sensitive information type

Follow these steps to create an EDM classifier in the Microsoft Purview portal:

1. Open **Data classification** > **EDM classifiers**.
2. Make sure the **New EDM experience** toggle is set to **On**.
3. Choose + **Create EDM classifier**.
4. Name the SIT and add a description. The system uses this name, appended with the word *schema*, for the associated schema it generates.
5. Choose **Next**.
6. Select the method you want to use for your schema: either **Upload a file containing sample data**, or **Manually define your data structure**. (Best practice is to upload a sample data file. The rest of this procedure assumes that you upload a sample data file.)

   In either case, you need the information discussed in Create an EDM SIT sample file (New experience) for your sample file.
7. Choose **Next**.
8. Select your sample file and then select **Upload file**. Choose **Next**.
   (If errors display during the upload, address them and then try again.)
9. On the **Select primary elements** page:
   a. In the **Primary element** column, select your primary element. Each primary element must be mapped to a SIT. Best practice is to select fields that show **Full match** under the **Match Validation** column.
   b. In the **Match mode** column for each field, designate which of the following matching options to apply:

      - **Option 1:** Do nothing to accept the system-suggested SIT.
      - **Option 2:** Expand the dropdown menu. Under **Sensitive Info type (SIT)**, choose the pencil (**Edit**) icon and then select another existing SIT.
      - **Option 3:** Under **Match mode** select **Single token**.
      - **Option 4:** Under **Match mode** select **Multi-token**.
10. Choose **Next**.

11. **Configure settings for data in selected columns**.

    - The toggle **Use the same settings for all columns** is set to **On** by default. If you want to use separate settings for each data field, set the toggle to **Off**.
    - The **Data in columns are case-insensitive** option is selected by default. To enforce **case-sensitive** detection, uncheck this box.
    - If needed, select the option to **Ignore delimiters and punctuation for data in all columns** You can then either select the delimiters and punctuation marks you want to ignore from a list or you can enter custom delimiters and punctuation marks to ignore.

    > ⓘ **Important**
    >
    > If you select the **Ignore Delimiters** option for the primary element column in your schema, make sure that the SIT you map it to is designed to match data both with and without the selected delimiters.

12. **Configure detection rules for primary elements**.

    - Specify the **proximity limit** in terms of the number of characters within which the supporting element should be found in proximity of the primary element. By default, it's set at 300.
    - Select the **proximity detection mode**.
        - **Wide Mode:** Matches the primary element against supporting elements from all primary matches within the defined proximity range. This mode provides a broader context, which can increase the likelihood of detecting relevant associations.
        - **Narrow Mode:** Matches the primary element only against its own supporting element within the proximity range. This mode focuses on tighter, more specific matches, reducing the chance of false positives.
    - Select the **unique counting mode**.
        - **Primary Match-level Counting:** Counts each matching combination of primary and supporting elements separately—even within the same row. Choose this if you want a detailed view of all sensitive matches.
        - **Row-level Counting:** Counts each row only once, even if multiple elements match in it. Use this when you care about how many unique records (rows) are sensitive, not how many elements matched.
    - Define the **detection rules for high, medium, and low confidence levels**. The high confidence rule should include more supporting elements within close proximity of the

primary element, whereas a low confidence rule would contain little to no supporting elements in close proximity

13. Choose **Submit**.

Once you're finished, EDM automatically generates one detection rule for each of the primary elements you identified. EDM also creates a *high-confidence* rule and a *medium-confidence* rule. High-confidence rules have more matching requirements than medium-confidence rules, which, in turn, have more requirements than low-confidence rules. (Low-confidence rules must be created manually.) You can review and edit these rules on the **Configure detection rules for primary elements** page.

> **Tip**
>
> Those elements that are not selected as primary can still be used as corroborative (supporting) evidence. The more supporting elements found that are in a defined proximity to primary elements, the higher the confidence that the match is a true positive.

Recommendations

Keep the following recommendations in mind when creating, downloading, and testing your EDM schema and SIT.

- Wait at least one hour after creating or editing a schema before downloading and using it for the EDM data upload. This helps ensure that the schema has synced with the system. If a schema is downloaded too soon, an error message might display when attempting to download the schema via the command line.

- Do not use the EDM Upload Agent to download, manually edit, and then re-upload a schema. Doing so results in an error because using the EDM Upload Agent to download a schema adds tags to the schema that don't pass schema creation checks.

- To help ensure that all corroborative evidence is detected, take one of the following actions:
  - Trim multi-token corroborative evidence fields to the maximum number of tokens supported by the multi-token feature (currently five tokens).
  - Map the multi-token field to a SIT that can fully detect the multi-token data. - After

creating or editing your EDM SIT, test it using the following PowerShell cmdlet and then wait 24 hours before testing it in a data loss prevention (DLP) policy solution.

PowerShell

```
Test-DataClassification  -ClassificationNames “[Your EDM sensitive info type]” -
TexttoClassify “[your own text to scan for matches]”
```

Next steps

After you create your EDM schema and rule package, hash and upload your sensitive data:

- Hash and upload the sensitive information source table for EDM SITs

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Use multi-token matching for corroborative evidence in EDM SITs

Article • 12/11/2023

To use multi-token matching to identify corroborative evidence for your EDM sensitive information types, you must explicitly opt in for multi-token support for each multi-token corroborative evidence field. You can do this either through the new EDM UI experience or through the schema XML update process. Following are the steps for each method.

Opt-in through the new EDM UI experience

The new EDM UI experience guides you through the process of creating or editing an EDM schema and uploading your data. You can access it from the Microsoft Purview Portal, under **Data classification** > **Sensitive info types** > **Exact data match**.

When you upload a sample file, EDM automatically maps the SIT that can best detect the sample data in each field. If none of the SITs in your environment can detect the sample data in a field, EDM defaults to single-token matching. Single-token matching means that EDM only compares the hashes of individual strings identified in DLP policy fields with each individual string in the content. The field selected as the primary element for your schema must be mapped to a SIT. In contrast, corroborative evidence fields can either be mapped to a SIT or can be configured for either single-token or multi-token matching.

However, if the sample data in a policy field contains multi-token values and if EDM is operating in single-token mode, a warning displays stating that using single-token matching might result in missed detections. In this case, you can switch to multi-token matching. When you use multi-token matching, EDM compares the hashes of consecutive tokens (strings) in the policy fields against consecutive tokens in the content, up to a maximum number of tokens supported by this feature. The current maximum is five (5) tokens. For example, if your field contains the value "Jane Doe", EDM compares the hashes of "Jane", "Doe", and "Jane Doe" with the content. It then produces a match if the hashes of any of these combinations are identical to the hashed value in your data source.

> **Tip**

To ensure multi-token match detection, trim multi-token corroborative evidence fields to the maximum number of tokens supported. Alternatively, map the multi-token fields to a SIT that can fully detect the multi-token data.

In contrast, if you select multi-token matching for a field that only contains single-token data, a warning displays stating that using multi-token matching might result in higher latency. Higher latency in matching occurs because multi-token matching is inefficient. Additionally, multi-token matching isn't required if the actual EDM data that is later hashed and uploaded is expected to be a single-token field. In general, if any of the fields you use for corroborative evidence can be mapped to a SIT, do that instead of relying on single-token or multi-token matching.

Make sure that your sample data is fully representative of your production data. If your sample file doesn't include multi-token values, but your production data does, you might not get a warning for multi-token matches. For instance, this can occur in cases where your sample data includes only single-token values for a `FirstName` field, but your production data includes multi-token values for the same field.

Opt-in through the EDM schema XML update process

You can also create or update an XML schema to implement multi-token matching for corroborative evidence using PowerShell.

Here's a sample XML schema with multi-token settings for protecting patient records:

XML

```
<EdmSchema xmlns="http://schemas.microsoft.com/office/2018/edm">
 <DataStore name="DemoMTCorroborativeEvidenceSchema"
            description="Test" version="1">
  <Field name="Name"
         searchable="false"
      caseInsensitive="true"
      ignoredDelimiters=""
      isMultiToken="false"/>
  <Field name="SSN"
         searchable="true"
      caseInsensitive="true"
      ignoredDelimiters=""
      isMultiToken="false"/>
  <Field name="PatientID"
      searchable="false"
      caseInsensitive="true"
      ignoredDelimiters=""
```

```
        isMultiToken="false"/>
    <Field name="MedicalCondition"
        searchable="false"
        caseInsensitive="true"
        ignoredDelimiters=""
        isMultiToken="true"/>
    <Field name="Address"
           searchable="false"
        caseInsensitive="true"
        ignoredDelimiters=""
        isMultiToken="true"/>[Learn about exact data match based sensitive
information types](sit-learn-about-exact-data-match-based-sits.md#learn-
about-exact-data-match-based-sensitive-information-types)
  </DataStore>
</EdmSchema>
```

Each corroborative evidence field can be configured for multi-token support through the new XML parameter `isMultiToken`. This parameter can be set to `true` or `false`. Any field with `isMultiToken` set to `true` is treated as a multi-token field and is compared with consecutive strings in the content, up to the maximum number of tokens supported by this feature. EDM treats fields with `isMultiToken` set to `false`, or that don't have this parameter, as single-token fields. These fields are only compared with individual strings in the content.

For more information about what an EDM schema is, see the Schema section of the Learn about exact data match based sensitive information types article.

For details about creating an EDM schema, see Create your EDM schema and SIT.

Feedback

Was this page helpful? Yes No

Test an exact data match sensitive information type

This article explains how to test an exact data match (EDM) sensitive information type (SIT) to verify that it correctly detects the sensitive data you want to protect. It covers testing methods available in the Microsoft Purview portal and PowerShell, as well as troubleshooting tips for common detection issues.

Products and services this EDM testing article applies to

This article covers testing for both EDM experience versions:

For instructions on creating an EDM SIT before testing, see the workflow for your experience:

- Create an EDM SIT using the new experience
- Create an EDM SIT using the classic experience

Overview of EDM SIT testing

This section explains how to verify that your exact data match (EDM) sensitive information type (SIT) correctly detects the information you want to protect. After your EDM SIT has been created, and an hour after verifying that your sensitive information table has finished uploading and indexing, you can test detection by using the **Test** function in the **Sensitive information types** section in the Microsoft Purview Portal.

> **Note**
>
> Changes in an existing EDM SIT can take some time to propagate across the system. If you're making changes to an EDM SIT in order to troubleshoot detection issues, make sure to wait at least one hour after making those changes before using the **Test** function to validate their impact.

Regardless of the method you use for testing, the test results include matches for both the specific EDM SIT and for the primary elements that are configured for that EDM SIT.

Methods for testing your EDM SIT

There are two methods that you can use to test your EDM SIT.

Expand table

Method	Available in New EDM experience	Available in New and Classic EDM experience
Sensitive information type (SIT) method	Yes	Yes
EDM classifiers method	Yes	No

> **Note**
>
> If you're using the Classic EDM experience, you must use the SIT method.

Testing an EDM SIT with the Sensitive Information Types method

To test an EDM SIT with the **Sensitive Information Types** method, take the following steps.

1. Sign in to the Microsoft Purview portal > **Information Protection** > **Classifiers** > **Sensitive info types**.
2. Select your EDM SIT from the list and then select the **Test** icon.
3. In the flyout pane, upload a file that contains data you want to detect. For example, create a file that contains a subset of the rows in your sensitive information table. If you used the *configurable match* feature in your schema to define ignored delimiters, make sure the sample file includes examples with and without those delimiters.
4. Choose **Test**.
5. After the file has been uploaded and scanned, check for matches to your EDM SIT.
6. If the **Test** function in the SIT detects a match, verify that the SIT isn't trimming it or extracting the matched item incorrectly. Common issues include SITs that:
    - Extract only a substring of the full string that should be detected
    - Pick up only the first word in a multi-word string

- Include extra symbols or characters in the extraction

For details about the regular expressions used in SIT pattern definitions and extraction logic, see the Regular Expression Language - Quick Reference.

Test your EDM SIT with the EDM Classifier method

To test an EDM SIT with the **EDM Classifier** method, follow these steps.

1. Sign in to the Microsoft Purview portal > **Information Protection** > **Classifiers** > **EDM classifiers**.
2. Make sure that the **New EDM experience** toggle is set to **On**.
3. Select your EDM SIT from the list and then select the **Test** icon.
4. Upload a file that contains data you want to detect. For example, create a file that contains a subset of the rows in your sensitive information table. If you used the *configurable* match feature in your schema to define ignored delimiters, make sure your sample file includes examples with and without those delimiters.
5. After the file has been uploaded and scanned, check for matches to your EDM SIT.
6. If the **Test** function in the SIT detects a match, verify that the SIT isn't trimming it or extracting the matched item incorrectly. Common issues include SITs that:
    - Extract only a substring of the full string that should be detected
    - Pick up only the first word in a multi-word string
    - Include extra symbols or characters in the extraction

Test your EDM SIT using PowerShell

Use the `Test-DataClassification` cmdlet to test your EDM sensitive information type against sample text and verify whether the classifier detects expected matches. For the `-ClassificationNames` parameter, supply the display name of the EDM sensitive information type you want to test. For `-TexttoClassify`, provide the sample text you want to scan for matches.

PowerShell

```
Test-DataClassification  -ClassificationNames “[Your EDM sensitive info type]” -TexttoClassify “[your own text to scan for matches]”
```

Regardless of the method you use for testing, the test results will include matches for both the specific EDM SIT and for the primary elements that are configured for that EDM SIT.

> **Note**
>
> An uploaded-file or text test doesn't refresh classification for existing SharePoint and OneDrive files. Changing an EDM SIT or its primary SIT doesn't by itself refresh classification for unchanged historical files. To update classification for those files, use **on-demand classification**. Review the scan scope, EDM scan-cost considerations, and active-policy effects before starting a scan.

Test your EDM SIT with information protection policies

You can see where your EDM SIT is being used, and how accurate it is in production, by using it in policies:

1. Create an auto-labeling policy and run it in **Simulation overview**.
2. Add some content that triggers the EDM SIT, along with content that won't trigger the EDM SIT, to a location that your policy is monitoring.
3. Open the **Items to review** tab to check the matches.
4. Based on the simulation results, tune your policy conditions and EDM SIT settings as appropriate to improve detection accuracy.

Once you're satisfied with the results of your testing and tuning, your EDM-based custom SIT is ready for use in information protection policies, for instance:

- DLP policies
- Auto-labeling policies
- Microsoft Defender for Cloud Apps

Troubleshooting tips

If your EDM SIT doesn't detect any matches in your data, the following tips might help you diagnose the problem.

Expand table

Issue	Troubleshooting tip
No matches found	Confirm that your sensitive data was uploaded correctly using the commands explained in Hash and upload the sensitive information source table for exact data match sensitive information types.
No matches found	Test the SIT you used when you configured the primary element in each of your patterns. This test verifies whether the SIT can match the examples in the item. Using an incorrectly defined SIT as the classification element of an EDM SIT is the most common cause for detection failures in EDM.
The SIT you selected for a primary element in the EDM type doesn't find a match in the item or finds fewer matches than you expected	Confirm that the SIT supports the separators and delimiters that occur in the content. Be sure to include the ignored delimiters defined in your schema.
The SIT associated with your primary element finds matches in your content, but the EDM SIT doesn't.	• Check whether your REGEX statements are catching whitespace delimiters at the start or end of an item you want to detect. For instance, look for statements that include the `\s` delimiter. If whitespace delimiters are included, the whitespace won't match the hashed value in the data table. Instead, use a word delimiter, such as `\b`. • Check your REGEX statements to ensure that they capture the entire string you want to detect, not just a substring. For example, consider this pattern for email addresses: `\b[a-zA-Z]{2,30}@[a-zA-Z]{2,20}.[a-zA-Z]{2,3}\b`. This pattern will correctly match *user@contoso.com*, but will only capture *user@contoso.co.jp* in an incomplete form.
An EDM SIT with primary elements, but no defined secondary elements, detects items but doesn't detect matches (or detects fewer matches than expected) when both primary and secondary elements are required.	If values in a column used for secondary evidence aren't composed of a single word or of strings that don't contain spaces, commas, or other word separators, there are two ways to test: 1. Select the multi-token matching option. 2. Associate the values with a SIT that uses either a REGEX designed to detect multi-word strings that follow the desired pattern (for example, a fixed number of consecutive words that start with an uppercase character), or a keyword dictionary that lists all of the unique values in that column. For example, if there's an additional evidence column for a person's city or residence, you can create a list with all the unique city names from the table and then use it to create a dictionary-based sensitive information type. Use this SIT as the classification element for the corresponding column in your EDM SIT by exporting and editing the EDM SIT definition in XML. For more information, see Create a rule package manually.

Issue	Troubleshooting tip
SIT test function doesn't detect any matches at all.	Verify that the SIT you selected includes requirements for additional keywords or other validations. For built-in SITs, see Sensitive information type entity definitions to determine what the minimum requirements are for matching each type.
The **Test** functionality works but your SharePoint or OneDrive items aren't being detected in DLP or auto-labeling rules	Check the policy scope and conditions, and whether classification for the existing files reflects the SIT definitions you intend to use. A direct test doesn't establish classification coverage for stored files. For auto-labeling, see Auto-labeling classification and simulation concepts; restarting simulation doesn't itself refresh classification. Use on-demand classification when classification needs updating, after reviewing scope, costs, and active-policy effects.
DLP or auto-labeling rules that require multiple matches don't trigger	Make sure that the proximity requirements for both your EDM SIT and the base SITs are met. For example, if the maximum distance between the primary element and supporting keywords is 300 characters, but the keywords are only present in the first row of a long table, only the first few rows of matching values are likely to meet the proximity requirements. Modify your SIT definitions to support more relaxed proximity rules or use the **Anywhere in the document** option for the additional evidence conditions.
Detection of an EDM SIT is inconsistent or erratic	Make sure that the SIT you used as the base for the primary element in your EDM SIT isn't detecting unnecessary content. Using a SIT that matches too much unrelated content, such as *any* word, *any* number, or *all* email addresses, can cause the service to ignore relevant matches. Check the number of content pieces that match the sensitive type you used for your primary elements in content explorer. To estimate whether the SIT is matching too much content: 1. Divide the number of content items in the Microsoft Purview Content Explorer by the number of days since the sensitive type was created. 2. If the number of matches per day is in the range of hundreds of thousands or more, it's possible that the primary SIT is too broad.

For recommendations and best practices on selecting the right sensitive information type for an EDM SIT, see Learn about exact data match based sensitive information types.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 10/14/2026

Create an exact data match sensitive information type (Classic experience workflow)

09/26/2025

Creating and making a sensitive information type (SIT) based on exact data matching criteria (an EDM SIT) is a multi-phase process. These SITs can be used in Microsoft Purview data loss prevention policies, auto-labeling, eDiscovery, and certain content governance tasks. This article outlines the workflow and links to the procedures for each phase using the classic experience.

Applies to

- Classic experience

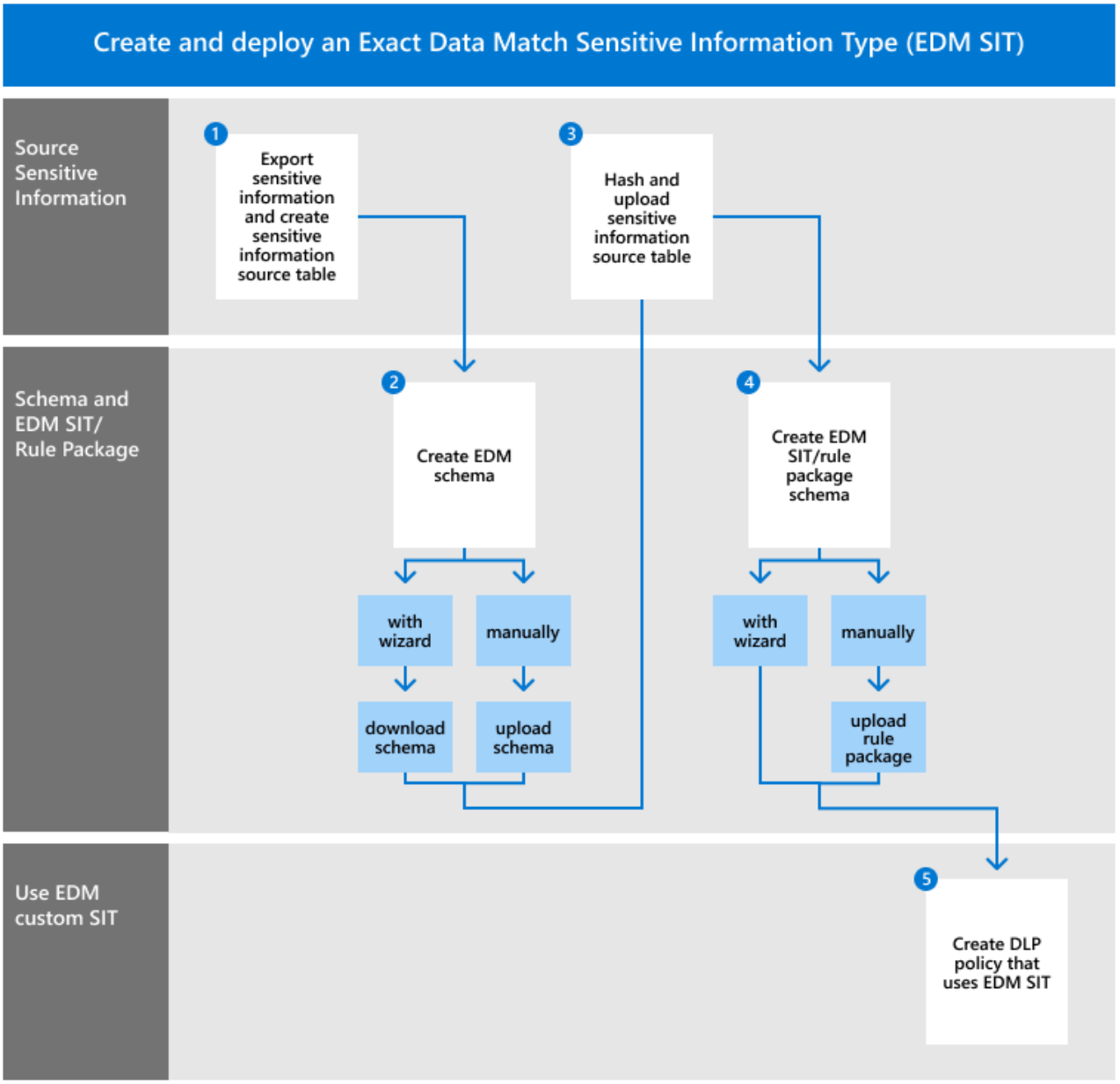
If you want to create an EDM SIT using the new experience see, Create an EDM SIT (New experience).

Before you begin

Make sure you review:

- Learn about EDM based SITs
- Get started with exact data match based sensitive information types overview

Overview of the EDM workflow (Classic experience)



Expand table

Phase	What's needed
Phase 1: Export source data for exact data match based sensitive information types	- Read access to the sensitive data
Phase 2: Create the schema for exact data match based sensitive information types	- Access to the sensitive information type wizard in the Microsoft Purviews portal - access to the Microsoft 365 admin center via Security & Compliance PowerShell
Phase 3: Hash and upload the sensitive information source table for exact data match sensitive information types	- Custom security group and user account - **Hash and upload from one computer**: local admin access to a computer with direct internet access and to host the EDM Upload Agent

Phase	What's needed
	- **Hash and upload from separate computers**: local admin access to a computer with direct internet access and host the EDM Upload Agent for the upload and local admin access to a secure computer to host the EDM Upload Agent to hash the sensitive information source table - Read access to the sensitive information source table file the schema file
Phase 4: Create exact data match sensitive information type/rule package	- Access to the Microsoft Purview portal
Test an exact data match sensitive information type	- Access to the Microsoft Purview portal

See also

- Learn about exact data match based sensitive information types
- Export source data for exact data match based sensitive information types

Export source data for exact data match based sensitive information types

This article explains how to prepare and export your sensitive data table for use with exact data match (EDM) based sensitive information types in Microsoft Purview. You learn how to choose primary and secondary fields, structure your data, and save it in a supported format so you can create an EDM schema.

Products and services this export procedure applies to

- New experience
- Classic experience

The sensitive data table is a text file containing rows of values against which you compare the content in your documents to identify sensitive data. These values might be personally identifiable information, product records, or other sensitive data in text form that you want to detect in your content and protect.

Once you export the data in your table (in one of the supported formats), you can create an EDM schema.

Define your EDM sensitive information type

When you define your EDM sensitive type, one of the most critical decisions is to define which fields are your primary fields. Primary fields need to follow a detectable pattern and be defined as searchable fields (columns) in your EDM schema. Secondary fields don't need to follow any pattern since they'll be compared against all the text surrounding matches to the primary fields.

Use these rules to help you decide which columns you should use as primary fields:

- If you must detect sensitive data based on the presence of a single value matching a field in your sensitive data table, regardless of the presence of any other sensitive data surrounding it, that column must be defined as a primary element for an EDM SIT.
- If multiple combinations of different fields in your sensitive data table must be detected in content, identify the columns that are common to most such combinations and designate them as primary elements. Designate combinations of the other fields as secondary elements.

- If a column you want to use as a primary element doesn't follow a detectable pattern, like `any text string` or follows detectable patterns that would be present somewhere in a large percentage of documents or emails, choose other, better structured, columns as primary elements.

For example, if you have the columns `full name`, `date of birth`, `account number`, and `Social Security Number`, even if the first and last names are the columns that are common to the different combinations of data you want to detect, such strings don't follow patterns that are easily identifiable and might be difficult to define as a sensitive information type. There are a number of reasons for this:

- some names might not start with an uppercase character
- some might be formed by two, three,or more words/strings
- some might contain numbers or other non-alphabetical characters. Dates of birth can be identified more easily but, since every email and most documents will contain at least one date, a `DateOfBirth` field is also not a good candidate. Instead, use fields such as Social Security numbers and account numbers, which are good candidates for primary fields.

Sample file templates

To make selecting your primary fields easier, we've put together some sample file templates for:

- U.S. Healthcare data
- U.S. Financial data
- U.S. Insurance data

These are comma separated value (.csv) files that have the most commonly used values across those industry verticals as column headers, along with Microsoft-generated synthetic values in the rows. Use the column headers to help you decide on your primary fields. Best practice is to export only the source data that is required. The column headers suggest the most relevant fields.

To learn how to use the sample file templates, see How to use the sample file templates.

Save sensitive data in *.csv*, *.tsv*, or *pipe-separated* format

Save your sensitive data as a supported text format by following these steps:

1. Identify the sensitive information you want to use. Export the data to an app such as Microsoft Excel and save the file as a text file. The file can be saved in any of the following formats: *.csv* (comma-separated values), *.tsv* (tab-separated values), or (|)(pipe-separated) format. The .tsv format is recommended in cases where your data values might include commas, such as street addresses. The data file can include a maximum of:

    - Up to 100 million rows of sensitive data
    - Up to 32 columns (fields) per data source
    - Up to 10 columns (fields) marked as searchable

2. Structure the sensitive data in the .csv or .tsv file such that the first row includes the names of the fields used for EDM-based classification. In your file you might have field names such as "ssn", "birthdate", "firstname", "lastname". The column header names can't include spaces or underscores. For example, a sample .csv file named *PatientRecords.csv*, and its columns include *PatientID*, *MRN*, *LastName*, *FirstName*, *SSN*, and more.

3. Pay attention to the format of the sensitive data fields; in particular, fields that might contain commas in their content. For example, a street address that contains the value "Seattle, WA" would be parsed as two separate fields if the .csv format is selected. To avoid this, use the .tsv format or surrounded the comma containing values by double quotes in the sensitive data table. If comma containing values also contain spaces, you need to create a custom SIT that matches the corresponding format. For example, a SIT that detects multi-word string with commas and spaces in it.

Next steps

Choose the next article based on the experience you're using:

- **For new experience**: Create EDM SIT sample file for the new experience

or

- **For classic experience**: Create the schema for exact data match based sensitive information types

See also

- Get started with exact data match based sensitive information types
- Learn about exact data match based sensitive information types

ⓘ**Note:** The author created this article with assistance from AI. Learn more

Last updated on 06/26/2026

Create the schema for exact data match based sensitive information types

Applicability for classic EDM schema creation

This procedure applies to the classic exact data match (EDM) sensitive information type (SIT) creation experience.

To simplify schema file creation, you can use the exact data match schema and sensitive information type pattern tool.

If you aren't familiar with EDM-based SITS or their implementation, you should familiarize yourself with:

- Learn about sensitive information types
- Learn about exact data match based sensitive information types
- Get started with exact data match based sensitive information types

A single EDM schema can be used in multiple sensitive information types that use the same sensitive data table. You can create up to 10 different EDM schemas in a Microsoft 365 tenant.

Use the Exact Data Match Schema and Sensitive Information Type Tool

You can use the exact data match schema and sensitive information type tool to help simplify the schema file creation process.

Prerequisites

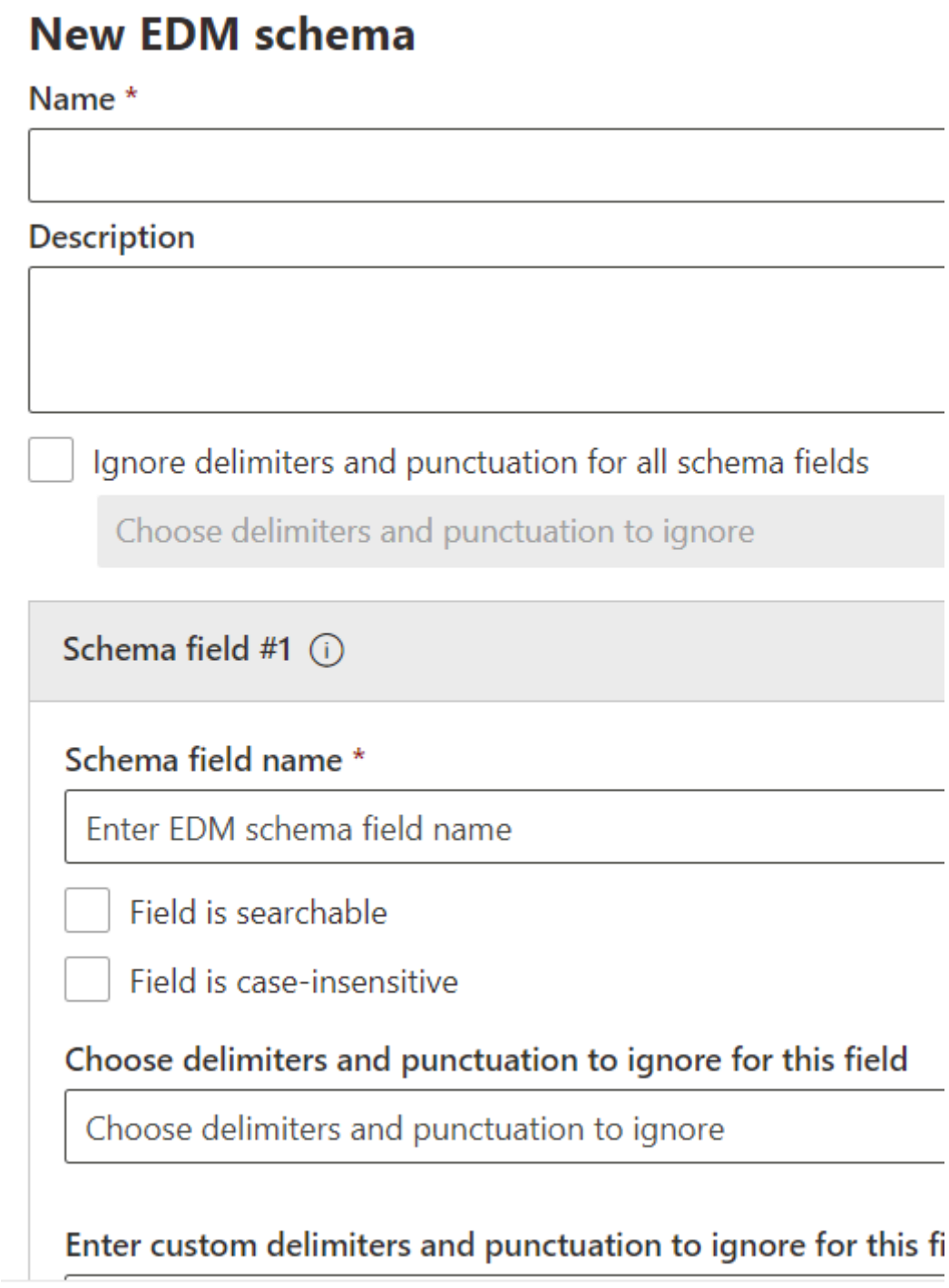
- Perform the steps in Export source data for exact data match based sensitive information type.

Use the exact data match schema and sensitive information type pattern tool

Perform the following steps to create an EDM schema by using the pattern tool:

1. Sign in to the Microsoft Purview portal > **Information Protection** > **Classifiers** > **EDM classifiers** > **EDM schemas** (available when the **New EMD experience** is toggled to **Off**).

2. Choose **Create EDM schema** to open the schema tool configuration flyout.



3. Fill in an appropriate **Name** and **Description**.

4. Choose **Ignore delimiters and punctuation for all schema fields** if you want to apply the **Ignore...** behavior for the entire schema. For more information about configuring EDM to ignore case or delimiters, see Using the caseInsensitive and ignoredDelimiters fields for more details on this feature.

5. Fill in your desired values for your **Schema field #1** and add more fields as needed. Each schema field must be identical to the column headers in your sensitive information source file.

6. If you want, set the per-field values for the following:

   - **Field is searchable**
   - **Field is case-insensitive**
   - **Choose delimiters and punctuation to ignore for this field**
   - **Enter custom delimiters and punctuation for this field**

   > ⓘ **Important**
   >
   > At least one, but no more than ten, of your schema fields must be designated as searchable.

7. Choose **Save**. Your schema is now listed and available for use.

   > ⓘ **Important**
   >
   > If you want to remove a schema that is already associated with an EDM SIT, you must first delete the EDM SIT. Deleting a schema that has a data store associated with it also deletes the data store within 24 hours.

Exporting the EDM schema file in XML format

If you created the EDM schema in the EDM schema tool, you must export the schema file in XML format. You'll need the XML file to complete the Hash and upload the sensitive information source table for exact data match sensitive information types phase.

1. Connect to Security & Compliance PowerShell.

2. To export the EDM schema to a local XML file so you can review or modify its field definitions, use this syntax:

   PowerShell

   ```
   $Schema = Get-DlpEdmSchema -Identity "[your EDM Schema name]"
   Set-Content -Path ".\Schemafile.xml" -Value $Schema.EdmSchemaXML
   ```

3. Save this file for later use.

Create and upload the exact data match schema file manually

As you create your schema file, your column headers (data fields) must adhere to the following naming requirements:

- Must start with a letter and must consist of at least three alphanumeric characters.
- Must include only alphanumeric characters.

Use the following syntax for each column/data field:

XML

```xml
<Field name="FieldName" searchable="true/false" caseInsensitive="true/false"
ignoredDelimiters="delimiter characters" />
```

Using the caseInsensitive and ignoredDelimiters fields

The following XML schema example uses the `caseInsensitive` and `ignoredDelimiters` fields to show how EDM handles case differences and delimiter characters.

When you include the `caseInsensitive` field set to the value of `true` in your schema definition, EDM won't exclude an item based on case differences. For example, EDM sees the values **FOO-1234** and **fOo-1234** as being identical for the `PatientID` field.

When you include the `ignoredDelimiters` field with supported characters, EDM ignores those characters. So EDM sees the values **FOO-1234** and **FOO#1234** as being identical for the `PatientID` field.

When both `caseInsensitive` and `ignoredDelimiters` are used together, EDM sees **FOO-1234** and **fOo#1234** as identical and classifies the item as a patient record sensitive information type.

Both `caseInsensitive` and `ignoredDelimiters` are used on a per-field basis.

> ⓘ **Important**
>
> If you configure *spaces* to be ignored, that configuration will only be effective for primary field columns for which a sensitive information type that can detect multi-word strings is

defined. Otherwise, the comparison will be made against each individual word in the content being analyzed.

The `ignoredDelimiters` flag supports any nonalphanumeric character, here are some examples:

- -
- /
- _
- *
- ^
- #
- [
-]
- {
- }
- \
- ~

The `ignoredDelimiters` flag doesn't support:

- characters 0-9
- A-Z
- a-z
- "
- ,
- Delimiters used as *Token Separators*

ⓘ Important

When defining your EDM sensitive information type, `ignoredDelimiters` will not affect how the Classification sensitive information type associated with the primary element in an EDM pattern identifies content in an item. So, if you configure `ignoredDelimiters` for a searchable field, you have to make sure the sensitive information type used for a primary element based on that field will pick strings both with and without those characters present.

The number of columns in your sensitive information source table and the number of fields in your schema must match, order doesn't matter.

The characters that are used as *token separators* behave differently than the other delimiters. Here are some examples:

- \ (space)
- \t
- ,
- .
- ;
- ?
- !
- \r
- \n

When you include a *token separator*, EDM breaks the token where the separator is. For example, EDM sees the value **Middle-Last Name** into **Middle-Last** and **Name** for the `LastName` field. If the `ignoredDelimiters` is included for the `LastName` field with the character '-', that action only happens after the value is broken. In the end, EDM would see the following values **MiddleLast** and **Name**.

To use the following characters as `ignoredDelimiters` and not *token separators*, a SIT that matches the corresponding format needs to be associated with the field. For example, a SIT that detects a multi-word string with dashes in it needs to be associated with the `LastName` field.

- .
- ;
- !
- ?
- \

It's possible to associate sensitive information types (SITs) with secondary elements using PowerShell.

1. Define the schema in XML format (similar to the following example). Name this schema file **edm.xml** and then configure it such that, for each column in the sensitive information source table, there's a line that uses the syntax:

   `\<Field name="" searchable=""/\>`.

   - Use column names for *Field name* values.

- Use `searchable="true"` for the fields that you want to be searchable and primary fields up to a maximum of five fields. At least one field must be searchable.

As an example, the following XML file defines the schema for a patient records database, with five fields specified as searchable: `PatientID`, `MRN`, `SSN`, `Phone`, and `DOB`.

(You can copy, modify, and use our example.)

XML

```
<EdmSchema xmlns="http://schemas.microsoft.com/office/2018/edm">
      <DataStore name="PatientRecords" description="Schema for patient records"
version="1">
            <Field name="PatientID" searchable="true" caseInsensitive="true"
ignoredDelimiters="-,/,*,#,^" />
            <Field name="MRN" searchable="true" />
            <Field name="FirstName" />
            <Field name="LastName" />
            <Field name="SSN" searchable="true" />
            <Field name="Phone" searchable="true" />
            <Field name="DOB" searchable="true" />
            <Field name="Gender" />
            <Field name="Address" />
      </DataStore>
</EdmSchema>
```

Once you have created the EDM schema file in XML format, you have to upload it to the cloud service.

2. Connect to Security & Compliance PowerShell.

3. To upload the database schema, run the following command:

PowerShell

```
New-DlpEdmSchema -FileData ([System.IO.File]::ReadAllBytes('.\\edm.xml')) -
Confirm:$true
```

You'll be prompted to confirm, as follows:

> **Confirm**
>
> **Are you sure you want to perform this action?**
>
> **New EDM Schema for the data store 'patientrecords' will be imported.**
>
> **[Y] Yes [A] Yes to All [N] No [L] No to All [?] Help (default is "Y"):**

> **Tip**
>
> If you want your changes to occur without confirmation, don't use `-Confirm:$true` in Step 3.

> **Note**
>
> It can take between 10-60 minutes to update the EDMSchema with additions. The update must complete before you execute steps that use the additions.

Next steps

After you create your EDM schema, proceed to hash and upload your sensitive data. See Hash and upload the sensitive information source table for exact data match sensitive information types.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Hash and upload the sensitive information source table for exact data match sensitive information types

Overview

This article shows you how to hash and upload your sensitive information source table as part of the exact data match (EDM) sensitive information type workflow. After you create an EDM schema and configure your sensitive information type, you use the EDM Upload Agent to hash your source data and securely upload it to Microsoft 365. The article covers prerequisites, data formatting requirements, and step-by-step instructions for both single-computer and two-computer upload methods. The procedures in this article apply to the following workflows:

Products and services this hashing and upload procedure applies to

- Create exact data match sensitive information type new experience
- Create exact data match sensitive information type classic experience

Hash and upload the sensitive information source table

In this phase, you:

1. Set up a custom security group and user account.
2. Set up the Exact Data Match (EDM) Upload Agent tool.
3. Use the EDM Upload Agent tool to hash, with a salt value, the sensitive information source table, and upload it.

To hash and upload your sensitive data, use either the **Two-computer method** or the **Single computer method** as described in Hash and upload your data. You can't use custom hashing for uploading data. You can only use these two methods as described in the documentation for hashing and uploading.

Best practice is to use two computers to separate the processes of hashing and uploading your sensitive data. Separating the steps across two computers helps ensure that your actual data is never available in clear text form on a computer that might be compromised due to its

connection to the internet. Separating hashing and uploading across two computers also makes any issues you encounter easier to isolate.

Prerequisites

Before you hash and upload your sensitive information source table, make sure you meet the following technology, account, schema, and data formatting requirements.

Technology requirements

To use the EDM Upload Agent to hash and upload your sensitive data, ensure that the following requirements are in place.

- A work or school account for Microsoft 365. This account must be added to the **EDM_DataUploaders** security group.
- A computer with one of the following operating systems. This computer runs the **EDM Upload Agent**.
  - Windows 10 and later versions
  - Windows server 2019 and later versions
  - Windows Server 2016 with .NET version 4.6.2
- A directory on the computer you use for uploading your data. This directory contains:
  - The EDM Upload Agent.
  - Your sensitive information data file in *.csv*, *.tsv*, or pipe (|) format. By default, the EDM Upload Agent expects your data file to be in .csv format. > [!TIP]

    > You can use a file with data that is separated by tabs or pipes (instead of commas), by indicating either the "(Tab)" or "(|)" options with the `/ColumnSeparated` parameter. For example: `EdmUploadAgent.exe /UploadData /DataStoreName PatientRecords /DataFile C:\Edm\Hash\PatientRecords.csv /HashLocation C:\Edm\Hash /Schema edm.xml /AllowedBadLinesPercentage 5`

  - The output hash and salt files that are created when completing the hash procedure.
  - The datastore name from the **edm.xml** file. Our example uses `PatientRecords`.

Security group and user account requirements

The account you use to upload your sensitive information source table must be a member of one of these roles to create and manage the exact data match sensitive information types.

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

Expand table

Role	Can create security groups and add members	Notes
Groups Administrator	Yes	Best least privilege option
User Administrator	Yes	Most commonly used role
Global Administrator	Yes	Grants full control, don't use

You can use an existing security group or create a new one for managing EDM data uploads.

1. Go to the admin center using the appropriate portal link for your subscription and create a security group called **EDM_DataUploaders**.
2. Add one or more users to the **EDM_DataUploaders** security group. These users manage the database of sensitive information.

Exact Data Match Schema

If you used the EDM schema and sensitive information type tool for the **new** experience or the EDM sensitive information type/rule package for the **classic** experience, you **must** download that schema to hash your sensitive information source table. For more information, see Exporting the EDM schema file in XML format.

To download the EDM schema file for your sensitive information type, open a Command Prompt window and run the following command:

```
EdmUploadAgent.exe /SaveSchema /DataStoreName <schema name> /OutputDir <path to output folder>
```

Data formatting requirements

Before you hash and upload your sensitive data, run a search for any special characters in the table that might cause problems in parsing the content.

To catch formatting or column-mapping issues before you upload, validate the data file against the EDM schema by using the EDM Upload Agent with the following syntax:

```
PowerShell
EdmUploadAgent.exe /ValidateData /DataFile [data file] /Schema [schema file]
```

Common formatting issues

The following formatting problems commonly cause validation or parsing errors when you prepare your sensitive information source table.

1. **Mismatched number of columns:** This issue can be due to the presence of commas or quote characters within values in the table that EDM interprets as column delimiters. Unless they're surrounding a whole value, single and double quotes can cause the tool to misidentify the start and end of individual columns.
2. **Single quote characters or commas inside a value:** For example, if a person's name includes a single quote such as *Tom O'Neil* or a city's name starts with an apostrophe such as *'s-Gravenhage*, you need to modify the data export process used to generate the sensitive information table and surround such columns with double quotes.
3. **Double quote characters inside values:** Best practice is to use the tab-delimited format for the table. Tab-delimited tables are less susceptible to such issues.

Hash and upload your data

Select the appropriate EDM Upload Agent for your organization.

Expand table

EDM Upload Agent version	Description
Commercial + GCC	Most commercial customers should use this option.
GCC-High	This option is specifically for high-security government cloud subscribers.
DoD	This option is specifically for United States Department of Defense cloud customers.

Each version of the EDM Upload Agent automatically adds a salt value to your hashed data. If you choose, you can provide your own salt value by using the `/Salt` parameter when running the EDM Upload Agent.

> **⚠ Warning**
>
> If you choose to use your own salt value, you can't go back to use the standard version of the EDM Upload Agent for your use case. If you want to hash and upload from one computer, the computer you use must be able to connect directly to your Microsoft 365 tenant. Additionally, your clear-text sensitive information source table file must reside on this computer so that it can be hashed.

Two-computer method

Your sensitive information source table is formatted as clear-text. By using one computer for the hash step and a different computer for the upload step, you protect your data from being exposed in clear text on a computer with a direct connection to your Microsoft 365 tenant.

> **ⓘ Important**
>
> The two-computer method requires that **the same version of the EDM Upload Agent must be installed on both computers**. You can then copy the hash file and the salt file from the secure machine to a computer that can connect directly to your Microsoft 365 tenant.

1. On the computer in the secure environment, run the following command in a Command Prompt window: `EdmUploadAgent.exe /CreateHash /DataFile [data file] /HashLocation [hash file location] /Schema [Schema file] /AllowedBadLinesPercentage [value]` For example: `EdmUploadAgent.exe /CreateHash /DataFile C:\Edm\Data\PatientRecords.csv /HashLocation C:\Edm\Hash /Schema edm.xml /AllowedBadLinesPercentage 5`

   This command outputs a hashed file and a salt file with these extensions if you didn't specify the `/Salt <saltvalue>` option:

   - EdmHash
   - EdmSalt

2. Securely copy these files to the computer you use to upload your sensitive information source table (for example, **PatientRecords.csv**) to your tenant.

3. Authorize the EDM Upload Agent:
   a. As an admin, open a Command Prompt window.
   b. Switch to the directory where the EDM Upload Agent is installed. (The recommended directory is **C:\EDM\Data**.)
   c. Run the following command:

```
EDM Upload Agent.exe /Authorize
```

> **Important**
>
> You must run the EDM Upload Agent from the folder where it's installed and you must sindicate the full path to your data files.

4. Sign in with your work or school Microsoft 365 account. (The account that was added to the **EDM_DataUploaders** security group). Your tenant information is extracted from the user account to make the connection.

5. To upload the hashed data, run the following command in a Command Prompt window:

```
EdmUploadAgent.exe /UploadHash /DataStoreName \<DataStoreName\> /HashFile \
<HashedSourceFilePath\ /ColumnSeparator ["{Tab}"|"|"]
```

   For example: `EdmUploadAgent.exe /UploadHash /DataStoreName PatientRecords /HashFile C:\\Edm\\Hash\\**PatientRecords.EdmHash**`

6. To verify that the upload of your sensitive data was successful, run the following command in a Command Prompt window:

```
EdmUploadAgent.exe /GetDataStore
```

   If the upload was successful, a list of data stores and when they were last updated displays.

7. To display all of the data uploads to a particular store, and when they were updated, run the following command in a Command Prompt window:

```
EdmUploadAgent.exe /GetSession /DataStoreName <DataStoreName>
```

> **Tip**

To automate the hash and upload process after you have created it the first time, see Refresh your exact data match sensitive information source table file.

Advanced multi-token configuration options

The following settings control how the EDM Upload Agent handles multi-token values during data upload. Configure these options in the **app.config** file.

Admins can upload upload values that exceed the multi-token limit by truncating the exceeding token to columns.

- **Configuration**: Controlled via **app.config**
- **Default Behavior**: Support is disabled by default

Admins can disable the minimum multi-token optimization feature.

- **Configuration**: Controlled via **app.config**
- **Default Behavior**: Optimization is enabled by default.

EDM and double-byte character set languages

Exact data match supports double-byte characters, such as those used in Chinese, Japanese, and Korean. However, it doesn't support string matches for corroborative evidence encoded as double byte characters. Neither does it match multi-token CJK text detected in the classified content, unless globalization for EDM is enabled as described later in this document. In all cases, a SIT must be mapped to any multi-token text, both for the primary field and for corroborative evidence fields.

To invoke exact data matching for double-byte characters, take the following steps:

1. Create an EDM Sensitive Information Type (SIT) configured to match on the double-byte character set language, such as Japanese kanji.
2. Ensure that you download and install version 17.01.0495.0 (or later) of the EDM Upload Agent
3. Update the **EdmUploadAgent.exe.config** file's globalization parameter to true: `<add key="IsGlobalizationEnabled" value="true">`
4. Hash and upload a source table with the data to be matched.

EDM Parse Quotes Scenarios

`ParseQuotesExhaustiveScenariosByDefault` is an optional configuration attribute that's available while uploading the data. It is designed to resolve data ingestion errors caused by complex formatting in source files, such as CSVs containing fields with embedded commas, inconsistent single/double quoting, or escaped characters. By default, the value of `ParseQuotesExhaustiveScenariosByDefault is false`. To enable this setting add the ParseQuotesExhaustiveScenariosByDefault="true" in the upload **EdmUploadAgent.exe.config** file.

- Double Quotes ("): If a field starts with a double quote, the parser looks for the matching closing double quote. Commas inside the quotes are treated as text, not separators. For example: Input row (Raw CSV) values are "ID",First Name,"Title, Role" then Parsed values (result) will be: `ID`, `First Name`, `Title, Role`. Double quotes have a higher hierarchy than single quotes. A single quote found inside a double-quoted string is treated as a literal character.

- Single Quotes ('): If a field starts with a single quote, the parser treats it as a delimited string similar to double quotes.

- No Quotes: If the field does not start with a quote, the data is ingested as-is until the next separator is found.

Next steps

After you upload the source table, continue with the appropriate next step for your experience.

- **For the new experience**: Test an exact data match sensitive information type
- **For the classic experience**: Create exact data match sensitive information type/rule package

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Create exact data match sensitive information type/rule package

Products and services this rule package procedure applies to

- Create an EDM SIT using the classic experience workflow

You can create an exact data match (EDM) SIT (SIT) by using the Use the Exact Data Match schema and SIT pattern tool in the Microsoft Purview Portal, or you can create the rule package manually as an XML file. You can also combine the two methods by using one method to create the schema and later using the other method to edit it.

If you are not familiar with EDM-based SITS or their implementation, you should familiarize yourself with:

- Learn about sensitive information types
- Learn about exact data match based sensitive information types
- Get started with exact data match based sensitive information types

Prerequisites

Perform the steps in these articles:

1. Export source data for exact data match based sensitive information types
2. Create the schema for exact data match based sensitive information types
3. Hash and upload the sensitive information source table for exact data match sensitive information types

- Whether you will be creating an EDM SIT using the tool or the rule package XML file via PowerShell, you must have Compliance administrator or Global administrator role to create, test, and deploy a custom SIT through the UI. See About admin roles in Office 365.

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. This helps improve security for your organization. Global Administrator is a highly privileged role that should

only be used in scenarios where a lesser privileged role can't be used.

- Identify one of the built-in SITs to use as the Primary elements SIT.
  - If none of the built-in SITs will match the data in the column you selected you will have to create a custom SIT that does.
  - If you selected the Ignored Delimiters option for the primary element column in your schema, make sure the custom SIT you create will match data with and without the selected delimiters.
  - If you use a built-in SIT, make sure it will detect exactly the strings you want to select, and not include any surrounding characters or exclude any valid part of the string as stored in your sensitive information table.

See Sensitive information type entity definitions and Create custom sensitive information types.

Use the Exact Data Match schema and SIT pattern tool

You can use the Exact Data Match schema and SIT pattern tool to create your SIT files to help simplify the process.

An EDM SIT is composed of one or more patterns. Each pattern describes a combination of fields from the schema that will be used to identify sensitive content in a document or email (evidence).

1. Sign in to the Microsoft Purview portal > **Information Protection** > **Classifiers** > **EDM classifiers**.
2. a. Set the **New EDM experience** toggle to **Off**
3. Choose **EDM sensitive info types** and **Create EDM sensitive info type** to open the Sensitive Information Type configuration tool.
4. Select **Choose an existing EDM schema** and pick the schema you created in Create the schema for exact data match based sensitive information types. Select **Add**.
5. Choose **Next** and choose **Create pattern**.
6. Pick the **Confidence level** and **Primary element**. To learn more about confidence levels, see Learn about sensitive information types.
7. Choose the **Primary element's sensitive info type** to associate it with to define what text in the document will be compared with all the values in the primary element field. See SIT Entity Definitions to learn more about the available sensitive information types.

> **Important**
>
> Select a SIT that closely matches the format of the content you want to find. Selecting a SIT that matches unnecessary content, such as one that matches *all* text strings, or *all* numbers can cause excessive load in the system, which can result in sensitive information remaining undetected.

8. Select your **Supporting elements** and match options.
9. Choose **Done**.
10. Choose **Create pattern** if you want to create additional patterns for your EDM SIT.
11. Select **Next**.
12. Choose your desired **Recommended confidence level** and **Character proximity**. This will be the default value for the whole EDM SIT. (For information on character proximity, see Understanding proximity). Select **Next**.
13. Choose **Next** and fill in a **Name** and **Description for admins**.

    As you create your schema file, your column headers (data fields) must adhere to the following naming requirements:
    - Must start with a letter and must consist of at least three alphanumeric characters.
    - Must include only alphanumeric characters.
14. Review and choose **Submit**.

Edit or delete a SIT pattern

To edit or delete an existing EDM SIT pattern, perform the following steps:

1. Sign in to the Microsoft Purview portal > **Information Protection**> **Classifiers** > **EDM classifiers**.
2. a. Set the **New EDM experience** toggle to **Off**
3. Choose **EDM sensitive info types**.
4. Pick the EDM SIT you want to edit.
5. Choose **Edit EDM sensitive info type** or **Delete EDM sensitive info type** from the flyout.

6. For the steps to edit an EDM sensitive info type, see Use the Exact Data Match schema and SIT pattern tool.

Working with specific types of data

For performance reasons, it is critical that you use patterns that will minimize the number of unnecessary matches. For example, you might use a SIT based on the regular expression.

`\b\w*\b`

This regular expression would match every individual word or number in any document or email. Matching every word or number would cause the service to be overloaded with matches and miss detecting true matches. Using more precise patterns can avoid this situation. Here are some recommendations for identifying the right configuration for some common types of data.

Email addresses: Email addresses can be easy to identify, but because they are so common in sensitive content, they might cause a significant load in the system if used as a primary field. Use email adderss only as secondary evidence. If they must be used as primary evidence, when you define your custom SIT use logic to exclude items where email addresses are used as `From` or `To` fields in emails. Also use logic to exclude email addresses from your company's domain to reduce the number of unnecessary strings that need to be matched.

Phone numbers: Phone numbers can come in many different formats, including or excluding country/region prefixes, area codes, and separators. To reduce the false negatives while keeping load to a minimum, use them only as secondary elements, exclude all likely separators, like parenthesis and dashes and only include in your sensitive data table the part that will always be present in the phone number.

People's names: Don't use people's names as primary elements if using a SIT based on a regular expression as the classification element for this EDM type, because they are difficult to distinguish from common words.

If you must use a primary element that is hard to identify with a specific pattern (such as a project code name), that could generate a high volume of matches to be processed, make sure you include keywords in the SIT you use as the classification element for your EDM type. For example, if using project code names that are also regular words, you can use the word `project` as required additional evidence in close proximity to the project name regular expression-based pattern in the SIT you use as the classification element for your EDM type. Or, you might consider using a SIT based on a regular dictionary as the classification element for your EDM SIT.

When trying to match numeric strings, specify the allowed ranges of numbers such as the number of digits or the starting digits, if known. If you need to match a relatively flexible range of numbers, you can use keywords in the base SIT to reduce the number of matches. For example, if trying to match account numbers consisting of 7-11 digits, add the words `account`, `customer`, `acct.` to the SIT as required additional evidence. This reduces the likelihood of unnecessary matches that could result in exceeding the limits of EDM matches that can be processed.

If a field you need to use as a primary element follows a simple pattern that might result in large numbers of matches, and you can't add the presence of keywords as additional evidence in the SIT, you can instead require a minimum number of occurrences of that pattern. For example, you could use a custom SIT defined in the following way to detect at least 29 other five-digit numbers surrounding a potential five-digit number to match against in your sensitive content:

XML

```
<Entity id="00aa00aa-bb11-cc22-dd33-44ee44ee44ee"
                patternsProximity="300"
                recommendedConfidence="85"
                relaxProximity="false">
                <Pattern confidenceLevel="85"
                           proximity="300">
                           <IdMatch idRef="MRN"/>
                           <Match idRef="30 AccountNrs"
                                 minCount="30"
                                 proximity="3000"
                                 uniqueResults="true"/>
                </Pattern>
     </Entity>
     <Regex id="30 AccountNrs">\d{5}</Regex>
```

In some cases, you might have to identify certain account or record identification numbers that for historical reasons don't follow a standardized pattern. For example, `Medical Record Numbers` can be composed of many different permutations of letters and numbers within the same organization. Even though it might be hard at first to identify a pattern, closer inspection often lets you narrow down a pattern that describes all valid values without causing an excessive number of invalid matches. For example, it might be detected that "all MRNs are at least seven characters in length, have at least two numerical digits in them, and if they have any letters in them, they start with one". Creating a regular expression based on such criteria should allow you to minimize unnecessary matches while capturing all the desired values, and further analysis might allow increased precision by defining separate patterns that describe different formats.

Create a rule package manually

This procedure shows you how to create a file in XML format called a rule package (with Unicode encoding), and then upload it into Microsoft Purview using Security & Compliance PowerShell cmdlets.

> **Note**
>
> If the SIT that you map to can detect multi-word corroborative evidence, the secondary elements you define in a manually created rule package can be mapped to the SIT. For example, the name `John Smith` would not match as a secondary element because we'd compare `John` and `Smith` found in the content separately to the term `John Smith` uploaded in one of the fields, if that corroborative evidence field wasn't mapped to a SIT that can detect that pattern.
>
> There's a limit of 10 rule packages in a Microsoft 365 tenant. Since a rule package can contain an arbitrary number of sensitive information types, you can avoid creating a new rule package each time you want to define a new SIT using this method, instead export an existing rule package and add your sensitive information types to the XML before re-uploading it.

1. Create a rule package in XML format (with Unicode encoding), similar to the following example. (You can copy, modify, and use our example.)

   When you set up your rule package, make sure to correctly reference your .csv, .tsv, or pipe (|) delimited sensitive information source table file and **edm.xml** schema file. You can copy, modify, and use our example. In this sample xml the following fields need to be customized to create your EDM sensitive type:

   - **RulePack id & ExactMatch id**: Use New-GUID to generate a GUID.
   - **Datastore**: This field specifies EDM lookup data store to be used. You provide the data source name of the configured EDM Schema.
   - **idMatch**: This field points to the primary element for EDM.
   - **Matches**: Specifies the field to be used in exact lookup. You provide a searchable field name in EDM Schema for the DataStore.
   - **Classification**: This field specifies the SIT match that triggers EDM lookup. You can use the name or GUID of an existing built-in or custom SIT.

> **Note**
>
> Be aware that any string that matches the SIT provided will be hashed and compared to every entry in the sensitive information source table. To avoid performance issues if you choose a custom SIT for the classification element, don't use one that will match a large percentage of content. For example one that matches "any number" or "any five-letter word". You can differentiate it by adding supporting keywords or including formatting in the definition of the custom classification SIT.

- **Match**: This field points to additional evidence found in proximity of idMatch.
- **Matches**: You provide any field name in EDM Schema for DataStore.
- **Resource idRef:** This section specifies the name and description for sensitive type in multiple locales
  - You provide GUID for ExactMatch ID.
  - **Name** & **description**: customize as required.

XML

```
<RulePackage xmlns="http://schemas.microsoft.com/office/2018/edm">
   <RulePack id="fd098e03-1796-41a5-8ab6-198c93c62b11">
     <Version build="0" major="2" minor="0" revision="0" />
     <Publisher id="..." />
     <Details defaultLangCode="en-us">
       <LocalizedDetails langcode="en-us">
         <PublisherName>IP DLP</PublisherName>
         <Name>Health Care EDM Rulepack</Name>
         <Description>This rule package contains the EDM sensitive type for
health care sensitive types.</Description>
       </LocalizedDetails>
     </Details>
   </RulePack>
   <Rules>
     <ExactMatch id = "E1CC861E-3FE9-4A58-82DF-4BD259EAB371"
patternsProximity = "300" dataStore ="PatientRecords" recommendedConfidence
= "65" >
       <Pattern confidenceLevel="65">
         <idMatch matches = "SSN" classification = "U.S. Social Security
Number (SSN)" />
       </Pattern>
       <Pattern confidenceLevel="75">
         <idMatch matches = "SSN" classification = "U.S. Social Security
Number (SSN)" />
         <Any minMatches ="3" maxMatches ="6">
           <match matches="PatientID" />
           <match matches="MRN"/>
```

```
            <match matches="FirstName"/>
            <match matches="LastName"/>
            <match matches="Phone"/>
            <match matches="DOB"/>
          </Any>
        </Pattern>
      </ExactMatch>
      <LocalizedStrings>
        <Resource idRef="E1CC861E-3FE9-4A58-82DF-4BD259EAB371">
          <Name default="true" langcode="en-us">Patient SSN Exact Match.
</Name>
          <Description default="true" langcode="en-us">EDM Sensitive type for
detecting Patient SSN.</Description>
        </Resource>
      </LocalizedStrings>
    </Rules>
</RulePackage>
```

2. Upload the rule package by running the following PowerShell command:

PowerShell

```
New-DlpSensitiveInformationTypeRulePackage -FileData
([System.IO.File]::ReadAllBytes('.\\rulepack.xml'))
```

> **Note**
>
> The syntax of the rule package file is the same as for other sensitive information types. For complete details on the syntax of the rule package file and for additional configuration options, and for instructions on modifying and deleting sensitive information types using PowerShell, Create a custom SIT using PowerShell.

Next step

After you create your rule package, continue with the following article:

- **For the Classic experience**: Test an exact data match SIT

Note: The author created this article with assistance from AI. Learn more

Last updated on 06/26/2026

Test an exact data match sensitive information type

This article explains how to test an exact data match (EDM) sensitive information type (SIT) to verify that it correctly detects the sensitive data you want to protect. It covers testing methods available in the Microsoft Purview portal and PowerShell, as well as troubleshooting tips for common detection issues.

Products and services this EDM testing article applies to

This article covers testing for both EDM experience versions:

For instructions on creating an EDM SIT before testing, see the workflow for your experience:

- Create an EDM SIT using the new experience
- Create an EDM SIT using the classic experience

Overview of EDM SIT testing

This section explains how to verify that your exact data match (EDM) sensitive information type (SIT) correctly detects the information you want to protect. After your EDM SIT has been created, and an hour after verifying that your sensitive information table has finished uploading and indexing, you can test detection by using the **Test** function in the **Sensitive information types** section in the Microsoft Purview Portal.

> **Note**
>
> Changes in an existing EDM SIT can take some time to propagate across the system. If you're making changes to an EDM SIT in order to troubleshoot detection issues, make sure to wait at least one hour after making those changes before using the **Test** function to validate their impact.

Regardless of the method you use for testing, the test results include matches for both the specific EDM SIT and for the primary elements that are configured for that EDM SIT.

Methods for testing your EDM SIT

There are two methods that you can use to test your EDM SIT.

Expand table

Method	Available in New EDM experience	Available in New and Classic EDM experience
Sensitive information type (SIT) method	Yes	Yes
EDM classifiers method	Yes	No

> **Note**
>
> If you're using the Classic EDM experience, you must use the SIT method.

Testing an EDM SIT with the Sensitive Information Types method

To test an EDM SIT with the **Sensitive Information Types** method, take the following steps.

1. Sign in to the Microsoft Purview portal > **Information Protection** > **Classifiers** > **Sensitive info types**.
2. Select your EDM SIT from the list and then select the **Test** icon.
3. In the flyout pane, upload a file that contains data you want to detect. For example, create a file that contains a subset of the rows in your sensitive information table. If you used the *configurable match* feature in your schema to define ignored delimiters, make sure the sample file includes examples with and without those delimiters.
4. Choose **Test**.
5. After the file has been uploaded and scanned, check for matches to your EDM SIT.
6. If the **Test** function in the SIT detects a match, verify that the SIT isn't trimming it or extracting the matched item incorrectly. Common issues include SITs that:
    - Extract only a substring of the full string that should be detected
    - Pick up only the first word in a multi-word string

- Include extra symbols or characters in the extraction

For details about the regular expressions used in SIT pattern definitions and extraction logic, see the Regular Expression Language - Quick Reference.

Test your EDM SIT with the EDM Classifier method

To test an EDM SIT with the **EDM Classifier** method, follow these steps.

1. Sign in to the Microsoft Purview portal > **Information Protection** > **Classifiers** > **EDM classifiers**.
2. Make sure that the **New EDM experience** toggle is set to **On**.
3. Select your EDM SIT from the list and then select the **Test** icon.
4. Upload a file that contains data you want to detect. For example, create a file that contains a subset of the rows in your sensitive information table. If you used the *configurable* match feature in your schema to define ignored delimiters, make sure your sample file includes examples with and without those delimiters.
5. After the file has been uploaded and scanned, check for matches to your EDM SIT.
6. If the **Test** function in the SIT detects a match, verify that the SIT isn't trimming it or extracting the matched item incorrectly. Common issues include SITs that:
    - Extract only a substring of the full string that should be detected
    - Pick up only the first word in a multi-word string
    - Include extra symbols or characters in the extraction

Test your EDM SIT using PowerShell

Use the `Test-DataClassification` cmdlet to test your EDM sensitive information type against sample text and verify whether the classifier detects expected matches. For the `-ClassificationNames` parameter, supply the display name of the EDM sensitive information type you want to test. For `-TexttoClassify`, provide the sample text you want to scan for matches.

PowerShell

```
Test-DataClassification  -ClassificationNames “[Your EDM sensitive info type]” -TexttoClassify “[your own text to scan for matches]”
```

Regardless of the method you use for testing, the test results will include matches for both the specific EDM SIT and for the primary elements that are configured for that EDM SIT.

> **Note**
>
> An uploaded-file or text test doesn't refresh classification for existing SharePoint and OneDrive files. Changing an EDM SIT or its primary SIT doesn't by itself refresh classification for unchanged historical files. To update classification for those files, use **on-demand classification**. Review the scan scope, EDM scan-cost considerations, and active-policy effects before starting a scan.

Test your EDM SIT with information protection policies

You can see where your EDM SIT is being used, and how accurate it is in production, by using it in policies:

1. Create an auto-labeling policy and run it in **Simulation overview**.
2. Add some content that triggers the EDM SIT, along with content that won't trigger the EDM SIT, to a location that your policy is monitoring.
3. Open the **Items to review** tab to check the matches.
4. Based on the simulation results, tune your policy conditions and EDM SIT settings as appropriate to improve detection accuracy.

Once you're satisfied with the results of your testing and tuning, your EDM-based custom SIT is ready for use in information protection policies, for instance:

- DLP policies
- Auto-labeling policies
- Microsoft Defender for Cloud Apps

Troubleshooting tips

If your EDM SIT doesn't detect any matches in your data, the following tips might help you diagnose the problem.

Expand table

Issue	Troubleshooting tip
No matches found	Confirm that your sensitive data was uploaded correctly using the commands explained in Hash and upload the sensitive information source table for exact data match sensitive information types.
No matches found	Test the SIT you used when you configured the primary element in each of your patterns. This test verifies whether the SIT can match the examples in the item. Using an incorrectly defined SIT as the classification element of an EDM SIT is the most common cause for detection failures in EDM.
The SIT you selected for a primary element in the EDM type doesn't find a match in the item or finds fewer matches than you expected	Confirm that the SIT supports the separators and delimiters that occur in the content. Be sure to include the ignored delimiters defined in your schema.
The SIT associated with your primary element finds matches in your content, but the EDM SIT doesn't.	• Check whether your REGEX statements are catching whitespace delimiters at the start or end of an item you want to detect. For instance, look for statements that include the `\s` delimiter. If whitespace delimiters are included, the whitespace won't match the hashed value in the data table. Instead, use a word delimiter, such as `\b`. • Check your REGEX statements to ensure that they capture the entire string you want to detect, not just a substring. For example, consider this pattern for email addresses: `\b[a-zA-Z]{2,30}@[a-zA-Z]{2,20}.[a-zA-Z]{2,3}\b`. This pattern will correctly match *user@contoso.com*, but will only capture *user@contoso.co.jp* in an incomplete form.
An EDM SIT with primary elements, but no defined secondary elements, detects items but doesn't detect matches (or detects fewer matches than expected) when both primary and secondary elements are required.	If values in a column used for secondary evidence aren't composed of a single word or of strings that don't contain spaces, commas, or other word separators, there are two ways to test: 1. Select the multi-token matching option. 2. Associate the values with a SIT that uses either a REGEX designed to detect multi-word strings that follow the desired pattern (for example, a fixed number of consecutive words that start with an uppercase character), or a keyword dictionary that lists all of the unique values in that column. For example, if there's an additional evidence column for a person's city or residence, you can create a list with all the unique city names from the table and then use it to create a dictionary-based sensitive information type. Use this SIT as the classification element for the corresponding column in your EDM SIT by exporting and editing the EDM SIT definition in XML. For more information, see Create a rule package manually.

Issue	Troubleshooting tip
SIT test function doesn't detect any matches at all.	Verify that the SIT you selected includes requirements for additional keywords or other validations. For built-in SITs, see Sensitive information type entity definitions to determine what the minimum requirements are for matching each type.
The **Test** functionality works but your SharePoint or OneDrive items aren't being detected in DLP or auto-labeling rules	Check the policy scope and conditions, and whether classification for the existing files reflects the SIT definitions you intend to use. A direct test doesn't establish classification coverage for stored files. For auto-labeling, see Auto-labeling classification and simulation concepts; restarting simulation doesn't itself refresh classification. Use on-demand classification when classification needs updating, after reviewing scope, costs, and active-policy effects.
DLP or auto-labeling rules that require multiple matches don't trigger	Make sure that the proximity requirements for both your EDM SIT and the base SITs are met. For example, if the maximum distance between the primary element and supporting keywords is 300 characters, but the keywords are only present in the first row of a long table, only the first few rows of matching values are likely to meet the proximity requirements. Modify your SIT definitions to support more relaxed proximity rules or use the **Anywhere in the document** option for the additional evidence conditions.
Detection of an EDM SIT is inconsistent or erratic	Make sure that the SIT you used as the base for the primary element in your EDM SIT isn't detecting unnecessary content. Using a SIT that matches too much unrelated content, such as *any* word, *any* number, or *all* email addresses, can cause the service to ignore relevant matches. Check the number of content pieces that match the sensitive type you used for your primary elements in content explorer. To estimate whether the SIT is matching too much content: 1. Divide the number of content items in the Microsoft Purview Content Explorer by the number of days since the sensitive type was created. 2. If the number of matches per day is in the range of hundreds of thousands or more, it's possible that the primary SIT is too broad.

For recommendations and best practices on selecting the right sensitive information type for an EDM SIT, see Learn about exact data match based sensitive information types.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 10/14/2026

Modify Exact Data Match schema to use configurable matching

This article walks administrators through modifying an existing Exact Data Match (EDM) schema to enable configurable matching using PowerShell. By adding the `caseInsensitive` and `ignoredDelimiters` flags to your schema, you can allow Microsoft Purview to match sensitive data while ignoring case differences and specified delimiter characters.

Applies to: EDM SIT creation with PowerShell

- Exact data match (EDM) sensitive information type (SIT) creation using PowerShell.

Exact Data Match (EDM) based classification enables you to create custom sensitive information types that refer to exact values in a database of sensitive information. When you need to allow for variants of an exact string, you can use *configurable matching* to tell Microsoft Purview to ignore case and some delimiters.

> **Important**
>
> Use this procedure to modify an existing EDM schema and data file to enable configurable matching using PowerShell. You must already have an EDM schema and uploaded data file before you begin.

1. Uninstall the **EdmUploadAgent.exe** from the computer that you use to connect to Microsoft 365 for EDM schema and data file upload purposes

2. Download the appropriate **EdmUploadAgent.exe** file for your subscription using the following links:

    - Download EDM Upload Agent for Commercial + GCC - Most commercial customers should use this option
    - Download EDM Upload Agent for GCC-High - This option is specifically for high-security government cloud subscribers
    - Download EDM Upload Agent for DoD - This option is specifically for United States Department of Defense cloud customers

3. Authorize the EDM Upload Agent so it can perform schema export and data upload operations in your environment. Open a Command Prompt window (as an administrator) and run the following command:

```dos
EdmUploadAgent.exe /Authorize
```

4. If you don't have a current copy of the existing schema, you need to download a copy of the existing schema. To do so, run this command:

```dos
EdmUploadAgent.exe /SaveSchema /DataStoreName <dataStoreName> [/OutputDir [Output dir location]]
```

5. Customize the schema so each column utilizes `caseInsensitive` and / or `ignoredDelimiters`. The default value for `caseInsensitive` is "false" and for `ignoredDelimiters`, the default is an empty string.

> **Note**
>
> The underlying custom sensitive information type or built-in sensitive information type used to detect the general regex pattern must support detection of the input variations listed with ignoredDelimiters. For example, the built-in U.S. Social Security Number (SSN) sensitive information type can detect variations in the data that include dashes, spaces, or lack of spaces between the grouped numbers that make up the SSN. As a result, the only delimiters that are relevant to include in EDM's `ignoredDelimiters` for SSN data are: dash and space.

Here's a sample schema that simulates case-insensitive matching by creating the extra columns needed to recognize case variations in the sensitive data.

```XML
<EdmSchema xmlns="http://schemas.microsoft.com/office/2018/edm">
  <DataStore name="PatientRecords" description="Schema for patient records policy" version="1">
            <Field name="PolicyNumber" searchable="true" />
            <Field name="PolicyNumberLowerCase" searchable="true" />
            <Field name="PolicyNumberUpperCase" searchable="true" />
            <Field name="PolicyNumberCapitalLetters" searchable="true" />
```

```
  </DataStore>
</EdmSchema>
```

In the sample schema with `PolicyNumberLowerCase`, `PolicyNumberUpperCase`, and `PolicyNumberCapitalLetters` columns, the variations of the original `PolicyNumber` column aren't necessary if both `caseInsensitive` and `ignoredDelimiters` are added.

To update this schema so that EDM uses configurable matching, use the `caseInsensitive` and `ignoredDelimiters` flags. Here's how that looks:

XML

```
<EdmSchema xmlns="http://schemas.microsoft.com/office/2018/edm">
  <DataStore name="PatientRecords" description="Schema for patient records
policy" version="1">
         <Field name="PolicyNumber" searchable="true" caseInsensitive="true"
ignoredDelimiters="-,/,*,#,^" />
  </DataStore>
</EdmSchema>
```

For information on the characters supported by the `ignoredDelimiters` flag, see Using the caseInsensitive and ignoredDelimiters fields.

6. Connect to Security & Compliance PowerShell

> **Note**
>
> If your organization has set up Customer Key for Microsoft 365 at the tenant level, Exact Data Match will make use of its encryption functionality automatically. This is available only to E5 licensed tenants in the Commercial cloud. For more information, see Overview of Customer Key.

7. Import the updated schema XML back into Microsoft Purview to apply the configurable matching changes to the EDM datastore. Run the following command:

PowerShell

```
Set-DlpEdmSchema -FileData ([System.IO.File]::ReadAllBytes('.\\edm.xml')) -
Confirm:$true
```

8. If necessary, update the data file to match the new schema version.

> **Tip**
>
> Optionally, you can run a validation against your CSV file before uploading it by running:
>
> `EdmUploadAgent.exe /ValidateData /DataFile [data file] /Schema [schema file]`
>
> For example: `EdmUploadAgent.exe /ValidateData /DataFile C:\data\testdelimiters.csv /Schema C:\EDM\patientrecords.xml`
>
> For more information on all of the parameters supported by EdmUploadAgent.exe, run
>
> `EdmUploadAgent.exe /?`

9. Open a Command Prompt window (as an administrator) and run the following command to hash and upload your sensitive data:

```dos
EdmUploadAgent.exe /UploadData /DataStoreName [DS Name] /DataFile [data file]
/HashLocation [hash file location] /Salt [custom salt] /Schema [Schema file]
```

Related articles

- Learn about exact data match based sensitive information types
- Sensitive information type-entity definitions
- Custom sensitive information types
- Learn about Microsoft Purview Data Loss Prevention
- Microsoft Defender for Cloud Apps
- New-DlpEdmSchema

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Create notifications for exact data match activities

When you create custom sensitive information types with exact data match (EDM), there are a number of activities that are created in the audit log. You can use the New-ProtectionAlert PowerShell cmdlet to create notifications that let you know when these activities occur:

- `CreateSchema`
- `EditSchema`
- `RemoveSchema`
- `UploadDataFailed`
- `UploadDataCompleted`

> **Note**
>
> The ability to create notifications for EDM activities is available for the World Wide and GCC clouds only.

Prerequisites

Your account must have one of the following roles:

- Global administrator
- Compliance administrator
- Exchange Online administrator

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. This helps improve security for your organization. Global Administrator is a highly privileged role that should only be used in scenarios where a lesser privileged role can't be used.

To learn more about DLP permissions, see Permissions in the Microsoft Purview portal.

EDM-based classification is included in these subscriptions:

- Office 365 E5

- Microsoft 365 E5
- Microsoft Purview Suite (formerly known as Microsoft 365 E5 Compliance)
- Microsoft E5/A5 Information Protection and Governance

To learn more about DLP licensing, see Microsoft 365 licensing guidance for security & compliance.

Before you begin, ensure you can connect to Security & Compliance PowerShell.

Configure notifications for EDM activities

1. Connect to Security & Compliance PowerShell.
2. Run the `New-ProtectionAlert` cmdlet using the activity that you want to create the notification for. For example, if you want to be notified when the `UploadDataCompleted` action occurred, run:

   PowerShell

   ```
   New-ProtectionAlert -Name "EdmUploadCompleteAlertPolicy" -Category Others -
   NotifyUser <address to send notification to> -ThreatType Activity -Operation
   UploadDataCompleted -Description "Custom alert policy to track when EDM upload
   Completed" -AggregationType None
   ```

   For the `UploadDataFailed` activity, you can run:

   PowerShell

   ```
   New-ProtectionAlert -Name "EdmUploadFailAlertPolicy" -Category Others -NotifyUser
   <SMTP address to send notification to> -ThreatType Activity -Operation
   UploadDataFailed -Description "Custom alert policy to track when EDM upload
   Failed" -AggregationType None -Severity High
   ```

Related articles

- Learn about exact data match based sensitive information types
- New-ProtectionAlert

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Manage your exact data match schema

Editing the schema for EDM-based classification manually

If you want to make changes to your EDM schema, for example the **edm.xml** file, such as changing which fields are used for EDM-based classification, follow these steps:

> **Tip**
>
> You can change your EDM schema and sensitive information table source file to take advantage of **configurable match**. When configured, EDM will ignore case differences and some delimiters when it evaluates an item. This makes defining your xml schema and your sensitive data files easier. To learn more see, Using the caseInsensitive and ignoredDelimiters fields.

1. Edit your **edm.xml** file (this file is discussed in the Create the schema for exact data match based sensitive information types.
2. Connect to Security & Compliance PowerShell.
3. To update your database schema, run the following command:

   PowerShell

   ```
   Set-DlpEdmSchema -FileData ([System.IO.File]::ReadAllBytes('.\\edm.xml')) -
   Confirm:$true
   ```

   You're prompted to confirm as follows:

   > Confirm
   >
   > Are you sure you want to perform this action?
   >
   > EDM Schema for the data store 'patientrecords' will be updated.
   >
   > [Y] Yes [A] Yes to All [N] No [L] No to All [?] Help (default is "Y"):

> **Tip**
>
> If you want your changes to occur without confirmation, don't use `-Confirm:$true` in Step 3.

> **Note**
>
> It can take between 10-60 minutes to update the EDMSchema with additions. The update must complete before you execute steps that use the additions.

Removing the schema for EDM-based classification manually

If you want to remove the schema you're using for EDM-based classification, follow these steps:

1. Connect to Security & Compliance PowerShell.
2. Run the following command, substituting the data store name of "patient records" with the one you want to remove (using the patientrecords store as an example):

```
PowerShell
Remove-DlpEdmSchema -Identity 'patientrecords'
```

You're prompted to confirm the following message:

> Confirm
>
> Are you sure you want to perform this action?
>
> EDM Schema for the data store 'patientrecords' will be removed.
>
> [Y] Yes [A] Yes to All [N] No [L] No to All [?] Help (default is "Y"):

> **Tip**
>
> If you want the schema removal to occur without confirmation, omit the `-Confirm:$true` parameter from the `Remove-DlpEdmSchema` command.

Edit or delete the EDM schema with the wizard

To edit or delete an EDM schema by using the wizard, follow these steps:

1. Open **Compliance center** > **Data classification** > **Classifiers**.
2. Choose **EDM classifiers**.
3. Double-click on the EDM SIT you want to edit.
4. Choose **Edit EDM sensitive info type** or **Delete EDM sensitive info type** from the flyout.
5. Follow the steps in the wizard to complete your edits or to delete the classifier.

> **Important**
>
> If you want to remove a schema, and it is already associated with an EDM sensitive info type, you must first delete the EDM sensitive info type, then you can delete the schema.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Refresh your exact data match sensitive information source table file

Refresh an exact data match sensitive information database

This article explains how to refresh the sensitive information source table used by exact data match (EDM) sensitive information types in Microsoft Purview. Use this procedure when your source data changes and you need to update the uploaded data. Compliance administrators can refresh the database up to 5 times in every 24-hour period by rehashing and reuploading the sensitive information source table. Each refresh overwrites the existing sensitive information source table with the new one.

1. Re-export the sensitive data to an app, such as Microsoft Excel, and save the file in .csv, .tsv format or pipe-delimited (|) format. Keep the same file name and location you used when you previously hashed and uploaded the file. See, Export source data for exact data match based sensitive information types for details on exporting your sensitive data and getting it into the correct format.

   > ⓘ **Note**
   >
   > If there are no changes to the structure (field names) of the sensitive information source table file, you won't need to make any changes to your database schema file when you refresh the data. But if you must make changes, make sure to edit the database schema and your rule package accordingly. See, Manage your exact data match schema for the steps to edit or remove a schema. To learn how to dit or remove your EDM SIT/rule package, see Create exact data match sensitive information type/rule package.

2. Use the procedures in Hash and upload the sensitive information source table for exact data match sensitive information types to upload your sensitive information table source file.

3. You can use Task Scheduler to automate the Hash and upload the sensitive information source table for exact data match sensitive information types procedure. You can schedule tasks using several methods:

Expand table

Method	What to do
PowerShell	See the ScheduledTasks documentation and the Example PowerShell script for Task Scheduler
Task Scheduler API	See the Task Scheduler documentation
Windows user interface	In Windows, click **Start**, and type Task Scheduler. Then, in the list of results, right-click **Task Scheduler**, and choose **Run as administrator**.

Example PowerShell script for Task Scheduler

The following examples include PowerShell scripts you can use to schedule tasks for hashing data and uploading hashed data:

Schedule hashing and upload in a combined step

Use the following PowerShell script to schedule hashing and upload as a single combined task:

PowerShell

```
param(\[string\]$dataStoreName,\[string\]$fileLocation)
\# Assuming current user is also the user context to run the task
$user = "$env:USERDOMAIN\\$env:USERNAME"
$edminstallpath = 'C:\\Program Files\\Microsoft\\EdmUploadAgent\\'
$edmuploader = $edminstallpath + 'EdmUploadAgent.exe'
$csvext = '.csv'
$schemaext = '.xml'
\# Assuming file name is same as data store name and file is in .csv format
$dataFile = "$fileLocation\\$dataStoreName$csvext"
\# Assuming location to store hash file is same as the location of csv file
$hashLocation = $fileLocation
\# Assuming Schema file name is same as data store name
$schemaFile = "$fileLocation\\$dataStoreName$schemaext"
$uploadDataArgs = '/UploadData /DataStoreName ' + $dataStoreName + ' /DataFile ' +
$dataFile + ' /HashLocation' + $hashLocation + ' /Schema ' + $schemaFile
\# Set up actions associated with the task
$actions = @()
$actions += New-ScheduledTaskAction -Execute $edmuploader -Argument $uploadDataArgs -
WorkingDirectory $edminstallpath
\# Set up trigger for the task
$trigger = New-ScheduledTaskTrigger -Weekly -DaysOfWeek Sunday -At 2am
\# Set up task settings
$principal = New-ScheduledTaskPrincipal -UserId $user -LogonType S4U -RunLevel
Highest
$settings = New-ScheduledTaskSettingsSet -RunOnlyIfNetworkAvailable -
StartWhenAvailable -WakeToRun
```

```
\# Create the scheduled task
$scheduledTask = New-ScheduledTask -Action $actions -Principal $principal -Trigger
$trigger -Settings $settings
\# Get credentials to run the task
$creds = Get-Credential -UserName $user -Message "Enter credentials to run the task"
$password=\[Runtime.InteropServices.Marshal\]::PtrToStringAuto(\
[Runtime.InteropServices.Marshal\]::SecureStringToBSTR($creds.Password))
\# Register the scheduled task
$taskName = 'EDMUpload\_' + $dataStoreName
Register-ScheduledTask -TaskName $taskName -InputObject $scheduledTask -User $user -
Password $password
```

Schedule hashing and upload as separate steps

Use the following PowerShell script to schedule hashing and upload as separate tasks:

PowerShell

```
param(\[string\]$dataStoreName,\[string\]$fileLocation)
\# Assuming current user is also the user context to run the task
$user = "$env:USERDOMAIN\\$env:USERNAME"
$edminstallpath = 'C:\\Program Files\\Microsoft\\EdmUploadAgent\\'
$edmuploader = $edminstallpath + 'EdmUploadAgent.exe'
$csvext = '.csv'
$edmext = '.EdmHash'
$schemaext = '.xml'
\# Assuming file name is same as data store name and file is in .csv format
$dataFile = "$fileLocation\\$dataStoreName$csvext"
$hashFile = "$fileLocation\\$dataStoreName$edmext"
\# Assuming Schema file name is same as data store name
$schemaFile = "$fileLocation\\$dataStoreName$schemaext "

\# Assuming location to store hash file is same as the location of csv file
$hashLocation = $fileLocation
$createHashArgs = '/CreateHash' + ' /DataFile ' + $dataFile + ' /HashLocation ' +
$hashLocation + ' /Schema ' + $schemaFile
$uploadHashArgs = '/UploadHash /DataStoreName ' + $dataStoreName + ' /HashFile ' +
$hashFile
\# Set up actions associated with the task
$actions = @()
$actions += New-ScheduledTaskAction -Execute $edmuploader -Argument $createHashArgs -
WorkingDirectory $edminstallpath
$actions += New-ScheduledTaskAction -Execute $edmuploader -Argument $uploadHashArgs -
WorkingDirectory $edminstallpath
\# Set up trigger for the task
$trigger = New-ScheduledTaskTrigger -Weekly -DaysOfWeek Sunday -At 2am
\# Set up task settings
$principal = New-ScheduledTaskPrincipal -UserId $user -LogonType S4U -RunLevel
Highest
$settings = New-ScheduledTaskSettingsSet -RunOnlyIfNetworkAvailable -
StartWhenAvailable -WakeToRun
\# Create the scheduled task
```

```
$scheduledTask = New-ScheduledTask -Action $actions -Principal $principal -Trigger
$trigger -Settings $settings
\# Get credentials to run the task
$creds = Get-Credential -UserName $user -Message "Enter credentials to run the task"
$password=\[Runtime.InteropServices.Marshal\]::PtrToStringAuto(\
[Runtime.InteropServices.Marshal\]::SecureStringToBSTR($creds.Password))
\# Register the scheduled task
$taskName = 'EDMUpload\_' + $dataStoreName
Register-ScheduledTask -TaskName $taskName -InputObject $scheduledTask -User $user -
Password $password
```

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Learn about trainable classifiers

Trainable Classifiers

This categorization method is well suited to content that can't be easily identified using either the manual or automated pattern-matching methods. A classifier learns how to identify a type of content by looking at hundreds of examples of the content you want to detect.

> ⓘ **Note**
>
> **In Preview:** You can view the trainable classifiers in content explorer by expanding **Trainable Classifiers** in the filters panel. The trainable classifiers automatically display the number of incidents found in SharePoint and Teams, without requiring any labeling. If you don't want to use this feature, you must file a request with Microsoft Support. This request disables the display of your sensitive data that isn't used in any labeling policies within Content Explorer. You can disable scanning of your data as well. If scanning is turned off, sensitivity labeling and DLP policies with those classifiers don't work.

Where you can use classifiers

Use classifiers as a condition for:

- Auto-labeling Office files with sensitivity labels
- Automatically applying a retention label policy based on a condition
- Communication compliance supports Microsoft provided trainable classifiers only.
- Sensitivity label conditions. (See Automatically apply a sensitivity label to Microsoft 365 data)
- Data loss prevention

> ⓘ **Important**
>
> Classifiers only work with items that aren't encrypted.

Types of classifiers

- **Microsoft provided Pretrained classifiers** - Microsoft created and pretrained multiple classifiers that you can start using without training them. These classifiers appear with the status of `Ready to use`. These classifiers evaluate items in all languages listed in the suppported languages columns.
- **Custom trainable classifiers** - If you need to identify and categorize your content beyond what the pretrained classifiers cover, you can create and train your own classifiers.

For a complete list of all pretrained classifiers, see Trainable classifiers definitions.

Custom classifiers

> **Important**
>
> **Language limitation:** Support for custom classifiers is limited to English.

When the Microsoft provided pretrained classifiers don't meet your needs, you can create and train your own classifiers. There's more work involved with creating your own, but they're better tailored to your organization's needs.

To create a custom trainable classifier, you start by feeding it one set of examples that are definitely in the category, and another set of examples that are definitely not in the category. Microsoft Purview processes those examples and the classifier then makes predictions as to whether any given item falls into the category you're building. You then confirm the results, sorting out the true positives, true negatives, false positives, and false negatives to help increase the accuracy of its predictions.

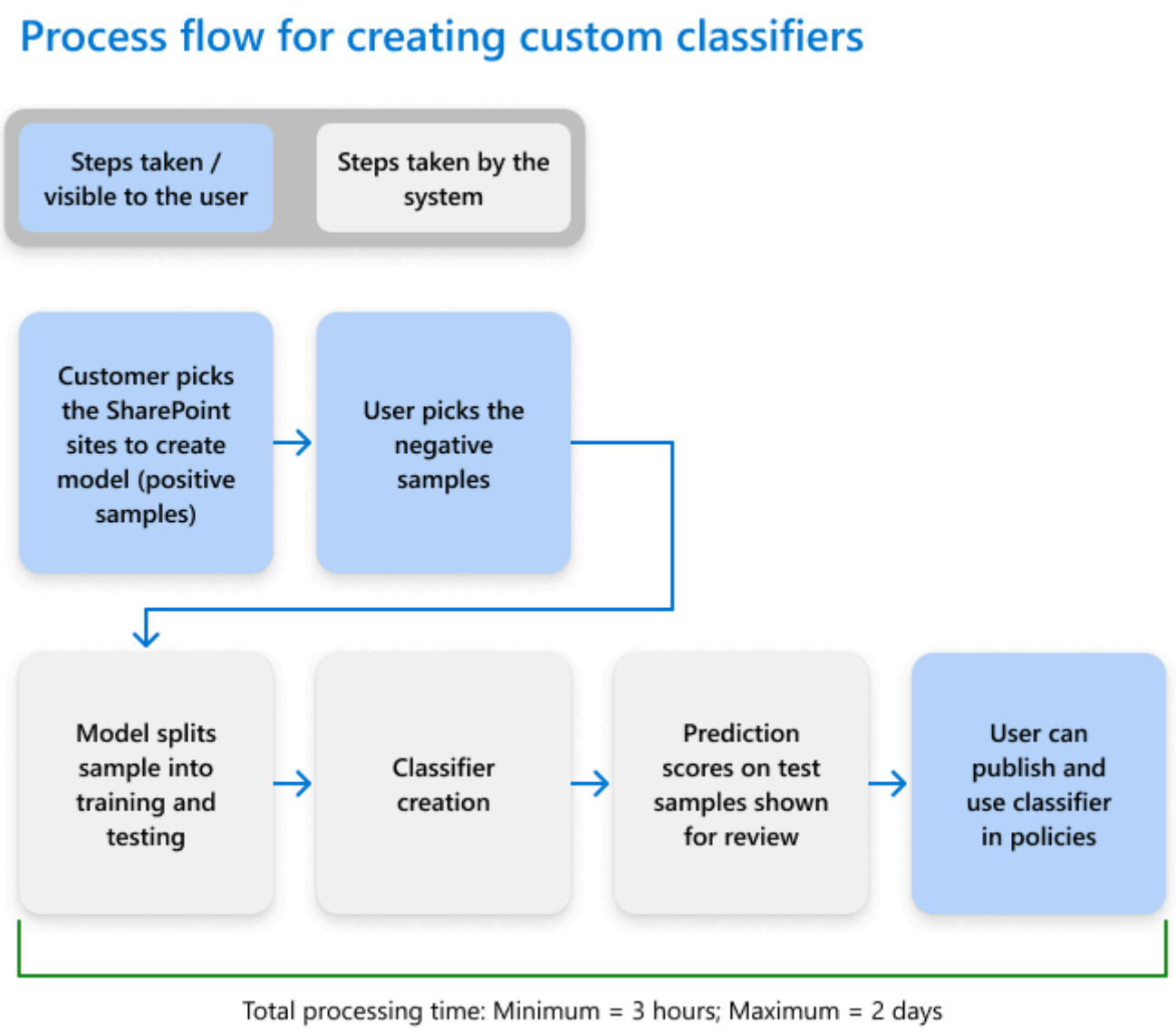
When you publish the classifier, it sorts through items in locations like SharePoint and Exchange, and classifies the content.

For example, you can create trainable classifiers for:

- Legal documents - such as attorney client privilege, closing sets, statement of work
- Strategic business documents - like press releases, merger and acquisition, deals, business or marketing plans, intellectual property, patents, design docs
- Pricing information - like invoices, price quotes, work orders, bidding documents
- Financial information - such as organizational investments, quarterly or annual results

Process flow for creating custom classifiers

The following flow diagram illustrates the process for creating and publishing a classifier for use in compliance solutions, such as retention policies and communication supervision. For more detail on creating a custom trainable classifier, see Get started with trainable classifiers.



> **Tip**
>
> If you create a new SharePoint site and folder for your seed data, allow at least an hour for that location to be indexed before creating the trainable classifier that uses that seed data.

Retraining classifiers

Retraining published custom classifiers isn't supported. If you need to improve the accuracy of a trainable classifier you published, remove the classifier and start over with larger sample sets.

To improve the accuracy of an unpublished classifier, review the test results, update the data set with additional data, and restart the training.

See also

- Retention labels
- Learn about data loss prevention
- Sensitivity labels
- Sensitive information type entity definitions
- Document finger printing
- Learn about exact data match based sensitive information types

---

Last updated on 06/01/2026

Get started with trainable classifiers

09/26/2025

A Microsoft Purview trainable classifier is a tool you can train to recognize various types of content by giving it samples to evaluate. Once trained, you can use it to identify items for application of Office sensitivity labels, Communications compliance policies, and retention label policies.

Implementing a custom trainable classifier requires two steps:

1. Provide two sets of sample data (selected by humans).
   a. A set that contains only items that belong in the category.
   b. A set that contains only items that **do not** belong in the category.
2. Test the classifier's ability to detect matches.

This article explains how to create and test a custom classifier.

For more information about the different types of classifiers, see Learn about trainable classifiers.

> **Important**
>
> Microsoft Purview Communication Compliance supports the use of the Microsoft provided trainable classifiers only. Custom trainable classifiers aren't supported.

Prerequisites

Licensing

For information on licensing, see

- Microsoft 365 Enterprise Plans
- Microsoft 365 Service Descriptions

Permissions

To use classifiers in the following scenarios, you need the following permissions:

Expand table

Scenario	Required Role Permissions
Retention label policy	Record Management Retention Management
Sensitivity label policy	Security Administrator Compliance Administrator Compliance Data Administrator
Communication compliance policy	Insider Risk Management Administrator Supervisory Review Administrator

> **Important**
>
> By default, only the user who creates a custom classifier can train it and review predictions made by that classifier.

Prepare for a custom trainable classifier

Before you create a custom trainable classifier, it's helpful to understand what's involved.

Overall workflow

For more information about the overall workflow of creating custom trainable classifiers, see the process flow for creating custom trainable classifiers.

Seed content

To ensure that your trainable classifier can independently and accurately identify that an item belongs to a particular category of content, you must present it with many samples of the type of content that is in the category. This feeding of samples to the trainable classifier is known as *seeding*. A human must select the seed content, and that content must include two sets of data: one set contains only items that strongly represent the content the classifier is designed to detect (positive samples) and a second set contains items that clearly don't belong (negative samples).

You need at least 50 positive samples (up to 500) and at least 150 negative samples (up to 1,500) to train a classifier. The more samples you provide, the more accurate the predictions the classifier makes will be. The trainable classifier processes up to the 2,000 most recently created samples (by file created date/time stamp).

> **Tip**
>
> For best results, have at least 200 items in your test sample set that includes at least 50 positive examples and at least 150 negative examples.

How to create a trainable classifier

In preview: The following process automates the testing of trainable classifiers and shortens the creation workflow from 12 days to two days. In some cases, the process can take only a few hours.

1. Collect between 50 and 500 seed content items that strongly represent the data you want the classifier to positively identify as being in the category. For a list of supported file types, see Default crawled file name extensions and parsed file types in SharePoint Server.

2. Collect a *second* set of seed content (from 150 to 1,500 items) that represents data that **don't** belong in the category.

3. Place the positive and negative seed content in separate SharePoint folders. Each folder must be dedicated to holding *only the seed content*. Make note of the site, library, and folder URL for each set.

   > **Tip**
   >
   > If you create a new SharePoint site and folder for your seed data, allow at least an hour for that location to be indexed before creating the trainable classifier that uses that seed data.

4. Sign in to the Microsoft Purview portal with either Compliance admin or Security admin role access and navigate to **Data loss prevention** > **Data classification** > **Classifiers**.

> **Important**
>
> The account you use must have access to the seed content folders in SharePoint.

1. Select the **Trainable classifiers** tab.

2. Select **Create trainable classifier**.

3. Add the source of your **positive** examples: select the SharePoint site, library, and folder URL for the seed content that the classifier should detect and then choose **Next**.

4. Add the source of your **negative** examples: select the SharePoint site, library, and folder URL for the seed content that the classifier should ignore and then choose **Next**.

5. Review the settings and select **Create trainable classifier**.

6. Within 24 hours or less, the trainable classifier processes the seed data and builds a prediction model. The classifier status is **In progress** while it processes the seed data. When the classifier finishes processing the seed data, the status changes to **Training is complete and items have been tested**.

7. When training completes and items are (automatically) tested, publish the classifier by choosing **Publish for use**.

After you publish your classifier, it's available as a condition in Office auto-labeling with sensitivity labels, autoapply retention label policy based on a condition, and in Communication compliance.

Test your classifier

After the trainable classifier processes enough positive and negative samples to build a prediction model, test the predictions it makes. When you test the classifier, you verify whether its predictions are correct. After the classifier processes all the data, you can go through the results to verify whether each prediction is correct, incorrect, or uncertain. Microsoft uses this feedback in aggregate to improve the prediction model.

See also

- Learn about data loss prevention
- Sensitivity labels
- Sensitive information type entity definitions
- Document finger printing
- Learn about exact data match based sensitive information types

Increase classifier accuracy

Classifiers, such as sensitive information types (SIT) and trainable classifiers, are used in various types of policies to identify sensitive information. Like most such models, sometimes they identify an item as being sensitive that isn't. Or, they may not identify an item as being sensitive when it actually is. These are called false positives and false negatives.

This article shows you how to confirm whether items matched by a classifier are true positives (a **Match**) or false positives (**Not a match**) and provide **Match**/**Not a match** feedback. You can use that feedback to tune your classifiers to increase accuracy. You can also send redacted versions of the document as well as the **Match**, **Not a Match** feedback to Microsoft if you want to help increase the accuracy of the classifiers that Microsoft provides.

The **Match**, **Not a match** and **Contextual Summary** experiences are available in:

- Data Explorer - for SharePoint sites, OneDrive sites
- Content Explorer - for SharePoint sites, OneDrive sites
- Sensitive Information Type Matched Items page - for SharePoint sites, OneDrive sites
- Trainable Classifier Matched Items page - for SharePoint sites, OneDrive sites
- Microsoft Purview Data Loss Prevention (DLP) Alerts page - for SharePoint sites, OneDrive, and emails in Exchange
- Microsoft Threat Protection (MTP) Alerts page - for SharePoint sites, OneDrive sites, and emails in Exchange

The standalone **Contextual Summary** capability is also available in the following location:

- Microsoft Purview Information Protection (MIP) Auto-labeling simulation matched items - for SharePoint sites, OneDrive sites

Supported classifiers and feedback capabilities

Expand table

Classifier	Contextual summary	Redacted preview panel	Match and Not a Match
SIT	Yes	Yes	Yes
Custom SIT	Yes	No	Yes
Fingerprint SIT	No	No	Yes

Classifier	Contextual summary	Redacted preview panel	Match and Not a Match
Exact data match SIT	No*	No	No
Named entities	No*	No	No
Credential scan	No	No	No
Built-in Trainable classifiers	Yes **	Yes	Yes
Custom trainable classifier	No	No	Yes

* These classifiers are supported in MIP Auto-labeling simulation matched items - for SharePoint sites and OneDrive sites.

** List of Built-in Trainable classifiers and which support contextual summary.

> **Important**
>
> The match/not a match feedback and contextual summary experience support items in: SharePoint sites & OneDrive sites - for Content Explorer, Sensitive Information Type and Trainable Classifier Matched Items, DLP Alerts and MTP Alerts. Emails in Exchange - for DLP Alerts and MTP Alerts. The contextual summary experience supports items in: SharePoint sites and OneDrive sites - for MIP simulation matched items

Licensing requirements for classifier accuracy tuning

For information on licensing, see

- Microsoft 365 Enterprise Plans
- Microsoft 365 Service Descriptions

Known limitations

The following limitations apply to the contextual summary and feedback experience:

- The contextual summary only shows a limited number of matches in any given item, not all matches.
- The contextual summary and feedback experience is only available for items created or updated after the feedback experience was enabled for the tenant. Items that were classified before the feature was enabled may not have the contextual summary and feedback experience available.

How to evaluate match accuracy and provide feedback

The contextual summary experience, where you indicate whether a matched item is a true positive (**Match**) or a false positive (**Not a match**), is similar in Content Explorer, Sensitive Information Type Matched Items, Trainable Classifier Matched Items, DLP Alerts, and MTP Alerts.

> **Important**
>
> You must have already deployed DLP policies that use either SITs or trainable classifiers to OneDrive sites, SharePoint sites, or Exchange mailboxes. You must also have had items match before any items appear in the **Contextual summary** page.

Using Content Explorer

This example shows you how to use the **Contextual Summary** tab to give feedback.

1. Sign in to the Microsoft Purview portal > **Solutions** > **Data Lifecycle Management** > **Explorers** > **Content explorer**.
2. Type the name of the SIT or trainable classifier that you want to check matches for in **Filter on labels, info types, or categories**.
3. Select the SIT.
4. Select the location and make sure that there's a non-zero value in the **Files** column. (The only supported locations are SharePoint and OneDrive.)
5. Open the folder and then select a document.
6. Select the link in the **Sensitive info type** column for the document to see which SITs the item matched and the confidence level.
7. Choose **Close**
8. Open a document and select the **Contextual Summary** tab.
9. Review the item and confirm whether or not it's a match.
10. If it's a match, choose **Close**. You're finished.
11. If it's not a match, choose **Not a match**.
12. If you make a mistake and chose the wrong option, select **Withdraw feedback** next to **Close**. This puts the item back into the **Not a match**/**Match** state.
13. Review the item and redact or un-redact any text.
14. Choose **Close**.

Using the Sensitive Information Type Matched Items page

You can access the same **Match**, **Not a match**, and **Contextual Summary** feedback options on the **Sensitive Info types** page.

1.Sign in to the Microsoft Purview portal > **Solutions** > **Data Lifecycle Management** > **Classifiers** > **Sensitive info types**.

1. In the **Search** field, enter the name of the SIT whose accuracy you want to check.
2. Open the SIT. This brings up **Overview** tab. Here you can see the count of the number of items that match, a count of the number of items that aren't a match, and the number of items with feedback.
3. Select the **Matched items** tab.
4. Open the folder and select a document. Only SharePoint, OneDrive are supported locations here. Make sure that there's a non-zero value in the **Files** column.
5. Select the link in the **Sensitive info type** column for an item to see which SITs the item matched and the confidence level.
6. Choose **Close**.
7. Open a document and then select the **Contextual Summary** tab.
8. Review the item and confirm whether it's a match.
9. If it's a match, choose **Match** and then **Close**.
10. If it isn't a match, choose **Not a Match**.
11. If you make a mistake and select the wrong option, select **Withdraw feedback** next to **Close**. This puts the item back into the **Not a match/Match** state.
12. Choose **Close**.

Using the Trainable Classifier Matched Items page

Use the Trainable Classifier Matched Items page to review matches and provide feedback:

1.Sign in to the Microsoft Purview portal > **Solutions** > **Data Lifecycle Management** > **Classifiers** > **Trainable classifiers**.

1. Select the trainable classifier whose accuracy you want to check.
2. Open the trainable classifier. This brings up **Overview** tab. Here you can see the count of the number of items that match, a count of the number of items that aren't a match, and the number of items with feedback.
3. Select the **Matched items** tab.
4. Open the folder and open a document. Only SharePoint, OneDrive are supported locations here. Make sure that there's a non-zero value in the **Files** column.
5. Open a document and then select the **Contextual Summary** tab.

6. Review the item and confirm whether it's a match.
7. If it's a match, choose **Match** and then choose **Close**.
8. If it isn't a match, choose **Not a Match**.
9. If you make a mistake and select the wrong option, choose **Withdraw feedback** next to **Close**. This puts the item back into the **Not a match/Match** state.
10. Choose **Close**.

Using the Data Loss Prevention Alerts page

Use a DLP alert to review the matched item and submit feedback:

1. Sign in to the Microsoft Purview portal > **Solutions** > **Data loss prevention** > **Alerts**.
2. Choose an alert.
3. Choose **View details**.
4. Choose the **Events** tab.
5. Maximize the **Details** tab.
6. Review the item and confirm whether it's a match.
7. Choose **Actions**.
8. If it's a match, close the window. You're finished.
9. If it's not a match, choose **Actions** and then **Not a match**.
10. Review the item and redact or un-redact any text.
11. Close the window.

Using the feedback to tune your classifiers

If your SITs or trainable classifiers are returning too many false positives based on the feedback, try some of these options to refine them and increase their accuracy.

Tune trainable classifiers with feedback

Retraining custom classifiers is no longer supported. If you need to improve the accuracy of the trainable classifiers you created, remove the classifier and begin fresh with larger sample sets. For more information, see Get started with trainable classifiers.

Tune sensitive information types with feedback

To improve sensitive information type accuracy, consider the following tuning options:

- Increase the thresholds of sensitive information types found to determine severity. It's okay to use different thresholds for individual classifiers.

- Understand confidence levels and how they're defined. Try using a low confidence with high instance count, or a higher confidence level with a low instance count.
- Clone and modify the built-in SITs to include other conditions, such as the presences of keywords, more stringent value matching, or stronger formatting requirements.
- Modify a custom SIT to exclude known prefixes, suffixes, or patterns. For example, a custom SIT to detect phone numbers might trigger for every email if your email signatures or document headers include phone numbers. Excluding your organization's phone number sequences from your custom SIT can prevent the rule from triggering for every email or document.
- Include more dictionary-based SITs as conditions to narrow down the matches to those items that talk about the relevant articles. For example, a rule for matching patient diagnostics may be enhanced by requiring the presence of words like diagnostic, diagnosis, condition, symptom, and patient.
- For named-entity SITs, like **All Full Names**, it's best to set a higher instance count threshold, like 10 or 50. If both the person names and the Social Security Numbers (SSNs) are detected together, it's more likely that the SSNs are truly SSNs, and we reduce the risk that the policy doesn't trigger because too few SSNs are detected.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Trainable classifiers definitions

Microsoft Purview comes with multiple pretrained classifiers. They appear in the **Microsoft Purview portal** > **Data classification** > **Trainable classifiers** view with the status of `Ready to use`.

All Microsoft provided classifiers evaluate items in all the languages listed in the Supported languages columns of the tables below. For more information on how to use trainable classifiers, see Learn about trainable classifiers.

> **Important**
>
> Built-in trainable and global classifiers don't provide an exhaustive or complete list of terms or language across these areas. Further, language and cultural standards continually change, and in light of these realities, Microsoft reserves the right to update these classifiers in its discretion. While classifiers can assist your organization in detecting these areas, classifiers are not intended to provide your organization's sole means of detecting or addressing the use of such language. Your organization, not Microsoft or its subsidiaries, remains responsible for all decisions related to monitoring, scanning, blocking, removal, and retention of any content identified by a pre-trained classifier, including compliance with local privacy and other applicable laws. Microsoft encourages consulting with legal counsel before deployment and use.

Actuary reports

Expand table

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
This classifier is used to identify reports prepared by an organization's actuary on the current and future conditions of funds, such as pension or insurance pools. These documents analyze the organization's loss experience using probability theory and statistical analysis to determine if the fund is on track to meet the needs of its dependents.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .one	English	Yes

Adult, racy, and gory images

Expand table

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects images that are potentially inappropriate. Scanning and detection are supported for Exchange Online email messages, and Microsoft Teams channels and chats.	Detects content in .jpeg, .png, .gif, and .bmp files.	N/A	N/A

> **Note**
>
> Images must be between 100 kilobytes (KB) and 4 megabytes (MB) in size and be greater than 50 x 50 pixels in height x width dimensions.

Agreements

Expand table

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects content related to legal agreements such as nondisclosure agreements, statements of work, loan and lease agreements, employment and noncompete agreements.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .one, .msg, .eml files.	English	Yes

Asset management

Expand table

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Identifies documents related to tracking fixed asset records. For example: asset inventory, long-term funding, life cycle costing, level of service, and criticality within an organization.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .one, .pptx, .pptm, .ppt, .potx, .potm, .pot, .ppsx, .ppsm, .pps, .ppam, .ppa		
English	Yes		

Bank statement

Expand table

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects items that contain a financial transaction of a bank account including account information, deposits, withdrawals, account balance, interest accrued, and bank charges within a given period.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt files.	English	Yes

Budget

Expand table

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects budget documents, budget forecasts and current budget statements including income and expenses of an organization.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .one, .eml, .pptx, .pptm, .ppt, .potx, .potm, .pot, .ppsx, .ppsm, .pps, .ppam, .ppa,	English	Yes

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
	.xlsx, .xlsm, .xlsb, .xls, .csv, .xltx, .xltm, .xlt, .xlam, .xla files.		

Business context

(Preview)

Expand table

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects business-related content such as organizational structure, policy updates, contracts, HR policies, and crucial financial data such as revenue and profits, healthcare forms, employee contracts, etc.	Detects content in docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .one, .pptx, .pptm, .ppt, .potx, .potm, .pot, .ppsx, .ppsm, .pps, .ppam, .ppa, .txt files.	English	N.A.

Business plan

Expand table

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects components of a business plan including business opportunity, plan of achieving the outcomes, market study, and competitor analysis.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .one, .eml, .pptx, .pptm, .ppt, .potx, .potm, .pot, .ppsx, .ppsm, .pps, .ppam, .ppa files.	English	No

Completion certificates

Expand table

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects official documents that are issued at the end of a project or work by a project manager or a contractor. This document is used to testify that work on a particular project has been completed as per a contract or an agreement.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt files.	English	Yes

Compliance policies

Expand table

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Identifies confidential documentation across various major compliance policies, including GDPR, HIPAA, ISO, PCI, SOC, and SSAE 18 compliance.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .txt	English	Yes

Construction specifications

Expand table

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects construction specifications for commercial and industrial projects like factories, plants, commercial offices, airports, roads.	Captures guidelines on the quality, quantity, types of building material, processes, etc. Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .one, .msg, .eml, .pptx, .pptm, .ppt, .potx, .potm, .pot, .ppsx, .ppsm, .pps, .ppam, .ppa files.	English	Yes

Control system and SCADA files

Expand table

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Identifies documents that set the framework for approval, updates, amendments, change tracking, publication (internal or external), and version control. It also identifies Supervisory Control and Data Acquisition (SCADA) files.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .txt	English	Yes

Corporate sabotage

Expand table

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects messages that mention acts to damage or destroy corporate assets or property. This classifier can help customers manage regulatory compliance obligations such as NERC Critical Infrastructure Protection standards or state by state regulations like Chapter 9.05 RCW in Washington state.	Detects content in .msg, .docx, .pdf, .txt, .rtf, .jpeg, .jpg, .png, .gif, .bmp, .svg files.	English	Yes

> **Important**
>
> This classifier can capture a large volume of bulk sender/newsletter content. In Communication Compliance, you can mitigate the detection of large volumes of bulk sender/newsletter content by selecting the **Filter email blasts** check box when you create the policy. You can also edit an existing Communication Compliance policy to turn on this feature.

Credit report

Expand table

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Identifies statements that inform about an individual's or organization's credit activity and current credit situation. This includes loan payment history and the status of credit accounts. It must include a credit report or credit score.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .one, .msg, .eml	English	Yes

Customer complaints

Expand table

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
The customer complaints classifier detects feedback and complaints made about your organization's products or services. This classifier can help you meet regulatory requirements on the detection and triage of complaints, like the Consumer Financial Protection Bureau and Food and Drug Administration requirements.	For Communications Compliance, it detects content in .msg, and .eml files. For the rest of Microsoft Purview Information Protection services, it detects content in .docx, .pdf, .txt, .rtf, .jpg, .jpeg, .png, .gif, .bmp, .svg files.	English	No

Customer files

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Identifies files related to customers, such as personal account files, account information files, grievance records, and feedback files.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .one, .pptx, .pptm, .ppt, .potx, .potm, .pot, .ppsx, .ppsm, .pps, .ppam, .ppa, .txt, .one, .msg, .eml	English	Yes

Discrimination

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects explicit discriminatory language and is sensitive to discriminatory language against the African American/Black communities when compared to other communities.	This applies to communications compliance, it's a text based classifier.	English	Yes

Employee disciplinary action

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects files relating to disciplinary action including a reprimand or corrective action in response to employee misconduct, rule violation, or poor performance.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .one, .msg, .eml files.	English	Yes

Employee insurance

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects documents pertaining to employee medical insurance and workplace disability insurance.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .one, .msg, .eml, .pptx, .pptm, .ppt, .potx, .potm, .pot, .ppsx, .ppsm, .pps, .ppam, .ppa files.	English	Yes

Employment agreement

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects employment agreement containing details like the starting date, salary, compensation, duties of employment.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt files.	English	Yes

Employee pension records

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects documents that are related to employee's pension records such as claim forms, declaration forms, schemes, and benefit statement.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .one, .pptx, .pptm, .ppt, .potx, .potm, .pot, .ppsx, .ppsm, .pps, .ppam, .ppa, .txt, .one, .msg, .eml files.	English	Yes

Employee stocks and financial bond records

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects documents that are related stock and financial bonds award by organization to employees. Identifies employee stocks and financial bonds details that fall under employee's payroll. Contains details like bond clause, allocations, equity.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .one, .pptx, .pptm, .ppt, .potx, .potm, .pot, .ppsx, .ppsm, .pps, .ppam, .ppa, .txt, .one, .msg, .eml files.	English	Yes

Enterprise risk management

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Enterprise risk management includes financial risks, strategic risks, operational risks, and risks associated with accidental losses. This category consists of methods used by organizations to manage risks and seize opportunities related to the achievement of their objectives.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt files.	English	Yes

Environmental permits and clearances

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
This classifier is used to identify documents related to the procedure of obtaining clearance from the government for the installation and modification (amendment) of certain projects.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .one, .pptx, .pptm, .ppt, .potx, .potm, .pot, .ppsx, .ppsm, .pps, .ppam, .ppa, .eml	English	Yes

Facility permits

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Identifies permits that include permission to construct or operate a facility. These include authorizations, licenses, or equivalent control documents issued by the relevant authority.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .txt	English	Yes

Factory incident investigation reports

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
This classifier is used to identify formal recordings of facts related to workplace accidents, injuries, or near misses. These reports aim to uncover the circumstances and conditions that led to the event to prevent future incidents.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .pptx, .pptm, .ppt, .potx, .potm, .pot, .ppsx, .ppsm, .pps, .ppam, .ppa, .txt, .one	English	Yes

Finance

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects content in corporate finance, accounting, economy, banking, and investment categories.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .one, .msg, .eml, .pptx, .pptm, .ppt, .potx, .potm, .pot, .ppsx, .ppsm, .pps, .ppam, .ppa, .xlsx, .xlsm, .xlsb, .xls, .csv, .xltx, .xltm, .xlt, .xlam, .xla files.	English	Yes

Financial audit

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects files, documents, and reports pertaining to financial audit, both external or internal audit undertaken in an organization.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .one, .msg, .eml files.	English	Yes

Financial statement

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects financial statements like income statement, balance sheet, cash flow statement, statement of changes in equity.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .xlsx, .xlsm, .xlsb, .xls, .csv, .xltx, .xltm, .xlt, .xlam, .xla files.	English	Yes

Freight documents

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects documents that authorize the export or import of a good in a specific quantity from source to destination. This model categorizes different documents including Bill of Ladings, Certificate of Origin, Commercial Invoice, Export import customs declaration, Importer Security Filing (ISF).	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .one, .pptx, .pptm, .ppt, .potx, .potm, .pot, .ppsx, .ppsm, .pps, .ppam, .ppa, .txt, .one files.	English	Yes

Garnishment

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Identifies documents related to garnishment notices, records, and earnings attachments. It covers the legal process of withholding money from paychecks to be sent to another party.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .one	English	Yes

Gifts & entertainment

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects messages that suggest exchanging gifts or entertainment in return for service, which violates regulations related to bribery. This classifier can help customers manage regulatory compliance obligations such as Foreign Corrupt Practices Act (FCPA), UK Bribery Act, and FINRA Rule 2320.	Detects content in .msg, .docx, .pdf, .txt, .rtf, .jpeg, .jpg, .png, .gif, .bmp, .svg files.	English	Yes

> **Important**
>
> This classifier can capture a large volume of bulk sender/newsletter content. In Communication Compliance, you can mitigate the detection of large volumes of bulk sender/newsletter content by selecting the **Filter email blasts** check box when you create the policy. You can also edit an existing Communication Compliance policy to turn on this feature.

Harassment

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects a specific category of offensive language text items related to offensive conduct targeting one or multiple individuals based on the following traits: race, ethnicity, religion, national origin, gender, sexual orientation, age, disability.	Detects content in .msg, .docx, .pdf, .txt, .rtf, .jpeg, .jpg, .png, .gif, .bmp, .svg files.	- Arabic - Chinese (Simplified) - Chinese (Traditional) - Dutch - English - French - German - Italian - Korean - Japanese - Portuguese - Spanish	Yes (English)

Health/Medical forms

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects various forms and files that are used for systematic documentation of a patient's admission details, medical history, patient information, and prior authorization request and are typically used in medical/health services.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .one, .msg, .eml, .pptx, .pptm, .ppt, .potx, .potm, .pot, .ppsx, .ppsm, .pps, .ppam, .ppa files.	English	Yes

Healthcare

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects content in medical and healthcare administration aspects such as medical services, diagnoses, treatment, claims, etc.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .one, .msg, .eml, .pptx, .pptm, .ppt, .potx, .potm, .pot, .ppsx, .ppsm, .pps, .ppam, .ppa, .xlsx, .xlsm, .xlsb, .xls, .csv, .xltx, .xltm, .xlt, .xlam, .xla files.	English	Yes

Human resources

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects content in human resources related categories of recruitment, interviewing, hiring, training, evaluating, warning, and termination.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .one, .msg, .eml, .pptx, .pptm, .ppt, .potx, .potm, .pot, .ppsx, .ppsm, .pps, .ppam, .ppa, .xlsx, .xlsm, .xlsb, .xls, .csv, .xltx, .xltm, .xlt, .xlam, .xla files.	English	Yes

Invoice

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects invoices containing an itemized summary of the purchase, the total balance owed, current payment due, and various payment methods.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .one, .eml, .xlsx, .xlsm, .xlsb, .xls, .csv, .xltx, .xltm, .xlt, .xlam, .xla files.	English	Yes

Intellectual property

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects content in intellectual property related categories such as trade secrets and similar confidential information.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .one, .msg, .eml, .pptx, .pptm, .ppt, .potx, .potm, .pot, .ppsx, .ppsm, .pps, .ppam, .ppa, .xlsx, .xlsm, .xlsb, .xls, .csv, .xltx, .xltm, .xlt, .xlam, .xla files.	English	Yes

Information technology

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects content in information technology and cybersecurity categories such as network settings, information security, hardware, and software.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .one, .msg, .eml, .pptx, .pptm, .ppt, .potx, .potm, .pot, .ppsx, .ppsm, .pps, .ppam, .ppa, .xlsx, .xlsm, .xlsb, .xls, .csv, .xltx, .xltm, .xlt, .xlam, .xla files.	English	Yes

IT infra and network security documents

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Identifies documents that outline rules for computer network access, policy enforcement, and basic architecture of a company's security or network security environment. It also covers documents related to access control, data security, storage policy, and maintenance plans.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .one, .pptx, .pptm, .ppt, .potx, .potm, .pot, .ppsx, .ppsm, .pps, .ppam, .ppa		
English	Yes		

Lease deeds

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
This classifier is used to identify documents detailing lease information such as period, amount, and agreements. It includes facility, factory, and real estate lease deeds, as well as responsibilities of the lessor and lessee, deposits, due dates, and consequences of lease violations.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .one, .pptx, .pptm, .ppt, .potx, .potm, .pot, .ppsx, .ppsm, .pps, .ppam, .ppa	English	Yes

Legal affairs

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects content in legal affairs-related categories such as litigation, legal process, legal obligation, legal terminology, law, and legislation.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .one, .msg, .eml files.	English	Yes

Legal agreements

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects various legally binding documents/ contracts/ agreements like Arbitration agreements, Power of Attorney, Purchase Agreements between two parties.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .txt files.	English	Yes

Letters of credit

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Identifies documents sent from banks or financial institutions that guarantee a seller will receive a buyer's payment on time and for the full amount.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .pptx, .pptm, .ppt, .potx, .potm, .pot, .ppsx, .ppsm, .pps, .ppam, .ppa, .txt, .one	English	Yes

License agreement

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects license agreements, contains terms and conditions for use and compensation for the licensor.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt files.	English	Yes

Loan agreements and offer letters

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects loan agreements, offer letters, and terms and conditions contained within the document.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .one, .msg, .eml files.	English	Yes

Manufacturing batch records

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects manufacturing batch documents that include details around the entire manufacturing process and the history of a product batch.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .one, .msg, .eml files.	English	Yes

Marketing collaterals

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
This classifier is used to identify various marketing materials utilized by organizations for outreach. This includes battle cards, fact sheets, product brochures, product launch materials, press releases, and product catalogs.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .txt	English	Yes

Merger and acquisition files

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects documents including letter of intent, term sheets, and related files.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .one, .msg, .eml files.	English	Yes

Meeting notes

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects documents and notes containing information specific to meetings.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .one, .msg, .eml, .pptx, .pptm, .ppt, .potx, .potm, .pot, .ppsx, .ppsm, .pps, .ppam, .ppa files.	English	Yes

Money laundering

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects signs that suggest money laundering or engagement in acts to conceal or disguise the origin or destination of proceeds. This classifier helps customers manage regulatory compliance obligations such as the Bank Secrecy Act, the USA Patriot Act, FINRA Rule 3310, and Anti-Money Laundering Act of 2020.	Detects content in .msg, .docx, .pdf, .txt, .rtf, .jpeg, .jpg, .png, .gif, .bmp, .svg files.	English	Yes

> **Important**
>
> This classifier can capture a large volume of bulk sender/newsletter content. In Communication Compliance, you can mitigate the detection of large volumes of bulk sender/newsletter content by selecting the **Filter email blasts** check box when you create the policy. You can also edit an existing Communication Compliance policy to turn on this feature.

MoU files (Memorandum of understanding)

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Identifies documents related to memoranda of understanding, annexures, corrigenda, and addenda. These documents are typically not legally binding.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .one, .pptx, .pptm, .ppt, .potx, .potm, .pot, .ppsx, .ppsm, .pps, .ppam, .ppa	English	Yes

Network design files

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects technical documentation about networks of computers. This includes various components of network, how they're connected, their architecture, how they perform, and where they troubleshoot.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .one, .msg, .eml, .pptx, .pptm, .ppt, .potx, .potm, .pot, .ppsx, .ppsm, .pps, .ppam, .ppa files.	English	Yes

Non-disclosure agreement

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects nondisclosure agreements (NDAs).	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt files.	English	Yes

OSHA records

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Identifies Occupational Safety and Health Administration (OSHA) recordkeeping forms. This includes the Log of Work-Related Injuries and Illnesses (OSHA Form 300), Summary of Work-Related Injuries and Illnesses (OSHA Form 300A), and the Injury and Illness Incident Report (OSHA Form 301).	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .pptx, .pptm, .ppt, .potx, .potm, .pot, .ppsx, .ppsm, .pps, .ppam, .ppa, .txt, .one	English	Yes

Paystub

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects paystub/salary statement files.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .xlsx, .xlsm, .xlsb, .xls, .csv, .xltx, .xltm, .xlt, .xlam, .xla files.	English	Yes

Personal financial information

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects documents related to different personal financial records consisting of financial statements, real estate planning, and retirement plans. Consists of details of all assets and liabilities held by an individual.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .one, .txt, .one files.	English	Yes

Procurement

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects content in categories of bidding, quoting, purchasing, and paying for supply of goods and services.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .one, .msg, .eml, .xlsx, .xlsm, .xlsb, .xls, .csv, .xltx, .xltm, .xlt, .xlam, .xla files.	English	No

Project documents

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects project reports and documents, which include project planning documents, project charter documents and schedules.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .one, .msg, .eml, .pptx, .pptm, .ppt, .potx, .potm, .pot, .ppsx, .ppsm, .pps, .ppam, .ppa files.	English	Yes

Profanity

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects a specific category of offensive language text items that contain expressions that embarrass most people.	Detects content in .msg, .docx, .pdf, .txt, .rtf, .jpeg, .jpg, .png, .gif, .bmp, .svg files.	- Arabic - Chinese (Simplified) - Chinese (Traditional) - Dutch - English - French - German - Italian - Korean - Japanese - Portuguese - Spanish	Yes (English)

Prompt Shields

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects adversarial user input attacks such as user prompt injection attacks (jailbreak). User prompt injection attacks will deliberately exploit system vulnerabilities to elicit unauthorized behavior from the LLM.	Generative AI prompts	English	No

> **Note**
>
> This classifier can be configured in Communication Compliance and evaluates generative AI prompts ONLY

Protected material

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects known text content that may be protected under copyright or branding laws. Detecting the display of protected material in generative AI responses ensures compliance with intellectual property laws and maintains content originality	Generative AI responses	English	No

> **Note**
>
> This classifier can be configured in Communication Compliance and evaluates generative AI responses ONLY

Quotation

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects documents that offer to sell goods or services for a set price, based on certain conditions. It contains a description of the goods or services, the price of the goods or rate of the service, the quantity, and a total cost.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .one, .eml, .xlsx, .xlsm, .xlsb, .xls, .csv, .xltx, .xltm, .xlt, .xlam, .xla files.	English	Yes

Regulatory collusion

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects messages that can violate regulatory anti-collusion requirements such as an attempted concealment of sensitive information. This classifier can help customers manage regulatory compliance obligations such as the Sherman Antitrust Act, Securities Exchange Act 1933, Securities Exchange Act of 1934, Investment Advisers Act of 1940, Federal Commission Act, and Robinson-Patman Act.	Detects content in .msg, .docx, .pdf, .txt, .rtf, .jpeg, .jpg, .png, .gif, .bmp, .svg files.	English	No

> **Important**
>
> This classifier can capture a large volume of bulk sender/newsletter content. In Communication Compliance, you can mitigate the detection of large volumes of bulk sender/newsletter content by selecting the **Filter email blasts** check box when you create the policy. You can also edit an existing Communication Compliance policy to turn on this feature.

Resume

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects a resume document that a job applicant provides an employer, which has a detailed statement of the candidate's prior work experience, education, and accomplishments.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .txt files.	English	Yes

Safety records

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects documents that are related to facility/factory safety. These documents can be facility safety plan, safety assessments and audit reports, emergency response and evacuation plan, and equipment's inspection reports concerning safety measurements.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .one, .pptx, .pptm, .ppt, .potx, .potm, .pot, .ppsx, .ppsm, .pps, .ppam, .ppa, .txt, .one, .eml files.	English	Yes

Sales and revenue

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects sales reports, revenue/income statement and sales/demand forecasting reports for organizations.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .one, .pptx, .pptm, .ppt, .potx, .potm, .pot, .ppsx, .ppsm, .pps, .ppam, .ppa files.	English	Yes

Software product development files

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects files used in software development; including product requirements document, product testing and planning, files including test cases, and test reports.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .one, .msg, .eml files.	English	Yes

Source code

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects items that contain a set of instructions and statements written computer programming languages on GitHub: ActionScript, C, C#, C++, Clojure, CoffeeScript, Go, Haskell, Java, JavaScript, Lua, MATLAB, Objective-C, Perl, PHP, Python, R, Ruby, Scala, Shell, Swift, TeX, Vim Script.	Detects content in .c, .h, .w, .cs, .cake, .csx, .cpp, .c++, .cc, .cp, .cxx, .hh, .hpp, .hxx, .java, .js, .m, .matlab, .pl, .perl, .pm, .prl, .ipb, .php, .php3, .php4, .php5, .py, .pyc, .pyo, .r, .rl, .rb, .irb, .swift, .as, .clj, .cljs, .cljc, .coffee, .Go, .hs, .hsc, .lua, .lub, .m, .mm, .scala, .sca, .Tex,T, .xs, . sh, .vim, .edn, .javac, .lhs, .mjs, .pod, .r, .rda, .RData, .rds, .rb, .bash, .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .one, .eml, .msg, .pptx, .pptm, .ppt, .potx, .potm, .pot, .ppsx, .ppsm, .pps, .ppam, .ppa, .xlsx, .xlsm, .xlsb, .xls, .csv, .xltx, .xltm, .xlt, .xlam, .xla, .sc, .litcoffee files.	N/A	No

> **Note**
>
> Source code is trained to detect when the bulk of the text is source code. It doesn't detect source code text that is interspersed with plain text.

Standard operating procedures and manuals

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects sets of documented instructions created to help workers perform routine operations or manufacturing tasks.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt files.	English	Yes

Statement of accounts

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
A statement of account is a detailed report of the contents of an account. Identifies documents related to statement of accounts, accounts payable, and accounts receivable.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .xlsx, .xlsm, .xlsb, .xls, .csv, .xltx, .xltm, .xlt, .xlam, .xla files.	English	Yes

Statement of work

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects statement of work (SOW) containing details like requirements, responsibilities, terms and conditions for both parties.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt files.	English	Yes

Stock manipulation

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects signs of possible stock manipulation, such as recommendations to buy, sell, or hold stocks that can suggest an attempt to manipulate the stock price. This classifier can help customers manage regulatory	Detects content in .msg, .docx, .pdf, .txt, .rtf, .jpeg, .jpg, .png,	English	Yes

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
compliance obligations such as the Securities Exchange Act of 1934, FINRA Rule 2372, and FINRA Rule 5270.	.gif, .bmp, .svg files.		

> **Important**
>
> This classifier can capture a large volume of bulk sender/newsletter content. In Communication Compliance, you can mitigate the detection of large volumes of bulk sender/newsletter content by selecting the **Filter email blasts** check box when you create the policy. You can also edit an existing Communication Compliance policy to turn on this feature.

Tax documents

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects tax related content such as tax planning, tax forms, tax filing, tax regulations.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt, .one, .msg, .eml, .pptx, .pptm, .ppt, .potx, .potm, .pot, .ppsx, .ppsm, .pps, .ppam, .ppa, .xlsx, .xlsm, .xlsb, .xls, .csv, .xltx, .xltm, .xlt, .xlam, xla files.	English	Yes

Threat

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects a specific category of offensive language text items related to threats to commit violence or do physical harm or damage to a person or property.	Detects content in .msg, .docx, .pdf, .txt, .rtf, .jpeg, .jpg, .png, .gif, .bmp, .svg files.	- Arabic - Chinese (Simplified) - Chinese	Yes (English)

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
		(Traditional) - Dutch - English - French - German - Italian - Korean - Japanese - Portuguese - Spanish	

Unauthorized disclosure

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Detects sharing of information containing content that is explicitly designated as confidential or internal to unauthorized individuals. This classifier can help customers manage regulatory compliance obligations such as FINRA Rule 2010 and SEC Rule 10b-5.	Detects content in .msg, .docx, .pdf, .txt, .rtf, .jpeg, .jpg, .png, .gif, .bmp, .svg files.	English	Yes

> **Important**
>
> This classifier can capture a large volume of bulk sender/newsletter content. In Communication Compliance, you can mitigate the detection of large volumes of bulk sender/newsletter content by selecting the **Filter email blasts** check box when you create the policy. You can also edit an existing Communication Compliance policy to turn on this feature.

Wire transfer

Description	File types	Supported Languages	Contextual summary and keyword highlighting summary
Wire transfer is a method of electronic funds transfer from one person or entity to another. The model captures all the wire transfer receipts and acknowledgments.	Detects content in .docx, .docm, .doc, .dotx, .dotm, .dot, .pdf, .rtf, .txt files.	English	Yes

Word count requirements

Pretrained classifiers require documents to contain at least 200 words for classification. Documents with fewer than 200 words aren't classified by these classifiers. This minimum doesn't apply to custom trainable classifiers during training or classification.

Communication compliance classifiers have separate minimum word count requirements for messages, as listed in the following table. To identify and take action on messages below the listed minimums, create a custom keyword dictionary for communication compliance policies that detect inappropriate language.

Classifier	Minimum word count	Language
Threat, Harassment, and Profanity	Six words	- Dutch - French - German - Italian - Japanese - Portuguese - Spanish
Threat, Harassment, and Profanity	12 words	- Arabic - Chinese Simplified - Chinese Traditional - Korean
Threat and Harassment	Three words	English
Profanity	Five words	English
Corporate sabotage - Customer complaints - Gifts & entertainment - Money laundering - Regulatory collusion	Six words	English

Classifier	Minimum word count	Language
- Stock manipulation - Unauthorized disclosure		

See also

- Retention labels
- Learn about data loss prevention
- Sensitivity labels
- Sensitive information type entity definitions
- Document finger printing
- Learn about exact data match based sensitive information types

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 10/12/2026

All credentials sensitive information types

"All credentials" is a bundled-entity sensitive information type (SIT). It detects credentials from all supported services and environments, which include Amazon, Azure, GitHub, Google, Microsoft general, Slack and more.

> **Important**
>
> Because the **All Credetials** SIT is a bundled SIT, it ***does not*** contain any keywords. The keywords that it detects are based on the individual SITs it contains. Refer to the individual SITs for keyword information.

For information about implementing bundled SITs, see Learn about named entities.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a bundled SIT in an endpoint DLP policy. This requirement is specific to the combination of bundled SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

> **Important**
>
> You must enable Advanced classification scanning and protection if you want to use a credential scanning SIT in an endpoint DLP policy. This requirement is specific to the combination of credential scanning SITS and endpoint DLP policies. The advanced classification article gives you information about enabling advanced classification, and about configuring endpoints and the supported file types.

Format

Various

Pattern

Various

Checksum

No

Keyword Highlighting

Not supported

When keyword highlighting is supported in the contextual summary for a sensitive information type or a trainable classifier, in the **Contextual Summary** view of activity explorer, the keywords in a document that were matched to a policy are highlighted.

Description

The All credentials SIT is a bundled SIT that scans for individual credential types. This includes checksum evaluated values and keyword/pattern combinations. The All credential entity scans for any of the underlying credential information type checksums and/or patterns are detected in a document.

Contains

This bundled entity SIT contains these individual SITs:

- Amazon S3 Client Secret Access Key
- ASP.NET Machine Key
- Microsoft Entra Client Access Token
- Microsoft Entra Client Secret
- Microsoft Entra user Credentials
- Azure App Service Deployment Password
- Azure Batch Shared Access Key
- Azure Bot Framework Secret Key
- Azure Bot Service App Secret
- Azure Cognitive Search API Key
- Azure Cognitive Service Key
- Azure Container Registry Access Key
- Azure COSMOS DB Account Access Key
- Azure Databricks Personal Access Token

- Azure DevOps App Secret
- Azure DevOps Personal Access Token
- Azure EventGrid Access Key
- Azure Function Master / API Key
- Azure IoT Shared Access Key
- Azure Logic App Shared Access Signature
- Azure Machine Learning Web Service API Key
- Azure Maps Subscription Key
- Azure Redis Cache Connection String Password
- Azure Service Bus Shared Access Signature
- Azure Shared Access Key / Web Hook Token
- Azure SignalR Access Key
- Azure SQL Connection String
- Azure Storage Account Access Key
- Azure Storage Account Shared Access Signature
- Azure Storage Account Shared Access Signature for High Risk Resources
- Azure Subscription Management Certificate
- Client Secret / API Key
- General Password

> **Tip**
>
> General Password detects password related keywords in English only.

- General Symmetric Key
- GitHub Personal Access Token
- Google API key
- Http Authorization Header
- Microsoft Bing Maps Key
- Slack Access Token
- User Login Credentials
- X.509 Certificate Private Key

Supported languages

- English
- Bulgarian

- Chinese
- Croatian
- Czech
- Danish
- Estonian
- Finnish
- French
- German
- Hungarian
- Icelandic
- Irish
- Italian
- Japanese
- Latvian
- Lithuanian
- Maltese
- Dutch
- Norwegian
- Polish
- Portuguese
- Romanian
- Slovak
- Slovenian
- Spanish
- Swedish
- Turkish

Last updated on 06/15/2026

Document fingerprinting

Article • 03/31/2025

Document fingerprinting is a Microsoft Purview feature that takes a standard form that you provide and creates a sensitive information type (SIT) based on that form. Document fingerprinting makes it easier for you to protect sensitive information by identifying standard forms that are used throughout your organization. This article describes the concepts behind document fingerprinting and how to create a document fingerprint using the user interface or using PowerShell.

Document fingerprinting includes the following benefits:

- SITs created from document fingerprinting can be used as a detection method in DLP policies scoped to Exchange, SharePoint, OneDrive, Teams, and Devices.
- MIP auto-labeling can use document fingerprinting as a detection method in Exchange, SharePoint, and OneDrive.
- Document fingerprint features can be managed through the Microsoft Purview user interface.
- Partial matching is supported.
- Exact matching is supported.
- Improved detection accuracy
- Support for detection in multiple languages, including dual-byte languages such as Chinese, Japanese, and Korean.

> **Important**
>
> If you are an E5 customer, we recommend updating your existing fingerprints to take advantage of the full document fingerprint feature set. If you are an E3 customer, we recommend upgrading to an E5 license. If you choose not to, you won't be able to modify existing fingerprints or create new ones after April, 2023.

Basic scenario for document fingerprinting

As mentioned, the document fingerprinting feature converts a standard form of information into a sensitive information type (SIT), which you can use in the rules of your DLP policies. For example, you can create a document fingerprint based on a blank patent template and then create a DLP policy that detects and blocks all outgoing patent templates with sensitive content filled in. Optionally, you can set up policy tips to notify senders that they might be sending sensitive information, and that the sender

should verify that the recipients are qualified to receive the patents. This process works with any text-based forms used in your organization. Other examples of forms that you can upload include:

- Government forms
- Health Insurance Portability and Accountability Act (HIPAA) compliance forms
- Employee information forms for Human Resources departments
- Custom forms created specifically for your organization

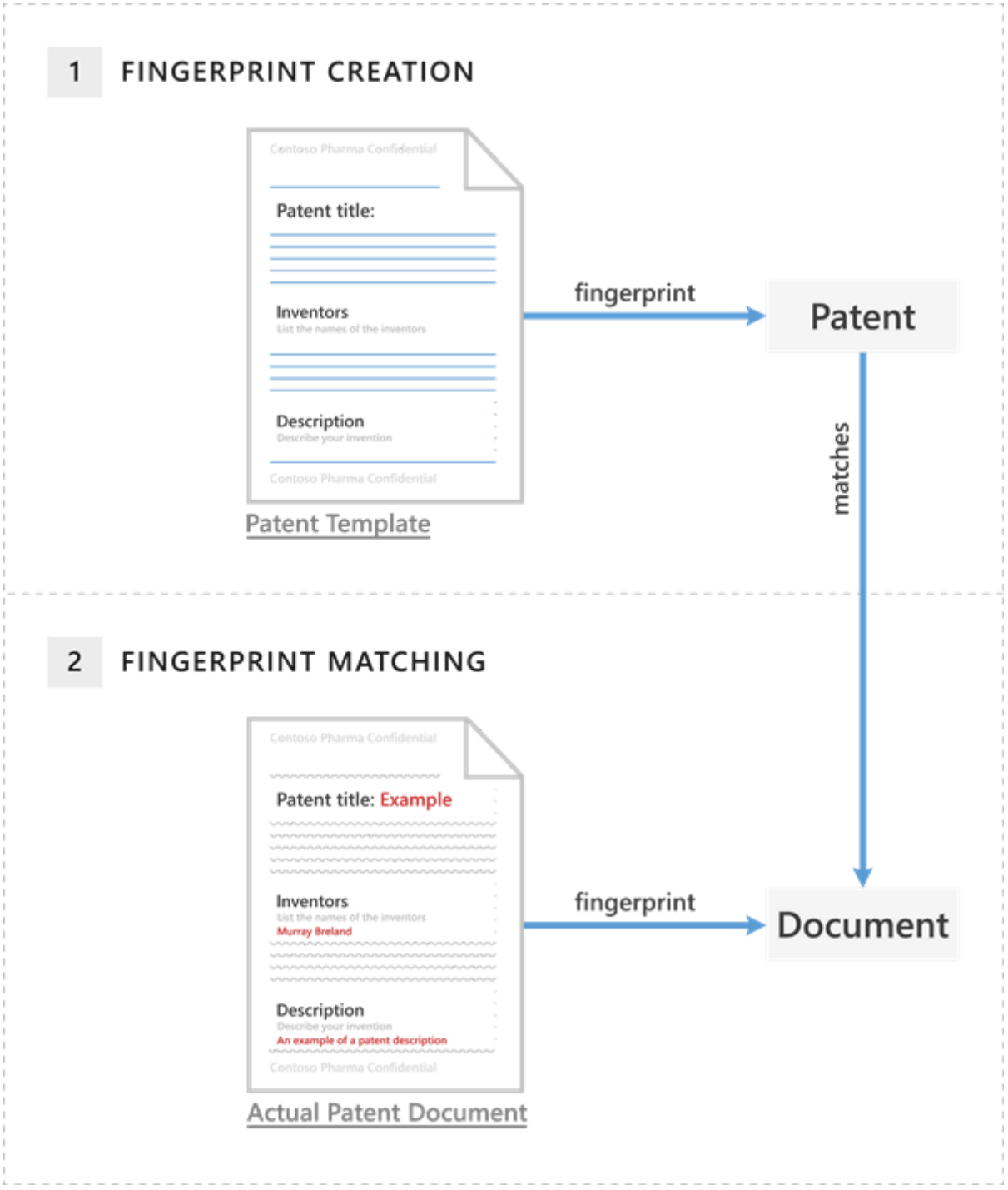
Ideally, your organization already has an established business practice of using certain forms to transmit sensitive information. To enable detection, upload an empty form to be converted to a document fingerprint. Next, set up a corresponding policy. Once you complete these steps, DLP detects any documents in outbound mail that match that fingerprint.

For more information on designing a DLP policy, see Design a data loss prevention policy.

For more informaiton on creating and deploying a DLP policy, see Create and Deploy data loss prevention policies.

How document fingerprinting works

You know documents don't have actual fingerprints, but the name helps explain the feature. In the same way a person's fingerprints have unique patterns, frequently used forms (templates) can have patterns of words that are unique to them. You can use the SIT that's based on this pattern to detect files that were created using the same template. This is why uploading a form or template creates the most effective type of document fingerprint. Everyone who fills out a form uses the same original set of words and then adds their own words to the document. Documents to be scanned can't be password protected and must contain all the text from the original form.



The patent template contains the blank fields **Patent title**, **Inventors**, and **Description**, along with descriptions for each of those fields — that's the word pattern. When you upload the original patent template, it's in one of the supported file types and in plain text. MIcrosoft Purview converts this word pattern into a document fingerprint, which is a small Unicode XML file containing a unique hash value that represents the original text. As a security measure, the original document itself isn't stored; only the hash value is stored. The original document can't be reconstructed from the hash value. The patent fingerprint is represented in a SIT that you can use as a condition in a DLP policy.

For example, if you set up a DLP policy that prevents regular employees from sending outgoing messages containing patents, DLP uses the patent fingerprint SIT to detect patents and block those emails. Alternatively, you might want to let your legal

department be able to send patents to other organizations because it has a business need for doing so. To allow specific departments to send sensitive information, create exceptions for those departments in your DLP policy. Alternatively, you can allow them to override a policy tip with a business justification.

> **Important**
>
> Text in embedded documents is not considered for fingerprint creation. You need to provide sample template files that don't contain embedded documents.

Limitations of document fingerprinting

Document fingerprinting doesn't detect sensitive information in the following cases:

- Password protected files
- Files that contain images only
- Documents that don't contain all the text from the original form used to create the document fingerprint
- Files larger than 4 MB

> **Note**
>
> To use document fingerprinting with devices, Advanced classification scanning and protection must be turned on.

Fingerprints are stored in a separate rule pack. This rule pack has a maximum size limit of 300 KB. Given this limit, you can create approximately 100 fingerprints per tenant.

> **Note**
>
> The template used to create a fingerprint should have at least 256 characters. The supported extracted text length for the fingerprint template must be between 256 and 204,800 characters.

The following examples show what happens if you create a document fingerprint based on a patent template. However, you can use any form as a basis for creating a document fingerprint.

Example: Create a patent document that matches the document fingerprint of a patent template

1. In the Microsoft Purview portal, navigate to **Data Loss Prevention** or **Information Protection** > **Classifiers** > **Sensitive info types**.
2. On the **Sensitive info types** page, choose **+ Create Fingerprint based SIT**.
3. Enter a name and description for your new SIT.
4. Upload the file you wish to use as the fingerprint template.
5. OPTIONAL: Adjust the requirements for each confidence level. (For more information, see Partial matching and Exact matching.)
6. Choose **Next**.
7. Review your settings and then choose **Create**.
8. When the confirmation page displays, choose **Done**.

PowerShell example of a patent document matching a document fingerprint of a patent template

PowerShell

```
>> $Patent_Form = ([System.IO.File]::ReadAllBytes('C:\My
Documents\patent.docx'))

>> New-DlpSensitiveInformationType -Name "Patent SIT" -FileData $Patent_Form
-ThresholdConfig @{low=40;medium=60;high=80} -IsExact $false -Description
"Contoso Patent Template"
```

Partial matching

To configure partial matching of a document fingerprint, when you are setting the configuration options during template upload, set the confidence level, choose **Low**, **Medium**, or **High**, and designate how much of the text in the file must match the fingerprint in terms of a percentage between 30% - 90%.

A high confidence level returns the fewest false positives but might result in more false negatives. Low or medium confidence levels return more false positives but few to zero false negatives.

- **low confidence**: Matched items contain the fewest false negatives but the most false positives. Low confidence returns all low, medium, and high confidence matches.
- **medium confidence**: Matched items contain an average number of false positives and false negatives. Medium confidence returns all medium, and high confidence

matches.

- **high confidence**: Matched items contain the fewest false positives but the most false negatives.

Exact matching

To configure exact matching of a document fingerprint, select **Exact** as the value for the high confidence level. When you set the high confidence level to **Exact**, only files that have exactly the same text as the fingerprint are detected. If the file has even a small deviation from the fingerprint, it won't be detected.

Already using fingerprint SITs?

Your existing fingerprints and policies/rules for those fingerprints should continue to work. If you don't want to use the latest fingerprint features, you don't have to do anything.

If you have an E5 license and want to use the latest fingerprint features, you have 2 choices:

- Create a new fingerprint.
- migrate a policy to the newer version.

> ⓘ **Note**
>
> Creating new fingerprints using the templates on which a fingerprint already exists is not supported.

Create a custom sensitive information type based on document fingerprinting using PowerShell

Currently, you can create a document fingerprint only in Security & Compliance PowerShell.

To create a custom SIT based on a document fingerprint, use the **New-DlpSensitiveInformationType** cmdlet. The following example creates a new document fingerprint named “Contoso Customer Confidential” based on the file C:\My Documents\Contoso Customer Form.docx.

PowerShell

```
$Employee_Form = ([System.IO.File]::ReadAllBytes('C:\My Documents\Contoso
Customer Form.docx'))

New-DlpSensitiveInformationType -Name "Contoso Customer Confidential" -
FileData $Employee_Form -ThresholdConfig @{low=40;medium=60;high=80} -
IsExact $false -Description "Message contains Contoso customer information."
```

Finally, add the "Contoso Customer Confidential" sensitive information type to a DLP policy in the Microsoft Purview portal. This example adds a rule to an existing DLP policy, named "ConfidentialPolicy".

PowerShell

```
New-DlpComplianceRule -Name "ContosoConfidentialRule" -Policy
"ConfidentialPolicy" -ContentContainsSensitiveInformation @{Name="Contoso
Customer Confidential"} -BlockAccess $True
```

You can also use the Fingerprint SIT in mail flow rules in Exchange, as shown in the following example. To run this command, you first need to connect to Exchange PowerShell. Also, note that it takes time for the SITs to sync with the Exchange admin center.

PowerShell

```
New-TransportRule -Name "Notify :External Recipient Contoso confidential" -
NotifySender NotifyOnly -Mode Enforce -SentToScope NotInOrganization -
MessageContainsDataClassification @{Name=" Contoso Customer Confidential"}
```

DLP can now detect documents that match the Contoso Customer Form.docx document fingerprint.

For syntax and parameter information, see:

- New-DlpFingerprint
- New-DlpSensitiveInformationType
- Remove-DlpSensitiveInformationType
- Set-DlpSensitiveInformationType
- Get-DlpSensitiveInformationType

Edit, test, or delete a document fingerprint

To do this in the Microsoft Purview portal, open the fingerprint SIT you want to edit, test, or delete and choose the appropriate icon.

To do this via PowerShell, run the following commands:

Edit a document fingerprint

```
PowerShell

>> Set-DlpSensitiveInformationType -Name "Fingerprint SIT" -FileData
([System.IO.File]::ReadAllBytes('C:\My Documents\file1.docx')) -
ThresholdConfig @{low=30;medium=50;high=80} -IsExact $false-Description "A
friendly Description"
```

Test a document fingerprint

```
PowerShell

>> $r = Test-DataClassification -TextToClassify "Credit card information
Visa: 4485 3647 3952 7352. Patient Identifier or SSN: 452-12-1232"
>> $r.ClassificationResults
```

Delete a document fingerprint

```
PowerShell

>> Remove-DlpSensitiveInformationType "Fingerprint SIT"
```

Migrate an existing fingerprint SIT to a via the Microsoft Purview Portal

1. Open the Microsoft Purview portal > **Information Protection** > **Classifiers** > **Sensitive info types**.
2. Open the SIT containing the fingerprint that you want to migrate.
3. Choose **Edit**.
4. Upload the same fingerprint file again.
5. Review the fingerprint settings > **Done**.

Migrate a fingerprint using PowerShell

Enter the following command:

```PowerShell
Set-DlpSensitiveInformationType -Name "Old Fingerprint" -FileData
([System.IO.File]::ReadAllBytes('C:\My Documents\file1.docx')) -
ThresholdConfig @{low=30;medium=50;high=80} -IsExact $false-Description "A
friendly Description"
```

Feedback

Was this page helpful? Yes No

Learn about optical character recognition in Microsoft Purview

By using optical character recognition (OCR) scanning, Microsoft Purview can scan images for sensitive information. OCR scanning is an optional feature that you must enable at the tenant level. After you enable it, select the locations where you want to scan images. You can scan images in Exchange, SharePoint, OneDrive, Teams, Windows, and macOS devices. When you configure the OCR settings, Microsoft Purview applies your existing policies for data loss prevention (DLP), records management, and insider risk management (IRM) to images and text-based content. For example, if you configure the DLP condition *content contains sensitive information* and include a data classifier such as the **Credit Card** sensitive information type (SIT), Microsoft Purview scans for credit card numbers in both text and images at all of the chosen locations.

Workflow at a glance

Expand table

Phase	What's needed
Create Azure subscription if needed	If your organization doesn't already have an Azure pay-as-you-go subscription for your tenant, your Global admin needs to start by creating an Azure account.
Estimate your OCR scanning charges	Use the OCR cost estimator to estimate the expected charges for your specific use cases.
Set up pay-as-you-go billing to enable OCR.	Your Global or SharePoint admin must follow the instructions in Set up Microsoft Syntex billing in Azure to add a subscription for OCR.
Configure OCR scanning settings	The Compliance admin for your organization configures the OCR settings for your tenant.

Prerequisites

To use OCR scanning, your organization's Global admin needs to verify that an Azure pay-as-you-go subscription is in place. If not, they need to set up the subscription by following the instructions in Create your initial Azure subscriptions.

Configure billing

When you enable OCR, all sensitive information types and trainable classifiers can detect characters that are in images.

Because it's an optional feature, your Global admin must set up pay-as-you-go billing to enable OCR. Refer to the instructions in Set up Microsoft Syntex billing in Azure to add a subscription for OCR.

> **Note**
>
> After you enter billing information in Microsoft Syntex, your Compliance admin can configure OCR in Microsoft Purview without any extra setup or licensing requirements.
>
> You can find OCR pay-as-you-go pricing information on the Set up Microsoft Syntex billing in Azure page.

Estimate your OCR scanning charges

Each image scanned counts as one transaction. This pricing means that stand-alone images (JPEG, JPG, PNG, BMP, or TIFF) each count as a single transaction. It also means that *each page* in a PDF file is charged separately. For example, if there are 10 pages in a PDF file, an OCR scan of the PDF file counts as 10 separate scans. For information on using the OCR cost estimator, see Estimate your OCR costs.

> **Note**
>
> To reduce your OCR costs, the service uses the following caching mechanisms: Small images, such as logos and signatures that are sent in email via Microsoft Exchange, are scanned and billed only once per unique image across all users of the tenant for a moving window of five days. For Endpoint, the cache is maintained for 30 days. Caching is local to each endpoint device and only classifiers identified on the image and image hash are stored. Customer data isn't stored. There's no caching mechanism for standalone images in SharePoint and OneDrive. However, in embedded file types, if only text is updated, images aren't scanned again.
>
> The service checks multiple parameters, including image stream hash and image size, to see if it can use the cache. If any parameter doesn't match, the service OCRs the image again.

Additionally, you can use each scanned image in any number of policies across data loss prevention, insider risk management, auto-labeling, and records management at no extra charge.

> **Important**
>
> For information about the Adobe requirements for using Microsoft Purview Data Loss Prevention (DLP) features with PDF files, see this article from Adobe: Microsoft Purview Information Protection Support in Acrobat.

Configure your OCR settings

To configure OCR scanning for your tenant, follow these steps:

1. Sign in to the Microsoft Purview portal.
2. Select **Settings**.
3. Select **Optical character recognition (OCR)** to enter your OCR configuration settings.
4. Select the locations where you want to scan images.
5. Select the groups that you want included or excluded from OCR scans.
6. Select **Done**.

For the full list of locations where OCR scans images and the solutions that act on the results, see Supported locations and solutions.

Permissions

To create and deploy policies, your account must be a member of one of these role groups:

- Compliance administrator
- Compliance data administrator
- Global administrator
- Information Protection
- Information Protection Admin

> **Note**
>
> In general, OCR settings take effect about an hour after you turn them on.

> **Note**
>
> For information on OCR functionality in Microsoft Purview Communication Compliance, see Create and manage communication compliance policies.

Supported locations and solutions

Expand table

Location	Supported Solutions
Exchange	Data loss prevention Information protection: Auto-labeling policies Records management: Autoapply retention label policies[1]
SharePoint sites	Data loss prevention Insider risk management[2] Records management: Autoapply retention label policies[1]
OneDrive accounts	Data loss prevention Records management: Autoapply retention label policies[1]
Teams chat and channel messages	Data loss prevention Insider risk management[2]
Devices	Data loss prevention Insider risk management[2]

[1] Supports keywords and sensitive information types.

[2] Considers sensitive information types and trainable classifiers present in images for risk scoring.

Supported file types

This functionality supports scanning images in the following file types, with the noted requirements:

Expand table

Locations	Supported file types
Exchange	JPEG, JPG, PNG, BMP, TIFF, and PDFs (scanned). Embedded images in DOCX, PPTX, XLSX, RAR, TAR, ZIP, 7z, and hybrid PDFs (containing searchable text and images) with a limit of 20 embedded images scanned per file.
SharePoint and OneDrive	BMP, PNG, JPEG, JPG, JFIF, ARW, CR2, CRW, ERF, GIF, MEF, MRW, NEF, NRW, ORF, PEF, RAW, RW2, RW1, SR2, TIF, TIFF, HEIC, HEIF, ARI, BAY, CAP, CR3, DCS, DCR, DRF, EIP, FFF, IIQ, K25, KDC, MOS, PTX, PXN, RAF, RWL, SRF, SRW, X3F, DNG, PDFs (scanned and hybrid containing searchable text and images) Embedded images in DOCX, PPTX, XLSX
Teams, Windows, and macOS endpoint	JPEG, JPG, PNG, BMP, TIFF, and PDF (image only)

Image requirements

Expand table

Requirement	Limit
File size (Exchange, Teams)	20 MB max
File size (SharePoint, OneDrive, Windows, and macOS endpoints)	50 MB max
Image resolution	50 × 50 px minimum, 16,000 × 16,000 px maximum

Important

- Only images uploaded after OCR is enabled are scanned.
- OCR extracts only the first 2 million characters of text.
- By default, incoming email (email from users outside the organization), internal mails (email shared within the users of the organization), and outgoing email (email sent to users outside the organization) are subject to OCR scanning. To exclude incoming mails from OCR scan, change the OCR settings from the default scope of **All sender groups** to the **Specific sender groups** and specify the internal groups that you want OCR to scan. To restrict OCR scans to mails sent outside the organization only, select the option under **Advanced Setting (Only Exchange)**. After selecting this checkbox, neither incoming mails nor any internal communications are OCRed. For information on changing the configurations, see Configure your OCR settings.
- Data loss prevention policy tips aren't supported for images in Exchange.

- If you **exclude a path** in the endpoint data loss prevention settings, OCR doesn't scan images in those folders.
- When OCR is turned on for Windows and macOS devices, the devices start sending messages to the cloud for scanning. The default bandwidth limit is **1,024 MB of data per device per day**. OCR stops scanning images once this daily limit is reached. If you want to continue scanning images, you can increase the bandwidth limit.
- For Endpoint Device, ensure that any network settings aren't obstructing the OCR, and a wildcard should be present allowing blob.core.windows.net endpoints.
- For Exchange, the feature supports embedded images in DOCX, PPTX, XLSX, RAR, TAR, ZIP, 7z, and hybrid PDFs (containing searchable text and images) with a limit of 20 embedded images scanned per file.

Supported languages

OCR scanning supports more than 150 languages.

Summary

- To use OCR, set up Microsoft Syntex pay-as-you-go billing. (You don't need to set up Microsoft Syntex itself.)
- Configure OCR at the tenant level, so once OCR is configured, it's available to the entire Microsoft Purview stack.
- You don't need to create separate data classifiers for OCR. Once OCR is configured, existing sensitive information types, exact data match based sensitive information types, trainable classifiers, and fingerprint SITs scan images as well as documents and emails.
- Microsoft Purview eDiscovery supports OCR at the case level. For more information, see Search and analytics settings in eDiscovery.

See also

- Learn about data loss prevention
- Learn about insider risk management
- Learn about records management

Last updated on 04/30/2026

Estimate your OCR costs

You can use the OCR cost estimator tool (preview) to estimate your optical character recognition (OCR) scanning costs. The tool scans your selected locations—such as Exchange, SharePoint, OneDrive, Teams, and Endpoint—and projects how much OCR processing will cost before you enable it. Use this article to run an estimation, review the results, and download cost reports.

Prerequisites

To enable the OCR cost estimator tool on Endpoint, you must first install the latest Windows Update released on February 29, 2024 or later.

Estimate your bill

To estimate your bill, follow these instructions:

1. Sign in to the Microsoft Purview portal.
2. Select **Settings**.
3. Select **Optical character recognition (OCR)**
4. Under **Estimate the cost of an OCR scan for free**, select **Try for free**.
5. In the flyout pane, select the locations where the files you want to scan are found and then choose **Run estimation**.
6. After **Estimation started successfully** displays in the flyout pane, select **View estimation** next to **Predict OCR cost estimate for free** to view the **Estimation** page.
7. To enable OCR scans from the **Estimation** page, select **Turn on OCR** in the flyout pane and then follow the steps to configure your OCR settings.

> **Note**
>
> Running an OCR estimation can take around 24 hours to complete the first time. After the first estimation completes, the dashboard is refreshed daily and can be viewed anytime.

Edit the scope of your estimate

To change which locations are included in the estimate, follow these steps:

1. On the **Estimation** page, select **Edit scope** from the dropdown menu next to **Enable OCR**.

2. In the flyout pane, edit the locations with files that you wish to scan and then select **Restart estimation**.

Download your estimate

To download your estimate reports, follow these steps:

1. On the **Estimation** page, select **Download Report** from the dropdown menu next to **Enable OCR**.
2. In the flyout pane, select **Download all**, or select the individual reports you want to download.

Cancel an in-progress estimate

To cancel an estimate that is currently running, follow these steps:

1. On the **Estimation** page, Select **Stop estimation** from the dropdown menu next to **Enable OCR**.
2. On the next page, confirm your choice.

Restart a stopped estimate

To restart a stopped estimate, follow these steps:

1. On the **Estimation** page, select **Start estimation** from the dropdown menu next to **Enable OCR**.
2. In the flyout pane, edit the locations with files that you wish to scan and then select **Start estimation**.

Limitations

Be aware of the following limitations when using the OCR cost estimator:

- OCR and OCR cost estimator can't run at the same time.
- OCR cost estimator can be used only once by a customer and can't be re-enabled. Once turned on, you will have 90 days to complete 30-day estimation. You may pause the estimation any number of times
- Reports must be downloaded after using estimator. The data from dashboards is deleted after 90 days from the date estimator is first turned on.
- Estimation for embedded and standalone image is only supported for the Exchange workload. Estimation for embedded images in SharePoint and OneDrive is not supported

yet.

- Estimation of images in pdf files is supported in Exchange and Teams. Estimation of images in pdf files isn't supported in SharePoint and OneDrive. As a result, your actual OCR bill for SharePoint and OneDrive may be higher than estimated, since OCR will run in SharePoint and OneDrive on all supported file types listed in Supported file types
- For Endpoint, each pdf file is estimated to have an image count of 4. This is the current average with OCR customers.
- In Exchange and Teams, there is the ability to cache small images, such as logos or email signatures, without being charged.
- Downloaded reports show the **To Date** in the **Estimation Timeframe** column as the date of latest scan (current date in case of ongoing scan) for all the rows in the report.
- All incoming emails and Teams messages from outside the organization are estimated if the scope is set to "ALL" when configuring OCR Cost Estimator for Exchange and Teams.
- All dashboards are reset when you restart estimation. Use the download report of the scan history before restarting an estimation. Even with the same location and scope, estimation can be restarted from scratch.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Create and analyze an on-demand classification scan

Microsoft Purview on-demand classification identifies and classifies sensitive content in historical data stored in SharePoint, OneDrive, and endpoints. On-demand classification extends the classification capabilities to files that aren't classified or modified for a long time, files that are never classified, or files that need updated classification on previously classified files.

On-demand classification updates classification results; matching policies determine any resulting actions. For SharePoint and OneDrive, updated results can trigger active policies, including policies that apply or remove sensitivity labels. On-demand classification isn't an auto-labeling simulation.

To decide whether existing files need updated classification or only policy reevaluation, see Auto-labeling classification and simulation concepts. Existing, correct classifications don't need to be refreshed for every auto-labeling policy change.

On-demand classification taken with Information Protection's continuous classification brings a two-pronged approach to keeping sensitive items aligned with your organization's latest security policies.

By using on-demand classification, organizations can:

- Extend protection to previously unclassified or inactive files, increasing overall coverage.
- Strengthen data protection across your environment without relying on end-user actions.
- Reduce the risk of AI tools surfacing unlabeled or unprotected information. Do it all natively, without exporting your data or relying on fragmented tools.

SKU and subscriptions licensing

For information on licensing, see

- Microsoft 365 Enterprise Plans
- Microsoft 365 Service Descriptions

Billing

This feature uses pay-as-you-go billing or per-user licensing for Microsoft Purview capabilities or both. To help you understand and manage your usage, Microsoft Purview provides a **Usage center** in the Microsoft Purview portal. For more information, see Manage pay-as-you-go and per-user licensing usage. To learn more about Purview billing, see Purview billing models..

> **Note**
>
> The actual billed amount may be lower than the estimated cost if some items or locations are skipped during the scan. Items can be skipped if they are empty files (excluded to avoid unnecessary billing), if the scan is canceled, or if the subscription expires before completion. Locations may be skipped if another scan is already running on the same location at the same time.

Permissions

To run a scan, you must be a member of the following role group:

- Compliance Administrator

To view classification results, you must be a member of one of these role groups:

- Content Explorer Content Viewer
- Content Explorer List Viewer

Create an on-demand classification scan

Before starting classification for SharePoint or OneDrive, review the scan scope, estimated charges, and active policies that might act on the updated results. Check that the file date range and selected classifiers cover the historical files you intend to classify. A policy you're testing can remain in simulation while other active policies act on those files.

1. Sign in to the Microsoft Purview portal.
2. Go to **Data loss prevention** > **Classifiers** > **On-demand classification** or **Information Protection** > **Classifiers** > **On-demand classification**.
3. Select **New scan**.
4. Follow the instructions in the wizard. During this process, you define the following settings:
    - Name and description

- Scope and location - You can choose to scan all SharePoint sites and OneDrive accounts, only specific ones, or skip certain sites and accounts from the scan. For endpoint devices, you can scope the scan by users.
- Classifiers to scan for
- File last modified date range you want to scan
- File extensions you want to scan

After you complete the wizard, the estimation process begins. The estimation duration depends on the scope of the scan.

> ⓘ **Note**
>
> By default, the scan includes all items created or modified within the past year, across all supported file extensions. It also includes every available classifier configured in your tenant. If you choose to scan for specific classifiers, you can select up to 50 at a time.

5. From the **On-demand classification** list view, select the scan you created.
6. Select **View estimation**.

> ⓘ **Note**
>
> After reviewing the estimates, you can edit the scan to narrow or expand the scope. Select **Edit scan** and rerun the estimation scan.

7. Select **Start classification.**

Analyze on-demand classification results

1. From **On-demand classification**, select a scan from the list.
2. Select **View estimation**.
3. On the **Estimation overview** tab, review scan results, including progress, items found, and estimated cost. To cancel in-progress scans, select **Cancel Scan**.
4. On the **Items for review** tab, review specific items found during the scan. You can filter and export the result.

Devices that don't respond are ones that don't send endpoint telemetry signals for over 72 hours. Devices with expired scan validity don't have a valid scan in the last 72 hours.

Additional considerations

- Classification can start up to 30 days after estimation, but minimizing the gap ensures greater accuracy in final counts and costs.
- Tenants using EDM classifiers should be aware that On-Demand Classification scans (including estimation and classification) rescan previously classified files, which can increase overall scan costs. To mitigate this impact, tenants can exclude the EDM classifier from the scan scope.
- Impact of classifier selection on scan cost: The classifiers selected for an on-demand scan influence how many files are identified for processing. Scan cost is determined by the total number of files scanned—not the number of classifiers. Selecting a broader set of classifiers may increase the number of files in scope and raise overall cost. However, this approach ensures files are assessed against all relevant classifiers in a single scan, reducing the risk of missed sensitive data and limiting the need for additional scans later.

Applies to SharePoint and OneDrive

- Each scan can process up to 150,000 locations and 100 million files. The system enforces these limits based on scan estimation results.

  > ⓘ **Note**
  >
  > In On-Demand Classification estimation scans (simulation), you see a difference between the number of matched locations and the number of users or accounts included in the scan scope. Simulation results are based on mailbox locations rather than unique users. A single user can have multiple mailboxes (for example, a primary mailbox along with archive or system mailboxes), and each mailbox is treated as a separate location during scanning. As a result, the total number of matched locations can exceed the number of users in scope, even when the scan is correctly applied at the user level.

- If you select specific classifiers during a scan, only those classifiers have their classification results updated for the scanned files. Other classifiers present in the files might remain unevaluated and might not trigger their associated policies.

- If a location is in scope of multiple active scans, only one scan processes it successfully and marks it skipped for others.

- Each file, once scanned, is evaluated against Data Loss Prevention (DLP), Information Protection (MIP), Data Lifecycle Management (DLM), and Insider Risk Management (IRM) policies, triggering appropriate actions as per matching policies.
- Content Explorer updates within seven days of scanning to reflect newly classified content.
- Adaptive scopes for SharePoint on-demand classification let administrators define scan scope dynamically based on site attributes (such as URL, name, or metadata) instead of selecting sites manually. The scan automatically includes all sites that match the criteria and updates as sites change, enabling scalable, accurate coverage with minimal manual maintenance. To learn more, see Adaptive scopes.
- Administrative units (AUs) for SharePoint on-demand classification let organizations scope and manage scans within defined administrative boundaries, enabling delegated administration across large or distributed environments. Administrators can create and manage scans only for sites within their assigned AUs, and access to scan configuration and results is limited to AU-scoped admins. This supports least-privilege access, simplifies ownership across regions or departments, and ensures alignment with existing governance in Microsoft Purview. For more information, see Administrative units in Microsoft Purview.

Applies to Windows 10/11 endpoints

- Ensure that the machine has a minimum four virtual processors and a minimum of 4,096 MB memory.
- Each device is limited to 2 GB of data discovery bandwidth per rolling 24-hour period (independent of DLP). Scanning pauses once the limit is reached and resumes after 24 hours.
- Only devices that successfully complete estimation are included in classification.
- In addition to DLP-excluded paths, certain system folders are automatically excluded from data discovery.

  `%WinDir%`

  `%ProgramFiles%`

  `%ProgramFiles(x86)%`

  `%SystemDrive%\Windows.old`

  `%SystemDrive%\Users\*(1)\Application Data`

  `%SystemDrive%\Users\*(1)\AppData\Local\Application Data`

  `%SystemDrive%\Users\*(1)\Local Settings\Application Data`

`%SystemDrive%\Documents and Settings\*(1)\Application Data`

`%SystemDrive%\Documents and Settings\*(1)\AppData\Local\Application Data`

`%SystemDrive%\Documents and Settings\*(1)\Local Settings\Application Data`

`%ProgramData%\Application Data`

`%ProgramData%\Microsoft\Windows\WER\`

See also

Learn about trainable classifiers

Learn about sensitive information types

Deploy an information protection solution with Microsoft Purview

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 10/14/2026

Get started with on-demand classification for endpoints

On-demand classification identifies and classifies sensitive content in files at rest on Windows endpoints. It's the endpoint workload of on-demand classification, which also covers SharePoint and OneDrive.

> **Note**
>
> This article covers on-demand classification for endpoints. For real-time endpoint protection, see Learn about Endpoint data loss prevention.

What on-demand classification covers on endpoints

Endpoint DLP classifies files in real time when they're accessed, modified, or shared. On-demand classification extends that coverage to inactive and historical files at rest that Endpoint DLP hasn't processed.

Expand table

Capability	Scope	Trigger
Endpoint DLP (real-time)	Active files (modified, accessed, shared)	Real-time, automatic
On-demand classification	Inactive/historical files at rest	Admin-initiated scan

Together, these capabilities help ensure that both active and inactive files on endpoints are classified according to your organization's current policies.

What on-demand classification doesn't do

Expand table

On-demand classification...	Details
Doesn't apply sensitivity labels	It classifies files but doesn't label them. Combine with auto-labeling policies to apply labels based on classification results.

On-demand classification...	Details
Doesn't enforce DLP policies directly	Classification results feed into policies, but enforcement happens through Endpoint DLP when files are accessed or shared.
Doesn't replace Endpoint DLP	Endpoint DLP handles real-time protection. On-demand classification covers historical files that Endpoint DLP hasn't processed.

How on-demand classification works

On-demand classification uses a two-phase approach: estimation followed by classification.

Phase 1: Estimation

After you create a scan, the estimation process begins on scoped devices. The endpoint agent performs a metadata-only pass without reading file contents or running classifiers:

- Enumerates files that match your scope criteria (selected users, file types, last modified date range).
- Evaluates file eligibility gates, including file size limits, supported file types, and time window filters.
- Checks eligibility for previously classified files by comparing classifier history and file modification state against previous scan results.
- Reports aggregated metadata to the Microsoft Purview service, including eligible file counts and estimated reusable files.

No file contents are read and no classifiers are executed during estimation.

> ⓘ **Note**
>
> Estimation numbers may differ from the final classification results. Estimation reflects a point-in-time metadata snapshot. Files may change, devices may go offline, or previously classified file eligibility may shift between estimation and classification. This is expected behavior.

Phase 2: Classification

After reviewing the estimation:

- Select **Start classification** to begin classification.

- Each file is evaluated against your sensitive information types (SITs) and trainable classifiers.
- Files are evaluated against DLP, Microsoft Information Protection (MIP), data lifecycle management (DLM), and insider risk management (IRM) policies simultaneously.
- Previously classified files with unchanged content and classifiers are retained without rescanning.

Phase 3: Review results

After classification completes:

- View results on the **Items for review** tab.
- Filter and export classified items.
- Monitor device status and scan progress.

Prerequisites

Devices must be onboarded to Microsoft Purview endpoint data loss prevention. For onboarding and device requirements, see Get started with Endpoint data loss prevention.

Expand table

Action	Required role
Create and run a scan	Compliance Administrator
View classification results	Content Explorer Content Viewer or Content Explorer List Viewer

For licensing requirements, see Microsoft 365 guidance for security & compliance.

Create your first endpoint scan

Step 1: Navigate to on-demand classification

1. Go to the Microsoft Purview portal.
2. In the left navigation, select **Data Loss Prevention** > **Classifiers**.
3. Select the **On-demand classification** tab.
4. Select **New scan**.

Step 2: Configure your scan

Expand table

Setting	Description	Guidance
Name & description	A name to identify this scan	Use a descriptive name, such as "Finance team endpoint scan, quarterly"
Scope & location	Select **Endpoints** and choose which users' devices to scan	Start with a small user group for your first scan
Classifiers	Select which SITs and trainable classifiers to scan for (up to 50)	Include only the classifiers you need. More classifiers means longer scan time.
File last modified range	Filter files by when they were last modified	Use this to focus on files from a specific time period
File extensions	Which file types to include	Leave as default (all supported types) unless you need to narrow scope

Endpoint scans are scoped by user. All onboarded devices for selected users are included. System folders (Windows, Program Files) are automatically excluded.

Step 3: Start estimation

1. Review your scan configuration.
2. Select **Create scan**.
3. The estimation phase begins automatically.

Step 4: Review estimation and start classification

Once estimation completes, review the estimated file count and projected scope. Select **Start classification** to begin.

You can start classification before estimation reaches 100%. Starting at 80–90% progress is recommended: the final 10–20% typically reflects slow-responding devices, and devices that have completed estimation begin classification immediately while remaining devices join as they finish.

Step 5: Review results

After classification completes:

1. Open your scan and select the **Items for review** tab.
2. Filter results by classifier, file type, or sensitivity.

Bandwidth throttling

Each device has a data discovery bandwidth limit per rolling 24-hour period. This limit is independent of other DLP bandwidth and prevents the scan from affecting device performance. When a device reaches the limit, scanning pauses automatically and resumes after the period resets.

Next steps

- Best practices for on-demand classification on endpoints and billing
- Monitor and troubleshoot endpoint scans
- FAQ and reference for endpoint scans

Last updated on 08/21/2026

Best practices for on-demand classification on endpoints and billing

Understand how to plan large-scale scans and how to optimize usage.

Billing model

On-demand classification uses a pay-as-you-go billing model. You're billed only for files that are actively classified. Previously classified files, skipped files, and failed files aren't billed. Usage is reported in the **Microsoft Purview Usage Center**, and charges appear on your Azure subscription bill.

For current pricing and licensing, see Microsoft Purview pricing.

Previously classified files

When a file has already been classified and neither the file content nor the classifiers have changed, the existing classification result is retained on subsequent scans without rescanning. This applies to files classified by prior on-demand scans, as well as files classified by real-time Endpoint DLP when advanced classification is enabled.

Conditions for previously classified files

> **Note**
>
> Reuse of previously classified verdicts requires that the earlier scan used advanced classification. For Endpoint DLP, advanced classification must be enabled on the DLP policy for its verdicts to be reusable by later on-demand scans.

A file qualifies as previously classified when all the following are true:

- File content unchanged since last verdict.
- The classifiers in the current scan are the same as, or a subset of, the classifiers used in the original scan, and those classifiers are unchanged.
- Classifier definitions unchanged.
- The device has the latest Windows cumulative updates installed.

Previously classified results apply across: on-demand → on-demand, on-demand → Endpoint DLP, and Endpoint DLP → on-demand.

Previously classified files are visible in results

Previously classified files appear in scan results and audit logs under classification status. The file was evaluated and the existing verdict was confirmed. It doesn't mean the file was skipped or removed from results; it only wasn't rescanned.

> **Important**
>
> Previously classified file recognition isn't retroactive. It only applies to scans run after the required Windows update version 10.8821 is installed. Scans completed before this update aren't eligible. Your first post-update scan establishes the baseline for all future scans.

What affects scan volume

Expand table

Reduces rescanning	Increases rescanning
Previously classified files on repeat scans	Changing classifiers between scans
Files already classified by Endpoint DLP (cross-path)	Large number of new or modified files
Narrowing file scope with date range and file type filters	Broad file scope (more file types, longer date ranges)

Estimation accuracy over time

An estimate stays representative for a limited window from the day estimation started. Within this window, the estimated file count closely reflects what classification will process. You can still start classification later, but expect higher variance in file counts as files on devices are added, modified, or deleted over time.

A common scenario: an administrator runs estimation, submits the projected scope for budget approval, and the approval cycle takes weeks. If estimation itself also drags out (for example, waiting for slow-responding devices to reach 100%), the combined delay can push classification past the window where the estimate stays representative.

How to keep the estimate accurate

- **Start with a pilot.** Run a small pilot scan (10–20 users) first to get per-user data for the approval process.
- **Start classification early.** Don't wait for estimation to reach 100%. Starting at 80–90% progress is recommended. Devices that have completed estimation begin classification immediately.
- **Run estimation and classification back-to-back** if your organization is comfortable with the projected scope.

Best practices for large-scale scans

Finalize classifiers first

Finalize your sensitive information types and trainable classifiers before starting a scan. Changing classifiers during or between scans invalidates previously classified files and increases rescanning.

> **Tip**
>
> Establish a change-freeze period for classifiers before large scans. Coordinate with your team to ensure that SIT and trainable classifier changes are paused until the scan completes.

Start with a phased rollout

Expand table

Phase	Scope	Goal
Phase 1	10–20 users	Validate the scan works and understand per-device timing. Complete the full cycle from start estimation to complete classification.
Phase 2	50–100 users	Identify device issues and gather per-user data. Complete the full cycle from start estimation to complete classification.
Phase 3	Full rollout	Scale to all target users with known expectations. Complete the full cycle from start estimation to complete classification.

The more devices in your scan, the more files are processed in parallel.

Batch large tenants

For organizations with thousands of users, consider batching scans into groups of users per scan. This approach keeps each scan within the estimation accuracy window, allows batches to run in

parallel, and makes it easier to track progress per batch.

> **Note**
>
> A device can only be part of one active scan at a time. Before adding a device to a new estimation or classification scan, ensure its previous scan has reached Classification Complete, Classification Cancelled, or Estimation Cancelled state.

Handle offline devices

In any large scan, some devices will be offline. Don't wait for 100% device coverage. Run a follow-up scan later to catch devices that missed the first round. Previously classified files aren't rescanned, so the follow-up scan only processes newly scanned files.

After the scan

- Download scan reports promptly; they're retained for a limited period of 30 days.
- Plan regular scans (monthly or quarterly) to keep classification current if the classifiers are changed or policy has been modified, since previously classified files aren't rescanned.

Key considerations

- The estimation phase reads only metadata. You can assess scope before committing to classification.
- Subsequent scans typically process fewer files than the initial scan because previously classified files with unchanged content and classifiers aren't rescanned.
- Classification runs on the device. Only classification results and metadata are sent to the Microsoft Purview service.
- On-demand classification uses the same classifiers and Microsoft Purview portal as your existing compliance solutions.

Next steps

- Get started with on-demand classification for endpoints
- Monitor and troubleshoot endpoint scans
- FAQ and reference for endpoint scans

Last updated on 08/21/2026

FAQ and reference for endpoint scans

Answers to common questions and key terms for on-demand classification for Windows endpoints in Microsoft Purview.

General

Does on-demand classification apply sensitivity labels to files?

No. On-demand classification identifies and classifies sensitive content but doesn't apply sensitivity labels. To apply labels based on classification results, configure auto-labeling policies.

Is on-demand classification required for Endpoint DLP to work?

No. Endpoint DLP works independently and classifies files in real time when they're created, modified, or used. On-demand classification is a separate, complementary capability that covers historical files at rest.

Does on-demand classification affect the user experience on endpoints?

On-demand classification runs in the background with minimal resource usage. A per-device bandwidth limit ensures scanning doesn't degrade device performance.

Does turning on a DLP policy automatically classify historical data?

No. DLP policies classify files when they're accessed, modified, or shared (real-time classification). Historical files that remain untouched aren't retroactively classified by DLP, however you can utilize JIT feature to prevent data loss. On-demand classification is required to cover those files.

Classification and labeling

Why was my file classified but not labeled?

Classification and labeling are separate steps. Classification identifies whether a file contains sensitive content. Labeling applies a sensitivity label based on classification results. For labels to be applied, configure auto-labeling policies.

Does opening or editing a file trigger reclassification?

Yes, but through Endpoint DLP (real-time classification), not on-demand classification. When a file is modified, the Endpoint DLP engine reclassifies it automatically. On-demand classification runs only when an admin starts a scan.

Do I need to rerun on-demand classification when I create a new SIT?

Yes, if you want historical files evaluated against the new SIT. Existing verdicts were created using the classifiers available at scan time. A new SIT isn't automatically applied to previously classified files. A new scan that includes the new SIT classifies the files against it.

Cost and billing

Is estimation billed?

No. Estimation is a metadata-only pass. The endpoint agent enumerates files and evaluates eligibility gates (file size, type, time window) and previously classified file potential, but doesn't read file contents or run classifiers. No charges are incurred during estimation.

What is the billing model?

On-demand classification uses a pay-as-you-go billing model. You're billed only for files that are actively classified. Previously classified files and failed files aren't billed. For current pricing and licensing, see Microsoft Purview pricing.

Will I be charged twice if I scan the same files again?

No. Files that haven't changed and use the same classifiers are recognized as previously classified files and aren't rescanned or billed.

Why do I have to pay if DLP has already scanned the file?

If the file qualifies as previously classified, you don't. Files classified by real-time Endpoint DLP are recognized as previously classified during on-demand scans at no charge, if the file content and classifiers haven't changed. This is cross-path recognition.

Why is my actual scan volume lower than the estimate?

The estimate shows an upper bound and assumes all files will be classified. Actual volume is lower because previously classified files aren't rescanned, some devices may be offline, and some files might be deleted before classification.

How do I see my usage and charges?

Usage is reported in the Microsoft Purview Usage Center. Charges appear on your Azure subscription bill.

Scans

Can I cancel a running scan?

Yes. Already-classified files remain classified. Canceling doesn't undo completed work.

Can I run multiple scans at the same time?

Yes. If a device is scoped by multiple active scans, the first scan is processed while the other scans are queued.

Can I start classification before estimation reaches 100%?

Yes. Starting at 80–90% progress is recommended. The final 10–20% typically reflects slow-responding devices, and waiting for them can push the overall estimation phase past the window where the estimate stays representative. Devices that have completed estimation begin classification immediately, and remaining devices join as they finish.

What happens if I change classifiers during a scan?

Changing sensitive information types or trainable classifiers during a scan invalidates the estimation. You should re-estimate before continuing to classification if you want the estimated and classified numbers to match.

What files are excluded from scanning?

System folders are automatically excluded: the Windows directory, Program Files, legacy application data folders, and Windows Error Reporting data. On-demand classification scans user data directories such as Documents, Desktop, and Downloads.

Devices

Why does my device show "Not responding"?

The device hasn't sent signals for an extended period, currently cutoff limit is 48 hours. This typically means the device is powered off, hibernating, or not connected to the network.

What are the device requirements?

Devices must be onboarded to Microsoft Purview endpoint DLP and have the latest Windows cumulative updates installed. For full device requirements, see Get started with Endpoint data loss prevention.

Does on-demand classification work on macOS?

Not currently. macOS support is planned but no release date is available.

Results

Where can I see which files failed?

Individual failed files aren't listed in the portal or audit logs. The portal shows the total count of failed files. Review the error categories in Monitor and troubleshoot endpoint scans for common causes.

Why don't audit log counts match Items for review counts?

Audit logs and the Items for review tab use different data sources and counting logic. Timing differences, deduplication, and filter differences can cause minor discrepancies. A small gap is normal.

What is the difference between "skipped" and "failed" files?

Expand table

Status	Meaning	Billed
Classified	File was scanned and classified	Yes
Previously classified	File was scanned previously and had a valid prior verdict	No
Skipped	Excluded by filters	No
Failed	Could not be scanned (encrypted, locked, access denied)	No

Previously classified files

When are files not recognized as previously classified?

Files aren't recognized as previously classified when the file was earlier not classified using advanced classification, has been modified, classifier definitions have changed, or new classifiers were added that weren't in the original classification.

Does this work between on-demand and Endpoint DLP?

Yes. Previously classified file recognition works across on-demand → on-demand, on-demand → Endpoint DLP, and Endpoint DLP → on-demand.

Is previously classified file recognition retroactive?

No. Previously classified file recognition only applies to scans run after the required Windows update is installed. Your first post-update scan establishes the baseline for all future scans.

Feature gaps

Can I mark false positives in the scan results?

There's no built-in workflow to mark items as false positives directly within on-demand classification results. Export results and manage them externally, or report false positive SIT matches through your normal classifier tuning process.

Does the portal show an estimated time remaining for scans?

No. The portal shows scan progress but doesn't provide an estimated time of completion.

The on-demand classification option isn't visible in my portal. What should I do?

In some environments, backend provisioning may be required. Verify licensing and endpoint DLP onboarding. If all prerequisites are met, contact Microsoft Support for backend enablement.

Key terms

Expand table

Term	Definition
On-demand classification	Admin-initiated scan that classifies sensitive content in files at rest on endpoints.
Estimation	First phase that enumerates files, evaluates eligibility gates, and checks previously classified file potential using metadata only. No file content is read. The estimate stays representative for a limited window from the day estimation started.
Classification	Second phase that scans files against classifiers and is billed under the pay-as-you-go model.
Verdict	Classification result stored on a file, recording which classifiers matched.
Previously classified files	Files with existing classification results from prior scans or classified by Endpoint DLP in real time. These files aren't rescanned or billed when content and classifiers are unchanged. Requires latest Windows updates.
Active data	Files frequently modified, accessed, or shared. Classified by Endpoint DLP in real time.
Inactive/historical data	Files at rest that haven't been recently modified. The target of on-demand classification.

Next steps

- Get started with on-demand classification for endpoints
- Best practices for on-demand classification on endpoints and billing
- Monitor and troubleshoot endpoint scans

Last updated on 08/21/2026

Monitor and troubleshoot endpoint scans

Track scan progress, understand device and scan statuses, and resolve common issues with on-demand classification for Windows endpoints.

Viewing scan progress

1. Go to the Microsoft Purview portal.
2. Navigate to **Data Loss Prevention** > **Classifiers** > **On-demand classification**.
3. Select your scan to view detailed progress.

The scan details page shows:

Expand table

Metric	Description
Scan status	Overall scan phase (estimation or classification) and progress
Files scanned	Number of files processed so far
Files in scope	Total estimated files from the estimation phase
Devices active	Devices currently scanning
Devices not responding	Devices that haven't reported recently
Devices where scan validity expired	Devices that exceeded the signal validity threshold and were removed from scan scope

> **Note**
>
> Scan progress in the portal isn't real-time. Updates arrive periodically as devices report their status.

Understanding scan statuses

Estimation phase

Expand table

Status	Meaning	Action
Estimation in progress	Devices are counting files and reporting scope	Wait for periodic progress updates
Estimation complete	Ready to start classification	Review results and start classification
Estimation cancelled	Admin cancelled the estimation	Create a new scan if needed
Estimation failed	An error occurred	Check device health and requirements

Classification phase

Expand table

Status	Meaning	Action
Classification in progress	Files are being scanned and classified	Monitor progress
Classification complete	All devices finished	Review results in the **Items for review** tab
Classification cancelled	Admin cancelled classification	Already-classified files are preserved
Classification failed	An error occurred	See troubleshooting sections below

Understanding device statuses

Expand table

Status	Meaning	Action needed
Active	Device responded recently, scanning normally	No
Not responding	No recent signals. The device is likely offline or powered down. A device that's once marked not responding won't participate any further in the scan.	No
Completed	Device finished its portion of the scan	No
Paused	Device reached the data discovery bandwidth limit	Resumes automatically after the period resets

Each device has a signal validity threshold. If a device doesn't send any event within this window, it's marked as "Not responding" and is out of scope for the scan.

> **Note**
>
> "Not responding" doesn't indicate a device failure. It typically means the device is offline.
> Common reasons: device powered off, user on leave, device on a different network.

Interpreting results

After classification completes:

Expand table

Metric	What it means
Files classified	Files actively scanned and classified (billable)
Previously classified files	Files with existing verdicts retained without rescanning, or already covered by another active scan
Files failed	Files that couldn't be classified (encrypted, locked). Not billed.
Files skipped	Files not found on the system (deleted before the scan moved to the classification phase), empty extracted text in small files, label- or password-protected files, and text extraction or format detection errors

Understanding result number differences

Expand table

Metric	What it counts	When it updates
Items found (estimation)	Total files matching your scope and filters	During estimation
New matches (classification)	Files actively scanned, a subset of items found	During classification
Items for review (portal tab)	Classified items that matched at least one classifier	During and after classification
Audit log entries	Events logged per classification action	After classification, with possible delay

> **Note**

Items found ≠ items classified. Items found is the estimation-phase count (all files in scope). Items classified is the classification-phase count (files actually scanned, minus previously classified and failed files). The difference is primarily due to previously classified files.

> **Note**
>
> Items for review ≠ audit log count. These use different data sources and may show slightly different numbers. A small discrepancy is normal.

Skipped file counts are only finalized during the classification phase, not during estimation.

> **Note**
>
> Endpoint on-demand classification scan results don't flow into Content Explorer. Review results from the **Items for review** tab on the scan.

Troubleshooting

Quick reference

Expand table

Symptom	Likely cause	Solution
Item count appears to drop or reset mid-scan	Previously classified files are deducted from the live classified count	Expected behavior. The final count reflects net new classifications, not a scan restart.
Device shows "Not responding"	No recent signals	Device is offline. This device won't be part of the current scan.
Estimation stuck at x%	Outdated Sense/Defender client on devices	Upgrade to the latest Sense/Defender client through Windows Update
Very low file count	Bandwidth limit reached	Expected. Wait for the bandwidth period to reset.
High failure count	File access errors	See error categories below
Unexpected rescanning	Classifiers changed	Finalize classifiers before starting scans

Classification errors

Failed files aren't billed. Common error categories:

Expand table

Category	Examples	Solution
Access errors	Access denied, cloud file placeholder not available locally	Check file permissions; sync cloud files locally
Encryption errors	File encrypted, password-protected, BitLocker locked volume	Expected. Encrypted files can't be classified.
Format errors	File type not supported, corrupt file	Check file extension filters; no action needed for corrupt files
Lock errors	File locked by another application	Transient. File is retried on the next scan.
I/O errors	File or path not found (deleted or moved during scan)	No action needed
System errors	Local scan database full, unexpected error	Check disk space; restart device and retry

> **Note**
>
> A small percentage of file failures is normal. Investigate only if failure counts indicate a systemic issue such as permissions problems across many files.

Estimation shows different numbers than classification

This is expected. Estimation counts all files in scope. Classification skips previously classified files that already have valid verdicts, and some devices may be offline during classification. Estimation provides an upper bound; the actual classification volume is typically lower.

Agent downgrades

If the endpoint agent (Defender/Sense client) is downgraded or rolled back while a scan is active, scan progress may be affected. Coordinate agent updates with your scan schedule.

Scan not appearing in the portal

In some environments, backend provisioning may be required before the feature appears. Verify your tenant's licensing and that endpoint DLP is enabled with devices onboarded. If all prerequisites are met, contact Microsoft Support for backend enablement.

Where to find diagnostic information

Expand table

Information	Location
Scan progress and status	Purview portal → On-demand classification → your scan
Classified items with matches	Items for review tab
Audit trail of scan events	Unified audit log
Failed file count	Scan summary

When to contact support

Contact Microsoft Support if:

- Scans fail repeatedly with system errors across multiple devices.
- Estimation doesn't progress despite devices being online and meeting requirements.
- Billing charges don't match expected classification volume.

Next steps

- Get started with on-demand classification for endpoints
- Best practices for on-demand classification on endpoints and billing
- FAQ and reference for endpoint scans

Last updated on 08/21/2026

Learn about sensitivity labels

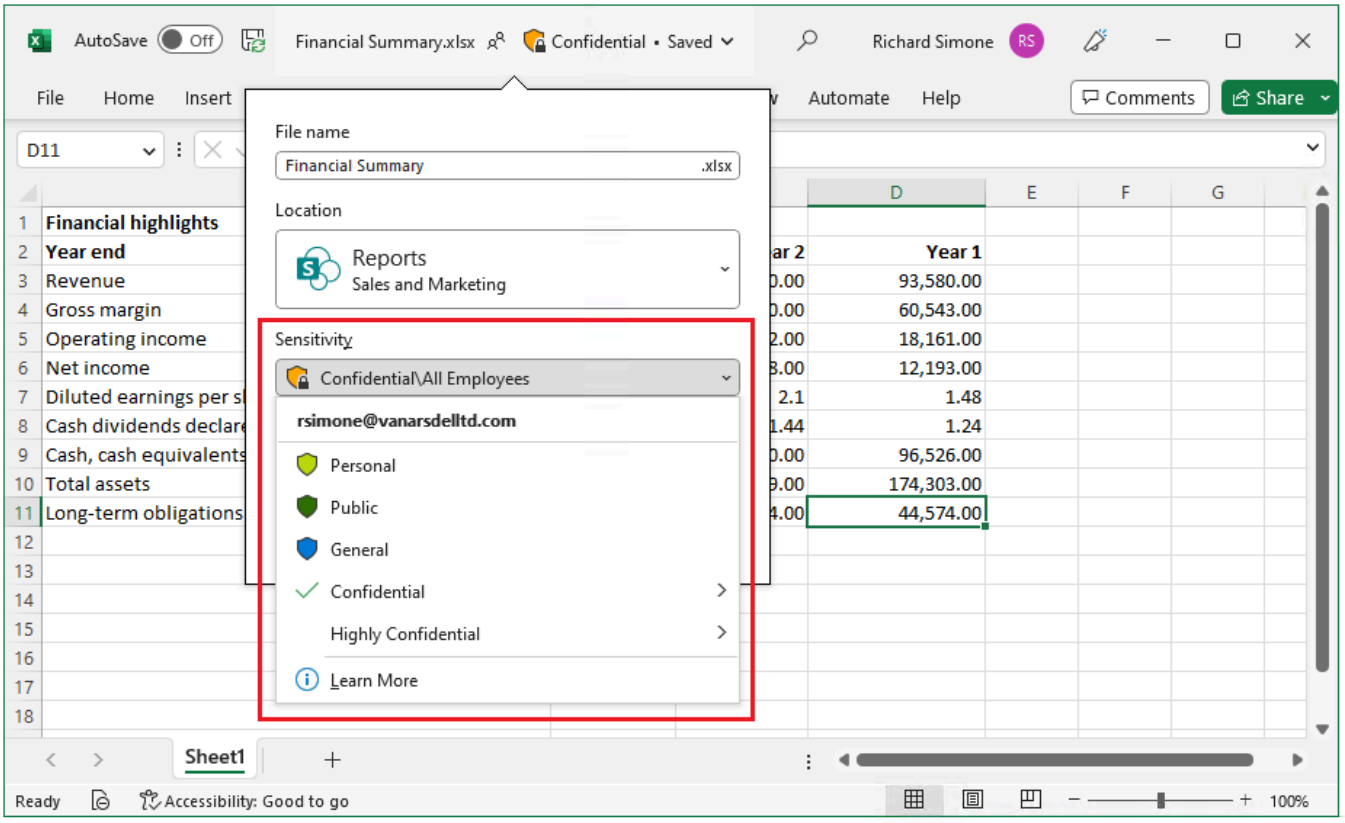
> *Microsoft Purview service description*

> **Note**
>
> If you're looking for information about sensitivity labels that you see and can select in your Office apps, check their documentation. For example, Apply sensitivity labels to your files.
>
> The information on this page is for IT administrators who can create and configure those labels by using the Microsoft Purview management solution.

To get their work done, people in your organization collaborate with others both inside and outside the organization. This collaboration means that content no longer stays behind a firewall—it can roam everywhere, across devices, apps, and services. When content roams, you want it to do so in a secure, protected way that meets your organization's business and compliance policies.

Sensitivity labels from Microsoft Purview Information Protection let you classify and protect your organization's data, while making sure that user productivity and their ability to collaborate isn't hindered.

The following example from Excel shows how users might see an applied sensitivity label from the window bar, and how they can easily change the label by using the sensitivity bar that's available with the latest versions of Office. The labels are also available from the **Sensitivity** button on the **Home** tab from the ribbon.



To apply sensitivity labels, users must be signed in with their Microsoft 365 work or school account.

> ⓘ Note
>
> For US Government tenants, sensitivity labels are supported for all platforms.
>
> If you use the Microsoft Purview Information Protection client and scanner, see the Azure Information Protection Premium Government Service Description.

You can use sensitivity labels to:

- **Provide protection settings that include encryption and content markings.** For example, apply a "Confidential" label to a document or email, and that label encrypts the content and applies a "Confidential" watermark. Content markings include headers and footers as well as watermarks, and encryption can also restrict what actions specified people can take on the content.

- **Extend SharePoint protection when files are downloaded** when you configure a default sensitivity label for SharePoint document libraries and select the option to extend protection for unencrypted files. Then, when these files are downloaded, the current SharePoint permissions travel with the labeled file.

- **Protect content in Office apps across different platforms and devices.** Supported by Word, Excel, PowerPoint, and Outlook on the Office desktop apps and Office for the web. Supported on Windows, macOS, iOS, and Android. For other apps, check their documentation.

- **Protect content in third-party apps and services** by using Microsoft Defender for Cloud Apps. With Defender for Cloud Apps, you can detect, classify, label, and protect content in third-party apps and services, such as SalesForce, Box, or DropBox, even if the third-party app or service doesn't read or support sensitivity labels.

- **Identify content for eDiscovery cases**. The condition builder to create search queries in eDiscovery supports sensitivity labels that are applied to content. For example, as part of your eDiscovery case, restrict content to files and emails that have a "Highly Confidential" sensitivity label. Or conversely, exclude content to files and emails that have a "Public" sensitivity label.

- **Protect containers** that include Teams, Microsoft 365 Groups, SharePoint sites, Viva Engage communities, and Loop workspaces. For example, set privacy settings, external user access and external sharing, access from unmanaged devices, and control how channels can be shared with other teams.

- **Protect meetings and chat** by labeling (and optionally, encrypting) meeting invites and any responses, and enforce Teams-specific options for the meeting and chat.

- **Extend sensitivity labels to Power BI**: When you turn on this capability, you can apply and view labels in Power BI, and protect data when it's saved outside the service.

- **Extend sensitivity labels to assets in Microsoft Purview Data Map**: When you turn on this capability, currently in preview, you can apply your sensitivity labels to files and schematized data assets in Microsoft Purview Data Map. The schematized data assets include SQL, Azure SQL, Azure Synapse, Azure Cosmos DB, and AWS RDS.

- **Extend sensitivity labels to third-party apps and services.** Using the Microsoft Information Protection SDK, third-party apps can read sensitivity labels and apply protection settings.

- **Label content without using any protection settings.** You can also just apply a label as a result of identifying the sensitivity of the data. This action provides users with a visual mapping of your organization's data sensitivity, and the labels can generate usage reports

and activity data for data that has different levels of sensitivity. Based on this information, you can always choose to apply protection settings later.

- **Protect data when Microsoft 365 Copilot is used.** Copilot and agents recognize and integrate sensitivity labels into the user interactions to help keep labeled data protected. For more information, see the following section Sensitivity labels for Microsoft 365 Copilot and Microsoft 365 Copilot Chat.

In all these cases, sensitivity labels from Microsoft Purview can help you take the right actions on the right content. With sensitivity labels, you can identify the sensitivity of data across your organization, and the label can enforce protection settings that are appropriate for the sensitivity of that data. That protection then stays with the content.

For more information about these and other scenarios that are supported by sensitivity labels, see Common scenarios for sensitivity labels. To see how your sensitivity labels are being used, view the reports from the Microsoft Purview portal.

New features are being developed all the time that support sensitivity labels, so you might also find it useful to check the Microsoft 365 roadmap.

What a sensitivity label is

When you assign a sensitivity label to content, it's like a stamp that you apply. The label is:

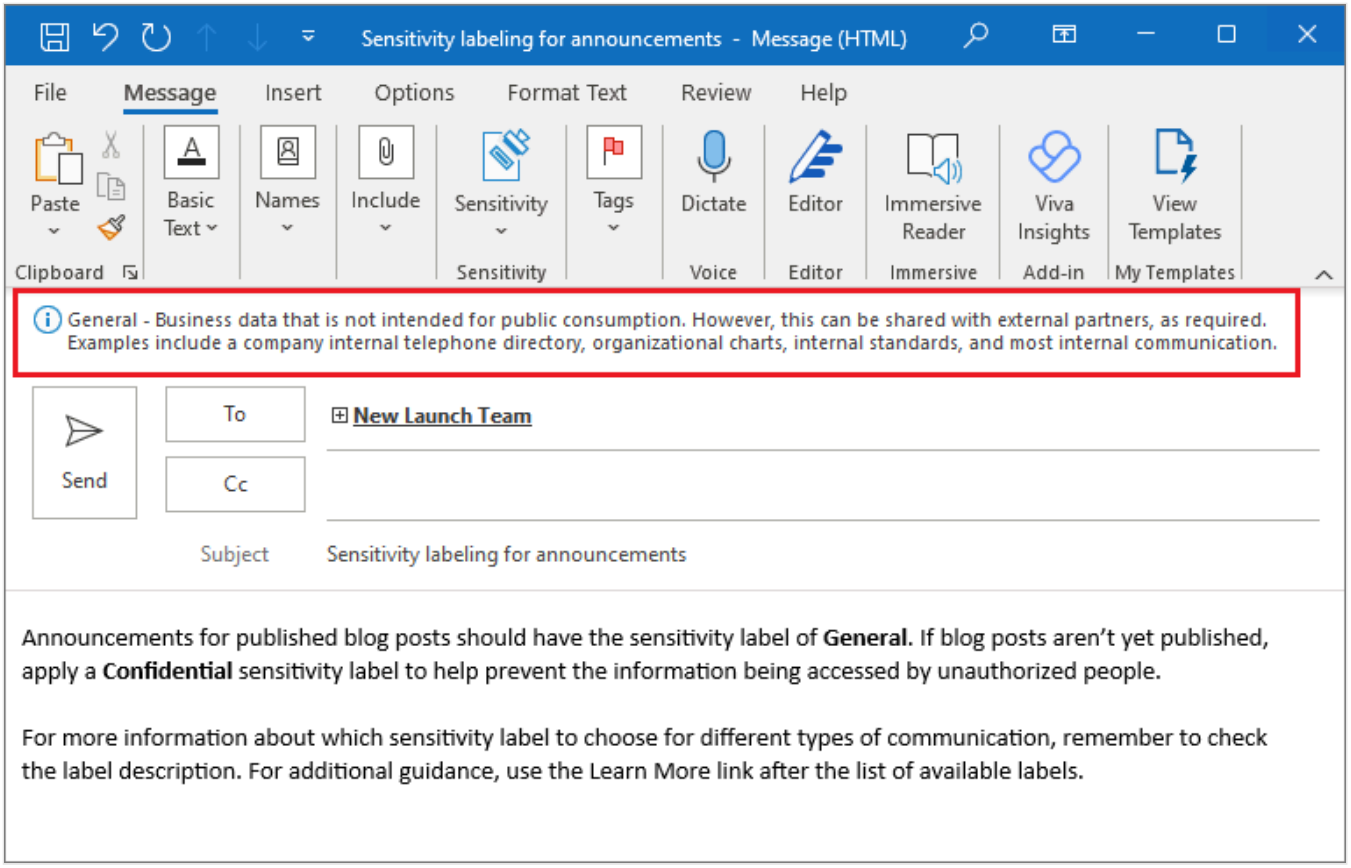
- **Customizable.** You can create categories that fit your organization's specific business needs for different levels of sensitive content. For example, you might use categories like Personal, Public, General, Confidential, and Highly Confidential.

- **Clear text.** Because the label is stored in clear text in the metadata for files and emails, third-party apps and services can read it and then apply their own protective actions, if required.

- **Persistent.** Because the label is stored in metadata for files and emails, the label stays with the content, no matter where you save or store it. The label identification that's unique to your organization becomes the basis for applying and enforcing policies that you configure.

When users in your organization view an applied sensitivity label, it appears like a tag on apps and can be easily integrated into their existing workflows. Users from other organizations or guests can't see your sensitivity labels in apps.

The following example shows an opened email where another user has applied the sensitivity label named **General**, which doesn't apply encryption. The label description supplied by the admin displays more detail to users about the category of data identified by this sensitivity label.



> **Note**
>
> Don't confuse sensitivity labels with Outlook's built-in sensitivity levels that indicate the sender's intention but can't provide data security.

Each item that supports sensitivity labels can have a single sensitivity label applied to it from your organization. Documents and emails can have both a sensitivity label and a retention label applied to them.

What sensitivity labels can do

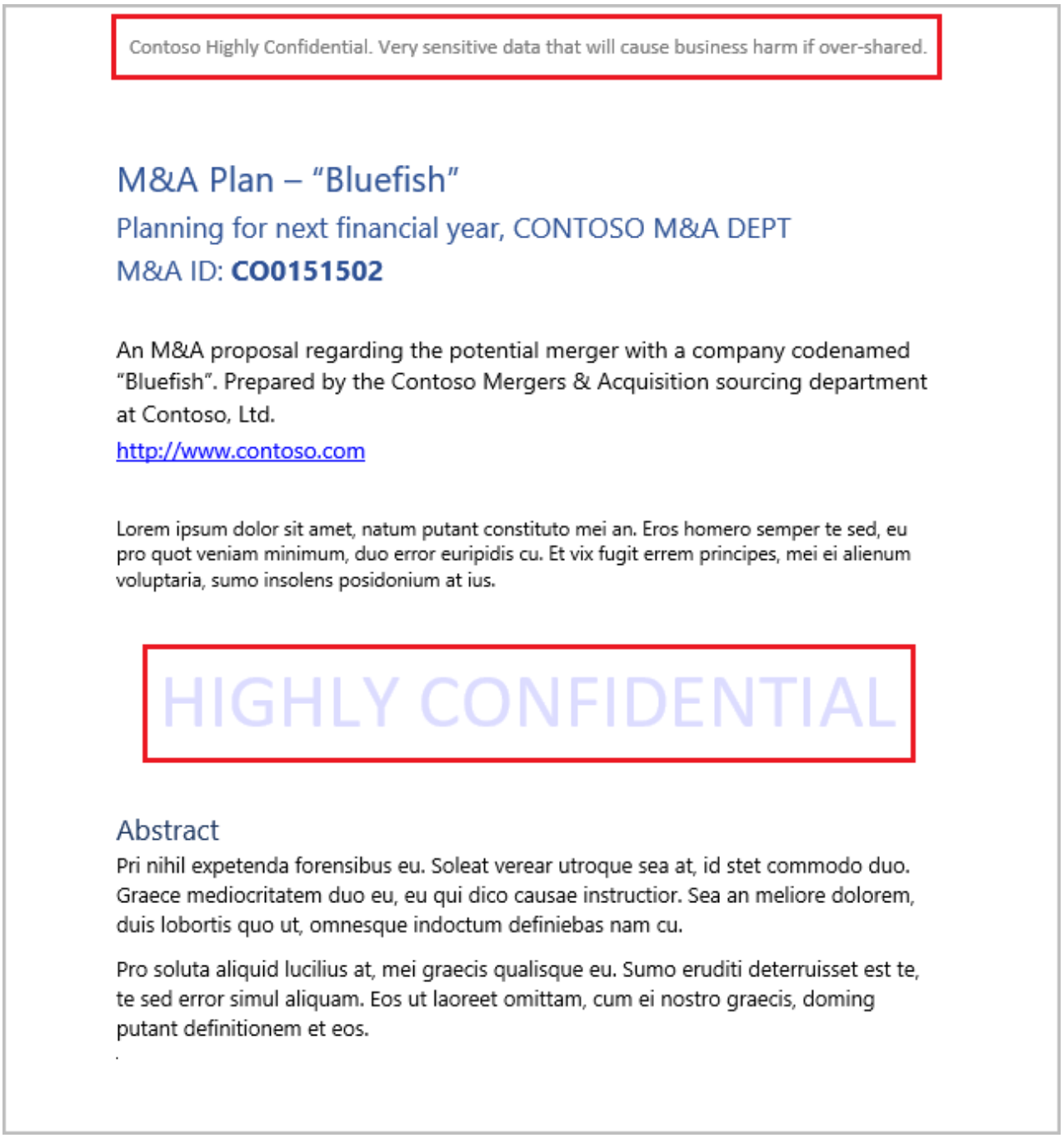
After you apply a sensitivity label to content, such as an email, meeting invite, document, or Loop page, the label enforces any configured protection settings on the content. You can configure a sensitivity label to:

- **Control access to content by using encryption** for emails, meeting invites, and documents to prevent unauthorized people from accessing this data. You can also choose

which users or groups have permissions to perform which actions and for how long. For example, you can choose to allow all users in your organization to modify a document while a specific group in another organization can only view it. Alternatively, instead of administrator-defined permissions, you can allow your users to assign permissions to the content when they apply the label.

For more information about the encryption-based access control settings when you create or edit a sensitivity label, see Restrict access to content by using encryption in sensitivity labels.

- **Mark the content** when you use Office apps, by adding watermarks, headers, or footers to email, meeting invites, or documents that have the label applied. You can apply watermarks to documents and Loop components and pages, but not to email or meeting invites. For other apps, check their documentation. Example header and watermark:



Content markings also support variables. For example, insert the label name or document name into the header, footer, or watermark. For more information, see Content markings with variables.

Need to check when content markings are applied? See When Office apps apply content marking and encryption.

If you have templates or workflows that are based on specific documents, test those documents with your chosen content markings before you make the label available for users. Some string length restrictions to be aware of:

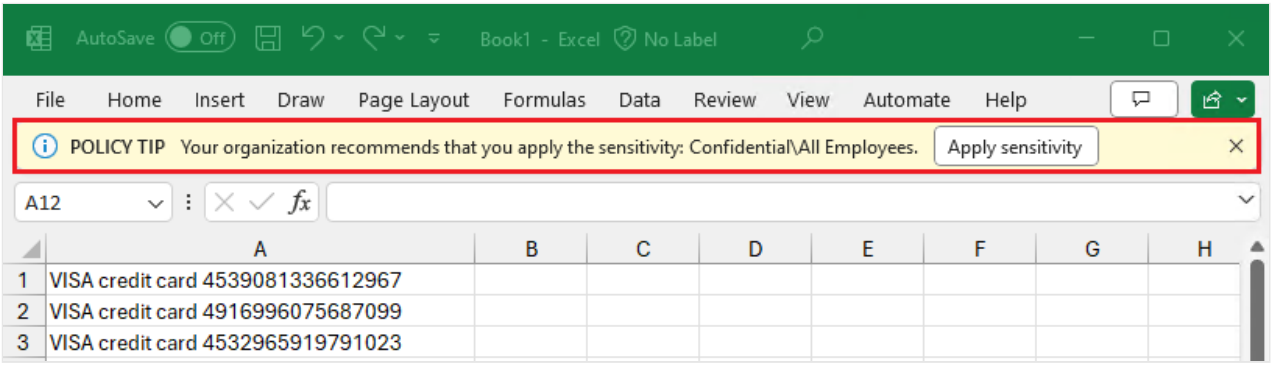
Watermarks are limited to 255 characters. Headers and footers are limited to 1,024 characters, except in Excel. Excel has a total limit of 255 characters for headers and footers

but this limit includes characters that aren't visible, such as formatting codes. If that limit is reached, the string you enter isn't displayed in Excel.

- **Protect content in containers such as sites and groups** when you enable the capability to use sensitivity labels with containers.

  You can't configure protection settings for groups and sites until you enable this capability. This label configuration doesn't result in individual items being automatically labeled but instead, the label settings protect content by controlling access to the container where content can be stored. These settings include privacy settings, external user access and external sharing, access from unmanaged devices, private teams discoverability, and sharing controls for channels.

- **Apply the label automatically to files and emails, or recommend a label.** Choose how to identify sensitive information that you want labeled, and the label can be applied automatically, or you can prompt users to apply the label that you recommend by using a policy tip. If you recommend a label, you can customize the prompt but the following example shows the automatically generated text:

  For more information about the **Auto-labeling for files and emails** settings when you create or edit a sensitivity label, see Automatically apply a sensitivity label to Microsoft 365 data for Office apps, and Labeling in Microsoft Purview Data Map.

- **Set the default sharing link type** for SharePoint sites and individual documents. To help prevent users from oversharing, set the default scope and permissions for when users share documents from SharePoint and OneDrive.

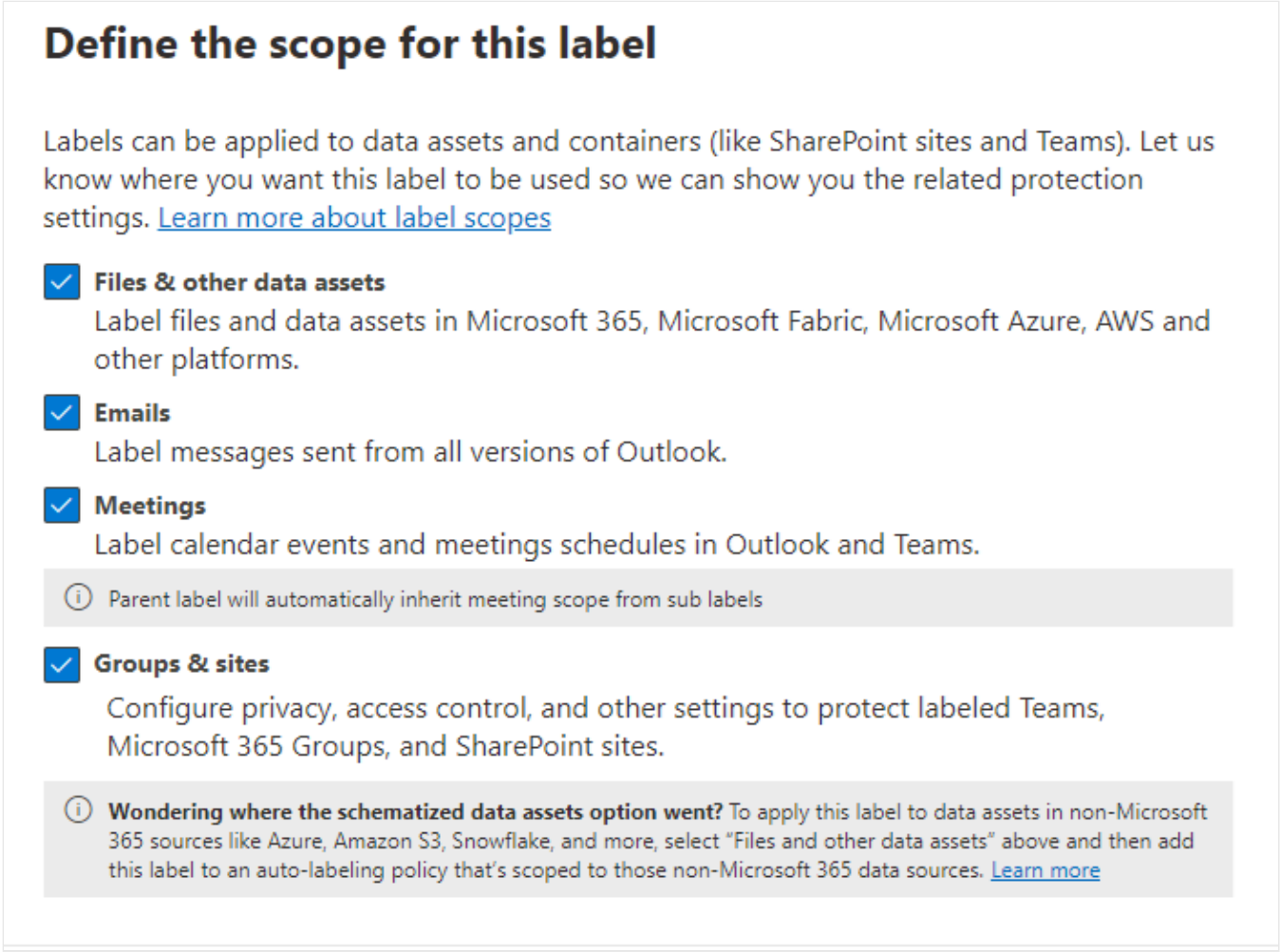
For more label configurations, see Manage sensitivity labels for Office apps.

Label scopes

When you create a sensitivity label, you choose the label's scope. The scope determines two things:

- Which label settings you can configure for that label
- The availability of the label to apps and services, which includes whether users can see and select the label

This scope configuration means you can have sensitivity labels that are just for items such as files, emails, and meetings, and can't be selected for groups and sites. Similarly, you can have sensitivity labels that are just for groups and sites and can't be selected for items.



By default, the **Files & other data assets** scope is always selected for a new label. As well as files for Office, Loop, and Power BI, it includes items from Microsoft Fabric, and data assets for Microsoft Purview Data Map when you extend your sensitivity labels beyond Microsoft 365. For more information about which items support sensitivity labels:

- For Office files: Office file types supported
- For Microsoft Loop: Sensitivity labels to protect Loop components and pages
- For Power BI: Supported export paths
- For Microsoft Fabric: Information protection in Microsoft Fabric
- For Data Map: Data sources that connect to Data Map

For other apps, check their documentation. Support for sensitivity labels is continually expanding.

Typically, you select the **Emails** scope together with **Files & other data assets**, because emails often include files as attachments and share the same sensitivity. Many labeling features require both options to be selected, but there might be times when you want a new label to be available for emails only. For more information, see Scope labels to just files or emails.

The scope for **Meetings** includes calendar events, Teams meetings options, and Team chat. You must also select the **Files & other data assets** scope and **Emails** scope for this option. For more information about this labeling scenario, see Use sensitivity labels to protect calendar items, Teams meetings, and chat.

The **Groups & sites** scope becomes available and selected by default when you enable sensitivity labels for containers and synchronize labels. This option lets you protect content in SharePoint sites, Teams sites, Viva Engage communities, and Loop workspaces by labeling those containers but doesn't label the items in them.

> ⓘ **Note**
>
> Items that were previously in the **Schematized data assets** scope are now included in **Files & other data assets**.

If you don't select one or more scopes, you see the first page of the configuration settings for these scopes, but you can't configure the settings. For these pages that have unavailable options, select **Next** to continue to configure settings for the next scope. Or, select **Back** to change the label's scope.

Label priority (order matters)

When you create your sensitivity labels in the Microsoft Purview portal, they appear in a list on the **Labels** page from **Information Protection**. In this list, the order of the labels is important because it sets their priority. You want your most restrictive sensitivity label, such as Highly Confidential, to appear at the **bottom** of the list, and your least restrictive sensitivity label, such as Personal or Public, to appear at the **top**.

You can apply just one sensitivity label to an item such as a document, email, or container. If you use the option that requires your users to provide a justification for changing a label to a lower sensitivity for files, emails, and meetings, the order of this list identifies the lower

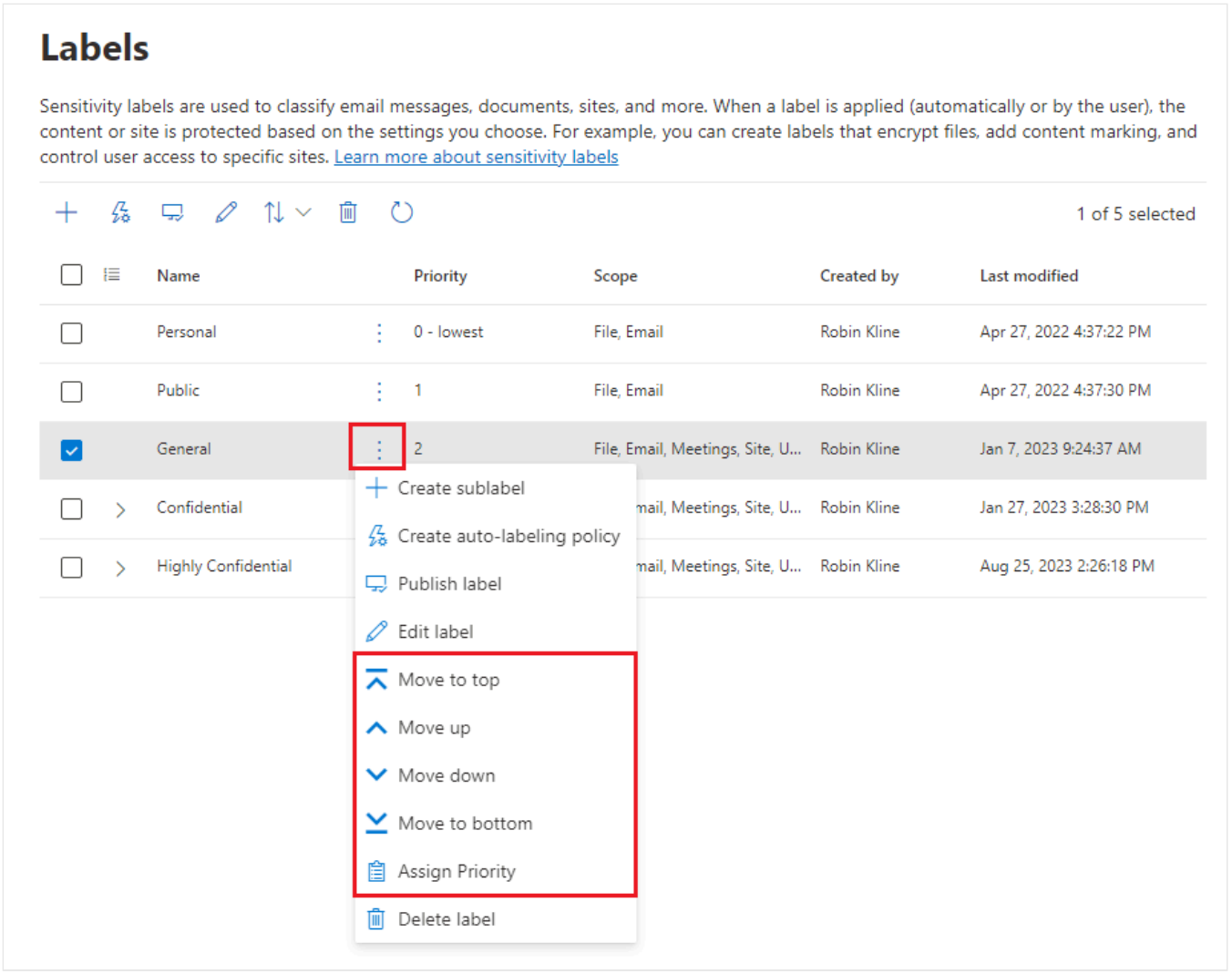
sensitivity. However, this option doesn't apply to sublabels that share the priority of their parent label, and it doesn't apply to data assets or labels that protect containers.

> **Note**
>
> If you're using the **modern label scheme** that replaces parent labels with label groups, sublabels in the same label group also don't support the justification setting.

The priority of sublabels is used with automatic labeling, though. When you configure auto-labeling policies, multiple matches can result for more than one label. Then, the last sensitive label is selected, and then if applicable, the last sublabel. When you configure the auto-labeling label setting (rather than auto-labeling policies) for sublabels themselves, the behavior is a little different when these labels share the same parent label or label group. For example, a sublabel configured for automatic labeling is preferred over a sublabel configured for recommended labeling. For more information, see How multiple conditions are evaluated when they apply to more than one label.

The priority of sublabels is also used with label inheritance from email attachments.

When you select a sensitivity label, you can change its priority by using the options to move it to the top or bottom of the list if it's not a sublabel, move it up or down by one label, or directly assign a priority number. For example:



Sublabels that use parent labels or label groups

You might often want to logically group sensitivity labels in a two-tier hierarchy, to denote the overall sensitivity level but provide labels with different settings. As an example from the default sensitivity labels:

Confidential

- Anyone (unrestricted)
- All Employees
- Trusted People

In apps, users first see **Confidential** and then must select one of the other sensitivity labels, such as **All Employees**. The applied sensitivity label is then **Confidential \ All Employees**.

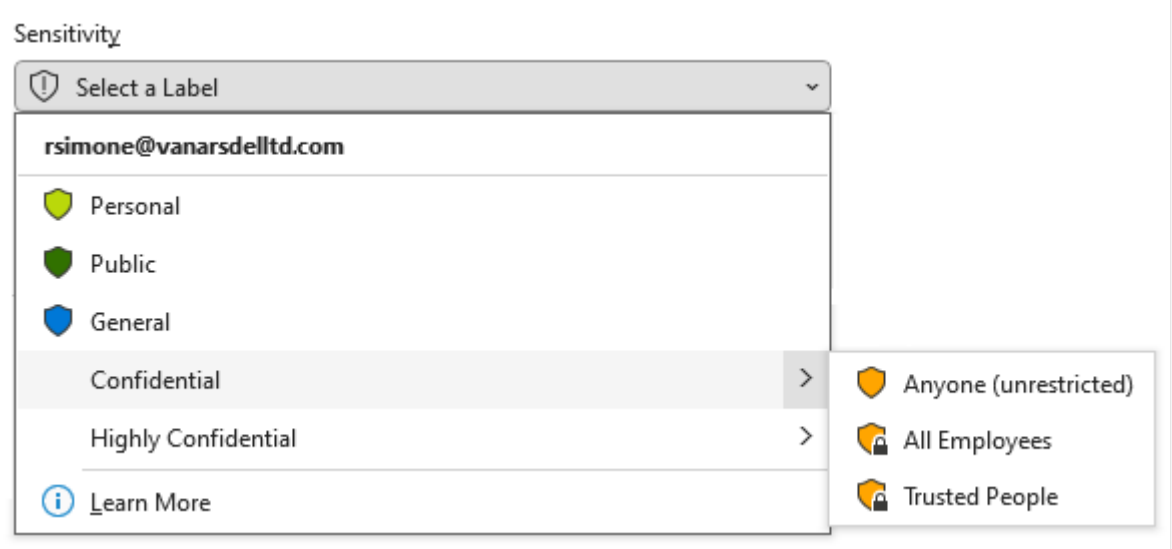
Sensitivity labels from the second tier don't inherit any settings from the first tier label, except the label color. In our example, the **Confidential** label is simply a text label with no protection settings, and can't be applied by itself to content.

The initial implementation to group labels in this two-tier hierarchy used parent labels and sublabels. The configuration of parent labels looked the same as the configuration of sublabels, but behaved differently when sublabels were published to users.

To reduce this complexity, label groups are now replacing parent labels. Label groups look and behave the same to users, but have configuration options only for the settings that they support: the label name, descriptions, color, and priority. You can't publish label groups themselves, only the labels within the groupings.

If you're still using parent labels with sublabels: Don't choose a parent label as the default label, or configure a parent label to be automatically applied (or recommended). If you do, the parent label can't be applied.

Example of how sublabels display for users:



If required, you can manually convert parent labels to label groups. For more information, see Migrate parent sensitivity labels to label groups.

Editing or deleting a sensitivity label

If you delete a sensitivity label from the Microsoft Purview portal, the label isn't automatically removed from content, and any protection settings continue to be enforced on content that had that label applied.

If you edit a sensitivity label, the version of the label that you applied to content is what's enforced on that content.

For detailed information about what happens when you delete a sensitivity label and how this action differs from removing it from a sensitivity label policy, see Removing and deleting labels.

Sensitivity label limitations per tenant

A single tenant can have a high number of sensitivity labels (1,000+), with one exception: If the label applies encryption that specifies the users and permissions, there's a maximum of 500 labels per tenant supported with this configuration. However, as a best practice to lower admin overhead and reduce complexity for your users, try to keep the number of labels to a minimum.

> **Tip**
>
> Real-world deployments show that effectiveness is noticeably reduced when users have more than five main labels or more than five sublabels per main label. You might also find that some applications can't display all your labels when too many are published to the same user.

What label policies can do

After you create your sensitivity labels, you need to publish them to make them available to people and services in your organization. The sensitivity labels can then be applied to Office documents and emails, and other items that support sensitivity labels.

Unlike retention labels, which are published to locations such as all Exchange mailboxes, sensitivity labels are published to users or groups. Apps that support sensitivity labels can then display them to those users and groups as labels that they can apply.

Although the default is to publish labels to all users in your organization, multiple label policies let you publish different sensitivity labels to different users if needed. For example, all users see labels that they can apply for Public, General, and Confidential, but only users in your legal department also see a Highly Confidential label that they can apply.

All users in the same organization can see the name of a sensitivity label applied to content, even if that label isn't published to them. They don't see sensitivity labels from other organizations.

When you configure a publishing label policy, you can:

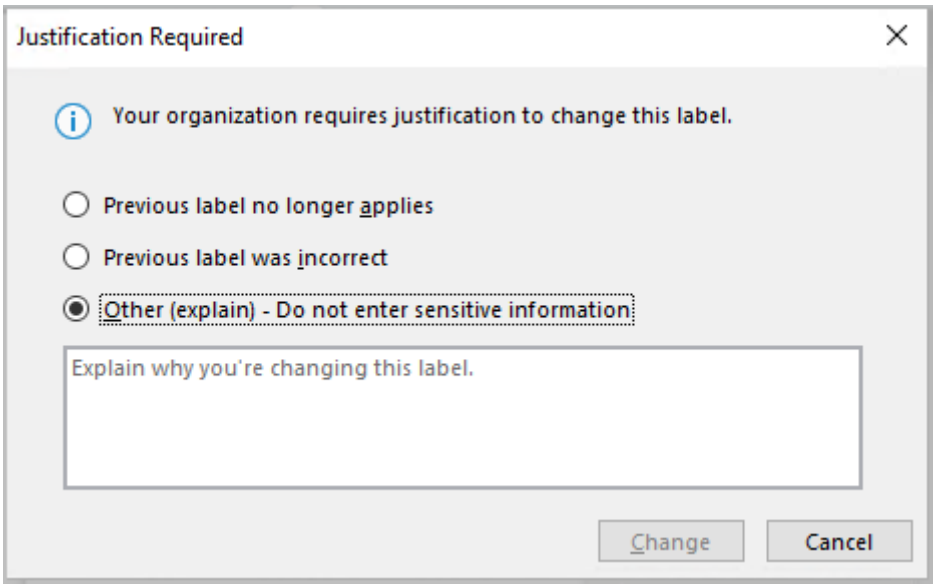
- **Choose which users and groups see the labels.** Labels can be published to any specific user or email-enabled security group, distribution group, or Microsoft 365 group (which can have dynamic membership) in Microsoft Entra ID.

- **Specify a default label** for unlabeled documents and Loop components and pages, emails and meeting invites, new containers (when you've enabled sensitivity labels for containers), and also a default label for Power BI content. You can specify the same label for all five types of items, or different labels. Users can change the applied default sensitivity label to better match the sensitivity of their content or container.

  Consider using a default label to set a base level of protection settings that you want applied to all your content. However, without user training and other controls, this setting can also result in inaccurate labeling. It's usually not a good idea to select a label that applies encryption as a default label to documents. For example, many organizations need to send and share documents with external users who might not have apps that support the encryption or they might not use an account that can be authorized. For more information about this scenario, see Sharing encrypted documents with external users.

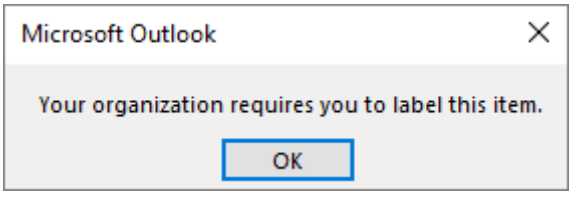
  > **ⓘ Important**
  >
  > When you have **sublabels**, be careful not to configure the parent label as a default label.

- **Require a justification for changing a label.** For files, emails, and meetings, but not for groups and sites (used by Teams and SharePoint), if a user tries to remove a label or replace it with a label that has a lower priority, by default the user must provide a justification to perform this action. For example, a user opens a document labeled Confidential (order number 3) and replaces that label with one named Public (order number 1). For Office apps, this justification prompt is triggered once per app session. When you use the Microsoft Purview Information Protection client, the prompt is triggered for each file. Administrators can read the justification reason along with the label change in activity explorer.



- **Require users to apply a label** for the different types of items and the containers that support sensitivity labels. Also known as mandatory labeling, these options ensure a label must be applied before users can save files and send emails or meeting invites, create new groups or sites, and when they use unlabeled content for Power BI.

  For documents and emails, a label can be assigned manually by the user, automatically as a result of a condition that you configure, or be assigned by default (the default label option previously described). An example prompt when a user is required to assign a label:

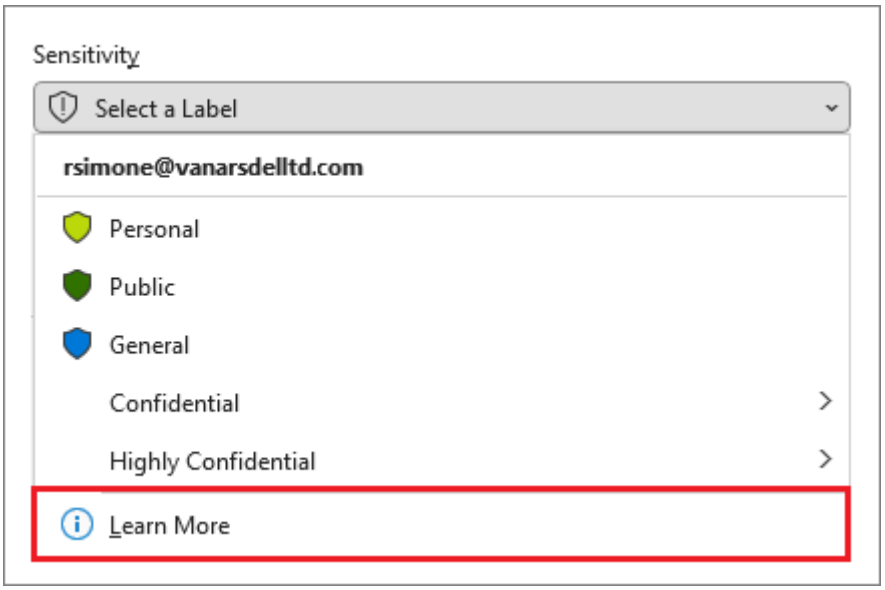
  For more information about mandatory labeling for documents and emails, see Require users to apply a label to their email and documents.

  For containers, a label must be assigned at the time the group or site is created.

  For more information about mandatory labeling for Power BI, see Mandatory label policy for Power BI.

  Consider using this option to help increase your labeling coverage. However, without user training, these settings can result in inaccurate labeling. In addition, unless you also set a corresponding default label, mandatory labeling can frustrate your users with the frequent prompts.

- **Provide help link to a custom help page.** If your users aren't sure what your sensitivity labels mean or how they should be used, you can provide a Learn More URL that appears after the list of available sensitivity labels in the Office apps. For example:



For more label policy configurations, see Manage sensitivity labels for Office apps.

After you create a publishing label policy that assigns new sensitivity labels to users and groups, users start to see those labels in their Office apps. Allow up to 24 hours for the latest changes to replicate throughout your organization.

Label policy priority (order matters)

You make your sensitivity labels available to users by publishing them in a sensitivity label policy that appears in a list on the **Label policies** page. Just like sensitivity labels (see Label priority (order matters)), the order of the sensitivity label policies is important because it reflects their priority: The label policy with lowest priority is shown at the top of the list with the **lowest** order number, and the label policy with the highest priority is shown at the bottom of the list with the **highest** order number.

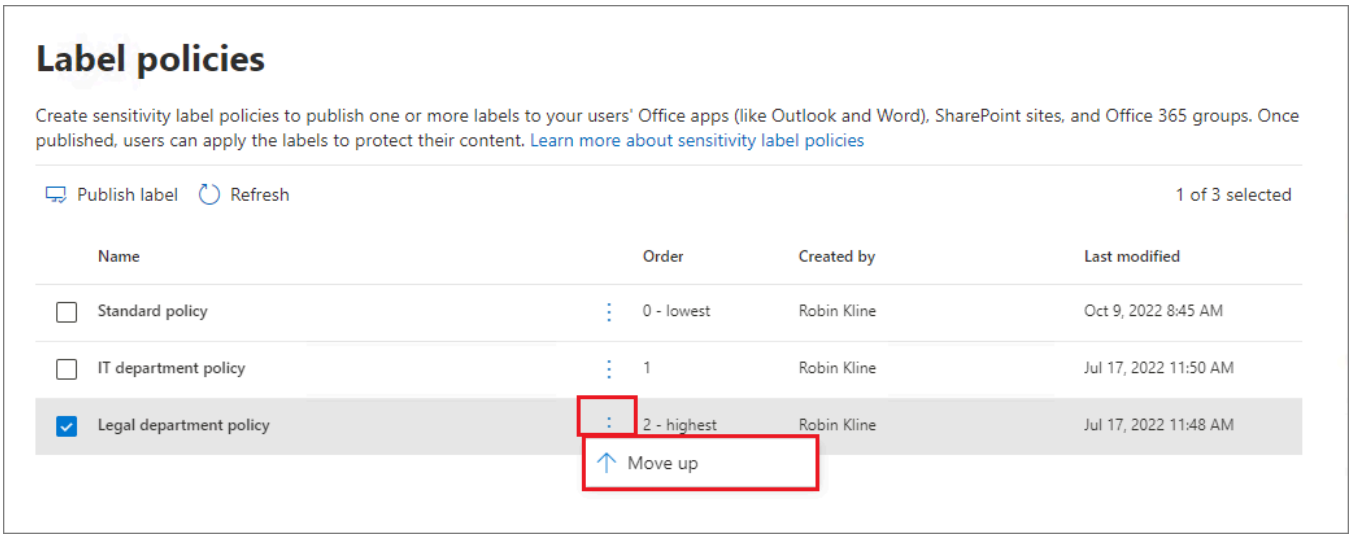
A label policy consists of:

- A set of labels.
- The users and groups that will be assigned the policy with labels.
- The scope of the policy and policy settings for that scope (such as default label for files and emails).

You can include a user in multiple label policies, and the user gets all the sensitivity labels and settings from those policies. If there's a conflict in settings from multiple policies, the settings

from the policy with the highest priority (highest order number) is applied. In other words, the highest priority wins for each setting.

If you don't see the label policy setting behavior that you expect for a user or group, check the order of the sensitivity label policies. You might need to move the policy down. To reorder the label policies, select a sensitivity label policy, choose the Actions ellipsis for that entry, and then select **Move down** or **Move up**. For example:



From the preceding screenshot example that shows three label policies, all users are assigned the standard label policy, so it's appropriate it has the lowest priority (lowest order number of 0). Only users in the IT department are assigned the second policy that has the order number 1. For these users, if there are any conflicts in settings between their policy and the standard policy, the settings from their policy wins because it has a higher order number.

Similarly for users in the legal department, who are assigned the third policy with distinct settings. It's likely these users have more stringent settings, so it's appropriate that their policy has the highest order number. It's unlikely a user from the legal department is in a group that's also assigned to the policy for the IT department. But if they are, the order number 2 (highest order number) ensures the settings from the legal department always take priority if there's a conflict.

> **Note**
>
> Remember: If there's a conflict of settings for a user who has multiple policies assigned to them, the setting from the assigned policy with the highest order number is applied.

Sensitivity labels for Microsoft 365 Copilot and Microsoft 365 Copilot Chat

Microsoft 365 Copilot, Microsoft 365 Copilot Chat, and Copilot agents recognize and use the sensitivity labels that you apply to protect your organization's data. These labels provide an extra layer of protection. For example, in Microsoft 365 Copilot Chat conversations that can reference data from different types of items, users see the sensitivity label with the highest priority (typically, the most restrictive label). Similarly, when Copilot and agents support sensitivity label inheritance, they select the sensitivity label with the highest priority.

If the labels apply encryption from Microsoft Purview Information Protection, Copilot and agents check the usage rights for the user. Then, Copilot and agents return data from an item only if the user is granted permissions to copy (the EXTRACT usage right) from that item.

For more detailed information about how sensitivity labels help protect your data when you use Copilot and other AI apps, see the following articles:

- Microsoft Purview data security and compliance protections for generative AI apps
- Use Microsoft Purview to manage data security & compliance for Microsoft 365 Copilot & Microsoft 365 Copilot Chat
- Considerations to manage Microsoft 365 Copilot for security and compliance

Sensitivity labels and Azure Information Protection

The older labeling client, the Azure Information Protection unified labeling client, is now replaced with the Microsoft Purview Information Protection client to extend labeling on Windows to File Explorer, PowerShell, the on-premises scanner, and provide a viewer for encrypted files.

Office apps support sensitivity labels with a subscription version of Office, such as Microsoft 365 Apps for enterprise.

Sensitivity labels and the Microsoft Information Protection SDK

Because a sensitivity label is stored in the metadata of a document, third-party apps and services can read from and write to this labeling metadata to supplement your labeling deployment. Additionally, software developers can use the Microsoft Information Protection

SDK to fully support labeling and encryption capabilities across multiple platforms. To learn more, see the General Availability announcement on the Tech Community blog.

You can also learn about partner solutions that are integrated with Microsoft Purview Information Protection.

Deployment guidance

For deployment planning and guidance that includes licensing information, permissions, deployment strategy, a list of supported scenarios, and end-user documentation, see Get started with sensitivity labels.

To learn how to use sensitivity labels to comply with data privacy regulations, see Deploy information protection for data privacy regulations with Microsoft 365.

Last updated on 04/15/2026

Get started with sensitivity labels

> *Microsoft Purview service description*

For information about what sensitivity labels are and how they can help you protect your organization's data, see Learn about sensitivity labels.

When you're ready to start protecting your organization's data by using sensitivity labels:

1. **Create the labels.** Create and name your sensitivity labels according to your organization's classification taxonomy for different sensitivity levels of content. Use common names or terms that make sense to your users. If you don't already have an established taxonomy, consider starting with label names such as Personal, Public, General, Confidential, and Highly Confidential. You can then use sublabels to group similar labels by category.

   For each label, specify a tooltip to help users select the appropriate label and consider including specific examples. However, don't make the tooltip so long that users won't read it, and be aware that some apps might truncate long tooltips.

   > ⓘ **Note**
   >
   > For some recommended examples, see the label names and descriptions for the default sensitivity labels. For more guidance about defining a classification taxonomy, see Data classification & sensitivity label taxonomy.

   Always test and tailor your sensitivity label names and tooltips with the people who need to apply them.

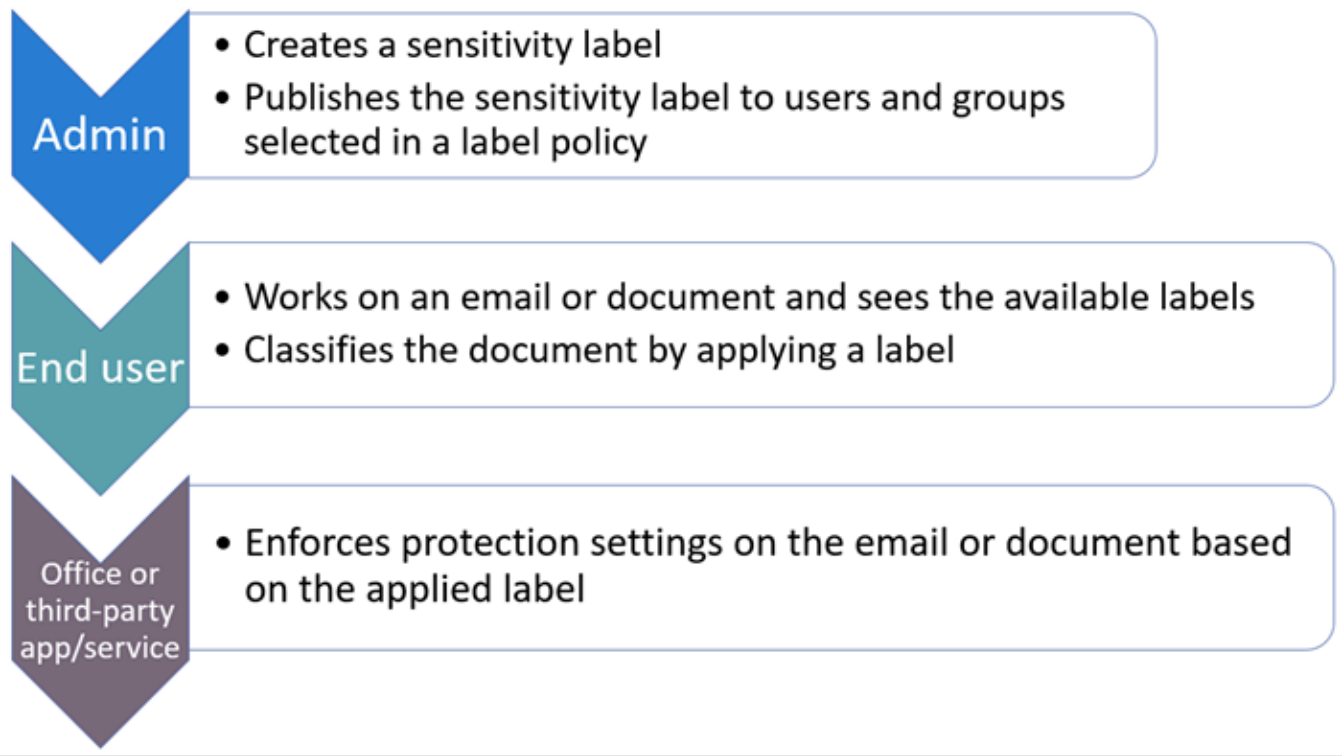
2. **Define what each label can do.** Configure the protection settings you want associated with each label. For example, you might want lower sensitivity content (such as a "General" label) to have just a header or footer applied, while higher sensitivity content (such as a "Confidential" label) should have a watermark and encryption.

3. **Publish the labels.** After your sensitivity labels are configured, publish them by using a label policy. Decide which users and groups should have the labels and what policy settings to use. A single label is reusable—you define it once, and then you can include it

in several label policies assigned to different users. So for example, you could pilot your sensitivity labels by assigning a label policy to just a few users. Then when you're ready to roll out the labels across your organization, you can create a new label policy for your labels and this time, specify all users.

> **Tip**
>
> You might be eligible for the automatic creation of default labels and a default label policy that takes care of steps 1-3 for you. For more information, see Default labels and policies for Microsoft Purview Information Protection.

The basic flow for deploying and applying sensitivity labels is summarized in the following diagram. It identifies the roles and actions of admins, end users, and then Office or third-party apps or services that respectively create labels, apply them, and enforce the label settings.



Licensing and billing requirements for sensitivity labels

A number of different per-user subscriptions support sensitivity labels for Microsoft 365 apps and services, and the licensing requirements for users depend on the features you use. Admins also need a license to manage sensitivity labels.

To see the options for licensing your users to benefit from Microsoft Purview features, see the Microsoft Purview service description. For sensitivity labels, see the Microsoft Purview Information Protection sensitivity labeling section and related PDF download for feature-level licensing requirements.

Specific to non-Microsoft 365 data sources, applying sensitivity labels in the Data Map and using protection policies requires you to set up pay-as-you-go billing in your organization.

> **Tip**
>
> Because features for this solutions use pay-as-you-go billing or per-user licensing for Microsoft Purview, you might find it helpful to use the **Usage center** to help you understand and manage your usage. For more information, see **Manage pay-as-you-go and per-user licensing usage**.

Permissions required to create and manage sensitivity labels

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview **roles and permissions**.

Members of your compliance team who will create and sensitivity labels need permissions to the Microsoft Purview portal. You can use the following role groups:

- **Information Protection**
- **Information Protection Admins**
- **Information Protection Analysts**
- **Information Protection Investigators**
- **Information Protection Readers**

For an explanation of each one, and the roles that they contain, select a role group in the Microsoft Purview portal, and then review the description in the flyout pane. Or, see Role groups in Microsoft Defender for Office 365 and Microsoft Purview compliance.

Alternatively to using the default role groups, you can create a new role group and add the **Sensitivity Label Administrator** role to this group. For a read-only role, use **Sensitivity Label Reader**.

Another option is to add users to the **Compliance Data Administrator**, **Compliance Administrator**, or **Security Administrator** role group.

For configuration instructions, see Permissions in the Microsoft Purview portal.

These permissions are required only to create and configure sensitivity labels and their label policies. They are not required to apply the labels in apps or services. If additional permissions are needed for specific configurations that relate to sensitivity labels, those permissions will be listed in their respective documentation instructions.

Support for administrative units

Sensitivity labels support administrative units that have been configured in Microsoft Entra ID:

- You can assign administrative units to members of role groups that are used with Microsoft Purview Information Protection. Edit these role groups and select individual members, and then the **Assign admin units** option to select administrative units from Microsoft Entra ID. These administrators are now restricted to managing just the users in those administrative units.

- You can define the initial scope of sensitivity label policies and auto-labeling policies when you create or edit these policies. When you select administrative units for Exchange and OneDrive, only the users in those administrative units will be eligible for the policy. When you select administrative units for SharePoint, only the sites in those administrative units will be eligible for the policy.

- Protection policies don't support administrative units.

The configuration of administrative units and the accuracy of their membership is a Microsoft Entra ID dependency. Although the main purpose of administrative units is to ensure the security best practice of least privileges, using administrative units for your labeling policies can simplify their configuration and maintenance.

For example, your organization has configured administrative units for specific countries, and you need to publish a new sensitivity label just to users in France and assign specific policy settings to these users:

1. You sign in to the Microsoft Purview portal. Your account is a member of the **Information Protection Admins** role group, and your account in that role group has been assigned administrative units for France, Germany, and Spain.

2. When you create the sensitivity label policy, you see just three administrative units and select the one for France, keeping the default of all users and groups.

   This configuration automatically scopes the policy to all users in France. You don't need to worry about which groups to select or manually select users. You also don't need to worry about changing the policy when there are new users in France, because this change is handled by the administrative unit in Microsoft Entra.

For more information about how Microsoft Purview supports administrative units, see Administrative units.

Deployment strategy for sensitivity labels

A successful strategy to deploy sensitivity labels for an organization is to create a working virtual team that identifies and manages the business and technical requirements, proof of concept testing, internal checkpoints and approvals, and final deployment for the production environment.

Using the table in the next section, we recommend identifying your top one or two scenarios that map to your most impactful business requirements. After these scenarios are deployed, return to the list to identify the next one or two priorities for deployment.

> **Tip**
>
> To help accelerate your deployment, use the **Overview** page from **Information Protection** in the **Microsoft Purview portal** to see labeling recommendations and reports specific to your organization, and more.

Common scenarios for sensitivity labels

All scenarios require you to Create and configure sensitivity labels and their policies.

Expand table

I want to ...	Documentation
Manage sensitivity labels for Office apps so that content is labeled as it's created—includes support for manual labeling on all platforms	Manage sensitivity labels in Office apps
Extend labeling to Windows File Explorer and PowerShell	Extend sensitivity labeling on Windows
Encrypt documents and emails with sensitivity labels and restrict who can access that content and how it can be used	Restrict access to content by using sensitivity labels to apply encryption
Ensure that documents from SharePoint sites retain their current permissions when files are downloaded, and protect them against move and copy	Configure SharePoint with a sensitivity label to extend permissions to downloaded documents
Protect Teams meetings, from meeting invites and responses, to protecting the meeting itself and related chat	Use sensitivity labels to protect calendar items, Teams meetings and chat
Protect Teams voicemail messages	Enable protected voicemail in your organization
Enable sensitivity labels for Office for the web, with support for coauthoring, eDiscovery, data loss prevention, search—even when documents are encrypted	Enable sensitivity labels for files in SharePoint and OneDrive
Use sensitivity labels with Microsoft Loop to protect Loop content	Use sensitivity labels with Microsoft Loop
Files in SharePoint to be automatically labeled with a default sensitivity label	Configure a default sensitivity label for a SharePoint document library
Use co-authoring and AutoSave in Office desktop apps when documents are encrypted	Enable co-authoring for files encrypted with sensitivity labels
Automatically apply sensitivity labels to documents and emails	Automatically apply a sensitivity label to Microsoft 365 data
Use sensitivity labels to protect content in Teams, SharePoint, Viva Engage, and Loop	Use sensitivity labels with containers (groups and sites)
Use sensitivity labels to configure the default sharing link type for sites and individual documents in SharePoint and OneDrive	Use sensitivity labels to set the default sharing link for sites and documents in SharePoint and OneDrive
Apply a sensitivity label to a document understanding model, so that identified documents in a SharePoint library are automatically classified and protected	Apply a sensitivity label to a model in Microsoft Syntex
Prevent or warn users about sharing files or emails with a specific sensitivity label	Use sensitivity labels as conditions in DLP policies

I want to ...	Documentation
Apply a sensitivity label to a file when I receive an alert that content containing personal data is being shared and needs protection	Investigate and remediate alerts in Privacy Risk Management
Apply a retention label to retain or delete files or emails that have a specific sensitivity label	Automatically apply a retention label to retain or delete content
Discover, label, and protect files stored in data stores that are on premises	Deploying the information protection scanner to automatically classify and protect files
Discover, label, and protect files stored in data stores that are in the cloud	Discover, classify, label, and protect regulated and sensitive data stored in the cloud
Label SQL database columns by using the same sensitivity labels as those used for files and emails so that the organization has a unified labeling solution that can continue to protect this structured data when it's exported	Data Discovery & Classification for Azure SQL Database, Azure SQL Managed Instance, and Azure Synapse Analytics SQL Data Discovery and Classification for SQL Server on-premises
Apply and view labels in Power BI, and protect data when it's saved outside the service	Sensitivity labels in Power BI
Label items in Microsoft Fabric, to restrict access and grant read-only or full control permissions to users	Create and manage protection policies for Fabric
Monitor and understand how sensitivity labels are being used in my organization	Reports overview Additionally, view the information protection reports from the Microsoft Purview portal
Audit sensitivity labeling in my organization	Audit log activities - Sensitivity label activities
Identify content for eDiscovery cases based on applied sensitivity labels, such as Confidential and Highly Confidential	Use the condition builder to create search queries in eDiscovery
Extend sensitivity labels to third-party apps and services	Microsoft Information Protection SDK
Extend sensitivity labels across content in my Microsoft Purview Data Map assets, such as Azure Blob Storage, Azure Files, Azure Data Lake Storage, and multi-cloud data sources	Labeling in Microsoft Purview Data Map

End-user documentation for sensitivity labels

The most effective end-user documentation will be customized guidance and instructions you provide for the label names and configurations you choose. You can use the label policy setting **Provide users with a link to a custom help page** to specify an internal link for this documentation.

In Office apps, users can then easily access your customized help from the **Sensitivity** button, **Learn More** menu option. If you don't specify a link in the policy setting, users don't see this menu option in their apps.

From the Microsoft Purview Information Protection file labler in Windows File Explorer, users can access the same customized help from **Help and Feedback** > **Tell Me More** in the file labeler dialog box.

You can use the following resources for basic instructions:

- Apply sensitivity labels to your files
  - Custom permissions for files
- Apply sensitivity labels to email messages in Outlook
- Known issues with sensitivity labels in Office
- Automatically apply or recommend sensitivity labels to your files and emails in Office
  - Known issues with automatically applying or recommending sensitivity labels
- Create protected PDFs from Office files
- Sensitivity labels for Teams meetings
- Apply sensitivity labels to your Outlook meetings.
- Sensitivity labels in Viva Engage
- Label and protect files in File Explorer in Windows

If your sensitivity labels apply encryption for PDF documents, these documents can be opened with Microsoft Edge on Windows or Mac. For more information, see View protected PDFs using Microsoft Edge on Windows or Mac.

Last updated on 04/16/2026

Create and configure sensitivity labels and their policies

> *Microsoft Purview service description*

All Microsoft Purview Information Protection solutions are implemented by using sensitivity labels. To create and publish these labels, use the Microsoft Purview portal.

First, create and configure the sensitivity labels that you want to make available for apps and other services. For example, the labels you want users to see and apply from Office apps.

Then, create one or more label policies that contain the labels and policy settings that you configure. When you publish the label policy to your chosen users:

- The labels become visible to these users in their apps that support sensitivity labels
- The policy settings are applied to these users

> **Note**
>
> You might be eligible for the automatic creation of default labels and a default label policy. Even if you have some labels, you might find it useful to see the configuration of these default labels that we're creating for new customers. For example, you can make the same manual configurations to help accelerate your own label deployment.
>
> For more information, see Default labels and policies for Microsoft Purview Information Protection.

Before you begin

To make sure you have permissions to create and manage sensitivity labels, see Permissions required to create and manage sensitivity labels.

Create and configure sensitivity labels

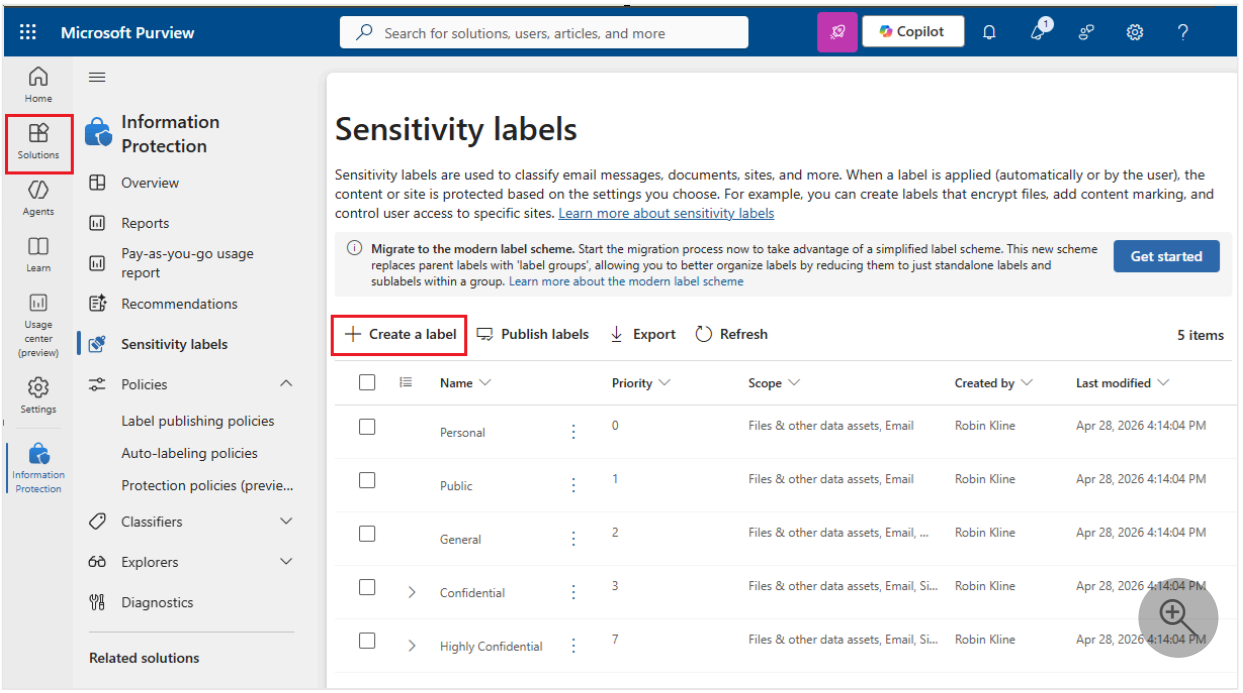
Choose the **Modern label scheme** tab if you don't already have any sensitivity labels or your tenant was created beginning October 1, 2025. Or, you manually migrated to the modern label

scheme. In both cases, you see a green message bar on the **Sensitivity labels** page that confirms your tenant is successfully migrated to the modern label scheme.

Classic label scheme

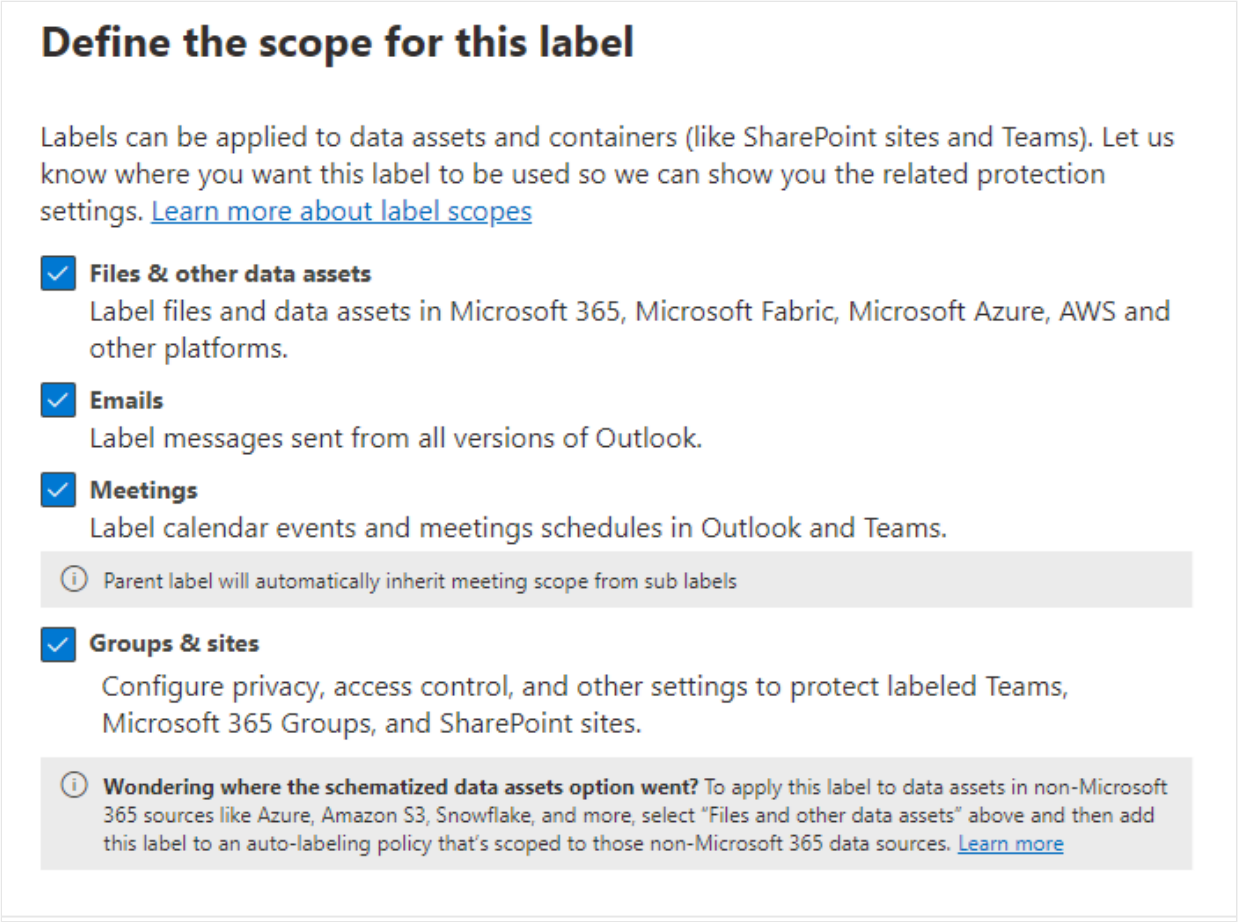
1. Sign in to the Microsoft Purview portal > **Solutions** > **Information Protection** > **Sensitivity labels**

2. On the **Sensitivity labels** page, select + **Create a label** to start the new sensitivity label configuration. For example, from the Microsoft Purview portal:



> **Note**
>
> If you need guidance to create new labels, see Get started with sensitivity labels.

3. On the **Define the scope for this label** page, the options selected determine the label's scope for the settings that you can configure and where they'll be visible when they're published:



You can't select **Meetings** unless **Files & other data assets** and **Emails** are also selected. For more information about this dependency and extending labels to include meetings, see Use sensitivity labels to protect calendar items, Teams meetings, and chat.

If scope options aren't selected, you see the first page to configure the settings for the scope, but you can't configure them and the labels won't be available for users to select in these apps.

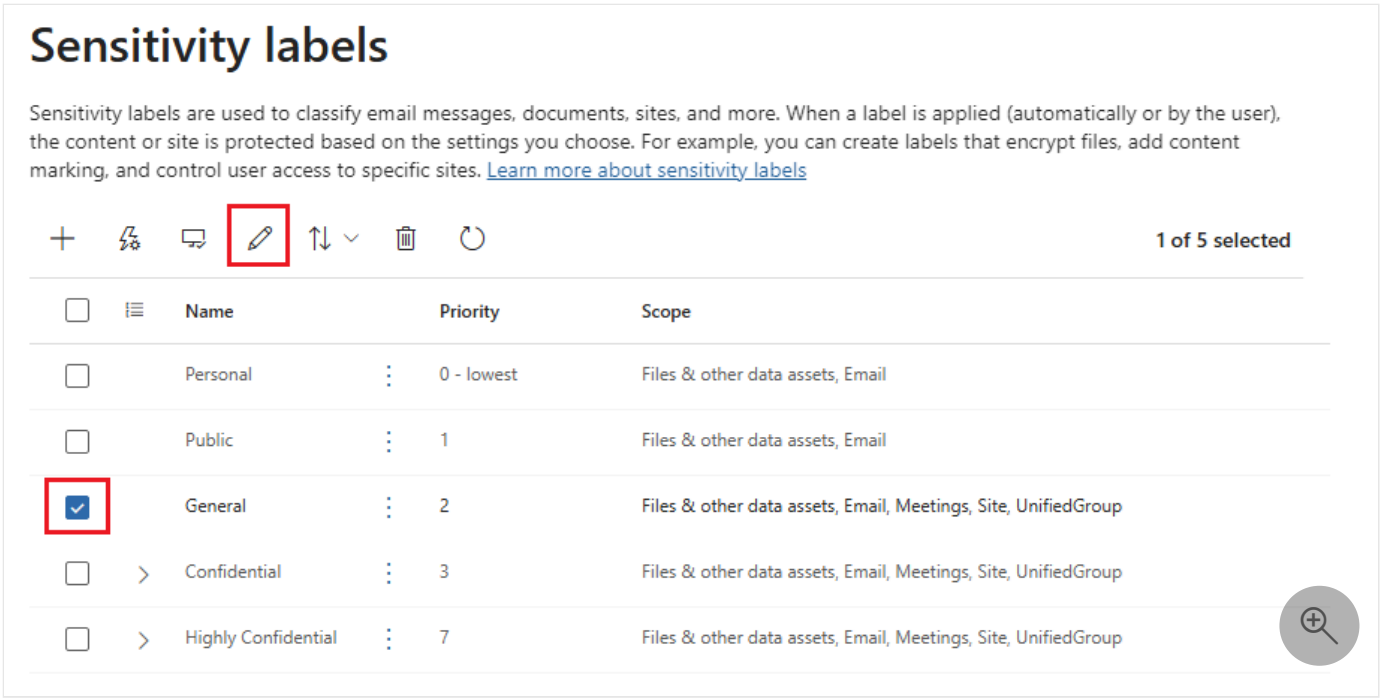
4. Follow the configuration prompts for the label settings.

   For more information about the label settings, see What sensitivity labels can do from the overview information and use the help in the UI for individual settings.

5. Repeat these steps to create more labels. However, if you want to create a sublabel, first select the parent label and select **...** for **Actions**, and then select **Create sublabel**.

6. When you've created all the labels you need, review their order and if necessary, move them up or down. To change the order of a label, select **...** for **Actions**, and then select one of the reordering options, such as **Move up** or **Move down**. For more information, see Label priority (order matters) from the overview information.

To edit an existing label, select it, and then select the **Edit label** icon. For example:



This action starts the **Edit sensitivity label** configuration, which lets you change all the label settings.

Don't delete a label unless you understand the impact for users. For more information, see the Removing and deleting labels section.

> **Note**
>
> If you edit a label that's already published by using a label policy, no extra steps are needed when you finish the configuration. For example, you don't need to add it to a new label policy for the changes to become available to the same users. However, allow up to 24 hours for the changes to replicate to all apps and services.

Until you publish your labels, they won't be available to select in apps or for services. To publish the labels, they must be added to a label policy.

> **Important**
>
> On this **Labels** tab, do not select the **Publish labels** tab (or the **Publish label** button when you edit a label) unless you need to create a new label policy. You need multiple label policies only if users need different labels or different policy settings. Aim to have as few

label policies as possible—it's not uncommon to have just one label policy for the organization.

Additional label settings with Security & Compliance PowerShell

Additional label settings are available with the Set-Label cmdlet from Security & Compliance PowerShell.

For example:

- Use the *LocaleSettings* parameter for multinational deployments so that users see the label name and tooltip in their local language. The following section has an example configuration that specifies the label name and tooltip text for French, Italian, and German.
- Advanced settings supported by built-in labeling are included in the PowerShell documentation. For more help in specifying these PowerShell advanced settings, see the PowerShell tips for specifying the advanced settings section. For additional advanced settings supported by the Microsoft Purview Information Protection client, see the separate documentation for these settings.

Example configuration to configure a sensitivity label for different languages

The following example shows the PowerShell configuration for a label named "Public" with placeholder text for the tooltip. In this example, the label name and tooltip text are configured for French, Italian, and German.

As a result of this configuration, users who have Office apps that use those display languages see their label names and tooltips in the same language. Similarly, if you have the Microsoft Purview Information Protection client installed to label files from File Explorer, users who have those language versions of Windows see their label names and tooltips in their local language when they use the right-click actions for labeling.

For the languages that you need to support, use the Office language identifiers (also known as language tags), and specify your own translation for the label name and tooltip.

Before you run the commands in PowerShell, you must first connect to Security & Compliance PowerShell.

PowerShell

```
$Languages = @("fr-fr","it-it","de-de")
$DisplayNames=@("Publique","Publico","Oeffentlich")
$Tooltips = @("Texte Français","Testo italiano","Deutscher text")
$label = "Public"
$DisplayNameLocaleSettings = [PSCustomObject]@{LocaleKey='DisplayName';
Settings=@(
@{key=$Languages[0];Value=$DisplayNames[0];}
@{key=$Languages[1];Value=$DisplayNames[1];}
@{key=$Languages[2];Value=$DisplayNames[2];})}
$TooltipLocaleSettings = [PSCustomObject]@{LocaleKey='Tooltip';
Settings=@(
@{key=$Languages[0];Value=$Tooltips[0];}
@{key=$Languages[1];Value=$Tooltips[1];}
@{key=$Languages[2];Value=$Tooltips[2];})}
Set-Label -Identity $Label -LocaleSettings (ConvertTo-Json $DisplayNameLocaleSettings
-Depth 3 -Compress),(ConvertTo-Json $TooltipLocaleSettings -Depth 3 -Compress)
```

PowerShell tips for specifying the advanced settings

Although you can specify a sensitivity label by its name, we recommend using the label GUID to avoid potential confusion over specifying the label name or display name. The label name is unique in your tenant, so you can be sure you're configuring the correct label. The display name isn't unique and could result in configuring the wrong label. To find the GUID and confirm the label's scope:

PowerShell

```
Get-Label | Format-Table -Property DisplayName, Name, Guid, ContentType
```

To remove an advanced setting from a sensitivity label, use the same AdvancedSettings parameter syntax, but specify a null string value. For example:

PowerShell

```
Set-Label -Identity 8faca7b8-8d20-48a3-8ea2-0f96310a848e -AdvancedSettings
@{DefaultSharingScope=""}
```

To check your label's configuration, including advanced settings, use the following syntax with your own label GUID:

PowerShell

```
(Get-Label -Identity 8faca7b8-8d20-48a3-8ea2-0f96310a848e).settings
```

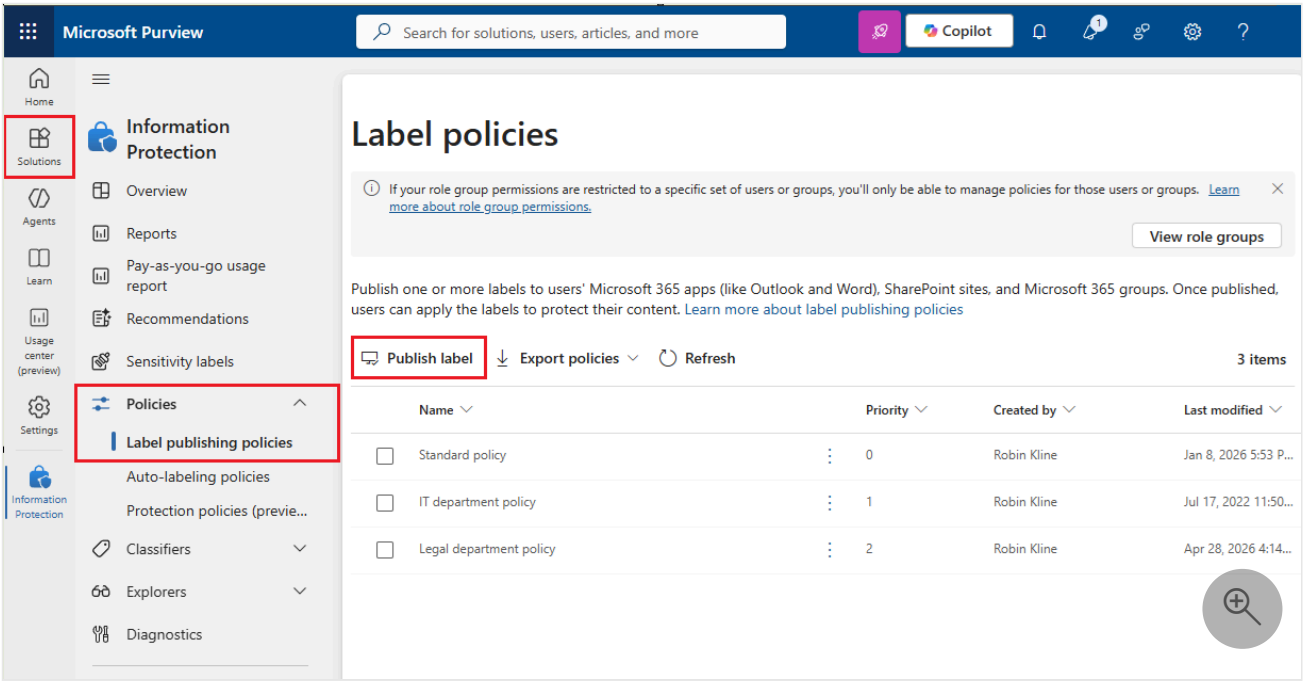
Publish sensitivity labels by creating a label policy

As part of the label policy, you select users and groups to have sensitivity labels and label policy settings. For lower administrative maintenance, groups are recommended rather than individual users. However, if a user is removed from a group you specify, the label policy is automatically removed for that user.

Groups supported: Email-enabled security groups, non-mail enabled security groups, distribution groups, Microsoft 365 groups (that can have dynamic membership) in Microsoft Entra ID.

1. Sign in to the Microsoft Purview portal > **Solutions** > **Information Protection** > **Publishing policies**.

2. On the **Label policies** page, select **Publish label** to start the **Create policy** configuration. For example, from the Microsoft Purview portal:



> **Note**
>
> For more information about the **Export policies** option that can capture detailed information about these policies and all sensitivity labels in your tenant, see Export policy configuration in Microsoft Purview.

If you had default sensitivity labels created for you, you also have a default label publishing policy, which you can modify or supplement with additional policies.

3. On the **Choose sensitivity labels to publish** page, select the **Choose sensitivity labels to publish** link. Select the labels that you want to make available in apps and to services, and then select **Add**.

   > **Important**
   >
   > If you're using the classic label scheme: When you select a sublabel, make sure you also select its parent label.

4. For the **Assign admin units**: If your organization is using administrative units in Microsoft Entra ID, the label policy can be automatically restricted to specific users by selecting administrative units. If your account has been assigned administrative units, you must select one or more administrative units.

   If you don't want to restrict the policy by using administrative units, or your organization hasn't configured administrative units, keep the default of **Full directory**.

5. Follow the prompts to complete the configuration.

   The policy settings that you see match the scope of the labels that you selected. For example, if you selected labels that have just the **Files & other data assets** scope, you don't see the policy settings **Apply this label by default to groups and sites** and **Require users to apply a label to their groups and sites**.

   For more information about these settings, see What label policies can do from the overview information and use the help in the UI for individual settings.

   For labels configured for **Microsoft Purview Data Map assets (preview)**: These labels don't have any associated publishing policy settings.

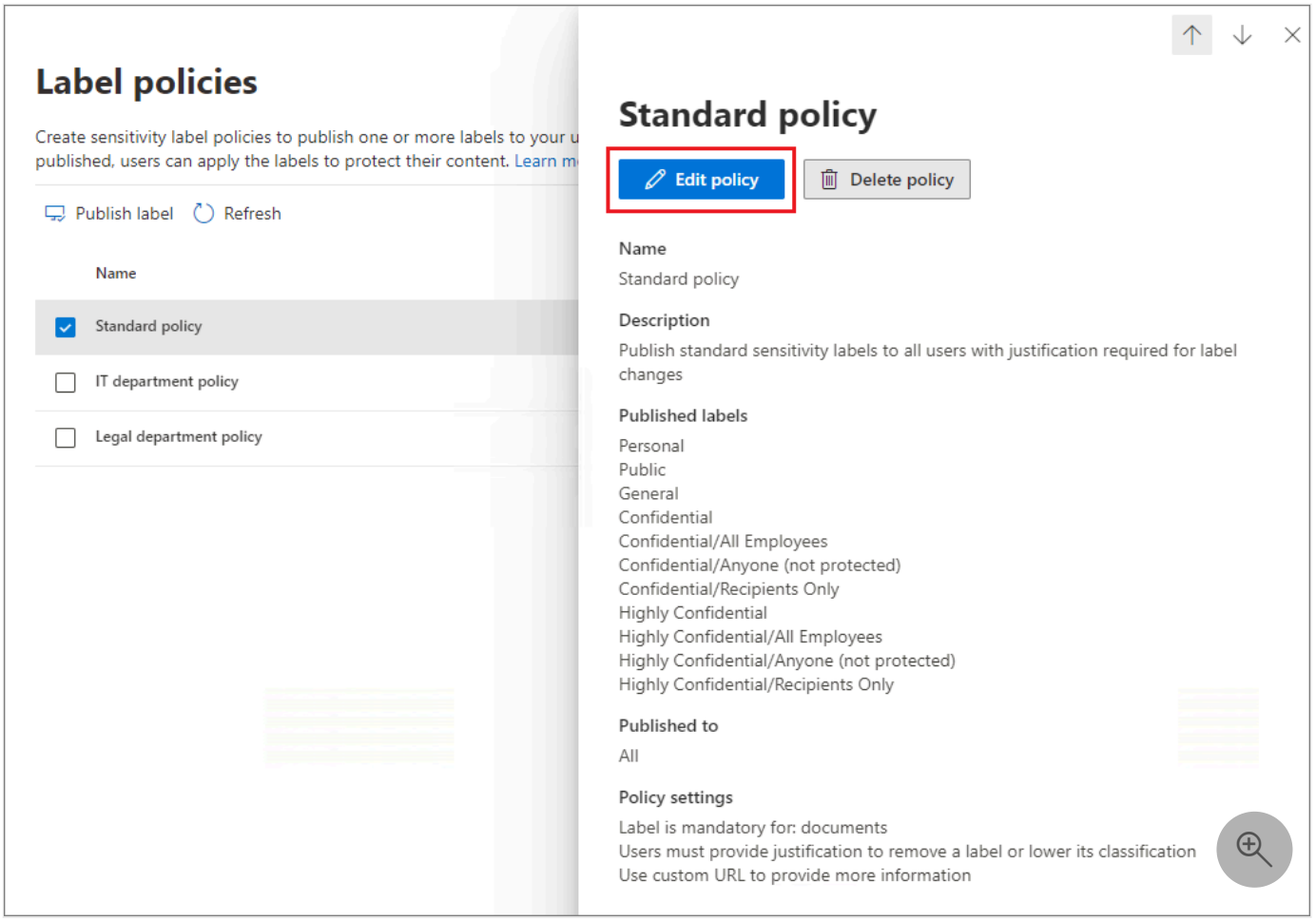
6. Repeat these steps if you need different publishing policy settings for different users or scopes. For example, you want additional labels for a group of users, or a different default label for a subset of users. Or, if you have configured labels to have different scopes.

7. If you have more than one label policy that might result in a conflict for a user, review the policy order and if necessary, move them up or down. To change the order of a label policy, select **...** for **Actions**, and then select one of the reordering options. For more information, see Label policy priority (order matters) from the overview information.

Completing the **Create policy** configuration automatically publishes the label policy. To make changes to a published policy, simply edit it. There's no specific publish or republish action for

you to select.

To edit an existing label policy, select it, and then select the **Edit Policy** button from the flyout pane:



This action starts the **Create policy** configuration, which lets you edit which labels are included and the label settings. When you complete the configuration, any changes are automatically replicated to the selected users and services.

Additional label policy settings with Security & Compliance PowerShell

Additional label policy settings are available with the Set-LabelPolicy cmdlet from Security & Compliance PowerShell.

The cmdlet documentation includes the advanced settings that are supported by built-in labeling. For additional advanced settings supported by the Microsoft Purview Information Protection client, see the separate documentation for these settings.

Create and configure protection policies

Protection policies are currently in preview and apply to items stored outside Microsoft 365. For example, in Microsoft Fabric, Microsoft Azure Storage, and Microsoft Data Lake Storage Gen 2.

For more information, see Authoring and publishing protection policies.

When to expect new labels and changes to take effect

For labels and label policy settings, allow 24 hours for the changes to propagate through the services. There are many external dependencies that each have their own timing cycles, so it's a good idea to wait this 24-hour time period before you spend time troubleshooting labels and label policies for recent changes.

However, there are some scenarios where label and label policy changes can take effect much faster or be longer than 24 hours. For example, for new and deleted sensitivity labels for Word, Excel, and PowerPoint on the web, you might see updates replicate within the hour. But for configurations that depend on populating a new group and group membership changes, or network replication latency and bandwidth restrictions, these changes might take 24-48 hours.

Use PowerShell for sensitivity labels and their policies

You can now use Security & Compliance PowerShell to create and configure all the settings you see in your labeling admin center. This means that in addition to using PowerShell for settings that aren't available in the labeling admin centers, you can now fully script the creation and maintenance of sensitivity labels and sensitivity label policies.

See the following documentation for supported parameters and values:

- New-Label
- New-LabelPolicy
- Set-Label
- Set-LabelPolicy

> **Tip**
>
> When you're configuring advanced settings for a sensitivity label, you might find it helpful to reference the **PowerShell tips for specifying the advanced settings** section on this page.

You can also use Remove-Label and Remove-LabelPolicy if you need to script the deletion of sensitivity labels or sensitivity label policies. However, before you delete sensitivity labels, make

sure you read the next section.

Removing and deleting labels

In a production environment, it's unlikely that you'll need to remove sensitivity labels from a label policy, or delete sensitivity labels. It's more likely that you might need to do one or either of these actions during an initial testing phase. Make sure you understand what happens when you do either of these actions.

Removing a label from a label policy is less risky than deleting it, and can always be added back later if needed. You can't delete a label if it's still in a label policy.

When you remove a label from a label policy so that the label is no longer published to the originally specified users, the next time the label policy is refreshed, users no longer see that label to select in their Office apps. If that label is already applied, the label isn't removed from the content or container. For example, users who are using built-in labeling in desktop apps for Word, Excel, and PowerPoint, still see the applied label name on the status bar. An applied container label continues to protect the Teams or SharePoint site.

In comparison, when you delete a label:

- If the label applied encryption, the underlying protection template is archived so that previously protected content can still be opened. Because of this archived protection template, you can't create a new label with the same name. Although it's possible to delete a protection template by using PowerShell, don't do this unless you're sure you don't need to open content that was encrypted with the archived template.

- For documents stored in SharePoint or OneDrive and you've enabled sensitivity labels for Office files: When you open the document in Office for the web, you won't see the label applied in the app, and the label name no longer displays in the **Sensitivity** column in SharePoint. If the deleted label applied encryption and the services can process the encrypted contents, the encryption is removed. Egress actions from these services result in the same outcome. For example, download, copy to, move to, and open with an Office desktop or mobile app. Although the label information remains in the file's metadata, apps can no longer map the label ID to a display name, so users will assume a file isn't labeled.

- For documents stored outside SharePoint and OneDrive or you haven't enabled sensitivity labels for Office files, and for emails: When you open the content, the label information in the metadata remains, but without the label ID to name mapping, users don't see the applied label name displayed (for example, on the status bar for desktop apps). If the

deleted label applied encryption, the encryption remains and users still see the name and description of the now archived protection template.

- For containers, such as sites in SharePoint and Teams: The label is removed and any settings that were configured with that label are no longer enforced. This action typically takes between 48-72 hours for SharePoint sites, and can be quicker for Teams and Microsoft 365 Groups. Until the label deletion process is complete, users might not be able to open the content that was previously protected by the sensitivity label.

- Be aware that without a GUID-to-name mapping available after you delete a label, deleted labels can display as GUIDs rather than label names in applications such as content explorer and activity explorer.

As with all label changes, removing a sensitivity label from a label policy or deleting a sensitivity label takes time to replicate to all users and services.

Next steps

To configure and use your sensitivity labels for specific scenarios, you might find the following articles helpful:

- Restrict access to content by using encryption in sensitivity labels
- Automatically apply a sensitivity label to Microsoft 365 data
- Use sensitivity labels with containers (groups and sites)
- Enable sensitivity labels for files in SharePoint and OneDrive

To monitor how your labels are being used, view the reports from the Microsoft Purview portal: **Solutions** > **Information Protection** > **Reports**

Last updated on 10/07/2026

Migrate parent sensitivity labels to label groups

> *Microsoft Purview service description*

> **Note**
>
> Migration is automatic if you have a new tenant beginning October 1, 2025 or the migration of parent labels won't create a new sublabel with the same name because you have applicable parent labels. In both these cases, your sensitivity labels will automatically use the modern label scheme to benefit from better organization and reduced deployment complexity.

Previously, parent labels were used to organize sensitivity labels in a two-tier hierarchy. To reduce complexity, their replacement, label groups, don't support settings other than name, description, color, and priority.

For more information about parent labels and label groups, see Sublabels that use parent labels or label groups.

To migrate your parent sensitivity labels to sensitivity label groups, use the Microsoft Purview portal if you see the information banner displayed in the following instructions.

> **Caution**
>
> Label migration is irreversible. We highly recommend that you first test the migration by using a test tenant that has the same sensitivity label configurations as your production tenant.

Before you begin

First, make sure you have permissions to create and manage sensitivity labels. For more information, see Permissions required to create and manage sensitivity labels.

Then, understand the impact of parent labels that if applicable, will be migrated to a sublabel with the same name, in addition to a label group. If this scenario applies to you, you must decide whether you want to leave these new sublabels published after the migration, or you want to unpublish them:

- If you decide to unpublish these new sublabels, you can do this manually, or use PowerShell scripts to automatically unpublish them.
- If you leave these new sublabels published, decide whether you will need to edit them post migration and before users can select them.

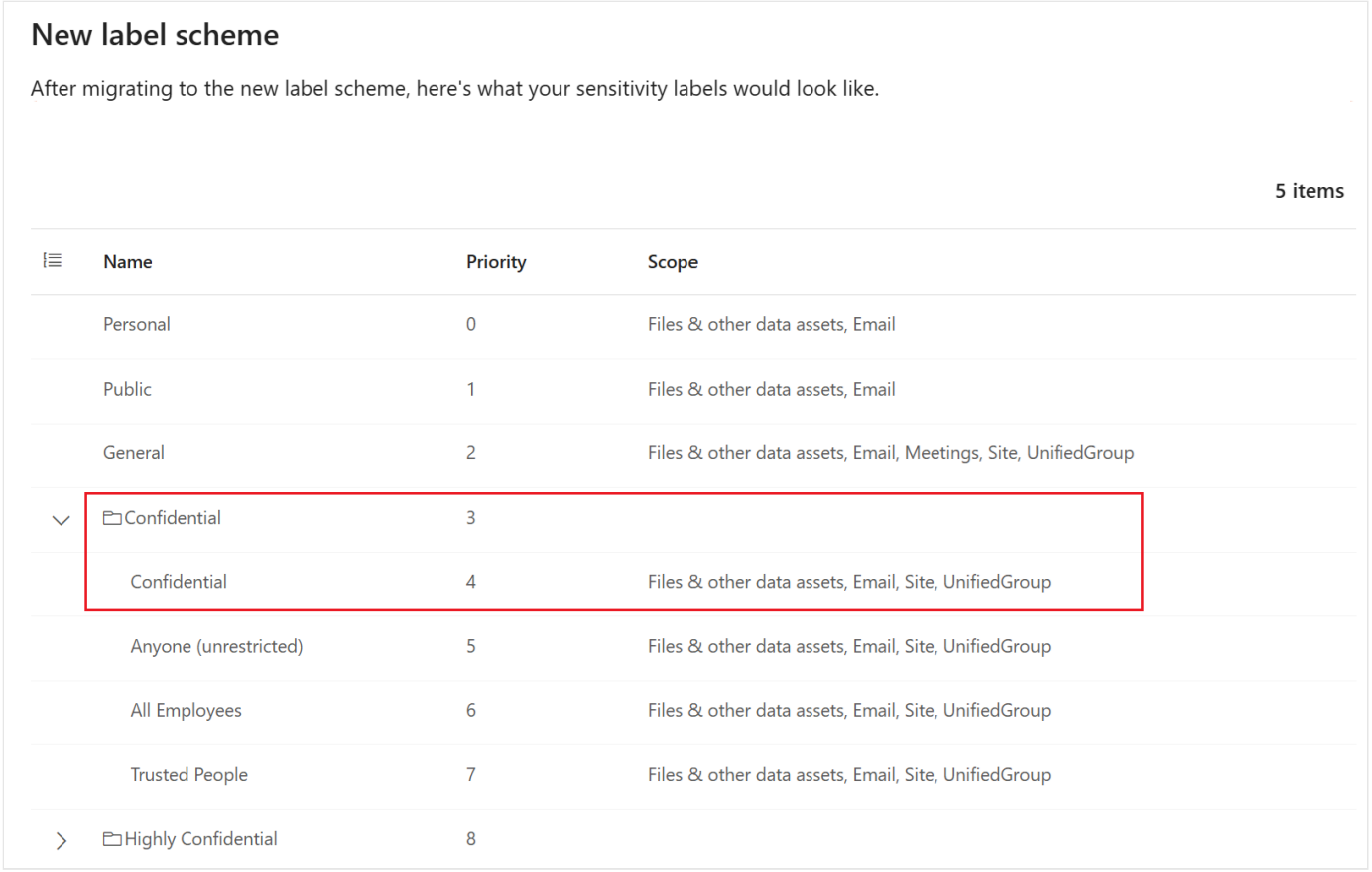
This new sublabel creation ensures that all users and services that might have accessed the parent label as a standalone label will continue to be able to access it after the migration.

- It's configured for label actions that include encryption, content marking, auto-labeling, and PowerShell advanced settings.
- It's included in a different publishing policy from its sublabels.
- It has a different label scope from its sublabels in one or more of the following ways:
  - A sublabel is configured for a label scope that the parent label isn't configured for.
  - Within the same label policy, the parent label has one or more scopes that aren't configured for any of its sublabels.

The resulting sublabel that's created has the same name, label GUID, and settings as the original parent label. Without reconfiguration, this new sublabel will be available to select in a policy where previously it couldn't be selected as a parent label.

> ⓘ **Note**
>
> Because label GUIDs are unique in your tenant, they are used by backend services and configurations to reliably identify a sensitivity label. For example, we recommend you **use the GUID rather than the label name to configure advanced settings**.
>
> To ensure business continuity post migration, the new sublabel inherits the GUID from the applicable parent label, and the new label group is given a new GUID.

You can confirm whether you have applicable parent labels before the labels are migrated when you review the proposed new label scheme as part of the migration process. In the following example, the parent label **Confidential** is identified as an applicable parent label, with the result that a new sublabel of the same name will be created:



PowerShell scripts for unpublishing sublabels from applicable parent labels

These PowerShell commands use Security & Compliance PowerShell.

Script One: Before the migration, run the following command to fetch all policies that publish parent labels.

PowerShell

```
$parents = (Get-Label | ? { $_.ParentId -ne $null } | %{ echo $_.ParentId } | Sort-Object -Unique)
$parentLabels = ($parents | % { Get-Label -Identity $_.Guid })
$policyPublishingParent = @{}
$policies = Get-LabelPolicy; $parentLabels | % { $parent = $_.Name; $policies | ? { $_.Labels -contains $parent } |
% { echo "$($_.Name) policy has published $parent"; $(If ($policyPublishingParent.Keys -notcontains $_.Name) {
$policyPublishingParent[$_.Name] = @() }); $policyPublishingParent[$_.Name] += $parent } }
```

Don't close your PowerShell session but keep it open during the label migration.

Script Two: As soon as migration is complete, run the following script to unpublish the newly created sublabels from applicable parent labels:

PowerShell

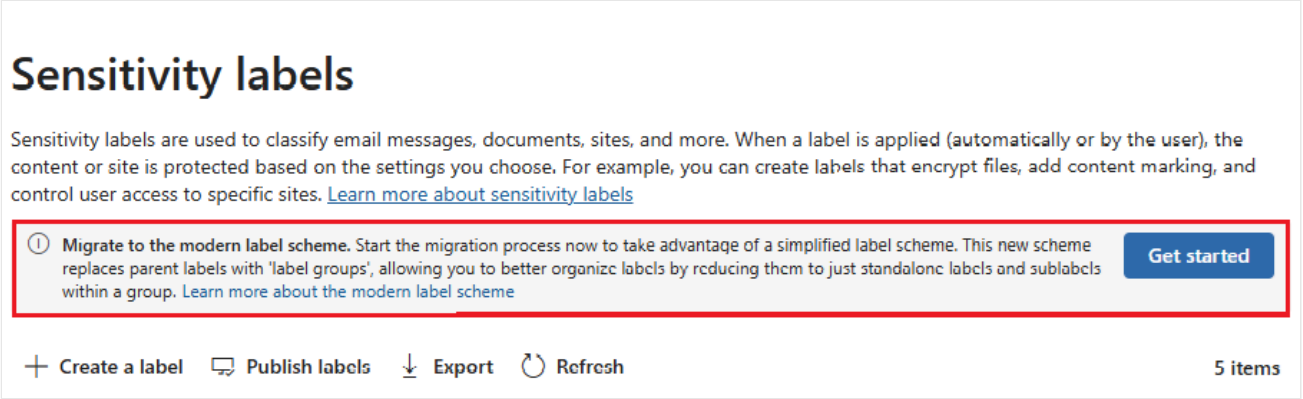
```
$policyPublishingParent.Keys | % { $policyToRemove = $_;
$policyPublishingParent.Item($policyToRemove) | % {
Set-LabelPolicy -Identity $policyToRemove -RemoveLabels $_ }
```

Migrate to the modern label scheme

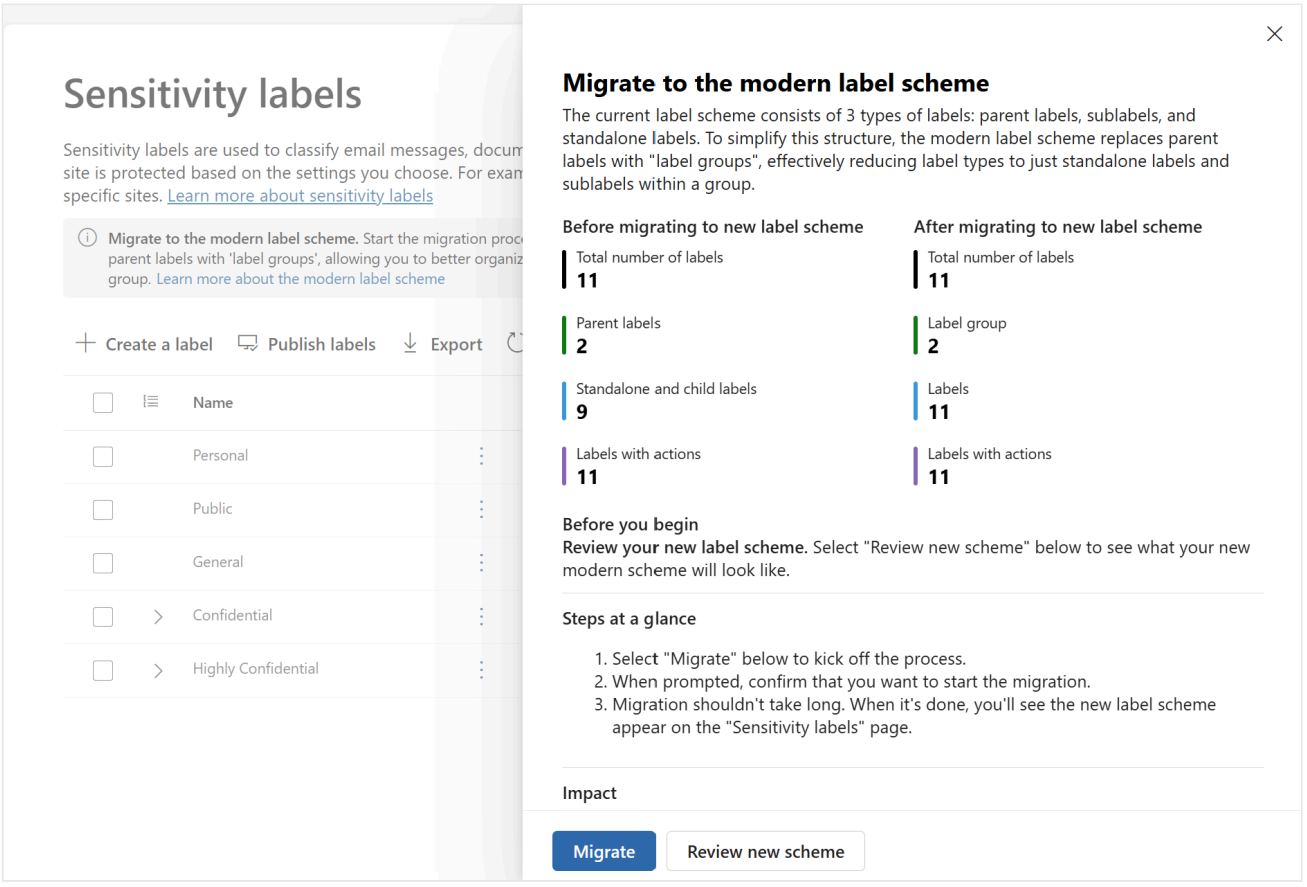
Migration is gradually rolling out, so migrating your parent labels to label groups might not yet be available for your tenant.

1. Sign in to the Microsoft Purview portal > **Solutions** > **Information Protection** > **Sensitivity labels**

2. On the **Sensitivity labels** page, if you see an information banner **Migrate to the modern label scheme**, label migration is available to your tenant. From this banner, select **Get started** to start the migration:



3. On the **Migrate to the modern label scheme** flyout pane, you see a summary of your sensitivity labels before and after the migration. For more details, select the **Review new scheme** button at the bottom of the pane. For example:



4. The **New label scheme** page shows you how your migrated labels will look in the Microsoft Purview portal, with a folder icon next to what used to be parent labels. This page will also alert you to any naming conflicts that need to be resolved before the migration. For example, the **General** parent label has a sublabel that's also named **General**. All sublabels must have a unique name from their parent label.

   You can use the **Download** button to save a copy to a CSV file. The default name is NewLabelSchemecsv.csv.

> **ⓘ Important**
>
> Even if you don't have alerts to resolve, use this opportunity to carefully review the labels and consider making changes to them before you migrate so you can remove or minimize the applicable parent labels that will create a new sublabel if they're not needed. After making changes, restart the migration and check the post-migration information again from this page.
>
> For example, you might have parent labels that you've never used as standalone labels, but are configured for encryption or content markings. Label actions are one of the reasons why migration creates a new sublabel that you might not need.

5. When you're ready to migrate your labels, select the **Migrate** button.

> **Note**
>
> If you have applicable parent labels that will create new sublabels and you decided to use a script to automatically unpublish them, run Script One before you select **Migrate**.

The migration process doesn't take long, typically just a few minutes to an hour. However, as a precaution, allow up to 24 hours. While the migration is in progress, we don't recommend making any changes to the label configuration because those changes will likely be lost.

When the migration is complete, the browser refreshes, and the new label scheme is displayed with a message to confirm that the migration is complete.

> **Note**
>
> If you have applicable parent labels that will create new sublabels and you decided to use a script to automatically unpublish them, now run Script Two.

Next steps

For instructions to manage your migrated labels, see Create and configure sensitivity labels. Review your new label groups and migrated labels, and make any changes that you need. For example, reconfigure or unpublish any new labels that were created from parent labels.

For parent labels that were migrated to label groups, and for existing sublabels, users won't see or experience any changes in their apps. For example, priority ordering behaves the same and a default sublabel continues to be honored.

Last updated on 04/17/2026

Restrict access to content by using sensitivity labels to apply encryption

> *Microsoft Purview service description*

When you create a sensitivity label, you can restrict access to content that the label will be applied to. For example, with the encryption settings for a sensitivity label, you can protect content so that:

- Only users within your organization can open a confidential document or email.
- Only users in the marketing department can edit and print the promotion announcement document or email, while all other users in your organization can only read it.
- Users can't forward an email or copy information from it that contains news about an internal reorganization.
- The current price list that is sent to business partners can't be opened after a specified date.
- Only the people sent a meeting invite to kick off a confidential project can open the meeting invite and they can't forward it to others.

When a document, email, or meeting invite is encrypted, access to the content is restricted, so that it:

- Can be decrypted only by users authorized by the label's encryption settings.
- Remains encrypted no matter where it resides, inside or outside your organization, even if the file's renamed.
- Is encrypted both at rest (for example, in a OneDrive account) and in transit (for example, email as it traverses the internet).

Finally, as an admin, when you configure a sensitivity label to apply encryption, you can choose either to:

- **Assign permissions now**, so that you determine exactly which users get which permissions to content with that label.
- **Let users assign permissions** when they apply the label to content. This way, you can allow people in your organization some flexibility that they might need to collaborate and get their work done.

The encryption settings are available when you create or edit a sensitivity label in the Microsoft Purview portal.

> **Note**
>
> A sensitivity label in Outlook can apply S/MIME protection rather than encryption and permissions from the Azure Rights Management service. For more information, see Configure a label to apply S/MIME protection in Outlook.

Understand how the encryption works

Unless you're using S/MIME for Outlook, encryption that's applied by sensitivity labels to documents, emails, and meeting invites all use the Azure Rights Management service (Azure RMS) from Microsoft Purview Information Protection. This protection solution uses encryption, identity, and authorization policies. To learn more, see What is Azure Rights Management?.

When you use this encryption solution, the **super user** feature ensures that authorized people and services can always read and inspect the data that has been encrypted for your organization. If necessary, the encryption can then be removed or changed. For more information, see Configure Azure Rights Management super users for discovery services or data recovery.

> **Important**
>
> You can also use sensitivity labels to apply encryption to audio and video streams for Teams meetings, but this uses a different method of encryption and not the Azure Rights Management service that's used for emails, meeting invites, and documents. For more information about the encryption used for Teams meetings, see the Media encryption from the Teams security guide.

Important prerequisites

Before you can use encryption, you might need to do some configuration tasks. When you configure encryption settings, there's no check to validate that these prerequisites are met.

- Activate Azure Rights Management

For sensitivity labels to apply encryption with rights management, the Azure Rights Management service from Microsoft Purview Information Protection must be activated for your tenant. In newer tenants, this is the default setting, but you might need to manually activate the service. For more information, see Activate the Azure Rights Management service.

- Check for network requirements

  You might need to make some changes on your network devices such as firewalls. For details, see Firewalls and network infrastructure.

- Check your Microsoft Entra configuration

  There are some Microsoft Entra configurations that can prevent authorized access to encrypted content. For example, cross-tenant access settings and Conditional Access policies. For more information, see Microsoft Entra configuration for encrypted content.

- Configure Exchange for Azure Rights Management

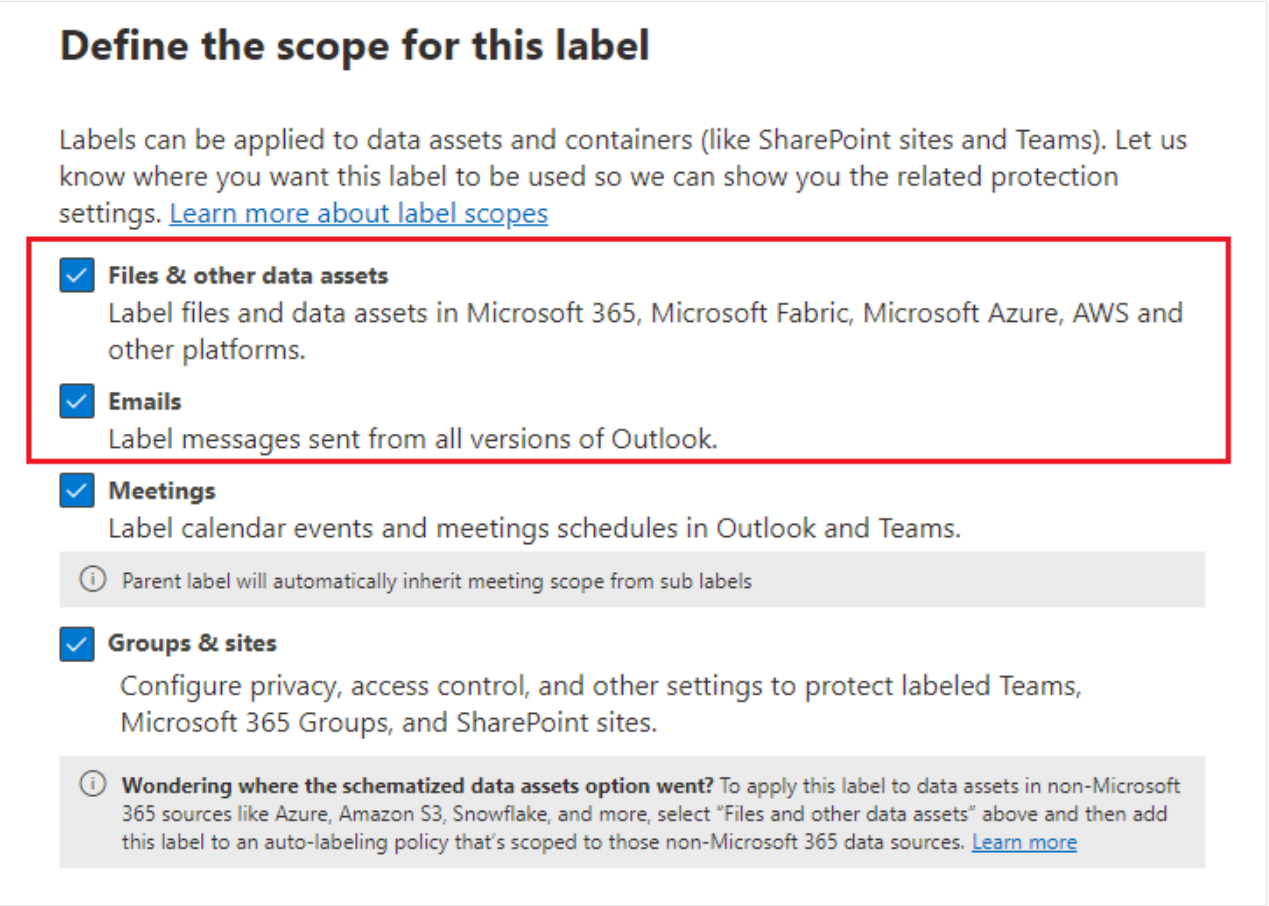
  Exchange doesn't have to be configured for Azure Rights Management before users can apply labels in Outlook to encrypt their emails. However, until Exchange is configured for Azure Rights Management, you don't get the full functionality of encryption with rights management.

  For example, users can't view encrypted emails or encrypted meeting invites on mobile phones or with Outlook on the web, encrypted emails can't be indexed for search, and you can't configure Exchange Online DLP for Rights Management protection.

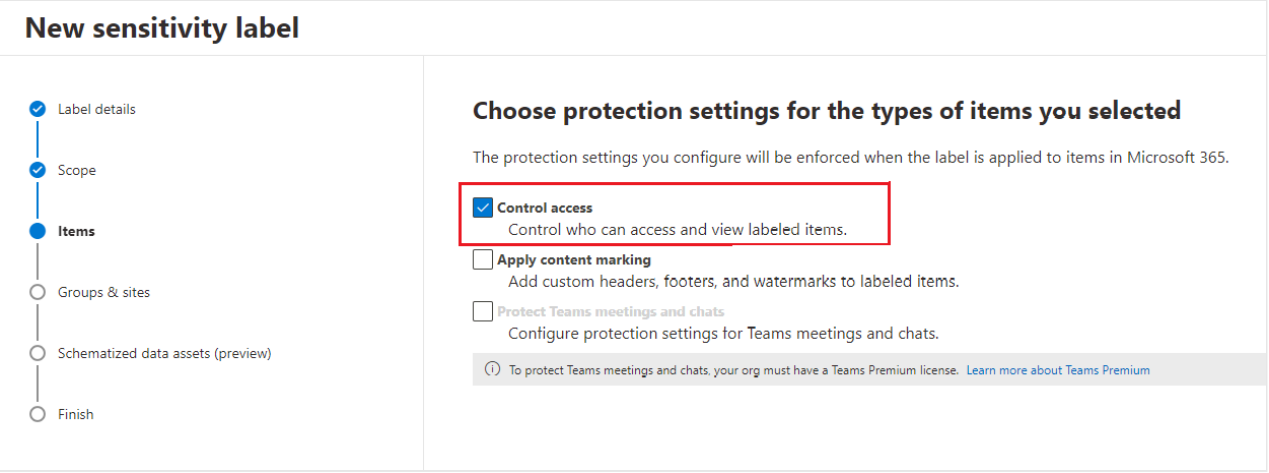
  To ensure that Exchange can support these additional scenarios:
  - For Exchange Online, see the instructions for Exchange Online: IRM Configuration.
  - For Exchange on-premises, you must deploy the RMS connector and configure your Exchange servers.

How to configure a label for encryption

1. Follow the general instructions to create or edit a sensitivity label and make sure the **Files & other data assets** option is selected for the label's scope:



2. Then, on the **Choose protection settings for the types of items you selected** page, make sure you select **Control access**.



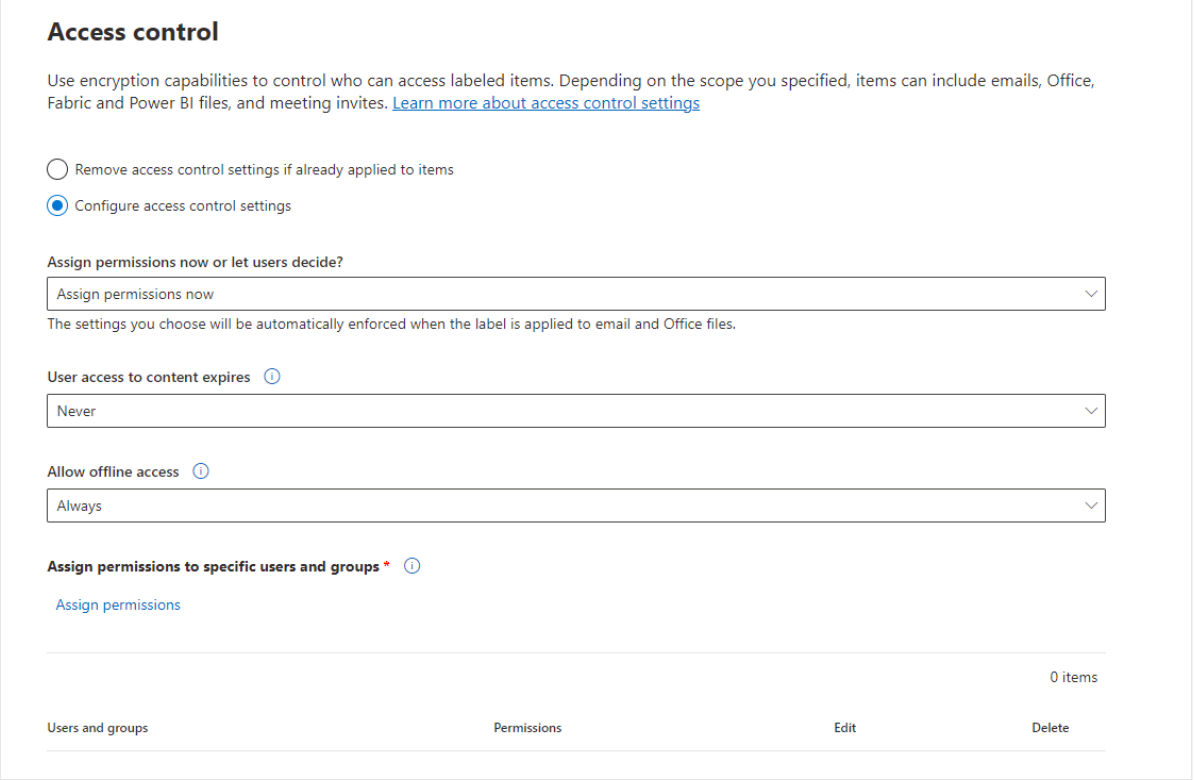
3. On the **Access control** page, select one of the following options:

   - **Remove access control settings if already applied to items**: When you select this option, applying the label removes existing encryption, even if it was applied independently from a sensitivity label.

     It's important to understand that this setting can result in a sensitivity label that users might not be able to apply when they don't have sufficient permissions to

remove the existing encryption. For more information about this scenario, see the What happens to existing encryption when a label's applied section.

- **Configure access control settings**: Turns on encryption with rights management and makes the following settings visible:



Instructions for these settings are in the following Configure encryption settings section.

Editing labels to newly apply encryption or change existing encryption settings

It's a common deployment strategy to initially configure sensitivity labels that don't apply encryption, and later edit some of the existing labels to apply encryption. When the label changes reach your apps, the labels that you edited will apply that encryption for newly labeled items.

For files accessed in SharePoint and OneDrive when you've enabled SharePoint and OneDrive for sensitivity labels the new encryption status of files automatically change when these files are next accessed. For example, they're opened in Office for the web, or downloaded. There's no need to remove the label and reapply it. For example, files that were previously unencrypted become encrypted.

For other items that are already labeled, they retain their previous encryption status unless you remove the label and reapply it. For example, after changing the sensitivity label to now apply encryption, when you open previously labeled and unencrypted documents or emails in Office Desktop or Office Mobile, those items remain unencrypted unless you remove the label and reapply it. Similarly, if you change the sensitivity label settings to not apply encryption (remove the option for access control), those items remain encrypted unless you remove the label and reapply it.

For items that are already labeled with encryption and the assign permissions now option, when you change the encryption settings for users or permissions, the new settings will be applied for existing items when users authenticate with the encryption service. In most cases, there's no need to remove and reapply the label. However, if users have already opened an encrypted document or email, they won't get the new settings until their use license has expired and they must reauthenticate. For more information about this scenario, see the related frequently asked question for how the encryption works.

Whenever you change the encryption options for letting users assign permissions, that change only applies to newly labeled or relabeled items. For example:

- You change the label from assigning permissions now to let users assign permissions, or the other way around
- You change the label from Do Not Forward to Encrypt-Only, or the other way around

What happens to existing encryption when a label's applied

If a sensitivity label is applied to unencrypted content, the outcome for the encryption options you can select is self-explanatory. For example, if you didn't select **Control access**, the content remains unencrypted.

However, the content might be already encrypted. For example, another user might have applied:

- Their own permissions, which include user-defined permissions when prompted by a label, custom permissions by the Microsoft Purview Information Protection client, and the **Restricted Access** document protection from within an Office app.
- A rights management template that encrypts the content independently from a label. This category includes mail flow rules that apply encryption by using rights protection.
- A label that applies encryption with permissions assigned by the administrator.

The following table identifies what happens to existing encryption when a sensitivity label is applied to that content:

Expand table

	Encryption: Not selected	Encryption: Configured	Encryption: Remove
Permissions specified by a user	Original encryption is removed	New label encryption is applied	Original encryption is removed
Rights management template	Original encryption is preserved	New label encryption is applied	Original encryption is removed
Label with administator-defined permissions	Original encryption is removed	New label encryption is applied	Original encryption is removed

In the cases where the new label encryption is applied or the original encryption is removed, this happens only if the user who applies the label has a usage right or role that supports this action:

- The usage right Export or Full Control.
- The role of Rights Management issuer or Rights Management owner, or super user.

If the user doesn't have one of these rights or roles, the label can't be applied and so the original encryption is preserved. The user sees the following message: **You don't have permission to make this change to the sensitivity label. Please contact the content owner.**

For example, the person who applies Do Not Forward to an email message can relabel the thread to replace the encryption or remove it, because they're the Rights Management owner for the email. But except for super users, recipients of this email can't relabel it because they don't have the required usage rights.

Email attachments for encrypted email messages and meeting invites

When an email message or meeting invite is encrypted by any method, any unencrypted Office documents that are attached to the email or invite automatically inherit the same encryption settings. You can't turn off this encryption inheritance, for example, with a setting or label scoping.

Any sensitivity label from the email message or meeting invite is not inherited by the attachment but could be automatically applied when a user in the same tenant opens the document. For more information, see Encryption-based label matching for documents.

Documents that are already encrypted and then added as attachments always preserve their original encryption.

Configure encryption settings

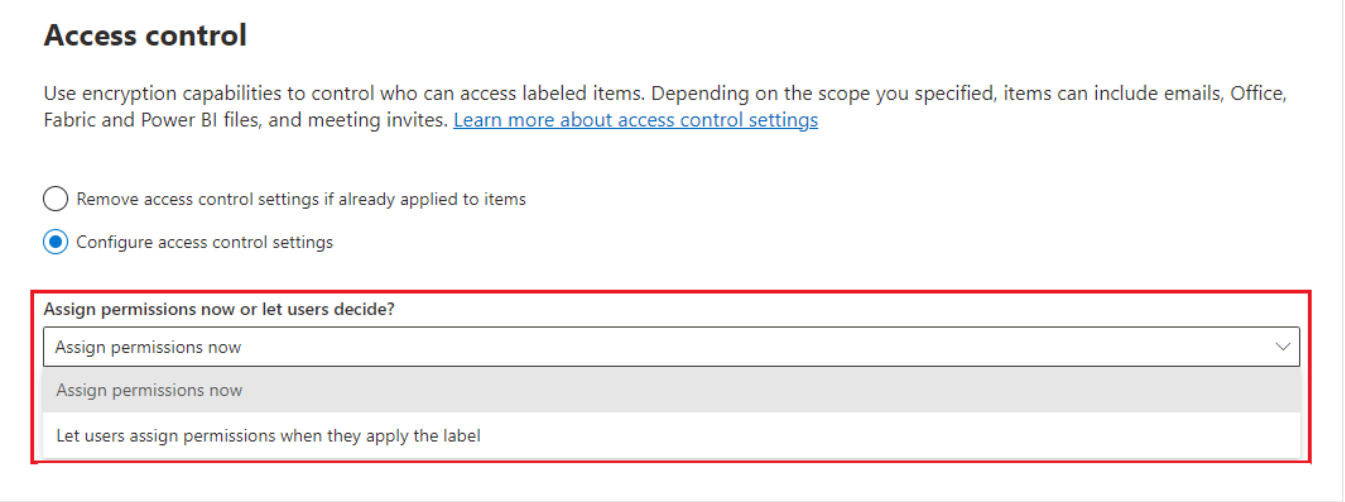
When you select **Configure access control settings** on the **Access control** page to create or edit a sensitivity label, choose one of the following options:

- **Assign permissions now**, so that you can determine exactly which users get which permissions to content that has the label applied. For more information, see the next section Assign permissions now.
- **Let users assign permissions** when your users apply the label to content. With this option, you can allow people in your organization some flexibility that they might need to collaborate and get their work done. For more information, see the Let users assign permissions section on this page.

For example, if you have a sensitivity label named **Highly Confidential** that will be applied to your most sensitive content, you might want to decide now who gets what type of permissions to that content.

Alternatively, if you have a sensitivity label named **Business Contracts**, and your organization's workflow requires that your people collaborate on this content with different people on an unplanned basis, you might want to allow your users to decide who gets permissions when they assign the label. This flexibility both helps your users' productivity and reduces the requests for your admins to update or create new sensitivity labels to address specific scenarios.

Choosing whether to assign permissions now or let users assign permissions:



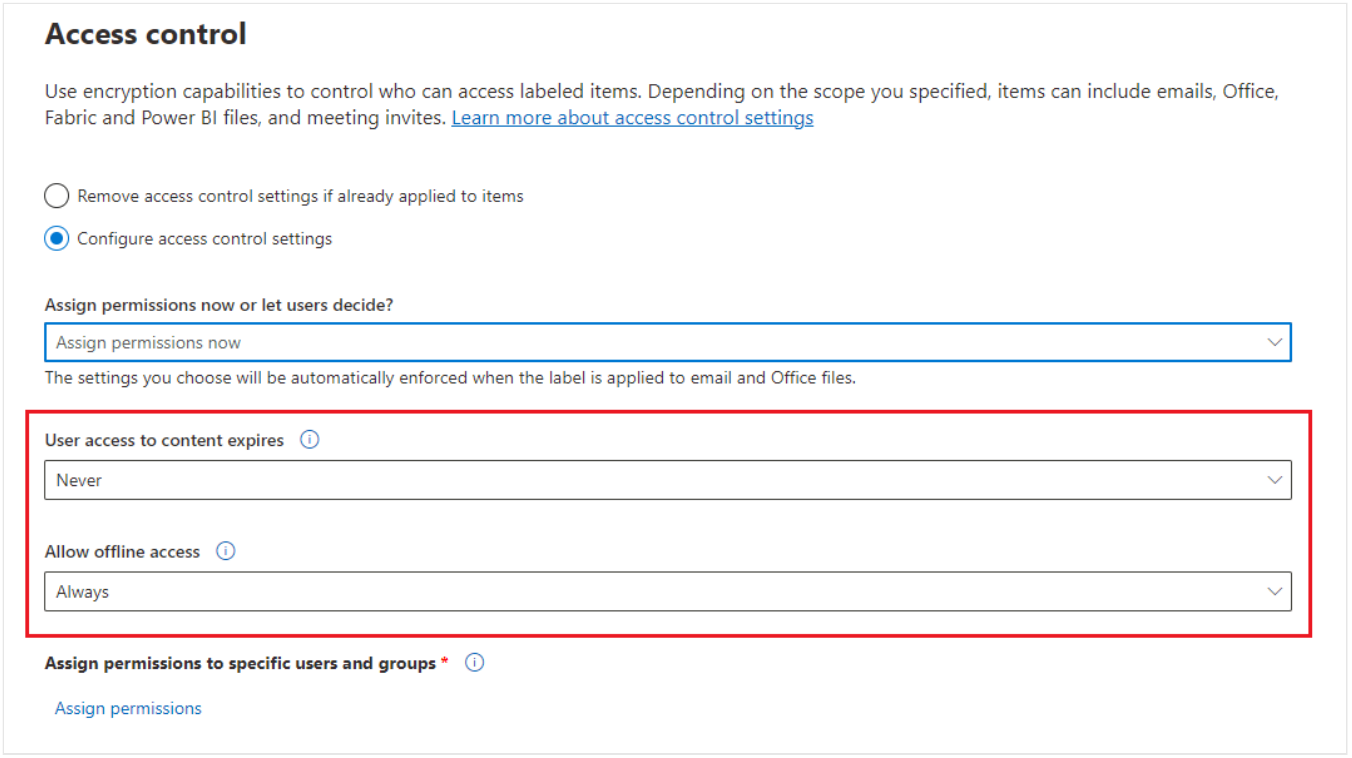
Assign permissions now

Use the following options to control who can access email, meeting invites (if enabled), or documents to which this label is applied. You can:

- **Allow access to labeled content to expire**, either on a specific date or after a specific number of days after the label is applied. After this time, users won't be able to open the labeled item. If you specify a date, it's effective midnight on that date in your current time zone. Some email clients might not enforce expiration and show emails past their expiration date, due to their caching mechanisms.

- **Allow offline access** never, always, or for a specific number of days after the label is applied. Use this setting to balance any security requirements you have with the ability for users to open encrypted content when they don't have an internet connection. If you restrict offline access to never or a number of days, when that threshold is reached, users must be reauthenticated and their access is logged. For more information about how this process works, see the following section about the Rights Management use license.

Access control settings for encrypted content:



Recommendations for the expiry and offline access settings:

Expand table

Setting	Recommended setting
User access to content expires	**Never** unless the content has a specific time-bound requirement.
Allow offline access	Depends on the sensitivity of the content: - **Only for a number of days** = **7** for sensitive business data that could cause damage to the business if shared with unauthorized people. This recommendation offers a balanced compromise between flexibility and security. Examples include contracts, security reports, forecast summaries, and sales account data. - **Never** for when the label is configured for dynamic watermarks and any very sensitive business data that would cause damage to the business if it was shared with unauthorized people. This recommendation prioritizes security over flexibility, and ensures that if you remove one or more users' access to the document, they won't be able to open it. Examples include employee and customer information, passwords, source code, and pre-announced financial reports. - **Always** for less sensitive content where it doesn't matter if users can continue to open encrypted content for up to 30 days (or the configured use license validity period for the tenant) after their access is removed and they have previously opened the encrypted content.

Only labels that are configured to assign permissions now support different values for offline access. Labels that let users assign the permissions automatically use the tenant's Rights Management use license validity period. For example, labels that are configured for Do Not Forward, Encrypt-Only, and prompt users to specify their own permissions. The default value for this setting is 30 days.

Rights Management use license for offline access

> **Note**
>
> Although you can configure the encryption setting to allow offline access, some apps might not support offline access for encrypted content. For example, labeled and encrypted files in Power BI Desktop won't open if you're offline.

When a user opens an item that's been protected by encryption from the Azure Rights Management service, an Azure Rights Management use license for that content is granted to the user. This use license is a certificate that contains the user's usage rights for the document

or email, and the encryption key that was used to encrypt the content. The use license also contains an expiration date if this has been set, and how long the use license is valid.

If no expiration date has been set, the default use license validity period for a tenant is 30 days. For the duration of the use license, the user isn't reauthenticated or reauthorized for the content. This process lets the user continue to open the protected document or email without an internet connection. When the use license validity period expires, the next time the user accesses the protected document or email, the user must be reauthenticated and reauthorized.

In addition to reauthentication, the encryption settings and user group membership is reevaluated. This means that users could experience different access results for the same item if there are changes in the encryption settings or group membership from when they last accessed the content.

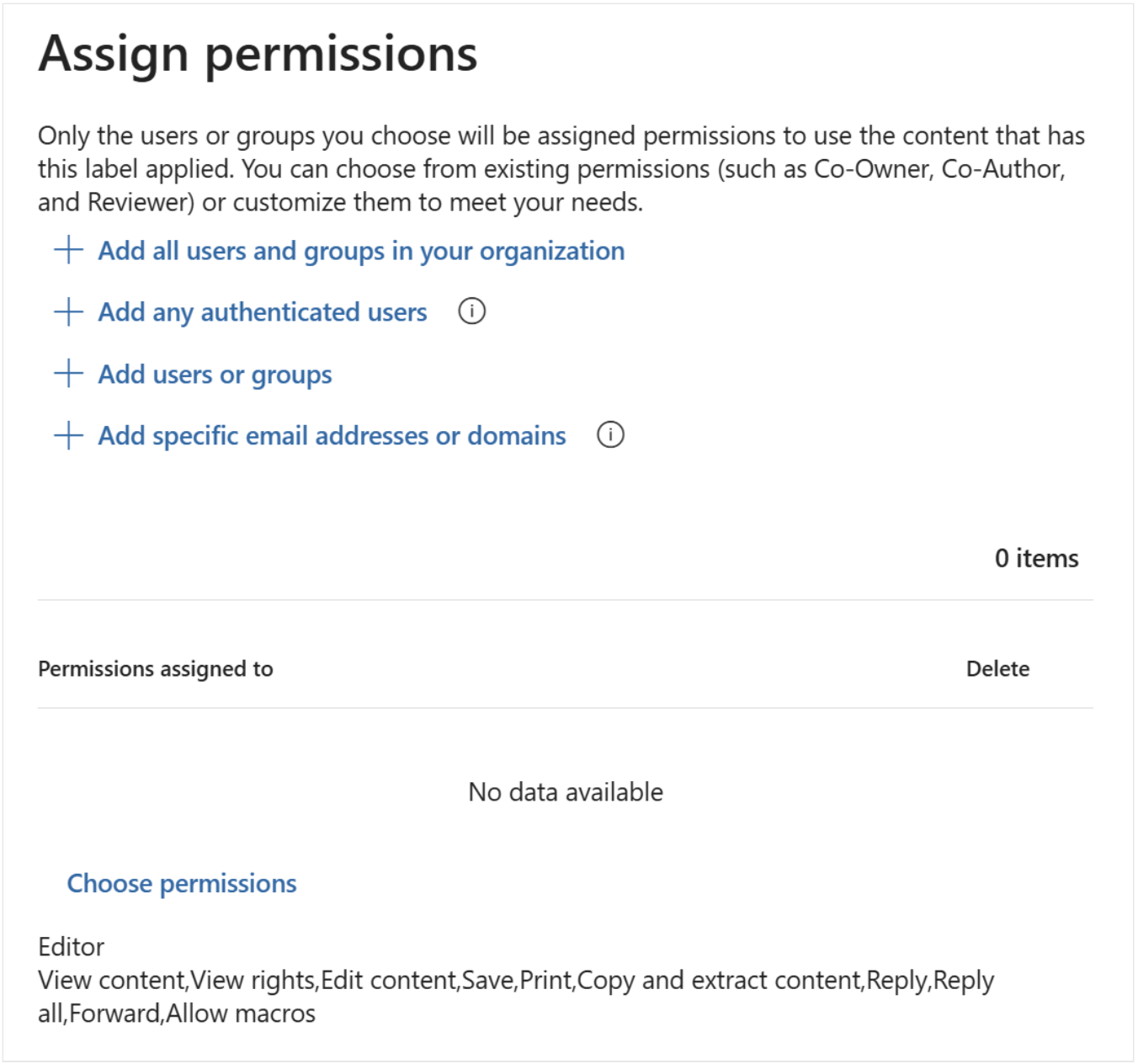
To learn how to change the default 30-day setting, see Rights Management use license.

Assign permissions to specific users or groups

You can grant permissions to specific people so that only they can interact with the labeled content:

1. First, add users or groups that will be assigned permissions to the labeled content.
2. Then, choose which permissions those users should have for the labeled content.

Assigning permissions:



Add users or groups

When you assign permissions, you can choose:

- Everyone in your organization (all tenant members). This setting excludes guest accounts.
- Any authenticated users. Make sure you understand the requirements and limitations of this setting before selecting it.
- Any specific user or email-enabled security group, distribution group, or Microsoft 365 group in Microsoft Entra ID. The Microsoft 365 group can have static or dynamic membership. You can't use a dynamic distribution group from Exchange because this group type isn't synchronized to Microsoft Entra ID. You also can't use a security group that isn't email-enabled.

Although you can specify groups that contain mail contacts as a convenient method to grant access to multiple people outside your organization, there's currently a known issue with this configuration. For more information, see Mail contacts in groups have intermittent access to encrypted content.

- Any email address or domain. Use this option to specify all users in another organization who uses Microsoft Entra ID, by entering any domain name from that organization. You can also use this option for social providers, by entering their domain name such as **gmail.com**, **hotmail.com**, or **outlook.com**.

  > **Note**
  >
  > If you specify a domain from an organization that uses Microsoft Entra ID, you can't restrict access to that specific domain. Instead, all verified domains in Microsoft Entra ID are automatically included for the tenant that owns the domain name you specify.

When you choose all users and groups in your organization or browse the directory, the users or groups must have an email address.

As a best practice, use groups rather than users. This strategy keeps your configuration simpler.

Requirements and limitations for "Add any authenticated users"

This setting doesn't restrict who can access the content that the label encrypts, while still encrypting the content and providing you with options to restrict how the content can be used (permissions), and accessed (expiry and offline access). However, the application opening the encrypted content must be able to support the authentication being used. For this reason, federated social providers such as Google, and onetime passcode authentication work only for email and meeting invites, and only when you use Exchange Online. Microsoft accounts can be used with Office 365 apps and the Microsoft Purview Information Protection viewer.

> **Note**
>
> Consider using this setting with SharePoint and OneDrive integration with Microsoft Entra B2B when sensitivity labels are enabled for Office files in SharePoint and OneDrive.

Some typical scenarios for any authenticated users setting:

- You don't mind who views the content, but you want to restrict how it's used. For example, you don't want the content to be edited, copied, or printed.
- You don't need to restrict who accesses the content, but you want to be able to confirm who opens it.
- You have a requirement that the content must be encrypted at rest and in transit, but it doesn't require access controls.

Choose permissions

When you choose which permissions to allow for those users or groups, you can select either:

- A predefined permissions level with a preset group of rights, such as Editor or Restricted Editor.
- Custom permissions, where you choose one or more usage rights.

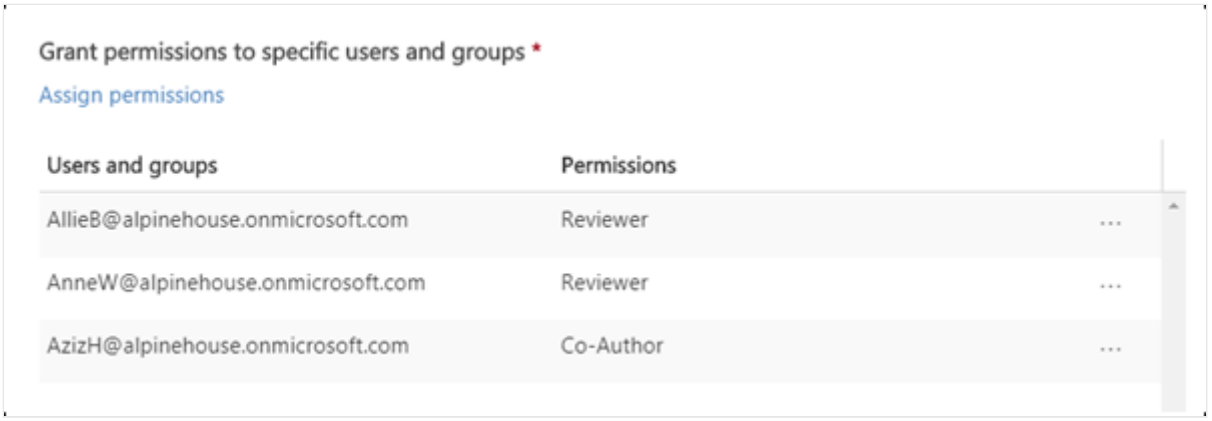
For more information to help you select the appropriate permissions, see Usage rights and descriptions.



Note that the same label can grant different permissions to different users. For example, a single label can assign some users as Restricted Editor and a different user as Owner, as shown in the following screenshot.

To do this, add users or groups, assign them permissions, and save those settings. Then repeat these steps, adding users and assigning them permissions, saving the settings each time. You

can repeat this configuration as often as necessary, to define different permissions for different users.



Rights Management issuer (user applying the sensitivity label) always has Full Control

By default, encryption for a sensitivity label uses the Azure Rights Management service from Microsoft Purview Information Protection. When a user applies a sensitivity label to protect a document or email by using encryption, that user becomes the Rights Management issuer for that content.

The Rights Management issuer is always granted Full Control permissions for the document or email, and in addition:

- If the encryption settings include an expiration date, the Rights Management issuer can still open and edit the document or email after that date.
- The Rights Management issuer can always access the document or email offline.
- The Rights Management issuer can still open a document after it's revoked.

For more information, see Rights Management issuer and Rights Management owner.

Dynamic watermarks

Identify the minimum versions of Office apps required by using the capabilities tables and the rows for **Dynamic watermarks**.

When a user accesses a file with this sensitivity label applied, by default, their Universal Principal Name (UPN) is dynamically inserted as a watermark on each page of the file. Typically, a user's UPN is the same as their email address. If you select **Customize text (optional)**, you can specify a custom string that also supports including the date and time.

This watermark is highly visible when viewing the file on a device, and persists when printed, although not when exported. This watermarking is more secure than the standard content markings for a label, because the user can't easily manually remove or change it.

The watermark is removed when a user relabels the document, choosing a sensitivity label that applies a different dynamic watermark, or no dynamic watermark. However, as with all encrypted content, a user must have the usage right of Export or Full Control to remove the existing encryption.

This protective watermarking is supported only for sensitivity labels that apply encryption and we also recommend that the encryption option of **Allow offline access** is set to **Never**. As with other encryption settings, dynamic watermarks are supported across tenants. Use it for your most sensitive documents as a visual deterrent against screen captures by the person who opened the document. However, be aware of the following considerations.

Limitations in Office where dynamic watermarks are not yet supported, so the document opens without apply the watermark:

- File preview for attachments in Outlook
- PowerPoint Live in Microsoft Teams

Unless opening the document falls into one of the exceptions previously listed, one of the following must be true to open a labeled document that applies dynamic watermarks:

- The user is the Rights Management owner, which most of the time, means they applied the label themselves. For more information, see Rights Management issuer and Rights Management owner.
- The user opens the document with an app that understands dynamic watermarks. Use the capabilities tables and the row **Dynamic watermarks** to confirm supported versions.
- The user opens the document in Office on the web.

If none of these conditions are met when a user opens the document in an Office app, and they're not one of the listed exceptions, the document won't open. Make sure you understand this restriction before you use this label configuration.

> **⊗ Caution**
>
> Except for the listed limitations, Microsoft Office apps will either deny access if they don't support dynamic watermarks, or enforce dynamic watermarks if they do.

Considerations for previously labeled documents:

- Like all changed encryption settings, a newly configured dynamic watermark doesn't immediately get applied when you already have a valid Rights Management use license for a document. After the use license expires, you must reauthenticate to the Azure Rights Management service to open the document and then the dynamic watermark setting gets applied.
- There's no automatic detection and resolution if the sensitivity label previously applied a standard watermark as a content marking.

Double Key Encryption

> ⓘ **Note**
>
> For built-in labeling, identify the minimum versions required by using the capabilities tables and the row **Double Key Encryption (DKE)**.

Select the **Double Key Encryption** label option only after you've configured the Double Key Encryption service and you need to use this double key encryption for files and emails that will have this label applied. After the label is configured and saved, you won't be able to edit it.

For more information, prerequisites, and configuration instructions, see Double Key Encryption (DKE).

Let users assign permissions

> ⓘ **Important**
>
> Not all labeling clients support all the options that let users assign their own permissions. Use this section to learn more.

Often referred to as "user-defined permissions", you can use the following options to let users assign permissions when they manually apply a sensitivity label to content:

- In Outlook, a user can select restrictions equivalent to the Do Not Forward option or Encrypt-only for their chosen recipients.

The Do Not Forward option is supported by all email clients that support sensitivity labels. However, applying the **Encrypt-Only** option with a sensitivity label is a more recent release. For email clients that don't support this capability, the label won't be visible.

To check the minimum versions of Outlook apps that use built-in labeling to support applying the Encrypt-Only option with a sensitivity label, use the capabilities table for Outlook and the row **Let users assign permissions: - Encrypt-Only**.

- In Word, PowerPoint, and Excel, unless the label is configured to extend permissions for a SharePoint document library, a user is prompted to select their own permissions for specific users, groups, or organizations.

  For Office apps that don't support this capability, the label either won't be visible for users, or the label is visible for consistency but it can't be applied with an explanation message to users.

  To check which Office apps support this option, use the capabilities table for Word, Excel, and PowerPoint and the rows for **Let users assign permissions**.

> **Note**
>
> You won't be able to use these configurations if the label scope excludes email (for Do Not Forward and Encrypt-Only) or excludes files (for prompting users in Word, PowerPoint, and Excel). For more information, see Scope labels to just files or emails.

When the options are supported, use the following table to identify when users see the sensitivity label:

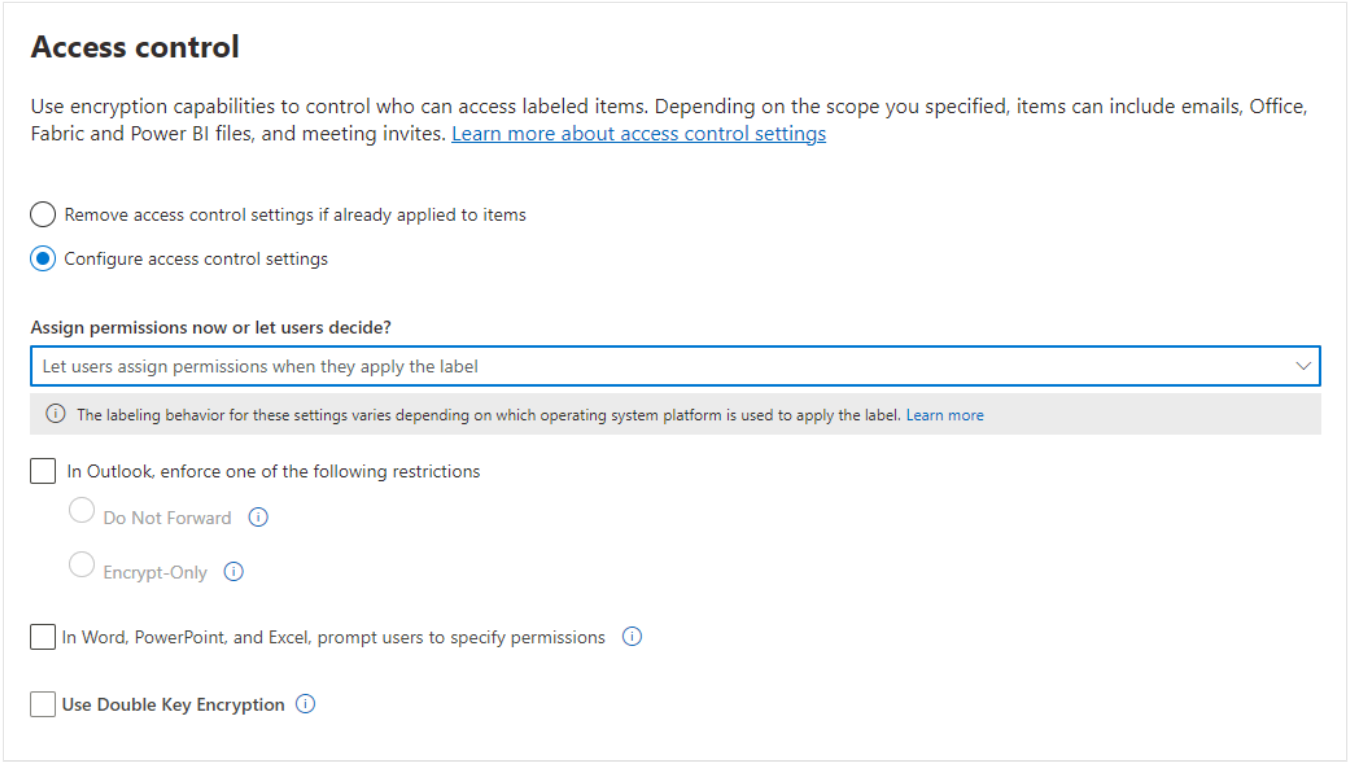
Expand table

Setting	Label visible in Outlook	Label visible in Word, Excel, PowerPoint
In Outlook, enforce restrictions with the Do Not Forward or Encrypt-Only option	Yes	No
In Word, PowerPoint, and Excel, prompt users to specify permissions	No	Yes

When both settings are selected, the label is therefore visible in both Outlook and in Word, Excel, and PowerPoint.

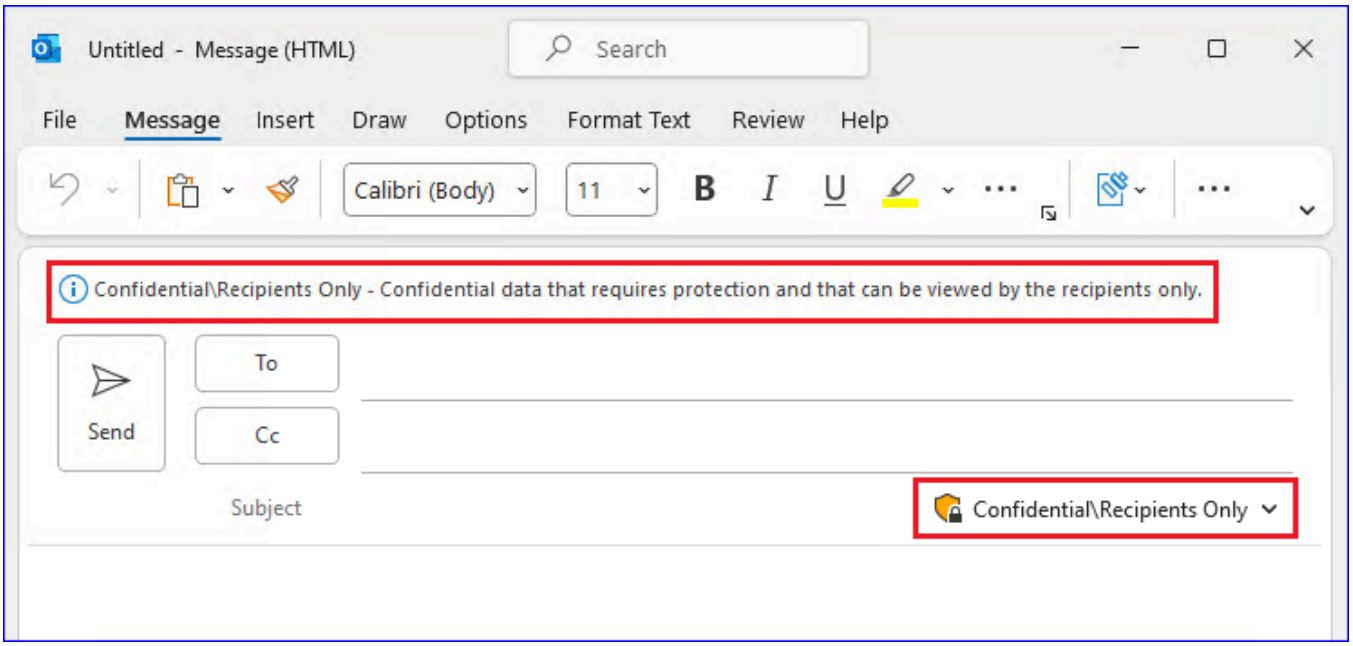
A sensitivity label that lets users assign permissions can be recommended to users, but can only be automatically applied for the Do Not Forward and Encrypt-Only options.

Configuring the user-assigned permissions:



Outlook restrictions

In Outlook, when a user applies a sensitivity label that lets them assign permissions to a message, you can choose the **Do Not Forward option** or **Encrypt-Only**. The user will see the label name and description at the top of the message, which indicates the content's being protected. Unlike Word, PowerPoint, and Excel (see the next section), users aren't prompted to select specific permissions. For this configuration, the administrator controls the permissions, but not who has access.



When either of these options are applied to an email, the email is encrypted and recipients must be authenticated. Then, the recipients automatically have restricted usage rights:

- **Do Not Forward**: Recipients can't forward the email, print it, or copy from it. For example, in the Outlook client, the Forward button isn't available, the Save As and Print menu options aren't available, and you can't add or change recipients in the To, Cc, or Bcc boxes.

  For more information about how this option works, see Do Not Forward option for emails.

- **Encrypt-Only**: Recipients have all usage rights except Save As, Export and Full Control. This combination of usage rights means that the recipients have no restrictions except that they can't remove the protection. For example, a recipient can copy from the email, print it, and forward it.

  For more information about how this option works, see Encrypt-only option for emails.

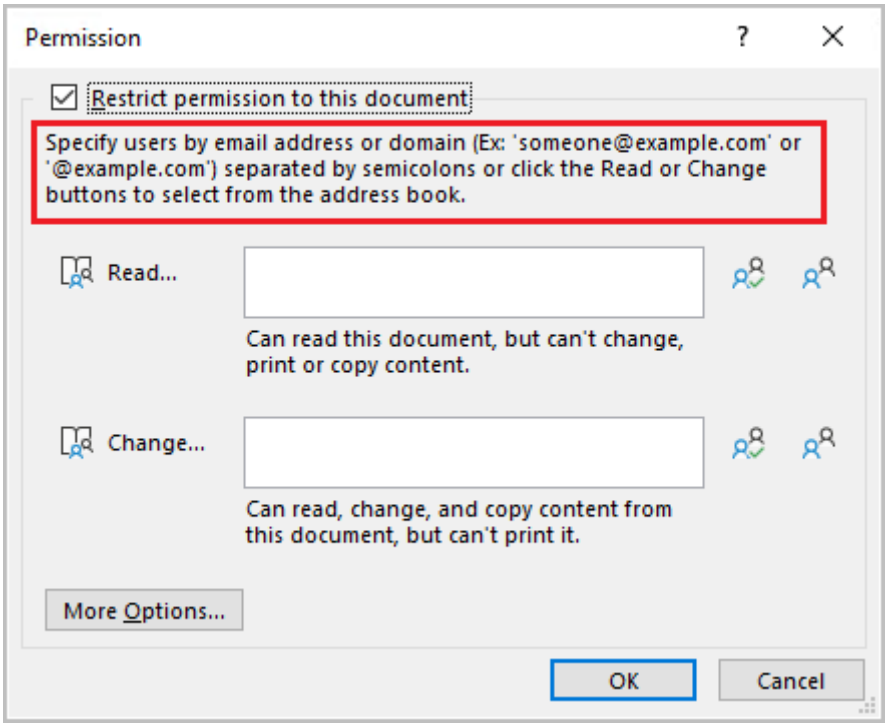
Unencrypted Office documents that are attached to the email or meeting invite automatically inherit the same restrictions. For Do Not Forward, the usage rights applied to these documents are Edit Content, Edit; Save; View, Open, Read; and Allow Macros. If the user wants different usage rights for an attachment, or the attachment isn't an Office document that supports this inherited protection, the user needs to encrypt the file before attaching it to the email or meeting invite.

Word, PowerPoint, and Excel permissions

In Word, PowerPoint, and Excel, when a user applies a sensitivity label that lets them assign permissions to a document, the user is prompted to specify their choice of users and permissions for the encryption. For this configuration, the user and not the administrator controls both who can access the document, and what permissions they have.

Support for organization-wide custom permissions

For labeling in Windows and using Office for the web, users can additionally specify a domain name when they're prompted to specify their choice of users and permissions. When a domain name is entered, the permissions will apply to all users in an organization that owns the domain and it is in Microsoft Entra ID. To identify the minimum versions that support this setting, use the capabilities table and the row **Let users assign permissions:- Prompt users for custom permissions (users, groups, and organizations)**.

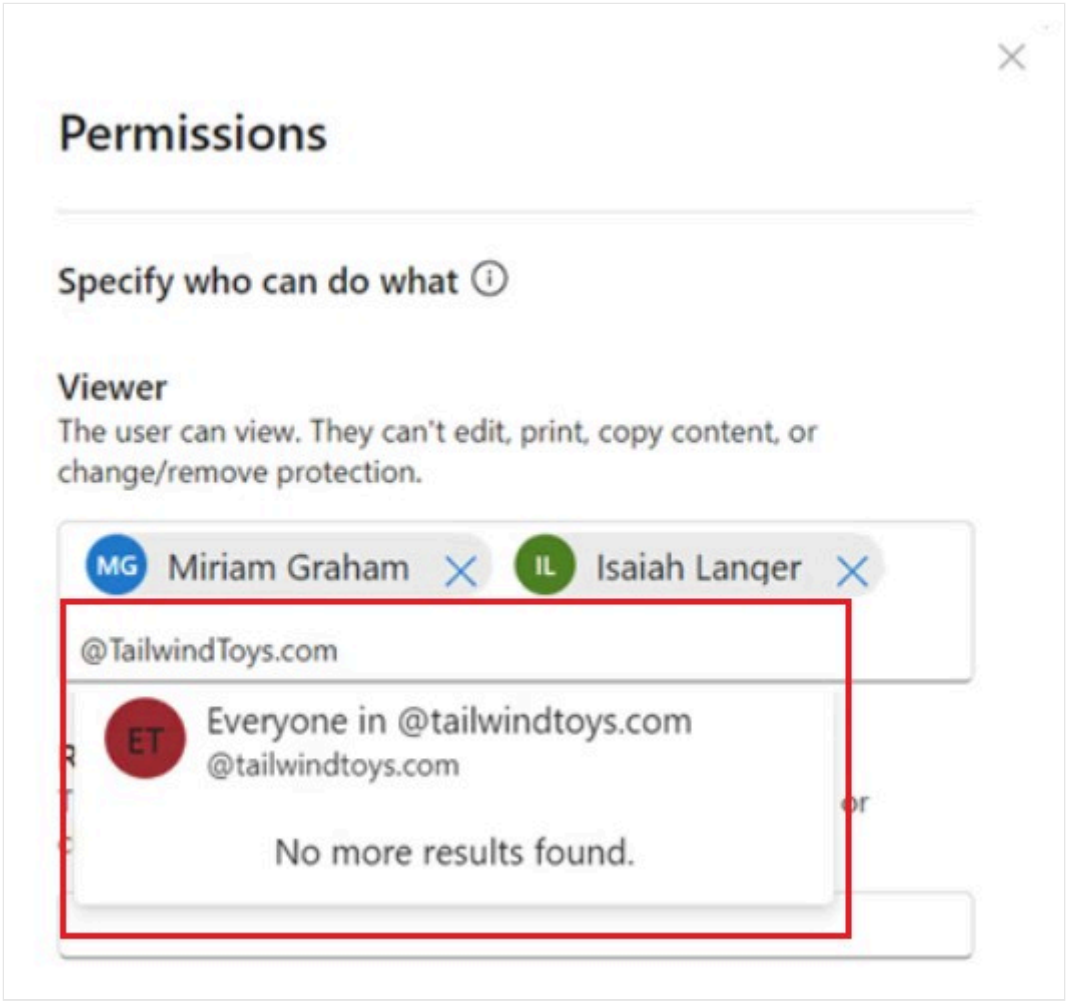


> **Note**
>
> The displayed dialog box is updated in the latest versions of Microsoft 365 for Windows apps, but the support for organization-wide custom permissions remains when a domain name is entered. Information about this support is included in the popup information box.

For example, a user types **@contoso.com** (or **contoso.com**) and grants read access. Because Contoso Corporation owns the contoso.com domain, all users in that domain and all other domains that the organization owns in Microsoft Entra ID will be granted read access.

> **Note**
>
> When you specify these values, don't surround them with quotation marks.

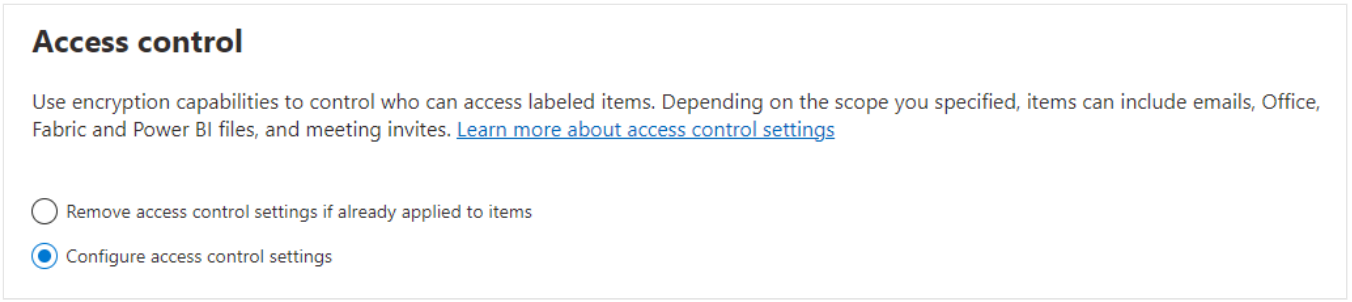
When you use Office for the web, typing the @ character automatically displays a selectable entry for everybody in the organization. For example:



It's important to let users know that access isn't restricted to just the users in the domain specified. For example, **@sales.contoso.com** wouldn't restrict access to users in just the sales subdomain, but also grant access to users in the marketing.contoso.com domain, and even users with a disjoint namespace in the same Microsoft Entra tenant. For user documentation, see Custom permissions for files.

Example configurations for the encryption settings

For each example that follows, do the configuration from the **Access control** page when the **Configure access control settings** option is selected:



Example 1: Label that applies Do Not Forward to send an encrypted email to a Gmail account

This label displays only in Outlook and Outlook on the web, and you must use Exchange Online. Instruct users to select this label when they need to send an encrypted email to people using a Gmail account (or any other email account outside your organization).

Your users type the Gmail email address in the **To** box. Then, they select the label and the Do Not Forward option is automatically added to the email. The result is that recipients can't forward the email, or print it, copy from it, or save the email outside their mailbox by using the **Save As** option.

1. On the **Access control** page: For **Assign permissions now or let users decide?** select **Let users assign permissions when they apply the label**.
2. Select the checkbox: **In Outlook, enforce restrictions equivalent to the Do Not Forward option**.
3. If selected, clear the checkbox: **In Word, PowerPoint, and Excel, prompt users to specify permissions**.
4. Select **Next** and complete the configuration.

Example 2: Label that restricts read-only permission to all users in another organization

This label is suitable for sharing very sensitive documents as read-only, and the documents always require an internet connection to view them.

This label isn't suitable for emails.

1. On the **Access control* page: For **Assign permissions now or let users decide?** select **Assign permissions now**.
2. For **Allow offline access**, select **Never**.

3. Select **Assign permissions**.
4. On the **Assign permissions** pane, select **Add specific email addresses or domains**.
5. In the text box, enter the name of a domain from the other organization, for example, **fabrikam.com**. Then select **Add**.
6. Select **Choose permissions**.
7. On the **Choose permissions** pane, select the dropdown box, select **Viewer**, and then select **Save**.
8. Back on the **Assign Permissions** pane, select **Save**.
9. On the **Access control** page, select **Next** and complete the configuration.

Example 3: Add external users to an existing label that encrypts content

The new users that you add will be able open documents and emails that have already been protected with this label. The permissions that you grant these users can be different from the permissions that the existing users have.

1. On the **Access control** page: For **Assign permissions now or let users decide?** make sure **Assign permissions now** is selected.
2. Select **Assign permissions**.
3. On the **Assign permissions** pane, select **Add specific email addresses or domains**.
4. In the text box, enter the email address of the first user (or group) to add, and then select **Add**.
5. Select **Choose permissions**.
6. On the **Choose permissions** pane, select the permissions for this user (or group), and then select **Save**.
7. Back on the **Assign Permissions** pane, repeat steps 3 through 6 for each user (or group) that you want to add to this label. Then click **Save**.
8. On the **Access control** page, select **Next** and complete the configuration.

Example 4: Label that encrypts content but doesn't restrict who can access it

This configuration has the advantage that you don't need to specify users, groups, or domains to encrypt an email or document. The content will still be encrypted and you can still specify usage rights, an expiry date, and offline access.

Use this configuration only when you don't need to restrict who can open the protected document or email. See more information about this setting.

1. On the **Access control** page: For **Assign permissions now or let users decide?** make sure **Assign permissions now** is selected.
2. Configure settings for **User access to content expires** and **Allow offline access** as required.
3. Select **Assign permissions**.
4. On the **Assign permissions** pane, select **Add any authenticated users**.

   For **Users and groups**, you see **Authenticated users** automatically added. You can't change this value, only delete it, which cancels the **Add any authenticated users** selection.
5. Select **Choose permissions**.
6. On the **Choose permissions** pane, select the dropdown box, select the permissions you want, and then select **Save**.
7. Back on the **Assign Permissions** pane, select **Save**.
8. On the **Access control** page, select **Next** and complete the configuration.

Considerations for encrypted content

Encrypting your most sensitive documents and emails helps to ensure that only authorized people can access this data. However, there are some considerations to take into account:

- If your organization hasn't enabled sensitivity labels for Office files in SharePoint and OneDrive:
  - Search, eDiscovery, and Delve won't work for encrypted files.

    - DLP policies work for the metadata of these encrypted files (including retention label information) but not the content of these files (such as credit card numbers within files).
    - Users can't open encrypted files using Office for the web. When sensitivity labels for Office files in SharePoint and OneDrive are enabled, users can use Office for the web to open encrypted files, with some limitations that include encryption that has been applied with an on-premises key (known as "hold your own key", or HYOK), double key encryption, and encryption that has been applied independently from a sensitivity label.

- If you share encrypted documents with people outside your organization, you might need to create guest accounts and modify Conditional Access policies. For more information, see Sharing encrypted documents with external users.

- When authorized users open encrypted documents in their Office apps, they see the label name and description in a yellow message bar at the top of their app. When the encryption permissions extend to people outside your organization, carefully review the label names and descriptions that will be visible in this message bar when the document is opened.

- For multiple users to edit an encrypted file at the same time, they must all be using Office for the web or you've enabled co-authoring for files encrypted with sensitivity labels and all users have Office apps that support this feature. If this isn't the case, and the file is already open:
    - In Office apps (Windows, Mac, Android, and iOS), users see a **File In Use** message with the name of the person who has checked out the file. They can then view a read-only copy or save and edit a copy of the file, and receive notification when the file is available.
    - In Office for the web, users see an error message that they can't edit the document with other people. They can then select **Open in Reading View**.

- The AutoSave functionality in Office apps is disabled for encrypted files if you haven't enabled co-authoring for files encrypted with sensitivity labels. Users see a message that the file has restricted permissions that must be removed before AutoSave can be turned on.

- Office for Windows supports labels that apply encryption when users aren't connected to the internet. But for the other platforms (macOS, iOS, Android), users must be online to apply these labels in Office apps. The Microsoft Purview Information Protection client

must also be online to apply these labels in File Explorer and PowerShell. Users don't have to be online to open encrypted content. For more information about offline access, see the Rights Management use license for offline access section.

- Encrypted files might take longer to open in Office apps (Windows, Mac, Android, and iOS).
- If a label that applies encryption is added by using an Office app when the document is checked out in SharePoint, and the user then discards the checkout, the document remains labeled and encrypted.
- Unless you've enabled co-authoring for files encrypted with sensitivity labels, the following actions for encrypted files aren't supported from Office apps (Windows, Mac, Android, and iOS), and users see an error message that something went wrong. However, SharePoint functionality can be used as an alternative:
  - View, restore, and save copies of previous versions. As an alternative, users can do these actions using Office for the web when you enable and configure versioning for a list or library.
  - Change the name or location of files. As an alternative, users can rename a file, folder, or link in a document library in SharePoint.

For the best collaboration experience for files that are encrypted by a sensitivity label, we recommend you use sensitivity labels for Office files in SharePoint and OneDrive and Office for the web.

Next steps

Need to share your labeled and encrypted documents with people outside your organization? See Sharing encrypted documents with external users.

To use sensitivity labels to encrypt video and audio streams for Teams meetings, see Use sensitivity labels to protect calendar items, Teams meetings and chat.

Last updated on 04/15/2026

Automatically apply a sensitivity label to Microsoft 365 data

> *Microsoft Purview service description*

> **Note**
>
> For information about automatically applying a sensitivity label to data stored outside Microsoft 365 and in the data map, see Labeling in Microsoft Purview Data Map.

> **Tip**
>
> **Start here — pick your task**
>
> - **I'm setting up my first auto-labeling policy.** Review the Prerequisites for auto-labeling policies first, then follow Creating an auto-labeling policy.
> - **I want to label Exchange email based on the subject, sender, or recipient.** See Example: Apply a label to Exchange email based on the subject.
> - **I'm deciding between labeling in Office apps and an auto-labeling policy.** See Compare auto-labeling for Office apps with auto-labeling policies.
> - **My simulation results look wrong or incomplete.** See Auto-labeling classification and simulation concepts to distinguish available classification, policy matches, and actual labeling. For simulation requirements and workflow, see Learn about simulation mode.
> - **I turned on a policy after simulation and want to review coverage.** See Review coverage by simulation context.
> - **I want SharePoint default library labels to apply to existing files.** See Apply default sensitivity labels from SharePoint document libraries to existing files.
> - **A file wasn't labeled, or has an unexpected label.** Start with Why was — or wasn't — a specific file labeled? to choose the checks for your file. For a reported failure, see Resolve auto-labeling failures in SharePoint and OneDrive files.
> - **My active policy has stopped matching Exchange email.** See Policy rule fails to load.
> - **I'm hitting a limit** (500,000 files/day, 100 policies, 1,000 explicitly selected SharePoint sites, or 20,000,000 files in simulation). See How to configure auto-labeling policies

for SharePoint, OneDrive, and Exchange, and Use PowerShell for auto-labeling policies to target up to 50,000 SharePoint sites by using an adaptive scope.

- **Auto-labeling isn't available in my region.** The **Auto-labeling** page isn't visible in unsupported regions. See the regional availability prerequisite.
- **I want to set a default label for Teams meetings, sites, groups, or Microsoft 365 Copilot.** This article covers files and emails. For container-level default labels (including Teams instant meetings and Meet Now), see Use sensitivity labels with Microsoft Teams, Microsoft 365 groups, and SharePoint sites. For Copilot default labels, see Considerations for Microsoft 365 Copilot.

When you create a sensitivity label, you can automatically assign that label to Microsoft 365 items such as files and emails when the data matches conditions that you specify.

This ability to apply sensitivity labels to content automatically is important because:

- You don't need to train your users when to use each of your classifications.
- You don't need to rely on users to classify all content correctly.
- Users no longer need to know about your policies—they can instead focus on their work.

There are two different methods for automatically applying a sensitivity label to content in Microsoft 365:

- **Client-side labeling when users edit documents or compose (also reply or forward) emails**: Use a label that's configured for auto-labeling for files and emails (includes Word, Excel, PowerPoint, and Outlook).

  This method supports recommending a label to users, as well as automatically applying a label. But in both cases, the user decides whether to accept or reject the label, to help ensure the correct labeling of content. This client-side labeling has minimal delay for documents because the label can be applied even before the document is saved. However, not all client apps support auto-labeling.

  For configuration instructions, see How to configure auto-labeling for Office apps on this page.

- **Service-side labeling when content is already saved (in SharePoint or OneDrive) or emailed (processed by Exchange Online)**: Use an auto-labeling policy.

You might also hear this method referred to as auto-labeling for data at rest (documents in SharePoint and OneDrive) and data in transit (email that is sent or received by Exchange). For Exchange, it doesn't include emails at rest (mailboxes).

Auto-labeling policies run in the service, so they don't depend on which app version a user has installed. They're available across your organization from the moment you turn a policy on, which makes them suitable for labeling at scale. Because the user doesn't interact with the labeling process, auto-labeling policies don't support *recommended* labeling — instead, run the policy in simulation before you turn it on to help ensure correct labeling.

For configuration instructions, see How to configure auto-labeling policies for SharePoint, OneDrive, and Exchange on this page.

Auto-labeling for SharePoint and OneDrive

- PDF documents and Office files for Word (.docx), PowerPoint (.pptx), and Excel (.xlsx) are supported.
  - These files can be auto-labeled at rest before or after the auto-labeling policies are created. Files can't be auto-labeled if they're part of an open session (the file is open).
  - Currently, attachments to list items aren't supported and won't be auto-labeled.

  > **Important**
  >
  > PDF labeling must be enabled for your tenant before an auto-labeling policy can process PDFs in SharePoint and OneDrive. Enable it by selecting the **Protect PDFs with Auto-labeling** banner in the Microsoft Purview portal, or by using PowerShell. If PDF labeling isn't enabled, PDFs are skipped entirely: PDFs aren't evaluated or sent for labeling, and no label is applied. For instructions, see Adding support for PDF.
- Maximum of 500,000 automatically labeled files in your tenant per day.
- Maximum of 100 auto-labeling policies per tenant. In the portal, each policy can target up to 1,000 explicitly included or excluded SharePoint sites, or up to 100 explicitly included or excluded OneDrive individual users or groups. If you keep the default configuration of **All**, this configuration is exempt from these maximums.
- For SharePoint, an auto-labeling policy can use an adaptive scope to target up to 50,000 sites. For configuration instructions, see Use PowerShell for auto-labeling policies.
- Existing values for modified, modified by, and the date aren't changed as a result of auto-labeling policies—for both simulation mode and when labels are applied.

- When the label applies encryption, the Rights Management issuer and Rights Management owner is the account that last modified the file.
- To apply default sensitivity labels from SharePoint document libraries to existing files, see Apply default sensitivity labels from SharePoint document libraries to existing files.

Auto-labeling for Exchange

- PDF attachments and Office attachments are scanned for the conditions you specify in your auto-labeling policy. When there's a match, the email is labeled but not the attachment.
  - For PDF files, if the label applies encryption, these files are encrypted by using Message encryption when your tenant is enabled for PDF attachments.
  - For these Office files, Word, PowerPoint, and Excel are supported. If the label applies encryption and these files are unencrypted, they're now encrypted by using Message encryption. The encryption settings are inherited from the email.
- To label and protect emails that contain Teams voicemail messages, see the configuration instructions in Enable protected voicemail in your organization.
- If you have Exchange mail flow rules or Microsoft Purview Data Loss Prevention (DLP) policies that apply IRM encryption: When content is identified by these rules or policies and an auto-labeling policy, the label is applied. If that label applies encryption, the IRM settings from the Exchange mail flow rules or DLP policies are ignored. However, if that label doesn't apply encryption, the IRM settings from the mail flow rules or DLP policies are applied in addition to the label.
- Email that has IRM encryption with no label will be replaced by a label with any encryption settings when there's a match by using auto-labeling.
- Incoming email is labeled when there's a match with your auto-labeling conditions. For this outcome to apply to senders outside your organization, the Exchange location must be set to **All** included and **None** excluded. If the label is configured for encryption:
  - That encryption is always applied when the sender is from your organization.
  - By default, that encryption isn't applied when the sender is outside your organization but can be applied by configuring **Additional settings for email** and specifying a Rights Management owner.
- When the label applies encryption, the Rights Management issuer and Rights Management owner is the person who sends the email when the sender is from your own organization. When the sender is outside your organization, you can specify a Rights Management owner for incoming email that's labeled and encrypted by your policy.
- If the label is configured to apply content markings with variables, be aware that for incoming email, this configuration can result in displaying the names of people outside

your organization.

> **Note**
>
> For some new customers, we're offering the automatic configuration of default auto-labeling settings for both client-side labeling and service-side labeling. Even if you're not eligible for this automatic configuration, you might find it useful to reference their configuration. For example, you can manually configure existing labels and create your own auto-labeling policies with the same settings to help accelerate your labeling deployment.
>
> For more information, see Default labels and policies for Microsoft Purview Information Protection.

Compare auto-labeling for Office apps with auto-labeling policies

Use the following table to help you identify the differences in behavior for the two complementary automatic labeling methods:

Expand table

Feature or behavior	Label setting: Auto-labeling for files and emails	Policy: Auto-labeling
App dependency	Yes (minimum versions)	No *
Restrict by location	No	Yes
Conditions: Sharing options and additional options for email	No	Yes
Conditions: Exceptions	No	Yes
Conditions: Nested rule logic (AND/OR/NOT)	No	Yes
Support for PDF files	No	Yes
Support for images	No	Yes
Recommendations, policy tooltip, and user overrides	Yes	No
Simulation mode	No	Yes
Exchange attachments checked for conditions	No	Yes

Feature or behavior	Label setting: Auto-labeling for files and emails	Policy: Auto-labeling
Apply visual markings	Yes	Yes (email only)
Override IRM encryption applied without a label	Yes if the user has the minimum usage right of Export	Yes (email only)
Label incoming email	No	Yes
Assign a Rights Management owner for emails sent from another organization	No	Yes
Replace manually applied label that has lower priority	No	Yes (configurable)
Remove existing label in SharePoint or OneDrive	No	Yes

* Auto-labeling isn't currently available in all regions because of a backend Azure dependency. If your tenant can't support this functionality, the **Auto-labeling** page isn't visible in the Microsoft Purview portal. For more information, see Azure dependency availability by country.

Licensing and where to find each surface

Auto-labeling is gated by license, and the two surfaces (client-side and service-side) appear in different places in the portal. If you expect to see auto-labeling and it isn't there, verify the license first.

Expand table

Surface	Configured in	Included with
Client-side auto-labeling (Office apps)	Sensitivity label configuration → **Auto-labeling for files and emails** page	Microsoft 365 E5, Microsoft 365 E5 Compliance, Microsoft 365 E5 Information Protection & Governance, Azure Information Protection Premium P2 (standalone)
Service-side auto-labeling (SharePoint, OneDrive, Exchange)	Microsoft Purview portal → **Information Protection** → **Policies** → **Auto-labeling policies**	Same E5 / P2-tier SKUs as above
Container-level default labels (Teams, groups, sites)	Documented in Use sensitivity labels with Microsoft Teams, Microsoft 365 groups, and SharePoint sites	Not covered on this page

For the authoritative, per-SKU feature matrix (including current entitlements for Microsoft 365 Business Premium, Frontline, Education, and Government clouds), see Microsoft 365 licensing guidance for security and compliance.

Common licensing symptoms:

- **You have E3 or Business Premium.** Auto-labeling of files and emails isn't included at this SKU. Sensitivity labels and manual labeling are — auto-labeling requires an E5 / P2-tier add-on.
- **You have E5 but don't see auto-labeling policies in the portal.** Confirm you're in the Microsoft Purview portal (not the legacy Compliance Center), that your role includes **Compliance Administrator** or **Compliance Data Administrator**, and that your region supports auto-labeling (Azure dependency availability by country).
- **You have E5 and see the page, but saving a new policy fails with "No matching database is available."** This is a known transient backend condition. Retry the save; if it fails repeatedly over 24 hours, open a support case with the policy configuration.

How multiple conditions are evaluated when they apply to more than one label

The labels are ordered for evaluation according to their position that you specify in the Microsoft Purview portal: The label positioned first has the lowest position (least sensitive, so lowest priority) and the label positioned last has the highest position (most sensitive, so highest priority). The label with the highest order number is selected.

This behavior is also true for service-side auto-labeling (auto-labeling policies) when sublabels share the same parent label: If after evaluation and ordering, more than one sublabel from the same parent label meets the auto-labeling conditions, the sublabel with the highest order number is selected and applied.

However, the behavior is a little different for client-side auto-labeling (auto-labeling settings in the label). If multiple sublabels from the same parent label match the conditions:

- If a file is not already labeled, the highest order sublabel that's configured for automatic labeling is always selected, rather than the highest order sublabel that's configured for recommended labeling. If none of these sublabels are configured for automatic labeling but only recommended labeling, the highest order sublabel is selected and recommended.

- If a file is already labeled with a sublabel from the same parent, no action is taken and the existing sublabel remains. This behavior applies even if the existing sublabel was a default

label or automatically applied.

For more information about label priority, see Label priority (order matters).

Considerations for label configurations

The following considerations apply to both client-side labeling and service-side labeling.

Don't configure a parent label to be applied automatically or recommended

If you're using parent labels with sublabels, rather than label groups: You can't apply a parent label (a label with sublabels) to content. Make sure that you don't configure a parent label to be auto-applied or recommended in Office apps, and don't select a parent label for an auto-labeling policy. If you do, the parent label won't be applied to content.

To use automatic labeling with sublabels that have parent labels, make sure you publish both the parent label and the sublabel.

For more information about parent labels, label groups, and sublabels, see Sublabels that use parent labels or label groups.

Label scoping that excludes files or emails

To automatically apply a sensitivity label to content, the label's scope must include **Files & other data assets** to automatically apply a label to documents, and **Emails** to automatically apply a label to emails.

For more information when you select just one of these label scopes, see Scope labels to just files or emails.

Will an existing label be overridden?

An auto-labeling policy that applies a sensitivity label won't remove an existing sensitivity label to leave content unlabeled. However, for SharePoint and OneDrive, an auto-labeling policy configured to automatically remove a label can remove the selected label, including when the label was manually applied, and leave the content unlabeled.

Default behavior whether automatic labeling will override an existing label:

- When content has been manually labeled, that label won't be replaced by automatic labeling.
- Automatic labeling will replace a lower priority sensitivity label that was automatically applied, but not a higher priority label.

  > **Tip**
  >
  > For example, the sensitivity label at the top of the list in the Microsoft Purview portal is named **Public** with an order number (priority) of 0, and the sensitivity label at the bottom of the list is named **Highly Confidential** with an order number (priority of 4). The **Highly Confidential** label can override the **Public** label but not the other way around.

For auto-labeling policies that apply a sensitivity label, you can select a setting to override a manually applied label that has lower priority. This setting isn't required for an auto-labeling policy configured to automatically remove a label.

Expand table

Existing label	Override with label setting: Auto-labeling for files and emails	Override with policy: Auto-labeling
Manually applied, lower priority	No	No by default, but configurable
Manually applied, higher priority	No	No
Automatically applied or default label, lower priority	Yes *	Yes
Automatically applied or default label, higher priority	No	No

* There's an exception for sublabels that share the same parent label

The configurable setting for auto-labeling policies is on the **Additional label settings** page.

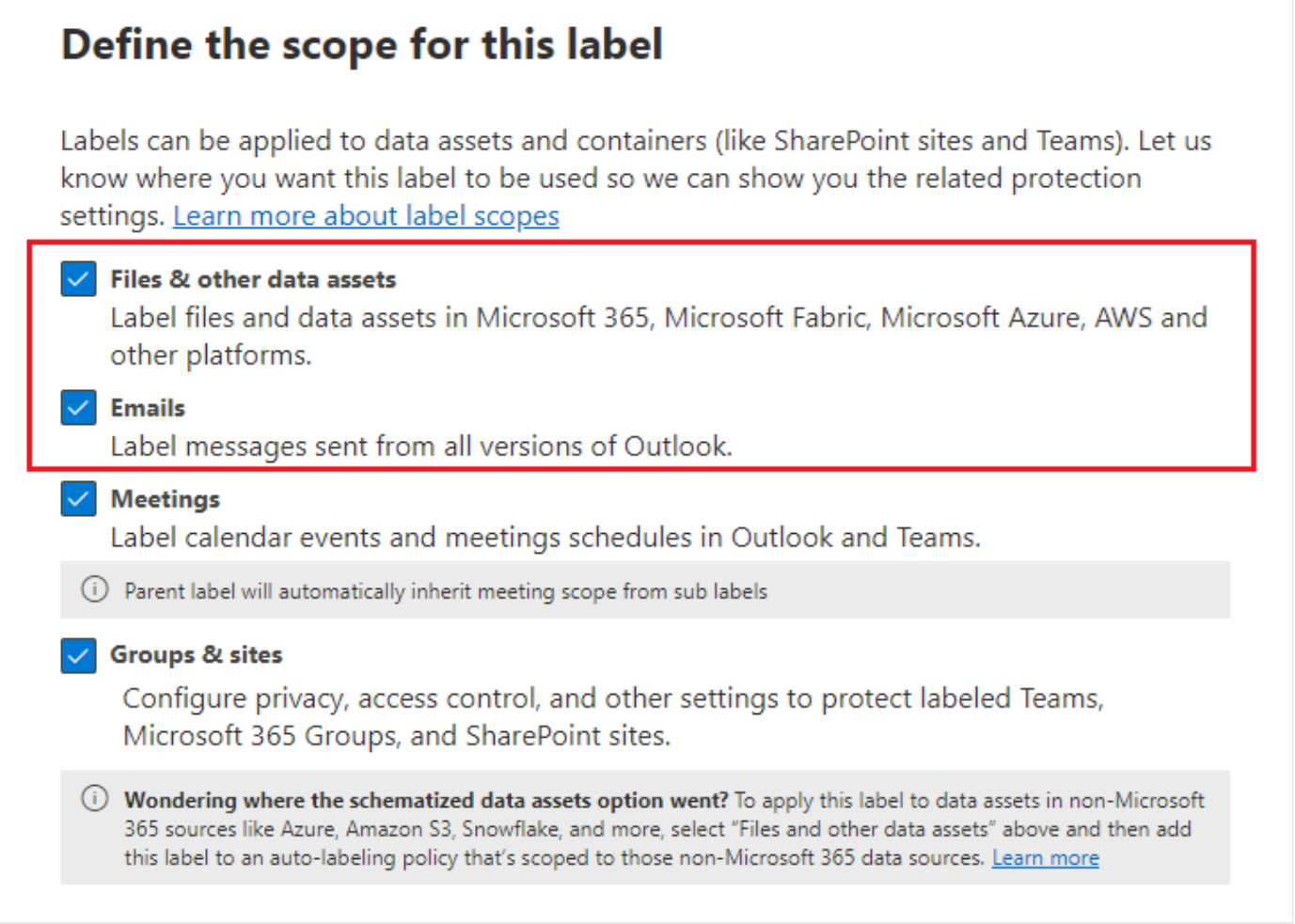
How to configure auto-labeling for Office apps

> **Tip**

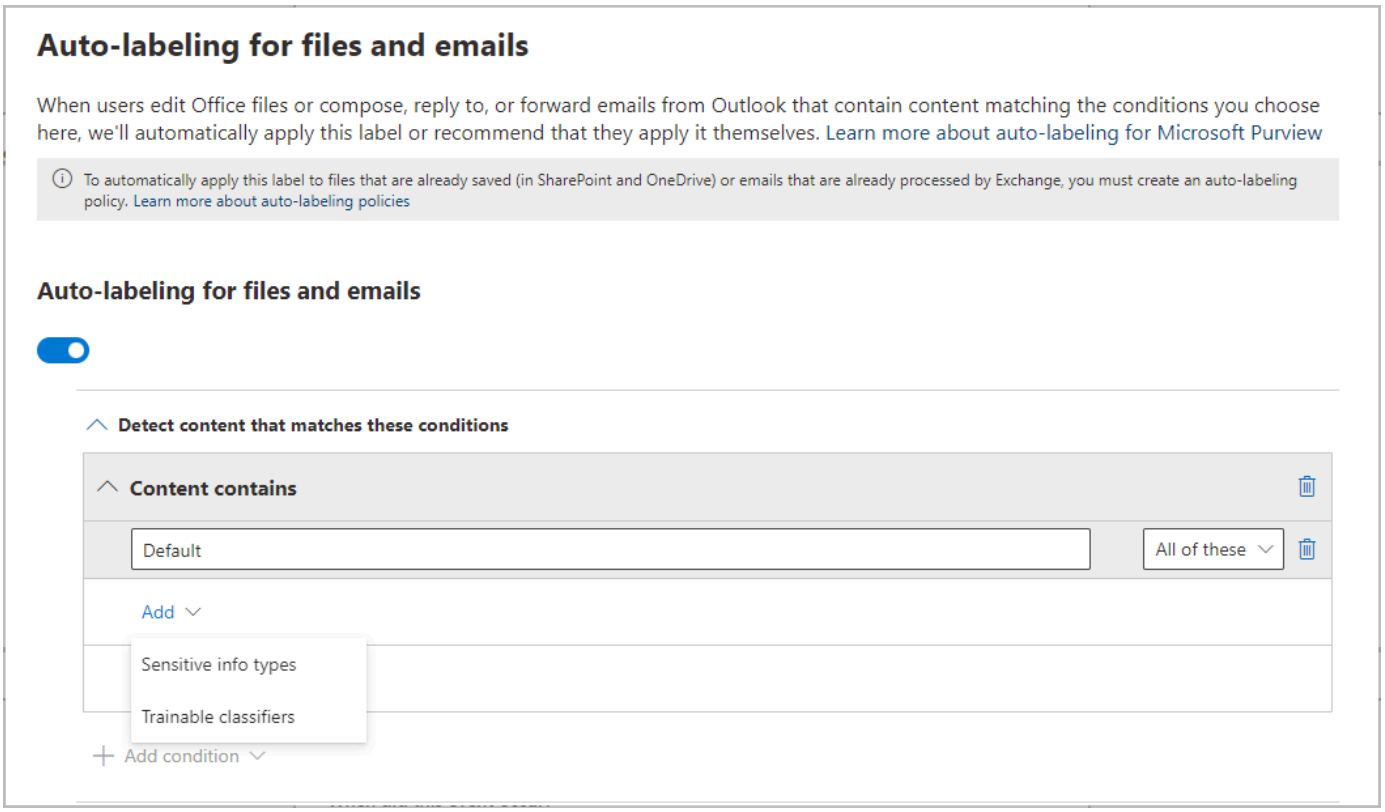
Looking for **auto-labeling policies for SharePoint, OneDrive, or Exchange** (service-side auto-labeling for files at rest and emails in transit)? Skip ahead to **How to configure auto-labeling policies for SharePoint, OneDrive, and Exchange**. The section on this page is for client-side auto-labeling in Word, Excel, PowerPoint, and Outlook.

Check the minimum versions required for automatic labeling in Office apps.

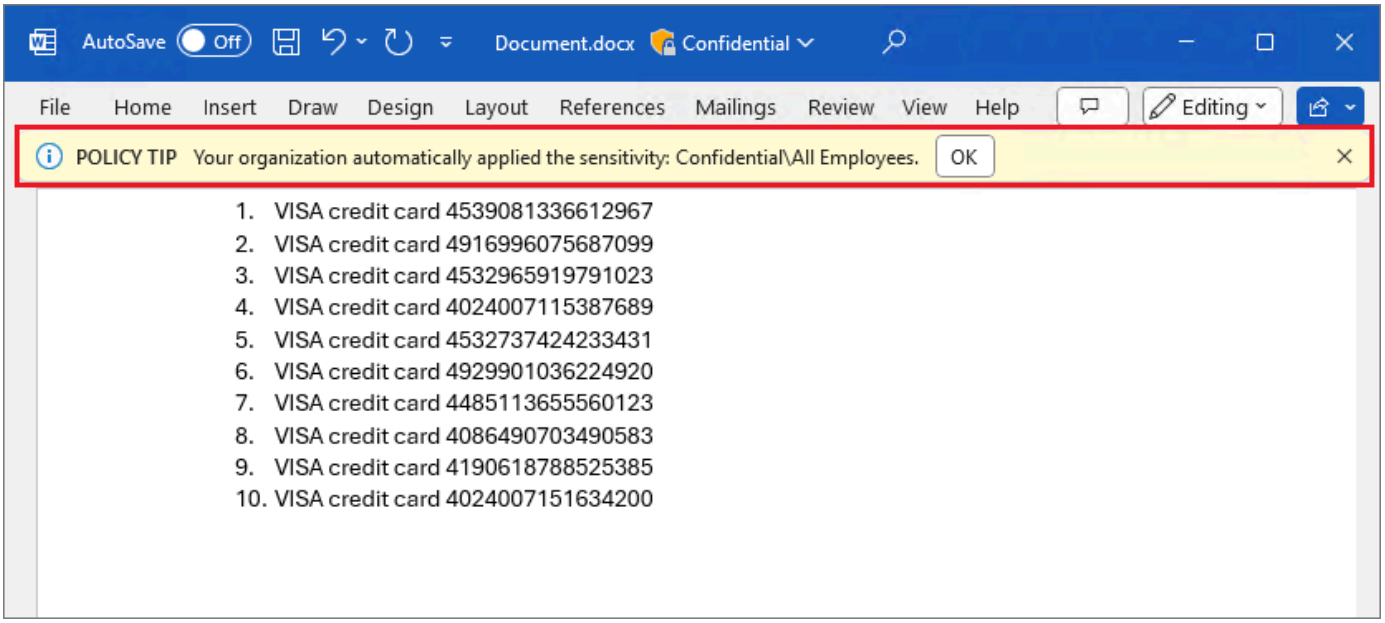
The auto-labeling settings for Office apps are available when you create or edit a sensitivity label. Make sure the label scope **Files & other data assets** is selected to auto-label documents, and **Emails** is selected to auto-label emails. For example:



As you move through the configuration, you see the **Auto-labeling for files and emails** page where you can choose from a list of sensitive info types or trainable classifiers:



When this sensitivity label is automatically applied, the user sees a notification in their Office app. For example:

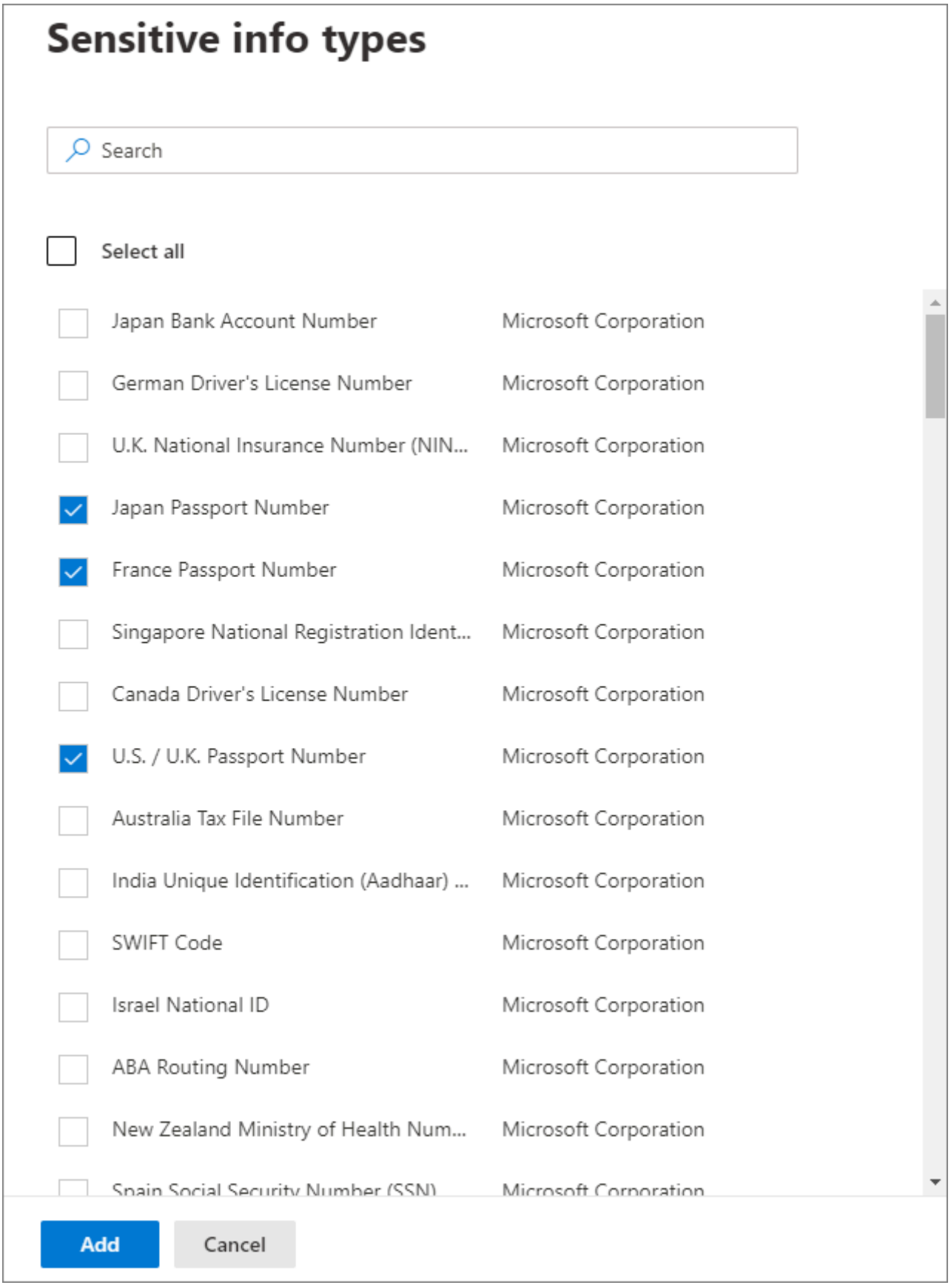


Because of a timing issue, auto-labeling for Office apps might not trigger if you also use mandatory labeling. For details, see the additional information for configuring the policy setting Require users to apply a label to their email and documents.

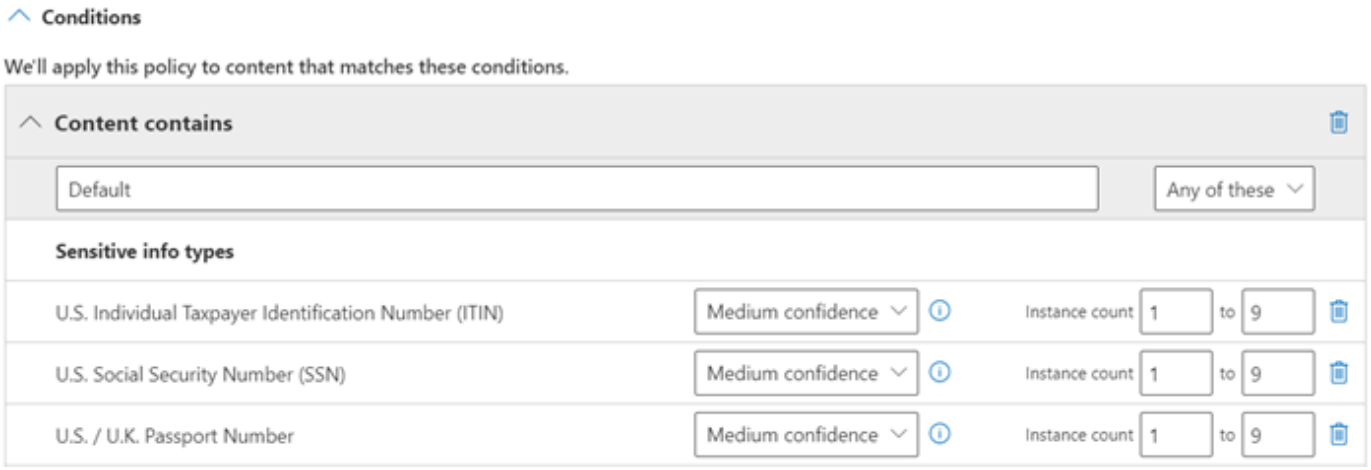
Configuring sensitive info types for a label

When you select the **Sensitive info types** option, you see the same list of sensitive information types as when you create a data loss prevention (DLP) policy. So you can, for example,

automatically apply a Highly Confidential label to any content that contains customers' personal information, such as credit card numbers, social security numbers, or passport numbers:



Similarly to when you configure DLP policies, you can then refine your condition by changing the instance count and match accuracy. For example:



You can learn more about confidence levels from the DLP documentation: More on confidence levels

> **Important**
>
> Sensitive information types have two different ways of defining the max unique instance count parameters. To learn more, see **Instance count supported values for SIT**.

Also similarly to DLP policy configuration, you can choose whether a condition must detect all sensitive information types, or just one of them. And to make your conditions more flexible or complex, you can add groups and use logical operators between the groups.

Custom sensitive information types with Exact Data Match

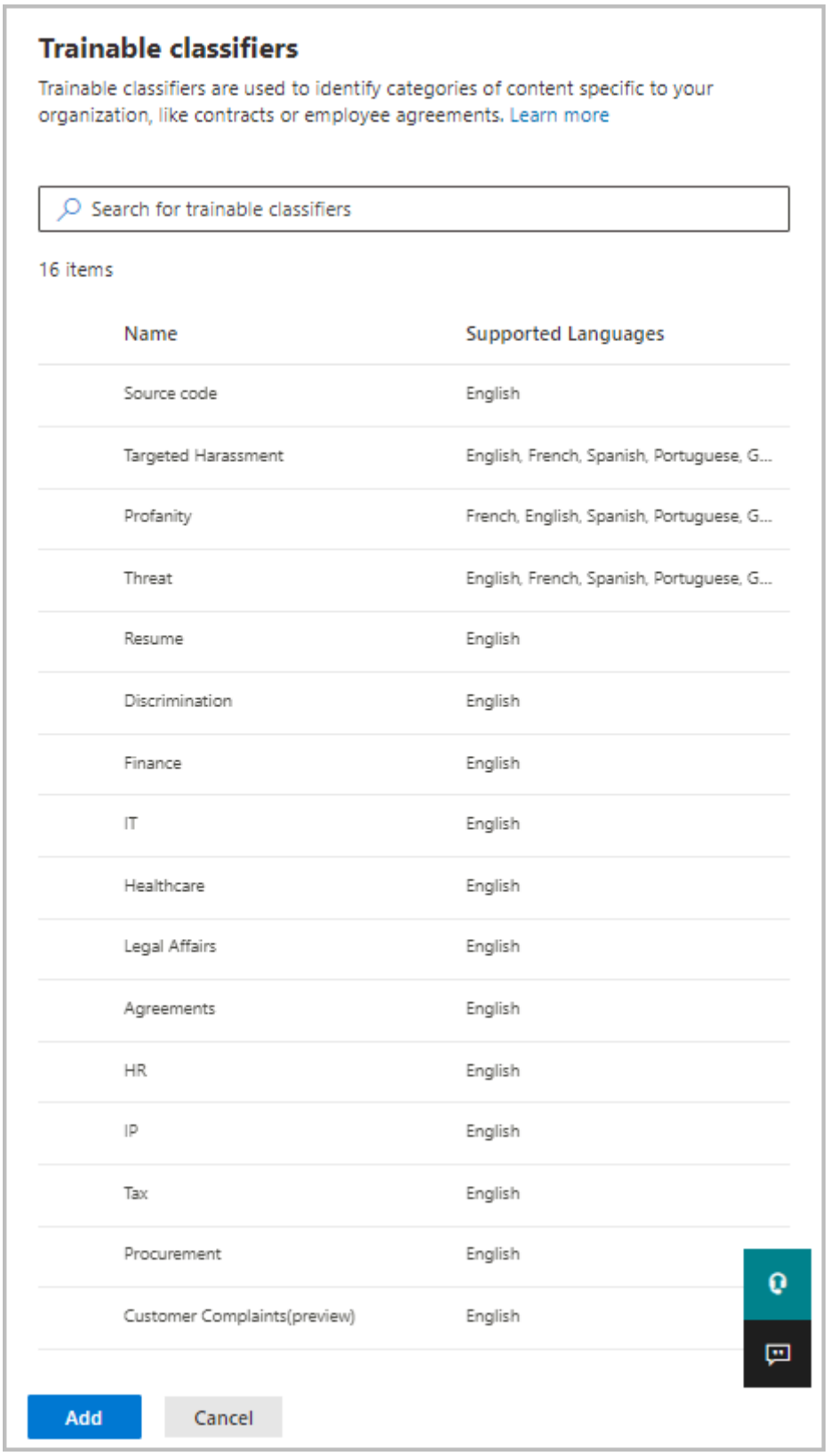
You can configure a sensitivity label to use exact data match based sensitive information types for custom sensitive information types. However, currently, you must also specify at least one sensitive information type that doesn't use EDM. For example, one of the built-in sensitive information types, such as **Credit card number**.

If you configure a sensitivity label with only EDM for your sensitive information type conditions, the auto-labeling setting is automatically turned off for the label.

Configuring trainable classifiers for a label

If you use this option with Microsoft 365 Apps for Windows version 2106 or lower, or Microsoft 365 Apps for Mac version 16.50 or lower, make sure you've published in your tenant at least one other sensitivity label that's configured for auto-labeling and the sensitive info types option. This requirement isn't necessary when you use later versions on these platforms.

When you select the **Trainable classifiers** option, select one or more of the pretrained or custom trainable classifiers:

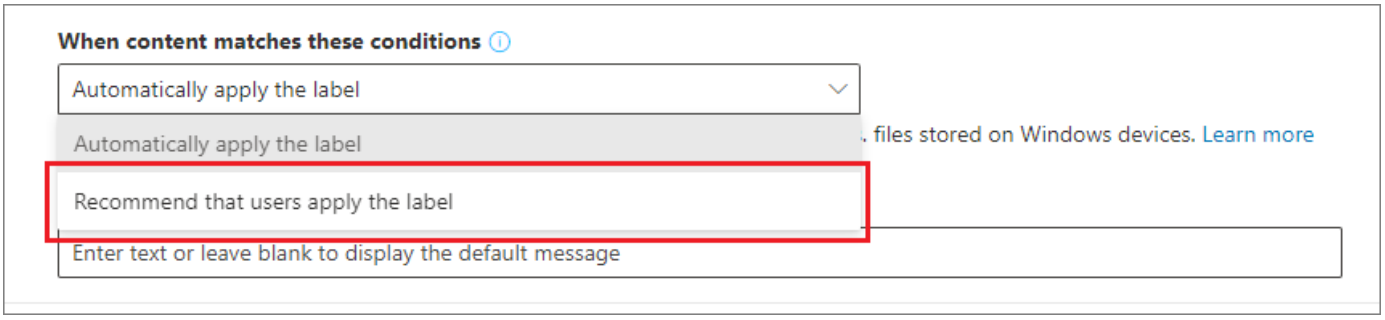


The available pretrained classifiers are often updated, so there might be more entries to select than the ones displayed in this screenshot.

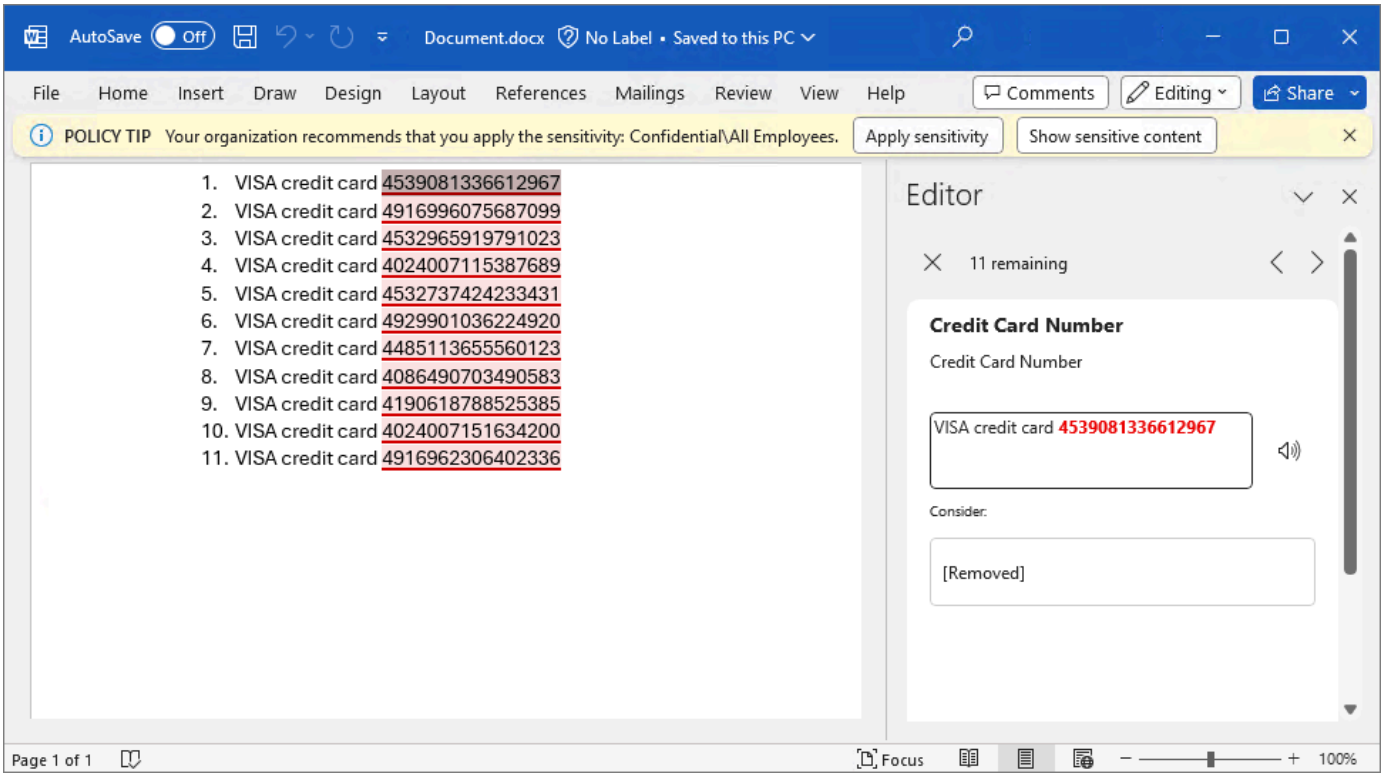
For more information about these classifiers, see Learn about trainable classifiers.

Recommend that the user applies a sensitivity label

If you prefer, you can recommend to your users that they apply the label. With this option, your users can accept the classification and any associated protection, or dismiss the recommendation if the label isn't suitable for their content.



When you use built-in labeling with the desktop versions of Word, users have an additional option to **Show sensitive content** with the recommended label prompt. When they select this button, the Editor pane steps the user through each detection. The user can then remove the sensitive data or leave it with a better understanding of why the sensitivity label was recommended. When they have this extra information, users have more confidence to select the **Apply sensitivity** button. For example:



This example shows the default recommended label prompt, but as with automatic labeling, you can customize this text to be more meaningful or specific for your users. For example, include your organization's name or reference your IT department to increase visibility and give users more confidence that this isn't a generic message that might not be applicable to them.

> **Tip**
>
> Although recommending a sensitivity label interrupts a user's workflow, it's a very efficient way to educate users in the moment about sensitive data that they work with. To see this in action, watch the video: **Automatically Classify & Protect Documents & Data**
>
> Recommended labeling is particularly powerful when it's coupled with this option to step the users through each instance of sensitive content that's detected. It can lead to more accurate labeling, not just for the immediate item, but also for future items that require manual labeling, or for modified items that might now need relabeling.

When automatic or recommended labels are applied

Not all Office apps support automatic and recommended labeling. For more information, see Support for sensitivity label capabilities in apps.

Other considerations:

- You can't use automatic labeling for documents and emails that were previously manually labeled, or previously automatically labeled with a higher sensitivity. Remember, you can only apply a single sensitivity label to a document or email (in addition to a single retention label).
- You can't use recommended labeling for documents or emails that were previously labeled with a higher sensitivity. When the content's already labeled with a higher sensitivity, the user won't see the prompt with the recommendation and policy tip.
- For recommended labels in the desktop versions of Word, the sensitive content that triggered the recommendation is flagged so that users can review and remove the sensitive content instead of applying the recommended sensitivity label.

For details about how these labels are applied in Office apps, example screenshots, and how sensitive information is detected, see Automatically apply or recommend sensitivity labels to your files and emails in Office.

Convert your label settings into an auto-labeling policy

If the label includes sensitive info types for the configured conditions, you'll see an option at the end of the label creation or editing process to automatically create an auto-labeling policy that's based on the same auto-labeling settings.

However, if the label contains trainable classifiers as a label condition:

- When the label conditions contain just trainable classifiers, you won't see the option to automatically create an auto-labeling policy.
- When the label conditions contain trainable classifiers and sensitivity info types, an auto-labeling policy will be created for just the sensitive info types.

Although an auto-labeling policy is automatically created for you by auto-populating the values that you would have to select manually if you created the policy from scratch, you can still view and edit the values before they are saved.

By default, all locations for SharePoint, OneDrive, and Exchange are included in the auto-label policy, and when the policy is saved, it runs in simulation mode. There's no check that you've enabled sensitivity labels for Office files in SharePoint and OneDrive, which is one of the prerequisites for auto-labeling to apply to content in SharePoint and OneDrive.

How to configure auto-labeling policies for SharePoint, OneDrive, and Exchange

> **Note**
>
> Don't use Exchange auto-labeling policies to send encrypted emails for mass mailing distributions. These policies aren't designed for this purpose and can result in sending failures and non-delivery receipts. For this scenario, the label setting to automatically send emails is more suitable.

Make sure you're aware of the prerequisites before you configure auto-labeling policies. Use the Pre-flight checklist to confirm each one in a single scan, then review the full Prerequisites for auto-labeling policies below for details.

Pre-flight checklist

Before you create an auto-labeling policy, confirm each item below. Missing any one of them typically causes the policy to run without labeling any files, with no error surfaced in the portal. Each item links to the full guidance further down this article.

- [ ] **Auditing is turned on for Microsoft 365.** Required for simulation. See Turn audit log search on or off.

- [ ] **Sensitivity labels are enabled for Office files in SharePoint and OneDrive.** Required before any label can be applied to files in those locations. See Enable sensitivity labels for Office files in SharePoint and OneDrive.
- [ ] **Auto-labeling is available in your region.** If the **Auto-labeling** page isn't visible in the Microsoft Purview portal, your tenant is in an unsupported region. See Azure dependency availability by country.
- [ ] **The label you plan to apply is not a parent label.** Parent labels (labels that have sublabels) can't be applied to content. If you select one, the policy will run but nothing will be labeled. See Don't configure a parent label to be applied automatically or recommended.
- [ ] **The label's scope matches your target content.** To label files, the label scope must include **Files & other data assets**. To label emails, it must include **Emails**. See Label scoping that excludes files or emails.
- [ ] **If the label applies encryption to SharePoint or OneDrive content:** the label is configured with **Assign permissions now**, and **User access to content expires** is set to **Never**. See the encryption details in Prerequisites for auto-labeling policies.
- [ ] **Classification covers the sensitive information types you intend to use.** Policies can evaluate existing files using available classification results, but changing a custom SIT doesn't by itself refresh those results for unchanged files. Review when historical files need updated classification, including the effects of on-demand classification on active policies.
- [ ] **You have the roles required to review simulation results.** Viewing file contents in the source view requires the **Data Classification Content Viewer** role (included in the Content Explorer Content Viewer, Information Protection, and Information Protection Investigators role groups). Global admins don't have this role by default.
- [ ] **You have the roles required to turn the policy on after simulation.** Turning a policy from **Ready to turn on** to **On** requires **Compliance Administrator** or **Compliance Data Administrator**. Without one of these roles, the **Turn on policy** button appears greyed out even after a successful simulation.

When every item is confirmed, continue to Prerequisites for auto-labeling policies for the full details, then to Creating an auto-labeling policy.

Prerequisites — detailed reference

Every prerequisite for an auto-labeling policy is listed below. Confirm each item before you configure a policy — missing any one typically causes the policy to run without labeling any files, with no in-portal error surfaced.

- Simulation mode:

    - Auditing for Microsoft 365 must be turned on. If you need to turn on auditing or you're not sure whether auditing is already on, see Turn audit log search on or off.
    - To view file or email contents in the source view, you must have the **Data Classification Content Viewer** role, which is included in the **Content Explorer Content Viewer** role group, or **Information Protection** and **Information Protection Investigators** role groups. Without the required role, you don't see the preview pane when you select an item from the **Items to review** tab. Global admins don't have this role by default.

- To auto-label files in SharePoint and OneDrive:

    - You have enabled sensitivity labels for Office files in SharePoint and OneDrive.

    - If you want the policy to process PDFs, you have enabled PDF support by using the portal banner or PowerShell. Otherwise, PDFs are skipped without being evaluated or sent for labeling.

    - For OneDrive, the OneDrive account must be associated with the corresponding user account. A OneDrive account that isn't associated with its user account (sometimes called a detached OneDrive account) isn't supported for auto-labeling or simulation. Use a properly associated OneDrive account, or move the content to a supported SharePoint or OneDrive location, and then rerun simulation.

    - At the time the auto-labeling policy runs, the file mustn't be open by another process or user. A file that's checked out for editing falls into this category, and similarly, all files in a library that's configured to require documents to be checked out.

    - If historical files haven't been classified or need updated classification, use on-demand classification. This isn't required for every policy change. Review scan scope, estimates, and active-policy effects before starting classification.

    - If you plan to use sensitive information types:
        - Policy evaluation uses available classification results for existing SharePoint and OneDrive files. Changing a custom SIT definition doesn't by itself reclassify unchanged historical files, and a file's last-modified date alone doesn't establish its classification time. See Policy changes aren't classifier changes.
        - Validate custom SITs with representative matching and nonmatching files by using Test sensitive information types. Then review policy simulation separately; a direct SIT test doesn't establish classification coverage for existing files.

- If you plan to use document properties as a condition (**Document property is**), this option uses SharePoint managed properties in the same way as they are used for DLP policies. Use exact string matches; regex patterns aren't supported. For more information about managed properties as a search method, see Manage the search schema in SharePoint.

  > **Important**
  >
  > When a document property is mapped to a SharePoint managed property, such as `RefinableString11`, values containing an ampersand (`&`) might not match an auto-labeling rule that uses **Document property is**. For example, documents with the mapped value `Billing & payment document` might remain unlabeled. Plan around this known limitation rather than relying on a future fix. Consider using a classification value without ampersands, keeping the document metadata and rule value consistent. Before deploying broadly, review simulation results and confirm actual label application in a limited scope.

- If you have an auto-labeling policy configured to replace a manually applied label but another active auto-labeling policy is configured to apply a label without this option, the label won't be replaced. Because simulation shows the result of one policy, simulation results for the replace label policy show the label as replaced but the conflict is resolved when all policies run.

- If you have an auto-labeling policy configured to remove a label but another active auto-labeling policy is configured to apply a label and the conditions are matched, the label will be applied rather than removed. This behavior applies even when the apply-label policy applies a different label than the label selected for removal. Because simulation shows the result of one policy, simulation results for the label removal policy show the label as removed but the conflict is resolved when all policies run.

- If you choose to remove a label that applies encryption, encryption is automatically removed with the label.

- One or more sensitivity labels created and published (to at least one user) that you can select for your auto-labeling policies. For these labels:
  - It doesn't matter if the auto-labeling in Office apps label setting is turned on or off, because that label setting supplements auto-labeling policies, as explained in the introduction.

- If the labels you select are configured to use visual markings (headers, footers, watermarks), these markings aren't applied to documents when you use auto-labeling.
- If the labels apply encryption:
  - When the auto-labeling policy includes locations for SharePoint or OneDrive, the label must be configured for the **Assign permissions now** setting, and **User access to content expires** must be set to **Never**.
  - When the auto-labeling policy is just for Exchange, the label can be configured for either **Assign permissions now** or **Let users assign permissions** (for the Do Not Forward or Encrypt-Only options). You can't auto-apply a label that's configured to apply S/MIME protection.

Scoping a policy with users and groups

Scope options differ by workload. For supported selections and Exchange sender-scope behavior, see **Choose locations** in Creating an auto-labeling policy.

Learn about simulation mode

Simulation mode is supported for auto-labeling policies and woven into the workflow. You can't automatically label documents and emails, or remove labels, until your policy has run at least one simulation.

For existing SharePoint and OneDrive files, simulation evaluates policy matches using available classification results. Restarting simulation doesn't itself reclassify the content. See Auto-labeling classification and simulation concepts for when to change a policy and when to refresh classification.

> ⓘ **Important**
>
> **What "simulation mode" does — and doesn't — do**
>
> Simulation is designed to evaluate and log matches without applying (or removing) labels. It is not a fully silent dry run. Before you run a simulation, know these behaviors that customers frequently misdiagnose as bugs:
>
> - **Simulation still generates activity alerts.** If a Microsoft Purview alert policy is configured on auto-labeling or sensitive-information-type activity, matches surfaced during simulation trigger those alerts. To avoid mass alert email while validating a policy, temporarily scope or disable the alert policy for the simulation window.

- **Policy actions depend on status and permissions.** If **Turn on policy**, **Edit**, or **Delete** is unavailable, check the current policy status and your permissions. For **Turn on policy**, also check the simulation status. Review any displayed error before retrying.
- **A duplicated policy lands in simulation by default and doesn't notify you.** Duplicating an existing policy creates a new policy in simulation regardless of the source policy's state. If you expect the duplicate to inherit the source's active state, it won't.
- **Simulation and DLP share the same test-mode surface.** Some **Items to review** views can mix auto-labeling and DLP results and can fail to load if either signal is under load. If a view returns an error, retry a few hours later before you open a support case.
- **Simulation limits are documented separately.** The following guidance covers item caps (100 per site displayed, 50,000-record CSV export ceiling, and 20,000,000-item scan limit), the 12-hour target completion time, and stall detection.

Simulation mode supports up to 20,000,000 matched files. If more than this number of files are matched from an auto-labeling policy, you can't turn on the policy to apply the labels. In this case, you must reconfigure the auto-labeling policy so that fewer files are matched, and rerun simulation. This maximum of 20,000,000 matched files applies to simulation mode only and not to an auto-labeling policy that's already turned on to apply sensitivity labels.

> ⓘ **Note**
>
> Simulation results can differ from what happens when the policy is turned on:
>
> - Simulation shows the result of a *single* policy. When more than one auto-labeling policy applies to the same content (for example, one policy replaces an existing label and another applies a label without that option), the conflict is resolved only when *all* policies run—so the enforced result can differ from any single policy's simulation.
> - Classification results or file contents can change after simulation. A simulation match doesn't prove that classification used the latest SIT definition or that enforcement will successfully label the file. See Classification, policy evaluation, and enforcement.
> - For Exchange, simulation evaluates email that's sent and received during the run, so email results aren't consistent unless the same messages are sent and received again.

Workflow for an auto-labeling policy:

> ⓘ **Important**

Before you edit an auto-labeling policy or run simulation again, export any available results that you need for later comparison. See **Preserve results before editing or rerunning**.

1. Create and configure an auto-labeling policy.
2. Run the policy in simulation mode, which can take 12 hours to complete. The completed simulation triggers an email notification that's sent to the user configured to receive activity alerts.
3. Review the results, and if necessary, refine your policy. For example, you might need to edit the policy rules to reduce false positives, or remove some sites so that the number of matched files doesn't exceed 20,000,000. Rerun simulation mode and wait for it to complete again.
4. Repeat step 3 as needed.
5. Deploy in production.

The simulated deployment runs like the WhatIf parameter for PowerShell, for a specific point in time. You see results reported as if the auto-labeling policy had applied your selected label, using the rules that you defined. You can then refine your rules for accuracy if needed, and rerun the simulation. However, because auto-labeling for Exchange applies to emails that are sent and received, rather than emails stored in mailboxes, don't expect results for email in a simulation to be consistent unless you can send and receive the exact same email messages.

Simulation mode also lets you gradually increase the scope of your auto-labeling policy before deployment. For example, you might start with a single location, such as a SharePoint site, with a single document library. Then, with iterative changes, increase the scope to multiple sites, and then to another location, such as OneDrive.

For details about the newer simulation overview, see Review simulation results for auto-labeling policies. **Sample items**, when displayed, shows the sampled item count. **Contextual summary** is available in supported portal experiences; viewing matched content requires the **Data Classification Content Viewer** role.

Finally, you can use simulation mode to provide an approximation of the time needed to run your auto-labeling policy, to help you plan and schedule when to run it without simulation mode.

Note

If simulation doesn't include expected files, check the policy's scope and conditions, whether files changed after the run, and whether classification reflects the SIT definitions you intend to use. Rerun simulation to review updated policy matches. Rerunning doesn't refresh classification or confirm successful labeling; see **When historical files need updated classification**.

Apply default sensitivity labels from SharePoint document libraries to existing files

SharePoint document libraries can be configured with a default sensitivity label. SharePoint applies this label to new files and to files that users edit in the library. Files that already existed in the library before the default label was configured might remain unlabeled.

To extend your default library label deployment to existing files, create an auto-labeling policy that applies the default sensitivity label from each selected SharePoint document library to eligible existing files at rest. This policy type doesn't inspect file content to choose a label. The label comes from the SharePoint document library's default sensitivity label setting.

Use this configuration when:

- You already use default sensitivity labels for SharePoint document libraries.
- You want the default library label to apply to files that already exist in those libraries.
- You want service-side labeling for files at rest, instead of relying on users to open or edit each file.
- You want to apply the label selected by the SharePoint library configuration, rather than build separate sensitive information type or classifier conditions.

Before you create this type of policy, confirm these requirements:

- Sensitivity labels are enabled for Office files in SharePoint and OneDrive. For more information, see Enable sensitivity labels for Office files in SharePoint and OneDrive.
- The sensitivity labels that can be used as default library labels are created and published to the appropriate users or groups.
- One or more SharePoint document libraries in your target sites have a default sensitivity label configured.
- You have permissions to create and manage auto-labeling policies.
- You have permission to read the SharePoint site and document library metadata required to identify libraries with default sensitivity labels. For the currently supported configuration

path, use Compliance Administrator or an existing SharePoint or Microsoft Entra role with this access, such as SharePoint Administrator.

When you create the policy, select the option to apply default sensitivity labels from SharePoint document libraries to existing files. Select specific SharePoint sites rather than **All**. This policy type doesn't support administrative-unit scoping, so the simplified wizard doesn't include the **Assign admin units** step.

In portal experiences that validate default library labels, each selected SharePoint site needs at least one library with a default sensitivity label. If a site can't be added, use the displayed error to choose your next step:

- **No default library label is configured:** Choose another site, or configure a default sensitivity label for a document library in that site, then try again.
- **Default labels couldn't be verified:** Check your access to the site and its library metadata, then try again. This message doesn't mean that the site has no default library label.

For example:

1. A SharePoint document library has the default sensitivity label **Confidential**.
2. The library contains files that were uploaded before **Confidential** was configured as the library default.
3. You create an auto-labeling policy that uses the default SharePoint document library label configuration.
4. The policy evaluates existing files in that library.
5. Eligible files are labeled **Confidential** by the policy.

For this policy type, simulation identifies files in document libraries that have a default sensitivity label configured. Review the simulation results before you turn on the policy to confirm that the expected SharePoint sites, document libraries, default library labels, and files are included.

If the simulation results don't include files you expected, confirm that the document library has a default sensitivity label configured, the SharePoint site is included in the policy scope, and sensitivity labels are enabled for Office files in SharePoint and OneDrive.

Currently, you can create one auto-labeling policy of this type per tenant, and include up to 10 SharePoint sites in the policy. The policy evaluates only SharePoint document libraries that have a default sensitivity label configured, and applies labels only to eligible supported files. Files might not be labeled if they're open, checked out, locked, unsupported, encrypted in an unsupported way, or otherwise unavailable when the policy evaluates them.

If you need to label existing files across more than 10 SharePoint sites, create the policy for up to 10 sites first. After labeling for those sites is complete and you've reviewed the results, edit the policy to replace the site selection with the next batch of up to 10 sites.

Creating an auto-labeling policy

To create a policy, follow these steps. Page order and available settings can vary by portal experience; follow the settings shown in your wizard.

1. Sign in to the Microsoft Purview portal > **Solutions** > **Information Protection** > **Policies** > **Auto-labeling policies**.

   > **Note**
   >
   > If you don't see the option for auto-labeling, this functionality isn't currently available in your region because of a backend Azure dependency. For more information, see Azure dependency availability by country.

2. Select + **Create auto-labeling policy**. Choose whether you want to apply a sensitivity label, or remove a label. Then follow the instructions according to the option you chose.

Automatically apply label only

3. For the page **Choose info you want this label applied to**: Select one of the categories, such as **Financial**, and then select a corresponding regulation, such as **Canadian Financial Data**. You can refine your search by using the template search or dropdown box for countries or regions. Or, select **Custom** > **Custom policy** if the templates don't meet your requirements. Select **Next**.

4. For the page **Name your auto-labeling policy**: Provide a unique name, and optionally a description to help identify the automatically applied label, locations, and conditions that identify the content to label.

5. For the **Choose a label to auto-apply** page: Select + **Choose a label**, select and add a label from the **Choose a sensitivity label** pane, and then select **Next**.

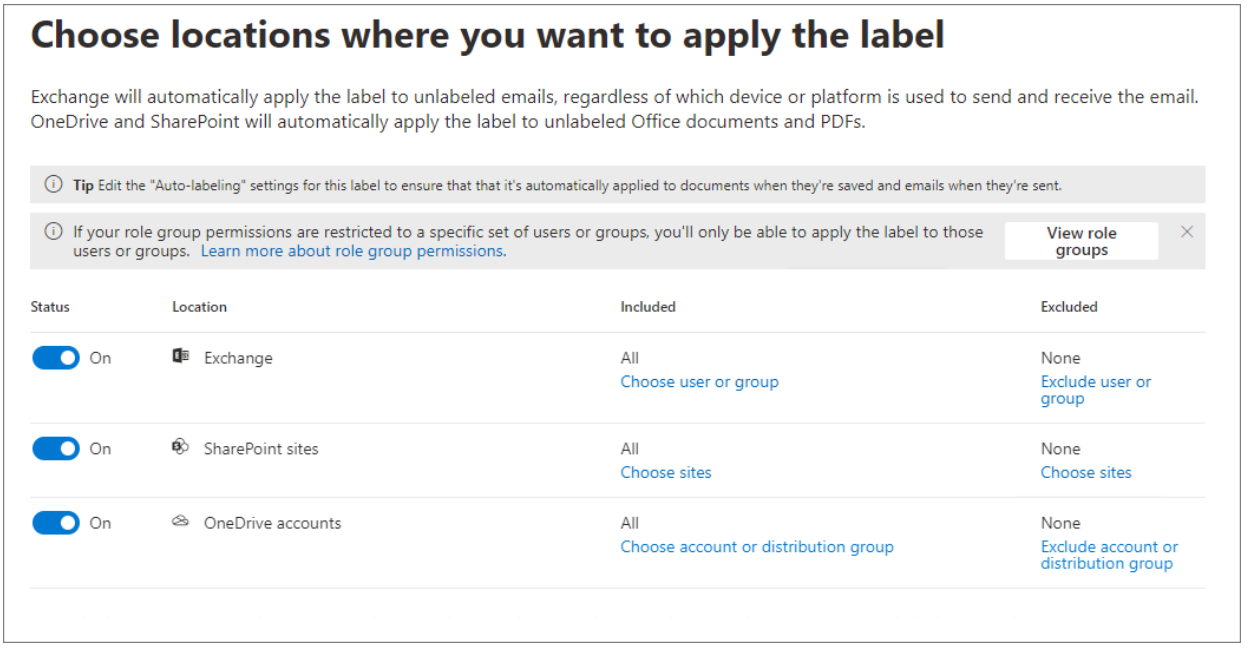
6. For the page **Assign admin units**: If your organization is using administrative units in Microsoft Entra ID, auto-labeling policies for Exchange and OneDrive can be automatically restricted to specific users and auto-labeling policies for SharePoint can

be automatically restricted to the specific sites configured for the selected administrative units. If your account has been assigned administrative units, you must select one or more administrative units.

If you don't want to restrict the policy by using administrative units, or your organization hasn't configured administrative units, keep the default of **Full directory**.

> **Note**
>
> If you are editing an existing policy and change the administrative units, you must now reconfigure the locations in the next step.

7. For the page **Choose locations where you want to apply the label**: Select and specify locations for Exchange, SharePoint, and OneDrive. If you don't want to keep the default of **All** included for your chosen locations, select the link to choose specific instances to include, or select the link to choose specific instances to exclude. Then select **Next**.



> **Note**
>
> For organizations that are using administrative units and selected in the policy:
>
> - For SharePoint, you can't use the **Included** or **Excluded** options and must use the **All** default selection.

- For Exchange and OneDrive, you can use the **Included** or **Excluded** options, but see and can select only users from the administrative units selected in the previous step.

If you use the **Included** or **Excluded** options:

- For the **Exchange** location, the policy is applied according to the sender address of the recipients specified. Most of the time, you'll want to keep the default of **All** included with **None** excluded. This configuration is suitable even if you're testing for a subset of users. Instead of specifying your subset of users here, use the advanced rules in the next step to configure conditions to include or exclude recipients in your organization. Otherwise, when you change the default settings here:
  - If you change the default of **All** included and instead, choose specific users or groups, email sent from outside your organization will be exempt from the policy.
  - If you keep the default of **All** included but specify users or groups to exclude, email that these excluded users send will be exempt from the policy, but not email that they receive.

- For the OneDrive location, you must specify users or groups. Previously, you had to specify sites by URLs. Any existing OneDrive URL sites in auto-labeling policies will continue to work but before you can specify new OneDrive locations, or for restricted admins, you must first delete any existing site URLs. Groups supported: distribution groups, Microsoft 365 groups, mail-enabled security groups, and security groups.

- For the SharePoint location, you can explicitly select up to 1,000 individual sites. To dynamically target up to 50,000 sites, use a SharePoint adaptive scope. For example, an adaptive scope can use the `SiteTemplate` property to include or exclude communication sites, group-connected sites, Teams private channel sites, classic team sites, or OneDrive sites. For instructions to associate an adaptive scope with an auto-labeling policy, see Use PowerShell for auto-labeling policies.

8. For the **Set up common or advanced rules** page: Keep the default of **Common rules** to define rules that identify content to label across all your selected locations. If you need different rules per location, including some rules that are only available for Exchange, or SharePoint sites and OneDrive accounts, select **Advanced rules**. Then select **Next**.

The rules use conditions that include sensitive information types, trainable classifiers, sharing options, and other conditions:

- To select a sensitive information type or trainable classifier as a condition, under **Content contains**, select **Add**, and then choose **Sensitive info types** or **Trainable classifiers**.
- To select sharing options as a condition, under **Content is shared**, choose either **only with people inside my organization** or **with people outside my organization**.
- Other conditions that you can select:
  - Attachment or file extension is
  - Attachment or document name matches words or phrases
  - Attachment or document property is
  - Attachment or document size equals or is greater than

In **Advanced rules**, the name condition is **Attachment name matches words or phrases** for Exchange, or **Document name contains words or phrases** for SharePoint and OneDrive.

If your location is **Exchange** and you selected **Advanced rules**, there are additional conditions that you can select:

- Sender IP address is
- Recipient domain is
- Recipient is
- Attachment is password protected
- Any email attachment's content could not be scanned
- Any email attachment's content didn't complete scanning
- Header matches patterns
- Subject matches patterns
- Recipient address contains words
- Recipient address matches patterns
- Sender address matches patterns
- Sender domain is
- Recipient is a member of
- Sender is

If your location is **SharePoint sites** or **OneDrive accounts** and you selected **Advanced rules**, there is one other condition that you can select:

- Document created by

9. Depending on your previous choices, you'll now have an opportunity to create new rules by using conditions and exceptions.

   The configuration options for sensitive information types are the same as those you select for auto-labeling for Office apps. If you need more information, see Configuring sensitive info types for a label.

   When you've defined the rules and confirmed their status is on, select **Next** and complete the remaining settings shown by the wizard.

10. Specify optional configurations on the **Additional label settings** page:

    - **Emails only**: Applicable for both incoming and outgoing emails. When you select this setting, an existing sensitivity label will always be overriden for emails if the sensitivity label you chose for the policy has a higher priority. Use this setting to override manually labeled emails with a higher priority label when sensitive content is detected.

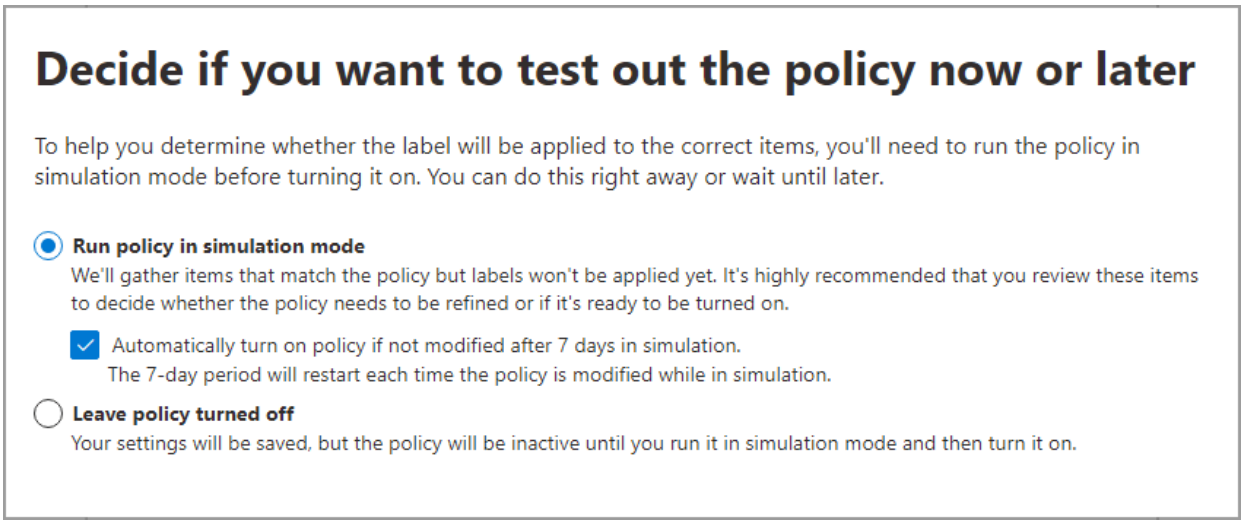
    - **All locations**: Applicable for emails and for files in SharePoint and OneDrive. When you select this setting, an existing sensitivity label will always be overriden if the sensitivity label you chose for the policy has a higher priority. Use this setting to override manually labeled emails and files with a higher priority label when sensitive content is detected.

    - **Apply encryption to email received from outside your organization**: This option is visible only when the selected sensitivity label applies encryption. If you then select this option to apply encryption to email received from outside your organization, you must assign a Rights Management owner. This configuration ensures that an authorized person in your organization has Full Control usage rights for these emails sent from your outside your organization and now encrypted. This role might be needed to later remove the encryption, or assign different usage rights for users in your organization.

      For **Assign a Rights Management owner**, specify a single user by an email address that's owned by your organization. Don't specify a mail contact, a shared mailbox, or any group type, because these aren't supported for this role.

11. For the **Decide if you want to test out the policy now or later** page: Select **Run policy in simulation mode** if you're ready to run the auto-labeling policy now, in simulation

mode. Then decide whether to automatically turn on the policy if it's not edited for 7 days:



If you're not ready to run simulation, select **Leave policy turned off**.

12. For the **Summary** page: Review the configuration of your auto-labeling policy and make any changes that you need, and complete the configuration.

Example: Apply a label to Exchange email based on the subject

In this example, an auto-labeling policy applies the **High Confidence** sensitivity label when an email subject matches the fictional project codename `Project Atlas`.

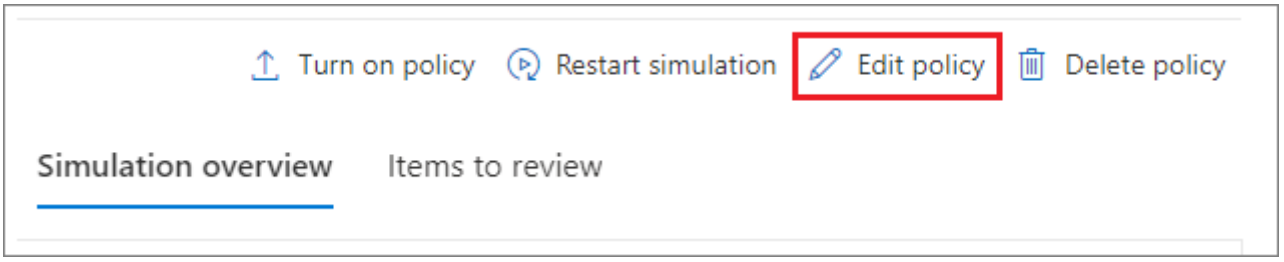
1. Make sure the **High Confidence** sensitivity label is published to at least one user and its scope includes **Emails**.
2. Follow the preceding steps to create a policy that automatically applies a label. Select **Custom** > **Custom policy**, and then select the **High Confidence** label.
3. On the **Choose locations where you want to apply the label** page, select **Exchange**. Keep the default of **All** included and **None** excluded if the policy must evaluate incoming email from outside your organization.
4. On the **Set up common or advanced rules** page, select **Advanced rules**. Create a rule for Exchange, add the **Subject matches patterns** condition, and enter a regular expression that matches `Project Atlas` in the subject. For information about regular expressions for this condition, see SubjectMatchesPatterns.
5. Run the policy in simulation mode and send representative test messages while simulation is running. Review the matching messages on the **Items to review** tab, and then turn on the policy.

> **Note**
>
> You don't need an Exchange mail flow rule or a data loss prevention (DLP) policy for this configuration. Exchange auto-labeling policies evaluate messages in transit as they're sent or received; they don't evaluate existing messages stored in mailboxes. Similarly, simulation evaluates messages sent or received while the simulation is running.

Now on the **Information protection** > **Auto-labeling** page, you see your auto-labeling policy in the **Simulation** or **Off** section, depending on whether you chose to run it in simulation mode or not. Select your policy to see the details of the configuration and status (for example, **Policy simulation is still running**). For policies in simulation mode, select the **Items to review** tab to see which emails or documents matched the rules that you specified.

Before selecting **Edit policy**, preserve simulation results you need for comparison. You can modify your policy directly from this interface:

- For a policy in the **Off** section, select the **Edit policy** button.
- For policy in the **Simulation** section, select the **Edit policy** option at the top of the page, from either tab.



When you're ready to run the policy without simulation, select the **Turn on policy** option.

> **Note**
>
> Files with no content (i.e. zero bytes in size) may not be labeled.

Auto-labeling policies run continuously until they're deleted. For example, new and modified files will be included with the current policy settings.

Removing or downgrading sensitivity labels with an auto-labeling policy

An auto-labeling policy can also be used to **remove** a sensitivity label from files that match its conditions — for example, to strip an encryption-enforcing label before content is migrated, shared externally, or reclassified. This mode is powerful, and it also has documented behaviors that customers frequently misdiagnose as a bug when a large removal run appears to do nothing after activation.

Before you turn on a removal or downgrade policy, know these behaviors:

- **A simulation match doesn't guarantee successful label removal.** Review actual removal results and failures after the policy is turned on. See Learn about simulation mode for reasons simulation and enforcement results can differ.
- **A removal policy can remove a label that applies Microsoft Purview encryption.** The associated encryption is also removed. Password-protected files and files protected with external or non-Microsoft encryption aren't supported and might report `EncryptedFileNotSupported`. Simulation confirms that files match the policy conditions, but doesn't guarantee that enforcement can successfully modify them.
- **After cross-tenant migration, the source-tenant label often isn't recognized in the target tenant.** The label's tenant-scoped identity moves with the file's metadata but the target tenant's label set doesn't include it, so a target-tenant removal policy can't find the label to strip. Republish the label in the target tenant (or use PowerShell to strip metadata directly) before you run a removal policy.
- **A conflicting active policy that applies a label wins.** If another auto-labeling policy has conditions that match the same files and applies a label, the label-apply wins over the label-remove even when the policy applies a different label than the label selected for removal. The label removal policy might show these files as removed in simulation because simulation evaluates one policy at a time, but after all active policies run, matching files stay labeled instead of becoming unlabeled. See the Prerequisites — detailed reference note beginning "If you have an auto-labeling policy configured to remove a label" for the full conflict-resolution behavior.
- **Removing a label that applies encryption also removes the encryption.** This is intentional but frequently surprising to admins who expected the file to become "unlabeled but still protected." If you need to retain the existing label-applied encryption, don't remove that label with an auto-labeling policy.

To create a removal policy, use the **Automatically remove label only** tab in Creating an auto-labeling policy above.

Monitoring your auto-labeling policy

After your auto-labeling policy is turned on, you can view the labeling progress for files in your chosen SharePoint and OneDrive locations. Emails aren't included in the labeling progress because they're automatically labeled as they're sent.

> **Note**
>
> Exchange email isn't included in policy-level labeling progress, **Insights** enforcement metrics, or the **Labeled items** view. These views report on files in SharePoint and OneDrive.
>
> To verify labeling for an Exchange email, use Activity explorer and filter for the **Sensitivity label applied** activity. Narrow the results by date, sensitivity label, location, and item details, and review **How applied** to verify that the label was applied automatically. Allow 60 to 90 minutes for the activity to appear. Activity explorer identifies the resulting label and how it was applied, but doesn't identify the specific auto-labeling policy or rule that applied the label to an Exchange message.

The labeling progress includes the files to be labeled or unlabeled by the policy, the affected files for the last seven days, and the total files affected. Because of the maximum of labeling 100,000 files a day, this information provides you with visibility into the current labeling progress for your policy and how many files are still to be labeled.

When you first turn on your policy, you initially see a value of 0 for files to be labeled until the latest data is retrieved. This progress information updates every 48 hours, so you can expect to see the most current data about every other day. In the earlier policy flyout experience, selecting a policy shows labeling progress by the top 10 sites. The newer full **Overview** page shows **Top 15 sites** instead. The information on the flyout pane might be more current than the aggregated policy information displayed on the **Auto-labeling** main page. To have visibility on all auto-labeling policy progress, use activity explorer for the most up to date information for files and emails that were auto-labeled.

You can also see the results of your auto-labeling policy by using content explorer when you have the appropriate permissions:

- **Content Explorer List Viewer** role group lets you see a file's label but not the file's contents.
- **Content Explorer Content Viewer** role group, and **Information Protection** and **Information Protection Investigators** role groups let you see the file's contents.

Content explorer provides a current snapshot of labeled items and doesn't identify the auto-labeling policy that applied a label. Counts can take up to seven days to update, so use Activity

explorer instead for recent labeling activity.

However currently, restricted admins won't be able to see labeling activities for OneDrive in activity explorer.

> **Tip**
>
> You can also use content explorer to identify locations that have documents with sensitive information, but are unlabeled. Using this information, consider adding these locations to your auto-labeling policy, and include the identified sensitive information types as rules.

Policy-level labeling activity for SharePoint and OneDrive

For auto-labeling policies that are turned on and target SharePoint and OneDrive, you can monitor them right from the main **Auto-labeling** page and use the per-policy review pages. Use these pages to monitor day-to-day labeling activity, spot-check labeled files, and investigate labeling failures.

To view the review pages, you need one of the following roles:

- Compliance Administrator
- Compliance Data Administrator
- Information Protection Admin
- Information Protection Analyst

Contextual summary is available in supported portal experiences. To view matched text inside a file, you also need the **Data Classification Content Viewer** role.

To open the newer review pages, go to **Solutions** > **Information Protection** > **Policies** > **Auto-labeling policies**. Select a policy, then select **View details** in the policy details panel. Earlier portal experiences might open Activity explorer instead. Most data on the review pages reflects activity from the last 30 days.

The newer review page has separate **Overview**, **Labeled items**, and **Labeling failures** tabs. For policies that remove labels, the item tabs are **Label removed items** and **Label removal failures**.

The **Overview** tab shows labeled and failed file counts and **Top 15 sites**. Its **Labeling activity overview** chart has separate successful and failed series for the last 30 days. For removal policies, the chart is **Label removal activity overview**. The **Configuration details** section shows the policy's locations, exclusions, and rule summary.

The policy details panel's **Insights** tab summarizes the sensitive information types detected in labeled files. A **Labeling failures** card shows the total failure count and the top three failure reasons. For details, see Use the Insights tab to analyze auto-labeling policies.

Use **Labeled items** to review successfully labeled files. Use **Labeling failures**, or **Review items** on the **Labeling failures** card, to investigate failures. Select a file and review its failure reason. In experiences with a **Details** tab, open that tab to find **Failure reason**. File previews and contextual summaries depend on the portal experience and your permissions.

Review the failure reason even when few files fail. Transient infrastructure failures are retried automatically, but file format, protection, or label configuration failures might require action regardless of volume. Some unsupported files require no action. Follow the reason-specific guidance in Resolve auto-labeling failures in SharePoint and OneDrive files.

Review coverage by simulation context

In supported portal experiences, **Coverage by simulation context** (preview) reports on SharePoint and OneDrive files. Run the policy in simulation mode, then turn it on before reviewing coverage. The coverage report helps you see which files were labeled, pending, or not labeled, and whether those files were found in the latest simulation run.

View coverage can be hidden if the policy isn't turned on or simulation completion information is missing. If the entry is greyed out, point to it or move keyboard focus to it to read the tooltip. The tooltip distinguishes an expired simulation from an unsupported Exchange-only policy. Rerunning an Exchange-only policy in simulation doesn't enable file coverage.

Use this report to:

- Review how labeling activity is distributed after the latest simulation run.
- Identify files that were found in simulation and then labeled during enforcement.
- Review files that are pending, not yet attempted, or failed to apply a label.
- Investigate files that don't have a match in the latest simulation run.

Coverage is based on labeling activity after the latest simulation run and is available for 30 days. After 30 days, **View coverage** is unavailable for that simulation run. To view coverage again, run the policy in simulation mode again and then review the updated coverage report.

The coverage report isn't an all-time policy total and isn't intended to be an exact reconciliation between simulation and enforcement. All-time labeled totals can include activity from before or after the coverage window, so they might not match the counts shown in the coverage report.

To open the coverage report:

1. In the Microsoft Purview portal, go to **Information Protection** > **Policies** > **Auto-labeling policies**.
2. Select an auto-labeling policy that includes SharePoint or OneDrive, has completed simulation, and is turned on.
3. In the **Coverage** column, select **View coverage**.

The report groups files by labeling outcome and simulation context.

Expand table

Area	What it means
Observed files	Files included in this report because they were found during simulation, enforcement, or both.
Labeled during enforcement	Files that had a label applied during enforcement.
Not labeled during enforcement	Files that don't have a successful label application in this report.
Found in simulation	Files that were found in the latest simulation run.
No simulation match found	Files that weren't matched to the latest simulation run.
Pending	Files that don't have a final enforcement outcome yet.
Not yet attempted	Files that enforcement hasn't attempted to label yet.
Failed to apply label	Files that enforcement attempted to label, but the attempt failed.

Select a bucket in the chart to open an item list for that bucket. Depending on the bucket, item details can include the file name, location, workload, label, last activity time, status, and failure reason when available.

Files under **No simulation match found** can be files that were created, changed, or classified after the latest simulation run completed. They can also appear when the report doesn't have enough information to identify a more specific reason. Files under **Failed to apply label** were attempted during enforcement, but the label wasn't applied. Open the item list and review the failure reason when available.

Why was — or wasn't — a specific file labeled?

For files in SharePoint and OneDrive, first check what the policy reports. Select the policy, then **View details** in the policy details panel. In the newer experience, use **Labeled items** to review successful labeling or **Labeling failures** to investigate a failure. Earlier portal experiences might open Activity explorer instead.

A simulation match doesn't confirm that a label was applied. A file missing from **Labeling failures** isn't proof of success either. Choose the guidance that matches what you see:

- **The policy is on, but the file isn't labeled.** Check the policy's locations and conditions, then the file's eligibility. Follow Policy is on but files aren't labeled.
- **The file has an unexpected label or protection.** Review the policy's conditions and exceptions, label replacement rules, and other matching policies before changing anything. Follow Existing label or protection isn't what you expect.
- **The reports show different numbers.** Check which time period and locations each report covers. SharePoint and OneDrive file views don't include Exchange email. Follow Numbers don't agree or Exchange activity is missing.

If the file reports a failure, review **Failure reason** (on the **Details** tab, if present). For reason-specific actions, see Resolve auto-labeling failures in SharePoint and OneDrive files.

A new or changed policy can evaluate existing files using available classification. Changing a custom sensitive information type doesn't by itself reclassify unchanged files. A file's last-modified date doesn't establish its classification time. For classification refresh and policy evaluation checks, see Auto-labeling classification and simulation concepts.

If your rule uses **Document property is** with a mapped SharePoint managed property, check whether the value contains an ampersand (`&`). See the mapped-property limitation in Prerequisites — detailed reference.

Monitoring your daily labeling volume

Use the main **Auto-labeling** page to review labeling activity across your policies. Depending on your portal experience, you might see **Files Labeled Today** or the newer **Files labeled in the last 30 days** chart. The chart shows daily counts and an average calculated over all 30 days.

These views count label-applied and label-changed activity for SharePoint and OneDrive files associated with your auto-labeling policies. They don't include Exchange email. Use them to spot activity trends, not to measure remaining daily capacity or the number of unique files currently labeled or protected.

Policy rule fails to load

In rare cases, a malformed rule in an auto-labeling policy can prevent the policy from loading correctly. The most common symptom is that Exchange messages stop matching an otherwise enabled auto-labeling policy, and there might not be any item-level labeling failure to review because the rule never loaded in the first place.

This is different from the SharePoint and OneDrive item-level labeling failures described earlier in this section. Those failures mean a specific file was evaluated against the policy but couldn't be labeled (for example, because of file format or existing protection). A rule-load failure means the policy's rule couldn't be compiled or loaded at all, so *no* content is evaluated against that rule until the problem is fixed.

If you suspect a policy has stopped matching content and there's no corresponding item-level failure, contact Microsoft Support. Support can check whether a rule failed to load and, if so, identify which specific rule is affected.

> **Important**
>
> Only delete and recreate a rule after Microsoft Support has identified the specific rule that failed to load. Don't delete rules from a policy as a general troubleshooting step, because this removes their enforcement.

If Support confirms that a specific rule needs to be recreated, use the following sequence:

1. Record the current configuration of the affected rule before you delete it, including the policy name, rule name, target label, workload/scope, conditions, exceptions, priority, and enabled state. You can use Get-AutoSensitivityLabelRule to retrieve and save these details.
2. Delete only the rule that Microsoft Support identified, using Remove-AutoSensitivityLabelRule. Don't delete other rules in the policy.
3. Recreate the rule with the same configuration you recorded, using New-AutoSensitivityLabelRule.
4. Run the policy in simulation mode to confirm the recreated rule matches content as expected.
5. For Exchange, validate the rule by using newly sent or received messages, because auto-labeling for Exchange doesn't apply to messages already at rest in a mailbox.
6. Turn on the policy and confirm that matching activity resumes as expected.

> ⚠ **Warning**
>
> After you delete the affected rule, enforcement from that rule is paused until it's recreated and the change propagates. Plan for a change window and expect a delay before the recreated rule is fully in effect.

Exchange auto-labeling rule limits

When you configure an Exchange auto-labeling rule in the Microsoft Purview portal, the **Sender domain is** condition accepts up to **600 domain entries**. If you exceed this limit, saving the policy fails with an error that identifies the maximum of 600.

This limit describes the domain entries in the condition. It doesn't establish a combined domain limit across all rules in a policy.

For Exchange auto-labeling policies, check the parameter-specific limits in New-AutoSensitivityLabelRule and Set-AutoSensitivityLabelRule. These limits are separate from the number of policies or locations that you can configure.

For example, the FromAddressMatchesPatterns parameter documentation specifies a maximum of 10 regular expressions, with up to 128 characters per expression. These figures describe the values for that parameter, not a combined limit for all conditions in a policy.

Support for multiple values doesn't mean an unlimited number of values. If the parameter reference doesn't specify a maximum for your configuration, contact Microsoft Support for guidance.

Use PowerShell for auto-labeling policies

You can use Security & Compliance PowerShell to create and configure auto-labeling policies. This means you can fully script the creation and maintenance of your auto-labeling policies, which also provides a more efficient method of specifying multiple locations for SharePoint and OneDrive.

To target more than 1,000 SharePoint sites, associate an existing SharePoint adaptive scope with the policy by using the `-SharePointAdaptiveScopes` parameter (or `-SharePointAdaptiveScopesException` for exclusions) with New-AutoSensitivityLabelPolicy. An adaptive scope supports up to 50,000 SharePoint sites per policy. This doesn't increase the portal limit; instead, it replaces static site enumeration with dynamic adaptive-scope membership.

Before rerunning simulation, preserve simulation results you need for comparison.

Before you run the commands in PowerShell, you must first connect to Security & Compliance PowerShell. To use PowerShell to start or rerun an auto-labeling policy simulation, connect by using Connect-IPPSSession with the *EnableSearchOnlySession* switch:

PowerShell

```
Connect-IPPSSession -EnableSearchOnlySession -UserPrincipalName <admin-UPN>
```

The *EnableSearchOnlySession* switch is required so that the simulation job can be submitted correctly. If a simulation started from PowerShell remains in progress, reconnect in a new PowerShell session with this switch and rerun the simulation, or use the Microsoft Purview portal.

To create a new auto-labeling policy:

PowerShell

```
New-AutoSensitivityLabelPolicy -Name <AutoLabelingPolicyName> -SharePointLocation "
<SharePointSiteLocation>" -ApplySensitivityLabel <Label> -Mode
TestWithoutNotifications
```

This command creates an auto-labeling policy for a SharePoint site that you specify. For a OneDrive location, use the *OneDriveLocation* parameter, instead.

To add more sites to an existing auto-labeling policy:

PowerShell

```
$spoLocations = @("<SharePointSiteLocation1>","<SharePointSiteLocation2>")
Set-AutoSensitivityLabelPolicy -Identity <AutoLabelingPolicyName> -
AddSharePointLocation $spoLocations -ApplySensitivityLabel <Label> -Mode
TestWithoutNotifications
```

This command specifies the new SharePoint URLs in a variable that is then added to an existing auto-labeling policy. To add OneDrive locations instead, use the *AddOneDriveLocation* parameter with a different variable, such as *$OneDriveLocations*.

To create a new auto-labeling policy rule:

PowerShell

```
New-AutoSensitivityLabelRule -Policy <AutoLabelingPolicyName> -Name
<AutoLabelingRuleName> -ContentContainsSensitiveInformation @{"name"= "00aa00aa-bb11-
```

```
cc22-dd33-44ee44ee44ee"; "mincount" = "2"} -Workload SharePoint
```

For an existing auto-labeling policy, this command creates a new policy rule to detect the sensitive information type of **U.S. social security number (SSN)**, which has an entity ID of 00aa00aa-bb11-cc22-dd33-44ee44ee44ee. To find the entity IDs for other sensitive information types, refer to Sensitive information type entity definitions.

For more information about the PowerShell cmdlets that support auto-labeling policies, their available parameters and some examples, see the following cmdlet help:

- Get-AutoSensitivityLabelPolicy
- New-AutoSensitivityLabelPolicy
- New-AutoSensitivityLabelRule
- Remove-AutoSensitivityLabelPolicy
- Remove-AutoSensitivityLabelRule
- Set-AutoSensitivityLabelPolicy
- Set-AutoSensitivityLabelRule

Tips to increase labeling reach

Although auto-labeling is one of the most efficient ways to classify, label, and protect Office and PDF files that your organization owns, check whether you can supplement it with any of the following methods to increase your labeling reach:

- For SharePoint document libraries, you can apply a default sensitivity label for new and edited files. To extend the same default label to existing files, use an auto-labeling policy that applies default sensitivity labels from SharePoint document libraries to existing files. For more information, see Configure a default sensitivity label for a SharePoint document library.

- With Microsoft Syntex, you can apply a sensitivity label to a document understanding model, so that identified documents in a SharePoint document library are automatically labeled.

- For Outlook messages, you can apply a sensitivity label based on attachments that are labeled.

- When you use the Microsoft Purview Information Protection client:
  - For files in on-premises data stores, such as network shares and SharePoint Server libraries: Use the scanner to discover sensitive information in these files and label them

appropriately. If you're planning to migrate or upload these files to SharePoint in Microsoft 365, use the scanner to label the files before you move them to the cloud.

- Encourage manual labeling after providing users with training which sensitivity labels to apply. When you're confident that users understand which label to apply, consider configuring a default label and mandatory labeling as policy settings.

Additionally, consider marking new files as sensitive by default in SharePoint to prevent guests from accessing newly added files until at least one DLP policy scans the content of the file.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 10/14/2026

Microsoft Purview auto-labeling classification and simulation concepts

In Microsoft Purview, service-side auto-labeling is a policy-based capability that applies or removes sensitivity labels. For existing files in SharePoint and OneDrive, classification, policy evaluation, and label enforcement are distinct stages. These distinctions help you interpret unexpected matches and decide whether to change a policy, update classification, or investigate a labeling failure.

This article focuses on policies that use sensitive information types (SITs). A new or changed policy can evaluate existing files using available classification results. Changing a custom SIT or restarting simulation doesn't, by itself, refresh classification for unchanged historical files.

Classification, policy evaluation, and enforcement

Each stage answers a different question. A policy match isn't evidence that the content was freshly classified or that a label was successfully applied.

Expand table

Stage	What it does	What the result means
Classification	Identifies sensitive content and records classification results, such as the SIT, match count, and confidence.	The results describe what classification detected when it processed the file.
Policy evaluation and simulation	Evaluates the policy's scope and conditions using available classification results for SIT conditions. Simulation reports matches without applying or removing labels.	The file matched the evaluated policy. The match doesn't prove that classification used the latest SIT definition.
Enforcement	Applies or removes a label according to the active policies, file eligibility, and applicable label rules.	The actual labeling outcome must be checked separately from the simulation match.

Policy changes aren't classifier changes

Changing the conditions that a policy evaluates isn't the same as changing how a SIT detects sensitive content.

Expand table

Change or action	Effect on existing SharePoint and OneDrive files
Create or change an auto-labeling policy or rule	Can evaluate existing in-scope files using available classification results. Existing, correct classifications don't need to be refreshed for every policy change.
Change a custom SIT definition	Doesn't by itself reclassify unchanged historical files. Available results can still reflect the earlier definition.
Restart auto-labeling simulation	Reevaluates policy matches. It doesn't itself reclassify the underlying content.
Run on-demand classification	Classifies files within the scan scope. If you select specific classifiers, only those classifiers have their results updated for the scanned files.

Creating a policy that references a new SIT isn't proof that all historical files have been classified with that SIT. Review classification coverage rather than assuming that either the policy creation date or the SIT creation date establishes coverage.

Why an older file can appear in simulation

Supported files can be eligible for auto-labeling whether they existed before or after the policy was created. An older file can match a new policy because classification results are already available for it. The file's last-modified date alone doesn't establish when it was last classified.

For example, a file already has classification results for a SIT. You create a policy that uses that SIT, or change the policy's required match count. The file can match the policy without being edited or freshly classified.

In a different example, you narrow a custom SIT definition after a file was classified. If the file's classification hasn't been refreshed, simulation can still match using the earlier results. This is one possible explanation for a discrepancy, not a diagnosis for every unexpected match.

Uploaded-file tests and policy simulation answer different questions

The SIT **Test** function evaluates an uploaded file against the SIT. Auto-labeling simulation evaluates the policy's scope and conditions using available classification results for existing SharePoint and OneDrive files. A successful uploaded-file test doesn't establish that historical files have the same classification results.

Use Test sensitive information types with representative positive and negative examples to validate your custom detector. Your organization is responsible for deciding whether a custom SIT meets its detection requirements and obligations. A successful test or simulation doesn't certify regulatory coverage.

When a direct test and simulation disagree, consider the SIT definition used for classification as well as the policy's conditions and scope. Don't assume that restarting simulation refreshes classification.

When historical files need updated classification

Use on-demand classification when relevant historical files haven't been classified or need updated classification, such as after a custom SIT change. Scope the scan to the relevant locations, file date range, file types, and classifiers. Check that the date range includes the historical files you intend to evaluate. Review estimates and charges before starting classification.

> **Important**
>
> On-demand classification isn't an auto-labeling simulation. Updated classification can trigger actions from other active matching policies, including labeling, label removal, and other protection or lifecycle actions. Review active policies before starting the scan, even if the auto-labeling policy you're testing is still in simulation.

After classification completes, review auto-labeling simulation results before deciding whether to turn on the policy you're testing. Refreshing classification doesn't itself undo existing labels; label changes depend on the applicable policies or an authorized labeling action.

Simulation doesn't guarantee enforcement

Simulation evaluates one policy. When policies are active, competing policies and label replacement rules can affect the outcome. File eligibility, protection, and labeling failures can also prevent a matching file from being labeled.

Use policy-level labeling activity to review actual outcomes and failures. Coverage by simulation context helps compare those outcomes with the latest simulation; it isn't an exact reconciliation or an all-time policy total.

These distinctions don't explain a stalled simulation, a source view that fails to open, or results that remain inconsistent after the relevant checks. Use the auto-labeling troubleshooting

guidance and contact Microsoft Support if the issue persists.

Related content

- Configure auto-labeling policies
- Create and analyze an on-demand classification scan
- Resolve auto-labeling failures in SharePoint and OneDrive files

ⓘ**Note:** This article contains content created with AI. Learn more

---

Last updated on 10/14/2026

Review simulation results for auto-labeling policies in Microsoft Purview

Before you enforce an auto-labeling policy, run it in simulation mode. Simulation reports policy matches without applying or removing labels. Review the match results and source distribution to decide whether the policy is ready for enforcement.

This article describes the newer simulation overview page. Available fields and review options can vary by portal experience. To review policy results after you turn on a policy, see Monitor your auto-labeling policy.

For existing SharePoint and OneDrive files, simulation uses available classification results. Restarting simulation reevaluates policy matches; it doesn't itself reclassify the content after a SIT change. See Auto-labeling classification and simulation concepts to distinguish direct SIT tests, classification refresh, and enforcement.

> **Important**
>
> Before you begin, make sure that you have the required roles and permissions. For more information, see Policy-level labeling activity for SharePoint and OneDrive.
>
> To view matched text in a file on **Items to review**, you also need the **Data Classification Content Viewer** role.

Review the simulation overview

Make sure the policy is running in simulation mode and has started evaluating the configured locations. Then follow these steps:

1. Sign in to the Microsoft Purview portal.
2. Select **Solutions** > **Information Protection** > **Policies** > **Auto-labeling policies**.
3. Select the name of a policy that is running in simulation mode.
4. In the policy details panel, select **Summary** > **View details** to open the **Simulation overview** page.

On the **Simulation overview** page, the following key metrics about the policy are available:

Expand table

Field	Description
Label	The sensitivity label that the policy applies to matched items.
Status	Simulation status: **Simulation in progress**, **Simulation almost complete**, **Simulation complete**, **Simulation expiring soon**, **Expired**, or **Policy can't be published**.
Duration	The length of time the simulation has been running in days and hours.
Matched items	The total number of items that matched the policy conditions across all workloads.
Sample items	When displayed, the number of sampled items available for review. Samples are retained for 30 days. See Preserve results before editing or rerunning.
Policy matches per rule	Shows details about the policy rules and their matches: - **Total Rules:** the number of rules in the policy, not the number of rules with matches. - **Total Matches:** the number of items that matched rules. - **Rule name:** the name of the rule in the policy. - **Matches:** the number of items that matched this rule. - **Location:** The workload where matches occurred (Exchange, SharePoint, or OneDrive). Use the per-rule match counts to identify which rules matched and whether each rule produces the expected results.
Matches by source	A donut chart showing matches distributed across workloads, with an item count and percentage for each of the following workloads: - **Exchange**: Email messages that matched the policy. Matched Exchange items are based on sampled data and might not represent all mailboxes in the policy scope. - **SharePoint**: Files in SharePoint sites that matched. - **OneDrive**: Files in OneDrive accounts that matched.
Sensitive info types	This table lists the sensitive information types (SITs) detected across the matched items and the number of matches for each type. Use it to identify which SITs produce the most matches.
Configuration details	Shows the policy's locations, exclusions, and rule configuration. Review these details alongside the match results to check the policy scope.
Review sample items	Select **Review sample items**, when displayed, to open **Items to review**. For each item, you can: - Preview the file or email content, where supported and with the required permissions. - View properties such as the file location, modified date, and detected SITs. - Read a **Contextual summary**, in supported portal experiences, that explains why the item matched. Viewing matched content requires the **Data Classification Content Viewer** role.

> **Note**
>
> The **Matched items** count is an **estimate**, not a count of successfully labeled files. Enforcement results can differ because files or classification results change, other active policies affect the outcome, or labeling fails. For details, see Simulation doesn't guarantee enforcement.

Preserve results before editing or rerunning

Before you edit an auto-labeling policy or restart simulation, export any available results that you need for later comparison. Check the exported file before you make changes.

Don't assume that sample retention guarantees access to results from earlier runs in the policy view or export. Export-size limits are separate from prior-run availability. Don't rely on filters or a larger export limit to recover results that are no longer available.

Next steps

Based on the simulation results, choose the appropriate next step:

- **Turn on policy** when the policy is eligible and you've reviewed the results and warnings. **Simulation complete**, **Simulation almost complete**, and **Simulation expiring soon** can be eligible states, but eligibility doesn't replace checking that the matches meet your requirements.
- **Edit the policy** when its scope or conditions don't meet your requirements. First, preserve results you need for comparison. Review sample items to distinguish false positives from valid matches, edit the conditions as needed, and restart simulation to reevaluate policy matches.
- **Refresh classification when needed** if historical files haven't been classified or their classification needs to reflect a changed SIT. Review on-demand classification scope, costs, and active-policy effects before starting a scan. A simulation restart alone doesn't refresh classification.

> **Important**

If you configure a scheduled turn-on date, the policy is scheduled to turn on automatically without waiting for your review. The displayed date isn't confirmation that activation succeeded. Cancel the scheduled turn-on date if you need more time.

Related articles

- Automatically apply a sensitivity label to Microsoft 365 data
- Monitor your auto-labeling policy
- Resolve auto-labeling failures in SharePoint and OneDrive files
- Learn about sensitivity labels

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 10/15/2026

Use the Insights tab to analyze auto-labeling policies in Microsoft Purview

The **Insights** tab provides an at-a-glance view of an auto-labeling policy's performance directly in the policy details panel. Its contents vary depending on whether the policy is running in simulation or enforcement mode, helping you understand how the policy identifies or labels content.

This article describes the newer Insights layout. Available fields and review options can vary by portal experience.

Availability also depends on the policy's mode and locations:

- **Simulation:** The tab is available for policies that include Exchange, SharePoint, or OneDrive.
- **Enforcement:** The tab is available only when the policy includes SharePoint or OneDrive.

The tab isn't available for disabled policies or Exchange-only policies that are turned on. For an Exchange-only policy that's turned on, use Activity explorer to review email labeling activity.

For more information on simulation, see Review simulation results for auto-labeling policies. For more information on enforcement, see Monitor your auto-labeling policy.

> **Important**
>
> Before you begin, make sure that you have the required roles and permissions. For more information, see **Policy-level labeling activity for SharePoint and OneDrive**.

To open the Insights tab:

1. Sign in to the Microsoft Purview portal.
2. Select **Solutions** > **Information Protection** > **Policies** > **Auto-labeling policies**.
3. Select a policy that is running in simulation mode or is turned on and includes SharePoint or OneDrive.
4. In the policy details panel, select **Insights**.
5. (Optional) Select **View details** to open the full policy overview. For policies in enforcement mode, earlier portal experiences might open Activity explorer instead.

Simulation mode: When the policy runs in simulation mode, the Insights tab shows a lightweight summary of the simulation findings. These findings include:

- General information about the policy, including **Label**, **Status**, **Duration**, **Results** (or total matches), and **Sample items**, when displayed.
- **Turn on date**, which shows scheduled automatic activation when configured; otherwise, no date is shown. This date isn't confirmation that activation succeeded. The separate **Start date/time** field shows the displayed simulation start date and time.
- **Rules and matches**, which shows the total number of rules in the policy and the total matches found. The rule total doesn't count only rules with matches. To identify which rules matched, select **View details** and review the per-rule match counts on the full simulation overview.
- **Matches by source**, which is a donut chart of matches across Exchange (purple), SharePoint (green), and OneDrive (blue).
- **Matches by sensitive info type**, which is a table listing each detected sensitive information type (SIT) and its match count.

> ⓘ Note
>
> Exchange matched-item counts are estimates based on sampled data.

Enforcement mode: When the policy is actively labeling files, the Insights tab shows 30-day operational metrics. **Insights** shows the following metrics about the policy:

> ⓘ Note
>
> Policy-level enforcement metrics and the **Labeled items** view report only files in SharePoint and OneDrive. They don't include Exchange email, including voicemail messages stored in Exchange. To verify labeling for an Exchange email, use Activity explorer and allow 60 to 90 minutes for the activity to appear.

Expand table

Field	Description
Sensitivity label	The sensitivity label applied by the policy.
Status	The current status of the policy.

Field	Description
Turn on date	The activation date and time displayed by the portal. Don't use this field as an immutable record of the policy's first activation.
Total labeled to date	The total number of SharePoint and OneDrive files labeled by the policy.
Duration	The number of days calculated from the displayed enforcement turn-on date.
Last simulated results	The results of the most recent simulation before enforcement was turned on.
Completed on	The date when the most recent simulation completed.
Count of auto-labeled and failed files	A donut chart showing files processed in the last 30 days: - **Labeled (green):** Files that were successfully labeled. - **Failed (red):** Files that matched but couldn't be labeled.
Labeling failures	The card shows the total number of failed items and the **Top 3 failure reasons**, sorted by volume. For more information, see Resolve auto-labeling failures in SharePoint and OneDrive files.
Top 10 sensitive info types	The table lists the sensitive information types (SITs) most frequently detected when this policy applied labels during the last 30 days. It includes the **Files labeled** count for each SIT.
Labels per source	To identify whether failures are concentrated to a specific workload, **Labels per source** shows a breakdown of labeled and failed item counts for SharePoint and OneDrive.

For the daily trend, select **View details** and open **Overview** in the newer full review experience. **Labeling activity overview** shows separate successful and failed series for the last 30 days. For removal policies, the chart is **Label removal activity overview**. This chart isn't part of the newer Insights layout.

Review failure reasons even when few files fail. Transient infrastructure failures are retried automatically, but configuration, protection, or file format failures might require action regardless of volume. Some unsupported files require no action. Follow the reason-specific guidance in Resolve auto-labeling failures in SharePoint and OneDrive files.

> ⓘ **Note**
>
> The Insights tab provides less data for non-Microsoft locations.

Related articles

- Review simulation results for auto-labeling policies
- Monitor your auto-labeling policy
- Resolve auto-labeling failures in SharePoint and OneDrive files
- Automatically apply a sensitivity label to Microsoft 365 data

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 10/12/2026

Use sensitivity labels to protect collaborative workspaces (groups and sites)

> *Microsoft Purview service description*

In addition to using sensitivity labels to protect items such as documents and emails, you can also use sensitivity labels to protect content in the following containers: Microsoft Teams sites, Microsoft 365 groups, SharePoint sites, Viva Engage communities, and Loop workspaces. To protect these collaborative workspaces, use the following label settings:

- Privacy (public or private)
- External user access
- External sharing from SharePoint sites
- Access from unmanaged devices
- Authentication contexts
- Prevent discovery of private teams for users who have this capability
- Shared channels control for team invitations
- Default sharing link for a SharePoint site (PowerShell-only configuration)
- Site sharing settings (PowerShell-only configuration)
- Default label for channel meetings

> **Important**
>
> The settings for unmanaged devices and authentication contexts work in conjunction with Microsoft Entra Conditional Access. You must configure this dependent feature if you want to use a sensitivity label for these settings. Additional information is included in the instructions that follow.

Support for container-level protection by using sensitivity labels is growing all the time. Refer to the documentation for each collaborative workspace and the Microsoft public roadmap to learn about new capabilities.

When you apply a sensitivity label to a supported container, the label automatically applies the sensitivity category and configured protection settings to the workspace.

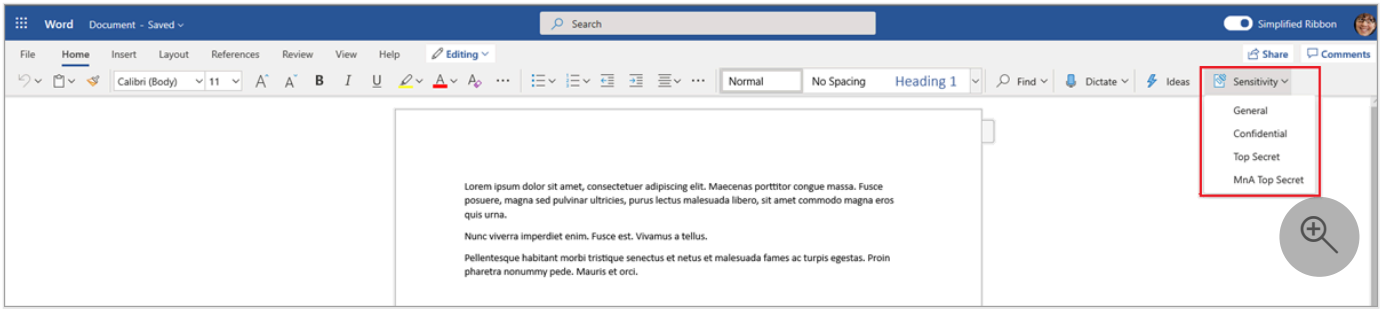
Be aware that some label options can extend configuration settings to site owners that are otherwise restricted to administrators. When you configure and publish the label settings for external sharing options and the authentication context, a site owner can now set and change these options for a site by applying or changing the sensitivity label for a team or site. Don't configure these specific label settings if you don't want site owners to be able to make these changes.

Items in these containers however, do not inherit the labels and therefore don't apply any item-level label settings, such as content markings and encryption. So that users can label their documents in SharePoint sites or team sites, make sure you've enabled sensitivity labels for Office files in SharePoint and OneDrive.

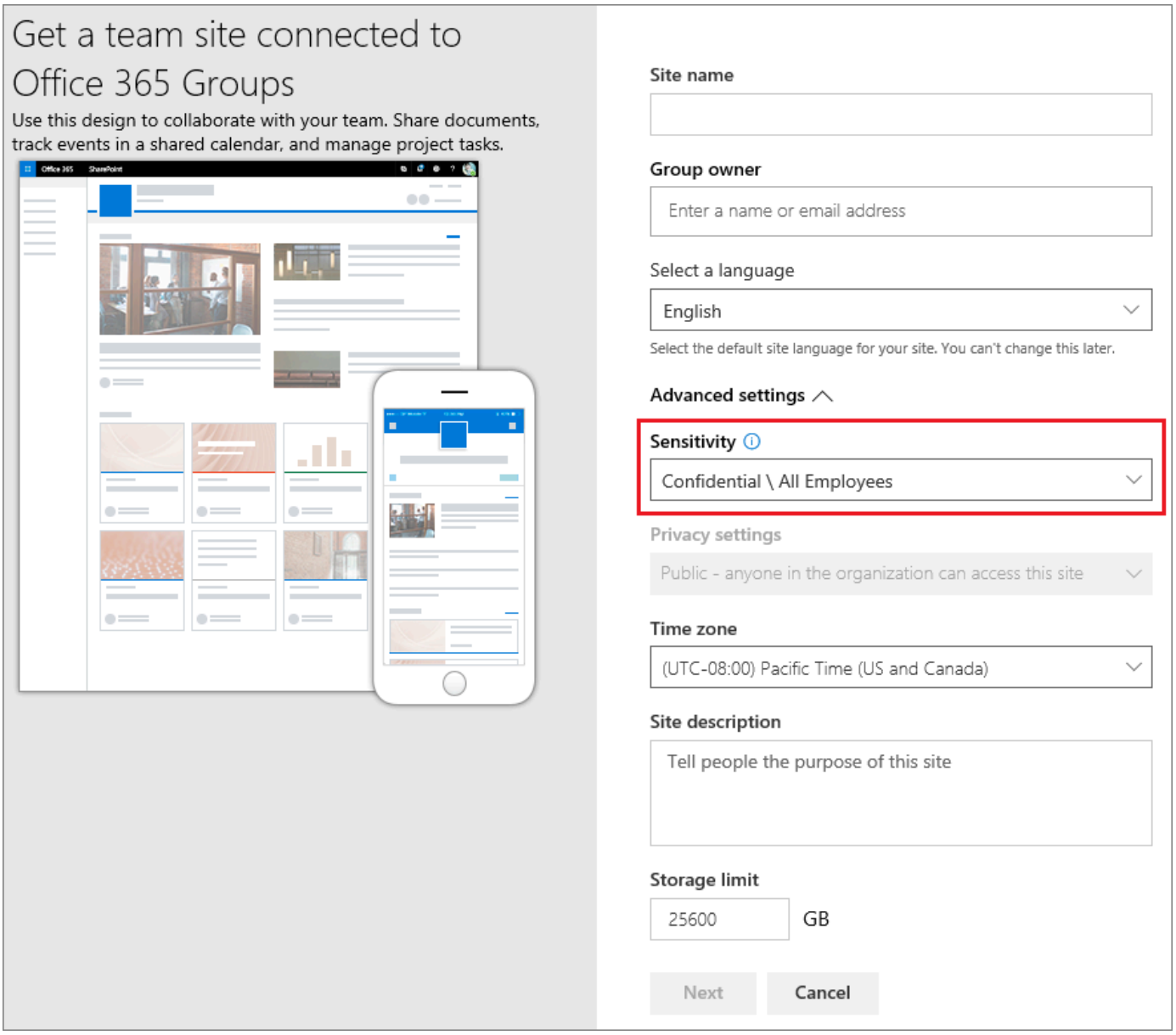
Container labels don't support displaying other languages and display the original language only for the label name and description.

Using sensitivity labels for containers

Before you enable and configure sensitivity labels for containers, users can see and apply sensitivity labels to items in their apps. For example, from Word:



After you enable and configure sensitivity labels for containers, users can additionally see and apply sensitivity labels to containers that support them. These include Microsoft team sites, Microsoft 365 groups, SharePoint sites, Viva Engage communities, and Loop workspaces. For example, when you create a new team site from SharePoint:



After a sensitivity label has been applied to a site, you must have the following role to change that label in SharePoint or Teams:

- For a group-connect site: Microsoft 365 group Owners
- For a site that isn't group-connected: SharePoint site admin

> **Note**
>
> Sensitivity labels for containers support Teams shared channels. If a team has any shared channels, they automatically inherit sensitivity label settings from their parent team, and that label can't be removed or replaced with a different label.

How to enable sensitivity labels for containers and synchronize labels

If you haven't yet enabled sensitivity labels for containers, do the following set of steps as a one-time procedure:

1. Because this feature uses Microsoft Entra functionality, follow the instructions from the Microsoft Entra documentation to enable sensitivity label support: Assign sensitivity labels to Microsoft 365 groups in Microsoft Entra ID.

2. You now need to synchronize your sensitivity labels to Microsoft Entra ID. First, connect to Security & Compliance PowerShell.

   For example, in a PowerShell session that you run as administrator, sign in with a global administrator account.

3. Then run the following command to ensure your sensitivity labels can be used with Microsoft 365 groups:

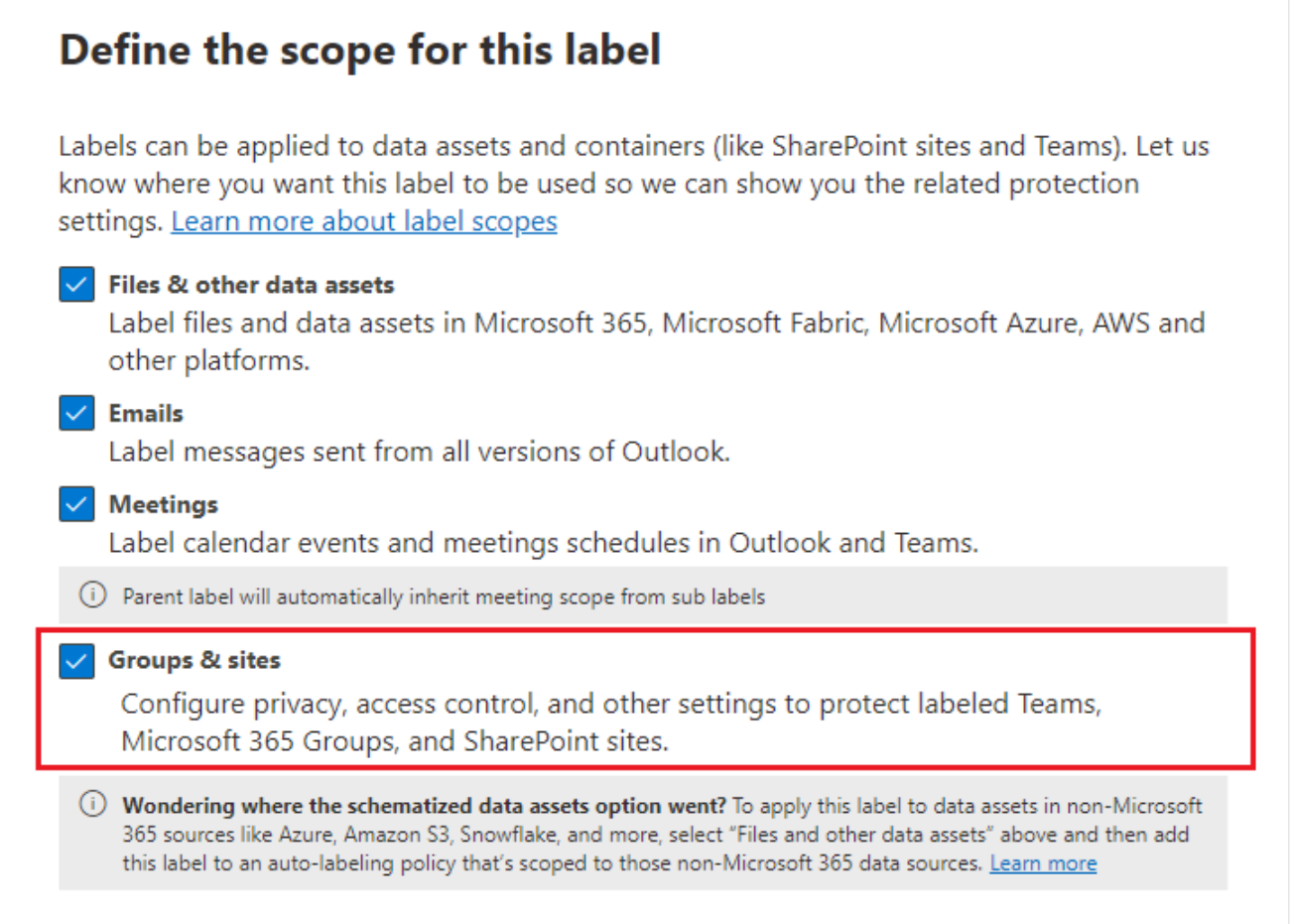
   PowerShell

   ```
   Execute-AzureAdLabelSync
   ```

How to configure groups and site settings for containers

After sensitivity labels are enabled for containers as described in the previous section, you can then configure protection settings for groups and sites in the sensitivity labeling configuration. Until sensitivity labels are enabled for containers, the settings are visible but you can't configure them.

1. Follow the general instructions to create or edit a sensitivity label and make sure you select **Groups & sites** for the label's scope:



When only this scope is selected for the label, the label isn't displayed in Office apps that support sensitivity labels and can't be applied to files and emails. Having this separation of labels can be helpful for both users and administrators, but can also add to the complexity of your label deployment.

For example, you need to carefully review your label ordering because SharePoint detects when a labeled document is uploaded to a labeled site. In this scenario, an audit event and email are automatically generated when the document has a higher priority sensitivity label than the site's label. For more information, see the Auditing sensitivity label activities section on this page.

2. Then, on the **Define protection settings for groups and sites** page, select the options you want to configure:

   - **Privacy and external user access settings** to configure the **Privacy** and **External users access** settings.
   - **External sharing and Conditional Access settings** to configure the **Control external sharing from labeled SharePoint sites** and **Use Microsoft Entra Conditional Access to protect labeled SharePoint sites** setting.
   - **Private teams discoverability and shared channel controls** to configure the settings to prevent users who can discover private teams from finding a private team with this

sensitivity label applied, and channel sharing controls for invitations to other teams.

Apply a label to channel meetings: This additional option is applicable only if you are editing an existing label where the scope includes meetings, and you've configured labels to protect meetings. Select a sensitivity label to automatically apply to channel meetings and all channel chats. For non-channel meetings, you can select a default label as a policy setting.

3. If you selected **Privacy and external user access settings**, now configure the following settings:

   - **Privacy**: Keep the default of **Public** if you want anyone in your organization to access the container where this label is applied.

     Select **Private** if you want access to be restricted to only approved members in your organization.

     Select **None** when you want to protect content in the container by using the sensitivity label, but still let users configure the privacy setting themselves.

     The settings of **Public** or **Private** set and lock the privacy setting when you apply this label to the container. Your chosen setting replaces any previous privacy setting that might be configured for the container, and locks the privacy value so it can be changed only by first removing the sensitivity label from the container. After you remove the sensitivity label, the privacy setting from the label remains and users can now change it again.

   - **External user access**: Control whether the owner can add guests to the container, similar to Manage guest access in Microsoft 365 groups.

4. If you selected **External sharing and Conditional Access settings**, now configure the following settings:

   - **Control external sharing from labeled SharePoint sites**: Select this option to then select either external sharing for anyone, new and existing guests, existing guests, or only people in your organization. For more information about this configuration and settings, see the SharePoint documentation, Turn external sharing on or off for a site.

   - **Use Microsoft Entra Conditional Access to protect labeled SharePoint sites**: Select this option only if your organization has configured and is using Microsoft Entra Conditional Access. Then, select one of the following settings:

- **Determine whether users can access SharePoint sites from unmanaged devices**: This option uses the SharePoint feature that uses Microsoft Entra Conditional Access to block or limit access to SharePoint and OneDrive content from unmanaged devices. For more information, see Control access from unmanaged devices from the SharePoint documentation. The option you specify for this label setting is the equivalent of running a PowerShell command for a site, as described in steps 3-5 from the Block or limit access to a specific SharePoint site or OneDrive section from the SharePoint instructions.

  For additional configuration information, see More information about the dependencies for the unmanaged devices option at the end of this section.

- **Choose an existing authentication context**: This option lets you enforce more stringent access conditions when users access SharePoint sites that have this label applied. These conditions are enforced when you select an existing authentication context that has been created and published for your organization's Conditional Access deployment. If users don't meet the configured conditions or if they use apps that don't support authentication contexts, they are denied access.

  For additional configuration information, see More information about the dependencies for the authentication context option at the end of this section.

  Examples for this label configuration:

  - You choose an authentication context that is configured to require multifactor authentication (MFA). This label is then applied to a SharePoint site that contains highly confidential items. As a result, when users from an untrusted network attempt to access a document in this site, they see the MFA prompt that they must complete before they can access the document.

  - You choose an authentication context that is configured for terms-of-use (ToU) policies. This label is then applied to a SharePoint site that contains items that require a terms-of-use acceptance for legal or compliance reasons. As a result, when users attempt to access a document in this site, they see a terms-of-use document that they must accept before they can access the original document.

5. If you selected **Private teams discoverability and shared channel controls**:

- For **Private teams discoverability**, use the **Allow users to discover private teams that have this label applied** checkbox when you've configured a Teams policy that allows

private teams discovery:

- When the checkbox is selected (the default setting), a private team with the sensitivity label applied will be discoverable for a user who is allowed to discover private teams.
- When the checkbox is cleared, a private team with the sensitivity label applied will remain hidden and won't be discoverable for all users.

- For **Teams shared channels**, when a team has a sensitivity label applied, you can allow or prevent other teams from being invited to the original team's shared channels. For more information about shared channels, see Shared channels in Microsoft Teams.

  > **Important**
  >
  > These options for Teams shared channels have a dependency on the settings on the previous **Privacy and external user access** page. If you select an option that's not compatible with these previous settings, you see a validation message to change your selection. Alternatively, you can go back in the configuration to change the dependent setting.

  Options include **Internal only**, **Same label only**, and **Private team only**. Only the last option can potentially remove previously invited teams, and none of the options affect invitations to individual users.

The site and group settings take effect when you apply the label to a container that supports sensitivity labels. If the label's scope includes files and emails, other label settings such as encryption and content marking aren't applied to the items within the container.

If your sensitivity label isn't already published, now publish it by adding it to a sensitivity label policy. The users who are assigned a sensitivity label policy that includes this label will be able to select it for sites and groups.

More information about the dependencies for the unmanaged devices option

If you don't configure the dependent conditional access policy for SharePoint as documented in Use app-enforced restrictions, the option you specify here will have no effect. Additionally, it will have no effect if it's less restrictive than a configured setting at the tenant level. If you have configured an organization-wide setting for unmanaged devices, choose a label setting that's either the same or more restrictive

For example, if your tenant is configured for **Allow limited, web-only access**, the label setting that allows full access will have no effect because it's less restrictive. For this tenant-level setting, choose the label setting to block access (more restrictive) or the label setting for limited access (the same as the tenant setting).

Because you can configure the SharePoint settings separately from the label configuration, there's no check in the sensitivity label configuration that the dependencies are in place. These dependencies can be configured after the label is created and published, and even after the label is applied. However, if the label is already applied, the label setting won't take effect until after the user next authenticates.

More information about the dependencies for the authentication context option

To display in the drop-down list for selection, authentication contexts must be created, configured, and published as part of your Microsoft Entra Conditional Access configuration. For more information and instructions, see the Configure authentication contexts section from the Microsoft Entra Conditional Access documentation.

Not all apps support authentication contexts. If a user with an unsupported app connects to the site that's configured for an authentication context, they see either an access denied message or they are prompted to authenticate but rejected. The apps that currently support authentication contexts:

- Office for the web, which includes Outlook for the web
- Microsoft Teams for Windows and macOS (excludes Teams web app)
- Microsoft Planner
- Microsoft 365 Apps for Word, Excel, and PowerPoint; minimum versions:
  - Windows: 2103
  - macOS: 16.45.1202
  - iOS: 2.48.303
  - Android: 16.0.13924.10000
- Microsoft 365 Apps for Outlook; minimum versions:
  - Windows: 2103
  - macOS: 16.45.1202
  - iOS: 4.2109.0
  - Android: 4.2025.1

- OneDrive sync app, minimum versions:
  - Windows: 21.002
  - macOS: 21.002
  - iOS: 12.30
  - Android: Not yet supported

Known limitations:

- For the OneDrive sync app, supported for OneDrive only and not for other sites.
- The following features and apps might be incompatible with authentication contexts, so we encourage you to check that these continue to work after a user successfully accesses a site by using an authentication context:
  - Workflows that use Power Apps or Power Automate
  - Third-party apps

Configure settings for the default sharing link type for a site by using PowerShell advanced settings

In addition to the label settings for sites and groups that you can configure from the Microsoft Purview portal, you can also configure the default sharing link type for a site. Sensitivity labels for documents can also be configured for a default sharing link type. These settings that help to prevent over-sharing are automatically selected when users select the **Share** button in their Office apps.

For more information and instructions, see Use sensitivity labels to configure the default sharing link type for sites and documents in SharePoint and OneDrive.

Configure site sharing permissions by using PowerShell advanced settings

Another PowerShell advanced setting that you can configure for the sensitivity label to be applied to a SharePoint site is **MembersCanShare**. This setting is the equivalent configuration that you can set from the SharePoint admin center > **Site permissions** > **Site Sharing** > **Change how members can share** > **Sharing permissions**.

The three options are listed with the equivalent values for the PowerShell advanced setting **MembersCanShare**:

Expand table

Option from the SharePoint admin center	Equivalent PowerShell value for MembersCanShare
Site owners and members can share files, folders, and the site. People with Edit permissions can share files and folders.	MemberShareAll
Site owners and members, and people with Edit permissions can share files and folders, but only site owners can share the site.	MemberShareFileAndFolder
Only site owners can share files, folders, and the site.	MemberShareNone

For more information about these configuration options, see Change how members can share from the SharePoint community documentation.

Example, where the sensitivity label GUID is **8faca7b8-8d20-48a3-8ea2-0f96310a848e**:

PowerShell

```
Set-Label -Identity 8faca7b8-8d20-48a3-8ea2-0f96310a848e -AdvancedSettings
@{MembersCanShare="MemberShareNone"}
```

For more help in specifying PowerShell advanced settings, see PowerShell tips for specifying the advanced settings.

Sensitivity label management

Use the following guidance for when you create, modify, or delete sensitivity labels that are configured for sites and groups.

Creating and publishing labels that are configured for sites and groups

Use the following guidance to publish a label for your users when that label is configured for site and group settings:

1. After you create and configure the sensitivity label, add this label to a label policy that applies to just a few test users.

2. Wait for the change to replicate:

    - New label: Wait at least one hour, unless your configured settings include Teams shared channel controls. If that's the case, wait at least 24 hours.
    - Existing label: Wait at least 24 hours.

For more information about the timing of labels, see When to expect new labels and changes to take effect.

3. After this wait period, use one of the test user accounts to create a team, Microsoft 365 group, or SharePoint site with the label that you created in step 1.

4. If there are no errors during this creation operation, you know it's safe to publish the label to all users in your tenant.

Modifying published labels that are configured for sites and groups

As a best practice, don't change the site and group settings for a sensitivity label after the label has been applied to teams, groups, or sites. If you do, remember to wait at least 24 hours for the changes to replicate to all containers that have the label applied.

In addition, if your changes include the **External users access** setting:

- The new setting applies to new users but not to existing users. For example, if this setting was previously selected and as a result, guest users accessed the site, these guest users can still access the site after this setting is cleared in the label configuration.

- The privacy settings for the group properties hiddenMembership and roleEnabled aren't updated.

Deleting published labels that are configured for sites and groups

If you delete a sensitivity label that has the site and group settings enabled, and that label is included in one or more label policies, this action can result in creation failures for new teams, groups, and sites. To avoid this situation, use the following guidance:

1. Remove the sensitivity label from all label policies that include the label.

2. Wait at least one hour.

3. After this wait period, try creating a team, group, or site and confirm that the label is no longer visible.

4. If the sensitivity label isn't visible, you can now safely delete the label.

How to apply sensitivity labels to containers

You're now ready to apply the sensitivity label or labels to the following containers:

- Microsoft 365 group in Microsoft Entra ID
- Microsoft Teams team site
- Microsoft 365 group in Outlook on the web
- SharePoint site
- Loop workspaces
- Viva Engage communities

You can use PowerShell if you need to apply a sensitivity label to multiple sites.

Apply sensitivity labels to Microsoft 365 groups

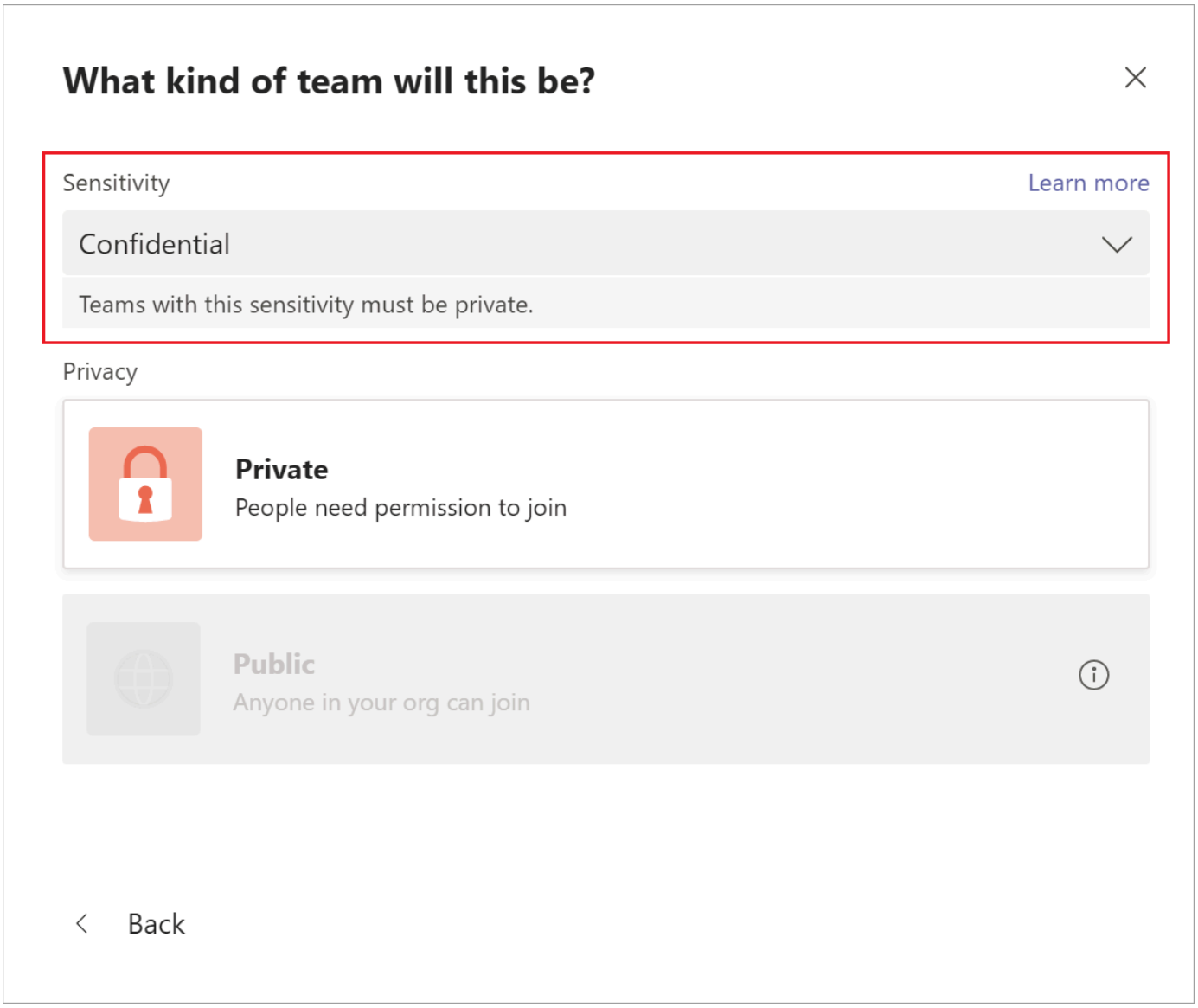
You're now ready to apply the sensitivity label or labels to Microsoft 365 groups. Return to the Microsoft Entra documentation for instructions:

- Assign a label to a new group in Azure portal
- Assign a label to an existing group in Azure portal
- Remove a label from an existing group in Azure portal.

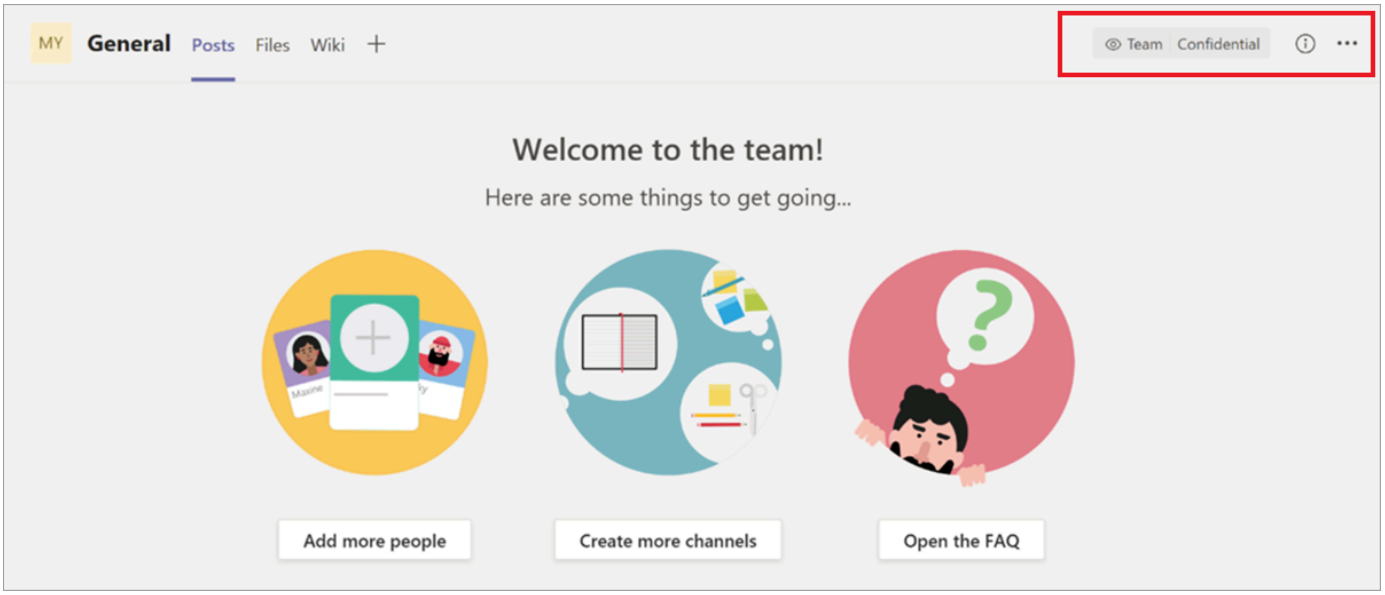
Apply a sensitivity label to a new team

Users can select sensitivity labels when they create new teams in Microsoft Teams. When they select the label from the **Sensitivity** dropdown, the privacy setting might change to reflect the label configuration. Depending on the external users access setting you selected for the label, users can or can't add people outside the organization to the team.

Learn more about sensitivity labels for Teams



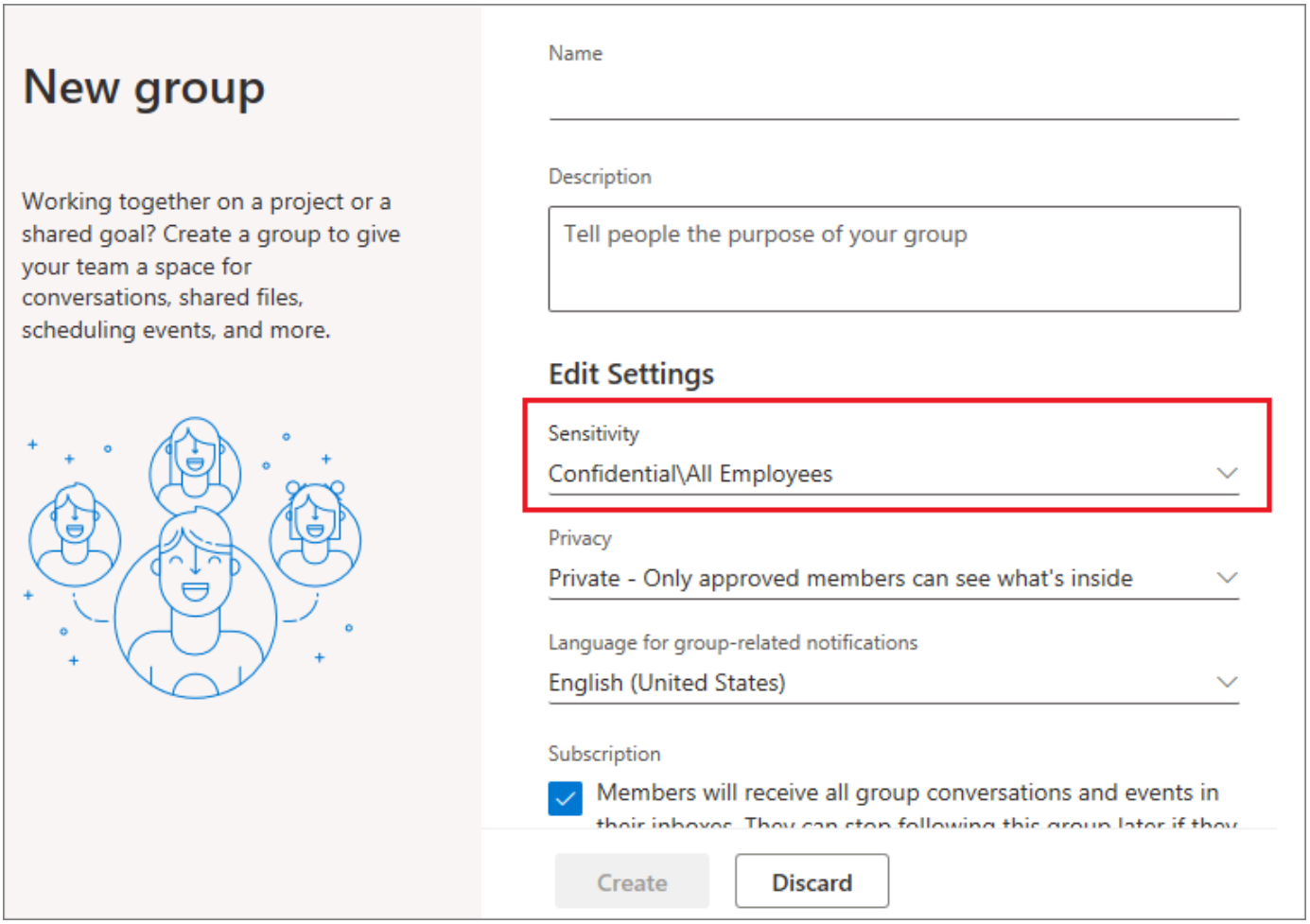
After you create the team, the sensitivity label appears in the upper-right corner of all channels.



The service automatically applies the same sensitivity label to the Microsoft 365 group and the connected SharePoint team site.

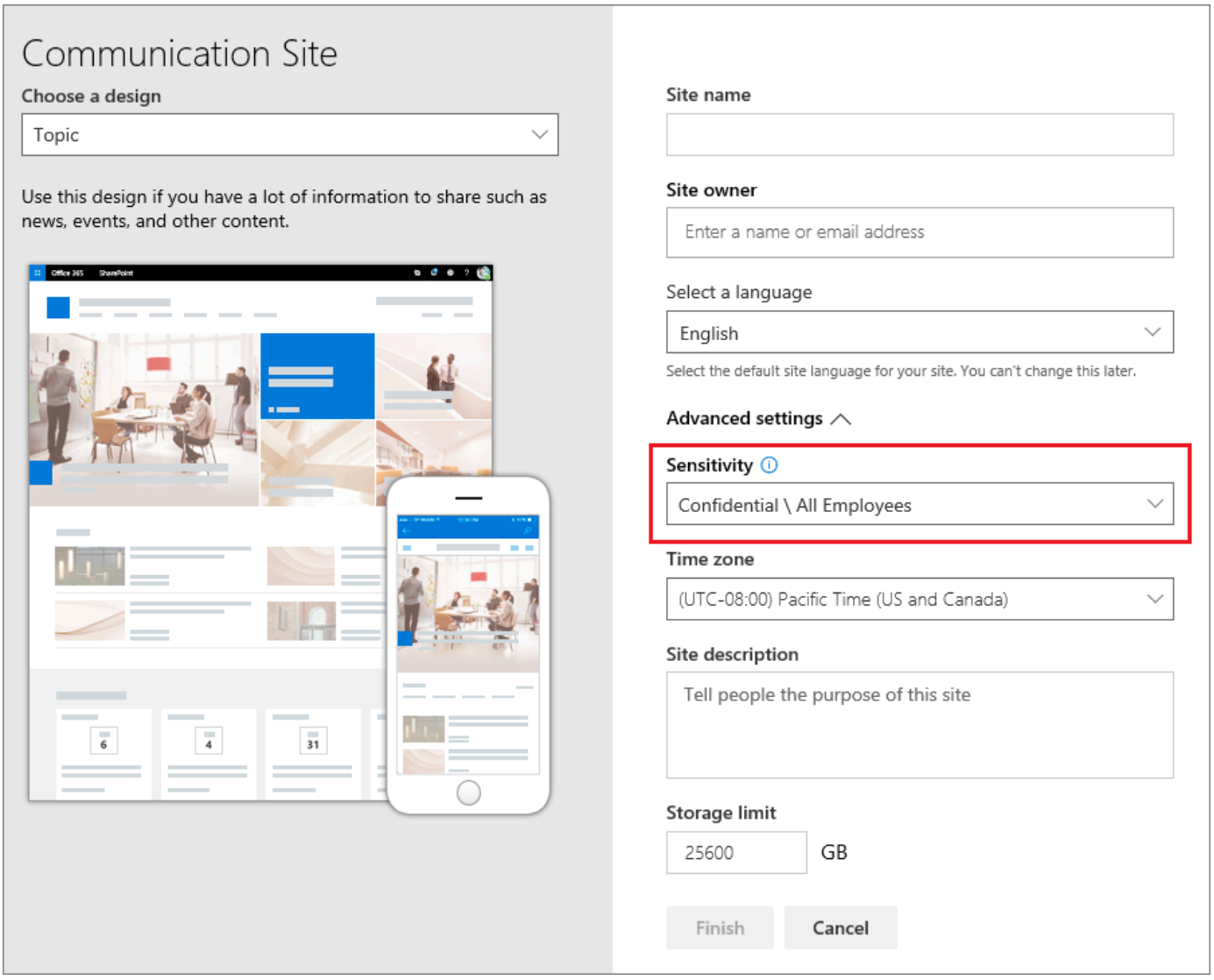
Apply a sensitivity label to a new group in Outlook on the web

In Outlook on the web, when you create a new group, you can select or change the **Sensitivity** option for published labels:



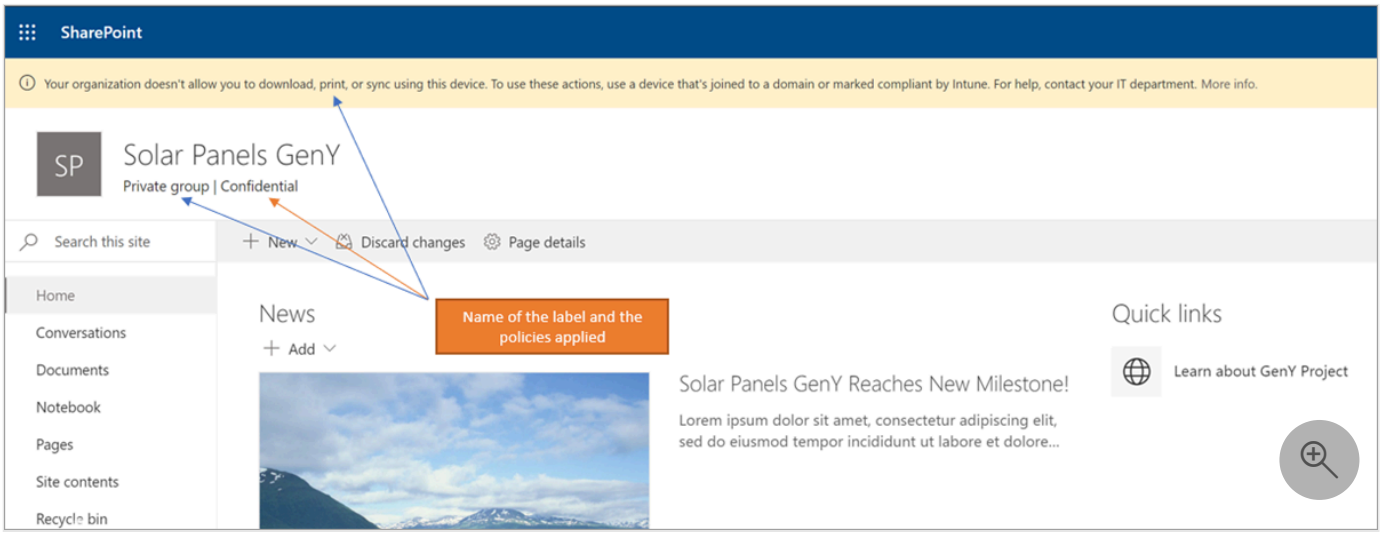
Apply a sensitivity label to a new site

Admins and end users can select sensitivity labels when they create modern team sites and communication sites, and expand **Advanced settings**:



The dropdown box displays the label names for the selection, and the help icon displays all the label names with their tooltip, which can help users determine the correct label to apply.

When the label is applied, and users browse to the site, they see the name of the label and applied policies. For example, this site has been labeled as **Confidential**, and the privacy setting is set to **Private**:



Use PowerShell to apply a sensitivity label to multiple sites

You can use the Set-SPOSite and Set-SPOTenant cmdlet with the *SensitivityLabel* parameter from the current SharePoint Online Management Shell to apply a sensitivity label to many sites. You can use the same procedure to replace an existing label. The sites can be any SharePoint site collection, or a OneDrive site.

Make sure you have version 16.0.19418.12000 or later of the SharePoint Online Management Shell.

1. Open a PowerShell session with the **Run as Administrator** option.
2. If you don't know your label GUID: Connect to Security & Compliance PowerShell and get the list of sensitivity labels and their GUIDs.

   PowerShell

   ```
   Get-Label |ft Name, Guid
   ```

3. Now connect to SharePoint Online PowerShell and store your label GUID as a variable. For example:

   PowerShell

   ```
   $Id = [GUID]("e48058ea-98e8-4940-8db0-ba1310fd955e")
   ```

4. Create a new variable that identifies multiple sites that have an identifying string in common in their URL. For example:

   PowerShell

   ```
   $sites = Get-SPOSite -IncludePersonalSite $true -Limit all -Filter "Url -like 'documents"
   ```

5. Run the following command to apply the label to these sites. Using our examples:

   PowerShell

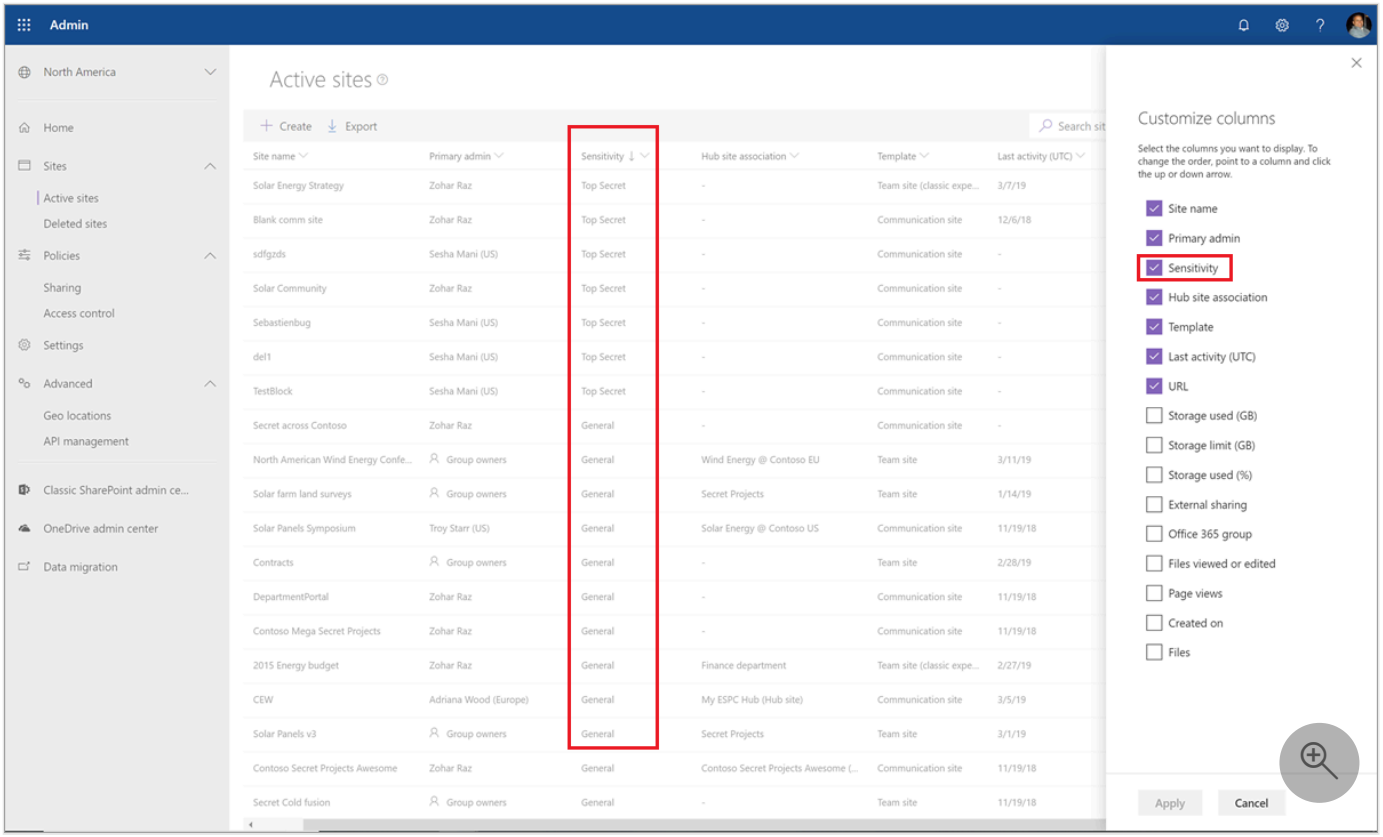
   ```
   $sites | ForEach-Object {Set-SPOTenant $_.url -SensitivityLabel $Id}
   ```

This series of commands lets you label multiple sites across your tenant with the same sensitivity label, which is why you use the Set-SPOTenant cmdlet, rather than the Set-SPOSite cmdlet that's

for per-site configuration. However, use the Set-SPOSite cmdlet when you need to apply a different label to specific sites by repeating the following command for each of these sites: `Set-SPOSite -Identity <URL> -SensitivityLabel "<labelguid>"`

View and manage sensitivity labels in the SharePoint admin center

To view, sort, and search the applied sensitivity labels, use **Active sites** in the SharePoint admin center. You might need to first add the **Sensitivity** column:



For more information about managing sites from the Active sites page, including how to add a column, see Manage sites in the SharePoint admin center.

You can also change and apply a label from this page:

1. Select the site name to open the details pane.
2. Select the **Policies** tab, and then select **Edit** for the **Sensitivity** setting.
3. From the **Edit sensitivity setting** pane, select the sensitivity label you want to apply to the site. Unlike user apps, where sensitivity labels can be assigned to specific users, the admin center displays all container labels for your tenant. After you've chosen a sensitivity label, select **Save**.

For information about managing sensitivity labels for containers in SharePoint Embedded, which are the equivalent of sites, see Manage SharePoint Embedded containers in SharePoint Admin Center.

Support for sensitivity labels

When you use admin centers that support sensitivity labels, with the exception of the Microsoft Entra admin center, you see all sensitivity labels for your tenant. In comparison, user apps and services that filter sensitivity labels according to publishing policies can result in you seeing a subset of those labels. The Microsoft Entra admin center also filters the labels according to publishing policies.

The following apps and services support sensitivity labels configured for sites and group settings:

- Admin centers:
  - SharePoint admin center
  - Teams admin center
  - Microsoft 365 admin center
  - Microsoft Purview portal
- User apps and services:
  - SharePoint
  - Teams
  - Outlook on the web and for Windows, macOS, iOS, and Android
  - Forms
  - Stream
  - Planner
  - Loop
  - Viva Engage

The following apps and services don't currently support sensitivity labels configured for sites and group settings:

- Admin centers:
  - Exchange admin center
- User apps and services:
  - Dynamics 365
  - Project
  - Power BI

- My Apps portal

Auditing sensitivity label activities

> **Important**
>
> If you use label separation by selecting just the **Groups & sites** scope for labels that protect containers: Because of the **Detected document sensitivity mismatch** audit event and email described in this section, consider ordering labels before labels that have a scope for **Files & other data assets**.

If somebody uploads a document to a site that's protected with a sensitivity label and their document has a higher priority sensitivity label than the sensitivity label applied to the site, this action isn't blocked. For example, you've applied the **General** label to a SharePoint site, and somebody uploads to this site a document labeled **Confidential**. Because a sensitivity label with a higher priority identifies content that's more sensitive than content that has a lower priority order, this situation could be a security concern.

Although the action isn't blocked, it automatically generates an email that's sent to the person who uploaded the document, and also sent to site owners and site admins (maximum of 100 in total). As a result, both the user and these administrators can identify documents that have this misalignment of label priority and take action if needed. For example, delete or move the uploaded document from the site.

> **Important**
>
> If the file is in the Preservation Hold library, it's not supported to interact with files in this location. Files can automatically move into the Preservation Hold library as a result of compliance requirements to automatically retain files that users delete or are cloud attachments. For more information, see Learn about retention for SharePoint and OneDrive.

It wouldn't be a security concern if the document has a lower priority sensitivity label than the sensitivity label applied to the site. For example, a document labeled **General** is uploaded to a site labeled **Confidential**. In this scenario, an auditing event and email aren't generated.

> **Note**

Just as for the policy option that requires users to provide a justification for changing a label to a lower classification, sublabels for the same parent label are all considered to have the same priority.

To search the audit log for this event, look for **Detected document sensitivity mismatch** from the **File and page activities** category.

The automatically generated email has the subject **Incompatible sensitivity label detected** and the email message explains the labeling mismatch with a link to the uploaded document and site. It also contains a line for your own internal documentation: **HelpLink : Troubleshooting Guide**. You must configure the hyperlink for the troubleshooting guide by using the Set-SPOTenant cmdlet with the *LabelMismatchEmailHelpLink* parameter. For example:

PowerShell

```
Set-SPOTenant -LabelMismatchEmailHelpLink "https://support.contoso.com"
```

The email message also has a Microsoft documentation link that provides basic information for users to change the sensitivity label: Apply sensitivity labels to your files and email in Office

Except for the internal URL that you must specify, these automated emails can't be customized. However, you can prevent them from being sent when you use the following PowerShell command from Set-SPOTenant:

PowerShell

```
Set-SPOTenant -BlockSendLabelMismatchEmail $True
```

When somebody adds or removes a sensitivity label to or from a site or group, these activities are also audited but without automatically generating an email.

All these auditing events can be found in the Sensitivity label activities section from the audit log activities documentation.

How to disable sensitivity labels for containers

You can turn off sensitivity labels for Microsoft Teams, Microsoft 365 groups, and SharePoint sites by using the same instructions from Enable sensitivity label support in PowerShell. However, to disable the feature, in step 5, specify `$setting["EnableMIPLabels"] = "False"`.

In addition to making all the settings unavailable for groups and sites when you create or edit sensitivity labels, this action reverts which property the containers use for their configuration. Enabling sensitivity labels for Microsoft Teams, Microsoft 365 groups, and SharePoint sites switches the property used from **Classification** to **Sensitivity**. When you disable sensitivity labels for containers, the containers ignore the Sensitivity property and use the Classification property again.

This means that any label settings from sites and groups previously applied to containers won't be enforced, and containers no longer display the labels.

If these containers have Microsoft Entra classification values applied to them, the containers revert to using the classifications again. Be aware that any new sites or groups that were created after enabling the feature won't display a label or have a classification. For these containers, and any new containers, you can now apply classification values. For more information, see SharePoint "modern" sites classification and Create classifications for Office groups in your organization.

Additional resources

- Overview of Teams and SharePoint integration
- Use sensitivity labels with Microsoft Loop

Last updated on 05/07/2026

Use sensitivity labels to protect calendar items, Teams meetings, and chat

> *Microsoft Purview service description*

> **Note**
>
> You won't be able to configure all the options referenced on this page if a Teams Premium license isn't found for your tenant. For those settings, you'll see an information bar in the Microsoft Purview portal that your organization doesn't have this license.

In addition to using sensitivity labels to protect documents and emails, you can extend sensitivity labels to protect meeting invites and responses that use Outlook and Teams, and to protect Teams meetings and chat:

- Meeting organizers apply a sensitivity label to meeting invites from Outlook or Teams. Recipients in your organization see the sensitivity label and all recipients see any headers or footers as configured content markings. Optionally, the meeting invite can be encrypted so only authorized people can see it and access the meeting link. Usage rights can further restrict access, for example, preventing the invite from being forwarded.

  The encryption works similarly to how you can apply a label with encryption to an email. As with email encryption, if recipients aren't using an email client that can decrypt this meeting invite, they're directed to Outlook on the web or the encryption portal where they can read the invite and access the meeting link.

- The sensitivity label from the meeting invite is automatically inherited by the meeting and protects it even before the meeting starts. For example, prevent copying from the chat if this has been configured for the label. When the meeting starts, it will enforce any meeting options, such as who can bypass the lobby and present.

- Participants from your organization see the label during the Teams meetings. Any meeting settings applied by the label can't be changed directly in Teams. When a label enforces these settings, users see a lock and message that explains the settings are applied by a sensitivity label.

Meeting settings that you can apply with a sensitivity label:

- Who can bypass the lobby
- Who can present
- Who can record and transcribe
- Encryption for meeting video and audio
- Automatically record
- Video watermark for screen sharing and camera streams
- Prevent copying and forwarding of meeting chat, live caption, and transcripts
- Detect sensitive content during screen sharing

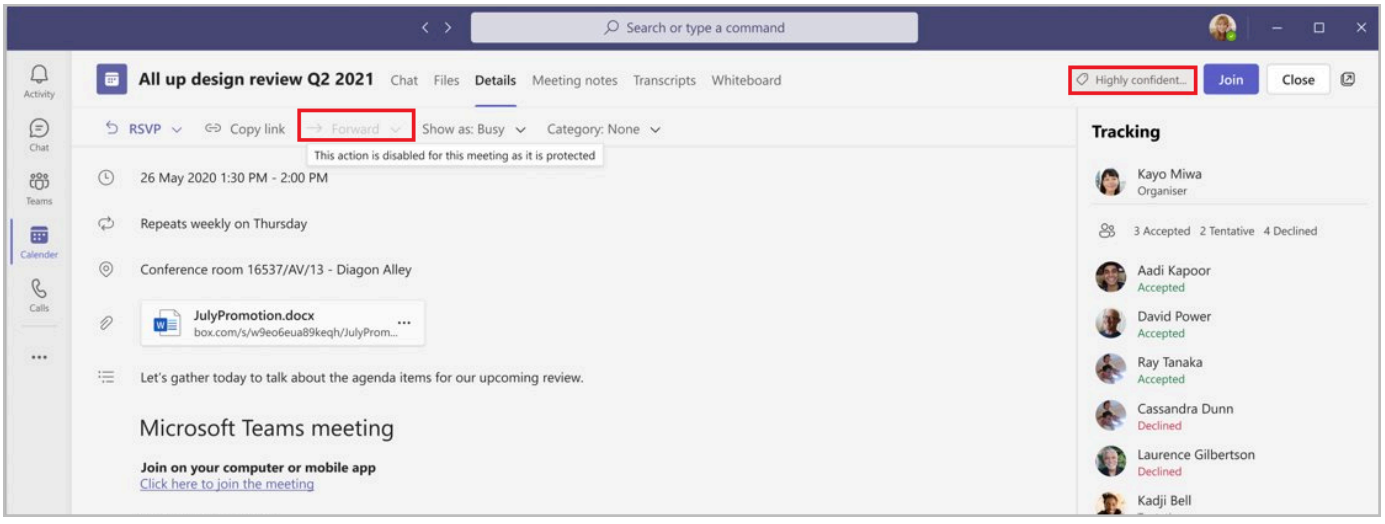
Users can apply sensitivity labels to meeting invites when they use Outlook or Teams:

- For Outlook, applying, changing, and viewing sensitivity labels for meeting invites and appointments are supported with Outlook on the web, and with minimum versions of Windows, macOS, iOS and Android. For other versions, platforms, and email clients, encrypted meeting invites can be opened in the same way as encrypted email.

- For Teams, applying, changing, and viewing sensitivity labels for meeting invites are supported with Teams on the web, and with minimum versions of Windows, macOS, iOS and Android.

For more information about those minimum versions, see the Requirements section.

Example showing a Teams meeting invite that has the label **Highly confidential**, and the label applies encryption settings that prevent the invite from being forwarded:



Requirements

To apply a sensitivity label to meeting invites and appointments using Outlook, users must use Outlook on the web or use built-in labeling from Microsoft 365 Apps for enterprise. For supported versions across Windows, macOS, iOS, and Android, see the capabilities table for Outlook and the row that specifies labeling for **Calendar items**.

To apply a sensitivity label to meeting invites using Teams, enforce meeting options and chat, and display and change sensitivity labels in Teams non-channel meetings, users must use Teams on the web with Edge or Chrome, or one of the following Teams clients:

- **Teams client for Windows**: 1.5.00.32158+
- **Teams client for macOS**: 1.5.00.32154+
- **Teams mobile app for iOS**: 5.23.0+
- **Teams mobile app for Android**: 1416/1.0.0.2023235001+

> ⓘ **Important**
>
> If meeting participants or organizers don't use a Teams client that meets these requirements, the configured label meeting options can't be enforced for these users.

Sensitivity labels that apply S/MIME protection or Double Key Encryption (DKE) can't be used to protect calendar items, Teams meetings, and chat.

Limitations

- Preventing copying chat and meeting-related content has some limitations that are listed in the Prevent copying and forwarding for chats and meeting content section on this page.
- Calendar items don't support label encryption that uses S/MIME or Double Key Encryption. You must use the default encryption that uses the Azure Rights Management service, with either admin-defined permissions (the option **Assign permissions now**) or user-defined permissions (the option **Let users assign permissions**).
- Automatic and recommended labeling isn't supported.
- No support for mailboxes on-premises; user mailboxes must be in Exchange Online.

Specific to Outlook:

- No support for group calendar meeting invites; participants must be specific users.

- As with labeled and encrypted email, if somebody forwards a meeting invite from an email client other than Outlook, any applied encryption persists but information about the sensitivity label is removed from the email headers.

- The Share to Teams option doesn't support sensitivity labels. If the selected sensitivity label applies encryption, the action is prevented with a message that the content is restricted. If the selected sensitivity label doesn't apply encryption, the action isn't prevented but the sensitivity label isn't displayed in Teams.

- For Outlook for Windows, when a labeled and encrypted meeting invite or response is sent, the following configuration setting must be enabled: **Outlook Options** > **Advanced** > **International options** > **Automatically select encoding for outgoing messages**

- Shared calendars:
  - When a user shares their calendar with others, only the calendar owner can create and change sensitivity labels for meetings in that calendar.
  - A shared mailbox calendar doesn't support applying and changing sensitivity labels for meetings.

- For a meeting series:
  - The organizer can apply a sensitivity label for the meeting series but not to individual occurrences.
  - Exceptions for a meeting series will be labeled (and encrypted if applicable) only if they're created after the label is applied to the meeting series. Existing exceptions, including those in the future, won't be labeled.
  - If the organizer changes the meeting body or attachments for a meeting series after it's encrypted by a sensitivity label, existing exceptions won't be updated with those changes.
  - If the organizer changes any meeting attributes (such as start date, end date, location, the meeting body, or attachments) for a specific occurrence in a series after it's encrypted by a sensitivity label, this action automatically creates an exception for the series with the same label for the meeting series.

Specific to Teams:

- While a meeting is in session, a sensitivity label can't prevent participants from inviting people to join by copying the meeting link, or by sharing via default email or the Outlook calendar. These choices are from the Teams meeting **Share invite** option.

- If there's a label change made in Teams while the meeting is progress, any changes to the following meeting options won't take effect for the current meeting unless the organizer ends and restarts the meeting:
  - Who can record and transcribe
  - Encryption for meeting video and audio
  - Automatically record and transcribe
  - Video watermark for screen sharing and camera streams

- The following meeting options won't take effect for a Meet Now meeting:
  - Who can record and transcribe
  - Encryption for meeting video and audio
  - Automatically record and transcribe
  - Video watermark for screen sharing and camera streams

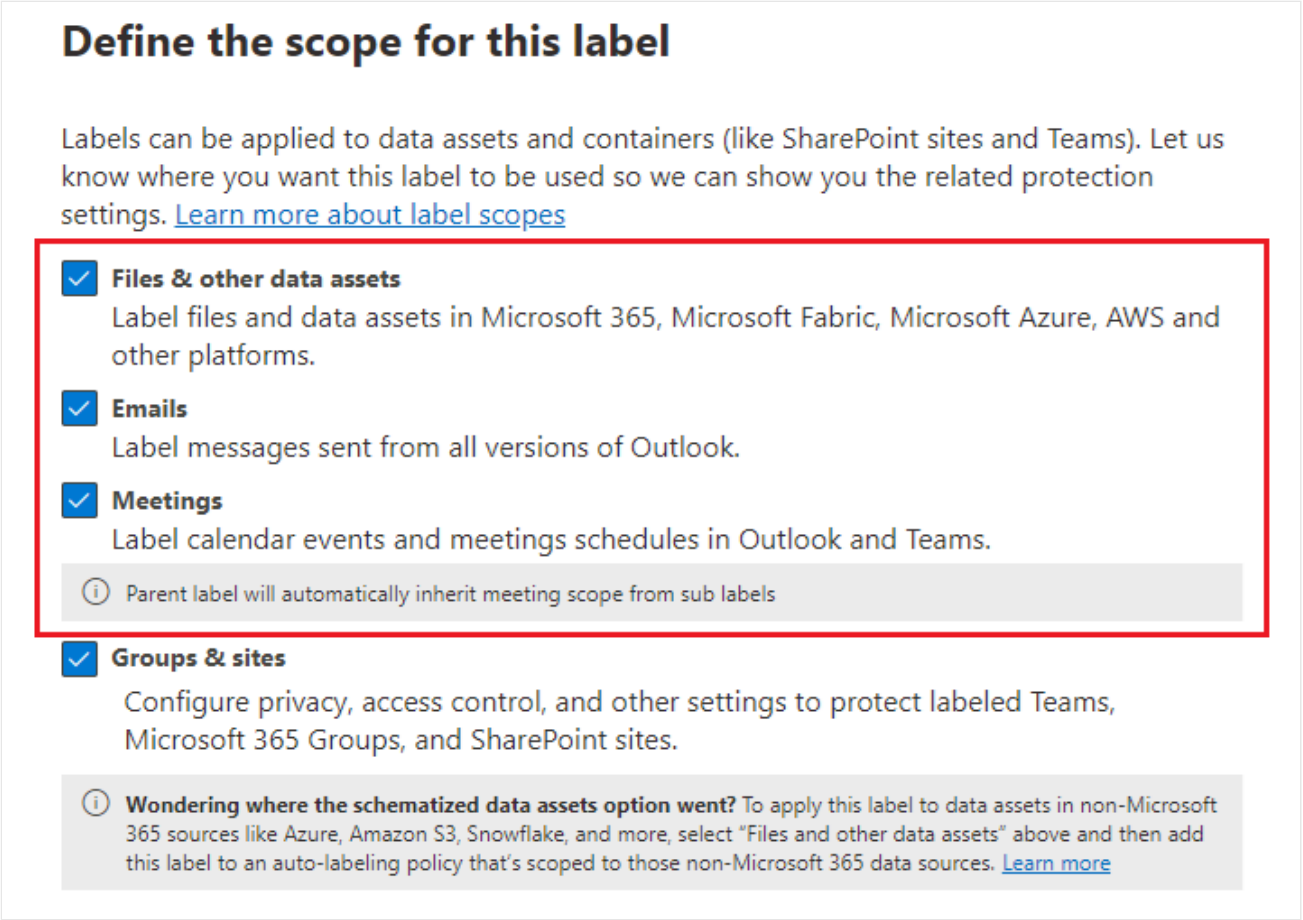
- Sensitivity labels can't be applied to live events.

- Labeling meeting invites with Graph APIs isn't supported.

How to configure a sensitivity label to protect calendar items, Teams meetings, and chat

1. Follow the general instructions to create or edit a sensitivity label and make sure the following label scopes are selected:

   - **Files & other data assets**
   - **Emails**
   - **Meetings**



All three options are needed because of their interdependency. A meeting consists of a calendar event, the invitation and responses, related attachments, related emails, and the actual Teams meeting. As a result, all these items need the same sensitivity label applied to protect the meeting.

2. On the **Choose protection settings for labeled items** page:

   - Select **Apply or remove encryption** if you want to encrypt meeting invites and responses, and any Office attachment in that calendar item
   - Select **Apply content marking** if you want to add headers or footers to meeting invites and responses
   - Select **Protect Teams meetings and chat** to display the label in Teams meetings and enforce Teams-specific settings for the Teams meeting itself and chat messages

3. On the subsequent pages, configure settings for the options you've selected.

   For more information about configuring encryption for meeting invites, see Restrict access to content by using sensitivity labels to apply encryption.

   For more information about the headers and footers that you can apply to meeting invites, see the **Mark the content** information from What sensitivity labels can do.

For more information about the settings that you can configure for Teams meetings and chat:

- Use the Teams documentation for an explanation of each setting. For information about how the settings interact with other configuration methods for Teams, see Use Teams meeting templates, sensitivity labels, and admin policies together for sensitive meetings.
- Specific to the option to prevent copy in chat, see the next section on this page, Prevent copying chat to the clipboard label setting.

4. Complete the label configuration and then create or configure other sensitivity labels if you need different settings for other labels.

Prevent copying and forwarding for chats and meeting content

The label setting **Prevent copying and forwarding of meeting chat, live caption, and transcripts** also extends to AI-generated meeting recap content.

The setting is enforced for all channel chats, even outside channel meetings. For non-channel meetings, it's enforced only for meetings.

The methods supported to prevent these actions:

- Select the text and then right-click > **Copy** or Ctrl+C
- Forward messages
- Share to Outlook
- Copy link

This setting can't prevent screen captures and doesn't apply to developer tools or third-party apps.

Configure a corresponding sensitivity label publishing policy

1. Make sure the label is published to meeting organizers and participants.
2. Configure label policy settings that are specific just to calendar items, Teams meetings, and chat:
   - **Apply a default label to meetings and calendar events**

- For Outlook: The default label is applied to new calendar events and to existing unlabeled calendar events when they are updated.
- For Teams: The default label is applied to new calendar events but won't be automatically applied when an existing unlabeled meeting is updated.

> ⓘ **Note**
>
> These settings are independent from, rather than override, the similar policy settings for emails. For example, if you don't select the setting to apply a default label to meetings and calendar events, but do select the setting to apply a default label for emails, no default label is applied to calendar events.

- **Require users to apply a label to their meetings and calendar events**
- **Inherit label from files shared to meetings**

  For Teams meetings, automatically apply or recommend that the meeting is labeled with the highest priority sensitivity label from files shared for the meeting when that label is published to the meeting organizer.

  These files can be shared in the chat or uploaded in the **Shared** tab of the meeting, or presented via Live share. Files shared via screen share or Window share, or shared any other way won't be inherited to meetings.

  A meeting that already has a sensitivity label with a higher priority label won't be overriden. However, a manually applied sensitivity label with a lower priority can be overriden.

  For Teams desktop and web clients, a notification event is sent to the meeting organizer if label inheritance happens before the meeting starts.

- **Apply meeting label to artifacts**

  This setting is in preview and for MP4 files, requires you to enable sensitivity labeling support for MP4.

  > ⓘ **Important**
  >
  > If MP4 support isn't enabled, these files aren't labeled when you select this option.

For Teams meetings, automatically apply the meeting's sensitivity label to recordings and their transcripts (.mp4 files), and to meeting notes (.loop files) at the time the artifact is saved to SharePoint or OneDrive. If the meeting label changes later, the inherited label doesn't automatically change.

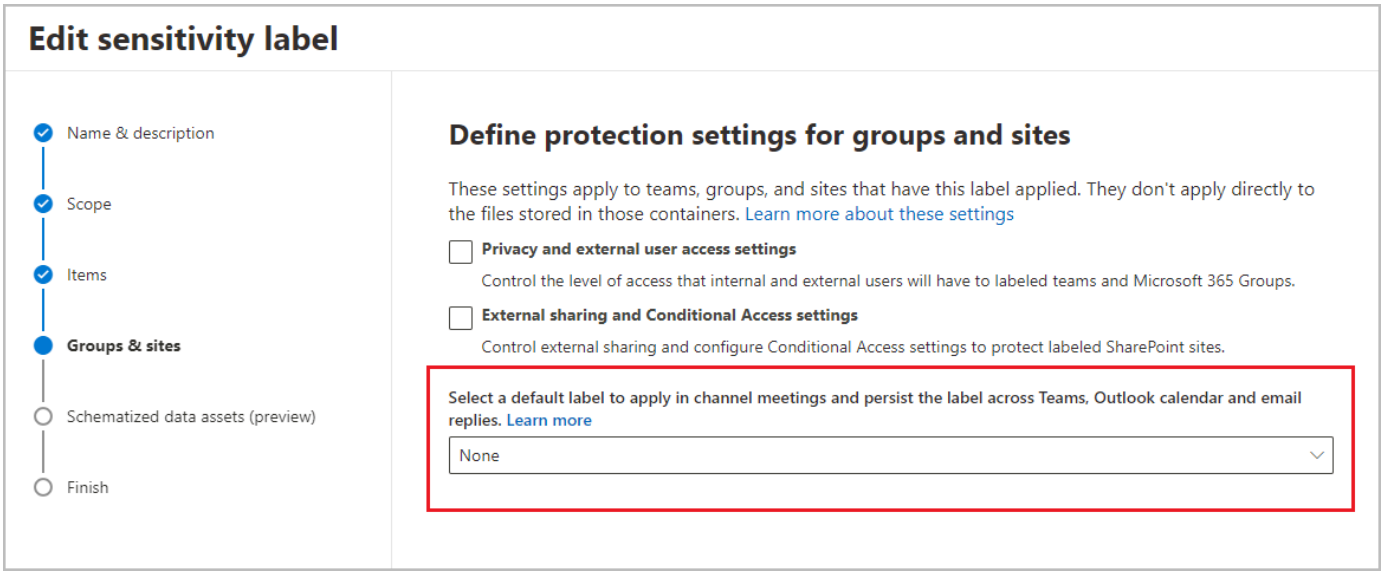
Recordings respect any access permissions and usage rights from label encryption settings and can be played directly in a web browser and supported experiences such as SharePoint, Teams, or Clipchamp. However, if a recording and transcript has a label that applies encryption, you can't download them. You can download them if the label doesn't apply encryption, but a downloaded file in this case doesn't retain the labeling information.

You can, if required, manually relabel these files in SharePoint or OneDrive, by using the file's **Details** pane. For example, change the label to one that doesn't apply encryption so that you can then download and share with others. As with all sensitivity labels, you must have the appropriate permissions to change a sensitivity label that applies encryption, which is always the case for the meeting owner.

How to configure and apply a label for channel meetings

To apply your label meeting options to channel meetings, the sensitivity label must be associated with a container label for the channel. Then, all channel meetings and chats inherit the associated label, which can't be changed by users.

This association is a setting that's available only when you edit an existing label for groups and sites. Then, you can select a sensitivity label that you've configured in the previous procedure:



> **Note**
>
> The option to select this default sensitivity label to protect channel meetings and chat is available only when both of the following conditions apply:
>
> - You're editing an existing sensitivity label, rather than creating a new label.
> - The label's scope includes files, emails, and meetings.

If you select a label that applies encryption, it can be a label with admin-defined permissions (**Assign permissions now**) or user-defined permissions (**Do Not Forward** and **Encrypt-Only**).

End-user documentation

To help end users understand how to apply and change sensitivity labels in Teams and Outlook, see the following articles:

- Sensitivity labels for Teams meetings.
- Apply sensitivity labels to your Outlook meetings.

Remember to provide your own guidance which named label to apply for different types of meetings. Then, users can focus on the label name instead of the individual settings applied by the label.

Related articles

- Learn about Double Key Encryption

Last updated on 06/10/2026

Enable sensitivity labels for files in SharePoint and OneDrive

> *Microsoft Purview service description*

Enable built-in labeling for supported Office files and PDF files in SharePoint and OneDrive so that users can apply your sensitivity labels in Office for the web. When this feature is enabled, users see the **Sensitivity** button on the ribbon so they can apply labels, and see any applied label name on the status bar.

Users can also see and apply sensitivity labels without having to first open the files, by using the details pane in SharePoint and OneDrive. This labeling method is also available from the **Files** tab in Microsoft Teams.

Enabling this feature also results in SharePoint and OneDrive being able to process the contents of Office files and optionally, PDF documents that have been encrypted by using a sensitivity label. The label can be applied in Office for the web, or in Office desktop apps and uploaded or saved in SharePoint and OneDrive. Until you enable this feature, these services can't process encrypted files, which means that coauthoring, eDiscovery, data loss prevention, search, and other collaborative features won't work for these files.

After you enable sensitivity labels for these files in SharePoint and OneDrive, for new and changed files that have a sensitivity label that applies encryption with a cloud-based key (and doesn't use Double Key Encryption):

- For Word, Excel, and PowerPoint files, and uploaded PDF files, SharePoint and OneDrive recognize the label and can now process the contents of the encrypted file.
- When users download or access these files from SharePoint or OneDrive, the sensitivity label and any encryption settings from the label are enforced and remain with the file, wherever it's stored. Ensure you provide user guidance to use only labels to protect documents. For more information, see Information Rights Management (IRM) options and sensitivity labels.
- When users upload labeled and encrypted files to SharePoint or OneDrive, they must have at least View usage rights to those files. For example, they can open the files outside

SharePoint. If they don't have this minimum usage right, the upload is successful but the service doesn't recognize the label and can't process the file contents.

- Use Office for the web (Word, Excel, PowerPoint) to open and edit Office files that have sensitivity labels that apply encryption. The permissions that were assigned with the encryption are enforced. You can also use auto-labeling for these documents.
- External users can access documents that are labeled with encryption by using guest accounts. For more information, see Support for external users and labeled content.
- eDiscovery supports full-text search for these files and data loss prevention (DLP) policies support content in these files.

> **Note**
>
> If encryption has been applied with an on-premises key (a key management topology often referred to as "hold your own key" or HYOK), or by using **Double Key Encryption**, the service behavior for processing the file contents doesn't change. So for these files, coauthoring, eDiscovery, data loss prevention, search, and other collaborative features won't work.
>
> The SharePoint and OneDrive behavior also doesn't change for existing files in these locations that are labeled with encryption using a single Azure-based key. For these files to benefit from the new capabilities after you enable sensitivity labels for Office files in SharePoint and OneDrive, the files must be either downloaded and uploaded again, or edited.

After you enable sensitivity labels for Office files in SharePoint and OneDrive, three new audit events are available for monitoring sensitivity labels that are applied to documents in SharePoint and OneDrive:

- **Applied sensitivity label to file**
- **Changed sensitivity label applied to file**
- **Removed sensitivity label from file**

You can disable sensitivity labels for Office files in SharePoint and OneDrive (opt-out) at any time.

If you're currently protecting documents in SharePoint by using SharePoint Information Rights Management (IRM), be sure to check the SharePoint Information Rights Management (IRM) and sensitivity labels section on this page.

Requirements

These new capabilities work with sensitivity labels only.

Use the OneDrive sync app version 19.002.0121.0008 or later on Windows, and version 19.002.0107.0008 or later on macOS. Both these versions were released January 28, 2019. For more information, see the OneDrive release notes. After you enable sensitivity labels for SharePoint and OneDrive, users who run an older version of the sync app see a prompt to update it.

Supported file types

After you enable sensitivity labels for SharePoint and OneDrive, the following Office file types support sensitivity labeling scenarios.

Applying a sensitivity label in Office for the web or in SharePoint:

- **Word**: .docx, .docm
- **Excel**: .xlsx, .xlsm, .xlsb
- **PowerPoint**: .pptx, .ppsx

Uploading a labeled Office document, and then extracting and displaying that sensitivity label:

- **Word**: .doc, .docx, .docm, .dot, .dotx, .dotm
- **Excel**: .xls, .xlt, .xla, .xlc, .xlm, .xlw, .xlsx, .xltx, .xlsm, .xltm, .xlam, .xlsb
- **PowerPoint**: .ppt, .pot, .pps, .ppa, .pptx, .ppsx, .ppsxm, .potx, .ppam, .pptm, .potm, .ppsm

Video support (MP4 files)

This support requires you to first enable it by using PowerShell.

MP4 files support manually applying sensitivity labels, but can't be labeled automatically or by default. Additionally, MP4 files can automatically inherit a sensitivity label from a Teams meeting.

To turn on support for MP4 video files, use the Set-SPOTenant cmdlet with the *EnableSensitivityLabelForVideoFiles* parameter:

PowerShell

```
Set-SPOTenant -EnableSensitivityLabelForVideoFiles $true
```

For Microsoft 365 Multi-Geo, you must connect to each of your geo-locations and run the command separately.

When MP4 files have a sensitivity label that applies encryption, they can't be downloaded. All labeled MP4 files support playback in SharePoint, Teams, and Clipchamp.

Sensitivity labels that apply encryption must be configured with the Assign permissions now option, rather than user-defined permissions.

Adding support for PDF

You can enable support for PDFs for the following scenarios:

- Applying a sensitivity label in Office for the web
- Uploading a labeled document, and then extracting and displaying that sensitivity label
- Search, eDiscovery, and data loss prevention
- Auto-labeling policies and default sensitivity labels for SharePoint document libraries

Sensitivity labels don't support signed PDFs.

> **Important**
>
> Enabling PDF support can increase the number of files that get automatically labeled by existing auto-labeling policies, which support a maximum of 100,000 files a day.

To turn on support from the Microsoft Purview portal:

- Sign in to the Microsoft Purview portal > **Solutions** > **Information Protection** > **Sensitivity labels** > **Policies** > **Auto-labeling policies**

  Then, if you see a message **Protect PDFs with Auto-labeling**, select this banner to confirm you want to turn on PDF protection for files in SharePoint and OneDrive.

Alternatively, you can turn on PDF support by using PowerShell. Use the Set-SPOTenant cmdlet with the *EnableSensitivityLabelforPDF* parameter:

PowerShell

```
Set-SPOTenant -EnableSensitivityLabelforPDF $true
```

This PDF parameter requires a minimum version of 16.0.24211.12000 for the SharePoint Online Management Shell. For more information about how to install this PowerShell module or run the cmdlets, see the section on this page to enable support for sensitivity labels.

For Microsoft 365 Multi-Geo: Similarly to the instructions to run the PowerShell command to enable support for sensitivity labels, you must connect to each of your geo-locations, and then run the command to enable support for PDFs.

Support for labels configured for user-defined permissions

User-defined permissions refers to an encryption configuration and the setting **Let users assign permissions when they apply the label** and the checkbox **In Word, PowerPoint, and Excel, prompt users to specify permissions** is selected.

After you've enabled sensitivity labels for SharePoint and OneDrive, content labeled with user-defined permissions can be inspected for search, data loss prevention, and eDiscovery. This functionality applies to only newly uploaded and edited files.

The following actions support labels configured for user-defined permissions:

- When a document is labeled with user-defined permissions and uploaded to SharePoint or OneDrive, these services can now process the document so that it can be opened and edited in Office for the web, and the label name is displayed in the **Sensitivity** column.
- Labels with this configuration are displayed in Office for the web. Before users select these labels to specify users and permissions, or change the label applied, make sure you've enabled co-authoring for files encrypted with sensitivity labels.

To support AutoSave and co-authoring for these encrypted files by using a desktop app, co-authoring for files encrypted with sensitivity labels must first be enabled.

This functionality comes with the following change in behavior for encryption:

- If an existing file that's labeled with user-defined permissions is relabeled with a sensitivity label that doesn't apply encryption, the original encryption is removed rather than preserved. For more information, see What happens to existing encryption when a label's applied.

Limitations

- SharePoint and OneDrive can't process some files that are labeled and encrypted from Office desktop apps when these files contain Power Query data, data stored by custom add-ins, or custom XML parts such as Cover Page Properties, content type schemas, custom Document Information Panel, and Custom XSN. This limitation also applies to files that include a bibliography, and to files that have a Document ID added when they're uploaded.

  For these files, either apply a label without encryption so that users can later open them in Office for the web, or instruct users to open the files in their desktop apps. Files that are labeled and encrypted only in Office for the web aren't affected.

- SharePoint and OneDrive don't support sensitivity labels if you applied the labels before enabling these services for sensitivity labels. The labels aren't recognized and if the labels applied encryption, the contents aren't processed. For the labels and encryption to be supported by SharePoint and OneDrive, download these files and then upload them to their original location.

- SharePoint and OneDrive can't process encrypted files when the label that applied the encryption has either of the following configurations for encryption:

  - **User access to content expires** is set to a value other than **Never**.

  - **Double Key Encryption** is selected.

    For labels with either of these encryption configurations, the labels aren't displayed to users in Office for the web. If they're parent labels, this means that users don't see that label's sublabels, even if the sublabels aren't configured to apply encryption.

    Additionally, the new capabilities can't be used with labeled documents that already have these encryption settings. For example, these documents aren't returned in search results, even if they're updated.

- SharePoint and OneDrive can't read or apply sensitivity labels when the document is password-protected.

- SharePoint and OneDrive prevents downloading MP4 files with sensitivity labels that apply encryption. If these files have labels that don't apply encryption, they can be downloaded but the labeling information is lost.

- For performance reasons, when you upload or save a document to SharePoint and the file's label doesn't apply encryption, the **Sensitivity** column in the document library can take a

while to display the label name. Factor in this delay if you use scripts or automation that depend on the label name in this column.

- If a document is labeled while it's checked out in SharePoint, the **Sensitivity** column in the document library doesn't display the label name until the document is checked in and next opened in SharePoint.

- If a labeled and encrypted document is downloaded from SharePoint or OneDrive by an app or service that uses a service principal name, and then uploaded again with a label that applies different encryption settings, the upload fails. An example scenario is Microsoft Defender for Cloud Apps changes a sensitivity label on a file from **Confidential** to **Highly Confidential**, or from **Confidential** to **General**.

  The upload doesn't fail if the app or service first runs the Unlock-SPOSensitivityLabelEncryptedFile cmdlet, as explained in the Remove encryption for a labeled document section. Or, before the upload, the original file is deleted, or the file name is changed.

- Users might experience delays in being able to open encrypted documents in the following Save As scenario: With a desktop version of Office, a user chooses Save As for a document that has a sensitivity label that applies encryption. The user selects SharePoint or OneDrive for the location, and then immediately tries to open that document in Office for the web. If the service is still processing the encryption, the user sees a message that the document must be opened in their desktop app. If they try again in a couple of minutes, the document successfully opens in Office for the web.

- For encrypted documents, printing, downloading, exporting, and creating a copy aren't supported in Office for the web.

- For encrypted documents in Office for the web, screen captures aren't prevented. However, when documents are labeled and encrypted, and the **Copy** usage right isn't granted, Office for the web prevents copying to clipboard in the same way as desktop apps prevent this action.

- By default, Office desktop apps and mobile apps don't support co-authoring for files that are labeled with encryption. These apps continue to open labeled and encrypted files in exclusive editing mode. To change the default behavior, see Enable co-authoring for files encrypted with sensitivity labels.

- If an admin changes settings for a published label that's already applied to files downloaded to users' sync client, users might be unable to save changes they make to the file in their OneDrive Sync folder. This scenario applies to files that are labeled with encryption, and also when the label change is from a label that didn't apply encryption to a label that does apply encryption. Users see a red circle with a white cross icon error, and they're asked to save new changes as a separate copy. Instead, they can close and reopen the file, or use Office for the web.

- Sensitivity labels that are configured for automatic labeling are supported for Office for the web when the label settings for conditions are for sensitive information types only. Automatic labeling isn't supported for Office for the web when the conditions include trainable classifiers.

- Users can experience save problems after going offline or into a sleep mode when instead of using Office for the web, they use the desktop and mobile apps for Word, Excel, or PowerPoint. For these users, when they resume their Office app session and try to save changes, they see an upload failure message with an option to save a copy instead of saving the original file.

- Documents that are encrypted in the following ways can't be opened in Office for the web:
  - Encryption that uses an on-premises key ("hold your own key" or HYOK)
  - Encryption that was applied by using Double Key Encryption
  - Encryption that was applied independently from a label, for example, by directly applying a Rights Management protection template.

- Labels configured for other languages aren't supported and display the original language only.

- If you delete a label that's been applied to a document in SharePoint or OneDrive, rather than remove the label from the applicable label policy, the document when downloaded isn't labeled or encrypted. In comparison, if the labeled document is stored outside SharePoint or OneDrive, the document remains encrypted if the label is deleted. Although you might delete labels during a testing phase, it's very rare to delete a label in a production environment.

- When a labeled and encrypted Office file larger than 12 MB is copied or moved to a different site, SharePoint can no longer process this file. As a result, the **Sensitivity** column doesn't display the label name and users can't open the file using Office for the web. Users can successfully open the file when they use an Office client app.

- In the following scenario, users from an external tenant can't open files protected with a user-defined permissions sensitivity label in Office for the web. The same limitation also applies when users open the file from SharePoint or OneDrive with Microsoft Office desktop applications.
  - A user (UserA) from an external tenant (fabrikam) is invited as a guest user to the contoso tenant.
  - The sensitivity label applied to the file grants permissions to a group in the external fabrikam tenant, and UserA is a member of that group.
  - The file is uploaded to SharePoint in the contoso tenant, and UserA attempts to open the file in Office for the web.

  To allow an external tenant user to open a file protected with a user-defined permissions sensitivity label, use one of the following methods:
  - Grant permissions directly to the external user in the sensitivity label.
  - Add the external user to a group in the contoso tenant, and grant permissions to that contoso tenant group in the sensitivity label.

How to enable sensitivity labels for SharePoint and OneDrive (opt-in)

> ⓘ **Note**
>
> Enabling sensitivity labels for SharePoint and OneDrive automatically enables **sensitivity labels for Loop components and pages**. However, enabling sensitivity labels for OneNote requires an additional step that uses PowerShell.

You can enable the new capabilities by using the Microsoft Purview portal, or by using PowerShell. See the following sections for instructions.

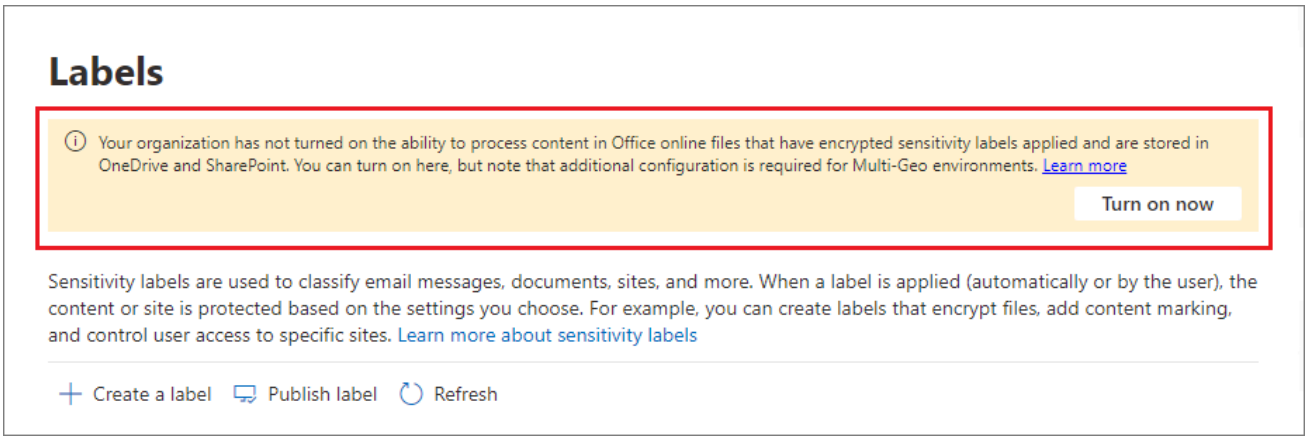
As with all tenant-level configuration changes for SharePoint and OneDrive, it takes about 15 minutes for the change to take effect.

Use the portal to enable support for sensitivity labels

This option is the easiest way to enable sensitivity labels for SharePoint and OneDrive. You must sign in as a global administrator for your tenant.

1. Sign in to the Microsoft Purview portal > **Solutions** > **Information Protection** > **Sensitivity labels**.

2. If you see a message to turn on the ability to process content in Office online files, select **Turn on now**:



   The command runs immediately and when the page is next refreshed, you no longer see the message or button.

> **Note**
>
> If you have Microsoft 365 Multi-Geo, you must use PowerShell to enable these capabilities for all your geo-locations. See the next section for details.

Use PowerShell to enable support for sensitivity labels

As an alternative to using the Microsoft Purview portal, you can enable support for sensitivity labels by using the Set-SPOTenant cmdlet from SharePoint Online PowerShell. The same Set-SPOTenant cmdlet can then be run again to enable sensitivity labels for OneDrive.

If you have Microsoft 365 Multi-Geo, you must use PowerShell to enable this support for all your geo-locations.

Prepare the SharePoint Online Management Shell

Before you run the PowerShell command to enable sensitivity labels for Office files in SharePoint and OneDrive, ensure that you're running SharePoint Online Management Shell version 16.0.19418.12000 or later. To additionally support sensitivity labels for OneNote, you need SharePoint Online Management Shell version 16.0.26914.12004 or later. If you already have the latest version, you can skip to next procedure to run the PowerShell command.

1. If you installed a previous version of the SharePoint Online Management Shell from the PowerShell gallery, update the module by running the following cmdlet.

   ```
   PowerShell
   Update-Module -Name Microsoft.Online.SharePoint.PowerShell
   ```

2. Alternatively, if you have installed a previous version of the SharePoint Online Management Shell from the Microsoft Download Center, you can also go to **Add or remove programs** and uninstall the SharePoint Online Management Shell.

3. In a web browser, go to the Download Center page and Download the latest SharePoint Online Management Shell.

4. Select your language and then select **Download**.

5. Choose between the x64 and x86 .msi file. Download the x64 file if you run the 64-bit version of Windows or the x86 file if you run the 32-bit version. If you don't know, see Which version of Windows operating system am I running?

6. After you download the file, run the file and follow the steps in the Setup configuration.

Run the PowerShell command to enable support for sensitivity labels

To enable the new capabilities, use the Set-SPOTenant cmdlet with the *EnableAIPIntegration* parameter:

1. Using a work or school account that has SharePoint admin privileges in Microsoft 365, connect to SharePoint. To learn how, see Getting started with SharePoint Online Management Shell.

   > **Note**
   >
   > If you have Microsoft 365 Multi-Geo, use the -Url parameter with Connect-SPOService, and specify the SharePoint Online Administration Center site URL for one of your geo-locations.

2. Run the following command and press **Y** to confirm:

   ```
   PowerShell
   ```

```
Set-SPOTenant -EnableAIPIntegration $true
```

3. For Microsoft 365 Multi-Geo: Repeat steps 1 and 2 for each of your remaining geo-locations.

4. If you want to additionally enable sensitivity labels for OneNote (supported at the section level), run *Set-SPOTenant* again, but use the following command:

   PowerShell

   ```
   Set-SPOTenant -EnableSensitivityLabelforOneNote $true
   ```

   This command currently doesn't support Microsoft 365 Multi-Geo.

Publishing and changing sensitivity labels

When you use sensitivity labels with SharePoint and OneDrive, keep in mind that you need to allow for replication time when you publish new sensitivity labels or update existing sensitivity labels. This replication time is especially important for new labels that apply encryption.

For example, you create and publish a new sensitivity label that applies encryption and it very quickly appears in a user's desktop app. The user applies this label to a document and then uploads it to SharePoint or OneDrive. If the label replication hasn't completed for the service, the new capabilities aren't applied to that document on upload. As a result, the document isn't returned in search or for eDiscovery, and the document can't be opened in Office for the web.

For more information about the timing of labels, see When to expect new labels and changes to take effect.

As a safeguard, publish new labels to just a few test users first, wait for at least one hour, and then verify the label behavior on SharePoint and OneDrive. Wait at least a day before you make the label available to more users by either adding more users to the existing label policy, or adding the label to an existing label policy for your standard users. By the time your standard users see the label, it has already synchronized to SharePoint and OneDrive.

SharePoint Information Rights Management (IRM) and sensitivity labels

SharePoint Information Rights Management (IRM) is an older technology to protect files at the list and library level by applying encryption and restrictions when users download files. This older

protection technology is designed to prevent unauthorized users from opening the file while it's outside SharePoint.

In comparison, sensitivity labels provide the protection settings of visual markings (headers, footers, watermarks) in addition to encryption. The encryption settings support the full range of usage rights to restrict what users can do with the content, and the same sensitivity labels are supported for many scenarios. Using the same protection method with consistent settings across workloads and apps results in a consistent protection strategy.

However, you can use both protection solutions together and the behavior is as follows:

- If you upload a file with a sensitivity label that applies encryption, SharePoint can't process the content of these files so coauthoring, eDiscovery, DLP, and search aren't supported for these files.
- If you label a file using Office for the web, any encryption settings from the label are enforced. For these files, coauthoring, eDiscovery, DLP, and search are supported.
- If you download a file that's labeled by using Office for the web, the label is retained and any encryption settings from the label are enforced rather than the IRM restriction settings.
- If you download an Office or PDF file that isn't encrypted with a sensitivity label, IRM settings are applied.
- If you enabled any of the additional IRM library settings, which include preventing users from uploading documents that don't support IRM, these settings are enforced.

With this behavior, you can be assured that all Office and PDF files are protected from unauthorized access if they're downloaded, even if they aren't labeled. However, labeled files that are uploaded don't benefit from the new capabilities.

Search for documents by sensitivity label

Use the managed property **InformationProtectionLabelId** to find all documents in SharePoint or OneDrive that have a specific sensitivity label. Use the following syntax:
`InformationProtectionLabelId:<GUID>`

For example, to search for all documents that are labeled as "Confidential" and that label has a GUID of `8faca7b8-8d20-48a3-8ea2-0f96310a848e`, type the following text in the search box:

```
```

```
InformationProtectionLabelId:8faca7b8-8d20-48a3-8ea2-0f96310a848e
```

Search can't find labeled documents in a compressed file, such as a .zip file.

To get the GUIDs for your sensitivity labels, use the Get-Label cmdlet:

1. First, connect to Office 365 Security & Compliance PowerShell.

   For example, in a PowerShell session that you run as administrator, sign in by using a compliance administrator account.

2. Then run the following command:

   PowerShell

   ```
   Get-Label |ft Name, Guid
   ```

For more information about using managed properties, see Manage the search schema in SharePoint.

Remove encryption for a labeled document

There might be rare occasions when a SharePoint administrator needs to remove encryption from a document stored in SharePoint. Any user who has the Rights Management usage right of Export or Full Control assigned to them for that document can remove encryption that was applied by the Azure Rights Management service from Microsoft Purview Information Protection. For example, users with either of these usage rights can replace a label that applies encryption with a label without encryption. A super user could also download the file and save a local copy without the encryption.

As an alternative, a SharePoint admin can run the Unlock-SPOSensitivityLabelEncryptedFile cmdlet, which removes both the sensitivity label and the encryption. This cmdlet runs even if the admin doesn't have access permissions to the site or file, or if the Azure Rights Management service is unavailable.

For example:

PowerShell

```
Unlock-SPOSensitivityLabelEncryptedFile -FileUrl
"https://contoso.com/sites/Marketing/Shared Documents/Doc1.docx" -JustificationText
"Need to decrypt this file"
```

Requirements:

- SharePoint Online Management Shell version 16.0.20616.12000 or later.
- The encryption is applied by a sensitivity label with admin-defined encryption settings (the Assign permissions now label settings). Double Key Encryption isn't supported for this cmdlet.

The justification text is added to the audit event of **Removed sensitivity label from file**, and the decryption action is also recorded in the usage logging for the Azure Rights Management service.

How to disable sensitivity labels for SharePoint and OneDrive (opt-out)

After you disable these new capabilities, SharePoint and OneDrive can't process newly labeled and encrypted files. As a result, coauthoring, eDiscovery, data loss prevention, search, and other collaborative features won't work for these files. However, these capabilities continue to be supported for files that were labeled and encrypted before you disabled sensitivity labels.

This mixed protection state can result in higher administrative overheads with help desk incidents to investigate the inconsistent behavior. It might also be a problem for compliance requirements.

To better understand the behavior of labeled files before and after disabling sensitivity labels for SharePoint and OneDrive, use the **Label behavior** tab.

Disable instructions

To disable these new capabilities, you must use PowerShell. By using the SharePoint Online Management Shell and the Set-SPOTenant cmdlet, specify the same *EnableAIPIntegration* parameter as described in the Use PowerShell to enable support for sensitivity labels section. But this time, set the parameter value to false and press **Y** to confirm:

PowerShell

```
Set-SPOTenant -EnableAIPIntegration $false
```

If you have Microsoft 365 Multi-Geo, you must run this command for each of your geo-locations.

Next steps

After you enable sensitivity labels for files in SharePoint and OneDrive, consider automatically labeling files by using either, or both, of the following labeling methods:

- Applying a default sensitivity label for document libraries, for new and edited files in SharePoint. For more information, see Configure a default sensitivity label for a SharePoint document library.
- Auto-labeling policies that use content inspection for files in SharePoint and OneDrive. For more information, see Automatically apply a sensitivity label to Microsoft 365 data.

Do you need to share your labeled and encrypted documents with people outside your organization? See Sharing encrypted documents with external users.

- Learn about Double Key Encryption

Last updated on 08/07/2026

Use sensitivity labels with Microsoft Loop

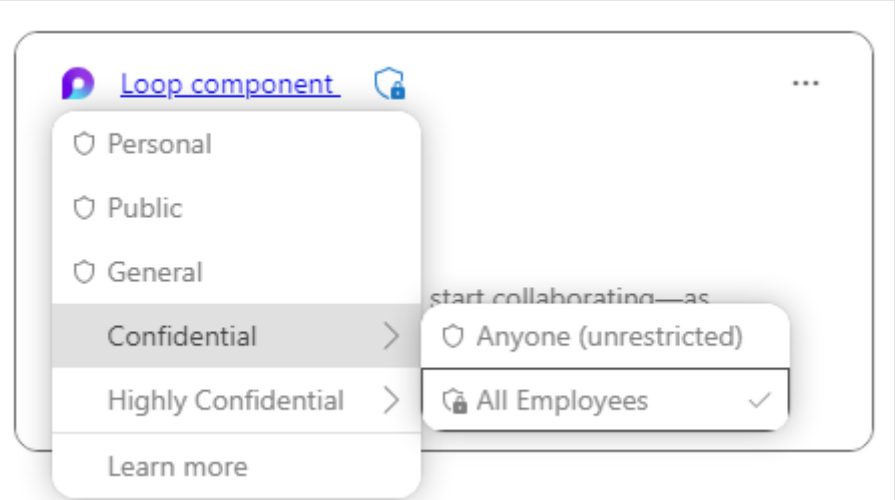
> *Microsoft Purview service description*

Use your sensitivity labels with Microsoft Loop to add an extra layer of protection to the Loop-related data that users create with this app. This protection extends to Microsoft 365 Copilot.

Loop supports sensitivity labels at the file level, to help protect components and pages:

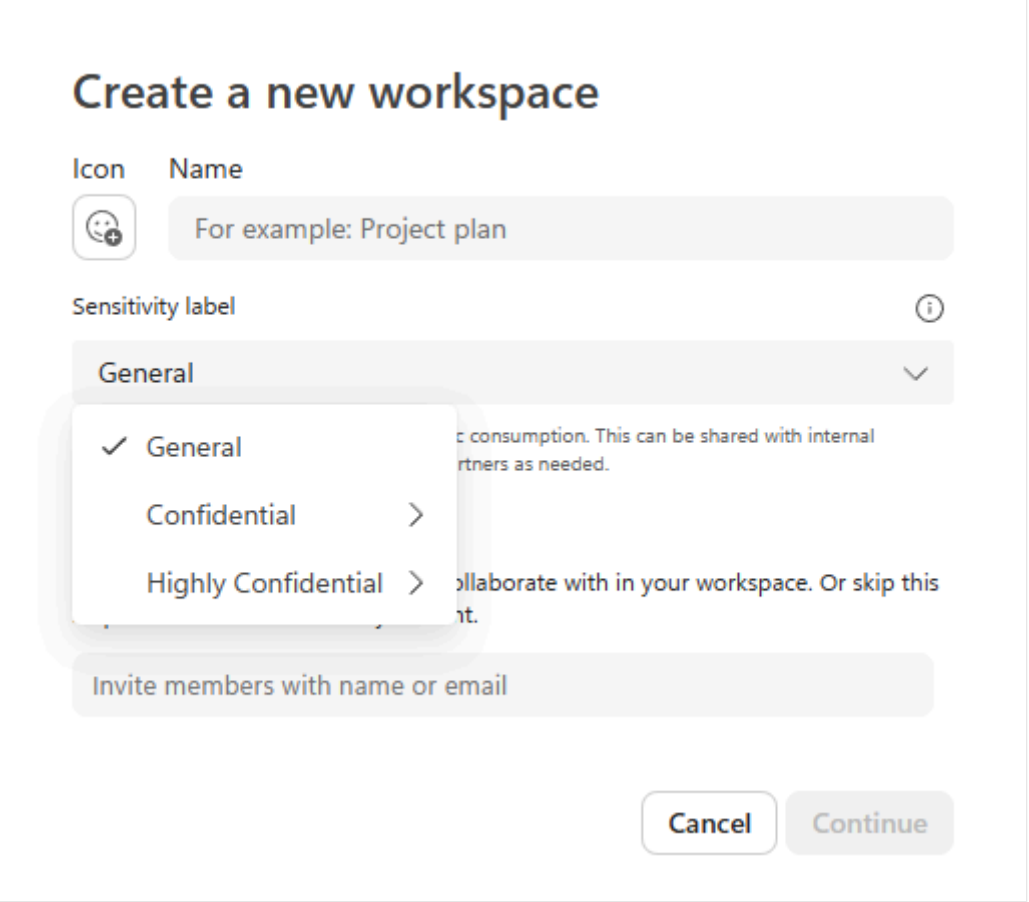
- Users can manually apply sensitivity labels that are published to them with the **Files & other data assets** label scope, including sensitivity labels that apply encryption configured to apply permissions now.
- Watermarks as a label setting are supported but not variables in the text string.
- Mandatory labeling and a default document label are supported as settings from a label publishing policy.
- Other policy settings supported:
  - Justification to remove a label or lower its priority
  - The "Learn more" link

Example Loop component that's currently labeled with the **Confidential \ All Employees** sensitivity label that applies encryption, displaying other available labels:



Loop supports sensitivity labels at the container level, to help protect access to workspaces:

- This functionality is supported for tenant-owned workspaces on the web version of the Loop app.
- You can apply a sensitivity label when you create a new workspace, or apply or change a sensitivity label as a More (...) workspace configuration.
- Using the SharePoint admin center with the SharePoint Embedded administrator role, you can see and change applied sensitivity labels from **Containers** > **Active containers**.



See the following sections for more details about requirements, supported scenarios, and limitations when you use sensitivity labels with Loop.

For information where Loop-related content is stored, see Overview of Loop storage.

Sensitivity labels to protect Loop components and pages

The ability to apply sensitivity labels to Loop components and pages requires that SharePoint and OneDrive is enabled for sensitivity labels. Then, you can apply sensitivity labels to loop components and pages in the following apps:

- The Loop app (web and mobile)
- Loop in Outlook (desktop, mobile, web)
- Microsoft 365 Copilot Chat in Microsoft Copilot Pages

When Loop components and pages are encrypted by sensitivity labels, Microsoft 365 Copilot honors the EXTRACT usage right. Microsoft 365 Copilot won't summarize this Loop data unless the user is granted this usage right.

As with sensitivity labels that are applied to documents in SharePoint and OneDrive, the following auditing events are supported:

- **Applied sensitivity label to file**
- **Changed sensitivity label applied to file**
- **Removed sensitivity label from file**

Sensitivity label inheritance for components and pages

Any unlabeled Loop components on the page inherit the sensitivity label applied to the page. That label is retained if you move it to another Loop page.

A Loop page can inherit a higher priority sensitivity label from Copilot in the Loop app when referenced files are labeled. If more than one file is referenced and labeled, the page inherits the label with the highest priority.

Limitations for Loop components and pages

The following labeling capabilities aren't supported:

- Label names and descriptions for other languages
- Headers and footers
- Variables in watermarks and dynamic watermarks
- Automatic and recommended labeling
- Sensitivity labels that are configured for any of the following encryption settings:
  - **Let users assign permissions**, sometimes known as "user-defined permissions"
  - **User access to content expires** is set to a value other than **Never**
  - **Double Key Encryption** is selected

Because the task management tools in Microsoft Planner and Microsoft To Do don't support encryption, they stop synchronizing a Loop component or page that's encrypted by a sensitivity label.

Sensitivity labels to protect Loop workspaces

The ability to apply sensitivity labels to Loop workspaces at the tenant-level requires that sensitivity labels are enabled for containers and that the label scope includes **Groups & sites**. Then, you can apply sensitivity labels to Loop workspaces in the Loop web app so that content in those workspaces is protected by the following label settings documented in Use sensitivity labels to protect collaborative workspaces (groups and sites):

- External user access
- External sharing from SharePoint sites
- Access from unmanaged devices
- Authentication contexts

- Default sharing link for a SharePoint site (PowerShell-only configuration)
- Site sharing settings (PowerShell-only configuration)

Although Loop workspaces are stored in SharePoint Embedded containers rather than SharePoint sites, the listed sensitivity label settings for sites are functionally the same.

Limitations for Loop workspaces

As with sensitivity labels that you apply to Loop components and pages, sensitivity labels for Loop workspaces don't support label colors or label names and descriptions for other languages.

The following Loop content isn't protected by the container label settings:

- Linked files
- Files uploaded to a workspace because these files are stored in the user's OneDrive
- Videos that are created in a workspace because these files are stored in the user's OneDrive

---

Last updated on 03/31/2026

Configure a default sensitivity label for a SharePoint document library

06/10/2025

> *Microsoft 365 licensing guidance for security & compliance*

When SharePoint is enabled for sensitivity labels, you can configure a default label for document libraries. Then, any new files uploaded to that library, or existing files edited in the library will have that label applied if they don't already have a sensitivity label, or they have a sensitivity label but with lower priority.

For example, you configure the **Confidential** label as the default sensitivity label for a document library. A user who has **General** as their policy default label saves a new file in that library. SharePoint will label this file as **Confidential** because of that label's higher priority. For a quick summary of the possible outcomes, see Will an existing label be overridden on this page.

A default label offers a baseline level of protection and a form of automatic labeling without content inspection. To help you distinguish between this feature's default label with the default label in label policies:

- **Default sensitivity label for a document library**: Location-based labeling, applicable only for SharePoint. Overrides a lower-priority label unless manually applied.
- **Default sensitivity label from a policy**: Always applicable for all locations. Never overrides an existing label.

When you use Office for the web to create or edit a file, the default sensitivity label for a document library can be applied without delays. However, labeling is not immediate if you upload a file or create it using Microsoft 365 Apps on Windows, macOS, iOS or Android, and then save to SharePoint:

- File upload: it can take a few minutes for the label to be applied.
- Microsoft 365 Apps: the label is applied after the app is closed.

To read the preview announcement for this feature, see the blog post.

Tip

If you're not an E5 customer, use the 90-day Microsoft Purview solutions trial to explore how additional Purview capabilities can help your organization manage data security and

compliance needs. Start now at the Microsoft Purview trials hub. Learn details about signing up and trial terms.

Will an existing label be overridden?

Summary of outcomes:

Expand table

Existing label	Override with library default label
Manually applied, any priority	No
Automatically applied, lower priority	Yes
Automatically applied, higher priority	No
Default label from policy, lower priority	Yes
Default label from policy, higher priority	No

Requirements

- You've created and published sensitivity labels that include the label scope of **Files & other data assets** and these labels are published to the users who will select a default sensitivity label for a SharePoint document library.
- You've enabled sensitivity labels for Office files in SharePoint and OneDrive. To check this status, you can run `(Get-SPOTenant).EnableAIPIntegration` from the SharePoint Online Management Shell to confirm the value is set to **True**.
- To support sensitivity labels for PDFs, you've added support for PDFs in SharePoint. To check this status, you can run `(Get-SPOTenant).EnableSensitivityLabelforPDF` from the SharePoint Online Management Shell to confirm the value is set to **True**.
- SharePoint Information Rights Management (IRM) is not enabled for the library. This older technology isn't compatible with using a default sensitivity label for a SharePoint document library. If a library is enabled for IRM, you won't be able to select a default sensitivity label.
- Site admin permissions are needed to apply and change the sensitivity label in SharePoint.

- Files must contain content to be labeled. Empty files are labeled when they are updated with content.

If you need to review a list of file types that are supported by sensitivity labels in SharePoint, see Supported file types.

Limitations

- Doesn't apply to existing files at rest in SharePoint.
- Unless you've enabled co-authoring for files encrypted with sensitivity labels, you'll see a delay in applying the default sensitivity label for a document library when users select the **File** > **Save as** option.
- As with sensitivity labels for Office for the web, some label configurations that apply encryption aren't suitable for SharePoint, and so don't support a default sensitivity label for a SharePoint document library:
  - **Let users assign permissions when they apply the label** and the checkbox **In Word, PowerPoint, and Excel, prompt users to specify permissions** is selected. This setting is sometimes referred to as "user-defined permissions".
  - **User access to content expires** is set to a value other than **Never**.
  - **Double Key Encryption** is selected.

How to configure a default sensitivity label for a SharePoint document library

This configuration is done by a SharePoint site admin: Add a sensitivity label to SharePoint document library.

> **Important**
>
> From the **Library settings** pane, select a sensitivity label and don't select the checkbox **Extend protection on download, copy, or move**. This setting is currently in preview and supports a different labeling scenario. For more information, see Configure SharePoint with a sensitivity label to extend permissions to downloaded documents.

Monitoring application of library default sensitivity labels

Use the SharePoint **Sensitivity** column to see the names of sensitivity labels applied to files. When the label has been applied by this features, the tooltip for the label name displays **This file has been automatically labeled**. However, this tooltip isn't exclusive to the default sensitivity label for a document library. It also displays when sensitivity labels are applied by using auto-labeling policies or as a result of a user's default label from sensitivity label policies.

To specifically identify when the label was applied because of the library's default sensitivity label, use the audit log search tool and the **Applied sensitivity label file** auditing event from the **Sensitivity label activities** group. Then:

1. Select an entry to view the details in a flyout pane.
2. From the details pane, scroll to the **SensitivityLabelEventData section**, and identify the value for **ActionScourceDetails**.
3. A value of **6** is used for when the label was applied because of the default sensitivity label for the document library.

To audit the configuration setting for this feature, use the **Updated list** auditing event from the **SharePoint list activities** group. In the details flyout pane for the document library, scroll to the **SensitivityLabelEventData** section where **OldSensitivityLabelId** and **SensitivityLabelId** can reflect three changes of states:

- Sensitivity label applied
- Sensitivity label changed from one label to another
- Sensitivity label removed

To map sensitivity label GUIDs to label names, use the Get-Label cmdlet:

1. First, connect to Security & Compliance PowerShell.
2. Then run the following command, where you specify the GUID:

```
PowerShell

Get-Label -Identity "<GUID>" | Name
```

How to turn off this feature

If you need to, you can turn off this feature that supports a default sensitivity label for SharePoint document libraries. This is a tenant-level setting that requires you to use the Set-SPOTenant cmdlet with the *DisableDocumentLibraryDefaultLabeling* parameter set to **True** by using the current SharePoint Online Management Shell:

PowerShell

```
Set-SpoTenant -DisableDocumentLibraryDefaultLabeling $true
```

After you run this command, you won't see the option to configure **Default sensitivity labels** for a document library. You won't be able to select a default sensitivity label for new or existing libraries.

For document libraries that were previously configured for a default sensitivity label:

- The label selection for the document library remains but is deactivated so new files won't have the selected sensitivity label applied.
- Sensitivity labels that were applied as a default label aren't removed.

As with all tenant-level configuration changes for SharePoint, it takes about 15 minutes for the change to take effect.

Next steps

Default labeling ensures a minimum level of protection but doesn't take into account the file contents that might require a higher level of protection. Consider supplementing this labeling method with automatic labeling that uses content inspection, and encourage manual labeling for users to replace the default label when needed.

Configure SharePoint with a sensitivity label to extend permissions to downloaded documents

Summarize this article for me

> *Microsoft Purview service description*

When SharePoint is enabled for sensitivity labels, you can configure document libraries to extend the existing SharePoint site permissions to documents when they're downloaded from the library. Then, any previously unlabeled files from that library continue to be protected with the current SharePoint permissions for the user, even though the files have left the original SharePoint boundary.

For example, you select the **Confidential \ Trusted People** label as the sensitivity label to extend permissions for a document library. In that library, a site is configured to allow only read permissions to a specfiic group of users. A user from that group downloads a file from the site, and that file is now labeled as **Confidential \ Trusted People**. Even though the file is no longer in SharePoint, that user can still only view and not edit the content or remove the label. Only users with access to the file in SharePoint have access to the downloaded file. Users without access to the file in SharePoint won't be able to open that downloaded file, no matter where it's stored.

Additionally, the file has a just-in-time layer of protection because the user who downloaded the file also won't be able to open it under the following circumstances:

- Permissions for that user are removed from the file
- The file is deleted from the site
- The site is no longer active
- The file is moved to a different site

If the SharePoint permissions for the user are changed, those changes are reflected for the downloaded file.

When these labeled files are in the document library, they're also protected from copy and move actions:

- The files can't be copied to or moved to a different site.
- The files can be copied to or moved to a different document library within the same site only if the user has SharePoint permissions to create or delete lists.

The specified sensitivity label is applied to all files that are unlabeled, and to files that are labeled but the label configuration doesn't apply encryption. Files that are synchronized with OneDrive also get labeled.

For existing files that are labeled but not encrypted, a manually applied label is also replaced with the specified label. For a quick summary of the possible outcomes, see Will an existing label be overridden on this page.

Because unencrypted files can be relabeled, this labeling configuration is well-suited to organizations that are early in their labeling deployment and haven't yet labeled files in SharePoint sites by using other methods.

Will an existing label be overridden?

Summary of outcomes:

Expand table

Existing label	Override with library label to extend permissions
Label applied using any methods (manual, automatic, default label from policy) and label configuration doesn't apply encryption, any priority	Yes
Label applied using any methods (manual, automatic, default label from policy) and label configuration does apply encryption, any priority	No

Requirements

- You've created and published sensitivity labels, published to the users who will select the sensitivity label for a SharePoint document library. Labels require the following configuration:
  - The label scope of **Files and other data assets**
  - Access control is selected with the encryption setting of **Let users assign permissions when they apply the label** and the checkbox **In Word, PowerPoint, and Excel, prompt users to specify permissions** is selected. This setting is sometimes referred to as "user-defined permissions."

    > **Note**

> In this application of the label, users won't be prompted for permissions but their SharePoint permissions will be automatically applied when a file is downloaded, copied, or moved from the site.

- You've enabled sensitivity labels for Office files in SharePoint and OneDrive. To check this status, you can run `(Get-SPOTenant).EnableAIPIntegration` from the SharePoint Online Management Shell to confirm the value is set to **True**.
- Your tenant has been enabled for co-authoring for files encrypted with sensitivity labels.
- To support sensitivity labels for PDFs, you've added support for PDFs in SharePoint. To check this status, you can run `(Get-SPOTenant).EnableSensitivityLabelforPDF` from the SharePoint Online Management Shell to confirm the value is set to **True**.
- Windows apps from Microsoft 365 Apps for enterprise that download files require a minimum version of 2402 from Current Channel, Monthly Enterprise Channel, or Semi-Annual Enterprise Channel.
- SharePoint Information Rights Management (IRM) isn't enabled for the library. This older technology isn't compatible with using a sensitivity label for a SharePoint document library. If a library is enabled for IRM, you won't be able to select a sensitivity label.
- Site admin permissions are needed to apply and change the sensitivity label in SharePoint.
- Files must contain content to be labeled.

If you need to review a list of file types that are supported by sensitivity labels in SharePoint, see Supported file types.

Mapping of SharePoint permissions to usage rights

Use the following table to understand how the SharePoint permissions for a user are extended to rights management usage rights and permission levels when a file is downloaded, or copied and moved outside the site.

Expand table

SharePoint permission	Usage rights applied	Permission levels
Owner	Full control over the content, all permissions, and the sensitivity label applied:	Owner

SharePoint permission	Usage rights applied	Permission levels
	VIEW, EXTRACT, DOCEDIT, EDIT, EXPORT, COMMENT, PRINT, FORWARD, REPLY, REPLYALL, VIEWRIGHTSDATA, EDITRIGHTSDATA, OBJMODEL *, OWNER *	
Edit	Full control over the content but can't change the sensitivity label applied: VIEW, EXTRACT, DOCEDIT, EDIT, EXPORT, COMMENT, PRINT, FORWARD, REPLY, REPLYALL, VIEWRIGHTSDATA	Editor
Read	Can view and copy the content but can't change the content or the sensitivity label applied: VIEW, VIEWRIGHTSDATA, EXTRACT	No equivalent permission level, Viewer is the closest

* Currently, these usage rights aren't applied and as a result, the user can change the sensitivity label, but can't remove it.

Limitations

The following limitations apply when you use this configuration:

- Users can't manually apply sensitivity labels that aren't configured to apply encryption.
- Users won't be able to open downloaded files offline; they must be able to connect to the original site.
- Users won't be able to open downloaded files if the original SharePoint site, folder, or file is deleted.
- Files labeled with this configuration can't be moved or copied to another site.
- Files labeled with this configuration don't support the **Save As** option, and won't save with this option unless they are the minimum versions that include client-side improvements:
  - Windows: Version 2601+ in Current Channel and Monthly Enterprise Channel
  - macOS: Version 16.103+
  - iOS: Version 2.103+
  - Android: Version 16.0.19426+
- Files labeled with this configuration can only be moved or copied to another document library within the same site if users have SharePoint permissions to create or delete lists. The label isn't retained for the copied or moved file.

- This configuration can override a previously manually applied label if the label isn't configured to apply encryption.
- Files labeled with this configuration currently aren't displayed as labeled in content explorer.
- Microsoft 365 Copilot can reference the labeled files if users have SharePoint read permissions to them, but won't summarize these files. Because files can't be summarized, they also can't be used by Copilot to generate new content.

> **Note**
>
> As with all files that are encrypted with Azure Rights Management, a **super user** can open encrypted documents if that becomes necessary because the original location is no longer accessible.

How to configure a SharePoint document library for a sensitivity label that extends permissions to downloaded documents

This configuration first requires that you enable the capability for the tenant by using PowerShell with the SharePoint Online Management Shell. Then, a new checkbox becomes available for document library settings.

Run the PowerShell command to enable support to extend SharePoint permissions

Ensure that you're running SharePoint Online Management Shell version 16.0.25430.12000 or later.

To enable the new capabilities, use the Set-SPOTenant cmdlet with the *ExtendPermissionsToUnprotectedFiles* parameter:

1. Using a work or school account that has SharePoint admin privileges in Microsoft 365, connect to SharePoint. To learn how, see Getting started with SharePoint Online Management Shell.

   > **Note**

If you have Microsoft 365 Multi-Geo, use the -Url parameter with **Connect-SPOService**, and specify the SharePoint Online Administration Center site URL for one of your geo-locations.

2. Run the following command and press **Y** to confirm:

```
PowerShell
Set-SPOTenant -ExtendPermissionsToUnprotectedFiles $true
```

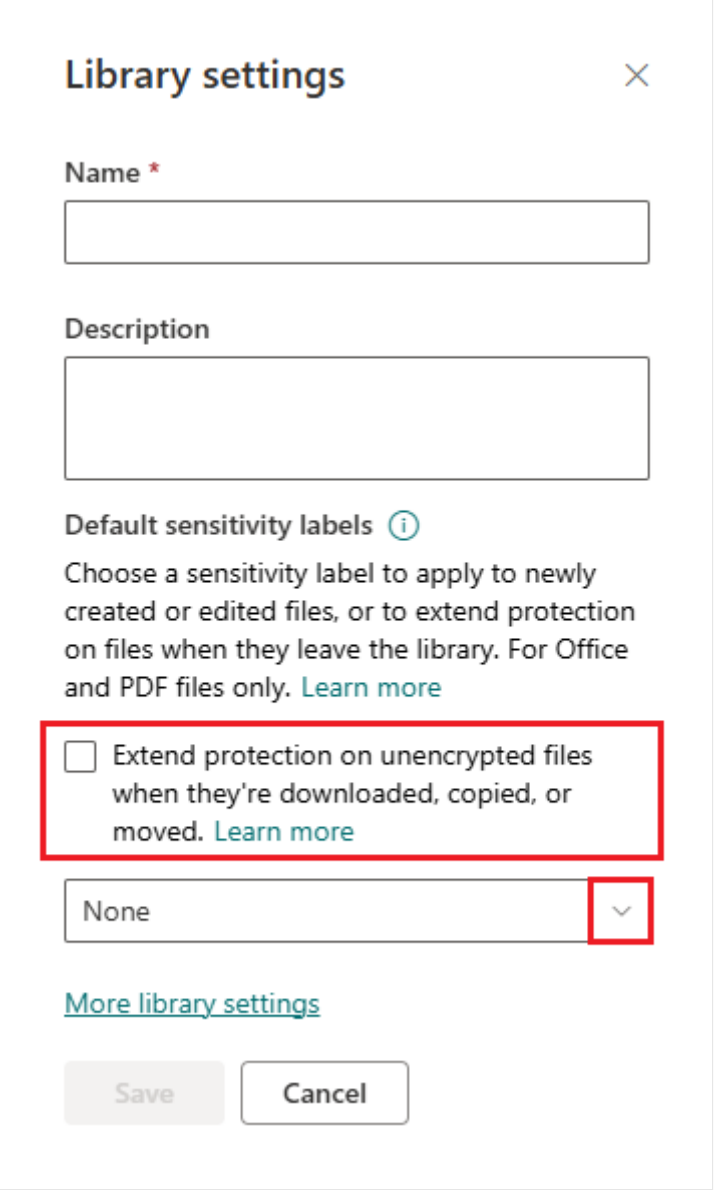
   If you see an error informing you that the feature isn't enabled, it's most likely because the rollout hasn't reached your tenant yet.

3. For Microsoft 365 Multi-Geo: Repeat steps 1 and 2 for each of your remaining geo-locations.

As with all tenant-level configuration changes for SharePoint, it takes about 15 minutes for the change to take effect.

Configure SharePoint document libraries with a default label to extend permissions

For each SharePoint document library that you want to have this configuration, follow the instructions for Add a sensitivity label to SharePoint document library, and select the checkbox **Extend protection on download, copy or move**:



You won't see this checkbox until the previous PowerShell command has completed.

After you've selected a sensitivity label that applies encryption with user-defined permissions, save the configuration.

> **Note**
>
> This feature is mutually exclusive with the option to select a default sensitivity label for a SharePoint document library that supports sensitivity labels without encryption, and sensitivity labels that are configured with the option **Assign permissions now** (sometimes referred to as "admin-defined permissions").

The selected sensitivity label will be applied to all files that are unlabeled, and files that are labeled but the label configuration doesn't apply encryption. The **Sensitivity** column for the document library displays your selected label for existing, new, and edited files. Users see the selected label displayed when they open the file for editing but won't experience any changes in permissions as a result of the label.

After the library is configured with the sensitivity label, all existing files will be resynchronized if the library is synced via the OneDrive sync client. The resynchronization process can take a while and until it's complete, the extended protection won't be applied.

> **Important**
>
> In the SharePoint document library that you've configured for the sensitivity label, users can't remove the applied sensitivity label in their Office apps and can change it only if the replacement label applies encryption.

Monitoring application of the sensitivity label that extends SharePoint permissions

Because this configuration extends the functionality of a default sensitivity label for a document library, you monitor its configuration in the same way. The sensitivity label GUID will identify the encryption configuration for user-defined permissions. There's no separate labeling auditing event for when a file is downloaded, copied, or moved.

How to turn off this feature

If you want a specific SharePoint document library to no longer extend permissions with a sensitivity label, clear the **Extend protection on download, copy or move** checkbox as a SharePoint document library setting. Then:

- Files in the document library that were previously labeled with this configuration will revert to either their original label that didn't apply encryption, or unlabeled if that was their original status before the checkbox was selected.
- Files synchronized with OneDrive will be resynchronized and similarly revert to their previous label status. The resynchronization process can take a while and until it's complete, the extended protection will remain.
- Labeled files that are now downloaded retain their label.

If you want to turn off this feature at the tenant-level when it has been previously enabled and configured, first clear the **Extend protection on download, copy or move** checkbox from all document libraries where it's been selected. Then, use the same Set-SPOTenant cmdlet with the *ExtendPermissionsToUnprotectedFiles* parameter, but set the value to **false**. Then:

- The **Extend protection on download, copy or move** checkbox is no longer visible as a SharePoint document library setting.

If you turn off the feature at the tenant level without first clearing the checkbox for a document library, the configuration remains but without the checkbox to turn it off.

In this scenario, to turn off the configuration, run the PowerShell command to enable support to extend SharePoint permissions to display the checkbox again, so you can clear it. If you don't want to reenable support at the tenant level, contact Microsoft Support to talk you through PowerShell commands to remove the configuration for a specific document library.

Next steps

Although this configuration provides immediate protection at scale for files stored in SharePoint, it doesn't take into account the file contents that might require a higher level of protection. Consider supplementing this labeling method with automatic labeling that uses content inspection to apply sensitivity labels with encryption.

Related resources

Microsoft Syntex Advanced Management overview

Last updated on 02/23/2026

Enable co-authoring for files encrypted with sensitivity labels

> *Microsoft 365 licensing guidance for security & compliance*.

Enable the setting to support co-authoring for Office desktop apps so that when documents are labeled and encrypted by sensitivity labels, multiple users can edit these documents at the same time.

Without this setting enabled for your tenant, users must check out an encrypted document stored in SharePoint or OneDrive when they use Office desktop apps. As a result, they can't collaborate in real time. Or, they must use Office for the web when sensitivity labels are enabled for Office files in SharePoint and OneDrive.

In addition, enabling this functionality results in the AutoSave functionality being supported for these labeled and encrypted files.

For the release announcement, see the blog post Co-authoring on Microsoft Information Protection encrypted documents is now generally available.

Metadata changes for sensitivity labels

> **Important**
>
> After you enable the setting for co-authoring, labeling information for unencrypted files is no longer saved in custom properties.
>
> Don't enable this setting if you use any apps, services, scripts, or tools that read or write labeling metadata to the old location.

Before you enable the setting to support co-authoring for Office desktop apps, understand that this action changes the labeling metadata that Office files save and read.

The labeling metadata includes information that identifies your tenant and applied sensitivity label. The change that this setting makes is the metadata format and location for Word, Excel, and PowerPoint files. You don't need to take any action for encrypted files or emails because the metadata change for encrypted files is backward-compatible, and there are no changes for emails. However, you do need to be aware of the metadata changes for encrypted files that can be automatically upgraded but aren't backward-compatible.

This change affects both newly labeled files and already labeled files. When you use apps and services that support the co-authoring setting:

- For files that are newly labeled, only the new format and location is used for the labeling metadata.
- For files that are already labeled, the next time the file is opened and saved, if the file has metadata in the old format and location, that information is copied to the new format and location.

For more information about this metadata change, see the following resources:

- Blog post: Upcoming Changes to Microsoft Information Protection Metadata Storage
- Open Specifications: 2.6.3 LabelInfo versus Custom Document Properties

Because of these changes, don't enable this setting if your organization has any apps, services, scripts, or tools that read or write labeling metadata to the old location. If you enable this setting, some example consequences include:

- A document that is labeled appears to users as unlabeled.
- A document displays an out-of-date label to users.
- Co-authoring and AutoSave don't work for a labeled and encrypted document if another user has it open in a Office desktop app that doesn't support the new labeling metadata. Be aware that this scenario can also occur for users outside your organization if external users and invited guests have the file open.
- An Exchange Online mail flow rule that identifies labels as custom properties in Office attachments fails to encrypt the email and attachment, or incorrectly encrypts them.

For a list of apps and services that support this setting and the changes to the labeling metadata, see the following section.

Prerequisites

Make sure you understand the following prerequisites before you turn on this feature.

- To turn this capability on or off, you need at least the DLP Compliance Management or Information Protection Admin role. Alternative roles include the Security Administrator and Compliance Administrator roles.
- Sensitivity labels must be enabled for Office files in SharePoint and OneDrive for the tenant. If this feature isn't already enabled, it is automatically enabled when you select the setting to turn on co-authoring for files with sensitivity labels.

- Microsoft 365 Apps for enterprise:
  - **Windows**: Minimum version 2107 from Current Channel or Monthly Enterprise Channel, or minimum version 2202 from Semi-Annual Enterprise Channel
  - **macOS**: Minimum version 16.51
  - **iOS**: Minimum version 2.58
  - **Android**: Minimum version 16.0.14931

- All apps, services, and operational tools in your tenant must support the new labeling metadata. If you use any of the following apps or services, check the minimum versions required:

  - **Microsoft Purview Information Protection client and scanner:**
    - Minimum version of 3.0 that you can install from the Microsoft Download Center, or minimum version of 2.12.62.0 for the older Azure Information Protection unified labeling client and scanner

  - **OneDrive sync app for Windows or macOS:**
    - Minimum version of 19.002.0121.0008

  - **Endpoint data loss prevention (Endpoint DLP):**
    - Windows 10 1809 with KB 4601383
    - Windows 10 1903 and 1909 with KB 4601380
    - Windows 10 2004 with KB 4601382
    - Windows versions later than Windows 10 2004 are supported without KB updates

  - **Apps and services that use the Microsoft Information Protection SDK:**
    - Minimum version of 1.7

Microsoft 365 services automatically support the new labeling metadata when you turn on this feature. For example:

- Auto-labeling policies
- DLP policies that use sensitivity labels as conditions
- Microsoft Defender for Cloud Apps configured to apply sensitivity labels

Limitations

Before you enable the tenant setting for co-authoring for files encrypted with sensitivity labels, make sure you understand the following limitations of this feature.

- Because of the labeling metadata changes, all apps, services, and operational tools in your tenant must support the new labeling metadata for a consistent and reliable labeling experience.

Specific to Excel: Metadata for a sensitivity label that doesn't apply encryption can be deleted from a file if somebody edits and saves that file by using a version of Excel that doesn't support the metadata changes for sensitivity labels.

- Co-authoring and AutoSave aren't supported for Office documents that use the label encryption configuration **User access to content expires** when it's set to a value other than **Never**, or **Double Key Encryption** is configured.

  Labels with these configurations do still display in Office apps. However, when users select these labels and nobody else is editing the document, they're warned that co-authoring and AutoSave won't be available. If somebody else is editing the document, users see a message that the labels can't be applied.

- Co-authoring and AutoSave are supported with specific versions and limitations for sensitivity labels that are configured for user-defined permissions. This encryption configuration refers to the setting **Let users assign permissions when they apply the label** and the checkbox **In Word, PowerPoint, and Excel, prompt users to specify permissions** is selected.

How to enable co-authoring for files with sensitivity labels

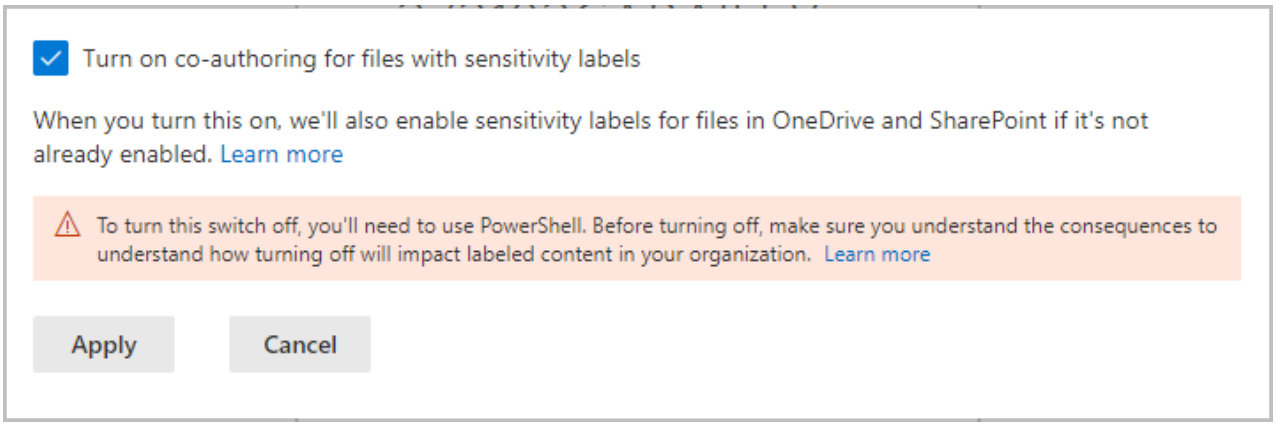
> ⊗ **Caution**
>
> Enable this option only after you read and understand the metadata changes, prerequisites, limitations, and any known issues documented on this page.

For this configuration, use an account that has at least the DLP Compliance Management or Information Protection Admin role.

1. Sign in to the Microsoft Purview portal > **Settings** > **Solution settings** > **Information Protection** > **Co-authoring for files with sensitivity labels**.

2. On the **Co-authoring for files with sensitivity labels** page, read the summary description, prerequisites, and what to expect.

   Then select **Turn on co-authoring for files with sensitivity labels**, and **Apply**:



3. Wait 24 hours for this setting to replicate across your environment before you use this new feature for co-authoring.

If you need to disable this feature

> **Important**
>
> If you do need to disable this feature, for example, you've discovered some apps don't support the metadata changes and you can't immediately update these apps, be aware that labeling information can be lost.

After you enable co-authoring for files with sensitivity labels for your tenant, you can't disable this setting in the Microsoft Purview portal. This action is supported only by using PowerShell. This setting isn't one that you casually disable. Check and understand the prerequisites, consequences, and limitations before you enable the setting.

If you do disable this setting for your tenant:

- For apps and services that support the new labeling metadata, they now revert to the original metadata format and location when labels are read or saved.
- The new metadata format and location for Office documents that were used while the setting was enabled won't be copied to the original format and location. As a result, this labeling information for unencrypted Word, Excel, and PowerPoint files will be lost.
- Co-authoring and AutoSave no longer work in your tenant for labeled and encrypted documents.
- Sensitivity labels remain enabled for Office files in OneDrive and SharePoint.

To disable co-authoring for your tenant

Use the Set-PolicyConfig cmdlet with the *EnableLabelCoauth* parameter.

1. In a PowerShell session, connect to Security & Compliance PowerShell.
2. Run the following command:

```
PowerShell
Set-PolicyConfig -EnableLabelCoauth:$false
```

The command completes without a prompt or confirmation.

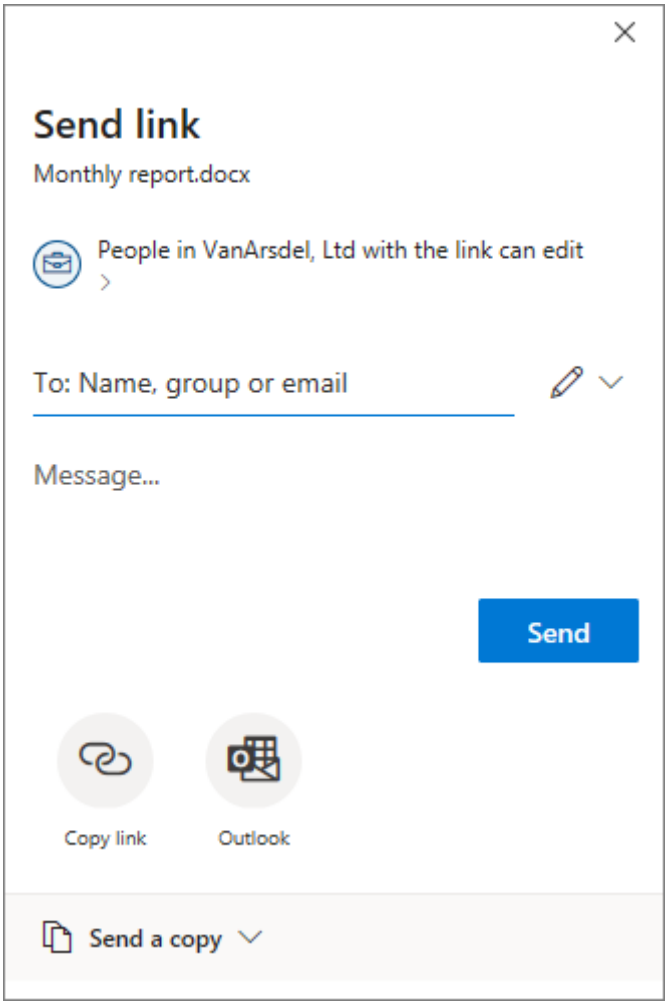
Last updated on 01/14/2026

Use sensitivity labels to configure the default sharing link type for sites and documents in SharePoint and OneDrive

09/26/2025

> *Microsoft 365 licensing guidance for security & compliance*.

As an additional configuration to the settings you see in the Microsoft Purview portal for sensitivity labels, you can use these labels to configure settings for the default sharing link type for a SharePoint site or OneDrive account, and for individual documents. These settings are automatically selected, but not highly visible to users when they select the **Share** button in their Office apps. As an example:



The default sharing link type sets the scope (who) and permissions (view or edit) that are automatically selected when users share files and folders. Although users can always override these default settings before sending the sharing link, the settings you choose provide a safe baseline. Typically, users don't change the settings before sharing.

At the site level (SharePoint site or OneDrive account), sensitivity labels provide a convenient alternative for setting the default sharing link type that can be configured for a site in the SharePoint admin center. For more information, see Change the default link type for a site from the SharePoint documentation.

This site-level configuration works well for SharePoint sites that have documents all with the same level of sensitivity. But if sites contain some documents that have a higher level of sensitivity that require more restrictive settings, you can configure a sensitivity label with different settings for the default sharing link type, and then apply this label to documents.

In this scenario where the site has default sharing link type settings, and a document in that site has different default link type settings, the more restrictive scope settings will be applied at the time the user selects the sharing option for the document. For example:

- The default sharing link type for the site is scoped to anybody in your organization. A document in that site is labeled with the default sharing link type set to specific people. When a user shares that document, the default sharing link type selected will be scoped to specific people.

- The default sharing link type for the site is scoped to specific people, with edit permissions. A document in that site is labeled with the default sharing link type set to anybody in the organization, with view permissions. When a user shares that document, the default sharing link type selected will be scoped to specific people with edit permissions.

Configuring the default link type for documents might also be appropriate without the site-level setting. For example, although SharePoint sites are typically organized to host the same type of documents, that isn't the case for OneDrive accounts. Users typically save a wide range of files to OneDrive, often including a mix of personal and business documents. Setting a default link type for all documents for a user's OneDrive account is probably not practical, but individual documents can still benefit from these settings. For example:

- Documents labeled **Highly Confidential** have a default sharing link type that restricts sharing to specific people rather than anybody in the organization.
- Documents labeled **General** have a default sharing link type that restricts sharing to people in your organization.
- Documents labeled **Personal** have a default sharing link type that allows sharing to anyone with the link.

Prerequisites

To apply the default sharing link type for sites, sensitivity labels must be enabled for containers. If this capability isn't yet enabled for your tenant, see How to enable sensitivity labels for containers and synchronize labels.

To apply the default sharing link type for documents in SharePoint and OneDrive, sensitivity labels must be enabled for these services. If this capability isn't yet enabled for your tenant, see How to enable sensitivity labels for SharePoint and OneDrive (opt-in).

In a PowerShell session, you must connect to Security & Compliance PowerShell to configure the settings for the default sharing link type.

> **Note**
>
> Although not required, it's easiest to first create and configure sensitivity labels in the Microsoft Purview portal, and then modify these labels with the settings that configure the default sharing link type.

How to configure settings for the default sharing link type

The configuration settings for the default sharing link type use the PowerShell *AdvancedSettings* parameter with the Set-Label and New-Label cmdlets from Security & Compliance PowerShell:

- **DefaultSharingScope**: The available values are:
  - **SpecificPeople**: Sets the default sharing link to specific people (only the people the user specifies)
  - **Organization**: Sets the default sharing link for only people in your organization
  - **Anyone**: Sets the default sharing link to anyone with the link, which is equivalent to anonymous access

- **DefaultShareLinkPermission**: The available values are:
  - **View**: Sets the default link permission to view permissions
  - **Edit**: Sets the default link permission to edit permissions

These two settings and values are the equivalent of the parameters *DefaultSharingScope* and *DefaultShareLinkPermission* from the Set-SPOSite cmdlet.

Another configuration for the default sharing link type is to use the **DefaultShareLinkToExistingAccess** advanced setting, which is the equivalent of the parameter

DefaultLinkToExistingAccess from the Set-SPOSite cmdlet. When you set this value to **True**, it overrides the other two advanced settings and their values.

PowerShell examples, where the sensitivity label GUID is **8faca7b8-8d20-48a3-8ea2-0f96310a848e**:

- To set the default sharing link type to SpecificPeople:

  PowerShell

  ```
  Set-Label -Identity 8faca7b8-8d20-48a3-8ea2-0f96310a848e -AdvancedSettings
  @{DefaultSharingScope="SpecificPeople"}
  ```

- To set the default sharing link type permissions to Edit:

  PowerShell

  ```
  Set-Label -Identity 8faca7b8-8d20-48a3-8ea2-0f96310a848e -AdvancedSettings
  @{DefaultShareLinkPermission="Edit"}
  ```

- To set the default sharing link type to people with existing access:

  PowerShell

  ```
  Set-Label -Identity 8faca7b8-8d20-48a3-8ea2-0f96310a848e -AdvancedSettings
  @{DefaultShareLinkToExistingAccess="True"}
  ```

For more help to specify PowerShell advanced settings, see PowerShell tips for specifying the advanced settings.

To configure the settings for the default sharing link type for a site, the scope of the sensitivity label must include **Groups & sites** when you create the sensitivity label in the Microsoft Purview portal. After it's created, you see this displayed as **Site, UnifiedGroup** in the **Scope** column on the **Labels** page, and the PowerShell *ContentType* setting also displays this same value. For documents, the scope must include **Files & other data assets**, which displays as **File, Email**. Then:

- When the scope includes **Groups & sites**, you can apply the label to a site, which sets the default sharing link type for that site. For information how to apply a sensitivity label to a site, see How to apply sensitivity labels to containers.

- When the scope of the sensitivity label includes **Files & other data assets**, you can apply the label to documents, which sets the default sharing link type for that document. The label can be applied manually or automatically.

Tip

You can also specify that the label is the default sensitivity label to be applied for new sites or new documents, as a **label policy setting**.

Manage sensitivity labels in Office apps

> *Microsoft Purview service description*

When you have published sensitivity labels from the Microsoft Purview portal, they start to appear in Office apps for users to classify and protect data as it's created or edited.

This article helps you successfully manage sensitivity labels in Office apps. For example, you can find additional configuration information for specific labeling features.

Sensitivity labeling support in apps

To use sensitivity labels in Office apps, you must use a subscription edition of Office. Use the licensing link at the top of this page to identify eligible plans. Sensitivity labels aren't supported for standalone editions of Office, sometimes called "Office Perpetual".

For Outlook (Windows, macOS, iOS, Android, and on the web), the mailbox must be hosted in Exchange Online. Sensitivity labels aren't supported for mailboxes that are hosted on-premises. This on-premises limitation also applies to shared mailboxes, even if the users accessing them have mailboxes in Exchange Online.

Support for sensitivity label capabilities in apps

Use the tables in Minimum versions for sensitivity labels in Office apps to identify the minimum Office version that introduced specific capabilities for sensitivity labels in Office apps, or whether the label capability is in public preview.

In addition to listing the minimum versions for Windows, macOS, iOS, and Android, the tables also include whether the capability is supported for Office for the web:

- Capabilities table for Word, Excel, and PowerPoint
- Capabilities table for Outlook

Office for iOS and Office for Android: Sensitivity labels are built into the Office app.

For information about Loop, see Use sensitivity labels with Microsoft Loop.

> **Tip**

Label and policy settings that refer to Office documents also apply to Loop components and pages.

Office built-in labeling and the Azure Information Protection client

The Windows Office add-in from the Azure Information Protection unified labeling client is now retired and no longer supported. It's replaced by sensitivity labels that are built into Office apps that support labeling.

Because the add-in is no longer supported, the Office policy setting **Use the Azure Information Protection add-in for sensitivity labeling** must be set to **Not configured** (the default), or **Disabled**. If this setting is configured for **Enabled**, you won't be able to use sensitivity labeling in Office apps.

To check for this setting:

- If you're using Group Policy, use the latest Microsoft 365 Apps for enterprise administrative templates and navigate to this setting from **User Configuration/Administrative Templates/Microsoft Office 2016/Security Settings**.
- If you're using the Cloud Policy service for Microsoft 365, search for the setting by name.

The other labeling capabilities from this older labeling client are supported with the Microsoft Purview Information Protection client. For more information, see Extend sensitivity labeling on Windows.

If you need to turn off built-in labeling in Office apps on Windows

To use sensitivity labeling in Office apps, you must have one or more label policies published to users from the Microsoft Purview portal, and a supported version of Office.

If both of these conditions are met but you need to turn off sensitivity labels in Windows Office apps, use the Office policy setting **Use the Sensitivity feature in Office to apply and view sensitivity labels**. Set the value to 0 by selecting **Disabled**.

For Group Policy and Microsoft 365 Apps for enterprise administrative templates, navigate to this setting from **User Configuration/Administrative Templates/Microsoft Office 2016/Security**

Settings. If you're using the Cloud Policy service for Microsoft 365, search for this setting by name. The setting takes effect when these Office apps restart.

If you later need to revert this configuration, change the value to 1 by selecting **Enabled**. You might also need to enable this setting if the **Sensitivity** button isn't displayed on the ribbon as expected. For example, a previous administrator turned this labeling setting off.

Because this setting is specific to Windows Office apps, it has no impact on other apps on Windows that support sensitivity labels (such as Power BI) or other platforms (such as macOS, mobile devices, and Office for the web). If you don't want some or all users to see and use sensitivity labels across all apps and all platforms, don't assign a sensitivity label policy to those users.

> **Tip**
>
> If you want to display built-in labels for only Outlook or for only Word, Excel, and PowerPoint, you can achieve this outcome with a per-label setting. For more information, see Scope labels to just files or emails.

Office file types supported

Generally, Office apps that have built-in labeling for Word, Excel, and PowerPoint files support the Open XML format (such as .docx and .xlsx) but not the Microsoft Office 97-2003 format (such as .doc and .xls), Open Document Format (such as .odt and .ods), or other formats.

> **Note**
>
> You can prevent users from opening or saving the older file types for Word, Excel, and PowerPoint in Windows. For instructions, see Block specific file format types in Office 2016.

When a file type isn't supported for sensitivity labels, the **Sensitivity** button isn't available in the Office app.

Supported file types for Office apps on Windows, macOS, iOS, and Android:

- **Word**: .docx, .docm, .dotx, .dotm
- **Excel**: .xlsx, .xlsb; .xlsm, .xltm, .xltx
- **PowerPoint**: .pptx, .pptm, .potx, .potm, .ppsx, .ppsm

For file types supported for SharePoint and OneDrive when these services are enabled for sensitivity labels, see Enable sensitivity labels for files in SharePoint and OneDrive.

Labeling for embedded files isn't supported, even with the listed file types.

For labeling scenarios outside Office apps, check their documentation for file types supported. For example, the additional file types supported by the Microsoft Purview information protection client and scanner.

Protection templates and sensitivity labels

Administrator-defined protection templates, such as those you define for Microsoft Purview Message Encryption, aren't visible in Office apps when you're using built-in labeling. This simplified experience reflects that there's no need to select a protection template, because the same settings are included with sensitivity labels that have encryption enabled.

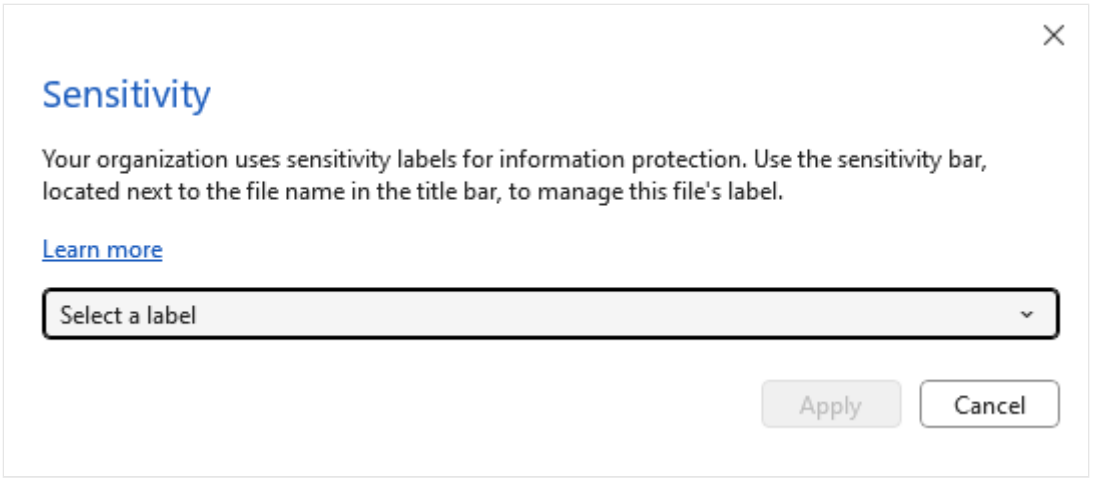
You can convert an existing template into a sensitivity label when you use the New-Label cmdlet with the *EncryptionTemplateId* parameter.

Information Rights Management (IRM) options and sensitivity labels

Sensitivity labels that you configure to apply encryption remove the complexity from users to specify their own encryption settings. In Office apps, these individual encryption settings might still be manually configured by users by using Information Rights Management (IRM) options. For example, for Windows apps:

- For a document: **File** > **Info** > **Protect Document** > **Restrict Access**
- For an email: From the **Options** tab > **Encrypt**

The latest Word, Excel, and PowerPoint apps for Windows and Mac no longer let users select Information Rights Management (IRM) options when they have sensitivity labels. Instead, users see a dialog box that prompts them to use a sensitivity label to apply information protection. The message also explains how to later access the sensitivity bar to select and change labels:



For Outlook and older Office apps that let users select Information Rights Management options, users can override settings from an applied sensitivity label with their own encryption settings. For example:

- A user applies the **Confidential \ All Employees** label to a document and this label is configured to apply encryption settings for all users in the organization. This user then manually configures the IRM settings to restrict access to a user outside your organization. The end result is a document that's labeled **Confidential \ All Employees** and encrypted, but users in your organization can't open it as expected.

- A user applies the **Confidential \ Recipients Only** label to an email and this email is configured to apply the encryption setting of **Do Not Forward**. In the Outlook app, this user then manually selects the IRM setting for Encrypt-Only. The end result is that while the email does remain encrypted, it can be forwarded by recipients, despite having the **Confidential \ Recipients Only** label.

  As an exception, for Outlook on the web, the options from the **Encrypt** menu aren't available for a user to select when the currently selected label applies encryption.

- A user applies the **General** label to a document, and this label isn't configured to apply encryption. This user then manually configures the IRM settings to restrict access to the document. The end result is a document that's labeled **General** but that also applies encryption so that some users can't open it as expected.

If the document or email is already labeled, a user can do any of these actions if the content isn't already encrypted, or they have the usage right Export or Full Control.

Guidance for a consistent label experience

For a more consistent label experience with meaningful reporting, provide appropriate labels and guidance for users to apply only labels to protect documents and emails. For example:

- For exception cases where users must assign their own permissions, provide labels that let users assign their own permissions.
- Instead of users manually removing encryption after selecting a label that applies encryption, provide a sublabel alternative when users need a label with the same classification, but no encryption. Such as:
  - **Confidential \ All Employees**
  - **Confidential \ Anyone (no encryption)**
- Upgrade to the latest versions that don't let users select IRM settings for Word, Excel, and PowerPoint. For Outlook, consider disabling IRM settings to prevent users from selecting them:
  - Outlook for Windows:
    - Registry keys `DWORD:00000001` *DisableDNF* and *DisableEO* from `HKEY_CURRENT_USER\Software\Microsoft\Office\16.0\Common\DRM`
    - Make sure that the Group Policy setting **Configure default encryption option for the Encrypt button** isn't configured
  - Outlook for Mac:
    - Keys *DisableEncryptOnly* and *DisableDoNotForward* security settings documented in Set preferences for Outlook for Mac
  - Outlook on the web:
    - Parameters *SimplifiedClientAccessDoNotForwardDisabled* and *SimplifiedClientAccessEncryptOnlyDisabled* documented for Set-IRMConfiguration
  - Outlook for iOS and Android: These apps don't support users applying encryption without labels, so nothing to disable.

> ⓘ **Note**
>
> If users manually remove encryption from a labeled document that's stored in SharePoint or OneDrive and you've **enabled sensitivity labels for Office files in SharePoint and OneDrive**, the label encryption will be automatically restored the next time the document is accessed or downloaded.

Encryption-based label matching for documents

When a document has been encrypted with administrator-defined permissions, the encryption policy is embedded in the document. This happens independently from labeling. For example, when an Office attachment inherits encryption from an email message, or a user has applied a rights management template by using Information Rights Management (IRM) in their Office app. If a sensitivity label in the tenant matches that same encryption policy, Office apps will automatically assign that matching label to the document.

In this scenario, the matching sensitivity label can label an unlabeled document, and replace an existing label that doesn't apply encryption. For example, the **General** label is replaced with **Confidential / All Employees**. Content markings from the matching label aren't automatically applied.

This scenario helps to move older encryption solutions from rights management templates to sensitivity labels that apply encryption with rights management.

However, you'll also see this behavior with a labeling scenario for email and meetings attachments when they are opened by the recipient. For example:

1. A user creates an email and attaches an unencrypted Office document, and then applies a label to the email.

   The label applies encryption with permissions that are set by the administrator, rather than the Do Not Forward or Encrypt-Only options. For example, for the label configuration, the admin selects **Assign permissions now**, and specifies all employees have read access.

2. When the email is sent, the attachment automatically inherits the encryption with rights management, but not the label.

3. When a recipient in the same tenant opens the encrypted document, a matching label for the admin-defined permissions is automatically displayed for the document, and persists if the document is saved.

   As an auditing event that's displayed in activity explorer, this user applied the label, not the email sender.

Encryption-based label matching works only within the tenant, and all the following conditions apply:

- The document is unlabeled or the existing label doesn't apply encryption
- The document is encrypted with admin-defined permissions
- The matching sensitivity label is published to the user who opens the document

- The label scope of the matching sensitivity label includes **Files & other data assets**

The matching label persists if the document is saved.

Sensitivity label compatibility

With RMS-enlightened apps: If you open a labeled and encrypted document or email in an application that supports the Microsoft rights management (RMS) technology but doesn't support sensitivity labels, the app still enforces encryption and rights management.

With the Microsoft Purview Information Protection client: You can view and change sensitivity labels that you apply to documents with Office apps by using the Microsoft Purview Information Protection client, and you can view and change labels applied by the Microsoft Purview Information Protection client in Office apps.

With other versions of Office: Any authorized user can open labeled documents and emails in other versions of Office. However, you can only view or change the label in supported Office versions or by using the Microsoft Purview Information Protection client. Supported Office app versions are listed in the previous section.

Support for SharePoint and OneDrive files protected by sensitivity labels

To use the Office built-in labeling client with Office for the web for documents in SharePoint or OneDrive, make sure you've enabled sensitivity labels for Office files in SharePoint and OneDrive.

Support for external users and labeled content

When you label a document or email, the label is stored as metadata that includes your tenant and a label GUID. When a labeled document or email is opened by an Office app that supports sensitivity labels, this metadata is read and only if the user belongs to the same tenant, the label displays in their app. For example, for built-in labeling for Word, PowerPoint, and Excel, the label name displays on the status bar. This behavior excludes guest accounts.

This implementation means that if you share documents with another organization that uses different label names, each organization can apply and see their own label applied to the document. Guest users won't see your labels in their apps.

> ⓘ **Note**

Two exceptions where another organization can't apply their own sensitivity labels:

- Using Office for the web, external users connect to your SharePoint sites or OneDrive locations and don't see their sensitivity labels because the site is owned by another organization.
- Using co-authoring from desktop or mobile apps, external users won't be able to apply their own sensitivity labels that are configured to apply encryption.

The same is true for email (and labeled calendar events) sent by Outlook. However, email clients other than Outlook might not retain the label metadata in the email headers. For example, users replying or forwarding from another organization that doesn't use Outlook will likely result in the original email label no longer visible to the original organization because the label metadata hasn't been retained. If that label applied encryption, the encryption persists to protect the contents.

The following elements from an applied label are visible to users outside your organization:

- Content markings. When a label applies content markings that include a header, footer, or watermark, these are added directly to the content and remain visible until somebody modifies or deletes them.
- The name and description of the underlying protection template from a label that applied encryption. This information can be displayed in a message bar at the top of the content, to provide information about who is authorized to view the content, and their usage rights for that content.

Sharing encrypted documents with external users

Although you can restrict access to users in your own organization, you can also extend access to any other user who has an account in Microsoft Entra ID. By default, these external users will be authenticated without any additional configuration. However, there might be additional configuration required for Microsoft Entra External Identities cross-tenant access settings and Conditional Access.

If external users don't have an account in Microsoft Entra ID, they can authenticate by using guest accounts in your tenant. These guest accounts can also be used to access shared documents in SharePoint or OneDrive when you have enabled sensitivity labels for Office files in SharePoint and OneDrive.

For more information about the optional Microsoft Entra features and using guest accounts for authentication requirements, see Microsoft Entra configuration for encryption content.

All Office apps and other applications that understand rights management technology from Microsoft can open encrypted documents after the user has successfully authenticated.

Because dynamic watermarks are an encryption setting, they're supported across organizations and will display the name of the person who opens the document as a result of a sensitivity label that's published in another tenant.

When Office apps apply content marking and encryption

Office apps apply content marking and encryption with a sensitivity label differently, depending on the app you use.

Expand table

App	Content marking	Encryption
Word, Excel, PowerPoint on all platforms	Immediately	Immediately
Outlook (classic) for PC	After Exchange Online sends the email or meeting invite	Immediately
Outlook for Mac *, new Outlook for Windows, Outlook on the web, iOS, and Android	After Exchange Online sends the email or meeting invite	After Exchange Online sends the email or meeting invite

* Version 16.79+ with the new Outlook for Mac. Earlier versions encrypt immediately.

For emails, content marking can be applied more than once if the message is processed by the Exchange transport service more than once. For example, you're using a third-party signature service for emails that are sent outside the organization.

Because Outlook for Mac, new Outlook for Windows, Outlook on the web, iOS, and Android don't encrypt items locally, copies stored in the Sent folder also aren't encrypted. If users send items from their Sent folders (for example, to forward or Reply All), instruct them to first remove and then reapply the sensitivity label that applies encryption

Solutions that apply sensitivity labels to files outside Office apps do so by applying labeling metadata to the file. In this scenario, content marking from the label's configuration isn't inserted into the file but encryption is applied. The content markings aren't automatically applied when you open those documents in Office apps. Similarly, the content markings aren't removed when a label is removed outside Office apps, even if those documents are then later opened in an Office app.

Scenarios that include applying a sensitivity label outside Office apps include:

- The scanner, File Explorer, and PowerShell from the Microsoft Purview Information Protection client
- Auto-labeling policies for SharePoint and OneDrive
- Exported labeled and encrypted data from Power BI
- Automatically applied labels from Microsoft Defender for Cloud Apps

For these scenarios, using their Office apps, a user can apply the label's content markings by temporarily removing or replacing the current label and then reapplying the original label.

Content markings with variables

> **Important**
>
> If your Office apps don't support this capability, they apply the markings as the original text specified in the label configuration, rather than resolving the variables. See the tables in Minimum versions for sensitivity labels in Office apps.

When you configure a sensitivity label for content markings, you can use the following variables in the text string for your header, footer, or watermark:

Expand table

Variable	Description	Example when label applied
`${Item.Label}`	Label display name of the label applied	**General**
`${Item.Name}`	File name or email subject of the content being labeled	**Sales.docx**
`${Item.Location}`	Path and file name of the document being labeled, or the email subject for an email being	**\\Sales\2020\Q3\Report.docx**

Variable	Description	Example when label applied
	labeled	
`${User.Name}`	Display name of the user applying the label	**Richard Simone**
`${User.PrincipalName}`	Microsoft Entra user principal name (UPN) of the user applying the label	**rsimone@contoso.com**
`${Event.DateTime}`	Date and time when the content is labeled, in the local time zone of the user applying the label in Microsoft 365 apps, or UTC (Coordinated Universal Time) for Office Online and auto-labeling policies	**8/10/2020 1:30 PM**

> **Note**
>
> The syntax for these variables is case-sensitive.

Dynamic watermarks also support inserting the email address of a user. However, content marking with a variable uses the email address of the user applying the label, which is suitable for accountability. Dynamic watermarks display the email address of the user who is currently opening the document, which acts as a visual deterrent against data leakage.

Additional considerations:

- All content markings with or without variables can be applied with labels that don't apply encryption. Dynamic watermarks can only be applied with labels that apply encryption.
- All content markings with or without dynamic variables can be changed or removed by the user after it's applied to the labeled content. Dynamic watermarks, once applied, can't be changed or removed without removing the label or exporting (if allowed) the content.
- All content markings with or without dynamic variables can be customized with different fonts, colors, orientation. Dynamic watermarks don't support this customization.

The restrictions that come with dynamic watermarks provide more security, but could also prevent the document from opening in a desktop version of Office that doesn't support this capability. Use dynamic watermarks only when you need this level of protection, and standard content marking when you don't. For detailed information, see the Dynamic watermarks documentation.

Setting different content markings for Word, Excel, PowerPoint, and Outlook

As an additional variable, you can configure content markings per Office application type by using an "If.App" variable statement in the text string, and identify the application type by using the values **Word**, **Excel**, **PowerPoint**, or **Outlook**. You can also abbreviate these values, which is necessary if you want to specify more than one in the same If.App statement.

Use the following syntax:

```
${If.App.<application type>}<your visual markings text> ${If.End}
```

As with the other content markings, the syntax is case-sensitive, which includes the abbreviations for each application type (WXPO).

Examples:

- **Set header text for Word documents only:**

  `${If.App.Word}This Word document is sensitive ${If.End}`

  In Word document headers only, the label applies the header text "This Word document is sensitive". No header text is applied to other Office applications.

- **Set footer text for Word, Excel, and Outlook, and different footer text for PowerPoint:**

  `${If.App.WXO}This content is confidential. ${If.End}${If.App.PowerPoint}This presentation is confidential. ${If.End}`

  In Word, Excel, and Outlook, the label applies the footer text "This content is confidential." In PowerPoint, the label applies the footer text "This presentation is confidential."

- **Set specific watermark text for Word and PowerPoint, and then watermark text for Word, Excel, and PowerPoint:**

  `${If.App.WP}This content is ${If.End}Confidential`

  In Word and PowerPoint, the label applies the watermark text "This content is Confidential". In Excel, the label applies the watermark text "Confidential". In Outlook, the label doesn't apply any watermark text because watermarks as content markings aren't supported for Outlook.

Require users to apply a label to their email and documents

> **Important**
>
> This configuration is sometimes referred to as mandatory labeling. To check for supported Office versions, see the tables in Minimum versions for sensitivity labels in Office apps.
>
> To use mandatory labeling for documents but not emails, see the instructions in the next section that explains how to configure Outlook-specific options.
>
> To use mandatory labeling for Power BI, see Mandatory label policy for Power BI.

When the policy setting **Require users to apply a label to their email and documents** is selected, users assigned the policy must select and apply a sensitivity label in their Office apps:

- For documents (Word, Excel, PowerPoint): When an unlabeled document is opened or saved.
- For emails (Outlook): At the time users send an unlabeled email message. For Outlook Mobile, this can be changed to when the email message is first composed.

Additional information:

- Unlabeled means a sensitivity label published from the user's tenant isn't applied to the content. Sensitivity labels applied by other tenants are ignored by this policy setting.
- When users are prompted to add a sensitivity label because they open an unlabeled document, they can add a label or choose to open the document in read-only mode.
- When mandatory labeling is in effect, users can't remove sensitivity labels from documents, but can change an existing label.
- When mandatory labeling is in effect, the print to PDF option will be unavailable when a document is labeled or encrypted. For more information, see the PDF support section on this page.
- When mandatory labeling is configured together with auto-labeling for Office apps for the same user, the user can be prompted to apply a label before auto-labeling is triggered. Because auto-labeling doesn't override a manually applied sensitivity label but can override

a default label, a potential workaround for this timing issue is to configure a default label for the user who is configured for both mandatory labeling and auto-labeling for Office apps.

For guidance about when to use this setting for mandatory labeling, see the information about policy settings.

> **Note**
>
> If you use the default label policy setting for documents and emails in addition to mandatory labeling:
>
> The default label always takes priority over mandatory labeling. However, if you use a version of Office that doesn't yet support a default label for existing documents, users will be prompted to apply a sensitivity label for each new document.
>
> Identify the minimum versions of Word, Excel, and PowerPoint that support a default label for existing documents by using the capabilities table and the row **Apply a default label to existing documents**.

For Outlook Mobile, change when users are prompted for a label

This setting requires a minimum version of 4.2316.0 for Outlook for Android and Outlook for iOS.

You can use a Microsoft Intune Managed apps app configuration policy to configure a setting from the Intune App Software Development Kit (SDK) that changes when users are prompted to select a sensitivity label for Outlook Mobile.

Instead of prompting for a label on send when mandatory labeling is configuring for emails, this configuration results in prompting for a label when a user first composes a message.

This configuration requires you to specify the following key/value pair as a general configuration setting in the policy:

Expand table

Key	Value
com.microsoft.outlook.Mail.LouderMandatoryLabelEnabled	true

Outlook-specific options for default label and mandatory labeling

Identify the minimum versions of Outlook that support these features by using the capabilities table for Outlook and the row **Different settings for default label and mandatory labeling**.

When the Outlook app supports a default label setting that's different from the default label setting for documents:

- In the label policy configuration from the Microsoft Purview portal, on the **Apply a default label to emails** page: You can specify your choice of sensitivity label that will be applied to all unlabeled emails, or no default label. This setting is independent from the **Apply this label by default to documents** setting on the previous **Policy settings for documents** page of the configuration.

When the Outlook app doesn't support a default label setting that's different from the default label setting for documents: Outlook will always use the value you specify for **Apply this label by default to documents** on the **Policy settings for documents** page of the label policy configuration.

When the Outlook app supports turning off mandatory labeling:

- In the label policy configuration from the Microsoft Purview portal, on the **Policy settings** page: Select **Require users to apply a label to their email or documents**. Then select **Next** > **Next** and clear the checkbox **Require users to apply a label to their emails**. Keep the checkbox selected if you want mandatory labeling to apply to emails as well as to documents.

When the Outlook app doesn't support turning off mandatory labeling: If you select **Require users to apply a label to their email or documents** as a policy setting, Outlook will always prompt users to select a label for unlabeled emails.

Scope labels to just files or emails

> **Note**
>
> Identify the minimum versions that support this feature by using the capabilities tables, and the row **Scope labels to files or emails**.

Unless this capability is supported on all the platforms used by your users, they will have an inconsistent labeling experience. For example, Word on one platform doesn't display a label that they see on a different platform.

This configuration is available when you create or edit a sensitivity label in the Microsoft Purview portal:

- To scope labels to just Word, Excel, and PowerPoint: Make sure the option for **Files & other data assets** is selected, and not the option for **Emails**.
- To scope labels to just Outlook, make sure the option for **Emails** is selected, and not the option for **Files & other data assets**.

> ⚠ **Warning**
>
> Although you can edit an existing label and remove the **Files & other data assets** scope, we don't recommend you do this because existing configurations might no longer work as expected. For example, a SharePoint site admin wouldn't understand why a sensitivity label that they selected as a default label for a document library no longer applies the label.
>
> If you want a sensitivity label just for emails, create a new label with just the **Emails** scope rather than edit an existing label.

Make sure both options are selected if you don't need to scope the labels to just Word, Excel, and PowerPoint, or to just Outlook.

Remember that other label configurations can also influence whether sensitivity labels are visible in apps. Check the documentation for the label configurations that you use.

> ⓘ **Note**
>
> The **Files & other data assets** option can include items other than Office files, for example, Power BI files. Check the application's documentation to verify, and remember to test all labeling apps and services used by your organization.

Be aware that this configuration affects both client apps and services, manual labeling and automatic labeling. For example:

- Default labels:
  - If the scope doesn't include email, a configured default label for email won't be applied.

    - If the scope doesn't include files, a configured default label for files won't be applied and can't be selected as a default sensitivity label for a SharePoint document library.

- Auto-labeling policies:
    - If the scope doesn't include email, you can't select the label for an auto-labeling policy that includes the Exchange location.
    - If the scope doesn't include files, you can't select the label for an auto-labeling policy that includes the SharePoint and OneDrive locations.

- Encryption that lets users assign permissions:
    - If the scope doesn't include email, you won't be able to select the encryption options of **Do Not Forward** or **Encrypt-Only**.
    - If the scope doesn't include files, you won't be able to select the encryption option **In Word, PowerPoint, and Excel, prompt users to specify permissions**.

- Label inheritance from email attachments:
    - For this configuration, the label must be scoped to both files and emails.

In addition, if a label has been previously applied but then removed from one of the scopes, users will no longer see that label applied for the scope in the apps that support this feature.

Because of the impact of scoping labels to just files or emails, some existing labeling configurations will prevent you from removing the scope options for **Files & other data assets** and **Emails**:

- Default label in label policies
- Default label to apply in channel meetings
- Label selected for auto-labeling policies

Before you can scope a label to just files or emails, you must first remove it if it's configured as one of these default labels, and remove it from any auto-labeling policies.

Limitations:

- If you use any label policy settings with a sensitivity label that's scoped to just **Files & other data assets** or scoped to just **Emails**, the same policy must also include at least one label with both scope options.

- If the label is configured as the default label in one or more label policies, and Outlook isn't configured with its own default label in the same policy, you can't remove the scope for **Email**. As a workaround, first remove this label as the default label. You'll then be able to

remove the email scope. Finally, reselect the now modified label as the default label for documents.

Configure a label to apply S/MIME protection in Outlook

> **Note**
>
> Identify the minimum versions of Outlook that support this feature by using the capabilities table for Outlook and the row **Apply S/MIME protection**.
>
> If you configure a label to apply S/MIME protection but your version of Outlook for Windows doesn't yet support it, the label is still displayed and can be applied, but the S/MIME settings are ignored. You won't be able to select this label for Exchange auto-labeling policies or configure it to protect calendar items, Teams meetings, and chat.

This configuration is not available in the Microsoft Purview portal. You must use PowerShell advanced settings with the Set-Label or New-Label cmdlet after you've connected to Security & Compliance PowerShell.

Use these settings only when you have a working S/MIME deployment and want a label to automatically apply this protection method for emails rather than the default protection that uses Rights Management encryption from Azure Information Protection. The resulting protection will be the same as when a user manually selects S/MIME options from Outlook.

Expand table

Configuration	Advanced setting key/value
S/MIME digital signature	SMimeSign="True"
S/MIME encryption	SMimeEncrypt="True"

The label you configure for these settings doesn't have to be configured for encryption in the Microsoft Purview portal. But if it is, S/MIME protection replaces the Rights Management encryption only in Outlook. For other apps, the label applies the encryption settings specified in the Microsoft Purview portal.

Example PowerShell commands, where the sensitivity label GUID is **8faca7b8-8d20-48a3-8ea2-0f96310a848e**:

PowerShell

```
Set-Label -Identity "8faca7b8-8d20-48a3-8ea2-0f96310a848e" -AdvancedSettings
@{SMimeSign="True"}

Set-Label -Identity "8faca7b8-8d20-48a3-8ea2-0f96310a848e" -AdvancedSettings
@{SMimeEncrypt="True"}
```

For more help in specifying PowerShell advanced settings, see PowerShell tips for specifying the advanced settings.

Configure label inheritance from email attachments

> **Note**
>
> Identify the minimum versions of Outlook that support this feature by using the capabilities table for Outlook, and the row **Label inheritance from email attachments**.

Turn on email inheritance for when users attach labeled documents to an email message that isn't manually labeled. With this configuration, a sensitivity label is dynamically selected for the email message, based on the sensitivity labels that are applied to the attachments and published to the user. The highest priority label is dynamically selected when it's supported by Outlook.

Whether this label inheritance will override an existing label on the email message:

- When an email message has been manually labeled, that label won't be replaced by label inheritance from email attachments.
- Label inheritance from email attachments will replace a lower priority sensitivity label that is automatically applied or applied as a default label, but won't override a higher priority label.

You configure this setting in the sensitivity label policy, on the **Default settings for emails** page. For the section **Inherit label from attachments**, select the checkbox **Email inherits highest priority label from attachments**. The attachment must be a physical file, and can't be a link to a file (for example, a link to a file on Microsoft SharePoint or OneDrive).

When you select this checkbox, you can then further select the following option: **Recommend users apply the attachments label instead of automatically applying it.** Without this selection, the label is automatically applied but users can still remove the label or select a different label before sending the email.

By default, if the automatically selected label applies encryption, the same encryption is applied to the email. For example, if the highest priority label applies encryption with Full Control to the Marketing group, the email will be protected with Full Control to the Marketing group. If the highest priority label applies the encryption option of Do Not Forward, the email message is also labeled and encrypted with Do Not Forward.

However, consider the outcome when an email client doesn't support a specific protection action that's been applied to an attachment:

- **Double Key Encryption**: Behavior depends on whether Outlook supports this encryption method. Use the capabilities tables and the row **Double Key Encryption (DKE)** to confirm support for your version.
    - When Outlook supports DKE: If the highest priority label applies the encryption setting for Double Key Encryption and **Assign permissions now**, Outlook for Windows applies that label and protection to the email message. The label and protection isn't applied if the label is configured for **Let users assign permissions when they apply the label**.
    - When Outlook doesn't support DKE: If the highest priority label applies Double Key Encryption, no label or encryption is selected for the email message in Outlook for Windows.
- **Custom permissions for Word, PowerPoint, and Excel**: If the highest priority label applies just user-defined permissions for Word, PowerPoint, and Excel (the option **Let users assign permissions when they apply the label** and **In Word, PowerPoint, and Excel, prompt users to specify permissions**), no label or protection is selected for the email message because Outlook doesn't support this label configuration.

PDF support

Use the tables in Minimum versions for sensitivity labels in Office apps to identify supported versions.

Office for Windows: Word, Excel, and PowerPoint support the following methods to convert an Office document into a PDF document:

- File > Save As > PDF
- File > Export > PDF
- Share > Send a Copy > PDF

This action is logged with the **Renamed file** audit event from the File and page activities auditing group. In the auditing search results in the Microsoft Purview portal, you'll see the details of this auditing event display **SensitivityLabeledFileRenamed** for the **Activity** field.

Office for iOS: Word, Excel, and PowerPoint support the following methods to convert an Office document into a PDF document:

- Export > PDF
- Send a copy > Send as a PDF

Office for Android: Word, Excel, and PowerPoint support the following methods to convert an Office document into a PDF document:

- Share > Send a Copy -> PDF

Office for the web: You must download the file from the browser. The following methods are supported to convert an Office online document into a PDF document:

- Word and PowerPoint on the web:
  - Save as > Download as PDF > Download
- Excel on the web:
  - Export > Download as PDF > Download
  - Print > Print > Download as PDF > Download

When the PDF is created, it inherits the label with any content markings. For Windows, if the label applied encryption, that encryption is also inherited. Encrypted PDFs can be opened with Microsoft Edge on Windows or Mac. For more information, see View protected PDFs using Microsoft Edge on Windows or Mac.

SharePoint and OneDrive support the following PDF scenarios:

- When you've enabled SharePoint and OneDrive for sensitivity labels and added PDF support. Then, PDFs are supported when you upload a labeled PDF document with or without encryption applied, these services can process the file such that search, eDiscovery, and data loss prevention can inspect the contents, and the sensitivity label name is displayed for users.

- Auto-labeling policies can apply a sensitivity label and encryption (if configured) to PDF documents when PDF support is added.

Outlook doesn't currently support PDF attachments inheriting encryption from a labeled message. However, Outlook now does support warning or blocking users from printing to PDF, as

described next.

PDF scenarios not supported:

- Print to PDF

  If users select this option, they are warned that the document or email will lose the protection of the label, and encryption (if applied), and must confirm to continue. If your sensitivity label policy requires justification to remove a label or lower its classification, they see this prompt.

  Because this option removes the sensitivity label, this option won't be available to users if you're using mandatory labeling. This configuration refers to the sensitivity label policy setting that requires users to apply a label to their emails and documents.

- PDF/A format and encryption

  This PDF format designed for long-term archiving isn't supported when the label applies encryption and will prevent users from converting Office documents to PDF. For configuration information, see the Group Policy documentation for Enforce PDF compliance with ISO 19005-1 (PDF/A).

- Password protection and encryption

  The option **File** > **Info** > **Protect Document** > **Encrypt with Password** isn't supported when the document's label applies encryption. In this scenario, the encrypt with password option becomes unavailable for users.

For more information about exporting to PDF, see the announcement Apply sensitivity labels to PDFs created with Office apps.

For end user documentation, see Create protected PDFs from Office files and View protected PDFs using Microsoft Edge on Windows or Mac.

Disabling PDF support

If you need to disable the PDF support in Office apps for Word, Excel, and PowerPoint, you can do so by using a Group Policy Office setting under **User Configuration/Administrative Templates/Microsoft Office 2016/Microsoft Save As PDF and Save As XPS add-ins**:

- **Use the Sensitivity feature in Office to apply sensitivity labels to PDFs**

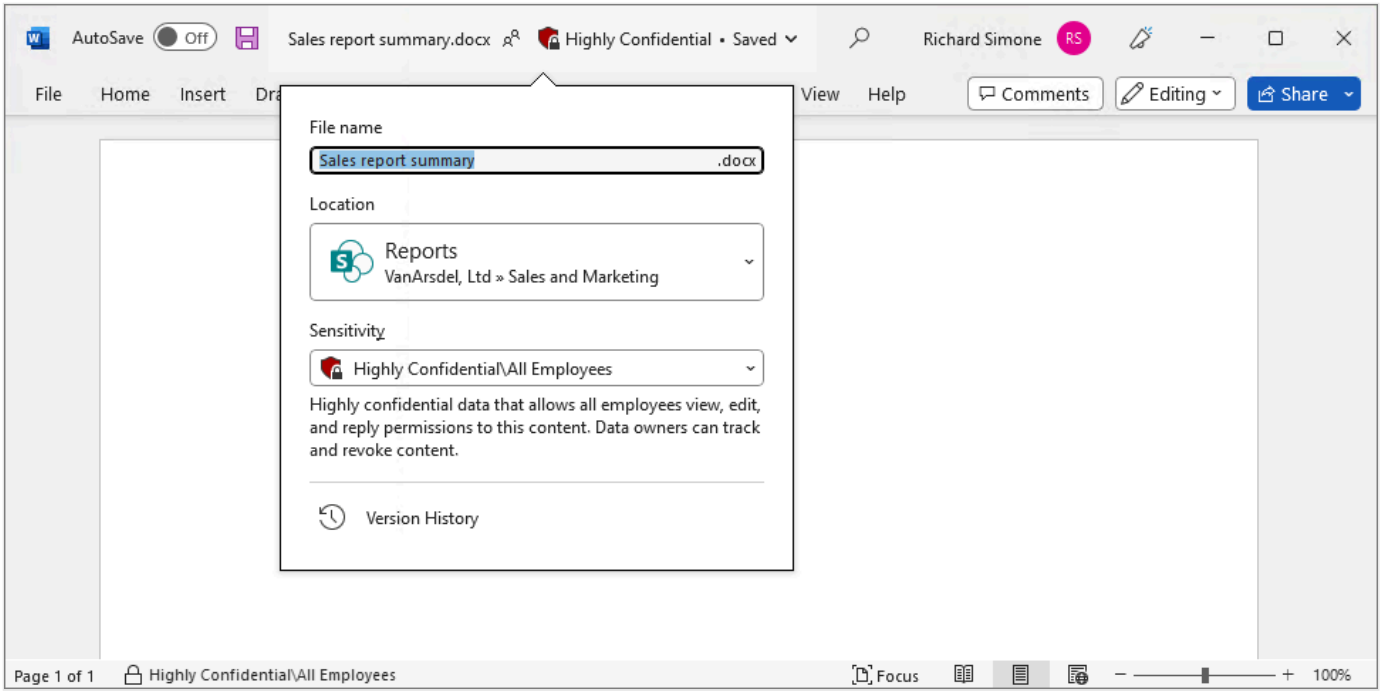
Configure this setting to be **Disabled**.

Deploy this setting by using Group Policy, or by using the Cloud Policy service for Microsoft 365.

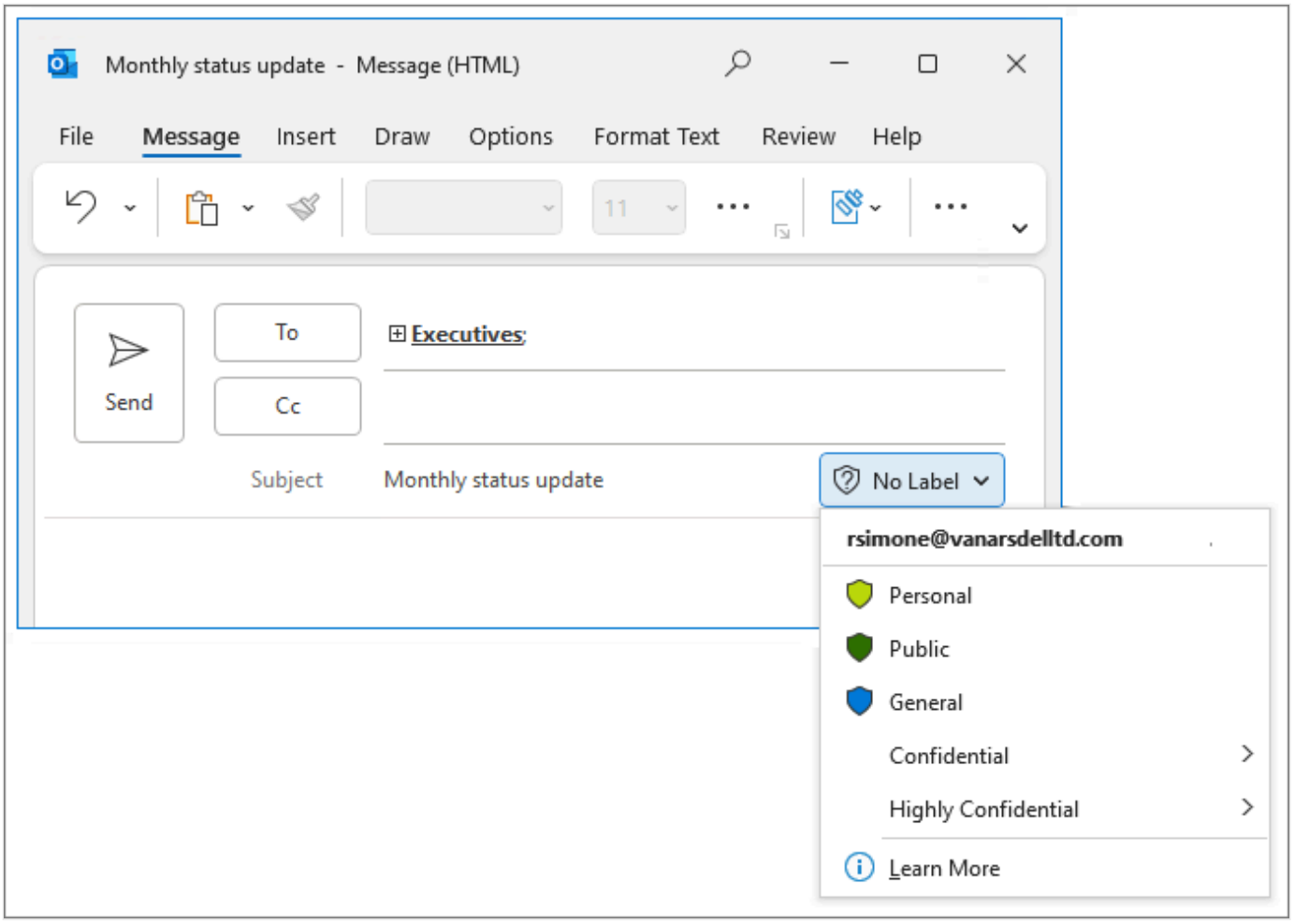
Sensitivity bar

Use the tables in Minimum versions for sensitivity labels in Office apps to identify which Office versions support the sensitivity bar.

When Word, Excel, and PowerPoint support this feature, sensitivity labels are displayed in a sensitivity bar next to the file name on the top window bar. For example:



When Outlook supports this feature, the sensitivity bar is displayed on the **Subject** line of the email. For example:



Information about the labels and the ability to select or change a label are also integrated into user workflows that includes save and rename, export, share, print, and convert to PDF. For more information and example screenshots, see the blog post announcement, New sensitivity bar in Office for Windows.

As part of this high visibility, these labels also support colors. For more information, see the next section.

Label colors

If your labeling apps don't support this capability, they don't display the configured label colors. To identify supported versions for Office apps, see the tables in Minimum versions for sensitivity labels in Office apps.

Newly created labels don't have a color by default. Use the Microsoft Purview portal to select one of 10 standard colors for sensitivity labels. The **Label color** configuration is on the first page of the label configuration after the label name and description.

You can't select colors for sublabels because they automatically inherit the label color from their [parent label or group label](sensitivity-labels.md#sublabels-that-use-parent-labels-or label-

groups).

If a label is configured for a different color from one of the 10 default colors, you see a **Use previously assigned custom color** check box selected, and the standard color options aren't available. You can change the custom color to one of the standard colors by first clearing the checkbox, and then you can select one of the standard colors.

You can't use the Microsoft Purview portal to configure a different custom color. Instead, use PowerShell, as described in the next section.

Configuring custom colors by using PowerShell

You can use the Security & Compliance PowerShell advanced setting **color** to set a color for a sensitivity label. This configuration supports colors that you can't configure in the Microsoft Purview portal.

To specify your choice of color, use a hex triplet code for the red, green, and blue (RGB) components of the color. For example, #40e0d0 is the RGB hex value for turquoise.

For more information about these codes, see the <color> page from the MSDN web docs, and you might also find RapidTables helpful. You can identify these codes in many applications that let you edit pictures. For example, Microsoft Paint lets you choose a custom color from a palette and the RGB values are automatically displayed, which you can then copy.

Example PowerShell command, where the sensitivity label GUID is **8faca7b8-8d20-48a3-8ea2-0f96310a848e**

PowerShell

```
Set-Label -Identity 8faca7b8-8d20-48a3-8ea2-0f96310a848e -AdvancedSettings
@{color="#40e0d0"}
```

For more information to help you specify PowerShell advanced settings for sensitivity labels, see PowerShell tips for specifying the advanced settings.

Specify a default sublabel for a parent label or label group

Note

Identify the minimum versions of Office that support this setting by using the capabilities tables and the row **Default sublabel for parent or group label**.

This configuration is not available in the Microsoft Purview portal. You must use the PowerShell advanced setting *DefaultSubLabelId* with the Set-Label or New-Label cmdlet after you've connected to Security & Compliance PowerShell.

By default, users select a [parent label or group label](sensitivity-labels.md#sublabels-that-use-parent-labels-or label-groups) to see the sublabels that they can apply, and then select one of those sublabels. If you specify a default sublabel for a parent label or label group, when users select the parent label or label group, a sublabel is automatically selected and applied for them.

As an example, the parent label **Confidential** is configured with a default sublabel, **All Employees**. The next parent label, **Highly Confidential**, doesn't have a default sublabel configured. You can tell the difference with the end of label bar for **Confidential** that isn't visible for **Highly Confidential**:



When users select to the left of that vertical bar, they are selecting **Confidential\All Employees** automatically with a single selection. If they need a different sublabel, they must select to the right of the vertical bar to select the label expansion that then displays all the sublabels for selection. In comparison, if they select **Highly Confidential**, the sublabels for that label are always displayed for selection.

After you configure this label setting, remember to update your end-user documentation accordingly.

Example PowerShell command, where the parent sensitivity label or label group GUID is **8faca7b8-8d20-48a3-8ea2-0f96310a848e** and its sublabel that you want to specify as the default is **1ace2cc3-14bc-4142-9125-bf946a70542c**:

PowerShell

```
Set-Label -Identity "8faca7b8-8d20-48a3-8ea2-0f96310a848e" -AdvancedSettings
```

```
@{DefaultSubLabelId="1ace2cc3-14bc-4142-9125-bf946a70542c"}
```

For more help in specifying PowerShell advanced settings, see PowerShell tips for specifying the advanced settings.

Prevent connected experiences that analyze content

> **Note**
>
> Support for this setting differs for clients and apps. Identify the minimum versions of Office that support this setting by using the capabilities tables and the rows **Prevent some connected experiences that analyze content** and **Prevent all connected experiences that analyze content**.

This configuration isn't available in the Microsoft Purview portal. You must use the PowerShell advanced setting *BlockContentAnalysisServices* with the Set-Label or New-Label cmdlet after you've connected to Security & Compliance PowerShell.

The setting lets you prevent content in Word, Excel, PowerPoint, and Outlook (for **Prevent all connected experiences that analyze content**) from being sent to Microsoft for content analysis as a privacy control. However, when it's set, it means that some services won't work as designed, such as data loss prevention policy tips for Outlook, automatic and recommended labeling, and Microsoft 365 Copilot.

> **Important**
>
> Although content with the configured sensitivity label will be excluded from Microsoft 365 Copilot in the named Office apps, the content remains available to Microsoft 365 Copilot for other scenarios. For example, in Teams, and in Microsoft 365 Copilot Chat.

Example PowerShell command, where the sensitivity label **Confidential\Project Onyx** has a GUID of **8faca7b8-8d20-48a3-8ea2-0f96310a848e** and you want to prevent all documents and emails with this label in Office apps from being sent to Microsoft connected experiences for analysis:

PowerShell

```
Set-Label -Identity "8faca7b8-8d20-48a3-8ea2-0f96310a848e" -AdvancedSettings
@{BlockContentAnalysisServices="True"}
```

To return the setting back to the default of sending the labeled content to Microsoft connected experiences for analysis, remove the setting or set the value to False.

For more help in specifying PowerShell advanced settings, see PowerShell tips for specifying the advanced settings.

On newer Windows clients and on all supported MacOS and mobile client versions, the setting affects all connected experiences that analyze content in Word, Excel, and PowerPoint. Supported versions are listed in the **Prevent all connected experiences that analyze content** row.

On older Windows versions of Word, Excel, PowerPoint, and Outlook that are listed in the **Prevent some connected experiences that analyze content** row, the following subset of connected experiences are impacted by this setting:

Expand table

Connected experiences	Apps
Automatic alt txt	Word, PowerPoint, Excel, Outlook
Automatically apply or recommend sensitivity labels	Word, PowerPoint, Excel, Outlook
Microsoft 365 Copilot	Word, PowerPoint, Excel, Outlook
Microsoft Purview Data Loss Prevention policy tips	Outlook
PowerPoint Designer	PowerPoint
Similarity checker	Word
Translator	Word, PowerPoint, Excel, Outlook

Auditing labeling activities

For information about the auditing events that are generated by sensitivity label activities, see the Sensitivity label activities section from the audit log activities documentation.

This auditing information is visually represented in content explorer and activity explorer to help you understand how your sensitivity labels are being used and where this labeled content is located.

You can also create custom reports with your choice of security information and event management (SIEM) software when you export and configure the audit log records. For larger-scale reporting solutions, see the Office 365 Management Activity API reference.

> **Tip**
>
> To help create custom reports, see the following blog posts:
>
> - Microsoft Purview audit log activities via O365 Management API - Part 1
> - Microsoft Purview audit log activities via O365 Management API - Part 2

End-user documentation

- Apply sensitivity labels to your files and email in Office
  - Known issues with sensitivity labels in Office
- Automatically apply or recommend sensitivity labels to your files and emails in Office
  - Known issues with automatically applying or recommending sensitivity labels
- Create protected PDFs from Office files

Last updated on 06/30/2026

Minimum versions for sensitivity labels in Office apps

> *Microsoft Purview service description*

The capabilities tables on this page supplement Manage sensitivity labels in Office apps by listing the minimum Office version that introduced specific capabilities for sensitivity labels built in to Word, Excel, PowerPoint, and Outlook. Alternatively, the tabels indicate if the label capability is in public preview or not available.

For other apps that support sensitivity labels, check their documentation.

New versions of Office apps are made available at different times for different update channels. For Windows, you'll get the new capabilities earlier when you're on the Current Channel or Monthly Enterprise Channel, rather than Semi-Annual Enterprise Channel. The minimum version numbers can also be different from one update channel to the next. For more information, see Overview of update channels for Microsoft 365 Apps and Update history for Microsoft 365 Apps.

New capabilities that are in private preview aren't included in the tables but you might be able to join these previews by nominating your organization for the Microsoft Information Protection private preview program.

Tip

When you compare the minimum versions in the tables with the versions you have, remember the common practice of release versions to omit leading zeros.

For example, you have version 4.2128.0 and read that 4.7.1+ is the minimum version. For easier comparison, read 4.7.1 (no leading zeros) as 4.**0007**.1 (and not 4.** 7000**.1). Your version of 4.2128.0 is higher than 4.0007.1; so, your version is supported.

Sensitivity label capabilities in Word, Excel, and PowerPoint

The numbers listed are the minimum Office application versions required for each capability.

> **Note**
>
> For Windows and the Semi-Annual Enterprise Channel, the minimum supported version numbers might not yet be released. Learn more.

Expand table

Capability	Windows	Mac	iOS	Android	Web
Manually apply, change, or remove label	Current Channel: 1910+ Monthly Enterprise Channel: 1910+ Semi-Annual Enterprise Channel: 2002+	16.21+	2.21+	16.0.11231+	Yes - opt-in
Multi-language support	Current Channel: 1910+ Monthly Enterprise Channel: 1910+ Semi-Annual Enterprise Channel: 2002+	16.21+	2.21+	16.0.11231+	Not available
Apply a default label to new documents	Current Channel: 1910+ Monthly Enterprise Channel: 1910+ Semi-Annual Enterprise Channel: 2002+	16.21+	2.21+	16.0.11231+	Yes - opt-in
Apply a default label to existing documents	Current Channel: 2208+ Monthly Enterprise Channel: 2207+ Semi-Annual	16.63+	Not available	Not available	Yes - opt-in

Capability	Windows	Mac	iOS	Android	Web
	Enterprise Channel: 2302+				
Require a justification to change a label	Current Channel: 1910+ Monthly Enterprise Channel: 1910+ Semi-Annual Enterprise Channel: 2002+	16.21+	2.21+	16.0.11231+	Yes - opt-in
Provide help link to a custom help page	Current Channel: 1910+ Monthly Enterprise Channel: 1910+ Semi-Annual Enterprise Channel: 2002+	16.21+	2.21+	16.0.11231+	Yes - opt-in
Mark the content	Current Channel: 1910+ Monthly Enterprise Channel: 1910+ Semi-Annual Enterprise Channel: 2002+	16.21+	2.21+	16.0.11231+	Yes - opt-in
Markings with variables	Current Channel: 2010+ Monthly Enterprise Channel: 2010+ Semi-Annual Enterprise Channel: 2102+	16.42+	2.42+	16.0.13328+	Yes - opt-in
Assign permissions now	Current Channel: 1910+ Monthly Enterprise Channel: 1910+	16.21+	2.21+	16.0.11231+	Yes - opt-in

Capability	Windows	Mac	iOS	Android	Web
	Semi-Annual Enterprise Channel: 2002+				
Let users assign permissions: - Prompt users for custom permissions (users and groups)	Current Channel: 2004+ Monthly Enterprise Channel: 2004+ Semi-Annual Enterprise Channel: 2008+	16.35+	Not available	Not available	Yes - opt-in
Let users assign permissions: - Prompt users for custom permissions (users, groups, and organizations)	Current Channel: 2212+ Monthly Enterprise Channel: 2302+ Semi-Annual Enterprise Channel: 2302+	Not available	Not available	Not available	Yes - opt-in
Audit label-related user activity: - Excludes encryption details	Current Channel: 2011+ Monthly Enterprise Channel: 2011+ Semi-Annual Enterprise Channel: 2108+	16.43+	2.46+	16.0.13628+	Yes
Audit label-related user activity: - Includes encryption details	Current Channel: 2301+ Monthly Enterprise Channel: 2302+ Semi-Annual Enterprise Channel: 2302+	16.70+	2.70+	16.0.16130+	Not available
Require users to apply a label to their email and documents	Current Channel: 2101+ Monthly Enterprise Channel: 2101+	16.45+	2.47+	16.0.13628+	Yes - opt-in

Capability	Windows	Mac	iOS	Android	Web
	Semi-Annual Enterprise Channel: 2108+				
Apply a sensitivity label to files automatically - Using sensitive info types	Current Channel: 2009+ Monthly Enterprise Channel: 2009+ Semi-Annual Enterprise Channel: 2102+	16.44+	Not available	Not available	Yes - opt-in
Apply a sensitivity label to files automatically - Using trainable classifiers	Current Channel: 2105+ Monthly Enterprise Channel: 2105+ Semi-Annual Enterprise Channel: 2108+	16.49+	Not available	Not available	Not available
Support co-authoring and AutoSave for labeled and encrypted documents	Current Channel: 2107+ Monthly Enterprise Channel: 2107+ Semi-Annual Enterprise Channel: 2202+	16.51+	2.58+	16.0.14931+	Yes - opt-in
PDF support	Current Channel: 2208+ Monthly Enterprise Channel: 2209+ Semi-Annual Enterprise Channel: 2302+	Not available	2.82+	16.0.17224+	Yes - opt-in
Sensitivity bar and display label color	Current Channel: 2302+ Monthly Enterprise	16.71+	Not available	Not available	Yes - opt-in

Capability	Windows	Mac	iOS	Android	Web
	Channel: 2303+ Semi-Annual Enterprise Channel: 2302+				
Default sublabel for parent or group label	Current Channel: 2302+ Monthly Enterprise Channel: 2302+ Semi-Annual Enterprise Channel: 2302+	Not available	Not available	Not available	Not available
Scope labels to files or emails	Current Channel: 2301+ Monthly Enterprise Channel: 2302+ Semi-Annual Enterprise Channel: 2302+	16.70+	2.70+	16.0.16130+	Yes - opt-in
Double Key Encryption (DKE)	Current Channel: 2307+ Monthly Enterprise Channel: 2309+ Semi-Annual Enterprise Channel: 2308+	16.85+	2.85+	16.0.18227+	Not available
Document tracking and revocation	Current Channel: 2402+ Monthly Enterprise Channel: 2402+ Semi-Annual Enterprise Channel: 2402+	Not available	Not available	Not available	Not available

Capability	Windows	Mac	iOS	Android	Web
Prevent some connected experiences that analyze content	Current Channel: 2406+ Monthly Enterprise Channel: 2406+ Semi-Annual Enterprise Channel: 2402+	Not available	Not available	Not available	Not available
Prevent all connected experiences that analyze content	Rolling out to Current Channel: 2606+ Rolling out to Monthly Enterprise Channel: 2606+	Rolling out: 16.110+	Rolling out: 2.111+	Rolling out: 16.0.20026+	Not available
Dynamic watermarks - Encryption option: Assign permissions now	Current Channel: 2411+ Monthly Enterprise Channel: 2412+	16.94+	2.94+	16.0.18526+	Yes - opt-in
Dynamic watermarks - Encryption option: Let users assign permissions	Current Channel: 2503+	16.96+	2.96+	16.0.18730+	Yes - opt-in

Sensitivity label capabilities in Outlook

The numbers listed are the minimum Office application versions required for each capability.

> **Note**
>
> For Windows and the Semi-Annual Enterprise Channel, the minimum supported version numbers might not yet be released. Learn more

Expand table

Capability	Outlook for Windows	Outlook for Mac	Outlook on iOS	Outlook on Android	Outlook on the web	New Outlook for Windows*
Manually apply, change, or remove label - Files and emails	Current Channel: 1910+ Monthly Enterprise Channel: 1910+ Semi-Annual Enterprise Channel: 2002+	16.21+	4.7.1+	4.0.39+	Yes	Yes
Manually apply, change, or remove label - Calendar items	Current Channel: 2302+ Monthly Enterprise Channel: 2302+ Semi-Annual Enterprise Channel: 2402+	16.70+ **	4.2331.0+	4.2326.1+	Yes	Yes
Multi-language support	Current Channel: 1910+ Monthly Enterprise Channel: 1910+ Semi-Annual Enterprise Channel: 2002+	16.21+	4.7.1+	4.0.39+	Yes	Yes
Apply a default label	Current Channel: 1910+	16.21+	4.7.1+	4.0.39+	Yes	Yes

Capability	Outlook for Windows	Outlook for Mac	Outlook on iOS	Outlook on Android	Outlook on the web	New Outlook for Windows*
	Monthly Enterprise Channel: 1910+ Semi-Annual Enterprise Channel: 2002+					
Require a justification to change a label	Current Channel: 1910+ Monthly Enterprise Channel: 1910+ Semi-Annual Enterprise Channel: 2002+	16.21+	4.7.1+	4.0.39+	Yes	Yes
Provide help link to a custom help page	Current Channel: 1910+ Monthly Enterprise Channel: 1910+ Semi-Annual Enterprise Channel: 2002+	16.21+	4.7.1+	4.0.39+	Yes	Yes
Mark the content	Current Channel: 1910+ Monthly Enterprise Channel: 1910+	16.21+	4.7.1+	4.0.39+	Yes	Yes

Capability	Outlook for Windows	Outlook for Mac	Outlook on iOS	Outlook on Android	Outlook on the web	New Outlook for Windows*
	Semi-Annual Enterprise Channel: 2002+					
Content markings with variables	Current Channel: 1910+ Monthly Enterprise Channel: 1910+ Semi-Annual Enterprise Channel: 2002+	16.21+	4.7.1+	4.0.39+	Yes	Yes
Assign permissions now	Current Channel: 1910+ Monthly Enterprise Channel: 1910+ Semi-Annual Enterprise Channel: 2002+	16.21+	4.7.1+	4.0.39+	Yes	Yes
Let users assign permissions: - Do Not Forward	Current Channel: 1910+ Monthly Enterprise Channel: 1910+ Semi-Annual Enterprise	16.21+	4.7.1+	4.0.39+	Yes	Yes

Capability	Outlook for Windows	Outlook for Mac	Outlook on iOS	Outlook on Android	Outlook on the web	New Outlook for Windows*
	Channel: 2002+					
Let users assign permissions: - Encrypt-Only	Current Channel: 2011+ Monthly Enterprise Channel: 2011+ Semi-Annual Enterprise Channel: 2108+	16.48+ **	4.2112.0+	4.2112.0+	Yes	Yes
Require users to apply a label to their email and documents	Current Channel: 2101+ Monthly Enterprise Channel: 2101+ Semi-Annual Enterprise Channel: 2108+	16.43+ **	4.2111+	4.2111+	Yes	Yes
Audit label-related user activity: - Excludes encryption details	Current Channel: 2011+ Monthly Enterprise Channel: 2011+ Semi-Annual Enterprise Channel: 2108+	16.51+ **	4.2126+	4.2126+	Yes	Yes

Capability	Outlook for Windows	Outlook for Mac	Outlook on iOS	Outlook on Android	Outlook on the web	New Outlook for Windows*
Audit label-related user activity: - Includes encryption details	Current Channel: 2301+ Monthly Enterprise Channel: 2302+ Semi-Annual Enterprise Channel: 2302+	16.70+ **	4.2309+	4.2309+	Not available	Not available
Apply a sensitivity label to emails automatically - Using sensitive info types	Current Channel: 2009+ Monthly Enterprise Channel: 2009+ Semi-Annual Enterprise Channel: 2102+	16.44+ **	Not available	Not available	Yes	Yes
Apply a sensitivity label to emails automatically - Using trainable classifiers	Current Channel: 2105+ Monthly Enterprise Channel: 2105+ Semi-Annual Enterprise Channel: 2108+	16.49+	Not available	Not available	Yes	Yes
Different settings for default label and mandatory labeling	Current Channel: 2105+	16.43+ **	4.2111+	4.2111+	Yes	Yes

Capability	Outlook for Windows	Outlook for Mac	Outlook on iOS	Outlook on Android	Outlook on the web	New Outlook for Windows*
	Monthly Enterprise Channel: 2105+ Semi-Annual Enterprise Channel: 2108+					
PDF support	Current Channel: 2205+ Monthly Enterprise Channel: 2205+ Semi-Annual Enterprise Channel: 2302+	Not available	Not available	Not available	Not available	Not available
Apply S/MIME protection	Current Channel: 2211+ Monthly Enterprise Channel: 2211+ Semi-Annual Enterprise Channel: 2302+	16.61+ **	4.2226+	4.2203+	Yes	Yes
Sensitivity bar	Current Channel: 2302+ Monthly Enterprise Channel: 2303+	Not available	4.2316.0+	4.2316.0+	Yes	Yes

Capability	Outlook for Windows	Outlook for Mac	Outlook on iOS	Outlook on Android	Outlook on the web	New Outlook for Windows*
	Semi-Annual Enterprise Channel: 2302+					
Display label color	Current Channel: 2302+ Monthly Enterprise Channel: 2303+ Semi-Annual Enterprise Channel: 2302+	16.71+	4.2316.0+	4.2316.0+	Yes	Yes
Default sublabel for parent or group label	Current Channel: 2302+ Monthly Enterprise Channel: 2302+ Semi-Annual Enterprise Channel: 2302+	Not available	Not available	Not available	Not available	Not available
Scope labels to files or emails	Current Channel: 2302+ Monthly Enterprise Channel: 2302+ Semi-Annual Enterprise	16.70+ **	4.2309+	4.2309+	Yes	Yes

Capability	Outlook for Windows	Outlook for Mac	Outlook on iOS	Outlook on Android	Outlook on the web	New Outlook for Windows*
	Channel: 2302+					
Label inheritance from email attachments	Current Channel: 2302+ Monthly Enterprise Channel: 2302+ Semi-Annual Enterprise Channel: 2302+	Not available	Not available	Not available	Yes	Yes
Double Key Encryption (DKE)	Current Channel: 2307+ Monthly Enterprise Channel: 2309+ Semi-Annual Enterprise Channel: 2308+	Not available	Not available	Not available	Not available	Not available
Prevent some connected experiences that analyze content	Current Channel: 2406+ Monthly Enterprise Channel: 2406+ Semi-Annual Enterprise Channel: 2402+	Not available	Not available	Not available	Not available	Not available

Sensitivity label capabilities in OneNote

The numbers listed are the minimum OneNote application versions required for each capability.

> **Note**
>
> Tenant admins must opt in to enable sensitivity labels in OneNote within their tenants. Learn more

Expand table

Capability	Windows	Mac	iOS	Android	Web
Manually apply, change, or remove label	Current Channel: 16.0.19426+ Monthly Enterprise Channel: 16.0.19426+	16.103+	16.103+	16.0.19426+	Yes
Multi-language support	Current Channel: 16.0.19426+ Monthly Enterprise Channel: 16.0.19426+	16.103+	16.103+	16.0.19426+	Yes
Apply a default label to new documents	Not available	Not available	Not available	Not available	Not available
Apply a default label to existing documents	Not available	Not available	Not available	Not available	Not available
Require a justification to change a label	Current Channel: 16.0.19426+ Monthly Enterprise Channel: 16.0.19426+	16.103+	16.103+	16.0.19426+	Yes
Provide help link to a custom help page	Current Channel: 16.0.19426+ Monthly Enterprise Channel: 16.0.19426+	16.103+	16.103+	16.0.19426+	Yes

Capability	Windows	Mac	iOS	Android	Web
Mark the content	Current Channel: 16.0.19426+ Monthly Enterprise Channel: 16.0.19426+	16.103+	16.103+	16.0.19426+	Yes
Markings with variables	Current Channel: 16.0.19426+ Monthly Enterprise Channel: 16.0.19426+	16.103+	16.103+	16.0.19426+	Yes
Assign permissions now	Not available	Not available	Not available	Not available	Not available
Let users assign permissions: - Prompt users for custom permissions (users and groups)	Not available	Not available	Not available	Not available	Not available
Let users assign permissions: - Prompt users for custom permissions (users, groups, and organizations)	Not available	Not available	Not available	Not available	Not available
Audit label-related user activity: - Excludes encryption details	Current Channel: 16.0.19426+ Monthly Enterprise Channel: 16.0.19426+	16.103+	16.103+	16.0.19426+	Yes
Audit label-related user activity: - Includes encryption details	Current Channel: 16.0.19426+ Monthly Enterprise Channel: 16.0.19426+	16.103+	16.103+	16.0.19426+	Yes
Require users to apply a label to their email and documents	Not available	Not available	Not available	Not available	Not available
Apply a sensitivity label to files automatically - Using sensitive info types	Not available	Not available	Not available	Not available	Not available
Apply a sensitivity label to files automatically	Not available	Not available	Not available	Not available	Not available

Capability	Windows	Mac	iOS	Android	Web
- Using trainable classifiers					
Support co-authoring and AutoSave for labeled and encrypted documents	Current Channel: 16.0.19426+ Monthly Enterprise Channel: 16.0.19426+	16.103+	16.103+	16.0.19426+	Yes
PDF support	Current Channel: 16.0.19426+ Monthly Enterprise Channel: 16.0.19426+	16.103+	16.103+	16.0.19426+	Not available
Sensitivity bar and display label color	Current Channel: 16.0.19426+ Monthly Enterprise Channel: 16.0.19426+	16.103+	16.103+	16.0.19426+	Yes
Default sublabel for parent label	Current Channel: 16.0.19426+ Monthly Enterprise Channel: 16.0.19426+	16.103+	16.103+	16.0.19426+	Yes
Scope labels to files or emails	Current Channel: 16.0.19426+ Monthly Enterprise Channel: 16.0.19426+	16.103+	16.103+	16.0.19426+	Yes
Double Key Encryption (DKE)	Not available	Not available	Not available	Not available	Not available
Document tracking and revocation	Not available	Not available	Not available	Not available	Not available
Prevent connected experiences that analyze content	Not available	Not available	Not available	Not available	Not available
Dynamic watermarks	Not available	Not available	Not available	Not available	Not available

Footnotes:

* For information about the new Outlook for Windows, see Getting started with the new Outlook for Windows

** Requires the new Outlook for Mac

Future releases

For information about new labeling capabilities that are planned for future releases, use the Microsoft 365 roadmap.

Last updated on 08/14/2026

Extend sensitivity labeling on Windows

The Microsoft Purview Information Protection client extends sensitivity labels beyond labels that are built into Microsoft 365 apps and services, and supports a wider range of file types.

This client runs on Windows only and replaces the Azure Information Protection (AIP) unified labeling client. It has the following components:

Expand table

Component	Description
Information protection scanner	Used to discover, label, and encrypt files on data stores such as network shares and SharePoint Server libraries.
Information protection file labeler	Used to apply sensitivity labels and encryption using File Explorer.
Information protection viewer	Used to view files that are encrypted.
Microsoft Purview Information Protection PowerShell module	Used to adjust sensitivity labels on files, and install and configure Microsoft Purview Information Protection scanner.

There's no Office Add-in with the Microsoft Purview Information Protection client because this functionality is replaced with sensitivity labels that are built into Office.

For the latest release information and support timelines for each client version, see Microsoft Purview Information Protection client - Release management and supportability.

Requirements for deploying the information protection client

To use the Microsoft Purview Information Protection client, install this client on Windows computers where you want to use the client components.

You also must meet the following requirements:

Operating Systems

The following operating systems support the Microsoft Purview Information Protection client:

- **Windows 11**, including Windows 11 Enterprise multi-session
- **Windows 10 (x64)** (Handwriting isn't supported in the Windows 10 RS4 build and later)
- **Windows Server 2019**
- **Windows Server 2016**

ARM64 is supported for the viewer and file labeler, but isn't supported for the scanner or PowerShell.

Install or upgrade the information protection client

If you install the Microsoft Purview Information client interactively and the Azure Information Protection (AIP) unified labeling client is detected, you can upgrade the older client after you acknowledge the AIP Add-in for Office will be removed.

> ⓘ **Note**
>
> Upgrades by using Microsoft Update Catalog or any other non-interactive installation require a registry key configuration if any Azure Information Protection client versions are present on the local computer.

There are two options for installing the information protection client:

- Installing the information protection client using the **.exe** installer
- Installing the information protection client using the **.msi** installer

If you're upgrading the information protection scanner, either from the Azure Information Protection (AIP) unified labeling client or a previous version of the information protection client, see Upgrade the Microsoft Purview Information Protection scanner before using these client installation instructions.

Install using the .exe installer

To install the information protection client with the .exe file:

1. Download the executable version of the Microsoft Purview Information Protection client from the Microsoft Download Center. For example, **PurviewInfoProtection.exe**.

   > ⓘ **Important**

If there is a preview version available, use that version only for testing. It is not intended for end users in a production environment.

2. For a default installation, simply run the executable. To view all installation options, first run the executable with **/help**. For example:

   `PurviewInfoProtection.exe **/help**`

   a. To silently install the client, run:

      `PurviewInfoProtection.exe /quiet`

   b. To silently install only the PowerShell cmdlets, run:

      `PurviewInfoProtection.exe PowerShellOnly=true /quiet`

   > **Note**
   >
   > By default, the option to send usage statistics to Microsoft is enabled. To disable this option, make sure to take one of the following steps:
   >
   > - During installation, specify **AllowTelemetry=0**
   > - After installation, update the registry key as follows: **EnableTelemetry=0**.

3. To complete the installation, restart any instances of File Explorer.

4. Confirm that the installation was successful by checking the install log file, which is created in the **%temp%** folder by default.

   The install log file has the following naming format:

   `Microsoft_Azure_Information_Protection_<number>_<number>_MSIP.Setup.Main.msi.log`

   For example:

   Microsoft_Azure_Information_Protection_20161201093652_000_MSIP.Setup.Main.msi.log

   In the log file, search for the following string: **Product: Microsoft Purview Information Protection--Installation completed successfully**. If the installation failed, this log file contains details to help you identify and resolve any problems.

   > **Tip**
   >
   > You can change the location of the installation log file with the **/log** installation parameter.

Log file locations

Client and scanner log files are located in the following locations on the Windows computer:

- **\ProgramFiles (x86)\Microsoft Purview Information Protection** (64-bit operating systems only)
- **\Program Files\Microsoft Purview Information Protection** (32-bit operating systems only)
- **%localappdata%\Microsoft\MSIP**

Supported languages

The information protection client supports the same languages that Office 365 supports. For a list of these languages, see the International availability page from Office.

For these languages, menu options, dialog boxes, and messages from the Microsoft Purview Information Protection client display in the user's language. There's a single installer that detects the language, so no additional configuration is required to install the information protection client for different languages.

However, label names and descriptions that you specify aren't automatically translated when you configure labels in the admin portal. For users to see labels in their preferred language, provide your own translations and configure them for the labels by using PowerShell and the *LocaleSettings* parameter for Set-Label. For more information, see Example configuration to configure a sensitivity label for different languages.

Supported file types

This section lists the file types supported by the Microsoft Purview Information Protection client. For the listed file types, WebDav locations aren't supported.

> **Tip**
>
> When you encrypt file types that don't have built-in support for encryption and so use generic encryption, we recommend that you assign the permission of co-owner to these files.

Sensitivity labels without encryption

The following file types can be labeled without encryption.

- **Adobe Portable Document Format:** .pdf
- **Microsoft Project:** .mpp, .mpt
- **Microsoft Publisher:** .pub
- **Microsoft XPS:** .xps .oxps
- **Images:** .jpg, .jpe, .jpeg, .jif, .jfif, .jfi. png, .tif, .tiff
- **Autodesk Design Review 2013:** .dwfx
- **Adobe Photoshop:** .psd
- **Digital Negative:** .dng
- **Microsoft Office:** The following file types, including 97-2003 file formats and Office Open XML formats for Word, Excel, and PowerPoint.

Expand table

Word	Excel	PowerPoint	Visio
.doc	.xls	.potm	.vdw
.docm	.xlsb	.potx	.vsd
.docx	.xlsm	.pps	.vsdm
.dot	.xlsx	.ppsm	.vsdx
.dotx	.xltm	.ppsx	.vss
	.xltx	.ppt	.vssm
		.pptx	.vst
			.vstm
			.vssx
			.vstx

Excluded folders and file types

To help prevent users from changing files that are critical for computer operations, some file types and folders are automatically excluded from classifying and labeling. If users try to label these files by using the information protection client, they see a message that those files are excluded.

Excluded folders

The following folders are excluded from classifying and labeling by the information protection client:

- Windows
- Program Files (\Program Files and \Program Files (x86))
- \ProgramData
- \AppData (for all users)

File types that can't be encrypted by default

Any file that is password-protected can't be natively encrypted by the client unless the file is currently open in the application that applies the encryption. You most often see PDF files that are password-protected but other applications, such as Office apps, also offer this functionality.

File types supported for inspection

The information protection client uses Windows IFilter to inspect the contents of documents. Windows IFilter is used by Windows Search for indexing. As a result, the following file types can be inspected when you use the Set-FileLabel -Autolabel PowerShell command.

Expand table

Application type	File type
Word	.doc, .docx, .docm, .dot, .dotx
Excel	.xls, .xlt, .xlsx, .xlsm, .xlsb
PowerPoint	.ppt, .pps, .pot, .pptx
PDF	.pdf
Text	.txt, .xml, .csv

Scanning .ZIP files

You can use the information protection scanner or the Set-FileLabel PowerShell command to inspect .zip files.

> **Note**
>
> When your information protection scanner is installed on a Windows server computer, you must also install the Microsoft Office iFilter in order to scan .zip files for sensitive information types. For more information, see the Microsoft download site.

After finding sensitive information, if the .zip file should be labeled and encrypted with a label, from the scanner deployment instructions, specify the .zip file name extension with the PowerShell **PFileSupportedExtensions** advanced setting.

Example scenario:

A file named *accounts.zip* contains Excel spreadsheets with credit card numbers. You have a sensitivity label named *Confidential \ Finance*, which is configured to discover credit card numbers and automatically apply the label with encryption that restricts access to the Finance group.

After inspecting the file, the client from your PowerShell session labels this file as *Confidential \ Finance*. Next, the client applies generic encryption to the file so that only members of the Finance groups can unzip it, and renames the file *accounts.zip.pfile*.

Integration with Endpoint data loss prevention

The Endpoint data loss prevention (DLP) setting **Advanced label-based protection for all files on devices** enhances productivity for users who label and encrypt file types other than those for Office and PDF.

For these files, although access and any usage restrictions from the encryption settings are enforced by DLP when the label is applied, the encryption is applied when the files leave the computer. As a result, the file name extension for the local files don't change and users can continue to view and edit the files in their standard application.

For more information, see Learn about Advanced Label Based Protection.

Support for disconnected computers

By default, the information protection client automatically tries to connect to the internet to download sensitivity labels and sensitivity label policy settings from Microsoft Purview.

If you have computers that can't connect to the internet for a period of time, you can export and copy files that manually manages the policy for the information protection client.

To support disconnected computers from the information protection client:

1. Choose or create a user account in Microsoft Entra ID that you will use to download labels and policy settings that you want to use on your disconnected computer.

2. As an additional label policy setting for this account, turn off sending audit data to Microsoft Purview by using the *EnableAudit* PowerShell advanced setting with Set-LabelPolicy from Security & Compliance PowerShell.

   We recommend this step because if the disconnected computer does have periodic internet connectivity, it will send logging information to Microsoft Purview that includes the user name from step 1. That user account might be different from the local account you're using on the disconnected computer.

3. From a computer with internet connectivity that has the information protection client installed and signed in with the user account from step 1, download the labels and policy settings.

4. From this computer, export the log files.

   For example, run the Export-DebugLogs cmdlet, or use the **Export Logs** option from the client's **Help and Feedback** dialog box from file labeler.

   The log files are exported as a single compressed file.

5. Open the compressed file, and from the MSIP folder, copy any files that have an .xml file name extension.

6. Paste these files into the **%localappdata%\Microsoft\MSIP** folder on the disconnected computer.

7. If your chosen user account is one that usually connects to the internet, enable sending audit data again, by setting the **EnableAudit** value to **True**.

Be aware that if a user on this computer selects the **Reset Settings** option from **Help and feedback** in the file labeler, this action deletes the policy files and leaves the client inoperable

until you manually replace the files or the client connects to the internet so it can download the files it needs.

The Information Protection client requires an active internet connection to apply encryption; therefore, the disconnected computers solution is supported only for labels that do not apply encryption.

If your disconnected computer is running the information protection scanner, there are additional configuration steps you must take. For more information, see Restriction: The scanner server cannot have internet connectivity from the scanner deployment instructions.

Supported customizations

The information protection client supports PowerShell advanced settings and some registry settings that might be needed for specific scenarios or users.

For the PowerShell advanced settings that are supported with New-Label or Set-Label, and New-LabelPolicy or Set-LabelPolicy from Security & Compliance PowerShell, see Advanced settings for Microsoft Purview Information Protection client.

Use the following sections to help you configure the registry for supported customizations.

Enable non-interactive upgrade from the Azure Information Protection client

If you install the Microsoft Purview Information Protection client and the Azure Information Protection unified labeling client is detected, an interactive installation of the client requires you to acknowledge that the AIP Add-in for Office from the older client will be removed.

To use a non-interactive installation for the client, such as Microsoft Update Catalog, Group Policy, or scripting, you must either first uninstall the Azure Information Protection client, or create and configure the following registry key for the local computer:

HKEY_LOCAL_MACHINE\SOFTWARE\WOW6432Node\Microsoft\MSIP\AllowMajorVersionUpgrade (DWORD)

Set the value to **1** to silently allow the upgrade and uninstall the AIP Office add-in; **0** blocks the upgrade if the Azure Information Protection client is installed.

Change the local logging level

By default, the Purview Information Protection client writes client log files to the **%localappdata%\Microsoft\MSIP** folder. These files are intended for troubleshooting by Microsoft Support.

To change the logging level for these files, locate the following value name in the registry and set the value data to the required logging level:

HKEY_CURRENT_USER\SOFTWARE\Microsoft\MSIP\LogLevel

Set the logging level to one of the following values:

- **Off**: No local logging.
- **Error**: Errors only.
- **Warn**: Errors and warnings.
- **Info**: Minimum logging, which includes no event IDs (the default setting for the scanner).
- **Debug**: Full information.
- **Trace**: Detailed logging (the default setting for clients).

This registry setting doesn't change the information that's sent to Microsoft Purview auditing.

Enable data boundary settings

Following Microsoft's commitment to EU data boundary, EU customers who use the Microsoft Purview Information Protection client can send their data to the EU to be stored and processed.

Turn on this feature in the information protection client by changing the following registry key that specifies the location to send events:

- Registry key: **HKEY_LOCAL_MACHINE\SOFTWARE\WOW6432Node\Microsoft\MSIP\DataBoundary (DWORD)**
- Values:
  - `Default = 0`
  - `North_America = 1`
  - `European_Union = 2`

Enable system default browser for authentication

Use the system default browser for authentication in Microsoft Purview Information Protection client. By default, the information protection client opens Microsoft Edge for authentication.

Turn on this feature in the information protection client by enabling the following registry key:

- Registry key: **HKEY_LOCAL_MACHINE\SOFTWARE\WOW6432Node\Microsoft\MSIP\MSALUseSytemDefaultBrowserAuth (DWORD)**
- Values:
  - `Disabled = 0`
  - `Enabled = 1`

Hide the "Apply sensitivity label with Microsoft Purview" menu option in File Explorer

To hide the **Apply sensitivity label with Microsoft Purview** right-click menu option in File Explorer, create the following DWORD registry key that has the value name of **LegacyDisable** and any value data:

HKEY_CLASSES_ROOT\AllFilesystemObjects\shell\Microsoft.Azip.RightClick

How to apply sensitivity labels using the information protection client

After the Microsoft Purview Information Protection client is installed, you can apply sensitivity labels that you've created and published, by using:

- File Explorer
- PowerShell
- The information protection scanner, after installing and configuring it

> **Important**
>
> If you are upgrading from the Azure Information Protection (AIP) unified labeling client, see **Upgrade the scanner from the Azure Information Protection client (version 2.x)**.

To view encrypted documents, see View protected files with the Microsoft Purview Information Protection viewer.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/22/2026

Use Microsoft Purview sensitivity label policies for non-Microsoft connected apps (preview)

> **Important**
>
> This feature is currently in preview. The Supplemental Terms of Use for Microsoft Azure Previews include additional legal terms that apply to Azure features that are in beta, in preview, or otherwise not yet released into general availability.

Auto-labeling isn't limited to Microsoft 365 locations. Through the integration between Microsoft Purview Information Protection and Microsoft Defender for Cloud Apps, you can automatically apply your Microsoft Purview sensitivity labels to files stored in non-Microsoft connected apps, such as Box and Google Workspace. Because these labels use the same sensitive information types and classifiers similar to Microsoft 365, your sensitive content is classified and protected consistently, wherever it's stored.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Supported non-Microsoft apps

Microsoft Purview supports auto-labeling policies for the following non-Microsoft apps:

- Box
- Google Workspace

> **Note**
>
> These apps are rolling out in phases. Not all apps may be available in your tenant at the same time. Check the Microsoft 365 roadmap for the latest availability information.

Before you begin

Licensing requirements

For information on licensing, see

- Microsoft 365 Enterprise Plans
- Microsoft 365 Service Descriptions

Permissions

The account you use to create and deploy policies must be a member of one of these role groups:

- Compliance administrator
- Compliance data administrator
- Information Protection
- Information Protection Admin
- Security administrator

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. This helps improve security for your organization. Global Administrator is a highly privileged role that should only be used in scenarios where a lesser privileged role can't be used.

Publish labels

To select sensitivity labels when setting up auto-labeling policies, the labels must first be published. For more information on creating and publishing labels, see Create and configure sensitivity labels and their policies.

Set up a Microsoft Defender for Cloud Apps connector

Before you can apply auto-labeling policies to a non-Microsoft app, you must connect that app to Microsoft Defender for Cloud Apps using an app connector. Purview uses the existing Microsoft Defender for Cloud Apps connectors to access capabilities in the non-Microsoft app.

For instructions on setting up app connectors, see Connect apps to get visibility and control with Microsoft Defender for Cloud Apps.

After you connect your cloud apps to Defender for Cloud Apps, you can create auto-labeling policies for them in Microsoft Purview.

Create an auto-labeling policy for non-Microsoft connected apps

You can use auto-labeling policies to automatically apply or remove sensitivity labels on data stored in non-Microsoft connected apps. For the general procedure to create an auto-labeling policy, see Automatically apply a sensitivity label to Microsoft 365 data.

To create an auto-labeling policy scoped to non-Microsoft connected apps:

1. Sign in to the Microsoft Purview portal.
2. Go to **Solutions** > **Information protection** > **Policies** > **Auto-labeling policies** > **+ Create auto-labeling policy**.
3. On **What type of auto-labeling policy do you want to create?**, choose **Automatically apply labels only**.
4. Select **Custom** > **Custom policy** > **Next**.

   > **Important**
   >
   > Non-Microsoft connected app policies are only supported with the **Custom** policy template. The predefined Financial, Medical and health, and Privacy templates don't support non-Microsoft app locations.

5. Enter a name and description for the auto-labeling policy, then select **Next**.
6. On **Choose a label to auto-apply**, choose **Choose a label**. On **Choose a sensitivity label**, select all labels to include in the policy, and then select **Add** > **Next**.
7. On **Assign admin units**, the scope is set to **Full directory**. Administrative units aren't supported for non-Microsoft app locations.
8. On **Choose locations where you want to apply the label**, select one or more of the supported non-Microsoft apps (such as **Box** or **Google Workspace**).

   > **Important**

Non-Microsoft app locations can be selected together, but they can't be combined with other locations like SharePoint, Exchange, SQL Server, Azure Storage or OneDrive in the same policy.

9. By default, the policy applies to all instances of the selected app. To scope the policy to specific tenants, select **Edit**.
   a. To exclude instances, in **All instances**, select **Exclude instances** > **+ Exclude instance**. From the list of instances, select instances to exclude from the policy, and then select **Done** > **Done**.
   b. To apply the policy to specific instances, select **Specific instances** > **+ Include instance**. From the list of instances, select instances to include, then select **Done** > **Done**.
10. Select **Next** to proceed.
11. On **Set up common or advanced rules**, begin creating rules by selecting **Advanced Rules**. Select **Next**.
12. Select **+ New rule**, and provide a **Name** and **Description**.
13. Configure **Conditions** for the rule. The conditions you can use vary by app.
14. Select **Save** to save the rule, then select **Next**.
15. On **Additional label settings**, optionally select **All locations** to replace existing labels with lower priority in the selected applications. Select **Next**.
16. On **Decide if you want to test out the policy now or later**, choose either **Turn on policy** or **Leave policy turned off**. Select **Next**.

    > ⓘ **Important**
    >
    > **Simulation mode** isn't supported for non-Microsoft connected app policies.

17. Review your settings and select **Create policy** to create the policy.

Attribute availability by app

The conditions and actions you can use in a policy depend on the item attributes that each app makes available. Most attributes are available for all supported apps, but some aren't available for

certain apps. Use the following table to understand which attributes you can rely on when you build rules for a specific app.

Expand table

Attribute	Description	Availability
Name	The display name of the item.	Available in all apps
File access level	The item's access level, which indicates its sharing scope.	Available in all apps
Created date	The date the item was created.	Not available
Modified date	The date the item was last modified.	Available in all apps
Collaborators	The users the item is shared with, including their roles.	Not available
Parent folders	The parent folders that contain the item.	Available in all apps

Monitoring and reporting

Activity explorer

Activity explorer supports non-Microsoft connected app policies. You can monitor labeling policy activity for your non-Microsoft apps in the same way you monitor Microsoft 365 locations.

Content explorer

Content explorer isn't supported for non-Microsoft connected apps. You can't browse or inspect content stored in these apps through Content explorer.

Known issues

The following is a list of known issues when using non-Microsoft connected apps for sensitivity label policies:

- **Encrypted files** – Encrypted files stored in non-Microsoft apps aren't currently supported for classification.
- **PDF files** – PDF files stored in non-Microsoft apps aren't currently supported for classification.
- In Microsoft Purview for Microsoft 365, when a sensitivity label is applied automatically through an auto-labeling policy, a higher-priority label replaces any existing lower-priority label. However, for Box and Google Workspace, if a label has already been applied to a file

by an auto-labeling policy, it cannot be replaced by a higher-priority label from **another** auto-labeling policy. For more information, see Automatically apply a sensitivity label to Microsoft 365 data.

Related content

- Automatically apply a sensitivity label to Microsoft 365 data
- Connect apps to get visibility and control with Microsoft Defender for Cloud Apps

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 08/28/2026

Learn about the information protection scanner

> **Note**
>
> There's a new version of the information protection scanner. For more information, see Upgrade the Microsoft Purview Information Protection scanner.

Use the information in this section to learn about the Microsoft Purview Information Protection scanner, and then how to successfully install, configure, run and if necessary, troubleshoot it.

This scanner runs as a service on Windows Server and lets you discover, classify, and protect files on the following data stores:

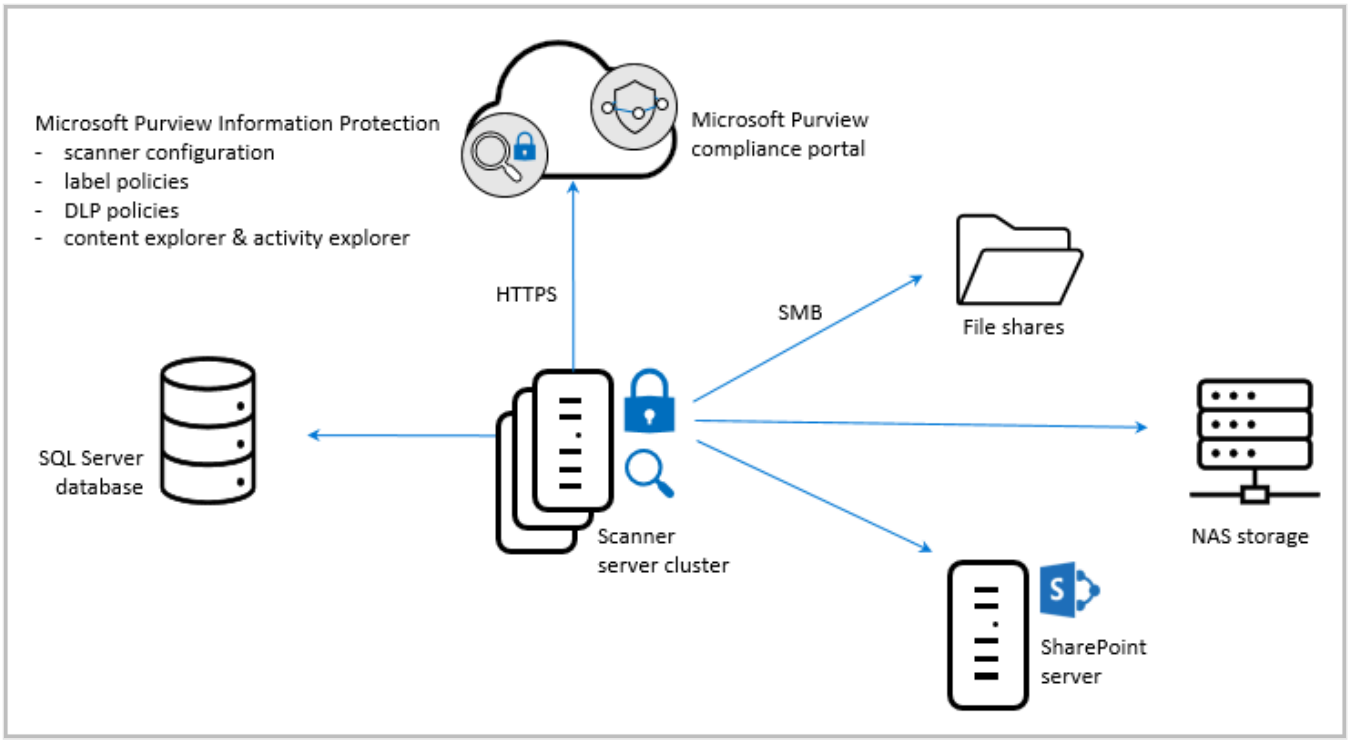
- **UNC paths** for network shares that use the SMB or NFS (Preview) protocols.
- **SharePoint document libraries and folder** for SharePoint Server (versions currently in mainstream support). For more information see the SharePoint Server Support Policy updated product servicing policy for SharePoint Server SE.

To classify and protect your files, the scanner uses sensitivity labels configured in the Microsoft Purview portal.

Overview of the scanner

The information protection scanner can inspect any files that Windows can index. If you configure sensitivity labels to apply automatic classification, the scanner can label discovered files to apply that classification, and optionally apply or remove protection. For information about the sensitive information types (SITs) supported by the information protection scanner, see Sensitive Information Types supported by Microsoft Purview Information Protection scanner.

The following image shows the scanner architecture, where the scanner discovers files across your on-premises and SharePoint servers.



To inspect your files, the scanner uses IFilters installed on the computer. To determine whether the files need labeling, the scanner uses sensitive information types and pattern detection, or regex patterns.

The scanner uses the Microsoft Purview Information Protection client, and can classify and protect the same types of files as the client. For more information, see Supported file types.

Do any of the following to configure your scans as needed:

- **Run the scanner in discovery mode only** to create reports that check to see what happens when your files are labeled.
- **Run the scanner to discover files with sensitive information**, without configuring labels that apply automatic classification.
- **Run the scanner automatically** to apply labels as configured.
- **Define a file types list** to specify specific files to scan or to exclude.

> **Note**
>
> The scanner does not discover and label in real time. It systematically crawls through files on data stores that you specify. Configure this cycle to run once, or repeatedly.

> **Tip**
>
> The scanner supports scanner clusters with multiple nodes, enabling your organization to scale out, achieving faster scan times and broader scope.

Deploy multiple nodes right from the start, or start with a single-node cluster and add additional nodes later on as you grow. Deploy multiple nodes by using the same cluster name and database for the **Install-Scanner** cmdlet.

The scanning process

When scanning files, the information protection scanner runs through the following steps:

1. Determine whether files are included or excluded for scanning.

2. Inspect and label files.

3. Label files that can't be inspected.

For more information, see Files not labeled by the scanner.

1. Determine whether files are included or excluded for scanning

The scanner automatically skips files that are excluded from classification and protection, such as executable files and system files. For more information, see Excluded file types.

The scanner also considers any file lists explicitly defined to scan, or exclude from scanning. File lists apply for all data repositories by default, and can also be defined for specific repositories only.

To define file lists for scanning or exclusion, use the **File types to scan** setting in the content scan job. For example:



For more information, see Deploying the scanner to automatically classify and protect files.

2. Inspect and label files

After identifying excluded files, the information protection scanner filters again to identify files supported for inspection.

These filters are the same ones used by the operating system for Windows Search and indexing, and require no extra configuration. Windows IFilter is also used to scan file types that are used by Word, Excel, and PowerPoint, and for PDF documents and text files.

For a full list of file types supported for inspection, and other instructions for configuring filters to include .zip and .tiff files, see File types supported for inspection.

After inspection, supported file types are labeled using the conditions specified for your labels. If you're using discovery mode, these files can either be reported to contain the conditions specified for your labels, or reported to contain any known sensitive information types.

Stopped scanner processes

If the scanner stops before completing a scan for a large number of files in your repository, you may need to increase the number of dynamic ports for the operating system that hosts the files.

For example, server hardening for SharePoint is one reason why the scanner would exceed the number of allowed network connections, and therefore stop.

To check whether server hardening for SharePoint is the cause of the scanner stopping, check for the following error message in the scanner logs at **%localappdata%\Microsoft\MSIP\Logs\MSIPScanner.iplog** (multiple logs are compressed into a zip file):

```
Unable to connect to the remote server ---> System.Net.Sockets.SocketException: Only one usage of each socket address (protocol/network address/port) is normally permitted IP:port
```

For more information about how to view the current port range and increase it if needed, see Settings that can be modified to improve network performance.

> **Tip**
>
> For large SharePoint farms, you may need to increase the list view threshold, which has a default of **5,000**.
>
> For more information, see the **Manage large lists and libraries in SharePoint**.

3. Label files that can't be inspected

For any file types that can't be inspected, the scanner applies the default label from its sensitivity label policy, or the default label configured for the scanner.

Files not labeled by the scanner

The scanner can't label files under the following circumstances:

- When the file type doesn't support labeling without encryption. For a list of file types that are supported for this label configuration, see Sensitivity labels without encryption.
- When the label applies encryption, but the scanner doesn't support the file type.

  By default, the scanner encrypts just Office file types, and PDF files when they are protected by using the ISO standard for PDF encryption.

  Other types of files can be added for encryption when you change the types of files to protect.

Example: After inspecting .txt files, the scanner can't apply a label that doesn't apply encryption, because the .txt file type doesn't support sensitivity labels without encryption.

However, if the label is configured to apply encryption, and the .txt file type is included for the scanner to protect, the scanner can label the file.

Next steps

For more information about deploying the scanner, see the following articles:

- Scanner deployment prerequisites
- Configuring and installing the scanner
- Running scans using the scanner

You can also use PowerShell to interactively classify and protect files from your desktop computer. For more information about this and other scenarios that use PowerShell, see Set up the information protection client using PowerShell.

Last updated on 03/26/2026

Get started with the information protection scanner

Article • 04/17/2025

> **Note**
>
> There's a new version of the information protection scanner. For more information, see Upgrade the Microsoft Purview Information Protection scanner.

Before installing the scanner from Microsoft Purview Information Protection, make sure that your system complies with basic Azure Information Protection requirements.

Additionally, the following requirements are specific for the scanner:

- Windows Server requirements
- Service account requirements
- SQL server requirements
- Information protection client requirements
- Label configuration requirements
- SharePoint requirements
- Microsoft Office requirements
- File path requirements

If you can't meet all the requirements listed for the scanner because they are prohibited by your organization's policies, see the alternative configurations section.

When deploying the scanner in production or testing the performance for multiple scanners, see Storage requirements and capacity planning for SQL Server.

When you're ready to start installing and deploying your scanner, continue with Configuring and installing the information protection scanner.

Windows Server requirements

You must have a Windows Server computer to run the scanner, which has the following system specifications:

Expand table

Specification	Details
Processor	4 core processors
RAM	8 GB
Disk space	10-GB free space (average) for temporary files. The scanner requires sufficient disk space to create temporary files for each file that it scans, four files per core. The recommended disk space of 10 GB allows for 4 core processors scanning 16 files that each have a file size of 625 MB.
Operating system	64-bit versions of: - Windows Server 2025 - Windows Server 2022 - Windows Server 2019 - Windows Server 2016 **Note**: For testing or evaluation purposes in a non-production environment, you can also use any Windows operating system that is supported by the information protection client. Server Core and Nano Server aren't supported.
- Network connectivity	Your scanner computer can be a physical or virtual computer with a fast and reliable network connection to the data stores to be scanned. If internet connectivity is not possible because of your organization policies, see Deploying the scanner with alternative configurations. Otherwise, make sure that this computer has internet connectivity that allows the following URLs over HTTPS (port 443): - *.aadrm.com - *.azurerms.com - *.informationprotection.azure.com - informationprotection.hosting.portal.azure.net - *.aria.microsoft.com - *.protection.outlook.com
NFS shares	To support scans on NFS shares, services for NFS must be deployed on the scanner machine. On your machine, navigate to the **Windows Features (Turn Windows features on or off)** settings dialog, and select the following items: **Services for NFS** > **Administrative Tools** and **Client for NFS**.

Specification	Details
Microsoft Office iFilter	When your scanner is installed on a Windows server machine, you must also install the Microsoft Office iFilter in order to scan .zip files for sensitive information types. For more information, see the Microsoft download site .

Service account requirements

You must have a service account to run the scanner service on the Windows Server computer, as well as authenticate to Microsoft Entra ID and download the scanner's policy.

Your service account must be an Active Directory account and synchronized to Microsoft Entra ID.

If you cannot synchronize this account because of your organization policies, see Deploying the scanner with alternative configurations.

This service account has the following requirements:

Expand table

Requirement	Details
Log on locally user right assignment	Required to install and configure the scanner, but not required to run scans. Once you've confirmed that the scanner can discover, classify, and protect files, you can remove this right from the service account. If granting this right even for a short period of time is not possible because of your organization policies, see Deploying the scanner with alternative configurations.
Log on as a service user right assignment.	This right is automatically granted to the service account during the scanner installation and this right is required for the installation, configuration, and operation of the scanner.
Permissions to the data repositories	- **File shares or local files**: Grant **Read**, **Write**, and **Modify** permissions for scanning the files and then applying classification and protection as configured. - **SharePoint**: You must grant **Full Control** permissions for scanning the files and then applying classification and protection to the files that meet the conditions in the Azure Information Protection policy. - **Discovery mode**: To run the scanner in discovery mode only, **Read** permission is sufficient.

Requirement	Details
For labels that reprotect or remove protection	To ensure that the scanner always has access to encrypted files, make this account a super user for Azure Information Protection, and ensure that the super user feature is enabled. Additionally, if you've implemented onboarding controls for a phased deployment, make sure that the service account is included in the onboarding controls you've configured.
Specific URL level scanning	To scan and discover sites and subsites under a specific URL, grant **Site Collector Auditor** rights to the scanner account on the farm level.
License for information protection	Required to provide file classification, labeling, or protection capabilities to the scanner service account. For more information, see the Microsoft 365 guidance for security & compliance.

SQL server requirements

To store the scanner configuration data, use an SQL server with the following requirements:

- **A local or remote instance.**

  We recommend hosting the SQL server and the scanner service on different machines, unless you're working with a small deployment. Additionally, we recommend having a dedicated SQL instance that serves the scanner database only, and that is not shared with other applications.

  If you're working on a shared server, make sure that the recommended number of cores are free for the scanner database to work.

  SQL Server 2016 is the minimum version for the following editions:

  - SQL Server Enterprise
  - SQL Server Standard
  - SQL Server Express (recommended for test environments only)

- **An account with Sysadmin role to install the scanner.**

  The Sysadmin role enables the installation process to automatically create the scanner configuration database and grant the required **db_owner** role to the service account that runs the scanner.

If you cannot be granted the Sysadmin role or your organization policies require databases to be created and configured manually, see Deploying the scanner with alternative configurations.

- **Capacity.** For capacity guidance, see Storage requirements and capacity planning for SQL Server.
- **Case insensitive collation**.

> **Note**
>
> Multiple configuration databases on the same SQL server are supported when you specify a custom cluster name for the scanner, or when you use the preview version of the scanner.

Storage requirements and capacity planning for SQL Server

The amount of disk space required for the scanner's configuration database and the specification of the computer running SQL Server can vary for each environment, so we encourage you to do your own testing. Use the following guidance as a starting point.

For more information, see Optimizing the performance of the scanner.

The disk size for the scanner configuration database will vary for each deployment. Use the following equation as guidance:

```cli
100 KB + <file count> *(1000 + 4* <average file name length>)
```

For example, to scan 1 million files that have an average file name length of 250 bytes, allocate 2-GB disk space.

For multiple scanners:

- **Up to 10 scanners**, use:
  - 4 core processors
  - 8-GB RAM recommended
- **More than 10 scanners** (maximum 40), use:
  - 8 core processes
  - 16-GB RAM recommended

Information protection client requirements

For a production network, you must have the current general availability version of the Microsoft Purview Information Protection client installed on the Windows Server computer.

For more information, see the Install or upgrade the information protection client.

> **Important**
>
> You must install the full client for the scanner. Do not install the client with just the PowerShell module.

Label configuration requirements

You must have at least one sensitivity label configured in the Microsoft Purview portal for the scanner account, to apply classification and, optionally, encryption.

The *scanner account* is the account that you'll specify in the **DelegatedUser** parameter of the Set-Authentication cmdlet, run when configuring your scanner.

If your labels don't have auto-labeling conditions, see the instructions for alternative configurations below.

For more information, see:

- Learn about sensitivity labels
- Automatically apply a sensitivity label to Microsoft 365 data
- Restrict access to content by using encryption in sensitivity labels
- Configuring and installing the information protection scanner

SharePoint requirements

To scan SharePoint document libraries and folders, ensure that your SharePoint server complies with the following requirements:

Expand table

Requirement	Description
Supported versions	Supported versions include: SharePoint 2019, SharePoint 2016, and SharePoint 2013. Other versions of SharePoint are not supported for the scanner.

Requirement	Description
Versioning	When you use versioning, the scanner inspects and labels the last published version. If the scanner labels a file and content approval is required, that labeled file must be approved to be available for users.
Large SharePoint farms	For large SharePoint farms, check whether you need to increase the list view threshold (by default, 5,000) for the scanner to access all files. For more information, see Manage large lists and libraries in SharePoint.
Long file paths	If you have long file paths in SharePoint, ensure that your SharePoint server's httpRuntime.maxUrlLength value is larger than the default 260 characters. For more information, see the next section, Avoid scanner timeouts in SharePoint.

Avoid scanner timeouts in SharePoint

If you have long file paths in SharePoint version 2013 or higher, ensure that your SharePoint server's httpRuntime.maxUrlLength value is larger than the default 260 characters.

This value is defined in the **HttpRuntimeSection** class of the `ASP.NET` configuration.

To update the HttpRuntimeSection class:

1. Back up your **web.config** configuration.
2. Update the **maxUrlLength** value as needed. For example:

   ```
   c#
   ```

   ```
   <httpRuntime maxRequestLength="51200" requestValidationMode="2.0"
   maxUrlLength="5000"  />
   ```

3. Restart your SharePoint web server and verify that it loads correctly.

   For example, in Windows Internet Information Servers (IIS) Manager, select your site, and then under **Manage Website**, select **Restart**.

Microsoft Office requirements

To scan Office documents, your documents must be in one of the following formats:

- Microsoft Office 97-2003
- Office Open XML formats for Word, Excel, and PowerPoint

For more information, see Supported file types.

File path requirements

By default, to scan files, your file paths must have a maximum of 260 characters.

To scan files with file paths of more than 260 characters, install the scanner on a computer with one of the following Windows versions, and configure the computer as needed:

Expand table

Windows version	Description
Windows 2016 or later	Configure the computer to support long paths
Windows 10 or Windows Server 2016	Define the following group policy setting: **Local Computer Policy** > **Computer Configuration** > **Administrative Templates** > **All Settings** > **Enable Win32 long paths**. For more information long file path support in these versions, see the Maximum Path Length Limitation section from the Windows 10 developer documentation.
Windows 10, version 1607 or later	Opt in for the updated **MAX_PATH** functionality. For more information, see Enable Long Paths in Windows 10 versions 1607 and later.

Deploying the scanner with alternative configurations

The prerequisites listed above are the default requirements for the scanner deployment, and recommended because they support the simplest scanner configuration.

The default requirements should be suitable for initial testing, so that you can check the capabilities of the scanner.

However, in a production environment, your organization's policies may be different than the default requirements. The scanner can accommodate the following changes with additional configuration:

- Discover and scan all sites and subsites under a specific URL
- Restriction: The scanner server cannot have internet connectivity

- Restriction: The scanner service account cannot be synchronized to Microsoft Entra ID but the server has internet connectivity
- Restriction: The service account for the scanner cannot be granted the **Log on locally** right
- Restriction: You cannot be granted Sysadmin or databases must be created and configured manually
- Restriction: Your labels do not have auto-labeling conditions

Discover and scan all SharePoint sites and subsites under a specific URL

The scanner can discover and scan all SharePoint sites and subsites under a specific URL with the following configuration:

1. Start **SharePoint Central Administration**.
2. On the **SharePoint Central Administration** website, in the **Application Management** section, click **Manage web applications**.
3. Click to highlight the web application whose permission policy level you want to manage.
4. Choose the relevant farm and then select **Manage Permissions Policy Levels**.
5. Select **Site Collection Auditor** in the **Site Collection Permissions** options, then grant **View Application Pages** in the Permissions list, and finally, name the new policy level **Scanner site collection auditor and viewer**.
6. Add your scanner user to the new policy and grant **Site collection** in the Permissions list.
7. Add a URL of the SharePoint that hosts sites or subsites that need to be scanned. For more information, see Configure the scanner settings.

To learn more about how to manage your SharePoint policy levels see, manage permission policies for a web application.

Restriction: The scanner server cannot have internet connectivity

While the information protection client can't apply encryption without an internet connection, the scanner can still apply labels based on imported policies.

To support a disconnected computer, use one of the following methods:

- Use the Microsoft Purview portal (recommended when possible)
- Use PowerShell

Use the Microsoft Purview portal with a disconnected computer

To support a computer that can't connect to the Microsoft Purview portal, perform the following steps:

1. Configure labels in your policy, and then use the procedure to support disconnected computers to enable offline classification and labeling.
2. Enable offline management for content jobs as follows:

   Enable offline management for content scan jobs:

   a. Set the scanner to function in **offline** mode, using the Set-ScannerConfiguration cmdlet.

   b. Configure the scanner in the Microsoft Purview portal by creating a scanner cluster. For more information, see Configure the scanner settings.

   c. Export your content job from the **Information protection - Content scan jobs** pane using the **Export** option.

   d. Import the policy using the Import-ScannerConfiguration cmdlet.

   Results for offline content scan jobs are located at:
   %localappdata%\Microsoft\MSIP\Scanner\Reports

Use PowerShell with a disconnected computer

Perform the following procedure to support a disconnected computer using PowerShell only.

> **Important**
>
> Admins of Azure China 21Vianet scanner servers *must* use this procedure in order to manage their content scan jobs.

Manage your content scan jobs using PowerShell only:

1. Set the scanner to function in **offline** mode, using the Set-ScannerConfiguration cmdlet.

2. Create a new content scan job using the Set-ScannerContentScan cmdlet, making sure to use the mandatory `-Enforce On` parameter.

3. Add your repositories using the Add-ScannerRepository cmdlet, with the path to the repository you want to add.

   > **Tip**
   >
   > To prevent the repository from inheriting settings from your content scan job, add the `OverrideContentScanJob On` parameter, as well as values for additional settings.
   >
   > To edit details for an existing repository, use the **Set-ScannerRepository** command.

4. Use the Get-ScannerContentScan and Get-ScannerRepository cmdlets to return information about your content scan job's current settings.

5. Use the Set-ScannerRepository command to update details for an existing repository.

6. Run your content scan job immediately if needed, using the Start-Scan cmdlet.

   Results for offline content scan jobs are located at:
   %localappdata%\Microsoft\MSIP\Scanner\Reports

7. If you need to remove a repository or an entire content scan job, use the following cmdlets:

   - Remove-ScannerContentScan
   - Remove-ScannerRepository

Restriction: You cannot be granted Sysadmin or databases must be created and configured manually

Use the following procedures to manually create databases and grant the **db_owner** role, as needed.

- Procedure for the scanner database

If you can be granted the Sysadmin role *temporarily* to install the scanner, you can remove this role when the scanner installation is complete.

Do one of the following, depending on your organization's requirements:

Expand table

Restriction	Description
You can have the Sysadmin role temporarily	If you temporarily have the Sysadmin role, the database is automatically created for you and the service account for the scanner is automatically granted the required permissions. However, the user account that configures the scanner still requires the **db_owner** role for the scanner configuration database. If you only have the Sysadmin role until the scanner installation is complete, grant the **db_owner** role to the user account manually.
You cannot have the Sysadmin role at all	If you cannot be granted the Sysadmin role even temporarily, you must ask a user with Sysadmin rights to manually create a database before you install the scanner. For this configuration, the **db_owner** role must be assigned to the following accounts: - Service account for the scanner - User account for the scanner installation - User account for scanner configuration Typically, you will use the same user account to install and configure the scanner. If you use different accounts, they both require the **db_owner** role for the scanner configuration database. Create this user and rights as needed. If you specify your own cluster name, the configuration database is named **AIPScannerUL_<cluster_name>**.

Additionally:

- You must be a local administrator on the server that will run the scanner
- The service account that will run the scanner must be granted Full Control permissions to the following registry keys:
  - `HKEY_LOCAL_MACHINE\SOFTWARE\WOW6432Node\Microsoft\MSIPC\Server`
  - `HKEY_LOCAL_MACHINE\SOFTWARE\Microsoft\MSIPC\Server`

If, after configuring these permissions, you see an error when you install the scanner, the error can be ignored and you can manually start the scanner service.

Manually create a database and user for the scanner, and grant db_owner rights

If you need to manually create your scanner database and/or create a user and grant **db_owner** rights on the database, ask your Sysadmin to perform the following steps:

1. Create a database for scanner:

```
SQL
```

```
**CREATE DATABASE AIPScannerUL_[clustername]**

**ALTER DATABASE AIPScannerUL_[clustername] SET TRUSTWORTHY ON**
```

2. Grant rights to the user that runs the installation command and is used to run scanner management commands. Use the following script:

```
SQL

if not exists(select * from master.sys.server_principals where sid =
SUSER_SID('domain\user')) BEGIN declare @T nvarchar(500) Set @T = 'CREATE
LOGIN ' + quotename('domain\user') + ' FROM WINDOWS ' exec(@T) END
USE DBName IF NOT EXISTS (select * from sys.database_principals where sid =
SUSER_SID('domain\user')) BEGIN declare @X nvarchar(500) Set @X = 'CREATE
USER ' + quotename('domain\user') + ' FROM LOGIN ' +
quotename('domain\user'); exec sp_addrolemember 'db_owner', 'domain\user'
exec(@X) END
```

3. Grant rights to scanner service account. Use the following script:

```
SQL

if not exists(select * from master.sys.server_principals where sid =
SUSER_SID('domain\user')) BEGIN declare @T nvarchar(500) Set @T = 'CREATE
LOGIN ' + quotename('domain\user') + ' FROM WINDOWS ' exec(@T) END
```

Restriction: The service account for the scanner cannot be granted the Log on locally right

If your organization policies prohibit the **Log on locally** right for service accounts, use the *OnBehalfOf* parameter with Set-Authentication.

For more information, see Run information protection labeling cmdlets unattended.

Restriction: The scanner service account cannot be synchronized to Microsoft Entra ID but the server has internet connectivity

You can have one account to run the scanner service and use another account to authenticate to Microsoft Entra ID:

- **For the scanner service account**, use a local Windows account or an Active Directory account.

- **For the Microsoft Entra account**, specify the Microsoft Entra user in the Set-Authentication cmdlet, in the *DelegatedUser* parameter.

  If you are running the scan under any user other than the scanner account, make sure to specify the scanner account in *OnBehalfOf* parameter as well.

  For more information, see Run information protection labeling cmdlets unattended.

Restriction: Your labels do not have auto-labeling conditions

If your labels do not have any auto-labeling conditions, plan to use one of the following options when configuring your scanner:

Expand table

Option	Description
Discover all info types	In your content scan job, set the **Info types to be discovered** option to **All**. This option sets the content scan job to scan your content for all sensitive information types.
Use recommended labeling	In your content scan job, set the **Treat recommended labeling as automatic** option to **On**. This setting configures the scanner to automatically apply all recommended labels on your content.
Define a default label	Define a default label in your policy, content scan job, or repository. In this case the scanner applies the default label on all files found.

Next steps

Once you've confirmed that your system complies with the scanner prerequisites, continue with Configuring and installing the information protection scanner.

For an overview about the scanner, see Learn about the information protection scanner.

Configure and install the information protection scanner

Summarize this article for me

> **Note**
>
> There's a new version of the information protection scanner. For more information, see Upgrade the Microsoft Purview Information Protection scanner.

This article describes how to configure and install the Microsoft Purview Information Protection scanner, formerly named *Azure Information Protection unified labeling scanner*, or the on-premises scanner.

> **Tip**
>
> While most customers will perform these procedures in the admin portal, you may need to work in PowerShell only.
>
> For example, if you are working in an environment without access to the admin portal, such as Azure China 21Vianet, follow the instructions in Use PowerShell to configure the scanner.

Overview

Before you start, verify that your system complies with the required prerequisites.

Then, use the following steps to configure and install the scanner:

1. Configure the scanner settings
2. Install the scanner
3. Get a Microsoft Entra token for the scanner
4. Configure the scanner to apply classification and protection

Next, perform the following configuration procedures as needed for your system:

Expand table

Procedure	Description
Change which file types to protect	You may want to scan, classify, or protect different file types than the default. For more information, see The scanning process.
Upgrading your scanner	Upgrade your scanner to use the latest features and improvements.
Editing data repository settings in bulk	Use import and export options to make changes in bulk for multiple data repositories.
Use the scanner with alternative configurations	Use the scanner without configuring labels with any conditions
Optimize performance	Guidance to optimize your scanner performance

If you don't have access to the scanner pages in the Microsoft Purview portal, configure any scanner settings in PowerShell only. For more information, see Use PowerShell to configure the scanner and Supported PowerShell cmdlets.

Configure the scanner settings

Before you install the scanner, or upgrade it from an older general availability version, configure or verify your scanner settings. For this configuration, you can use either the Microsoft Purview portal.

To configure your scanner in the Microsoft Purview portal:

1. Sign in using one of the following roles:

- **Compliance Administrator**
- **Compliance Data Administrator**
- **Security Administrator**
- **Organization Management**

2. Sign in to the Microsoft Purview portal > **Settings** card > **Information Protection** > **Information protection scanner**.

3. Create a scanner cluster. This cluster defines your scanner and is used to identify the scanner instance, such as during installation, upgrades, and other processes.

4. Create a content scan job to define the repositories you want to scan.

Create a scanner cluster

To create a scanner cluster in the Microsoft Purview portal:

1. From the tabs on the **Information protection scanner** page, select **Clusters**.
2. On the **Clusters** tab, select **Add** +.
3. On the **New cluster** pane, enter a meaningful name for the scanner, and an optional description.

   The cluster name is used to identify the scanner's configurations and repositories. For example, you might enter **Europe** to identify the geographical locations of the data repositories you want to scan.

   You'll use this name later on to identify where you want to install or upgrade your scanner.
4. Select **Save** to save your changes.

Create a content scan job

Deep dive into your content to scan specific repositories for sensitive content.

To create your content scan job on the Microsoft Purview portal:

1. From the tabs on the **Information protection scanner** page, select **Content scan jobs**.
2. On the **Content scan jobs** pane, select **Add** +.
3. For this initial configuration, configure the following settings, and then select **Save**.

Expand table

Setting	Description
Content scan job settings	- **Schedule**: Keep the default of **Manual** - **Info types to be discovered**: Change to **Policy only**
DLP policy	If you're using a data loss prevention policy, set **Enable DLP rules** to **On**. For more information, see Use a DLP policy.
Sensitivity policy	- **Enforce sensitivity labeling policy**: Select **Off** - **Label files based on content**: Keep the default of **On** - **Default label**: Keep the default of **Policy default** - **Relabel files**: Keep the default of **Off**
Configure file settings	- **Preserve "Date modified", "Last modified" and "Modified by"**: Keep the default of **On** - **File types to scan**: Keep the default file types for **Exclude** - **Default owner**: Keep the default of **Scanner Account** - **Set repository owner**: Use this option only when using a DLP policy.

Setting	Description

4. Open the content scan job that was saved, and select the **Repositories** tab to specify the data stores to be scanned.

   Specify UNC paths and SharePoint Server URLs for SharePoint on-premises document libraries and folders.

   > ⓘ Note
   >
   > SharePoint Server 2019, SharePoint Server 2016, and SharePoint Server 2013 are supported for SharePoint. SharePoint Server 2010 is also supported when you have extended support for this version of SharePoint.

   To add your first data store, while on the **Repositories** tab:

   a. On the **Repositories** pane, select **Add**:

   b. On the **Repository** pane, specify the path for the data repository, and then select **Save**.

   - For a network share, use `\\Server\Folder`.
   - For a SharePoint library, use `http://sharepoint.contoso.com/Shared%20Documents/Folder`.
   - For a local path: `C:\Folder`
   - For a UNC path: `\\Server\Folder`

   > ⓘ Note
   >
   > Wildcards are not supported and WebDav locations are not supported. Scanning of OneDrive locations as repositories is not supported.

   If you add a SharePoint path for **Shared Documents**:

   - Specify **Shared Documents** in the path when you want to scan all documents and all folders from Shared Documents. For example: `http://sp2013/SharedDocuments`
   - Specify **Documents** in the path when you want to scan all documents and all folders from a subfolder under Shared Documents. For example: `http://sp2013/Documents/SalesReports`
   - Or, specify only the **FQDN** of your SharePoint, for example `http://sp2013` to discover and scan all SharePoint sites and subsites under a specific URL and subtitles under this URL. Grant scanner **Site Collector Auditor** rights to enable this.

For the remaining settings on this pane, don't change them for this initial configuration, but keep them as **Content scan job default**. The default setting means that the data repository inherits the settings from the content scan job.

Use the following syntax when adding SharePoint paths:

Expand table

Path	Syntax
Root path	`http://<SharePoint server name>` Scans all sites, including any site collections allowed for the scanner user. Requires additional permissions to automatically discover root content
Specific SharePoint subsite or collection	One of the following: - `http://<SharePoint server name>/<subsite name>` - `http://SharePoint server name>/<site collection name>/<site name>` Requires additional permissions to automatically discover site collection content
Specific SharePoint library	One of the following: - `http://<SharePoint server name>/<library name>` - `http://SharePoint server name>/.../<library name>`
Specific SharePoint folder	`http://<SharePoint server name>/.../<folder name>`

5. Repeat the previous steps to add as many repositories as needed.

You're now ready to install the scanner with the content scanner job that you've created. Continue with Install the scanner.

Install the scanner

After you've configured the scanner, perform the following steps to install the scanner. This procedure is performed fully in PowerShell.

1. Sign in to the Windows Server computer that will run the scanner. Use an account that has local administrator rights and that has permissions to write to the SQL Server master database.

> **ⓘ Important**
>
> You must have the information protection client installed on your machine before installing the scanner.
>
> For more information, see Prerequisites for installing and deploying the information protection scanner.

2. Open a Windows PowerShell session with the **Run as an administrator** option.

3. Run the Install-Scanner cmdlet, specifying your SQL Server instance on which to create a database for the information protection scanner, and the scanner cluster name that you specified in the preceding section:

   ```PowerShell
   Install-Scanner -SqlServerInstance <name> -Cluster <cluster name>
   ```

   Examples, using the scanner cluster name of **Europe**:

   - For a default instance: `Install-Scanner -SqlServerInstance SQLSERVER1 -Cluster Europe`
   - For a named instance: `Install-Scanner -SqlServerInstance SQLSERVER1\SCANNER -Cluster Europe`
   - For SQL Server Express: `Install-Scanner -SqlServerInstance SQLSERVER1\SQLEXPRESS -Cluster Europe`

   When you're prompted, provide the Active Directory credentials for the scanner service account.

   Use the following syntax: `\<domain\user name>`. For example: `contoso\scanneraccount`

4. Verify that the service is now installed by using **Administrative Tools** > **Services**.

   The installed service is named **Microsoft Purview Information Protection Scanner** and is configured to run by using the scanner service account that you created.

Now that you've installed the scanner, you need to get a Microsoft Entra token for the scanner service account to authenticate, so that the scanner can run unattended.

Get a Microsoft Entra token for the scanner

A Microsoft Entra token allows the scanner to authenticate to the Microsoft Purview Information Protection Scanner service, enabling the scanner to run unattended.

For more information, see Run information protection labeling cmdlets unattended.

> **Note**
>
> Support for certificate-based authentication is in public preview. For more information, see Certificate-based authentication

To get a Microsoft Entra token:

1. Navigate to the Azure Portal and proceed to the **Microsoft Entra ID** Blade.
2. In the Microsoft Entra ID side pane, click **App Registrations**.
3. At the top, go ahead and click **+ New registration**.
4. In the Name section type in **InformationProtectionScanner**.
5. Leave **Supported account types** as default.
6. For the Redirect URI, leave the type as Web but type in **http://localhost** for the entry portion and click **Register**.
7. On the Overview page of this application, note down in your text editor of choice the following IDs: **Application (client) ID** and **Directory (tenant) ID**. You will need this later when setting up the Set-AIPAuthentication command.
8. On the side pane, navigate to **Certificates and Secrets**.
9. Click on **+ New client secret**.
10. In the dialog box that shows up, enter a description for your secret and set it to Expire **In 1 year** and then Add the secret.
11. You should see now under the client secrets section that there is an entry with the **Secret Value**. Go ahead and copy this value and store it in the file where you saved the Client ID and Tenant ID. This is the only time you will be able to see the secret value, it will not be recoverable if you don't copy it at this time.
12. On the side pane, navigate to **API Permissions**.
13. Go ahead and select **Add a permission**.

14. When the screen shows, select **Azure Rights Management Service**. Then select **Application Permissions**.

15. Click the drop down for **Content** and put checkmarks down for **Content.DelegatedReader** and **Content.DelegatedWriter**. Then at the bottom of the screen, click **Add Permissions**.

16. Navigate back the **API Permissions** section and add another permission.

17. This time, for the Select an API section, click on **APIs my organization uses**. In the search bar, type in **Microsoft Information Protection Sync Service** and select it.

18. Select **Application Permissions** and then in the **Unified Policy** drop down, checkmark the permission **UnifiedPolicy.Tenant.Read**. Then at the bottom of the screen, click **Add Permissions**.

19. Back on the **API Permissions** screen, click **Grant Admin Consent** and look for the operation being successful (signified by a green checkmark).

20. From the Windows Server computer, if your scanner service account has been granted the **Log on locally** right for the installation, sign in with this account and start a PowerShell session.

    Run Set-Authentication, specifying the values that you copied from the previous step:

    PowerShell

    ```
    Set-Authentication -AppId <ID of the registered app> -AppSecret <client secret
    sting> -TenantId <your tenant ID> -DelegatedUser <Azure AD account>
    ```

    For example:

    PowerShell

    ```
    $pscreds = Get-Credential CONTOSO\scanner
    Set-Authentication -AppId "77c3c1c3-abf9-404e-8b2b-4652836c8c66" -AppSecret
    "OAkk+rnuYc/u+]ah2kNxVbtrDGbS47L4" -DelegatedUser scanner@contoso.com -
    TenantId "9c11c87a-ac8b-46a3-8d5c-f4d0b72ee29a" -OnBehalfOf $pscreds
    Acquired application access token on behalf of CONTOSO\scanner.
    ```

Tip

If your scanner service account cannot be granted the **Log on locally** right for the installation, use the *OnBehalfOf* parameter with **Set-Authentication**, as described in as described in **Run information protection labeling cmdlets unattended**.

The scanner now has a token to authenticate to Microsoft Entra ID. This token is valid for one year, two years, or never, according to your configuration of the **Web app /API** client secret in Microsoft Entra ID. When the token expires, you must repeat this procedure.

Continue using one of the following steps, depending on whether you're using the Microsoft Purview portal to configure your scanner, or PowerShell only:

Admin portal only

You're now ready to run your first scan in discovery mode. For more information, see Run a discovery cycle and view reports for the scanner.

Once you've run your initial discovery scan, continue with Configure the scanner to apply classification and protection.

For more information, see Run information protection labeling cmdlets unattended.

Configure the scanner to apply classification and protection

The default settings configure the scanner to run once, and in reporting-only mode. To change these settings, edit the content scan job.

Tip

If you're working in PowerShell only, see **Configure the scanner to apply classification and protection - PowerShell only**.

To configure the scanner to apply classification and protection in the Microsoft Purview portal::

1. In the Microsoft Purview portal, on the **Content scan jobs** tab, select a specific content scan job to edit it.

2. Select the content scan job, change the following, and then select **Save**:

    - From the **Content scan job** section: Change the **Schedule** to **Always**

- From the **Enforce sensitivity labeling policy** section: Change the radio button to **On**

3. Make sure a node for the content scan job is online, then start the content scan job again by selecting **Scan now**. The **Scan now** button only appears when a node for the selected content scan job is online.

The scanner is now scheduled to run continuously. When the scanner works its way through all configured files, it automatically starts a new cycle so that any new and changed files are discovered.

Use a DLP policy

Using a data loss prevention policy enables the scanner to detect potential data leaks by matching DLP rules to files stored in file shares and SharePoint Server.

- **Enable DLP rules in your content scan job** to reduce the exposure of any files that match your DLP policies. When your DLP rules are enabled, the scanner may reduce file access to data owners only, or reduce exposure to network-wide groups, such as **Everyone**, **Authenticated Users**, or **Domain Users**.
- **In the Microsoft Purview portal**, determine whether you are just testing your DLP policy or whether you want your rules enforced and your file permissions changed according to those rules. For more information, see Create and Deploy data loss prevention policies

DLP policies are configured in the Microsoft Purview portal. For more information about DLP licensing, see Get started with the data loss prevention on-premises scanner.

> **Tip**
>
> Scanning your files, even when just testing the DLP policy, also creates file permission reports. Query these reports to investigate specific file exposures or explore the exposure of a specific user to scanned files.
>
> To use PowerShell only, see **Use a DLP policy with the scanner - PowerShell only**.

To use a DLP policy with the scanner in the Microsoft Purview portal:

1. In the Microsoft Purview portal, navigate to the **Content scan jobs** tab and select a specific content scan job. For more information, see Create a content scan job.

2. Under **Enable DLP policy rules**, set the radio button to **On**.

   > **Important**

> Do not set **Enable DLP rules** to **On** unless you actually have a DLP policy configured in Microsoft 365.
>
> Turning this feature on without a DLP policy will cause the scanner to generate errors.

3. (Optional) Set the **Set repository owner** to **On**, and define a specific user as the repository owner.

   This option enables the scanner to reduce the exposure of any files found in this repository, which match the DLP policy, to the repository owner defined.

DLP policies and *make private* actions

If you are using a DLP policy with a *make private* action, and are also planning to use the scanner to automatically label your files, we recommend that you also define the unified labeling client's **UseCopyAndPreserveNTFSOwner** advanced setting.

This setting ensures that the original owners retain access to their files.

For more information, see Create a content scan job and Automatically apply a sensitivity label to Microsoft 365 data.

Change which file types to protect

By default the scanner protects Office file types and PDF files only.

Use PowerShell commands to change this behavior as needed, such as to configure the scanner to protect all file types, just as the client does, or to protect additional, specific file types.

For a label policy that applies to the user account downloading labels for the scanner, specify a PowerShell advanced setting named **PFileSupportedExtensions**.

For a scanner that has access to the internet, this user account is the account that you specify for the *DelegatedUser* parameter with the Set-Authentication command.

Example 1: PowerShell command for the scanner to protect all file types, where your label policy is named "Scanner":

PowerShell

```
Set-LabelPolicy -Identity Scanner -AdvancedSettings @{PFileSupportedExtensions="*"}
```

Example 2: PowerShell command for the scanner to protect .xml files and .tiff files in addition to Office files and PDF files, where your label policy is named "Scanner":

PowerShell

```
Set-LabelPolicy -Identity Scanner -AdvancedSettings
@{PFileSupportedExtensions=ConvertTo-Json(".xml", ".tiff")}
```

For more information, see Change which file types to protect.

Upgrade your scanner

If you've previously installed the scanner and want to upgrade, use the instructions described in Upgrade the Microsoft Purview Information Protection scanner.

Then, configure and use your scanner as usual, skipping the steps to install your scanner.

Edit data repository settings in bulk

Use the **Export** and **Import** buttons to make changes for your scanner across several repositories.

This way, you don't need to make the same changes several times, manually, in the Microsoft Purview portal.

For example, if you've a new file type on several SharePoint data repositories, you may want to update the settings for those repositories in bulk.

To make changes in bulk across repositories in the Microsoft Purview portal:

1. In the Microsoft Purview portal, select a specific content scan job and navigate to the **Repositories** tab within the pane. Select the **Export** option.

2. Manually edit the exported file to make your change.

3. Use the **Import** option on the same page to import the updates back across your repositories.

Use the scanner with alternative configurations

The scanner usually looks for conditions specified for your labels in order to classify and protect your content as needed.

In the following scenarios, the scanner is also able to scan your content and manage labels, without any conditions configured:

- Apply a default label to all files in a data repository
- Remove existing labels from all files in a data repository
- Identify all custom conditions and known sensitive information types

Apply a default label to all files in a data repository

In this configuration, all unlabeled files in the repository are labeled with the default label specified for the repository or the content scan job. Files are labeled without inspection.

Configure the following settings:

Expand table

Setting	Description
Label files based on content	Set to **Off**
Default label	Set to **Custom**, and then select the label to use
Enforce default label	Select to have the default label applied to all files, even if they're already labeled by turning **Relabel files** and **Enforce default label** on

Remove existing labels from all files in a data repository

In this configuration, all existing labels are removed, including protection, if protection was applied with the label. Protection applied independently of a label is retained.

Configure the following settings:

Expand table

Setting	Description
Label files based on content	Set to **Off**
Default label	Set to **None**
Relabel files	Set to **On**, with the **Enforce default label** set to **On**

Identify all custom conditions and known sensitive information types

This configuration enables you to find sensitive information that you might not realize you had, at the expense of scanning rates for the scanner.

Set the **Info types to be discovered** to **All**.

To identify conditions and information types for labeling, the scanner uses any custom sensitive information types specified, and the list of built-in sensitive information types that are available to select, as defined in your labeling management center.

Optimize scanner performance

> **Note**
>
> If you're looking to improve the responsiveness of the scanner computer rather than the scanner performance, use an advanced client setting to limit the number of threads used by the scanner.

Use the following options and guidance to help you optimize scanner performance:

Expand table

Option	Description
Have a high speed and reliable network connection between the scanner computer and the scanned data store	For example, place the scanner computer in the same LAN, or preferably, in the same network segment as the scanned data store. The quality of the network connection affects the scanner performance because, to inspect the files, the scanner transfers the contents of the files to the computer running the scanner Microsoft Purview Information Protection Scanner service. Reducing or eliminating the network hops required for the data to travel also reduces the load on your network.
Make sure the scanner computer has available processor resources	Inspecting the file contents and encrypting and decrypting files are processor-intensive actions. Monitor the typical scanning cycles for your specified data stores to identify whether a lack of processor resources is negatively affecting the scanner performance.

Option	Description
Install multiple instances of the scanner	The scanner supports multiple configuration databases on the same SQL server instance when you specify a custom cluster name for the scanner. **Tip**: Multiple scanners can also share the same cluster, resulting in quicker scanning times. If you plan to install the scanner on multiple machines with the same database instance, and want your scanners to run in parallel, you must install all your scanners using the same cluster name.
Check your alternative configuration usage	The scanner runs more quickly when you use the alternative configuration to apply a default label to all files because the scanner doesn't inspect the file contents. The scanner runs more slowly when you use the alternative configuration to identify all custom conditions and known sensitive information types.

Additional factors that affect performance

Additional factors that affect the scanner performance include:

Expand table

Factor	Description
Load/response times	The current load and response times of the data stores that contain the files to scan will also affect scanner performance.
Scanner mode (Discovery / Enforce)	Discovery mode typically has a higher scanning rate than enforce mode. Discovery requires a single file read action, whereas enforce mode requires read and write actions.
Policy changes	Your scanner performance may be affected if you've made changes to the autolabeling in the label policy. Your first scan cycle, when the scanner must inspect every file, will take longer than subsequent scan cycles that by default, inspect only new and changed files. If you change the conditions or autolabeling settings, all files are scanned again. For more information, see Rescanning files.
Regex constructions	Scanner performance is affected by how your regex expressions for custom conditions are constructed.

Factor	Description
	To avoid heavy memory consumption and the risk of timeouts (15 minutes per file), review your regex expressions for efficient pattern matching. For example: - Avoid greedy quantifiers - Use non-capturing groups such as `(?:expression)` instead of `(expression)`
Log level	Log level options include **Debug**, **Info**, **Error** and **Off** for the scanner reports. - **Off** results in the best performance - **Debug** considerably slows down the scanner and should be used only for troubleshooting. For more information, see the *ReportLevel* parameter for the Set-ScannerConfiguration cmdlet.
Files being scanned	- With the exception of Excel files, Office files are more quickly scanned than PDF files. - Unprotected files are quicker to scan than protected files. - Large files obviously take longer to scan than small files.

Use PowerShell to configure the scanner

This section describes the steps required to configure and install the scanner when you don't have access to the scanner pages in the Microsoft Purview portal, and must use PowerShell only.

> ⓘ **Important**
>
> - Some steps require Powershell whether or not you are able to access the scanner pages in the Microsoft Purview portal, and are identical. For these steps, see the earlier instructions in this article as indicated.
> - If you're working with the scanner for Azure China 21Vianet, additional steps are required in addition to the instructions detailed here. For more information, see Microsoft Purview Information Protection for Office 365 operated by 21Vianet.

For more information, see Supported PowerShell cmdlets.

To configure and install your scanner:

1. Start with PowerShell closed. If you've previously installed the information protection client and scanner, make sure that the **Microsoft Purview Information Protection Scanner** service is stopped.

2. Open a Windows PowerShell session with the **Run as an administrator** option.

3. Run the Install-Scanner command to install your scanner on your SQL server instance, with the **Cluster** parameter to define your cluster name.

   This step is identical whether or not you're able to access the scanner pages in the Microsoft Purview portal. For more information, see the earlier instructions in this article: Install the scanner

4. Get an Azure token to use with your scanner, and then reauthenticate.

   This step is identical whether or not you're able to access the scanner pages in the Microsoft Purview portal. For more information, see the earlier instructions in this article: Get a Microsoft Entra token for the scanner.

5. Run the Set-ScannerConfiguration cmdlet to set the scanner to function in offline mode. Run:

   PowerShell

   ```
   Set-ScannerConfiguration -OnlineConfiguration Off
   ```

6. Run the Set-ScannerContentScan cmdlet to create a default content scan job.

   The only required parameter in the **Set-ScannerContentScan** cmdlet is **Enforce**. However, you might want to define other settings for your content scan job at this time. For example:

   PowerShell

   ```
   Set-ScannerContentScan -Schedule Manual -DiscoverInformationTypes PolicyOnly -
   Enforce Off -DefaultLabelType PolicyDefault -RelabelFiles Off -
   PreserveFileDetails On -IncludeFileTypes '' -ExcludeFileTypes '.msg,.tmp' -
   DefaultOwner <account running the scanner>
   ```

   The syntax above configures the following settings while you continue the configuration:

   - Keeps the scanner run scheduling to *manual*
   - Sets the information types to be discovered based on the sensitivity label policy
   - Does *not* enforce a sensitivity label policy

- Automatically labels files based on content, using the default label defined for the sensitivity label policy
- Does *not* allow for relabeling files
- Preserves file details while scanning and auto-labeling, including *date modified*, *last modified*, and *modified by* values
- Sets the scanner to exclude .msg and .tmp files when running
- Sets the default owner to the account you want to use when running the scanner

7. Use the Add-ScannerRepository cmdlet to define the repositories you want to scan in your content scan job. For example, run:

```
PowerShell
Add-ScannerRepository -OverrideContentScanJob Off -Path 'c:\repoToScan'
```

Use one of the following syntaxes, depending on the type of repository you're adding:

- For a network share, use `\\Server\Folder`.
- For a SharePoint library, use `http://sharepoint.contoso.com/Shared%20Documents/Folder`.
- For a local path: `C:\Folder`
- For a UNC path: `\\Server\Folder`

> ⓘ **Note**
>
> Wildcards are not supported and WebDav locations are not supported.
>
> To modify the repository later on, use the Set-ScannerRepository cmdlet instead.

If you add a SharePoint path for **Shared Documents**:

- Specify **Shared Documents** in the path when you want to scan all documents and all folders from Shared Documents. For example: `http://sp2013/SharedDocuments`
- Specify **Documents** in the path when you want to scan all documents and all folders from a subfolder under Shared Documents. For example: `http://sp2013/Documents/SalesReports`
- Or, specify only the **FQDN** of your SharePoint, for example `http://sp2013` to discover and scan all SharePoint sites and subsites under a specific URL and subtitles under this URL. Grant scanner **Site Collector Auditor** rights to enable this.

Use the following syntax when adding SharePoint paths:

Expand table

Path	Syntax
Root path	`http://<SharePoint server name>` Scans all sites, including any site collections allowed for the scanner user.
Specific SharePoint subsite or collection	One of the following: - `http://<SharePoint server name>/<subsite name>` - `http://SharePoint server name>/<site collection name>/<site name>`
Specific SharePoint library	One of the following: - `http://<SharePoint server name>/<library name>` - `http://SharePoint server name>/.../<library name>`
Specific SharePoint folder	`http://<SharePoint server name>/.../<folder name>`

Continue with the following steps as needed:

- Configuration for customers in China
- Run a discovery cycle and view reports for the scanner
- Use PowerShell to configure the scanner to apply classification and protection
- Use PowerShell to configure a DLP policy with the scanner

Use PowerShell to configure the scanner to apply classification and protection

1. Run the Set-ScannerContentScan cmdlet to update your content scan job to set your scheduling to always and enforce your sensitivity policy.

   PowerShell

   ```
   Set-ScannerContentScan -Schedule Always -Enforce On
   ```

   > **Tip**
   >
   > You may want to change other settings on this pane, such as whether file attributes are changed and whether the scanner can relabel files. For more information about the settings available, see the full Set-ScannerContentScan documentation.

2. Run the Start-Scan cmdlet to run your content scan job:

PowerShell

```
Start-Scan
```

The scanner is now scheduled to run continuously. When the scanner works its way through all configured files, it automatically starts a new cycle so that any new and changed files are discovered.

Use PowerShell to configure a DLP policy with the scanner

Run the Set-ScannerContentScan cmdlet again with the **-EnableDLP** parameter set to **On**, and with a specific repository owner defined.

For example:

PowerShell

```
Set-ScannerContentScan -EnableDLP On -RepositoryOwner 'domain\user'
```

Certificate-based authentication (Public Preview)

Using certificate-based authentication with the information protection scanner

This section describes how to configure the Microsoft Purview Information Protection scanner to authenticate using a certificate instead of a client secret.

Prerequisites

Before configuring certificate-based authentication, ensure you have:

- The Microsoft Purview Information Protection scanner installed
- Administrative access to the Windows Server running the scanner
- Access to the Microsoft Entra app registration used by the scanner
- Windows PowerShell 5.1

Step 1: Obtain or create a certificate

You have two options for obtaining a certificate:

1. **CA-issued certificate**: Request a certificate from your organization's Certificate Authority.

2. **Self-signed certificate**: Create a self-signed certificate.

Option A: Using a CA-issued certificate (production)

Using a CA-issued certificate provides:

- Automatic trust via the existing PKI infrastructure
- Centralized certificate lifecycle management
- Compliance with organizational security policies
- No need to skip certificate chain validation

Certificate requirements

When requesting a certificate from your CA, ensure it meets these requirements:

Expand table

Requirement	Value
Key Algorithm	RSA
Key Length	2048 bits minimum (4096 recommended)
Key Usage	Digital Signature
Private Key	Exportable (for backup) or marked for the machine
Validity	Per your organization's policy (typically 1-2 years)

Review documentation from your CA for specific instructions on requesting a certificate with these parameters.

After obtaining the key from your enterprise CA, install the certificate in the Local Machine store on each Windows Server running the scanner. Make sure to install the certificate in the **Personal** store of the Local Machine.

Exporting the public key for Microsoft Entra

After the certificate is issued and installed, export the public key:

PowerShell

```
# Find the certificate (adjust the subject filter as needed)
$cert = Get-ChildItem Cert:\LocalMachine\My | Where-Object { $_.Subject -like
```

```
"*PurviewScanner*" }
Export-Certificate -Cert $cert -FilePath "C:\temp\PurviewScanner.cer"
```

Option B: Using a self-signed certificate

A self-signed certificate can be created on a single server, then imported to servers across the scanner cluster.

PowerShell

```powershell
# Create certificate in Local Machine store with RSA provider
$cert = New-SelfSignedCertificate `
    -Subject "CN=PurviewScanner" `
    -CertStoreLocation Cert:\LocalMachine\My `
    -KeyExportPolicy Exportable `
    -KeySpec Signature `
    -KeyLength 2048 `
    -KeyAlgorithm RSA `
    -HashAlgorithm SHA256 `
    -NotAfter (Get-Date).AddYears(2) `
    -Provider "Microsoft Enhanced RSA and AES Cryptographic Provider"

# Display the certificate details
Write-Host "Certificate created successfully"
Write-Host "Thumbprint: $($cert.Thumbprint)"
Write-Host "Subject: $($cert.Subject)"
```

> **ⓘ Important**
>
> The `-Provider "Microsoft Enhanced RSA and AES Cryptographic Provider"` parameter is required. Using CNG (Cryptography Next Generation) providers will result in "Keyset does not exist" errors when the scanner attempts to authenticate.

Export the public key

Export the certificate's public key to upload to Microsoft Entra:

PowerShell

```powershell
# Export public key certificate (.cer file)
Export-Certificate -Cert $cert -FilePath "C:\temp\PurviewScanner.cer"
```

Step 2: Grant private key permissions to the scanner service account

The scanner service account must have permission to read the certificate's private key. Without this permission, authentication will fail with "Keyset does not exist" errors.

Using the Certificate Manager UI

1. Open the Local Machine certificate store:
    - Press `Win + R`, type `certlm.msc`, and press Enter
2. Navigate to **Certificates - Local Computer**
3. Right-click on the certificate you created (e.g., "PurviewScanner") and select **All Tasks** > **Manage Private Keys**
4. Click **Add** and enter the scanner service account name (e.g., `CONTOSO\ScannerService`)
5. Select the account and ensure **Read** permission is checked
6. Click **OK** to save

Using PowerShell

Alternatively, grant permissions using PowerShell. Make sure to replace the service account name with your actual scanner service account.

PowerShell

```
# Get the certificate
$cert = Get-ChildItem Cert:\LocalMachine\My | Where-Object { $_.Subject -eq
"CN=PurviewScanner" }

# Get the private key file path
$privateKey =
[System.Security.Cryptography.X509Certificates.RSACertificateExtensions]::GetRSAPri
vateKey($cert)
$keyPath = $privateKey.Key.UniqueName
$fullKeyPath = "$env:ALLUSERSPROFILE\Microsoft\Crypto\RSA\MachineKeys\$keyPath"

# Get current ACL
$acl = Get-Acl $fullKeyPath

# Add read permission for the scanner service account
$serviceAccount = "CONTOSO\ScannerService"  # Replace with your scanner service
account
$accessRule = New-Object
System.Security.AccessControl.FileSystemAccessRule($serviceAccount, "Read",
"Allow")
$acl.AddAccessRule($accessRule)
```

```
# Apply the updated ACL
Set-Acl $fullKeyPath $acl
```

Step 3: Upload the certificate to Microsoft Entra

Register the certificate's public key with the app registration used by the scanner.

For detailed instructions on uploading a certificate to a Microsoft Entra app registration, see Register a certificate with Microsoft Entra.

After uploading, note the certificate thumbprint displayed in the portal - it should match the thumbprint from Step 1.

Step 4: Configure the scanner to use certificate authentication

Use the `Set-Authentication` cmdlet with certificate parameters instead of the `-AppSecret` parameter.

PowerShell

```
# Get credentials for the scanner service account
$pscreds = Get-Credential CONTOSO\ScannerService

# Set authentication using certificate thumbprint
Set-Authentication `
    -AppId "your-app-id-guid" `
    -TenantId "your-tenant-id-guid" `
    -DelegatedUser "scanner@contoso.com" `
    -CertificateThumbprint "your-certificate-thumbprint" `
    -CertificateStoreLocation LocalMachine `
    -CertificateStoreName My `
    -OnBehalfOf $pscreds
```

The certificate chain is validated automatically for certificates issued by a trusted CA. If validation fails, ensure:

- The CA's root certificate is in the Trusted Root Certification Authorities store
- Any intermediate CA certificates are in the Intermediate Certification Authorities store

> ⓘ **Note**
>
> If you're using a self-signed certificate, add the `-SkipCertificateChainValidation` parameter to bypass chain validation, which will fail for self-signed certificates.

Authenticate using certificate subject name

Alternatively, you can specify the certificate by its subject name:

PowerShell

```powershell
# Get credentials for the scanner service account
$pscreds = Get-Credential CONTOSO\ScannerService

# Set authentication using certificate subject name
Set-Authentication `
    -AppId "your-app-id-guid" `
    -TenantId "your-tenant-id-guid" `
    -DelegatedUser "scanner@contoso.com" `
    -CertificateSubjectName "PurviewScanner" `
    -CertificateStoreLocation LocalMachine `
    -CertificateStoreName My `
    -OnBehalfOf $pscreds
```

> **Note**
>
> If you're using a self-signed certificate, add the `-SkipCertificateChainValidation` parameter to bypass chain validation, which will fail for self-signed certificates.

Parameter reference

Expand table

Parameter	Description
`-CertificateThumbprint`	The SHA-1 thumbprint of the certificate to use for authentication. Cannot be used with `-CertificateSubjectName`.
`-CertificateSubjectName`	The subject name (CN) of the certificate to use for authentication. Cannot be used with `-CertificateThumbprint`.
`-CertificateStoreLocation`	The certificate store location. Use `LocalMachine` for the scanner service. Valid values: `CurrentUser`, `LocalMachine`. Default: `CurrentUser`.
`-CertificateStoreName`	The certificate store name. Default: `My` (Personal store).
`-SkipCertificateChainValidation`	Skips certificate chain validation. Required for self-signed certificates.

Troubleshooting

"Keyset does not exist" error

This error typically indicates one of the following issues:

1. **Private key permissions**: The scanner service account doesn't have read access to the certificate's private key. See Step 2 to grant permissions.
2. **CNG provider**: The certificate was created with a CNG provider instead of an RSA provider. Recreate the certificate using the `-Provider "Microsoft Enhanced RSA and AES Cryptographic Provider"` parameter.

"Certificate not found" error

Verify that:

- The thumbprint or subject name is correct
- The certificate is in the specified store (`LocalMachine\My` by default)
- The certificate has not expired

"A certificate chain could not be built" error

For self-signed certificates, use the `-SkipCertificateChainValidation` parameter. For CA-issued certificates, ensure the full certificate chain (including intermediate CAs) is installed on the server.

401 Unauthorized or "invalid_client" error

Verify that:

- The certificate's public key is uploaded to the correct Microsoft Entra app registration
- The app registration ID matches the `-AppId` parameter
- The certificate has not expired
- The tenant ID is correct

Migrating from secret to certificate authentication

To migrate an existing scanner from secret-based to certificate-based authentication:

1. Create and configure the certificate as described above
2. Run `Set-Authentication` with the certificate parameters
3. The new authentication configuration will overwrite the previous secret-based configuration

4. Restart the scanner service

```powershell
# Restart the scanner service after changing authentication
Restart-Service -Name "Microsoft Purview Information Protection Scanner"
```

See also

- Run information protection labeling cmdlets unattended
- Register a certificate with Microsoft Entra

Supported PowerShell cmdlets

This section lists PowerShell cmdlets supported for the information protection scanner and instructions for configuring and installing the scanner with PowerShell only.

Supported cmdlets for the scanner include:

- Add-ScannerRepository
- Clear-Authentication
- Export-DebugLogs
- Get-FileStatus
- Get-ScannerConfiguration
- Get-ScannerContentScan
- Get-ScannerRepository
- Get-ScanStatus
- Import-ScannerConfiguration
- Install-Scanner
- New-CustomPermissions
- Remove-FileLabel
- Remove-ScannerContentScan
- Remove-ScannerRepository

- Set-Authentication
- Set-FileLabel
- Set-ScannerConfiguration
- Set-ScannerContentScan
- Set-ScannerDatabase
- Set-ScannerRepository
- Start-Scan
- Start-ScannerDiagnostics
- Stop-Scan
- Uninstall-Scanner
- Update-ScannerDatabase

Next steps

Once you've installed and configured your scanner, start scanning your files.

Last updated on 02/20/2026

Running the information protection scanner

Article • 03/31/2025

> **Note**
>
> There's a new version of the information protection scanner. For more information, see Upgrade the Microsoft Purview Information Protection scanner.

Once you've confirmed your system requirements and configured and installed your scanner, run a discovery scan to get started.

Use other steps detailed below to manage your scans moving forward.

- Stop a scan
- Rescanning files

For more information, see Learn about the Microsoft Purview Information Protection scanner.

> **Tip**
>
> While most customers will perform these procedures in the admin portal, you may need to work in PowerShell only.
>
> For example, if you are working in an environment without access to the Azure portal, such as Azure China 21Vianet scanner servers, authenticate to the PurviewInformationProtection PowerShell module, and then continue with instructions in this article for PowerShell only.

Run a discovery cycle and view reports for the scanner

Use the following procedure after you've configured and installed your scanner to get an initial understanding of your content.

Perform these steps again as needed when your content changes.

1. Start a scan on your content scan job.

Do either of the following to start a content scan job:

- **Use the Microsoft Purview portal.** On the **Information protection scanner - Content scan jobs** pane, select your content scan jobs, and then select the **Scan now** option. The **Scan now** option only appears once a content scan job is selected.
- **Use a PowerShell command.** Run `Start-Scan` to start the scan.

2. Wait for the scanner to complete its cycle. The scan completes when the scanner has crawled through all the files in the specified data stores.

   Do either of the following to monitor scanner progress:

   - **Use the Microsoft Purview portal.** On the **Information protection scanner - Content scan jobs** pane, select **Refresh**.

     Wait until you see values for the **LAST SCAN RESULTS** column and the **LAST SCAN (END TIME)** column.

   - **Use a PowerShell command.** Run `Get-ScanStatus` to monitor the status change.

3. When the scan is complete, review the reports stored in the **%*localappdata*%\Microsoft\MSIP\Scanner\Reports** directory.

   - The .txt summary files include the time taken to scan, the number of scanned files, and how many files had a match for the information types.
   - The .csv files have more details for each file. This folder stores up to 60 reports for each scanning cycle and all but the latest report is compressed to help minimize the required disk space.

     When a scan is completed, a `Summary_<x>.txt` file is created with the scan summary.

> **Note**
>
> Scanners send collected data information to Microsoft Purview Information Protection every five minutes, so that you can view the results in near real time from the admin portal. For more information, see Analytics and central reporting for Azure Information Protection.
>
> The admin portal displays information about the last scan only. If you need to see the results of previous scans, return to the reports that are stored on the scanner

computer, in the %*localappdata*%\Microsoft\MSIP\Scanner\Reports folder.

Initial configurations instruct you to set the **Info types to be discovered** to **Policy only**. This configuration means that only files that meet the conditions you've configured for automatic classification are included in the detailed reports.

If you don't see any labels applied, check that your label configuration includes automatic rather than recommended classification, or enable **Treat recommended labeling as automatic** (available in scanner version 2.7.x.x and above).

If the results are still not as you expect, you might need to reconfigure the conditions that you specified for your labels. If that's the case, reconfigure the conditions as needed, and repeat this procedure until you're satisfied with the results. Then, update your configuration automatically, and optionally protection.

Changing log levels or locations

Change the level of logging by using the *ReportLevel* parameter with Set-ScannerConfiguration.

The report folder location or name can't be changed. If you want to store reports in a different location, consider using a directory junction for the folder.

For example, use the Mklink command: `mklink /j D:\Scanner_reports C:\Users\aipscannersvc\AppData\Local\Microsoft\MSIP\Scanner\Reports`

If you've performed these steps after an initial configuration and installation, continue with Configure the scanner to apply classification and protection.

Stopping a scan

To stop a currently running scan before it's complete, use either of the following methods:

- **Microsoft Purview portal.** Select **Stop scan**:
- **Run a PowerShell command.** Run the following command:

```
PowerShell

Stop-Scan
```

Rescanning files

For the first scan cycle, the scanner inspects all files in the configured data stores. For subsequent scans, only new or modified files are inspected.

It's useful to inspect all files again when you want the reports to include all files, when you have changes that you want to apply across all files, and when the scanner runs in discovery mode.

To manually run a full rescan in the Microsoft Purview portal:

1. Navigate to the **Information protection scanner - Content scan jobs** pane in the Microsoft Purview portal.
2. Select your content scan job from the list, and then select the **Rescan all files** option:

When a full scan is complete, the scan type automatically changes to incremental so that for subsequent scans, only new or modified files are scanned again.

> **Tip**
>
> If you've made changes to your content scan job, the portal will prompt you to skip a full rescan. To ensure that your rescan occurs, make sure to select **No** in the prompt that appears.

Trigger a full rescan by modifying your settings

Earlier versions of the scanner scanned all files whenever the scanner detected new or changed settings for automatic and recommended labeling. The scanner automatically refreshed the policy every four hours.

In scanner versions 2.8.85.0 or later, the information protection scanner skips the full rescan for updated settings to ensure consistent performance. Make sure that you run a full rescan manually as needed.

For example, if you've changed **Sensitivity policy** settings from **Enforce** = **Off** to **Enforce** = **On**, make sure to run a full rescan to apply your labels across your content.

> **Note**

In scanner version 2.7.101.0 and lower, you may want to refresh the policy sooner than every four hours, such as while testing. In such cases, manually delete the contents of the **%LocalAppData%\Microsoft\MSIP\mip<processname>\mip** directory and restart the Microsoft Purview Information Protection Scanner service.

If you've also changed encryption settings for your sensitivity labels, wait an extra 15 minutes from when you saved the updated encryption settings before restarting the Microsoft Purview Information Protection Scanner service.

Next steps

You can also use PowerShell to interactively classify and protect files from your desktop computer. For more information about this and other scenarios that use PowerShell, see Set up the information protection client using PowerShell.

Feedback

Was this page helpful? Yes No

Microsoft Purview Information Protection scanner feature control (preview)

Scanner administrators can enable, disable, and configure Microsoft Purview Information Protection scanner features at the cluster level. Use PowerShell to manage feature settings across all nodes in a cluster.

Admin-controlled feature configuration is available with Microsoft Purview Information Protection client and scanner version 3.2.89.0 (preview) or later.

How feature control works

Scanner features are configured once per scanner cluster. The settings are stored centrally in the shared scanner cluster database, which means:

- An administrator runs a configuration cmdlet once from any node in the cluster, and all nodes pick up the change on the next scan cycle.
- No service restart is required.
- Nodes can't get into mismatched feature states.
- Disabling a feature stops the action that the feature performs (for example, writing to a new database table or producing a new file). Existing data isn't deleted, so a disabled feature can be re-enabled later without data loss.
- Features that aren't explicitly configured use their default state (typically **off**). Defaults for each feature are listed under Currently supported features.

Custom Reporting is the first feature delivered through admin-controlled feature configuration. For details about that feature, see Custom Reporting (preview).

Evolution of control: PowerShell and the Microsoft Purview portal

Admin-controlled feature configuration is designed to evolve over time. Supported features are initially configured by using PowerShell. As features become available in the Microsoft Purview portal, the portal becomes the source of truth for those features.

The following rules apply when both surfaces are available for a feature:

- **PowerShell first.** Until a feature is exposed in the Microsoft Purview portal, administrators configure it from PowerShell on any node in the cluster. Settings configured this way are tracked as managed by PowerShell.
- **Portal takes precedence.** When a feature has been configured from the Microsoft Purview portal, the portal becomes the source of truth for that feature. PowerShell attempts to change a portal-managed feature are ignored with a warning, and the cluster state remains unchanged. Other features in the same `-FeatureSettings` call that aren't portal-managed are still applied.
- **No synchronization.** Settings aren't synchronized between PowerShell and the portal. Not every scanner feature will end up being configurable from the portal. When a feature *is* available in the portal and has been configured there, that portal-configured setting takes precedence and prevents updates from PowerShell. There's a single source of truth for each feature at any time.
- **Get always reflects state.** Regardless of where a feature was configured, Get-ScannerConfiguration returns the current state and source for every configured feature. Administrators can audit the cluster from PowerShell at any time.

This PowerShell-to-portal management model means scanner administrators can adopt new features through PowerShell as soon as they reach public preview, and switch to portal-based management at their own pace as features become available there.

Supported cmdlets and the `-FeatureSettings` parameter

The following cmdlets accept the `-FeatureSettings` parameter:

Expand table

Cmdlet	Purpose
Install-Scanner	Configure features as part of a new scanner node installation.
Set-ScannerConfiguration	Configure features on an existing scanner cluster. The change applies to every node in the cluster.
Get-ScannerConfiguration	Return the current state and source for every configured feature, including features that were configured from the Microsoft Purview portal.

`-FeatureSettings` accepts a PowerShell **hashtable** of feature name and setting pairs, using the same pattern as `-AdvancedSettings`:

- The hashtable key is the feature name (for example, `CustomReporting`).
- The hashtable value is the setting for that feature. Supported values for each feature are listed in the Currently supported features table.

Both the feature name and its value are validated before any change is written. If any entry is invalid, the cmdlet terminates with an error and no entries from the hashtable are applied.

For background on the `-AdvancedSettings` pattern, including PowerShell quoting and hashtable tips, see PowerShell tips for specifying the advanced settings and the -AdvancedSettings parameter on **Set-Label**.

Example: Enable Custom Reporting on an existing cluster

Run the following from any node in the scanner cluster:

PowerShell

```
Set-ScannerConfiguration -FeatureSettings @{CustomReporting=$true}
```

All nodes in the cluster pick up the change on their next scan cycle.

Example: Enable Custom Reporting at install time

Use the following command to enable Custom Reporting when you install a new scanner node:

PowerShell

```
Install-Scanner -SqlServerInstance SQLSERVER1 -Cluster Europe -FeatureSettings
@{CustomReporting=$true}
```

Example: Review the current feature state

Run the following command to review the current feature configuration for the cluster:

PowerShell

```
Get-ScannerConfiguration
```

For each configured feature, the `Features` line in the output shows the current state and which surface configured it (PowerShell or the Microsoft Purview portal). For example, with Custom Reporting enabled from PowerShell:

```
Features             : {CustomReporting: True (Source: PowerShell)}
```

Features that have never been configured don't appear and use their default state.

Disable a feature

To turn Custom Reporting off, set its value to `$false`. Existing data that the feature wrote isn't deleted; the scanner stops performing the feature's action on the next scan cycle.

PowerShell

```powershell
Set-ScannerConfiguration -FeatureSettings @{CustomReporting=$false}
```

Currently supported features

The following table lists the currently supported scanner features and their configuration values:

Expand table

Feature name	Default	Supported values	Description
`CustomReporting`	Off	`$true`, `$false`	Populates the Custom Reporting tables and columns in the scanner cluster database. See Custom Reporting (preview).

Additional features will be added over time and listed in the Microsoft Purview Information Protection client release notes.

Related content

- Custom Reporting (preview)
- Configure and install the Microsoft Purview Information Protection scanner
- Microsoft Purview Information Protection client release notes
- Install-Scanner
- Set-ScannerConfiguration
- Get-ScannerConfiguration

ⓘ**Note:** The author created this article with assistance from AI. Learn more

Last updated on 06/26/2026

Microsoft Purview Information Protection scanner Custom Reporting (preview)

This article describes the Microsoft Purview Information Protection scanner **Custom Reporting** feature (preview). Custom Reporting gives scanner administrators the data they need in the scanner cluster database to build their own reports against scan results—including labeling, protection state, and matched sensitive information types (SITs).

Custom Reporting is available with Microsoft Purview Information Protection client and scanner version 3.2.89.0 or later. Custom Reporting is enabled by an administrator through Scanner feature control.

What Custom Reporting enables

Without Custom Reporting, the scanner produces per-scan CSV and TXT reports and stores a limited set of operational data in the cluster database. To build a complete picture of file state—what changed, what's labeled, what's protected, and what kinds of sensitive data exist where—administrators have to combine multiple CSV exports across scan cycles and load them into a separate reporting tool.

Custom Reporting moves file-state and SIT-match data into the scanner cluster database so that administrators can:

- Query the present and previous label, protection state, and SIT counts for every scanned file in a repository.
- Compute deltas between scans (for example, the change in number of matched SITs per file).
- See which sensitive information types were matched on which files, and how many matches per type.
- Connect the scanner cluster database to a reporting tool of their choice (for example, Power BI, an enterprise reporting warehouse, or a SQL-based dashboarding tool) without first having to stitch CSV exports together.

The scanner remains the source of truth. Custom Reporting writes the additional reporting data to the scanner cluster database on the next scan cycle after the feature is enabled.

Database schema

The reporting tables and columns required by Custom Reporting were first added to the scanner cluster database schema in client version 3.2.57.0. Until Custom Reporting is enabled, the new columns and tables exist but aren't populated. Existing scanner functionality is unchanged.

> **Note**
>
> You don't need to install client version 3.2.57.0 first. Whether you're installing the scanner fresh or upgrading from any earlier version, the scanner deploys the full database schema (including the Custom Reporting tables and columns) on install or upgrade.

When Custom Reporting is enabled by using Scanner feature control, the scanner starts populating the following data on the next scan cycle.

Additions to `dbo.ScannerFiles`

`dbo.ScannerFiles` continues to hold one row per scanned file. Custom Reporting populates the following additional columns to capture file state and the delta between the current and previous scan:

Expand table

Column	Type	Description
`LabelName`	`NVARCHAR(MAX)`, nullable	Current label name applied to the file. `NULL` if unlabeled.
`PrevLabelId`	`NVARCHAR(MAX)`, nullable	Label ID applied at the previous scan, stored as a string. `NULL` if previously unlabeled.
`PrevLabelName`	`NVARCHAR(MAX)`, nullable	Label name applied at the previous scan. `NULL` if previously unlabeled.
`ProtectionState`	`NVARCHAR(MAX)`, nullable	Current protection state for the file at the end of the scan.
`PrevProtectionState`	`NVARCHAR(MAX)`, nullable	Protection state recorded at the previous scan.
`ClassificationCount`	`INT`, default `0`	Count of sensitive information type matches on the file at the current scan.
`LatestScanSessionId`	`UNIQUEIDENTIFIER`, nullable	Identifies the most recent scan session that touched the file. Used to compute deltas across scan cycles.

Column	Type	Description
`FileStatus`	`NVARCHAR(MAX)`, nullable	Final disposition of the file in the scan cycle (for example, the requires-justification status set by `ProcessJob`, or `Failed` for files that couldn't be processed).

New table: `dbo.MatchedClassificationAction`

`dbo.MatchedClassificationAction` is a new table that stores the matched sensitive information types per file per scan. Each row represents one matched SIT for a single file in a single scan session.

Expand table

Column	Type	Description
`Id`	`BIGINT IDENTITY` (primary key)	Surrogate key for the row.
`FilePath`	`NVARCHAR(MAX)`, nullable	Full path of the file that the SIT was matched on.
`FileHashPath`	`BINARY(64)`, nullable	Hash of the file path. Joins to `dbo.ScannerFiles.HashPath` and is indexed for join performance.
`ScanSessionId`	`UNIQUEIDENTIFIER`, nullable	Identifies the scan session in which the SIT was matched. Joins to `dbo.ScannerFiles.ScanSessionId` and is indexed.
`MatchedInformationTypeName`	`NVARCHAR(MAX)`, nullable	Display name of the matched sensitive information type (for example, `U.S. social security number (SSN)`).
`MatchedInformationTypeId`	`UNIQUEIDENTIFIER`, nullable	GUID of the matched SIT. The same GUID is used in the local Microsoft Purview Information Protection logs for `Workload=OnPremisesFileShareScanner`.
`MatchedInformationTypeCount`	`INT`, default `0`	Number of matches for this SIT in the file.
`ConfidenceScore`	`INT`, default `0`	Confidence score of the match.

New table: `dbo.ScannedFilesArchive`

`dbo.ScannedFilesArchive` is a new table that stores a historical archive of files the scanner processed, with rows tied to the scan session that created them. Unlike the live `dbo.ScannerFiles`

table, which only keeps the latest state of each file, this table preserves prior archived states for change tracking, reporting, and auditing.

Files that the scanner skips because nothing has changed since the previous scan session aren't re-inserted, so later scan sessions don't contain a complete snapshot of all files. Instead, only the existing archive entry from the session that last processed the file remains, which keeps the archive lean while preserving a history of meaningful changes.

Expand table

Column	Type	Description
`Id`	`BIGINT IDENTITY` (primary key)	Surrogate key for the row.
`HashPath`	`BINARY(64)`	Hash of the file path. Joins to `dbo.ScannerFiles.HashPath`.
`FullPath`	`NVARCHAR(MAX)`	Full path of the file at the time of the scan session.
`ScanSessionId`	`UNIQUEIDENTIFIER`	The scan session that produced this archive row.
`LastModifiedTime`	`DATETIME2`, nullable	Last-modified timestamp of the file as observed during the scan session.
`LabelId`	`UNIQUEIDENTIFIER`, nullable	Label ID applied to the file at the time of the scan session.
`IssueCount`	`INT`, default `0`	Number of issues recorded against the file in this scan session.
`EndScan`	`DATETIME2`, nullable	Time the file completed processing in this scan session.
`JobStartTime`	`DATETIME2`, nullable	Time the processing job for this file started.
`LabelName`	`NVARCHAR(MAX)`, nullable	Label name applied at the time of this scan session.
`PrevLabelId`	`NVARCHAR(MAX)`, nullable	Label ID at the previous scan session, stored as a string.
`PrevLabelName`	`NVARCHAR(MAX)`, nullable	Label name at the previous scan session.
`ProtectionState`	`NVARCHAR(MAX)`, nullable	Protection state at the time of this scan session.
`PrevProtectionState`	`NVARCHAR(MAX)`, nullable	Protection state at the previous scan session.
`ClassificationCount`	`INT`, default `0`, nullable	Count of sensitive information type matches in this scan session.
`LatestScanSessionId`	`UNIQUEIDENTIFIER`, nullable	The latest scan session known to have touched the file at the time this archive row was recorded.

Column	Type	Description
`FileStatus`	`NVARCHAR(MAX)`, nullable	Disposition of the file in the scan session (for example, `Modified` or `Deleted`).

`dbo.ScanSummary` (per-scan totals) is unchanged.

Example questions Custom Reporting can answer

Once Custom Reporting is populating the database, administrators can run queries such as:

- Which repositories have the highest concentration of matched SITs, and how has that concentration changed since the last scan?
- Which files were labeled or relabeled during the most recent scan cycle, and what was the previous label?
- Which files are still unlabeled but contain matches for one or more sensitive information types?
- Which sensitive information types are most prevalent in a given repository, and at what confidence?
- Which files transitioned from unprotected to protected (or the reverse) since the last scan?

Enable Custom Reporting

Custom Reporting is turned on through admin-controlled feature configuration. From any node in the scanner cluster, run:

PowerShell

```powershell
Set-ScannerConfiguration -FeatureSettings @{CustomReporting=$true}
```

To enable Custom Reporting at install time on a new scanner node, use the `-FeatureSettings` parameter with Install-Scanner:

PowerShell

```powershell
Install-Scanner -SqlServerInstance SQLSERVER1 -Cluster Europe -FeatureSettings
@{CustomReporting=$true}
```

To confirm the current state, run:

PowerShell

```
Get-ScannerConfiguration
```

The change takes effect on every node in the cluster on the next scan cycle. No service restart is required.

To stop populating the Custom Reporting columns and table, run:

PowerShell

```
Set-ScannerConfiguration -FeatureSettings @{CustomReporting=$false}
```

Disabling Custom Reporting stops new writes. Data already written to the reporting columns and tables isn't deleted, so the feature can be re-enabled later without data loss.

Plan your scanner cluster database for reporting

When you turn on Custom Reporting, the scanner cluster database stores significantly more data:

- Additional columns per file in `dbo.ScannerFiles`
- One row per matched SIT per file per scan in `dbo.MatchedClassificationAction`
- Archive rows in `dbo.ScannedFilesArchive` for files the scanner processes and archives during a scan session (not for every unchanged file in every cycle)

Building reports against the scanner cluster database also adds a read workload that runs alongside the scanner's operational reads and writes.

Before you enable Custom Reporting in production, consider:

- **Cluster size and scan volume.** Estimate the number of files scanned per cycle, the typical number of SIT matches per file, and how often you scan. Use those numbers to size the database.
- **Operational impact.** Reporting queries against the same database that the scanner is writing to can compete with the scanner for resources during a scan cycle.
- **The reporting workload you actually want.** Interactive dashboards, scheduled extracts, and ad-hoc queries each have very different impacts.

For most production deployments, we recommend hosting the scanner cluster database on SQL Server Enterprise so that you can dedicate a read-only replica to reporting. With a SQL Server Always On availability group, you can route reporting traffic to a readable secondary replica. Reporting queries then run on the secondary replica instead of competing with the scanner's operational workload on the primary.

This separation lets reporting tools (for example, Power BI) connect to the read-only replica and refresh on their own cadence without affecting scan throughput on the primary.

> **Note**
>
> The scanner itself always reads from and writes to the primary database. Only your custom reporting workload should be pointed at a read-only replica.

Limitations during preview

- Custom Reporting doesn't include a built-in dashboard in this preview. Customers build their own reports against the scanner cluster database.
- Settings configured by admin-controlled feature configuration aren't synchronized with the Microsoft Purview portal. Not every scanner feature will end up being configurable from the portal. For features that are also configurable from the Microsoft Purview portal, a setting configured in the portal takes precedence and prevents updates from PowerShell.

Related content

- Scanner feature control (preview)
- Configure and install the Microsoft Purview Information Protection scanner
- Microsoft Purview Information Protection client release notes
- Always On availability groups: readable secondary replicas

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 05/04/2026

Sensitive Information Types supported by Microsoft Purview Information Protection scanner

Article • 07/01/2024

The Microsoft Purview Information Protection scanner supports the following sensitive information types (SITs).

- ABA Routing Number
- Argentina National Identity (DNI) Number
- Argentina Unique Tax identification Key (CUIT/CUIL)
- Australia Bank Account Number
- Australia Driver's License Number
- Australia Medical Account Number
- Australia Passport Number
- Australia Tax File Number
- Australia Business Number
- Australia Company Number
- Austria Driver's License Number
- Austria Identity Card
- Austria Passport Number
- Austria Social Security Number
- Austria Tax identification Number
- Austria Value Added Tax (VAT) Number
- Azure DocumentDB Auth Key
- Azure IAAS Database Connection String and Azure SQL Connection String
- Azure IoT Connection String
- Azure Publish Setting Password
- Azure Redis Cache Connection String
- Azure SAS
- Azure Service Bus Connection String
- Azure Storage Account Key
- Azure Storage Account Key (Generic)
- Belgium Driver's License Number
- Belgium National Number
- Belgium Passport Number
- Belgium Value Added Tax Number
- Brazil CPF Number

- Brazil Legal Entity Number (CNPJ)
- Brazil National ID Card (RG)
- Bulgaria Driver's License Number
- Bulgaria Passport Number
- Bulgaria Uniform Civil Number
- Canada Bank Account Number
- Canada Driver's License Number
- Canada Health Service Number
- Canada Passport Number
- Canada Personal Health identification Number (PHIN)
- Canada Social Insurance Number
- Chile Identity Card Number
- China Resident Identity Card (PRC) Number
- Credit Card Number
- Croatia Driver's License Number
- Croatia Identity Card Number
- Croatia Passport Number
- Croatia Personal identification (OIB) Number
- Cyprus Driver's License Number
- Cyprus Identity Card
- Cyprus Passport Number
- Cyprus Tax identification Number
- Czech Driver's License Number
- Czech Personal Identity Number
- Czech Passport Number
- Denmark Driver's License Number
- Denmark Passport Number
- Denmark Personal identification Number
- Drug Enforcement Agency (DEA) Number
- EU Debit Card Number
- EU Driver's License Number
- EU National identification Number
- EU Passport Number
- EU Social Security Number (SSN) or Equivalent identification
- EU Tax identification Number (TIN)
- Estonia Driver's License Number
- Estonia Passport Number
- Estonia Personal identification Code
- Finland Driver's License Number
- Finland European Health Insurance Number
- Finland National ID

- Finland Passport Number
- France Driver's License Number
- France Health Insurance Number
- France National ID Card (CNI)
- France Passport Number
- France Social Security Number (INSEE)
- France Tax identification Number
- France Value Added Tax Number
- Germany Driver's License Number
- Germany Passport Number
- Germany Identity Card Number
- Germany Tax identification Number
- Germany Value Added Tax Number
- Greece Driver's License Number
- Greece National ID Card
- Greece Passport Number
- Greece Social Security Number (AMKA)
- Greece Tax identification Number
- Hong Kong Identity Card (HKID) Number
- Hungary Driver's License Number
- Hungary Passport Number
- Hungary Personal identification Number
- Hungary Social Security Number (TAJ)
- Hungary Tax identification Number
- Hungary Value Added Tax Number
- IP Address
- IP Address v4
- IP Address v6
- India Driver's License Number
- India GST Number
- India Permanent Account Number (PAN)
- India Unique identification (Aadhaar) Number
- India Voter Id Card
- Indonesia Identity Card (KTP) Number
- International Banking Account Number (IBAN)
- International Classification of Diseases (ICD-9-CM)
- International Classification of Diseases (ICD-10-CM)
- Ireland Driver's License Number
- Ireland Passport Number
- Ireland Personal Public Service (PPS) Number
- Israel Bank Account Number

- Israel National identification number
- Italy Driver's License Number
- Italy Fiscal Code
- Italy Passport Number
- Italy Value Added Tax Number
- Japan Bank Account Number
- Japan Driver's License Number
- Japan My Number Corporate
- Japan My Number Personal
- Japan Passport Number
- Japan Residence Card Number
- Japan Resident Registration Number
- Japan Social Insurance Number (SIN)
- Latvia Driver's License Number
- Latvia Passport Number
- Latvia Personal Code
- Lithuania Driver's License Number
- Lithuania Passport Number
- Lithuania Personal Code
- Luxembourg Driver's License Number
- Luxembourg National identification Number (Natural persons)
- Luxembourg National identification Number (Non-natural persons)
- Luxembourg Passport Number
- Malaysia Identity Card Number
- Malaysia Passport Number
- Malta Driver's License Number
- Malta Identity Card Number
- Malta Passport Number
- Malta Tax ID Number
- Medicare Beneficiary Identifier (MBI) card
- Mexico Unique Population Registry Code (CURP)
- Netherlands Citizen's Service (BSN) Number
- Netherlands Driver's License Number
- Netherlands Passport Number
- Netherlands Tax identification Number
- Netherlands Value Added Tax Number
- New Zealand Driver License Number
- New Zealand Inland Revenue number
- New Zealand Ministry of Health Number
- New Zealand Social Welfare Number
- New Zealand bank account number

- Norway Identity Number
- Philippines Passport Number
- Philippines Unified Multi-Purpose ID Number
- Poland Driver's License Number
- Poland Identity Card
- Poland National ID (PESEL)
- Poland Passport Number
- Poland REGON Number
- Poland Tax identification Number
- Portugal Citizen Card Number
- Portugal Driver's License Number
- Portugal Passport Number
- Portugal Tax identification Number
- Qatari ID Card Number
- Romania Driver's License Number
- Romania Passport Number
- Romania Personal Numerical Code (CNP)
- Russia Passport Number (Domestic)
- Russia Passport Number (International)
- SQL Server Connection String
- SWIFT Code
- Saudi Arabia National ID
- Singapore Driver's License Number
- Singapore National Registration Identity Card (NRIC) Number
- Singapore Passport Number
- Slovakia Driver's License Number
- Slovakia Passport Number
- Slovakia Personal Number
- Slovenia Driver's License Number
- Slovenia Passport Number
- Slovenia Tax identification Number
- Slovenia Unique Master Citizen Number
- South Africa identification Number
- South Korea Driver's License Number
- South Korea Passport Number
- South Korea Resident Registration Number
- Spain DNI
- Spain Driver's License Number
- Spain Passport Number
- Spain Social Security Number (SSN)
- Spain Tax identification Number

- Sweden Driver's License Number
- Sweden National ID
- Sweden Passport Number
- Sweden Tax identification Number
- Switzerland Social Security Number AHV
- Taiwan National ID
- Taiwan Passport Number
- Taiwan Resident Certificate (ARC/TARC)
- Thai Population identification Code
- Turkey National identification number
- U.K. Driver's License Number
- U.K. Electoral Roll Number
- U.K. National Health Service Number
- U.K. National Insurance Number (NINO)
- U.K. Unique Taxpayer Reference Number
- U.S. / U.K. Passport Number
- U.S. Bank Account Number
- U.S. Driver's License Number
- U.S. Individual Taxpayer identification Number (ITIN)
- U.S. Social Security Number (SSN)
- UAE Identity Card Number
- UAE Passport Number
- Ukraine Passport Number (Domestic)
- Ukraine Passport Number (International)

Feedback

Was this page helpful? Yes No

Upgrade the Microsoft Purview Information Protection scanner

Article • 06/28/2024

Use the following instructions to upgrade the information protection scanner to the latest version from the Microsoft Purview Information Protection client:

- From the Azure Information Protection unified labeling client (2.x)
- From a previous version of the scanner from the Microsoft Purview Information Protection client (3.x)

For the latest release information and support timelines for each version, see Microsoft Purview Information Protection client - Release management and supportability.

Upgrade the scanner from the Azure Information Protection client (version 2.x)

> **ⓘ Important**
>
> Upgrading the Microsoft Purview Information Protection scanner from client version 2.x to version 3.x requires additional and specific steps because service names and other components are renamed.
>
> To avoid misconfigurations that can result in the scanner no longer working, make sure you follow each step that follows.

1. On any scanner node in the cluster, run the following command and make a note of the SQL database name and cluster:

   PowerShell

   ```
   Get-AIPScannerConfiguration
   ```

2. Get ready to install the new version of the scanner on each node in the cluster by stopping the scanner service:

   PowerShell

```
Net Stop AIPScanner
```

3. Uninstall the existing scanner service:

```PowerShell
Uninstall-AIPScanner
```

4. Restart the Windows server computer to completely remove the service.

5. Uninstall version 2.x of the Azure Information Protection unified labeling client.

6. Install version 3.x of the Microsoft Purview Information Protection client

7. Install the scanner from the information protection client:

```PowerShell
Install-Scanner -SqlServerInstance <SQL\InstanceName> -Cluster
<cluster-name>
```

   For example, to upgrade the scanner that uses the SQL Server instance named SCANNER that runs on the server named SQLSERVER1 and uses the cluster name of EU as the database for the scanner configuration:

```PowerShell
Install-Scanner -SqlServerInstance SQLSERVER1\SCANNER -Cluster EU
```

   For more details and other examples that include using the SQL Server default instance and SQL Server Express, see the examples from Install-Scanner.

8. Stop the scanner service:

```PowerShell
Net Stop MIPScanner
```

9. Update the scanner database: On one scanner node, run the following command:

```PowerShell
Update-ScannerDatabase
```

10. Start the scanner service on the Windows server computer.

11. For each node in the cluster, start the scanner service:

```
PowerShell

Net Start MIPScanner
```

The scanner is now upgraded for the Microsoft Purview Information Protection client and ready to resume any scanning jobs that are configured.

Upgrade a previous version of the scanner from the information protection client (version 3.x)

Use the following instructions to upgrade the scanner from an earlier version from the Microsoft Purview Information Protection client (version 3.x).

For example, you're upgrading from a preview version of the Microsoft Purview Information Protection client to a general availability (GA) version. Or, you're upgrading from a GA version of the Microsoft Purview Information Protection client to the latest version of this client.

1. Get ready to install the new version of the scanner on each node in the cluster by stopping the scanner service:

```
PowerShell

Net Stop MIPScanner
```

2. Install the latest version of the Microsoft Purview Information Protection client.

3. Update the scanner database: On one scanner node, run the following command:

```
PowerShell

Update-ScannerDatabase
```

4. Start the scanner service on the Windows server computer.

5. For each node in the cluster, start the scanner service:

```
PowerShell
```

```
Net Start MIPScanner
```

The scanner is now upgraded to the latest version from the Microsoft Purview Information Protection client and ready to resume any scanning jobs that are configured.

Feedback

Was this page helpful? Yes No

Learn about data loss prevention

Organizations control sensitive information like:

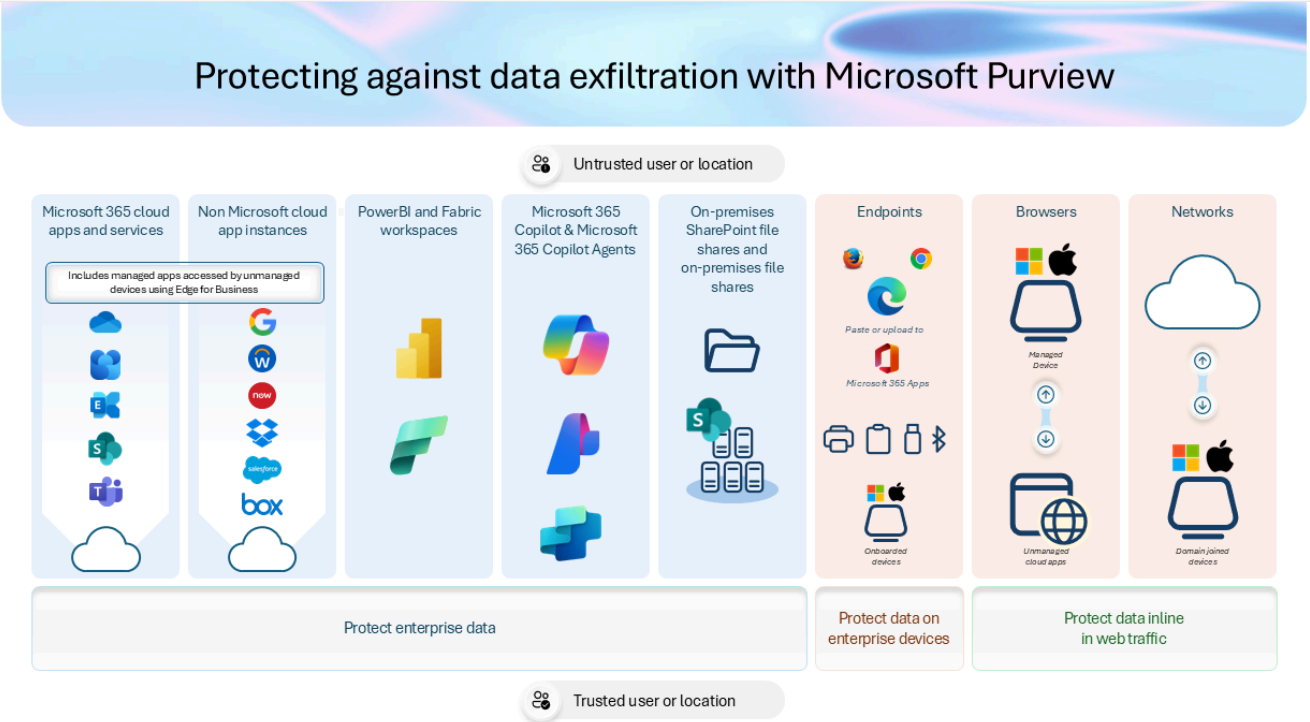
- financial data
- proprietary data
- credit card numbers
- health records
- Social Security numbers

To help protect this sensitive data, and to reduce the risk from oversharing, they need a way to help prevent their users from inappropriately sharing sensitive data with people who shouldn't have it. This practice is called data loss prevention (DLP).

In Microsoft Purview, you implement data loss prevention by defining and applying DLP policies. A DLP policy can help you identify, monitor, and automatically protect sensitive in **Enterprise applications & devices** and **Inline web traffic** data. DLP policies act on a variety of locations, methods of data transmission, and types of user activities.

DLP uses deep content analysis—not a simple text scan. It analyzes content:

- For primary data matches to keywords
- By the evaluation of regular expressions
- By internal function validation
- By secondary data matches that are in proximity to the primary data match
- DLP also uses machine learning algorithms and other methods to detect content that matches your DLP policies
- Inline by Microsoft Edge for business for Windows devices that haven't been onboarded into Microsoft Purview (preview) and Use Network Data Security to help prevent sharing sensitive information with unmanaged AI (preview)



Enterprise applications and devices

DLP monitors and protects against oversharing in enterprise apps and on devices. It targets Microsoft 365 locations, like Exchange and SharePoint, and locations you add, like on-premises file shares, endpoint devices, and non-Microsoft cloud apps. These locations and sources include:

- Microsoft 365 services, like Exchange, SharePoint, OneDrive accounts, and Teams chat and channel messages
- Office applications, such as Word, Excel, and PowerPoint
- Devices running Windows 10, Windows 11, and the three most recent versions of macOS
- Non-Microsoft cloud apps
- On-premises file shares and on-premises SharePoint
- Microsoft Fabric and Power BI workspaces
- Microsoft 365 Copilot and Copilot chat (preview)
- Managed cloud apps

Create DLP policies for **Enterprise applications & devices** to cover these locations.

Inline web traffic

DLP, with collection policies, monitors and protects against oversharing to **Unmanaged cloud apps** by targeting data transmitted on your network and in Microsoft Edge for Business. Create

policies that target Inline web traffic (preview) and Network activity (preview) to cover locations like:

- OpenAI ChatGPT—for **Edge for Business** and **Network** options
- Google Gemini—for **Edge for Business** and **Network** options
- DeepSeek—for **Edge for Business** and **Network** options
- Microsoft Copilot—for **Edge for Business** and **Network** options
- Over 34,000 cloud apps in the Microsoft Defender for Cloud Apps cloud app catalog—**Network** option only

Before you start

1. Administrative units
2. Learn about Microsoft Purview Data Loss Prevention
3. Plan for data loss prevention (DLP) - by working through this article you will:
   a. Identify stakeholders
   b. Describe the categories of sensitive information to protect
   c. Set goals and strategy
4. Collection Policies solution overview
5. Collection policy reference
6. Data Loss Prevention policy reference - this article introduces all the components of a DLP policy and how each one influences the behavior of a policy
7. Design a DLP policy - this article walks you through creating a policy intent statement and mapping it to a specific policy configuration.
8. Create and Deploy data loss prevention policies - This article presents some common policy intent scenarios that you map to configuration options, then it walks you through configuring those options.
9. Learn about investigating data loss prevention alerts - This article introduces you to the lifecycle of alerts from creation, through final remediation and policy tuning. It also introduces you to the tools you use to investigate alerts.

Microsoft Purview licensing

For information on licensing, see

- Microsoft 365 Enterprise Plans
- Microsoft 365 Service Descriptions

DLP in Microsoft Purview

DLP is just one of the Microsoft Purview tools that you use to help protect your sensitive items wherever they live or travel. You should understand the other tools in the Microsoft Purview tool set, how they interrelate, and work better together. See, Microsoft Purview tools to learn more about the information protection process.

Protective actions of DLP policies

DLP policies monitor the activities that users take on sensitive items and, if the policy conditions are met, take protective actions. For example, when a user attempts a prohibited action, like copying a sensitive item to an unapproved location or sharing medical information in an email, DLP can:

- show a pop-up policy tip to the user that warns them that they might be trying to share a sensitive item inappropriately
- block the sharing and, via a policy tip, allow the user to override the block and capture the users' justification
- block the sharing without the override option
- for data at rest, sensitive items can be locked and moved to a secure quarantine location
- for Teams chat, the sensitive information won't be displayed

All DLP monitored activities are recorded to the Microsoft 365 Audit log by default and routed to Activity explorer.

DLP lifecycle

A DLP implementation typically follows these major phases.

- Plan for DLP
- Prepare for DLP
- Deploy your policies in production

Plan for DLP

DLP monitoring and protection are native to the applications that users use every day. This helps to protect your organization's sensitive items from risky activities, even if your users are unaccustomed to data loss prevention thinking and practices. If your organization and your users are new to data loss prevention practices, the adoption of DLP might require a change to your business processes, and there will be a culture shift for your users. But, with proper planning,

testing and tuning, your DLP policies protect your sensitive items while minimizing any potential business process disruptions.

Technology planning for DLP

Keep in mind that DLP as a technology can monitor and protect your data at rest, data in use and data in motion across Microsoft 365 services, Windows 10, Windows 11, and macOS (three latest released versions) devices, on-premises file shares, and on-premises SharePoint. There are planning implications for the different locations, the type of data you want to monitor and protect, and the actions to be taken when a policy match occurs.

Business processes planning for DLP

DLP policies can block users from performing prohibited activities, like inappropriate sharing of sensitive information via email. As you plan your DLP policies, you must identify the business processes that touch your sensitive items. The business process owners can help you identify appropriate user behaviors that should be allowed and inappropriate user behaviors that should be protected against. You should plan your policies and deploy them in simulation mode, and evaluate their impact, before running them in more restrictive modes.

Organizational culture planning for DLP

A successful DLP implementation is as much dependent on getting your users trained and acclimated to data loss prevention practices as it is on well planned and tuned policies. Since your users are heavily involved, be sure to plan for training for them too. You can strategically use policy tips to raise awareness with your users before changing the policy status from simulation mode to more restrictive modes.

Prepare for DLP

You can apply DLP policies to data at rest, data in use, and data in motion in locations such as:

- Exchange Online email
- SharePoint sites
- OneDrive accounts
- Teams chat and channel messages
- Instances: Microsoft Defender for Cloud Apps
- Devices: Windows 10, Windows 11, and macOS (three latest released versions)
- On-premises repositories
- Fabric and Power BI workspaces

- Microsoft 365 Copilot (preview)

Each one has different prerequisites. Sensitive items in some locations, like Exchange online, can be brought under the DLP umbrella by just configuring a policy that applies to them. Others, such as on-premises file repositories, require a deployment of Microsoft Purview Information Protection scanner. You'll need to prepare your environment, code draft policies, and test them thoroughly before activating any blocking actions.

Deploy your policies in production

Design your policies

Start by defining your control objectives, and how they apply across each respective workload. Draft a policy that embodies your objectives. Feel free to start with one workload at a time, or across all workloads - there's no impact yet. For more information, see Create and deploy data loss prevention policies.

Implement policy in simulation mode

Evaluate the impact of the controls by implementing them with a DLP policy in simulation mode. Actions defined in a policy aren't applied while the policy is in simulation mode. It's ok to apply the policy to all workloads in simulation mode, so that you can get the full breadth of results, but you can start with one workload if you need to. For more information, see Policy Deployment.

Monitor outcomes and fine-tune the policy

While in simulation mode, monitor the outcomes of the policy and fine-tune it so that it meets your control objectives while ensuring you aren't adversely or inadvertently impacting valid user workflows and productivity. Here are some examples of things to fine-tune:

- Adjust the locations and people/places that are in or out of scope
- Tune the conditions that are used to determine if an item and what is being done with it matches the policy
- Refine the sensitive information definitions
- Add new controls
- Add new people
- Add new restricted apps
- Add new restricted sites

> ⓘ **Note**

Stop processing more rules doesn't work in simulation mode, even when it's turned on.

Enable the control and tune your policies

Once the policy meets all your objectives, turn it on. Continue to monitor the outcomes of the policy application and tune as needed.

> **Note**
>
> In general, policies take effect about an hour after being turned on.

DLP policy configuration overview

You have flexibility in how you create and configure your DLP policies. You can start from a predefined template and create a policy in just a few clicks or you can design your own from the ground up. No matter which you choose, all DLP policies require the same information from you.

1. **Choose what you want to monitor** - DLP comes with many predefined policy templates to help you get started or you can create a custom policy.

    - A predefined policy template, such as Financial data, Medical and health data, Privacy data all for various countries and regions.
    - A custom policy that uses the available sensitive information types (SIT), retention labels, and sensitivity labels.

2. **Choose administrative scoping** - DLP supports assigning Administrative Units to some **Enterprise applications & devices** policies. Administrators who are assigned to an administrative unit can only create and manage policies for the users, groups, distribution groups, accounts, and sites that they're assigned to. So, policies can be applied to all users, groups, and sites by an unrestricted administrator, or they can be scoped to administrative units. See, Policy Scoping for more DLP specific details. See, Administrative units for the details on administrative units across Microsoft Purview Information Protection.

3. **Choose where you want to monitor** - You pick one or more locations that you want DLP to monitor for sensitive information. You can monitor:

Expand table

location	include/exclude by
Exchange email	distribution groups
SharePoint sites	sites
OneDrive accounts	accounts or distribution groups
Teams chat and channel messages	account or distribution group
Windows 10, Windows 11, and macOS (three latest released versions) devices	users and groups + devices and device groups
Microsoft Cloud App Security	instance
On-premises repositories	repository file path
Fabric and Power BI	workspaces
Microsoft 365 Copilot (preview)	account or distribution group

> ⓘ **Note**
>
> The users and groups mentioned above should be Online users and M365, Exchange online, and Microsoft Entra groups

4. **Choose the conditions that must be matched for a policy to be applied to an item** - You can accept preconfigured conditions or you can define custom conditions. Some examples are:

    - item contains a specified type of sensitive information that is being used in a certain context. For example, 95 social security numbers being emailed to recipient outside your org.
    - item has a specified sensitivity label
    - item with sensitive information is shared either internally or externally

5. **Choose the action to take when the policy conditions are met** - The actions depend on the location where the activity is happening. Some examples are:

    - SharePoint/Exchange/OneDrive: Block people who are outside your organization from accessing the content. Show the user a tip and send them an email notification that they're taking an action that is prohibited by the DLP policy.
    - Teams Chat and Channel: Block sensitive information from being shared in the chat or channel.

- Windows 10, Windows 11, and macOS (three latest released versions) Devices: Audit or restrict copying a sensitive item to a removable USB device.
- Office Apps: Show a popup notifying the user that they're engaging in a risky behavior and block or block but allow override.
- On-premises file shares: move the file from where it's stored to a quarantine folder.

> **Note**
>
> The conditions and the actions to take are defined in an object called a *rule*.

Create and deploy a DLP policy

All DLP policies are created and maintained in the Microsoft Purview portal. See, Create and Deploy data loss prevention policies for more information.

After you create a DLP policy, it's stored in a central policy store, and then synced to the various content sources, including:

- Exchange, and from there to Outlook on the web and Outlook
- OneDrive
- SharePoint sites
- Office desktop programs (Excel, PowerPoint, and Word)
- Microsoft Teams channels and chat messages

After the policy is synced to the right locations, it starts to evaluate content and enforce actions.

Viewing policy application results

DLP reports a vast amount of information to Microsoft Purview from monitoring policy matches and actions, to user activities. You need to consume and act on that information to tune your policies and triage actions taken on sensitive items. The telemetry goes into the Microsoft 365 audit Logs first, is processed, and makes its way to different reporting tools. Each reporting tool has a different purpose.

Overview page

The DLP **Overview** page gives you quick access to important information about your DLP policies, including:

- **Policy sync status**
- **Device status**
- **Top activites detected**
- **Device overall health**

You can investigate incidents for Microsoft Purview Data Loss Prevention (DLP) from the Microsoft Defender portal **Incidents & alerts** > **Incidents**. See Investigate data loss incidents with Microsoft Defender XDR and Investigate alerts in Microsoft Defender XDR.

DLP Alerts

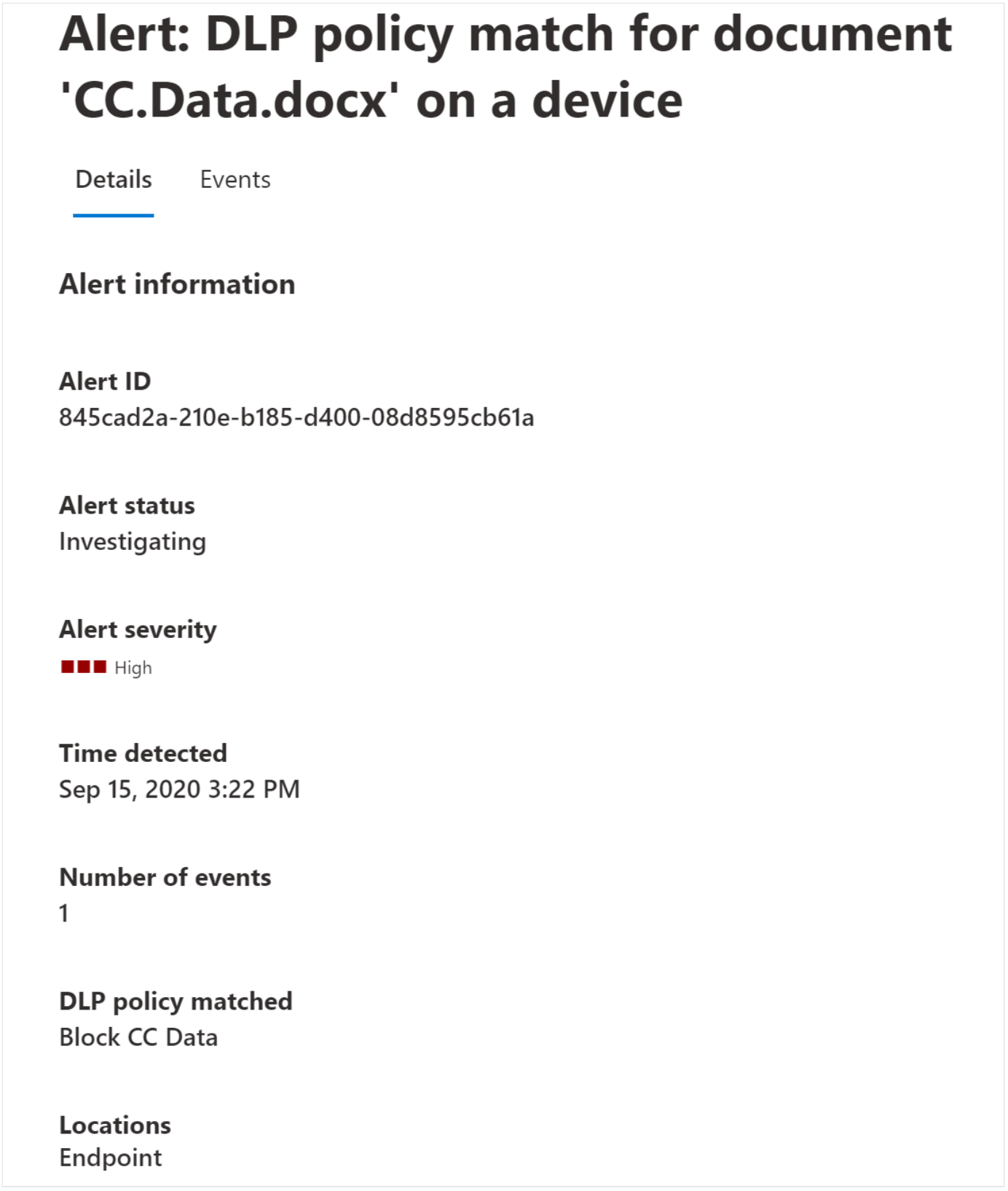
DLP can generate alerts when a user(s) performs an activity that meets the criteria of a rule in a DLP policy, and you have incident reports configured to generate alerts. Depending on your subscription level, alerts can be aggregated on a time window/rule basis or on a time windows/user basis(preview).

DLP posts the alert for investigation in the DLP Alerts dashboard. Use the DLP Alerts dashboard to view alerts, triage them, set investigation status, and track resolution. Alerts are also routed to Microsoft Defender portal where you can do all the alert dashboard tasks plus more.

DLP alerts are available in the Microsoft Defender portal for six months. They're only available in the Microsoft Purview DLP alerts dashboard for 30 days.

If you're an administrative unit restricted admin, you'll only see the DLP alerts for your administrative unit.

Here's an example of alerts generated by policy matches and activities from Windows 10 devices.



You can also view details of the associated event with rich metadata in the same dashboard.

Event: Sensitive info in 'CC.Data.docx' - File copied to clipboard

Details | Sensitive info types

Event details

ID

Location
Endpoint

Time of activity
Sep 15, 2020 3:19 PM

Impacted entities

User



Hostname

IP address

File
CC.Data.docx

File path
C:\DLP.test.files\CC.Data.docx

Policy details

DLP policy matched
Block CC Data

Rule matched
Rule to stop sharing credit card

Sensitive info types detected
Credit Card Number (1, 85%)

Violating action
File copied to clipboard

Note

Alerts are generated differently for emails than they are for SharePoint or OneDrive items. In SharePoint and OneDrive, DLP scans existing items as well as new ones and generates an alert whenever a match is found. In Exchange, new email messages are scanned and an alert is generated if there's a policy match. DLP ***does not*** scan or match previously existing email items that are stored in a mailbox or archive.

For more information on Alerts, see:

- Alerts in DLP policies: Describes alerts in the context of a DLP policy.
- Get started with data loss prevention alerts: Covers the necessary liscensing, permissions, and prerequisites for DLP alerts and alert reference details.
- Create and deploy data loss prevention policies: Includes guidance on alert configuration in the context of creating a DLP policy.
- Learn about investigating data loss prevention alerts: Covers the various methods for investigating of DLP alerts.
- Investigate data loss incidents with Microsoft Defender XDR: How to investigate DLP alerts in Microsoft Defender portal.

DLP Activity Explorer and reports

The Activity explorer tab on the DLP page has multiple filters you can use to view DLP events. Use this tool to review activity related to content that contains sensitive info or has labels applied, such as what labels were changed, files were modified, and matched a rule.

You can view the last 30 days of DLP information in Activity Explorer using these preconfigured filters:

- Endpoint DLP activities
- Files containing sensitive info types
- Egress activities
- DLP policies that detected activities
- DLP policy rules that detected activities

Expand table

To see this information	Select this activity
User overrides	**DLP rule undo**
Items that match a DLP rule	**DLP rule matched**

You can also access DLP report using via these cmdlets in the Security & Compliance PowerShell.

1. Connect to Security & Compliance PowerShell

Use these cmdlets:

- Get-DlpDetailReport

- Get-DlpDetectionsReport
- Get-DlpSiDetectionsReport

However, DLP reports need to pull data from across Microsoft 365, including Exchange. For this reason, the following cmdlets for DLP reports are available in Exchange PowerShell. To use the cmdlets for these DLP reports, take the following steps:

1. Connect to Exchange PowerShell

Use these cmdlets:

- Get-DlpDetailReport
- Get-MailDetailDlpPolicyReport

DLP Rule Undo event

A DLP Rule Undo event occurs when a previously matched DLP rule no longer applies to a document, causing the rule to be effectively reversed. This can happen in two main scenarios: when a user modifies content so it no longer meets the rule conditions, or when a policy or rule is changed and triggers re-evaluation of already-classified content. During this re-evaluation, if a document no longer satisfies the rule criteria, a DLP Rule Undo event is generated. For example, if a policy is configured to detect documents containing at least five Social Security numbers and a user uploads a document with ten, the rule will initially match. If the user later removes the sensitive data so the document no longer meets the threshold, the rule is undone. Similarly, if an administrator updates the policy to require fifteen Social Security numbers, the same document will no longer meet the criteria upon re-evaluation, resulting in a DLP Rule Undo event.

Contextual summary

You can see the text that surrounds the matched content, like a credit card number in a **DLPRuleMatch** event in Activity explorer.

DLPRuleMatch events are paired with user egress activities such as "CopyToClipboard" or "CloudEgress". They should be right next to (or at least very close to) each other in Activity explorer. You want to look at both because the **user activity** contains details about the matched policy and the **DLPRuleMatch** event contains the details about the text that surrounds the matched content.

For endpoint, be sure that you have applied KB5016688 for Windows 10 devices and KB5016691 for Windows 11 devices or above.

For more information, see Get started with activity explorer.

To learn more about Microsoft Purview DLP, see:

- Learn about Endpoint data loss prevention
- Learn about the default data loss prevention policy in Microsoft Teams (preview)
- Learn about data loss prevention on-premises scanner
- Learn about the Microsoft Compliance Extension
- Get started with the data loss prevention Alerts dashboard

To learn how to use data loss prevention to comply with data privacy regulations, see Deploy information protection for data privacy regulations with Microsoft Purview (aka.ms/m365dataprivacy).

Last updated on 06/26/2026

Plan for data loss prevention (DLP)

Every organization plans for and implements data loss prevention (DLP) differently. Why? Because every organization's business needs, goals, resources, and situation are unique. However, there are elements that are common to all successful DLP implementations. This article presents the best practices for planning a DLP deployment.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Before you begin

1. Administrative units
2. Learn about Microsoft Purview Data Loss Prevention
3. Plan for data loss prevention (DLP) - by working through this article you will:
   a. Identify stakeholders
   b. Describe the categories of sensitive information to protect
   c. Set goals and strategy
4. Collection Policies solution overview
5. Collection policy reference
6. Data Loss Prevention policy reference - this article introduces all the components of a DLP policy and how each one influences the behavior of a policy
7. Design a DLP policy - this article walks you through creating a policy intent statement and mapping it to a specific policy configuration.
8. Create and Deploy data loss prevention policies - This article presents some common policy intent scenarios that you map to configuration options, then it walks you through configuring those options.
9. Learn about investigating data loss prevention alerts - This article introduces you to the lifecycle of alerts from creation, through final remediation and policy tuning. It also introduces you to the tools you use to investigate alerts.

Multiple starting points

Many organizations choose to implement DLP to comply with various governmental or industry regulations. For example, the European Union's General Data Protection Regulation (GDPR), or the Health Insurance Portability and Accountability Act (HIPAA), or the California Consumer Privacy Act (CCPA). They also implement data loss prevention to protect their intellectual property. However, the starting place and ultimate destination in the DLP journey vary.

Organizations can start their DLP journey from several different points:

- With a platform focus, like wanting to protect information in Teams Chat and Channel messages or on Windows 10 or 11 devices
- Knowing the sensitive information they want to prioritize protecting, such as health care records, and going straight to defining policies to protect it
- Without knowing what their sensitive information is, where it is, or who is doing what with it; so, they start with discovery and categorization and take a more methodical approach
- Without knowing what their sensitive information is, where it is, or who is doing what with it, they go straight to defining policies and then using those outcomes to refine them
- Knowing that they need to implement the full Microsoft Purview Information Protection stack, and a plan to take a longer term, methodical approach

These are just some examples of how customers can approach DLP. It doesn't matter where you start from; DLP is flexible enough to accommodate various types of information protection journeys from start to a fully realized data loss prevention strategy.

Overview of planning process

The Learn about Microsoft Purview Data Loss Prevention introduces the three different aspects of the DLP planning process. We go into more detail here on the elements that are common to all DLP plans.

Identify stakeholders

When implemented, DLP policies can be applied across large portions of your organization. Your IT department can't develop a broad ranging plan on their own without negative consequences. You need to identify the stakeholders who can:

- identify the regulations, laws, and industry standards your organization is subject to
- identify the categories of sensitive items to be protected
- identify the business processes they're used in
- identify the risky behavior that should be limited

- prioritize which data should be protected first, based on the sensitivity of the items and risk involved
- outline the DLP policy match event review and remediation process

In general, these needs tend to be 85% regulatory and compliance protection, and 15% intellectual property protection. Here are some suggestions on roles to include in your planning process:

- Regulatory and compliance officers
- Chief risk officer
- Legal officers
- Security and compliance officers
- Business owners for the data items
- Business users
- IT

Describe the categories of sensitive information to protect

Once identified, the stakeholders then describe the categories of sensitive information to be protected and the business processes that they're used in. For example, DLP defines these categories:

- Financial
- Medical and health information
- Privacy
- Custom

Stakeholders might identify the sensitive information as "We are a data processor, so we have to implement privacy protections on data subject information and financial information".

Set goals and strategy

Once you have identified your stakeholders, know which sensitive information needs protection, and where it's used, the stakeholders can set their protection goals and IT can develop an implementation plan.

Set implementation plan

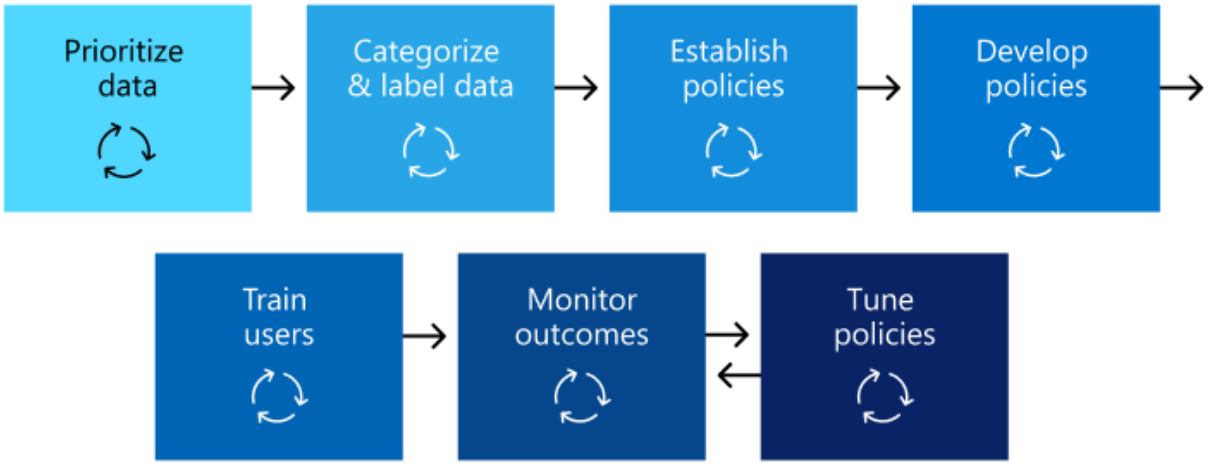
Your implementation plan should include:

- a map of your starting state, desired end state, and the steps to get from one to the other

- a plan for how you'll address discovery of sensitive items
- a plan for developing policies and the order in which policies you'll implement them
- a plan for how you'll address any prerequisites
- a plan for how you'll simulate policies before implementing them for enforcement
- a plan for how you'll train your end users
- a plan for how you'll tune your policies
- a plan for how you'll review and update your data loss prevention strategy based on changing regulatory, legal, industry standard, or intellectual property protection and business needs

Map out path from start to desired end state

Documenting how your organization is going to get from its starting state to the desired end state is essential to communicating with your stakeholders and setting the project scope. Here's a set of steps that are commonly used to deploy DLP. You'll want more detail than the graphic shows, but you can use it to frame your DLP adoption path.



Sensitive item discovery

There are multiple ways to discover what individual sensitive items are and where they're located. You may have sensitivity labels already deployed or you may have decided to deploy a broad DLP policy to all locations that only discovers and audits items. To learn more, see Know your data.

Policy planning

As you begin your DLP adoption, you can use these questions to focus your policy design and implementation efforts.

What laws, regulations, and industry standards must your organization comply with?

Because many organizations come to DLP with the goal of regulatory compliance, answering this question is a natural starting place for planning your DLP implementation. But, as the IT implementer, you're probably not positioned to answer it. Instead, you should consult your legal team and business executives for an answer.

Example Your organization is subject to U.K. financial regulations.

What sensitive items must your organization protect from leakage?

Once you know where your organization stands in terms of regulatory compliance needs, you'll have some idea of what sensitive items need to be protected from leakage. You'll also have an idea about how you want to prioritize policy implementation to protect those items. This knowledge helps you choose the most appropriate DLP policy templates. Microsoft Purview comes with preconfigured DLP templates for Financial, Medical and health, and Privacy policy templates. Additionally, you can build your own using the Custom template. As you design and create your actual DLP policies, having a working knowledge of the sensitive items you need to protect helps you choose the right sensitive information type.

Example To get started quickly, you might pick the preconfigured `U.K. Financial Data` policy template, which includes the `Credit Card Number`, `EU Debit Card Number`, and `SWIFT Code` sensitive information types.

How you want your policies scoped

If your organization has implemented administrative units, you can scope your DLP policies by administrative unit or leave the scope default, which applies policies to the full directory. For more information, see Policy Scoping.

Where are the sensitive items and what business processes are they involved in?

The items that contain your organization's sensitive information are used every day in the course of doing business. You need to know where instances of that sensitive information may occur and what business processes they're used in. Knowing this helps you choose the right locations to apply your DLP policies. DLP policies can be applied to the following locations:

- Exchange online email
- SharePoint sites

- OneDrive accounts
- Teams chat and channel messages
- Windows 10, 11 and macOS Devices
- Microsoft Defender for Cloud Apps
- On-premises repositories

Example Your organization's internal auditors are tracking a set of credit card numbers. They keep a spreadsheet of them in a secure SharePoint site. Several of the employees make copies and save them to their work OneDrive site, which is synced to their Windows 10 device. One of these employees pastes a list of 14 credit card numbers into an email and tries to send it to the outside auditors for review. In this case, you'd want to apply the policy to the secure SharePoint site, all the internal auditors OneDrive accounts, their Windows 10 devices, and Exchange email.

What is your organization's tolerance for leakage?

Different groups in your organization may have different views on what counts as an acceptable level of sensitive item leakage. Achieving the perfection of zero leakage may come at too high a cost to the business.

Example Your organization's security group and legal team both feel that there should be no sharing of credit card numbers with anyone outside the org. They insist on zero leakage. However, as part of their regular review of credit card number activity, the internal auditors must share some credit card numbers with third-party auditors. If your DLP policy prohibits all sharing of credit card numbers outside the org, there will be a significant business process disruption and added cost to mitigate the disruption in order for the internal auditors to complete their tracking. This extra cost is unacceptable to the executive leadership. To resolve this issue, there needs to be an internal conversation to decide an acceptable level of leakage. Once that is decided, the policy can provide exceptions for certain individuals to share the information, or, it can be applied in audit-only mode.

> **Important**
>
> To learn how to create a policy intent statement and map it to policy configurations see, Design a data loss prevention policy

Planning for prerequisites

Before you can monitor some DLP locations, there are prerequisites that must be met. See the **Before you begin** sections of the following articles:

- Get started with the data loss prevention on-premises scanner
- Get started with Endpoint data loss prevention
- Get started with the Microsoft compliance extension
- Use data loss prevention policies for non-Microsoft cloud apps

Policy deployment

When you create your DLP policies, you should consider rolling them out gradually, so you can assess their impact and test their effectiveness before fully enforcing them. For example, you don't want a new DLP policy to unintentionally block access to thousands of documents or to break an existing business process.

If you're creating DLP policies with a large potential impact, we recommend following this sequence:

1. **Run the policy in simulation mode, without Policy Tips** and then use the DLP reports, and any incident reports, to assess the impact of the policies. You can use DLP reports to view the number, location, type, and severity of policy matches. Based on the results, you can fine-tune your policies as needed. In simulation mode, DLP policies won't impact the productivity of people working in your organization. It's also good to use this stage for testing out your workflow for DLP event review and issue remediation.

2. **Run the policy in simulation mode with notifications and Policy Tips** so that you can begin to teach users about your compliance policies and prepare them for when the policies are applied. It's useful to have a link to an organization policy page that provides more details about the policy in the Policy Tip. At this stage, you can also ask users to report false positives, so that you can further refine the conditions and reduce the number of false positives. Move to this stage once you have confidence that the results of applying the policies match what they stakeholders had in mind.

3. **Start full policy enforcement** so that the actions in the rules are applied and the content is protected. Continue to monitor the DLP reports and any incident reports or notifications to make sure that the results are what you intend.

End-user training

You can configure your policies so that, when a DLP policy is triggered, email notifications are sent automatically, and policy tips are shown to admins and end users. Policy tips are useful ways to raise awareness of risky behaviors on sensitive items and train users to avoid those behaviors in the future.

Review DLP requirements and update strategy

The regulations, laws, and industry standards that your organization is subject to will change over time, as will your business goals for DLP. Be sure to include regular reviews of all these areas so that your organization stays in compliance and so your DLP implementation continues to meet your business needs.

Approaches to deployment

Expand table

Customer business needs description	Approach
Contoso Bank is in a highly regulated industry and has many different types of sensitive items in many different locations. Contoso: - knows which types of sensitive information are top priority - must minimize business disruption as policies are rolled out - has involved business process owners - has IT resources and can hire experts to help plan, design, and deploy - has a premier support contract with Microsoft	- Take time to understand what regulations they must comply with and how they're going to comply. - Take time to understand the "better together" value of the Microsoft Purview Information Protection stack - Develop a sensitivity labeling scheme for prioritized items and apply it - Design and code policies, deploy them in simulation mode, and train users - Repeat and refine policies
TailSpin Toys doesn't know what sensitive data they have or where it is, and they have little to no resource depth. They use Teams, OneDrive, and Exchange extensively.	- Start with simple policies on the prioritized locations. - Monitor what gets identified - Apply sensitivity labels accordingly - Refine policies and train users
Fabrikam is a small startup. They want to protect their intellectual property and must move quickly. They're willing to dedicate some resources, but can't afford to hire outside experts. Other considerations: - Sensitive items are all in Microsoft 365 OneDrive / SharePoint - Adoption of OneDrive and SharePoint is slow. Many employees still use DropBox and Google drive to store and share items - Employees value speed of work over data protection discipline - All 18 employees have new Windows devices	- Take advantage of the default DLP policy in Teams - Use the "restricted by default" setting for SharePoint items - Deploy policies that prevent external sharing - Deploy policies to prioritized locations - Deploy policies to Windows devices - Block uploads to cloud storage solutions other than OneDrive

Next steps

> **Important**

To learn more about DLP policy deployment, see **Deployment**

See also

- Learn about data loss prevention

Last updated on 06/26/2026

Design a data loss prevention policy

Taking the time to design a policy before you implement it gets you to the desired results faster, with fewer unintended issues, than creating it and then tuning by trial and error alone. Having your policy designs documented will help you in communications, policy reviews, troubleshooting, and further tuning.

If you're new to Microsoft Purview DLP, it's helpful to work through these articles before you start designing a policy:

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about **Microsoft Security Copilot in Microsoft Purview**.

Before you begin

1. Administrative units
2. Learn about Microsoft Purview Data Loss Prevention
3. Plan for data loss prevention (DLP) - by working through this article you will:
   a. Identify stakeholders
   b. Describe the categories of sensitive information to protect
   c. Set goals and strategy
4. Collection Policies solution overview
5. Collection policy reference
6. Data Loss Prevention policy reference - this article introduces all the components of a DLP policy and how each one influences the behavior of a policy
7. Design a DLP policy - this article walks you through creating a policy intent statement and mapping it to a specific policy configuration.
8. Create and Deploy data loss prevention policies - This article presents some common policy intent scenarios that you map to configuration options, then it walks you through configuring those options.
9. Learn about investigating data loss prevention alerts - This article introduces you to the lifecycle of alerts from creation, through final remediation and policy tuning. It also introduces you to the tools you use to investigate alerts.

Policy design overview

Designing a policy is mostly about clearly defining your business needs, documenting them in a policy intent statement, and then mapping those needs to policy configuration. You use the decisions you made in your planning phase to inform some of your policy design decisions.

Define intent for the policy

You should be able to summarize, in a single statement, the business intent for every policy you have. Developing this statement drives conversations in your organization and, when fully fleshed out, this statement directly links the policy to a business purpose and provides a roadmap for policy design. The steps in the Plan for data loss prevention (DLP) article helps you get started on your policy intent statement.

Remember that, as described in the DLP policy configuration overview, all DLP policies require you to:

- Choose what you want to monitor
- Choose the Policy Scoping.
- Choose where you want to monitor .
- Choose the conditions that must be matched for a policy to be applied to an item.
- Choose the action to take when the policy conditions are met.

For example, here's a fictitious first draft of an intent statement that provides answers to all five questions:

"We're a U.S. based organization, and we need to detect Office documents that contain sensitive health care information covered by HIPAA that are stored in OneDrive/SharePoint and to protect against that information being shared in Teams chat and channel messages and restrict everyone from sharing them with unauthorized third parties".

As you develop a policy design, you'll likely modify and extend the statement.

Map business needs to policy configuration

Let's break down the example draft statement and map it to DLP policy configuration points. This example assumes that you're using an unrestricted DLP admin account and that administrative units aren't configured.

> **Important**

Be sure you understand the difference between an *unrestricted administrator* and an *administrative unit restricted administrator* by reading Administrative units before you start.

Expand table

Statement	Configuration question answered and configuration mapping
"We're a U.S. based organization, and we need to detect Office documents that contain sensitive health care information covered by HIPAA...	- **What to monitor**: Office docs, use the U.S. Health Insurance Act (HIPAA) template - **Conditions for a match**: (preconfigured but editable) - item contains U.S. SSN and Drug Enforcement Agency (DEA) number, International Classification of Diseases (ICD-9-CM), International Classification of Diseases (ICD-10-CM), content is shared with people outside my organization - drives conversations to clarify the triggering threshold for detection like confidence levels, and instance count (called leakage tolerance).
...that are stored in OneDrive/SharePoint and protect against that information being shared in Teams chat and channel messages...	- **Where to monitor**: Location scoping by including or excluding OneDrive and SharePoint sites and Teams chat/channel accounts or distribution groups. **Policy scoping** (preview): Full directory
...and restrict everyone from sharing those items with unauthorized third parties."	- **Actions to take**: You add *Restrict access or encrypt the content in Microsoft 365 locations* - drives conversation on what actions to take when a policy is triggered including protective actions like sharing restrictions, awareness actions like notifications and alerts, and user empowerment actions like allow user overrides of a blocking action

This example doesn't cover all the configuration points of a DLP policy; it would need to be expanded. However, it should get you thinking in the right direction as you develop your own DLP policy intent statements.

Important

Keep in mind that the location(s) you pick impact whether you can use sensitive information types, sensitivity labels, and retention labels. The location(s) you pick also impact the actions that are available. See Data Loss Prevention policy reference for more information.

Complex rule design

The above HIPAA content in SharePoint and OneDrive is a simple example of a DLP policy. The DLP rule builder supports boolean logic (AND, OR, NOT) and nested groups.

> **Important**
>
> - All existing **Exceptions** are replaced with a NOT condition in a nested group inside of the **Conditions**.
> - You need to create groups in order to use multiple operators.

> **Important**
>
> When an action in Office desktop client apps, (Word, Outlook, Excel, and PowerPoint) matches a policy that uses complex conditions, the user will only see policy tips for rules that use the **Content contains sensitive information** condition.

- *Example 1 We need to block emails to all recipients that contain credit card numbers, OR that have the 'highly confidential' sensitivity label applied, but do NOT block the email if it is sent from someone on the finance team to adele.vance@contoso.com*
- *Example 2 Contoso needs to block all emails that contain a password protected file OR a zip document file extension ('zip' or '7z'), but do NOT block the email if the recipient is in the contoso.com domain OR the fabrikam.com domain, OR the sender is a member of the Contoso HR group.*

https://learn-video.azurefd.net/vod/player?id=0edf38a9-89fb-4bc5-95f4-48c70053e031&locale=en-us&embedUrl=%2Fpurview%2Fdlp-policy-design

> **Important**
>
> - Using the NOT condition in a nested group replaces the **Exceptions** functionality.
> - You need to create groups in order to use multiple operators.

> **Important**
>
> When an action in Office desktop client apps (Word, Outlook, Excel, and PowerPoint) matches a policy that uses complex conditions, the user will only see policy tips for rules that use the **Content contains sensitive information** condition.

Policy Design Process

1. Complete the steps in Plan for data loss prevention (DLP) - by working through this article you will:
   a. Identify your stakeholders
   b. Describe the categories of sensitive information to protect
   c. Set goals and strategy
   d. Define your policy deployment plan

2. Familiarize yourself with Data Loss Prevention policy reference so that you understand all the components of a DLP policy and how each one influences the behavior of a policy.

3. Familiarize yourself with what the DLP policy templates include.

4. Develop your policy intent statement with your key stakeholders. Refer to the example earlier in this article.

5. Determine how this policy fits into your overall DLP policy strategy.

6. Map the items in your policy intent statement to configuration options.

7. Decide which policy template you start from: predefined or custom.

8. Go through the template and assemble all information required before you create the policy. It's likely that you will find that there are some configuration points that aren't covered in your policy intent statement. That's ok. Go back to your stakeholders to iron out the requirements for any missing configuration points.

9. Document the configuration of all the policy settings and review them with your stakeholders. You can reuse your policy intent statement mapping to configuration points, which is now fully fleshed out.

10. Create a draft policy and refer back to your policy deployment plan.

See Also

- Learn about data loss prevention
- Plan for data loss prevention (DLP)
- Data Loss Prevention policy reference
- Data Loss Prevention policy tips reference
- Create and Deploy data loss prevention policies

Last updated on 04/08/2025

Create and deploy data loss prevention policies

Microsoft Purview Data Loss Prevention (DLP) policies include many configuration options. Each option changes the policy's behavior. The articles in this series cover some of the most common DLP policy scenarios. They walk you through configuring those options to give you hands-on experience with the DLP policy creation process. When you familiarize yourself with these scenarios, you gain the foundational skills that you need to use the DLP policy creation UX to create your own policies.

How you deploy a policy is as important policy design. You have multiple options to control policy deployment, including policy state, actions, and scope, as described in the Deploy and manage DLP policies section of this article. This article shows you how to use these options so that the policy achieves your intent while avoiding costly business disruptions.

In preview You can change the display name of DLP policies and rules. Once you rename a policy or a rule, any existing records retain their previous name in activity explorer evetns, in alerts and in audit records. New records will reflect the new name in activity explorer events, in alerts and in audit records. These names will remain until the items age out of the system.

Orient yourself to DLP

1. Administrative units
2. Learn about Microsoft Purview Data Loss Prevention
3. Plan for data loss prevention (DLP) - by working through this article you will:
    a. Identify stakeholders
    b. Describe the categories of sensitive information to protect
    c. Set goals and strategy
4. Collection Policies solution overview
5. Collection policy reference
6. Data Loss Prevention policy reference - this article introduces all the components of a DLP policy and how each one influences the behavior of a policy
7. Design a DLP policy - this article walks you through creating a policy intent statement and mapping it to a specific policy configuration.
8. Create and Deploy data loss prevention policies - This article presents some common policy intent scenarios that you map to configuration options, then it walks you through

configuring those options.

9. Learn about investigating data loss prevention alerts - This article introduces you to the lifecycle of alerts from creation, through final remediation and policy tuning. It also introduces you to the tools you use to investigate alerts.

SKU/subscriptions licensing

For information on licensing, see

- Microsoft 365 Enterprise Plans
- Microsoft 365 Service Descriptions

Permissions

The account you use to create and deploy policies must be a member of one of these role groups:

- Compliance administrator
- Compliance data administrator
- Information Protection
- Information Protection Admin
- Security administrator

> **Important**
>
> Before you start, make sure you understand the difference between an *unrestricted administrator* and an *administrative unit restricted administrator* by reading Administrative units.

Granular Roles and Role Groups

You can use these roles and role groups to fine tune your access controls.

Here's a list of applicable roles. To learn more, see Permissions in the Microsoft Purview portal.

- DLP Compliance Management
- Information Protection Admin
- Information Protection Analyst
- Information Protection Investigator
- Information Protection Reader

Here's a list of applicable role groups. To learn more, see Permissions in the Microsoft Purview portal.

- Information Protection
- Information Protection Admins
- Information Protection Analysts
- Information Protection Investigators
- Information Protection Readers

Policy creation scenarios for Enterprise applications & devices

The previous article, Design a DLP policy, introduces you to the methodology of creating a policy intent statement and then mapping that intent statement to policy configuration options.

- Help prevent sharing credit card numbers through email
- Help prevent sharing sensitive items via SharePoint and OneDrive with external users
- Help protect files that Endpoint Data Loss Prevention fails to scan
- Help protect files that Endpoint Data Loss Prevention doesn't scan
- Help protect against sharing of a defined set of unsupported files
- Disable Microsoft Purview data loss prevention scanning for some supported files and apply controls
- Help prevent sharing Power BI reports with credit card numbers

Policy creation scenarios for Inline web traffic

The following scenarios show how to create DLP policies for inline web traffic.

- Help prevent sharing via Microsoft Edge for Business to unmanaged AI apps from managed devices
- Help Prevent Users from Sharing Sensitive Info with Cloud Apps in Edge for Business
- Help prevent sharing sensitive information with unmanaged AI apps via network data security

Deploy and manage DLP policies

A successful policy deployment isn't just about getting the policy into your environment to enforce controls on user actions. A haphazard, rushed deployment can negatively impact business processes and annoy your users. Those consequences slow acceptance of DLP

technology in your organization and the safer behaviors it promotes. Ultimately, those consequences make your sensitive items less safe in the long run.

Before you start your deployment, make sure you read through Policy deployment. It gives you a broad overview of the policy deployment process and general guidance.

The following guidance covers the three types of deployment controls—state, actions, and scope—that you use in concert to manage your policies in production. You can change any of these controls at any time, not just during policy creation.

Three axes of deployment management

Use three axes to control the policy deployment process: scope, state, and actions. Always take an incremental approach to deploying a policy, starting from the least impactful DLP simulation mode through full enforcement.

Recommended deployment control configurations

The following table summarizes recommended combinations of policy state, scope, and action impact during deployment.

Expand table

When your policy state is	Your policy scope can be	Impact of policy actions
Run the policy in simulation mode	Policy scope of locations can be narrow or broad	- You can configure any action - No user impact from configured actions - Admin sees alerts and can track activities
Run the policy in simulation mode with policy tips	Policy should be scoped to target a pilot group and then expand the scope as you tune the policy	- You can configure any action - No user impact from configured actions - Users can receive policy tips and alerts - Admin sees alerts and can track activities
Turn it on	All targeted location instances	- All configured actions are enforced on user activities - Admin sees alerts and can track activities

When your policy state is	Your policy scope can be	Impact of policy actions
Keep it off	n/a	n/a

Policy state

State is the primary control you use to roll out a policy. When you finish creating your policy, set the state of the policy to **Keep it off**. Leave it in this state while you work on the policy configuration and until you get a final review and sign off. Set the state to:

- **Run the policy in simulation mode**: No policy actions are enforced, events are audited. While in this state, you can monitor the impact of the policy in the DLP simulation mode overview and the DLP **Activity explorer** console.
- **Run the policy in simulation mode and show policy tips while in simulation mode**: No actions are enforced, but users receive policy tips and notification emails to raise their awareness and educate them.
- **Turn it on right away**: This is full enforcement mode.
- **Keep it off**: The policy is inactive. Use this state while developing and reviewing your policy before deployment.

You can change the state of a policy at any time.

Policy actions

Actions are what a policy does in response to user activities on sensitive items. Because you can change these actions at any time, you can start with the least impactful, *Allow* (for devices) and **Audit only** (for all other locations), gather and review the audit data, and use it to tune the policy before moving to more restrictive actions.

- **Allow**: The user activity is allowed to occur, so no business processes are impacted. You get audit data and there aren't any user notifications or alerts.

  > ⓘ **Note**
  >
  > The **Allow** action is only available for policies that are scoped to the **Devices** location.

- **Audit only**: The user activity is allowed to occur, so no business processes are impacted. You get audit data and you can add notifications and alerts to raise awareness and train your users to know that what they're doing is a risky behavior. If your organization intends to enforce more restrictive actions later on, you can tell your users that too.

- **Block with override**: The user activity is blocked by default. You can audit the event, raise alerts and notifications. This action impacts the business process, but your users are given the option to override the block and provide a reason for the override. Because you get direct feedback from your users, this action can help you identify false positive matches, which you can use to further tune the policy.

  > ⓘ **Note**
  >
  > For Exchange online and SharePoint in Microsoft 365, you configure overrides in the user notification section.

- **Block**: The user activity is blocked no matter what. You can audit the event, raise alerts and notifications.

Define policy scope

Every policy is scoped to one or more locations, such as Exchange, SharePoint in Microsoft 365, Teams, and Devices. By default, when you select a location, all instances of that location fall under the scope and none are excluded. You can further refine which instances of the location (such as sites, groups, accounts, distribution groups, mailboxes, and devices) that the policy is applied to by configuring the include/exclude options for the location. To learn more about your include/exclude scoping options, see DLP policy locations.

In general, you have more flexibility with scoping while the policy is in **Run the policy in simulation mode** state because no actions are taken. You can start with just the scope you designed the policy for or go broad to see how the policy would impact sensitive items in other locations.

When you change the state to **Run the policy in simulation mode and show policy tips**, narrow your scope to a pilot group that can give you feedback and be early adopters who can be a resource for others when they come onboard.

When you move the policy to **Turn it on right away**, you broaden the scope to include all the instances of locations that you intended when the policy was designed.

Policy deployment steps

Use the following steps to roll out a DLP policy safely from initial review through full enforcement.

1. After you've created the policy and set its state to **Keep it off**, do a final review with your stakeholders.
2. Change the state to **Run the policy in simulation mode**. The location scope can be broad at this point so you can gather data on the behavior of the policy across multiple locations or just start with a single location.
3. Tune the policy based on the behavior data so that it better meets the business intent.
4. Change the state to **Run the policy in simulation mode and show policy tips**. Refine the scope of locations to support a pilot group if needed and make use of includes/excludes so that the policy is first rolled out to that pilot group.
5. Gather user feedback and alert and event data. If needed, tune the policy and your plans. Make sure you address all the issues that your users bring up. Your users will most likely encounter issues and raise questions about things that you didn't think of during the design phase. Develop a group of super users at this point. They can be a resource to help train other users as the scope of the policy is increased and more users come onboard. Before you go to the next stage of deployment, make sure that the policy is achieved your control objectives.
6. Change the state to **Turn it on right away**. The policy is fully deployed. Monitor DLP alerts and DLP activity explorer. Address alerts.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Test your Data Loss Prevention policies

You should test and tune the behavior of your Microsoft Purview Data Loss Prevention (DLP) policies as part of your DLP policy deployment. This article introduces you to two of the basic methods you can use to test policies in your DLP environment.

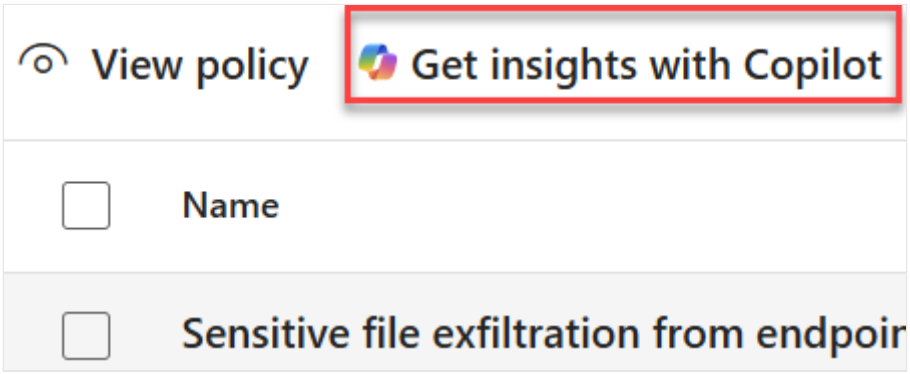
Simulation mode

When you deploy a new policy or need to modify an existing one, you should run it in simulation mode, and then review the alerts to assess its accuracy. Simulation mode allows you to see how an individual policy affects all the items that are in the policy's scope without actual enforcement. You use it to find out what items match a policy.

Get insights with Microsoft Security Copilot

> **Note**
>
> The **Get insights with Copilot** feature is in preview.

Get insights with Copilot is a Security Copilot skill that is embedded in the Microsoft Purview Data Loss Prevention policy page. It can help you understand what you policies are doing in your organization, and where they're active.



Select one or multiple policies, then select the **Get insights with Copilot** button. Copilot then generates a response that gives you information about the selected policies like:

- Where your policies are looking for sensitive information.
- What kind of sensitive information your policies are looking for.
- What scenarios trigger the policies.

- How end users are impacted by the policies.
- How the admins are notified.

You can further pivot your investigation by administrative units, locations DLP protects and, types of data that's been classified

Prerequisites

- Your organization must be licensed for Microsoft Security Copilot.
- The account you use to access the **Get insights with Copilot** feature must be in the Information protection admin role.

Use the Test-DlpPolicies cmdlet

Test-DlpPolicies is a cmdlet that allows you to see which DLP policies scoped to SharePoint and OneDrive match (or don't match) an individual item in SharePoint or OneDrive.

Before you begin

- You must be able to connect to Exchange Online PowerShell.
- You must have a valid SMTP address to send the report to. For example: `dlp_admin@contoso.com`
- You must have the site ID where the item is located.
- You must have the direct link path to the item.

> **Important**
>
> - Test-DlpPolicies only works for items that are in SharePoint or OneDrive.
> - Test-DlpPolices only reports results for policies that include SharePoint alone, OneDrive alone, or both SharePoint and OneDrive in their scope.
> - Test-DlpPolices works only with simple conditions. It doesn't work with complex, grouped, or nested conditions.

Use Test-DlpPolicies

To see which DLP policies an item matches, follow these steps:

Get the direct link path to the item

To get the direct link path for the item, follow these steps:

1. Open the SharePoint or OneDrive folder in a browser.
2. Select the file's ellipsis and select **details**.
3. In the details pane, scroll down and select **Path**. Copy the direct link and save it.

   For example:

`https://contoso.sharepoint.com/personal/user_contoso_com/Documents/test.docx`

Get the site ID

To get the site ID, follow these steps:

1. Connect to Exchange Online PowerShell.
2. For SharePoint, use the following syntax to get the site ID and save it:

PowerShell

```
$reportAddress = "email@contoso.com"

$siteName = "SITENAME@TENANT.onmicrosoft.com"

$filePath =
"https://Contoso.sharepoint.com/sites/SOMESITENAME/Shared%20Documents/TESTFILE.ppt
x"

$r = Get-Mailbox -Identity $siteName -GroupMailbox

$e = $r.EmailAddresses | Where-Object {$_ -like '*SPO*'}

Test-DlpPolicies -SiteId $e.Substring(8,36) -FileUrl $filePath -Workload SPO -
SendReportTo $reportAddress
```

3. For OneDrive, use the following syntax to get the site ID and save it.

PowerShell

```
$reportAddress = "email@contoso.com"

$odbUser = "USER@TENANT.onmicrosoft.com"

$filePath = "https://contoso-
```

```
my.sharepoint.com/personal/userid_contoso_onmicrosoft_com/Documents/TESTFILE.docx
"

$r = Get-Mailbox -Identity $odbUser

$e = $r.EmailAddresses | Where-Object {$_ -like '*SPO*'}

Test-DlpPolicies -SiteId $e.Substring(8,36) -FileUrl $filePath -Workload ODB -
SendReportTo $reportAddress
```

Here's an example of a returned value:

`36ca70ab-6f38-7f3c-515f-a71e59ca6276`

Run Test-DlpPolicies

After you have the site ID and the direct link path, run the cmdlet to test your policies.

- Run the following syntax in the PowerShell window:

PowerShell

```
Test-DlpPolicies -workload <workload> -Fileurl <path/direct link> -SendReportTo
<smtpaddress>
```

For example:

`Test-DlpPolicies -workload <ODB> -Fileurl <https://contoso.sharepoint.com/personal/user_contoso_com/Documents/test.docx> -SendReportTo <dlp_admin@contoso.com>`

Interpret the report

The report is sent to the SMTP address specified in the `-SendReportTo` parameter of the `Test-DlpPolicies` command. There are multiple fields. Here are explanations of the most important ones.

Expand table

Field name	Means
Classification ID	The sensitive information type (SIT) the item is categorized as
Confidence	The confidence level of the SIT

Field name	Means
Count	The total number of times the SIT value was found in the item, this includes duplicates
Unique Count	The number SIT values found in the item with duplicates eliminated
Policy Details	The name and GUID of the policy that was evaluated
Rules - Rule Details	The DLP rule name and GUID
Rules - Predicates - Name	The condition defined in the DLP rule
Rules - Predicates - IsMatch	Whether the item matched the conditions
Predicates - Past Actions	Any actions, like notify user, block, block with override that 's been taken on the item
Predicates - Rule's Actions	The action defined in the DLP rule
Predicates - IsMatched	Whether the item matched the rule
IsMatched	Whether the item matched the overall policy

Related content

- Test-DataClassification explains how to use the PowerShell cmdlet `Test-DataClassification`.
- Test-Message explains how to use the PowerShell cmdlet `Test-Message`.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Learn about using regular expressions (regex) in data loss prevention policies

A *regular expression*, commonly referred to as a *regex*, is a sequence of characters that defines a search pattern. Regular expressions are primarily used for pattern matching with strings and in string matching; for example, in "find and replace" operations. You can use a regex in Microsoft Purview Data Loss Prevention (DLP) to define patterns that help you identify and classify sensitive data, or to help detect patterns in content. The most common regex uses in Microsoft Purview DLP are:

- Defining a custom sensitive information types.
- Leveraging the `SubjectOrBodyMatchesPatterns` condition in a DLP rule (Read more here.)

This article describes common issues that occur when working with regular expressions and how you can resolve them.

Potential validation issue when using a regex with DLP

- Basic units of the pattern, such as literal characters, digits, whitespace, and punctuation marks can be represented by themselves or by special symbols called metacharacters, such as `\d` for any digit, `\s` for any whitespace, or `\.` for a literal dot.
- The basic units, when combined with quantifiers, specify how many times they can or must occur in a match. For example, `*` means zero or more, `+` means one or more, `?` means zero or one, and `{n,m}` means between `n` and `m` times. For example, `\d+` means one or more digits, `\s?` means optional whitespace, and `a{3,5}` means between three and five instances of the literal character **a**.
- A regex either uses a positive lookbehind or a negative lookbehind. A *lookbehind* is used to check whether there's a match before a certain position in the input string, without including the actual characters in the match. A positive lookbehind is used to match when the lookbehind pattern **is** present, while a negative lookbehind is used to match when the lookbehind pattern **is not** present.
- Consider this example: `(?<=^|\s|_)`. This example shows a lookbehind that includes three possibilities:

    1. `^` asserts the position. In this case, it requires the pattern matching to begin at the start of the line.

2. `\s` detects any whitespace characters as a match.
3. `_` matches the literal underscore character (_).

- In the previous example, possibilities #2 and #3 will each match a single character. However, possibility #1 only indicates where the matching should start. It will not produce results with respect to any character matches.
- Take a second example, `^\d+$`. This regex will only detect a string composed entirely of digits, from start to end.
- Regex pattern validations in conditions are case insensitive by default. To honor the case sensitivity please add (?-i) to the pattern

How to get extracted text

A regex is matched on the extracted text of the content, rather than on the content itself. So, even when the pattern appears to be on the content, it might not match when evaluating a DLP policy.

To ensure that you capture the appropriate matches, take the following steps:

1. Use the Test-TextExtraction cmdlet to get the extracted text, which will consist of a stream of strings.
2. Next, use the extracted text for matching the regular expression.

For instance:

PowerShell

```
$data = ([System.IO.File]::ReadAllBytes('<FilePath>'))
$tr = Test-TextExtraction -FileData $data
$tr.ExtractedResults.ExtractedStreamText | Format-List
```

How to verify sensitive information type detection

To verify sensitive information type (SIT) detection, we need to take the text we just extracted and then run the Test-DataClassification cmdlet on it to verify detection. The results of running the cmdlet will indicate whether there are any SIT matches for the regex.

For example:

PowerShell

```
$textStream = $tr.ExtractedResults.ExtractedStreamText | Out-String
$result = Test-DataClassification -TextToClassify $textStream
```

```
$result.ClassificationResults | Format-List
```

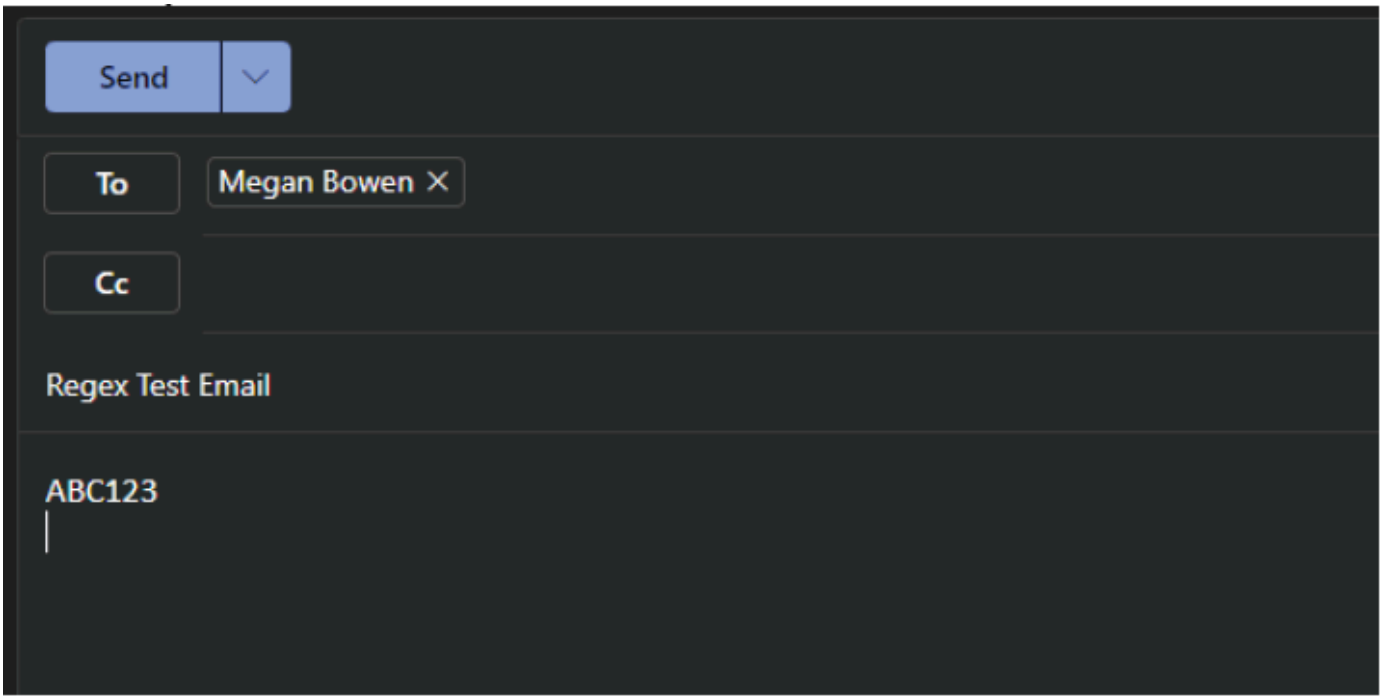
Example of using a regex in a DLP policy rule

In this example, we will block email that contains strings starting with *ABC* followed by a number.

Regex used: `^ABC\d`

DLP rule sample: `New-DlpComplianceRule -Name "Rule_00" -Policy "Policy_00" -SubjectOrBodyMatchesPatterns "^ABC\d" - BlockAccess $True`

Sample email



While it seems that the regex will detect a match with this mail item, the extracted text looks like this:

```
Regex Test Email ABC123
```

As you can see, the extracted text begins with the content in the subject line of the email, rather than with the content in the body of the email. However, the inclusion of the assertion character, `^`, at the start of the regex requires that the *ABC*... string must be at the start of the extracted text for a match to be detected.

To resolve this issue, you can change the regex to `ABC\d`.

Regex requirements for custom sensitive information types

When you define a regex for a custom sensitive information type, the pattern must follow a specific structure to work correctly with Microsoft Purview classification. This requirement applies whether you create the SIT through the Microsoft Purview portal or PowerShell.

The required structure is:

`(?:prefix)(primary capturing group)(?:suffix)`

Where:

- Use **exactly one** capturing group for the primary match. Use noncapturing groups `(?: ... )` for any prefix/suffix or other grouping.
- All match variants must be handled **inside** that single capturing group using `|` (alternation).
- Multiple capturing groups separated by `|` at the top level are **not supported** and are blocked during validation.

If your regex contains multiple top-level capturing groups, the SIT fails to classify content on Endpoint DLP and other workloads.

Example of an invalid pattern:

`(?:[\s,;])([A-Z]{2}[0-9]{3})(?:[\s,;])|(?:[\s,;])([A-Z]{2}[A-Z]{4}[0-9])(?:[\s,;])|(?:[\s,;])([A-Z]{2}[0-9]{5})(?:[\s,;])`

This pattern has three separate capturing groups at the top level. Instead, consolidate all variants into one capturing group:

Example of a valid pattern:

`(?:[\s,;])([A-Z]{2}[0-9]{3}|[A-Z]{2}[A-Z]{4}[0-9]|[A-Z]{2}[0-9]{5})(?:[\s,;])`

For more information on limits, see Sensitive information type limits.

Learn more about regular expressions

- Regular Expression Language - Quick Reference
- Regex101 (an interactive tool for testing and learning regular expressions)
- "Regular Expressions" from Mozilla Developer Network (a comprehensive guide to regular expressions)

- Boost Article on *Lookbehind*

Last updated on 06/26/2026

Learn about Adaptive Protection in Data Loss Prevention

Article • 03/31/2025

Adaptive Protection in Microsoft Purview integrates Microsoft Purview Insider Risk Management with Microsoft Purview Data Loss Prevention (DLP). When insider risk identifies a user who is engaging in risky behavior, they are dynamically assigned to a inside risk level. Then Adaptive Protection can automatically create a DLP policy to help protect the organization against the risky behavior that's associated with that inside risk level. As users insider risk levels change in insider risk management, the DLP policies applied to users can adjust.

You can manually create DLP policies that help protect against risky behaviors that insider risk identifies too.

Refer to Help dynamically mitigate risks with Adaptive Protection to learn about Adaptive Protection and how to configure it.

How Adaptive Protection shows up in DLP policies

If you're unfamiliar with DLP policies, you should review these articles before working with Adaptive Protection:

- Learn about data loss prevention
- Plan for data loss prevention (DLP)
- Data Loss Prevention policy reference
- Design a data loss prevention policy

Once Adaptive Protection is configured in insider risk, a condition called **User's risk level for Adaptive Protection is** will be available to use in rules that are configured for policies scoped to Exchange Online, Devices, and Teams locations.

The condition **Insider risk level for Adaptive Protection is** has three values:

- **Elevated risk level**
- **Moderate risk level**
- **Minor risk level**

These insider risk level profiles are defined in insider risk. You can select one, two or all three in a policy rule. Learn more about insider risk levels.

You can manually configure DLP policies that are part of Adaptive Protection and also use the quick setup configuration in insider risk to create DLP policies automatically from a template.

Manual configuration

You manually configure an Adaptive Protection DLP policy just like you would configure any other policy. Just select the **Insider risk level for Adaptive Protection is** condition and the insider risk level profiles that you want, configure all the other policy options and deploy the policy according to your normal procedures.

Quick setup configuration

If quick setup is used to configure Adaptive Protection in insider risk, DLP policies are created automatically, so you should be on the lookout for them. Quick setup will create one policy for Teams and Exchange Online with two rules, one for the elevated risk profile and one for the moderate and minor insider risk levels. It will also create one policy for Devices with two rules, one for the elevated risk profile and one for the moderate and minor insider risk levels.

> **Tip**
>
> Insider risk presents a view of just the DLP policies that use the **Insider risk level for Adaptive Protection is** condition. Open **Microsoft Purview portal** > **Insider risk management** > **Adaptive protection** to see the list. You'll need DLP to be in one of these roles to access the insider risk node:

- Compliance administrator
- Compliance Data administrator
- Organization management (Users who are not global administrators must be Exchange administrators to see and take action on devices that are managed by Basic Mobility and Security for Microsoft 365)
- Global administrator
- DLP compliance management
- View-only DLP compliance management

> **Important**

Microsoft recommends that you use roles with the fewest permissions. This helps improve security for your organization. Global Administrator is a highly privileged role that should only be used in scenarios where a lesser privileged role can't be used.

Policy values for Teams and Exchange online DLP policy

This is the configuration for the Teams and Exchange DLP policy created during Quick Setup. The policy name is **Adaptive Protection policy for Teams and Exchange DLP**.

Rule: Adaptive Protection block rule for Teams and Exchange DLP

Expand table

DLP policy element	Configured value
Conditions	**Insider risk level for Adaptive Protection is** **- Elevated Risk Level** AND **- Content is Shared from Microsoft 365 With people outside my organization**
Actions	**Restrict access or encrypt the content in Microsoft 365 locations** **- Block only people outside your organization**
User Notification	**On** **- Notify user with a policy tip** **– Notify the user who sent, shared, or last modified the content**
User Override	**Off**
Incident reports	**On** **- Severity Level – Low** **- Send alert every time an activity matches the rule**
Additional Options	**Off**
Status	**Run the policy in simulation mode** **- Policy Tips** not selected

Rule: Adaptive Protection audit rule for Teams and Exchange DLP

Expand table

DLP policy element	Configured value
Conditions	**Insider risk level for Adaptive Protection is** **- Moderate Risk Level, Minor Risk Level** AND **- Content is Shared from Microsoft 365 With people outside my organization**
Actions	None
User Notification	**On** **- Notify user with a policy tip** **- Notify the user who sent, shared, or last modified the content**
User Override	**Off**
Incident reports	**On** **- Severity Level – Low** **- Send alert every time an activity matches the rule**
Additional Options	**Off**
Status	**Run the policy in simulation mode** - **Policy tips** not selected

Policy values for Devices DLP policy

This is the configuration for the Devices DLP policy created during Quick Setup. The policy name is **Adaptive Protection policy for Endpoint DLP**.

> **ⓘ Important**
>
> For Adaptive Protection to work on Devices, you must either enable Advanced classification scanning and protection or if you are manually creating the Adaptive Protection policy, select the **File Type is** condition.

> **ⓘ Important**
>
> If a user is targeted by a default Adaptive Protection Device DLP policy and is targeted by an independent Device DLP policy, only the actions of the *most restrictive* policy will be applied.

Rule: Adaptive Protection block rule for Endpoint DLP

Expand table

DLP policy element	Configured value
Conditions	**Insider risk level for Adaptive Protection is** **- Elevated Risk Level** AND **- File Type is** **- Word processing** **- Spreadsheet** **- Presentation** **- Archive** **- Mail**
Actions	**Audit or Restrict activities on Devices** **- Upload to a restricted cloud service domain or access from unallowed browsers - Block** **File activities for all apps** **- Apply restrictions to specific activity** **- Copy to clipboard – Block** **- Copy to removable USB device – Block** **- Copy to network share – Block** **- Print – Block** **Restricted App activities - Access by restricted apps - Block**
User Notification	**Off**
User Override	**Off**
Incident reports	**On** **- Severity Level – Low** **- Send alert every time an activity matches the rule**
Additional Options	**Off**
Status	**Run the policy in simulation mode** **- Policy Tips** option not selected

Rule: Adaptive Protection rule for Endpoint DLP

Expand table

DLP policy element	Configured value
Conditions	**Insider risk level for Adaptive Protection is** **- Moderate Risk Level, Minor Risk Level**

DLP policy element	Configured value
	AND - **File Type is** - **Word processing** - **Spreadsheet** - **Presentation** - **Archive** - **Mail**
Actions	**Audit or Restrict activities on Devices** **- Upload to a restricted cloud service domain or access from unallowed browsers – Audit** **File activities for all apps** **- Apply restrictions to specific activity** **- Copy to clipboard – Audit** **- Copy to removable USB device – Audit** **- Copy to network share – Audit** **-Print – Audit** **Restricted App activities** **- Access by restricted apps - Audit**
User Notification	**Off**
User Override	**Off**
Incident reports	**On** **- Severity Level – Low** **- Send alert every time an activity matches the rule**
Additional Options	**Off**
Status	**Run the policy in simulation mode** - **Policy tips** option not selected

See Also

- Adaptive Protection in Microsoft Purview

Feedback

Was this page helpful? Yes No

Learn about data loss prevention simulation mode

Article • 03/14/2025

Run the policy in simulation mode, otherwise known as simulation mode, for Microsoft Purview Data Loss Prevention (DLP) policies replaces the **Test** and **Test with policy tips** policy states. When a policy is in simulation mode, it's run as if it were being enforced, without any actual enforcement. Unlike the **Test** modes, all matched items and alerts are reported in a separate dashboard. This makes it easy to see the impact of the policy before you enforce it by keeping all the simulation results separate from the results of policies that are being enforced.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Simulation mode provides:

- An isolated experience to run and assess policies.
- A summary dashboard that gives you visibility into the impact of the policies across different locations and shows which items were matched.
- A flat list of matched items at a policy level.

Simulation mode for DLP policies is a tool you can use for tuning your data loss prevention policies at any time. You should incorporate it into your policy creation and deployment process. Using simulation mode to tune a policy reduces false positives without impact to your users or business processes. Use it as part of your deployment process for new policies, use it to test changes to existing policies before enforcing those changes in production.

For example:

DLP policy *Protect Credit Cards v1* is in production, but is throwing too many false positives. You think you know what is wrong, but you don't want to experiment with changes to it in production to find out. You can make a copy of policy *Protect Credit Cards v1*, call it *Protect Credit Cards v2*, make tuning changes, and then run *v2* in simulation mode. If it the changes have the desired result, then you can turn off *v1* and set *v2* to enforce mode.

> **Important**
>
> Simulations can run for up to 15 days so the results are not a point in time snapshot. For SharePoint Online and OneDrive for Business locations, all existing and new/changed items are evaluated. For Exchange, Teams, Devices locations only items that are new during simulation are evaluated. Data from a simulation run is kept for 30 days.

Placing a policy in simulation mode

See Get started with the data loss prevention simulation mode for the procedures to manage simulation mode for a policy.

Simulation results

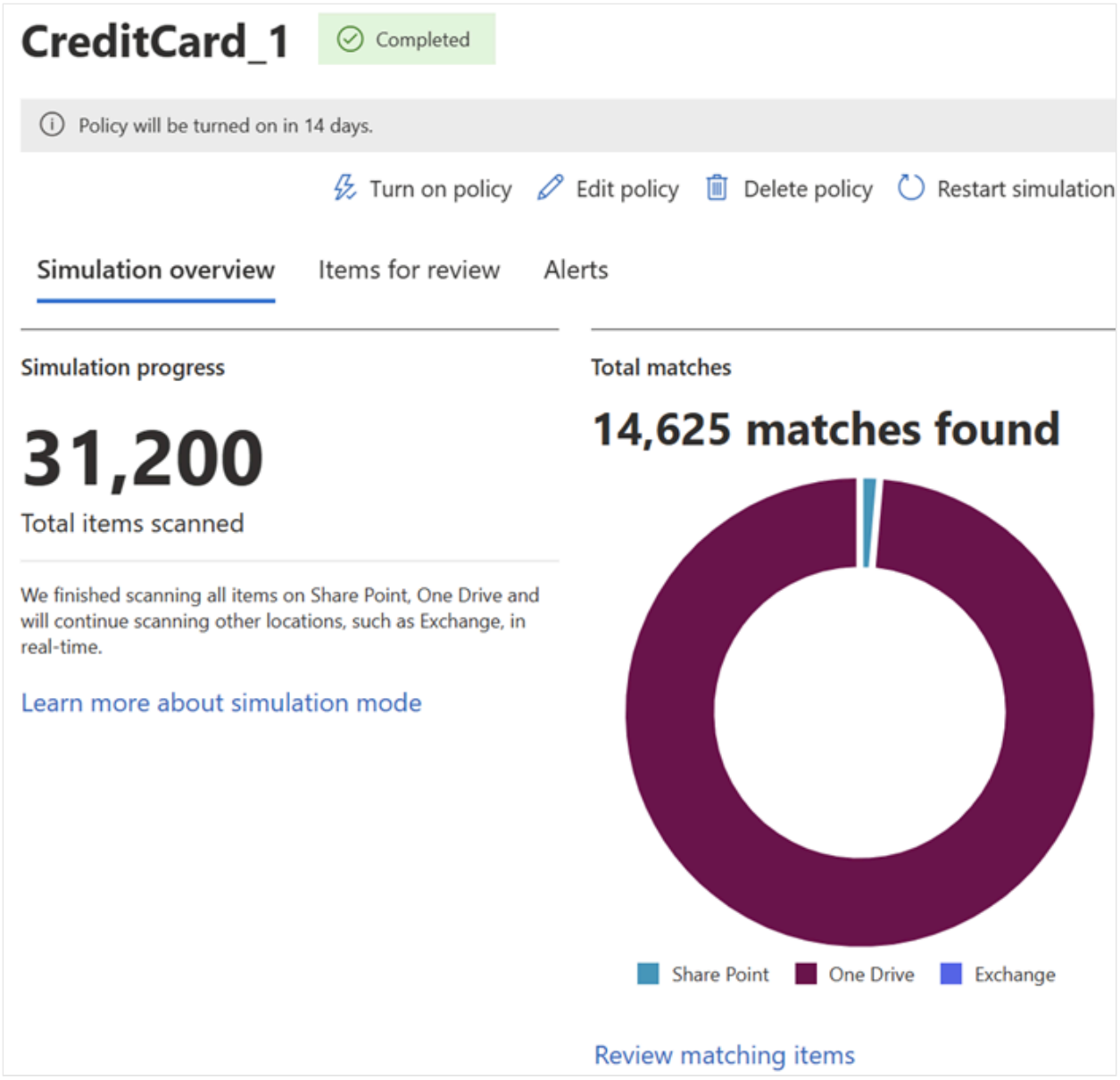
See Get started with the data loss prevention simulation mode for the steps to access simulation results. Simulation results are presented across three pages.

Simulation overview

The **Simulation overview** tab gives you basic information on:

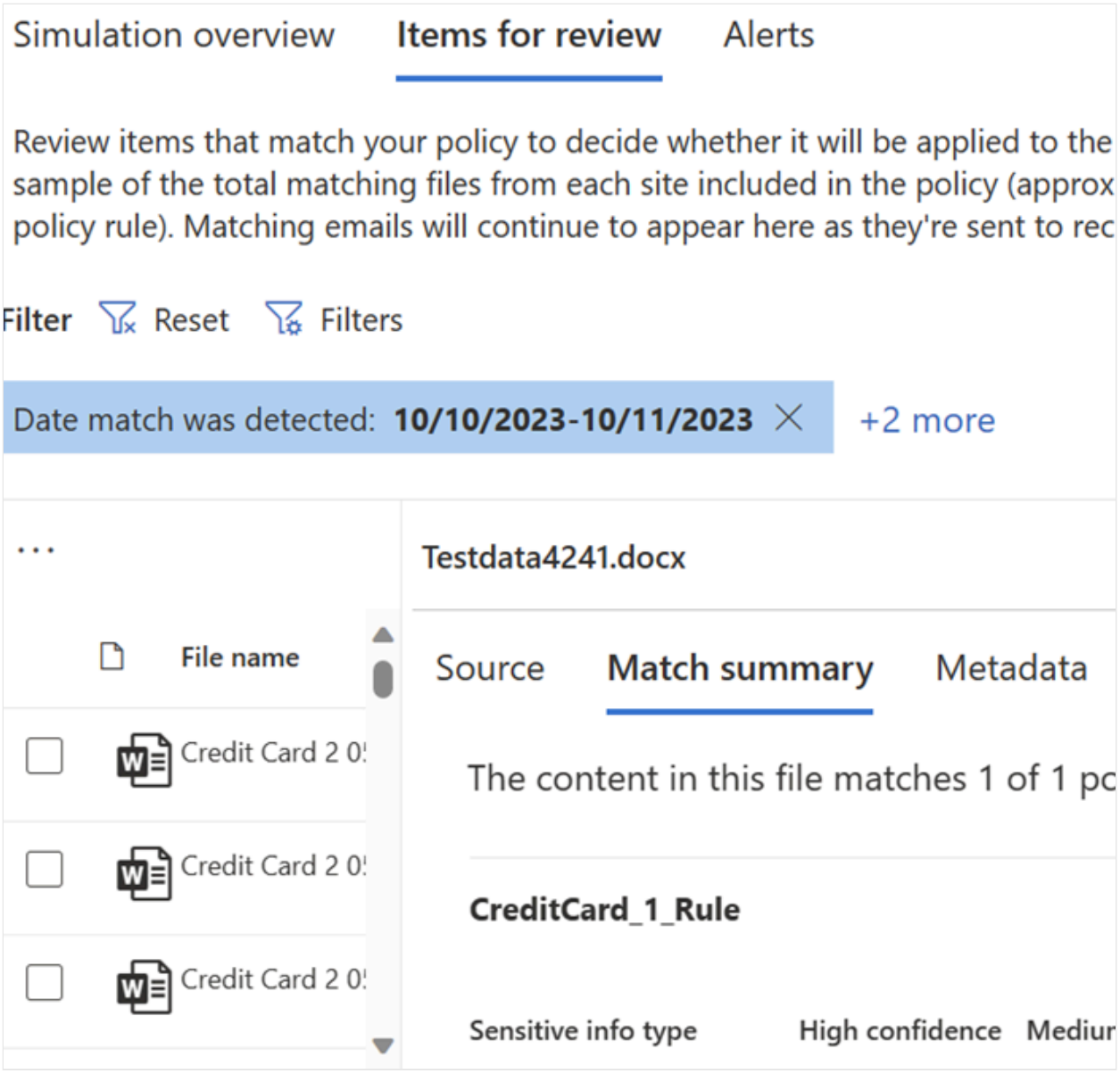
- the status of the policy simulation (complete, in progress or expired)
- status of scanning per location
- total items scanned
- total matches found
- matched items for each location
- links to the matched items

and other information.



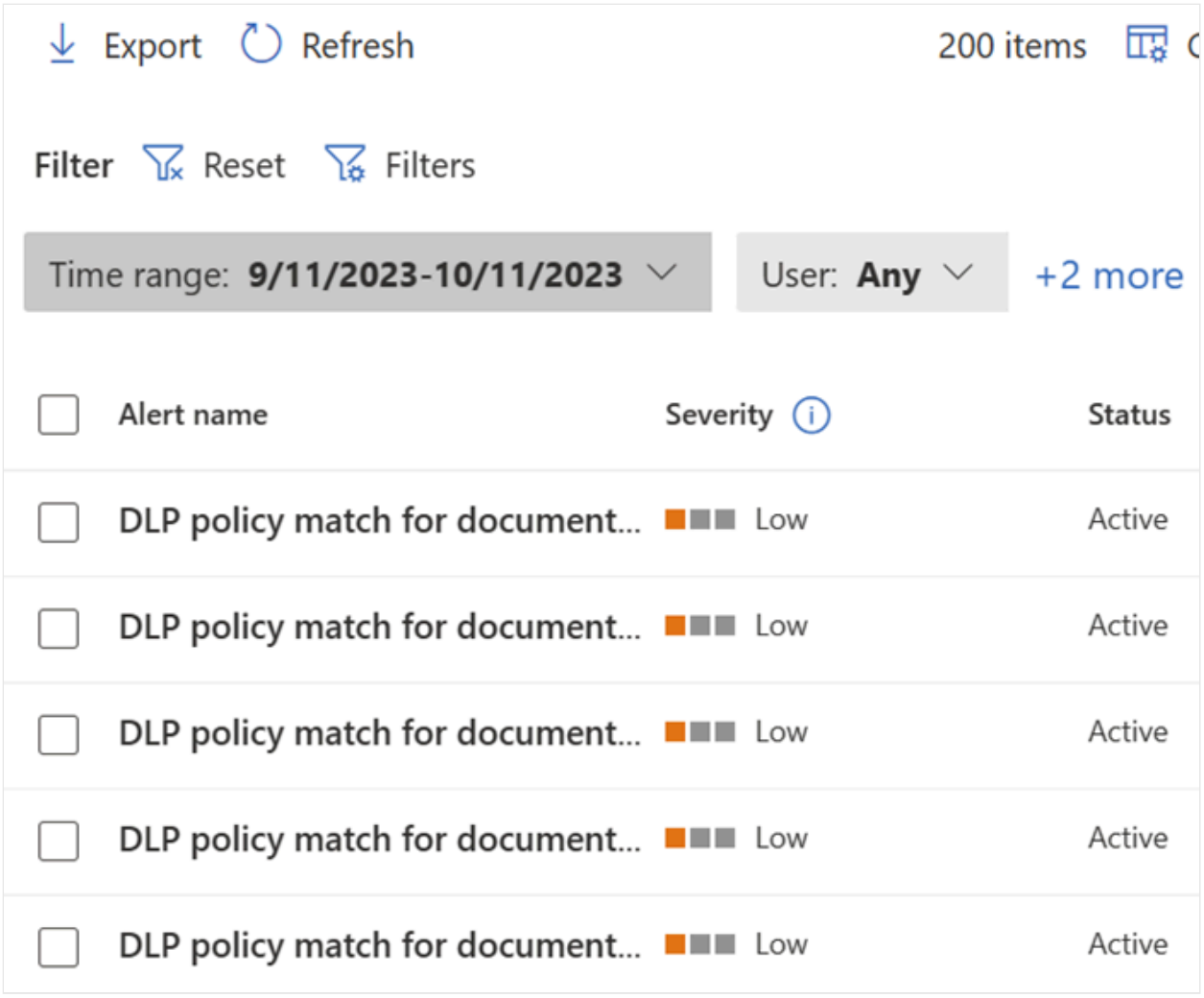
Items for review

The **Items for review** tab shows a flat list of the items that matched the policy in simulation and lets you read the source item along with item metadata much like Content explorer.



Alerts

The simulation mode alerts tab shows a flat list of all the alerts that were generated when an item matched the policy in simulation. It has the same format as the DLP alerts console.



Important

While a policy is in simulation mode, all simulation alerts are only surfaced in the simulation alerts tab. They don't show up in the DLP alerts console and they don't flow into the Microsoft Defender portal.

Feedback

Was this page helpful? Yes No

Get started with the data loss prevention simulation mode

You can use Microsoft Purview Data Loss Prevention (DLP) simulation mode to see:

- The impact of a policy on your production environment without enforcement.
- All the items that would be matched by a policy if it were enforced.

This article walks you through simulation mode prerequisites, configuration options and how to view simulation results.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Before you begin

Licensing

For information on licensing, see

- Microsoft 365 Enterprise Plans
- Microsoft 365 Service Descriptions

Permissions

The account you use to interact with simulation mode must be in the Information Protection admin role. For more information on the roles and role groups necessary to use simulation mode, see Permissions. For more information on roles and role groups in Microsoft Purview compliance, see Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview compliance

Configure system settings for DLP simulation mode

To see matched items from endpoint devices in their native application on the **Items for review**, you must configure evidence collection for file activities on devices.

Manage DLP simulation mode

You can set a policy to be in simulation mode when you create it or after it's been created. You can also turn off simulation mode for a policy that's already in simulation mode.

1. Use the steps in Create and Deploy data loss prevention policies to create a new policy or edit an existing policy.
2. The last step in the policy configuration workflow is **Simulate or turn on the policy**. Select **Run the policy in simulation mode** to enable simulation mode. Select either **Turn it on right away** or **Keep it off** to disable simulation mode. You can further select:
   a. **Show policy tips with in simulation mode** to help educate your users when they take actions that might trigger policy actions.
   b. **Turn the policy on if it's not edited within fifteen days of the simulation** to turn the policy on without further interaction.
3. Select **Next** and **Submit**.

After disabling simulation mode, it can take up to 24 hours for the insights to stop appearing on the Overview page.

Viewing DLP policies in simulation mode

Perform the following steps to view a policy that is running in simulation mode.

1. Sign in to the Microsoft Purview portal > **Data Loss Prevention** > **Policies**.
2. Select a policy with a status of **In simulation** or **In simulation with notifications** to open the fly-out pane.
3. Select **View simulation** to see the **Simulation overview**, **Items for review**, and **Alerts** tabs.

Simulation mode status messages

> **Note**
>
> Scan results from running a policy in simulation mode are saved for 30 days. You can continue to run the policy in simulation mode longer than that, but only the results for the most recent 30-day period are displayed.

When a policy runs in simulation mode, content scanning in most locations occurs in real time, that is, it occurs when an email or Teams message is sent. Content scanning for the SharePoint and OneDrive locations works a little differently. In addition to scanning content when documents are uploaded to these locations, policies running in simulation mode can display three additional progress messages: *Completed*, *In progress*, and *Expired*. The following table identifies and describes the status messages available for each location:

Simulation mode scan status by location

The following table lists the scan status messages by location.

Expand table

Location	Status message descriptions
Exchange	**Real-time** - Content is scanned when message is sent
SharePoint	**Completed** - Signifies that scanning the existing content in Sharepoint and/or OneDrive is complete. This status message only displays when SharePoint and/or OneDrive are the only locations in scope. **In progress** - Signifies that the simulation is running. Results are updated as more matches are found.
OneDrive	**Completed** - Signifies that scanning the existing content in Sharepoint and/or OneDrive is complete. This status message only displays when SharePoint and/or OneDrive are the only locations in scope. **In progress** - Signifies that the simulation is running. Results are updated as more matches are found.
Teams chat and channel messages	**Real-time** - Content is scanned when message is sent
Devices	**Real-time** - Content is scanned when transferred off the device
Fabric and Power BI	**Real-time** - Content is scanned when uploaded

The following table explains the status messages related to the progress of the overall policy simulation.

Overall simulation mode status

The following table explains the overall simulation status values.

Expand table

Simulation status	Description
Completed	Only available when a DLP policy scope is limited to SharePoint, OneDrive, or both. Indicates that all data at rest has been scanned.
In progress	Simulation scanning is underway. Results are updated as additional policy matches are detected.
Expired	The policy being simulated is more than 30 days old. Microsoft Purview retains simulation data for only 30 days. To get scan results for data at rest scanned prior to the most recent 30-day window, re-run the scan.

Only the first 100 items matched in the SharePoint and OneDrive locations are displayed for review. This might differ from the total number of matched items.

Simulation events are displayed up in activity explorer. You can filter on *Policy mode*, which has *TestWithNotifyUser*, *TestWithoutNotifyUser* and *enforce* values.

See also

- Design a DLP policy
- Create and deploy a DLP policy

ⓘ**Note:** The author created this article with assistance from AI. Learn more

Last updated on 06/26/2026

Learn about Advanced label-based protection for all files on devices

07/22/2025

This setting enhances productivity for users that label and encrypt file types other than those for Office and PDF by using the Microsoft Purview Information Protection client. Typical examples of file types other than Office and PDF include .txt, .jpg, .csv, and files from third-party applications.

When users label and encrypt these files with the information protection client and without this setting for advanced label-based protection:

- The encrypted file changes its file name extension and becomes read-only, and can no longer be opened by apps that support the original file name extension. Instead, the file opens in the Microsoft Purview Information Protection viewer. To make any changes to the file, users must manually remove the label with encryption.
- Although the encryption can include restrictive permissions, such as not allowing copy and save for specific users, not all files can support these restrictions. As a result, when these files are opened in the information protection viewer, users are informed of the configured permissions, but the permissions can't be enforced. This type of encryption is referred to as generic, rather than native.

For a better understanding of the file name extension changes for encrypted files, see Supported file types from the information protection client documentation. Also be aware of the standard client exceptions for files that are critical for computer operations.

When you configure users for the endpoint data loss prevention setting **Advanced label-based protection for all files on devices** and use the information protection client, the following changes occur for file types that aren't supported by Office or that are PDF files:

- When a user selects a sensitivity label that applies encryption, the file name extension now doesn't change. As a result, there's no change to the user workflow because they can continue to view and edit the file in their standard application. Endpoint DLP tracks and monitors the file, enforcing configured permissions without requiring the information protection viewer.

> **Note**
>
> Because the file name extension hasn't changed, users can confirm that the label is applied by using the Microsoft Purview Information Protection File Labeler.

- Only if the user copies or moves the file from the computer, does the file name change:
  - If the file is copied or moved to a network drive or USB device, the file there has the changed file name extension that requires the information protection viewer to open it.
  - If the file is copied or moved to any other location (Bluetooth device, over remote desktop, or uploaded to the cloud), a local copy of the file is created with the changed file name extension that requires the information protection viewer to open it. Endpoint DLP informs the user that they must use this version of the file to manually copy or move it to the new location.
  - If the operation to encrypt the file fails, the copy or move activity is blocked to ensure that the unencrypted file remains on the device and isn't exfiltrated.

Here's a summary of the differences when you label files other than Office and PDF and you use just the information protection client rather than the client with endpoint DLP advanced label-based protection:

Expand table

Labeling behavior	Just the client	The client with endpoint DLP setting enabled
Files labeled with encryption retain their original file name extension	No The file name extension always changes	Yes The file extension is changed when egress happens.
Files labeled with encryption can be opened and edited with the original application or others that support the original file type	No Always requires the information protection viewer	Yes The file extension is changed when egress happens.
Configured permissions for labeled and encrypted files are enforced	Native encryption: Yes Generic encryption: No	Yes The following permissions are enforced until egress: view, extract, print

Prepare your Windows devices

Make sure that the Windows devices that you need to onboard meet these requirements:

- Anti-malware Client Version 4.18.25050 or newer is required.

- Window build:
  - Windows 10: See update details
  - Windows 11 22H2: See update details
  - Windows 11 23H2: See update details
  - Windows 11 24H2: See update details
- Information protection client version: 3.1.309 or newer is required.

User experience

Expand table

Labeling scenario	Activity	User experience	Activity explorer
File being labeled	Information protection File Labeler	A user right-clicks to open Information Protection Viewer to label specific file. The user can assign a label through the Viewer, Endpoint DLP will assign the label to the file but keep the file in original extension.	Label applied, no DLP Rule Matched, no Alert
Labeled file being moved off of the device	File copied to removable media, File copied to network share	A user moves a labeled file to a removable media or network share, the file will be encrypted, e.g. .ptxt file extension, when land on the network share or removable media with a Windows toast. This version of the file must be opened with the information protection viewer and can't be monitored by DLP.	File copied to USB removable device or File copied to a network share, no DLP Rule Matched, no Alert
Labeled file being moved off of the device	File transferred through Bluetooth, File transferred over remote desktop, or File uploaded to the cloud	A user uploads a labeled file out of the machine through listed channels, the activity will be blocked, a copy of the file will be created, and the copied file will be encrypted, e.g. .ptxt file extension, a toast will pop up asking the user to retry the operation with the labeled new copy. This version of the file must be opened with the information protection viewer and can't be monitored by DLP. The original file on the computer retains the .txt file extension and can still be opened and edited with any text editor. DLP continues to monitor this file.	File copied or moved using unallowed Bluetooth app or File copied or moved using RDP or File copied to cloud app

Labeling scenario	Activity	User experience	Activity explorer
Encrypted file being moved on the device that is using this feature, for example .ptxt	Information protection Viewer	A user moves an encrypted labeled file into a machine that has this feature onboarded, the file will be kept in encrypted extension, the user can double-click to open the Information Protection Viewer. The Information Protection Viewer will pop up a notification, user can click Continue button to decrypt the file to original file extension with label assigned.	
Encrypted file being moved on the device that is using this feature onboarded, for example .ptxt	Information protection File Labeler	A user moves an encrypted labeled file into a machine that has this feature onboarded, the file will be kept in encrypted extension, the user can right-click to open the Information Protection Viewer. The user can change the label through the Viewer, the file will be kept in encrypted format.	Label changed, Label removed, Protection applied, Protection removed.
Labeled file being opened through a native application	N/A	A user opens a labeled file through a native application, and perform Print or Copy activity which violate the permission associated with the label, the activity will be blocked with a toast.	File copied to clipboard or File printed, no DLP Rule Matched, no Alert

Considerations

- Supported only for sensitivity labels that apply encryption. For sensitivity labels without encryption, the behavior and supported file types are the same as for the information protection client.
- Not supported for labeling files on network locations or USB drives.
- The applied sensitivity labels won't be detected as conditions in DLP policies before egress.
- When multiple files are selected for labeling, and they include Office files or PDF files, the Office files, and PDF files won't be labeled or encrypted. User will see failure message on the Information Protection File Labeler application.
- Endpoint DLP doesn't enforce all rights management usage rights before egress: VIEW, EXTRACT, and PRINT are supported.
- The Save As option doesn't automatically inherit the current label and encryption settings. The user must label the new file.

- The Microsoft Purview Information Protection PowerShell module doesn't recognize the DLP setting. Labeling files using PowerShell will encrypt and change the file extension. Only use the Microsoft Purview Information Protection File Labeler when you enable the DLP setting.

Configure device proxy and internet connection settings for Microsoft Purview Endpoint DLP

Microsoft Endpoint technologies use Microsoft Windows HTTP (WinHTTP) to report data and communicate with the Microsoft endpoint cloud service. The embedded service runs in system context using the LocalSystem account.

> **Tip**
>
> For organizations that use forward proxies as a gateway to the Internet, you can use network protection to investigate behind a proxy. For more information, see **Investigate connection events that occur behind forward proxies**.

The WinHTTP configuration setting is independent of the Windows Internet (WinINet) Internet browsing proxy settings and can only discover a proxy server by using the following auto discovery methods:

- Transparent proxy
- Web Proxy Auto-discovery Protocol (WPAD)

  > **Note**
  >
  > If you're using Transparent proxy or WPAD in your network topology, you don't need special configuration settings. For more information on Defender for Endpoint URL exclusions in the proxy, see **Enable access to Endpoint DLP cloud service URLs in the proxy server**.

- Manual static proxy configuration:
  - Registry-based configuration
  - WinHTTP configured using netsh command – Suitable only for desktops in a stable topology (for example: a desktop in a corporate network behind the same proxy)

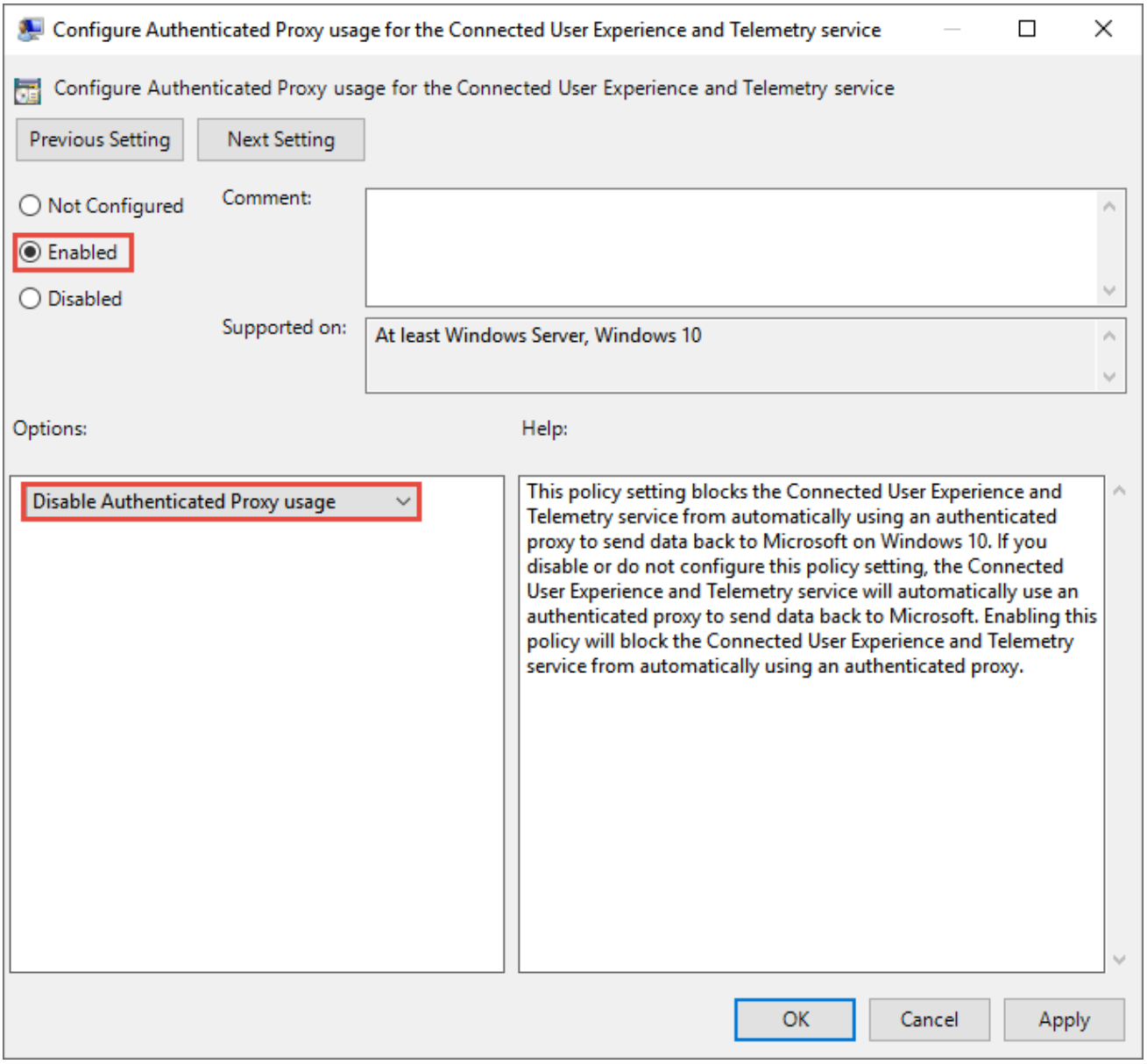
Configure the proxy server manually using a registry-based static proxy

For endpoint devices that aren't permitted to connect to the Internet, you need to configure a registry-based static proxy. You need to configure the registry-based static proxy to allow only Microsoft Endpoint DLP to report diagnostic data and communicate with Microsoft endpoint cloud service.

The static proxy is configurable through Group Policy (GP). The Group Policy settings can be found under:

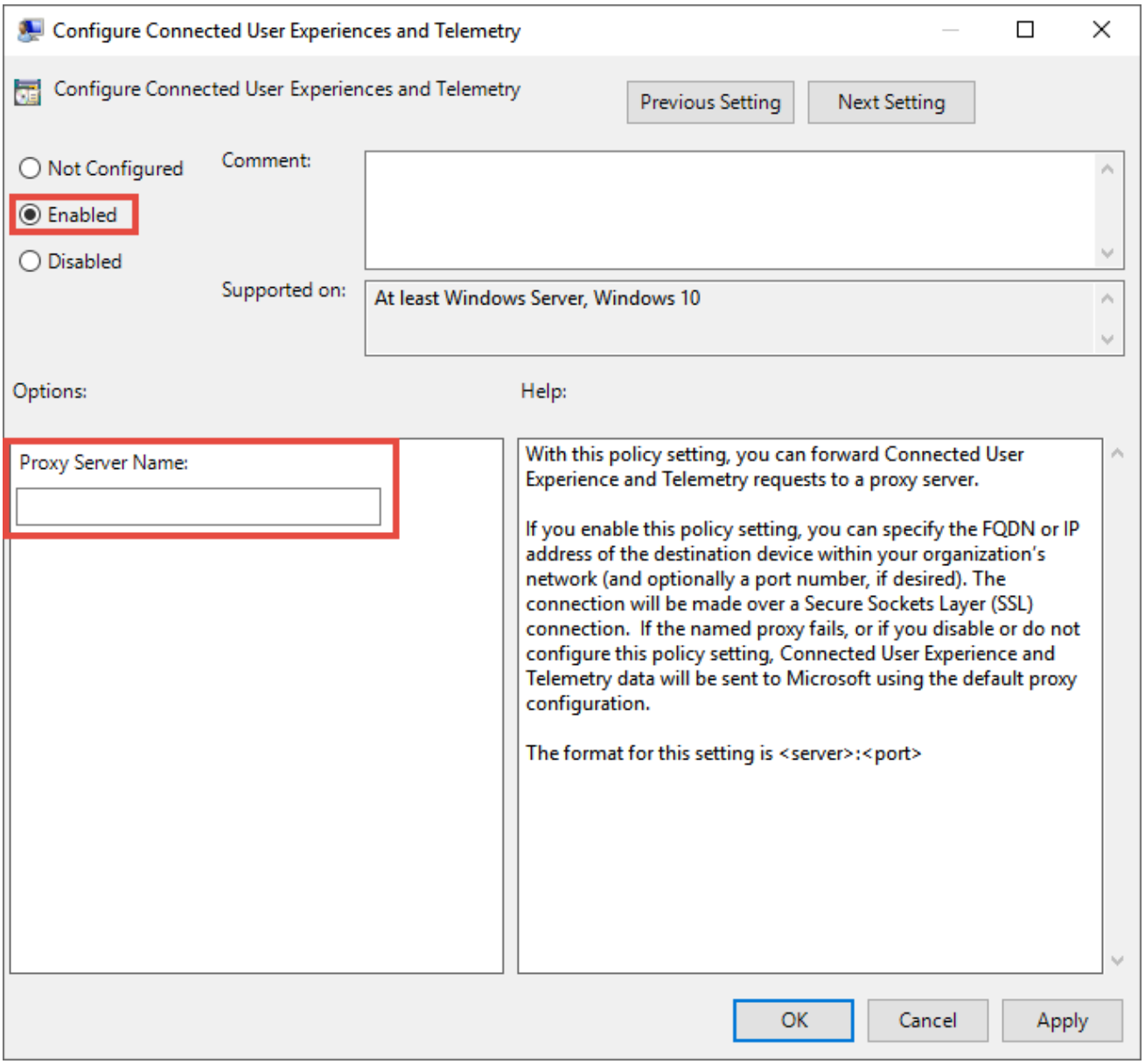
1. Open **Administrative Templates** > **Windows Components** > **Data Collection and Preview Builds** > **Configure Authenticated Proxy usage for the Connected User Experience and Telemetry Service**

2. Set it to **Enabled** and select **Disable Authenticated Proxy usage**:



3. Open **Administrative Templates** > **Windows Components** > **Data Collection and Preview Builds** > **Configure connected user experiences and telemetry**:

   Configure the proxy



   The policy sets two registry values `TelemetryProxyServer` as REG_SZ and `DisableEnterpriseAuthProxy` as REG_DWORD under the registry key `HKLM\Software\Policies\Microsoft\Windows\DataCollection`.

   The registry value TelemetryProxyServer is in this format <server name or ip>:<port>. For example: **10.0.0.6:8080**

   The registry value `DisableEnterpriseAuthProxy` should be set to 1.

Configure the proxy server manually using "netsh" command

Use netsh to configure a system-wide static proxy.

> **Note**
>
> This will affect all applications including Windows services which use WinHTTP with default proxy. - Laptops that are changing topology (for example: from office to home) will malfunction with netsh. Use the registry-based static proxy configuration.

1. Open an elevated command line:
   a. Go to **Start** and type **cmd**
   b. Right-click **Command prompt** and select **Run as administrator**.

2. Enter the following command and press **Enter**:

   ```
   netsh winhttp set proxy <proxy>:<port>
   ```

   For example: **netsh winhttp set proxy 10.0.0.6:8080**

3. To reset the winhttp proxy, enter the following command and press **Enter**:

   ```
   netsh winhttp reset proxy
   ```

See Netsh Command Syntax, Contexts, and Formatting to learn more.

Enable access to Endpoint DLP cloud service URLs in the proxy server

If a proxy or firewall is blocking all traffic by default and allowing only specific domains through, add the domains listed in the Microsoft Defender for Endpoint service URLs spreadsheet to the allowed domains list.

This Microsoft Defender for Endpoint service URLs spreadsheet lists the services and their associated URLs that your network must be able to connect to. You should ensure that there are no firewall or network filtering rules that would deny access to these URLs, or you may need to create an allow rule specifically for them.

If a proxy or firewall has HTTPS scanning (SSL inspection) enabled, exclude the domains listed in the Microsoft Defender for Endpoint service URLs spreadsheet from HTTPS scanning. If a proxy or firewall is blocking anonymous traffic, as Endpoint DLP is connecting from system context, make sure anonymous traffic is permitted in the previously listed URLs.

Verify client connectivity to Microsoft cloud service URLs

Verify the proxy configuration completed successfully, that WinHTTP can discover and communicate through the proxy server in your environment, and that the proxy server allows traffic to the Defender for Endpoint service URLs.

1. Download the MDATP Client Analyzer tool to the PC where Endpoint DLP is running on.

2. Extract the contents of MDATPClientAnalyzer.zip on the device.

3. Open an elevated command line:
   a. Go to **Start** and type **cmd**.
   b. Right-click **Command prompt** and select **Run as administrator**.

4. Enter the following command and press **Enter**:

   ```
   HardDrivePath\MDATPClientAnalyzer.cmd
   ```

   Replace *HardDrivePath* with the path where the MDATPClientAnalyzer tool was downloaded to, for example

   C:\Work\tools\MDATPClientAnalyzer\MDATPClientAnalyzer.cmd

5. Extract the **MDATPClientAnalyzerResult.zip*** file created by tool in the folder used in the *HardDrivePath*.

6. Open **MDATPClientAnalyzerResult.txt** and verify that you have performed the proxy configuration steps to enable server discovery and access to the service URLs. The tool checks the connectivity of Defender for Endpoint service URLs that Defender for Endpoint client is configured to interact with. It then prints the results into the **MDATPClientAnalyzerResult.txt** file for each URL that can potentially be used to communicate with the Defender for Endpoint services. For example:

   Windows Command Prompt

   ```
   Testing URL: https://xxx.microsoft.com/xxx
   1 - Default proxy: Succeeded (200)
   2 - Proxy auto discovery (WPAD): Succeeded (200)
   3 - Proxy disabled: Succeeded (200)
   4 - Named proxy: Doesn't exist
   5 - Command-line proxy: Doesn't exist
   ```

If at least one of the connectivity options returns a (200) status, then the Defender for Endpoint client can communicate with the tested URL properly using this connectivity method.

However, if the connectivity check results indicate a failure, an HTTP error is displayed (see HTTP Status Codes). You can then use the URLs listed in Enable access to Endpoint DLP cloud service URLs in the proxy server. The URLs you'll use will depend on the region selected during the onboarding procedure.

> **Note**
>
> The Connectivity Analyzer tool is not compatible with attack surface reduction rule **Block process creations originating from PSExec and WMI commands**. You will need to temporarily disable this rule to run the connectivity tool.
>
> When the TelemetryProxyServer is set, in Registry or via Group Policy, Defender for Endpoint will fall back to direct if it can't access the defined proxy. Related topics:
>
> - Onboard Windows 10 devices
> - Troubleshoot Microsoft Endpoint DLP onboarding issues

See also

For related information, see the following articles:

- Learn about Endpoint data loss prevention
- Using Endpoint data loss prevention
- Learn about data loss prevention
- Create and Deploy data loss prevention policies
- Get started with Activity explorer
- Microsoft Defender for Endpoint
- Onboarding tools and methods for Windows 10 machines
- Microsoft 365 subscription
- Microsoft Entra joined devices
- Download the new Microsoft Edge based on Chromium

Note: The author created this article with assistance from AI. Learn more

Last updated on 06/26/2026

Get started with Power Automate integration in Microsoft Purview DLP

Microsoft Power Automate is a workflow service that automates actions across applications and services. By using flows from templates or created manually, you can automate common tasks associated with these applications and services. Microsoft Purview Data Loss Prevention (DLP) can take various actions across locations on sensitive content identified as per DLP policies. With Power Automate integration, you can now trigger custom Power Automate workflows as a DLP rule action.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Before you begin

Before you start using Power Automate integration in Microsoft Purview DLP, read the following information on subscription, licensing, permissions, and supported locations.

Subscriptions & licensing

Review the following subscription and licensing requirements for Microsoft Purview DLP and Power Automate.

Purview DLP

Power Automate integration in Microsoft Purview DLP is available for all existing DLP customers with no other Purview license requirements.

For information on licensing, see

- Microsoft 365 Enterprise Plans
- Microsoft 365 Service Descriptions

Power Automate

The access premium connectors when configuring custom flows you need a Power Automate plan. A Power Automate licenses is not needed to access the templates.

Permissions

The following permissions apply when using Power Automate integration in Microsoft Purview DLP:

- There is no change to required role for DLP policy authoring in Microsoft Purview portal to use Power Automate integration as a DLP rule action.
- See Get started with Power Automate for information about Power Automate roles.

Supported locations in policy

Power Automate integration for DLP supports the following policy locations:

- Exchange emails
- Files in SharePoint Online & OneDrive
- Windows devices (24H2 and above)

Create a workflow

Create a custom workflow using the DLP trigger provided as part of Microsoft Purview connector. You can create a workflow from any of the following experiences:

- Power Automate portal
- DLP rule authoring page: Add **Start a Power Automate workflow** as an action, and then create the workflow using a template or a custom workflow using DLP trigger.

You can select or update the Power Automate workflow linked to a DLP rule just like any other rule action.

Use a template

The following Power Automate template is available to support process automation for DLP rule violations:

Notify manager when a DLP policy is violated: Some organizations might need to have immediate management notification when a user violates a policy. When the **Notify manager when a DLP policy is violated** flow is configured and selected, the manager of the violating user is sent an email message with the following information about the event:

- User name
- Policy name

You can customize the sender and add additional information by editing the workflow in Power Automate.

Create a custom Power Automate flow

Some processes and workflows for your organization might not fit an available template. If no available template fits your organization's processes, you can create custom Power Automate flows to meet your specific needs. Power Automate flows are flexible and support extensive customization, but can require configuration to integrate with data loss prevention policies:

1. Create a flow that performs one or more tasks after a data loss prevention event. For details on how to create an automated flow, see Create a flow in Power Automate.
2. Select the Microsoft Purview connector: Search for and select the Microsoft Purview connector. For more information on connectors, see: Connector reference overview.
3. Choose data loss prevention triggers for your flow.
4. Choose from available Power Automate connectors for your flow.

Share a Power Automate flow

By default, a Power Automate flow is only available to the user who created it. For other administrators to access and use a flow, the flow creator must share it. For information on how to share a flow, see: Share a cloud flow

Known issues and limitations

For files in SharePoint and OneDrive, Power Automate workflows are executed only for new or modified content. Preexisting files matching the rules don't trigger a flow.

Related content

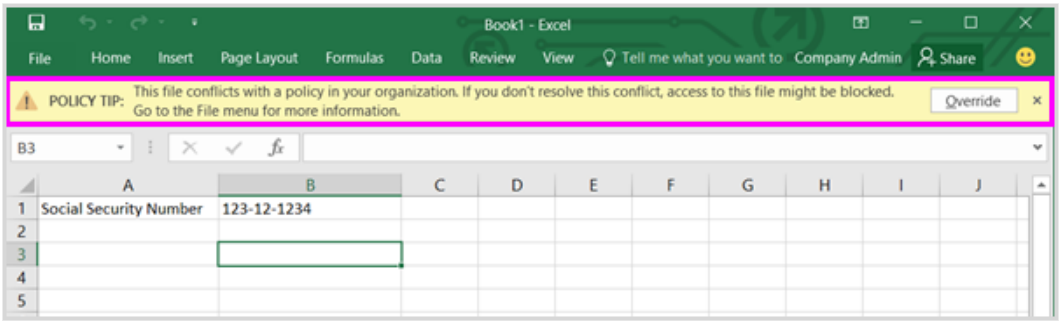
- Learn about data loss prevention
- Create and Deploy data loss prevention policies
- Create a flow in Power Automate
- Microsoft 365 subscription

ⓘ**Note:** The author created this article with assistance from AI. Learn more

Last updated on 06/26/2026

Send email notifications and show policy tips for DLP policies

You can use a Microsoft Purview Data Loss Prevention (DLP) policy to identify, monitor, and protect sensitive information across Microsoft 365. You want people who work with sensitive information to stay compliant with your DLP policies without being blocked from getting their work done. This is where email notifications and policy tips can help.



When you create a DLP policy, you can configure the user notifications to:

- Send an email notification to the people you choose that describes the issue.
- Display a policy tip for content that conflicts with the DLP policy:
  - For Outlook on the web and Outlook 2013 and later, the policy tip appears at the top of a message above the recipients while the message is being composed.
  - For documents in a OneDrive account or SharePoint site, the policy tip shows as a warning icon that appears on the item. To view more information, you can select an item and then choose **Information** in the upper-right corner of the page to open the details pane.
  - For Excel, PowerPoint, and Word documents that are stored on a OneDrive site or SharePoint site that's in scope of a DLP policy, the policy tip appears on the Message Bar and the Backstage view (**File** menu > **Info**).

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

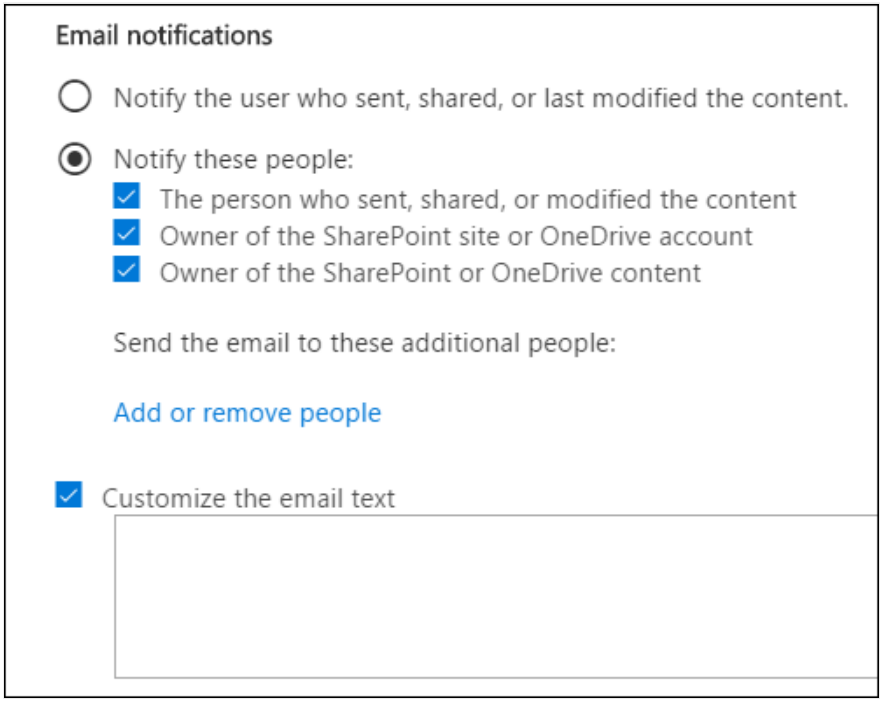
> **Note**
>
> Notification emails are sent unprotected.

Options for configuring email notifications

For each rule in a DLP policy, you can:

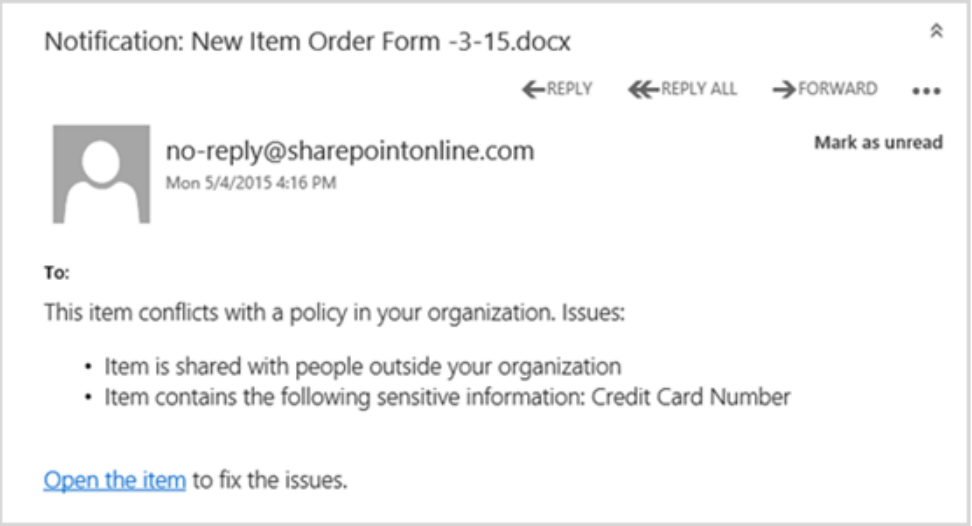
- Send the notification to the people you choose. These people can include the owner of the content, the person who last modified the content, the owner of the site where the content is stored, or a specific user.
- Customize the text that's included in the notification by using HTML or tokens. For more information, see Custom email notifications.

> **Note**
>
> - You can send email notifications only to individual recipients, not to groups or distribution lists.
> - Only new content triggers an email notification. Editing existing content triggers policy tips, but not email notifications.
> - External senders receive only a templatized notification without full details to prevent any unintended loss of information about the policy configuration.



Default email notification

Notifications have a **Subject** line that begins with the action taken, such as `Notification`, `Message Blocked` for email, or `Access Blocked` for documents. If the notification is about a document, the notification message body includes a link. That link takes you to the site where the document's stored and opens the policy tip for the document, where you can resolve any issues. If the notification is about a message, the notification includes as an attachment the message that matches a DLP policy.



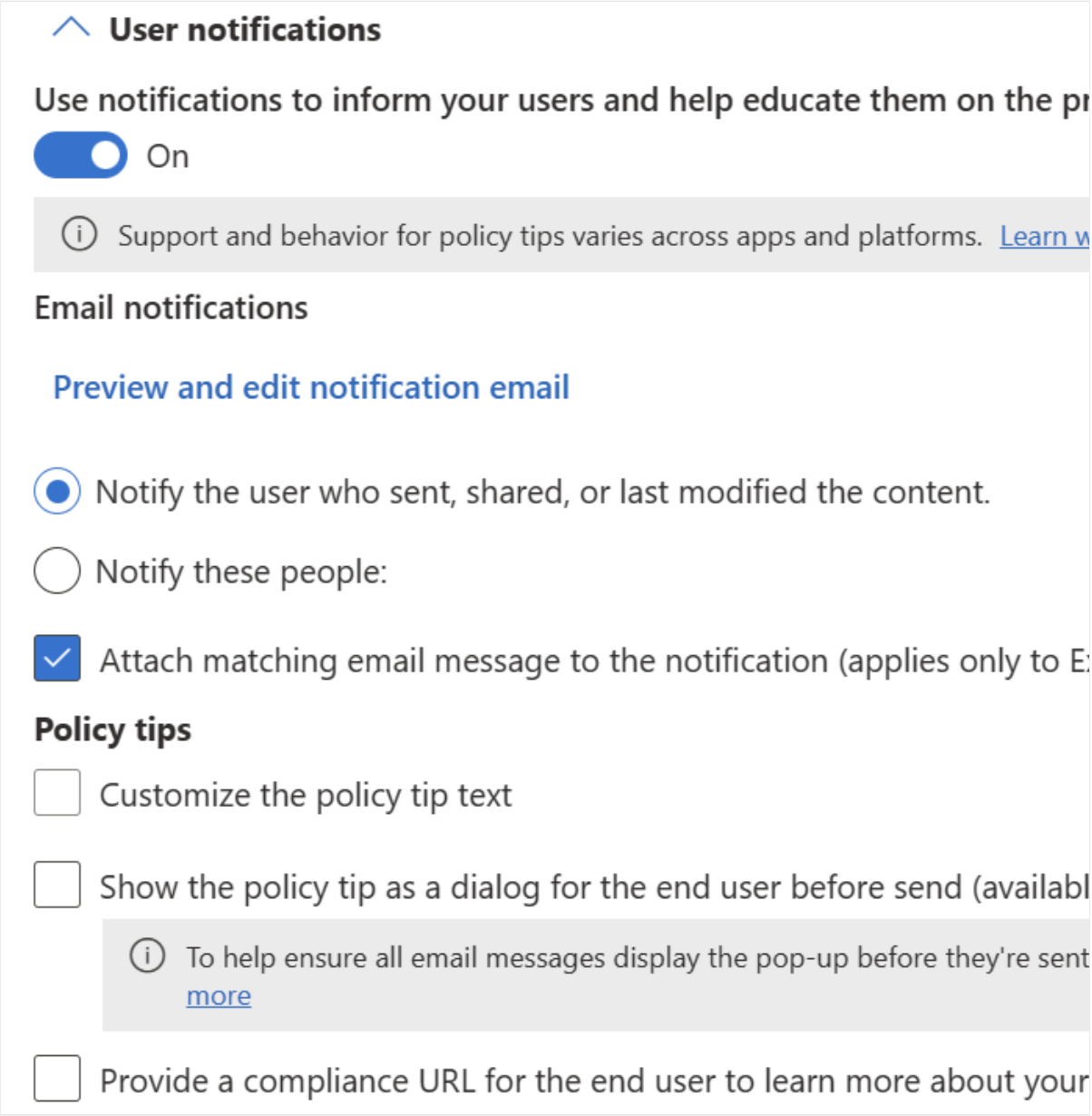
By default, notifications display text similar to the following for an item on a site. You configure the notification text separately for each rule, so the text that's displayed differs depending on which rule is matched.

Expand table

If the DLP policy rule does this...	Then the default notification for SharePoint or OneDrive documents says this...	Then the default notification for Outlook messages says this...
Sends a notification but doesn't allow override	This item conflicts with a policy in your organization.	Your email message conflicts with a policy in your organization.
Blocks access, sends a notification, and allows override	This item conflicts with a policy in your organization. If you don't resolve this conflict, access to this file might be blocked.	Your email message conflicts with a policy in your organization. The message wasn't delivered to all recipients.
Blocks access and sends a notification	This item conflicts with a policy in your organization. Access to this item is blocked for everyone except its owner, last modifier, and the primary site admin.	Your email message conflicts with a policy in your organization. The message wasn't delivered to all recipients.

Custom email notifications

You can create a custom email notification template for end user email notifications for each rule. Then send it instead of sending the default email notification. This option is available for policies scoped to the Exchange, SharePoint, and OneDrive locations.



Select **Preview and edit notification email** to create a customized template when you edit or create a rule in a policy.

The custom email notification supports HTML and customization of these fields:

- Sender display name: Define a custom value that suits your organization's needs. There's a 70 character limit. Setting this value doesn't change the sender's email address.
- Email subject: Define a custom value to make the email subject more meaningful to your end users. There's a 256 character limit.
- Email body: Define a custom value that has greater meaning to your users. There's a 5120 character limit. You can use HTML or markdown to include images, formatting, and other branding in the notification. Custom definition for the email body supports the use of inline styling. If you add actions that users can take to email, only markdown can be used to format the email body.

> ⓘ Note

If you configure a sender email alias in Exchange Admin Center mailboxes, it overwrites the **sender display name** that you configure here.

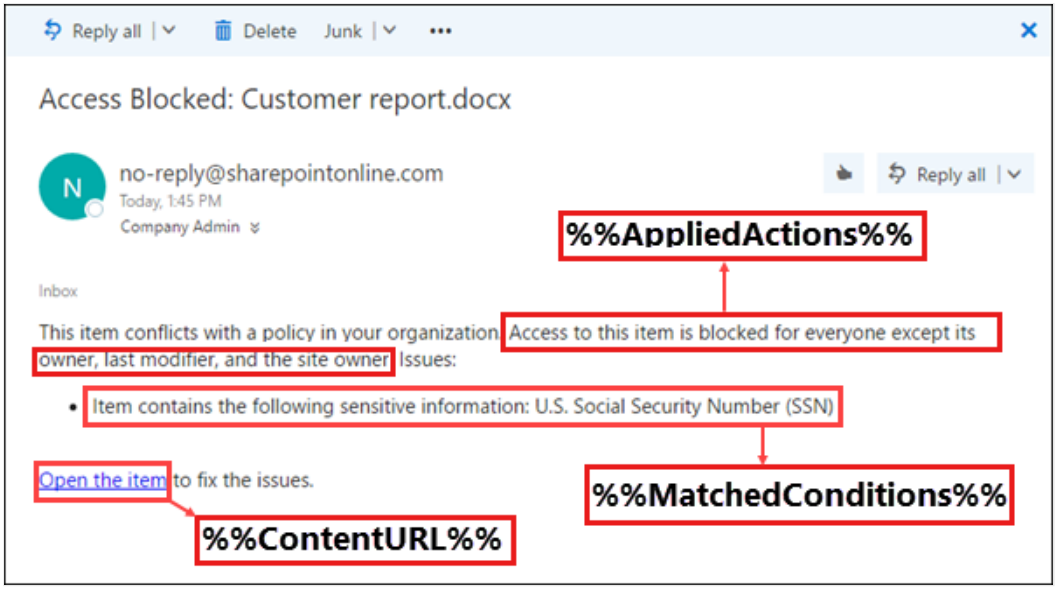
You can also use the following tokens to help customize the body of the email. Tokens are variables that get replaced by specific information when the notification is sent. For example, the %%ContentURL%% token is replaced by the URL of the document on the SharePoint site or OneDrive site.

Expand table

Token	Description	Available for Exchange	Available for SharePoint	Av for On
%%MatchedConditions%%	The conditions that the content matched. Use this token to inform people of possible issues with the content.	Yes	Yes	Yes
%%ContentURL%%	The URL of the document on the SharePoint site or OneDrive site.	No	Yes	Yes
%%AppliedActions%%	The actions applied to the content. This token is populated only when the action - **Restrict access or encrypt the content in Microsoft 365 locations** is selected in the DLP rule configuration	Yes	Yes	Yes
%%BlockedMessageInfo%%	The details of the message that was blocked. Use this token to inform people of the details of the message that was blocked. This token is populated only when the action - **Restrict access or encrypt the content in Microsoft 365 locations** is selected in the DLP rule configuration	Yes	No	No
%%ContentId%%	The unique identifier of the message.	Yes	No	No
%%TimestampForIncidentOccurrence%%	Timestamp in UTC of when the DLP policy conditions were matched.	Yes	Yes	Yes
%%MatchedConditionsAndValues%%	The matched DLP condition and values. This token doesn't cover the *content contains sensitive info* condition. For matched SITs and redacted values, see %%MatchedSITAndSurroundingcontext%%	Yes	Yes	Yes
%%Filename%%	For SharePoint and OneDrive matches, this token shows the document name. For Exchange, it shows the email attachment name.	Yes	Yes	Yes
%%PolicyName%%	The matched DLP policy name.	Yes	Yes	Yes
%%PolicyRule%%	The matched DLP rule name.	Yes	Yes	Yes

Token	Description	Available for Exchange	Available for SharePoint	Av for On
%%Workload%%	The workload name where the match occurred.	Yes	Yes	Yes
%%MatchedSITAndSurroundingcontext%%	The matched SITs and the redacted values.	Yes	Yes	Yes
%%UserEmail%%	The email address of the end user associated with the matched content.	Yes	Yes	Yes
%%SiteAdmin%%	For SharePoint sites, this token shows the email address of the site administrator.	No	Yes	No
%%EmailSubject%%	For Exchange, this token shows the email subject.	Yes	No	No
%%EmailRecipients%%	For Exchange, this token shows the list of email recipients.	Yes	No	No
%%FileOwner%%	For SharePoint and OneDrive matches, this token shows the name of file owner.	No	Yes	Yes

> ⓘ Note
>
> Use the HTML tag to style the tokens. For example, <div\ "style="color:blue; font-size: 12px;"> %%MatchedConditions%% renders the token in font size 12 px and font color blue.



Actionable email notifications for OneDrive and SharePoint workloads

> ⓘ Note

Actionable email notifications for OneDrive and SharePoint workloads are in preview.

Admins can add controls to customized email notifications. These controls enable users to fix the issue that caused the DLP policy match or otherwise remediate issues on OneDrive folders or SharePoint sites directly from the email. This streamlines the remediation process. The controls are:

- **Stop sharing file**: Removes any sharing access and links on the file.
- **Delete file**: Deletes the file.
- **Apply sensitivity or retention label on the file**: Opens the OneDrive or SharePoint folder or site where the file is located so the user can apply the sensitivity or the retention label.
- **Override the policy**: Requires the end user to add business justification and overrides the policy.
- **Report false positive**: Requires the end user to enter the justification and notifies the admin via audit log.
- **Report unable to take action**: Requires the end user to enter the justification and notifies the admin via audit log.

To make these options available to your users in notification emails, select them when you configure the email notification.

> ⓘ **Note**
>
> Only markdown mode can be used to format the email body if you add actions to the email.

Actions details in unified audit logs:

- Delete, Stop share, Apply label: Record types = **SharepointFileOperation**
- Override, Report false positive, Report unable to take action: Activities - operation names = **DLPInfo**, Record types = **ComplianceDLPSharePoint**

Options for configuring policy tips

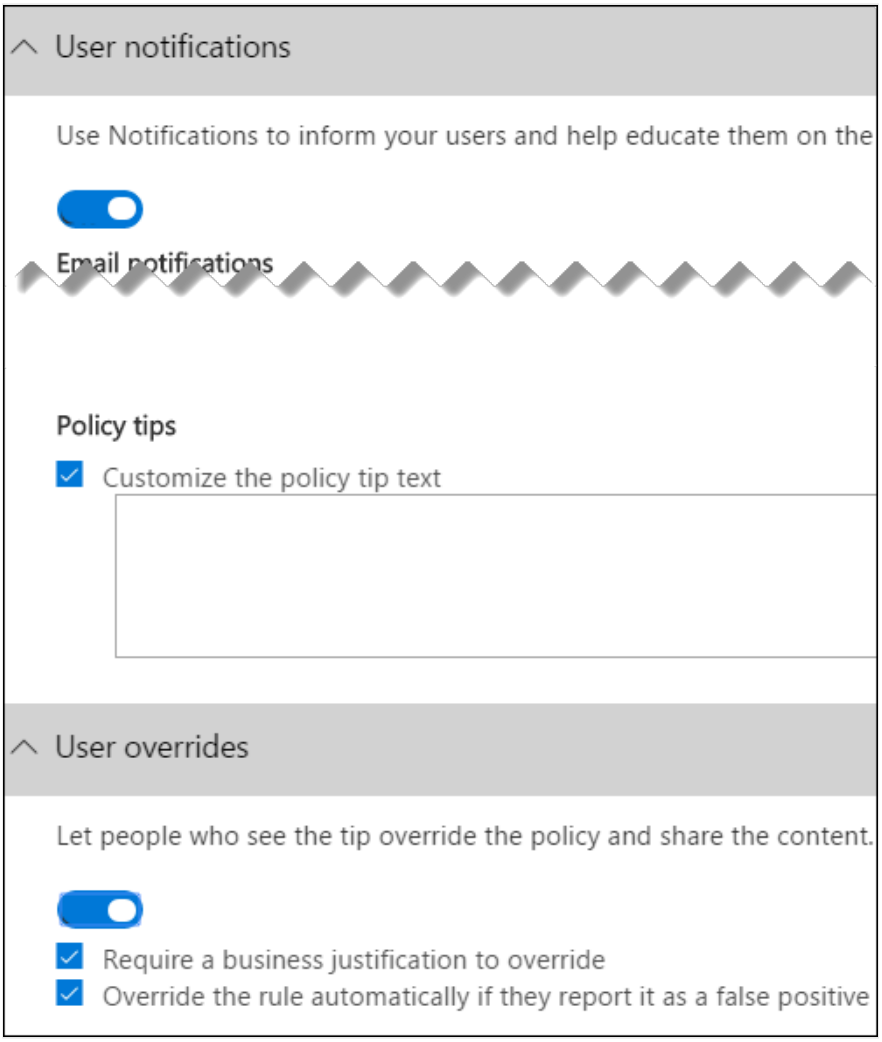
For each rule in a DLP policy, you can configure policy tips to:

- Notify the person that the content conflicts with a DLP policy, so they can take action to resolve the conflict. You can use the default text (see the tables below) or enter custom text about your organization's specific policies.

- Allow the person to override the DLP policy. Optionally, you can:

    - Require the person to enter a business justification for overriding the policy. This information is logged and you can view it in the DLP reports in the **Reports** section of the portal.

- Allow the person to report a false positive and override the DLP policy. This information is also logged for reporting, so that you can use false positives to fine tune your rules.



For example, say you have a DLP policy applied to OneDrive sites that detects personal data, and this policy has three rules:

1. First rule: If fewer than five instances of this sensitive information are detected in a document, and the document is shared with people inside the organization, the **Send a notification** action displays a policy tip. For policy tips, no override options are necessary because this rule is simply notifying people and not blocking access.

2. Second rule: If more than five instances of this sensitive information are detected in a document, and the document is shared with people inside the organization, the **Block access to content** action restricts the permissions for the file, and the **Send a notification** action allows people to override the actions in this rule by providing a business justification. Your organization's business sometimes requires internal people to share personal data, and you don't want your DLP policy to block this work.

3. Third rule: If more than five instances of this sensitive information are detected in a document, and the document is shared with people outside the organization, the **Block access to content** action restricts the permissions for the file, and the **Send a notification** action doesn't allow people to override the actions in this rule because the information is

shared externally. Under no circumstances should people in your organization be allowed to share personal data outside the organization.

User override support

The override option is per rule, and it overrides all of the actions in the rule.

Content can match several rules in a DLP policy or several different DLP policies, but only the policy tip from the most restrictive rule that has the highest priority is shown. This includes policies that are in simulation mode. This setup prevents people from seeing a cascade of policy tips.

For example, if a document matches three rules in a DLP policy, and the rules have these configurations, rule LMN is applied and the policy tip from rule LMN is shown to the user.

Expand table

Rule name	Rule priority	Rule action	Rule action applied
ABC	1	Send notification	No
LMN	2	Block access and send notification	Yes
XYZ	3	Block	No

If the policy tips in the most restrictive rule allow people to override the rule, then overriding this rule also overrides any other rules that the content matched.

If you set the NotifyAllowOverride action to WithoutJustification or WithJustification or FalsePositives, make sure BlockAccess is set to true and BlockAccessScope has appropriate value. Otherwise, the policy tip comes up but the user doesn't find an option to override the email with justification.

To see overrides in policy tips in Outlook on the web, the policy must be set to the *Turn it on* state. The policy action must also be configured to block with override.

Availability of override

The following table shows when override is available for each notification and block-action combination.

Expand table

Notification Rule	Notify/Block action	Override available	Require Justification
Notify only	Notify	No	No
Notify + AllowOverride	Notify	No	No

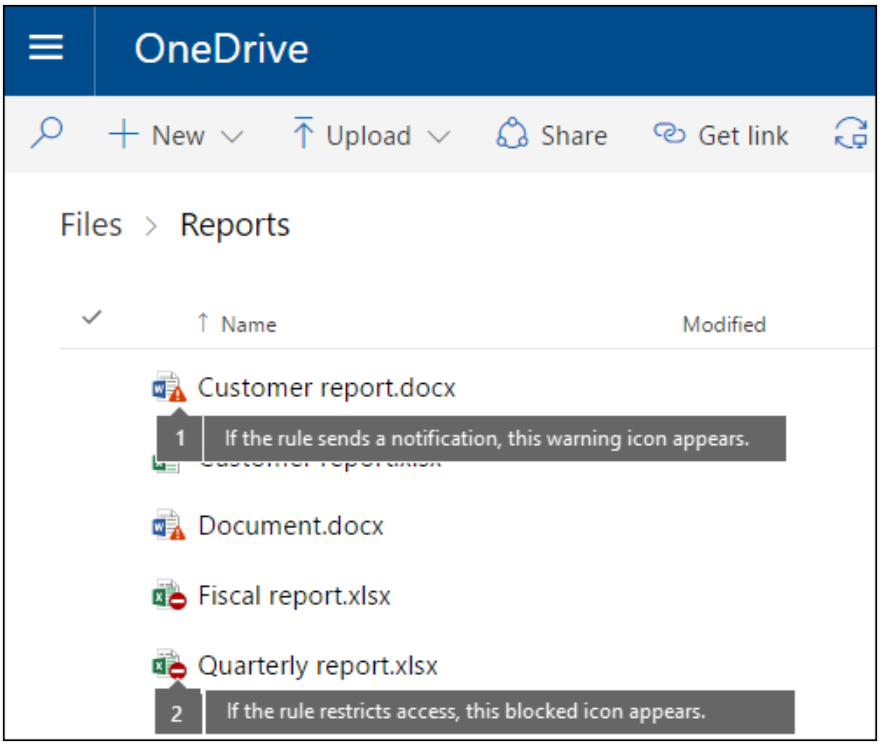
Notification Rule	Notify/Block action	Override available	Require Justification
Notify + AllowOverride + False positive	Notify	No	No
Notify + AllowOverride + With justification	Notify	No	No
Notify + AllowOverride + False positive + Without justification	Notify	No	No
Notify + AllowOverride + False positive + With justification	Notify	No	No
Notify + Block	Block	No	No
Notify + Block + AllowOverride	Block	Yes	No
Notify + Block + AllowOverride + False positive	Block	Yes	No
Notify + Block + AllowOverride + With justification	Block	Yes	Yes
Notify + Block + AllowOverride + False positive + Without justification	Block	Yes	No
Notify + Block + AllowOverride + False positive + With justification	Block	Yes	Yes

Policy tips on OneDrive sites and SharePoint sites

When a document on a OneDrive site or SharePoint site matches a rule in a DLP policy that uses policy tips, special icons appear on the document:

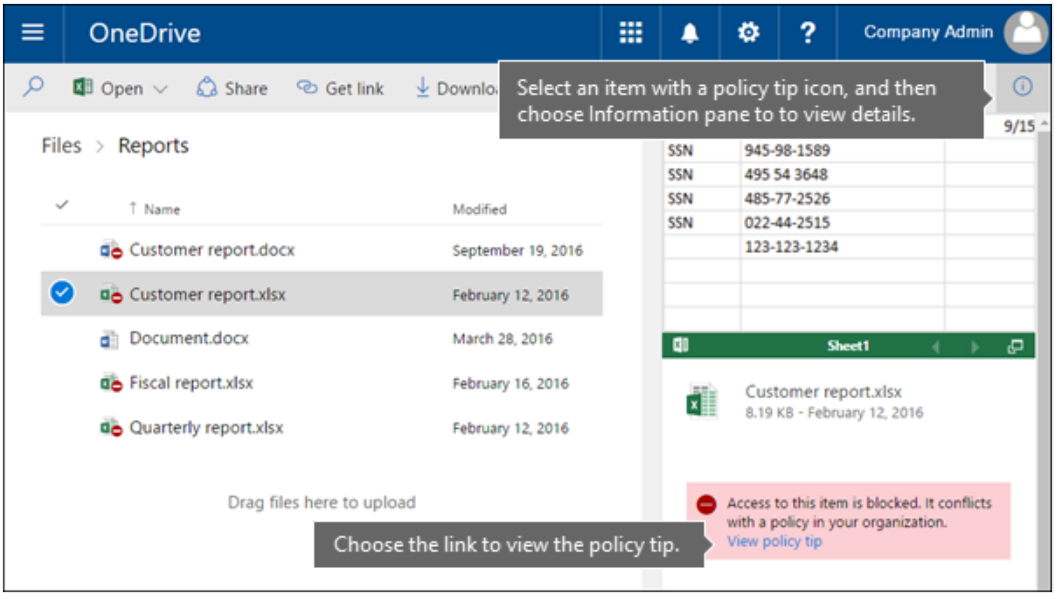
1. If the rule sends a notification about the file, the warning icon appears.

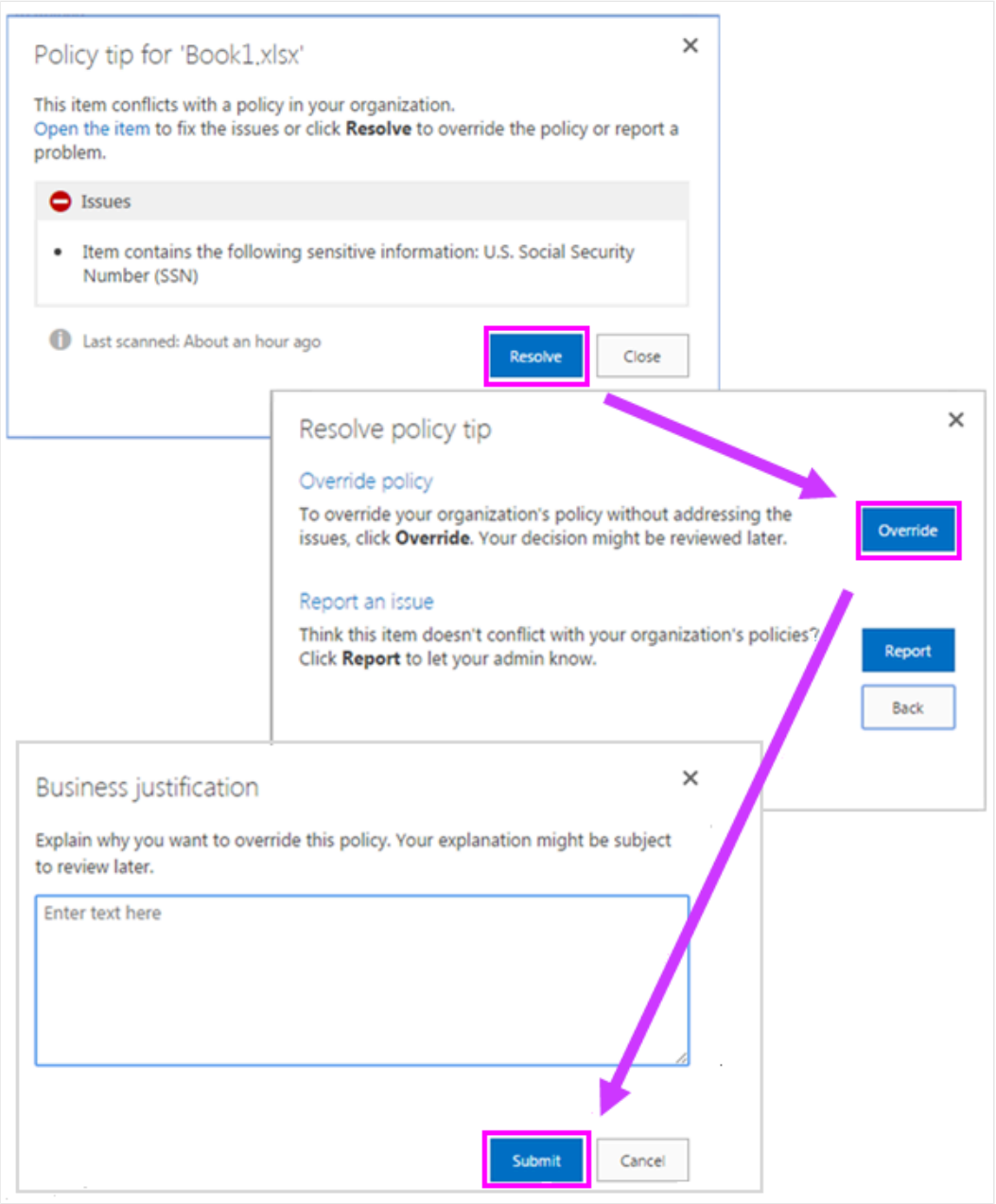
2. If the rule blocks access to the document, the blocked icon appears.



To take action on a document, select an item, choose **Information** in the upper-right corner of the page to open the details pane, then select **View policy tip**.

The policy tip lists the issues with the content. If the policy tips are configured with these options, you can choose **Resolve**, then **Override** the policy tip or **Report** a false positive.





DLP policies sync to sites and content is evaluated against them periodically and asynchronously, so there might be a short delay between the time you create the DLP policy and the time you begin to see policy tips. There might be a similar delay from when you resolve or override a policy tip to when the icon on the document on the site goes away.

Default text for policy tips on sites

By default, policy tips display text similar to the following for an item on a site. The notification text is configured separately for each rule, so the text that displays differs depending on which rule is matched.

Expand table

If the DLP policy rule does this...	Then the default policy tip says this...
Sends a notification but	This item conflicts with a policy in your organization.

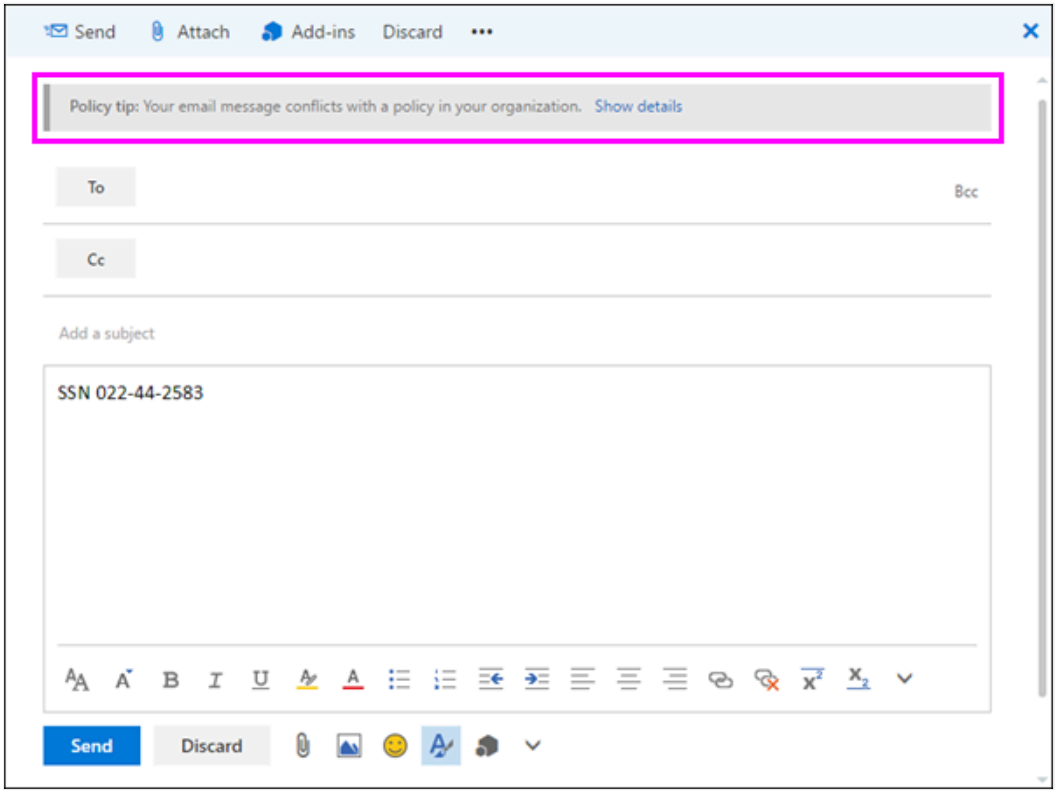
If the DLP policy rule does this...	Then the default policy tip says this...
doesn't allow override	
Blocks access, sends a notification, and allows override	This item conflicts with a policy in your organization. If you don't resolve this conflict, access to this file might be blocked.
Blocks access and sends a notification	This item conflicts with a policy in your organization. Access to this item is blocked for everyone except its owner, last modifier, and the primary site admin.

Custom text for policy tips on sites

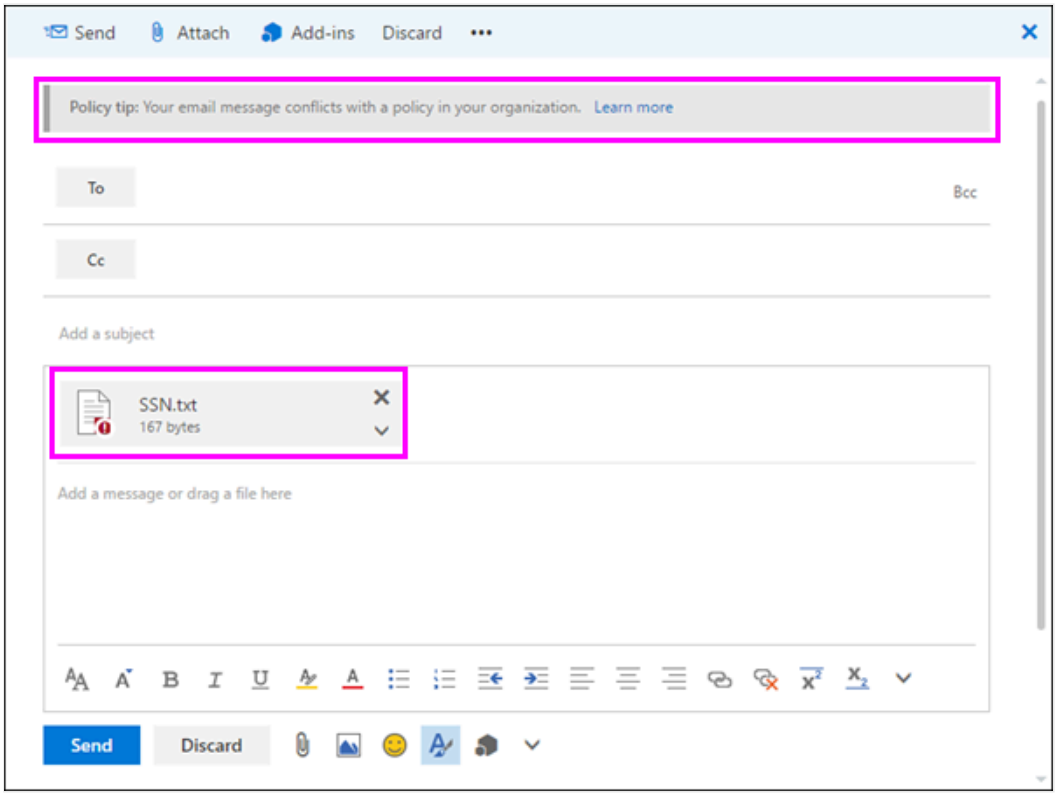
You can customize the text for policy tips separately from the email notification. Unlike custom text for email notifications (described in Custom email notifications), custom text for policy tips doesn't accept HTML or tokens. Instead, custom text for policy tips is plain text only with a 256-character limit.

Policy tips in Outlook on the web and Outlook 2013 and later

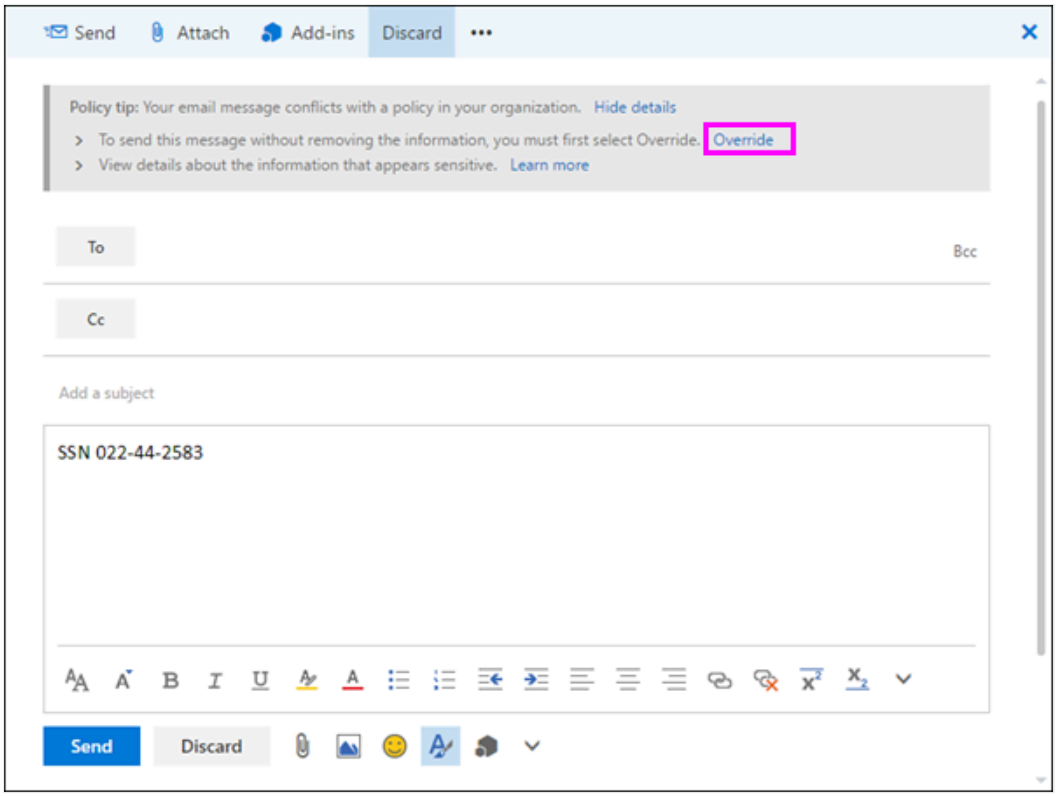
When you compose a new email in Outlook on the web and Outlook 2013 and later, you see a policy tip if you add content that matches a rule in a DLP policy, and that rule uses policy tips. The policy tip appears at the top of the message, above the recipients, while the message is being composed.

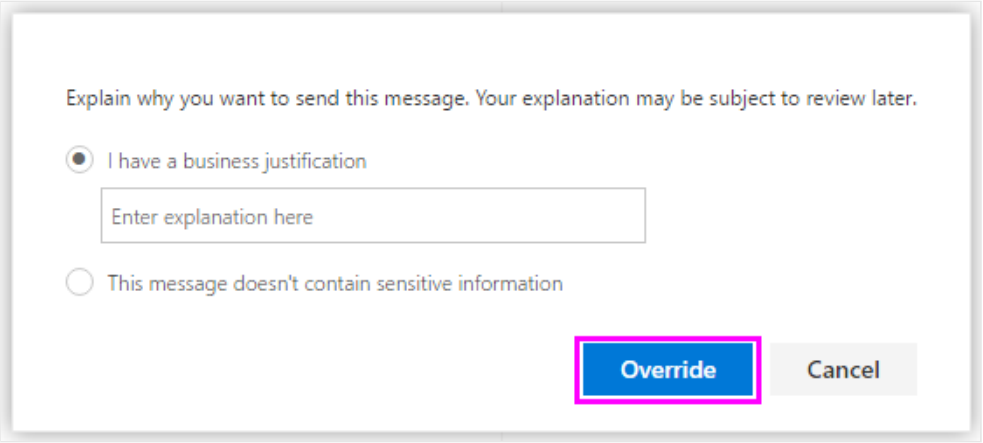


Policy tips work whether the sensitive information appears in the message body, subject line, or even a message attachment as shown here.



If the policy tips are configured to allow override, you can choose **Show Details** > **Override** > enter a business justification or report a false positive > **Override**.





When you add sensitive information to an email, there might be latency between when you add the sensitive information and when the policy tip appears. When emails are encrypted with Microsoft Purview Message Encryption and the policy used to detect them uses the detected encryption condition, policy tips don't appear.

Policy tips in the Exchange admin center vs. the Microsoft Purview portal

Policy tips can work either with DLP policies and mail flow rules created in the Exchange admin center, or with DLP policies, but not both. This limitation exists because these policies are stored in different locations, but policy tips can draw only from a single location.

If you configure policy tips in the Exchange admin center, any policy tips that you configure in the Purview portal don't appear to users in Outlook on the web and Outlook 2013 and later until you turn off the tips in the Exchange admin center. This behavior ensures that your current Exchange mail flow rules (also known as transport rules) continue to work until you choose to switch over to the Purview portal.

While policy tips can draw only from a single location, email notifications are always sent, even if you're using DLP policies in both the Purview portal and the Exchange admin center.

Default text for policy tips in email

By default, policy tips display text similar to the following text for email.

Expand table

If the DLP policy rule does this...	Then the default policy tip says this...
Sends a notification but doesn't allow override	Your email conflicts with a policy in your organization.
Blocks access, sends a notification, and allows override	Your email conflicts with a policy in your organization.
Blocks access and sends a notification	Your email conflicts with a policy in your organization.

Policy tips in Excel, PowerPoint, and Word

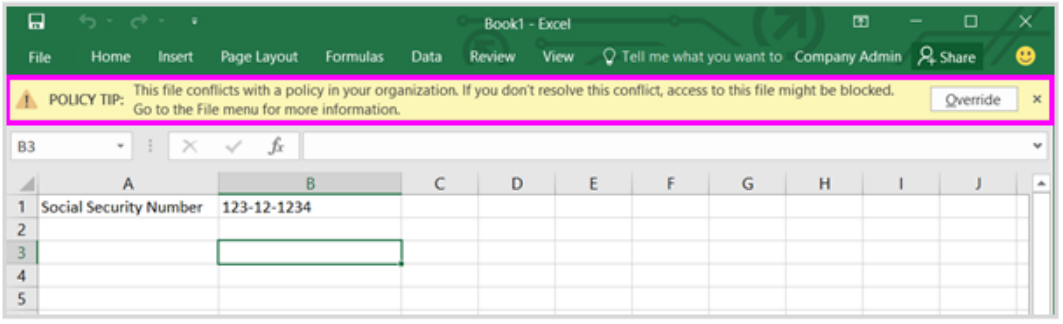
When people work with sensitive content in the desktop versions of Excel, PowerPoint, and Word, policy tips notify them that the content conflicts with a DLP policy. This feature requires that:

- The Office document is stored on a OneDrive site or SharePoint site.
- The site is included in a DLP policy that's configured to use policy tips.

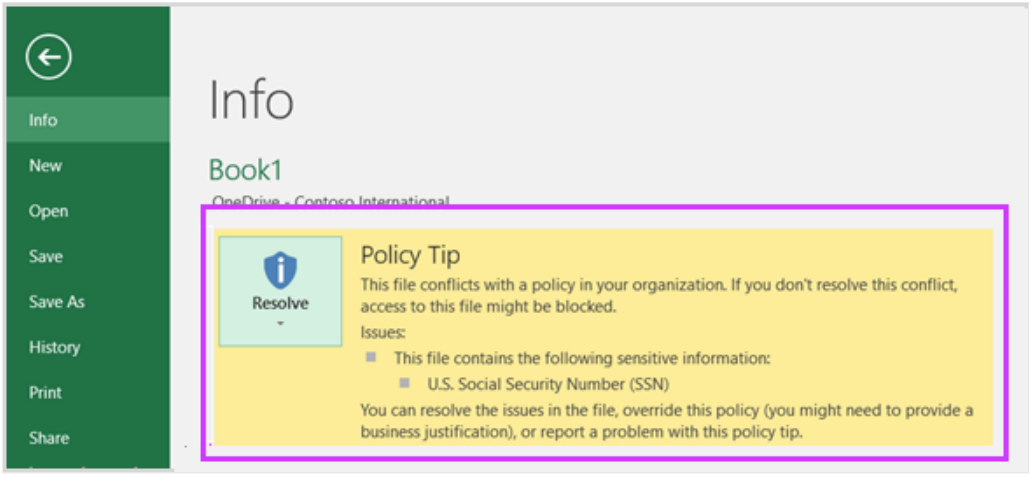
Office desktop programs automatically sync DLP policies directly from Microsoft 365, and then scan your documents to ensure that they don't conflict with your DLP policies.

Depending on how you configure the policy tips in the DLP policy, people can choose to ignore the policy tip, override the policy with or without a business justification, or report a false positive.

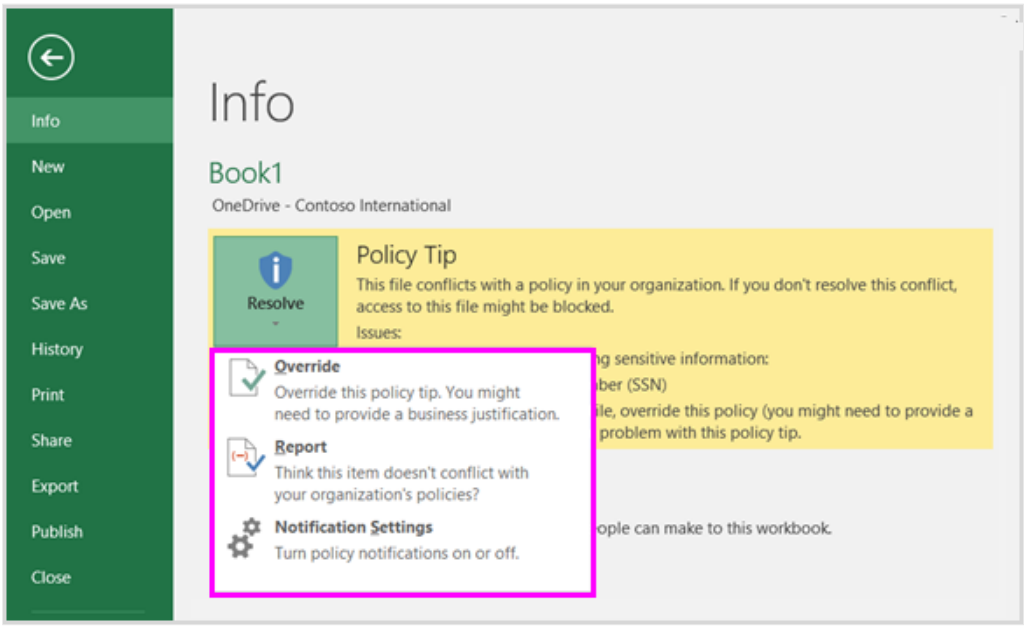
Policy tips appear on the Message Bar.



Policy tips also appear in the Backstage view (on the **File** tab).



If you configure policy tips in the DLP policy with these options, you can choose **Resolve** to **Override** a policy tip or **Report** a false positive.



In each of these Office desktop programs, people can choose to turn off policy tips. If turned off, policy tips that are simple notifications don't appear on the Message Bar or Backstage view (on the **File** tab). However, policy tips about blocking and overriding still appear, and the user still receives the email notification. In addition, turning off policy tips doesn't exempt the document from any DLP policies that you applied to it.

Default text for policy tips in Excel 2016, PowerPoint 2016, and Word 2016

By default, policy tips display text similar to the following text on the Message Bar and Backstage view of an open document. You configure the notification text separately for each rule, so the text that displays differs depending on which rule is matched.

Expand table

If the DLP policy rule does this...	Then the default policy tip says this...
Sends a notification but doesn't allow override	This file conflicts with a policy in your organization. Go to the **File** menu for more information.
Blocks access, sends a notification, and allows override	This file conflicts with a policy in your organization. If you don't resolve this conflict, access to this file might be blocked. Go to the **File** menu for more information.
Blocks access and sends a notification	This file conflicts with a policy in your organization. If you don't resolve this conflict, access to this file might be blocked. Go to the **File** menu for more information.

Custom text for policy tips in Excel, PowerPoint, and Word

You can customize the text for policy tips separately from the email notification. Unlike custom text for email notifications (described in Custom email notifications), custom text for policy tips

doesn't accept HTML or tokens. Instead, custom text for policy tips is plain text only with a 256-character limit.

Related content

- Learn about data loss prevention
- Create and deploy data loss prevention policies
- DLP policy conditions, exceptions, and actions (preview)
- What the DLP policy templates include
- Sensitive information type entity definitions

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

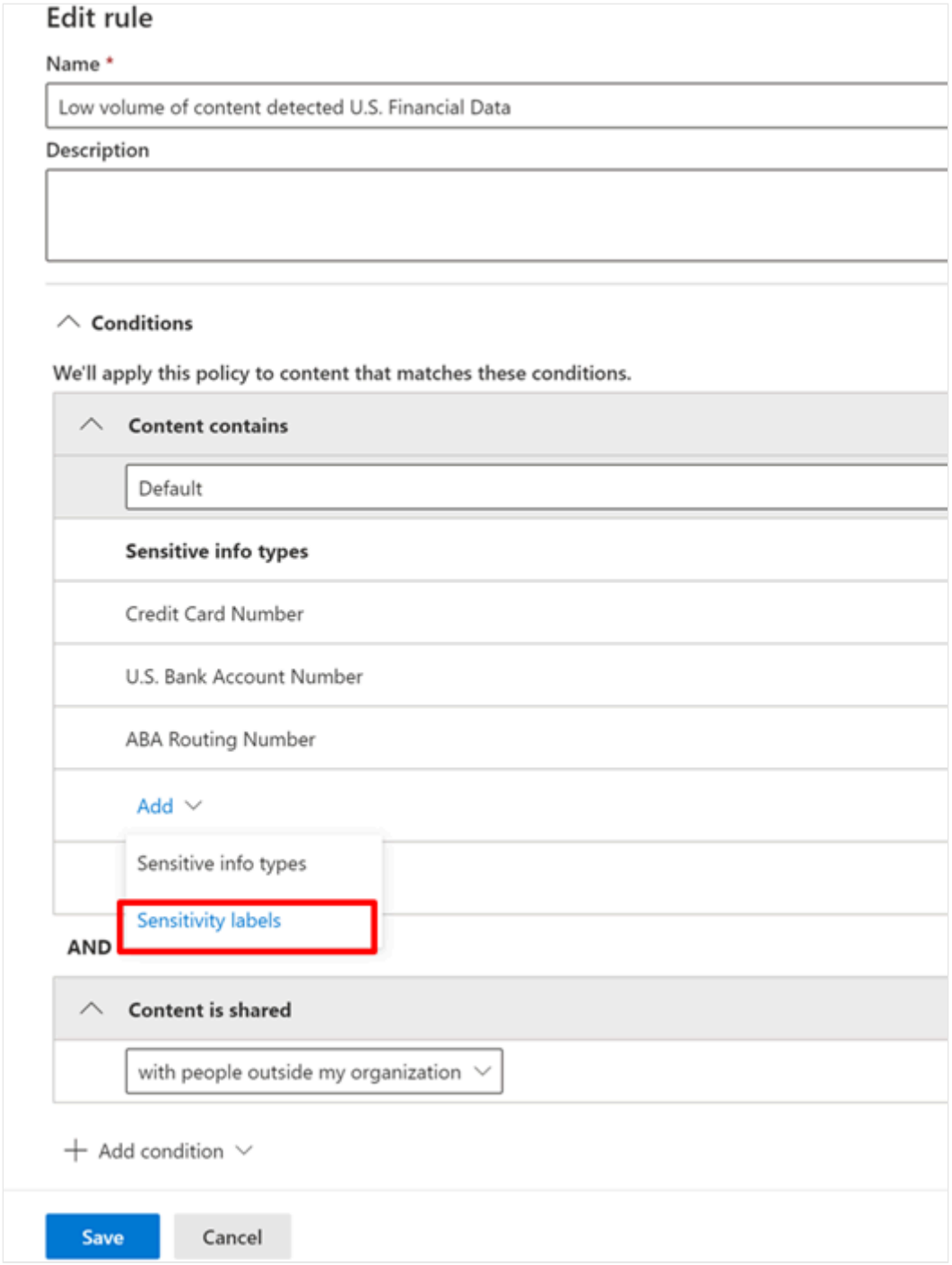
Last updated on 06/22/2026

Use sensitivity labels as conditions in DLP policies

You can use sensitivity labels as a condition in DLP policies for these locations:

- Exchange email messages
- SharePoint
- OneDrive
- Devices
- Instances
- On-premises repositories
- Microsoft 365 Copilot
- Fabric and Power BI workspaces
- Inline web traffic (Unmanaged cloud apps and adaptive app scopes)

Sensitivity labels appear as an option in the **Content contains** condition list.



> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Supported items, file types, scenarios, and policy tips

You can use sensitivity labels as conditions on these items and in the following scenarios.

Supported items

Expand table

Service	Item type	Available to policy tip	Enforceable	Notes
Exchange	email message	Yes	Yes	
Exchange	email attachment	Yes	Yes	
SharePoint	items in SharePoint	Yes	Yes	
OneDrive	items	Yes	Yes	
Teams	Teams and channel messages	Not applicable	Not applicable	
Teams	attachments	Yes	Yes	If SharePoint or OneDrive are included as locations in your DLP policy, you don't need to set Teams as a location. This is because attachments sent in Teams over 1:1 chat or channels are automatically uploaded to OneDrive and SharePoint and will be evaluated by DLP then.
Devices	items	Yes	Yes	
MCAS (preview)	items	Yes	Yes	
Inline web traffic	Uploaded and downloaded files	No	Yes	DLP policies don't support custom policy tips for inline web traffic.

> **Note**
>
> DLP's ability to detect sensitivity labels in SharePoint and OneDrive is limited. For more information, see Enable sensitivity labels for files in SharePoint and OneDrive.

Supported file types

Expand table

Workload	File types supported
Exchange emails	.docx, .xlsx, .pptx, .pdf, .pfile
SharePoint	.docx, .xlsx, .pptx, .pdf
OneDrive	.docx, .xlsx, .pptx, .pdf
endpoint devices	.docx, .xlsx, .pptx, .pdf, .pfile
Inline web traffic	.docx, .xlsx, .pptx, .pdf, .pfile, .eml - legacy Office file types, .doc, .xls, and .ppt, aren't supported.

Supported scenarios

- DLP Admin can see a list of all sensitivity labels in the tenant when they choose to include one or more sensitivity labels as a condition.
- All workloads support using sensitivity labels as a condition, as shown in the support matrix.
- DLP policy tips continue to show across workloads for DLP policies that contain one or more sensitivity labels as a condition.
- Sensitivity labels appear as part of the incident report email if a DLP policy with one or more sensitivity labels as a condition is matched.
- Sensitivity label details show in the DLP rule match audit log for DLP policy matches that contain a sensitivity label as a condition.
- For policies that apply to the Inline web traffic location, content that's encrypted can't be extracted for evaluation by DLP. The only conditions that work are metadata based ones, for example:
- **Content contains** > **Sensitivity label** returns the label name
- **File size** returns the size of the encrypted file
- **File extension is** returns the extension of the encrypted file Conditions that rely on the content being evaluated don't work, for example:
- **Content contains** > **Sensitive information type**
- **Content contains** > **Trainable classifiers**
- **File type is**

Support policy tips

Expand table

Workload	Policy tips supported/not supported
OWA	supported
Outlook for Windows	supported
SharePoint	supported
OneDrive	supported
endpoint devices	not supported
Inline web traffic	not supported

> **Note**
>
> If you use sensitivity labels as a condition and policy tips in a DLP policy for Exchange, and, at the same time, you have mandatory labeling for sensitivity labels configured, when a user selects a label through the Mandatory Labeling prompt, the client does not re-invoke the DLP Policy Tips evaluation gate with the newly applied label. In result, the policy tip is **not ** going to show, but the action configured in the DLP policy will be enforced.

Last updated on 08/17/2026

Learn about DLP file quarantine for SharePoint and OneDrive

> **Note**
>
> The features described in this article are in preview. Organizations sometimes need to go beyond blocking access to sensitive files. When a file contains highly sensitive content that shouldn't exist in a particular SharePoint or OneDrive location, blocking access alone might not be enough — the file itself needs to be removed from the location entirely. Data loss prevention (DLP) file quarantine for SharePoint and OneDrive gives administrators the ability to automatically move files that match a DLP policy rule to an admin-controlled quarantine location.

What is file quarantine?

File quarantine is a DLP policy action for SharePoint and OneDrive locations. When you configure a DLP policy with the quarantine action, files that match the policy rule are:

- Removed from the original SharePoint or OneDrive location.
- Moved to an admin-defined quarantine site in SharePoint.
- Replaced at the original location with a tombstone text file that contains an admin-configured message.

File quarantine is the most restrictive DLP action available for SharePoint and OneDrive. It removes access from everyone, including the file owner.

When to use file quarantine

Use file quarantine when your organization needs to:

- **Isolate highly sensitive content** — For example, a file containing confidential business data, such as revenue numbers for an unannounced project, is stored in a SharePoint location where it shouldn't be. The file needs to be removed from that location and placed somewhere that only authorized administrators can access.
- **Remove access from the file owner** — In some cases, the file owner shouldn't retain access to the content. For example, the file was placed in an insecure location and the owner needs

to work with an administrator to find a more secure way to store and share it.
- **Provide stronger protection than blocking alone** — Blocking prevents sharing but leaves the file in place. Quarantine removes the file entirely from the original location.

How file quarantine works

When a file in a SharePoint or OneDrive location matches a DLP policy rule that has the quarantine action configured, the following sequence occurs:

1. **Permission removal** — All existing permissions and sharing links on the file are removed.
2. **File move** — A system account moves the file from the original location to the configured admin quarantine site. The folder structure of the original location is preserved in the quarantine site.
3. **Tombstone file creation** — A text (.txt) file with the same name as the original file replaces it at the source location. The tombstone file displays an admin-configured message and includes the relative path of the quarantined file. The full quarantine folder path isn't disclosed to the file owner.
4. **Audit and alerting** — An audit record is created that includes the quarantine file location. A DLP rule match event appears in Activity Explorer, and DLP alerts include the quarantine file location, the file owner, and the original file path.

File identity

Each file in SharePoint and OneDrive is uniquely identified by its doc ID, not by its file name. DLP uses the doc ID to track files through the quarantine and restore process.

Quarantine site exclusion

Once you designate a SharePoint site as the quarantine location, the contents of that site are automatically excluded from DLP rule evaluation. You don't need to manually add the quarantine site to an exclusion list. This automatic exclusion prevents conflicts where a quarantined file could match another DLP rule and create circular enforcement behavior.

Files eligible for quarantine

The quarantine action only applies to files that are created or modified after the policy with the quarantine action is turned on. Files that existed before the policy was activated aren't evaluated for the quarantine action. This design prevents a situation where a broadly scoped policy

accidentally quarantines a large number of existing files, which would be difficult to reverse because file restore is a manual process.

> **Important**
>
> This behavior is specific to the quarantine action. Other DLP actions, such as block, apply to all files in scope, regardless of when they were created. If your organization uses on-demand classification, files that are scanned and classified by an on-demand scan can trigger the quarantine action if the classification result matches a DLP policy rule for the first time.

> **Note**
>
> The quarantine action is not available as an on-demand classification action. On-demand classification only scans and classifies files — it doesn't take enforcement actions.

Supported file types

File quarantine supports all file types that DLP supports for SharePoint and OneDrive locations.

Tombstone files

When a file is quarantined, a tombstone text file replaces it at the original location. The tombstone file:

- Uses the original file's name with a `.txt` extension.
- Displays an admin-configured message. Administrators configure this message in DLP settings. For example, the message might instruct the file owner to contact a specific email address for assistance.
- Includes the relative path of the quarantined file within the quarantine site. The full quarantine folder path isn't provided, because the quarantine location is a confidential admin-controlled site.

The file owner can view the tombstone file to understand why their file was removed and who to contact.

Manual file restore

There's no automated file restore for quarantined files. All file restore is a manual process. Administrators decide when and whether to restore a quarantined file.

Admins might restore files when:

- The sensitive content is determined to be harmless in context.
- The file usage is approved for valid business reasons.
- The policy match is a false positive.

To restore a quarantined file:

1. Locate the quarantined file in the quarantine site.
2. Identify the original file path by using audit logs.
3. Move the file from the quarantine site back to the original location. The quarantine site admin must have access to the original location.
4. Delete the tombstone file from the original location.

Post-restore behavior

After a file is restored:

- **Original sharing permissions aren't restored.** The file owner or admin must reconfigure sharing permissions.
- **Only the latest file version is restored.** Previous versions aren't preserved through the quarantine and restore process.
- **The same DLP rule won't re-quarantine the restored file**, even if the file is modified or the policy rule is updated. This behavior is intentional to prevent quarantine loops. Other DLP rules continue to apply normally to the restored file.

Known limitations

- **No automated restore** — All file restore is manual. There's no restore action available in the Microsoft Purview portal.
- **Sharing permissions not preserved** — Original sharing permissions and links aren't retained through the quarantine and restore cycle.
- A maximum of **200,000 items per tenant within a 24-hour period** may be processed for quarantine. Items exceeding this limit may be processed at a later time.

Relationship to other quarantine features

DLP file quarantine for SharePoint and OneDrive is separate from these related features:

- **Endpoint DLP auto-quarantine** — Applies only to files on Windows and macOS devices. It moves files to a local quarantine folder on the device. SharePoint and OneDrive file quarantine operates on cloud-hosted files.
- **Microsoft Defender for Cloud Apps file quarantine** — A similar quarantine capability that exists in Microsoft Defender for Cloud Apps (MDA). DLP file quarantine for SharePoint and OneDrive is designed for organizations that want to manage quarantine directly from Microsoft Purview.

See also

- Get started with DLP file quarantine for SharePoint and OneDrive
- Create a DLP policy to quarantine files in SharePoint and OneDrive
- Learn about data loss prevention
- Create and deploy data loss prevention policies
- [Learn about investigating data loss prevention alerts](dlp-alert-investigation-learn.md

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 07/17/2026

Help prevent sharing credit card numbers through email

This section shows how to create a Microsoft Purview Data Loss Prevention (DLP) policy that blocks emails containing credit card numbers. It follows the process described in Design a data loss prevention policy. Work through this scenario in your test environment to familiarize yourself with the policy creation UI.

> **Important**
>
> This article presents a hypothetical scenario with hypothetical values. It's only for illustrative purposes. Substitute your own sensitive information types, sensitivity labels, distribution groups, and users.

How you deploy a policy is as important as policy design. This article shows you how to use the deployment options so that the policy achieves your intent while avoiding costly business disruptions.

Prerequisites and assumptions

This scenario uses the *Highly confidential* sensitivity label, so it requires that you create and publish sensitivity labels. To learn more, see:

- Learn about sensitivity labels
- Get started with sensitivity labels
- Create and configure sensitivity labels and their policies

This procedure uses a hypothetical distribution group *Finance team* at Contoso.com and a hypothetical SMTP recipient *adele.vance@fabrikam.com*.

This procedure uses alerts, see: Get started with the data loss prevention alerts

Policy intent statement and mapping

We need to block emails to all recipients that contain credit card numbers or that have the 'highly confidential' sensitivity label applied except if the email is sent from someone on the finance team to adele.vance@fabrikam.com. We want to notify the compliance admin every time an email is

blocked and notify the user who sent the item and no one can be allowed to override the block. Track all occurrences of this high risk event in the log and we want the details of any events captured and made available for investigation

Expand table

Statement	Configuration question answered and configuration mapping
"We need to block emails to all recipients..."	- Where to monitor: **Enterprise applications & devices** and **Exchange** - Administrative scope: **Full directory** - Action: **Restrict access or encrypt the content in Microsoft 365 locations** > **Block users from receiving email or accessing shared SharePoint, OneDrive, and Teams files** > **Block everyone**
"...that contain credit card numbers or have the 'highly confidential' sensitivity label applied..."	- What to monitor: use the **Custom template** - Conditions for match: edit it to add the *highly confidential* sensitivity label
"...except if..."	- Condition group configuration: Create a nested boolean NOT condition group joined to the first conditions using a boolean AND
"...the email is sent from someone on the finance team..."	- Condition for match: **Sender is a member of**
"...and..."	- Condition for match: add a second condition to the NOT group
"...to adele.vance@fabrikam.com..."	- Condition for match: **Recipient is**
"...Notify..."	- User notifications: **enabled** - Policy tips: **enabled**
"...the compliance admin every time an email is blocked and notify the user who sent the item..."	- Policy tips: **enabled** - Notify these people: **selected** - The person who sent, shared, or modified the content: **selected** - Send the email to these additional people: add the email address of the compliance administrator
"...and no one can be allowed to override the block...	- **Allow overrides from M365 Services**: not selected
"...Track all occurrences of this high risk event in the log and we want the details of any events captured and made available for investigation."	- Use this severity level in admin alerts and reports: **high** - Send an alert to admins when a rule match occurs: **selected**

Statement	Configuration question answered and configuration mapping
	- Send alert every time an activity matches the rule: **selected**

Steps to create the policy

> **Important**
>
> For this policy creation procedure, accept the default include and exclude values and leave the policy turned off. Change these values when you deploy the policy.

1. Sign in to the Microsoft Purview portal.
2. Open the **Data loss prevention** solution and go to **Policies** > **+ Create policy**.
3. Select **Enterprise applications & devices**.
4. Select **Custom** from the **Categories** list.
5. Select **Custom** from the **Regulations** list.
6. Select **Next**.
7. Enter a **Name** and a **Description** for the policy. You can use the policy intent statement here.

   > **Important**
   >
   > You can't rename policies.

8. Select **Next**.
9. **Assign admin units**. To apply the policy to all users, accept the default setting.
10. Select **Next**.
11. Choose where to apply the policy. Select only the **Exchange email** location. Deselect all the other locations.
12. Select **Next**.

13. On the **Define policy settings** page, the **Create or customize advanced DLP rules** option should already be selected.
14. Select **Next**.
15. Select **Create rule**. Name the rule and provide a description.
16. Under **Conditions**, select **Add condition** > **Content contains**.
17. (Optional) Enter a **Group name**.
18. (Optional) Select a **Group operator**.
19. Select **Add** > **Sensitive info types** > **Credit Card Number**.
20. Select **Add**.
21. Still within the **Content contains** section, select **Add** > **Sensitivity labels** > **Highly confidential** and then select **Add**.
22. Next, beneath the **Content contains** section, select **Add group**.
23. Leave the Boolean operator set to **AND**, then set the toggle to **NOT**.
24. Select **Add condition**.
25. Select **Sender is a member of**.
26. Select **Add or remove distribution groups**.
27. Select **Finance Team** and then select **Add**.
28. Select **Add condition** > **Recipient is**.
29. In the email field, enter *adele.vance@fabrikam.com* and select **Add** .
30. Under **Actions**, select **Add an action** > **Restrict access or encrypt the content in Microsoft 365 locations**.
31. Select **Block users from receiving email or accessing shared SharePoint, OneDrive, and Teams files**, then select **Block everyone**.
32. Set the **User notifications** toggle to **On**.
33. Select **Email notifications** > **Notify the person who sent, shared, or last modified the content**.

34. Choose whether or not to **Attach matching email message to the notification**.
35. Choose whether or not to add **Policy tips**.
36. Under **User ovverides**, make sure that **Allow overrides from Microsoft 365 apps and services ...** is *NOT* selected.
37. Under **Incident reports**, set **Use this severity level in admin alerts and reports** to **High**.
38. Set **Send alert every time an activity matches the rule** toggle to **On**.
39. Select **Save**.
40. Select **Next**, then select **Run the policy in simulation mode**.
41. Select **Next** and then select **Submit**.
42. Select **Done**.

See also

- Create and deploy a data loss prevention policy

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Configure default protection for classification failures in Data loss prevention for Exchange (preview)

Microsoft Purview data loss prevention (DLP) policies for Exchange Online can detect and act on classification failures—cases where the classification service can't fully scan a document. This article shows you how to enable classification failure detection and use the improved conditions to apply protection actions for specific failure types, such as timeouts and throttling.

Enable classification failure detection

To enable classification failure detection, you can update the setting in the Purview portal or use PowerShell. Allow up to one hour for the setting change to synchronize across the service before validating policy behavior.

> **Important**
>
> This setting applies to Exchange Online only. It doesn't affect SharePoint, OneDrive, Teams, or Endpoint DLP policies.

- **Purview Portal**: To access the setting, from the Microsoft Purview portal, go to **Settings** on the left navigation menu > **Data loss prevention** > **Scan failure settings for Exchange**. In **Scan failure settings for Exchange** turn on the toggle.
- **PowerShell**: Connect to Security & Compliance PowerShell and run the following commands.

```powershell
Connect-IPPSSession
$json = '{"Classification":{"State":1}}'  # To disable, set State to 0
Set-PolicyConfig -DlpErrorHandlingConfig $json
```

Improved conditions for file attachments

By default, these conditions detect only text extraction failures. After you enable classification failure detection, they also detect classification failures such as timeouts, throttling, and other errors.

The following table compares each condition's capabilities before and after you enable classification failure detection.

Expand table

Condition	Default capabilities	Capabilities after you enable detection
Document couldn't be scanned	Text extraction failures (for example, the file couldn't be opened, the format is unsupported, or the file is password protected).	All text extraction failures, plus Classification failures (for example, when the classification service can't complete scanning of a document because of a timeout, throttling, or another error).
Document didn't complete scanning	Partial text extraction (for example, content exceeded extraction limits).	Partial text extraction, plus partial classification failures—for example, when at least one classifier (sensitive information type, exact data match, or trainable classifier) succeeded, but one or more of the others failed because of a timeout, throttling, or another error.

> **Note**
>
> If your tenant already has rules that use these conditions, you might see an increase in DLP rule matches after you enable detection.

To take different actions for specific classification failure types—**Timeout**, **Throttled**, or **Other**—use the **DocumentScanFailures** condition. Combine it with either **Document couldn't be scanned** or **Document didn't complete scanning** using AND; **DocumentScanFailures** can't be used on its own.

For the complete list of attachment conditions, see Attachments.

Rule ordering

Scan-failure rules detect failures only from classifiers that DLP has already evaluated. If a content-detection rule is placed after a scan-failure rule and its classifier fails, the failure won't be detected.

To avoid missed detections, order your rules so that content-detection rules (sensitive information types, exact data matches, and trainable classifiers) have higher priority than scan-failure rules (Timeout, Throttled, Other):

- Content rules at priorities 1–3.
- Scan-failure rules at priorities 4–6.

The following table shows an example that follows this pattern.

Expand table

Priority	Rule	Condition	Action
1	Detect credit cards	Content contains Credit Card Number	Block
2	Detect account numbers	Content contains Account Number	Block
3	Detect EDM match	Content contains EDM SIT	Block
4	Full scan timeout	DocumentIsUnsupported AND DocumentScanFailures = Timeout	Audit
5	Partial scan throttling	ProcessingLimitExceeded AND DocumentScanFailures = Throttled	Audit
6	Other scan errors	DocumentIsUnsupported AND DocumentScanFailures = Other	Audit

See Also

- Learn about data loss prevention
- Data Loss Prevention policy reference
- DLP Exchange conditions and actions reference

Last updated on 07/17/2026

Help prevent sharing sensitive items via SharePoint and OneDrive with external users

This article uses the process you learned in Design a data loss prevention policy to show you how to create a Microsoft Purview Data Loss Prevention (DLP) policy that helps prevent users from sharing sensitive items via SharePoint and OneDrive with external users. Work through these scenarios in your test environment to familiarize yourself with the policy creation UI.

For SharePoint and OneDrive, you create a policy to help prevent sharing of sensitive items with external users via SharePoint and OneDrive.

> **Important**
>
> This article presents a hypothetical scenario with hypothetical values. It's only for illustrative purposes. Substitute your own sensitive information types, sensitivity labels, distribution groups, and users.

How you deploy a policy is as important policy design. This article shows you how to use the deployment options so that the policy achieves your intent while avoiding costly business disruptions.

Prerequisites and assumptions

This scenario uses the *Confidential* sensitivity label, so it requires you to create and publish sensitivity labels. To learn more, see:

- Learn about sensitivity labels
- Get started with sensitivity labels
- Create and configure sensitivity labels and their policies

This procedure uses a hypothetical distribution group *Human Resources* and a distribution group for the security team at Contoso.com.

The policy creation procedure in this article uses alerts. To learn more, see Get started with the data loss prevention alerts

Policy intent statement and mapping

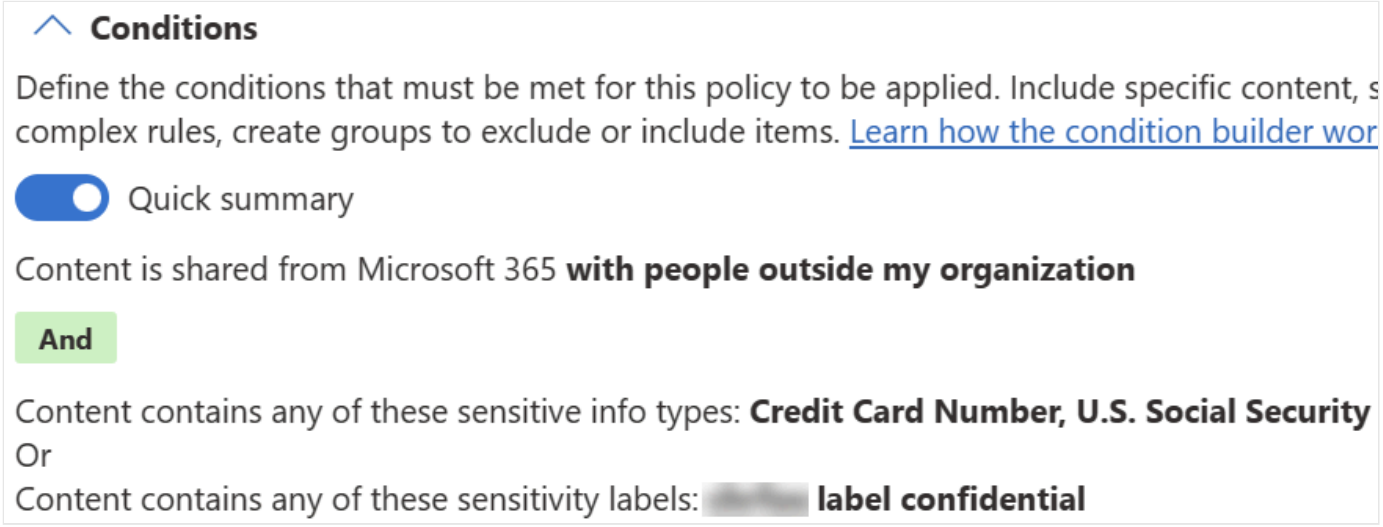
We need to block all sharing of SharePoint and OneDrive items to all external recipients that contain social security numbers, credit card data or have the "Confidential" sensitivity label. We don't want this to apply to anyone on the Human Resources team. We also have to meet alerting requirements. We want to notify our security team with an email every time a file is shared and then blocked. In addition, we want the user to be alerted via email and within the interface if possible. Lastly, we don't want any exceptions to the policy and need to be able to see this activity within the system.

Expand table

Statement	Configuration question answered and configuration mapping
"We need to block all sharing of SharePoint and OneDrive items to all external recipients..."	- Administrative scope: **Full directory** - Where to monitor: **Enterprise applications & devices**, **SharePoint sites**, **OneDrive accounts** - Conditions for a match: First **Condition** > **Shared outside my org** - Action: **Restrict access or encrypt the content in Microsoft 365 locations** > **Block users from receiving email or accessing shared SharePoint, OneDrive** > **Block only people outside your organization**
"...that contain social security numbers, credit card data or have the "Confidential" sensitivity label..."	- What to monitor: use the **Custom** template - Condition for a match: Create a second condition that is joined to the first condition with a boolean AND - Conditions for a match: Second condition, first condition group > **Content contains Sensitive info types** > **U.S. Social Security Number (SSN)**, **Credit Card Number** - **Condition group configuration** Create a second Condition group connected to the first by boolean OR - Condition for a match: Second condition group, second condition > **Content contains any of these sensitivity labels** > **Confidential**.
"...We don't want this to apply to anyone on the Human Resources team..."	- Where to apply: exclude the Human Resources Team OneDrive accounts
"...We want to notify our Security team with an email every time a file is shared and then blocked..."	- Incident reports: **Send an alert to admins when a rule match occurs** - Send email alerts to these people (optional): add the Security team - Send an alert every time an activity matches the rule: **selected** - Use email incident reports to notify you when a policy match occurs: **On** - Send notifications to these people: add individual admins as desired

Statement	Configuration question answered and configuration mapping
	- You can also include the following information in the report: select **all options**
"...In addition, we want the user to be alerted via email and within the interface if possible..."	- User notifications: **On** - Notify users in Office 365 with a policy tip: **selected**
"...Lastly, we don't want any exceptions to the policy and need to be able to see this activity within the system..."	-User overrides: **not selected**

When you configure the conditions, the summary looks like this:



Steps to create policy

> **Important**
>
> For this policy creation procedure, leave the policy turned off. Change these settings when you deploy the policy.

1. Sign in to the Microsoft Purview portal.
2. Select **Data loss prevention** > **Policies** > **+ Create policy**.
3. Select **Enterprise applications & devices**.
4. Select **Custom** from both the **Categories** list and the **Regulations** list.
5. Select **Next**.
6. Enter a **Name** and a **Description** for the policy. You can use the policy intent statement here.

> **Important**

You can't rename policies.

1. Select **Next**.
2. Accept the default **Full directory** on the **Assign admin units** page.
3. Select **Next**.
4. Choose where to apply the policy.
   a. Ensure that the **SharePoint sites** and **OneDrive accounts** locations are selected.
   b. Deselect all other locations.
   c. Select **Edit** in the **Actions** column next to **OneDrive accounts**.
   d. Select **All users and groups** and then select **Exclude users and groups**.
   e. Select **+Exclude** and then **Exclude groups**.
   f. Select **Human Resources**.
5. Select **Done** and then select **Next**.
6. On the **Define policy settings** page, the **Create or customize advanced DLP rules** option should already be selected. Select **Next**.
7. On the **Customize advanced DLP rules** page, select **+ Create rule**.
8. Give the rule a **Name** and a **description**.
9. Select **Add condition** and use these values:
   a. Select **Content is shared from Microsoft 365**.
   b. Select **with people outside my organization**.
10. Select **Add condition** to create a second condition and use these values.
    a. Select **Content contains**.
11. Select **Add** > **Sensitivity labels** > and then **Confidential**.
12. Select **Add**.
13. Under **Actions**, add an action with these values:
    a. **Restrict access or encrypt the content in Microsoft 365 locations**.
    b. **Block only people outside your organization**.
14. Set the **User Notifications** toggle to **On**.
15. Select **Notify users in Office 365 services with a policy tip** and then select **Notify the user who sent, shared, or last modified the content**.
16. Under **User overrides**, make sure that **Allow override from M365 services** is *NOT* selected.
17. Under **Incident reports**:
    a. Set **Use this severity level in admin alerts and reports** to **Low**.
    b. Set the toggle for **Send an alert to admins when a rule match occurs** to **On**.
18. Under **Send email alerts to these people (optional)**, choose **+ Add or remove users** and then add the email address of the security team.
19. Select **Save** and then select **Next**.

20. On the **Policy mode** page, choose **Run the policy in simulation mode** and **Show policy tips while in simulation mode**.
21. Select **Next** and then select **Submit**.
22. Select **Done**.

Variant: Block specific external domains or users (public preview)

This variant uses the same scenario as the main SharePoint and OneDrive external sharing policy but swaps the action to demonstrate the **Block access for specific external domains or users** suboption, available in public preview for SharePoint and OneDrive only. Use this variant when you need to block access for a specific list of external domains or user SMTPs rather than blocking all external recipients.

Policy intent statement (variant)

We need to block sharing of SharePoint and OneDrive items containing social security numbers, credit card data, or the "Confidential" sensitivity label to people on our list of restricted external domains (for example, `adatum.com`) and to specific guest SMTPs (for example, `j.doe@firstupconsultants.com`). All other external recipients are allowed. We don't want this to apply to anyone on the Human Resources team. Alerting is supported for this action. User notification and override capability are not supported for this action.

Configuration mapping (variant)

The following table maps the variant policy intent to the corresponding configuration choices.

Expand table

Statement	Configuration question answered and configuration mapping
"...block sharing... to people on our list of restricted external domains... and to specific guest SMTPs..."	- Action: **Restrict access or encrypt the content in Microsoft 365 locations** > **Block users from receiving email or accessing shared SharePoint, OneDrive** > **Block access for specific external domains or users** - Specify each restricted domain with **Domain IS** (for example, `adatum.com`). - Specify each restricted user with **User IS** (for example, `j.doe@firstupconsultants.com`). - (Optional) Use **Domain IS NOT** or **User IS NOT** to explicitly allow specific domains or users that would otherwise match a block list.

All other mapping rows from the main SharePoint and OneDrive external sharing scenario — administrative scope, conditions, exclusions, incident reports, user notifications, and user overrides — are unchanged.

Steps to create the variant policy

Follow the same procedure in Steps to create policy, with one change at the **Actions** step:

- Under **Actions**, add an action with these values:

    1. **Restrict access or encrypt the content in Microsoft 365 locations**.
    2. **Block access for specific external domains or users**.
    3. Add each restricted domain by using **Domain IS** (for example, `adatum.com`).
    4. Add each restricted user by using **User IS** (for example, `j.doe@firstupconsultants.com`).
    5. (Optional) Use **Domain IS NOT** or **User IS NOT** to add allow entries that override a block entry.

> ⓘ **Note**
>
> When you use **Block access for specific external domains or users**: if a user or domain appears in both allow and block lists, the block takes effect (most restrictive wins). If a file matches both an allow rule and a block rule, evaluation is per rule — allowed users and domains are permitted, blocked users and domains are denied, and users in neither list are blocked by default. Internal users and domains can't be blocked with this suboption; use **Block everyone** for internal users.

> ⓘ **Important**
>
> Public preview limitations: image files aren't protected by this suboption, and multiple audit records can be generated for certain actions on a blocked file by a blocked user.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Help prevent sharing Power BI reports with credit card numbers

Overview

> **Note**
>
> The features used in this scenario are in preview.

This article uses the process you learned in Design a data loss prevention policy to show you how to create a Microsoft Purview data loss prevention (DLP) policy that helps protect against unintentional sharing of Power BI reports that contain credit card numbers. Work through this scenario in your test environment to familiarize yourself with the policy creation UI.

> **Important**
>
> This article presents a hypothetical scenario with hypothetical values. It's only for illustrative purposes. Substitute your own sensitive information types, sensitivity labels, distribution groups, and users.

How you deploy a policy is as important as policy design. This article shows you how to use the deployment options so that the policy achieves your intent while avoiding costly business disruptions.

Prerequisites and assumptions

This scenario uses the *Highly Confidential - Internal* sensitivity label, so it requires that you create and publish sensitivity labels. To learn more, see:

- Learn about sensitivity labels
- Get started with sensitivity labels
- Create and configure sensitivity labels and their policies

The following policy creation procedure uses alerts. For setup information, see Get started with the data loss prevention alerts.

Policy intent statement and mapping

We need to block external users from reports containing credit card numbers, unless the data is labeled with the 'Highly Confidential - Internal' sensitivity lable, in which case a protection policy restricts access to select security groups. We want to notify the compliance admin to know whenever a semantic model is blocked and the data owner to be aware the restriction took place. We also want internal users to be aware that the data is highly confidential and that they shouldn't share it outside the organization.

Expand table

Statement	**Configuration question answered and configuration mapping**
"We need to block external users..."	- Where to monitor: **Fabric and Power BI** - Administrative scope: **Full directory** - Action: **Restrict access or encrypt the content in Microsoft 365 locations** > **Block users from receiving email or accessing shared SharePoint, OneDrive, and Teams files, and Power BI items** > **Block only people outside your organization**
"...from reports containing credit card numbers..."	- What to monitor: use the **Custom template** - Conditions for match: edit it to add the *Credit Card Number* sensitive info type.
"...except if the data is labeled with the *Highly Confidential - Internal* sensitivity label..."	- Condition group configuration: Create a nested boolean NOT condition group joined to the first condition using a boolean AND - Condition for match: edit it to add the *Highly Confidential - Internal* sensitivity label.
"We want to notify the compliance admin to know whenever a semantic model is blocked..."	-Incident reports: **Send an alert to admins when a rule match occurs**: **On** - Send an alert every time an activity matches the rule: **selected**
"...the data owner to be aware the restriction took place. We also want internal users to be aware that the data is highly confidential and that they shouldn't share it outside the organization."	- User notifications: **On** - Microsoft 365 files and Microsoft Fabric items: Notify users in Office 365 service with a policy tip or email notifications: **selected** - Policy tips: Customize the policy tip text: **selected**. Add text in the text box explaining the rules governing sharing highly confidential data.

Steps to create the policy

> **Important**
>
> For this policy creation procedure, accept the default include and exclude values and leave the policy turned off. Change these values when you deploy the policy.

1. Sign in to the Microsoft Purview portal.
2. Open the **Data loss prevention** solution and go to **Policies**.
3. Select **Create policy**.
4. Select **Enterprise applications & devices**.
5. Select **Custom** from the **Categories** list.
6. Select **Custom policy** from the **Regulations** list.
7. Select **Next**.
8. Enter a name and description for the policy. You can use the policy intent statement here.

   > **Important**
   >
   > Policies can't be renamed

9. Select **Next**.
10. Accept the **Full directory** default under **Admin units**.
11. Select **Next**.
12. Choose where to apply the policy. Select only the **Fabric and Power BI workspaces** location.
13. Select **Next**.
14. On the **Define policy settings** page, make sure the **Create or customize advanced DLP rules** option is selected.
15. Select **Next**.
16. Select **+ Create rule**. Name the rule and provide a description.
17. Under **Conditions**, select **Add condition** > **Content contains** > **Add** > **Sensitive info types**.
18. Select the info types that apply to credit card numbers.
19. Select **Add**.
20. Next, under the **Content contains** section, select **Add group**.
21. Leave the Boolean operator set to **AND**, then set the toggle to **NOT**.
22. In the line below the toggle, select **Add condition** > **Content contains** > **Add** > **Sensitivity labels**.
23. Select the sensitivity label **Highly Confidential - Internal**.
24. Select **Add**.

25. Under **Actions**, select **Add an action** > **Restrict access or encrypt the content in Microsoft 365 locations** > **Block users from receiving email or accessing shared SharePoint, OneDrive, and Teams files, and Power BI items** > **Block only people outside your organization**.
26. Under **User notifications**, set the toggle to **On**.
27. Select **Notify users in Office 365 service with a policy tip or email notifications**, and then select **Customize the policy tip text**.
28. Provide a policy tip that explains that the data in the report is highly confidential is not to be shared outside the organization.
29. Under **Incident reports**, set **Use this severity level in admin alerts and reports** to **High**.
30. Make sure the **Send an alert to admins when a rule match occurs** toggle is set to **On**.
31. Make sure the **Send alert every time an activity matches the rule** radio button is selected.
32. Select **Save**.
33. Review the rule, then select **Next**.
34. Make sure the **Run the policy in simulation mode** radio button and the **Show policy tips while in simulation mode** checkbox are selected.
35. Select **Next**.
36. Review the policy, then select **Submit**.
37. Select **Done**.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

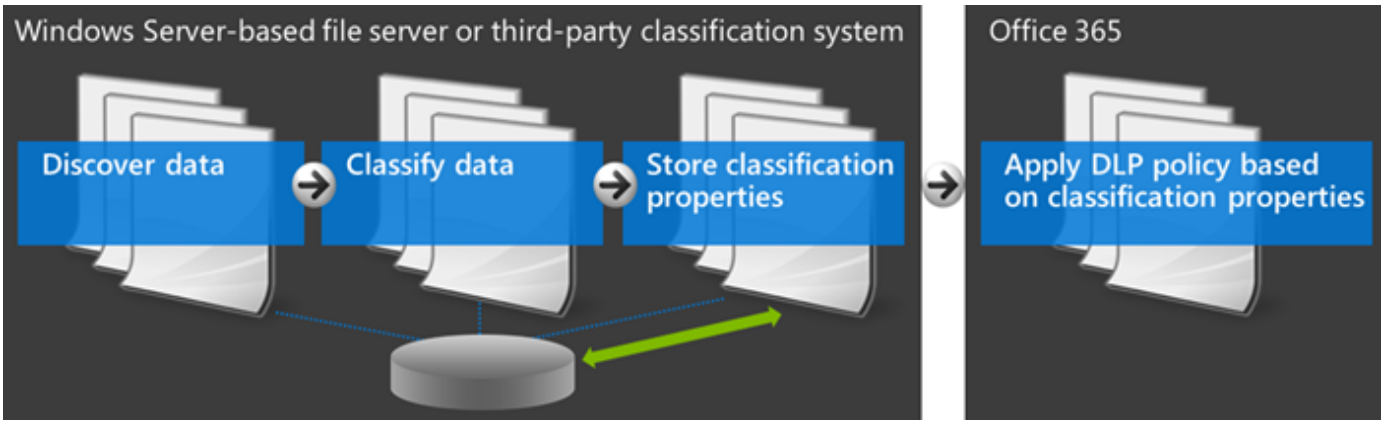
---

Last updated on 06/26/2026

Create a DLP policy to protect documents with FCI or other properties

Microsoft Purview Data Loss Prevention (DLP) policies can use classification properties or item properties to identify sensitive items. For example you can use:

- Windows Server File Classification infrastructure (FCI) properties
- SharePoint document properties
- third-party system document properties



For example, your organization might use Windows Server FCI to identify items with personal data, such as social security numbers, and then classify those documents by setting the **Personally Identifiable Information** property to **High**, **Moderate**, **Low**, **Public**, or **Not PII** based on the type and number of occurrences of personal data found in each document.

In Microsoft 365, you can create a DLP policy that identifies documents that have that property set to specific values, such as **High** and **Medium**, and then takes an action such as blocking access to those files. The same policy can have another rule that takes a different action if the property is set to **Low**, such as sending an email notification. This way, DLP integrates with Windows Server FCI and can help protect Office documents uploaded or shared to Microsoft 365 from Windows Server-based file servers.

A DLP policy simply looks for a specific property name/value pair. Any document property can be used, as long as the property has a corresponding managed property for SharePoint search. For example, a SharePoint site collection might use a content type named **Trip Report** with a required field named **Customer**. Whenever a person creates a trip report, they must enter the customer name. This property name/value pair can also be used in a DLP policy—for example, if you want a rule that blocks access to the document for guests when the **Customer** field contains **Contoso**.

If you want to apply your DLP policy to content with specific Microsoft 365 labels, don't follow the steps here. Instead, refer to lCreate and Deploy data loss prevention policies.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about **Microsoft Security Copilot in Microsoft Purview**.

Before you create the DLP policy

Before you can use a Windows Server FCI property or other property in a DLP policy, you need to create a managed property in the SharePoint admin center. Here's why.

In SharePoint and OneDrive, the search index is built up by crawling the content on your sites. The crawler picks up content and metadata from the documents in the form of crawled properties. The search schema helps the crawler decide what content and metadata to pick up. (Examples of metadata are the author and the title of a document.) However, to get the content and metadata from the documents into the search index, the crawled properties must be mapped to managed properties. Only managed properties are kept in the index. For example, a crawled property related to author is mapped to a managed property related to author.

> **Note**
>
> Be sure to use a managed property name, not a crawled property name, when creating DLP rules using the `ContentPropertyContainsWords` condition. This is important because DLP uses the search crawler to identify and classify sensitive information on your sites, and then stores that sensitive information in a secure portion of the search index. When you upload a document to Office 365, SharePoint automatically creates crawled properties based on the document properties. However, to use an FCI or other property in a DLP policy, that crawled property needs to be mapped to a managed property so that the content with that property is kept in the index.

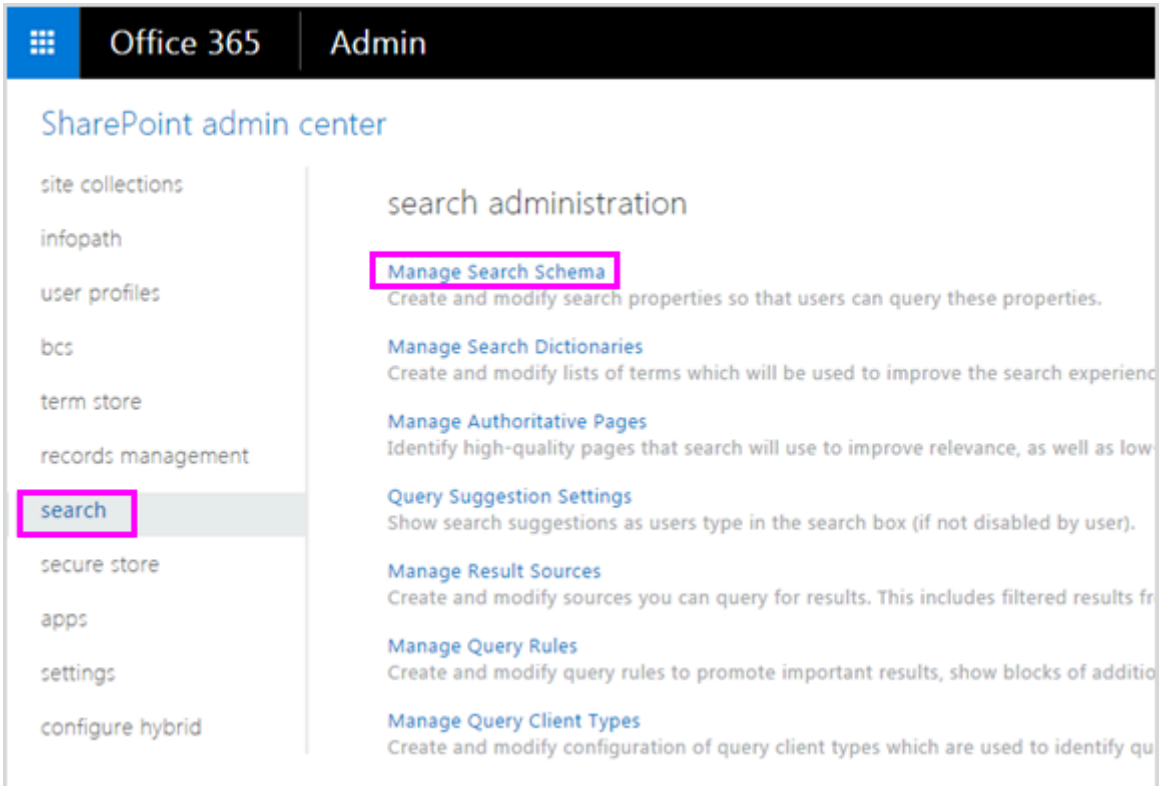
For more information on search and managed properties, see Manage the search schema in SharePoint Online.

Step 1: Upload a document with the needed property to Office 365

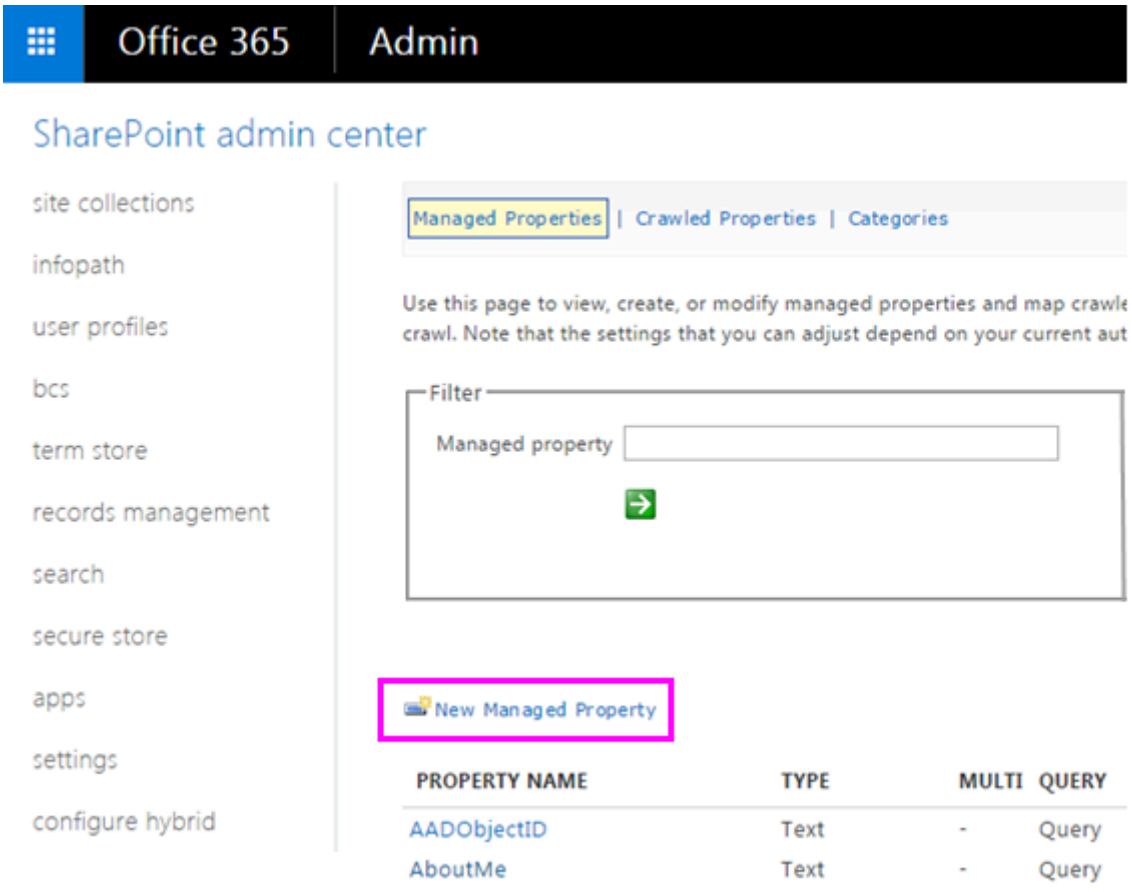
First, you need to upload a document with the property that you want to reference in your DLP policy. Microsoft 365 detects the property and automatically create a crawled property from it. In the next step, you'll create a managed property, and then map the managed property to this crawled property.

Step 2: Create a managed property

1. Sign in to the Microsoft 365 admin center.
2. In the left navigation pane, choose **Admin centers** > **SharePoint**. You're now in the SharePoint admin center.
3. In the left navigation pane, choose **search**. On the **search administration** page, choose **Manage Search Schema**.



4. On the **Managed Properties** page > **New Managed Property**.



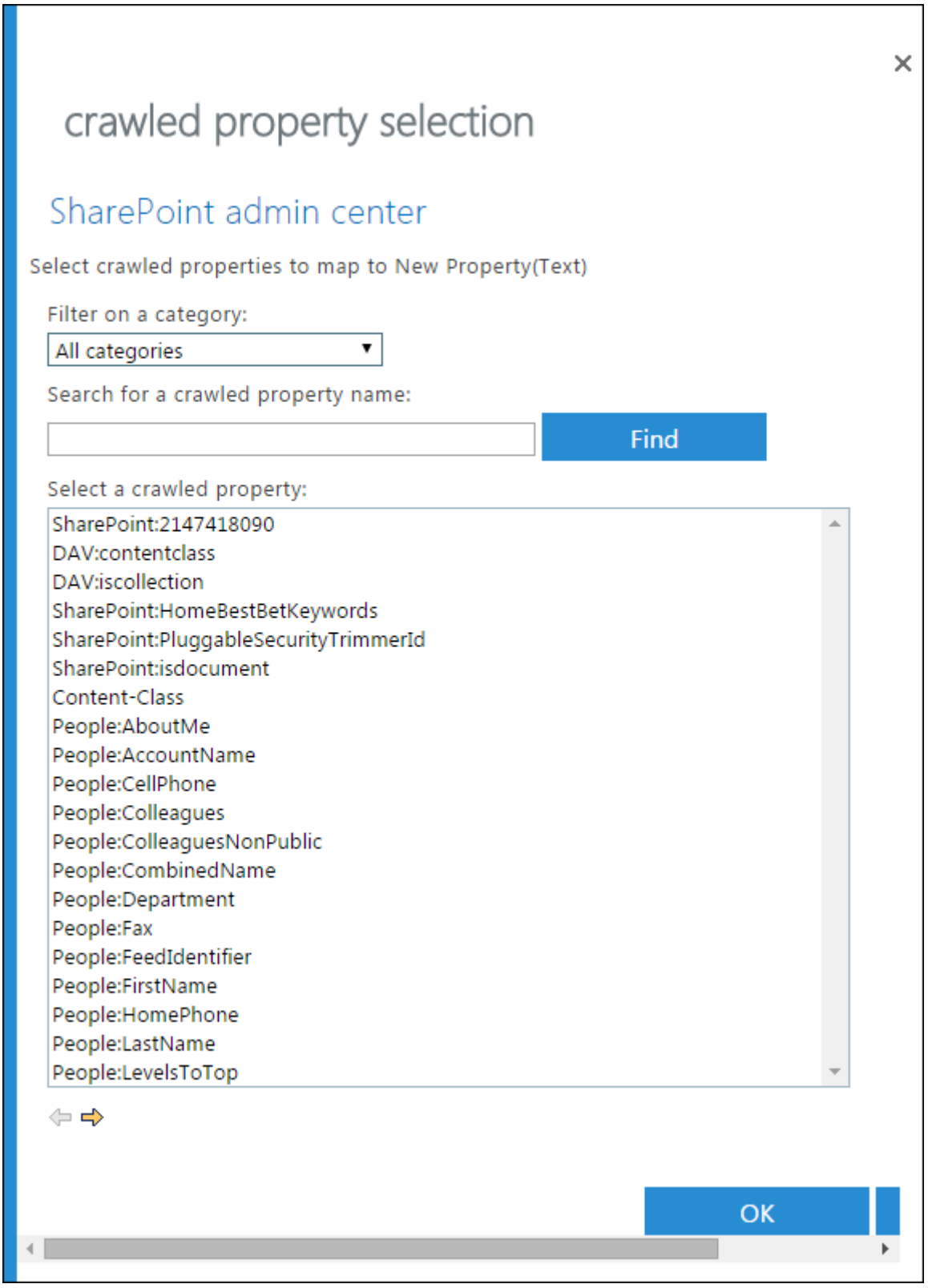
5. Enter a name and description for the property. This name is what will appear in your DLP policies.

6. For **Type**, choose **Text**.

7. Under **Main characteristics**, select **Queryable** and **Retrievable**.

8. Under **Mappings to crawled properties** choose **Add a mapping**.

9. In the **crawled property selection** dialog box, find and select the crawled property that corresponds to the Windows Server FCI property or other property that you'll use in your DLP policy, then choose **OK**.



10. At the bottom of the page choose **OK**.

Create a DLP policy that uses an FCI property or other property

In this example, an organization is using FCI on its Windows Server-based file servers; specifically, they're using the FCI classification property named **Personally Identifiable Information** with

possible values of **High**, **Moderate**, **Low**, **Public**, and **Not PII**. Now they want to use their existing FCI classification in their DLP policies in Office 365.

To begin, they follow the steps above to create a managed property in SharePoint Online, which maps to the crawled property created automatically from the FCI property.

Next, they create a DLP policy with two rules that both use the condition **Document properties contain any of these values**:

- **FCI PII content - High, Moderate** The first rule restricts access to the document if the FCI classification property **Personally Identifiable Information** equals **High** or **Moderate** and the document is shared with people outside the organization.
- **FCI PII content - Low** The second rule sends a notification to the document owner if the FCI classification property **Personally Identifiable Information** equals **Low** and the document is shared with people outside the organization.

After you create the DLP policy

Completing the steps in the previous sections creates a DLP policy that will quickly detect content with that property, but only if that content is newly uploaded (so that the content's indexed), or if that content is old but edited (so that the content's reindexed).

To detect content with that property everywhere, you need to have your library, site, or site collection reindexed, so that the DLP policy is aware of all the content with that property. In SharePoint, content is automatically crawled when content is edited. Specific SharePoint sites can't be manually reindexed.

> ⊗ **Caution**
>
> Reindexing a site for DLP scenarios isn't possible.

For more information, see Manually request crawling and reindexing of a site, a library, or a list.

Reindex a site (optional)

1. On the site, choose **Settings** (gear icon in upper right) > **Site Settings**.
2. Under **Search**, choose **Search and offline availability** > **Reindex site**.

More information

- Learn about data loss prevention
- Create and Deploy data loss prevention policies
- Send notifications and show policy tips for DLP policies
- What the DLP policy templates include
- Sensitive information type entity definitions

Last updated on 05/12/2025

Learn about the data loss prevention on-premises repositories

Article • 03/31/2025

When you select the **On-premises repositories** location, Microsoft Purview Data Loss Prevention (DLP) can enforce protective actions on on-premises data-at-rest in file shares and SharePoint document libraries and folders. This gives you the visibility and control you need to ensure that sensitive items are used and protected properly, and to help prevent risky behavior that might compromise them. The DLP detects sensitive information by using built-in or custom sensitive information types, sensitivity labels or file properties. The information about what users are doing with sensitive items is made visible in activity explorer and you can enforce protective actions on those items via DLP policies.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about **Microsoft Security Copilot in Microsoft Purview**.

DLP relies on Microsoft Purview Information Protection scanner

DLP relies on a full implementation of the Microsoft Purview Information Protection scanner to monitor, label, and protect sensitive items. If you haven't implemented Information Protection scanner, you must do so before you can use DLP. For more information, read these articles:

- What is Azure Information Protection
- Learn about the information protection scanner
- Get started with the information protection scanner
- Configuring and installing the information protection scanner
- Microsoft Purview Information Protection client - Release management and supportability

DLP On-premises repository actions

DLP detects files in on-premises repositories by looking for the following:

- sensitive information types
- sensitivity labels
- file extension
- custom document properties on Office files only

When a detected file matches a DLP policy, DLP can take one of the following four actions.

Expand table

Action	Description
Block people from accessing file stored in on-premises scanner - Block everyone	When enforced, this action blocks access to all accounts except the content owner, the account that last modified the item, and the administrator. It does this by removing all accounts from NTFS/SharePoint permissions at the file level except for the file owner, repository owner (set in the Use a DLP policy setting in content scan job), last modifier (can be identified in SharePoint only), and admin. The scanner account is also granted FC rights on the file.
Block only people who have access to your on-premises network and users in your organization who weren't granted explicit access to the files from accessing file	When enforced, this action removes the ***Everyone***, ***NT AUTHORITY\authenticated users***, and ***Domain Users*** SIDs from the file access control list (ACL). Only users and groups that have been explicitly granted rights to the file or parent folder will be able to access the file.
Set permissions on the file (permissions will be inherited from the parent folder)	When enforced, this action forces the file to inherit the permissions of its parent folder. By default, this action will only be enforced if the permissions on the parent folder are more restrictive than the permissions that are already on the file. For example, if the ACL on the file is set to allow only ***specific users*** and the parent folder is configured to allow the ***Domain Users*** group, the parent folder permissions wouldn't be inherited by the file. You can override this behavior by selecting the **Inherit even if parent permissions are less restrictive** option.
Remove the file from improper location	When enforced, this action replaces the original file with a stub file with .txt extension and places a copy of the original file in a quarantine folder.

What's different in the on-premises scanner

There are a few extra concepts that you need to be aware of before you dig into the on-premises scanner.

Scanner repositories and content scan jobs

You must create a content scan job for the information protection scanner and identify the repositories that host the files that you want to DLP to evaluate. Make sure you enable DLP rules in the created content scan job.

Policy tips

Policy tips aren't available in on-premises scanner.

Viewing DLP on-premises scanner events

You view DLP data in the Microsoft Purview portal activity explorer.

Next steps

Now that you've learned about the Information Protection on-premises scanner, your next steps are:

1. Get started with the On-premises repositories location
2. Use the DLP on-premises scanner

See also

- Getting started with the data loss prevention on-premises scanner
- Use the data loss prevention on-premises scanner
- Learn about data loss prevention
- Create and Deploy data loss prevention policies
- Get started with Activity explorer

Feedback

Was this page helpful? Yes No

Learn about the default data loss prevention policy in Microsoft Teams

Article • 03/28/2025

Microsoft Purview Data Loss Prevention capabilities have been extended to include Microsoft Teams chat and channel messages, including private channel messages. As a part of this release, we created a default DLP policy for Microsoft Teams for first-time customers to the Microsoft Purview portal.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about **Microsoft Security Copilot in Microsoft Purview**.

Licensing

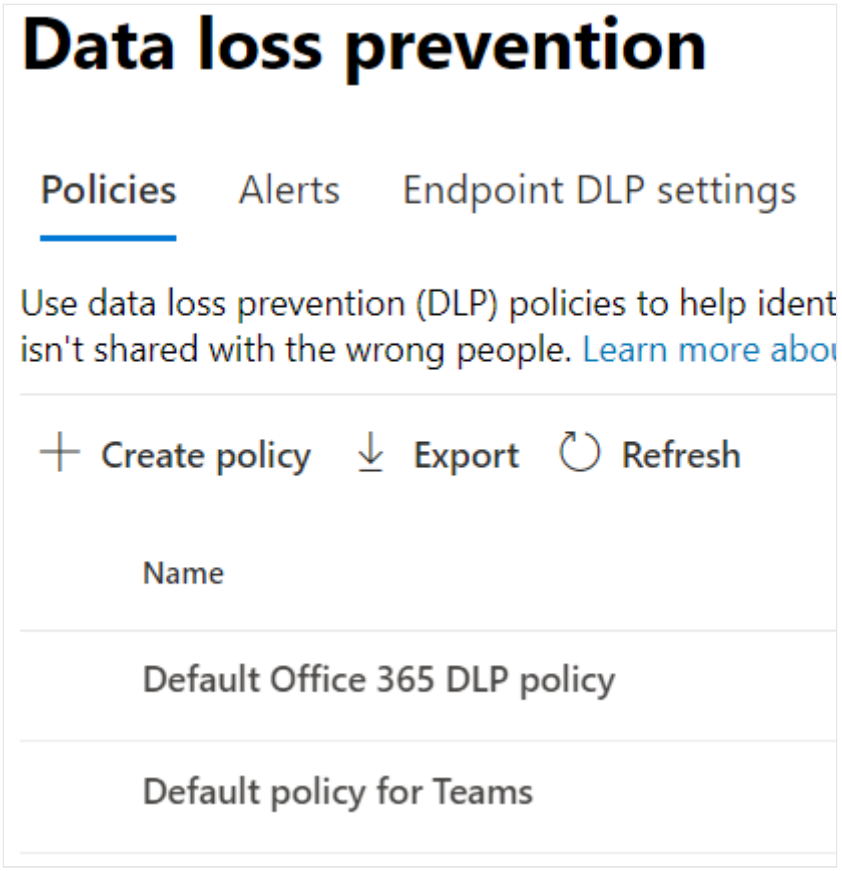
For information on licensing, see

- Microsoft 365 Enterprise Plans
- Microsoft 365 Service Descriptions

What does the default policy do?

The default DLP policy for Teams tracks all the credit card numbers shared internally and externally to the organization. This policy is on by default for all users of the tenant. It does not generate any policy tips for end users but does generate an Alert event and also triggers a low severity email to the admin (added in the policy). Administrator can view the activities and edit the policies details by logging into the Microsoft Purview portal.

Admins can view this policy in the Microsoft Purview portal. Navigate to **Data Loss prevention** > **Policies**.



Edit or delete the default policy

To edit the default policy for better performance or to delete it, just use an account with **DLP Compliance Management** permissions. For more information, see, Permissions.

Feedback

Was this page helpful? Yes No

Data loss prevention and Microsoft Teams

If your organization has Microsoft Purview Data Loss Prevention (DLP), you can define policies that help prevent people from sharing sensitive information in a Microsoft Teams channel or chat session. Here are some example scenarios:

- **Protecting sensitive information in messages**. Suppose that someone attempts to share sensitive information in a Teams chat or channel with guests (external users). If you have a DLP policy defined in your tenant to prevent this, messages with sensitive information that are sent to external users are deleted. However, your tenant must be the initiator of the chat or the thread. This happens automatically, and within seconds, according to how your DLP policy is configured.

- **Protecting sensitive information in documents**. Suppose that someone attempts to share a document with guests in a Microsoft Teams channel or chat, and the document contains sensitive information. If you have a DLP policy defined to prevent this, the document won't open for those users. Your DLP policy must include SharePoint and OneDrive in order for protection to be enforced. This is an example of DLP for SharePoint that shows up in Microsoft Teams, and therefore requires that users are licensed for Office 365 DLP (included in Office 365 E3), but doesn't require that users be licensed for Office 365 Advanced Compliance.

- **Protecting communications in Teams Shared Channels**. Shared channels are built on Microsoft Entra B2B direct connect, so each participant remains signed in to their own (home) tenant and doesn't get an identity object in the host tenant. Because DLP evaluation is driven by the sender's identity, each user's messages are evaluated against the DLP policies of their own tenant, regardless of which tenant owns the shared channel. For example, let's say there's a shared channel owned by Team A of Contoso, where Contoso has a DLP policy P1 and Fabrikam has a DLP policy P2. There are three ways to share the channel:
  - **Share with member**: You invite User1 from Contoso to join the shared channel without making them a member of Team A. User1's messages are evaluated against Contoso's DLP policies (P1), the same policies that apply to User1 elsewhere in Contoso.
  - **Share with team (internally)**: You share the channel with another team within Contoso, Team B. Messages from Team A and Team B users are evaluated against Contoso's DLP policies. If Team A and Team B are scoped to different policies, each user is still evaluated against the policies that target them.

    - **Share with team (cross tenant)**: You share the channel with a team, Team F, in Fabrikam. Contoso's DLP policies (P1) apply to messages sent by Team A (Contoso) users. Fabrikam's DLP policies (P2) apply to messages sent by Team F (Fabrikam) users. Contoso can't enforce its DLP policies on Fabrikam users' messages, and vice versa.

- **Protecting communications when chatting with external users in Microsoft Teams**.

People from different Microsoft 365 organizations who use the *external* access feature can participate in the same chat session. DLP policy enforcement in these cross-tenant conversations is based on the hosting tenant, which is the tenant that initiated the chat or thread.

For example, suppose UserA, UserB, and UserC are from Contoso, while UserX, UserY, and UserZ are from Fabrikam, and they participate in the same Teams group chat. Contoso's DLP policies apply to messages sent by Contoso users only when Contoso is the hosting tenant, such as when UserA, UserB, or UserC created the chat. In this case, Fabrikam's DLP policies aren't evaluated or enforced on messages sent by UserX, UserY, or UserZ, and Contoso's DLP policies aren't applied to those Fabrikam users.

Therefore, in a cross-tenant conversation, a user's home-tenant DLP policies don't automatically follow the user into a conversation hosted by another tenant.

For more information on using Microsoft Teams to chat with people external to your organization, see Manage external meetings and chat with people and organizations using Microsoft identities

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about **Microsoft Security Copilot in Microsoft Purview**.

DLP Licensing for Microsoft Teams

Data loss prevention capabilities include Microsoft Teams chat and channel messages, **including private channel messages** for:

- Office 365 E5/A5/G5
- Microsoft 365 E5/A5/G5
- Microsoft 365 E5/A5/G5 Information Protection and Governance
- Microsoft 365 E5/A5/G5/F5 Compliance and F5 Security & Compliance

Office 365 and Microsoft 365 E3 include DLP protection for SharePoint, OneDrive, and Exchange. This also includes files that are shared through Teams because Teams uses SharePoint and OneDrive to share files.

Support for DLP protection in Teams Chat requires an E5 license.

> **Important**
>
> For information on licensing, see Microsoft 365, Office 365, Enterprise Mobility + Security, and Windows 11 Subscriptions for Enterprises.

Scope of DLP protection

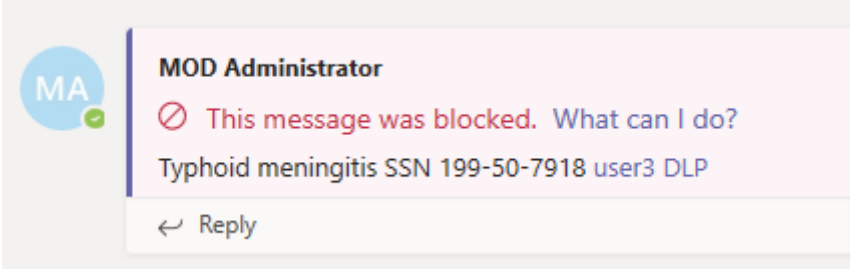
DLP protection is applied differently to Teams entities and is described in the table that follows.

Expand table

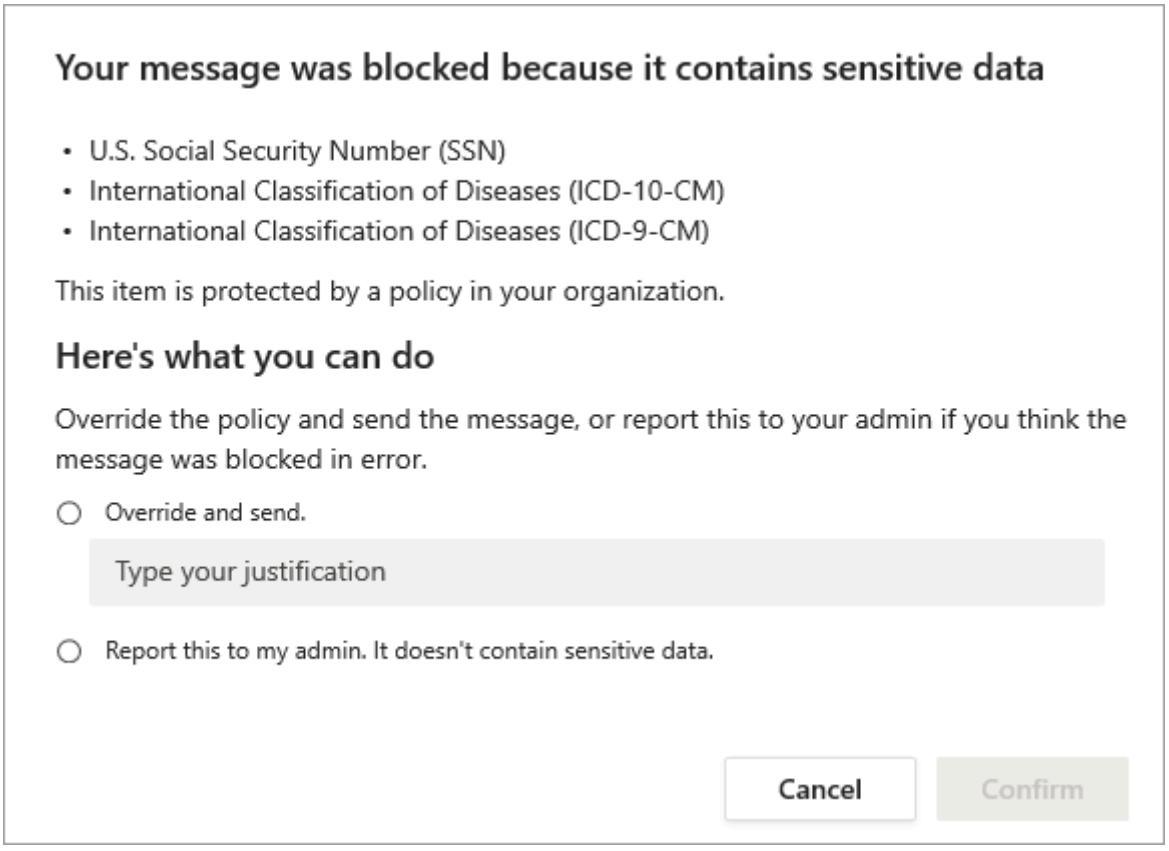
Policy scope	Teams Entities	DLP Protection
Individual user accounts	1:1/n chats	Yes
Individual user accounts	Standard, private and shared channel messages	No
Security groups/distribution group/non-mail-enabled security group	1:1/n chats	Yes
Security groups/distribution group/non-mail enabled security group	Standard, private and shared channel messages	Yes
Microsoft 365 groups	1:1/n chats	Yes
Microsoft 365 groups	Standard, private and shared channel messages	Yes

Policy tips help educate users

Similar to how DLP policy tips work in Exchange, Outlook, SharePoint, OneDrive, and on Windows Devices, policy tips in Teams appear when an action triggers with a DLP policy. Here's an example of a policy tip:

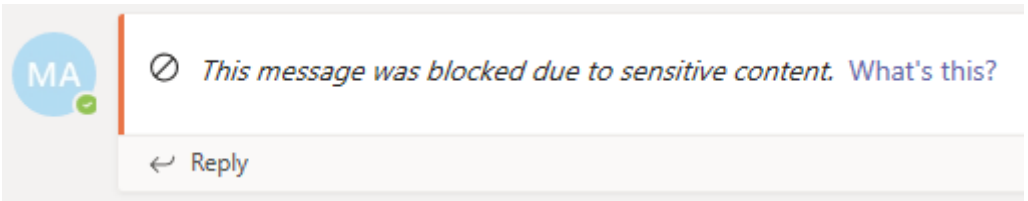


Here, the sender attempted to share a social security number in a Microsoft Teams channel. The **What can I do?** Link opens a dialog box that provides options for the sender to resolve the issue. Notice that the sender can opt to override the policy or notify an admin to review and resolve the issue.



You can choose to allow users in your organization to override a DLP policy. When you configure your DLP policies, you can use the default policy tips, or customize policy tips for your organization.

Returning to our example, when a sender has shares a social security number in a Teams channel, here's what the recipient sees:



DLP protections are applied to the actual messages in the chat or channel thread. Message information is also shown in the short preview in Activity notifications that are created from chat or channel message. When there's a match to a DLP policy, the corresponding preview is hidden, and a "Preview Unavailable" message is shown instead of the blocked preview. An Activity entry is also generated for the sender if the message sent matched a DLP policy. Flagged and blocked messages, activities are created stating "Your message has been blocked".

DLP for Microsoft Teams does not send user notification emails like DLP for Exchange, SharePoint, and OneDrive. Instead, users only receive message flags as notifications.

To customize policy tips

To perform this task, you must be assigned a role that has permissions to edit DLP policies. To learn more, see Permissions in the Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal >**Data loss prevention** > **Policies**.
2. Select a policy, and then choose **Edit policy** (the pencil icon).
3. Navigate through the tool until you come to the **Customize advanced DLP rules** screen.
4. Either create a new rule, or edit an existing rule for the policy.
5. Scroll down to **User notifications** and set the **Use notifications to inform your users and help educate them on the proper use of sensitive info** toggle to **On**.
6. Under **Microsoft 365 services**, select **Notify users in Office 365 service with a policy tip**.
7. Under **Policy tips**, select **Customize the policy tip text**.
8. Specify the text you want to use for the policy tip.
9. If the policy tip applies to user activities in Microsoft Exchange and you want a dialog box to display the tip before an email is sent, select **Show the policy tip as a dialog for the end user before send**.
10. Choose **Save** and then **Next**.
11. On the **Policy mode** page, check the box next to **Show policy tips while in simulation mode** if desired.
12. Choose **Next**, choose **Submit**, and then choose **Done**.

Add Microsoft Teams as a location to existing DLP policies

To perform this task, you must be assigned a role that has permissions to edit DLP policies. To learn more, see Permissions in the Microsoft Purview portal

1. Sign in to the Microsoft Purview portal > **Data loss prevention** > **Policies**.
2. Select a policy, and then choose **Edit policy** (the pencil icon).
3. Navigate through the tool until you arrive at the **Choose locations to apply the policy** page.
4. Select **Teams chat and channel messages**.
5. Choose **Next** and work your way through to the end of the process.
6. Choose **Submit**.

Allow approximately one hour for your changes to work their way through your data center and sync to user accounts.

Define a new DLP policy for Microsoft Teams

For information about how to create and implement a new DLP policy, see Create and Deploy data loss prevention policies.

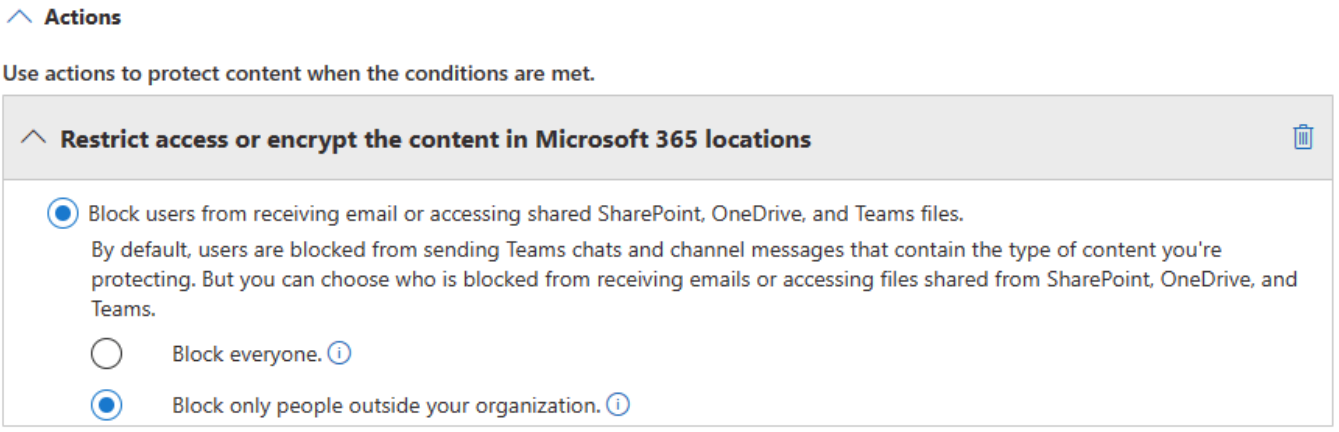
Prevent external access to sensitive documents

You can ensure that documents are protected until DLP scans and marks them as safe to share by marking new files as sensitive by default.

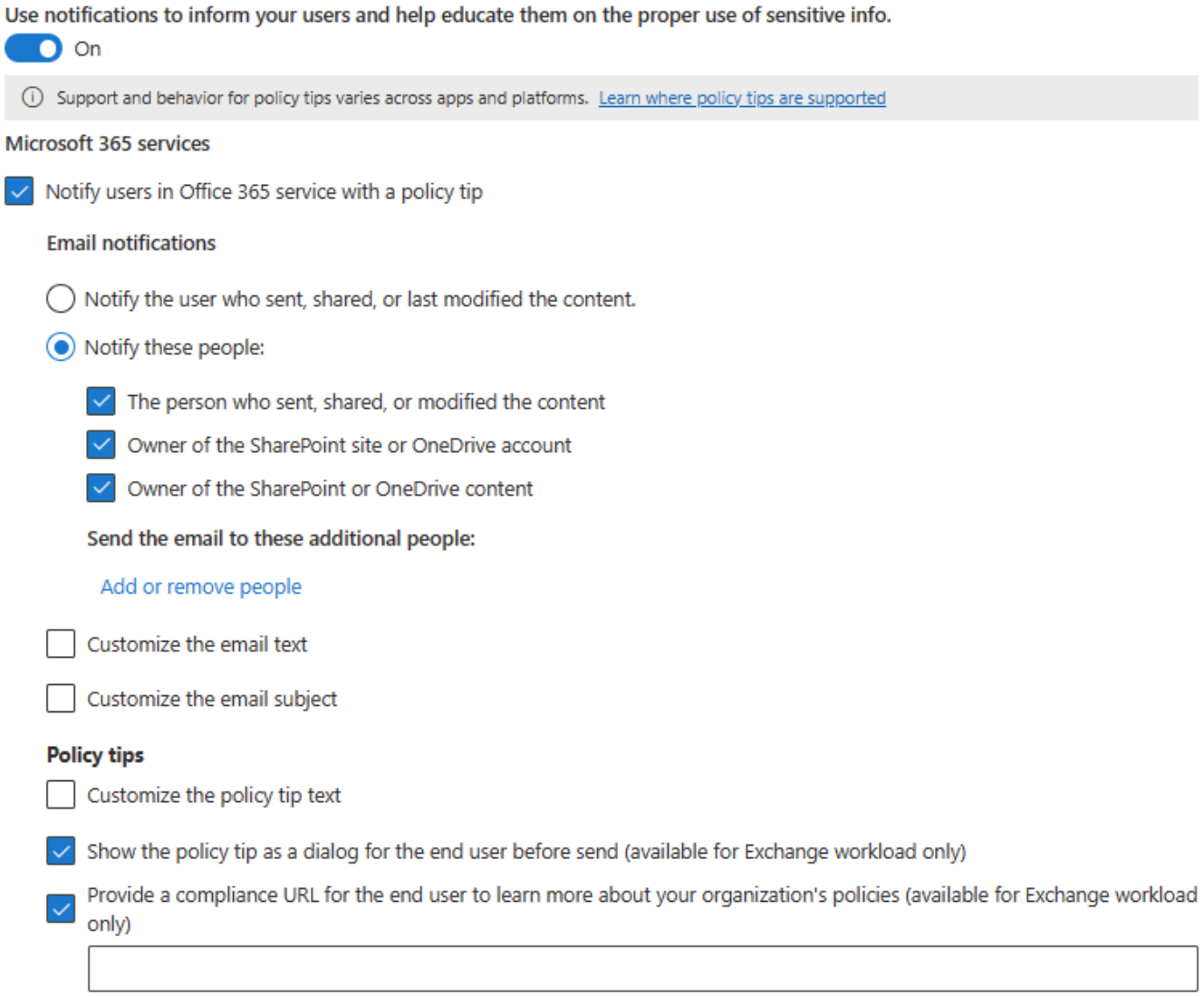
Recommended DLP policy structure

- **Conditions**
- **Content contains any of these sensitive information types**: [Select all that apply]
- **Content is shared from Microsoft 365** > **with people outside my organization**
- **Actions**
  - **Add an action**

- Restrict access or encrypt the content in Microsoft 365 locations > Block only people outside your organization



- Use notifications to inform your users and help educate them on the proper use of sensitive info > Notify users in Office 365 with a policy tip
- **Notify these people** [Select all that apply]
- **Policy tips** [Select all that apply]



> **Note**
>
> The sender address for incident reports is *no-reply@teams.mail.microsoft.com*.

Related articles

- Create and Deploy data loss prevention policies
- Send email notifications and show policy tips for DLP policies

Last updated on 08/28/2026

Learn about using Microsoft Purview Data Loss Prevention to protect interactions with Microsoft 365 Copilot, Copilot Chat, and Cowork

Microsoft Purview Data Loss Prevention (DLP) can help you protect interactions with Microsoft 365 Copilot, Copilot Chat, and Cowork in the following ways:

- **Restrict Microsoft 365 Copilot, Copilot Chat, and Cowork from using external web search when prompts contain sensitive data.** You can use DLP policies to prevent these Copilot experiences from sending sensitive information to external web services. When a prompt contains sensitive information types (SITs)—such as credit card numbers, passport numbers, Social Security numbers, or custom SITs defined by your organization—these experiences automatically block the use of external web search as a grounding source for that prompt. Instead, they continue to generate responses using permitted internal Microsoft 365 data sources. This prevents prompt text that matches the configured SITs from being sent to external search providers for that request.

- **Restrict Microsoft 365 Copilot, Copilot Chat, and Cowork from processing sensitive prompts.** You can create a DLP policy to help protect against the use of sensitive information types (SITs), such as credit card numbers, passport numbers, or Social Security numbers in prompts submitted in these Copilot experiences. This includes Microsoft-provided SITs and custom SITs that you create. This real-time control helps organizations reduce data leakage and oversharing risks. It prevents these experiences from returning a response when prompts contain sensitive data and from using that sensitive data for both internal and external web searches.

- **Restrict Microsoft 365 Copilot, Copilot Chat, and Cowork from processing sensitive files and emails.** You can create a DLP policy to prevent these Copilot experiences from using files and emails that have sensitivity labels when generating responses.

- **Restrict Microsoft 365 Copilot and Copilot Chat from processing external email (preview).** You can prevent Microsoft 365 Copilot and Copilot Chat from using emails sent from external domains as grounding data for responses. When this control is enabled, Copilot excludes external emails received by users from being referenced or summarized during

prompt processing, while continuing to use internal Microsoft 365 data sources where permitted. This protection helps organizations reduce the risk of prompt injection and untrusted data influence. Copilot responses are grounded only in content from trusted internal sources.

> **Important**
>
> You can't use both content contains sensitive info types and content contains sensitivity labels conditions in the same rule. You can create a rule for each condition in the same policy, but not in the same rule.

Licensing

For information on licensing, see

- Microsoft 365 Enterprise Plans
- Microsoft 365 Service Descriptions

Permissions

Accounts with one of the following roles or role groups can create or edit a DLP policy for the Microsoft 365 Copilot and Copilot Chat location:

- **Microsoft Entra AI Admin** - Role for managing all aspects of Microsoft 365 Copilot and AI-related enterprise services in Microsoft 365.
- **Purview Data Security AI Admin** - Role for editing Data Loss Prevention policies related to Copilot and viewing AI content in Data Security Posture Management. This role doesn't have access to read prompts and responses of AI interactions.
- **Purview Data Security AI Admins** - Use this group to assign editing capabilities for Data Loss Prevention policies related to Copilot and viewing AI content in Data Security Posture Management. Review the role description for access details. It contains the Data Security AI Admin role.
- **Purview Compliance Administrator**
- **Purview Compliance Data Administrator**
- **Purview Information Protection**
- **Purview Information Protection Admin**
- **Purview Security Administrator**

- **Microsoft Entra Global Admin** - Role for managing all aspects of Microsoft Entra ID and Microsoft services that use Microsoft Entra identities.

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

Block sensitive information types in web search

To prevent Microsoft 365 Copilot, Copilot Chat, and Cowork from sending sensitive data to external web search providers when grounding responses:

1. Sign in to the Microsoft Purview portal.
2. Go to **Data Loss Prevention** > **Policies** and select **+ Create policy**.
3. Select the **Custom** template, then **Custom policy**.
4. On the **Locations** page, set the **Microsoft 365 Copilot and Copilot Chat** location to on.
5. Add a rule with the **Content contains** > **Sensitive information types** condition and choose the SITs you want to detect.
6. In the same rule, set the action to **Prevent Copilot from processing content** > **Performing Web Searches**.
7. Save and turn on the policy.

When a user prompt contains the configured SITs, Copilot and Cowork block external web search as a grounding source and continue to generate the response using allowed internal Microsoft 365 data sources.

Block SITs in web search use case example

Contoso wants employees to use Microsoft 365 Copilot for productivity, but doesn't want sensitive customer data sent to external web search providers when Copilot grounds responses on the web.

To meet this business need, Contoso configures a DLP policy as follows:

1. Targets the **Microsoft 365 Copilot and Copilot Chat** policy location.

2. Adds a rule with the **Content contains** > **Sensitive information types** condition that detects the SITs representing sensitive customer data.

3. Sets the rule action to **Prevent Copilot from processing content** > **Performing Web Searches**.

When a user submits a prompt containing those SITs, Copilot doesn't send the prompt to external web search providers. Copilot still returns a response grounded in internal Microsoft 365 data sources where the user has access.

Block sensitive information types in prompts

This feature is in preview and is rolling out to all tenants with access to Microsoft 365 Copilot and Copilot Chat. Check whether rollout has reached your tenant. It's available in Microsoft 365 Copilot, Copilot Chat, Cowork, and Copilot in Word, Excel, and PowerPoint.

To set this up, create DLP policies that use the **Microsoft 365 Copilot and Copilot Chat** policy location with the **Content contains** > **Sensitive information types** condition. This policy prevents Copilot and Cowork from returning a response when prompts contain sensitive data. Only out-of-the-box and custom SITs are currently supported.

> **Tip**
>
> During preview, the user messaging in Word, Excel, PowerPoint might not clearly state that the interaction with Copilot in those products is blocked due to an organizational policy. The sensitive prompt is still restricted, and Copilot won't provide a response.

Block SITs in prompts use case example

Contoso encourages their employees to use Microsoft 365 Copilot to enhance productivity, but they don't want their users placing Canada physical addresses or EU debit card numbers into prompts.

To meet this business need, Contoso creates a DLP policy that targets the Microsoft 365 Copilot and Copilot Chat location and has a rule that uses the **Content contains** > **Sensitive information types** > **Canada physical addresses** or **EU debit card numbers** condition to identify prompts that contain those SITs. The actions in the rule are configured to **Restrict Copilot from processing content** > **Processing prompts**.

When a user attempts to submit a prompt that contains either of these sensitive information types, they receive a message indicating that the request can't be completed because it contains

sensitive information that the organization has blocked Microsoft 365 Copilot from using.

Block files and emails with sensitivity labels from being processed

This feature is available in Microsoft 365 Copilot, Copilot Chat, Cowork, and Copilot in Word, Excel, and PowerPoint.

To set this up, create DLP policies that use the **Microsoft 365 Copilot and Copilot Chat** policy location with the **Content contains** > **Sensitivity labels** condition to exclude items from being processed. Identified items still appear in the citations of the response, but Copilot and Cowork don't access or use the content of the item in the response.

Block items with sensitivity labels example use case

Contoso establishes and applies a sensitivity label taxonomy to their data. The taxonomy includes these labels:

- **Highly Confidential**
- **Confidential**
- **Internal**
- **Public**
- **Personal**

They deploy Microsoft 365 Copilot and Copilot Chat to help users find and use Contoso enterprise information in their organization. They want to minimize the risk of General Data Protection Regulation (GDPR) data being included in Microsoft 365 Copilot and Copilot Chat summaries and also exclude private information from summaries. They create a DLP policy that uses the **Microsoft 365 Copilot and Copilot Chat** policy location with the **Content contains** > **Sensitivity labels** condition to exclude items that have the **Personal** sensitivity label from being processed in the response summary and also to exclude items that have the **Highly Confidential** sensitivity label from being processed in the response summary.

Coverage of email and file content

The DLP for **Microsoft 365 Copilot and Copilot Chat** policy location supports specific content that Copilot processes across various experiences.

Microsoft 365 Copilot and Copilot Chat rule configured to protect items with sensitivity labels supports:

- File items, which are stored and items that are actively open. For more information on supported file types, see: file types supported by sensitivity labels.
- Emails sent on or after January 1, 2025.
- Calendar invites aren't supported.

> **Note**
>
> When a file is open in Word, Excel, or PowerPoint and has a sensitivity label for which DLP policy is configured to prevent processing by Microsoft 365 Copilot and Copilot Chat, the skills in these apps are disabled. Certain experiences that don't reference file content or that aren't using any large language models aren't currently blocked on the user experience.

Block external email from being processed (preview)

This feature is in preview. It applies to Microsoft 365 Copilot and Copilot Chat, including email summarization and reasoning experiences that use email data.

To set this up, create DLP policies that use the **Microsoft 365 Copilot and Copilot Chat** policy location with the **Email is received from** > **External users** condition. When the policy detects that an email was received from a sender outside your organization's accepted domains, Copilot excludes that email from grounding, summarization, and citation. Copilot continues to use internal Microsoft 365 data sources where permitted, and the user's access to the email itself isn't affected.

The policy evaluates email metadata only - specifically, the sender domain compared against your tenant's accepted domains. The body of the email isn't inspected.

Block external email use case example

Contoso wants employees to keep using Microsoft 365 Copilot for productivity, but is concerned that external email could carry untrusted instructions or prompt-injection content. They want Copilot to ground responses only in trusted internal data.

To meet this need, Contoso creates a DLP policy that targets the Microsoft 365 Copilot and Copilot Chat location with a rule that uses the **Email is received from** > **External users** condition. The action is configured to **Prevent Copilot from processing content**.

When a user prompts Copilot to summarize their inbox or reason over recent email, external emails are excluded from grounding. Copilot returns a response based on internal email and other permitted Microsoft 365 sources, and the user sees a message that some content was excluded by an organizational policy.

Files uploaded in prompts

You can upload files when you craft a prompt for Microsoft 365 Copilot to analyze. DLP can't scan the contents of files that you upload directly into prompts, so evaluation of the uploaded file for sensitive data doesn't occur. DLP only checks the text you type into the prompt itself.

Supported conditions and actions

The **Microsoft 365 Copilot and Copilot Chat** policy location supports the following conditions and actions:

Expand table

Conditions	Supported policy actions	Description
Content contains > Sensitivity labels	**Prevent Copilot from processing content**	Detects when a file or an email in Exchange has a chosen sensitivity label. Copilot and Cowork don't process the content of the item or use it in the response summary, but the item could be available in the citations of the response.
Content contains > Sensitive information types	**Prevent Copilot from processing content > Processing prompts**	Detects when the text entered directly into a Copilot or Cowork prompt contains chosen sensitive information types. The experience doesn't respond to the prompt. The prompt isn't used for internal or web searches.
Content contains > Sensitive information types	**Prevent Copilot from processing content > Performing Web Searches**	Detects when the text entered directly into a Copilot or Cowork prompt contains chosen sensitive information types. The experience blocks the use of external web search as a grounding source for that prompt.
Email is received from > External users	**Prevent Copilot from processing content**	Detects when an email was received from a sender outside your organization's accepted domains. The external email is excluded from being used by Copilot for grounding, summarization, or citation. Email body content isn't inspected; only sender metadata is evaluated.

> **Note**

All Microsoft 365 Copilot prompts run in the security context of the user who initiates the prompt. This means for a user to see an item in a prompt response, they must first have the necessary permissions to access the content of the item. You can then use the **Microsoft 365 Copilot and Copilot Chat** policy location feature to exclude items from being processed in the response summary.

In Word, Excel, and PowerPoint the **Microsoft 365 Copilot and Copilot Chat** policy to prevent Copilot from processing content is evaluated at file open. If a sensitivity label is applied mid-session, the policy will be enforced starting the next time the file is opened.

Customize the link in Copilot policy tips

When a DLP rule blocks a Copilot interaction, the standard Copilot message includes a **Learn about access restrictions** link. By default, the link opens Microsoft Learn. You can configure the rule so that the link opens an existing page that explains your organization's policies or tells users where to get help.

> **Tip**
>
> Using a URL that clearly explains why the policy is in place can help users understand the reason for a block.

You don't need to create a new website for this setting. You can use an existing SharePoint page, help desk page, compliance site, or other webpage that the affected users have permission to access. Microsoft Purview doesn't create or host the destination page.

To customize the link:

1. Create or edit a DLP policy for the **Microsoft Copilot and Chat** location.
2. In the rule's user notification settings, turn on **Policy Tips**.
3. Select **Provide a compliance URL for the end user to learn more about your organization's policies**.
4. Enter the full URL of the existing page, including `https://` and the domain name. For example, enter `https://contoso.sharepoint.com/sites/compliance/data-handling`, not `/sites/compliance/data-handling`.
5. Save the rule.

When the rule blocks an interaction, the **Learn about access restrictions** link opens the page that you specified. Only the link destination changes. The standard Copilot block message, the link text, and the DLP enforcement action remain unchanged. If you don't configure a compliance URL, the link continues to open Microsoft Learn.

Availability

- The **Microsoft 365 Copilot and Copilot Chat** policy location is only available in the **Custom** policy template.
- When you select the **Microsoft 365 Copilot and Copilot Chat** policy location, all other locations for that policy are disabled.
- DLP alerts, DLP notifications, and policy simulation mode are supported.
- Updates to a DLP policy can take up to four hours to reflect in Microsoft 365 Copilot and Copilot Chat experience.

Admin units

The **Microsoft 365 Copilot and Copilot Chat** policy location doesn't support **Admin units**.

Related content

- Learn about data loss prevention
- Plan for data loss prevention (DLP)
- Design a data loss prevention policy
- Create and Deploy data loss prevention policies

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 10/05/2026

Learn about the default data loss prevention policy for Microsoft 365 Copilot location

To help you get started with Microsoft Purview data loss prevention (DLP) for the Microsoft 365 Copilot location, Microsoft created a default DLP policy called **Default DLP policy - Protect sensitive M365 Copilot interactions**. The policy is referenced in the **Safeguard sensitive data in Copilot interactions** banner.

> ⓘ **Note**
>
> See these message center announcements MC1182689 and MC1181998.

> ⓘ **Important**
>
> The default policy detects sensitive information types (SIT) in user prompts. The **Default DLP policy - Protect sensitive M365 Copilot interactions** is initially configured to **Run the policy in simulation mode** > **Show policy tips while in simulation mode**. This means that the event will only be logged. If you want to block processing of the prompt, you must change the state of the policy to enforce mode.

You can edit this policy to better suit your organization's needs. To edit the default policy for better performance or to delete it, use an account with **DLP Compliance Management** permissions. For more information, see Permissions. For example, you can:

- Change the **Policy mode** to **Turn the policy on immediately** which enforces the block prompt processing action or you can **Leave the policy turned off**.
- Delete it if you don't need it.
- Add or remove sensitive information types to the policy.
- Add or remove users and groups.
- Remove sensitive information types that aren't relevant to your organization to reduce false positives and improve performance.

You can perform these actions through both the simplified experience by selecting **Get started** in the banner and through the full policy editor flow by selecting the policy in the list of policies and selecting the edit button.

Microsoft recommends that you at least edit the policy to turn on **Incident reports** and add your security team as recipients of the email notifications, so they're aware of potential

sensitive information being shared in the prompts.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Licensing

For information on licensing, see

- Microsoft 365 Enterprise Plans
- Microsoft 365 Service Descriptions

What does the policy do?

The policy is in simulation mode by default, which means it doesn't block the prompt but generates an alert and sends an email to the admin (if configured). This approach helps you understand the impact of the policy and make necessary adjustments before enforcing it. It initially includes all tenant users and groups in scope and detects these SITs:

- ABA Routing Number
- Argentina National Identity (DNI) Number
- Argentina Unique Tax Identification Key (CUIT/CUIL)
- Australia Bank Account Number
- Australia Driver's License Number
- Australia Medical Account Number
- Australia Passport Number
- Australia Tax File Number
- Australian Business Number
- Australian Company Number
- Austria Tax Identification Number
- Azure SAS
- Azure Storage Account Key (Generic)
- Belgium National Number
- Brazil CPF Number
- Brazil Legal Entity Number (CNPJ)
- Brazil National ID Card (RG)
- Canada Bank Account Number

- Canada Driver's License Number
- Canada Health Service Number
- Canada Personal Health Identification Number (PHIN)
- Canada Social Insurance Number
- Chile Identity Card Number
- Colombia National ID
- Credit Card Number
- Croatia Personal Identification (OIB) Number
- Czech Personal Identity Number
- Denmark Personal Identification Number
- Drug Enforcement Agency (DEA) Number
- Estonia Personal Identification Code
- Finland National ID
- France National ID Card (CNI)
- France Social Security Number (INSEE)
- Germany Identity Card Number
- Germany Tax Identification Number
- Germany Value Added Tax Number
- Greek Tax identification Number
- Hungarian Social Security Number (TAJ)
- India GST Number
- India Permanent Account Number (PAN)
- India Unique Identification (Aadhaar) Number
- International Banking Account Number (IBAN)
- IP Address
- IP Address v4
- IP Address v6
- Ireland Passport Number
- Israel Bank Account Number
- Italy Fiscal Code
- Italy Passport Number
- Japan Bank Account Number
- Japan Passport Number
- Japanese My Number Personal
- Latvia Personal Code
- Lithuania Personal Code
- Luxembourg National Identification Number (Non-natural persons)
- Malaysia Identity Card Number
- Medicare Beneficiary Identifier (MBI) card
- Mexico Unique Population Registry Code (CURP)
- Netherlands Citizen's Service (BSN) Number

- Netherlands Tax Identification Number
- Netherlands Value Added Tax Number
- New Zealand bank account number
- New Zealand Inland Revenue number
- New Zealand Ministry of Health Number
- New Zealand Social Welfare Number
- Norway Identity Number
- Poland National ID (PESEL)
- Poland Tax Identification Number
- Polish REGON Number
- Portugal Tax Identification Number
- Saudi Arabia National ID
- Singapore National Registration Identity Card (NRIC) Number
- Slovakia Personal Number
- Slovenia Tax Identification Number
- South Africa Identification Number
- Spain DNI
- Spain Driver's License Number
- Spain Tax Identification Number
- Sweden National ID
- SWIFT Code
- Taiwan Passport Number
- Thai Population Identification Code
- Turkish National Identification number
- U.K. National Health Service Number
- U.K. National Insurance Number (NINO)
- U.K. Unique Taxpayer Reference Number
- U.S. / U.K. Passport Number
- U.S. Bank Account Number
- U.S. Driver's License Number
- U.S. Individual Taxpayer Identification Number (ITIN)
- U.S. Social Security Number (SSN)

The policy doesn't generate any policy tips for end users, but it does generate an Alert event. It also triggers a low severity email to the admin (added in the policy). Administrators can view the activities and edit the policy details by signing in to the Microsoft Purview portal.

Admins can view this policy in the Microsoft Purview portal. Go to **Data Loss prevention** > **Policies**.

See also

- Learn about using Microsoft Purview Data Loss Prevention to protect interactions with Microsoft 365 Copilot and Copilot Chat

---

Last updated on 03/24/2026

Get started with data loss prevention for on-premises repositories

> **Note**
>
> There's a new version of the information protection scanner. For more information, see Upgrade the Microsoft Purview Information Protection scanner.

This article walks you through the prerequisites and configuration for using the Microsoft Purview Data Loss Prevention on-premises repositories location in a DLP policy.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Before you begin

Licensing

For information on licensing, see

- Microsoft 365 Enterprise Plans
- Microsoft 365 Service Descriptions

> **Important**
>
> All users who contribute to the scanned location, either by adding files or consuming files, need to have a license, not just the scanner user.

Permissions

Data from DLP can be viewed in activity explorer. There are four roles that grant permission to activity explorer, the account you use for accessing the data must be a member of any one of them.

- Compliance administrator
- Security administrator
- Compliance data administrator
- Global administrator

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

Roles and Role Groups

There are roles and role groups in the Microsoft Purview portal that you can test to fine-tune your access controls.

Here's a list of applicable roles. To learn more about them, see Permissions in the Microsoft Purview portal.

- Information Protection Admin
- Information Protection Analyst
- Information Protection Investigator
- Information Protection Reader

Here's a list of applicable role groups. To learn more, see Permissions in the Microsoft Purview portal.

- Information Protection
- Information Protection Admins
- Information Protection Analysts
- Information Protection Investigators
- Information Protection Readers

DLP on-premises repositories prerequisites

Before you deploy DLP on-premises repositories, ensure that the following prerequisites are met:

- The Microsoft Purview Information Protection scanner implements DLP policy matching and policy enforcement. The scanner is installed as part of the information protection client, so

your installation must meet all the prerequisites for the rights management encryption service, the information protection client, and the information protection scanner.

- Deploy the client and scanner. For more information, see, Install or upgrade the information protection client and, Configuring and installing the information protection scanner.
- There must be at least one label and policy published in the tenant, even if all your detection rules are based on sensitive information types only.

Deploy the DLP on-premises scanner

Perform the following steps to deploy the DLP on-premises scanner:

1. Sign in to the Microsoft Purview portal.

2. Follow the procedures in Install or upgrade the information protection client.

3. Follow the procedures in Configuring and installing the information protection scanner to complete the scanner installation.
   a. You must create a content scan job and specify the repositories that host the files to be evaluated by the DLP engine.
   b. Enable DLP rules in the created content scan job, and set the **Enforce** option to **Off** (unless you want to proceed directly to the DLP enforcement stage).

4. Verify that your content scan job is assigned to the right cluster. If you haven't created a content scan job, create a new one and assign it to the cluster that contains the scanner nodes.

5. Connect to the Microsoft Purview portal and add your repositories to the content scan job that will perform the scan.

6. Do one of the following to run your scan:
   a. Set the scanner schedule
   b. Use the manual **Scan Now** option in the portal
   c. Run **Start-Scan** PowerShell cmdlet

   > **Important**
   >
   > Remember that the scanner runs a delta scan of the repository by default and files that were scanned in the previous scan cycle will be skipped, unless the file was changed or you initiated a full rescan. A full rescan can be initiated by using the **Rescan all files** option in the UI or by running **Start-Scan -Reset**.

7. Open the Microsoft Purview portal > Data Loss Prevention > Policies.

8. Choose + **Create policy** and create a test DLP policy. See Create and Deploy data loss prevention policies if you need help with creating a policy. Be sure to **Run the policy in simulation mode** until you're comfortable with this feature. Use these parameters for your policy:
   a. Scope the DLP on-premises repositories rule to specific locations if needed. If you scope **locations** to **All**, all files scanned will be subject to the DLP rule matching and enforcement.
   b. When specifying the locations, you can use either exclusion or inclusion list. You can either define that the rule is relevant only to paths matching one of the patterns listed in inclusion list or, all files, except the files matching the pattern listed in inclusion list. No local paths are supported. Here are some examples of valid paths:

      - \\server\share
      - \\server\share\folder1\subfolderabc
      - *\folder1
      - *secret*.docx
      - *secret*.*
      - https:// sp2010.local/sites/HR
      - https://*/HR

   a. Here are some examples of unacceptable values use:

      - *
      - *\a
      - Aaa
      - c:\
      - C:\test

> **Important**
>
> The *exclusion* list takes precedence over the *inclusions* list.

Viewing DLP alerts

Use the following steps to view alerts for your on-premises DLP policies:

1. Open Data loss prevention in the Microsoft Purview portal and select **Alerts**.

2. Refer to the procedures in Get started with the data loss prevention Alerts dashboard and Investigate data loss incidents with Microsoft Defender XDR to view alerts for your on-premises DLP policies.

Viewing DLP data in activity explorer and audit log

> **Note**
>
> The Information Protection scanner requires that auditing be enabled. Auditing is enabled by default in Microsoft 365.

1. Open the Data classification page for your domain in the Microsoft Purview portal and select **Activity explorer**.
2. Refer to the procedures in Get started with Activity explorer to access and filter all the data for your on-premises scanner locations.
3. Open the Audit log. The DLP rule matches are available in the Audit log UI or accessible by Search-UnifiedAuditLog in PowerShell.

Next steps

Now that you've deployed a test policy for DLP on-premises locations and can view the activity data in Activity explorer, create production DLP policies to protect your sensitive items.

- Using DLP on-premises

See also

- Learn about the data loss prevention on-premises repositories
- Use the data loss prevention on-premises repositories
- Learn about data loss prevention
- Create and Deploy data loss prevention policies
- Get started with Activity explorer
- Microsoft 365 subscription

Note: The author created this article with assistance from AI. Learn more

Last updated on 06/26/2026

Use the data loss prevention on-premises repositories location

To help familiarize you with Microsoft Purview Data Loss Prevention on-premises features and how they surface in DLP policies, we've put together a couple of scenarios for you to follow. These scenarios show you how to discover files that match DLP rules across Activity Explorer, the audit log, and the information protection scanner, and how to enforce protective actions on on-premises file shares and SharePoint document libraries. Specifically, the first scenario covers where DLP rule-match data appears when scanning on-premises repositories, and the second scenario walks through enabling policy enforcement so that protective actions are applied to scanned files.

> **Important**
>
> These DLP on-premises scenarios are not the official procedures for creating and tuning DLP policies. Refer to the following topics when you need to work with DLP policies in general situations:
>
> - Learn about data loss prevention
> - Create and Deploy data loss prevention policies

Scenario: Discover files matching DLP rules

Data from DLP surfaces in several areas:

View on-premises repository events in Activity Explorer

DLP reports rule matches are available in Activity Explorer.

View on-premises repository events in the Microsoft 365 audit log

The DLP rule matches are also available in the Audit log UI (see Search the audit log) and are accessible via PowerShell through theSearch-UnifiedAuditLog.

Use the information protection scanner to inspect matching files

Discovery data is available in a local report in .csv format and is stored under:

%localappdata%\Microsoft\MSIP\Scanner\Reports\DetailedReport_%timestamp%.csv report.

Look for the following columns:

- DLP Mode
- DLP Status
- DLP Comment
- DLP Rule Name
- DLP Actions
- Owner
- Current NTFS Permissions (SDDL)
- Applied NTFS Permissions (SDDL)
- NTFS permissions type

Scenario: Enforce DLP rule

If you want to enforce DLP rules on scanned files, enforcement must be enabled both on the content scan job and at the policy level in DLP.

Configure DLP to enforce policy actions

To turn on enforcement for your DLP policy targeting on-premises repositories, complete the following steps:

1. Sign in to the Microsoft Purview portal
2. Navigate to **Data loss prevention** > **Policies**.
3. Select the DLP policy that is targeted to the on-premises location repositories you have configured for the scanner.
4. Edit the policy.
5. On the **Policy Mode** page, select **Turn the policy on immediately**.
6. Choose **Next** and then choose **Submit**.
7. Choose **Done**.

See also

- Learn about the data loss prevention on-premises repositories
- Get started with data loss prevention on-premises repositories
- Learn about data loss prevention
- Create and Deploy data loss prevention policies
- Get started with Activity explorer

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Get started with data loss prevention policies for Fabric and Power BI

This article provides a general overview of Microsoft Purview Data Loss Prevention (DLP) policies for Fabric and Power BI. The target audience is Fabric administrators, security and compliance teams, and Fabric data owners. If you're a data owner and want to know how to respond when a policy tip tells you your item has a DLP policy match, see Respond to a DLP policy match in Fabric. If you're a Fabric admin or a security and compliance admin and need to audit alerts on DLP policy matches, see Monitor DLP policy matches in Fabric.

Overview

To help organizations detect and protect their sensitive data, Fabric supports Microsoft Purview Data Loss Prevention (DLP) policies. When a DLP policy for Fabric detects a supported item type containing sensitive information, the policy triggers the configured actions. These actions can include:

- Attaching a policy tip to the item that explains the nature of the sensitive content.
- Registering an alert for administrators on the data loss prevention **Alerts** page in the Microsoft Purview portal.
- Sending email alerts to administrators and specified users.
- Restricting access to the item.

For more detail, see How do DLP policies for Fabric and Power BI work.

For DLP policies, a *member of the organization* means a user with UserType = Member. This includes users who sign in with an external account. Because their UserType is Member, they're treated as internal users. In contrast, users with UserType = Guest are considered external. The **Block only people outside your organization** setting applies to these Guest users.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Licensing and permissions

Licensing

For information on licensing, see

- Microsoft 365 Enterprise Plans
- Microsoft 365 Service Descriptions

Permissions

You can view data from DLP for Fabric and Power BI in Activity explorer. Four roles grant permission to Activity explorer; the account you use for accessing the data must be a member of any one of them.

To view the Activity explorer, your account must be a member of any of the following roles:

- Compliance administrator
- Security administrator
- Compliance data administrator

Billing

DLP evaluation workloads impact capacity consumption. For information about how this consumption is metered and billed, see Learn about Microsoft Purview billing models.

How DLP policies for Fabric and Power BI work

You define a DLP policy in the data loss prevention section of the Microsoft Purview portal. In the policy, you specify the conditions, such as which sensitivity labels and sensitive info types you want to detect. You also specify the actions that the system takes when it detects a policy match.

When DLP policies evaluate a supported item type, they check if it matches the conditions in a DLP policy. If it does, the policy's specified actions take place. DLP policies start with the following actions:

Semantic models

DLP policies evaluate a semantic model whenever one of the following events occurs:

- Publish
- Republish
- On-demand refresh

- Scheduled refresh

> **Note**
>
> DLP evaluation of the semantic model doesn't occur if either of the following is true:
>
> - An account using service principal authentication initiates the event (publish, republish, on-demand refresh, scheduled refresh).
> - The semantic model owner is a service principal.

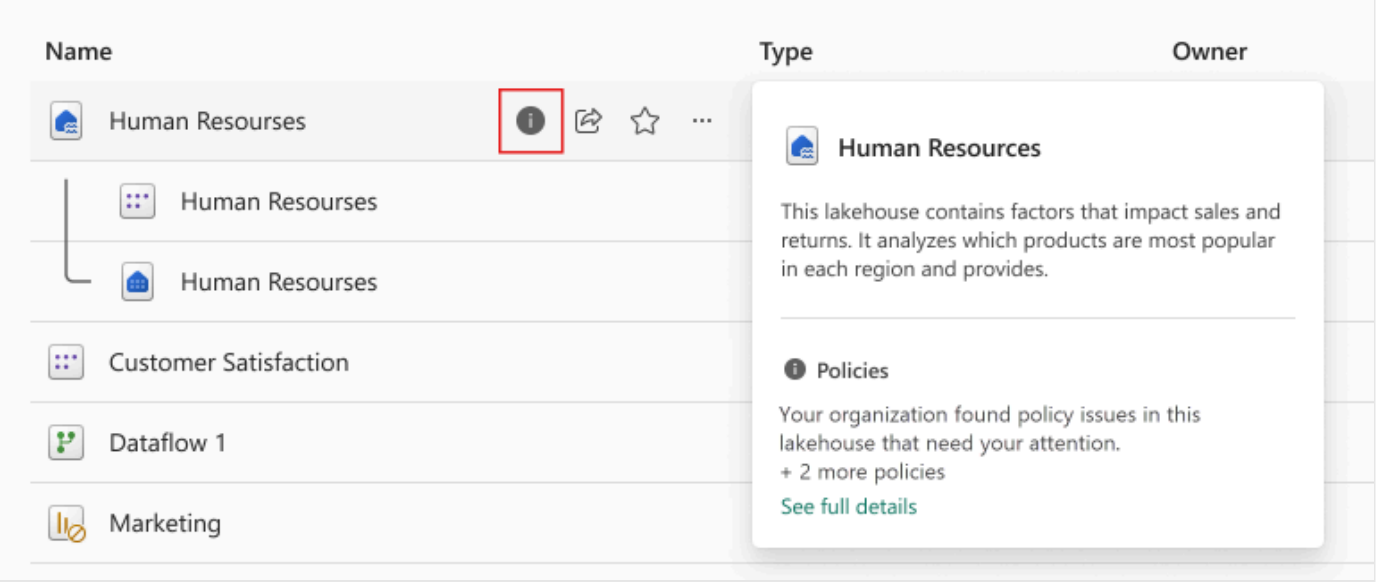
Fabric Items

DLP policies evaluate a Fabric item, such as a lakehouse, SQL database, or Mirrored database, when the data within it changes. Changes include getting new data, connecting a new source, adding tables, updating existing tables, and more.

What happens when a Fabric DLP policy flags an item

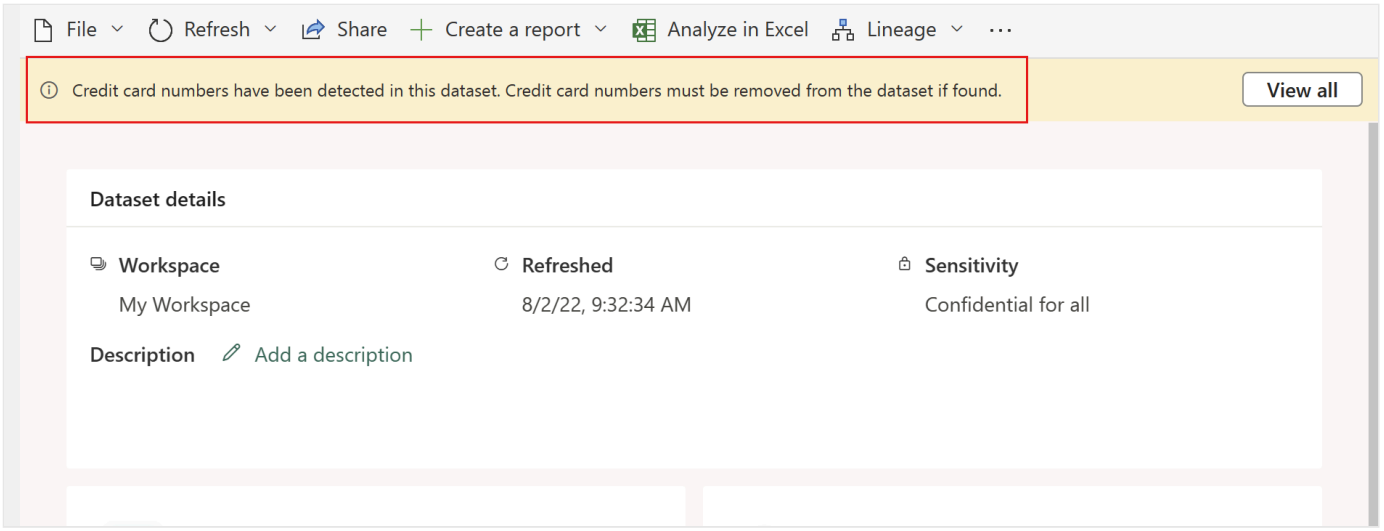
When a DLP policy detects an issue with an item, it takes the following actions:

If you enable **user notification** in the policy, Fabric marks the item with an icon that indicates a DLP policy detected an issue with the item. When you hover over the icon, a hover card appears that provides an option to see the full details in a side panel. For more information about what you see in the side panel, see Respond to a DLP violation in Fabric.



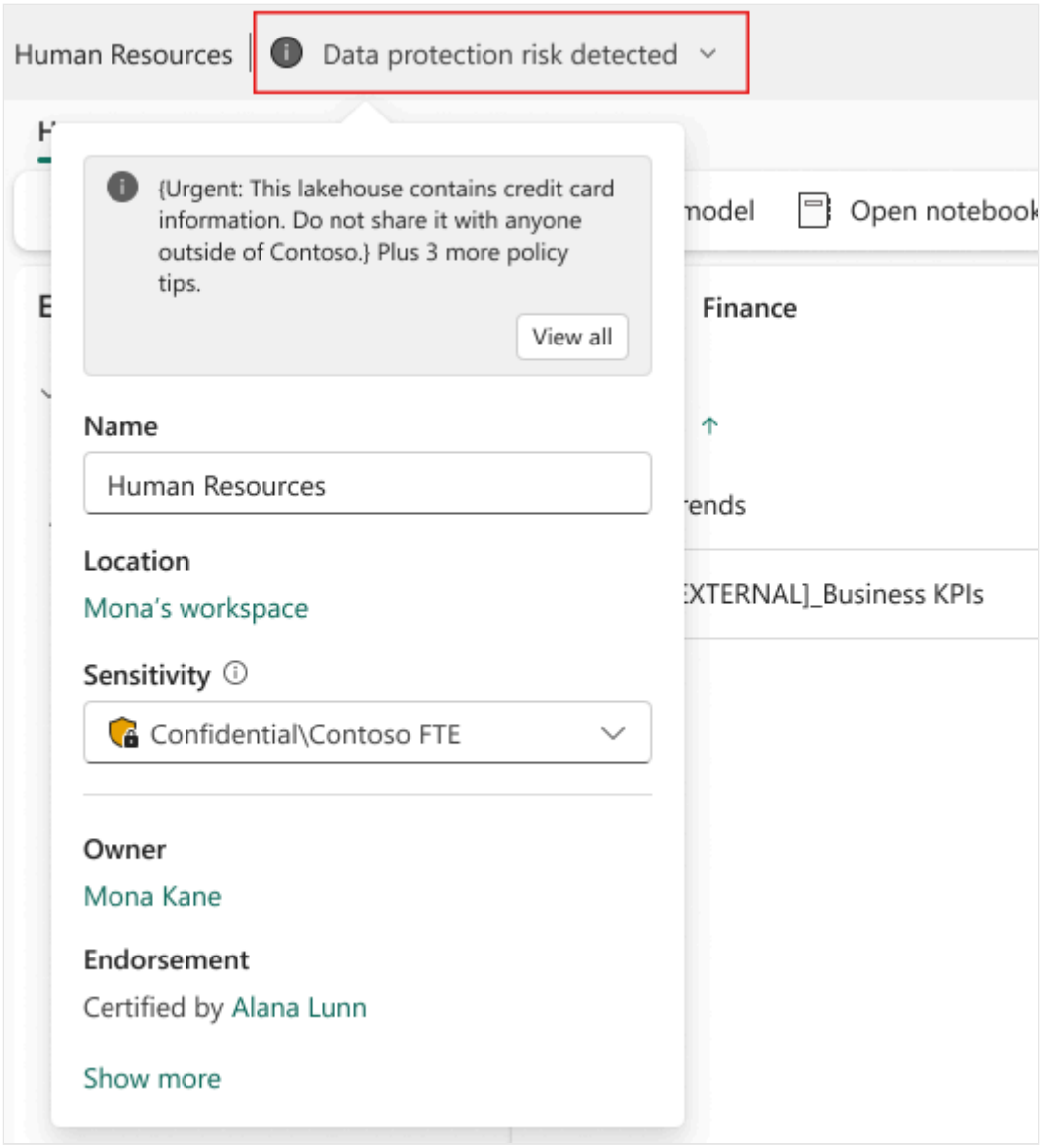
For semantic models, opening the details page shows a policy tip that explains the policy violation and how to handle the type of sensitive information detected. Selecting **View all** opens

a side panel with all the policy details.



> **Note**
>
> If you hide the policy tip, it doesn't get deleted. It appears the next time you visit the page.

For Fabric items, such as lakehouses, the indication appears in the header in edit mode. Opening the flyout makes it possible to see more details about the policy tips affecting the item. Selecting **View all** opens a side panel with all the policy details.



If you enable alerts in the policy, an alert records on the data loss prevention **Alerts** page in the Microsoft Purview portal. If configured, an email is sent to administrators and/or specified users. For more information, see Monitor and manage DLP policy violations.

Supported actions

When Fabric items or semantic models meet the conditions in a DLP policy, the actions defined in the policy are applied. DLP policies for Fabric and Power BI support three actions:

- Notify the user through a policy tip.
- Generate an alert. Administrators and users can receive alert emails. Additionally, administrators can monitor and manage alerts on the **Alerts** tab in the Purview portal.
- Restrict access (preview). When you configure a policy with the restrict access action and a policy match occurs, the policy restricts access to the data owners or to members of the organization, depending on the policy configuration. All other users lose access to the item.

For information about what triggers DLP evaluation, see How DLP policies for Fabric and Power BI work.

Supported item types

DLP policies for Fabric and Power BI currently support the following item types.

- Semantic models
- Lakehouses
- Warehouses
- KQL databases
- Mirrored databases
- SQL databases
- Cosmos databases

See Considerations and limitations for exceptions.

Supported condition types

DLP policy rules for Fabric and Power BI support sensitivity labels and a subset of sensitive info types (see considerations and limitations) as conditions. Learn more about Purview Sensitive Info Types and confidence levels.

Configure a DLP policy for Fabric and Power BI

For information about creating a DLP policy for Fabric or Power BI, see Help prevent sharing Power BI reports with credit card numbers and adapt it to your own scenario.

Gain visibility into DLP activity and adoption

After deploying DLP policies, administrators can use Fabric governance experiences in the OneLake catalog to understand how data loss prevention policies are being used across their organization. These insights can help organizations:

- Identify workspaces and items evaluated by DLP policies.
- Understand where sensitive information is being detected.
- Track policy adoption across business units and workspaces.
- Prioritize remediation of high-risk data assets.
- Measure the effectiveness of DLP investments and governance programs.

Together, DLP insights help security and compliance teams move beyond individual policy violations and understand their overall data protection posture.

Considerations and limitations

- DLP policies apply to workspaces. Only workspaces hosted in Fabric or Premium capacities are supported. For more information, see Microsoft Fabric concepts and licenses.
- DLP policy templates aren't supported for Fabric DLP policies. When creating a DLP policy for Fabric, choose the custom policy option.
- DLP policies for Fabric aren't supported for sample semantic models, streaming datasets, or semantic models that connect to their data source via DirectQuery or live connection. This limitation includes semantic models with mixed storage, where some of the data comes via import-mode and some comes via DirectQuery.
- DLP policies for Fabric apply only on data in tables stored in Delta format.
- Delta tables are treated as text files for the purpose of scanning considerations.
- DLP policies for Fabric support all the primitive Delta types except *timestamp_ntz*.
- DLP policies for Fabric aren't supported for the following Delta Parquet data types:
  - Binary, timestamp_ntz, Struct, Array, List, Map, Json, Enum, Interval, Void.
  - Data with LZ4, Zstd, and Gzip compression codecs.
- Advanced classifiers, such as exact data match (EDM), trainable classifiers, credential classifiers, and named entities aren't supported by DLP for Fabric. If you select an advanced classifier in the condition of a policy, the policy returns no results even if the item contains data that satisfies the advanced classifier. Other classifiers specified in the policy return results if any.
- DLP policies for Fabric aren't supported in the China North region. See How to find the default region for your organization to learn how to find your organization's default data region.
- Onboarding a new tenant to DLP can take a few hours, depending on the number of supported workspaces that are being onboarded.

See also

- Learn about data loss prevention
- Sensitivity labels in Power BI
- Audit schema for sensitivity labels in Power BI
- Configure a DLP policy for Fabric.
- Respond to a DLP policy violation in Fabric.
- Monitor and manage DLP policy violations

- Learn about data loss prevention

Last updated on 07/22/2026

Get started with oversharing pop-ups in Outlook

When you configure appropriate Microsoft Purview Data Loss Prevention (DLP) policy, DLP checks email messages before they're sent for any labeled or sensitive information and apply the actions defined in the DLP policy. Oversharing pop-ups require a Microsoft 365 **E5 subscription**, along with a version of Outlook that supports oversharing pop-ups. For more information in required version, see Oversharing dialog for Outlook for Microsoft 365.

> **Important**
>
> The following is a hypothetical scenario with hypothetical values. It's only for illustrative purposes. You should substitute your own sensitive information types, sensitivity labels, distribution groups, and users when implementing this feature.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Before you begin

Licensing

For information on licensing, see

- Microsoft 365 Enterprise Plans
- Microsoft 365 Service Descriptions

Permissions

The account you use to create and deploy policies must be a member of one of the following role groups

- Compliance administrator
- Compliance data administrator

- Information Protection
- Information Protection Admin
- Security administrator

> **Important**
>
> Read **Administrative units** before you start to make sure you understand the difference between an *unrestricted administrator* and an *administrative unit restricted administrator*.

Granular roles and role groups

There are several roles and role groups that you can use to fine tune access controls for Microsoft Purview DLP policy management.

Here's a list of applicable roles. To learn more, see Permissions in the Microsoft Purview portal.

- DLP Compliance Management
- Information Protection Admin
- Information Protection Analyst
- Information Protection Investigator
- Information Protection Reader

Here's a list of applicable role groups. To learn more about them, see Permissions in the Microsoft Purview portal.

- Information Protection
- Information Protection Admins
- Information Protection Analysts
- Information Protection Investigators
- Information Protection Readers

Prerequisites and assumptions

In Outlook for Microsoft 365, an oversharing pop-up displays a pop-up before a message is sent. To enable these pop-ups, first scope your policy to the Exchange location and then select the **Show policy tip as a dialog for the user before send** option in the policy tip when you create a data loss prevention (DLP) rule for that policy.

Our example scenario uses the *Highly confidential* sensitivity label, so it requires that you have created and published sensitivity labels. To learn more, see:

- Learn about sensitivity labels
- Get started with sensitivity labels
- Create and configure sensitivity labels and their policies

This procedure uses *contoso.com*. a hypothetical company domain.

Policy intent statement and mapping

For this example, our policy intent statement is:

We need to block emails to all recipients that have the 'highly confidential' sensitivity label applied, unless the recipient domain is contoso.com. We want to notify the user with a pop-up dialogue when they send the email. No users can be allowed to override the block.

Expand table

Statement	Configuration question answered and configuration mapping
"We need to block emails to all recipients..."	- **Where to monitor**: Exchange - **Administrative scope**: Full directory - **Action**: Restrict access or encrypt the content in Microsoft 365 locations > Block users from receiving email or accessing shared SharePoint, OneDrive, and Teams files > Block everyone
"...that have the 'highly confidential' sensitivity label applied..."	- **What to monitor**: use the Custom template - **Conditions for a match**: edit it to add the *highly confidential* sensitivity label
"...unless..."	**Condition group configuration** - Create a nested boolean NOT condition group joined to the first conditions using a boolean AND
"...the recipient domain is contoso.com."	**Condition for match**: Recipient domain is
"...Notify..."	**User notifications**: enabled
"...the user with a pop-up dialogue when they send..."	**Policy tips**: selected - **Show policy tip as a dialog for the end user before send**: selected
"...No users can be allowed to override the block...	**Allow overrides from M365 Services**: not selected

To configure oversharing pop-ups with default text, the DLP rule must include these conditions:

- **Content contains** > **Sensitivity labels** > **choose your sensitivity label(s)**

and one or more of the following recipient-based conditions

- Recipient is
- Recipient is a member of
- Recipient domain is

When these conditions are met, the policy tip displays untrusted recipients while the user is writing the mail in Outlook, before they send it.

If the block rules defined in the DLP policy are met, the message will be blocked regardless of how **Policy tips** is configured.

Steps to configure "wait on send"

Optionally, you can set the *dlpwaitonsendtimeout Regkey (Value in dword)* on all the devices where you want to implement "wait on send" for oversharing pop-ups. This registry key (*RegKey*) defines the maximum amount of time to hold the email when a user selects **Send**. This allows the system to complete the DLP policy evaluation for labeled or sensitive content. You can find this RegKey under:

Software\Policies\Microsoft\office\16.0\Outlook\options\Mail

You can set this *RegKey* via group policy, **Specify wait time to evaluate sensitive content**, script, or other mechanism for configuring registry keys.

If you're using Group Policy, make sure you download the most recent version of the Group Policy Administrative Template files for Microsoft 365 Apps for enterprise and then navigate to **User Configuration** > **Administrative Templates** > **Microsoft Office 2016** > **Security Settings** > **Specify wait time to evaluate sensitive content**. If you're using the Cloud Policy service for Microsoft 365, search for the setting by name to configure it.

When this value is set and the DLP policy is configured, email messages are checked for sensitive information before they're sent. If a message contains a match to the conditions defined in the policy, a policy tip notification appears before the user select **Send**..

The *dlpwaitonsendtimeout* registry key allows you to specify the **wait on send** behavior for your Outlook clients.

Here's what each *dlpwaitonsendtimeout* setting means:

Not configured or **Disabled**: This is the default. When *dlpwaitonsendtimeout* isn't configured, the message isn't checked before the user sends it. The email message is sent immediately once

Send is clicked. The DLP data classification service evaluates the message and apply the actions defined in the DLP policy.

Enabled: The email message is checked when the **Send** is clicked but before the message is sent. You can set a time limit on how long to wait for DLP policy evaluation to complete (**T** value, in seconds). If the policy evaluation doesn't complete in the specified time a **Send anyway** button appears allowing the user to bypass the presend check. The **T** value range is 0 to 9999 seconds.

> **Important**
>
> If the **T** value is greater than 9999, it's replaced with 10000 and the **Send Anyway** button doesn't appear. This holds the message until the policy evaluation completes and doesn't provide the user with an *override* option. The duration for completing the evaluation can vary depending on factors such as internet speed, content length, and the number of policies defined. Some users may encounter policy evaluation messages more frequently than others, depending on the policies that are deployed on their mailbox.

To learn more about configuring and using GPO see, Administer Group Policy in a Microsoft Entra Domain Services managed domain.

Steps to create a DLP policy for an oversharing pop-up

Perform the following steps to create a DLP policy that triggers an oversharing pop-up in Outlook.

1. Sign in to the Microsoft Purview portal > **Data loss prevention** > **Policies**
2. Choose + **Create policy**.
3. Select **Custom** from the **Categories** list.
4. Select **Custom** from the **Regulations** list.
5. Give the policy a name.
6. Fill in a description. You can use the policy intent statement here.
7. Select **Next**.
8. Select **Full directory** under **Admin units**.
9. Select only the **Exchange email** location.

10. Select **Next**.
11. On the **Define policy settings** page, select **Create or customize advanced DLP rules**.
12. The **Create or customize advanced DLP rules** option should already be selected.
13. Select **Next**.
14. Select **Create rule**. Name the rule and provide a description.
15. Select **Add condition** > **Content contains** > **Add** > **Sensitivity labels** > **Highly confidential**. Choose **Add**.
16. Select **Add group** > **AND** > **NOT** > **Add condition**.
17. Select **Recipient domain is** > **contoso.com**. Choose **Add**.

    > **Tip**
    >
    > You can also use **Recipient is** or **Recipient is a member of** instead of **Recipient domain is** to trigger an oversharing pop-up.

18. Select **Add an action** > **Restrict access or encrypt the content in Microsoft 365 locations**.
19. Select **Block users from receiving email or accessing shared SharePoint, OneDrive, and Teams files, and Power BI items**.
20. Select **Block everyone**.
21. Set the **User notifications** toggle to **On**.
22. Select **Policy tips** > **Show the policy tip as a dialog for the end user before send (available for Exchange workload only)**.
23. *If already selected*, uncheck the **Allow override from M365 services** option.
24. Choose **Save**.
25. Change the **Status** toggle to **On** and then choose **Next**.
26. On the **Policy mode** page, select **Run the policy in test mode** and check the box for the **Show policy tips while in simulation mode** option.
27. Choose **Next** and then choose **Submit**.

28. Choose **Done**.

PowerShell steps to create policy

DLP policies and rules can also be configured in PowerShell. To configure oversharing pop-ups using PowerShell, first you create a DLP policy (using PowerShell) and add DLP rules for each warn, justify or block pop-up type.

You'll configure and scope your DLP Policy using New-DlpCompliancePolicy. Then, you configure each oversharing rule using New-DlpComplianceRule

To configure a new DLP policy for the oversharing pop-up scenario use this code snippet:

PowerShell

```
PS C:\> New-DlpCompliancePolicy -Name <DLP Policy Name> -ExchangeLocation All
```

This sample DLP policy is scoped to all users in your organization. Scope your DLP Policies using `-ExchangeSenderMemberOf` and `-ExchangeSenderMemberOfException`.

Expand table

Parameter	Configuration
-ContentContainsSensitiveInformation	Configures one or more sensitivity label conditions. This sample includes one. At least one label is mandatory.
-ExceptIfRecipientDomainIs	List of trusted domains.
-NotifyAllowOverride	"WithJustification" enables justification radio buttons, "WithoutJustification" disables them.
-NotifyOverrideRequirements	"WithAcknowledgement" enables the new acknowledgment option. This is optional.

To configure a new DLP rule to generate a *warn* pop-up using trusted domains, run the following PowerShell code:

PowerShell

```
PS C:\> New-DlpComplianceRule -Name <DLP Rule Name> -Policy <DLP Policy Name> -
NotifyUser Owner -NotifyPolicyTipDisplayOption "Dialog" -
ContentContainsSensitiveInformation @(@{operator = "And"; groups =
@(@{operator="Or";name="Default";labels=@(@{name=<Label GUID>;type="Sensitivity"})})})
-ExceptIfRecipientDomainIs @("contoso.com","microsoft.com")
```

To configure a new DLP rule to generate a *justify* pop-up using trusted domains, run this PowerShell code:

PowerShell

```
PS C:\> New-DlpComplianceRule -Name <DLP Rule Name> -Policy <DLP Policy Name> -
NotifyUser Owner -NotifyPolicyTipDisplayOption "Dialog" -BlockAccess $true -
ContentContainsSensitiveInformation @(@{operator = "And"; groups = @(@{operator =
"Or"; name = "Default"; labels = @(@{name=<Label GUID 1>;type="Sensitivity"},@{name=
<Label GUID 2>;type="Sensitivity"})})}) -ExceptIfRecipientDomainIs
@("contoso.com","microsoft.com") -NotifyAllowOverride "WithJustification"
```

To configure a new DLP rule to generate a *block* pop-up using trusted domains, run this PowerShell code:

PowerShell

```
PS C:\> New-DlpComplianceRule -Name <DLP Rule Name> -Policy <DLP Policy Name> -
NotifyUser Owner -NotifyPolicyTipDisplayOption "Dialog" -BlockAccess $true -
ContentContainsSensitiveInformation @(@{operator = "And"; groups = @(@{operator =
"Or"; name = "Default"; labels = @(@{name=<Label GUID 1>;type="Sensitivity"},@{name=
<Label GUID 2>;type="Sensitivity"})})}) -ExceptIfRecipientDomainIs
@("contoso.com","microsoft.com")
```

For more information, see Business justification X-Header.

See also

For related guidance, see the following articles:

- Get started with the data loss prevention Alerts dashboard
- Create and Deploy data loss prevention policies

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Use data loss prevention policies for non-Microsoft cloud apps

> **Important**
>
> The Instances policy location in Microsoft Purview Data Loss Prevention (DLP) and policy creation experiences are being retired on January 6th, 2027, as part of the retirement of Microsoft Defender for Cloud Apps file policies. Recreate your policies in Purview using dedicated policy locations for supported non-Microsoft connected apps before this date. For more information, see Migrate file policies to Microsoft Purview.

You can scope DLP policies to **Instances** to monitor, detect, and take actions when sensitive items are used and shared via non-Microsoft cloud apps.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Before you begin

For information on licensing, see

- Microsoft 365 Enterprise Plans
- Microsoft 365 Service Descriptions

Permissions

The account you use to create and deploy policies must be a member of one of these role groups

- Compliance administrator
- Compliance data administrator
- Information Protection
- Information Protection Admin
- Security administrator

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. This helps improve security for your organization. Global Administrator is a highly privileged role that should only be used in scenarios where a lesser privileged role can't be used.

Prepare your Defender for Cloud Apps environment

Before you configure DLP policies scoped to **Instances**, you must prepare your Defender for Cloud Apps environment. For instructions, see Quickstart: Get started with Microsoft Defender for Cloud Apps.

Connect a non-Microsoft cloud app

To use a DLP policy that's scoped to a specific non-Microsoft cloud app, the app must be connected to Defender for Cloud Apps. For information, see:

- Connect Box
- Connect Dropbox
- Connect Google Workspace
- Connect Salesforce
- Connect Cisco Webex

After you connect your cloud apps to Defender for Cloud Apps, you can create DLP policies for those apps.

Create a DLP policy scoped to a non-Microsoft cloud app

Refer to Create and Deploy data loss prevention policies for the procedures to create a DLP policy. Keep these points in mind as you configure your policy:

1. Sign in to the Microsoft Purview portal.
2. Open **Data loss prevention** > **Policies** > **+ Create policy**.
3. Select **Custom** > **Custom policy**.
4. Name the policy and add a description.

5. Set the **Admin units** scope as desired.

6. On the **Locations** page, toggle the **Instances** location to on.

7. To select a specific app or instance, select **Edit** > **Specific instances** > **+ Include instances** and then the instances (cloud apps) you want to include. If you don't select a instance, the policy will be scoped to all connected apps in your Microsoft Defender for Cloud Apps tenant.

8. **Create rule** with the desired settings. For more information on policy creation, see Create and Deploy data loss prevention policies.
   a. Under **Actions** select **Restrict Third Party Apps** and select any of the apps listed.

> **Note**
>
> When you create a DLP policy that is scoped to **Instances**, the same policy will be automatically created in Microsoft Defender for Cloud Apps.

Related content

- Create and Deploy data loss prevention policies

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 08/28/2026

Use Microsoft Purview data loss prevention policies for non-Microsoft connected apps (preview)

> **Important**
>
> This feature is currently in preview. The Supplemental Terms of Use for Microsoft Azure Previews include additional legal terms that apply to Azure features that are in beta, in preview, or otherwise not yet released into general availability.

Microsoft Purview Data Loss Prevention (DLP) extend to non-Microsoft connected apps, allowing you to detect, monitor, and protect sensitive data at rest in non-Microsoft SaaS applications. You can create DLP policies that apply to data stored in these non-Microsoft apps, using the same classification engine and policy framework available for Microsoft 365 locations.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Supported non-Microsoft apps

Microsoft Purview supports DLP policies for the following non-Microsoft apps:

- Box
- Dropbox
- Google Workspace
- Salesforce

> **Note**
>
> These apps are rolling out in phases. Not all apps may be available in your tenant at the same time. Check the Microsoft 365 roadmap for the latest availability information.

Before you begin

Licensing requirements

For information on licensing, see

- Microsoft 365 Enterprise Plans
- Microsoft 365 Service Descriptions

Permissions

The account you use to create and deploy policies must be a member of one of these role groups:

- Compliance administrator
- Compliance data administrator
- Information Protection
- Information Protection Admin
- Security administrator

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. This helps improve security for your organization. Global Administrator is a highly privileged role that should only be used in scenarios where a lesser privileged role can't be used.

Set up a Microsoft Defender for Cloud Apps connector

Before you can apply DLP policies to a non-Microsoft app, you must connect that app to Microsoft Defender for Cloud Apps using an app connector. Purview uses the existing Microsoft Defender for Cloud Apps connectors to access capabilities in the non-Microsoft app.

For instructions on setting up app connectors, see Connect apps to get visibility and control with Microsoft Defender for Cloud Apps.

After you connect your cloud apps to Defender for Cloud Apps, you can create DLP policies for them in Microsoft Purview.

Create a DLP policy for non-Microsoft connected apps

To create a DLP policy scoped to non-Microsoft connected apps:

1. Sign in to the Microsoft Purview portal.
2. Go to **Solutions** > **Data Loss Prevention** > **Policies** > **+ Create policy**.
3. On **What info do you want to protect?**, choose the **Enterprise applications & devices** option.
4. Select **Custom** > **Custom policy**, and then select **Next**.

   > **Important**
   >
   > Non-Microsoft connected app policies are only supported with the **Custom** policy template. The predefined Financial, Medical and health, and Privacy templates don't support non-Microsoft app locations.

5. Enter a name and description for the DLP policy, then select **Next**.
6. On **Assign admin units**, the scope is set to **Full directory**. Administrative units aren't supported for non-Microsoft app locations.
7. On **Choose where to apply the policy**, select one or more of the supported non-Microsoft apps (such as **Box**, **Dropbox**, **Google Workspace**, and **Salesforce**.

   > **Important**
   >
   > Non-Microsoft app locations can be selected together, but they can't be combined with other locations like SharePoint, OneDrive, Exchange, Fabric, or Devices in the same policy.

8. By default, the policy applies to all instances of the selected app. To scope the policy to specific tenants, select **Edit**.
   a. To exclude instances, in **All instances**, select **Exclude instances** > **+ Exclude instance**. From the list of instances, select instances to exclude from the policy, and then select **Done** > **Done**.
   b. To apply the policy to specific instances, select **Specific instances** > **+ Include instance**. From the list of instances, select instances to include, then select **Done** > **Done**.
9. Select **Next** to proceed.

10. On **Define policy settings**, select **Create or customize advanced DLP rules** to define an advanced DLP rule. Only advanced DLP rules are supported for non-Microsoft connected apps. Select **Next**.
11. Select + **Create rule**, and provide a **Name** and **Description**.
12. Configure **Conditions** for the rule. The conditions you can use vary by app.
13. Configure **Actions** for the rule. The available actions vary by app.
14. Configure **Notifications**. Note the following limitations for non-Microsoft apps:
    - **Policy tips** aren't supported.
    - **User overrides** aren't supported.
15. Select **Save** to save the rule, then select **Next**.
16. On **Policy mode**, choose either **Turn the policy on immediately** or **Leave the policy turned off**. Select **Next**.

    > ⓘ **Important**
    >
    > **Simulation mode** isn't supported for non-Microsoft connected app policies.

17. Review your settings and select **Submit** to create the policy.

Attribute availability by app

The conditions and actions you can use in a policy depend on the item attributes that each app makes available. Most attributes are available for all supported apps, but some aren't available for certain apps. Use the following table to understand which attributes you can rely on when you build rules for a specific app.

Expand table

Attribute	Description	Availability
Name	The display name of the item.	Available in all apps
File access level	The item's access level, which indicates its sharing scope.	Available in all apps
Created date	The date the item was created.	Not available in Dropbox

Attribute	Description	Availability
Modified date	The date the item was last modified.	Available in all apps
Collaborators	The users the item is shared with, including their roles.	Available in all apps
Parent folders	The parent folders that contain the item.	Available in all apps

Monitoring and reporting

Activity explorer

Activity explorer supports non-Microsoft connected app policies. You can monitor DLP policy activity for your non-Microsoft apps in the same way you monitor Microsoft 365 locations.

Content explorer

Content explorer isn't supported for non-Microsoft connected apps. You can't browse or inspect content stored in these apps through Content explorer.

Alerts

DLP alerts are supported for non-Microsoft connected app policies. Alerts appear in the DLP alerts dashboard alongside alerts from other locations.

Known issues

The following is a list of known issues when using non-Microsoft connected apps for DLP policies:

- **Encrypted files** – Encrypted files stored in non-Microsoft apps aren't currently supported for classification.
- **PDF files** – PDF files stored in non-Microsoft apps aren't currently supported for classification.

Related content

- Learn about data loss prevention
- Data Loss Prevention policy reference
- Create and deploy data loss prevention policies
- Use data loss prevention policies for non-Microsoft cloud apps
- Connect apps to get visibility and control with Microsoft Defender for Cloud Apps

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 08/28/2026

Learn about the default Office 365 DLP policy

To help you protect sensitive data in Microsoft Office 365 apps, Microsoft created a default Microsoft Purview data loss prevention (DLP) policy called **Default Office 365 DLP policy**.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Licensing

For information on licensing, see

- Microsoft 365 Enterprise Plans
- Microsoft 365 Service Descriptions

What does the default policy do?

This default policy looks for items with the sensitive information type Credit card number It's configured like this.

- Admin units: **Full directory**
- Choose where to apply the policy:
  - **Exchange email** > **All groups**
  - **SharePoint sites** > **All sites**
  - **OneDrive accounts** > **All users and groups**
- Conditions:
  - **Content is shared from Microsoft 365** > **with people outside my organization** AND
  - **Content contains** > **Credit card number**
- User Notifications:
  - **On**
  - **Notify these people** > **The person who sent, shared, or modified the content** and **Owner of the SharePoint site or OneDrive account** and **Owner of the SharePoint or OneDrive content**.
  - **Attach matching email message to the notification (applies only to Exchange)**
  - With customized policy tip text.

- Policy mode: **Turn the policy on immediately**

Administrator can view the activities and edit the policies details by logging into the Microsoft Purview portal.

Admins can view this policy in the Microsoft Purview portal. Navigate to **Data Loss prevention** > **Policies**.

Edit or delete the default policy

To edit the default policy for better performance or to delete it, just use an account with **DLP Compliance Management** permissions. For more information, see, Permissions.

Last updated on 03/24/2026

Get started with DLP file quarantine for SharePoint and OneDrive

> **Note**
>
> The features described in this article are in preview. This article walks you through the prerequisites and configuration for using the DLP file quarantine action for SharePoint and OneDrive locations.

Before you begin

Licensing

For information on licensing, see:

- Microsoft 365 Enterprise Plans
- Microsoft 365 Service Descriptions

DLP file quarantine for SharePoint and OneDrive is only available for E5 licenses.

Permissions

To configure and manage DLP file quarantine, you need one of the following Microsoft Entra ID roles:

- Compliance administrator
- Security administrator
- Compliance data administrator

Recommended permissions for quarantine site management

Because the quarantine site contains highly sensitive files, the administrator who manages it should also have the following data classification permissions:

- **Data Classification List Viewer** — Allows viewing items listed in Data Explorer.
- **Data Classification Content Viewer** — Allows viewing the actual contents of items in Data Explorer.

> ![CAUTION] These permissions aren't technically required to configure or use file quarantine. However, you should use the same criteria when selecting accounts for administering this feature as you do for selecting accounts to be members of the **Data Classification List Viewer** and **Data Classification Content Viewer** roles. You should apply the same account-selection criteria because the quarantine site can contain files with sensitive content that the admin might need to inspect before deciding whether to restore or delete them. To learn more, see Get started with Data Explorer.

Roles and role groups

Here's a list of applicable roles. To learn more about them, see Permissions in the Microsoft Purview portal.

- Information Protection Admin
- Information Protection Analyst
- Information Protection Investigator
- Information Protection Reader

Here's a list of applicable role groups. To learn more, see Permissions in the Microsoft Purview portal.

- Information Protection
- Information Protection Admins
- Information Protection Analysts
- Information Protection Investigators
- Information Protection Readers

Deploy a DLP policy with the quarantine action

For step-by-step instructions on how to configure quarantine settings and create and deploy a policy see, Create a DLP policy to quarantine files in SharePoint and OneDrive.

View quarantine alerts

Use the following steps to view alerts for quarantined files:

1. Open the Microsoft Purview portal and go to **Data loss prevention** > **Alerts**.
2. Review alerts for policies that use the quarantine action. DLP alerts and DLP rule match events include:

- The quarantine file location.
- The file owner.
- The original file path.

3. For more information, see Get started with the data loss prevention alerts dashboard and Investigate data loss incidents with Microsoft Defender XDR.

View quarantine activity in Data Explorer

Use Activity explorer to review quarantine-related DLP events:

1. Open the Microsoft Purview portal and go to **Data classification** > **Activity explorer**.
2. Filter for DLP policy rule match events related to your quarantine policy.
3. For more information, see Get started with Activity explorer.

Restore a quarantined file

File restore is a manual process.

> **Important**
>
> After restore, original sharing permissions aren't restored. Only the latest file version is preserved. The same DLP rule won't re-quarantine the restored file, even if the file or policy is modified. Other DLP rules continue to apply.

To restore a quarantined file:

1. Locate the quarantined file in the quarantine site.
2. Identify the original file path by using audit logs or the DLP alert details.
3. Move the file from the quarantine site back to the original location.
4. Delete the tombstone text file from the original location.

Next steps

Now that you've configured DLP file quarantine settings, you're ready to create a policy:

- Create a DLP policy to quarantine files in SharePoint and OneDrive

See also

- Learn about DLP file quarantine for SharePoint and OneDrive
- Learn about data loss prevention
- Create and deploy data loss prevention policies
- Get started with Activity explorer
- Protect files with admin quarantine - Microsoft Defender for Cloud Apps

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Create a DLP policy to quarantine files in SharePoint and OneDrive

> **Note**
>
> The features described in this article are in preview. This article uses the process you learned in Design a data loss prevention policy to show you how to create a Microsoft Purview data loss prevention (DLP) policy that automatically quarantines files containing sensitive information in SharePoint and OneDrive. Work through this scenario in your test environment to familiarize yourself with the quarantine action and policy creation UI.

> **Important**
>
> This article presents a hypothetical scenario with hypothetical values. It's only for illustrative purposes. Substitute your own sensitive information types, sensitivity labels, distribution groups, and users. How you deploy a policy is as important as policy design. This article shows you how to use the deployment options so that the policy achieves your intent while avoiding costly business disruptions.

Prerequisites and assumptions

Before you start, make sure you've completed the following:

Create a quarantine site in SharePoint

The quarantine destination must be a SharePoint site within your tenant. You can use any SharePoint site template. The quarantine site can't be a OneDrive location.

When you set up your quarantine site:

- **Restrict access.** Make sure the site is accessible only by the administrators who manage quarantine operations. Don't share it with information workers or general users.
- **Don't use a site that contains active business content.** Create a dedicated site for quarantine purposes.

- **Manage the lifecycle.** Administrators are responsible for access control and lifecycle management of the quarantine site.

> **Tip**
>
> If your organization uses on-demand classification, it's recommended that you exclude the quarantine site from on-demand classification scans. While scanning the quarantine site doesn't compromise security, excluding it avoids unnecessary processing.

Configure quarantine settings in Microsoft Purview

Before you can select the quarantine action in a DLP policy, you must configure the quarantine folder path and tombstone file message in DLP settings.

> **Important**
>
> You can't select the quarantine action in a DLP policy until you complete this configuration. The policy creation wizard prompts you to configure these settings first if they aren't already set.

1. Sign in to the Microsoft Purview portal.
2. Go to **Data loss prevention** > **Settings**.
3. Select **File quarantine**.
4. Under **Quarantine folder path**, select a SharePoint site from the list of available sites.

   > **Note**
   >
   > The site list is automatically populated from your tenant's SharePoint sites. You can't manually enter a path. If you recently created a new SharePoint site, it might take time for the site to appear in the list. If the site you need isn't listed, wait and check back later.

5. Under **Replacement file message**, provide the text content for the tombstone file that replaces each quarantined file. For example:

"This file has been moved to a secure quarantine location because it matched a data loss prevention policy. Contact your compliance administrator at compliance@contoso.com for assistance."

6. Save your settings.

> **Important**
>
> You can't select the quarantine action in a DLP policy until you complete this configuration. The policy creation wizard prompts you to configure these settings first if they aren't already set.

Sensitivity label requirements

This scenario uses the *Highly Confidential* sensitivity label. To learn more, see:

- Learn about sensitivity labels
- Get started with sensitivity labels
- Create and configure sensitivity labels and their policies

Alert configuration

This procedure uses alerts. See Get started with the data loss prevention alerts.

Policy intent statement and mapping

We need to quarantine files in SharePoint and OneDrive that contain financial data classified as "Highly Confidential," so that the files are removed from their original location and placed in a secure admin-controlled quarantine site. We want a tombstone file left in the original location to notify the file owner. We want the compliance admin alerted whenever a file is quarantined.

Expand table

Statement	Configuration question answered and configuration mapping
"We need to quarantine files in SharePoint and OneDrive..."	- **Where to monitor:** SharePoint sites, OneDrive accounts - **Administrative scope:** Full directory - **Action:** Restrict access or encrypt the content in Microsoft 365 locations > **Block everyone and move file to quarantine location**

Statement	Configuration question answered and configuration mapping
"...that contain financial data classified as 'Highly Confidential'..."	- **What to monitor:** Use the **Custom template** - **Conditions for match:** Add the sensitive info types that apply to the financial data your organization needs to protect (for example, *Credit Card Number*, *U.S. Bank Account Number*). - **Additional condition:** Content contains > Sensitivity labels > *Highly Confidential*
"We want the compliance admin alerted whenever a file is quarantined."	- **Incident reports:** Send an alert to admins when a rule match occurs: **On** - Send alert every time an activity matches the rule: **selected** - Severity level: **High**
"We want a tombstone file left in the original location to notify the file owner."	- Tombstone file content is configured in **DLP Settings** > **File quarantine** (not in the policy itself). Verify that the tombstone message is configured before you create this policy.

Steps to create the policy

> **Important**
>
> For this policy creation procedure, accept the default include and exclude values and leave the policy turned off. Change these values when you deploy the policy.

Create the policy

Perform the following steps to create the DLP policy in the Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal.
2. Open the **Data loss prevention** solution and go to **Policies**.
3. Select **Create policy**.
4. Select **Enterprise applications & devices**.
5. Select **Custom** from the **Categories** list.
6. Select **Custom policy** from the **Regulations** list.
7. Select **Next**.
8. Enter a name and description for the policy. You can use the policy intent statement here.

> **ⓘ Important**
>
> Policies can't be renamed.

9. Select **Next**.
10. Accept the **Full directory** default under **Admin units**.
11. Select **Next**.
12. Choose where to apply the policy. Select only the following locations:
    - **SharePoint sites**
    - **OneDrive accounts**
13. Select **Next**.
14. On the **Define policy settings** page, make sure the **Create or customize advanced DLP rules** option is selected.
15. Select **Next**.
16. Select **+ Create rule**. Name the rule and provide a description.
17. Under **Conditions**, select **Add condition** > **Content contains** > **Add** > **Sensitive info types**.
18. Select the sensitive info types that apply to the financial data you want to protect (for example, *Credit Card Number*, *U.S. Bank Account Number*).
19. Select **Add**.
20. Optionally, add a sensitivity label condition:
    a. Under the **Content contains** section, select **Add group**.
    b. Leave the Boolean operator set to **AND**.
    c. Select **Add condition** > **Content contains** > **Add** > **Sensitivity labels**.
    d. Select the **Highly Confidential** sensitivity label.
    e. Select **Add**.
21. Under **Actions**, select **Add an action** > **Restrict access or encrypt the content in Microsoft 365 locations**.
22. Select **Block everyone and move file to quarantine location**.

> **Note**
>
> If quarantine settings aren't configured, you're prompted to configure them. Select the link to open the **File quarantine** settings page in a new tab, complete the configuration, and then return to the policy creation wizard.

23. Under **User notifications**, set the toggle to **On** if you want to notify users through policy tips.
24. Under **Incident reports**, set **Use this severity level in admin alerts and reports** to **High**.
25. Make sure the **Send an alert to admins when a rule match occurs** toggle is set to **On**.
26. Select **Send alert every time an activity matches the rule**.
27. Select **Save**.
28. Review the rule, then select **Next**.
29. Select the **Run the policy in simulation mode** radio button.

> **Tip**
>
> Simulation mode is supported for the quarantine action. Use simulation mode to evaluate the impact of the policy before you turn it on. Review the policy matches in Activity Explorer and the DLP alerts dashboard to verify that the policy targets the files you expect.

30. Select **Next**.
31. Review the policy, then select **Submit**.
32. Select **Done**.

Validate the policy

After the policy is deployed, validate the following:

Expand table

Scenario	Expected result
A file matching the policy is uploaded to a SharePoint site in scope.	The file is moved to the quarantine site. A tombstone file appears at the original location. A DLP alert is generated.
Check the tombstone file at the original location.	The tombstone file displays the admin-configured message and the relative path of the quarantined file.
Review the DLP alert.	The alert shows the file owner, original file path, and quarantine file location.
Review Activity Explorer.	A DLP rule match event appears with quarantine file location details.

See also

- Learn about DLP file quarantine for SharePoint and OneDrive
- Get started with DLP file quarantine for SharePoint and OneDrive
- Design a data loss prevention policy
- Create and deploy data loss prevention policies
- Data loss prevention policy reference

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Learn about Endpoint data loss prevention

You can use Microsoft Purview Data Loss Prevention (DLP) to monitor the actions that are being taken on items you've determined to be sensitive and to help prevent the unintentional sharing of those items.

Endpoint data loss prevention (Endpoint DLP) extends the activity monitoring and protection capabilities of DLP to Windows 10/11, macOS (the three latest released major versions) devices, and Windows certain server versions. Once devices are onboarded into the Microsoft Purview solutions, the information about what users are doing with sensitive items is made visible in activity explorer. You can then enforce protective actions on those items via DLP policies.

When are files classified by Endpoint DLP?

Endpoint DLP classifies files under the following conditions:

- **Creation or Modification:** Every time a file is created or modified, it is fully scanned for sensitive information types (SITs) and labels. The file is then evaluated for matches against DLP policies and rules.
- **Reading an Already Classified File:** When an already classified file is read, Endpoint DLP checks for any changes to the policies, rules, or SITs. If there is a change in the DLP configuration, it re-evaluates the policies and rules, but it will not re-extract the text.

> **Tip**
>
> If you're looking for device control for removable storage, see **Microsoft Defender for Endpoint Device Control Removable Storage Access Control**.

> **Note**
>
> Endpoint DLP cannot detect the sensitivity label from another tenant on a document.

Endpoint DLP Windows 10/11 and macOS support

Endpoint DLP allows you to onboard devices running Windows Server 2019 and later versions.

> **Note**
>
> Installing the supported Windows Server KBs disables the *Classification* feature on the server. This means that Endpoint DLP won't classify files on the server. However, Endpoint DLP will still protect those files on the server that were classified before those KBs were installed on server. To ensure this protection, install Microsoft Defender version 4.18.23100 (October 2023) or later.

By default, Endpoint DLP isn't enabled for Windows servers when they're initially onboarded. Before you can see Endpoint DLP events for your servers in Activity Explorer, you must first Enable Endpoint DLP for Windows Servers.

Once properly configured, the same data loss protection policies can be automatically applied to both Windows PCs and Windows servers.

Expand table

Setting	Subsetting	Windows 10, 1809 and later, Windows 11, Windows Server 2019 and later versions for Endpoints (X64)	macOS (three latest released versions)	Notes
Advanced classification scanning and protection	**Allocated bandwidth limits**	Supported	Supported	Advanced classification enables these features for macOS: - Document Fingerprinting - Exact data match based sensitive information types - Trainable classifiers - Learn about named entities
File path exclusions for Windows	n/a	Supported	n/a	
File path exclusions for Mac	n/a	n/a	Supported	macOS includes a recommended list of exclusions that is on by default

Setting	Subsetting	Windows 10, 1809 and later, Windows 11, Windows Server 2019 and later versions for Endpoints (X64)	macOS (three latest released versions)	Notes
Setup evidence collection for file activities on devices	Set evidence cache on device	Supported	Preview	
Network share coverage and exclusions	n/a	Supported	Supported	
Restricted apps and app groups	Restricted app groups	Supported	Supported	
Restricted apps and app groups	Restricted apps	Supported	Supported	
Access by apps not included in restricted apps list or any restricted app groups added to rule	n/a	Supported	Not supported	
Restricted apps and app groups	Auto-quarantine settings	Supported	Supported	
Unallowed Bluetooth apps	n/a	Supported	Supported	
Browser and domain restrictions to sensitive data	Unallowed browsers	Supported	Supported	
Browser and domain restrictions to sensitive data	Service domains	Supported	Supported	
Browser and domain restrictions to sensitive data	Sensitive service domain groups	Supported	Supported	
Additional settings for Endpoint DLP	Business justification in policy tips	Supported	Supported	

Setting	Subsetting	Windows 10, 1809 and later, Windows 11, Windows Server 2019 and later versions for Endpoints (X64)	macOS (three latest released versions)	Notes
Always audit file activity for devices	n/a	Supported	Supported	
Printer groups	n/a	Supported	Supported	Following parameters are only supported on Windows: - Print to file - Universal print deployed on a printer - Corporate printer - Print to local
Removable USB device groups	n/a	Supported	Supported	
Network share groups	n/a	Supported	Supported	
VPN settings	n/a	Supported	Not supported	

Scoping DLP policies for endpoints

Microsoft Purview Data Loss Prevention (DLP) policies can be applied to multiple locations, including endpoint devices. When a DLP policy includes the **Devices** location, the policy can be scoped by both users and devices. For a DLP policy to be enforced on an endpoint, both the user and the device must be included in the policy scope.

Expand table

Device in scope	User in scope	Policy applied
Yes	Yes	Yes
No	Yes	No
No	Yes	No

This enables more precise data protection controls based on who is accesing data and from which device. For more information, see Device scoping.

For implementation details see Create policy using device scoping.

Other settings

Expand table

Setting	Windows 10/11, Windows 10, 1809 and later, Windows 11	Windows Server 2019 and later versions for Endpoints (X64)	macOS (three latest released versions)
Archive file	Supported	Supported	Supported
File type and File extension	Supported	Supported	Supported
Enable Endpoint DLP for Windows Servers	Not supported	Supported	Not supported

> **Note**
>
> Endpoint DLP is not supported on Windows Servers configured as Domain Controllers or Windows Servers with the Core Server option selected during installation.

Endpoint activities you can monitor and take action on

Endpoint DLP enables you to audit and manage the following types of activities users take on sensitive items that are physically stored Windows 10, Windows 11, or macOS devices.

Expand table

Activity	Description	Windows 10 (21H2, 22H2), Windows 11 (21H2, 22H2), Windows Server 2019 and later versions for Endpoints (X64)	Windows 11 (21H2, 22H2) for Endpoints (ARM64)	macOS three latest released versions	Auditable/ Restrictable
Upload to a restricted	Detects when a user attempts to upload an item to a restricted	Supported	Supported	Supported	Auditable and

cloud service domain or access from an unallowed browser	service domain or access an item through a browser. If they're using an unallowed browser, the upload activity is blocked and the user is redirected to use Microsoft Edge. Microsoft Edge then either allows or blocks the upload or access based on the DLP policy configuration. You can block, warn, or audit when protected files can be uploaded or prevented from being uploaded to cloud services based on the allow/unallowed domains list in **Data loss prevention settings**. When the configured action is set to warn or block, other browsers (defined on the unallowed browsers list under **Data loss prevention settings**) are blocked from accessing the file.				restrictable
Paste to supported browsers	Detects when a user attempts to paste content to a restricted service domain. Evaluation is performed on the content that is being pasted. This evaluation is independent of how the source item that the content came from is classified.	Supported	Supported	Preview	Auditable and restrictable
Copy to clipboard	When a user attempts to copy content from a protected file, you can block, block with override, or audit the copying of protected files to a clipboard on an endpoint device. If the rule is configured to **Block** or **Block with override** copying is blocked when the source content is sensitive except when the destination is within the same Microsoft 365 Office app. This activity also applies to redirected clipboards when using Azure Virtual Desktop with Windows 365.	Supported	Supported	Supported	Auditable and restrictable
Copy to USB	When this activity is detected, you can block, warn, or audit the	Supported	Supported	Supported	Auditable and

removable device	copying or moving of protected files from an endpoint device to USB removable media.				restrictable
Copy to a network share	When this activity is detected, you can block, warn, or audit the copying or moving of protected files from an endpoint device to any network share, including redirected USB devices that are displayed as network shares on an Azure Virtual Desktop with Windows 365.	Supported	Supported	Supported	Auditable and restrictable
Print	When this activity is detected, you can block, warn, or audit the printing of protected files from an endpoint device. This activity also applies to redirected printers when using Azure Virtual Desktop together with Windows 365.	Supported	Supported	Supported	Auditable and restrictable
Copy or move using unallowed Bluetooth app	Detects when a user attempts to copy an item to an unallowed Bluetooth app (as defined in the list of unallowed Bluetooth apps in **Data loss prevention settings** > **Endpoint settings**).	Supported	Supported	Supported	Auditable and restrictable
Copy or move using RDP	Detects when a user attempts to copy an item to a remote desktop session.	Supported	Supported	Not supported	Auditable and restrictable
Create an item	Detects the creation of an item.	Supported	Supported	Supported	Auditable
Rename an item	Detects the renaming of an item.	Supported	Supported	Supported	Auditable
Access by restricted apps	Detects when an application that is on the restricted apps list (as defined in restricted apps and app groups) attempts to access protected files on an endpoint device.	Supported	Supported	Supported	
Create Windows Recall snapshots (in preview)	Detects when there is an item or Teams message that contains a sensitive information type or has a senstivity label applied that	Supported	Not supported	Not supported	Auditable and restrictable

would be included in a Windows Recall snapshot.

Copy to clipboard behavior

When you configure a rule to **Block** or **Block with override** when a user attempts the Copy to clipboard activity on content from a file that matches the policy, end users see this behavior with these configurations:

- Word file 123 contains sensitive information that matches the copy to clipboard Block rule.
- Excel file 123 contains sensitive information that matches the copy to clipboard Block rule.
- PowerPoint file 123 contains sensitive information that matches the copy to clipboard Block rule.
- Word file 789 doesn't contain sensitive information.
- Excel file 789 doesn't contain sensitive information.
- PowerPoint file 789 doesn't contain sensitive information.
- Notepad (or any non Microsoft Office based app or process) file XYZ contains sensitive information that matches the copy to clipboard Block rule.
- Notepad (or any non Microsoft Office based app or process) file ABC doesn't contain sensitive information.

Expand table

Source	Destination	Behavior
Word file 123/Excel file 123/PowerPoint file 123	Word file 123/Excel file 123/PowerPoint file 123	copy and paste are allowed, in other words intra file copy and paste is allowed.
Word file 123/Excel File 123/PowerPoint file 123	Word file 789/Excel file 789/PowerPoint file 789	copy and paste are blocked, in other words inter file copy and paste is blocked.
Word file 789/Excel file 789/PowerPoint file 789	Word file 123/Excel File 123/PowerPoint file 123	copy and paste are allowed
Word file 123/Excel file 123/PowerPoint file 123	Notepad file ABC	copy and paste are blocked
Notepad file XYZ	any	copy is blocked

Source	Destination	Behavior
Notepad file ABC	any	copy and paste are allowed

> **Note**
>
> When a DLP rule that blocks copying is applied to an open file, copying from any other file within the same application (even files with no DLP rules applied) is restricted while the DLP blocked file is open.

Endpoint DLP and Unsaved Files

If data is never saved to a file on the local device, Endpoint DLP cannot scan or classify it. For example, if a user opens a document in Word and saves it directly to a USB device without first storing it locally, Endpoint DLP cannot inspect or block that action.

Best practice for endpoint DLP policies

Say you want to block all items that contain credit card numbers from leaving endpoints of Finance department users. We recommend:

- Create a policy and scope it to endpoints and to that group of users.
- Create a rule in the policy that detects the type of information that you want to protect. In this case, set **content contains** to *Sensitive information type**, and select **Credit Card**.
- Set the actions for each activity to **Block**.

For more information on designing your DLP policies, see Design a data loss prevention policy .

> **Note**
>
> In Microsoft Purview, DLP policy evaluation of sensitive items occurs centrally, so there's no time lag for policies and policy updates to be distributed to individual devices. When a policy is updated in Microsoft Purview portal, it generally takes about an hour for those updates to be synchronized across the service. Once policy updates are synchronized, items on targeted devices are automatically reevaluated the next time they're accessed or modified. In preview, you can also use **on-demand classification** to identify and classifies sensitive content in historical data stored in SharePoint and OneDrive. For Authorized Groups changes, the policy needs 24 hours to sync.

Monitored files

Files monitored via policy

Endpoint DLP monitors these file types policy in Windows 10, 11:

Expand table

File Type	Format	Monitored file extensions
Word processing	Word	`.doc, .docx, .docm, .dot, .dotx, .dotm, .docb, .obd, obt`
Spreadsheet	Excel, CSV, TSV	`.xls, .xlsx, .xlt, .xlm, .xlsm, .xlsb, .xlb, .xlc, .csv, .tsv`
Presentation	PowerPoint	`.ppt, .pptx, .pptm, .pps, .ppsx, .ppsm, .pot, .potx, .potm, .ppam, .pos`
Archive	Zip, RAR, 7z, TAR	`.zip, .rar, .7z, .tar, .gz, .arj, .bz2, .cab, .chm, .lzh, .lzma, .mhtml, .xar, .xz`
Adobe PDF	PDF	`.pdf`
Text	Text	`.asm, .bat, .c, .cmd, .cpp, .cs, .csv, .cxx, .def, .dic, .h, .hpp, .hxx, .idl, .inc, .inx, .java, .m3u, .mpx, .php, .pl, .pos, .txt, .vcf, .vcs`
HTML	HTML	`.ascx, .asp, .aspx, .hta, .htm, .htw, .htx, .jhtml, .html`
JSON	JSON	`.json, .adaptivecard, .messagecard`
Mail	Mail	`.eml, .msg, .nws`
XML	XML	`.infopathml, .jsp, .mspx`
Microsoft Office xml	Microsoft Office xml	`.excelml, .powerpointml, .wordml`
Protected Files	PFile	`.pbmp, .pgif, .pjfif, .pjpe, .pjpeg, .pjpg, .ppng, .ptif, .ptiff, .ptxt, .pxla, .pxlam, .pxml`
Other	Other	`.dfx, .dxf, .fluid, .mime, .pointpub, .rtf, .vtt`

Endpoint DLP monitors these file types policy in the latest three major releases of macOS:

Expand table

File Type	Format	Monitored file extensions
Word processing	Word	`.doc, .docx, .docm, .dot, .dotx, .dotm, .docb`
Spreadsheet	Excel, CSV, TSV	`.xls, .xlsx, .xlt, .xltx, .xltm, .xlw, .xlm, .xlsm, .xlsb, .xlb, .xlc, .csv, .tsv`
Presentation	PowerPoint	`.ppt, .pptx, .pptm, .pps, .ppsx, .ppsm, .pot, .potx, .potm, .ppam, .pos`
Archive	Zip, RAR, 7z, TAR	`.zip, .zipx, .rar, .7z, .tar`
Adobe PDF	PDF	`.pdf`
Text	Text	`.asm, .bat, .c, .cmd, .cpp, .cs, .csv, .cxx, .dic, .h, .hpp, .hxx, .idl, .inc, .inx, .inf, .inx, .java, .m3u, .mpx, .php, .pl, .pos, .txt, .vcf, .vcs`
HTML	HTML	`.ascx, .asp, .aspx, .css, .hta, .htm, .htw, .htx, .jhtml, .html, .mht`
JSON	JSON	`.json, .adaptivecard, .messagecard`
Mail	Mail	`.eml, .msg, .nws`
XML	XML	`.infopathml, .jsp, .mspx`
Microsoft Office xml	Microsoft Office xml	`.excelml, .powerpointml, .wordml`
Other	Other	`.pointpub, .rtf, .vtt`

These file types can be monitored through policy settings in Windows 10, 11 and macOS devices, if OCR is enabled:

`.jpg, .png, .tif, .tiff, .bmp, .jpeg`

Endpoint DLP scans files with the specified extensions listed above. If you need to protect or monitor files not listed, use the 'Apply restrictions to only unsupported file extensions' feature. The key differences between this feature and the **File extension is** condition are:

- Endpoint DLP scans content for the **File extension is** condition, allowing you to see Sensitive info type values in events or alerts. In contrast, the **Apply restrictions to only unsupported file extensions** feature does not scan file content.
- The **File extension is** condition triggers content scanning, which may consume higher machine resources like CPU and memory, potentially causing performance issues for some

file types. For more details on the **Apply restrictions to only unsupported file extensions** feature, refer to the following resources:

- Help protect files that Endpoint Data Loss Prevention doesn't scan
- Help protect against sharing of a defined set of unsupported files

Files audited regardless of policy match

Activities can be audited on these file types in Windows 10, 11, and in the latest three major releases of macOS, even if no policy match exists:

Expand table

Windows 10, 11	macOS
`.doc, .docx, .docm, .dot, .dotx, .dotm, .docb, .xls, .xlsx, .xlt, .xlm, .xlsm, .xltx, .xltm, .xlsb, .xlw, .ppt, .pptx, .pos, .pps, .pptm, .potx, .potm, .ppam, .ppsx, .pbix, .pdf, .csv, .tsv, .zip, .rar, .7z, .tar, .war, .gz, .dlp`	`.doc, .docx, .docm, .dot, .dotx, .dotm, .docb, .xls, .xlsx, .xlt, .xlm, .xlsm, .xltx, .xltm, .xlsb, .xlw, .ppt, .pptx, .pos, .pps, .pptm, .potx, .potm, .ppam, .ppsx, .pbix, .pdf, .csv, .tsv`

> **Note**
>
> These file types can be audited, regardless of a policy match, in Windows 10, 11 and macOS devices, so long as OCR is enabled:
>
> `.jpg, .png, .tif, .tiff, .bmp, .jpeg`

> **Important**
>
> For information about the Adobe requirements for using Microsoft Purview Data Loss Prevention (DLP) features with PDF files, see this article from Adobe: Microsoft Purview Information Protection Support in Acrobat.

If you only want to monitor data from policy matches, you can turn off the **Always audit file activity for devices** in the **Data loss prevention settings** > **Endpoint settings**.

If the **Always audit file activity for devices** setting is on, activities on any Word, PowerPoint, Excel, PDF, and .csv files are always audited, even if the device isn't targeted by any policy.

To ensure activities are audited for all supported file types, create a custom DLP policy.

Endpoint DLP monitors activity-based on MIME type, so activities are captured, even if the file extension is changed, for these files types:

After the extension is changed to any other file extension:

- `.doc`
- `.docx`
- `.xls`
- `.xlsx`
- `.ppt`
- `.pptx`
- `.pdf`

If the extension is changed only to supported file extensions:

- `.txt`
- `.msg`
- `.rtf`
- `.c`
- `.cpp`
- `.h`
- `.cs`
- `.java`
- `.tsv`

Unsupported File types

Endpoint DLP does not support monitoring for the following file types:

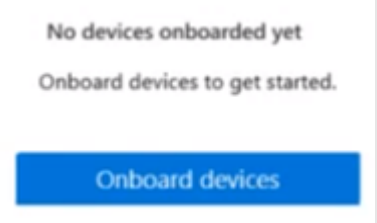
- `.exe`
- `.dll`
- `.sys`
- `.lib`
- `.obj`
- `.mui`
- `.spl`
- `.drv`
- `.pf`
- `.crdownload`
- `.ini`

What's different in Endpoint DLP

There are a few extra concepts that you need to be aware of before you dig into Endpoint DLP.

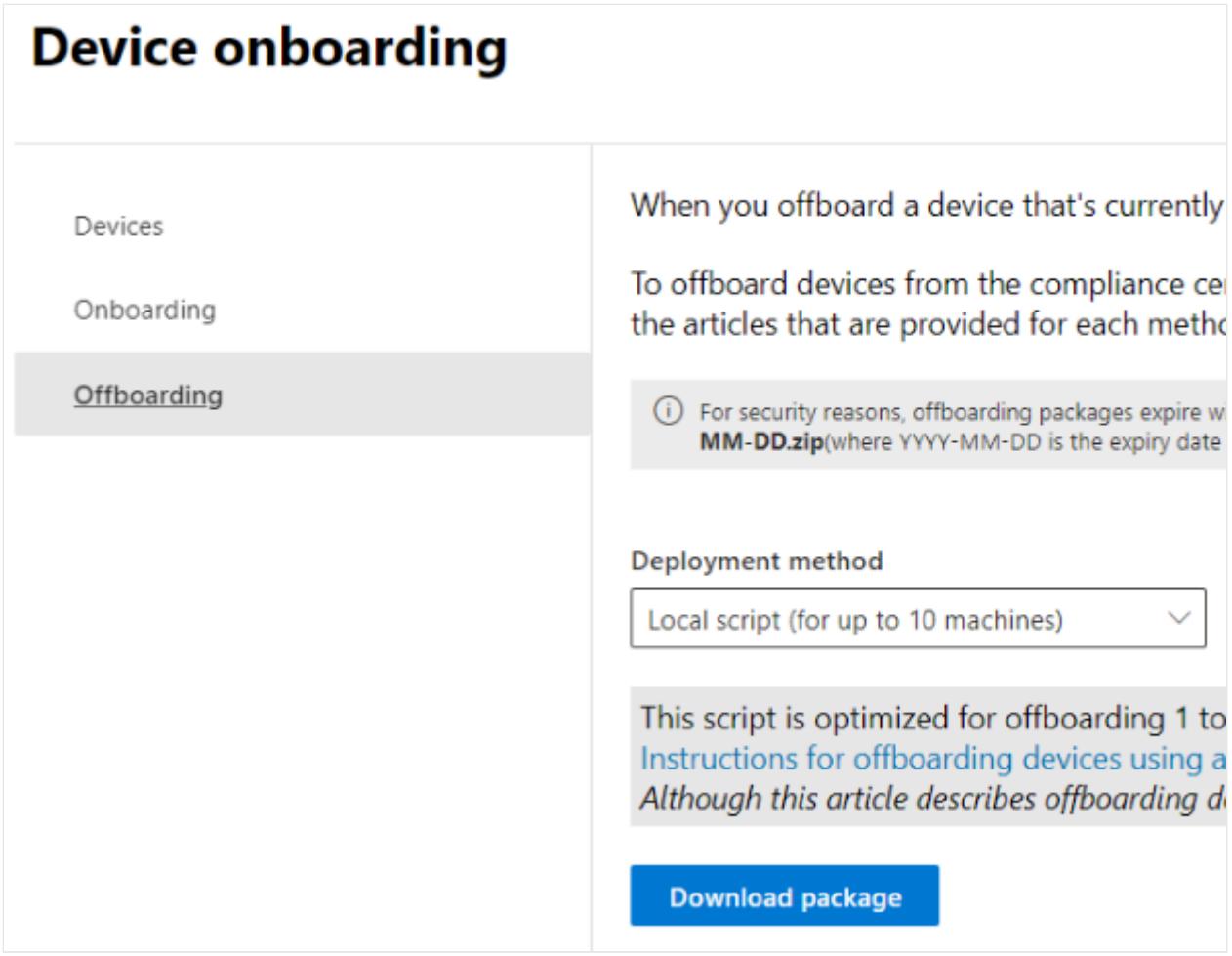
Enabling Device management

Device management is the functionality that enables the collection of telemetry from devices and brings it into Microsoft Purview solutions like Endpoint DLP and insider risk management. You need to onboard all the devices you want to use as locations in your DLP policies.



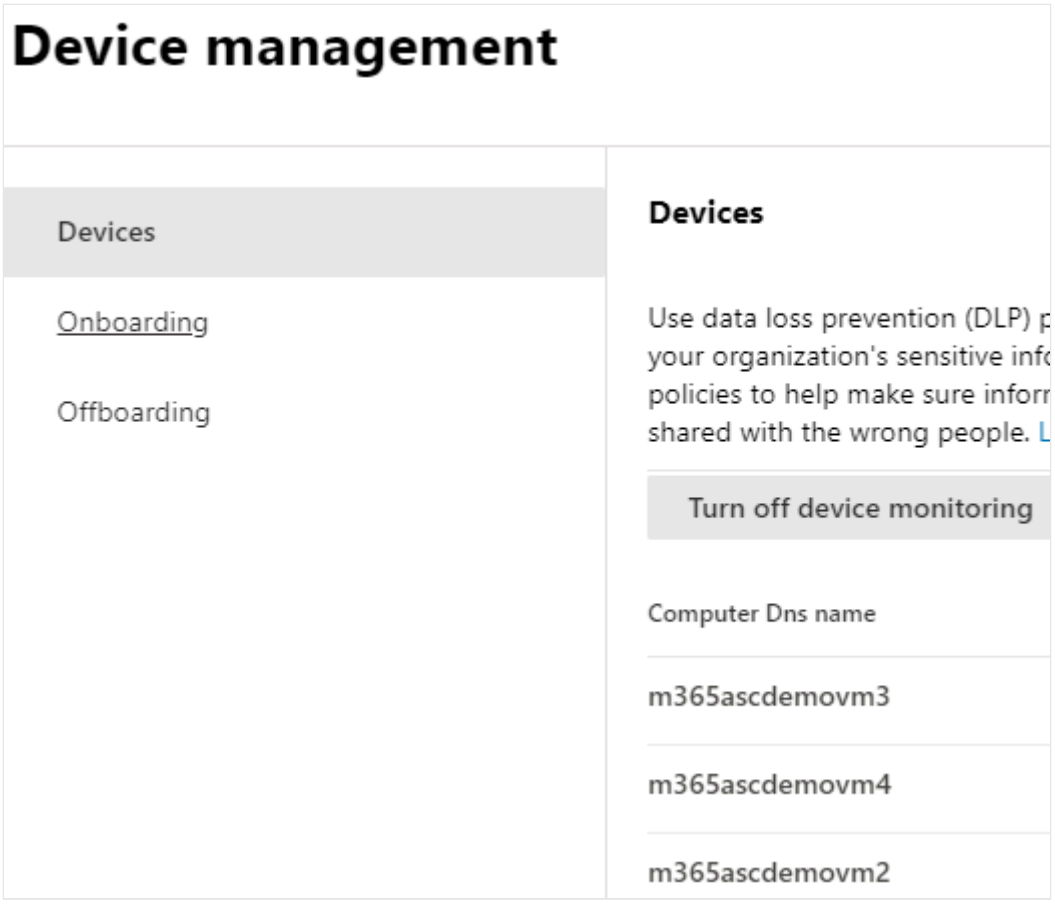
Onboarding and offboarding are handled via scripts that you download from the device management center. The device management center has custom scripts for each of the following deployment methods:

- Local script (up to 10 machines)
- Group policy
- System Center Configuration Manager (version 1610 or later)
- Mobile Device Management/Microsoft Intune
- VDI onboarding scripts for non-persistent machines



Use the procedures in Getting started with Microsoft 365 Endpoint DLP to onboard devices.

Onboarding devices to Defender also onboards them to DLP. So, if you have onboarded devices through Microsoft Defender for Endpoint, those devices show up automatically in the list of devices. You need only **Turn on device monitoring** to use endpoint DLP.

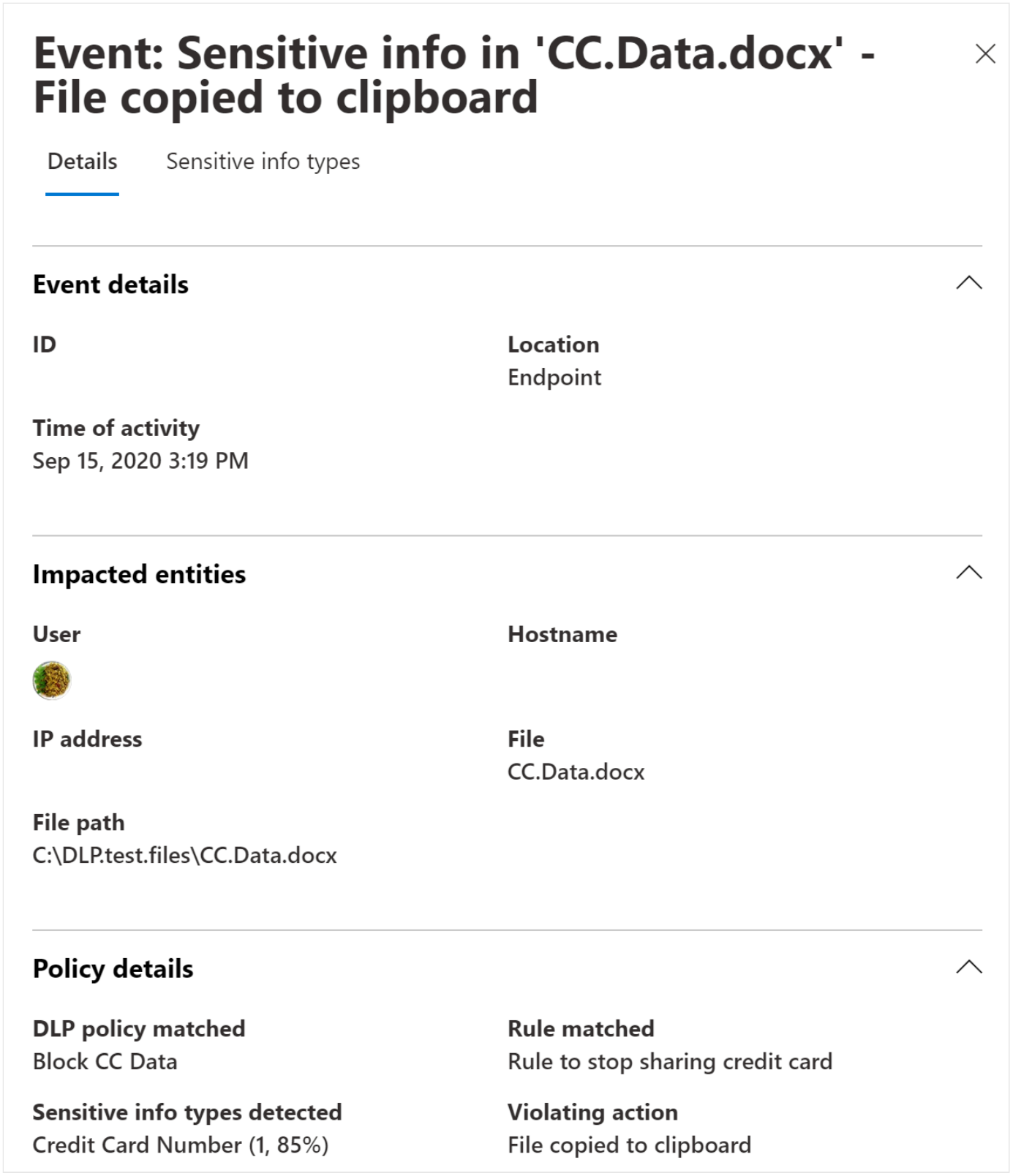


Viewing Endpoint DLP data

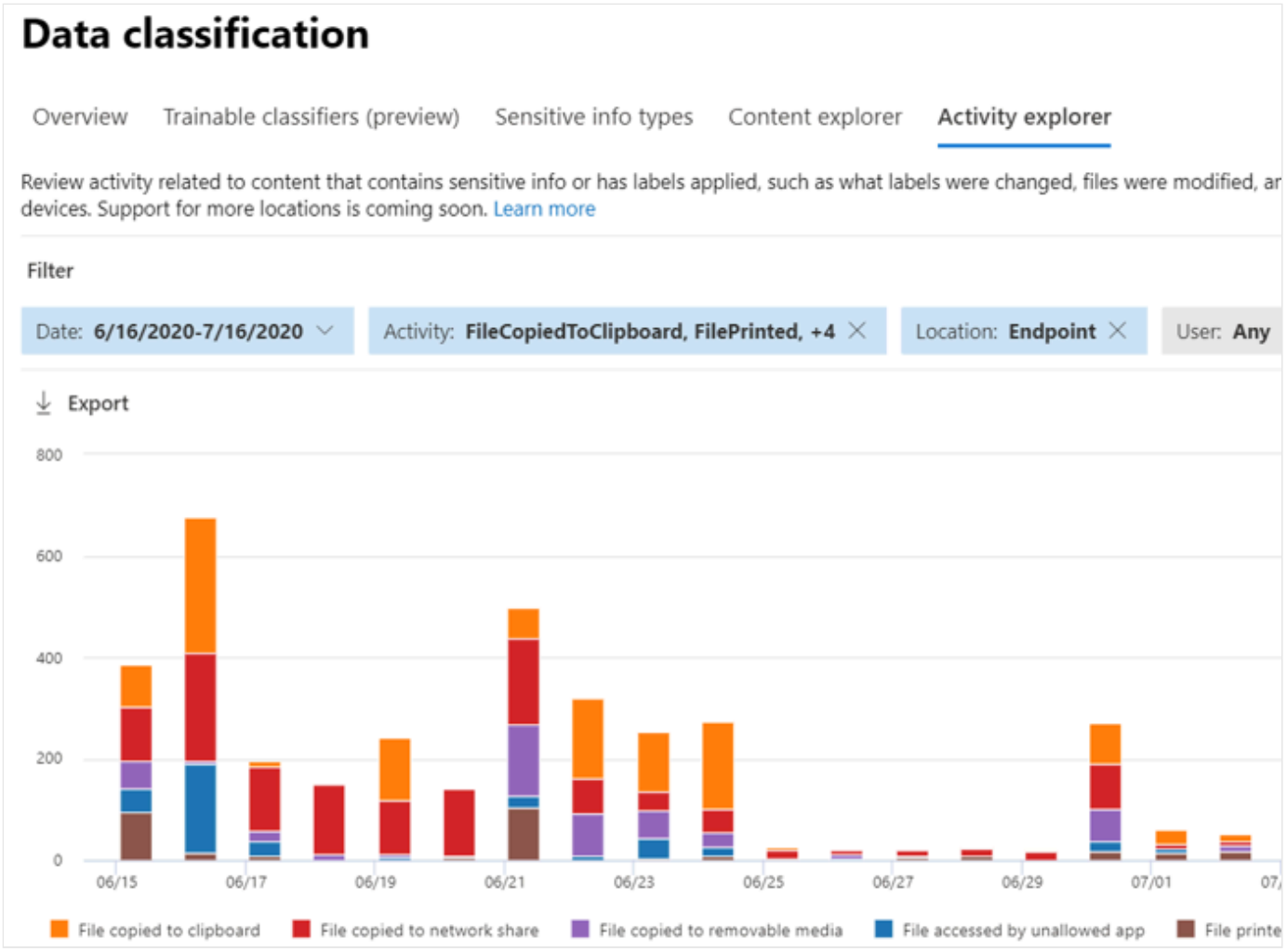
You can view alerts related to DLP policies enforced on endpoint devices by going to the DLP Alerts Management Dashboard and Investigate data loss incidents with Microsoft Defender XDR.



You can also view details of the associated event, with rich metadata, in the same dashboard



Once a device is onboarded, information about audited activities flows into Activity explorer even before you configure and deploy any DLP policies that have devices as a location.

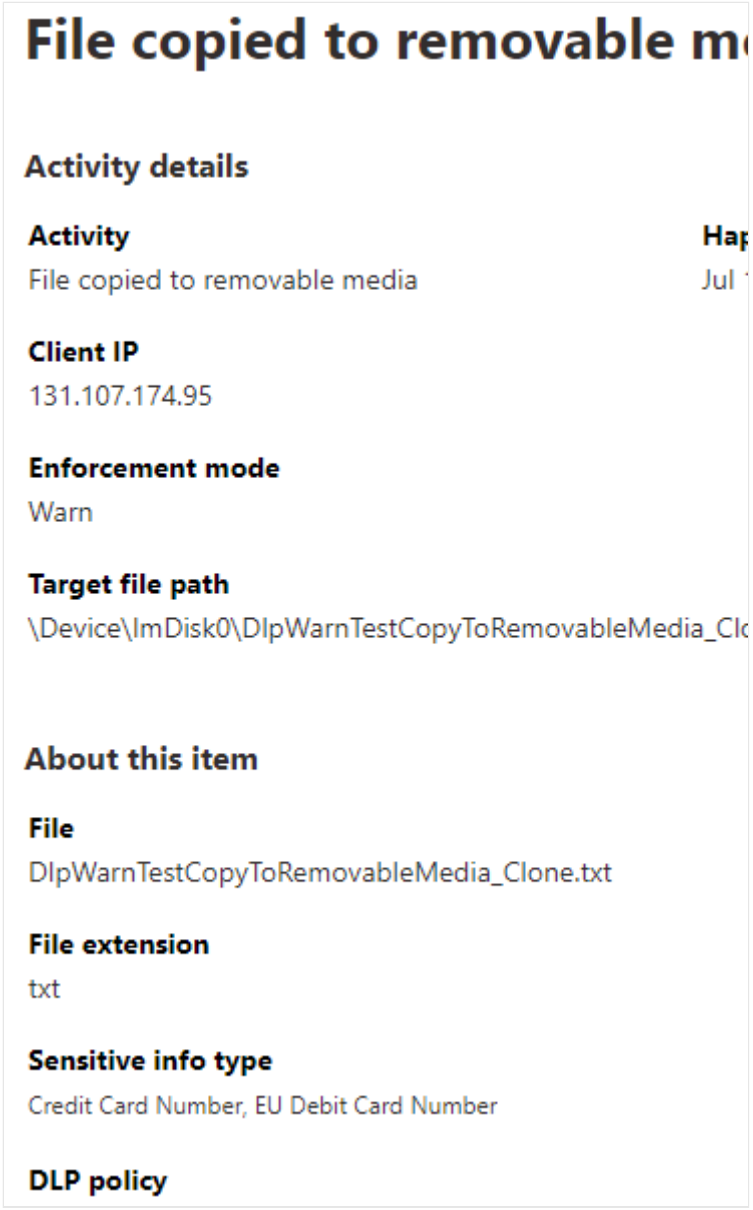


Endpoint DLP collects extensive information on audited activity.

For example, if a file is copied to removable USB media, you'd see these attributes in the activity details:

- activity type
- client IP
- target file path
- happened timestamp
- file name
- user
- file extension
- file size
- sensitive information type (if applicable)
- sha1 value
- sha256 value
- previous file name
- location
- parent

- filepath
- source location type
- platform
- device name
- destination location type
- application that performed the copy
- Microsoft Defender for Endpoint device ID (if applicable)
- removable media device manufacturer
- removable media device model
- removable media device serial number



Why are all endpoint policies on all onboarded devices?

A device can support multiple user accounts. Because endpoint policies are scoped to users, different users of the device may have different policies scoped to them. Therefore, the device

must have access to all endpoint DLP policies in order to evaluate items. All devices that are onboarded into Purview are scanned regardless if the user is in scope.

Endpoint DLP and offline devices

When a Windows endpoint device is offline, existing policies continue to be enforced on existing files. Additionally, with just-in-time protection enabled and in "block" mode, when a new *file* is created on an offline device, the file is still prevented from being shared until the device connects to the data classification service and evaluation completes. If a new *policy* is created on the server, or an existing policy is modified, those changes are updated on the device once it reconnects to the internet.

> **Important**
>
> This functionality isn't supported on macOS endpoint devices.

Consider the following use cases.

1. Policies that have been pushed to a device will continue to be applied to files already classified as sensitive even after the device goes offline.
2. Policies that are updated while a device is offline won't be pushed to that device. Similarly, such policies won't be enforced on that device, until the device is back online. However, the outdated policy that exists on the offline device will still be enforced. Just-in-time protection

If notifications are configured to display, they'll always display when DLP policies are triggered, regardless of whether or not the device is online.

> **Note**
>
> While policies that have already been pushed to an offline device are enforced, the enforcement events don't appear in activity explorer until the device is back online.

DLP policies regularly sync to endpoint devices. If a device is offline, the policies can't be synchronized. In this case, the **Devices list** reflects that the device is out of sync with the policies in on the server.

Just-in-time protection

Just-in-time protection blocks egress activities on these monitored files until policy evaluation completes successfully:

- Items that have never been evaluated.
- Items on which the evaluation has gone stale. These are previously evaluated items that haven't been reevaluated by the current, updated cloud versions of the policies.

For more information on how just-in-time protection works,see Learn about just-in-time protection, and Get started with Microsoft Purview Data Loss Prevention just-in-time protection.

Next steps

Now that you've learned about Endpoint DLP, your next steps are:

1. Onboard Windows 10 or Windows 11 devices into Microsoft Purview overview
2. Onboard macOS devices into Microsoft Purview overview
3. Configure endpoint data loss prevention settings
4. Using Endpoint data loss prevention

See also

- Getting started with Microsoft Endpoint data loss prevention
- Using Microsoft Endpoint data loss prevention
- Learn about data loss prevention
- Investigate endpoint DLP evidence in Data Security Investigations (preview)
- Create and Deploy data loss prevention policies
- Get started with Activity explorer
- Microsoft Defender for Endpoint
- Insider risk management

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 08/05/2026

Get started with endpoint data loss prevention

Endpoint data loss prevention (Endpoint DLP) is part of the Microsoft Purview Data Loss Prevention (DLP) suite of features you can use to discover and protect sensitive items across Microsoft 365 services. For more information about all of Microsoft's DLP offerings, see Learn about data loss prevention. To learn more about Endpoint DLP, see Learn about Endpoint data loss prevention

Microsoft Endpoint DLP allows you to monitor onboarded Windows 10, and Windows 11 and onboarded macOS devices running any of the three latest released versions. Once a device is onboarded, DLP detects when sensitive items are used and shared. This gives you the visibility and control you need to ensure that they're used and protected properly, and to help prevent risky behavior that might compromise them.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about **Microsoft Security Copilot in Microsoft Purview**.

Before you begin

SKU/subscriptions licensing

For information on licensing, see

- Microsoft 365 Enterprise Plans
- Microsoft 365 Service Descriptions

> **Important**
>
> Endpoint DLP policy targeting can be configured using a combination of users and devices. If a policy is scoped to devices where the signed-in users do not meet the **required criteria**, ensure those users or devices or both are explicitly excluded from the policy. Failure to do so may lead to unintended policy enforcement behavior.

Configure proxy on the Windows 10 or Windows 11 device

If you're onboarding Windows 10 or Windows 11 devices, check to make sure that the device can communicate with the cloud DLP service. For more information, see, Configure device proxy and internet connection settings for Microsoft Purview Endpoint DLP.

Windows 10 and Windows 11 Onboarding procedures

For a general introduction to onboarding Windows devices, see:

- Onboard Windows devices into Microsoft 365 overview

For specific guidance to onboarding Windows devices, see:

Expand table

Article	Description
Onboard Windows 10 or 11 devices using Group Policy	Use Group Policy to deploy the configuration package on devices.
Onboard Windows 10 or 11 devices using Microsoft Endpoint Configuration Manager	You can use either use Microsoft Endpoint Configuration Manager (current branch) version 1606 or Microsoft Endpoint Configuration Manager (current branch) version 1602 or earlier to deploy the configuration package on devices.
Onboard Windows 10 or 11 devices using Microsoft Intune	Use Microsoft Intune to deploy the configuration package on device.
Onboard Windows 10 or 11 devices using a local script	Learn how to use the local script to deploy the configuration package on endpoints.
Onboard non-persistent virtual desktop infrastructure (VDI) devices	Learn how to use the configuration package to configure VDI devices.

Endpoint DLP support for virtualized environments

You can onboard virtual machines as monitored devices in Microsoft Purview portal. There's no change to the onboarding procedures listed above.

The table that follows lists the virtual operating systems that are supported by virtualization environments.

Expand table

Virtualization platform	Windows 10	Windows 11	Windows Server 2019	Windows Server 2022 21H2, 22H2, Data Center
Azure virtual desktop (AVD)	• Single session supported for 21H2, 22H2 • Multi session supported for 21H2, 22H2	• Single session supported for 21H2, 22H2 • Multi session supported for 21H2, 22H2	Single session and Multi session supported.	Supported
Windows 365	• Supported for 21H2, 22H2	• Supported for 21H2, 22H2	Not applicable	Not applicable
Citrix Virtual Apps and Desktops 7 (2209 and higher)	• Single session supported for 21H2, 22H2 • Multi session supported for 21H2, 22H2	• Single session supported for 21H2, 22H2 • Multi session supported for 21H2, 22H2	Supported	Supported
Amazon workspaces	• Single session supported for 21H2, 22H2	Not applicable	• Windows 10 powered by Windows Server 2019	Not applicable
Hyper-V	• Single session supported for 21H2, 22H2 • Multi session with Hybrid AD join supported for 21H2, 22H2	• Single session supported for 21H2, 22H2 • Multi session with Hybrid AD join supported for 21H2, 22H2	Supported with Hybrid AD join	Supported with Hybrid AD join

Known issues

1. You can't monitor **Copy to Clipboard** and **Enforcing Endpoint DLP** on Azure Virtual Desktop environments via browsers. However, the same egress operation will be monitored by **Endpoint DLP for actions via Remote Desktop Session (RDP)**.
2. Citrix XenApp doesn't support access by restricted app monitoring.

Limitations

1. Handling of USBs in virtualized environments: USB storage devices are treated as network shares. You need to include the **Copy to network share** activity to monitor **Copy to a USB device**. All activity explorer events for virtual devices and incident alerts show the **Copy to a network share** activity for all copy to USB events.

macOS onboarding procedures

For a general introduction to onboarding macOS devices, see:

- Onboard macOS devices into Microsoft Purview

For specific guidance to onboarding macOS devices, see:

> **Important**
>
> There are important permission changes for Microsoft Purview Endpoint DLP in macOS 27. If you're onboarding or upgrading devices that run macOS 27, review the updated guidance before deployment. For more information, see Microsoft Purview Endpoint DLP on macOS 27 (Preview).

Expand table

Article	Description
Intune	For macOS devices that are managed through Intune
Intune for Microsoft Defender for Endpoint customers	For macOS devices that are managed through Intune and that have Microsoft Defender for Endpoint (MDE) deployed to them
JAMF Pro)	For macOS devices that are managed through JAMF Pro
JAMF Pro for Microsoft Defender for Endpoint customers)	For macOS devices that are managed through JAMF Pro and that have Microsoft Defender for Endpoint (MDE) deployed to them

Once a device is onboarded, it should be visible in the devices list, and should start reporting audit activity to Activity explorer.

See also

- Learn about Endpoint data loss prevention
- Using Endpoint data loss prevention
- Learn about data loss prevention

- Create and Deploy data loss prevention policies
- Get started with Activity explorer
- Microsoft Defender for Endpoint
- Onboarding tools and methods for Windows machines
- Microsoft 365 subscription
- Microsoft Entra joined devices
- Download the new Microsoft Edge based on Chromium

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 08/20/2026

Microsoft Purview Endpoint DLP changes in macOS 27 overview (preview)

Microsoft Purview Endpoint Data Loss Prevention (Endpoint DLP) uses macOS permissions to identify websites, applications, files, and active content involved in user activities. This information helps Endpoint DLP evaluate policies and apply protection actions for activities such as uploading sensitive data to websites and cloud services or copying sensitive content to a browser.

Before macOS 27 (Golden Gate), Endpoint DLP relied on the **Accessibility** permission for browser and website visibility. In macOS 27, **Device Control and Data Access** manages this capability. Users can disable this permission after it's deployed through mobile device management (MDM), which can affect Endpoint DLP's ability to get browser and website context. To help maintain data protection coverage, administrators can configure protection modes, monitor permission status, and guide users to restore the required permission when needed.

This article describes the changes introduced in macOS 27, their impact on Endpoint DLP, and how administrators can prepare their organization. Administrators can configure protection modes, monitor affected devices, and validate deployment readiness.

> **Important**
>
> Use Endpoint DLP client version 101.26072 or later with this macOS 27 guidance.

Prerequisites

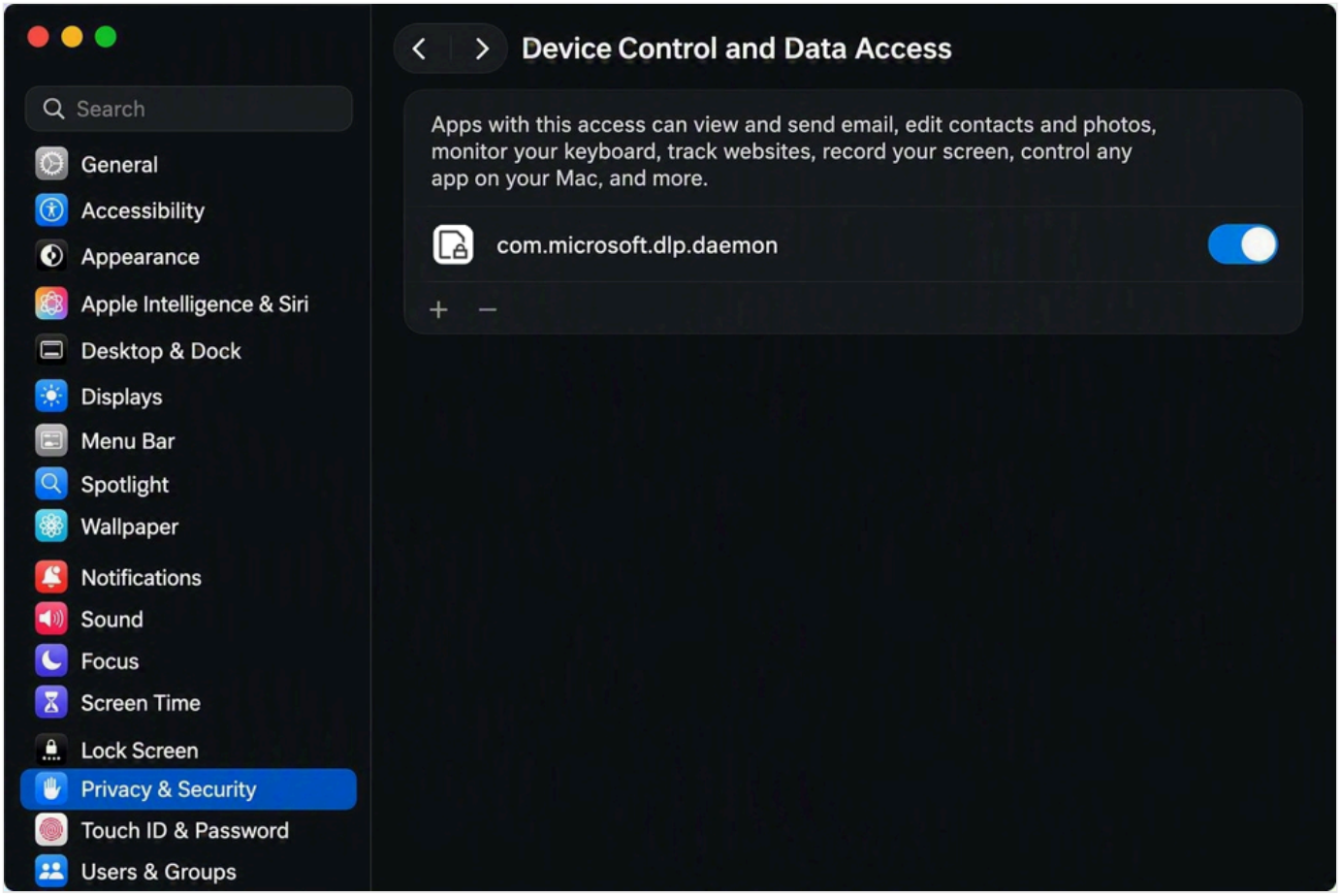
- Endpoint DLP client version 101.26072 or later.
- Familiarity with onboarding macOS devices into Microsoft Purview.

Permission changes in macOS 27

On macOS 27, Endpoint DLP uses **Device Control and Data Access** permission to support browser visibility, activity monitoring, and policy enforcement. The permission lets Endpoint DLP identify the active browser, website, application, and file associated with user activity. This context supports policies for activities such as uploading sensitive files and pasting sensitive content into browsers.

In macOS 27, the permission is under **Settings** > **Privacy & Security** > **Device Control and Data Access**. In earlier macOS versions, the permission is under **Settings** > **Privacy & Security** > **Accessibility**.

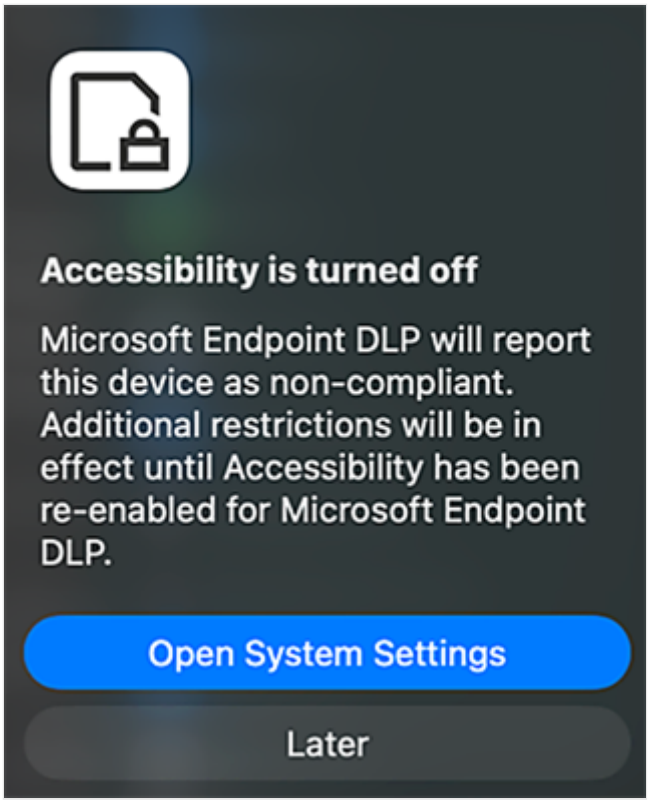
In some macOS interfaces, management profiles, and diagnostic logs, Microsoft Endpoint DLP might appear as `com.microsoft.dlp.daemon.app` or `com.microsoft.dlp.daemon`. These identifiers refer to the Endpoint DLP application.



Behavior when permission is turned off

When Endpoint DLP detects that a user turned off the Microsoft Purview Endpoint DLP permission under **Device Control and Data Access**, Endpoint DLP can no longer get URL and browser context for Safari, Chrome, and Firefox. Endpoint DLP responds by:

1. Displaying a notification that prompts the user to restore the permission (see the following screenshot).



2. Applying administrator-configured protection settings when browser URL context is unavailable.

Administrators can configure separate protection modes for:

- Uploading sensitive files to websites and cloud services
- Copying sensitive data to a browser

Protection modes

Endpoint DLP provides **Enforce**, **Audit**, and **Off** modes for cloud egress and paste-to-browser activities when the permission is unavailable. The default for both settings is **Enforce**.

Uploading sensitive files to websites and cloud services

Setting name: Restricted protection mode for cloud egress

This setting controls how Endpoint DLP handles uploads of sensitive files when browser and website context are unavailable.

- **Enforce**: Applies upload protection. Endpoint DLP treats supported non-Edge browsers, such as Safari, Chrome, and Firefox, as restricted destinations and enforces the most restrictive applicable policy action.
- **Audit**: Records and monitors the condition without enforcing upload restrictions.

- **Off**: Doesn't apply permission-related upload protection.

Copying sensitive data to a browser

Setting name: Restricted protection mode for paste to browse

This setting controls how Endpoint DLP handles copy-and-paste activities into browsers when browser and website context are unavailable.

- **Enforce**: Applies policy actions, such as block or warn.
- **Audit**: Records events without enforcing restrictions.
- **Off**: Doesn't apply permission-related browser copy-and-paste protection.

Permission notification

Setting name: Permission notification (`notification`)

Endpoint DLP can notify users when the permission is turned off. The notification explains why permission is required and prompts the user to restore access. Administrators can customize the notification message.

The profile settings map to the following configuration keys:

Expand table

Customer scenario	Setting name	Configuration key
Uploading sensitive files to websites and cloud services	Restricted protection mode for cloud egress	`unallowedBrowserMode`
Copying sensitive data to a browser	Restricted protection mode for paste to browser	`pasteToBrowserMode`
Permission notification	Permission notification	`notification`

The following source fragment sets both protection modes to Enforce:

XML

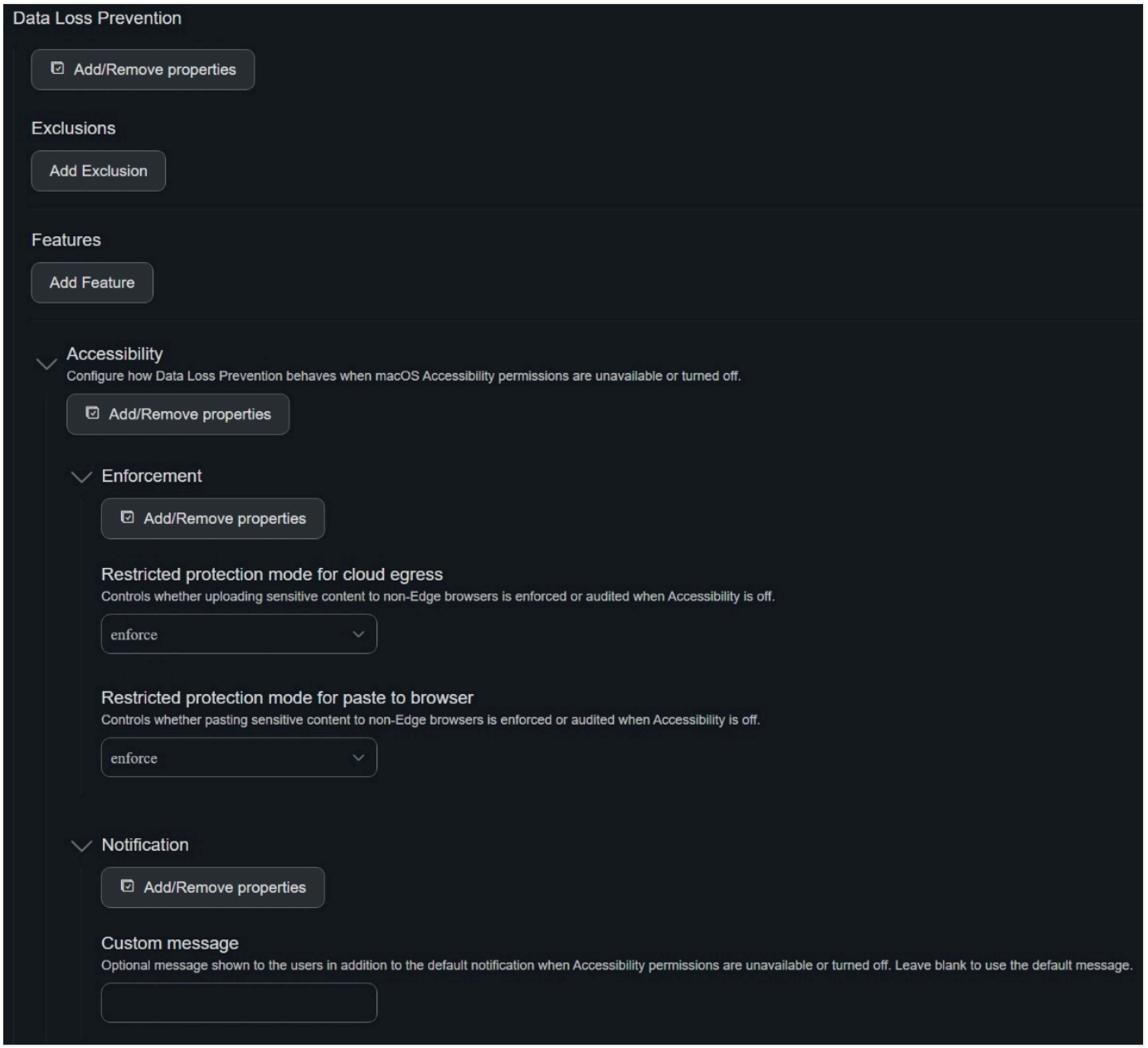
```
<key>accessibility</key>
                    <dict>
                        <key>enforcement</key>
                        <dict>
                            <key>unallowedBrowserMode</key>
                            <string>enforce</string>
                            <key>pasteToBrowserMode</key>
```

```
                            <string>enforce</string>
                        </dict>
                        <key>notification</key>
                        <dict>
                            <key>Microsoft Endpoint DLP will report this device as
non-compliant. Additional restrictions will be in effect until Accessibility has been
re-enabled for Microsoft Endpoint DLP</key>
                            <string/>
                        </dict>
                    </dict>
```

MDM configuration options

Microsoft provides configuration samples for Microsoft Intune and Jamf Pro. These options configure the protection modes and the message that prompts users to turn on the permission.

- **Microsoft Intune**: Use the sample com.microsoft.wdav.mobileconfig configuration profile as the starting point for deployment.
- **Jamf Pro with `schema.json`**: Update the existing Microsoft Defender for Endpoint preferences configuration profile to use the latest schema.json. The configuration is under **Data Loss Prevention** > **Accessibility**.
- **Jamf Pro with a property list**: Use the sample com.microsoft.wdav.plist.



Audit mode

Audit mode lets administrators assess the effect of unavailable permissions before they enable enforcement. Endpoint DLP continues to detect when the permission is turned off and can notify users to restore access, but it doesn't apply the additional restricted protection behavior for cloud egress or paste-to-browser activities.

Audit mode can reduce user disruption while administrators identify affected devices. During this assessment, controls that would normally apply when browser context is unavailable aren't enforced. This condition can leave a temporary protection gap until the permission is restored or enforcement is enabled.

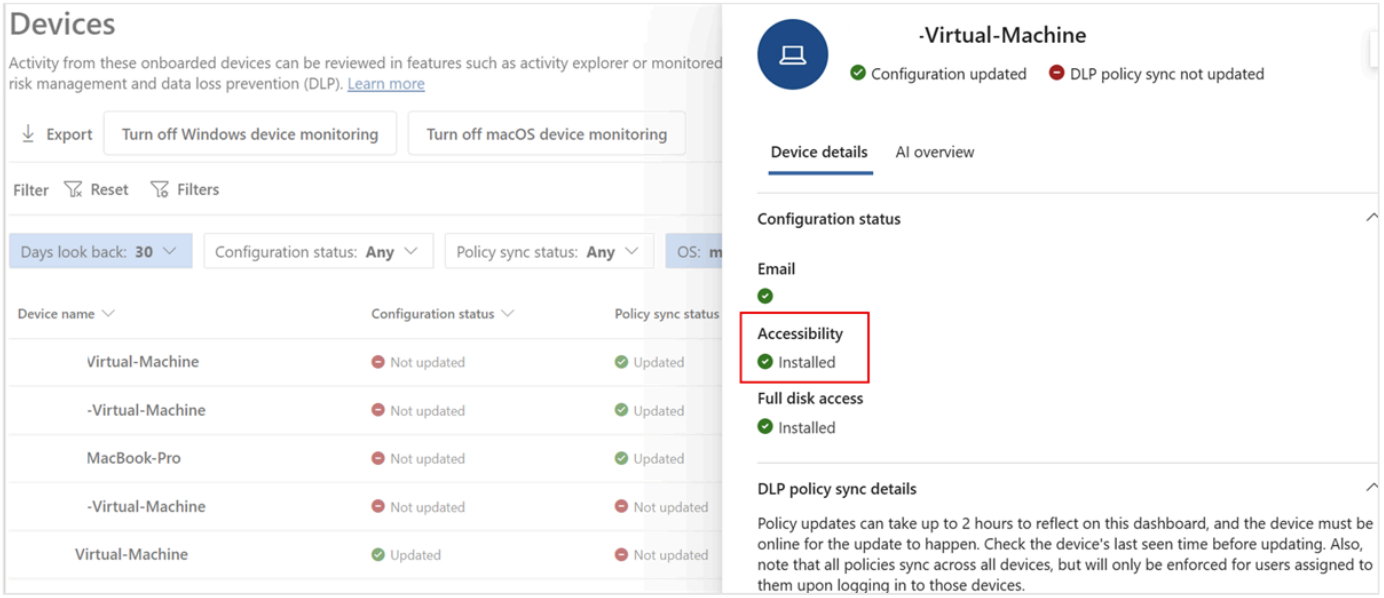
The following source fragment sets both protection modes to Audit:

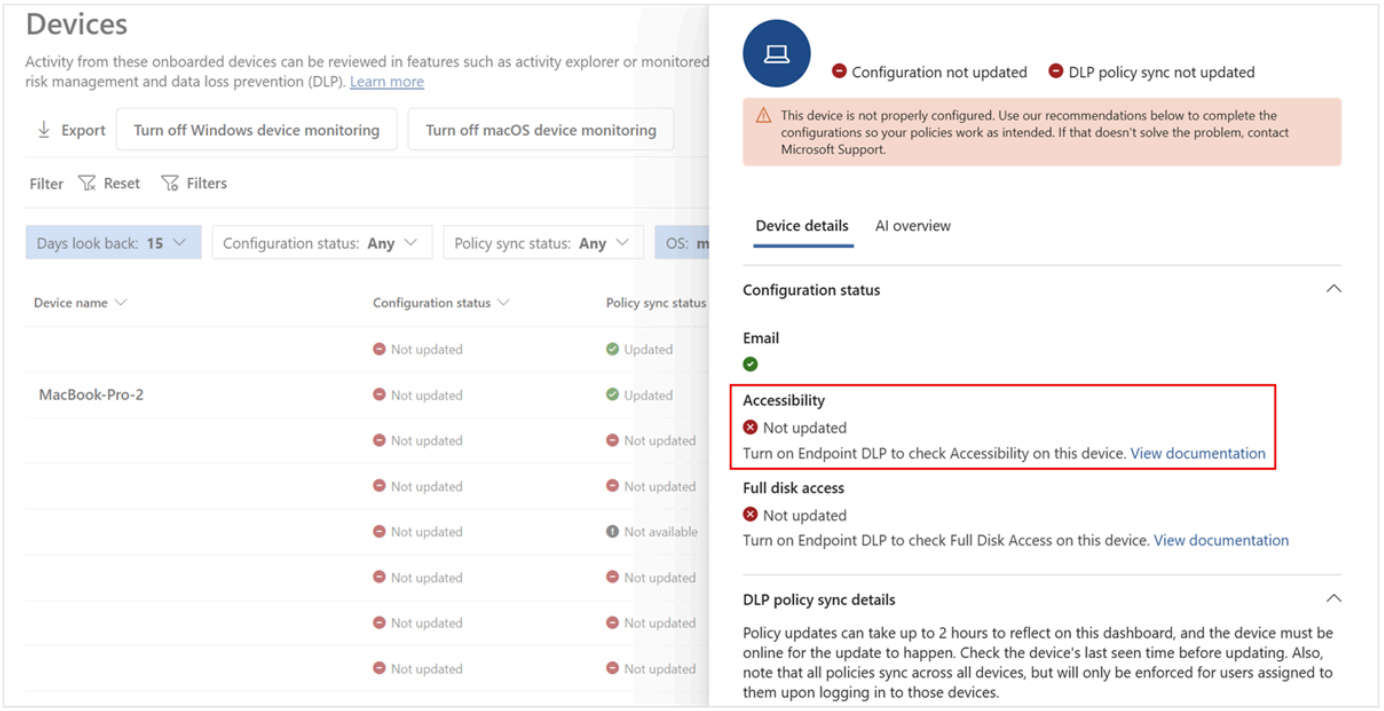
XML

```
<key>accessibility</key>
                    <dict>
                        <key>enforcement</key>
                        <dict>
                            <key>unallowedBrowserMode</key>
                            <string>audit</string>
                            <key>pasteToBrowserMode</key>
                            <string>audit</string>
                        </dict>
                        <key>notification</key>
                        <dict>
                            <key>Microsoft Endpoint DLP will report this device as
non-compliant. Additional restrictions will be in effect until Accessibility has been
re-enabled for Microsoft Endpoint DLP</key>
                            <string/>
                        </dict>
                    </dict>
```

Permission monitoring

Monitor Endpoint DLP permission status in Microsoft Purview. The device details pane shows the **Accessibility** configuration status for an onboarded macOS device. A status of **Installed** indicates the required permission is configured, while **Not updated** indicates the permission isn't available to Endpoint DLP and requires administrator attention. See the following screenshot examples.





Administrator readiness

Preparing Endpoint DLP for macOS 27 includes the following activities:

1. Review the updated macOS deployment guidance for Endpoint DLP.
2. Deploy or validate the required Endpoint DLP permission configuration on managed macOS devices.
3. Upgrade to Endpoint DLP client version 101.26072 or later.
4. Monitor the **Device Onboarding** page for devices that report the permission as off.
5. Configure user notifications and educate users to restore the permission if it's turned off.

Related content

- Onboard macOS devices into Microsoft 365 overview
- Get started with Endpoint data loss prevention
- Learn about Endpoint data loss prevention

ⓘ**Note:** This article contains content created with AI. Learn more

---

Last updated on 08/26/2026

Configure endpoint data loss prevention settings

Centrally configured settings control many aspects of endpoint data loss prevention (DLP) behavior. These settings apply to all DLP policies for devices. Use these settings to control the following behaviors:

- Cloud egress restrictions
- Various types of restrictive actions on user activities per application
- File path exclusions for Windows and macOS devices
- Browser and domain restrictions
- Appearance of business justifications for overriding policies in policy tips
- Whether actions performed on Office, PDF, and CSV files are automatically audited

To access these settings, from the Microsoft Purview portal, go to **Data loss prevention** > **Overview** > **Data loss prevention settings** > **Endpoint settings**.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

> **Important**
>
> For information about the Adobe requirements for using Microsoft Purview Data Loss Prevention (DLP) features with PDF files, see this article from Adobe: Microsoft Purview Information Protection Support in Acrobat.

Advanced classification scanning and protection

When you turn on advanced classification scanning and protection, the Microsoft Purview cloud-based data classification service scans items, classifies them, and returns the results to the local machine. Because advanced classification runs in the cloud, you can take advantage of classification techniques such as exact data match classification, trainable classifiers, credential classifiers, and named entities in your DLP policies.

> **Note**
>
> The **Paste to browser** action doesn't support advanced classification.

When you turn on advanced classification, the local device sends content to the cloud services for scanning and classification. If you're concerned about bandwidth usage, set a limit on how much bandwidth can be used in a rolling 24-hour period. You configure the limit in **Endpoint DLP settings** and it applies per device. If you set a bandwidth usage limit and that usage limit is exceeded, DLP stops sending the user content to the cloud. At that point, data classification continues locally on the device but classification by using exact data match, named entities, trainable classifiers, and credential classifiers isn't available. When the cumulative bandwidth usage drops below the rolling 24-hour limit, communication with the cloud services resumes.

If bandwidth usage isn't a concern, select **Do not limit bandwidth. Unlimited** to allow unlimited bandwidth use.

Advanced classification file scanning size limits

Even if you enable **Do not limit bandwidth. Unlimited** for advanced classification, limits still exist on the size of individual files that the system can scan.

- There's a 64-MB limit on text files.
- There's a 50-MB limit on image files when Optical Character Recognition (OCR) is enabled.

Advanced classification doesn't work for text files larger than 64 MB, even if you set the bandwidth limit to **Do not limit bandwidth. Unlimited**.

The following Windows versions (and later versions) support advanced classification scanning and protection.

- All Windows 11 versions
- Windows 10 versions 20H1/21H1 or higher (KB 5006738)
- Windows 10 RS5 (KB 5006744)

> **Note**
>
> - Support for advanced classification is available for Office (Word, Excel, PowerPoint) and PDF file types.
> - DLP policy evaluation always occurs in the cloud, even if user content isn't being sent.

> **Tip**
>
> To use advanced classification for Windows 10 devices, you must install KB5016688. To use advanced classification for Windows 11 devices, you must install KB5016691 on those Windows 11 devices. Additionally, you must enable advanced classification before **Activity explorer** displays contextual text for DLP rule-matched events. To learn more about contextual text, see the "Contextual summary" section in Learn about data loss prevention.

Advanced label-based protection for all files on devices

When you turn on advanced label-based protection, users can work on files - including files other than Office and PDF files - that have sensitivity labels applying access control settings in an unencrypted state, on their devices. Endpoint DLP continues to monitor and enforce access control and label-based protections on these files even in an unencrypted state. It automatically encrypts them before they're transferred outside from a user's device. For more information about advanced label-based protection, see Learn about Advanced Label Based Protection.

> **Note**
>
> Advanced label-based protection is supported only on onboarded Windows devices.

File path exclusions

To exclude certain paths from DLP monitoring, DLP alerts, and DLP policy enforcement on your devices, set up file path exclusions to turn off those configuration settings. Files in excluded locations aren't audited. Any files that you create or modify in those locations aren't subject to DLP policy enforcement. To configure path exclusions in DLP settings, go to Microsoft Purview portal > **Data loss prevention** > **Overview** > **Data loss prevention settings** > **Endpoint settings** > **File path exclusions for Windows**.

File path exclusions for Windows

Use the following logic to construct your exclusion paths for Windows 10 and Windows 11 devices:

- A valid file path that ends with `\` means only files directly under the specified folder are excluded.
  Example: `C:\Temp\`
- A valid file path that ends with `\*` means only files within subfolders of the specified folder are excluded. Files directly under the specified folder itself aren't excluded.
  Example: `C:\Temp\*`
- A valid file path that ends without `\` or `\*` means all files directly under the specified folder and all of its subfolders are excluded.
  Example: `C:\Temp`
- A path with a wildcard between `\` from each side.
  Example: `C:\Users\*\Desktop\`
- A path with a wildcard between `\` from each side and with `(number)` to specify the exact number of subfolders to exclude.
  Example: `C:\Users\*(1)\Downloads\`
- A path with SYSTEM environment variables.
  Example: `%SystemDrive%\Test\*`
- A mix of all the patterns described here.
  Example: `%SystemDrive%\Users\*\Documents\*(2)\Sub\`

Windows file paths excluded by default

- `%SystemDrive%\\Users\\*(1)\\AppData\\Roaming`
- `%SystemDrive%\\Users\\*(1)\\AppData\\Local`

File path exclusions for Mac

You can also add your own exclusions for macOS devices.

- File path definitions are case insensitive, so `User` is the same as `user`.
- Wildcard values are supported. So a path definition can contain an asterisk (`*`) in the middle of the path or at the end of the path.
  Example: `/Users/*/Library/Application Support/Microsoft/Teams/*`

macOS file paths excluded by default

`/System`

Recommended file path exclusions for macOS

For performance reasons, Endpoint DLP includes a list of recommended file path exclusions for macOS devices. If you set the **Include recommended file path exclusions for Mac** toggle to **On**, the following paths are also excluded:

- `/Applications`
- `/usr`
- `/Library`
- `/private`
- `/opt`
- `/Users/*/Library/Logs`
- `/Users/*/Library/Containers`

- `/Users/*/Library/Application Support`
- `/Users/*/Library/Group Containers`
- `/Users/*/Library/Caches`
- `/Users/*/Library/Developer`

Keep the **Include recommended file path exclusions for Mac** toggle set to **On**. However, you can stop excluding these paths by setting the **Include recommended file path exclusions for Mac** toggle to **Off**.

Set up evidence collection for file activities on devices

When DLP identifies items that match policies on devices, it can copy them to an Azure storage account. Copying matched items to an Azure storage account is useful for auditing policy activity and troubleshooting specific matches. Use the **Set up evidence collection for file activities on devices** settings to add the name and URL of the storage account.

> **Note**
>
> Before you enable evidence collection, create an Azure storage account and a container in that storage account. You must also configure permissions for the account. As you set up your Azure storage account, keep in mind that you probably want to use a storage account that's in the same Azure region or geopolitical boundary as your tenant. Consider configuring Azure storage account access tiers and Azure storage account pricing.

- For more information on copying matched items to Azure storage, see Learn about collecting files that match data loss prevention policies from devices.
- For more information on how to configure this feature, see Get started with collecting files that match data loss prevention policies from devices.

Network share coverage and exclusions

Network share coverage and exclusions extends endpoint DLP policies and actions to new and edited files on network shares and mapped network drives. If just in time protection is also enabled, just in time protection coverage and exclusions are extended to network shares and mapped drives. To exclude a specific network path for all monitored devices, add the path value in **Exclude these network share paths**.

> **Important**
>
> To use **Network share coverage and exclusions**, devices must have the following updates applied:
>
> - Windows 10 - March 21, 2023—KB5023773 (OS Builds 19042.2788, 19044.2788, and 19045.2788) Preview, March 28, 2023—KB5023774 (OS Build 22000.1761) Preview
> - Windows 11 - March 28, 2023—KB5023778 (OS Build 22621.1485) Preview
> - Microsoft Defender April-2023 (Platform: 4.18.2304.8 | Engine: 1.1.20300.3)
> - macOS Last 3 OS versions supported; Minimum Defender app version supported - 101.24122.0005

The following table shows the default settings for network share coverage and exclusions.

Expand table

Network share coverage and exclusions	Just in time protection	Resultant behavior
Enabled	Disabled	- DLP policies scoped to Devices are applied to all network shares and mapped drives that the device connects to. Supported actions: Devices
Disabled	Enabled	- Just-in-time protection is applied only to the files on storage devices that are local to the endpoint.
Enabled	Enabled	- DLP policies scoped to Devices are applied to all network shares and mapped drives that the device connects to. Supported actions: Devices - Just-in-time protection is applied to all network shares and mapped drives that the device connects to.

Network share coverage and exclusions complements DLP On-premises repository actions. The following table shows the exclusion settings and the resulting behavior depending on whether DLP is enabled or disabled for on-premises repositories.

Expand table

Network share coverage and exclusions	DLP on-premises repositories	Resultant behavior
Enabled	Disabled	- DLP policies scoped to Devices are applied to all network shares and mapped drives that the device connects to. Supported actions: Devices
Disabled	Enabled	- Policies that are scoped to On-premises repositories can enforce protective actions on on-premises data-at-rest in file shares and SharePoint document libraries and folders. DLP On-premises repository actions
Enabled	Enabled	- DLP policies scoped to Devices are applied to all network shares and mapped drives that the device connects to. Supported actions: Devices - Policies that are scoped to On-premises repositories can enforce protective actions on on-premises data-at-rest in file shares and SharePoint document libraries and folders. DLP On-premises repository actions

Restricted apps and app groups

Restricted apps

The **Restricted apps** list is a custom list of applications that you create. You configure what actions DLP takes when someone uses an app on the list to ***access*** a DLP-protected file on a device. The **Restricted apps** list is available for Windows 10/11 and macOS devices running any of the three latest macOS releases.

Some apps have a web-based interface in addition to a locally installed version of the application. In preview, when you add an app that users can access both locally and through a web-based interface to a **Restricted app group** or as a **Restricted app**, DLP enforces any applicable policies for accessing a protected file through Edge for the browser app interface and on the device for the application-based interface.

> **Important**
>
> - Don't include the path to the executable for Windows devices. Include only the executable name, such as `browser.exe`.

- The action (`audit`, `block with override`, or `block`) that you define for apps on the restricted apps list only applies when a user attempts to ***access*** a protected item.

When you select **Access by restricted apps** in a policy and a user uses an app on the restricted apps list to access a protected file, the activity is `audited`, `blocked`, or `blocked with override`, depending on how you configured the **Restricted apps** list. If an app on the **Restricted apps** list is also a member of a **Restricted app group**, the actions you configure for activities in the **Restricted app group** override the actions you configure for the **Restricted apps** list. All activity is audited and available for review in activity explorer.

> **Important**
>
> Endpoint DLP protections trigger when a restricted app tries to access a file, either by a user or by the application itself. To exclude a path from enforcement, add it to the exclusion list. This path is where the files identified in Activity Explorer are located.

Restricted app groups

Restricted app groups are collections of apps that you create in DLP settings and then add to a rule in a policy. When you add a restricted app group to a policy, you can take the actions defined in the following table.

Expand table

Restricted App group option	What it allows you to do
Don't restrict file activity	Tells DLP to allow users to access DLP protected items using apps in the app group without taking any action when the user attempts to **Copy to clipboard**, **Copy to a USB removable drive**, **Copy to a network drive**, or **Print** from the app.
Apply a restriction to all activity	Tells DLP to `Audit only`, `Block with override`, or `Block` when a user attempts to access a DLP-protected item using an app that's in the relevant app group
Apply restrictions to a specific activity	This setting allows a user to access a DLP-protected item using an app that is in the app group. It also allows you to select a default action (`Audit only`, `Block`, or `Block with override`) for DLP to take when a user attempts to **Copy to clipboard**, **Copy to a USB removable drive**, **Copy to a network drive**, and **Print**.

You can add a maximum of 50 apps into a single group and you can create a maximum of 10 groups. This gives a maximum of 500 apps that the policy actions can be assigned to.

> **Important**
>
> Settings in a restricted app *group* override any restrictions set in the restricted apps *list* when they are in the same rule. So, if an app is on the restricted apps list and is also a member of a restricted apps group, the settings of the restricted apps group is applied.

Block all apps except for a list of allowed apps

You can create a list of allowed applications and block all others. By creating a list of allowed applications and blocking all others, you don't need to create and manage a comprehensive list of untrusted applications. This feature helps simplify policy management and enhances your control over app-based file activities.

> **Note**
>
> Common background applications such as **teamsupdate.exe** or **svchost.exe** are preconfigured to bypass enforcement to prevent unintentional interference with essential

operations.

In this procedure, you apply the restriction level of **Allow** to explicitly allow activity for a defined app group, and then block any apps that aren't on this list. Because the policy blocks apps that aren't in the allowed group, apps that have no restriction level defined are effectively blocked, and apps that have a restriction level defined as **Allow** are explicitly allowed. You define a restricted app group to allow that app group, but you also define the restricted app group to block any apps that have no defined restrictions.

1. Go to endpoint DLP settings.
2. Define allowed or sanctioned apps in the **Restricted Apps and app groups** list.
3. In your existing or new endpoint DLP policy, locate the **File activities for apps in restricted app groups** setting.
4. Add the desired restricted app group.
5. Select **Apply restriction to all activity**, and select **Allow**. Audit and Off will not work.
6. For all other apps, set the **Access by apps that aren't on the 'unallowed apps' list** setting to **Block**.

How DLP applies restrictions to activities

Interactions between **File activities for apps in restricted app groups**, **File activities for all apps**, and the **Restricted app activities** list are scoped to the same rule.

Restricted app groups overrides

Configurations defined in **File activities for apps in restricted app groups** override the configurations in the **Restricted app activities** list and **File activities for all apps** in the same rule.

Restricted app activities and File activities for all apps

The configurations of **Restricted app activities** and **File activities for all apps** work in concert if the action defined for **Restricted app activities** is either `Audit only`, or `Block with override` in the same rule. Why? Actions defined for **Restricted app activities** only apply when a user accesses a file using an app that's on the list. Once the user has access, the actions defined for activities in **File activities for all apps** apply.

For example, you add Notepad.exe to **Restricted apps** and configure **File activities for all apps** to **Apply restrictions to specific activity**. You configure both as shown in the following table:

Expand table

Setting in policy	App name	User activity	DLP action to take
Restricted app activities	Notepad	Access a DLP protected item	Audit only
File activities for all apps	All apps	Copy to clipboard	Audit only
File activities for all apps	All apps	Copy to a USB removeable device	Block
File activities for all apps	All apps	Copy to a network share	Audit only
File activities for all apps	All apps	Print	Block
File activities for all apps	All apps	Copy or move using unallowed Bluetooth app	Blocked
File activities for all apps	All apps	Remote desktop services	Block with override

When User A opens a DLP-protected file using Notepad, DLP allows the access and audits the activity. While still in Notepad, User A then tries to copy content from the protected item to the

clipboard. This action is successful, and DLP audits the activity. User A then tries to print the protected item from Notepad and the activity is blocked.

> **Note**
>
> When the DLP action to take in **Restricted app activities** is set to `block`, all access is blocked and the user can't perform any activities on the file.

File activities for all apps only

If an app *isn't* in the **File activities for apps in restricted app groups** or the **Restricted app activities** list, or *is* in the **Restricted app activities** list, with an action of either `Audit only`, or `Block with override`, any restrictions defined in the **File activities for all apps** are applied in the same rule.

macOS devices

You can also prevent macOS apps from accessing sensitive data by adding them to the **Restricted app activities** list.

> **Note**
>
> Enter cross-platform apps with their unique paths for the OS they're running on.

To find the full path of Mac apps:

1. On the macOS device, open **Activity Monitor**. Find and double-click the process you want to restrict.
2. Select the **Open Files and Ports** tab.
3. Make a note of the full path name, including the name of the app. For example,

- `/System/Applications/TextEdit.app/Contents/MacOS/TextEdit`

Auto-quarantine

To prevent sensitive items from syncing to the cloud by cloud sync apps such as *onedrive.exe*, add the cloud sync app to the **Restricted apps** list with **Auto-quarantine**.

When enabled, autoquarantine triggers when a restricted app attempts to access a DLP-protected sensitive item. Auto-quarantine moves the sensitive item to an admin-configured folder. If configured to do so, autoquarantine can leave a placeholder (`.txt`) file in place of the original. You can configure the text in the placeholder file to tell users the new location of the item, and other pertinent information.

Use the autoquarantine feature when an unallowed cloud-sync app tries to access an item that is protected by a blocking DLP policy. DLP might generate repeated notifications. You can avoid these repeated notifications by enabling **Auto-quarantine**.

You can also use autoquarantine to prevent an endless chain of DLP notifications for the user and admins. For more information, see Help prevent exposure of sensitive files by configuring auto-quarantine for OneDrive sync

Unsupported file extension exclusions

Use the **Document could not be scanned** condition together with **Apply restrictions to only unsupported file extensions** in your DLP policies to restrict activities involving files with extensions that aren't supported by endpoint DLP. Because this condition can potentially include many unsupported file extensions, refine detection by adding unsupported extensions to exclude.

For more information, see Help protect files that Endpoint Data Loss Prevention doesn't scan and Help protect against sharing of a defined set of unsupported files.

> **Note**
>
> Don't add a `.` when you add an extension. Use the latest anti-malware client version.

> **Important**
>
> When you select **Actions** > **Audit or restrict activities on devices** rule configuration, the **Apply restrictions to only unsupported file extensions** option appears. The **Apply restrictions to only unsupported file extensions** configuration option doesn't support scoping by **Device and device groups** in the policy location setting.

Blocking specific file extensions in DLP policies can lead to unexpected behavior if an application that is marked as unallowed needs to access files with those extensions as part of its normal operation. For example, certain apps might read or temporarily open files, like `.dll`, `.json`, `.tmp` during routine processes such as rendering, caching, or validating content. If you block these extensions, the app might fail to function properly, causing errors, incomplete workflows, or enforcement pop-ups unrelated to user intent. Before implementing extension-based restrictions, make sure you know which apps interact with these file types during standard operations, and whether alternative controls, such as app restrictions or contextual rules, can achieve the security goal without disrupting functionality.

Unallowed Bluetooth apps

To prevent people from transferring files protected by your policies via specific Bluetooth apps, add those apps to the **Unallowed Bluetooth apps** list in Endpoint DLP settings.

Browser and domain restrictions to sensitive data

Restrict sensitive files that match your policies from being shared with unrestricted cloud service domains.

Unallowed browsers

For Windows devices, you can restrict the use of specified web browsers by their executable names. The specified browsers can't access files that match the conditions of an enforced DLP policy where the upload-to-cloud services restriction is set to `block` or `block override`. When these browsers are blocked from accessing a file, end users see a toast notification asking them to open the file through Microsoft Edge.

For macOS devices, you must add the full file path. To find the full path of Mac apps:

1. On the macOS device, open **Activity Monitor**. Find and double-click the process you want to restrict.
2. Choose the **Open Files and Ports** tab.
3. Make sure to make a note of the full path name, including the name of the app.

Service domains

The **Service domains** setting works with the **Audit or restrict activities on devices** setting in the workflow for creating a rule within a DLP policy.

When you create a rule, use actions to protect your content when certain conditions are met. When creating rules for endpoint devices, choose the **Audit or restrict activities on devices** option, and select one of these options:

- Audit only
- Block with override
- Block

To control whether sensitive files that your policies protect can be uploaded to specific service domains, go to **Endpoint DLP Settings** > **Browser and domain restrictions to sensitive data** and choose whether to *block* or *allow* **Service domains** by default.

> ⓘ **Note**
>
> The **Service domains** setting only applies to files uploaded by using Microsoft Edge, or by using instances of Google Chrome or Mozilla Firefox that have the Microsoft Purview Chrome Extension installed.

Block service domains

When you set the **Service domains** list to **Block**, use the **Add cloud service domain** to specify domains that you want to block. All other service domains are allowed. When the Service domains list is set to Block, DLP policy restrictions only apply when a user attempts to upload a sensitive file to any of the domains on the list.

For example, consider the following configurations:

- A DLP policy is configured to detect sensitive items that contain physical addresses and the **Audit or restrict activities on devices** option is set to **Audit only**.
- The **Service domains** setting is set to **Block**.
- contoso.com IS NOT on the list.
- wingtiptoys.com IS on the list.

In this case, if a user attempts to upload a sensitive file with physical addresses to contoso.com, the upload is allowed to complete and an audit event is generated but no alert is triggered.

In contrast, if a user attempts to upload a sensitive file with credit card numbers to wingtiptoys.com, the user activity - the upload - is also allowed to complete and both an audit event and an alert are generated.

Another example, consider the following configuration:

- A DLP policy is configured to detect sensitive items that contain physical addresses and the **Audit or restrict activities on devices** option is set to **Block**.
- The **Service domains** setting is set to **Block**.
- contoso.com IS NOT on the list.
- wingtiptoys.com IS on the list.

In this case, if a user attempts to upload a sensitive file with physical addresses to contoso.com, the upload is allowed to complete and an audit event is generated but no alert is triggered.

In contrast, if a user attempts to upload a sensitive file with credit card numbers to wingtiptoys.com, the user activity - the upload - is blocked and both an audit event and an alert are generated.

Allow service domains

When you set the **Service domains** list to **Allow**, use the **Add cloud service domain** to specify domains that are allowed. All other service domains have DLP policy restrictions enforced. When the Service domains list is set to Allow, DLP policies only apply when a user attempts to upload a sensitive file to any of the listed domains.

For example, here are two starting configurations:

- A DLP policy is configured to detect sensitive items that contain credit card numbers and the **Audit or restrict activities on devices** option is set to **Block with override**.
- The **Service domains** setting is set to **Allow**.
- contoso.com IS NOT on the **Allow** list.
- wingtiptoys.com IS on the **Allow** list.

In this case, if a user attempts to upload a sensitive file with credit card numbers to contoso.com, the upload is blocked, a warning displays, giving the user the option to override the block. If the user chooses to override the block, an audit event is generated and an alert is triggered.

However, if a user attempts to upload a sensitive file with credit card numbers to wingtiptoys.com, the policy restriction *isn't* applied. The upload is allowed to complete, and an audit event is generated but no alert is triggered.

- A DLP policy is configured to detect sensitive items that contain physical addresses and the **Audit or restrict activities on devices** option is set to **Audit only**.
- The **Service domains** setting is set to **Allow**.
- contoso.com IS NOT on the list.
- wingtiptoys.com IS on the list.

In this case, if a user attempts to upload a sensitive file with physical addresses to contoso.com, the upload is allowed to complete and both an audit event and an alert are generated.

In contrast, if a user attempts to upload a sensitive file with credit card numbers to wingtiptoys.com, the user activity - the upload - is also allowed to complete, an audit event is generated but no alert is triggered.

> **Important**
>
> When you set the service restriction mode to **Allow**, you must configure at least one service domain before the system enforces restrictions.

Summary table: Allow or block behavior

The following table shows how the system behaves depending on the settings listed.

Expand table

Endpoint DLP Service domain setting	DLP policy rule Audit or restrict activities on devices setting	User goes to a listed site	User goes to a site NOT listed
Allow	Audit only	- User activity is audited - No alert is generated - No DLP policies are applied	- User activity is audited - An alert is generated - DLP policies are applied in Audit mode
Allow	Block with override	- User activity is audited - No alert is generated - No DLP policies are applied	- User activity is audited - An alert is generated - DLP policies are applied in Block with override mode
Allow	Block	- User activity is audited - No alert is generated - No DLP policies are applied	- User activity is audited - An alert is generated - DLP policies are applied in Block mode
Block	Audit only	- User activity is audited - An alert is generated - DLP policies are applied in Audit mode	- User activity is audited - No alert is generated - No DLP policies are applied

Endpoint DLP Service domain setting	DLP policy rule Audit or restrict activities on devices setting	User goes to a listed site	User goes to a site NOT listed
Block	Block with override	- User activity is audited - An alert is generated - DLP policies are applied in Block with override mode	- User activity is audited - No alert is generated - No DLP policies are applied
Block	Block	- User activity is audited - An alert is generated - DLP policies are applied in Block mode	- User activity is audited - No alert is generated - No DLP policies are applied

When you add a domain to the list, use the FQDN format of the service domain without the ending period (`.`). Use `*.` as a wildcard to specify all domains or subdomains. Exclude the protocol (anything before `//`) from the domain. Only include the host name, without any subsites.

For example:

Expand table

Input	URL matching behavior
contoso.com	**Matches the specified domain name, and any subsite**: ://contoso.com ://contoso.com/ ://contoso.com/anysubsite1 ://contoso.com/anysubsite1/anysubsite2 (etc.) **Does not match sub-domains or unspecified domains**: ://anysubdomain.contoso.com ://anysubdomain.contoso.com.AU
*.contoso.com	**Matches the specified domain name, any subdomain, and any site**: ://contoso.com ://contoso.com/anysubsite ://contoso.com/anysubsite1/anysubsite2 ://anysubdomain.contoso.com/ ://anysubdomain.contoso.com/anysubsite/ ://anysubdomain1.anysubdomain2.contoso.com/anysubsite/ ://anysubdomain1.anysubdomain2.contoso.com/anysubsite1/anysubsite2 (etc.) **Does not match unspecified domains** ://anysubdomain.contoso.com.AU/
`www.contoso.com`	**Matches the specified domain name**: `www.contoso.com` **Does not match unspecified domains or subdomains** ://anysubdomain.contoso.com/, in this case, you have to put the FQDN domain name itself `www.contoso.com`

You can configure up to 50 domains under **Sensitive Service domains**.

> **Note**

The Service domains list setting only applies to file uploads to websites. Actions like pasting into a browser don't follow the Service Domain list.

Sensitive service domain groups

When you add a website to **Sensitive service domains**, you can set the policy to `audit`, `block with override`, or fully `block` user activity when users attempt to take any of the following actions:

- print from a website
- copy data from a website
- save a website as local files
- upload or drag/drop a sensitive file to an excluded website
- paste sensitive data to an excluded website

The following table shows which browsers support these features:

Expand table

Browser	Supported Feature
Microsoft Edge	- Print the site - Copy data from the site - Save the site as local files (save-as) - Paste to supported browsers - Upload to a restricted cloud service domain
Google Chrome (with the Microsoft Purview extension)	- Paste to supported browsers - Upload to a restricted cloud service domain
Mozilla Firefox (with the Microsoft Purview extension)	- Upload to a restricted cloud service - Paste to supported browsers

For the **Paste to supported browsers** action, there might be a brief time lag between when the user attempts to paste text into a web page and when the system finishes classifying it and responds. If this classification latency happens, you might see both policy-evaluation and check-complete notifications in Microsoft Edge or policy-evaluation toast on Chrome and Firefox. Here are some tips for minimizing the number of notifications:

1. Notifications trigger when a policy for the target website is configured to **Block** or **Block with override** the **Paste to supported browsers** for that user. You can configure the overall action to **Audit** and then use the exceptions to **Block** the target websites. Alternately, you can set the overall action to **Block** and then use the exceptions to **Audit** the secure websites.
2. Use the latest Antimalware client version.
3. Ensure your version of Microsoft Edge is 120 or higher.
4. Install these Windows KBs:
   a. Windows 10: KB5032278, KB5023773
   b. Windows 11 21H2: KB5023774
   c. Win 11 22H2: KB5032288, KB5023778
5. On macOS, ensure your antimalware Client version is 101.25022.0003 or later.

The **Paste to supported browsers** action doesn't follow the behavior defined in the Service Domain list. However, if Sensitive Service Domain Groups are configured on the rule for Paste To Browser, the system honors the configured Sensitive Service Domain Groups.

> **Note**
>
> The **Service domains** setting only applies to files uploaded by using Microsoft Edge or an instance of Google Chrome or Mozilla Firefox that has the Microsoft Purview Chrome Extension installed. The Generative AI Websites group contains these supported sites. The

group is used for default policies within Data Security Posture Management for AI and can't be edited or deleted.

For devices, you must configure the **Sensitive service domains** list to use the **Upload to a restricted cloud service domain** action in a DLP policy. You can also define website groups that you want to assign policy actions to that are different from the global website group actions. You can add a maximum of 100 websites into a single group and you can create a maximum of 150 groups. This configuration gives a maximum of 15,000 websites that the policy actions can be assigned to. For more information, see Help prevent risky user activity by monitoring or restricting access to sensitive service domains.

> **Important**
>
> Regarding the **Paste to supported browser** action. If 'Collect original file as evidence for all selected file activities on Endpoint' is enabled on the rule for this feature, garbage characters might appear in the source text if the user's Windows device doesn't have Antimalware Client Version 4.18.23110 or newer installed. Select **Actions** > **Download** to view the actual content.

For more information, see Help prevent leakage of sensitive content by restricting paste actions into browsers.

Supported syntax for designating websites in a website group

If you use URLs to identify websites, don't include the networking protocol as part of the URL (for instance, *https://* or *file://*). Instead, use a flexible syntax to include and exclude domains, subdomains, websites, and subsites in your website groups. For example,

- Use `*.` as a wildcard to specify all domains or all subdomains.
- Use `/` as a terminator at the end of a URL to scope to that specific site only.

When you add a URL without a terminating slash mark (`/`), that URL is scoped to that site and all subsites. You can only add `*.` to the beginning of a domain. The `/` terminator is only supported at the end of a domain.

This syntax applies to all http/https websites. Here are some examples:

Expand table

URL added to the website group	URL matches	URL doesn't match
contoso.com	http://contoso.com https://contoso.com https://contoso.com/ https://contoso.com/allsubsites1 https://contoso.com/allsubsites1/allsubsites2	https://allsubdomains.contoso.com https://allsubdomains.contoso.com.a https://www.contoso.com
contoso.com/	http://contoso.com https://contoso.com https://contoso.com/	https://contoso.com/allsubsites1 https://contoso.com/allsubsites1/alls https://allsubdomains.contoso.com https://allsubdomains.contoso.com/ https://www.contoso.com
*.contoso.com	http://contoso.com https://contoso.com https://www.contoso.com https://www.contoso.com/allsubsites https://contoso.com/allsubsites1/allsubsites2 https://allsubdomains.contoso.com https://allsubdomains.contoso.com/allsubsites https://allsubdomains1.allsubdomains2.contoso.com/allsubsites1/allsubsites2	https://allsubdomains.contoso.com.a
*.contoso.com/xyz	http://contoso.com/xyz/ https://contoso.com/xyz/	https://contoso.com/xyz https://allsubdomains.contoso.com/

URL added to the website group	URL matches	URL doesn't match
	https://contoso.com/xyz/allsubsites/ https://allsubdomains.contoso.com/xyz/ https://allsubdomains.contoso.com/xyz/allsubsites https://allsubdomains1.allsubdomains2.contoso.com/xyz/allsubsites https://allsubdomains1.allsubdomains2.contoso.com/xyz/allsubsites1/allsubsites2	
*.contoso.com/xyz/	http://contoso.com/xyz https://contoso.com/xyz https://contoso.com/xyz/ https://allsubdomains.contoso.com/xyz https://allsubdomains.contoso.com/xyz/	https://contoso.com https://contoso.com/xyz/allsubsites/ https://allsubdomains.contoso.com/ https://allsubdomains1.allsubdomair https://allsubdomains1.allsubdomair

Supported syntax for designating IP ranges or IP addresses (preview)

In preview, DLP supports using IP address and address ranges to identify websites. Set the Match type to *IP address* or *IP address range* and then enter a specific IP address or an IP range in the **Sensitive service domain** field. Select **Add site** to add the selection to the Sensitive service domain group.

Examples of supported syntax:

- 1.1.1.1
- 1.1.1.1-2.2.2.2
- 2001:0db8:85a3:0000:0000:8a2e:0370:7334
- 2001:0db8:85a3:0000:0000:8a2e:0370:7320-2001:0db8:85a3:0000:0000:8a2e:0370:7334

> **Important**
>
> URLs support these actions:

- **Print the site**
- **Copy data from the site**
- **Save the site as local files (save-as)**
- **Paste to supported browsers**
- **Upload to a restricted cloud service domain**

> IP address and IP address range support these actions:

- **Print the site**
- **Copy data from the site**
- **Save the site as local files (save-as)**
- **Upload to a restricted cloud service domain** (Windows only)

Sensitive service domain groups contain a preconfigured group for **Generative AI websites**. For a list of all the websites in this group, see List of AI sites supported by Microsoft Purview Data Security Posture Management for AI.

Additional settings for Endpoint DLP

Business justification in policy tips

You can control how users interact with the business justification option in Options for configuring policy tips. This option appears when users perform an activity that's protected by the **Block with override** setting in a DLP policy. The business justification setting is global. Choose from one of the following options:

- **Show default options and custom text box**: By default, users can select either a built-in justification or enter their own text.

- **Only show default options**: Users can only select from a list of built-in justifications.
- **Only show custom text box**: Users can only enter a custom justification. The text box appears in the end-user policy tip notification, without a list of options.

Customizing the options in the drop-down menu

You can create up to five customized options that appear when users interact with the policy notification tip by selecting **Customize the options drop-down menu**.

Expand table

Option	Default text
option 1	**This is part of an established business workflow** or enter customized text
option 2	**My manager has approved this action** or enter customized text
option 3	**Urgent access required; I'll notify my manager separately** or enter customized text
Show false positive option	**The information in these files is not sensitive** or enter customized text
option 5	**Other** or enter customized text

Turn on automatic diagnostic logging for endpoint DLP

Microsoft Purview Always-on diagnostics feature automatically records comprehensive trace logs, saving you time and enabling faster troubleshooting. For more information, see Always-on diagnostics for endpoint DLP.

Enable Endpoint DLP for Windows Servers

Endpoint DLP supports Windows Server 2019 and later versions.

After you onboard a Windows Server, turn on Endpoint DLP support to apply endpoint protection.

To work with the DLP alert management dashboard:

1. In the Microsoft Purview portal, go to **Data loss prevention** > **Overview**.
2. Choose **Settings** in the upper right corner.
3. On **Settings**, select **Endpoint settings** and expand **Endpoint DLP support for onboarded servers**.
4. Set the toggle to **On**.

Always audit file activity for devices

By default, when you onboard devices, the system automatically audits activity for Office, PDF, and CSV files. You can review this activity in activity explorer. Turn off this feature if you want to audit this activity only when onboarded devices are included in an active policy. The **Always audit file activity for devices** setting enables the auditing of file activities for documents where a DLP Rule didn't match: File Created, File Modified, File Renamed, File created on removable media, and File created on network share.

The system always audits file activity for onboarded devices, regardless of whether they're included in an active policy.

Printer groups

Use this setting to define groups of printers that you want to assign policy actions to that are different from the global printing actions.

The most common use case for creating printer groups is to use them for limiting the printing of contracts to only those printers in an organization's Legal department. After you define a printer group here, you can use it in all of your policies that are scoped to **Devices**. For more information on configuring policy actions to use authorization groups, see Create policy to manage printer access using authorization groups.

You can create a maximum of 20 printer groups. Each group can contain a maximum of 50 printers.

> ⓘ **Note**
>
> Printer groups are available for devices running any of the following Windows versions:
>
> - Windows 10 and later (21H1, 21H2, and later) - KB5020030
> - Win 11 21H2 - KB5019157
> - Win 11 22H2 - KB5020044
> - Windows Server 2022 - KB5020032

Let's look at an example. Say you want your DLP policy to block printing of contracts to all printers except for those that are in the legal department.

1. Use the following parameters to assign printers in each group.

    - **Friendly printer** name - the Friendly printer name is the value showing on the UX when you select the Printer.
    - **USB printer** - A printer connected through a computer's USB port. Select this option if you want to enforce any USB printer while leaving the USB product ID and USB vendor ID unselected. You can also assign a specific USB printer by specifying its USB product ID and USB vendor ID.
    - **USB product ID** - Get the **Device Instance** path value from the printer device property details in device manager. Convert that value the to Product ID and Vendor ID format. For more information, see Standard USB identifiers.
    - **USB vendor ID** - Get the **Device Instance** path value from the printer device property details in device manager. Convert that value to the Product ID and Vendor ID format. For more information, see Standard USB identifiers.
    - **IP range**
    - **Print to file** - Microsoft Print to PDF or Microsoft XPS Document Writer. If you only want to enforce Microsoft Print to PDF, you should use Friendly printer name with 'Microsoft Print to PDF'.
    - **Universal print deployed on a printer** - For more information on universal printers, see Set up Universal Print.
    - **Corporate printer** - is a print queue shared through on-premises Windows print server in your domain. Its path might look like this: \print-server\contoso.com\legal_printer_001.
    - **Print to local** - Any printer connecting through Microsoft print port but not any of above types. For example: print through remote desktop or redirect printer.

> ⓘ **Note**

You shouldn't use multiple parameters of **USB printer**, **IP range**, **Print to file**, **Universal print deployed on a printer**, **Corporate printer**, and **Print to local**.

2. Assign each printer in the group a **Display name**. These names appear only in the Microsoft Purview console.
3. Create a printer group named **Legal printers** and add individual printers (with an alias) by their friendly name; for instance: `legal_printer_001`, `legal_printer_002`, and `legal_color_printer`. (You can select multiple parameters at once to help you unambiguously identify a specific printer.)
4. Assign the policy actions to the group in a DLP policy:
   - `Allow` (audit with no user notifications or alerts)
   - `Audit only` (you can add notifications and alerts)
   - `Block with override` (blocks the action, but the user can override)
   - Preview, a fix to address the unnecessary resubmission to queue the print job after initial override, has been implemented.
   - `Block` (blocks no matter what)

Create a printer group

1. Open the Microsoft Purview portal and go to **Data Loss Prevention** > **Overview** > settings gear icon in the upper right corner > **Data Loss Prevention** > **Endpoint DLP settings** > **Printer groups**.
2. Select + **Create printer group**.
3. Enter a name for the group.
4. Select **Add printer**.
5. Enter a **Friendly name** for the printer. Make sure the name matches the value from the printer's device property details in Device Manager.
6. Select the parameters and provide the values to unambiguously identify the specific printer.
7. Select **Add**.
8. Add other printers as needed.
9. Select **Save** and then **Close**.

File extension groups

Use this setting to define groups of file extensions that you want to assign policy actions to. For example, you can apply a **File could not be scanned** policy only to file extensions in the created groups.

> ⓘ **Note**
>
> Don't add a `.` when you add an extension.

Disable classification

Use this setting to exclude specific file extensions from Endpoint DLP classification.

For files that are on the Monitored files list, you can disable classification through this setting. When you add a file extension to this setting, Endpoint DLP doesn't scan content in files with this extension. As a result, Endpoint DLP doesn't perform policy evaluation based on the content of those files. You can't see content information for the purposes of conducting investigations.

> ⓘ **Note**

Don't add a `.` when you add an extension.

Removable USB device groups

Use this setting to define groups of removable storage devices, such as USB thumb drives, that you want to assign policy actions to that are different from the global removable storage actions. For example, say you want your DLP policy to block items with engineering specifications from being copied to removable storage devices, except for designated USB-connected hard drives that are used to back up data for offsite storage.

You can create up to 20 groups, with up to 50 removable storage devices in each group.

> **Note**
>
> Removable USB device groups are available for devices running any of the following Windows versions:
>
> - Windows 10 and later (21H1, 21H2) with KB 5018482
> - Windows 11 21H2, 22H2 with KB 5018483
> - Windows 10 RS5 (KB 5006744) and Windows Server 2022

Use the following parameters to define your removable storage devices.

- **Storage device friendly name** - Get the Friendly name value from the storage device property details in device manager. Wildcard values (*) are supported.
- **USB product ID** - Get the Device Instance path value from the USB device property details in device manager. Convert it to Product ID and Vendor ID format. For more information, see Standard USB identifiers.
- **USB vendor ID** - Get the Device Instance path value from the USB device property details in device manager. Convert it to Product ID and Vendor ID format. For more information, see Standard USB identifiers.
- **Serial number ID** - Get the serial number ID value from the storage device property details in device manager. Wildcard values (*) are supported.
- **Device ID** - Get the device ID value from the storage device property details in device manager. Wildcard values (*) are supported.
- **Instance path ID** - Get the device ID value from the storage device property details in device manager. Wildcard values (*) are supported.
- **Hardware ID** - Get the hardware ID value from the storage device property details in device manager. Wildcard values (*) are supported.

You assign each removable storage device in the group an **Alias**. The alias is a friendly name that only appears in the Microsoft Purview console. So, continuing with the example, you create a removable storage device group named **Backup** and add individual devices (with an alias) by their friendly name, like `backup_drive_001`, and `backup_drive_002`.

You can multiselect the parameters and then the printer group includes all devices that satisfy those parameters.

You can assign these policy actions to the group in a DLP policy:

- `Allow` (audit with no user notifications or alerts)
- `Audit only` (you can add notifications and alerts)
- `Block with` override (blocks the action, but the user can override)
- `Block` (blocks no matter what)

Create a removable USB device group

1. Open the Microsoft Purview portal and go to **Data Loss Prevention** > **Overview** > settings gear icon in the upper right corner > **Data Loss Prevention** > **Endpoint DLP settings** > **Removable USB device groups**.
2. Select + **Create removable storage device group**.
3. Enter a **Group name**.
4. Select **Add removable storage device**.
5. Enter an **Alias**.
6. Select the parameters and enter the values to clearly identify the specific device.
7. Select **Add**.
8. Add other devices to the group as needed.
9. Select **Save** and then **Close**.

The most common use case for creating removable storage groups is to specify which removable storage devices users can copy files to. Generally, copying is only allowed for devices in a designated **Backup** group.

After you define a removable storage device group, you can use it in all of your policies that are scoped to **Devices**. For more information about configuring policy actions to use authorization groups, see Create policy to manage printer access using authorization groups.

Network share groups

Use this setting to define groups of network share paths that you want to assign policy actions to that are different from the global network share path actions. For example, say you want your DLP policy to prevent users from saving or copying protected files to network shares except the network shares in a particular group.

> **Note**
>
> Network share groups are available for devices running any of the following Windows versions:
>
> - Windows 10 and later (21H1, 21H2) with KB 5018482
> - Windows 11 21H2, 22H2 with KB 5018483
> - Windows 10 RS5 (KB 5006744) and Windows Server 2022

To include network share paths in a group, define the prefix that they all the shares start with. For example:

- '\Library' matches:
  - \Library folder and all its subfolders.
- You can use wildcards, for example '\Users*\Desktop' matches:
  - '\Users\user1\Desktop'
  - '\Users\user1\user2\Desktop'
  - '\Users*\Desktop'
- You can also use environmental variables, for example:
  - %AppData%\app123
- Wildcard values are supported. So a path definition can contain an asterisk (`*`) in the middle of the path or at the end of the path.
  Example: `\\Lib*` covers `\\Libray`

You can assign the following policy actions to the group in a DLP policy:

- `Allow` (audit with no user notifications or alerts)
- `Audit only` (you can add notifications and alerts)

- `Block with override` (blocks the action, but the user can override)
- `Block` (blocks no matter what)

After you define a network share group, you can use it in all of your DLP policies that are scoped to **Devices**. For more information about configuring policy actions to use authorization groups, see Create policy to manage printer access using authorization groups.

Create a Network Share group

1. Open Microsoft Purview portal and go to **Data Loss Prevention** > **Overview** > settings gear icon in the upper right corner > **Data Loss Prevention** > **Endpoint DLP settings** > **Network share groups**.
2. Select + **Create network share group**.
3. Enter a **Group name**.
4. Add the file path to the share.
5. Select **Add**.
6. Add other share paths to the group as needed.
7. Select **Save** and then **Close**.

VPN settings

Use the VPN list to control only those actions that are carried out over that VPN.

> **Note**
>
> VPN settings are available for devices running any of these versions of Windows:
>
> - Windows 10 and later (21H1, 21H2) with KB 5018482
> - Windows 11 21H2, 22H2 with KB 5018483
> - Windows 10 RS5 (KB 5006744)

When you list a VPN in **VPN Settings**, you can assign the following policy actions to it:

- `Allow` (audit with no user notifications or alerts)
- `Audit only` (you can add notifications and alerts)
- `Block with override` (blocks the action, but the user can override)
- `Block` (blocks no matter what)

You can apply these actions individually or collectively to the following user activities:

- Copy to clipboard
- Copy to a USB removable device
- Copy to a network share
- Print
- Copy or move using unallowed (restricted) Bluetooth app
- Copy or move using RDP

When you configure a DLP policy to restrict activity on devices, you can control what happens to each activity performed when users are connected to your organization within any of the VPNs listed.

Use the **Server address** or **Network address** parameters to define the VPN allowed.

Get the server address or network address

1. On a DLP monitored Windows device, open a **Windows PowerShell** window as an administrator.
2. Run the following cmdlet. It returns multiple fields and values.

```
```

PowerShell

```
Get-VpnConnection
```

1. Among the results of the cmdlet, find the **ServerAddress** field and record that value. Use the **ServerAddress** when you create a VPN entry in the VPN list.
2. Find the **Name** field and record that value. The **Name** field maps to the **Network address** field when you create a VPN entry in the VPN list.

Add a VPN

1. Open Microsoft Purview portal and go to **Data Loss Prevention** > **Overview** > settings gear icon in the upper right corner > **Data Loss Prevention** > **Endpoint DLP settings** > **VPN settings**.
2. Select **Add or edit VPN addresses**.
3. Enter either the **Server address** or **Network address** that you recorded after running `Get-VpnConnection`.
4. Select **Save**.
5. Close the item.

> **Important**
>
> Under the **Network restrictions** setting, you also see **Corporate network** as an option. **Corporate network** connections are all connections to your organization's resources. You can see if a device is using a **Corporate network** by running the `Get-NetConnectionProfile` cmdlet as an administrator. If the `NetworkCategoryId` in the output is `DomainAuthenticated`, it means the machine is connected to the **Corporate network**. If the output is anything else, the machine isn't. In some cases, a machine can be both VPN connected and **Corporate network** connected. If both are selected under the **Network restrictions**, Endpoint DLP applies the action based on the order. If you want the action for VPN to be the one that's applied, move the VPN entry above **Corporate network** to have higher priority than the action for **Corporate network**.

> See Create policy to manage file activities by implementing network exceptions for more information on configuring policy actions to use network exceptions.

See also

- Learn about Endpoint data loss prevention
- Get started with Endpoint data loss prevention
- Learn about data loss prevention
- Get started with Activity explorer
- Microsoft Defender for Endpoint
- Onboard Windows 10 and Windows 11 devices into Microsoft Purview overview
- Microsoft 365 subscription
- Microsoft Entra joined
- Download the new Microsoft Edge based on Chromium
- Create and Deploy data loss prevention policies

Note: The author created this article with assistance from AI. Learn more

Last updated on 06/26/2026

Protect FTP and SFTP transfers with Microsoft Purview Endpoint DLP

Endpoint data loss prevention (DLP) in Microsoft Purview can monitor and protect sensitive data transferred over FTP and SFTP from managed Windows devices. Endpoint DLP evaluates these transfers against your policy rules, just as it does other device activities such as copying to USB devices or network shares. You can review the resulting events in Activity explorer.

This protection isn't turned on by default. This article explains how to add FTP and SFTP transfers to an Endpoint DLP policy rule and review the resulting events.

Before you begin, review the device onboarding and permissions requirements in Prerequisites.

Prerequisites

- Devices onboarded to Microsoft Purview.
- Permissions to create and deploy DLP policies and view Activity explorer. Use the least-privileged roles needed for these tasks.

Create the policy

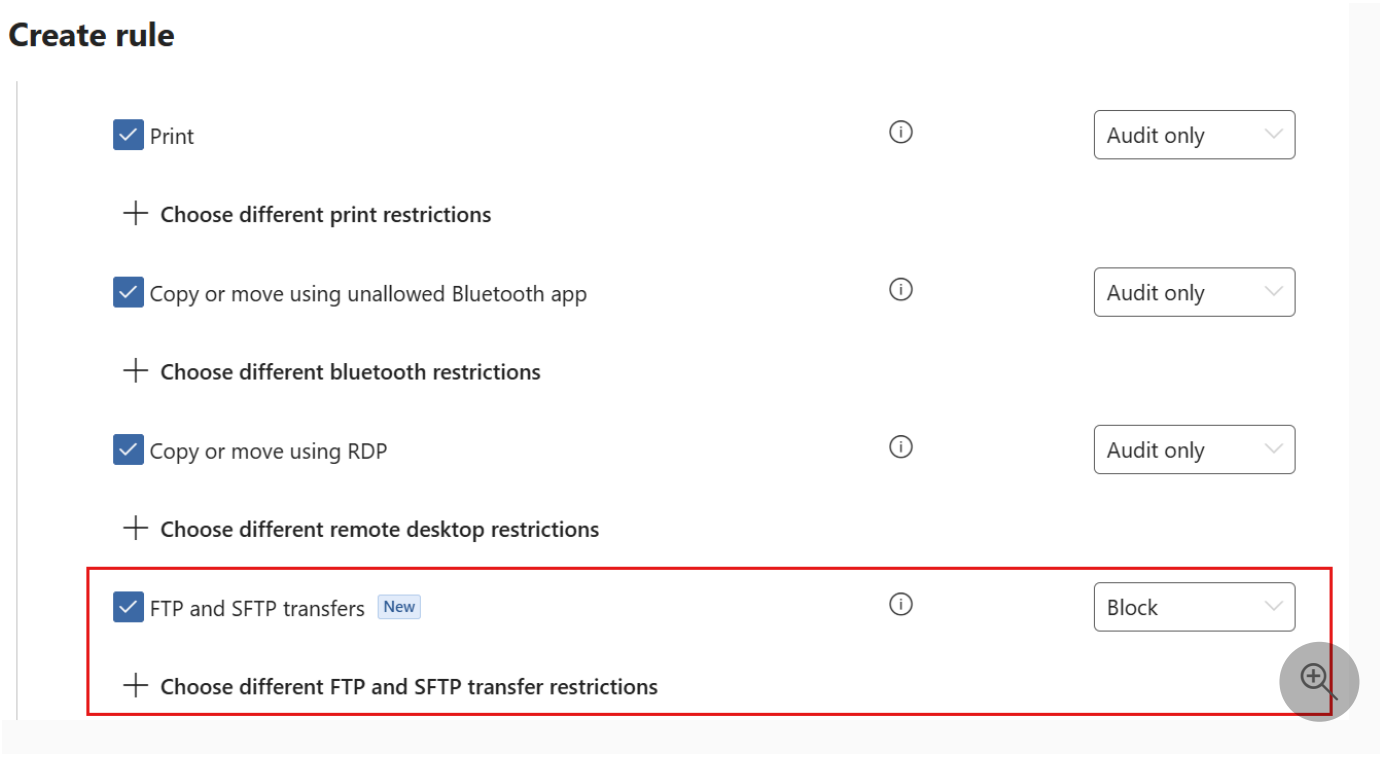
We recommend starting with a small pilot group and setting the action to Audit only. Review Activity explorer to understand how users transfer files over FTP/SFTP before you deploy the policy broadly.

To create the policy:

1. In the Microsoft Purview portal, go to Data loss prevention > Policies > Create policy.
2. On the Locations page, select Devices, and then scope the policy to the users or device groups you want.
3. Create a rule, and add the conditions you want to detect, such as sensitive information types or sensitivity labels.
4. Under Actions, select Audit or restrict activities on devices.
5. In File activities for all apps, select Apply restrictions to specific activity.
6. Select the FTP and SFTP transfers checkbox, and then choose an action. Supported actions are Audit only, Block, Block with override, and Allow.
7. (Optional) Turn on policy tips under User notifications.

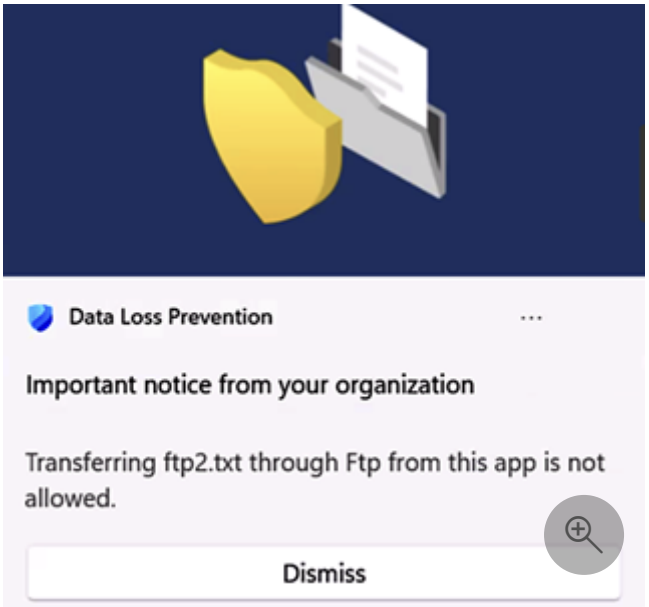
8. Review and publish the policy.

The Create rule page shows FTP and SFTP transfers selected with its action set to Block.



End-user notification

Example of an end-user notification for a blocked FTP transfer:



Review activity

Use Activity explorer to review the details of each transfer.

1. Open Activity explorer.

2. Filter Activity to File transfer by FTP/SFTP. Review the user, device, file, matched policy, and action for each transfer.

Collect full file evidence

FTP and SFTP transfers support full file evidence (FFE) collection. Enable evidence collection in Endpoint DLP settings and configure it in your DLP policy. When a user attempts to transfer a matched file, a complete copy can be saved.

For requirements and setup, see Get started with evidence collection for file activities on devices.

Related content

- Learn about Endpoint data loss prevention
- Configure endpoint DLP settings

ⓘ**Note:** This article contains content created with AI. Learn more

---

Last updated on 10/16/2026

Endpoint DLP policy scenarios in Microsoft Purview

This article provides guidance on using Endpoint Data Loss Prevention (DLP) capabilities within Microsoft Purview to help safeguard sensitive information across organizational devices. It focuses on practical, scenario-based approaches for creating and modifying DLP policies, including auditing, alerting, blocking actions, and enforcing controls on how sensitive data is accessed, shared, or transferred. By addressing common administrative use cases, this guidance helps define how policies can be applied across endpoints and integrated services to monitor and regulate user activity effectively.

By applying these scenarios, administrators can reduce data loss risks, support compliance, and gain better visibility into user activity. This scenario-based approach helps strengthen data governance and improves the organization's overall security posture.

> **Important**
>
> These Endpoint DLP scenarios aren't the official procedures for creating and tuning DLP policies. For general situations, see the following topics:
>
> - Learn about Microsoft Purview Data Loss Prevention
> - Create and Deploy data loss prevention policies

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Before you begin

SKU and subscription licensing

For information on licensing, see

- Microsoft 365 Enterprise Plans
- Microsoft 365 Service Descriptions

These scenarios require that you already onboard devices and that they report to Activity explorer. If you didn't onboard devices yet, see Get started with Endpoint data loss prevention.

> **Important**
>
> Before you start, make sure you understand the difference between an unrestricted administrator and an administrative unit restricted administrator. For more information, see Administrative units.

Endpoint DLP policy scenarios

The following scenarios show common Endpoint DLP policy configurations you can implement:

- Create policy to audit activities using a template (audit only mode)
- Create policy to manage printer access using authorization groups
- Create policy to detect and alert on U.S. PII data exposure
- Help prevent unauthorized sensitive data sharing with block actions and allow overrides
- Help prevent risky user activity by monitoring or restricting access to sensitive service domains
- Help prevent leakage of sensitive content by restricting paste actions into browsers
- Help prevent exposure of sensitive files by configuring auto-quarantine for OneDrive sync
- Help prevent sharing of sensitive items with unauthorized cloud apps and services
- Create policy to manage file activities by implementing network exceptions

See also

- Learn about Endpoint data loss prevention
- Get started with Endpoint data loss prevention
- Learn about data loss prevention
- Get started with Activity explorer
- Microsoft Defender for Endpoint
- Onboard Windows 10 and Windows 11 devices into Microsoft Purview overview
- Microsoft 365 subscription
- Microsoft Entra joined
- Download the new Microsoft Edge based on Chromium
- Create and Deploy data loss prevention policies

Note: The author created this article with assistance from AI. Learn more

Last updated on 06/26/2026

Create policy to audit activities using a template (audit only mode)

This scenario shows how to create a Microsoft Purview DLP policy using the U.S. PII Data Enhanced template to monitor sensitive data on endpoint devices. The policy runs in audit-only (simulation) mode, enabling visibility into user activity and data risk without impacting users, while using policy tips and Activity explorer for validation and analysis.

This scenario assumes an unrestricted administrator creating a policy with full directory scope.

Prerequisites and assumptions

This article uses the process you learned in Design a data loss prevention policy to show you how to create a Microsoft Purview Data Loss Prevention (DLP) policy. Work through these scenarios in your test environment to familiarize yourself with the policy creation UI.

> **Important**
>
> This article presents a hypothetical scenario with hypothetical values. It's only for illustrative purposes. Substitute your own sensitive information types, sensitivity labels, distribution groups, and users.

How you deploy a policy is as important policy design. This article shows you how to use the deployment options so that the policy achieves your intent while avoiding costly business disruptions.

This scenario uses the *Confidential* sensitivity label, so you must create and publish sensitivity labels before proceeding. To learn more, see:

- Learn about sensitivity labels
- Get started with sensitivity labels
- Create and configure sensitivity labels and their policies

This procedure uses a hypothetical distribution group *Human Resources* and a distribution group for the security team at Contoso.com.

This policy-creation scenario uses alerts. To learn more, see Get started with the data loss prevention alerts.

Policy intent statement and mapping

We, Contoso, want to begin evaluating how sensitive information—specifically U.S. Personally Identifiable Information (PII)—is being handled on endpoint devices across our organization. Before enforcing restrictions, we want visibility into user activity and potential data risk without impacting productivity.

To achieve this, we will create a policy that monitors activity on devices, detects when U.S. PII is involved, and records those events for analysis. When users perform actions that match the policy conditions, we want to show policy tips to raise awareness and educate them, but we do not want to block or restrict their actions at this stage.

Expand table

Statement	Configuration question answered and configuration mapping
"We want to evaluate how sensitive (U.S. PII) data is being used on endpoint devices…"	- Administrative scope: Full directory - Where to monitor: Devices - Scope: All users, devices
"We want to detect content that matches U.S. Personally Identifiable Information (PII) definitions…"	- Template selection: U.S. Personally Identifiable Information (PII) Data Enhanced - Conditions for a match: Built-in sensitive information types included in the template
"We want to observe activity without impacting users or blocking business processes…"	- Protection actions: No restrictive actions configured - Mode: Audit-only behavior via simulation mode
"We want to capture and review events when policy conditions are met…"	- Monitoring and reporting: Activity explorer used to review matched events - Alerts/events generated for matching activities
"We want to educate users at the time of activity without enforcing restrictions…"	- User experience: Policy tips enabled (Show policy tips while in simulation mode) - No blocking or override prompts configured
"We want to safely test the policy before enforcing it in production…"	- Policy mode: Run in simulation mode - Deployment approach: Validate behavior before switching to enforcement

Steps to create policy

1. Sign in to the Microsoft Purview portal. > **Data loss prevention** > **Policies**.
2. Choose + **Create policy**.
3. Select **Data stored in connected sources**.
4. For this scenario, under **Categories**, choose **Privacy**.
5. Under **Regulations** choose **U.S. Personally Identifiable Information (PII) Data Enhanced**, and then choose **Next**.
6. Give your new policy a **Name** and **Description**, and then choose **Next**.
7. Accept the default **Full directory** under **Admin units**. Select **Next**.
8. Select the **Devices** location and deselect all other locations and then choose **Next**.
9. On the **Define policy settings** page, choose **Next**.
10. On the **Info to protect page**, choose **Next**.
11. On the **Protection actions** page, choose **Next** again.
12. On the **Customize access and override settings** page, choose **Audit or restrict activities on Devices** and then choose **Next**.
13. On the **Policy mode** page, accept the default **Run the policy in simulation mode** and select **Show policy tips while in simulation mode**.
14. Choose **Next**.
15. Choose **Submit**.
16. Once the new policy is created, choose **Done**.

Now it's time to test your policy:

1. On the **Data Loss Prevention** home page, open **Activity explorer**.
2. Try to share a test item from your Endpoint DLP device that contains content matching the sensitive information types in the **U.S. Personally Identifiable Information (PII) Data Enhanced** template. Sharing the test item should trigger the policy.

3. Back in **Activity explorer**, check the list of activities for the policy-match event generated by the test share action.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Create policy to manage printer access using authorization groups

This scenario demonstrates how to use **authorization groups** in Microsoft Purview DLP to enforce a default block on sensitive data actions while allowing controlled exceptions. In this example, printing of documents classified as legal content is restricted across all devices, except for a defined set of approved **legal department printers**.

Combining **device-level print restrictions** with a printer allowlist enables organizations to protect sensitive information while still supporting legitimate business workflows. This scenario also illustrates how authorization groups can be reused for other scenarios, such as removable storage devices or network shares.

This scenario is for an unrestricted admin creating a full directory policy.

The procedures in this article require that you already have devices onboarded and reporting into Activity explorer. If you haven't onboarded devices yet, see Get started with Endpoint data loss prevention.

Authorization groups are mostly used as allowlists. You assigned policy actions to the group that are different than the global policy actions. In this scenario, we go through defining a printer group and then configuring a policy with block actions for all print activities except for the printers in the group. The procedures for defining a printer group and configuring policy exceptions are essentially the same for **Removeable storage device groups**, and **Network share groups**.

In this scenario, we define a group of printers that the legal department uses for printing contracts. Printing contracts to any other printers is blocked.

Prerequisites and assumptions

This article uses the DLP policy design process described in Design a data loss prevention policy to show you how to create a Microsoft Purview Data Loss Prevention (DLP) policy. Work through these scenarios in your test environment to familiarize yourself with the policy creation UI.

> **Important**

This article presents a hypothetical scenario with hypothetical values. It's only for illustrative purposes. Substitute your own sensitive information types, sensitivity labels, distribution groups, and users.

How you deploy a policy is as important policy design. This article shows you how to use the deployment options so that the policy achieves your intent while avoiding costly business disruptions.

This scenario uses the *Confidential* sensitivity label, so it requires you to create and publish sensitivity labels. To learn more, see:

- Learn about sensitivity labels
- Get started with sensitivity labels
- Create and configure sensitivity labels and their policies

This procedure uses a hypothetical distribution group *Human Resources* and a distribution group for the security team at Contoso.com.

The policy-creation procedure in this article uses alerts, see: Get started with the data loss prevention alerts

Policy intent statement and mapping

*We, Contoso, want to prevent users from printing sensitive legal content to unauthorized printers while still allowing approved business workflows for the legal department. In this scenario, documents that match the **Legal Affairs** trainable classifier should be protected from being printed broadly across the organization, but users in Legal must be able to print those documents to a defined set of approved printers.*

*To achieve this, we will create a policy that applies print restrictions on devices for content classified as legal in nature. Printing will be blocked by default, but an **authorization group** of approved printers will be configured as an exception and allowed. This enables a strong default-protect posture while supporting legitimate business needs through a controlled allowlist model.*

Expand table

Statement	Configuration question answered and configuration mapping
"We want to protect sensitive legal documents from being printed to	- Administrative scope: Full directory - Where to monitor: Devices only

Statement	Configuration question answered and configuration mapping
unauthorized printers…"	- Policy scope: All users/devices covered by the policy
"We want to identify documents that contain legal-sensitive content…"	- Condition: **Content contains** = Trainable classifiers, **Legal Affairs**
"We want to restrict printing of that sensitive content across endpoint devices…"	- Action: **Audit or restrict activities on devices** - Activity type: **File activities on all apps** - Restriction model: **Apply restrictions to specific activity**
"We want printing to be blocked by default unless explicitly allowed…"	- Activity restriction: **Print** = **Block**
"We want approved legal department printers to be exempt from the default block…"	- Endpoint settings: Create a **Printer group** named **Legal printers** - Group members can be defined by friendly printer name, USB product/vendor ID, IP range, print-to-file, Universal Print, corporate printer, or print to local
"We want different policy behavior for approved printers than the global policy action…"	- Authorization group behavior: **Choose different print restrictions** - Printer group restriction: Add **Legal printers** - Group-specific action: **Allow**
"We want approved printing activity to remain visible for audit purposes without creating unnecessary alerts…"	- Allow action behavior: Allowed print activity is recorded in the audit log - No alert or user notification is generated for the allowlisted printer action
"We want to test the policy safely before enforcement…"	- Policy mode: **Run the policy in simulation mode** - User experience: **Show policy tips while in simulation mode**
"We want to use this same design pattern for other authorized destinations in the future…"	- Reusable authorization group model: Same approach applies to **Removable storage device groups** and **Network share groups**

Steps to create policy

Create and use printer groups

Perform the following steps to create a printer group for approved legal printers:

1. Sign in to the Microsoft Purview portal > **Data loss prevention** > **Settings** (gear in the upper left hand corner) > **Data Loss Prevention** > **Endpoint settings** > **Printer groups**.

2. Select **Create printer group** and enter a **Group a name**. In this scenario, we use `Legal printers`.

3. Select **Add printer** and provide a name. You can define printers by:
   a. Friendly printer name
   b. USB product ID
   c. USB vendor ID
   d. IP range
   e. Print to file
   f. Universal print deployed on a printer
   g. Corporate printer
   h. Print to local

4. Select **Close**.

Configure policy printing actions

Perform the following steps to configure the DLP policy's printing actions and assign the printer group exception:

1. Sign in to the Microsoft Purview portal.

2. Navigate to **Data loss prevention** > **Policies**.

3. Select **Create policy**.

4. **Data stored in connected sources**.

5. Select the **Custom** from the **Categories** then **Custom policy** template from **Regulations**.

6. Give your new policy a **Name** and **Description**.

7. Accept the default **Full directory** under **Admin units**.

8. Scope the location to only the **Devices** location.

9. Create a rule with the following values:
   a. Add a **Condition**: **Content contains** = **Trainable classifiers**, **Legal Affairs**
   b. **Actions** = **Audit or restrict activities on devices**
   c. Then pick **File activities on all apps**
   d. The select **Apply restrictions to specific activity**
   e. Select **Print** = **Block**

10. Select **Choose different print restrictions**
11. Under **Printer group restrictions**, select **Add group** and select **Legal printers**.
12. Set **Action** = **Allow**.

    > **Tip**
    >
    > The **Allow** action wil record and audit event to the audit log, but not generate an alert or notification.

13. Select **Save** and then **Next**.
14. Accept the default **Run the policy in simulation mode** value and choose **Show policy tips while in simulaiton mode**. Choose **Next**.
15. Review your settings and choose **Submit**.
16. The new DLP policy appears in the **Data loss prevention** > **Policies** list.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Create policy to detect and alert on U.S. PII data exposure

This scenario shows how to enhance an existing Microsoft Purview DLP policy by enabling **real-time alerting** for policy matches involving U.S. PII data. Alerts are configured to notify administrators each time a match occurs, improving visibility and response readiness while the policy continues to run without restricting user activity.

This scenario is for an unrestricted admin modifying a full directory scoped policy.

Prerequisites and assumptions

This article applies the policy design process from Design a data loss prevention policy to show you how to modify an existing Microsoft Purview Data Loss Prevention (DLP) policy. Work through these scenarios in your test environment to familiarize yourself with the policy creation UI.

> **Important**
>
> This article presents a hypothetical scenario with hypothetical values. It's only for illustrative purposes. Substitute your own sensitive information types, sensitivity labels, distribution groups, and users.

How you deploy a policy is as important policy design. This article shows you how to use the deployment options so that the policy achieves your intent while avoiding costly business disruptions.

This scenario uses the *Confidential* sensitivity label, so it requires you to create and publish sensitivity labels. To learn more, see:

- Learn about sensitivity labels
- Get started with sensitivity labels
- Create and configure sensitivity labels and their policies

This procedure uses a hypothetical distribution group *Human Resources* and a distribution group for the security team at Contoso.com.

This procedure uses alerts, see: Get started with the data loss prevention alerts

Policy intent statement and mapping

We, Contoso, have already deployed a policy to monitor U.S. Personally Identifiable Information (PII) activity on endpoint devices in audit mode. After reviewing initial results, we now want to enhance our visibility by ensuring that administrators are proactively notified when policy matches occur.

To achieve this, we will modify the existing policy to generate alerts whenever low-volume PII-related activity is detected. These alerts will be sent in real time to administrators so they can quickly review and respond to potential risks. The policy will continue to operate without restricting user activity, maintaining the audit-first approach while improving incident awareness and response readiness.

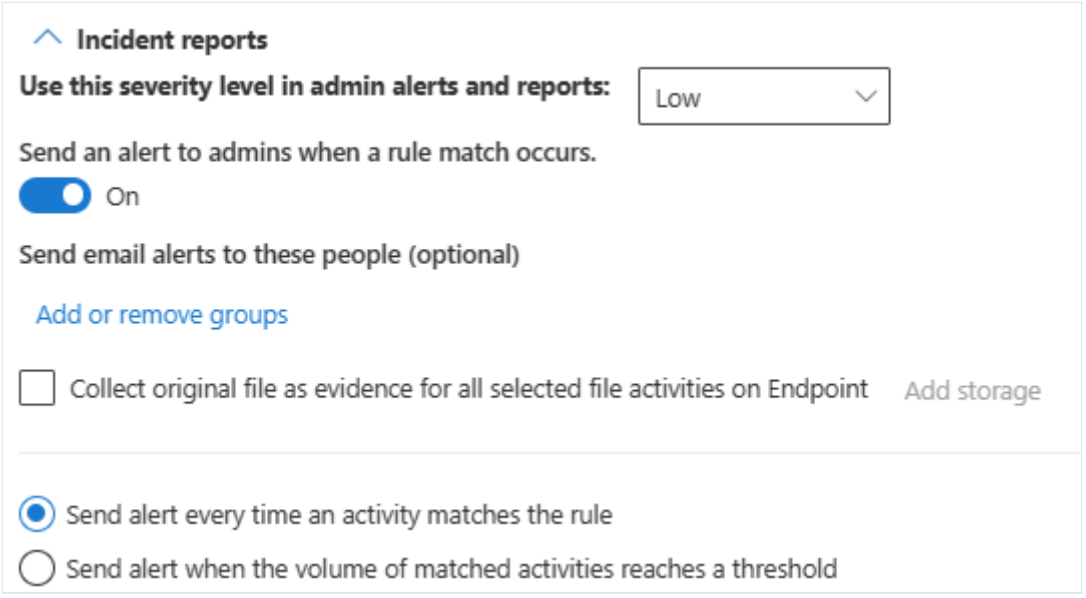
Expand table

Statement	Configuration question answered and configuration mapping
"We want to enhance our existing monitoring policy without changing its scope…"	- Administrative scope: Full directory (unchanged) - Where to monitor: Devices (unchanged) - Existing policy reused and edited
"We want to generate alerts when U.S. PII policy conditions are matched…"	- Rule edited: Low volume of content detected scenarios U.S. PII Data Enhanced - Condition: Existing PII detection logic from template retained
"We want administrators to be notified automatically when matches occur…"	- Incident reports: Enable 'Send an alert to admins when a rule match occurs' - Recipients: Default admin and additional configured recipients
"We want immediate visibility into every matching event…"	- Alert frequency: Send alert every time an activity matches the rule
"We want to improve response readiness without disrupting users…"	- Protection actions: No blocking or restriction changes - Policy continues in audit/simulation behavior
"We want to validate alerting behavior alongside activity tracking…"	- Monitoring: Activity explorer used to correlate events and alerts - Testing: Trigger policy using test file with matching PII content

Steps to create policy

Perform the following steps to edit the existing policy and enable alerting:

1. Sign in to the Microsoft Purview portal > **Data loss prevention** > **Policies**.
2. Choose the **U.S. Personally Identifiable Information (PII) Data Enhanced** policy that you created in Create an endpoint DLP policy to audit activity on devices. This policy must already exist in your environment before you proceed.
3. Choose **Edit policy**.
4. Go to the **Customize advanced DLP rules** page and edit the **Low volume of content detected scenarios U.S. PII Data Enhanced**.
5. Scroll down to the **Incident reports** section and toggle **Send an alert to admins when a rule match occurs** to **On**. Email alerts are automatically sent to the administrator and anyone else you add to the list of recipients.



6. For the purposes of this scenario, choose **Send alert every time an activity matches the rule**.
7. Choose **Save**.
8. Leave all other policy and rule settings unchanged. Choose **Next** through the remaining wizard pages, and then select **Submit** to save the policy changes.
9. Try to share a test item containing content that will trigger the U.S. Personally Identifiable Information (PII) Data Enhanced condition. Sharing the test item should trigger a policy

match and generate an alert.

10. Check the activity explorer for the event.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Help prevent unauthorized sensitive data sharing with block actions and allow overrides

This scenario demonstrates how to transition an existing Microsoft Purview DLP policy from monitoring to enforcement by blocking the sharing of sensitive U.S. PII data. The policy is updated to use **block with override**, preventing high-risk actions while still allowing users to proceed when justified.

This approach strengthens data protection by reducing the risk of data exfiltration, while maintaining business continuity through controlled user overrides and continued visibility into user activity via Activity explorer.

This scenario is for an unrestricted admin modifying a full directory policy.

Prerequisites and assumptions

This article uses the process you learned in Design a data loss prevention policy to show you how to create a Microsoft Purview Data Loss Prevention (DLP) policy. Work through these scenarios in your test environment to familiarize yourself with the policy creation UI.

> ⓘ **Important**
>
> This article presents a hypothetical scenario with hypothetical values. It's only for illustrative purposes. Substitute your own sensitive information types, sensitivity labels, distribution groups, and users.

How you deploy a policy is as important policy design. This article shows you how to use the deployment options so that the policy achieves your intent while avoiding costly business disruptions.

This scenario uses the *Confidential* sensitivity label, so it requires you to create and publish sensitivity labels. To learn more, see:

- Learn about sensitivity labels
- Get started with sensitivity labels

- Create and configure sensitivity labels and their policies

This procedure uses a hypothetical distribution group *Human Resources* and a distribution group for the security team at Contoso.com.

This procedure generates data loss prevention alerts when policy rules are triggered. To learn how to configure and review these alerts, see Get started with the data loss prevention alerts.

Policy intent statement and mapping

We, Contoso, have already deployed and tuned a policy that detects U.S. Personally Identifiable Information (PII) on endpoint devices and provides visibility and alerting. After validating the policy behavior, we now want to move to enforcement by preventing users from sharing sensitive information outside the organization.

To achieve this, we will modify the existing policy so that when PII is detected in user actions involving service domains or browsers, the action is blocked. However, to maintain business flexibility, users will be allowed to override the block with justification.

This ensures strong data protection while still enabling users to proceed in legitimate business scenarios when necessary.

Expand table

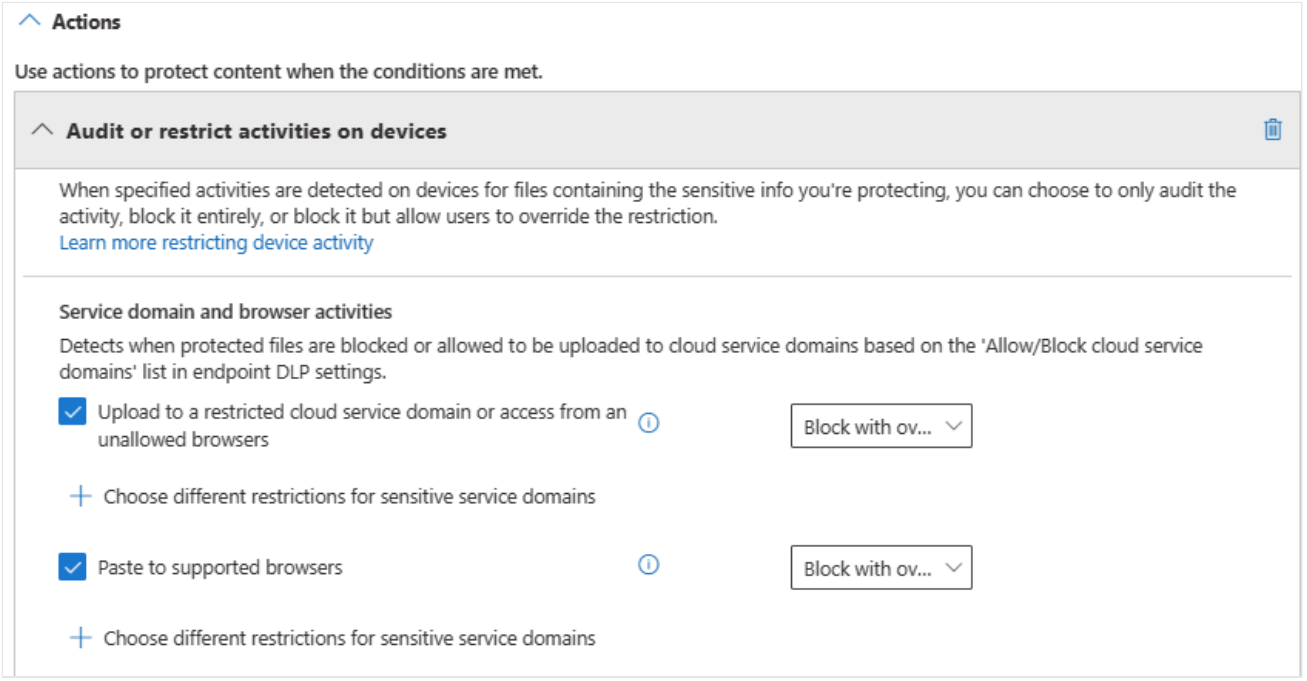
Statement	Configuration question answered and configuration mapping
"We want to transition from monitoring to enforcing protection of U.S. PII on endpoint devices..."	- Administrative scope: Full directory (unchanged) - Where to monitor: Devices (unchanged) - Existing policy reused and updated for enforcement
"We want to prevent users from sharing sensitive data outside the organization..."	- Actions: Audit or restrict activities on Windows devices - Activity types: Service domain and browser activities
"We want to block high-risk actions involving sensitive data..."	- Action configuration: Block with override enabled for service domain and browser activities
"We want to allow users to proceed in legitimate cases with accountability..."	- Override behavior: Allow override with user justification captured - User interaction: Prompt displayed on endpoint when action is blocked
"We want consistent enforcement across different sensitivity levels..."	- Rule coverage: Apply same block with override settings to both low-volume and high-volume

Statement	Configuration question answered and configuration mapping
	detection rules
"We want to maintain visibility into enforcement events and user behavior..."	- Monitoring: Activity explorer logs blocked and overridden events - Alerts/events: Captured for investigation and auditing
"We want to validate enforcement behavior through testing..."	- Testing: Trigger policy using test file containing PII data and attempt external sharing - Expected result: User receives block prompt with override option

Steps to create policy

Perform the following steps to edit the existing policy and enable block with override behavior.

1. Sign in to the Microsoft Purview portal > **Data loss prevention** > **Policies**.
2. Choose your existing **U.S. Personally Identifiable Information (PII) Data** policy that you previously created for monitoring on endpoint devices.
3. Choose **Edit policy**.
4. Go to the **Customize advanced DLP rules** page and edit the **Low volume of content detected scenarios U.S. PII Data Enhanced**.
5. Scroll down to the **Actions** > **Audit or restrict activities on Windows device** section and set both options under the **Service domain and browser activities** to **Block with override**.



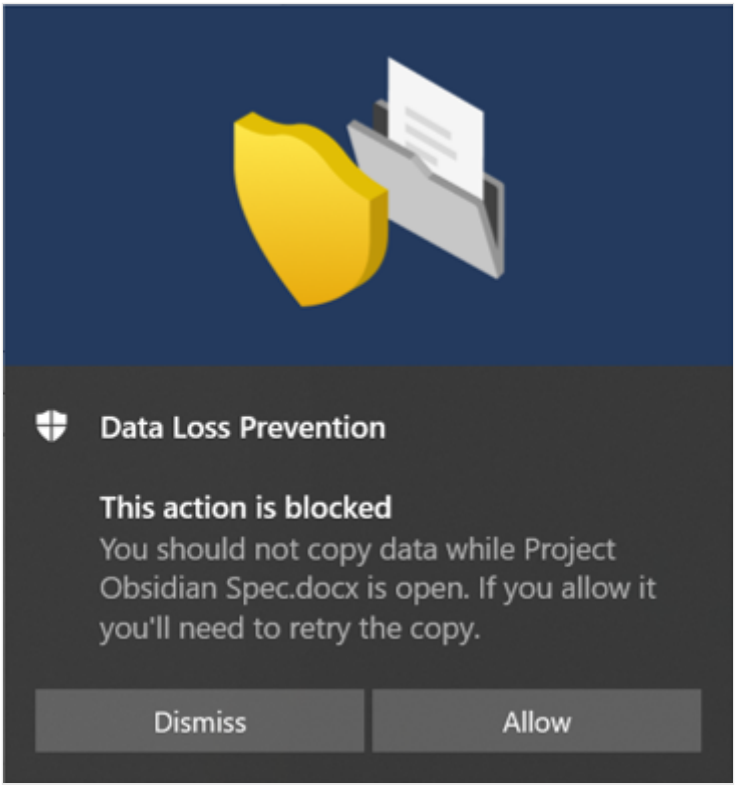
6. Choose **Save**.

7. Edit the **High volume of content detected scenarios U.S. PII Data Enhanced** rule. Scroll down to **Actions** > **Audit or restrict activities on Windows device**, set both options under **Service domain and browser activities** to **Block with override**, and then choose **Save**.

8. Leave all other existing policy settings unchanged by choosing **Next** through the rest of the wizard, and then select **Submit** to apply the policy changes.

9. Attempt to share a test item that contains content that will trigger the U.S. Personally Identifiable Information (PII) Data condition with someone outside your organization. This sharing attempt should trigger the policy.

   You see a popup like this on the client device:



10. Check Activity explorer for the blocked or overridden sharing event generated by the test action.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Help prevent risky user activity by monitoring or restricting access to sensitive service domains

Overview

This scenario demonstrates how to monitor and control user interactions with sensitive websites using Microsoft Purview DLP. By defining **sensitive service domains**, organizations can audit or restrict actions such as copying, printing, or saving web content when users access these sites.

Policy enforcement occurs in a supported, DLP-aware browser environment, ensuring consistent control and user guidance. This enables organizations to protect sensitive data accessed through web applications while maintaining flexibility to tailor controls based on business needs.

> ⓘ **Note**
>
> The following web browsers are supported:
>
> - Microsoft Edge (Win/macOS)
> - Chrome (Win/macOS)- Microsoft Purview extension for Chrome Windows only
> - Firefox (Win/macOS)- Microsoft Purview extension for Firefox Windows only
> - Safari (macOS only)

Prerequisites and assumptions

This article uses the process you learned in Design a data loss prevention policy to show you how to create a Microsoft Purview Data Loss Prevention (DLP) policy. Work through these scenarios in your test environment to familiarize yourself with the policy creation UI.

> ⓘ **Important**
>
> This article presents a hypothetical scenario with hypothetical values. It's only for illustrative purposes. Substitute your own sensitive information types, sensitivity labels, distribution groups, and users.

How you deploy a policy is as important policy design. This article shows you how to use the deployment options so that the policy achieves your intent while avoiding costly business disruptions.

This scenario uses the *Confidential* sensitivity label, so it requires you to create and publish sensitivity labels. To learn more, see:

- Learn about sensitivity labels
- Get started with sensitivity labels
- Create and configure sensitivity labels and their policies

This procedure uses a hypothetical distribution group *Human Resources* and a distribution group for the security team at Contoso.com.

The policy configuration in this article uses alerts. To learn more, see Get started with the data loss prevention alerts

Policy intent statement and mapping

We, Contoso, want to monitor and control how users interact with sensitive websites to reduce the risk of data exposure through browser-based activities. Specifically, we want to audit or restrict actions such as copying data, printing content, or saving web pages locally when users access designated sensitive service domains.

*To achieve this, we will define a set of **sensitive service domains** and configure a policy that detects when users access these sites through supported browsers. When a match occurs, we will apply activity-specific controls—such as auditing or blocking printing, copying, or file-saving actions—while ensuring users are redirected to supported browsers (such as Microsoft Edge) where these controls can be enforced consistently.*

This approach enables granular control of in-browser data handling for sensitive sites, helping protect organizational data while maintaining flexibility in how policies are applied and tuned over time.

Expand table

Statement	Configuration question answered and configuration mapping
"We want to monitor and control user interactions with designated sensitive websites..."	- Administrative scope: Full directory - Where to monitor: Devices only - Policy scope: All users/devices (or scoped test users)

Statement	Configuration question answered and configuration mapping
"We want to define which websites are considered sensitive for monitoring or restriction..."	- Endpoint settings: Configure **Service domains** and **Sensitive service domain groups** - Domains defined using URL/IP/IP range with wildcard support
"We want to detect when users access these sensitive sites in supported browsers..."	- Condition: **User accessed a sensitive site from Edge** - Browser coverage: Supported browsers with Microsoft Purview extensions where applicable
"We want to ensure policy enforcement occurs in a DLP-aware browser..."	- Browser handling: Redirect unsupported browsers (e.g., Chrome/Firefox) to **Microsoft Edge** when accessing sensitive domains
"We want to audit or restrict specific user activities performed on sensitive websites..."	- Actions: **Audit or restrict activities when users access sensitive sites** - Activity controls include: copy, print, save as local files
"We want fine-grained control over which activities are monitored or blocked..."	- Action configuration: Select individual user activities (copy, print, save) - Choose audit, block, or block with override depending on requirement
"We want to flexibly apply controls across different sets of sensitive domains..."	- Rule configuration: Assign one or more **Sensitive service domain groups** - Reusable group design for scalability
"We want to finalize and deploy the policy for immediate use..."	- Policy mode: Submit and enable policy - Deployment: Active policy with configured monitoring/restriction behavior

Steps to create the sensitive service domains policy

Perform the following steps to configure sensitive service domains and create the DLP policy.

Configure Sensitive service domains

Use the following steps to define sensitive service domains and create a DLP policy that monitors or restricts user activity on those domains.

1. Sign in to the Microsoft Purview portal > **Data loss prevention** > **Settings** (gear icon in the upper left hand corner) > **Data Loss Prevention** > **Endpoint settings** > **Browser and domain restrictions to sensitive data** > **Service domains**.

2. Set **Service domains** to **Block**.

3. To control whether sensitive files can be uploaded to specific domains, select **Add cloud service domain**.
4. Enter the domain that you want to audit or block and choose the + button. Repeat for any additional domains. Choose **Save**.
5. Under **Sensitive service domain groups**, choose **Create sensitive service domain group**.
6. Give the group a name, select the **Match type** you want (you can select from **URL**, **IP address**, **IP address range**), and enter the URL, IP address, or IP address range to be audited or blocked. When matching a URL, you can add multiple websites to a group and use wildcards to cover subdomains. For example, `www.contoso.com` for just the top level website or *.contoso.com for corp.contoso.com, hr.contoso.com, fin.contoso.com.
7. Select **Save**.
8. In the left navigation pane, select **Data loss prevention** > **Policies**.
9. On the policy template page, select **Data stored in connected sources**.
10. Create and scope a policy that is applied only to the **Devices** location. For detailed steps on scoping a policy to the Devices location, see Create and Deploy data loss prevention policies. Be sure to scope the **Admin units** to **Full directory**.
11. Create a rule that uses the condition **the user accessed a sensitive site from Edge**, and the action **Audit or restrict activities when users access sensitive sites in Microsoft Edge browser on Windows devices**.
12. In the action, under **Sensitive Site Restrictions**, select **Add or remove Sensitive site groups**.
13. Create and/or select the **Sensitive site groups** you want. Any website under the group(s) you select here will be redirected to Microsoft Edge when opened in Chrome or Firefox (so long as the Microsoft Purview extension is installed).
14. Select **Add**.
15. Select the user activities you want to monitor or restrict and the actions you want Microsoft Purview to take in response to those activities.
16. Finish configuring the rule and policy and choose **Submit** and then **Done**.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

Last updated on 06/26/2026

Help prevent leakage of sensitive content by restricting paste actions into browsers

This scenario demonstrates how to restrict users from pasting sensitive content into browser-based applications using Microsoft Purview DLP. By evaluating content at the moment it is pasted, organizations can detect and control sensitive information in real time, regardless of where the content originated.

Using **Paste to browser** controls with **Sensitive service domain groups**, this approach enables flexible enforcement—such as auditing, warning, or blocking paste actions—based on the destination website. This approach helps reduce the risk of accidental data exposure while allowing policies to be tailored for different levels of risk.

The **Paste to browser** activity operates independently of the source file's classification and requires rules configured with **Sensitive service domain groups**. It is not supported with the **Service domains** endpoint DLP setting. For more information, see Endpoint activities you can monitor and take action on

> ⓘ **Note**
>
> The following web browsers are supported:
>
> - Microsoft Edge (Win/macOS)
> - Chrome (Win/macOS)- Microsoft Purview extension for Chrome Windows only
> - Firefox (Win/macOS)- Microsoft Purview extension for Firefox Windows only
> - Safari (macOS only)

> ⓘ **Important**
>
> - If you have configured evidence collection for file activities on devices and your Antimalware Client Version on the device is older than 4.18.23110, when you implement this scenario, *Restrict pasting sensitive content into a browser*, you will see random characters when you attempt to view the source file in Alert details. To see the actual source file text, you should download the file.

- If **Document could not be scanned** is selected as the action for the matched rule, the evidence file will not be captured.

Prerequisites and assumptions

This article uses the process you learned in Design a data loss prevention policy to show you how to create a Microsoft Purview Data Loss Prevention (DLP) policy. Work through these scenarios in your test environment to familiarize yourself with the policy creation UI.

> **Important**
>
> This article presents a hypothetical scenario with hypothetical values. It's only for illustrative purposes. Substitute your own sensitive information types, sensitivity labels, distribution groups, and users.

How you deploy a policy is as important policy design. This article shows you how to use the deployment options so that the policy achieves your intent while avoiding costly business disruptions.

This scenario uses the *Confidential* sensitivity label, so it requires you to create and publish sensitivity labels. To learn more, see:

- Learn about sensitivity labels
- Get started with sensitivity labels
- Create and configure sensitivity labels and their policies

This procedure uses a hypothetical distribution group *Human Resources* and a distribution group for the security team at Contoso.com.

The policy configuration procedure uses alerts, see: Get started with the data loss prevention alerts

Policy intent statement and mapping

We, Contoso, want to prevent users from unintentionally exposing sensitive information by pasting it into web forms or browser-based applications. Sensitive data—such as personally identifiable information or other regulated content—can be copied from local files or applications and pasted into websites, creating a potential data exfiltration risk.

*To address this, we will create a policy that evaluates content dynamically at the moment it is pasted into a supported browser. Based on the sensitivity of the pasted content and the destination website, we will audit, warn, or block the action. We will also use **Sensitive service domain groups** to apply different levels of enforcement depending on the target websites. This allows us to balance security and usability by applying stricter controls to higher-risk domains while allowing flexibility for trusted sites.*

Expand table

Statement	Configuration question answered and configuration mapping
"We want to prevent sensitive data from being pasted into web forms or browser fields..."	- Administrative scope: Full directory - Where to monitor: Devices only - Policy scope: All users/devices or targeted users
"We want to evaluate content at the time it is pasted, regardless of its source..."	- Condition: Content contains selected **Sensitive info types** - Evaluation: Real-time classification at paste event (independent of source file classification)
"We want to control paste actions based on the sensitivity of the content..."	- Action: **Audit or restrict activities on devices** - Activity type: **Paste to supported browsers**
"We want different levels of enforcement depending on the destination website..."	- Endpoint settings: Create **Sensitive service domain groups** - Rule design: Associate paste restrictions with specific domain groups
"We want to audit, warn, or block paste actions based on risk level..."	- Action configuration: Set to **Audit**, **Block with override**, or **Block** depending on enforcement needs
"We want flexibility to tailor restrictions across different categories of websites..."	- Policy design: Multiple URL/domain groups (e.g., trusted vs untrusted sites) - Use exceptions or multiple rules for granular control
"We want to ensure enforcement occurs consistently across supported browsers..."	- Browser support: Microsoft Edge, Chrome, Firefox (with extensions), Safari - Endpoint DLP integration ensures policy enforcement in supported browsers
"We want users to be informed when paste actions are evaluated or restricted..."	- User experience: Policy tips or notifications triggered during paste evaluation - Behavior note: Possible brief delay while classification completes
"We want to deploy and test the policy with appropriate enforcement level..."	- Policy mode: Configurable (test, audit, or enforce) - Deployment: Submit policy and validate behavior with

Statement	Configuration question answered and configuration mapping
	test scenarios

Steps to create policy

You can set up different levels of enforcement when it comes to blocking data from being pasted into a browser. To do this, create different URL groups. For instance, you can create a policy that warns users against posting U.S. Social Security Numbers (SSN) to any website, and that triggers an audit action for websites in Group A. You can create another policy that completely blocks the paste action--without giving a warning--for all of the websites in Group B.

Create a URL group

Perform the following steps to create a sensitive service domain group for browser paste restrictions.

1. Sign in to the Microsoft Purview portal > **Data loss prevention** > **Settings** (gear icon in the upper left hand corner) > **Data Loss Prevention** > **Endpoint settings**, and scroll down to **Browser and domain restrictions to sensitive data**. Expand the section.

2. Scroll down to **Sensitive service domain groups**.

3. Choose **Create sensitive service domain group**.
   a. Enter a **Group name**.
   b. In the **Sensitive service domain** field, enter the URL for the first website you want to monitor and then choose **Add site**.
   c. Continue adding URLs for the rest of the websites you want to monitor in this group.
   d. When you are finished adding all URLs to your group, choose **Save**.

4. Create as many separate groups of URLs as you need.

Restrict pasting content into a browser

Use the following steps to create a DLP policy that restricts pasting sensitive content into a browser.

1. Sign in to the Microsoft Purview portal > **Data loss prevention** > **Policies**.

2. Under **Data stored in connected sources**, select **Devices**.

3. Create a DLP policy scoped to **Devices**. For information on how to create a DLP policy, see Create and Deploy data loss prevention policies.
4. On the Define policy settings page in the DLP policy creation flow, select **Create or customize advanced DLP rules** and then choose **Next**.
5. On the **Customize advanced DLP rules** page, choose **Create rule**.
6. Enter a name and description for the rule.
7. Expand **Conditions**, choose **Add condition**, and then select the **Sensitive info types**.
8. Under **Content Contains**, scroll down and select the new sensitive information type that you previously chose or created.
9. Scroll down to the **Actions** section, and choose **Add an action**.
10. Choose **Audit or restrict activities on devices**
11. In the **Actions** section, under **Service domain and browser activities**, select **Paste to supported browsers**.
12. Set the restriction to **Audit**, **Block with override**, or **Block**, and then choose **Add**.
13. Choose **Save**.
14. Choose **Next**
15. Choose whether you want to test your policy, turn it on right away, or keep it off, and then choose **Next**.
16. Choose **Submit**.

> **Important**
>
> There may be a brief time lag between when the user attempts to paste text into a web page and when the system finishes classifying it and responds. If this classification latency happens, you may see both policy-evaluation and check-complete notifications in Edge or policy-evaluation toast on Chrome and Firefox. Here are some tips for minimizing the number of notifications:
>
> 1. Notifications are triggered when policy for the target website is configured to **Block** or **Block with override** paste to browser for that user. You can configure setting the

overall action to **Audit** and then list the target websites using the exceptions as **Block**. Alternately, you can set the overall action to **Block** and then list secure websites using the exceptions as **Audit**.

2. Use latest Antimalware client version.
3. Use latest Edge browser version, especially Edge 120.
4. Install these Windows KBs

- Windows 10: KB5032278
- Windows 11: KB5032288

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Help prevent exposure of sensitive files by configuring auto-quarantine for OneDrive sync

This scenario shows how to configure Endpoint DLP to prevent synchronization of **Highly Confidential** files to cloud sync apps like OneDrive while avoiding repeated notifications. By using **auto-quarantine** and restricted app controls, sensitive files are automatically moved to a secure location and replaced with a placeholder, enabling effective protection without creating a poor user experience from repeated sync attempts.

This scenario is for an unrestricted admin creating a full directory policy.

Prerequisites and assumptions

This article uses the process you learned in Design a data loss prevention policy to show you how to create a Microsoft Purview Data Loss Prevention (DLP) policy. Work through these scenarios in your test environment to familiarize yourself with the policy creation UI.

> **Important**
>
> This article presents a hypothetical scenario with hypothetical values. It's only for illustrative purposes. Substitute your own sensitive information types, sensitivity labels, distribution groups, and users.

How you deploy a policy is as important policy design. This article shows you how to use the deployment options so that the policy achieves your intent while avoiding costly business disruptions.

This scenario uses the *Confidential* sensitivity label, so it requires you to create and publish sensitivity labels. To learn more, see:

- Learn about sensitivity labels
- Get started with sensitivity labels
- Create and configure sensitivity labels and their policies

This procedure uses a hypothetical distribution group *Human Resources* and a distribution group for the security team at Contoso.com.

This scenario uses alerts. For more information, see Get started with the data loss prevention alerts

Policy intent statement and mapping

We, Contoso, want to prevent users from synchronizing highly sensitive files to cloud synchronization applications such as OneDrive from endpoint devices. In this scenario, files labeled ***Highly Confidential*** *must not be synchronized to cloud storage through restricted sync apps.*

At the same time, we want to avoid repeated DLP notifications and poor user experience caused by failed sync retries. To do that, we will use ***Endpoint DLP auto-quarantine*** *so that when a restricted sync app attempts to process a protected file, the original file is automatically moved to a quarantine location and replaced with a placeholder text file. This both prevents exfiltration and stops the sync client from repeatedly retrying the same blocked file.*

We also want users to be informed when the action is blocked, and we want administrators to be able to validate and investigate the resulting events through ***Activity explorer****.*

Expand table

Statement	Configuration question answered and configuration mapping
"We want to prevent users from synchronizing files labeled Highly Confidential to cloud sync apps such as OneDrive..."	- Administrative scope: Full directory - Where to monitor: Devices - Policy scope: Devices only, with optional targeting to specific test users - Condition: Content contains the sensitivity label **Highly Confidential**
"We want to treat cloud synchronization apps as restricted applications for these sensitive files..."	- Endpoint settings: Create a restricted app group named **Cloud Sync apps** - Restricted apps include OneDrive executables for Windows and macOS - Restricted app group is referenced by the DLP rule
"We want to block restricted apps from performing file activities on these sensitive items..."	- Rule action: **Audit or restrict activities on devices** - Activity type: **File activities for apps in restricted app groups** - Restriction: **Apply a restriction to all activity** > **Block**

Statement	Configuration question answered and configuration mapping
“We want to avoid looping DLP notifications and repeated cloud sync retries for blocked sensitive files…”	- Endpoint DLP setting: Enable **Auto-quarantine for unallowed apps** - Quarantine behavior: Move the original file to a configured quarantine folder - File handling: Append date/time stamp to quarantined file name
“We want the user to understand what happened to the blocked file…”	- Auto-quarantine response: Replace the original file with a **.txt placeholder file** containing a custom message - Placeholder message includes tokens such as **%%FileName%%**, **%%PolicyName%%**, and **%%QuarantinePath%%**
“We want users to be notified in real time when the sync action is blocked…”	- User notifications: Turn **User notifications** On - Endpoint behavior: Show users a **policy tip notification** when the blocked activity occurs
“We want this protection to apply immediately once the rule is active…”	- Policy creation: Custom policy with advanced DLP rules - Deployment mode: **Turn it on right away**
“We want to validate that blocking, quarantine, and user messaging all work together correctly…”	- Test procedure: Create a Word document, apply the **Highly Confidential** label, and copy it into the local OneDrive sync folder - Expected outcome: User sees a notification toast, the source file is replaced by a placeholder text file, and the original file appears in the quarantine folder
“We want administrators to be able to review the enforcement event after testing…”	- Monitoring and investigation: Use **Activity explorer** - Filters: Set location to **Devices** and filter by the policy name to review the policy match and action outcome

Steps to create policy

Complete the following procedures to configure Endpoint DLP settings, create the policy, and test auto-quarantine.

Before you begin

In this Endpoint DLP auto-quarantine scenario, synchronizing files with the **Highly Confidential** sensitivity label to OneDrive is blocked. This is a complex scenario with multiple components and procedures. You need:

- A Microsoft Entra user account to target and an onboarded Windows 10/11 or macOS computer that is already synchronizing a local OneDrive folder with OneDrive cloud storage.

- Sensitivity labels configured and published—see Get started with sensitivity labels and Create and configure sensitivity labels and their policies.

There are three procedures.

1. Configure the Endpoint DLP Autoquarantine settings.
2. Create a policy that blocks sensitive items that have the **Highly Confidential** sensitivity label.
3. Create a Word document on the Windows 10/11 or macOS device that the policy targets, apply the label, and copy the document to the user account's local OneDrive folder that is being synchronized.

Configure Endpoint DLP unallowed app and Autoquarantine settings

Perform the following steps to create a restricted app group and configure auto-quarantine settings for cloud sync apps.

1. Sign in to the Microsoft Purview portal > **Data loss prevention** > **Settings** (gear in the upper left hand corner) > **Data loss prevention** > **Endpoint settings**.

2. Expand **Restricted apps and app groups**.

3. Under **Restricted app groups**, choose **Add restricted app group**. Enter *Cloud Sync apps* as the group name.

4. Select the **Auto-quarantine** box.

5. For the **App name**, enter *OneDrive*. For the Windows **Executable name**, enter *onedrive.exe*, then choose the + button. For the macOS **Executable name**, enter */Applications/OneDrive.app/Contents/MacOS/OneDrive*, then choose the + button. This disallows onedrive application on devices from accessing items with the **Highly Confidential** label.

   > **Note**
   >
   > To restrict other common cloud sync app on macOS devices, add the below path
   >
   > - Box: `/Applications/Box Sync.app/Contents/MacOS/Box Sync`
   > - Dropbox: `/Applications/Dropbox.app/Contents/MacOS/Dropbox`
   > - Google Drive: `/Applications/Google Drive.app/Contents/MacOS/Google Drive`

- iCloud: `System/Library/PrivateFrameworks/CloudDocsDaemon.framework/Versions/A/Support/bird`

6. Choose **Save**.
7. Under **Auto-quarantine settings** choose **Edit auto-quarantine settings**.
8. Enable **Auto-quarantine for unallowed apps**.
9. Enter the path to the folder on local machines where you want the original sensitive files to be moved to. For example:

> `%homedrive%%homepath%\Microsoft DLP\Quarantine` for the username *Isaiah Langer* will place the moved items in a folder named:
>
> *C:\Users\IsaiahLanger\Microsoft DLP\Quarantine\OneDrive*
>
> A separate path for the macOS device is required, such as `/System/Applications/Microsoft DLP/Quarantine`

1. Append a date and time stamp to the original file name.

> **ⓘ Note**
>
> DLP auto-quarantine will create sub-folders for the files for each unallowed app. So if you have both *Notepad* and *OneDrive* in your unallowed apps list, a sub-folder will be created for **\OneDrive** and another sub-folder for **\Notepad**.

2. Choose **Replace the files with a .txt file that contains the following text** and enter the text you want in the placeholder file. For example for a file named *auto quar 1.docx*, you could enter:

> %%FileName%% contains sensitive info that your organization is protecting with the data loss prevention (DLP) policy %%PolicyName%%. The file was moved to the quarantine folder: %%QuarantinePath%%

will leave a text file that contains this message:

> auto quar 1.docx contains sensitive info that your organization is protecting with the data loss prevention (DLP) policy. It was moved to the quarantine folder: C:\Users\IsaiahLanger\Microsoft DLP\Quarantine\OneDrive\auto quar 1.docx.

3. Choose **Save**.

Configure a policy to block OneDrive synchronization of files with the sensitivity label "Highly Confidential"

Use the following steps to create a DLP policy that blocks OneDrive synchronization for files labeled **Highly Confidential**.

1. Sign in to the Microsoft Purview portal > **Data loss prevention** > **Policies**.
2. Select **Create policy**.
3. **Data stored in connected sources**.
4. For this scenario, choose **Custom**, then **Custom policy**. Choose **Next**.
5. Fill in the **Name** and **Description** fields, choose **Next**.
6. Select **Full directory** under **Admin units**.
7. Toggle the **Status** field to off for all locations except **Devices**. If you have a specific end user account that you want to test this from, be sure to select it in the scope. Choose **Next**.
8. Accept the default **Create or customize advanced DLP rules** selection and choose **Next**.
9. Create a rule with these values:
   a. **Name** > *Scenario 4 Autoquarantine*.
   b. Under **Conditions** choose **Add condition** and then **Content Contains**.
   c. Enter a group name, for example *Highly Confidential Sensitivity Labels* and then choose **Add**.
   d. Select **Sensitivity labels** then **Highly Confidential** and choose **Add**.
   e. Under **Actions** choose **Add an action**.
   f. Select **Audit or restrict activities on devices** > **File activities for apps in restricted app groups**.
   g. Choose **Add restricted app group** then choose the *Cloud Sync Apps* group you created previously.
   h. Choose **Apply a restriction to all activity** > **Block**. For the purposes of this scenario, clear all the other activities.

i. Under **User notifications**, toggle **User notifications** to **On** and under **Endpoint devices** choose **Show users a policy tip notification when an activity** if not already enabled.

10. Choose **Save** and **Next**.
11. Choose **Turn it on right away**. Choose **Next**.
12. Review your settings and choose **Submit**.

> **Note**
>
> Allow at least an hour for the new policy to be replicated and applied to the target Windows 10 computer.

13. The new DLP policy appears in the policy list.

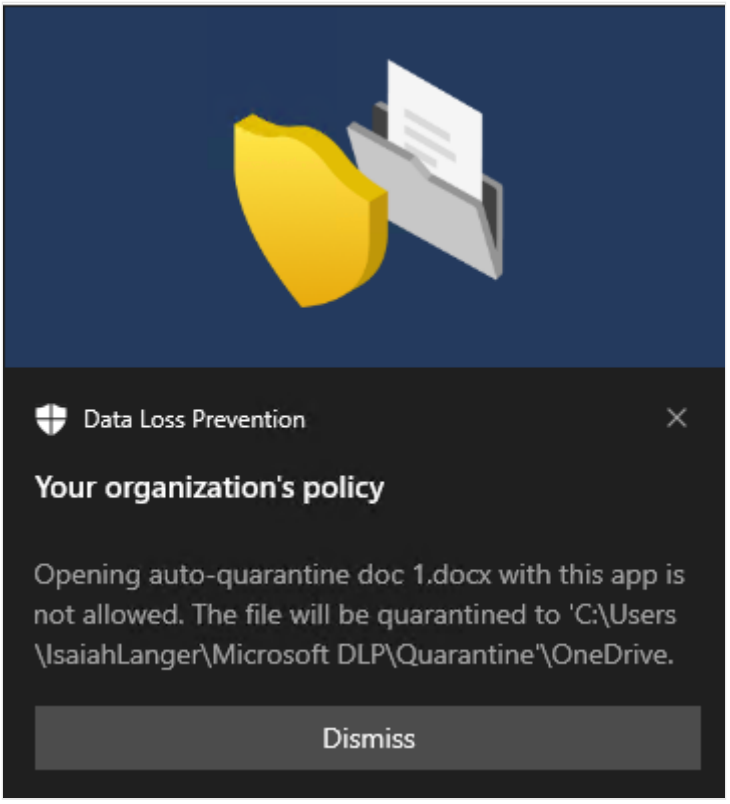
Test Autoquarantine on the Windows 10/11 device

Perform the following steps to verify that auto-quarantine blocks the sync attempt and replaces the file with a placeholder. Before testing, wait at least an hour for the new policy to replicate and apply to the target device.

1. Sign in to the Windows 10/11 computer with the user account you included in the **Devices** scope for the policy.
2. Create a folder whose contents won't be synchronized to OneDrive. For example:

   > *C:\auto-quarantine source folder*

3. Open Microsoft Word and create a file in the autoquarantine source folder. Apply the **Highly confidential** sensitivity label; see Apply sensitivity labels to your files and email in Office.
4. Copy the file you created to your OneDrive synchronization folder. A user notification toast should appear telling you that the action isn't allowed and that the file will be quarantined. For example, for user name *Isaiah Langer*, and a document titled *autoquarantine doc 1.docx* you would see this message:



The message reads:

> Opening auto-quarantine doc 1.docx with this app is not allowed. The file will be quarantined to *C:\Users\IsaiahLanger\Microsoft DLP\OneDrive*

5. Choose **Dismiss**.
6. Open the place holder text file. It is named **auto-quarantine doc 1.docx_*date_time*.txt**.
7. Open the quarantine folder and confirm that the original file is there.
8. Check Activity explorer for data from the monitored endpoints. Set the location filter for devices and add the policy, then filter by policy name to see the effect of the *Scenario 4 Autoquarantine* policy. For information on using activity explorer, see Get started with activity explorer.
9. Check Activity explorer for the event.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Help prevent sharing of sensitive items with unauthorized cloud apps and services

This scenario demonstrates how to restrict unintentional sharing of sensitive information to unapproved cloud applications and services using Microsoft Purview DLP. By defining **sensitive service domains** and enforcing controls through supported browsers, organizations can monitor and guide how sensitive data is uploaded or accessed.

> **Note**
>
> The following web browsers are supported:
>
> - Microsoft Edge (Win/macOS)
> - Chrome (Win/macOS)- Microsoft Purview extension for Chrome Windows only
> - Firefox (Win/macOS)- Microsoft Purview extension for Firefox Windows only
> - Safari (macOS only)

When a policy is configured for the **Devices** location, unsupported browsers are prevented from accessing sensitive content, and users are redirected to Microsoft Edge, where DLP controls can block or restrict actions based on policy conditions. This browser-aware enforcement helps reduce the risk of data exfiltration while maintaining a consistent and guided user experience.

To implement this approach, you define restricted destinations (domains, services, or IPs), specify unsupported browsers, and configure DLP rules that detect sensitive content and apply controls such as **Upload to cloud services** and **Access from unallowed browser**.

The sensitive service domain and browser-restriction configuration enables organizations to audit user behavior, refine policies, and progressively enforce stricter controls as needed, while minimizing disruption to legitimate business activities.

Prerequisites and assumptions

This article uses the process you learned in Design a data loss prevention policy to show you how to create a Microsoft Purview Data Loss Prevention (DLP) policy. Work through these scenarios in your test environment to familiarize yourself with the policy creation UI.

> **Important**
>
> This article presents a hypothetical scenario with hypothetical values. It's only for illustrative purposes. Substitute your own sensitive information types, sensitivity labels, distribution groups, and users.

How you deploy a policy is as important policy design. This article shows you how to use the deployment options so that the policy achieves your intent while avoiding costly business disruptions.

This scenario uses the *Confidential* sensitivity label, so it requires you to create and publish sensitivity labels. To learn more, see:

- Learn about sensitivity labels
- Get started with sensitivity labels
- Create and configure sensitivity labels and their policies

This procedure uses a hypothetical distribution group *Human Resources* and a distribution group for the security team at Contoso.com.

The policy-creation procedure in this article uses alerts. For more information, see Get started with the data loss prevention alerts

Policy intent statement and mapping

We, Contoso, want to prevent users from unintentionally sharing sensitive information to unapproved cloud applications and services from endpoint devices. At the same time, we want to ensure users can continue to access and work with non-sensitive data without unnecessary restrictions.

*To achieve this, we will define a set of **restricted cloud service domains** and enforce controls when sensitive information is detected in user activity. When users attempt to upload sensitive content to these unapproved services—or access such content through unsupported browsers—we will audit the activity and guide users toward supported, compliant workflows (such as using Microsoft Edge).*

Using restricted cloud service domains with audit-first enforcement allows us to gradually enforce controls by first auditing user behavior, understanding risk patterns, and refining the policy before moving to stricter enforcement if needed.

Expand table

Statement	Configuration question answered and configuration mapping
"We want to prevent users from sharing sensitive information to unapproved cloud apps and services..."	- Administrative scope: Full directory - Where to monitor: Devices only - Policy scope: All users/devices (or targeted users for testing)
"We want to define which cloud services are considered unallowed for sensitive data sharing..."	- Endpoint settings: Create **Sensitive service domain group** - Domains defined using URL/IP/IP range (with wildcard support) - Group reused in policy rules
"We want to detect sensitive content being shared to these services..."	- Conditions: Content contains selected **Sensitive info types** - Detection logic: Built-in or custom sensitive information types
"We want to monitor attempts to upload sensitive content or access it through unsupported browsers..."	- Actions: **Upload to a restricted cloud service domain or access from an unallowed browser** - Browser control integrated with Endpoint DLP
"We want to initially observe user behavior without blocking business processes..."	- Action mode: **Audit only** for service domain and browser activities - No blocking or override enforced at this stage
"We want to redirect users toward supported, policy-aware browsers..."	- Endpoint behavior: Unallowed browsers are prevented from accessing sensitive content - User experience: Redirection to Microsoft Edge where DLP controls are enforced
"We want to maintain flexibility to extend protections over time..."	- Design capability: Add more domain groups, apps, and policies as needed - Policy extensibility: Supports future transition to block or block with override
"We want to monitor and optionally control other file activities across apps..."	- Additional actions: Configure **File activities for all apps** as needed - Fine-grained monitoring or restriction of endpoint behaviors
"We want this policy to be active immediately for evaluation..."	- Policy mode: **Turn it on right away** - Deployment: Immediate enforcement in audit mode

Steps to create policy

Perform the following steps to create the sensitive service domain group and configure the DLP policy.

1. Sign in to the Microsoft Purview portal > **Data loss prevention** > **Settings** (gear icon in the upper left hand corner) > **Data Loss Prevention** > **Endpoint DLP settings** > **Browser and domain restrictions to sensitive data** > **Sensitive service domain groups**.
2. Select **Create sensitive service domain group**.
3. Name the group.
4. Enter the **Sensitive service domain** for the group. You can add multiple websites to a group and use wildcards to cover subdomains. For example, `www.contoso.com` for just the top level website, or: *.contoso.com for corp.contoso.com, hr.contoso.com, fin.contoso.com.
5. Select the **Match type** you want. You can select from **URL**, **IP address**, **IP address range**.
6. Select **Save**.
7. In the left navigation pane, select **Data loss prevention** > **Policies**.
8. On the **Choose locations to apply the policy** page, select **Data stored in connected sources** and then choose **Next**.
9. Create and scope a policy that is applied only to the **Devices** location. For more information on how to create a policy, see Create and Deploy data loss prevention policies. Be sure to scope the **Admin units** to **Full directory**.
10. On the **Define policy settings** page, select **Create or customize advanced DLP rules** and choose **Next**.
11. Create a rule, as follows:
    a. Under **Conditions**, select + **Add condition** and select **Content contains** from the drop-down menu.
    b. Give the group a name.
    c. Choose **Add** and then select **Sensitive info types**.
    d. Select a **Sensitive info type** from the flyout pane, then choose **Add**.
    e. Add the action **Audit or restrict activities on devices**.
    f. Under **Service domain and browser activities**, choose **Upload to a restricted cloud service domain or access from an unallowed browser** and set the action to **Audit only**.
    g. Select + **Choose different restrictions for sensitive service domains** and then choose **Add group**.

h. On **the Choose sensitive service domain groups** flyout, select the sensitive service domain group(s) you want, choose **Add** and then choose **Save**.
i. Under **File activities for all apps**, select the user activities you want to monitor or restrict and the actions for DLP to take in response to those activities.
j. Finish creating the rule and choose **Save** and then **Next**.
k. On the confirmation page, choose **Done**.
l. On the **Policy mode** page, choose **Turn it on right away**. Choose **Next** and then **Submit**.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Create policy to manage file activities by implementing network exceptions

This scenario shows how to use **network-based conditions** in Microsoft Purview DLP to apply different protections based on where users access sensitive data. By defining VPN connections and configuring **network exceptions**, the policy adjusts actions—such as auditing or blocking clipboard activity—based on network context, enabling more precise control for hybrid work environments without uniformly restricting user activity.

This scenario is for an unrestricted admin creating a full directory policy.

Prerequisites and assumptions

This scenario requires that you already have devices onboarded and reporting into Activity explorer. If you haven't onboarded devices yet, see Get started with Endpoint data loss prevention.

In this scenario, we define a list of VPNs that hybrid workers use for accessing organization resources.

This article uses the process you learned in Design a data loss prevention policy to show you how to create a Microsoft Purview Data Loss Prevention (DLP) policy. Work through these scenarios in your test environment to familiarize yourself with the policy creation UI.

> **Important**
>
> This article presents a hypothetical scenario with hypothetical values. It's only for illustrative purposes. Substitute your own sensitive information types, sensitivity labels, distribution groups, and users.

How you deploy a policy is as important policy design. This article shows you how to use the deployment options so that the policy achieves your intent while avoiding costly business disruptions.

This scenario uses the *Confidential* sensitivity label, so it requires you to create and publish sensitivity labels. To learn more, see:

- Learn about sensitivity labels
- Get started with sensitivity labels
- Create and configure sensitivity labels and their policies

This procedure uses a hypothetical distribution group *Human Resources* and a distribution group for the security team at Contoso.com.

This procedure uses alerts, see: Get started with the data loss prevention alerts

Policy intent statement and mapping

We, Contoso, want to control how sensitive legal content is handled based on the network context in which users are operating. In particular, we want to apply different levels of restriction depending on whether users are connected through trusted corporate networks or VPN connections used by hybrid workers.

*To achieve this, we will define known VPN connections and use them in **network exception rules** within a DLP policy. This allows us to enforce stricter controls—such as blocking clipboard activity with override—when users are connected through specific VPNs, while maintaining less restrictive (audit-only) behavior in other contexts. This approach enables contextual data protection that adapts to how and where users access sensitive data.*

Expand table

Statement	Configuration question answered and configuration mapping
"We want to apply different data protection controls depending on the network users are connected from…"	- Administrative scope: Full directory - Where to monitor: Devices only - Policy scope: All users/devices or targeted users
"We want to identify VPN connections used by hybrid workers…"	- Endpoint settings: Configure **VPN settings** using Server address or Network address - Data gathered via PowerShell commands (Get-VpnConnection, Get-NetConnectionProfile)
"We want to detect sensitive legal content handled on endpoints…"	- Condition: **Content contains** = Trainable classifiers, **Legal Affairs**
"We want to control specific user activities involving sensitive content…"	- Action: **Audit or restrict activities on devices** - Activity type: **File activities on all apps** - Specific activity: **Copy to clipboard** (extendable to others like print or USB copy)

Statement	Configuration question answered and configuration mapping
"We want less restrictive monitoring under normal conditions..."	- Default activity action: **Audit only** for copy to clipboard
"We want stricter controls when users are connected through defined VPN networks..."	- Network exception: Select **VPN** and set action to **Block with override** - Priority: VPN must be set as top priority in network exception configuration
"We want to support user productivity while maintaining accountability..."	- Override capability: Users can proceed with justification when blocked under VPN conditions
"We want to ensure correct precedence of network-based rules..."	- Configuration behavior: VPN rules take precedence over Corporate network settings when ordered correctly - Caution: 'Apply to all activities' can overwrite other activity-specific configurations
"We want to safely test the policy behavior before full enforcement..."	- Policy mode: **Run in simulation mode** - User experience: **Show policy tips while in simulation mode**
"We want to validate policy behavior through monitoring and testing..."	- Monitoring: Use **Activity explorer** to review policy matches - Testing: Perform clipboard copy action under different network conditions (VPN vs non-VPN)

Steps to create policy

The following procedures walk you through defining VPN settings, verifying corporate network connectivity, and creating the DLP policy with network-based controls.

Create and use a Network exception

Network exceptions enable you to configure Allow, Audit only, Block with override, and Block actions to file activities based on the network from which users access sensitive files. You can select from the VPN settings list you've defined and use the **Corporate network** option. The actions can be applied individually or collectively to these user activities:

- Copy to clipboard
- Copy to a USB removable device
- Copy to a network share
- Print

- Copy or move using unallowed Bluetooth app
- Copy or move using RDP

Get the Server address or Network address

Use the following steps to retrieve the Server address or Network address values needed to define a VPN entry in the endpoint DLP settings.

1. On a DLP monitored Windows device, open a **Windows PowerShell** window as an administrator.
2. Run this cmdlet:

```
PowerShell
Get-VpnConnection
```

3. Running this cmdlet returns multiple fields and values.
4. Find the **ServerAddress** field and record that value. You use this when you create a VPN entry in the VPN list.
5. Find the **Name** field and record that value. The **Name** field maps to the **Network address** field when you create a VPN entry in the VPN list.

Determin if the device is connected through a Corporate network

Use the following steps to verify whether the device is currently connected through a corporate network.

1. On a DLP monitored Windows device, open a **Windows PowerShell** window as an administrator.
2. Run this cmdlet:

```
PowerShell
Get-NetConnectionProfile
```

3. If the **NetworkCategory** field is **DomainAuthenticated**, then the device is connected to a corporate network. If its anything else, the device's connection is not through a corporate network.

Add a VPN

After you collect the server or network address, use the following steps to add the VPN to the Endpoint DLP settings.

1. Sign in to the Microsoft Purview portal.
2. Open **Settings** > **Data Loss Prevention** > **Endpoint settings** > **VPN settings**.
3. Select **Add or edit VPN addresses**.
4. Provide either the **Server address** or **Network address** value that you recorded in the Get the Server address or Network address procedure.
5. Select **Save**.
6. Close the item.

Configure policy actions

Use the following steps to create the DLP policy and configure its network-based controls for file activities.

1. Sign in to the Microsoft Purview portal.
2. Open **Data Loss Prevention** > **Policies**.
3. Select **Create policy**
4. On the **Choose what to protect** page, select **Data stored in connected sources**.
5. Select **Custom** from the **Categories** list, then select the **Custom policy** template from **Regulations**.
6. Name your new policy and provide a description.
7. Select **Full directory** under **Admin units**.
8. Scope the location to **Devices** only.
9. Create a rule where:
   a. **Content contains** = **Trainable classifiers**, **Legal Affairs**
   b. **Actions** = **Audit or restrict activities on devices**
   c. Then pick **File activities on all apps**

d. The select **Apply restrictions to specific activity**

e. Select the actions that you want to configure **Network exceptions** for.

10. Select **Copy to clipboard** and the **Audit only** action

11. Select **Choose different copy to clipboard restrictions**.

12. Select **VPN** and set the action to **Block with override**.

> ⓘ **Important**
>
> When you want to control the activities of a user when they're connected through a VPN *you must* select the VPN and make the VPN the top priority in the **Network exceptions** configuration. Otherwise, if the **Corporate network** option is selected, then that action defined for the **Corporate network** entry will be enforced.

> ⊗ **Caution**
>
> The **Apply to all activities** option will copy the network exceptions that are defined here and apply them to all the other configured specific activities, like **Print**, and **Copy to a network share**. ***This will overwrite the network exceptions on the other activities The last saved configuration wins.***

13. **Save**.

14. Accept the default **Run the policy in simulation mode** value and choose **Show policy tips while in simulation mode**. Choose **Next**.

15. Review your settings and choose **Submit** and then **Done**.

16. The new DLP policy appears in the policy list.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Create a DLP policy that uses device scoping

This scenario shows how to scope a Microsoft Purview Endpoint DLP policy to specific device groups. For example, you can enforce a policy only when Finance users access data from Windows devices while excluding macOS devices. Device groups are commonly defined using dynamic device groups in Microsoft Entra ID.

This scenario is for an unrestricted admin creating a full directory policy.

This scenario requires that you already have devices onboarded and reporting into activity explorer. If you haven't onboarded devices yet, see Get started with Endpoint data loss prevention.

Prerequisites and assumptions

This article uses the process you learned in Design a data loss prevention policy to show you how to create a Microsoft Purview Data Loss Prevention (DLP) policy. Work through these scenarios in your test environment to familiarize yourself with the policy creation UI.

> **Important**
>
> This article presents a hypothetical scenario with hypothetical values. It's only for illustrative purposes. Substitute your own sensitive information types, sensitivity labels, distribution groups, and users.

How you deploy a policy is as important as policy design. Deploy a DLP policy shows you how to use the deployment options so that the policy achieves your intent while avoiding costly business disruptions.

You've created a dynamic device group.

Policy intent statement and mapping

An organization wants to enforce a DLP policy that blocks copying sensitive items with credit card data to USB devices when Finance users access data from Windows devices, and does not apply when the same users work from macOS devices.

Expand table

Statement	Configuration question answered and configuration mapping
"An organization wants to enforce a DLP policy that blocks copying sensitive items with credit card data to USB devices..."	- Administrative scope: Full directory - Where to monitor: Devices only - Policy scope: All users/Finance Windows Devices group
"when Finance users access data from Windows devices..."	- Condition: **Content contains** = **Credit Card Number**
"...and does not apply when the same users work from macOS devices..."	- Action: **Audit or restrict activities on devices** - Activity type: **File activities on all apps** - Restriction model: **Apply restrictions to specific activity** - **Copy to a removable USB device** = **Block**

Create a device group

If you're unfamiliar with dynamic device groups, see:

- Create or update a dynamic membership group in Microsoft Entra ID
- Manage rules for dynamic membership groups in Microsoft Entra ID.

To create the device group for this scenario:

1. Create a new dynamic device group named **Finance Windows Device** for the Windows devices that the finance team uses.

Create a policy scoped to the dynamic device group

1. Sign in to the Microsoft Purview portal > **Data loss prevention** > **policies**.
2. Select **Create policy**.
3. Select **Enterprise applications & devices**.
4. Select the **Custom** from the **Categories** then **Custom policy** template from **Regulations**. Select **Next**.
5. Give your new policy a **Name** and **Description**.

6. Accept the default **Full directory** under **Admin units**.
7. Scope the location to only the **Devices** location.
8. Select **Edit** on the **Devices** location.
9. Accept the default **All users, groups, and adaptive scopes**
10. Select **Specific devices and device groups** and add the **Finance Windows Device** that you created.
11. Create a rule with the following values:
    - **Condition**: **Content contains** = **Sensitive Info types** = **Credit Card Number**
    - **Action**: **Audit or restrict activities on devices** > **File activities for all apps** > **Apply restrictions to specific activity** > **Copy to a removable USB device** = **Block**.
12. Select **Save** and then **Next**.
13. Accept the default **Run the policy in simulation mode** value and choose **Show policy tips while in simulation mode**. Choose **Next**.
14. Review your settings and choose **Submit**.

The new DLP policy appears in the policy list.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Troubleshooting endpoint data loss prevention configuration and policy sync

This article provides detailed instructions for:

1. Determining the device configuration and policy sync status values for Windows devices and macOS devices that are successfully onboarded into Microsoft Purview Data Loss Prevention (DLP).
2. Identifying and resolving any issues with the configuration status and the policy sync status.
3. Reviewing and understanding the device attribute that are available for each device and their meaning.

Device configuration and policy sync status values

Configuration status and the **Policy sync status** of all your onboarded devices have three possible values.

The **Configuration status** value shows you if the device is configured correctly, is sending a heartbeat signal to Purview, and the last time the configuration was validated. For Windows devices, configuration includes checking the status of Microsoft Defender Antivirus always-on protection and behavior monitoring.

The **Policy sync status** shows you if the device received the latest policy version, or if the corresponding policies synced successfully to the device.

Expand table

Field value	Configuration status	Policy sync status
Updated	Device health parameters are enabled and correctly set. This status indicates that the device's configuration is up to date with the recommended settings.	Device is up to date with the current versions of policies.
Not updated	Certain settings need attention. Follow the steps in the workflow diagram to address issues. You might need to enable the configuration settings for this device. Follow	This device isn't synced the latest policy updates. It might take up to 2 hours for the status in the devices list to update. Follow the steps in the workflow diagram to address issues.

Field value	Configuration status	Policy sync status
	the procedures in Microsoft Defender Antivirus always-on protection	
Not available	Device properties aren't available in the device list. This condition might be because the device doesn't meet the minimum OS version to provide visibility into its properties, or configuration, or if the device was just onboarded. Follow the steps in the workflow to address issues.	Device properties aren't available in the device list. This condition might be because the device doesn't meet the minimum OS version to provide visibility into its properties, or configuration, or if the device was just onboarded. Follow the steps in the workflow to address issues. System shows **Not available** if there is no Endpoint DLP policy.

> **Important**
>
> Devices must be online for the policy update to happen. If the status isn't updating, check the last time the device was seen.

Device attribute details

To maintain overall device health from a DLP perspective, go beyond determining the configuration and policy sync status and troubleshooting any issues found. You need to understand the attributes of an onboarded device. The values for these attributes can provide useful information to help you track the device health.

Expand table

Device attribute	Note
Last seen	The most recent time that the device was determined to be online.
Last policy sync time	The timestamp of the previous instance when the device downloaded the latest policy versions.
OS	The current operating system.
Defender engine version	The version of the antivirus engine on the device.
Defender Mocamp version	The version of the Defender client.
MDATP device ID	The unique identifier assigned to this device.

Device attribute	Note
Valid user	This indicates if the currently logged on user has a corresponding Entra ID account and is in scope of a DLP policy that's targeted at Devices.
Sensitive Data Activity	This provides a view all sensitive data activity for this device for the last 30 days.
Advanced classification bandwidth usage exceeded	This attribute shows if the bandwidth usage limit for Advanced Classification has been exceeded in the past 24 hours.
Endpoint DLP status	Shows if Endpoint DLP is enabled or disabled for the device.

Access device attribute data using Advanced Hunting

In addition to viewing device attributes in the Microsoft Purview portal, you can access the same Endpoint DLP device data through the DeviceInfo table in Advanced hunting.

Previously, you could get device attribute data manually through the **Export** functionality on the Device onboarding page in the Microsoft Purview portal.

With the `DeviceInfo` table, you can:

- Query device attribute data using KQL
- Retrieve up-to-date information without manual exports
- Analyze device status across your environment
- Integrate device data into custom dashboards and third-party reporting platforms

Access device data

To retrieve device attribute data:

1. Go to the **Microsoft Defender portal**
2. Navigate to **Investigation & response** > **Hunting** > **Advanced hunting**
3. Query the **`DeviceInfo`** table
4. Expand the **`DlpInfo`** column to view Endpoint DLP device details

Sample query:

Kusto

```
DeviceInfo
| where DlpInfo != ""
| project DlpInfo
```

DlpInfo

The **DlpInfo** column contains Endpoint Data Loss Prevention device information in JSON format. The JSON object can be used to query and report on device configuration status, policy synchronization status, Endpoint DLP enablement, and other configuration details for onboarded devices. These values correspond to the device information available in the DLP device onboarding page.

Example value:

JSON


```
{
 "IsDlpConfigurationValid": true,
 "DlpPolicyLastModifiedTimeUTC": "2026-01-14T20:00:47Z",
 "IsDlpEnabled": true,
 "IsDefenderRealTimeProtectionEnabled": true,
 "IsDefenderBehaviorMonitoringEnabled": true,
 "HasDlpACBandwidthExceeded": false,
 "HasDlpValidUpn": true,
 "DlpUpn": "user@contoso.com"
}
```


The following properties can be returned in the `DlpInfo` object:

Expand table

Property	Data type	Description
IsDlpConfigurationValid	`boolean`	Indicates whether the device configuration health is valid.
DlpPolicyLastModifiedTimeUTC	`datetime`	Timestamp of the most recent DLP policy modification associated with the device, in Coordinated Universal Time (UTC).
IsDlpEnabled	`boolean`	Indicates whether Endpoint DLP is enabled on the device.
IsDefenderRealTimeProtectionEnabled	`boolean`	Indicates whether Microsoft Defender Real-time Protection is enabled on the device.

Property	Data type	Description
IsDefenderBehaviorMonitoringEnabled	`boolean`	Indicates whether Microsoft Defender Behavior Monitoring is enabled on the device.
HasDlpACBandwidthExceeded	`boolean`	Indicates whether the device has exceeded the DLP Advanced Classification bandwidth threshold.
HasDlpValidUpn	`boolean`	Indicates whether the device is associated with a valid user principal name (UPN).
DlpUpn	`string`	User principal name (UPN) associated with the device.

Relationship to device attribute details

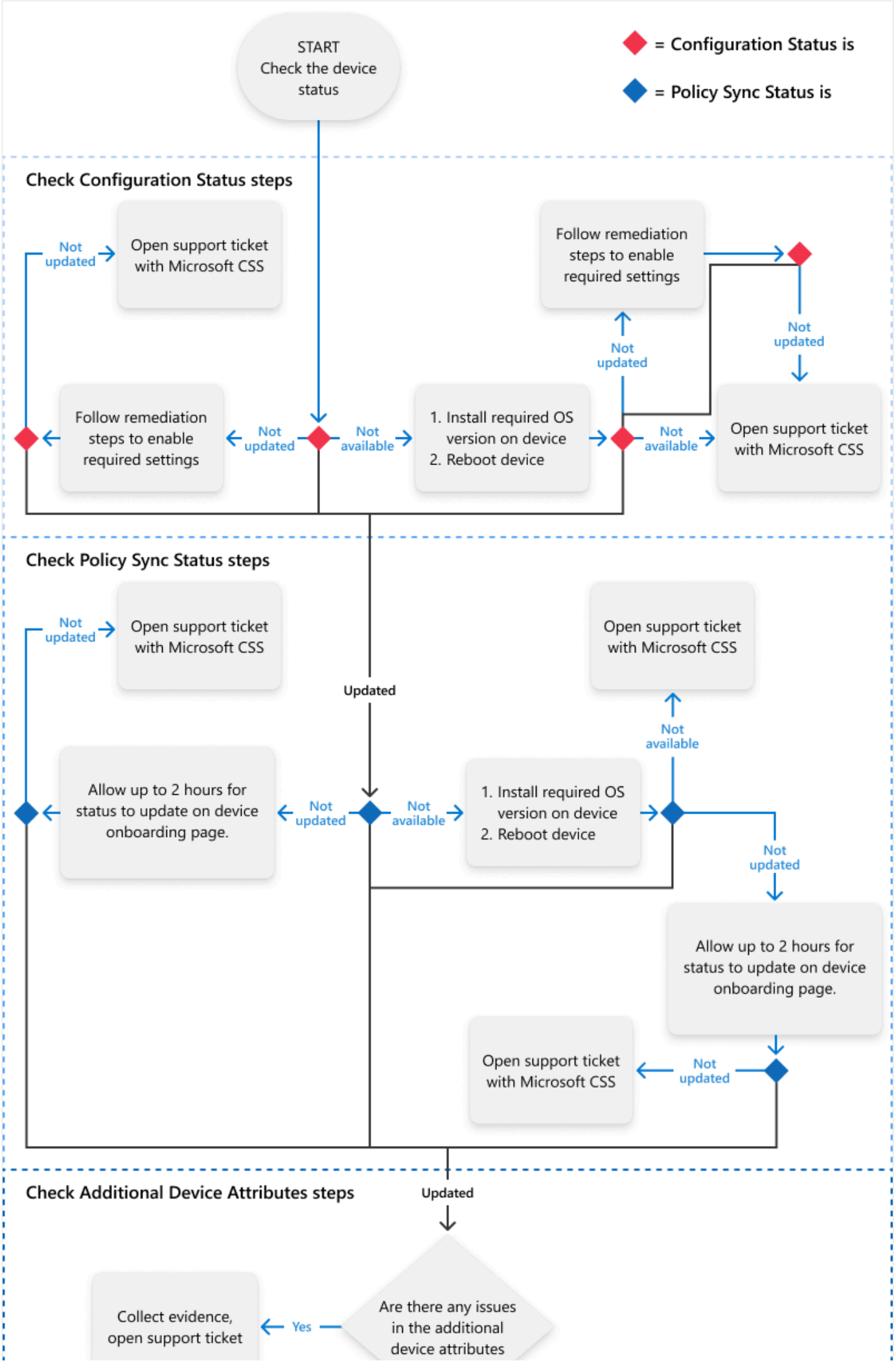
The fields in the `DlpInfo` column correspond directly to the device attributes described earlier. This correspondence enables you to investigate configuration and policy sync issues across multiple devices without relying on point-in-time exports.

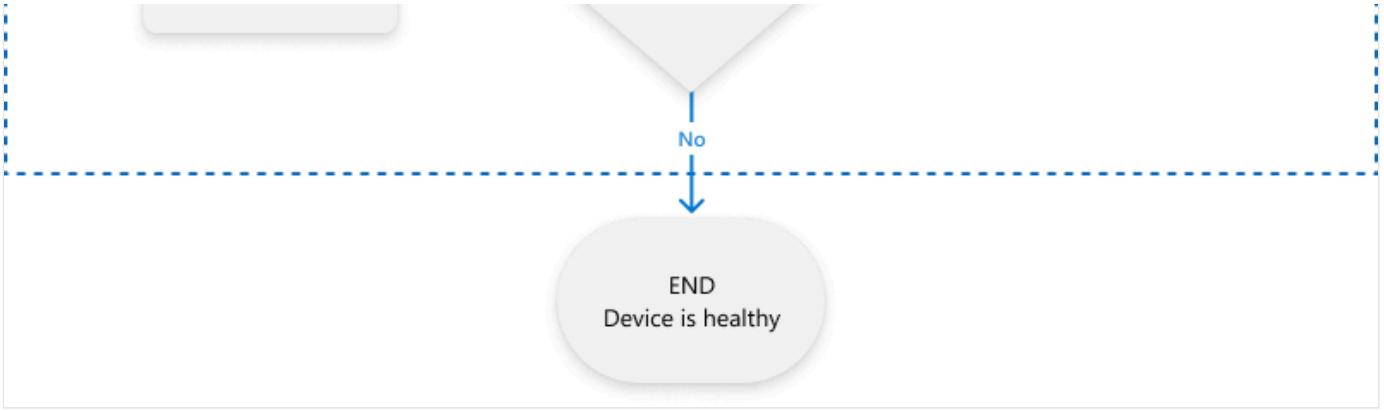
Use this data to:

- Identify devices with invalid configurations
- Detect devices that aren't ready for Endpoint DLP enforcement
- Perform large-scale analysis beyond what the portal UI offers

Configuration and policy sync troubleshooting workflow

This diagram provides a workflow that walks you through the steps for diagnosing and resolving configuration and policy synchronization status for onboarded devices.


START
Check the device status
= Configuration Status is
= Policy Sync Status is
Check Configuration Status steps
Not updated
Open support ticket with Microsoft CSS
Follow remediation steps to enable required settings
Not updated
Not updated
Follow remediation steps to enable required settings
Not updated
Not available
1. Install required OS version on device
2. Reboot device
Not available
Open support ticket with Microsoft CSS
Check Policy Sync Status steps
Not updated
Open support ticket with Microsoft CSS
Updated
Open support ticket with Microsoft CSS
Not available
Allow up to 2 hours for status to update on device onboarding page.
Not updated
Not available
1. Install required OS version on device
2. Reboot device
Not updated
Allow up to 2 hours for status to update on device onboarding page.
Open support ticket with Microsoft CSS
Not updated
Check Additional Device Attributes steps
Updated
Collect evidence, open support ticket
Yes
Are there any issues in the additional device attributes



Check configuration status and resolve issues

1. Sign in to the Microsoft Purview portal > **Settings** (gear icon in the upper right corner) > **Device onboarding** > **Devices**.
2. Apply filters to narrow down the list of devices and simplify your investigation.
3. Select a device to open the details pane for more information on the configuration status.
4. If the status is **Updated**, the device is configured correctly. No further action is required. You can move on to Check policy sync status and resolve issues.
5. If the status is **Not available** or **Not updated**, follow the remediation steps in the details pane and the steps in the workflow diagram.

Check policy sync status and resolve issues

1. Sign in to the Microsoft Purview portal > **Settings** (gear icon in the upper right corner) > **Device onboarding** > **Devices**.
2. Apply filters to narrow down the list of devices and simplify your investigation.
3. Select a device to open the details pane for more information on the policy sync status.
4. If the status is **Updated**, the device successfully received the latest policy version. No further action is required. You can move on to Check device details.
5. If the status is **Not updated** or **Not available**, follow the remediation steps in the details pane and the steps in the workflow diagram.

> **Tip**
>
> You can see the overall status of how policy sync to devices is working on the **Policy status** report. The **Policy status** report is available in the Microsoft Purview compliance portal on the > **Data loss prevention** > **Overview** page.

Check device details

1. Sign in to the Microsoft Purview portal > **Settings** (gear icon in the upper right corner) > **Device onboarding** > **Devices**.
2. Apply filters to narrow down the list of devices and simplify your investigation.
3. Select a device to open the details pane for more information on the specific device attributes under **Device details**.

Collect evidence for a support ticket

If self-remediation isn't successful, gather evidence and open a support ticket for comprehensive support analysis.

From the **Device details** section, record the values for these fields:

- **OS**
- **Defender engine version**
- **Defender client version**
- **MDATP device ID**
- **Valid user**

For more guidance, see:

- Run the client analyzer on Windows
- Run the client analyzer on macOS and Linux

Last updated on 06/29/2026

Learn about Microsoft Purview Data Loss Prevention just-in-time protection

Use endpoint data loss prevention (DLP) just-in-time (JIT) protection to detect and block egress activities on monitored files while policy evaluation completes.

JIT audits and blocks these user egress activities on protected items:

- Copy to a removable media
- Copy to a network share
- Print
- Copy or move using Remote Desktop Protocol (RDP)
- Copy or move using a blocked Bluetooth app
- Copy to clipboard: JIT Audit by default
- Upload to a restricted cloud service domain

Terms

This article uses the following terms:

- **stale classification**: A classification that isn't produced by the most current version of a DLP policy. If you update a policy after an item is evaluated and classified, the item has a stale classification until the policy re-evaluates it.
- **JIT candidate file**: Files that DLP hasn't evaluated or that have a stale classification.
- **JIT audit**: After you enable JIT, Endpoint DLP generates an event in activity explorer for every JIT candidate file. In the JIT activity explorer event, the **JIT triggered** field has the value `true`, and the **Enforcement mode** value is `Audit`.
- **JIT block**: After you enable JIT, Endpoint DLP blocks the activity and generates an event in activity explorer for every JIT candidate file. In the JIT activity explorer event, the **JIT triggered** field has the value `true`, and the **Enforcement mode** field has the value `Block`.

> **Note**
>
> Endpoint DLP doesn't generate a `DLPRuleMatch` event or an alert.

- **JIT in progress notification**: When users who are in scope for JIT attempt an egress activity on a JIT candidate file, Endpoint DLP might block the egress activity and display a toast notification. This toast is called the JIT in progress toast.
- **JIT evaluation complete notification**: When Endpoint DLP finishes policy evaluation for a JIT candidate file, Endpoint DLP shows a toast notification to let the user know. This notification is called the JIT evaluation complete toast.
- **JIT event**: Endpoint DLP records and shows a JIT event in activity explorer when JIT audit or JIT block actions trigger. The event has the `JIT triggered` value set to `true`.
- **Fallback action in case of failure**: This configuration specifies the enforcement mode that DLP should apply when the policy evaluation doesn't complete. No matter which value you select, the relevant telemetry shows in activity explorer.

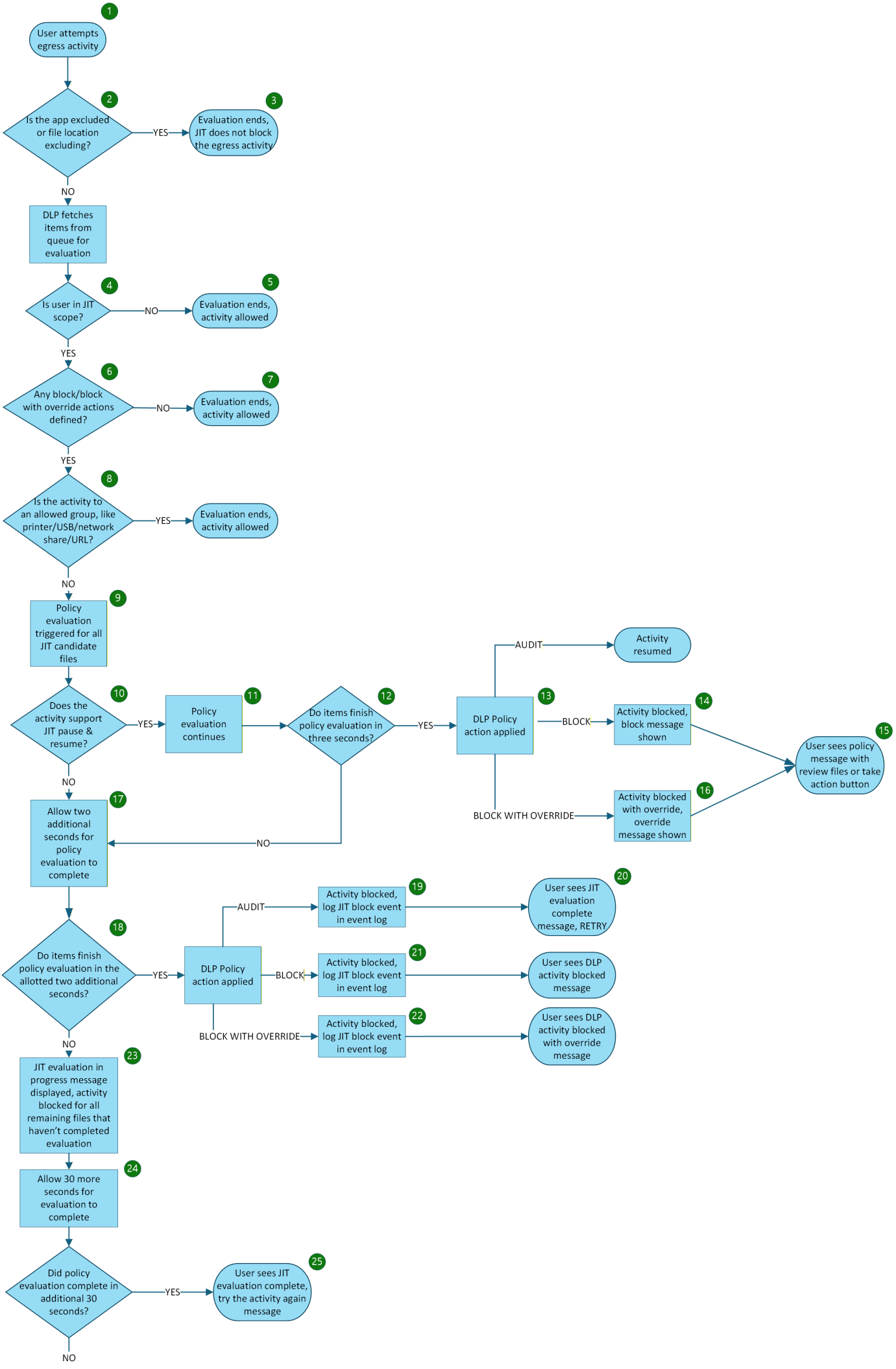


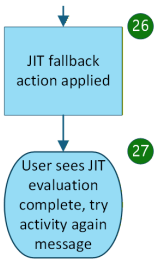
How JIT protection works

JIT protection blocks the egress activity when all the following conditions are true:

- A user attempts an egress activity on an item that has never been classified or that is classified with a stale policy. A stale policy means the file was classified with a policy that has since been updated, and the file hasn't been reclassified with the updated policy.
- The user is in the scope of JIT.
- Microsoft Purview Data Loss Prevention (DLP) policies exist that block or block with override for the *egress activity*.
- The egress activity isn't to an allowed location. For example, an allowed printer, USB device, URL, or network share.
- The egress activity doesn't support JIT pause and resume.
- The DLP policy evaluation *doesn't* complete in five seconds.

JIT workflow


1
User attempts egress activity
2
Is the app excluded or file location excluding?
YES
3
Evaluation ends, JIT does not block the egress activity
NO
DLP fetches items from queue for evaluation
4
Is user in JIT scope?
NO
5
Evaluation ends, activity allowed
YES
6
Any block/block with override actions defined?
NO
7
Evaluation ends, activity allowed
YES
8
Is the activity to an allowed group, like printer/USB/network share/URL?
YES
Evaluation ends, activity allowed
NO
9
Policy evaluation triggered for all JIT candidate files
10
Does the activity support JIT pause & resume?
YES
11
Policy evaluation continues
12
Do items finish policy evaluation in three seconds?
YES
13
DLP Policy action applied
AUDIT
Activity resumed
BLOCK
14
Activity blocked, block message shown
BLOCK WITH OVERRIDE
16
Activity blocked with override, override message shown
15
User sees policy message with review files or take action button
NO
17
Allow two additional seconds for policy evaluation to complete
18
Do items finish policy evaluation in the allotted two additional seconds?
YES
DLP Policy action applied
AUDIT
19
Activity blocked, log JIT block event in event log
20
User sees JIT evaluation complete message, RETRY
BLOCK
21
Activity blocked, log JIT block event in event log
User sees DLP activity blocked message
BLOCK WITH OVERRIDE
22
Activity blocked, log JIT block event in event log
User sees DLP activity blocked with override message
NO
23
JIT evaluation in progress message displayed, activity blocked for all remaining files that haven't completed evaluation
24
Allow 30 more seconds for evaluation to complete
Did policy evaluation complete in additional 30 seconds?
YES
25
User sees JIT evaluation complete, try the activity again message
NO



1. A user attempts an egress activity on an onboarded device for one or more JIT candidate items.
2. If the activity involves an app on the excluded app list, an excluded file path location, or an excluded file extension, the process ends.
3. The evaluation ends and the activity isn't blocked by JIT. No notification message is shown to the user and no JIT audit event is logged.
4. DLP confirms that the user is in scope for JIT. If yes, evaluation continues. If not, the process ends at step 5.
5. JIT doesn't block the activity. No notification is shown, and a JIT audit event is logged.
6. DLP checks if any DLP rule defines **Block** or **Block with override** actions for the attempted activity. If yes, evaluation continues. If not, the process ends with the behavior noted in step 5.
7. JIT doesn't block the activity. No notification is shown, and a JIT audit event is logged.
8. DLP checks if the activity is to an allowed printer group, USB group, network share, or URL. If yes, JIT doesn't block the activity and policy evaluation ends.
9. JIT evaluation is triggered.
10. JIT checks to see if the activity supports JIT pause and resume.
11. If yes, evaluation continues.
12. For all the activities that complete policy evaluation within three seconds, the policy action is applied.
13. For **audit** action, the activity resumes, there's no message shown to the user and a JIT audit event is logged in the audit log.
14. For **block** action, the activity is blocked. A blocked activity message is shown to the user and a JIT audit event is logged in the audit log.

15. User sees policy message with **Review files** button or **Take action** button. Policy evaluation ends.

16. For **block with override** action, the activity is blocked, the block with override message is shown to the user so they can override the block as needed, and both JIT audit event and the block with override event are logged in the audit log.

17. For all the items that didn't complete DLP policy evaluation within three seconds OR don't support resume, JIT evaluation allows two more seconds.

18. For all activities that complete DLP policy evaluation within the allotted two seconds, the policy action is applied.

19. For **audit** policy action, the activity is blocked because the activity doesn't support resume and a JIT block event is logged in the audit log.

20. User sees the policy evaluation complete message and is told to retry. Policy evaluation ends.

21. For **block** policy action, the activity is blocked, a blocked activity message is shown, and a JIT block event is logged in the audit log and evaluation ends.

22. For **block with override** policy action, the activity is blocked, the block with override message is shown so the user can override the block as needed, and both JIT block event and block with override event are logged in the audit log and evaluation ends.

23. For all the items that didn't complete DLP policy evaluation within the total five seconds, the activity is blocked, the **Just-in-time in progress** message is shown to the user. A JIT block event is logged in the audit log.

24. The JIT in progress allows 30 more seconds for the policy evaluation to complete while the activity is blocked.

25. If the DLP policy evaluation completes within these 30 seconds, the user is shown the **Just-in-time evaluation complete** message and the user is asked to retry the activity again.

26. If DLP policy evaluation doesn't complete within the 30 seconds, the JIT fallback action is applied.

27. The **JIT policy evaluation complete** message is shown to the user and evaluation ends. The user is prompted to try the activity again.

User experience of just-in-time protection

This section describes the user experience with antimalware client version **4.18.25080 or later**.

Resume support for each activity

If the policy evaluation finishes within 3 seconds, Endpoint DLP automatically resumes these activities:

- Copy to a removable media
- Copy to a network share

If the policy evaluation takes longer than 3 seconds, you need to repeat the activity after the JIT policy evaluation complete notification appears.

Repeat these activities after Endpoint DLP completes the policy evaluation:

- Print
- Copy or move using Remote Desktop Protocol (RDP)
- Copy or move using unallowed Bluetooth app
- Copy to clipboard: JIT Audit by default

Perform an activity on a single file

When a user performs an activity on a single file, Endpoint DLP takes the JIT audit action when:

- the user isn't in the JIT Scope setting
- there's no **Block** or **Block with override** for the activity
- the activity is to an allowed printer, removable media, network share, or website
- the policy evaluation for the file completes within 5 seconds for activities that support JIT resume, or completes within seconds for activities that don't support JIT resume.

Endpoint DLP blocks the activity with a notification (no alert) and applies JIT block only when the policy evaluation takes more than 5 seconds.

Perform an activity on multiple files

When a user performs an activity on multiple files simultaneously, Endpoint DLP takes the JIT audit action when:

- the user isn't in the JIT Scope setting

- there's no **Block** or **Block with override** for the performed activity
- the activity is to an allowed printer, or to an allowed removable media, or to an allowed network share

For JIT candidate files, Endpoint DLP triggers policy evaluation, consolidates notifications for files that finish within 5 seconds for activities that support resume, and automatically resumes the activity. If the activity doesn't support resume, Endpoint DLP triggers policy evaluation and consolidates notifications for files that finish within 2 seconds. In both cases, Endpoint DLP doesn't raise a JIT in progress toast. It only shows the final policy verdict in the consolidated toast.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/22/2026

Get started with Microsoft Purview Data Loss Prevention just-in-time protection

Use endpoint data loss prevention (DLP) just-in-time (JIT) protection to detect and block egress activities on monitored files while policy evaluation completes.

Applies to

JIT protection for Endpoint DLP supports these devices:

- Windows 10
- Windows 11
- macOS (the three most recent versions)

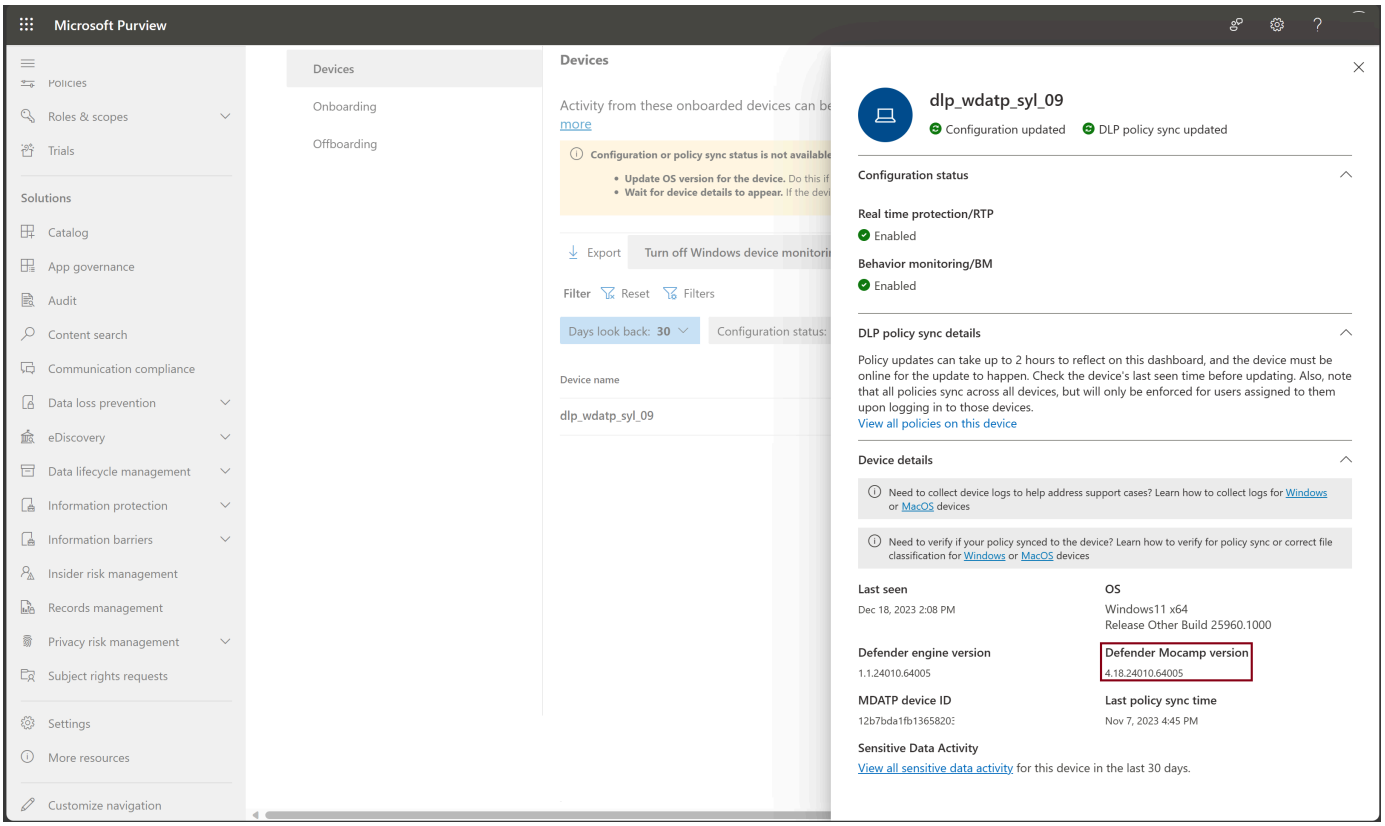
Best practice for deploying Just-in-time protection

> **Note**
>
> Allow at least an hour for JIT setting updates, including disabling JIT, to be pushed to client devices.

Step 1: Prepare your environment

Before you can deploy just-in-time protection, you must first deploy antimalware Client version 4.18.23080 or later. The just-in-time protection end user experience is improved in version 4.18.25080 or later.

> **Tip**
>
> To maximize user productivity, configure and deploy your Endpoint DLP policies to your devices before enabling just-in-time protection. This approach prevents unnecessary blocking of user activity during policy evaluation.



> **Note**
>
> For machines with an outdated version of the antimalware client, disable just-in-time protection by installing one of the following KBs:
>
> - Windows 10 - KB5032278
> - Windows 11 - KB5032288

1. To find out which devices have the necessary antimalware client, go to **Security portal** > **Investigation & response** > **Advanced hunting**, and run this query.

Kusto

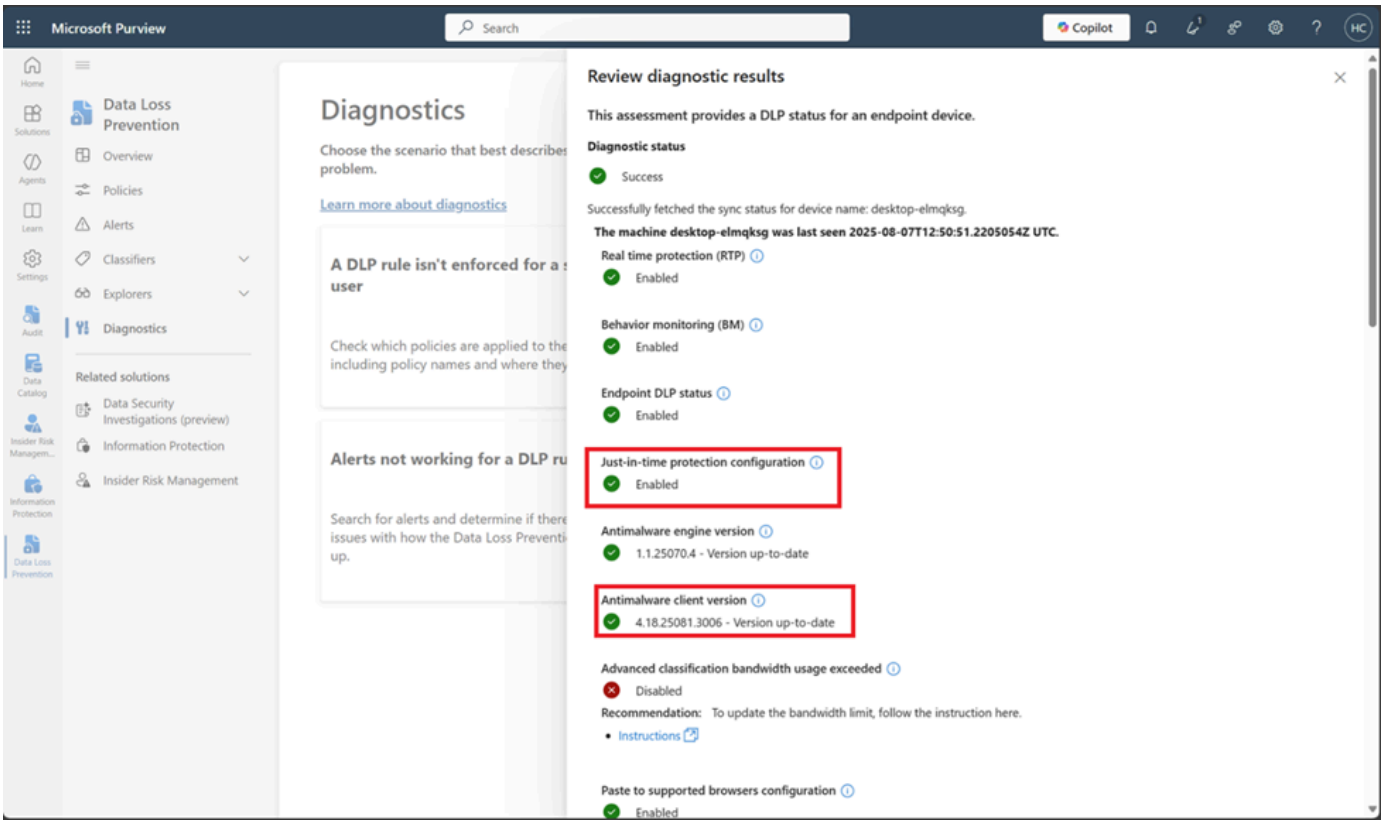
```
`DeviceRegistryEvents`
| where InitiatingProcessVersionInfoInternalFileName == "MsMpEng.exe" and Timestamp >= ago(60d)
| summarize arg_max(Timestamp, *) by DeviceId
| distinct DeviceName, DeviceId, vTimeStamp = Timestamp, AntiMalwareClientVersion = InitiatingProcessVersionInfoProductVersion
| extend Meet_Minimum_JIT_Version = strcmp(AntiMalwareClientVersion, "4.18.23080") // whether the device has required minimum JIT version
| extend Meet_Latest_JIT_Version = strcmp(AntiMalwareClientVersion, "4.18.25080") // whether the device has latest JIT improvement
| project DeviceId, Meet_Latest_JIT_Version, Meet_Minimum_JIT_Version, AntiMalwareClientVersion
| summarize dcount(DeviceId) by AntiMalwareClientVersion // distribution of AntiMalwareClientVersion
```

```
// | summarize dcount(DeviceId) by Meet_Minimum_JIT_Version //how many devices meet
minimum JIT version
// | summarize dcount(DeviceId) by Meet_Latest_JIT_Version //how many devices meet
latst JIT improvements
| order by dcount_DeviceId desc
```

Here's an example of the output of the query.

AntiMalwareClientVersion	dcount_DeviceId
4.18.25090.3009	6
4.18.25080.5	3
4.18.25101.3017	2
4.18.25100.9002	2
4.18.25111.3003	1
4.18.25090.3008	1
4.18.25111.3001	1
4.18.25101.3024	1
4.18.25101.3006	1
4.18.25101.3028	1

You can also go to **Data Loss Prevention** > **Diagnostics page**, and select **Endpoint DLP not working** card to check whether a specific device meets JIT prerequisite.



Step 2: Deploy JIT protection

1. Sign in to the Microsoft Purview portal.
2. Select **Settings** > **Data Loss Prevention** > **Just-in-time protection**.
3. Under **Choose which locations to monitor**, select the checkbox next to **Devices**.
4. Under **Fallback action in case of failure**, select **Allow users to complete actions**. This option lets the user action complete if the classification fails.

> **⊗ Caution**
>
> Don't choose the **Block users from completing actions** option until you fully understand the impact of this feature.

Selecting **Allow users to complete actions** or **Block users from completing actions** doesn't change whether *JIT Block* is triggered. Blocking by JIT is applied if the user is in scope. The **Allow users to complete actions** or **Block users from completing actions** setting controls Endpoint DLP enforcement when classification fails.

If you select **Allow users to complete actions**, Endpoint DLP allows the egress activity when classification fails. If you select **Block users to complete actions**, Endpoint DLP blocks the egress

activity when classification fails.

Step 3: Estimate the number of JIT protection events for your deployment

Validate the JIT protection scope and fallback action settings after each rollout stage until the number of events is stable. Make sure you understand the possible size of the user group you want to enforce the policy on, based on the following telemetry calculations.

Estimate the impact of deploying JIT protection by performing the following calculation based on the events in activity explorer:

- N = The number of unique machines firing JIT events.
- S = The total number of machines within the scope of your deployment.

`N/S` yields the percentage of machines that might experience a JIT protection block event.

With this information, you should know how many machines will be affected by implementing the JIT Block mode when you expand the scope, and how many possible support tickets you may see. Then, you can decide whether or not to expand the scope.

Step 4: Fine-tune JIT protection through other Additional settings

In addition to the **Fallback action in case of failure** setting, you can also use the following settings to fine-tune JIT protection:

- **Control copying to clipboard**: Turn on this setting to prevent users from copying content to the clipboard while JIT protection is evaluating the file.

> ⓘ **Note**
>
> Turning on **Control copying to clipboard** might impact user's productivity. Be sure to test the impact on productivity before turning on this setting.

- **App exclusions for Windows**: You can exclude maximum 50 apps here from JIT protection on Windows devices.
- **App exclusions for Mac**: You can exclude maximum 50 apps here from JIT protection on Mac devices.

- **File extensions exclusions**: Files with extensions you add here aren't evaluated by JIT protection.
- **File path exclusions for Windows**: Files in these locations aren't evaluated by JIT protection.
- **File path exclusions for Mac**: Files in these locations aren't evaluated by JIT protection.

If you want to change the scope of JIT protection after tuning all these settings, return to **Settings** > **Data Loss Prevention** > **Just-in-time protection** in the Microsoft Purview portal and update the monitored locations.

Exclusions for Just-in-time protection

File path exclusion settings in JIT are different from **File path exclusions for Windows** setting found via **Data loss prevention** > **Settings** > **Endpoint settings** > **File path exclusions for Windows**.

- **File path exclusions** in JIT only excludes specific file paths from JIT protection. In all other cases, Microsoft Purview still applies Endpoint DLP classification and protection for files in those folders.
- **File path exclusions for Windows** setting prevents Purview from applying Endpoint DLP classification and protection for files under the specified folders.
- **File extension exclusions**: Files with these extensions aren't evaluated by JIT protection.

Step 5: Deploy JIT protection in 'Block users from completing actions' for the 'Fallback action in case of failure' setting

The **Block users from completing actions** option for **Fallback action in case of failure** controls the enforcement mode that DLP applies when classification fails. This setting doesn't control JIT Block or JIT Audit for JIT candidate files. JIT Block or JIT Audit is controlled by how the policy is scoped. No matter which value you select here, the relevant telemetry displays in activity explorer.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Help protect files that Endpoint Data Loss Prevention fails to scan

This article uses the process you learned in Design a data loss prevention policy to show you how to create a Microsoft Purview Data Loss Prevention (DLP) policy that helps protect files on supported Windows and macOS devices where the scanning process fails. Work through this scenario in your test environment to familiarize yourself with the policy creation UI.

> **Note**
>
> Applying controls to supported files that fail scanning is currently in preview.

Use this scenario when you want to apply Audit, Block, or Block with override controls on user activities for files that are on the Monitored files list but fail to be scanned by Endpoint DLP.

> **Important**
>
> This article presents a hypothetical scenario with hypothetical values. It's only for illustrative purposes. Substitute your own sensitive information types, sensitivity labels, distribution groups, and users.

How you deploy a policy is as important policy design. This article shows you how to use the deployment options so that the policy achieves your intent while avoiding costly business disruptions.

Prerequisites and assumptions

This scenario requires that you already have devices onboarded and reporting into activity explorer. If you haven't onboarded devices yet, see Get started with Endpoint data loss prevention.

> **Important**
>
> This feature only supports following action types:
>
> - Upload to a restricted cloud service domain

- Copy to a removable USB device
- Copy to a network share
- Print

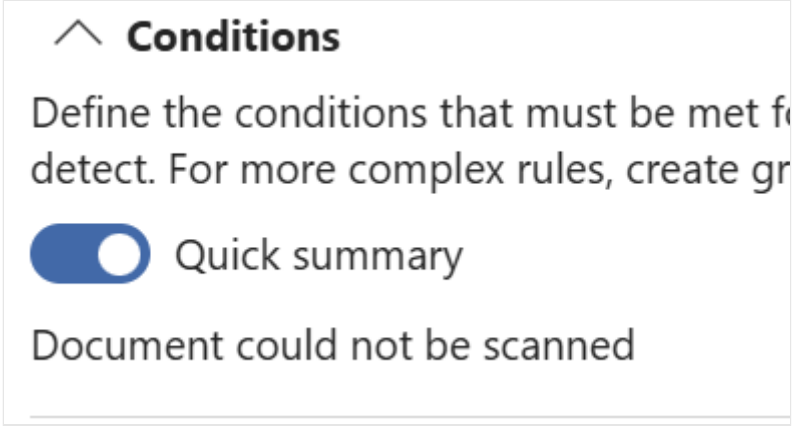
Policy intent statement and mapping

Sometimes scanning fails for files that endpoint DLP supports. There might be sensitive information in those files, but we don't know. When scanning fails for a file on any of our onboarded devices, we want to prevent users from copying that file to a USB device or to a network share.

Expand table

Statement	Configuration question answered and configuration mapping
"When scanning fails for a file on any of our onboarded devices ..."	- Administrative scope: **Full directory** - Where to monitor: **Enteprise applications & devices**, **Devices** -Scope: **Allusers, groups, devices, device groups** - Conditions: **Document could not be scanned**
"...we want to prevent users from copying that file to a USB device or to a network share."	- Action : select **audit or restrict activities on devices** - clear **Upload to a restriced cloud service domain or access from an unallowed browser** - select **Apply restrictions to specific activity** - select **Copy to a removable USB device**, and **Block with override** - select **Copy to a network share** and **Block with override** clear **Copy to clipboard**, **Print**, **Copy or move using unallowed Bluetooth app**, and **Copy or move using RDP**

When you configure the conditions, the summary looks like this:



Steps to configure policy actions

Perform the following steps to configure the policy actions for this scenario.

> **Important**
>
> For the purposes of this policy creation procedure, leave the policy turned off. Change the policy activation settings when you deploy the policy.

1. Sign in to the Microsoft Purview portal.
2. Open **Data Loss Prevention** > **Policies**.
3. Select **Create policy**.
4. Select **Enterprise applications & devices**.
5. Select the **Custom** from the **Categories** then **Custom policy** template from **Regulations**.
6. Name your new policy and provide a description.
7. Select **Full directory** under **Admin units**.
8. Scope the location to **Devices** only.
9. Create a rule where:
    a. In **Conditions**:
        i. Select **Document could not be scanned**
    b. In **Actions**:
        i. Select **Audit or restrict activities on devices**.
        ii. Clear **Upload to a restriced cloud service domain or access from an unallowed browser**
        iii. Select **Apply restrictions to specific activity**
        iv. Select **Copy to a removable USB device**, and **Block with override**
        v. Select **Copy to a network share**, and **Block with override**
        vi. Clear **Copy to clipboard**, **Print**, **Copy or move using unallowed Bluetooth app**, and **Copy or move using RDP**
10. **Save**.
11. Choose **Turn on the policy immediately**. Choose **Next**.
12. Review your settings and choose **Submit**.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Help protect files that Endpoint Data Loss Prevention doesn't scan

This article uses the process you learned in Design a data loss prevention policy to show you how to create a Microsoft Purview Data Loss Prevention (DLP) policy that helps protect files that Endpoint DLP doesn't support. Work through these scenarios in your test environment to familiarize yourself with the policy creation UI.

Apply controls to all unsupported files, is in *preview*.

Use the "Apply controls to all unsupported files" scenario to apply Audit, Block, or Block with override controls on user activities for files that aren't on the Monitored files list without enumerating all the file extensions through the **File extension is** condition. Use the **Apply restrictions to only unsupported file extensions** configuration to create a blanket policy to place controls on files like .mp3, .wav, and .dat.

> ⊗ **Caution**
>
> The ability to apply controls to all files is a powerful capability. If you implement this capability without proper caution, the policy might have unintended consequences. Test this policy in a nonproduction environment before deploying it to your production environment. The unsupported-files policy example also shows how you can (optionally) exclude specific file extensions from the scope of the policy.

> ⓘ **Important**
>
> When you select the **Audit or restrict activities on devices** action in the policy configuration **Actions**, the **Apply restrictions to only unsupported file extensions** shows up. **Apply restrictions to only unsupported file extensions** configuration option ***DOES NOT*** support scoping by **Device and device groups** in the policy location setting.

These actions are supported:

- Upload to a restricted cloud service domain
- Copy to a removable USB device
- Copy to a network share

- Print

> **Important**
>
> This article presents a hypothetical scenario with hypothetical values. It's only for illustrative purposes. Substitute your own sensitive information types, sensitivity labels, distribution groups, and users.

How you deploy a policy is as important policy design. This article shows you how to use the deployment options so that the policy achieves your intent while avoiding costly business disruptions.

Policy intent statement and mapping

*There are many types of files that endpoint DLP doesn't scan, more than we can track. We don't know if there is sensitive information in those files, and so we want to have a checkpoint in place before our user try to copy those files to a USB device or to a network share. We don't want to disrupt users workflow for the **abc** file type which is well known to us and poses no threat of data leakage.*

Expand table

Statement	Configuration question answered and configuration mapping
“There are many types of files that endpoint DLP doesn't scan,...”	- Administrative scope: **Full directory** - Where to monitor: **Enterprise applications & devices**, **Devices** - Scope: **Allusers, groups, devices, device groups**
"...We don't want to disrupt users workflow for the **abc** file type which is well known to us and poses no threat of data leakage."	- Endpoint settings: create a Unsupported file extension exclusions list and add the file extension **abc** to the list.
"...more than we can track. We don't know if there is sensitive information in those files, and so we want to have a checkpoint in place before they try to copy those files to a USB device or to a network share..."	- Conditions for a match: **Document could not be scanned** action : select **audit or restrict activities on devices** - clear **Upload to a restriced cloud service domain or access from an unallowed browser**

Statement	Configuration question answered and configuration mapping
	- select **Apply restrictions to specific activity** - select **Copy to a removable USB device**, and **Block with override** - select **Copy to a network share and Block with override** - clear **Copy to clipboard**, **Print**, **Copy or move using unallowed Bluetooth app**, and **Copy or move using RDP** - select **file could not be scanned**. - select **apply restrictions to only unsupported file extensions**.

> **Important**
>
> The difference between the **Apply restrictions to only unsupported file extensions** option and the **File extension is** condition is:
>
> - Endpoint DLP scans content for **File extension is** condition. For example, you see **Sensitive info type** value on the event or alert; on the other hand, the **Apply restrictions to only unsupported file extensions** option doesn't scan file content.
> - **File extension is** condition triggering content scanning might consume higher machine resource, like CPU and memory, and might cause application performance issue for some file types.

Steps to add a file type to unsupported file extension exclusions

Use the **Unsupported file extension exclusions** setting to exclude file extensions from the policy.

1. Sign in to the Microsoft Purview portal
2. Open **Settings** > **Data Loss Prevention** > **Endpoint DLP settings** > **Unsupported file extension exclusions**.
3. Select **Add file extensions**.
4. Provide extensions.
5. Select **Save**.
6. Close the item.

Steps to configure policy actions

1. Sign in to the Microsoft Purview portal.
2. Open **Data Loss Prevention** > **Policies**.
3. Select **Enterprise applications & devices**.
4. Select **Create policy** and select the **Custom** from the **Categories** then **Custom policy** template from **Regulations**.
5. Name your new policy and provide a description.
6. Select **Full directory** under **Admin units**.
7. Scope the location to **Devices** only.
8. Create a rule where:
    a. In **Conditions**:
        i. Select **Document could not be scanned**.
    b. In **Actions**:
        i. Select **Audit or restrict activities on devices**.
        ii. Clear **Upload to a restriced cloud service domain or access from an unallowed browser**.
        iii. Select **Apply restrictions to specific activity**.
        iv. Select **Copy to a removable USB device**, and **Block with override**.
        v. Select **Copy to a network share**, and **Block with override**.
        vi. Clear **Copy to clipboard**, **Print**, **Copy or move using unallowed Bluetooth app**, and **Copy or move using RDP**.
        vii. Select **Apply restrictions to only unsupported file extensions**.
9. **Save**.
10. Choose **Turn on the policy immediately**. Choose **Next**.
11. Review your settings and choose **Submit**.

> **Important**
>
> You can't use **Document could not be scanned** together with other conditions in the "Apply controls to all unsupported files" scenario.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Help protect against sharing of a defined set of unsupported files

Create a Microsoft Purview Data Loss Prevention (DLP) policy that helps protect against sharing of a defined set of unsupported files. This article walks you through the configuration steps using the process described in Design a data loss prevention policy. Work through these scenarios in your test environment to familiarize yourself with the policy creation UI.

How you deploy a policy is as important policy design. This article shows you how to use the deployment options so that the policy achieves your intent while avoiding costly business disruptions.

Apply controls to some file types is in *preview*.

Use the unsupported file extensions configuration to apply Audit, Block with override, or Block controls to a list of file types only. For example, you can apply controls to video files like .mp3 files (which aren't monitored by endpoint Data Loss Prevention) only.

> **ⓘ Important**
>
> When you select the **Audit or restrict activities on devices** action in the policy configuration **Actions**, the **Apply restrictions to only unsupported file extensions** shows up. **Apply restrictions to only unsupported file extensions** configuration option ***DOES NOT*** support scoping by **Device and device groups** in the policy location setting.

The following device actions are supported for unsupported file extension restrictions:

- Upload to a restricted cloud service domain
- Copy to a removable USB device
- Copy to a network share
- Print

The unsupported file extensions policy configuration uses a combination of a **File extension group** and the **Document could not be scanned** condition. The unsupported file extensions configuration doesn't use the **File extension is** condition. This policy configuration means that Microsoft Purview endpoint Data Loss Prevention (DLP) doesn't scan the content of the files that

you include in the **File extension group** and you don't see values for **Sensitive info type** in events or alerts that are generated by policy matches.

> **Important**
>
> This article presents a hypothetical scenario with hypothetical values. It's only for illustrative purposes. Substitute your own sensitive information types, sensitivity labels, distribution groups, and users.

Policy intent statement and mapping

We, Wingtip Toys, have a list of file types on devices which are NOT on the monitored files list and that we want to apply controls to. We know that they're not on the monitored files list for Endpoint DLP. We want to prevent users from copying those files to a USB device or to a network share. When they do try, we want to let them know, to educate them, that they are attempting a prohibited action.

Expand table

Statement	Configuration question answered and configuration mapping
“We have a list of file types on devices that are *NOT* on the Monitored files list and that we want to apply controls to....”	- Administrative scope: **Full directory** - Where to monitor: **Devices** -Scope: **Allusers, groups, devices, device groups**
"we want to apply controls to a list of file types on devices that are on the Monitored files list but that is not covered by Endpoint DLP scanning..."	- Conditions for a match: **Document could not be scanned** - Action: select **audit or restrict activities on devices** - clear **Upload to a restriced cloud service domain or access from an unallowed browser** - select **Apply restrictions to specific activity** - select **Copy to a removable USB device**, > **Block** - **Copy to a network share** > **Block** - clear **Copy to clipboard**, **Print**, **Copy or move using unallowed Bluetooth app**, and **Copy or move using RDP** - select **file could not be scanned** - select **apply restrictions to only unsupported file extensions**.
"We know that they're not on the monitored files list for Endpoint DLP."	- Endpoint settings: create a **File extension group**
"...We need to prevent users from copying those files to a USB device or to a network share."	- Conditions for a match: **Document could not be scanned** - Action: select **audit or restrict activities on devices** - clear **Upload to a restricted cloud service domain or access from**

Statement	Configuration question answered and configuration mapping
	an unallowed browser - select **Apply restrictions to specific activity** - select **Copy to a removable USB device**, > **Block** - **Copy to a network share** > **Block** - clear **Copy to clipboard**, **Print**, **Copy or move using unallowed Bluetooth app**, and **Copy or move using RDP** - select **file could not be scanned** - select **apply restrictions to only unsupported file extensions**.
"When they do try, we want to let them know, to educate them, that they are attempting a prohibited action."	- Use notifications to inform your users and help educate them on the proper use of sensitive info: **On** - Endpoint devices > Show users a policy tip notification when an activity is restricted...: **selected** - Customize the notification: **selected** > Notification Title: **Wingtip toys don't copy files** > **Notification Content**: FYI, Wingtip Toy policy doesn't let you copy that type of file to USB device or a network share. *(preview) Ability to add hyperlinks in Endpoint policy tip notifications*

Steps to create a file extension group

The following steps use the Microsoft Purview portal. To learn more, see Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal.
2. Open **Settings** > **Data Loss Prevention** > **Endpoint DLP settings** > **File extension groups**.
3. Select **Create file extension group** and enter a **Group name**. In this scenario, use `Non-classified file extensions`.
4. Provide extensions.
5. Select **Save**.
6. Close the item.

Steps to configure policy actions

Perform the following steps to configure the policy actions for this scenario.

1. Sign in to the Microsoft Purview portal.
2. Open **Data Loss Prevention** > **Policies**.
3. Select **Create policy**.
4. Select **Enterprise applications & devices**.
5. Select **Custom** from the **Categories** then select **Custom policy** template from **Regulations**.
6. Name your new policy and provide a description.

7. Select **Full directory** under **Admin units**.
8. Scope the location to **Devices** only.
9. Create a rule where:
    a. In **Conditions**.
        i. **Document could not be scanned**.
    b. In **Actions**:
        i. Select: **Audit or restrict activities on devices**.
        ii. Clear: **Upload to a restricted cloud service domain or access from an unallowed browser**.
        iii. Select: **Apply restrictions to specific activity**.
        iv. Clear: **Copy to clipboard**.
        v. Select: **Copy to a removable USB device** > **Block**.
        vi. Select: **Copy to a network share** > **Block**.
        vii. Clear: **Print**.
        viii. Clear: **Copy or move using unallowed Bluetooth app**.
        ix. Clear: **Copy or move using RDP**.
        x. Clear: **Access by restricted apps**.
        xi. Select: **Apply restrictions to only unsupported file extensions**.
        xii. Select: **Add file extension group** and select `Non-classified file extensions`.
10. Set **User notifications** to **On**.
    a. Under **Endpoint devices** select **Show users a policy tip notification when an activity is restricted...**.
    b. Select **Customize the notification**.
        i. Enter `Wingtip toys don't copy files` in the **Notification Title**.
        ii. Enter `FYI, Wingtip Toy policy doesn't let you copy that type of file to USB device or a network share` in the **Notification Content**.
11. **Save**.
12. Choose **Turn the policy on immediately**. Choose **Next**.
13. Review your settings and choose **Submit**.

> **Important**
>
> You can't use the **Document could not be scanned** condition with other conditions in this scenario.

Note: The author created this article with assistance from AI. Learn more

Last updated on 06/26/2026

Disable Microsoft Purview data loss prevention scanning for some supported files and apply controls

This article shows you how to create a Microsoft Purview data loss prevention (DLP) policy that disables scanning for selected supported file types and applies audit or restriction controls to those files, using the process described in Design a data loss prevention policy. Work through this scenario in your test environment to familiarize yourself with the policy creation UI.

Disable scanning for some supported files and apply controls is in *preview*.

Use the disable-classification configuration to save local resource consumption by disabling scanning of some file types on the Monitored files list. You can apply Audit, Block, or Block with override controls to those file types.

> **Important**
>
> This feature only supports the following action types:
>
> - Upload to a restricted cloud service domain
> - Copy to a removable USB device
> - Copy to a network share
> - Print

> **Important**
>
> This article presents a hypothetical scenario with hypothetical values. It's only for illustrative purposes. Substitute your own sensitive information types, sensitivity labels, distribution groups, and users.

How you deploy a policy is as important as policy design. This article shows you how to use the deployment options so that the policy achieves your intent while avoiding costly business disruptions.

Policy intent statement and mapping

Bellows College needs to conserve resources on all our users Windows devices and cutting down on the scanning of files by DLP would be a big help. We have a list of file types that we know the contents of because they are auto generated. These file types are on the supported file types list. There's no need to scan these auto generated files, but we do want to prevent users from copying them to a USB device or to a network share. When they do try, we want to let them know, to educate them, that they are attempting a prohibited action.

Expand table

Statement	**Configuration question answered and configuration mapping**
"Bellows College needs to conserve resources on all our users Windows devices and cutting down on the scanning of files by DLP would be a big help..."	- Administrative scope: **Full directory** - Where to monitor: **Enterprise applications & devices**, **Devices** -Scope: **Allusers, groups, devices, device groups**
"...We have a list of file types that we know the contents of because they are auto generated. These file types are on the supported file types list. There's no need to scan these auto generated files..."	- Endpoint settings: create a **File extension group** - Disable classification
...but we do want to prevent users from copying them to a USB device or to a network share...	- Conditions for a match: **Document could not be scanned** Action : select **audit or restrict activities on devices** - clear **Upload to a restriced cloud service domain or access from an unallowed browser** - select **Apply restrictions to specific activity** - select **Copy to a removable USB device**, > **Block** - **Copy to a network share** > **Block** - clear **Copy to clipboard**, **Print**, **Copy or move using unallowed Bluetooth app**, and **Copy or move using RDP** - select **file could not be scanned**.
"...When they do try, we want to let them know, to educate them, that they are attempting a prohibited action..."	- Use notifications to inform your users and help educate them on the proper use of sensitive info: **On** - Endpoint devices > Show users a policy tip notification when an activity is restricted...: **selected** - Customize the notification: **selected** > **Notication Title**: Bellows College IT don't copy files > **Notification Content**: FYI, Bellows College data loss prevention policies don't let you copy that type of file to USB device or a network share

Create a File extension group

Select the appropriate tab for the portal you're using. To learn more about the Microsoft Purview portal, see Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal.
2. Open **Settings** > **Data Loss Prevention** > **Endpoint DLP settings** > **File extension groups**.
3. Select **Create file extension group** and enter a **Group name**. In this scenario, we use `Student Class Registration file extensions`.
4. Enter the extensions.
5. Select **Save**.
6. Close the item.

Disable classification

Use the **Disable classification** setting to exclude specific file extensions from Endpoint DLP classification.

1. Sign in to the Microsoft Purview portal.
2. Open **Settings** > **Data Loss Prevention** > **Endpoint DLP settings** > **Disable classification**.
3. Select **Add or edit file extensions**.
4. Enter the extensions.
5. Select **Save**.
6. Close the item.

Configure policy actions

Perform the following steps to configure the policy actions for this scenario.

1. Sign in to the Microsoft Purview portal.
2. Open **Data Loss Prevention** > **Policies**.
3. Select **Create policy**.
4. Select **Enterprise applications & devices**.
5. Select **Custom** from the **Categories** then select **Custom policy** template from **Regulations**.
6. Name your new policy and provide a description.
7. Select **Full directory** under **Admin units**.
8. Scope the location to **Devices** only.
9. Create a rule where:
    a. In **Conditions**:
        i. **Document could not be scanned**.
    b. In **Actions**:

i. Select: **Audit or restrict activities on devices**.
ii. Select: **Apply restrictions to specific activity**.
iii. Clear: **Copy to clipboard**.
iv. Select: **Copy to a removable USB device** > **Block**.
v. Select: **Copy to a network share** > **Block**.
vi. Clear: **Print**.
vii. Clear: **Copy or move using unallowed Bluetooth app**.
viii. Clear: **Copy or move using RDP**.
ix. Clear: **Access by restricted apps**.
x. Select: **Apply restrictions to only unsupported file extensions**.
xi. Select: **Add file extension group** and select `Student Class Registration file extensions`.

c. **Save**.

10. Choose **Turn it on right away**. Choose **Next**.
11. Review your settings and choose **Submit**.

> **Important**
>
> You can't use the **Document could not be scanned** condition with other conditions in this policy scenario.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Always-on diagnostics for endpoint DLP

The Always-on diagnostics feature in Microsoft Purview enables continuous, automated trace logging for endpoint data loss prevention (DLP). This functionality significantly reduces administrative overhead by eliminating the need for manual log configuration and issue reproduction.

When a support case is opened with Microsoft, diagnostic logs are often required to assist with root cause analysis. With Always-on diagnostics enabled, comprehensive telemetry is already available, allowing administrators to bypass traditional data collection steps. This not only accelerates the support workflow but also improves accuracy in identifying and resolving issues.

Diagnostic logs can be securely uploaded directly to Microsoft Support, streamlining case submission and accelerating time-to-resolution. This proactive logging approach enhances operational efficiency and improves support responsiveness.

Trace logs

When an admin turns on Always-on diagnostics, the feature is enabled on all onboarded Windows devices, including servers, and all onboarded macOS devices. When an Endpoint DLP policy does not behave as expected for an in-scope user or device, the customer opens a Microsoft support case. To expedite troubleshooting, Microsoft Support requests diagnostic trace logs, which the customer administrator collects from the affected device using documented procedures. The collected data is protected by Microsoft through restricted access controls and a proprietary log format.

Understanding upload behavior and operational constraints

Trace uploads typically occur within a 24-hour polling window and require the device to remain online for the upload to complete. Only one active collection request is supported per device at any given time, and additional requests can be initiated only after the prior request completes, fails, or is deleted. Restart-based upload acceleration is supported, but restarting the device isn't required to complete an upload. It's intended only to speed up the process. Collection requests may remain pending if the device is offline, wiped, reimaged, reassigned, or otherwise unable to respond.

Admins are responsible for monitoring the status of collection requests and for taking action when failures occur. If a request fails, admins must retry or re-initiate the request as needed to ensure successful data collection.

Understanding log access, security, and data storage compliance

Logs aren't available for direct download by admins and are stored in a proprietary Microsoft-internal format that can only be decoded using Microsoft tools. Access to logs is restricted to Microsoft personnel who are members of designated access groups, and all such access is fully audited to ensure accountability and security.

Logs are stored in Microsoft-managed Azure Blob Storage and comply with Microsoft security and privacy standards, including GDPR. Data remains within the tenant's data residency region and is retained for 180 days unless manually deleted earlier. After the retention period, logs are automatically purged and can't be recovered, and no backup or extended retention is supported.

> **Important**
>
> This feature is only for authorized eDLP troubleshooting scenarios (whether functioning or not). It must not be used for investigative or non-support purposes, including any activity outside approved troubleshooting workflows.
>
> This feature is supported on Windows and macOS.

Permissions required for always-on diagnostics

Roles required to view and create log collection requests

There are multiple roles that allow you to view and create log collection requests. The account you use must be a member of any one of them.

Microsoft Entra roles

- Compliance Administrator
- Security Administrator
- Global Administrator

> **Important**

Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

Purview roles for viewing and creating log collection requests

Must be assigned at the tenant level; scoped admins are not supported:

- OrganizationConfiguration
- ManageAlerts
- ViewOnlyManageAlerts
- InformationProtectionAdmin
- InformationProtectionAnalyst
- InformationProtectionInvestigator

Roles required to enable the feature

The account that you use must be a member of any one of the following roles to turn on Always-on diagnostics.

Microsoft Entra roles for enabling always-on diagnostics

- Compliance Administrator
- Security Administrator
- Global Administrator

Important

Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

Purview roles (tenant-level only)

The following tenant-level Purview roles can enable Always-on diagnostics:

- OrganizationConfiguration
- ComplianceAdmin
- SecurityAdmin

- DLPComplianceManagement
- InformationProtectionAdmin

Prerequisites

The device must be onboarded to Microsoft Purview and actively reporting with Always-on Diagnostics enabled. It must maintain continuous network connectivity and be able to reach Microsoft upload endpoints, including *.blob.core.windows.net. Outbound HTTPS traffic to Microsoft services must not be blocked by firewalls, proxies, or security policies, and any proxy configuration must allow outbound HTTPS traffic without interception or modification that would prevent uploads.

If you are using Windows Server, you must enable endpoint DLP for Windows Servers before turning on Always-on diagnostics, because endpoint DLP is disabled by default on Windows Server.

If you are using macOS, the device must be onboarded to Microsoft Purview through Intune, JAMF Pro, or another MDM solution. For more information, see Onboard macOS devices into Microsoft Purview.

Supported Windows operating systems

The following table lists the supported Windows client and server versions for Always-on diagnostics.

Expand table

OS	Version	Minimum Build
Windows 11	24H2	Build 26100.4202
Windows 11	23H2	Build 22621.5039 and 22631.5039
Windows 11	22H2	Build 22621.5039 and 22631.5039
Windows 10	22H2	Build 19045.5917
Windows 10	21H2	Build 19045.5917
Windows Server 2019	-	Build 17763.7434
Windows Server 2022	-	Build 20348.3807
Windows Server 2025	-	Build 26100.4349

> **Important**
>
> Endpoint DLP is disabled by default on Windows Server. You must enable endpoint DLP for Windows Servers to support always-on diagnostics.

Supported macOS versions

Always-on diagnostics is supported on onboarded macOS devices running any of the three latest released versions. Always-on diagnostics requires build 101.26042.xxxx or later.

Supported processors

macOS devices with x64 and M1, M2, and M3 (ARM64) processors are supported.

Turn on Always-on diagnostics and enable upload

Perform the following steps to turn on Always-on diagnostics and enable log upload. These settings apply to both Windows and macOS devices.

1. Sign in to the Microsoft Purview portal.
2. Navigate to **Settings** > **Data Loss Prevention** > **Always-on diagnostics (preview)**.
3. Select **On**.
4. Set the cache storage period. A 90-day cache storage period is recommended.
5. Set the maximum storage for the device. The value must be between 500 and 1,500 MB.
6. Select **Save**.
7. To enable upload, under **Automatically upload device logs** select **Share diagnostics with Microsoft**.

Request device logs

When you identify an issue and open a support case with Microsoft, you can request that the log files be sent to Microsoft Support. The procedure is the same for Windows and macOS devices.

1. In Purview, select one of the locations to initiate a request for log files.
   a. From **Settings** > **Device onboarding** > **Devices**, select a device from the list.
   b. From **Data Loss Prevention** > **Alerts** > **Events**, select an event from the list.
   c. From **Data Loss Prevention** > **Explorers** > **Activity explorer**, select an alert from the list.
2. Based on the location you chose, under **Always-on Diagnostics** select **Request device log**.

3. Select the date range and provide a short description.
4. Select **Submit collection request**.
5. After you submit the request, you must wait until the request is complete. Navigate to **Settings** > **Data Loss Prevention** > **Always-on diagnostics (preview)**.
6. From the list, identify the device under investigation. When the status is complete, provide the associated request number to Microsoft Support.

See also

Self-help diagnostics for Microsoft Purview

Collect endpoint DLP diagnostic logs

Analyze endpoint DLP diagnostic logs

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 09/02/2026

Learn about the Microsoft Purview extension for Chrome

10/07/2025

Endpoint data loss prevention (endpoint DLP) extends the activity monitoring and protection capabilities of Microsoft Purview Data Loss Prevention (DLP) to sensitive items on Windows 10/11 devices. Once you onboard devices into the Microsoft Purview solutions, information about what users are doing with sensitive items is made visible in activity explorer. Then you can enforce protective actions on those items using data loss prevention policies.

After you install the Microsoft Purview extension for Chrome on a Windows 10/11 device, your organization can monitor attempts to access or upload a sensitive item to a cloud service in Google Chrome, and enforce protective actions via DLP.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Activities you can monitor and take action on

The extension enables you to audit and manage the following types of activities users take on sensitive items on devices running Windows 10/11.

Expand table

activity	description	supported policy actions
file copied to cloud	Detects when a user attempts to upload a sensitive item to a restricted service domain through the Chrome browser	audit, block with override, block
file printed	Detects when a user attempts to print a sensitive item from the Chrome browser to a local or network printer	audit, block with override, block
file copied to clipboard	Detects when a user attempts to copy & paste information from a sensitive item in the Chrome browser into another app, process, or item.	audit, block with override, block
file copied to removable storage	Detects when a user attempts to copy a sensitive item or information from the Chrome browser to removable media or	audit, block with override, block

activity	description	supported policy actions
	USB device	
file copied to network share	Detects when a user attempts to copy a sensitive item or information from the Chrome browser to a network share or mapped network drive.	audit, block with override, block

> **Note**
>
> Microsoft Purview Insider Risk Management requires that this extension be installed to monitor and take action on activities in the Chrome browser. For more information, see Learn about and configure Insider Risk Management browser signal detection.

Purview Chrome extension behavior

The DLPEngine caches all files for about 15 minutes by default to prevent repeated upload attempts. During the cache period, the file is blocked from being uploaded to any domain, so if you upload to an unallowed domain and then immediately to an allowed domain, the cached response may block the upload.

> **Tip**
>
> Microsoft recommends that you use Microsoft Edge if you need to upload files in quick succession.

Deployment process

1. Get started with endpoint data loss prevention
2. Onboarding tools and methods for Windows 10/11 devices
3. Install the extension for Chrome on your Windows 10/11 devices

   > **Note**
   >
   > Command access for the user is a prerequisite for the Microsoft Purview Extension extension for Chrome to work.

4. Create and Deploy data loss prevention policies that restrict upload to cloud service, or access by unallowed browsers actions and apply them to your Windows 10/11 devices

Next steps

See Get started with the Microsoft Purview extension for Chrome for complete deployment procedures and scenarios.

See also

- Get started with Microsoft Purview extension for Chrome
- Learn about Endpoint data loss prevention
- Getting started with Endpoint data loss prevention
- Using Endpoint data loss prevention
- Learn about data loss prevention
- Create and Deploy data loss prevention policies
- Get started with Activity explorer
- Microsoft Defender for Endpoint
- Insider risk management

Get started with the Microsoft Purview extension for Chrome

Use these procedures to roll out the Microsoft Purview extension for Chrome.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about **Microsoft Security Copilot in Microsoft Purview**.

> **Note**
>
> The Microsoft Purview extension for Chrome only works on Windows devices. You don't need the extension to enforce data loss prevention on macOS devices.

Before you begin

To use the Microsoft Purview extension for Chrome, you must onboard the device into Endpoint data loss prevention (DLP). Review these articles if you're new to DLP or Endpoint DLP:

- Learn about the Microsoft Purview extension for Chrome
- Learn about Microsoft Purview Data Loss Prevention
- Create and Deploy data loss prevention policies
- Learn about Endpoint data loss prevention
- Get started with Endpoint data loss prevention
- Onboarding tools and methods for Windows 10/11 devices
- Configure device proxy and internet connection settings for Microsoft Purview Endpoint DLP
- Using Endpoint data loss prevention

Licensing

For information on licensing, see

- Microsoft 365 Enterprise Plans
- Microsoft 365 Service Descriptions

- Your organization must be licensed for Endpoint DLP.
- Your devices must run Windows 10 x64 (build 1809 or later).
- The device must have Antimalware Client version 4.18.2202.x or later. Check your current version by opening the **Windows Security** app, select the **Settings** icon, and then select **About**.

Permissions

You can view data from Endpoint DLP in Activity explorer. Seven roles grant permission to view and interact with the activity explorer. The account you use to access the data must be a member of at least one of these roles.

- Global administrator
- Compliance admin
- Security admin
- Compliance data admin
- Global reader
- Security reader
- Reports reader

> ⓘ **Important**
>
> Use roles with the fewest permissions to improve security for your organization. Global Administrator is a highly privileged role that you should only use in scenarios where a lesser privileged role can't be used.

Roles and role groups

You can use roles and role groups to fine tune your access controls.

Here's a list of applicable roles. To learn more about them, see Permissions in the Microsoft Purview portal.

- Information Protection Admin
- Information Protection Analyst
- Information Protection Investigator
- Information Protection Reader

Here's a list of applicable role groups. To learn more about them, see Permissions in the Microsoft Purview portal.

- Information Protection
- Information Protection Admins
- Information Protection Analysts
- Information Protection Investigators
- Information Protection Readers

Overall installation workflow

Deploying the extension is a multistage process. You can install it on one machine at a time, or you can use Microsoft Intune or Group Policy for organization-wide deployments.

1. Prepare your devices.
2. Basic Setup Single Machine Selfhost
3. Deploy using Microsoft Intune
4. Deploy using Group Policy
5. Test the extension
6. Use the Alerts Management Dashboard to view Chrome DLP alerts
7. Viewing Chrome DLP data in activity explorer

Prepare infrastructure

If you're rolling out the extension to all your monitored Windows 10/11 devices, you should remove Google Chrome from the unallowed app and unallowed browser lists. For more information, see Unallowed browsers. If you're only rolling it out to a few devices, you can leave Chrome on the unallowed browser or unallowed app lists. The extension bypasses the restrictions of both lists for those computers where it's installed.

> **Important**
>
> If you configure a Chrome NativeMessageBlocklist, you must configure a NativeMessageAllowlist that includes the Chrome extension.

Prepare your devices

1. Use the procedures in these articles to onboard your devices:

a. Get started with Endpoint data loss prevention

b. Onboarding Windows 10 and Windows 11 devices

c. Configure device proxy and internet connection settings for Microsoft Purview Endpoint DLP

> **Important**
>
> Microsoft Purview upgraded Purview Chrome extension to Manifest V3. If you already have Purview Chrome extension installed, you should see automatic upgrade on your machine to 3.0.0.239 or higher.

Set up the extension on a single self-hosted machine

Installing the extension from the Chrome Web Store is the recommended method for a single machine.

1. Go to Microsoft Purview Extension - Chrome Web Store (google.com).
2. Install the extension by following the instructions on the Chrome Web Store page.

Deploy using Microsoft Intune

Use the Microsoft Intune deployment method for organization-wide deployments.

Force-install the Chrome extension with Microsoft Intune

Using the settings catalog, follow these steps to manage Chrome extensions:

1. Sign in to the Microsoft Intune admin center.
2. Go to **Configuration Profiles**.
3. Select **Create Profile**.
4. Select **Windows 10 and later** as the platform.
5. Select **Settings catalog** as the profile type.
6. Select **Custom** as the template name.
7. Select **Create**.

8. Enter a name and optional description on the **Basics** tab and select **Next**.
9. Select **Add settings** on the **Configuration settings** tab.
10. Select **Google** > **Google Chrome** > **Extensions**.
11. Select **Configure the list of force-installed apps and extensions**.
12. Change the toggle to **Enabled**.
13. Enter the following value for the extensions and app IDs and update URL: `echcggldkblhodogklpincgchnpgcdco;https://clients2.google.com/service/update2/crx`.
14. Select **Next**.
15. Add or edit scope tags on the **Scope tags** tab as needed and select **Next**.
16. Add the required deployment users, devices, and groups on the **Assignments** tab and select **Next**.
17. Add applicability rules on the **Applicability Rules** tab as required and select **Next**.
18. Select **Create**.

Deploy by using Group Policy

If you don't want to use Microsoft Intune, you can use group policies to deploy the extension across your organization.

Add the Chrome extension to the ForceInstall list

Perform the following steps to add the extension to Chrome's force-install list through Group Policy:

1. In the Group Policy Management Editor, go to your OU.
2. Expand the following path: **Computer/User configuration** > **Policies** > **Administrative templates** > **Classic administrative templates** > **Google** > **Google Chrome** > **Extensions**. This path might vary depending on your configuration.
3. Select **Configure the list of force-installed extensions**.
4. Right-click and select **Edit**.
5. Select **Enabled**.

6. Select **Show**.
7. Under **Value**, add the following entry:
   `echcggldkblhodogklpincgchnpgcdco;https://clients2.google.com/service/update2/crx`
8. Select **OK** and then **Apply**.

Test the extension

Test uploads to cloud services and access from unallowed browsers

Use the following test to verify that Chrome DLP blocks uploads to restricted cloud destinations:

1. Create or get a sensitive item and try to upload a file to one of your organization's restricted service domains. The sensitive data must match one of our built-in Sensitive Info Types, or one of your organization's sensitive information types. You should get a DLP toast notification on the device you're testing from that shows that uploading the file isn't allowed when the file is open.

Simulate other DLP scenarios in Chrome

Now that you removed Chrome from the disallowed browsers/apps list, you can run simulation scenarios to confirm that the behavior meets your organization's requirements:

- Copy data from a sensitive item to another document using the Clipboard
  - To test, open a file that is protected against copy to clipboard actions in the Chrome browser and attempt to copy data from the file.
  - Expected result: A DLP toast notification showing that copying data to the clipboard isn't allowed when the file is open.
- Print a document
  - To test, open a file that is protected against print actions in the Chrome browser and attempt to print the file.
  - Expected result: A DLP toast notification showing that printing the file isn't allowed when the file is open.
- Copy to USB Removable Media
  - To test, try to save the file to a removable media storage.
  - Expected result: A DLP toast notification showing that saving the file to removable media isn't allowed when the file is open.
- Copy to Network Share
  - To test, try to save the file to a network share.

- Expected result: A DLP toast notification showing that saving the file to a network share isn't allowed when the file is open.

Use the Alerts Management Dashboard to view Chrome DLP alerts

Use the following steps to view Chrome DLP alerts in the Alerts dashboard:

1. Open the **Data loss prevention** page in the Microsoft Purview portal and select **Alerts**.
2. Refer to the procedures in Get started with the data loss prevention Alerts dashboard and Investigate data loss incidents with Microsoft Defender XDR to view alerts for your Endpoint DLP policies.

View Chrome DLP data in Activity Explorer

Use the following steps to view Chrome DLP activity in Activity explorer:

1. Open the Microsoft Purview portal > **Information protection** > **Explorers** > **Activity explorer**.
2. Follow the procedures in Get started with Activity explorer to access and filter all the data for your Endpoint devices.

Next steps

After you onboard devices and view the activity data in Activity explorer, create DLP policies that protect your sensitive items.

- Using Endpoint data loss prevention

See also

- Learn about Endpoint data loss prevention
- Using Endpoint data loss prevention
- Learn about data loss prevention
- Create and Deploy data loss prevention policies
- Get started with Activity explorer
- Microsoft Defender for Endpoint
- Onboarding tools and methods for Windows 10 machines
- Microsoft 365 subscription

- Microsoft Entra joined devices
- Download the new Microsoft Edge based on Chromium

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Learn about the Microsoft Purview extension for Firefox

Article • 08/21/2024

Endpoint data loss prevention (endpoint DLP) extends the activity monitoring and protection capabilities of Microsoft Purview data loss prevention (DLP) to sensitive items that are on Windows 10/11 devices. Once you onboard devices to the Microsoft Purview solutions, the information about what users are doing with sensitive items is made visible in activity explorer. Then you can enforce protective actions on those items via DLP policies.

After you install the Firefox extension on a Windows 10/11 device, your organization can monitor when a user attempts to access or upload a sensitive item to a cloud service using Mozilla Firefox, and enforce protective actions via DLP.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about **Microsoft Security Copilot in Microsoft Purview**.

Activities you can monitor and take action on

The extension enables you to audit and manage the following types of activities users take on sensitive items on devices running Mozilla Firefox in Windows 10.

Expand table

activity	description	supported policy actions
file copied to cloud	Detects when a user attempts to upload a sensitive item to a restricted service domain through the Firefox browser	audit, block with override, block
file printed	Detects when a user attempts to print a sensitive item that is open in the Firefox browser to a local or network printer	audit, block with override, block
file copied to clipboard	Detects when a user attempts to copy information from a sensitive item in the Firefox browser into another app,	audit, block with override, block

activity	description	supported policy actions
	process, or item.	
file copied to removable storage	Detects when a user attempts to copy a sensitive item or information from a sensitive item in the Firefox browser to removable media or USB device	audit, block with override, block
file copied to network share	Detects when a user attempts to copy a sensitive item or information from the Firefox browser to a network share or mapped network drive.	audit, block with override, block

Deployment process

1. Get started with endpoint data loss prevention
2. Onboarding tools and methods for Windows 10 devices
3. Install the extension for Firefox extension on your Windows 10 devices
4. Create and Deploy data loss prevention policies that restrict upload to cloud service, or access by unallowed browsers actions and apply them to your Windows 10 devices

Next steps

See Get started with the Microsoft Purview extension for Firefox for complete deployment procedures and scenarios.

See also

- Get started with Microsoft Purview extension for Firefox
- Learn about Endpoint data loss prevention
- Getting started with Endpoint data loss prevention
- Using Endpoint data loss prevention
- Learn about data loss prevention
- Create and Deploy data loss prevention policies
- Get started with Activity explorer
- Microsoft Defender for Endpoint
- Insider risk management

Feedback

Was this page helpful? Yes No

Get started with the Microsoft Purview extension for Firefox

Use these procedures to roll out the Microsoft Purview extension for Firefox.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about **Microsoft Security Copilot in Microsoft Purview**.

Recommended reading

To use the Microsoft Purview extension for Firefox, the device must be onboarded into endpoint DLP. Review these articles if you're new to DLP or endpoint DLP

If you're new to Microsoft Purview DLP, here's a list of the core articles you'll need as you implement DLP:

1. Administrative units
2. Learn about Microsoft Purview Data Loss Prevention - the article you're reading now introduces you to the data loss prevention discipline and Microsoft's implementation of DLP
3. Plan for data loss prevention (DLP) - by working through this article you will:
   a. Identify stakeholders
   b. Describe the categories of sensitive information to protect
   c. Set goals and strategy
4. Data Loss Prevention policy reference - this article introduces all the components of a DLP policy and how each one influences the behavior of a policy
5. Design a DLP policy - this article walks you through creating a policy intent statement and mapping it to a specific policy configuration.
6. Create and Deploy data loss prevention policies - This article presents some common policy intent scenarios that you'll map to configuration options, then it walks you through configuring those options.
7. Learn about investigating data loss prevention alerts - This article introduces you to the lifecycle of alerts from creation, through final remediation and policy tuning. It also introduces you to the tools you use to investigate alerts.

Additionally, you should be familiar with the information in these articles:

- Learn about Microsoft Purview extension for Firefox
- Learn about endpoint data loss prevention
- Get started with Endpoint data loss prevention
- Onboarding tools and methods for Windows 10 devices
- Configure device proxy and internet connection settings for Microsoft Purview Endpoint DLP
- Using Endpoint data loss prevention

Licensing and subscription requirements

For information on licensing, see

- Microsoft 365 Enterprise Plans
- Microsoft 365 Service Descriptions

- Your organization must be licensed for Endpoint DLP
- Your devices must be running Windows 10 x64 build 1809 or later.
- The device must have Antimalware Client Version is 4.18.2202.x or later. Check your current version by opening **Windows Security** app, select the **Settings** icon, and then select **About**.

Permissions

Data from Endpoint DLP can be viewed in Activity explorer. There are seven roles that grant permission to activity explorer, the account you use for accessing the data must be a member of any one of them.

- Compliance admin
- Security admin
- Compliance data admin
- Global reader
- Security reader
- Reports reader
- Global administrator

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. This helps improve security for your organization. Global Administrator is a highly privileged role that should

only be used in scenarios where a lesser privileged role can't be used.

Roles and Role Groups

There are roles and role groups that you can use to fine tune your access controls.

Here's a list of applicable roles. To learn more about them, see Permissions in the Microsoft Purview portal.

- Information Protection Admin
- Information Protection Analyst
- Information Protection Investigator
- Information Protection Reader

Here's a list of applicable role groups. To learn more about these role groups, see Permissions in the Microsoft Purview portal.

- Information Protection
- Information Protection Admins
- Information Protection Analysts
- Information Protection Investigators
- Information Protection Readers

Overall installation workflow

Deploying the extension is a multi-phase process. You can choose to install on one machine at a time, or use Microsoft Intune or Group Policy for organization-wide deployments.

1. Prepare your devices.
2. Basic Setup Single Machine Selfhost
3. Deploy using Microsoft Intune
4. Deploy using Group Policy
5. Test the extension
6. Use the Alerts Management Dashboard to view Firefox DLP alerts
7. Viewing Firefox DLP data in activity explorer

Prepare infrastructure

If you're rolling out the extension to all your monitored Windows 10 devices, you should remove Mozilla Firefox from the unallowed app and unallowed browser lists. For more information, see

Unallowed browsers. If you're only rolling it out to a few devices, you can leave Firefox on the unallowed browser or unallowed app lists. The extension bypasses the restrictions of both lists for those computers where it's installed.

Prepare your devices

1. Use the procedures in these articles to onboard your devices:
   a. Get started with Endpoint data loss prevention
   b. Onboarding Windows 10 and Windows 11 devices
   c. Configure device proxy and internet connection settings for Microsoft Purview Endpoint DLP

Set up a basic single-machine self-hosted deployment

The single-machine self-host installation is the recommended method.

1. Download the initial XPI file.
2. Locate the extension in your file explorer and drag the file into an open Mozilla Firefox window.
3. Confirm the installation.

Deploy using Microsoft Intune

Use the Microsoft Intune setup method for organization-wide deployments.

Configure force install with Microsoft Intune

Before adding the extension to the list of force-installed extensions, it's important to ingest the Firefox ADMX. The following steps explain how to ingest the Firefox ADMX into Microsoft Intune. Before beginning these steps, please ensure you have downloaded the latest Firefox ADMX from the Firefox GitHub.

The following steps can be followed to ingest Firefox ADMX.

1. Sign in to the Microsoft Intune Admin Center.
2. Navigate to **Devices** then to **Configuration**.
3. Select **Create New Policy**.

4. Select **Windows 10 and later** as the platform.
5. Select **Templates** and **Custom** as profile type then click **Create**.
6. Enter a descriptive name like *Firefox ADMX* and an optional description.
7. Click **Add** *OMA-URI Settings* and enter the following policy information.

   Name: Descriptive name.

   Description: Optional description

   OMA-URI:
   `./Device/Vendor/MSFT/Policy/ConfigOperations/ADMXInstall/Firefox/Policy/FirefoxAdmx`

   Data type: `String`

   Value: Copy all of the text from the downloaded *firefox.admx* file into the *Value* field
8. Select create.

After ingesting the Firefox ADMX, follow the Microsoft Intune procedure in this section to create a configuration profile for the Microsoft Purview extension for Firefox.

1. Sign in to the Microsoft Intune admin center.
2. Navigate to Configuration Profiles.
3. Select **Create Profile**.
4. Select **Windows 10** as the platform.
5. Select **Custom** as profile type.
6. Select the **Settings** tab.
7. Select **Add**.
8. Enter the following policy information.

   OMA-URI:
   `./Device/Vendor/MSFT/Policy/Config/Firefox~Policy~firefox~Extensions/ExtensionSettings`
   Data type: `String`
   Value: `<enabled/><data id="ExtensionSettings"`
   `value='{"microsoft.defender.browser_extension.native_message_host@microsoft.com":`

`{"installation_mode": "force_installed", "install_url": "https://github.com/microsoft/purview/raw/main/endpointDLP/browser_extension/prod-1.1.0.212.xpi","updates_disabled":false}}'/>`

9. Note: It's critical that `updates_disabled` is set to `false` so that the extension can automatically update over time.

10. Select create.

Deploy using Group Policy

If you don't want to use Microsoft Intune, you can use group policies to deploy the extension across your organization.

Add the Firefox extension to the force-install list

Use the following steps to add the Firefox extension to the force-install list in Group Policy.

1. In the Group Policy Management Editor, navigate to your organizational unit (OU).

2. Expand the following path **Computer/User configuration** > **Policies** > **Administrative templates** > **Classic administrative templates** > **Firefox** > **Extensions**. This path may vary depending on your configuration.

3. Select **Extensions to install**.

4. Right click and select **Edit**.

5. Select **Enabled**.

6. Select **Show**.

7. Under **Value**, add the following entry:
   `https://github.com/microsoft/purview/raw/main/endpointDLP/browser_extension/prod-1.1.0.212.xpi`

8. Select **OK** and then **Apply**.

Test the extension

Use the following procedures to validate that the Microsoft Purview extension for Firefox is working correctly in common DLP scenarios.

Test uploads to cloud services and access from blocked browsers

Use the following test to verify cloud egress blocking behavior in Firefox.

1. Create or get a sensitive item and, try to upload a file to one of your organization's restricted service domains. The sensitive data must match one of our built-in Sensitive Info Types, or one of your organization's sensitive information types. You should get a DLP toast notification on the device you're testing from that shows that this action isn't allowed when the file is open.

Testing other DLP scenarios in Firefox

Now that you've removed Firefox from the disallowed browsers/apps list, you can run simulations on the policy for these scenarios below to confirm the behavior meets your organization's requirements:

- Copy data from a sensitive item to another document using the Clipboard
  - To test, open a file that is protected against copy to clipboard actions in the Firefox browser and attempt to copy data from the file.
  - Expected Result: A DLP toast notification showing that this action isn't allowed when the file is open.
- Print a document
  - To test, open a file that is protected against print actions in the Firefox browser and attempt to print the file.
  - Expected Result: A DLP toast notification showing that this action isn't allowed when the file is open.
- Copy to USB Removable Media
  - To test, try to save the file to a removable media storage.
  - Expected Result: A DLP toast notification showing that this action isn't allowed when the file is open.
- Copy to Network Share
  - To test, try to save the file to a network share.
  - Expected Result: A DLP toast notification showing that this action isn't allowed when the file is open.

Use the Alerts Management Dashboard to view Firefox DLP alerts

Use the following steps to review Firefox DLP alerts in the Alerts dashboard.

1. Open the **Data loss prevention** page in the Microsoft Purview portal and select **Alerts**.
2. Refer to the procedures in Get started with the data loss prevention Alerts dashboard and Investigate data loss incidents with Microsoft Defender XDR to view alerts for your Endpoint DLP policies.

Use Activity Explorer to view Firefox DLP data

Use Activity explorer to review and filter Firefox DLP activity data for onboarded devices.

1. Open **Activity explorer**.
2. Refer to the procedures in Get started with Activity explorer to access and filter all the data for your Endpoint devices.

Known issues and limitations

Be aware of the following limitations when deploying the Microsoft Purview extension for Firefox.

1. Incognito mode isn't supported and must be disabled.

Next steps

After onboarding devices and verifying activity data in Activity explorer, create DLP policies that protect your sensitive items.

- Using Endpoint data loss prevention

See also

- Learn about Endpoint data loss prevention
- Using Endpoint data loss prevention
- Learn about data loss prevention
- Create and Deploy data loss prevention policies
- Get started with Activity explorer
- Microsoft Defender for Endpoint
- Onboarding tools and methods for Windows 10 machines
- Microsoft 365 subscription
- Microsoft Entra joined devices
- Download the new Microsoft Edge based on Chromium

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Learn about evidence collection for file activities on devices

When you're investigating a Microsoft Purview Data Loss Prevention (DLP) incident or troubleshooting a DLP policy, it can be helpful to have a complete copy of the item that matched the policy to refer to. DLP can copy the item that matches a DLP policy from onboarded Windows devices or macOS devices (preview) to an Azure storage account. DLP incident investigators and administrators that have been granted the appropriate permissions on the Azure storage blob can then access the files.

To get started configuring and using the feature, see Get started with collecting files that match data loss prevention policies from devices.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about **Microsoft Security Copilot in Microsoft Purview**.

If you're new to Microsoft Purview DLP, here's a list of the core articles you need as you implement DLP:

1. Administrative units
2. Learn about Microsoft Purview Data Loss Prevention - This article introduces you to the data loss prevention discipline and Microsoft's implementation of DLP.
3. Plan for data loss prevention (DLP) - by working through this article you will:
   a. Identify stakeholders
   b. Describe the categories of sensitive information to protect
   c. Set goals and strategy
4. Data Loss Prevention policy reference - This article introduces all the components of a DLP policy and how each one influences the behavior of a policy.
5. Design a DLP policy - This article walks you through creating a policy intent statement and mapping it to a specific policy configuration.
6. Create and Deploy data loss prevention policies - This article presents some common policy intent scenarios that you'll map to configuration options, then it walks you through configuring those options.

Where evidence collection for file activities on devices fits in Purview

Endpoint DLP is part of the larger DLP offering and part of the larger range of services offered in Microsoft Purview. You should understand how evidence collection for file activities on devices fits into the larger set of service offerings.

Evidence collection for file activities on devices and eDiscovery

This feature makes copies of items that match DLP policies on onboarded Windows devices or macOS devices (preview) and places those copies in an Azure storage account. These copies aren't held in a changeless state and aren't evidence in the legal sense of the term. If you need to find and hold items for legal purposes, you should use the Microsoft Purview eDiscovery solutions. Electronic discovery, or eDiscovery, is the process of identifying and delivering electronic information that can be used as evidence in legal cases.

Evidence collection for file activities on devices and contextual summary

When an item and the activity that a user takes on that item match the conditions defined in a DLP policy, a **DLPRuleMatch** event shows up in Activity explorer. This is true for every location that DLP supports. The **DLPRuleMatch** event contains a limited amount of the text that surrounds the matched content. This limited amount of text is called contextual summary.

It's important to understand the difference between evidence collection for file activities on devices and a contextual summary. Evidence collection for file activities on devices are available for onboarded Windows devices or macOS devices (preview). It saves a copy of the entire item that matched a policy to the Azure storage account. A contextual summary is captured for every DLP policy rule match and only contains a limited amount of the text that surrounds the target text that triggered the match.

Evidence collection for file activities on devices and Data Security Investigations

Microsoft Purview Data Security Investigations can query endpoint DLP evidence collection data as an additional data source for security investigations. Investigators can use the endpoint DLP evidence collection (preview) search option in Data Security Investigations to aggregate endpoint events and analyze them collectively using AI-powered tools, rather than triaging individual alerts

one by one. Data Security Investigations queries whatever data is already collected by your endpoint DLP policies — no additional configuration is required beyond enabling evidence collection.

Covered user activities

You can configure evidence collection for file activities on devices to save a copy of a matched item to the Azure storage account when a user attempts to do one of these activities on a matched item:

- Copy to a removable USB
- Copy to network share
- Print
- Copy or move using unallowed Bluetooth app
- Copy or move through RDP
- Upload to cloud service domains or access from an unallowed browser
- Paste to supported browsers
- Access by restricted apps

The detection of these activities is configured in the DLP policy. For more information on how to create a DLP policy, see, Create and Deploy data loss prevention policies and Using Endpoint data loss prevention.

Covered actions

When you enable evidence collection for file activities on devices in Endpoint DLP settings and configure a DLP policy to use this feature, it saves a copy of a matched item for these actions:

- Audit only
- Block with override
- Block

These actions are configured in the DLP policy. For more information on how to create a DLP policy, see Create and Deploy data loss prevention policies and Using Endpoint data loss prevention.

Design considerations

Regions for your Azure Storage accounts

To comply with regulatory requirements, make sure that the Azure storage accounts that you use are in the same geopolitical or regulatory boundaries as the devices that they're being copied from. Also, be aware of the geopolitical location of the DLP investigators who will access the sensitive items once they're saved. Consider using Administrative units to scope the administration of the users and devices appropriately for each DLP policy. To learn how to use data loss prevention to comply with data privacy regulations, see Deploy information protection for data privacy regulations with Microsoft Purview. Evidence collection for file activities on devices supports up to 10 Azure storage accounts.

To learn how to use data loss prevention to comply with data privacy regulations, see Deploy information protection for data privacy regulations with Microsoft Purview.

Local storage and bandwidth

By default, copies of matched items are saved asynchronously to the configured Azure storage account over the existing network connection. If the device doesn't have connectivity, matched items are save locally, up to the 500-MB limit. You can save items locally up to 60 days.

While the device has connectivity to the Azure storage account URL, there's no limit on bandwidth usage. The bandwidth that evidence collection for file activities on devices uses doesn't affect the default or configured bandwidth limits for Advanced classification scanning and protection.

Azure storage accounts

Customers are responsible for creating and managing their own Azure storage accounts. If you're new to Azure storage, see:

- What is Azure Blob storage
- Introduction to Azure Storage
- Create a storage account

Items that match a policy are copied from the users' device to the Azure storage account blob in the security context of the logged in user. So, all users who are in-scope for the policy must have read and write permission to the blob storage. For more information, see Get started with collecting files that match data loss prevention policies from devices

Similarly, all administrators who are reviewing the saved items must have `read` permission to the Azure storage account blob. For more information, see Get started with collecting files that match data loss prevention policies from devices.

Storing evidence when sensitive information is detected

Supported file types

For more information on supported file types, see Supported file types for storing and previewing evidence.

Supported storage types

You have two options for storing the evidence Purview collects when it detects sensitive information in your content. You can use a **Customer-managed** data store, or a **Microsoft Managed** data store. The option you should use depends on your requirements and your use cases. To help you decide, review the comparison table that follows.

Storage type comparison

Matched files continue to be included in alert results even after changing your storage type as long as the role-based access control (RBAC) permissions remain intact. Since customer managed storage is owned by customers, DLP admins can continue to download files directly from the storage on a per-file basis.

The following table identifies the differences between customer-managed storage and Microsoft-managed storage for collecting evidence of the sensitive information detected in your content.

Expand table

Feature Element	Customer Managed	Microsoft Managed
File retention	You can keep files as long as you need/want to.	Files are retained for a maximum of 180 days.
Upload limit	You can upload files based on your blob storage size settings.	You can upload upto 5 GB per day at the overall tenant level.
Endpoint settings	You must add blob storage (container URLs) in the endpoint settings and then use the Microsoft Entra admin center to configure explicit user permissions on the blob for in-scope users.	All configuration and permissions are handled with a single click when configuring your endpoint settings.
Policy and location	You must add and configure storage blobs on a per-policy	No storage selection is needed for specific policy locations.

Feature Element	Customer Managed	Microsoft Managed
configuration	basis for each location where a policy is applied.	
Access to storage	Users with relevant permissions to storage blob can access the evidence	No human access to storage. Evidence is uploaded and rendered by the system.
Data store location/region	Chosen by the customer	The same region as your Microsoft Purview tenant.
Charges	Storage costs are charged in addition to the cost of your Entra subscription.	Storage cost today is included in E5. However, Microsoft will monitor storage usage and may charge additionally based on excessive usage. This will be communicated to customers separately should there be a change in the business model.
Network configuration	You must allow the container URLs for your storage blobs to pass through your network firewall.	You must include **compliancedrive.microsoft.com** on an "allow" list, so that it can pass through the network firewall.
Retention	Customers can selectively archive and purge to free up storage)	30 days on Purview and 180 days on Defender XDR
Preview and Download Permissions	Read access on storage blobs, RBAC roles **Data Classification Content Viewer** to preview the evidence and **Data Classification Content Download** to download the evidence	RBAC roles **Data Classification Content Viewer** to preview the evidence and **Data Classification Content Download** to download the evidence

Changing storage types

Customers can switch between storage types at any time. However, best practice is to carefully plan for the type of storage you'll need over the long term and select the appropriate option for your use case. For more information on the differences between the two types of storage, see the Storage type comparison table.

> ⓘ **Note**
>
> When switching storage types, you'll need to refresh your policies to ensure that they are applied to the files in the new data store.

Impact of changing storage types on evidence files

Matched files continue to be included in alert results, even after you change the storage management type, provided that the role-based access control (RBAC) permissions do not change.

Because you own your customer-managed storage solution, your DLP admins can continue to download files directly on a per-file basis after they have been moved to the Microsoft-managed storage solution.

Next step

Your next step is to configure evidence collection for file activities on devices.

For more information, see Get started with collecting files that match data loss prevention policies from devices.

Last updated on 06/22/2026

Get started with collecting files that match data loss prevention policies from devices

This article walks you through the prerequisites and configuration steps for evidence collection for file activities on devices and introduces how to view the items that are copied and saved.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about **Microsoft Security Copilot in Microsoft Purview**.

Here are the high level steps for configuring and using evidence collection for file activities on devices.

1. Onboard devices
2. Understand your requirements
3. Create your managed Azure storage account
4. Add an Azure storage blob to your account
5. Enable and configure evidence collection on a storage account managed by Microsoft
6. Configure your DLP policy
7. View usage of Microsoft managed storage

Before you begin

Before you start these procedures, you should review Learn about evidence collection for file activities on devices.

Licensing requirements for evidence collection

For information on licensing, see

- Microsoft 365 Enterprise Plans
- Microsoft 365 Service Descriptions

See the prerequisites licensing requirements for Microsoft Entra ID P1 or P2 needed to create custom role-based access control (RBAC).

Permissions

Standard Microsoft Purview Data Loss Prevention (DLP) permissions are required. For more information, see Permissions.

Permissions for preview or download of evidence

When you want to preview or download the evidence, the account used needs to be in these roles:

- Preview the evidence: Data Classification Content Viewer
- Download the evidence: Data Classification Content Download

The Data Classification Content Download role is preassigned to these built-in role groups:

- Data Security Management
- Information Protection
- Information Protection Investigators

If you need to manage users, roles and role groups, see Permissions in the Microsoft Purview portal.

Onboard devices

Before you can use copy matched items you have to onboard Windows 10/11 devices or macOS devices (Preview) into Purview, see Onboard Windows devices into Microsoft 365 overview.

Understand your requirements

> **Important**
>
> Each container inherits the permissions of the storage account that it is in. You can't set different permissions per container. If you need to configure different permissions for different regions, you must create multiple storage accounts, not multiple containers.

You should have answers to the following question before setting up your Azure storage and scoping the feature to users.

Do you need to compartmentalize items and access along role or departmental lines?

For example, if your organization wants to have one set of administrators or DLP event investigators who can view saved files from your senior leadership and another set of administrators or DLP event investigators for saved items from human resources, you should create one Azure storage account for your organization's senior leadership and another for the Human Resources department. This ensures that the Azure storage admins or DLP event investigators can only see the items that matched DLP policies from their respective groups.

Do you want to use containers to organize saved items?

You can create multiple evidence containers within the same storage account for sorting saved files. For example, one for files saved from the HR department and another for those from the IT department.

What is your strategy for protecting against saved item deletion or modification?

In Azure Storage, data protection refers to both the strategies for protecting the storage account and data within it from being deleted or modified and to restoring data after it's deleted or modified. Azure storage also offers options for disaster recovery, including multiple levels of redundancy, to protect your data from service outages due to hardware problems or natural disasters. It can also protect your data using customer-managed failover if the data center in the primary region becomes unavailable. For more information, see Data protection overview.

You can also configure immutability policies for your blob data that protects against the saved items being overwritten or deleted. For more information, see Store business-critical blob data with immutable storage

Supported file types for storing and previewing evidence

The following table shows which file types can be stored and previewed as evidence.

Expand table

Can be Stored	Can be Previewed
All file types provided they match the policy	All file types supported for previewing files in OneDrive, SharePoint, and Teams

> ⓘ Note
>
> To understand how to create policies for different file types see Create and deploy DLP policies

Save matched items to your preferred storage

To save the evidence that Microsoft Purview detects when your data loss prevention policies are applied, you need to set up storage. There are two ways to set up storage for evidence collection:

1. Create customer-managed storage
2. Create Microsoft-managed storage

For more information and a comparison of these two types of storage, see Storing evidence when sensitive information is detected.

Create customer-managed storage

The procedures for setting up your Azure storage account, container, and blobs are documented in the Azure Blob Storage documentation. Here are links to relevant articles you can refer to help you get started:

1. Introduction to Azure Blob Storage
2. Create a storage account
3. Default to and authorize access to blobs using Microsoft Entra ID
4. Manage blob containers using the Azure portal
5. Manage block blobs with PowerShell

> ⓘ Note
>
> Make sure to select **Enable public access from all networks while creating the storage account**. Support for Virtual networks and IP addresses and use of private access isn't available

Be sure to save the name and URL of the Azure blob container. To view the URL, open the Azure storage portal > **Home** > **Storage Accounts** > **Container** > **Properties**

The format for the Azure blob container URL is: `https://storageAccountName.blob.core.windows.net/containerName`.

Add an Azure storage blob to your account

There are several ways you can add an Azure storage blob to your account. Chose *one* of the methods below.

1. Sign in to the Microsoft Purview portal and choose the **Settings** gear in the menu bar.
2. Choose **Data Loss Prevention**.
3. Select **Endpoint DLP settings**.
4. Expand **Setup evidence collection for file activities on devices**.
5. Change the toggle from **Off** to **On**.
6. In the **Set evidence cache on device** field, select the amount of time evidence should be saved locally when the device is offline. You can choose **7**, **30**, or **60** days.
7. Select a **storage type** (*Customer managed store* or *Microsoft managed store*) and then select **+ Add storage**.
    a. For **Customer-managed storage**:
        i. Choose **Customer managed store**: and then choose **+ Add storage**.
        ii. Enter give the account a name and enter the **URL** for the storage blob.
        iii. Choose **Save**.
    b. For **Microsoft-managed storage**:
        i. Choose **Microsoft managed store**

To add Azure blob storage using the DLP policy creation workflow in the Microsoft Purview portal:

1. Sign in to the Microsoft Purview portal and choose **Data Loss Prevention**.
2. Follow the steps to create a new policy.
3. At the **Locations** step, make sure that only the **Devices** location is selected, and then choose **Next**.
4. At the Policy settings step, select Create or customize advanced DLP rules.
5. Select **+Create rule** and add **Conditions**, **Actions**, **User notifications** and **User overrides** with values of your choosing.
6. In the **Incident reports** section, the **Send an alert to admins when a rule match occurs** toggle should be **On** by default. (If it isn't, turn it on.)
7. Select the checkbox next to **Collect original file as evidence for all selected file activities on Endpoint**.
8. Select the activities you want to collect evidence for.
9. Select **Add storage** next to the checkbox item.

10. On the **Endpoint DLP settings** page, expand **Setup evidence collection for file activities on devices** and make sure that the option is toggled to **On**.

Set permissions on the Azure blob storage

Using Microsoft Entra authorization, you must configure two sets of permissions (role groups) on the blobs:

1. One for the administrators and investigators so they can view and manage evidence
2. One for users who need to upload items to Azure from their devices

Best practice is to enforce *least privilege* for all users, regardless of role. By enforcing least privilege, you ensure that user permissions are limited to only those permissions necessary for their role. To configure user permissions, create roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview.

Permissions on Azure blob for administrators and investigators

Once you create the role group for DLP incident investigators, you must configure the permissions described in the Investigator actions and Investigator data actions sections that follow.

For more information on configuring blob access, see these articles:

- How to authorize access to blob data in the Azure portal
- Assign share-level permissions.

Investigator actions

Configure these object and action permissions for the investigator role:

Expand table

Object	Permissions
Microsoft.Storage/storageAccounts/blobServices	Read: List Blob Services
Microsoft.Storage/storageAccounts/blobServices	Read: Get blob service properties or statistics
Microsoft.Storage/storageAccounts/blobServices/containers	Read: Get blob container

Object	Permissions
Microsoft.Storage/storageAccounts/blobServices/containers	Read: List of blob containers
Microsoft.Storage/storageAccounts/blobServices/containers/blobs	Read: Read blob
Microsoft.Storage/storageAccounts/blobServices/generateUserDelegationKey/action	Other: Generate a user delegation key

Investigator data actions

Assign the following data actions to the investigator role.

Expand table

Object	Permissions
Microsoft.Storage/storageAccounts/blobServices/containers/blobs	Read: Read Blob

Your JSON for the investigator role group should look like this:

JSON

```
"permissions": [
            {
                "actions": [
                 "Microsoft.Storage/storageAccounts/blobServices/read",
                 "Microsoft.Storage/storageAccounts/blobServices/containers/read",

"Microsoft.Storage/storageAccounts/blobServices/generateUserDelegationKey/action"
                ],
                "notActions": [],
                "dataActions": [

"Microsoft.Storage/storageAccounts/blobServices/containers/blobs/read"
                ],
                "notDataActions": []
            }
        ]
```

Permissions on Azure blob for users

Assign these object and action permissions to the Azure blob for the user role:

User actions

Assign the following object and action permissions to the user role.

Expand table

Object	Permissions
Microsoft.Storage/storageAccounts/blobServices	Read: List Blob Services
Microsoft.Storage/storageAccounts/blobServices/containers	Read: Get blob container
Microsoft.Storage/storageAccounts/blobServices/containers	Write: Put blob container

User data actions

Assign the following data actions to the user role.

Expand table

Object	Permissions
Microsoft.Storage/storageAccounts/blobServices/containers/blobs	Write: Write Blob
Microsoft.Storage/storageAccounts/blobServices/containers/blobs	Other: Add blob content

The JSON for user role group should look like this:

JSON

```
"permissions": [
  {
     "actions": [
       "Microsoft.Storage/storageAccounts/blobServices/containers/read",
       "Microsoft.Storage/storageAccounts/blobServices/containers/write",
       "Microsoft.Storage/storageAccounts/blobServices/read"
     ],
     "notActions": [],
     "dataActions": [
         "Microsoft.Storage/storageAccounts/blobServices/containers/blobs/add/action",
         "Microsoft.Storage/storageAccounts/blobServices/containers/blobs/write"
     ],
     "notDataActions": []
    }
 ]
```

Enable and configure evidence collection on a storage account managed by Microsoft

To enable and configure evidence collection on a storage account managed by Microsoft from within the Microsoft Purview portal:

1. Sign in to the Microsoft Purview portal > **Settings** gear in the menu bar.
2. Choose **Data Loss Prevention**.
3. Select **Endpoint DLP settings**.
4. Expand **Setup evidence collection for file activities on devices** and set the toggle to **On**.
5. Under **Select storage type**, choose **Microsoft managed storage**.

> **Tip**
>
> After you configure Microsoft-managed storage, you can use **Data Security Investigations** to query and analyze evidence collected from endpoint DLP policies across multiple users and time periods.

Configure your DLP Policy

Create a DLP policy as you normally would. For policy configuration examples, see Create and Deploy data loss prevention policies.

Configure your policy using these settings:

- Make sure that **Devices** is the only location selected.
- In **Incident reports**, toggle **Send an alert to admins when a rule match occurs** to **On**.
- In **Incident reports**, select **Collect original file as evidence for all selected file activities on Endpoint**.
- Select the storage account you want.
- Select the activities for which you want to copy matched items to Azure storage, such as:
  - **Access by restricted apps (preview)**
  - **Paste to supported browsers**
  - **Upload to cloud service domains or access unallowed browsers**
  - **Copy to a removable USB device**
  - **Copy to a network share**
  - **Print**
  - **Copy or move using an unallowed Bluetooth app**
  - **Copy or move using RDP**

View usage of Microsoft Managed Storage

You can view and export the daily usage for past 6 months for Microsoft Managed storage.

1. Sign in to the Microsoft Purview portal > **Settings** gear in the menu bar.

2. Choose **Data Loss Prevention**.
3. Select **Endpoint DLP settings**.
4. Expand **Setup evidence collection for file activities on devices**.
5. Click on **View Usage** button, a panel will open. Here you can see month wise graph of daily usage and also export the usage data.

Preview and Download the evidence

There are different ways to preview your evidence, depending on which type of storage you select.

Expand table

Storage Type	Preview options
Customer-managed	- Use activity explorer - Use the Microsoft Purview portal
Microsoft Managed	- Use activity explorer - Use the Purview portal

> **Note**
>
> Reminder, the account you use to preview and download needs to be in these roles:
>
> - Preview the evidence: Data Classification Content Viewer
> - Download the evidence: Data Classification Content Download

Preview evidence via Activity explorer

Use the following steps to preview evidence in Activity explorer.

1. Sign in to the Microsoft Purview portal > **Data loss prevention** > **Activity explorer**.
2. Using the **Date** dropdown, select the **Start** and **End** dates for the period you're interested in.
3. In the list of results, double-click the line item of the activity you want to investigate.
4. In the flyout pane, the link to the Azure blob where the evidence is stored appears under **Evidence file**.
5. Select the Azure blob storage link to display the file that was matched.

Preview evidence via the Purview portal Alerts page

Use the following steps to preview evidence from the Alerts page in the Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal > **Data loss prevention** > **Alerts**.
2. Using the **Date** dropdown, select the **Start** and **End** dates for the period you're interested in.
3. In the list of results, double-click the line item of the activity you want to investigate.
4. In the flyout pane, select **View details**.
5. Select the **Events** tab.
6. In the **Detail** pane, select the **Source tab**. The file that was matched displays.

> **Note**
>
> If the file that was matched already exists in the Azure storage blob, it won't be uploaded again until changes are made to the file and a user takes an action on it.

Known issues

- The maximum size for files that can be uploaded from a device is 500 MB.
- If Just-in-Time Protection is triggered on a scanned file, OR if the file is stored on a network share, the evidence file is not collected.
- When multiple files are opened in the same process (non-office apps) and one of the files matching a policy is egressed, DLP events are triggered for all files. No evidence is captured.
- If multiple policy rules are detected in a single file, the evidence file is only stored if the most restrictive policy rule is configured to collect evidence.

Note: The author created this article with assistance from AI. Learn more

Last updated on 06/26/2026

Get started with data loss prevention protections for Recall

> **Note**
>
> Microsoft Purview Data Loss Prevention (DLP) protection for Recall snapshots feature is *in preview*

Recall is a Copilot+ PC feature that allows users to search locally saved and locally analyzed snapshots of their screen using natural language. To help protect against the inclusion of sensitive content in these snapshots, you can use DLP policies. The following sections walk you through the prerequisites and configuration steps you perform to start using DLP protections for Recall.

DLP protections for Recall are supported for sensitive information types and sensitivity labeled items in:

- Sensitivity labeled Teams channels
- Sensitivity labeled Teams meeting chats
- SIT or sensitivity labeled files that are opened with Microsoft 365 Copilot App (Word, Excel PowerPoint web apps) using Microsoft Edge for Business
- Labeled emails Outlook
- Locally stored files with sensitivity labels or SITs
- Files stored in the cloud that with sensitivity labels or SITS that get opened in Office apps (Word, Excel, PowerPoint) on the device

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about **Microsoft Security Copilot in Microsoft Purview**.

Before you begin

IT admins can set policies that give Copilot+ PC users the option to enable saving snapshots. The user must then opt in to taking snapshots. For more information, see Manage Recall.

Prerequisites

Before you configure DLP protections for Recall, ensure the following prerequisites are met:

- meet all device prerequisites for Recall as described in Manage Recall.
- Microsoft Purview endpoint DLP is enabled for the tenant the Copilot+ PC is onboarded to endpoint data loss prevention.
- Microsoft Intune is running on the tenant to create Windows tenant policies.
- Copilot+ PC must be running
  - Windows build greater than 262100.67225
  - Anti-malware Client version greater than 4.18.2508.5 or higher
  - Teams version 25290.205.4069.4 or higher
  - Office desktop apps (Word, Excel, PowerPoint) version 19127 or higher
  - Microsoft Edge for Business. For more information, see Automatic activation of your Microsoft Purview DLP policy in Microsoft Edge

Set the following Group Policy settings on the Copilot+ PC to enable DLP protections for Recall using either Microsoft Intune or Local Group Policy Editor.

Expand table

Local Group Policy	Set to	Required	Notes
WindowsAI > Turn off saving snapshots for use with Recall	Disable	Yes	This policy makes the Recall feature available to users to opt in. This doesn't automatically start Recall for users. Users must launch recall and opt-in to taking snapshots on their machines.
WindowsAI > Set maximum duration for storing snapshots used by Recall	30-180	Optional	This is the retention period for snapshots in days

Local Group Policy	Set to	Required	Notes
WindowsAI/Set maximum storage for snapshots used by Recall	10-150	Optional	This is the storage limit. For more information, see, Manage Recall.
WindowsAI > SetDataLossPreventionProvider policy	HKEY_LOCAL_MACHINE\software\microsoft\windows defender; value:InstallLocation; binary:endpointdlp.dll; minversion:4.18.25080.5	Yes	This registry value sets the DLP provider (Microsoft Purview) to be used with Recall. This capability is implemented as a public API.

For more information, see: SetDataLossPreventionProvide

Licensing

For information on licensing, see

- Microsoft 365 Enterprise Plans
- Microsoft 365 Service Descriptions

Permissions

For information on the permissions required to create and manage DLP policies, see Permissions for creating and deploying DLP policies.

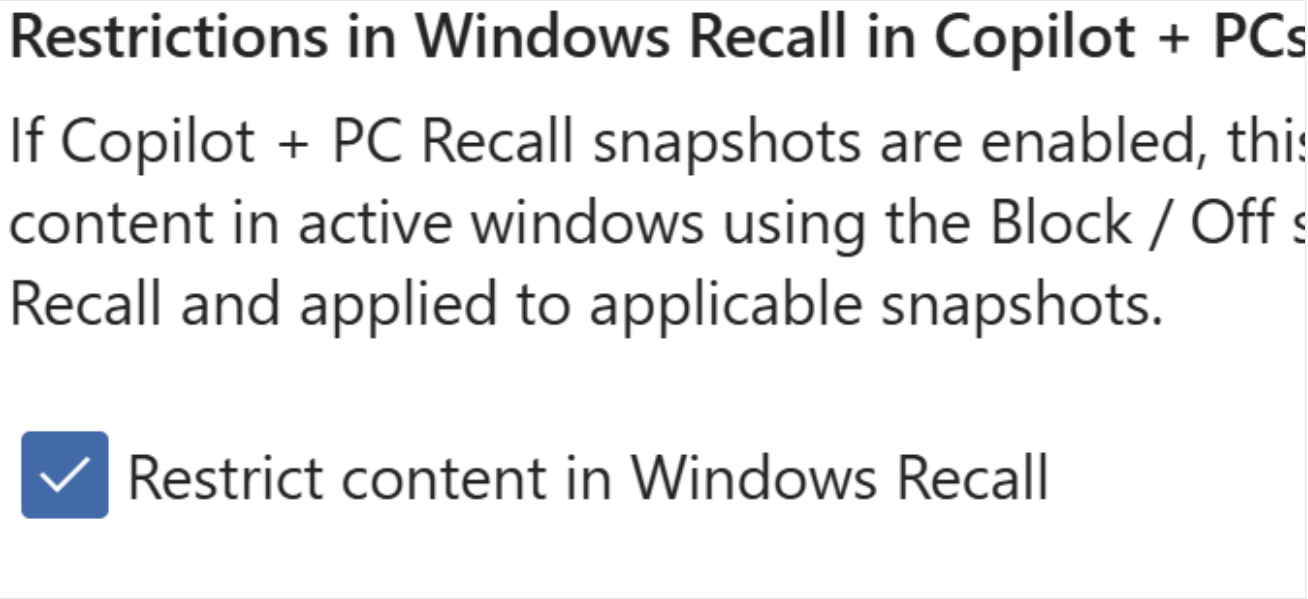
Configuring DLP policy protections for Recall

When you configure a Copilot+ PC so that the user can opt in to Recall, and the user does opt in, you can create new DLP policies or edit existing policies to include protections for Recall.

For general information on creating DLP policies, see Create and deploy a data loss prevention policy.

The setting to enable DLP protections for Recall is available when you create or edit a DLP policy in **Enterprise Applications & devices** policies. In the policy editing workflow under **Actions** >

Audit or restrict activities on devices > **Restrictions in Windows Recall in Copilot+ PCs** > **Restrict content in Windows Recall**. You can select **Audit only** or **Block**.



If you select **Audit only**, then the user doesn't see any notification when a snapshot is taken that contains sensitive information. However, the event is logged in the DLP reports.

If you select **Block**, then when a snapshot is taken that contains sensitive information, then the sensitive content isn't included in the snapshot.

View Recall related events in Microsoft Purview Activity Explorer.

See also

- Learn about Endpoint data loss prevention
- Create and deploy data loss prevention policies
- Recall

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/22/2026

Learn about the default DLP policy for devices

To help you protect sensitive data on endpoint devices, Microsoft provides a default Microsoft Purview data loss prevention (DLP) policy called **Default policy for devices**.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Licensing

For information on licensing, see

- Microsoft 365 Enterprise Plans
- Microsoft 365 Service Descriptions

What does the default policy do?

This default policy looks for items with the sensitive information type Credit card number It's configured like this.

- Admin units: **Full directory**
- Choose where to apply the policy: **Devices** > **All users, groups, devices, device groups**
  - **SharePoint sites** > **All sites**
  - **OneDrive accounts** > **All users and groups**
- Conditions:
  - **Content contains** > **Credit card number**
- Action:
  - **Audit or restrict activities on devices**
    - **Upload to a restricted cloud service domain or access from an unallowed browser** > **Audit only**
    - **Copy to clipboard** > **Audit only**
    - **Copy to a removable USB device** > **Audit only**
    - **Copy to a network share** > **Audit only**
    - **Print** > **Audit only**
    - **Copy or move using unallowed Bluetooth app** > **Audit only**

    - **Copy or move using RDP** > **Audit only**
  - **App access restrictions** > **Access by restricted apps** > **Audit only**
- Policy mode: **Turn the policy on immediately**

Administrator can view the activities and edit the policies details by logging into the Microsoft Purview portal.

Admins can view this policy in the Microsoft Purview portal. Navigate to **Data Loss prevention** > **Policies**.

Edit or delete the default policy

To edit the default policy for better performance or to delete it, just use an account with **DLP Compliance Management** permissions. For more information, see, Permissions.

Last updated on 03/24/2026

Learn about Data Loss Prevention for Cloud Apps in Edge for Business

Microsoft Purview Data Loss Prevention (DLP) monitoring and protection are built right into the Microsoft Edge for Business browser. You don't need to onboard the device into Microsoft Purview. This integration helps you stop users from sharing sensitive information to and from cloud apps by using Edge for Business.

Before you begin

If you're new to Microsoft Purview collection policies, Microsoft Purview pay-as-you-go billing models, or Microsoft Purview DLP, familiarize yourself with the information in these articles:

- Collection policies
- Learn about data loss prevention
- Learn about Microsoft Purview billing models
- Get started with activity explorer

Licensing

For information on licensing, see

- Microsoft 365 Enterprise Plans
- Microsoft 365 Service Descriptions

Protecting data shared from a managed device to an unmanaged app in Edge for Business is a pay-as-you-go feature. For more information, see Learn more about pay-as-you-go-capabilities and request definitions. For information on setting up the pay-as-you-go billing model, see Enable Microsoft Purview pay-as-you-go features for new customers.

Scenarios where data in Microsoft Entra-registered (managed) apps is protected while using Edge for Business are included in a Microsoft 365 E5 or equivalent license.

Permissions

Permissions to create and deploy Microsoft Purview DLP policies are found here.

You also need permissions for prerequisites and configurations outside of Microsoft Purview. For more information on required permissions, see Supported cloud apps.

Managed devices

You can protect Windows 10 and Windows 11 devices that are managed by Microsoft Intune. Users must sign in to the device using their work or school account.

On these devices, Edge for Business connects directly with Microsoft Purview and Microsoft Edge services to get policy updates and apply protections. Microsoft Edge configuration policies block users from using protected unmanaged cloud apps in unprotected browsers. If users try to access an unmanaged app in an unprotected browser, they're blocked and must use Edge for Business.

Microsoft Purview DLP policies can help prevent sharing via Edge for Business from managed devices to unmanaged AI apps, and Microsoft Purview collection policies can be applied to interactions with unmanaged AI apps from managed devices. Policies targeting unmanaged apps on managed devices apply across all Edge browser profiles (work profile, personal profile and InPrivate).

Unmanaged devices

Unmanaged devices aren't connected to Intune or joined to your organization using Microsoft Entra. Users don't sign into the device with their work or school account. Instead, they sign into their Edge for Business work profile to access organization managed apps from their unmanaged device.

Edge for Business applies DLP policies for unmanaged devices to the work profile identity. These policies don't apply when users choose a personal or InPrivate profile, except when the work profile used to open the InPrivate window has a policy applied. When you target policies to managed apps on unmanaged devices, you must enforce the work profile in Edge for Business. Protections for unmanaged devices help prevent users from sharing sensitive information with cloud apps in Edge for Business.

Supported cloud apps

Microsoft Entra connected (managed) apps

Microsoft Entra connected (managed) apps are business apps set up for Microsoft Entra single sign-on (SSO). Policies apply when users access them in Edge for Business by using work or

school account credentials. Policies for managed apps in Edge for Business are supported on unmanaged devices and managed devices and policies apply only in Edge work profile.

To activate policies that apply to managed apps, you need additional permissions for Conditional Access administration and Microsoft Defender for Edge In-Browser protection.

Unmanaged cloud apps

Your organization doesn't manage these apps. Users access them without signing in by using their Microsoft work or school account. You can use policies for unmanaged cloud apps in Edge for Business on Intune-managed devices and the policies apply across all Edge browser profiles (work profile, personal profile and InPrivate).

> **Important**
>
> Microsoft Purview browser and network data security policies don't apply to B2B guest users.

To fully activate DLP policies for unmanaged apps, additional groups and policies are required outside of Purview. This behavior is automated and triggered by the admin action in Purview. After you save your first Purview collection or DLP policy targeting unmanaged apps in Edge for Business, the Microsoft Edge management service automatically creates and manages the required configurations and policies outside of Purview. To learn more, see Automatic activation of your Microsoft Purview DLP policy in Microsoft Edge.

To fully activate your Purview policies in Edge for Business, the admin needs these permissions:

- Directory Reader
- Microsoft Edge administration
- Microsoft Intune administration

Browser policies in Edge for Business support these unmanaged apps:

- Adobe Firefly
- CapCut
- ChatGPT (consumer)
- Cohere
- DeepAI
- DeepSeek

- Google Gemini
- Grok (xAI)
- Meta AI
- Microsoft Copilot 365 Chat
- Notion AI
- Otter.ai
- Perplexity AI
- QwenAI
- Qwen Chat
- Runway
- Textcortex
- Textcortex Zenochat
- You (You.com)

> **Note**
>
> The unmanaged cloud app features apply to the consumer versions of Microsoft 365 Copilot and ChatGPT. For more information on the enterprise versions of these apps and available Purview features see, Learn more about Microsoft 365 Copilot Enterprise protections and Use Microsoft Purview to manage data security & compliance for ChatGPT Enterprise.

Supported browsers

DLP policies for cloud apps in the browser work directly in Edge for Business.

Edge for Business

These features are available in the two latest stable versions of Edge for Business, starting with version 144. For more information on Edge for Business versions, see Microsoft Edge Releases.

> **Important**
>
> Microsoft Purview browser data security policies don't apply to B2B guest users.

> **Tip**

Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Activities you can monitor and take action on

You can audit and manage these activities on sensitive items in the browser:

Expand table

Activity	Device Type	App Type	Supported Policy Actions
Upload text	Managed	Unmanaged	allow, block, both actions audited
Upload file	Managed, Unmanaged	Managed, Unmanaged	allow, block, both actions audited
Download file	Managed, Unmanaged	Managed	allow, block, both actions audited
Cut/copy data	Managed, Unmanaged	Managed	allow, block, both actions audited
Paste data	Managed, Unmanaged	Managed	allow, block, both actions audited
Print data	Managed, Unmanaged	Managed	allow, block, both actions audited
Protected clipboard (Preview)	Managed, Unmanaged	Managed	See Microsoft Edge Protected Clipboard (preview)
Screen capture (Preview)	Managed, Unmanaged	Managed	See Microsoft Edge Protected Clipboard (preview)

Some activities have limitations:

- Cut/copy data, paste data, and print data activities can only be used with the managed or unmanaged devices condition.
- Download file activity isn't supported for apps that don't follow Microsoft Edge for Business's download pipeline.

> **Note**
>
> Policy sync status displays N/A for DLP policies that use the built-in protections in Edge for Business. These policies use a different activation and delivery path than other DLP locations

and don't report through the same policy sync status field. Activation or configuration failures are surfaced in the Microsoft Purview portal.

Policies for managed app interactions

DLP policies that target managed apps in the browser apply to Edge for Business in Windows 10/11 and macOS desktop devices when the user signs in to their Edge for Business work profile.

Edge for Business automatically disables developer tools and blocks the apps from opening in native clients when policies apply to managed apps (in both audit and block modes).

To activate protections in the Edge for Business work profile for managed apps:

- Onboard apps to Conditional Access app control.
- Import user groups from connected apps. This step allows you to target Groups in your Entra Conditional Access policy and your Purview policies for managed apps in Edge for Business.
- Set up a Microsoft Entra Conditional Access policy with custom session controls.
- Configure Edge for Business in-browser protection.
- Create a Microsoft Purview DLP policy that targets managed app interactions.
- (Optional) Activate protected clipboard and screen capture prevention capabilities in the Microsoft Admin Portal settings for Microsoft Edge.

For full implementation details, see Help Prevent Users from Sharing Sensitive Info with Cloud Apps in Edge for Business.

> **Important**
>
> Protections might not apply in Edge for Business to managed apps included in a Microsoft Purview browser policy if the user is in scope for both a Microsoft Purview managed cloud app DLP policy *and* a Microsoft Defender session policy *or* Microsoft Purview endpoint DLP policy. You must remove or exclude the users from the Microsoft Defender and the endpoint DLP policies for the managed cloud apps in Edge for Business policy to apply.

When you add users to policies for the first time, the policy might not be applied right away if they're already signed in to the app. The policy applies after their token expires and they sign in

again. You can change the sign-in frequency by using conditional access session controls to shorten the wait time.

Some known limitations in Conditional Access app control can impact Microsoft Purview policies that target managed apps in the browser. For more information, see known limitations in Conditional Access app control.

Accessing data from managed app interactions

You can view policy data and alerts in Defender XDR investigations.

Policies for unmanaged app interactions

DLP policies that target unmanaged apps in the browser apply to Microsoft Edge for Business on Windows 10 and Windows 11 desktop devices managed by Microsoft Intune. For setup details, see Help prevent sharing via Edge for Business to unmanaged AI apps from managed devices.

> **Important**
>
> Microsoft Purview inline protection in Microsoft Edge for Business for unmanaged apps doesn't support tenants that use multi-admin approval in Microsoft Intune. For more information, see **Multi-admin approval**.

When you activate a Microsoft Purview DLP policy for unmanaged cloud apps, Purview automatically creates the required Microsoft Edge configuration policies and Microsoft Intune policies (outside of Purview) and assigns the included users.

If you configure the DLP policy to block, users in scope can't use unprotected browsers where the policy doesn't apply. This restriction doesn't affect their experience in Microsoft Edge for Business.

For more information, see Activate your Microsoft Purview policy in Microsoft Edge.

Activating protections in Microsoft Edge for Business follows these phases:

- Create a Microsoft Purview DLP policy targeting unmanaged app interactions.
- Microsoft Edge management service automatically creates the configuration policies that activate DLP policies in Microsoft Edge for Business. The configuration policies use Microsoft Intune policies to activate your Microsoft Purview policies in Microsoft Edge for Business.

- Create a collection policy targeting unmanaged apps in the browser to identify additional sensitive data sharing that might be happening across your organization.

> **Important**
>
> If the **automatic behaviors** fail to sync, Microsoft Purview shows an error message. Purview doesn't enforce DLP and collection policies for unmanaged apps in Edge for Business until you resolve the error. An admin with the **required permissions** must resync to resolve the error. For more information, see **activate your Microsoft Purview policy in Microsoft Edge**.

Accessing data from unmanaged app interactions

You can view activities and audit log entries in activity explorer, audit logs, and Defender XDR investigations. In activity explorer, filter by enforcement plane set to browser. Data specific to AI apps is also visible in DSPM for AI.

Considerations for unmanaged cloud apps policies

Review this information when you configure policies that include unmanaged cloud apps, or when you troubleshoot unexpected policy behavior:

- When multiple catalog entries exist for the same app with differences in the entry name (for example, QwenAI and Qwen Chat), include all entries for the app to avoid unintended coverage gaps.
- Some unmanaged AI apps like Runway and Meta AI might intermittently send content in encoded form to dynamically generated endpoints, which can impact policy enforcement.
- Policies that target unmanaged apps can capture interactions from both the consumer and enterprise versions of an app when the app's URL is shared across instances (for example, ChatGPT consumer and ChatGPT enterprise).
- When you target an unmanaged cloud app in an Edge for Business browser policy, detection is based on the destination app's traffic, which isn't always exactly the same as the originating app.
- Inline evaluation size limits: Microsoft Purview evaluates up to 4 MB of content for uploadText, and files up to 3 MB for uploadFile.

Default policies for unmanaged AI apps from Microsoft Data Security Posture Management for AI

Microsoft Purview Data Security Posture Management for AI (DSPM for AI) offers recommended policies to monitor and block supported unmanaged generative AI apps. Use one-click policies in DSPM for AI to apply them.

Working with Microsoft Purview Endpoint data loss prevention

Endpoint DLP policies are prioritized over inline DLP policies for cloud apps in Edge for Business if the same user, context, and action are configured in both policies. For example, if you have an Endpoint DLP policy that blocks a file uploads, and you also have an inline policy for cloud apps in Edge for Business that monitors file uploads. the Endpoint DLP policy is prioritized and applied. For more information, see Learn about Endpoint data loss prevention.

See also

- Learn about data loss prevention
- Create and deploy data loss prevention policies
- Get started with Activity explorer

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 09/15/2026

Learn about Microsoft Purview Network Data Security

Microsoft Purview network data security helps organizations monitor, classify, and apply protections to unmanaged and untrusted cloud app traffic, including generative AI apps. Purview provides these capabilities through integrations with one or more secure access service edge (SASE) or secure browser solutions:

- **Microsoft Entra Global Secure Access**: To learn how to configure content policies that integrate with Purview, see Create a content policy to filter network content.
- **Non-Microsoft SASE and secure browser solutions** (general availability or preview, depending on the integration): Find these solutions in the Security Store in Microsoft Purview Data Loss Prevention.

Network data security utilizes the same classifiers already configured in other Microsoft Purview policies, extending discovery and protection to the network layer through collection policies for data discovery, data loss prevention policies for protection, or both, depending on your organization's needs.

With network data security, you can identify, block, and alert on sensitive content that's shared through these interactions:

- Interactions with generative AI through browsers, apps, and add-ins, such as ChatGPT, Gemini, and Claude.
- Files uploaded to unsanctioned cloud storage providers, including Dropbox, Box, and Google Drive.
- Emails and file attachments shared with cloud email providers, such as Gmail.
- Form submissions through online form services, including Google Forms.
- Social media posts on common services like Facebook and X.

Before you begin

If you're new to Microsoft Purview collection policies, Microsoft Purview pay-as-you-go billing models, or Microsoft Purview DLP, you should familiarize yourself with the information in these articles:

- Collection Policies solution overview

- Learn about data loss prevention
- Learn about Microsoft Purview billing models
- Get started with activity explorer

Licensing

For information on licensing, see

- Microsoft 365 Enterprise Plans
- Microsoft 365 Service Descriptions

- **Network data security with Microsoft Entra Global Secure Access** requires **one** of the following license combinations:
  - Microsoft 365 E7 per-seat licenses
  - Purview E5 (or equivalent) per-seat licenses *and* Entra Internet Access (or equivalent) licenses

- **Network data security with non-Microsoft SASE and secure browser solutions** requires:
  - Purview E5 (or equivalent) per-seat licenses *and* the Microsoft Purview pay-as-you-go billing model. Learn about Microsoft Purview billing models

> **Note**
>
> You must configure pay-as-you-go for your Microsoft 365 tenant to use Purview network data security with non-Microsoft SASE and secure browser solutions. For information on setting up the pay-as-you-go billing model, see **Enable Microsoft Purview pay-as-you-go features for new customers**.
>
> Pay-as-you-go is not required to be set up for integration with Microsoft Entra Global Secure Access.

How network data security works

From a broad perspective, the Microsoft Purview network data security solution combines two components:

Network security solution

The network data security solution integrates secure access service edge (SASE) and non-Microsoft secure browser solutions directly into Microsoft Purview. The integrated solutions

monitor network traffic and send the data to Microsoft Purview for classification and policy evaluation. When you apply protections via the use of DLP policies, the communication between the integrated solution and Microsoft Purview is in real time. If you are using network data security for monitoring only via collection policies, the communication is asynchronous.

For more information on which integrations are supported, see the Security Store in Microsoft Purview Data Loss Prevention.

> **Important**
>
> If you choose to integrate with non-Microsoft partners, they will be able to access and possibly store some policy configuration, including user identifiers. Their terms, conditions, and privacy policy will govern the usage and storage of this data.

Microsoft Purview

You configure the integration between Microsoft Purview and the network security solution using the Security Store. This integration establishes the bidirectional communication channel between the network security solution and Microsoft Purview.

Next, configure a collection policy or data loss prevention policy that defines the conditions, activities, and cloud apps that you want the network security solution to collect and send to Microsoft Purview.

- For more information on how to create a collection policy for network data security, see Scenario 1 Detect sensitive data shared with unmanaged cloud apps via network.
- For more information on how to create a data loss prevention policy for network data security, see Use Network Data Security to help prevent sharing sensitive information with unmanaged AI

Microsoft Purview shares the appropriate DLP and collection policy configuration with your integrated network security solution, and the solution sends any matching network data to Microsoft Purview for classification and policy evaluation. If you configure content capture in the collection policy, the full conversation that happens between the user and the AI app is captured and sent to Microsoft Purview as well.

After the data is classified, it's available in activity explorer and activity explorer in Data Security Posture Management (DSPM). If a data loss prevention policy is matched, and you have configured alerts, they will be available in DLP alerts.

After you configure the integration between Microsoft Purview and your network security solution, allow up to 24 hours for your policies to be distributed to the network security solution and for the first data to show up. Once the two services fully communicate with each other, it can take up to 30 minutes for data about a request from a client to a website or cloud app to appear in the audit log and activity explorer.

> **Important**
>
> Purview network data security policies don't apply to B2B guest users.

Supported network data security collection policy configuration

The Microsoft Purview side of the configuration is done via a collection policy. Here are the configuration options that are supported:

- **Conditions** - The conditions you can use in a network data security collection policy are the same as the conditions you can use in other Microsoft Purview policies. For example, you can use the **Content contains** > **Sensitive information types** condition to classify sensitive items that are being shared with generative AI and other unmanaged cloud apps.

- **Activities** - Network data security supports four activities:
  - **Text sent to or shared with cloud or AI app**.
  - **File uploaded to or shared with cloud or AI app**.
  - **Text received from cloud or AI app**.
  - **File downloaded from cloud or AI app**.

> **Note**
>
> The activities supported may differ depending on integrated SASE solution. Check with your SASE solution provider for details on supported activities. Microsoft Entra Global Secure Access supports all supported activities.

- **Data sources** - These are the locations that the endpoint device is communicating with.
  - **Unmanaged cloud apps** - Network data security collection policies support all the sources that are in the Microsoft Defender for Cloud Apps Cloud app catalog which includes over 35,000 discoverable cloud apps.
  - **Adaptive app scopes** - all apps in multiple categories including generative AI, cloud storage, collaboration, social network, and webmail.

Supported network data security data loss prevention policy configuration

The Microsoft Purview side of the configuration is done via a data loss prevention policy. Here are the configuration options that are supported:

- **Data sources** - These are the locations that the endpoint device is communicating with.
  - **Unmanaged cloud apps** - Network data security collection policies support all the sources that are in the Microsoft Defender for Cloud Apps Cloud app catalog which includes over 35,000 discoverable cloud apps.
  - **Adaptive app scopes** - all apps in multiple categories including generative AI, cloud storage, collaboration, social network, and webmail.

- **Conditions** - The conditions you can use in a network data security collection policy are the same as the conditions you can use in other Microsoft Purview policies. For example, you can use the **Content contains** > **Sensitive information types** condition to classify sensitive items that are being shared with generative AI and other unmanaged cloud apps.

- **Activities** - Network data security supports four activities:
  - **Text sent to or shared with cloud or AI app**.
  - **File uploaded to or shared with cloud or AI app**.
  - **Text received from cloud or AI app**.
  - **File downloaded from cloud or AI app**.

- **Actions** - Network data security supports **Audit only** and **Block** actions.

> ⓘ **Note**
>
> The activities and actions supported may differ depending on integrated solution. Check with your solution provider for details on supported activities. Microsoft Entra Global Secure Access supports all supported activities and actions.

Default policy from Microsoft Purview Data Security Posture Management

Microsoft Purview Data Security Posture Management (DSPM) offers recommendations to help monitor communications with generative AI apps. Select the recommendation **Extend insights into sensitive data in AI app interactions** to create a one-click policy named **DSPM for AI - Detect sensitive info shared with AI via network**. After it's created, you can edit this default policy for network data security as you would any collection policy.

Considerations for unmanaged cloud apps policies

Review this information when you configure policies that include unmanaged cloud apps, or when you troubleshoot unexpected policy behavior:

- When multiple catalog entries exist for the same app with differences in the entry name (for example, QwenAI and Qwen Chat), include all entries for the app to avoid unintended coverage gaps.
- Some unmanaged AI apps like Runway and Meta AI might intermittently send content in encoded form to dynamically generated endpoints, which can impact policy enforcement.
- Policies that target unmanaged apps can capture interactions from both the consumer and enterprise versions of an app when the app's URL is shared across instances (for example, ChatGPT consumer and ChatGPT enterprise).
- Inline evaluation size limits: Microsoft Purview evaluates up to 4 MB of content for uploadText and downloadText, and files up to 3 MB for uploadFile and downloadFile.

> **Note**
>
> The unmanaged cloud app features only apply to the consumer version of Microsoft Copilot. For more information on the enterprise version and available Purview features, see Learn more about Microsoft Copilot Enterprise protections.

Supported network protocols

All network data security integrations support classifying traffic sent and received from an endpoint device over HTTP and HTTPS protocols to websites, cloud apps, and generative AI services. Support for WebSocket protocol may differ depending on the integrated SASE or non-Microsoft secure browser solution. Check with your solution provider for details on supported protocols.

Accessing network data security data

Data from network data security appear in activity explorer, Data Security Posture Management activity explorer events, and if alerts are enabled, DLP alerts.

Activity explorer

In activity explorer, you can filter events with **enforcement plane** set to *network*. This filter shows you activities that network data security policies generate.

Billing model

Network data security uses the *request* as unit of measure for pay-as-you-go billing purposes when using non-Microsoft SASE and secure browser integrations. A request is each network call made from a device or browser to a website or API. This definition doesn't include the responses to the requests. For more information on pay-as-you-go billing for network data security, see Other Microsoft Purview solutions that use pay-as-you-go pricing and Requests.

Here are some examples:

Expand table

Activity	Data type	Example
Text sent to or shared with cloud or AI app	Human readable strings transmitted inline	- submitting a form with textual information - sending raw text or a prompt to a generative AI - the body of an email - sending JSON data to an API
File uploaded to or shared with cloud or AI app	Byte streams, including text based file, binary files, txt files, source code, documents, images, videos, .exe's, .pdf's, archive files	- Uploading a profile picture to social media - sending a document or .pdf file as an email attachment - sharing a document with generative AI - transferring a document or .zip files to a cloud storage solution

Next steps

- Scenario 1 Detect sensitive data shared with unmanaged cloud apps via network
- Use Network Data Security to help prevent sharing sensitive information with unmanaged AI
- If using Microsoft Entra Global Secure Access, see Create a content policy to filter network content

Last updated on 09/16/2026

Use Network Data Security to help prevent sharing sensitive information with unmanaged AI

> **Note**
>
> Some features used in this scenario are in preview.

This article uses the process you learned in Design a data loss prevention policy to show you how to create a Microsoft Purview data loss prevention (DLP) policy that helps prevent sharing sensitive information with unmanaged AI apps through network data security using an integrated SASE solution, such as Microsoft Entra Global Secure Access, or non-Microsoft secure browser solution. Work through this scenario in your test environment to familiarize yourself with the policy creation UI.

> **Important**
>
> This article presents a hypothetical scenario with hypothetical values. It's only for illustrative purposes. Substitute your own sensitive information types, sensitivity labels, distribution groups, and users.

How you deploy this DLP policy scenario is as important as policy design. Create and deploy a data loss prevention policy shows you how to use the deployment options so that the policy achieves your intent while avoiding costly business disruptions.

Prerequisites and assumptions

> **Important**
>
> Before you start this procedure, read Learn about Microsoft Purview Network Data Security. It provides important information about the prerequisites and assumptions for this scenario, including considerations for unmanaged apps policies.

Before you begin, make sure the following prerequisites and assumptions are met:

- You have an supported SASE or secure enterprise browser integration. For more information on which integrations are supported, see the Security Store in Microsoft Purview Data Loss Prevention.
- If you're using Microsoft Entra Global Secure Access
  - The device must be Entra ID joined.
  - You must configure a Transport Layer Security inspection policy
  - You must configure a content policy in Microsoft Entra GSA Internet Access.
- If you are using a non-Microsoft SASE or secure enterprise browser provider
  - Ensure you've configured Purview pay-as-you-go billing. For more information, see Enable Microsoft Purview pay-as-you-go features for new customers.
  - Ensure you've completed the integration steps outlined in your provider's documentation.
- This procedure uses hypothetical distribution groups, one named *Finance Team*, and another group for the *Security Team*. You have to create these groups in your environment or substitute your own groups.

Policy intent statement and mapping

We want to embrace responsible AI use within our organization while protecting customer and organizational sensitive data from intentional or accidental exposure. Because the Finance department regularly handles highly sensitive information, we need to create a policy that blocks Finance users from sharing sensitive content with any unsanctioned generative AI application. Since many of our users have flexibility in which browser they choose to use, as well as the ability to use locally installed AI apps and Office add-ins, we need to ensure broad endpoint coverage through network data security protection. Given the sensitivity of the Finance department's access to data, we need to prevent financial data, personally identifiable information (PII), or confidential documents from being shared through text prompts or uploaded files. Then, to ensure efficient and quick incident management, we must generate alerts and ensure our security team is notified via email any time a block occurs. Lastly, we want this policy to take effect immediately to begin protecting company data as soon as possible.

Expand table

Statement	Configuration question answered and configuration mapping
Because the Finance department regularly handles highly sensitive information, we need to create a policy that blocks Finance users from sharing sensitive content with any unsanctioned generative AI application.	Choose cloud apps to apply the policy: - Select + **Add cloud apps** - Select the **Adaptive app scopes** tab - Choose **All unmanaged AI apps** and select **Add** - Select **Edit scope** - Select **Include only specific** - Choose the **Finance department** group
Since many of our users have flexibility in which browser they choose to use, as well as the ability to use locally installed AI apps and Office add-ins, we need to ensure broad endpoint coverage through network data security protection.	Choose where to enforce the policy: - Enable **Network and non-Microsoft secure browsers**
Given the sensitivity of the Finance department's access to data, we need to prevent financial data, personally identifiable information (PII), or confidential documents...	Customize advanced DLP rules: - Select + **Create rule** - Select + **Add condition** and choose **Content contains** - Select **Add** and chooose **Sensitive info types** - Select **ABA Routing Number**, **U.S. Bank Account Number**, **Credit Card Number**, **U.S. Social Security Number (SSN)** and select **Add** - Select **Add** and choose **Sensitivity labels** - Choose **Confidential** and select **Add** - Ensure the **Group operator** is set to **Any of these**
...from being shared through text prompts or uploaded files.	Actions: - Select + **Add an action** and choose **Restrict browser and network activities** - Select **Text sent to or shared with cloud or AI apps** and **File uploaded to or shared with cloud or AI apps** - Select **Block** for both
Then, to ensure efficient and quick incident management, we must generate alerts and ensure our security team is notified via email any time a block occurs.	Incident reports: - Set the severity level to **High** -Ensure **Send an alert to admins when a rule match occurs** is **On** - Select + **Add or remove users** and choose the **SecurityOps distribution group** - Ensure **Send alert every time an**

Statement	Configuration question answered and configuration mapping
	activity matches the rule is selected, and then **Save** the rule
Lastly, we want this policy to take effect immediately to begin protecting company data as soon as possible.	Policy mode: - Select **Turn the policy on immediately**

Configure SASE provider integration

> **Important**
>
> You must have a Secure Access Service Edge (SASE) or non-Microsoft secure browser provider integrated with Purview to begin discovering and protecting content shared over the network.

To integrate your provider with Purview, complete the following steps:

1. Sign in to the Microsoft Purview portal.
2. Open the **Data Loss Prevention** solution and select **Security Store** from the left navigation bar.
3. Select **Get solution** for the provider you're using.
4. Complete the setup steps provided by your integrated SASE or secure browser provider.

Steps to create policy

To create the DLP policy, complete the following steps:

1. Sign in to the Microsoft Purview portal.
2. Select **Data loss prevention** > **Policies** > **+ Create policy**.
3. Select **Inline web traffic**.
4. Select **Custom** from the **Categories** list and then select **Custom policy** from the **Regulations** list.
5. Choose **Next**.
6. Enter a policy name and provide an optional description. You can use the policy intent statement here.
7. Choose **Next**.
8. Select **+ Add cloud apps**.

9. Select the **Adaptive app scopes** tab, then choose **All unmanaged AI apps** and select **Add (1)**.
10. Select **Edit scope** on **All unmanaged AI apps**.
11. Select the **Include only specific** option on the **Add or edit the scope for All Unmanaged AI Apps**.
12. Select **+ Add inclusions**.
13. Select **Specific users and groups**.
14. Search for and select the **Finance department** group, then select **Add**.
15. Select **Save and close**.
16. Choose **Next**.
17. On the **Choose where to enforce the policy** page, ensure **Network and non-Microsoft secure browsers** is **On**, then select **Next**.

> ⓘ **Note**
>
> If using a non-Microsoft SASE or secure enterprise browser integration, you must have pay-as-you-go configured. Learn more about pay-as-you-go billing.

18. On the **Define policy settings** page, ensure **Create or customize advanced DLP rules** is selected and select **Next**.
19. On the **Customize advanced DLP rules** page, select **+ Create rule**.
20. Give the rule a unique **name** and optional description.
21. Under **Conditions**
    a. Select **+ Add condition** then choose **Content contains**.
    b. Select **Add** and choose **Sensitive info types**.
    c. Choose **ABA Routing Number**, **U.S. Bank Account Number**, **Credit Card Number**, **U.S. Social Security Number (SSN)** and select **Add**.
    d. Select **Add** and choose **Sensitivity labels**
    e. Choose **Confidential** and select **Add**.
    f. Ensure the **Group operator** is set to **Any of these** for the **Content contains group**.
22. Under **Actions**
    a. Select **+ Add an action** and choose **Restrict browser and network activities**.
    b. Select **Text sent to or shared with cloud or AI apps** and **File uploaded to or shared with cloud or AI apps**.
    c. Select **Block** for both.
23. Under **Incident reports**
    a. Set the **Severity level in admin alerts and reports** to **High**

    b. Ensure **Send an alert to admins when a rule match occurs** is **On**
    c. Select **+ Add or remove users** and choose the **SecurityOps distribution group**
    d. Ensure **Send alert every time an activity matches the rule** is selected
24. Select **Save**.
25. Review the rule configuration, ensure it's status is **On**, and select **Next**.
26. On the **Policy mode** page choose **Turn the policy on immediately** and then select **Next**.
27. Review the policy information then select **Submit** to create the policy.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 09/16/2026

Help prevent sharing via Microsoft Edge for Business to unmanaged AI apps from managed devices

This article uses the process you learned in Design a data loss prevention policy to show you how to create a Microsoft Purview data loss prevention (DLP) policy that helps prevent sharing sensitive information from organization managed devices to unmanaged AI apps in Microsoft Edge for Business. Work through this scenario in your test environment to familiarize yourself with the policy creation UI.

Use this scenario to use the Edge for Business browser as the control point to block the sharing of sensitive information to unmanaged AI apps. The policy that's created here applies across all Edge browser profiles (work profile, personal profile and InPrivate) on Windows devices that are managed by Microsoft Intune.

> **Important**
>
> Microsoft Purview inline protection in Microsoft Edge for Business for unmanaged apps doesn't support tenants that use multi-admin approval in Microsoft Intune. For more information, see Multi-admin approval.

> **Important**
>
> This article presents a hypothetical scenario with hypothetical values. It's only for illustrative purposes. Substitute your own sensitive information types, sensitivity labels, distribution groups, and users.

How you deploy a policy is as important as policy design. This article shows you how to use the deployment options so that the policy achieves your intent while avoiding costly business disruptions.

Prerequisites and assumptions

This procedure uses hypothetical distribution groups, one named *Finance Team*, and another group for the *Security Team*.

> **Important**
>
> Before you start this procedure, read Learn about Data Loss Prevention for Cloud Apps in Edge for Business. It provides important information about the prerequisites and assumptions for this scenario.

Implementing DLP policies for unmanaged apps in the browser follows these phases:

1. Create a DLP policy targeting unmanaged apps in the browser (the procedure in this article) in the Microsoft Purview Data Loss Prevention portal.
2. The creation of the DLP policy triggers the Microsoft Edge management service to automatically create the Microsoft Edge configuration policies required to activate DLP policies in Edge for Business. The configuration policies use Microsoft Intune policies to fully activate your Microsoft Purview policies in Microsoft Edge.
3. (optional) Create a collection policy that targets unmanaged apps in the browser to identify other sensitive data sharing across your organization.

> **Important**
>
> Activating a Purview DLP policy for unmanaged cloud apps automatically adds included users to the required Microsoft Edge configuration policies. When the Purview DLP policy is set to block, included users are blocked from using unprotected browsers where the policy doesn't apply. The experience in Edge for Business isn't impacted. For more information see: activate your Microsoft Purview policy in Microsoft Edge.

Billing and licensing requirements

This feature uses pay-as-you-go billing or per-user licensing for Microsoft Purview capabilities or both. To help you understand and manage your usage, Microsoft Purview provides a **Usage center** in the Microsoft Purview portal. For more information, see Manage pay-as-you-go and per-user licensing usage. To learn more about Purview billing, see Purview billing models..

Policy intent statement and mapping

We need to block members of the finance team from sharing sensitive information to unmanaged AI apps in Edge for Business. Other teams don't have access to this highly sensitive information, so

the block only needs to apply to this team. When their prompts contain information like bank account, routing, or the SWIFT (Society for Worldwide Interbank Financial Telecommunication) numbers of our international customers, the sharing is blocked. We also have to meet alerting requirements. We want to notify our security team with an email every time there's a match to the policy. Lastly, we want the policy to take effect as soon as possible after testing, and we need to be able to see related activity within the system.

Expand table

Statement	Configuration question answered and configuration mapping
We need to block members of the finance team from sharing sensitive information to unmanaged AI apps via Edge...	- Choose where to apply the policy: **Edge for Business** -Administrative scope: **Full directory** - Where to apply the policy: **Adaptive app scopes** > **All unmanaged AI apps**
Other teams don't have access to this information, so the block only needs to apply to this team...	- scope for each app" **specific users and groups, Include users and groups** > **Finance Team**
When their text prompts contain information like bank account, routing or the SWIFT numbers of our international customers, the sharing should be blocked.	What to monitor: - use the **custom policy template** - Conditions for a match: **Content contains Sensitive info types** > **ABA Routing Number**, **Australia Bank Account Number**, **Canada Bank Account Number**, **International Bank Account Number (IBAN)**, **Israel Bank Account Number**, **Japan Bank Account Number**, **New Zealand bank account number**, **SWIFT Code**, **U.S Bank Account Number** - Action: **Restrict browser and network activities** > **Text upload** > **Block** and **File upload** > **Block**
We also have to meet alerting requirements. We want to notify our security team with an email every time there's a match to the policy.	- Incident reports: **Send an alert to admins when a rule match occurs** - Send email alerts to these people (optional): add the **Security team** - Send an alert every time an activity matches the rule: **selected** - Use email incident reports to notify you when a policy match occurs: **On** - Send notifications to these people: add individual admins as desired - You can also include the following information in the report: **select all options**
...Lastly, we want the policy to take effect as soon as possible after testing and need to be able to see related activity within the system....	Policy mode: **on in simulation**

Create the policy for unmanaged AI apps

1. Sign in to the Microsoft Purview portal.
2. Select **Data loss prevention** > **Policies** > **+ Create policy**.
3. Select **Inline web traffic**.
4. Select **Custom** from the **Categories** list and then select **Custom policy** from the **Regulations** list.
5. Choose **Next**.
6. Enter a policy name and provide a description. You can use the policy intent statement here.

> **Important**
>
> You can't rename policies.

1. Choose **Next**.
2. Select **+ Add cloud apps** > **Adaptive app scopes**
   a. Select **All unmanaged AI apps** > **Add.** Unmanaged apps in the Generative AI category are included in this scope.

> **Important**
>
> Not all unmanaged AI apps are supported in Edge for Business. Adaptive app scopes apply only to the supported unmanaged AI apps in Edge for Business. For more information, see Supported apps for browser DLP.

1. Select **Edit scope**
2. Select **Include only specific**
3. Choose **+ Add inclusions**
4. Select **Finance Team**
5. Select **Add** and then select **Save and Close** and then select **Next**.
6. On the **Choose where to enforce the policy** step, select **Edge for Business.**
7. On the **Define policy settings** page, the **Create or customize advanced DLP rules** option should already be selected.
8. Choose **Next**.
9. On the **Customize advanced DLP rules** page, select **+ Create rule**.

10. Enter a name and description for the rule.

11. Select **Add condition** and use these values:
    a. Select **Content contains**.
    b. Select **Add** > **Sensitive information types** > **Sensitive info types** > **ABA Routing Number**, **Australia Bank Account Number**, **Canada Bank Account Number**, **International Bank Account Number (IBAN)**, **Israel Bank Account Number**, **Japan Bank Account Number**, **New Zealand bank account number**, **SWIFT Code**, **U.S Bank Account Number**.

12. Select **Add**.

13. Under **Actions**, add an action with these values:

    a. **Restrict browser and network activities**

    b. **Text sent to or shared with cloud or AI apps** > **Block**

    c. **File uploaded to or shared with cloud or AI app** > **Block**

14. Under **Incident reports** select:
    a. Set **Use this severity level in admin alerts and reports** to **Low**.
    b. Set the toggle for **Send an alert to admins when a rule match occurs** to **On**.
    c. Under **Send email alerts to these people (optional)**, select **+ Add or remove users** and then add the email address of the security team.

15. Select **Save** and then select **Next**.

16. On the **Policy mode** page, select **Run the policy in simulation mode**.

17. Select **Next** and then select **Submit**.

18. Select **Done**.

19. The Microsoft Edge management service automatically creates the required configuration policies and required Microsoft Intune policies to activate your Microsoft Purview policies in Microsoft Edge. Additional permissions are required for Microsoft Intune administration and Microsoft Edge administration for this step.

> ⓘ **Important**

If the automatic Purview DLP configuration sync in Microsoft Edge fails, Microsoft Purview shows an error message. Purview doesn't enforce DLP and collection policies for unmanaged apps in Edge for Business until you resolve the error. An admin with the required permissions for unmanaged cloud apps must resync to resolve the error. For more information, see activate your Microsoft Purview policy in Microsoft Edge.

> **Note**
>
> When the action of the policy is configured to **Block**, users are blocked at the device level from opening Firefox and other browsers. They are also blocked from opening Chrome if Microsoft Purview extension for Chrome isn't installed or is out of date. In Chrome with Microsoft Purview extension, access to a dynamic set of generative AI apps is blocked. For more information, see: Activate your Microsoft Purview policies in Microsoft Edge

Steps to review outcomes of the policy in Microsoft Purview activity explorer

See Activity explorer overview for instructions on how to monitor and view activities in the Activity explorer tool.

Steps to review outcomes of Microsoft Purview policy in Microsoft Defender

See Investigate and respond in Microsoft Defender XDR for instructions on how to investigate and respond using Microsoft Defender.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Help prevent users from sharing sensitive information with cloud apps in Microsoft Edge for Business

Use this procedure to create a Microsoft Purview data loss prevention (DLP) policy that blocks users from exfiltrating sensitive information from business-managed apps (such as ServiceNow and Workday) in Microsoft Edge for Business. This procedure follows the policy design process described in Design a data loss prevention policy. Work through this scenario in your test environment to familiarize yourself with the policy creation UI.

Use the managed-app browser DLP policy scenario to block users from exfiltrating sensitive information from business-managed apps such as ServiceNow and Workday through Microsoft Edge.

Prerequisites:

- Onboard the apps to Microsoft Entra Conditional Access with session controls applied.
- Block access to the protected apps outside the Edge work profile.

When the policy applies:

This policy takes effect when users access the protected apps in Edge for Business and sign in with their work or school account for both the Edge work profile and the app. Edge for Business applies DLP policies for unmanaged devices only to the work profile. These policies don't apply to personal or InPrivate profiles unless the work profile that opens the InPrivate window has a policy applied. To target managed apps on unmanaged devices, enforce the work profile in Edge for Business.

> **Important**
>
> This article presents a hypothetical scenario with hypothetical values. It's only for illustrative purposes. Substitute your own sensitive information types, sensitivity labels, distribution groups, and users.

How you deploy a policy is as important as policy design. This article shows you how to use the deployment options so that the policy achieves your intent while avoiding costly business

disruptions.

Prerequisites and assumptions

The following policy-creation procedure uses a hypothetical distribution group named *External*. For more information on DLP protections applying in the browser, see browser data security resources.

> **Important**
>
> Read **Learn about Data Loss Prevention for Cloud Apps in Edge for Business** before you start this procedure. It provides important information about the prerequisites and assumptions for this scenario.

Setting up a policy to protect data sharing with Microsoft Entra-managed apps in the browser follows these phases:

1. Onboard apps to Conditional Access app control.
2. Import user groups from connected apps. Importing user groups is necessary for you to target groups in your Entra Conditional Access policy and your Purview policies for managed apps in Edge for Business.
3. Create a Microsoft Entra Conditional Access policy that's configured with custom session controls enabled and using the client apps condition for browser.
4. Enable Edge for Business in-browser protection, configured to enforce Edge work profile sign in for business apps on all devices.
5. Create a Microsoft Purview DLP policy targeting user interactions with managed apps.
6. (optional) Activate Protected Clipboard and Screen Capture prevention capabilities in the Microsoft Admin Portal settings for Microsoft Edge. See the Get started with Protected Clipboard for Microsoft Edge for Business instructions.
7. Review outcomes of Microsoft Purview policy in Microsoft Defender.

> **Important**

The user and the app must both be in scope for all the prerequisite protections for the policy to apply to the user in Edge.

Policy intent statement and mapping

We need to allow users to access resources in business apps from their BYOD and personal devices, but block them from downloading the data to those devices. Contractors and vendor employees use these types of devices to collaborate for their work activities. When they attempt to download files containing sensitive information, like customer banking information, the action should be blocked. We also have to meet alerting requirements. Lastly, we want this to take effect as soon as possible.

Expand table

Statement	Configuration question answered and configuration mapping
We need to allow users to access resources in business apps like Workday from their BYOD and personal devices, but block them from downloading the data to their devices ...	- Choose where to apply the policy: **Enterprise apps and Devices** -Administrative scope: **Full directory** - Where to apply the policy: **Managed apps** > **Workday**
Contractors and vendor employees use these types of devices to collaborate for their work activities....	- scope alongside apps" **specific users and groups, Include users and groups** > **External**
When they attempt to download files containing sensitive information, like customer banking information, the action should be blocked.	What to monitor: - use the **custom policy template** - Conditions for a match: **Content contains Sensitive info types** > **ABA Routing Number**, **Australia Bank Account Number**, **Canada Bank Account Number**, **International Bank Account Number (IBAN)**, **Israel Bank Account Number**, **Japan Bank Account Number**, **New Zealand bank account number**, **SWIFT Code**, **U.S Bank Account Number** - Action: **Restrict browser and network activities** > **File download ** > **Block**.
We also have to meet alerting requirements. Our security team must have a way to investigate and take actions against the policy match outcomes.	- Incident reports: **Send an alert to admins when a rule match occurs** is on by default
...Lastly, we want this to take effect as soon as possible....	Policy mode: **on **

Steps to set up prerequisites

Complete the following steps to set up the prerequisites before you create the DLP policy:

1. Onboard apps to Conditional Access app control. For more information, please see Onboard non-Microsoft IdP catalog apps for Conditional Access app control and Onboard non-Microsoft IdP custom apps for Conditional Access app control.

2. Import user groups. Sign into the Microsoft Defender XDR, select **Settings** > **Cloud Apps** > **System** > **User groups** > **+ Import.** For more information, please see Import user groups from connected apps.

3. Sign into the Microsoft Entra admin center.

4. Create a new Conditional Access policy that targets cloud apps with session controls and using the and using the client apps condition for browser..
   a. **Custom policy** must be selected in the session controls dropdown.
   b. **Browser** must be selected in **Conditions** > **Client Apps**

5. In the Microsoft Defender portal at https://security.microsoft.com), go to System > Settings > Cloud apps > Conditional Access App Control section > Edge for Business protection. Or, to go directly to the Edge for Business protection page, use https://security.microsoft.com/cloudapps/settings?tabid=edgeIntegration .

6. Configure in-browser protection for Edge on, selecting **Allow access only from Edge** and **All devices**

> **Important**
>
> To include groups in policies targeting managed cloud apps, you must first **import user groups from connected apps**.

Steps to create the policy

1. Sign in to the Microsoft Purview portal.
2. Select **Data loss prevention** > **Policies** > **+ Create policy**.
3. Select **Enterprise apps and Devices**.
4. Select **Custom** from the **Categories** list and then **Custom policy** from the **Regulations** list.
5. Choose **Next**.
6. Provide a policy name and give a description. You can use the policy intent statement here.

7. Select **Next**.
8. Accept the default **Full directory** on the **Assign admin units** page.
9. Choose **Next**.
10. Select **Managed cloud apps**
11. Select **Specific users and groups**.
12. Choose + **Include** and then **Include groups**.
13. Select **External**.
14. Choose **Done** and then choose **Next**.
15. Select + **Include managed apps**.
16. Select **Workday**
17. Choose **Done**
18. On the **Define policy settings** page, the **Create or customize advanced DLP rules** option should already be selected.
19. Choose **Next**.
20. On the **Customize advanced DLP rules** page, select + **Create rule**.
21. Name the rule and give a description.
22. Select **Add condition** and use these values:
    a. Select **Content contains**.
    b. Select **Add** > **Sensitive information types** > **Sensitive info types** > **ABA Routing Number**, **Australia Bank Account Number**, **Canada Bank Account Number**, **International Bank Account Number (IBAN)**, **Israel Bank Account Number**, **Japan Bank Account Number**, **New Zealand bank account number**, **SWIFT Code**, **U.S Bank Account Number**.
23. Choose **Add**.
24. Under **Actions**, add an action with these values:
    a. **Restrict browser and network activities**
    b. **File download** > **Block**
25. Under **Incident reports** select: 2. The toggle for **Send an alert to admins when a rule match occurs** defaults to **On**.
26. Choose **Save** and then choose **Next**.
27. On the **Policy mode** page, choose **On**.
28. Choose **Next** and then choose **Submit**.
29. Choose **Done**.

> **Important**
>
> The DLP policy won't be applied in Edge until all requirements are met.

Steps to activate Microsoft Edge for Business Protected Clipboard and Screen Capture protections (optional)

See the Get started with Protected Clipboard for Microsoft Edge for Business instructions.

Steps to review outcomes of a Microsoft Purview policy in Microsoft Defender

See Investigate and respond in Microsoft Defender XDR for instructions on how to review and act on policy match outcomes.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Learn about investigating data loss prevention alerts

This article introduces you to the alert investigation flow and the tools you can use to investigate DLP alerts.

Before you begin

1. Administrative units
2. Learn about Microsoft Purview Data Loss Prevention
3. Plan for data loss prevention (DLP) - by working through this article you will:
   a. Identify stakeholders
   b. Describe the categories of sensitive information to protect
   c. Set goals and strategy
4. Collection Policies solution overview
5. Collection policy reference
6. Data Loss Prevention policy reference - this article introduces all the components of a DLP policy and how each one influences the behavior of a policy
7. Design a DLP policy - this article walks you through creating a policy intent statement and mapping it to a specific policy configuration.
8. Create and Deploy data loss prevention policies - This article presents some common policy intent scenarios that you map to configuration options, then it walks you through configuring those options.
9. Learn about investigating data loss prevention alerts - This article introduces you to the lifecycle of alerts from creation, through final remediation and policy tuning. It also introduces you to the tools you use to investigate alerts.

The lifecycle of a DLP alert

All alerts and your interaction with them go through these six steps:

- Trigger
- Notify
- Triage
- Investigate
- Remediate
- Tune

Trigger

The life of a Microsoft Purview Data Loss Prevention (DLP) alert starts when the conditions defined in the policy are matched. When a policy match occurs, the actions defined in the policy are triggered, which can include generating an alert if the policy is configured to do so.

DLP policies are typically configured to monitor for and generate alerts when:

- Sensitive information, such as personally identifying data or intellectual property, is exfiltrated from your organization.
- Sensitive information is shared inappropriately with people outside or inside your organization.
- Users engage in risky activities, such as downloading sensitive information to removable media.

Notify

When an alert is generated, it's sent to the Microsoft Defender portal as an incident and the DLP alert management dashboard. DLP policies can be configured to send notifications to users, administrators, and other stakeholders via email.

In the notify phase Microsoft Purview:

- Reports on DLP policy matches and user overrides.
- You can use Activity explorer to view DLP related activities and filter for report generation purposes.

To export activity data for reporting use Export-ActivityExplorerData (ExchangePowerShell) | Microsoft Doc by using O365 Management Activity API or Incident API.

> **Note**
>
> The Microsoft Defender portal retains incidents for six months. The DLP alert management dashboard retains alerts for 30 days.

Triage

In this step, you analyze an alert and any associated logs and decide if the alert is a true positive or a false positive. If it's a true positive, you set the priority of the alert based on the severity of the issue and its impact on your organization and assign an owner. If it's a false positive, you can unblock the user and move on to the next alert.

Defender portal groups DLP events into incidents. Incidents are a collection of related alerts that are grouped together based on all the other signals that Defender is receiving. For

example, when you have a DLP policy configured to monitor and alert on sensitive files on SharePoint sites, and a user downloads a file from a SharePoint site and then uploads it to a personal OneDrive, and then shares it with an external user, Defender groups all of those alerts into a single incident. This is a powerful feature that allows you to focus on the most important alerts first.

In the Defender portal you can immediately start triaging incidents and use tags, comments, and other features to structure your incident management. You should be utilizing the Incidents page in the Microsoft Defender portal to manage your DLP alerts. You can filter the Incidents queue to view all incidents with Microsoft Purview DLP alerts by selecting **Filters** and choose **Service Source: Data Loss Prevention**.

If you have enabled sharing of insider risk management data with Microsoft Defender XDR (preview) - you'll see the severity level of the Insider Risk Management policy that is associated with a user in the DLP alerts page. Insider Risk Management severity levels are: **Low**, **Medium**, **High**, and **None**. You can use this information to prioritize your investigation and remediation efforts. This information will also be available in the Microsoft 365 Defender portal in the details of the incident.

Triage with Microsoft Security Copilot

Another tool you can use to help triage alerts it Microsoft Security Copilot. Security Copilot is a cloud-based AI platform that can assist security and compliance professionals in protecting their organization's data. It is available in the DLP Alerts dashboard and in Data Security Posture Management.

Triage with Microsoft Security Copilot DLP Alert Triage Agent (preview)

In preview, Microsoft Security Copilot for Purview supports the DLP Alert Triage Agent. Microsoft Security Copilot agents are AI powered processes that are designed to help you with specific role based tasks.

Investigate

The main goal of the investigation stage is for the assigned owner to correlate evidence, determine the cause and full impact of the alert and decide on a remediation plan. The assigned owner is responsible for deeper investigation and remediation of the alert. The primary alert investigation tools are the Microsoft Defender portal and the DLP alert management dashboard. You might also use Activity explorer to investigate alerts. You can also share alert events with other users in your organization.

You can take advantage of DLP features like:

- Evidence collection for file activities on devices makes files like email and documents that matched a policy easily accessible.
- Use Content explorer to deeply investigate the content in the incident.

You can use both Microsoft Defender portal and Purview tools to triage and investigate alerts, but the Microsoft Defender portal provides more capabilities for managing alerts and incidents, such as:

- View all your DLP alerts grouped under incidents in the Microsoft Defender XDR incident queue.
- View intelligent inter-solution (DLP-MDE, DLP-MDO) and intra-solution (DLP-DLP) correlated alerts under a single incident.
- Hunt for compliance logs along with security under Advanced Hunting.
- In-place admin remediation actions on user, file, and device.
- Associate custom tags to DLP incidents and filter by them.
- Filter by DLP policy name, tag, Date, service source, incident status, and user on the unified incident queue.

If you're sharing insider risk management data with Defender (preview), you can see the **User activity summary** of all the exfiltration activities the user has engaged in up to the past 120 days.

Remediate

Your remediation plan is unique to your organization's policies, the industry, the geopolitical regulations it must comply with, and business practices. How your organization chooses to respond to an alert revolves around the accuracy of the alert (true positive, false positive, false negative), the severity of the issue, and the impact on your organization.

Remediation actions can include:

- Monitor only, no further action required.
- No further action required because the actions take by the policy sufficiently mitigated the risk.
- Risk mitigated by automated policy actions but user education is necessary.
- The issue wasn't fully mitigated by the policy, so further clean up and risk mitigation is required along with more user training.
- Via Adaptive Protection in Data Loss Prevention (preview) where DLP integrates with Insider Risk Management, you can assign a risk level to the user for further monitoring and actions.

With the Defender portal, you can immediately take remediation actions on alerts and incidents. For example:

- Reset password
- Disable account
- View user activity
- Actions on DLP detections
- Remove document
- Apply sensitivity label
- Unshare
- Download email
- Advanced Hunting
- Isolate Device
- Collect investigation pack from Device
- Run AV Scan
- Quarantine file
- Disable user
- Reset pwd
- Delete email
- Move mail to other mailbox folder
- Download file

Tune

Based on the accuracy and effectiveness of your policy, you might need to update it so it remains effective. You've already tuned your policy during the policy creation and deployment process, but as your data estate and business needs change, policies have to be updated to continue to be effective. These changes are best tracked in the policy intent statement and the policy configuration.

Items you tune:

- The scope of the policy.
- The conditions required for a policy match.
- The actions taken when a policy match occurs.
- Notifications sent to users and administrators.

For more information about mapping business needs to policy design and testing policies, see:

- Design a data loss prevention policy
- Test your Data Loss Prevention policies

Toolsets

There are multiple tools that you can use to investigate and manage Microsoft Purview Data Loss Prevention (DLP) alerts. There are:

- Microsoft Defender portal
- Microsoft Purview portal Alerts dashboard
- Activity explorer
- Content explorer
- Key features in the embedded experience
- Key features in the standalone experience
- Data Security Posture Management

Microsoft recommends using the unified incident queue in Microsoft Defender portal to manage your DLP alerts. However, your organization may have needs that can be met by using the DLP alert management dashboard in addition to the Microsoft Defender portal.

Microsoft Defender portal

- DLP alerts are integrated with other events and alerts into a single incident queue, which provides a more complete picture of the incident.
- Six months of incident history is available.
- Advanced hunting is available.
- Investigate data loss incidents with Microsoft Defender XDR - You can manage DLP incidents along with security incidents from **Incidents & alerts** > **Incidents** on the quick launch of the Microsoft Defender portal.
- Responding to your first incident in Microsoft Defender XDR
- Incident response with Microsoft Defender XDR
- Prioritize incidents in Microsoft Defender XDR
- Manage incidents in Microsoft Defender XDR
- Investigate incidents in Microsoft Defender XDR
- Investigate alerts in Microsoft Defender XDR
- Proactively hunt for threats with advanced hunting in Microsoft Defender XDR

Microsoft Purview portal

- Alerts dashboard, Activity explorer, and Content explorer are all available in the Microsoft Purview portal. You can summarize alerts using Microsoft Security Copilot Investigate a DLP alert
- You can set an alert status to **Investigating**.
- You can share alert events with other users in your organization.

- Download files from OneDrive and SharePoint (data classification content viewer role is required for this action)

If you're new to using the DLP Alerts dashboard, you should read through these articles to help you get started.

- Get started with the data loss prevention Alerts dashboard
- Share data loss prevention alert events (preview)
- Get started with data loss prevention alerts
- Get started with activity explorer
- Get started with content explorer

Next steps

- Investigate data loss prevention alerts with Microsoft Defender XDR
- Get started with the data loss prevention Alerts dashboard

Last updated on 11/18/2025

Get started with the Data Loss Prevention Alerts dashboard

Microsoft Purview Data Loss Prevention (DLP) policies can take protective actions to prevent unintentional sharing of sensitive items. You can be notified when an action is taken on a sensitive item by configuring alerts for DLP.

You use the DLP alert management dashboard in the Microsoft Purview portal to view alerts, events, and associated metadata for DLP policy violations.

If you're new to DLP alerts, you should review Get started with the data loss prevention alerts.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about **Microsoft Security Copilot in Microsoft Purview**.

Microsoft Purview portal shows alerts for DLP policies that are enforced on the following workloads:

- Exchange email
- SharePoint sites
- OneDrive accounts
- Teams chat and channel messages
- Devices
- Instances
- On-premises repositories
- Fabric and Power BI

Prerequisites and permissions

Make sure you have the necessary prerequisites:

For information on licensing, see

- Microsoft 365 Enterprise Plans
- Microsoft 365 Service Descriptions

Roles and Role Groups

If you want to view the DLP alert management dashboard or to edit the alert configuration options in a DLP policy, you must be a member of one of these role groups:

- Compliance Administrator
- Compliance Data Administrator
- Security Administrator
- Security Operator
- Security Reader
- Information Protection Admin
- Information Protection Analyst
- Information Protection Investigator

To learn more about these role groups, see Permissions in the Microsoft Purview portal.

Here's a list of applicable role groups:

- Information Protection
- Information Protection Admins
- Information Protection Analysts
- Information Protection Investigators

To access the DLP alert management dashboard, you need the *Manage alerts* role and either of these two roles:

- DLP Compliance Management
- View-Only DLP Compliance Management

To access the Content preview feature and the Matched sensitive content and context features, you must be a member of the *Content Explorer Content Viewer* role group, which has the *Data classification content viewer* role preassigned.

Configure DLP alerts

To learn how to configure an alert in your DLP policy, see Create and Deploy data loss prevention policies.

> **Important**

Your organization's audit log retention policy configuration controls how long an alert remains visible in the console. See Manage audit log retention policies for more information.

Investigate a DLP alert

There are three tools you can use to investigate DLP alerts in the DLP Alerts dashboard:

- The **DLP Alerts** dashboard in standard view
- The **DLP Alerts** dashboard in Alert Triage Agent (Preview)
- Microsoft Security Copilot

Standard view

To investigate an alert using the standard DLP Alerts dashboard view, follow these steps:

1. Sign in to the Microsoft Purview portal > **Data loss prevention**.
2. Select **Alerts** to view the **DLP Alerts** dashboard.
3. Use the **Filter** fields to refine the list of alerts.
4. Choose **Customize columns** to list the properties you want to see.
5. To sort the results in ascending or descending order, double-click the column header.
6. Double-click an alert for more information about it.
7. The **Details** tab opens by default and provides high-level information about the alert.
8. Select **Summarize with Copilot**. This causes the Microsoft Security Copilot to generate a summary of the alert. The alert summary will contain the:

    - alert severity
    - alert title
    - the name of the policy that was matched
    - the name of the file involved and a link to the file
    - alert status
    - the email address of the user who performed the action that matched the policy

9. Select the ellipsis in the **Security Copilot** summary to:

    - copy the summary to the clipboard
    - regenerate the summary
    - open the alert in the Security Copilot standalone experience.

10. Select **View details** to open the **Overview** tab. The **Overview** tab provides a summary of the following information:

    - What happened
    - Who performed the actions that caused the policy match
    - Additional information about the policy match

11. The **Events** tab lists all of the events associated with the alert. Select any event in the list to get detailed information about the event. For each event, choose the **Actions** drop down for a list of actions you can take on the alert, such as verifying whether or not the alert has identified a true match or a false positive.
12. The **User activity summary** tab requires that sharing be turned *On* in the **insider risk management** settings Once enabled, the **User activity summary** tab provides all the exfiltration activities the user has engaged in (up to the past 120 days). Users must be in scope of an insider risk management policy to see the **User activity summary** tab.
13. After you investigate the alert, return to the **Overview** tab where you can **View details** to triage and manage the disposition of the alert, add comments and assign ownership of the alert. (To see the history of workflow management.)
14. After you take the required action for the alert, set the **Status** of the alert to **Resolved**.

Alert Triage Agent

In preview, Microsoft Security Copilot in Purview supports Security Copilot Agents (preview). The Alert Triage Agent helps reduce the time it takes to triage alerts by reasoning over the alerts and identifying the ones that pose the greatest risk to your organization.

1. Sign in to the Microsoft Purview portal > **Data loss prevention** with appropriate permissions, see DLP agent permissions and roles for Data Loss Prevention workloads or IRM agent permissions and roles for Insider Risk Management workloads.

2. Select **Alerts** to view the **DLP Alerts** dashboard.

3. Set the Alerts page view toggle to **Alert Triage Agent**.

4. Choose the category (**All**, **Needs attention**, **Less urgent**, or **Not categorized**) to see the alerts that are relevant to your task.

5. Select the triaged alert to view the **Agent summary**, an **Overview**, and **Events** associated with the alert.

> **Note**
>
> If an alert contains multiple assets, the agent will summarize the top 10 assets that contributed to the summary of findings and the decision on which category to assign the alert to.

6. After you review the summary, you can select **Manage alert** to assign it to an analyst, change the status of the alert or add more comments.

Security Copilot

Security Copilot is a cloud-based AI platform that can assist security and compliance professionals in protecting their organization's data. You can use it to summarize Microsoft Purview alerts, triage alerts and to drill down into your Microsoft Purview data. It is available in the DLP Alerts dashboard and in Data Security Posture Management.

Create a shareable link to an alert event

Users with the appropriate permissions can view Microsoft Purview Data Loss Prevention (DLP) alerts in the DLP Alerts console. However, as alerts are triaged and investigated, you may want to share alert events with users who don't have access to DLP and the alerts console. You can create a shareable link to an alert event:

1. In the Purview Portal, navigate to **Data Loss Prevention**, and then select **Alerts**.
2. Select an alert, and then select **View details**.
3. On the Alert details screen, select the **Events** tab just below the alert title to view the events contained in this alert.
4. Select the event you want to share, and then click the **Actions** button at the bottom of the screen to see the **Actions** menu.
5. Select Copy event link, and then select Copy to copy the shareable link to the event.
6. Now you can paste the link from your clipboard to share the link via chat, email, or other means.

View other matched conditions in DLP events

Microsoft Purview supports showing matched conditions in a DLP event to reveal the exact cause for a flagged DLP policy. This information shows up in:

- DLP Alerts console

- Activity explorer
- Microsoft Defender for Business portal

In the **Events** tab, open **Details** to see **Other matched conditions**.

Enhanced matched conditions for Exchange Online (preview)

For Exchange Online DLP events, matched conditions include enhanced detail for non-sensitive information type (SIT) conditions in addition to SIT matches. For the full list of supported condition categories and detail levels, see Enhanced matched conditions for Exchange DLP events.

Prerequisites

Make sure the following requirements are met before you use matched conditions on endpoints:

- Must be running Windows 10 x64 (build 1809 or later) or Windows 11.
  - See March 21, 2023—KB5023773 (OS Builds 19042.2788, 19044.2788, and 19045.2788) Preview for required minimum Windows Operating System builds.
- Matched conditions data is available for valid E3 and E5 license holders.
- Enable Auditing.
- Enable Advanced classification scanning and protection.

Matched events information is supported for these conditions

Expand table

Condition	Exchange	Sharepoint	Teams	Endpoint
Sender is	Yes	No	Yes	No
Sender domain is	Yes	No	Yes	No
Sender address contains words	Yes	No	No	No
Sender address matches patterns	Yes	No	No	No
Sender is a member of	Yes	No	No	No
Sender IP address is	Yes	No	No	No
Has sender overridden the policy tip	Yes	No	No	No
SenderAdAttribute Contains words	Yes	No	No	No
SenderAdAttribute Matches patterns	Yes	No	No	No

Condition	Exchange	Sharepoint	Teams	Endpoint
Recipient is	Yes	No	Yes	No
Recipient domain is	Yes	No	Yes	No
Recipient address contains words	Yes	No	No	No
Recipient address matches patterns	Yes	No	No	No
Recipient is a member of	Yes	No	No	No
RecipientAdAttribute Contains words	Yes	No	No	No
RecipientAdAttribute Matches patterns	Yes	No	No	No
Document is password protected	Yes	No	No	No
Document could not be scanned	Yes	No	No	No
Document did not complete scanning	Yes	No	No	No
Document name contains words	Yes	Yes	No	No
Document name matches patterns	Yes	No	No	No
Document property is	Yes	Yes	No	No
Document size over	Yes	Yes	No	No
Document content contains words	Yes	No	No	No
Document content matches patterns	Yes	No	No	No
Document type is	No	No	No	Yes
Document extension is	Yes	Yes	No	Yes
Content is shared from M365	Yes	Yes	Yes	No
Content is received from	Yes	No	No	No
Content character set contains words	Yes	No	No	No
Subject contains words	Yes	No	No	No
Subject matches patterns	Yes	No	No	No
Subject or body contains words	Yes	No	No	No
Subject or body matches patterns	Yes	No	No	No
Header contains words	Yes	No	No	No

Condition	Exchange	Sharepoint	Teams	Endpoint
Header matches patterns	Yes	No	No	No
Message size over	Yes	No	No	No
Message type is	Yes	No	No	No
Message importance is	Yes	No	No	No

Limitation when downloading emails from within a DLP alert

In general, when using the DLP alert management dashboard, you can download specific emails from within an alert. However, emails that have been deleted in any of the following scenarios can't be downloaded.

Expand table

Sender	Recipient	Email Status
Internal	External	Deleted by sender
External	Internal	Deleted by recipient
Internal	Internal	Deleted by both parties

Related content

Use the following resources to learn more about DLP alerts, activity exploration, and data security posture management:

- Get started with data loss prevention alerts
- Share data loss prevention alerts
- Get started with activity explorer
- Learn about Data Security Posture Management

ⓘ**Note:** The author created this article with assistance from AI. Learn more

Last updated on 06/26/2026

Get started with the data loss prevention alerts

Your Microsoft Purview Data Loss Prevention (DLP) policies can be configured to generate alerts when the conditions in a rule are matched. Alerts are configured in DLP policy rules.

For a brief overview of alerts see:

- DLP Alerts
- Incident reports

This article includes the licensing and permission details and other crucial information you need as you work with alerts.

DLP alerts can be investigated and managed in the Microsoft Defender XDR dashboard and in the Microsoft Purview portal. The Microsoft Defender XDR dashboard is the recommended location for investigating and managing DLP alerts. The Microsoft Purview portal is the recommended location for creating and editing DLP policies.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about **Microsoft Security Copilot in Microsoft Purview**.

Types of DLP alerts

Alerts can be sent every time an activity matches a rule or they can be aggregated to cut down on the noise. There are two types of alerts that can be configured in DLP policies.

Single-event alerts Single-event alerts are typically used in policies that monitor for highly sensitive events that occur in a low volume, like a single email with 10 or more customer credit card numbers being sent outside your organization. By default when events from a single rule occur in a one minute window of each other they are aggregated for E5 license and 15 mins for E3 license. In preview, single-event alerts can be aggregated on per rule per user basis. This is called user and rule-based alert aggregation.

Aggregate-event alerts are typically used in policies that monitor for events that occur in a higher volume over a period of time. For example, an aggregate alert can be triggered when 10 individual emails each with one customer credit card number is sent outside your org over 48 hours.

Before you begin

Before you begin, make sure you have these prerequisites:

Licensing for alert configuration options

The following licensing requirements apply to each alert configuration option:

- **Single-event alert configuration**: All organizations with a DLP subscription (E1, E3, E5, F1, G1, E3, G3, E5, G5) can configure policies to generate an alert for every time a triggering activity occurs.
- **Aggregated alert configuration**: To configure aggregate alert policies based on a threshold, you must have either of the following configurations:
  - An A5 subscription
  - An E5 or G5 subscription
  - An E1, F1, or G1 subscription or an E3 or G3 subscription that includes one of the following features:
    - Office 365 Advanced Threat Protection Plan 2
    - Microsoft Purview Suite (formerly known as Microsoft 365 E5 Compliance)
    - Microsoft 365 eDiscovery and Audit add-on license

Customers who use endpoint DLP and who are eligible for Teams DLP will see their endpoint DLP policy alerts and Teams DLP policy alerts in the DLP alert management dashboard.

Roles and Role Groups

If you want to view the DLP alert management dashboard or to edit the alert configuration options in a DLP policy, you must be a member of one of these role groups:

- Compliance Administrator
- Compliance Data Administrator
- Security Administrator
- Security Operator
- Security Reader
- Information Protection Admin

- Information Protection Analyst
- Information Protection Investigator

To learn more about these role groups, see Permissions in the Microsoft Purview portal

Here's a list of applicable role groups. To learn more about these role groups, see Permissions in the Microsoft Purview portal.

- Information Protection
- Information Protection Admins
- Information Protection Analysts
- Information Protection Investigators

To access the DLP alert management dashboard, you need the *Manage alerts* role and either of these two roles:

- DLP Compliance Management
- View-Only DLP Compliance Management

To access the Content preview feature and the Matched sensitive content and context features, you must be a member of the *Content Explorer Content Viewer* role group, which has the *Data classification content viewer* role preassigned.

> **Tip**
>
> If the admin requires access to alerts but not contextual/sensitive information, you can create and assign a custom role that does not include Data Classification Content Viewer permission.

Configure DLP alerts

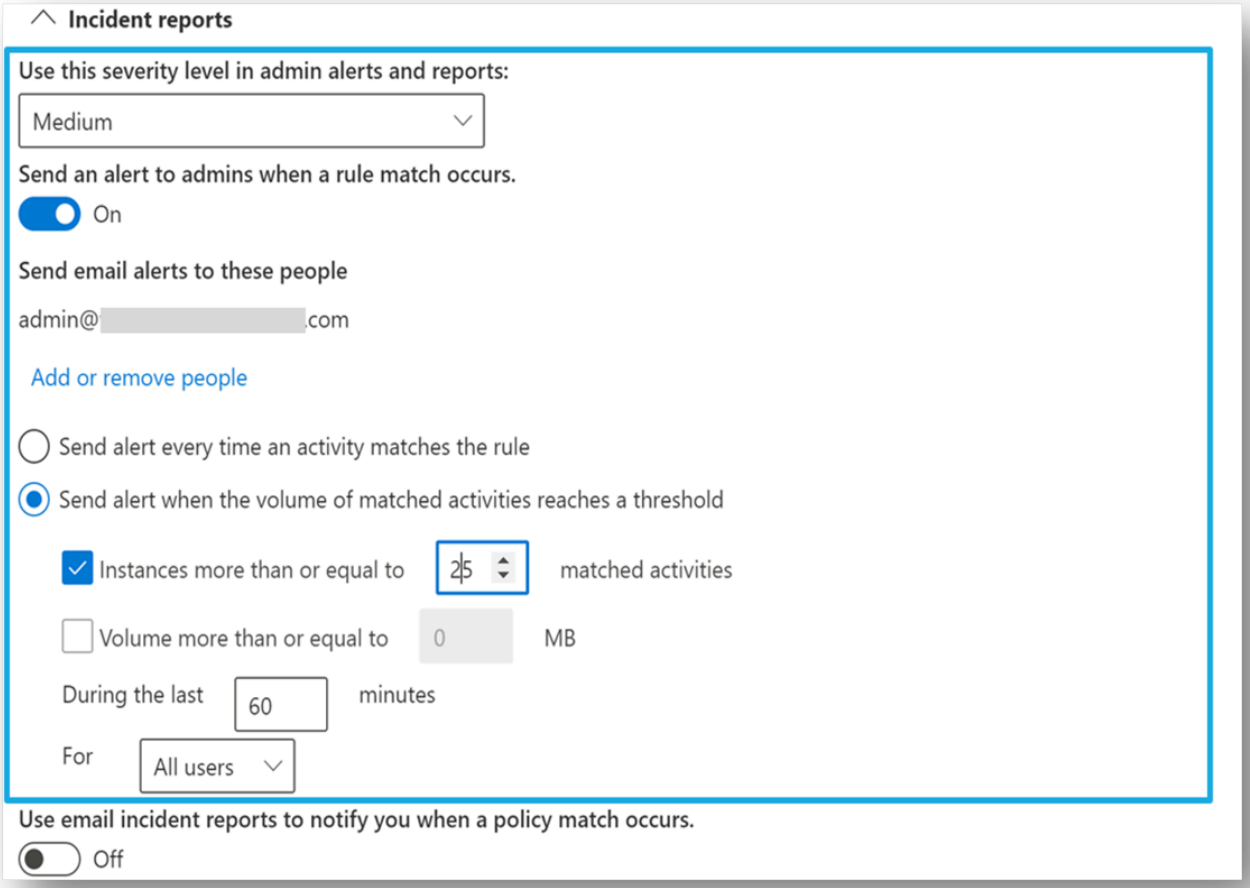
To learn how to configure an alert in your DLP policy, see Create and Deploy data loss prevention policies. There are different alert configuration experiences depending on your licensing.

> **Note**
>
> It may take up to 3 hours to generate alerts after you configure or modify existing alerts in a DLP policy.

An Alert email, Incident Report email, and User Notification will only be sent once per document. If a document with a **Content is Shared** condition is shared twice, there will still be only one notification.

Aggregate event alert configuration

If you're licensed for aggregated alert configuration options, then you'll see options to generate alerts for each rule match, when thresholds are met, and within a defined time window when you create or edit a DLP policy.

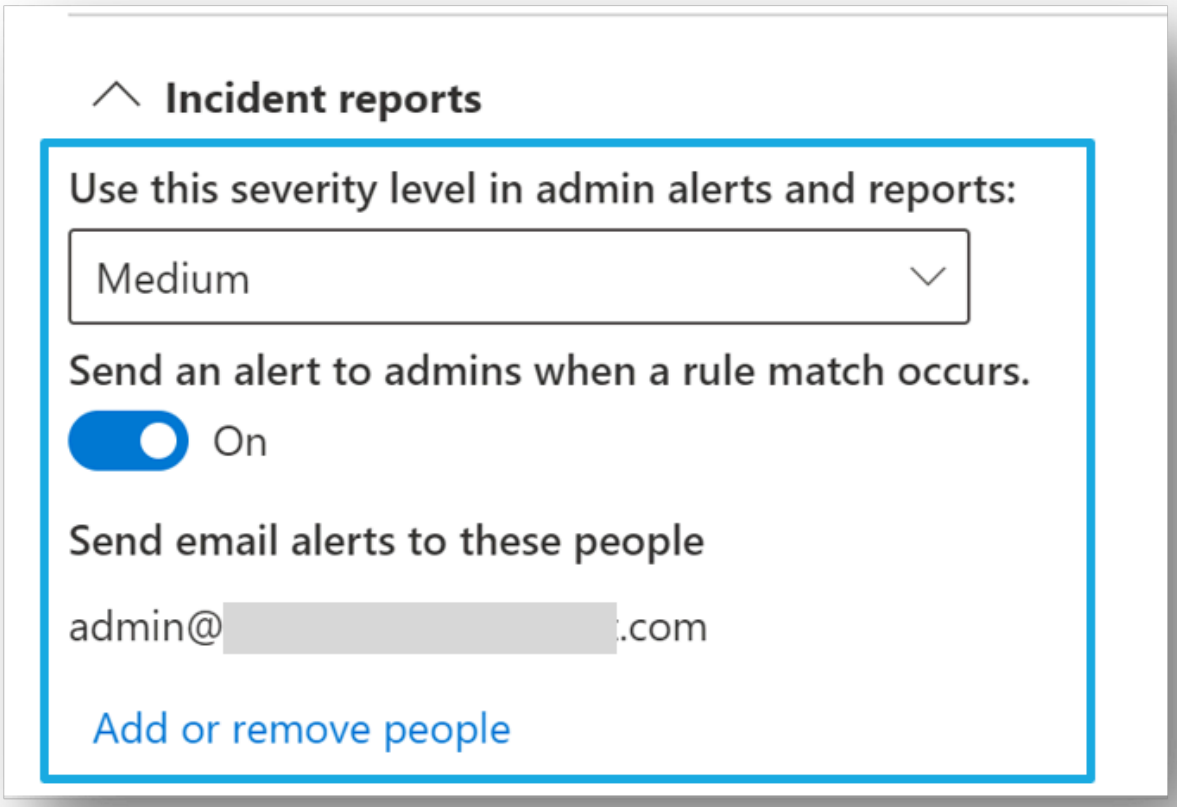


This configuration allows you to set up a policy to generate an alert:

- every time an activity matches the conditions in a rule for default aggregation based on a rule or of user and rule based aggregation.
- when the defined threshold is met or exceeded based on the number of matches or the volume of or on the volume of exfiltrated data
- for activities that meet the criteria that you define within a speficied time window

Single event alert configuration

If you're licensed for single-event alert configuration options, then you'll see the single-event alert settings when you create or edit a DLP policy. Use this option to create an alert that's raised every time a DLP rule match happens.



User and rule based alert aggregation (preview)

When you enable user based alert aggration in DLP setting at the tenant level, single event alerts are aggregated based on users. The rule match events must occur within the configurable time window (15, 30, 45 and 60 minutes). Alerts are generated per rule match event per user.

Compare alert aggregation options

Expand table

I want DLP to...	Time window	Aggregation type	Notes
...generate single alert when an email containing credit card information is sent by any number of users.	These windows *are not* configurable by the admin -E5: 60 seconds -E3: 15 minutes	- User based alert aggregation is set to **Off** at the tenant level. - Single event alert aggregation configured	- Applies to multiple users for one rule.

I want DLP to...	Time window	Aggregation type	Notes
		at the rule level and is available if match occurs within time window. -Aggregate DLP rule matches of multiple users into one alert.	
...generate one alert for each email sender when an email containing credit card information is sent out.	This time window *is configurable* by the admin at the tenant level. - 15 - 60 minutes	- User based alert aggregation is set to **On** at the tenant level. - Single event alert aggregation configured at the rule level. -Aggregate DLP rule matches by single user into one alert.	- For single user per rule. - Expect an increase in alert volume. - If the alert is closed within the aggregation time window and a new match occurs for the same user and rule, then the new rule match will be aggregated into the same alert.
...to generate an alert when more than 100 sensitive files are accessed by multiple users within 60 minutes.	- Configured at the rule level, 60-999 minutes.	- Threshold based aggregation - Aggregate DLP rule matches based on count of matches. - User based alert aggregation not applicable.	- pivots on rules - Use this for multiple users for one rule. - Only works for the **All users** option.
...to generate an alert when more than 25 MB data is exfiltrated by multiple users within 60 minutes.	Configured at the rule level, 60-999 minutes.	-Threshold based aggregateon based on the volume of data. - user based alert aggregation not applicable.	- pivots on rules - Use this for multiple users for one rule. - Only works for the **All users** option.

Types of DLP alert events

Here are some of the events associated with an alert. In the Alert dashboard, you can choose a particular event to view its details.

Review event details

The following table describes the event details shown for alert events.

Expand table

Property name	Description	Event types
ID	unique ID associated with the event	all events
Location	workload where the event was detected	all events
time of activity	time of the user activity that matched the criteria of the DLP policy	

Review affected entities in an alert

The following table lists the entities that can be associated with a DLP alert event.

Expand table

Property name	Description	Event types
user	user who took the action that caused the policy match	all events
hostname	host name of the computer where the DLP policy match occurred	device events
IP address	IP address of the computer where the DLP policy match occurred	device events
sha1	SHA-1 hash of the file	device events
sha256	SHA-256 hash of the file	device events
MDATP device ID	endpoint device MDATP ID	
file size	size of the file	SharePoint, OneDrive, and device events
file path	the absolute path of the item involved with the DLP policy match	SharePoint, OneDrive, and devices events
email recipients	if an email was the sensitive item that matched the DLP policy, this field includes the recipients of that email	Exchange events
email subject	subject of the email that matched the DLP policy	Exchange events
email attachments	names of the attachments in the email that matched the DLP policy	Exchange events
site owner	name of the site owner	SharePoint and OneDrive events
site URL	full of the URL of the SharePoint or OneDrive site where the	SharePoint and OneDrive

Property name	Description	Event types
	DLP policy match occurred	events
file created	time of creation of the file that matched the DLP policy	SharePoint and OneDrive events
file last modified	the last time that the file that matched the DLP policy was changed	SharePoint and OneDrive events
file size	size of the file that matched the DLP policy	SharePoint and OneDrive events
file owner	owner of the file that matched the DLP policy	SharePoint and OneDrive events

Review policy details for an alert

The following table describes the policy-related details available for alert events.

Expand table

Property name	Description	Event types
DLP policy matched	name of the matched DLP policy	all events
rule matched	name of the matched DLP policy rule	all events
sensitive information types (SIT) detected	SITs that were detected as part of the DLP policy match	all events
actions taken	actions that were taken that caused the DLP policy match	all events
violating action	action on the endpoint device that raised the DLP alert	device events
user overrode policy	did the user override the policy via a policy tip	all events
use override justification	the text of the reason provided by the user for the override	all events

Important

Your organization's audit log retention policy configuration controls for how long an alert remains visible in the console. For more information, see **Manage audit log retention policies**.

See also

- Alerts in DLP policies: Describes alerts in the context of a DLP policy.
- Get started with data loss prevention alerts: Covers the necessary liscensing, permissions, and prerequisites for DLP alerts and alert reference details.
- Create and deploy data loss prevention policies: Includes guidance on alert configuration in the context of creating a DLP policy.
- Learn about investigating data loss prevention alerts: Covers the various methods for investigating of DLP alerts.
- Investigate data loss incidents with Microsoft Defender XDR: How to investigate DLP alerts in Microsoft Defender portal.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Get started with the data loss prevention analytics

Microsoft Purview data loss prevention (DLP) analytics helps customers understand top data protection risks, blind spots, and policy and posture improvement opportunities in their organization. It can help you investigate these risks using intelligent Purview features, and mitigate them in a few simple steps.

This section introduces DLP analytics concepts, prerequisites, and configuration steps to start using it.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

DLP analytics shows customers their top data loss risks and vulnerabilities and how to mitigate them through new or improved policies. DLP analytics helps mitigate risks in three ways:

- It communicates the top item oversharing risks through the **Risk Spotlighting** card on the DLP overview page in the Microsoft Purview portal. Analytics reports on top risks, blind spots, and policy improvement opportunities based on past 30 days data.
- Help prevent users from sharing additional sensitive information externally by creating new DLP policies with one click. Policy creation recommendations are based on industry best practices and risks found in the tenant.
- Improve the accuracy of existing policies via the **Policy Improvement** card with one click.

Risks and recommendations are refreshed every week.

When DLP analytics is enabled, it scans signals on user activity, sensitive data sharing patterns, and policy information to generate insights that help you set up and refine DLP policies. It takes seven days to generate recommendations after you turn on DLP analytics.

Before you begin

Licensing

For information on licensing, see

- Microsoft 365 Enterprise Plans
- Microsoft 365 Service Descriptions

Permissions

To see DLP analytics your account must be one of these two roles:

- Global admin
- Compliance Administrator

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. This helps improve security for your organization. Global Administrator is a highly privileged role that should only be used in scenarios where a lesser privileged role can't be used.

Enable DLP risk-detection analytics

DLP analytics is turned **Off** by default. You need to explicitly opt in. Scans run every hour on your data present in audit logs and policy data.

> **Note**
>
> It takes seven days to start seeing analytics and recommendations in your tenant once analytics are re-enabled.

1. Sign in to the Microsoft Purview portal > **Data loss prevention** > **Overview**
2. Check the **Turn on analytics for risk detection and policy refinement opportunities (preview)** option.
3. Select **Turn on analytics**

Disable DLP analytics

After disabling DLP analytics, it can take up to 24 hours for the insights to stop appearing on the Data loss prevention **Overview** page.

1. Sign in to the Microsoft Purview portal > **Settings** (gear in the menu bar) > **Data Loss Prevention** > **Analytics (preview)**.
2. Set the **Activate analytics** toggle to **Off**.

View policies created by DLP analytics

To identify policies created by DLP analytics, perform the following steps:

1. Sign in to the Microsoft Purview portal > **Data Loss Prevention** > **Policies**.
2. Look for policy names in this format: *RiskSpotlighting-YYYY-MM-DD*.

Review policies updated by DLP analytics

When you tell DLP analytics to update an existing policy, it makes all updates in a new version of the policy. The name of the original policy is appended with *_copy* and the original policy is turned off. The new version is then deployed. For example:

1. There's a policy named *All credit cards* that is generating too many false positives.
2. DLP analytics suggests changes to the policy to reduce false positives.
3. You tell DLP analytics to update the policy.
4. DLP analytics creates a new version of the policy named *All credit cards*
5. DLP analytics renames the original version of the policy to *All credit cards_copy* and sets its status to **Keep it off**.
6. DLP analytics deploys the new version of the policy named *All credit cards* and sets it status to **Turn it on**.

Related content

- Design a DLP policy
- Create and deploy a DLP policy

ⓘ**Note:** The author created this article with assistance from AI. Learn more

Last updated on 06/26/2026

Design a data loss prevention policy

Taking the time to design a policy before you implement it gets you to the desired results faster, with fewer unintended issues, than creating it and then tuning by trial and error alone. Having your policy designs documented will help you in communications, policy reviews, troubleshooting, and further tuning.

If you're new to Microsoft Purview DLP, it's helpful to work through these articles before you start designing a policy:

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about **Microsoft Security Copilot in Microsoft Purview**.

Before you begin

1. Administrative units
2. Learn about Microsoft Purview Data Loss Prevention
3. Plan for data loss prevention (DLP) - by working through this article you will:
   a. Identify stakeholders
   b. Describe the categories of sensitive information to protect
   c. Set goals and strategy
4. Collection Policies solution overview
5. Collection policy reference
6. Data Loss Prevention policy reference - this article introduces all the components of a DLP policy and how each one influences the behavior of a policy
7. Design a DLP policy - this article walks you through creating a policy intent statement and mapping it to a specific policy configuration.
8. Create and Deploy data loss prevention policies - This article presents some common policy intent scenarios that you map to configuration options, then it walks you through configuring those options.
9. Learn about investigating data loss prevention alerts - This article introduces you to the lifecycle of alerts from creation, through final remediation and policy tuning. It also introduces you to the tools you use to investigate alerts.

Policy design overview

Designing a policy is mostly about clearly defining your business needs, documenting them in a policy intent statement, and then mapping those needs to policy configuration. You use the decisions you made in your planning phase to inform some of your policy design decisions.

Define intent for the policy

You should be able to summarize, in a single statement, the business intent for every policy you have. Developing this statement drives conversations in your organization and, when fully fleshed out, this statement directly links the policy to a business purpose and provides a roadmap for policy design. The steps in the Plan for data loss prevention (DLP) article helps you get started on your policy intent statement.

Remember that, as described in the DLP policy configuration overview, all DLP policies require you to:

- Choose what you want to monitor
- Choose the Policy Scoping.
- Choose where you want to monitor .
- Choose the conditions that must be matched for a policy to be applied to an item.
- Choose the action to take when the policy conditions are met.

For example, here's a fictitious first draft of an intent statement that provides answers to all five questions:

"We're a U.S. based organization, and we need to detect Office documents that contain sensitive health care information covered by HIPAA that are stored in OneDrive/SharePoint and to protect against that information being shared in Teams chat and channel messages and restrict everyone from sharing them with unauthorized third parties".

As you develop a policy design, you'll likely modify and extend the statement.

Map business needs to policy configuration

Let's break down the example draft statement and map it to DLP policy configuration points. This example assumes that you're using an unrestricted DLP admin account and that administrative units aren't configured.

> **Important**

Be sure you understand the difference between an *unrestricted administrator* and an *administrative unit restricted administrator* by reading Administrative units before you start.

Expand table

Statement	Configuration question answered and configuration mapping
"We're a U.S. based organization, and we need to detect Office documents that contain sensitive health care information covered by HIPAA...	- **What to monitor**: Office docs, use the U.S. Health Insurance Act (HIPAA) template - **Conditions for a match**: (preconfigured but editable) - item contains U.S. SSN and Drug Enforcement Agency (DEA) number, International Classification of Diseases (ICD-9-CM), International Classification of Diseases (ICD-10-CM), content is shared with people outside my organization - drives conversations to clarify the triggering threshold for detection like confidence levels, and instance count (called leakage tolerance).
...that are stored in OneDrive/SharePoint and protect against that information being shared in Teams chat and channel messages...	- **Where to monitor**: Location scoping by including or excluding OneDrive and SharePoint sites and Teams chat/channel accounts or distribution groups. **Policy scoping** (preview): Full directory
...and restrict everyone from sharing those items with unauthorized third parties."	- **Actions to take**: You add *Restrict access or encrypt the content in Microsoft 365 locations* - drives conversation on what actions to take when a policy is triggered including protective actions like sharing restrictions, awareness actions like notifications and alerts, and user empowerment actions like allow user overrides of a blocking action

This example doesn't cover all the configuration points of a DLP policy; it would need to be expanded. However, it should get you thinking in the right direction as you develop your own DLP policy intent statements.

Important

Keep in mind that the location(s) you pick impact whether you can use sensitive information types, sensitivity labels, and retention labels. The location(s) you pick also impact the actions that are available. See Data Loss Prevention policy reference for more information.

Complex rule design

The above HIPAA content in SharePoint and OneDrive is a simple example of a DLP policy. The DLP rule builder supports boolean logic (AND, OR, NOT) and nested groups.

> **Important**
>
> - All existing **Exceptions** are replaced with a NOT condition in a nested group inside of the **Conditions**.
> - You need to create groups in order to use multiple operators.

> **Important**
>
> When an action in Office desktop client apps, (Word, Outlook, Excel, and PowerPoint) matches a policy that uses complex conditions, the user will only see policy tips for rules that use the **Content contains sensitive information** condition.

- *Example 1 We need to block emails to all recipients that contain credit card numbers, OR that have the 'highly confidential' sensitivity label applied, but do NOT block the email if it is sent from someone on the finance team to adele.vance@contoso.com*
- *Example 2 Contoso needs to block all emails that contain a password protected file OR a zip document file extension ('zip' or '7z'), but do NOT block the email if the recipient is in the contoso.com domain OR the fabrikam.com domain, OR the sender is a member of the Contoso HR group.*

https://learn-video.azurefd.net/vod/player?id=0edf38a9-89fb-4bc5-95f4-48c70053e031&locale=en-us&embedUrl=%2Fpurview%2Fdlp-policy-design

> **Important**
>
> - Using the NOT condition in a nested group replaces the **Exceptions** functionality.
> - You need to create groups in order to use multiple operators.

> **Important**
>
> When an action in Office desktop client apps (Word, Outlook, Excel, and PowerPoint) matches a policy that uses complex conditions, the user will only see policy tips for rules that use the **Content contains sensitive information** condition.

Policy Design Process

1. Complete the steps in Plan for data loss prevention (DLP) - by working through this article you will:
   a. Identify your stakeholders
   b. Describe the categories of sensitive information to protect
   c. Set goals and strategy
   d. Define your policy deployment plan

2. Familiarize yourself with Data Loss Prevention policy reference so that you understand all the components of a DLP policy and how each one influences the behavior of a policy.

3. Familiarize yourself with what the DLP policy templates include.

4. Develop your policy intent statement with your key stakeholders. Refer to the example earlier in this article.

5. Determine how this policy fits into your overall DLP policy strategy.

6. Map the items in your policy intent statement to configuration options.

7. Decide which policy template you start from: predefined or custom.

8. Go through the template and assemble all information required before you create the policy. It's likely that you will find that there are some configuration points that aren't covered in your policy intent statement. That's ok. Go back to your stakeholders to iron out the requirements for any missing configuration points.

9. Document the configuration of all the policy settings and review them with your stakeholders. You can reuse your policy intent statement mapping to configuration points, which is now fully fleshed out.

10. Create a draft policy and refer back to your policy deployment plan.

See Also

- Learn about data loss prevention
- Plan for data loss prevention (DLP)
- Data Loss Prevention policy reference
- Data Loss Prevention policy tips reference
- Create and Deploy data loss prevention policies

Last updated on 04/08/2025

Learn about the Microsoft Purview Data Loss Prevention migration assistant for Symantec and Forcepoint

Article • 01/28/2025

This article helps you to learn about the Microsoft Purview Data Loss Prevention migration assistant for Symantec and Forcepoint.

The migration assistant is a Windows-based desktop application for migrating your Symantec and Forcepoint data loss prevention (DLP) policies to Microsoft Purview Data Loss Prevention. This article takes you through the five-step migration process. The migration assistant accepts Symantec DLP policy XML exports and Forcepoint DLP policy backup, performs mapping, and creates equivalent DLP policies through PowerShell scripts. You can use the migration assistant to create DLP policies in **Run the policy in simulation mode**. Policies in simulation mode won't affect your live data or impact your existing business processes.

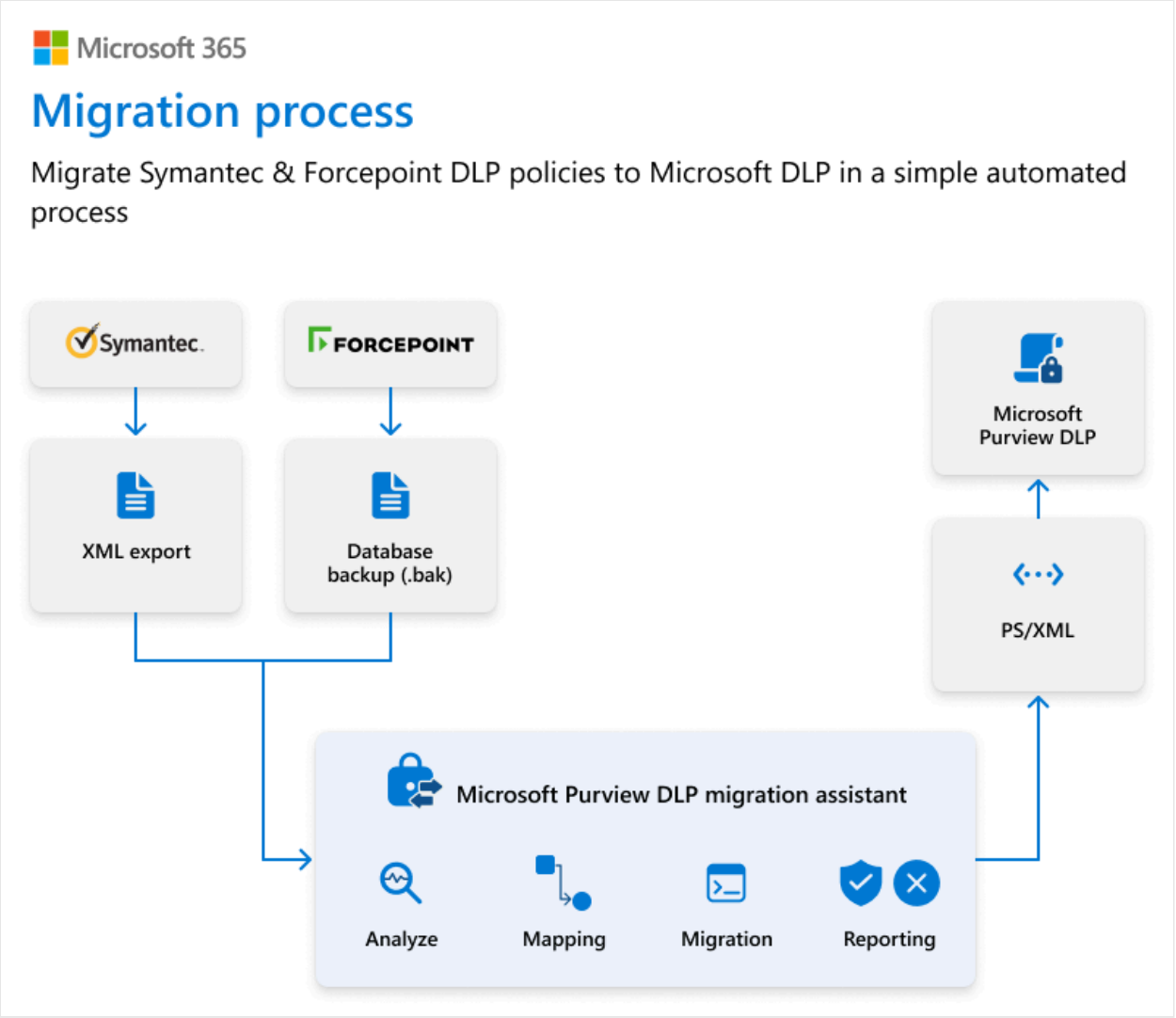
What can the migration assistant help with?

The migration assistant helps with some of the tasks involved in a DLP migration project:

- In a manual migration scenario, you need to perform a feasibility analysis between the source and target DLP platforms, map the features, migrate policies manually, and test and tweak DLP policies. With the migration assistant, your migrated DLP policies can be up and running within minutes of starting the migration assistant process.
- With migration assistant, you can quickly scale up your migration project. You can start by moving a single policy manually to multiple policies at the same time.
- The migration assistant automatically identifies sensitive information types (SITs) or Data Identifiers in source policies and creates custom SITs in your Microsoft tenant. It also moves all of your custom regular expressions and keywords in a few clicks.
- The migration assistant detects which conditions, exclusions, and actions are currently used in source policies and automatically creates new rules with the same conditions and actions.
- The migration assistant provides you with a detailed migration report that includes the migration status and recommendations at the policy level.
- The migration assistant ensures that your DLP policy migration project is private and takes place within the boundaries of your organization.

How does the migration assistant for Symantec and Forcepoint work?

Here's how the migration process works:



Each time the migration assistant runs, it performs the following steps:

- **Input:** The migration assistant ingests one or more Symantec DLP policy XML files or Forcepoint DLP policy backup (.bak) file.
- **Analyze:** The migration assistant interprets the files and identifies Symantec and Forcepoint DLP policy constructs.
- **Rationalize:** The migration assistant maps the identified Symantec and Forcepoint DLP policy constructs to Microsoft DLP capabilities. It performs validations for Microsoft DLP platform limitations.
- **Migrate:** The migration assistant runs PowerShell scripts for the DLP scenarios identified and supported by the Microsoft Purview DLP platform.
- **Report:** The migration assistant reports which policies were migrated successfully, which were partially migration, and which ones couldn't be migrated. It also provides recommendations to improve future migrations.

Understand mapping of Symantec and Forcepoint DLP elements to Microsoft Purview DLP elements

The migration assistant translates different policy elements from Symantec DLP to Microsoft Purview DLP.

Symantec DLP supported versions

The migration assistant supports migrating DLP policies from Symantec versions 15.0 through 16.0, maintenance packs included.

Forcepoint DLP supported versions

The migration assistant supports migrating DLP policies from Forcepoint versions 8.0 through 10.0.

Supported Workloads

The migration assistant migrates policies into these workloads:

Expand table

Workload	Migration assistant support
Exchange	Yes
SharePoint	Yes
OneDrive	Yes
Teams chat and channel messages	Yes
Endpoint devices	Yes

> **Tip**
>
> You can use the migration assistant to extend your policy to more workloads than the ones detected in the input Symantec or Forcepoint DLP policy.

Classification Elements

Here's how the migration assistant maps Symantec and Forcepoint elements to Purview DLP elements.

Expand table

Symantec Classification Element	Microsoft Purview DLP Classification Element
Regular Expression	Create a new custom sensitive information type (SIT) with the regular expression.
Keyword	Create a new custom SIT with a keyword list or keyword dictionary.
Keyword Pair	Create a new custom SIT with first keyword list as primary element & second keyword list as a supporting element with 300 char proximity.
Data Identifier	Map to a preconfigured SIT if an equivalent is available, else create a new custom SIT.

Here are the mapping details of optional validators for sensitive information types (also known as Data Identifiers in Symantec DLP) that the migration assistant uses while translating Symantec DLP policies:

Expand table

Symantec Optional Validators	Microsoft Purview DLP Optional Validators
Exclude exact match	Exclude specific matches
Exact Match Data Identifier Check	N/A
Exclude beginning characters	Starts or doesn't start with characters
Exclude ending characters	Ends or doesn't end with characters
Exclude prefix	Include or Exclude prefixes
Exclude suffix	Include or Exclude prefixes
Number Delimiter	N/A
Require beginning characters	Starts or doesn't start with characters
Exact Match	N/A
Duplicate digits	Exclude duplicate characters
Require ending characters	Ends or doesn't end with characters
Find keywords	Available as both primary & supporting elements

Regular Expressions – Potential validation issues to be aware of

When you upload your rule package XML file, the system validates the XML and checks for known bad patterns and obvious performance issues. Here are known issues that the validation process checks a regular expression for.

- Can't begin or end with alternator "|", which matches everything because it's considered an empty match.
  - For example, "|a" or "b|" won't pass validation.
- Can't begin or end with a ".{0,m}" pattern, which has no functional purpose and only impairs performance.
  - For example, ".{0,50}ASDF" or "ASDF.{0,50}" won't pass validation.
- Can't have ".{0,m}" or ".{1,m}" in groups, and can't have ".*" or ".+" in groups.
  - For example, "(.{0,50000})" won't pass validation.
- Can't have any character with "{0,m}" or "{1,m}" repeaters in groups.
  - For example, "(a*)" won't pass validation.
- Can't begin or end with ".{1,m}"; instead, use just "."
  - For example, ".{1,m}asdf" won't pass validation; instead, use just ".asdf".
- Can't have an unbounded repeater (such as "*" or "+") on a group.
  - For example, "(xx)*" and "(xx)+" won't pass validation.

Condition and Exception Mapping

Here's how the migration assistant maps Symantec and Forcepoint, condition and exception elements for various workloads to Microsoft Purview DLP conditions.

Exchange Workload

Expand table

Condition/Exception in Symantec	Condition in Microsoft Purview DLP
Content Matches Regular Expression	Content contains SIT
Content Matches Keyword	Content contains SIT
Content Matches Data Identifier	Content contains SIT
Content Matches Classification	Not supported
File Properties • File name	One or more of the following: • Document name is

Condition/Exception in Symantec	Condition in Microsoft Purview DLP
• File type	• File extension is
Message Attachment or File Type Match	One or more of the following: • Attachment is password protected • Attachment's file extension is
Message Attachment or File Size Match	Document size equals or is greater than
Message Attachment or File Name Match	One or more of the following: • Document name contains words or phrases • Document name matches patterns
Message/Email Properties and Attributes	One or more of the following: • Email subject contains
Sender/User Matches Pattern	One or more of the following: • Sender is • Sender is a member of • Sender domain is • Sender address contains words • Sender address matches patterns • Sender IP address is
Recipient Matches Pattern	One or more of the following: • Recipient is a member of • Recipient domain is • Recipient is • Recipient address contains words • Recipient address matches patterns
Sender/User based on a Directory Server Group	Not supported
Recipient based on a Directory Server Group	Not supported
Content Matches Exact Data from an Exact Data Profile (EDM)	Not supported
Content Matches Document Signature from an Indexed Document Profile (IDM)	Not supported
Detect using Vector Machine Learning profile (VML)	Not supported

Condition/Exception in Symantec	Condition in Microsoft Purview DLP
Protocol Monitoring • SMTP protocol	Exchange (EXO) DLP policy

Endpoint Devices, SharePoint Online, OneDrive, and other workloads

Expand table

Condition/Exception in Symantec	Condition in Microsoft Purview DLP
Content Matches Regular Expression	Content contains SIT
Content Matches Keyword	Content contains SIT
Content Matches Data Identifier	Content contains SIT
Message Attachment or File Type Match	Document's file extension is
Protocol Monitoring • HTTP • HTTPS • FTP	Cross-workload DLP policy(s)
Protocol Monitoring: Endpoint Device Type • CD/DVD • Removable storage • Copy to network share • Printer/Fax • Clipboard • Cloud storage • Application File Access • SEP Intensive Protection	One or more of the following (Devices): • Copy to USB removable media • Copy to network share • Copy to clipboard • Print • Upload to cloud service domains or access by browsers that aren't allowed

Response Rules

Here's how the migration assistant maps Symantec and Forcepoint response rules to Microsoft Purview DLP actions.

Expand table

Symantec Response Rule	Microsoft Purview DLP Action
Generate DLP Incident	Generate Alert
Logging (Syslog)	Audit logs
Network Prevent: Modify SMTP Message • Modify email subject • Modify header	One or more of the following: • Prepend subject • Set headers
Network Prevent: Block SMTP Message • Bounce message to sender • Redirect message to this address	One or more of the following: • Block / Restrict access • Send user notification • Redirect message to
Send Email Notification	Send User Notification
Endpoint Prevent • Notify • Notify with Cancel • Block	One or more of the following (Endpoint Devices) • Notify • Block • Audit
User Cancel	One or more of the following: • Block / Restrict access • User Overrides

Next steps

Now that you've learned about the Microsoft Purview Data Loss Prevention migration assistant for Symantec, your next steps are to:

1. Get started with the Microsoft Purview Data Loss Prevention migration assistant for Symantec and Forcepoint
2. Use the Microsoft Purview Data Loss Prevention migration assistant for Symantec and Forcepoint

---

Feedback

Was this page helpful? Yes No

Get started with the Microsoft Purview Data Loss Prevention migration assistant for Symantec and Forcepoint

This article walks you through the prerequisites and installation of the Microsoft Purview Data Loss Prevention migration assistant.

Before you begin

If you're using the Microsoft Purview Data Loss Prevention migration assistant for the first time, ensure the following prerequisites are met:

> **Tip**
>
> If the application won't launch after completing all the steps in this article, refer to Troubleshoot the DLP migration assistant.

1. Have an appropriate Microsoft 365 subscription

You'll need the appropriate DLP licensing for the locations that the migrated policies are scoped to. Check the DLP licensing requirements.

2. Have appropriate user role and privileges

You need to have either the *Global Administrator or Compliance Administrator* role to be able to use the migration assistant.

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. This helps improve security for your organization. Global Administrator is a highly privileged role that should only be used in scenarios where a lesser privileged role can't be used.

3. Check your Operating System

You can only install and run the migration assistant on these operating systems:

- Windows 11,10,8.1,7
- Windows Vista
- Windows Server 2008+
- Windows Server 2003 (excluding IA-64)

4. Install PowerShell

You must have PowerShell v7.2.6 or higher installed on the host machine. Download Powershell-7.2.6-win-x64.msi core on your machine.

5. Install .NET

You must have .NET desktop runtime v6.0.13 or higher installed on the host machine. Download .NET 6.0 Desktop Runtime (v6.0.13) - Windows x64 Installer (microsoft.com).

Export DLP policies

Before you begin the migration process, you need to export your DLP policies from your current DLP solution.

Symantec DLP

You need to export your Symantec DLP policies as XML files from Symantec DLP. Export the policies using the Symantec DLP policy export procedures.

Forcepoint DLP

You need to export your Forcepoint DLP policies using the backup capability from Forcepoint DLP. Export the policy backup using the Forcepoint DLP policy backup procedures.

Installation steps

Follow these steps to install the Microsoft Purview Data Loss Prevention migration assistant for Symantec:

1. Download and launch the **Microsoft Purview Data Loss Prevention migration assistant** file.
2. The following dialog box will open. Select **Next**.
3. Review the End-User License Agreement.

4. After reviewing, accept the License Agreement and select **Next**.
5. Select the location where you want to install the migration assistant and select **Next**.
6. Once installer is ready, select **Install** to begin installation.
7. Wait while the migration assistant is being installed.
8. Once migration assistant is installed, select **Finish** and close the dialog box.

Next steps

Now that you have installed Microsoft Purview Data Loss Prevention migration assistant for Symantec and Forcepoint, you're ready to move on to your next step where you use the migration assistant.

- Using the Microsoft Purview Data Loss Prevention migration assistant for Symantec and Forcepoint

See also

- Learn about the Microsoft Purview Data Loss Prevention migration assistant for Symantec and Forcepoint
- Learn about data loss prevention
- Create and Deploy data loss prevention policies
- Microsoft 365 subscription

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Use the Microsoft Purview Data Loss Prevention migration assistant for Symantec and Forcepoint

The Microsoft Purview Data Loss Prevention migration assistant for Symantec and Forcepoint helps you migrate your existing Symantec or Forcepoint DLP policies to Microsoft Purview DLP. The assistant maps data identifiers, keywords, and regular expressions to Microsoft sensitive information types (SITs), and migrates rules, conditions, and actions. By the end of this procedure, your policies are active in Microsoft Purview DLP and ready for testing.

Before you start, make sure you've met these prerequisites:

- Complete the steps in the Before you begin section.
- Export the required DLP policy files from your current DLP solution.

After a policy is migrated, you can test and fine-tune it in Microsoft Purview DLP.

Steps for migration

Use these steps to perform a DLP policy migration:

- Step 1: Sign in to your account
- Step 2: Upload the policies that are being migrated
- Step 3: Edit policy settings
- Step 4: Review premigration feasibility report
- Step 5: Merge policies and select policy mode
- Step 6: Migration in progress
- Step 7: View the migration report
- Next steps: After policy import

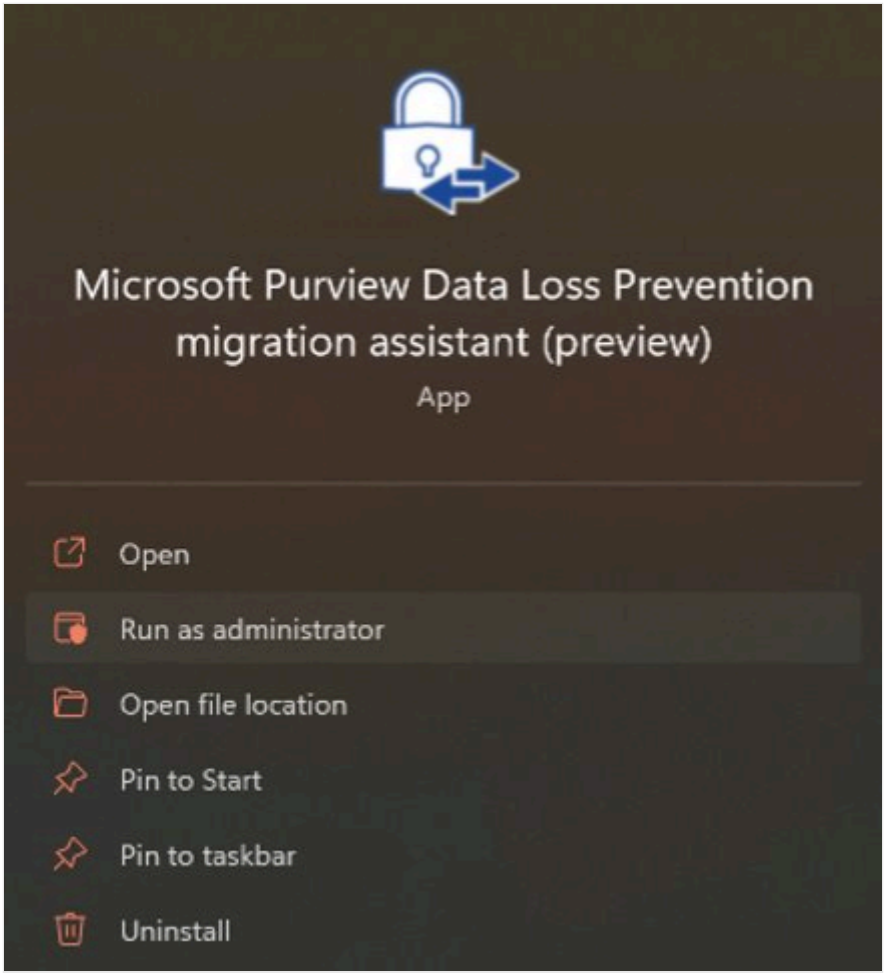
Interactive guide

Check out the Use the Microsoft Purview DLP migration assistant interactive guide for a visual walkthrough of the migration process.

Step 1: Sign in to your account

After you install and launch the migration assistant, sign in to your account.

> **Important**
>
> The first time you launch the migration assistant, choose the **Run as administrator** option. Running as administrator is required because the migration assistant might need to install additional components if they aren't already available on your machine.
>
> For all subsequent launches, you can launch the migration assistant normally and don't need to run as administrator.



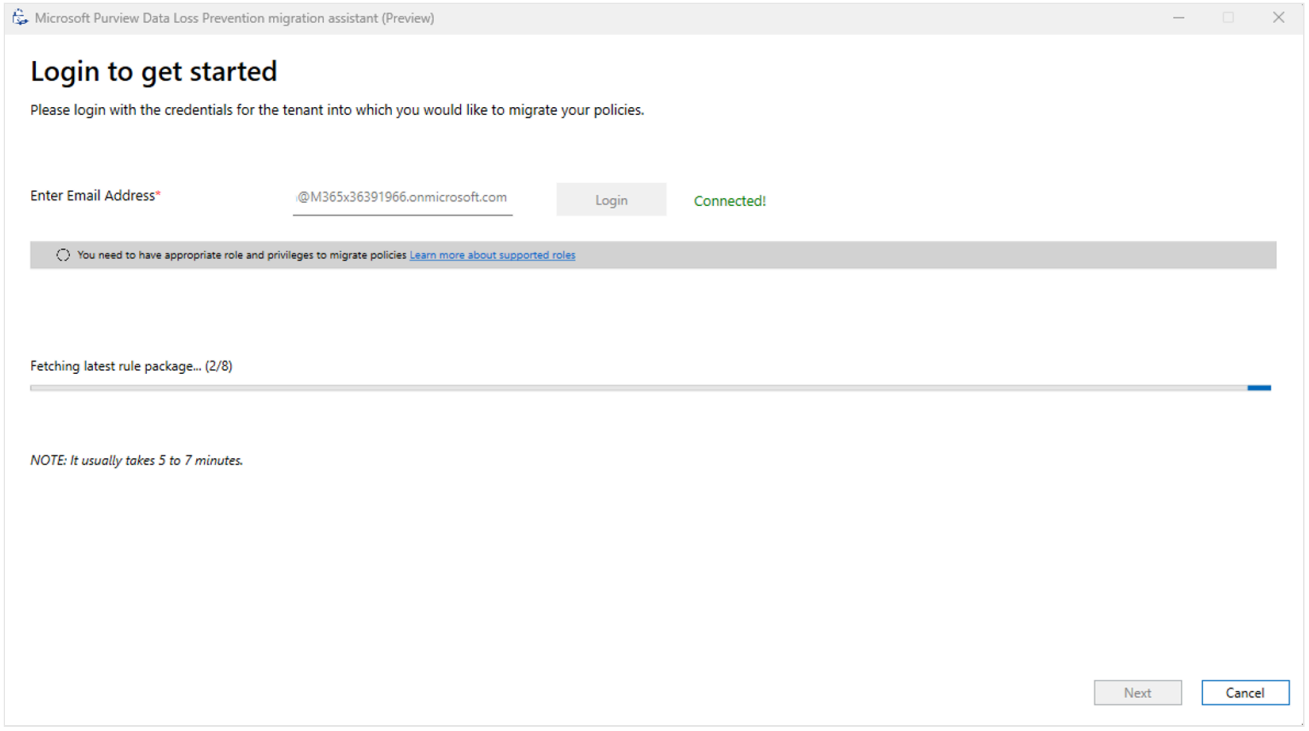
To sign in to the migration assistant:

1. You're greeted with a welcome screen.
   a. Select **Get Started** and the migration assistant checks if your environment is set up correctly.
   b. Select **Next**.
2. Enter your username and select **Login**.

   - Enter your password in the browser window that opens and select **Sign in**.

> **Note**
>
> This application uses the Exchange Online PowerShell module. Basic authentication must be enabled in WinRM on the local computer. For more information, see Prerequisites for the Exchange Online PowerShell module.

3. Wait for sign-in validation. The migration assistant fetches the information required for later stages of the migration process.



4. After you're signed in, choose **Next**.

Step 2: Upload the policies that are being migrated

You need to upload your current DLP policy exports, which act as input for the migration assistant. The policies you upload will be migrated to the Microsoft Purview DLP platform.

1. Select which DLP solution you want to migrate policies from.

2. To upload the files, select **Browse**.

3. Select the required policy files in the File Explorer pop-up window and choose **Open**.

    a. Symantec DLP

        i. For migrating policies from Symantec DLP, you need to upload the XML files that you have exported from Symantec DLP.

        ii. You can select more than one XML file to migrate multiple policies at a time. Migrate one to three policies at a time to avoid confusion during later stages of the migration

process.

> **Note**
>
> Ensure that the XML files you upload are Symantec DLP policy exports only and no other kind of XML.

b. Forcepoint DLP

i. For migrating policies from Forcepoint DLP, you need to upload the policy backup (.bak) file from Forcepoint DLP.

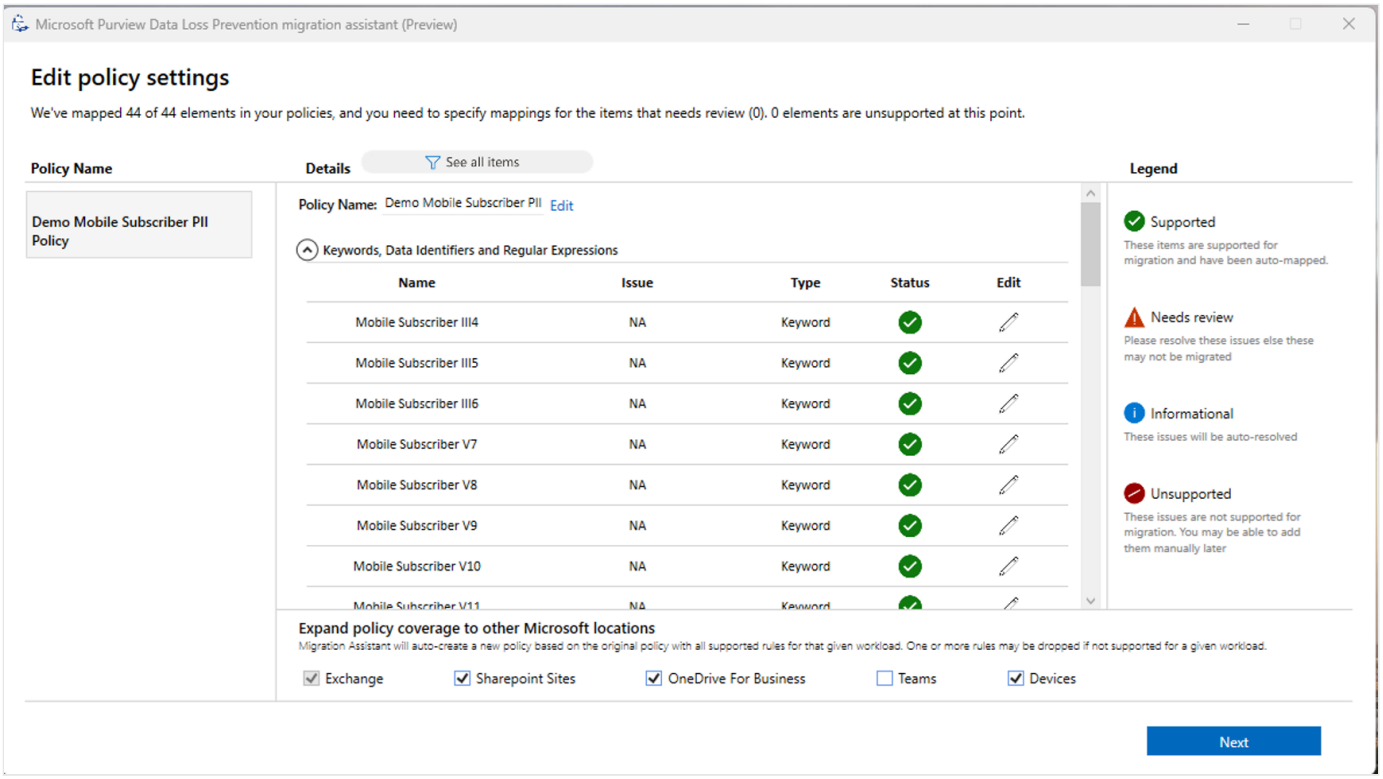
4. The tool shows you a list of your selected input policy files. You can modify your selection from the policy list.

5. After you complete the selection of policy files you want to migrate, select **Next** to move to the next step.

Step 3: Edit policy settings

After you input the policies you want to migrate, the migration assistant processes those files and maps Symantec and Forcepoint DLP policy elements to Microsoft DLP elements.

> **Important**
>
> There may be some items that may need your review and will be marked with 'Needs review' warning symbol.



Keywords, data identifiers, and regular expressions

Symantec DLP and Forcepoint DLP differ from Microsoft Purview Information Protection in how they allow users to define sensitive information that needs to be protected.

Microsoft Purview Information Protection allows users to define sensitive items that need protection as sensitive information types (SITs), or through trainable classifiers. Microsoft provides many commonly used SITs like Credit Card Number that are preconfigured. If the preconfigured SITs don't meet your needs, you can create your own custom SITs.

The most common ways Symantec and Forcepoint users specify the types of sensitive information that need protection are:

- Using out-of-box (OOB) data identifiers
- Customizing OOB data identifiers
- Defining regular expressions or keywords in DLP rules

The migration assistant handles OOB data identifiers, customized OOB data identifiers, and regular expressions or keywords in one of two ways:

- **Map to an existing OOB SIT:** For each sensitive data type that has an equivalent SIT in Microsoft DLP, the migration assistant attempts to create a 1:1 mapping. It automatically maps OOB Symantec or Forcepoint data identifiers to preconfigured Microsoft SITs, if an equivalent exists. If you want to bring the Symantec or Forcepoint data identifier over as-is, you can create a new SIT as described in the next step.

- **Migrate as a new custom SIT:** For each sensitive data type that doesn't have an equivalent SIT in Microsoft DLP, the migration assistant automatically creates a new SIT. Similarly, any regular expressions or keywords defined directly in rules are brought over as a new custom SIT.

> **Note**
>
> Regular expressions or keywords defined directly at the rule-level of Symantec or Forcepoint policies take on the name of the rule itself and appear in the **Source** column. If there are multiple regular expressions or keywords, they take the rule name followed by roman numerals.
>
> Each of these items is migrated separately as a custom SIT. Migrating each regular expression or keyword as a separate custom SIT can lead to confusion. Review and rename these SITs as soon as possible.
>
> You can't edit the name of these SITs within the migration assistant. You can edit the names of these custom SITs from the Microsoft Purview portal or via PowerShell after the policy migration is complete.

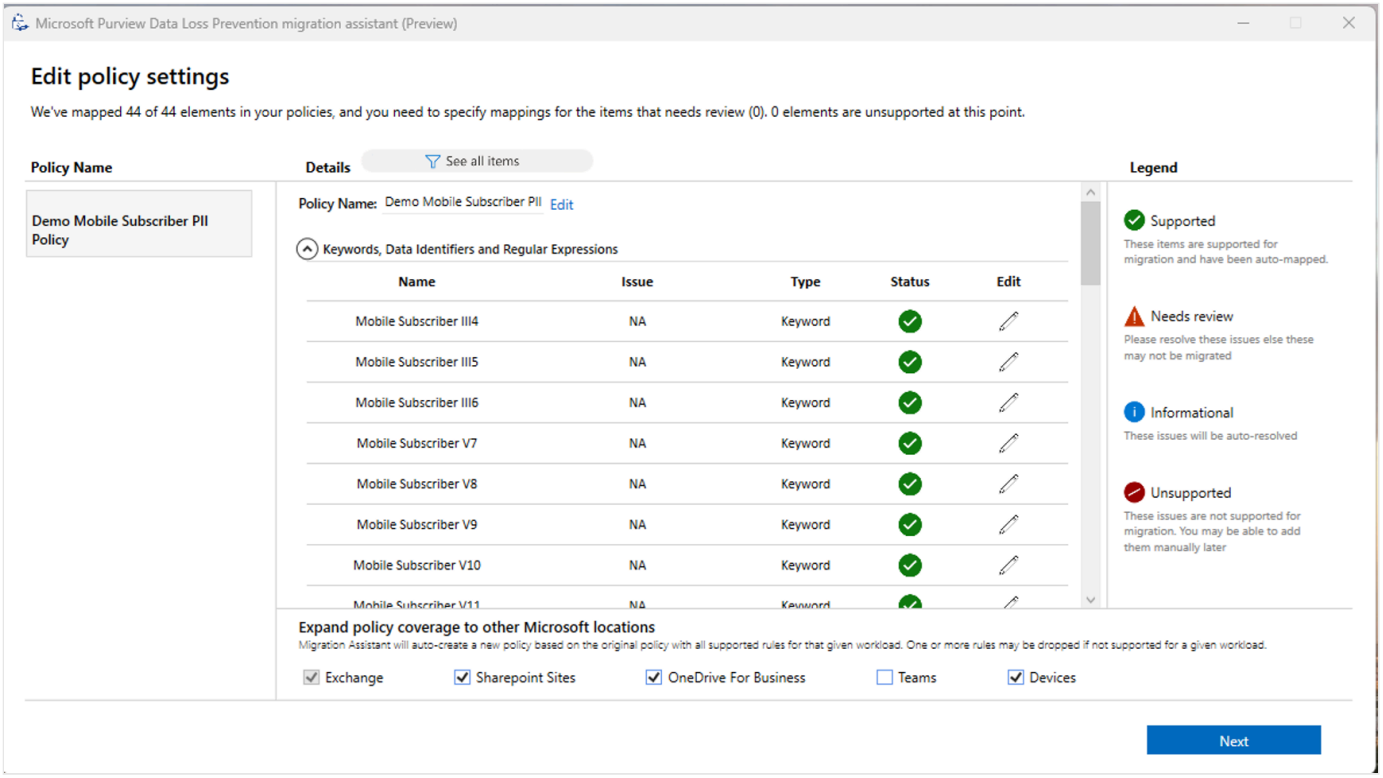
Inclusions, exclusions, and response rules

The current version of the migration assistant brings over policies with '*Generate Incident Report*' as a default action. Also, DLP policies in Microsoft DLP automatically log events in Unified Audit Log and won't need a separate action equivalent to 'Syslog' in Symantec DLP.

In addition to the automatically migrated response rules (or DLP actions), you can manually add (or remove) actions to the policies after the migration assistant has successfully migrated the policies.

You can see a list of all your policies and rules along with their status. Select different policies from the left column to see details for each policy. By default, the tool displays all items that need review. You can toggle to see all items in a given policy by choosing the **Show All Items** button at

the top of the Details section.



Policy details

Policy Name - You can edit the name of the policy before migration.

Each policy is divided into two sections:

- Keywords, data identifiers, and regular expressions — These are migrated as sensitive information types (SITs) in Microsoft DLP.

- Rules — These are automatically mapped to different Microsoft DLP conditions. Each row in the table shows:
  - The name that the SIT will have when migrated
  - Potential issues that may affect migration
  - Type of rule as detected from the input policy
  - Status:
    - Blank/empty status — This row element is migrated without issues.
    - Needs Review — This row element might have one or more issues and might require input from you.
    - Informational — This row element might have one or more changes needed for migration but will be autoresolved.
    - Unsupported — This row element isn't supported for migration by the tool and might need to be migrated manually after the tool exits.
  - Edit button

Extend policy coverage to other Microsoft locations: - You can extend your current Symantec policies to other Microsoft locations in addition to the original scope of the policy.

For example, a Purview DLP policy that's scoped to email can be extended to SharePoint, OneDrive, Teams, and Endpoint Devices.

The migration assistant auto-creates a new policy based on the original policy with all supported rules for that given workload. One or more rules may be dropped if not supported for a given workload.

For example, the **Email subject is** condition may be dropped while extending an email (Exchange) DLP policy to OneDrive.

Some checkboxes may be disabled by default if there are no supported conditions available in extended locations.

Editing a row element — When you select a row element to edit, you're taken to an **Edit** screen with more details. If there are issues, they're reported through a yellow banner at the top of the screen. Make changes to the content in the editable sections. Those changes are incorporated at the time of migration. After you resolve the issues, the yellow banner disappears.

Use existing SITs from the tenant to automatically replace the current SIT. You can replace the current SIT (which is being edited) with another SIT from your tenant.

You can manually change any of the mappings by selecting the corresponding row in the 'Target' column. This opens a drop-down list with all the OOB SITs and any custom SITs you previously created. Select the SIT you want to map to the source item. Alternatively, you can choose **New SIT** from the drop-down. The migration assistant then brings over the source SIT as a new custom SIT.

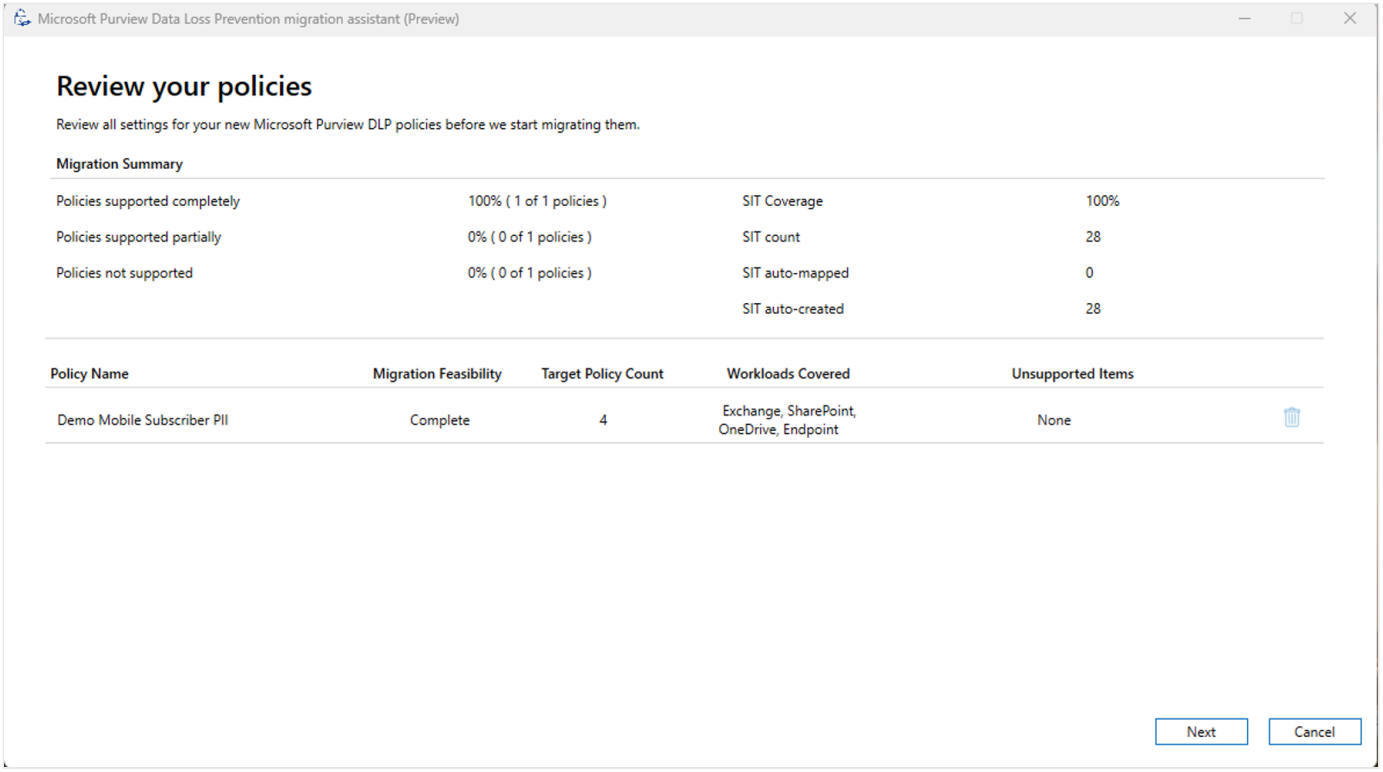
We strongly recommend using existing SITs to replace current SITs wherever possible. Reusing existing SITs helps reduce creation of duplicate SITs and reduces the effort of optimizing multiple custom SITs of the same kind. Learn more about sensitive information type entity definitions.

> **⚠ Warning**
>
> Microsoft DLP platform has a threshold for up to 10 rule packages per tenant. This limit is enough for most customers, but the creation of many duplicate custom SITs may quickly lead you to hitting this threshold without the ability to create any new custom SITs.

After you've reviewed all the policies and the rules within them, select **Next**. If one or more of the policies contain at least one element with *Needs Review* status, then you'll see a **Continue with errors** button instead of **Next**.

Step 4: Review premigration feasibility report

The premigration feasibility report shows you how policies will be migrated. Review this report and make any necessary adjustments before starting a migration run.



Review the details and choose **Next**.

Step 5: Merge policies and select policy mode

Each DLP policy is split into multiple rules depending on the workloads selected and support for underlying conditions and exceptions for those workloads.

The tool allows you to merge each of these rules with compatible target policies for the same workload. These target policies can be existing Microsoft Purview DLP policies you created earlier, or new policies that you're importing along with the policy you want to merge.

Merging policies together creates a single Microsoft Purview DLP policy with multiple rules. Consolidating multiple rules into a single policy allows for better management of policies.

For each DLP policy that is created, you can set it to one of three states:

- On (**Yes, turn it on right away**)
- Simulate (**Run the policy in simulation mode** and **Show policy tips**)
- Off (**No, keep it off. I'll turn it on later**)

Set the state in the migration assistant before migrating a policy:

1. Choose whether to turn on or off from the following three options:

   - Turn on policy immediately.
   - Turn on policy in simulation mode first. Remove from simulation mode later manually.
   - Keep it off. Turn it on later manually.

   We recommend bringing over the policies first in simulation mode. You can monitor the alerts that the policy generates and fine-tune the policy as required by your organization. After your policy is fine-tuned, you can turn it on.

2. Select **Start Migration** to import your policy. A new PowerShell window opens asking you to sign in again.

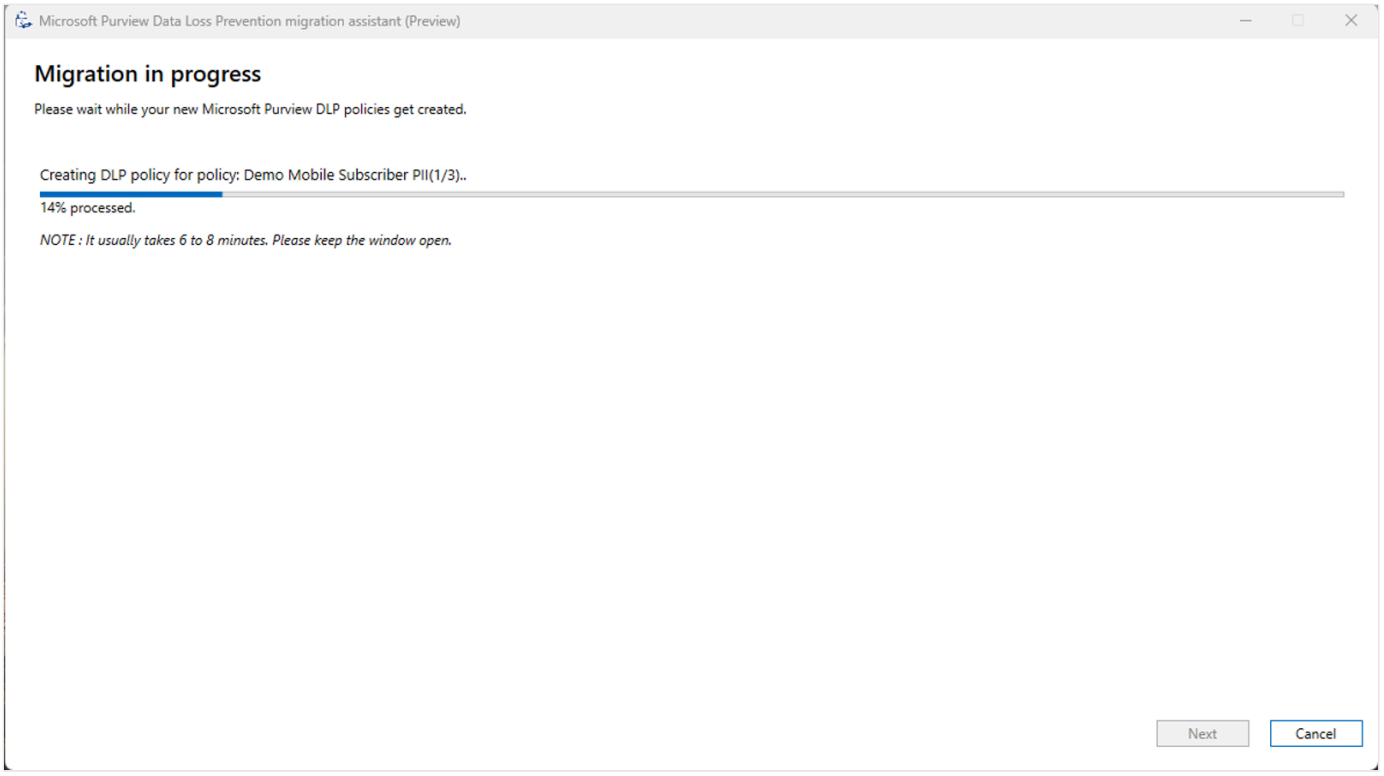
   After you sign in, PowerShell scripts run to create new policies in Microsoft DLP with all the data from the input policy files, and any settings you configured during the previous migration assistant steps.

   Wait until the script finishes with a Success/Failure message. Then, new SITs and policies start appearing in the Purview portal.

Step 6: Migration in progress

In this step, the migration tool creates DLP policies in the Microsoft Purview portal.

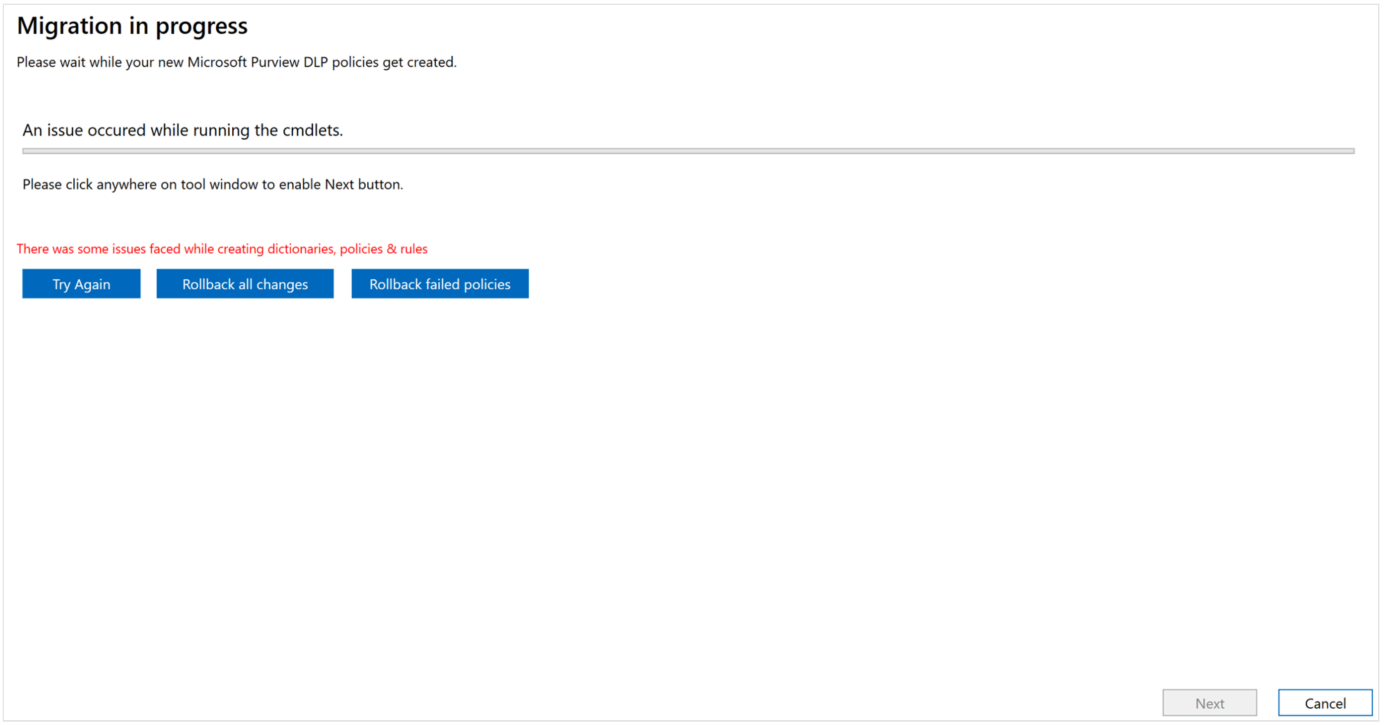
> ⓘ **Note**
>
> Don't close the tool window while policies are being created. Closing the tool might cause incomplete SITs and policies, which could require manual cleanup.



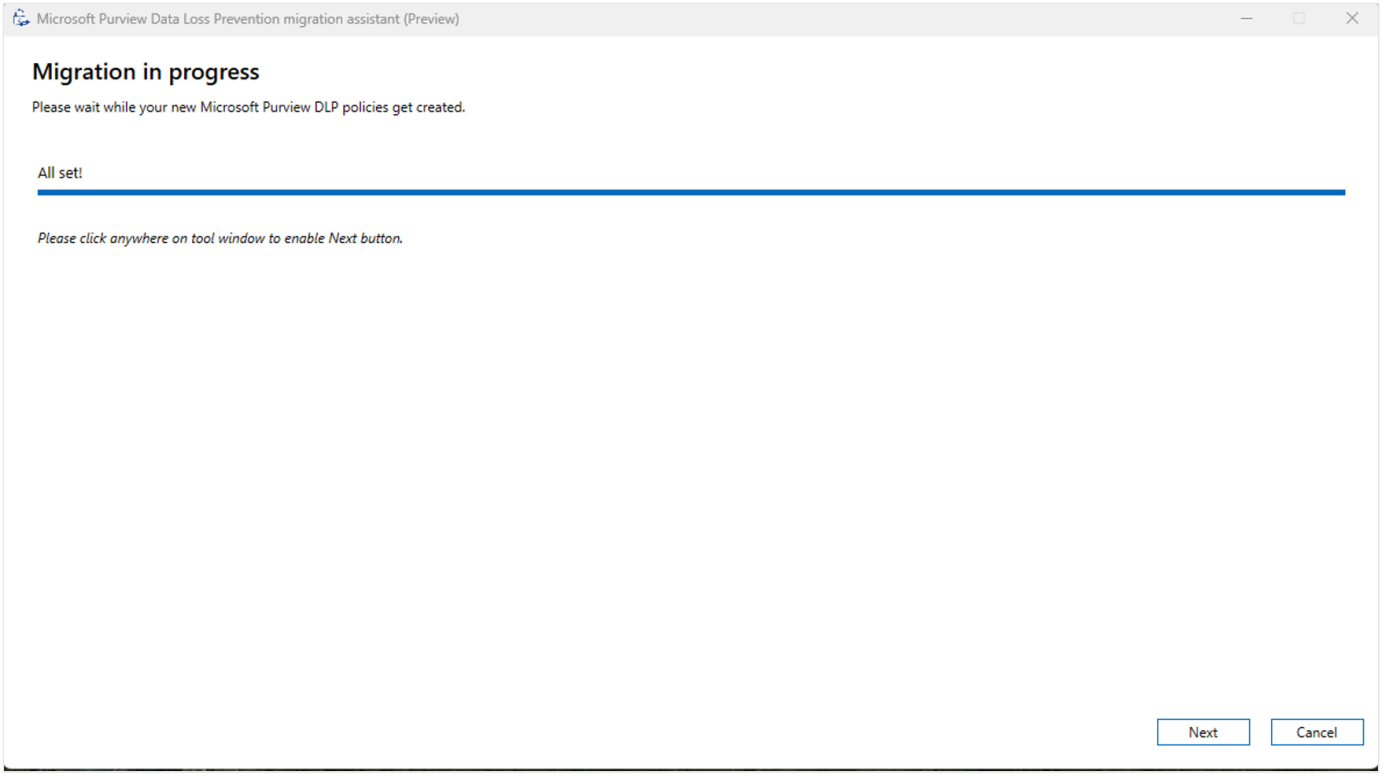
If an error occurs during migration, you can choose one of these actions:

- **Try Again** — Policy creation is attempted again.
- **Rollback all changes** — All SITs and policies for that session are deleted.
- **Rollback failed policies** — Only the failed SITs and policies for that session are deleted.

> **Important**
>
> When you choose to rollback any changes, the rollback can take two to four hours to complete. Keep the migration assistant tool window open for the entire duration for a successful rollback.

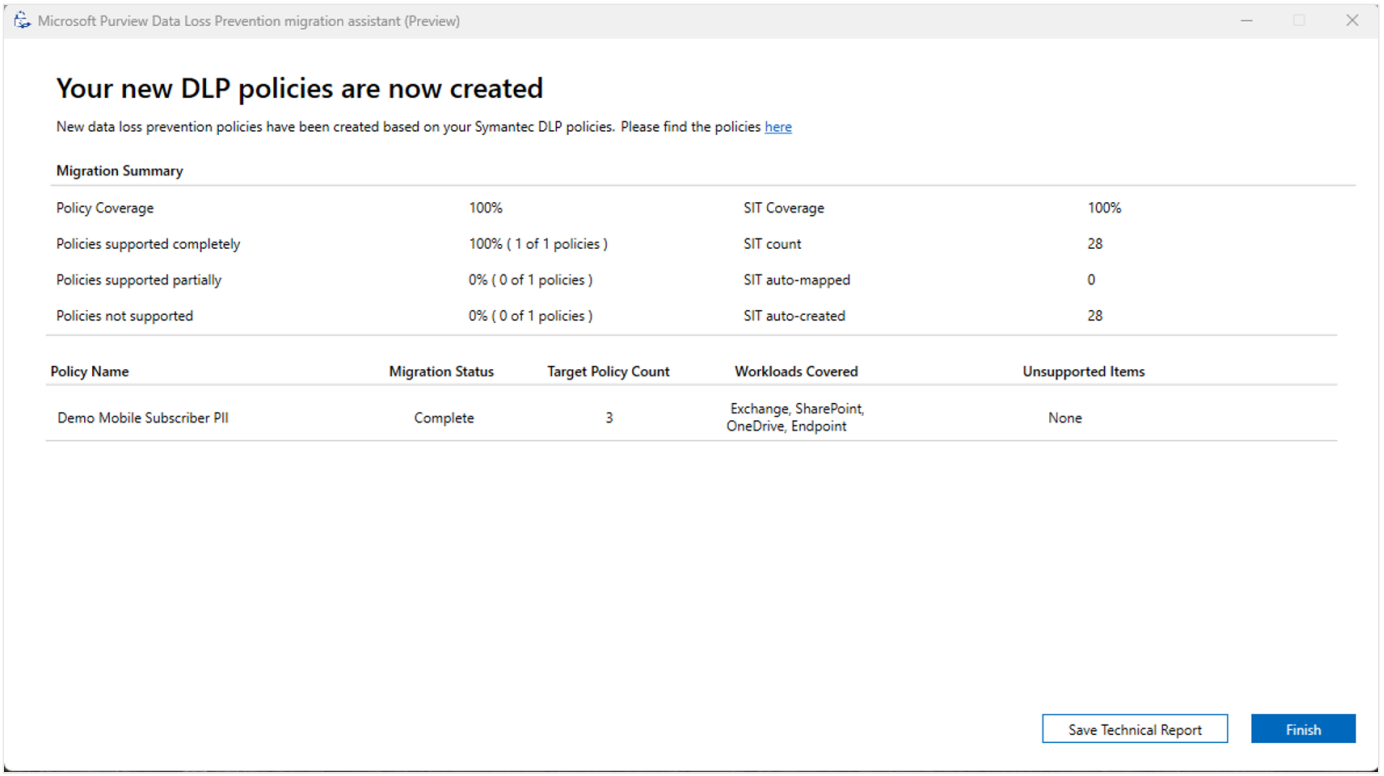


After the policies are migrated, select **Next** to view the migration report.



Step 7: View the migration report

You can view the migration report after your policies are imported and the migration process is complete.



Each session gets its own report. A session begins at the time you launch the app and ends when you exit the app or when migration process is completed.

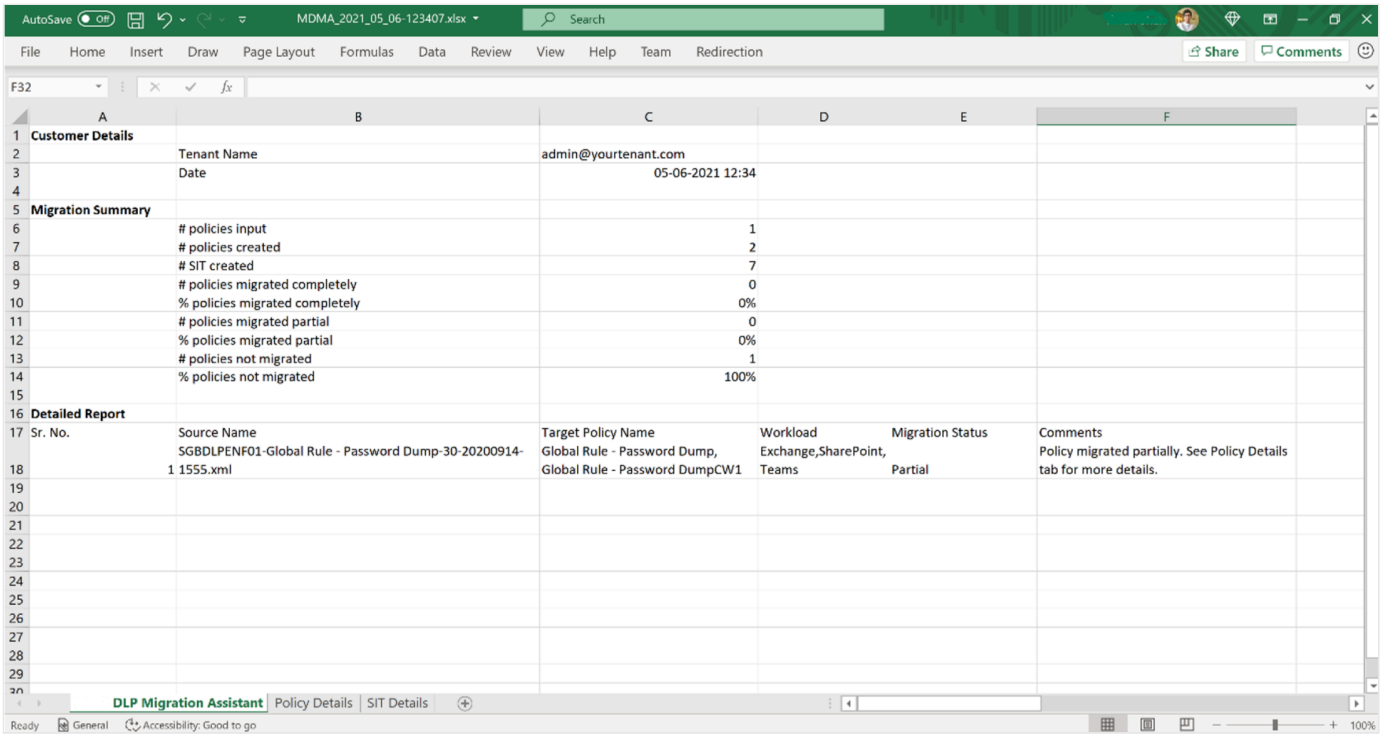
Technical report

You can select the **Save Technical Report** button to save a more detailed Excel-based report that is divided into three sheets:

- Overview
- Policy Details
- SIT Details

Overview Sheet — This sheet provides an overview of the migration session with the following details:

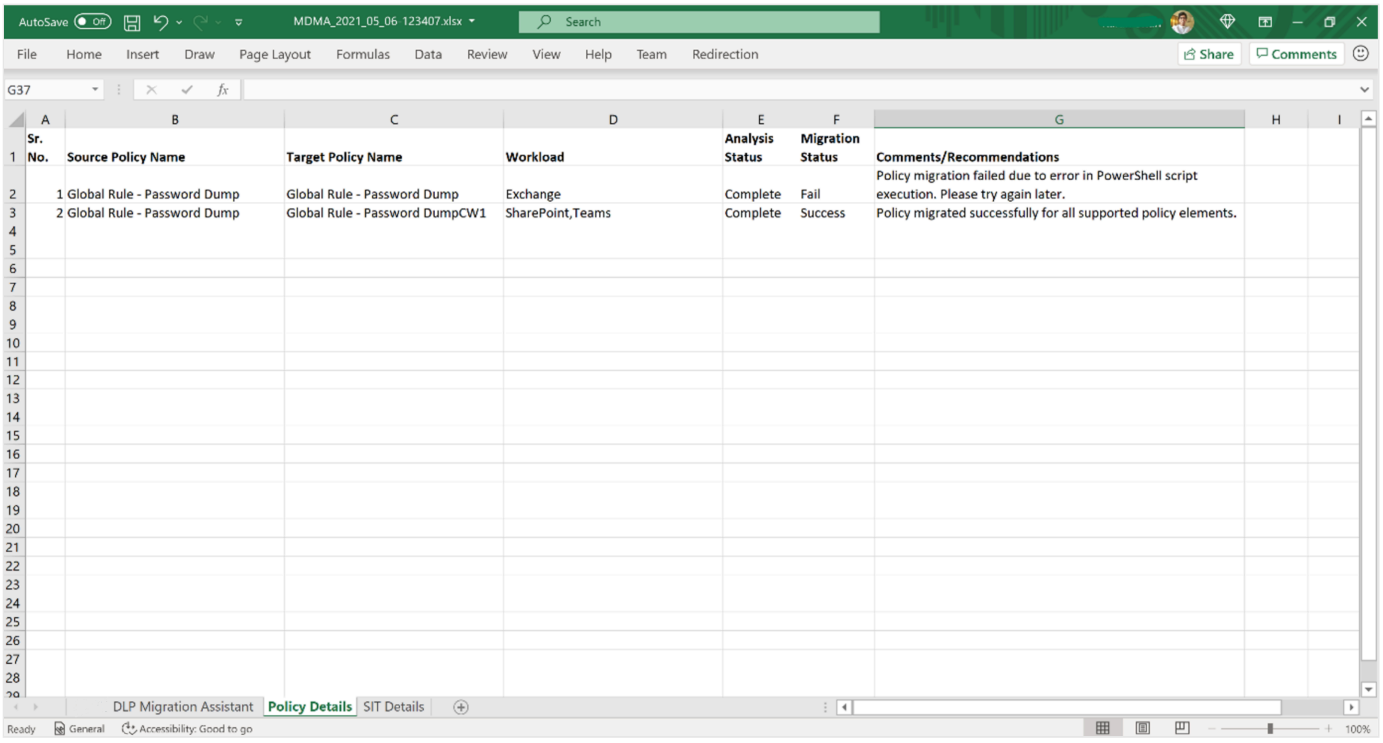
- Tenant's name.
- Timestamp of session.
- Overall summary stats for that session.
- Input policy level details, migration status and comments/recommendations.



	A	B	C	D	E	F
1	**Customer Details**					
2		Tenant Name	admin@yourtenant.com			
3		Date	05-06-2021 12:34			
4						
5	**Migration Summary**					
6		# policies input	1			
7		# policies created	2			
8		# SIT created	7			
9		# policies migrated completely	0			
10		% policies migrated completely	0%			
11		# policies migrated partial	0			
12		% policies migrated partial	0%			
13		# policies not migrated	1			
14		% policies not migrated	100%			
15						
16	**Detailed Report**					
17	Sr. No.	Source Name	Target Policy Name	Workload	Migration Status	Comments
18	1	SGBDLPENF01-Global Rule - Password Dump-30-20200914-1555.xml	Global Rule - Password Dump, Global Rule - Password DumpCW1	Exchange,SharePoint, Teams	Partial	Policy migrated partially. See Policy Details tab for more details.

Policy Details — This sheet provides a detailed view of each migrated policy created or not created with the following information:

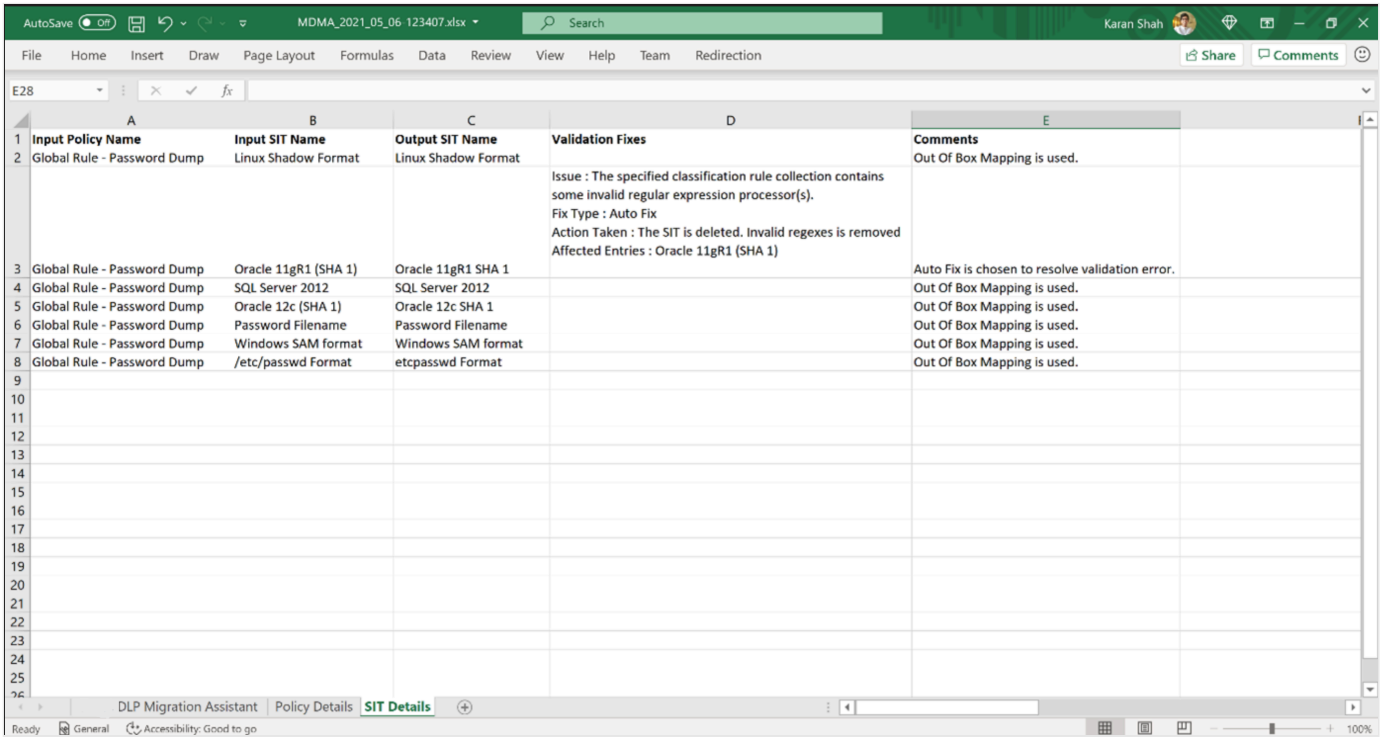
- Mapping of source policy and target policy(s) created.
- List of workloads each policy is applied to.
- Analysis status stating if the policy is migrated completely, partially or can't migrate.
    - For workloads other than Exchange, this would typically show as *Complete* since we create a policy with the *Content contains* condition, which is supported across all workloads.
- Migration status describing if the policy migration was a success or failure.
- Comments/recommendations with more details of that policy.



Sr. No.	Source Policy Name	Target Policy Name	Workload	Analysis Status	Migration Status	Comments/Recommendations
1	Global Rule - Password Dump	Global Rule - Password Dump	Exchange	Complete	Fail	Policy migration failed due to error in PowerShell script execution. Please try again later.
2	Global Rule - Password Dump	Global Rule - Password DumpCW1	SharePoint,Teams	Complete	Success	Policy migrated successfully for all supported policy elements.

SIT Details — This sheet provides information about all the sensitive information types (SITs) that were migrated with the following information:

- Policy-wise mapping of Input and Output SITs created.
- Validation fixes with information about validation errors that occurred during the migration process.
- Comments about SIT automapping, and remediation steps.



Input Policy Name	Input SIT Name	Output SIT Name	Validation Fixes	Comments
Global Rule - Password Dump	Linux Shadow Format	Linux Shadow Format		Out Of Box Mapping is used.
Global Rule - Password Dump	Oracle 11gR1 (SHA 1)	Oracle 11gR1 SHA 1	Issue : The specified classification rule collection contains some invalid regular expression processor(s). Fix Type : Auto Fix Action Taken : The SIT is deleted. Invalid regexes is removed Affected Entries : Oracle 11gR1 (SHA 1)	Auto Fix is chosen to resolve validation error.
Global Rule - Password Dump	SQL Server 2012	SQL Server 2012		Out Of Box Mapping is used.
Global Rule - Password Dump	Oracle 12c (SHA 1)	Oracle 12c SHA 1		Out Of Box Mapping is used.
Global Rule - Password Dump	Password Filename	Password Filename		Out Of Box Mapping is used.
Global Rule - Password Dump	Windows SAM format	Windows SAM format		Out Of Box Mapping is used.
Global Rule - Password Dump	/etc/passwd Format	etcpasswd Format		Out Of Box Mapping is used.

Next steps: After policy import

Visit the Purview portal and validate the policies you migrated.

Check sensitive information types

Use the following steps to verify and tune the migrated SITs:

1. To validate that the SITs were created, navigate to **Data Classification** > **Classifiers** > **Sensitive Info Types** and look for the SITs. You can also sort the list on **Publisher** and check for SITs with a publisher name of “DLP Migration Tool”.

2. Rename SITs as needed. For many SITs, you may notice there are similar names, often followed by roman numerals. To avoid confusion and duplication post-migration, you should rename these SITs. This is true for cases where your regular expressions and keywords are defined directly in rules within your input Symantec or Forcepoint DLP policies.

3. Test and fine-tune SITs as needed. You should test and fine-tune the migrated SITs. The migration assistant creates new SITs with a few standard settings, which might not be right for your tenant so look out for the following:
   a. Regular expressions: Unsupported or deleted regexes (during migration)
   b. Keywords
      i. Case sensitive versus insensitive keywords
      ii. String versus word match
      iii. Proximity
      iv. Optional validators

Check DLP policies

Use the following checks to validate and finalize the migrated DLP policies:

1. **Validate the DLP policies created**

   Select **Data Loss Prevention** from the left panel and check if new policies are created.

2. **Add any missing policy elements**

   While most Symantec and Forcepoint DLP policy elements (like conditions, exclusions, or actions) are migrated, some elements from your input policy might be dropped during the migration process. Dropping some policy elements during migration is a known limitation of the migration assistant. In this scenario, manually add these elements to the policy, as they're supported by the Microsoft Purview DLP platform.

3. **Test and fine-tune the policy**

   You can test the policy and fine-tune it for the needs of your organization.

4. **Turn the policy on**

   After fine-tuning, you can turn on the policy or put it in production mode.

5. **Bring over any remaining policies**

   You can go back to bring over the next policy or next batch of policies using the migration assistant.

Troubleshooting

If you see an error on welcome screen after selecting **Get Started**, follow these steps:

1. Confirm that you've installed all the prerequisites using the links/versions mentioned in Before you start.
2. Ensure that you've restarted the machine after installing the prerequisites.
3. Ensure that you're running the tool in *admin mode* using the **Run as administrator** option when starting the application.
4. Ensure that your PowerShell module path is set correctly using these steps:
   a. Go to **Edit system environment variables**.
   b. Add this path in the **PsModulePath** user variable: `C:\Program Files\PowerShell\7\Modules`.
   c. Move this up and keep at the top.
   d. Restart the tool in administrator mode.

If you're unable to install/uninstall due to another app/process installation, follow these steps:

1. Right click on the task bar to open Task Manager. If necessary, choose **More Details**.
2. On the **Details** tab, look for *msiexec.exe* and select **End Task**.
3. Try to install or uninstall again, or, wait until the installation is finished.

To report issues to Microsoft:

1. **Collect logs** - Find the logs generated by the migration assistant on your local machine at the following path: `C:\Users\<username>\AppData\Local\Temp\MigrationAssistantforMicrosoftPurviewDLP\Logs`.
2. **Send email** - Send us the details about your issue along with the latest logs at dlpmigrations@microsoft.com or contact your CXE / FastTrack / Microsoft partner to share

your feedback and suggestions.

Report errors and provide feedback

To report errors, raise feature requests, or share more feedback, reach out to us at dlpmigrations@microsoft.com or contact your CXE / FastTrack / Microsoft partner.

Telemetry notice

Data Collection: This software might collect information about you and your use of the software, and send it to Microsoft. Microsoft might use this information to improve our products and services. If you want to turn off telemetry, you can reach out to us, and we provide you with a separate version of the tool with telemetry turned off.

There are also some features in the software that might enable you and Microsoft to collect data from users of your applications. If you use these features, you must comply with applicable law, including providing appropriate notices to users of your applications with a copy of Microsoft's privacy statement. Our privacy statement is located at Microsoft Privacy Statement. You can learn more about data collection and use in the help documentation and our privacy statement. Your use of the software operates as your consent to these practices.

See also

For related information, see the following resources:

- Learn about Microsoft Purview Data Loss Prevention migration assistant for Symantec and Forcepoint
- Get started with Microsoft Purview Data Loss Prevention migration assistant for Symantec and Forcepoint
- Learn about data loss prevention
- Create and Deploy data loss prevention policies

ⓘ**Note:** The author created this article with assistance from AI. Learn more

Last updated on 06/26/2026

Data Loss Prevention policy reference

Microsoft Purview Data Loss Prevention (DLP) policies have many components to configure. To create an effective policy, you need to understand what the purpose of each component is and how its configuration alters the behavior of the policy. This article provides a detailed anatomy of a DLP policy.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Recommended reading

1. Administrative units
2. Learn about Microsoft Purview Data Loss Prevention
3. Plan for data loss prevention (DLP) - by working through this article you will:
   a. Identify stakeholders
   b. Describe the categories of sensitive information to protect
   c. Set goals and strategy
4. Collection Policies solution overview
5. Collection policy reference
6. Data Loss Prevention policy reference - this article introduces all the components of a DLP policy and how each one influences the behavior of a policy
7. Design a DLP policy - this article walks you through creating a policy intent statement and mapping it to a specific policy configuration.
8. Create and Deploy data loss prevention policies - This article presents some common policy intent scenarios that you map to configuration options, then it walks you through configuring those options.
9. Learn about investigating data loss prevention alerts - This article introduces you to the lifecycle of alerts from creation, through final remediation and policy tuning. It also introduces you to the tools you use to investigate alerts.

DLP platform considerations

Also, you need to be aware of these constraints of the platform:

- Maximum number of MIP + MIG policies in a tenant: 10,000
- Maximum size of a DLP policy (100 KB)
- Maximum number of DLP rules:
  - In a policy: Limited by the size of the policy
  - In a tenant: 600
- Maximum size of an individual DLP rule: 100 KB (102,400 characters)
- Generate Incident Report evidence limit: 100, with each SIT evidence, in proportion of occurrence
- Maximum size of text scanned from a file: The first 2 million characters (~2 MB) of extractable text. If a file exceeds this limit, first two million characters are scanned, and a "Document didn't complete scanning" signal is emitted.
- Maximum number of nested levels of data scanned from a file: The first three levels. If a file exceeds this limit, data in the first three levels are scanned, and a *Document didn't complete scanning* event is emitted.
- Regex size limit for all matches predicted: 20 KB
- Policy name length limit: 64 characters
- Policy rule length limit: 64 characters
- Comment length limit: 1,024 characters
- Description length limit: 1,024 characters
- Maximum size of Endpoint DLP Settings: 16,384 characters

Default DLP policies

There are four default DLP policies that are created in every tenant. These policies are designed to help you get started with DLP and provide a baseline level of protection for your organization. You can edit or delete these policies as needed.

Default Office 365 DLP policy

This policy detects the presence of credit card numbers in externally shared documents and emails. For more information, see Learn about the default data loss prevention policy for Microsoft 365 Copilot location

Default policy for Teams

The default DLP policy for Teams tracks all the credit card numbers shared internally and externally to the organization. For more information see, Learn about the default data loss prevention policy in Microsoft Teams

Defaul policy for devices

This policy detects the presence of credit card numbers in files on devices when users perform specific activities (such as printing a file). When detected, the activity is only audited (not blocked). Admins will receive an alert, but policy tips won't be displayed to users. You can edit these actions at any time. For more information, see Learn about the default DLP policy for device.

Default DLP policy for Microsoft 365 Copilot location

Prevent data leakage and oversharing by restricting Microsoft 365 Copilot and agents from using sensitive info in Copilot interactions. For more information, see Learn about the default DLP policy for Microsoft 365 Copilot location.

Policy templates

There are five types of DLP policy templates across two categories.

Enterprise applications & devices

- enhanced policies
- policies that can detect and protect types of **Financial** information.
- policies that can detect and protect types of **Medical and health** information.
- policies that can detect and protect types of **Privacy** information.
- A **Custom** policy template that you can use to build your own policy if none of the others meet your organization's needs.

The following table lists all policy templates and the sensitive information types (SIT) that they cover.

Expand table

Category	Template	SIT
Financial	Australia Financial Data	- SWIFT code - Australia tax file number - Australia bank account number - Credit card number
Financial	Canada Financial data	- Credit card number - Canada bank account number

Category	Template	SIT
Financial	France Financial data	- Credit card number - EU debit card number
Financial	Germany Financial Data	- Credit card number - EU debit card number
Financial	Israel Financial Data	- Israel bank account number - SWIFT code - Credit card number
Financial	Japan Financial Data	- Japan bank account number - Credit card number
Financial	PCI Data Security Standard (PCI DSS)	- Credit card number
Financial	Saudi Arabia Anti-Cyber Crime Law	- SWIFT code - International banking account number (IBAN)
Financial	Saudi Arabia Financial Data	- Credit card number - SWIFT code - International banking account number (IBAN)
Financial	UK Financial Data	- Credit card number - EU debit card number - SWIFT code
Financial	US Financial Data	- Credit card number - U.S. bank account number - ABA Routing Number
Financial	U.S. Federal Trade Commission (FTC) Consumer Rules	- Credit card number - U.S. bank account number - ABA Routing Number
Financial	U.S. Gramm-Leach-Bliley Act (GLBA) Enhanced	- Credit card number - U.S. bank account number - U.S. Individual Taxpayer Identification Number (ITIN) - U.S. social security number (SSN) - U.S./U.K. passport number -U.S. driver's license number - All Full Names - U.S. Physical Addresses
Financial	U.S. Gramm-Leach-Bliley Act (GLBA)	- Credit card number - U.S. bank account number - U.S. Individual Taxpayer Identification

Category	Template	SIT
		Number (ITIN) - U.S. social security number (SSN)
Medical and health	Australia Health Records Act (HRIP Act) Enhanced	- Australia tax file number - Australia medical account number - All Full Names - All Medical Terms And Conditions - Australia Physical Addresses
Medical and health	Australia Health Records Act (HRIP Act)	- Australia tax file number - Australia medical account number
Medical and health	Canada Health Information Act (HIA)	- Canada passport number - Canada social insurance number - Canada health service number - Canada Personal Health Identification Number
Medical and health	Canada Personal Health Information Act (PHIA) Manitoba	- Canada social insurance number - Canada health service number - Canada Personal Health Identification Number
Medical and health	Canada Personal Health Act (PHIPA) Ontario	- Canada passport number - Canada social insurance number - Canada health service number - Canada Personal Health Identification Number
Medical and health	U.K. Access to Medical Reports Act	- U.K. national health service number - U.K. national insurance number (NINO)
Medical and health	U.S. Health Insurance Act (HIPAA) Enhanced	- International classification of diseases (ICD-9-CM) - International classification of diseases (ICD-10-CM) - All Full Names - All Medical Terms And Conditions - U.S. Physical Addresses
Medical and health	U.S. Health Insurance Act (HIPAA)	- International classification of diseases (ICD-9-CM) - International classification of diseases (ICD-10-CM)
Privacy	Australia Privacy Act Enhanced	- Australia driver's license number - Australia passport number - All Full Names

Category	Template	SIT
		- All Medical Terms And Conditions - Australia Physical Addresses
Privacy	Australia Privacy Act	- Australia drivers license number - Australia passport number
Privacy	Australia Personally Identifiable Information (PII) Data	- Australia tax file number - Australia driver's license number
Privacy	Canada Personally Identifiable Information (PII) Data	- Canada driver's license number - Canada bank account number - Canada passport number - Canada social insurance number - Canada health service number - Canada Personal Health Identification Number
Privacy	Canada Personal Information Protection Act (PIPA)	- Canada passport number - Canada social insurance number - Canada health service number - Canada Personal Health Identification Number
Privacy	Canada Personal Information Protection Act (PIPEDA)	- Canada driver's license number - Canada bank account number - Canada passport number - Canada social insurance number - Canada health service number - Canada Personal Health Identification Number
Privacy	France Data Protection Act	- France national ID card (CNI) - France social security number (INSEE)
Privacy	France Personally Identifiable Information (PII) Data	- France social security number (INSEE) - France driver's license number - France passport number - France national ID card (CNI)
Privacy	General Data Protection Regulation (GDPR) Enhanced	- Austria Physical Addresses - Belgium Physical Addresses - Bulgaria Physical Addresses - Croatia Physical Addresses - Cyprus Physical Addresses - Czech Republic Physical Addresses - Denmark Physical Addresses - Estonia Physical Addresses - Finland Physical Addresses - France Physical Addresses

Category	Template	SIT
		- Germany Physical Addresses - Greece Physical Addresses - Hungary Physical Addresses - Ireland Physical Addresses - Italy Physical Addresses - Latvia Physical Addresses - Lithuania Physical Addresses - Luxembourg Physical Addresses - Malta Physical Addresses - Netherlands Physical Addresses - Poland Physical Addresses - Portuguese Physical Addresses - Romania Physical Addresses - Slovakia Physical Addresses - Slovenia Physical Addresses - Spain Physical Addresses - Sweden Physical Addresses - Austria Social Security Number - France Social Security Number (INSEE) - Greece Social Security Number (AMKA) - Hungarian Social Security Number (TAJ) - Spain Social Security Number (SSN) - Austria Identity Card - Cyprus Identity Card - Germany Identity Card Number - Malta Identity Card Number - France National ID Card (CNI) - Greece National ID Card - Finland National ID - Poland National ID (PESEL) - Sweden National ID - Croatia Personal Identification (OIB) Number - Czech Personal Identity Number - Denmark Personal Identification Number - Estonia Personal Identification Code - Hungary Personal Identification Number - Luxemburg National Identification Number natural persons - Luxemburg National Identification Number (Non-natural persons) - Italy Fiscal Code - Latvia Personal Code - Lithuania Personal Code - Romania Personal Numerical Code (CNP) - Netherlands Citizen's Service (BSN) Number - Ireland Personal Public Service (PPS)

Category	Template	SIT
		Number - Bulgaria Uniform Civil Number - Belgium National Number - Spain DNI - Slovenia Unique Master Citizen Number - Slovakia Personal Number - Portugal Citizen Card Number - Malta Tax ID Number - Austria Tax Identification Number - Cyprus Tax Identification Number -France Tax Identification Number (numéro SPI.) - Germany Tax Identification Number - Greek Tax identification Number - Hungary Tax identification Number - Netherlands Tax Identification Number - Poland Tax Identification Number - Portugal Tax Identification Number - Slovenia Tax Identification Number - Spain Tax Identification Number - Sweden Tax Identification Number - Austria Driver's License - Belgium Driver's License Number - Bulgaria Driver's License Number - Croatia Driver's License Number - Cyprus Driver's License Number - Czech Driver's License Number - Denmark Driver's License Number - Estonia Driver's License Number - Finland Driver's License Number - France Driver's License Number - German Driver's License Number - Greece Driver's License Number - Hungary Driver's License Number - Ireland Driver's License Number - Italy Driver's License Number - Latvia Driver's License Number - Lithuania Driver's License Number - Luxemburg Driver's License Number - Malta Driver's License Number - Netherlands Driver's License Number - Poland Driver's License Number - Portugal Driver's License Number - Romania Driver's License Number - Slovakia Driver's License Number - Slovenia Driver's License Number - Spain Driver's License Number - Sweden Driver's License Number

Category	Template	SIT
		- Austria Passport Number - Belgium Passport Number - Bulgaria Passport Number - Croatia Passport Number - Cyprus Passport Number - Czech Republic Passport Number - Denmark Passport Number - Estonia Passport Number - Finland Passport Number - France Passport Number - German Passport Number - Greece Passport Number - Hungary Passport Number - Ireland Passport Number - Italy Passport Number - Latvia Passport Number - Lithuania Passport Number - Luxemburg Passport Number - Malta Passport Number - Netherlands Passport Number - Poland Passport - Portugal Passport Number - Romania Passport Number - Slovakia Passport Number - Slovenia Passport Number - Spain Passport Number - Sweden Passport Number - EU Debit Card Number - All Full Names
Privacy	General Data Protection Regulation (GDPR)	- EU debit card number - EU driver's license number - EU national identification number - EU passport number - EU social security number or equivalent identification - EU Tax identification number
Privacy	Germany Personally Identifiable Information (PII) Data	- Germany driver's license number - Germany passport number
Privacy	Israel Personally Identifiable Information (PII) Data	- Israel national identification number
Privacy	Israel Protection of Privacy	- Israel national identification number - Israel bank account number

Category	Template	SIT
Privacy	Japan Personally Identifiable Information (PII) Data enhanced	- Japan Social Insurance Number (SIN) - Japan My Number - Personal - Japan passport number - Japan driver's license number - All Full Names - Japan Physical Addresses
Privacy	Japan Personally Identifiable Information (PII) Data	- Japan resident registration number - Japan Social Insurance Number (SIN)
Privacy	Japan Protection of Personal Information Enhanced	- Japan Social Insurance Number (SIN) - Japan My Number - Personal - Japan passport number - Japan driver's license number - All Full Names - Japan Physical Addresses
Privacy	Japan Protection of Personal Information	- Japan resident registration number - Japan Social Insurance Number (SIN)
Privacy	Saudi Arabia Personally Identifiable (PII) Data	- Saudi Arabia National ID
Privacy	U.K. Data Protection Act	- U.K. national insurance number (NINO) - U.S./U.K. passport number - SWIFT code
Privacy	U.K. Privacy and Electronic Communications Regulations	- SWIFT code
Privacy	U.K. Personally Identifiable Information (PII) Data	- U.K. national insurance number (NINO) - U.S./U.K. passport number
Privacy	U.K. Personal Information Online Code of Practice (PIOCP)	- U.K. national insurance number (NINO) - U.K. national health service number - SWIFT code
Privacy	U.S Patriot Act Enhanced	- Credit card number - U.S. bank account number - U.S. Individual Taxpayer Identification Number (ITIN) - U.S. social security number (SSN) - All Full Names - U.S. Physical Addresses
Privacy	U.S. Patriot Act	- Credit card number - U.S. bank account number - U.S. Individual Taxpayer Identification

Category	Template	SIT
		Number (ITIN) - U.S. social security number (SSN)
Privacy	U.S. Personally Identifiable Information (PII) Data Enhanced	- U.S. Individual Taxpayer Identification Number (ITIN) - U.S. social security number (SSN) - U.S./U.K. passport number - All Full Names - U.S. Physical Addresses
Privacy	U.S. Personally Identifiable Information (PII) Data	- U.S. Individual Taxpayer Identification Number (ITIN) - U.S. social security number (SSN) - U.S./U.K. passport number
Privacy	U.S. State Breach Notification Laws Enhanced	- Credit card number - U.S. bank account number -U.S. driver's license number - U.S. social security number (SSN) - All Full Names - U.S./U.K. passport number - All Medical Terms And Conditions
Privacy	U.S. State Breach Notification Laws	- Credit card number - U.S. bank account number -U.S. driver's license number - U.S. social security number (SSN)
Privacy	U.S. State Social Security Number Confidentiality Laws	- U.S. social security number (SSN)

Inline web traffic policies

- A **Custom** policy template that you can use to build your own policy.

Policy scoping

See, Administrative units to make sure you understand the difference between an unrestricted admin and an administrative unit restricted admin.

DLP policies are scoped at two different levels. The first level applies unrestricted admin scope policies to all of the following in your organization (depending on the locations that are selected) or to subgroups of your organization, called Administrative Unit restricted policies:

- users

- groups
- distribution groups
- accounts
- sites
- cloud app instances
- on-premises repositories
- Fabric and Power BI workspaces

At this level, an administrative unit restricted admin will only be able to pick from the administrative units that they're assigned to. DLP support admin unit scoping for some of the locations protected under Enterprise applications & devices.

The second level of DLP policy scoping is by the locations that DLP supports. At this level, both unrestricted and administrative unit restricted administrators see only the users, distribution groups, groups, and accounts that were included in the first level of policy scoping and that are available for that location.

Support for adding SharePoint sites to Administrative Units

Microsoft Purview supports adding SharePoint sites to existing administrative units. When you assign a DLP policy for the SharePoint location to an administrative unit, the policy will only apply to the sites that are part of that administrative unit. The option to further edit the scope to include or exclude specific sites isn't available. The policy applies to all sites that are part of the administrative unit.

Here's an example use case:

Contoso has created an Entra ID administrative unit for the engineering department and has assigned certain administrators to manage the users and groups for that department. The engineering department has a SharePoint site that is used to store sensitive information. Contoso wants to ensure that the DLP policy for the engineering department SharePoint sites only applies to the SharePoint site that is part of the administrative unit. By assigning the DLP policy to the administrative unit, the policy will only apply to the SharePoint site that is part of that administrative unit. Also, the administrative unit restricted administrator will only be able to manage the DLP policy for that site and only see policy match result data for the administrative unit in activity explorer and the alert dashboard.

Unrestricted policies

Unrestricted policies are created and managed by users in these role groups:

- Compliance administrator
- Compliance data administrator
- Information Protection
- Information Protection Admin
- Security administrator

See, the Permissions article for more details.

Unrestricted administrators can manage all policies and see all the alerts and events that flow from policy matches into the Alerts dashboard and DLP Activity Explorer.

Administrative unit restricted policies

Administrative units are subsets of your Microsoft Entra ID directory and are created for the purposes of managing collections of users, groups, distribution groups, and accounts. These collections are typically created along business group lines or geopolitical areas. Administrative units have a delegated administrator who is associated with an administrative unit in the role group. These are called administrative unit restricted admins.

DLP supports associating policies with administrative units. See Administrative units for implementation details in the Microsoft Purview portal. Administrative unit admins need to be assigned to one of the same roles or role groups as administrators of unrestricted DLP policies in order to create and manage DLP policies for their administrative unit.

Expand table

DLP Administrative Role Group	Can
Unrestricted administrator	- create and scope DLP policies to entire organization - edit all DLP policies - create and scope DLP policies to administrative units - view all alerts and events from all DLP policies
Administrative Unit Restricted administrator - must be a member of/assigned to a role group/role that can administer DLP	- create and scope DLP policies only to the administrative unit that they're assigned to - edit DLP policies that are associated to their administrative unit - view alerts and events only from the DLP policies that are scoped to their administrative unit

Locations

A DLP policy can find and protect items that contain sensitive information across multiple locations.

Expand table

Location	Supports Administrative Units	Include/Exclude scope	Data state	Other prerequisites
Exchange Online	Yes	- distribution groups assigned or dynamic - security groups - non email enabled security groups assigned or dynamic) - Microsoft 365 groups assigned or dynamic	data-in-motion	No
SharePoint	Yes	site locations and, in preview, adaptive scopes at the policy level. If the policy is scoped to an administrative unit that includes SharePoint sites, the policy will only apply to all sites in the administrative unit, no further scoping is possible	data-at-rest data-in-use	No
OneDrive	Yes	- Distribution groups - Security groups - Non email enabled security groups - Microsoft 365 groups (Group members only, not the group as an entity)	data-at-rest data-in-use	No
Teams chat and channel messages	Yes	- Distribution groups - Security groups - Mail-enabled security groups - Microsoft 365 groups (Group members only, not the group as an entity)	data-in-motion data-in-use	See Scope of DLP protection
Instances	No	Cloud app instance	data-at-rest	- Use data loss prevention policies for non-Microsoft cloud apps
Devices	Yes	- Distribution groups - Security groups - Non-email enabled security groups	data-in-use data-	- Learn about Endpoint data loss prevention - Get started with

Location	Supports Administrative Units	Include/Exclude scope	Data state	Other prerequisites
		- Microsoft 365 groups (Group members only, not the group as an entity) - Dynamic groups	in-motion	Endpoint data loss prevention - Configure device proxy and internet connection settings for Microsoft Purview Endpoint DLP
On-premises repositories (file shares and SharePoint)	No	Repository	data-at-rest	- Learn about the data loss prevention on-premises repositories - Get started with data loss prevention for on-premises repositories
Fabric and Power BI	No	Workspaces	data-in-use	No
Non-Microsoft connected apps (preview)	No	Cloud app instance	data-at-rest	- Use DLP and sensitivity label policies for non-Microsoft connected apps - Only available in the **Custom** policy template - Set up a Microsoft Defender for Cloud Apps connector
Microsoft 365 Copilot (preview)	No	Account or Distribution group	data-at-rest data-in-use	- Only available in the **Custom** policy template
Managed cloud apps	No	Account or Distribution group	data-in-motion	- Only available in the **Custom** policy template
Unmanaged cloud apps	No	Account or Distribution group	data-in-motion	- Only available in the **Custom** policy template

> **Note**
>
> In the **Devices** location, the **Devices and device groups** setting supports only **Audit or restrict activities on devices**. You can't use it with **Audit or restrict activities when users access sensitive sites in the Microsoft Edge browser on Windows devices** under **Rule Actions**.

Exchange location scoping

If you choose to include specific distribution groups in Exchange, the DLP policy is scoped to the emails sent by members of that group or sent to members of that group. Similarly, excluding a distribution group excludes all the emails sent by the members of that distribution group or from policy evaluation.

Expand table

Group Type	Membership Type	Supported during Policy Creation	Supported during Rule Evaluation	Notes
Non-Mail Enabled Security Groups	Assigned	Yes	No	Enabled for specific customers only
Non-Mail Enabled Security Groups	Dynamic	Yes	No	
Mail-Enabled Security Groups	Assigned	Yes	Yes	
Distribution Groups	Assigned	Yes	Yes	
Distribution Groups	Dynamic	Yes	Yes	
Microsoft 365 Groups	Assigned	Yes	Yes	
Microsoft 365 Groups	Dynamic	Yes	Yes	
Adaptive Scopes	Dynamic	No	No	

Expand table

Sender is	Recipient is	Resultant behavior
In scope	N/A	Policy is applied

Sender is	Recipient is	Resultant behavior
Out of scope	In scope	Policy isn't applied

Exchange location scope calculation

Here's an example of how Exchange location scope is calculated:

Say you have four users in your org and two distribution groups that you use for defining Exchange location inclusion and exclusion scopes. Group membership is set up like this:

Expand table

Distribution Group	Membership
Group1	User1, User2
Group2	User2, User3
No group	User4

Expand table

Include setting	Exclude setting	Policy applies to	Policy doesn't apply to	Explanation of behavior
All	None	All senders in the Exchange org (User1, User2, User3, User4)	N/A	When neither are defined, all senders are included
Group1	None	Member senders of Group1 (User1, User2)	All senders who aren't members of Group1 (User3, User4)	When one setting is defined and the other isn't, the defined setting is used
All	Group2	All senders in the Exchange org who aren't members of Group2 (User1, User4)	All senders who are members of Group2 (User2, User3)	When one setting is defined and the other isn't, the defined setting is used
Group1	Group2	User1	User2, User3, User4	Exclude overrides include

You can choose to scope a policy to the members of distribution lists, dynamic distribution groups, and security groups. A DLP policy can contain no more than 50 such inclusions and exclusions.

OneDrive location scoping

When scoping a policy for OneDrive locations, in addition to applying your DLP policies to all users and groups in your organization, you can limit the scope of a policy to specific users and groups. DLP supports scoping policies to up to 100 individual users.

For instance, if you want to include more than 100 users, you must first put those users in distribution groups or security groups, as appropriate. You can then scope your policy to up to 50 groups.

In some cases, you might want to apply a policy to one or two groups, plus two or three individual users who don't belong to either of those groups. Here, the best practice is to **put those two or three individuals into a group of their own**. This is the only way to make sure that the policy is scoped to all intended users.

The reason for this is that, when you list only users, DLP adds all of the users specified to the policy scope. Similarly, when you add only groups, DLP adds all the members of all the groups to the policy scope.

Say you have the following groups and users:

Expand table

Distribution Group	Membership
Group1	User1, User2
Group2	User2, User3

If you limit the scope of a policy to *only* users or *only* groups, the DLP applies the policy to users as illustrated in the following table:

Expand table

Specified scope	DLP Scope evaluation behavior	Users in scope
(Users only) User1 User2	DLP takes the union of the specified users	User1, User2
(Groups only) Group1 Group2	DLP takes the union of the specified groups	User1, User2, User3

However, when users and groups are mixed in the scoping configuration, things get complicated. Here's why: DLP only scopes policies to users to **the intersection** of the listed groups and users.

DLP uses the following order of operations when determining which users and groups to include in the scope:

1. Evaluate the union of group membership
2. Evaluate the union of users
3. Evaluate the intersection of group members and users, that is, where the results overlap

It then applies the scope of the policy to the intersection of group members and users.

Let's extend our example, working with the same set of groups, and let's add User4, who isn't in a group:

Expand table

Distribution Group	Membership
Group1	User1, User2
Group2	User2, User3
No group	User 4

The following table explains how policy scoping works in cases where users and groups are both included in the scoping instructions.

Expand table

Specified scope	DLP Scope evaluation behavior	Users in scope

Group1
Group2
User3
User4
	First evaluation: Union of groups:
	`(Group1 + Group2) = User1, User2, User3`
	
	Second evaluation: Union of users:
	`(User3 + User4) = User3, User4`
	
	Third evaluation: Intersection of groups and Users (the overlap):
	
	`(Group1 + Group2) = User1, User2, User3`
	
	`(User3 + User4) = User3, User4`
		User3
		(User3 is the only user that appears in the results of both the first and second evaluations.)

Specified scope	DLP Scope evaluation behavior	Users in scope

Group1
Group2
User1
User3
User4
	First evaluation: Union of groups: `(Group1 + Group2) = User1, User2, User3`
	
	Second evaluation: Union of users: `(User1 + User3 + User4) = User1, User3, User4`
	
	Third evaluation: Intersection of groups and Users (the overlap):
	
	`(Group1 + Group2) = User1, User3`
	
	`(User1 + User3, User4) = User1, User3, User4`
		User1, **User3**
		(These are the only users that appear in the results of both the first and second evaluations.)

SharePoint location scoping

When scoping a policy for SharePoint locations, you can limit the scope of a policy to specific SharePoint sites and adaptive scopes. DLP supports scoping policies to up to 100 sites.

Device scoping

Microsoft Purview Data Loss Prevention (DLP) policies that include the **Devices** location can be scoped by **users** and by **devices**. A policy is enforced on an endpoint only when *both* the user and the device are included in the policy scope. If a user is in scope but the device is not, the policy isn't applied. Similarly, if a device is in scope but the user is not, the policy isn't applied.

Use 101.25072 or higher build for this feature support on macOS.

Device scoping does not support Microsoft Entra registered.

Here's how to configure the scope of a DLP policy for different outcomes.

Expand table

If you want to target the policy to...	Set user scope to...	Set device scope to...	Example use case
All users on all	**All users and groups**	**All devices and device groups**	Use this for general enforcement of DLP policies on all devices in your

If you want to target the policy to...	Set user scope to...	Set device scope to...	Example use case
onboarded devices			organization. This is the default setting for DLP policies.
All users on specific devices	**All users and groups**	either **All devices and device groups** with **Exclude devices and device groups** and add the devices to be excluded or **Specific devices and device groups** and add the devices to be included	Use this to apply a restrictive policy to kiosk devices that will be used by multiple users.
Specific users on all onboarded devices	either **All users and groups** with **Exclude users and groups** and add the users to be excluded or **Specific users and groups** and add the users to be included	**All devices and device groups**	Use this to help control data leakage by specific users on all devices in your organization.
Specific users on specific devices	**Specific users and groups**	**Specific devices and device groups**	Say you have special use devices in payroll that are used for printing checks and there are only a few accounts that are allowed to use those devices for printing checks. You can scope a restrictive endpoint DLP policy to those user accounts on those specific devices

Location support for how content can be defined

DLP policies detect sensitive items by matching them to a sensitive information type (SIT), or to a sensitivity label or a retention label. Each location supports different methods of defining sensitive content. How content can be defined when you combine locations in a policy, can change from how it can be defined when limited to a single location.

> **Important**
>
> When you select multiple locations for a policy, a "no" value for a content definition category takes precedence over "yes" value. For example, when you select SharePoint sites

only, the policy supports detecting sensitive items by one or more of SIT, by sensitivity label or by retention label. But, when you select SharePoint sites ***and*** Teams chat and channel messages locations, the policy will only support detecting sensitive items by SIT.

Expand table

Location	Content can be defined by SIT	Content can be defined sensitivity label	Content can be defined by retention label
Exchange email online	Yes	Yes	No
SharePoint in Microsoft 365 sites	Yes	Yes	Yes
OneDrive for work or school accounts	Yes	Yes	Yes
Teams Chat and Channel messages	Yes	No	No
Devices	Yes	Yes	No
Instances	Yes	Yes	Yes
On-premises repositories	Yes	Yes	No
Fabric and Power BI	Yes	Yes	No
Microsoft 365 Copilot (preview)	No	Yes	No
Managed cloud apps	Yes	No	No
Unmanaged cloud apps	Yes	Yes	No

DLP supports using trainable classifiers as a condition to detect sensitive information. Content can be defined by trainable classifiers in Exchange, SharePoint sites, OneDrive accounts, Teams Chat and Channels, Devices, and unmanaged cloud apps. For more information, see Trainable Classifiers.

Note

DLP supports detecting sensitivity labels on emails and attachments. For more information, see Use sensitivity labels as conditions in DLP policies.

Rules

Rules are the business logic of DLP policies. They consist of:

- **Conditions** that when matched, trigger the policy
- **Actions** that determine the activities included and outcomes of a match.
- **User notifications** to inform your users when they're doing something that triggers a policy and help educate them on how your organization wants sensitive information treated
- **User Overrides** when configured by an admin, allow users to selectively override a blocking action
- **Incident reports** that notify admins and other key stakeholders when a rule match occurs
- **Additional options** which define the priority for rule evaluation and can stop further rule and policy processing.

A policy contains one or more rules. Rules are executed sequentially, starting with the highest-priority rule in each policy.

How DLP classification works

DLP evaluates an item for sensitive information when the item is created, read, or modified. Evaluation is also initiated by on demand classification. However, events such as **DLP Rule Matched** only appear in Audit log or activity explorer when a user attempts an activity that matches a DLP policy.

Here's a list of some of the user activities that DLP can monitor and take actions on:

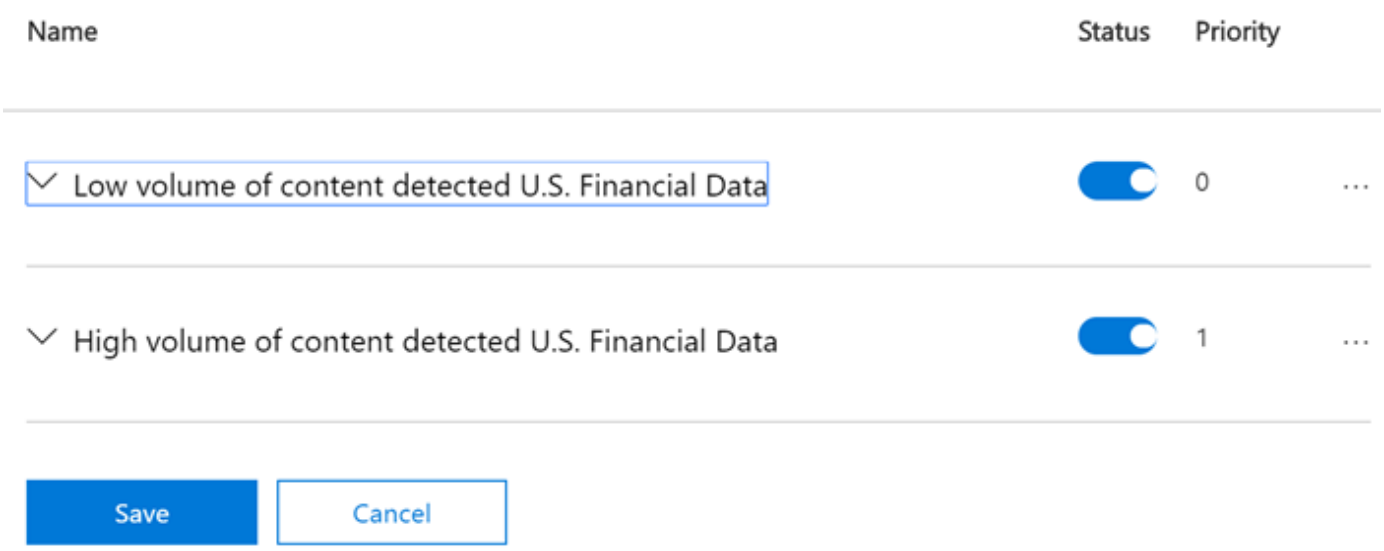
- Text upload via Microsoft Edge browser using integrated capabilities
- File upload via Microsoft Edge browser using integrated capabilities
- File download via Microsoft Edge browser using integrated capabilities
- Cut/copy data via Microsoft Edge browser using integrated capabilities
- Paste data via Microsoft Edge browser using integrated capabilities
- Print data via Microsoft Edge browser using integrated capabilities
- Print data from other locations

- Copy to removable media
- Copy to network share
- Copy to clipboard
- Transfer using Bluetooth
- File accessed by unallowed app
- Paste to browsers other than Microsoft Edge
- Print data
- Transfer using remote desktop
- An item that is created, read, or modified will match a DLP rule and policy on the client if the conditions and user activity are met. This is audited as file activity, such as FileRead, or FileRenamed.
- If an activity is met, then a DLP rule match event appears in activity explorer as a 'DLP Rule Matched' event. An event describing the mode of egress will also be generated.
- Policies take actions and actions are different from conditions. A rule can match on a file even if no actions are performed.

The priority by which rules are evaluated and applied

Hosted service locations

For the hosted service locations, like Exchange, SharePoint, and OneDrive, each rule is assigned a priority in the order in which it's created. This means that the rule created first has first priority, the rule created second has second priority, and so on.



When content is evaluated against rules, the rules are processed in priority order. If content matches multiple rules, the first rule evaluated that has the *most* restrictive action is enforced. For example, if content matches all of the following rules, *Rule 3* is enforced because it's the highest priority, most restrictive rule:

- Rule 1: only notifies users
- Rule 2: notifies users, restricts access, and allows user overrides
- *Rule 3: notifies users, restricts access, and doesn't allow user overrides*
- Rule 4: restricts access

Rules 1, 2, and 4 would be evaluated, but not applied. In this example, matches for all of the rules are recorded in the audit logs and shown in the DLP reports, even though only the most restrictive rule is applied.

You can use a rule to meet a specific protection requirement, and then use a DLP policy to group together common protection requirements, such as all of the rules needed to comply with a specific regulation.

For example, you might have a DLP policy that helps you detect the presence of information subject to the Health Insurance Portability and Accountability Act (HIPAA). This DLP policy could help protect HIPAA data (the "what") across all SharePoint sites and all OneDrive sites (the "where") by finding any document containing this sensitive information that's shared with people outside your organization (the conditions) and then blocking access to the document and sending a notification (the actions). These requirements are stored as individual rules and grouped together as a DLP policy to simplify management and reporting.



For endpoints

When an item matches multiple DLP rules, DLP goes uses through a complex algorithm to decide which actions to apply. Endpoint DLP applies the aggregate or sum of most restrictive actions. DLP uses these factors when making the calculation.

Policy priority order When an item matches multiple policies and those policies have identical actions, the actions from the highest priority policy are applied.

Rule priority order When an item matches multiple rules in a policy and those rules have identical actions, the actions from the highest priority rule are applied.

Mode of the policy When an item matches multiple policies and those policies have identical actions, the actions from all policies that are in **Turn it on** state (enforce mode) are applied preferentially over the policies in **Run the policy in simulation mode with policy tips** and **Run the policy in simulation mode** state.

Run the policy in simulaiton mode changes the action that's applied at runtime

When policies are matched and in the **Run the policy in simulation** state, they apply different actions than they would if they are in the **Turn it on right away** state.

Expand table

Action configured in the policy is	**When policy state is** Run the policy in simulation mode **and** Show policy tips in simulation mode **is off, the applied action is**	**When policy state is** Run the policy in simulation mode **and** Show policy tips while in simulation mode **is on, the applied action is**
Block	Audit	Block with override
Block with override	Audit	Block with override
Audit only	Audit	Audit
Allow	Allow	Allow

Action When an item matches multiple policies and those policies differ in actions, the aggregate or sum of the most restrictive actions are applied.

Authorization groups configuration When an item matches multiple policies and those policies differ in action, the aggregate, or sum, of the most restrictive actions are applied.

override options When an item matches multiple policies and those policies differ in the override option, actions are applied in this order:

No override > *Allow override*

Here are scenarios that illustrate the runtime behavior. For the first three scenarios, you have three DLP policies configured like this:

Expand table

Policy name	**Condition to match**	**Action**	**Policy priority**
ABC	Content contains credit card number	Block print, audit all other user egress activities	0
MNO	Content contains credit card number	Block copy to USB, audit all other user egress activities	1
XYZ	Content contains U.S. social security number	Block copy to clipboard, audit all other user egress activities	2

Item contains credit card numbers

An item on a monitored device contains credit card numbers, so it matches policy ABC and policy MNO. Both ABC and MNO are in *Turn it on* mode.

Policy	Cloud egress action	Copy to clipboard action	Copy to USB action	Copy to network share action	Unallowed apps action	Print action	Copy via Bluetooth action	Copy to remote desktop action
ABC	Audit	Audit	Audit	Audit	Audit	**Block**	Audit	Audit
MNO	Audit	Audit	**Block**	Audit	Audit	Audit	Audit	Audit
Actions applied at runtime	Audit	Audit	**Block**	Audit	Audit	**Block**	Audit	Audit

Item contains credit card numbers and U.S. social security numbers

An item on a monitored device contains credit card numbers and U.S. social security numbers, so this item matches policy ABC, policy MNO, and policy XYZ. All three policies are in *Turn it on* mode.

Policy	Cloud egress action	Copy to clipboard action	Copy to USB action	Copy to network share action	Unallowed apps action	Print action	Copy via Bluetooth action	Copy to remote desktop action
ABC	Audit	Audit	Audit	Audit	Audit	**Block**	Audit	Audit
MNO	Audit	Audit	**Block**	Audit	Audit	Audit	Audit	Audit
XYZ	Audit	**Block**	Audit	Audit	Audit	**Block**	Audit	Audit
Actions applied at runtime	Audit	**Block**	**Block**	Audit	Audit	**Block**	Audit	Audit

Item contains credit card numbers, different policy state

An item on a monitored device contains credit card number, so it matches policy ABC and policy MNO. Policy ABC is in *Turn it on* mode and policy *MNO* is in **Run the policy in simulation mode** state.

Policy	Cloud egress action	Copy to clipboard action	Copy to USB action	Copy to network share action	Unallowed apps action	Print action	Copy via Bluetooth action	Copy to remote desktop action
ABC	Audit	Audit	Audit	Audit	Audit	**Block**	Audit	Audit

Policy	Cloud egress action	Copy to clipboard action	Copy to USB action	Copy to network share action	Unallowed apps action	Print action	Copy via Bluetooth action	Copy to remote desktop action
MNO	Audit	Audit	**Block**	Audit	Audit	Audit	Audit	Audit
Actions applied at runtime	Audit	Audit	Audit	Audit	Audit	**Block**	Audit	Audit

Item contains credit card numbers, different override configuration

An item on a monitored device contains credit card number, so it matches policy ABC and policy MNO. Policy ABC is in *Turn it on* state and policy *MNO* is in *Turn it on* state. They have different *Override* actions configured.

Policy	Cloud egress action	Copy to clipboard action	Copy to USB action	Copy to network share action	Unallowed apps action	Print action	Copy via Bluetooth action	Copy to remote desktop action
ABC	Audit	Audit	**Block with override**	Audit	Audit	**Block**	Audit	Audit
MNO	Audit	Audit	**Block without override**	Audit	Audit	Audit	Audit	Audit
Actions applied at runtime	Audit	Audit	**Block without override**	Audit	Audit	**Block**	Audit	Audit

Item contains credit card numbers, different authorization groups configuration

An item on a monitored device contains credit card number, so it matches policy ABC and policy MNO. Policy ABC is in *Turn it on* state and policy *MNO* is in *Turn it on* state. They have different *authorization group* actions configured.

Policy	Cloud egress action	Copy to clipboard action	Copy to USB action	Copy to network share action	Unallowed apps action	Print action	Copy via Bluetooth action	Copy to remote desktop action
ABC	Audit	Audit	**Auth group A - Block**	Audit	Audit	**Auth group A - Block**	Audit	Audit
MNO	Audit	Audit	**Auth group A - Block with override**	Audit	Audit	**Auth group B - block**	Audit	Audit
Actions applied at runtime	Audit	Audit	**Auth group A - Block**	Audit	Audit	**Auth group A - Block, Auth group B - Block**	Audit	Audit

Conditions

Conditions are where you define what you want the rule to look for and the context in which those items are being used. They tell the rule: when you find an item that looks like *this* and is being used like *that*—it's a match and the rest of the actions in the policy should be taken on it. You can use conditions to assign different actions to different risk levels. For example, sensitive content shared internally might be lower risk and require fewer actions than sensitive content shared with people outside the organization.

> ⓘ **Note**
>
> Users who have non guest accounts in a host organization's Active Directory or Microsoft Entra tenant are considered as people inside the organization.

Content contains

All locations support the **Content contains** condition. You can select multiple instances of each content type and further refine the conditions by using the **Any of these** (logical OR) or **All of these** (logical AND) operators:

- sensitive information types
- sensitivity labels
- retention labels
- Trainable Classifiers

The rule will only look for the presence of any **sensitivity labels** and **retention labels** you pick.

SITs have a predefined **confidence level** which you can alter if needed. For more information, see More on confidence levels.

> **Important**
>
> SITs have two different ways of defining the maximum unique instance count parameters. To learn more, see **Instance count supported values for SIT**.

> Duplicate SIT words across different attachments in a single email are only counted once toward the unique count.

Adaptive Protection in Microsoft Purview

Adaptive protection integrates Microsoft Purview Insider Risk Management risk profiles into DLP policies so that DLP can help protect against dynamically identified risky behaviors. When configured in insider risk management, the **Insider risk level for Adaptive Protection is** will show up as condition for Exchange Online, Devices, Teams, and unmanaged cloud apps locations. Refer to Learn about Adaptive Protection in Data Loss Prevention for more details.

Conditions that Adaptive Protection supports

- Insider risk level for Adaptive Protection is supported with these values:
- Elevated risk level
- Moderate risk level
- Minor risk level

Condition context

The available context options change depending on which location you choose. If you select multiple locations, only the conditions that the locations have in common are available.

Conditions Exchange supports

> **Note**
>
> DLP policies for Exchange scan non system generated emails and journaling emails.

- Content contains
- Insider risk level for Adaptive Protection is
- Content isn't labeled
- Content is shared from Microsoft 365
- Content is received from
- Sender IP address is
- Header contains words or phrases
- Sender AD Attribute contains words or phrases
- Content character set contains words
- Header matches patterns
- Sender AD Attribute matches patterns
- Recipient AD Attribute contains words or phrases
- Recipient AD Attribute matches patterns
- Recipient is member of
- Document property is
- Any email attachment's content could not be scanned
- Document or attachment is password protected
- Has sender overridden the policy tip
- Sender is a member of
- Any email attachment's content didn't complete scanning
- Recipient address contains words
- File extension is
- Recipient domain is
- Recipient is
- Sender is
- Sender domain is
- Recipient address matches patterns
- Document name contains words or phrases
- Document name matches patterns
- Subject contains words or phrases
- Subject matches patterns

- Subject or body contains words or phrases
- Subject or body matches patterns
- Sender address contains words
- Sender address matches patterns
- Document size equals or is greater than
- Document content contains words or phrases
- Document content matches patterns
- Message size equals or is greater than
- Message type is
- Message importance is

> **Tip**
>
> For more information on the conditions that Exchange supports, including PowerShell values, see: **Data loss prevention Exchange conditions and actions reference**.

Conditions SharePoint supports

- Content contains
- Content is shared from Microsoft 365
- Document property is
- Document could not be scanned
- Document or attachment is password protected
- Document didn't complete scanning
- File extension is
- Document name contains words or phrases
- Document size equals or is greater than
- Document created by
- Document creation date is on or after
- Document creation date is on or before

- Document modification date is on or after
- Document modification date is on or before

Conditions OneDrive accounts support

- Content contains
- Content is shared from Microsoft 365
- Document property is
- Document could not be scanned
- Document or attachment is password protected
- Document didn't complete scanning
- File extension is
- Document name contains words or phrases
- Document size equals or is greater than
- Document created by
- Document is shared
- Document creation date is on or after
- Document creation date is on or before
- Document modification date is on or after
- Document modification date is on or before

Conditions Teams chat and channel messages support

- Content contains
- Insider risk level for Adaptive Protection is
- Content is shared from Microsoft 365
- Recipient domain is
- Recipient is
- Sender is
- Sender domain is

Conditions managed cloud apps support

- Content contains
- File extension is
- Document size equals or is greater than
- Managed or unmanaged devices (the cut/copy data, paste data, and print data activities only support the managed or unmanaged devices condition).

Conditions unmanaged cloud apps support

- Content contains
- Content not fully scanned
- File cannot be scanned
- File extension is
- File is not labeled
- File is password protected
- File size is greater than
- File type is
- Insider risk level for Adaptive Protection is
- URL contains text (preview)

(Preview) The **URL contains text** condition detects when the URL of the unmanaged cloud app contains specified text strings. You can use it as a condition to scope DLP rules to specific URLs, or as an exception to exclude specific URLs from policy enforcement. For example, you can create a rule that blocks uploading sensitive data to any GitHub URL that doesn't contain your organization's name while allowing uploads to your organization's GitHub repositories. This condition is supported for browser and network detection and doesn't require devices to be onboarded to endpoint DLP.

> **Important**
>
> Encrypted files can't be inspected for sensitive info types or trainable classifiers.

Conditions supported for Endpoints

- **Content contains:** Specifies content to be detected. For details on supported file types, see Files scanned for content.

- **Content is not labeled:** Detects content that doesn't have a sensitivity label applied. To help ensure only supported file types are detected, you should use this condition with the **File extension is** or **File type is** conditions. PDF and Office files are supported:

File Type	Format	Monitored file extensions
Word processing	Word, PDF	.doc, .docx, .docm, .dot, dotx, .dotm, .docb, .pdf
Spreadsheet	Excel, CSV, TSV	.xls, .xlsx, .xlt, .xlm, .xlsm, xltx, xltm, xlsb, .xlw, .csv, .tsv
Presentation	PowerPoint	.ppt, .pptx, .pos, .pps, .pptm, .potx, .potm, .ppam, .ppsx

- **Document could not be scanned:** Applies to files that can't be scanned for one of the following reasons:
  - File contains one or more transient text-extraction errors
  - File is password-protected
  - File size exceeds the supported limit (Maximum file sizes: 64 MB for uncompressed files; 256 MB for compressed files)
  - Microsoft classification engine (MCE) timeout or failure

- **Document name contains words or phrases:** Detects documents with file names that contain any of the words or phrases you specify, for example: `file`, `credit card`, `patent`, etc.

- **Document name matches patterns:** Detects documents where the file name matches specific patterns. The evaluation considers the entire path of the document, not just the document's name. The pattern is checked as a string match, meaning it can match any part of the document path. To define the patterns, use wild cards.

> ⓘ **Note**
>
> Due to potential performance issues, this condition will gradually be phased out from Purview Endpoint DLP. We recommend using the 'Document name contains words or phrases' condition instead.

- **Document or attachment is password protected:** Detects only protected files that are open. The following files are supported:
- Archive files (ZIP, .7z, RAR)
  - Office files
  - PDFs
  - Symantec PGP encrypted files

- **Document size equals or is greater than:** Detects documents with file sizes that are equal to or greater than the specified value. DLP only supports content inspection for files less than 64 MB.

  > ⓘ **Important**
  >
  > We recommend setting this condition to detect items that are larger than 10KB

- **File type is:** Detects the following file types:

File type	Apps	Monitored file extensions
Word processing	Word, PDF	doc, .docx, .docm, .dot, dotx, .dotm, .docb, .pdf
Spreadsheet	Excel, CSV, TSV	.xls, .xlsx, .xlt, .xlm, .xlsm, xltx, xltm, xlsb, .xlw, .csv, .tsv
Presentation	PowerPoint	.ppt, .pptx, .pos, .pps, .pptm, .potx, .potm, .ppam, .ppsx
Email	Outlook	.msg

  > ⓘ **Important**
  >
  > The **file extensions** and **file types** options *can't* be used as conditions in the same rule. If you want to use them as conditions in the same policy, they must be in separate rules.

To use the **File type is** condition, you must have one of the following versions of Windows:

- Windows Endpoints (X64):
  - Windows 10 (21H2, 22H2)
    - Windows 10 21H2 update details
    - Windows 10 22H2 update details

- Windows Endpoints (ARM64):
  - Windows 11 (21H2, 22H2)
    - Windows 11 21H2 update details
    - Windows 11 22H2 update details

- **File extension is:** In addition to detecting sensitive information in files with the same extensions as those covered by the **File type is** condition, you can use the **File extension is** condition to detect sensitive information in files with any file extension you need to monitor. To do so, add the necessary file extensions, separated by commas to a rule in your

policy. The **File extension is** condition is supported only for those versions of Windows that support the **File type is** condition. **File extension is** doesn't support archive file types.

> ⚠ **Warning**
>
> Including any of the following file extensions in your policy rules might significantly increase the CPU load: .dll, .exe, .mui, .ost, .pf, .pst.

- **Scanning did not complete:** Applies when the scanning of a file started, but stopped before the entire file was scanned. The primary reason for an incomplete scan is that extracted text within the file exceeds the maximum size allowed. (Maximum sizes for extracted text: Uncompressed files: The first 4 MB of extractable text; Compressed files: N=1000 / Extraction Time = 5 minutes.)

- **Document property is:** Detects documents with custom properties matching specified values. For example: `Department = 'Marketing'`, `Project = 'Secret'`. To specify multiple values for a custom property, use double quotes. For example, "Department: Marketing, Sales". Supported file types are Office and PDF:

File Type	Format	Monitored file extensions
Word processing	Word, PDF	.doc, .docx, .docm, .dot, dotx, .dotm, .docb, .pdf
Spreadsheet	Excel, CSV, TSV	.xls, .xlsx, .xlt, .xlm, .xlsm, xltx, xltm, xlsb, .xlw, .csv, .tsv
Presentation	PowerPoint	.ppt, .pptx, .pos, .pps, .pptm, .potx, .potm, .ppam, .ppsx

- **The user accessed a sensitive website from Microsoft Edge:** For more information, see Help prevent risky user activity by monitoring or restricting access to sensitive service domains.

- **Insider risk level for Adaptive Protection is:** Detects the insider risk level.

See also: Endpoint activities you can monitor and take action on.

Operating system requirements for five conditions

- Document could not be scanned
- Document name contains words or phrases
- Document name matches patterns
- Document size equals or is greater than
- Scanning did not complete

To use any of these conditions, your endpoint devices must be running one of the following operating systems:

- **Windows 11 23H2:** December 4, 2023—KB5032288 (OS Builds 22621.2792 and 22631.2792) Preview
- **Windows 11 22H2:** December 4, 2023—KB5032288 (OS Builds 22621.2792 and 22631.2792) Preview - Microsoft Support
- **Windows 11 21H2:** December 12, 2023—KB5033369 (OS Build 22000.2652) - Microsoft Support
- **Windows 10 22H2:** November 30, 2023—KB5032278 (OS Build 19045.3758) Preview - Microsoft Support
- **Windows 10 21H2:** November 30, 2023—KB5032278 (OS Build 19045.3758) Preview - Microsoft Support
- **Windows Server 2022/2019:** November 14, 2023—KB5032198 (OS Build 20348.2113) - Microsoft Support (or later)

Operating system requirements for Condition 'Document Property is'

- **Windows 11:** February 29, 2024—KB5034848 (OS Builds 22621.3235 and 22631.3235) Preview - Microsoft Support (or later)
- **Windows 10:** February 29, 2024—KB5034843 (OS Build 19045.4123) Preview - Microsoft Support (or later)

> **Important**
>
> For information about the Adobe requirements for using Microsoft Purview Data Loss Prevention (DLP) features with PDF files, see this article from Adobe: **Microsoft Purview Information Protection Support in Acrobat**.

Conditions Instances supports

- Content contains
- Content is shared from Microsoft 365

Conditions On-premises repositories support

- Content contains
- File extension is
- Document property is

Conditions Fabric and Power BI support

- Content contains

Conditions Microsoft 365 Copilot supports

This feature is in preview.

- Content contains (sensitivity labels)
- Content contains (sensitive information types)
- Email is received from (External users) (preview)

Condition groups

Sometimes you need a rule to identify only one thing, such as all content that contains a U.S. Social Security Number, which is defined by a single SIT. However, in many scenarios where the types of items you're trying to identify are more complex and therefore harder to define, more flexibility in defining conditions is required.

For example, to identify content subject to the U.S. Health Insurance Act (HIPAA), you need to look for:

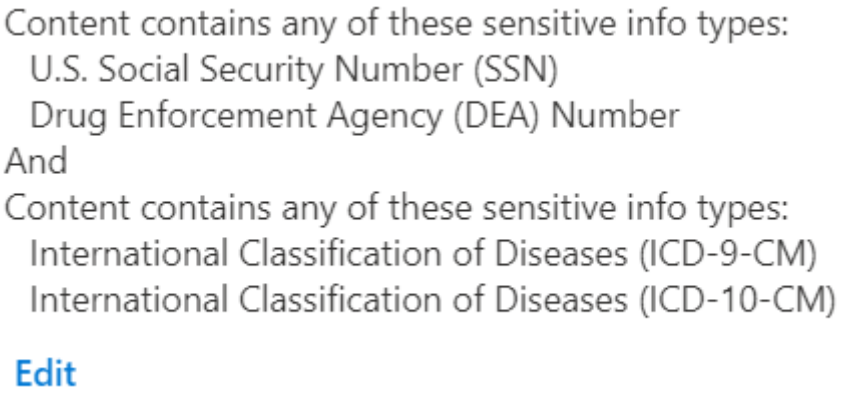
- Content that contains specific types of sensitive information, such as a U.S. Social Security Number or Drug Enforcement Agency (DEA) Number.

  AND

- Content that's more difficult to identify, such as communications about a patient's care or descriptions of medical services provided. Identifying this content requires matching keywords from large keyword lists, such as the International Classification of Diseases (ICD-9-CM or ICD-10-CM).

You can identify this type of data by grouping conditions and using logical operators (AND, OR) between the groups.

For the **U.S. Health Insurance Act (HIPAA)**, conditions are grouped like this:



The first group contains the SITs that identify an individual and the second group contains the SITs that identify medical diagnosis.

Conditions can be grouped and joined by boolean operators (AND, OR, NOT) so that you define a rule by stating what should be included and then defining exclusions in a different group joined to the first by a NOT. To learn more about how Purview DLP implements booleans and nested groups see, Complex rule design.

DLP platform limitations for conditions

Condition	Workload	Limit	Cost of Evaluation
Content Contains	EXO/SPO/ODB	125 SITs per rule	High
Content is shared from Microsoft 365	EXO/SPO/ODB	-	High
Sender IP address is	EXO	Individual range length <= 128; Count <= 600	Low
Has sender overridden the policy tip	EXO	-	Low
Sender is	EXO	Individual email length <= 256; Count <= 600	Medium
Sender is a member of	EXO	Count <= 600	High
Sender domain is	EXO	Domain name length <= 67; Count <= 600	Low
Sender address contains words	EXO	Individual word length <= 128; Count <= 600	Low
Sender address matches patterns	EXO	Regex length <= 128 char; Count <= 600	Low

Condition	Workload	Limit	Cost of Evaluation
Sender AD attribute contains words	EXO	Individual word length <= 128; Count <= 600	Medium
Sender AD attribute matches patterns	EXO	Regex length <= 128 char; Count <= 600	Medium
Content of email attachment(s) can't be scanned	EXO	Supported file types	Low
Incomplete scan of email attachment content	EXO	Extracted content size > 2 MB (2 million characters)	Low
Attachment is password-protected	EXO	File types: Office files, .PDF, .ZIP, and 7z	Low
Attachment's file extension is	EXO/SPO/ODB	Count <= 600 per rule	High
Recipient is a member of	EXO	Count <= 600	High
Recipient domain is	EXO	Domain name length <= 67; Count <= 5000	Low
Recipient is	EXO	Individual email length <= 256; Count <= 600	Low
Recipient address contains words	EXO	Individual word length <= 128; Count <= 600	Low
Recipient address matches patterns	EXO	Count <= 300	Low
Document name contains words or phrases	EXO	Individual word length <= 128; Count <=600	Low
Document Name matches patterns	EXO	Regex length <= 128 char; Count <= 300	Low
Document property is	EXO/SPO/ODB	-	Low
Document size equals or is greater than	EXO	-	Low
Subject contains words or phrases	EXO	Individual word length <= 128; Count <= 600	Low
Header contains words or phrases	EXO	Individual word length <= 128; Count <= 600	Low

Condition	Workload	Limit	Cost of Evaluation
Subject or body contains words or phrases	EXO	Individual word length <= 128; Count <= 600	Low
Content character set contains words	EXO	Count <= 600	Low
Header matches patterns	EXO	Regex length <= 128 char; Count <= 300	Low
Subject matches patterns	EXO	Regex length <= 128 char; Count <= 300	Low
Subject or body matches patterns	EXO	Regex length <= 128 char; Count <= 300	Low
Message type is	EXO	-	Low
Message size over	EXO	-	Low
With importance	EXO	-	Low
Sender AD attribute contains words	EXO	Each attribute key value pair: has Regex length <= 128 char; Count <= 600	Medium
Sender AD attribute matches patterns	EXO	Each attribute key value pair: has Regex length <= 128 char; Count <= 300	Medium
Document contains words	EXO	Individual word length <= 128; Count <= 600	Medium
Document matches patterns	EXO	Regex length <= 128 char; Count <= 300	Medium

Actions

Any item that makes it through the ***conditions*** filter has any ***actions*** that are defined in the rule applied to it. You have to configure the required options to support the action. For example, if you select Exchange with the **Restrict access or encrypt the content in Microsoft 365 locations** action, you need to choose from these options:

- Block users from accessing shared SharePoint, OneDrive, and Teams content
  - Block everyone. Only the content owner, last modifier, and site admin will continue to have access
  - Block only people from outside your organization. Users inside your organization continue to have access.

- Encrypt email messages (applies only to content in Exchange)

The actions that are available in a rule depend on the locations that have been selected. The available actions for each individual location are listed below.

> **Important**
>
> For SharePoint and OneDrive locations, documents will be proactively blocked right after detection of sensitive information (regardless of whether the document is shared or not) for all guests; internal users continue to have access to the document.

Supported actions: Exchange

When DLP policy rules are applied in Exchange, they may be **halting**, **non-halting**, or **neither**. Most of the rules that Exchange supports are non halting. Non halting actions are applied after processing the subsequent rules and policies.

DLP actions are taken on inbound encrypted emails that are in scope of a policy, such as block, *but* to maintain the confidentiality of the encryption, the event won't appear in Activity Explorer or in the Alert and the content of the message won't be accessible to anyone other than the recipient.

However, when a *halting* action is triggered by a DLP policy rule, Purview stops processing any subsequent rules. For instance, when the *Restrict access or encrypt the content in Microsoft 365 locations* action is triggered, no further rules or policies are processed.

If an action is *neither* halting nor non halting, Purview waits for the result of the action to occur before continuing. So, when an outgoing email triggers the *Forward the message for approval to sender's manager* action, Purview waits to get the manager's decision on whether or not the email may be sent. If the manager approves, the action behaves as a non halting action and the subsequent rules are processed. In contrast, if the manager rejects sending the email, *Forward the message for approval to sender's manager* behaves as a halting action and blocks sending the email; no subsequent rules or policies are processed.

The following table lists the actions that Exchange supports, and indicates whether they're halting or non halting.

Action	Halting / Non halting
Restrict access or encrypt the content in Microsoft 365 locations (Block Everyone, Block only people outside your organization)	Halting
Restrict access or encrypt the content in Microsoft 365 locations (Encrypt Email Messages)	Non - Halting
Set headers	Non halting
Remove header	Non halting
Redirect the message to specific users	Non halting
Forward the message for approval to sender's manager	Neither
Forward the message for approval to specific approvers	Neither
Add recipient to the To box	Non halting
Add recipient to the Cc box	Non halting
Add recipient to the Bcc box	Non halting
Add the sender's manager as recipient	Non halting
Remove message encryption and rights protection	Non halting
Prepend Email Subject	Non halting
Add HTML Disclaimer	Non halting
Modify Email Subject	Non halting
Deliver the message to the hosted quarantine	Halting
Apply branding to encrypted messages	Non halting

Tip

For the **Apply branding to encrypted messages** action, if you already have Microsoft Purview Message Encryption implemented, the templates automatically show up in the drop-down list. If you want to implement Microsoft Purview Message Encryption, see **Add your organization's brand to your Microsoft Purview Message Encryption encrypted messages** for background on message encryption and how to create and configure your branding templates.

For more information on the actions that Exchange supports, including PowerShell values, see: Data loss prevention Exchange conditions and actions reference.

Supported actions: SharePoint

- Restrict access or encrypt the content in Microsoft 365 locations
  - Block everyone
  - Block only people from outside your organization
  - Block access for specific external domains or users (in public preview)

The **Block access for specific external domains or users** sub-option lets you block external access by domain (for example, `partner.com`) or by user SMTP (for example, `user@example.com`). You can also specify allow lists by using **Domain IS NOT** or **User IS NOT**. Internal users and domains can't be blocked with this sub-option; continue to use **Block everyone** for internal users.

> **Note**
>
> When you use **Block access for specific external domains or users**: if a user or domain appears in both allow and block lists, the block takes effect (most restrictive wins). If a file matches both an allow rule and a block rule, evaluation is across all matching rules — allowed users and domains are permitted, blocked users and domains are denied, and users in neither list are blocked by default.

Supported actions: OneDrive

- Restrict access or encrypt the content in Microsoft 365 locations
  - Block everyone
  - Block only people from outside your organization
  - Block access for specific external domains or users (in public preview)

The **Block access for specific external domains or users** sub-option lets you block external access by domain (for example, `partner.com`) or by user SMTP (for example, `user@example.com`). You can also specify allow lists by using **Domain IS NOT** or **User IS NOT**. Internal users and domains can't be blocked with this sub-option; continue to use **Block everyone** for internal users.

> **Note**
>
> When you use **Block access for specific external domains or users**: if a user or domain appears in both allow and block lists, the block takes effect (most restrictive wins). If a file

matches both an allow rule and a block rule, evaluation is per rule — allowed users and domains are permitted, blocked users and domains are denied, and users in neither list are blocked by default.

Supported actions: Teams Chat and Channel Messages

- Restrict access or encrypt the content in Microsoft 365 locations

Supported actions: Devices

The **Devices** location supports these actions:

- Restrict access or encrypt the content in Microsoft 365 locations
  - Block users from receiving email, or accessing shared SharePoint, OneDrive, Teams files, and Power BI items
- Audit or restrict activities when users access sensitive sites in Microsoft Edge browsers on Windows devices
  - Sensitive site restrictions
- Audit or restrict activities on devices
  - Upload to a restricted cloud service domain or access from an unallowed browser
  - Paste to supported browsers
- File activities for all apps
  - Copy to clipboard
  - Copy to removable USB device
  - Copy to a network share
  - Print
  - Copy or move using unallowed Bluetooth app
  - Copy or move using RDP
  - File activities for apps in restricted app groups
  - App access restrictions
    - Access by restricted apps
    - Access by apps not included in restricted apps list or any restricted app groups added to rule
    - *in preview* Restrictions in Windows Recall in Copilot+ PCs
    - Apply restrictions to only unsupported file extensions Start a Power Automate flow

Important

When you select the **Audit or restrict activities on devices** action, the **Apply restrictions to only unsupported file extensions** shows up. **Apply restrictions to only unsupported file extensions** configuration option ***DOES NOT*** support scoping by **Device and device groups** in the policy location setting.

You can tell DLP to **Allow**, **Audit only**, **Block with override**, or **Block** (the actions) these user activities for onboarded Windows devices.

You can tell DLP to **Audit only**, **Block with override**, or **Block** (the actions) these user activities for onboarded macOS devices.

- **Block**: User related activity is blocked, and auditing is enabled. Admins may optionally see alerts.
- **Block with override**: This option acts as a standard block but permits users to bypass it. By clicking the 'Allow' button on the toast notification or the 'Ok' button on the Microsoft Edge notification, users can proceed. Once allowed, Endpoint DLP will automatically resume for actions including 'Copy to a network share', 'Copy to a removable USB device', and 'Print'. For other actions, users will need to repeat the process after clicking 'Allow' to bypass the policy.
- **Audit**: No blocking of activities, but auditing is enabled, and admins may optionally see alerts.
- **Allow**: Activities are allowed without triggering alerts, but auditing is still enabled.
- **Off**: No blocking or auditing of activities.

Enforcement mode	Block user	Alert generation	Auditing record generation
Block	Yes	Yes if alert is turned on for the DLP rule	Yes
Block with override	Yes	Yes if alert is turned on for the DLP rule	Yes
Audit	No	Yes if alert is turned on for the DLP rule	Yes
Allow	No	Never triggered	Yes
Off	No	No	No

Supported actions: Managed cloud apps

You can tell DLP to **Audit only** or **Block** (the actions) for user activities in managed cloud apps on Windows and macOS devices.

Supported actions: Unmanaged cloud apps

You can tell DLP to **Audit only** or **Block** (the actions) for user activities in unmanaged cloud apps on Windows.

More information on supported actions

You can find more details here about actions:

- Restrict access or encrypt content in Microsoft 365 locations
- Audit or restrict activities when users access sensitive sites in Microsoft Edge browsers on Windows devices
- Audit or restrict activities on devices
- Service domain and browser activities
- File activities for all apps
- Restricted app activities
- File activities for apps in restricted app groups

Restrict access or encrypt content in Microsoft 365 locations

Use this to block users from receiving email, or accessing shared SharePoint, OneDrive, Teams files, and Power BI items. This action can block everyone or block only people who are outside your organization. For SharePoint and OneDrive locations only, you can also block access for specific external domains or user SMTPs (in public preview).

Audit or restrict activities when users access sensitive sites in Microsoft Edge browsers on Windows devices

Use this action to control when users attempt to:

Activity	Description/options
Print the site	Detects when users try to print a protected site from an onboarded device.
Copy data from the site	Detects when users try to copy data from a protected site from an onboarded device.
Save the site as local files (Save-As)	Detects when users try to save a protected site as local files from an onboarded device.

Audit or restrict activities on devices

Use this to restrict user activities by Service domain and browser activities, File activities for all apps, Restricted app activities. To use **Audit or restrict activities on devices**, you have to configure options in **DLP settings** and in the policy in which you want to use them. See, Restricted apps and app groups for more information.

DLP rules with the action **Audit or restrict activities on devices** can have **Block with override** configured. When this rule is applied to a file, any attempt to perform a restricted action on the file is blocked. A notification is displayed with the option to override the restriction. If the user chooses to override, the action is permitted for a period of 1 minute, during which the user can retry the action without restriction. The exception to this behavior is when a sensitive file is dragged and dropped into Microsoft Edge, which will immediately attach the file if the rule is overridden.

Service domain and browser activities

When you configure the **Allow/Block cloud service domains** and the **Unallowed browsers** list (see Browser and domain restrictions to sensitive data) and a user attempts to upload a protected file to a cloud service domain or access it from an unallowed browser, you can configure the policy action to `Audit only`, `Block with override`, or `Block` the activity.

Activity	Description/options
Upload to a restricted cloud services domain or access from an unallowed app	Detects when protected files are blocked or allowed to be uploaded to cloud service domains. See, Browser and domain restrictions to sensitive data and Help prevent risky user activity by monitoring or restricting access to sensitive service domains.
Paste to supported browsers	Detects when users paste sensitive information into a text field or web form using Microsoft Edge, Google Chrome (with Microsoft Purview extension), or Mozilla Firefox (with Microsoft Purview extension). Evaluation is independent of the classification of the source file. For more information, see: Endpoint activities you can monitor and take action on.

Paste to Browser limitations

Only certain rule conditions work with Paste to Browser events due to the fact that the rules are being evaluated only on the clipboard data. Paste to Browser won't evaluate based on where the text is being copied from.

Rule conditions that work with Paste to Browser:

- Content contains
- Content is not labeled

Additional notes:

- Paste to Browser supports SITs, not Sensitivity Labels.
- Paste to Browser doesn't evaluate on text smaller than 30 characters.
- Advanced Classification isn't supported.
- Contextual Summary doesn't show for Paste to Browser events.
- Paste to Browser takes 2 seconds to evaluate before allowing the paste action.
- IF JIT is configured to block on fallback, it blocks pasting.
- Paste to Browser only classifies the first 4 MB of text from the clipboard

File activities for all apps

With the **File activities for all apps** option, you select either **Don't restrict file activities** or **Apply restrictions to specific activities**. When you select **Apply restrictions to specific activities**, the actions that you select here are applied when a user has accessed a DLP protected item.

Activity	Description/options
Copy to clipboard	Detects when protected files are copied to the clipboard on an onboarded device. For more information, see Endpoint activities you can monitor and take action on and Copy to clipboard behavior
Copy to a removable device	Detects when protected files are copied or moved from an onboarded device to a removable USB device. For more information, see Removable USB device groups.
Copy to a network share	Detects when protected files are copied or moved from an onboarded device to any network share. For more information, see Endpoint activities you can monitor and take action on and Network share coverage and exclusions.
Print	Detects when a protected file is printed from an onboarded device. For more information, see Printer groups.
Copy or move using unallowed Bluetooth app	Detects when a protected file is copied or moved from an onboarded Windows device using an unallowed Bluetooth app. For more information, see Unallowed Bluetooth apps. This isn't supported for macOS.
Copy or move using RDP	Detects when users copy or move protected files from an onboarded Windows device to another location using RDP. This isn't supported for macOS.

Restricted app activities

Previously called Unallowed apps, *restricted app activities* are apps that you want to place restrictions on. You define these apps in a list in Endpoint DLP settings. When a user attempts to access a DLP protected file using an app that is on the list, you can either `Audit only`, `Block with override`, or `Block` the activity. DLP actions defined in **Restricted app activities** are overridden if

the app is a member of restricted app group. Then the actions defined in the restricted app group are applied.

Activity	Description/options
Access by restricted apps	Detects when unallowed apps try to access protected files on an onboarded Windows device. For more information, see Restricted apps and app groups.

File activities for apps in restricted app groups

You define your restricted app groups in Endpoint DLP settings and add restricted app groups to your policies. When you add a restricted app group to a policy, you must select one of these options:

- Don't restrict file activity
- Apply restrictions to all activity
- Apply restrictions to specific activity

When you select either of the *Apply restrictions* options, and a user attempts to access a DLP protected file using an app that is in the restricted app group, you can either `Audit only`, `Block with override`, or `Block` by activity. DLP actions that you define here override actions defined in **Restricted app activities** and **File activities for all apps** for the app.

For more information, see Restricted apps and app groups.

> **Note**
>
> The devices location provides many subactivities (conditions) and actions. To learn more, see Endpoint activities you can monitor and take action on.

> **Important**
>
> The **Copy to clipboard** condition detects when a user copies information from a *protected file* to the clipboard. Use **Copy to clipboard** to block, block with override, or audit when users copy information from a protected file.
>
> The **Paste to supported browsers** condition detects when a user attempts to paste sensitive text into a text field or web form using Microsoft Edge, Google Chrome with Microsoft Purview extension, or Mozilla Firefox with Microsoft Purview extension *regardless of where*

that information came from. Use **Paste to supported browsers** to block, block with override, or audit when users paste sensitive information into a text field or web form.

Instances actions

- Restrict access or encrypt the content in Microsoft 365 locations
- Restrict Third Party Apps

On-premises repositories actions

- Restrict access or remove on-premises files.
  - Block people from accessing files stored in on-premises repositories
  - Set permissions on the file (permissions inherited from the parent folder)
  - Move file from where it's stored to a quarantine folder

See, DLP On-premises repository actions for full details.

Fabric and Power BI actions

- Notify users with email and policy tips
- Send alerts to Administrator
- Restrict access

> **Note**
>
> Applicable to supported item types only.

Microsoft 365 Copilot actions

This feature is in preview.

- Exclude content in the location

Managed cloud apps actions

- Restrict browser and network activities

Unmanaged cloud apps actions

- Restrict browser and network activities

Actions available when you combine locations

If you select Exchange and any other single location for the policy to be applied to, the

- Restrict access or encrypt the content in Microsoft 365 locations and all actions for the non-Exchange location actions are available.

If you select two or more non-Exchange locations for the policy to be applied to, the

- Restrict access or encrypt the content in Microsoft 365 locations and all actions for non-Exchange locations actions are available.

For example, if you select the Exchange and Devices locations, these actions are available:

- Restrict access or encrypt the content in Microsoft 365 locations
- Audit or restrict activities on Windows devices

If you select Devices and Instances, these actions are available:

- Restrict access or encrypt the content in Microsoft 365 locations
- Audit or restrict activities on Windows devices
- Restrict Third Party Apps

Whether an action takes effect or not depends on how you configure the mode of the policy. You can choose to run the policy in simulation mode with or without showing policy tip by selecting the **Run the policy in simulation mode** option. You choose to run the policy as soon as an hour after it's created by selecting the **Turn it on right away** option, or you can choose to just save it and come back to it later by selecting the **Keep it off** option.

DLP platform limitations for actions

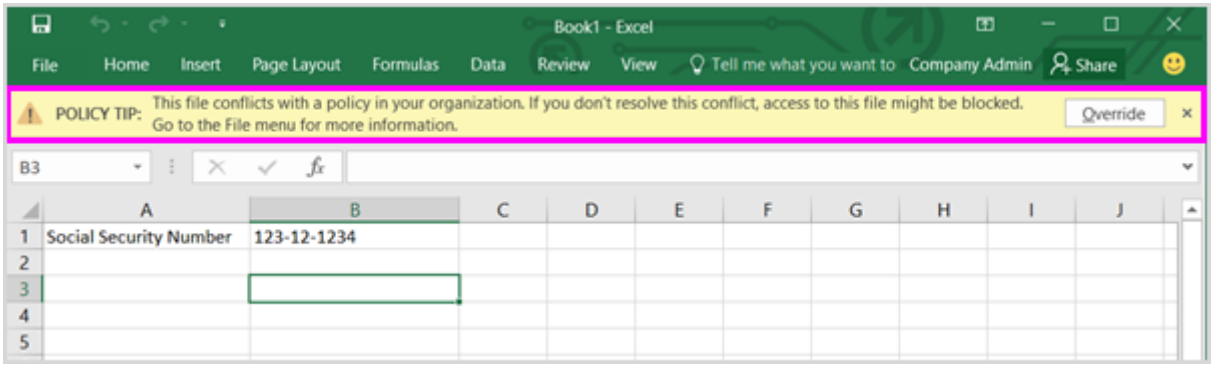
Action Name	Workload	Limits
Restrict access or encrypt content in Microsoft 365	EXO/SPO/ODB	**Block access for specific external domains or users** sub-option: image files aren't protected; multiple audit records can be generated for certain actions on a blocked file by a blocked user (public preview, SPO/ODB only).
Set headers	EXO	
Remove header	EXO	
Redirect the message to specific users	EXO	Total of 100 across all DLP rules. Can't be DL/SG

Action Name	Workload	Limits
Forward the message for approval to sender's manager	EXO	Manager should be defined in AD
Forward the message for approval to specific approvers	EXO	Groups aren't supported
Add recipient to the **To** box	EXO	Recipient count <= 10; Can't be DL/SG
Add recipient to the **Cc** box	EXO	Recipient count <= 10; Can't be DL/SG
Add recipient to the **Bcc** box	EXO	Recipient count <= 10; Can't be DL/SG
Add the sender's manager as recipient	EXO	Manager attribute should be defined in AD
Apply HTML disclaimer	EXO	
Prepend subject	EXO	
Apply message encryption	EXO	
Remove message encryption	EXO	
(preview) Exclude content in Copilot location	Microsoft 365 Copilot (preview)	Only content in SharePoint and OneDrive for Business can be excluded from being processed by Microsoft 365 Copilot

User notifications and policy tips

When a user attempts an activity on a sensitive item in a context that meets the conditions of a rule (for example, content such as an Excel workbook on a OneDrive site that contains personal data (PII) and is shared with a guest), you can let them know about it through user notification emails and in-context policy tip popups. These notifications are useful because they increase awareness and help educate people about your organization's DLP policies.

An Alert email, Incident Report email, and User Notification will only be sent once per document. If a document with a 'Content is Shared' condition is shared twice, there will still be only one notification.



> **Important**
>
> - Notification emails are sent unprotected.
> - Email notifications are only supported for the Microsoft 365 services.

Email notifications support by selected location

Selected location	Email notifications supported
Devices	- Not supported
Exchange + Devices	- Supported for Exchange - Not supported for Devices
Exchange	- Supported
SharePoint + Devices	- Supported for SharePoint - Not supported for Devices
SharePoint	- Supported
Exchange + SharePoint	- Supported for Exchange - Supported for SharePoint
Devices + SharePoint + Exchange	- Not supported for Devices - Supported for SharePoint Supported for Exchange
Teams	- Not supported
OneDrive	- Supported for OneDrive for work or school - Not supported for Devices
Fabric and Power-BI	- Not supported
Instances	- Not supported
On-premises repositories	- Not supported

Selected location	Email notifications supported
Exchange + SharePoint + OneDrive	- Supported for Exchange - Supported for SharePoint - Supported for OneDrive
Microsoft 365 Copilot (preview)	- Not supported
Managed cloud apps	- Not supported
Unmanaged cloud apps	- (Preview) Supported for email notifications. See Email notifications for browser and network (preview).

Email notifications for browser and network (preview)

When a DLP rule is configured for unmanaged cloud apps (browser and network), you can enable email notifications to notify end users via email when their activity is blocked. This is different from the inline pop-up notifications that appear directly in the browser, such as in Microsoft Edge for Business. Email notifications are particularly useful for integrations that don't support pop-up notifications, such as the Android Global Secure Access integration.

You can configure email notifications for:

- The person who performed the blocked activity.
- Additional recipients, such as admins or compliance officers.

Email notification batching

Email notifications are batched to prevent a flood of communications:

- The **first** policy match sends an email notification immediately (within seconds of the activity).
- Any additional policy matches within a rolling **10-minute window** are batched into a single email notification that lists all blocked activities.
- After 10 minutes with no further policy matches, the window resets. The next policy match sends an immediate notification and starts a new 10-minute window.

The batched email lists all blocked activities as individual entries, regardless of whether they were triggered by the same or different DLP rules for the same user.

You can also give people the option to override the policy, so that they're not blocked if they have a valid business need or if the policy is detecting a false positive.

Policy tips and user notifications configuration options

The user notifications and policy tips configuration options vary depending on the monitoring locations you've selected. If you selected:

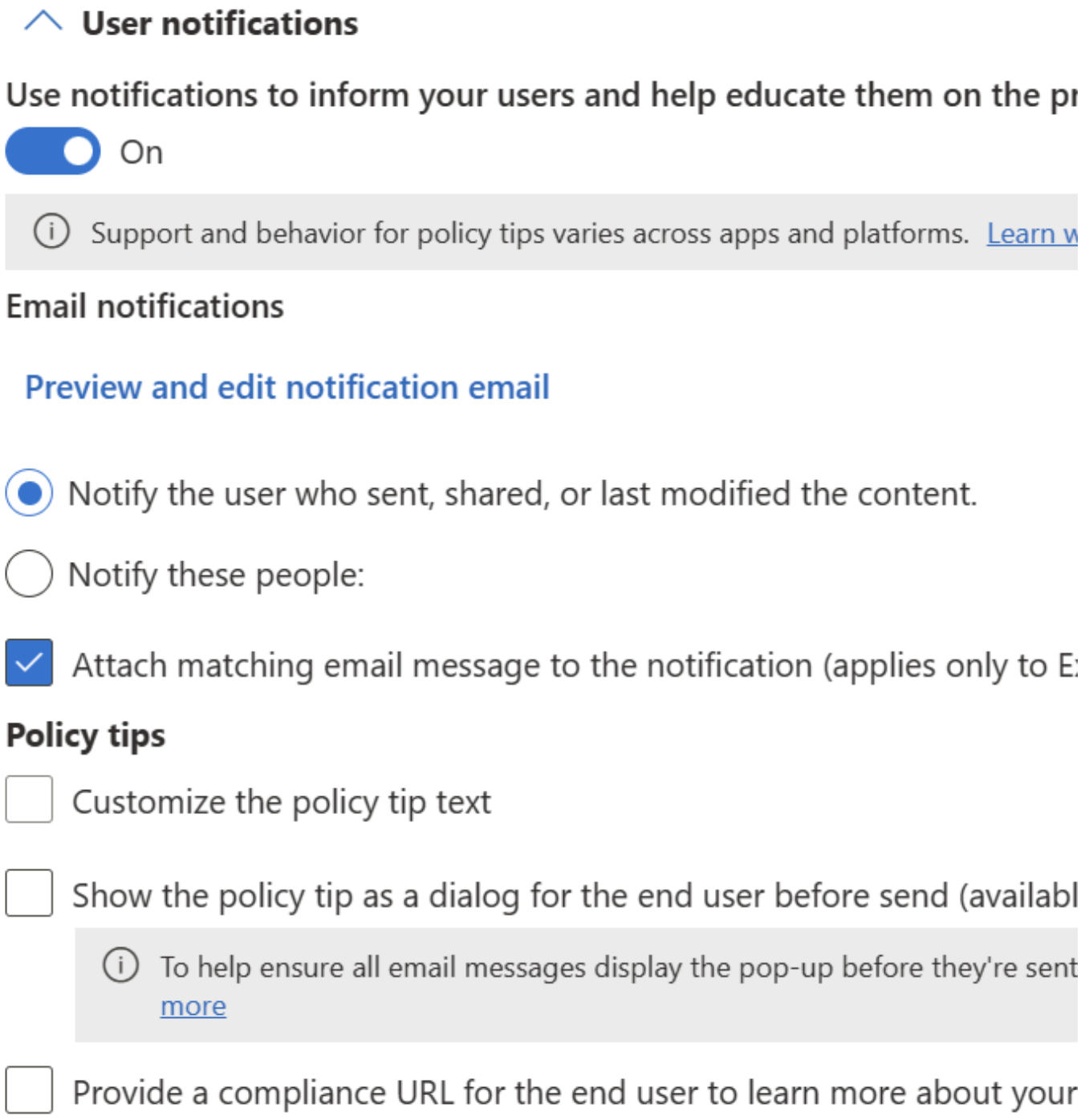
- Exchange
- SharePoint
- OneDrive
- Teams Chat and Channel
- Instances

You can enable/disable user notifications for various Microsoft apps, see Data Loss Prevention policy tips reference.

You can also enable/disable notifications with a policy tip.

- email notifications to the user who sent, shared, or last modified the content OR
- notify specific people

Furthermore, you can customize the email text, subject, and the policy tip text.



For detailed information on customizing end user notification emails, see Custom email notifications.

If you selected Devices only, you get all the same options that are available for Exchange, SharePoint, OneDrive, Teams Chat and Channel, and Instances, plus the option to customize the notification title, content, and add a hyperlink in the form of a *Get Support* button that appears on the Windows 10/11 device.

Custom policy tip notifications character limits

Policy tip notifications are subject to the following character limits:

Variable	Character Limit
`TITLE`	120
`CONTENT`	250
`JUSTIFICATION`	250

Hyperlink has no character limit but is limited to the remaining space available in the entire DLP package. The hyperlink must be a resolvable URL, and it's abstracted behind a selectable control. The more information hyperlink is available in Microsoft 365 Office apps.

You can customize the title and body of text using the following parameters.

Common name	Parameter	Example
file name	%%FileName%%	Contoso doc 1
process name	%%ProcessName%%	Word
policy name	%%PolicyName%%	Contoso highly confidential
action	%%AppliedActions%%	pasting document content from the clipboard to another app

%%AppliedActions%% substitutes these values into the message body:

action common name	value substituted in for %%AppliedActions%% parameter
copy to removeable storage	*writing to removable storage*
copy to network share	*writing to a network share*
print	*printing*
paste from clipboard	*pasting from the clipboard*
copy via bluetooth	*transferring via Bluetooth*
open with an unallowed app	*opening with this app*
copy to a remote desktop (RDP)	*transferring to remote desktop*
uploading to an unallowed website	*uploading to this site*
accessing the item via an unallowed browser	*opening with this browser*

Using this customized text

%%AppliedActions%% File name %%FileName%% via %%ProcessName%% isn't allowed by your organization. Select 'Allow' if you want to bypass the policy %%PolicyName%%

produces this text in the customized notification:

pasting from the clipboard File Name: Contoso doc 1 via WINWORD.EXE isn't allowed by your organization. Select the 'Allow' button if you want to bypass the policy Contoso highly confidential

You can localize your custom policy tips by using the Set-DlpComplianceRule -NotifyPolicyTipCustomTextTranslations cmdlet.

Custom Policy Tips show for the most restrictive rule, not necessarily the rule that's performing the restriction.

> ⓘ **Note**
>
> User notifications and policy tips aren't available for the On-premises location
>
> Only the policy tip from the highest priority, most restrictive rule is shown. For example, a policy tip from a rule that blocks access to content will be shown over a policy tip from a rule that simply sends a notification. This prevents people from seeing a cascade of policy tips.

To learn more about user notification and policy tip configuration and use, including how to customize the notification and tip text, see Send email notifications and show policy tips for DLP policies.

Policy tip considerations

- Policy tips aren't generated if the sensitivity label is in a compressed file.
- DLP has difficulty generating policy tips for encrypted files.

Policy tip references

Details on support for policy tips and notifications for different apps can be found here:

- Data loss prevention policy tip reference for Outlook for Microsoft 365
- Data loss prevention policy tip reference for Outlook on the Web
- Data loss prevention policy tip reference SharePoint in Microsoft 365 and OneDrive. web client

Block users in Exchange

> ⓘ **Note**

If you have policy tips enabled when the policy is configured to **Block only users outside your organization** is met, the policy tips notification blocks you from sending a message at all if there are external recipients. Therefore, you have to remove any external recipients before you can send the message to internal recipients.

Blocking and notifications in SharePoint in Microsoft 365 and OneDrive

The following table shows the DLP blocking and notification behavior for policies that are scoped to SharePoint in Microsoft 365 and OneDrive. This isn't intended to be an exhaustive list, and there are more settings that aren't in scope for this article.

Conditions	Restrict access setting	Blocking and notification behavior
Content is shared from Microsoft 365- with people outside my organization	Not configured	User notification emails, policy tips, DLP alerts, and incident reports are sent when a file is shared with a guest or when a guest accesses the file.
Content is shared from Microsoft 365- only with people inside my organization	Not configured	User notification emails, policy tips, DLP alerts, and incident reports are sent when sensitive information is detected in the document.
Content is shared from Microsoft 365- only with people inside my organization	- **Restrict access or encrypt the content in Microsoft 365 locations** - **Block users from receiving email or accessing shared SharePoint, OneDrive, and Teams files** -**Block everyone**	- Access to sensitive files is blocked as soon as they're uploaded. - User notification emails, policy tips, DLP alerts, and incident reports are sent when a file is uploaded
Content is shared from Microsoft 365 - with people outside my organization	- **Restrict access or encrypt the content in Microsoft 365 locations** - **Block users from receiving email or accessing shared SharePoint, OneDrive, and Teams files**	- Access to a sensitive file is blocked as soon as it's uploaded, regardless of whether the document is shared or not for all guests. Policy Tips will be sent when the sensitive information is detected. - If the sensitive information is added to a file after it's shared and accessed by a user outside the organization, then alerts and incident reports will be sent. - If the document contains sensitive information before it's uploaded, external sharing will be blocked proactively. Because external sharing in this scenario is blocked when the

Conditions	Restrict access setting	Blocking and notification behavior
	- **Block only people outside your organization**	file is uploaded, no alerts or incident reports are sent. Suppression of the alerts and incident reports is designed to prevent a flood of alerts to the user for each blocked file. User notification emails, policy tips, DLP alerts, and incident reports are sent when a file is uploaded or sensitive information is added to the file. - Proactive blocking will show up as event in the Audit Log and Activity Explorer.
- **Content is shared from Microsoft 365** - **with people outside my organization**	- **Restrict access or encrypt the content in Microsoft 365 locations** - **Block users from receiving email or accessing shared SharePoint, OneDrive, and Teams files** - **Block everyone**	- When the first user outside the organization access the document, the event causes the document to be blocked. - It's expected that for a short time, the document is accessible by guests who have the link to the file. - User notification emails, policy tips, DLP alerts, and incident reports are sent when a file is shared with a guest or when a guest accesses that file.
Content is shared from Microsoft 365 - with people outside my organization	- **Restrict access or encrypt the content in Microsoft 365 locations** - **Block users from receiving email or accessing shared SharePoint, OneDrive, and Teams files** - **Block access for specific external domains or users** (in public preview, SPO/ODB only)	- Access is denied when a configured external user or domain attempts to open, preview, or download the file. - An audit record is generated per blocked access attempt (subject to preview limitations). - Events can be reviewed in DLP Alerts, Activity Explorer, and Audit logs. The event contains the blocked user's email. - Internal users and domains can't be blocked with this sub-option; use **Block everyone** to block internal users.

Learn more URL

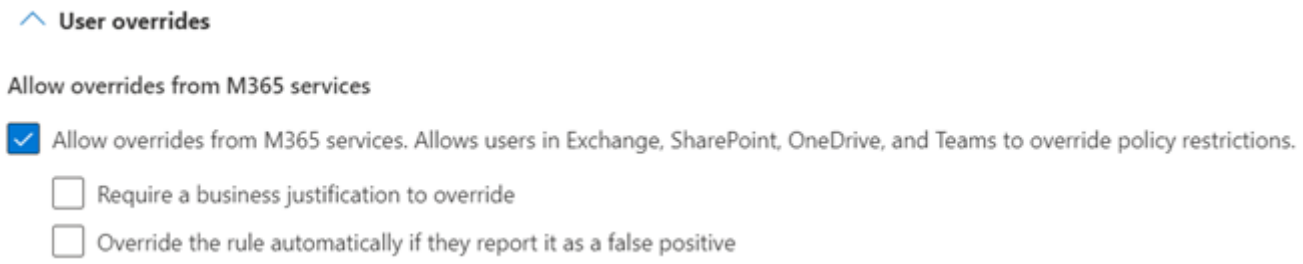
Users may want to learn why their activity is being blocked. You can configure a site or a page that explains more about your policies. In the rule's user notification settings, select **Provide a compliance URL for the end user to learn more about your organization's policies**, and then enter the URL.

- For Outlook Win32, the *Learn more* link in the policy tip points to the site URL that you provide. This URL has priority over the global compliance URL configured with Set-PolicyConfig -ComplianceURL.
- For Microsoft Copilot and Copilot Chat, the **Learn about access restrictions** link in the standard block message points to the URL that you provide. Only the link destination changes; the message text and enforcement action remain unchanged. For more information, see Customize the link in Copilot policy tips.

> **Important**
>
> You must configure the site or page that the link points to from scratch. Microsoft Purview doesn't provide this functionality out of the box.

User overrides

The intent of **User overrides** is to give users a way to bypass, with justification, DLP policy blocking actions on sensitive items in Exchange, SharePoint, OneDrive, or Teams, so that they can continue their work. User overrides are enabled only when **Notify users in Office 365 services with a policy tip** is enabled, so user overrides go hand-in-hand with Notifications and Policy tips.



> **Note**
>
> User overrides aren't available for the On-premises repositories location.

Typically, user overrides are useful when your organization is first rolling out a policy. The feedback that you get from any override justifications and identifying false positives helps in tuning the policy.

- If the policy tips in the most restrictive rule allow people to override the rule, then overriding this rule also overrides any other rules that the content matched.

User override behavior

When a user selects the **Allow** option to override a block action for these activities:

- Print
- Copy to a removable USB device
- Copy to a network share

Within 30 seconds of the popup notification showing, the activity is allowed to continue. If the user doesn't select the **Allow** option within 30 seconds, the activity is blocked.

For all other activities, the user must repeat the activity after selecting the **Allow** option in order for it to complete

Business justification X-Header

When a user overrides a block with override action on an email, the override option and the text that they provide are stored in the Audit log and in the email X-header. To view the business justification overrides, search the audit log for `ExceptionInfo` value for the details. Here's an example of the audit log values:

XML

```
{
    "FalsePositive"; false,
    "Justification"; My manager approved sharing of this content",
    "Reason"; "Override",
    "Rules": [
         "<message guid>"
    ]
}
```

If you have an automated process that makes use of the business justification values, the process can access that information programmatically in the email X-header data.

> ⓘ Note
>
> The `msip_justification` values are stored in the following order:
>
> `False Positive; Recipient Entitled; Manager Approved; I Acknowledge; JustificationText_[free text]`.
>
> Notice that the values are separated by semicolons. The maximum free text allowed is 500 characters.

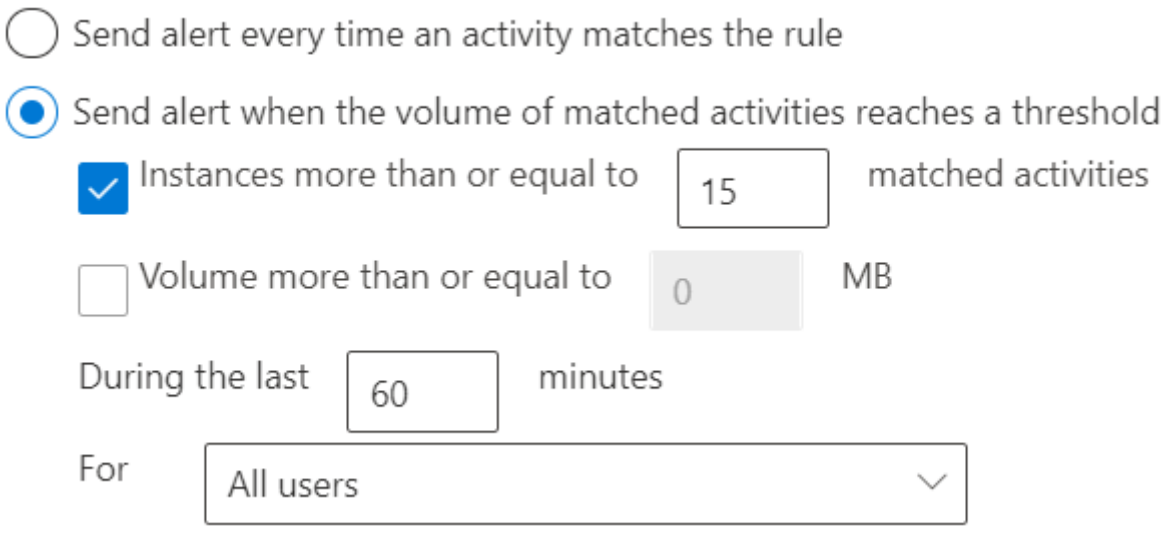
Incident reports

When a rule is matched, you can send an *Alert* email to your compliance officer (or any people you choose) with details of the event and you can view them in the Microsoft Purview Data Loss Prevention **Alerts** dashboard and in the Microsoft Defender XDR portal. An alert includes information about the item that was matched, the actual content that matched the rule, and the name of the person who last modified the content.

In preview, admin alert emails include details such as:

- The alert severity
- The time the alert occurred
- The activity.
- The sensitive data that were detected.
- The alias of the user whose activity triggered the alert.
- The policy that was matched.
- The alert ID
- The endpoint operation that was attempted if the **Devices** location is in the scope of the policy.
- The app that was being used.
- The device name, if the match occurred on an endpoint device.

DLP feeds incident information to other Microsoft Purview Information Protection services, like insider risk management. In order to get incident information to insider risk management, you must set the **Incident reports** severity level to **High**.



Alert types

Alerts can be sent every time an activity matches a rule, which can be noisy. To help cut down on the noise, they can be aggregated based on number of matches or volume of items over a set period of time. There are two types of alerts that can be configured in DLP policies.

- **Single-event alerts** are typically used in policies that monitor for highly sensitive events that occur in a low volume, like a single email with 10 or more customer credit card numbers being sent outside your organization. In preview user based alert aggregation (preview) modifies the behavior of single event alerts.
- **Aggregate-event alerts** are typically used in policies that monitor for events that occur in a higher volume over a period of time. For example, an aggregate alert can be triggered when 10 individual emails each with one customer credit card number is sent outside your org over 48 hours.

> **Note**
>
> For rules with Alerts configured on Sharepoint, or OneDrive workloads we only send one alert per file per rule. This is true, even if the same violation has been committed by multiple users.

Other alert options

When you select **Use email incident reports to notify you when a policy match occurs** you can choose to include:

- The name of the person who last modified the content.
- The types of sensitive content that matched the rule.
- The rule's severity level.
- The content that matched the rule, including the surrounding text.
- The item containing the content that matched the rule.

For more information on alerts, see:

- Alerts in DLP policies: Describes alerts in the context of a DLP policy.
- Get started with data loss prevention alerts: Covers the necessary licensing, permissions, and prerequisites for DLP alerts and alert reference details.
- Create and deploy data loss prevention policies: Includes guidance on alert configuration in the context of creating a DLP policy.

- Learn about investigating data loss prevention alerts: Covers the various methods for investigating of DLP alerts.
- Investigate data loss incidents with Microsoft Defender XDR: How to investigate DLP alerts in Microsoft Defender portal.

Evidence collection for file activities on devices

If you've enabled **Setup evidence collection for file activities on devices** and added Azure storage accounts, you can select **Collect original file as evidence for all selected file activities on Endpoint** and the Azure storage account you want to copy the items to. You must also choose the activities you want to copy items for. For example, if you select **Print** but not **Copy to a network share**, then only items that are printed from monitored devices will be copied to the Azure storage account.

Additional options

If you have multiple rules in a policy, you can use the **Additional options** to control further rule processing if there's a match to the rule you're editing as well as setting the priority for evaluation of the rule. This is only supported for Exchange and Teams locations

See also

- Learn about data loss prevention
- Plan for data loss prevention (DLP)
- Create and Deploy data loss prevention policies

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 10/05/2026

What the DLP policy templates include

07/18/2025

Microsoft Purview Data Loss Prevention (DLP) in the Microsoft Purview portal includes ready-to-use policy templates that address common compliance requirements, such as helping you to protect sensitive information subject to the U.S. Health Insurance Act (HIPAA), U.S. Gramm-Leach-Bliley Act (GLBA), or U.S. Patriot Act. This article lists all of the policy templates, what types of sensitive information they look for, and what the default conditions and actions are. This article doesn't include every detail of how each policy template is configured; instead, the article presents with you enough information to help you decide which template is the best starting point for your scenario. Remember, you can customize these policy templates to meet your specific requirements.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Australia Financial Data

Expand table

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Australia Financial: Scan content shared outside - low count	Content contains sensitive information: SWIFT Code - Min count 1, Max count 9 Australia Tax File Number - Min count 1, Max count 9 Australia Bank Account Number - Min count 1, Max count 9 Credit Card Number - Min count 1, Max count 9 Content is shared with:	Send a notification

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
		People outside my organization	
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Australia Financial: Scan content shared outside - high count	Content contains sensitive information: SWIFT Code - Min count 10, Max count 500 Australia Tax File Number - Min count 10, Max count 500 Australia Bank Account Number - Min count 10, Max count 500 Credit Card Number - Min count 10, Max count 500 Content is shared with: People outside my organization	Block access to content Send a notification Allow override Require business justification Send incident report

Australia Health Records Act (HRIP Act)

Expand table

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Australia HRIP: Scan content shared outside - low count	Content contains sensitive information: Australia Tax File Number - Min count 1, Max count 9 Australia Medical Account Number - Min count 1, Max count 9 Australia Physical Address - Min count 1, Max count 9 Content is shared with: People outside my organization	Send a notification

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Australia HRIP: Scan content shared outside - high count	Content contains sensitive information: Australia Tax File Number - Min count 10, Max count 500 Australia Medical Account Number - Min count 10, Max count 500 Australia Physical Address - Min count 10, Max count 500 Content is shared with: People outside my organization	Block access to content Send a notification Allow override Require business justification Send incident report

Australia Personally Identifiable Information (PII) Data

Expand table

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Australia PII: Scan content shared outside - low count	Content contains sensitive information: Australia Tax File Number - Min count 1, Max count 9 Australia Driver's License Number - Min count 1, Max count 9 Content is shared with: People outside my organization	Send a notification
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Australia PII: Scan content shared outside - high count	Content contains sensitive information: Australia Tax File Number - Min count 10, Max count 500 Australia Driver's	Block access to content Send a notification Allow override Require

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
		License Number - Min count 10, Max count 500 Content is shared with: People outside my organization	business justification Send incident report

Australia Privacy Act Enhanced

Expand table

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Australia Privacy: Scan content shared outside - low count	Content contains sensitive information: Australia Driver's License Number - Min count 1, Max count 9 Australia Passport Number - Min count 1, Max count 9 All Medical Terms and Conditions - Min count 1, Max count 9 Australia Physical Addresses - Min count 1, Max count 9 Content is shared with: People outside my organization	Send a notification
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Australia Privacy: Scan content shared outside - high count	Content contains sensitive information: Australia Driver's License Number - Min count 10, Max count 500 Australia Passport Number - Min count 10, Max count 500 All Medical Terms and Conditions - Min count	Block access to content Send a notification Allow override Require business justification Send incident report

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
		10, Max count 500 Australia Physical Addresses - Min count 10, Max count 500 Content is shared with: People outside my organization	

Canada Financial Data

Expand table

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Canada Financial Data: Scan content shared outside - low count	Content contains sensitive information: Credit Card Number - Min count 1, Max count 9 Canada Bank Account Number - Min count 1, Max count 9 Content is shared with: People outside my organization	Send a notification
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Canada Financial Data: Scan content shared outside - high count	Content contains sensitive information: Credit Card Number - Min count 10, Max count 500 Canada Bank Account Number - Min count 10, Max count 500 Content is shared with: People outside my organization	Block access to content Send a notification Allow override Require business justification Send incident report

Canada Health Information Act (HIA)

Expand table

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Canada HIA: Scan content shared outside - low count	Content contains sensitive information: Canada Passport Number - Min count 1, Max count 9 Canada Social Insurance Number - Min count 1, Max count 9 Canada Health Service Number - Min count 1, Max count 9 Canada Personal Health Identification Number (PHIN) - Min count 1, Max count 9 Content is shared with: People outside my organization	Send a notification
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Canada HIA: Scan content shared outside - high count	Content contains sensitive information: Canada Passport Number - Min count 10, Max count 500 Canada Social Insurance Number - Min count 10, Max count 500 Canada Health Service Number - Min count 10, Max count 500 Canada Personal Health Identification Number (PHIN) - Min count 10, Max count 500 Content is shared with: People outside my organization	Block access to content Send a notification Allow override Require business justification Send incident report

Canada Personal Health Act (PHIPA) - Ontario

Expand table

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Canada PHIPA: Scan content shared outside - low count	Content contains sensitive information: Canada Passport Number - Min count 1, Max count 9 Canada Social Insurance Number - Min count 1, Max count 9 Canada Health Service Number - Min count 1, Max count 9 Canada Personal Health Identification Number (PHIN) - Min count 1, Max count 9 Content is shared with: People outside my organization	Send a notification
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Canada PHIPA: Scan content shared outside - high count	Content contains sensitive information: Canada Passport Number - Min count 10, Max count 500 Canada Social Insurance Number - Min count 10, Max count 500 Canada Health Service Number - Min count 10, Max count 500 Canada Personal Health Identification Number (PHIN) - Min count 10, Max count 500 Content is shared with: People outside my organization	Block access to content Send a notification Allow override Require business justification Send incident report

Canada Personal Health Information Act (PHIA) - Manitoba

Expand table

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Canada PHIA: Scan content shared outside - low count	Content contains sensitive information: Canada Social Insurance Number - Min count 1, Max count 9 Canada Health Service Number - Min count 1, Max count 9 Canada Personal Health Identification Number (PHIN) - Min count 1, Max count 9 Content is shared with: People outside my organization	Send a notification
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Canada PHIA: Scan content shared outside - high count	Content contains sensitive information: Canada Social Insurance Number - Min count 10, Max count 500 Canada Health Service Number - Min count 10, Max count 500 Canada Personal Health Identification Number (PHIN) - Min count 10, Max count 500 Content is shared with: People outside my organization	Block access to content Send a notification Allow override Require business justification Send incident report

Canada Personal Information Protection Act (PIPA)

Expand table

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and	Canada PIPA: Scan content	Content contains sensitive information: Canada Passport Number -	Send a notification

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
channel messages, Devices, Instances, On-premises repositories	shared outside - low count	Min count 1, Max count 9 Canada Social Insurance Number - Min count 1, Max count 9 Canada Health Service Number - Min count 1, Max count 9 Canada Personal Health Identification Number (PHIN) - Min count 1, Max count 9 Content is shared with: People outside my organization	
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Canada PIPA: Scan content shared outside - high count	Content contains sensitive information: Canada Passport Number - Min count 10, Max count 500 Canada Social Insurance Number - Min count 10, Max count 500 Canada Health Service Number - Min count 10, Max count 500 Canada Personal Health Identification Number (PHIN) - Min count 10, Max count 500 Content is shared with: People outside my organization	Block access to content Send a notification Allow override Require business justification Send incident report

Canada Personal Information Protection Act (PIPEDA)

Expand table

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Canada PIPEDA: Scan content shared outside - low count	Content contains sensitive information: Canada Driver's License Number - Min count 1, Max count 9 Canada Bank Account Number - Min count 1, Max count 9 Canada Passport Number - Min count 1, Max count 9 Canada Social Insurance Number - Min count 1, Max count 9 Canada Health Service Number - Min count 1, Max count 9 Canada Personal Health Identification Number (PHIN) - Min count 1, Max count 9 Content is shared with: People outside my organization	Send a notification
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Canada PIPEDA: Scan content shared outside - high count	Content contains sensitive information: Canada Driver's License Number - Min count 10, Max count 500 Canada Bank Account Number - Min count 10, Max count 500 Canada Passport Number - Min count 10, Max count 500 Canada Social Insurance Number - Min count 10, Max count 500 Canada Health Service Number - Min count 10, Max count 500 Canada Personal Health Identification Number (PHIN) - Min count 10, Max count 500	Block access to content Send a notification Allow override Require business justification Send incident report

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
		Content is shared with: People outside my organization	

Canada Personally Identifiable Information (PII) Data

Expand table

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Canada PII: Scan content shared outside - low count	Content contains sensitive information: Canada Driver's License Number - Min count 1, Max count 9 Canada Bank Account Number - Min count 1, Max count 9 Canada Passport Number - Min count 1, Max count 9 Canada Social Insurance Number - Min count 1, Max count 9 Canada Health Service Number - Min count 1, Max count 9 Canada Personal Health Identification Number (PHIN) - Min count 1, Max count 9 Content is shared with: People outside my organization	Send a notification
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Canada PII: Scan content shared outside - high count	Content contains sensitive information: Canada Driver's License Number - Min count 10, Max count 500 Canada Bank Account	Block access to content Send a notification Allow override Require

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
		Number - Min count 10, Max count 500 Canada Passport Number - Min count 10, Max count 500 Canada Social Insurance Number - Min count 10, Max count 500 Canada Health Service Number - Min count 10, Max count 500 Canada Personal Health Identification Number (PHIN) - Min count 10, Max count 500 Content is shared with: People outside my organization	business justification Send incident report

France Data Protection Act

Expand table

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	France DPA: Scan content shared outside - low count	Content contains sensitive information: France National ID Card (CNI) - Min count 1, Max count 9 France Social Security Number (INSEE) - Min count 1, Max count 9 Content is shared with: People outside my organization	Send a notification
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	France DPA: Scan content shared outside - high count	Content contains sensitive information: France National ID Card (CNI) - Min count 10, Max count 500	Block access to content Send a notification Allow override

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
		France Social Security Number (INSEE) - Min count 10, Max count 500 Content is shared with: People outside my organization	Require business justification Send incident report

France Financial Data

Expand table

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	France Financial: Scan content shared outside - low count	Content contains sensitive information: Credit Card Number - Min count 1, Max count 9 EU Debit Card Number - Min count 1, Max count 9 Content is shared with: People outside my organization	Send a notification
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	France Financial: Scan content shared outside - high count	Content contains sensitive information: Credit Card Number - Min count 10, Max count 500 EU Debit Card Number - Min count 10, Max count 500 Content is shared with: People outside my organization	Block access to content Send a notification Allow override Require business justification Send incident report

France Personally Identifiable Information (PII) Data

Expand table

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	France PII: Scan content shared outside - low count	Content contains sensitive information: France Social Security Number (INSEE) - Min count 1, Max count 9 France Driver's License Number - Min count 1, Max count 9 France Passport Number - Min count 1, Max count 9 France National ID Card (CNI) - Min count 1, Max count 9 Content is shared with: People outside my organization	Send a notification
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	France PII: Scan content shared outside - high count	Content contains sensitive information: France Social Security Number (INSEE) - Min count 10, Max count 500 France Driver's License Number - Min count 10, Max count 500 France Passport Number - Min count 10, Max count 500 France National ID Card (CNI) - Min count 10, Max count 500 Content is shared with: People outside my organization	Block access to content Send a notification Allow override Require business justification Send incident report

Enhanced General Data Protection Regulation (GDPR)

Expand table

Supported Location	Rule name	Conditions (including sensitive information types & trainable classifiers)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Low volume content detected GDPR Enhanced	Content contains sensitive information: Austria Physical Addresses - Min count 1, Max count 9 Belgium Physical Addresses - Min count 1, Max count 9 Bulgaria Physical Addresses - Min count 1, Max count 9 Croatia Physical Addresses - Min count 1, Max count 9 Cyprus Physical Addresses - Min count 1, Max count 9 Czech Republic Physical Addresses - Min count 1, Max count 9 Denmark Physical Addresses - Min count 1, Max count 9 Estonia Physical Addresses - Min count 1, Max count 9 Finland Physical Addresses - Min count 1, Max count 9 France Physical Addresses - Min count 1, Max count 9 Germany Physical Addresses - Min count 1, Max count 9 Greece Physical Addresses - Min count 1, Max count 9 Hungary Physical Addresses - Min count 1, Max count 9 Ireland Physical Addresses - Min count 1, Max count 9 Italy Physical Addresses - Min count 1, Max count 9 Latvia Physical Addresses - Min count 1, Max count 9 Lithuania Physical Addresses - Min count 1, Max count 9 Luxembourg Physical	Send incident reports to Administrator

Supported Location	Rule name	Conditions (including sensitive information types & trainable classifiers)	Actions
		Addresses - Min count 1, Max count 9 Malta Physical Addresses - Min count 1, Max count 9 Netherlands Physical Addresses - Min count 1, Max count 9 Poland Physical Addresses - Min count 1, Max count 9 Portuguese Physical Addresses - Min count 1, Max count 9 Romania Physical Addresses - Min count 1, Max count 9 Slovakia Physical Addresses - Min count 1, Max count 9 Slovenia Physical Addresses - Min count 1, Max count 9 Spain Physical Addresses - Min count 1, Max count 9 Sweden Physical Addresses - Min count 1, Max count 9 Austria Social Security Number - Min count 1, Max count 9 France Social Security Number (INSEE) - Min count 1, Max count 9 Greece Social Security Number (AMKA) - Min count 1, Max count 9 Hungarian Social Security Number (TAJ) - Min count 1, Max count 9 Spain Social Security Number (SSN) - Min count 1, Max count 9 Austria Identity Card - Min count 1, Max count 9 Cyprus Identity Card - Min count 1, Max count 9 Germany Identity Card Number - Min count 1, Max count 9 Malta Identity Card Number	

Supported Location	Rule name	Conditions (including sensitive information types & trainable classifiers)	Actions
		- Min count 1, Max count 9 France National ID Card (CNI) - Min count 1, Max count 9 Greece National ID Card - Min count 1, Max count 9 Finland National ID - Min count 1, Max count 9 Poland Identity Card - Min count 1, Max count 9 Sweden National ID - Min count 1, Max count 9 Croatia Personal Identification (OIB) Number - Min count 1, Max count 9 Czech Personal Identity Number - Min count 1, Max count 9 Denmark Personal Identification Number - Min count 1, Max count 9 Estonia Personal Identification Code - Min count 1, Max count 9 Hungary Personal Identification Number - Min count 1, Max count 9 Luxembourg National Identification Number (Natural persons) - Min count 1, Max count 9 Luxembourg National Identification Number (Non-natural persons) - Min count 1, Max count 9 Italy Fiscal Code - Min count 1, Max count 9 Latvia Personal Code - Min count 1, Max count 9 Lithuania Personal Code - Min count 1, Max count 9 Romania Personal Numerical Code (CNP) - Min count 1, Max count 9 Netherlands Citizen's Service	

Supported Location	Rule name	Conditions (including sensitive information types & trainable classifiers)	Actions
		(BSN) Number - Min count 1, Max count 9 Ireland Personal Public Service (PPS) Number - Min count 1, Max count 9 Bulgaria Uniform Civil Number - Min count 1, Max count 9 Belgium National Number - Min count 1, Max count 9 Spain DNI - Min count 1, Max count 9 Slovenia Unique Master Citizen Number - Min count 1, Max count 9 Slovakia Personal Number - Min count 1, Max count 9 Portugal Citizen Card Number - Min count 1, Max count 9 Malta Tax ID Number - Min count 1, Max count 9 Austria Tax Identification Number - Min count 1, Max count 9 Cyprus Tax Identification Number - Min count 1, Max count 9 France Tax Identification Number (numéro SPI.) - Min count 1, Max count 9 Germany Tax Identification Number - Min count 1, Max count 9 Greek Tax identification Number - Min count 1, Max count 9 Hungary Tax identification Number - Min count 1, Max count 9 Netherlands Tax Identification Number - Min count 1, Max count 9 Poland Tax Identification Number - Min count 1, Max	

Supported Location	Rule name	Conditions (including sensitive information types & trainable classifiers)	Actions
		count 9 Portugal Tax Identification Number - Min count 1, Max count 9 Slovenia Tax Identification Number - Min count 1, Max count 9 Spain Tax Identification Number - Min count 1, Max count 9 Sweden Tax Identification Number - Min count 1, Max count 9 Contains content matching any of the following trainable classifiers: HR Tax Invoice Healthcare Health/Medical Forms Employee disciplinary action files Legal affairs Agreements Content is shared with: People outside my organization	
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	High volume content detected GDPR Enhanced	Content contains sensitive information: Austria Physical Addresses - Min count 10, Max count 500 Belgium Physical Addresses - Min count 10, Max count 500 Bulgaria Physical Addresses - Min count 10, Max count 500 Croatia Physical Addresses - Min count 10, Max count 500 Cyprus Physical Addresses - Min count 10, Max count 500 Czech Republic Physical Addresses - Min count 10, Max count 500 Denmark Physical Addresses	Restrict access to the content for external users Notify users with email and policy tips Allow override Require business justification Send incident reports to Administrator

Supported Location	Rule name	Conditions (including sensitive information types & trainable classifiers)	Actions
		- Min count 10, Max count 500 Estonia Physical Addresses - Min count 10, Max count 500 Finland Physical Addresses - Min count 10, Max count 500 France Physical Addresses - Min count 10, Max count 500 Germany Physical Addresses - Min count 10, Max count 500 Greece Physical Addresses - Min count 10, Max count 500 Hungary Physical Addresses - Min count 10, Max count 500 Ireland Physical Addresses - Min count 10, Max count 500 Italy Physical Addresses - Min count 10, Max count 500 Latvia Physical Addresses - Min count 10, Max count 500 Lithuania Physical Addresses - Min count 10, Max count 500 Luxembourg Physical Addresses - Min count 10, Max count 500 Malta Physical Addresses - Min count 10, Max count 500 Netherlands Physical Addresses - Min count 10, Max count 500 Poland Physical Addresses - Min count 10, Max count 500 Portuguese Physical Addresses - Min count 10, Max count 500 Romania Physical Addresses - Min count 10, Max count 500 Slovakia Physical Addresses - Min count 10, Max count 500 Slovenia Physical Addresses - Min count 10, Max count 500	

Supported Location	Rule name	Conditions (including sensitive information types & trainable classifiers)	Actions
		Spain Physical Addresses - Min count 10, Max count 500 Sweden Physical Addresses - Min count 10, Max count 500 Austria Social Security Number - Min count 10, Max count 500 France Social Security Number (INSEE) - Min count 10, Max count 500 Greece Social Security Number (AMKA) - Min count 10, Max count 500 Hungarian Social Security Number (TAJ) - Min count 10, Max count 500 Spain Social Security Number (SSN) - Min count 10, Max count 500 Austria Identity Card - Min count 10, Max count 500 Cyprus Identity Card - Min count 10, Max count 500 Germany Identity Card Number - Min count 10, Max count 500 Malta Identity Card Number - Min count 10, Max count 500 France National ID Card (CNI) - Min count 10, Max count 500 Greece National ID Card - Min count 10, Max count 500 Finland National ID - Min count 10, Max count 500 Poland Identity Card - Min count 10, Max count 500 Sweden National ID - Min count 10, Max count 500 Croatia Personal Identification (OIB) Number - Min count 10, Max count 500 Czech Personal Identity Number - Min count 10, Max	

Supported Location	Rule name	Conditions (including sensitive information types & trainable classifiers)	Actions
		count 500 Denmark Personal Identification Number - Min count 10, Max count 500 Estonia Personal Identification Code - Min count 10, Max count 500 Hungary Personal Identification Number - Min count 10, Max count 500 Luxembourg National Identification Number (Natural persons) - Min count 10, Max count 500 Luxembourg National Identification Number (Non-natural persons) - Min count 10, Max count 500 Italy Fiscal Code - Min count 10, Max count 500 Latvia Personal Code - Min count 10, Max count 500 Lithuania Personal Code - Min count 10, Max count 500 Romania Personal Numerical Code (CNP) - Min count 10, Max count 500 Netherlands Citizen's Service (BSN) Number - Min count 10, Max count 500 Ireland Personal Public Service (PPS) Number - Min count 10, Max count 500 Bulgaria Uniform Civil Number - Min count 10, Max count 500 Belgium National Number - Min count 10, Max count 500 Spain DNI - Min count 10, Max count 500 Slovenia Unique Master Citizen Number - Min count 10, Max count 500 Slovakia Personal Number - Min count 10, Max count 500	

Supported Location	Rule name	Conditions (including sensitive information types & trainable classifiers)	Actions
		Portugal Citizen Card Number - Min count 10, Max count 500 Malta Tax ID Number - Min count 10, Max count 500 Austria Tax Identification Number - Min count 10, Max count 500 Cyprus Tax Identification Number - Min count 10, Max count 500 France Tax Identification Number (numéro SPI.) - Min count 10, Max count 500 Germany Tax Identification Number - Min count 10, Max count 500 Greek Tax identification Number - Min count 10, Max count 500 Hungary Tax identification Number - Min count 10, Max count 500 Netherlands Tax Identification Number - Min count 10, Max count 500 Poland Tax Identification Number - Min count 10, Max count 500 Portugal Tax Identification Number - Min count 10, Max count 500 Slovenia Tax Identification Number - Min count 10, Max count 500 Spain Tax Identification Number - Min count 10, Max count 500 Sweden Tax Identification Number - Min count 10, Max count 500 Contains content matching any of the following trainable classifiers: HR	

Supported Location	Rule name	Conditions (including sensitive information types & trainable classifiers)	Actions
		Tax Invoice Healthcare Health/Medical Forms Employee disciplinary action files Legal affairs Agreements Content is shared with: People outside my organization	
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Low volume EU Sensitive Enhanced content found	Content contains sensitive information: Austria Physical Addresses - Min count 1, Max count 9 Belgium Physical Addresses - Min count 1, Max count 9 Bulgaria Physical Addresses - Min count 1, Max count 9 Croatia Physical Addresses - Min count 1, Max count 9 Cyprus Physical Addresses - Min count 1, Max count 9 Czech Republic Physical Addresses - Min count 1, Max count 9 Denmark Physical Addresses - Min count 1, Max count 9 Estonia Physical Addresses - Min count 1, Max count 9 Finland Physical Addresses - Min count 1, Max count 9 France Physical Addresses - Min count 1, Max count 9 Germany Physical Addresses - Min count 1, Max count 9 Greece Physical Addresses - Min count 1, Max count 9 Hungary Physical Addresses - Min count 1, Max count 9 Ireland Physical Addresses - Min count 1, Max count 9 Italy Physical Addresses - Min count 1, Max count 9	Send incident reports to Administrator

Supported Location	Rule name	Conditions (including sensitive information types & trainable classifiers)	Actions
		Latvia Physical Addresses - Min count 1, Max count 9 Lithuania Physical Addresses - Min count 1, Max count 9 Luxembourg Physical Addresses - Min count 1, Max count 9 Malta Physical Addresses - Min count 1, Max count 9 Netherlands Physical Addresses - Min count 1, Max count 9 Poland Physical Addresses - Min count 1, Max count 9 Portuguese Physical Addresses - Min count 1, Max count 9 Romania Physical Addresses - Min count 1, Max count 9 Slovakia Physical Addresses - Min count 1, Max count 9 Slovenia Physical Addresses - Min count 1, Max count 9 Spain Physical Addresses - Min count 1, Max count 9 Sweden Physical Addresses - Min count 1, Max count 9 Austria Social Security Number - Min count 1, Max count 9 France Social Security Number (INSEE) - Min count 1, Max count 9 Greece Social Security Number (AMKA) - Min count 1, Max count 9 Hungarian Social Security Number (TAJ) - Min count 1, Max count 9 Spain Social Security Number (SSN) - Min count 1, Max count 9 Austria Identity Card - Min count 1, Max count 9 Cyprus Identity Card - Min	

Supported Location	Rule name	Conditions (including sensitive information types & trainable classifiers)	Actions
		count 1, Max count 9 Germany Identity Card Number - Min count 1, Max count 9 Malta Identity Card Number - Min count 1, Max count 9 France National ID Card (CNI) - Min count 1, Max count 9 Greece National ID Card - Min count 1, Max count 9 Finland National ID - Min count 1, Max count 9 Poland Identity Card - Min count 1, Max count 9 Sweden National ID - Min count 1, Max count 9 Croatia Personal Identification (OIB) Number - Min count 1, Max count 9 Czech Personal Identity Number - Min count 1, Max count 9 Denmark Personal Identification Number - Min count 1, Max count 9 Estonia Personal Identification Code - Min count 1, Max count 9 Hungary Personal Identification Number - Min count 1, Max count 9 Luxembourg National Identification Number (Natural persons) - Min count 1, Max count 9 Luxembourg National Identification Number (Non-natural persons) - Min count 1, Max count 9 Italy Fiscal Code - Min count 1, Max count 9 Latvia Personal Code - Min count 1, Max count 9 Lithuania Personal Code -	

Supported Location	Rule name	Conditions (including sensitive information types & trainable classifiers)	Actions
		Min count 1, Max count 9 Romania Personal Numerical Code (CNP) - Min count 1, Max count 9 Netherlands Citizen's Service (BSN) Number - Min count 1, Max count 9 Ireland Personal Public Service (PPS) Number - Min count 1, Max count 9 Bulgaria Uniform Civil Number - Min count 1, Max count 9 Belgium National Number - Min count 1, Max count 9 Spain DNI - Min count 1, Max count 9 Slovenia Unique Master Citizen Number - Min count 1, Max count 9 Slovakia Personal Number - Min count 1, Max count 9 Portugal Citizen Card Number - Min count 1, Max count 9 Malta Tax ID Number - Min count 1, Max count 9 Austria Tax Identification Number - Min count 1, Max count 9 Cyprus Tax Identification Number - Min count 1, Max count 9 France Tax Identification Number (numéro SPI.) - Min count 1, Max count 9 Germany Tax Identification Number - Min count 1, Max count 9 Greek Tax identification Number - Min count 1, Max count 9 Hungary Tax identification Number - Min count 1, Max count 9	

Supported Location	Rule name	Conditions (including sensitive information types & trainable classifiers)	Actions
		Netherlands Tax Identification Number - Min count 1, Max count 9 Poland Tax Identification Number - Min count 1, Max count 9 Portugal Tax Identification Number - Min count 1, Max count 9 Slovenia Tax Identification Number - Min count 1, Max count 9 Spain Tax Identification Number - Min count 1, Max count 9 Sweden Tax Identification Number - Min count 1, Max count 9 EU Debit Card Number - Min count 1, Max count 9 Contains content matching any of the following trainable classifiers: HR Tax Invoice Healthcare Health/Medical Forms Employee disciplinary action files Legal affairs Agreements Content is shared with: People outside my organization	
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	High volume of EU Sensitive Enhanced content found	Content contains sensitive information: Austria Physical Addresses - Min count 10, Max count 500 Belgium Physical Addresses - Min count 10, Max count 500 Bulgaria Physical Addresses - Min count 10, Max count 500 Croatia Physical Addresses -	Restrict access to the content for external users Notify users with email and policy tips Allow override Require business justification

Supported Location	Rule name	Conditions (including sensitive information types & trainable classifiers)	Actions
		Min count 10, Max count 500 Cyprus Physical Addresses - Min count 10, Max count 500 Czech Republic Physical Addresses - Min count 10, Max count 500 Denmark Physical Addresses - Min count 10, Max count 500 Estonia Physical Addresses - Min count 10, Max count 500 Finland Physical Addresses - Min count 10, Max count 500 France Physical Addresses - Min count 10, Max count 500 Germany Physical Addresses - Min count 10, Max count 500 Greece Physical Addresses - Min count 10, Max count 500 Hungary Physical Addresses - Min count 10, Max count 500 Ireland Physical Addresses - Min count 10, Max count 500 Italy Physical Addresses - Min count 10, Max count 500 Latvia Physical Addresses - Min count 10, Max count 500 Lithuania Physical Addresses - Min count 10, Max count 500 Luxembourg Physical Addresses - Min count 10, Max count 500 Malta Physical Addresses - Min count 10, Max count 500 Netherlands Physical Addresses - Min count 10, Max count 500 Poland Physical Addresses - Min count 10, Max count 500 Portuguese Physical Addresses - Min count 10, Max count 500	Send incident reports to Administrator

Supported Location	Rule name	Conditions (including sensitive information types & trainable classifiers)	Actions
		Romania Physical Addresses - Min count 10, Max count 500 Slovakia Physical Addresses - Min count 10, Max count 500 Slovenia Physical Addresses - Min count 10, Max count 500 Spain Physical Addresses - Min count 10, Max count 500 Sweden Physical Addresses - Min count 10, Max count 500 Austria Social Security Number - Min count 10, Max count 500 France Social Security Number (INSEE) - Min count 10, Max count 500 Greece Social Security Number (AMKA) - Min count 10, Max count 500 Hungarian Social Security Number (TAJ) - Min count 10, Max count 500 Spain Social Security Number (SSN) - Min count 10, Max count 500 Austria Identity Card - Min count 10, Max count 500 Cyprus Identity Card - Min count 10, Max count 500 Germany Identity Card Number - Min count 10, Max count 500 Malta Identity Card Number - Min count 10, Max count 500 France National ID Card (CNI) - Min count 10, Max count 500 Greece National ID Card - Min count 10, Max count 500 Finland National ID - Min count 10, Max count 500 Poland Identity Card - Min count 10, Max count 500	

Supported Location	Rule name	Conditions (including sensitive information types & trainable classifiers)	Actions
		Sweden National ID - Min count 10, Max count 500 Croatia Personal Identification (OIB) Number - Min count 10, Max count 500 Czech Personal Identity Number - Min count 10, Max count 500 Denmark Personal Identification Number - Min count 10, Max count 500 Estonia Personal Identification Code - Min count 10, Max count 500 Hungary Personal Identification Number - Min count 10, Max count 500 Luxembourg National Identification Number (Natural persons) - Min count 10, Max count 500 Luxembourg National Identification Number (Non-natural persons) - Min count 10, Max count 500 Italy Fiscal Code - Min count 10, Max count 500 Latvia Personal Code - Min count 10, Max count 500 Lithuania Personal Code - Min count 10, Max count 500 Romania Personal Numerical Code (CNP) - Min count 10, Max count 500 Netherlands Citizen's Service (BSN) Number - Min count 10, Max count 500 Ireland Personal Public Service (PPS) Number - Min count 10, Max count 500 Bulgaria Uniform Civil Number - Min count 10, Max count 500 Belgium National Number - Min count 10, Max count 500	

Supported Location	Rule name	Conditions (including sensitive information types & trainable classifiers)	Actions
		Spain DNI - Min count 10, Max count 500 Slovenia Unique Master Citizen Number - Min count 10, Max count 500 Slovakia Personal Number - Min count 10, Max count 500 Portugal Citizen Card Number - Min count 10, Max count 500 Malta Tax ID Number - Min count 10, Max count 500 Austria Tax Identification Number - Min count 10, Max count 500 Cyprus Tax Identification Number - Min count 10, Max count 500 France Tax Identification Number (numéro SPI.) - Min count 10, Max count 500 Germany Tax Identification Number - Min count 10, Max count 500 Greek Tax identification Number - Min count 10, Max count 500 Hungary Tax identification Number - Min count 10, Max count 500 Netherlands Tax Identification Number - Min count 10, Max count 500 Poland Tax Identification Number - Min count 10, Max count 500 Portugal Tax Identification Number - Min count 10, Max count 500 Slovenia Tax Identification Number - Min count 10, Max count 500 Spain Tax Identification Number - Min count 10, Max count 500	

Supported Location	Rule name	Conditions (including sensitive information types & trainable classifiers)	Actions
		Sweden Tax Identification Number - Min count 10, Max count 500 EU Debit Card Number - Min count 10, Max count 500 Contains content matching any of the following trainable classifiers: HR Tax Invoice Healthcare Health/Medical Forms Employee disciplinary action files Legal affairs Agreements Content is shared with: People outside my organization	

Germany Financial Data

Expand table

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Germany Financial Data: Scan content shared outside - low count	Content contains sensitive information: Credit Card Number - Min count 1, Max count 9 EU Debit Card Number - Min count 1, Max count 9 Content is shared with: People outside my organization	Send a notification

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Germany Financial Data: Scan content shared outside - high count	Content contains sensitive information: Credit Card Number - Min count 10, Max count 500 EU Debit Card Number - Min count 10, Max count 500 Content is shared with: People outside my organization	Block access to content Send a notification Allow override Require business justification Send incident report

Germany Personally Identifiable Information (PII) Data

Expand table

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Germany PII: Scan content shared outside - low count	Content contains sensitive information: German Driver's License Number - Min count 1, Max count 9 German Passport Number - Min count 1, Max count 9 Content is shared with: People outside my organization	Send a notification
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Germany PII: Scan content shared outside - high count	Content contains sensitive information: German Driver's License Number - Min count 10, Max count 500 German Passport Number - Min count 10, Max count 500 Content is shared with:	Block access to content Send a notification Allow override Require business justification

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
		People outside my organization	Send incident report

Israel Financial Data

Expand table

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Israel Financial Data: Scan content shared outside - low count	Content contains sensitive information: Israel Bank Account Number - Min count 1, Max count 9 SWIFT Code - Min count 1, Max count 9 Credit Card Number - Min count 1, Max count 9 Content is shared with: People outside my organization	Send a notification
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Israel Financial Data: Scan content shared outside - high count	Content contains sensitive information: Israel Bank Account Number - Min count 10, Max count 500 SWIFT Code - Min count 10, Max count 500 Credit Card Number - Min count 10, Max count 500 Content is shared with: People outside my organization	Block access to content Send a notification Allow override Require business justification Send incident report

Israel Personally Identifiable Information (PII) Data

Expand table

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Israel PII: Scan content shared outside - low count	Content contains sensitive information: Israel National ID - Min count 1, Max count 9 Content is shared with: People outside my organization	Send a notification
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Israel PII: Scan content shared outside - high count	Content contains sensitive information: Israel National ID - Min count 10, Max count 500 Content is shared with: People outside my organization	Block access to content Send a notification Allow override Require business justification Send incident report

Israel Protection of Privacy

Expand table

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Israel Privacy: Scan content shared outside - low count	Content contains sensitive information: Israel National ID - Min count 1, Max count 9 Israel Bank Account Number - Min count 1, Max count 9 Content is shared with: People outside my organization	Send a notification
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and	Israel Privacy: Scan content shared	Content contains sensitive information:	Block access to content

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
channel messages, Devices, Instances, On-premises repositories	outside - high count	Israel National ID - Min count 10, Max count 500 Israel Bank Account Number - Min count 10, Max count 500 Content is shared with: People outside my organization	Send a notification Allow override Require business justification Send incident report

Japan Financial Data

Expand table

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Japan Financial: Scan content shared outside - low count	Content contains sensitive information: Japan Bank Account Number - Min count 1, Max count 9 Credit Card Number - Min count 1, Max count 9 Content is shared with: People outside my organization	Send a notification
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Japan Financial: Scan content shared outside - high count	Content contains sensitive information: Japan Bank Account Number - Min count 10, Max count 500 Credit Card Number - Min count 10, Max count 500 Content is shared with: People outside my organization	Block access to content Send a notification Allow override Require business justification Send incident report

Japan Personally Identifiable Information (PII) Data

Expand table

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Japan PII: Scan content shared outside - low count	Content contains sensitive information: Japan Social Insurance Number (SIN) - Min count 1, Max count 9 Japanese My Number Personal - Min count 1, Max count 9 Japan Passport Number - Min count 1, Max count 9 Japan Driver's License Number - Min count 1, Max count 9 Japan Physical Addresses - Min count 1, Max count 9 Content is shared with: People outside my organization	Send a notification
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Japan PII: Scan content shared outside - high count	Content contains sensitive information: Japan Social Insurance Number (SIN) - Min count 10, Max count 500 Japanese My Number Personal - Min count 10, Max count 500 Japan Passport Number - Min count 10, Max count 500 Japan Driver's License Number - Min count 10, Max count 500 Japan Physical Addresses - Min count 10, Max count 500 Content is shared with: People outside my organization	Block access to content Send a notification Allow override Require business justification Send incident report

Japan Protection of Personal Information Enhanced

Expand table

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Japan PPI: Scan content shared outside - low count	Content contains sensitive information: Japan Social Insurance Number (SIN) - Min count 1, Max count 9 Japanese My Number Personal - Min count 1, Max count 9 Japan Passport Number - Min count 1, Max count 9 Japan Driver's License Number - Min count 1, Max count 9 Japan Physical Addresses - Min count 1, Max count 9 Content is shared with: People outside my organization	Send a notification
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Japan PPI: Scan content shared outside - high count	Content contains sensitive information: Japan Social Insurance Number (SIN) - Min count 10, Max count 500 Japanese My Number Personal - Min count 10, Max count 500 Japan Passport Number - Min count 10, Max count 500 Japan Driver's License Number - Min count 10, Max count 500 Japan Physical Addresses - Min count 10, Max count 500 Content is shared with: People outside my organization	Block access to content Send a notification Allow override Require business justification Send incident report

PCI Data Security Standard (PCI DSS)

Expand table

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	PCI DSS: Scan content shared outside - low count	Content contains sensitive information: Credit Card Number - Min count 1, Max count 9 Content is shared with: People outside my organization	Send a notification
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	PCI DSS: Scan content shared outside - high count	Content contains sensitive information: Credit Card Number - Min count 10, Max count 500 Content is shared with: People outside my organization	Block access to content Send a notification Allow override Require business justification Send incident report

Saudi Arabia - Anti-Cyber Crime Law

Expand table

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Saudi Arabia ACC: Scan content shared outside - low count	Content contains sensitive information: SWIFT Code - Min count 1, Max count 9 International Banking Account Number (IBAN) - Min count 1, Max count 9 Content is shared with:	Send a notification

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
		People outside my organization	
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Saudi Arabia ACC: Scan content shared outside - high count	Content contains sensitive information: SWIFT Code - Min count 10, Max count 500 International Banking Account Number (IBAN) - Min count 10, Max count 500 Content is shared with: People outside my organization	Block access to content Send a notification Allow override Require business justification Send incident report

Saudi Arabia Financial Data

Expand table

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Saudi Arabia Financial: Scan content shared outside - low count	Content contains sensitive information: Credit Card Number - Min count 1, Max count 9 SWIFT Code - Min count 1, Max count 9 International Banking Account Number (IBAN) - Min count 1, Max count 9 Content is shared with: People outside my organization	Send a notification
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Saudi Arabia Financial: Scan content shared outside - high count	Content contains sensitive information: Credit Card Number - Min count 10, Max count 500 SWIFT Code - Min count 10, Max count 500 International Banking	Block access to content Send a notification Allow override Require business justification

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
		Account Number (IBAN) - Min count 10, Max count 500 Content is shared with: People outside my organization	Send incident report

Saudi Arabia Personally Identifiable Information (PII) Data

Expand table

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Saudi Arabia PII: Scan content shared outside - low count	Content contains sensitive information: Saudi Arabia National ID - Min count 1, Max count 9 Content is shared with: People outside my organization	Send a notification
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Saudi Arabia PII: Scan content shared outside - high count	Content contains sensitive information: Saudi Arabia National ID - Min count 10, Max count 500 Content is shared with: People outside my organization	Block access to content Send a notification Allow override Require business justification Send incident report

U.K. Access to Medical Reports Act

Expand table

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	U.K. AMRA: Scan content shared outside - low count	Content contains sensitive information: U.K. National Health Service Number - Min count 1, Max count 9 U.K. National Insurance Number (NINO) - Min count 1, Max count 9 Content is shared with: People outside my organization	Send a notification
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	U.K. AMRA: Scan content shared outside - high count	Content contains sensitive information: U.K. National Health Service Number - Min count 10, Max count 500 U.K. National Insurance Number (NINO) - Min count 10, Max count 500 Content is shared with: People outside my organization	Block access to content Send a notification Allow override Require business justification Send incident report

U.K. Data Protection Act

Expand table

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	U.K. DPA: Scan content shared outside - low count	Content contains sensitive information: U.K. National Insurance Number (NINO) - Min count 1, Max count 9 U.S. / U.K. Passport Number - Min count 1, Max count 9 SWIFT Code - Min count 1, Max count 9 Content is shared with:	Send a notification

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
		People outside my organization	
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	U.K. DPA: Scan content shared outside - high count	Content contains sensitive information: U.K. National Insurance Number (NINO) - Min count 10, Max count 500 U.S. / U.K. Passport Number - Min count 10, Max count 500 SWIFT Code - Min count 10, Max count 500 Content is shared with: People outside my organization	Block access to content Send a notification Allow override Require business justification Send incident report

U.K. Financial Data

Expand table

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	U.K. Financial: Scan content shared outside - low count	Content contains sensitive information: Credit Card Number - Min count 1, Max count 9 EU Debit Card Number - Min count 1, Max count 9 SWIFT Code -Min count 1, Max count 9 Content is shared with: People outside my organization	Send a notification
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	U.K. Financial: Scan content shared outside - high count	Content contains sensitive information: Credit Card Number - Min count 10, Max	Block access to content Send a notification

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
		count 500 EU Debit Card Number - Min count 10, Max count 500 SWIFT Code - Min count 10, Max count 500 Content is shared with: People outside my organization	Allow override Require business justification Send incident report

U.K. Personal Information Online Code of Practice (PIOCP)

Expand table

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	U.K. PIOCP: Scan content shared outside - low count	Content contains sensitive information: U.K. National Insurance Number (NINO) - Min count 1, Max count 9 U.K. National Health Service Number - Min count 1, Max count 9 SWIFT Code - Min count 1, Max count 9 Content is shared with: People outside my organization	Send a notification
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	U.K. PIOCP: Scan content shared outside - high count	Content contains sensitive information: U.K. National Insurance Number (NINO) - Min count 10, Max count 500 U.K. National Health Service Number - Min count 10, Max count 500	Block access to content Send a notification Allow override Require business justification

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
		SWIFT Code - Min count 10, Max count 500 Content is shared with: People outside my organization	Send incident report

U.K. Personally Identifiable Information (PII) Data

Expand table

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	U.K. PII: Scan content shared outside - low count	Content contains sensitive information: U.K. National Insurance Number (NINO) - Min count 1, Max count 9 U.S. / U.K. Passport Number - Min count 1, Max count 9 Content is shared with: People outside my organization	Send a notification
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	U.K. PII: Scan content shared outside - high count	Content contains sensitive information: U.K. National Insurance Number (NINO) - Min count 10, Max count 500 U.S. / U.K. Passport Number - Min count 10, Max count 500 Content is shared with: People outside my organization	Block access to content Send a notification Allow override Require business justification Send incident report

U.K. Privacy and Electronic Communications Regulations

Expand table

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	U.K. PECR: Scan content shared outside - low count	Content contains sensitive information: SWIFT Code - Min count 1, Max count 9 Content is shared with: People outside my organization	Send a notification
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	U.K. PECR: Scan content shared outside - high count	Content contains sensitive information: SWIFT Code - Min count 10, Max count 500 Content is shared with: People outside my organization	Block access to content Send a notification Allow override Require business justification Send incident report

U.S. Federal Trade Commission (FTC) Consumer Rules

Expand table

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	U.S. FTC Rules: Scan content shared outside - low count	Content contains sensitive information: Credit Card Number - Min count 1, Max count 9 U.S. Bank Account Number - Min count 1, Max count 9 ABA Routing Number - Min count 1, Max	Send a notification

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
		count 9 Content is shared with: People outside my organization	
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	U.S. FTC Rules: Scan content shared outside - high count	Content contains sensitive information: Credit Card Number - Min count 10, Max count 500 U.S. Bank Account Number - Min count 10, Max count 500 ABA Routing Number - Min count 10, Max count 500 Content is shared with: People outside my organization	Block access to content Send a notification Allow override Require business justification Send incident report

U.S. Financial Data

Expand table

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	U.S. Financial: Scan content shared outside - low count	Content contains sensitive information: Credit Card Number - Min count 1, Max count 9 U.S. Bank Account Number - Min count 1, Max count 9 ABA Routing Number - Min count 1, Max count 9 Content is shared with: People outside my organization	Send a notification

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	U.S. Financial: Scan content shared outside - high count	Content contains sensitive information: Credit Card Number - Min count 10, Max count 500 U.S. Bank Account Number - Min count 10, Max count 500 ABA Routing Number - Min count 10, Max count 500 Content is shared with: People outside my organization	Block access to content Send a notification Allow override Require business justification Send incident report

U.S. Gramm-Leach-Bliley Act (GLBA) Enhanced

Expand table

Supported Location	Rule name	Conditions (including sensitive information types & trainable classifiers)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	U.S. GLBA: Scan content shared outside - low count	Content contains sensitive information: Credit Card Number - Min count 1, Max count 9 U.S. Bank Account Number - Min count 1, Max count 9 U.S. Individual Taxpayer Identification Number (ITIN) - Min count 1, Max count 9 U.S. Social Security Number (SSN) - Min count 1, Max count 9 U.S. Driver's License Number - Min count 1, Max count 9 U.S. / U.K. Passport Number - Min count 1, Max count 9 U.S. Physical Addresses - Min count 1, Max count 9	Send a notification

Supported Location	Rule name	Conditions (including sensitive information types & trainable classifiers)	Actions
		Contains content matching any of the following trainable classifiers: Tax Finance Budget Content is shared with: People outside my organization	
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	U.S. GLBA: Scan content shared outside - high count	Content contains sensitive information: Credit Card Number - Min count 10, Max count 500 U.S. Bank Account Number - Min count 10, Max count 500 U.S. Individual Taxpayer Identification Number (ITIN) - Min count 10, Max count 500 U.S. Social Security Number (SSN) - Min count 10, Max count 500 U.S. Driver's License Number - Min count 10, Max count 500 U.S. / U.K. Passport Number - Min count 10, Max count 500 U.S. Physical Addresses - Min count 10, Max count 500 Contains content matching any of the following trainable classifiers: Tax Finance Budget Content is shared with: People outside my organization	Block access to content Send a notification Allow override Require business justification Send incident report

U.S. Health Insurance Act (HIPAA) Enhanced

Expand table

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	Content matches U.S. HIPAA	Contains any of the following sensitive information: U.S. Social Security Number (SSN) - Min count 1, Max count any Drug Enforcement Agency (DEA) Number - Min count 1, Max count any U.S. Physical Addresses - Min count 1, Max count any **AND** Content contains any of these terms: International Classification of Diseases (ICD-9-CM) - Min count 1, Max count any International Classification of Diseases (ICD-10-CM) - Min count 1, Max count any All Medical Terms And Conditions - Min count 1, Max count any Contains content matching any of the following trainable classifiers: Healthcare Employee Insurance Files Health/Medical Forms Content is shared with: People outside my organization	Send a notification

U.S. Patriot Act

Expand table

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	U.S. Patriot Act: Scan content shared outside - low count	Content contains sensitive information: Credit Card Number - Min count 1, Max count 9 U.S. Bank Account Number - Min count 1, Max count 9 U.S. Individual Taxpayer Identification Number (ITIN) - Min count 1, Max count 9 U.S. Social Security Number (SSN) - Min count 1, Max count 9 U.S. Physical Addresses - Min count 1, Max count 9 Content is shared with: People outside my organization	Send a notification
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	U.S. Patriot Act: Scan content shared outside - high count	Content contains sensitive information: Credit Card Number - Min count 10, Max count 500 U.S. Bank Account Number - Min count 10, Max count 500 U.S. Individual Taxpayer Identification Number (ITIN) - Min count 10, Max count 500 U.S. Social Security Number (SSN) - Min count 10, Max count 500 U.S. Physical Addresses - Min count 10, Max count 500 Content is shared with: People outside my organization	Block access to content Send a notification Allow override Require business justification Send incident report

U.S. Personally Identifiable Information (PII) Data Enhanced

Expand table

Supported Location	Rule name	Conditions (including sensitive information types & trainable classifiers)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	U.S. PII: Scan content shared outside - low count	Content contains sensitive information: U.S. Individual Taxpayer Identification Number (ITIN) - Min count 1, Max count 9 U.S. Social Security Number (SSN) - Min count 1, Max count 9 U.S. / U.K. Passport Number - Min count 1, Max count 9 U.S. Physical Addresses - Min count 1, Max count 9 Contains content matching any of the following trainable classifiers: HR Tax Invoice Healthcare Health/Medical Forms Employee disciplinary action files Legal affairs Agreements Content is shared with: People outside my organization	Send a notification
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	U.S. PII: Scan content shared outside - high count	Content contains sensitive information: U.S. Individual Taxpayer Identification Number (ITIN) - Min count 10, Max count 500 U.S. Social Security Number (SSN) - Min count 10, Max count 500 U.S. / U.K. Passport Number - Min count 10, Max count 500 U.S. Physical Addresses - Min count 10, Max count 500 Contains content matching any of the following trainable	Block access to content Send a notification Allow override Require business justification Send incident report

Supported Location	Rule name	Conditions (including sensitive information types & trainable classifiers)	Actions
		classifiers: HR Tax Invoice Healthcare Health/Medical Forms Employee disciplinary action files Legal affairs Agreements Content is shared with: People outside my organization	

U.S. State Breach Notification Laws

Expand table

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	U.S. State Breach: Scan content shared outside - low count	Content contains sensitive information: Credit Card Number - Min count 1, Max count 9 U.S. Bank Account Number - Min count 1, Max count 9 U.S. Driver's License Number - Min count 1, Max count 9 U.S. Social Security Number (SSN) - Min count 1, Max count 9 U.S. / U.K. Passport Number - Min count 1, Max count 9 All Medical Terms And Conditions - Min count 1, Max count 9 Content is shared with:	Send a notification

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
		People outside my organization	
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	U.S. State Breach: Scan content shared outside - high count	Content contains sensitive information: Credit Card Number - Min count 10, Max count 500 U.S. Bank Account Number - Min count 10, Max count 500 U.S. Driver's License Number - Min count 10, Max count 500 U.S. Social Security Number (SSN) - Min count 10, Max count 500 U.S. / U.K. Passport Number - Min count 10, Max count 500 All Medical Terms And Conditions - Min count 10, Max count 500 Content is shared with: People outside my organization	Block access to content Send a notification Allow override Require business justification Send incident report

U.S. State Social Security Number Confidentiality Laws

Expand table

Supported Location	Rule name	Conditions (including sensitive information types)	Actions
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	U.S. SSN Laws: Scan content shared outside - low count	Content contains sensitive information: U.S. Social Security Number (SSN) - Min count 1, Max count 9 Content is shared with:	Send a notification

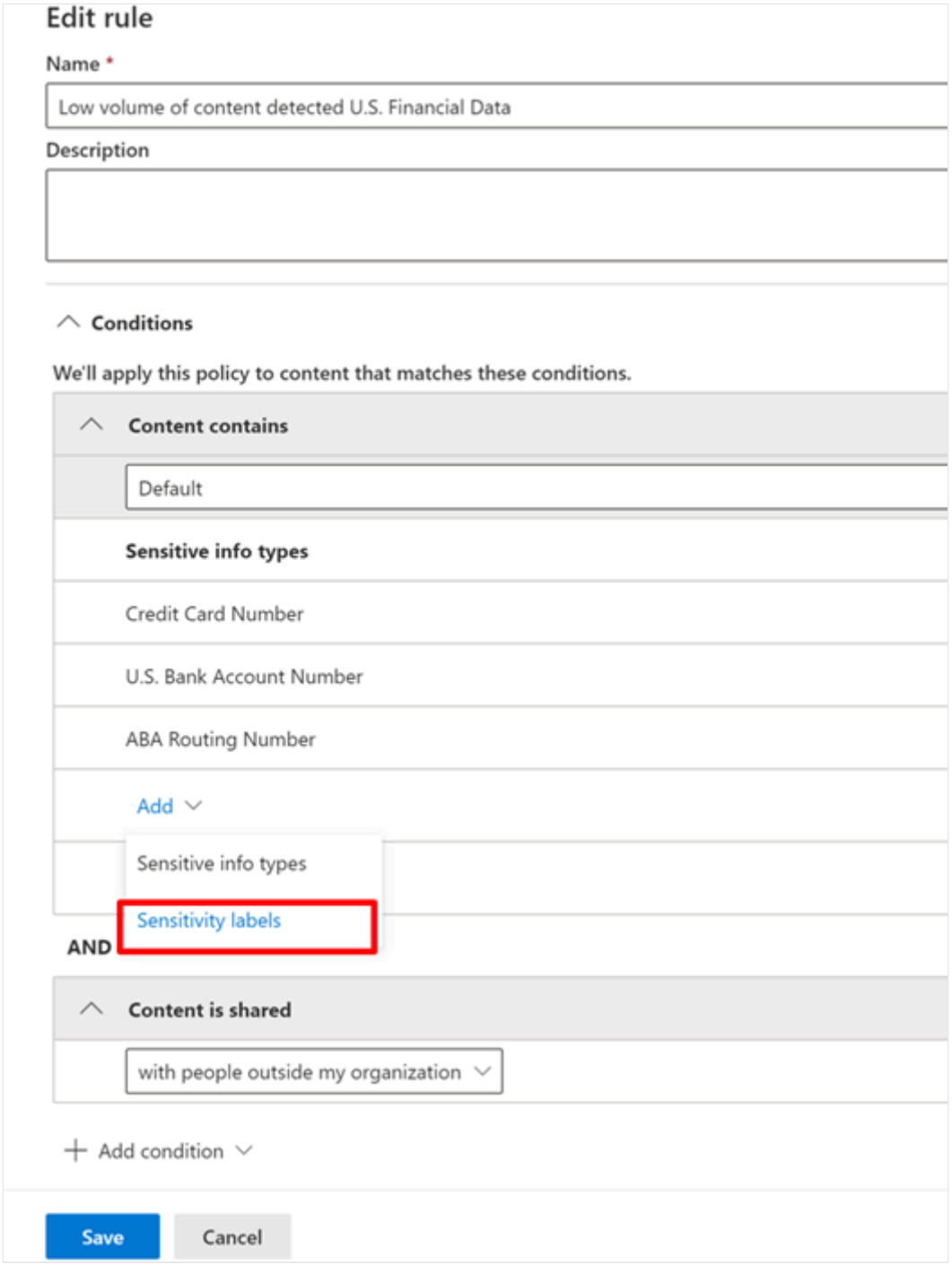
Supported Location	Rule name	Conditions (including sensitive information types)	Actions
		People outside my organization	
Exchange email, SharePoint sites, OneDrive accounts, Teams chat and channel messages, Devices, Instances, On-premises repositories	U.S. SSN Laws: Scan content shared outside - high count	Content contains sensitive information: U.S. Social Security Number (SSN) - Min count 10, Max count 500 Content is shared with: People outside my organization	Block access to content Send a notification Allow override Require business justification Send incident report

Use sensitivity labels as conditions in DLP policies

You can use sensitivity labels as a condition in DLP policies for these locations:

- Exchange email messages
- SharePoint
- OneDrive
- Devices
- Instances
- On-premises repositories
- Microsoft 365 Copilot
- Fabric and Power BI workspaces
- Inline web traffic (Unmanaged cloud apps and adaptive app scopes)

Sensitivity labels appear as an option in the **Content contains** condition list.



> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Supported items, file types, scenarios, and policy tips

You can use sensitivity labels as conditions on these items and in the following scenarios.

Supported items

Expand table

Service	Item type	Available to policy tip	Enforceable	Notes
Exchange	email message	Yes	Yes	
Exchange	email attachment	Yes	Yes	
SharePoint	items in SharePoint	Yes	Yes	
OneDrive	items	Yes	Yes	
Teams	Teams and channel messages	Not applicable	Not applicable	
Teams	attachments	Yes	Yes	If SharePoint or OneDrive are included as locations in your DLP policy, you don't need to set Teams as a location. This is because attachments sent in Teams over 1:1 chat or channels are automatically uploaded to OneDrive and SharePoint and will be evaluated by DLP then.
Devices	items	Yes	Yes	
MCAS (preview)	items	Yes	Yes	
Inline web traffic	Uploaded and downloaded files	No	Yes	DLP policies don't support custom policy tips for inline web traffic.

> **Note**
>
> DLP's ability to detect sensitivity labels in SharePoint and OneDrive is limited. For more information, see Enable sensitivity labels for files in SharePoint and OneDrive.

Supported file types

Expand table

Workload	File types supported
Exchange emails	.docx, .xlsx, .pptx, .pdf, .pfile
SharePoint	.docx, .xlsx, .pptx, .pdf
OneDrive	.docx, .xlsx, .pptx, .pdf
endpoint devices	.docx, .xlsx, .pptx, .pdf, .pfile
Inline web traffic	.docx, .xlsx, .pptx, .pdf, .pfile, .eml - legacy Office file types, .doc, .xls, and .ppt, aren't supported.

Supported scenarios

- DLP Admin can see a list of all sensitivity labels in the tenant when they choose to include one or more sensitivity labels as a condition.
- All workloads support using sensitivity labels as a condition, as shown in the support matrix.
- DLP policy tips continue to show across workloads for DLP policies that contain one or more sensitivity labels as a condition.
- Sensitivity labels appear as part of the incident report email if a DLP policy with one or more sensitivity labels as a condition is matched.
- Sensitivity label details show in the DLP rule match audit log for DLP policy matches that contain a sensitivity label as a condition.
- For policies that apply to the Inline web traffic location, content that's encrypted can't be extracted for evaluation by DLP. The only conditions that work are metadata based ones, for example:
- **Content contains** > **Sensitivity label** returns the label name
- **File size** returns the size of the encrypted file
- **File extension is** returns the extension of the encrypted file Conditions that rely on the content being evaluated don't work, for example:
- **Content contains** > **Sensitive information type**
- **Content contains** > **Trainable classifiers**
- **File type is**

Support policy tips

Expand table

Workload	Policy tips supported/not supported
OWA	supported
Outlook for Windows	supported
SharePoint	supported
OneDrive	supported
endpoint devices	not supported
Inline web traffic	not supported

> **Note**
>
> If you use sensitivity labels as a condition and policy tips in a DLP policy for Exchange, and, at the same time, you have mandatory labeling for sensitivity labels configured, when a user selects a label through the Mandatory Labeling prompt, the client does not re-invoke the DLP Policy Tips evaluation gate with the newly applied label. In result, the policy tip is **not ** going to show, but the action configured in the DLP policy will be enforced.

Last updated on 08/17/2026

Data loss prevention Exchange conditions and actions reference

Conditions in Microsoft Purview Data Loss Prevention (DLP) policies identify sensitive items that the policy is applied to. Actions define what happens as a consequence of a condition being met.

- Conditions define what to include
- Actions define what happens as a consequence of condition being met

Most conditions have one property that supports one or more values. For example, if the DLP policy is being applied to Exchange emails, the **The sender is** condition requires the sender of the message. Some conditions have two properties. For example, the **A message header includes any of these words** condition requires one property to specify the message header field, and a second property to specify the text to look for in the header field. Some conditions or exceptions don't have any properties. For example, the **Attachment is password protected** condition simply looks for attachments in messages that are password protected.

Actions typically require additional properties. For example, when the DLP policy rule redirects a message, you need to specify where the message is redirected to.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

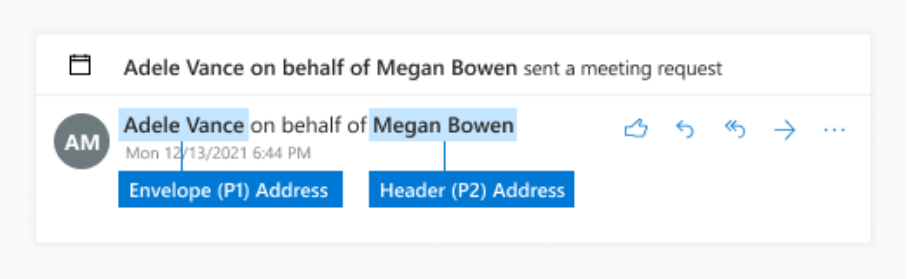
Exchange conditions for DLP policies

The tables in the following sections describe the conditions and exceptions that are available in DLP.

- Senders
- Recipients
- Message subject or body
- Attachments
- Message headers
- Message properties

Senders

If you use the sender address as a condition the actual field where the value is looked for varies depending on the sender address location configured. By default, DLP rules use the Header address as the sender address.



At the tenant level, you can configure a sender address location to be used across all rules, unless overridden by a single rule. To set tenant DLP policy configuration to evaluate the sender address from the Envelope across all rules, you can run the following command:

PowerShell

```
Set-PolicyConfig -SenderAddressLocation Envelope
```

To configure the sender address location at a DLP rule level, the parameter is *SenderAddressLocation*. The available values are:

- **Header**: Only examine senders in the message headers (for example, the **From**, **Sender**, or **Reply-To** fields). This is the default value.
- **Envelope**: Only examine senders from the message envelope (the **MAIL FROM** value that was used in the SMTP transmission, which is typically stored in the **Return-Path** field).
- **Header or envelope** (`HeaderOrEnvelope`) Examine senders in the message header and the message envelope.

Expand table

Condition or Exception in DLP	Condition/Exception parameters in Security & Compliance PowerShell	Property type	Description
Sender is	Condition: *From* Exception: *ExceptIfFrom*	*Addresses*	Messages sent by the specified mailboxes, mail users, mail contacts.
The sender is a member of	Condition: *FromMemberOf* Exception: *ExceptIfFromMemberOf*	*Addresses*	Messages sent by a member of the specified distribution group whether assigned or dynamic, mail-enabled security group, or Microsoft 365 group whether assigned or dynamic.
Sender IP address is	Condition: *SenderIPRanges* Exception: *ExceptIfSenderIPRanges*	*IPAddressRanges*	Messages where the sender's IP address matches the specified IP address, or falls within the specified IP address range.

Condition or Exception in DLP	Condition/Exception parameters in Security & Compliance PowerShell	Property type	Description
Sender address contains words	Condition: *FromAddressContainsWords* Exception: *ExceptIfFromAddressContainsWords*	*Words*	Messages that contain the specified words in the sender's email address.
Sender address matches patterns	Condition: *FromAddressMatchesPatterns* Exception: *ExceptIfFromAddressMatchesPatterns*	*Patterns*	Messages where the sender's email address contains text patterns that match the specified regular expressions.
Sender domain is	Condition: *SenderDomainIs* Exception: *ExceptIfSenderDomainIs*	*DomainName*	Messages where the domain of the sender's email address matches the specified value. If you need to find sender domains that *contain* the specified domain (for example, any subdomain of a domain), use **The sender address matches pattern** (*FromAddressMatchesPatterns*) condition and specify the domain by using the syntax: `'\.domain\.com$'`.
Sender scope	Condition: *FromScope* Exception: *ExceptIfFromScope*	*UserScopeFrom*	Messages sent by either internal or external senders.
The sender's specified properties include any of these words	Condition: *SenderADAttributeContainsWords* Exception: *ExceptIfSenderADAttributeContainsWords*	First property: *ADAttribute* Second property: *Words*	Messages where the specified Microsoft Entra ID attribute of the sender contains any of the specified words.
The sender's specified properties match these text patterns	Condition: *SenderADAttributeMatchesPatterns* Exception: *ExceptIfSenderADAttributeMatchesPatterns*	First property: *ADAttribute* Second property: *Patterns*	Messages where the specified Microsoft Entra ID attribute of the sender contains text patterns that match the specified regular expressions.

Recipients

When an email is sent to multiple recipients and the DLP policy rules allow only some of those emails to be delivered, the email might get *bifurcated*. For example, say that your DLP policy rules

allow emails to be sent to email addresses within your organization and blocks emails from being sent to external email addresses.

There are several policy conditions that cause bifurcation; that allowing an email to be sent to some users but not to others. For more information about bifurcation and the details about how bifurcation works, see the article on Bifurcation.

Expand table

Condition or Exception in DLP	Condition/Exception parameters in Security & Compliance PowerShell	Property type	Description	Bifurcating?
Recipient is	Condition: *SentTo* Exception: *ExceptIfSentTo*	*Addresses*	Messages where one of the recipients is the specified mailbox, mail user, or mail contact in the organization. The recipients can be in the **To**, **Cc**, or **Bcc** fields of the message.	Yes
Recipient domain is	Condition: *RecipientDomainIs* Exception: *ExceptIfRecipientDomainIs*	*DomainName*	Messages where the domain of the recipient's email address matches the specified value.	Yes
Recipient address contains words	Condition: *AnyOfRecipientAddressContainsWords* Exception: *ExceptIfAnyOfRecipientAddressContainsWords*	*Words*	Messages that contain the specified words in the recipient's email address. **Note**: This condition doesn't consider messages that are sent	No

Condition or Exception in DLP	Condition/Exception parameters in Security & Compliance PowerShell	Property type	Description	Bifurcating?
			to recipient proxy addresses. It only matches messages that are sent to the recipient's primary email address.	
Recipient address matches patterns	Condition: *AnyOfRecipientAddressMatchesPatterns* Exception: *ExceptIfAnyOfRecipientAddressMatchesPatterns*	*Patterns*	Messages where a recipient's email address contains text patterns that match the specified regular expressions. **Note**: This condition doesn't consider messages that are sent to recipient proxy addresses. It only matches messages that are sent to the recipient's primary email address.	No
Sent to member of	Condition: *SentToMemberOf* Exception: *ExceptIfSentToMemberOf*	*Addresses*	Messages that contain recipients who are members of the specified distribution group whether assigned or	Yes

Condition or Exception in DLP	Condition/Exception parameters in Security & Compliance PowerShell	Property type	Description	Bifurcating?
			dynamic, mail-enabled security group, or Microsoft 365 group whether assigned or dynamic. The group can be in the **To**, **Cc**, or **Bcc** fields of the message.	
The recipient's specified properties include any of these words	Condition: *RecipientADAttributeContainsWords* Exception: *ExceptIfRecipientADAttributeContainsWords*	First property: *ADAttribute* Second property: *Words*	Messages where the specified Microsoft Entra ID attribute of a recipient contains any of the specified words. Note that the **Country** attribute requires the two-letter country code value (for example, DE for Germany).	Yes
The recipient's specified properties match these text patterns	Condition: *RecipientADAttributeMatchesPatterns* *ExceptIfRecipientADAttributeMatchesPatterns*	First property: *ADAttribute* Second property: *Patterns*	Messages where the specified Entra ID attribute of a recipient contains text patterns that match the specified regular expressions.	Yes

Condition or Exception in DLP	Condition/Exception parameters in Security & Compliance PowerShell	Property type	Description	Bifurcating?
Recipient scope/Content is shared with	Condition: *AccessScope* Exception: *ExceptIfAccessScope*	*UserScopeFrom*	Messages that are received by either internal or external recipients.	Yes
Unique recipient count over	Condition: *RecipientCountOver*	*Number*	Messages where the count of recipients is greater than the specified value. Groups are expanded and final recipient count is compared.	No
Unique recipient domain over	Condition: *DomainCountOver*	*Number*	Messages where the count of recipients' domains is greater than the specified value.	No

> **Note**
>
> DLP supports only mail enabled security groups, distribution groups assigned, and dynamic, and Microsoft 365 Groups, both assigned and dynamic for sender and recipient based conditions that accept groups as input.

Message subject or body

Expand table

Condition or Exception in DLP	Condition/Exception parameters in Security & Compliance PowerShell	Property type	Description
Subject contains	Condition: *SubjectContainsWords* Exception: *ExceptIf SubjectContainsWords*	*Words*	Messages that have the specified words in the

Condition or Exception in DLP	Condition/Exception parameters in Security & Compliance PowerShell	Property type	Description
words or phrases			Subject field. This condition has limited support for non-english multibyte characters.
Subject matches patterns	Condition: *SubjectMatchesPatterns* Exception: *ExceptIf SubjectMatchesPatterns*	*Patterns*	Messages where the Subject field contains text patterns that match the specified regular expressions.
Content contains	Condition: *ContentContainsSensitiveInformation* Exception: *ExceptIfContentContainsSensitiveInformation*	*SensitiveInformationTypes*	Messages or documents that contain sensitive information as defined by Microsoft Purview Data Loss Prevention (DLP) policies.
Content is not labeled	Condition: *ContentIsNotLabeled* Exception:*ExceptIfContentIsNotLabeled*	*Sensitivity Labels*	Messages where neither the email nor the attached documents contain any sensitivity labels as defined by Microsoft Purview Data Loss Prevention (DLP) policies.
Subject or Body matches pattern	Condition: *SubjectOrBodyMatchesPatterns* Exception: *ExceptIfSubjectOrBodyMatchesPatterns*	*Patterns*	Messages where the subject field or message body contains text patterns that match the specified regular expressions.
Subject or Body contains words	Condition: *SubjectOrBodyContainsWords* Exception: *ExceptIfSubjectOrBodyContainsWords*	*Words*	Messages that have the specified words in the subject field or message body.

Condition or Exception in DLP	Condition/Exception parameters in Security & Compliance PowerShell	Property type	Description
			This condition has limited support for non-english multibyte characters.

Attachments

Expand table

Condition or Exception in DLP	Condition/Exception parameters in Security & Compliance PowerShell	Property type	Description
Attachment is password protected	Condition: *DocumentIsPasswordProtected* Exception: *ExceptIfDocumentIsPasswordProtected*	None	Messages where an attachment is password protected (and therefore can't be scanned). Password detection works for Office documents, compressed files (.zip, .7z), and .pdf files.
Attachment's file extension is	Condition: *ContentExtensionMatchesWords* Exception: *ExceptIfContentExtensionMatchesWords*	*Words*	Messages where an attachment's visible or the actual file extension matches any of the specified words. To detect renamed .7z files, you must configure configure the rule to protect files with the .7zip extension.
Any email attachment's content could not be scanned	Condition: *DocumentIsUnsupported* Exception: *ExceptIf DocumentIsUnsupported*	N/A	Messages where an attachment isn't recognized by Exchange Online.
Any email attachment's content didn't complete scanning	Condition: *ProcessingLimitExceeded* Exception: *ExceptIfProcessingLimitExceeded*	N/A	Messages where the rules engine couldn't complete the scanning of the attachments. You can use this condition to create rules that work together to identify and process messages where the content couldn't be fully scanned.
Document scan failures	Condition: *DocumentScanFailures*	N/A	Recommend using with **Document could not be scanned** or **Document didn't complete scanning** conditions to enforce actions on granular scanning error scenarios across

Condition or Exception in DLP	Condition/Exception parameters in Security & Compliance PowerShell	Property type	Description
			classification timeout, throttling or other failures.
Document name contains words	Condition: *DocumentNameMatchesWords* Exception: *ExceptIfDocumentNameMatchesWords*	*Words*	Messages where an attachment's file name matches any of the specified words that are delimited between the start of the name, any non-alphanumeric character that appears on the keyboard, or end of the name.
Document name matches patterns	Condition: *DocumentNameMatchesPatterns* Exception: *ExceptIfDocumentNameMatchesPatterns*	*Patterns*	Messages where an attachment's file name contains text patterns that match the specified regular expressions. This has been discontinued for SharePoint and OneDrive workloads. Existing rules can't be modified and new rules can't be created. Existing customers can continue to use this condition.
Document property is	Condition: *ContentPropertyContainsWords* Exception: *ExceptIfContentPropertyContainsWords*	*Words*	Messages with documents where an attachment's custom property matches the given value.
Document size equals or is greater than	Condition: *DocumentSizeOver* Exception: *ExceptIfDocumentSizeOver*	Size	Messages where any attachment is greater than or equal to the specified value.
Any attachment's content includes any of these words	Condition: *DocumentContainsWords* Exception: *ExceptIfDocumentContainsWords*	*Words*	Messages where an attachment contains the specified words.
Any attachments content matches these text patterns	Condition: *DocumentMatchesPatterns* Exception: *ExceptIfDocumentMatchesPatterns*	*Patterns*	Messages where an attachment contains text patterns that match the specified regular expressions.

Message Headers

Expand table

Condition or Exception in DLP	Condition/Exception parameters in Security & Compliance PowerShell	Property type	Description
Header contains words or phrases	Condition: *HeaderContainsWords* Exception: *ExceptIfHeaderContainsWords*	*Hash Table*	Messages that contain the specified header field, and the value of that header field contains the specified words.
Header matches patterns	Condition: *HeaderMatchesPatterns* Exception: *ExceptIfHeaderMatchesPatterns*	*Hash Table*	Messages that contain the specified header field, and the value of that header field contains the specified regular expressions.

Message properties

Expand table

Condition or Exception in DLP	Condition/Exception parameters in Security & Compliance PowerShell	Property type	Description
With importance	Condition: *WithImportance* Exception: *ExceptIfWithImportance*	*Importance*	Messages that are marked with the specified importance level.
Content character set contains words	Condition: *ContentCharacterSetContainsWords* Exception: *ExceptIfContentCharacterSetContainsWords*	*CharacterSets*	Messages that have any of the specified character set names.
Has sender override	Condition: *HasSenderOverride* Exception: *ExceptIfHasSenderOverride*	N/A	Messages where the sender has chosen to override a data loss prevention (DLP) policy. For more information, see Learn about data loss prevention
Message type matches	Condition: *MessageTypeMatches* Exception: *ExceptIfMessageTypeMatches*	*MessageType*	Messages of the specified type. **Note**: The available message types are Automatic reply, Auto-forward, Encrypted (S/MIME), Calendaring, Permission controlled (rights management), Voicemail, Signed, Read receipt, and Approval request.
The message size is greater than or equal to	Condition: *MessageSizeOver* Exception: *ExceptIfMessageSizeOver*	*Size*	Messages where the total size (message plus attachments) is greater than or equal to the specified value. **Note**: Message size limits on mailboxes are

Condition or Exception in DLP	Condition/Exception parameters in Security & Compliance PowerShell	Property type	Description
			evaluated before mail flow rules. A message that's too large for a mailbox will be rejected before a rule with this condition is able to act on the message.

Enhanced matched conditions in alerts and Activity Explorer (preview)

When an Exchange Online DLP policy matches a message, the DLP alert and the corresponding Activity Explorer event surface enhanced detail for every non-sensitive information type (SIT) condition that contributed to the match. This applies to all condition categories listed in the preceding sections: Senders, Recipients, Message subject or body, Attachments, Message headers, and Message properties. For the full list of supported conditions and detail levels, see Enhanced matched conditions for Exchange DLP events.

Actions for DLP policies

This table describes the actions that are available in DLP.

Expand table

Action in DLP	Action parameters in Security & Compliance PowerShell	Property type	Description
Restrict access or encrypt content in Microsoft 365 locations	BlockAccess	First property: *Boolean* Second property: *BlockAccessScope*	This allows you to block the access or encrypt the content to the specified users using RMS template
Set header	SetHeader	First property: *Header Name* Second property: *Header Value*	The `SetHeader` parameter specifies an action for th rule that adds or modifies a header field and value message header. This parameter uses the syntax `"HeaderName:HeaderValue"`. You can specify multipl header name and value pairs separated by comma
Remove header	RemoveHeader	First property: *MessageHeaderField* Second property: *String*	The `RemoveHeader` parameter specifies an action fo DLP rule that removes a header field from the mes header. This parameter uses the syntax `HeaderName` `"HeaderName:HeaderValue`. You can specify multiple

Action in DLP	Action parameters in Security & Compliance PowerShell	Property type	Description
			header names or header name and value pairs sep by commas
Redirect the message to specific users	RedirectMessageTo	*Addresses*	Redirects the message to the specified recipients. T message isn't delivered to the original recipients, a notification is sent to the sender or the original recipients.
Forward the message for approval to sender's manager	Moderate	First property: *ModerateMessageByManager* Second property: *Boolean* `$true`	The *Moderate* parameter specifies an action for the rule that sends the email message to a moderator user's manager or specified approvers). To forward message to the user's manager for approval, use th syntax: `@{ModerateMessageByManager = $true}`
Forward the message for approval to specific approvers	Moderate	First property: *ModerateMessageByManager* Second property: *Boolean* `$false` Third property: *ModerateMessageByUser* Fourth property: *Addresses*	The *Moderate* parameter specifies an action for the rule that sends the email message to a moderator user's manager or specified approvers). To forward message to specified recipients for approval, use th syntax: `@{ModerateMessageByManager = $false; ModerateMessageByUser = @("emailaddress1","emailaddress2",..."emailaddre`
Add recipient	AddRecipients	First property: *Field* Second property: *Addresses*	Adds one or more recipients to the **To/Cc/Bcc** fielc message. This parameter uses the syntax: `@{<AddToRecipients \<CopyTo | BlindCopyTo\> = "emailaddress"}`
Add the sender's manager as recipient	AddRecipients	First property: *AddedManagerAction* Second property: *Field*	Adds the sender's manager to the message as the specified recipient type (**To**, **Cc**, **Bcc**), or redirects th message to the sender's manager without notifyin sender or the recipient. This action only works if th sender's `Manager` attribute is defined in the Micros Entra ID. This parameter uses the syntax: `@{AddManagerAsRecipientType = "\<To | Cc | Bc`
Prepend subject	PrependSubject	*String*	Adds the specified text to the beginning of the **Sul** field of the message. Consider using a space or a c as the last character of the specified text to differer it from the original subject text. To prevent the same string from being added to messages that already contain the text in the subje example, replies), add the *The subject contains wor* (`ExceptIfSubjectContainsWords`) exception to the r

Action in DLP	Action parameters in Security & Compliance PowerShell	Property type	Description
Apply HTML disclaimer	ApplyHtmlDisclaimer	First property: *Text* Second property: *Location* Third property: *Fallback action*	Applies the specified HTML disclaimer to the requi location of the message. This parameter uses the syntax: `@{Text = " " ; Lo = \<Append | Prepend\>; FallbackAction = \<Wrap Ignore | Reject\>}`
Remove message encryption and rights protection	RemoveRMSTemplate	N/A	Removes message encryption applied on an email
Apply Branding to encrypted messages	ApplyBrandingTemplate	*String*	The `ApplyBrandingTemplate` parameter specifies an for the DLP rule that applies a custom branding ter for messages encrypted by Microsoft Purview Mes Encryption. You identify the custom branding temp name. If the name contains spaces, enclose the nar quotation marks (").
Make external recipients open mail in encrypted message portal	EnforcePortalAccess	*Boolean*	The `EnforcePortalAccess` parameter controls whet external users are required to use the encrypted m portal to view encrypted messages
Deliver the message to the hosted quarantine	*Quarantine*	n/a	Delivers the message to the quarantine in Microso For more information, see Quarantined email mess
Modify Subject	ModifySubject	*PswsHashTable*	Remove text from the subject line that matches a s pattern and replace it with different text. See the e below. You can: - **Replace** all matches in the subject with the replac text - **Append** to remove all matches in the subject anc inserts the replacement text at the end of the subje - **Prepend** to remove all matches and inserts the replacement text at the beginning of the subject. F more information, see the ModifySubject paramete description in the New-DlpComplianceRule referer article.

Last updated on 06/15/2026

Data loss prevention policy tips reference

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

DLP Policy Tips support matrix for Outlook clients highlights

This short table highlights some common DLP support for policy tips across Outlook clients. For complete information, see:

- Data loss prevention policy tip reference for Outlook for Microsoft 365
- Data loss prevention policy tip reference for Outlook on the Web
- Data loss prevention policy tip reference for new Outlook for Windows

Expand table

Feature	Outlook for Microsoft 365 (ver. 2105 build 14026.20000 and semi-annual channel ver. 2102 build 13801.20862)	Outlook On the Web	new Outlook for Windows
Policy tips supported	Yes	Yes	Yes
Top 10 conditions - Content is shared from M365 - Content contains SITs - Content contains sensitivity label - Subject or Body contains words or phrases - Sender is - Sender is a member of - Sender domain is - Recipient is - Recipient domain is - Recipient is a member of	Yes	Yes	Yes

Feature	Outlook for Microsoft 365 (ver. 2105 build 14026.20000 and semi-annual channel ver. 2102 build 13801.20862)	Outlook On the Web	new Outlook for Windows
Default Oversharing Dialog	Yes	No	Yes
Customized Oversharing Dialog	Yes	No	No
Wait on Send dialog support for Oversharing for Outlook for Microsoft 365	Yes	No	No
Microsoft provided sensitive information types (SIT)	Yes	Yes	Yes
Custom SITs	Yes	Yes	Yes
Exact data match SITs	Yes	Yes	Yes
Custom compliance URL to further educate end users	Yes	No	No

Support Matrix for DLP policy tips across Microsoft apps

Expand table

App and platform	DLP policy tip support	Sensitive information types supported	Conditions and actions supported	Comments
Outlook On the Web	✓	Subset. See Sensitive information types that support policy tips in Outlook on the Web	Subset. See Conditions that support policy tips in Outlook on the Web and Actions that support policy tips in Outlook on the Web	For full details, see Data loss prevention policy tip reference for Outlook on the Web
Outlook for Microsoft 365 (ver. 2105 build	✓	Varies by version and	Subset. See Data loss	For full details, see Data loss prevention policy tip reference for

App and platform	DLP policy tip support	Sensitive information types supported	Conditions and actions supported	Comments
14026.20000 and semi-annual channel ver. 2102 build 13801.20862)		license. See Data loss prevention policy tip reference for Outlook for Microsoft 365	prevention policy tip reference for Outlook for Microsoft 365	Outlook for Microsoft 365
Outlook Mobile (iOS, Android)	✓	Varies by version and license. See Data loss prevention policy tip reference for Outlook for Microsoft 365	Subset. See Data loss prevention policy tip reference for Outlook for Microsoft 365	For full details, see Data loss prevention policy tip reference for Outlook for Android, iOS, and macOS
Outlook Mac	✓	Varies by version and license. See Data loss prevention policy tip reference for Outlook for Microsoft 365	Subset. See Data loss prevention policy tip reference for Outlook for Microsoft 365	For full details, see Data loss prevention policy tip reference for Outlook for Android, iOS, and macOS
SharePoint/OneDrive Web client	✓	All.	All.	See Data loss prevention policy tip reference for SharePoint and OneDrive web client
SharePoint Win32/ OneDrive Win32 client	✗	None	None	DLP policy tips aren't supported on SharePoint or OneDrive desktop client apps
Word, Excel, PowerPoint Web Client	✓	All	All SPO/ODB predicates and actions in DLP	DLP policy tip s are supported if the document is hosted on SharePoint or the OneDrive web app and the DLP policy is already stamped.
Word, Excel, PowerPoint Mobile Client	✗	None	None	DLP policy tips aren't supported in mobile apps for Office.

App and platform	DLP policy tip support	Sensitive information types supported	Conditions and actions supported	Comments
Teams Web/ Teams Desktop/ Teams Mobile/ Teams Mac	✓	All	All Teams predicates in DLP policy	Policy tips will show when a message is flagged as "This message has been flagged. What can I do?" When clicking the link, the user can review the sensitive info types detected and override or report an issue if allowed by the admin. No policy tips are shown for files. When the recipient tries to access the document, they might get access denied if not allowed.
Win32 Endpoint Devices	✓	Subset	All Endpoint DLP predicates and actions in DLP policy	For more information, see Data Loss Prevention on Endpoint supports policy tips for only some sensitive information types
macOS devices	✓	All	DLP policy tips are supported for all actions that are supported for macOS.	For more information, see Endpoint DLP Windows 10/11 and macOS support
3rd party cloud apps	✗	None	None	DLP policy tips aren't supported on third party cloud apps
On-prem	✗	None	None	
Word, Excel, PowerPoint Win32 Client	✗	Subset	Subset	See Data loss prevention policy tip reference for Outlook for Microsoft 365 for the list of sensitive information types supported Policy tips for WXP client apps will work for documents stored on SharePoint or OneDrive Sites for all DLP policies that have exactly the below or a subset of conditions or actions in the DLP policy: • Content contains sensitive information types

App and platform	DLP policy tip support	Sensitive information types supported	Conditions and actions supported	Comments
				• Access Scope (Content is shared internally/externally) • Notify User (policy tips/user notifications) • Block everyone • Incident reports If any other condition or action is present, the DLP policy tip for that policy won't appear in the desktop apps of Word, Excel or PowerPoint. See Policy tips in Excel, PowerPoint, and Word for more details
Fabric and Power BI	✓	- Sensitivity labels - Sensitive information types (excluding exact data match sensitive information types)	Subset	Data loss prevention policies in Power BI and Fabric are generally available. Policy tips and admin alerts are supported.

Data loss prevention on endpoint devices supports policy tips for only some sensitive information types

These sensitive information types (SIT)s can be detected on endpoint devices:

- ABA routing number
- Argentina national identity (DNI) number
- Australia bank account number
- Australia medical account number
- Australia passport number
- Australia tax file number
- Australia business number
- Australia Company Number

- Austria Driver's License Number
- Austria Identity Card
- Austria Passport Number
- Austria Social Security Number
- Austria Tax Identification Number
- Austria value added tax
- Azure DocumentDB auth key
- Azure IAAS database connection string and Azure SQL connection string
- Azure IoT connection string
- Azure publish setting password
- Azure Redis cache connection string
- Azure SAS
- Azure service bus connection string
- Azure storage account key
- Azure Storage account key (generic)
- Azure Storage account Shared Access Signature
- Belgium drivers license number
- Belgium National Number
- Belgium passport number
- Belgium value added tax number
- Brazil CPF number
- Brazil legal entity number (CNPJ)
- Brazil national identification card (RG)
- Bulgaria drivers license number
- Bulgaria passport number
- Bulgaria Uniform Civil Number
- Canada bank account number
- Canada drivers license number
- Canada health service number
- Canada Passport Number
- Canada personal health identification number (PHIN)
- Canada Social Insurance Number
- Chile Identity Card Number
- China resident identity card (PRC) number
- Credit Card Number
- Croatia Driver's License Number
- Croatia Identity Card Number

- Croatia Passport Number
- Croatia Personal Identification (OIB) Number
- General Symmetric Key
- Cyprus Driver's License Number
- Cyprus Identity Card
- Cyprus Passport Number
- Cyprus Tax Identification Number
- Czech drivers license number
- Czech Personal Identity Number
- Czech passport number
- Denmark Driver's License Number
- Denmark Passport Number
- Denmark Personal Identification Number
- Drug Enforcement Agency (DEA) number
- Estonia drivers license number
- Estonia Passport Number
- Estonia Personal Identification Code
- EU Debit Card Number
- EU Driver's License Number
- EU National Identification Number
- EU Passport Number
- EU Social Security Number (SSN) or Equivalent ID
- EU Tax Identification Number (TIN)
- Finland Driver's License Number
- Finland European Health Insurance Number
- Finland National ID
- Finland Passport Number
- France Driver's License Number
- France Health Insurance Number
- France National ID Card (CNI)
- France Passport Number
- France Social Security Number (INSEE)
- France Tax Identification Number (numéro SPI.)
- France Value Added Tax Number
- German Driver's License Number
- German Passport Number
- German Identity Card Number

- German Tax Identification Number
- German Value Added Tax Number
- Greece Driver's License Number
- Greece National ID Card
- Greece Passport Number
- Greece Social Security Number (AMKA)
- Greek Tax identification Number
- Hong Kong Identity Card (HKID) Number
- Hungary Social Security Number (TAJ)
- Hungary Value Added Tax Number
- Hungary Driver's License Number
- Hungary Passport Number
- Hungary Personal Identification Number
- Hungary Tax identification Number
- India Permanent Account Number (PAN)
- India Unique Identification (Aadhaar) Number
- Indonesia Identity Card (KTP) Number
- International Banking Account Number (IBAN)
- International Classification of Diseases (ICD-10-CM)
- International Classification of Diseases (ICD-9-CM)
- IP Address
- Ireland Driver's License Number
- Ireland Passport Number
- Ireland Personal Public Service (PPS) Number
- Israel Bank Account Number
- Israel National ID
- Italy Driver's License Number
- Italy Fiscal Code
- Italy Passport Number
- Italy Value Added Tax Number
- Japan Bank Account Number
- Japan Driver's License Number
- Japan Passport Number
- Japan Resident Registration Number
- Japan Social Insurance Number (SIN)
- Japan My Number Corporate
- Japan My Number Personal

- Japan Residence Card Number
- Latvia Driver's License Number
- Latvia Passport Number
- Latvia Personal Code
- Lithuania Driver's License Number
- Lithuania Passport Number
- Lithuania Personal Code
- Luxemburg drivers license number
- Luxemburg national identification number (natural persons)
- Luxemburg national identification number (non-natural persons)
- Luxemburg passport number
- Malaysia Identity Card Number
- Malta Driver's License Number
- Malta Identity Card Number
- Malta Passport Number
- Malta Tax ID Number
- Netherlands Citizen's Service (BSN) Number
- Netherlands Driver's License Number
- Netherlands Passport Number
- Netherlands Tax Identification Number
- Netherlands Value Added Tax Number
- New Zealand bank account Number
- New Zealand Driver License Number
- New Zealand Inland Revenue Number
- New Zealand Ministry of Health Number
- New Zealand Social Welfare Number
- Norway Identity Number
- Philippines Unified Multi-Purpose ID Number
- Poland Driver's License Number
- Poland Identity Card
- Poland National ID (PESEL)
- Poland Passport
- Poland Tax Identification Number
- Polish REGON Number
- Portugal Citizen Card Number
- Portugal Driver's License Number
- Portugal Passport Number

- Portugal Tax Identification Number
- Romania Driver's License Number
- Romania Passport Number
- Romania Personal Numerical Code (CNP)
- Russian Passport Number (Domestic)
- Russian Passport Number (International)
- Saudi Arabia National ID
- Singapore national registration identity card (NRIC) number
- Slovakia Driver's License Number
- Slovakia Passport Number
- Slovakia Personal Number
- Slovenia Driver's License Number
- Slovenia Passport Number
- Slovenia Tax Identification Number
- Slovenia Unique Master Citizen Number
- South Africa Identification Number
- South Korea Resident Registration Number
- Spain DNI
- Spain Driver's License Number
- Spain Passport Number
- Spain Social Security Number (SSN)
- Spain Tax Identification Number
- SQL Server Connection String
- Sweden Driver's License Number
- Sweden National ID
- Sweden Passport Number
- Sweden Tax Identification Number
- SWIFT Code
- Swiss Social Security Number AHV
- Taiwan National ID
- Taiwan Passport Number
- Taiwan Resident Certificate (ARC/TARC)
- Thai Population Identification Code
- Turkish National Identification number
- U.K. Driver's License Number
- U.K. Electoral Roll Number
- U.K. National Health Service Number

- U.K. National Insurance Number (NINO)
- U.K. Unique Taxpayer Reference Number
- U.S./U.K. passport number
- U.S. bank account number
- U.S. drivers license number
- U.S. individual taxpayer identification number (ITIN)
- U.S. social security number (SSN)
- Ukraine passport domestic
- Ukraine passport international

Custom sensitive information types will also be detected in addition to the above out-of-the-box sensitive information types.

Last updated on 08/19/2026

Data loss prevention policy tip reference for Outlook for Microsoft 365

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Mailboxes must be hosted in Exchange Online. For more information, see: Learn about data loss prevention.

Microsoft Purview Data Loss Prevention (DLP) will only process the first 4 MB on message content for policy tip in Outlook for Microsoft 365 and only classify up to 2 MB of attachments.

> **Important**
>
> Policy Tips won't be displayed in the classic Outlook client for DLP policies scoped to nonmail enabled security groups but rule actions are enforced.

Licensing

For information on licensing, see

- Microsoft 365 Enterprise Plans
- Microsoft 365 Service Descriptions

Conditions that support policy tips for Outlook perpetual users

Expand table

For Outlook perpetual versions and users	These conditions apply
-E3 users -E5 users	- For email and unencrypted Microsoft 365 files **Content contains built-in/custom sensitive information types** - **Content is shared from Microsoft 365**

Conditions that support policy tips for Outlook for Microsoft 365 users

Expand table

For Outlook for Microsoft versions and users	These conditions apply
- All E3 users - All offline E5 users - All E5 users with connected experience disabled - All online E5 users with production version builds lower than 2303 (Build 16.0.16216.10000) - All online E5 users with Semi-Annual Enterprise Channel version builds lower than 2302 (Build 16.0.16130.20478)	- Content contains built-in/custom sensitive information types - Content is shared from Microsoft 365
- All online E5 users with connected experience enabled in WW commercial and GCC/GCC-H/DoD clouds - production version 2303 & Build 16.0 16216.10000 or higher - Semi-Annual Enterprise Channel version 2302 & Build 16.0.16130.20478 or higher - Users without a Microsoft 365 E5 license must be assigned a license that includes the Information Protection for Office 365 - Premium service plan to support DLP Policy Tips in Outlook Classic.	- Content contains built-in/custom sensitive information types (works for email and unencrypted Microsoft 365 and PDF files) - Message (includes email subject) contains built-in/custom sensitive information types - Attachment contains built-in/custom sensitive information types - Content contains sensitivity labels (Works for email and Office & PDF file types) - Content is shared - Sender is - Sender is member of (Only Distribution lists, Azure-based Dynamic Distribution groups, and email-enabled Security groups are supported.) - Sender domain is - Recipient is - Recipient is a member of (Only Distribution lists, Azure-based Dynamic Distribution groups, and email-enabled Security groups are supported.) - Recipient domain is - Subject contains words - Content isn't labeled - Message isn't labeled - Attachment isn't labeled - File extension is

Actions that support policy tips

All Exchange actions support policy tips

- Restrict access or encrypt the content in Microsoft 365 locations
- Set headers
- Remove header
- Redirect the message to specific users
- Forward the message for approval to sender's manager
- Forward the message for approval to specific approvers
- Add recipient to the To box
- Add recipient to the Cc box
- Add recipient to the Bcc box
- Add the sender's manager as recipient
- Removed O365 Message Encryption and rights protection
- Prepend Email Subject
- Add HTML Disclaimer
- Modify Email Subject
- Deliver the message to the hosted quarantine

Sensitive information types that support policy tips for Outlook perpetual users

For Outlook perpetual version E3 and E5 users these built-in sensitive information types and custom sensitive information types support policy tips:

- ABA routing number
- Argentina national identity (DNI) number
- Australia bank account number
- Australia medical account number
- Australia passport number
- Australia tax file number
- Azure DocumentDB auth key
- Azure IAAS database connection string and Azure SQL connection string
- Azure IoT connection string
- Azure publish setting password
- Azure Redis cache connection string
- Azure SAS
- Azure service bus connection string
- Azure storage account key
- Azure Storage account key (generic)

- Belgium national number
- Brazil CPF number
- Brazil legal entity number (CNPJ)
- Brazil national identification card (RG)
- Canada bank account number
- Canada driver's license number
- Canada health service number
- Canada passport number
- Canada personal health identification number (PHIN)
- Canada social insurance number
- Chile identity card number
- China resident identity card (PRC) number
- Credit card number
- Croatia identity card number
- Croatia personal identification (OIB) number
- Czech personal identity number
- Denmark personal identification number
- Drug Enforcement Agency (DEA) number
- EU debit card number
- EU driver's license number
- EU national identification number
- EU passport number
- EU social security number or equivalent identification
- EU Tax identification number
- Finland national ID
- Finland passport number
- France driver's license number
- France national id card (CNI)
- France passport number
- France social security number (INSEE)
- Germany driver's license number
- Germany passport number
- Germany identity card number
- Greece national ID card
- Hong Kong identity card (HKID) number
- India permanent account number (PAN)
- India unique identification (Aadhaar) number

- Indonesia identity card (KTP) number
- International banking account number (IBAN)
- International classification of diseases (ICD-10-CM)
- International classification of diseases (ICD-9-CM)
- IP address
- Ireland personal public service (PPS) number
- Israel bank account number
- Israel national identification number
- Italy driver's license number
- Japan bank account number
- Japan driver's license number
- Japan passport number
- Japan resident registration number
- Japan social insurance number (SIN)
- Japan residence card number
- Malaysia identification card number
- Netherlands citizens service (BSN) number
- New Zealand ministry of health number
- Norway identification number
- Philippines unified multi-purpose identification number
- Poland identity card
- Poland national ID (PESEL)
- Poland passport number
- Portugal citizen card number
- Saudi Arabia National ID
- Singapore national registration identity card (NRIC) number
- South Africa identification number
- South Korea resident registration number
- Spain social security number (SSN)
- SQL Server connection string
- Sweden national ID
- Sweden passport number
- SWIFT code
- Taiwan national identification number
- Taiwan passport number
- Taiwan-resident certificate (ARC/TARC) number
- Thai population identification code

- Turkey national identification number
- U.K. drivers license number
- U.K. electoral roll number
- U.K. national health service number
- U.K. national insurance number (NINO)
- U.S./U.K. passport number
- U.S. bank account number
- U.S. driver's license number
- U.S. individual taxpayer identification number (ITIN)
- U.S. social security number (SSN)

Sensitive information types that support policy tips for Outlook for Microsoft 365 users

For these Outlook versions and users

- All Microsoft 365 for Enterprise E3 users
- All Microsoft 365 for Enterprise offline E5 users
- All Microsoft 365 for Enterprise E5 users with connected experience disabled
- All Microsoft 365 for Enterprise online E5 users with production version lower than 2303 (Build 16.0.16216.10000)
- All Microsoft 365 for Enterprise online E5 users with Semi-Annual Enterprise Channel version lower than 2302 (Build 16.0.16130.20478)

These OOB sensitive information types support policy tips along with any custom and advanced sensitive information types

- ABA routing number
- Argentina national identity (DNI) number
- Australia bank account number
- Australia medical account number
- Australia passport number
- Australia tax file number
- Azure DocumentDB auth key
- Azure IAAS database connection string and Azure SQL connection string
- Azure IoT connection string
- Azure publish setting password
- Azure Redis cache connection string

- Azure SAS
- Azure service bus connection string
- Azure storage account key
- Azure Storage account key (generic)
- Belgium national number
- Brazil CPF number
- Brazil legal entity number (CNPJ)
- Brazil national identification card (RG)
- Canada bank account number
- Canada driver's license number
- Canada health service number
- Canada passport number
- Canada personal health identification number (PHIN)
- Canada social insurance number
- Chile identity card number
- China resident identity card (PRC) number
- Credit card number
- Croatia identity card number
- Croatia personal identification (OIB) number
- Czech personal identity number
- Denmark personal identification number
- Drug Enforcement Agency (DEA) number
- EU debit card number
- EU driver's license number
- EU national identification number
- EU passport number
- EU social security number or equivalent identification
- EU Tax identification number
- Finland national ID
- Finland passport number
- France driver's license number
- France national ID card (CNI)
- France passport number
- France social security number (INSEE)
- Germany driver's license number
- Germany passport number
- Germany identity card number

- Greece national ID card
- Hong Kong identity card (HKID) number
- India permanent account number (PAN)
- India unique identification (Aadhaar) number
- Indonesia identity card (KTP) number
- International banking account number (IBAN)
- International classification of diseases (ICD-10-CM)
- International classification of diseases (ICD-9-CM)
- IP address
- Ireland personal public service (PPS) number
- Israel bank account number
- Israel national identification number
- Italy driver's license number
- Japan bank account number
- Japan driver's license number
- Japan passport number
- Japan resident registration number
- Japan social insurance number (SIN)
- Japan residence card number
- Malaysia identification card number
- Netherlands citizens service (BSN) number
- New Zealand ministry of health number
- Norway identification number
- Philippines unified multi-purpose identification number
- Poland identity card
- Poland national ID (PESEL)
- Poland passport number
- Portugal citizen card number
- Saudi Arabia National ID
- Singapore national registration identity card (NRIC) number
- South Africa identification number
- South Korea resident registration number
- Spain social security number (SSN)
- SQL Server connection string
- Sweden national ID
- Sweden passport number
- SWIFT code

- Taiwan national identification number
- Taiwan passport number
- Taiwan-resident certificate (ARC/TARC) number
- Thai population identification code
- Turkey national identification number
- U.K. drivers license number
- U.K. electoral roll number
- U.K. national health service number
- U.K. national insurance number (NINO)
- U.S./U.K. passport number
- U.S. bank account number
- U.S. driver's license number
- U.S. individual taxpayer identification number (ITIN)
- U.S. social security number (SSN)

For these Outlook versions and users:

- Online E5 users with connected experience enabled.
- Production version 2303 (Build 16.0.16216.10000) or higher.
- Semi-Annual Enterprise Channel version 2302 (Build 16.0.16130.20478) or higher.

These sensitive information types support policy tips:

- ABA Routing Number
- All Credential Types
- All full names
- All medical terms and conditions
- All Physical Addresses
- Amazon S3 Client Secret Access Key
- Argentina national identity (DNI) number
- Argentina Unique Tax Identification Key (CUIT/CUIL)
- ASP.NET machine Key
- Australia bank account number
- Australian Business Number
- Australian Company Number
- Australia Driver's License Number
- Australia medical account number
- Australia passport number
- Australia physical addresses

- Australia tax file number
- Austria Driver's License Number
- Austria identity card
- Austria passport number
- Austria physical addresses
- Austria social security number
- Austria tax identification number
- Austria Value Added Tax (VAT) Number
- Azure App Service deployment password
- Azure Batch shared access key
- Azure Bot Framework secret key
- Azure Bot service app secret
- Azure Cognitive Search API key
- Azure Cognitive Service key
- Azure Container Registry access key
- Azure Cosmos DB account access key
- Azure Databricks personal access token
- Azure DevOps app secret
- Azure DevOps personal access token
- Azure DocumentDB auth key
- Azure EventGrid access key
- Azure Function Master / API key
- Azure IAAS database connection string and Azure SQL connection string
- Azure IoT connection string
- Azure IoT shared access key
- Azure Logic app shared access signature
- Azure Machine Learning web service API key
- Azure Maps subscription key
- Azure publish setting password
- Azure Redis cache connection string
- Azure Redis cache connection string password
- Azure SAS
- Azure service bus connection string
- Azure service bus shared access signature
- Azure Shared Access key / Web Hook token
- Azure SignalR access key
- Azure SQL connection string

- Azure storage account access key
- Azure storage account key
- Azure Storage account key (generic)
- Azure Storage account shared access signature
- Azure Storage account shared access signature for high risk resources
- Azure subscription management certificate
- Belgium driver's license number
- Belgium national number
- Belgium passport number
- Belgium physical addresses
- Belgium value added tax number
- Blood test terms
- Brand medication names
- Brazil CPF number
- Brazil legal entity number (CNPJ)
- Brazil National ID Card (RG)
- Brazil physical addresses
- Bulgaria driver's license number
- Bulgaria passport number
- Bulgaria physical addresses
- Bulgaria uniform civil number
- Canada bank account number
- Canada driver's license number
- Canada health service number
- Canada passport number
- Canada personal health identification number (PHIN)
- Canada physical addresses
- Canada social insurance number
- Chile identity card number
- China resident identity card (PRC) number
- China physical addresses
- Client secret / API key
- Credit card number
- Colombia national ID
- Colombia tax identification number
- Croatia driver's license number
- Croatia identity card number

- Croatia passport number
- Croatia personal identification (OIB) number
- Croatia physical addresses
- Cyprus Driver's License Number
- Cyprus identity card
- Cyprus passport number
- Cyprus physical addresses
- Cyprus tax identification number
- Czech driver's license number
- Czech passport number
- Czech personal identity number
- Czech Republic physical addresses
- Denmark driver's license number
- Denmark passport number
- Denmark personal identification number
- Denmark physical addresses
- Diseases
- Drug Enforcement Agency (DEA) number
- Ecuador Unique Identification Number
- Estonia driver's license number
- Estonia passport number
- Estonia Personal Identification Code
- Estonia physical addresses
- EU debit card number
- EU driver's license number
- EU national identification number
- EU passport number
- EU Social Security Number (SSN) or Equivalent ID
- EU Tax Identification Number (TIN)
- Finland driver's license number
- Finland European health insurance number
- Finland national ID
- Finland passport number
- Finland physical addresses
- France driver's license number
- France health insurance number
- France national id card (CNI)

- France passport number
- France physical addresses
- France social security number (INSEE)
- France Tax Identification Number (numéro SPI.)
- France value added tax number
- General password
- General Symmetric key
- Generic medication names
- German Driver's License Number
- Germany identity card number
- German Passport Number
- Germany physical addresses
- Germany tax identification number
- Germany value added tax number
- GitHub Personal Access Token
- Google API key
- Greenland physical addresses
- Greece driver's license number
- Greece national ID card
- Greece passport number
- Greece physical addresses
- Greece Social Security Number (AMKA)
- Greek Tax Identification Number
- Hong Kong identity card (HKID) number
- Http authorization header
- Hungary driver's license number
- Hungary passport number
- Hungary personal identification number
- Hungary physical addresses
- Hungarian Social Security Number (TAJ)
- Hungary tax identification number
- Hungarian Value Added Tax Number
- Iceland physical addresses
- Impairments Listed In The U.S. Disability Evaluation Under Social Security
- India driver's License Number
- India GST Number
- India permanent account number (PAN)

- India unique identification (Aadhaar) number
- India Voter Id Card
- Indonesia Drivers License Number
- Indonesia identity card (KTP) number
- Indonesia passport number
- International banking account number (IBAN)
- International classification of diseases (ICD-10-CM)
- International classification of diseases (ICD-9-CM)
- IP address
- IP Address v4
- IP Address v6
- Ireland driver's license number
- Ireland passport number
- Ireland personal public service (PPS) number
- Ireland physical addresses
- Israel bank account number
- Israel National ID
- Italy driver's license number
- Italy fiscal code
- Italy passport number
- Italy physical addresses
- Italy value added tax number
- Japan bank account number
- Japan driver's license number
- Japanese My Number Corporate
- Japanese My Number Personal
- Japan passport number
- Japan physical addresses
- Japanese Residence Card Number
- Japan resident registration number
- Japan social insurance number (SIN)
- Lab test terms
- Latvia driver's license number
- Latvia passport number
- Latvia personal code
- Latvia physical addresses
- Liechtenstein physical addresses

- Lifestyles that relate to medical conditions
- Lithuania driver's license number
- Lithuania passport number
- Lithuania personal code
- Lithuania physical addresses
- Luxembourg Driver's License Number
- Luxembourg National Identification Number (Natural persons)
- Luxembourg National Identification Number (Non-natural persons)
- Luxembourg Passport Number
- Luxembourg Physical Addresses
- Malaysia Identity Card Number
- Malaysia passport number
- Malta driver's license number
- Malta identity card number
- Malta passport number
- Malta physical addresses
- Malta Tax ID Number
- Medical Specialities
- Medicare Beneficiary Identifier (MBI) card
- Mexico Unique Population Registry Code (CURP)
- Microsoft Bing maps key
- Microsoft Entra client access token
- Microsoft Entra client secret
- Microsoft Entra user Credentials
- Netherlands citizen's service (BSN) number
- Netherlands driver's license number
- Netherlands passport number
- Netherlands physical addresses
- Netherlands tax identification number
- Netherlands value added tax number
- New Zealand bank account number
- New Zealand Driver License Number
- New Zealand inland revenue number
- New Zealand ministry of health number
- New Zealand physical addresses
- New Zealand social welfare number
- Norway Identity Number

- Norway physical addresses
- Philippines National ID
- Philippines passport number
- Philippines Unified Multi-Purpose ID Number
- Poland driver's license number
- Poland identity card
- Poland national ID (PESEL)
- Poland Passport Number
- Poland physical addresses
- Poland REGON number
- Poland tax identification number
- Portugal citizen card number
- Portugal driver's license number
- Portugal passport number
- Portugal physical addresses
- Portugal tax identification number
- Qatari ID Card Number
- Romania driver's license number
- Romania passport number
- Romania personal numeric code (CNP)
- Romania physical addresses
- Russian Passport Number (Domestic)
- Russian Passport Number (International)
- Russia physical addresses
- Russia taxpayer identification number
- Saudi Arabia National ID
- Singapore driver's license number
- Singapore passport number
- Singapore national registration identity card (NRIC) number
- Singapore physical addresses
- Slack access token
- Slovakia driver's license number
- Slovakia passport number
- Slovakia personal number
- Slovakia physical addresses
- Slovenia driver's license number
- Slovenia passport number

- Slovenia physical addresses
- Slovenia tax identification number
- Slovenia Unique Master Citizen Number
- South Africa identification number
- South Africa physical addresses
- South Korea driver's license number
- South Korea passport number
- South Korea resident registration number
- Spain DNI
- Spain driver's license number
- Spain passport number
- Spain physical addresses
- Spain social security number (SSN)
- Spain tax identification number
- SQL Server connection string
- Surgical procedures
- Sweden driver's license number
- Sweden national ID
- Sweden passport number
- Sweden physical addresses
- Sweden tax identification number
- SWIFT code
- Switzerland physical addresses
- Swiss Social Security Number AHV
- Taiwan National ID
- Taiwan passport number
- Taiwan Resident Certificate (ARC/TARC)
- Thai population identification code
- Turkey national identification number
- Turkey physical addresses
- Types of medication
- U.A.E. identity card number
- U.A.E. passport number
- U.K. driver's license number
- U.K. electoral roll number
- U.K. national health service number
- U.K. national insurance number (NINO)

- U.K. physical addresses
- U.K. Unique Taxpayer Reference Number
- U.S. bank account number
- U.S. driver's license number
- U.S. individual taxpayer identification number (ITIN)
- U.S. physical addresses
- U.S. social security number (SSN)
- U.S./U.K. passport number
- Ukraine Passport Number (Domestic)
- Ukraine Passport Number (International)
- Ukraine physical addresses
- User login credentials
- X.509 certificate private key

Exact Data Match sensitive information types that support policy tips Outlook for Microsoft 365

Applies to

- Online E5 users with connected experience enabled
- Production version 2303 (Build 16.0.16216.10000) or higher.
- Semi-Annual Enterprise Channel version 2302 (Build 16.0.16130.20478) or higher.

For more information on exact data match based SITs, see Learn about exact data match based sensitive information types.

Trainable classifiers support for Outlook for Microsoft 365

> **Important**
>
> Trainable classifiers aren't supported for the following conditions:
>
> - For **new** conditions:
> - `Message contains`
> - `Attachment contains`

- For the **existing** `Content contains` condition when used in combination with `Evaluate rule per component` condition. (Preview limitation)

Applies to:

For these Outlook versions and users:

- Online E5 users with connected experience enabled
- Production version 2303 (Build 16.0.16216.10000) or higher.
- Semi-Annual Enterprise Channel version 2302 (Build 16.0.16130.20478) or higher.

These trainable classifiers are supported:

- Adult, racy, and gory images
- Agreements
- Bank statement
- Budget
- Business plan
- Completion certificates
- Construction specifications
- Corporate sabotage
- Customer complaints
- Discrimination
- Employee disciplinary action
- Employee insurance
- Employment agreement
- Employee Pension Records
- Employee Stocks and Financial Bond Records
- Enterprise Risk Management
- Finance
- Financial audit
- Financial statement
- Freight Documents
- Gifts & entertainment
- Harassment
- Health/Medical forms
- Healthcare
- Human resources

- Invoice
- Intellectual property
- Information technology
- Legal affairs
- Legal Agreements
- License agreement
- Loan agreements and offer letters
- Merger and acquisition files
- Manufacturing batch records
- Meeting notes
- Money laundering
- Network design files
- Non-disclosure agreement
- Paystub
- Personal Financial Information
- Procurement
- Project documents
- Profanity
- Quotation
- Regulatory collusion
- Resume
- Safety records
- Sales and revenue
- Software product development files
- Source code
- Standard Operating Procedures and Manuals
- Statement of Accounts
- Statement of work
- Stock manipulation
- Tax documents
- Threat
- Unauthorized disclosure
- Wire Transfer
- Word count requirements

For more information on trainable classifiers, see Learn about trainable classifiers.

Sensitivity label support for Outlook for Microsoft 365

Applies to:

- Online E5 users with connected experience enabled
- Production version 2303 (Build 16.0.16216.10000) or higher.
- Semi-Annual Enterprise Channel version 2302 (Build 16.0.16130.20478) or higher.

For more information on sensitivity labels, see Learn about sensitivity labels.

Oversharing dialog for Outlook for Microsoft 365

The oversharing dialog is available in DLP for Outlook desktop for E5 users. When enabled in a DLP rule, this feature displays popups for warning, override, or block actions to end users who are sharing labeled or sensitive emails in Outlook desktop.

For information about legacy AIP Add-in, see admin guide for the AIP client.

There are multiple types of oversharing dialogs that can be presented to your users.

Default Oversharing Dialog

Applies to:

- Online E5 users with connected experience enabled
- Production version 2303 (Build 16.0.16216.10000) or higher.
- Semi-Annual Enterprise Channel version 2308 (Build 16.0.16731.20716) or higher.

This dialog uses the exact same text as the policy tip (default or custom) and when applicable, a noncustomizable set of justification options to override the policy.

Customized Oversharing Dialog

Applies to:

- Online E5 users with connected experience enabled
- Production version 2404 (Build 16.0.17531.20000) or higher.
- Semi-Annual Enterprise Channel version 2408 (Build TBD) or higher.

You can choose to tailor your oversharing dialog with a customized title, body, and dynamic variables like %%MatchedRecipientsList%%, and justification options.

For customized dialog, create a JSON file like this and ensure the following:

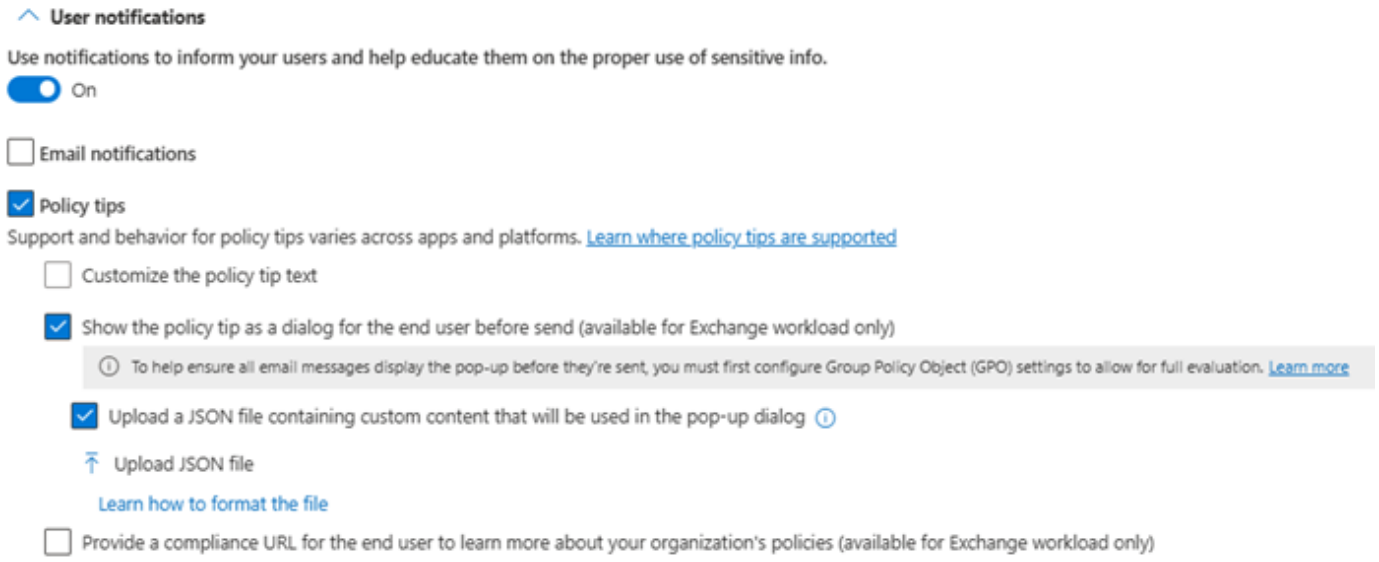
- The file is UTF-8 encoded.
- The content is plain text.
- No comments are included.

JSON

```
{
    "LocalizationData": [
        {
            "Language": "en-us",
            "Title": "WARNING: A Sensitivity Label Not for External Use was
Detected.",
            "Body": "The following classification(s) have been detected on this email
or its attachments. <LineBreak /><LineBreak /><Bold>%%MatchedLabelName%%</Bold>
<LineBreak /><LineBreak />The email cannot be sent until either the following issues
are corrected or a justification is provided. <LineBreak /><LineBreak />Attachment(s)
needing attention (if applicable): <LineBreak />%%MatchedAttachmentName%% <LineBreak
/><LineBreak />List of external recipients: <LineBreak />%%MatchedRecipientsList%%
<LineBreak /><LineBreak />",
            "Options": [
                "The recipients have signed an NDA",
                "Manager has approved this email",
                "Organization required this content to be shared"
            ]
        },
{
            "Language": "es-es",
            "Title": "ADVERTENCIA: Etiqueta de sensibilidad no para uso externo
detectada.",
            "Body": "Se ha detectado la(s) siguiente(s) clasificación(es) en este
correo electrónico o sus archivos adjuntos. <LineBreak /><LineBreak />
<Bold>%%MatchedLabelName%%</Bold><LineBreak /><LineBreak />El correo electrónico no se
puede enviar hasta que se corrijan los siguientes problemas o se proporcione una
justificación. <LineBreak /><LineBreak />Archivos adjuntos que necesitan atención (si
corresponde): <LineBreak />%%MatchedAttachmentName%% <LineBreak /><LineBreak />Lista
de destinatarios externos: <LineBreak />%%MatchedRecipientsList%% <LineBreak />
<LineBreak />",
            "Options": [
                "Los destinatarios han firmado un NDA",
                "El gerente ha aprobado este correo electrónico",
                "La organización requirió que se compartiera este contenido"
            ]
        }
    ],
    "DefaultLanguage": "en-us"
}
```

The above JSON content can be uploaded in a DLP rule using below options:

UX instructions



PowerShell

PowerShell

```
$content = Get-Content "path to the JSON file" -Encoding utf8| Out-String
```

PowerShell

```
New-DlpComplianceRule -Name "<Rule_name>" -Policy "<Policy_name>" `-<Any_Condition>
<Condition_Value> `-NotifyPolicyTipCustomDialog $content `
-NotifyPolicyTipDisplayOption Dialog
```

PowerShell

```
  Set-DlpComplianceRule -Identity "<Rule_name>" `-NotifyPolicyTipCustomDialog
$content `-NotifyPolicyTipDisplayOption Dialog
```

When you run the cmdlet, validation checks are run on the content. The validation checks include, the JSON character limit, formatting and, mandatory presence of one default language validation. The administrator is notified of any errors so that they can be corrected

Features and limitations

The dialog title, body, and override justifications options can be customized using the JSON file. You can apply bold, underline, italic formatting, and line breaks. There can be up to three justifications options plus an option for free-text input.

The text for **Acknowledgement** and **False positive overrides** isn't customizable.

Here's the required structure of the JSON files. You use this to customize the dialog for matched rules. The keys are **all case sensitive**. Formatting and dynamic tokens for matched conditions can only be used in the *Body* key.

Expand table

Keys	Mandatory?	Rules/Notes
{}	Y	Container
LocalizationData	Y	Array that contains all the language options.
Language	Y	Specify language code, limited to 10 languages.
Title	Y	Specify the title for the dialog. Limited to 75 characters.
Body	Y	Specify the body for the dialog. Limited to 800 characters. Dynamic tokens for matched conditions can be added in the body.
HasFreeTextOption	N	This can be set to true or false. True will display a text box as a last option in the dialog.
DefaultLanguage	Y	One of the languages must be defined as DefaultLanguage within the LocalizationData key.

Dynamic tokens and text formatting in custom Oversharing dialog

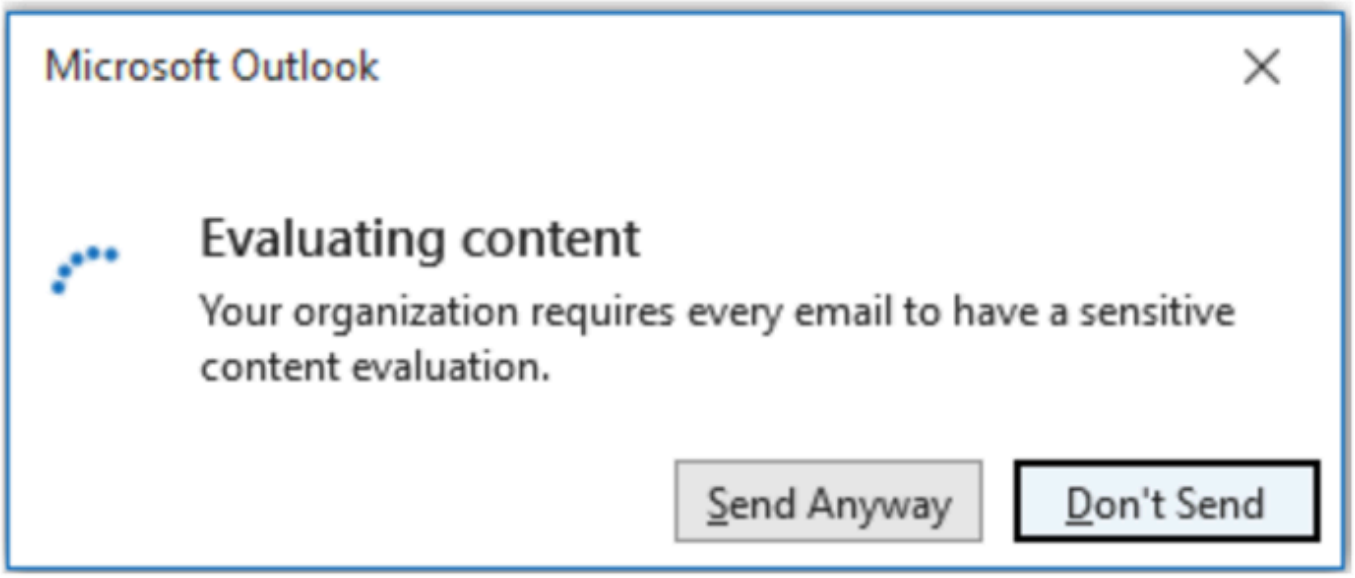
Expand table

Placeholder	Explanation
%%MatchedRecipientsList%%	Display the matched recipients for a given DLP rule for these conditions: - **Recipient is** - **Recipient domain is** - **Recipient is a member of** - **Content is shared from Microsoft 365**
%%MatchedLabelName%%	Display the matched labels for a given DLP rule for this condition: - **Content contains sensitivity label**
%%MatchedAttachmentName%%	Display the matched attachments for a given DLP rule for these conditions: - **Content contains sensitive information** - **Content contains sensitivity label** - **Attachment is not labeled** - **File extension is**
<Bold>lorem ipsum</Bold>	Bold format
<Italic>lorem ipsum</Italic>	Italic format
<Underline>lorem ipsum</Underline>	Underline

Placeholder	Explanation
<Linebreak /> or	Introduce a line break

Wait on Send dialog support for Oversharing for Outlook for Microsoft 365

After you've configured the Oversharing dialog, you can optionally enable the **Wait on send** dialog feature using the *dlpwaitonsendtimeout Registry key (DWORD value)*. Do this on all devices where you want to ensure sensitive emails are evaluated according to your DLP policies before they're sent. This registry key defines the maximum amount of time to hold an email when the user selects **Send**. It allows DLP policy evaluation to complete for labeled or sensitive content. The **Wait on send** dialog is noncustomizable.



Applies to:

- Online E5 users with connected experience enabled
- Production version 2303 (Build 16.0.16216.10000) or higher.
- Semi-Annual Enterprise Channel version 2302 (Build 16.0.16130.20478) or higher.

Configure Wait on send

To configure **Wait on send** see, Steps to configure Wait on send.

> **Important**
>
> Policy tips aren't supported for meeting invitations.

Last updated on 08/04/2026

Data loss prevention policy tip reference for Outlook on the Web

Mailboxes must be hosted in Exchange Online. For more information, see: Learn about data loss prevention.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about **Microsoft Security Copilot in Microsoft Purview**.

DLP policy tips supported

Yes.

Email notification supported for Outlook on the Web

Yes.

Conditions that support policy tips in Outlook on the Web

- Content contains (SITs, sensitivity labels)
- Content is shared from M365
- Sender is a member of
- Recipient Domain Is
- Recipient is
- Subject Contains Words or phrases
- Sender is
- Sender domain is
- File extension is
- Subject matches patterns
- Subject or Body contains words or phrases
- Doc or Attachment is password protected
- Document name contains words or phrases

- Sender address contains words
- Document size equals or is greater than
- Subject or Body matches patterns
- Recipient address contains words
- Sender address matches patterns
- Document name matches patterns
- Recipient address matches patterns
- Message importance is
- Document property is
- Content is not labeled

Actions that support policy tips in Outlook on the Web

- Restrict access or encrypt the content in Microsoft 365 locations
- Set headers
- Remove header
- Redirect the message to specific users
- Forward the message for approval to sender's manager
- Forward the message for approval to specific approvers
- Add recipient to the To box
- Add recipient to the Cc box
- Add recipient to the Bcc box
- Add the sender's manager as recipient
- Removed O365 Message Encryption and rights protection
- Prepend Email Subject
- Add HTML Disclaimer
- Modify Email Subject
- Deliver the message to the hosted quarantine

Sensitive information types that support policy tips in Outlook on the Web

These preconfigured sensitive information types (SITs) support policy tips in Outlook on the Web.

- ABA routing number
- Argentina national identity (DNI) number
- Australia bank account number
- Australia medical account number

- Australia passport number
- Australia tax file number
- Azure DocumentDB auth key
- Azure IAAS database connection string and Azure SQL connection string
- Azure IoT connection string
- Azure publish setting password
- Azure Redis cache connection string
- Azure SAS
- Azure service bus connection string
- Azure storage account key
- Azure Storage account key (generic)
- Belgium national number
- Brazil CPF number
- Brazil legal entity number (CNPJ)
- Brazil national identification card (RG)
- Canada bank account number
- Canada driver's license number
- Canada health service number
- Canada passport number
- Canada personal health identification number (PHIN)
- Canada social insurance number
- Chile identity card number
- China resident identity card (PRC) number
- Credit card number
- Croatia identity card number
- Croatia personal identification (OIB) number
- Czech personal identity number
- Denmark personal identification number
- Drug Enforcement Agency (DEA) number
- EU debit card number
- EU driver's license number
- EU national identification number
- EU passport number
- EU social security number or equivalent identification
- EU Tax identification number
- Finland national ID
- Finland passport number

- France driver's license number
- France national id card (CNI)
- France passport number
- France social security number (INSEE)
- Germany driver's license number
- Germany passport number
- Germany identity card number
- Greece national ID card
- Hong Kong identity card (HKID) number
- India permanent account number (PAN)
- India unique identification (Aadhaar) number
- Indonesia identity card (KTP) number
- International banking account number (IBAN)
- International classification of diseases (ICD-10-CM)
- International classification of diseases (ICD-9-CM)
- IP address
- IP Address v4
- IP Address v6
- Ireland personal public service (PPS) number
- Israel bank account number
- Israel national identification number
- Italy driver's license number
- Japan bank account number
- Japan driver's license number
- Japan passport number
- Japan resident registration number
- Japan social insurance number (SIN)
- Japan residence card number
- Malaysia identification card number
- Netherlands citizens service (BSN) number
- New Zealand ministry of health number
- Norway identification number
- Philippines unified multi-purpose identification number
- Poland identity card
- Poland national ID (PESEL)
- Poland passport number
- Portugal citizen card number

- Saudi Arabia National ID
- Singapore national registration identity card (NRIC) number
- South Africa identification number
- South Korea resident registration number
- Spain social security number (SSN)
- SQL Server connection string
- Sweden national ID
- Sweden passport number
- SWIFT code
- Taiwan national identification number
- Taiwan passport number
- Taiwan-resident certificate (ARC/TARC) number
- Thai population identification code
- Turkey national identification number
- U.K. drivers license number
- U.K. electoral roll number
- U.K. national health service number
- U.K. national insurance number (NINO)
- U.S./U.K. passport number
- U.S. bank account number
- U.S. driver's license number
- U.S. individual taxpayer identification number (ITIN)
- U.S. social security number (SSN)

Exact data match sensitive information types support for policy tips in Outlook on the Web

Yes.

All exact data match sensitive information types are custom created. For more information, see Learn about exact data match based sensitive information types.

Custom sensitive information types support for policy tips in Outlook on the Web

Yes.

Custom sensitive information types that use REGEX, functions, keyword lists, and keyword dictionaries support policy tips in Outlook on the Web. For more information, see Create custom sensitive information types and Create a custom sensitive information type using PowerShell.

Sensitivity labels support for policy tips in Outlook on the Web

Yes.

Retention labels support for policy tips in Outlook on the Web

No.

Trainable classifiers support for policy tips in Outlook on the Web

No.

Important information

- When emails are encrypted with Microsoft Purview Message Encryption and the policy used to detect them uses the *detect encryption* condition, policy tips will not appear.
- Administrators can use the DlpViaDcsEnabled parameter to enable exact same features and conditions that New Outlook does, see: Learn about Organisation configuration.

Last updated on 06/17/2026

Data loss prevention policy tip reference for SharePoint in Microsoft 365 and OneDrive for work or school web client

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

DLP policy tips supported

Yes.

> **Important**
>
> When emails are encrypted with Microsoft Purview Message Encryption and the policy used to detect them uses the detect encryption condition policy tips will not appear.

For more information on blocking and notifications in SharePoint in Microsoft 365 and OneDrive for work or school, see Blocking and notifications in SharePoint in Microsoft 365 and OneDrive for work or school

Conditions that support policy tips in SharePoint in Microsoft 365 and OneDrive for work or school web client

- Content contains
- Content is shared from Microsoft 365
- Document property is
- File extension is
- Document name contains words or phrases
- Document size equals or is greater than
- Document created by
- Document created by member of

- Document is shared (OneDrive for work or school web client only)

Actions that support policy tips in SharePoint in Microsoft 365 and OneDrive for work or school web client

- Restrict access or encrypt the content in Microsoft 365 locations

Sensitive information types that support policy tips in SharePoint in Microsoft 365 and OneDrive for work or school web client

These preconfigured sensitive information types (SIT) support policy tips in SharePoint in Microsoft 365 and OneDrive for work or school web client

- ABA routing number
- All credentials
- All full names
- All medical terms and conditions
- All Physical Addresses
- Amazon S3 Client Secret Access Key
- Argentina national identity (DNI) number
- Argentina Unique Tax Identification Key (CUIT/CUIL)
- ASP.NET machine Key
- Australia bank account number
- Australia business number
- Australia company number
- Australia drivers license number
- Australia medical account number
- Australia passport number
- Australia physical addresses
- Australia tax file number
- Austria drivers license number
- Austria identity card
- Austria passport number
- Austria physical addresses
- Austria social security number

- Austria tax identification number
- Austria value added tax
- Azure App Service deployment password
- Azure Batch shared access key
- Azure Bot Framework secret key
- Azure Bot service app secret
- Azure Cognitive Search API key
- Azure Cognitive Service key
- Azure Container Registry access key
- Azure Cosmos DB account access key
- Azure Databricks personal access token
- Azure DevOps app secret
- Azure DevOps personal access token
- Azure DocumentDB auth key
- Azure EventGrid access key
- Azure Function Master / API key
- Azure IAAS database connection string and Azure SQL connection string
- Azure IoT connection string
- Azure IoT shared access key
- Azure Logic app shared access signature
- Azure Machine Learning web service API key
- Azure Maps subscription key
- Azure publish setting password
- Azure Redis cache connection string
- Azure Redis cache connection string password
- Azure SAS
- Azure service bus connection string
- Azure service bus shared access signature
- Azure Shared Access key / Web Hook token
- Azure SignalR access key
- Azure SQL connection string
- Azure storage account access key
- Azure storage account key
- Azure Storage account key (generic)
- Azure Storage account shared access signature
- Azure Storage account shared access signature for high risk resources
- Azure subscription management certificate

- Belgium driver's license number
- Belgium national number
- Belgium passport number
- Belgium physical addresses
- Belgium value added tax number
- Blood test terms
- Brand medication names
- Brazil CPF number
- Brazil legal entity number (CNPJ)
- Brazil national identification card (RG)
- Brazil physical addresses
- Bulgaria driver's license number
- Bulgaria passport number
- Bulgaria physical addresses
- Bulgaria uniform civil number
- Canada bank account number
- Canada driver's license number
- Canada health service number
- Canada passport number
- Canada personal health identification number (PHIN)
- Canada physical addresses
- Canada social insurance number
- Chile identity card number
- China resident identity card (PRC) number
- China physical addresses
- Client secret / API key
- Credit card number
- Colombia national ID
- Colombia tax identification number
- Croatia driver's license number
- Croatia identity card number
- Croatia passport number
- Croatia personal identification (OIB) number
- Croatia physical addresses
- Cyprus drivers license number
- Cyprus identity card
- Cyprus passport number

- Cyprus physical addresses
- Cyprus tax identification number
- Czech driver's license number
- Czech passport number
- Czech personal identity number
- Czech Republic physical addresses
- Denmark driver's license number
- Denmark passport number
- Denmark personal identification number
- Denmark physical addresses
- Diseases
- Drug Enforcement Agency (DEA) number
- Ecuador Unique Identification Number
- Estonia driver's license number
- Estonia passport number
- Estonia Personal Identification Code
- Estonia physical addresses
- EU debit card number
- EU driver's license number
- EU national identification number
- EU passport number
- EU social security number or equivalent identification
- EU Tax identification number
- Finland driver's license number
- Finland European health insurance number
- Finland national ID
- Finland passport number
- Finland physical addresses
- France driver's license number
- France health insurance number
- France national id card (CNI)
- France passport number
- France physical addresses
- France social security number (INSEE)
- France tax identification number
- France value added tax number
- General password

- General Symmetric key
- Generic medication names
- German Driver's License Number
- Germany identity card number
- German Passport Number
- Germany physical addresses
- Germany tax identification number
- Germany value added tax number
- GitHub Personal Access Token
- Google API key
- Greenland physical addresses
- Greece driver's license number
- Greece national ID card
- Greece passport number
- Greece physical addresses
- Greece Social Security Number (AMKA)
- Greek Tax Identification Number
- Hong Kong identity card (HKID) number
- Http authorization header
- Hungary driver's license number
- Hungary passport number
- Hungary personal identification number
- Hungary physical addresses
- Hungary social security number (TAJ)
- Hungary tax identification number
- Hungary value added tax number
- Iceland physical addresses
- Impairments Listed In The U.S. Disability Evaluation Under Social Security
- India driver's License Number
- India GST Number
- India permanent account number (PAN)
- India unique identification (Aadhaar) number
- India Voter Id Card
- Indonesia driver's license number
- Indonesia identity card (KTP) number
- Indonesia passport number
- International banking account number (IBAN)

- International classification of diseases (ICD-10-CM)
- International classification of diseases (ICD-9-CM)
- IP address
- IP Address v4
- IP Address v6
- Ireland driver's license number
- Ireland passport number
- Ireland personal public service (PPS) number
- Ireland physical addresses
- Israel bank account number
- Israel national identification number
- Italy driver's license number
- Italy fiscal code
- Italy passport number
- Italy physical addresses
- Italy value added tax number
- Japan bank account number
- Japan driver's license number
- Japanese My Number Corporate
- Japanese My Number Personal
- Japan passport number
- Japan physical addresses
- Japanese Residence Card Number
- Japan resident registration number
- Japan social insurance number (SIN)
- Lab test terms
- Latvia driver's license number
- Latvia passport number
- Latvia personal code
- Latvia physical addresses
- Liechtenstein physical addresses
- Lifestyles that relate to medical conditions
- Lithuania driver's license number
- Lithuania passport number
- Lithuania personal code
- Lithuania physical addresses
- Luxemburg driver's license number

- Luxemburg national identification number (natural persons)
- Luxemburg national identification number (non-natural persons)
- Luxemburg passport number
- Luxemburg physical addresses
- Malaysia identification card number
- Malaysia passport number
- Malta driver's license number
- Malta identity card number
- Malta passport number
- Malta physical addresses
- Malta tax identification number
- Medical specialities
- Medicare Beneficiary Identifier (MBI) card
- Mexico Unique Population Registry Code (CURP)
- Microsoft Bing maps key
- Microsoft Entra client access token
- Microsoft Entra client secret
- Microsoft Entra user Credentials
- Netherlands citizen's service (BSN) number
- Netherlands driver's license number
- Netherlands passport number
- Netherlands physical addresses
- Netherlands tax identification number
- Netherlands value added tax number
- New Zealand bank account number
- New Zealand driver's license number
- New Zealand inland revenue number
- New Zealand ministry of health number
- New Zealand physical addresses
- New Zealand social welfare number
- Norway identification number
- Norway physical addresses
- Philippines national identification number
- Philippines passport number
- Philippines unified multi-purpose identification number
- Poland driver's license number
- Poland identity card

- Poland national ID (PESEL)
- Poland Passport Number
- Poland physical addresses
- Poland REGON number
- Poland tax identification number
- Portugal citizen card number
- Portugal driver's license number
- Portugal passport number
- Portugal physical addresses
- Portugal tax identification number
- Qatari ID Card Number
- Romania driver's license number
- Romania passport number
- Romania personal numeric code (CNP)
- Romania physical addresses
- Russia passport number domestic
- Russia passport number international
- Russia physical addresses
- Russia taxpayer identification number
- Saudi Arabia National ID
- Singapore driver's license number
- Singapore passport number
- Singapore national registration identity card (NRIC) number
- Singapore physical addresses
- Slack access token
- Slovakia driver's license number
- Slovakia passport number
- Slovakia personal number
- Slovakia physical addresses
- Slovenia driver's license number
- Slovenia passport number
- Slovenia physical addresses
- Slovenia tax identification number
- Slovenia Unique Master Citizen Number
- South Africa identification number
- South Africa physical addresses
- South Korea driver's license number

- South Korea passport number
- South Korea resident registration number
- Spain DNI
- Spain driver's license number
- Spain passport number
- Spain physical addresses
- Spain social security number (SSN)
- Spain tax identification number
- SQL Server connection string
- Surgical procedures
- Sweden driver's license number
- Sweden national ID
- Sweden passport number
- Sweden physical addresses
- Sweden tax identification number
- SWIFT code
- Switzerland physical addresses
- Swiss Social Security Number AHV
- Taiwan National ID
- Taiwan passport number
- Taiwan-resident certificate (ARC/TARC) number
- Thai population identification code
- Turkey national identification number
- Turkey physical addresses
- Types of medication
- U.A.E. identity card number
- U.A.E. passport number
- U.K. driver's license number
- U.K. electoral roll number
- U.K. national health service number
- U.K. national insurance number (NINO)
- U.K. physical addresses
- U.K. Unique Taxpayer Reference Number
- U.S. bank account number
- U.S. driver's license number
- U.S. individual taxpayer identification number (ITIN)
- U.S. physical addresses

- U.S. social security number (SSN)
- U.S./U.K. passport number
- Ukraine passport domestic
- Ukraine passport international
- Ukraine physical addresses
- User login credentials
- X.509 certificate private key

Exact data match sensitive information types support for policy tips in SharePoint in Microsoft 365 and OneDrive for work or school web client

Yes.

All exact data match sensitive information types are custom created. For more details, see Learn about exact data match based sensitive information types.

Custom sensitive information types support for policy tips in SharePoint in Microsoft 365 and OneDrive for work or school web client

Yes.

Custom sensitive information types that use REGEX, functions, keyword lists and keyword dictionaries support policy tips in Outlook on the Web. For more details, see Create custom sensitive information types and Create a custom sensitive information type using PowerShell.

Sensitivity labels support for policy tips in SharePoint in Microsoft 365 and OneDrive for work or school web client

Yes.

Retention labels support for policy tips in SharePoint in Microsoft 365 and OneDrive for work or school web client

No.

Trainable classifiers support for policy tips in SharePoint in Microsoft 365 and OneDrive for work or school web client

No.

Last updated on 06/15/2026

Data loss prevention policy tip reference for new Outlook for Windows

New Outlook now supports data loss prevention (DLP) policy tips with commonly used predicates and exceptions, advanced classifiers, and override capabilities.

Licensing requirements

For information on licensing, see

- Microsoft 365 Enterprise Plans
- Microsoft 365 Service Descriptions

> **Note**
>
> Features are enabled based on Licenses and connected experience settings. Review license requirements in. For any of the following conditions to work, connected experience must be turned on.

Expand table

License	These conditions apply
E3 and equivalent licenses	- Content contains Microsoft built-in or custom sensitive information types - Content is shared from Microsoft 365
E5 and equivalent licenses	- Content contains built-in or custom sensitive information types - Content is shared - Content contains sensitivity labels (works for email and Office and PDF file types) - Sender is - Sender is member of (Only Distribution lists, Azure-based Dynamic Distribution groups, and email-enabled Security groups are supported) - Sender domain is - Recipient is - Recipient is a member of (Only Distribution lists, Azure-based Dynamic Distribution groups, and email-enabled Security groups are supported) - Recipient domain is - Subject contains words - Content isn't labeled - Message isn't labeled

License	These conditions apply
	- Attachment isn't labeled - File extension is

> **Note**
>
> Conditions related to attachments are supported, except for messages (.msg) and emails (.eml), and encrypted attachments, which can't be scanned for content. See the new Outlook roadmap for details on future updates.

Actions that support policy tips

Refer to the Exchange actions support policy tips for details on actions that support policy tips.

> **Note**
>
> Offline evaluation for data loss prevention isn't available for new Outlook.

Sensitive information types that support policy tips

Refer to the Sensitive information types for details on sensitive information types that support policy tips for Outlook perpetual users and Microsoft 365 users.

Oversharing dialog for New Outlook for Windows

The oversharing dialog is available in DLP for new Outlook for E5 users. When enabled in a DLP rule, this feature displays popups for warning, override, or block actions to end users who are sharing labeled or sensitive emails in Outlook desktop.

For information about legacy AIP Add-in, see admin guide for the AIP client.

Default Oversharing Dialog

This dialog uses the exact same text as the policy tip (default or custom) and when applicable, a noncustomizable set of justification options to override the policy.

Customized Oversharing Dialog

You can choose to tailor your oversharing dialog with a customized title, body, and dynamic variables like %%MatchedRecipientsList%%, and justification options.

For customized dialog, create a JSON file like this and ensure the following:

- The file is UTF-8 encoded.
- The content is plain text.
- No comments are included.

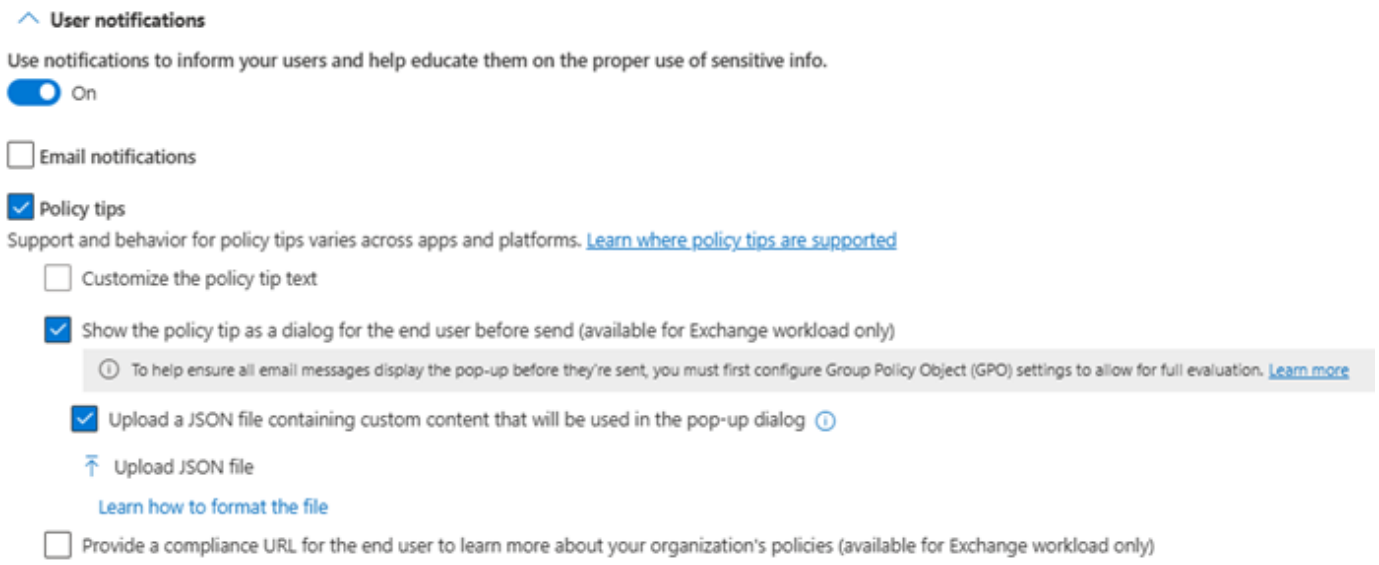
JSON

```
{
    "LocalizationData": [
        {
            "Language": "en-us",
            "Title": "WARNING: A Sensitivity Label Not for External Use was
Detected.",
            "Body": "The following classification(s) have been detected on this email
or its attachments. <LineBreak /><LineBreak /><Bold>%%MatchedLabelName%%</Bold>
<LineBreak /><LineBreak />The email cannot be sent until either the following issues
are corrected or a justification is provided. <LineBreak /><LineBreak />Attachment(s)
needing attention (if applicable): <LineBreak />%%MatchedAttachmentName%% <LineBreak
/><LineBreak />List of external recipients: <LineBreak />%%MatchedRecipientsList%%
<LineBreak /><LineBreak />",
            "Options": [
                "The recipients have signed an NDA",
                "Manager has approved this email",
                "Organization required this content to be shared"
            ]
        },
{
            "Language": "es-es",
            "Title": "ADVERTENCIA: Etiqueta de sensibilidad no para uso externo
detectada.",
            "Body": "Se ha detectado la(s) siguiente(s) clasificación(es) en este
correo electrónico o sus archivos adjuntos. <LineBreak /><LineBreak />
<Bold>%%MatchedLabelName%%</Bold><LineBreak /><LineBreak />El correo electrónico no se
puede enviar hasta que se corrijan los siguientes problemas o se proporcione una
justificación. <LineBreak /><LineBreak />Archivos adjuntos que necesitan atención (si
corresponde): <LineBreak />%%MatchedAttachmentName%% <LineBreak /><LineBreak />Lista
de destinatarios externos: <LineBreak />%%MatchedRecipientsList%% <LineBreak />
<LineBreak />",
            "Options": [
                "Los destinatarios han firmado un NDA",
                "El gerente ha aprobado este correo electrónico",
                "La organización requirió que se compartiera este contenido"
            ]
        }
    ],
```

```
    "DefaultLanguage": "en-us"
}
```

The above JSON content can be uploaded in a DLP rule using below options:

UX instructions



PowerShell

PowerShell

```
$content = Get-Content "path to the JSON file" -Encoding utf8| Out-String
```

PowerShell

```
New-DlpComplianceRule -Name "<Rule_name>" -Policy "<Policy_name>" `-<Any_Condition>
<Condition_Value> `-NotifyPolicyTipCustomDialog $content `
-NotifyPolicyTipDisplayOption Dialog
```

PowerShell

```
  Set-DlpComplianceRule -Identity "<Rule_name>" `-NotifyPolicyTipCustomDialog
$content `-NotifyPolicyTipDisplayOption Dialog
```

When you run the cmdlet, validation checks are run on the content. The validation checks include, the JSON character limit, formatting and, mandatory presence of one default language validation. The administrator is notified of any errors so that they can be corrected

Features and limitations

The dialog title, body, and override justifications options can be customized using the JSON file. You can apply bold, underline, italic formatting, and line breaks. There can be up to three justifications options plus an option for free-text input.

The text for **Acknowledgement** and **False positive overrides** isn't customizable.

Here's the required structure of the JSON files. You use this to customize the dialog for matched rules. The keys are **all case sensitive**. Formatting and dynamic tokens for matched conditions can only be used in the *Body* key.

Expand table

Keys	Mandatory?	Rules/Notes
{}	Y	Container
LocalizationData	Y	Array that contains all the language options.
Language	Y	Specify language code, limited to 10 languages.
Title	Y	Specify the title for the dialog. Limited to 75 characters.
Body	Y	Specify the body for the dialog. Limited to 800 characters. Dynamic tokens for matched conditions can be added in the body.
HasFreeTextOption	N	This can be set to true or false. True will display a text box as a last option in the dialog.
DefaultLanguage	Y	One of the languages must be defined as DefaultLanguage within the LocalizationData key.

Dynamic tokens and text formatting in custom Oversharing dialog

Expand table

Placeholder	Explanation
%%MatchedRecipientsList%%	Display the matched recipients for a given DLP rule for these conditions: - **Recipient is** - **Recipient domain is** - **Recipient is a member of** - **Content is shared from Microsoft 365**
%%MatchedLabelName%%	Display the matched labels for a given DLP rule for this condition: - **Content contains sensitivity label**
%%MatchedAttachmentName%%	Display the matched attachments for a given DLP rule for these conditions: - **Content contains sensitive information** - **Content contains sensitivity label** - **Attachment is not labeled** - **File extension is**

Placeholder	Explanation
<Bold>lorem ipsum</Bold>	Bold format
<Italic>lorem ipsum</Italic>	Italic format
<Underline>lorem ipsum</Underline>	Underline
<Linebreak /> or	Introduce a line break

Wait on send dialog support for oversharing

New Outlook now supports *DLP Wait-to-Send* through centrally managed Exchange Online mailbox settings. The new experience allows for multi-tasking and doesn't block the user to use New Outlook even when mail is in evaluation.

During evaluation, a toast briefly notifies the user that Outlook hasn't sent the email yet and is evaluating the content. The toast gets autodismissed while evaluation continues in the background.

The wait-to-send experience is controlled solely through the following Exchange mailbox parameters:

- **DLPWaitOnSendEnabled** *boolean*: true or false.
  Default: false When true, it enables evaluation of the mail for Data loss prevention before mail is sent.

- **DLPWaitOnSendTimeout** *Integer*: range 0–10,000 seconds.
  Default:__ 9999 Specifies how long Outlook waits for DLP evaluation to complete before offering override. For Example – If the value is set to 25 seconds, Outlook waits to complete the evaluation for 25 seconds before it allows user to override and allow to send the mail without completing evaluation.

Note that this doesn't mean that evaluation is stopped after the set time, it just means that it allows for overriding the evaluation.

- When DLPWaitOnSendTimeout = 0, user will immediately see “Send Anyway” in the infobar of mail.

- When DLPWaitOnSendTimeout >= 9999, user will never be allowed to send the email without DLP evaluation getting completed.

Refer this page for Exchange mailbox parameters - Set-OrganizationConfig (ExchangePowerShell) | Microsoft Learn

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/30/2026

Data loss prevention policy tip reference for Outlook for Android, iOS, and macOS

This feature is in ***preview***.

Outlook for Android, iOS, and macOS now supports data loss prevention (DLP) policy tips with commonly used predicates and exceptions, advanced classifiers, and override capabilities.

Licensing requirements

Policy tips supported in Outlook for Android, iOS, and macOS

Enable Policy Tips UI to Appear

Policy Tip UI is turned off by default in Outlook for Android, iOS, and macOS to allow for administrators to test and perform user training prior to deployment.

To enable Policy Tip UI to appear in Outlook for Android, iOS, and macOS, you must enable the policy listed in the following section in the Microsoft 365 Apps Admin Center. To access this policy setting, go to the Microsoft 365 Admin Center, expand "All Admin Centers" to find the "Microsoft 365 app" admin center.

You can access this policy under **Customization** and **Policy Management**. You can target policies to individuals or groups, so you can test and deploy policy tip UI as needed through your organization.

Expand table

Policy Name	Value
Enable Purview Data Loss Prevention (DLP) policy tips in Outlook	(default) **Disabled** policy tips aren't shown **Enabled** policy tips are shown

Supported conditions and actions

For details on actions that support policy tips, see Exchange actions support policy tips.

Supported SITs

For details on sensitive information types that support policy tips for Microsoft 365 users, see Sensitive information types.

Supported custom SITs

All custom SITs are supported.

Supported EDM SITs

All EDM SITs are supported.

Supported sensitivity labels

All sensitivity labels are supported.

Supported retention labels

All retention labels are supported.

Supported trainable classifiers

All trainable classifiers are supported.

Actions and information types that support policy tips

All actions and information types for email communications are supported.

For details on creating and using all policy tip types, see DLP Policy Reference.

Oversharing dialog for Outlook for Android and iOS

The oversharing dialog is available in DLP for Outlook for Android and iOS users when used with the necessary licenses and configuration. When you enable this feature in a DLP rule, it displays popups for warning, override, or block actions to end users who are sharing labeled or sensitive emails in Outlook for Android and iOS.

Default oversharing dialog

This dialog uses the exact same text as the policy tip (default or custom) and, when applicable, a noncustomizable set of justification options to override the policy.

Customized oversharing dialog

You can choose to tailor your oversharing dialog with a customized title, body, and dynamic variables like %%MatchedRecipientsList%%, and justification options. For more information, see Customized Oversharing Dialog.

Wait on send dialog support for oversharing

Outlook for Android and iOS doesn't yet support *DLP Wait-to-Send, which is* centrally managed Exchange Online mailbox settings. For future announcements, see the Microsoft 365 Roadmap, Microsoft 365 Roadmap.

For Exchange mailbox parameters, see Set-OrganizationConfig (ExchangePowerShell) | Microsoft Learn.

ⓘ **Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 04/16/2026

Learn about Data Security Investigations

Microsoft Purview Data Security Investigations helps cybersecurity teams in your organization use generative artificial intelligence (AI) to analyze and respond to data security incidents, risky insiders, and data breaches. Investigations help you quickly identify risks from sensitive data exposure and more effectively collaborate with your partner teams to remediate the issues. It simplifies tasks that traditionally are time consuming and complex.

https://learn-video.azurefd.net/vod/player?id=cf1d0239-cb00-4db6-a925-dfdb9e605dd2&locale=en-us&embedUrl=%2Fpurview%2Fdata-security-investigations

Analysts in your organization can use Data Security Investigations features to:

- Quickly and efficiently search, discover, and identify impacted data.
- Use deep content AI analysis to discover exact data risks hidden in data.
- Take action to reduce the impact of data security incidents and quickly mitigate ongoing risks.
- Collaborate with internal and external stakeholders on investigation details.
- Receive notifications through the Microsoft Purview Notification Center and email when setup completes and investigations are ready to use.

For more information about how Data Security Investigations can help you respond to data security incidents, watch the following videos:

- Data Security Investigations in Microsoft Purview
- Enhance data security investigations with Microsoft Purview
- Introducing Microsoft Purview Data Security Investigations
- Unlock the Power of Data Security Investigations with Microsoft Purview
- Microsoft Purview Data Security Investigations Ask Me Anything

AI integration

Data Security Investigations uses generative AI to conduct deep content analysis and uncover key security and sensitive data risks for data included in investigations. AI helps analysts quickly analyze large volumes of data with high accuracy, saving critical time for triage, review, and mitigation actions.

There are three main AI-related investigative capabilities:

- **Vector search**: Vector-based semantic search enables similarity-based information retrieval and understands user intent beyond literal words. Analysts query impacted data to find all assets related to a particular subject, even if keywords are missing.
- **Categorization**: To get an initial understanding of incident severity, analysts can use **Standard** or **Advanced** AI options to categorize impacted data, narrowing focus to high-risk assets. Data Security Investigations automatically sorts data into default, custom, or AI-suggested categories. Categorized items also include grouping by subject matter and risk level.
- **Examination**: Data Security Investigations enables analysts to easily address the number one priority for most data security incidents - finding security risks buried within impacted data. By examining impacted data for security risks, analysts can find compromised credentials, network risks, personal data exposure, or evidence of threat actor discussions associated with security incidents.

For more information about AI integration and these tools, see Learn about AI analysis in Data Security Investigations.

Common data breach actions

Most organizations are concerned with the increasing risks and impact related to internal and external data security breaches. Your cybersecurity team is responsible for identifying impacted data during a data breach scenario and reducing the time required to mitigate the risk associated with vulnerabilities exposed within a data breach scenario. When a data breach occurs, you must respond quickly and effectively to mitigate risks and protect sensitive information.

Your cybersecurity team needs to:

- **Assess the data breach**: Analysts must understand the scope and impact of the data breach. This understanding includes identifying all affected systems, the data impacted, and any external and internal users associated with the data breach.
- **Containment and communication**: Analysts must isolate affected systems to minimize and prevent further damage. Additionally, internal stakeholders and departments in your organization (IT, Legal, Executive Management) need notification and the details about the data breach.
- **Forensics and evidence preservation**: Analysts might need support when engaging with data forensics experts. System and status details must be retained (logs, system snapshots, network traffic) for potential legal and regulatory requirements.

- **Risk assessment and mitigation**: Analysts must evaluate potential risks and prioritize mitigation actions based on risk severity. This information helps with implementing configuration and policy updates, including implementing applicable security patches.
- **Data purge and protection**: Analysts might need to remove unnecessary or compromised data and delete redundant copies. Data Security Investigations supports soft purge and hard purge options to remove or permanently delete items from data sources. Other tasks might include strengthening access controls, data encryption, and monitoring.
- **Reporting and compliance**: Depending on the nature of the data involved and the jurisdictions affected, analysts might need to comply with regulatory requirements for breach notification. These requirements might include informing affected individuals, regulatory bodies, and other stakeholders about the nature of the breach and the steps taken to address it.

Common scenarios

Investigate a data breach

After a data security incident, it can be difficult to understand the impact of sensitive data exposure. As part of your mitigation strategy, you need to identify the intellectual property, personal data, and financial information that might be compromised. Additionally, some organizations build and maintain custom tools to investigate these issues.

When you use Microsoft Defender XDR to investigate data security incidents, you might discover a document containing unfiled patents. When you create an investigation from the incident in Data Security Investigations, further analysis shows which users downloaded the document and if it was accessed from a risky IP address. This information gives you critical context to take mitigation steps, such as protecting or purging the sensitive document from your organization.

Investigate a potential data leak from a risky insider

When you use Insider Risk Management to investigate data leak alerts for a risky user in your organization, you might learn that the user shared emails and files with an external storage solution. With built-in integration, you can create a new investigation in Data Security Investigations from the Insider Risk Management case to quickly review preliminary insights about the shared content. Before committing any additional compute resources or charges, you can decide if this content needs further analysis.

If you add the shared items to the scope of the investigation, processing automatically starts so you can dig further into the details of the content. When processing completes, all the data is

categorized and assessed for risk severity. Analysts can quickly review the highest severity items and identify the items needing a closer look. Details about usernames, passwords, documentation content, and source code are all available for review as applicable. Analysts can then work with other stakeholders in your organization to take action on the most critical items to mitigate ongoing risks from the data leak.

Confidential data leak

In scenarios where confidential, sensitive, or protected information is intentionally or unintentionally exposed outside its intended environment, you can use Data Security Investigations to further investigate a potential data leak, better understand incident impact, and take necessary next steps.

Proactive data security assessment

You can also use Data Security Investigations to proactively assess your data estate for data risk. You can analyze a sample of high-value data sources and users for data risk or run proactive scans of specific high-value data sources or a broader organization-wide scope. The assessment helps to identify opportunities for refining policies or organizational changes to strengthen security practices and potentially prevent or reduce the impact of future data security incidents.

When you enable proactive AI insights from Data Security Posture Management (preview), Data Security Investigations automatically analyzes recently exfiltrated sensitive data across five risk categories, providing continuous visibility into data exposure without manual investigation creation.

Integration with other Microsoft platforms and solutions

Data Security Investigations is a unified, purpose-built solution that's tightly integrated with other Microsoft security services. Integrated with the Microsoft Graph-enabled approach to security, Data Security Investigations draws correlations between impacted data, users, and activities. This integration helps you quickly identify impacted data, investigate incidents using generative AI, and securely collaborate with other members of your security team to mitigate risk.

Microsoft Defender XDR

Microsoft Defender XDR is a unified pre- and post-breach enterprise defense suite that natively coordinates detection, prevention, investigation, and response across endpoints, identities, email,

and applications to provide integrated protection against sophisticated attacks. With built-in integration with Data Security Investigations, you can extend the analysis of threat signals and dive deeper into data associated with attacks and security incidents.

Using the Microsoft Defender portal, you can assess alerts and quickly open an investigation in Data Security Investigations for deeper insights about the data and users associated with the incident. The investigation helps you analyze the alert details, understand what they mean, and collect and examine data for deeper review and mitigation efforts.

For more information about Microsoft Defender XDR features, see Get started with Microsoft Defender XDR.

Microsoft Purview Insider Risk Management

By combining the integration and investigative capabilities of Data Security Investigations with the proactive risk management features of Microsoft Purview Insider Risk Management, your organization has a holistic view of potential risks and the ability to take timely actions to protect sensitive data. Insider Risk Management investigators might not have important insights into the content contained in files associated with risky alerts. This gap might lead to the mishandling of alerts and ultimately hinder your effectiveness at mitigating insider threats.

The integration between Data Security Investigations and Insider Risk Management enables seamless collaboration between your cybersecurity and risk management teams. This collaboration allows for more efficient identification and a quicker mitigation of insider threats. For more information about Insider Risk Management features, see the following articles:

- Get started with Insider Risk Management
- Create and manage Insider Risk Management policies
- Investigate insider risk activities
- Take action on insider risk cases

Microsoft Purview Data Security Posture Management

If you're using Microsoft Purview Data Security Posture Management (DSPM), you can escalate potential data exfiltration findings to Data Security Investigations for deeper, AI-powered content analysis. You can manually create an investigation from a DSPM (preview) insight, or enable proactive AI insights that automatically create and refresh an investigation for your tenant every 24 hours. Investigations created from DSPM are automatically scoped to all sensitive data recently exfiltrated from your organization, allowing your team to quickly understand what was exposed, assess its sensitivity, and take informed remediation actions.

Unified audit log

Microsoft Purview auditing solutions provide an integrated solution to help organizations effectively respond to security events, forensic investigations, internal investigations, and compliance obligations. Your organization's unified audit log captures, records, and retains thousands of user and admin operations performed in dozens of Microsoft services and solutions.

Data Security Investigations integrates with the unified audit log in two ways:

- **Audit search**: Use audit search to identify and collect content based on user activities recorded in the unified audit log. Audit search queries the unified audit log for content-associated activities, such as file access, downloads, and sensitivity label changes, and adds the related content directly to your investigation scope.
- **Activity logging**: Data Security Investigations provides powerful investigation tools that might require the ability to audit activities taken to ensure secure and approved use of the solution. To meet these requirements, Data Security Investigations activities are automatically logged in the unified audit log.

For more information about these activities, see the Data Security Investigations activities section in the Audit log activities article.

Microsoft Purview Endpoint Data Loss Prevention

Microsoft Purview Endpoint Data Loss Prevention (DLP) monitors and protects sensitive items on onboarded Windows and macOS devices. For a general overview of DLP capabilities, see Learn about data loss prevention. When endpoint DLP policies are configured with evidence collection for file activities on devices, copies of files that match policies are captured and stored for review.

Data Security Investigations integrates with endpoint DLP evidence collection (preview), allowing investigators to query captured endpoint data and add it to an investigation scope for AI-powered analysis. Instead of triaging individual endpoint DLP alerts one by one, investigators can aggregate endpoint events and analyze them collectively to detect patterns such as mass exfiltration or coordinated data theft.

For more information about using endpoint DLP data in investigations, see Endpoint DLP evidence collection (preview).

Billing models

Data Security Investigations helps you identify data risks for breached content by using large language models (LLMs) and Microsoft Copilot for Security. Because Data Security Investigations provides on-demand AI analysis and integrates with other Microsoft Purview solutions, it uses pay-as-you-go and capacity billing models. This billing structure means you pay for the storage and AI capacity features you use in the solution. However, you don't need a dedicated enterprise plan or license to use Data Security Investigations.

If you enable proactive AI insights from DSPM, storage and AI capacity costs are incurred continuously while the toggle remains enabled, as the auto-created investigation refreshes every 24 hours.

For more information about billing, see Billing models in Data Security Investigations.

Ready to get started?

- Learn about the Data Security Investigations workflow
- Get started with Data Security Investigations

ⓘ**Note:** The author created this article with assistance from AI. Learn more

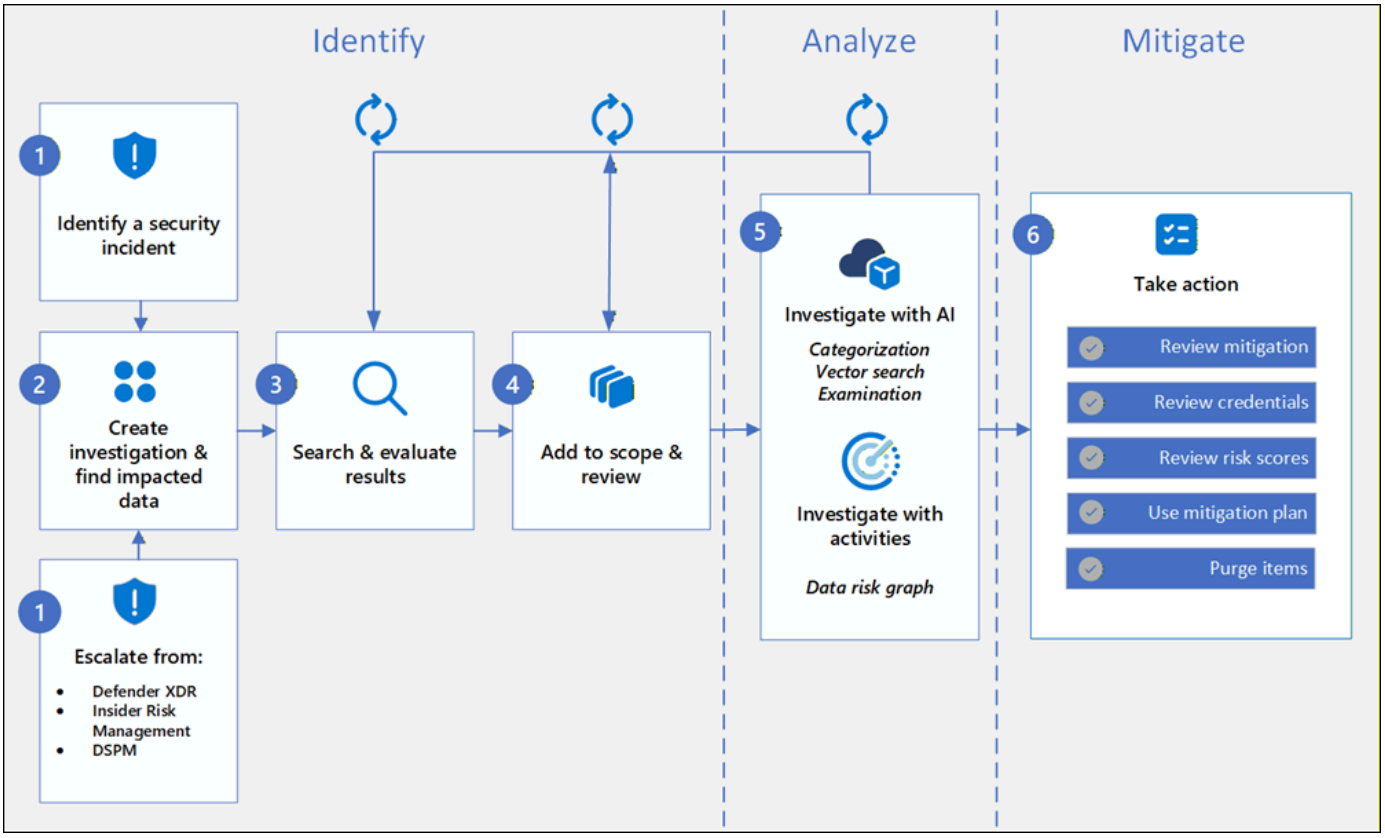
---

Last updated on 06/22/2026

Learn about the Data Security Investigations workflow

The Data Security Investigations workflow helps you quickly identify, investigate, and take action on data associated with security and data breach incidents. This workflow isn't a linear process. It includes significant iteration requirements for several of the steps to fine tune searches, evidence gathering, classification, and investigation by using AI and activities.

Identifying and taking action on data and access uses the following workflow:



Step 1: Identify and escalate incidents

Identify a data security incident

Data breaches and other data security incidents require quick action to identify and contain potential risks to your organization. You need to quickly identify these incidents and streamline an integrated response. Investigating a data security incident can be daunting. It might include inefficient workflows across multiple tools, manual work, and extra complexity as the investigation grows in size. You might face labor-intensive reviews of impacted data and increased costs.

Data Security Investigations helps you investigate and mitigate data security incidents and accelerates the time to resolution. After you identify a data security incident, create a new investigation to enable your data security team to identify incident-related data, conduct deep content analysis, and mitigate risk within one unified solution.

Escalate a data security incident from Microsoft Defender XDR

If you're already using Microsoft Defender XDR in your organization, the integration with Data Security Investigations allows you to quickly and seamlessly create a new investigation. The investigation automatically includes all incident-related data items from the Defender XDR incident node.

Escalate a data security incident from Insider Risk Management

If you're already using Insider Risk Management in your organization, the integration with Data Security Investigations allows you to quickly create a new investigation for identified risky user activities.

Escalate from an insight in Data Security Posture Management

If you're already using Data Security Posture Management in your organization, the integration with Data Security Investigations allows you to quickly create a new investigation from potential data exfiltration findings and insights.

Step 2: Create an investigation and find impacted data

Creating an investigation in Data Security Investigations is quick and easy. Depending on your scenario, you can create investigations from:

- **Microsoft Defender XDR incidents**: Create an investigation from a Defender XDR incident.
- **Microsoft Purview Insider Risk Management cases**: Create an investigation from Insider Risk Management.
- **Data Security Posture Management insights**: Create an investigation from Data Security Posture Management exfiltration findings, or enable proactive AI insights that auto-create an investigation.
- **Manually by using full draft mode**: Create an investigation by using the full draft mode option to configure specific data sources and search conditions.

Step 3: Search, evaluate results, and review

After creating an investigation, you can review and update data sources. Use search tools to identify items related to the data security incident. This review includes items from the following Microsoft 365 services:

Expand table

Data source	Content type searched	Risky content examples
Endpoint Data Loss Prevention (DLP)	Files captured by endpoint DLP evidence collection	Files copied to USB, printed, uploaded to cloud, or accessed by unauthorized apps on onboarded devices.
Exchange Online	Emails and attachments in user mailboxes	Credentials or secrets sent in email, sensitive files shared externally.
Microsoft Copilot	AI prompts and responses	Sensitive data in Copilot or AI prompts.
Microsoft Teams	Chat messages (1:1 and group chats) and channel posts	Secrets included in chats, confidential information in Teams conversations.
OneDrive	User files	User files with secure access keys, export of databases, and more.
SharePoint	Documents and files in sites	Documents containing passwords, customer data, or confidential plans.

You can create and run different searches that are associated with an investigation. Use conditions such as keywords, file types, incidents, and more in the query builder to create custom search queries. These queries return results with the data most likely relevant to the data security incident. You can also use audit search to identify content based on user activities recorded in the unified audit log, such as files accessed, downloads, and sensitivity label changes. Additionally, you can use endpoint DLP evidence collection to query data captured by endpoint DLP policies on onboarded devices.

Step 4: Add data to the investigation scope

After you review and refine your search query, add all relevant data items to the investigation scope. In this step, you filter and review specific data items. You also identify any items you don't want to review by using AI tools in Data Security Investigations.

Step 5(a): Investigate items

After adding data items to the investigation scope, start refining and narrowing the data to the most relevant items for your investigation. Narrowing the items to the smallest amount of applicable data helps increase the speed and reduce costs associated with AI processing.

In this step, perform the following actions:

- Select specific items to add directly to your mitigation plan if applicable.
- Identify or exclude specific items from AI processing.

After data items are added to the investigation scope, the system automatically prepares all items. This automatic vectorization enables semantic searches and categorization of data.

Step 5(b): Investigate with AI

After Add to scope completes for the data items in the investigation, start using AI-related tools to help narrow the focus of your investigation to only the most impactful and critical items.

Use the following tools and actions in your review to identify and take action on specific data items:

- Use vector search for natural language queries to identify items for review.
- Use Search with AI (preview) with natural language queries and keywords to summarize search results for review.
- Define and configure AI-driven categories for the content.
- Use built-in and custom examination areas for selected items.
- Select items to add to the mitigation plan.

> **Important**
>
> Storage and AI capacity cost considerations are associated with the usage of each of the AI tools in Data Security Investigations. For more information, see **Billing models in Data Security Investigations** and **Use AI analysis in Data Security Investigations**.

Step 6: Take actions to mitigate

After you identify the most relevant and impactful items associated with the data security incident, take specific actions to help mitigate risks.

- **Review mitigation recommendations**: When you select items for examination and choose *Mitigation* as the focus area, the solution creates mitigation recommendations. Automated

AI processing identifies related threats and recommends mitigation steps.

- **Review credentials examinations**: When you select items for examination and choose *Credentials* as the focus area, the solution identifies credentials and other access asset examinations. AI analysis processing automatically identifies and generates credential details, type, and specific recommendations.
- **Review risk examinations**: When you select items for examination and choose *Risks* as the focus area, the solution automatically creates a security risk score and examination. The risk score helps you prioritize mitigation actions for the most impactful and risky data items.
- **Review personal data examinations**: When you select items for examination and choose *Personal Data* as the focus area, the solution identifies and extracts personal data from selected items. AI analysis processing automatically identifies PII types, extracts values, and provides surrounding context.
- **Use the mitigation plan**: After reviewing examinations and recommendations, add specific data items from the investigation scope to the mitigation plan. This plan helps you manage and track the mitigation status for each data item.
- **Purge items if needed**: Create a search query and save it as a purge query to help you identify and remove items in your organization. You can soft purge items to move them to the recoverable items folder or hard purge items to permanently delete them.

Best practices

- **Iterate and use AI analysis wisely**: Investigations rarely finish in one step. Use the AI suggestions in categorization and vector search to uncover hidden issues. Always validate any critical findings by using human judgment. The investigator's judgment is key in critical decision-making.
- **Keep the investigation scope focused**: It's easy to get overwhelmed as you discover more data. Use categorization to exclude irrelevant information and concentrate primarily on any risky items.
- **Engage the right teams in your organization early**: Data security incidents often span multiple areas. These areas might include security, IT, and compliance in your organization. Share with partner teams as soon as you identify areas needing action.
- **Use integrated Microsoft services**: Data Security Investigations is part of the broader Microsoft Purview and Defender ecosystem. If your investigation suggests an insider risk, open an Insider Risk Management case. If it started from a Microsoft Defender XDR alert, update that incident with your findings. The goal is an end-to-end view of the incident in your organization.

Email and portal notifications

Data Security Investigations sends notifications to inform investigators about setup status and keep them updated on investigation progress. Notifications are delivered through both the Microsoft Purview Notification Center (bell icon in the Microsoft Purview portal) and email. Email notifications include the investigation name as the header and a button that links directly to the relevant page in the Microsoft Purview portal.

The following table describes the notifications associated with setup and onboarding:

Expand table

Notification	Trigger	Recipients	Description
Setup complete	Initial setup is completed	All users with Data Security Investigations roles	Notifies all assigned users that Data Security Investigations is configured and ready for use.
Integration investigation ready	Initial setup is completed and user previously created investigations from Microsoft Defender XDR, Insider Risk Management, or Data Security Posture Management	Users who created investigations before setup completed	Directs users back to their pre-existing investigation to begin AI analysis.
Setup reminder	User created an investigation but initial setup isn't completed after 7 days	User who created the investigation	Reminds the user or their admin to complete setup so the investigation can proceed.
Permissions provisioning complete	1 hour after permissions assignment when final onboarding step isn't completed	User completing onboarding	Reminds the user to complete the final onboarding step after permissions are provisioned.

> **Note**
>
> Notification preference settings aren't currently available. All users with Data Security Investigations roles receive notifications for the scenarios listed in this section.

Ready to get started?

- Get started with Data Security Investigations

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 09/24/2026

Learn about AI analysis in Data Security Investigations

> **Important**
>
> Data Security Investigations uses generative artificial intelligence (AI), large language models, and orchestration in the analysis of data in your organization. Results generated by AI might not always be accurate or complete. While we strive to provide reliable and helpful information, AI systems can produce incorrect or false results. It is important to verify the information and use it with caution. Microsoft makes no warranties, express, implied, or statutory, as to the information provided by AI systems.

Data Security Investigations uses AI services and tools to help you quickly review and take action on items associated with security incidents. AI-related services include the following tools:

- Vector search
- Categorization
- Examination

Vector search

Vector search in Data Security Investigations provides a way to contextually search through data that you add to the investigation scope by using advanced orchestration and embeddings. Vector search is a search engine technology that focuses on understanding the meaning and context behind words and phrases in a query, rather than just matching keywords.

Some key aspects of vector search are:

- **Contextual understanding**: Vector search interprets the context of your search terms, considering factors like your organization, search history, and the overall meaning of the query.
- **Intent recognition**: Vector search works to understand your intent, whether you're looking for information, trying to take an action, or seeking a specific type of content associated with a search.

- **Relevance and accuracy**: By focusing on the semantics (the meaning and intent of words in your query), vector search provides more accurate and relevant results and improves the overall search experience.

- **Suggested search (preview)**: Vector search also enables customized suggested search queries for you to select based on the investigation context or any previous searches in an investigation. If your investigation doesn't have a context defined and you don't have any previous searches, you can choose from the following default suggested searches:
  - Look for anything containing credentials or passwords
  - Find all content containing sensitive information
  - List all legal and financial documents

When investigators in your organization investigate compromised data sets, vector search in Data Security Investigations can significantly enhance your investigation by addressing several key challenges:

- **Identifying relevant information**: Vector search understands the context and intent behind your queries. This focus helps you quickly locate relevant documents, emails, or records, even if they don't contain the exact keywords you used.
- **Handling ambiguity**: Vector search disambiguates terms that have multiple meanings, ensuring that you get results that are contextually appropriate to your investigation.
- **Reducing noise**: Vector search filters out irrelevant information, so you can focus on the most pertinent data and reduce the time spent sifting through unrelated results.
- **Improving efficiency**: Vector search streamlines the search process, making your investigation more efficient and effective by quickly surfacing the most relevant information.

How it works

When you create an investigation and define the scope, the system automatically prepares data for AI. Starting June 2026, add to scope automatically encompasses the previous **Prepare for AI** action for all new investigations. You can then run vector searches over the data set. While previous steps of the process allow for simple keyword, metadata, and date range search, vector search uses AI embeddings to contextually search through data. This process allows investigators to find items without knowing their exact content.

> **Important**
>
> Items are vectorized if they contain text or supported image content with extractable text. For JPEG, PNG, and scanned PDF files, Optical Character Recognition (OCR) automatically

extracts text from images before vectorization. Meeting invites and calendar notifications without text in the body are excluded from vectorization.

Vector search works by first running all scoped data in an investigation through an AI embeddings model. This model extracts semantic meaning from every item in your data set and breaks them into smaller parts. This process is called embedding and allows Data Security Investigations to use dimension values to understand your data contextually. A semantic search index is built from these values that you can query.

When you create a vector search query in an investigation, AI automatically expands and broadens your query and runs the query through the semantic search index. Data Security Investigations then matches the semantic meaning of your query with the semantic meaning of your content and returns all contextually relevant items.

For example, if you search for "Confidential data included in the Contoso Security project", the vector search engine understands that you're looking for confidential data in this specific project rather than simply matching keywords (*confidential*, *data*, *Contoso*, etc.) contained in the search query. Using vector search, you can query impacted data to find all data items related to a particular subject, even if keywords are missing.

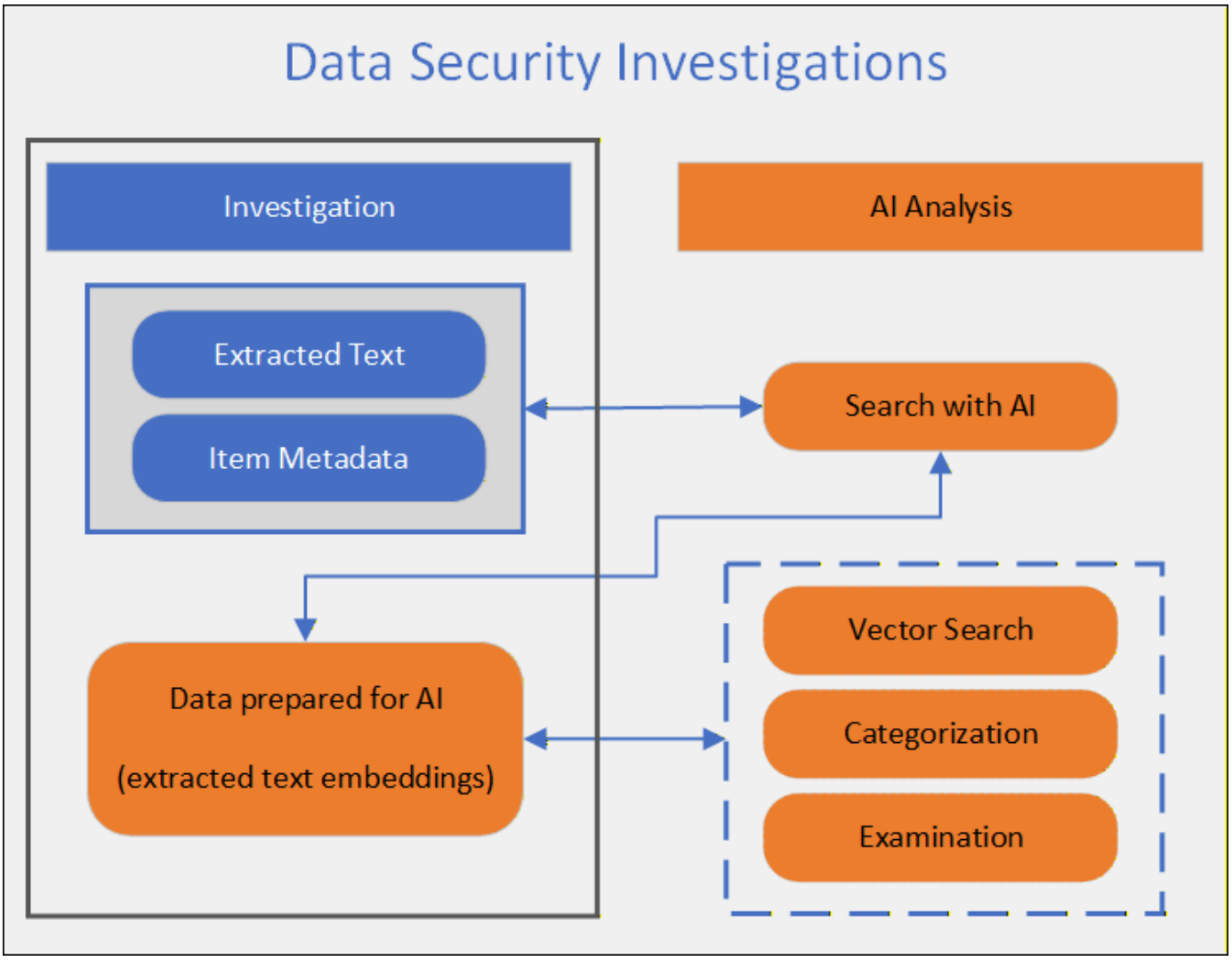
Additionally, a search relevance score is automatically assigned to each item returned. The search relevance score helps you quickly determine the confidence level of the connection between your search and the items identified by vector search.

For more information about vector search concepts, see the concepts section in the Vectors in Azure AI Search article.

Search with AI (preview)

Building on vector search, Search with AI (preview) incorporates retrieval-augmented generation (RAG) capabilities and optimizes the search results model. This feature includes search relevant items based on item file types, sizes, versions, and more.

This capability expansion to vector search also includes a search summary based on individual items returned in the search, including citation links to specific search results samples, to help you quickly evaluate if the search meets your needs.



Categorization

When your organization is breached and you identify the impacted data, investigators need to prioritize data to start identifying security risks. Data Security Investigations categorization speeds up investigations by automatically grouping content around the data security risks you care about, so analysts can narrow thousands of files down to a small set of prioritized, meaningful clusters - without manual review.

You can use **Standard** or **Advanced** AI-powered categorization in Data Security Investigations to more quickly reason through and prioritize potentially impacted data. To categorize data, you can select all or some default category options, use AI-suggested categories based on their investigation, or create your own custom categories.

You can also add categories incrementally after an initial categorization run completes. When you add new categories, only the net new categories are processed and previously categorized data is preserved. You can also change existing categories between **Standard** and **Advanced** processing levels.

The AI-generated categories are enriched with additional information for subject-level content in scope:

- **Name**: The name of the category or area based on the content
- **Summary**: A short description of the underlying content

Within each category, you can use vector search and examination tools on any content.

How categorization processes data

Categorization is designed to prioritize the most relevant content for each selected category rather than analyze every item in the investigation scope. When you run categorization, the system uses AI to evaluate content against your selected categories, scores each content segment for relevance, and retains the highest-confidence results. Content that doesn't meet the relevance threshold for a category isn't included in the results for that category.

Because categorization is optimized for speed and cost efficiency, be aware of the following considerations:

- **Results represent prioritized content**: Categorization surfaces a prioritized subset of the most relevant content for each category, not an exhaustive analysis of all items in your data set.
- **Content volume can affect results**: In large datasets, documents with extensive content might be more strongly represented in results. Some relevant items might not appear if they score below the relevance threshold.
- **Categories are processed independently**: Each selected category is evaluated separately. The number of categories you select affects processing time and compute unit costs but doesn't change how relevance is determined for individual categories.

When you select **Standard** categorization, the system groups content into your selected categories based on relevance scoring. When you select **Advanced** categorization, the system performs additional processing to identify and organize content into specific topics within each category, providing more granular groupings for deeper analysis.

> **Tip**
>
> If your investigation requires comprehensive analysis of all content in scope, use examination tools instead of or in addition to categorization. Categorization is best for fast

prioritization and triage of large datasets, while examination provides targeted, item-level analysis for selected items.

Default categories

Data Security Investigations includes default categories to categorize items in your investigation scope. When you run categorization, you can select all default categories or only the default categories that apply to the scope of your review. The analysis ignores unselected default categories, and you can't review results for these categories.

AI processing determines the initial default categories for content items:

- **Business information**: General business information. This category typically contains a large number of items. Some example areas in this category might include digital engagement and analysis, user and human resources, routine administrative communication, customer engagement/experience, and more.
- **Communication records**: General communication information. This category also typically contains a large number of items. Use this category to see your investigations based on areas of communications. Some example areas in this category might include client complaints, holiday greetings, internal memos, project updates, and more.
- **Credentials and access information**: Information related to access assets in investigations. This information helps identify potentially risky data and communications in your organization. Some example areas in this category might include user credentials, unauthorized database access, data exposure, and more.
- **Customer information**: Information shared with customers. Use this category to understand what customer data might be at risk. Some example areas in this category might include payment confirmations, customer experience improvement, delivery information, and more.
- **User information**: Information related to users in your organization. This category also typically contains a large number of items. Some example areas in this category might include user employment information, user retention strategies, specialized group memberships, and more.
- **Financial information**: Financial information in an investigation. Some example areas in this category might include financial planning, grant opportunities, budgets, financial statements, and more.
- **Health information**: Health and medical-related items in an investigation. Some example areas in this category might include wellness and health records, COVID-19 safety protocol updates, health claims and incident reports, and more.

- **Incident and investigation information**: Items about incidents and investigations in an investigation. This category includes security incidents and investigations within your organization. Some example areas in this category might include data breach, health records incidents, high-risk client account monitoring, and more.
- **Intellectual property**: Intellectual property (IP) data in an investigation. Some example areas in this category might include future patent applications, research and development work, experiment result metrics, and more.
- **Marketing information**: Marketing data in an investigation. Some example areas in this category might include press releases, advertising campaigns, marketing and sales plans, or strategies and more.
- **Operational information**: Your organization's operational data. Some example areas in this category might include logistics, shipping, inventory, compliance, tax records, and more.
- **Personally identifiable information**: Group personal data and related items in an investigation. Some example areas in this category might include event guest lists, staff and training sessions, employee personal information, and more.
- **Regulated data**: Regulated data in an investigation. Some example areas in this category might include regulatory, data protection, regulatory records, and more.

> ⓘ Note
>
> Investigations auto-created by proactive AI insights from Data Security Posture Management (DSPM) use five fixed risk categories: Credentials and access information, Personally identifiable information, Financial information, Confidential information, and Intellectual property. These categories are separate from the 13 default categories listed here and aren't configurable in the current preview.

Suggested categories

Data Security Investigations also provides AI-generated suggested categories based on the content analyzed in your investigation scope. These suggested categories automatically help investigations review items grouped in unexpected or unknown areas. Depending on the type of content included, the suggested categories vary.

If the content analyzed primarily focuses on a specific subject area outside of the default category areas, the suggested categories customize to that specific content area. For example, if the analyzed content focuses on a highly confidential subject with terms and concepts specific solely to your organization, the suggested categories are automatically created for these areas. These categories are unique to your organization and the content analyzed.

Custom categories

Data Security Investigations enables you to manually create custom categories for the generative AI process to use when analyzing content. By defining categories that best fit your investigation needs, you can save time and let the AI process automatically categorize items based on these custom categories.

Custom categories can be specific words or phrases that capture the specific nature of content you're interested in during the investigation. For example, custom categories might include *Security vulnerability*, *Bug fix*, specific project code names, or custom intellectual property like R&D related to a specific medicine or drug candidates.

Some additional custom categories that might be helpful in your investigations:

- Access codes
- API access documentation
- API authentication keys
- API tokens
- Application configuration files
- Certificate Authority (CA) details
- Certificates
- Database credentials
- Digital certificates
- Disaster recovery plans
- Employee location data
- Encryption keys
- Environment configuration files
- Incident Management Logs
- Integration tokens
- JWT tokens
- Key management policies
- Multifactor authentication (MFA) backup codes
- Personal Identification Numbers (PINs)
- Privileged account details
- Secure strings
- Security logs
- Security policies
- Session

- SSH private keys
- Third-party API secrets
- User credentials
- Vulnerability assessments

Examination

As you identify items that require deeper analysis, Data Security Investigations provides AI-based examination capabilities to help you focus on key security and sensitive data risks.

- **Credentials**: Use this examination focus area to scan and extract credentials from all selected items in an investigation scope. This information provides investigators with a quick way to understand which accounts and credentials are associated with a security incident and that might be potentially exfiltrated.

- **Risk**: Use this examination focus area to score all the risk areas in selected files to help investigators focus and prioritize investigations. This tool provides the overall risk for each item, if the item is privileged content, and other specific risks for the item.

  Types of risk areas include:
  - Asset identifiers
  - Credentials and secrets
  - Evidence of threat actor and breach discussions
  - Urgent security incidents
  - Vulnerability and security hygiene
  - Personal and sensitive content
  - Network and access information
  - Policy compliance and data protection
  - Infrastructure information
  - Customer information
  - Government information
  - Privileged information
  - Trade secrets

- **Mitigation**: Use this examination focus area to score the risk for selected files and enable Data Security Investigations to provide you with mitigation instructions for what to do next. Selected files get a risk score, risk summary, and detailed mitigation recommendations to prevent more harm from a content breach.

- **Personal data**: Use this examination focus area to identify and extract personal data from selected items in an investigation scope. This information provides investigators with a quick way to understand what personal data types are present in data associated with a security incident, including names, email addresses, employee IDs, IP addresses, and other PII.

- **Custom**: Use this examination focus area to define a custom analysis using a natural language prompt. Describe what you want the AI to look for in selected items, up to 2,000 characters. Custom examinations produce the same structured results as built-in focus areas, with findings validated by safety guardrails. For more information about creating custom examinations, see Create a custom examination.

AI analysis recommendations

The following table outlines recommendations, example scenarios, and best practices when using the AI analysis tools in Data Security Investigations.

Expand table

Recommendations	Vector search	Categorization	Examination
When to use	Look for examples of specific items within a vectorized data set (invoices, bug fixes, and more) to confirm and further investigation hypotheses. Use vector search for quick interactive analysis, results are quickly populated.	Quickly prioritize large amounts of data into default, custom, or AI-generated categories by sensitivity and severity. Categorization surfaces the most relevant content for each category rather than analyzing every item. Depending on the size of the data set, categorization might take some time to complete.	Comprehensive, targeted analysis at the item level for all selected data items. Helps extract detailed insights from confirmed data assets for next steps. Use examination to identify items for mitigation.
Example scenario	Evaluation of potentially fraudulent activity.	Prioritization of items for analysis after large breach.	Extraction of credentials and personal data from a validated data set and recommended mitigation steps.
Analysis scope	All vectorized content in the investigation scope.	Prioritized subset of the most relevant content for each selected category.	All selected items in the investigation scope.

Recommendations	Vector search	Categorization	Examination
Best practices	Search across all vectorized content for items of interest to generate more meaningful AI suggested categories.	Select one or multiple categories and use vector search to search within the category. Review AI-generated areas within each category to understand specific content within the data set.	Use examination to drill into specific items with high sensitivity to get individual scores and results.

Ready to get started?

- Learn about the Data Security Investigations workflow
- Get started with Data Security Investigations
- Application card for Data Security Investigations

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/11/2026

Get started with Data Security Investigations

Use Data Security Investigations to identify, investigate, and mitigate data risk either reactively when an incident occurs or proactively to improve data security hygiene for your organization. Complete the following steps in the workflow to set up prerequisites and configure Data Security Investigations.

For more information about how Data Security Investigations can help you with security incidents in your organization, see Learn about Data Security Investigations.

Step 1: Read and agree to terms

The first time you access Data Security Investigations in the Microsoft Purview portal, you must read and agree to the terms of the Privacy Statement.

Confirm that you accept the terms and select **Get started**.

Step 2: Configure permissions

To allow users to access Data Security Investigations tools in the Microsoft Purview portal, assign the users the appropriate permissions. The easiest way to assign roles is to use the setup task or add the user the appropriate role group on the **Role groups** page in the Microsoft Purview portal.

Members of the *Compliance Administrator* and *Organization Management* role groups automatically have Data Security Investigations administrative and contributor access. Members of the *Data Security Management* and *Insider Risk Management* role groups automatically have contributor access. For more information about access levels, see Additional role groups with Data Security Investigations access.

> **Important**
>
> To configure the AI capacity for Data Security Investigations, you must be a member of the *Data Security Investigations Admins* role group.

For step-by-step guidance, see Assign permissions in Data Security Investigations.

Step 3: Configure billing and usage

To get started with Data Security Investigations, configure billing and usage settings for data storage and AI analysis features. Data Security Investigations uses a payment model based on how much data you add to investigations and how much AI capacity you use for analysis of that data.

For step-by-step guidance, see Billing models in Data Security Investigations.

If you plan to enable proactive AI insights from Data Security Posture Management (DSPM), configure billing before enabling the toggle because the auto-created investigation incurs storage and AI costs immediately.

> **Tip**
>
> After setup completes, all users with Data Security Investigations roles receive a notification through the Microsoft Purview Notification Center and email confirming the solution is ready. For more information, see **Email and portal notifications**.

Step 4: Create an investigation

To get started with Data Security Investigations, you must create an investigation and configure investigation settings.

You can create investigations in several ways. Choose from the following methods to create a new investigation:

- From a Microsoft Defender XDR incident.
- From an Insider Risk Management case.
- From Data Security Posture Management insights.
- Manually using the full draft mode.

The user who creates the investigation is automatically added as a member. Members of the investigation can access the investigation in the Microsoft Purview portal and perform Data Security Investigations tasks.

To configure user access and permissions for specific investigations, see Configure investigation access and permission settings in Data Security Investigations.

Step 5: Search and evaluate results

After creating an investigation, use search tools to identify content such as email, documents, and instant messaging conversations in your organization that are relevant to a data security incident.

> **Tip**
>
> Refining and narrowing searches to key areas of your investigation is an iterative process. Narrow the results as much as possible before adding data items to the investigation scope. Use filters to identify non-relevant items, select the items, and mark the items as excluded.

Search tools allow you to:

- Define custom data sources
- Create custom searches with the query builder
- View scope and sample dashboards for search result items

Step 6: Add items to an investigation scope

After you add data items from your search queries to the scope of the investigation, the system automatically vectorizes items for AI analysis.

You can group and view items in the investigation scope, including file detail views for each data item. After automatic vectorization completes, you can use AI-powered tools for advanced analysis.

Step 7: Investigate items

Use generative AI processing to help you quickly identify and review data items in the investigation scope. These tools help you decide which items need closer examination and if they should be added to the mitigation plan.

- Vector search
- Categorization
- Examination

Step 8: Take mitigation actions

Create detailed examinations and review risk examinations and recommendations for items in the investigation scope:

- Review mitigation recommendations

- Review credentials examinations
- Review risk examinations
- Review personal data examinations
- Use the mitigation plan

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/11/2026

Get started with the Data Security Posture agent (preview) in Data Security Investigations

> **Important**
>
> Data Security Investigations uses generative artificial intelligence (AI), large language models, and orchestration in the analysis of data in your organization. Results generated by AI might not always be accurate or complete. While we strive to provide reliable and helpful information, AI systems can produce incorrect or false results. It is important to verify the information and use it with caution. Microsoft makes no warranties, express, implied, or statutory, as to the information provided by AI systems.

The Data Security Posture agent (preview) in Data Security Investigations helps your organization proactively surface credentials buried in data across your organization at scale. Rather than manually investigating individual items, the Posture agent automates credential scanning across Microsoft 365 data locations, including SharePoint sites, OneDrive accounts, Exchange mailboxes, and Teams content. After the agent completes scanning, it produces AI-generated risk assessments with confidence scores and supporting reasoning to help your security team quickly prioritize and remediate credential exposure risks.

> **Important**
>
> The Posture agent in Data Security Investigations is currently in preview. Features and functionality might change before general availability.

How the Posture agent works

The Posture agent in Data Security Investigations uses the same agent as Data Security Posture Management (DSPM) (preview). When you enable the Posture agent, you make it available in both DSPM and Data Security Investigations. If you already enabled the Posture agent in DSPM, you don't need to onboard again to use it in Data Security Investigations.

By using the Posture agent in Data Security Investigations, you can:

- **Scan for credentials at scale**: Run tenant-wide or scoped credential scanning tasks to discover exposed credentials across your organization's Microsoft 365 data.

- **Track tasks on a task board**: Monitor the progress of scanning tasks through a Kanban-style task board with statuses like *In progress* and *Ready for review*.
- **Review AI-generated risk assessments**: Examine AI-produced risk scores, confidence levels, and reasoning for flagged items to help prioritize remediation efforts.
- **Explore data with KQL queries**: Use the Data Explorer to run Kusto Query Language (KQL) queries against credential findings for advanced analysis and investigation.
- **Download reports**: Generate and download reports of credential scan results for offline review and compliance reporting.

Prerequisites

Before you can use the Posture agent in Data Security Investigations, confirm the following prerequisites are in place:

- **Permissions**: You must be assigned to one of the role groups that includes Posture agent access. For more information, see Permissions for deploying, running and viewing results from the Posture Agent.
- **Agent enabled**: You must enable the Posture agent in the Microsoft Purview portal. For more information, see Enable the Posture agent.
- **Security Compute Units (SCUs)**: Your organization must have Security Copilot SCUs provisioned for the Posture agent to run.

> **Tip**
>
> You don't need to configure or enable Data Security Investigations to use the Posture agent. The Posture agent is independent of the Data Security Investigations solution and doesn't require Data Security Investigations billing, permissions, or setup to be completed. As long as the Posture agent is enabled and your organization has SCUs provisioned, the Posture agent experience is available in the Data Security Investigations navigation. Other Data Security Investigations features, such as creating investigations and AI analysis, still require separate Data Security Investigations configuration and billing.

Enable the Posture agent

If your organization hasn't already enabled the Posture agent, use the following steps to enable it:

1. Go to the Microsoft Purview portal and sign in by using the credentials for a user account assigned the appropriate permissions.
2. Select **Agents** in the left navigation.

3. Select **Explore agents**, and then select **View details** for the Posture agent in Data Security Investigations.
4. Select **Set up** to enable the agent for your organization.

After you enable the Posture agent, it appears in Agents, DSPM, and Data Security Investigations. In Data Security Investigations, you can access the agent by selecting **Posture agent** in the left navigation.

> **Note**
>
> If you already enabled the Posture agent in DSPM, you don't need to complete these steps. The agent is automatically available in Data Security Investigations.

Create a credential scanning task

Create a credential scanning task so the Posture agent can scan your organization's data for exposed credentials. You can scope the scan to your entire tenant or narrow it to specific users, sites, or groups.

Complete the following steps to create a credential scanning task:

1. Go to Data Security Investigations in the Microsoft Purview portal and sign in by using the credentials for a user account assigned Data Security Investigations permissions.
2. Select **Posture agent** in the left navigation.
3. Select **Assign to agent** to create a credential scanning task.
4. For the **Credential Scanning Task**, enter a name for the task in the **Task name** field.
5. Configure the data sources for the scan. You can select a tenant-wide scan or narrow the scope to your organization, specific sites, users, mailboxes, or groups.
6. Select **Save**.
7. In the **Additional context for AI**, use natural language guidance to help AI scope the agents to applicable areas.
8. Select **Create** to start the task.

After you create the task, it appears in the **In progress** column on the task board.

> **Tip**
>
> If you realize you scoped a task incorrectly after it starts, you can stop the task and create a new one with the correct scope.

Monitor tasks on the task board

The Posture agent uses a Kanban-style task board to help you track the progress of credential scanning tasks. Each task appears in one of the following columns:

- **In progress**: The agent actively scans the specified data locations and analyzes content.
- **Ready for review**: The agent finished scanning and the results are ready for a human reviewer to examine.
- **Closed**: A human reviewer has marked a tasked as completed.

The task board displays key information for each task, including the scan scope, the number of locations scanned, and the number of items analyzed. You can also assign a security operations center (SOC) administrator to review the results for a specific task.

Review scan results

After a credential scanning task finishes and moves to the **Ready for review** column, you can review the results to identify and prioritize credential exposure risks.

Summary view

The summary view provides a high-level overview of the credential scan results. In the summary view, you can:

- **Sort by risk level**: View results sorted by *High*, *Medium*, or *Low* risk to prioritize the most critical findings.
- **Group by credential category**: Organize results by credential type to understand the distribution of exposed credentials.
- **Review AI reasoning**: For each flagged item, the AI provides its reasoning for the assigned risk level, helping you understand why a specific item is flagged as high, medium, or low risk.
- **Download reports**: Generate a report of the scan results for offline review, compliance reporting, or sharing with stakeholders.

Item details

Select a specific item in the scan results to view detailed information, including:

- The detected credential type.
- The data location where the credential was found.
- The AI-generated risk and confidence scores.

- The reasoning behind the risk classification.

Explore data with Data Explorer

Data Explorer provides advanced analysis capabilities by allowing you to run Kusto Query Language (KQL) queries against the credential findings from the Posture agent scans. This experience is similar to advanced hunting in Microsoft Defender XDR and provides a schema-based metadata framework for querying credential data.

To use Data Explorer, complete the following steps:

1. Go to Data Security Investigations in the Microsoft Purview portal.
2. Select **Posture agent** in the left navigation.
3. Select a completed task, and then select the **Data Explorer** tab.
4. Write and run KQL queries against the credential findings schema.

> **Tip**
>
> If you're new to KQL, see **Kusto Query Language overview** to learn the basics of writing queries.

ⓘ **Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 03/23/2026

Assign permissions in Data Security Investigations

To enable users to access Data Security Investigations capabilities and features in the Microsoft Purview portal, assign users the appropriate permissions.

This article describes the permissions required to perform specific Data Security Investigations tasks. For more information about Microsoft Purview role groups and permissions, see Permissions in the Microsoft Purview portal.

Roles and role groups

> **Important**
>
> After you configure your role groups, it might take up to 30 minutes for the role group permissions to apply to assigned users across your organization.

To quickly assign the *Admin* or *Investigator* roles to users, use the **Assign roles to your team members** setup task. Complete the following steps to assign these roles:

1. Go to Data Security Investigations in the Microsoft Purview portal with an account assigned the *Global Administrator* role.
2. Select **Overview** in the left navigation.
3. Select **Assign roles to your team members** in the **Setup tasks** section.
4. Select users in the **Admin** field on the **Role assignment** flyout pane.
5. Select users in the **Investigators** field.
6. Select **Confirm** to assign users these roles.

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

You can also add users to the appropriate role group on the **Role groups** page in the Microsoft Purview portal. To continue with these configuration steps in the Microsoft Purview portal, you

must be assigned to the *Data Security Investigations Admins* role group.

Configure permissions

To manage Data Security Investigations workflow and investigations, assign users to specific role groups to manage different sets of Data Security Investigations features. You can assign users with different security responsibilities to specific role groups to manage different areas of Data Security Investigations features. Or, assign all user accounts for designated admins, investigators, and reviewers to the *Data Security Investigations Admins* role group. Use a single role group or multiple role groups to best fit your security management requirements.

> **Tip**
>
> Always have at least one user in the *Data Security Investigations Admins* role group so that your Data Security Investigations configuration doesn't get in to a 'zero administrator' scenario if specific users leave your organization.

Choose from these solution role group options when configuring and managing Data Security Investigations:

Expand table

Actions	Data Security Investigations Admins	Data Security Investigations Investigators	Data Security Investigations Reviewers
Add, delete, manage items for a mitigation plan	Yes	Yes	Yes
Create and manage all investigations	Yes	No	No
Create and manage assigned investigations	Yes	Yes	No
Create searches and add items to an investigation	Yes	Yes	No
Estimate and preview search results	Yes	Yes	No
Manage investigation scope	Yes	Yes	No

Actions	Data Security Investigations Admins	Data Security Investigations Investigators	Data Security Investigations Reviewers
Run categorization activities	Yes	Yes	Yes
Run examination activities	Yes	Yes	Yes
Run vector searches	Yes	Yes	Yes
Create and run purge queries	Yes	Yes	No
View data risk graphs	Yes	Yes	Yes
View the Pay-as-you-go usage dashboard	Yes	No	No

> **Important**
>
> Global admins must be assigned to one of the listed role groups for Data Security Investigations to create and access investigations.

Additional role groups with Data Security Investigations access

In addition to the dedicated Data Security Investigations role groups, the following Microsoft Purview role groups automatically include Data Security Investigations roles. Members of these role groups have Data Security Investigations access without needing explicit assignment to a Data Security Investigations-specific role group.

Expand table

Role group	Data Security Investigations role included	Access level
Compliance Administrator	Data Security Investigation Admin	Full administrative access
Organization Management	Data Security Investigation Admin, Data Security Investigation Contributor	Administrative and contributor access
Data Security Management	Data Security Investigation Contributor	Contributor access
Insider Risk Management	Data Security Investigation Contributor	Contributor access

Members with the *Data Security Investigation Contributor* role can create and manage investigations they own or are members of. This role doesn't include access to data storage, AI job execution, or the Pay-as-you-go usage dashboard.

For more information about these role groups and their complete role assignments, see Roles and role groups in Microsoft Defender for Office 365 and Microsoft Purview.

Option 1: Assign multiple users to the Data Security Investigations Admins role group

Complete these steps to give multiple users administrative access and responsibilities for Data Security Investigations in your organization.

1. Go to the Microsoft Purview portal with an admin account in your Microsoft 365 organization.
2. Select **Settings** in the upper-right corner of the page, select **Roles and groups**, then select **Role groups** in the left navigation.
3. Select the *Data Security Investigations Admins* role group, then select **Edit**.
4. Select **Choose users**, then select the checkboxes for all the users you want to add to the role group.
5. Select **Select**, then select **Next**.
6. Select **Save** to add the users to the role group, then select **Done**.

Option 2: Assign users to specific Data Security Investigations role groups

Complete these steps to assign users to specific role groups to segment Data Security Investigations access and responsibilities among different users in your organization.

1. Go to the Microsoft Purview portal with an admin account in your Microsoft 365 organization.
2. Select **Settings** in the upper-right corner of the page, then select **Role groups** in the left navigation.
3. Select one of the Data Security Investigations role groups, then select **Edit**.
4. Select **Choose users**, then select the checkboxes for all the users you want to add to the role group.
5. Select **Select**, then select **Next**.
6. Select **Save** to add the users to the role group.

7. Select the next Data Security Investigations role group, then repeat the previous steps for each required role group.
8. Select **Close** when you're done.

ⓘ **Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 04/23/2026

Billing in Data Security Investigations

Billing in Data Security Investigations is based solely on the combination of the amount of stored data and the associated computing capacity needed for AI analysis. It isn't based on a dedicated enterprise plan or license.

This billing combination is based on the **storage meter** and **AI capacity** models. The Data Security Investigations non-AI processing and storage meter lets you store data related to an investigation. Data Security Investigations Compute Units (compute unit) measure the computational capacity needed to run the AI analysis within Data Security Investigations.

Storage meter

The storage meter for Data Security Investigations uses the pay-as-you-go billing model. You pay for storage on a gigabyte (GB) per month basis for stored data associated with all investigations in your organization.

For example, you have three investigations in your organization that have the following data associated with the items included in the investigation scopes for the current month:

Expand table

Investigation	Items	Total data
Data breach 02/12/2025	3,766	200 GB
Potential data theft 03/03/2025	1,982	10 GB
Risky user activity 04/06/2025	287	5 GB

For the current month, your total storage meter is 215 GB and you pay for that amount at the current per gigabyte rate. When you delete an investigation, you stop paying for the associated data storage amount. The charges are prorated to the number of days the data was stored in the billing cycle.

Some other important considerations for managing storage charges:

- If you create an investigation, use data storage, and then delete the investigation in a single 24-hour period, you pay for the storage amount.

- When you delete an investigation anytime after 00:00 UTC, you pay for storage for the full day (24 hours).
- Investigations auto-created by proactive AI insights from DSPM consume storage continuously while the toggle is enabled. Because the investigation refreshes every 24 hours, storage charges are ongoing.

For more data storage rates and pricing, see Data Security Investigations storage meter pricing.

AI capacity

> **Important**
>
> To configure the AI capacity for Data Security Investigations, you must be a member of the *Data Security Investigations Admins* role group.

Data Security Investigations uses the pay-as-you-go billing model for AI capacity. You need compute units for deep analysis and AI-related processing in your investigations. You can change the maximum limit of compute units ready to use at any time. You pay only for the compute units you use.

Proactive AI insights from DSPM also consume compute units. Each 24-hour refresh cycle uses compute units to generate AI insights across five fixed risk categories. For more information, see Enable proactive AI insights from DSPM.

For guidance on calculating the compute unit needs for your organization, see the Estimate and manage usage of Data Security Investigations Compute Units section in this article.

Locations for compute units and AI processing

Compute unit processing uses GPU resources in Azure datacenters protected by Azure security and privacy controls. You can choose where to use compute units from any of the following locations:

- Australia/New Zealand (ANZ)
- Europe (EU)
- United Kingdom (UK)
- United States (US)

You can opt in to having compute units provisioned and AI analysis processing done anywhere in the world to mitigate potential disruptions in case your primary location experiences high activity.

Microsoft recommends having compute units and AI processing done anywhere with available GPU capacity. This recommendation enables Data Security Investigations and Azure to determine the optimal location based on load, latency, and responsiveness.

Best practices for managing investigation costs

Planning and managing costs is an essential part of running effective Data Security Investigations. By estimating costs up front, applying safeguards during configuration, and continuously monitoring usage, you can control spending without limiting your ability to respond to security incidents.

The following best practices help ensure investigations remain cost-efficient while still providing the flexibility required for thorough and timely analysis.

1. **Use the in-product cost estimator before starting investigations**: Before launching an investigation, use the in-product cost estimator to anticipate costs based on investigation size and the analyses you plan to conduct. This process helps data security teams plan budgets with precision and prevent unforeseen expenses.

2. **Use built-in spending controls in Azure**: Use the powerful built-in spend controls included with your Azure subscription, including:

    - Create budgets for the Data Security Investigations resource group and meter level.
    - Configure alerts when spending approaches defined thresholds (for example, 50%, 75%, or 90%).
    - Use Azure Cost Analysis for drill-downs and anomaly detection.

3. **Scope investigations tightly**: Focus on the investigation scope to lower storage and compute costs. Start with targeted searches and expand only as needed based on initial findings.

4. **Delete completed investigations promptly**: When you delete an investigation, you no longer incur costs for using data storage. Charges are prorated to the number of days the data was stored in the billing cycle. Regularly review and remove investigations that you no longer need to minimize ongoing storage costs.

5. **Monitor usage with the usage dashboard**: Use the usage dashboard to track storage and compute unit spending within Data Security Investigations in real time. This granular visibility helps identify cost-saving opportunities and optimize resource allocation across investigations.

6. **Right-size your AI analysis**: Not every investigation requires the full suite of AI capabilities. Consider which analyses (categorization, vector search, examination) are truly necessary for each investigation to optimize compute unit consumption.

7. **Add categories incrementally**: When using categorization, start with a focused set of categories and add more as your investigation progresses. When you add new categories to an existing categorization, you're only charged for the compute units needed to process the net new categories. This approach helps you manage costs effectively while expanding your analysis over time.

8. **Budget for proactive AI costs from DSPM**: If you enable proactive AI insights from DSPM, the auto-created investigation refreshes every 24 hours, incurring ongoing storage and AI compute charges. Monitor usage regularly and factor these recurring costs into your monthly budget.

Estimate your investigation data storage and AI costs

Before configuring billing, use the estimate costs tool to calculate an estimate of potential monthly data storage and AI costs for investigations in Data Security Investigations. Use the estimator tool to plan your costs before setting up data storage and compute units for AI features.

Use the tool to estimate individual investigation costs to create an average investigation cost or use the tool to estimate the costs of all anticipated investigations. This estimate is based on current storage and AI usage billing meters for a 30-day period.

Use the guidelines in the Investigation size and complexity section in this article to help complete the fields in the estimator tool.

To estimate investigations costs, complete the following steps:

1. Go to Data Security Investigations in the Microsoft Purview portal and sign in with the credentials for a user account assigned Data Security Investigations permissions.
2. On the **Estimate storage and AI usage costs** card, select **Estimate costs**.
3. Complete the required and optional fields for a single investigation to calculate an estimated average cost or for the total number of investigations in your organization.
4. Select **Calculate**.
5. Review the estimated costs and recalculate as needed.

Configure billing

Before you can use Data Security Investigations in your organization, you must configure both billing models. Storage meter billing requires a link to an Azure subscription for pay-as-you-go billing. Capacity billing requires dedicated compute unit configuration.

Prerequisites

To configure the storage meter and capacity for Data Security Investigations, you must have:

- An Azure subscription in the same tenant as Microsoft Purview.
- An Azure resource group in that subscription.
- The *Global Administrator* role.
- The resource group *Owner* or *Contributor* role.

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

Configure storage meter

The Azure subscription you configure for Data Security Investigations manages all pay-as-you-go capabilities across Microsoft Purview solutions. If you previously configured an Azure subscription for other Microsoft Purview solutions, you don't need to configure an additional subscription.

To configure the storage meter for Data Security Investigations, complete the following steps:

1. Go to Data Security Investigations in the Microsoft Purview portal and sign in with the credentials for a user account assigned Data Security Investigations permissions.
2. Select **Setup pay-as-you-go-billing for data storage** in the **Setup tasks** section.
3. On the **Setup pay-as-you-go-billing for data storage** flyout pane, complete the following fields for the Azure subscription:
   a. **Azure subscription**: The Azure subscription you want to assign to the storage meter.
   b. **Resource group**: The resource group in your Azure subscription you want to assign to the storage meter.
4. Select **Enable**.

Configure AI capacity

When you configure compute units, you enable the AI capacity for investigations in your organization. To configure AI capacity for Data Security Investigations, complete the following steps:

1. Go to Data Security Investigations in the Microsoft Purview portal and sign in with the credentials for a user account assigned Data Security Investigations permissions.
2. Select **Setup pay-as-you-go to use AI** in the **Setup tasks** section.
3. On the **Set up AI capacity** pane, complete the following fields:
   a. **AI analysis data processing location**: The primary country/region. Select from the available locations.
4. In the **Number of compute units** section, select one of the following options:
   a. **No limit**.
   b. **Set maximum limit**.
5. Select the check box to confirm that you read, understood, and agreed to the terms and conditions.
6. Select **Confirm**.

Estimate and manage usage of compute units

To manage the usage of compute units in Data Security Investigations, estimate your likely compute unit usage and the general size of the investigation. While it's challenging to estimate the exact amount of these variables, having a general sense of the investigation processing, size, and complexity helps you make informed decisions about provisioning and scaling compute unit usage.

Investigation processing

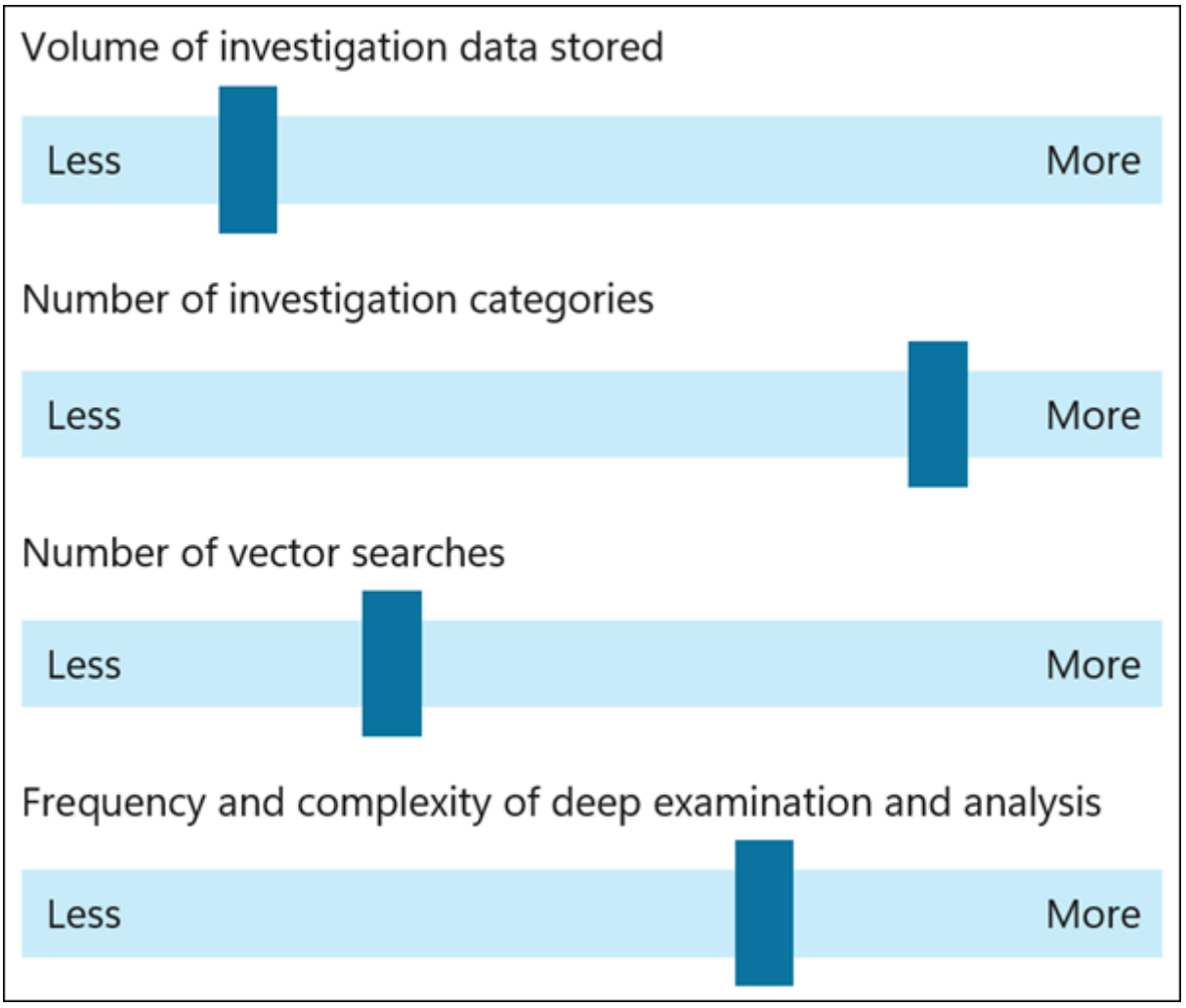
Data Security Investigations supports various processing tools to help you identify and take action on data security incidents. Each processing tool has different variables that you should consider when estimating support for each processing type:

Expand table

Process type example	Scale units or variables
Add to scope (including data preparation for AI analysis)	Size of data and extracted text
Vector searches	Size of data searched

Process type example	Scale units or variables
Categorization	Number of new categories selected. Adding categories to an existing categorization only processes net new categories.
Examination	Number of files analyzed

These variables primarily influence the cost for compute unit billing. They work together to determine the exact amount of compute unit usage per investigation. Each investigation might have a different combination of these cost influencers, so determining the exact compute unit billing estimation is challenging.



Investigation size and complexity

Each investigation uses different amounts of storage and processing power. To estimate the compute unit usage for an investigation, consider the overall size of the investigation. In general, investigations fall into one of the following sizes:

- **Small investigation**: A small investigation to test usage as a proof of concept in your organization.

- **Medium investigation**: A medium investigation, such as a data audit or responding to a minor data security incident.
- **Large investigation**: A large investigation for a significant data breach or leak incident.
- **Extra large investigation**: An extra large investigation for a major data security breach, leak, or incident.

To estimate the total compute units your organization needs for specific types of investigations, use the investigation processing, size, and complexity guidelines in this section. See the following sections during your planning and configuration:

- Add to scope
- Vector search and compute units
- Categorization and compute units
- Examination and compute units

> ⓘ **Important**
>
> You aren't billed for failed compute unit operations but are billed for partial completions. For example, you pay when you successfully vectorize partial data in an investigation scope. If you cancel an in-progress job (such as **Add to scope**, Categorization, or Examination), you still pay for any compute units consumed before the cancellation.

Pay-as-you-go usage dashboard

Use the **Pay-as-you-go usage dashboard** to quickly view and track data storage and compute unit usage for each investigation in your organization. The current data storage and compute unit totals for each investigation help you manage usage and cost per investigation. This information is helpful during reactive investigation surges associated with severe data breaches and ongoing data storage trends.

> ⓘ **Note**
>
> It might take up to 48 hours for the latest billing data to show up on the **Pay-as-you-go usage dashboard** for your organization. Only users assigned the *Data Security Investigations Admin* role can access the **Pay-as-you-go usage dashboard**.

To view the **Pay-as-you-go usage dashboard**, complete the following steps:

1. Go to Data Security Investigations in the Microsoft Purview portal and sign in with the credentials for a user account assigned to the Data Security Investigations Administrator role group.
2. Select **Usage dashboard** in the left-navigation.

From the **Pay-as-you-go usage dashboard**, you can review:

- The data storage usage for each investigation and the total for each process.
- The total amount of compute units used by Data Security Investigations processes.
- The compute units per day, for the last 7, 14, or 30 day timescale.
- The compute unit usage for each investigation and the total for each process.

Each investigation has the total data storage for each of the following processes:

- **Add to scope**: Storage and compute supporting processes when adding items to the investigation scope, including automatic vectorization.
- **Categorization**: Storage and compute supporting categorization processes for an investigation.
- **Examination**: Storage and compute supporting examination processes for an investigation.
- **Interactive AI**: Storage and compute supporting semantic search and AI suggested categories.

To export the data storage and compute information for all investigations, select **Export**. The export file lists the following information:

- **ID**: The ID of the investigation.
- **Date**: The date the report was created for each investigation.
- **Storage**: The total data storage for each process in the investigation.
- **InvestigationName**: The friendly name of the investigation.
- **JobType**: The type of process associated with the data storage or compute usage.
- **JobID**: The ID for the process.
- **ConsumedUnits**: The amount of compute units used for the process.

Stop billing

To stop billing for Data Security Investigations, disable both storage meter and AI capacity support. This action deletes all investigations in Data Security Investigations and deletes compute units for your resource group.

> ⊗ **Caution**

Deleting investigations and your compute unit resource group is irreversible. Make sure you export or preserve any investigation results or reports before stopping billing.

Stop storage meter billing

Data storage charges start when you add items to the scope of an investigation. To stop storage metering, complete the following steps:

1. Go to Data Security Investigations in the Microsoft Purview portal and sign in with the credentials for a user account assigned to the Data Security Investigations Administrators role group.
2. Delete all existing investigations in your organization.
3. Verify that no active or pending investigations remain in your organization.

When you delete all investigations, you remove the associated data and stop storage billing for Data Security Investigations.

Stop AI capacity billing

AI analysis is billed through compute units in Microsoft Azure. To stop compute unit billing, complete the following steps:

1. Go to Data Security Investigations in the Microsoft Purview portal and sign in with a user account that has either the **Owner** or **Contributor** role for the Azure subscription associated with the Data Security Investigations compute unit resource.
2. Go to the **Subscription** where you deployed your Data Security Investigations resource.
3. Select the associated **Resource group** for Data Security Investigations.
4. Select and delete the **DSI compute unit capacity resource** (for example, `DataSecurityInvestigation-GUID`).

When you delete the compute unit resource, all AI capacity and compute billing stops for Data Security Investigations.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/11/2026

Create an investigation in Data Security Investigations

After configuring billing and usage options for Data Security Investigations and assigning permissions to analysts in your organization that manage investigations, create and manage an investigation for collection and analysis activities.

Investigations dashboard

Depending on your permissions, the **Investigations** dashboard in the Microsoft Purview portal displays investigations in your organization. If you're not assigned to the *Data Security Investigations Admins* role group, you can only view investigations you're assigned to as an investigator or reviewer.

The **Investigations** dashboard lets you quickly see the investigations in your organization that you have access to, the description of the investigations, and important information associated with each investigation.

- **Name**: The name of the investigation. The investigation name must be unique in your organization.
- **Description**: The description entered when the investigation was created.
- **Created on**: The date and time in Coordinated Universal Time (UTC) when the investigation was created.
- **Last modified**: The date and time in Coordinated Universal Time (UTC) when the investigation was last modified.

Select **Edit columns** to customize the displayed columns and the order of the columns in the **Investigations** dashboard. Choose the columns to display or drag and drop columns to reorder.

AI context for investigations

Additional context for AI helps focus AI-provided results in an investigation to specific areas, subjects, and issues to help eliminate noise and decrease the time it takes to discover and mitigate relevant items. Without additional context, AI analysis might return nonrelated categories and risk examinations results for data prepared for AI in your investigation. By providing detailed context for an investigation, categorization and risk examination results more accurately represent the areas you need to investigate.

For example, if you want an investigation to focus primarily on bribery issues, consider including *Find instances of bribery* in the **Additional context for AI** field on the **Create an investigation** dialog when creating new investigations in full-draft mode, from Defender XDR, or from Insider Risk Management. The more details that you provide as the context for the investigation, the more focused the AI analysis results are in the investigation.

> **Tip**
>
> Use natural language to provide the context and details you're looking for in the investigation. The context is automatically processed to remove typos, incorrect grammar or punctuation, symbols, and more to help improve clarity and readability for backend AI processing. This context review doesn't change the original intent of the provided context.

Any AI-related analysis for an investigation includes the context you provide for a new or existing investigation. Context-related categories are automatically included in AI-suggest categories and risk examinations for investigations including a custom context statement. You can add or edit context for an investigation at any time, but you must rerun categorization and risk examinations after any context changes. All changes to context statements are automatically recorded as extended properties for an investigation.

Create an investigation

Depending on your scenario, you can create investigations by using each of the following methods:

- **From Microsoft Defender XDR incidents**: Create an investigation from a Defender XDR incident.
- **From Microsoft Insider Risk Management cases**: Create an investigation from an Insider Risk Management case.
- **From Data Security Posture Management insights**: Create an investigation from exfiltration activities identified in Data Security Posture Management.
- **Manually by using full draft mode**: Create an investigation by using the *full draft mode* option to configure specific data sources and search conditions.

Create an investigation from Defender XDR

With the integration with the Microsoft Defender XDR solution, you can quickly open an investigation in Data Security Investigations as part of your response to data breach incidents.

You can create an investigation from Defender XDR incidents containing mailbox, email message, or file nodes. To create investigations in Data Security Investigations in the Microsoft Defender portal, you must have the following roles assigned:

- Security Administrator
- Security Operator

Create a new investigation from a Defender XDR incident by using one of the following methods:

- Select **Create investigation** from the Data Security Investigations banner at the top of any incident in Defender XDR. This banner automatically displays if the incident might contain sensitive information. The new investigation contains all nodes included in the Defender XDR incident.
- Select **Create Data Security Investigation** from the ellipsis control in the top right of an incident page. The new investigation contains all nodes included in the Defender XDR incident.
- Select any individual mailbox, email message, or file node in the Defender XDR incident to display the node menu and select **Create Data Security Incident**. The investigation contains only the selected node from the Defender XDR incident.

After you select one of the previous options, the **Create a data security investigation** dialog is displayed. Complete the following steps:

1. In the **Name** field, give the investigation a name (required). The investigation name must be unique in your organization. If the name you enter isn't unique, you're notified when you select **Create**.
2. In the **Description** field, add an optional description to help others understand this investigation.
3. In the **Scope** field, choose the items from the incident to include in the investigation.
4. In the **Additional context for AI** field, enter additional context to help AI focus your investigation on specific areas and issues.

> ⓘ **Important**
>
> You can't include mailboxes in the scope of an investigation and also include files or email messages. Investigations created from incidents including mailboxes must be standalone investigations. Select either mailboxes in the scope area OR email messages and files. Both email messages and files from an incident can be included in the same investigation.

1. Select **Create** to create the investigation.

The items included in the scope of the investigation are automatically included as data sources and you're ready to start reviewing and adding these items to the investigation scope in Data Security Investigations.

> **Note**
>
> Sometimes the data sources for SharePoint site or files might not autopopulate in an investigation. If this situation occurs, manually add the data sources and use the suggested query.

> **Important**
>
> Investigations created from Defender XDR incidents that contain more than approximately 3,000 items might not return complete results. The exact limit depends on the file names and paths of the returned items. For more information about limits, see Limits in Data Security Investigations.

Create an investigation from Insider Risk Management

With the integration of Microsoft Purview Insider Risk Management, you can quickly open an investigation in Data Security Investigations from a case and for potentially risky activities. To create investigations in Data Security Investigations in Insider Risk Management, you must have at least one of the following roles assigned:

- Insider Risk Management
- Insider Risk Management Investigator

To create a new investigation from an Insider Risk Management case, complete the following steps:

1. In the Microsoft Purview portal, go to **Insider Risk Management** > **Cases** and select a case.
2. On the **Cases** page, select **Case actions**.
3. Select **Investigate data security with AI**.
4. In the **Create a data security investigations** dialog box, complete the following fields:

a. In the **Name** field, enter a name for the investigation (required). The investigation name must be unique in your organization. If the name you enter isn't unique, you're notified when you select **Create investigation**.
b. In the **Description** field, add an optional description to help others understand this investigation.
c. In the **Investigation scope** field, choose the items from the case to include in the investigation.
d. In the **Additional context for AI** field, enter additional context to help AI focus your investigation on specific areas and issues.

5. Select **Create investigation** to create the investigation.

The items included in the case are automatically included as data sources. You're ready to start reviewing and adding these items to the investigation scope in Data Security Investigations.

> ⓘ **Important**
>
> Investigations created from Insider Risk Management cases that contain more than approximately 3,000 items might not return complete results. The exact limit depends on the file names and paths of the returned items. For more information about limits, see Limits in Data Security Investigations.

Create an investigation from Data Security Posture Management

By using the integration with the Microsoft Purview Data Security Posture Management (DSPM) solution, you can investigate data exposure insights surfaced by DSPM exfiltration objectives. The **Prevent exfiltration to risky destinations** objective card supports two investigation modes:

- **Manual investigation**: Create an investigation on demand from the exfiltration objective card to analyze specific data exposure insights.
- **Proactive AI insights (preview)**: Enable an opt-in toggle that automatically creates and refreshes a single investigation for your tenant every 24 hours. This investigation continuously analyzes recently exfiltrated sensitive data across five fixed risk categories: credentials and access information, personally identifiable information (PII), financial information, confidential information, and intellectual property (IP). The DSPM exfiltration objective card displays the count of items identified in each risk category.

Investigations created from DSPM automatically scope to all sensitive data exfiltrated from your organization in the last 30 days. These investigations enable deeper, AI-powered content analysis of recently exfiltrated data, giving security teams actionable insights without extra configuration.

For more information about the roles and permissions needed to create an investigation in Data Security Investigations from DSPM, see Permissions for Microsoft Purview Data Security Posture Management.

Create a manual investigation from DSPM

To manually create a new investigation from DSPM, complete the following steps:

1. In the Microsoft Purview portal, go to **DSPM**.
2. In **Objectives**, select the **Prevent exfiltration to risky destinations** card.
3. Select **View objective** to view exfiltration details and proactive insights.

   > **Note**
   >
   > It might take 24 hours or more for insights to display for the objective.

4. On the **Exfiltration risk patterns** card, select **Create investigation**.
5. In the **Create a data security investigation** dialog, complete the following fields:

   - **Name**: Enter a unique name for the investigation (required).
   - **Description**: Provide a description to help others understand the purpose of the investigation (optional).
   - **Add context for AI**: Provide context for the investigation (optional).

6. Select **Create investigation** to create the investigation in Data Security Investigations.

After the investigation is created, it automatically includes all scoped data items as data sources. You're ready to start reviewing, analyzing, and taking mitigation actions by using the full capabilities of Data Security Investigations. If you didn't configure Data Security Investigations before creating the investigation, the investigation is created but you need to configure Data Security Investigations before you can view the investigation.

The most recent manual investigation that you created from DSPM (preview) is always available for selection on the **Exfiltration risk patterns** card in the **Prevent exfiltration to risky destinations** objective.

Enable proactive AI insights from DSPM

To enable proactive AI insights, complete the following steps:

1. In the Microsoft Purview portal, go to **DSPM (preview)**.
2. In **Objectives**, select the **Prevent exfiltration to risky destinations** card.
3. Select **View objective** to view exfiltration details.
4. On the **Exfiltration risk patterns** card, turn on the **Proactive AI insights** toggle.

After you enable the toggle, the first investigation is created automatically. Insights might take up to 24 hours to appear after initial activation. The investigation refreshes automatically every 24 hours while the toggle remains enabled.

> **Important**
>
> Proactive AI insights consume Data Security Investigations storage and AI analysis meters while the toggle is enabled. In the current preview, proactive AI insights use five fixed risk categories. Custom category templates and per-item deletion within the auto-created investigation aren't supported. For more information about billing, see Data Security Investigations billing.

Create an investigation using full draft mode

In some scenarios, you might prefer to create a new investigation that isn't associated with a Microsoft Purview Insider Risk Management or Microsoft Defender XDR incident. By using full draft mode, create a new investigation that starts with new data sources and queries.

Manage investigation settings

Investigation settings include investigation information and access permissions. You can access settings for a specific investigation by selecting **Investigation settings** after selecting an investigation.

For more information about investigation settings, see:

- Investigation settings
- Access and permissions settings

Note: The author created this article with assistance from AI. Learn more

Last updated on 06/22/2026

Create an investigation using full draft mode in Data Security Investigations

When you create an investigation by using full draft mode in Data Security Investigations, you get the most flexibility and options for analyzing a data security incident. With this process, you can customize the data sources and query in your investigation.

Use full draft mode

Complete the following steps to create an investigation and configure the investigation scope by using the *full draft mode*. The user who creates the investigation is automatically added as a member. Members of the investigation can access the investigation in the Microsoft Purview portal and perform Data Security Investigations tasks.

1. Go to Data Security Investigations in the Microsoft Purview portal and sign in with the credentials for a user account assigned Data Security Investigations permissions. Members of the *Compliance Administrator*, *Organization Management*, *Data Security Management*, and *Insider Risk Management* role groups can also create investigations.

2. Select **Investigations** in the left navigation.

3. Select **Create investigation**.

4. On the **Create an investigation** dialog, complete the following fields:

    - **Title**: Enter a name for the investigation (required). The investigation name must be unique in your organization.
    - **Description**: Add an optional description to help others understand this investigation.
    - **Additional context for AI**: Enter additional context to help AI focus your investigation on specific areas and issues.

5. Select **Switch to full draft mode**.

6. Select **Add data sources**

7. On the **Manage data sources** flyout pane, search and add data sources for your investigation query. You can filter to scope data sources to help you choose one or more users or group sources to add to the investigation.

Use one or more of the following options in the **Search for people, groups, locations, or tenant locations** field to select from the following default filters:

- *All people and groups (default)*
- *All apps*
- *All public folders*

If needed, enter specific people, groups, locations, or tenant locations to reduce the scope of sources to specific resources in your organization.

8. Select **Save**.
9. Use the query builder to create a new search query for the investigation.
10. Select **Review scope** to review statistics about the results from the query
11. Select **Add to scope** to add data items to the investigation scope.

Next steps

After you create your investigation and add data items to the investigation scope, the system automatically vectorizes the items for AI analysis.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/11/2026

Investigation settings in Data Security Investigations

Investigation details let you view and update investigation information. You can take action on specific investigations.

Investigation details

The investigation settings display the following information for the selected investigation:

- **ID**: Shows the investigation ID number. You can review this ID but can't change it after the investigation is created.
- **Investigation name**: Shows the investigation name. This field is required. To change the investigation name, enter the new investigation name and select **Actions** > **Save**.
- **Investigation description**: Shows an optional description to help others understand this investigation. To change the investigation description, enter the new investigation description and select **Actions** > **Save**.
- **Investigation status**: Shows the current investigation status.
- **Investigation created**: Shows the date and time when the investigation was created.

Delete an investigation

Delete investigations when they're no longer needed for your organization. When you delete an investigation, you also delete all components associated with the investigation, such as searches, investigation scopes, and mitigation plans. You can't recover a deleted investigation.

To delete an investigation, complete the following steps:

1. Go to Data Security Investigations in the Microsoft Purview portal and sign in with the credentials for a user account assigned Data Security Investigations permissions.
2. Select **Investigations** in the left navigation.
3. Select an investigation, then select **Investigation settings**.
4. On the **Investigation settings** page, select **Actions**, then select **Delete investigations**.

ⓘ **Note:** The author created this article with assistance from AI. Learn more

Last updated on 04/29/2026

Configure investigation access and permission settings in Data Security Investigations

Access and permission settings in Data Security Investigations allow you to add or remove users to an investigation, manage role group membership for an investigation, and add people outside your organization as guest users for an investigation.

Add or remove users from an investigation

You can add or remove users to manage who can access the investigation. However, before a user can access an investigation (and perform tasks in the investigation), you must add the user to the *Data Security Investigations Manager* role group in the Microsoft Purview portal. For more information, see Assign Data Security Investigations permissions.

Add users to an investigation

Complete the following steps to add users to an investigation:

1. Go to Data Security Investigations in the Microsoft Purview portal and sign in using the credentials for a user account assigned Data Security Investigations permissions.
2. Select **Investigations** in the left nav.
3. Select an investigation, then select **Investigation settings**.
4. On the **Investigation settings** page, select **Permissions**.
5. Under **Users**, select **Add** to add users to the investigation. You can also choose to add a role group to the investigation by selecting **Add** under **Role groups**.
6. In the list of users or role groups that can be added to the investigation, select the check box next to the names of the users or role groups that you want to add.

> **Note**

When adding a role group to an investigation, you can only add the role groups that you're a member of.

7. After selecting the people or role groups to add as members of the investigation, select **Add**. The selected users are added to the investigation.

> **Important**
>
> If a role is added or removed from a role group that you added as a member of an investigation, then the role group is automatically removed as a member of the investigation (or any investigation the role group is a member of). The reason for this process is to protect your organization from inadvertently providing additional permissions to members of an investigation. If a role group is deleted, it's removed from all investigations it's a member of. For more information, see Assign Data Security Investigations permissions.

Remove users from an investigation

Only a Data Security Investigations Administrator can remove users from an investigation. Even if you're assigned to the *Data Security Investigations Manager* role group or initially created the investigation, you aren't able to remove yourself or other members from an investigation unless you're also a Data Security Investigations Administrator. To remove yourself or other members from an investigation, contact a Data Security Investigations Administrator in your organization.

Complete the following steps to remove users from an investigation:

1. Go to Data Security Investigations in the Microsoft Purview portal and sign in using the credentials for a user account assigned Data Security Investigations permissions.
2. Select **Investigations** in the left nav.
3. Select an investigation, then select **Investigation settings**.
4. On the **Investigation settings** page, select **Permissions**.
5. Under **Users**, select **Remove** to remove users from the investigation. You can also remove users from a role group for the investigation by selecting **Remove** under **Role groups**.
6. In the list of users or role groups that can be removed from the investigation, select the check box next to the names of the users or role groups that you want to remove.
7. After selecting the people or role groups to add as members of the investigation, select **Remove**. The selected users are removed from the investigation.

Last updated on 04/01/2026

Search, review, and refine results in Data Security Investigations

Use search in Data Security Investigations to find Microsoft 365 content such as email, documents, and instant messaging conversations in your organization that are relevant to a security incident. Use search to find content in these cloud-based Microsoft 365 data sources:

Expand table

Data source	Content type searched	Risky content examples
Endpoint Data Loss Prevention (DLP)	Files captured by endpoint DLP evidence collection	Files copied to USB, printed, uploaded to cloud, or accessed by unauthorized apps on onboarded devices.
Exchange Online	Emails and attachments in user mailboxes	Credentials or secrets sent in email, sensitive files shared externally.
Microsoft Copilot	AI prompts and responses	Sensitive data in Copilot or AI prompts.
Microsoft Teams	Chat messages (1:1 and group chats) and channel posts	Secrets included in chats, confidential information in Teams conversations.
OneDrive	User files	User files with secure access keys, export of databases, and more.
SharePoint	Documents and files in sites	Documents containing passwords, customer data, or confidential plans.

Create and run different searches that are associated with an investigation. Use conditions such as keywords, file types, incidents, and more to build search queries that return search results with the data most relevant to the investigation. You can also:

- View search statistics that help you refine a search query to narrow results.
- Preview the search results to quickly verify whether you're finding the relevant data.
- Revise a query and rerun the search.
- Use audit search to identify content based on user activities recorded in the unified audit log.
- Use endpoint DLP evidence collection (preview) to query data captured by endpoint DLP policies on onboarded devices.

Consider the activities that triggered the investigation. For example, reported password sharing or leak, a suspicious file, work with an external development team, or an alert about data exfiltration or specific individuals or teams who may be involved. Use this information to help develop your initial search query. For example, if an administrator reported that passwords might be stored in plain text in SharePoint, start by searching for known password keywords or file names across SharePoint and OneDrive, or specific SharePoint sites managed by engineering teams.

When you're satisfied with the results of a search and ready to review and analyze the results, add them to an investigation scope in the investigation. Adding copies of the original data to an investigation scope also facilitates the AI analysis and review process by providing you with advanced categorization, examination, and vector search tools.

Access search tools

Select **Summary** from the navigation options at the top of any page within a specific investigation to access search tools on the Identify page.

The Identify page is organized into the following tabs:

- **Search**: Create search queries to identify content relevant to the investigation. The **Search** tab includes three options:
  - **Query builder**: Search for content based on keywords, file types, dates, and other conditions.
  - **Audit**: Use audit search to identify content based on user activities recorded in the unified audit log.
  - **Endpoint DLP**: Use endpoint DLP evidence collection (preview) to query data captured by endpoint DLP policies on onboarded devices.
- **Review estimates**: Review scope and sample statistics for search results to refine your query before adding results to an investigation scope.

You can refine search query data sources and conditions at any time during the investigation and add the results to an investigation scope.

Data sources

In Microsoft 365, data is stored across three platforms: Exchange, Teams, and SharePoint. These platforms serve as the backbone for organizing and managing data within Microsoft 365 applications.

> **Important**
>
> If you set a site status to *Locked* and set the site to *NoAccess*, Data Security Investigations doesn't return search results. For more information about managing site lock statuses, see Lock and unlock sites. In addition, if you restrict a document in the content store, the Data Security Investigations investigator must be a site admin, site owner, or the owner or last modifier of the restricted document to access it in Data Security Investigations. For more information about restricted content, see Data Loss Prevention policy reference.

Most Microsoft 365 apps store data in one or more of the following containers:

- **Users**: Data associated with individual users, such as their mail, 1:1 Teams messages, and OneDrive files.
- **Groups**: Data owned by the organization or a group of users within an organization. These groups are often referred to as Unified Groups or Teams.

In Data Security Investigations, the concept of *data sources* streamlines the process of identifying and managing data across Microsoft 365 platforms. Analysts select a user or group and scope searches to those data sources. Analysts can refine the scope by selecting or excluding specific locations.

Analysts can also use organization-wide sources to perform search across your organization. Organization-wide sources include:

- **All people and groups**: Includes all users and all groups in your organization.
- **All public folders**: Includes all content in Exchange public folders mailboxes.

> **Important**
>
> If your organization has configured compliance boundaries using search permissions filters, Data Security Investigations searches respect those boundaries. Investigators whose accounts are scoped by a compliance boundary only see search results for content locations within their permitted boundary. For more information about how compliance boundaries and search permissions filters work, see Set up compliance boundaries in eDiscovery and Configure search permissions filtering in eDiscovery.

Query builder

The **Query builder** option on the **Search** tab provides a visual filtering experience when you build search queries in Data Security Investigations. Since you'll use AI analysis tools to quickly identify applicable data in items returned by your search, one strategy is to broaden your search scope to include more sources of data. This approach helps reduce the chance of excluding any relevant content for review.

Use the query builder to construct complex queries with additional functionality, including AND, OR, and grouping of conditions. These features in the query builder help you build queries more effectively, provide a visual interface for grouping subqueries, and provide additional space for complex keyword queries to be constructed and reviewed.

Use the query builder

To create a query and custom filtering for your search, use the following controls:

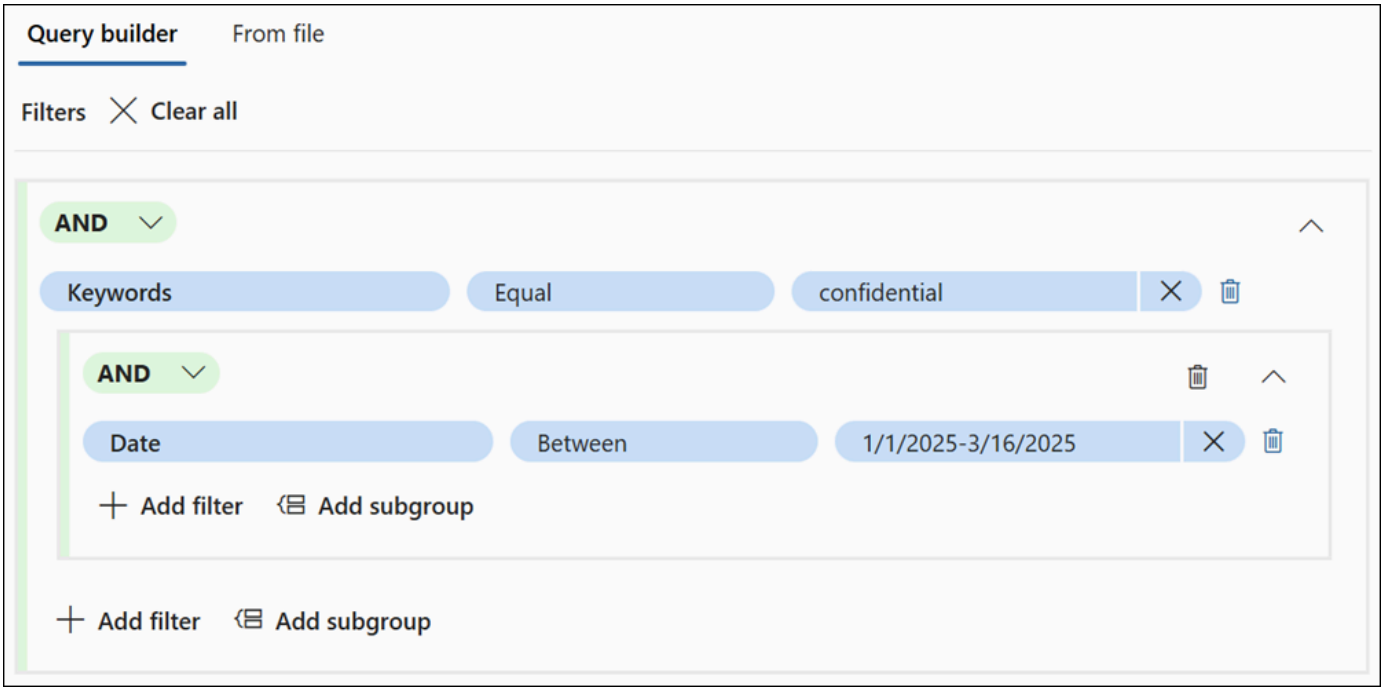
- **AND/OR**: These conditional logical operators help you choose the query condition that applies to specific filters and filter subgroups. Use these operators to include multiple filters or subgroups connected to a single filter in your query.
- **Select a filter**: Choose filters for the specific data sources and location content you select for the collection.
- **Add filter**: Add multiple filters to your query. This option becomes available after you define at least one query filter.
- **Select an operator**: Depending on the selected filter, choose from compatible operators. For example, if you select the *Date* filter, you can choose from *Before*, *After*, and *Between*. If you select the *Size (in bytes)* filter, you can choose from *Greater than*, *Greater or equal*, *Less than*, *Less or equal*, *Between*, and *Equal*.
- **Value**: Depending on the selected filter, enter compatible values. Some filters support multiple values, while others support a single specific value. For example, if you select the *Date* filter, enter date values. If you select the *Size (in bytes)* filter, enter a value for bytes.
- **Add subgroup**: After you define a filter, add a subgroup to refine the results returned by the filter. You can also add a subgroup to a subgroup for multilayered query refinement.
- **Remove a filter condition**: To remove an individual filter or subgroup, select the remove icon to the right of each filter line or subgroup.
- **Clear all**: To clear the entire query of all filters and subgroups, select **Clear all**.

Scenario example

A Data Security Investigations analyst needs to create a query for any item that includes the keyword *confidential* used between January 1, 2025, and March 16, 2025*. For this example, the

analyst creates the following query using the query builder:

1. For the first filter, the analyst selects *Keyword*, then selects the *Equal* operator, then enters *confidential* in the **Value** control.
2. Next, the analyst selects **Add subgroup** and the **AND** operator, then the **Add filter**.
3. The analyst selects the *Date* filter, the *Between* operator, and start and ending dates for the *Value*.
4. The analyst selects **Save** to save the query, then **Review scope** to run the search query.



Example search queries

You can use Keyword Query Language (KeyQL) syntax in the query builder to create targeted search queries for common investigation scenarios. Use these examples as starting points and adjust the date ranges, paths, and keywords to match your investigation.

> **Tip**
>
> Since you can use AI analysis tools to quickly identify applicable data in items returned by your search, consider broadening your search scope to include more sources of data. This helps reduce the chance of excluding any relevant content for review.

Exchange Online

Search for emails only

To search mailboxes for email messages only and exclude other items such as calendar invites, tasks, and to-do items, use the `Kind` property. This example returns emails sent after a specified date.

Data sources: All people and groups

```
(Kind=email) AND (Date>2025-01-01)
```

Microsoft Teams

Search all Teams content

To search for all Teams content, including chat messages, channel posts, meeting entries, and calls, use the `Kind=microsoftTeams` filter.

Data sources: All people and groups

```
(Kind=microsoftTeams) AND (Date>2025-01-01)
```

Search chat messages only

To narrow your search to Teams chat messages only and exclude meetings and channel posts, use `Kind=IM`.

Data sources: All people and groups

```
(Kind=IM)
```

SharePoint Online

Search a specific SharePoint site

To search all content in a specific SharePoint site, use the `Path` property.

```
(Path="https://contoso.sharepoint.com/sites/ProjectAlpha")
```

Search a site and exclude system-generated content

To search a SharePoint site for documents created after a specified date while excluding system-generated content, use the `Author` property with the NOT (`-`) operator.

text

```
((Date>2025-01-01) AND
(Path="https://contoso.sharepoint.com/sites/EngineeringTeam/Shared Documents") AND ((-
Author="System Account") AND (-Author="SHAREPOINT\System")))
```

Search multiple sites for specific file names

To search across multiple SharePoint sites for files with names that begin with a specific term, use the `OR` operator to combine paths and the `filename` property to filter by name.

text

```
(((Path="https://contoso.sharepoint.com/sites/ProjectAlpha") OR
(Path="https://contoso.sharepoint.com/sites/ProjectBeta")) AND (filename="proposal*"))
```

Search for specific files

To search for known files by their full path, use the `Identifier` property and the `OR` operator to combine multiple file paths.

text

```
((Identifier="https://contoso.sharepoint.com/sites/EngineeringTeam/Docs/SecurityReport
.docx") OR
(Identifier="https://contoso.sharepoint.com/sites/EngineeringTeam/Docs/IncidentLog.xls
x"))
```

Microsoft Copilot

Search Copilot prompts and responses

Copilot interactions are stored in users' Exchange Online mailboxes. To search for Copilot prompts and responses after a specified date, use the `ItemClass` property with the Copilot item class value.

Data sources: All people and groups

text

```
(ItemClass=IPM.SkypeTeams.Message.Copilot.*) AND (Date>2025-01-01)
```

For more information about Copilot item classes and searching for Copilot data, see Search for and delete Copilot and other AI data in eDiscovery.

Audit search

Use audit search in Data Security Investigations to identify and collect content based on user activities recorded in the Microsoft Purview unified audit log. While the query builder searches Microsoft 365 content directly, audit search queries the unified audit log for specific user activities, such as accessing, copying, or downloading files, and pulls the associated content into your investigation. You can also filter audit search results by keywords, specific users, and file or site locations.

Example scenario: Departing employee data exfiltration

A security team receives an alert that a departing employee might have exfiltrated sensitive files before their last day. Using audit search, the investigator queries for **Downloaded file** and **Sent email** activities scoped to that user over their final two weeks. Data Security Investigations pulls the associated content directly into the investigation scope - no manual file hunting required - so the team can quickly assess what data was accessed and whether anything was shared externally.

Get started with Audit search

When you create an audit search, Data Security Investigations queries the unified audit log for the specified activities and adds the associated content directly to your investigation scope. Audit search uses the permissions and role-based access control requirements from the unified audit log.

Audit search requires that your organization is configured for audit log and audit search. For more information about audit log requirements, see Search the audit log.

Supported activities

Audit search supports a curated set of content-associated activities from the unified audit log. You can only audit search activities that are associated with identifiable content items, such as files and emails. Audit search doesn't support activities that aren't associated with specific content, such as signing in to a solution.

The following activities are supported for audit search:

Expand table

Activity	Description
Accessed file	A user accessed a file in SharePoint or OneDrive.
Copied file	A user copied a file in SharePoint or OneDrive.
Deleted file	A user deleted a file from SharePoint or OneDrive.
Downloaded file	A user downloaded a file from SharePoint or OneDrive.
Moved file	A user moved a file between folders in SharePoint or OneDrive.
Renamed file	A user renamed a file in SharePoint or OneDrive.
Changed sensitivity label applied to file	A user changed the sensitivity label on a file.
Removed sensitivity label from file	A user removed a sensitivity label from a file.
Uploaded file	A user uploaded a file to SharePoint or OneDrive.
Sent email	A user sent an email message from Exchange Online.

Create an audit search

To create an audit search and add the associated content to your investigation scope, complete the following steps:

1. Go to Data Security Investigations in the Microsoft Purview portal and open an investigation.
2. Select **Summary** from the navigation options at the top of the page.
3. On the **Search** tab, select **Audit**.
4. Configure the search parameters:
    - **Date and time range (UTC)** (required): Specify a **Start** and **End** date and hour for the activity search.
    - **Activities - friendly names**: Select one or more supported activities by their display name from the activities list.
    - **Activities - operation names**: Select one or more supported activities by their operation name.

- **Users**: Select users to search for activities performed by specific users.
- **ObjectID (file, folder, or site)**: Enter a file, folder, or site path to narrow the search to activities associated with a specific location.
- **Keyword search**: Enter keywords to search for in audit log entries.

5. Select **Search** to run the audit query. Data Security Investigations queries the unified audit log and displays the results in the results section at the bottom of the Audit page. Review the results to verify the query returns the expected data. To discard the search without adding results, select **Cancel**.

6. Select **Add to scope** to add the content associated with the search results to your investigation scope.

After you add to scope, Data Security Investigations starts processing the results. You can monitor the progress of audit search jobs on the Activities page for the investigation. Previously submitted audit searches are also listed in the audit search history at the bottom of the Audit page.

> ⓘ Note
>
> The **Review estimates** and **Save** buttons on the command bar are visible but unavailable when the **Audit** option is selected. These options are only available for the **query builder**. To preview audit search results before adding them to an investigation scope use the **Audit solution** to test your queries separately.

Audit search jobs and query chunking

Audit search uses *query chunking* to process results. This method creates multiple jobs to ingest data into the investigation scope. When you submit an audit search, the following jobs appear on the Activities page:

- **One audit search job**: The primary job that queries the unified audit log for the specified activities.
- **Multiple add-to-scope jobs**: Additional jobs that process and add the associated content to your investigation scope. The number of add-to-scope jobs depends on the volume of data returned by the search.

After all jobs complete, the system automatically vectorizes the data for AI analysis. You can continue investigating the data after Add to scope completes.

> **Tip**
>
> After submitting an audit search, you see an *Audit Search* activity and multiple *Add to scope* activities on the **Activities** page. This result indicates that query chunking is processing the results.

Audit search and the unified audit log

Audit searches you create in Data Security Investigations also appear in the search history of the Microsoft Purview unified audit log. This behavior is expected because Data Security Investigations queries the unified audit log directly as the data source for audit searches.

> **Important**
>
> Data Security Investigations supports up to 10 concurrent audit searches per user. Audit searches that return more than approximately 3,000 items might not return complete results. The exact limit depends on the file names and paths of the returned items. For more information about limits, see Limits in Data Security Investigations.

Endpoint DLP evidence collection (preview)

You can use endpoint DLP evidence collection (preview) in Data Security Investigations to query data captured by endpoint data loss prevention policies on onboarded devices and add the associated content to your investigation scope. Instead of triaging individual endpoint DLP alerts one by one, investigators can aggregate endpoint events and analyze them collectively using Data Security Investigations AI tools.

Endpoint DLP evidence collection queries the data stored by the evidence collection for file activities on devices feature. This data includes files that matched endpoint DLP policies when users attempted activities such as copying to USB, printing, uploading to cloud services, or accessing files with unauthorized apps.

Prerequisites

Before you can use endpoint DLP evidence collection (preview) in Data Security Investigations, ensure the following requirements are met:

- Devices are onboarded to Microsoft Purview.

- An endpoint DLP policy is configured and active in your organization.
- The endpoint DLP policy has evidence collection for file activities on devices enabled with the **Collect original file as evidence for all selected file activities on endpoint** option selected.
- Microsoft-managed storage is configured for evidence collection.

> **Note**
>
> In addition to the standard Data Security Investigations permissions, you must also be a member of the Insider Risk Management role group and View-Only DLP Compliance Management role to query from endpoint evidence collection.

Key considerations

Keep the following considerations in mind when using endpoint DLP evidence collection (preview) in Data Security Investigations:

- **Data availability depends on evidence collection retention**: If endpoint DLP evidence collection data is no longer available in the evidence store due to retention expiration, Data Security Investigations can't query it. The same query run at different times might yield different results based on the current state of the evidence store.
- **Data isn't immutable before ingestion**: Endpoint DLP evidence collection data can change in the evidence store before you add it to your investigation scope. After you add items to scope, the data is static within the investigation.
- **File content only**: Data Security Investigations ingests the file content associated with endpoint DLP events. Some contextual metadata from endpoint DLP, such as per-event device details, isn't preserved in the investigation scope. For detailed per-event context, use the DLP Alerts dashboard or Activity explorer.
- **Data availability latency**: Evidence collection data might not appear immediately after a policy match. Availability depends on device connectivity and upload timing.
- **Endpoint DLP only**: This feature queries data from endpoint DLP policies only. Browser DLP, network DLP, and non-endpoint DLP scenarios aren't supported for this search option.

Create an endpoint DLP search (preview)

To query endpoint DLP evidence collection data and add the associated content to your investigation scope, complete the following steps:

1. Go to Data Security Investigations in the Microsoft Purview portal and open an investigation.
2. Select **Summary** from the navigation options at the top of the page.
3. On the **Search** tab, select **Endpoint DLP**.
4. Configure the search parameters:
    - **Date range** (required): Specify a start and end date for the endpoint DLP events to query.
    - **Users** (optional): Select one or more users to narrow the search to endpoint DLP events for specific users. If you don't select any users, the search returns results for all users in your organization.
5. Select **Search** to run the query. Data Security Investigations queries the endpoint DLP evidence collection data and displays the results.
6. Review the pre-ingestion estimate, which includes a summary of matched alerts, users, actions, and policies.
7. Select **Add to scope** to add the content associated with the search results to your investigation scope.

After you add items to scope, Data Security Investigations starts processing the results. You can monitor the progress on the Activities page for the investigation.

> **Important**
>
> Endpoint DLP evidence collection searches that return more than approximately 3,000 items might not return complete results. To ensure complete coverage, narrow your search by using a shorter date range or by selecting specific users. For more information about limits, see Limits in Data Security Investigations.

Work with large audit search results

When an investigation involves a large volume of user activity, audit searches might exceed the approximately 3,000-item limit. To work with large result sets, use the Audit solution to analyze the full scope of results. Then, split the data into smaller time-based slices that fit within the Data Security Investigations audit search limit.

Analyze audit results outside Data Security Investigations

The Audit solution doesn't have the same per-search item limit as Data Security Investigations audit search. Use the Audit solution to run the full query first and determine the total volume of results before splitting it into smaller searches.

To analyze audit results, complete the following steps:

1. Go to Audit in the Microsoft Purview portal.
2. Configure the same search parameters you plan to use in Data Security Investigations:
   - **Date and time range**: Specify the full date and time range for the investigation.
   - **Activities**: Select the same activities you need for the investigation.
   - **Users**: Specify the same users you plan to search.
3. Run the search and export the results to a CSV file.

Identify time slices for audit searches

After you export the audit results, use the CSV data to identify time periods that contain fewer than approximately 3,000 items each. These time periods become the individual audit searches you run in Data Security Investigations.

To identify time slices, complete the following steps:

1. Open the exported CSV file in Microsoft Excel.
2. Add helper columns to break the `CreationDate` timestamp into components for grouping:

Expand table

Column name	Formula	Description
Date	`=LEFT(B2,10)`	Extracts the date portion (YYYY-MM-DD).
Hour	`=MID(B2,12,2)`	Extracts the hour (HH).
Day-Hour	`=I2&" "&J2&":00"`	Combines date and hour for hourly grouping.
Day-Hour-Minute	`=I2&" "&J2&":"&MID(B2,15,2)`	Combines date, hour, and minute for finer grouping.

> **Note**

> Replace `B2` with the column reference for `CreationDate` in your export, and `I2`/`J2` with the column references for your Date and Hour helper columns. Adjust column references based on your export format.

3. Create a PivotTable to count records by user and time period:
    - Set **Rows** to the user column and the *Day-Hour* helper column.
    - Set **Values** to count of records.
    - Identify time groups that contain fewer than approximately 3,000 items.
4. For time groups that exceed approximately 3,000 items, apply finer grouping:
    - Break hourly groups into minute-level slices using the *Day-Hour-Minute* column.
    - Optionally, split by activity type if a single time period still exceeds the limit.

Run sliced audit searches in Data Security Investigations

After you identify the time slices, run separate audit searches in Data Security Investigations for each slice and add all results to the same investigation scope.

To run sliced audit searches, complete the following steps:

1. Go to Data Security Investigations in the Microsoft Purview portal and open an investigation.
2. Select **Summary** from the navigation options at the top of the page.
3. On the **Search** tab, select **Audit**.
4. Configure the search parameters for the first time slice:
    - **Date and time range (UTC)**: Enter the start and end times for the slice.
    - **Activities**: Select the same activities as the original query.
    - **Users**: Select the same users as the original query.
5. Select **Search** to run the audit query.
6. Select **Add to scope** to add the results to your investigation scope.
7. Repeat for each remaining time slice until all data is ingested into the investigation.

> **Tip**

Data Security Investigations automatically deduplicates items across audit searches. If time slices overlap, duplicate items aren't added to the investigation scope more than once. You can also run up to 10 concurrent audit searches per user to speed up the ingestion process.

Scope dashboard

The **Review estimates** tab displays statistics and metrics for the data results included in the search query. This view helps you determine if the search query results are ready for adding to the investigation scope or if you need to refine your query for broader or narrower results.

The search results for the **Scope** dashboard are included in the following sections:

- **Summary**: Shows the number of search hits, locations, data sources, and the total file size of partially indexed items.
  - **Total matches**: Displays the total search hit count and volume from all items matching the query criteria from locations searched.
  - **Locations**: Displays the fraction of locations with hits out of all locations searched. The numerator shows the locations with hits and denominator shows the number of locations searched. Locations with errors appear in red. To view full details on all the locations and associated hits and errors, select **Download report** to download the full .csv report.
  - **Data sources**: Displays the fraction of data sources with hits out of all data sources searched. The numerator shows the data sources with hits and denominator shows the number of data sources included in the search. This data source is consistent with the data source in the search design flow and should match the number of people or groups included in the search. A tenant-wide data source of *All people and all groups* counts as a single data source.
  - **Partially indexed items** or "Advanced indexed items hits": Displays the count and volume of partially and unindexed items returned as part of the search. The advanced indexed hit count comes from a statistic sample on the partially indexed items, actual hits might be more and you can confirm by using the *Add to a review set* and *Export search results* actions.
  - **Top data sources**: Displays the top five data sources that make up the most search hits matching your query. The names of these data sources (names of users, groups, or organization-wide locations) are listed with the hit count. These data sources should match what you selected in the data sources workflow when building the search query.
  - **Indexing status**: Breakdown of unindexed (including partially indexed) and fully indexed data items.
  - **Top location type**: Hit count by location type (mailbox versus site).

Select **Regenerate view** to rerun the query and to review the most current results. Select **Download report** to combine all **Scope** results into a single .csv file. When viewing the top 100 results for any trend area, select **Download report** for a .csv file of the top 100 results of the selected hit trend.

Samples dashboard

Samples let you inspect a representative subset of individual items and details for each item returned for the search. The number of samples per location and the number of sample locations defined in the search determine the number of sample items and location representation in the sample items.

The search results for the **Samples** dashboard columns contain the following information for each item:

- **Subject/Title**: The subject or title of the items included in the sample.
- **Date**: The date the item was created or sent.
- **Sender/Author**: The sender or author of the item.

Select a sample item to view the Source information for the item. If available for the item, this view displays a rich view of a selected item so that you can evaluate the relevancy of the item as it relates to the defined search data source and conditions.

Select **Download reports** to combine all **Sample** results into a single .csv file. Select **View settings** to view the settings applied to the sample view generation.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/22/2026

Data risk graph in Data Security Investigations

> **Important**
>
> We are retiring Data risk graph (Explore insights) in Microsoft Purview Data Security Investigations (DSI). **Beginning September 30, 2026**, organizations will no longer be able to configure Data risk graph. The feature will be fully retired and no longer available after **November 30, 2026**.

Data risk graphs in Data Security Investigations provide a visual investigation experience that combines asset and activity data into a single view. Powered by Microsoft Sentinel integration, it summarizes the previous 30 days of activity for any file scoped in your investigation. It helps you identify risky user accounts that interacted with content of interest and analyze the sequence of events leading up to recent incidents. By enriching analysis with activity insights, data risk graphs can help you resolve your data security incidents faster and with greater confidence.

For more information about Microsoft Sentinel integration, see Learn about Microsoft Sentinel in Microsoft Purview.

> **Note**
>
> Admin units aren't supported in the data risk graph. If you're scoped to an admin unit, data doesn't appear in data risk graphs.

Supported activities in data risk graph

The data risk graph currently supports the following exfiltration activities:

- Anonymous links created in SharePoint or OneDrive
- Anonymous links used via SharePoint or OneDrive
- Company links created in SharePoint or OneDrive
- File downloads from SharePoint and OneDrive
- Files renamed in SharePoint and OneDrive

Before you begin

Before you can use data risk graphs in Data Security Investigations, complete the following steps:

- Learn about pay-as-you-go billing in Microsoft Sentinel for your organization.
- Configure the prerequisites for Microsoft Sentinel data lake and Microsoft Sentinel graph.
- Review the changes made when onboarding to Microsoft Sentinel data lake and Microsoft Sentinel graph.
- Create an investigation in Data Security Investigations.

Configure data risk graph

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview **roles and permissions**.

After you complete the prerequisites, complete the following steps to configure data risk graph in Data Security Investigations:

1. Go to Data Security Investigations in the Microsoft Purview portal with an account assigned the *Global Administrator* or *Security Administrator* role.

2. Select **Overview** in the left navigation.

3. In the **Setup tasks** section, select **Set up data lake and data risk graph**.

   If you don't have the correct permissions to set up the data lake, a notification appears to contact a global or billing administrator in your organization.

4. On the **Setup** tab on **Set up data lake & risk graph (preview)**, configure the following settings:
   a. **Subscription**: Enter the Azure subscription you want to use. Your selected subscription must be valid and accessible. You must be a subscription owner.
   b. **Resource group**: Enter the resource group you want to use. Your selected resource group must be valid and accessible.

   > **Important**

After the data lake is provisioned for a specific Azure subscription and resource group, you can't migrate it to a different subscription or resource group.

5. Select **Setup**.

> **Important**
>
> When the data risk graph is first created, it includes the most recent seven days of data. As the data risk graph refreshes over time, the look-back window expands up to a maximum of 30 days. It might take some time for the data risk graph to reach the full 30-day window, so expect gradual growth after initial setup.

The set up process begins and onboarding might take up to 60 minutes to complete. You can close the setup panel while the process is running. After the onboarding process completes, you see *Complete* on **Set up data lake & risk graph (preview)**. The data risk graph shows events that occur after your Microsoft Sentinel data lake was created or onboarded.

Initial processing of data and data risk graph availability for an investigation might take 24-48 hours.

> **Important**
>
> You always have one data lake. If you already onboarded to the data lake with another Microsoft service, your existing data lake is used. If you never onboarded to the data lake, onboarding in Data Security Investigations enables Microsoft Sentinel data lake and graph in the Defender portal.

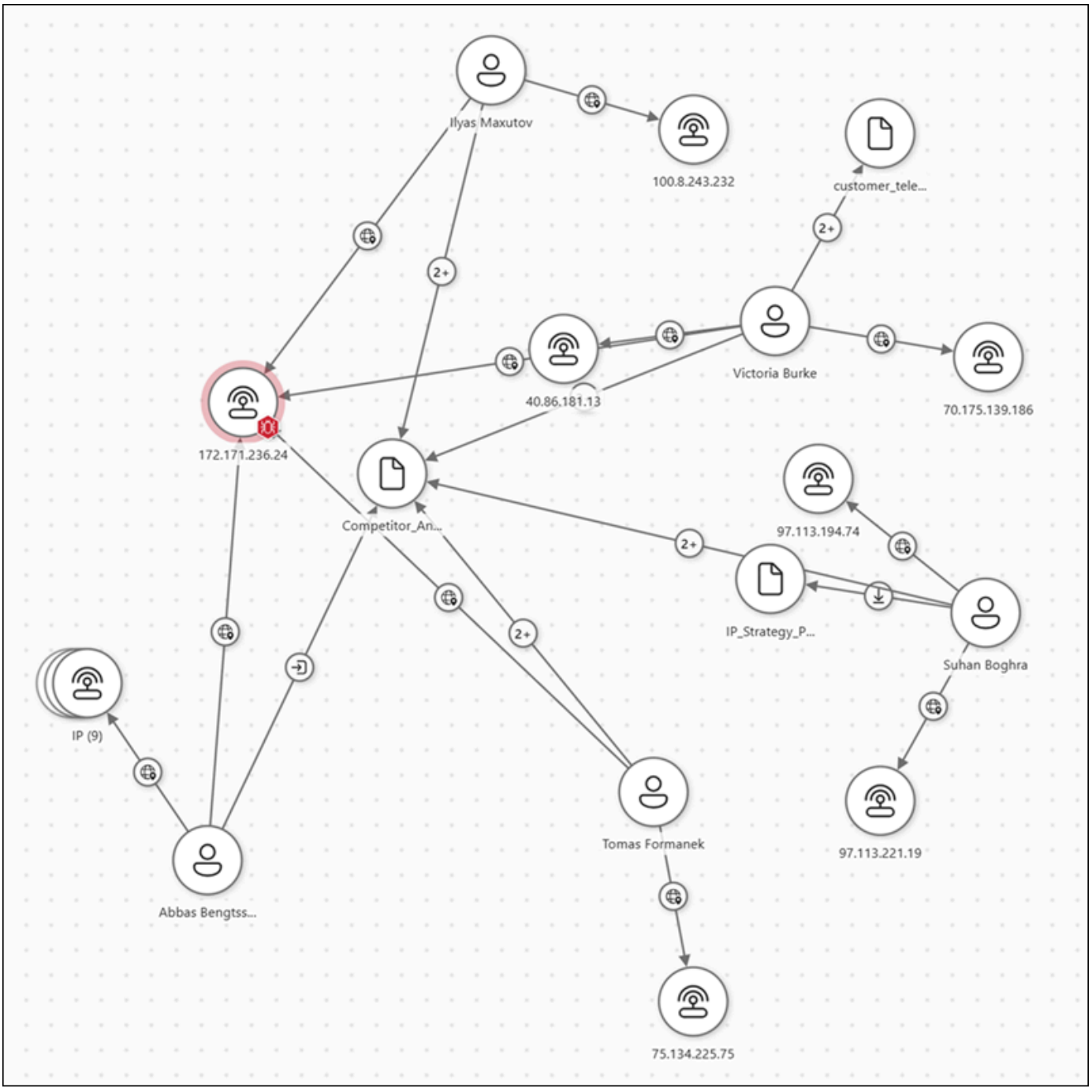
Use data risk graph

Data risk graph in Data Security Investigations uniquely visualizes correlations between impacted data, users, and their activities. It provides critical context to guide mitigation and next steps. For example, upon uncovering a highly sensitive document, data risk graphs give you visibility into which users downloaded it or if they accessed it from a risky IP address. This visibility lets you uncover new nodes to a data security incident, like additional users or new content that requires investigation.

To view the data risk graph for an investigation, you must be assigned to at least one built-in Data Security Investigations role group:

- *Data Security Investigations Admins*
- *Data Security Investigations Investigators*
- *Data Security Investigations Reviewers*

For more information about role groups, see Assign permissions in Data Security Investigations.



To use the data risk graph in Data Security Investigations, complete the following steps:

1. Go to Data Security Investigations in the Microsoft Purview portal with an account assigned to a Data Security Investigations role group.

2. Select **Investigations** in the left navigation.

3. Select an investigation to review, then select one of the following options:

a. On **Summary**, select **Manage scope** on the **Investigation scope** card.
b. Select the **Analysis** tab to list items included in the **Investigation scope**.

4. Select items (up to 100) included in the scope, and then select **Explore insights**.

> ⓘ Note
>
> **Explore insights** isn't available unless you configure Data Security Investigations for Microsoft Sentinel data lake and graph.

5. Filter the time as applicable and select individual data risk graph nodes for more information about each connection during triage.

6. Expand relationships in the graph by selecting the + icon on users or assets.

The data risk graph contains several nodes. Selecting a node shows details about the node:

- **User details**: Shows information about users associated with the investigation. This information includes organization information for each user, including User Principal Name (UPN), labels, location, job title, and more.
- **Relationships**: Shows information about starting and ending connections between nodes. This information includes starting and last connection dates, relationship types, and more.
- **IP addresses**: Shows information about each IP address node.
- **Data details**: Shows information about files and sites. This information includes the object location, type, labels, and more.

You can view data risk graph information for users based on the last 30 days of activity. This duration is fixed and can't be extended.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 09/02/2026

Manage investigation scopes in Data Security Investigations

After you add data items from your search queries to the scope of the investigation, you can start working with your data and validate items before preparing for AI analysis. You can add items to an investigation scope by using any of the following search tools:

- Query builder: Search by keywords, file types, and conditions.
- Audit search: Identify content based on user activities in the unified audit log.
- Endpoint DLP: Query data captured by endpoint data loss prevention policies.

An investigation scope contains all the data items from your search results that you want to analyze and review. Analysts in your organization can now access this data for AI preparation and further review.

Access the investigation scope

Select **Analysis** from the navigation options at the top of any page within a specific investigation to access the investigation scope. The investigation scope includes a dashboard of scoped items and includes filter and viewer tools to help with item review.

Investigation scope dashboard

The investigation scope dashboard lists all the scoped data items in an investigation. This dashboard allows you to review details about individual data items and displays information about each item.

The investigation scope dashboard contains the following information and controls:

- **Subject/Title**: The subject or title of the data item.
- **Exclude**: The exclude status of a data item. Values are *Yes* or blank.
- **Vectorized**: The vectorization status of a data item. Values are *Yes* or blank.
- **Date**: The date and time the data item was last modified.

To customize the columns display on the investigation scope dashboard, select **Customize columns** to choose the columns to display or drag and drop the columns in the list to reorder.

To download the list of data items and the column information, select **Download list** to create a .csv file containing this information.

Grouping data items

When you view a large list of data items in an investigation, it's often helpful to group the investigation scope by family or conversations and related items. Select **Group** to group data items by this criteria.

Review data items

To view detailed information about individual data items, select the item in the investigation scope dashboard or use **Filters** to narrow items based on specific attributes.

Filter items

Filters are automatically created based on the data items you add to the scope of the investigation. Typical filter options include attributes such as *Author*, *File class*, *Location name*, *Native file extension*, *Sender*, and many more.

Expand a filter area and select one or more filters to refine the data items included in the investigation scope dashboard view. Each filter attribute lists the total number of data items matching the attribute.

For example, to view only email and email attachments, expand the **File class** filter and select the **Email** and **Attachment** filters. The displayed data items are automatically filtered by the selected attributes.

View item details

Select a data item to display more information about the item. Use the investigation scope viewers to examine source, plain text, metadata, and audit details for the selected item.

Item actions

For each data item, you can choose to exclude the item from AI analysis, add the item to your mitigation plan, or purge the item from its source location. Select **Actions** on the command bar for these options. To purge items directly from the investigation scope, see Purge a single item. To purge items by using a search query, see Purge items with a search query. Review and exclude any redundant items for your investigation, such as system generated emails or meeting invites.

Reducing redundant data in your investigation helps reduce data preparation time and improves overall AI insight quality.

All data items in the investigation scope are prepared for AI analysis unless you explicitly exclude them. You can change the status of any excluded data item later if you want to have the data item processed for AI analysis.

Automatic vectorization

After you add data items to the investigation scope, the system automatically prepares all items for AI analysis. This automatic vectorization process runs in the background during each **Add to scope** action and enables you to use AI-based categorization and examination tools on these items. Automatic data preparation utilizes the latest Azure OpenAI embedding models to help improve the cost and performance of your AI results.

Automatic data preparation doesn't require many compute units, even for larger amounts of data included in an investigation scope. The system doesn't revectorize previously vectorized items. The system doesn't vectorize items that don't contain text. Some examples of these items include:

- Calendar notifications
- Meeting invites

You can see the Add to scope processing status of data items at the top of the investigation analysis page. Depending on the amount of scoped data, Add to scope can take some time to complete. If you need to cancel an in-progress **Add to scope** job, select **Cancel** on the **Analysis** tab. You still pay for any compute units consumed before the cancellation. For more information about billing, see Billing in Data Security Investigations.

Existing investigations with unvectorized data

Starting June 2026, all new investigations automatically vectorize data when items are added to the investigation scope. The **Prepare for AI** action is no longer needed or available for new investigations created after this date.

For investigations created before June 2026, the following transition applies:

- **June–August 2026**: Existing investigations that already have scoped data retain the **Prepare for AI** experience. You can continue to use **Prepare for AI** to vectorize data in these investigations during this transition period.

- **After August 2026**: The **Prepare for AI** action is removed from all investigations. All existing investigations will vectorize data automatically during the Add to scope process. If you have an existing investigation with data that was never prepared for AI, you must rerun **Add to scope** to trigger automatic vectorization for those items.

> **Important**
>
> Rerunning **Add to scope** on previously scoped data doesn't create duplicate copies or incur additional storage charges. The system recognizes data that is already stored and only initiates the vectorization process for items that haven't been prepared yet.

Next steps

After Add to scope is complete, you're ready to review and use AI analytics tools for the data in the investigation scope.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 06/22/2026

Group and view items in an investigation scope in Data Security Investigations

Data Security Investigations investigation scopes display content by using different grouping options. They include specialized viewers that you can use to examine details about individual items.

Grouping

Use the **Group** control in the command bar of an investigation scope to view review content grouped by the following options:

- **Group by families**: All items related to a specific file are grouped together by using the same Group ID. For example, if you have a PowerPoint file in the investigation scope that includes embedded images or .zip files, these images and files are grouped with the PowerPoint file and shown as nested items with the file in the item list view. Microsoft 365 Copilot and Microsoft 365 Copilot Chat user prompts and responses are also grouped together by using the same Group ID when there are multiple prompts and responses associated with a file.
- **Group by conversations**: All email messages, Teams conversations, and Viva Engage conversations are grouped by using the same Thread ID and appear as nested items. Additionally, all associated content for these messages and conversations is also grouped together. For example, if you have an email conversation that includes several email messages, some of which include attachments and some that include embedded images, all of the email messages, attachments, and images are grouped together in the investigation scope list view under an applicable item.

Source view

The **Source** viewer displays the richest view of a selected document. It supports hundreds of file types and is meant to display the truest to native experience possible. For Microsoft Office files, the viewer uses the web version of Office apps to display content such as document comments, Microsoft Teams chats, Excel formulas, hidden rows and columns, and PowerPoint notes.

Plain text view

The **Plain text** viewer provides a view of the extracted text of a selected file. It ignores any embedded images and formatting but is useful if you're trying to understand the content quickly. A line counter makes it easier to reference specific portions of a document.

For image-based items, the evidence detail pane displays text extracted through Optical Character Recognition (OCR) alongside the original image. This view allows you to verify the accuracy of the text extraction and review content identified through OCR processing.

Metadata view

The panel in the **Metadata** view displays various metadata associated with the selected document. Search for specific metadata by using the **Search metadata name** field.

AI Summary view

> **Note**
>
> This viewer is only available for items returned as results when you use Search with AI (preview).

The AI Summary viewer provides information about the relevance categorization for an item and provides an AI summary and explanation for the reasoning behind the categorization.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 05/29/2026

Use AI analysis in Data Security Investigations

After automatic vectorization completes for your investigation scope, you're ready to review and use AI analytics tools for the data in the investigation. Generative AI processing conducts a deep content analysis of selected items and can uncover key security and sensitive data risks within impacted data.

To get started with AI analysis in an investigation, complete the following steps:

1. Go to Data Security Investigations in the Microsoft Purview portal and sign in with the credentials for a user account assigned Data Security Investigations permissions.
2. Select **Investigations** in the left navigation.
3. Select an investigation, then select **Analysis** on the navigation bar.

> **Tip**
>
> Consider increasing the default item display from 50 to 1,000 items for easier selection of multiple items to exclude from the investigation scope.

Use categorization

Categorization considerations

Keep the following considerations in mind when configuring categorization and reviewing the results:

- **Not all items may appear in results**: Categorization uses relevance scoring to prioritize content. Items that score below the relevance threshold for a category aren't included in the results for that category, even if the items exist in the investigation scope.
- **Results may change when categories are updated**: When you add new categories or change existing categories between Standard and Advanced processing, the system reprocesses those categories. Because relevance is evaluated per category, results may differ based on the categories you select.
- **Large documents may be more strongly represented**: Documents with extensive content may contribute more content segments to categorization results, which can affect the

distribution of results across documents.

- **Use examination for comprehensive analysis**: If your investigation requires analysis of all items in scope rather than a prioritized subset, use examination tools to perform item-level analysis on selected data items.

Configure categorization

When you first open the **Analysis** page, items aren't categorized. Configure categorization first because it's helpful to start your triage by grouping items by risk.

Categorization takes time to complete. The completion time depends on the data volume and consumes AI capacity (compute units). To get an initial understanding of incident severity, use AI to categorize impacted data and narrow the focus to high-risk assets. Data Security Investigations sorts data into default, custom, or AI-generated categories, including by subject matter and risk.

> **Note**
>
> Investigations auto-created by proactive AI insights from Data Security Posture Management (DSPM) arrive pre-categorized with five fixed risk categories. You can run additional categorization with default, suggested, or custom categories on these investigations.

To categorize data items in the investigation scope, complete the following steps.

> **Important**
>
> Add to scope must complete before you can configure categorization.

1. Go to Data Security Investigations in the Microsoft Purview portal and sign in with the credentials for a user account assigned Data Security Investigations permissions.
2. Select **Investigations** in the left navigation.
3. Select an investigation, then select **Analysis**.
4. Select **Categorize**.
5. Select the **Standard** or **Advanced** categorization option.

- **Standard**: AI groups your scope into the selected categories. This option can significantly reduce the time it takes to complete processing and the amount of Data Security Investigations Compute Units (compute unit) needed for categorization.
- **Advanced**: AI further groups your scope into various topics within each category. This grouping takes additional time and compute units to process.

> **Important**
>
> Categorization prioritizes the most relevant content for each selected category rather than analyzing every item in the investigation scope. Results represent the highest-confidence content segments for each category. If your investigation requires comprehensive analysis of all items, use **examination tools** instead of or in addition to categorization. For more information, see **How categorization processes data**.

6. In the **Categorize with AI** dialog, complete the following areas to customize your categories as applicable:

   - **Default categories**: Select one or more default categories.
   - **Suggested categories**: Select one or more AI suggested categories. Suggested categories are generated based on the most recent vector search. If searches aren't run, no categories are suggested.
   - **Custom categories**: Select **Create category** and enter a theme or area to include. Select **Save** for the custom category.

7. After configuring your categorization settings, select **Save**.

Categorization processing can take some time depending on data volume. If you need to cancel an in-progress categorization job, select **Cancel** on the **Analysis** tab. You still pay for any compute units consumed before the cancellation. For more information, see Billing in Data Security Investigations.

After categorization processing completes, select categorization areas or individual subject areas within a category to filter data items for review. Examining the categories helps you quickly identify incident severity and scope.

When you select a specific subject area, a summary for the subject area displays with the following information:

- **Topic name**: The name of the subject area in the category.
- **Topic description**: The description of the subject area generated from AI processing.

- **Topic impact score**: The impact score related to potential risk generated from AI processing.
- **Total documents in sample**: The total number of data items that match the subject area in the investigation scope.

Add or update categories

After the initial categorization run finishes, you can add new categories or update existing categories without reprocessing the entire categorization job. When you add categories, the system processes only the new categories, and you pay only for the compute units needed to process those newly added categories. The system preserves previously categorized data unless you select an existing category in a new categorization run. Selecting existing categories in a new run reprocesses the category.

You can also change existing categories from **Standard** to **Advanced** or from **Advanced** to **Standard**. When you change the categorization option for existing categories, the system reprocesses those categories with the new option.

To add or update categories for an existing categorization, complete the following steps:

1. Go to Data Security Investigations in the Microsoft Purview portal and sign in with the credentials for a user account assigned Data Security Investigations permissions.
2. Select **Investigations** in the left navigation.
3. Select an investigation, then select **Analysis**.
4. Select **Categorize**.
5. Add new categories or update existing categories as applicable:
    - To add new default, suggested, or custom categories, select the categories you want to add to the existing categorization.
    - To change the categorization option for existing categories, select **Standard** or **Advanced** to update the processing level.
6. Select **Save**.

> **Tip**

Adding categories incrementally helps you manage compute unit costs effectively. Start with a focused set of categories and expand as your investigation progresses, rather than selecting all categories upfront.

Categorization and compute units

For example, you might see the following categories in your investigation:

- **Credentials (712 items)**: This category identifies documents or emails containing passwords or API keys.
- **Operational Information (356 items)**: This category identifies items that contain app credentials in a SharePoint site used by a team in your organization.
- **Internal Communications (122 items)**: This category identifies chat logs or emails that include shared credentials.

From an incident response perspective in this example, exposed credentials are typically the highest risk and the priority to triage, then the other areas for follow-up investigation.

Using categorization in Data Security Investigations might require a significant number of compute units, even for smaller amounts of data included in an investigation scope. The compute unit requirements are directly proportional to the number of categories selected, not the size of the data. When you add new categories to an existing categorization, you pay only for the compute units needed to process the new categories. Previously processed categories don't incur additional compute unit costs.

For more information about compute unit capacity and billing, see Billing models in Data Security Investigations.

Use vector search

Use vector search to describe what you're looking for in the vectorized data items in the investigation scope. Vector-based semantic search enables similarity-based information retrieval and understands user intent beyond literal words. You can query your impacted data to find all assets related to a particular subject, even if keywords are missing. For example, a pharmaceutical company might use vector search to find all emails, documents, Copilot prompts and responses, and Microsoft Teams messages related to vaccine trials to identify relevant assets that don't mention the words *vaccine* or *trial* but remain pertinent to the investigation.

Vector search also includes content from image-based items that were processed through Optical Character Recognition (OCR). You can find text from screenshots and scanned documents alongside text-based content.

Additionally, Data Security Investigations supports searching and returning results across multiple languages. You can create a vector search in one language, and vector search can also identify items with the same terms in other languages. For example, if you search for *shared passwords* in English, vector search might identify content in French containing *mot de passe* or content in Spanish containing *contrasena*.

Use natural language to ask a question or enter phrases with specific focus to narrow down items for review. Vector search queries don't add any extra compute unit related capacity costs because the system already processes these scoped items.

To create a vector search, complete the following steps:

> **Important**
>
> Add to scope must complete before you can use vector search.

1. In an investigation, select the **Analyze** card or the **Analysis** tab.
2. Select **Standard** mode.
3. Describe what you're looking for in the search field or select one of the suggested searches.
4. Select the search arrow or press **Enter**.

The vector or suggested search starts and lists data items associated with your query in the items area. Each item includes a search relevance score. The results are listed from high to low by default, with the most relevant items listed first. The search relevance measures how closely each result matches the terms of the search that you provided. The search relevance score indicates the confidence level of the connection and helps give you a sense of confidence about how well each result fits your search. The score applies only to the current vector search, and item scores can change based on the search performed.

Search relevance scores are as follows:

- **High**: Strong connection signals, highly relevant vector search results.
- **Medium**: Moderate connection signals, likely relevant vector search results.
- **Low**: Weak connection signals, possibly less relevant vector search results.

> **Important**
>
> Search relevance scores show only when vector search items are returned.

Vector search and compute units

Using vector search in Data Security Investigations doesn't require many compute units, even for larger amounts of data included in an investigation scope.

For more information about compute unit capacity and billing, see Billing models in Data Security Investigations.

> **Tip**
>
> Consider adding **context to the investigation** to help focus categories on specific areas or issues.

Use Search with AI (preview)

Use Search with AI (preview) to ask natural language questions or enter keywords with a specific focus to narrow down items for review. Search with AI (preview) supplements vector search and extends AI capabilities when analyzing your data. Search with AI (preview) now includes metadata for data items, which helps you narrow down relevant items based on item file types, sizes, versions, and more.

In addition to relevant items, search results also include a high-level summary of all results. This summary helps you quickly determine if the search items are relevant to your search question or keywords. The summary includes citations to specific items returned by the search and relevance scores for each item.

To use Search with AI (preview), complete the following steps:

> **Important**
>
> Add to scope must complete before you can use Search with AI (preview).

1. In an investigation, select the **Analyze** card or the **Analysis** tab.
2. Select **Ask AI (preview)** mode.

3. In the **Search with AI (preview)** pane, enter your question about the data or enter keywords.
4. After the search completes, review the search summary, item results, and item AI search details.

The item detail pane displays all items matching the context related to your AI search. Use filters to help focus the results by document and sender or author. Select **Document** to view a list of cited items included in the results to automatically filter by relevant items. Use actions on the command bar to examine, categorize, or add one or more items to your mitigation plan.

Select an item in the results and select the **AI summary** view to review the relevance categorization score and a snippet with an extracted example that matches the intent of the search.

Use examination tools

> **Tip**
>
> Consider adding **context to the investigation** to help focus examination results on specific areas or issues.

Use examination to run deep content analysis with AI on selected data items. This examination helps you find security risks buried within impacted data. By examining impacted data for security risks, you can find credentials, network risks, or evidence of threat actor discussion. After you identify security risks, you can scan for sensitive data, like personal data, financial, or health information.

In addition to summarizing risks, Data Security Investigations provides mitigation steps and the thought process to explain the assessment. From here, you can add open issues to the mitigation plan, connecting analysis to mitigation. This analysis helps you identify data relevant to your investigation and quickly take action to minimize the impact.

You can choose examination processing for the following focus areas:

- **Credentials**: Credentials processing examines and extracts credentials and access assets included in selected data items.
- **Risks**: Risks processing analyzes and scores selected data items for active risks.
- **Mitigation**: Mitigation processing identifies specific threats and recommends mitigation steps for selected data items.

- **Personal data**: Personal data processing identifies and extracts personal data from selected data items, including names, email addresses, employee IDs, IP addresses, and other PII.
- **Custom**: Define a custom examination focus area using a natural language prompt to describe what you want AI to analyze. Custom examinations produce the same structured results as built-in focus areas. For more information about creating and using custom examinations, see Create a custom examination.

Examination processing can take some time depending on data volume and the focus area you select. If you need to cancel an in-progress examination job, select **Cancel** on the **Analysis** tab. You still pay for any compute units consumed before the cancellation. For more information, see Billing in Data Security Investigations.

When the examination process completes for a focus area, select **Probing history** from the command bar on the right side of the investigation scope page. In the **Probing history** pane, select **View details** for a specific examination process.

Create a custom examination

In addition to the built-in focus areas, you can create a custom examination to analyze data items for specific scenarios unique to your investigation. Custom examinations use a natural language prompt that you define to instruct the AI on what to look for.

Complete the following steps to create a custom examination:

1. On the investigation scope page, select one or more data items.
2. Select **Examine** from the command bar.
3. Select **Custom** as the examination type.
4. In the prompt field, enter a natural language description of what you want the AI to analyze. Prompts can be up to 2,000 characters.
5. Select **Verify prompt**.
6. Verify the prompt is understood with the same intent.
7. Select **Examine** to start the examination.

After the examination completes, results appear in **Probing history** with the same structured output as built-in focus areas, including findings, severity, surrounding snippets, and thought process explanations.

To edit and re-run a custom examination prompt, select the previously created custom examination type and select **Edit**.

Example prompts

The following examples illustrate custom examination prompts for common investigation scenarios:

- **Protected Health Information (PHI)**: "Identify any Protected Health Information (PHI) including patient names, medical record numbers, diagnosis codes, treatment information, and insurance identifiers."
- **Insider trading**: "Identify references to specific stock tickers, trade volumes, non-public earnings data, or language suggesting intent to trade on insider information."
- **IP exfiltration**: "Identify references to project codenames, patent application numbers, unreleased product specifications, or proprietary algorithms."

Examination and compute units

Using examination in Data Security Investigations might require a significant number of compute units, even for smaller amounts of data included in an investigation scope. The compute unit requirements are directly proportional to the size of the data categorized and each examination option you select.

For more information about compute unit capacity and billing, see Billing models in Data Security Investigations.

Examination process information

Select **Probing history** from the far-right command bar to display a list of the examination activities for the investigation scope.

The list shows the following summary information for each examination process:

- **Name**: The name of the examination.
- **Created by**: The user principal name (UPN) of the user that created the examination process.
- **Probe**: The probing area selected for the examination.
- **Scope**: The number of items selected for examination.
- **Date**: The creation date of the examination process.
- **Status**: The process status. Values include *In progress*, *Successful*, *Failed*, or *Cancelled*.

Select **View details** when the process finishes to view the examination report and recommendations.

Next steps

After the examination process finishes, review the recommendation summaries that you selected:

- Review credentials examinations.
- Review risk examinations.
- Review mitigation recommendations.
- Review personal data examinations.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/11/2026

Review mitigation recommendations in Data Security Investigations

Analyze data for mitigation information

Complete the following steps to identify risks in data items included in the investigation scope:

> **Important**
>
> Add to scope must complete before you can configure examinations.

1. Go to Data Security Investigations in the Microsoft Purview portal and sign in with the credentials for a user account assigned Data Security Investigations permissions.
2. Select **Investigations** in the left navigation.
3. Select an investigation, then select **Analysis** on the navigation bar.
4. Use vector search or categorization tools to identify data for credentials examination.
5. Select one or more items, then select **Examine** on the command bar.
6. In the **Examine** dialog, enter a name for your examination process in the **Name** field.
7. Enter a description for the examination process in the **Description** field.
8. Select *Mitigation: Identify threats and recommend mitigation steps in selected evidence* in the **Choose a focus area** field.
9. Select **Examine** to start the AI analysis.

> **Note**
>
> The time estimates for the process to complete are based on the amount and size of the selected data. To reduce processing time, filter and exclude data not applicable to the investigation.

Mitigation recommendation report

After the AI processing completes for the selected data items, you can review the mitigation plan to identify details and recommended mitigation steps for each item.

The mitigation recommendation reports include the following information for each item:

- **Subject/Title**: The subject or title of the data item.
- **Risk score**: The risk score of the data item. Values range from 0.00 to 0.10.
- **Risk summary**: A summary of the risks associated with an item. Includes information about content types, users, senders/recipients, details about communications, and more.
- **Mitigation steps**: Guidance for recommended steps to help mitigate the risk associated with the item. Includes audit log reviews, location tracing, access reviews, monitoring suggestions, and more.
- **Thought process**: A summary of the reasoning as to why the risks associated with the item are important. Includes security, legal, financial justifications, and more.
- **Risk implications**: A summary of the implications if the identified risks aren't addressed.
- **Errors**: A summary of any processing errors encountered when the AI process runs.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/11/2026

Review credentials examinations in Data Security Investigations

Understanding how credentials and other asset access information associate with a data security incident is an important part of identifying and mitigating risks.

For example, if you discover credentials within impacted data associated with the data security incident, a Microsoft Entra admin in your organization can join the investigation, securely view the extracted credentials, and take necessary next steps to reset the accounts. You can also use investigation learnings to refine existing account management policies to strengthen your organization's security practices.

Analyze data for credentials and asset access information

Complete the following steps to identify credentials and asset access data in items included in the investigation scope:

> **Important**
>
> Add to scope must complete before you can configure examinations.

1. Go to Data Security Investigations in the Microsoft Purview portal and sign in by using the credentials for a user account assigned Data Security Investigations permissions.
2. Select **Investigations** in the left navigation.
3. Select an investigation, then select **Analysis** on the navigation bar.
4. Use vector search or categorization tools to identify data for credentials examination.
5. Select one or more items, then select **Examine** on the command bar.
6. In the **Examine** dialog, enter a name for your examination process in the **Name** field.
7. Enter a description for the examination process in the **Description** field.
8. Select *Credentials: Extract the credentials and access assets in selected evidence* in the **Choose a focus area** field.
9. Select **Examine** to start the AI analysis.

> **Note**

The time estimates for the process to complete is based on the amount and size of the selected data. To reduce processing time, filter and exclude data not applicable to the investigation.

Credentials examinations

After the AI processing completes for the selected data items, you can review credentials examinations to identify credential and asset access details for each item.

The credentials examinations include the following information for each item:

- **Subject/Title**: The subject or title of the data item.
- **Extracted credentials**: The credential details included in the data item. This information includes user names, passwords, and more.
- **Credential type**: The credential type. Might include user credentials, API tokens, MFA backup codes, and more.
- **Surrounding snippet**: The text or string values surrounding the credential details. This information helps with determining the context in which the credential is used in the data item.
- **Thought process**: A summary of the reasoning as to why the credentials associated with the item are important.
- **Errors**: A summary of any processing errors encountered when the AI process runs.

Credentials examination example

The output of the examination might be a summary or annotation for each identified item. The examination highlights sentences or pieces of data in the content that are likely sensitive.

For example, the examination might return *In File X, the AI found what looks like 10 user passwords and 5 API keys* in a document. Or for an email, *Email Y contains a conversation about sharing database credentials*. This summary saves you from manually reading every line for each item and helps you triage risky content very quickly. Here's an example of a deep examination output for an item with Credential Risk:

Expand table

File	String	Type	Extracted Value	Analysis
EV-1.docx	T9!vX3p@7qLz	LoosePassword	T9!vX3p@7qLz	The string matches a typical password pattern, featuring uppercase, lowercase, numbers, and special characters. There are no other indicators, so it's classified as a `LoosePassword`.

An item with cleartext passwords might be tagged as `Critical risk: Credentials exposed` by the system. These examinations are important and they align with what a human analyst would consider severe and they also include explanations.

You should also take the following actions:

- **Manually verify critical details**: While the examination process does the heavy lifting, you should manually verify the most critical pieces. If a specific password is identified in a document, open the document to confirm the context. You should verify that it's an actual active password, and whose password it is. For instance, the string `P@ssw0rd123` might be identified and it's labeled as a password. But by opening the file, you verify it's in a table titled *VPN Credentials* and not in the content only as an example placeholder. This confirms it's actionable and gives you the information needed to act.
- **Expand the scope**: Sometimes examinations reveal new clues that require pulling in more data. For example, you discover references to a project name or a user account you didn't search for initially. You should consider another search for the project or user account to see if there are related items. A vector search for the project name might find related files that didn't explicitly mention passwords, but are related to the security incident.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/11/2026

Review risk examinations in Data Security Investigations

By examining impacted data for security risks, you can find network risks or data that presents continued security risks in your organization. Risk examinations provide reasoning for risk scores and recommended mitigation steps for each item so you can take the appropriate mitigation steps. Data Security Investigations summarizes scores and explanations for the risk focus areas.

Analyze data for risk information

Complete the following steps to identify risks in data items included in the investigation scope:

> **Important**
>
> Add to scope must complete before you can configure examinations.

1. Go to Data Security Investigations in the Microsoft Purview portal and sign in with the credentials for a user account assigned Data Security Investigations permissions.
2. Select **Investigations** in the left navigation.
3. Select an investigation, then select **Analysis** on the navigation bar.
4. Use vector search or categorization tools to identify data for credentials examination.
5. Select one or more items, then select **Examine** on the command bar.
6. In the **Examine** dialog, enter a name for your examination process in the **Name** field.
7. Enter a description for the examination process in the **Description** field.
8. Select *Risk: Analyze and score all active risks in selected evidence* in the **Choose a focus area** field.
9. Select **Examine** to start the AI analysis.

> **Note**
>
> The time estimates for the process to complete are based on the amount and size of the selected data. To reduce processing time, filter and exclude data not applicable to the investigation.

Risk examination scores range between 0 (low risk) to 4 (critical risk) and include the following information for each item:

- **Subject/Title**: The subject or title of the data item.
- **Privileged content**: Indication if the data item is privileged content. Values are *Yes* or *No*.
- **Overall security risk score**: The risk score of the data item. Values from 0 to 4, higher values indicate higher risk items.
- **Overall security risk explanation**: A summary of the reasoning as to why the risks associated with the item are important.
- **Errors**: A summary of any processing errors encountered when the AI process runs.

The following risk examination areas, with scores and explanations, are available for each item:

- **Asset identifiers score**: Score reflecting the presence and severity of information related to identities or internal assets within emails and files.
- **Credentials and secrets exposure score**: Score reflecting the discussion of credential exposure within emails and files.
- **Customer intel risk score**: Score capturing the risk associated with confidential information related to business deals, trade secrets, corporate strategy, and cybersecurity vulnerabilities.
- **Government intel risk score**: Score capturing the risk associated with confidential government information related to security of infrastructure and products, domestic and foreign security, and domestic and foreign affairs.
- **Infrastructure risk score**: Score assessing the risk related to infrastructure exploits, critical gaps related to infrastructure, and potential infrastructure related discussions.
- **Network and access risk score**: Score reflecting the risk associated with network security and access control, including exposed ports and anonymous connections.
- **Personal sensitive content score**: Score reflecting the level of personal or sensitive information within emails and files, such as medical details, financial information, or confidential topics.
- **Policy compliance and data protection risk score**: Score capturing the risk associated with policy deviations, compliance, and data protection practices.
- **Privileged information risk score**: Score capturing the risk associated with exposure of privileged and confidential information and/or attorney communication for business matters.
- **Threat actor or breach discussion score**: Score reflecting the discussion of threat actors or sensitive topics within emails and files, including past breaches, ongoing investigations, or internally reported risks.

- **Trade secret risk score**: Score capturing the risk associated with policy deviations, compliance, and data protection practices.
- **Urgent security incident score**: Score reflecting the discussion of any urgent, immediately actionable security incidents that threat actors might find valuable.
- **Vulnerability and security hygiene information score**: Score reflecting the discussion of vulnerabilities within emails and files, such as zero-day exploits, security hygiene practices, internet-exposed endpoints, and over-provisioned accounts.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/11/2026

Review personal data examinations in Data Security Investigations

Understanding what personal data is present in data associated with a security incident is a critical part of investigating and responding to data breaches. After a data security incident, one of the most time-consuming and high-stakes tasks for investigators is determining which personal data elements were exposed and whether regulatory notification obligations exist under frameworks such as the General Data Protection Regulation (GDPR), the Health Insurance Portability and Accountability Act (HIPAA), or state-level privacy laws.

For example, if you discover personal data within impacted data associated with a data security incident, you can quickly assess potential privacy exposure, support compliance reporting, and build defensible evidence chains from detection to action. You can also use investigation learnings to inform regulatory notification decisions and strengthen your organization's data protection practices.

Personal data examinations support common investigation scenarios, including:

- **Data breach privacy impact assessment**: Run personal data examinations on impacted files to quickly determine which personal data types were exposed and which individuals might be affected.
- **Regulatory compliance evidence**: Generate structured, auditable results showing what personal data was found, in which files, at what risk level, with AI-generated reasoning to support breach notification reporting.
- **Insider risk investigation**: Identify the types and risk levels of personal data present in data associated with risky user activities, helping prioritize which files require immediate attention and remediation.
- **Cross-functional incident response**: Provide security, legal, and privacy teams with a shared, actionable summary of personal data exposure that all teams can reference during incident response.
- **Post-incident audit**: Conduct a systematic review of past incidents to ensure personal data exposure was accounted for, validating or supplementing earlier manual review efforts.

Analyze data for personal data

Complete the following steps to identify personal data in items included in the investigation scope:

> **Important**
>
> Add to scope must complete before you can configure examinations.

1. Go to Data Security Investigations in the Microsoft Purview portal and sign in by using the credentials for a user account assigned Data Security Investigations permissions.
2. Select **Investigations** in the left navigation.
3. Select an investigation, and then select **Analysis** on the navigation bar.
4. Use vector search or categorization tools to identify data for personal data examination.
5. Select one or more items, and then select **Examine** on the command bar.
6. In the **Examine** dialog, enter a name for your examination process in the **Name** field.
7. Enter a description for the examination process in the **Description** field.
8. Select *Personal data: Identify and extract the personally identifiable information in selected evidence* in the **Choose a focus area** field.
9. Select **Examine** to start the AI analysis.

> **Note**
>
> The time estimates for the process to complete are based on the amount and size of the selected data. To reduce processing time, filter and exclude data that aren't applicable to the investigation.

Personal data examinations

After the AI finishes processing the selected data items, you can review the personal data examinations to identify personal data details for each item.

The personal data examinations include the following information for each item:

- **Subject/Title**: The subject or title of the data item.
- **Risk level**: The risk levels for all personal data found in the item. Values include *High*, *Medium*, and *Low*.
- **Contains Personal Data**: Indication if the data item contains personal data. Values are *Yes* or *No*.

- **Type of Personal Data**: A list of personal data types found in the data item. For more information, see Personal data types and risk level classification.
- **Personal Data Found**: A list of extracted personal data values from the data item.
- **Surrounding Snippet**: The text or string values surrounding the personal data details. This information helps with determining the context in which the personal data appears in the data item.
- **Thought Process**: A summary of the reasoning as to why the personal data associated with the item was classified and its risk level rationale.
- **Errors**: A summary of any processing errors encountered when the AI process runs.

Personal data types and risk level classification

Each personal data finding is assigned a risk level based on the type of personal data identified. The risk level helps you prioritize which findings require immediate attention and remediation. The following table includes examples of personal data types for each risk level. These examples aren't a comprehensive list of all personal data types detected by the examination.

Expand table

Risk level	Example personal data types	Description
High	SSN/National ID, bank account numbers, financial account numbers, routing numbers	Financial identifiers and government-issued identification numbers that pose the highest risk if exposed.
Medium	Email addresses, phone numbers, dates of birth, medical record numbers, insurance IDs	Contact information that can be used to identify or reach individuals.
Low	Full names, addresses, device identifiers	General identifying information that presents lower risk when exposed individually.

> **Note**
>
> A single data item might contain multiple personal data instances of different types and risk levels. The process captures and reports personal data instances individually.

Personal data examination example

The output of the examination might be a summary or annotation for each identified item. The examination highlights personal data elements in the content that are likely sensitive.

For example, the examination might return *In File X, what looks like 3 Social Security numbers and 5 email addresses were found* in a document. Or for an email, *Email Y contains a conversation where bank account details were shared*. This summary saves you from manually reading every line for each item and helps you triage risky content very quickly. Here's an example of a personal data examination output for an item:

Expand table

File	String	Type	Risk level	Analysis
EV-2.docx	555-12-3456	SSN/National ID	High	The string matches a Social Security number pattern (XXX-XX-XXXX) found in a table titled *User Records*. It's associated with a full name in the adjacent column, confirming it as identifiable PII.
EV-3.msg	user1@contoso.com	EmailAddress	Medium	The email address appears in the body of the message as a recipient of shared financial documents, making it contextually relevant to the investigation.

Take the following actions:

- **Manually verify critical details**: While the examination process does the heavy lifting, manually verify the most critical pieces. If the process identifies a specific personal data value in a document, open the document to confirm the context. For instance, the process might identify a Social Security number in a file, but by opening the file, you verify it's in a table titled *User Records* and not only in the content as an example placeholder. This verification confirms the finding is actionable and gives you the information needed to act.
- **Expand the scope**: Sometimes examinations reveal new clues that require pulling in more data. For example, you discover references to a user account or a department you didn't search for initially. Consider another search for the user account or department to see if there are related items. A vector search for the department name might find related files that didn't explicitly mention personal data but are related to the security incident.
- **Connect findings to the mitigation plan**: After reviewing personal data examination results, add specific data items to the mitigation plan to manage and track the mitigation status for each item. This approach creates an auditable chain from detection to remediation.

- **Consider regulatory implications**: Personal data exposure might trigger notification obligations under privacy regulations. Use the structured examination results, including risk levels, personal data types, and AI-generated reasoning, to support compliance reporting and regulatory notification decisions.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/11/2026

Take mitigation actions in Data Security Investigations

Data Security Investigations provides several features to help you quickly mitigate the impact of data security incidents in your organization.

Mitigation plan

The mitigation plan helps you connect analysis to specific mitigation actions from examination tools and analyst reviews. Use the mitigation plan like a to-do list to manage and track data risk mitigation actions in your investigation. This plan helps analysts centralize all items that require attention or action to mitigate risks associated with the investigation and data security incident.

Using the mitigation plan, take the following actions:

- **Prioritize by risk level**: Immediately address anything that might lead to an active security breach if you don't fix it.
- **Mark items for follow-up**: Use the mitigation plan to track follow-up activities.
- **Consult or involve relevant stakeholders**: Make sure your plan addresses responsibilities for the following areas:
  - **Credentials**: Identify who resets credentials and by when.
  - **Data exposure**: Identify who removes external sharing links or moves files to a secure location.
  - **User education or enforcement**: Identify who follows up with training or disciplinary measures.
  - **Further investigation**: Identify who is responsible for any item that suggests a larger security issue in your organization.

Mitigation plan dashboard

The **Mitigation plan** displays all items you add to the mitigation plan in your investigation. By using this view, you can quickly see all the items that need action in your investigation, the status for each item, and other important information about each item in the plan.

- **Name**: The subject or title of the data item.
- **Type**: The file type of the data item.

- **Description**: The description of the data item.
- **Status**: The current status of the data item. Values include *In-progress* or *Complete*.
- **Last modified by**: The name of the analyst or investigator that last updated information about the data item.
- **Last updated (UTC)**: The date and time in Coordinated Universal Time (UTC) that the data item was last updated.

Select **Customize columns** to customize the displayed columns and the order of the columns in the mitigation plan. Choose the columns to display or hide. Select **Add filter** and choose an available filter to filter items in the mitigation plan.

Select **Download list** to create a .csv file containing the list of data items and the column information in the mitigation plan.

Item detailed view

Analysts can view details about each data item included in the mitigation plan as part of the review and evaluation process. After you select a data item, you see an item summary and viewers to view metadata, source, and other information.

Update item status

As you review data items and work to mitigate risks for each data item, manage the current status for each item. Update the status for individual items or bulk update the status for multiple data items.

To update the status for a data item, complete the following steps:

1. Go to Data Security Investigations in the Microsoft Purview portal and sign in with the credentials for a user account assigned Data Security Investigations permissions.
2. Select the **Data Security Investigations** solution card, then select **Investigations** in the left navigation.
3. Select an investigation, then select **Mitigation**.
4. Select one or more data items in the mitigation plan.
5. Select **Update status**, then select the applicable status.

You can also change the status of a data item from the item detail view. From the detailed view for an item, select **Complete** to change the status of the data item.

Implement the mitigation plan

To implement your mitigation plan, work through the following mitigation steps:

- **Revoke or change credentials immediately**: For each exposed password or secret, ensure the responsible owner resets it. This action might involve forcing password resets for user accounts, deleting or regenerating access keys, or changing service account credentials.
- **Contain data exposure**: If files were accessible publicly or to too many people, lock down permissions immediately. Remove external guest access on SharePoint sites if a sensitive file was shared. If an email with secrets was sent externally, reach out to the recipient to delete it (if appropriate) or use message recall if possible.
- **Remediate documents**: For files that shouldn't contain certain information, decide whether to delete them, redact them, or secure them. Often, removing the sensitive content from the source is advised (or encrypting it). Consider applying sensitivity labels or Data Loss Prevention policies to prevent files from being shared again.
- **Track activities and insights**: Use the activity timeline for items. This information might identify related user activities. Check if any risky user behavior is ongoing and if further investigation of that user might be warranted.
- **Document every action**: In the activity history, every search, analysis, and mitigation action is logged. This information is useful for audit and post-incident review. Ensure the log shows that all risky items are addressed.
- **Communicate and resolve the incident**: Update all stakeholders on the following information:
  - Confirm that critical holes (credentials, public links, and other vulnerabilities) are closed.
  - If necessary, prepare any external notifications. For example, informing customers of a data breach and informing applicable regulators.
  - Conduct a debrief with all stakeholders. The insights and lessons learned can be used to improve policies and prevent future incidents.

Purge

Data Security Investigations includes purge features to help you identify and remove Exchange mailbox and Microsoft Teams items in your organization. This need might include identifying items associated with highly confidential or sensitive information about financial, marketing, or sales projects, or other proprietary property. By using the built-in search capability for investigations, you can quickly create a new search query and save it as a purge query that you can review and run whenever needed.

> **Tip**
>
> Use soft purge as the default purge method for most investigation-driven workflows. Soft purge lets you quickly remove sensitive content from a user's view while preserving the ability to restore items if needed. Reserve hard purge for scenarios where you have high confidence that permanent deletion is required and items are correctly identified.

Data Security Investigations supports the following purge options:

- **Soft purge**: Items are moved to the recoverable items folder and retained based on your retention settings. Soft purge is only supported for Exchange mailbox items. Teams messages and OneDrive and SharePoint files aren't supported for soft purge.
- **Hard purge**: Items are permanently deleted from their data sources. This action can't be undone and search queries can't be edited after the purge is complete.

> **⚠ Warning**
>
> Hard purge permanently deletes items with no recovery path. Only use hard purge when you have high confidence that the items are correctly identified and permanent deletion is explicitly required.

Soft purge is useful for scenarios such as user terminations or active data-leak containment, where you need to quickly remove sensitive content from a user's view while preserving the ability to restore items if the investigation scope changes or the content is needed for further evidence review.

Both soft purge and hard purge respect litigation holds, eDiscovery holds, and retention policies. If items are subject to a hold or retention policy, purged items are preserved in the recoverable items folder and can't be permanently deleted until the hold is removed or the retention period expires. For more information, see The principles of retention, or what takes precedence?.

> **ⓘ Important**
>
> Purging removes items from their original source locations but doesn't remove items from the investigation scope in Data Security Investigations. To remove all items from Data Security Investigations, **delete the investigation**.

You pay for the processing needed to purge the volume of data resulting from your search query and uses the new Data Security Investigations compute unit. For more information about the compute unit, see Billing in Data Security Investigations.

Purge queue dashboard

The **Purge queue** displays all the saved purge queries in your investigation. By using this view, you can quickly see all the purge queries in your investigation, the status for each query, and other important information about each query in the queue.

- **Name**: The name of the purge query.
- **Status**: The current status of the purge query. Values include *Not started*, *In-progress*, *Failed*, or *Complete*.
- **Added by**: The name of the investigator that added the purge query.
- **Added date (UTC)**: The date and time in Coordinated Universal Time (UTC) that the purge query was added to the purge queue.

Select **Customize columns** to customize the displayed columns and the order of the columns in the purge queue. Choose the columns to display or hide. Select **Add filter** and choose an available filter to filter items in the purge queue.

Select **Download list** to create a .csv file containing the list of data items and the column information in the purge queue.

Choose a purge method

Use the following table to determine the best purge method for your scenario:

Expand table

Consideration	Soft purge	Hard purge
Action	Moves items to the recoverable items folder	Permanently deletes items from data sources
Recoverable	Yes, based on retention settings.	No, action can't be undone.
Supported data sources	Exchange mailbox items only.	Exchange mailbox items and Teams messages.
Risk level	Lower risk - items can be restored if needed.	High risk - irreversible, no recovery path.

Consideration	Soft purge	Hard purge
Recommended use	Default for most investigation-driven purges, bulk or search-based purges, and scenarios where query scope might be broad.	Only when you have high confidence that items are correctly identified and permanent deletion is required.
Example scenarios	User terminations, active data-leak containment, removing sensitive content while preserving evidence.	Confirmed malicious content, prohibited material with no compliance retention requirement.

Purge items with a search query

To purge specific items in your organization, complete the following steps:

1. Go to Data Security Investigations in the Microsoft Purview portal and sign in with the credentials for a user account assigned Data Security Investigations permissions.
2. Select **Investigations** in the left navigation.
3. Select an investigation or select **Create investigation** to create a new investigation.
4. On the **Search** tab, select the data sources that contain the items you want to purge.

   > **Important**
   >
   > Only items in Exchange mailboxes and Microsoft Teams currently support purge. Selecting SharePoint sites or OneDrive accounts disables purge actions. Soft purge is only supported for Exchange mailbox items. If the search query includes Teams messages or OneDrive and SharePoint files, the soft purge option isn't available.

5. Create a search to identify the items you want to purge, and then select **Review for purge**.

   > **Important**
   >
   > *FileID* isn't available for purged items not added to a review set or when exported. Use *ImmutableID* if needed.

6. On the **Review for purge** page, verify that the items returned in the estimates align with the items you want to purge. You can purge up to 10,000 items per search.

   If you need to edit the search, select **Search** and update the conditions of the search. To view details for search matches, select the ellipsis and **View samples** or select **Add to scope**.

> **ⓘ Important**
>
> If you add items to an investigation scope, a copy of the items is added to the Azure storage account associated with Data Security Investigations in your organization. Running the purge query in the **Purge queue** removes items from all their current source locations, but not from the investigation scope in Data Security Investigations. To remove all items from Data Security Investigations, delete the investigation.

7. Select **Save for purge** and enter a unique name for the purge query in the **Query name** field.

8. Select **Save** to save the query and add the query to the purge queue. This step only saves the query as a purge query and doesn't purge any items.

   > **⚠ Warning**
   >
   > The item count displayed during review for purge is an *estimate* based on the search results at the time of review. If you don't run the purge query immediately, the items matching the query can change significantly over time. Running a saved purge query at a later date could result in purging substantially more items than originally estimated, including items you didn't intend to delete. To avoid unintended data loss, always review for purge again immediately before running the purge query.

9. On the **Mitigation** tab for the investigation, you're automatically directed to the **Purge queue** dashboard.

10. Select the purge query you want to run.

11. On the purge query details page, review the query details and select **Run**.

12. On the purge confirmation dialog, select the purge method for the query:

    - **Soft purge**: Moves items to the recoverable items folder and retains them based on your retention settings. This option isn't supported for Teams messages and OneDrive and SharePoint files. If the search query includes unsupported items, the soft purge option isn't available.
    - **Hard purge**: Permanently deletes items from their data sources. This action can't be undone.

> ⓘ Note
>
> If items are under a litigation hold, eDiscovery hold, or retention policy, the purge process preserves the items and can't permanently delete them until the hold is removed or the retention period expires.

13. Select **Confirm purge**.
14. To view the status of the purge query, select the **Activities** tab.

After the purge is complete, you can download a report as a .zip file that includes item, location, settings, and summary information about the purged items.

Purge a single item

To purge a single item from an investigation, complete the following steps:

1. Go to Data Security Investigations in the Microsoft Purview portal and sign in with the credentials for a user account assigned Data Security Investigations permissions.
2. Select **Investigations** in the left navigation.
3. Select an investigation, then select **Analysis** to access the investigation scope.
4. Select the item you want to purge from the investigation scope.
5. Select **Add to mitigation**, and then select **Mitigation** to see the Mitigation list.
6. Select the item from the Mitigation list.
7. Select **Add to purge queue**, then select **Purge queue** to view the Purge queue.
8. Select the item from the Purge queue, then select **Purge**.
9. Select the purge method for the item:
    - **Soft purge**: Moves the item to the recoverable items folder and retains it based on your retention settings. This option is only available for Exchange mailbox items.
    - **Hard purge**: Permanently deletes the item from its data source. This action can't be undone.

> ⓘ Note

If items are under a litigation hold, eDiscovery hold, or retention policy, the purge process preserves the items and can't permanently delete them until the hold is removed or the retention period expires.

10. Select **Confirm purge**.

11. To view the status of the purge, select the **Activities** tab.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 06/22/2026

Use the Activity history in Data Security Investigations

The Activity history for Data Security Investigations shows information about activities you perform in each area of your investigation. It includes activities associated with searches, investigation scopes, AI analysis, and mitigation activities.

Each time you create or request an activity in a workflow area, the system automatically starts and logs specific activities in the Activity history. For example, if you search for data and select **Estimate scope** in an investigation, the **Generate statistics** and **Generate sample** activities appear in the Activity history list for the investigation.

To view and manage activities, complete the following steps:

1. Go to Data Security Investigations in the Microsoft Purview portal and sign in with the credentials for a user account assigned Data Security Investigations permissions.
2. Select **Investigations** in the left navigation.
3. Select an investigation, then select the **Activities** tab.

Activity history dashboard

The Activity history dashboard lists all the activities for an investigation and shows the following information for each activity:

- **Activity type**: The type of activity.
- **Status**: The status of the activity.
- **Export name**: The name of the export. For non-export activity types, this value is blank.
- **Created**: The date and time the activity was created.
- **Completed**: The date and time the activity was completed.
- **Duration**: The duration of the activity.
- **Created by**: The user that created the activity.

To customize the columns display for the Activity history, select **Customize columns** to choose the columns to display or drag and drop the columns in the list to reorder. To download the list of activities and the column information, select **Download list** to create a .csv file containing this information.

Group activities

When you view a large list of activities for an area, it's often helpful to group activities. Select **Group** to group review sets by the following parameters:

- **None**: Activities are ungrouped in the Activity history.
- **Activity type**: Activities grouped by the *Activity type* values.
- **Created by**: Activities grouped by the *Created by* values.

Activity details

Select an activity listed in the Activity history for more information in the following areas:

- **Overview** tab: Lists a summary of information for the activity. Includes a calculation for the time remaining to completion for in-progress activities. Select **Download report** to combine all **Overview** information into a single .zip file.
- **Settings** tab: Summarizes the **Statistics view** setting options selected for the investigation.

Progress bar

When you select an activity in the Activity history, a progress bar appears that shows details about the current state of the activity. For long-running activities, the progress bar has three phases:

- **Assessment phase**: In this phase, the progress bar submits the progress and backend services assess the scope of the activity. This activity includes examining the input conditions based on data sources, query, and relevant activity settings. The service analyzes the scope of the work and identifies the number of locations and number of items for activities. During this phase, a completion estimate isn't available and a flashing progress bar is displayed. The progress bar displays the status of the assessment while the scope is calculated. When all locations and items are identified, the activity transitions to the next phase.
- **In progress phase**: In this phase, the progress bar displays the current status, number of locations, and number of items for the activity. An estimated completion and current elapsed time are displayed in hours, minutes, and seconds.
- **Catch-up phase (optional)**: If there are locations or items with errors, activities are retried. The progress of any retries is displayed.

Download activity list

To download a .csv file of the activity items on the Activity history dashboard, select **Download list** on the dashboard command bar.

Activity types and descriptions

These activities start when users take actions to use and manage searches, investigation scopes, AI activities, and more.

Expand table

Activity type	Area	Description
Add to scope	Investigation scope	A user adds search results to an investigation scope. The process copies the items to an Azure Storage location and then reindexes those items in the Azure Storage location. This new index enables querying and analyzing items in the data set.
Audit search	Search	A user submits an audit search to query the unified audit log for content-associated activities. This activity queries the audit log and identifies content items associated with the specified activities.
Categorize	AI analysis	A user categorizes items in an investigation scope.
Examine	AI analysis	A user analyzes data in an investigation scope by running AI examination and recommendation tools, such as Credentials, Mitigation, Risks.
Generate sample	Search	After a user creates and runs a new search (or reruns an existing search), the search tool prepares a sample subset of items that match the search query for preview. Previewing search results helps you determine the effectiveness of the search.
Generate scope	Search	After a user creates and runs or reruns a search, the search tool searches the index for items that match the search query and prepares a scope estimate that includes the number and total size of all items by the search, the number of data sources searched, and other relevant pattern to the search hits like keywords, sensitive information types, and more.

> **Note**
>
> When an audit search is submitted, query chunking creates an *Add to scope* activity and multiple additional activities to process the results. Seeing multiple *Add to scope* activities after an audit search is expected behavior.

Activity status

The following table describes the different status states for activities:

Expand table

Status	Description
Abandoned	The system automatically stops the activity. This stoppage might occur due to excessive activities time or other reasons.
Canceled	The user cancels the activity.
Complete	The activity successfully completes.
Failed	The activity encounters an error and fails to complete. This error might be caused by a wrongly formatted query.
In progress	The activity is in progress. You can monitor the progress of the activity in the progress bar in the applicable area.

> **Note**
>
> In addition to tracking activities on this page, Data Security Investigations sends notifications through the Microsoft Purview Notification Center and email when setup completes and investigations are ready to use. For more information, see Email and portal notifications.

Note: The author created this article with assistance from AI. Learn more

Last updated on 06/11/2026

Application card: Microsoft Purview Data Security Investigations

What is an Application or Platform Card?

Microsoft's Application and Platform cards are intended to help you understand how our AI technology works, the choices application owners can make that influence application performance and behavior, and the importance of considering the whole application, including the technology, the people, and the environment. Application cards are created for AI applications and platform cards are created for AI platform services. These resources can support the development or deployment of your own applications and can be shared with users or stakeholders impacted by them.

As part of its commitment to responsible AI, Microsoft adheres to six core principles: fairness, reliability and safety, privacy and security, inclusiveness, transparency, and accountability. These principles are embedded in the Responsible AI Standard, which guides teams in designing, building, and testing AI applications. Application and Platform Cards play a key role in operationalizing these principles by offering transparency around capabilities, intended uses, and limitations. For further insight, readers are encouraged to explore Microsoft's Responsible AI Transparency Report and Code of Conduct, which outline how enterprise customers and individuals can engage with AI responsibly.

Overview

Microsoft Purview Data Security Investigations is an AI-enabled investigation application for cybersecurity and data security teams. It helps organizations analyze data associated with potential breaches, insider incidents, and exfiltration events by combining search, AI analysis, and mitigation planning in one workflow. The application uses generative AI to identify sensitive content risks, prioritize high-impact items, and produce recommendations that support containment and remediation.

The application is designed to reduce time and complexity in post-breach and high-risk investigations. Instead of manually reviewing very large sets of files, emails, and messages, analysts can use vector search, categorization, and examination tools to focus on the most relevant and risky content first. This improves triage efficiency, supports cross-team coordination,

and helps organizations take timely actions such as investigation-driven mitigation planning and purge operations.

Data Security Investigations is intended for enterprise customers and downstream users such as security analysts, incident responders, threat investigators, and data protection teams. Key documentation includes Learn about Data Security Investigations amd Learn about AI analysis in Data Security Investigations.

Key terms

The following list provides a glossary of key terms related to Microsoft Purview Data Security Investigations:

- **Categorization**: An AI capability that groups scoped investigation content into default, custom, or AI-suggested categories to speed prioritization. Standard and Advanced processing options help analysts move from broad triage to deeper topic-level clustering.
- **Examination**: An AI analysis mode that performs focused item-level review for areas such as credentials, risk, and mitigation guidance. Examination outputs help investigators understand severity and recommended next actions.
- **Investigation scope**: The set of content items selected for review and analysis in an investigation.
- **Mitigation plan**: A structured set of actions for items identified as high priority during investigation. It helps analysts track and execute remediation and risk-reduction tasks.
- **Purge (soft/hard)**: A mitigation action to remove risky items identified in an investigation. Soft purge moves items to recoverable locations, while hard purge permanently deletes items.
- **Risk score**: A score generated during examination that helps analysts prioritize content based on potential security impact. It supports triage and escalation decisions.
- **AI Search (preview)**: A retrieval-augmented search experience that combines semantic retrieval with AI-generated summaries and citations. It helps analysts assess whether search results match investigative intent.
- **Vector search**: A semantic search method that uses embeddings to retrieve contextually relevant content, even when exact keywords are missing. It supports intent-aware queries and relevance-based ranking.

Key features or capabilities

The key features and capabilities outlined here describe what Microsoft Purview Data Security Investigations is designed to do and how it performs across supported tasks.

- **Integrated incident-to-investigation workflow**: The application supports investigation creation from Microsoft Defender XDR incidents, Insider Risk Management cases, Data Security Posture Management (preview) insights, or manual draft mode. This helps security teams move from alerting to data-focused analysis without switching tools. It also improves collaboration between cybersecurity, compliance, and risk teams.
- **AI-powered semantic retrieval**: Vector search enables intent-aware retrieval across scoped, text-based content. Analysts can find relevant items even when content does not contain exact query terms. This improves recall and reduces manual keyword trial-and-error.
- **AI-assisted prioritization through categorization**: Categorization organizes data into default, custom, and AI-suggested categories so analysts can quickly focus on higher-risk areas. Advanced categorization provides additional topic-level organization for deeper review. This reduces review effort in large investigations by surfacing prioritized content instead of requiring full manual review first.
- **Targeted risk and credential examination**: Examination tools extract credentials, assess item risk, and generate mitigation-oriented outputs for selected content. Analysts get a faster understanding of what might be exposed and what to do next. This supports quicker containment and remediation in active incidents.
- **Mitigation and purge actions**: Investigators can convert findings into concrete actions through mitigation plans and purge workflows. This ties analysis outputs directly to risk-reduction execution. Teams can track progress and status of mitigation decisions within the investigation lifecycle.
- **Audit and ecosystem integration**: The solution integrates with Microsoft Purview audit workflows and Microsoft security solutions, helping teams preserve evidence context and align response actions. It also supports investigation activity logging in the unified audit log to strengthen operational accountability.

Intended uses

Microsoft Purview Data Security Investigations can be used in multiple scenarios. Some examples of use cases include:

- **Breach response from a Defender XDR incident**: An analyst can create new Data Security Investigation directly from a Defender XDR incident to assess the blast radius of a compromised user. AI-powered examination surfaces credentials, sensitive documents, and communications the attacker could have accessed, with regulated and customer-impacting

data prioritized first. Mitigation actions are taken within the same investigation, producing a documented response for audit and follow-up.

- **Tracing employee data exfiltration through the audit log**: When an employee is suspected of exfiltrating a large volume of data, an investigator queries the Unified Audit Log through Data Security Investigations to trace files downloaded and accessed. AI categorization separates routine activity from suspicious content and focuses attention on the highest-risk data. AI enriched findings are handed off to legal and privacy teams for further actions.
- **Hunting sensitive data in an overexposed SharePoint site**: After learning that a SharePoint site containing sensitive content is broadly overshared, an investigator scopes an investigation to that site and runs AI analysis to identify what is actually at risk. Vector search and categorization surface concentrations of sensitive content that would be impractical to review manually.
- **Reviewing risky AI interactions**: An investigator scopes a DSI investigation to AI prompts and responses involving sensitive data, whether from Microsoft Copilot or third-party AI tools. Examinations identify where regulated content, credentials, or PII appeared in interactions, and categorization highlights recurring patterns of risky use. Findings inform both immediate remediation on exposed content and broader follow-up with the users and tools involved.

Models and training data

Microsoft Purview Data Security Investigations leverages a variety of AI models to power the experience that users see. Some examples include Azure OpenAI Service models and Microsoft Security Copilot platform capabilities used for vector embeddings, categorization, and examination-related AI outputs. To learn more about the data used to train the foundation models behind Microsoft Purview Data Security Investigations, refer to the linked model cards to find the relevant data cards.

- Transparency note for Azure OpenAI
- Privacy and data security in Microsoft Security Copilot
- Responsible AI practices at Microsoft

Performance

Microsoft Purview Data Security Investigations is designed to perform reliably in enterprise post-breach and data risk investigation workflows where investigation scope is properly defined and users are assigned the required role groups. In these conditions, analysts can iteratively refine

search results, prioritize data through categorization, and run examination on selected items to support mitigation decisions.

The intended inputs are primarily enterprise content and investigator prompts. Inputs include files, emails, and messages scoped from supported Microsoft 365 data sources; metadata and audit-linked context; and natural language or keyword queries used in vector search and Search with AI (preview). The expected outputs include relevance-ranked search results, category groupings, risk and credential examination outputs, and mitigation-oriented recommendations.

The primary evaluated and documented AI interaction modality is text. Vectorization and semantic retrieval depend on text-bearing content. Some content types that do not include usable text are excluded from vectorization workflows, so investigators should account for this during scope planning.

Performance and usefulness are influenced by investigation setup quality, including scope precision, category selection, and iterative human review. In large datasets, categorization returns prioritized subsets rather than exhaustive coverage for each category. For comprehensive item-level analysis, teams should combine categorization with examination and investigator validation.

Limitations

Understanding Microsoft Purview Data Security Investigations's limitations is crucial to determine if it is used within safe and effective boundaries. While we encourage customers to leverage Microsoft Purview Data Security Investigations in their innovative solutions or applications, it's important to note that Microsoft Purview Data Security Investigations was not designed for every possible scenario. We encourage users to refer to either the Microsoft Enterprise AI Services Code of Conduct (for organizations) or the Code Conduct section in the Microsoft Services Agreement (for individuals) as well as the following considerations when choosing a use case:

- **AI output reliability constraints**: Results generated by generative AI might not always be fully accurate or complete. This limitation is important because investigators may otherwise overestimate confidence and miss context that changes incident severity. Users should validate critical findings, especially before legal, compliance, or irreversible technical actions. Human judgment remains required in final decisions.
- **Text-content dependency for semantic tooling**: Vector search and related semantic workflows require text-bearing content. Certain content types, such as image-only items or content without extractable text, may be excluded from vectorization and therefore from some AI-assisted retrieval workflows. Teams should evaluate whether key evidence might

require additional review methods outside semantic search. This helps avoid false assumptions about complete coverage.

- **Prioritized-not-exhaustive categorization behavior**: Categorization is optimized to surface high-relevance content for selected categories, not to guarantee full review of every item in scope per category. In large datasets, relevant lower-scoring items might not appear in category outputs. Users should use examination for comprehensive item-level analysis where complete coverage is required. This is especially important in high-impact investigations.
- **Operational limits and capacity boundaries**: Investigation and feature limits apply, including limits on total documents, total file size per investigation, and limits in audit search and purge workflows. These constraints can affect how teams segment work and sequence large investigations. Users should plan scope and processing strategy around published limits to avoid delays. Capacity-aware planning improves both cost control and investigative continuity.
- **Not intended for unsupported scenarios or prohibited use**: Data Security Investigations is intended for data security incident investigation and mitigation contexts. Using it as a sole decision-maker, outside governance and security workflows, or in ways prohibited by terms and code-of-conduct requirements can create safety and compliance risk. Organizations should keep the tool within intended operational boundaries and apply layered controls. This includes legal and policy review before high-consequence actions.

Evaluations

Performance and safety evaluations assess whether AI applications are operating reliably and securely by examining factors like groundedness, relevance, and coherence while identifying the risks of generating harmful content. The following evaluations were conducted with safety components already in place, which are also described in 9. Safety Components and Mitigations.

Performance and quality evaluations

Performance and quality evaluations for AI features in Microsoft Purview Data Security Investigations focus on practical investigation outcomes, including vector search relevance, categorization accuracy, and examination output usefulness. Based on product documentation, evaluation activities include pre-release testing and scenario-based validation to assess whether AI-assisted discovery, prioritization, and analysis features help investigators act on identified risks.

For current AI experiences, the primary evaluated modality is text. Ideal outcomes are accurate, relevant findings that help analysts quickly identify high-risk items and take informed

investigative actions. Suboptimal outcomes include results that are incomplete, irrelevant, or require substantial additional analysis and validation.

Risk and safety evaluations

Our evaluation data is custom-built to assess AI application performance across key areas of safety and quality, simulating real-world scenarios and risks. We begin by identifying relevant evaluation aspects of concern based on multi-disciplinary research and expert input. These concerns are translated into targeted evaluation objectives and guide formulation of evaluation metrics. For safety, we create adversarial prompts to elicit undesirable or edge-case responses, which are then scored using AI-assisted annotators trained to assess alignment with Microsoft's safety standards. For quality, we craft rubric-based prompts relevant to scenarios including evaluating retrieval-augmented generation (RAG) applications and agents. Datasets are curated from diverse sources including synthetic and public datasets to simulate real-world user scenarios. Using the curated datasets, both evaluations undergo iterative refinement and human alignment to improve metric efficacy and reliability. This methodology forms the foundation of repeatable, rigorous assessments that reflect how customers use evaluations to build better and safer AI.

Custom evaluations

Data Security Investigations was evaluated using multiple quality and reliability indicators, including AI accuracy, result relevance, and output clarity. Based on documented practices, evaluations use a combination of manual red teaming and human grading to test whether AI-assisted outputs are useful and dependable for incident investigation workflows. The primary evaluated modality for current AI experiences is text.

Evaluation scenarios include semantic retrieval, categorization relevance, and examination outputs used to prioritize and mitigate risk. Ideal results are contextually relevant findings that help analysts quickly identify high-risk items and take informed next actions. Suboptimal results include incomplete, unclear, or less relevant outputs that require additional iteration and validation by investigators.

Operational feedback is also used to improve performance over time. The product collects selected quality-focused metrics about AI inputs and outputs to help tune relevance and usefulness, while documented privacy controls describe what content is and is not collected. This continuous loop supports ongoing quality improvements after release.

Safety components & mitigations

- **Role-based access controls**: Data Security Investigations uses dedicated role groups (Admins, Investigators, Reviewers) to control feature access and responsibilities. This helps limit who can run sensitive activities such as purge and administrative configuration. Clear separation of duties reduces misuse risk and strengthens accountability in investigation workflows.
- **Human-in-the-loop decision making**: AI outputs are designed to assist analysts, not replace investigator judgment. Teams are expected to validate high-impact findings and mitigation recommendations before taking consequential action. This mitigates overreliance and reduces the chance of acting on incomplete AI output.
- **Scope refinement and exclusion controls**: Investigators can iteratively narrow investigation scope and exclude irrelevant items before advanced AI analysis. This improves relevance, reduces noise, and lowers exposure to unnecessary content processing. Focused scope management also supports cost control and faster incident response.
- **Audit logging and evidentiary traceability**: Investigation-related activities are logged through integrated auditing workflows. Logging supports forensic traceability, operational review, and compliance accountability. This is a key mitigation for governance and post-incident validation requirements.
- **Privacy and data handling safeguards**: Documented controls describe tenant-isolated investigation storage and privacy protections for AI processing. Data sharing, logging, and scanning controls in integrated processing paths are documented as off by default for specific Copilot processing contexts. These safeguards help reduce privacy risk and support enterprise compliance expectations.

Best practices for deploying and adopting Microsoft Purview Data Security Investigations

Responsible AI is a shared commitment between Microsoft and its customers. While Microsoft builds AI applications with safety, fairness, and transparency at the core, customers play a critical role in deploying and using these technologies responsibly within their own contexts. To support this partnership, we offer the following best practices for deployers and end users to help customers implement responsible AI effectively. Deployers and end-users (admins) should:

- Exercise caution and evaluate outcomes when using Microsoft Purview Data Security Investigations for consequential decisions or in sensitive domains: Consequential decisions are those that may have a legal or significant impact on a person's access to education, employment, financial platforms, government benefits, healthcare, housing, insurance, legal platforms, or that could result in physical, psychological, or financial harm. Sensitive

domains—such as financial platforms, healthcare, and housing—require particular care due to the potential for disproportionate impact on different groups of people. When using AI for decisions in these areas, make sure that impacted stakeholders can understand how decisions are made, appeal decisions, and update any relevant input data.

- Evaluate legal and regulatory considerations: Customers need to evaluate potential specific legal and regulatory obligations when using any AI platforms and solutions, which may not be appropriate for use in every industry or scenario. Additionally, AI platforms or solutions are not designed for and may not be used in ways prohibited in applicable terms of service and relevant codes of conduct.
- **Plan investigations for iterative analysis**: Data Security Investigations workflows are intentionally iterative, especially for search refinement, scope tuning, and AI-assisted triage. Teams should define repeatable checkpoints for when to run vector search, categorization, and examination. This improves consistency, reduces unnecessary compute use, and strengthens investigation quality.
- **Coordinate cross-functional response early**: Security incidents often involve security, IT, legal, compliance, and business owners. Bring these stakeholders into the investigation process as soon as high-risk findings are identified. Early coordination improves containment speed and reduces downstream remediation friction.

End-users should:

- **Use clear and focused investigation prompts**: For vector search and Search with AI (preview), prompts should include context, subject, and intent (for example, credentials in a named project or access-risk content for a specific incident). Clear prompts produce more relevant retrieval and reduce investigative noise. Users should iterate and compare outputs before drawing conclusions.
- **Validate AI findings with supporting evidence**: Review citations, item details, and activity context before escalating or mitigating. AI outputs can accelerate triage but may miss edge cases or include lower-confidence results. Cross-checking with search, scope views, and stakeholder input reduces error risk.
- Exercise human oversight when appropriate: Human oversight is an important safeguard when interacting with AI applications. While we continuously improve our AI applications, AI might still make mistakes. The outputs generated may be inaccurate, incomplete, biased, misaligned, or irrelevant to your intended goals. This could happen due to various reasons, such as ambiguity in the inputs or limitations of the underlying models. As such, users should review the responses generated by Microsoft Purview Data Security Investigations and verify that they match their expectations and requirements.

- Be aware of the risk of overreliance: Overreliance on AI happens when users accept incorrect or incomplete AI outputs, mainly because mistakes in AI outputs may be hard to detect. For the end-user, overreliance could result in decreased productivity, loss of trust, application abandonment, financial loss, psychological harm, physical harm, among others. (e.g. a doctor accepts an incorrect AI output).
- Exercise caution when designing agentic AI in sensitive domains: Users should exercise caution when designing and/or deploying agentic AI applications in sensitive domains where agent actions are irreversible or highly consequential.
- Additional precautions should also be taken when creating autonomous agentic AI as described further in either the Microsoft Enterprise AI Services Code of Conduct (for organizations) or the Code Conduct section in the Microsoft Services Agreement (for individuals).

Deployers should:

- **Configure permissions and ownership model upfront**: Assign Data Security Investigations role groups based on least privilege and operational responsibility before incidents occur. Include backup admins and clear role coverage to avoid response delays. This reduces risk from overbroad access and strengthens continuity during active incidents.
- **Design for limits, costs, and throughput**: Investigation, audit search, and purge limits should inform how teams segment large incidents. Teams should also monitor capacity usage and stage AI processing to align with urgency and budget. This avoids bottlenecks and supports predictable response performance.
- **Test mitigation workflows and escalation paths**: Run tabletop or pilot investigations using representative data to validate query patterns, categorization strategy, and mitigation handoffs. Pre-tested workflows reduce confusion during real incidents and improve time to action. Include checks for legal and compliance notification requirements as part of this testing.

Learn more about Microsoft Purview Data Security Investigations

For additional guidance or to learn more about the responsible use of Microsoft Purview Data Security Investigations, we recommend reviewing the following documentation:

- Learn about Data Security Investigations
- Get started with Data Security Investigations
- Learn about AI analysis in Data Security Investigations

- Use AI analysis in Data Security Investigations
- Learn about the Data Security Investigations workflow
- Assign permissions in Data Security Investigations
- Limits in Data Security Investigations
- Privacy FAQ for Data Security Investigations
- Microsoft Purview privacy statement
- Microsoft Trust Center

Learn more about responsible AI

- Microsoft AI principles
- Microsoft responsible AI resources
- Microsoft Azure Learning courses on responsible AI

Last updated on 06/17/2026

Privacy FAQ for Data Security Investigations

How is my data stored in Microsoft Purview Data Security Investigations?

When you use Data Security Investigations, Microsoft copies user queried organization data from the Microsoft 365 application storage to tenant-isolated regional investigation storage. This data is stored until Data Security Investigations administrators or investigators delete the investigation. Only Data Security Investigations Administrators, Investigators, and Reviewers have access to this data in your organization.

How is my data processed for Data Security Investigations AI features?

Data Security Investigations partners closely with Microsoft Security Copilot platform to power Data Security Investigations AI features. This means that data leaves Data Security Investigations (Microsoft 365 compliance area) and travels to Microsoft Security Copilot (Microsoft Azure compliance area) for up to 48 hours and return with AI generated content and insights. This generates the vector embeddings for vector search, category and subject generation, and examination results.

To ensure user privacy, all data sharing, logging, and scanning is turned off by default for Microsoft Security Copilot processing of Data Security Investigations data. For more information on Microsoft Security Copilot, see Privacy and data security in Microsoft Security Copilot

What feedback information does Data Security Investigations collect?

Data Security Investigations collects customer feedback to enhance its AI features and improve user experience. This feedback is gathered during the public preview stage and includes specific metrics related to the AI features' inputs and outputs.

- **What we collect**: We collect metrics such as the number of files returned to a vector search query, the total number of vectorized files in the user's data set, category impact scores, AI generated risk and relevance scores of files, and the number of files selected for AI

examination. These metrics help us understand the effectiveness and relevance of our AI features.

- **What we don't collect**: We don't collect customer content or any personal data from the files themselves. Our focus is solely on collecting metrics and qualitative data about the inputs and outputs of our AI services to improve their performance and accuracy.

By collecting these metrics, Data Security Investigations aims to continuously improve its AI features and provide users with more accurate and relevant results. Your feedback is invaluable in helping us achieve this goal.

Last updated on 06/11/2026

Limits in Data Security Investigations

This article describes the limits in Microsoft Purview Data Security Investigations.

Investigation limits

The following table lists the limits for investigations in Data Security Investigations.

Expand table

Description of limit	Data Security Investigations feature support
Total number of documents that you can add to an investigation.	40 million
Total file size per investigation.	1 TB

Purge limits

The following table lists the limits for purge in Data Security Investigations.

Expand table

Description of limit	Data Security Investigations feature support
Maximum number of items per purge search query.	10,000

Search and integration limits

The following table lists the search and integration limits for Data Security Investigations. These limits apply to searches and investigations created from the following solutions:

- Audit search
- Endpoint DLP evidence collection. For more information about endpoint DLP, see Learn about endpoint data loss prevention.
- Insider Risk Management
- Microsoft Defender XDR

Expand table

Description of limit	Data Security Investigations feature support
Maximum items returned per search or integration query.	Approximately 3,000
Maximum concurrent audit searches per user.	10

> **Tip**
>
> If your investigation involves more than approximately 3,000 audit items, you can split the data into smaller time-based slices. For more information, see Work with large audit search results.

Note: The author created this article with assistance from AI. Learn more

Last updated on 09/24/2026

Partially indexed items in Data Security Investigations

Summarize this article for me

Data Security Investigations searches automatically include partially indexed items in the estimated search results. Partially indexed items are Exchange mailbox items and documents on SharePoint and OneDrive sites that aren't completely indexed for search. In Exchange, a partially indexed item typically contains a file (of a file type that can't be indexed) attached to an email message.

The following list describes other reasons why items can't be indexed for search and are returned as partially indexed items:

- The file type is unrecognized or unsupported for indexing.
- Messages have an attached file that can't be opened. This issue is the most common cause of partially indexed email items.
- The file type is supported for indexing but an indexing error occurred for a specific file.
- Too many files attached to an email message.
- A file attached to an email message is too large.
- A file is encrypted with non-Microsoft technologies.
- A file is password-protected.

> **Note**
>
> Most organizations have less than 1% of content by volume and less than 12% by size that is partially indexed. The reason for the difference between volume and size is that larger files have a higher probability of containing content that can't be fully indexed.

File types not indexed for search

Certain types of files, such as Bitmap (.bmp) or MP3 (.mp3) files, don't contain content that can be indexed. The search indexing servers in Exchange and SharePoint don't perform full-text indexing on these types of files and they're considered as unsupported file types. There are also file types for which full-text indexing is disabled, either by default or by an administrator. Unsupported and disabled file types are labeled as unindexed items in searches.

For a list of supported and disabled file formats, see the following articles:

- **Exchange**: File formats indexed by Exchange Search
- **Exchange**: Get-SearchDocumentFormat

- **SharePoint**: Default crawled file name extensions and parsed file types in SharePoint

Messages and documents with partially indexed file types returned in search results

Not every email message with a partially indexed file attachment or every partially indexed SharePoint document is automatically returned as a partially indexed item. That's because other message or document properties, such as the **Subject** property in email messages and the **Title** or **Author** properties for documents, are indexed and available to be searched. For example, a keyword search for *financial* returns items with a partially indexed file attachment if that keyword appears in the subject of an email message or in the file name or title of a document. However, if the keyword appears only in the body of the file, the message or document is returned as a partially indexed item.

Similarly, messages with partially indexed file attachments and documents of a partially indexed file type are included in search results when other message or document properties, which are indexed and searchable, match the search criteria. Message properties that are indexed for search include sent and received dates, sender and recipient, the file name of an attachment, and text in the message body. Document properties indexed for search include created and modified dates. So even though a message attachment might be a partially indexed item, the message is included in the regular search results if the value of other message or document properties matches the search criteria.

> **Note**
>
> If a mailbox item is moved from a folder that is indexed to a folder that isn't indexed, a flag is set to unindex the item and the item is removed from the index and isn't searchable. Later, if that same item is moved back to a folder that is indexed, the flag isn't reset. That means the item remains unindexed, and not searchable.

To perform a deeper analysis of partially indexed items, use Purview eDiscovery to export and review partially indexed items. For more information, see Partially indexed items in eDiscovery.

More information about partially indexed items

- Because message and document properties and their metadata are indexed, a keyword search might return results if that keyword appears in the indexed metadata. However, that same keyword search might not return the same item if the keyword only appears in the content of an item with an unsupported file type. In this case, the item is returned as a partially indexed item.

- Although a file type is supported for indexing and is indexed, indexing or search errors can cause a file to be returned as a partially indexed item. For example, searching a large Excel file might be partially successful (because the first 4 MB are indexed), but then fails because the file size limit is exceeded. In this case, it's possible that the same file is returned with the search results and as a partially indexed item.
- Files that are encrypted with Microsoft encryption technologies (and stored in SharePoint or OneDrive) are partially indexed.
- Email messages encrypted with S/MIME are partially indexed. This limitation includes encrypted messages with or without file attachments.
- Email messages protected using Azure Rights Management are indexed and are included in the search results if they match the search query.

Last updated on 02/18/2026

Supported file types in Data Security Investigations

Microsoft Purview Data Security Investigations supports Optical Character Recognition (OCR) text extraction for supported image file types. The following table lists the currently supported image file types and indicates if a file type is supported for file identification, metadata extraction, and OCR text extraction.

Image

Extraction of images is part of adding items to an investigation scope and automatic OCR processing isn't an additional charge for your organization.

Expand table

Mime type	File identification	Metadata extraction	OCR text extraction	Possible Extensions
image/bmp	Yes	Yes	Yes	.bmp
image/emf	Yes	Yes	Yes	.emf
image/gif	Yes	Yes	Yes	.gif
image/jpeg	Yes	Yes	Yes	.jpeg; .jpg
image/png	Yes	Yes	Yes	.png
image/svg+xml	Yes	Yes	Yes	.svg
image/tiff	Yes	Yes	Yes	.tif
image/vnd.dwg	Yes	Yes	Yes	.dwg; .dxf
image/wmf	Yes	Yes	Yes	.wmf

> **Note**
>
> The **OCR text extraction** column indicates that you can extract text from these image formats when data is automatically vectorized. OCR text extraction occurs automatically during data preparation and the extracted text is vectorized for use in AI-based analysis tools.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/11/2026

Insider Risk Management

09/26/2025

> **Important**
>
> Microsoft Purview Insider Risk Management correlates various signals to identify potential malicious or inadvertent insider risks, such as IP theft, data leakage, and security violations. Insider Risk Management enables customers to create policies to manage security and compliance. Built with privacy by design, users are pseudonymized by default, and role-based access controls and audit logs are in place to help ensure user-level privacy.

Employees now have more access to create, manage, and share data across a broad spectrum of platforms and services. In most cases, organizations have limited resources and tools to identify and mitigate organization-wide risks while also meeting compliance requirements and employee privacy standards. These risks include potential data theft by departing employees and risk of data leaks of information outside your organization by accidental oversharing or malicious intent.

Microsoft Purview Insider Risk Management uses the full breadth of service and third-party indicators to help you quickly identify, triage, and act on potentially risky activity. By using logs from Microsoft 365 and Microsoft Graph, Insider Risk Management allows you to define specific policies to identify risk indicators. After identifying the risks, you can take action to mitigate these risks, and if necessary open investigation cases and take appropriate legal action.

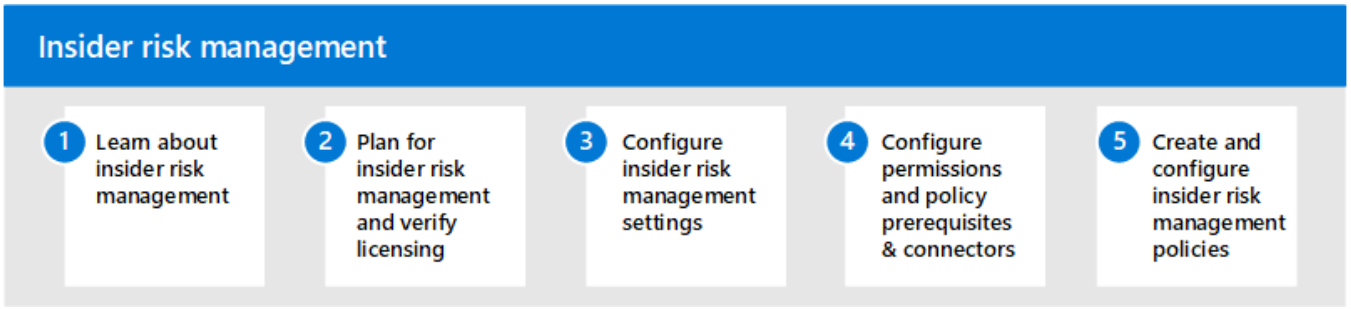
Watch the following videos to learn how Insider Risk Management can help your organization prevent, detect, and contain risks:

Insider Risk Management solution & development:
https://learn-video.azurefd.net/vod/player?id=13a0ac77-ac1a-49e0-9b8b-4af3e6df6053&locale=en-us&embedUrl=%2Fpurview%2Finsider-risk-management-solution-overview

Insider Risk Management workflow:
https://learn-video.azurefd.net/vod/player?id=bf70614d-1ba7-4f60-978d-080ae514eb6f&locale=en-us&embedUrl=%2Fpurview%2Finsider-risk-management-solution-overview

Configure Insider Risk Management

Use the following steps to configure Insider Risk Management for your organization:



1. Learn about Insider Risk Management
2. Plan for Insider Risk Management and verify licensing
3. Configure Insider Risk Management settings
4. Configure permissions and policy prerequisites & connectors
5. Create and configure Insider Risk Management policies

More information about Insider Risk Management

- Manage insider risk policies
- Investigate insider risk activities
- Act on insider risk cases

Learn about Insider Risk Management

> **Important**
>
> Microsoft Purview Insider Risk Management correlates various signals to identify potential malicious or inadvertent insider risks, such as IP theft, data leakage, and security violations. Insider Risk Management enables customers to create policies to manage security and compliance. Built with privacy by design, users are pseudonymized by default, and role-based access controls and audit logs are in place to help ensure user-level privacy.

Microsoft Purview Insider Risk Management is a compliance solution that helps minimize internal risks by enabling you to detect, investigate, and act on malicious and inadvertent activities in your organization. With insider risk policies, you can define the types of risks to identify and detect in your organization. You can also set up processes for acting on cases and escalating cases to Microsoft eDiscovery (Premium) if needed. Risk analysts in your organization can quickly take appropriate actions to make sure users are compliant with your organization's compliance standards.

For more information and an overview of the planning process to address potentially risky activities in your organization that might lead to a security incident, see Starting an Insider Risk Management program.

Watch the following videos to learn how Insider Risk Management can help your organization prevent, detect, and contain risks while prioritizing your organization values, culture, and user experience:

Insider Risk Management solution & development:

https://learn-video.azurefd.net/vod/player?id=13a0ac77-ac1a-49e0-9b8b-4af3e6df6053&locale=en-us&embedUrl=%2Fpurview%2Finsider-risk-management

Insider Risk Management workflow:

https://learn-video.azurefd.net/vod/player?id=bf70614d-1ba7-4f60-978d-080ae514eb6f&locale=en-us&embedUrl=%2Fpurview%2Finsider-risk-management

Check out the Microsoft Mechanics video on how Insider Risk Management and Communication Compliance work together to help minimize data risks from users in your organization.

> **Important**
>
> Insider Risk Management is currently available in tenants hosted in geographical regions and countries supported by Azure service dependencies. To verify that Insider Risk Management is supported for your organization, see Azure dependency availability by country/region.

Modern risk pain points

Managing and minimizing risk in your organization starts with understanding the types of risks found in the modern workplace. Some risks come from external events and factors that you can't directly control. Other risks come from internal events and user actions that you can minimize and avoid. Some examples are risks from illegal, inappropriate, unauthorized, or unethical behavior and actions by users in your organization. These behaviors include a broad range of internal risks from users:

- Leaks of sensitive data and data spillage
- Confidentiality violations
- Intellectual property (IP) theft
- Fraud
- Insider trading
- Regulatory compliance violations

Users in the modern workplace can create, manage, and share data across a broad spectrum of platforms and services. In most cases, organizations have limited resources and tools to identify and mitigate organization-wide risks while also meeting user privacy standards.

Insider Risk Management uses the full breadth of service and third-party indicators to help you quickly identify, triage, and act on risk activity. By using logs from Microsoft 365 and Microsoft Graph, Insider Risk Management enables you to define specific policies to identify risk indicators. These policies help you identify risky activities and act to mitigate these risks.

Insider Risk Management centers around the following principles:

- **Transparency**: Balance user privacy versus organization risk with privacy-by-design architecture.
- **Configurable**: Configurable policies based on industry, geographical, and business groups.
- **Integrated**: Integrated workflow across Microsoft Purview solutions.
- **Actionable**: Provides insights to enable reviewer notifications, data investigations, and user investigations.

Identifying potential risks with analytics

Insider risk analytics enables you to evaluate potential insider risks in your organization without configuring any insider risk policies. This evaluation can help your organization identify potential areas of higher user risk and help determine the type and scope of Insider Risk Management policies you might consider configuring. This evaluation can also help you determine needs for additional licensing or future optimization of existing insider risk policies.

For more information about insider risk analytics, see Insider Risk Management settings: Analytics.

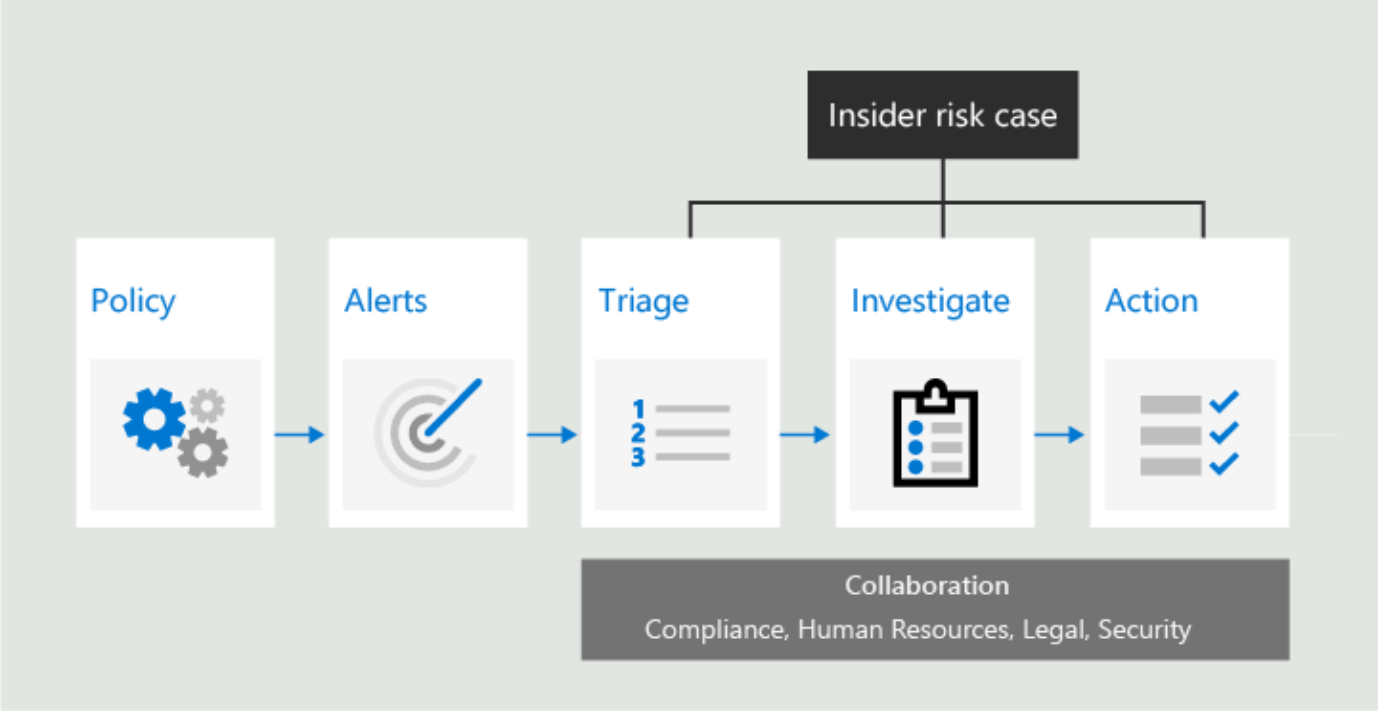
Get started with recommended actions

Whether you're setting up Insider Risk Management for the first time or getting started with creating new policies, the new recommended actions experience can help you get the most out of Insider Risk Management capabilities. Recommended actions include setting up permissions, choosing policy indicators, creating a policy, and more.

Workflow

The Insider Risk Management workflow helps you identify, investigate, and take action to address internal risks in your organization. With focused policy templates, comprehensive activity signaling across the Microsoft 365 service, and alert and case management tools, you can use actionable insights to quickly identify and act on risky behavior.

Identifying and resolving internal risk activities and compliance issues with Insider Risk Management uses the following workflow:



Policies

You create Insider Risk Management policies by using predefined templates and policy conditions that define which triggering events and risk indicators to examine in your organization. These conditions include how risk indicators are used for alerts, which users are included in the policy, which services are prioritized, and the detection time period.

To quickly get started with Insider Risk Management, select from the following policy templates:

- Data theft by departing users
- Data leaks
- Data leaks by priority users
- Data leaks by risky users
- Patient data misuse (preview)
- Risky Agents
- Risky AI usage
- Risky browser usage (preview)
- Security policy violations
- Security policy violations by departing users
- Security policy violations by risky users
- Security policy violations by priority users

Alerts

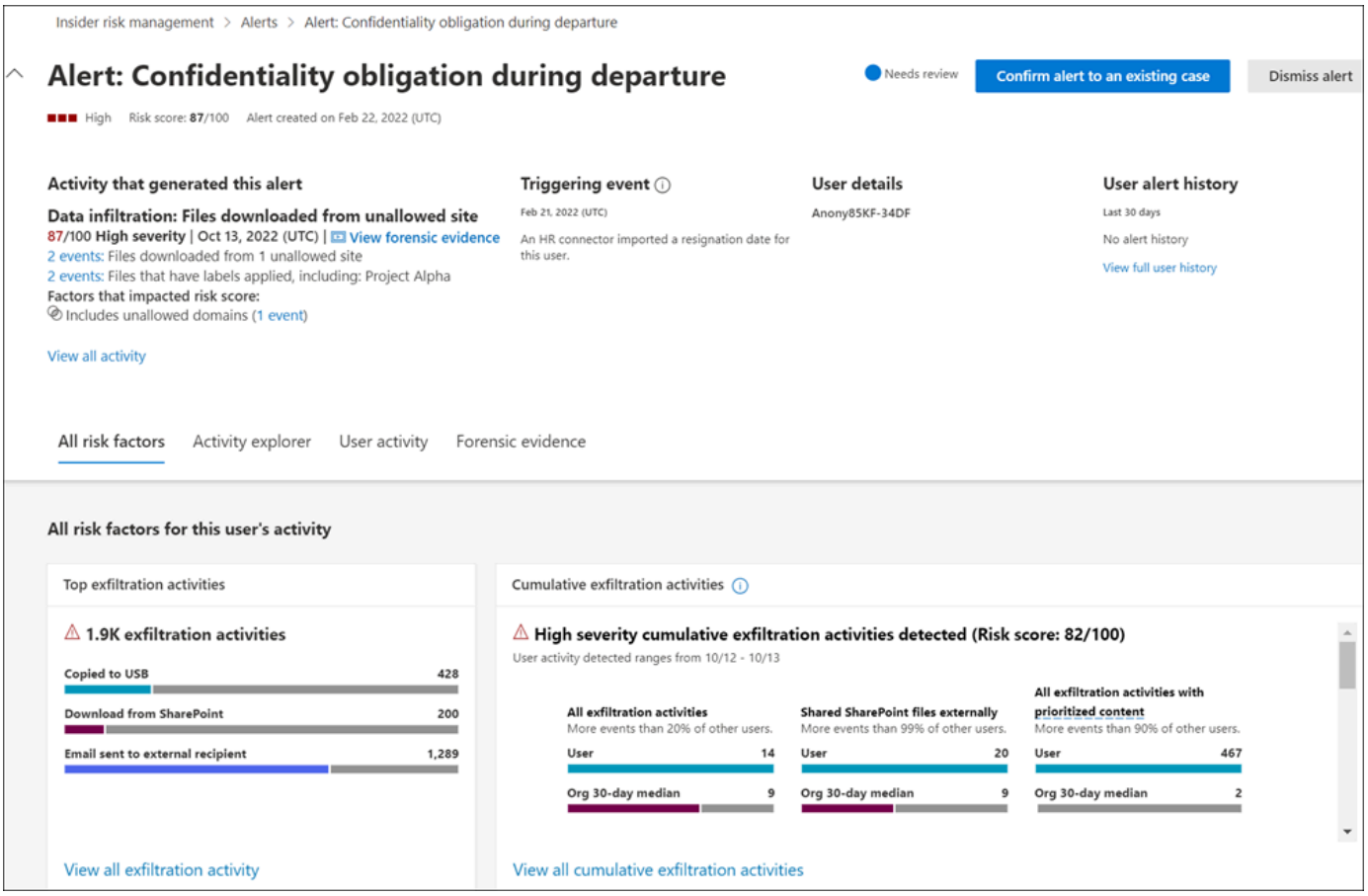
Risk indicators automatically generate alerts when they match policy conditions. You can see these alerts in the Alerts dashboard or the Triage Agent dashboard. These dashboards provide a quick view of all alerts that need review, open alerts over time, and alert statistics for your organization. All policy alerts display the following information to help you quickly identify the status of existing alerts and new alerts that need action:

- ID
- Users
- Alert
- Status
- Alert severity
- Time detected
- Case
- Case status
- Risk factors

Triage

New user activities that need investigation automatically generate alerts and assign them a *Needs review* status. Reviewers can quickly identify, review, evaluate, and triage these alerts.

Resolve alerts by opening a new case, assigning the alert to an existing case, or dismissing the alert. With alert filters, you can quickly identify alerts by status, severity, or time detected. As part of the triage process, reviewers can view alert details for the activities identified by the policy, view user activity associated with the policy match, see the severity of the alert, and review user profile information.



Investigate

Quickly investigate all risky activities for a selected user with User activity reports (preview). These reports let investigators in your organization examine activities for specific users during a defined time period without temporarily or explicitly assigning them to an Insider Risk Management policy. After examining activities for a user, investigators can dismiss individual activities as benign, share or email a link to the report with other investigators, or choose to assign the user temporarily or explicitly to an Insider Risk Management policy.

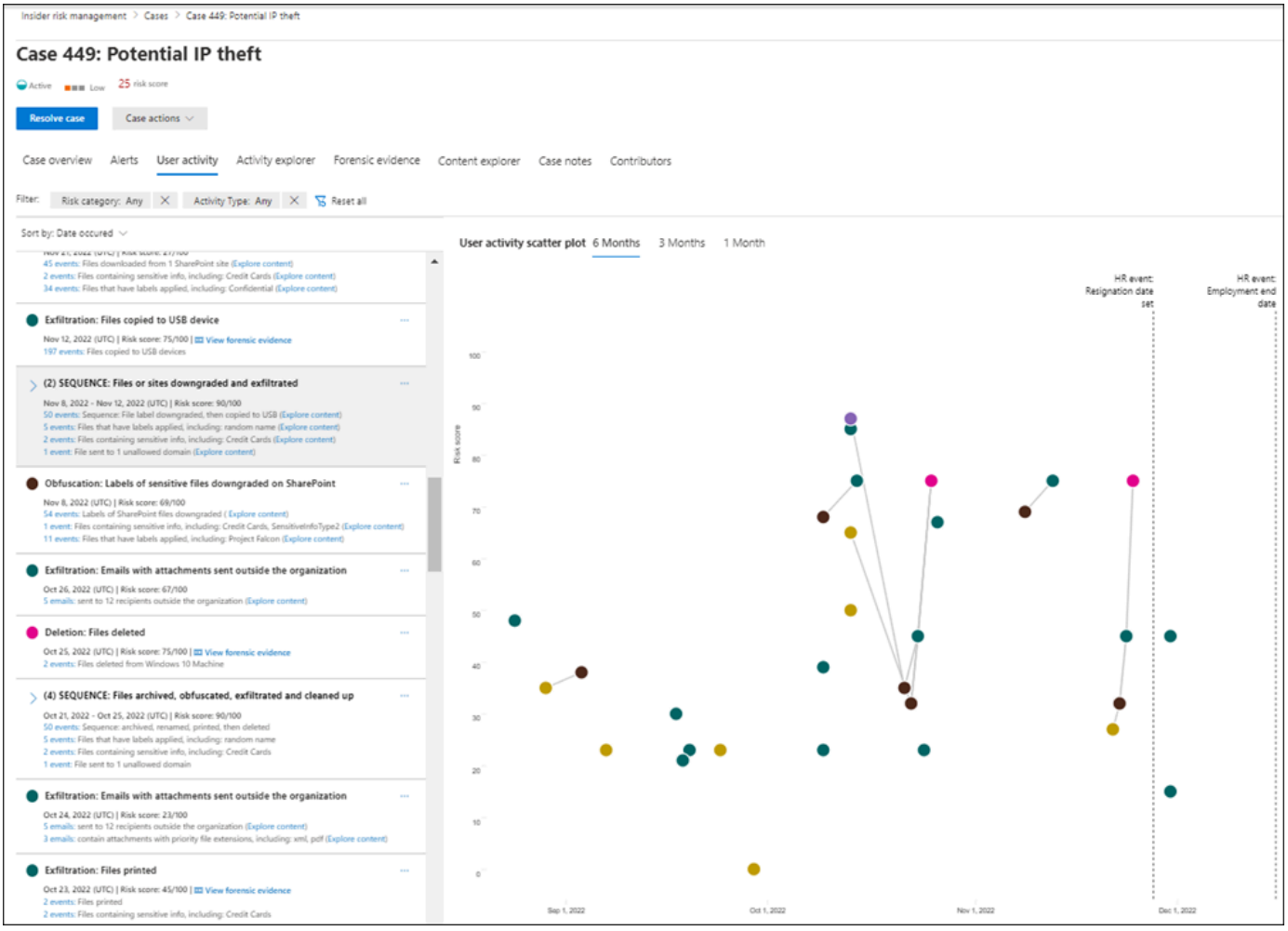
Cases are created for alerts that require deeper review and investigation of the activity details and circumstances around the policy match. The **Case dashboard** provides an all-up view of all active cases, open cases over time, and case statistics for your organization. Reviewers can quickly filter cases by status, the date the case was opened, and the date the case was last updated.

Selecting a case on the case dashboard opens the case for investigation and review. This step is the heart of the Insider Risk Management workflow. This area synthesizes risk activities, policy conditions, alert details, and user details into an integrated view for reviewers. The primary investigation tools in this area are:

- **User activity**: An interactive chart automatically displays user risk activity. It plots activities over time and by risk level for current or past risk activities. Reviewers can quickly filter and

view the entire risk history for the user and drill into specific activities for more details.

- **Content explorer**: The Content explorer automatically captures and displays all data files and email messages associated with alert activities. Reviewers can filter and view files and messages by data source, file type, tags, conversation, and many more attributes.
- **Case notes**: Reviewers can provide notes for a case in the Case Notes section. This list consolidates all notes in a central view and includes reviewer and date submitted information.



Additionally, the new Audit log (preview) enables you to stay informed of the actions taken on Insider Risk Management features. This resource allows an independent review of the actions taken by users assigned to one or more Insider Risk Management role groups.

Action

After investigating cases, reviewers can quickly act to resolve the case or collaborate with other risk stakeholders in your organization. If users accidentally or inadvertently violate policy conditions, reviewers can send a simple reminder notice to the user from notice templates you can customize for your organization. These notices might serve as simple reminders or might

direct the user to refresher training or guidance to help prevent future risky behavior. For more information, see Insider Risk Management notice templates.

In more serious situations, you might need to share the Insider Risk Management case information with other reviewers or services in your organization. Insider Risk Management tightly integrates with other Microsoft Purview solutions to help you with end-to-end risk resolution.

- **eDiscovery (Premium)**: Escalating a case for investigation allows you to transfer data and management of the case to Microsoft Purview eDiscovery (Premium). eDiscovery (Premium) provides an end-to-end workflow to preserve, collect, review, analyze, and export content that's responsive to your organization's internal and external investigations. It allows legal teams to manage the entire legal hold notification workflow. To learn more about eDiscovery (Premium) cases, see Overview of Microsoft Purview eDiscovery (Premium).
- **Office 365 Management APIs integration (preview)**: Insider Risk Management supports exporting alert information to security information and event management (SIEM) services via the Office 365 Management APIs. Having access to alert information in the platform that best fits your organization's risk processes gives you more flexibility in how to act on risk activities. To learn more about exporting alert information with Office 365 Management APIs, see Export alerts.

Scenarios

Insider Risk Management can help you detect, investigate, and take action to mitigate internal risks in your organization in several common scenarios:

Data theft by departing users

When users leave an organization, either voluntarily or as the result of termination, legitimate concerns often arise that company, customer, and user data are at risk. Users might innocently assume that project data isn't proprietary, or they might be tempted to take company data for personal gain and in violation of company policy and legal standards. Insider Risk Management policies that use the Data theft by departing users policy template automatically detect activities typically associated with this type of theft. With this policy, you automatically receive alerts for suspicious activities associated with data theft by departing users so you can take appropriate investigative actions. You need to configure a Microsoft 365 HR connector for your organization for this policy template.

Intentional or unintentional leak of sensitive or confidential information

In most cases, users try their best to properly handle sensitive or confidential information. But occasionally users make mistakes and information is accidentally shared outside your organization or in violation of your information protection policies. In other circumstances, users might intentionally leak or share sensitive and confidential information with malicious intent and for potential personal gain. Insider Risk Management policies created using the following policy templates automatically detect activities typically associated with sharing sensitive or confidential information:

- Data leaks
- Data leaks by priority users
- Data leaks by risky users
- Risk AI usage

Intentional or unintentional security policy violations (preview)

Users typically have a large degree of control when managing their devices in the modern workplace. This control might include permissions to install or uninstall applications needed in the performance of their duties or the ability to temporarily disable device security features. Whether this risk activity is inadvertent, accidental, or malicious, this conduct can pose risk to your organization and is important to identify and act to minimize. To help identify these risky security activities, the following Insider Risk Management security policy violation templates scores security risk indicators and use Microsoft Defender for Endpoint alerts to provide insights for security-related activities:

- Security policy violations
- Security policy violations by departing users
- Security policy violations by priority users
- Security policy violations by risky users

Policies for users based on position, access level, or risk history

Users in your organization might have different levels of risk depending on their position, level of access to sensitive information, or risk history. This structure might include members of your organization's executive leadership team, IT administrators that have extensive data and network access privileges, or users with a past history of risky activities. In these circumstances, closer inspection and more aggressive risk scoring are important to help surface alerts for investigation

and quick action. To help identify risky activities for these types of users, you can create priority user groups and create policies from the following policy templates:

- Security policy violations by priority users
- Data leaks by priority users

Healthcare (preview)

For organizations in the healthcare industry, recent studies found a very high rate of insider-related data breaches. Detecting misuse of patient data and health record information is a critical component of safeguarding patient privacy and complying with compliance regulation such as the Health Insurance Portability and Accountability Act (HIPAA) and the Health Information Technology for Economic and Clinical Health (HITECH) Act. Patient data misuse can range from accessing privileged patient records to accessing records of patients from family or neighbors with malicious intent. To help identify these types of risky activities, the following Insider Risk Management policy template uses the Microsoft 365 HR connector and a healthcare-specific data connector to start scoring risk indicators relating to behaviors that can occur within your electronic heath record (EHR) systems:

- Patient data misuse (preview)

Actions and behaviors by risky users

Employment stressor events can impact user behavior in several ways that relate to insider risks. These stressors might be a poor performance review, a position demotion, or the user being placed on a performance review plan. Stressors might also result in potentially inappropriate behavior such as users sending potentially threatening, harassing, or discriminatory language in email and other messages. Though most users don't respond maliciously to these events, the stress of these actions can result in some users behaving in ways that they might not normally consider during normal circumstances. To help identify these types of potentially risky activities, the following Insider Risk Management policy templates can use the HR connector and/or integration with a dedicated Communication Compliance policy to bring users into scope for Insider Risk Management policies and start scoring risk indicators relating to behaviors that might occur:

- Data leaks by risky users
- Risk AI usage
- Risky browser usage (preview)
- Security policy violations by risky users

Visual context for potentially risky user activities with forensic evidence

Having visual context is crucial for security teams during forensic investigations to get better insights into potentially risky user activities that might lead to a security incident. This insight might include visual capturing of these activities to help evaluate if they're indeed risky or taken out of context and not potentially risky. For activities that are determined to be risky, having forensic evidence captures can help investigators and your organization better mitigate, understand, and respond to these activities. To help with this scenario, enable forensic evidence capturing for online and offline devices in your organization.

Ready to get started?

- See Plan for Insider Risk Management to prepare for enabling Insider Risk Management policies in your organization.
- See Get started with Insider Risk Management settings to configure global settings for insider risk policies.
- See Get started with Insider Risk Management to configure prerequisites, create policies, and start receiving alerts.

Last updated on 06/26/2026

Plan for Insider Risk Management

08/20/2025

> **Important**
>
> Microsoft Purview Insider Risk Management correlates various signals to identify potential malicious or inadvertent insider risks, such as IP theft, data leakage, and security violations. Insider Risk Management enables customers to create policies to manage security and compliance. Built with privacy by design, users are pseudonymized by default, and role-based access controls and audit logs are in place to help ensure user-level privacy.

Before you get started with Insider Risk Management in your organization, review important planning activities and considerations with your information technology and compliance management teams. Thoroughly understanding and planning for deployment in the following areas helps ensure that your implementation and use of Insider Risk Management features goes smoothly and aligns with best practices.

For more information and an overview of the planning process to address risky activities in your organization, see Starting an Insider Risk Management program.

To learn how the Insider Risk Management workflow can help your organization prevent, detect, and contain risks while prioritizing your organization values, culture, and user experience, watch the following video:
https://learn-video.azurefd.net/vod/player?id=bf70614d-1ba7-4f60-978d-080ae514eb6f&locale=en-us&embedUrl=%2Fpurview%2Finsider-risk-management-plan

Check out the Microsoft Mechanics video on how Insider Risk Management and Communication Compliance work together to help minimize data risks from users in your organization.

> **Tip**
>
> If you're not an E5 customer, use the 90-day Microsoft Purview solutions trial to explore how additional Purview capabilities can help your organization manage data security and compliance needs. Start now at the Microsoft Purview trials hub. Learn details about signing up and trial terms.

Work with stakeholders in your organization

Identify the appropriate stakeholders in your organization to collaborate for taking actions on Insider Risk Management alerts and cases. Consider including the following areas of your organization in initial planning and the end-to-end Insider Risk Management workflow:

- Information technology
- Compliance
- Privacy
- Security
- Human resources
- Legal

Determine regional compliance requirements

Different geographic and organizational areas might have compliance and privacy requirements that differ from other areas of your organization. Work with the stakeholders in these areas to ensure they understand the compliance and privacy controls in Insider Risk Management and how to use them across different areas of your organization. In some scenarios, compliance and privacy requirements might require policies that designate or restrict some stakeholders from investigations and cases based on the case for a user or regulatory or policy requirements for the area.

If you have requirements for specific stakeholders to be involved in case investigations that involve users in certain regions, roles, or divisions, consider implementing separate (even if identical) Insider Risk Management policies that target the different regions and populations. This configuration makes it easier for the right stakeholders to triage and manage cases that are relevant to their roles and regions. Consider creating processes and policies for regions where investigators and reviewers speak the same language as the users. This approach can help streamline the escalation process for Insider Risk Management alerts and cases.

Plan permissions to support the review and investigation workflow

Depending on how you want to manage Insider Risk Management policies and alerts, assign users to specific role groups to manage different sets of Insider Risk Management features. You can assign users with different compliance responsibilities to specific role groups to manage different areas of Insider Risk Management features. Or, assign all user accounts for designated administrators, analysts, investigators, and viewers to the *Insider Risk Management* role group. For more information, see Assign Insider Risk Management permissions.

Understand requirements and dependencies

Depending on how you plan to implement Insider Risk Management policies, you need the proper Microsoft 365 licensing subscriptions. You also need to understand and plan for some solution prerequisites.

Licensing: Microsoft offers Insider Risk Management as part of a wide selection of Microsoft 365 licensing subscriptions. For details, see the Getting started with Insider Risk Management article.

> **Important**
>
> Microsoft currently offers Insider Risk Management in tenants hosted in geographical country/regions supported by Azure service dependencies. To verify that Microsoft supports Insider Risk Management for your organization, see **Azure dependency availability by country/region**.

If you don't have an existing Microsoft 365 Enterprise E5 plan and want to try Insider Risk Management, you can add Microsoft 365 to your existing subscription or sign up for a trial of Microsoft 365 Enterprise E5.

Policy template requirements: Depending on the policy template you choose, understand the following requirements and plan accordingly prior to configuring Insider Risk Management in your organization:

- For the **Data theft by departing users** template, configure a Microsoft 365 HR connector to periodically import resignation and termination date information for users in your organization. For step-by-step guidance to configure the Microsoft 365 HR connector, see the Import data with the HR connector article.
- For the **Data leaks** template, configure at least one Microsoft Purview Data Loss Prevention (DLP) policy to define sensitive information in your organization and to receive insider risk alerts for High Severity DLP policy alerts. For step-by-step guidance to configure DLP policies, see the Create and Deploy data loss prevention policies article.
- For the **Security policy violation** template, enable Microsoft Defender for Endpoint for Insider Risk Management integration in the Defender Security Center to import security violation alerts. For step-by-step guidance to enable Defender for Endpoint integration with Insider Risk Management, see Configure advanced features in Microsoft Defender for Endpoint.
- For the **Risky user** template, configure a Microsoft 365 HR connector to periodically import performance or demotion status information for users in your organization. For

step-by-step guidance to configure the Microsoft 365 HR connector, see the Import data with the HR connector article.

Test with a small group of users in a production environment

Before enabling this solution broadly in your production environment, consider testing the policies with a small set of production users so you can complete the necessary compliance, privacy, and legal reviews in your organization. Evaluating Insider Risk Management in a test environment requires that you generate simulated user actions and other signals to create alerts for triage and cases for processing. This approach might not be practical for many organizations, so we recommend that you test Insider Risk Management with a small group of users in a production environment.

Keep the anonymization feature in policy settings enabled to anonymize user display names in the Insider Risk Management console during this testing to maintain privacy within the tool. This setting helps protect the privacy of users that have policy matches and can help promote objectivity in data investigation and analysis reviews for insider risk alerts.

If you don't see any alerts immediately after configuring an Insider Risk Management policy, it might mean the minimum risk threshold isn't met yet. Check the **Users** page to verify that the policy is triggered and working as expected and to see if users are in-scope for the policy.

Migrating between Microsoft 365 US Government Cloud and the commercial cloud

If you migrate your organization from the Microsoft 365 US Government Cloud to the worldwide commercial cloud or from the worldwide commercial cloud to the Government Cloud, active cases and alerts don't migrate. Close any alerts and cases before starting the migration.

Resources for stakeholders

Share Insider Risk Management documentation with the stakeholders in your organization that are included in your management and remediation workflow:

- Create and manage insider risk policies
- Investigate insider risk activities
- Take action on insider risk cases
- Review case data with the insider risk Content explorer

- Create insider risk notice templates

Ready to get started?

Ready to configure Insider Risk Management for your organization? We recommend that you review the following articles:

- Get started with Insider Risk Management settings to configure global policy settings.
- Get started with Insider Risk Management to configure prerequisites, create policies, and start receiving alerts.
- Get started with Insider Risk Management forensic evidence for step-by-step guidance to configure forensic evidence capturing in your organization.

Get started with Insider Risk Management

> **Important**
>
> Microsoft Purview Insider Risk Management correlates various signals to identify potential malicious or inadvertent insider risks, such as IP theft, data leakage, and security violations. Insider Risk Management enables customers to create policies to manage security and compliance. Built with privacy by design, users are pseudonymized by default, and role-based access controls and audit logs are in place to help ensure user-level privacy.

Use Insider Risk Management policies to identify risky activities and management tools to act on risk alerts in your organization. Complete the following steps to set up prerequisites and configure an Insider Risk Management policy.

> **Important**
>
> The Insider Risk Management solution provides a tenant-level option to help customers facilitate internal governance at the user level. Tenant-level administrators can set up permissions to provide access to this solution for members of your organization and set up data connectors in the Microsoft Purview portal to import relevant data to support user-level identification of potentially risky activity. Customers acknowledge insights related to the individual user's behavior, character, or performance materially related to employment can be calculated by the administrator and made available to others in the organization. In addition, customers acknowledge that they must conduct their own full investigation related to the individual user's behavior, character, or performance materially related to employment, and not just rely on insights from the Insider Risk Management service. Customers are solely responsible for using the Insider Risk Management service, and any associated feature or service in compliance with all applicable laws, including laws relating to individual user identification and any remediation actions.

For more information about how insider risk policies can help you manage risk in your organization, see Learn about Insider Risk Management.

Subscriptions and licensing

Before getting started with Insider Risk Management, confirm your Microsoft 365 subscription and any add-ons. To access and use Insider Risk Management, administrators need to verify their organization has a supported subscription and assign the appropriate

licenses to users. For more information about subscriptions and licensing, see the subscription requirements for Insider Risk Management.

> **Important**
>
> Insider Risk Management is currently available in tenants hosted in geographical regions and countries supported by Azure service dependencies. To verify that Insider Risk Management is supported for your organization, see **Azure dependency availability by country/region**.

If you don't have an existing Microsoft 365 Enterprise E5 plan and want to try Insider Risk Management, you can add Microsoft 365 to your existing subscription or sign up for a trial of Microsoft 365 Enterprise E5.

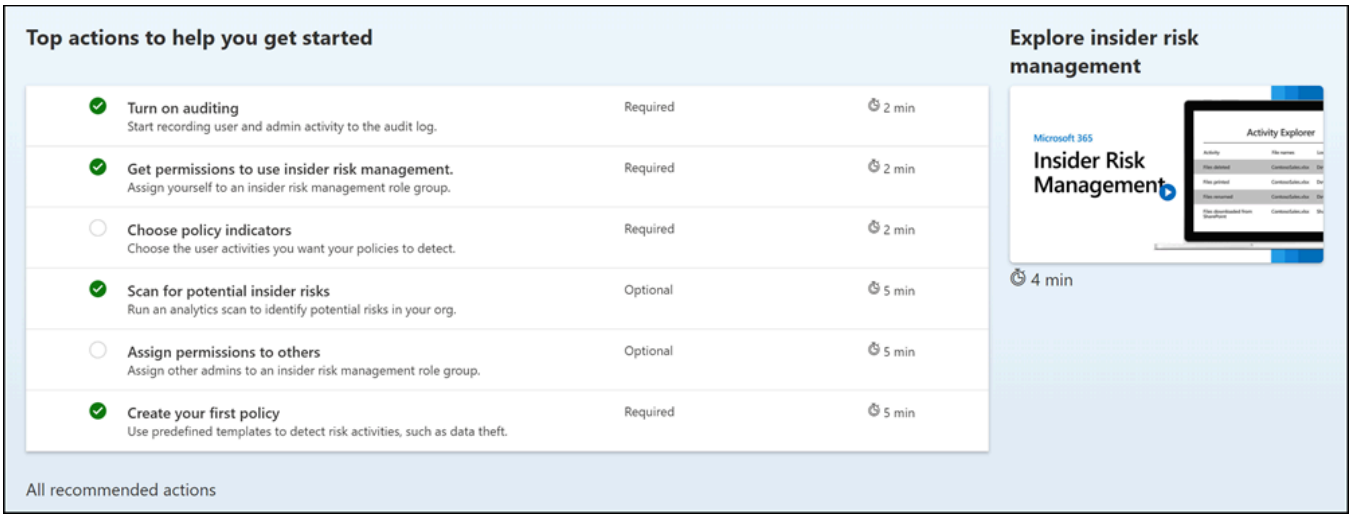
Pay-as-you-go billing

Some indicators included in Insider Risk Management are only available if you enable the pay-as-you-go billing model for your organization. For more information, see Configure policy indicators in Insider Risk Management.

This feature uses pay-as-you-go billing or per-user licensing for Microsoft Purview capabilities or both. To help you understand and manage your usage, Microsoft Purview provides a **Usage center** in the Microsoft Purview portal. For more information, see Manage pay-as-you-go and per-user licensing usage.

Recommended actions

Recommended actions can help your organization quickly get started with Insider Risk Management. Recommended actions, which are included on the **Overview** page, help guide you through the steps to configure and deploy policies.



The following recommendations help you get started with or maximize your Insider Risk Management configuration:

- **Turn on auditing**: When turned on, Microsoft 365 records user and admin activity in your organization to the audit log. Insider risk policies and analytics scans use this log to detect risk activities.
- **Get permissions to use Insider Risk Management**: The level of access you have to Insider Risk Management features depends on which role group you're assigned. To access and configure recommended actions, users must be assigned to the *Insider Risk Management* or *Insider Risk Management Admins* role groups.
- **Choose policy indicators**: Indicators are essentially the risk management activities you want to detect and investigate. You can choose indicators to detect activity across several Microsoft 365 locations and services.
- **Scan for potential insider risks**: Run an analytics scan to discover potential insider risks occurring in your organization. After evaluating results, review recommended policies to set up.
- **Assign permissions to others**: If there are additional team members who are responsible for managing insider risk features, assign them to the appropriate role groups.
- **Create your first policy**: To receive alerts on potentially risky activities, set up policies based on predefined templates that define the user activities you want to detect and investigate.
- **Configure data risk graph**: Use data risk graph to view connections between assets, users, and their activities using an interactive graph experience.

Each recommended action included in this experience has four attributes:

- **Action**: Name and description of the recommended action.
- **Status**: Status of the recommended action. Values are *Not started*, *In progress*, *Saved for later*, or *Completed*.
- **Required or optional**: Whether the recommended action is required or optional for Insider Risk Management features to function as expected.

- **Estimated time to complete**: Estimated time to complete the recommended action in minutes.

Select a recommendation from the list to get started with configuring Insider Risk Management. Each recommended action guides you through the required action for the recommendation, including any requirements, what to expect, and the impact of configuring the feature in your organization. Each recommended action is automatically marked as complete when configured or you can manually select the action as complete when configured.

Step 1 (required): Assign permissions for Insider Risk Management

To use any of the Insider Risk Management-related tools in the Microsoft Purview portal, users need the appropriate permissions. The easiest way to assign roles is to add the user to the appropriate role group on the **Role groups** page in the Microsoft Purview portal.

For step-by-step guidance, see Assign permissions in Insider Risk Management.

Step 2 (required): Enable the Microsoft 365 audit log

Insider Risk Management uses Microsoft 365 audit logs for user insights and risk management activities identified in policies and analytics insights. The Microsoft 365 audit logs are a summary of all activities within your organization and Insider Risk Management policies might use these activities for generating policy insights.

Auditing is enabled for Microsoft 365 organizations by default. Some organizations might have disabled auditing for specific reasons. If auditing is disabled for your organization, it might be because another administrator has turned it off. We recommend confirming that it's OK to turn auditing back on when completing this step.

For step-by-step instructions to turn on auditing, see Turn audit log search on or off. After you turn on auditing, a message is displayed that says the audit log is being prepared and that you can run a search in a couple of hours after the preparation is complete. You only have to do this action once. For more information about the using the Microsoft 365 audit log, see Search the audit log.

Step 3 (optional): Enable and view insider risk analytics insights

If you enable Insider Risk Management analytics, you can:

- **Scan for potential insider risks before creating policies.** You can conduct an evaluation of potential insider risks in your organization without configuring any insider risk policies. This evaluation can help your organization identify potential areas of higher user risk and help determine the type and scope of Insider Risk Management policies you might want to configure. This evaluation might also help you determine needs for additional licensing or future optimization of existing policies. Analytics scan results might take up to 48 hours before insights are available as reports for review. To learn more about analytics insights, see Insider Risk Management settings: Analytics and check out the Insider Risk Management Analytics video to help understand how analytics can help accelerate the identification of potential insider risks and help you to quickly take action.
- **Receive real-time recommendations for indicator threshold settings.** Manually tuning policies to reduce "noise" can be a very time-consuming experience that requires you to do a lot of trial and error to determine the desired configuration for your policies. If analytics is turned on, Insider Risk Management can provide real-time recommendations for indicator thresholds. You can also manually adjust the provided recommendations and see in real time how many users are brought into scope of the policy based on the changes you make. Learn more about real-time indicator threshold recommendations

> **Note**
>
> To enable insider risk analytics, you must be a member of the *Insider Risk Management*, *Insider Risk Management Admins*, or Microsoft 365 *Global admin* role group.

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

Enable insider risk analytics

1. Sign in to the Microsoft Purview portal with an admin account in your Microsoft 365 organization.
2. Go to the **Insider Risk Management** solution.

3. On the **Overview** tab, on the **Scan for insider risks in your organization** card, select **Run scan**. This action turns on analytics scanning for your organization. You can also turn on scanning by going to **Insider risk settings** > **Analytics** and enabling **Scan your tenant's user activity to identify potential insider risks**.
4. On the **Analytics details** pane, select **Run scan to start the scan for your organization**. Analytics scan results might take up to 48 hours before insights are available as reports for review.

After reviewing the analytics insights, choose the insider risk policies and configure the associated prerequisites that best meet your organization's insider risk mitigation strategy.

Step 4 (recommended): Configure prerequisites for policies

Most Insider Risk Management policies have prerequisites that you must configure for policy indicators to generate relevant activity alerts. Configure the appropriate prerequisites depending on the policies you plan to configure for your organization.

Connect to cloud apps in Microsoft Defender

Insider Risk Management includes the following cloud indicators (preview):

- Cloud storage indicators, including Google Drive, Box, and Dropbox
- Cloud service indicators, including Amazon S3 and Azure (Storage and SQL Server)

Cloud storage indicators

Use cloud storage indicators to detect the following activities in Google Drive, Box, and Dropbox:

- **Discovery:** Techniques used to figure out the environment
- **Collection:** Techniques used to gather data of interest
- **Exfiltration:** Techniques used to steal data, such as sensitive documents
- **Deletion (impact):** Techniques used to disrupt the availability or compromise the integrity of a system

Cloud service indicators

Use cloud service indicators to detect the following activities in Amazon S3 and Azure:

- **Defense evasion:** Techniques used to avoid detection of risky activities by disabling trace logs or by updating or deleting SQL Server firewall rules
- **Exfiltration:** Techniques used to steal data, such as sensitive documents
- **Deletion (impact):** Techniques used to disrupt the availability or compromise the integrity of a system
- **Privilege escalation:** Techniques used to gain higher-level permissions to systems and data

Prerequisites for accessing cloud indicators

To select from cloud indicators in Insider Risk Management settings and policies, you must first connect to the relevant cloud apps in Microsoft Defender, if you didn't already.

After connecting to the apps, you can find the indicators on the Policy indicators settings page and from individual policies.

Configure Insider Risk Indicator (preview) connector

You can extend Insider Risk Management by importing detections for non-Microsoft (third-party) workloads. For example, you might want to extend your detections to include Salesforce and Dropbox activities and use them alongside the built-in detections provided by the Insider Risk Management solution, which focuses on Microsoft services like SharePoint Online and Exchange Online.

To bring your own detections to the Insider Risk Management solution, import preprocessed, aggregated detections from security information and event management (SIEM) solutions such as Microsoft Sentinel or Splunk. Import a sample file into the Insider Risk Indicators connector workflow. The connector workflow analyzes the sample file and configures the required schema.

> **Note**
>
> Currently, you can't import "raw" detection signals into Insider Risk Management. You can only import preprocessed aggregations as a file.

You can use a custom indicator as:

- A trigger to bring a user into the scope of a policy.
- A policy indicator to score the user for risk.

For step-by-step guidance to configure the Insider Risk Indicators connector for your organization, see the Insider Risk Indicators connector article. After you configure the

connector, return to these configuration steps.

Configure Microsoft 365 HR connector

Insider Risk Management supports importing user and log data imported from third-party risk management and human resources platforms. The Microsoft 365 Human Resources (HR) data connector allows you to pull in human resources data from CSV files, including user termination dates, last employment dates, performance improvement plan notifications, performance review actions, and job level change status. This data helps drive alert indicators in Insider Risk Management policies and is an important part of configuring full risk management coverage in your organization. If you configure more than one HR connector for your organization, Insider Risk Management automatically pulls indicators from all HR connectors.

The Microsoft 365 HR connector is required when using the following policy templates:

- Data leaks by risky users
- Departing user data theft
- Patient data misuse
- Security policy violations by departing users
- Security policy violations by risky users

For step-by-step guidance to configure the Microsoft 365 HR connector for your organization, see the Set up a connector to import HR data article. After you configure the HR connector, return to these configuration steps.

Configure a Microsoft Healthcare connector

Insider Risk Management supports importing user and log data from third-party electronic medical record (EMR) systems. The Microsoft Healthcare connector enables you to bring in activity data from your EMR system through CSV files, including improper patient record access, suspicious volume activities, and editing and exporting activities. This data helps drive alert indicators in Insider Risk Management policies and plays an important role in configuring full risk management coverage in your organization.

If you configure more than one Healthcare connector for your organization, Insider Risk Management automatically supports event and activity signals from all Healthcare connectors. The Microsoft 365 Healthcare connector is required when using the following policy template:

- Patient data misuse

For step-by-step guidance to configure a healthcare-specific connector for your organization, see Set up a connector to import healthcare data. After you configure a connector, return to

these configuration steps.

Configure data loss prevention (DLP) policies

Insider Risk Management supports using DLP policies to help identify the intentional or accidental exposure of sensitive information to unwanted parties for high severity level DLP alerts. When you configure an Insider Risk Management policy with any of the **Data leaks** templates, you can assign a specific DLP policy to the policy for these types of alerts.

> **Tip**
>
> You can also use Adaptive Protection in Insider Risk Management to dynamically apply DLP protection controls to high-risk users while maintaining productivity for lower-risk users. **Learn more about Adaptive Protection**.

Data loss policies help identify users to activate risk scoring in Insider Risk Management for high severity DLP alerts for sensitive information. They're an important part of configuring full risk management coverage in your organization. For more information about Insider Risk Management and DLP policy integration and planning considerations, see Insider Risk Management policies.

> **Important**
>
> Make sure you complete the following steps:
>
> - Understand and properly configure the in-scope users in both the DLP and Insider Risk Management policies to produce the policy coverage you expect.
> - Configure the **Incident reports** setting in the DLP policy for Insider Risk Management used with these templates for *High* severity level alerts. Insider Risk Management alerts aren't generated from DLP policies with the **Incident reports** field set at *Low* or *Medium*.

A DLP policy is optional when using the following policy templates:

- Data leaks
- Data leaks by priority users

For step-by-step guidance to configure DLP policies for your organization, see Create and Deploy data loss prevention policies. After you configure a DLP policy, return to these configuration steps.

> **Note**
>
> Endpoint DLP now supports virtualized environments, which means that the Insider Risk Management solution supports virtualized environments through endpoint DLP. Learn more about support for virtualized environments in endpoint DLP.

Configure sharing of insider risk levels with Microsoft Defender and DLP alerts

You can share insider risk levels from Insider Risk Management (preview) to bring unique user context to Microsoft Defender and DLP alerts. Insider Risk Management analyzes user activities over a period of 90-120 days and looks for anomalous behavior over that period. Adding this data to Microsoft Defender and DLP alerts enhances the data available in those solutions to help analysts prioritize alerts. Learn more about sharing user risk severity levels with Microsoft Defender and DLP alerts.

Sharing Insider Risk Management user risk severity levels also enhances the Microsoft Security Copilot. For example, in Security Copilot, you might want to start by asking Copilot to summarize a DLP alert, then ask Copilot to show the insider risk level associated with the user flagged in the alert. Or you might want to ask why the user is considered a high-risk user. The user risk information in this case comes from Insider Risk Management. Security Copilot seamlessly integrates Insider Risk Management with DLP to assist with investigations. Learn more about using the standalone version of Copilot for combined DLP/Insider Risk Management investigations.

Configure priority user groups

Insider Risk Management supports assigning priority user groups to policies to help identify unique risk activities for users with critical positions, high levels of data and network access, or a past history of risk behavior. Creating a priority user group and assigning users to the group helps scope policies to the unique circumstances presented by these users.

To enable the *priority user groups* risk score booster, go to the *Insider Risk Management settings* page, then select **Policy indicators** and **Risk score boosters**. Analysts and investigators can review and prioritize these users' risk severity to help triage alerts in accordance with your organization's risk policies and standards.

A priority user group is required when using the following policy templates:

- Security policy violations by priority users
- Data leaks by priority users

See the Getting started with Insider Risk Management settings article for step-by-step configuration guidance.

Configure Physical badging connector

Insider Risk Management supports importing user and log data from physical control and access platforms. The Physical badging connector lets you pull in access data from JSON files, including user IDs, access point IDs, access time and dates, and access status. This data helps drive alert indicators in Insider Risk Management policies and is an important part of configuring full risk management coverage in your organization. If you configure more than one Physical badging connector for your organization, Insider Risk Management automatically pulls indicators from all Physical badging connectors. Information from the Physical badging connector supplements other insider risk signals when using all insider risk policy templates.

> **Important**
>
> For Insider Risk Management policies to use and correlate signal data related to departing and terminated users with event data from your physical control and access platforms, you must also configure the Microsoft 365 HR connector. If you enable the Physical badging connector without enabling the Microsoft 365 HR connector, Insider Risk Management policies only process events for unauthorized physical access for users in your organization.

See the Set up a connector to import physical badging data article for step-by-step guidance to configure the Physical badging connector for your organization. After you configure the connector, return to these configuration steps.

Configure Microsoft Defender for Endpoint

Microsoft Defender for Endpoint is an enterprise endpoint security platform designed to help enterprise networks prevent, detect, investigate, and respond to advanced threats. To gain better visibility of security violations in your organization, you can import and filter Defender for Endpoint alerts for activities used in policies created from Insider Risk Management security violation policy templates.

If you create security violation policies, you need to configure Microsoft Defender for Endpoint in your organization and enable Defender for Endpoint for Insider Risk Management integration in the Defender Security Center to import security violation alerts. For more information about requirements, see the Minimum requirements for Microsoft Defender for Endpoint article.

For step-by-step guidance to configure Defender for Endpoint for Insider Risk Management integration, see the Configure advanced features in Defender for Endpoint article. After you configure Microsoft Defender for Endpoint, return to these configuration steps.

Configure forensic evidence

Having visual context is crucial for security teams during forensic investigations to get better insights into risky user activities that might lead to a security incident. With customizable event triggers and built-in user privacy protection controls, forensic evidence enables customizable capturing across devices to help your organization better mitigate, understand, and respond to potential data risks like unauthorized data exfiltration of sensitive data.

For step-by-step guidance to configure forensic evidence for your organization, see the Get started with Insider Risk Management forensic evidence article.

Configure optical character recognition

Microsoft Purview can scan for sensitive content in documents to help protect those documents from inappropriate exposure. When you enable optical character recognition (OCR) in Microsoft Purview, data classifiers, such as sensitive information types and trainable classifiers, can also detect characters in stand-alone images. After configuring OCR settings (preview), your existing insider risk policies apply to both images and documents.

For the OCR preview, Insider Risk Management supports scanning in the following locations: Windows endpoint devices, SharePoint Online, and Teams. Exchange Online and OneDrive aren't supported for the preview.

OCR settings don't apply to forensic evidence clips in Insider Risk Management.

Learn more about setting up OCR scanning and pay-as-you-go billing.

Step 5 (required): Configure insider risk settings

Insider risk settings apply to all Insider Risk Management policies, regardless of the template you choose when creating a policy. You configure settings by using **Settings** located at the top of Insider Risk Management pages. These settings control privacy, indicators, global exclusions, detection groups, intelligent detections, and more. Learn more about settings to consider before you creating a policy.

Step 6 (required): Create an Insider Risk Management policy

Insider Risk Management policies include assigned users and define which types of risk indicators are configured for alerts. Before potentially risky activities can trigger alerts, you must configure a policy. Use the policy workflow to create new Insider Risk Management policies.

> **Note**
>
> To create a custom trigger or indicator for a non-Microsoft workload, see Custom indicators.

Create policy

1. Sign in to the Microsoft Purview portal with an admin account in your Microsoft 365 organization.
2. Go to the **Insider Risk Management** solution.
3. Select **Policies** in the left navigation.
4. Select **Create policy** to open the policy workflow.
5. On the **Policy template** page, choose a policy category and then select the template for the new policy. These templates consist of conditions and indicators that define the risk activities you want to detect and investigate. Review the template prerequisites, triggering events, and detected activities to confirm this policy template fits your needs.

   > **Important**
   >
   > Some policy templates have prerequisites that you must configure for the policy to generate relevant alerts. If you didn't configure the applicable policy prerequisites, see Step 4.

   Select **Next** to continue.
6. On the **Name and description** page, complete the following fields:

   - **Name (required)**: Enter a friendly name for the policy. You can't change this name after you create the policy.

- **Description (optional)**: Enter a description for the policy.

Select **Next** to continue.

7. If admin units exist for your tenant, you see the **Admin units** page. Otherwise, you see the **Users and groups** page and can skip to the next step.

   If you want to scope the policy to one or more admin units, select **Add admin units**, then select the admin units that you want to apply to the policy.

   > ⓘ Note
   >
   > You can only see the admin units that are scoped to your role. If you're an unrestricted administrator, you can see all admin units for the tenant. To view a summary of the role groups and admin units that you're assigned to, select **View my permissions**.

   Select **Next** to continue.

8. On the **Users and groups** page, select one of the following options:

   - **Include all users and groups**. When you select this option, Insider Risk Management looks for triggering events for all users and groups in your organization to start assigning risk scores for the policy.

     If your policy is scoped by one or more admin units, this option selects all users and groups within the administrative units.

     > ⓘ Note
     >
     > To take advantage of real-time analytics (preview) for indicator threshold settings, you must scope your policy to **Include all users and groups**. Real-time analytics enables you to see estimates of the number of users that could potentially match a given set of policy conditions in real time. This feature helps you efficiently adjust the selection of indicators and thresholds of activity occurrence so you don't have too few or too many policy alerts. Scoping your policy to **Include all users and groups** also provides better overall protection across your tenant. For more information on real-time analytics for indicator threshold settings, see Indicator level settings.

     > ⓘ Note

> Insider Risk Management supports selection of following types of groups in policies - Microsoft 365 groups, Distribution groups, and Security groups (both mail-enabled and non-mail-enabled types) with a limitation that non-mail-enabled Security groups cannot be selected if your policy is **scoped by one or more admin units**. For more information on different types of groups, see **Compare Groups**.

- **Include specific users and groups**. Select this option to define which users or groups are included in the policy.

  If your policy is scoped by one or more admin units, you can only choose users within the admin unit scope.

  > ⓘ **Note**
  >
  > Guest accounts aren't supported.

- **Adaptive scope**. This option appears if you select the **Include specific users and groups** option. Select **Add or edit adaptive scopes** to apply an adaptive scope to the policy. You must create the adaptive scope before you create or edit the policy. If the policy is also scoped by one or more admin units, the adaptive scopes available to you are limited by the admin units. Learn how adaptive scopes work together with admin units.

- **Add or edit priority user groups**. This option appears only if you choose the *Data leaks by priority users* template. Select this option, then add or edit priority user groups.

  > ⓘ **Note**
  >
  > If the policy template is based on priority user groups, you can't select an **admin unit** to scope the policy. Priority user groups aren't currently supported for use with admin units.

Select **Next** to continue.

9. Use the **Exclude users and groups (optional) (preview)** page if you want to exclude certain users or groups from the policy scope. For example, you might want to create a policy that detects potentially risky actions for people in the whole organization but excludes executive-level sales managers. Select **Add users to exclude** or **Add groups to exclude** to select the users or groups that you want to exclude. If your policy is scoped by

one or more admin units, you can only exclude users or groups that are within the admin unit scope.

Select **Next** to continue.

10. On the **Content to prioritize** page, assign the sources to prioritize, if needed. Prioritizing these sources increases the chance of generating a high severity alert for these sources. Select one of the following choices:

    - **I want to prioritize content**. Select this option to prioritize *SharePoint sites*, *Sensitivity labels*, *Sensitive info types*, and *File extensions* content types. If you choose this option, you must select at least one priority content type.
    - **I don't want to specify priority content right now**. Select this option to skip the priority content detail pages in the workflow.

    Select **Next** to continue.

11. If you select **I want to prioritize content** in the previous step, you see the detail pages for *SharePoint sites*, *sensitivity labels*, *sensitive info types*, *file extensions*, and *Scoring*. Use these detail pages to define the SharePoint, sensitive info types, sensitivity labels, trainable classifiers, and file extensions to prioritize in the policy. The *Scoring* detail page allows you to scope the policy to only assign risk scores and generate alerts for specified activities that include priority content.

    - **SharePoint sites**: Select **Add SharePoint site** and select the SharePoint sites you have access to and want to prioritize. For example, "*group1@contoso.sharepoint.com/sites/group1*".

      > ⓘ **Note**
      >
      > If your policy is scoped by one or more admin units, you still see all SharePoint sites, not just SharePoint sites scoped to your admin units since admin units don't support SharePoint sites.

    - **Sensitive info type**: Select **Add sensitive info type** and select the sensitivity types you want to prioritize. For example, "*U.S. Bank Account Number*" and "*Credit Card Number*".
    - **Sensitivity labels**: Select **Add sensitivity label** and select the labels you want to prioritize. For example, "*Confidential*" and "*Secret*".
    - **Trainable classifiers**: Select **Add trainable classifier** and select the trainable classifiers you want to prioritize. For example, *Source code*.

- **File extensions**: Add up to 50 file extensions. You can include or omit the '.' with the file extension. For example, *.py* or *py* would prioritize Python files.
- **Scoring**: Decide whether to assign risk scores to all risk management activities detected by this policy or only for activities that include priority content. Choose **Get alerts for all activity** or **Get alerts only for activity that includes priority content**.

Select **Next** to continue.

12. If you select the *Data leaks* or *Data leaks by priority users* templates, you see options on the **Triggers for this policy** page for custom-triggering events and policy indicators. You can select a DLP policy or indicators for triggering events that bring users assigned to the policy in-scope for activity scoring. If you select the **User matches a data loss prevention (DLP) policy triggering event** option, you must select a DLP policy from the DLP policy dropdown list to enable triggering indicators for the DLP Policy for this Insider Risk Management policy. If you select the **User performs an exfiltration activity triggering event** option, you must select one or more of the listed indicators for the policy triggering event.

    > ⓘ **Note**
    >
    > Priority user groups aren't currently supported for admin units. If you're creating a policy based on the *Data leaks by priority users* template or the *Security policy violations by priority users* template, you can't select admin units for scoping the policy. Unrestricted administrators can select priority user groups without selecting admin units, but restricted or scoped administrators can't create these policies at all.

    > ⓘ **Important**
    >
    > If you're unable to select a listed indicator or sequence, it's because they aren't currently enabled for your organization. To make them available to select and assign to the policy, select the **Turn on indicators** prompt.

    If you select other policy templates, custom triggering events aren't supported. The built-in policy triggering events apply. Skip to Step 15 without defining policy attributes.

13. If you select the *Data leaks by risky users* or *Security policy violations by risky users* templates, you see options on the **Triggers for this policy** page for integration with Communication Compliance and HR data connector events. You can assign risk scores when users send messages that contain potentially threatening, harassing, or discriminatory language or bring users into the policy scope after risky user events are

reported in your HR system. If you select the **Risk triggers from Communication Compliance (preview)** option, you can accept the default Communication Compliance policy (automatically created), choose a previously created policy scope for this trigger, or create another scoped policy. If you select **HR data connector events**, you must configure an HR data connector for your organization.

Select **Next** to continue.

14. If you select the *Data leaks* or *Data leaks by priority users* templates and select the **User performs an exfiltration activity and associated indicators**, you can choose custom or default thresholds for the indicator triggering events that you select. Choose either the **Use default thresholds (Recommended)** or **Use custom thresholds for the triggering events**.

    Select **Next** to continue.

15. If you select **Use custom thresholds for the triggering events**, for each triggering event indicator that you selected in Step 10, choose the appropriate level to generate the desired level of activity alerts. You can use the recommended thresholds, custom thresholds, or thresholds based on anomalous activities (for certain indicators) over the daily norm for users.

    Select **Next** to continue.

16. On the **Policy indicators** page, you see the indicators that you define as available on the **Insider risk settings** > **Indicators** page which include indicator variants if you have defined any. Select the indicators you want to apply to the policy.

    > **ⓘ Important**
    >
    > If indicators on this page can't be selected, select the indicators you want to enable for all policies. You can use **Turn on indicators** in the workflow or select indicators on the **Insider Risk Management** > **Settings** > **Policy indicators** page.

    If you select at least one *Office* or *Device* indicator, select the **Risk score boosters** as appropriate. Risk score boosters only apply for selected indicators. If you select a *Data theft* or *Data leaks* policy template, select one or more **Sequence detection** methods and a **Cumulative exfiltration detection** method to apply to the policy. If you select the *Risky browser usage* policy template, select one or more of the **Browsing indicators**.

    Select **Next** to continue.

17. On the **Decide whether to use default or custom indicator thresholds** page, choose custom or default thresholds for the policy indicators that you select. Choose either the **Use default thresholds for all indicators** or **Specify custom thresholds** for the selected policy indicators. If you select **Specify custom thresholds**, choose the appropriate level to generate the desired level of activity alerts for each policy indicator.

    > **Tip**
    >
    > To view a graph that helps you determine appropriate threshold settings, select the **View impact** link in the insight for each set of threshold settings. **Learn more about manually customizing thresholds**.

    Select **Next** to continue.

18. On the **Review** page, review the settings you chose for the policy and any suggestions or warnings for your selections. Select **Edit** to change any of the policy values or select **Submit** to create and activate the policy.

Next steps

After you complete these steps to create your first Insider Risk Management policy, you start receiving alerts from activity indicators after about 24 hours. Configure additional policies as needed by using the guidance in Step 4 of this article or the steps in Create a new insider risk policy.

To learn more about investigating insider risk alerts and the **Alerts dashboard** or the **Triage Agent dashboard**, see Insider Risk Management activities.

Last updated on 11/14/2025

Assign permissions in Insider Risk Management

09/30/2025

Use six role groups to configure Insider Risk Management features. To make **Insider Risk Management** available as a menu option in Microsoft Purview and to continue with these configuration steps, you must be assigned to one of the following roles or role groups:

- Microsoft Entra ID *Global Administrator* role
- Microsoft Entra ID *Compliance Administrator* role
- Microsoft Purview *Organization Management* role group
- Microsoft Purview *Compliance Administrator* role group
- *Insider Risk Management* role group
- *Insider Risk Management Admins* role group

Depending on how you want to manage Insider Risk Management policies and alerts, assign users to specific role groups to manage different sets of Insider Risk Management features. You can assign users with different compliance responsibilities to specific role groups to manage different areas of Insider Risk Management features. Or you might decide to assign all user accounts for designated administrators, analysts, investigators, and viewers to the *Insider Risk Management* role group. Use a single role group or multiple role groups to best fit your compliance management requirements.

> **Important**
>
> After configuring your role groups, it might take up to 30 minutes for the role group permissions to apply to assigned users across your organization.

Choose from these role group options and solution actions when working with Insider Risk Management:

Expand table

Actions	Insider Risk Management	Insider Risk Management Admins	Insider Risk Management Analysts	Insider Risk Management Investigators	Insider Risk Management Auditors	Insider Risk Management Approvers
Access & investigate alerts	Yes	No	Yes	Yes	No	No

Actions	Insider Risk Management	Insider Risk Management Admins	Insider Risk Management Analysts	Insider Risk Management Investigators	Insider Risk Management Auditors	Insider Risk Management Approvers
Access & investigate cases	Yes	No	Yes	Yes	No	No
Access & view forensic evidence captures	Yes	No	No	Yes	No	No
Access & view the Content Explorer	Yes	No	No	Yes	No	No
Access analytics insights	Yes	Yes	Yes	No	No	No
Approve forensic evidence capturing requests	Yes	No	No	No	No	Yes
Configure Adaptive Protection	Yes	Yes	No	No	No	No
Configure notice templates	Yes	No	Yes	Yes	No	No
Configure policies and settings	Yes	Yes	No	No	No	No
Create forensic evidence capturing request	Yes	Yes	No	No	No	No
View & export audit logs	Yes	No	No	No	Yes	No

Actions	Insider Risk Management	Insider Risk Management Admins	Insider Risk Management Analysts	Insider Risk Management Investigators	Insider Risk Management Auditors	Insider Risk Management Approvers
View Adaptive Protection users tab	Yes	No	Yes	Yes	No	No
View alert and case reports	Yes	Yes	Yes	Yes	No	No
View data risk graphs for alerts	Yes	No	Yes	Yes	No	No

> **Important**
>
> Make sure you always have at least one user in the built-in *Insider Risk Management* or *Insider Risk Management Admins* role groups (depending on the option you choose) so that your Insider Risk Management configuration doesn't get into a 'zero administrator' scenario if specific users leave your organization.

Members of the following roles can assign users to Insider Risk Management role groups and have the same solution permissions included with the *Insider Risk Management Admins* role group:

- Microsoft Entra ID *Global Administrator*
- Microsoft Entra ID *Compliance Administrator*
- Microsoft Purview *Organization Management*
- Microsoft Purview *Compliance Administrator*

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

Add users to the Insider Risk Management role group

1. Sign in to the Microsoft Purview portal with an admin account in your Microsoft 365 organization.
2. Select **Settings** in the upper-right corner of the page, select **Roles and groups**, then select **Roles groups** in the left navigation pane.
3. Select the *Insider Risk Management* role group, then select **Edit**.
4. Select **Choose users**, then select the checkboxes for all the users you want to add to the role group.
5. Select **Select**, then select **Next**.
6. Select **Save** to add the users to the role group, then select **Done**.

Consider administrative units if you want to scope user permissions to a region or department

You can use administrative units in Insider Risk Management to scope user permissions to a particular geography or department. For example, a global company that has subsidiaries throughout the world might want to create an admin unit that provides a German scope for investigators so that they only see user activity for German users.

To use admin units in Insider Risk Management, you must first create the admin units if they aren't already created, then assign the admin units to members of role groups. After you assign admin units to members of role groups, those members become *restricted administrators* and have limited access to Insider Risk Management settings, policies, and user data in the organization. Members who aren't assigned administrative units are *unrestricted administrators* and have access to all settings, policies, and user data.

> **Important**
>
> Restricted administrators can't access alerts for the users assigned to them through security groups or distribution groups added in administrative units. Such user alerts are visible only to unrestricted administrators. Microsoft recommends adding users directly to administrative units to ensure their alerts are also visible to restricted administrators with administrative units assigned.

Effect of admin unit scoping on Insider Risk Management roles

The following table shows how admin units, when enforced, affect each combination of Insider Risk Management task and role.

> **Note**

Scoped in the following table means that the admin actions for that role are limited by their assigned admin unit.

Expand table

Task	Scoped Insider Risk Management	Scoped Insider Risk Management Admin	Scoped Insider Risk Management Analysts	Scoped Insider Risk Management Investigators	Scoped Insider Risk Management Approvers
Access analytics insights	Not allowed, if scoped	Not allowed, if scoped	Not allowed, if scoped	Never allowed	Never allowed
Access and investigate alerts	Scoped	Never allowed	Scoped	Scoped	Never allowed
Access and investigate cases	Scoped	Never allowed	Scoped	Scoped	Never allowed
Access and view forensic evidence captures	Not allowed, if scoped	Never allowed	Never allowed	Not allowed, if scoped	Never allowed
Access and view the Content explorer	Unrestricted	Never allowed	Never allowed	Unrestricted	Never allowed
Approve forensic evidence capturing requests	Not allowed, if scoped	Never allowed	Never allowed	Never allowed	Not allowed, if scoped
Assign or reassign alerts	Not allowed, if scoped	Never allowed	Not allowed, if scoped	Not allowed, if scoped	Never allowed
Assign or reassign cases	Not allowed, if scoped	Never allowed	Not allowed, if scoped	Not allowed, if scoped	Never allowed
Configure Adaptive Protection	Not allowed, if scoped	Not allowed, if scoped	Never allowed	Never allowed	Never allowed

Task	Scoped Insider Risk Management	Scoped Insider Risk Management Admin	Scoped Insider Risk Management Analysts	Scoped Insider Risk Management Investigators	Scoped Insider Risk Management Approvers
Configure global settings	Unrestricted	Unrestricted	Never allowed	Never allowed	Never allowed
Configure notice templates	Unrestricted	Never allowed	Unrestricted	Unrestricted	Never allowed
Configure policies	Scoped	Scoped	Never allowed	Never allowed	Never allowed
Configure priority user groups	Not allowed, if scoped	Not allowed, if scoped	Never allowed	Never allowed	Never allowed
Configure priority-user-specific policies	Not allowed, if scoped	Not allowed, if scoped	Never allowed	Never allowed	Never allowed
Create forensic evidence capturing request	Not allowed, if scoped	Not allowed, if scoped	Never allowed	Never allowed	Never allowed
Create quick policies	Not allowed, if scoped	Not allowed, if scoped	Never allowed	Never allowed	Never allowed
Investigate user activity	Scoped	Never allowed	Never allowed	Scoped	Never allowed
Start scoring activity for users	Scoped	Scoped	Scoped	Scoped	Never allowed
View Adaptive Protection users tab	Not allowed, if scoped	Never allowed	Not allowed, if scoped	Not allowed, if scoped	Never allowed
View alert and case reports	Scoped	Scoped	Scoped	Scoped	Scoped
View data risk graphs	Not supported	Not supported	Not supported	Not supported	Not supported

Task	Scoped Insider Risk Management	Scoped Insider Risk Management Admin	Scoped Insider Risk Management Analysts	Scoped Insider Risk Management Investigators	Scoped Insider Risk Management Approvers
for alerts					
View device health report	Not allowed, if scoped	Not allowed, if scoped	Never allowed	Never allowed	Never allowed

> **Note**
>
> You can use adaptive scopes together with admin units. If one or more admin units scope role groups for your organization, the admin units limit the adaptive scopes that you can select when you create or edit a policy.

Microsoft Purview Insider Risk Management and Communication Compliance privacy guide

08/20/2025

Microsoft Purview insider risk solutions give organizations the tools to help detect and reduce potential risks and policy violations. The Microsoft Purview insider risk solutions include:

- **Microsoft Purview Insider Risk Management** uses different signals to find potential malicious or accidental insider risks, such as intellectual property theft, data leakage, and security violations. Insider risk management lets customers create policies to manage security and compliance.
- **Microsoft Purview Communication Compliance** provides tools to help organizations detect potential regulatory compliance issues (for example SEC or FINRA) and business conduct violations such as sensitive or confidential information, harassing or threatening language, and sharing of adult content.

Insider risk management and communication compliance are built with privacy by design and balance user privacy with tools that help detect and reduce organizational risks. We're committed to protecting user trust and maintaining user-level privacy through our core privacy principles:

- Pseudonymization
- Role-based access controls
- Admin explicit opt-in
- Audit logs

Pseudonymization

Pseudonymization helps protect end-user privacy by removing identifiable user details like user name or email address. Pseudonymization also helps prevent potential bias and conflicts of interest by removing identifiable user details (name, email) and personal data (title, department, or location) exposed in the solution. For example, a user named John Smith is pseudonymized into a nonpersonal identifier such as ANON2340. Pseudonyms are on by default for specific roles such as *Insider Risk Management Analysts* and *Insider Risk Management Investigators* (review alerts and take action respectively) and *Communication Compliance Analysts* (review policy alerts).

Role-based access controls

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

We recommend using stringent role-based access controls so only authorized insider risk management and communication compliance roles can use and access alerts and insights into potential policy violations. By default, global administrators don't have access to insider risk management and communication compliance features. This default setting helps ensure that only the appropriate stakeholders can access the solution and details specific to their role permissions. Organizations can assign users to specific role groups to manage different sets of features based on their responsibilities. For example, insider risk management and communication compliance admins can create, configure, and delete policies but can't access or investigate alerts or cases. On the other hand, insider risk management and communication compliance investigators can access and investigate alerts and cases but can't configure policies.

> **Note**
>
> Insider risk management admins can allow investigators and analysts to make edits to policy indicators and thresholds by using the inline alert customization setting.

Whether your organization chooses a single role group or multiple role groups to fit your organization's compliance and privacy requirements, both insider risk management and communication compliance allow admins to choose from predefined role group options within each solution.

Learn more about role group options for each solution:

- Get started with insider risk management
- Get started with communication compliance

Admin explicit opt-in

Insider risk management and communication compliance policies detect risky activities and communications and potential policy violations that could result in a security incident. An admin with the right permissions can explicitly scope employees into a policy.

Additionally, by default, insider risk management and communication compliance indicators that help detect risky activities and communications that might lead to potential data security incidents are disabled. For example, indicators like "downloading content from OneDrive", "sharing SharePoint files with people outside the organization", or "sending sensitive information or harassing messages" are off by default. Insider risk management and communication compliance don't detect those activities without an admin's explicit opt-in. Admins with the right permissions must explicitly select and opt in to one or more indicators in settings before a policy can detect those activities.

Admin explicit opt-in controls help safeguard end-user privacy by ensuring the solutions only flag alerts and policy violations for users and indicators specified in the policies.

Audit logs

Microsoft Purview insider risk solutions audit logs record all admin actions. With these logs, organizations can stay informed about all actions taken within Microsoft Purview insider risk solutions, including when an admin creates or edits a policy, adds a user, views user activity insights, or adds indicators.

Audit logs are enabled by default for all Microsoft 365 organizations. With these logs, organizations can audit privileged admins' actions and meet compliance and privacy requirements.

For more information about audit logs capabilities for each solution, see:

- Review activities with the insider risk management audit log
- Use communication compliance reports and audits

Protect user trust and build a holistic insider risk program

User privacy and trust are essential for organizations to establish a holistic insider risk program. The right set of tools can help you address risks in a way that meets security needs. Learn how to build a holistic insider risk management program with five elements that help companies have stronger data protection while ensuring user trust.

Monitoring agents in Insider Risk Management

> **Important**
>
> Microsoft Purview Insider Risk Management correlates various signals to identify potential malicious or inadvertent insider risks, such as IP theft, data leakage, and security violations. Insider Risk Management enables customers to create policies to manage security and compliance. Built with privacy by design, users are pseudonymized by default, and role-based access controls and audit logs are in place to help ensure user-level privacy.

Microsoft Purview Insider Risk Management enables organizations to identify, investigate, and respond to potential insider risks across users and agents. As organizations increasingly adopt AI agents, Insider Risk Management provides built-in capabilities to help monitor agent activity and surface risky behaviors that could lead to data exposure, compliance violations, or security concerns.

Before you begin monitoring agents in your organization, you must complete the steps provided in Get started with Insider Risk Management. For information about subscriptions, licensing, and permissions, see Subscriptions and licensing and Permissions in Insider Risk Management.

Risky Agents policy

The Risky Agents (preview) policy provides out-of-the-box coverage for common risk scenarios associated with AI agents. The policy is designed to help surface behaviors that may indicate inappropriate or unsafe agent usage, including situations where agents interact with sensitive data, generate risky outputs, or share information outside organizational boundaries.

By default, the policy detects:

- Risky prompts
- Agents generating sensitive responses
- Agents accessing sensitive data
- Accessing risky websites

These detections help security, compliance, and risk teams gain visibility into how agents are being used and where potential risks may exist.

Supported agents

The Risky Agents (preview) policy supports the following agent types:

- Copilot Studio agents
- Microsoft Foundry agents
- Agents built using the P4AI SDK

Default policy availability and behavior

The Risky Agents (preview) policy is available by default to all organizations with supported licenses. When the service is set up, this policy is automatically present and ready to generate alerts based on observed agent activity.

This default configuration enables organizations to:

- Begin monitoring agent activity immediately.
- Gain early visibility into potential agent-related risks.
- Investigate agent behaviors using established Insider Risk Management workflows.

Customize agent risk monitoring

While the default Risky Agents (preview) policy provides broad coverage for common scenarios, some organizations may require more specific or tailored monitoring.

If the default policy doesn't meet your organizational requirements, you can create a custom Insider Risk Management policy to better align with your risk posture or investigation needs. For more information, see Create a new policy.

Investigate agent-related alerts

When the Risky Agents (preview) policy detects relevant activity, alerts are generated within Insider Risk Management. These alerts can be reviewed and investigated using the standard alert investigation workflow.

Through alert investigation, analysts can:

- Review details of the detected activity.
- Understand the context of agent behavior.
- Determine whether further action is required.

> **Note**
>
> When guest users are interacting with agents, the full prompt and response content isn't available for investigation. Instead, only indicators of sensitive information detected in those prompts and responses are shown.

> **Note**
>
> Admins units are currently not supported for agent alerts. If a scoped analyst is signing in, then they can't see agent related alerts.

For detailed guidance on reviewing and managing alerts, see Triaging alerts.

Share agent risk with other Microsoft solutions

Agent risk calculated in Microsoft Purview Insider Risk Management is shared with the following solutions:

- Data Security Posture Management for AI
- Microsoft 365 admin center

Related content

- Learn about Insider Risk Management
- Insider Risk Management policy templates
- Get started with Insider Risk Management
- Get started with the Microsoft Purview Triage Agent in Insider Risk Management

---

Last updated on 08/04/2026

Learn about Insider Risk Management settings

08/20/2025

> **Important**
>
> Microsoft Purview Insider Risk Management correlates various signals to identify potential malicious or inadvertent insider risks, such as IP theft, data leakage, and security violations. Insider Risk Management enables customers to create policies to manage security and compliance. Built with privacy by design, users are pseudonymized by default, and role-based access controls and audit logs are in place to help ensure user-level privacy.

Before you get started with Insider Risk Management policies, understand and choose the Insider Risk Management settings that best meet the compliance needs for your organization. Insider Risk Management settings apply to all Insider Risk Management policies, regardless of the template you choose when creating a policy.

> **Note**
>
> Use **Settings** at the top of any Insider Risk Management page to make settings changes.

The following table describes each Insider Risk Management setting and provides a link to learn more about the setting.

Expand table

Setting	Description
Admin notifications	Use the **Admin notifications** setting to automatically send an email notification to selectable Insider Risk Management role groups. You can: - Send a notification email when the first alert is generated for a new policy - Send a daily email when new high-severity alerts are generated - Send a weekly email summarizing policies that have unresolved warnings
Analytics	Insider risk analytics enables you to conduct an evaluation of potential insider risks in your organization without configuring any insider risk policies. This evaluation can help your organization identify potential areas of higher user risk and help determine the type and scope of Insider Risk Management policies you might consider configuring.
Data sharing	Use the **Data sharing** setting to do either of the following: 1) Export Insider Risk Management alert information to SIEM solutions by using the Office 365 Management

Setting	Description
	Activity API schema; 2) Share Insider Risk Management user risk levels with Microsoft Defender and DLP alerts.
Detection groups	Use the **Detection groups** setting to create variants of built-in indicators if you want to tailor detections for different sets of users. Creating detection groups helps to reduce false positives.
Export alerts	Insider Risk Management alert information is exportable to security information and event management (SIEM) and security orchestration automated response (SOAR) solutions by using the Office 365 Management Activity API schema. You can use the Office 365 Management Activity APIs to export alert information to other applications your organization might use to manage or aggregate insider risk information.
Global exclusions	Use the **Global exclusions** setting to specify global exclusions that won't be scored by your Insider Risk Management policies.
Inline alert customization	Inline alert customization allows you to quickly tune an Insider Risk Management policy directly from the **Alerts dashboard** while reviewing the alert. Alerts are generated when a risk management activity meets the thresholds configured in the related policy. To reduce the number of alerts you get from this type of activity, you can change the thresholds or remove the risk management activity from the policy altogether.
Intelligent detections	Use the **Intelligent detections** setting to boost the score for unusual download activity, control alert volume, import and filter Microsoft Defender for Endpoint alerts, and specify unallowed and third-party domains for risk scoring.
Microsoft Teams (preview)	You can enable Microsoft Teams support so that compliance analysts and investigators can use Teams to collaborate on Insider Risk Management cases. Use Teams to: - Coordinate and review response activities for cases in private Teams channels - Securely share and store files and evidence related to individual cases - Detect and review response activities by analysts and investigators
Policy indicators	Each Insider Risk Management policy template is based on specific indicators that correspond to specific triggers and risk activities. All global indicators are disabled by default; **you must select one or more indicators to configure an Insider Risk Management policy**. Indicator level settings help you control how the number of occurrences of risk events in your organization affect the risk score.
Policy timeframes	The **Policy timeframes** setting allows you to define past and future review periods that are triggered after policy matches based on events and activities for the Insider Risk Management policy templates.
Power Automate flows (preview)	Microsoft Power Automate is a workflow service that automates actions across applications and services. By using flows from templates or created manually, you can automate common tasks associated with these applications and services. When you enable Power Automate flows for Insider Risk Management, you can automate important tasks for cases and users. You can configure Power Automate flows to retrieve user, alert, and case information and share this information with stakeholders and other applications, as well as automate actions in Insider Risk Management, such

Setting	Description
	as posting to case notes. Power Automate flows are applicable for cases and any user in scope for a policy.
Priority physical assets (preview)	Identifying access to priority physical assets and correlating access activity to user events is an important component of your compliance infrastructure. These physical assets represent priority locations in your organization, such as company buildings, data centers, or server rooms. Insider risk activities might be associated with users working unusual hours, attempting to access these unauthorized sensitive or secure areas, and requests for access to high-level areas without legitimate needs.
Priority user groups	Users in your organization might have different levels of risk depending on their position, level of access to sensitive information, or risk history. Prioritizing the examination and scoring of the activities of these users can help alert you to potential risks that might have higher consequences for your organization. Use the **Priority user groups** setting to define the users in your organization that need closer inspection and more sensitive risk scoring.
Privacy	Choose whether to display usernames or anonymized versions of usernames for all current and past policy matches for alerts and cases.

Manage username privacy in Insider Risk Management

> **Important**
>
> Microsoft Purview Insider Risk Management correlates various signals to identify potential malicious or inadvertent insider risks, such as IP theft, data leakage, and security violations. Insider Risk Management enables customers to create policies to manage security and compliance. Built with privacy by design, users are pseudonymized by default, and role-based access controls and audit logs are in place to help ensure user-level privacy.

Protecting the privacy of users that have policy matches in Microsoft Purview Insider Risk Management is important. It helps promote objectivity in data investigation and analysis reviews for insider risk alerts. For users with an insider risk policy match, choose one of the following settings:

> **Important**
>
> With the anonymized versions of usernames setting enabled, you cannot use data risk graph.

- **Show anonymized versions of usernames**: The solution anonymizes the names of users to prevent admins, data investigators, and reviewers from seeing who is associated with policy alerts. For example, a user 'Grace Taylor' appears with a randomized pseudonym such as 'AnonIS8-988' in all areas of the Insider Risk Management experience. Choosing this setting anonymizes all users with current and past policy matches and applies to all policies.

  > **Important**
  >
  > If you're using the Triage Agent for Insider Risk Management, user names in prioritized alerts aren't anonymized in the Triage Agent dashboard if **Show anonymized versions of usernames** is enabled.

  User profile information in the insider risk alert and case details won't be available when this option is chosen. However, names of users that appear in the metadata of an activity, in the

metadata of files in Content explorer, and within content of files in Content explorer **aren't anonymized**. Usernames are displayed when adding new users to existing policies or when assigning users to new policies. If you choose to turn off this setting, usernames are displayed for all users that have current or past policy matches.

> **Important**
>
> To maintain referential integrity for users who have insider risk alerts or cases in Microsoft 365 or other systems, anonymization of usernames isn't preserved for exported alerts when using the exporting API or when exporting to Microsoft Purview eDiscovery solutions. Exported alerts display usernames for each alert in this case. If you export to CSV files from alerts or cases, anonymization *is* preserved.

- **Do not show anonymized versions of usernames**: Usernames are displayed for all current and past policy matches for alerts and cases. User profile information (the name, title, alias, and organization or department) is displayed for the user for all Insider Risk Management alerts and cases.

Change the privacy setting

Perform the following steps to change the privacy setting for displayed usernames:

1. Sign in to the Microsoft Purview portal with an admin account in your Microsoft 365 organization.
2. Select **Settings** in the upper-right corner of the page.
3. Select **Insider Risk Management** to go to the Insider Risk Management settings.
4. Select **Privacy**.
5. Select the option you want, to show anonymized names or not.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Configure policy indicators in Insider Risk Management

> **Important**
>
> Microsoft Purview Insider Risk Management correlates various signals to identify potential malicious or inadvertent insider risks, such as IP theft, data leakage, and security violations. Insider Risk Management enables customers to create policies to manage security and compliance. Built with privacy by design, users are pseudonymized by default, and role-based access controls and audit logs are in place to help ensure user-level privacy.

Insider risk policy templates in Microsoft Purview Insider Risk Management define the type of risk activities that you want to detect and investigate. Each policy template is based on specific indicators that correspond to specific triggers and risk activities. All global indicators are disabled by default; **you must select one or more indicators to configure an Insider Risk Management policy**.

Policies collect signals and trigger alerts when users perform activities related to the indicators.

Types of events and indicators

Insider Risk Management uses different types of events and indicators to collect signals and create alerts:

- **Triggering events**: Events that determine if a user is active in an Insider Risk Management policy. If you add a user to an Insider Risk Management policy that doesn't have a triggering event, the policy doesn't evaluate the user as a potential risk. For example, User A is added to a policy created from the *Data theft by departing users* policy template and the policy and Microsoft 365 HR connector are properly configured. Until User A has a termination date reported by the HR connector, the policy doesn't evaluate User A for potential risk. Another example of a triggering event is if a user has a *High* severity data loss prevention (DLP) policy alert when using *Data leaks* policies.
- **Global settings indicators**: Indicators enabled in global settings for Insider Risk Management define both the indicators available for configuration in policies and the types of events signals collected by Insider Risk Management. For example, if a user copies data

to personal cloud storage services or portable storage devices and you select these indicators only in global settings, you can review the user's potentially risky activity in the Activity explorer. If you don't define this user in an Insider Risk Management policy, the policy doesn't evaluate the user as a potential risk and therefore doesn't assign a risk score or generate an alert.

- **Policy indicators**: Indicators included in Insider Risk Management policies determine a risk score for an in-scope user. You enable policy indicators from indicators defined in global settings. The policy indicators activate only after a triggering event occurs for a user. Examples of policy indicators include:
  - A user copies data to personal cloud storage services or portable storage devices.
  - A user account is removed from Microsoft Entra ID.
  - A user shares internal files and folders with unauthorized external parties.

You can use certain policy indicators and sequences to customize triggering events for specific policy templates. When you configure these indicators or sequences in the policy workflow for the *General data leaks* or *Data leaks by priority users* templates, you get more flexibility and customization for your policies and when users are in-scope for a policy. You can also define risk management activity thresholds for these triggering indicators for more fine-grained control in a policy.

Define the insider risk policy indicators that are enabled in all insider risk policies

To enable policy indicators for all insider risk policies, complete the following steps:

1. Select **Settings**, then select **Policy indicators**.
2. Select one or more policy indicators. The indicators you select on the **Policy indicators** settings page can't be individually configured when creating or editing an insider risk policy in the policy workflow.

   > ⓘ **Note**
   >
   > It might take several hours for new manually added users to appear in the **Users dashboard**. Activities for the previous 90 days for these users might take up to 24 hours to display. To view activities for manually added users, select the user on the **Users dashboard** and open the **User activity** tab in the details pane.

Two types of policy indicators: built-in indicators and custom indicators

Indicators and pay-as-you-go billing

Some indicators included in Insider Risk Management require that you enable the pay-as-you-go billing model for your organization. Depending on your configured billing model, a notification might be displayed prompting you to configure pay-as-you-go billing to use these indicators.

Built-in indicators vs. custom indicators

Policy indicators are organized into two tabs:

- **Built-in indicators**: Insider Risk Management includes many built-in indicators for various scenarios that you can use right away in your policies. Choose the indicators that you want to activate, then customize indicator thresholds for each indicator level when you create an insider risk policy. This article describes the built-in indicators in more detail.
- **Custom indicators**: Use custom indicators together with the Insider Risk Indicators (preview) connector to bring non-Microsoft detections to Insider Risk Management. For example, you might want to extend your detections to include Salesforce and Dropbox and use them alongside the built-in detections provided by the Insider Risk Management solution, which is focused on Microsoft workloads (SharePoint Online and Exchange Online, for example). Create a custom indicator

Built-in indicators

Insider Risk Management includes the following built-in indicators.

Office indicators

These indicators include policy indicators for SharePoint sites, Microsoft Teams, and email messaging.

Cloud storage indicators

> ⓘ **Important**
>
> To use this indicator, enable pay-as-you-go billing in your organization.

These indicators include policy indicators for Google Drive, Box, and Dropbox that you can use to detect techniques used to determine the environment, gather and steal data, and disrupt the availability or compromise the integrity of a system. To select from cloud storage indicators, you must connect cloud storage apps in Microsoft Defender.

After configuring the cloud storage indicators, you can turn off indicators for the apps you don't want to use in settings. For example, you can select a content download indicator for Box and Google Drive, but not Dropbox.

Cloud service indicators

> **Important**
>
> To use this indicator, enable pay-as-you-go billing in your organization.

These indicators include policy indicators for Amazon S3 and Azure (SQL Server and Storage) that you can use to detect techniques used to avoid detection or risky activities. These techniques might include:

- Disabling trace logs
- Updating or deleting SQL Server firewall rules
- Techniques used to steal data, such as sensitive documents
- Techniques used to disrupt the availability or compromise the integrity of a system
- Techniques used to gain higher-level permissions to systems and data.

To select from cloud service indicators, you must connect source service apps in Microsoft Defender.

Network indicators

> **Important**
>
> To use this indicator, enable pay-as-you-go billing in your organization.

These indicators include HTTP and HTTPS network traffic from third party network security solutions. You can identify sensitive items that are being shared through these interactions. Detection of network traffic activities requires creation of a collection policy as a prerequisite. Learn more about network data security.

Microsoft Entra ID indicators

These indicators include risk detections from Microsoft Entra ID Protection. Risk detections are a powerful resource that can include any suspicious or anomalous activity related to a user account in the directory. Microsoft Entra ID Protection risk detections can be linked to an individual user or sign-in event.

User risk detections might flag a legitimate user account as at risk when a potential threat actor gains access to an account by compromising their credentials or when they detect some type of anomalous user activity. Sign-in risk detections represent the probability that a given authentication request isn't the authorized owner of the account.

To maintain relevance of the indicators to Insider Risk Management policies, only Microsoft Entra alerts in a *ConfirmedCompromised* or *Remediated* status state are evaluated. To learn more about the risk detections in Microsoft Entra ID Protection, see Risks detections in Microsoft Entra ID Protection.

Microsoft Fabric indicators

> **Important**
>
> To use the Microsoft Fabric indicators, enable pay-as-you-go billing in your organization.

These indicators include policy indicators for Microsoft Fabric workloads such as Power BI and Lakehouse. They help you detect techniques used to:

- Figure out the environment (for example, viewing Power BI reports and dashboards).
- Gather data of interest (for example, downloading Power BI reports).
- Obfuscate the data gathered or change protection (for example, downgrading or removing sensitivity labels of Power BI or Lakehouse assets).
- Exfiltrate the data (for example, sharing Lakehouse data with people outside the organization).

Generative AI apps indicators

These indicators include policy indicators for numerous generative AI applications. Use these indicators in policies to analyze interactions (prompts and responses) entered into these applications and help detect inappropriate or risky interactions or sharing of confidential information. These indicators include the following generative AI applications:

- **Microsoft Copilot experiences**: Support for user interactions in Copilot in Microsoft Fabric, Microsoft Security Copilot, Microsoft Copilot Studio, any connected or cloud AI application.

  > **ⓘ Important**
  >
  > To use the Microsoft Copilot experiences indicator for non-Microsoft 365 AI data, enable pay-as-you-go billing in your organization. Non-Microsoft 365 AI data includes information from other generative AI applications from Microsoft and other connected external AI applications. This data type includes Copilot in Microsoft Fabric, Microsoft Security Copilot, Microsoft Copilot Studio, and any connected or cloud AI application. There aren't any pay-as-you-go billing requirements or charges for Microsoft 365 detecting inappropriate or risky interaction for Microsoft 365 Copilot data.

- **Enterprise AI apps**: Non-Copilot AI applications connected using Microsoft Entra and Microsoft Purview Data Map connectors.

  > **ⓘ Important**
  >
  > To use the Enterprise AI apps indicator, enable pay-as-you-go billing in your organization.

- **Other AI applications**: AI applications that users in your organization discover from their browser activity.

  > **ⓘ Important**
  >
  > To use this indicator, enable pay-as-you-go billing in your organization.

- **Azure AI Content Safety indicators**: Support for Communication Compliance indicators to identify prompts and responses matching classifiers provided by Azure AI Content Safety like Prompt shields and Protected Materials.

  > **ⓘ Important**
  >
  > When you select the Azure AI Content Safety indicator, you create a Communication Compliance policy. If you modify this policy in Communication Compliance, you might need to pay for pay-as-you-go billing.

For Microsoft Copilot experiences and Enterprise AI apps, you can select which specific generative AI apps you want to monitor. By default, all available apps are selected. You can deselect apps that aren't relevant to your organization to reduce alert noise and avoid unnecessary pay-as-you-go billing charges. When an app is deselected, signals from that app aren't processed and no cost is incurred.

Select generative AI apps to monitor

> **Note**
>
> For **Microsoft Copilot experiences** and **Enterprise AI apps** only, you can review and deselect apps to optimize costs and reduce alert noise.

1. In the Microsoft Purview portal, go to **Insider Risk Management** > **Settings** > **Policy indicators**.
2. Expand **Generative AI apps indicators**.
3. For **Microsoft Copilot experiences** and **Enterprise AI apps**, select **Choose apps to monitor**.
4. From the side panel, you can select or clear the checkbox for each app to include or exclude it from monitoring.
5. Select **Save**.

Communication Compliance indicators

These indicators include policy indicators that detect employment stressor events, such as emotional outbursts, bullying, failure to take criticism, inability to work or communicate with a team or group, discrimination, violent threats, extremist behavior, and so on. Insider Risk Management works together with the Microsoft Purview Communication Compliance solution to detect these types of stressors that indicate an unhealthy workplace environment. Employment stressor events can impact user behavior for risky personas (whether initiators or targets of bad behavior) in several ways that relate to insider risks. Counterproductive work behavior can be a precursor to more serious violations, such as sabotaging company assets or leaking sensitive information.

Additionally, you can choose to detect messages matching specific sensitivity information types (SITs). Including sensitive information inadvertently or maliciously included in messages to user risk scores and their activity history provides investigators with more information to help quickly take actions to mitigate potential data leakage. You can select up to 30 SITs for a policy. Some scenarios might include helping to detect:

- Foreign recruitment
- State actor poaching
- Sharing sensitive information like secret formulas, financial reports, and other proprietary property
- Sharing passwords

> **Note**
>
> You can also use a Communication Compliance policy as a trigger.

How it works

You can choose from these Communication Compliance indicators:

- Sending inappropriate content
- Sending financial regulatory text that might be risky
- Sending inappropriate images

You can also choose to detect sensitive information types included in messages.

When you select **Create policy** from the **Communication Compliance indicators** section:

- A single policy is created in Communication Compliance that detects messages in Microsoft Exchange Online, Microsoft Teams, Microsoft Viva Engage, and Microsoft 365 Copilot and Microsoft 365 Copilot Chat. The Communication Compliance policy is based on the indicators and SITs you select. Each indicator is associated with specific trainable classifiers used by Communication Compliance. For more information, see Content safety classifiers based on large language models.

  > **Tip**
  >
  > Select the information icon next to each indicator to see the trainable classifiers that the indicator uses.

  In Communication Compliance, the trainable classifiers and SITs are listed as conditions for the policy.

- The Communication Compliance policy is named "Insider risk indicator" plus the timestamp, for example: "Insider risk indicator 24-05-01T09.27.17Z" or "Insider risk SIT indicator 24-05-01T09.27.17Z".

- Anyone with the *Insider Risk Investigators* role in Insider Risk Management is automatically added as a reviewer for the Communication Compliance policy.

  > ⓘ **Note**
  >
  > After creating the Communication Compliance policy, to add a reviewer to the policy, you must add the reviewer manually to the ***Communication Compliance Investigators* role group**.

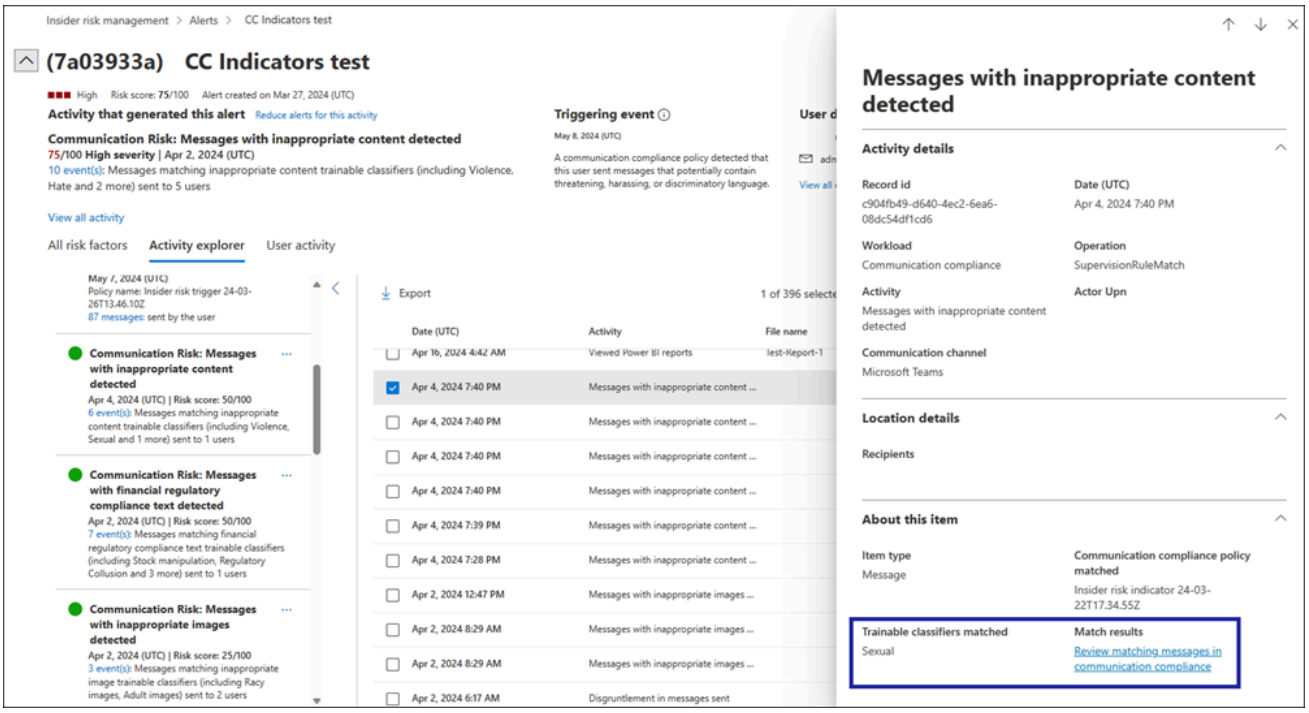
- If you turn off all of the indicators in the **Policy indicators** setting, you pause the Communication Compliance policy. The policy is reenabled if you turn any of the indicators back on.

- You make the Communication Compliance indicators available for new and existing policies in Insider Risk Management that are based on the *Data theft* or *Data leaks* templates.

- If content sent in a message matches any of the trainable classifiers, it results in a policy match in Communication Compliance that can be remediated. For details, see Investigate and remediate Communication Compliance alerts.

- Indicators included in Insider Risk Management policies determine a risk score for an in-scope user. They activate only after a triggering event occurs for a user.

- *Communication Risk* insights appear in the **Activity explorer** and **User activity** tabs in Insider Risk Management. If you drill down into a policy match from the **Activity explorer** or **User activity** tab, you can learn more about the activity and access a link that opens the Communication Compliance policy. In the Communication Compliance policy, you can see the content of the messages that were sent.



> ⓘ Note
>
> You must have the *Communication Compliance* role or the *Communication Compliance Investigators* role to access the Communication Compliance link.

Enable Communication Compliance indicators in Insider Risk Management

To enable Communication Compliance indicators in Insider Risk Management, complete the following steps:

1. In Insider Risk Management, go to **Settings** > **Policy indicators**, then scroll to the **Communication Compliance indicators (preview)** section.

2. Under **Detect messages matching specific trainable classifiers (preview)**, select **Create policy**.

   The policy you create is in Communication Compliance, and the Communication Compliance indicators become available in the **Policy indicators** setting.

> ⓘ Note
>
> If you already created a Communication Compliance policy but paused it, selecting **Create policy** resumes the Communication Compliance policy. In this case, the **Status** column in the Communication Compliance **Policies** list shows "Resuming".

3. Select one or more of the Communication Compliance indicators in the **Policy indicators** setting.

   > **Note**
   >
   > If you already created a Communication Compliance policy and you select different indicators, the Communication Compliance policy changes to reflect the appropriate trainable classifiers. Turning off all indicators pauses the Communication Compliance policy.

4. Select **Save**.

5. To use the indicators, create an insider risk management policy or update an existing policy. The indicators appear on the **Indicators** page of the policy workflow. You can adjust thresholds for the indicators as you would for any other indicators in an Insider Risk Management policy.

   > **Note**
   >
   > At this time, real-time analytics for indicator threshold settings aren't available for the Communication Compliance indicators.

Data loss prevention alerts indicators

These indicators include policy indicators that integrate with data loss prevention (DLP) policies. By selecting DLP policies as indicators in Insider Risk Management policies, you can automatically detect if a user has existing alerts in connected DLP policies. DLP policies help protect sensitive information and reduce the risks of oversharing data with inappropriate users or organizations.

When an Insider Risk Management alert is generated for a user, you can quickly determine if the user has any high risk alerts associated with DLP policies in your organization without having to navigate to DLP solution in the Microsoft Purview portal. You can review and evaluate the Insider Risk Management activity and associated DLP alerts within Insider Risk Management in a unified view.

Supported DLP workloads

The High Severity DLP Alert indicator currently supports alerts generated by DLP policies scoped to the following workloads:

Expand table

Supported workload	Description
Exchange Online	DLP policy matches on email messages
SharePoint Online	DLP policy matches on documents stored in SharePoint sites
OneDrive for Business	DLP policy matches on documents stored in OneDrive accounts

> **Important**
>
> The following DLP workloads are **not** currently supported by the High Severity DLP Alert indicator. Alerts generated by DLP policies that apply exclusively to these workloads won't trigger the indicator and won't appear in Insider Risk Management alerts, cases, or activity views:
>
> - **Endpoint DLP**: DLP policy matches on Windows and macOS devices
> - **Microsoft Teams**: DLP policy matches in Teams chat and channel messages
> - **Microsoft 365 Copilot**: DLP policy matches in Copilot interactions
> - **On-premises repositories**: DLP policy matches on on-premises file shares and SharePoint Server
> - **Power BI**: DLP policy matches in Power BI datasets
>
> This is by design. Only DLP alerts written to the Microsoft 365 audit log by the supported workloads listed above are evaluated by Insider Risk Management. If your DLP policy spans multiple workloads (for example, Exchange + Endpoint), only the alerts from the supported workloads (Exchange Online, SharePoint Online, and OneDrive for Business) are processed.

Configure DLP alerts indicators

Step 1: To enable the DLP alerts as indicators, complete the following steps:

1. In Insider Risk Management settings, select **Policy indicators** and then select the **Built-in Indicators** tab.
2. Navigate to **Data loss prevention (DLP) indicators**
3. Select **Add DLP policies**
4. Select the DLP policies that you want to see alerts for in Insider Risk Management.

5. Select **Add**.

6. Select the **Generating alerts from selected DLP policies** checkbox.

7. Select **Save**.

Step 2: To assign DLP alerts indicators to a specific Insider Risk Management policy, complete the following steps:

1. Create a custom policy using one of the following templates:

    - Data theft by departing users
    - Data leaks
    - Data leaks by priority users
    - Data leaks by risky users
    - Risky AI usage

2. Configure the policy as applicable until you reach the **Indicators** page.

3. On the **Indicators** page, navigate to **Data loss prevention (DLP) indicators**.

4. Select the **Generating alerts from selected DLP policies** checkbox.

5. Complete the policy configuration workflow and save the new policy.

Device indicators

These policy indicators include activities such as sharing files over the network or with devices. Indicators include activities involving all file types, excluding executable (.exe) and dynamic link library (.dll) file activity. If you select *Device indicators*, the system processes activity for devices with Windows 10 Build 1809 or higher and macOS (three latest released versions) devices. For both Windows and macOS devices, you must onboard devices. Device indicators also include browser signal detection to help your organization detect and act on exfiltration signals for nonexecutable files viewed, copied, shared, or printed in Microsoft Edge and Google Chrome. For more information on configuring Windows devices for integration with insider risk, see Enable device indicators and onboard Windows devices. For more information on configuring macOS devices for integration with insider risk, see Enable device indicators and onboard macOS devices. For more information about browser signal detection, see Learn about and configure Insider Risk Management browser signal detection.

> **Important**

Device indicators are included in collection policy evaluations. When you configure and deploy a collection policy, if there's a mismatch between an Insider Risk Management policy that includes device indicators and a collection policy in your organization, the collection policy configuration takes precedence. This configuration means that if you configure an Insider Risk Management policy to monitor a specific activity for devices, but you configure the collection policy to filter out that device activity, the device activity isn't collected and isn't available for review in Insider Risk Management.

Microsoft Defender for Endpoint indicators (preview)

These indicators come from Microsoft Defender for Endpoint and relate to unapproved or malicious software installation or bypassing security controls. To receive alerts in Insider Risk Management, you must have an active Defender for Endpoint license and insider risk integration enabled. For more information on configuring Defender for Endpoint for Insider Risk Management integration, see Configure advanced features in Microsoft Defender for Endpoint.

Health record access indicators

These policy indicators cover patient medical record access. For example, attempted access to patient medical records in your electronic medical records (EMR) system logs can be shared with Insider Risk Management healthcare policies. To receive these types of alerts in Insider Risk Management, you must have a healthcare-specific data connector and the HR data connector configured.

Physical access indicators

These policy indicators cover physical access to sensitive assets. For example, attempted access to a restricted area in your physical badging system logs can be shared with Insider Risk Management policies. To receive these types of alerts in Insider Risk Management, you must have priority physical assets enabled in Insider Risk Management and the Physical badging data connector configured. To learn more about configuring physical access, see Configure priority physical assets.

Microsoft Defender for Cloud Apps indicators

These policy indicators come from shared alerts from Defender for Cloud Apps. Automatically enabled anomaly detection in Defender for Cloud Apps immediately starts detecting and collating results, targeting numerous behavioral anomalies across your users and the machines and devices connected to your network. To include these activities in Insider Risk Management

policy alerts, select one or more indicators in this section. To learn more about Defender for Cloud Apps analytics and anomaly detection, see Get behavioral analytics and anomaly detection.

Risky Agents indicators (preview)

These policy indicators cover agent interactions with users and sensitive or risky resources. Risky agent prompts, agent generated sensitive responses, accessing sensitive or priority SharePoint files, agents accessing risky websites, or agents using tools with sensitive information are included.

Risky AI usage indicators (preview)

These policy indicators cover Microsoft AI tools and applications. Risky prompt behavior from users and AI-generated responses that include sensitive information are both included in these indicators. For example, attempted sharing of sensitive information in an AI tool or application by a user is considered risky activity. Similarly, an AI tool or application returning a response that contains sensitive information is also considered risky behavior.

Risky browsing indicators (preview)

These policy indicators cover browsing activity related to websites that are considered malicious or risky and pose potential insider risk that might lead to a security or compliance incident. Risky browsing activity refers to users who visit potentially risky websites, such as those associated with malware, pornography, violence, and other unallowed activities. To include these risk management activities in policy alerts, select one or more indicators in this section. To learn about configuring browser exfiltration signals, see Insider Risk Management browser signal detection.

Cumulative exfiltration detection indicators

These indicators detect when a user's exfiltration activities across all exfiltration channels over the last 30 days exceed organization or peer group norms. For example, if a user is in a sales role and communicates regularly with customers and partners outside of the organization, their external email activity is likely higher than the organization's average. However, the user's activity might not be unusual compared to the user's teammates, or others with similar job titles. A risk score is assigned if the user's cumulative exfiltration activity is unusual and exceeds organization or peer group norms.

> **Note**

Peer groups are defined based on organization hierarchy, access to shared SharePoint resources, and job titles in Microsoft Entra ID. If you enable cumulative exfiltration detection, your organization agrees to share Microsoft Entra data with the Microsoft Purview portal, including organization hierarchy and job titles. If your organization doesn't use Microsoft Entra ID to maintain this information, detection might be less accurate.

Risk score boosters

These indicators raise the risk score for activity for the following reasons:

- *Activity that is above the user's usual activity for that day*: Scores are boosted if the detected activity deviates from the user's typical behavior.
- *User had a previous case resolved as a policy violation*: Scores are boosted if the user had a previous case in Insider Risk Management that was resolved as a policy violation.
- *User is a member of a priority user group*: Scores are boosted if the user is a member of a priority user group.
- *User is detected as a potential high impact user*: When you enable this indicator, users are automatically flagged as potential high-impact users based on the following criteria:
  - The user interacts with more sensitive content compared to others in the organization.
  - The user's level in the organization's Microsoft Entra hierarchy.
  - The total number of users reporting to the user based on the Microsoft Entra hierarchy.
  - The user is a member of a Microsoft Entra built-in role with elevated permissions.

  > **Note**
  >
  > When you enable the potential high impact user risk score booster, you agree to share Microsoft Entra data with the Microsoft Purview portal. If your organization doesn't use sensitivity labels or has not configured organization hierarchy in Microsoft Entra ID, this detection might be less accurate. If a user is detected as both a member of a priority user group and also a potential high-impact user, their risk score is only boosted once.

In some cases, you might want to limit the insider risk policy indicators that apply to insider risk policies in your organization. You can turn off the policy indicators for specific areas by disabling them from all insider risk policies in global settings. You can only modify triggering events for policies created from the *Data leaks* or *Data leaks by priority users* templates. Policies created from all other templates don't have customizable triggering indicators or events.

Custom indicators

Use the **Custom Indicators** tab to create a custom indicator to use as a trigger or as a policy indicator in your policies.

> ⓘ **Note**
>
> To create a custom indicator to import third-party indicator data, you must first create an Insider Risk Indicators connector (preview).

1. In Insider Risk Management settings, select **Policy indicators** and then select the **Custom Indicators** tab.
2. Select **Add custom indicator**.
3. Enter an indicator name and a description (optional).
4. In the **Data connector** list, select the Insider Risk Indicator connector that you created previously.

   When you select a data connector:

   - The name of the source column that you select when you create the connector appears in the **Source column from mapping file** field. If you don't select a source column when you create the connector, **None** appears in this field and you don't need to make a selection.
   - In the **Values in source column** list, select the value that you want to assign to the custom indicator. These values relate to the source column that you specified when you created the connector. For example, if you create a single connector that includes data for two indicators (Salesforce and Dropbox), you see those values in the list.
5. If you want to use a column to set threshold values, in the **Data from mapping file** list, select the column that you want to use for the threshold setting; otherwise, select the **Use only as a triggering event without any thresholds** option.

   > ⓘ **Note**

Only fields that have a *Number* data type appear in the **Data from mapping file** list, since a *Number* data type is required to set a threshold value. The data type is specified when you set up the connector.

6. Select **Add indicator**. The indicator is added to the **Custom Indicators** list.

Now you can add the custom indicator to an insider risk policy in any *Data theft* or *Data leaks* policies that you create or edit.

- If you use the custom indicator as a trigger, select your custom trigger on the **Triggers** page when you create or edit the policy.
- If you use the custom indicator as a policy indicator, select your custom indicator on the **Indicators** page when you create or edit the policy.

> **Note**
>
> After selecting your custom trigger or indicator, make sure to set a custom threshold (don't use the default thresholds). You can't set trigger thresholds on a custom indicator if you select the **Use only as a triggering event without any thresholds** option.

After adding the custom indicator to your policies, the triggers and insights generated based on the custom indicators appear in the **Alerts dashboard**, **Activity explorer**, and **User timeline**.

> **Important**
>
> Wait 24 hours before running the sample script to upload your data after you update the custom indicators and the associated policies. It can take several hours to sync all components. If you immediately upload the data while the updates are syncing, some data might not be scored for risk.

Create a variant of a built-in indicator

You can create detection groups and use them with variants of built-in indicators to tailor detections for different sets of users. For example, to reduce the number of false positives for email activities, you might want to create a variant of the *Sending email with attachments to recipients outside the organization* built-in indicator to only detect email sent to personal domains. A variant inherits all the properties of the built-in indicator. You can modify the variant with exclusions or inclusions.

1. In Insider Risk Management settings, select **Policy indicators**.
2. Select **New indicator variant (preview)**. This step opens the **New indicator variant (preview)** pane on the right side of the screen.
3. In the **Base indicator** list, select the indicator that you want to create a variant for.

   > ⓘ **Note**
   >
   > You can create up to ten variants for each built-in indicator and a total of 100 variants across all indicators. If you already created ten variants for a particular built-in indicator, the built-in indicator appears grayed out in the list. Some built-in indicators (Microsoft Defender for Endpoint indicators, for example) don't support variants.

4. Add a name for the variant (or accept the suggested name). A variant name can't be more than 110 characters.
5. Add a description for the variant (optional). The description appears in the policy to help you differentiate it from other indicators or indicator variants. A description for a variant can't be more than 256 characters.
6. Under **Detection group**, select one of the following options:

   - **Ignore activity involving items in selected groups**. Select this option if you want to capture everything except for a few *exclusions*. For example, you might want to use this option to capture all outgoing email except for email sent to specific domains.
   - **Only detect activity involving items in selected groups**. Select this option if you want to specify *inclusions* to capture. For example, select this option if you want to capture only email sent to certain domains.

   > ⓘ **Note**
   >
   > If you didn't already create a detection group, you can't select an option in the **Detection group** section.

7. In the **Select one or more detection groups** list, select the detection groups that you want to apply to the variant. Detection groups are listed under the appropriate detection type heading to help you find the appropriate group. For a single variant, you can add up to five

detection groups of a single type. For example, you can add up to five groups of domains, five groups of file types, and so on.

> **Note**
>
> Only detection groups that are applicable to the variant appear in the list. For example, a file type detection group won't appear for the *Sharing SharePoint folders with people outside the organization* indicator since it's not applicable.

8. Select **Save**.

9. In the **Next steps** dialog box, if you want to apply the new variant to a specific policy, select the **Policies page** link.

> **Tip**
>
> To make sure that a variant captures all the important activities that you want to detect, you can apply the built-in indicator and the variant of the built-in indicator in the same policy. You can then observe the activities that each indicator captures in alerts and then use only the variant indicator after making sure everything is detected.

Use a variant in a policy

To use an indicator variant in a policy, complete the following steps:

1. Go to the **Indicators** page of the policy workflow.

2. Find the built-in indicator that includes one or more variants. A small blue box in the variant checkbox marks built-in indicators that have variants. A list appears at the end of the indicator descriptor text to show the number of selected variants. Open the list to see the variants.

   > **Note**
   >
   > If you select one or more checkboxes in the variant list, the first-level checkbox for the built-in indicator becomes a solid blue checkbox. If you don't select any of the boxes in the variant list, the first-level checkbox is blank.

3. Select **Next**.

4. In the **Customize thresholds** page, you can customize threshold values for variants individually.

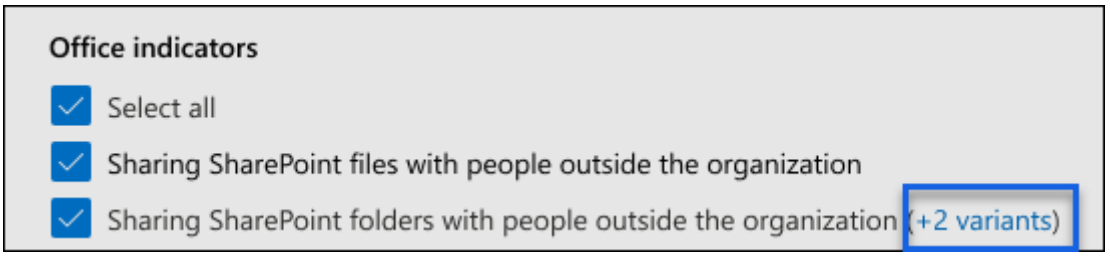
Investigate insights provided by variants

When you add variants to policies, the dashboard generates alerts. An investigator can view more details in the **Activity explorer** and **User activity** tabs.

Edit a variant

To edit an existing variant, complete the following steps:

1. Select the blue text at the end of the indicator description text. For example, select **+2** variants as shown in the following screenshot example.



2. In the **View/edit indicators** page, select **Edit**.

3. Make your changes.

Variant limitations

- You can create up to ten variants for each built-in indicator and a total of 100 variants across all indicators.
- You can add up to five detection groups of a single type for a single variant. For example, you can add a maximum of five groups of domains, five groups of file types, and so on.
- Variants don't support sequences, cumulative exfiltration activities, the risk score booster, or real-time analytics for the detections group preview.

How variants are prioritized against global exclusions and priority content

Insider Risk Management scopes activities in the following priority order:

1. Global exclusions
2. Variant scoping exclusion/inclusion

3. Priority content

Enable device indicators and onboard Windows devices

To enable the detection of risk activities on Windows devices and include policy indicators for these activities, your Windows devices must meet the following requirements, and you must complete the following onboarding steps. Learn more about device onboarding requirements

Step 1: Prepare your endpoints

Make sure that the Windows 10 devices you plan to report in Insider Risk Management meet these requirements.

1. The device must run Windows 10 x64 build 1809 or later and have the Windows 10 update (OS Build 17763.1075) from February 20, 2020 installed.
2. The user account used to sign in to the Windows 10 device must be an active Microsoft Entra account. The Windows 10 device might be Microsoft Entra ID, Microsoft Entra hybrid, joined, or registered.
3. Install the Microsoft Edge browser on the endpoint device to detect actions for the cloud upload activity. See Download the new Microsoft Edge based on Chromium.

> **Note**
>
> Endpoint DLP now supports virtualized environments, which means that the Insider Risk Management solution supports virtualized environments through endpoint DLP. **Learn more about support for virtualized environments in endpoint DLP**

Step 2: Onboard devices

To detect Insider Risk Management activities on a device, you must enable device checking and onboard your endpoints. You perform both actions in Microsoft Purview.

To enable devices that you didn't onboard yet, download the appropriate script and deploy it as outlined in this article.

If you already onboarded devices into Microsoft Defender for Endpoint, they appear in the managed devices list.

Onboard devices

Use this deployment scenario to enable devices that aren't onboarded yet when you want to detect insider risk activities on Windows devices.

1. Sign in to the Microsoft Purview portal with an admin account in your Microsoft 365 organization.
2. Select **Settings** in the upper-right corner of the page.
3. Under **Device onboarding**, select **Devices**. The list is empty until you onboard devices.
4. Select **Turn on device onboarding**.

   > **Note**
   >
   > While it usually takes about 60 seconds to enable device onboarding, allow up to 30 minutes before engaging with Microsoft Support.

5. Select how you want to deploy to these devices from the **Deployment method** list, then select **Download package**.
6. Follow the appropriate procedures in Onboarding tools and methods for Windows machines. This link takes you to a landing page where you can access Microsoft Defender for Endpoint procedures that match the deployment package you selected in step 5:

   - Onboard Windows machines using Group Policy
   - Onboard Windows machines using Microsoft Endpoint Configuration Manager
   - Onboard Windows machines using Mobile Device Management tools
   - Onboard Windows machines using a local script
   - Onboard non-persistent virtual desktop infrastructure (VDI) machines

When you finish and onboard the endpoint device, it appears in the devices list. The endpoint device starts reporting audit activity logs to Insider Risk Management.

> **Note**
>
> This experience is under license enforcement. Without the required license, data isn't visible or accessible.

If devices are already onboarded to Microsoft Defender for Endpoint

If you already deployed Microsoft Defender for Endpoint and endpoint devices are reporting in, the endpoint devices appear in the managed devices list. To expand coverage, onboard new devices into Insider Risk Management by going to Step 2: Onboarding devices.

Enable device indicators and onboard macOS devices

You can onboard macOS devices (Catalina 10.15 or later) into Microsoft 365 to support Insider Risk Management policies by using either Intune or JAMF Pro. For more information and configuration guidance, see Onboard macOS devices into Microsoft 365 overview (preview).

Indicator level settings

When you create a policy by using the policy workflow, you can configure how the daily number of risk events influences the risk score for insider risk alerts. These indicator settings help you control how the number of occurrences of risk events in your organization affect the risk score and the associated alert severity for these events.

For example, suppose you decide to enable SharePoint indicators in the insider risk policy settings and select custom thresholds for SharePoint events when configuring indicators for a new insider risk *Data leaks* policy. In the insider risk policy workflow, you configure three different daily event levels for each SharePoint indicator to influence the risk score for alerts associated with these events.

For the first daily event level, set the threshold to:

- *10 or more events per day* for a lower impact to the risk score for the events
- *20 or more events per day* for a medium impact to the risk score for the events
- *30 or more events per day* for a higher impact to the risk score for the events

These settings mean:

- If there are 1-9 SharePoint events that take place after the triggering event, risk scores are minimally impacted and tend not to generate an alert.
- If there are 10-19 SharePoint events that take place after a triggering event, the risk score is inherently lower and alert severity levels tend to be at a low level.
- If there are 20-29 SharePoint events that take place after a triggering event, the risk score is inherently higher and alert severity levels tend to be at a medium level.
- If there are 30 or more SharePoint events that take place after a triggering event, the risk score is inherently higher and alert severity levels tend to be at a high level.

Another option for policy thresholds is to assign the policy triggering event to risk management activity that is over the typical daily number of users. Instead of being defined by specific threshold settings, each threshold is dynamically customized for anomalous activities detected for in-scope policy users. If threshold activity for anomalous activities is supported for an individual indicator, you can select **Activity is above user's usual activity for the day** in the policy workflow for that indicator. If this option isn't listed, anomalous activity triggering isn't available for the indicator. If the **Activity is above user's usual activity for the day** option is listed for an indicator, but isn't selectable, you need to enable this option in **Insider risk settings** > **Policy indicators**.

Use real-time analytics recommendations to set thresholds

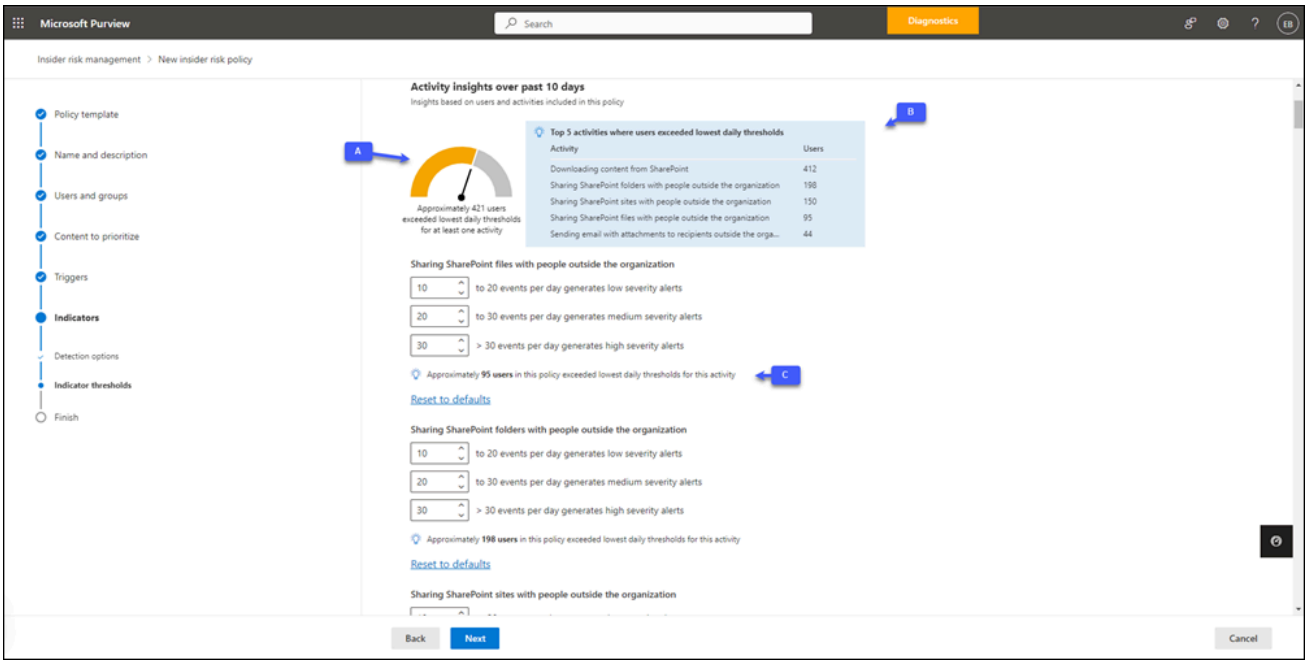
Use real-time analytics (preview) to get a guided, data-driven threshold configuration experience. With this experience, you can quickly select the right thresholds for policy indicators. The guided experience helps you efficiently adjust the selection of indicators and thresholds for activity occurrences so you don't have too few or too many policy alerts.

When you turn on analytics:

- The **Apply thresholds specific to your users' activity** option is enabled on the **Indicators** page of the policy workflow. Select this option if you want Insider Risk Management to provide indicator threshold recommendations based on the previous 10 days of user activity in your organization.

  > ⓘ **Note**
  >
  > To use this option, you must select at least one built-in policy indicator. Insider Risk Management doesn't provide recommended thresholds for custom indicators or variants of built-in indicators.

- If you select the **Choose your own thresholds** option on the **Indicators** page of the policy workflow, the defaults for the threshold settings are based on recommended threshold values (based on activity in your organization) instead of the built-in default values. You also see a gauge, a list of the top five indicators, and insights for each indicator.



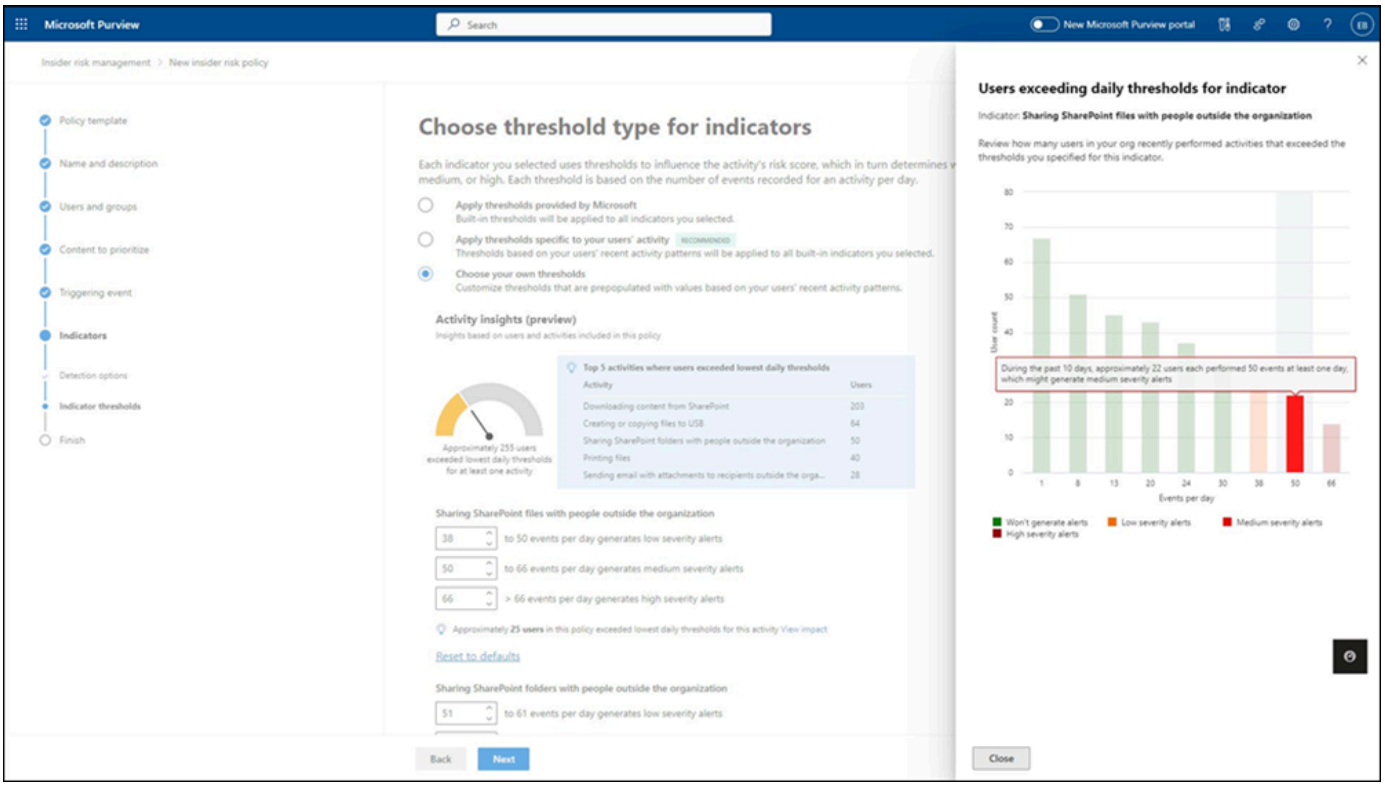
- **A.** The gauge shows the approximate number of scoped users whose activities from the previous 10 days exceeded the *lowest daily thresholds* for at least one of the selected built-in indicators for the policy. This gauge helps you estimate the number of alerts that might be generated if all users included in the policy were assigned risk scores.
- **B.** The top five indicators list is sorted by the number of users exceeding the *lowest daily thresholds*. If your policies generate too many alerts, focus on these indicators to reduce "noise."
- **C.** Insights for each indicator are displayed for the set of threshold settings for that indicator. The insight shows the approximate number of users whose activities from the previous 10 days exceeded the specified *low thresholds* for the indicator. For example, if the low threshold setting for *Downloading content from SharePoint* is set to 100, the insight shows the number of users in the policy who performed more than 100 download activities on average in the previous 10 days.

> ⓘ Note
>
> Global exclusions (intelligent detections) are taken into account for real-time analytics.

Adjust threshold settings manually

If you select the **Choose your own thresholds** option and manually adjust a threshold setting for a specific indicator, the insight for the indicator updates in real time. This feature helps you configure the appropriate thresholds for each indicator to achieve the highest level of alert effectiveness before activating your policies.

To save time and make it easier to understand the impact of manual changes to threshold values, select the **View impact** link in the insight to display the **Users exceeding daily thresholds for indicator** graph. This graph provides sensitivity analysis for each policy indicator.



> **Important**
>
> This graph isn't available if you select the **Include specific users** option when you create the policy. You must select the **Include all users and groups** option.

Use this graph to analyze the activity patterns of users in your organization for the selected indicator. For example, in the preceding illustration, the threshold indicator for *Sharing SharePoint files with people outside the organization* is set to **38**. The graph shows how many users performed actions that exceeded that threshold value and the distribution of low, medium, and high severity alerts for those actions. Select a bar to see insights for each value. For example, in the previous illustration, the bar for the **50** value is selected and the insight shows that at that threshold value, approximately 22 users each performed at least 50 events on at least one day in the past 10 days.

Prerequisites for using real-time analytics

To use real-time analytics (preview), you must enable insider risk analytics insights. After you enable analytics, it can take 24 to 48 hours for insights and recommendations to appear.

Note: The author created this article with assistance from AI. Learn more

Last updated on 06/22/2026

Fine-tune exclusions in Insider Risk Management by creating detection groups (preview)

Overview

> **Important**
>
> Microsoft Purview Insider Risk Management correlates various signals to identify potential malicious or inadvertent insider risks, such as IP theft, data leakage, and security violations. Insider Risk Management enables customers to create policies to manage security and compliance. Built with privacy by design, users are pseudonymized by default, and role-based access controls and audit logs are in place to help ensure user-level privacy.

You can exclude items from being scored by Insider Risk Management policies by using the Global exclusions setting. These types of exclusions apply to every trigger and indicator for all policies that you create within a tenant. By using detection groups, you can modify a built-in global exclusion to make it specific for your organization's needs.

You can also use detection groups to modify the built-in insider risk indicators to tailor detections for different sets of users at the policy level. For example, to reduce the number of false positives for email activities, you might want to create a variant of the *Sending email with attachments to recipients outside the organization* built-in indicator to only detect email sent to personal domains.

Process for creating and using detection groups

To use a detection group as part of a global exclusion, create the detection group as described in this article, then select it in the Global exclusions settings.

Using a detection group to modify a built-in indicator includes the following steps:

1. Create the detection group as described in the Create a detection group section of this article. You select this detection group when you create the variant.
2. Create the variant.

3. Use the variant in a new or existing policy.
4. Review alerts related to the activities specified in the variant.

Create a detection group

A detection group helps you scope a built-in indicator or a global exclusion to focus on the high-value activities important to your organization.

Detection types for policy indicators

Every policy indicator includes certain detection types that are applicable to that indicator. For example, for the *Sharing SharePoint files with people outside the organization* indicator, one of the detection types is *domains*. When you share a SharePoint file, you choose a domain to share the file with. Not all indicators have the same detection types. For example, *domains* is a detection type of the *Sharing SharePoint files with people outside the organization* indicator, but it's not a detection type of the *Creating or copying files to USB* indicator since it's not applicable in that case.

Insider Risk Management currently supports seven detection types:

- Domains
- File paths
- File types
- Keywords
- Sensitive info types
- SharePoint sites
- Trainable classifiers

> **Tip**
>
> When you create a detection group, you can see all the applicable indicators for a particular detection by selecting the **View applicable indicators** link in the introductory text for the group type.

Use the procedures in this section to create a detection group for each supported group type: domains, file paths, file types, keywords, sensitive info types, SharePoint sites, and trainable classifiers.

Create a domains detection group

To create a domains detection group, complete these steps:

1. In Insider Risk Management settings, select **Detection groups (preview)**.
2. In the panel to the right, under **Type**, select **Domains**.
3. In the panel to the right, select **New domain group**.
4. In the **New domain group** pane, enter a name for the group (or accept the suggested name) and an optional description.
5. In the **Add domains** field, enter a domain, then press Enter. Continue adding more domains in the same way. To add a long list of domains, import them as a CSV file by selecting **Import domains from CSV file**. The domains you add appear at the bottom of the pane. You can create up to 10 domain detection groups, and each group can have up to 500 items.

   > **Note**
   >
   > To specify multi-level subdomains for a root domain, select the **Include multi-level subdomains** checkbox, add a domain, then press Enter to add the domain to the list. Any subdomains included within that domain are included. Repeat the same process to add more domains, then select **Add domains** when you're done.

   > **Tip**
   >
   > Use wildcards to help match variations of root domains or subdomains. For example, to specify `sales.wingtiptoys.com` and `support.wingtiptoys.com`, use the wildcard entry `*.wingtiptoys.com` to match these subdomains (and any other subdomain at the same level).

6. Select **Save**. You see a **Next steps** dialog box that advises you on the next step in the process, which includes using this detection group in a variant.

How the *Free public domains* domain group detects exfiltration of business data to personal email domains

The Domains detection group includes a special domain group called *Free public domains* that consists of a list of free email providers (such as `gmail.com` or `yahoo.com`) that users can use to create a personal email ID. This list automatically highlights exfiltration of business data to

personal email domains for the email insight on the **User activity** and **Activity explorer** tabs. The insight lists the number of emails sent to free public domains by an in-scope user. The list also supports the algorithm that identifies email sent to self (sent to the in-scope user's personal email account). The email insight lists the number of emails sent to self.



You can use this built-in domain group like any other domain group to create a variant of a built-in indicator for your policies. This domain group only applies to the *Sending email with attachments to recipients outside the organization* indicator. For more information, see Create a variant of a built-in indicator.

Create a file paths detection group

To create a file paths detection group, complete these steps:

1. In Insider Risk Management settings, select **Detection groups (preview)**.
2. In the panel to the right, under **Type**, select **File paths**.
3. In the panel to the right, select **New file path group**.
4. In the **New file path group** pane, add a name for the group or accept the suggested name, and add a description (optional).
5. Select **Add file paths**, select the file paths, then select **Add**. You can create up to 10 file path detection groups and each group can have up to 500 items.
6. Select **Save**. You see a **Next steps** dialog box that advises you on the next step in the process, which includes using this detection group in a variant.

Create a file types detection group

To create a file types detection group, complete these steps:

1. In Insider Risk Management settings, select **Detection groups (preview)**.
2. In the panel to the right, under **Type**, select **File types**.
3. In the panel to the right, select **New file type group**.
4. In the **New file type group** pane, add a name for the group or accept the suggested name, and add a description (optional).
5. In the **Add file type** field, enter a file extension, then press Enter. Continue adding more file extensions the same way. The extensions that you add are listed at the bottom of the pane. You can create up to 10 file types detection groups and each group can have up to 500 items.
6. Select **Save**. You see a **Next steps** dialog box that advises you on the next step in the process, which includes using this detection group in a variant.

Create a keywords detection group

To create a keywords detection group, complete these steps:

1. In Insider Risk Management settings, select **Detection groups (preview)**.
2. In the panel to the right, under **Type**, select **Keywords**.
3. In the panel to the right, select **New keywords group**.
4. In the **New keywords group** pane, add a name for the group or accept the suggested name, and add a description (optional).
5. In the **Add keywords** field, enter a keyword, then press Enter. Repeat this process for each keyword you want to add. The keywords that you add are listed at the bottom of the pane. You can create up to 10 keywords detection groups and each group can have up to 500 items.
6. Select **Save**. You see a **Next steps** dialog box that advises you on the next step in the process, which includes using this detection group in a variant.

Create a sensitive info types detection group

> ⓘ **Note**
>
> The exclusion list of sensitive info types takes precedence over the priority content list.

1. In Insider Risk Management settings, select **Detection groups (preview)**.
2. In the panel to the right, under **Type**, select **Sensitive info types**.

3. In the panel to the right, select **New sensitive info type group**.
4. In the **New sensitive info type group** pane, enter a name for the group or accept the suggested name. Optionally, add a description.
5. Select **Add sensitive info types**, select the sensitive info types, then select **Add**. Default SIT exclusions are not available for selection. You can create up to 10 sensitive info type groups, and each group can have up to 500 items.
6. Select **Save**. You see a **Next steps** dialog box that advises you on the next step in the process, which includes using this detection group in a variant.

Create a SharePoint sites detection group

To create a SharePoint sites detection group, complete these steps:

1. In Insider Risk Management settings, select **Detection groups (preview)**.
2. In the panel to the right, under **Type**, select **SharePoint sites**.
3. In the panel to the right, select **New SharePoint site group**.
4. In the **New SharePoint site group** pane, enter a name for the group or accept the suggested name. Optionally, add a description.
5. Select **Add sites**, select the SharePoint sites, then select **Add**. You can create up to 10 SharePoint sites detection groups, and each group can have up to 500 items.
6. Select **Save**. You see a **Next steps** dialog box that advises you on the next step in the process, which includes using this detection group in a variant.

Create a trainable classifier detection group

To create a trainable classifier detection group, complete these steps:

1. In Insider Risk Management settings, select **Detection groups (preview)**.
2. In the panel to the right, under **Type**, select **Trainable classifiers**.
3. In the panel to the right, select **New trainable classifiers group**.
4. In the **New trainable classifiers group** pane, enter a name for the group or accept the suggested name. Optionally, add a description.
5. Select **Add trainable classifiers**, select the trainable classifiers, then select **Add**. You can create up to 10 trainable classifier detection groups, and each group can have up to 500 items.
6. Select **Save**. You see a **Next steps** dialog box that advises you on the next step in the process, which includes using this detection group in a variant.

See also

- Create a variant of a built-in indicator.
- Use a variant in a new or existing policy.
- Investigate alerts related to the activities specified in a variant.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Set up global exclusions for Insider Risk Management policies

> **Important**
>
> Microsoft Purview Insider Risk Management correlates various signals to identify potential malicious or inadvertent insider risks, such as IP theft, data leakage, and security violations. Insider Risk Management enables customers to create policies to manage security and compliance. Built with privacy by design, users are pseudonymized by default, and role-based access controls and audit logs are in place to help ensure user-level privacy.

Use the **Global exclusions** setting in Microsoft Purview Insider Risk Management to exclude items from being scored by Insider Risk Management policies. For example, to reduce "noise" in your policies, you might want to exclude certain file types or domains from being scored for risk if those file types or domains don't present risk to your organization.

When you set up global exclusions, consider using detection groups to tailor detections for different sets of users. Detection groups can help you reduce false positives for your policies.

Configure a domain exclusion

You can exclude specific domains related to user activities from being scored by your Insider Risk Management polices. These activities include:

- Email sent to external domains.
- Files, folders, and sites shared with external domains.
- Files uploaded to external domains (using the Microsoft Edge browser).

Allowed domains are ignored by your policies and don't generate alerts.

Set up a domain exclusion

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.

2. Select **Settings** in the upper-right corner of the page, then select **Insider Risk Management** to go to the Insider Risk Management settings.

3. Under **Insider risk settings**, select **Global exclusions**.

4. In the **Type** panel to the right, select **Domains**.

5. In the **Domains** panel to the right, select one of the following tabs:

   - **Individual domains**. To add domains one at a time:

     a. Select **Add domains**, then enter an exact domain in the **Domain** field or use a wildcard asterisk (*) to detect a subdomain (for example, *.contoso.com). Preceding the domain name with a wildcard detects a maximum of one subdomain (for example, support.contoso.com).

     b. To include all subdomains within the domain, select the **Include multi-level subdomains** checkbox.

     > **Note**
     >
     > You can use wildcards to help match variations of root domains or subdomains. For example, to specify sales.wingtiptoys.com and support.wingtiptoys.com, use the wildcard entry "*.wingtiptoys.com" to match these subdomains (and any other subdomain at the same level). To specify multi-level subdomains for a root domain, you must select the **Include multi-level subdomains** checkbox.

     c. Press Enter.

     d. Repeat this process for each domain that you want to add.

     Each domain that you enter is added to the **Domain** column and **Yes** or **No** is added to the **Multi-level subdomains included** list.

     > **Tip**
     >
     > If you don't want to add domains one at a time, you can import a list of domains from a CSV file by selecting **Import domains from CSV file** on the previous page.

     e. Select **Add domains**.

   - **Domain groups**. To select a domain detection group that you already created:

a. Select **Add domain group**.

b. Select the appropriate domain groups from the list. The number of domains included in each domain group is listed in the **Included domains** column.

c. Select **Save**.

Configure an email signature attachments exclusion

One of the main sources of noise in Insider Risk Management policies is images in email signatures. The system often detects these images as attachments in emails. If you select the *Sending email with attachments to recipients outside the organization* indicator, the system scores the attachment like any other email attachment sent outside the organization, even if the only thing in the attachment is the email signature. You can use the **Ignore email signature attachments (preview)** setting to exclude these attachments from scoring.

Turning on this setting significantly reduces noise from email signature attachments, but it doesn't completely eliminate all noise. This limitation exists because the system excludes from scoring only the email signature attachment of the email sender (the person who initiates the email or replies to the email). The system still scores a signature attachment for anyone on the To, CC, or BCC line. Also, if someone changes their email signature, the system needs to profile the new signature, which can cause alert noise for a short period of time.

Configure an email signature attachments exclusion

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.
2. Select **Settings** in the upper-right corner of the page, then select **Insider Risk Management** to go to the Insider Risk Management settings.
3. Under **Insider risk settings**, select **Global exclusions**.
4. In the **Type** panel to the right, select **Email signature attachments**.
5. In the **Ignore email signature attachments (preview)** panel to the right, turn the setting **On**.
6. Select **Save**.

Configure a file path exclusion

When you exclude file paths, user activities that map to specific indicators and that occur in those file path locations don't generate policy alerts. Examples include copying or moving files to a system folder or network share path. You can enter up to 500 file paths for exclusion.

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.

2. Select **Settings** in the upper-right corner of the page, then select **Insider Risk Management** to go to the Insider Risk Management settings.

3. Under **Insider risk settings**, select **Global exclusions**.

4. In the **Type** panel to the right, select **File paths**.

5. In the **File paths** panel to the right, select one of the following tabs:

   - **Individual file paths**. To add an individual file path, select **Add file paths to exclude**, enter an exact network share or device file path, then select **Add file paths to exclude**. Repeat this process for each file path that you want to exclude. Examples:

     Expand table

Example	Description
\ms.temp\LocalFolder\ or C:\temp	Excludes files directly under the folder and all subfolders for every file path starting with the entered prefix.
\public\local	Excludes files from every file path containing the entered value. Matches with 'C:\Users\Public\local\', 'C:\Users\User1\Public\local', and '\ms.temp\Public\local'.
C:\Users*\Desktop	Matches with 'C:\Users\user1\Desktop' and 'C:\Users\user2\Desktop'.
C:\Users*(2)\Desktop	Wilcards with numbers are supported. Matches with 'C:\Users\user1\Desktop' and 'C:\Users\user2\Shared\Desktop'.

     Each file path that you enter is added to the **File path column** on the page.

     > **Note**
     >
     > See Default file paths for file paths that are automatically excluded from generating policy alerts.

   - **File path groups**. To select a file path detection group that you have already created, select **Add file path group**, then select the appropriate file path groups. The number of file paths included in each group is listed in the **Included file paths** column.

6. Select **Save**.

Default file paths

By default, the system automatically excludes several file paths from generating policy alerts. Activities in these file paths are typically benign and could potentially increase the volume of non-actionable alerts. If needed, you can cancel the selection for these default file path exclusions to enable risk scoring for activities in these locations.

The default file path exclusions are:

- \Users*\AppData
- \Users*\AppData\Local
- \Users*\AppData\Roaming
- \Users*\AppData\Local\Temp

The wildcards in these paths denote that all folder levels between the \Users and \AppData are included in the exclusion. For example, activities in *C:\Users\Test1\AppData\Local* and *C:\Users\Test2\AppData\Local*, *C:\Users\Test3\AppData\Local* (and so on) are all included and not scored for risk as part of the *\Users*\AppData\Local* exclusion selection.

Configure a file type exclusion

You can exclude specific file types from all Insider Risk Management policy matching. For example, you might want to exclude all .wav files. Files with that extension are ignored for risk scoring by all Insider Risk Management policies.

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.

2. Select **Settings** in the upper-right corner of the page, then select **Insider Risk Management** to go to the Insider Risk Management settings.

3. Under **Insider risk settings**, select **Global exclusions**.

4. In the **Type** panel to the right, select **File types**.

5. In the **File types** panel to the right, select one of the following tabs:

    - **Individual file types**. To add file types one at a time, select **Add file type**, enter a file type, and then press Enter. Repeat this process for each file type that you want to add. Each file type that you enter is added to the page.

- **File type groups**. To select a file type detection group that you have already created, select **Add File type group**, then select the appropriate file type groups from the list.

6. Select **Save**.

Configure a keyword exclusion

You can configure exclusions for keywords that appear in file names, file paths, or email message subject lines. This configuration provides flexibility for organizations that need to reduce potential alert frequency due to flagging of benign terms. Your Insider Risk Management policies ignore activities related to files or email subjects containing the keyword and don't generate alerts. You can exclude up to 500 keywords.

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.

2. Select **Settings** in the upper-right corner of the page, then select **Insider Risk Management** to go to the Insider Risk Management settings.

3. Under **Insider risk settings**, select **Global exclusions**.

4. In the **Type** panel to the right, select **Keywords**.

5. In the **Keywords** panel to the right, select one of the following tabs:

   - **Individual keywords**. To add keywords one at a time, enter a keyword in the **Exclude these keywords** field, then press Enter. Repeat this process for each keyword that you want to add. To delete a keyword that you added to the list, select the *X* next to the keyword.

     > **Tip**
     >
     > If you want to exclude a keyword from scoring, but you want to score that keyword when it's used in combination with other keywords or a phrase, enter the keyword you want to exclude in the **Exclude these keywords** field, then enter the word or words that are part of the phrase that you do want to score in the **Exclude the above keywords only if these terms aren't also present** field. For example, if you add the keyword "compliance" to the **Exclude these keywords** field, but enter the keyword "training" in the **Exclude the above keywords only if**

these terms aren't also present field, the word "compliance" by itself is excluded from scoring, but the phrase "compliance training" is scored.

- **Keyword groups**. To select a keyword detection group that you already created, select **Add keyword group**, then select the appropriate keyword groups from the list. The number of keywords included in each group displays under the **Included keywords** heading.

6. Select **Save**.

Configure a sensitive info type (preview) exclusion

Excluded sensitive info types map to indicators and triggers involving file-related activities for Endpoint, SharePoint, Teams, OneDrive, and Exchange. The system treats these excluded types as non-sensitive info types. If a file contains any sensitive info type you identify in this section, the system risk scores the file but doesn't show activities involving content related to sensitive info types. For a complete list of sensitive info types, see Sensitive information type entity definitions.

You can select sensitive info types to exclude from the list of all available (out-of-box and custom) types available in your organization. You can choose up to 500 sensitive info types.

> **Note**
>
> The exclusion list of sensitive info types takes precedence over the **priority content** list.

Configure a sensitive info type exclusion

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.

2. Select **Settings** in the upper-right corner of the page, then select **Insider Risk Management** to go to the Insider Risk Management settings.

3. Under **Insider risk settings**, select **Global exclusions**.

4. In the **Type** panel to the right, select **Sensitive info types**.

5. In the **Sensitive info types** panel to the right, select one of the following tabs:

- **Individual sensitive info types**. To add sensitive info types one at a time, select **Add or edit sensitive info types to exclude**, select the sensitive info types you want to exclude, and then select **Add**. You can also use the search box to search for a sensitive info type.
- **SIT groups**. To select a sensitive info type detection group that you already created, select **Add sensitive info type group**, then select the appropriate groups from the list. The number of sensitive info types included in a sensitive info type group displays under the **Included sensitive info types** heading. Select **Save** when you're done.

Configure a Sharepoint site exclusion

You can configure SharePoint site exclusions to prevent activities that occur in SharePoint (and SharePoint sites associated with Teams channel sites) from generating policy alerts. For example, you might want to exclude sites or channels that contain non-sensitive files and data that you can share with stakeholders or the public. You can enter up to 500 SharePoint site URL paths.

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.

2. Select **Settings** in the upper-right corner of the page, then select **Insider Risk Management** to go to the Insider Risk Management settings.

3. Under **Insider risk settings**, select **Global exclusions**.

4. In the **Type** panel to the right, select **SharePoint sites**.

5. In the **SharePoint sites** panel to the right, select one of the following tabs:

   - **Individual sites**. To add SharePoint sites one at a time, select **Add or edit sites to exclude**, select the sites you want to exclude, then select **Edit**. You can also use the search box to search for a site.
   - **Site groups**. To select a SharePoint site detection group that you already created, select **Add SharePoint site group**, then select the appropriate groups from the list. The number of sites included in a site group displays under the **Included sites** heading. Select **Save** when you're done.

Set up a trainable classifier (preview) exclusion

Excluded trainable classifiers map to indicators and triggers that involve file-related activities for SharePoint, Teams, OneDrive, and Exchange. If any file contains a trainable classifier identified as

an exclusion, the file is risk scored but not shown as activity involving content related to trainable classifiers. For a complete list of pre-trained classifiers, see Trainable classifiers definitions.

You can select trainable classifiers to exclude from the list of all available (out-of-box and custom) types in your organization. Insider Risk Management excludes some trainable classifiers by default, including Threat, Profanity, Targeted harassment, Offensive language, and Discrimination. You can choose up to 500 trainable classifiers.

> **Note**
>
> Optionally, you can choose trainable classifiers to include in the priority content list.

Configure a trainable classifier exclusion

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.
2. Select **Settings** in the upper-right corner of the page, then select **Insider Risk Management** to go to the Insider Risk Management settings.
3. Under **Insider risk settings**, select **Global exclusions**.
4. In the **Type** panel to the right, select **Trainable classifiers**.
5. In the **Trainable classifiers** panel to the right, select one of the following tabs:
   - **Individual trainable classifier**. To select trainable classifiers one at a time, select **Add or edit trainable classifiers to exclude**, select the appropriate trainable classifiers from the list, and then select **Add**. You can use the search box to search for a trainable classifier.
   - **Trainable classifier groups**. To select a trainable classifier detection group that you already created, select **Add trainable classifier group**, then select the appropriate groups from the list. The number of trainable classifiers included in a trainable classifier group displays under the **Included trainable classifiers** heading. Select **Save** when you're done.

Last updated on 07/21/2026

Set policy timeframes in Insider Risk Management

> **Important**
>
> Microsoft Purview Insider Risk Management correlates various signals to identify potential malicious or inadvertent insider risks, such as IP theft, data leakage, and security violations. Insider Risk Management enables customers to create policies to manage security and compliance. Built with privacy by design, users are pseudonymized by default, and role-based access controls and audit logs are in place to help ensure user-level privacy.

Policy timeframes in Microsoft Purview Insider Risk Management let you set past and future review periods that start after policy matches events and activities for the Insider Risk Management policy templates. Depending on the policy template you choose, you can set the following policy timeframes:

- **Activation window**: Available for all policy templates, *Activation window* is the number of days that the window activates **after** a triggering event. The window activates for 1 to 30 days after a triggering event occurs for any user assigned to the policy. For example, you configure an Insider Risk Management policy and set *Activation window* to 30 days. Several months pass since you configured the policy, and a triggering event occurs for one of the users included in the policy. The triggering event activates *Activation window* and the policy is active for that user for 30 days after the triggering event occurred.

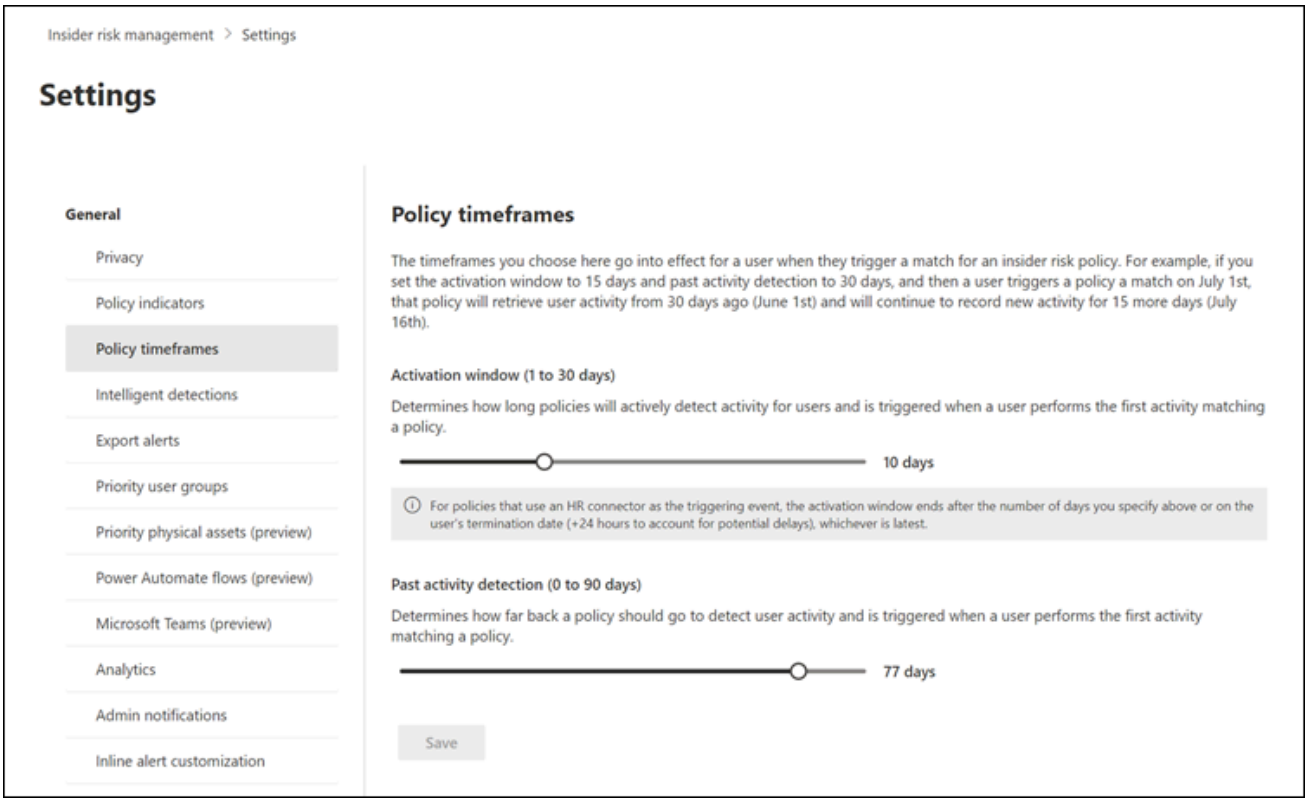
- **Past activity detection**: Available for all policy templates, *Past activity detection* is the number of days that the window activates **before** a triggering event. For activities in the audit log, the window activates for 0 to 90 days before a triggering event occurs for any user assigned to the policy. For example, you configure an Insider Risk Management policy and set *Past activity detection* to 90 days. Several months pass since you configured the policy, and a triggering event occurs for one of the users included in the policy. The triggering event activates *Past activity detection* and the policy gathers historic activities for that user for 90 days prior to the triggering event.

  > **Note**

For email activities, the past activity detection period is 10 days.

Configure policy timeframes

1. In Insider Risk Management settings, select **Policy timeframes**.
2. Move the slider for **Activation window** and **Past activity detection** to the number of days that you want.



ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/25/2026

Configure intelligent detections in Insider Risk Management

> **Important**
>
> Microsoft Purview Insider Risk Management correlates various signals to identify potential malicious or inadvertent insider risks, such as IP theft, data leakage, and security violations. Insider Risk Management enables customers to create policies to manage security and compliance. Built with privacy by design, users are pseudonymized by default, and role-based access controls and audit logs are in place to help ensure user-level privacy.

> **Note**
>
> Global exclusions settings that the **Intelligent detections** setting previously included now appear in the Global exclusions (preview) setting.

Use the **Intelligent detections** setting in Microsoft Purview Insider Risk Management to:

- Boost the score for unusual file download activities by entering a minimum number of daily events.
- Increase or decrease the volume and distribution of high, medium, and low alerts.
- Import and filter Defender for Endpoint alerts for activities used in policies created from Insider Risk Management templates.
- Specify unallowed domains to boost the risk score for potentially risky activity.
- Specify third-party domains to generate alerts for potentially risky download activity.

File activity detection

Use this section to specify the number of daily events required to boost the risk score for download activity that's unusual for a user. For example, if you enter "25", and a user downloads 10 files on average over the previous 30 days, but a policy detects that they downloaded 20 files on one day, the score for that activity doesn't get boosted even though it's unusual for that user because the number of files they downloaded that day is less than 25.

Configure alert volume levels

Insider risk policies assign a specific risk score to potentially risky activities they detect. This risk score determines the alert severity: low, medium, or high. By default, Insider Risk Management generates a certain number of low, medium, and high severity alerts. You can increase or decrease the volume of a specific level of alerts to suit your needs.

To adjust the volume of alerts for all Insider Risk Management policies, choose one of the following settings:

- **Fewer alerts**: You see all high-severity alerts, fewer medium-severity alerts, and no low-severity alerts. You might miss some true positives if you choose this setting.
- **Default volume**: You see all high-severity alerts and a balanced amount of medium-severity and low-severity alerts.
- **More alerts**: You see all medium-severity and high-severity alerts and most low-severity alerts. This setting might result in more false positives.

Microsoft Defender for Endpoint alert statuses

> **Important**
>
> To import security violation alerts, you must configure Microsoft Defender for Endpoint in your organization and enable Defender for Endpoint for Insider Risk Management integration in the Defender Security Center. For more information on configuring Defender for Endpoint for Insider Risk Management integration, see Configure advanced features in Defender for Endpoint.

Microsoft Defender for Endpoint is an enterprise endpoint security platform designed to help enterprise networks prevent, detect, investigate, and respond to advanced threats. To have better visibility of security violations in your organization, you can import and filter Defender for Endpoint alerts for activities used in policies created from Insider Risk Management security violation policy templates.

Depending on the types of signals you're interested in, you can choose to import alerts to Insider Risk Management based on the Defender for Endpoint alert triage status. You can define one or more of the following alert triage statuses in the global settings to import:

- Unknown
- New
- In progress
- Resolved

Alerts from Defender for Endpoint are imported daily. Depending on the triage status you choose, you might see multiple user activities for the same alert as the triage status changes in Defender for Endpoint.

For example, if you select *New*, *In progress*, and *Resolved* as the alert triage statuses to import, when a Microsoft Defender for Endpoint alert is generated and the status is *New*, an initial alert activity is imported for the user in Insider Risk Management. When the Defender for Endpoint triage status changes to *In progress*, a second activity for this alert is imported. When the final Defender for Endpoint triage status of *Resolved* is set, a third activity for this alert is imported. This functionality allows investigators to follow the progression of the Defender for Endpoint alerts and choose the level of visibility that their investigation requires.

Configure unallowed and third-party domains

To boost your detections, specify unallowed and third-party domains:

- **Unallowed domains:** When you specify an unallowed domain, risk management activity that takes place with that domain has a higher risk score. Optionally, use it as a setting in policy triggers. For example, you might want to specify activities that involve sharing content with someone (such as sending email to someone with a gmail.com address) or activities that involve users downloading content to a device from an unallowed domain. You can add up to 500 unallowed domains.
- **Third-party domains:** If your organization uses third-party domains for business purposes (such as cloud storage), include them in the **Third-party domains** section to receive alerts for potentially risky activity related to the device indicator *Use a browser to download content from a third-party site*. You can add up to 500 third-party domains.

> **Tip**
>
> You can also specify domains to exclude from being scored by Insider Risk Management policies.

Add an unallowed domain

To add an unallowed domain to the Intelligent detections settings, perform the following steps:

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.

2. Select **Settings** in the upper-right corner of the page, and then select **Insider Risk Management** to go to the Insider Risk Management settings.
3. Under **Insider risk settings**, select **Intelligent detections**.
4. Scroll down to the **Unallowed domains** section, and then select **Add domains**.
5. Enter a domain.

   > **Tip**
   >
   > To add multiple domains at once, import them as a CSV file by selecting **Import domains from CSV file** on the previous page.

6. To include all subdomains within the domain you entered, select the **Include multi-level subdomains** checkbox.

   > **Note**
   >
   > Use wildcards to help match variations of root domains or subdomains. For example, to specify sales.wingtiptoys.com and support.wingtiptoys.com, use the wildcard entry "*.wingtiptoys.com" to match these subdomains (and any other subdomain at the same level). To specify multi-level subdomains for a root domain, select the **Include multi-level subdomains** checkbox.

7. Press Enter. Repeat steps 5 and 6 for each additional domain that you want to add.
8. Select **Add domains**.

Add a third-party domain

To add a third-party domain to the Intelligent detections settings, perform the following steps:

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.
2. Select **Settings** in the upper-right corner of the page, and then select **Insider Risk Management** to go to the Insider Risk Management settings.
3. Under **Insider risk settings**, select **Intelligent detections**.

4. Scroll down to the **Third-party domains** section, and then select **Add domains**.

5. Enter a domain.

   > **Tip**
   >
   > To add multiple domains at once, import them as a CSV file by selecting **Import domains from CSV file** on the previous page.

6. To include all subdomains within the domain you entered, select the **Include multi-level subdomains** checkbox.

   > **Note**
   >
   > Use wildcards to help match variations of root domains or subdomains. For example, to specify sales.wingtiptoys.com and support.wingtiptoys.com, use the wildcard entry "*.wingtiptoys.com" to match these subdomains (and any other subdomain at the same level). To specify multi-level subdomains for a root domain, select the **Include multi-level subdomains** checkbox.

7. Press Enter. Repeat steps 5 and 6 for each additional domain that you want to add.

8. Select **Add domains**.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Share Insider Risk Management data with other solutions

> **Important**
>
> Microsoft Purview Insider Risk Management correlates various signals to identify potential malicious or inadvertent insider risks, such as IP theft, data leakage, and security violations. Insider Risk Management enables customers to create policies to manage security and compliance. Built with privacy by design, users are pseudonymized by default, and role-based access controls and audit logs are in place to help ensure user-level privacy.

You can share data from Insider Risk Management in the following ways:

- Export alert information to SIEM solutions.
- Share alerts and user risk severity levels with Microsoft Defender XDR.
- Share user risk severity levels with data loss prevention (DLP) alerts.

Export alert information to SIEM solutions

You can export Microsoft Purview Insider Risk Management alert information to security information and event management (SIEM) and security orchestration automated response (SOAR) solutions by using the Office 365 Management Activity API schema. Use the Office 365 Management Activity APIs to export alert information to other applications your organization uses to manage or aggregate insider risk information. Alert information is exported and available every 60 minutes through the Office 365 Management Activity APIs.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

If your organization uses Microsoft Sentinel, you can also use the out-of-the-box Insider Risk Management data connector to import insider risk alert information to Microsoft Sentinel. For more information, see Insider Risk Management in the Microsoft Sentinel article.

> **Important**
>
> To maintain referential integrity for users who have insider risk alerts or cases in Microsoft 365 or other systems, anonymization of usernames isn't preserved for exported alerts when using the exporting API or when exporting to Microsoft Purview eDiscovery solutions. Exported alerts display usernames for each alert in this case. If you export to CSV files from alerts or cases, anonymization *is* preserved.

Use the APIs to review insider risk alert information

Use the following steps to enable alert export by using the APIs:

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.
2. Select **Settings** in the upper-right corner of the page.
3. Select **Insider Risk Management** to go to the Insider Risk Management settings.
4. Select **Export alerts**. By default, this setting is disabled for your Microsoft 365 organization.
5. Turn the setting **On**.
6. Filter the common Office 365 audit activities by *SecurityComplianceAlerts*.
7. Filter *SecurityComplianceAlerts* by the *InsiderRiskManagement* category.

Alert information contains information from the Office 365 Management Activity API Security and Compliance Alerts schema and common schema. The following fields and values are exported for Insider Risk Management alerts for the common schema:

- CreationTime
- ID
- Operation
- OrganizationId
- RecordType
- UserId
- UserKey
- UserType

Share alert severity levels with other Microsoft security solutions

You can share alert severity levels from Insider Risk Management to bring unique user context to alert investigation experiences in the following Microsoft security solutions:

- Microsoft Defender XDR
- Microsoft Purview Communication Compliance
- Microsoft Purview Data Loss Prevention (DLP)

Insider Risk Management analyzes user activities over a period of 90 to 120 days and looks for anomalous behavior over that period. Adding this data to other security solutions enhances the data available in those solutions to help analysts prioritize alerts.

> **Tip**
>
> *Alert severity levels* in Insider Risk Management are different from *insider risk levels* that are defined in Adaptive Protection.
>
> - Alert severity levels (Low, Medium, or High) are assigned to users based on activity detected in Insider Risk Management policies. The system calculates these levels from alert risk scores assigned to all active alerts associated with the user. These levels help insider risk analysts and investigators prioritize and respond to user activity accordingly.
> - Insider risk levels (Elevated, Moderate, or Minor) in Adaptive Protection measure risk determined by admin-defined conditions such as the number of exfiltration activities users perform in a day or whether their activity generated a high-severity insider risk alert.

Prerequisites

To share Insider Risk Management user risk levels with other Microsoft security solutions, the user must meet the following conditions:

- Be part of an Insider Risk Management policy.
- Perform exfiltration activities that bring the user into the scope of the policy.
- **For sharing with DLP**: Have DLP alert permissions. After the **Data sharing** setting is turned on, users with DLP alert permissions can access Insider Risk Management context for DLP alert investigation and for the Microsoft Defender XDR Users page. Users with Insider Risk Management permissions can also access this data.
- **For sharing with Communication Compliance**: Be assigned either the *Communication Compliance Analyst* or *Communication Compliance Investigator* roles to view user risk

severity levels and activity history in Communication Compliance.

> **Tip**
>
> If you have access to DLP alerts in Microsoft Purview and/or Microsoft Defender, you can view user context from Insider Risk Management shared with those solutions.

Share data with other Microsoft security solutions

You can share Insider Risk Management alert severity levels with other Microsoft security solutions by turning on a single setting.

1. In Insider Risk Management settings, select the **Data sharing** setting.
2. Under the **Sharing data with other Microsoft security solutions** section, turn on the setting.

> **Note**
>
> If you don't turn on this setting, the value displayed in the DLP alerts **Insider risk severity** column is "User data isn't available" and is displayed as "Insider Risk Activity not available" in Communication Compliance.

What happens when you share Insider Risk Management alert severity levels?

When you turn on data sharing, Insider Risk Management severity information becomes available in the following Microsoft security experiences:

In Microsoft Defender XDR

Sharing alert data in the Microsoft Defender portal is vital for protecting your organization's sensitive information and maintaining security. Security operations center (SOC) analysts can:

- Investigate Insider Risk Management alerts in the Microsoft Defender alert queue.
- Investigate Insider Risk Management alerts correlated with alerts from other detection sources such as Data Loss Prevention, Microsoft Defender for Identity, Microsoft Defender for Office, and more in the Microsoft Defender XDR incident queue.
- Run Advanced Hunting queries on two new tables containing Insider Risk Management alert data.
- Export rich Insider Risk Management alert data through Microsoft Graph API.

- View Insider Risk Management severity for a user during your alert triage. An **Insider risk severity** field is added to the **Users** page for users that have a **High**, **Medium**, or **Low** risk level in Insider Risk Management. This data is available for any users with an active Insider Risk Management alert. An insider risk activity summary and activity timeline for that user appears on the right side of the **Users** page.

Migrating from Office 365 management activity APIs to Microsoft Graph APIs

You can access Insider Risk Management alert metadata through Office 365 management activity API. However, Graph API provides richer metadata and bi-directional support. We recommend migrating to Graph API to integrate Insider Risk Management data with your enterprise systems.

The following table summarizes the most common Office 365 management activity API parameters and the equivalent Microsoft Graph API parameter:

Expand table

Office 365 management activity API parameter	Microsoft Graph API parameter
Alert parameter	No equivalent parameter
AlertId	ID
Category	ServiceSource
Comments	No equivalent parameter
Data	Evidence.useraccount.userPrincipalName(along with suffix)
Name	AlertPolicyName
PolicyId	AlertPolicyId
Severity	Severity
Source	DetectionSource
Status	Status
Version	No equivalent parameter

For detailed schema mapping, see the Defender XDR schema in the Office 365 management activity API and Microsoft Graph API.

Integrating Insider Risk Management with ServiceNow Security Operations

ServiceNow provides integrations to connect with Microsoft security technologies using Microsoft Graph. These solutions include Microsoft Sentinel, Microsoft Defender Advanced Threat Protection, and Azure Advanced Threat Protection. The ServiceNow Microsoft Graph integration allows you to access valuable insights from Microsoft products and helps you efficiently manage and respond to security incidents centrally using the ServiceNow platform. Insider Risk Management alert data is accessed outside the Microsoft Purview environment through the Microsoft Graph security API.

To share Insider Risk Management alert data with ServiceNow, complete the following steps:

1. Get the Microsoft Graph Security API Alert Ingestion Integration For Security Operations
2. Follow the ServiceNow documentation to set up or create a profile for the Microsoft Graph Security API alert ingestion integration.

For more information, see Investigate insider risk threats in the Microsoft Defender portal.

In Communication Compliance alerts

For every Communication Compliance policy match, you can view the user risk severity associated with the sender. View this information on the **User activity** tab in communication for the alert. This view provides risk profile, policy matches, and user activities captured by Insider Risk Management and Communication Compliance.

The severity levels are categorized as *High*, *Medium*, *Low*, or *None*.

For risk severity levels of *None*, the reason might be for any of the following scenarios:

- The user isn't included in an insider risk policy.
- The user's activities aren't assigned a risk score, meaning the user isn't in the active scope for the policy.
- The user is included in an Insider Risk Management policy, but didn't engage in any risky activity.
- The organization doesn't have an active Insider Risk Management policy.

If the user risk severity isn't available, data sharing isn't enabled from Insider Risk Management.

You can view insider risk activities for up to 120 days in the **View Details** section on the **User history** tab in Insider Risk Management. Currently, the user activity summary in Communication Compliance shows only data from exfiltration indicators.

In DLP alerts

In DLP alerts, shared Insider Risk Management data appears in the following ways:

- For the Insider Risk Management policy associated with the DLP alert, the DLP alerts queue includes an **Insider risk severity** column with values of *High*, *Medium*, *Low*, or *None*. If multiple users have activities that match the policy, the queue displays the user with the highest insider risk level.

  A value of **None** can mean either of the following:

  - The user isn't part of any Insider Risk Management policy.
  - The user is part of an Insider Risk Management policy, but they didn't perform risky activities that bring them into the scope of the policy (there's no exfiltration data).

- You can select the insider risk level in the DLP alerts queue to access the **User activity summary** tab, which shows a timeline of all exfiltration activities for that user for the past 90-120 days. Like the DLP alerts queue, the **User activity summary** tab shows the user with the highest insider risk level. This deep context into what a user did over the past 90 to 120 days provides a wider view of the risks presented by that user.

  Only data from exfiltration indicators appear in the user activity summary. Data from other sensitive indicators, such as HR, browsing, and so on, don't share with DLP alerts.

- An **Actor details** section is added to the DLP Alert details page. You can use this page to see *all* users involved in the specific DLP alert. For each user involved in the DLP alert, you can view all the exfiltration activities for the past 90 to 120 days.

- If you select **Get a summary from Security Copilot** in a DLP alert, the alert summary provided by Microsoft Security Copilot includes the Insider Risk Management severity level in addition to the DLP summary info, if the user is in scope of an Insider Risk Management policy.

  > **Tip**
  >
  > You can also use Security Copilot to investigate DLP alerts. If the Insider Risk Management **Data sharing** setting is turned on, you can then do a combined DLP/Insider Risk Management investigation. For example, you might want to start by asking Copilot to summarize a DLP alert, then ask Copilot to show the insider risk level associated with the user flagged in the alert. Or you might want to ask why the user is

considered a high-risk user. The user risk information in this case comes from Insider Risk Management. Security Copilot seamlessly integrates Insider Risk Management with DLP to assist with investigations. **Learn more about using the standalone version of Copilot for combined DLP/Insider Risk Management investigations**.

See also

- Investigate data loss prevention alerts with Microsoft 365 Defender XDR
- Learn about investigating data loss prevention alerts
- Get started with the data loss prevention Alerts dashboard

ⓘ**Note:** The author created this article with assistance from AI. Learn more

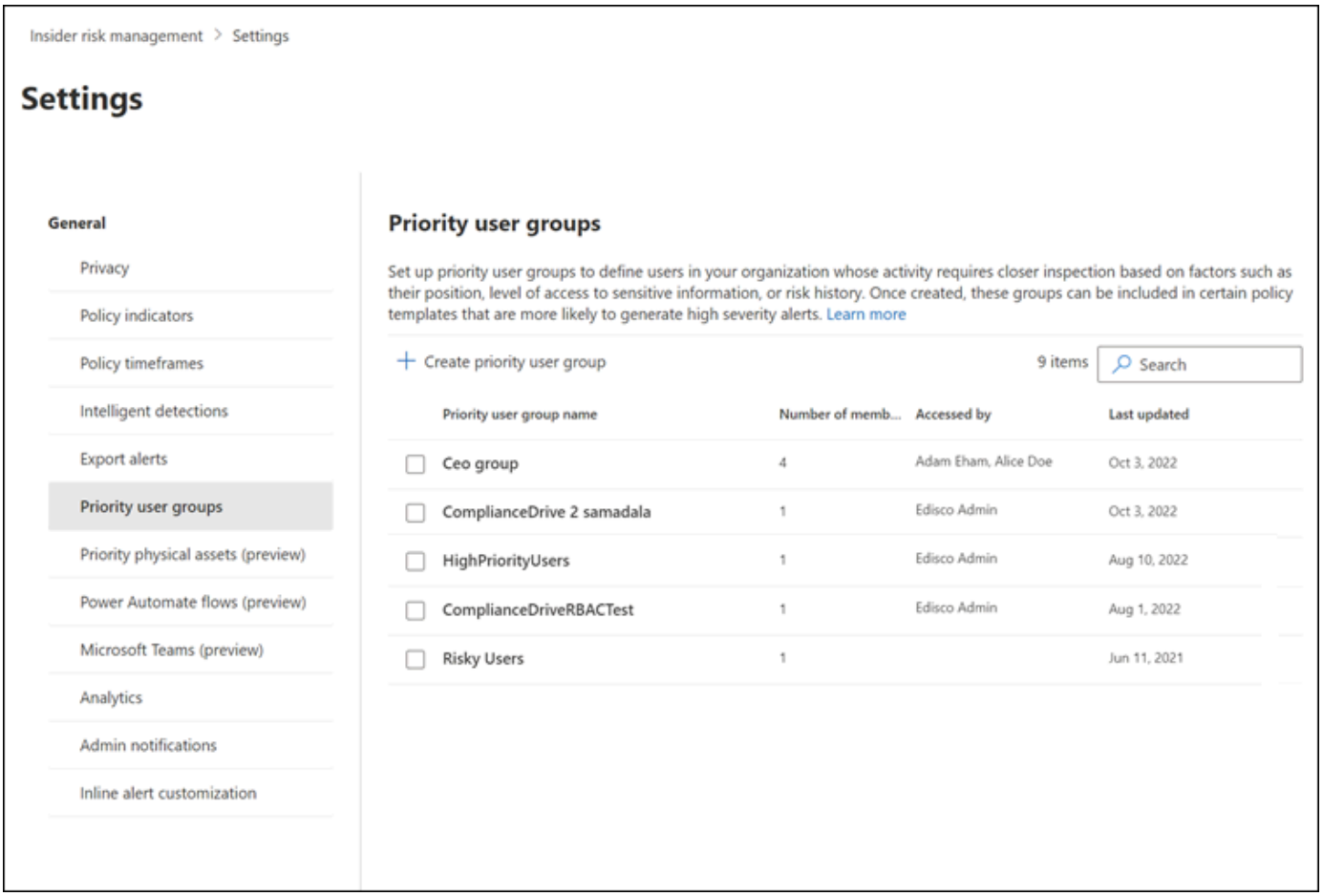
---

Last updated on 06/22/2026

Prioritize user groups for Insider Risk Management policies

> **Important**
>
> Microsoft Purview Insider Risk Management correlates various signals to identify potential malicious or inadvertent insider risks, such as IP theft, data leakage, and security violations. Insider Risk Management enables customers to create policies to manage security and compliance. Built with privacy by design, users are pseudonymized by default, and role-based access controls and audit logs are in place to help ensure user-level privacy.

Users in your organization might have different levels of risk depending on their position, level of access to sensitive information, or risk history. Prioritizing the examination and scoring of the activities of these users can help alert you to potential risks that might have higher consequences for your organization. Priority user groups in Microsoft Purview Insider Risk Management help define the users in your organization that need closer inspection and more sensitive risk scoring. Coupled with the *Security policy violations by priority users* and *Data leaks by priority users* policy templates, users you add to a priority user group have an increased likelihood of insider risk alerts and alerts with higher severity levels.



Instead of being open to review by all analysts and investigators, you might need to restrict review activities for priority user groups to specific users or insider risk role groups. You can assign individual users and role groups to review users, alerts, cases, and reports for each priority user group. You can assign review permissions for priority user groups to the built-in *Insider Risk Management*, *Insider Risk Management Analysts*, and *Insider Risk Management Investigators* role groups, one or more of these role groups, or to a custom selection of users.

For example, you need to protect against data leaks for a highly confidential project where users have access to sensitive information. You choose to create the *Confidential Project Users* priority user group for users in your organization that work on this project. Also, you don't want all the default Insider Risk Management admins, analysts, and investigators to see users, alerts, cases, and reports associated with this priority user group. In settings, you create the *Confidential Project Users* priority user group and assign two users as reviewers that can view data related to the group. Use the policy workflow and the *Data leaks by priority users* policy template to create a new policy and assign the *Confidential Project Users* priority users group to the policy. Activities examined by the policy for members of the *Confidential Project Users* priority user group are more sensitive to risk and activities by these users are more likely to generate alerts and have alerts with higher severity levels.

Create a priority user group

You must be a member of the *Insider Risk Management* or *Insider Risk Management Admins* role group to create a priority user group.

1. Sign in to the Microsoft Purview portal as a member of the *Insider Risk Management* or *Insider Risk Management Admins* role group.
2. Select **Settings** in the upper-right corner of the page.
3. Select **Insider Risk Management** to go to the Insider Risk Management settings.
4. Select **Priority user groups**.
5. On the **Priority user groups** page, select **Create priority user group** to start the group creation workflow.
6. On the **Name and describe the priority user group** page, complete the following fields:
   - **Name (required)**: Enter a friendly name for the priority user group. You can't change the name of the priority user group after you complete the workflow.
   - **Description (optional)**: Enter a description for the priority user group.
7. Select **Next** to continue.
8. On the **Choose members** page, select **Choose members** to search and select which mail-enabled user accounts are included in the group or select the **Select all** checkbox to add all users in your organization to the group. Select **Add** to continue.

   > **Note**
   >
   > You can add up to 10,000 users to a priority user group.

9. Select **Upload** to add members by uploading a CSV file if needed. The .csv file must have a column titled *user principal name* with the list of users you want to add.
10. Select **Next** to continue.
11. On the **Choose who can view this group** page, you must define who can review users, alerts, cases, and reports for the priority user group. Assign at least one user or Insider Risk Management role group. Select **Choose users and role groups**, and then select the users or Insider Risk Management role groups you want to assign to the priority user group. Select **Add** to assign the selected users or role groups to the group.

12. Select **Next** to continue.
13. On the **Review** page, review the settings you chose for the priority user group. Select the **Edit** links to change any of the group values or select **Submit** to create and activate the priority user group.
14. On the confirmation page, select **Done**.

Update a priority user group

To update a priority user group, you must be a member of the *Insider Risk Management* or *Insider Risk Management Admins* role group.

1. Sign in to the Microsoft Purview portal as a member of the *Insider Risk Management* or *Insider Risk Management Admins* role group.
2. Select **Settings** in the upper-right corner of the page.
3. Select **Insider Risk Management** to go to the Insider Risk Management settings.
4. Select **Priority user groups**.
5. Select the priority user group you want to edit, then select **Edit group**.
6. On the **Name and describe** page, update the **Description** field if needed. You can't update the name of the priority user group. Select **Next** to continue.
7. On the **Choose members** page, add new members to the group using the **Choose members** control. To remove a user from the group, select the 'X' next to the user you want to remove. Select **Next** to continue.
8. On the **Choose who can view this group** page, add or remove users or role groups that can review users, alerts, cases, and reports for the priority user group.
9. Select **Next** to continue.
10. On the **Review** page, review the update settings for the priority user group. Select the **Edit** links to change any of the group values or select **Submit** to update the priority user group.
11. On the confirmation page, select **Done**.

Delete a priority user group

> **Important**
>
> Deleting a priority user group removes it from any active policy to which it's assigned. If you delete a priority user group that an active policy uses, the policy has no included users. The policy is idle and doesn't create alerts.

To delete a priority user group, you must be a member of the *Insider Risk Management* or *Insider Risk Management Admin* role group.

1. Sign in to the Microsoft Purview portal as a member of the *Insider Risk Management* or *Insider Risk Management Admin* role group.
2. Select **Settings** in the upper-right corner of the page.
3. Select **Insider Risk Management** to go to the Insider Risk Management settings.
4. Select **Priority user groups**.
5. Select the priority user group you want to delete, then select **Delete**.
6. In the **Delete** dialog box, select **Yes**.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/22/2026

Identify priority physical assets for Insider Risk Management policies

> **Important**
>
> Microsoft Purview Insider Risk Management correlates various signals to identify potential malicious or inadvertent insider risks, such as IP theft, data leakage, and security violations. Insider Risk Management enables customers to create policies to manage security and compliance. Built with privacy by design, users are pseudonymized by default, and role-based access controls and audit logs are in place to help ensure user-level privacy.

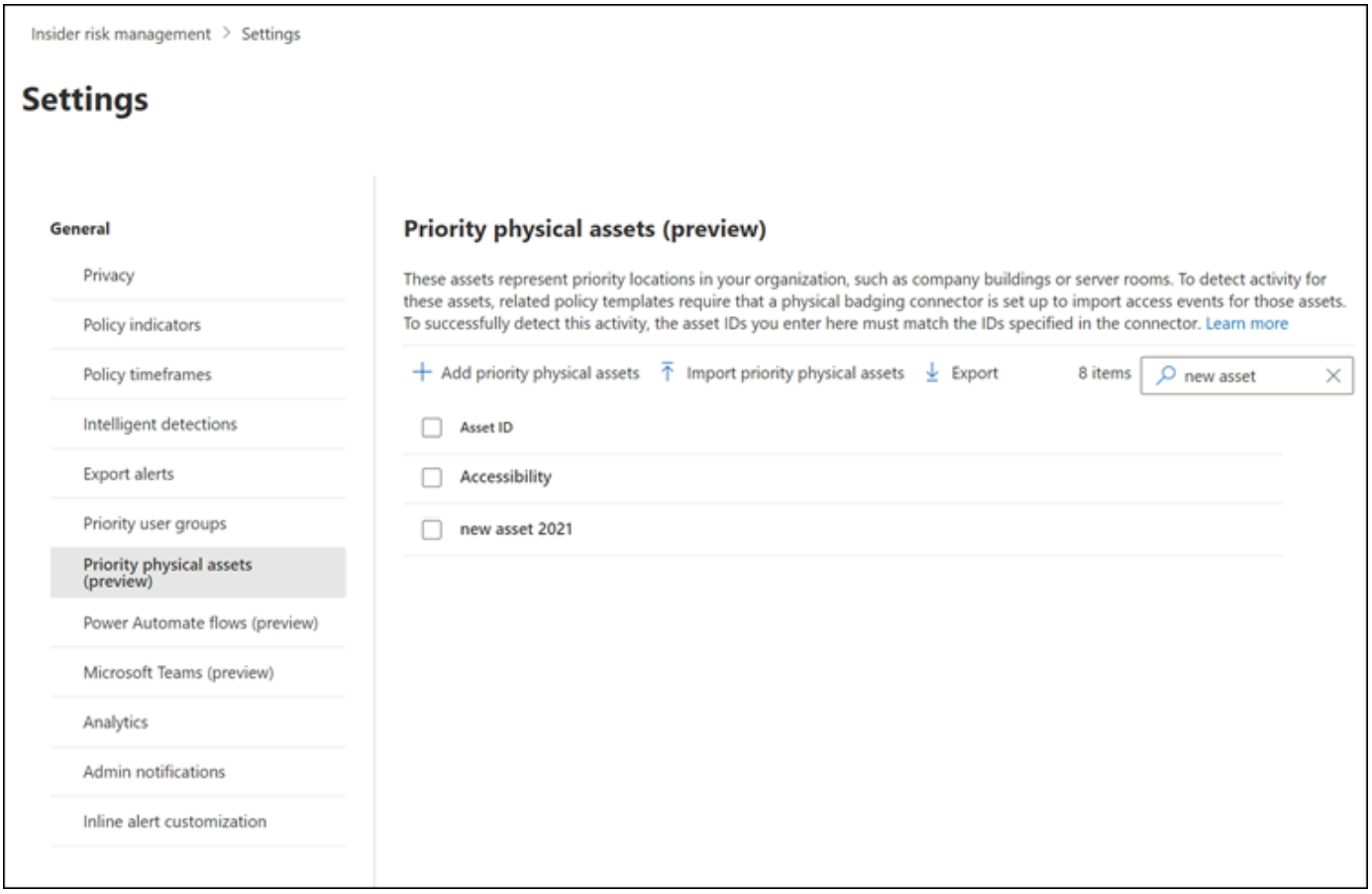
Identifying access to priority physical assets and correlating access activity to user events is an important component of your Microsoft Purview Insider Risk Management compliance infrastructure. These physical assets represent priority locations in your organization, such as company buildings, data centers, or server rooms. Insider risk activities might be associated with users working unusual hours, attempting to access unauthorized sensitive or secure areas, and requests for access to high-level areas without legitimate needs.

When you enable priority physical assets and configure the Physical badging data connector, Insider Risk Management integrates signals from your physical control and access systems with other user risk activities. By examining patterns of behavior across physical access systems and correlating these activities with other insider risk events, Insider Risk Management helps compliance investigators and analysts make more informed response decisions for alerts.

Insider Risk Management scores and identifies access to priority physical assets in insights differently from access to non-priority assets. For example, your organization has a badging system for users that governs and approves physical access to normal working and sensitive project areas. You have several users working on a sensitive project and these users return to other areas of your organization when the project is completed. As the sensitive project nears completion, you want to make sure that the project work remains confidential and that access to the project areas is tightly controlled.

You enable the Physical badging data connector in Microsoft 365 to import access information from your physical badging system and specify priority physical assets in Insider Risk Management. By importing information from your badging system and correlating physical access information with other risk activities identified in Insider Risk Management, you notice that

one of the users on the project is accessing the project offices after normal working hours and is also exporting large amounts of data to a personal cloud storage service from their normal work area. This physical access activity associated with the online activity might point to possible data theft and compliance investigators and analysts can take appropriate actions as dictated by the circumstances for this user.



Configure priority physical assets

To configure priority physical assets, configure the Physical badging connector and use setting controls in the Insider Risk Management solution. To configure priority physical assets, you must be a member of the *Insider Risk Management* or *Insider Risk Management Admin* role group.

> **Important**
>
> For Insider Risk Management policies to use and correlate signal data related to departing and terminated users with event data from your physical control and access platforms, also configure the Microsoft 365 HR connector. If you enable the Physical badging connector without enabling the Microsoft 365 HR connector, Insider Risk Management policies only process events for physical access activities for users in your organization.

1. Follow the configuration steps for Insider Risk Management in the Getting started with Insider Risk Management article. In Step 3, make sure to configure the Physical badging connector.

2. Sign in to the Microsoft Purview portal using credentials for an admin account in your Microsoft 365 organization, select **Settings** in the upper-right corner of the page, and select **Insider Risk Management** to open the Insider Risk Management settings. Select **Priority physical assets**.

3. On the **Priority physical assets** page, you can either manually add the physical asset IDs imported by the Physical badging connector or import a CSV file of all physical asset IDs imported by the Physical badging connector:
   a. To manually add physical asset IDs, choose **Add priority physical assets**, enter a physical asset ID, and then select **Add**. Enter other physical asset IDs, then select **Add priority physical assets** to save all the assets you entered.
   b. To add a list of physical asset IDs from a CSV file, choose **Import priority physical assets**. In the file explorer dialog box, select the CSV file you want to import, then select **Open**. The physical asset IDs from the CSV files are added to the list.

4. In settings, select **Policy indicators**.

5. On the **Policy indicators** page, go to the **Physical access indicators** section, then select the **Physical access after termination or failed access to sensitive asset** check box.

6. Select **Save**.

Delete a priority physical asset

To delete a priority physical asset, you must be a member of the *Insider Risk Management* or *Insider Risk Management Admins* role group.

> **Important**
>
> When you delete a priority physical asset, you remove it from examination by any active policy that includes it. However, alerts generated by activities associated with the priority physical asset aren't deleted.

1. Sign in to the Microsoft Purview portal as a member of the *Insider Risk Management* or *Insider Risk Management Admins* role group.
2. Select **Settings** in the upper-right corner of the page.

3. Select **Insider Risk Management** to go to the Insider Risk Management settings.
4. Select **Priority physical assets**.
5. On the **Priority physical assets** page, select the asset you want to delete.
6. Select **Delete**.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/22/2026

Automate Insider Risk Management actions with Microsoft Power Automate flows

> **Important**
>
> Microsoft Purview Insider Risk Management correlates various signals to identify potential malicious or inadvertent insider risks, such as IP theft, data leakage, and security violations. Insider Risk Management enables customers to create policies to manage security and compliance. Built with privacy by design, users are pseudonymized by default, and role-based access controls and audit logs are in place to help ensure user-level privacy.

Microsoft Power Automate is a workflow service that automates actions across applications and services. By using flows from templates or created manually, you can automate common tasks associated with these applications and services. When you enable Power Automate flows for Microsoft Purview Insider Risk Management, you can automate important tasks for cases and users. You can configure Power Automate flows to retrieve user, alert, and case information and share this information with stakeholders and other applications, as well as automate actions in Insider Risk Management, such as posting to case notes. Power Automate flows are applicable for cases and any user in scope for a policy.

Customers with Microsoft 365 subscriptions that include Insider Risk Management don't need extra Power Automate licenses to use the recommended Insider Risk Management Power Automate templates. These templates can be customized to support your organization and cover core Insider Risk Management scenarios. If you choose to use premium Power Automate features in these templates, create a custom template by using the Microsoft Purview connector, or use Power Automate templates for other compliance areas in Microsoft 365, you might need more Power Automate licenses.

The following Power Automate templates support process automation for Microsoft Purview Insider Risk Management users and cases:

- **Notify users when they're added to an insider risk policy**: Use this template if your organization has internal policies, privacy requirements, or regulatory requirements that users must be notified when they're subject to Microsoft Purview Insider Risk Management policies. When you configure this flow and select it for a user, the user and their manager receive an email message when the user is added to a Microsoft Purview Insider Risk

Management policy. This template also supports updating a SharePoint list hosted on a SharePoint site to detect notification message details like date/time and the message recipient. If you choose to anonymize users through the **Privacy** setting, flows created from this template don't function as intended so that user privacy is maintained.

Power Automate flows using this template are available on the **Users dashboard**.

- **Request information from HR or business about a user in an insider risk case**: When acting on a case, insider risk analysts and investigators might need to consult with HR or other stakeholders to understand the context of the case activities. When you configure this flow and select it for a case, analysts and investigators send an email message to HR and business stakeholders configured for this flow. Each recipient receives a message with preconfigured or customizable response options. When recipients select a response option, the response is recorded as a case note and includes recipient and date/time information. If you choose to anonymize users through the **Privacy** setting, flows created from this template don't function as intended so that user privacy is maintained.

  Power Automate flows using this template are available on the **Cases dashboard**.

- **Notify manager when a user has an insider risk alert**: Some organizations might need to have immediate management notification when a user has a Microsoft Purview Insider Risk Management alert. When you configure and select this flow, the manager for the case user receives an email message with the following information about all case alerts:
  - Applicable policy for the alert
  - Date/time of the alert
  - Severity level of the alert

  The flow automatically updates the case notes that the message was sent and that the flow was activated. If you choose to anonymize users through the **Privacy** setting, flows created from this template don't function as intended so that user privacy is maintained.

  Power Automate flows using this template are available on the **Cases dashboard**.

- **Create record for insider risk case in ServiceNow**: Use this template if your organization wants to use their ServiceNow solution to keep track of Microsoft Purview Insider Risk Management cases. When in a case, insider risk analysts and investigators can create a record for the case in ServiceNow. You can customize this template to populate selected fields in ServiceNow based on your organization's requirements. For more information on available ServiceNow fields, see the ServiceNow Connector reference article.

Power Automate flows using this template are available on the **Cases dashboard**.

Create a Power Automate flow from an Insider Risk Management template

> **Important**
>
> Currently, only Power Automate flows that you create in the default environment work with Insider Risk Management.

To create a Power Automate flow in the settings area, you must be a member of the *Insider Risk Management* or *Insider Risk Management Admins* role group. To create a Power Automate flow with the **Manage Power Automate flows** option, you must be a member of at least one Insider Risk Management role group.

1. Perform one of the following actions:

   - Sign in to the Microsoft Purview portal, select **Settings** in the upper-right corner of the page, select **Insider Risk Management** to go to the Insider Risk Management settings, then select **Power Automate flows**.
   - In the **Cases dashboard** or the **Users dashboard**, select **Automate** > **Manage Power Automate flows**.

2. On the **Power Automate flows** page, select a recommended template from the **Insider Risk Management templates you may like** section.
3. Review the flow's embedded connections and check their statuses. If needed, update any connections that aren't displayed as available, then select **Continue**.
4. By default, the recommended flows come pre-configured with the recommended Insider Risk Management and Microsoft 365 service data fields required to complete the assigned task for the flow. If needed, customize the flow components by using the **Show advanced options** control and configuring the available properties for the flow component.
5. If needed, add any other steps to the flow by selecting **New step**. In most cases, you shouldn't need to add steps for the recommended default templates.
6. Select **Save draft** to save the flow for further configuration or select **Save** to complete the configuration for the flow.
7. Select **Close** to return to the **Power Automate flows** page. The new template appears as a flow on the **My flows** tabs and is automatically available from the **Automate** dropdown control when working with Insider Risk Management cases for the user creating the flow.

> **Important**
>
> To give other users in your organization access to the flow, you must share the flow.

Create a custom Power Automate flow for Insider Risk Management

> **Important**
>
> Currently, only Power Automate flows that you create in the default environment work with Insider Risk Management.

Some processes and workflows for your organization might be outside the recommended Insider Risk Management flow templates. In these cases, you might need to create custom Power Automate flows for Insider Risk Management areas. Power Automate flows are flexible and support extensive customization, but they have required steps to integrate with Insider Risk Management features.

To create a custom Power Automate template for Insider Risk Management:

1. **Check your Power Automate flow license**: To create customized Power Automate flows that use Insider Risk Management triggers, you need a Power Automate license. The recommended Insider Risk Management flow templates don't require extra licensing and are included as part of your Insider Risk Management license.
2. **Create an automated flow**: Create a flow that performs one or more tasks after it's triggered by an Insider Risk Management event. For details on how to create an automated flow, see Create a flow in Power Automate.
3. **Select the Microsoft Purview connector**: Search for and select the Microsoft Purview connector. This connector enables Insider Risk Management triggers and actions. For more information on connectors, see the Connector reference overview article.
4. **Choose Insider Risk Management triggers for your flow**: Insider Risk Management has two triggers available for custom Power Automate flows:

    - **For a selected Insider Risk Management case**: You can select flows with this trigger from the Insider Risk Management **Cases dashboard**.
    - **For a selected Insider Risk Management user**: You can select flows with this trigger from the Insider Risk Management **Users dashboard**.

5. **Choose Insider Risk Management actions for your flow:** You can choose from several actions for Insider Risk Management to include in your custom flow:

    - Get Insider Risk Management alert
    - Get Insider Risk Management case
    - Get Insider Risk Management user
    - Get Insider Risk Management alerts for a case
    - Add Insider Risk Management case note

Share a Power Automate flow with other users

By default, only the creator can access Power Automate flows they create. For other Insider Risk Management users to access and use a flow, the flow creator must share it. To share a flow, use the settings controls in the Insider Risk Management solution or the **Manage Power Automate flows** option from the **Automate** control when working directly in the **Cases dashboard** or **Users dashboard**. After you share a flow, everyone you share it with can access the flow in the **Automate** control dropdown in the **Cases dashboard** and **Users dashboard**.

To share a Power Automate flow in the settings area, you must be a member of the *Insider Risk Management* or *Insider Risk Management Admins* role group. To share a Power Automate flow with the **Manage Power Automate flows** option, you must be a member of at least one Insider Risk Management role group.

Share a Power Automate flow

1. Perform one of the following actions:

    - Sign in to the Microsoft Purview portal, select **Settings** in the upper-right corner of the page, select **Insider Risk Management** to go to the Insider Risk Management settings, then select **Power Automate flows**.
    - In the **Cases dashboard** or the **Users dashboard**, select **Automate** > **Manage Power Automate flows**.

2. On the **Power Automate flows** page, select the **My flows** or **Team flows** tab.
3. Select the flow to share, then select **Share** from the flow options menu.
4. On the flow sharing page, enter the name of the user or group you want to add as an owner for the flow.
5. In the **Connection Used** dialog box, select **OK** to acknowledge that the added user or group has full access to the flow.

Edit a Power Automate flow

To edit a Power Automate flow in the settings area, you must be a member of the *Insider Risk Management* or *Insider Risk Management Admins* role group. To edit a Power Automate flow with the **Manage Power Automate flows** option, you must be a member of at least one Insider Risk Management role group.

1. Perform one of the following actions:

    - Sign in to the Microsoft Purview portal, select **Settings** in the upper-right corner of the page, select **Insider Risk Management** to go to the Insider Risk Management settings, then select **Power Automate flows (preview)**.
    - In the **Cases dashboard** or the **Users dashboard**, select **Automate** > **Manage Power Automate flows**.

2. On the **Power Automate flows** page, select a flow to edit, then select **Edit** from the flow control menu.
3. Select **ellipsis** > **Settings** to change a flow component setting or **ellipsis** > **Delete** to delete a flow component.
4. Select **Save**, then select **Close** to complete editing the flow.

Delete a Power Automate flow

> **Note**
>
> When you delete a flow, you remove it as an option for all users.

To delete a Power Automate flow in the settings area, you must be a member of the *Insider Risk Management* or *Insider Risk Management Admins* role group. To delete a Power Automate flow with the **Manage Power Automate flows** option, you must be a member of at least one Insider Risk Management role group.

1. Perform one of the following actions:

    - Sign in to the Microsoft Purview portal, select **Settings** in the upper-right corner of the page, select **Insider Risk Management** to go to the Insider Risk Management settings, then select **Power Automate flows**.
    - In the **Cases dashboard** or the **Users dashboard**, select **Automate** > **Manage Power Automate flows**.

2. On the **Power Automate flows** page, select a flow to delete, then select **Delete** from the flow control menu.
3. In the deletion confirmation dialog box, select **Delete**.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/22/2026

Enable Microsoft Teams for collaborating on Insider Risk Management cases

> **Important**
>
> Microsoft Purview Insider Risk Management correlates various signals to identify potential malicious or inadvertent insider risks, such as IP theft, data leakage, and security violations. Insider Risk Management enables customers to create policies to manage security and compliance. Built with privacy by design, users are pseudonymized by default, and role-based access controls and audit logs are in place to help ensure user-level privacy.

Compliance analysts and investigators can use Microsoft Teams to collaborate on Microsoft Purview Insider Risk Management cases. They can communicate with other stakeholders on Teams to:

- Coordinate and review response activities for cases in private Teams channels
- Securely share and store files and evidence related to individual cases
- Detect and review response activities by analysts and investigators

When you enable Teams for Insider Risk Management, the solution creates a dedicated team every time you confirm an alert and create a case. By default, the team automatically includes all members of the *Insider Risk Management*, *Insider Risk Management Analysts*, and *Insider Risk Management Investigators* role groups (up to 100 initial users). You can add additional organization contributors to the team after it's created and as appropriate.

For existing cases created before enabling Teams, analysts and investigators can choose to create a new team when working on a case, if needed. When you resolve the Insider Risk Management case linked to that team, the solution automatically archives the team (moves to hidden and read-only).

Learn more: Overview of teams and channels in Microsoft Teams.

Enable Teams support

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.
2. Select **Settings** in the upper-right corner of the page.

3. Select **Insider Risk Management** to go to the Insider Risk Management settings.
4. Select **Microsoft Teams (preview)**.
5. Turn the setting on.
6. Select **Save**.

Create a team for existing cases

If you enable Teams support for Insider Risk Management after you have existing cases, you need to manually create a team for each case, as needed. Users need permission to create Microsoft 365 Groups in your organization to create a team from a case. For more information about managing permissions for Microsoft 365 Groups, see Manage who can create Microsoft 365 Groups.

> **Note**
>
> After enabling Teams support in Insider Risk Management settings, when you create a new case, the system automatically creates a new team.

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.
2. Go to the **Insider Risk Management** solution.
3. Select **Cases** in the left navigation, then select an existing case.
4. On the case action menu, select **Create Microsoft Team**.
5. In the **Team name** field, enter a name for the new Microsoft Teams team.
6. Select **Create Microsoft team**, then select **Close**.

Depending on the number of users assigned to Insider Risk Management role groups, it might take 15 minutes for all investigators and analysts to be added to the newly created Microsoft Teams team for the case.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 06/22/2026

Enable analytics in Insider Risk Management

> **Important**
>
> Microsoft Purview Insider Risk Management correlates various signals to identify potential malicious or inadvertent insider risks, such as IP theft, data leakage, and security violations. Insider Risk Management enables customers to create policies to manage security and compliance. Built with privacy by design, users are pseudonymized by default, and role-based access controls and audit logs are in place to help ensure user-level privacy.

When you enable Microsoft Purview Insider Risk Management analytics, you get several important benefits. You can:

- Evaluate potential insider risks in your organization without configuring any insider risk policies.
- Get real-time guidance on configuring indicator threshold settings.
- Generate insider risk severity and user activity summary for the users not part of policies. Share this summary information with Data Loss Prevention, Communication Compliance, and Microsoft Defender.

Evaluate insider risks in your organization

Microsoft Purview Insider Risk Management analytics enables you to evaluate potential insider risks in your organization without configuring any insider risk policies. This evaluation can help your organization identify potential areas of higher user risk and help determine what types of Insider Risk Management policies to configure and how broadly to scope them. Analytics scans offer the following advantages for your organization:

- **Easy to configure**: Select **Run scan** when prompted by the analytics recommendation or go to **Insider risk settings** > **Analytics** and enable analytics.
- **Privacy by design**: Scanned results and insights are returned as aggregated and anonymized user activity. Reviewers can't identify individual usernames. Since Insider Risk Management doesn't classify any identity in the organization for analytics, the solution accounts for all the UPNs/identities that might be involved in data leaving the organization boundary. Accounting for all relevant identities might involve user accounts, system accounts, guest accounts, and non-person accounts.

- **Understand potential risks through consolidated insights**: Scan results help you quickly identify potential risk areas for your users and which policy best helps mitigate these risks.

To learn how analytics can help accelerate the identification of potential insider risks, see the Insider Risk Management Analytics video.

Areas scanned

Analytics scans several sources for risk management activity to help you identify insights into potential areas of risk. Depending on your current configuration, analytics looks for qualifying risk activities in the following areas:

- **Microsoft 365 audit logs**: Included in all scans, this primary source helps you identify most of the potentially risky activities.
- **Exchange Online**: Included in all scans, Exchange Online activity helps you identify activities where data in attachments are emailed to external contacts or services.
- **Microsoft Entra ID**: Included in all scans, Microsoft Entra history helps you identify risky activities associated with users with deleted user accounts.
- **Microsoft 365 HR data connector**: If configured, HR connector events help you identify risky activities associated with users that have resignation or upcoming termination dates.

Analytics insights from scans use the same risk management activity signals as Insider Risk Management policies. They report results based on both single and sequence user activities. However, the risk scoring for analytics uses up to 10 days of activity while insider risk policies use daily activity for insights. When you first enable and run analytics in your organization, you see the scan results for one day. If you leave analytics enabled, you see the results of each daily scan added to the insight reports for a maximum range of the previous 10 days of activity.

Receive real-time guidance on configuring indicator threshold settings

Manually tuning policies to reduce "noise" can be a time-consuming experience that requires a lot of trial and error testing to determine the desired configuration for your policies. If you turn on analytics, you can get real-time insights to help you efficiently adjust the selection of indicators and thresholds of activity occurrence so that you don't receive too few or too many policy alerts. Use real-time analytics recommendations to set thresholds.

Enable analytics and start a scan of potential insider risks in your organization

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

To enable insider risk analytics, you must be a member of the *Insider Risk Management*, *Insider Risk Management Admins*, or *Microsoft 365 Global admin* role group.

1. Sign in to the Microsoft Purview portal with an admin account in your Microsoft 365 organization.
2. Go to the **Insider Risk Management** solution.
3. On the **Overview** tab, scroll down to the **Insider risk analytics** card. Selecting **Run scan** turns on analytics scanning for your organization, and scan results can take up to 48 hours before insights are available as reports for review. Under **Scan for insider risks in your organization**, select **Run scan**.

> **Tip**
>
> You can also turn on scanning in your organization through **Settings** at the top of any Insider Risk Management page. Select **Analytics**, then turn on the setting.

View analytics insights after the first analytics scan

To view the results of analytics scanning in your organization, go to **Insider Risk Management** > **Reports** > **Analytics**. For more information about reports, see Use reports in Insider Risk Management.

Turn off analytics

To turn off insider risk analytics, you must be a member of the *Insider Risk Management*, *Insider Risk Management Admins*, or Microsoft 365 *Global admin* role group.

> **Important**

Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview **roles and permissions**.

1. Sign in to the Microsoft Purview portal with an admin account in your Microsoft 365 organization.
2. Select **Settings** in the upper-right corner of the page.
3. Select **Insider Risk Management** to go to the Insider Risk Management settings.
4. Under **Insider risk settings**, select **Analytics**, and then turn off the setting.

When you disable analytics:

- Analytics insight reports stop updating and don't show new risks.
- You can't see real-time analytics when you customize indicator threshold settings for your policies.

Enable user-level analytics in your organization

Insider Risk Management generates user activity summaries and insider risk severity insights for potentially risky behaviors at a user level. You can view these insights in the following experiences:

- Data Loss Prevention (DLP) alerts
- Communication Compliance policies
- Microsoft Defender user entity pages and alerts

User analytics cover all eligible users in your organization, including users not in the scope of any Insider Risk Management policy. When investigating alerts in Microsoft Defender, DLP, or Communication Compliance, analysts automatically have access to user information that helps improve risk assessments. Analysts can also use insider risk severity to quickly identify the most egregious and time-sensitive alerts and prioritize these alerts for investigation and remediation.

To enable insider risk user-level analytics, you must be a member of the *Insider Risk Management*, *Insider Risk Management Admins*, or Microsoft 365 *Global admin* role group.

Important

Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview **roles and permissions**.

1. Sign in to the Microsoft Purview portal with an admin account in your Microsoft 365 organization.
2. Select **Settings** in the upper-right corner of the page.
3. Select **Insider Risk Management** to go to the Insider Risk Management settings.
4. Under **Insider Risk Management settings**, select **Analytics** and turn on the **Show insights at user level** setting.
5. Under **Insider Risk Management settings**, select **Data sharing** and turn on the **Share user risk details with other security solutions** setting.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/22/2026

Enable admin notifications in Insider Risk Management

> **Important**
>
> Microsoft Purview Insider Risk Management correlates various signals to identify potential malicious or inadvertent insider risks, such as IP theft, data leakage, and security violations. Insider Risk Management enables customers to create policies to manage security and compliance. Built with privacy by design, users are pseudonymized by default, and role-based access controls and audit logs are in place to help ensure user-level privacy.

Admin notifications

When you turn on the **Admin notifications** setting in Microsoft Purview Insider Risk Management, the system automatically sends email notifications to selected role groups. You can send email notifications for the following scenarios:

- Send a notification email when the first alert is generated for a new policy. The system checks policies every 24 hours for first-time alerts and doesn't send notifications on subsequent alerts for the policy.
- Send a daily email when new high-severity alerts are generated. The system checks policies every 24 hours for high severity alerts.
- Send a weekly email summarizing policies that have unresolved warnings.

If you enable Insider Risk Management analytics for your organization, members of the *Insider Risk Management Admins* role group automatically receive an email notification for initial analytics insights for data leaks, theft, and exfiltration activities.

> **Note**
>
> Add individuals to the role group to receive notifications. Email notifications don't support Admin Units, Priority Users Groups, or Entra ID Security Groups.

Disable admin and analytics notifications

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.
2. Select **Settings** in the upper-right corner of the page, and then select **Insider Risk Management** to go to the Insider Risk Management settings.
3. Under **Insider risk settings**, select **Admin notifications**.
4. Clear the check boxes for the following options, as applicable:
    - **Send a notification email when the first alert is generated for a new policy**
    - **Send an email when an analytics scan detects an insight for the first time**
    - **Send an email when analytics is turned off for your organization**
5. Select **Save**.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/22/2026

Configure inline alert customization in Insider Risk Management

> **Important**
>
> Microsoft Purview Insider Risk Management correlates various signals to identify potential malicious or inadvertent insider risks, such as IP theft, data leakage, and security violations. Insider Risk Management enables customers to create policies to manage security and compliance. Built with privacy by design, users are pseudonymized by default, and role-based access controls and audit logs are in place to help ensure user-level privacy.

Inline alert customization in Microsoft Purview Insider Risk Management enables you to quickly tune an Insider Risk Management policy directly from the **Alert dashboard** while reviewing the alert. Alerts are generated when a risk management activity meets the thresholds configured in the related policy. To reduce the number of alerts you get from the risk management activity associated with the alert, you can change the thresholds or remove the risk management activity from the policy altogether.

You can enable inline alert customization to allow users assigned to the *Insider Risk Management Analysts* and *Insider Risk Management Investigators* role groups to edit policy thresholds and to disable specific indicators. If you don't enable inline alert customization, only users assigned to the *Insider Risk Management Admins* or *Insider Risk Management* role groups can edit these policy conditions. Inline alert customization supports alerts regardless of the current alert status, so analysts and investigators can update policies for *Dismissed* and *Resolved* alerts if needed.

When enabled, analysts and investigators can select **Reduce alerts for this activity** for an alert on the **Alert dashboard** and view details about the risk management activity and indicators associated with the alert. The portal displays the current policy thresholds for the number of events used to create low, medium, and high severity alerts. If you select **Reduce alerts for this activity** and a previous policy edit changes the threshold or removes the associated indicator, you see a notification message detailing previous changes to the policy.

Analysts and investigators can choose from the following options on the **Reduce alerts for this activity** pane to quickly edit the policy that created the alert:

- **Reduce alerts using Microsoft's recommended thresholds**: This option automatically increases the thresholds in the policy for you. You can review the new recommended threshold settings before changing the policy.
- **Reduce alerts by choosing your own thresholds**: This option lets you manually increase the thresholds for the risk management activity associated with the alert for the current and future alerts. You can review the current threshold settings and configure the new threshold settings before changing the policy.
- **Stop getting alerts for this activity**: This option removes the indicator associated with the selected alert from the policy and the risk management activity is no longer detected by the policy. This option applies to all indicators, regardless of whether the indicator is threshold-based.

After choosing an option, analysts and investigators can choose two options to update the policy:

- **Save and dismiss alert**: Saves the changes to the policy and updates the alert status to *Resolved*.
- **Save only**: Saves the changes to the policy, but the alert status remains the same.

Enable inline alert customization

Perform the following steps to enable inline alert customization:

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.
2. Select **Settings** in the upper-right corner of the page.
3. Select **Insider Risk Management** to go to the Insider Risk Management settings.
4. Under **Insider risk settings**, select **Inline alert customization**, and then turn on the setting.
5. Select **Save**.

> **Note**
>
> After turning on the **Inline alert customization** setting, it takes approximately one hour before inline alert customization is available in new and existing policy alerts.

Note: The author created this article with assistance from AI. Learn more

Last updated on 06/26/2026

Learn about Insider Risk Management policy templates

> **Important**
>
> Microsoft Purview Insider Risk Management correlates various signals to identify potential malicious or inadvertent insider risks, such as IP theft, data leakage, and security violations. Insider Risk Management enables customers to create policies to manage security and compliance. Built with privacy by design, users are pseudonymized by default, and role-based access controls and audit logs are in place to help ensure user-level privacy.

Policy templates

Insider Risk Management templates are predefined policy conditions that define the types of risk indicators and risk scoring model used by the policy. Each policy must have an assigned template in the policy creation workflow before the policy is created. Insider Risk Management supports up to 100 policies for each policy template. When you create a new insider risk policy with the policy workflow, choose from one of the following policy templates:

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Data theft by departing users

When users leave your organization, there are specific risk indicators typically associated with potential data theft by departing users. This policy template uses exfiltration indicators for risk scoring and focuses on detection and alerts in this risk area. Data theft for departing users might include downloading files from SharePoint Online, printing files, and copying data to personal cloud messaging and storage services near their employment resignation and end dates. When you use either the Microsoft HR connector or the option to automatically check for user account deletion in Microsoft Entra for your organization, this template starts scoring for risk indicators relating to these activities and how they correlate with user employment status. This can include cloud indicators for services such as Box, Dropbox, Google Drive, Amazon S3, and Azure.

> **Important**
>
> When you use this template, you can configure a Microsoft 365 HR connector to periodically import resignation and termination date information for users in your organization. See the Import data with the HR connector article for step-by-step guidance to configure the Microsoft 365 HR connector. If you choose not to use the HR connector, select the User account deleted from Microsoft Entra option when configuring trigger events in the policy workflow.

Data leaks

Protecting data and preventing data leaks is a constant challenge for most organizations, particularly with the rapid growth of new data created by users, devices, and services. Users can create, store, and share information across services and devices, which makes managing data leaks increasingly complex and difficult. Data leaks can include accidental oversharing of information outside your organization or data theft with malicious intent. This template scores real-time detections of suspicious SharePoint Online data downloads, file and folder sharing, printing files, and copying data to personal cloud messaging and storage services.

When you use a *Data leaks* template, you can assign a DLP policy to trigger indicators in the insider risk policy for high severity alerts in your organization. Whenever a high severity alert is generated by a DLP policy rule and added to the Office 365 audit log, insider risk policies created with this template automatically examine the high severity DLP alert. If the alert contains an in-scope user defined in the insider risk policy, the alert is processed by the insider risk policy as a new alert and assigned an insider risk severity and risk score. You can also choose to assign selected indicators as triggering events for a policy. This flexibility and customization helps scope the policy to only the activities covered by the indicators. This policy allows you to evaluate this alert in context with other activities included in the case.

This template can also include cloud indicators for services such as Box, Dropbox, Google Drive, Amazon S3, and Azure.

Data leaks policy guidelines

When creating or modifying data loss prevention policies for use with Insider Risk Management policies, consider the following guidelines:

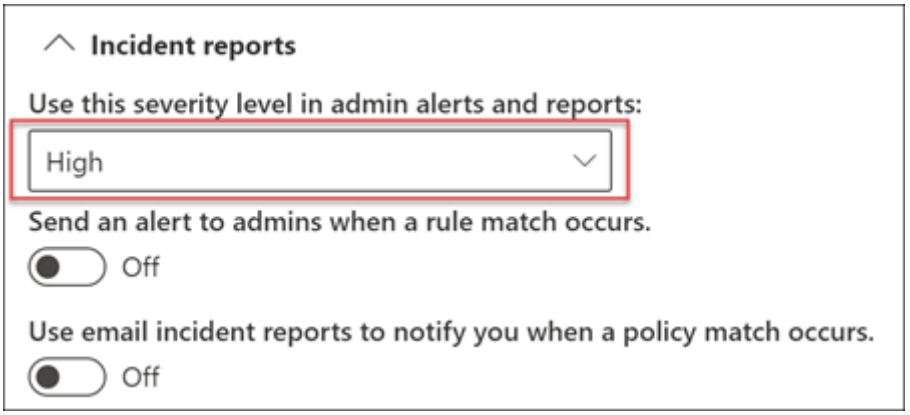
- Prioritize data exfiltration events and be selective when assigning **Incident reports** settings to *High* when configuring rules in your DLP policies. For example, emailing sensitive documents to a known competitor should be a *High* alert level exfiltration event. Over-assigning the *High* level in the **Incident reports** settings in other DLP policy rules

can increase the noise in the Insider Risk Management alert workflow and make it more difficult for your data investigators and analysts to properly evaluate these alerts. For example, assigning *High* alert levels to access denial activities in DLP policies makes it more challenging to evaluate truly risky user behavior and activities.

- When using a DLP policy as the triggering event, make sure you understand and properly configure the in-scope users in both the DLP and Insider Risk Management policies. Only users defined as in-scope for Insider Risk Management policies using the **Data leaks** template have high severity DLP policy alerts processed. Additionally, only users defined as in-scope in a rule for a high severity DLP alert are analyzed by the Insider Risk Management policy for consideration. Don't unknowingly configure in-scope users in both your DLP and insider risk policies in a conflicting manner.

  For example, if your DLP policy rules are scoped to only users on the Sales Team and the insider risk policy created from the **Data leaks** template has defined all users as in-scope, the insider risk policy only processes high severity DLP alerts for the users on the Sales Team. The insider risk policy doesn't receive any high priority DLP alerts for users to process that aren't defined in the DLP rules in this example. Conversely, if your Insider Risk Management policy created from **Data leaks** templates is scoped to only users on the Sales Team and the assigned DLP policy is scoped to all users, the insider risk policy only processes high severity DLP alerts for members of the Sales Team. The Insider Risk Management policy ignores high severity DLP alerts for all users not on the Sales Team.

- Make sure the **Incident reports** rule setting in the DLP policy used for this Insider Risk Management template is configured for *High* severity level alerts. The *High* severity level is the triggering events and Insider Risk Management alerts aren't generated from rules in DLP policies with the **Incident reports** field set at *Low* or *Medium*.



> ⓘ **Note**

When creating a new DLP policy using the built-in templates, select the **Create or customize advanced DLP rules** option to configure the **Incident reports** setting for the *High* severity level.

Each Insider Risk Management policy created from the **Data leaks** template can only have one DLP policy assigned when using this triggering event option. Consider creating a dedicated DLP policy that combines the different activities you want to detect and act as triggering events for insider risk policies that use the **Data leaks** template.

For step-by-step guidance to configure DLP policies for your organization, see Create and Deploy data loss prevention policies.

Data leaks by priority users

Protecting data and preventing data leaks for users in your organization might depend on their position, level of access to sensitive information, or risk history. Data leaks can include accidental oversharing of highly sensitive information outside your organization or data theft with malicious intent. With an assigned data loss prevention (DLP) policy as a triggering event option, this template starts scoring real-time detections of suspicious activity and results in an increased likelihood of insider risk alerts and alerts with higher severity levels. Priority users are defined in priority user groups configured in the Insider Risk Management settings area.

As with the **Data leaks** template, you can choose a DLP policy to trigger indicators in the insider risk policy for high severity alerts in your organization. Follow the Data leaks policy guidelines for DLP policies when creating a policy with the DLP option when using this template. You can also choose to assign selected indicators as triggering events for a policy. This flexibility and customization help scope the policy to only the activities covered by the indicators. Additionally, you need to assign priority user groups created in **Insider Risk Management** > **Settings** > **Priority user groups** to the policy.

Data leaks by risky users

When users experience employment stressors, they might become risky users, which might increase the chances of insider risk activity. This template starts scoring user activity when an indicator associated with risky user is identified. Examples include performance improvement notifications, poor performance reviews, changes to job level status, or email and other messages that might signal risk activities. Data leaks for risky users might include downloading files from SharePoint Online and copying data to personal cloud messaging and storage services.

When using this template, you must either configure an HR connector, select the option to integrate Communication Compliance risk signals from user messages, or do both. The HR connector enables the periodic import of performance improvement notifications, poor performance review statuses, or job level change information for users in your organization. Communication Compliance risk integration imports signals for user messages that might contain potentially threatening, harassing, or discriminatory text content. Associated alerts generated in Communication Compliance don't need to be triaged, remediated, or changed in status to be integrated with the Insider Risk Management policy.

To configure an HR connector, see the Import data with the HR connector article. To configure integration with Communication Compliance, select this option in the workflow when you configure the policy.

Patient data misuse (preview)

Protecting healthcare record data and preventing the misuse of patient personal data is a significant concern for organizations in the healthcare industry. This misuse might include confidential data leaks to unauthorized persons, fraudulent modification of patient records, or the theft of patient healthcare records. Preventing this misuse of patient data helps meet the regulatory requirements of the Health Insurance Portability and Accountability Act (HIPAA) and the Health Information Technology for Economic and Clinical Health (HITECH) Act. Both of these acts establish the requirements for safeguarding patient protected health information (PHI).

This policy template enables risk scoring for internal users that detects suspicious activities associated with records hosted on existing electronic medical record (EMR) systems. Detection focuses on unauthorized access, viewing, modification, and export of patient data. You need to configure a connector the Microsoft Healthcare connector to support detection of access, exfiltration, or obfuscation activities in your EMR system.

When using this template, you must also configure a Microsoft HR connector to periodically import organization profile data for users in your organization. See the Set up a connector to import HR data article for step-by-step guidance to configure the Microsoft 365 HR connector.

Risky Agents (preview)

This template helps you detect activities of agents hosted on Microsoft Copilot Studio and Microsoft Foundry. The use of agents within your organization introduces new risk opportunities. Agents interact with other users, other agents, and tools. Misconfigured agents could send sensitive data externally, or sensitive data could be shared at a user's request.

This policy detects risky prompts, agents generating sensitive responses, agents accessing sensitive or priority SharePoint files, agents accessing risky websites, and agents sharing SharePoint files with people outside of the organization. It also tracks activity above the agent's baseline.

This policy template is applied by default for all organizations.

Risky AI usage

The use of AI tools and applications in many organizations presents new risk opportunities that might be difficult to identify and help mitigate. Users might inadvertently or purposefully create prompts in Microsoft AI tools that might contain sensitive information or risky intent. Additionally, AI tools might inappropriately use sensitive information to provide responses to user prompts that might expose your organization to increased risks.

To help protect against these risks, this policy can help detect and enable risk scoring for these types of prompts and responses across AI tools in your organization. Detection focuses on user browsing activities to generative AI websites, user prompts and AI responses containing sensitive information in Microsoft 365 Copilot, Microsoft Copilot, and agents. Detection of risky activity by this policy also contributes to the user risk scoring in Adaptive Protection.

When using this template, you must configure the Microsoft Insider risk extension (on Microsoft Edge browser) or Microsoft Purview extension (on Chrome browser) on the user devices. We also recommend optionally configuring a Microsoft HR connector to detect departing users and configuring a Communication Compliance policy to detect inappropriate content in messages for wider detection coverage. This policy can also be created as a one-click policy in Microsoft Purview Data Security Posture Management (DSPM) for AI when you take action on the recommendation **Detect risky interactions in AI apps**.

Risky browser usage (preview)

Identifying user visits to potentially inappropriate or unacceptable websites on organization devices and networks is an important part of minimizing security, legal, and regulatory risks. Users who inadvertently or purposefully visit these types of websites might expose the organization to legal actions from other users, violate regulatory requirements, elevate network security risks, or jeopardize current and future business operations and opportunities. This misuse is often defined in an organization's acceptable use policy for user devices and organization network resources but is often difficult to quickly identify and act on.

To help protect against these risks, this policy detects and enables risk scoring for web browsing that might violate your organization's acceptable use policy, such as visiting sites that pose a threat (phishing sites, for example) or that contain adult content. Several types of

categories are available for automatic categorization of web browsing activities by in-scope users.

This policy template has several prerequisites. For more information, see Learn about and configure Insider Risk Management browser signal detection.

Security policy violations

In many organizations, users have permission to install software on their devices or to modify device settings to help with their tasks. Either inadvertently or with malicious intent, users might install malware or disable important security features that help protect information on their device or on your network resources. This policy template uses security alerts from Microsoft Defender for Endpoint to start scoring these activities and focus detection and alerts to this risk area. Use this template to provide insights for security policy violations in scenarios when users might have a history of security policy violations that might be an indicator of insider risk.

You need to configure Microsoft Defender for Endpoint in your organization and enable Defender for Endpoint for Insider Risk Management integration in the Defender Security Center to import security violation alerts. For more information on configuring Defender for Endpoint for Insider Risk Management integration, see Configure advanced features in Defender for Endpoint.

Security policy violations by departing users

Departing users, whether they're leaving on positive or negative terms, might pose higher risks for security policy violations. To help protect against inadvertent or malicious security violations for departing users, this policy template uses Defender for Endpoint alerts to provide insights into security-related activities. These activities include the user installing malware or other potentially harmful applications and disabling security features on their devices. When you use the Microsoft HR connector or the option to automatically check for user account deletion in Microsoft Entra for your organization, this template starts scoring for risk indicators relating to these security activities and how they correlate with user employment status.

You need to configure Microsoft Defender for Endpoint in your organization and enable Defender for Endpoint for Insider Risk Management integration in the Defender Security Center to import security violation alerts. For more information on configuring Defender for Endpoint for Insider Risk Management integration, see Configure advanced features in Defender for Endpoint.

Security policy violations by priority users

Protecting against security violations for users in your organization might depend on their position, level of access to sensitive information, or risk history. Because security violations by priority users might have a significant impact on your organization's critical areas, this policy template starts scoring on these indicators, and uses Microsoft Defender for Endpoint alerts to provide insights into security-related activities for these users. These activities might include the priority users installing malware or other potentially harmful applications and disabling security features on their devices. Priority users are defined in priority user groups configured in the Insider Risk Management settings area.

You need to configure Microsoft Defender for Endpoint in your organization and enable Defender for Endpoint for Insider Risk Management integration in the Defender Security Center to import security violation alerts. For more information on configuring Defender for Endpoint for Insider Risk Management integration, see Configure advanced features in Defender for Endpoint. Additionally, you need to assign priority user groups created in **Insider Risk Management** > **Settings** > **Priority user groups** to the policy.

Security policy violations by risky users

Users who experience employment stressors might be at higher risk for inadvertent or malicious security policy violations. These stressors could include being placed on a performance improvement plan, receiving a poor performance review, or experiencing a demotion. This policy template starts risk scoring based on these indicators and activities associated with these types of events.

When you use this template, you must configure an HR connector, or select the option to integrate Communication Compliance risk signals from user messages, or both. The HR connector enables the periodic import of performance improvement notifications, poor performance review statuses, or job level change information for users in your organization. Communication Compliance risk integration imports signals for user messages that might contain potentially threatening, harassing, or discriminatory text content. Associated alerts generated in Communication Compliance don't need to be triaged, remediated, or changed in status to be integrated with the Insider Risk Management policy. To configure an HR connector, see the Import data with the HR connector article. To configure integration with Communication Compliance, select this option in the workflow when you configure the policy.

You also need to configure Microsoft Defender for Endpoint in your organization and enable Defender for Endpoint for Insider Risk Management integration in the Defender Security Center to import security violation alerts. For more information on configuring Defender for Endpoint for Insider Risk Management integration, see Configure advanced features in Defender for Endpoint.

Policy template prerequisites and triggering events

Depending on the template you choose for an Insider Risk Management policy, the triggering events and policy prerequisites vary. Triggering events are prerequisites that determine if a user is active for an Insider Risk Management policy. If you add a user to an Insider Risk Management policy but the user doesn't have a triggering event, the policy doesn't evaluate the user's activity unless you manually add the user in the Users dashboard. Policy prerequisites are required items so that the policy receives the signals or activities necessary to evaluate risk.

The following table lists the triggering events and prerequisites for policies created from each Insider Risk Management policy template:

Expand table

Policy template	Triggering events for policies	Prerequisites
Data leaks by priority users	Data leak policy activity that creates a *High severity* alert or built-in exfiltration event triggers	DLP policy configured for *High severity* alerts OR Customized triggering indicators Priority user groups configured in insider risk settings
Data leaks by risky users	- Performance improvement, poor performance, or job level change indicators from HR connector. - Messages containing potentially threatening, harassing, or discriminatory language	Microsoft 365 HR connector configured for disgruntlement indicators AND/OR Communication Compliance integration and dedicated disgruntlement policy
Data leaks	Data leak policy activity that creates a *High severity* alert or built-in exfiltration event triggers	DLP policy configured for *High severity* alerts OR Customized triggering indicators
Data theft by departing users	Resignation or termination date indicator from HR connector or Microsoft Entra account deletion	(optional) Microsoft 365 HR connector configured for termination and resignation date indicators
Patient data misuse	Defense evasion of security controls from EMR systems User and patient address matching indicators from HR systems	Healthcare access indicators selected in policy or insider risk settings Microsoft 365 HR connector configured for address matching

Policy template	Triggering events for policies	Prerequisites
		Microsoft Healthcare or Epic connector configured
Risky Agents	- Exposing agent to risky prompts - Agent generating sensitive responses - Accessing sensitive or priority SharePoint files - Accessing risky websites - Sharing SharePoint files with people outside the organization - Activity is above agent's usual activity for the day	Copilot Studio agents Microsoft Foundry agents
Risky AI usage		Microsoft Insider risk extension (on Microsoft Edge browser) or Microsoft Purview extension (on Chrome browser) must be installed on the user's devices At least one browsing indicator selected in the policy AND (optional) Microsoft 365 HR connector configured for termination and resignation date indicators (optional) A Communication Compliance policy detecting inappropriate content in messages
Risky browser usage	User browsing activity related to security that matches at least one selected *Browsing indicator*	See the complete list of prerequisites in the browser signal detection article
Security policy violations by departing users	Resignation or termination date indicators from HR connector or Microsoft Entra account deletion	(optional) Microsoft 365 HR connector configured for termination and resignation date indicators Active Microsoft Defender for Endpoint subscription Configuration of Microsoft Defender for Endpoint integration with Microsoft Purview portal
Security policy violations by	Defense evasion of security controls or unwanted software	Active Microsoft Defender for Endpoint subscription

Policy template	Triggering events for policies	Prerequisites
priority users	detected by Microsoft Defender for Endpoint	Configuration of Microsoft Defender for Endpoint integration with Microsoft Purview portal Priority user groups configured in insider risk settings
Security policy violations by risky users	- Performance improvement, poor performance, or job level change indicators from HR connector. - Messages containing potentially threatening, harassing, or discriminatory language	Microsoft 365 HR connector configured for risk indicators AND/OR Communication Compliance integration and dedicated risky user policy AND Active Microsoft Defender for Endpoint subscription Configuration of Microsoft Defender for Endpoint integration with Microsoft Purview portal
Security policy violations	Defense evasion of security controls or unwanted software detected by Microsoft Defender for Endpoint	Active Microsoft Defender for Endpoint subscription Configuration of Microsoft Defender for Endpoint integration with Microsoft Purview portal

Policy template limits

Insider Risk Management policy templates use limits to manage the volume and rate of processing for in-scope user risk activities and for integration with supporting Microsoft 365 services. Each policy template supports a maximum number of users who can actively receive risk scores for the policy. The policy can effectively process and report potentially risky activities. In-scope users are users with triggering events for the policy.

The limit for each policy calculates the total number of unique users receiving risk scores per policy template type. If the number of users for a policy template type nears or exceeds the user limit, the policy performance reduces. To view the current number of users for a policy, go to the Policy tab and the Users in scope column. These maximum limits apply to users across all policies using a given policy template.

For user scope limits for specific policy templates, see Limits in Insider Risk Management.

Integration with collection policies

Depending on your configuration, collection policies might impact Insider Risk Management policy behavior. When you configure and deploy a collection policy, the collection policy configuration takes precedence if there's a mismatch between an Insider Risk Managment policy that includes device indicators and a collection policy in your organization. This configuration means that if you configure an Insider Risk Management policy to monitor a specific activity for devices, but you configure the collection policy to filter out that device activity, the device activity isn't collected and isn't available for review in Insider Risk Management.

Last updated on 11/21/2025

Create and manage Insider Risk Management policies

> **Important**
>
> Microsoft Purview Insider Risk Management correlates various signals to identify potential malicious or inadvertent insider risks, such as IP theft, data leakage, and security violations. Insider Risk Management enables customers to create policies to manage security and compliance. Built with privacy by design, users are pseudonymized by default, and role-based access controls and audit logs are in place to help ensure user-level privacy.

Insider Risk Management policies determine which users are in scope and which types of risk indicators are configured for alerts. You can quickly create a security policy that applies to all users in your organization or define individual users or groups for management in a policy. Policies support content priorities to focus policy conditions on multiple or specific Microsoft Teams, SharePoint sites, data sensitivity types, and data labels. By using templates, you can select specific risk indicators and customize event thresholds for policy indicators, effectively customizing risk scores, and level and frequency of alerts.

You can also configure quick data leak and data theft policies by departing user policies that automatically define policy conditions based on results from the latest analytics. Also, risk score boosters and anomaly detections help identify potentially risky user activity that is of higher importance or unusual. Policy windows allow you to define the time frame to apply the policy to alert activities and are used to determine the duration of the policy once activated.

For an overview of how policies created with built-in policy templates can help you quickly act on potential risks, see the Insider Risk Management Policies Configuration video.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Policy dashboard

The **Policy dashboard** lets you quickly see the user and agent policies in your organization, the health of each policy, manually add users to security policies, and view the status of alerts associated with each policy. Your policies are separated into two categories: **User policy** and **Agent policy**. Switch between the two to locate relevant policy details. **Agent policy** covers agents from Microsoft Copilot Studio and Microsoft Foundry.

- **Policy name**: Name assigned to the policy in the policy workflow.
- **Status**: Health status for each policy. Displays number of policy warnings and recommendations, or a status of *Healthy* for policies without issues. You can select the policy to see the health status details for any warnings or recommendations.
- **Active alerts**: Number of active alerts for each policy.
- **Confirmed alerts**: Total number of alerts that resulted in cases from the policy in the last 365 days.
- **Actions taken on alerts**: Total number of alerts that were confirmed or dismissed for the last 365 days.
- **Policy alert effectiveness**: Percentage determined by total confirmed alerts divided by total actions taken on alerts (which is the sum of alerts that were confirmed or dismissed over the past year).



Policy recommendations from analytics

Insider risk analytics gives you an aggregate view of anonymized user activities related to security and compliance. With this view, you can evaluate potential insider risks in your organization without configuring any insider risk policies. This evaluation helps your organization identify potential areas of higher risk and helps determine the type and scope of Insider Risk Management policies you might consider configuring. If you decide to act on analytics scan results for data leaks or data theft by departing users policies, you can even configure a quick policy based on these results.

For more information about insider risk analytics and policy recommendations, see Insider Risk Management settings: Analytics.

> **Note**
>
> You must be an unrestricted administrator to access analytics insights. Learn how administrative groups affect permissions.

Quick policies

For many organizations, getting started with an initial policy can be a challenge. If you're new to Insider Risk Management or are using the recommended actions to get started, use a quick policy to create and configure a new policy. Quick policy settings automatically populate from recommended best practices or on results from the latest analytics scan in your organization. For example, if the analytics check detected potential data leak activities in your organization, the *Date leaks* quick policy automatically includes the indicators used to detect those activities.

You can choose from the following quick policies:

- **Critical assets protection**: Detects activities involving your organization's most valuable assets. Loss of these assets could result in legal liability, financial loss, or reputational damage.
- **Data leaks**: Detect potential data leaks from all users in your organization, which can range from accidental oversharing of sensitive info to data theft with malicious intent.
- **Data theft from Microsoft 365 apps by users leaving your organization**: Detects potential data theft from Microsoft 365 cloud apps by users leaving your organization or whose account was deleted from Microsoft Entra ID.
- **Data theft from non-Microsoft 365 apps by users leaving your organization**: Detects potential data theft from non-Microsoft 365 cloud apps, including Microsoft Fabric, by users leaving your organization or whose account was deleted from Microsoft Entra ID.
- **Email exfiltration**: Detects when users email sensitive assets outside your organization. For example, users emailing sensitive assets to their personal email address.

To get started, go to **Insider Risk Management** > **Policies** and select **Create policy** > **Quick policy**. If you're reviewing analytics reports, select **View details** > **Get started** to get started with a quick policy for the applicable area.

When you start the quick policy workflow, review the policy settings and configure the policy with a single selection. If you need to customize a quick policy, change the conditions after the policy

is created. Stay up to date with the detection results for a quick policy by configuring email notifications each time you have a policy warning or each time the policy generates a high severity alert.

> **Note**
>
> You must be an unrestricted administrator to create quick policies. Learn how administrative groups affect permissions.

Prioritize content in policies

Insider Risk Management policies support specifying a higher priority for content depending on where you store it, the type of content, or how you classify it. You can also choose whether to assign risk scores to all activities detected by a policy or only activities that include priority content. Specifying content as a priority increases the risk score for any associated activity, which in turn increases the chance of generating a high severity alert. However, some activities don't generate an alert at all unless the related content contains built-in or custom sensitive info types or you specify it as a priority in the policy.

For example, your organization has a dedicated SharePoint site for a highly confidential project. Data leaks for information in this SharePoint site could compromise the project and would have a significant impact on its success. By prioritizing this SharePoint site in a Data leaks policy, you automatically increase risk scores for qualifying activities. This prioritization increases the likelihood that these activities generate an insider risk alert and raises the severity level for the alert.

Additionally, you can choose to focus this policy for SharePoint site activity that only includes priority content for this project. Risk scores are assigned and alerts are generated only when specified activities include priority content. Activities without priority content aren't scored, but you can still review them if an alert is generated.

> **Important**
>
> Cumulative exfiltration activity always receives a risk score, regardless of whether or not it contains prioritized content.

The following policy templates support priority content selection:

- Data leaks
- Data leaks by priority users
- Data leaks by risky users
- Data theft by departing users
- Risky AI usage

> **Note**
>
> If you configure a policy to generate alerts only for activity that includes priority content, no changes are applied to risk score boosters.

When you create an Insider Risk Management policy in the policy workflow, you can choose from the following priorities:

- **SharePoint sites**: Assign a higher risk score to any activity associated with all file types in defined SharePoint sites. Users configuring the policy and selecting priority SharePoint sites can select SharePoint sites that they have permission to access. If SharePoint sites aren't available for selection in the policy by the current user, another user with the required permissions can select the sites for the policy later, or the current user should be given access to the required sites.
- **Sensitive information types**: Assign a higher risk score to any activity associated with content that contains sensitive information types.
- **Sensitivity labels**: Assign a higher risk score to any activity associated with content that has specific sensitivity labels applied.
- **File extensions**: Assign a higher risk score to any activity associated with content that has specific file extensions. Users configuring a data theft/leak policy that selects **File extensions to prioritize** in the policy workflow can define up to 50 file extensions to prioritize in the policy. Entered extensions can include or omit a '.' as the first character of the prioritized extension.
- **Trainable classifiers**: Assign a higher risk score to any activity associated with content that is included in a trainable classifier. Users configuring a policy that selects Trainable classifiers in the policy workflow can select up to 5 trainable classifiers to apply to the policy. These classifiers can be existing classifiers that identify patterns of sensitive information like social security, credit card, or bank account numbers or custom classifiers created in your organization.

Sequence detection

Risk management activities might not occur as isolated events. These risks are frequently part of a larger sequence of events. A sequence is a group of two or more potentially risky activities performed one after the other that might suggest an elevated risk. Identifying these related user activities is an important part of evaluating overall risk. When you select sequence detection for data theft or data leaks policies, you see insights from sequence information activities on the **User activity** tab within an Insider Risk Management case. The following policy templates support sequence detection:

- Data leaks
- Data leaks by priority users
- Data leaks by risky users
- Data theft by departing users
- Risky AI usage

These Insider Risk Management policies can use specific indicators and the order that they occur to detect each step in a sequence of risk. Selected sequence detections only detect sequences of risk in the order displayed in policy setup workflow. For policies created from the *Data leaks* and *Data leaks by priority user* templates, you can also select which sequences trigger the policy. File names are used when mapping activities across a sequence. These risks are organized into four main categories of activity:

- **Collection**: Detects download activities by in-scope policy users. Example risk management activities include downloading files from SharePoint sites, third-party cloud services, unallowed domains, or moving files into a compressed folder.
- **Exfiltration**: Detects sharing or extraction activities to internal and external sources by in-scope policy users. An example risk management activity includes sending emails with attachments from your organization to external recipients.
- **Obfuscation**: Detects the masking of potentially risky activities by in-scope policy users. An example risk management activity includes renaming files on a device.
- **Clean-up**: Detects deletion activities by in-scope policy users. An example risk management activity includes deleting files from a device.

> ⓘ **Note**
>
> Sequence detection uses indicators that you enable in the global settings for Insider Risk Management. If you don't select appropriate indicators, you can turn on these indicators in the sequence detection step in the policy workflow.

You can customize individual threshold settings for each sequence detection type when you configure it in the policy. These threshold settings adjust alerts based on the volume of files associated with the sequence type.

> **Note**
>
> A **sequence** might contain one or more events that are excluded from risk scoring based on your settings configuration. For example, your organization might use the **Global exclusions setting** to exclude .png files from risk scoring since .png files aren't normally risky. But a .png file could be used to obfuscate a malicious activity. For this reason, if an event that's excluded from risk scoring is part of a sequence due to an obfuscation activity, the event is included in the sequence since it might be interesting in the context of the sequence. **Learn more about how exclusions that are part of a sequence are shown in the Activity explorer**.

To learn more about sequence detection management in the **User activity** view, see Insider Risk Management cases: User activity.

Cumulative exfiltration detection

With privacy on by default, insider risk indicators help identify unusual levels of risk activities when evaluated daily for users that are in-scope for insider risk policies. Cumulative exfiltration detection uses machine learning models to help you identify when exfiltration activities that a user performs over a certain time exceed the normal amount performed by users in your organization for the past 30 days over multiple exfiltration activity types. For example, if a user shared more files than most users over the past month, this activity is detected and classified as a cumulative exfiltration activity.

Insider Risk Management analysts and investigators might use cumulative exfiltration detection insights to help identify exfiltration activities that might not typically generate alerts but are more than what is typical for their organization. Some examples might be departing users slowly exfiltrate data across a range of days, or when users repeatedly share data across multiple channels more than usual for data sharing for your organization, or compared to their peer groups.

> **Note**
>
> By default, cumulative exfiltration detection generates risk scores based on a user's cumulative exfiltration activity compared to their organization norms. You can enable

Cumulative exfiltration detection options in the **Policy indicators** section of the Insider Risk Management settings page.

Higher risk scores are assigned to cumulative exfiltration activities for SharePoint sites, sensitive information types, and content with sensitivity labels configured as priority content in a policy or for activity involving labels configured as high priority in Microsoft Purview Information Protection.

Cumulative exfiltration detection is enabled by default when using the following policy templates:

- Data leaks
- Data leaks by priority users
- Data leaks by risky users
- Data theft by departing users

> **Important**
>
> The option to only score activities containing priority content doesn't apply to cumulative exfiltration activities. Cumulative exfiltration activity always receives a risk score, regardless of whether or not it contains prioritized content.

Peer groups for cumulative exfiltration detection

Insider Risk Management identifies three types of peer groups for analyzing exfiltration activity performed by users. It defines peer groups for users based on the following criteria:

SharePoint sites: Insider Risk Management identifies peer groups based on users who access similar SharePoint sites. **Similar organization**: Users with reports and team members based on organization hierarchy. This option requires that your organization uses Microsoft Entra ID to maintain organization hierarchy. **Similar job title**: Users with a combination of organizational distance and similar job titles. For example, a user with a Senior Sales Manager title with a similar role designation as a Lead Sales Manager in the same organization is identified as similar job title. This option requires that your organization uses Microsoft Entra ID to maintain organization hierarchy, role designations, and job titles. If you don't have Microsoft Entra ID configured for organization structure and job titles, then Insider Risk Management identifies peer groups based on common SharePoint sites.

When you enable cumulative exfiltration detection, your organization agrees to share Microsoft Entra data with the Microsoft Purview portal, including organization hierarchy and job titles. If

your organization doesn't use Microsoft Entra ID to maintain this information, detection might be less accurate.

> **Note**
>
> Cumulative exfiltration detection uses exfiltration indicators that you enable in the global settings for Insider Risk Management and exfiltration indicators that you select in a policy. As such, cumulative exfiltration detection only evaluates the necessary exfiltration indicators. Cumulative exfiltration activities for sensitivity labels configured in priority content generate higher risk scores.

When you enable cumulative exfiltration detection for data theft or data leak policies, insights from cumulative exfiltration activities appear on the **User activity** tab within an Insider Risk Management case. For more information about user activity management, see Insider Risk Management cases: User activities.

Policy health

> **Note**
>
> If your policy is scoped by one or more administrative units, you can only view policy health for the policies you're scoped for. If you're an unrestricted administrator, you can view policy health for all policies in the tenant.

The policy health status gives you insights into potential issues with your Insider Risk Management policies. The **Status** column on the **Policies** tab alerts you to policy issues that might prevent user activity from being reported or explain why the number of activity alerts is unusual. The policy health status also confirms that the policy is healthy and doesn't need attention or configuration changes.

> **Important**
>
> You must have the *Insider Risk Management* or the *Insider Risk Management Admins* role to access policy health.

If there are issues with a policy, the policy health status displays notification warnings and recommendations to help you take action to resolve policy issues. These notifications can help

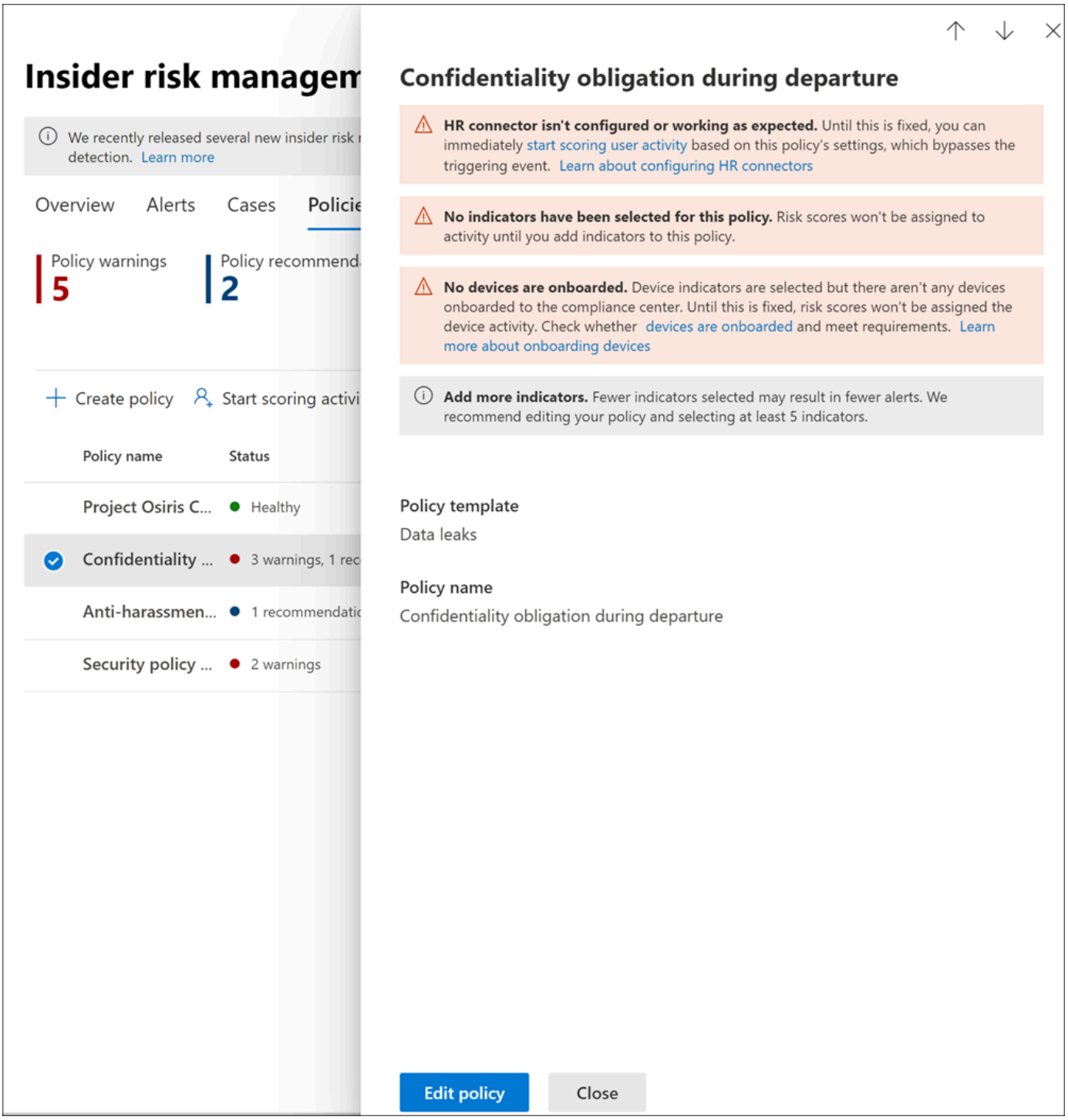
you resolve the following issues:

- **Policies with incomplete configuration**. These issues might include missing users or groups in the policy or other incomplete policy configuration steps.
- **Policies with indicator configuration issues**. Indicators are an important part of each policy. If you don't configure indicators, or if you select too few indicators, the policy might not evaluate risky activities as expected.
- **Policy triggers aren't working, or policy trigger requirements aren't properly configured**. Policy functionality might depend on other services or configuration requirements to effectively detect triggering events to activate risk score assignment to users in the policy. These dependencies might include issues with connector configuration, Microsoft Defender for Endpoint alert sharing, or data loss prevention policy configuration settings.
- **Volume limits are nearing or over limits**. Insider Risk Management policies use numerous Microsoft 365 services and endpoints to aggregate risk activity signals. Depending on the number of users in your policies, volume limits might delay identification and reporting of risk activities. Learn more about these limits in the Policy template limits section.

To quickly view the health status for a policy, go to the **Policy** tab and check the **Status** column. You see the following policy health status options for each policy:

- *Healthy*: No issues are identified with the policy.
- *Recommendations*: An issue with the policy that might prevent the policy from operating as expected.
- *Warnings*: An issue with the policy that might prevent it from identifying potentially risky activities.

For more details about any recommendations or warnings, select a policy on the **Policy** tab to open the policy details card. The **Notifications** section of the details card displays more information about the recommendations and warnings, including guidance on how to address these issues.



Notification messages

Use the following table to learn more about recommendations and warning notifications and actions to take to resolve potential issues.

Expand table

Notification messages	Policy templates	Causes / Try this action to fix
DLP policy doesn't meet requirements	- Data leaks - Data leaks by priority users	DLP policies used as triggering events must be configured to generate high severity alerts.

Notification messages	Policy templates	Causes / Try this action to fix
		1. Edit your DLP policy to assign applicable alerts as *High severity*. OR 2. Edit this policy and select *User performs an exfiltration activity* as the triggering event.
DLP policy isn't selected as the triggering event	- Data leaks - Data leaks by priority users	A DLP policy wasn't selected as a triggering event or the selected DLP policy was deleted. Edit the policy and either select an active DLP policy or 'User performs an exfiltration activity' as the triggering event in the policy configuration.
DLP policy used in this policy is turned off	- Data leaks - Data leaks by priority users	DLP policy used in this policy is turned off. 1. Turn the DLP policy assigned to this policy on. OR 2. Edit this policy and either select a new DLP policy or 'User performs an exfiltration activity' as the triggering event in the policy configuration.
HR connector hasn't uploaded data recently	- Data theft by departing user - Security policy violations by departing user - Data leaks by risky users - Security policy violations by risky users	HR connector didn't import data in more than 7 days. Check that your HR connector is configured correctly and sending data.
HR connector isn't configured or working as expected	- Data theft by departing user - Security policy violations by departing user - Data leaks by risky users - Security policy violations by risky users	There's an issue with the HR connector. 1. If you're using an HR connector, check that your HR connector is sending correct data OR 2. Select the Microsoft Entra account deleted triggering event.
Microsoft Defender for Endpoint alerts aren't being	- Security policy violations	Microsoft Defender for Endpoint alerts aren't being shared with the Microsoft Purview portal.

Notification messages	Policy templates	Causes / Try this action to fix
shared with the Microsoft Purview portal	- Security policy violations by departing users - Security policy violations by risky users - Security policy violations by priority users	Configure sharing of Microsoft Defender for Endpoint alerts.
No devices are onboarded	- Data theft by departing users - Data leaks - Data leaks by risky users - Data Leaks by priority users	Device indicators are selected but there aren't any devices onboarded to the Microsoft Purview portal Check whether devices are onboarded and meet requirements.
No indicators have been selected for this policy	All policy templates	Indicators weren't selected for the policy Edit your policy and select appropriate policy indicators for the policy.
No priority user groups are included in this policy	- Data leaks by priority users - Security policy violations by priority users	Priority user groups aren't assigned to the policy. Configure priority user groups in Insider Risk Management settings and assign priority user groups to the policy.
No triggering event has been selected for this policy	All policy templates	A triggering event isn't configured for the policy Risk scores won't be assigned to user activities until you edit the policy and select a triggering event.
No users or groups are included in this policy	All policy templates	Users or groups aren't assigned to the policy. Edit your policy and select users or groups for the policy.
Policy hasn't generated any alerts	All policy templates	You might want to review your policy configuration so that you're analyzing the most relevant scoring activity. 1. Confirm that you selected indicators that you want to score. The more indicators selected, the more activities are assigned risk scores. 2. Review threshold customization for policy. If the thresholds selected don't align with your organization's risk tolerance, adjust the selections so

Notification messages	Policy templates	Causes / Try this action to fix
		that alerts are created based on your preferred thresholds. 3. Review the users and groups selected for the policy. Confirm you selected all of the applicable users and groups. 4. For security violation policies, confirm you selected the alert triage status that you want to score for Microsoft Defender for Endpoint alerts in Intelligent Detections in settings.
Policy isn't assigning risk scores to activity	All policy templates	You might want to review your policy scope and triggering event configuration so that the policy can assign risk scores to activities 1. Review the users that are selected for the policy. If you have few users selected, you might want to select additional users. 2. If you're using an HR connector, check that your HR connector is sending the correct data. 3. If you're using a DLP policy as your triggering event, check your DLP policy configuration to ensure it's configured to be used in this policy. 4. For security violation policies, review the Microsoft Defender for Endpoint alert triage status selected in Insider risk settings > Intelligent detections. Confirm that the alert filter isn't too narrow.
Triggering event is repeatedly occurring for over 15% of users in this policy	All policy templates	Adjust the triggering event to help reduce how often users are brought into the policy scope.
We're unable to check the status of your HR connector right now, please check again later	- Data theft by departing user - Security policy violations by departing user - Data leaks by risky users - Security policy violations by risky users	The Insider Risk Management solution is unable to check the status of your HR connector. Check that your HR connector is configured correctly and sending data, or come back and check the policy status.
You're approaching the maximum limit of users being actively scored for this policy template	All policy templates	Each policy template has a maximum number of included users. See the Policy template limits section for details. Review the users in the Users tab and remove any users who don't need to be scored anymore.

Notification messages	Policy templates	Causes / Try this action to fix
Your organization doesn't have a Microsoft Defender for Endpoint subscription	- Security policy violations - Security policy violations by departing users - Security policy violations by risky users - Security policy violations by priority users	An active Microsoft Defender for Endpoint subscription wasn't detected for your organization. Until a Microsoft Defender for Endpoint subscription is added, these policies won't assign risk scores to user activity.

Create a new policy

To create a new Insider Risk Management policy, generally use the policy workflow in the **Insider Risk Management** solution in the Microsoft Purview portal. You can also create quick policies for general data leaks and data theft by departing users from Analytics checks if applicable.

Complete Create an Insider Risk Management policy to configure new insider risk policies.

Update a policy

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.
2. Go to the **Insider Risk Management** solution.
3. Select **Policies** in the left navigation.
4. On the policy dashboard, select the policy you want to update.
5. On the policy details page, select **Edit policy**.
6. On the **Name and description** page, update the description for the policy if you want.

   > **Note**
   >
   > You can't edit the **Policy template** or **Name** field.

7. Select **Next** to continue.
8. Follow the remaining steps to configure your Insider Risk Management policy.

Copy a policy

You might need to create a new policy that's similar to an existing policy but needs just a few configuration changes. Instead of creating a new policy from scratch, you can copy an existing policy and then modify the areas that need to be updated in the new policy.

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.
2. Go to the **Insider Risk Management** solution.
3. Select **Policies** in the left navigation.
4. On the policy dashboard, select the policy you want to copy.
5. On the policy details page, select **Copy**.
6. In the policy workflow, name the new policy, and then update the policy configuration as needed.

Immediately start scoring user activity

Some scenarios require assigning risk scores to users with insider risk policies outside of the Insider Risk Management triggering event workflow. Use **Start scoring activity for users** on the **Policies** tab to manually add one or more users to one or more insider risk policies for a specific amount of time. This action starts assigning risk scores to their activity and bypasses the requirement for a user to have a triggering indicator, like a DLP policy match or an Employment End Date from the HR Connector.

The value in the **Reason for scoring activity** field appears on the users' activity timeline. The **Users** dashboard displays users you manually add to policies, and alerts are created if the activity meets the policy alert thresholds. You can have up to 4,000 users in scope that you manually add by using the **Start scoring activity for users** feature.

Some scenarios where you might want to immediately start scoring user activities include:

- You identify users with risk concerns and want to immediately start assigning risk scores to their activity for one or more of your policies.
- There's an incident that might require you to immediately start assigning risk scores to involved users' activity for one or more of your policies.
- You haven't configured your HR connector yet, but you want to start assigning risk scores to user activities for HR events by uploading a .csv file.

> **Note**

It might take several hours for manually added users to appear in the **Users** dashboard. Activities for the previous 90 days for these users might take up to 24 hours to display. To view activities for manually added users, go to the **Users** tab, select the user on the **Users** dashboard, and then open the **User activity** tab on the details pane.

Manually start scoring activity for users in one or more Insider Risk Management policies

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.
2. Go to the **Insider Risk Management** solution.
3. Select **Policies** in the left navigation.
4. On the policy dashboard, select the policies where you want to add users.
5. Select **Start scoring activity for users**.
6. In the **Add users to multiple policies** pane, enter a reason for adding the users in the **Reason** field.
7. Define the number of days to score the user's activity in the **This should last for (choose between 5 and 30 days)** field.
8. Enter the name of the user you want to add, or use the **Search user to add to policies** field to search for a user, then select the user name. Repeat this process to assign additional users. The list of users you select appears in the users section of the **Add users to multiple policies** pane.

   > **Note**
   >
   > If the policy is scoped by one or more administrative units, you can only see users that you're scoped for.

   Select **Import** to import a .csv (comma-separated values) file to import a list of users. The file must be in the following format and must list the user principal names:

   ```
   csv
   ```

```
user principal name
user1@domain.com
user2@domain.com
```

9. Select **Add users to policies** to accept the changes.

Stop scoring users in a policy

To stop scoring users in a policy, see the Insider Risk Management users: Remove users from in-scope assignment to policies article.

Delete a policy

> **Important**
>
> You can't undo a policy deletion.

When you delete a policy, you have two options. You can:

- Delete just the policy.
- Delete the policy and all associated alerts and users.

If you choose the second option:

- All alerts generated by that policy are deleted unless they're associated with a case. Associated cases are never deleted when you delete a policy.
- Any user associated with an alert from that policy is removed from the **Users** page.
- If a user is in scope of more than one policy, you remove the user only from the policy that you're deleting. You don't remove the user from other active policies.

For example, you might create a policy for test purposes before rolling it out to your organization. After you finish testing, you can quickly delete the policy and all associated test data so that you can start fresh when you're ready to push the policy live.

It can take up to 72 hours to complete a policy deletion.

> **Note**
>
> If you delete a policy associated with Insider Risk Management Adaptive Protection, you see a warning that Adaptive Protection stops assigning insider risk levels to users until you

choose a different policy in Adaptive Protection. This warning appears because Adaptive Protection must be associated with a policy to be in effect.

To delete a policy

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.
2. Go to the **Insider Risk Management** solution.
3. Select **Policies** in the left navigation.
4. On the policy dashboard, select the policy you want to delete.
5. Select **Delete** on the dashboard toolbar.
6. Choose one of the following options:

   - Select **Delete only the policy**.
   - Select **Delete the policy and all associated alerts and users**.

   > **Important**
   >
   > You can't undo a policy deletion.

7. Select **Confirm**.

   You see a message in the upper part of the screen that tells you if the deletion was successful or whether it's pending.

Note: The author created this article with assistance from AI. Learn more

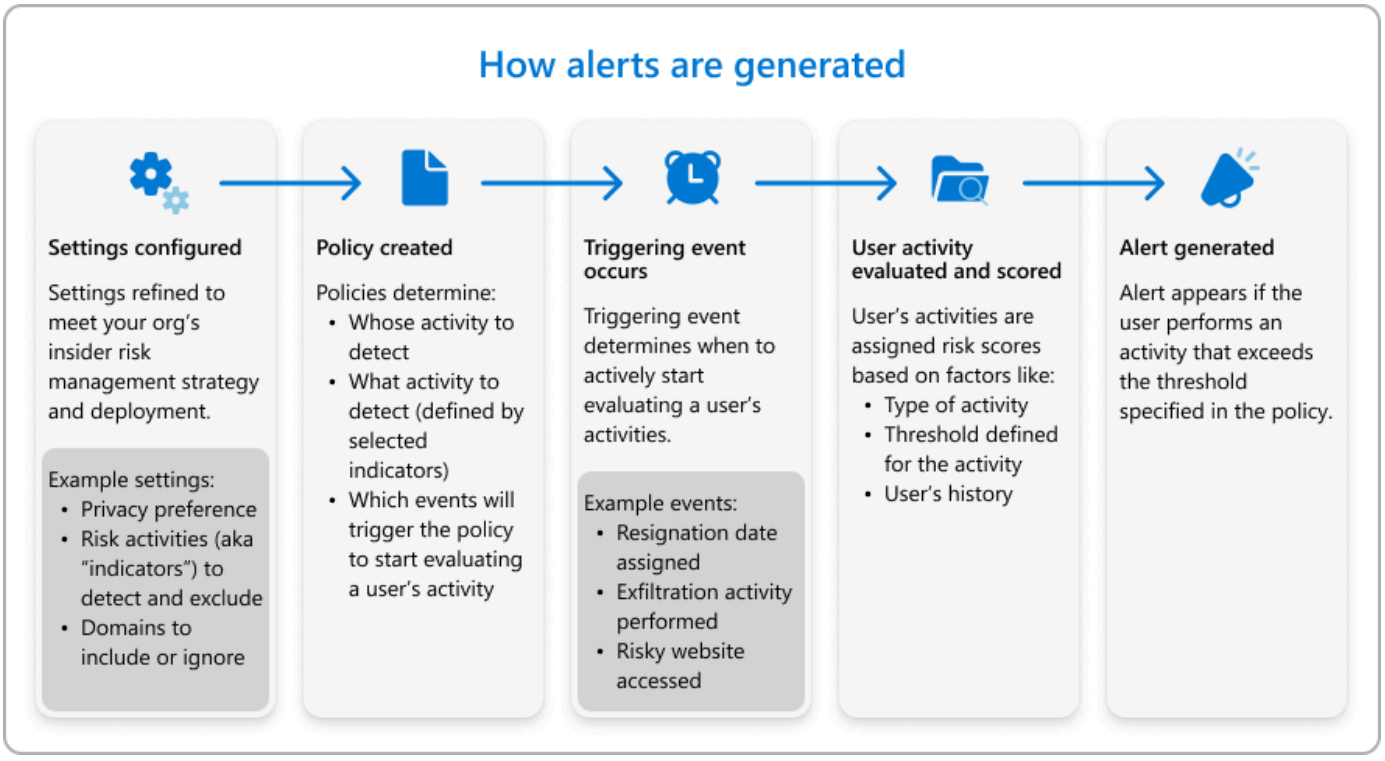
---

Last updated on 06/25/2026

Investigate Microsoft Purview Insider Risk Management activities

> **Important**
>
> Microsoft Purview Insider Risk Management correlates various signals to identify potential malicious or inadvertent insider risks, such as IP theft, data leakage, and security violations. Insider Risk Management enables customers to create policies to manage security and compliance. Built with privacy by design, users are pseudonymized by default, and role-based access controls and audit logs are in place to help ensure user-level privacy.

To minimize insider risks in your organization, investigate potentially risky user activities. These risks might be activities that generate alerts from Insider Risk Management policies. They can also be risks from compliance-related activities that the policies detect but don't immediately create Insider Risk Management alerts for users.



You can investigate these types of activities by using User activity reports or with the **Triage Agent in Insider Risk Management** and **Standard alert** dashboards.

Triaging alerts

Investigate and act on alerts in Insider Risk Management by following these steps:

1. **Review the dashboards for alerts**. On the Standard dashboard, filter alerts by alert *Status* to locate *Needs review* alerts. You can also use spotlighted alerts on the dashboard to quickly see prioritized alerts. On the Alert Triage Agent dashboard, select the **Needs attention** filter to view alerts with the highest prioritization.

2. **Start with the alerts with the highest severity**. Filter alerts by alert *Severity* if needed to help locate these types of alerts.

3. **Select an alert to discover more information and to review the alert details**. Review details and connections associated with the alert:
   a. View the agent summary tab to review intelligently distilled user activity and risk pattern narratives.
   b. Use the Activity explorer tab to review a timeline of the associated potentially risky behavior, and to identify all risk activities for the alert.
   c. Use the Data risk graph to review connections and details about users and files for the alert.

   > **Tip**
   >
   > If **Content preview (preview)** is available for the content within the alert, you can review relevant files from SharePoint, Exchange, and OneDrive for Business in Activity explorer to identify false positives, confirm that sensitive data is present, and quickly decide whether the alert warrants escalation.

4. **Act on the alert**. You can either confirm and create a case for an alert or dismiss and resolve the alert.

You can triage alerts by going to the **Alert details** page for an alert in either dashboard. On the **Alert details** page, you can review information about the alert. You can confirm the alert and create a new case, confirm the alert and add to an existing case, or dismiss the alert.

The **Alert details** page also includes the current status for the alert and the alert risk severity level, listed as *High*, *Medium*, or *Low*. The severity level might increase or decrease over time if you don't triage the alert. For false positives, you can select multiple alerts and dismiss them in bulk by selecting **Dismiss alerts**.

Use Copilot to summarize an alert

Select **Summarize with Copilot** or the Copilot icon to quickly summarize an alert and prioritize the alerts that need further investigation. You can summarize selected alerts without opening the alert or after viewing the details of the alert. When you summarize an alert with Microsoft Copilot in Microsoft Purview, a **Copilot** pane appears on the right side of the screen with an alert summary.

The alert summary includes all the essential details about the alert, such as the policy that triggered the alert, the activity that generated the alert, the triggering event, the user involved, their last working date (if applicable), any key user attributes, and the user's top risk factors. Copilot in Microsoft Purview consolidates information about the user from all their alerts and in-scope policies and emphasizes the user's top risk factors.

Suggested prompts automatically show to help further refine your summary and provide additional insights to the activities associated with the alert. Choose from the following suggested prompts:

- *List all the data exfiltration activities involving this user.*
- *List all the sequential activities involving this user.*
- *Did the user engage in any unusual behavior?*
- *Show key actions performed by the user in the last 10 days.*
- *Summarize user's last 30 days of activity.*

> **Tip**
>
> You can also use the standalone version of Microsoft Security Copilot to investigate Insider Risk Management, Microsoft Purview Data Loss Prevention (DLP), and Microsoft Defender XDR alerts.

Alerts (preview)

The new unified alert experience combines the Triage Agent and classic alert dashboards into a single alerts list page. This unified view lets you manage both classic and agent-triaged alerts from one location, making it easier to investigate and act on alerts without toggling between dashboards.

The unified alert experience includes the following capabilities:

- **Preview agent summaries, alert, and user details directly on the alerts list page** without opening the full alert details.

- **Filter by all classic and agent attributes on one page**, including a filter for agent-triaged alerts that need attention.
- **Manage the Triage Agent directly from the alerts list page**, including viewing agent categorization details.
- **New Agent Categorization column** added to the alerts list to quickly identify how the agent has categorized each alert.

> **Important**
>
> The following changes apply to the unified alert experience:
>
> - Alert **Spotlight** is no longer available in the unified experience. To prioritize agent-triaged alerts, use the **Needs Attention** filter instead.
> - The ability to toggle between the Triage Agent and classic alert dashboards is no longer available. Both alert types appear in a single unified view.
> - The classic and new alert experiences will both be available for at least 60 days. After August 31, only the unified experience will be supported.

To access the unified alert experience, select the **Alerts (preview)** tab in the left navigation of the Insider Risk Management solution.

System-generated notes for alerts (preview)

In the unified alert experience, system-generated notes are automatically added to alerts when there's a change in alert status, a change in the assigned user, or closure of an alert. These notes appear in the **Notes** tab within the alert details panel, helping analysts and investigators stay up to date with all activity on an alert without needing to manually track changes. Some system-generated notes might not appear for activity prior to July 1. For information about system-generated notes for cases, see Case notes.

Alert notes (preview)

Analysts and investigators can add and view notes directly on alerts within the Purview portal.

> **Important**
>
> - You can add up to 50 manual notes per alert or case (not including system-generated notes).

- Notes can't be edited or deleted at this time.
- There's no sync between alert notes in Insider Risk Management and Microsoft Defender.

User profile details (preview)

In the unified alert experience, you can view expanded user profile details directly from an alert. The expanded user profile aggregates all Insider Risk Management user signals into a unified view, making it easier to understand complete user risk at a glance.

To access expanded user profile details:

1. Select **Alerts (preview)** in the left navigation.
2. Select an alert.
3. Scroll down to **User details**.
4. Select **View user details**.

The expanded user profile includes the following additional signals from the user's Microsoft Entra profile:

- **Office location**
- **Employee type**
- **Department**
- **Last working date**
- **Entra account deletion date**

The profile also aggregates the following Insider Risk Management signals:

- **Past alert and case history**
- **Priority user group status**
- **Policy inclusion**
- **HR Connector last working date**

> ⓘ **Note**
>
> When pseudo-anonymization is enabled, the expanded user profile details aren't visible.

For more information about user profile details, see View user details.

Use Spotlight to prioritize alerts (preview)

> **Note**
>
> Alert Spotlight will be removed in the unified alert experience (preview) on August 31, 2026. To prioritize alerts in the new experience, use the **Needs Attention** filter for agent-triaged alerts instead.

The alert **Spotlight** on the **Alerts** dashboard helps you prioritize alerts to triage. Every generated alert has a risk score, a list of activities performed, tags, and triggers. Alert spotlighting uses this information to determine if an alert is spotlighted.

An alert is automatically spotlighted if it has a risk score of 85 or higher and if at least three of the following conditions are met:

- One insight from the following table matches the highest insight of the alert.
- The alert contains priority content or the user is detected as a potential high impact user.
- The user is manually brought into scope.
- Alert activity contains two or more high-confidence insights from the following table.

Expand table

Category	Indicator (high confidence insights)
Device indicator	- Creating or transferring files to a network share - Creating or copying files to USB - Using a browser to upload files to the web
Office indicator	- Sharing SharePoint files with people outside the organization - Sharing SharePoint folders with people outside the organization - Sharing file links with people outside organization in a Teams chat
Policy	- Content to prioritize - Start scoring activity for users
Risk score booster	- User is detected as a potential high impact user
Sequence detection	- Download from Microsoft 365 location then exfiltrate - Download from Microsoft 365 location, exfiltrate, then delete - Download from Microsoft 365 location, obfuscate, then exfiltrate

Triage Agent in Insider Risk Management dashboard

When you enable the Triage Agent in Insider Risk Management in your organization, the agent reviews alerts and displays them on the Triage Agent dashboard. This dashboard automatically includes prioritized alerts and filters to help Insider Risk Management analysts quickly investigate and resolve issues.

You can customize the dashboard columns and filter alerts triaged by the agent by priority, date, alert status, or alert scope. To view the Triage Agent dashboard, select **Triage Agent** at the top of the dashboard page.

Select an alert to display the Agent summary and overview details for the alert. Select **View details** to view details for the alert and for access to the details sections like the risk factors, risk pattern narratives (preview), the Activity explorer, and more.

If you disagree with the **Agent categorization**, select **Is this incorrect?** (preview) to provide feedback on why the categorization is incorrect.

> **Important**
>
> The file risk section of the Triage Agent is deprecated.

For an overview of how alerts provide details, context, and related content for risky activity and how to make your investigation process more effective, see the Insider Risk Management Alerts Triage Experience video.

Understand the Agent summary tab

The **Agent summary** section in the **Alert details** page is only available when selecting an alert from the Triage Agent in Insider Risk Management dashboard. Agent summary information includes details on the categorization for the alert and details about the associated risks used in the triage process. This view provides you with the following information:

- **Agent categorization**: Alerts are categorized as **Needs attention** or **Less urgent**.
- **User info**: Provides the **Alert history**, **Title**, **Organization** and **Last working date** for the user. Select the tile to view additional details (preview).
- **Risk patterns**: Provides a narrative summary of each risk associated with the alert (preview).

Select the narrative summary (preview) for each listed risk to review the following details:

- **Summary**: Each risk pattern includes a narrative summary detailing the risky activity, the sensitive data involved, and the affected files.

- **Actors**: The device involved in the alert and the client IP address, if available.
- **Timeframe**: Provides the start and end dates and times in UTC.
- **Sensitive data involved**: Lists the sensitive information types, classifiers, and file sensitivity labels involved in the alert.
- **Activity names**: The types of actions taken by the user.
- **Activity details**: Explains key aspects of the observed activity, such as the scope of events, sensitive data and files involved, and other contextual signals that contribute to the risk assessment.

Select **Filter activity** for a filtered view of the relevant information in the Activity explorer that contributed to the selected narrative.

> ⓘ **Note**
>
> The Triage Agent does the following:
>
> - Performs an in-depth analysis of the most recent 30,000 activity events associated with the user referenced in the alert.
> - Evaluates all recorded user activities to assess risk, rather than limiting the analysis to only the activities defined in the policy that triggered the alert.
> - Relies on underlying activity log details to determine risk, and when specific information isn't available, it clearly indicates this by marking the detail as "Not found by agent".

Standard alert dashboard

Insider Risk Management policies automatically generate alerts based on risk indicators you define, and display them on the Standard alert dashboard. These alerts give compliance analysts and investigators an all-up view of the current risk status and allow your organization to triage and take actions for discovered potential risks. By default, policies generate a certain number of low, medium, and high severity alerts, but you can increase or decrease the alert volume to suit your needs. Additionally, you can configure the alert threshold for policy indicators when creating a new policy with the policy creation tool. To view the Standard dashboard, select **Standard** at the top of the dashboard page.

> ⓘ **Note**

For any generated alerts, Insider Risk Management generates a single aggregated alert per user. The system adds any new insights for that user to the same alert.

Important

If you scope your policies by one or more administrative units, you can only see alerts for the users you're scoped for. For example, if an administrative scope applies to just users in Germany, you can only see alerts for users in Germany. Unrestricted administrators can see all alerts for all users in the organization.

Restricted administrators can't access alerts for the users assigned to them through security groups or distribution groups added in administrative units. Such user alerts are visible only to unrestricted administrators. Microsoft recommends adding users directly to administrative units to ensure their alerts are also visible to restricted administrators with administrative units assigned.

Filter alerts, save a view of a filter set, customize columns, or search for alerts

Depending on the number and type of active Insider Risk Management policies in your organization, reviewing a large queue of alerts can be challenging. To help you keep track of alerts, you can:

- Filter alerts by various attributes.
- Save a view of a filter set to reuse later.
- Display or hide columns.
- Search for an alert.
- View alert reports.

Filter alerts

To narrow the alerts shown on the dashboard, follow these steps:

1. Select **Add filter**.

2. Select one or more of the following attributes:

Expand table

Attribute	Description
Activity that generated the alert	Displays the top potentially risky activity and policy match during the activity evaluation period that leads to the alert. This value can update over time.
Alert dismissal reason	The reason for dismissing the alert.
Assigned to	The admin that the alert is assigned to for triaging (if assigned).
Policy	The name of the policy.
Risk factors	The risk factors that help determine how risky a user's activity might be. The possible values are *Cumulative exfiltration activities*, *Activities include priority content*, *Sequence activities*, *Activities include unallowed domains*, *Member of a priority user group*, and *Potential high impact user*.
Severity	The user's risk severity level. The options are *High*, *Medium*, and *Low*.
Status	Status of the alert. The options are *Confirmed*, *Dismissed*, *Needs review*, and *Resolved*.
Time detected (UTC)	The start and end dates for when the alert was created. The filter searches for alerts between UTC 00:00 on the start date and UTC 00:00 on the end date.
Triggering event	The event that brought the user into scope of the policy. The triggering event can change over time.

The attributes that you select are added to the filter bar.

3. Select an attribute in the filter bar, and then select a value to filter by. For example, select the **Time detected (UTC)** attribute, enter or select the dates in the **Start date** and **End date** fields, and then select **Apply**.

> **Tip**
>
> To start over at any point, select **Reset all** on the filter bar.

Save a view of a filter set to reuse later

To save the current filter set for later reuse, follow these steps:

1. After applying the filters as described in the preceding procedure, select **Save**, enter a name for the filter set, and then select **Save**.

The filter set is added as a card. It includes a number that shows the count of alerts that meet the criteria in the filter set.

> **Note**
>
> You can save up to five filter sets. To delete a filter set, select the ellipsis (three dots) in the upper-right corner of the card, and then select **Delete**.

2. To reapply a saved filter set, select the card for the filter set.

Display or hide columns

To customize which columns appear in the alerts view, follow these steps:

1. On the right side of the page, select **Customize columns**.
2. Select or clear the checkboxes for the columns you want to display or hide.

The column settings are saved across sessions and across browsers.

Search for alerts

Use the **Search** control to search for a user principal name (UPN), an assigned admin name, or an Alert ID.

View alert reports

Go to **Insider Risk Management** > **Reports** > **Alerts** to view Insider Risk Management alert reports for generated alerts, alerts by region, alerts by triggering event, and more.

Header/summary section of the Alert details page

The **Header/summary** section in the **Alert details** page is only available when selecting an alert from the Standard dashboard. It contains general information about the user and alert. Use this information for context while reviewing detailed information about the detected risk management activity included in the alert for the user:

- **Activity that generated this alert**: Displays the top potentially risky activity and policy match during the activity evaluation period that led to the alert being generated.
- **Triggering event**: Displays the most recent triggering event that prompted the policy to start assigning risk scores to the user's activity. If you configure integration with

Communication Compliance for *Data leaks by risky users* or *Security policy violations by risky users* policies, the triggering event for these alerts is scoped to Communication Compliance activity.

- **User details**: Displays general information about the user assigned to the alert. If anonymization is enabled, the username, email address, alias, and organization fields are anonymized.
- **User alert history**: Displays a list of alerts for the user for the last 30 days. Includes a link to view the complete alert history for the user.

> **Note**
>
> When a user is detected as a potential high impact user, this information is highlighted in the alert header in the **User details** page. The user details also include a summary with the reasons the user is detected as such. To learn more about setting policy indicators for potential high impact users, see Insider Risk Management settings.

Alerts generated from policies scoped to only activities that include priority content include the *Only activity with priority content was scored for this alert* notification in the **Header/summary** section.

> **Tip**
>
> To get a quick overview of an alert, select **Summarize** on the alert details page. When you select **Summarize**, a **Copilot** pane appears on the right side of the page with an alert summary. The alert summary includes all the essential details about the alert, such as the policy that was triggered, the activity that generated the alert, the triggering event, the user involved, their last working date (if applicable), any key user attributes, and the user's top risk factors. Copilot in Microsoft Purview consolidates information about the user from all their alerts and in-scope policies and emphasizes the user's top risk factors. You can also summarize the alert from the Alerts queue without having to open the alert by using Copilot. Or use the standalone version of Microsoft Security Copilot to investigate Insider Risk Management, Microsoft Purview Data Loss Prevention (DLP), and Microsoft Defender XDR alerts.

All risk factors tab

You can access this tab in **Alert details** for alerts in both dashboard views. It opens the summary of risk factors for the user's alert activity. Risk factors help you determine how risky the user's risk management activity is during your review. The risk factors include summaries for:

- **Cumulative exfiltration activities**: Events associated with cumulative exfiltration activities.
- **Health record access**: Potentially risky activities for events associated with accessing health records.
- **Priority content**: Potentially risky activities associated with priority content.
- **Risky browser usage**: Potentially risky activities for events associated with browsing to potentially inappropriate websites.
- **Sequences of activities**: Detected potentially risky activities associated with risk sequences.
- **Top exfiltration activities**: Exfiltration activities with the highest number of events for the alert.
- **Unallowed domains**: Potentially risky activities for events associated with unallowed domains.
- **Unusual activity for this user**: Specific activities for the user that are considered potentially risky, as they're unusual and a departure from their typical activities.

With these filters, you see only alerts with these risk factors, but the activity that generated an alert might not fall into any of these categories. For example, an alert containing sequence activities might be generated simply because the user copied a file to a USB device.

Content detected

The **Content detected** section on the **All risk factors** tab includes content associated with the risk activities for the alert and summarizes activity events by key areas. When you select an activity link, you open the Activity explorer and see more details about the activity.

Activity explorer tab

> ⓘ **Note**
>
> Users can access Activity explorer in the alert management area when their organization enables the feature and after the users trigger events.

The **Activity explorer** tab is available for alerts in both dashboard views. It provides risk investigators and analysts with a comprehensive analytics tool that provides detailed information about alerts. With the Activity explorer, reviewers can quickly review a timeline of detected potentially risky activity and identify and filter all risk activities associated with alerts.

Content preview (preview)

For supported activities, investigators can preview related content directly within Activity explorer to quickly validate risk without creating a case. Content preview (preview) enables early triage by helping investigators determine whether an activity contains sensitive data, represents a false positive, or requires escalation to a full investigation. Content preview is supported for activities that involve accessing or transmitting content.

Supported Content preview (preview) activities

Content preview is supported for activities that involve accessing or transmitting content, including:

- **SharePoint**: File accessed, file download, full file sync download
- **Exchange**: Email sent (hygiene events), Data loss prevention rule matches
- **OneDrive for Business**: File accessed, file download, full file sync download

Limitations of Content preview (preview)

Content preview isn't supported for activities that represent state changes, permission changes, deletions, or signal-only events. Examples include:

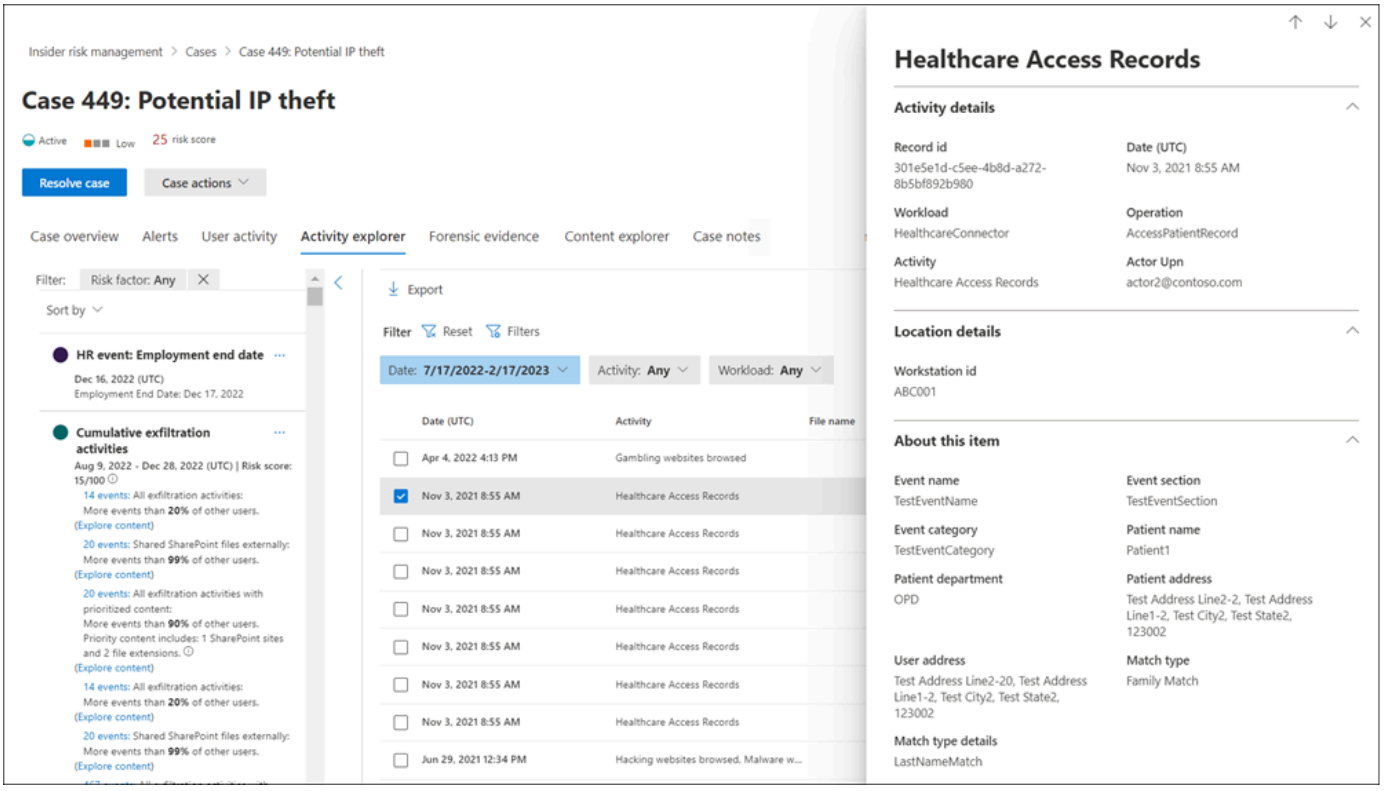
- File or folder deletion and recycle bin events
- Endpoint activities (such as file deletion, copying, printing, or USB transfers)
- Browser or removable media events
- Metadata-only events (user or group changes, Power BI export or deletion events)
- Files that have been renamed

Use the Activity explorer

When reviewing activities in the Activity explorer, investigators and analysts can select a specific activity and open the activity details pane. The pane displays detailed information about the activity that investigators and analysts can use during the alert triage process. Detailed information might provide context for the alert and assist with identifying the full scope of the risk activity that triggered the alert.

When selecting an activity's events from the activity timeline, the number of activities displayed in the explorer might not match the number of activity events listed in the timeline. Examples of why this difference might occur include:

- **Cumulative exfiltration detection**: Cumulative exfiltration detection analyzes event logs but applies a model that includes deduplicating similar activities to compute cumulative exfiltration risk. Additionally, you might see a difference in the number of potentially risky activities displayed in the Activity explorer if you make changes to your existing policy or settings. For example, if you modify allowed/unallowed domains or add new file type exclusions after a policy is created and potentially risky activity matches occur, the cumulative exfiltration detection activities differ from the results before the policy or settings changes. Cumulative exfiltration detection activity totals are based on the policy and settings configuration at the time of computation and don't include activities prior to the policy and settings changes.
- **Emails to external recipients**: Potentially risky activity for emails sent to external recipients is assigned a risk score based on the number of emails sent, which might not match the activity event logs.



Sequences that contain events excluded from risk scoring

A sequence detection might contain one or more events that are excluded from risk scoring based on your settings configuration. For example, your organization might use the **Global exclusions** setting to exclude .png files from risk scoring since .png files aren't normally risky. But a .png file could be used to obfuscate a malicious activity. For this reason, if an event that's excluded from risk scoring is part of a sequence due to an obfuscation activity, the event is included in the sequence since it might be interesting in the context of the sequence.

The Activity explorer displays the following information for excluded events in sequences:

- If a sequence contains a step where *all* events are excluded, the insight includes just the activity name and date. Select the **View the excluded events** link to filter for the excluded events in the Activity explorer. The User activity scatter plot icon has a risk score of 0 if all events are excluded.
- If a sequence has an insight where *some* events are excluded, the event information for the nonexcluded events is displayed, but the event count doesn't include the excluded events. Select the **View the excluded events** link to filter for the excluded events in the Activity explorer.
- If you select a *sequence link* for an insight, you can drill down into the sequence of events in the activity details pane, including any events that were excluded from scoring. An event excluded from scoring is marked as **Excluded**.

Filter alerts in the Activity explorer

To filter alerts in the Activity explorer for column information, select **Filters**. You can filter alerts by one or more attributes listed in the details pane for the alert. Activity explorer also supports customizable columns to help investigators and analysts focus the dashboard on the information most important to them.

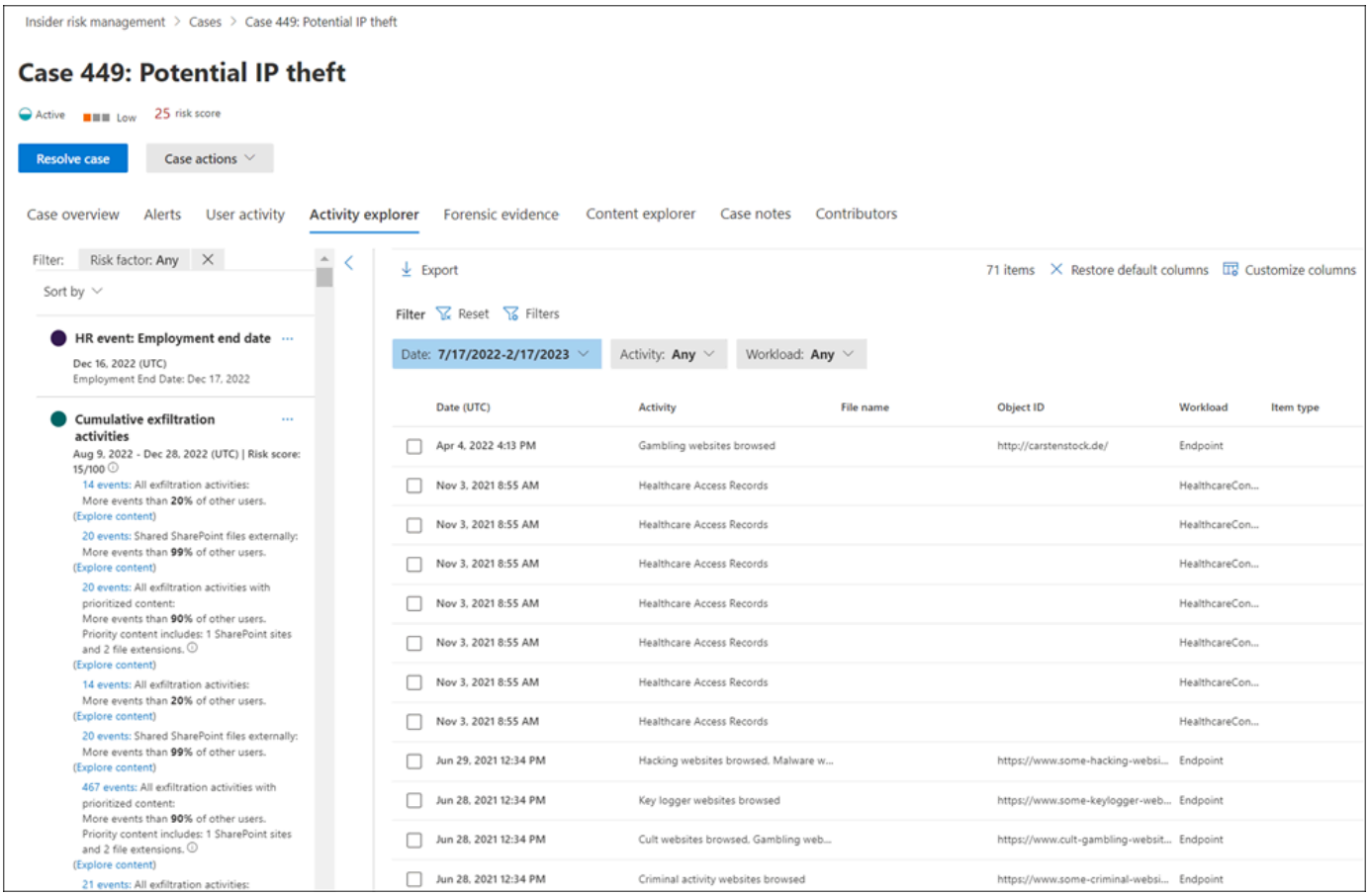
Use the *Activity scope*, *Risk factor*, and *Review status* filters to display and sort activities and insights for the following areas.

- **Activity scope**: Filters all scored activities for the user.
  - All scored activity for this user
  - Only scored activity in this alert

- **Risk factor**: Filters for risk factor activity applicable for all policies assigning risk scores. This filter includes all activity for all policies for included users.
  - Unusual activity
  - Includes events with priority content
  - Includes events with unallowed domain
  - Sequence activities
  - Cumulative exfiltration activities
  - Health record access activities
  - Risky browser usage

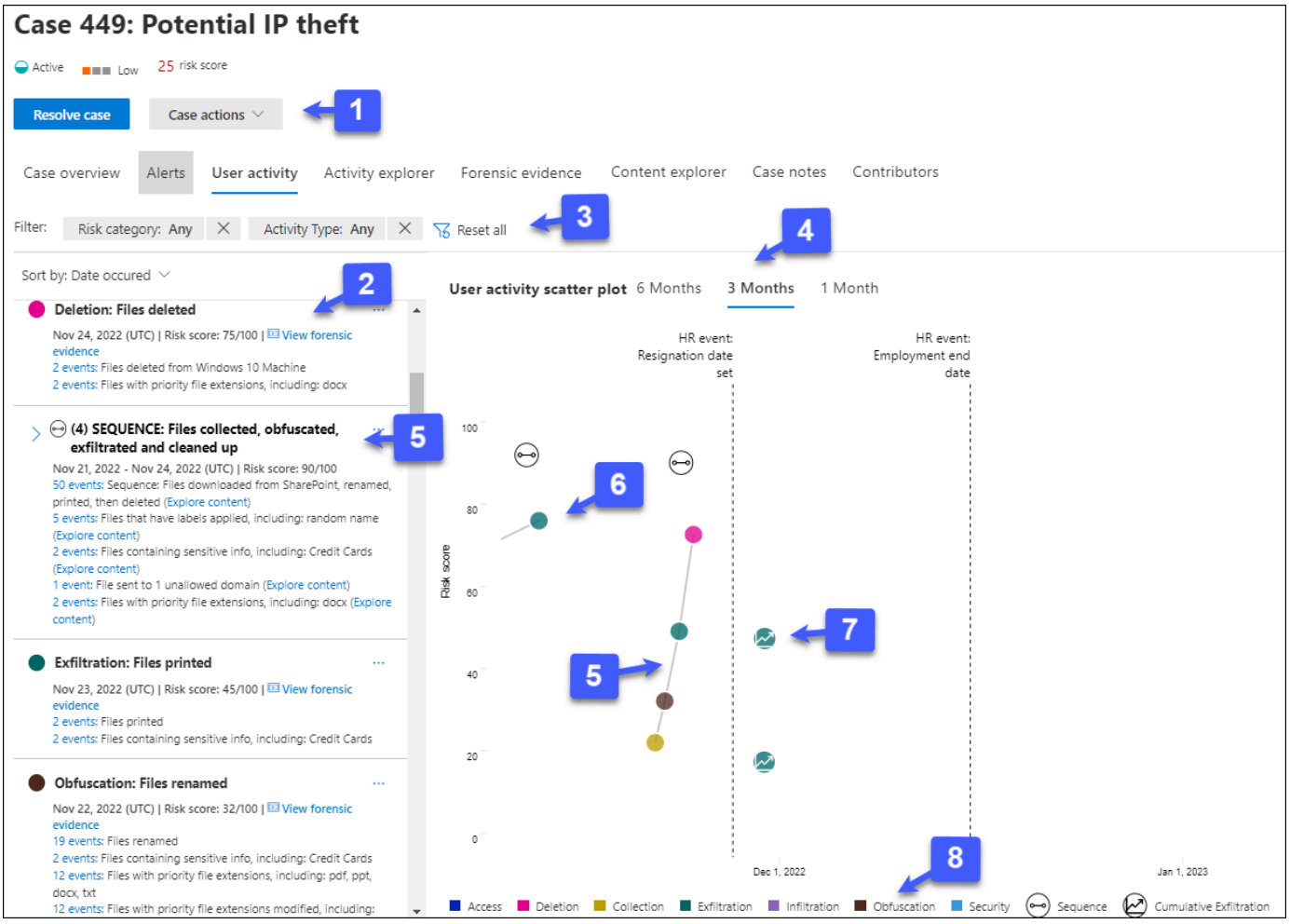
- **Review status**: Filters activity review status.
  - All

- Not yet reviewed (filters out any activity that was part of a dismissed or resolved alert)



Review the User activity tab

Both dashboard views provide the **User activity** tab for alerts. It's one of the most powerful tools for internal risk analysis and investigation for alerts and cases in the Insider Risk Management solution. You can quickly review all activities for a user, including a historical timeline of all alerts, alert details, the current risk score for the user, and the sequence of risk events.



1. **Case actions**: The case action toolbar provides options for resolving the case. When viewing a case, you can resolve the case, send an email notice to the user, or escalate the case for a data or user investigation.

2. **Risk activity chronology**: The full chronology of all risk alerts associated with the case, including all the details available in the corresponding alert bubble.

3. **Filters and sorting (preview)**:

    - **Risk category**: Filter activities by the following risk categories: *Activities with risk scores > 15 (unless in a sequence)* and *Sequence activities*.
    - **Activity Type**: Filter activities by the following types: *Access*, *Deletion*, *Collection*, *Exfiltration*, *Infiltration*, *Obfuscation*, *Security*, *Custom Indicator*, *Defense Evasion*, *Privilege Escalation*, *Communication Risk*, *User Compromise Risk*, and *AI Usage*.
    - **Sort by**: List the timeline of potentially risky activities by *Date occurred* or *Risk score*.

4. **Time filters**: By default, the User activity chart displays the last three months of potentially risky activities. You can easily filter the chart view by selecting the *6 Months*, *3 Months*, or *1 Month* tabs on the bubble chart.

5. **Risk sequence**: The chronological order of potentially risky activities is an important aspect of risk investigation. Identifying these related activities is an important part of evaluating overall risk for your organization. Alert activities that are related are displayed with connecting lines to highlight that these activities are associated with a larger risk area. An icon positioned over the sequence activities identifies sequences in this view relative to the risk score for the sequence. Hover over the icon to see the date and time of the risky activity associated with this sequence. This view of activities can help investigators 'connect the dots' for risk activities that could have been viewed as isolated or one-off events. Select the icon or any bubble in the sequence to display details for all the associated risk activities. Details include:

    - **Name** of the sequence.
    - **Date** or **Date range** of the sequence.
    - **Risk score** for the sequence. This score is the numerical score for the sequence of the combined alert risk severity levels for each related activity in the sequence.
    - **Number of events associated with each alert in the sequence**. Links to each file or email associated with each potentially risky activity are also available.
    - **Show activities in sequence**. Displays the sequence as a highlight line on the bubble chart and expands the alert details to display all related alerts in the sequence.

6. **Risk alert activity and details**: The User activity chart visually displays potentially risky activities as colored bubbles. The chart creates bubbles for different categories of risk. Select a bubble to display the details for each potentially risky activity. Details include:

    - **Date** of the risk activity.
    - The **risk activity category**. For example, *Email(s) with attachments sent outside the organization* or *File(s) downloaded from SharePoint Online*.
    - **Risk score** for the alert. This score is the numerical score for the alert risk severity level.
    - Number of events associated with the alert. Links to each file or email associated with the risk activity are also available.

7. **Cumulative exfiltration activities**: Select to view a visual chart of how activity is building over time for the user.

8. **Risk activity legend**: Across the bottom of the user activity chart, a color-coded legend helps you quickly determine risk category for each alert.

Use the Data risk graph tab

> **Important**
>
> **Data risk graph is being deprecated.** The data risk graph in Microsoft Purview Insider Risk Management will be retired on **November 24, 2026**. After that date, the data risk graph experience will no longer be available.
>
> You can't use this feature with the anonymized usernames privacy setting enabled.

Both dashboard views provide access to the **Data risk graph** tab for alerts. By integrating with Microsoft Sentinel, the data risk graph helps you view connections between impacted assets, users, and their activities in an interactive graph experience.

For more information about data risk graphs, see Data risk graph in Insider Risk Management.

Save a view of a filter to reuse later

If you create a filter and customize columns for the filter, you can save a view of your changes so that you or others can quickly filter for the same changes again later. When you save a view, you save both the filters and columns. When you load the view, it loads both saved filters and columns.

1. Create a filter and customize columns.

   > **Tip**
   >
   > To start over at any point, select **Reset**. To change columns that you customized, select **Reset columns**.

2. When you have the filter the way you want it, select **Save this view**, enter a name for the view, and then select **Save**.

   > **Note**
   >
   > The maximum length for a view name is 40 characters and you can't use any special characters.

3. To reuse the view of the filter later, select **Views**, and then select the view you want to open from the **Recommended views** tab (shows the most-used views) or the **Custom views** tab (displays the most frequently used filters).

When you select a view this way, it resets all the existing filters and replaces them with the view that you selected.

Alert status and severity

> **Note**
>
> Insider Risk Management throttles trigger processing to help protect and optimize your risk investigation and review experience. This throttling guards against issues that might result in an overload of policy alerts, such as misconfigured data connectors or data loss prevention policies. Insider Risk Management doesn't process signals received beyond the throttling limits. Learn more about limits in Insider Risk Management.

You can triage alerts into one of the following statuses:

- **Confirmed**: An alert you confirm and assign to a new or existing case.
- **Dismissed**: An alert you dismiss as benign in the triage process. You can provide a reason for the alert dismissal and include notes that are available in the user's alert history to provide additional context for future reference or for other reviewers. Reasons could range from expected activities, nonimpactful events, simply reducing the number of alert activities for the user, or a reason related to the alert notes. Reason classification choices include *Activity is expected for this user*, *Activity is impactful enough for me to investigate further*, and *Alerts for this user contain too much activity*.
- **Needs review**: A new alert where triage actions haven't yet been taken.
- **Resolved**: An alert that's part of a closed and resolved case.

The system automatically calculates alert risk scores from several risk activity indicators. These indicators include the type of risk activity, the number and frequency of the activity occurrence, the history of users' risk activity, and the addition of activity risks that might boost the seriousness of the potentially risky activity. The alert risk score drives the programmatic assignment of a risk severity level for each alert and can't be customized. If you don't triage alerts and risk activities continue to accrue to the alert, the risk severity level can increase. Risk analysts and investigators can use alert risk severity to help triage alerts in accordance with your organization's risk policies and standards.

Alert risk severity levels are:

- **High severity**: The potentially risky activities and indicators for the alert pose significant risk. The associated risk activities are serious, repetitive, and correlate strongly to other

significant risk factors.

- **Medium severity**: The potentially risky activities and indicators for the alert pose a moderate risk. The associated risk activities are moderate, frequent, and have some correlation to other risk factors.
- **Low severity**: The potentially risky activities and indicators for the alert pose a minor risk. The associated risk activities are minor, more infrequent, and don't correlate to other significant risk factors.

Dismiss multiple alerts (preview)

To save triage time, analysts and investigators can dismiss multiple alerts at once. With the **Dismiss alerts** command bar option, you can select one or more alerts with a *Needs review* status on the dashboard and quickly dismiss these alerts as benign. You can select up to 400 alerts to dismiss at one time.

Dismiss an insider risk alert

To dismiss alerts, you must be a member of the *Insider Risk Management*, *Insider Risk Management Analysts*, or *Insider Risk Management Investigators* role group.

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.
2. Go to the **Insider Risk Management** solution.
3. Select **Alerts** in the left navigation.
4. On the **Alerts dashboard**, select the alerts that have a *Needs review* status.
5. On the Alerts command bar, select **Dismiss alerts**.
6. On the **Dismiss alerts** detail pane, review the user and policy details associated with the selected alerts.
7. Select **Dismiss alerts** to resolve the alerts as benign.

Reports for alerts

To see reports for alerts, go to the **Reports** page. Each report widget on the **Reports** page displays information for the last 30 days:

- **Total alerts that need review**: The total number of alerts needing review and triage, including a breakdown by alert severity.
- **Open alerts over past 30 days**: The total number of alerts created by policy matches over the last 30 days, sorted by high, medium, and low alert severity levels.

- **Average time to resolve alerts**: A summary of useful alert statistics:
  - Average time to resolve high severity alerts, listed in hours, days, or months.
  - Average time to resolve medium severity alerts, listed in hours, days, or months.
  - Average time to resolve low severity alerts, listed in hours, days, or months.

Assign an alert

If you're an administrator and a member of the *Insider Risk Management*, *Insider Risk Management Analysts*, or *Insider Risk Management Investigators* role group, you can assign ownership of an alert to yourself or to an Insider Risk Management user with one of the same roles. After you assign an alert, you can reassign it to a user with any of the same roles. You can only assign an alert to one admin at a time.

> **Note**
>
> If you **scope your policies by one or more administrative units**, you can only give ownership of an alert to Insider Risk Management users with the appropriate role group permissions. The user highlighted in the alert must be in scope of the admin unit. For example, if an administrative scope applies to just users in Germany, the Insider Risk Management user can only see alerts for users in Germany. Unrestricted administrators can see all alerts for all users in the organization.

After you assign an admin, you can search by admin.

> **Note**
>
> Admins contained within a Microsoft Entra security group aren't supported for alert assignment. Admins must be directly assigned to one of the required roles.
>
> If you're using a custom group, make sure that the custom group contains the ***Case management* role**. The *Insider Risk Management Analysts* and the *Insider Risk Management Investigators* role groups both contain the *Case management* role. If you're using a custom group, you must explicitly add the *Case management* role to the group.

Assign an alert from the Alerts dashboard

To assign an alert from the Alerts dashboard, complete the following steps:

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.
2. Go to the **Insider Risk Management** solution.
3. Select **Alerts** in the left navigation.
4. On the **Alerts dashboard**, select the alerts that you want to assign.
5. In the command bar over the alerts queue, select **Assign**.
6. In the **Assign owner** pane on the right side of the screen, search for an admin with the appropriate permissions, and then select the checkbox for that admin.
7. Select **Assign**.

Assign an alert from the Alerts detail page

To assign an alert from the alert details page, follow these steps:

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.
2. Go to the **Insider Risk Management** solution.
3. Select **Alerts** in the left navigation.
4. Select an alert.
5. In the detail pane for the alert, in the upper-right corner of the page, select **Assign**.
6. In the **Suggested contacts** list, select the appropriate admin.

Create a case for an alert

Create a case for an alert to investigate potentially risky activity. You must be a member of the *Insider Risk Management* or *Insider Risk Management Investigators* role group to create a case.

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.
2. Go to the **Insider Risk Management** solution.
3. Select **Alerts** in the left navigation.
4. On the **Alerts dashboard**, select the alert you want to confirm and create a new case for.
5. On the **Alerts details pane**, select **Actions** > **Confirm alerts & create case**.
6. In the **Confirm alert and create insider risk case** dialog box, enter a name for the case, select users to add as contributors, and add comments as applicable. Comments automatically add to the case as a case note.
7. To download the content for the case, select **Create case with content download on** (preview). After you reach the content download limit, you can still create cases, but you

must leave the box unchecked. If you don't enable content download when you create the case, you can enable downloads when the case is active.

8. Select **Create case** to create a new case.

After you create the case, investigators and analysts can manage and act on the case. For more information, see the Insider Risk Management case article.

Retention and item limits

As Insider Risk Management alerts age, their value to minimize potentially risky activity diminishes for most organizations. Conversely, active cases and associated artifacts (alerts, insights, activities) always provide value and don't have an automatic expiration date. This retention policy includes all future alerts and artifacts in an active status for any user associated with an active case.

To minimize the number of older items that provide limited current value, the following retention and limits apply for Insider Risk Management alerts, cases, and user reports:

Expand table

Item	Retention/Limit
Active cases (and associated artifacts)	2,000 active cases, retained indefinitely and never expire
Alerts with Needs review status	120 days from alert creation, then automatically deleted
Maximum number of active cases	100
Resolved cases (and associated artifacts)	120 days from case resolution, then automatically deleted
User activities reports	120 days from report creation, then automatically deleted

Best practices for managing your alert volume

To minimize insider risks in your organization, regularly review, investigate, and act on potentially risky insider alerts. Quickly taking action to reduce these risks can potentially save your organization time, money, and potential regulatory or legal consequences. Learn about best practices for managing your Insider Risk Management alert queue.

Related content

- Best practices for managing your Insider Risk Management alert queue

- Take action on insider risk cases

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 08/27/2026

Data risk graph in Insider Risk Management

> **Important**
>
> Microsoft Purview Insider Risk Management correlates various signals to identify potential malicious or inadvertent insider risks, such as IP theft, data leakage, and security violations. Insider Risk Management enables customers to create policies to manage security and compliance. Built with privacy by design, users are pseudonymized by default, and role-based access controls and audit logs are in place to help ensure user-level privacy.

Data risk graphs in Insider Risk Management provide a visual investigation experience that combines asset and activity data into a single view. Powered by Microsoft Sentinel integration, it summarizes alert-related activity by a potential insider over the past 30 days, helping analysts triage alerts by revealing hidden contextual connections.

> **Important**
>
> **Data risk graph is being deprecated.** The data risk graph in Microsoft Purview Insider Risk Management will be retired on **November 24, 2026**. After that date, the data risk graph experience will no longer be available.

For more information about Microsoft Sentinel integration, see Learn about Microsoft Sentinel in Microsoft Purview.

> **Important**
>
> You cannot use this feature with the anonymized usernames privacy setting enabled. If you have anonymized user names, the graph will not load.

> **Note**
>
> Admin units aren't supported in the data risk graph. If you're scoped to an admin unit, data doesn't appear in data risk graphs.

Supported activities in data risk graph

The data risk graph currently supports the following exfiltration activities:

- Anonymous links created in SharePoint or OneDrive
- Anonymous links used via SharePoint or OneDrive
- Company links created in SharePoint or OneDrive
- File downloads from SharePoint and OneDrive
- Files renamed in SharePoint and OneDrive

For more information about exfiltration detection indicators, see Configure policy indicators in Insider Risk Management.

Before you begin

Before you can use data risk graphs in Insider Risk Management, complete the following steps:

- Learn about pay-as-you-go billing in Microsoft Sentinel for your organization.
- Configure the prerequisites for Microsoft Sentinel data lake and Microsoft Sentinel graph
- Review the changes made when onboarding to Microsoft Sentinel data lake and Microsoft Sentinel graph
- Configure at least one Insider Risk Management policy for supported data exfiltration activities.

Configure data risk graph

> **Important**
>
> You always have one data lake. If you already onboarded to the data lake with another Microsoft service, your existing data lake is used. If you never onboarded to the data lake, onboarding in Insider Risk Management enables Microsoft Sentinel data lake and graph in the Defender portal.

After you complete the prerequisites and understand the onboarding changes that Microsoft Sentinel data lake and graph make to your organization, complete the following steps to configure data risk graph in Insider Risk Management:

> **Important**

Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

1. Go to the Microsoft Purview portal with an account assigned the *Global Administrator* or *Security Administrator* role.
2. Select the **Insider Risk Management** solution card, then select **Recommended actions** in the left nav.
3. In the **Comprehensive protection** section, select **Set up data lake and data risk graph**.

   If you don't have the correct permissions to set up the data lake, a notification appears to contact a global or billing administrator in your organization.
4. On the **Setup** tab on **Set up data lake & risk graph**, configure the following settings:
   a. **Subscription**: Enter the Azure subscription you want to use. Your selected subscription must be valid and accessible and you must be the subscription owner.
   b. **Resource group**: Enter the resource group you want to use. Your selected resource group must be valid and accessible.

   > **Important**
   >
   > After the data lake is provisioned for a specific Azure subscription and resource group, you can't migrate it to a different subscription or resource group.
5. Select **Setup**.

The data lake and data risk graph setup process begins, and onboarding might take up to 60 minutes to complete. You can close the setup panel while the process is running. After the onboarding process completes, you see *Complete* on **Set up data lake & risk graph**. The data risk graph shows events that occur after your Microsoft Sentinel data lake was created or onboarded.

Initial processing of data and data risk graph availability for an investigation can take 24–48 hours.

> **Important**
>
> When the data risk graph is first created, it includes the most recent seven days of data. As the data risk graph refreshes over time, the look-back window expands up to a maximum of

30 days. It might take some time for the data risk graph to reach the full 30-day window, so expect gradual growth after initial setup.

Use data risk graph

Data risk graph in Insider Risk Management uniquely visualizes correlations between impacted data, users, and their activities. It provides critical context to guide mitigation and next steps. For example, when you uncover an alert related to a highly sensitive document, data risk graphs give you visibility into which users downloaded it or if they accessed it from an IP address flagged as risky by your organization's security signals. This visibility lets you uncover new nodes to a data security incident, like additional users or new content related to an alert.

To view the data risk graph for an alert, you must be assigned the *Insider Risk Management Graph Reader* role. This role is included by default in several Insider Risk Management built-in role groups. For more information, see Assign permissions in Insider Risk Management.

> **Important**
>
> All user names are visible in data risk graph, even if pseudonymization is enabled in Insider Risk Management privacy settings.

To use the data risk graph in Insider Risk Management, complete the following steps:

1. Go to the Microsoft Purview portal with an account assigned the *Insider Risk Management Graph Reader* role.
2. Select the **Insider Risk Management** solution card, then select **Alerts** in the left nav.
3. Select an alert to review, then select the **Data risk graph** tab.
4. Filter the time as applicable and select individual data risk graph nodes for more information about each connection during triage.
5. Expand relationships in the graph by selecting the + icon on users or assets.

The data risk graph contains several nodes. Selecting a node shows details about the node:

- **User details**: Shows information about users associated with the alert. This information includes organization information for each user, including User Principal Name (UPN), labels, location, job title, and more.
- **Data details**: Shows information about files and sites. This information includes the object location, type, labels, and more.

You can view data risk graph information for users based on the last 30 days of activity. This fixed duration isn't extendable.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 08/27/2026

Best practices for managing your alert volume in Insider Risk Management

08/20/2025

> **Important**
>
> Microsoft Purview Insider Risk Management correlates various signals to identify potential malicious or inadvertent insider risks, such as IP theft, data leakage, and security violations. Insider Risk Management enables customers to create policies to manage security and compliance. Built with privacy by design, users are pseudonymized by default, and role-based access controls and audit logs are in place to help ensure user-level privacy.

Reviewing, investigating, and acting on potentially risky insider alerts are important parts of minimizing insider risks in your organization. When you quickly take action to minimize the impact of these risks, you can potentially save time, money, and regulatory or legal ramifications for your organization. In this remediation process, the first step of reviewing alerts can seem like the most difficult task for many analysts and investigators.

This article provides best practices for managing the volume of alerts in your organization so that you don't have too many or too few alerts. For a general discussion of how alerts are generated and the tools for managing alerts, see Investigate insider risk activities.

Too few alerts to review

If you're receiving too few Insider Risk Management alerts, try the following solutions:

- **Update your settings:** Changes that you make to settings apply globally across all of your policies.
  - **Enable more indicators:** When you select more indicators, your policies can detect a larger group of activities.

    How to: Go to **Settings** > **Policy indicators**, and then enable all available and relevant indicators.
  - **Adjust the Alert volume slider:** Use this slider to see all medium and high severity alerts, and most low severity alerts. **Note:** Adjusting the slider might result in more false positives.

How to: Go to **Settings** > **Intelligent detections** > **Alert volume**, and then move the slider to **More alerts**.

- **Modify the policy:** Find the policy that isn't generating enough alerts, and then consider the following actions:

  - **Increase the user coverage in the policy:** Policies with few users included in the scope are less likely to generate alerts. If applicable, consider increasing the number of users in scope for your policy.

    How to: Select a specific policy on the **Policies** page, select **Edit policy**, and then go to the **Users and groups** page to increase the number of in-scope users.

  - **Lower the trigger thresholds:** For policies based on the *Data leaks* and *Risky Browser Usage (preview)* templates, you can customize some trigger thresholds. These thresholds define when you start to detect user activities. If you lower trigger thresholds, you lower the criteria for a user to start being evaluated for risky activity. **Note:** If a user doesn't appear in the **Users and groups** page, it means that triggering event criteria isn't met yet.

    How to: Go to the specific policy on the **Policies** page, select **Edit policy**, go to the **Trigger thresholds** page, select the **Use custom thresholds** option, and then adjust your thresholds.

  - **Add more indicators:** Indicators are the activities a user must do to be considered risky. If you don't have many indicators (activities considered to be risky) selected in your policy, alerts are less likely to be generated.

    How to: Go to the specific policy on the **Policies** page, select **Edit policy**, go to the **Indicators** page, and then select more indicators.

  - **Lower indicator thresholds:** After your users begin getting evaluated (have a triggering event), an alert is generated only if they do activities over a certain threshold that might indicate that their activity is risky. When you lower the indicator thresholds, you lower the threshold that users must surpass to generate an alert.

    How to: Go to the specific policy on the **Policies** page, select **Edit policy**, go to the **Indicator thresholds** page, select the **Customize thresholds** option, and then set your thresholds. Learn about indicator threshold recommendations

Too many alerts to review

If you receive too many valid alerts or have too many stale low-risk alerts, consider taking the following actions:

- **Enable analytics:** When you enable analytics, you can quickly identify potential risk areas for your users. Analytics helps you determine the type and scope of Insider Risk Management policies that you might want to configure.

  How to: Go to **Settings** > **Analytics**.

- **Get real-time insights:** If you want to take advantage of threshold recommendations, get real-time insights from analytics. These insights can help you efficiently adjust the selection of indicators and thresholds of activity occurrence so that you don't receive too few or too many policy alerts.

  How to: See Use real-time analytics to help manage alert volume.

- **Adjust your policies:** The most basic method to address the type and volume of alerts is selecting and configuring the correct insider risk policy. Starting with the appropriate policy template helps you focus the types of risk activities and alerts you see. Other factors that might impact alert volume are the size of the in-scope user and groups and the content and channels that you prioritize. Consider adjusting policies to refine these areas to what is most important for your organization.

  How to: Select a specific policy on the **Policies** page, then select **Edit policy**.

- **Modify your insider risk settings:** Insider risk settings include a wide variety of configuration options that can impact the volume and types of alerts you receive. Review and understand the following settings to filter out alert noise:
  - Policy indicators and indicator thresholds
  - Global exclusions
  - Detection groups
  - Variants of built-in indicators
  - Intelligent detections
  - Policy timeframes

- **Enable inline alert customization:** When you enable inline alert customization, analysts and investigators can quickly edit policies when reviewing alerts. They can update thresholds for activity detection with Microsoft recommendations, configure custom thresholds, or choose to ignore the type of activity that created the alert. If you don't enable this setting, only users assigned to the *Insider Risk Management* role group can use inline alert customization.

  How to: Go to **Settings** > **Inline alert customization**.

- **Bulk delete alerts, where applicable:** To save triage time for your analysts and investigators, consider immediately dismissing multiple alerts in bulk. You can select up to 400 alerts to dismiss at one time.

Managing resource constraints in your organization

Modern workplace users often have a wide variety of responsibilities and demands on their time. To address resource constraints, take these actions:

- **Focus analyst and investigator efforts on the highest risk alerts first:** Depending on your policies, you might capture user activities and generate alerts with varying degrees of potential impact on your risk mitigation efforts. Filter alerts by severity and prioritize *High severity* alerts.
- **Use Microsoft Copilot**: You can use Microsoft Copilot in Microsoft Purview to summarize an alert without even opening it. This approach can speed up the triaging experience.
- **Assign users as analysts and investigators:** Having the right user assigned to the proper roles is an important part of the insider risk alert review process. Make sure you assign the appropriate users to the *Insider Risk Management Analysts* and *Insider Risk Management Investigators* role groups.
- **Use automated insider risk features to help discover the highest risk activities:**
  - Use Insider Risk Management sequence detection and cumulative exfiltration detection to quickly discover harder to find risks in your organization.
  - Consider fine-tuning your risk score boosters, and use global exclusions, detection groups, and variants.
  - Fine tune minimum indicator threshold settings for your policies.

See also

Investigate Insider Risk Management activities

Take action on Insider Risk Management cases

> **Important**
>
> Microsoft Purview Insider Risk Management correlates various signals to identify potential malicious or inadvertent insider risks, such as IP theft, data leakage, and security violations. Insider Risk Management enables customers to create policies to manage security and compliance. Built with privacy by design, users are pseudonymized by default, and role-based access controls and audit logs are in place to help ensure user-level privacy.

Cases are the core of Insider Risk Management. They let you investigate and act on issues that risk indicators in your policies generate. Create cases manually from alerts when you need to take extra steps to address a compliance-related issue for a user. Each case focuses on one user. You can add multiple alerts for that user to an existing case or start a new case.

After you investigate the details of a case, take action by:

- Sending the user a notice
- Resolving the case as benign
- Sharing the case with your ServiceNow instance or with an email recipient
- Escalating the case for an eDiscovery (Premium) investigation

For an overview of how you investigate and manage cases in Insider Risk Management, see the Insider Risk Management Investigation and Escalation video.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Use the Cases dashboard

> **Tip**

To see reports for cases, go to the **Reports** page. Each report widget on the **Reports** page displays information for the last 30 days:

- **Active cases**: The total number of active cases under investigation.
- **Cases over past 30 days**: The total number of cases created, sorted by *Active* and *Closed* status.
- **Statistics**: Average time of active cases, listed in hours, days, or months.

The Insider Risk Management **Cases dashboard** enables you to view and act on cases. The case queue lists all active and closed cases for your organization, along with the current status of the following case attributes:

- **Case ID**: The ID of the case.
- **Case name**: The name of the case, defined when you confirm an alert and create the case.
- **Status**: The status of the case, either *Active* or *Closed*.
- **User**: The user for the case. If you enable anonymization for usernames, the portal displays anonymized information.
- **Content download**: This field shows whether content download is on or off for the case.
- **Time case opened**: The time that passes since the case is opened.
- **Total policy alerts**: The number of policy matches included in the case. This number might increase if you add new alerts to the case.
- **Last activity date**: The time that passes since you add a case note or change the case state.
- **Last updated by**: The name of the Insider Risk Management analyst or investigator that last updated the case.

> ⓘ **Note**
>
> If you scope your policies by one or more administrative units, you can only give ownership of a case to Insider Risk Management users with the appropriate role group permissions. The user highlighted in the alert must be in scope of the admin unit. For example, if an administrative scope applies to just users in Germany, the Insider Risk Management user can only see alerts for users in Germany. Unrestricted administrators can see all cases for all users in the organization.

Use the **Search** control to search for a Case ID or to search for specific text in case names. Use the case filter to sort cases by the following attributes:

- Status

- Time case opened, start date, and end date
- Last updated, start date, and end date

Assign a case

If you're an administrator with the appropriate permissions, you can assign ownership of a case to yourself or to an Insider Risk Management user with the Insider Risk Management, Insider Risk Management Analyst, or Insider Risk Management Investigator role. After you assign a case, you can reassign it to a user with any of the same roles. You can only assign a case to one admin at a time.

If you assign an admin to a case, you can filter by admin.

Assign a case from the Cases dashboard

To assign a case from the Cases dashboard, complete the following steps:

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.
2. Go to the **Insider Risk Management** solution.
3. Select **Cases** in the left navigation.
4. On the **Cases dashboard**, select the cases that you want to assign.
5. In the command bar over the cases queue, select **Assign**.
6. On the **Assign owner** pane on the right side of the screen, search for an admin with the appropriate permissions, and then select the checkbox for that admin.
7. Select **Assign**.

Assign a case from the Cases detail page

To assign a case from the case detail page, complete the following steps:

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.
2. Go to the **Insider Risk Management** solution.
3. Select **Cases** in the left navigation.
4. Select a case.
5. In the detail pane for the case, select **Assign**.
6. In the **Suggested contacts** list, select the appropriate admin.

Filter cases

Depending on the number and type of active Insider Risk Management policies in your organization, reviewing a large queue of cases can be challenging. By using case filters, analysts and investigators can sort cases by several attributes. To filter alerts on the **Cases dashboard**, select the **Filter** control. You can filter cases by one or more attributes:

- **Status**: Select one or more status values to filter the case list. The options are *Active* and *Closed*.
- **Time case opened**: Select the start and end dates for when the case was opened.
- **Last updated**: Select the start and end dates for when the case was updated.

Filter cases, save views, customize columns, or search cases

Depending on the number and type of active Insider Risk Management policies in your organization, reviewing a large queue of cases can be challenging. To help you keep track of cases, you can:

- Filter cases by various attributes.
- Save a view of a filter set to reuse later.
- Display or hide columns.
- Search for an alert.

Filter cases

To filter cases on the Cases dashboard, complete the following steps:

1. Select **Add filter**.
2. Select one or more of the following attributes:

Expand table

Attribute	Description
Assigned to	The admin that the alert is assigned to for triaging (if assigned).
Last activity date	The start and end dates for when the case was last updated.
Status	Current status of the case. The options are *Active* and *Closed*.
Time case opened	The start and end dates for when the case was opened.

The filter bar shows the attributes that you select.

3. Select an attribute in the filter bar, and then select a value to filter by. For example, select the **Last activity date** attribute, enter or select the dates in the **Start date** and **End date** fields, and then select **Apply**.

   > **Tip**
   >
   > To start over at any point, select **Reset all** on the filter bar.

Save a view of a filter set to reuse later

To save a filter set for reuse, complete the following steps:

1. After applying the case filters described in Filter cases, select **Save** over the filter bar, enter a name for the filter set, and then select **Save**.

   The filter set appears as a card over the filter bar. It includes a number that shows the count of cases that meet the criteria in the filter set.

   > **Note**
   >
   > You can save up to five filter sets. To delete a filter set, select the ellipsis (three dots) in the upper-right corner of the card, and then select **Delete**.

2. To reapply a saved filter set, select the card for the filter set.

Display or hide columns

To display or hide columns on the Cases dashboard, complete the following steps:

1. On the right side of the page, select **Customize columns**.
2. Select or clear the checkboxes for the columns you want to display or hide.

The portal saves the column settings across sessions and across browsers.

Search for cases

Use the **Search** control to search for a user principal name (UPN), an assigned admin name, or an Alert ID.

Investigate a case

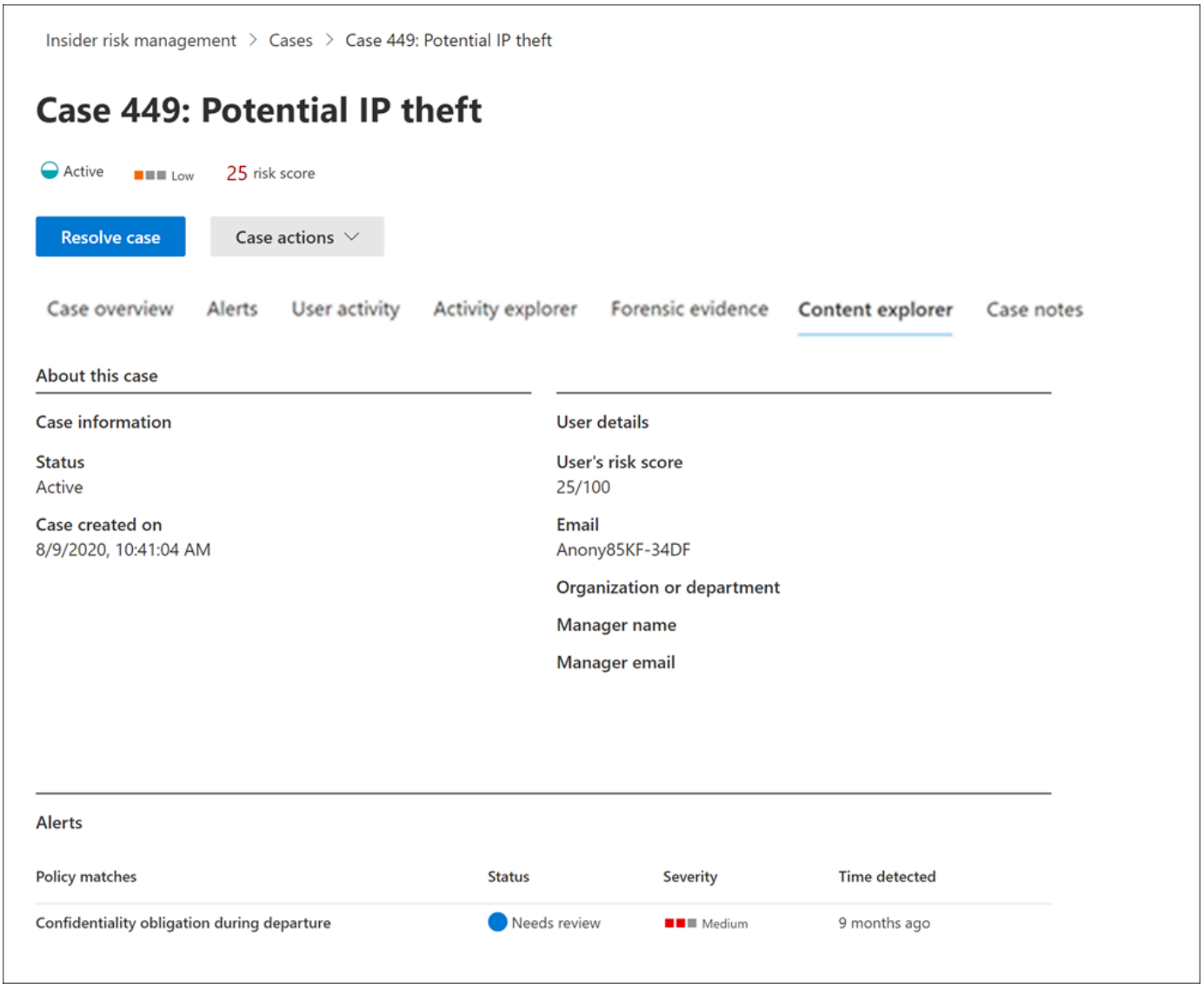
Deeper investigation into Insider Risk Management alerts is critical to taking proper corrective actions. Insider Risk Management cases are the central management tool to dive deeper into user risk activity history, alert details, the sequence of risk events, and to explore the content and messages exposed to risks. Risk analysts and investigators also use cases to centralize review feedback and notes and to process case resolution.

Selecting a case opens the case management tools and allows analysts and investigators to dig into the details of cases.

Review the case overview

The **Case overview** tab summarizes the case details for risk analysts and investigators. It includes the following information in the **About this case** area:

- **Case ID**: The ID of the case.
- **Status**: The current status of the case, either Active or Closed.
- **Case created on**: The date and time the case was created.
- **User's risk score**: The current calculated risk level of the user for the case. The system calculates this score every 24 hours and uses alert risk scores from all active alerts associated to the user. When *User is detected as a potential high impact user* or *User is a member of a priority user group* risk booster is enabled as **Risk score boosters** in the **Policy indicators** section of the **Insider Risk Management settings** page, the **User details** page includes detailed information about the user's calculated risk level.
- **Email**: The email alias of the user for the case.
- **Organization or department**: The organization or department that the user is assigned to.
- **Manager name**: The name of the user's manager.
- **Manager email**: The email alias of the user's manager.



The **Case overview** tab also includes an **Alerts** section that includes the following information about policy match alerts associated with the case:

- **Policy matches**: The name of the Insider Risk Management policy associated with the match alerts for potentially risky user activity that might lead to a security incident.
- **Status**: Status of the alert.
- **Severity**: Severity of the alert.
- **Time detected**: The time that has passed since the alert was generated.

Review alerts in a case

The **Alerts** tab summarizes the current alerts included in the case. You can add new alerts to an existing case. When you assign new alerts, the **Alert** queue includes them. The queue lists the following alert attributes:

- Alert
- Alert ID

- Status
- Severity
- Time detected

Select an alert from the queue to display the **Alert detail** page.

Use the search control to search for an Alert ID or specific text in alert names. Use the alert filter to sort cases by the following attributes:

- Status
- Severity
- Time detected, start date, and end date

Use the filter control to filter alerts by several attributes, including:

- **Status**: Select one or more status values to filter the alert list. The options are *Confirmed*, *Dismissed*, *Needs review*, and *Resolved*.
- **Severity**: Select one or more alert risk severity levels to filter the alert list. The options are *High*, *Medium*, and *Low*.
- **Time detected**: Select the start and end dates for when the alert was created.
- **Policy**: Select one or more policies to filter the alerts generated by the selected policies.

Review user activity in a case

Risk analysts and investigators use the **User activity** tab to review user activity details. It provides a visual representation of all the potentially risky activities associated with risk alerts and cases. Use this information to determine whether those risky activities might lead to a security incident. For example, as part of the alert triage process, analysts might need to review all the risk activities associated with the case for more details. In cases, risk investigators can review user activity details and the bubble chart to help understand the overall scope of the risk activities associated with the case. For more information about the User activity chart, see the Insider Risk Management activities article.

Activity explorer (preview)

The **Activity explorer** tab helps risk analysts and investigators review case activity details associated with risk alerts. For example, as part of the case management actions, investigators and analysts might need to review all the risk activities associated with the case for more details. By using the **Activity explorer**, reviewers can quickly examine a timeline of detected potentially risky activity and identify and filter all risk activities associated with alerts.

For more information about the Activity explorer, see the Insider Risk Management activities article.

Review forensic evidence

The **Forensic evidence** tab helps risk investigators review visual captures associated with risk activities included in cases. For example, as part of the case management actions, investigators might need to clarify the context of the user activity under review. Viewing the actual clips of the activity can help the investigator determine if the user activity is potentially risky and might lead to a security incident.

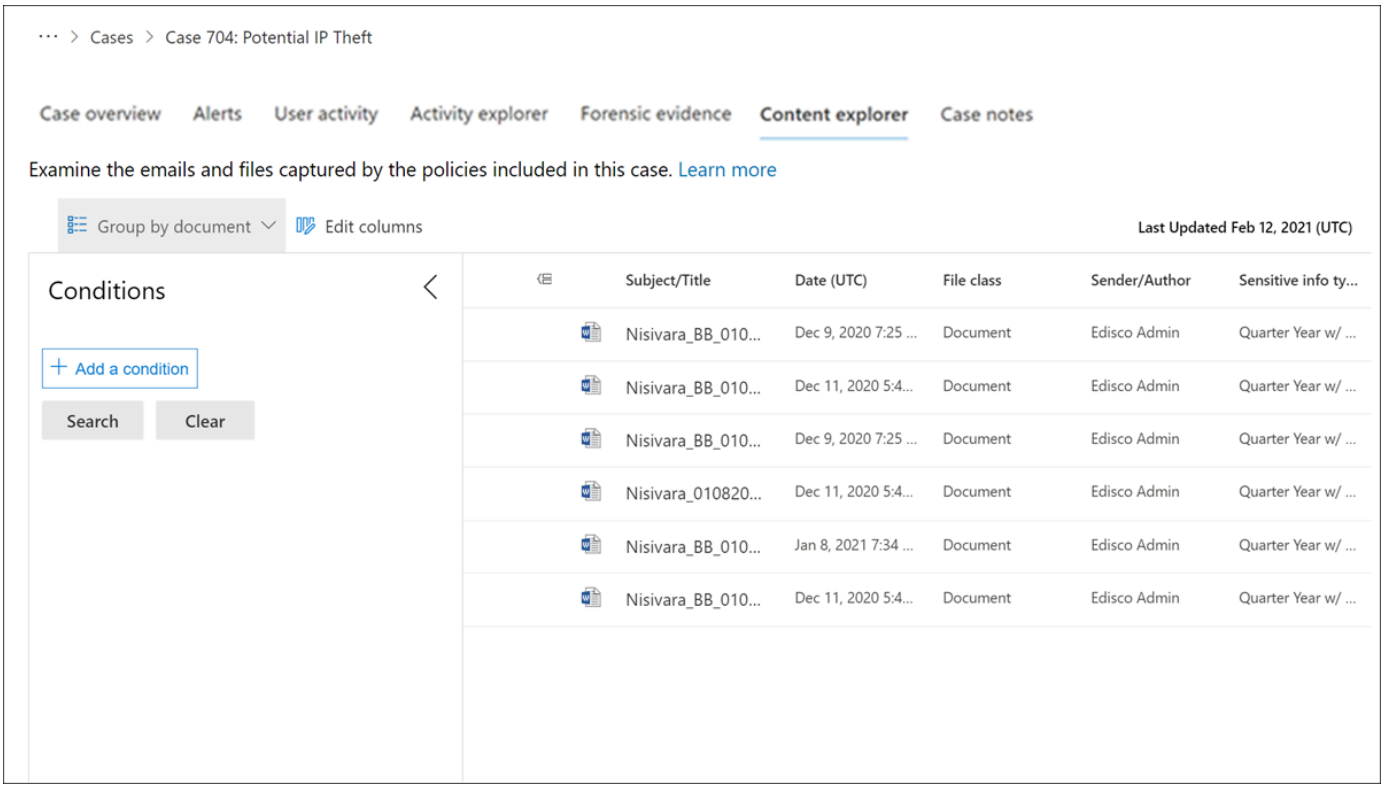
For more information about forensic evidence, see the Learn about Insider Risk Management forensic evidence article.

Use Content explorer for case evidence

The **Content explorer** tab helps risk investigators review copies of all individual files and email messages associated with risk alerts. For example, if an alert is created when a user downloads hundreds of files from SharePoint Online and the activity triggers a policy alert, the case captures and copies all the downloaded files for the alert to the Insider Risk Management case from original storage sources.

If content download is off for the case, select **Add content** from the Content explorer tab. If you've reached the maximum number cases with content downloaded limit, you must resolve cases to download additional content.

The Content explorer is a powerful tool with basic and advanced search and filtering features. To learn more about using the Content explorer, see Insider Risk Management Content explorer.



Manage case notes

Risk analysts and investigators use the **Case notes** tab in a case to share comments, feedback, and insights about their work for the case. Notes are permanent additions to a case. After saving a note, you can't edit or delete it. When you create a case from an alert, the comments you enter in the **Confirm alert and create insider risk case** dialog box automatically become a case note.

The case notes dashboard shows notes by the user who created the note and the time that passes since the note is saved. To search the case note text field for a specific keyword, use **Search** on the case dashboard and enter a specific keyword.

System-generated notes (preview)

System-generated notes are automatically added to cases when there's a change in status or assignment. The following changes trigger a system-generated note:

- A change in case status
- A change in the assigned user
- Closure of a case
- A case escalation

System-generated notes appear alongside analyst-created notes in the **Notes** tab within the case details panel, helping you stay up to date with all activity on a case.

Add a case note

To add a note to a case, complete the following steps:

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.
2. Go to the **Insider Risk Management** solution.
3. Select **Cases** in the left navigation.
4. Select a case, then select the **Case notes** tab.
5. Select **Add case note**.
6. In the **Add case note** dialog box, type the note.
7. Select **Save** to add the note to the case.

Manage case contributors

Risk analysts and investigators can add other reviewers to the case through the **Contributors** tab. Users must be assigned the **Role Management** role to add contributors to the case. By default, all users assigned the **Insider Risk Management Investigators** and the **Insider Risk Management** roles appear as contributors for each active and closed case.

You can grant temporary access to a case by adding a user as a contributor, but with the following restrictions:

- Analysts and investigators can add contributors.
- You can't add analysts as contributors.
- Contributors can't add contributors.

Contributors have all case management control on the specific case except for:

- Permission to confirm or dismiss alerts.
- Permission to edit the contributors for cases.

Add a contributor to a case

To add a contributor to a case, complete the following steps:

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.
2. Go to the **Insider Risk Management** solution.
3. Select **Cases** in the left navigation.
4. Select a case, and then select the **Contributors** tab.
5. Select **Add contributor**.

6. In the **Add contributor** dialog box, start typing the name of the user you want to add, and then select the user from the suggested user list. This list comes from the Microsoft Entra ID of your tenant subscription.
7. Select **Add** to add the user as a contributor.

Case actions

Risk investigators can take action on a case in one of several methods. The method depends on the severity of the case, the history of risk of the user, and the risk guidelines of your organization. In some situations, you might need to escalate a case to a user or data investigation to collaborate with other areas of your organization and to dive deeper into risk activities. Insider Risk Management is tightly integrated with other Microsoft Purview solutions to help you with end-to-end resolution management.

Send email notice

In most cases, user actions that create insider risk alerts are inadvertent or accidental. Sending a reminder notice to the user by email is an effective method for documenting case review and action. This method reminds users of corporate policies or points them to refresher training. You generate notices from notice templates that you create for your Insider Risk Management infrastructure.

Remember that sending an email notice to a user ***does not*** resolve the case as *Closed*. In some cases, you might want to leave a case open after sending a notice to a user to look for more risk activities without opening a new case. If you want to resolve a case after sending a notice, select **Resolve case** as a follow-on step after sending a notice.

Send a notice to the user assigned to a case

To send a notice to the user assigned to a case, complete the following steps:

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.
2. Go to the **Insider Risk Management** solution.
3. Select **Cases** in the left navigation.
4. Select a case, and then select **Send email notice** on the case action toolbar.
5. In the **Send e-mail notice** dialog box, select the **Choose a notice template** dropdown control to select the notice template for the notice. This selection pre-fills the other fields in the notice.

6. Review the notice fields and update as appropriate. The values you enter override the values in the template.
7. Select **Send** to send the notice to the user. All sent notices are added to the case notes queue on the **Case notes** dashboard.

Escalate for investigation

Escalate the case for user investigation when you need extra legal review for the user's risk activity. This escalation opens a new Microsoft Purview eDiscovery (Premium) case in your Microsoft 365 organization. eDiscovery (Premium) provides an end-to-end workflow to preserve, collect, review, analyze, and export content that's responsive to your organization's internal and external legal investigations. It also lets your legal team manage the entire legal hold notification workflow to communicate with custodians involved in a case. Escalating to an eDiscovery (Premium) case from an Insider Risk Management case helps your legal team take appropriate action and manage content preservation. For more information about eDiscovery (Premium) cases, see Overview of Microsoft Purview eDiscovery (Premium).

Escalate a case to a user investigation

To escalate a case to a user investigation, complete the following steps:

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.
2. Go to the **Insider Risk Management** solution.
3. Select **Cases** in the left navigation.
4. Select a case, and then select **Escalate for investigation** on the case action toolbar.
5. In the **Escalate for investigation** dialog box, enter a name for the new user investigation. If needed, enter notes about the case, and then select **Escalate**.
6. Review the notice fields and update as appropriate. The values you enter override the values on the template.
7. Select **Confirm** to create the user investigation case.

After you escalate the Insider Risk Management case to a new user investigation case, you can review the new case in the **eDiscovery** > **Advanced** area in the Microsoft Purview portal.

Run automated tasks by using Power Automate flows for the case

By using recommended Power Automate flows, risk investigators and analysts can quickly take action to:

- Request information from HR or business about a user in an insider risk case.
- Notify manager when a user has an insider risk alert.
- Create a record for an Insider Risk Management case in ServiceNow.
- Notify users when they're added to an insider risk policy.

To run, manage, or create Power Automate flows for an Insider Risk Management case:

1. Select **Automate** on the case action toolbar.
2. Choose the Power Automate flow to run, and then select **Run flow**.
3. After the flow completes, select **Done**.

For more information about Power Automate flows for Insider Risk Management, see Getting started with Insider Risk Management settings.

View or create a Microsoft Teams team for the case

When you enable Microsoft Teams integration for Insider Risk Management in settings, the solution automatically creates a Microsoft Teams team every time you confirm an alert and create a case. Risk investigators and analysts can quickly open Microsoft Teams and navigate directly to the team for a case by selecting **View Microsoft Teams team** on the case action toolbar.

For cases you open before enabling Microsoft Team integration, risk investigators and analysts can create a new Microsoft Teams team for a case by selecting **Create Microsoft Teams team** on the case action toolbar.

When you resolve a case, the solution automatically archives the associated Microsoft Team (hides it and turns it to read-only).

For more information about Microsoft Teams for Insider Risk Management, see Getting started with Insider Risk Management settings.

Resolve the case

After risk analysts and investigators complete their review and investigation, resolve a case to act on all the alerts currently included in the case. Resolving a case adds a resolution classification, changes the case status to *Closed*, and automatically adds the resolution action reasons to the case notes queue on the **Case notes** dashboard. Resolve cases as either:

- **Benign**: The classification for cases where policy match alerts are evaluated as low risk, non-serious, or false positive.
- **Confirmed policy violation**: The classification for cases where policy match alerts are evaluated as risky, serious, or the result of malicious intent.

> **Note**
>
> If you select the option to delete the case and associated content immediately when resolving the case, you permanently remove all related content, alerts, and insights.

Resolve a case

To resolve a case, complete the following steps:

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.
2. Go to the **Insider Risk Management** solution.
3. Select **Cases** in the left navigation.
4. Select a case, and then select **Resolve case** on the case action toolbar.
5. In the **Resolve case** dialog box, select the **Resolve as** dropdown control to select the resolution classification for the case. The options are **Benign** or **Confirmed policy violation**.
6. In the **Resolve case** dialog box, enter the reasons for the resolution classification in the **Action taken** text field.
7. Select **Resolve** to close the case.

Note: The author created this article with assistance from AI. Learn more

---

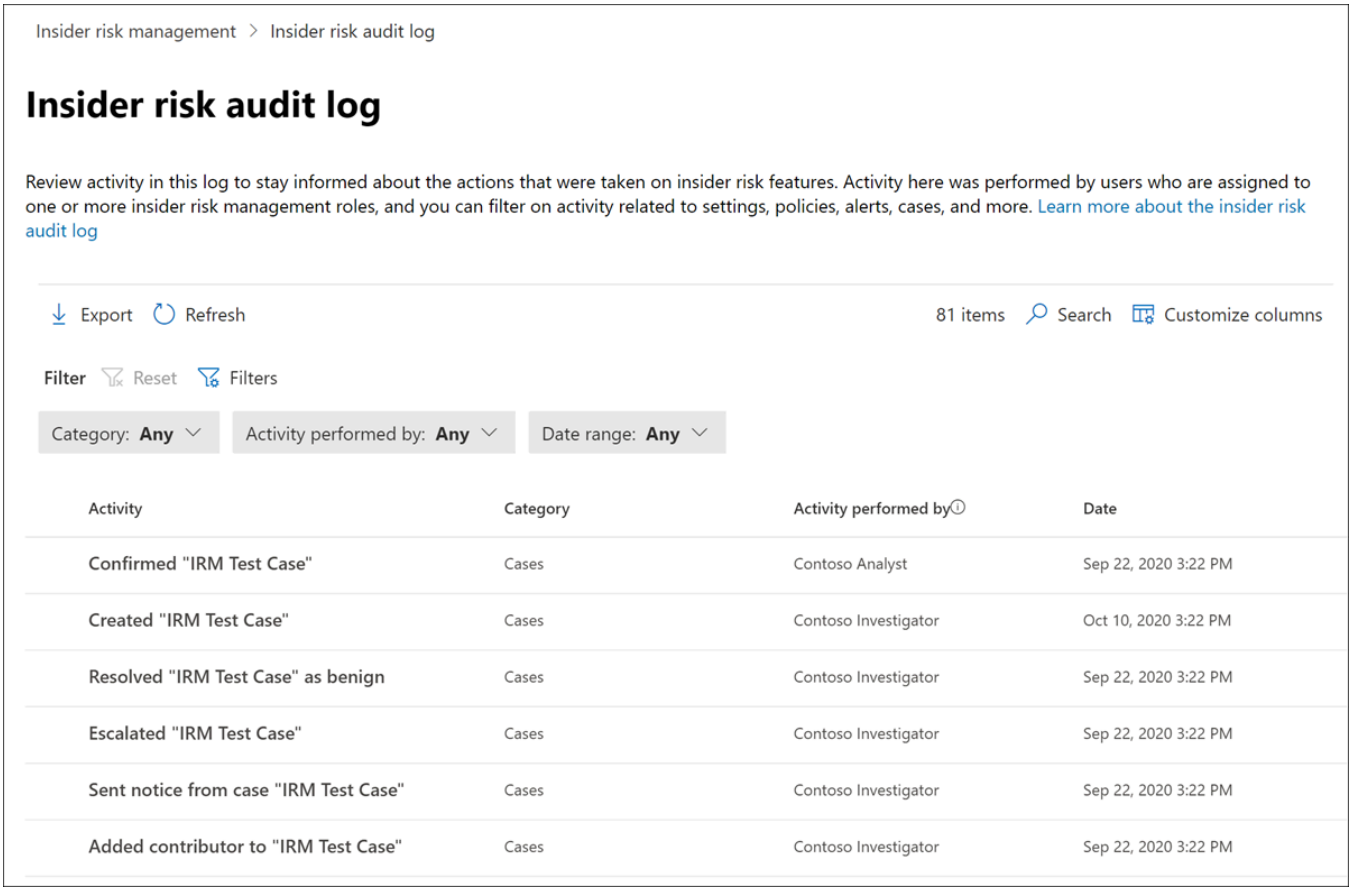
Last updated on 06/26/2026

Review activities with the Insider Risk Management audit log

08/20/2025

> **Important**
>
> Microsoft Purview Insider Risk Management correlates various signals to identify potential malicious or inadvertent insider risks, such as IP theft, data leakage, and security violations. Insider Risk Management enables customers to create policies to manage security and compliance. Built with privacy by design, users are pseudonymized by default, and role-based access controls and audit logs are in place to help ensure user-level privacy.

The Insider Risk Management audit log helps you stay informed about actions taken on Insider Risk Management features. This log makes it possible to independently review the actions of users assigned to one or more Insider Risk Management role groups. Your organization automatically enables the Insider Risk Management audit log and you can't disable it.



The audit log automatically and immediately updates whenever it detects identified risk activities. The audit log keeps information for 180 days (about six months). After 180 days, the log permanently deletes data.

Areas in identified risk activity detection include:

- Policies
- Cases
- Alerts
- Settings
- Users
- Notice templates

To view and export data from the audit log, assign users to the *Insider Risk Management* or *Insider Risk Management Auditors* role groups. For more information about Insider Risk Management role groups, see Assign Insider Risk Management permissions.

> **Note**
>
> The Insider Risk Management audit log isn't associated with the Microsoft 365 audit log, as they're independent auditing systems that capture information on separate areas. Disabling Microsoft 365 auditing doesn't impact activity auditing within Insider Risk Management.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

View activity in the insider risk audit log

To view feature activity detected for Insider Risk Management, go to the **Insider risk audit log** link in the top-right area of any Insider Risk Management tab. By default, you see the following information displayed for Insider Risk Management activities:

- **Activity:** A description of the identified risk activity taken within the Insider Risk Management solution by a user.
- **Category:** The area or item where the identified risk activity was performed. For example, you see *Policies* as the category when policy change activities were performed.
- **Activity performed by:** The user name of the user that performed the identified risk activity.
- **Date:** The date and time the identified risk activity was performed. The date and time are the local date and time for your organization.

For more information about a logged activity, select the activity to display the activity details pane. This pane includes additional information about the identified risk activity.

Columns and filtering

To make it easier for auditors to review audit logs, filtering is supported in the **Insider risk audit log**. For basic filtering, queue columns are available to add to the view to provide different pivots on the files and messages. You can filter identified risk activities by the **Category, Date range,** and **Activity performed by** fields.

To add or remove column headings for the queue, use the **Customize columns** control and select from the column options. These columns map to common conditions supported in the **Insider risk audit log** and are listed later in this article.

Audit log export

Users assigned to the *Insider Risk Management* or *Insider Risk Management Auditors* role groups can export audit log activity to a .csv (comma-separated values) file by selecting **Export** on the **Insider risk audit log** page. Depending on the audit log activity, some fields might not be included in the filtered queue, and those fields appear as blank in the exported file.

The file contains audit log activity information for the following fields:

- **Activity performed by:** Name of the user modifying an item value. Users listed here are assigned to one or more of the following Insider Risk Management role groups: *Insider Risk Management*, *Insider Risk Management Admins*, *Insider Risk Management Analysts*, *Insider Risk Management Investigators*. Each role group has different permission levels for managing insider risk features.
- **Activity:** Type of activity taken on an item. Values are *Viewed, Deleted, Added, Edited policy, Case, User, Alert,* and *Settings*.
- **Added**: Objects you add during the identified risk activity, such as users, file types, or domains.
- **Alert volume**: Level of alert volume defined in Insider Risk Management settings.
- **Amount**: Currently selected custom indicator amounts for a policy.
- **Asset ID**: Asset ID of the priority physical asset the activity was performed on.
- **Category:** Category of the item modified. Values are *Policies, Cases, Users, Alerts, Settings,* and *Notice templates*.
- **Date:** Date and time, listed in your organization's local date and time.
- **Description**: Description input by the user for the object being acted on (such as a policy or a priority user group).

- **DLP policy**: The Microsoft Purview Data Loss Prevention (DLP) policy selected to trigger inclusion in an Insider Risk Management policy.
- **Indicator**: Indicator in the within insider risk settings that the activity was performed on (such as adding or removing an indicator).
- **Notice template**: Notice template that the identified risk activity was performed on.
- **Number of days**: Policy activation window defined in insider risk settings.
- **Number of files**: File volume limit defined in Insider Risk Management settings.
- **Policy template**: Policy template that the indicators acted on belongs to.
- **Previous amount**: Previously selected custom indicator amounts for a policy.
- **Priority user group**: Priority user group the identified risk activity was performed on.
- **Removed**: Objects you remove during the identified risk activity, such as users, file types, or domains.
- **Sender**: Sender field of the notice template the identified risk activity was performed on.
- **Target policy**: The policy the identified risk activity was performed on (such as adding a user to or removing a user from).
- **Template message body**: The message body of the notice template the identified risk activity was performed on.
- **Template subject**: The subject field of the notice template the identified risk activity was performed on.
- **User:** User the identified risk activity was performed on.

Review data with the Insider Risk Management content explorer

10/01/2025

> **Important**
>
> Microsoft Purview Insider Risk Management correlates various signals to identify potential malicious or inadvertent insider risks, such as IP theft, data leakage, and security violations. Insider Risk Management enables customers to create policies to manage security and compliance. Built with privacy by design, users are pseudonymized by default, and role-based access controls and audit logs are in place to help ensure user-level privacy.

The Insider Risk Management **Content explorer** enables users assigned the *Insider Risk Management Investigators* role to examine the context and details of content associated with activity in alerts. The case data in Content explorer refreshes daily to include new risk activity. For all alerts that you confirm to a case, the solution archives copies of data and message files as a snapshot in time of the items, while maintaining the original files and messages in the storage sources. If needed, you can export case data files as a portable document file (PDF) or in the original file format.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Use the Content explorer to view details for a specific case

To examine the emails and files captured by the policies included in a specific case, go to the Insider Risk Management **Cases** page and select the row of the *Case name* in the list for the case you want to view details for. Then in the case details page, select the **Content explorer** tab to open the Content explorer.

> **Important**

After you confirm an alert to a case, Content explorer doesn't display any details for that case if the organization hasn't assigned a user to either the *Insider Risk Management Investigators* or *Insider Risk Management* role group.

Content explorer includes user activities related to Microsoft 365 service files, such as user activity on SharePoint, Exchange, and OneDrive for Business. For new cases, it usually takes about an hour for content to populate in Content explorer. For cases with large amounts of content, it might take longer to create a snapshot. If content is still loading in Content explorer, you see a progress indicator that displays the completion percentage.

In some cases, you might not be able to review data associated with a case as a snapshot in Content explorer. This situation might occur when you delete or move case data, or when a temporary error occurs when processing case data. If this situation occurs, select **View files** in the warning bar to view the file names, file path, and reason for the failure for each file. If needed, you can export this information to a .csv (comma-separated values) file.

If the content includes Information Rights Management permissions, these permissions are maintained for the copied content and users assigned the *Insider Risk Management Investigators* role need these permissions and rights if they need to open and view the files. Each file and message are automatically assigned a unique file ID in the Insider Risk Management case for management purposes. Documents associated with device indicator activities aren't included in Content explorer.

Note

After a case is active for over 30 days, Content explorer isn't updated to reflect new activities. To update content explorer with new activities for a user in this scenario, resolve the case and open a new case for the user.

View options

Source view

The **Source** viewer displays the richest view of a selected document. It supports hundreds of file types and is meant to display the truest to native experience possible. For Microsoft Office files, the viewer uses the web version of Office apps to display content such as document comments, Microsoft Teams chats, Excel formulas, hidden rows and columns, and PowerPoint notes.

Plain text view

The **Plain text** viewer (preview) provides a view of the extracted text of a selected file. It ignores any embedded images and formatting but is useful if you want to understand the content quickly. The text view also includes these features:

- Line counter makes it easier to reference specific portions of a document
- Search hit highlighting that highlights terms within the document and in the scrollbar

Annotate view

The **Annotate** view (preview) provides features that let you apply markup on a selected document, including the ability to:

- **Select annotations**: Select annotations on a document to delete
- **Select text**: Select text on the document to delete
- **Area redactions**: Draw a box on the document to hide sensitive content
- **Pencil**: Free-hand draw in selectable colors on a document to bring attention to certain portions of a document
- **Toggle annotation transparency**: Make annotations semi-transparent to view the content behind the annotation
- **Previous page**: Navigate to previous page
- **Next page**: Navigate to the next page
- **Go to page**: Enter a specific page number to navigate to
- **Zoom**: Set the zoom level for annotate view
- **Rotate**: Rotate the document clockwise

Metadata view

The panel in the **Metadata** view (preview) displays various metadata associated with the selected document. Search for specific metadata by using the **Search metadata name** field.

You can toggle the panel in the **Metadata** view on or off to display various metadata associated with the selected document. Although you can customize the search results grid to display specific metadata, there are instances where scrolling horizontally can be difficult while reviewing data. The File metadata panel allows you to toggle on a view within the viewer.

Group options

Use the **Group** control to view content grouped by the following options:

- **Group by families**: All items related to a specific file are grouped together using the same Group ID. For example, if you have a PowerPoint file in the alert that includes imbedded images or .zip files, these images and files are grouped with the PowerPoint file and shown as nested items with the file in the item list view. Microsoft 365 Copilot and Microsoft 365 Copilot Chat user prompts and responses are also grouped together using the same Group ID when there are multiple prompts and responses associated with a file.
- **Group by conversations and related items**: All email messages, Teams conversations, and Viva Engage conversations are grouped using the same Thread ID and appear as nested items. Additionally, all associated content for these messages and conversations is also grouped together. For example, if you have an email conversation that includes several email messages, some of which include attachments and some that include embedded images, all of the email messages, attachments, and images are grouped together in the alert list view under an applicable item.

Column options

To help risk analysts and investigators review captured data and messages, the Content explorer includes several filtering and sorting tools. For basic sorting, the **Date** and **File class** columns support sorting by using the column titles in the content queue pane. You can add other queue columns to the view to provide different pivots on the files and messages.

To add, remove, or reorder column headings for the content queue, use the **Customize columns** control and select from the following column options. These columns map to the common, email, and document property conditions supported in the Content explorer.

View results in UTC or your local time by toggling the **Time zone** option in the action bar.

Expand table

Column option	Description
Author	The author field from Office documents, which persists if a document is copied. For example, if a user creates a document and the emails it to someone else who then uploads it to SharePoint, the document still retains the original author.
Bcc	Available for email messages, the users in the Bcc message field.
Cc	Available for email messages, the users in the Cc message field.
Compound path	Human readable path that describes the source of the item.
Conversation ID	Conversation ID from the message.
Conversation index	Conversation index from the message.

Column option	Description
Created time	The time the file or email message was created.
Date (UTC)	For email, the date a message was received by a recipient or sent by the sender. For documents, the date a document was last modified. Date is in Coordinated Universal Time (UTC).
Dominant theme	Dominant theme as calculated for analytics.
Email set ID	Group ID for all messages in the same email set.
Family ID	Family ID groups together all items; for email, this column includes the message and all attachments; for documents, this column includes the document and any embedded items.
File class	For content from SharePoint and OneDrive: **Document**; for content from Exchange: **Email** or **Attachment**.
File ID	Document identifier unique within the case.
File type icon	The extension of a file; for example, docx, one, pptx, or xlsx. This field is the same property as the FileExtension site property.
ID	The GUID identifier for the file.
Immutable ID	Immutable ID as stored in Office 365.
Inclusive type	Inclusive type calculated for analytics: **0** - not inclusive; **1** - inclusive; **2** - inclusive minus; **3** - inclusive copy.
Last modified	The date that a document was last changed.
Marked as representative	One document from each set of exact duplicates is marked as representatives.
Message kind	The type of email message to search for. Possible values: contacts, docs, email, external data, faxes, im, journals, meetings, microsoft teams (returns items from chats, meetings, and calls in Microsoft Teams), notes, posts, RSS feeds, tasks, voicemail
Participants	List of all participants of a message; for example, Sender, To, Cc, Bcc.
Pivot ID	The ID of a pivot.
Received	The date that an email message was received by a recipient. This field is the same property as the Received email property.
Recipients	All recipient fields in an email message. These fields are To, Cc, and Bcc.
Representative ID	Numeric identifier of each set of exact duplicates.

Column option	Description
Sender	The sender of an email message.
Sender/Author	For email, the person who sent a message. For documents, the person cited in the author field from Office documents. You can type more than one name, separated by commas. Two or more values are logically connected by the OR operator.
Sensitive info types	The sensitive info types identified in content.
Sensitivity labels	The sensitivity labels applied to the content.
Sent	The date that an email message was sent by the sender. This field is the same property as the Sent email property.
Size	For both email and documents, the size of the item (in bytes).
Subject	The text in the subject line of an email message.
Subject/Title	For email, the text in the subject line of a message. For documents, the title of the document. As previously explained, the Title property is metadata specified in Microsoft Office documents. You can type the name of more than one subject/title, separated by commas. Two or more values are logically connected by the OR operator.
Themes list	Themes list as calculated for analytics.
Title	The title of the document. The Title property is metadata that's specified in Office documents. It's different than the file name of the document.
To	The recipient of an email message in the To field.

Filtering

Use one or more filters to narrow the scope of a search and return a more refined set of results. To set a filter, select **Query** (preview) at the top of the content queue. Many filters include additional conditions to help narrow the results returned by the filter. For example, the *Date* filter includes controls to configure a *Start date* and *Ending date* for the **Date** filter. Save queries for future use by selecting **Save filters** and following the prompts. Select one or more filter items from the following example categories:

Common filters

Expand table

Filter	Description
Date (UTC)	For email, the date a message was received by a recipient or sent by the sender. For documents, the date a document was last modified.
Sender/Author	For email, the person who sent a message. For documents, the person cited in the *Author* field from Office documents. Type more than one name, separated by commas.
Source	The location of the document in your organization. For example, a specific SharePoint site location.
Subject/Title	For email, the text in the subject line of a message. For documents, the title of the document. The Title property in documents is metadata specified in Microsoft Office documents. Type the name of more than one subject/title, separated by commas. Two or more values are logically connected by the OR operator.

Email filters

Expand table

Filter	Description
Bcc	The Bcc field of an email message.
Cc	The Cc field of an email message.
Has attachment	Indicates whether a message has an attachment. Values are listed as **true** or **false**.
Is email attachment	If the document is an attachment, the value is listed as **Yes**.
Is embedded document	If the document is embedded in the email message, the value is listed as **Yes**.
Is inline attachment	If the document is an inline attachment in the email message, the value is listed as **Yes**.
Participants	All the people fields in an email message. These fields are From, To, Cc, and Bcc.
Received	The date that an email message was received by a recipient.
Recipient domains	List of all domains of recipients of a message.
Recipients	The email message recipients.
Sender domain	Domain of the sender.
Sender	Sender (From) field for message types. Format is **DisplayName <SmtpAddress>**.

Filter	Description
To	The To field of an email message.
Unique in email set	False if there's a duplicate of the attachment in its email set.

Document filters

Expand table

Filters	Description
Compliance labels	Compliance labels applied in Microsoft 365.
Created time (UTC)	The date and time the file or email message was created. The date and time are in Coordinated Universal Time (UTC).
File extension	The extension type of the file.
Last modified date (UTC)	The date that a document was last changed. The date and time are in Coordinated Universal Time (UTC).
User activity events	Activity for items related to specific user activity in a case. For example, when you select a link to 'Explore Content' for an activity in the **User Activity** page of a case, this filter is used to display items related to that activity.
Work product	The type of work product for the document. For example, annotations or tags in the document.

Advanced filters

Advanced filters (preview) let you build more flexible and granular filters. This advanced filtering capability enables you to:

- Quickly search for filter conditions.
- Create complex filters using subgroups, *AND*, or *OR* conditions.
- Easily change your queries with **Clear all** and **Reset all** controls.
- Manage saved filters without having to navigate to another area.
- Use *Is empty* and *Is not empty* conditions for each filter.

Filter types

Every searchable field in an alert has a corresponding filter that you can use for filter items based on a specific field.

There are multiple types of filters:

- **Freetext**: A freetext filter is applied to text fields such as *Subject*. You can list multiple search terms by separating them with a comma.
- **Date**: A date filter is used for date fields such as *Last modified date*.
- **Search options**: A search options filter provides a list of possible values (each value is displayed with a checkbox that you can select) for particular fields in the review. This filter is used for fields, such as *Sender*, where there's a finite number of possible values in the alert.
- **Keyword**: A keyword condition is a specific instance of freetext condition that you can use to search for terms. You can also use Keyword Query language (KeyQL) in this type of filter. For more information, see the [Query language](#
) and [Advanced query builder](#
) sections in this article.

Advanced filter controls

To create and custom filtering for your search results, use the following controls:

- **AND/OR**: These conditional logical operators allow you to select the query condition that applies to specific filters and filter subgroups. These operators allow you to use multiple filters or subgroups connected to a single filter in your query.
- **Add conditions**: Allows you to add multiple filtering conditions to your query. Is available after you've defined at least one query filter.
- **Select an operator**: Depending on the selected filter, the operators compatible for the filter are available to select. For example, if the *Date* filter is selected, the available operators are *Before*, *After*, and *Between*. If the *Size (in bytes)* filter is selected, the available operators are *Greater than*, *Greater or equal*, *Less than*, *Less or equal*, *Between*, and *Equal*. If it's a file name, ID, or keyword, the available operators are *Equal any of*, *Equal none of*, *Contains all of*, *Contains any of*, *Contains none of*, *Is empty*, or *Is not empty*.
- **Value**: Depending on the selected filter, the values compatible for the filter are available. Additionally, some filters support multiple values and some filters support one specific value. For example, if the *Date* filter is selected, select date values. If the *Size (in bytes)* filter is selected, select a value for bytes.
- **Add subgroup**: After you've defined a filter, you can add a subgroup to refine the results returned by the filter. You can also add a subgroup to a subgroup for multi-layered query refinement.
- **Save**: Allows you to save the current set of filters, conditions, and operators for future use. These saved queries are user-specific.
- **Remove a filter condition**: To remove an individual filter or subgroup, select the remove icon to the right of each filter line or subgroup.

- **Clear all**: To clear the entire query of all filters and subgroups, select **Clear all**.

Help dynamically mitigate risks with Adaptive Protection

> **Important**
>
> Microsoft Purview Insider Risk Management correlates various signals to identify potential malicious or inadvertent insider risks, such as IP theft, data leakage, and security violations. Insider Risk Management enables customers to create policies to manage security and compliance. Built with privacy by design, users are pseudonymized by default, and role-based access controls and audit logs are in place to help ensure user-level privacy.

Adaptive protection in Microsoft Purview uses machine learning to identify the most critical risks and proactively and dynamically apply protection controls from:

- Microsoft Purview Data Loss Prevention (DLP)
- Microsoft Purview Data Lifecycle Management (preview)
- Microsoft Entra Conditional Access (preview)

Integration with data loss prevention, data lifecycle management, and Conditional Access helps organizations automate their response to insider risks and reduces the time required to identify and remediate potential threats. By using the capabilities of all four solutions, organizations can create a more comprehensive security framework that addresses both internal and external threats.

> **Important**
>
> To learn about licensing requirements for using Adaptive Protection with data loss prevention, data lifecycle management, and Conditional Access, see Microsoft 365, Office 365, Enterprise Mobility + Security, and Windows 11 Subscriptions.

Adaptive protection helps mitigate potential risks by using:

- **Context-aware detection**. Helps identify the most critical risks with ML-driven analysis of both content and user activities.
- **Dynamic controls**. Helps enforce effective controls on high-risk users while others maintain productivity.

- **Automated mitigation**. Helps minimize the impact of potential data security incidents and reduce admin overhead.

Adaptive protection dynamically assigns appropriate data loss prevention, data lifecycle management, and Conditional Access policies to users based on the insider risk levels defined and analyzed by the machine learning models in Insider Risk Management. Policies become adaptive based on user context, ensuring that the most effective policy, such as blocking data sharing through data loss prevention or blocking application access through Conditional Access, is applied only to high-risk users while low-risk users maintain productivity. Data loss prevention and Conditional Access policy controls constantly adjust, so when a user's insider risk level changes, an appropriate policy dynamically applies to match the insider new risk level. In the case of data lifecycle management, Insider Risk Management detects *Elevated risk level* users and preserves their deleted data for 120 days by using an automatically created data lifecycle management retention label policy.

> **Important**
>
> Insider Risk Management is currently available in tenants hosted in geographical regions and countries supported by Azure service dependencies. To verify that the Insider Risk Management solution is supported for your organization, see Azure dependency availability by country/region. Insider Risk Management is available for commercial clouds but isn't available for US Government cloud programs at this time.

Watch the following video for a summary of how Adaptive Protection can help identify and mitigate the most critical risks in your organization:

https://learn-video.azurefd.net/vod/player?id=9487d97b-4306-4a17-abb8-1008d4c63325&locale=en-us&embedUrl=%2Fpurview%2Finsider-risk-management-adaptive-protection

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Insider risk levels and preventive controls

With Adaptive Protection, administrators can configure risk factors or activities for customizable insider risk levels based on your organization's needs. The insider risk levels for Adaptive Protection update continuously and automatically based on the users' risk factors and insights. When users' data security risks increase or decrease, their insider risk levels adjust accordingly. Based on the insider risk levels, data loss prevention policies and Conditional Access policies automatically apply the right level of preventative controls as configured by admins, such as *block*, *block with override*, or *warning*.

> **Tip**
>
> *Insider risk levels* in Adaptive Protection differ from *alert severity levels* that assign Insider Risk Management users based on activity detected in Insider Risk Management policies.
>
> - Insider risk levels (Elevated, Moderate, or Minor), as described in this article, measure risk determined by admin-defined conditions such as the number of exfiltration activities users perform in a day or whether their activity generated a high-severity insider risk alert.
> - Alert severity levels (Low, Medium, or High) assign to users based on activity detected in Insider Risk Management policies. These levels calculate based on alert risk scores assigned to all active alerts associated with the user. These levels help insider risk analysts and investigators prioritize and respond to user activity accordingly.

For data lifecycle management, the data lifecycle management policy monitors users that assign an *Elevated risk level* by Insider Risk Management Adaptive Protection. When a risky user deletes any content from SharePoint, OneDrive, or Exchange Online, the contents are automatically preserved for 120 days. Admins can contact Microsoft support to restore any preserved content.

Depending on the Insider Risk Management policy assigned in Adaptive Protection, different criteria (users, groups, indicators, thresholds, etc.) are used to determine applicable insider risk levels. Insider risk levels are based on user insights, not solely on the number of instances of specific user activities. Insights are a calculation of the aggregate number of activities and the severity level of these activities.

For example, insider risk levels for User A wouldn't be determined by User A performing a potentially risky activity more than three times. The insider risk levels for User A would be determined by an insight of the aggregate number of activities and risk scores would be assigned to the activity based on the thresholds configured in the selected policy.

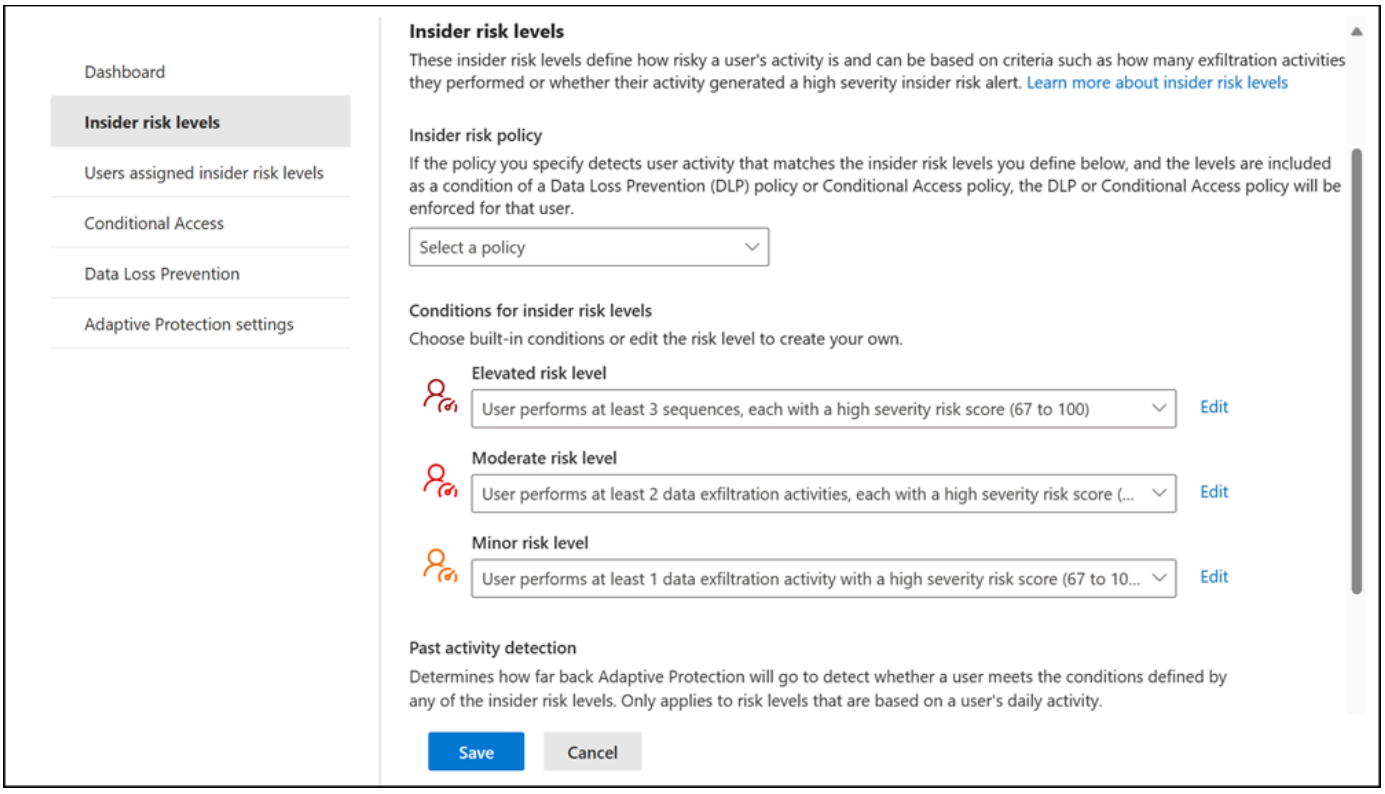
Insider risk levels

Insider risk levels in Adaptive Protection define how risky a user's activity is and can base on criteria such as how many exfiltration activities they performed or whether their activity generated a high severity insider risk alert. These insider risk levels have built-in insider risk level definitions, but you can customize these definitions as needed:

- **Elevated risk level**: This risk level includes built-in definitions for users with high severity alerts, users with at least three sequence insights that each have a high severity alert for specific risk activities, or one or more confirmed high severity alerts.
- **Moderate risk level**: This risk level includes built-in definitions for users with medium severity alerts or users with at least two data exfiltration activities with high severity scores.
- **Minor risk level**: This risk level includes built-in definitions for users with low severity alerts or users with at least one data exfiltration activity with a high severity score.

For an insider risk level to assign to a user, the number of insights and the severity assigned to the activity need to match the definition for the insider risk level. The number of activities for an insight might be a single activity or multiple activities accruing to the single insight. The number of insights evaluates for the insider risk level definition, not the number of activities contained in an insight.

For example, suppose the conditions in the Insider Risk Management policy assigned to Adaptive Protection scope for identifying downloads from SharePoint sites in your organization. If the policy detects that a user downloaded 10 files from a SharePoint site in a single day that are determined to be high severity, this counts as a single insight that consists of 10 activity events. To qualify for assigning an *Elevated risk level* to the user, two additional insights (with high severity) require for the user. The additional insights might or might not contain one or more activities.



Customizing insider risk levels

Custom insider risk levels enable you to create insider risk levels that fit your organization's needs. You can customize the criteria for the insider risk level and define conditions for when to assign the insider risk level to users.

Consider the following examples for using Adaptive Protection together with data loss prevention, data lifecycle management, and Conditional Access policies.

- **Data loss prevention policies**:
  - Allow users with the *Minor risk level* or *Medium risk level* to receive policy tips and education on best practices for handling sensitive data. With this approach, you can influence positive behavior changes over time and reduce organizational data risks.
  - Block users with the *Elevated risk level* from saving or sharing sensitive data to minimize the impact of potential data incidents.
- **Data lifecycle management policy**:
  - Preserve any SharePoint, OneDrive, or Exchange Online content deleted by a risky user for 120 days. A risky user in this case is a user that Adaptive Protection assigns the *Elevated risk level*.
- **Conditional Access policies**:
  - Require *Minor risk level* users to acknowledge Terms of Use before using an application.
  - Block *Medium risk level* users from accessing certain applications.

- Completely block *Elevated risk level* users from using *any* applications. Learn more about commonly applied Conditional Access policies

Insider risk level criteria and conditions

You can customize insider risk level criteria and conditions based on the following areas:

- **Alerts generated or confirmed for a user**: Choose conditions based on the severity level for alerts that the selected Insider Risk Management policy generates or confirms for a user. Conditions for alerts aren't additive. The insider risk level is assigned to a user if one of the conditions is met.
- **Specific user activity**: Choose conditions for activity to detect, its severity, and the number of daily occurrences during the past activity detection window (optional). Conditions for user activity are additive. The insider risk level is assigned to a user only if all the conditions are met.

Past activity detection

The **Past activity detection** setting determines how many days back Adaptive Protection examines to detect whether a user meets the conditions defined by any of the insider risk levels. The default setting is 7 days, but you can choose between 5 and 30 days of previous activity to apply insider risk level conditions. This setting only applies to insider risk levels that are based on a user's daily activity and excludes insider risk levels based on alerts.

The following example illustrates how past activity detection settings and insider risk levels interact to determine if a user's past activity is in-scope:

- **Elevated risk level** setting: User performs at least three sequences, each with a high severity risk score (67 to 100)
- **Past activity detection** setting: 3 days

Expand table

User activity	Activity in-scope for insider risk level
User has 1 high severity sequence each day on Day T-3, T-2, T-1	Yes
User has 1 high severity sequence on Day T-4 and 2 high severity sequences on Day T-3	No
User has 3 high severity sequences on Day T-3	Yes

Insider risk level timeframe

The **Insider risk level timeframe** setting determines how long an insider risk level remains assigned to a user before it's automatically reset. The default setting is 7 days, but you can choose between 5 and 30 days before resetting the insider risk level for a user.

Insider risk levels also reset for a user when:

- The associated alert for the user is dismissed.
- The associated case for the user is resolved.
- The insider risk level end date is manually expired.

> **Note**
>
> If a user is currently assigned an insider risk level and that user meets the criteria again for that insider risk level, the insider risk level timeframe extends for the defined number of days for the user.

Risk level expiration options

When you enable the automatic expiration option (enabled by default), Adaptive Protection risk levels for a user automatically expire when the associated alert for a user is dismissed or the associated case for the user is closed. If you want to keep the Adaptive Protection risk level for a user even if the alert is dismissed or the case is closed, disable the option.

Permissions for Adaptive Protection

> **Important**
>
> To learn about licensing requirements for using Adaptive Protection with data loss prevention, data lifecycle management, and Conditional Access, see Microsoft 365, Office 365, Enterprise Mobility + Security, and Windows 11 Subscriptions.

Depending on how you use Insider Risk Management built-in role groups and role groups for data loss prevention or Conditional Access, you might need to update Insider Risk Management permissions for administrators, analysts, and investigators in your organization.

The following table describes the permissions required for specific Adaptive Protection tasks.

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

Expand table

Task	Required role group
Configure Adaptive Protection and update settings	*Insider Risk Management* or *Insider Risk Management Admins*
Create and manage Conditional Access policies with the Adaptive Protection condition	One of the following: *Global administrator*, *Conditional Access Administrator*, *Security Administrator*
Create and manage data loss prevention policies with the Adaptive Protection condition	One of the following: *Compliance Administrator*, *Compliance Data Administrator*, *DLP Compliance Management*, *Global Administrator*
View details on users' or agents' assigned insider risk levels	*Insider Risk Management*, *Insider Risk Management Analysts*, or *Insider Risk Management Investigators*

> **Important**
>
> The four categories of role groups correspond to the following tabs on the Adaptive Protection page: **Insider risk levels**, **Users assigned insider risk levels**, **Data Loss Prevention**, **Conditional Access**. If you're not assigned to the appropriate role group, the tab doesn't appear on the **Adaptive protection** page.

Learn more about role groups in Microsoft Defender for Office 365 and Microsoft Purview compliance

Configure Adaptive Protection

Depending on the needs of your organization or where you're currently configured with Insider Risk Management, data loss prevention, data lifecycle management, and Conditional Access, you have two options to get started with Adaptive Protection:

- Quick setup

- Custom setup

Quick setup

The quick setup option is the fastest way to get started with Adaptive Protection. With this option, you don't need any preexisting Insider Risk Management, data loss prevention, data lifecycle management, or Conditional Access policies. You also don't need to preconfigure any settings or features. If your organization doesn't have a current subscription or license that supports Insider Risk Management, data loss prevention, or data lifecycle management, sign up for a Microsoft Purview risk and compliance solutions trial before starting the quick setup process. You can also sign up for a Microsoft Entra trial for Conditional Access.

To get started, select **Turn on Adaptive Protection** from the Adaptive Protection cards on the Microsoft Purview portal home page or the data loss prevention Overview page. You can also start the quick setup process by going to **Insider Risk Management** > **Adaptive protection** > **Dashboard** > **Quick setup**.

> **Note**
>
> If you're already a scoped admin for Microsoft Purview, you can't turn on quick setup.

Here's what the quick setup process configures for Adaptive Protection:

Expand table

Area	Configuration
A new insider risk policy	- Policy template: Data leaks - Policy name: Adaptive Protection policy for Insider Risk Management - Policy scope for users and groups: All users and groups - Priority content: None - Triggering events: Selected exfiltration events (you can view in Insider Risk Management settings) - Policy indicators: A subset of Office indicators (you can view in Insider Risk Management settings) - Risk score boosters: Activity is over user's usual activity for that day
Adaptive protection insider risk levels	- *Elevated risk level*: Users must have at least three high severity exfiltration sequences - *Moderate risk level*: Users must have at least two high severity activities (excluding some types of downloads) - *Minor risk level*: Users must have at least one high severity activity (excluding some types of downloads)

Area	Configuration
Insider risk settings (if already configured)	- Policy indicators: Office indicators not already configured (you can view in Insider Risk Management settings) - All other settings previously configured aren't updated or changed. - Analytics: On (thresholds for triggering events in policies are the default settings determined by Analytics recommendations)
Insider risk settings (if not already configured)	- Privacy: Show anonymized versions of user names. **Note:** User names aren't anonymized in Conditional Access or data loss prevention. - Policy timeframes: Defaults - Policy indicators: A subset of Office indicators (you can view in Insider Risk Management settings) - Risk score boosters: All - Intelligent Detections: Alert Volume = Default volume - Analytics: On - Admin notifications: Send notification email when first alert is generated to all
New Conditional Access policy (created in Report-only mode so users won't be blocked)	**1-Block access for users with Insider Risk (Preview)** - Included users: All users - Excluded guest or external users: B2bDirectConnect User; OtherExternalUser; ServiceProvider - Cloud apps: Office 365 apps - Insider risk levels: *Elevated* - Block access: Selected
New data lifecycle management policy	An organization-wide autoapply label policy that monitors users assigned the *Elevated risk level*. Any SharePoint, OneDrive, or Exchange Online content deleted by a risky user is preserved for 120 days. If you haven't already turned on Adaptive Protection, the policy is automatically created and applied when you turn on Adaptive Protection. If you have already turned on Adaptive Protection, you need to explicitly opt in to the data lifecycle management policy that's created automatically.
Two new data loss prevention policies	Adaptive Protection policy for Endpoint DLP - *Elevated risk level* rule: Blocked - *Moderate/Minor risk level* rule: Audit - Policy starts in test mode (audit only) Adaptive Protection policy for Teams and Exchange DLP - *Elevated risk level* rule: Blocked - *Moderate/Minor risk level* rules: Audit - Policy starts in test mode (audit only)

After you start the quick setup process, it can take up to 72 hours before the following items are completed:

- Analytics
- Associated Insider Risk Management, data loss prevention, data lifecycle management, and Conditional Access policies
- Adaptive Protection insider risk levels, data loss prevention, data lifecycle management, and Conditional Access actions applied to applicable user activities.

> **Note**
>
> Once you enable Adaptive Protection, don't disable it before the setup process completes. Disabling Adaptive Protection before the setup process completes might lead to policy errors.

Administrators receive a notification email when the quick setup process completes.

Custom setup

The custom setup option lets you customize the Insider Risk Management policy, the insider risk levels, and the data loss prevention and Conditional Access policies configured for Adaptive Protection.

> **Note**
>
> For data lifecycle management, if you don't turn on Adaptive Protection, the policy isn't created and applied automatically. If you turn on Adaptive Protection for your organization, you need to **explicitly opt in to apply the data lifecycle management policy**.

The custom setup option also lets you configure these items before enabling the Adaptive Protection connections between Insider Risk Management and data loss prevention. In most cases, use this option if your organization already has Insider Risk Management and data loss prevention policies in place.

Complete the following steps to configure Adaptive Protection by using the custom setup.

Step 1: Create Insider Risk Management policy

Assign insider risk levels to users or agents when a policy assigned in Adaptive Protection either detects user or agent activity or generates alerts that match the insider risk level conditions you define in the next step. If you don't want to use an existing Insider Risk Management policy,

create a new Insider Risk Management policy. Your Insider Risk Management policy for Adaptive Protection should include:

- **Users whose activity you want to detect**. This group can include all users and groups in your organization or just a subset for specific risk mitigation scenarios or testing purposes.
- **Activities you consider risky and custom thresholds that influence an activity's risk score**. Risky activities might include emailing people outside your organization or copying files to USB devices.

Select **Create insider risk policy** to start the new policy workflow. The Data leaks policy template is automatically selected in the workflow, but you can select any policy template.

> **Important**
>
> Depending on the policy template you select, you might need to configure additional settings for the policy to properly detect potentially risky activities and to create applicable alerts.

Step 2: Configure insider risk level settings

Select the **Insider risk levels** tab. Start by selecting the Insider Risk Management policy you want to use for Adaptive Protection. This policy can be the new policy you created in Step 1 or an existing policy or policies that you already configured.

Next, accept the applicable built-in insider risk level conditions or create your own. Depending on the type of policy you select, the insider risk level conditions reflect the applicable conditions associated with indicators and activities you configure in the policy.

For example, if you choose a policy based on the Data leaks policy template, the built-in insider risk level condition choices apply to indicators and activities available in that policy. If you select a policy based on the Security policy violations policy template, the built-in insider risk level conditions automatically scope to indicators and activities available in that policy.

Customize an insider risk level for your policy

1. Sign in to the Microsoft Purview portal by using credentials for an admin account in your Microsoft 365 organization.
2. Go to the **Insider Risk Management** solution.
3. Select **Adaptive protection** in the left navigation, then select **Insider risk levels**.

4. On the **Insider risk levels** page, select **Edit** for the insider risk level you want to customize (*Elevated*, *Moderate*, or *Minor*).

5. On the **Custom insider risk level** pane, select an option in the **Insider risk level based on** section:

   - **Alert generated or confirmed for a user**
   - **Specific user activity**

6. If you select the **Alert generated or confirmed for a user** option, choose the severity levels for alerts that are generated or confirmed for a user that should use this insider risk level. You can keep the *Severity for generated alerts* and the *Severity for confirmed alerts* conditions or remove one of these conditions if you only want to use one. If you need to add one of these conditions back, select **Add condition**, then select the condition. For each condition, choose the severity level that should be applied for the condition (*High*, *Medium*, or *Low*). If the **any** of the conditions are met, the insider risk level is assigned to a user.

7. If you select the **Specific user activity** option, choose the activity to detect, its severity, and number of daily occurrences during the past activity detection window. You must configure the *Activities*, *Activity severity*, and *Activity occurrences during detection window* conditions for this insider risk level.

   For the *Activities* condition, the options you can choose from automatically update for the types of activities you define with the indicators configured in the associated policy. If needed, select the **Assign this insider risk level to any user who has a future alert confirmed, even if conditions above aren't met** checkbox. If the **all** of the conditions are met, the insider risk level is assigned to a user.

   For the *Activity severity* condition, specify the severity level for the activities included in the daily activity insight. The options are *High*, *Medium*, and *Low*, and are based on risk score ranges.

   For the *Activity occurrences during detection window* condition, specify the number of times selected activities must be detected within the specified *Past activity detection* period. This number isn't related to the number of events that might occur for an activity. For example, if the policy detects that a user downloaded 20 files from SharePoint in one day, that count is one daily activity insight consisting of 20 events.

8. Select **Confirm** to apply the custom insider risk level conditions.

How insider risk level is assigned if a user is in scope for multiple policies

If a user is in scope for multiple policies and receives alerts of different severity levels, the user is assigned the highest severity level received. For example, consider a policy that assigns an *Elevated risk level* if users receive a high severity alert. If a user receives a low-severity alert from policy 1, a medium-severity alert from policy 2, and a high-severity alert from policy 3, the user is assigned an *Elevated risk level* - the level for the highest alert severity received.

Insider risk level conditions must be present in selected policies to be detected. For example, if you select the *Copy to USB* activity to assign a *Moderate risk level*, but the activity is selected in just one of three selected policies, then only activity from that one policy assigns a *Moderate risk level* for that activity.

Step 3: Create or edit a data loss prevention policy

Next, create a new data loss prevention policy or edit an existing policy to restrict actions for users who match your insider risk level conditions in Adaptive Protection. Use the following guidelines for your data loss prevention policy configuration:

- You must include the *User's insider risk level for Adaptive Protection is* condition in your data loss prevention policy. This data loss prevention policy can include other conditions as needed.
- Although you can include other locations in the data loss prevention policy, Adaptive Protection currently supports only Exchange, Microsoft Teams, and devices.

Select **Create policy** to start the data loss prevention policy workflow and create a new data loss prevention policy. If you have an existing data loss prevention policy that you'd like to configure for Adaptive Protection, go to **Data loss prevention** > **Policies** in the Microsoft Purview portal and select the data loss prevention policy you'd like to update for Adaptive Protection. For guidance on how to configure a new data loss prevention policy or update an existing data loss prevention policy for Adaptive Protection, see Learn about Adaptive Protection in data loss prevention: Manual configuration.

> **Tip**
>
> We recommend testing the data loss prevention policy (with policy tips) so you can review data loss prevention alerts to verify that the policy is working as expected before enabling Adaptive Protection.

Step 4: Create or edit a Conditional Access policy

Next, create a new Conditional Access policy or edit an existing policy to restrict actions for users who match your insider risk level conditions in Adaptive Protection. Use the following guidelines for your Conditional Access policy configuration:

- On the Conditional Access page where you control access based on signals from conditions, set the *Insider risk* condition to *Yes*, and then select an insider risk level (*Elevated*, *Moderate*, or *Minor*). This is the insider risk level that users must have in order for the policy to be enforced.

Select **Create policy** to start the Conditional Access policy workflow and create a new Conditional Access policy. If you have an existing Conditional Access policy that you'd like to configure for Adaptive Protection, go to **Protection** > **Conditional Access** in the Microsoft Entra admin center and select the Conditional Access policy you'd like to update for Adaptive Protection. For guidance on how to configure a new Conditional Access policy or update an existing Conditional Access policy for Adaptive Protection, see Common Conditional Access policy: Insider risk based policy.

Step 5: Turn on Adaptive Protection

After you complete the Insider Risk Management policy, insider risk level, DLP policy, and Conditional Access policy configuration steps, you're ready to enable Adaptive Protection. When you turn on Adaptive Protection:

- The Insider Risk Management policy starts looking for user activity matching your insider risk level conditions. If detected, the insider risk levels are assigned to users.
- Users who are assigned insider risk levels appear on the **Users assigned insider risk levels** tab in Adaptive Protection.
- The data loss prevention policy applies protection actions for any user assigned to insider risk levels included in the data loss prevention policy. The data loss prevention policy is added to the **Data Loss Prevention** tab in **Adaptive protection**. You can view details about the data loss prevention policy and edit policy conditions from the dashboard.
- The data lifecycle management policy, which is automatically created, applies protection actions for any user assigned to the *Elevated risk level*. The following message appears with a green background on the **Adaptive Protection** tab to let users know that proactive data preservation is being applied: "Your organization is also being dynamically protected from users who might potentially delete critical data." The message also provides a link to the

data lifecycle management setting where you can turn off the proactive data preservation feature, if desired.

- The Conditional Access policy applies protection actions for any user assigned to insider risk levels included in the Conditional Access policy. The Conditional Access policy is added to the **Conditional Access policies** tab in **Adaptive protection**. You can view details about the Conditional Access policy and edit policy conditions from the dashboard.

To enable Adaptive Protection, select the **Adaptive Protection settings** tab, and then turn **Adaptive Protection** to *On*. It might take up to 36 hours before you can expect to see Adaptive Protection insider risk levels and data loss prevention, data lifecycle management, and Conditional Access actions applied to applicable user activities.

Watch the following video on the Microsoft Mechanics channel to see how Adaptive Protection can automatically adjust the strength of data protection based on calculated data security insider risk levels of users.

Manage Adaptive Protection

When you enable Adaptive Protection and configure your Insider Risk Management, data loss prevention, and Conditional Access policies, you can access information about policy metrics, users currently included, and insider risk levels currently in scope.

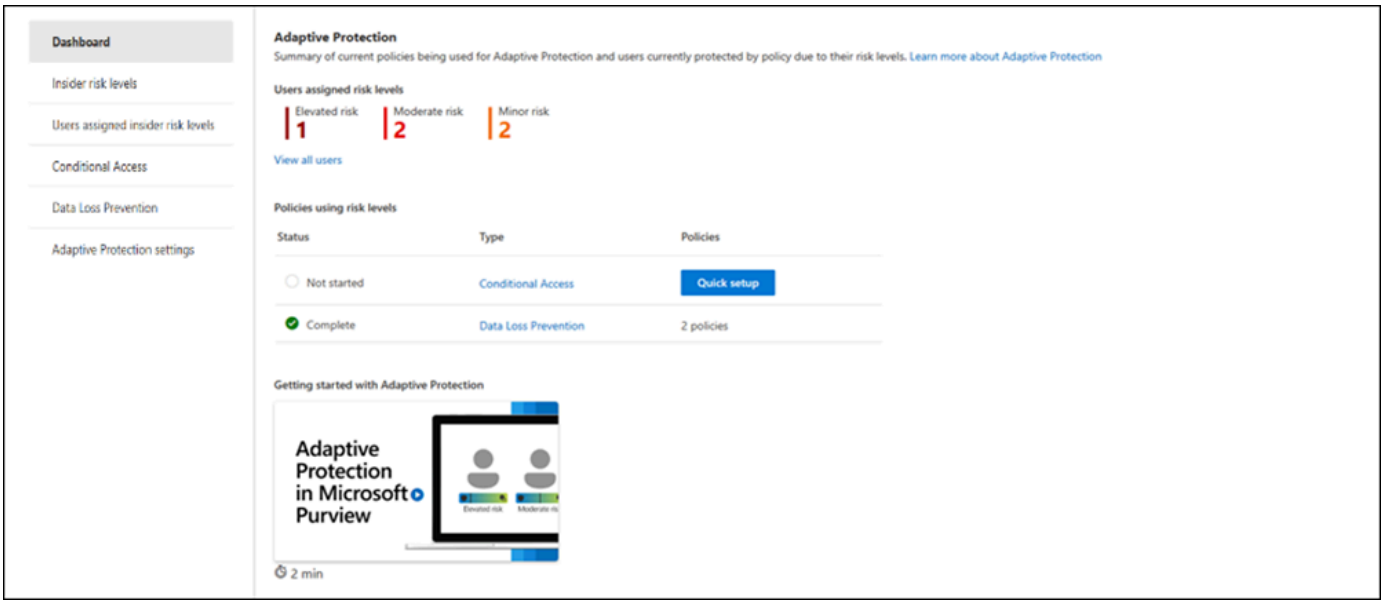
> **Note**
>
> Data lifecycle management metrics don't currently appear on the dashboard. However, if proactive data preservation is turned on, you see the following message on the **Adaptive Protection** tab: "Your organization is also being dynamically protected from users who might potentially delete critical data."

Use the Adaptive Protection dashboard

After you complete either the Quick setup or Custom setup process, the **Dashboard** tab in **Adaptive protection** displays widgets for summary information about user insider risk levels, Conditional Access policies, and data loss prevention policies.

- **Users assigned insider risk levels**: Shows the number of users for each insider risk level (*Elevated*, *Moderate*, and *Minor*).
- **Policies using insider risk levels**: Shows the status of policies (*Not started* or *Complete*), the type of policy (*Conditional Access* or *Data Loss Prevention*), and the number of configured

policies for each type of policy. If a policy type isn't configured, select **Quick setup** to configure the policy.



View users assigned insider risk levels

The **Users assigned insider risk levels** tab shows users assigned an insider risk level in Adaptive Protection. You can review the following information for each user:

- **Users**: Lists the user name. For data loss prevention policies, if the Show anonymized versions of usernames option is selected in Insider Risk Management settings, you see anonymized user names. For Conditional Access policies, user names aren't anonymized even if the **Show anonymized versions of usernames** setting is selected.

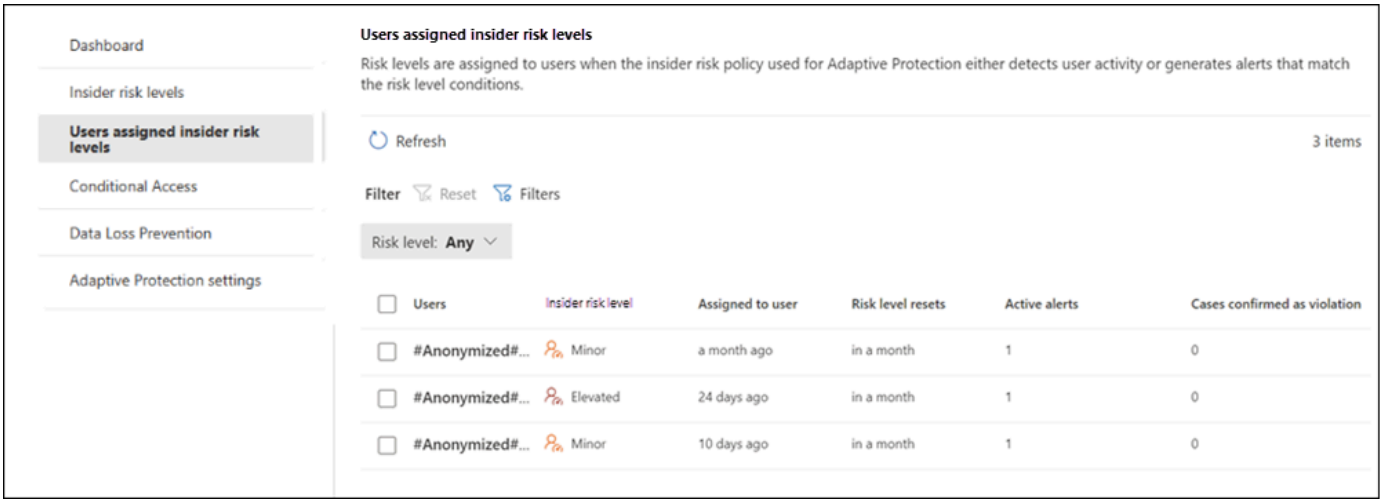
  > **Important**
  >
  > To maintain referential integrity, anonymization of usernames (if turned on) isn't preserved for users from Adaptive Protection who have alerts or activity that appear outside Insider Risk Management. Actual usernames appear in related data loss prevention alerts and the activity explorer.

- **Insider risk level**: The current insider risk level assigned to the user.
- **Assigned to user**: The number of days or months that pass since assigning an insider risk level to the user.
- **Insider risk level resets**: The number of days until the insider risk level automatically resets for the user.

To manually reset the insider risk level for a user, select the user, then select **Expire**. This user no longer has an assigned insider risk level. Existing alerts or cases for this user aren't removed. If this user is included in the selected Insider Risk Management policy, an insider risk level is assigned again if a triggering event is detected.

- **Active alerts**: The number of current Insider Risk Management alerts for the user.
- **Cases confirmed as violation**: The number of confirmed cases for the user.
- **Case**: The name of the case.

You can filter users by *Insider risk level*.

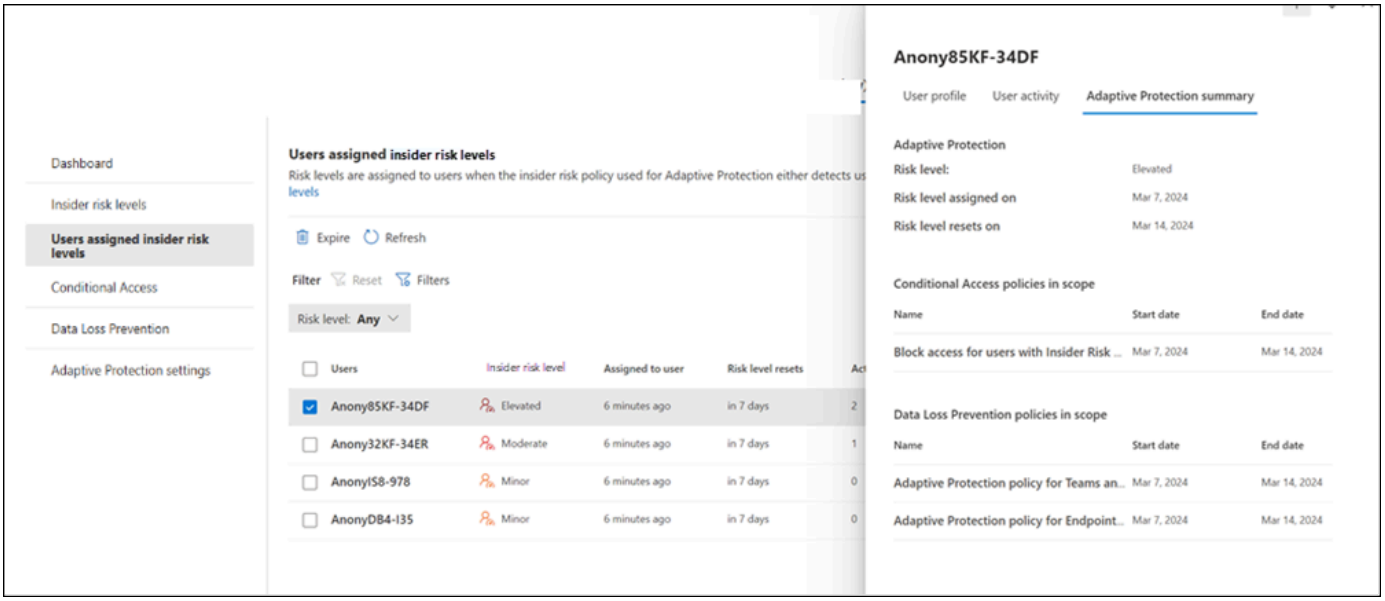


To view detailed insider risk and Adaptive Protection information for a specific user, select the user to open the user details pane. The details pane contains three tabs: **User profile**, **User activity**, and **Adaptive protection summary**. For information about the **User profile** and **User activity** tabs, see View user details.

The **Adaptive protection summary** tab aggregates information in four sections:

- **Adaptive protection**: Displays information about the current *Risk level*, *Risk level assigned on*, and *Risk level reset on* for the user.
- **Data Loss Prevention policies in scope (dynamic)**: Displays all the data loss prevention policies currently in scope for the user and the start and end date for the policy. This information is based on the insider risk level for the user and data loss prevention policy configuration for insider risk levels. For example, if a user has activities that are defined as *Elevated risk level* for Insider Risk Management policies, and two data loss prevention policies are configured with the *Elevated risk level* condition, these two data loss prevention policies are displayed here for the user.

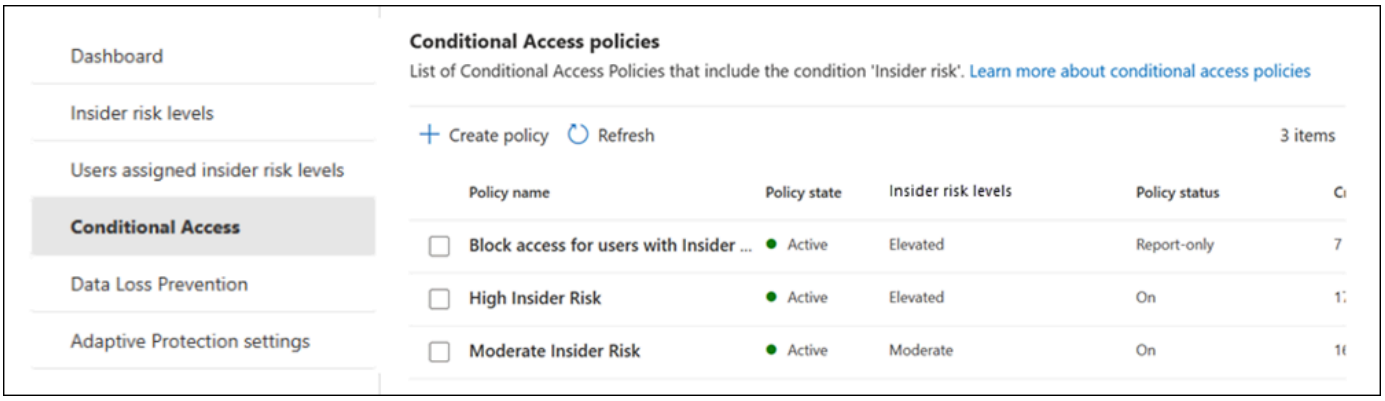
- **Conditional Access policies in scope (dynamic)**: Displays all the Conditional Access policies currently in scope for the user and the start and end date for the policy. This information is based on the insider risk level for the user and Conditional Access policy configuration for insider risk levels. For example, if a user has activities that are defined as *Elevated risk level* for Insider Risk Management policies, and a Conditional Access policy is configured with the *Elevated risk level* condition, the Conditional Access policy is displayed here for the user.
- **Insider risk policy for Adaptive protection**: Displays any Insider Risk Management policy where the user is currently in-scope.



Manage Conditional Access policies for Adaptive Protection

The **Conditional Access policies** page shows all Conditional Access policies that use the **Insider risk** condition. You can review the following information for each policy:

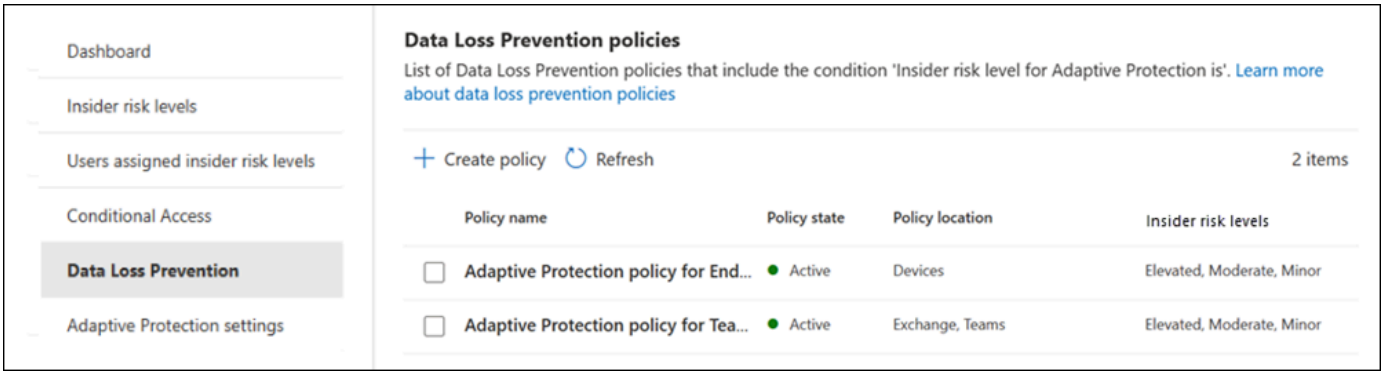
- **Policy name**: The name of the Conditional Access policy.
- **Policy state**: The current state of the policy. Values are *Active* or *Inactive*.
- **Insider risk levels**: The insider risk levels included in the Conditional Access policy by using the **Insider Risk** condition. The options are *Elevated*, *Moderate*, or *Minor*.
- **Policy status**: The current status of the Conditional Access policy. The options are *On* or *Test with notifications*.
- **Created on**: The date you created the Conditional Access policy.
- **Last modified**: The date you last edited the Conditional policy.



Manage Data Loss Prevention policies for Adaptive Protection

The **Data Loss Prevention policies** page shows all data loss prevention policies that use the **User's insider risk level for Adaptive Protection is** condition. You can review the following information for each policy:

- **Policy name**: The name of the data loss prevention policy.
- **Policy state**: The current state of the policy. Values are *Active* or *Inactive*.
- **Policy location**: The DLP policy locations included in the data loss prevention policy. Currently, Adaptive Protection supports Exchange, Teams, and Devices.
- **Insider risk levels**: The insider risk levels included in the data loss prevention policy by using the **Insider risk level for Adaptive Protection is** condition. The options are *Elevated*, *Moderate*, or *Minor*.
- **Policy status**: The current status of the data loss prevention policy. The options are *On* or *Test with notifications*.
- **Created**: The date you created the data loss prevention policy.
- **Last modified**: The date you last edited the data loss prevention policy.



Tune your insider risk level settings

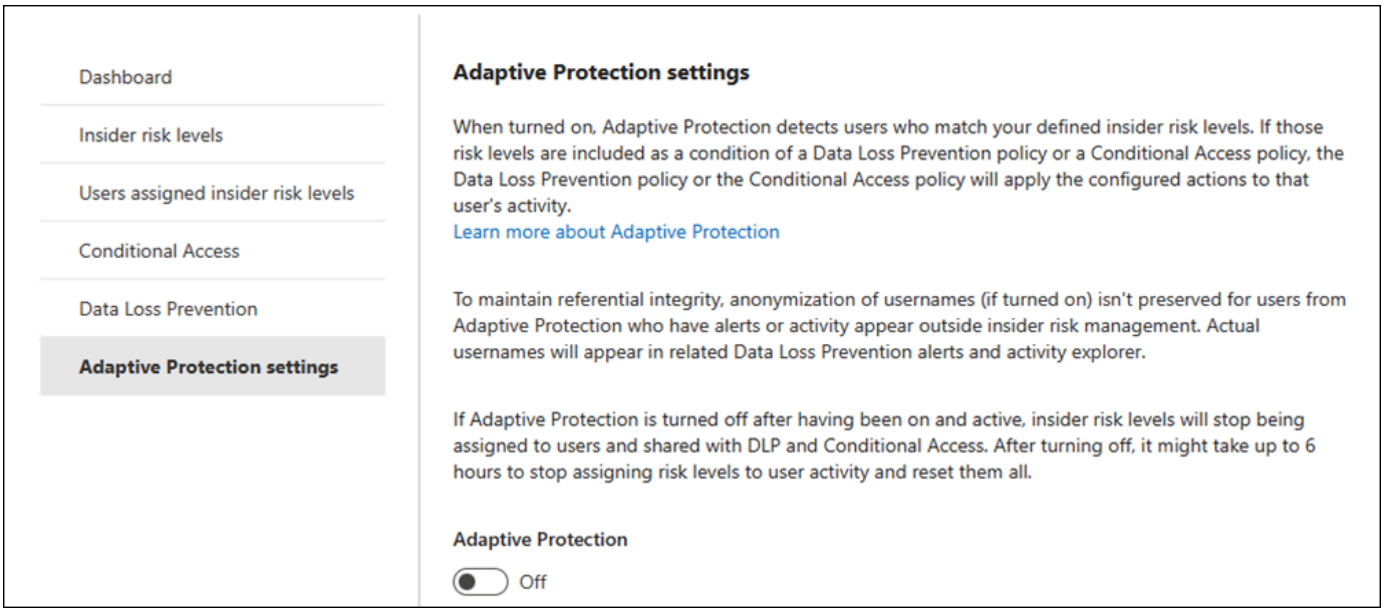
After reviewing users with insider risk levels, you might find too many or too few users assigned an insider risk level. Use these two methods to tune your policy configurations and adjust the number of users assigned insider risk levels:

- Modify insider risk level settings. Adjust your thresholds to assign an insider risk level to a user:
  - Increase or decrease the severity of the activity required to assign an insider risk level. For example, reduce the activity or alert severity if you're seeing too few users with insider risk levels.
  - If the insider risk level is based on a specific user activity, increase or decrease the activity occurrences during the detection window. For example, reduce the activity occurrences if you're seeing too few users with insider risk levels.
  - Change what the insider risk level is based on. For example, assign an insider risk level only if the alert is confirmed to decrease the number of users if you're seeing too many users with insider risk levels.
- Modify policy thresholds. Since policy detections assign insider risk levels, modify your policy to change the requirements to assign an insider risk level. Modify a policy by increasing or decreasing the policy thresholds that lead to high, medium, or low severity activities and alerts.

Disable Adaptive Protection

In certain scenarios, you might need to temporarily disable Adaptive Protection. To disable Adaptive Protection, select the **Adaptive Protection settings** tab and turn **Adaptive Protection** *Off*.

If you turn off Adaptive Protection after it's been on and active, the system stops assigning insider risk levels to users and stops sharing them with data loss prevention, data lifecycle management, and Conditional Access. All existing insider risk levels for users are reset. After turning off Adaptive Protection, it might take up to six hours to stop assigning insider risk levels to user activity and reset them all. The Insider Risk Management, data loss prevention, data lifecycle management, and Conditional Access policies aren't automatically deleted.



> **Note**
>
> You can opt out of data lifecycle management protection without disabling Insider Risk Management Adaptive Protection by turning off the **Adaptive protection in Data Lifecycle Management** setting in data lifecycle management. If you turn off this setting, the data lifecycle management policy is deleted. The setting isn't enabled again unless you turn it back on. Learn more about the Adaptive protection in Data Lifecycle Management setting

Note: The author created this article with assistance from AI. Learn more

Last updated on 06/26/2026

Learn about Insider Risk Management forensic evidence

08/20/2025

> **Important**
>
> Forensic evidence is an opt-in add-on feature in Microsoft Purview Insider Risk Management that gives security teams visual insights into potential insider data security incidents, with user privacy built in. Forensic evidence includes customizable event triggers and built-in user privacy protection controls, enabling security teams to better investigate, understand and respond to potential insider data risks like unauthorized data exfiltration of sensitive data.
>
> Organizations set the right policies for themselves, including what risky events are highest priority for capturing forensic evidence and what data is most sensitive. Forensic evidence is off by default, policy creation requires dual authorization and usernames can be masked with pseudonymization (which is on by default for Insider Risk Management). Setting up policies and reviewing security alerts within Insider Risk Management leverages strong role-based access controls (RBAC), ensuring that the designated individuals in the organization are taking the right actions with additional auditing capabilities.

> **Important**
>
> Microsoft Purview Insider Risk Management correlates various signals to identify potential malicious or inadvertent insider risks, such as IP theft, data leakage, and security violations. Insider Risk Management enables customers to create policies to manage security and compliance. Built with privacy by design, users are pseudonymized by default, and role-based access controls and audit logs are in place to help ensure user-level privacy.

Visual context is crucial for security teams during forensic investigations to get better insights into potentially risky security-related user activities. With customizable event triggers and built-in user privacy protection controls, forensic evidence enables customizable visual activity capturing across devices to help your organization better mitigate, understand, and respond to potential data risks like unauthorized data exfiltration of sensitive data. You set the right policies for your organization, including what risky events are the highest priority for capturing forensic evidence, what data is most sensitive, and whether users are notified when forensic

capturing is activated. Forensic evidence capturing is off by default and policy creation requires dual authorization.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Feature capabilities

- **Visual capturing** enables your organization to capture clips of key security-related user activities, so you can have more secure or compliant visibility that meets your organizational needs.
- **Include or exclude desktop applications and websites** to configure a recording policy that focuses on the applications and websites that present the most risk. This approach preserves storage space and user privacy. For example, you can exclude personal email and social media accounts.
- **Enhanced phishing protection (preview)** enables your organization to capture clips related to Enhanced Phishing Protection in Microsoft Defender SmartScreen. For example, you can capture when a user enters the Microsoft password they used to sign into their Windows 11 device on a phishing site or application connecting to a phishing site. Learn more about Enhanced Phishing Protection in Microsoft Defender SmartScreen
- **Protected user privacy** through multiple levels of approval for the activation of the capturing feature.
- **Customizable triggers and capturing options** mean that security teams can set up forensic evidence to meet their needs, whether it be based on incidents (for example, *Capture 5 min before and 10 min after a user has downloaded 'SecretResearchPlans.docx'*), or based on continuous capturing needs.
- **User-centric policy targeting** means that security and compliance teams can focus on activity by user, not device, for better contextual insights.
- **Strong role-based access controls (RBAC)** mean that the ability to set up and review forensic clips is tightly controlled and only available to individuals in the organization with the right permissions.
- **Deep integration with current Insider Risk Management features**, making for easier onboarding and more familiar workflows for Insider Risk Management administrators and a trusted single-platform approach.
- **Trial capacity (up to 20 GB)** for captured clips, with quick access to capacity utilization and the ability to purchase additional capacity.

Device and configuration requirements

The following tables include the supported minimum requirements for utilizing Insider Risk Management forensic evidence.

Supported platforms

Expand table

Operating system	SKU	Processor
Windows 10 (including Windows 365)	Enterprise	64-bit (Intel or AMD)
Windows 11 (including Windows 365)	Enterprise	64-bit (Intel or AMD)

Physical devices

Expand table

Hardware	Minimum requirement
RAM	Minimum of 8 GB (at least 2 GB should be available for client usage)
CPU processor	Intel i5 or higher and AMD Ryzen 5 or higher
Graphics card	Compatible with DirectX 11 or later, with a WDDM 1.0 driver or later (currently only integrated graphics cards supported)
Disk space	Minimum of 10 GB of disk storage
Display	Minimum screen resolution of 1920 x 1080

Hyper-V and virtual machines

Expand table

Hardware	Minimum requirement
RAM	Minimum of 16 GB (at least 2 GB should be available for client usage)
CPU processor	Minimum of eight vCPU processors or equivalent
Disk space	Minimum of 10 GB of disk storage

Hardware	Minimum requirement
Display	Minimum screen resolution of 1920 x 1080

> **Important**
>
> If you don't meet the minimum requirements, you might run into Microsoft Purview client issues and the quality of forensic captures might not be reliable.

Capturing options

Triggering events, global indicators, and policy indicators play an important role in all Insider Risk Management policies, including forensic evidence policies. Triggering events are user actions that determine if users are brought into scope for evaluation in Insider Risk Management policies. Global settings indicators determine what activities Insider Risk Management collects. Policy indicators determine a risk score for an in-scope user.

Depending on how your organization decides to configure forensic evidence, you can choose from two capturing options:

- **Specific activities**: This policy option captures activity only when a triggering event brings an approved user into scope for the forensic evidence policy and when the conditions for a policy indicator are detected for the user. For example, a user approved for forensic evidence capturing is brought in-scope to the forensic evidence policy and the user copies data to personal cloud storage services or portable storage devices. Capturing is scoped only to the configured time frame when the user is copying the data to the personal cloud storage service or portable storage device. You can review captures for this option on the **Forensic evidence** tab on the **Alerts** dashboard.
- **All activities**: This policy option captures any activity performed by users. For example, your organization has a time-sensitive need for capturing activities for an approved user that is actively involved in potentially risky activities that might lead to a security incident. Policy indicators might not have reached the threshold for an alert to be generated by the policy and the potentially risky activity might not be documented. Continuous capturing help prevents the potentially risky activity from being missed or going undetected. You can review captures for this option on the **Forensic evidence** tab on the **User activity reports (preview)** dashboard.

> **Important**

Forensic evidence clips are deleted 120 days after they're captured or at the end of the preview period, whichever is sooner. You can download or transfer forensic evidence clips before they're deleted.

Workflow

The overall workflow for detecting, investigating, and remediating alerts that contain clip capturing follows the same basic steps as other Insider Risk Management policies. However, some notable differences exist for forensic evidence when configured in your organization:

- **Users subject to capturing must have explicit capturing requests and approvals**: This extra process isn't included as a part of configuring other Insider Risk Management policies. Users assigned to the *Insider Risk Management* or *Insider Risk Management Admins* role groups must submit a request to users assigned to the *Insider Risk Management Approvers* role group before any user in your organization is eligible for any clip capturing options. For example, this requirement helps support organizational scenarios where your Insider Risk Management admins must get explicit approval from your designated legal or human resources personnel before capturing for any user is enabled.
- **Devices must be onboarded and have the Microsoft Purview client installed**: Before forensic evidence can collect and store clips captured for eligible users, their devices must be onboarded to the Microsoft Purview portal. Additionally, each device must have the Microsoft Purview Client installed. These prerequisites enable support for both online and offline device capturing.

Ready to get started?

- For step-by-step guidance to configure forensic evidence capturing in your organization, see Get started with Insider Risk Management forensic evidence.
- To configure prerequisites, create policies, and start receiving alerts, see Get started with Insider Risk Management.

Get started with Insider Risk Management forensic evidence

08/20/2025

> **Important**
>
> Forensic evidence is an opt-in add-on feature in Microsoft Purview Insider Risk Management that gives security teams visual insights into potential insider data security incidents, with user privacy built in. Forensic evidence includes customizable event triggers and built-in user privacy protection controls, enabling security teams to better investigate, understand and respond to potential insider data risks like unauthorized data exfiltration of sensitive data.
>
> Organizations set the right policies for themselves, including what risky events are highest priority for capturing forensic evidence and what data is most sensitive. Forensic evidence is off by default, policy creation requires dual authorization and usernames can be masked with pseudonymization (which is on by default for Insider Risk Management). Setting up policies and reviewing security alerts within Insider Risk Management leverages strong role-based access controls (RBAC), ensuring that the designated individuals in the organization are taking the right actions with additional auditing capabilities.

> **Important**
>
> Microsoft Purview Insider Risk Management correlates various signals to identify potential malicious or inadvertent insider risks, such as IP theft, data leakage, and security violations. Insider Risk Management enables customers to create policies to manage security and compliance. Built with privacy by design, users are pseudonymized by default, and role-based access controls and audit logs are in place to help ensure user-level privacy.

Configuring forensic evidence in your organization is similar to configuring other policies from Insider Risk Management policy templates. In general, you follow the same basic configuration steps to set up forensic evidence, but a few areas require feature-specific configuration actions before you get started with the basic configuration steps.

> **Tip**

Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Step 1: Confirm your subscription and configure data storage access

Before you get started with forensic evidence, confirm your Insider Risk Management subscription and any add-ons.

Also, add the following domains to your firewall and proxy server allowlist to support forensic evidence capture storage for your organization:

Worldwide - Domain

```
- compliancedrive.microsoft.com
- *.events.data.microsoft.com
- officeclient.microsoft.com
- odc.officeapps.live.com
- hrd.svc.cloud.microsoft
```

GCC/GCC High - Domain

```
- *.compliancedrive.microsoft.us
- tb.events.data.microsoft.com
- officeclient.microsoft.com
- odc.officeapps.live.com
- hrd.svc.cloud.microsoft
```

DOD - Domain

```
- dod.compliancedrive.apps.mil
- pf.events.data.microsoft.com
- officeclient.microsoft.com
- odc.officeapps.live.com
- hrd.svc.cloud.microsoft
```

For more information on these endpoints, see Microsoft 365 endpoints.

> **Note**
>
> Captures and capture data are stored at these domains and are assigned only to your organization. No other Microsoft 365 organization has access to forensic evidence captures for your organization.
>
> Forensic evidence data is stored in one region where your Exchange Online region is set.

Step 2: Configure supported devices

To capture forensic evidence, eligible user devices must be onboarded to the Microsoft Purview portal and have the Microsoft Purview Client installed.

> **Important**
>
> The Microsoft Purview Client automatically collects general diagnostic data related to device configuration and performance metrics. This data includes information on critical errors, RAM consumption, process failures, and other data. This data helps us assess the client's health and identify any issues. For more details about how diagnostic data might be used, see the Use of Software with Online Services on the **Microsoft Product Terms**.

For a list of device and configuration requirements, see Learn about forensic evidence. To onboard supported devices, complete the steps outlined in the Onboard Windows 10 and Windows 11 devices into Microsoft 365 overview article.

Install the Microsoft Purview client

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.
2. Go to the **Insider Risk Management** solution.
3. Select **Forensic Evidence** in the left navigation, then select **Client installation**.
4. Select **Download installer package (x64 version)** to download the installation package for Windows.
5. After downloading the installation package, use your preferred method to install the client on users' devices. These options might include manually installing the client on

devices or using tools to help automate the client installation:

- **Microsoft Intune**: Microsoft Intune is an integrated solution for managing all of your devices. Microsoft brings together Configuration Manager and Intune, without a complex migration, and with simplified licensing.
- **Third-party device management solutions**: If your organization uses third-party device management solutions, see the documentation for these tools to install the client.

Step 3: Configure settings

Forensic evidence has several configuration settings that provide flexibility for the types of security-related user activity captured, capturing parameters, bandwidth limits, and offline capturing options. By configuring forensic evidence capturing, you can create policies based on your requirements. Adding users to a policy requires dual authorization.

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.
2. Go to the **Insider Risk Management** solution.
3. Select **Forensic evidence** in the left navigation, then select **Forensic evidence settings**.
4. Select **Forensic evidence capturing** to enable capturing support in your forensic evidence policies. If you turn this setting off, it removes all previously added users for forensic evidence policies.

   > **Important**
   >
   > The Microsoft Purview Client used to capture activity on users' devices is licensed under the Use of Software with the Online Services on the Microsoft Product Terms. Customers are solely responsible for using the Insider Risk Management solution, including the Microsoft Purview Client, in compliance with all applicable laws.

5. In the **Capturing window** section, define when to start and stop activity capturing. Available values are *10 seconds*, *30 seconds*, *1 minute*, *3 minutes*, or *5 minutes*.
6. In the **Upload bandwidth limit** section, define the amount of capture data to upload into your data storage account per user, per day. Available values are *100 MB*, *250 MB*, *500 MB*, *1 GB*, or *2 GB*.

7. In the **Offline capturing cache limit** section, define the maximum cache size to store on users' devices when offline capturing is enabled. Available values are *100 MB*, *250 MB*, *500 MB*, *1 GB*, or *2 GB*.

8. Select **Save**.

Step 4: Create a policy

Forensic evidence policies define the scope of security-related user activity to capture for configured devices. You have two options for capturing forensic evidence:

- **Capture only specific activities (such as printing or exfiltrating files).** With this option, you choose the device activities to capture. You can also choose to capture activity for specific desktop apps and websites. This option lets you focus on just the activities, apps, and websites that present risk.
- **Capture all activities that approved users perform on their devices.** This option typically covers a specific period of time, such as when a particular user is potentially involved in risky activity that might lead to a security incident. All forensic evidence indicators are automatically included, such as device indicators and Enhanced Phishing Protection indicators. To preserve capacity and user privacy, you can choose to exclude specific desktop apps and websites from the capture.

After you create a policy, include it in forensic evidence requests to control what activity to capture for users whose requests are approved.

> **Note**
>
> Continuous forensic policies (capturing all activities) take precedence over selective forensic evidence policies (capturing only specific activities).

Capture only specific activities

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.

2. Go to the **Insider Risk Management** solution.

3. Select **Forensic evidence** in the left navigation, then select **Forensic evidence policies**.

4. Select **Create forensic evidence policy**.

5. On the **Scope** page, select **Specific activities**. This option only captures activities detected by policies that users are included in. These activities are defined by the indicators selected in forensic evidence policies. You can review captures for this option on the **Forensic evidence (preview)** tab on the **Alerts** or **Cases** dashboard.

6. Select **Next**.

7. On the **Name and description** page, complete the following fields:

   - **Name (required)**: Enter a friendly name for the forensic evidence policy. You can't change this name after the policy is created.
   - **Description (optional)**: Enter a description for the forensic evidence policy.

8. Select **Next**.

9. On the **Choose device activities to capture** page:
   a. Select any device activities that you want to capture. The policy only captures the selected activities.

   > ⓘ **Note**
   >
   > If the indicators aren't selectable, you're prompted to turn them on.

10. To capture activity for particular desktop apps and websites in your policy, select the **Opening a specific app or website** check box under **App and web browsing activities to capture**.

    > ⓘ **Important**
    >
    > To capture browsing activities (to include or exclude specific URLs in your forensic evidence policies), make sure to install the necessary browser extensions. You also need to turn on at least one browsing indicator. If you haven't already turned on one or more browsing indicators, you're prompted to do so if you choose to include or exclude desktop apps or websites. The triggering event for capturing browsing activities is a URL update in the URL bar that contains the specified URL.

    Select **Next**.

11. (Optional) If you chose to capture activity for particular desktop apps and websites, in the **Add apps and websites you want to capture activity for** page:
    a. To add a desktop app, select **Add desktop apps**, enter the name of an executable file (for example, teams.exe), then select **Add**. Repeat this process for each desktop app

that you want to add (up to 25 apps). To find the name of an executable file for the app, open the Task Manager, then view the properties for the app. Here's a list of exe names for some of the common applications: Microsoft Edge (msedge.exe), Microsoft Excel (Excel.exe), the Snipping tool (SnippingTool.exe), Microsoft Teams (Teams.exe), Microsoft Word (WinWord.exe), and Microsoft Remote Desktop Connection (mstsc.exe).

> **Note**
>
> Sometimes, the exe names for an app might differ based on the device and the permissions with which the app was opened. For example, on a Windows 11 enterprise device, when Windows PowerShell is opened without administrator permissions, the exe name is WindowsTerminal.exe but when opened with administrator permissions, the exe name changes to powershell.exe. Make sure to include/exclude both exe names in such scenarios.

a. To add a web app or website, select **Add web apps and websites**, enter a URL (for example, https://teams.microsoft.com), then select **Add**. Repeat this process for each web app or website that you want to add. You can add up to 25 URLs with a character length of 100 for each URL.

> **Tip**
>
> If an app has a desktop and web version, add both the desktop executable and the web URL to make sure you capture activity for both.

a. Select **Next**.

12. On the **Review settings and finish** page, review the settings you chose for the policy and any suggestions or warnings for your selections. Edit any of the policy values or select **Submit** to create and activate the policy.

13. After you complete the policy configuration steps, continue to Step 5.

Capture all activities

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.

2. Go to the **Insider Risk Management** solution.

3. Select **Forensic evidence** in the left navigation, then select **Forensic evidence policies**.

4. Select **Create forensic evidence policy**.
5. On the **Scope** page, select **All activities**. This option captures any activity performed by users. You can review captures for this option on the **Forensic evidence (preview)** tab on the **User activity reports (preview)** dashboard.
6. Select **Next**.
7. On the **Name and description** page, complete the following fields:
    - **Name (required)**: Enter a friendly name for the forensic evidence policy. You can't change this name after the policy is created.
    - **Description (optional)**: Enter a description for the forensic evidence policy.
8. Select **Next**.
9. On the **Choose device activities to capture** page, if you want to exclude certain desktop apps and/or web apps or websites from the capture, under **App and web browsing activities to capture**, select the **Exclude specific apps or websites** check box.
10. Select **Next**.
11. If you chose to exclude particular desktop apps and websites from the capture, in the **Exclude applications/URLs** page:
    - To exclude a desktop app from the capture, select **Exclude desktop apps**, enter the name of an executable file (for example, teams.exe), and then select **Add**. Repeat this process for each desktop app that you want to exclude (up to 25 apps). To find the name of an executable file for an app, open the Task Manager, and then view the properties for the app.
    - To exclude a web app or website, select **Exclude web apps and websites**, enter a URL (for example, https://teams.microsoft.com), and then select **Add**. Repeat this process for each web app or website that you want to exclude. You can exclude up to 25 URLs with a character length of 100 for each URL.

    > **Tip**
    >
    > If an app has a desktop and web version, add both the desktop executable and the web URL to make sure you exclude both.

12. On the **Review settings and finish** page, review the settings you chose for the policy and any suggestions or warnings for your selections. Edit any of the policy values or select **Submit** to create and activate the policy.

13. After you complete the policy configuration steps, continue to Step 5.

Step 5: Define and approve users for capturing

Before you can capture security-related user activities, admins must follow the dual authorization process in forensic evidence. This process requires that applicable people in your organization both define and approve the enabling of visual capturing for specific users.

You must request to enable forensic evidence capturing for specific users. When you submit a request, approvers in your organization receive an email notification and can approve or reject the request. If approved, the user appears on the **Approved users** tab and is eligible for capturing. For more information, see these links:

- Request approval for forensic evidence capturing.
- Approve (or reject) requests for forensic evidence capturing.

Next steps

After you configure your forensic evidence policy, it can take up to two hours for policy enforcement, given that the forensic evidence clients (installed at the endpoint devices) are online. When a client captures a clip, it can take up to one hour before the clip is available to review. For more information about managing forensic evidence and reviewing clip captures, see the Manage information risk management forensic evidence article.

Manage Insider Risk Management forensic evidence

> **Important**
>
> Forensic evidence is an opt-in add-on feature in Microsoft Purview Insider Risk Management that gives security teams visual insights into potential insider data security incidents, with user privacy built in. Forensic evidence includes customizable event triggers and built-in user privacy protection controls, enabling security teams to better investigate, understand and respond to potential insider data risks like unauthorized data exfiltration of sensitive data.
>
> Organizations set the right policies for themselves, including what risky events are highest priority for capturing forensic evidence and what data is most sensitive. Forensic evidence is off by default, policy creation requires dual authorization and usernames can be masked with pseudonymization (which is on by default for Insider Risk Management). Setting up policies and reviewing security alerts within Insider Risk Management leverages strong role-based access controls (RBAC), ensuring that the designated individuals in the organization are taking the right actions with additional auditing capabilities.

> **Important**
>
> Microsoft Purview Insider Risk Management correlates various signals to identify potential malicious or inadvertent insider risks, such as IP theft, data leakage, and security violations. Insider Risk Management enables customers to create policies to manage security and compliance. Built with privacy by design, users are pseudonymized by default, and role-based access controls and audit logs are in place to help ensure user-level privacy.

When you complete the configuration steps and create your forensic evidence policy, you start seeing alerts for potentially risky security-related user activities that meet the conditions for indicators defined in the policy.

> **Tip**

Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Use the forensic evidence dashboard

The forensic evidence dashboard provides a summary view of key areas of the forensic evidence configuration in your organization. For this preview release, the dashboard includes only a **Forensic evidence device health** section. Select **View device health report** to open the **Device health** tab and report.

Manage forensic evidence users and approvals

You must request and approve specific users before they're eligible for forensic evidence capturing. Simply adding users to a forensic evidence policy doesn't automatically make those users eligible for capturing. You can request and approve users before or after creating forensic evidence policies, but the clip captures associated with policy indicators are created and available for reviewing only after approving the users.

Users assigned to the *Insider Risk Management* or *Insider Risk Management Admins* role groups can submit approval requests to users assigned to the *Insider Risk Management Approvers* role group.

Request capturing approvals

Request turning on forensic evidence to capture information for specific users. When you submit a request, approvers in your organization receive an email notification and can approve or reject the request. If approved, the user appears on the **Approved users** tab and is eligible for capturing. If not addressed, the request expires six months from the day it was submitted.

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.
2. Go to the **Insider Risk Management** solution.
3. Select **Forensic evidence** in the left navigation, then select **User management**.
4. Select the **Manage forensic evidence requests** tab.
5. Select **Create request**.

6. On the **Users** page, select **Add users**.
7. Use **Search** to locate a specific user or select one or more users from the list. Select **Add**, then select **Next**.
8. On the **Forensic evidence policy** page, select a forensic evidence policy for the added users. The policy you choose determines the scope of activity to capture for users.
9. Select **Next**.
10. On the **Justification** page, explain why you're requesting that capturing be enabled for the users you added in the **Justification for turning on forensic evidence capturing** text box. This is a required field. When complete, select **Next**.
11. On the **Email notifications** page, use a notification template to send an email to users letting them know that forensic evidence capturing is turned on for their device in accordance with your organization's policies. The email is sent to users only if their request is approved.

    Select the **Send an email notification to approved users** check box. Choose an existing template or create a new one. To create a new template, select **Create a notification template** and complete the required fields in the **New email notification template** pane.
12. Select **Next**.
13. On the **Finish** page, review your settings before submitting the request. Select **Edit users** or **Edit justification** to change any of the request values or select **Submit** to create and send the request to reviewers.

To view pending approval requests, go to **Insider Risk Management** > **Forensic evidence** > **Pending requests**. Here you see the users with pending requests, their email address, the request submission date, and who submitted the approval request. If no users are displayed, there aren't any pending approval requests for any users.

Users assigned to the *Insider Risk Management Approvers* role group can select a user on the **Forensic evidence request** tab and review the request. After reviewing the request, these users can approve or reject the forensic evidence capturing request. Approving or rejecting the capturing request removes the pending request for users from this view.

Approve or reject capturing requests

After a capturing approval request is submitted, users assigned to the *Insider Risk Management Approvers* role group receive an email notification for the approval request.

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.
2. Go to the **Insider Risk Management** solution.
3. Select **Forensic evidence** in the left navigation, then select **Pending requests**.
4. Select a user to review.
5. On the **Review forensic evidence request (preview)** pane, review the justification submitted by the requestor. Select **Approve** or **Reject** as applicable.
6. On the **Request approved** or **Request rejected** page, select **Close**.

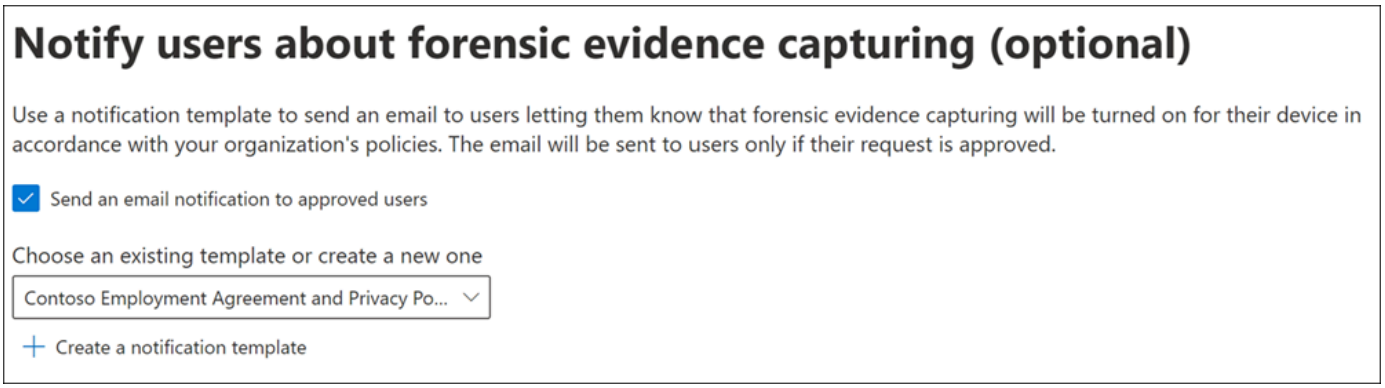
Revoke capturing approvals

If needed, you can revoke approval for specific users and exclude them from forensic evidence capturing. Revoking approval doesn't delete or remove any existing captures for these users. It only disables future capturing of activity for these users.

1. Sign in to the Microsoft Purview portal as a member of the *Insider Risk Management Approvers* role group.
2. Go to the **Insider Risk Management** solution.
3. Select **Forensic evidence** in the left navigation, then select **User management**.
4. Select the **Approved users** tab.
5. Select a user, then select **Remove**.
6. On the removal confirmation page, select **Remove** to revoke capturing approval.

Create and manage forensic evidence notification templates

You can create and use a notification template to send an email to users, letting them know that forensic evidence capturing is turned on for their device in accordance with your organization's policies. The email is sent to users only if their request is approved.



Create a new notification template

To create a notification template, complete the following steps:

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.
2. Go to the **Insider Risk Management** solution.
3. Select **Forensic evidence** in the left navigation, then select **Notification templates**.
4. Select **Create notification template**.
5. On the **New email notification template** pane, enter the following required fields:
    - Template name
    - Send from
    - Subject
    - Message body
6. Select **Save**.

> **Note**
>
> To delete an existing notification template, select the template, then select **Delete**.

View and manage captured clips

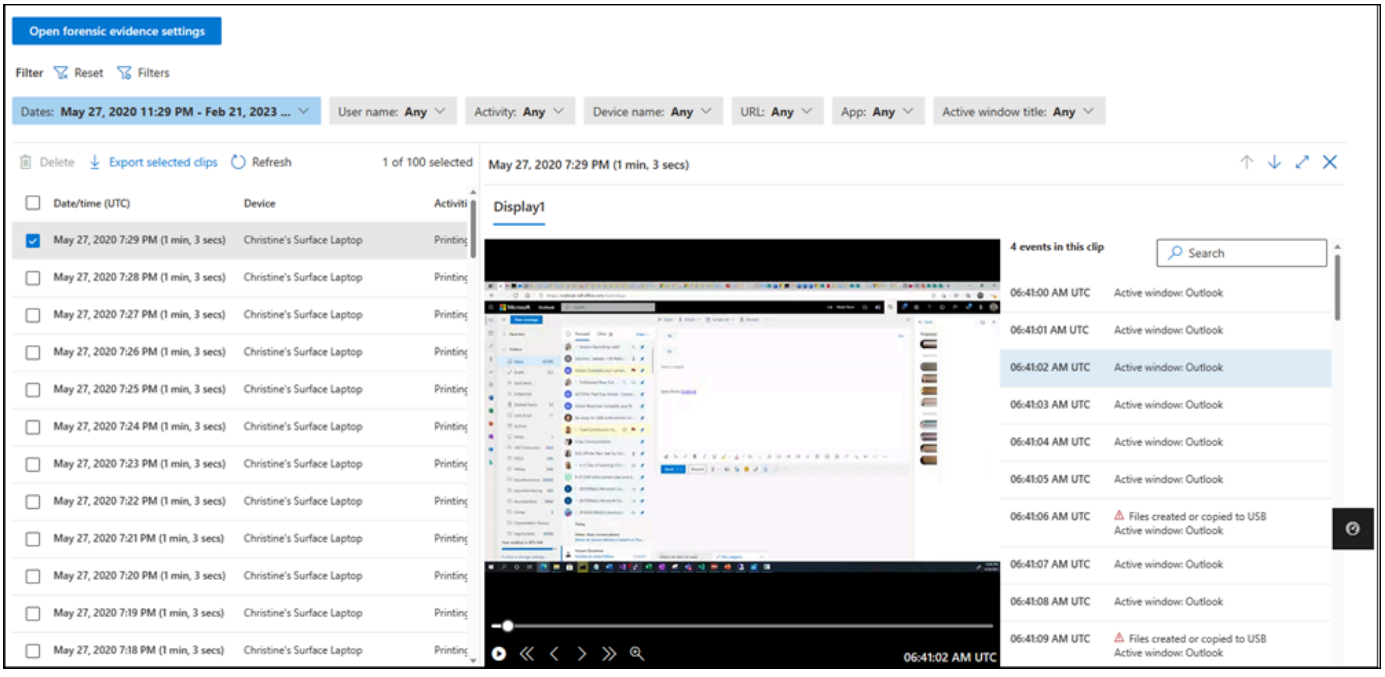
You can view and explore captured clips by selecting the **Forensics evidence** tab when you open Microsoft Purview Insider Risk Management. You can also select the **Forensics evidence** tab from other areas in the solution to view a list of captured clips in context:

- **Alerts dashboard**. Clips accessible from the **Alerts** dashboard correspond to the option to capture **specific user activities** when the forensic evidence policy is created. The captured clips are defined by indicators you select in the forensic evidence policy.

- **User activity reports**. Clips accessible from **User activity** reports correspond to the option to capture **any security-related activity** performed by users included in forensic evidence policies.
- **Cases dashboard**. Clips accessible from the **Cases** dashboard are alerts that you escalated to cases.

> **Note**
>
> If you're a member of both the *Insider Risk Management Investigators* role group and the *Insider Risk Management Admins* role group, you see **Review captured clips** when you open Microsoft Purview Insider Risk Management. If you select **Review captured clips**, it changes to **Open forensic evidence settings**. Selecting **Review captured clips** or **Open forensic evidence settings** lets you switch between the list of captured clips and settings if you have both roles. Learn more about role groups.

When you select the **Forensics evidence** tab, captured clips and associated information display in a list. If you select a captured clip in the list, a video player appears in the center of the screen, and a transcript of activities and events from the clip displays to the right of the video player.



Each captured clip includes the following information:

- **Date/time (UTC)**: The date, time (UTC), and duration of the capture. The duration of the capture is the total time spanned by the capture. The actual length of the capture might be shorter since Insider Risk Management automatically eliminates identical frames.
- **Device**: The name of the device in Windows 10/11.

- **Activities**: The Insider Risk Management activity type included in the capture. These activities are based on global and policy indicators assigned to the associated policy.
- **User**: The name of the user.
- **URL** (if applicable): The URL that the user was accessing when the activity took place.
- **Application** (if applicable): The application that the user was accessing when the activity took place.
- **Active window title**: The title of the window that the user was accessing when the activity took place.

To view a captured clip:

1. If needed, configure the filters at the top of the list.
2. Select a clip from the list.
3. Using the video player controls, select the *Play control* to review the entire clip from beginning to end.
4. To scope the review to a specific activity or event in the clip, select the activity or event in the transcript. You can also use the search box to search for specific activities or events.

   > **Note**
   >
   > A red triangle in the transcript denotes an activity.

Filtering the captured clips list

You can use the filters for the captured clips list to filter for specific activities and information. Each filter supports up to 10 unique IDs so, for example, you can filter on up to 10 users at one time. The following table describes the different filters.

Expand table

Filter name	Description
Active window title	Filter for the active window title. You can select **Contains any of** or **Contains all of** to filter in the same way as the **App** filter.
Activity	Select a specific activity to filter on, such as *Printing files* or *Password used to sign into the device has been reused*.

Filter name	Description
App	Filter by app name. You can select **Contains any of** or **Contains all of** with this filter. Example: if you select **Contains any of** and enter "Contoso.com,Contoso2.com", you could have one clip that captures Contoso.com and another that captures Contoso2.com. If you select **Contains all of** and enter "Contoso.com,Contoso2.com", any captures would have to contain both domains.
Device name	Filter on any device name.
URL	Filter for a domain name or search for any keyword after filtering for a domain name. Example: entering "SharePoint" as a keyword returns any URL that includes "SharePoint" anywhere in the URL.
User name	Filter on any user name.

Deleting clips

Users assigned to the *Insider Risk Management Investigators* role group can delete individual clips from the captured clips list. To do this:

1. Select the check box next to the capture.
2. Select **Delete**.

Users assigned to the *Insider Risk Management Admins* role group can do bulk deletions through settings.

Make a bulk deletion

To delete clips in bulk for a specific user, complete the following steps:

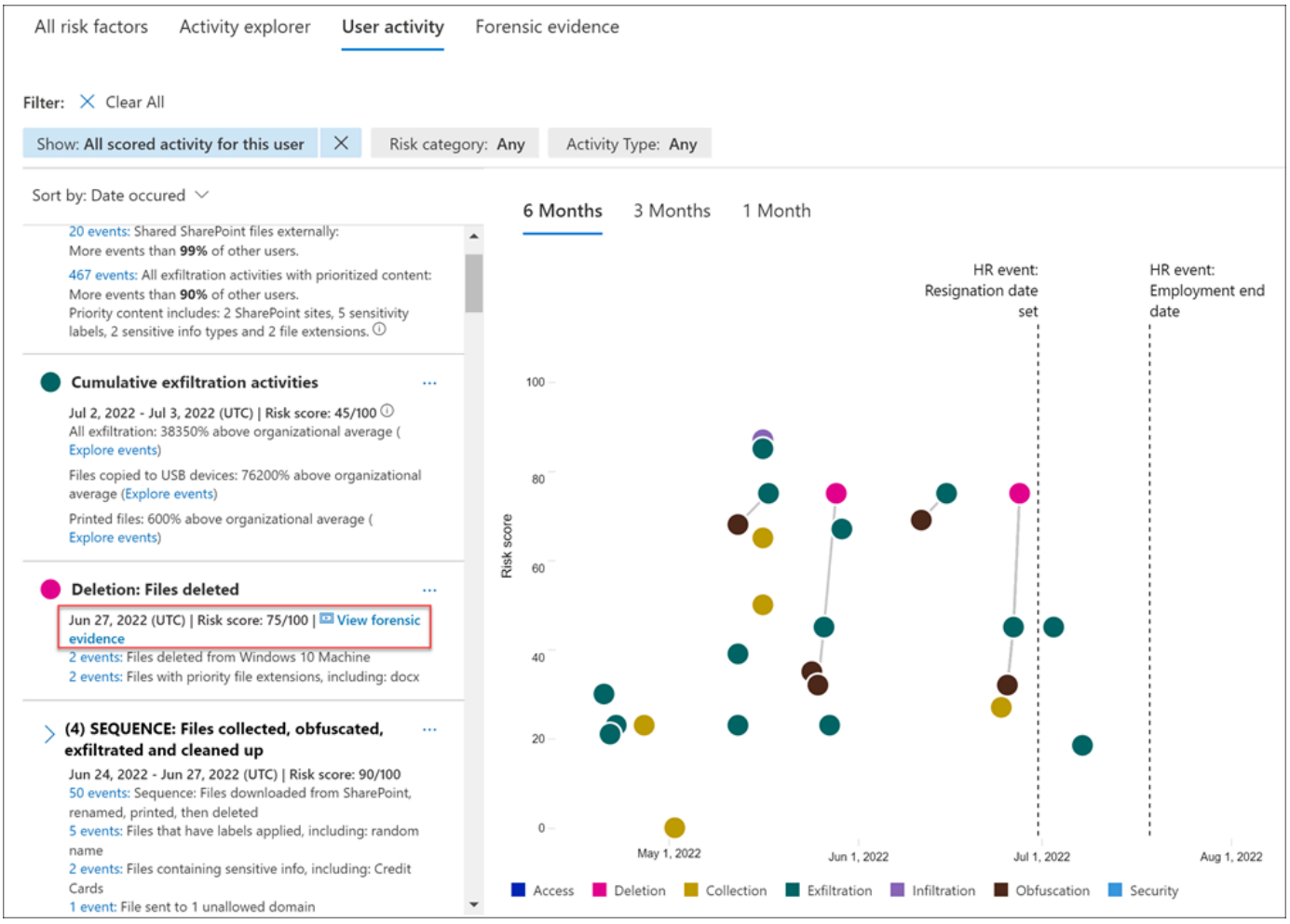
1. Sign in to the Microsoft Purview portal as a member of the *Insider Risk Management Admins* role group.
2. Go to the **Insider Risk Management** solution.
3. Select **Forensic evidence** in the left navigation, then select **Forensic evidence settings**.
4. Make sure that the **Allow deletion of forensic user data by an Administrator or Investigator** option is set to **On**.
5. Under **Delete a user's** data, choose **Select a user**, then select the user that you want to delete clips for.

> **Important**

Forensic evidence clips are deleted 120 days after they're captured. You can export or transfer forensic evidence clips before they're deleted.

View captured clips in the Alerts dashboard

For alerts generated by policies, you can review forensic evidence captures on the **Forensic evidence** tab on the **Alerts** dashboard. If one or more captures are available for the alert, you also see a **View forensic evidence** notification link in the activity that generates an alert header section. You can select the notification link or the **Forensic evidence** tab to review a list of activity captures.



Reviewing an alert for potentially risky activity that might contain forensic evidence captures is essentially the same as reviewing an alert without forensic evidence captures. The significant difference is the inclusion of any applicable captures. The **Forensic evidence** tab provides access to all available captures associated with the alert.

View captured clips in the Cases dashboard

If you escalate alerts to cases, include all associated forensic evidence captures as part of the case. Reviewing forensic evidence captures for cases follows the same process as reviewing captures for alerts.

View captured clips in user activity reports

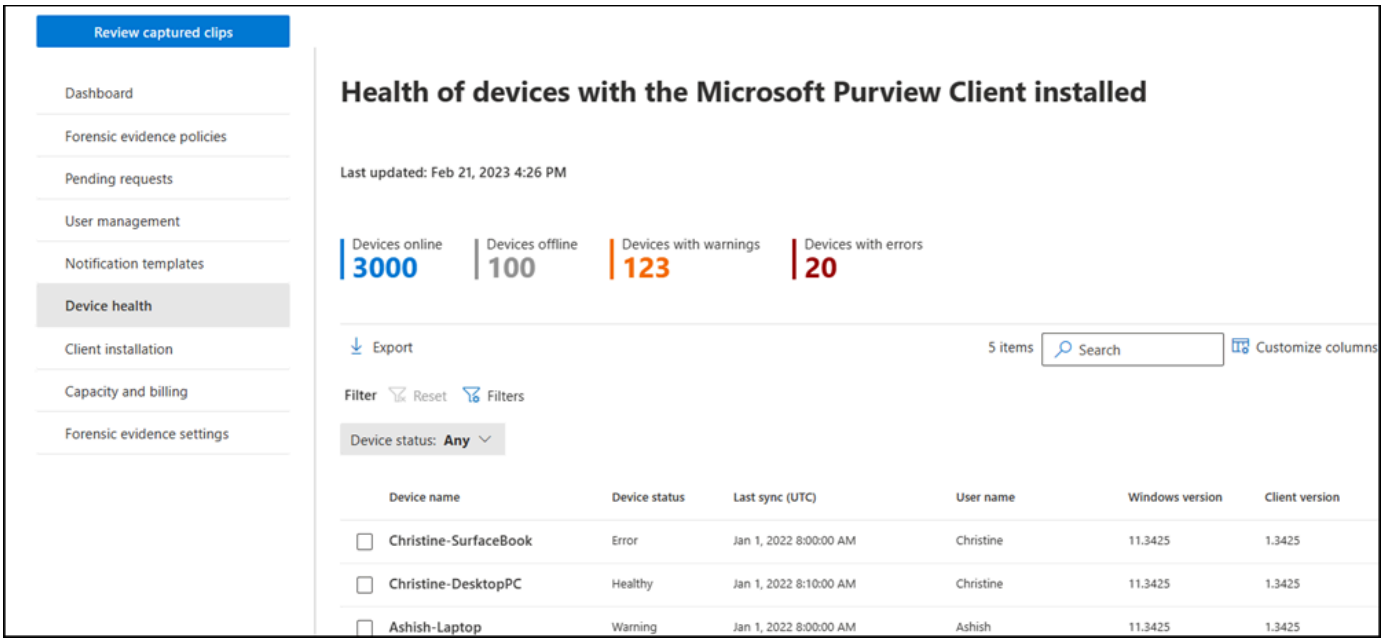
User activity reports allow you to examine activities for specific users for a defined time period without having to assign them temporarily or explicitly to an Insider Risk Management policy. If these user activities include activities supported by forensic evidence capturing, the report includes clips with the user activity.

If you configure forensic evidence to capture all security-related user activity, regardless of whether they're included in a forensic evidence policy, review these captures by following these steps:

1. Select **Insider Risk Management** > **Overview**.
2. At the bottom of the **Overview** screen, under **Investigate user activity**, select **Manage reports**.
3. Select a specific user, then select the **Forensic evidence** tab.
4. Select a clip from the list, then use the video player controls to review the clip. To scope the review to a specific activity or event, select the activity or event in the transcript.

Device health report (preview)

After you configure devices to support forensic evidence, you can review the Microsoft Purview Client health status for all devices in your organization by going to **Insider Risk Management** > **Forensic evidence** > **Device health**.



For a list of minimum device and configuration requirements, see Learn about forensic evidence. To onboard supported devices, follow the steps in the Onboard Windows 10 and Windows 11 devices into Microsoft 365 overview article.

The Device health report lets you view the status and health of all devices that have the forensic evidence agent installed. Each report widget on the report displays information for last 24 hours.

- **Devices online**: The total number of devices currently online.
- **Devices offline**: The total number of devices currently offline.
- **Devices with warnings**: The total number of devices with a warning.
- **Devices with errors**: The total number of devices with an error.

The device health queue lists all the devices configured for forensic evidence in your organization. In addition, the report lists the status of the following device attributes:

- **Device name**: The name of the device, defined by the *ComputerName* attribute of the device.
- **Device status**: The status of the Microsoft Purview Client on the device. Status values are as follows:
  - ***Healthy***: The client on the device is working properly and forensic evidence capture features are fully supported.
  - ***Warning***: The client on the device has a warning and forensic evidence capture features might not be fully supported.
  - ***Error***: The client on the device has an error and forensic evidence capture features are disabled or not fully supported.
- **Status details**: More information about the device status.
- **Last sync (UTC)**: Date and time of the last status sync for the device.

- **User name**: The user name for the user logged into the device when the status sync was performed.
- **Windows version**: The version of Microsoft Windows installed on the device.
- **Client version**: The version of the Microsoft Purview Client installed on the device.

The device health status gives you insights into potential issues with your devices and the Microsoft Purview Client. The **Device status** column on the **Device health** page can alert you to device issues that might prevent user activity from being captured or why the volume of forensic evidence capturing is unusual. The device health status can also confirm that the devices included in forensic evidence capturing are healthy and don't need attention or configuration changes. The following table lists potential status detail messages and recommended actions you can take to address warnings and errors:

Expand table

Status details	Status	Suggested action
An internal server error occurred. As a result, capture data might be missing.	Error	Create a support ticket with Microsoft for further investigation
CPU usage on the device has exceeded the maximum threshold.	Error	The capture process has been stopped and will restart in a few minutes.
Device is unable to access the directory that stores forensic evidence captures.	Error	Reinstall the client on this device
Encoder initialization failed.	Error	Reinstall the client on this device.
GPU usage on the device has exceeded the maximum threshold.	Error	The capture process has been stopped and will restart in a few minutes.
Memory usage on the device has exceeded the maximum threshold.	Error	The capture process has been stopped and will restart in a few minutes.
Offline storage has reached 90% of the configured limit on this device. Captures might be overwritten soon.	Warning	Increase the offline capturing cache limit on the Forensic evidence settings page.
The configured offline storage limit has been reached on this device. As a result, offline captures are being overwritten.	Error	Increase the offline capturing cache limit on the Forensic evidence settings page.
The configured upload bandwidth limit has been reached on this device. No more captures are uploaded	Error	Increase the upload bandwidth limit on the Forensic evidence settings

Status details	Status	Suggested action
for the day.		page.
The Microsoft Purview Client installed on the device hasn't synced with the forensic evidence policy in over 24 hours.	Error	Connect to network & reinstall client
The Microsoft Purview Client installed on the device in unable to sync with the forensic evidence policy.	Warning	Connect to network & reinstall client
The Microsoft Purview Client is unable to capture activity because display monitors on this device were turned off or disconnected.	Error	Connect/Turn on display monitors for the device
The Microsoft Purview Client is unable to capture activity because no display monitors are detected on this device.	Error	Add display monitors for this device
The Microsoft Purview Client is unable to capture activity because no graphics card is detected on this device.	Error	Add a graphics card or replace the device with one that has a graphics card
Upload bandwidth has reached 90% of the configured limit on this device. Captures might be overwritten soon.	Warning	Increase the upload bandwidth limit on the Forensic evidence settings page.

Contact Microsoft Support if the recommended actions don't resolve issues with the client.

Capacity and billing

When you configure forensic evidence, you can choose to purchase the forensic evidence add-on for Insider Risk Management for your captured clips. Organizations with any of the following licenses can get the add-on: Microsoft 365 E5, Microsoft Purview Suite (formerly known as Microsoft 365 E5 Compliance), or Microsoft 365 E5 Insider Risk Management.

Organizations can purchase the add-on in units of 100 GB per month. Purchased capacity applies to the ingestion of forensic evidence starting on the date of purchase and resets on the first of the month. Unused capacity doesn't carry over. We recommend purchasing the license at the beginning of the month to maximize the value of the license. One hundred gigabytes roughly equals around 1,100 hours of forensic evidence captures per tenant, at a video resolution of 1080p. You can download the Forensic evidence capacity calculator to help estimate the number of GBs needed per month.

Once you ingest the forensic evidence, the system retains it for 120 days. You can export forensic evidence if needed after the 120-day retention period.

Compare forensic evidence payment plans

When you purchase the add-on through the Microsoft 365 admin center, you have two payment plans available:

- **Pay yearly (available in all channels).** With the annual commitment option, you buy the number of licenses you specify each month for 12 months. This option is suitable for customers who want to ensure they have capacity available each month to ingest forensic evidence without interruption. This payment plan automatically replenishes the number of licenses purchased each month. Purchased capacity applies to the ingestion of forensic evidence starting on the date of purchase and resets on the first of the month. Unused capacity doesn't carry over. Customers can choose to be billed one time or split the bill into 12 monthly payments.
- **Pay monthly (only available in web direct).** If you don't want to make an annual commitment, you can buy the number of licenses needed each month. Purchased capacity applies to the ingestion of forensic evidence starting on the date of purchase and resets on the first of the month. Unused capacity doesn't carry over.

Can I try the forensic capability before purchasing it?

Each tenant with a Microsoft 365 E5, Microsoft Purview Suite (formerly known as Microsoft 365 E5 Compliance), or Microsoft 365 E5 Insider Risk Management license can sign up for a 20-GB trial license to try the forensic evidence capability.

> ⓘ **Note**
>
> The 20-GB trial license is only available for customers on the legacy commerce platform.

The trial license provides 20 GB of capacity until you use it up or until one year from the activation date, whichever comes first. If you purchase a forensic evidence add-on license before using all the trial capacity, you can use the remaining trial capacity before the system starts metering the purchased capacity.

If you use up the 20 GB of trial capacity without purchasing the forensic add-on for Insider Risk Management, you can view any ingested clips but can't ingest new clips.

Sign up for the 20-GB trial license

To claim the 20-GB trial license, complete the following steps:

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.
2. Go to the **Insider Risk Management** solution.
3. Select **Forensic evidence** in the left navigation, then select **Capacity and billing**.

   > **Note**
   >
   > You can also sign up for the trial license from the **Insider Risk Management** > **Forensic evidence** > **Dashboard** tab.

4. Select **Claim 20 GB of capacity**.
5. Follow the prompts in the Microsoft 365 admin center.

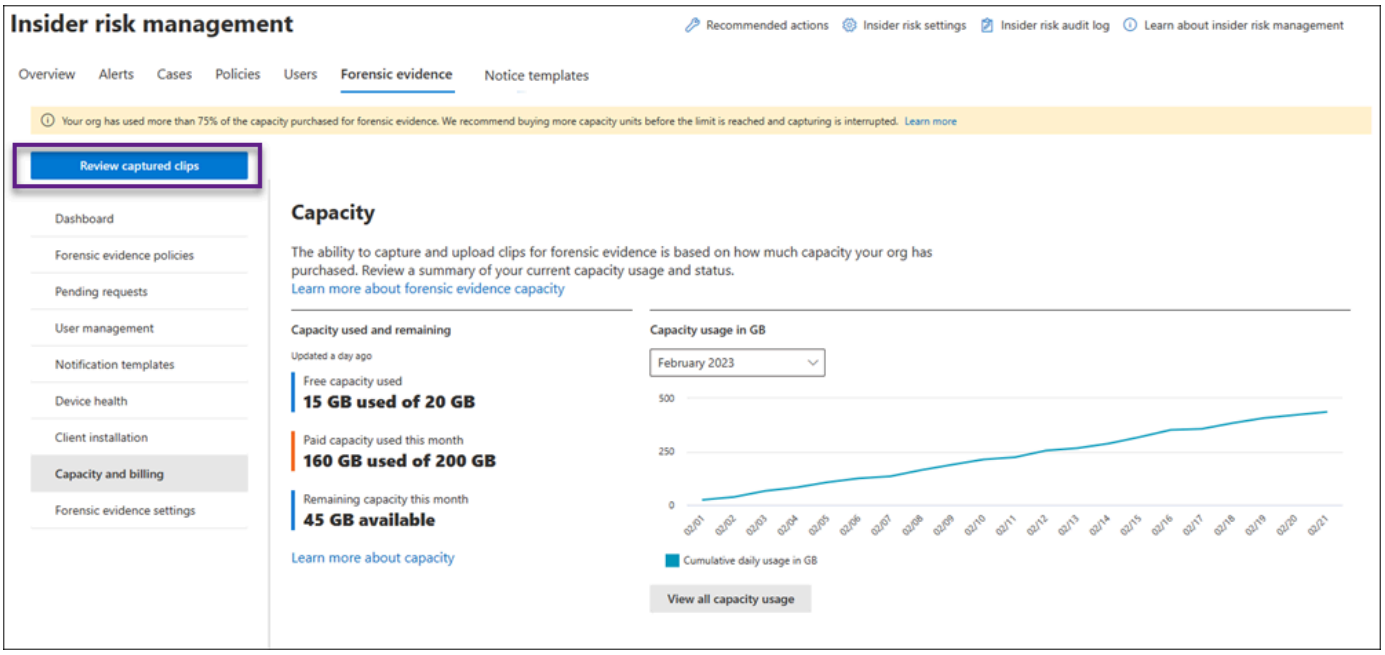
Purchase the forensic add-on for Insider Risk Management

To purchase the forensic evidence add-on, complete the following steps:

1. Go to **Microsoft 365 admin center** > **Marketplace** > **All products**.
2. Search for "forensic evidence".

Analyze your capacity

After purchasing capacity or signing up for the 20-GB trial license, use the **Capacity** page to check how much capacity you used and how much capacity remains. You can also see how much capacity you use each month by selecting an option from the **Capacity usage in GB** list or by selecting **View all capacity usage**.



> **Note**
>
> The commerce platform has a legacy billing platform and a modern billing platform. Insider Risk Management billing works with the modern billing platform. Purchased capacity applies to the ingestion of forensic evidence each month, starting on the date of purchase and resetting on the first day of every month. Any purchased capacity can be fully used in that month and resets on the first day of the next month. You can use the capacity until the date the license expires. For more information, see New commerce experience for license-based services.

Note: The author created this article with assistance from AI. Learn more

Last updated on 06/26/2026

Use reports in Insider Risk Management

> **Important**
>
> Microsoft Purview Insider Risk Management correlates various signals to identify potential malicious or inadvertent insider risks, such as IP theft, data leakage, and security violations. Insider Risk Management enables customers to create policies to manage security and compliance. Built with privacy by design, users are pseudonymized by default, and role-based access controls and audit logs are in place to help ensure user-level privacy.

Use reports in Microsoft Purview Insider Risk Management to understand the landscape of your insider risk program. Reports provide deeper insights and summary information for all areas relating to insider risk, including alerts, cases, user activity, and analytics generated from qualifying risk activities across Microsoft services to help identify insights into potential areas of risk.

The following reports are available by navigating to **Insider Risk Management** > **Reports** in the Microsoft Purview portal:

- **Alerts (preview)**: View alert report trends for generated alerts, actions taken on alerts over time, and alerts by policy.
- **Analytics**: View an analytics reports summary of anonymized user activities detected in your organization.
- **Cases (preview)**: View case report trends for created cases, actions taken on cases over time, and case status by policy and region.
- **User activity**: Review recent user activity reports, regardless if the user is included in a policy or alert.
- **Pay-as-you-go**: Review pay-as-you-go usage reports to identify cost-saving opportunities and fine-tune your policy configurations.

Work with reports

To get started and view Insider Risk Management reports, you must be a member of the applicable role or role group. Permission requirements differ for viewing reports for alerts and cases and the permissions required to view reports for analytics. If your organization uses administrative units, access might vary for these reports. For more information, see:

- Permissions for Insider Risk Management
- Permissions for Insider Risk Management and administrative units

Customize charts

For each report area, predefined report filter controls and a date selector help you quickly scope insights to specific criteria or date ranges. Select a filter at the top of the page for each area to quickly scope all the reports in the area to a filter setting. For reporting dates, select a specific month or select *Last 3 months* to view all data for the report area.

You might also want to pick and choose which reports to display by default in each report area. To customize the reports shown in each area, select **Customize view**. On the **Customize view** flyout pane, you can choose which reports to display and which reports to hide in each area.

Chart details

To explore the results in a report, select **View details** for the report. In the details view, you can filter the information and change your view by dates or policies. To export data in a report, select **Export** to download either an image of the report chart or a .csv file with the report values.

Alert reports (preview)

Investigating potentially risky user activities is an important step in minimizing insider risks for your organization. These risks might be activities that generate alerts from Insider Risk Management policies. Alert reports provide insights into alert volumes, progress of managing alerts, common alert sources, and time spent triaging. You can quickly filter all alert reports by *All alerts*, *Confirmed alerts*, and *Dismissed alerts*.

Expand table

Report type	Report	Description
Summary	Alerts generated	Shows the number of low, medium, and high severity alerts generated during the specified date range.
	Alerts acted on	Shows the number of alerts confirmed to a case or dismissed during the specified date range.
	Dismissal reasons	Shows...
Demographics	Top countries/regions with alerts	Shows the number of alerts generated for users based on what country or region they're in. *Not specified* means their country or region isn't specified in Microsoft Entra ID.

Report type	Report	Description
	Alerts generated by department	Shows the number of alerts generated for users based on what department they're in. *Not specified* means their department isn't specified in Microsoft Entra ID.
Insights	Alerts by top triggering event	Shows the number of alerts based on top triggering events detected during the specified date range.
	Alerts by top activity that generated the alert	Shows the number of alerts based on top activities detected during the specified date range.
Productivity	Alert status by assignment	Shows the number of alerts assigned to specific admins, broken down by alert status.
	Average days alerts are in *Needs Review*	Shows the average number of days that alerts remained in a *Needs review* status during the specified date range.

Analytics reports

After the first analytics scan completes for your organization, members of the *Insider Risk Management Admins* role group automatically receive an email notification. They can view the initial insights and recommendations for potentially risky activities by your users. Daily scans continue unless you turn off analytics for your organization. Admins receive email notifications for each of the three in-scope categories for analytics (data leaks, theft, and exfiltration) after the first instance of potentially risky activity in your organization. Admins don't receive email notifications for follow-up risk management activity detection resulting from the daily scans.

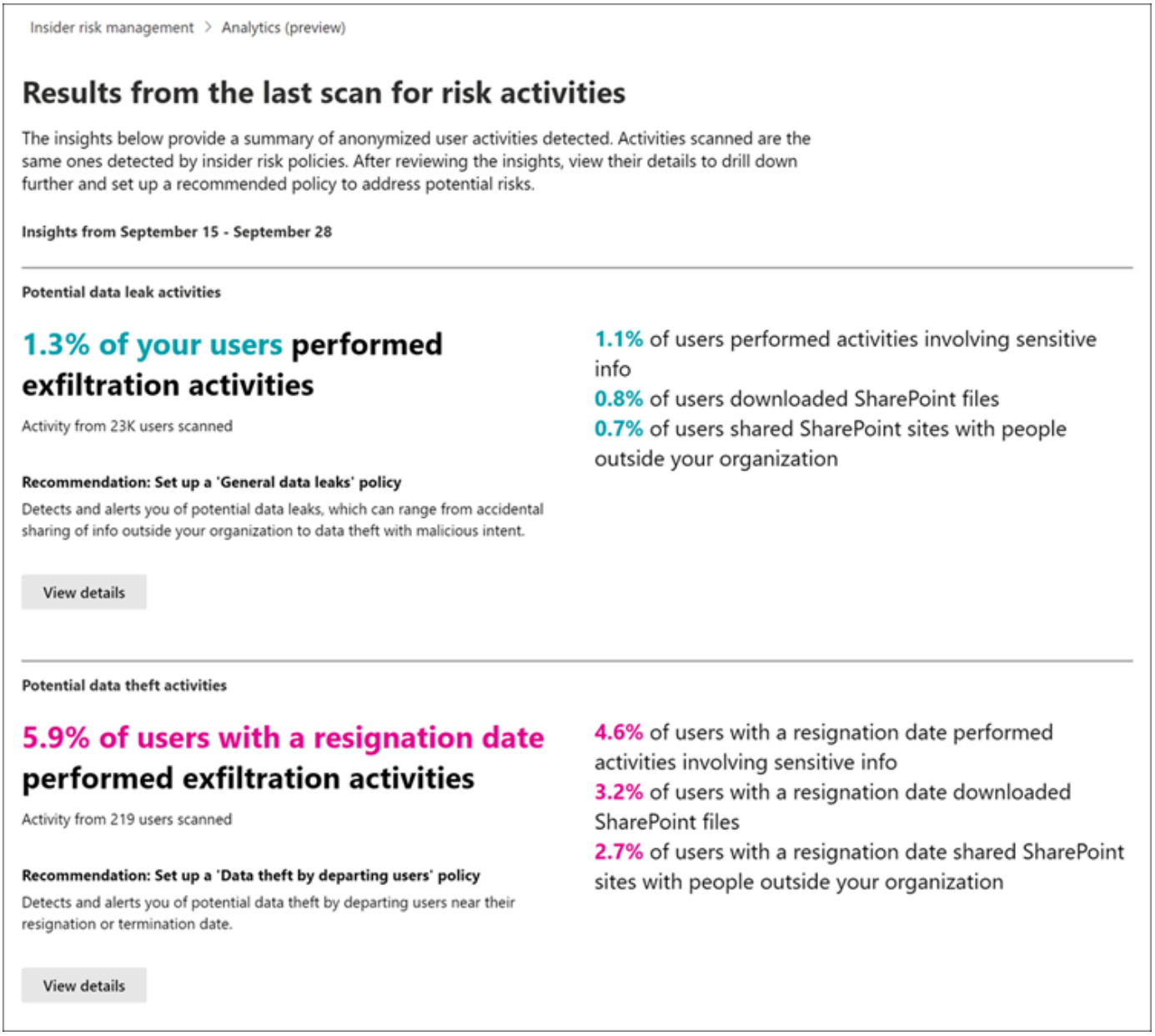
> ⓘ **Note**
>
> If you disable and then re-enable the **Analytics** setting, you reset automatic email notifications. Members of the *Insider Risk Management Admins* role group receive email notifications for new scanning insights.

To view potential risks for your organization, go to **Insider Risk Management** > **Reports** > **Analytics**.

> ⓘ **Note**
>
> If the scan for your organization isn't complete, you see a message that the scan is still active.

For completed analyses, you see the potential risks discovered in your organization and insights and recommendations to address these risks. Reports include identified risks and specific insights grouped by area, the total number of users (all types of Microsoft Entra accounts, including user, guest, system, and so on) with identified risks, the percentage of these users with potentially risky activities, and a recommended insider risk policy to help mitigate these risks. The reports include:

- **Data leaks insights**: For all users that might include accidental oversharing of information outside your organization or data leaks by users with malicious intent.
- **Data theft insights**: For departing users or users with deleted Microsoft Entra accounts that might include risky sharing of information outside your organization or data theft by users with malicious intent.
- **Top exfiltration insights**: For all users that might include sharing data outside of your organization.



To display more information for an insight, select **View details** to display the details pane for the insight. The details pane includes the complete insight results, an insider risk policy recommendation, and **Create policy** to quickly help you create the recommended policy. Selecting **Create policy** takes you to the policy workflow and automatically selects the recommended policy template related to the insight. For example, if the analytics insight is for *Data Theft* activity, the *Data Theft* policy template is pre-selected in the policy workflow for you.

Potential data theft activities

The exfiltration activities below might be related to data theft by departing users near their resignation or termination date. After reviewing them, consider setting up the recommended policy to help address potential risks.

What we detected

The following is recent activity based on a scan of 219 users who are leaving your organization.

5.9% of users with a resignation date performed exfiltration activities

4.6% of users with a resignation date performed activities involving sensitive info
3.2% of users with a resignation date downloaded SharePoint files
2.7% of users with a resignation date shared SharePoint sites with people outside your organization
2.3% of users with a resignation date shared SharePoint folders with people outside your organization
2.3% of users with a resignation date emailed people outside your organization
1.8% of users with a resignation date copied content to USB
1.8% of users with a resignation date printed a large number of files
1.4% of users with a resignation date shared SharePoint files with people outside your organization
1.4% of users with a resignation date copied sensitive content to personal cloud
0.9% of users with a resignation date shared files across network

Recommendation

Create a 'Data theft by departing users' policy that detects data theft by users near their resignation or termination date, which can range from accidental sharing of info outside your organization to data theft with malicious intent.



Case reports (preview)

Cases help you investigate and act on issues generated by risk indicators defined in your policies. You manually create cases from alerts when you need to take further action to address a compliance-related issue for a user. Case reports provide trend information to help you track the type of cases, the current status of cases, cases by region or assignment, and more. You can quickly filter all case reports by *All cases*, *Confirmed cases*, and *Benign cases*.

Expand table

Report type	Report	Description
Summary	Cases created	Shows the number of cases created during the specified date range.
	Actioned cases	Shows the number of cases with activity during the specified date range.
Demographics	Top countries/regions with cases	Shows the number of cases created for users based on their country or region. *Not specified* means their country or region isn't specified in Microsoft Entra ID.
	Top departments with cases	Shows the number of cases created for users based on their department. *Not specified* means their country or region isn't specified in Microsoft Entra ID.
Productivity	Case status by assignment	Shows the number of cases assigned to specific admins, broken down by alert status.
	Average days with active status	Shows the average number of days that cases stayed in an *Active* status during the specified date range.

Pay-as-you-go usage reports

Pay-as-you-go usage reports provide transparency and enable more accurate budget planning and policy tuning. The reports let you explore granular breakdowns of billed processing usage across different data source categories and activity indicators over time. View your usage and trends, and understand activity patterns. This information empowers you to identify cost-saving opportunities and fine-tune your policy configurations.

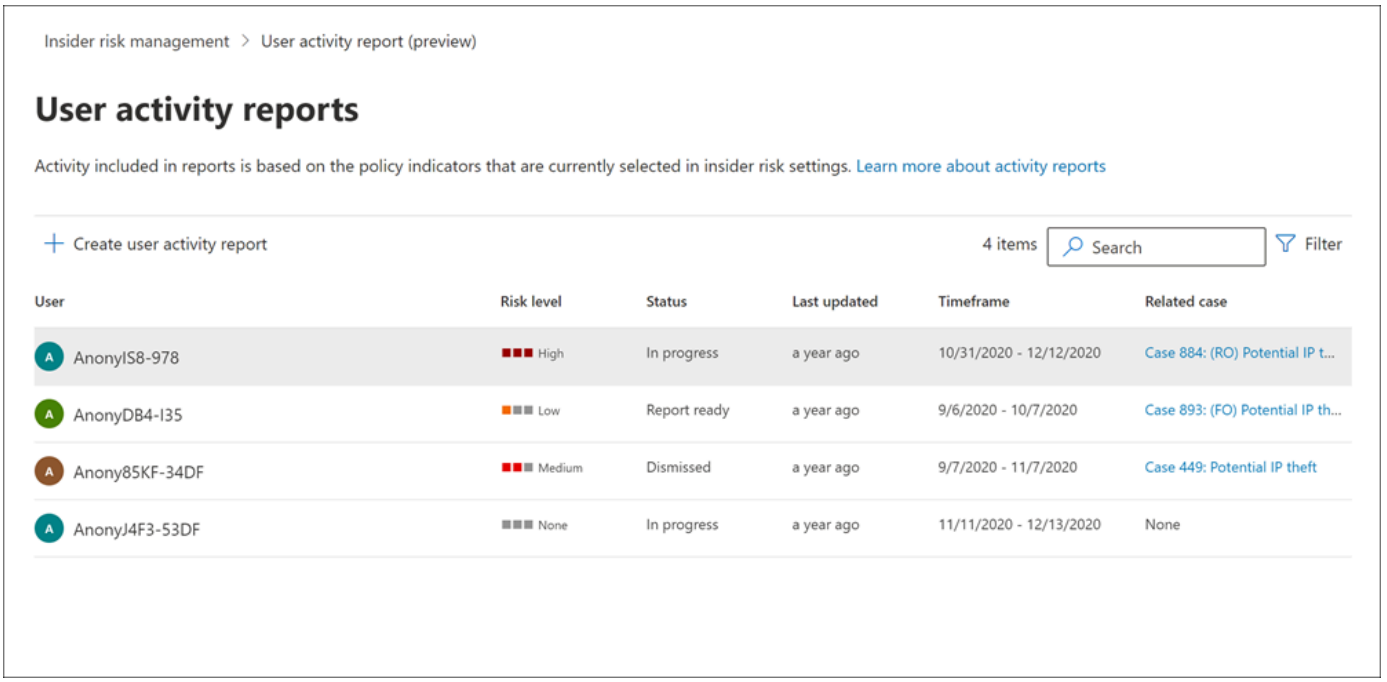
This report is available in the **Reports** page by default. For prerequisites and billing details, see Pay-as-you-go billing model.

User activity reports

User activity reports let you examine potentially risky activities for specific users during a defined time period without assigning these activities, temporarily or explicitly, to an Insider Risk Management policy. In most Insider Risk Management scenarios, you explicitly define users in policies. These users might have policy alerts (depending on triggering events) and risk scores associated with the activities. But in some scenarios, you might want to examine the activities for users that you don't explicitly define in a policy. These activities might be for users that you receive a tip about the user and potentially risky activities, or users that typically don't need to be assigned to an Insider Risk Management policy.

After you configure indicators on the Insider Risk Management **Settings** page, the solution detects user activity for potentially risky activity associated with the selected indicators. This configuration makes all detected activity for users available for review, regardless if it has a triggering event or if it creates an alert. Reports are created on a per-user basis and can include all activities for a custom 90-day period. The solution doesn't support multiple reports for the same user.

After examining potentially risky activities, investigators can dismiss individual user's activities as benign. They can also share or email a link to the report with other investigators, or choose to assign users (temporarily or explicitly) to an Insider Risk Management policy. To view the **User activity reports** page, you must assign users to the *Insider Risk Management Investigators* role group.



Create a user activity report

To create a user activity report, complete the following steps:

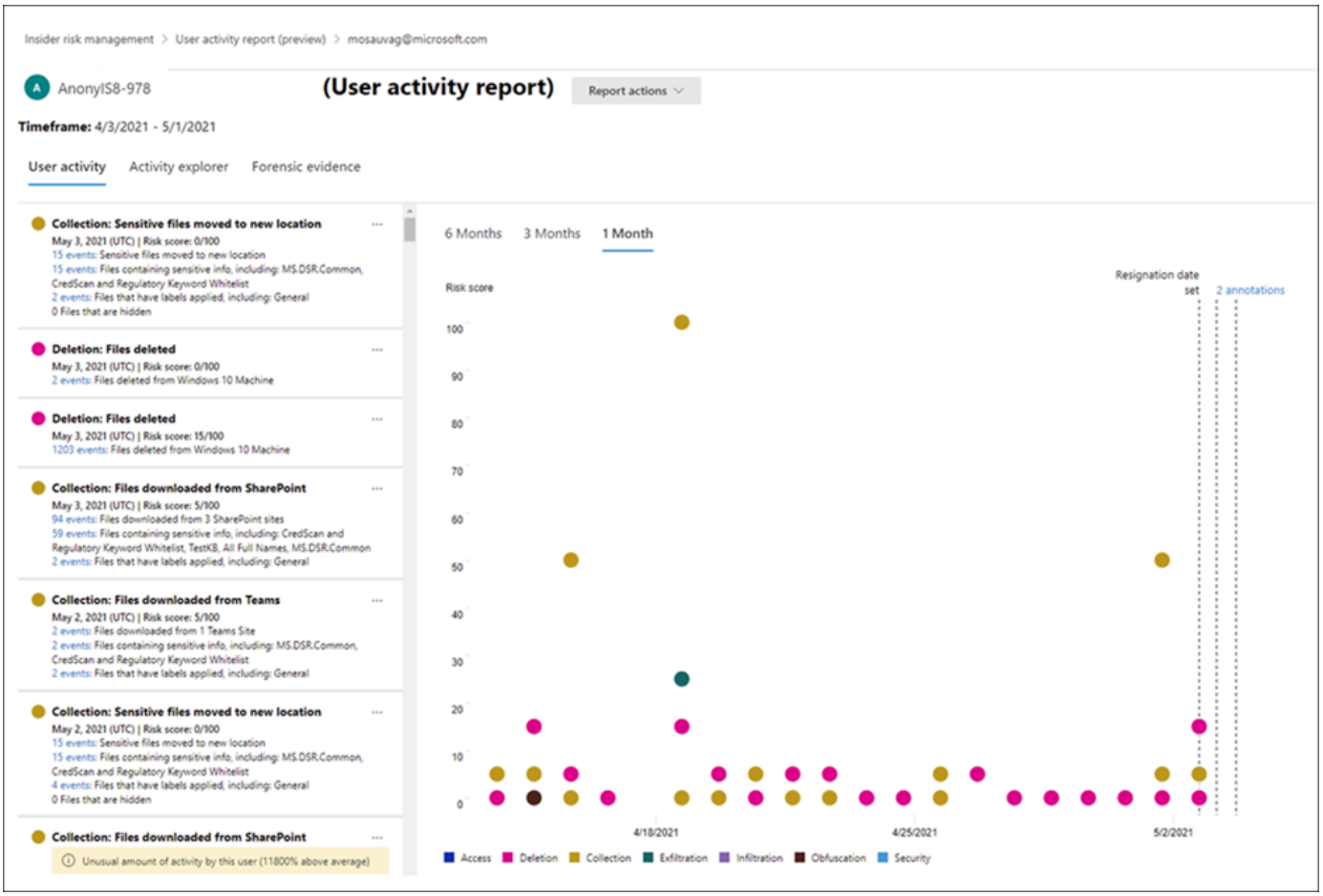
1. Go to **Insider Risk Management** > **Reports** > **User activity**.
2. Select **Create user activity report**.
3. Select a date for the **Start date**.
4. Select a date for the **End date**.
5. In the **Users** field, search for one or more users by name or user principal name.
6. Select **Create report**.

User activity data becomes available for reporting about 48 hours after the activity occurs. For example, to review user activity data for December 1, you need to wait at least 48 hours before creating the report (you can create a report on December 3 at the earliest).

New reports typically take up to 10 hours before they're ready for review. When the report is ready, *Report ready* appears in the **Status** column on the User activity report page.

View a user activity report

Select the user to view the detailed report:



The **User activity report** for the selected user contains the **User activity**, **Activity explorer**, and **Forensic evidence** tabs:

- **User activity**: Use this chart view to investigate potentially risky activities and view potentially related activities that occur in sequences. This tab is structured to enable quick review of a case, including a historical timeline of all activities, activity details, the current risk score for the user in the case, the sequence of risk events, and filtering controls to help with investigative efforts. For more information, see User activity.
- **Activity explorer**: This tab provides risk investigators with a comprehensive analytics tool that provides detailed information about activities. With the Activity explorer, reviewers can quickly review a timeline of detected risky activity and identify and filter all potentially risky activities associated with alerts. For more information, see Activity explorer.
- **Forensic evidence**: This tab allows risk investigators to review visual captures associated with risk activities included in cases from forensic evidence detection.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Learn about and configure Insider Risk Management browser signal detection

08/20/2025

> **Important**
>
> Microsoft Purview Insider Risk Management correlates various signals to identify potential malicious or inadvertent insider risks, such as IP theft, data leakage, and security violations. Insider Risk Management enables customers to create policies to manage security and compliance. Built with privacy by design, users are pseudonymized by default, and role-based access controls and audit logs are in place to help ensure user-level privacy.

In Microsoft Purview Insider Risk Management, you can use browser signal detection with the following templates:

- Data theft by departing users
- Data leaks
- Forensic evidence
- Risky browser usage (preview)

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Browsers and templates

Users often use web browsers to access both sensitive and nonsensitive files within an organization. Insider Risk Management enables your organization to detect and act on browser exfiltration signals for all nonexecutable files viewed in Microsoft Edge and Google Chrome browsers. With these signals, analysts and investigators can quickly act when in-scope policy users perform any of the following risky activities while using these browsers:

- Copying files to personal cloud storage
- Printing files to local or network devices
- Transferring or copying files to a network share

- Copying files to USB devices
- Browsing potentially risky websites

Microsoft Edge detects signals for these events by using built-in browser capabilities and the *Microsoft Compliance Extension* add-on. In Google Chrome, customers use the *Microsoft Compliance Extension* for signal detection.

The following table summarizes identified risk activities and extension support for each browser:

Expand table

Detected activities	Microsoft Edge	Google Chrome
Browsing potentially risky websites	Extension	Extension
Files copied to personal cloud storage	Native	Extension
Files copied to USB devices	Extension	Extension
Files printed to local or network devices	Native	Extension
Files transferred or copied to a network share	Extension	Extension

The following table summarizes activities by template:

Expand table

Detected activities	Data theft by departing users	Data leaks	Risky browser usage
Browsing potentially risky websites	No	No	Yes
Files copied to personal cloud storage	Yes	Yes	No
Files copied to USB devices	Yes	Yes	No
Files printed to local or network devices	Yes	Yes	No
Files transferred or copied to a network share	Yes	Yes	No

Tip

If you're not an E5 customer, use the 90-day Microsoft Purview solutions trial to explore how additional Purview capabilities can help your organization manage data security and

compliance needs. Start now at the Microsoft Purview trials hub. Learn details about signing up and trial terms.

Forensic evidence

For forensic evidence, you can capture all types of browsing activities; you're not limited to the browsing indicators of the Risky browser usage template. You can specify the desktop apps and websites that you want to include or exclude. To capture browsing activity for forensic evidence, you must install the extensions described in this article, and you must also turn on at least one risky browsing indicator in the insider risk settings.

Common requirements

Before installing the Microsoft Edge add-on or Google Chrome extension, ensure that devices for in-scope policy users meet the following requirements:

- Latest Windows 10 x64 build is recommended, minimum Windows 10 x64 build 1809 for signal detection support. Browser signal detection isn't currently supported on non-Windows devices.
- Current Microsoft 365 subscription with Insider Risk Management support.
- Devices must be onboarded to the Microsoft Purview portal.

For specific browser configuration requirements, see the Microsoft Edge and Google Chrome sections later in this article.

Additional requirements

If you use policies based on the *Risky browser usage* template, select at least one *Browsing indicator* in **Insider Risk Management** > **Settings** > **Policy indicators**.

Configure browser signal detection for Microsoft Edge

Microsoft Edge browser requirements

- Meet the common requirements
- Latest Microsoft Edge x64, version 91.0.864.41 or higher
- Latest *Microsoft Compliance Extension* add-on, version 1.0.0.44 or higher

- Edge.exe isn't configured as an unallowed browser

Option 1: Basic setup (recommended for testing with Microsoft Edge)

Use this option to configure a single-machine self-host for each device in your organization when testing browser signal detection.

1. Go to Microsoft Compliance Extension.
2. Install the extension.

Option 2: Intune setup for Microsoft Edge

Use this option to configure the extension and requirements for your organization by using Intune.

1. Sign in to the Microsoft Intune admin center.
2. Go to **Devices** > **Configuration Profiles**.
3. Select **Create Profile**.
4. Select **Windows 10 and later** as the platform.
5. Select **Settings catalog** as the profile type.
6. Select **Create**.
7. Enter a name and optional description on the **Basics** tab, then select **Next**.
8. Select **Add settings** on the **Configuration settings** tab.
9. Select **Administrative Templates** > **Microsoft Edge** > **Extensions**.
10. Select **Control which extensions are installed silently**.
11. Change the toggle to **Enabled**.
12. Enter the following value for the extensions and app IDs and update URL: `lcmcgbabdcbngcbcfabdncmoppkajglo`.
13. Select **Next**.
14. Add or edit scope tags on the **Scope tags** tab as needed, then select **Next**.
15. Add the required deployment users, devices, and groups on the **Assignments** tab, then select **Next**.
16. Select **Create**.

Option 3: Group Policy setup for Microsoft Edge

Use this option to configure the extension and requirements organization-wide by using Group Policy.

Step 1: Import the latest Microsoft Edge Administrative Template (.admx) file.

Devices must be manageable by using Group Policies. Import all Microsoft Edge Administrative Templates into the Group Policy Central Store. For more information, see How to create and manage the Central Store for Group Policy Administrative Templates in Windows.

Step 2: Add the *Microsoft Compliance Extension* add-on to the *Force Install* list.

1. In the **Group Policy Management Editor**, go to your Organizational Unit (OU).
2. Expand the following path **Computer/User configuration** > **Policies** > **Administrative templates** > **Classic administrative templates** > **Microsoft Edge** > **Extensions**. This path might vary depending on the configuration of your organization.
3. Select **Configure which extensions are installed silently.**
4. Right-click and select **Edit**.
5. Check **Enabled**.
6. Select **Show**.
7. For **Value**, add the following entry:
   lcmcgbabdcbngcbcfabdncmoppkajglo;https://edge.microsoft.com/extensionwebstorebase/v1/crx
8. Select **OK**, then select **Apply**.

Configure browser signal detection for Google Chrome

Insider Risk Management browser signal detection support for Google Chrome is enabled through the *Microsoft Compliance Extension*. This extension also supports Endpoint DLP on Chrome. For more information about Endpoint DLP support, see Get started with the Microsoft Compliance Extension (preview).

Google Chrome browser requirements

- Meet common requirements
- Latest version of Google Chrome x64
- Latest *Microsoft Compliance Extension* version (2.0.0.183 or higher)
- Chrome.exe isn't configured as an unallowed browser

Option 1: Basic setup (recommended for testing with Chrome)

Use this option to configure a single-machine self-host for each device in your organization when testing browser signal detection.

1. Go to Microsoft Compliance Extension.
2. Install the extension.

Option 2: Intune setup for Chrome

Use this option to configure the extension and requirements for your organization by using Intune.

1. Sign in to the Microsoft Intune admin center.
2. Go to **Devices** > **Configuration Profiles**.
3. Select **Create Profile**.
4. Select **Windows 10 and later** as the platform.
5. Select **Settings catalog** as the profile type.
6. Select **Custom** as the template name.
7. Select **Create**.
8. Enter a name and optional description on the **Basics** tab, then select **Next**.
9. Select **Add settings** on the **Configuration settings** tab.
10. Select **Administrative Templates** > **Google** > **Google Chrome** > **Extensions**.
11. Select **Configure the list of force-installed apps and extensions**.
12. Change the toggle to **Enabled**.
13. Enter the following value for the extensions and app IDs and update URL:
    `echcggldkblhodogklpincgchnpgcdco;https://clients2.google.com/service/update2/crx`
14. Select **Next**.
15. Add or edit scope tags on the **Scope tags** tab as needed, then select **Next**.
16. Add the required deployment users, devices, and groups on the **Assignments** tab, then select **Next**.
17. Select **Create**.

Option 3: Group Policy setup for Chrome

Use this option to configure the extension and requirements organization-wide by using Group Policy.

Step 1: Import the Chrome Administrative Template file

Your devices must be manageable by using Group Policy. Import all Chrome Administrative Templates into the Group Policy Central Store. For more information, see How to create and manage the Central Store for Group Policy Administrative Templates in Windows.

Step 2: Add the Chrome extension to the Force Install list

1. In the **Group Policy Management Editor**, go to your organizational unit (OU).
2. Expand the following path **Computer/User configuration** > **Policies** > **Administrative templates** > **Classic administrative templates** > **Google** > **Google Chrome** > **Extensions**.

This path might vary depending on the configuration for your organization.

3. Select **Configure the list of force installed extensions**.
4. Right-click and select **Edit**.
5. Select **Enabled**.
6. Select **Show**.
7. For **Value**, add the following entry:
   echcggldkblhodogklpincgchnpgcdco;https://clients2.google.com/service/update2/crx
8. Select **OK**, then select **Apply**.

Test and verify Insider Risk Management browser signal detections

1. Create an Insider Risk Management policy with device indicators enabled.

2. To test signal detection for files copied to personal cloud storage, complete the following steps from a supported Windows device:

   - Open a file sharing website (Microsoft OneDrive, Google Drive, etc.) with the browser type that you configured for signal detection.
   - With the browser, upload a nonexecutable file to the website.

3. To test signal detection for files printed to local or network devices, files transferred or copied to a network share, and files copied to USB devices, complete the following steps from a supported Windows device:

   - Open a nonexecutable file directly in the browser. You must open the file directly through File Explorer or open it in a new browser tab for viewing rather than a webpage.
   - Print the file.
   - Save the file to a USB device.
   - Save the file to a network drive.

4. After you create your first Insider Risk Management policy, you start receiving alerts from activity indicators after about 24 hours. Check the Alerts dashboard for Insider Risk Management alerts for the tested activities.

Manage the workflow with the Insider Risk Management users dashboard

> **Important**
>
> Microsoft Purview Insider Risk Management correlates various signals to identify potential malicious or inadvertent insider risks, such as IP theft, data leakage, and security violations. Insider Risk Management enables customers to create policies to manage security and compliance. Built with privacy by design, users are pseudonymized by default, and role-based access controls and audit logs are in place to help ensure user-level privacy.

The **Users dashboard** is an important tool in the Insider Risk Management workflow. It helps investigators and analysts have a more complete understanding of risk activities. This dashboard offers views and management features to meet administrative needs between creating Insider Risk Management policies and managing Insider Risk Management cases.

After users are added to Insider Risk Management policies, background processes automatically evaluate user activities for triggering indicators. After the process identifies triggering indicators, it assigns risk scores to user activities. Some of these activities result in an insider risk alert, but some activities don't meet a minimum risk score level and an insider risk alert isn't created. The **Users dashboard** allows you to view users with these types of indicators and risk scores, as well as users that have active insider risk alerts.

Learn more about how the Users dashboard displays users in the following scenarios:

- Users with active insider risk policy alerts
- Users with triggering events
- Users identified as potential high impact users or in priority user groups
- Users added temporarily to policies

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Users with active insider risk policy alerts

The **Users dashboard** automatically displays all users with active insider risk policy alerts. These users with alerts have both a triggering indicator and an activity risk score that meets the requirements for creating an insider risk alert. You can view activities for these users by selecting the user in the **Users dashboard** and navigating to the **User activity** tab.

Users with triggering events

The **Users dashboard** automatically displays all users with triggering events, but that don't have an activity risk score that would create an insider risk alert. For example, a user with a reported resignation date is displayed because this activity is a triggering event but isn't an activity that has a risk score. You can view activities for these users by selecting the user in the **Users dashboard** and navigating to the **User activity** tab.

Users added temporarily to policies

The **Users dashboard** shows users you add to Insider Risk Management policies after an unusual event outside of the Insider Risk Management workflow. Temporarily adding users (from the Policies dashboard) is also a way to start scoring user activity for an Insider Risk Management policy for testing the policy, even if you don't configure a required connector.

When you manually add a user to a policy, the solution scores user activities for the previous 90 days and adds them to the **User activity** timeline. For example, you have a user not currently being assigned risk scores for an insider risk policy and the user has data leak activities reported to the legal department in your organization. The legal department recommends that you configure new short-term detection requirements for the user. You can temporarily assign the user to your *Data leaks* policy for a designated length of time (activation window). All users added temporarily are displayed in the **Users dashboard** because triggering event requirements are waived.

> **Note**
>
> It might take several hours for new manually added users to appear in the **Users dashboard**. Activities for the previous 90 days for these users might take up to 24 hours to display. To view activities for manually added users, select the user on the **Users dashboard** and open the **User activity** tab on the details pane.

The user is automatically removed from the **Users dashboard** and scoring stops when the time defined in the **Activation window** expires if:

- the user doesn't have any additional triggering events or insider risk policy alerts, and
- the manually defined **Activation window** duration is longer than the global policy **Activation window** duration.

The **Activation window** setting with the longest duration always overrides the **Activation window** setting with a shorter duration. For example, you configure the **Activation window** on the global **Policy timeframes** tab in the Insider Risk Management global settings for 15 days, which automatically applies to all your insider risk policies.

You temporarily add a user to your *Data leaks* insider risk policy and define 30 days as the **Activation window** for this user. The global **Activation window** setting of 15 days is overridden by defining the **Activation window** setting of 30 days for the temporarily added user. The temporarily added user remains in the **Users dashboard** and is in-scope for the policy for 30 days.

In the opposite scenario where the global **Activation window** setting is longer than the **Activation window** setting defined for a temporarily added user, the global **Activation window** setting overrides the **Activation window** setting for the temporarily added user. The temporarily added user remains in the **Users dashboard** and is in-scope for the policy for the number of days defined in the global **Activation window** settings.

View user information on the Users dashboard

Each user displayed in the **Users dashboard** shows the following information:

- **Users**: Username for a user. This field is anonymized if the global anonymization setting for Insider Risk Management is enabled.
- **Alert severity level**: Current calculated risk level of the user. The system calculates this score every 24 hours and uses the alert risk scores from all active alerts associated with the user. For users with only triggering indicators, the alert severity level is zero.
- **Active alerts**: Number of active alerts for all policies.
- **Confirmed violations**: Number of cases resolved as *confirmed policy violation* for the user.
- **Case**: Current active case for the user.

To quickly locate a specific user, use **Search** at the top right of the Users dashboard. When searching for users, you must use the user principal name (UPN). For example, when searching for

a user named 'Tiara Hidayah' that has a UPN of 'thidayah' in your organization, enter 'thidayah' or some part of the UPN in **Search**.



> **Note**
>
> The number of users displayed on the **Users dashboard** might be limited in some instances, depending on the volume of active alerts and matching policies. The **Users dashboard** displays users with active alerts as the system generates the alerts. In rare cases, the maximum number of displayed users is reached. If the maximum number of displayed users is reached, the **Users dashboard** adds users with active alerts who aren't displayed as existing user alerts are triaged.

View user details

To view more details about risk activity for a user, open the user details pane by double-clicking a user in the **Users dashboard**. On the details pane, you can view the following information:

- **User profile** tab
  - **Name and title**: Name and position title for the user from Microsoft Entra ID. These user fields are anonymized or empty if you enable the global anonymization setting for Insider Risk Management.
  - **User details**: Shows whether the user is a potential high impact user or if the user is in priority user groups.
  - **Alert and activity summary**: Shows active user alerts and open cases.
  - **User email**: Email address for the user.
  - **Alias**: Network alias for the user.
  - **Organization or department**: Organization or department for the user.
  - **In scope**: Shows in-scope assignment of the user to policies.

The following additional user profile signals are available in the unified alert experience (preview). To access, select an alert on the **Alerts (preview)** tab, scroll down to **User details**, and select **View user details**:

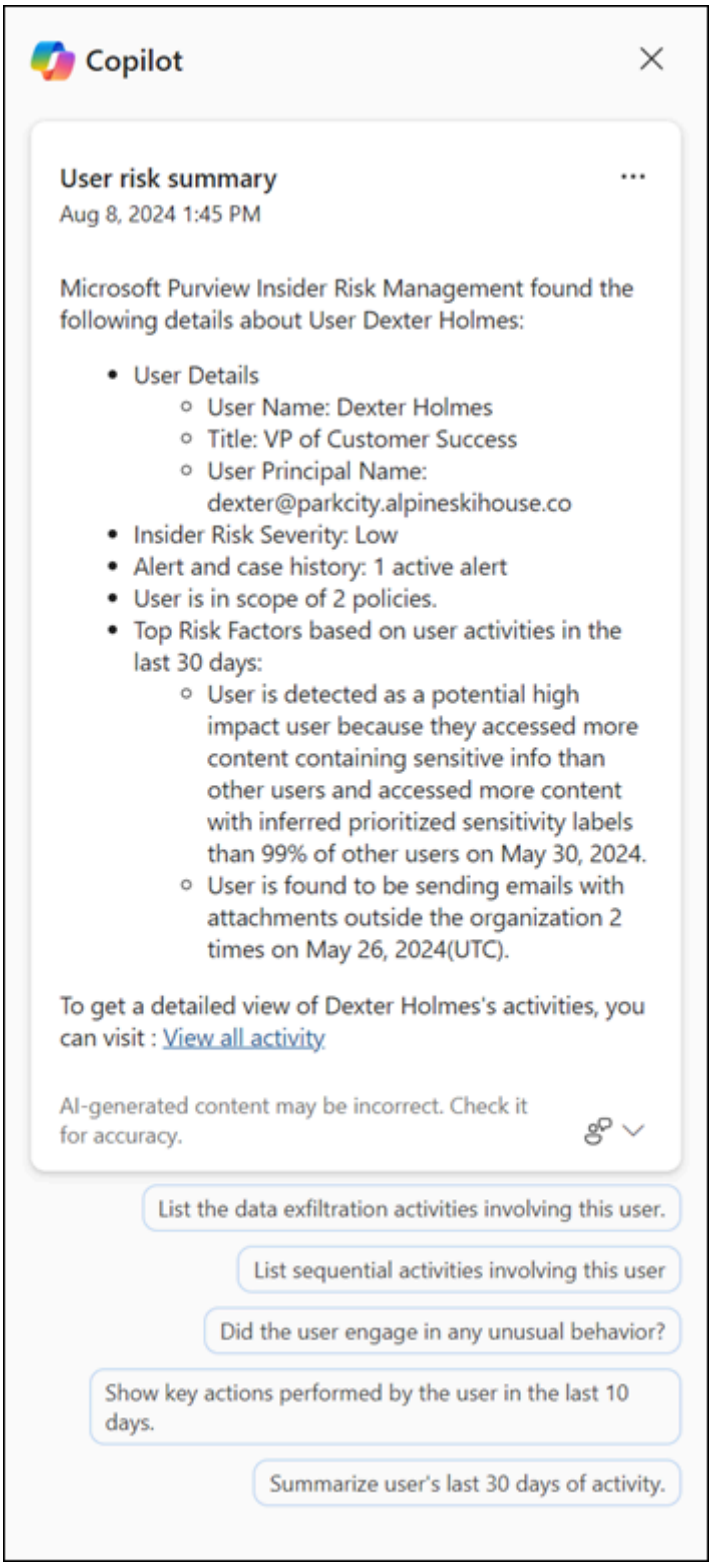
- **Office location**: The user's office location from their Entra profile.
- **Employee type**: The user's employee type from their Entra profile.
- **Department**: The user's department from their Entra profile.
- **Last working date**: The user's last scheduled date of work, if available.
- **Past alert and case history**: A summary of the user's previous alerts and cases across Insider Risk Management.
- **Priority user group status**: Shows whether the user belongs to any priority user groups.
- **Policy inclusion**: Lists the policies in which the user is currently in scope.

> ⓘ **Note**
>
> When pseudo-anonymization is enabled, the expanded user profile details aren't visible.

- **User activity** tab
  - **History of recent user activity**: Shows both triggering indicators and insider risk indicators for risk activities up to the last 90 days. All risk activities pertinent to insider risk indicators are also scored, though the activities might or might not have generated an insider risk alert. Triggering indicator examples might be a resignation date or the last scheduled date of work for the user. Insider risk indicators are activities determined to have an element of risk, which might potentially lead to a security incident, and are defined in policies that the user is included in. Event and risk activities are listed with the most recent item listed first.

Use Copilot to summarize user activities (preview)

Select **Summarize with Copilot** on the user details pane to quickly summarize the activities for users that might need further investigation. When you summarize user risk activities with Microsoft Copilot in Microsoft Purview, a **Copilot** pane appears on the right side of the screen with a user summary.



The user summary includes essential details such as the user's name, title, user principal name, alert and case history, and top risk factors. The top risk factors provide crucial insights into the user roles, permissions, involvement in exfiltration activities, sequential patterns, or any unusual behavior. This view helps you identify users who might pose the highest insider risk to your organization.

Suggested prompts are automatically listed to help further refine your summary and to help provide additional insights to the activities associated with the user. Choose from the following suggested prompts:

- *List all the data exfiltration activities involving this user.*
- *List all the sequential activities involving this user.*
- *Did the user engage in any unusual behavior?*
- *Show key actions performed by the user in the last 10 days.*
- *Summarize user's last 30 days of activity.*

> **Tip**
>
> You can also use the standalone version of Microsoft Security Copilot to investigate Insider Risk Management, Microsoft Purview data loss prevention (DLP), and Microsoft Defender XDR alerts.

Remove users from in-scope assignment to policies

You might need to stop assigning risk scores to a user in Insider Risk Management policies. Use **Stop scoring activity for users** on the **Users dashboard** to stop assigning risk scores for a user from all Insider Risk Management policies that they're currently in scope for. Stopping scoring activity doesn't remove the user from the overall policy assignment (when you add users or groups to a policy configuration), but simply removes the user from active processing by policies after current triggering events. If the user has another triggering event in the future, risk scores from policies automatically begin to be assigned to the user again. Any existing alerts or cases for this user don't remove.

Remove a user from in-scope status in all Insider Risk Management policies

To remove a user from in-scope status in all Insider Risk Management policies, complete the following steps:

1. Sign in to the Microsoft Purview portal with an admin account in your Microsoft 365 organization.
2. Go to the **Insider Risk Management** solution.
3. Select **Users** in the left navigation.
4. On the **Users dashboard**, select the users you want to stop scoring activity for.
5. Select **Stop scoring activity for users**.

> **Note**

Removing a user from in-scope status might take several minutes. When the process completes, the user doesn't appear on the **Users dashboard**. If the removed user has active alerts or cases, the user remains on the **Users dashboard** and the user details show that they are no longer in-scope for a policy.

Run automated tasks with Power Automate flows for a user

Using recommended Power Automate flows, risk investigators and analysts can quickly take action to notify users when they're added to an insider risk policy.

To run, manage, and create Power Automate flows for Insider Risk Management users:

1. Select **Automate** on the user action toolbar.
2. Choose the Power Automate flow to run, then select **Run flow**.
3. After the flow completes, select **Done**.

For more information about Power Automate flows for Insider Risk Management, see Getting started with Insider Risk Management settings.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/22/2026

Create Insider Risk Management notice templates

> **Important**
>
> Microsoft Purview Insider Risk Management correlates various signals to identify potential malicious or inadvertent insider risks, such as IP theft, data leakage, and security violations. Insider Risk Management enables customers to create policies to manage security and compliance. Built with privacy by design, users are pseudonymized by default, and role-based access controls and audit logs are in place to help ensure user-level privacy.

Insider Risk Management notice templates enable you to automatically send email messages to users when you create a case for risk activities that generate a policy match and confirmed alert. For most alerts that generate cases, user actions result from mistakes or inadvertent actions without ill intent. Use notices as simple reminders for users to be more careful, to provide links to information for refresher training, or to corporate policy resources. Notices can be an important part of your internal compliance training program and can help create a documented audit trail for users with recurring risk activities.

Create notice templates if you want to send users an email reminder notice for policy matches as part of the case resolution process. You can only send notices to the user email address associated with the specific case being reviewed. When selecting a notice template to apply to a policy match, you can choose to accept the field values defined in the template or overwrite the fields as needed.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Notice templates dashboard

The **Notices templates dashboard** displays a list of configured notice templates and lets you create new notice templates. The portal lists the notice templates in reverse date order with the most recent notice template listed first.



HTML for notices

If you want to create more than a simple text-based email message for notifications, you can create a more detailed message by using HTML in the message body field of a notice template. The following example provides the message body format for a basic HTML-based email notification template:

HTML

```
<!DOCTYPE html>
<html>
<body>
<h2>Action Required: Contoso User Code of Conduct Policy Training</h2>
<p>A recent activity you've performed has generated a risk alert prohibited by the
Contoso User <a href='https://www.contoso.com'>Code of Conduct Policy</a>.</p>
<p>You are required to attend the Contoso User Code of Conduct <a
href='https://www.contoso.com'>training</a> within the next 14 days. Please contact
<a href='mailto:hr@contoso.com'>Human Resources</a> with any questions about this
training request.</p>
<p>Thank you,</p>
<p><em>Human Resources</em></p>
</body>
</html>
```

> **Note**
>
> In Insider Risk Management notice templates, the HTML href attribute implementation currently supports only single quotation marks instead of double quotation marks for URL references.

Create a new notice template

To create a new Microsoft Purview Insider Risk Management notice template, use the notice creation tool in the **Insider Risk Management** solution in the Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.
2. Go to the **Insider Risk Management** solution.
3. Select **Notification templates** in the left navigation.
4. Select **Create notification template** to open the notice creation tool.
5. On the **New email notification template** page, complete the following fields:

    - **Template name**: Enter a friendly name for the notice. This name appears on the list of notices on the notice dashboard and in the notice selection list when sending notices from a case.
    - **Send from**: Enter the sender email address for the notice. This address appears in the **From:** field in all notices sent to users unless changed when sending a notice from a case.
    - **Subject**: Enter the information that appears in the subject line of the message. You can enter text values.
    - **Message body**: Enter the information that appears in the message body. You can enter text or HTML values.

6. Select **Create** to create and save the notice template.

Update a notice template

To update an existing notice template, complete the following steps:

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.
2. Go to the **Insider Risk Management** solution.
3. Select **Notification templates** in the left navigation.
4. In the **Email notification templates** list, select the notification template that you want to manage.
5. On the notice details page, select **Edit**.
6. On the **Edit** page, edit the following fields:

    - **Template name**: Enter a new friendly name for the notice. This name appears on the list of notices on the notice dashboard and in the notice selection list when sending notices from a case.
    - **Send from**: Update the sender email address for the notice. This address appears in the **From:** field in all notices sent to users unless changed when sending a notice from a case.

- **Subject**: Update the information that appears in the subject line of the message. You can enter text characters.
- **Message body**: Update the information that appears in the message body. You can enter text characters or HTML values.

7. Select **Save** to update and save the notice.

Delete a notice template

To delete a notice template that you no longer need, complete the following steps:

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.
2. Go to the **Insider Risk Management** solution.
3. Select **Notification templates** in the left navigation.
4. In the **Email notification templates** list, select the notification template that you want to delete.
5. Select the **Delete** icon on the toolbar.
6. Select **Yes**.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Limits in Insider Risk Management

08/20/2025

> **Important**
>
> Microsoft Purview Insider Risk Management correlates various signals to identify potential malicious or inadvertent insider risks, such as IP theft, data leakage, and security violations. Insider Risk Management enables customers to create policies to manage security and compliance. Built with privacy by design, users are pseudonymized by default, and role-based access controls and audit logs are in place to help ensure user-level privacy.

This article describes the limits in Microsoft Purview Insider Risk Management.

Global settings

Expand table

Item	Limit
Lookback period limits (Exchange Online).	10 days
Lookback period limits for all other signals.	90 days
Maximum number of custom indicators.	10
Maximum number of fields per custom indicator.	20
Maximum number of items in a detection group	200
Maximum number of items in each global exclusion list (Domains, SharePoint sites, File paths, Keywords, and File types).	500 for each list
Maximum number of users that you can add to a priority user group.	10,000
Maximum number of variants per indicator.	3

Triggers (per UTC calendar day)

Expand table

Item	Limit
Custom indicators	15,000
All other triggers	5,000
All signals collected through the HR connector	15,000
Maximum trigger volume for an organization	50,000
User account deleted from Microsoft Entra ID	15,000

> **Note**
>
> Limitations apply to each individual trigger type.

Maximum number of users in scope for a policy template

Expand table

Template name	Limit
Data leaks by priority users	1,000
Data leaks by risky users	7,500
Data leaks	15,000
Data theft by departing users	20,000
Forensic evidence	Unlimited
Patient data misuse (preview)	5,000
Risky AI usage	10,000
Risky browser usage (preview)	7,000
Security policy violations by departing users	15,000
Security policy violations by priority users	1,000
Security policy violations by risky users	7,500
Security policy violations	1,000

> **Note**
>
> You can add any number of users to a policy. The limit applies to users *in scope* of a policy template (users brought in scope after a triggering event).

Other policy limits

Expand table

Item	Limit
Maximum number of policies that you can create per template type.	100
Maximum number of priority file extensions.	50
Maximum number of priority sensitive info types.	50
Maximum number of priority sensitivity labels.	50
Maximum number of priority sites.	50
Maximum number of priority trainable classifiers.	5

> **Note**
>
> The **Users in scope** column on the **Policies** tab shows the number of in-scope users for a policy.

Adaptive protection

Expand table

Item	Limit
Maximum number of users that you can include in a Data Loss Prevention (DLP) policy for each risk level.	10,000

Manual user scoring

Expand table

Item	Limit
Maximum number of users that you can score manually.	4,000

Cases

Expand table

Item	Limit
Maximum number of active cases.	100

Exporting

Expand table

Item	Limit
Maximum number of alerts that you can export from the **Alerts** page.	1,000
Maximum number of logs that you can export to a CSV file from Activity explorer.	100,000
Maximum number of users that you can export from the **Users** page.	1,000

Retention limits for alerts, cases, and associated artifacts

As Insider Risk Management alerts age, their value to minimize potentially risky activity diminishes for most organizations. Conversely, active cases and associated artifacts (alerts, insights, activities) are always valuable to organizations and shouldn't have an automatic expiration date. This value includes all future alerts and artifacts in an active status for any user associated with an active case.

To help minimize the number of older items that provide limited current value, the following retention limits apply for Insider Risk Management alerts, cases, and user reports:

Expand table

Item	Retention period
Active cases (and associated artifacts)	Indefinite retention, never expire.

Item	Retention period
Alerts with Needs review or Dismissed status	120 days from alert creation, then automatically deleted.
Resolved cases (and associated artifacts)	120 days from case resolution, then automatically deleted.
User activities reports	120 days from report creation, then automatically deleted.

Connectors

Expand table

Item	Limit
Maximum number of records in the JSON file that the API processes	50,000

Learn about Information Barriers

Microsoft Purview Information Barriers (IB) is a compliance solution that restricts two-way communication and collaboration between groups and users in Microsoft Teams, SharePoint, and OneDrive. Often used in highly regulated industries, IB helps avoid conflicts of interest and safeguards internal information between users and organizational areas.

When you create IB policies, users who can't communicate or share files with other specific users can't find, select, chat, or call those users. IB policies automatically put checks in place to detect and prevent unauthorized communication and collaboration among defined groups and users. IB policies are independent from compliance boundaries for eDiscovery investigations that control user content locations that eDiscovery managers can search.

IB policies can allow or prevent communication and collaboration between groups and users for the following example scenarios:

- Users in the *Day Trader* group can't communicate or share files with the *Marketing Team*.
- Instructors in one school can't communicate or share files with students in another school in the same school district.
- Finance personnel working on confidential company information can't communicate or share files with certain groups within their organization.
- An internal team with trade secret material can't call or chat online with users in certain groups within their organization.
- A research team can only call or chat online with a product development team.
- A SharePoint site for *Day Trader* group can't be shared or accessed by anyone outside of the *Day Trader* group.
- A lawyer's data obtained from one client can't be accessed by a lawyer at the same firm who represents a different client.
- Government information access and control are limited across departments and groups.
- A group of people in a company can only chat with a client or a specific customer via guest access during a customer engagement.

> **Important**

Information Barriers **only supports** two-way communication and collaboration restrictions. For example, a scenario where Marketing can communicate and collaborate with Day Traders, but Day Traders can't communicate and collaborate with Marketing **isn't supported**.

> **Important**
>
> Information barrier (IB) policies can't restrict communication and collaboration between groups and users in email messages including Exchange Online. If your organization needs to define and control email communications, consider using Exchange mail flow rules.

Information Barriers and Microsoft Teams

In Microsoft Teams, IB policies determine and prevent the following kinds of unauthorized communication and collaboration:

- Searching for a user
- Adding a member to a team
- Starting a chat session with someone
- Starting a group chat
- Inviting someone to join a meeting
- Sharing a screen
- Placing a call
- Sharing a file with another user
- Accessing a file through a sharing link

If the users conducting these activities in Microsoft Teams are included in an IB policy to prevent the activity, they can't proceed. In addition, everyone included in an IB policy can be potentially blocked from communicating with other users in Microsoft Teams. When users affected by IB policies are part of the same team or group chat, they might be removed from those chat sessions and further communication with the group might not be allowed.

For more information, see Information Barriers in Microsoft Teams.

Information Barriers and SharePoint and OneDrive

In SharePoint and OneDrive, IB policies detect and prevent the following kinds of unauthorized collaboration:

- Adding a member to a site
- Accessing site or content by a user
- Sharing site or content with another user
- Searching a site

For more information, see Information Barriers in SharePoint and Information Barriers in OneDrive.

Information Barriers and Microsoft Planner

As a work management tool, Microsoft Planner enables users to collaborate on plans and tasks. If your compliance admin configures IB policies to restrict communication and collaboration between user segments, Microsoft Planner supports these restrictions.

IB policies allow administrators to enable or disable search restrictions in the people picker. By using IB support in Planner, when a user searches for others in the people picker to share a plan or to assign a task, they don't see users from segments they're restricted from communicating with. This restriction prevents users from one segment from sharing plans or assigning tasks to users in another segment.

IB support in Microsoft Planner is available only for basic plans in the following applications:

- Planner Web
- Planner in Teams web
- Planner in Teams desktop
- Planner in Teams mobile app

Policy behavior

When IB policy administrators create a new policy or modify an existing policy, users can still access existing plans shared with them or already assigned tasks. For any subsequent plan sharing or task assignment, an IB policy check is triggered and collaboration is permitted or restricted as defined by the policy.

Information Barriers and Exchange Online

Information barrier (IB) policies can't restrict communication and collaboration between groups and users in email messages. If your organization needs to define and control email communications, consider using Exchange mail flow rules. **Enabling Information Barriers in**

environments without configured ABPs can result in the loss of Address List visibility between all users.

The following table summarizes the key differences between IB modes for Exchange Online Address Book Policies (ABPs):

Expand table

Feature	Single and multi-segment modes	Legacy mode
ABP management	Automatic; no reliance on existing ABPs	Automatic; existing ABPs must be removed first
ABP for unsegmented users	System creates ABP with empty address lists	Not applicable
Custom ABP changes	Allowed; keep consistent with IB segments	Not supported; IB controls all ABPs
Hierarchical address book	Available for users not in an IB segment	Available for users not in an IB segment
Key prerequisite	None for ABPs	Remove all existing ABPs before enabling IB

Information Barriers and Exchange for single and multi-segment modes

If your organization uses *single* or multi-segment mode, Information Barriers no longer relies on Exchange Online Address Book Policies (ABPs). Information barrier (IB) policies can't restrict communication in email messages. If your organization needs to define and control email communications, consider using Exchange mail flow rules.

Enabling Information Barriers doesn't affect organizations that use ABPs. If users don't have an ABP defined with associated IB segments and policies, the system **automatically creates an ABP with empty address lists for these users**. You can change these ABPs as needed, they are not automatically adjusted based on your IB Segmentation. Keep your ABPs consistent with the segments you configure in Information Barriers to avoid losing proper Address List visibility. Avoid user visibility differences between your existing ABPs and your new Information Barriers configuration.

Information Barriers and Exchange for legacy mode

If your organization uses *legacy* mode, IB policies rely on Exchange Online Address Book Policies (ABPs). ABPs let organizations virtually assign users into specific groups to provide customized views of the organization's global address list (GAL). When you create IB policies, the system automatically creates ABPs for the policies. As you add IB policies in your organization, the structure and behavior of your GAL changes to comply with IB policies.

Before you define and apply IB policies, remove all existing Exchange address book policies in your organization. IB policies rely on address book policies, and existing ABPs aren't compatible with the ABPs that IB creates. To remove your existing address book policies, see Remove an address book policy in Exchange Online. When you enable IB policies and enable hierarchical address book, all users not included in an IB segment see the hierarchical address book in Exchange Online.

Ready to get started?

- Get started with Information Barriers: Define segments and create your first IB policies.
- Information Barriers attributes: Review the Microsoft Entra attributes available for defining user segments.
- Use multi-segment support in Information Barriers: Assign users to multiple segments for complex organizational scenarios.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 09/17/2026

Get started with Information Barriers

This article describes how to configure Information Barriers (IB) policies in your organization. Several steps are involved, so review the entire process before you begin configuring IB policies.

Use the Microsoft Purview portal or Security & Compliance PowerShell to configure IB in your organization. For organizations configuring IB for the first time, use the **Information Barriers** solution in the Microsoft Purview portal. If you're managing an existing IB configuration and you're comfortable using PowerShell, you can still use that option.

For more information about IB scenarios and features, see Learn about Information Barriers.

> **Tip**
>
> To help you prepare your plan, this article includes an **example scenario**.

Required subscriptions and permissions

Before you get started with IB, confirm your Microsoft 365 subscription and any add-ons. To access and use IB, your organization must have supporting subscriptions or add-ons. For more information, see the subscription requirements for Information Barriers.

To manage IB policies, you must be assigned one of the following roles:

- Microsoft 365 global administrator
- Office 365 global administrator
- Compliance administrator
- IB Compliance Management

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview **roles and permissions**.

Configuration concepts

When you configure IB, you work with several objects and concepts.

- **User account attributes**: Microsoft Entra ID (or Exchange Online) defines these attributes. These attributes can include department, job title, location, team name, and other job profile details. You assign users or groups to segments by using these attributes.
- **Segments**: You define these sets of groups or users in the Microsoft Purview portal or by using PowerShell. They use selected group or user account attributes.

  Your organization can have up to 5,000 segments, and users can be assigned to a maximum of 10 segments. See the list of IB supported attributes for details.

  > **ⓘ Important**
  >
  > Support for 5,000 segments and assigning users to multiple segments is only available when your organization isn't in *Legacy* mode. Assigning users to multiple segments requires extra steps to change the Information Barriers mode for your organization. For more information, see **Use multi-segment support in Information Barriers**.
  >
  > For organizations in *Legacy* mode, the maximum number of segments supported is 250, and users are restricted to being assigned to only one segment. Organizations in *Legacy* mode are eligible to upgrade to the newest version of Information Barriers in the future. For more information, see the **Information Barriers roadmap**.

- **IB policies**: These policies determine communication limits or restrictions. When you define IB policies, you choose from two kinds of policies:
  - *Block* policies prevent one segment from communicating with another segment.
  - *Allow* policies allow one segment to communicate with only certain other segments.

    > **ⓘ Note**
    >
    > **For organizations in *Legacy* mode**: Non-IB groups and users **aren't visible** to users included in IB segments and policies for *allow* policies. If you need non-IB groups and users to be visible to users included in IB segments and policies, you must use *block* policies.
    >
    > **For organizations in *SingleSegment* or *MultiSegment* mode**: Non-IB groups and users are visible to users included in IB segments and policies.

To verify your IB mode, see Check the IB mode for your organization.

- **Policy application**: You apply all IB policies after you define them.
- **Visibility of non-IB users and groups**: Non-IB users and groups are users and groups excluded from IB segments and policies. Depending on when you configure IB policies in your organization and the type of IB policies (block or allow), the behavior for these users and groups differs in Microsoft Teams, SharePoint, OneDrive, and in your global address list.
  - **For organizations in *Legacy* mode**: For users defined in *allow* policies, non-IB groups and users aren't visible to users included in IB segments and policies. For users defined in *block* policies, non-IB groups and users are visible to users included in IB segments and policies.
  - **For organizations in *SingleSegment* or *MultiSegment* mode**: Non-IB groups and users are visible to users included in IB segments and policies.
- **Group support**: IB currently supports only Microsoft 365 Groups. Distribution Lists and Security Groups are treated as non-IB groups.
- **Hidden or disabled user and guest accounts**: For hidden or disabled user and guest accounts in your organization, the *HiddenFromAddressListEnabled* parameter is automatically set to *True* when you hide or disable user accounts, or when you create a guest. When the organization mode is *Legacy* for IB-enabled organizations, IB policies aren't enforced for these accounts in Microsoft Teams. Administrators can change this behavior by manually setting the *HiddenFromAddressListEnabled* parameter to *False*, which allows IB policy enforcement for these users. When the organization is in *SingleSegment* or *MultiSegment* mode, IB policies are enforced for hidden users.

Configuration overview

- Step 1: Make sure prerequisites are met: Verify required subscriptions, permissions, directory data for segmenting users, audit logging, IB mode, and Exchange address book policy configuration.
- Step 2: Segment users in your organization: Determine what policies you need, identify attributes to use, and define segments in terms of policy filters.
- Step 3: Create Information Barriers policies: Create your policies (don't apply them yet) and select from two kinds (block or allow).

- **Step 4: Apply Information Barriers policies**: Set policies to active status, run the policy application, and view policy status.
- **Step 5: Configure Information Barriers on SharePoint and OneDrive (optional)**: Configure IB for SharePoint and OneDrive.
- **Step 6: Information Barriers modes (optional)**: Update IB modes if applicable.
- **Step 7: Configure user discoverability (optional)**: Enable or restrict user discoverability in the people picker if applicable.

Step 1: Make sure prerequisites are met

In addition to the required subscriptions and permissions, make sure that you meet the following requirements before configuring IB:

- **Directory data**: Make sure that your organization's structure is reflected in directory data. To take this action, make sure that user account attributes (such as group membership, department name, and other attributes) are populated correctly in Microsoft Entra ID or Exchange Online. For more information, see the following resources:
  - Attributes for Information Barriers policies
  - Add or update a user's profile information using Microsoft Entra ID
  - Configure user account properties with Office 365 PowerShell
- **Scoped directory search**: Before you define your organization's first IB policy, you must enable scoped directory search in Microsoft Teams. Wait at least 24 hours after enabling scoped directory search before you set up or define IB policies.
- **Verify audit logging is enabled**: To look up the status of an IB policy application, audit logging must be turned on. Auditing is enabled for Microsoft 365 organizations by default. Some organizations might disable auditing for specific reasons. If auditing is disabled for your organization, it might be because another administrator turned it off. Confirm that it's OK to turn auditing back on when completing this step. For more information, see Turn the audit log search on or off.
- **Check the IB mode for your organization**: Support for multiple segments, people discoverability options, Exchange ABPs, and other features is determined by the IB mode for your organization. To verify the IB mode for your organization, see Check the IB mode for your organization.
- **Remove existing Exchange Online address book policies (optional)**:

- **For organizations in *Legacy* mode**: Before you define and apply IB policies, remove all existing Exchange Online address book policies in your organization. IB policies are based on address book policies and existing ABPs aren't compatible with the ABPs created by IB. To remove your existing address book policies, see Remove an address book policy in Exchange Online. For more information about IB policies and Exchange Online, see Information Barriers and Exchange Online.
- **For organizations in *SingleSegment* or *MultiSegment* mode**: Information Barriers no longer uses Exchange Online Address Book Policies (ABPs). Enabling Information Barriers doesn't affect existing ABPs.

- **Manage using PowerShell (optional)**: You can define and manage IB segments and policies in the Microsoft Purview portal, or you can use Security & Compliance PowerShell if preferred or needed. Although several examples are provided in this article, you need to be familiar with PowerShell cmdlets and parameters if you choose to use PowerShell to configure and manage IB segments and policies. You also need the following modules depending on your configuration:

Expand table

Module	Used for	Installation
Security & Compliance PowerShell	Segments, policies, and organization settings	Connect to Security & Compliance PowerShell
Microsoft Graph PowerShell SDK	User and group management	Install the Microsoft Graph PowerShell SDK
SharePoint Online Management Shell	SharePoint and OneDrive IB configuration	Get started with SharePoint Online Management Shell

When you meet all the prerequisites, proceed to the next step.

Step 2: Segment users in your organization

In this step, you determine what IB policies you need, make a list of segments to define, and define your segments. Defining segments doesn't affect users; it just sets the stage for defining and applying IB policies.

Determine what policies you need

Consider the needs of your organization and determine the groups within your organization that need IB policies. Ask yourself the following questions:

- Are there internal, legal, or industry regulations that require the restriction of communication and collaboration between groups and users in your organization?
- Are there any groups or users you should prevent from communicating with another group of users?
- Are there any groups or users you should allow to communicate only with one or two other groups of users?

Think about the policies you need as belonging to one of two types:

- *Block* policies prevent one group from communicating with another group.
- *Allow* policies allow a group to communicate with only specific groups.

When you have your initial list of needed groups and policies, proceed to identify the segments you need for the IB policies.

Identify segments

In addition to your initial list of policies, make a list of segments for your organization. Users included in IB policies should belong to at least one segment. Users can be assigned to multiple segments if needed. You can have up to 5,000 segments in your organization, and each segment can have only one IB policy applied.

> **Important**
>
> For organizations in *Legacy* or *SingleSegment* modes, a user can belong to only one segment. To verify your IB mode, see **Check the IB mode for your organization**.

Determine which attributes in your organization's directory data you'll use to define segments. You can use *Department*, *MemberOf*, or any of the supported IB attributes. Make sure that you have values in the attribute you select for users. For more information, see the supported attributes for IB.

> **Important**
>
> **Before you proceed to the next section, make sure your directory data has values for attributes that you can use to define segments**. If your directory data doesn't have values for the attributes you want to use, update the user accounts to include that information before you proceed with configuring IB. For help, see the following resources:

- Configure user account properties with Office 365 PowerShell
- Add or update a user's profile information using Microsoft Entra ID

Enable multiple segment support for users (optional)

Support for assigning users to multiple segments is available only when your organization isn't in *Legacy* mode. To support assigning users to multiple segments, see Use multi-segment support in Information Barriers.

Users can be assigned to only one segment for organizations in *Legacy* mode. Organizations in *Legacy* mode are eligible to upgrade to the newest version of Information Barriers in the future. For more information, see the Information Barriers roadmap.

Define segments using the portals

Complete the following steps to define segments:

1. Go to the Microsoft Purview portal with credentials for an admin account in your organization.
2. Select the **Information Barriers** solution card. If you don't see the **Information Barriers** solution card, select **View all solutions** and then select **Information Barriers** from the **Risk & Compliance** section.
3. Select **Segments**.
4. On the **Segments** page, select **New segment** to create and configure a new segment.
5. On the **Name** page, enter a name for the segment. You can't rename a segment after you create it.
6. Select **Next**.
7. On the **User group filter** page, select **Add** to configure the group and user attributes for the segment. Choose an attribute for the segment from the list of available attributes.
8. For the selected attribute, select either *Equal* or *Not equal* and then enter the value for the attribute. For example, if you select *Department* as the attribute and *Equals*, you could enter *Marketing* as the defined *Department* for this segment condition. Select **Add condition** to add extra conditions for an attribute. To delete an attribute or attribute condition, select the delete icon for the attribute or condition.
9. Add extra attributes as needed on the **User group filter** page, then select **Next**.
10. On the **Review your settings** page, review the settings you chose for the segment and any suggestions or warnings for your selections. Select **Edit** to change any of the segment attributes and conditions or select **Submit** to create the segment.

Define segments using PowerShell

To define segments by using PowerShell, complete the following steps:

1. Use the **New-OrganizationSegment** cmdlet with the **UserGroupFilter** parameter that corresponds to the attribute you want to use.

   Syntax:

   ```
   PowerShell
   New-OrganizationSegment -Name "segmentname" -UserGroupFilter "attribute -eq 'attributevalue'"
   ```

   Example:

   ```
   PowerShell
   New-OrganizationSegment -Name "HR" -UserGroupFilter "Department -eq 'HR'"
   ```

   In this example, you define a segment named *HR* by using *HR*, a value in the *Department* attribute. The **-eq** operator refers to "equals." (Alternately, you can use **-ne** to mean "not equals". See PowerShell reference: "Equals" and "not equals" in segment definitions.)

   After you run each cmdlet, you see a list of details about the new segment. Details include the segment's type, who created or last modified it, and so on.

2. Repeat this process for each segment you want to define.

After you define your segments, go to Step 3: Create IB policies.

Step 3: Create IB policies

When you create your IB policies, decide whether to prevent communications between certain segments or limit communications to certain segments. Use the minimum number of IB policies to ensure your organization complies with internal, legal, and industry requirements. Use the Microsoft Purview portal or PowerShell to create and apply IB policies.

> **Tip**
>
> For user experience consistency, use *Block* policies for most scenarios.

With your list of user segments and the IB policies you want to define, select a scenario and follow the steps.

- Scenario 1: Block communications between segments
- Scenario 2: Allow a segment to communicate only with one other segment

> **Important**
>
> **Don't assign more than one policy to a segment**. For example, if you define one policy for a segment called *Sales*, don't define an additional policy for the *Sales* segment.
>
> As you define IB policies, set those policies to inactive status until you're ready to apply them. Defining (or editing) policies doesn't affect users until you set those policies to active status and then apply them.

Scenario 1: Block communications between segments

When you want to block segments from communicating with each other, define two policies: one for each direction. Each policy blocks communication in one direction only.

For example, suppose you want to block communications between Segment A and Segment B. Define two policies:

- One policy preventing Segment A from communicating with Segment B
- A second policy to prevent Segment B from communicating with Segment A

Create policies using the portals for Scenario 1

Complete the following steps to create policies:

1. Go to the Microsoft Purview portal with credentials for an admin account in your organization.
2. Select the **Information Barriers** solution card. If you don't see the **Information Barriers** solution card, select **View all solutions** and then select **Information Barriers** from the **Risk & Compliance** section.
3. Select **Policies**.
4. On **Policies**, select **Create policy** to create and configure a new IB policy.

5. On **Name**, enter a name for the policy, and then select **Next**.
6. On **Assigned segment**, select **Choose segment**. Use the search box to search for a segment by name or scroll to select the segment from the displayed list. Select **Add** to add the selected segment to the policy. You can only select one segment.
7. Select **Next**.
8. On **Communication and collaboration**, select the policy type in the **Communication and collaboration** field. The policy options are either *Allowed* or *Blocked*. In this example scenario, select *Blocked* for the first policy.

   > **Important**
   >
   > You can't change the Allowed and Blocked status for segments after creating a policy. To change the status, delete the policy and create a new one.

9. Select **Choose segment** to define the actions for the target segment. You can assign more than one segment in this step. For example, if you want to block users in a segment called *Sales* from communicating with users in a segment called *Research*, define the *Sales* segment in Step 5 and assign *Research* in the **Choose segment** option in this step.
10. Select **Next**.
11. On **Policy status**, toggle the active policy status to **On**. Select **Next** to continue.
12. On **Review your settings**, review the settings you choose for the policy and any suggestions or warnings for your selections. Select **Edit** to change any of the policy segments and status or select **Submit** to create the policy.

In this example, repeat the previous steps to create a second *Block* policy to restrict users in a segment called *Research* from communicating with users in a segment called *Sales*. Define the *Research* segment in Step 5 and assign *Sales* (or multiple segments) in the **Choose segment** option.

Create policies using PowerShell for Scenario 1

To define policies by using PowerShell, complete the following steps:

1. To define your first blocking policy, use the **New-InformationBarrierPolicy** cmdlet with the **SegmentsBlocked** parameter.

Syntax:

```
PowerShell
New-InformationBarrierPolicy -Name "policyname" -AssignedSegment "segmentAname" -
SegmentsBlocked "segmentBname"
```

Example:

```
PowerShell
New-InformationBarrierPolicy -Name "Sales-Research" -AssignedSegment "Sales" -
SegmentsBlocked "Research" -State Inactive
```

In this example, you define a policy called *Sales-Research* for a segment called *Sales*. When active and applied, this policy prevents users in *Sales* from communicating with users in a segment called *Research*.

2. To define your second blocking segment, use the **New-InformationBarrierPolicy** cmdlet with the **SegmentsBlocked** parameter again, this time with the segments reversed.

   Example:

```
PowerShell
New-InformationBarrierPolicy -Name "Research-Sales" -AssignedSegment "Research" -
SegmentsBlocked "Sales" -State Inactive
```

   In this example, you define a policy called *Research-Sales* to prevent *Research* from communicating with *Sales*.

3. Proceed to one of the following actions:

   - (If needed) Define a policy to allow a segment to communicate only with one other segment
   - (After all your policies are defined) Apply IB policies

Scenario 2: Allow a segment to communicate only with one other segment

When you want to allow a segment to communicate with only one other segment, define two policies: one for each direction. Each policy allows communication in one direction only.

In this example, define two policies:

- One policy that allows Segment A to communicate with Segment B
- A second policy that allows Segment B to communicate with Segment A

Create policies using the portals for Scenario 2

Complete the following steps to create policies:

1. Go to the Microsoft Purview portal with credentials for an admin account in your organization.

2. Select the **Information Barriers** solution card. If you don't see the **Information Barriers** solution card, select **View all solutions** and then select **Information Barriers** from the **Risk & Compliance** section.

3. On **Policies**, select **Create policy** to create and configure a new IB policy.

4. On **Name**, enter a name for the policy, and then select **Next**.

5. On **Assigned segment**, select **Choose segment**. Use the search box to search for a segment by name or scroll to select the segment from the displayed list. Select **Add** to add the selected segment to the policy. You can only select one segment.

6. Select **Next**.

7. On the **Communication and collaboration** page, select the policy type in the **Communication and collaboration** field. The policy options are either *Allowed* or *Blocked*. In this example scenario, select *Allowed* for the policy.

   > **Important**
   >
   > You can't change the Allowed and Blocked status for segments after creating a policy. To change the status, delete the policy and create a new one.

8. Select **Choose segment** to define the actions for the target segment. You can assign more than one segment in this step. For example, if you want to allow users in a segment called *Manufacturing* to communicate with users in a segment called *HR*, define the *Manufacturing* segment in Step 5 and assign *HR* in the **Choose segment** option in this step.

9. Select **Next**.

10. On **Policy status**, toggle the active policy status to **On**. Select **Next** to continue.

11. On **Review your settings**, review the settings you choose for the policy and any suggestions or warnings for your selections. Select **Edit** to change any of the policy segments and status or select **Submit** to create the policy.

12. Repeat the previous steps to create a second *Allow policy* to allow users in a segment called *Research* to communicate with users in a segment called *Sales*. Define the *Research* segment in **Step 5** and assign *Sales* (or multiple segments) in the **Choose segment** option.

Create a policy using PowerShell for Scenario 2

To define policies by using PowerShell, complete the following steps:

1. To allow one segment to communicate with the other segment, use the **New-InformationBarrierPolicy** cmdlet with the **SegmentsAllowed** parameter.

   Syntax:

   ```
   PowerShell
   New-InformationBarrierPolicy -Name "policyname" -AssignedSegment "segmentAname" -SegmentsAllowed "segmentBname","segmentAname"
   ```

   Example:

   ```
   PowerShell
   New-InformationBarrierPolicy -Name "Manufacturing-HR" -AssignedSegment "Manufacturing" -SegmentsAllowed "HR","Manufacturing" -State Inactive
   ```

   In this example, you define a policy called *Manufacturing-HR* for a segment called *Manufacturing*. When active and applied, this policy allows users in *Manufacturing* to communicate only with users in a segment called *HR*. In this case, *Manufacturing* can't communicate with users who aren't part of *HR*.

   If needed, you can specify multiple segments by using this cmdlet, as shown in the following example.

   Syntax:

   ```
   PowerShell
   ```

```
New-InformationBarrierPolicy -Name "policyname" -AssignedSegment "segmentAname" -
SegmentsAllowed "segmentBname","segmentCname","segmentDname"
```

Example:

PowerShell

```
New-InformationBarrierPolicy -Name "Research-HRManufacturing" -AssignedSegment
"Research" -SegmentsAllowed "HR","Manufacturing","Research" -State Inactive
```

In this example, you define a policy that allows the *Research* segment to communicate with only *HR* and *Manufacturing*.

Repeat this step for each policy you want to define to allow specific segments to communicate with only certain other specific segments.

2. To define your second allowing segment, use the **New-InformationBarrierPolicy** cmdlet with the **SegmentsAllowed** parameter again, this time with the segments reversed.

   Example:

   PowerShell

   ```
   New-InformationBarrierPolicy -Name "Research-Sales" -AssignedSegment "Research" -
   SegmentsAllowed "Sales" -State Inactive
   ```

   In this example, you define a policy called *Research-Sales* to allow *Research* to communicate with *Sales*.

3. Proceed to one of the following actions:

   - (If needed) Define a policy to block communications between segments
   - (After all your policies are defined) Apply IB policies

Step 4: Apply IB policies

IB policies take effect when you set them to active status and apply the policies.

Apply policies using the portals

Complete the following steps to apply policies:

1. Go to the Microsoft Purview portal with credentials for an admin account in your organization.
2. Select the **Information Barriers** solution card. If you don't see the **Information Barriers** solution card, select **View all solutions** and then select **Information Barriers** from the **Risk & Compliance** section.
3. Select **Policy application**.
4. On the **Policies application** page, select **Apply all policies** to apply all IB policies in your organization.

> **Note**
>
> Allow 30 minutes for the system to start applying the policies. The system applies policies user by user and processes about 5,000 user accounts per hour.

Apply policies using PowerShell

To apply policies by using PowerShell, complete the following steps:

1. Use the **Get-InformationBarrierPolicy** cmdlet to see a list of defined policies. Note the status and identity (GUID) of each policy.

   ```PowerShell
   Get-InformationBarrierPolicy
   ```

2. Use the **Set-InformationBarrierPolicy** cmdlet with the **Identity** parameter and the **State** parameter set to **Active** to set a policy to active status.

   Syntax:

   ```PowerShell
   Set-InformationBarrierPolicy -Identity GUID -State Active
   ```

   Example:

   ```PowerShell
   ```

```
Set-InformationBarrierPolicy -Identity 43c37853-ea10-4b90-a23d-ab8c93772471 -
State Active
```

In this example, you set an IB policy that has the GUID *43c37853-ea10-4b90-a23d-ab8c93772471* to active status.

Repeat this step for each policy.

3. When you finish setting your IB policies to active status, use the **Start-InformationBarrierPoliciesApplication** cmdlet in Security & Compliance PowerShell.

```
PowerShell
Start-InformationBarrierPoliciesApplication
```

After you run `Start-InformationBarrierPoliciesApplication`, allow 30 minutes for the system to start applying the policies. The system applies policies user by user and processes about 5,000 user accounts per hour.

View status of user accounts, segments, policies, or policy application

By using PowerShell, you can view the status of user accounts, segments, policies, and policy application.

View user account IB relationships

Use the Get-ExoInformationBarrierRelationship cmdlet with RecipientId parameters. You can use any value that uniquely identifies each user, such as name, alias, distinguished name, canonical domain name, email address, or GUID.

```
PowerShell
Get-ExoInformationBarrierRelationship -RecipientId1 meganb -RecipientId2 alexw
```

In this example, you refer to two user accounts: *meganb* for *Megan*, and *alexw* for *Alex*. This cmdlet returns information about users, such as attribute values and any IB policies that are applied.

View all segments

PowerShell

```
Get-OrganizationSegment
```

This cmdlet displays a list of all segments defined for your organization.

View all IB policies

PowerShell

```
Get-InformationBarrierPolicy
```

This cmdlet displays a list of IB policies that you defined, and their status.

View the most recent IB policy application

PowerShell

```
Get-InformationBarrierPoliciesApplicationStatus
```

This cmdlet displays information about whether policy application completed, failed, or is in progress.

View all IB policy applications

PowerShell

```
Get-InformationBarrierPoliciesApplicationStatus -All
```

This cmdlet displays information about whether policy application completed, failed, or is in progress.

What if I need to remove or change policies?

You can find resources to help you manage your IB policies.

- To edit, stop, or remove IB policies, see Manage Information Barriers policies.
- If something goes wrong with IB, see Resolve Information Barriers issues.

Step 5: Configure Information Barriers on SharePoint and OneDrive

If you configure IB for SharePoint and OneDrive, you need to enable IB on these services. You also need to enable IB on these services if you configure IB for Microsoft Teams. When you create a team in Microsoft Teams, you automatically create and associate a SharePoint site with Microsoft Teams for the files experience. By default, IB policies don't apply to this new SharePoint site and files.

To enable IB in SharePoint and OneDrive, follow the guidance and steps in the Use Information Barriers with SharePoint article.

Step 6: Information Barriers modes (optional)

Modes help strengthen access, sharing, and membership of a Microsoft 365 resource based on the resource's IB mode. Microsoft 365 Groups, Microsoft Teams, OneDrive, and SharePoint sites support modes. Your new or existing IB configuration automatically enables modes.

The following IB modes support Microsoft 365 resources:

Expand table

Mode	Description	Example
Open	The Microsoft 365 resource has no IB policies or segments associated with it. Anyone can be invited to be a member of the resource.	A team site created for a picnic event for your organization.
Owner Moderated	The resource owner's IB policy determines the IB policy of the Microsoft 365 resource. The resource owners can invite any user to the resource based on their IB policies. This mode is useful when your company wants to allow collaboration among incompatible segment users that the owner moderates. Only the resource owner can add new members per their IB policy.	The VP of HR wants to collaborate with the VPs of Sales and Research. A new SharePoint site that is set with IB mode *Owner Moderated* to add both Sales and Research segment users to the same site. It's the responsibility of the owner to ensure appropriate members are added to the resource.
Implicit	The resource members' IB policy determines the IB policy or segments of the Microsoft 365 resource. The owner can add members as long as they're compatible with the existing members of the resource. This mode is the default IB mode for Microsoft Teams.	The Sales segment user creates a Microsoft Teams team to collaborate with other compatible segments in the organization.
Explicit	The segments associated with the resource determine the IB policy of the Microsoft 365 resource. The resource owner or SharePoint	A site created only for Sales segment members to collaborate by associating the Sales segment with the site.

Mode	Description	Example
	administrator can manage the segments on the resource.	
Mixed	Only applicable to OneDrive. The segments associated with the OneDrive determine the IB policy of the OneDrive. The resource owner or OneDrive administrator can manage the segments on the resource.	A OneDrive created for Sales segment members to collaborate is allowed to be shared with unsegmented users.

Implicit mode updates

Depending on when you enabled IB in your organization, your organization is in either *Legacy*, *SingleSegment*, or *MultiSegment* organization mode. To verify your mode, see Check the IB mode for your organization.

If your organization is in *SingleSegment* or *MultiSegment* mode and the Information Barriers mode of the Teams group is *Implicit*, the Teams-connected groups and sites don't have any segments associated with them.

For more information about IB modes and how to configure them across services, see the following articles:

- Information Barriers modes and Microsoft Teams
- Information Barriers modes and OneDrive
- Information Barriers modes and SharePoint

Step 7: Configure user discoverability for Information Barriers (optional)

By using Information Barriers policies, administrators can enable or disable search restrictions in the people picker. By default, the people picker restriction is enabled for IB policies. For example, IB policies that block two specific users from communication can also restrict the users from seeing each other when using the people picker.

> **Important**
>
> Support for enabling or disabling search restrictions is only available when your organization isn't in *Legacy* mode. Organizations in *Legacy* mode can't enable or disable search restrictions. Enabling or disabling search restrictions requires additional actions to change the Information Barriers mode for your organization. For more information, see Use multi-

segment support in Information Barriers.

Organizations in *Legacy* mode are eligible to upgrade to the newest version of Information Barriers in the future. For more information, see the Information Barriers roadmap.

To disable the people picker search restriction by using PowerShell, complete the following steps:

1. Use the **Set-PolicyConfig** cmdlet to disable the people picker restriction:

```
PowerShell
Set-PolicyConfig -InformationBarrierPeopleSearchRestriction 'Disabled'
```

Example scenario: Contoso's departments, segments, and policies

To see how an organization might approach defining segments and block-based policies, consider the following example scenario.

> **Tip**
>
> For an example of allow-based policies with multi-segment support, see the North School District example in **Use multi-segment support in Information Barriers**.

Contoso's departments and plan

Contoso has five departments: *HR*, *Sales*, *Marketing*, *Research*, and *Manufacturing*. To comply with industry regulations, users in some departments can't communicate with other departments, as shown in the following table:

Expand table

Segment	Can communicate with	Can't communicate with
HR	Everyone	(no restrictions)
Sales	HR, Marketing, Manufacturing	Research
Marketing	Everyone	(no restrictions)
Research	HR, Marketing, Manufacturing	Sales

Segment	Can communicate with	Can't communicate with
Manufacturing	HR, Marketing	Anyone other than HR or Marketing

For this structure, Contoso's plan includes three IB policies:

1. An IB policy that prevents Sales from communicating with Research.
2. An IB policy that prevents Research from communicating with Sales.
3. An IB policy that allows Manufacturing to communicate with HR and Marketing only.

For this scenario, you don't need to define IB policies for *HR* or *Marketing*.

Contoso's defined segments

Contoso uses the Department attribute in Microsoft Entra ID to define segments, as follows:

HR segment:

```PowerShell
New-OrganizationSegment -Name "HR" -UserGroupFilter "Department -eq 'HR'"
```

Sales segment:

```PowerShell
New-OrganizationSegment -Name "Sales" -UserGroupFilter "Department -eq 'Sales'"
```

Marketing segment:

```PowerShell
New-OrganizationSegment -Name "Marketing" -UserGroupFilter "Department -eq
'Marketing'"
```

Research segment:

```PowerShell
New-OrganizationSegment -Name "Research" -UserGroupFilter "Department -eq 'Research'"
```

Manufacturing segment:

PowerShell

```
New-OrganizationSegment -Name "Manufacturing" -UserGroupFilter "Department -eq
'Manufacturing'"
```

After defining the segments, Contoso defines the IB policies.

Contoso's IB policies

Contoso defines three IB policies:

Policy 1: Prevent Sales from communicating with Research.

PowerShell

```
New-InformationBarrierPolicy -Name "Sales-Research" -AssignedSegment "Sales" -
SegmentsBlocked "Research" -State Inactive
```

In this example, the IB policy is called *Sales-Research*. When you activate and apply this policy, it prevents users in the Sales segment from communicating with users in the Research segment. This policy is a one-way policy; it doesn't prevent Research from communicating with Sales. For that, Policy 2 is needed.

Policy 2: Prevent Research from communicating with Sales.

PowerShell

```
New-InformationBarrierPolicy -Name "Research-Sales" -AssignedSegment "Research" -
SegmentsBlocked "Sales" -State Inactive
```

In this example, the IB policy is called *Research-Sales*. When you activate and apply this policy, it prevents users in the Research segment from communicating with users in the Sales segment.

Policy 3: Allow Manufacturing to communicate with HR and Marketing only.

PowerShell

```
New-InformationBarrierPolicy -Name "Manufacturing-HRMarketing" -AssignedSegment
"Manufacturing" -SegmentsAllowed "HR","Marketing","Manufacturing" -State Inactive
```

In this case, the IB policy is called *Manufacturing-HRMarketing*. When you activate and apply this policy, Manufacturing can communicate only with HR and Marketing. HR and Marketing aren't restricted from communicating with other segments.

After defining the segments and policies, Contoso applies the policies by running the **Start-InformationBarrierPoliciesApplication** cmdlet.

When the cmdlet finishes, Contoso complies with industry requirements.

PowerShell reference: "Equals" and "not equals" in segment definitions

When you define segments by using PowerShell, use "equals" (`-eq`) and "not equals" (`-ne`) parameters. Also, use logical *AND* and *OR* operators for complex group filters.

Basic "equals" example:

```
PowerShell
New-OrganizationSegment -Name "HR" -UserGroupFilter "Department -eq 'HR'"
```

Defines a segment named *HR* by using *HR*, a value in the *Department* attribute. The **-eq** portion of the cmdlet refers to "equals".

Basic "not equals" example:

```
PowerShell
New-OrganizationSegment -Name "NotSales" -UserGroupFilter "Department -ne 'Sales'"
```

Defines a segment called *NotSales* that includes everyone who isn't in *Sales*. The **-ne** portion of the cmdlet refers to "not equals".

Using "AND" operators:

```
PowerShell
New-OrganizationSegment -Name "LocalFTE" -UserGroupFilter "Location -eq 'Local'" -and "Position -ne 'Temporary'"
```

Defines a segment called *LocalFTE* that includes users who are located locally and whose positions aren't listed as *Temporary*.

Using "AND" with MemberOf:

```
PowerShell
```

```
New-OrganizationSegment -Name "Segment1" -UserGroupFilter "MemberOf -eq
'group1@contoso.com'' -and MemberOf -ne 'group3@contoso.com'"
```

Defines a segment called *Segment1* that includes users who are members of group1@contoso.com and not members of group3@contoso.com.

Using "OR" with MemberOf:

PowerShell

```
New-OrganizationSegment -Name "Segment2" -UserGroupFilter "MemberOf -eq
'group2@contoso.com' -or MemberOf -ne 'group3@contoso.com'"
```

Defines a segment called *Segment2* that includes users who are members of group2@contoso.com and not members of group3@contoso.com.

Combining "OR" and "AND" operators:

PowerShell

```
New-OrganizationSegment -Name "Segment1and2" -UserGroupFilter "(MemberOf -eq
'group1@contoso.com' -or MemberOf -eq 'group2@contoso.com') -and MemberOf -ne
'group3@contoso.com'"
```

Defines a segment called *Segment1and2* that includes users in group1@contoso.com and group2@contoso.com and not members of group3@contoso.com.

> **Tip**
>
> If possible, use segment definitions that include `-eq` or `-ne`. Avoid defining complex segment definitions.

PowerShell cmdlet quick reference

The following table lists the PowerShell cmdlets you can use to manage Information Barriers segments and policies.

Expand table

Cmdlet	Module	Purpose
New-OrganizationSegment	Security & Compliance PowerShell	Create a new segment
Get-OrganizationSegment	Security & Compliance PowerShell	View all defined segments
Set-OrganizationSegment	Security & Compliance PowerShell	Edit a segment definition
Remove-OrganizationSegment	Security & Compliance PowerShell	Remove a segment
New-InformationBarrierPolicy	Security & Compliance PowerShell	Create a new IB policy
Get-InformationBarrierPolicy	Security & Compliance PowerShell	View IB policies and their status
Set-InformationBarrierPolicy	Security & Compliance PowerShell	Edit or activate a policy
Remove-InformationBarrierPolicy	Security & Compliance PowerShell	Remove a policy
Start-InformationBarrierPoliciesApplication	Security & Compliance PowerShell	Apply all active IB policies
Stop-InformationBarrierPoliciesApplication	Security & Compliance PowerShell	Stop applying IB policies
Get-InformationBarrierPoliciesApplicationStatus	Security & Compliance PowerShell	View policy application status
Get-ExoInformationBarrierRelationship	Security & Compliance PowerShell	View IB relationship between users
Get-PolicyConfig	Security & Compliance PowerShell	View IB organization mode
Set-PolicyConfig	Security & Compliance PowerShell	Configure IB organization settings

For managing Information Barriers in SharePoint and OneDrive, see the SharePoint Online Management Shell cmdlets documented in Use Information Barriers with SharePoint and Use Information Barriers with OneDrive.

Next steps

- Manage Information Barriers policies: Edit or remove segments and policies after initial configuration.
- Use Information Barriers with SharePoint: Configure IB modes and segment associations for SharePoint sites.
- Use Information Barriers with OneDrive: Configure IB modes and segment associations for OneDrive accounts.
- Information Barriers in Microsoft Teams: Understand how IB policies affect Teams chats, meetings, and channels.

ⓘ **Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 04/22/2026

Information Barriers attributes

You can use certain attributes in Microsoft Entra ID to segment users in Information Barriers (IB). After you define segments, use them as filters for IB policies. For example, you might use **Department** to define segments of users by department within your organization (assuming no single user works for two departments at the same time).

This article describes how to use attributes with Information Barriers and provides a list of attributes you can use. For more information about Information Barriers, see the following resources:

- Information Barriers
- Define policies for Information Barriers in Microsoft Teams
- Edit (or remove) IB policies

How to use attributes in IB policies

Use the attributes listed in this article to define or edit segments of users. Your defined segments serve as parameters (called *UserGroupFilter* values) in IB policies.

1. Decide which attribute you want to use to define segments. (See the Reference section in this article.)

2. Make sure the user accounts have values filled in for the attribute(s) you selected in Step 1. View user account details, and if necessary, edit user accounts to include attribute values.

    - To edit multiple accounts or use PowerShell to edit a single account, see Configure user account properties with Office 365 PowerShell.
    - To edit a single account, see Add or update a user's profile information using Microsoft Entra ID.

3. Define segments using PowerShell, similar to the following examples. Before running these cmdlets, connect to Security & Compliance PowerShell.

    > **Tip**

You can also define segments using the Microsoft Purview portal. For more information, see Define segments using the portals.

Define a segment using the Department attribute:

```
PowerShell
New-OrganizationSegment -Name "Segment1" -UserGroupFilter "Department -eq 'Department1'"
```

Define a segment using the MemberOf attribute (suppose this attribute contains group names, such as "BlueGroup"):

```
PowerShell
New-OrganizationSegment -Name "SegmentA" -UserGroupFilter "MemberOf -eq 'BlueGroup'"
```

Define a segment using ExtensionAttribute1 (suppose this attribute contains job titles, such as "DayTrader"):

```
PowerShell
New-OrganizationSegment -Name "DayTraders" -UserGroupFilter "ExtensionAttribute1 -eq 'DayTrader'"
```

Tip

When you define segments, use the same attribute for all your segments. For example, if you define some segments using *Department*, define all of the segments using *Department*. Don't define some segments using *Department* and others using *MemberOf*. Make sure your segments don't overlap; each user belongs to exactly one segment.

Reference

The following table lists the attributes that you can use with Information Barriers.

Tip

The most commonly used attributes for defining segments are **Department**, **MemberOf**, **Company**, and **ExtensionAttribute1–15**. For more information about defining segments, see Get started with Information Barriers.

Expand table

Microsoft Entra property name (LDAP display name)	Exchange property name
Co	Co
Company	Company
Department	Department
Description	Description
ExtensionAttribute1	CustomAttribute1
ExtensionAttribute2	CustomAttribute2
ExtensionAttribute3	CustomAttribute3
ExtensionAttribute4	CustomAttribute4
ExtensionAttribute5	CustomAttribute5
ExtensionAttribute6	CustomAttribute6
ExtensionAttribute7	CustomAttribute7
ExtensionAttribute8	CustomAttribute8
ExtensionAttribute9	CustomAttribute9
ExtensionAttribute10	CustomAttribute10
ExtensionAttribute11	CustomAttribute11
ExtensionAttribute12	CustomAttribute12
ExtensionAttribute13	CustomAttribute13
ExtensionAttribute14	CustomAttribute14
ExtensionAttribute15	CustomAttribute15
Mail	WindowsEmailAddress
MailNickname	Alias

Microsoft Entra property name (LDAP display name)	Exchange property name
MemberOf	MemberOfGroup
MSExchExtensionCustomAttribute1	ExtensionCustomAttribute1
MSExchExtensionCustomAttribute2	ExtensionCustomAttribute2
MSExchExtensionCustomAttribute3	ExtensionCustomAttribute3
MSExchExtensionCustomAttribute4	ExtensionCustomAttribute4
MSExchExtensionCustomAttribute5	ExtensionCustomAttribute5
PhysicalDeliveryOfficeName	Office
PostalCode	PostalCode
ProxyAddresses	EmailAddresses
StreetAddress	StreetAddress
TargetAddress	ExternalEmailAddress
UsageLocation	UsageLocation
UserPrincipalName	UserPrincipalName

Next steps

- Get started with Information Barriers: Use these attributes to define segments and create IB policies.
- Manage Information Barriers policies: Edit segments if you need to change the attributes used to define them.
- Use multi-segment support in Information Barriers: Assign users to multiple segments for complex organizational scenarios.

ⓘ **Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 04/14/2026

Use multi-segment support in Information Barriers

> **Important**
>
> Support for 5,000 segments and assigning users to multiple segments is only available when your organization isn't in *Legacy* mode. Assigning users to multiple segments requires extra steps to change the Information Barriers mode for your organization. For more information, see Use multi-segment support in Information Barriers.
>
> For organizations in *Legacy* mode, the maximum number of segments supported is 250, and users are restricted to being assigned to only one segment. Organizations in *Legacy* mode are eligible to upgrade to the newest version of Information Barriers in the future. For more information, see the Information Barriers roadmap.

By using the multi-segment mode, you can assign users in your organization to up to 10 segments in Information Barriers instead of being limited to just one segment. This support for more diverse communication rules between individuals and groups supports more complex organizational and operational scenarios. For organizations using multi-segment support, define all Information Barriers policies with an allow list.

When you configure multi-segment support, user compatibility depends on each user's assignment to a shared segment. If users share assignment to the same segment, they're compatible. For example, the following table shows that User A and User B aren't compatible because they don't share an assigned segment. However, User A is compatible with User C and User B is compatible with User C because they each have an assigned segment in common.

Expand table

User	Assigned segments
User A	Segment 1, Segment 2
User B	Segment 3, Segment 4
User C	Segment 2, Segment 4

Multi-segment example: North School District's schools, segments, and policies

The North School District has two schools, School 1 and School 2. The district policy is to allow students and teachers to communicate with each other only if they're both in the same school. For example, a student and teacher that are both in School 1 can communicate, but a student in School 1 can't communicate with a teacher in School 2. For this scenario, configure multiple segments to support the following district policy scenarios:

North School District's schools and plan

North School District has two schools:

Expand table

Segment	Allowed communication	Prevented communication
School 1	Students and teachers in School 1	Students and teachers in School 2
School 2	Students and teachers in School 2	Students and teachers in School 1

For this structure, North School District's plan includes three IB policies:

1. An IB policy that enables students and teachers in School 1 to communicate with each other.
2. An IB policy that enables students and teachers in School 2 to communicate with each other.
3. An IB policy that allows teachers in School 1 and School 2 to communicate with each other.

North School District's defined segments

North School District uses the *Department* attribute in Microsoft Entra ID to define segments, as follows:

School1 segment:

PowerShell

```
New-OrganizationSegment -Name "School1" -UserGroupFilter "Department -eq 'School1'"
```

School2 segment:

PowerShell

```
New-OrganizationSegment -Name "School2" -UserGroupFilter "Department -eq 'School2'"
```

AllTeachers segment:

```
PowerShell
New-OrganizationSegment -Name "AllTeachers" -UserGroupFilter "MemberOfGroup -eq
'AllTeachersgroup@northschoolsdistrict.com'"
```

After defining the segments, North School District defines the IB policies.

North School District's IB policies

North School District defines three IB policies:

Policy 1: Students and teachers in School 1 can communicate with each other.

```
PowerShell
New-InformationBarrierPolicy -Name School1Policy -SegmentsAllowed 'School1' -
AssignedSegment 'School1' -State Active
```

When you activate and apply this policy, students and teachers in School 1 can communicate with each other. This policy is a one-way policy; it doesn't prevent students and teachers in School 1 from communicating with School 2. For that, Policy 2 is needed.

Policy 2: Students and teachers in School 2 can communicate with each other.

```
PowerShell
New-InformationBarrierPolicy -Name School2Policy -SegmentsAllowed 'School2' -
AssignedSegment 'School2' -State Active
```

When you activate and apply this policy, students and teachers in School 2 can communicate with each other.

Policy 3: Teachers in different schools can communicate with each other.

```
PowerShell
New-InformationBarrierPolicy -Name AllTeachersPolicy -SegmentsAllowed 'AllTeachers'
-AssignedSegment 'AllTeachers' -State Active
```

When you activate and apply this policy, teachers in School 1 and School 2 can communicate with each other.

After defining segments and policies, North School District runs the **Start-InformationBarrierPoliciesApplication** cmdlet to apply the policies. When the cmdlet finishes, North School District implements its communication policy for students and teachers.

Check the IB mode for your organization

To support assigning users to multiple segments, verify that your IB organization supports multiple segments. Run the following cmdlet to verify your IB mode in the Security & Compliance PowerShell:

PowerShell

```
Get-PolicyConfig
```

If the value of the `InformationBarrierMode` property is *SingleSegment*, enable multi-segment support by following the guidance in the Enable multiple segment support for users section in this article. If the value of the `InformationBarrierMode` property is *MultiSegment*, you can skip enabling support for multi-segment because it's already enabled for your organization.

If the value of the `InformationBarrierMode` property is *Legacy*, enabling multi-segment isn't supported for your organization. *Legacy* organizations are eligible to upgrade to the newest version of Information Barriers in the future. For more information, see the Information Barriers roadmap.

Enable multiple segment support for users

To enable multiple segment support for organizations in *SingleSegment* mode, make sure that you don't currently have any IB segments or policies defined for your organization. Run the following cmdlet to enable multiple segment support in your organization:

PowerShell

```
Set-PolicyConfig -InformationBarrierMode 'MultiSegment'
```

> **Important**
>
> If you enable multiple segments and configure IB in your organization, don't revert to single segment support.

Multi-segment support for users in OneDrive

When you configure OneDrive for Information Barriers to support multiple segments, the following behavior applies:

- A multi-segment user's OneDrive is automatically set to *Owner Moderated* mode.
- In *Explicit* or *Mixed* mode, a multi-segment user can access a OneDrive if they belong to at least one of its segments and have site access permission. All other modes use the same access experience as single segment support.
- Sharing behavior follows the IB mode set for OneDrive. In *Explicit* mode, users can share with others who have the same segment. In *Open* or *Owner Moderated* mode, users can share with compatible users per IB policies.

For more information, see Use Information Barriers with OneDrive.

Multi-segment support for users in SharePoint Online

When you configure SharePoint for Information Barriers to support multiple segments, the following behavior applies:

- A user with multiple segments who creates a SharePoint site (either a Microsoft 365 group connected site or a non-group site) automatically uses *Owner Moderated* mode.
- In *Explicit* mode, users get access if they have at least one segment that matches the site's segment and they have site access permission. All other modes use the same access experience as single segment support.
- Sharing behavior follows the IB mode of the site. In *Explicit* mode, users can share with others who match the site's segment. In *Implicit* or *Owner Moderated* mode, users can share with existing members of the connected Microsoft 365 group. In *Open* mode, users can share with compatible users per IB policy.

For more information, see Use Information Barriers with SharePoint.

Multi-segment support for users in Microsoft Teams

When you configure Teams for Information Barriers to support multiple segments, the following behavior applies:

- A user with multiple segments who creates a team automatically uses *Implicit* mode.
- All users in the team must have one segment that is compatible with all other users.

For more information, see Use Information Barriers with Microsoft Teams.

Resources

- Learn about Information Barriers
- Get started with Information Barriers
- Manage Information Barriers policies
- Information Barriers in Microsoft Teams
- Information Barriers in SharePoint
- Information Barriers in OneDrive

Last updated on 03/30/2026

Manage Information Barriers policies

After you define Information Barriers (IB) policies, you might need to change those policies or your user segments as part of troubleshooting or regular maintenance.

Manage segments and policies

- **Edit user account attributes**: Fill in attributes in Microsoft Entra ID that you can use to define segments. Edit user account attributes when users aren't included in segments they should be, to change which segments users are in, or to define segments by using different attributes.
- **Edit a segment**: Edit segments when you want to change how you define a segment. For example, you might have originally defined segments by using *Department* and now want to use another attribute, such as *MemberOf*.
- **Edit a policy**: Edit an Information Barriers policy when you want to change how a policy works. For example, instead of blocking communications between two segments, you might decide you want to allow communications to occur only between certain segments.
- **Set a policy to inactive status**: Set a policy to inactive status when you want to make changes to a policy, or when you don't want a policy to be in effect.
- **Remove a policy**: Remove an Information Barriers policy when you no longer need a particular policy in place.
- **Remove a segment**: Remove an Information Barriers segment when you no longer need a particular segment.
- **Remove a policy and a segment**: Remove an Information Barriers policy and a segment at the same time.
- **Stop a policy application**: Stop the process of applying Information Barriers policies. Stopping a policy application isn't instant, and it doesn't undo policies that are already applied to users.
- **Enable or disable user discoverability**: Enable or disable if users are displayed in the people picker.
- **Define policies for Information Barriers**: Define an Information Barriers policy when you don't already have such policies in place, and you must restrict or limit communications between specific groups of users.
- **Resolve Information Barriers issues**: Refer to this article when you run into unexpected issues with Information Barriers.

> **ⓘ Important**
>
> To perform the tasks described in this article, you must be assigned an appropriate role, such as one of the following roles:
>
> - Microsoft 365 Enterprise Global Administrator
> - Global Administrator
> - Compliance Administrator
> - IB Compliance Management
>
> For more information about prerequisites for Information Barriers, see Prerequisites (for Information Barriers policies).
>
> Make sure to connect to the Security & Compliance PowerShell.

> **ⓘ Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

Edit user account attributes

Complete the following steps to edit attributes that segment users. For example, if you're using a Department attribute and one or more user accounts don't currently have any values listed for Department, edit those user accounts to include Department information. Use user account attributes to define segments so you can assign Information Barriers policies.

1. Use the **Get-ExoInformationBarrierRelationship** cmdlet to view details for a specific user account, such as attribute values and assigned segments. Use Identity parameters in Exchange Online PowerShell. You can use any value that uniquely identifies each user, such as name, alias, distinguished name, canonical domain name, email address, or GUID.

   Syntax:

   PowerShell

   ```
   Get-ExoInformationBarrierRelationship -RecipientId1 <value> -RecipientId2 <value>
   ```

Example:

```
PowerShell
Get-ExoInformationBarrierRelationship -Identity meganb -Identity2 alexw
```

In this example, two user accounts are referenced: *meganb* for *Megan*, and *alexw* for *Alex*.

2. Decide which attribute to edit for your user account profiles. For more information, see Attributes for Information Barriers policies.

3. Edit one or more user accounts to include values for the attribute you selected in the previous step. Use one of the following procedures:

   - To edit a single account, see Add or update a user's profile information using Microsoft Entra ID.
   - To edit multiple accounts (or use PowerShell to edit a single account), see Configure user account properties with Office 365 PowerShell.

Edit a segment

Complete the following steps to edit the definition of a user segment. For example, you might change the name of a segment or the filter that determines who is included in the segment.

1. Use the **Get-OrganizationSegment** cmdlet to view all existing segments.

```
PowerShell
Get-OrganizationSegment
```

You see a list of segments and details for each segment, such as the segment type, its UserGroupFilter value, who created or last modified it, GUID, and more.

> **Tip**
>
> Print or save your list of segments for reference. For example, if you want to edit a segment, you need to know its name or identifying value (used with the Identity parameter).

2. Use the **Set-OrganizationSegment** cmdlet with the **Identity** parameter and relevant details to edit a segment.

   Syntax:

   ```
   PowerShell
   Set-OrganizationSegment -Identity GUID -UserGroupFilter "attribute -eq 'attributevalue'"
   ```

   Example:

   ```
   PowerShell
   Set-OrganizationSegment -Identity c96e0837-c232-4a8a-841e-ef45787d8fcd -UserGroupFilter "Department -eq 'HRDept'"
   ```

   In this example, you update the department name to *HRDept* for the segment with GUID *c96e0837-c232-4a8a-841e-ef45787d8fcd*.

3. When you finish editing segments for your organization, you can either define or edit Information Barriers policies.

Edit a policy

1. Use the **Get-InformationBarrierPolicy** cmdlet to view a list of current Information Barriers policies.

   ```
   PowerShell
   Get-InformationBarrierPolicy
   ```

   In the list of results, find the policy that you want to change. Note the policy's GUID and name.

2. Use the **Set-InformationBarrierPolicy** cmdlet with an **Identity** parameter, and specify the changes you want to make.

   Example: Suppose you defined a policy to block the *Research* segment from communicating with the *Sales* and *Marketing* segments:

   ```
   PowerShell
   ```

```
New-InformationBarrierPolicy -Name "Research-SalesMarketing" -AssignedSegment
"Research" -SegmentsBlocked "Sales","Marketing"
```

Suppose you want to change it so that users in the *Research* segment can only communicate with users in the *HR* segment. To make this change:

PowerShell

```
Set-InformationBarrierPolicy -Identity 43c37853-ea10-4b90-a23d-ab8c93772471 -
SegmentsAllowed "HR"
```

In this example, you change *SegmentsBlocked* to *SegmentsAllowed* and specify the *HR* segment.

3. When you finish editing a policy, make sure to apply your changes. For more information, see Apply Information Barriers policies.

Set a policy to inactive status

1. Use the **Get-InformationBarrierPolicy** cmdlet to view a list of current Information Barriers policies.

PowerShell

```
Get-InformationBarrierPolicy
```

In the list of results, find the policy that you want to change or remove. Note the policy's GUID and name.

2. Use the **Set-InformationBarrierPolicy** cmdlet with the *Identity* parameter and the *State* parameter set to *Inactive* to set the policy's status to inactive.

Syntax:

PowerShell

```
Set-InformationBarrierPolicy -Identity GUID -State Inactive
```

Example:

PowerShell

```
Set-InformationBarrierPolicy -Identity 43c37853-ea10-4b90-a23d-ab8c9377247 -
State Inactive
```

In this example, the Information Barriers policy that has GUID *43c37853-ea10-4b90-a23d-ab8c9377247* is set to an inactive status.

3. Use the **Start-InformationBarrierPoliciesApplication** cmdlet to apply your changes.

PowerShell

```
Start-InformationBarrierPoliciesApplication
```

Changes are applied user-by-user for your organization. If your organization is large, it can take 24 hours or more for this process to complete. As a general guideline, it takes about an hour to process 5,000 user accounts.

4. One or more Information Barriers policies are set to inactive status. From here, you can take any of the following actions:

   - Keep it as is (a policy set to inactive status has no effect on users)
   - Edit a policy
   - Remove a policy

Remove a policy

1. Use the **Get-InformationBarrierPolicy** cmdlet to view a list of current Information Barriers policies.

PowerShell

```
Get-InformationBarrierPolicy
```

In the list of results, find the policy that you want to remove. Note the policy's GUID and name.

2. Make sure the policy is set to inactive status. To set the policy's status to inactive, use the **Set-InformationBarrierPolicy** cmdlet with an Identity parameter and the State parameter set to Inactive.

Syntax:

```
```

PowerShell

```
Set-InformationBarrierPolicy -Identity GUID -State Inactive
```

Example:

PowerShell

```
Set-InformationBarrierPolicy -Identity 43c37853-ea10-4b90-a23d-ab8c9377247 -
State Inactive
```

In this example, you set an Information Barriers policy that has GUID *43c37853-ea10-4b90-a23d-ab8c9377247* to an inactive status.

3. To apply your changes on the policy, use the **Start-InformationBarrierPoliciesApplication** cmdlet.

PowerShell

```
Start-InformationBarrierPoliciesApplication
```

Changes are applied user-by-user for your organization. If your organization is large, it can take 24 hours or more for this process to complete. As a general guideline, it takes about an hour to process 5,000 user accounts.

4. Use the **Remove-InformationBarrierPolicy** cmdlet with an Identity parameter.

Syntax:

PowerShell

```
Remove-InformationBarrierPolicy -Identity GUID
```

Example:

PowerShell

```
Remove-InformationBarrierPolicy -Identity 43c37853-ea10-4b90-a23d-ab8c93772471
```

In this example, you remove the policy that has GUID *43c37853-ea10-4b90-a23d-ab8c93772471*.

When prompted, confirm the change.

Remove a segment

1. Use the **Get-OrganizationSegment** cmdlet to view all existing segments.

   ```
   PowerShell
   Get-OrganizationSegment
   ```

   You see a list of segments and details for each segment, such as the segment type, its UserGroupFilter value, who created or last modified it, GUID, and more.

   > **Tip**
   >
   > Print or save your list of segments for reference later. For example, if you want to edit a segment, you need to know its name or identify value (this value works with the Identity parameter).

2. Identify the segment to remove and make sure you remove the IB policy associated with the segment. For more information, see the Remove a policy procedure.

3. Edit the segment to remove the relationship of users to that segment. This action updates the segment definition and removes all users from the segment. Use the UserGroupFilter parameter to disassociate users from the segment before removal.

   Use the **Set-OrganizationSegment** cmdlet with the *Identity* parameter and relevant details to edit a segment.

   Syntax:

   ```
   PowerShell
   Set-OrganizationSegment -Identity GUID -UserGroupFilter "attribute -eq
   'attributevalue'"
   ```

   Example:

   ```
   PowerShell
   Set-OrganizationSegment -Identity c96e0837-c232-4a8a-841e-ef45787d8fcd -
   UserGroupFilter "Department -eq 'FakeDept'"
   ```

In this example, for the segment that has the GUID c96e0837-c232-4a8a-841e-ef45787d8fcd, you define the department name as *FakeDept* to remove users from the segment. This example uses the *Department* attribute, but you can use other attributes as appropriate. The example uses *FakeDept* because this doesn't exist and is certain to not contain any users.

4. Use the **Start-InformationBarrierPoliciesApplication** cmdlet to apply your changes.

```
PowerShell
Start-InformationBarrierPoliciesApplication -CleanupGroupSegmentLink
```

> **Note**
>
> The *CleanupGroupSegmentLink* attribute removes group associations with the segment with no user associations.

Changes are applied user-by-user for your organization. If your organization is large, it can take 24 hours or more for this process to complete. As a general guideline, it takes about an hour to process 5,000 user accounts.

5. Use the **Remove-OrganizationSegment** cmdlet with the *Identity* parameter and relevant details to remove a segment.

Syntax:

```
PowerShell
Remove-OrganizationSegment -Identity GUID
```

Example:

```
PowerShell
Remove-OrganizationSegment -Identity c96e0837-c232-4a8a-841e-ef45787d8fcd
```

In this example, you remove the segment that has the GUID c96e0837-c232-4a8a-841e-ef45787d8fcd.

Remove a policy and segment

1. Use the **Get-InformationBarrierPolicy** cmdlet to view a list of current Information Barriers policies.

   ```PowerShell
   Get-InformationBarrierPolicy
   ```

   In the list of results, find the policy that you want to remove. Note the policy's GUID and name.

2. Use the **Get-OrganizationSegment** cmdlet to view all existing segments.

   ```PowerShell
   Get-OrganizationSegment
   ```

   You see a list of segments and details for each, such as segment type, its *UserGroupFilter* parameter value, who created or last modified it, GUID, and so on.

   > **Tip**
   >
   > Print or save your list of segments for reference later. For example, if you want to edit a segment, you need to know its name or identify value (this value works with the Identity parameter).

3. Set the policy to inactive status using the **Set-InformationBarrierPolicy** cmdlet.

   Syntax:

   ```PowerShell
   Set-InformationBarrierPolicy -Identity GUID -State Inactive
   ```

   Example:

   ```PowerShell
   Set-InformationBarrierPolicy -Identity 43c37853-ea10-4b90-a23d-ab8c93772471 -
   State Inactive
   ```

   In this example, this command sets an Information Barriers policy that has GUID 43c37853-ea10-4b90-a23d-ab8c93772471 to an inactive status.

4. Edit the segment that you want to remove to remove the relationship of users to that segment. This action updates the segment definition and removes all users from the segment. Use the *UserGroupFilter* parameter to disassociate users from the segment before removal.

   Syntax:

   ```PowerShell
   Set-OrganizationSegment -Identity GUID -UserGroupFilter "attribute -eq 'attributevalue'"
   ```

   Example:

   ```PowerShell
   Set-OrganizationSegment -Identity c96e0837-c232-4a8a-841e-ef45787d8fcd -UserGroupFilter "Department -eq 'FakeDept'"
   ```

   In this example, for the segment that has the GUID c96e0837-c232-4a8a-841e-ef45787d8fcd, the department name is updated to *FakeDept* to remove users from the segment. This example uses the *Department* attribute, but you can use other attributes as appropriate. The example uses *FakeDept* because this doesn't exist and is certain to contain no users.

5. Use the **Start-InformationBarrierPoliciesApplication** cmdlet to apply your changes.

   ```PowerShell
   Start-InformationBarrierPoliciesApplication -CleanupGroupSegmentLink
   ```

   > ⓘ **Note**
   >
   > The *CleanupGroupSegmentLink* attribute removes group associations with the segment with no user associations.

   Changes are applied user-by-user for your organization. If your organization is large, it can take 24 hours or more for this process to complete. As a general guideline, it takes about an hour to process 5,000 user accounts.

6. Remove the policy using the **Remove-InformationBarrierPolicy** cmdlet.

Syntax:

```
PowerShell
Remove-InformationBarrierPolicy -Identity GUID
```

Example:

```
PowerShell
Remove-InformationBarrierPolicy -Identity 43c37853-ea10-4b90-a23d-ab8c93772471
```

In this example, the policy that has GUID *43c37853-ea10-4b90-a23d-ab8c93772471* is removed.

When prompted, confirm the change.

7. Remove the segment using the **Remove-OrganizationSegment** cmdlet.

   Syntax:

   ```
   PowerShell
   Remove-OrganizationSegment -Identity GUID
   ```

   Example:

   ```
   PowerShell
   Remove-OrganizationSegment -Identity c96e0837-c232-4a8a-841e-ef45787d8fcd
   ```

   In this example, the segment with GUID c96e0837-c232-4a8a-841e-ef45787d8fcd was removed.

Stop applying a policy

If you started applying Information Barriers policies and want to stop applying those policies, complete the following steps. It takes about 30 to 35 minutes for the process to start.

1. Use the **Get-InformationBarrierPoliciesApplicationStatus** cmdlet to view the status of the most recent Information Barriers policy application.

   ```
   PowerShell
   ```

```
Get-InformationBarrierPoliciesApplicationStatus
```

Note the application's GUID.

2. Use the **Stop-InformationBarrierPoliciesApplication** cmdlet with an Identity parameter.

   Syntax:

   ```
   PowerShell
   Stop-InformationBarrierPoliciesApplication -Identity GUID
   ```

   Example:

   ```
   PowerShell
   Stop-InformationBarrierPoliciesApplication -Identity 46237888-12ca-42e3-a541-3fcb7b5231d1
   ```

   In this example, you're stopping the application of Information Barriers policies.

Enable or disable user discoverability

> **Important**
>
> Support for enabling or disabling search restrictions is only available when your organization isn't in *Legacy* mode. Organizations in *Legacy* mode can't enable or disable search restrictions. Enabling or disabling search restrictions requires additional actions to change the Information Barriers mode for your organization. For more information, see Use multi-segment support in Information Barriers.
>
> Organizations in *Legacy* mode are eligible to upgrade to the newest version of Information Barriers in the future. For more information, see the Information Barriers roadmap.

To enable the people picker search restriction by using PowerShell, complete the following steps:

1. Use the **Set-PolicyConfig** cmdlet to enable the people picker restriction:

PowerShell

```
Set-PolicyConfig -InformationBarrierPeopleSearchRestriction 'Enabled'
```

To disable the people picker search restriction by using PowerShell, complete the following steps:

1. Use the **Set-PolicyConfig** cmdlet to disable the people picker restriction:

PowerShell

```
Set-PolicyConfig -InformationBarrierPeopleSearchRestriction 'Disabled'
```

Next steps

- Get started with Information Barriers: Re-apply policies after editing segments or policies.
- Create an Information Barriers policy compliance report: Identify SharePoint and OneDrive sites that don't comply after policy changes.
- Information Barriers attributes: Review available attributes when redefining segment filters.

ⓘ **Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 04/14/2026

Change Information Barriers modes with a PowerShell script

Applies to: Microsoft Teams

Use this PowerShell script to update the Information Barriers (IB) mode for all Teams-connected groups in your tenant. You need to update the mode for these groups after you deploy Information Barriers. Groups created before you enable IB are assigned the *Open* mode. In *Open* mode, there aren't any applicable IB policies. After you enable IB, *Implicit* becomes the default mode for any new groups you create. However, existing groups still keep *Open* mode configuration. Run this script to change these existing groups to *Implicit* mode.

In this script, you use the Get-UnifiedGroup cmdlet, which is in the Exchange Online PowerShell module to update the mode. For more information about managing Teams using PowerShell, see Teams PowerShell overview.

Prerequisites

Before you run this script, make sure you have the following prerequisites:

- The Exchange Online PowerShell module installed.
- A work or school account that has the global administrator role for your tenant.

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

Run the script

Complete the following steps to update the IB mode for all existing Teams-connected groups:

1. Connect to Exchange Online PowerShell:

```
PowerShell
```

```
Connect-ExchangeOnline
```

2. Copy and run the following script. The script identifies all Teams-connected groups in *Open* mode and updates them to *Implicit* mode.

3. After the script finishes, review the output for the number of groups successfully updated. If any groups fail to update, check the `BackfillFailedTeams.csv` file generated in your current directory for details and retry the failed groups.

Sample script

PowerShell

```
<#
.SYNOPSIS
This script updates the IB mode for all Teams-connected groups in your tenant at
the same time.
.DESCRIPTION
Use this script to update the info barrier mode from open to implicit across the
groups in your tenant.
#>

$teams = Get-UnifiedGroup -Filter {ResourceProvisioningOptions -eq "Team"} -
ResultSize Unlimited

Write-Output ([string]::Format("Number of Teams = {0}", @($teams).Length))

$teamsToUpdate = New-Object System.Collections.ArrayList

foreach($team in $teams)
{
  if ($team.InformationBarrierMode -eq "Open")
  {
    $teamsToUpdate.Add($team.ExternalDirectoryObjectId) | out-null
  }
}

Write-Output ([string]::Format("Number of Teams to be backfilled = {0}",
@($teamsToUpdate).Length))

$outfile = "BackfillFailedTeams.csv"

if (!(Test-Path "$outfile"))
{
  $newcsv = {} | Select "ExternalDirectoryObjectId", "ExceptionDetails" | Export-
Csv $outfile -NoTypeInformation
}
else
{
  $dateTime = Get-Date
```

```
  $newEntry = "{0},{1}" -f "New session started", $dateTime
  $newEntry | add-content $outfile
}

$SuccessfullyBackfilledGroup = 0

for($i = 0; $i -lt @($teamsToUpdate).Length; $i++)
{
  Invoke-Command { Set-UnifiedGroup $teamsToUpdate[$i] -InformationBarrierMode
"Implicit" } -ErrorVariable ErrorOutput

  if ($ErrorOutput)
  {
    # saving the errors in a csv file
    $errorBody = $ErrorOutput[0].ToString() -replace "`n"," " -replace "`r"," " -
replace ",", " "
    $newEntry = "{0},{1}" -f $teamsToUpdate[$i].ToString(), '"' + $errorBody + '"'
    $newEntry | add-content $outfile
  }
  else
  {
    $SuccessfullyBackfilledGroup++
  }

  if (($i+1) % 100 -eq 0)
  {
    # print the number of teams backfilled after the batch of 100 updates
    Write-Output ([string]::Format("Number of Teams processed= {0}", $i+1))
  }
}

Write-Output ([string]::Format("Backfill completed. Groups backfilled: {0}, Groups
failed to backfill: {1}", $SuccessfullyBackfilledGroup, @($teamsToUpdate).Length -
$SuccessfullyBackfilledGroup))

if (!($SuccessfullyBackfilledGroup -eq @($teamsToUpdate).Length))
{
  Write-Output ([string]::Format("Check the failed teams in
BackfillFailedTeams.csv, retry to backfill the failed teams."))
}
```

Next steps

- Information Barriers in Microsoft Teams: Verify Teams IB behavior after updating group modes to Implicit.
- Use Information Barriers with SharePoint: Update IB modes for SharePoint sites connected to the updated Teams.
- Manage Information Barriers policies: Edit segments or policies if compliance issues arise after mode changes.

ⓘ **Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 04/14/2026

Information Barriers in Microsoft Teams

Applies to: Microsoft Teams

Microsoft Purview Information Barriers (IBs) are policies that prevent individuals or groups from communicating with each other. This article covers IB behavior in Microsoft Teams, including shared channels. For more information about shared channels and Information Barriers behavior, see Information Barriers and shared channels.

For Microsoft Teams, Information Barriers can determine and prevent the following kinds of unauthorized collaborations:

- Adding a user to a team or channel
- User access to team or channel content
- User access to 1:1 and group chats
- User access to meetings
- Prevents lookups and discovery, users aren't visible in the people picker.

> **Note**
>
> - You can't create IB groups across tenants.
> - Version 1 doesn't support using bots, Microsoft Entra apps, or APIs to send activity feed notifications, and it doesn't support some APIs to add users.
> - Private channels comply with Information Barriers policies that you configure.
> - For information about support for barriers for SharePoint sites that are connected to Teams, see Segments associated with Microsoft Teams sites.

Background

The primary driver for Information Barriers is the financial services industry, where the Financial Industry Regulatory Authority (FINRA) reviews IBs and conflicts of interest within member firms (Debt Research Regulatory Notice 15-31). However, IBs are useful across many industries including education, legal, government, and professional services. For more information about IB scenarios, see Learn about Information Barriers.

When to use Information Barriers

Use IBs in situations like these:

- You need to prevent a team from communicating or sharing data with a specific other team.
- You need to prevent a team from communicating or sharing data with anyone outside of the team.

The Information Barrier Policy Evaluation Service determines whether a communication complies with IB policies.

Manage segments and policies

Manage IB segments and policies in the Microsoft Purview portal or by using PowerShell cmdlets. For full guidance, see Get started with Information Barriers and Manage Information Barriers policies.

> **Important**
>
> Support for 5,000 segments and assigning users to multiple segments is only available when your organization isn't in *Legacy* mode. Assigning users to multiple segments requires extra steps to change the Information Barriers mode for your organization. For more information, see **Use multi-segment support in Information Barriers**.
>
> For organizations in *Legacy* mode, the maximum number of segments supported is 250, and users are restricted to being assigned to only one segment. Organizations in *Legacy* mode are eligible to upgrade to the newest version of Information Barriers in the future. For more information, see the **Information Barriers roadmap** .

> **Important**
>
> Before you set up or define policies, you must enable scoped directory search in Microsoft Teams. Wait at least a few hours after enabling scoped directory search before you set up or define policies for IB. For more information, see **Define IB policies**.

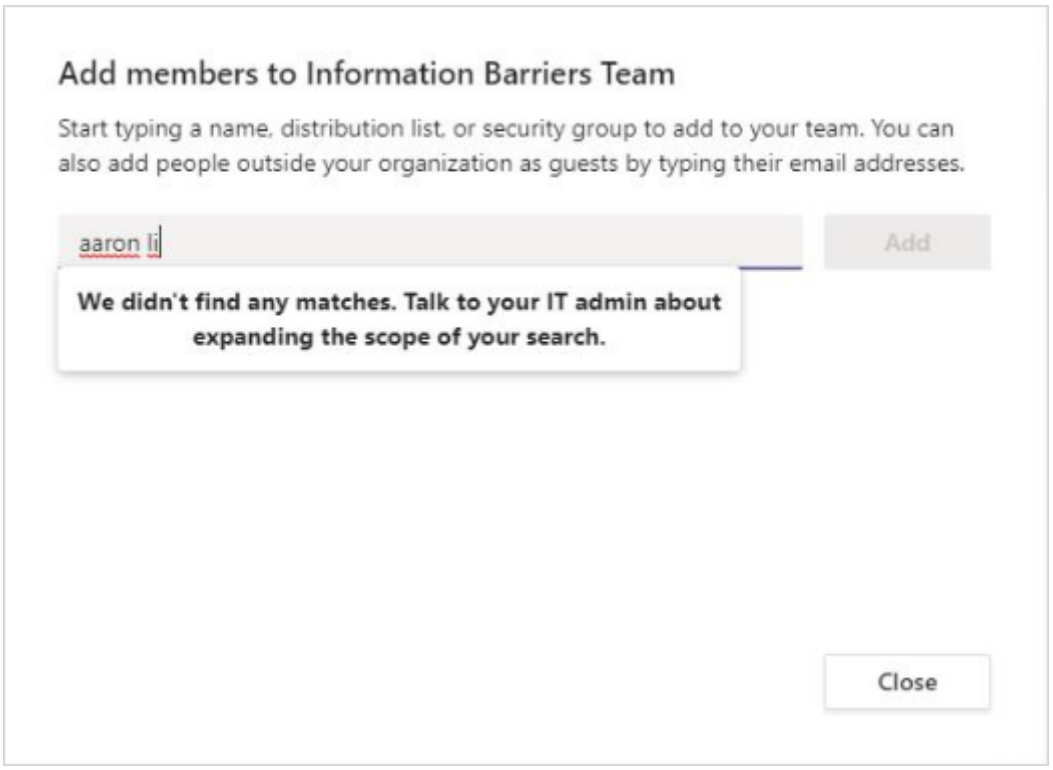
Information Barriers administrator role

The IB Compliance Management role is responsible for managing IB policies. For more information about this role, see Permissions in the Microsoft Purview portal.

Information barrier triggers

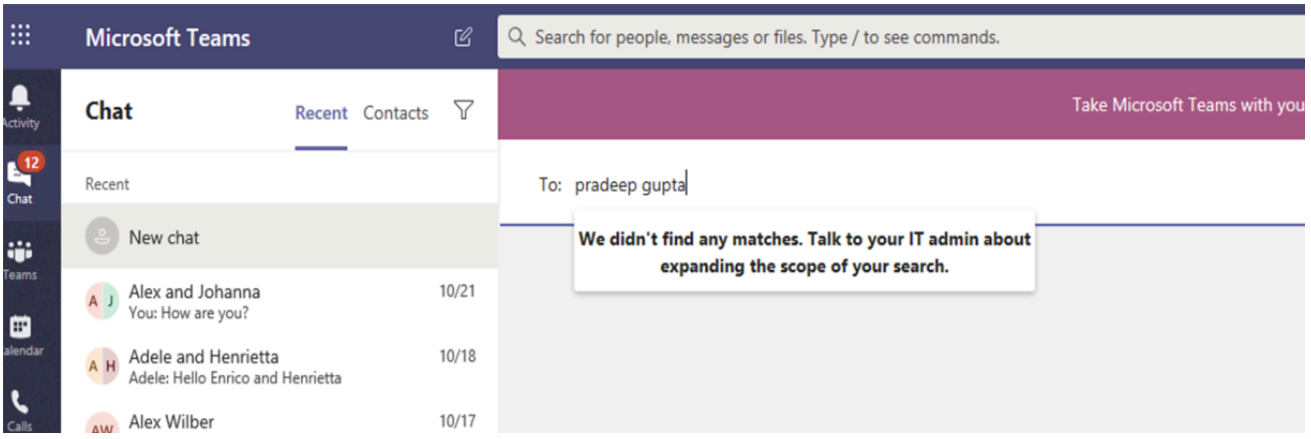
IB policies activate when the following Teams events take place:

- **Add members to a team**: Whenever you add a user to a team, the system evaluates the user's policy against the IB policies of other team members. After the user is successfully added, the user can perform all functions in the team without further checks. If the user's policy blocks them from being added to the team, the user doesn't show up in search.

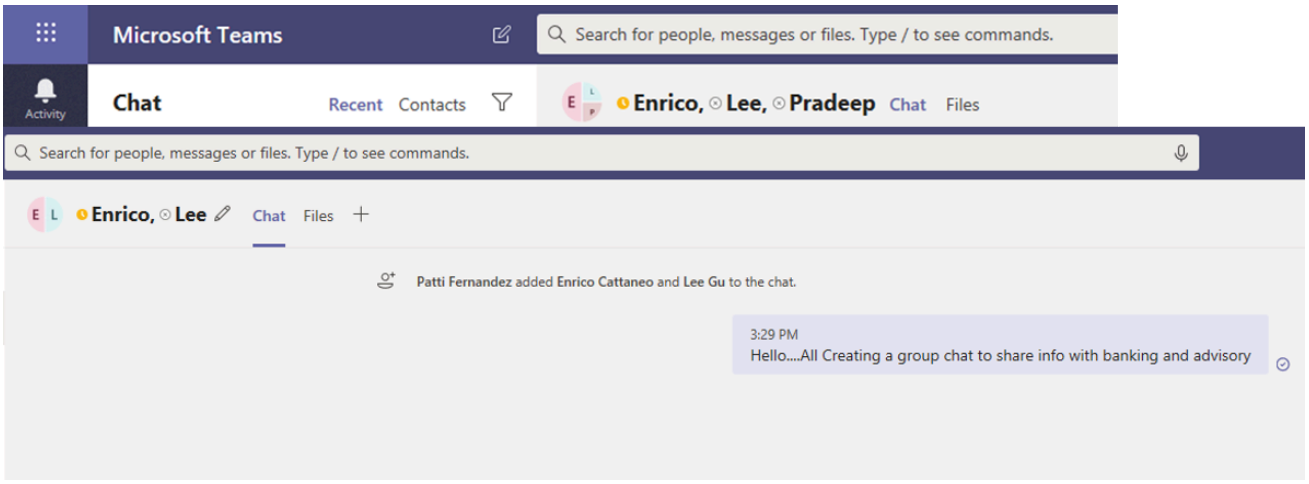


- **Request a new chat**: Each time that a user requests a new chat with one or more other users, the system evaluates the chat to make sure that it isn't violating any IB policies. If the conversation violates an IB policy, the conversation isn't started.

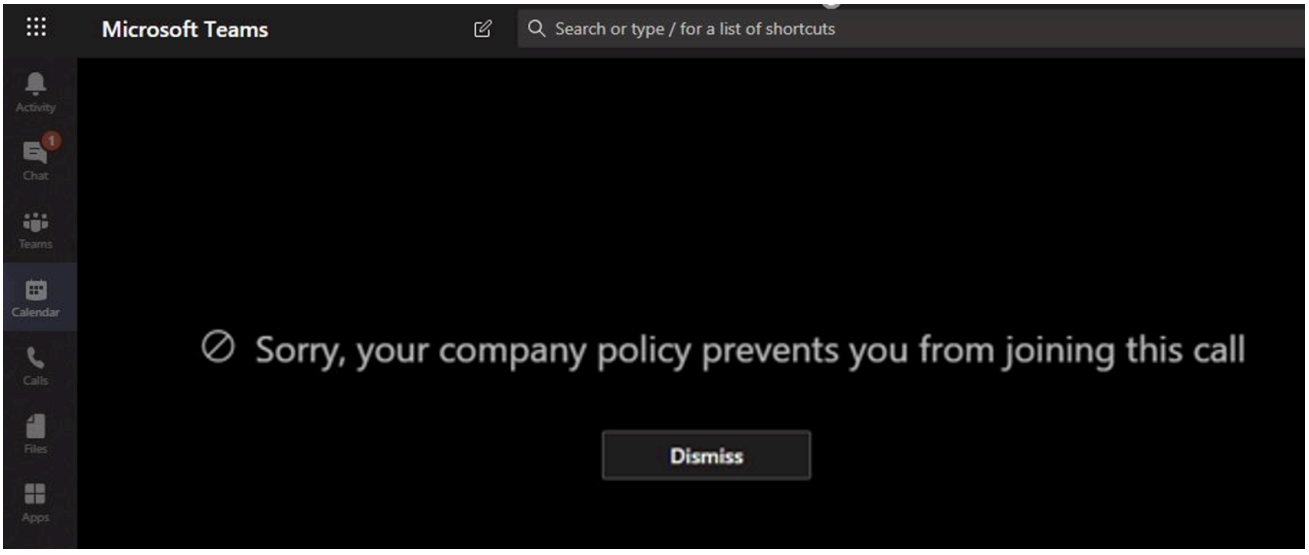
  Here's an example of a 1:1 chat.



  Here's an example of a group chat.

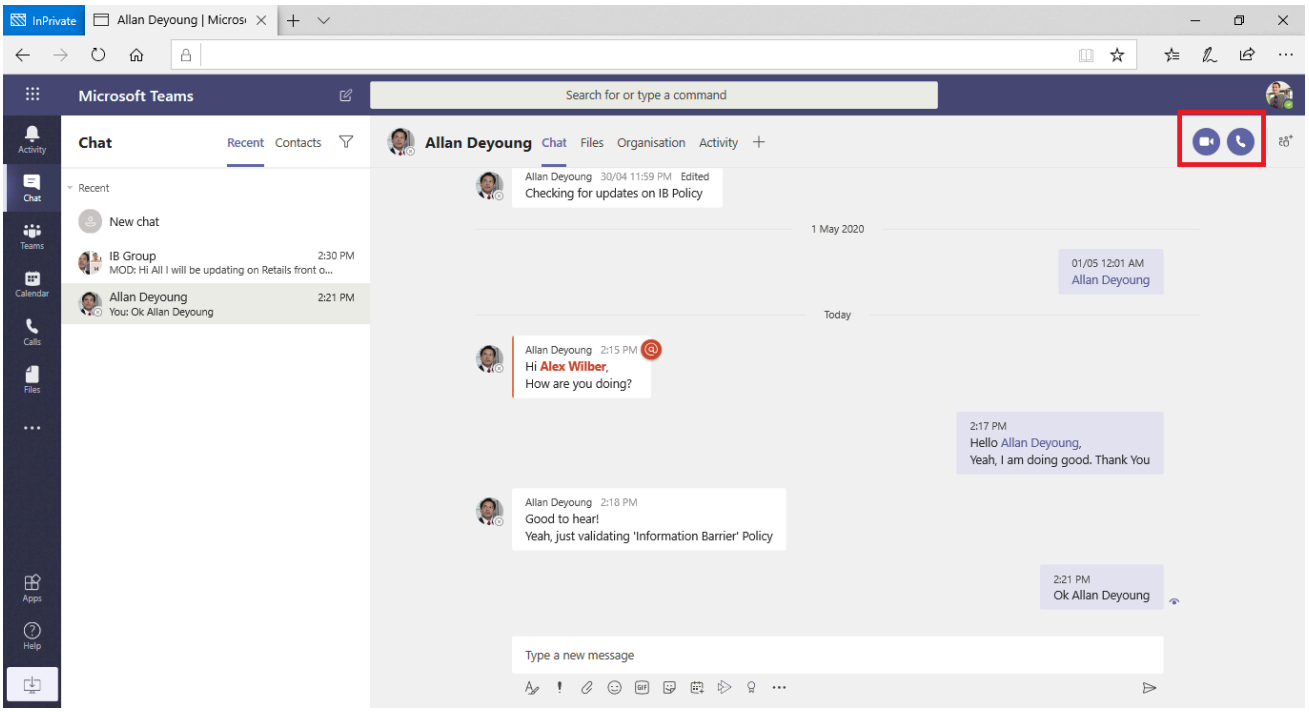


- **Invite a user to join a meeting**: When you invite a user to join a meeting, the system evaluates the IB policy that applies to the user against the IB policies that apply to the other team members. If there's a violation, the user isn't allowed to join the meeting.

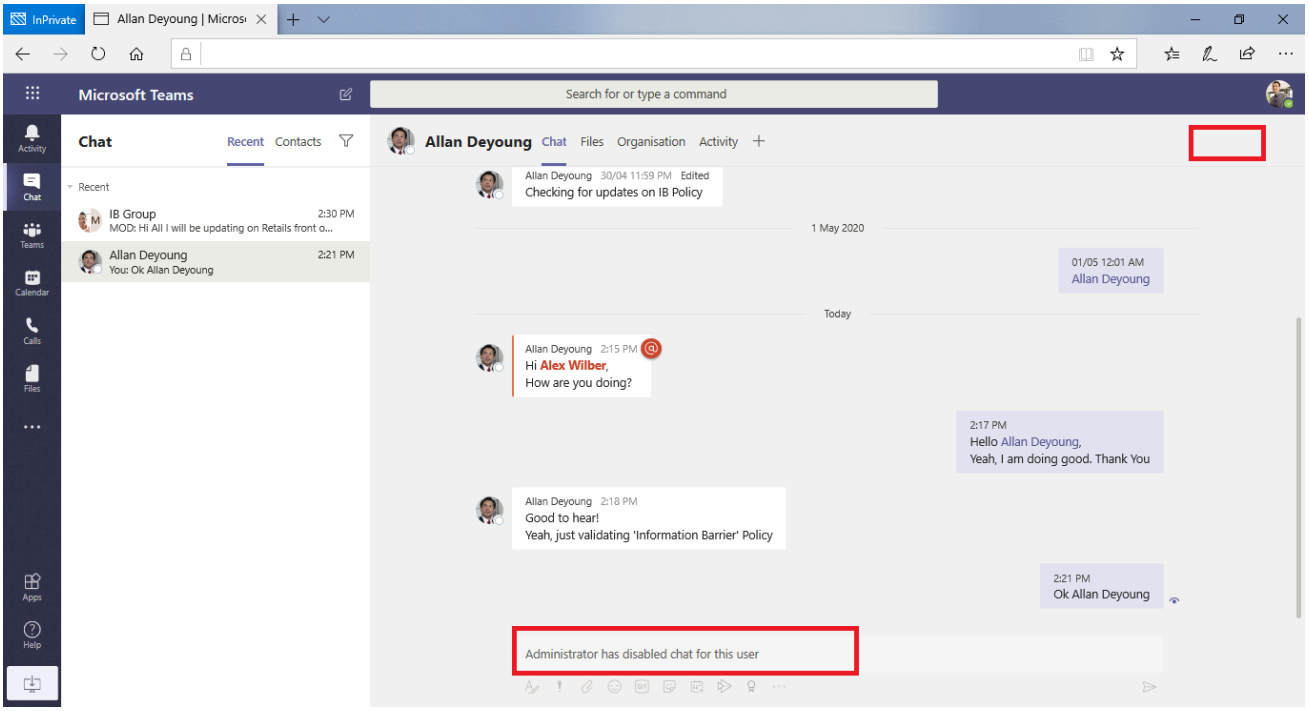


- **Share a screen between two or more users**: When a user shares a screen with other users, the system evaluates the sharing to make sure that it doesn't violate the IB policies of other users. If an IB policy is violated, the screen share isn't allowed.

  Here's an example of screen share before the policy is applied.



Here's an example of screen share after the policy is applied. The screen share and call icons aren't visible.



- **Place a phone call in Teams**: Whenever a user initiates a voice call (via VOIP) to another user or group of users, the system evaluates the call to make sure that it doesn't violate the IB policies of other team members. If there's any violation, the voice call is blocked.

- **Guests in Teams**: IB policies apply to guests in Teams, too. If guests need to be discoverable in your organization's global address list, see Manage guest access in Microsoft 365 Groups. After guests are discoverable, you can define IB policies.

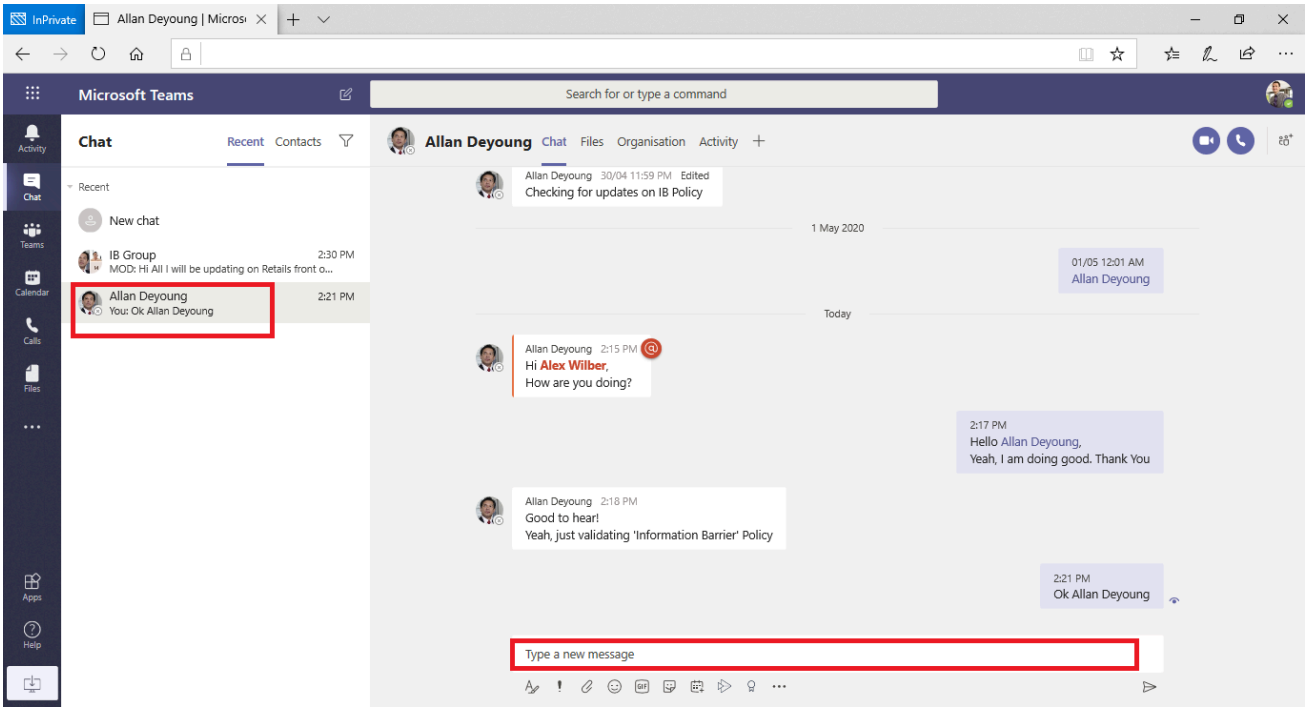
How policy changes impact existing chats

When the IB policy administrator changes a policy or activates a policy change because of a change to a user's profile (such as for a job change), the Information Barrier Policy Evaluation Service automatically searches the members to ensure that their membership in the team doesn't violate any policies.

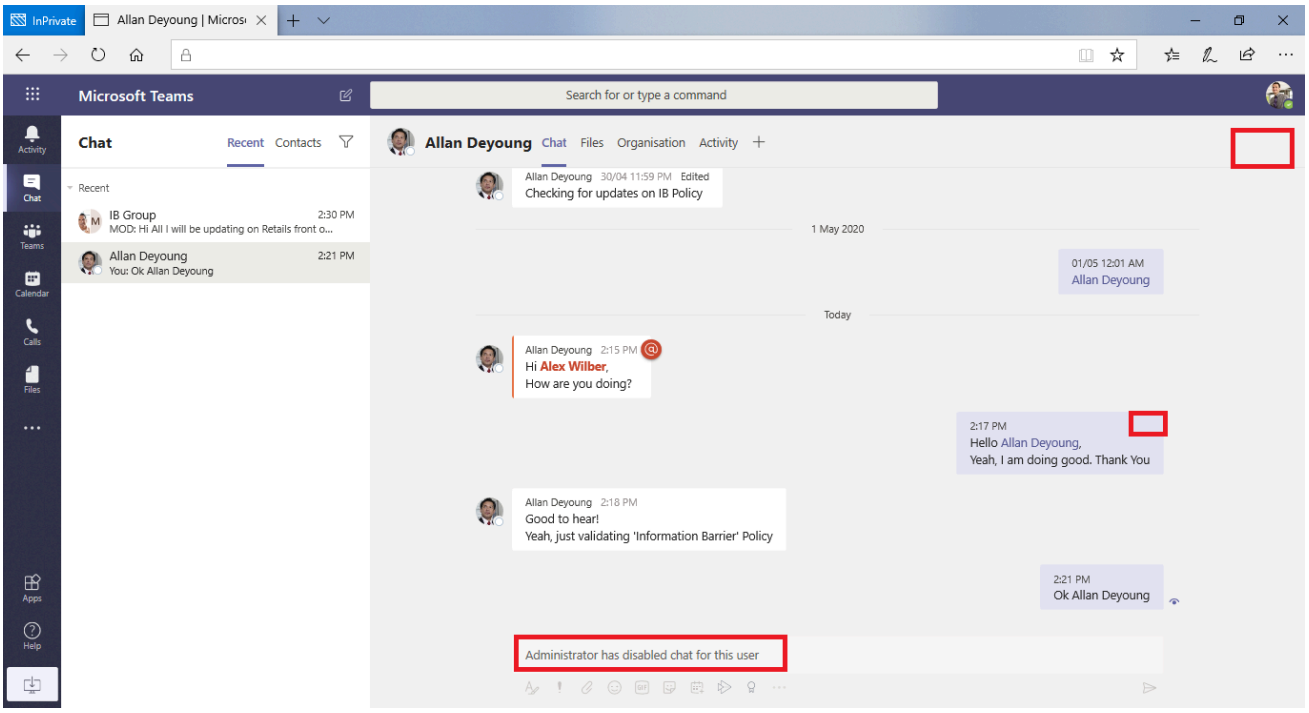
If users have an existing chat or other communication, and you set a new policy or change an existing policy, the service evaluates existing communications to make sure that the communications are still allowed.

- **1:1 chat**: If communication between two users is no longer allowed (because of application to one or both users of a policy that blocks communication), the service blocks further communication. Their existing chat conversations become read-only.

  Here's an example that shows the chat is visible.

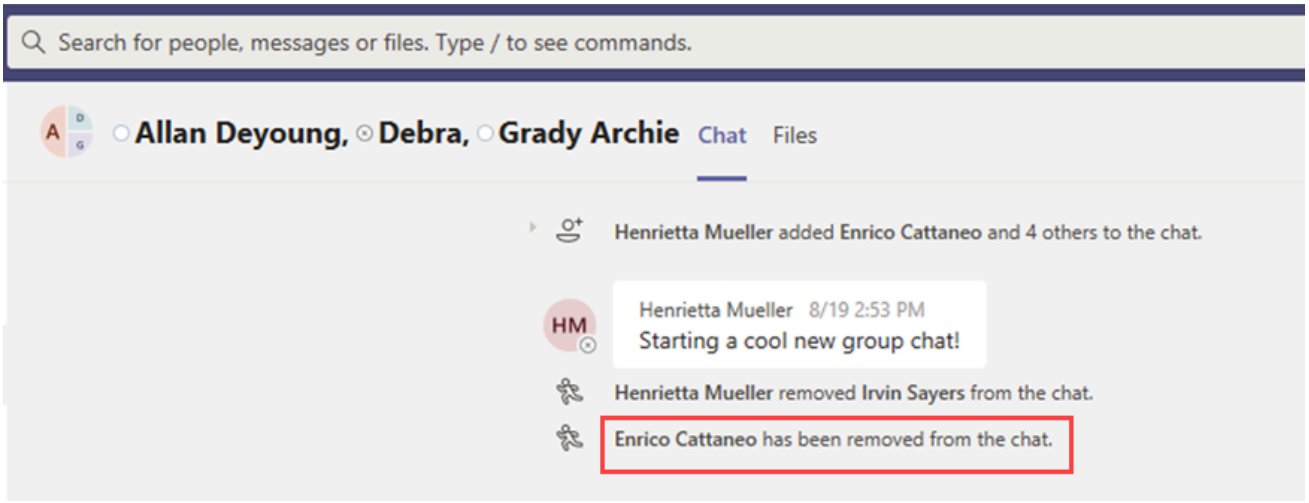


  Here's an example that shows the chat is disabled.

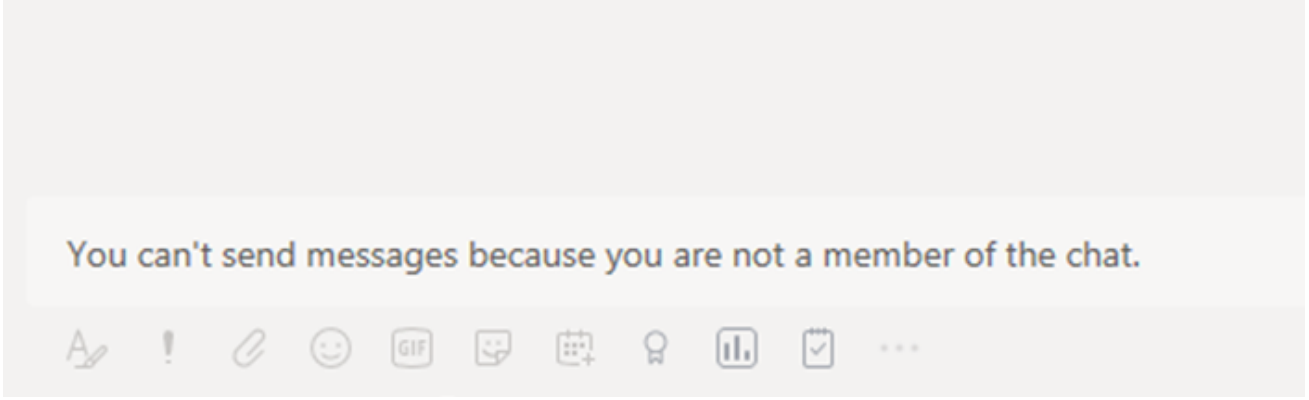


- **Group chat**: If communication from one user to a group is no longer allowed (for example, because a user changed jobs), the service removes the user, along with the other users whose participation violates the policy, from group chat, and blocks further communication with the group. The user can still see old conversations, but can't see or participate in any new conversations with the group. If the new or changed policy that prevents communication is applied to more than one user, the service removes the users who are affected by the policy from group chat. They can still see old conversations.

  In this example, Enrico moves to a different department within the organization and is removed from the group chat.



Enrico can no longer send messages to the group chat.

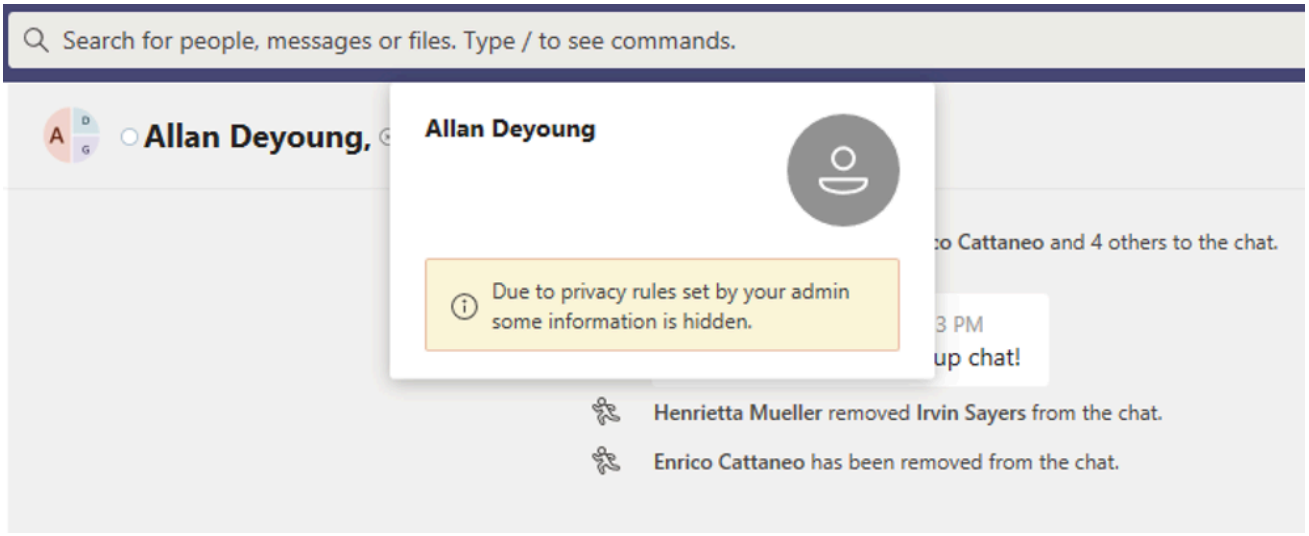


- **Team**: The service removes any users who are removed from the group from the team and they can't see or participate in existing or new conversations.

Scenario: A user in an existing chat becomes blocked

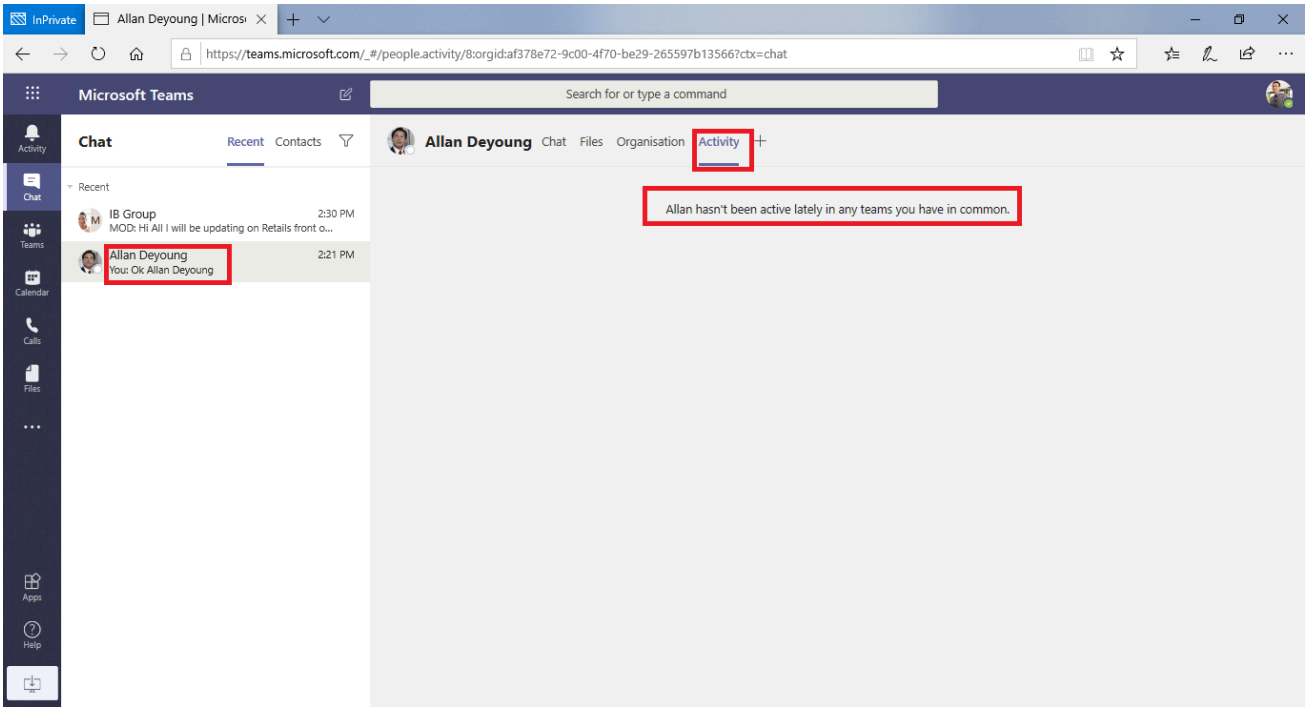
Currently, users experience the following scenarios if an IB policy blocks another user:

- **People tab**: A user can't see blocked users on the **People** tab.
- **People Picker**: Blocked users aren't visible in the people picker.



- **Activity tab**: If a user visits the **Activity** tab of a blocked user, no posts appear. (The **Activity** tab displays channel posts only, and there are no common channels between the two users.)

  Here's an example of the activity tab view that is blocked.



- **Org charts**: If a user accesses an org chart on which a blocked user appears, the blocked user doesn't appear on the org chart. Instead, an error message appears.

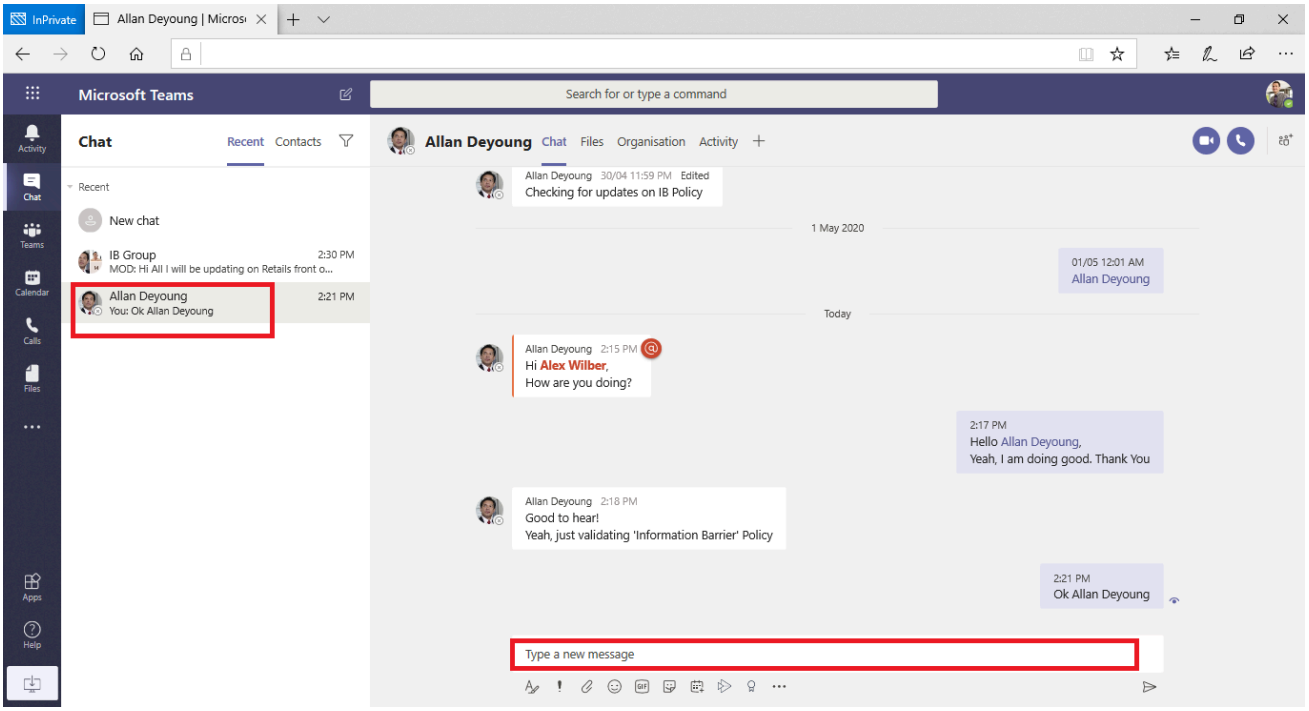
- **People card**: If a user participates in a conversation and the user is later blocked, other users see an error message instead of the people card when they hover over the blocked user's name. Actions listed on the card (such as calling and chat) are unavailable.

- **Suggested contacts**: Blocked users don't appear on the suggested contacts list (the initial contact list that appears for new users).

- **Chat contacts**: A user can see blocked users on the chats contact list, but the blocked users are identified. The only action that the user can perform on the blocked users is to delete them. The user can also select them to view their past conversation.

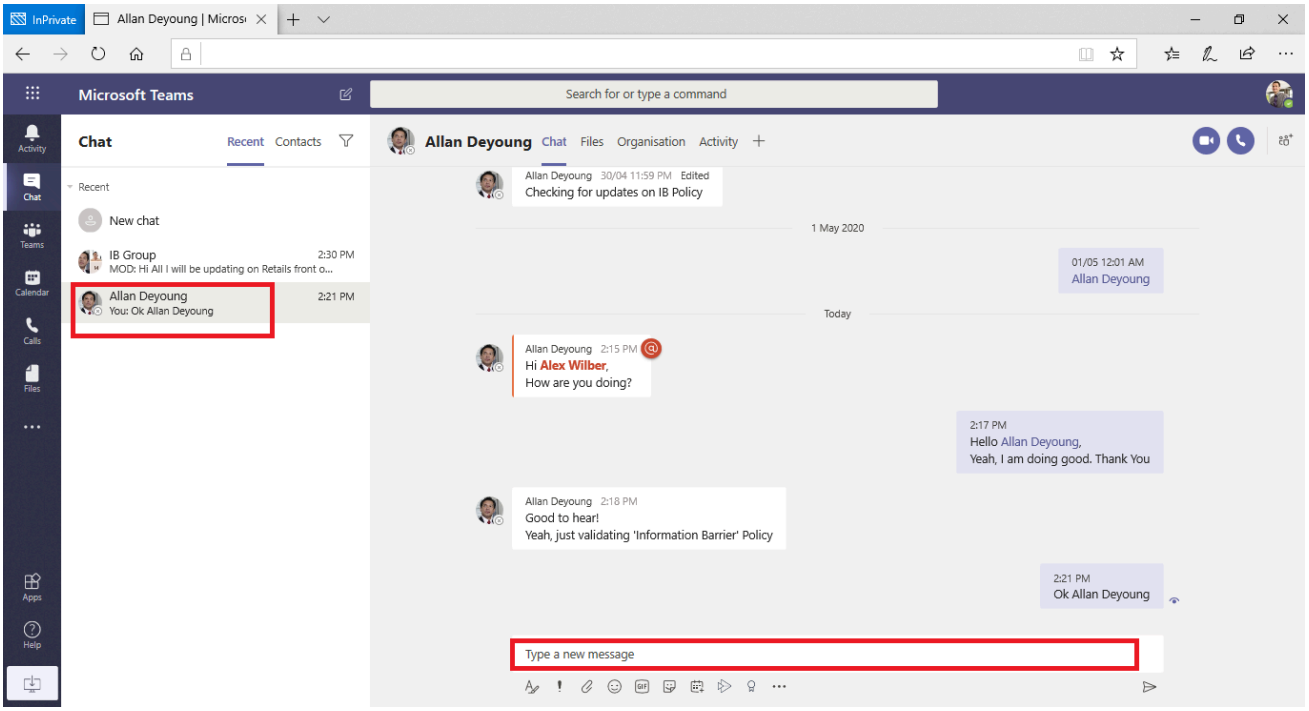
- **Calls contacts**: A user can see blocked users on the calls contact list, but the blocked users are identified. The only action that the user can perform on the blocked users is to delete them.

  Here's an example of a blocked user in the calls contact list.

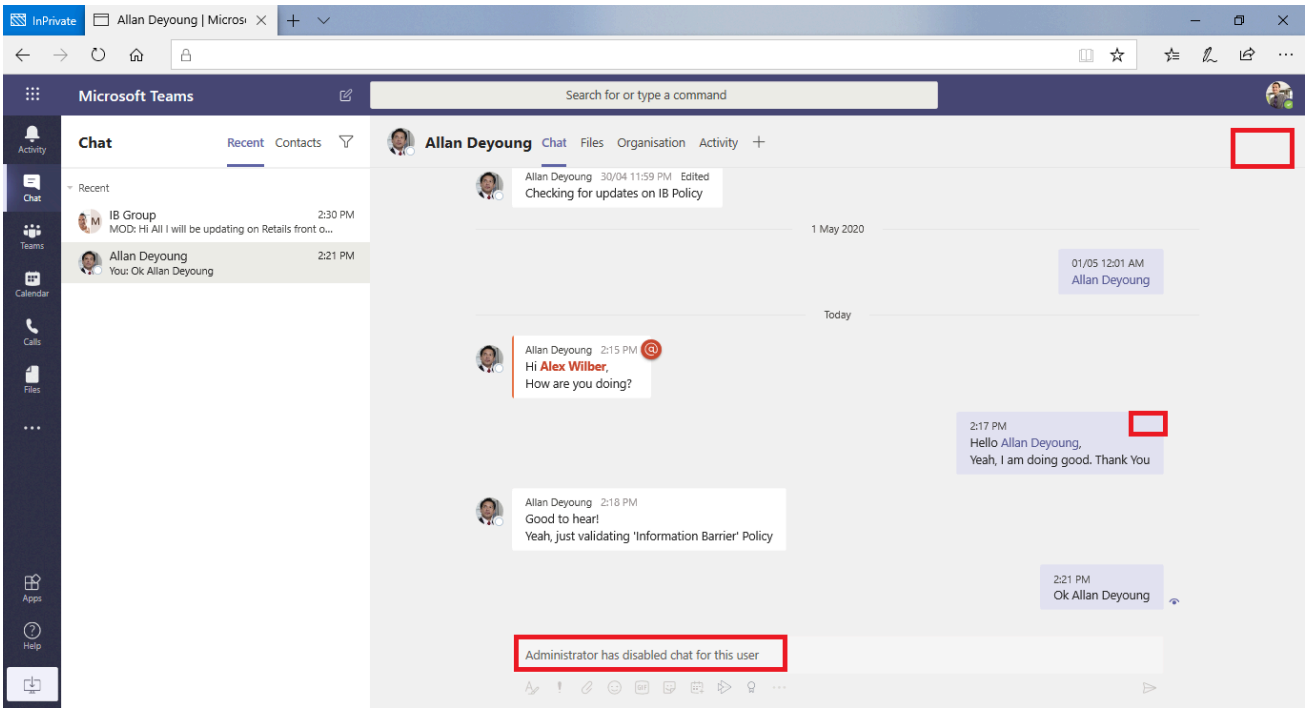


- **Skype to Teams migration**: During a migration from Skype for Business to Teams, the migration process includes all users—even those users who are blocked by IB policies. The migration process handles those users as described previously.

  Here's an example of a blocked user in the calls contact list.



Here's an example of the chat being disabled for a user on the calls contact list.



Teams policies and SharePoint sites

When you create a team, you also create and associate a SharePoint site with Microsoft Teams for the files experience. Information barrier policies don't apply to this SharePoint site and its files by default. To enable Information Barriers in SharePoint and OneDrive, follow the guidance and steps in the Use Information Barriers with SharePoint article.

Information barrier modes and Teams

Information barrier modes help control who you can add to or remove from a Team. When you use Information Barriers with Teams, the following IB modes are supported:

- **Open**: This configuration is the default IB mode for all existing groups that were provisioned before Information Barriers were enabled. In this mode, no IB policies apply.
- **Implicit**: This configuration is the default IB mode when you provision a Team after enabling Information Barriers. *Implicit* mode allows you to add all compatible users in the group.
- **Owner Moderated**: This mode is set on a team when you want to allow collaboration between incompatible segment users that the owner moderates. The team owner can add new members per their IB policy.

Teams created before activating an IB policy in your tenant are automatically set to *Open* mode by default. When you activate IB policies on your tenant, you must update mode of your existing teams to *Implicit* to ensure that existing teams are IB-compliant. For more information about updating modes, see Change Information Barriers modes with a PowerShell script.

Use the Set-UnifiedGroup cmdlet with the *InformationBarrierMode* parameter that corresponds to the mode you want to use for your segments. The allowed values for the *InformationBarrierMode* parameter are *Open*, *Implicit*, and *Owner Moderated*.

For example, to configure *Implicit* mode for a Microsoft 365 Group, use the following PowerShell command:

```
PowerShell
Set-UnifiedGroup -InformationBarrierMode Implicit
```

To update the mode from *Open* to *Implicit* for all existing teams, use this PowerShell script.

If you change the *Open* mode configuration on existing Teams-connected groups to meet compliance requirements for your organization, you need to update the IB modes for associated SharePoint sites connected to the Teams team.

IB policy application in Teams

The IB policy application is a background IB processor for Teams that gets a notification when there are changes to either users (policy or segment changes) or groups (mode changes). The following steps outline the processing flow:

- The policy application receives a group change notification when mode is updated and retrieves the message thread and Group IDs applicable to the update.
- If the message thread exists, the application schedules processing and fetches all members from the team. It sends the underlying group and members to downstream Teams components for IB evaluation.
- The application evaluates the mode on the group and the IB policies per user and sends the results to the policy application.
- The policy application removes the non-compliant users from the group and team.

Required licenses and permissions

For more information about licenses and permissions, plans, and pricing, see the subscription requirements for Information Barriers.

Usage notes

- **Users can't join any scheduled meetings**: If you enable IB policies, users can't join meetings if the meeting chat roster reaches its current limit of 1,000 users. IB checks rely on whether a user can be added to the meeting chat roster - only users who can be added are allowed to join the meeting. Every meeting in an IB-enabled tenant performs IB checks at the point of entry, even if the meeting organizer isn't explicitly covered by IB policies.

  Teams also enforces a separate meeting limit of 1,000 users. If you enable overflow, additional participants can join as view-only attendees. View-only participants aren't subject to IB policy checks. If you invite a large audience (for example, more than 1,000 users), Teams pre-adds 750 users to the meeting chat before the meeting starts. This process reserves room for about 250 additional users who actually join the meeting. Because pre-added users count toward the chat roster limit even if they never join the meeting, the chat roster can reach capacity before the meeting attendance limit is met. For recurring meetings, the roster can fill quickly, as any user who joins once remains on the roster for future occurrences. After the 1,000-user chat roster limit is reached, you can't add more users to the meeting.

  A short-term solution is to remove inactive members from the meeting chat roster to make space for new users or keep the chat disabled until the meeting begins and enable it only when needed. This solution helps align the chat participant count more closely with the meeting participant count. Regular Teams meetings work best for audiences of fewer than 1,000 expected attendees. For larger audiences, use Teams Town hall for a smoother experience.

- **Users can't join channel meetings**: If you enable IB policies, users can't join channel meetings if they're not a member of the team. The root cause is that IB checks rely on whether users can be added to a meeting chat roster, and only when they can be added to the roster are they allowed to join the meeting. The chat thread in a channel meeting is available to Team/Channel members only, and non-members can't see or access the chat thread. If you enable IB for the organization and a non-team member attempts to join a channel meeting, that user isn't allowed to join the meeting. However, if you don't enable IB for the organization and a nonteam member attempts to join a channel meeting, the user is allowed to join the meeting but they don't see the chat option in the meeting.

- **IB policies are supported in federated communication**: If you allow federation with external organizations, the users of those organizations aren't restricted by IB policies. If users of your organization join a group chat, meeting chat, or a call organized by external federated users, IB policies continue to restrict communication between users of your organization.

Next steps

- Use Information Barriers with SharePoint: Enable IB on SharePoint sites connected to Teams.
- Information Barriers and shared channels: Understand how IB policies apply when sharing channels with other teams.
- Change Information Barriers modes with a PowerShell script: Update existing Teams groups from Open to Implicit mode after enabling IB.
- Manage Information Barriers policies: Edit or remove IB policies and segments.

ⓘ **Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/01/2026

Information Barriers and moderated meetings

Applies to: Microsoft Teams

Moderated meetings in Information Barriers (IB) enable controlled interactions between users who can't communicate because of IB policies. This feature ensures that meetings occur with the presence of a designated moderator, such as a compliance admin, who oversees the interaction and ensures compliance with IB policies.

Moderated meetings in IB provide the following benefits:

- **Granular control**: Selectively allow meetings with IB-incompatible users, such as town halls or new hire orientations.
- **Moderator role**: Designated moderators facilitate the meeting, ensuring that all interactions comply with IB policies.
- **Flexible user discovery**: Users can look up and invite participants on either side of the IB during scheduling.
- **Compliance and security**: Adheres to regulatory requirements and provides a secure environment for sensitive discussions.

> **Important**
>
> Your organization must use the IBv2 architecture to use moderated meetings in IB. The IBv2 architecture supports flexible people discovery and the creation of a separate moderator segment. This support is required for necessary compliance and security measures in moderated interactions.

Set up moderators, segments, and policies

Setting up moderators and associated segments and policies is crucial for conducting moderated meetings where attendees from segments blocked from communicating with each other attend a common meeting.

To configure these settings, you must be assigned one of the following roles in your organization:

- Microsoft 365 global administrator
- Office 365 global administrator
- Compliance administrator
- IB Compliance Management

> **ⓘ Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

Additionally, all users included in Moderated Meeting must meet the following prerequisites:

- All users and the admin must have a Microsoft 365 E5 license
- All users scheduling a Moderated Meeting in Teams must have a Microsoft Teams Premium license.

Step 1: Define a moderators group in Microsoft Entra

Complete the following steps to define a moderators group in Microsoft Entra:

1. Go to the Microsoft Entra admin center, select **Expand groups**, select **Select all groups**, and then select **Add new group**.
2. For **Group type**, select *Microsoft 365*.
3. Enter a name for the group in the **Group name** field.
4. Enter *moderator* in the **Group email address** field. If you delete the moderators' group or this group name isn't available, use one of the following address names:

   - moderators1@<domain>.com
   - moderators2@<domain>.com

   > **ⓘ Important**
   >
   > The name of the group email ID must contain the keyword *moderator* and be in stated format.

5. Enter a description for the group in the **Group description** field.

6. Select **No** for **Microsoft Entra roles can be assigned to this group**.
7. Select *Assigned* for **Membership type**.
8. Add owners and moderators to this group.
9. Select **Create**.

You use the *Object ID* of the moderator group in the next step for the creation of a moderators segment in Microsoft Purview portal.

Step 2: Define the moderators segment

After you configure a Microsoft Entra group with all the needed moderators, use this group to configure a new moderator segment.

Complete the following steps to define a moderators group in the Microsoft Purview portal:

1. Go to the Microsoft Purview portal and navigate to **Information Barriers** > **Segments**.
2. Select **New segment**.
3. On the **Provide a segment name** page, enter a name for the segment and select **Next**.
4. On the **Add user group filter** page, select **Add**.
5. Select the *MemberOf* filter.
6. Copy the *Object ID* of the group you created in Step 1 in Microsoft Entra and paste it in the **Member of** condition field with the *Equal* operator. Select **Next**.
7. Review the segment details and select **Submit** to create the segment.

When you create the segment, add or remove owners and moderators by updating the Microsoft Entra group. This segment updates automatically if there are any membership changes in the Microsoft Entra group.

Step 3: Create a policy for the moderator segment

Create an *Allow* policy for the moderator segment. This policy enables moderators to facilitate meetings for participants from all segments, even if existing IB policies block those participants.

To create a policy for the moderator segment in the Microsoft Purview portal, complete the following steps:

1. Go to the Microsoft Purview portal and navigate to **Information Barriers** > **Policies** > **Policies**.
2. Select **Create policy**.

3. On the **Provide a policy name** page, enter a name for the policy and select **Next**.
4. On the **Add assigned segment details** page, select **Choose segment**.
5. Select the segment you created in Step 2, and then select **Add** and **Next**.
6. On the **Configure communication and collaboration details** page, select **Allowed**.
7. Select the **Allow moderation** checkbox.
8. Add all the segments in your organization that should participate in a moderated meeting. Add all segments in your organization, including the moderators segment you created in Step 2.
9. Select **Add** and **Next**.
10. On the **Policy status** page, enable **Set your policy to active status** and select **Next**.
11. Review the policy details and select **Submit** to create the policy.

To verify that the policy is applied in your organization, go to **Information Barriers** > **Policies** > **Policy applications** and select **Apply all policies**.

> ⓘ **Note**
>
> When you select **Apply all policies**, the system evaluates all existing and new policies. You might see some validation errors. Update policies to correct any errors as needed and rerun the evaluation. Applying all policies might take several hours, depending on the size of your organization.

Step 4: Setting policy configuration

> ⓘ **Important**
>
> Before you begin, make sure you have the following PowerShell modules installed:
>
> - Exchange Online Management module (`ExchangeOnlineManagement`)
> - Security & Compliance PowerShell (for `Connect-IPPSSession`)

Complete the following steps to change your IB mode to multi-segment and to disable the IB *InformationBarrierPeopleSearchRestriction* property.

1. Start PowerShell in admin mode and use the following command to change your IB mode to multi-segment:

```
PowerShell
```

```
Connect-IPPSSession -UserPrincipalName <compliance admin email id>
Import-Module ExchangeOnlineManagement
Get-PolicyConfig | fl "Information*"
Set-PolicyConfig -InformationBarrierMode 'MultiSegment'
```

> **Note**
>
> If you configure any Block policies in your organization, multi-segment doesn't work. For multi-segment IBv2 to work, configure your organization with Allow policies only. All segments that aren't set to allow are blocked.

2. Use the following cmdlet to disable the *InformationBarrierPeopleSearchRestriction* property:

PowerShell

```
Set-PolicyConfig -InformationBarrierPeopleSearchRestriction 'Disabled'
```

To verify that you configured multi-segment support and disabled the *InformationBarrierPeopleSearchRestriction* property, use the following cmdlet:

PowerShell

```
Get-PolicyConfig | fl "Information*"
```

Step 5: Apply policies

To apply the changes in Step 4, you must apply the policy changes again in your organization. To apply these changes in your organization, go to **Information Barriers** > **Policies** > **Policy applications** and select **Apply all policies**.

> **Note**
>
> When you select **Apply all policies**, the system evaluates all existing and new policies. You might see some validation errors. Update policies to correct any errors as needed and rerun the evaluation. Applying all policies might take several hours, depending on the size of your organization.

Next steps

- Information Barriers in Microsoft Teams: Understand how IB policies affect other Teams interactions beyond moderated meetings.
- Get started with Information Barriers: Apply updated policies after configuring moderator segments.
- Use multi-segment support in Information Barriers: Assign moderators to multiple segments for broader meeting facilitation.

ⓘ **Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 04/14/2026

Information Barriers and shared channels

Applies to: Microsoft Teams

Shared channels in Microsoft Teams create collaboration spaces where you can invite people who aren't in the team. You implement Microsoft Purview Information Barriers policies to restrict and prevent users and groups from communicating with each other within and outside of your organization.

Shared channels are enabled by default in Teams. You can decide if people can create shared channels, share them with people outside your organization, and participate in external shared channels by creating a channel policy. When you create Information Barriers policies in your organization, Teams performs checks when configuring shared channels to verify that none of the existing channel members and any new users added to the shared channel violate Information Barriers policy conditions.

Use the following table to understand how Information Barriers policies affect communications and result in specific behaviors when configuring shared channels:

Expand table

Scenario	Information Barriers behavior
Share a channel with a user in your organization	If an Information Barriers policy prevents the user from communicating with shared channel members, the user doesn't appear in the user search and the channel isn't shared with the team. If an Information Barriers policy prevents adding the user, you see the following message: *We didn't find any matches. Talk to your IT admin about expanding the scope of your search.*
Share a channel in your organization with another team that you own	If an Information Barriers policy prevents users in the other team from communicating with shared channel members, the channel isn't shared with the team. If an Information Barriers policy prevents communications, you see the following message: *The channel can't be shared with this team. Pick another team or contact your admin for more info.*
Share a channel in your organization with another team that you don't own	If an Information Barriers policy prevents the team owner or any users of the other team from communicating with shared channel members, the channel can't be shared with the team.

Scenario	Information Barriers behavior
	If an Information Barriers policy prevents communications, you see the following message: *The channel can't be shared with this team. Pick another team or contact your admin for more info.*
Add a new user to the team when the team has shared channels with other teams	If an Information Barriers policy prevents the new user from communicating with members of the shared channel team, the user can't be added to the team. When you add a user to a team with six or more shared channels, you immediately add the user to the channel and support sharing. Sharing for the team and the previously shared channels stops if an Information Barriers policy finds the new user to be non-compliant. If an Information Barriers policy prevents adding the user, you see the following message: *Unable to add user due to an Information Barriers policy.*
Share a channel with an external team	Information Barriers policies on internal and external organizations don't restrict communications between users from the different organizations. You can share shared channels with external guests.

Next steps

- Information Barriers in Microsoft Teams: Understand how IB policies affect chats, meetings, and standard channels.
- Manage Information Barriers policies: Edit policies if shared channel restrictions need adjustment.
- Use Information Barriers with SharePoint: Configure IB on SharePoint sites associated with shared channels.

ⓘ **Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 04/14/2026

Use Information Barriers with SharePoint

Microsoft Purview Information Barriers (IB) are policies that prevent users from communicating and collaborating with each other. This article covers IB behavior in SharePoint.

For SharePoint, Information Barriers can determine and prevent the following kinds of unauthorized collaborations:

- Adding a user to a site
- User access to a site or site content
- Sharing a site or site content with other users

Information Barriers modes and SharePoint sites

Information Barriers modes help strengthen access, sharing, and membership of a site based on its IB mode and segments associated with the site.

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

When you use Information Barriers with SharePoint, the following IB modes apply:

- **Open**: The site has no segments. Anyone can be invited to be a member. This mode is the default when a SharePoint site doesn't have segments.
- **Owner Moderated**: The site owner moderates collaboration between incompatible segments. Only the resource owner can add new members per their IB policy.
- **Implicit**: Microsoft Teams sets the site's IB mode when it provisions the site. A SharePoint Administrator or Global Administrator can't manage segments with the *Implicit* mode configuration.
- **Explicit**: Segments are associated with the site, either through user site creation or by a SharePoint Administrator. Only users matching the site's segments can access and share content.

Open mode

Open mode sharing behavior

When a site has no segments and you set the site's IB mode to *Open*:

- Users can share the site and its contents based on the IB policy applied to them. For example, if a user in HR is allowed to communicate with users in Research, the user can share the site with those users.

> **Tip**
>
> If you want to allow sharing of *Open* mode sites with mail-enabled security groups, see the **Allow sharing of Open mode sites with mail-enabled security groups** section in this article.

Open mode access behavior

For a user to access a SharePoint site that has no segment and the site's Information Barriers mode is set to *Open*:

- The user has site access permissions.

Owner Moderated mode

Owner Moderated sharing behavior

When you set a site's Information Barriers mode to *Owner Moderated*:

- The option to share with *Anyone with the link* is disabled.
- The option to share with *Company-wide link* is disabled.
- (For group connected sites) The site and its content can be shared with existing members.
- (For non-group connected sites) Only the site owner can share the site and its content, according to their IB policy.

Owner moderated access behavior

For a user to access a SharePoint site with the site's Information Barriers mode set to *Owner Moderated*:

- (For non-group connected sites) The user has site access permissions.
- (For group connected sites) The user must be a member of the Microsoft 365 group connected to the site.

Implicit mode

Implicit mode sharing behavior

When you set a site's Information Barriers mode to *Implicit*:

- The option to share with *Anyone with the link* is disabled.
- The option to share with *Company-wide link* is disabled.
- The site and its content can be shared with existing members via a sharing link.
- You can't add new users to the site directly. The Team owner should add users to the Team's group by using Microsoft Teams.

> **Note**
>
> If you enabled Information Barriers for SharePoint in your organization before March 15, 2022, see the **Enable SharePoint and OneDrive Information Barriers** section in this article.

Implicit mode access behavior

To access SharePoint sites that use Information Barriers mode set to *Implicit*:

- You're a member of the Microsoft 365 group connected to the site.
- If you're not a member of the Microsoft 365 group connected to the site, you can't access the site.
- The Information Barriers compliance assistant ensures the group membership is IB compliant.

> **Note**
>
> If you enabled Information Barriers for SharePoint in your organization before March 15, 2022, see the **Enable SharePoint and OneDrive Information Barriers** section in this article.

Explicit mode

Explicit mode sharing behavior

When you associate a site with segments and set the site's Information Barriers mode to *Explicit*:

- The option to share with *Anyone with the link* is disabled.
- The option to share with *Company-wide link* is disabled.
- You can share the site and its content only with users whose segment matches that of the site. For example, if you associate a site with the HR segment, you can share the site with just HR users (even though HR is compatible with both Sales and Research segments).
- You can add new users as site members only if their segment matches the segment of the site.

Explicit mode access behavior

To access SharePoint sites that use segments and site's Information Barriers mode is *Explicit*:

- Your segment matches a segment that's associated with the site.

  AND

- You have access permission to the site.

Non-segment users can't access a site associated with segments. They see an error message.

Allow apps running in app-only mode to access IB sites

Many organizations use applications running in an app-only context in their organization. To allow these apps running in app-only mode to access IB protected sites, SharePoint admins can enable the opt-in capability.

> **ⓘ Important**
>
> Information Barriers policies might impact the applications accessing sites in app-only mode. We recommend that you enable the policy and then test the experience for the apps used in your organization.

To enable applications running in app-only mode to access IB sites, run the following command:

PowerShell

```
Set-SPOTenant -AppBypassInformationBarriers $true
```

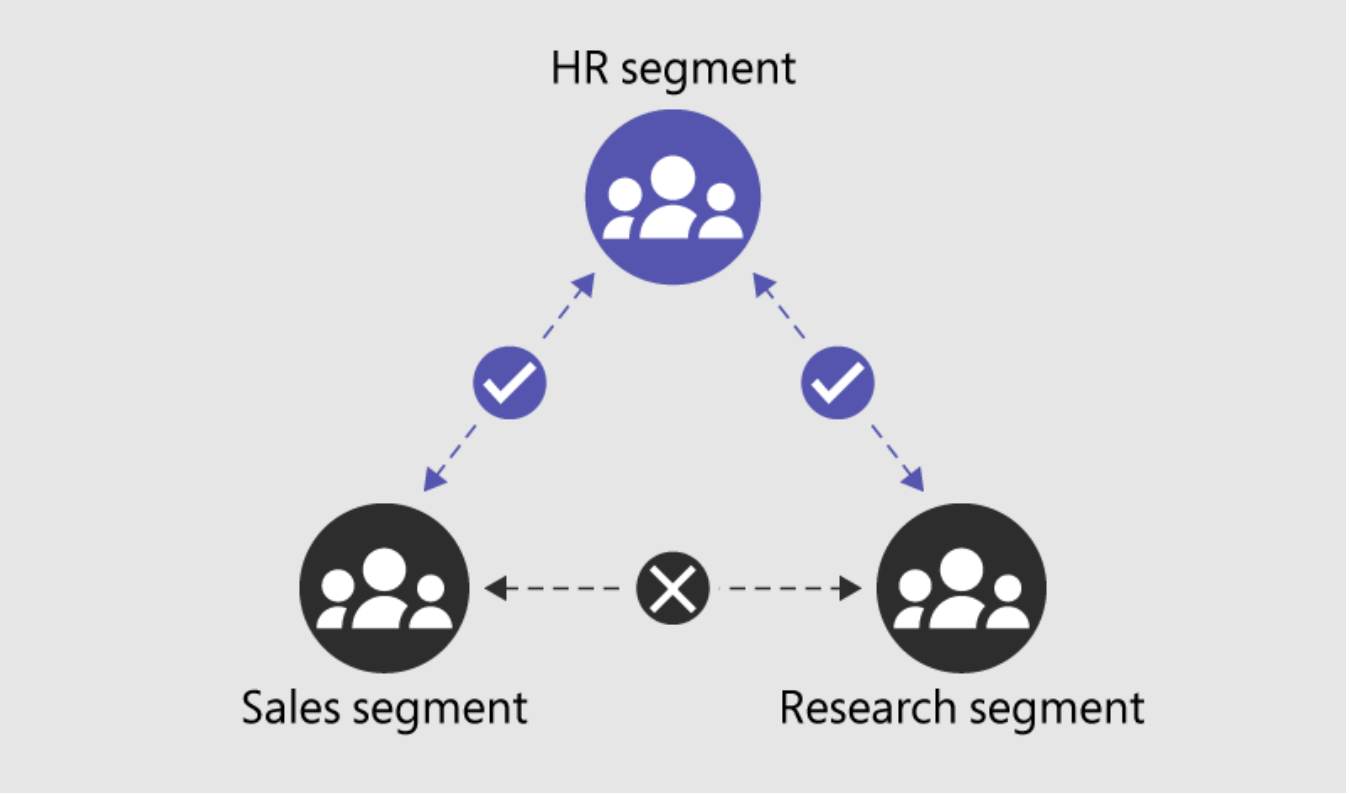
If you enable Teams Meeting Recording or EDU Assignment application in your organization, run the following command to allow these applications to interact with IB protected sites:

PowerShell

```
Set-SPOTenant -AppOnlyBypassPeoplePickerPolicies $true
```

Example scenario

The following example illustrates three segments in an organization: HR, Sales, and Research. An IB policy blocks communication and collaboration between the Sales and Research segments. These segments are incompatible.



By using SharePoint Information Barriers, a SharePoint Administrator or Global Administrator can associate segments to a site to prevent the site from being shared with or accessed by users outside the segments. You can associate up to 100 compatible segments with a site. You associate the segments at the site level (previously called site collection level). The Microsoft 365 group connected to the site is also associated with the site's segment.

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

In the previous example, the HR segment is compatible with both Sales and Research. However, because the Sales and Research segments are incompatible, you can't associate them with the same site.

Prerequisites

1. Make sure you meet the licensing requirements for Information Barriers.
2. Create IB policies that allow or block communication between the segments, and then set them to active. Create segments and define the users in each segment.
3. Wait 24 hours after configuring and activating your IB policies for the changes to propagate through your organization.
4. Complete the steps in the following sections to enable and manage SharePoint and OneDrive Information Barriers in your organization.

Enable SharePoint and OneDrive Information Barriers in your organization

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview **roles and permissions**.

SharePoint Administrators or Global Administrators can enable Information Barriers in SharePoint and OneDrive in your organization. Complete the following steps to enable Information Barriers for your organization:

1. You must use the latest version of the SharePoint Online Management Shell. If you installed a previous version of the SharePoint Online Management Shell, go to **Add or remove programs** and uninstall *SharePoint Online Management Shell*. Then, install the latest version. To learn more, see Get started with SharePoint Online Management Shell.

2. Connect to SharePoint Online as a Global Administrator or SharePoint Administrator in Microsoft 365. To learn how, see Getting started with SharePoint Online Management Shell.

3. Run the following command to enable Information Barriers in SharePoint and OneDrive:

   PowerShell

```
Set-SPOTenant -InformationBarriersSuspension $false
```

4. Wait for approximately 1 hour for the changes to take effect after you enable Information Barriers for SharePoint and OneDrive in your organization.

> **Note**
>
> If you enabled Information Barriers for SharePoint in your organization before March 15, 2022, the default access and sharing control for Implicit mode for Microsoft Teams-connected sites are based on the segments associated with the site.

To enable Microsoft 365 group-membership based access and sharing control for all Implicit mode Teams-connected sites in your tenant, run the following command:

PowerShell

```
Set-SPOTenant -IBImplicitGroupBased $true
```

> **Note**
>
> If you have Microsoft 365 Multi-Geo, run this command for each of your geolocations.

View and manage segments as an administrator

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

SharePoint Administrators or Global Administrators can view and manage segments on a SharePoint site. Your organization can have up to 5,000 segments, and users can be assigned to multiple segments.

> **Important**
>
> Support for 5,000 segments and assigning users to multiple segments is only available when your organization isn't in *Legacy* mode. Assigning users to multiple segments

requires extra steps to change the Information Barriers mode for your organization. For more information, see Use multi-segment support in Information Barriers.

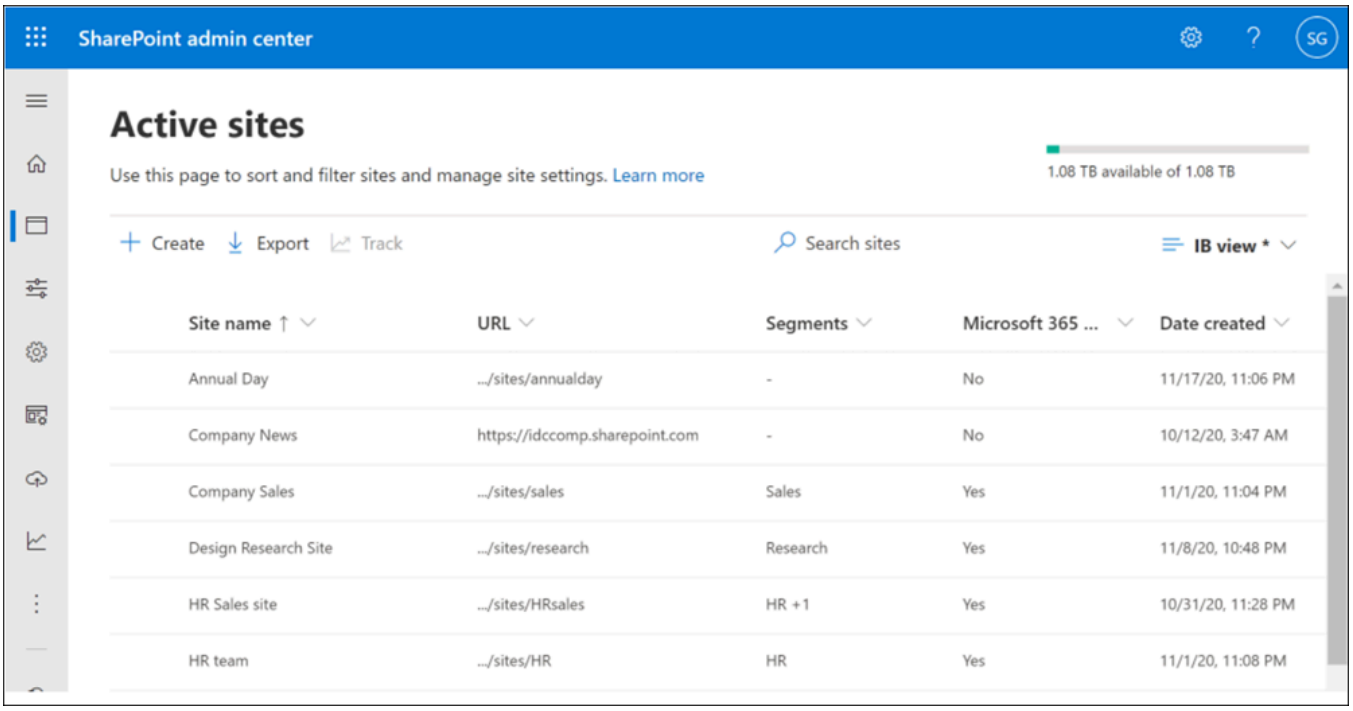
For organizations in *Legacy* mode, the maximum number of segments supported is 250, and users are restricted to being assigned to only one segment. Organizations in *Legacy* mode are eligible to upgrade to the newest version of Information Barriers in the future. For more information, see the Information Barriers roadmap.

View and manage Information Barriers segments as follows:

1. Use the SharePoint admin center to view and manage information segments

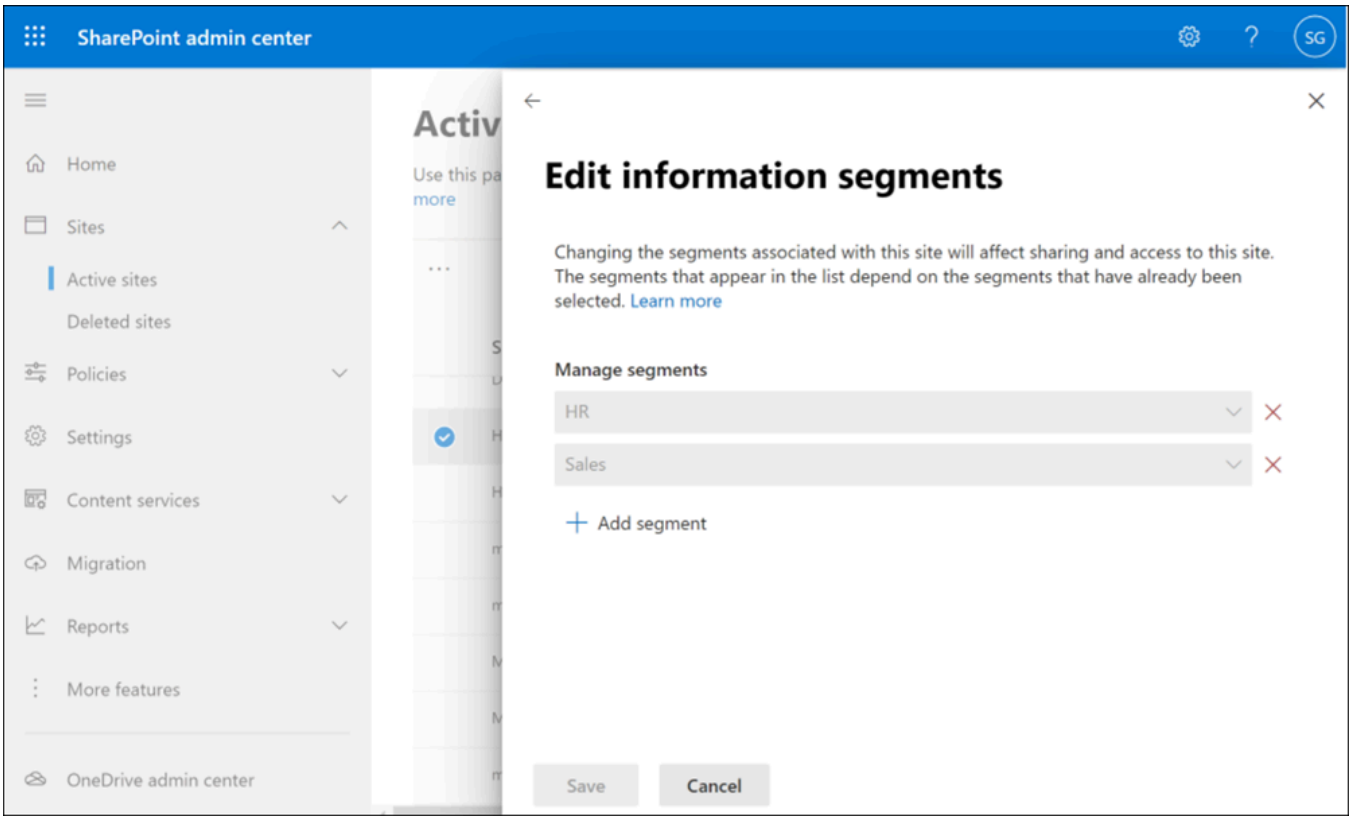
To view, edit, or remove information segments for a site, use **Active sites** in the SharePoint admin center.

The Segments column lists the first segment associated with the site and shows whether the site has other segments associated. Learn how to show or move this column.



To view the complete list of segments associated with a site, select the site name to open the details panel, and then select the **Settings** tab.

To edit the segments associated with the site, select **Edit**, add or remove segments, and then select **Save**.



2. Use SharePoint PowerShell to view and manage information segments on a site

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

1. Connect to Security & Compliance PowerShell as a Global Administrator.
2. Run the following command to get the list of segments and their GUIDs.

```
PowerShell
Get-OrganizationSegment | ft Name, EXOSegmentID
```

3. Save the list of segments.

Expand table

Name	EXOSegmentId
Sales	a9592060-c856-4301-b60f-bf9a04990d4d

Name	EXOSegmentId
Research	27d20a85-1c1b-4af2-bf45-a41093b5d111
HR	a17efb47-e3c9-4d85-a188-1cd59c83de32

4. You must use the latest version of the SharePoint Online Management Shell. If you installed a previous version of the SharePoint Online Management Shell, go to **Add or remove programs** and uninstall *SharePoint Online Management Shell*. Then, install the latest version. To learn more, see Get started with SharePoint Online Management Shell.

5. Connect to SharePoint Online as a Global Administrator or SharePoint Administrator in Microsoft 365. To learn how, see Getting started with SharePoint Online Management Shell.

6. Run the following command:

   PowerShell

   ```
   Set-SPOSite -Identity <site URL> -AddInformationSegment <segment GUID>
   ```

   For example:

   PowerShell

   ```
   Set-SPOSite -Identity https://contoso.sharepoint.com/sites/ResearchTeamSite -
   AddInformationSegment 27d20a85-1c1b-4af2-bf45-a41093b5d111
   ```

You see an error message if you attempt to associate a segment that isn't compatible with the site's existing segments.

> **Note**
>
> When you add a segment to a site, the site's IB mode is automatically updated to *Explicit*.

To remove a segment from a site, run the following command:

PowerShell

```
Set-SPOSite -Identity <site URL> -RemoveInformationSegment <segment GUID>
```

For example:

PowerShell

```
Set-SPOSite -Identity https://contoso.sharepoint.com/sites/ResearchTeamSite -
RemoveInformationSegment 27d20a85-1c1b-4af2-bf45-a41093b5d111
```

> **Note**
>
> When you remove all segments from a site, the site's IB mode is automatically updated to *Open*.

To view the segments of a site, run the following command to return the GUIDs of any segments associated with the site.

PowerShell

```
Get-SPOSite -Identity <site URL> | Select InformationSegment
```

3. Use the SharePoint REST API to view and manage information segments on a site

SharePoint includes a Representational State Transfer (REST) service that you can use to manage segments on a site. To access SharePoint resources and manage site segments by using REST, construct a RESTful HTTP request by using the OData standard. This request corresponds to the desired client object model application programming interface (API).

For more information about the SharePoint REST service, see Get to know the SharePoint REST service.

View and manage IB modes as an administrator with SharePoint PowerShell

To view the IB mode of a site, run the following command:

PowerShell

```
Get-SPOSite -Identity <site URL> | Select InformationBarriersMode
```

Owner Moderated mode scenario

You want to allow a Sales and Research user to collaborate on a SharePoint site in the presence of an HR user.

Owner Moderated is a mode applicable to sites (Teams-connected sites, non-group connected sites) that grants incompatible segment users access to the site. Only the site owner can invite incompatible segment users to the same site.

To update a site's mode to *Owner Moderated*, run the following PowerShell command:

PowerShell

```
Set-SPOSite -Identity <siteurl> -InformationBarriersMode OwnerModerated
```

You can't set the Owner Moderated IB mode on a site with segments. Remove the segments first before setting IB mode as Owner Moderated. Users who have site access permissions can access an Owner Moderated site. Only the site owner can share an Owner Moderated site and its contents according to their IB policy.

Auditing

You can view audit events in the Microsoft Purview portal to monitor IB activities. The system logs audit events for the following activities:

- Enabling Information Barriers for SharePoint and OneDrive
- Applying a segment to a site
- Changing the segment of a site
- Removing the segment of a site
- Applying Information Barriers mode to a site
- Changing Information Barriers mode of a site
- Disabling Information Barriers for SharePoint and OneDrive

For more information about SharePoint segment auditing in Microsoft 365, see Search the audit log in the Microsoft Purview portal.

Site creation and management by site owners

When a segmented user creates a SharePoint site, the site associates with the user's segment and the site's Information Barriers mode automatically sets to *Explicit*.

Site owners can add more segments to a SharePoint site that already has segments with the site's mode set as *Explicit*. Site owners can't remove added segments from sites. SharePoint Administrators need to remove added segments in your organization if needed.

When a non-segmented user creates a SharePoint site, the site doesn't associate with any segment and the site's Information Barriers mode automatically sets to *Open*.

When a SharePoint Administrator creates a SharePoint site from the SharePoint admin center, the site doesn't associate with any segment and the site's IB mode sets to *Open*.

To help site owners add a segment to a site, share the Associate information segments with SharePoint sites article with your SharePoint site owners.

Microsoft Teams sites

When you create a team in Microsoft Teams, you also automatically create a SharePoint site for the team's files. To protect the Microsoft Teams sites by using Information Barriers control, enable Information Barriers in SharePoint for your tenant.

Within 24 hours, the site's Information Barriers mode is automatically set to *Implicit* and segments associated with the team's members are associated with the site.

Microsoft Teams sites with the IB mode set to *Implicit* have site access and sharing based on Microsoft 365 group membership.

For example, users have access to the Microsoft Teams site if they're members of the Microsoft 365 group connected to the site. The Microsoft 365 group connected to the Team is IB compliant.

> **Note**
>
> If you enabled Information Barriers for SharePoint in your organization before March 15, 2022, the Teams-connected site's access and sharing is based on the segments of the site. For example:

- You can share the site and its content with users whose segment matches that of the site.
- Users can access the site and its content if they have the same segment as that of the site and have site access permissions.

To enable Microsoft 365 group membership-based access and sharing control for all *Implicit* mode sites in your organization, see Enable SharePoint and OneDrive Information Barriers in your organization.

Private channel and Information Barriers

When you enable SharePoint Information Barriers in your organization, any new private channel site automatically inherits its parent Microsoft Team's IB mode within 24 hours. The mode for a private channel is assigned as follows:

Expand table

Parent Team's IB mode	Private channel site's IB mode
Open	Open
Implicit or Owner Moderated	Implicit

The private channel site's IB mode governs access and sharing for the private channel site:

- Private channel site with *Open* Information Barriers mode
  - Access anyone with site access permissions
  - Sharing links are allowed per the site's existing sharing policy
  - People picker allows discoverability of user per the sharer's IB policy
- Private channel site with *Implicit* Information Barriers mode
  - Access only to users who are currently members of the private channel
  - Sharing by using **People with existing access link**

Private channel sites already configured in your organization have their Information Barriers mode set as *Open*. To configure existing private channel sites to *Implicit* mode, run the following cmdlet in the SharePoint PowerShell module:

PowerShell

```
Set-Sposite -Identity <site URL> -InformationBarriersMode Implicit
```

Learn more about managing Microsoft Teams connected teams sites.

Search

Users see search results from:

- **Segment associated sites**: When the site's segment matches the user's segment and the user has site access permission. For example, a site with *Explicit* mode.
- **Non-segmented sites**: When the user has existing access to the content or site. For example, sites with *Open*, *Owner Moderated*, or *Implicit* mode. When the user selects the search result to open the content in the site, the user is denied access if they don't match the site's IB policy.

Tenant wide search and Copilot experience

Users experience is as follows:

- **Segment associated sites**: When the site's segment matches the user's segment and the user has site access permission. For example, a site with *Explicit* mode.
- **Open mode sites:** When the user has existing access to the content or site.
- **Implicit and Owner moderated mode sites:**
  - When the user has existing access to the content or site.
  - If the user had access to the site and content prior to the policy application, they continue to see the site and its contents in search and Copilot results. However, when attempting to open the content, they're denied access if they don't comply with the site's IB policy.

Effects of changes to user segments

If a SharePoint site owner or site member's segment changes, they retain access to the site or content according to the site's IB mode:

- **Open mode**: User can access the site if they have existing site access permissions.
- **Owner Moderated**: User can access the site if they have existing site access permissions.
- **Implicit Mode**: If the user is a member of the Microsoft 365 group, they retain access to the site.
- **Explicit Mode**: If the user's new segment matches the site's segment and user has site access permissions, they retain access to the site.

Effects of changes to existing IB policies

If a compliance administrator changes an existing IB policy, the change might affect the compatibility of the segments associated with a site (in *Explicit* or *Implicit* mode*). For example, segments that were once compatible might no longer be compatible.

By using the Information Barriers policy compliance report, the SharePoint Administrator can view the list of sites where segments are no longer compatible. For more information, see Learn how to create an Information Barriers policy compliance report in PowerShell.

To manage out of compliance sites:

- In *Explicit* mode, a SharePoint Administrator must change the associated segments to bring them into IB compliance.
- In *Implicit* mode, a SharePoint Administrator can't manage segments directly. The Teams admin should manage the Team's membership to bring the Teams membership roster and segments into IB compliance.

How to suspend SharePoint and OneDrive Information Barriers in your organization

If your organization wants to temporarily suspend Information Barriers on SharePoint, use SharePoint Online Management Shell and the Set-Spotenant cmdlet.

To suspend Information Barriers, run the following command:

```
PowerShell
Set-SPOTenant -InformationBarriersSuspension $true
```

> **Note**
>
> If you have Microsoft 365 Multi-Geo, run this command for each of your geo-locations.

Allow sharing of Open mode sites with mail-enabled security groups

IB supports an opt-in capability available in the SharePoint PowerShell module for sites in *Open* mode to be shared with mail-enabled security groups for site permissions, sharing, and audience targeting. This capability is only supported in *Open* mode sites. SharePoint admins can enable this support in your organization. Ensure the security group membership is IB compliant.

Before enabling group support, verify that you meet the following prerequisites:

- Your organization has only IB *Block* policies.
- Your organization is enabled for SharePoint IB (see this section in this article).

To configure mail-enabled security group support in *Open* mode sites, run the following command:

```
PowerShell
Set-SPOTenant -ShowPeoplePickerGroupSuggestionsForIB $true
```

Resources

- Learn about Information Barriers

- Get started with Information Barriers
- Manage Information Barriers policies
- Information Barriers in Microsoft Teams
- Information Barriers in OneDrive

Last updated on 03/30/2026

Use Information Barriers with OneDrive

Microsoft Purview Information Barriers are policies that prevent users from communicating and collaborating with each other. This article covers IB behavior in OneDrive.

For OneDrive, Information Barriers can determine and prevent the following kinds of unauthorized collaborations:

- User access to OneDrive or stored content
- Sharing OneDrive or stored content with other users

Information Barriers modes and OneDrive

When you enable Information Barriers on SharePoint and OneDrive, you automatically protect the OneDrive of segmented users with IB policies. Information Barriers modes help strengthen access, sharing, and membership of a OneDrive site based on its IB mode and segments associated with the OneDrive.

When you use Information Barriers with OneDrive, the following IB modes apply:

- **Open**: The OneDrive has no segments. This mode is the default when a non-segmented user provisions their OneDrive.
- **Owner Moderated**: The OneDrive is used for collaboration with incompatible users in the presence of the site owner/moderator. See this section for details.
- **Explicit**: The user's segment and other compatible segments are associated with the OneDrive. This mode is the default when a segmented user provisions their OneDrive within 24 hours of enablement.
- **Mixed**: The segmented user's OneDrive is allowed to be shared with unsegmented users. This mode is an opt-in mode that the SharePoint admin can set on a OneDrive of a segmented user.

> ⓘ **Note**
>
> Starting July 12, 2022, *Inferred* mode changed to *Mixed* mode. The functionality for the mode remains the same.

Open mode

Open mode sharing behavior

When a OneDrive has no segments and IB mode as *Open*:

- The user can share files and folders based on the IB policy applied to the user and the sharing setting for the OneDrive.

Open mode access behavior

For a user to access content in a OneDrive with no segments and IB mode as *Open*:

- The files must be shared with the user.

Owner Moderated mode

Owner Moderated sharing behavior

When a OneDrive has Information Barriers mode set to *Owner Moderated*:

- The option to share with *Anyone with the link* is disabled.
- The option to share with *Company-wide link* is disabled.
- You can share the site and its content with existing members.
- Only the OneDrive owner can share the site and its content, following their IB policy.

Owner Moderated access behavior

For a user to access a OneDrive with Information Barriers mode set to *Owner Moderated*:

- The user has site access permissions.

Explicit mode

Explicit mode sharing behavior

When a OneDrive has Information Barriers segments and the mode is set to *Explicit*:

- The option to share with *Anyone with the link* is disabled.
- The option to share with *Company-wide link* is disabled.
- You can share files and folders only with users whose segment matches that of the OneDrive.

Explicit mode access behavior

For a user to access content in a OneDrive with segments and IB mode set to *Explicit*:

1. The user's segment matches a segment associated with the OneDrive.

   AND

2. The files are shared with the user.

> **Note**
>
> By default, non-segment users can access shared OneDrive files only from other non-segment users with IB modes set to *Open*. They can't access shared files from OneDrive that have segments applied and the IB mode is *Explicit*.

Mixed mode

Mixed mode sharing behavior

When a OneDrive uses Information Barriers segments and the mode is set to *Mixed*:

- The option to share with *Anyone with the link* is disabled.
- The option to share with *Company-wide link* is disabled.
- You can share files and folders with users whose segment matches that of the OneDrive and unsegmented users in the tenant.

Mixed mode access behavior

For a segmented user to access content in a OneDrive with segments and IB mode set to *Mixed*:

1. The user's segment matches a segment associated with the OneDrive.

   AND

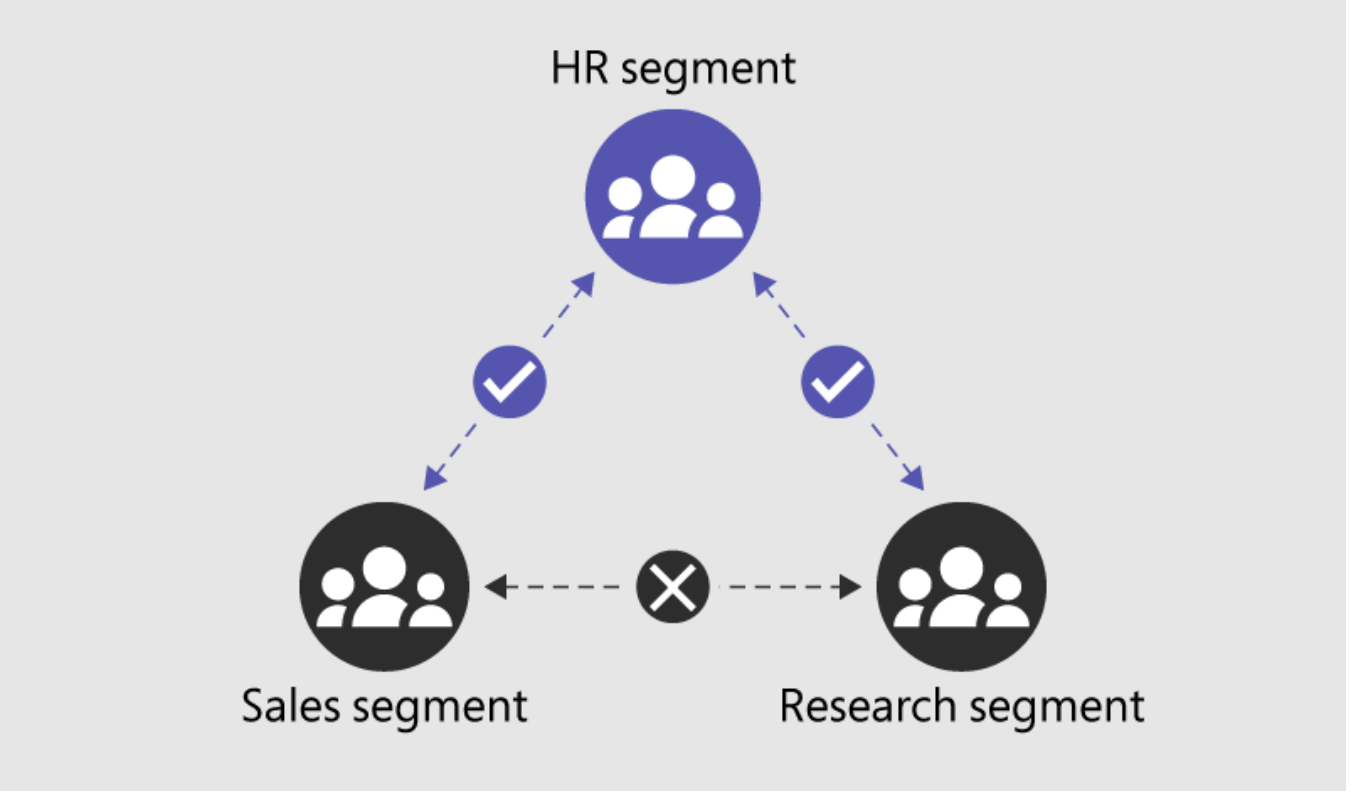
2. The files are shared with the user.

For an unsegmented user to access content in a OneDrive with segments and IB mode set to *Mixed*:

- The user has site access permissions.

Example scenario

The following example illustrates three segments in an organization: HR, Sales, and Research. You define an IB policy that blocks communication and collaboration between the Sales and Research segments.



By using Information Barriers in OneDrive, when you apply a segment to a user, the segment automatically associates with the user's OneDrive within 24 hours. The OneDrive also associates with other segments that are compatible with the user's segment and with each other. A OneDrive can have up to 100 associated segments. A global or SharePoint admin can manage these segments by using PowerShell, as described later in the section Associate or remove additional segments on a user's OneDrive.

The following table shows the effects of this example configuration:

Expand table

Components	HR users	Sales users	Research users	Non-segment users
Segments associated with OneDrive	HR	Sales, HR	Research, HR	None
IB mode on OneDrive	Explicit	Explicit	Explicit	Open
OneDrive content can be shared with	HR only	Sales and HR	Research and HR	Anyone based on the sharing settings selected
OneDrive content can be accessed by	HR only	Sales and HR	Research and HR	Anyone with whom the content is shared

Enable SharePoint and OneDrive Information Barriers in your organization

You configure enabling Information Barriers for SharePoint and OneDrive in a single action. You can't enable Information Barriers for the services separately. For more information, see Enable SharePoint and OneDrive Information Barriers in your organization. After you enable Information Barriers for SharePoint and OneDrive, continue with the OneDrive guidance in this article.

Prerequisites

1. Make sure you meet the licensing requirements for Information Barriers.
2. Create IB policies that allow or block communication between the segments and activate the policies. Create segments and define the users in each.
3. After you configure and activate your IB policies, wait 24 hours for the changes to propagate through your organization.
4. Enable Information Barriers for OneDrive. You configure enabling Information Barriers for SharePoint and OneDrive in a single action. You can't enable these services separately. For more information, see the guidance and steps in the Use Information Barriers with SharePoint article.
5. Complete the steps in the following sections to customize and manage Information Barriers for OneDrive in your organization.

Use PowerShell to view the segments associated with a OneDrive

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

A global admin or SharePoint admin can view and change the segments associated with a user's OneDrive. Your organization can have up to 5,000 segments, and users can be assigned to multiple segments.

> **Important**

Support for 5,000 segments and assigning users to multiple segments is only available when your organization isn't in *Legacy* mode. Assigning users to multiple segments requires extra steps to change the Information Barriers mode for your organization. For more information, see Use multi-segment support in Information Barriers.

For organizations in *Legacy* mode, the maximum number of segments supported is 250, and users are restricted to being assigned to only one segment. Organizations in *Legacy* mode are eligible to upgrade to the newest version of Information Barriers in the future. For more information, see the Information Barriers roadmap.

1. Connect to Security & Compliance PowerShell as a global admin.
2. Run the following command to get the list of segments and their GUIDs.

```
PowerShell
Get-OrganizationSegment | ft Name, EXOSegmentID
```

3. Save the list of segments.

Expand table

Name	EXOSegmentId
Sales	a9592060-c856-4301-b60f-bf9a04990d4d
Research	27d20a85-1c1b-4af2-bf45-a41093b5d111
HR	a17efb47-e3c9-4d85-a188-1cd59c83de32

4. You must use the latest version of the SharePoint Online Management Shell. If you installed a previous version of the SharePoint Online Management Shell, go to **Add or remove programs** and uninstall *SharePoint Online Management Shell*. Then, install the latest version. To learn more, see Get started with SharePoint Online Management Shell.
5. Connect to SharePoint as a global admin or SharePoint admin in Microsoft 365. To learn how, see Getting started with SharePoint Online Management Shell.
6. Run the following command:

```
PowerShell
Get-SPOSite -Identity <site URL> | Select InformationSegment
```

For example:

PowerShell

```
Get-SPOSite -Identity https://contoso-
my.sharepoint.com/personal/John_contoso_onmicrosoft_com | Select
InformationSegment
```

Manage segments on a user's OneDrive

> ⚠ **Warning**
>
> If the segments associated with a user's OneDrive don't match the segment you apply to the user, the user can't access their OneDrive. Don't associate any segments with the OneDrive of a non-segment user.

> ⓘ **Note**
>
> Any changes you make are overwritten if the user's segment changes.

To associate a segment with a OneDrive, run the following command in the SharePoint Online Management Shell.

PowerShell

```
Set-SPOSite -Identity <site URL> -AddInformationSegment <segment GUID>
```

For example:

PowerShell

```
Set-SPOSite -Identity https://contoso-
my.sharepoint.com/personal/John_contoso_onmicrosoft_com -AddInformationSegment
27d20a85-1c1b-4af2-bf45-a41093b5d111
```

When you add segments to a OneDrive, the site's IB mode automatically updates to *Explicit*. An error appears if you attempt to associate a segment that isn't compatible with the existing segments on the OneDrive.

To remove a segment from a OneDrive, run the following command.

PowerShell

```
Set-SPOSite -Identity <site URL> -RemoveInformationSegment <segment GUID>
```

For example:

PowerShell

```
Set-SPOSite -Identity https://contoso-
my.sharepoint.com/personal/John_contoso_onmicrosoft_com -RemoveInformationSegment
27d20a85-1c1b-4af2-bf45-a41093b5d111
```

If you remove all the segments of a OneDrive site, the IB mode of the OneDrive automatically updates to *Open*.

Manage the IB mode of a user's OneDrive (preview)

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

To view the IB mode of a OneDrive site, run the following command in the SharePoint Online Management Shell as a SharePoint admin or global administrator:

PowerShell

```
Get-SPOSite -Identity <site URL> | Select InformationBarriersMode
```

For example:

PowerShell

```
Get-SPOSite -Identity https://contoso-
my.sharepoint.com/personal/John_contoso_onmicrosoft_com | Select
InformationBarriersMode
```

A SharePoint admin or global administrator can also manage the IB mode of a OneDrive site to meet the needs of your organization by using new IB modes:

Owner Moderated mode example

Allow an incompatible segment user access to a OneDrive. For example, you want to allow HR user's OneDrive to be accessed by both Sales and Research segment users in your tenant.

Owner Moderated is a mode applicable to OneDrive sites that allows incompatible segment users to access OneDrive in the presence of a moderator/owner. Only the site owner has the capability to invite incompatible segment users on the same site.

To update a OneDrive site IB mode to *Owner Moderated*, run the following PowerShell command:

PowerShell

```
Set-SPOSite -Identity <siteurl> -InformationBarriersMode OwnerModerated
```

You can't set the Owner Moderated IB mode on a site with segments. Remove the segments before setting the IB mode as Owner Moderated. Users with site access permissions can access an Owner Moderated site. Only the site owner can share an Owner Moderated OneDrive and its contents per their IB policy.

Mixed mode example

Allow unsegmented users to access OneDrive associated with segments. For example, you want to allow HR users' OneDrive to be accessed by HR segment and unsegmented users in your tenant. Mixed mode applies to OneDrive sites that allow segmented and unsegmented users to access OneDrive.

To update a OneDrive site's IB Mode to Mixed, run the following PowerShell command:

PowerShell

```
Set-SPOSite -Identity <siteurl> -InformationBarriersMode Mixed
```

You can't set mixed IB mode on a site without segments. Add segments before setting the IB mode as mixed.

Effects of changes to user segments

If a user's segment changes, the OneDrive segment and IB mode automatically update within 24 hours as described in the OneDrive Information Barriers section.

Example 1: User's segment updated from Research to Sales, the user's OneDrive updates as follows within 24 hours:

- Segment: Sales, HR
- IB mode: *Explicit*

Example 2: User's segment updated from HR to None

- Segment: None
- IB mode: *Open*

Effects of changes to IB policies

If a compliance administrator changes an existing policy, the change might affect the compatibility of the segments associated with the OneDrive.

For example, segments that were once compatible might no longer be compatible. A SharePoint admin must change the segments associated with an affected site. For more information, see create an Information Barriers policy compliance report in PowerShell.

If you change a policy after users share files, the sharing links work only if the user attempting to access the shared files has a segment applied that matches a segment associated with the OneDrive.

Auditing

Audit events are available in the Microsoft Purview portal to help you monitor IB activities. The system logs audit events for the following activities:

- Enabling Information Barriers for SharePoint and OneDrive
- Applying segment to site
- Changing segment of site
- Removing segment of site
- Applying Information Barriers mode to site
- Changing Information Barriers mode of site
- Disabling Information Barriers for SharePoint and OneDrive

For more information about OneDrive segment auditing in Microsoft 365, see Search the audit log.

Resources

- Learn about Information Barriers
- Get started with Information Barriers
- Manage Information Barriers policies

- Information Barriers in Microsoft Teams
- Information Barriers in SharePoint

Last updated on 03/30/2026

SharePoint and OneDrive insights reports (preview)

The insights report in Information Barriers (IB) helps you identify and discover usage patterns across SharePoint sites and OneDrive accounts in your organization. As an administrator, you can use PowerShell to create and view reports that identify top sites and their modes, so you can apply suitable controls for the sites.

The insights report provides the following information for the top 100 most actively used SharePoint sites and OneDrive accounts in your organization:

Expand table

Report section	Description	Applies to
Explicit mode section	Top restrictive sites with highest collaboration between IB users	SharePoint sites and OneDrive accounts
Implicit mode section	Top sites with highest collaboration between compatible users	SharePoint sites only
Mixed mode section	Top accounts with collaboration between segmented and unsegmented users	OneDrive accounts only
Mode Distribution section	The number and percentage of modes across all sites and accounts	SharePoint sites and OneDrive accounts
Open mode section	Top least restrictive sites	SharePoint sites and OneDrive accounts
Owner Moderated mode section	Top sites with highest collaboration between IB users and non-IB users	SharePoint sites and OneDrive accounts

Prerequisites

- Your organization must have Information Barriers enabled for SharePoint and OneDrive to create the insights report. To enable IB in SharePoint and OneDrive for your organization, see Use Information Barriers with SharePoint.
- You must use the latest version of the SharePoint Online Management Shell. If you installed a previous version of the SharePoint Online Management Shell, go to **Add or remove programs** and uninstall *SharePoint Online Management Shell*. Then, install the latest version. To learn more, see Get started with SharePoint Online Management Shell.

Create the insights report

Before you can run specific queries for Information Barriers details, you must first create a report. After the report finishes successfully, you can run specific detail reports.

The system supports one insights report for every 24-hour period. When you create the next report, it overwrites the previous report. For example, if you create an insights report on April 25, 2023, at 4:10 PM, you can't create another insights report until after 4:10 PM on April 26, 2023.

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

To create the insights report, complete the following steps:

1. Connect to SharePoint Online as a global administrator or SharePoint administrator in Microsoft 365.
2. Run the following PowerShell cmdlet and accept the confirmation prompt to create the insights report:

   ```PowerShell
   Start-SPOInformationBarriersInsightsReport
   ```

   > **Note**
   >
   > Depending on the number of SharePoint sites and OneDrive accounts in your organization, it might take up to one hour to create this report.

   You can automate acceptance of the confirmation prompt when creating the insights report by appending the *-Yes* parameter to the cmdlet. For example, `Start-SPOInformationBarriersInsightsReport -Yes`.

3. Run the following PowerShell cmdlet to view the status of the insights report:

   ```PowerShell
   Get-SPOInformationBarriersInsightsReport
   ```

The following example shows the information returned for the insights report:

```text
State: Completed
Id: ec65a1cf-9b1a-48c2-a1b4-f923ac4c0776
StartTimeInUtc: 4/25/2023 4:10:16 PM
CompleteTimeInUtc: 4/25/2023 4:10:25 PM
QueuedTimeInUtc: 4/25/2023 4:06:47 PM
```

The report is ready to view or download when the *State* value is *Completed*. Other *State* values include:

- *Not Started*: State when the insights report hasn't started.
- *In Progress*: State when the insights report is in progress.
- *Error*: State when the insights report failed.

Insights reports for SharePoint sites

View a summary of modes with results for SharePoint sites

To view a summary of the modes with results for SharePoint sites, use the following cmdlet syntax to view insights from the report:

```PowerShell
Get-SPOInformationBarriersInsightsReport -reportId <ID>
```

Example:

```PowerShell
Get-SPOInformationBarriersInsightsReport -reportId ec65a1cf-9b1a-48c2-a1b4-
f923ac4c0776
```

Example results:

```text
Content: Explicit, Implicit, Open, OwnerModerated, ModeDistribution
State: Completed
Id: ec65a1cf-9b1a-48c2-a1b4-f923ac4c0776
StartTimeInUtc: 4/25/2023 4:10:16 PM
```

```
CompleteTimeInUtc: 4/25/2023 4:10:25 PM
QueuedTimeInUtc: 4/25/2023 4:06:47 PM
```

In this example, the insights report results are displayed for SharePoint sites included in the organization with an ID of *ec65a1cf-9b1a-48c2-a1b4-f923ac4c0776*. The values in the *Content* line represent the modes that have results in the report. If a mode (applicable to SharePoint) isn't listed, there aren't any SharePoint sites in the organization with that mode.

View the details for a specific mode for SharePoint sites

To view details about a specific mode with results for SharePoint sites, use the following cmdlet syntax to view insights from the insights report:

PowerShell

```
Get-SPOInformationBarriersInsightsReport -reportId <ID> -section <Mode>
```

Example:

PowerShell

```
Get-SPOInformationBarriersInsightsReport -reportId ec65a1cf-9b1a-48c2-a1b4-
f923ac4c0776 -section Explicit
```

Example results:

text

```
SiteName: Contoso Budget Planning
SiteURL: https://contoso.sharepoint.com/sites/ContosoBudgetPlanning
Site Owner: User1
LastActivity: 4/25/2023 4:10:16 PM
IBMode: Explicit

SiteName: Contoso Training Budgets
SiteURL: https://contoso.sharepoint.com/sites/ContosoTrainingBudgets
Site Owner: User2
LastActivity: 4/25/2023 4:10:16 PM
IBMode: Explicit

SiteName: Contoso Viewpoint Project
SiteURL: https://contoso.sharepoint.com/sites/ContosoViewpointProject
Site Owner: User3
LastActivity: 4/25/2023 4:10:16 PM
IBMode: Explicit
```

View the details for the mode distribution for SharePoint sites

To view details about the mode distribution with results for SharePoint sites, use the following cmdlet syntax to view insights from the insights report:

```PowerShell
Get-SPOInformationBarriersInsightsReport -reportId <ID> -section <ModeDistribution>
```

Example:

```PowerShell
Get-SPOInformationBarriersInsightsReport -reportId ec65a1cf-9b1a-48c2-a1b4-f923ac4c0776 -section ModeDistribution
```

Example results:

```text
Total site count: 10000
Owner Moderated SiteCount: 500 SitePercentage: 5
Open SiteCount: 1682 SitePercentage: 16.82
Explicit SiteCount: 3628 SitePercentage: 36.82
Implicit SiteCount: 4190 SitePercentage: 41.9
```

Insights reports for OneDrive accounts

View a summary of modes with results for OneDrive accounts

To view a summary of the modes with results for OneDrive accounts, use the following cmdlet syntax to view insights from the insights report:

```PowerShell
Get-SPOInformationBarriersInsightsReport -reportId <ID> -service <OneDrive>
```

Example:

```PowerShell
Get-SPOInformationBarriersInsightsReport -reportId ec65a1cf-9b1a-48c2-a1b4-f923ac4c0776 -service OneDrive
```

Example results:

```text
Content: Explicit, Mixed, Open, OwnerModerated, ModeDistribution
State: Completed
Id: ec65a1cf-9b1a-48c2-a1b4-f923ac4c0776
StartTimeInUtc: 4/25/2023 4:10:16 PM
CompleteTimeInUtc: 4/25/2023 4:10:25 PM
QueuedTimeInUtc: 4/25/2023 4:06:47 PM
```

In this example, you see the insights report results for OneDrive accounts included in the organization with an ID of *ec65a1cf-9b1a-48c2-a1b4-f923ac4c0776*. The values in the *Content* line represent the modes that have results in the report. If a mode (applicable to OneDrive) isn't listed, there aren't any OneDrive accounts in the organization with that mode.

View the details for a specific mode for OneDrive accounts

To view details about a specific mode with results for OneDrive accounts, use the following cmdlet syntax to view insights from the insights report:

```PowerShell
Get-SPOInformationBarriersInsightsReport -reportId <ID> -service OneDrive -section
<Mode>
```

Example:

```PowerShell
Get-SPOInformationBarriersInsightsReport -reportId ec65a1cf-9b1a-48c2-a1b4-
f923ac4c0776 -service OneDrive -section Open
```

Example results:

```text
SiteName: User1
SiteURL: https://spdfcontoso-
my.sharepoint.com/personal/user1_spdfcontoso_onmicrosoft_com
Site Owner: user1@spdfcontoso.onmicrosoft.com
LastActivity: 4/25/2023 4:10:16 PM
IBMode: Open

SiteName: User2
SiteURL: https://spdfcontoso-
my.sharepoint.com/personal/user2_spdfcontoso_onmicrosoft_com
```

```
Site Owner: user2@spdfcontoso.onmicrosoft.com
LastActivity: 4/25/2023 4:10:16 PM
IBMode: Open

SiteName: User3
SiteURL: https://spdfcontoso-
my.sharepoint.com/personal/user3_spdfcontoso_onmicrosoft_com
Site Owner: user3@spdfcontoso.onmicrosoft.com
LastActivity: 4/25/2023 4:10:16 PM
IBMode: Open
```

View the details for the mode distribution for OneDrive accounts

To view details about the mode distribution with results for OneDrive accounts, use the following cmdlet syntax to view insights from the insights report:

PowerShell

```
Get-SPOInformationBarriersInsightsReport -reportId <ID> -service OneDrive -section
<ModeDistribution>
```

Example:

PowerShell

```
Get-SPOInformationBarriersInsightsReport -reportId ec65a1cf-9b1a-48c2-a1b4-
f923ac4c0776 -service OneDrive -section ModeDistribution
```

Example results:

text

```
Total site count: 19305
Owner Moderated SiteCount: 167 SitePercentage: 0.87
Mixed SiteCount: 4881 SitePercentage: 25.28
Explicit SiteCount: 4910 SitePercentage: 25.43
Open SiteCount: 9347 SitePercentage: 48.42
```

Download the results of the insights report

To download the full results of a specific insights report for SharePoint sites, use the following cmdlet syntax:

```
```

PowerShell

```
Get-SPOInformationBarriersInsightsReport -reportId <ID> -section <report you want to download> -action Download -FullDetails
```

The report automatically exports as a .csv file with a file name of *SharePoint_the name of the report type*. For example, if you download the *Explicit* report, the file name is *SharePoint_Explicit.csv*. The *-FullDetails* property supports downloading up to 1 million records for SharePoint sites.

To download the results of a specific insights report for OneDrive accounts, use the following cmdlet syntax:

PowerShell

```
Get-SPOInformationBarriersInsightsReport -reportId <ID> -service OneDrive -section <report you want to download> -action Download
```

The report automatically exports as a .csv file with a file name of *OneDrive_the name of the report type*. For example, if you download the *Explicit* report, the file name is *OneDrive_Explicit.csv*.

To rename the report or export to custom locations, use standard PowerShell parameters. For more information, see the PowerShell documentation.

Audit activities for the insights report

You can view audit events for the insights report in the Microsoft Purview portal. For more information about the activities logged, see Information Barriers activities.

Resources

- Learn about Information Barriers
- Get started with Information Barriers
- Information Barriers in SharePoint
- Information Barriers in OneDrive

Last updated on 03/30/2026

Create an Information Barriers policy compliance report

If a compliance administrator changes an existing Information Barriers policy, the change might affect the compatibility of segments already associated with a site.

For example, a policy might allow communication and collaboration between the Sales and Research segments. Later, the policy might not allow communication and collaboration between these segments. The segments are incompatible and shouldn't be associated with the same site.

By using the SharePoint Information Barriers policy compliance report, SharePoint administrators can view the list of sites that don't comply with existing policies. The report covers these sites:

- Microsoft 365 group-connected team sites that aren't connected to Microsoft Teams
- Communication sites
- Modern team sites that aren't connected to Microsoft 365 groups
- OneDrive
- User-owned SharePoint Embedded containers

The report displays the list of sites that don't comply with the existing policies that you recently updated. For each noncompliant site, it shows compatible segments, incompatible segments, and invalid segments (those segments that no longer exist).

If a OneDrive or user-owned SharePoint Embedded container doesn't comply, this report lets you update its segments to be compliant with the latest IB policies in your organization.

> **Note**
>
> You only need to run this report if you change Information Barriers policies. Depending on the number of sites in your organization, it can take a long time for this report to run.

Run the report

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your

organization. Learn more about Microsoft Purview roles and permissions.

1. You must use the latest version of the SharePoint Online Management Shell. If you installed a previous version of the SharePoint Online Management Shell, go to **Add or remove programs** and uninstall *SharePoint Online Management Shell*. Then, install the latest version. To learn more, see Get started with SharePoint Online Management Shell.

2. Connect to SharePoint Online as a Global Administrator or SharePoint Administrator in Microsoft 365. For more information, see Getting started with SharePoint Online Management Shell.

3. Run the following command to build the report:

   PowerShell

   ```
   Start-SPOInformationBarriersPolicyComplianceReport
   ```

   Or, to automatically update any noncompliant OneDrive accounts when you build the report, run:

   PowerShell

   ```
   Start-SPOInformationBarriersPolicyComplianceReport -UpdateOneDriveSegments
   ```

   Or, to automatically update any noncompliant user-owned SharePoint Embedded containers when you build the report, run:

   PowerShell

   ```
   Start-SPOInformationBarriersPolicyComplianceReport -
   UpdateUserOwnedContainerSegments
   ```

4. Run the following command to view the status of the task:

   PowerShell

   ```
   Get-SPOInformationBarriersPolicyComplianceReport
   ```

   The command returns the following set of information:

   text

```
State: Completed
Id: beedfc3e-850c-4ca2-9ae0-7eacbe529d77
StartTimeInUtc: 9/21/2026 1:11:32 AM
CompleteTimeInUtc: 9/21/2026 1:11:33 AM
QueuedTimeInUtc: 9/21/2026 1:07:11 AM
UpdateOneDriveSegments: False
UpdateUserOwnedContainerSegments: True
```

5. Run the following command to view the report:

PowerShell

```
Get-SPOInformationBarriersPolicyComplianceReport -reportid <ID>
```

(Where *ID* is the report's ID from the previous step.)

The command returns the following set of information for each site:

text

```
Content: {8d7c4a12-5f6e-4b8d-9a31-c7e2f4a9d6b5}
HasNonCompliantSites: True
State: Completed
Id: beedfc3e-850c-4ca2-9ae0-7eacbe529d77
StartTimeInUtc: 9/21/2026 1:11:32 AM
CompleteTimeInUtc: 9/21/2026 1:11:33 AM
QueuedTimeInUtc: 9/21/2026 1:07:11 AM
UpdateOneDriveSegments: False
UpdateUserOwnedContainerSegments: True
```

The **Content** row lists the sites that aren't compliant. If all sites are compliant, the **Content** row is empty and **HasNonCompliantSites** is `False`.

6. Run the following command to view details about the noncompliant segments associated with each site:

PowerShell

```
$report = Get-SPOInformationBarriersPolicyComplianceReport -reportid <ID>
$report.Content
```

(Where *ID* is the report's ID from the previous step.)

The command returns the following set of information for each site:

text

```
SiteId: 3ef21e8a-69d9-4bf0-a70f-0328e5a18087
SiteUrl: https://contoso.sharepoint.com/sites/Research
SiteType: Group
ComplianceState: NonCompliant
CurrentSegments: Sales, Research
OriginalSegments: Sales, Research
InvalidIBSegments:
IncompatibleSegmentsPairs: <Sales, Research>
FailedToBeProcessed: False
```

> **Note**
>
> For info about removing incompatible segments, see Use Information Barriers with SharePoint. When you finish with a report, delete it by using `Remove-SPOInformationBarriersPolicyComplianceReport -reportid <>`.

Next steps

- Use Information Barriers with SharePoint: Remove incompatible segments from noncompliant SharePoint sites identified in the report.
- Use Information Barriers with OneDrive: Update segments on noncompliant OneDrive accounts identified in the report.
- Manage Information Barriers policies: Edit policies to resolve widespread compliance issues across sites.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 10/02/2026

Information Barriers compliance assistant (preview)

This article explains how to enable the Information Barriers (IB) compliance assistant for group-connected SharePoint sites. These sites are sites that don't have an associated team in Microsoft Teams. When you enable the IB compliance assistant, it automatically removes users who don't match the segments specified on this site to ensure group membership honors configured IB policies. This configuration helps ensure your organization remains compliant with standards, policies, and compliance regulations.

Prerequisites

1. Make sure you define policies for Information Barriers.
2. Configure IB segments on a SharePoint Site.
3. Install the Azure PowerShell module.
4. Install the Microsoft Graph PowerShell SDK.
5. PowerShell account must have directory administrator access for the tenant.

Enable the background compliance assistant

These steps create a new application in your organization's enterprise applications. For the compliance assistant to function properly, you must have explicitly added segments to a SharePoint site. Complete the following steps to enable the compliance assistant:

1. Run the following PowerShell cmdlets.

   PowerShell

   ```
   Connect-MgGraph
   Connect-AzAccount
   $appId="00001111-aaaa-2222-bbbb-3333cccc4444"
   $sp=Get-AzADServicePrincipal -ServicePrincipalName $appId
   if ($sp -eq $null) {New-AzADServicePrincipal -ApplicationId $appId}
   Start-Process "https://login.microsoftonline.com/common/adminconsent?
   client_id=$appId"
   ```

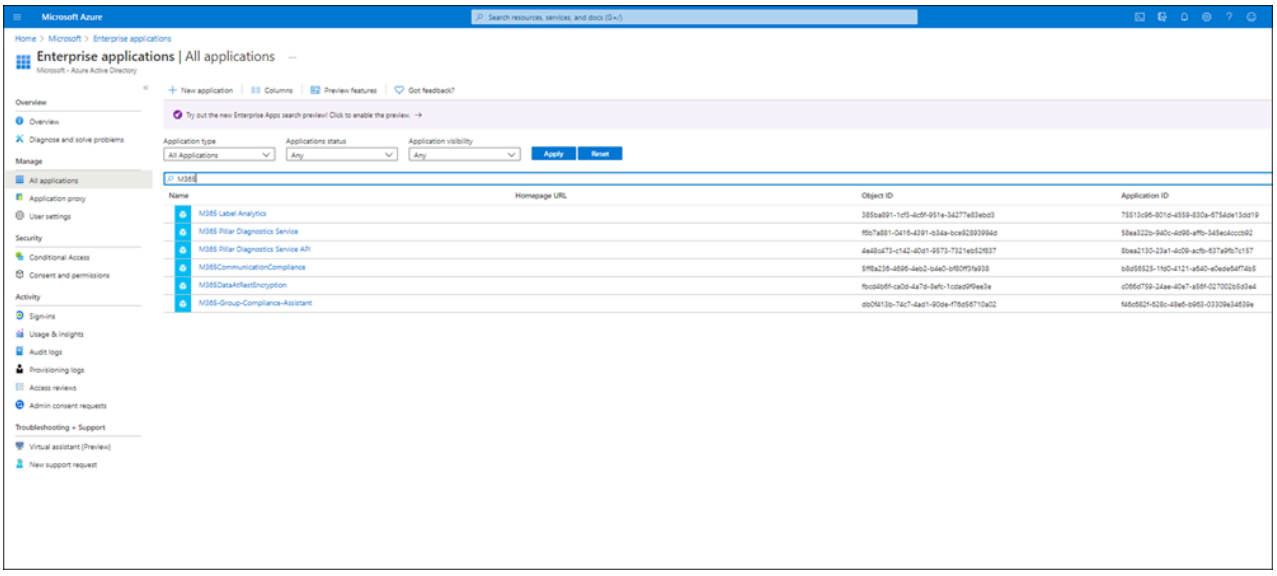
2. When prompted, sign in by using your Microsoft 365 work or school account.

3. In the **Permissions requested** dialog box, review the information, and select **Accept**. This action configures admin consent for the compliance assistant.

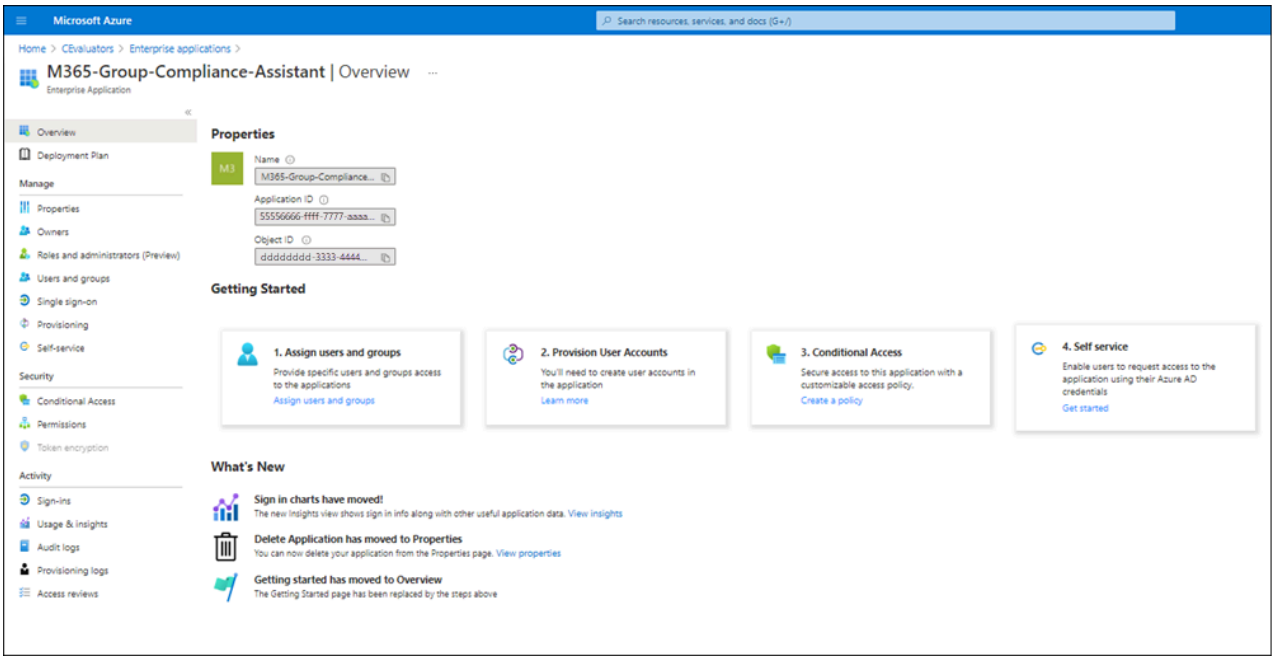
Verify you created a new application

To verify you created a new application in your organization's enterprise applications, complete the following steps:

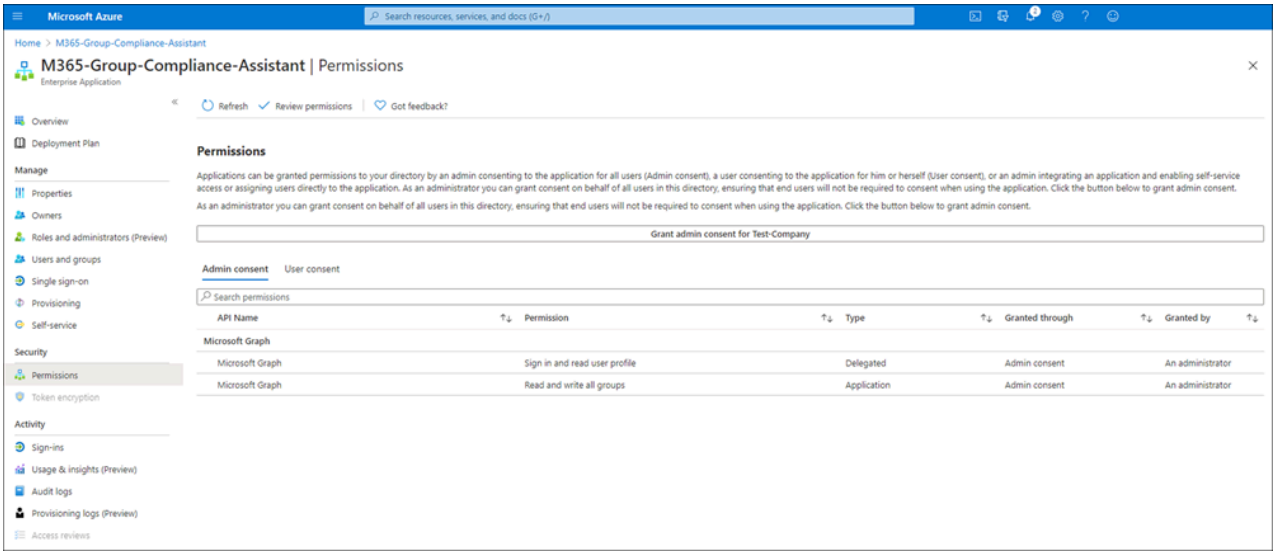
1. Sign in to portal.azure.com with directory administrator credentials.
2. Select **Manage Microsoft Entra ID**.
3. Select **Enterprise Applications** in the left nav.
4. Search for the compliance assistant by using 'M365' as the search term.



5. Select **M365-Group-Compliance-Assistant** from the list of search results.
6. On the **M365-Group-Compliance-Assistant overview** page, review the application properties.



7. Select **Permissions** in the left nav to review the permissions that the application is authorized for.



8. In this example, the **M365-Group-Compliance-Assistant** is authorized to add or remove noncompliant IB users from your Microsoft 365 groups.

You can use audit log search in the Microsoft Purview portal to search, review, and track audit log events for the M365-Group-Compliance-Assistant application. The audit activities associated with the compliance assistant are:

- **IB assistant removed group member**: The IB noncompliant group member was removed from the group by the compliance assistant.
- **IB assistant removed group owner**: The IB noncompliant owner was removed from the group by the compliance assistant.

- **Identified as IB non-compliant group**: The segments on the group are non-IB compliant with each other.

To search the audit log for Microsoft 365 Groups activities, see Search the audit log.

> **Note**
>
> The compliance assistant runs periodically (every 24 hours). The assistant runs on group-connected SharePoint sites that don't have an associated team in Microsoft Teams. To enable the compliance assistant for SharePoint sites connected to Microsoft Teams, follow the instructions in the **Define IB policies** article.

Next steps

- Use Information Barriers with SharePoint: Manage IB segments and modes on the SharePoint sites the assistant monitors.
- Create an Information Barriers policy compliance report: Identify noncompliant sites that need manual remediation beyond the assistant.
- Manage Information Barriers policies: Update segments or policies to resolve recurring compliance issues.

Note: The author created this article with assistance from AI. Learn more

---

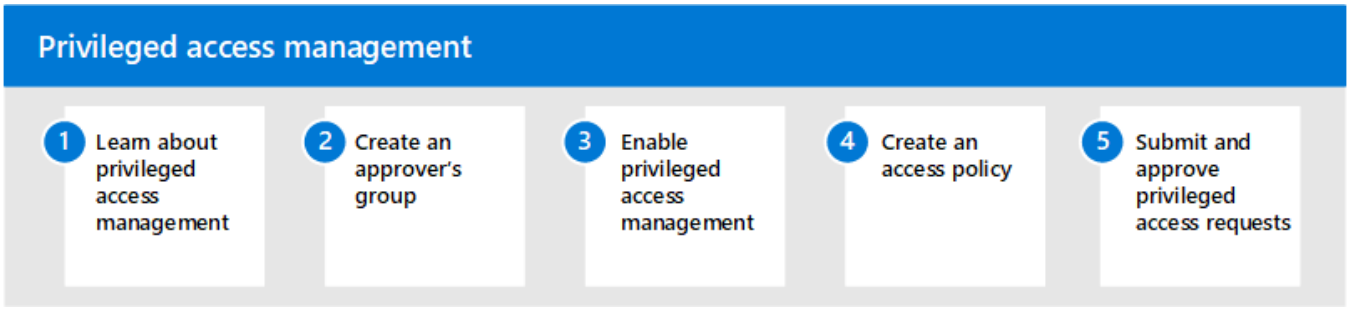
Last updated on 04/14/2026

Privileged access management

When users have standing access to sensitive information or critical network configuration settings in Microsoft Exchange Online, compromised accounts or internal threat activities can occur. Microsoft Purview Privileged Access Management helps protect your organization from breaches and helps you meet compliance best practices by limiting standing access to sensitive data or access to critical configuration settings. Instead of administrators having constant access, implement just-in-time access rules for tasks that need elevated permissions. When you enable privileged access management for Exchange Online in Microsoft 365, your organization operates with zero standing privileges and provides a layer of defense against standing administrative access vulnerabilities.

Configure privileged access management

Use the following steps to configure privileged access management for your organization:



1. Learn about privileged access management
2. Create an approver's group
3. Enable privileged access management
4. Create an access policy
5. Submit/approve privileged access requests

More information about privileged access management

- Frequently asked questions about privileged access management

---

Last updated on 03/24/2026

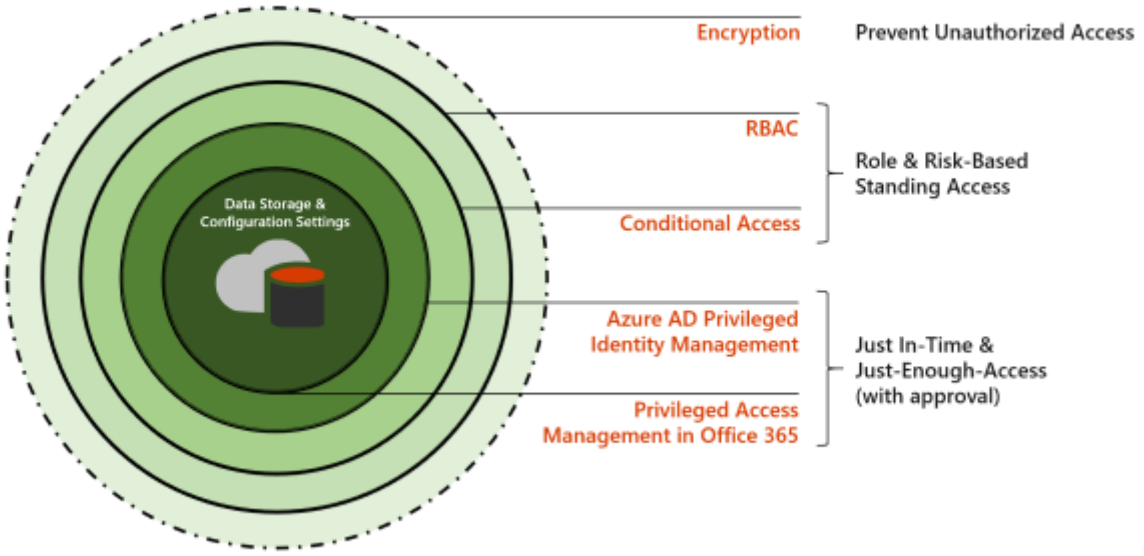
Learn about privileged access management

Microsoft Purview Privileged Access Management provides granular access control over privileged admin tasks in Microsoft 365. It helps protect your organization from breaches that use existing privileged admin accounts with standing access to sensitive data or access to critical configuration settings. Privileged access management requires users to request just-in-time access to complete elevated and privileged tasks through a highly scoped and time-bounded approval workflow. This configuration gives users just enough access to perform the task at hand, without risking exposure of sensitive data or critical configuration settings. When you enable privileged access management, your organization operates with zero standing privileges and provides a layer of defense against standing administrative access vulnerabilities.

For a quick overview of the integrated Customer Lockbox and privileged access management workflow, see this Customer Lockbox and privileged access management video.

Layers of protection

Privileged access management complements other data and access feature protections within the Microsoft 365 security architecture. Including privileged access management as part of an integrated and layered approach to security provides a security model that maximizes protection of sensitive information and Microsoft 365 configuration settings. As shown in the diagram, privileged access management builds on the protection provided by native encryption of Microsoft 365 data and the role-based access control security model of Microsoft 365 services. When used with Microsoft Entra Privileged Identity Management, these two features provide access control with just-in-time access at different scopes.



Privileged access management is defined and scoped at the **task** level, while Microsoft Entra Privileged Identity Management applies protection at the **role** level with the ability to execute multiple tasks. Microsoft Entra Privileged Identity Management primarily manages accesses for AD roles and role groups, while Microsoft Purview Privileged Access Management applies only at the task level.

- **Enable privileged access management while already using Microsoft Entra Privileged Identity Management:** Adding privileged access management provides another granular layer of protection and audit capabilities for privileged access to Microsoft 365 data.
- **Enable Microsoft Entra Privileged Identity Management while already using Microsoft Purview Privileged Access Management:** Adding Microsoft Entra Privileged Identity Management to Microsoft Purview Privileged Access Management can extend privileged access to data outside of Microsoft 365 that's primarily defined by user roles or identity.

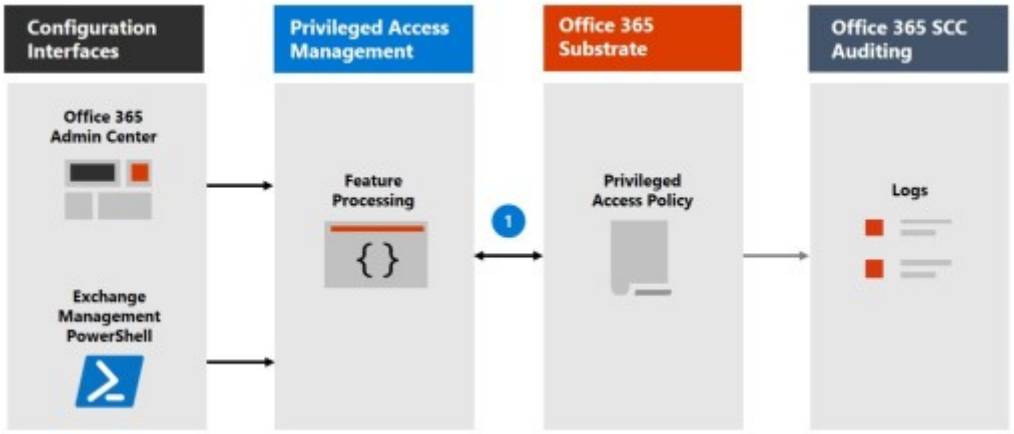
Privileged access management architecture and process flow

Each of the following process flows outlines the architecture of privileged access and how it interacts with the Microsoft 365 substrate, auditing, and the Exchange Management runspace.

Step 1: Configure a privileged access policy

When you configure a privileged access policy with the Microsoft 365 admin center or the Exchange Management PowerShell, you define the policy and the privileged access feature processes and the policy attributes in the Microsoft 365 substrate. The system logs the

activities in the audit log. The policy is now enabled and ready to handle incoming requests for approvals.



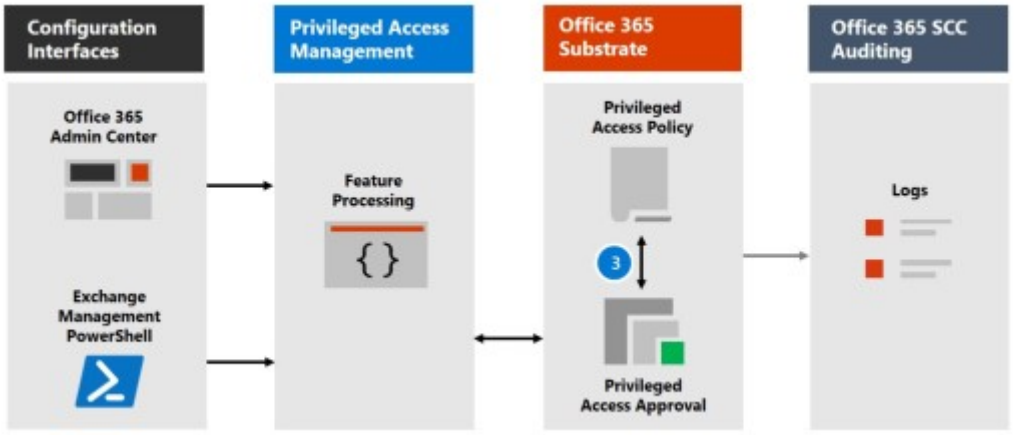
Step 2: Access request

In the Microsoft 365 admin center or by using Exchange Management PowerShell, users can request access to elevated or privileged tasks. The privileged access feature sends the request to the Microsoft 365 substrate for processing against the configured privileged access policy and records the activity in the audit logs.



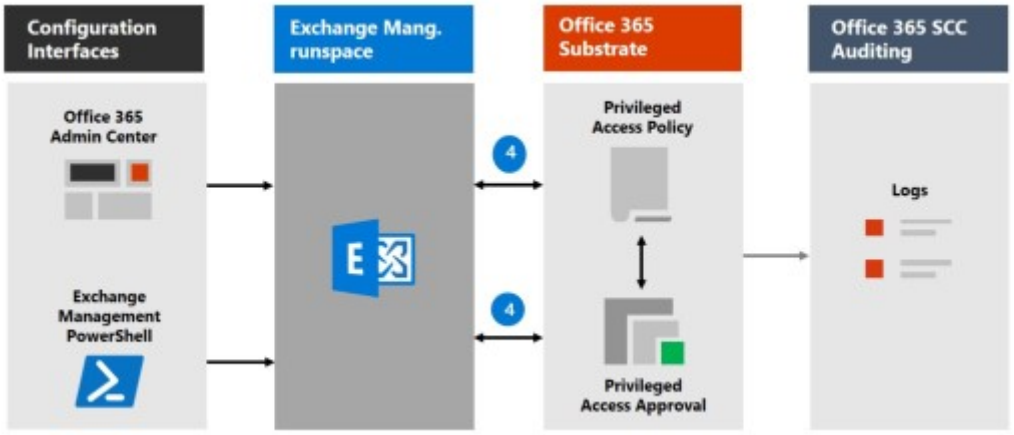
Step 3: Access approval

The privileged access feature generates an approval request and sends a pending request notification to approvers by email. If approvers grant approval, the privileged access feature processes the privileged access request as an approval and makes the task ready to complete. If approvers deny the request, the privileged access feature blocks the task and grants no access to the requestor. The requestor receives an email message that notifies them of the request approval or denial.



Step 4: Access processing

For an approved request, the Exchange Management runspace processes the task. The Microsoft 365 substrate checks the approval against the privileged access policy and processes it. The audit logs record all activity for the task.



Frequently asked questions

What SKUs can use privileged access in Microsoft 365?

You can use privileged access management with a wide selection of Microsoft 365 and Microsoft 365 subscriptions and add-ons. For details, see Get started with privileged access management.

When will privileged access management support Microsoft 365 workloads beyond Exchange?

For more information about the timeline for privileged access management support, see the Microsoft 365 Roadmap.

My organization needs more than 30 privileged access policies. Will this limit be increased?

Yes, Microsoft is working on increasing the current limit of 30 privileged access policies per organization.

Do I need to be a Global Admin to manage privileged access in Microsoft 365?

Yes, you need the *Global Admin* role assigned to accounts that manage privileged access in Microsoft 365. Users included in an approvers' group don't need to have the *Global Admin* or *Role Management* roles assigned to review and approve requests by using PowerShell. Users must have the *Exchange Administrator* role assigned to request, review, and approve privileged access requests in the Microsoft 365 admin center.

How is privileged access management related to Customer Lockbox?

Customer Lockbox provides access control for organizations when Microsoft accesses data. Privileged access management provides granular access control within an organization for all Microsoft 365 privileged tasks.

Ready to get started?

Start configuring your organization for privileged access management.

Learn more

Interactive guide: Monitor and control administrator tasks with privileged access management

ⓘ **Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 04/14/2026

Get started with privileged access management

This article guides you through enabling and configuring privileged access management in your organization. You can use either the Microsoft 365 admin center or Exchange Management PowerShell to manage and use privileged access.

Before you begin

Before you get started with privileged access management, confirm your Microsoft 365 subscription and any add-ons.

To access and use privileged access management, your organization must have supporting subscriptions or add-ons. For more information, see the subscription requirements for privileged access management.

If you don't have an existing Microsoft 365 Enterprise E5 plan and want to try privileged access management, you can add Microsoft 365 to your existing Microsoft 365 subscription or sign up for a trial of Microsoft 365 Enterprise E5.

Enable and configure privileged access management

Complete the following steps to set up and use privileged access in your organization:

- Step 1: Create an approver's group

  Before you start using privileged access, determine who needs approval authority for incoming requests for access to elevated and privileged tasks. Any user who is part of the Approvers' group can approve access requests. Enable this group by creating a mail-enabled security group in Microsoft 365.

- Step 2: Enable privileged access

  Explicitly enable privileged access in Microsoft 365 with the default approver group, including a set of system accounts that you want to exclude from the privileged access management access control.

- Step 3: Create an access policy

Define specific approval requirements scoped at individual tasks by creating an approval policy. The approval type options are **Auto** or **Manual**.

- Step 4: Submit and approve privileged access requests

  Once enabled, privileged access requires approvals for any task that has an associated approval policy defined. For tasks included in an approval policy, users must request and be granted access approval to have permissions necessary to execute the task.

After approval is granted, the requesting user can execute the intended task. Privileged access authorizes and executes the task on behalf of the user. The approval remains valid for the requested duration (the default duration is four hours), during which the requester can execute the intended task multiple times. All such executions are logged and made available for security and compliance auditing.

> **Note**
>
> To use Exchange Management PowerShell to enable and configure privileged access, follow the steps in **Connect to Exchange Online PowerShell using Multi-Factor authentication** to connect to Exchange Online PowerShell with your Microsoft 365 credentials. You don't need to enable multi-factor authentication for your organization to use the steps to enable privileged access while connecting to Exchange Online PowerShell. Connecting with multi-factor authentication creates an auth token that privileged access uses for signing your requests.

Step 1: Create an approver's group

1. Sign in to the Microsoft 365 admin center with credentials for an admin account in your organization.
2. In the admin center, go to **Groups** > **Add a group**.
3. Select **mail-enabled security group** and then enter the **Name**, **Group email address**, and **Description** for the new group.
4. Save the group. It might take a few minutes for the group to be fully configured and to appear in the Microsoft 365 admin center.
5. Select the new approvers group and select **edit** to add users to the group.

6. Save the group.

Step 2: Enable privileged access

In the Microsoft 365 admin center

1. Sign in to the Microsoft 365 admin center with credentials for an admin account in your organization.
2. In the admin center, go to **Settings** > **Org Settings** > **Security & Privacy** > **Privileged access**.
3. Turn on **Require approvals for privileged tasks**.
4. Assign the approvers group you created in Step 1 as the **Default approvers group**.
5. **Save** and **Close**.

In Exchange Management PowerShell

To enable privileged access and assign the approver group, run the following command in Exchange Online PowerShell:

PowerShell

```
Enable-ElevatedAccessControl -AdminGroup '<default approver group>' -SystemAccounts @('<systemAccountUPN1>','<systemAccountUPN2>')
```

Example:

PowerShell

```
Enable-ElevatedAccessControl -AdminGroup 'pamapprovers@fabrikam.onmicrosoft.com' -SystemAccounts @('sys1@fabrikamorg.onmicrosoft.com', 'sys2@fabrikamorg.onmicrosoft.com')
```

> **Note**
>
> The system accounts feature ensures that certain automations within your organizations work without privileged access. However, make such exclusions exceptional and regularly approve and audit those allowed.

Step 3: Create an access policy

You can create and configure up to 30 privileged access policies for your organization.

In the Microsoft 365 admin center

1. Sign in to the Microsoft 365 admin center with credentials for an admin account in your organization.
2. In the Admin Center, go to **Settings** > **Org Settings** > **Security & Privacy** ** > **Privileged access**.
3. Select **Manage access policies and requests**.
4. Select **Configure policies** and select **Add a policy**.
5. From the drop-down fields, select the appropriate values for your organization:
    - **Policy type**: Task, Role, or Role Group
    - **Policy scope**: Exchange
    - **Policy name**: Select from the available policies
    - **Approval type**: Manual or Auto
    - **Approval group**: Select the approvers group created in Step 1
6. Select **Create** and then **Close**. It might take a few minutes for the policy to be fully configured and enabled.

In Exchange Management PowerShell

To create and define an approval policy, run the following command in Exchange Online PowerShell:

PowerShell

```
New-ElevatedAccessApprovalPolicy -Task 'Exchange\<exchange management cmdlet name>'
-ApprovalType <Manual, Auto> -ApproverGroup '<default/custom approver group>'
```

Example:

PowerShell

```
New-ElevatedAccessApprovalPolicy -Task 'Exchange\New-MoveRequest' -ApprovalType
Manual -ApproverGroup 'mbmanagers@fabrikamorg.onmicrosoft.com'
```

Step 4: Submit or approve privileged access requests

Request elevation authorization to execute privileged tasks

Requests for privileged access are valid for up to 24 hours after submission. If approvers don't approve or deny the request within 24 hours, the request expires and access isn't granted.

In the Microsoft 365 Admin Center

1. Sign in to the Microsoft 365 Admin Center with your credentials.
2. In the Admin Center, go to **Settings** > **Org Settings** > **Security & Privacy** ** > **Privileged access**.
3. Select **Manage access policies and requests**.
4. Select **New request**. From the drop-down fields, select the appropriate values for your organization:
   - **Request type**: Task, Role, or Role Group
   - **Request scope**: Exchange
   - **Request for**: Select from the available policies
   - **Duration (hours)**: Number of hours of requested access. You can request any number of hours.
   - **Comments**: Text field for comments related to your access request
5. Select **Save** and then **Close**. The system sends your request to the approver group by email.

In Exchange Management PowerShell

Run the following command in Exchange Online PowerShell to create and submit an approval request to the approver group:

PowerShell

```
New-ElevatedAccessRequest -Task 'Exchange\<exchange management cmdlet name>' -
Reason '<appropriate reason>' -DurationHours <duration in hours>
```

Example:

PowerShell

```
New-ElevatedAccessRequest -Task 'Exchange\New-MoveRequest' -Reason 'Attempting to
fix the user mailbox error' -DurationHours 4
```

View status of elevation requests

After you create an approval request, you can check the status of the elevation request in the admin center or in Exchange Management PowerShell by using the associated request ID.

In the Microsoft 365 admin center

1. Sign in to the Microsoft 365 admin center with your credentials.
2. In the admin center, go to **Settings** > **Org Settings** > **Security & Privacy** > **Privileged access**.
3. Select **Manage access policies and requests**.
4. Select **View** to filter submitted requests by **Pending**, **Approved**, **Denied**, or **Customer Lockbox** status.

In Exchange Management PowerShell

Run the following command in Exchange Online PowerShell to view an approval request status for a specific request ID:

PowerShell

```
Get-ElevatedAccessRequest -Identity <request ID> | select RequestStatus
```

Example:

PowerShell

```
Get-ElevatedAccessRequest -Identity 28560ed0-419d-4cc3-8f5b-603911cbd450 | select
RequestStatus
```

Approving an elevation authorization request

When you create an approval request, members of the relevant approver group receive an email notification. They can approve the request by using the request ID. The requestor receives an email notification when you approve or deny the request.

In the Microsoft 365 admin center

1. Sign in to the Microsoft 365 admin center with your credentials.
2. In the admin center, go to **Settings** > **Org Settings** > **Security & Privacy** > **Privileged access**.
3. Select **Manage access policies and requests**.
4. Select a listed request to view the details and take action on the request.
5. Select **Approve** to approve the request or select **Deny** to deny the request. You can revoke access for previously approved requests by selecting **Revoke**.

In Exchange Management PowerShell

To approve an elevation authorization request, run the following command in Exchange Online PowerShell:

```
PowerShell
Approve-ElevatedAccessRequest -RequestId <request id> -Comment '<approval comment>'
```

Example:

```
PowerShell
Approve-ElevatedAccessRequest -RequestId a4bc1bdf-00a1-42b4-be65-b6c63d6be279 -
Comment '<approval comment>'
```

To deny an elevation authorization request, run the following command in Exchange Online PowerShell:

```
PowerShell
Deny-ElevatedAccessRequest -RequestId <request id> -Comment '<denial comment>'
```

Example:

```
PowerShell
Deny-ElevatedAccessRequest -RequestId a4bc1bdf-00a1-42b4-be65-b6c63d6be279 -Comment
'<denial comment>'
```

Delete a privileged access policy in Microsoft 365

If your organization no longer needs a privileged access policy, you can delete it.

In the Microsoft 365 admin center

1. Sign in to the Microsoft 365 admin center with credentials for an admin account in your organization.
2. In the admin center, go to **Settings** > **Org Settings** > **Security & Privacy** > **Privileged access**.
3. Select **Manage access policies and requests**.
4. Select **Configure policies**.
5. Select the policy you want to delete, and then select **Remove Policy**.
6. Select **Close**.

In Exchange Management PowerShell

To delete a privileged access policy, run the following command in Exchange Online PowerShell:

```
PowerShell
Remove-ElevatedAccessApprovalPolicy -Identity <identity GUID of the policy you want
to delete>
```

Disable privileged access in Microsoft 365

If needed, you can disable privileged access management for your organization. Disabling privileged access doesn't delete any associated approval policies or approver groups.

In the Microsoft 365 admin center

1. Sign in to the Microsoft 365 admin center with credentials for an admin account in your organization.
2. In the Admin Center, go to **Settings** > **Org Settings** > **Security & Privacy** ** > **Privileged access**.
3. Turn on **Require approvals for privileged access**.

In Exchange Management PowerShell

To disable privileged access, run the following command in Exchange Online PowerShell:

PowerShell

```
Disable-ElevatedAccessControl
```

ⓘ **Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 04/14/2026

Encryption

09/26/2025

Encryption is an important part of your file protection and information protection strategy. This article provides an overview of encryption for Microsoft 365. Get help with encryption tasks like how to set up encryption for your organization and how to password-protect Microsoft 365 documents.

- For information about certificates and technologies like TLS, see Technical reference details about encryption in Microsoft 365.
- For an overview of how to configure or set up encryption for your organization, see Set up encryption in Microsoft 365 Enterprise.

What is encryption, and how does it work in Microsoft 365?

The encryption process encodes your data (referred to as plaintext) into ciphertext. Unlike plaintext, ciphertext can't be used by people or computers unless and until the ciphertext is decrypted. Decryption requires an encryption key that only authorized users have. Encryption helps ensure that only authorized recipients can decrypt your content. Content includes files, email messages, calendar entries, and so on.

Encryption by itself doesn't prevent content interception. Encryption is part of a larger information protection strategy for your organization. By using encryption, you help ensure that only authorized parties can use the encrypted data.

You can have multiple layers of encryption in place at the same time. For example, you can encrypt email messages and also the communication channels through which your email flows. With Microsoft 365, your data is encrypted at rest and in transit, using several strong encryption protocols, and technologies that include Transport Layer Security/Secure Sockets Layer (TLS/SSL), Internet Protocol Security (IPSec), and Advanced Encryption Standard (AES).

Encryption for data at rest and data in transit

Examples of data at rest include files that you uploaded to a SharePoint library, Project Online data, documents that you uploaded in a Skype for Business meeting, email messages and attachments that you stored in folders in your mailbox, and files you uploaded to OneDrive.

Examples of data in transit include mail messages that are in the process of being delivered, or conversations that are taking place in an online meeting. In Microsoft 365, data is in transit

whenever a user's device is communicating with a Microsoft server, or when a Microsoft server is communicating with another server.

With Microsoft 365, multiple layers and kinds of encryption work together to secure your data. The following table includes some examples, with links to additional information.

Expand table

Kinds of Content	Encryption Technologies	Resources to learn more
Files on a device. These files can include email messages saved in a folder, documents saved on a computer, tablet, or phone, or data saved to the Microsoft cloud.	BitLocker in Microsoft data centers. BitLocker can also be used on client machines, such as Windows computers and tablets Distributed Key Manager (DKM) in Microsoft data centers Customer Key for Microsoft 365	Windows IT Center: BitLocker Microsoft Trust Center: Encryption Cloud security controls series: Encrypting Data at Rest How Exchange Online secures your email secrets Service encryption with Microsoft Purview Customer Key
Files in transit between users. These files can include Microsoft 365 documents or SharePoint list items shared between users.	TLS for files in transit	Data Encryption in OneDrive and SharePoint Skype for Business Online: Security and Archiving
Email in transit between recipients. This email includes email hosted by Exchange Online.	Microsoft Purview Message Encryption with Azure Rights Management, S/MIME, and TLS for email in transit	Message Encryption Email encryption in Microsoft 365 How Exchange Online uses TLS to secure email connections in Microsoft 365
Chats, messages, and files in transit between recipients using Microsoft Teams.	Teams uses TLS and MTLS to encrypt instant messages. Media traffic is encrypted using Secure RTP (SRTP). Teams uses FIPS (Federal Information Processing Standard) compliant algorithms for encryption key exchanges.	Encryption for Teams

Microsoft 365 Crypto Update

In late August 2023, Microsoft Purview Information Protection began to use Advanced Encryption Standard (AES) with 256-bit key length in Cipher Block Chaining mode (AES256-CBC). By October 2023, AES256-CBC became the default for encryption of Microsoft 365 Apps documents and emails. You might need to take action to support this change in your organization. For more information, see Technical reference details about encryption.

What if I need more control over encryption to meet security and compliance requirements?

Microsoft 365 provides Microsoft-managed solutions for volume encryption, file encryption, and mailbox encryption in Microsoft 365. In addition, Microsoft provides encryption solutions that you can manage and control. These encryption solutions are built on Azure.

To learn more, see the following resources:

- What is Azure Rights Management?
- Activate Rights Management in the admin center
- Set up Information Rights Management (IRM) in SharePoint admin center
- Overview of Customer Key
- Double Key Encryption

How do I...

Expand table

To do this task	See these resources
Set up encryption for my organization	Set up encryption in Microsoft 365 Enterprise
View details about certificates, technologies, and TLS cipher suites	Technical details about encryption
Work with encrypted messages on a mobile device	View encrypted messages on your Android device View encrypted messages on your iPhone or iPad
Encrypt a document using password protection. (Password protection isn't supported in a browser.	Add or remove protection in your document, workbook, or presentation . Choose an **Add**

To do this task	See these resources
Use desktop versions of Word, Excel, and PowerPoint for password protection.)	**protection** section, and then see **Encrypt with Password**.
Remove encryption from a document	Add or remove protection in your document, workbook, or presentation. Choose a **Remove protection** section, and then see **Remove password encryption**.

Related topics

Plan for Microsoft 365 security and information protection capabilities

Secure your business data with Microsoft 365 for business

Microsoft Stream Video level encryption and playback flow

Learn about the Azure Rights Management encryption service

09/03/2025

Azure Rights Management, sometimes abbreviated to Azure RMS, is the main cloud-based encryption service from Microsoft Purview Information Protection.

Azure Rights Management helps to protect items such as files and emails across multiple devices that include phones, tablets, and PCs by using encryption, identity, and authorization policies.

For example, when employees email a document to a partner company, or save a document to their cloud drive, the persistent encryption from Azure Rights Management helps secure the data.

- **Encryption settings remain with your data**, even when it leaves your organization's boundaries, keeping your content protected both within and outside your organization.
- **Encryption may be legally required** for compliance, legal discovery requirements, or best practices for information management.
- **Use Azure Rights Management with Microsoft 365 subscriptions or subscriptions for Microsoft Purview Information Protection**. For more information, see the Microsoft 365 licensing guidance for security & compliance page.

Azure Rights Management ensures that authorized people and services, such as search and indexing, can continue to read and inspect the encrypted data.

Ensuring ongoing access for authorized people and services, also known as "reasoning over data", is a crucial element in maintaining control of your organization's data. This capability may not be easily accomplished with other information protection solutions that use peer-to-peer encryption.

Protection features

Expand table

Feature	Description
Encrypt multiple file types	In early implementations of Rights Management, only Office files could be encrypted, using built-in Rights Management protection.

Feature	Description
	Azure Rights Management provides support for additional file types. For more information, see Supported file types.
Protect files anywhere	When a file is encrypted, this protection stays with the file, even if it's saved or copied to storage that isn't under the control of IT, such as a cloud storage service.

Collaboration features

Expand table

Feature	Description
Safely share information	Encrypted files are safe to share with others, such as an attachment to an email or a link to a SharePoint site. If the sensitive information is within an email message, encrypt the email, or use the **Do Not Forward** option from Outlook.
Support for business-to-business collaboration	Because Azure Rights Management is a cloud service, there's usually no need to explicitly configure trusts with other organizations before you can share encrypted content with them. By default, collaboration with other organizations that already have a Microsoft 365 or a Microsoft Entra directory is automatically supported. Some additional configuration might be needed for advanced configurations or specialized scenarios. For organizations without Microsoft 365 or a Microsoft Entra directory, users can sign up for the free RMS for individuals subscription, or use a Microsoft account for supported applications.

Tip

Attaching encrypted files, rather than encrypting an entire email message, enables you to keep the email text unencrypted.

For example, you may want to include instructions for first-time use if the email is being sent outside your organization. If you attach an encrypted file, the basic instructions can be read by anyone, but only authorized users will be able to open the document, even if the email or document is forwarded to other people.

Platform support features

The Azure Rights Management service supports a broad range of platforms and applications, including:

Expand table

Feature	Description
Commonly used devices not just Windows computers	Client devices include: - Windows computers and phones - Mac computers - iOS tablets and phones - Android tablets and phones
On-premises services	In addition to working seamlessly with Microsoft 365, you can use Azure Rights Management with the following on-premises services when you deploy the Microsoft Rights Management connector: - Exchange Server - SharePoint Server - Windows Server running File Classification Infrastructure
Application extensibility	Azure Rights Management has tight integration with Microsoft 365 applications and services, and extends support for other applications by using the Microsoft Purview Information Protection client. The Microsoft Information Protection SDK provides your internal developers and software vendors with APIs to write custom applications that support the Azure Rights Management service.

Infrastructure features

The Azure Rights Management service provides the following features to support IT departments and infrastructure organizations:

- Create simple and flexible policies
- Easy activation
- Auditing and monitoring services
- Ability to scale across your organization
- Maintain IT control over data

> **Note**
>
> Organizations always have the choice to stop using the Azure Rights Management service without losing access to content that was previously protected by Azure Rights

Management.

For more information, see Decommission and deactivate the Azure Rights Management service.

Create simple and flexible policies

Encryption settings that are applied with sensitivity labels provide a quick and easy method for administrators to apply information protection policies, and for users to apply the correct level of protection for each item as needed.

For example, for a company-wide strategy paper to be shared with all employees, apply a read-only policy to all internal employees. For a more sensitive document, such as a financial report, restrict access to executives only.

For more information, see Restrict access to content by using sensitivity labels to apply encryption.

Easy activation

For new subscriptions, activation is automatic. For existing subscriptions, activating the Rights Management service requires just two PowerShell commands.

Auditing and monitoring services

Audit and monitor usage of your encrypted files, even after these files leave your organization's boundaries.

For example, if a Contoso, Ltd employee works on a joint project with three people from Fabrikam, Inc, they might send their Fabrikam partners a document that's encrypted and restricted to *read-only*.

Azure Rights Management auditing can provide the following information:

- Whether the Fabrikam partners opened the document, and when.
- Whether other people who weren't specified, attempted, and failed to open the document. This might happen if the email was forwarded on, or saved to a shared location.

Administrators can track document usage and revoke access for Office files. Users can revoke access for their labeled and encrypted documents as needed.

Ability to scale across your organization

Because Azure Rights Management runs as a cloud service with the Azure elasticity to scale up and out, you don't have to provision or deploy additional on-premises servers.

Maintain IT control over data

Organizations can benefit from IT control features, such as:

Expand table

Feature	Description
Tenant key management	Use tenant key management solutions, such as Bring Your Own Key (BYOK) or Double Key Encryption (DKE). For more information, see: - Planning and implementing your Azure Rights Management tenant key - What is Double Key Encryption (DKE)?
Auditing and usage logging	Use auditing and usage logging to analyze for business insights, monitor for abuse, and perform forensic analysis for information leaks.
Access delegation	Delegate access with the super user feature, ensuring that IT can always access encrypted content, even if a document was encrypted by an employee who then leaves the organization. In comparison, peer-to-peer encryption solutions risk losing access to company data.
Active Directory synchronization	Synchronize just the directory attributes that Azure RMS needs to support a common identity for your on-premises Active Directory accounts, by using a hybrid identity solution, such as Microsoft Entra Connect.
Single-sign on	Enable single-sign on without replicating passwords to the cloud, by using AD FS.
Migration from AD RMS	If you've deployed Active Directory Rights Management Services (AD RMS), migrate to the Azure Rights Management service without losing access to data that was previously encrypted by AD RMS.

Security, compliance, and regulatory requirements

Azure Rights Management supports the following security, compliance, and regulatory requirements:

- **Use of industry-standard cryptography and supports FIPS 140-2.** For more information, see the Cryptographic controls: Algorithms and key lengths information.

- **Support for nCipher nShield hardware security module (HSM)** to store your tenant key in Microsoft Azure data centers.

  Azure Rights Management uses separate security worlds for its data centers in North America, EMEA (Europe, Middle East and Africa), and Asia, so your keys can be used only in your region.

- **Certification for the following standards**:
  - ISO/IEC 27001:2013 (./includes ISO/IEC 27018)
  - SOC 2 SSAE 16/ISAE 3402 attestations
  - HIPAA BAA
  - EU Model Clause
  - FedRAMP as part of Microsoft Entra ID in Office 365 certification, issued FedRAMP Agency Authority to Operate by HHS
  - PCI DSS Level 1

For more information about these external certifications, see the Microsoft Trust Center.

Frequently asked questions

Some of our more frequently asked questions about the Azure Rights Management encryption service from Microsoft Purview Information Protection:

Do files have to be in the cloud to be encrypted by the Azure Rights Management service?

No, this is a common misconception. The Azure Rights Management service (and Microsoft) doesn't see or store your data as part of the encryption process. Information that you encrypt is never sent to or stored in Azure unless you explicitly store it in Azure or use another cloud service that stores it in Azure.

For more information, see How the Azure Rights Management service works: Technical details to understand how a secret formula that's created and stored on-premises is encrypted by the Azure Rights Management service but remains on-premises.

What's the difference between Azure Rights Management encryption and encryption in other Microsoft cloud services?

Microsoft provides multiple encryption technologies that enable you to protect your data for different, and often complementary scenarios. For example, while Microsoft 365 offers encryption at-rest for data stored in Microsoft 365, the Azure Rights Management service from

Microsoft Purview Information Protection independently encrypts your data so that it's protected regardless of where it's located or how it's transmitted.

These encryption technologies are complementary and using them requires enabling and configuring them independently. When you do so, you might have the option to bring your own key for the encryption, a scenario also known as "BYOK." Enabling BYOK for one of these technologies doesn't affect the others. For example, you can use BYOK for the Azure Rights Management service and not use BYOK for other encryption technologies, and vice versa. The keys used by these different technologies might be the same or different, depending on how you configure the encryption options for each service.

I see Microsoft Rights Management Services is listed as an available cloud app for conditional access—how does this work?

Yes, you can configure Microsoft Entra Conditional Access for the Azure Rights Management service.

When a user opens a document that's encrypted by the Azure Rights Management service, administrators can block or grant access to users in their tenant, based on the standard conditional access controls. Requiring multifactor authentication (MFA) is one of the most commonly requested conditions. Another one is that devices must be compliant with your Intune policies so that, for example, mobile devices meet your password requirements and a minimum operating system version, and computers must be domain-joined.

For more information, see Conditional Access policies and encrypted documents.

Additional information:

Expand table

Topic	Details
Evaluation frequency	For Windows computers, the conditional access policies for the Azure Rights Management service are evaluated when the user environment is initialized (this process is also known as bootstrapping), and then every 30 days. To fine-tune how often your conditional access policies get evaluated, configure the token lifetime.
Administrator accounts	We recommend that you don't add administrator accounts to your conditional access policies because these accounts won't be able to access the Azure Rights Management service when you configure encryption settings for sensitivity labels in the Microsoft Purview portal.

Topic	Details
MFA and B2B collaboration	If you use MFA in your conditional access policies for collaborating with other organizations (B2B), you must use Microsoft Entra B2B collaboration and create guest accounts for the users you want to share with in the other organization.
Terms of Use prompts	You can prompt users to accept a terms of use before they open an encrypted document for the first time.
Cloud apps	If you use many cloud apps for conditional access, you might not see **Microsoft Information Protection Sync Service** and **Microsoft Rights Management Service** displayed in the list to select. In this case, use the search box at the top of the list. Start typing "Microsoft Information Protection Sync Service" and "Microsoft Rights Management Service" to filter the available apps. Providing you have a supported subscription; you'll then see these options and will be able to select them.

I've encrypted a document and now want to change the usage rights or add users—do I need to re-encrypt the document?

If the document was encrypted by using a sensitivity label or rights management template, there's no need to re-encrypt the document. Modify the sensitivity label or rights management template by making your changes to the usage rights or add new groups (or users), and then save these changes:

- When a user hasn't accessed the document before you made the changes, the changes take effect as soon as the user opens the document.
- When a user has already accessed the document, these changes take effect when their use license expires. Re-encrypt the document only if you can't wait for the use license to expire. For example, remove the sensitivity label that applied the encryption, save, and apply the label again. Re-encrypting effectively creates a new version of the document, and therefore a new use license for the user.

Alternatively, if you have already configured a group for the required permissions, you can change the group membership to include or exclude users and there isn't a need to change the sensitivity label or rights management template. There might be a small delay before the changes take effect because group membership is cached by the Azure Rights Management service.

If the document was encrypted by using user-defined permissions, such as the Let users assign permissions encryption setting for sensitivity labels, you can't change the permissions for the

existing document. You must encrypt the document again and specify all the users and all the usage rights that are required for this new version of the document. To re-encrypt an encrypted document, you must have the Full Control usage right.

If I use this encryption solution for my production environment, is my company then locked into the solution?

No, you always remain in control of your data and can continue to access it, even if you decide to no longer use the Azure Rights Management service. For more information, see Decommission and deactivate the Azure Rights Management service.

Can I control which of my users can use the Azure Rights Management service to encrypt content?

Yes, when you use sensitivity labels to encrypt content, label publishing policies define which users see the labels in their apps. If you don't want some users to encrypt content, create a separate label publishing policy for them and in the policy, include only labels that don't apply encryption. For more information, see Create and configure sensitivity labels and their policies.

When I share an encrypted document with somebody outside my company, how does that user get authenticated?

By default, the Azure Rights Management service uses a Microsoft Entra account and an associated email address for user authentication, which makes business-to-business collaboration seamless for administrators. If the other organization uses Azure services, users already have accounts in Microsoft Entra ID, even if these accounts are created and managed on-premises and then synchronized to Azure. If the organization has Microsoft 365, this service also uses Microsoft Entra ID for the user accounts. If the user's organization doesn't have managed accounts in Azure, they can be authenticated with a guest account. For more information, see Sharing encrypted documents with external users.

The authentication method for these accounts can vary, depending on how the administrator in the other organization has configured the Microsoft Entra accounts. For example, they could use passwords that were created for these accounts, federation, or passwords that were created in Active Directory Domain Services and then synchronized to Microsoft Entra ID.

Other authentication methods:

- If you encrypt an email with an Office document attachment to a user who doesn't have an account in Microsoft Entra ID, the authentication method changes. The Azure Rights Management service is federated with some popular social identity providers, such as

Gmail. If the user's email provider is supported, the user can sign in to that service and their email provider is responsible for authenticating them. If the user's email provider isn't supported, or as a preference, the user can apply for a one-time passcode that authenticates them and displays the email with the encrypted document in a web browser.

- The Azure Rights Management service can use Microsoft accounts for supported applications. However, not all applications can open encrypted content when a Microsoft account is used for authentication.

What type of groups can I use with the Azure Rights Management service?

For most scenarios, you can use any group type in Microsoft Entra ID that has an email address. This rule of thumb always applies when you assign usage rights but there are some exceptions for administering the Azure Rights Management service. For more information, see Azure Rights Management service requirements for group accounts.

How do I send an encrypted email to a Gmail or Hotmail account?

When you use Exchange Online and the Azure Rights Management service, you just send the email to the user as an encrypted message. For example, you can select a sensitivity label that automatically applies Do Not Forward for you.

The recipient sees an option to sign in to their Gmail, Yahoo, or Microsoft account, and then they can read the encrypted email. Alternatively, they can choose the option for a one-time passcode to read the email in a browser.

To support this scenario, Exchange Online must be enabled for the Azure Rights Management service and Microsoft Purview Message Encryption.

For more information about the capabilities that include supporting all email accounts on all devices, see the following blog post: Announcing new capabilities available in Office 365 Message Encryption.

Which file types are supported by the Azure Rights Management service?

The Azure Rights Management service can support all file types. For text, image, Microsoft Office (Word, Excel, PowerPoint) files, PDF files, and some other application file types, the Azure

Rights Management service supports native encryption that includes the enforcement of usage rights (permissions). For all other applications and file types, generic encryption provides file encapsulation and authentication to verify if a user is authorized to open the file.

For a list of file types supported for Office apps and services, see Office file types supported from the sensitivity labeling documentation.

For a list of additional file types that are supported by the Microsoft Purview Information Protection client, see Supported file types from the information protection client documentation.

When I open an Office document that's encrypted by the Azure Rights Management service, does the associated temporary file become encrypted by this service as well?

No. In this scenario, the associated temporary file doesn't contain data from the original document but instead, only what the user enters while the file is open. Unlike the original file, the temporary file is obviously not designed for sharing and would remain on the device, protected by local security controls, such as BitLocker and EFS.

How do we regain access to files that were encrypted by an employee who has now left the organization?

Use the super user feature, which grants the Full Control usage rights to authorized users for all items that are encrypted by your tenant. Super users can always read this encrypted content, and if necessary, remove the encryption or re-encrypt the item for different users. This same feature lets authorized services index and inspect items, as needed.

If your content is stored in SharePoint or OneDrive, admins can run the Unlock-SensitivityLabelEncryptedFile cmdlet, to remove both the sensitivity label and the encryption. For more information, see Remove encryption for a labeled document.

Can Rights Management prevent screen captures?

By not granting the **Copy** usage right, Rights Management can prevent screen captures from many of the commonly used screen capture tools on Windows platforms. In Office for Mac, screen captures can similarly be prevented in Office for Mac, for Word, Excel, and PowerPoint, but not Outlook.

However, for other apps on iOS and for Android, these operating systems don't allow apps to prevent screen captures. In addition, browsers other than Edge can't prevent screen captures.

Browser use includes Outlook on the web and Office for the web.

Preventing screen captures can help to avoid accidental or negligent disclosure of confidential or sensitive information. But there are many ways that a user can share data that is displayed on a screen, and taking a screenshot is only one method. For example, a user intent on sharing displayed information can take a picture of it using their camera phone, retype the data, or simply verbally relay it to somebody.

As these examples demonstrate, even if all platforms and all software supported the Rights Management APIs to block screen captures, technology alone can't always prevent users from sharing data that they should not. Rights Management can help to safeguard your important data by using authorization and usage policies, but this enterprise rights management solution should be used with other controls. For example, implement physical security, carefully screen and monitor people who have authorized access to your organization's data, and invest in user education so users understand what data shouldn't be shared.

What's the difference between a user encrypting an email with Do Not Forward and usage rights that don't include the Forward right?

Despite its name, **Do Not Forward** isn't the opposite of the Forward usage right, or a template. It's actually a set of rights that include restricting copying, printing, and saving the email outside the mailbox, in addition to restricting the forwarding of emails. The rights are dynamically applied to users via the chosen recipients, and not statically assigned by the administrator. For more information, see the Do Not Forward option for emails section in Configure usage rights for the Azure Rights Management service.

Next steps

For more technical information about how the Azure Rights Management service works, see How the Azure Rights Management service works: Technical details.

If you're ready to make Azure Rights Management encryption an integrated part of your information protection solution, see Deploy an information protection solution with Microsoft Purview.

Requirements for the Azure Rights Management service

08/31/2025

Before you start to encrypt items by using the Azure Rights Management service from Microsoft Purview Information Protection, make sure you meet all the requirements.

Firewalls and network infrastructure

If you have firewalls or similar intervening network devices that are configured to allow specific connections, the network connectivity requirements are listed in the following Microsoft 365 article: Microsoft 365 Common and Office Online.

The Azure Rights Management service has the following additional requirements:

- **If you use the Microsoft Purview Information Protection client**: To download sensitivity labels and label policies, allow the following URL over HTTPS: ***.protection.outlook.com**

- **If you use web proxies**: If your web proxy requires authentication, you must configure the proxy to use integrated Windows authentication with the user's Active Directory sign in credentials.

  To support **Proxy.pac** files when using a proxy to acquire a token, add the following new registry key:
  - **Path**: `Computer\HKEY_LOCAL_MACHINE\SOFTWARE\WOW6432Node\Microsoft\MSIP\`
  - **Key**: `UseDefaultCredentialsInProxy`
  - **Type**: `DWORD`
  - **Value**: `1`

- **TLS client-to-service connections**. Don't terminate any TLS client-to-service connections, for example to perform packet-level inspection, to the **aadrm.com** URL. Doing so breaks the certificate pinning that clients for the Azure Rights Management service use with Microsoft-managed Certificate Authorities (CAs) to help secure their communication with the Azure Rights Management service.

  To determine whether your client connection is terminated before it reaches the Azure Rights Management service, use the following PowerShell commands:

  PowerShell

  ```
  $request =
  [System.Net.HttpWebRequest]::Create("https://admin.na.aadrm.com/admin/admin.s
  ```

```
vc")
$request.GetResponse()
$request.ServicePoint.Certificate.Issuer
```

The result should show that the issuing CA is from a Microsoft CA, for example: `CN=Microsoft Secure Server CA 2011, O=Microsoft Corporation, L=Redmond, S=Washington, C=US`.

If you see an issuing CA name that isn't from Microsoft, it's likely that your secure client-to-service connection is being terminated and needs reconfiguration on your firewall.

- **Microsoft 365 Enhanced Configuration Service (ECS)**. The Azure Rights Management service must have access to the **config.edge.skype.com** URL, which is a Microsoft 365 Enhanced Configuration Service (ECS).

  ECS provides Microsoft the ability to reconfigure the Azure Rights Management service when it's necessary. For example, ECS is used to control the gradual rollout of features or updates, while the impact of the rollout is monitored from diagnostic data being collected.

  ECS is also used to mitigate security or performance issues with a feature or update. ECS also supports configuration changes related to diagnostic data, to help ensure that the appropriate events are being collected.

  Limiting the **config.edge.skype.com** URL may affect Microsoft's ability to mitigate errors and may affect your ability to test preview features.

  For more information, see Essential services for Office.

- **Audit logging URL network connectivity**. The Azure Rights Management service must be able to access the following URLs to support audit logs:
  - `https://*.events.data.microsoft.com`
  - `https://*.aria.microsoft.com` (Android device data only)

Coexistence with Active Directory Rights Management Services (AD RMS)

Using the on-premises version of Active Directory Rights Management Services (AD RMS) and the Azure Rights Management service side by side, in the same organization, to encrypt content by the same user in the same organization, isn't supported in AD RMS unless you're using the HYOK (hold your own key) configuration for the AD RMS root encryption key.

This scenario isn't supported for a migration to the Azure Rights Management Service. The supported migration paths are:

- From AD RMS to the Azure Rights Management service
- From the Azure Rights Management service to AD RMS

For other, non-migration scenarios, where both services are active in the same organization, both services must be configured so that only one of them allows any given user to encrypt content. Configure such scenarios as follows:

- Use redirections for an AD RMS to Azure Rights Management migration.
- If both services must be active for different users at the same time, use service-side configurations to enforce exclusivity. Use onboarding controls from the Azure Rights Management service, and an ACL on the Publish URL to set **Read-Only** mode for AD RMS.

Azure Network Security Groups and Service Tags

If you're using an Azure endpoint and an Azure Network Security Group (NSG), make sure to allow access to all ports for the following Service Tags:

- **AzureInformationProtection**
- **AzureActiveDirectory**
- **AzureFrontDoor.Frontend**

Additionally, in this case, the Azure Rights Management service also depends on the following IP addresses and ports:

- **13.107.9.198**
- **13.107.6.198**
- **2620:1ec:4::198**
- **2620:1ec:a92::198**
- **13.107.6.181**
- **13.107.9.181**
- **Port 443**, for HTTPS traffic

Make sure to create rules that allow outbound access to these specific IP addresses, and via this port.

Supported on-premises servers for the Azure Rights Management service

The following on-premises servers are supported with the Azure Rights Management service when you use the Microsoft Rights Management connector:

- Exchange Server
- SharePoint Server
- File servers that run Windows Server and use File Classification Infrastructure (FCI)

For supported versions, other requirements, and configuration steps for the connector, see Deploying the Microsoft Rights Management connector.

Supported operating systems

The following operating systems natively support the Azure Rights Management service. However, if you install an app that uses the Azure Rights Management service, such as Microsoft 365 Enterprise apps or the Microsoft Purview Information Protection client, check that app's requirements for supported operating systems.

Expand table

OS	Supported versions
Windows computers	- Windows 10 (x86, x64) - Windows 11 (x86, x64)
macOS	Minimum version of macOS 10.8 (Mountain Lion)
Android phones and tablets	Minimum version of Android 6.0
iPhone and iPad	Minimum version of iOS 11.0
Windows phones and tablets	Windows 10 Mobile

Microsoft Entra requirements

A Microsoft Entra directory is a requirement for using the Azure Rights Management service. Use an account from a Microsoft Entra directory to sign in to the Microsoft Purview portal.

If you have a subscription that includes Microsoft Purview Information Protection, your Microsoft Entra directory is automatically created for you if needed.

The following sections list additional Microsoft Entra requirements for specific scenarios.

Support for certificate-based authentication (CBA)

When iOS and Android apps support the Azure Rights Management service, they support certificate-based authentication.

For more information, see Get started with certificate-based authentication in Microsoft Entra ID with federation.

Multifactor authentication (MFA)

To use multifactor authentication (MFA) with the Azure Rights Management service, you must have at least one of the following installed:

- **Microsoft-managed tenants, with Microsoft Entra ID or Microsoft 365**. Configure Azure MFA to enforce MFA for users.

  For more information, see:
  - Plan a Microsoft Entra multifactor authentication deployment
  - How it works: Microsoft Entra multifactor authentication

- **Federated tenants, where federation servers operate on-premises**. Configure your federation servers for Microsoft Entra ID or Microsoft 365. For example, if you're using Active Directory Federation Services (AD FS), see Configure Additional Authentication Methods for AD FS.

Rights Management connector and MFA

The Rights Management connector and the Microsoft Purview Information Protection scanner don't support MFA.

If you deploy the connector or scanner, the following accounts mustn't require MFA:

- The account that installs and configures the connector.
- The service principal account in Microsoft Entra ID, **Aadrm_S-1-7-0**, that the connector creates.
- The service account that runs the scanner.

User UPN values don't match their email addresses

Configurations where users' UPN values don't match their email addresses isn't a recommended configuration, and doesn't support single-sign on for the Azure Rights Management service.

If you can't change the UPN value, configure alternate IDs for the relevant users, and instruct them how to sign in to their Office apps by using this alternate ID.

For more information, see Configuring Alternate Login ID.

> **Tip**
>
> If the domain name in the UPN value is a domain that is verified for your tenant, add the user's UPN value as another email address to the Microsoft Entra ID **proxyAddresses** attribute. This allows the user to be authorized for the Azure Rights Management service if their UPN value is specified at the time the usage rights are granted.

For more information, see Learn how user accounts and groups use the Azure Rights Management service.

Authenticating on-premises using another authentication provider

If you're using a mobile device or Mac computer that authenticates on-premises using a non-Microsoft authentication provider, it must support the OAuth 2.0 protocol.

Advanced configurations for Entra ID

For dependent configurations in Entra that can prevent or allow access to the Azure Rights Management service, see Microsoft Entra configuration for encrypted content.

How the Azure Rights Management service works: Technical details

> *Microsoft Purview service description*

Use the following information if you want to understand, in-depth, how the Azure Rights Management service works. You don't need to know this level of information to configure or apply the encryption settings to protect your data.

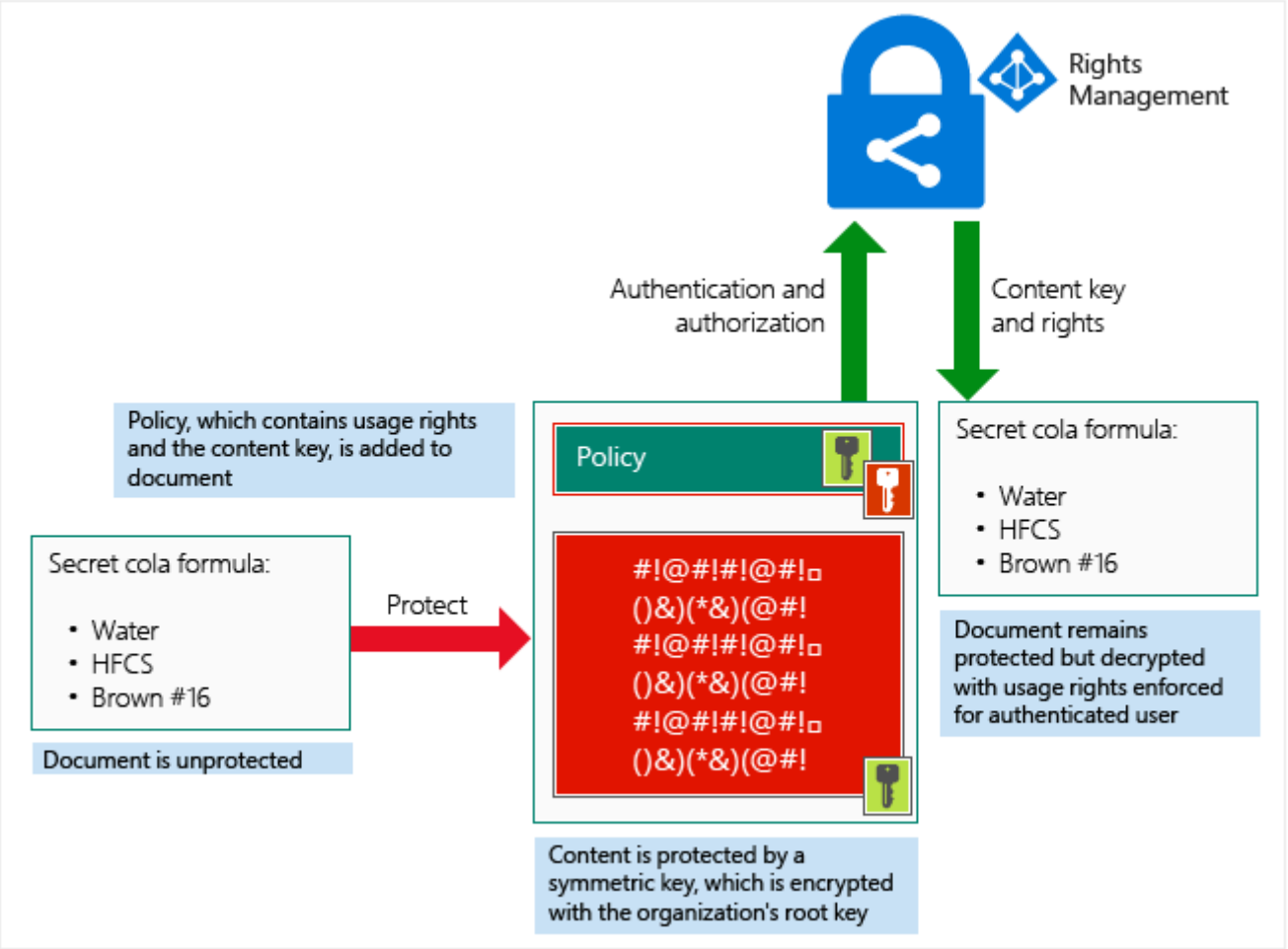
> **Note**
>
> When the Azure Rights Management service was available as a standalone product rather than part of Microsoft Purview Information Protection, it was often abbreviated to *Azure RMS*. You see this abbreviation in the pictures in this article.

An important concept to understand about the Azure Rights Management service is that this service doesn't see or store your data as part of the encryption process. Information that you encrypt is never sent to or stored in Azure, unless you explicitly store it in Azure or use another cloud service that stores it in Azure. The Azure Rights Management service simply makes the data in an item unreadable to anyone other than authorized users and services:

- The data is encrypted at the application level and includes a policy that defines the authorized use for that item.
- When an encrypted item is used by a legitimate user or it's processed by an authorized service, the data in the item is decrypted and the rights that are defined in the policy are enforced.

At a high level, you can see how this process works in the following picture. A document containing the secret formula is encrypted, and then successfully opened by an authorized user or service. The document is encrypted by a content key (the green key in this picture). The content key is unique for each document and is placed in the file header where it's encrypted by your Azure Rights Management tenant root key (the red key in this picture). Your tenant key can be generated and managed by Microsoft, or you can generate and manage your own tenant key.

Throughout the encrypted process when the Azure Rights Management service is encrypting and decrypting, authorizing, and enforcing restrictions, the secret formula is never sent to Azure.



For a detailed description of what's happening, see the Walkthrough of how the service works: First use, content encryption, content consumption section in this article.

For details about the algorithms and key lengths that the service uses, see Cryptographic controls: Algorithms and key lengths.

Cryptographic controls: Algorithms and key lengths

Even if you don't need to know in detail how this technology works, you might be asked about the cryptographic controls that it uses. For example, to confirm that the encryption is industry-standard.

Expand table

Cryptographic controls	Use in the Azure Rights Management service
Algorithm: AES Key length: 128 bits and 256 bits [1]	Content encryption

Cryptographic controls	Use in the Azure Rights Management service
Algorithm: RSA Key length: 2048 bits [2]	Key encryption
SHA-256	Certificate signing

Footnote 1: 256 bits is used by the Microsoft Purview Information Protection client in the following scenarios:

- Generic encryption (.pfile).
- Native encryption for PDF documents when the document has been encrypted with the ISO standard for PDF encryption, or the resulting encrypted document has a .ppdf file name extension.
- Native encryption for text or image files (such as .ptxt or .pjpg).

Footnote 2: 2048 bits is the key length when the Azure Rights Management service is activated. 1024 bits is supported for the following optional scenarios:

- During a migration from on-premises if the Active Directory Rights Management Services (AD RMS) cluster is running in Cryptographic Mode 1.
- For archived keys that were created on-premises before the migration, so that content that was previously encrypted by AD RMS can continue to be opened by the Azure Rights Management service post migration.

How the cryptographic keys are stored and secured

For each document or email that is protected by the Azure Rights Management service, the service creates a single AES key (the "content key"), and that key is embedded to the document, and persists through editions of the document.

The content key is encrypted with the organization's RSA key (the "Azure Rights Management tenant key") as part of the policy in the document, and the policy is also signed by the author of the document. This tenant key is common to all documents and emails that are protected by the Azure Rights Management service for the organization and the tenant key can only be changed by an administrator for the service if the organization is using a tenant key that is customer-managed (known as "bring your own key", or BYOK).

This tenant key is protected in Microsoft's online services, in a highly controlled environment and under close monitoring. When you use a customer-managed tenant key (BYOK), this security is enhanced by the use of an array of high-end hardware security modules (HSMs) in each Azure region, without the ability for the keys to be extracted, exported, or shared under any circumstances. For more information about the tenant key and BYOK, see Managing the root key for your Azure Rights Management service.

Licenses and certificates that are sent to a Windows device are encrypted with the client's device private key. This private key is created the first time a user on the device uses the Azure Rights Management service. This private key, in turn, is encrypted with DPAPI on the client, which protects these secrets by using a key derived from the user's password. On mobile devices, the keys are used only one time, so because they aren't stored on the clients, these keys don't need to be encrypted on the device.

Walkthrough of how the service works: First use, content encryption, content consumption

To understand in more detail how the Azure Rights Management service works, let's walk through a typical flow after the Azure Rights Management service is activated and when a user first uses the Rights Management service from their Windows computer (a process sometimes known as **initializing the user environment** or bootstrapping), **encrypts content** (a document or email), and then **consumes** (opens and uses) content that has been encrypted by somebody else.

After the user environment is initialized, that user can then encrypt documents or consume encrypted documents on that computer.

> **Note**
>
> If this user moves to another Windows computer, or another user uses this same Windows computer, the initialization process is repeated.

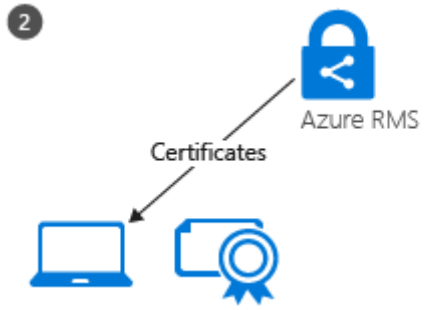
Initializing the user environment

Before a user can encrypt content or consume encrypted content on a Windows computer, the user environment must be prepared on the device. This is a one-time process and happens automatically without user intervention when a user tries to encrypt or consume encrypted content:



What's happening in step 1: The client for the Azure Rights Management service that runs on the computer first connects to the service, and the service authenticates the user by using their Microsoft Entra account.

When the user's account is federated with Microsoft Entra ID, this authentication is automatic and the user isn't prompted for credentials.



What's happening in step 2: After the user is authenticated, the connection is automatically redirected to the organization's tenant, which issues certificates that let the user authenticate to the Azure Rights Management service in order to consume encrypted content and to encrypt content offline.

One of these certificates is the rights account certificate. This certificate, often abbreviated as RAC, authenticates the user to Microsoft Entra ID and is valid for 31 days. The certificate is automatically renewed by the client, providing the user account is still in Microsoft Entra ID and the account is enabled. This certificate isn't configurable by an administrator.

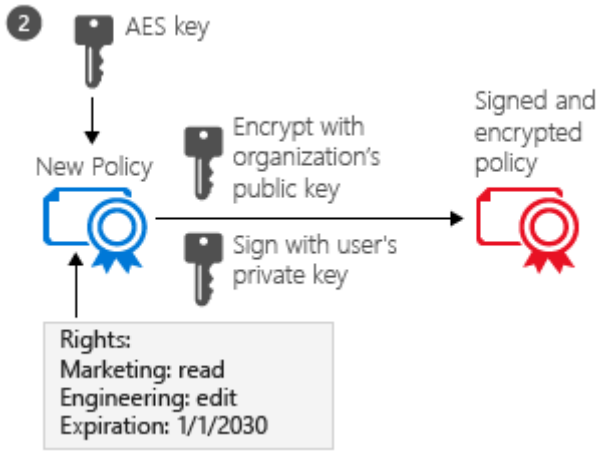
A copy of this certificate is stored in Azure so that if the user moves to another device, the certificates are created by using the same keys.

Content encryption

When a user encrypts a document, the client for the Azure Rights Management service takes the following actions on an unencrypted document:



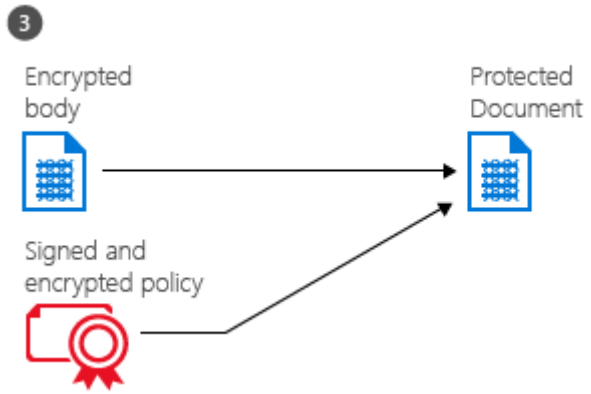
What's happening in step 1: The client creates a random key (the content key) and encrypts the document using this key with the AES symmetric encryption algorithm.



What's happening in step 2: The client then creates a certificate that includes a policy for the document that includes the usage rights for users or groups, and other restrictions, such as an expiration date. These settings can be defined with sensitivity label encryption settings that an administrator previously configured, or specified at the time the content is encrypted (sometimes referred to as "user-defined permissions" or an "ad hoc policy").

The main Microsoft Entra attribute used to identify the selected users and groups is the Microsoft Entra **ProxyAddresses** attribute, which stores all the email addresses for a user or group. However, if a user account doesn't have any values in the AD ProxyAddresses attribute, the user's **UserPrincipalName** value is used instead.

The client then uses the organization's key that was obtained when the user environment was initialized and uses this key to encrypt the policy and the symmetric content key. The client also signs the policy with the user's certificate that was obtained when the user environment was initialized.

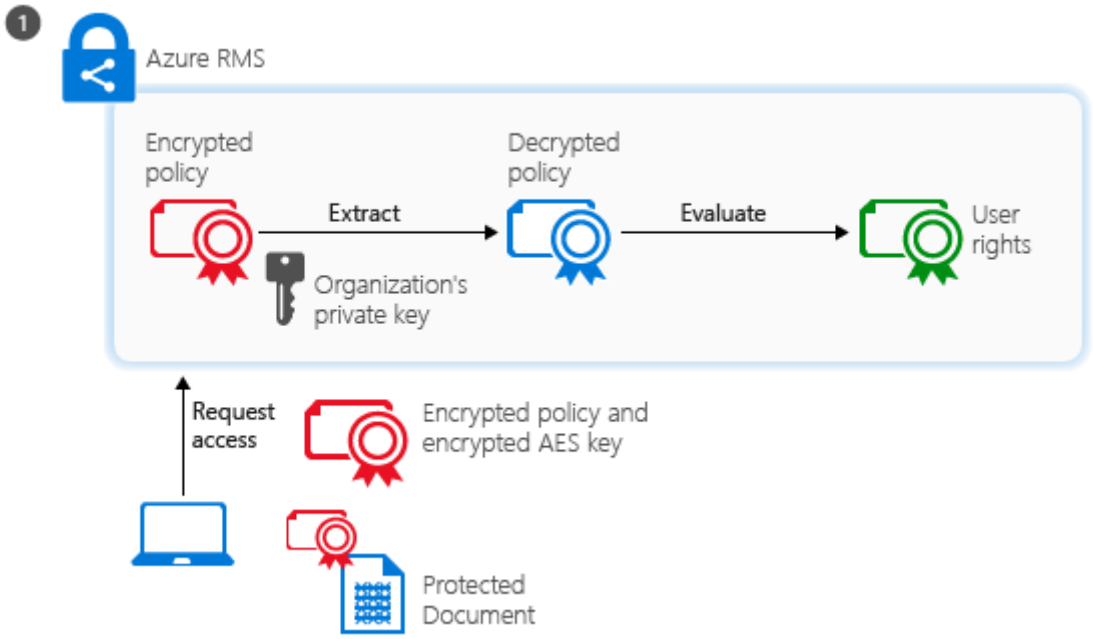


What's happening in step 3: Finally, the client embeds the policy into a file with the body of the document encrypted previously, which together comprise an encrypted document.

This document can be stored anywhere or shared by using any method, and the policy always stays with the encrypted document.

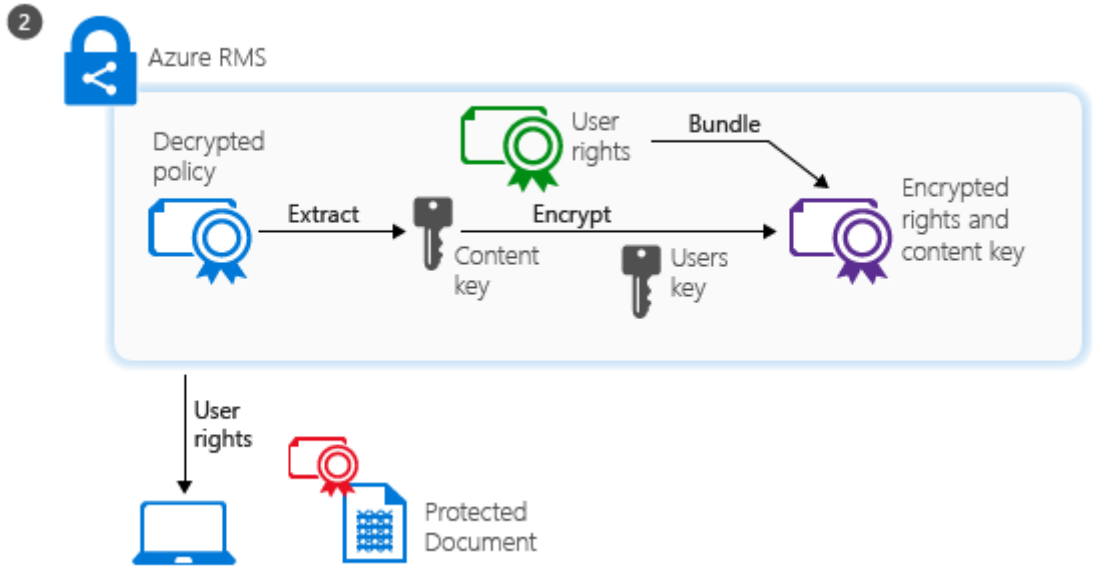
Content consumption

When a user wants to consume an encrypted document, the client starts by requesting access to the Azure Rights Management service:



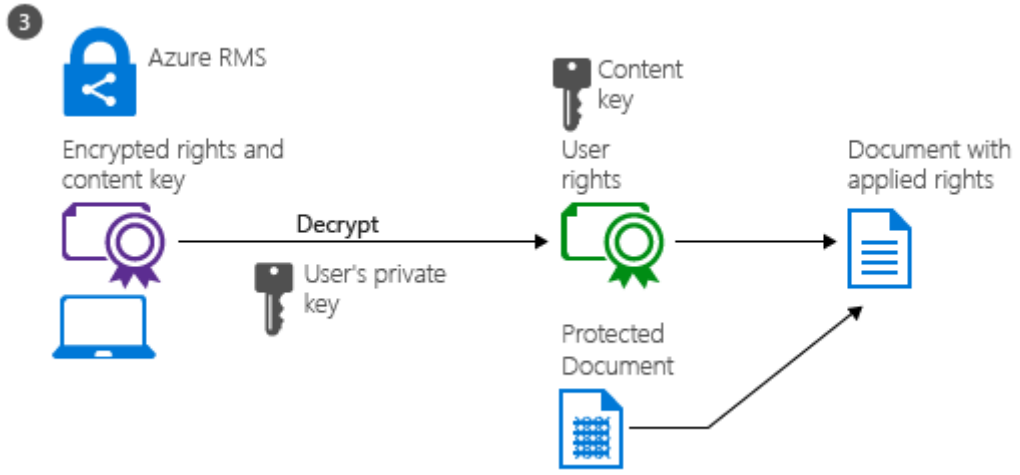
What's happening in step 1: The authenticated user sends the document policy and the user's certificates to the Azure Rights Management service. The service decrypts and evaluates the policy, and builds a list of rights (if any) the user has for the document. To identify the user, the Microsoft Entra ProxyAddresses attribute is used for the user's account and groups to which the user is a member. For performance reasons, group membership is cached. If the user account

hasn't any values for the Microsoft Entra ProxyAddresses attribute, the value in the Microsoft Entra UserPrincipalName is used instead.



What's happening in step 2: The service then extracts the AES content key from the decrypted policy. This key is then encrypted with the user's public RSA key that was obtained with the request.

The re-encrypted content key is then embedded into an encrypted use license with the list of user rights, which is then returned to the client.



What's happening in step 3: Finally, the client takes the encrypted use license and decrypts it with its own user private key. This lets the client decrypt the document's body as it is needed and render it on the screen.

The client also decrypts the rights list and passes them to the application, which enforces those rights in the application's user interface.

> **Note**
>
> When users who are external to your organization consume content that you have encrypted, the consumption flow is the same. What changes for this scenario, is how the user is authenticated. For more information, see **When I share an encrypted document with somebody outside my company, how does that user get authenticated?**

Variations

The preceding walkthroughs cover the standard scenarios but there are some variations:

- **Email encryption**: When Exchange Online and Microsoft Purview Message Encryption is used to encrypt email messages, authentication for consumption can also use federation with a social identity provider or by using a one-time passcode. Then, the process flows are very similar, except that content consumption happens service-side in a web browser session over a temporarily cached copy of the outbound email.

- **Mobile devices**: When mobile devices encrypt or consume files with the Azure Rights Management service, the process flows are simpler. Mobile devices don't first go through the user initialization process because instead, each transaction (to encrypt or consume content) is independent. As with Windows computers, mobile devices connect to the Azure Rights Management service and authenticate. To encrypt content, mobile devices submit a policy and the Azure Rights Management service sends them a publishing license and symmetric key to encrypt the document. To consume encrypted content, when mobile devices connect to the Azure Rights Management service and authenticate, they send the document policy to the Azure Rights Management service and request a use license to consume the document. In response, the Azure Rights Management service sends the necessary keys and restrictions to the mobile devices. Both processes use TLS to protect the key exchange and other communications.

- **Rights Management connector**: When the Azure Rights Management service is used with the Rights Management connector, the process flows remain the same. The only difference is that the connector acts as a relay between the on-premises services (such as Exchange Server and SharePoint Server) and the Azure Rights Management service. The connector itself doesn't perform any operations, such as the initialization of the user environment, or encryption or decryption. It simply relays the communication that would usually go to an AD RMS server, handling the translation between the protocols that are used on each side. This scenario lets you use the Azure Rights Management service with on-premises services.

- **Generic encryption (.pfile)**: When the Azure Rights Management service generically encrypts a file, the flow is basically the same for content encryption except that the client creates a policy that grants all rights. When the file is consumed, it's decrypted before it's passed to the target application. This scenario lets you encrypt all files, even if they don't natively support the Azure Rights Management service.

- **Microsoft accounts**: The Azure Rights Management service can authorize email addresses for consumption when they're authenticated with a Microsoft account. However, not all applications can open encrypted content when a Microsoft account is used for authentication. Supported scenarios for opening protected documents.

Next steps

For a less technical overview of the Azure Rights Management service, see Learn about the Azure Rights Management service.

If you're ready for deployment recommended steps that include encryption to protect your data, see Deploy an information protection solution with Microsoft Purview.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/22/2026

Compliance and supporting information for the Azure Rights Management service

> *Microsoft Purview service description*

The Azure Rights Management service supports other services and also relies on other services. If you're looking for information that is related to this encryption service but not about how to use the service, check the following resources:

Suitability for different countries/regions

Given the variability between laws and regulations in different countries/regions, different use cases and scenarios, and the varying requirements between different business sectors, consult your legal adviser to help you answer whether the Azure Rights Management service is suitable for your country/region.

However, some relevant information that can help your legal adviser make a determination:

- The Azure Rights Management service uses AES 256 and AES 128 to encrypt documents. For more information, see Cryptographic controls, algorithms, and key lengths.
- All encryption keys used by the Azure Rights Management service are protected with a customer-specific root key that uses RSA 2048 bits. RSA 1024 bits is also supported for backwards compatibility. For more information, see Cryptographic controls, algorithms, and key lengths.
- Customer-specific root keys are either managed by Microsoft or provisioned by the customer in a nCipher HSM by using "bring your own key" (BYOK). The Azure Rights Management service also supports features for on-premises encryption, for content that cannot be encrypted with a cloud-based key. For more information, see Managing the root key for your Azure Rights Management service.
- The Azure Rights Management service is hosted in regional data centers across the globe. The Azure Rights Management service keys always remain within the region in which the service was originally deployed.

- The Azure Rights Management service doesn't transmit document contents from clients to the Azure Rights Management service. Content encryption and decryption operations are performed in-place in the client device. Or, for service-based rendering, these operations are performed within the service that's rendering the content. For more information, see How the Azure Rights Management service works.

Legal and privacy

- For Microsoft Azure agreement information: Microsoft Azure Agreement
- For Microsoft Azure privacy information: Microsoft Azure Privacy Statement

Security, compliance, and auditing

See the Security, compliance, and regulatory requirements section in the Learn about the Azure Rights Management service article, for information about specific certifications for the Azure Rights Management service. In addition:

- For external certifications for the Azure Rights Management service: Microsoft Azure Trust Center
- For FIPS 140 information: FIPS 140 Validation

For in-depth technical information about how the encryption technology works, see How the Azure Rights Management service works: Technical details.

Service level agreements (SLAs)

The following service level agreement resources apply to the Azure Rights Management service and its dependencies:

- SLA for the Azure Rights Management service
- SLA for Microsoft Entra ID
- SLA for Azure Key Vault

Documentation

For related product documentation, see the following resources:

- Microsoft Entra documentation: Microsoft Entra ID

- Microsoft 365 documentation: Microsoft 365 for enterprise documentation and resources

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/22/2026

Microsoft Entra configuration for encrypted content

Service description

> *Microsoft Purview service description*

If you protect sensitive items such as emails and documents by using encryption from the Azure Rights Management service from Microsoft Purview Information Protection, there are some Microsoft Entra configurations that can prevent authorized access to this encrypted content.

Similarly, if your users receive encrypted email from another organization or collaborate with other organizations that encrypt documents by using the Azure Rights Management service, your users might not be able to open that email or document because of how their Microsoft Entra ID is configured.

For example:

- A user can't open encrypted email sent from another organization. Or, a user reports that the recipients in another organization can't open an encrypted email that they sent them.
- Your organization collaborates with another organization on a joint project, and project documents are protected by encrypting them, granting access by using groups in Microsoft Entra ID. Users can't open the documents encrypted by users in the other organization.
- Users can successfully open an encrypted document when they are in the office, but can't when they try to access this document remotely and they're prompted for multifactor authentication (MFA).

To ensure access to the encryption service isn't inadvertently blocked, use the following sections to help configure your organization's Microsoft Entra ID, or relay the information to a Microsoft Entra administrator in another organization. Without access to the Azure Rights Management service, users can't be authenticated and aren't authorized to open encrypted content.

Cross-tenant access settings and encrypted content

> **Important**

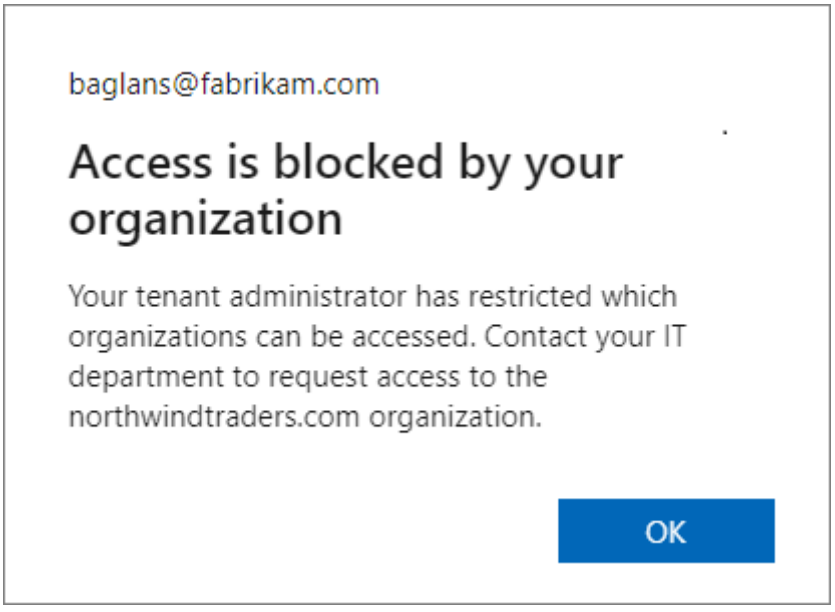
Another organization's cross-tenant access settings can be responsible for either their users being unable to open content that your users have encrypted, or your users being unable to open content encrypted by the other organization.

The message that users see indicates which organization blocked access. You might need to direct the Microsoft Entra admin from another organization to this Cross-tenant access settings and encrypted content section.

By default, there's nothing to configure for cross-tenant authentication to work when users protect content by using encryption from the Azure Rights Management service. However, your organization can restrict access by using Microsoft Entra External Identities cross-tenant access settings. Conversely, another organization can also configure these settings to restrict access with users in your organization. These settings affect opening any encrypted items, which include encrypted emails and encrypted documents.

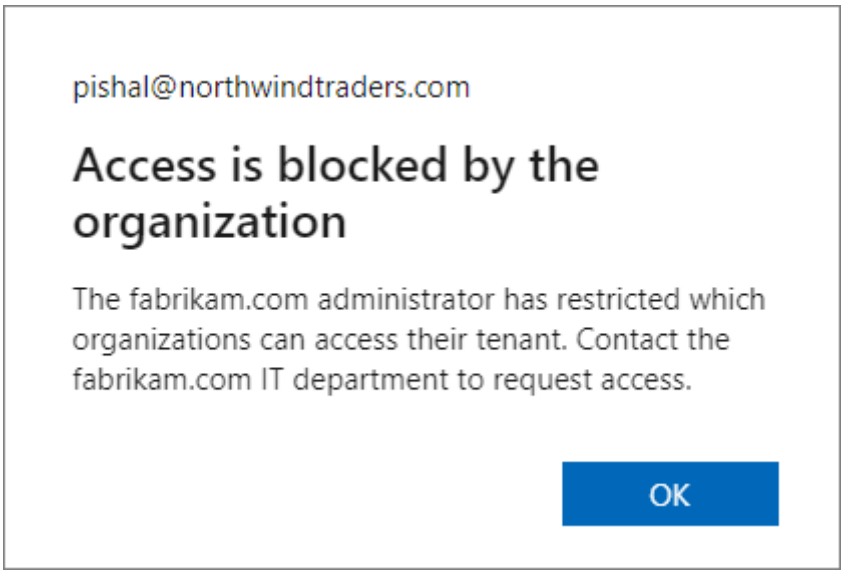
For example, another organization might have settings configured that prevent their users from opening content encrypted by your organization. In this scenario, until their Microsoft Entra admin reconfigures their cross-tenant settings, an external user attempting to open that content sees a message that informs them **Access is blocked by your organization** with a reference to **Your tenant administrator**.

Example message for the signed in user from the Fabrikam, Inc organization, when their local Microsoft Entra ID blocks access:



Your users will see a similar message when it's your Microsoft Entra configuration that blocks access.

From the perspective of the signed in user, if it's another Microsoft Entra organization that's responsible for blocking access, the message changes to **Access is blocked by the organization** and displays the domain name of that other organization in the body of the message. For example:



Whenever cross-tenant access settings restrict access by applications, they must be configured to allow access to the rights management service, which has the following app ID:

```plaintext
00000012-0000-0000-c000-000000000000
```

If access to the rights management service isn't allowed, users can't be authenticated and authorized to open encrypted content. This configuration can be set as a default setting and as an organizational setting:

- To permit sharing of encrypted content with another organization, create an inbound setting that allows access to Microsoft Rights Management Service (ID: 00000012-0000-0000-c000-000000000000).
- To permit access to encrypted content that users receive from other organizations, create an outbound setting that allows access to Microsoft Rights Management Service (ID: 00000012-0000-0000-c000-000000000000)

When these settings are configured for the encryption service, the application displays **Microsoft Rights Management Services**.

For instructions to configure these cross-tenant access settings, see Configure cross-tenant access settings for B2B collaboration.

If you configured Microsoft Entra Conditional Access policies that require multifactor authentication (MFA) for users, see Conditional Access policies and encrypted documents.

Conditional Access policies and encrypted documents

If your organization has implemented Microsoft Entra Conditional Access policies that include **Microsoft Rights Management Services** and the policy extends to external users who need to open documents encrypted by your organization:

- For external users who have a Microsoft Entra account in their own tenant, we recommend you use External Identities cross-tenant access settings to configure trust settings for MFA claims from one, many, or all external Microsoft Entra organizations.
- For external users not covered by the previous entry, for example, users who don't have a Microsoft Entra account or you haven't configured cross-tenant access settings for trust settings, these external users must have a guest account in your tenant.

Without one of these configurations, external users can't open the encrypted content and see an error message. The message text might inform them that their account needs to be added as an external user in the tenant, with the incorrect instruction for this scenario to **Sign out and sign in again with a different Microsoft Entra user account**.

If you can't meet these configuration requirements for external users who need to open content encrypted by your organization, you must either remove Microsoft Rights Management Services from the Conditional Access policies, or exclude external users from the policies.

For more information, see the frequently asked question, I see Microsoft Rights Management Services is listed as an available cloud app for conditional access—how does this work?

Guest accounts for external users to open encrypted documents

You might need guest accounts in your Microsoft Entra tenant for external users to open documents encrypted by your organization. Options to create the guest accounts:

- Create these guest accounts yourself. You can specify any email address that these users already use. For example, their Gmail address.

  The advantage of this option is that you can restrict access and rights to specific users by specifying their email address in the encryption settings. The downside is the administration

overhead for the account creation and coordination with the label configuration.

- Use SharePoint and OneDrive integration with Microsoft Entra B2B so that guest accounts are automatically created when your users share links.

  The advantage of this option is minimum administrative overhead because the accounts are created automatically, and simpler label configuration. For this scenario, you must select the encryption option Add any authenticated user because you won't know the email addresses in advance. The downside is that this setting doesn't let you restrict access and usage rights to specific users.

External users can also use a Microsoft account to open encrypted documents when they use Windows and Microsoft 365 Apps (formerly Office 365 apps; see Microsoft 365 Apps name change) or the standalone edition of Office 2019. More recently supported for other platforms, Microsoft accounts are also supported for opening encrypted documents on macOS (Microsoft 365 Apps, version 16.42+), Android (version 16.0.13029+), and iOS (version 2.42+).

For example, a user in your organization shares an encrypted document with a user outside your organization, and the encryption settings specify a Gmail email address for the external user. This external user can create their own Microsoft account that uses their Gmail email address. Then, after signing in with this account, they can open the document and edit it, according to the usage restrictions specified for them. For a walkthrough example of this scenario, see Opening and editing the protected document.

> **Note**
>
> The email address for the Microsoft account must match the email address specified to restrict access for the encryption settings.

When a user with a Microsoft account opens an encrypted document in this way, it automatically creates a guest account for the tenant if a guest account with the same name doesn't already exist. When the guest account exists, it can then be used to open documents in SharePoint and OneDrive by using Office for the web, in addition to opening encrypted documents from the supported desktop and mobile Office apps.

However, the automatic guest account isn't created immediately in this scenario, because of replication latency. If you specify personal email addresses as part of your encryption settings, we recommend that you create corresponding guest accounts in Microsoft Entra ID. Then, let these

users know that they must use this account to open an encrypted document from your organization.

> **Tip**
>
> Because you can't be sure that external users will be using a supported Office client app, sharing links from SharePoint and OneDrive after creating guest accounts (for specific users) or when you use **SharePoint and OneDrive integration with Microsoft Entra B2B** (for any authenticated user) is a more reliable method to support secure collaboration with external users.

Cross-cloud access settings and encrypted content

Cross-tenant access settings enable cross-cloud access to encrypted documents. As a result, users who are part of an organization in another cloud environment can open Word, Excel, and PowerPoint files that are encrypted by your organization. For example, these users might be in Microsoft Azure Government, or Microsoft Azure China (operated by 21Vianet).

Supported scenarios

The following scenarios are supported for cross-cloud access to encrypted content:

- Organizations in the Microsoft Azure Commercial cloud can allow access to organizations in the Microsoft Azure Government cloud or Microsoft Azure China (operated by 21Vianet) cloud without contacting Microsoft.
- Organizations in the Microsoft Azure China (operated by 21Vianet) cloud can allow access to organizations in the Microsoft Azure Commercial cloud without contacting Microsoft.
- Encrypted Word, Excel, and PowerPoint files can be shared with users in another cloud environment and viewed on Windows, MacOS, and mobile devices. This scenario doesn't apply to files shared via SharePoint or OneDrive sharing links, but rather to methods like direct email attachments, USB drive, file share, or shared and downloaded from SharePoint.
- Encrypted PDFs and generically encrypted files (.pfile extension) can be shared with users in another cloud environment and viewed using the Microsoft Purview Information Protection Viewer on Windows.
- The user can be, but isn't required to be, a guest in your organization.
- External users and domains can be configured for sensitivity labels that apply encryption when you use admin-defined permissions, or specified by users when the labels are configured for user-defined permissions.

Unsupported scenarios

The following scenarios are currently unsupported for cross-cloud access to encrypted content:

- Email isn't supported, including all Outlook clients.
- PDF viewers such as Microsoft Edge or Adobe Acrobat are currently unsupported.

Requirements

Before you configure cross-cloud access, ensure the following requirements are met:

- Microsoft 365 Apps version 2402 or later or Microsoft Office 2024.
- Microsoft Purview Information Protection client 3.1.310.0 or later.
- Administrators on both sides must:
  - Configure cross-tenant access settings to enable cross-cloud access
  - Add the partner organization as a permitted organization
- Users opening encrypted files must have an email address assigned.

Cross-tenant access settings: Enable cross-cloud collaboration

There's a one-time step required to enable collaboration across different cloud environments:

1. In the Microsoft Entra admin center, navigate to **External identities** > **Cross-tenant access settings**.
2. Select **Microsoft cloud settings**.
3. Select the cloud environment you need to enable for collaboration:

   - Organizations in the Commercial Cloud see both **Microsoft Azure Government** and **Microsoft Azure China (operated by 21Vianet)**.
   - Organizations in a sovereign cloud see only **Microsoft Azure Commercial Cloud**.

4. Select the cloud environment where your partner organization is hosted, and select **Save**.

> **Note**
>
> Both organizations must complete this step, and select the partner cloud.

Cross-tenant access settings: Configure organizational settings

To configure organizational settings:

1. In the Microsoft Entra admin center, navigate to **External identities** > **Cross-tenant access settings**.
2. Select **Organizational settings**.
3. Select **Add organization**.
4. Enter the **Microsoft Entra tenant ID** for the partner organization.
5. Select **Add**.

The tenant ID is displayed in the **Overview** section of the Microsoft Entra admin center and should be provided by the partner organization.

Cross-tenant access settings: Trust external MFA claims

If external users aren't invited as guests, inbound trust settings should be configured to trust external MFA claims. If users are invited as guests and you trust external MFA claims, it's recommended to exclude external users from MFA registration in your organization.

Review Manage cross-tenant access settings for B2B collaboration for configuration details.

Cross-tenant access settings: Inbound and outbound access settings

The steps to enable the partner cloud and add the partner organization are sufficient to enable both organizations to consume protected content from the partner organization. If you'd like to restrict the cross-cloud collab to only permit document decryption, excluding other services, or to make the relationship one-way, review Cross-tenant access settings and encrypted content.

Changes to domain-based sharing

When the encryption scenario is confined to the same cloud environment, specifying a single domain for the label encryption settings results in authenticating and authorizing all users from the organization that owns the specified domain, regardless of the users' email domain. For more information about this domain-based access, see the note in Restrict access to content by using sensitivity labels to apply encryption.

However, the cross-cloud functionality relies on data embedded in Entra access tokens. The feature makes use of two optional claims: `verified_primary_email` and `verified_secondary_email`.

During authorization, these claims are extracted and evaluated against the permissions defined in the label. When the rights management service evaluates domain-based access, it can evaluate

only the email suffixes that were provided in the access token. This results in a scenario where the admin-defined permissions or user-defined permissions must contain a domain name that maps exactly to a user's email suffix.

Cross-cloud common issues

External users see "Access is blocked by your organization" when they try to open encrypted files

Cross-tenant access settings aren't configured correctly on one or both sides. Review the Cross-tenant access settings and encrypted content section and the Cross-tenant access settings: Inbound and outbound access settings section.

AADSTS90072: Nonguest users see "User account {user@contoso.com} from identity provider 'microsoftonline.com' does not exist in..."

This error indicates that external MFA claims aren't trusted. For configuration details, see Manage cross-tenant access settings for B2B collaboration.

Next steps

For addition configurations you might need to make, see Restrict access to a tenant. Specific to network infrastructure configuration for the Azure Rights Management service, see Firewalls and network infrastructure.

If you use sensitivity labels to encrypt documents and emails, you might be interested in Support for external users and labeled content to understand which label settings apply across tenants. For configuration guidance for the label encryption settings, see Restrict access to content by using sensitivity labels to apply encryption.

Interested to learn how and when the encryption service is accessed? See Walkthrough of how the service works: First use, content encryption, content consumption.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/22/2026

Learn how user accounts and groups use the Azure Rights Management service

08/26/2025

Before you use the Azure Rights Management service to encrypt content for your organization, understand how the service works with accounts for users and groups in Microsoft Entra ID.

There are different ways to create these accounts for users and groups, which include:

- You create the users in the Microsoft 365 admin center, and the groups in the Exchange Online admin center.
- You create the users and groups in the Azure portal.
- You create the users and group by using PowerShell cmdlets.
- You create the users and groups in your on-premises Active Directory and synchronize them to Microsoft Entra ID.
- You create the users and groups in another directory and synchronize them to Microsoft Entra ID.

When you create users and groups by using the first three methods from this list, with one exception, they're automatically created in Microsoft Entra ID, and the Azure Rights Management service can use these accounts directly. However, some enterprise networks use an on-premises directory to create and manage users and groups. The Azure Rights Management service can't use these accounts directly; you must synchronize them to Microsoft Entra ID.

The exception referred to in the previous paragraph is dynamic distribution lists that you can create for Exchange Online. Unlike static distribution lists, these groups aren't replicated to Microsoft Entra ID and so can't be used by the Azure Rights Management service.

How users and groups are used by the Azure Rights Management service

There are two scenarios for using users and groups with the Azure Rights Management service:

For encryption settings when you use the Azure Rights Management service to encrypt documents and emails. Administrators and users can select users and groups that can open the encrypted content, and additionally:

- Usage rights that determine how they can use the content. For example, whether they can only read it, or read and print it, or read and edit it.
- Access controls include an expiry date and whether a connection to the internet is required for access.

For configuring the Azure Rights Management service to support specific scenarios, and therefore only administrators select these groups. Examples include configuring the following:

- Super users, so that designated services or people can open encrypted content if required for eDiscovery or data recovery.
- Delegated administration of the Azure Rights Management service.
- Onboarding controls to support a phased deployment.

Azure Rights Management service requirements for user accounts

For encryption settings, and configuring the Azure Rights Management service:

- To authorize users, two attributes in Microsoft Entra ID are used: **proxyAddresses** and **userPrincipalName**.
- The **Microsoft Entra proxyAddresses** attribute stores all email addresses for an account and can be populated in different ways. For example, a user in Microsoft 365 that has an Exchange Online mailbox automatically has an email address that is stored in this attribute. If you assign an alternative email address for a Microsoft 365 user, it is also saved in this attribute. It can also be populated by the email addresses that are synchronized from on-premises accounts.

  The Azure Rights Management service can use any value in this Microsoft Entra proxyAddresses attribute, providing the domain has been added to your tenant (a "verified domain"). For more information about verifying domains:

  - For Microsoft Entra ID: Add a custom domain name to Microsoft Entra ID
  - For Microsoft 365: Add a domain to Microsoft 365
- The **Microsoft Entra userPrincipalName** attribute is used only when an account in your tenant doesn't have values in the Microsoft Entra proxyAddresses attribute. For example, you create a user in the Azure portal, or create a user for Microsoft 365 that doesn't have a mailbox.

Encryption settings for external users

In addition to using the Microsoft Entra proxyAddresses and Microsoft Entra userPrincipalName for users in your tenant, the Azure Rights Management service also uses these attributes in the same way to authorize users from another tenant.

Other authorization methods:

- For email addresses that aren't in Microsoft Entra ID, the Azure Rights Management service can authorize these when they're authenticated with a Microsoft account. However, not all applications can open encrypted content when a Microsoft account is used for authentication.
- When an email is sent by using Microsoft Purview Message Encryption to a user who doesn't have an account in Microsoft Entra ID, the user is first authenticated by using federation with a social identity provider or by using a one-time passcode. Then the email address specified in the protected email is used to authorize the user.

Azure Rights Management service requirements for group accounts

For assigning usage rights and access controls:

- You can use any type of group in Microsoft Entra ID that has an email address that contains a verified domain for the user's tenant. A group that has an email address is often referred to as a mail-enabled group.

For configuring the Azure Rights Management service:

- You can use any type of group in Microsoft Entra ID that has an email address from a verified domain in your tenant, with one exception. That exception is when you configure onboarding controls to use a group, which must be a security group in Microsoft Entra ID for your tenant.
- You can use any group in Microsoft Entra ID (with or without an email address) from a verified domain in your tenant for delegated administration of the Azure Rights Management service.

Encryption settings for external groups

In addition to using the Microsoft Entra proxyAddresses for groups in your tenant, the Azure Rights Management service also uses this attribute in the same way to authorize groups from

another tenant.

Using accounts from Active Directory on-premises

If you have accounts that are managed on-premises that you want to use with the Azure Rights Management service, you must synchronize these to Microsoft Entra ID. For ease of deployment, we recommend that you use Microsoft Entra Connect. However, you can use any directory synchronization method that achieves the same result.

When you synchronize your accounts, you don't need to synchronize all attributes. For a list of the attributes that must be synchronized, see the Azure RMS section from the Microsoft Entra documentation.

From the attributes list for Azure Rights Management, you see that for users, the on-premises AD attributes of **mail**, **proxyAddresses**, and **userPrincipalName** are required for synchronization. Values for **mail** and **proxyAddresses** are synchronized to the Microsoft Entra proxyAddresses attribute. For more information, see How the proxyAddresses attribute is populated in Microsoft Entra ID

Considerations if email addresses change

If you change the email address of a user or group, we recommend that you add the old email address as a second email address (also known as a proxy address, alias, or alternate email address) to the user or group. When you do this, the old email address is added to the Microsoft Entra proxyAddresses attribute. This account administration ensures business continuity for any usage rights or other configurations there were saved when the old email address was in use.

If you can't do this, the user or group with the new email address risks being denied access to documents and emails that were previously protected with the old email address. In this case, you must repeat the protection configuration to save the new email address. For example, if the user or group was granted usage rights in templates or labels, edit those templates or labels and specify the new email address with same usage rights as you granted to the old email address.

Note that it's rare for a group to change its email address and if you assign usage rights to a group rather to than individual users, it doesn't matter if the user's email address changes. In this scenario, the usage rights are assigned to the group email address and not individual user email addresses. This is the most likely (and recommended) method for an administrator to configure usage rights that protect documents and emails. However, users might more typically assign custom permissions for individual users. Because you can't always know

whether a user account or group has been used to grant access, it's safest to always add the old email address as a second email address.

Group membership caching

For performance reasons, the Azure Rights Management service caches group membership. This caching means that any changes to group membership in Microsoft Entra ID can take up to three hours to take effect when these groups are used by the Azure Rights Management service and this time period is subject to change.

Remember to factor this delay into any changes or testing that you do when you use groups for granting usage rights or configuring the Azure Rights Management service.

Next steps

When you have confirmed that your users and groups can be used with the Azure Rights Management service, check whether you need to activate the Azure Rights Management service:

- Beginning with February 2018: If your subscription that includes Azure Rights Management or Microsoft Purview Information Protection was obtained during or after this month, the service is automatically activated for you.
- If your subscription was obtained before February 2018: You must activate the service yourself.

For more information, which includes checking the activation status, see Activate the Azure Rights Management service.

Administering the Azure Rights Management service by using PowerShell

> *Microsoft Purview service description*

PowerShell is required to manually activate the Azure Rights Management encryption service, and you also need PowerShell for some advanced configurations. Some configurations can be done from the Microsoft Purview portal, however, you might prefer to use PowerShell for more efficient command-line control and scripting.

The following table lists some of the advanced configuration scenarios that use PowerShell. When the configuration can also be completed without using PowerShell, this information is also included in the table.

For a complete list of the available cmdlets for this module, with more information about each one, see AIPService.

To install this PowerShell module, see Install the AIPService PowerShell module for the Azure Right Management service.

> **Tip**
>
> In addition to this service-side PowerShell module, the Microsoft Purview Information Protection client installs a supplemental PowerShell module, **PurviewInformationProtection**.
>
> This client module supports labeling and encrypting multiple files so that, for example, you can bulk-protect all files in a folder. For more information, see **Set up the information protection client using PowerShell**.

Cmdlets grouped by administration task

The following table groups common administration tasks by the cmdlets used to complete them.

Expand table

If you need to...	...use the following cmdlets
Migrate from on-premises Rights Management (AD RMS or Windows RMS) to the Azure Rights Management service from Microsoft Purview Information Protection.	Import-AipServiceTpd Set-AipServiceKeyProperties
Connect to or disconnect from the Rights Management service for your organization.	Connect-AipService Disconnect-AipServiceService
Generate and manage your own root key for the Azure Rights Management service – the bring your own key (BYOK) scenario.	Set-AipServiceKeyProperties Use-AipServiceKeyVaultKey Get-AipServiceKeys
Activate or deactivate the Azure Rights Management service for your organization.	Enable-AipService Disable-AipService
Configure onboarding controls for a phased deployment of the Azure Rights Management service.	Get-AipServiceOnboardingControlPolicy Set-AipServiceOnboardingControlPolicy
Create and manage Rights Management templates for your organization.	Add-AipServiceTemplate Export-AipServiceTemplate Get-AipServiceTemplate Get-AipServiceTemplateProperty Import-AipServiceTemplate New-AipServiceRightsDefinition Remove-AipServiceTemplate Set-AipServiceTemplateProperty
Configure the maximum number of days that content that your organization encrypts can be accessed without an internet connection (the use license validity period).	Get-AipServiceMaxUseLicenseValidityTime Set-AipServiceMaxUseLicenseValidityTime
Manage the super user feature of the Azure Rights Management service for your organization.	Enable-AipServiceSuperUserFeature Disable-AipServiceSuperUserFeature

If you need to...	...use the following cmdlets
	Add-AipServiceSuperUser Get-AipServiceSuperUser Remove-AipServiceSuperUser Set-AAipServiceSuperUserGroup Get-AipServiceSuperUserGroup Clear-AipServiceSuperUserGroup
Manage users and groups who are authorized to administer the Azure Rights Management service for your organization.	Add-Aip-ServiceRoleBasedAdministrator Get-Aip-ServiceRoleBasedAdministrator Remove-Aip-ServiceRoleBasedAdministrator
Get a log of the Azure Rights Management service administrative tasks for your organization.	Get-AipServiceAdminLog
Log and analyze usage logging for the Azure Rights Management service.	Get-AipServiceUserLog
Display the current configuration of the Azure Rights Management service for your organization.	Get-AipServiceConfiguration
Migrate your organization from the Azure Rights Management service to an on-premises AD RMS deployment.	Set-AipServiceMigrationUrl Get-AipServiceMigrationUrl
Manage the legacy document tracking site for rights-protected documents	Disable-AipServiceDocumentTrackingFeature Enable-AipServiceDocumentTrackingFeature Get-AipServiceDocumentTrackingFeature Set-AipServiceDoNotTrackUserGroup Clear-AipServiceDoNotTrackUserGroup Get-AipServiceDoNotTrackUserGroup

If you need to...	...use the following cmdlets
	Get-AipServiceTrackingLog Get-AipServiceDocumentLog

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/22/2026

Install the AIPService PowerShell module for the Azure Rights Management service

> *Microsoft Purview service description*

Use the following information to help you install the PowerShell module for the Azure Right Management encryption service from Microsoft Purview Information Protection. The name of this module is AIPService, and it replaces the previous version that was named AADRM.

You can use this PowerShell module to administer the Azure Rights Management encryption service from the command line by using any Windows computer that has an internet connection and that meets the prerequisites listed in the following table. PowerShell for the Azure Rights Management service supports scripting for automation or might be necessary for advanced configuration scenarios. For more information about the administration tasks and configurations that the AIPService module supports, see Administering the Azure Rights Management service by using PowerShell.

Prerequisites

This table lists the prerequisites to install and use the AIPService PowerShell module for the Azure Rights Management encryption service.

Expand table

Requirement	More information
Supported PowerShell versions	Windows PowerShell 5.1, or PowerShell 7.4 and later. You can confirm the version you're running by typing `$PSVersionTable` in a PowerShell session. The `PSEdition` property shows `Desktop` for Windows PowerShell and `Core` for PowerShell 7. To install PowerShell 7, see Installing PowerShell on Windows. PowerShell 7 requires AIPService version 3.1.0.0 or later. Earlier versions of the module support Windows PowerShell only.
Minimum .NET version	Windows PowerShell 5.1 requires Microsoft .NET Framework 4.8. This minimum version of the Microsoft .NET Framework is included with the later operating systems but if it's not already installed, you can download Microsoft .NET Framework 4.8. This minimum version of the Microsoft .NET Framework is required

Requirement	More information
	for some of the classes that the AIPService module uses. PowerShell 7.4 and later includes the .NET version that the module requires, so no separate installation is needed.
Required license	You must have a product license to use this feature. For more information, see Microsoft 365 licensing guidance for security & compliance - Service Descriptions.

PowerShell editions

AIPService version 3.1.0.0 supports both PowerShell editions on Windows from a single package:

Expand table

Edition	Host	Runtime
Desktop	Windows PowerShell 5.1	.NET Framework 4.8
Core	PowerShell 7.4 and later	.NET 8

The module selects the correct edition automatically when you import it, so the same `Install-Module` command works in both hosts. Cmdlet names, parameters, and output are the same in both editions.

Interactive sign-in differs between the two editions. Windows PowerShell 5.1 signs in through an embedded web view, while PowerShell 7 uses your default web browser and a local loopback address. On a computer with no default browser available, such as a Server Core installation, or in a remote session, use certificate-based or service principal authentication instead. For more information, see Connect-AipService.

To confirm which editions your installed version supports, run:

PowerShell

```
(Get-Module AIPService -ListAvailable).CompatiblePSEditions
```

If you have the AADRM module installed

The AIPService module replaces the older module, AADRM. If you have the older module installed, uninstall it and then install the AIPService module.

The AIPService module has aliases to the cmdlet names in the AADRM module so that any existing scripts will continue to work. The AADRM module fell out of support on July 15, 2020.

If you installed the AADRM module from the PowerShell Gallery, to uninstall it, start a PowerShell session with the **Run as Administrator** option, and type:

```
PowerShell
Uninstall-Module -Name AADRM
```

If you installed the AADRM module with the Azure Rights Management Administration Tool, use your Windows remove applications procedures to uninstall **Windows Azure AD Rights Management Administration**.

How to install the AIPService module

The AIPService module is on the PowerShell Gallery and is not available from the Microsoft Download Center.

To install the AIPService module from the PowerShell Gallery

If you're new to the PowerShell Gallery, see Get Started with the PowerShell Gallery. Follow the PowerShell Gallery setup instructions, including installing the PowerShellGet module and the NuGet provider.

To see details about the AIPService module on the PowerShell Gallery, visit the AIPService package in PowerShell Gallery.

To install the AIPService module, start a PowerShell session with the **Run as Administrator** option, and type:

```
PowerShell
Install-Module -Name AIPService
```

If you are warned about installing from an untrusted repository, you can press Y to confirm. Or, press N and configure the PowerShell Gallery as a trusted repository by using the command `Set-PSRepository -Name PSGallery -InstallationPolicy Trusted` and then rerun the command to install the AIPService module.

If you have a previous version of the AIPService module installed from the Gallery, update it to the latest by typing:

PowerShell

```
Update-Module -Name AIPService
```

> ⓘ Note
>
> To use the module in PowerShell 7, run `Install-Module -Name AIPService` from a PowerShell 7 session. Windows PowerShell and PowerShell 7 use separate per-user module folders, so a module that was installed for the current user in Windows PowerShell isn't available in PowerShell 7.

Next steps

In a PowerShell session, confirm the version of the installed module. This check is particularly important if you upgraded from an older version:

PowerShell

```
(Get-Module AIPService -ListAvailable).Version
```

> ⓘ Note
>
> If `(Get-Module AIPService -ListAvailable).Version` fails, first run **Import-Module AIPService**.

To see which cmdlets are available, type the following:

Powershell

```
Get-Command -Module AIPService
```

Use the `Get-Help <cmdlet_name>` command to see the Help for a specific cmdlet, and use the **-online** parameter to see the latest help on the Microsoft documentation site. For example:

PowerShell

```
Get-Help Connect-AipService -online
```

For more information:

- Full list of cmdlets available: AIPService Module
- List of main configuration scenarios that support PowerShell: Administering the Azure Rights Management service by using PowerShell

Before you can run any commands that configure the Azure Rights Management service, you must connect to the service by using the Connect-AipService cmdlet.

When you have finished running your configuration commands, as a best practice, disconnect from the service by using the Disconnect-AipService cmdlet. If you do not disconnect, the connection is automatically disconnected after a period of inactivity. Because of the automatic disconnection behavior, you might find that you need to occasionally reconnect in a PowerShell session.

> **Note**
>
> If the Azure Rights Management service is not yet activated, you can activate it after you have connected to the service by using the **Enable-AipService** cmdlet.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 09/02/2026

Managing the root key for your Azure Rights Management service

> *Microsoft Purview service description*

The Azure Rights Management tenant key is your organization's root key for the main encryption service for Microsoft Purview Information Protection. Other keys can be derived from this root key, including user keys, computer keys, or document encryption keys. Whenever the Azure Rights Management service uses these keys for your organization, they cryptographically chain to your tenant's root key for the Azure Rights Management service.

In addition to your tenant root key, your organization may require on-premises security for specific documents. On-premises key encryption is typically required only for a small amount of content, and therefore is configured together with a tenant root key.

Your Azure Rights Management tenant key can be Microsoft-managed (the default), customer-managed through Bring Your Own Key (BYOK), or supplemented with Double Key Encryption (DKE) for highly sensitive content. Each option provides different levels of control over key life-cycle operations such as rekeying, backup, export, and breach response. To understand your options, see Key types for the service and Operations for your tenant key.

If you're migrating across tenants, such as after a company merger, we recommend that you read our blog post on mergers and spin-offs for more information.

Key types for the service

The root key for your Azure Rights Management service can either be:

- Generated by Microsoft
- Generated by customers with Bring Your Own Key (BYOK)

If you have highly sensitive content that requires additional, on-premises protection, we recommend using Double Key Encryption (DKE).

Tenant root key generated by Microsoft

The default root key, automatically generated by Microsoft, is the default key used exclusively for the Azure Rights Management service to manage most aspects of your tenant key life cycle.

Continue using the default Microsoft key when you want to use the Azure Rights Management service quickly and without special hardware, software, or an Azure subscription. Examples include testing environments or organizations without regulatory requirements for key management.

For the default key, no configuration steps are required.

> **Note**
>
> The default root key generated by Microsoft is the simplest option with the lowest administrative overheads.
>
> In most cases, you may not even know that you have a tenant key, as you can configure encryption settings for Microsoft Purview Information Protection and the key management process is handled by Microsoft.

Bring Your Own Key (BYOK) option

BYOK uses keys that are created by customers, either in the Azure Key Vault or on-premises in the customer organization. These keys are then transferred to Azure Key Vault for further management.

Use BYOK when your organization has compliance regulations for key generation, including control over all life-cycle operations. For example, when your key must be protected by a hardware security module.

For more information, see Configure BYOK.

Double Key Encryption (DKE)

DKE provides additional security for your content by using two root keys that work together: one created and held by Microsoft in Azure, and another created and held on-premises by the customer.

DKE requires both keys to access encrypted content, which ensures that Microsoft and other third parties never have access to encrypted data on their own.

DKE can be deployed either in the cloud or on-premises, providing full flexibility for storage locations.

For more information, see Double key encryption.

Operations for your tenant key

Depending on your Azure Rights Management service key type, you have different levels of control and responsibility for your Azure Rights Management tenant key.

Terminology for key management operations:

- When you use a tenant key that's generated by Microsoft, the key is **Microsoft-managed**.
- When you use a key in Azure Key Vault that you generated with BYOK, the key is **customer-managed**.

Use the following table to identify the management operations that you can do, depending on whether your Azure Rights Management tenant key is Microsoft-managed or customer-managed:

Expand table

Life cycle operation	Microsoft-managed (default)	Customer-managed (BYOK)
Revoke your tenant key	✕ (automatic)	✓
Rekey your tenant key	✓	✓
Backup and recover your tenant key	✕	✓
Export your tenant key	✓	✕
Respond to a breach	✓	✓

For more information about revoke, rekey, backup, export, and breach response operations, see the **Microsoft-managed** or **Customer-managed (BYOK)** section for your tenant key type.

However, if you want to create an Azure Rights Management tenant key by importing a trusted publishing domain (TPD) from Active Directory Rights Management Services, this import operation is part of the migration from AD RMS to Azure Information Protection. As part of the design, an AD RMS TPD can only be imported to one tenant.

Microsoft-managed

The following sections describe key management operations when Microsoft manages your Azure Rights Management tenant key.

Revoke your tenant key

When you cancel the only or last subscription that includes the Azure Rights Management service, the service stops using your Azure Rights Management tenant key and no action is needed from you.

Rekey your tenant key

Rekeying is also known as rolling your key. When you rekey, the Azure Rights Management service stops using the existing tenant key to encrypt items, and starts to use a different key. Policies and rights management templates are immediately resigned but this changeover is gradual for existing clients and services using the Azure Rights Management service. So for some time, some new content continues to be encrypted with the old Azure Rights Management tenant key.

To rekey, you must configure the Azure Rights Management tenant key object and specify the alternative key to use. Then, the previously used key is automatically marked as archived for the Azure Rights Management service. Archiving the previously used key ensures that content encrypted with that key remains accessible.

Examples of when you might need to rekey for the Azure Rights Management service:

- You have migrated from Active Directory Rights Management Services (AD RMS) with a cryptographic mode 1 key. When the migration is complete, you want to change to using a key that uses cryptographic mode 2.
- Your company has split into two or more companies. When you rekey your Azure Rights Management tenant key, the new company won't have access to new content that your employees encrypt. They can access the old content if they have a copy of the old Azure Rights Management tenant key.
- You want to move from one key management topology to another.
- You believe the master copy of your Azure Rights Management tenant key is compromised.

To rekey, you can select a different Microsoft-managed key to become your Azure Rights Management tenant key, but you can't create a new Microsoft-managed key. To create a new key, you must change your key topology to be customer-managed (BYOK).

You have more than one Microsoft-managed key if you migrated from Active Directory Rights Management Services (AD RMS) and chose the Microsoft-managed key topology for the Azure Rights Management service. In this scenario, you have at least two Microsoft-

managed keys for your tenant. One key, or more, is the key or keys that you imported from AD RMS. You'll also have the default key that was automatically created for your Azure Rights Management tenant.

To select a different key to be your active tenant key for the Azure Rights Management service, use the Set-AipServiceKeyProperties cmdlet from the AIPService module. To help you identify which key to use, use the Get-AipServiceKeys cmdlet. You can identify the default key that was automatically created for your Azure Rights Management tenant by running the following command:

```PowerShell
(Get-AipServiceKeys) | Sort-Object CreationTime | Select-Object -First 1
```

To change your key topology to be customer-managed (BYOK), see Bring your own key for the Azure Rights Management service root key.

Backup and recover your tenant key

Microsoft is responsible for backing up your Azure Rights Management tenant key and no action is required from you.

Export your tenant key

You can export your Azure Rights Management service configuration and corresponding tenant key by following the instructions in the following three steps:

Step 1: Initiate export

Start the export process by opening a support case with Microsoft.

- Contact Microsoft Support to open an **Microsoft Purview Information Protection support case with a request for an Azure Rights Management key export**. You must prove you're a global administrator for your tenant, and understand that this process takes several days to confirm. Standard support charges apply; exporting your tenant key is not a free-of-charge support service.

Step 2: Wait for verification

After you submit the request, Microsoft reviews and verifies it.

- Microsoft verifies that your request to release your Azure Rights Management tenant key is legitimate. This process can take up to three weeks.

Step 3: Receive key instructions from CSS

After verification, Microsoft sends you the files and instructions needed to complete the export.

- Microsoft Customer Support Services sends you your Azure Rights Management service configuration and tenant key encrypted in a password-protected file. This file has a **.tpd** file name extension. To prepare the encrypted file, a Microsoft support engineer first sends you (as the person who initiated the export) a tool by email. You must run the tool from a command prompt as follows:

  PowerShell

  ```
  AadrmTpd.exe -createkey
  ```

  This generates an RSA key pair and saves the public and private halves as files in the current folder. For example: **PublicKey-FA29D0FE-5049-4C8E-931B-96C6152B0441.txt** and **PrivateKey-FA29D0FE-5049-4C8E-931B-96C6152B0441.txt**.

  Respond to the email from the support engineer, attaching the file that has a name that starts with **PublicKey**. Microsoft Support next sends you a TPD file as an .xml file that is encrypted with your RSA key. Copy this file to the same folder as you ran the AadrmTpd tool originally, and run the tool again, using your file that starts with **PrivateKey** and the file from Microsoft Support. For example:

  PowerShell

  ```
  AadrmTpd.exe -key PrivateKey-FA29D0FE-5049-4C8E-931B-96C6152B0441.txt -
  target TPD-77172C7B-8E21-48B7-9854-7A4CEAC474D0.xml
  ```

  The output of this command should be two files: a password text file containing the plain text password for the password-protected TPD, and an exported TPD XML file containing the password-protected TPD itself. The files have a new GUID, for example:

  - Password-5E4C2018-8C8C-4548-8705-E3218AA1544E.txt
  - ExportedTPD-5E4C2018-8C8C-4548-8705-E3218AA1544E.xml

Back up these files and store them safely to ensure that you can continue to decrypt content that is encrypted with this tenant key. In addition, if you're migrating to AD RMS, you can import this TPD file (the file that starts with **ExportedTDP**) to your AD RMS server.

Step 4: Ongoing: Protect your tenant key

After you receive your tenant key, keep it well-guarded, because if somebody gets access to it, they can decrypt all items that are encrypted by using that key.

If the reason to export your tenant key is because you no longer want to use the Azure Rights Management service, as a best practice, now deactivate the Azure Rights Management service in your tenant. Don't delay deactivating the Azure Rights Management service after you receive your tenant key, because deactivation helps to minimize the consequences if your tenant key is accessed by somebody who shouldn't have it. For instructions, see Decommission and deactivate the Azure Rights Management service.

Respond to a breach

No security system, no matter how strong, is complete without a breach response process. Your Azure Rights Management tenant key might be compromised or stolen. Even when it’s protected well, vulnerabilities might be found in current generation key technology or in current key lengths and algorithms.

Microsoft has a dedicated team to respond to security incidents in its products and services. As soon as there's a credible report of an incident, Microsoft's security incident response team engages to investigate the scope, root cause, and mitigations. If the investigated security incident affects your assets, Microsoft will notify the global administrators for your tenant by email.

If you have a breach, the best action that you or Microsoft can take depends on the scope of the breach; Microsoft works with you through this process. The following table shows some typical situations and the likely response, although the exact response depends on all the information that is revealed during the investigation.

Expand table

Incident description	Likely response
Your Azure Rights Management tenant key is leaked.	Rekey your tenant key. See the Rekey your tenant key section in this article.

Incident description	Likely response
An unauthorized individual or malware got rights to use your Azure Rights Management tenant key, but the key itself didn't leak.	Rekeying your tenant key doesn't help here and requires root-cause analysis. If a process or software bug was responsible for the unauthorized individual to get access, that situation must be resolved.
Vulnerability discovered in the RSA algorithm, or key length, or brute-force attacks become computationally feasible.	Microsoft must update the Azure Rights Management service to support new algorithms and longer key lengths that are resilient and instruct all customers to rekey their Azure Rights Management tenant key.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

Last updated on 06/26/2026

Bring your own key for the Azure Rights Management service root key

> *Microsoft Purview service description*

Use the information in this article if you've decided to configure your tenant to use your own root key for the Azure Rights Management service, instead of a default key generated by Microsoft. This configuration is often referred to as Bring Your Own Key (BYOK).

For more information about the root key options, see Key types for the service.

BYOK and usage logging work seamlessly with applications that integrate with the Azure Rights Management service used by Microsoft Purview Information Protection.

Supported applications include:

- **Cloud services**, such as Microsoft SharePoint or Microsoft 365
- **On-premises services** running Exchange and SharePoint applications that use the Azure Rights Management service by using the Rights Management connector
- **Client applications**, such as Office 2024 and Office 2021.

> **Tip**
>
> If needed, apply additional security to specific documents using an additional on-premises key. For more information, see **Double Key Encryption (DKE)**.

Azure Key Vault key storage

Customer-generated keys must be stored in the Azure Key Vault for BYOK.

> **Note**
>
> Using HSM-protected keys in the Azure Key Vault requires an Azure Key Vault Premium service tier, which incurs an additional monthly subscription fee.

Sharing key vaults and subscriptions

We recommend using a **dedicated key vault** for your Azure Rights Management tenant key. Dedicated key vaults help to ensure that calls by other services don't cause service limits to be exceeded. Exceeding service limits on the key vault where your tenant key is stored may cause response time throttling for the Azure Rights Management service.

Because different services have varying key management requirements, Microsoft also recommends using **a dedicated Azure subscription** for your key vault for the following reasons:

- Help safeguard against misconfigurations
- More secure when different services have different administrators

To share an Azure subscription with other services that use Azure Key Vault, make sure that the subscription shares a common set of administrators. Confirm that all administrators who use the subscription have a thorough understanding of every key they can manage. Administrators with this level of understanding are less likely to misconfigure your keys.

Example: Use a shared Azure subscription when the administrators for your Azure Rights Management tenant key are the same people that administer your keys for Microsoft Purview Customer Key and Dynamics 365 Online. If the key administrators for these services are different, we recommend using dedicated subscriptions.

Benefits of using Azure Key Vault

Azure Key Vault provides a centralized and consistent key management solution for many cloud-based and on-premises services that use encryption.

In addition to managing keys, Azure Key Vault offers your security administrators the same management experience to store, access, and manage certificates and secrets (such as passwords) for other services and applications that use encryption.

Storing your tenant key in the Azure Key Vault provides the following advantages:

Expand table

Advantage	Description
Built-in interfaces	Azure Key Vault supports a number of built-in interfaces for key management, including PowerShell, CLI, REST APIs, and the Azure portal. Other services and tools have integrated with Azure Key Vault for optimized capabilities

Advantage	Description
	for specific tasks, such as monitoring. For example, analyze your key usage logs with Azure Monitor Agent, set alerts when specified criteria are met, and so on.
Role separation	Azure Key Vault provides role separation as a recognized security best practice. Role separation ensures that administrators for the Azure Rights Management service can focus on their highest priorities, including managing data classification and protection, as well as encryption keys and policies for specific security or compliance requirements.
Master key location	Azure Key Vault is available in a variety of locations, and supports organizations with restrictions where master keys can live. For more information, see the Products available by region page on the Azure site.
Separated security domains	Azure Key Vault uses separate security domains for its data centers in regions such as North America, EMEA (Europe, Middle East and Africa), and Asia. Azure Key Vault also uses different instances of Azure, such as Microsoft Azure Germany, and Azure Government.
Unified experience	Azure Key Vault also enables security administrators to store, access, and manage certificates and secrets, such as passwords, for other services that use encryption. Using Azure Key Vault for your tenant keys provides a seamless user experience for administrators who manage all of these elements.

For the latest updates and to learn how other services use Azure Key Vault, visit the Azure Key Vault team blog.

Usage logging for BYOK

Usage logs are generated by every application that makes requests to the Azure Rights Management service.

Although usage logging is optional, we recommend using the near real-time usage logs for the Azure Rights Management service to see exactly how and when your Azure Rights Management tenant key is being used.

For more information about key usage logging for BYOK, see Usage logging for the Azure Rights Management service.

> **Tip**
>
> For additional assurance, this usage logging can be cross referenced with **Azure Key Vault logging**. Azure Key Vault logs provide a reliable method to independently monitor that your key is only used by the Azure Rights Management service.
>
> If necessary, immediately revoke access to your key by removing permissions on the key vault.

Options for creating and storing your key

> **Note**
>
> For general information about the Managed HSM offering, and how to set up a vault and a key, see the Azure Key Vault documentation.
>
> This section contains additional instructions about granting key authorization for the Azure Rights Management service.

BYOK supports keys that are created either in Azure Key Vault or on-premises.

If you create your key on-premises, you must then transfer or import it into your key vault and configure the Azure Rights Management service to use the key. Perform any additional key management from within Azure Key Vault.

Options to create and store your own key:

- **Created on-premises**. Create your key on-premises and transfer it to Azure Key Vault using one of the following options:
  - **HSM-protected key, transferred as an HSM-protected key**. The most typical method chosen.

    While this method has the most administrative overhead, it may be required for your organization to follow specific regulations. The HSMs used by Azure Key Vault have FIPS 140 validation.
  - **Software-protected key that is converted and transferred to Azure Key Vault as an HSM-protected key**. This method is supported only when migrating from Active Directory Rights Management Services (AD RMS).

- **Created on-premises as a software-protected key and transferred to Azure Key Vault as a software-protected key**. This method requires a .PFX certificate file.

For example, do the following to use a key created on-premises:

1. Generate your tenant key on your premises, in line with your organization's IT and security policies. This key is the master copy. It remains on-premises, and you're responsible for its backup.

2. Create a copy of the master key, and securely transfer it from your HSM to Azure Key Vault. Throughout this process, the master copy of the key never leaves the hardware protection boundary.

Once transferred, the copy of the key is protected by Azure Key Vault.

- **Created in Azure Key Vault**. Create and store your key in Azure Key Vault as an HSM-protected key or a software-protected key.

  > ⚠ **Warning**
  >
  > Keys that are generated directly in Azure Key Vault can't be exported for use outside Azure Key Vault. Most organizations that use BYOK have compliance requirements that mandate that they maintain their own copy of the key outside of Azure Key Vault. **Do not create the key directly in Azure Key Vault**; instead, create the key on-premises and transfer it to Azure Key Vault.
  >
  > If your organization requires that keys are exportable, and in your possession, you must create the key on-premises and import to Azure Key Vault and maintain your own backups of the key on-premises. Disaster recovery planning and testing should include measures to regularly test recovery of these keys. Keys can be backed up from Azure Key Vault but can only be imported to the original subscription.

Exporting your trusted publishing domain

If you ever decide to stop using the Azure Rights Management service, you'll need a trusted publishing domain (TPD) to decrypt content that was encrypted by the Azure Rights Management service.

However, exporting your TPD isn't supported if you're using BYOK for your Azure Rights Management tenant key.

To prepare for this scenario, make sure to create a suitable TPD ahead of time. For more information, see How to prepare an Azure Information Protection "Cloud Exit" plan.

Implementing BYOK for your Azure Rights Management tenant key

Use the following steps to implement BYOK:

1. Review BYOK prerequisites
2. Choose a key vault location
3. Create and configure your key

Prerequisites for BYOK

BYOK prerequisites vary, depending on your system configuration. Verify that your system complies with the following prerequisites as needed:

Expand table

Requirement	Description
Azure subscription	Required for all configurations. For more information, see Verifying that you have a BYOK-compatible Azure subscription.
AIPService PowerShell module for the Azure Rights Management service	Required for all configurations. For more information, see Install the AIPService PowerShell module for the Azure Rights Management service. **Note:** The AIPService module runs on Windows PowerShell 5.1, and on PowerShell 7.4 and later with module version 3.1.0.0 or later.
Azure Key Vault prerequisites for BYOK	If you're using an HSM-protected key that was created on-premises, ensure that you also comply with the prerequisites for BYOK listed in the Azure Key Vault documentation.
Thales firmware version 11.62	You must have a Thales firmware version of 11.62 if you're migrating from AD RMS to the Azure Rights Management service by using software key to hardware key and are using Thales firmware for your HSM.
Firewall bypass for trusted Microsoft services	If the key vault that contains your tenant key uses Virtual Network Service Endpoints for Azure Key Vault, you must allow trusted Microsoft services to bypass this firewall. For more information, see Virtual Network Service Endpoints for Azure Key Vault.

Verifying that you have a BYOK-compatible Azure subscription

Your Azure Rights Management tenant must have an Azure subscription. If you don't have one yet, you can sign up for a free account. However, to use an HSM-protected key, you must have the Azure Key Vault Premium service tier.

The free Azure subscription that provides access to Microsoft Entra configuration isn't sufficient for using Azure Key Vault.

To confirm whether you have an Azure subscription that is compatible with BYOK, do the following to verify, using Azure PowerShell cmdlets:

1. Start an Azure PowerShell session as an administrator.
2. Run `Connect-AzAccount` and sign in with a role that can access resources in Azure Key Vault.
3. Run `Get-AzSubscription` to confirm that the following values are displayed:

   - Your subscription name and ID
   - Your tenant ID for the Azure Rights Management service
   - Confirmation that the state is enabled

   If no values are displayed and you're returned to the prompt, you don't have an Azure subscription that can be used for BYOK.

4. Verify that you can use the Azure Key Vault Premium tier, which is required for HSM-protected keys. If you have an existing key vault, run the following command to check its SKU:

   PowerShell

   ```powershell
   Get-AzKeyVault -VaultName "YourVaultName" | Select-Object VaultName, Sku
   ```

   The **Sku** value must be **Premium** for HSM-protected keys. If you're creating a new key vault, ensure you select the Premium tier during creation. For pricing information, see Azure Key Vault pricing.

Choosing your key vault location

When you create a key vault to contain the key to be used as your tenant key for the Azure Rights Management service, you must specify a location. This location is an Azure region, or Azure instance.

Make your choice first for compliance, and then to minimize network latency:

- If you have chosen the BYOK key method for compliance reasons, those compliance requirements might also mandate which Azure region or instance can be used to store your Azure Rights Management tenant key.
- All cryptographic calls for protection chain to your Azure Rights Management key. Therefore, you may want to minimize the network latency these calls require by creating your key vault in the same Azure region or instance as your Azure Rights Management tenant.

To identify the location of your tenant that uses the Azure Rights Management service, use the Get-AipServiceConfiguration PowerShell cmdlet and identify the region from the URLs. For example:

PowerShell

```
LicensingIntranetDistributionPointUrl : https://5c6bb73b-1038-4eec-863d-
49bded473437.rms.na.aadrm.com/_wmcs/licensing
```

The region is identifiable from **rms.na.aadrm.com**, and for this example, it is in North America.

The following table lists recommended Azure regions and instances for minimizing network latency:

Expand table

Azure region or instance	Recommended location for your key vault
rms.**na**.aadrm.com	**North Central US** or **East US**
rms.**eu**.aadrm.com	**North Europe** or **West Europe**
rms.**ap**.aadrm.com	**East Asia** or **Southeast Asia**
rms.**sa**.aadrm.com	**West US** or **East US**
rms.**govus**.aadrm.com	**Central US** or **East US 2**
rms.**aadrm.us**	**US Gov Virginia** or **US Gov Arizona**
rms.**aadrm.cn**	**China East 2** or **China North 2**

Create and configure your key

> **Important**
>
> For information specific for Managed HSMs, see **Enabling key authorization for Managed HSM keys via Azure CLI**.

Create an Azure Key Vault and the key you want to use for the Azure Rights Management service. For more information, see the Azure Key Vault documentation.

> **Important**
>
> After creating the Azure Key Vault, immediately enable both soft delete and purge protection. This prevents accidental deletion of the vault and keys. Loss of the keys without sufficient backups results in complete data loss of encrypted files and emails. For details, see **Azure Key Vault: soft-delete overview**.

Note the following for configuring your Azure Key Vault and key for BYOK:

- Key length requirements
- Creating an HSM-protected key on-premises and transferring it to your key vault
- Configuring the Azure Rights Management service for your key ID
- Authorizing the Azure Rights Management service to use your key

Key length requirements

When you create your key, make sure that the key length is 2048 bits. Other key lengths aren't supported by the Azure Rights Management service.

Creating an HSM-protected key on-premises and transferring it to your key vault

To create an HSM-protected key on-premises and transfer it to your key vault as an HSM-protected key, follow the procedures in the Azure Key Vault documentation: How to generate and transfer HSM-protected keys for Azure Key Vault.

For the Azure Rights Management service to use the transferred key, the following Key Vault operations must be permitted for the key:

- get
- decrypt

- sign

To check the permitted operations for a specific key, run the following PowerShell command:

PowerShell

```
(Get-AzKeyVaultKey -VaultName <key vault name> -Name <key name>).Attributes.KeyOps
```

Configuring the Azure Rights Management service for your key ID

Keys stored in the Azure Key Vault each have a key ID.

The key ID is a URL that contains the name of the key vault, the keys container, the name of the key, and the key version. For example: `https://contosorms-kv.vault.azure.net/keys/contosorms-byok/aaaabbbbcccc111122223333`

Configure the Azure Rights Management service to use your key by specifying its key vault URL.

Authorizing the Azure Rights Management service to use your key

The Azure Rights Management service must be authorized to use your key through Azure RBAC. This requires creating a custom role with the necessary permissions and assigning it to the Rights Management service principal.

> **Note**
>
> This role, when assigned, allows the Azure Rights Management service to perform cryptographic operations with your key. Members assigned to this role can use all keys in the vault for these operations. Assign only to a vault dedicated to the Azure Rights Management service. Do not assign this role to other vaults or use the vault for other services.

Enabling key authorization using Azure CLI

1. Sign in to Azure CLI and get your key vault's subscription ID:

   Azure CLI

   ```
   az login
   az keyvault show --name <vault-name> --resource-group <resource-group-name> --
   query "id" -o tsv
   ```

The output contains the subscription ID in the format: `/subscriptions/<subscription-id>/resourceGroups/...`

2. Create a JSON file named `PurviewRMSCustomAKVRole.json` with the following content, replacing `<subscription-id>` with the subscription ID from the previous step:

JSON

```
{
  "Name": "Purview Rights Management key vault Encryption User",
  "IsCustom": true,
  "Description": "Grants access to use key vault keys for the Purview Rights
Management service.",
  "Actions": [],
  "NotActions": [],
  "DataActions": [
    "Microsoft.KeyVault/vaults/keys/read",
    "Microsoft.KeyVault/vaults/keys/sign/action",
    "Microsoft.KeyVault/vaults/keys/decrypt/action"
  ],
  "NotDataActions": [],
  "AssignableScopes": [
    "/subscriptions/<subscription-id>"
  ]
}
```

3. Create the custom role definition, changing the role definition file path if necessary:

Azure CLI

```
az role definition create --role-definition "./PurviewRMSCustomAKVRole.json"
```

> **Note**
>
> Run this command from the directory where you saved `PurviewRMSCustomAKVRole.json`, or replace `./PurviewRMSCustomAKVRole.json` with the full path to the file.

4. Get the key vault resource ID and assign the role to the Rights Management service principal in a single command:

Azure CLI

```
az role assignment create --role "Purview Rights Management key vault Encryption
User" --assignee "00000012-0000-0000-c000-000000000000" --scope $(az keyvault
```

```
show --name <vault-name> --resource-group <resource-group-name> --query "id" -o
tsv)
```

Enabling key authorization using Azure PowerShell

1. Sign in to Azure PowerShell and get your key vault's subscription ID:

PowerShell

```
Connect-AzAccount

# Define your key vault and resource group names
$vaultName = "<vault-name>"
$resourceGroupName = "<resource-group-name>"

# Get the key vault resource ID and subscription ID
$keyVault = Get-AzKeyVault -VaultName $vaultName -ResourceGroupName
$resourceGroupName
$resourceId = $keyVault.ResourceId
$subscriptionId = ($resourceId -split "/")[2]

# Create the custom role definition JSON file
$jsonRoleDefinition = @"
{
  "Name": "Purview Rights Management key vault Encryption User",
  "IsCustom": true,
  "Description": "Grants access to use key vault keys for the Purview Rights
Management service.",
  "Actions": [],
  "NotActions": [],
  "DataActions": [
    "Microsoft.KeyVault/vaults/keys/read",
    "Microsoft.KeyVault/vaults/keys/sign/action",
    "Microsoft.KeyVault/vaults/keys/decrypt/action"
  ],
  "NotDataActions": [],
  "AssignableScopes": [
    "/subscriptions/$($subscriptionId)"
  ]
}
"@

$jsonRoleDefinition | Out-File -FilePath "./PurviewRMSCustomAKVRole.json"

# Create the custom role definition
New-AzRoleDefinition -InputFile "./PurviewRMSCustomAKVRole.json"

# Assign the role to the Rights Management service principal
New-AzRoleAssignment -RoleDefinitionName "Purview Rights Management key vault
Encryption User" -ServicePrincipalName "00000012-0000-0000-c000-000000000000" -
Scope $resourceId
```

Enabling key authorization for Managed HSM keys via Azure CLI

To grant the Azure Rights Management service principal user permissions as a **Managed HSM Crypto** user, run the following command:

Azure CLI

```
az keyvault role assignment create --hsm-name "ContosoMHSM" --role "Managed HSM Crypto User" --assignee-principal-type ServicePrincipal --assignee https://aadrm.com/ --scope /keys/contosomhskey
```

Where:

- **ContosoMHSM** is a sample HSM name. When running this command, replace this value with your own HSM name.

The **Managed HSM Crypto User** user role allows the user to decrypt, sign, and get permissions to the key, which are all required for the Managed HSM functionality.

Configure the Azure Rights Management service to use your key

After you've completed the previous steps, you're ready to configure the Azure Rights Management service to use this key as your organization's tenant key.

First, copy the URL for your key from Azure Key Vault, including the key version. You can retrieve this using Azure CLI or PowerShell:

Azure CLI:

Azure CLI

```
az keyvault key show --vault-name <vault-name> --name <key-name> --query "key.kid" -o tsv
```

Azure PowerShell:

PowerShell

```
(Get-AzKeyVaultKey -VaultName "<vault-name>" -Name "<key-name>").Id
```

Configuring a new BYOK key

If this is a new key, use the following steps to configure the Azure Rights Management service:

1. Install and import the AIPService PowerShell module if you haven't already:

   > **Note**
   >
   > The AIPService module runs on Windows PowerShell 5.1, and on PowerShell 7.4 and later with module version 3.1.0.0 or later.

   PowerShell

   ```
   Install-Module -Name AIPService
   Import-Module -Name AIPService
   ```

2. Connect to the Azure Rights Management service:

   PowerShell

   ```
   Connect-AipService
   ```

3. Run the Use-AipServiceKeyVaultKey cmdlet, specifying the key URL:

   PowerShell

   ```
   Use-AipServiceKeyVaultKey -KeyVaultKeyUrl "https://contosorms-
   kv.vault.azure.net/keys/contosorms-byok/<key-version>"
   ```

   > **Important**
   >
   > In this example, `<key-version>` is the version of the key you want to use. You must specify the key version. If your key is later updated or renewed, the Azure Rights Management service will stop working for your tenant unless you update the configuration with the new key version.

4. List the keys configured for your tenant to get the key ID of the new key using `Get-AipServiceKeys`. Find the key with the matching KeyVaultKeyUrl. Copy the KeyIdentifier value:

   PowerShell

```
Get-AipServiceKeys | Select-Object KeyIdentifier, KeyVaultKeyUrl
```

5. Set the new key as the active tenant key using new key ID from the previous command:

PowerShell

```
Set-AipServiceKeyProperties -KeyIdentifier <key-id-from-previous-step> -Active $true
```

The Azure Rights Management service is now configured to use your key instead of the default Microsoft-created key that was automatically created for your tenant.

Migrating from Azure Key Vault hsmPlatform 1 to hsmPlatform 2

Azure Key Vault hsmPlatform 1 is being retired. For more information about this retirement, see Azure Key Vault hsmPlatform 1 retirement announcement.

If your Azure Rights Management tenant key is stored on a hsmPlatform 1 key vault, you must migrate to hsmPlatform 2 to continue using BYOK.

Determine if your key is on hsmPlatform 1

First, identify whether your existing Rights Management key is on hsmPlatform 1:

Using Azure CLI:

Azure CLI

```
az keyvault key show --vault-name <vault-name> --name <key-name> --query "attributes.hsmPlatform" -o tsv
```

Using Azure PowerShell:

PowerShell

```
(Get-AzKeyVaultKey -VaultName "<vault-name>" -Name "<key-name>").Attributes.HsmPlatform
```

- If the `hsmPlatform` value is `1`, your key is on hsmPlatform 1 and requires migration.
- If the `hsmPlatform` value is `2` or higher, your key is already on hsmPlatform 2 and no action is required.

Migration steps for keys on hsmPlatform 1

If your key is on hsmPlatform 1, follow these steps to migrate:

Step 1: Locate your on-premises HSM

You must have access to the on-premises HSM that was originally used to generate the key material, or a backup of that key that can be imported into a compatible HSM within your administrative boundary.

> **ⓘ Important**
>
> If you no longer have access to the original HSM or the key material, see **My organization has lost access to the original key material**.

Step 2: Re-import the key as a new version in the existing key vault

You don't need to create a new key vault for this migration. Package and re-import the same key material from your on-premises HSM into the existing Azure Key Vault as a **new key version**. For detailed instructions, see How to generate and transfer HSM-protected keys for Azure Key Vault.

Step 3: Validate that the new key version is on hsmPlatform 2

After the import completes, confirm that the new key version reports `hsmPlatform` `2` before you update the Rights Management service:

Using Azure CLI:

Azure CLI

```
az keyvault key show --vault-name <vault-name> --name <key-name> --version <new-key-version> --query "attributes.hsmPlatform" -o tsv
```

Using Azure PowerShell:

PowerShell

```
(Get-AzKeyVaultKey -VaultName "<vault-name>" -Name "<key-name>" -Version "<new-key-version>").Attributes.HsmPlatform
```

The returned value must be `2`. If it isn't, stop and verify that you imported the correct key material and that the import completed successfully.

Step 4: Verify that the public keys match

Before updating the Rights Management service configuration, verify that the public keys match across all locations:

1. **Get the public key from the new key version in Azure Key Vault:**

   Using Azure CLI:

   ```
   Azure CLI
   az keyvault key show --vault-name <vault-name> --name <key-name> --version <new-key-version> --query "key.n" -o tsv
   ```

   Using Azure PowerShell:

   ```
   PowerShell
   Get-AzKeyVaultKey -VaultName "<vault-name>" -Name "<key-name>" -Version "<new-key-version>" | Format-List Key
   ```

Review the **n** value from the output, which represents the public key modulus.

1. **Get the public key from the Rights Management service:**

   ```
   PowerShell
   Connect-AipService
   Get-AipServiceKeys | Select-Object KeyIdentifier, PublicKey
   ```

2. **Compare with your on-premises HSM copy** using your HSM vendor's tools to export or display the public key. Contact your HSM vendor for assistance if you're unable to locate the public key details. Microsoft Support can't assist with HSM-specific operations.

All three public keys must match exactly. If they don't match, verify you're importing the correct key from your HSM.

Step 5: Update the Azure Rights Management service to use the new key version

Run Convert-AipServiceKeyToKeyVault to update the key URL to the new key version:

PowerShell

```
Connect-AipService
Convert-AipServiceKeyToKeyVault -KeyIdentifier <existing-key-id> -KeyVaultKeyUrl
"https://<vault-name>.vault.azure.net/keys/<key-name>/<new-key-version>"
```

> ⓘ Note
>
> This command fails if the public keys don't match exactly between the existing BYOK key and the key in the new Azure Key Vault. Ensure you've completed the verification in Step 4 before running this command.

My organization has lost access to the original key material

If you no longer have access to the original on-premises HSM or the key material used to create your BYOK key, you must create a new key:

1. **Create a new key** following the steps in Create and configure your key.
2. **Configure access control** as described in Authorizing the Azure Rights Management service to use your key.
3. **Set the new key as active:**

   PowerShell

   ```
   Connect-AipService
   Use-AipServiceKeyVaultKey -KeyVaultKeyUrl "https://<vault-
   name>.vault.azure.net/keys/<key-name>/<key-version>"
   Set-AipServiceKeyProperties -KeyIdentifier <new-key-id> -Active $true
   ```

4. **Open a support case** with Microsoft Purview support to discuss next steps for content that was protected with the previous key. Go to Microsoft 365 admin center and open a service request, or contact your Microsoft support representative. For more information, see Get support for Microsoft 365 for business.

> ⚠ Warning
>
> If you aren't able to migrate the original key, **do not remove** the existing hsmPlatform 1 key after creating the new key. Content protected with the previous key will become inaccessible if the key is removed.

Next steps

If you haven't yet activated the Rights Management service, do this step now.

If the service was activated before you configured BYOK, users are gradually transitioned from the old key to the new key over the course of a few weeks. During this transition, documents and files that were encrypted with the old tenant key remain accessible to authorized users.

Use the Azure Rights Management usage logging to see every transaction that the Azure Rights Management service and your key performs. For example, from a log file displayed in Excel, the **KeyVaultDecryptRequest** and **KeyVaultSignRequest** request types show that the tenant key is being used.

date	time	row-id	request-type	user-id	result	correlation-id
7/18/2016	17:41:39	-	KeyVaultDecryptRequest	''	'Success'	b9e063ac-6208-4d2
7/17/2016	02:22:51	76445a86-3d	AcquireTemplates	-	'Success'	b01bc442-cde2-41e
7/18/2016	17:41:40	-	KeyVaultSignRequest	''	'Success'	b9e063ac-6208-4d2
7/18/2016	17:41:40	-	KeyVaultDecryptRequest	''	'Success'	b9e063ac-6208-4d2
7/17/2016	01:40:33	a5b1172b-c6	GetClientLicensorCert	'RmsTenantUser@db8048	'Success'	d6670bb5-4da7-4eb
7/17/2016	01:40:59	9d93a998-8c	Certify	'RmsTenantUser@db8048	'Success'	928a5df6-ccab-4143
7/17/2016	01:35:50	d5da5952-0b	AcquireLicense	'RmsTenantConsumer@d	'Success'	ae695fd1-a4e1-44e

Last updated on 08/26/2026

Activate the Azure Rights Management service

> *Microsoft Purview service description*

This article describes how administrators can activate the Azure Rights Management encryption service for Microsoft Purview Information Protection. When this encryption service is activated for your organization, administrators and users can start to protect important data by using apps and services that support this encryption solution. Administrators can also manage and monitor encrypted items that your organization owns.

The configuration information in this article is for administrators who are responsible for a service that applies to all users in an organization. If you are looking for user help and information to use the Rights Management functionality for a specific application or how to open a file or email that is rights-protected, use the help and guidance that accompanies your application.

Automatic activation for the Azure Rights Management service

When you have a service plan that includes the Azure Rights Management service, you may not have to activate the service:

- **If your subscription that includes the Azure Rights Management service, Azure Information Protection (former name), or Microsoft Purview Information Protection was obtained towards the end of February 2018 or later**: The service is automatically activated for you. You do not have to activate the service unless you or another administrator for your organization deactivated the Azure Rights Management service.

- **If your subscription that includes the Azure Rights Management service, Azure Information Protection (former name), or Microsoft Purview Information Protection was obtained before or during February 2018**: Microsoft activates the Azure Rights Management service for these subscriptions if your tenant is using Exchange Online. For these subscriptions, the service will be activated for you unless you see that **AutomaticServiceUpdateEnabled** is set to **false** when you run Get-IRMConfiguration.

If neither of the listed scenarios apply to you, you must manually activate the Azure Rights Management service.

How to activate or confirm the status of the encryption service

> **Important**
>
> Do not activate the Azure Rights Management service if you have Active Directory Rights Management Services (AD RMS) deployed for your organization. **Prepare your environment for AD RMS**.

To activate the Azure Rights Management service, your organization must have a service plan that includes the Azure Rights Management service from Microsoft Purview Information Protection. For details about required service plans, see Microsoft 365 licensing guidance for security & compliance.

When the Azure Rights Management service is activated, all users in your organization can apply encryption to items such as documents and emails, and all users can open (consume) items that have been encrypted by this service. However, if you prefer, you can restrict who can apply this encryption, by using onboarding controls for a phased deployment. For more information, see the Configuring onboarding controls for a phased deployment section in this article.

Activate the Azure Rights Management service via PowerShell

You must use PowerShell to activate the Azure Rights Management service. You can no longer activate or deactivate this service from admin portals.

1. Install the AIPService module to configure and manage the Azure Rights Management service. For instructions, see Install the AIPService PowerShell module for the Azure Right Management service.

2. Use the Connect-AipService cmdlet to connect your PowerShell session to the Azure Rights Management service. Run `Connect-AipService` and sign in with a role that has permissions to activate the Azure Rights Management service for your tenant. For example, the compliance administrator role, or compliance data administrator role.

3. Run Get-AipService to confirm whether the Azure Rights Management service is activated. A status of **Enabled** confirms activation; **Disabled** indicates that the service is deactivated.

4. To activate the service, run Enable-AipService.

Configuring onboarding controls for a phased deployment

If you don't want all users to be able to encrypt documents and emails immediately by using the Azure Rights Management service, you can configure user onboarding controls by using the Set-AipServiceOnboardingControlPolicy PowerShell command. You can run this command before or after you activate the Azure Rights Management service.

For example, if you initially want only administrators in the "IT department" group (that has an object ID of aaaaaaaa-0000-1111-2222-bbbbbbbbbbbb) to be able to protect content for testing purposes, use the following command:

PowerShell

```
Set-AipServiceOnboardingControlPolicy -UseRmsUserLicense $False -SecurityGroupObjectId
"aaaaaaaa-0000-1111-2222-bbbbbbbbbbbb"
```

For this configuration option, you must specify a group; you can't specify individual users. To get the object ID for the group, you can use the Microsoft Graph PowerShell—for example, for version 1.0 of the module, use the Get-MgGroup command. Or, you can copy the **Object ID** value of the group from the Azure portal.

Alternatively, if you want to ensure that only users who are correctly licensed to use the Azure Rights Management service can protect content:

PowerShell

```
Set-AipServiceOnboardingControlPolicy -UseRmsUserLicense $True
```

When you no longer need to use onboarding controls, whether you used the group or licensing option, run:

PowerShell

```
Set-AipServiceOnboardingControlPolicy -UseRmsUserLicense $False
```

For more information about this cmdlet and additional examples, see the Set-AipServiceOnboardingControlPolicy help.

When you use these onboarding controls, all users in the organization can always consume encrypted content that has been protected by your subset of users, but they won't be able to apply encryption themselves from client applications. Server-side applications, such as Exchange, can implement their own per-user controls to achieve the same result. For example, to prevent users from protecting emails in Outlook on the web, use Set-OwaMailboxPolicy to set the *IRMEnabled* parameter to *$false*.

Next steps

Now that the Azure Rights Management service is activated for your organization, apps and services can apply encryption to help protect your data. The easiest, and recommended way to apply encryption, is by using sensitivity labels from Microsoft Purview Information Protection. If you're ready to do that, see Get started with sensitivity labels.

To verify the service is working, try a simple test. Use one account to apply a sensitivity label that encrypts a document or email. Then use a different account on another computer to open that item.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/22/2026

Configure usage rights for the Azure Rights Management service

08/28/2025

This article describes usage rights you can configure to be automatically applied when a sensitivity label from Microsoft Purview Information Protection is selected by users, administrators, or configured services.

Usage rights are selected when you configure sensitivity labels or rights management templates for encryption. For example, you can select roles that configure a logical grouping of usage rights, or configure the individual rights separately. Alternatively users might select and apply the usage rights themselves.

> **Important**
>
> Use this article to understand how usage rights are designed to be interpreted by applications.
>
> Applications may vary in how they implement usage rights, and we recommend you consult your application's documentation and do your own testing to check application behavior before you deploy in production.

Usage rights and descriptions

The following table lists and describes the usage rights that Rights Management supports, and how they're used and interpreted. They're listed by their **common name**, which is typically how you might see the usage right displayed or referenced, as a more friendly version of the single-word value that is used in the code (the **Encoding in policy** value).

In this table:

- The **API Constant or Value** is the SDK name for an MSIPC or Microsoft Information Protection SDK API call, used when you write an application that checks for a usage right, or adds a usage right to a policy.
- The names from **AD RMS templates** are included for customers on or migrating from the on-premises rights management solution, Active Rights Rights Management Services.

Expand table

Usage right	Description	Implementation
Common name: **Edit Content, Edit** Encoding in policy: **DOCEDIT**	Allows the user to modify, rearrange, format, or sort the content inside the application, which includes Office on the web. It doesn't grant the right to save the edited copy.	Office custom rights: As part of the **Change** and **Full Control** options. Name in the Microsoft Purview portal: **Edit Content, Edit (DOCEDIT)** Name in AD RMS templates: **Edit** API constant or value: **MSIPC:** Not applicable. **MIP SDK:** `DOCEDIT`
Common name: **Save** Encoding in policy: **EDIT**	Allows the user to save the item to the current location. In Office on the web, it also allows the user to edit the content. In Office applications, this right allows the user to save the contents of the file to a new location and with a new name if the selected file format has built-in support for Rights Management. The list of file formats available is restricted to those that support Rights Management. This restriction ensures that the original encryption can't be removed from the file.	Office custom rights: As part of the **Change** and **Full Control** options. Name in the Microsoft Purview portal: **Save (EDIT)** Name in AD RMS templates: **Save** API constant or value: **MSIPC:** `IPC_GENERIC_WRITE L"EDIT"` **MIP SDK:** `EDIT`
Common name: **Comment** Encoding in policy: **COMMENT**	Enables the option to add annotations or comments to the content. This right is available in the SDK, is available as an ad-hoc policy in the PurviewInformationProtection PowerShell module, and has been implemented in some software vendor applications. However, it isn't widely used and isn't supported by Office applications.	Office custom rights: Not implemented. Name in the Microsoft Purview portal: Not implemented. Name in AD RMS templates: Not implemented. API constant or value: **MSIPC:** `IPC_GENERIC_COMMENT L"COMMENT` **MIP SDK:** `COMMENT`

Usage right	Description	Implementation
Common name: **Save As, Export** Encoding in policy: **EXPORT**	Enables the option to save the content to a different file name (Save As). This right also allows the user to perform other export options in applications, such as **Send to OneNote**.	Office custom rights: As part of the **Full Control** option. Name in the Microsoft Purview portal: **Save As, Export (EXPORT)** Name in AD RMS templates: **Export (Save As)** API constant or value: **MSIPC:** `IPC_GENERIC_EXPORT L"EXPORT"` **MIP SDK:** `EXPORT`
Common name: **Forward** Encoding in policy: **FORWARD**	Enables the option to forward an email message and to add recipients to the **To** and **Cc** lines. This right does not apply to documents; only email messages. Does not allow the forwarder to grant rights to other users as part of the forward action. When you grant this right, also grant the **Edit Content, Edit** right (common name), and additionally grant the **Save** right (common name) to ensure that the encrypted email message isn't delivered as an attachment. Also specify these rights when you send an email to another organization that uses the Outlook client or Outlook web app. Or, for users in your organization that are exempt from using Rights Management encryption because you have implemented onboarding controls.	Office custom rights: Denied when using the **Do Not Forward** standard policy. Name in the Microsoft Purview portal: **Forward (FORWARD)** Name in AD RMS templates: **Forward** API constant or value: **MSIPC:** `IPC_EMAIL_FORWARD L"FORWARD"` **MIP SDK:** `FORWARD`
Common name: **Full Control** Encoding in policy: **OWNER**	Grants all rights to the item and all available actions can be performed. Includes the ability to remove encryption and re-encrypt a document. Note that this usage right isn't the same as the Rights Management owner.	Office custom rights: As the **Full Control** custom option. Name in the Microsoft Purview portal: **Full Control (OWNER)** Name in AD RMS templates: **Full Control** API constant or value:

Usage right	Description	Implementation
		MSIPC: `IPC_GENERIC_ALL L"OWNER"` **MIP SDK:** `OWNER`
Common name: **Print** Encoding in policy: **PRINT**	Enables the options to print the content.	Office custom rights: As the **Print Content** option in custom permissions. Not a per-recipient setting. Name in the Microsoft Purview portal: **Print (PRINT)** Name in AD RMS templates: **Print** API constant or value: **MSIPC:** `IPC_GENERIC_PRINT L"PRINT"` **MIP SDK:** `PRINT`
Common name: **Reply** Encoding in policy: **REPLY**	Enables the **Reply** option in an email client, without allowing changes in the **To** or **Cc** lines. When you grant this right, also grant the **Edit Content, Edit** right (common name), and additionally grant the **Save** right (common name) to ensure that the encrypted email message isn't delivered as an attachment. Also specify these rights when you send an email to another organization that uses the Outlook client or Outlook web app. Or, for users in your organization that are exempt from using Rights Management protection because you have implemented onboarding controls.	Office custom rights: Not applicable. Name in AD RMS templates: **Reply** API constant or value: **MSIPC:** `IPC_EMAIL_REPLY` **MIP SDK:** `REPLY`
Common name: **Reply All** Encoding in policy: **REPLYALL**	Enables the **Reply All** option in an email client, but doesn't allow the user to add recipients to the **To** or **Cc** lines. When you grant this right, also grant the **Edit Content, Edit** right (common name), and additionally grant the **Save** right (common name) to ensure that the encrypted email message isn't delivered as an attachment. Also specify these rights when you send an email to another organization that uses the Outlook client or Outlook web app. Or, for users in your organization that are exempt from using the Azure Rights Management service because you have implemented onboarding controls.	Office custom rights: Not applicable. Name in the Microsoft Purview portal: **Reply All (REPLY ALL)** Name in AD RMS templates: **Reply All** API constant or value: **MSIPC:** `IPC_EMAIL_REPLYALL`

Usage right	Description	Implementation
		`L"REPLYALL"` **MIP SDK:** `REPLYALL`
Common name: **View, Open, Read** Encoding in policy: **VIEW**	Allows the user to open the document and see the content. Additionally: - In Microsoft 365 Copilot and other AI apps, allows the LLM to reference the item with a link but without summarizing the content, so the user can then open and view the content outside the AI app. In Excel, this right isn't sufficient to sort data, which requires the following right: **Edit Content, Edit**. To filter data in Excel, you need the following two rights: **Edit Content, Edit** and **Copy**.	Office custom rights: As the **Read** custom policy, **View** option. Name in the Microsoft Purview portal: **View, Open, Read (VIEW)** Name in AD RMS templates: **Read** API constant or value: **MSIPC:** `IPC_GENERIC_READ` `L"VIEW"` **MIP SDK:** `VIEW`

Common name: **Copy**
 	
Encoding in policy: **EXTRACT**
	Enables options to copy data (including screen captures) from the item into the same or another item. Additionally:
	- In Microsoft 365 Copilot and other AI apps, allows the LLM to access and summarize the content for the requesting user.
	- In some applications, this right also allows the entire document to be saved unencrypted.
	
	In Microsoft Teams and similar screen-sharing applications, the presenter must have this right to successfully present an encrypted document. If the presenter doesn't have this right, the attendees can't view the document and it displays as blacked out to them.
		Office custom rights: As the **Allow users with Read access to copy content** custom policy option.
		
		Name in the Microsoft Purview portal: **Copy (EXTRACT)**
		
		Name in AD RMS templates: **Extract**
		
		API constant or value:
		MSIPC:
		`IPC_GENERIC_EXTRACT`
		`L"EXTRACT"`
		MIP SDK: `EXTRACT`

Common name: **View Rights** Encoding in policy: **VIEWRIGHTSDATA**	Allows the user to see the policy that is applied to the document. Not supported by Office apps or by the Microsoft Purview Information Protection client.	Office custom rights: Not implemented. Name in the Microsoft Purview portal: **View Rights (VIEWRIGHTSDATA)**. Name in AD RMS templates: **View Rights**

Usage right	Description	Implementation
		API constant or value: **MSIPC:** `IPC_READ_RIGHTSL"VIEWRIGHTSDATA"` **MIP SDK:** `VIEWRIGHTSDATA`
Common name: **Change Rights** Encoding in policy: **EDITRIGHTSDATA**	Allows the user to change the policy that is applied to the document. Includes removing encryption. Not supported by Office apps or the Microsoft Purview Information Protection client.	Office custom rights: Not implemented. Name in the Microsoft Purview portal: **Edit Rights (EDITRIGHTSDATA)**. Name in AD RMS templates: **Edit Rights** API constant or value: **MSIPC:** `PC_WRITE_RIGHTSL"EDITRIGHTSDATA"` **MIP SDK:** `EDITRIGHTSDATA`
Common name: **Allow Macros** Encoding in policy: **OBJMODEL**	Enables the option to run macros or perform other programmatic or remote access to the content in a document.	Office custom rights: As the **Allow Programmatic Access** custom policy option. Not a per-recipient setting. Name in the Microsoft Purview portal: **Allow Macros (OBJMODEL)** Name in AD RMS templates: **Allow Macros** API constant or value: **MSIPC:** Not implemented. **MIP SDK:** `OBJMODEL`

Rights included in permissions levels

Some applications group usage rights together into permissions levels, to make it easier to select usage rights that are typically used together. These permissions levels help to abstract a level of complexity from users, so they can choose options that are role-based. For example, **Viewer** and **Restricted Editor**. Although these options often show users a summary of the rights, they might not include every right that is listed in the previous table.

Use the following table for a list of these permissions levels and a complete list of the usage rights that they contain. The usage rights are listed by their common name.

Expand table

Permissions level	Applications	Usage rights included
Viewer	Microsoft Purview portal Microsoft Purview Information Protection client	View, Open, Read; View Rights; Reply [1]; Reply All [1]; Allow Macros [2] Note: For emails, use Reviewer rather than this permission level to ensure that an email reply is received as an email message rather than an attachment. Reviewer is also required when you send an email to another organization that uses the Outlook client or Outlook web app. Or, for users in your organization that are exempt from using the Azure Rights Management service because you have implemented onboarding controls.
Restricted Editor (Previously **Reviewer** [3])	Microsoft Purview portal Microsoft Purview Information Protection client	View, Open, Read; Save; Edit Content, Edit; View Rights; Reply: Reply All; Forward; Allow Macros [2]
Editor (Previously **Co-Author** [3])	Microsoft Purview portal Microsoft Purview Information Protection client	View, Open, Read; Save; Edit Content, Edit; Copy; View Rights; Allow Macros [2]; Save As, Export [4]; Print; Reply; Reply All; Forward
Owner (Previously **Co-Owner** [3])	Microsoft Purview portal Microsoft Purview Information Protection client	View, Open, Read; Save; Edit Content, Edit; Copy; View Rights; Change Rights; Allow Macros; Save As, Export; Print; Reply; Reply All; Forward; Full Control

Footnote 1

Not included in the Microsoft Purview portal.

Footnote 2

Not included in the custom permissions dialog in Word, Excel, and PowerPoint for Windows (version 2408+).

Footnote 3

The Microsoft Purview portal and the custom permissions dialog in Word, Excel, and PowerPoint for Windows (version 2411+) have updated permissions level naming. **Reviewer** is now called **Restricted Editor**, **Co-Author** is now called **Editor**, and **Co-Owner** is now called **Owner**. Other applications continue to use the original permissions level naming.

Footnote 4

Not included in the Microsoft Purview portal.

Do Not Forward option for emails

Exchange clients and services (for example, the Outlook client, Outlook on the web, Exchange mail flow rules, and DLP actions for Exchange) have an additional information rights protection option for emails: **Do Not Forward**.

Although this option appears to users (and Exchange administrators) as if it's a default Rights Management template that they can select, **Do Not Forward** isn't a template. Instead, the **Do Not Forward** option is a set of usage rights that is dynamically applied by users to their email recipients.

When the **Do Not Forward** option is applied to an email, the email is encrypted and recipients must be authenticated. Then, the recipients can't forward it, print it, or copy from it. For example, in the Outlook client, the Forward button isn't available, the **Save As** and **Print** menu options are not available, and you can't add or change recipients in the **To**, **Cc**, or **Bcc** boxes.

Unencrypted Office documents that are attached to the email automatically inherit the same restrictions. The usage rights applied to these documents are **Edit Content, Edit**; **Save**; **View, Open, Read**; and **Allow Macros**. If you want different usage rights for an attachment, or your attachment isn't an Office document that supports this inherited encryption, encrypt the file before you attach it to the email. You can then assign the specific usage rights that you need for the file.

Difference between Do Not Forward and not granting the Forward usage right

There's an important distinction between applying the **Do Not Forward** option and applying a rights management template that doesn't grant the **Forward** usage right to an email: The **Do Not Forward** option uses a dynamic list of authorized users that is based on the user's chosen recipients of the original email; whereas the rights in the template have a static list of authorized users that the administrator has previously specified. What's the difference? Let's take an example:

A user wants to email some information to specific people in the Marketing department that shouldn't be shared with anybody else. Should they protect the email with a rights management template that restricts rights (viewing, replying, and saving) to the Marketing department? Or should they choose the **Do Not Forward** option? Both choices would result in the recipients not able to forward the email.

- If they applied the template, the recipients could still share the information with others in the marketing department. For example, a recipient could use Explorer to drag and drop the email to a shared location or a USB drive. Now, anybody from the marketing department (and the email owner) who has access to this location can view the information in the email.
- If they applied the **Do Not Forward** option, the recipients will not be able to share the information with anybody else in the marketing department by moving the email to another location. In this scenario, only the original recipients (and the email owner) will be able to view the information in the email.

> ⓘ **Note**
>
> Use **Do Not Forward** when it's important that only the recipients that the sender chooses should see the information in the email. Use a template for emails to restrict rights to a group of people that the administrator specifies in advance, independently from the sender's chosen recipients.

Encrypt-only option for emails

When Exchange Online uses Microsoft Purview Message Encryption, a new **Encrypt** email option becomes available, to encrypt the data without additional restrictions.

This option is available to tenants who use Exchange Online and can be selected as follows:

- **In Outlook on the web** with the **Encrypt** option or a sensitivity label that's configured for **Let users assign permissions** and the **Encrypt-Only** option
- **As another rights protection option** for a mail flow rule

- **As Microsoft Purview DLP action**
- **From the Outlook app** for desktops and mobile devices:
  - With a sensitivity label that's configured for **Let users assign permissions** and the **Encrypt-Only** option, for Windows, macOS, iOS and Android when you use sensitivity labels with the minimum versions listed in the table for sensitivity label capabilities in Outlook.
  - With the **Encrypt** option on Windows and macOS, for the versions listed in the table of supported versions for Microsoft 365 Apps by update channel.

For more information about the encrypt-only option, see the following blog post when it was first announced from the Office team: Encrypt only rolling out in Office 365 Message Encryption.

When this option is selected, the email is encrypted and recipients must be authenticated. Then, the recipients have all usage rights except **Save As, Export** and **Full Control**. This combination of usage rights means that the recipients have no restrictions except that they can't remove the encryption. For example, a recipient can copy from the email, print it, and forward it.

Similarly, by default, unencrypted Office documents that are attached to the email inherit the same permissions. These documents are automatically encryption and when they're downloaded, they can be saved, edited, copied, and printed from Office applications by the recipients. When the document is saved by a recipient, it can be saved to a new name and even a different format. However, only file formats that support Rights Management encryption are available so that the document can't be saved without the original encryption. If you want different usage rights for an attachment, or your attachment isn't an Office document that supports this inherited encryption, encrypt the file before you attach it to the email. You can then assign the specific usage rights that you need for the file.

Alternatively, you can change this encryption inheritance of documents by specifying `Set-IRMConfiguration -DecryptAttachmentForEncryptOnly $true` with Exchange Online PowerShell. Use this configuration when you don't need to retain the original encryption for the document after the user is authenticated. When recipients open the email message, the document isn't encrypted.

> **Note**
>
> If you see references to **DecryptAttachmentFromPortal**, this parameter is now deprecated for Set-IRMConfiguration. Unless you have previously set this parameter, it's not available.

Automatically encrypt PDF documents with Exchange Online

When Exchange Online uses Microsoft Purview Message Encryption, you can automatically encrypt PDF documents when they're attached to an encrypted email. The document inherits the same permissions as those for the email message. To enable this configuration, set **EnablePdfEncryption $True** with Set-IRMConfiguration.

Recipients can use Microsoft Edge to view the encrypted PDF document. Alternatively, recipients can read the encrypted PDF document in the OME portal.

Rights Management issuer and Rights Management owner

When an item is encrypted by using the Azure Rights Management service, the account that encrypts that content automatically becomes the Rights Management issuer for that content. This account is logged as the **issuer** field in the usage logs.

The Rights Management issuer is always granted the Full Control usage right for the item, and in addition:

- If the encryption settings include an expiry date, the Rights Management issuer can still open and edit the document or email after that date.
- The Rights Management issuer can always access the document or email offline.
- The Rights Management issuer can still open a document after it's revoked.

By default, this account is also the **Rights Management owner** for that content, which is the case when a user who created the item initiates the encryption. But there are some scenarios where an administrator or service can encrypt content on behalf of users. For example:

- An administrator bulk-encrypts files on a file share: The administrator account in Microsoft Entra ID encrypts the documents for the users.
- The Rights Management connector encrypts Office documents on a Windows Server folder: The service principal account in Microsoft Entra ID that is created for the Rights Management connector encrypts the documents for the users.

In these scenarios, the Rights Management issuer can assign the Rights Management owner to another account by using the Microsoft Information Protection SDKs or PowerShell. For example, when you use the Set-LabelFile PowerShell cmdlet with the Microsoft Purview

Information Protection client, you can specify the *Owner* parameter to assign the Rights Management owner to another account.

When the Rights Management issuer encrypts on behalf of users, assigning the Rights Management owner ensures that the original item owner has the same level of control for their encrypted content as if they initiated the encryption themselves.

For example, the user who created the document can print it, even though it's now labeled with encryption settings that don't include the Print usage right. The same user can always access their document, regardless of the offline access setting or expiry date that might have been configured in the encryption settings. In addition, because the Rights Management owner has the Full Control usage right, this user can also re-encrypt the document to grant additional users access (at which point the user then becomes the Rights Management issuer as well as the Rights Management owner), and this user can even remove the encryption. However, only the Rights Management issuer can track and revoke a document.

The Rights Management owner for an item is logged as the **owner-email** field in the usage logs.

> **Note**
>
> The Rights Management owner is independent from the Windows file system Owner. They're often the same but can be different, even if you don't use the SDKs or PowerShell.

Rights Management use license

When a user opens an item that has been encrypted by the Azure Rights Management service, a Rights Management use license for that content is granted to the user. This use license is a certificate that contains the user's usage rights for item, and the encryption key that was used to encrypt the content. The use license also contains an expiry date if this has been set, and how long the use license is valid.

A user must have a valid use license to open the content in addition to their rights account certificate (RAC), which is a certificate that's granted when the user environment is initialized and then renewed every 31 days.

For the duration of the use license, the user isn't reauthenticated or reauthorized for the content. This lets the user continue to open the encrypted item without an internet connection. When the use license validity period expires, the next time the user accesses the encrypted item, the user must be reauthenticated and reauthorized.

When items are encrypted by using a sensitivity label that defines the encryption settings, you can change these settings in your sensitivity label without having to encrypt the content again. If the user has already accessed the content, the changes take effect after their use license has expired. However, when users apply permissions themselves (also known as user-defined permissions, custom permissions, or an ad-hoc rights policy) and these permissions need to change after the item is encrypted, that content must be encrypted again with the new permissions. User-defined permissions for an email message are implemented with the Do Not Forward option.

The default use license validity period for a tenant is 30 days and you can configure this value by using the PowerShell cmdlet, Set-AipServiceMaxUseLicenseValidityTime. You can configure a more restrictive setting for when encryption is applied by using a sensitivity label that's configured to assign permissions now, or a rights management template:

- When you configure a sensitivity label, the use license validity period takes its value from the **Allow offline access** setting.

  For more information and guidance to configure this setting for a sensitivity label, see the recommendations table from the instructions how to configure permissions now for a sensitivity label.

- When you configure a rights management template by using PowerShell, the use license validity period takes its value from the *LicenseValidityDuration* parameter in the Set-AipServiceTemplateProperty and Add-AipServiceTemplate cmdlets.

  For more information and guidance to configure this setting by using PowerShell, see the help for each cmdlet.

Related documentation

- Restrict access to content by using sensitivity labels to apply encryption
- Configure Azure Rights Management super users for discovery services or data recovery

Manage personal data for the Azure Rights Management encryption service

> *Microsoft Purview service description*

When you configure and use the Azure Rights Management encryption service from Microsoft Purview Information Protection, email addresses and IP addresses are stored and used by this service. The personal data can be found in the following items:

- Super users and delegated administrators for the Azure Rights Management service
- Administration logs for the Azure Rights Management service
- Usage logs for the Azure Rights Management service
- Usage logs for the Microsoft Purview Information Protection client and RMS client

> **Note**
>
> This article provides steps for how to delete personal data from the device or service and can be used to support your obligations under the GDPR. If you're looking for general info about GDPR, see the GDPR section of the Service Trust portal.

Viewing personal data that the Azure Rights Management service uses

- **Microsoft Purview Information Protection client**:

  For the Microsoft Purview Information Protection client, sensitivity labels and label policies are configured in the Microsoft Purview portal. For more information, see Create and configure sensitivity labels and their policies.

> **Note**
>
> When the Microsoft Purview Information Protection client is used to encrypt documents and emails, email addresses and the users' IP addresses might be saved in log files.

Super users and delegated administrators for the Azure Rights Management service

Run the Get-AipServiceSuperUser cmdlet and get-aipservicerolebasedadministrator cmdlet to see which users have been assigned the super user role or global administrator role for the Azure Rights Management service. For users who have been assigned either of these roles, their email addresses are displayed.

Administration logs for the Azure Rights Management service

Run the Get-AipServiceAdminLog cmdlet to get a log of admin actions for the Azure Rights Management service. This log includes personal data in the form of email addresses and IP addresses. The log is in plaintext and after it is downloaded, the details of a specific administrator can be searched offline.

For example:

PowerShell

```
PS C:\Users> Get-AipServiceAdminLog -Path '.\Desktop\admin.log' -FromTime 4/1/2018 -
ToTime 4/30/2018 -Verbose
The Rights Management administration log was successfully generated and can be found
at .\Desktop\admin.log.
```

Usage logs for the Azure Rights Management service

Run the Get-AipServiceUserLog cmdlet to retrieve a log of end-user actions that use the Azure Rights Management service. The log could include personal data in the form of email addresses and IP addresses. The log is in plaintext and after it is downloaded, the details of a specific administrator can be searched offline.

For example:

PowerShell

```
PS C:\Users> Get-AipServiceUserLog -Path '.\Desktop\' -FromDate 4/1/2018 -ToDate
4/30/2018 -NumberOfThreads 10
Acquiring access to your user log…
Downloading the log for 2018-04-01.
Downloading the log for 2018-04-03.
Downloading the log for 2018-04-06.
Downloading the log for 2018-04-09.
Downloading the log for 2018-04-10.
Downloaded the log for 2018-04-01. The log is available at .\Desktop\rmslog-2018-04-
```

```
01.log.
Downloaded the log for 2018-04-03. The log is available at .\Desktop\rmslog-2018-04-
03.log.
Downloaded the log for 2018-04-06. The log is available at .\Desktop\rmslog-2018-04-
06.log.
Downloaded the log for 2018-04-09. The log is available at .\Desktop\rmslog-2018-04-
09.log.
Downloaded the log for 2018-04-10. The log is available at .\Desktop\rmslog-2018-04-
10.log.
Downloading the log for 2018-04-12.
Downloading the log for 2018-04-13.
Downloading the log for 2018-04-14.
Downloading the log for 2018-04-16.
Downloading the log for 2018-04-18.
Downloaded the log for 2018-04-12. The log is available at .\Desktop\rmslog-2018-04-
12.log.
Downloaded the log for 2018-04-13. The log is available at .\Desktop\rmslog-2018-04-
13.log.
Downloaded the log for 2018-04-14. The log is available at .\Desktop\rmslog-2018-04-
14.log.
Downloaded the log for 2018-04-16. The log is available at .\Desktop\rmslog-2018-04-
16.log.
Downloaded the log for 2018-04-18. The log is available at .\Desktop\rmslog-2018-04-
18.log.
Downloading the log for 2018-04-24.
Downloaded the log for 2018-04-24. The log is available at .\Desktop\rmslog-2018-04-
24.log.
```

Usage logs for the Microsoft Purview Information Protection client and RMS client

When labels and encryption are applied to documents and emails, email addresses and IP addresses can be stored in log files on a user's computer in the following locations:

- For the Microsoft Purview Information Protection client: **%localappdata%\Microsoft\MSIP\Logs**
- For the RMS client: **%localappdata%\Microsoft\MSIPC\msip\Logs**

In addition, the Microsoft Purview Information Protection client logs this personal data to the local Windows event log **Applications and Services Logs** > **Microsoft Purview Information Protection**.

When the Microsoft Purview Information Protection client runs the scanner, personal data is saved to **%localappdata%\Microsoft\MSIP\Scanner\Reports** on the Windows Server computer that runs the scanner.

You can turn off logging information for the Microsoft Purview Information Protection client and the scanner by using the following configurations:

- For the Microsoft Purview Information Protection client: Edit the registry to change the local logging level for the Information Protection client to **Off**.
- For the Microsoft Purview Information Protection scanner: Use the Set-ScannerConfiguration cmdlet to set the *ReportLevel* parameter to **Off**.

> **Note**
>
> If you're interested in viewing or deleting personal data, please review Microsoft's guidance in the Microsoft Purview Compliance Manager and in the GDPR section of the Microsoft 365 Enterprise Compliance site. If you're looking for general information about GDPR, see the GDPR section of the Service Trust portal.

Document tracking logs

Relevant for: The Azure Rights Management service with the legacy tracking portal

Run the Get-AipServiceDocumentLog cmdlet to retrieve information from the document tracking site about a specific user. To get tracking information associated with the document logs, use the Get-AipServiceTrackingLog cmdlet.

For example:

PowerShell

```
PS C:\Users> Get-AipServiceDocumentLog -UserEmail "admin@aip500.onmicrosoft.com"


ContentId             : 6326fcb2-c465-4c24-a7f6-1cace7a9cb6f
Issuer                : admin@aip500.onmicrosoft.com
Owner                 : admin@aip500.onmicrosoft.com
ContentName           :
CreatedTime           : 3/6/2018 10:24:00 PM
Recipients            : {
                        PrimaryEmail: johndoe@contoso.com
                        DisplayName: JOHNDOE@CONTOSO.COM
                        UserType: External,
                        PrimaryEmail: alice@contoso0110.onmicrosoft.com
                        DisplayName: ALICE@CONTOSO0110.ONMICROSOFT.COM
                        UserType: External
                        }
TemplateId            :
```

```
PolicyExpires         :
EULDuration           :
SendRegistrationEmail : True
NotificationInfo      : Enabled: False
                        DeniedOnly: False
                        Culture:
                        TimeZoneId:
                        TimeZoneOffset: 0
                        TimeZoneDaylightName:
                        TimeZoneStandardName:

RevocationInfo        : Revoked: False
                        RevokedTime:
                        RevokedBy:


PS C:\Users> Get-AipServiceTrackingLog -UserEmail "admin@aip500.onmicrosoft.com"

ContentId             : 6326fcb2-c465-4c24-a7f6-1cace7a9cb6f
Issuer                : admin@aip500.onmicrosoft.com
RequestTime           : 3/6/2018 10:45:57 PM
RequesterType         : External
RequesterEmail        : johndoe@contoso.com
RequesterDisplayName  : johndoe@contoso.com
RequesterLocation     : IP: 167.220.1.54
                        Country/region: US
                        City: redmond
                        Position: 47.6812453974602,-122.120736471666

Rights                : {VIEW,OBJMODEL}
Successful            : False
IsHiddenInfo          : False
```

There is no search by ObjectID. However, you are not restricted by the `-UserEmail` parameter and the email address you provide doesn't need to be part of your tenant. If the email address provided is stored anywhere in the document tracking logs, the document tracking entry is returned in the cmdlet output.

Securing and controlling access to personal information

Personal data that you view and specify in the Microsoft Purview portal is accessible only to users who have been assigned one of the following Microsoft Entra ID built-in administrator roles:

- **Azure Information Protection administrator**
- **Compliance administrator**
- **Compliance data administrator**

- **Security administrator**
- **Security reader**
- **Global administrator**
- **Global reader**

Personal data that you view and specify by using the AIPService module is accessible only to users who have been assigned the **Azure Information Protection administrator**, **Compliance administrator**, **Compliance data administrator**, or **Global Administrator** roles from Microsoft Entra ID, or the global administrator role for the Azure Rights Management service.

Updating personal data

Sensitivity labels and label policies are configured in the Microsoft Purview. For more information, see the Create and configure sensitivity labels and their policies.

For the encryption settings, you can update the related personal data by using PowerShell cmdlets from the AIPService module.

You cannot update email addresses for the super users and delegated administrators. Instead, remove the specified user account, and add the user account with the updated email address.

Super users and delegated administrators for the Azure Rights Management service

When you need update an email address for a super user:

1. Use Remove-AipServiceSuperUser to remove the user and old email address.
2. Use Add-AipServiceSuperUser to add the user and new email address.

When you need update an email address for a delegated administrator:

1. Use Remove-AipServiceRoleBasedAdministrator to remove the user and old email address.
2. Use Add-AipServiceRoleBasedAdministrator to add the user and new email address.

Deleting personal data

Sensitivity labels and label policies are configured in the Microsoft Purview portal. For more information, see Create and configure sensitivity labels and their policies.

For the encryption settings, you can delete the related personal data by using PowerShell cmdlets from the AIPService module.

To delete email addresses for super users and delegated administrators, remove these users by using the Remove-AipServiceSuperUser cmdlet and Remove-AipServiceRoleBasedAdministrator.

To delete personal data in document tracking logs, administration logs, or usage logs for the Azure Rights Management service, raise a request with Microsoft Support by using the procedure in To delete personal data with Microsoft Support.

To delete personal data in the client log files and scanner logs that are stored on computers, use any standard Windows tools to delete the files or personal data within the files.

To delete personal data with Microsoft Support

Use the following three steps to request that Microsoft deletes personal data in document tracking logs, administration logs, or usage logs for the Azure Rights Management service.

Before you submit a delete request, make sure you can prove that you are an administrator for your tenant. Also be aware that this process takes several days to confirm.

Step 1: Initiate delete request Contact Microsoft Support to open Microsoft Purview Information Protection support case with a request for deleting data from your tenant. When you submit your request, you need to provide additional information, depending on the data that needs to be deleted.

- To delete the administration log, provide the **end date**. All admin logs until that end date will be deleted.
- To delete the usage logs, provide the **end date**. All usage logs until that end date will be deleted.
- To delete the document tracking logs, provide the **end date** and **UserEmail**. All document tracking information relating to the UserEmail until that end date will be deleted. Regular expression (Perl format) for UserEmail is supported.

Deleting this data is a permanent action. There is no means to recover the data after a delete request has been processed. It's recommended that administrators export the required data before submitting a delete request.

Step 2: Wait for verification Microsoft will verify that your request to delete one or more logs is legitimate. This process can take up to five working days.

Step 3: Get confirmation of the deletion Microsoft Customer Support Services (CSS) will send you a confirmation email that the data has been deleted.

Exporting personal data

When you use the AIPService cmdlets, the personal data is made available for search and export as a PowerShell object. The PowerShell object can be converted into JSON and saved by using the `ConvertTo-Json` cmdlet.

Restricting the use of personal data for profiling or marketing without consent

The Azure Rights Management service follows Microsoft's Microsoft Privacy Statement for profiling or marketing based on personal data.

Auditing and reporting

Only users who have been assigned the required administrator roles for accessing personal information can use the AIPService for search and export of personal data. These operations are recorded in the administration log that can be downloaded.

For delete actions, the support request acts as the auditing and reporting trail for the actions performed by Microsoft. After deletion, the deleted data will not be available for search and export, and the administrator can verify this using the Get cmdlets from the AIPService module.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Track and revoke document access

> *Microsoft Purview service description*

Document tracking provides information for administrators about when a protected document was accessed. If necessary, both admins and users can revoke document access for tracked documents.

A document must be registered for tracking before an admin can track access details, including successful access events and denied attempts, and revoke access if needed. For minimum versions of Office apps that support file registration the next time they're opened, see Requirements.

> **Note**
>
> Track and revoke features are supported for Office file types only.

Requirements

Use the capabilities table and the row **Document tracking and revocation** to identify the minimum versions of Word, Excel, and PowerPoint that automatically register label-protected local Office documents (if not already registered) the next time they're opened.

PowerShell cmdlets in this article use the AIPService PowerShell module, which you can install from the PowerShell Gallery. You must run Connect-AipService to connect to your tenant before you run any of the documented cmdlets.

Limitations

The following limitations apply to track and revoke features:

- Password-protected documents aren't supported by track and revoke features.
- If you attach multiple documents to an email, and then protect the email and send it, each of the attachments gets the same ContentID value. This ContentID value will be returned only with the first file that had been opened. Searching for the other attachments won't return the ContentID value required to get tracking data.

Additionally, revoking access for one of the attachments also revokes access for the other attachments in the same protected email.

- Documents protected with admin-defined permissions that are uploaded to SharePoint or OneDrive lose their **ContentID** value, and access can't be tracked or revoked.

- If a user downloads a file protected with admin-defined permissions from SharePoint or OneDrive, a new **ContentID** is applied to the document. Using the original **ContentID** value to track data won't include any access performed for the user's downloaded file. Additionally, revoking access based on the original **ContentID** value won't revoke access for any of the downloaded files.

  If administrators have access to the downloaded files, they can use PowerShell to identify a document's **ContentID** for track and revoke actions.

Track document access

Admins can track access for protected documents via PowerShell using the **ContentID** generated for the protected document during registration.

To view document access details:

Use the following cmdlets to find details for the document you want to track:

1. Find the **ContentID** value for the document you want to track.

   Use the Get-AipServiceDocumentLog cmdlet to search for a document using the filename or the email address of the user who applied protection.

   For example;

   PowerShell

   ```
   Get-AipServiceDocumentLog -ContentName "test.docx" -Owner "alice@contoso.com" -
   FromTime "12/01/2020 00:00:00" -ToTime "12/31/2020 23:59:59"
   ```

   This command returns the **ContentID** for all matching, protected documents that are registered for tracking.

   > ⓘ Note

Protected documents are registered for tracking when they're first opened in an Office app that supports file registration. If this command does not return the ContentID for your protected file, open it in a supported Office app. For minimum supported versions, see Office apps that support file registration.

2. Use the Get-AipServiceTrackingLog cmdlet with your document's **ContentID** to return your tracking data.

   For example:

   ```
   PowerShell
   Get-AipServiceTrackingLog -ContentId c03bf90c-6e40-4f3f-9ba0-2bcd77524b87
   ```

   Tracking data is returned, including emails of users who attempted access, whether access was granted or denied, the time and date of the attempt, and the domain and location where the access attempt originated.

Revoke document access from PowerShell

Admins can revoke access for any protected document stored in their local content shares, using the Set-AIPServiceDocumentRevoked cmdlet.

> **Note**
>
> If offline access is allowed, users will continue to be able to access the documents that have been revoked until the offline policy period expires.

1. Find the **ContentID** value for the document you want to revoke access for.

   Use the Get-AipServiceDocumentLog cmdlet to search for a document using the filename or the email address of the user who applied protection.

   For example:

   ```
   PowerShell
   Get-AipServiceDocumentLog -ContentName "test.docx" -Owner "alice@contoso.com" -
   FromTime "12/01/2020 00:00:00" -ToTime "12/31/2020 23:59:59"
   ```

   The data returned includes the ContentID value for your document.

> **Tip**
>
> Only documents that have been protected and registered for tracking have a **ContentID** value. If your document has no **ContentID**, open it in a supported Office app. For minimum supported versions, see **Office apps that support file registration**.

2. Use the Set-AIPServiceDocumentRevoked with your document's ContentID to revoke access.

   For example:

   ```
   PowerShell
   Set-AipServiceDocumentRevoked -ContentId 0e421e6d-ea17-4fdb-8f01-93a3e71333b8 -
   IssuerName testIssuer
   ```

By using the **Sensitivity** menu in their Office apps, users can also revoke access for any documents that they protected.

Restore access

If you've accidentally revoked access to a specific document, use the same **ContentID** value with the Clear-AipServiceDocumentRevoked cmdlet to remove the revoked status so that authorized users can access the document again. Specify the same **IssuerName** value that was used when the document was revoked.

For example:

```
PowerShell
Clear-AipServiceDocumentRevoked -ContentId   0e421e6d-ea17-4fdb-8f01-93a3e71333b8 -
IssuerName testIssuer
```

Document access is granted to the user you defined in the **IssuerName** parameter.

Turn off track and revoke features for your tenant

If you need to turn off track and revoke features for your tenant, such as for privacy requirements in your organization or region, run the Disable-AipServiceDocumentTrackingFeature cmdlet.

Document tracking and options to revoke access are turned off for your tenant:

- Opening protected documents no longer registers the documents for track and revoke.
- Access logs aren't stored when protected documents that are already registered are opened. Access logs that were stored before turning off these features are still available.
- Admins won't be able to track or revoke access via PowerShell, and although end-users still see the **Revoke** menu option in their Office apps, the site displays a message that tracking and revocation has been disabled by their administrator.

If you need to turn track and revoke back on, run the Enable-AipServiceDocumentTrackingFeature cmdlet.

Remove the track and revoke options from the Sensitivity menu

If you need to remove the Track & Revoke button from the Sensitivity menu on Office clients, use PowerShell advanced settings with the Set-LabelPolicy cmdlet.

The following example disables revoke support in the client UI by setting the `EnableRevokeGuiSupport` advanced setting to `False` in the global label policy:

PowerShell

```
Set-LabelPolicy -Identity Global -AdvancedSettings @{EnableRevokeGuiSupport="False"}
```

End user instructions for Office apps

See Track and revoke access to your files.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Usage logging for the Azure Rights Management service

The Azure Rights Management service is a cloud-based encryption technology used by Microsoft Purview Information Protection. It helps protect files and emails across devices by applying usage rights and encryption policies.

Use this information to help you understand how you can use usage logging for the encryption service, Azure Rights Management from Microsoft Purview Information Protection. This encryption service provides additional data protection for your organization's items such as documents and emails, and it can log every request. These requests include:

- When users encrypt items to protect them, and decrypt to read them or remove encryption.
- Actions performed by your administrators to manage the Azure Rights Management service, and actions performed by Microsoft operators to support the service.

You can then use these usage logs to support the following business scenarios:

- **Analyze for business insights**

  The logs generated by Azure Rights Management can be imported into a repository of your choice (such as a database, an online analytical processing (OLAP) system, or a map-reduce system) to analyze the information and produce reports. As an example, you could identify who is accessing your encrypted data. You can determine what encrypted data people are accessing, and from what devices and from where. You can find out whether people can successfully read encrypted content. You can also identify which people have read an important document that was encrypted.

- **Monitor for abuse**

  Logging information about how the Azure Rights Management service is used is available to you in near-real time, so that you can continuously monitor your company's use of the service. 99.9% of logs are available within 15 minutes of an initiated action to the service.

  For example, you might want to be alerted if there's a sudden increase of people reading encrypted data outside standard working hours, which could indicate that a malicious user is collecting information to sell to competitors. Or, if the same user apparently accesses data from two different IP addresses within a short time frame, which could indicate that a user account has been compromised.

- **Perform forensic analysis**

  If you have an information leak, you're likely to be asked who recently accessed specific documents and what information did a suspected person access recently. You can answer these types of questions when you use Azure Rights Management usage logging because people who use encrypted content must always get a Rights Management use license to open items that are encrypted by Azure Rights Management, even if these items are moved by email or copied to USB drives or other storage devices. This means that you can use these logs as a definitive source of information for forensic analysis when you protect your data by using the Azure Rights Management service.

Additional logging options for the Azure Rights Management service:

Expand table

Logging option	Description
Admin log	Logs administrative tasks for the Azure Rights Management service. For example, if the service is deactivated, when the super user feature is enabled, and when users are

Logging option	Description
	delegated admin permissions to the service. For more information, see the PowerShell cmdlet, Get-AipServiceAdminLog.
Document tracking	Lets users track and revoke their documents that they have encrypted with the Microsoft Purview Information Protection client. The Compliance Data Administrator role can also track these documents on behalf of users. For more information, see Track and revoke document access.

The remainder of this article explains how to download and interpret Azure Rights Management usage logs, including log format details and a PowerShell reference.

How to access and use your Azure Rights Management usage logs

Azure Rights Management usage logging is enabled by default for all customers. There's no extra cost for the log storage or for the logging feature functionality.

The Azure Rights Management service writes logs as a series of blobs to an Azure storage account that it automatically creates for your tenant. Each blob contains one or more log records, in W3C extended log format. The blob names are numbers, in the order in which they were created. For more information about log contents and their creation, see How to interpret your Azure Rights Management usage logs.

It can take a while for logs to appear in your storage account after an Azure Rights Management action. Most logs appear within 15 minutes. Usage logs are only available when the "date" field name contains a value of a previous date (in UTC time). Usage logs from the current date aren't available. We recommend that you download the logs to local storage, such as a local folder, a database, or a big-data processing system.

To download your usage logs, you use the AIPService PowerShell module for Microsoft Purview Information Protection. For installation instructions, see Install the AIPService PowerShell module for the Azure Right Management service.

To download your usage logs by using PowerShell

Use the following steps to connect to the Azure Rights Management service and download your usage logs by using PowerShell.

1. Start PowerShell with the **Run as administrator** option and use the Connect-AipService cmdlet to connect to the Azure Rights Management service:

   ```PowerShell
   Connect-AipService
   ```

2. Run the following command to download the logs for a specific date:

   ```PowerShell
   Get-AipServiceUserLog -Path <location> -fordate <date>
   ```

   For example, after creating a folder called Logs on your E: drive:

   - To download logs for a specific date (such as 2/1/2025), run the following command: `Get-AipServiceUserLog -Path E:\Logs -fordate 2/1/2025`
   - To download logs for a date range (such as from 2/1/2025 through 2/14/2025), run the following command: `Get-AipServiceUserLog -Path E:\Logs -fromdate 2/1/2025 –`

`todate 2/14/2025`

When you specify the day only, as in our examples, the time is assumed to be 00:00:00 in your local time, and then converted to UTC. When you specify a time with your -fromdate or -todate parameters (for example, -fordate "2/1/2025 15:00:00"), that date and time is converted to UTC. The Get-AipServiceUserLog command then gets the logs for that UTC time period.

You can't specify less than a whole day to download.

By default, this cmdlet uses three threads to download the logs. If you have sufficient network bandwidth and want to decrease the time required to download the logs, use the -NumberOfThreads parameter, which supports a value from 1 through 32. For example, if you run the following command, the cmdlet spawns 10 threads to download the logs: `Get-AipServiceUserLog -Path E:\Logs -fromdate 2/1/2025 –todate 2/14/2025 -numberofthreads 10`

> **Tip**
>
> You can aggregate all your downloaded log files into a CSV format by using Microsoft's Log Parser, which is a tool to convert between various well-known log formats. You can also use this tool to convert data to SYSLOG format, or import it into a database. After you have installed the tool, run `LogParser.exe /?` for help and information to use this tool.
>
> For example, you might run the following command to import all information into a .log file format: `logparser –i:w3c –o:csv "SELECT * INTO AllLogs.csv FROM *.log"`

How to interpret your usage logs

This section explains the Azure Rights Management log sequence, blob format, log record fields, and common request types so you can parse and analyze your usage logs.

The log sequence

The Azure Rights Management service writes the logs as a series of blobs.

Each entry in the log has a UTC timestamp. Because the service runs on multiple servers across multiple data centers, sometimes the logs might seem to be out of sequence, even when they're sorted by their timestamp. However, the difference is small and usually within a minute. In most cases, this isn't an issue that would be a problem for log analysis.

The blob format

Each blob is in W3C extended log format. It starts with the following two lines:

#Software: RMS

#Version: 1.1

The first line identifies that these are usage logs from Azure Rights Management. The second line identifies that the rest of the blob follows the version 1.1 specification. We recommend that any applications that parse these logs verify these two lines before continuing to parse the rest of the blob.

The third line enumerates a list of field names that are separated by tabs:

#Fields: date time row-id request-type user-id result correlation-id content-id owner-email issuer template-id file-name date-published c-info c-ip admin-action acting-as-user

Each of the subsequent lines is a log record. The values of the fields are in the same order as the preceding line, and are separated by tabs. Use the following table to interpret the fields.

Expand table

Field name	W3C data type	Description	Example value
date	Date	UTC date when the request was served. The source is the local clock on the server that served the request.	2013-06-25
time	Time	UTC time in 24-hour format when the request was served. The source is the local clock on the server that served the request.	21:59:28
row-id	Text	Unique GUID for this log record. If a value isn't present, use the correlation-id value to identify the entry. This value is useful when you aggregate logs or copy logs into another format.	aaaa0000-bb11-2222-33cc-444444dddddd
request-type	Name	Name of the RMS API that was requested.	AcquireLicense
user-id	String	The user who made the request. The value is enclosed in single quotation marks. Calls from an Azure Rights Management tenant key that is managed by you (BYOK) have a value of '', which also applies when the request types are anonymous.	'joe@contoso.com'

Field name	W3C data type	Description	Example value
result	String	'Success' if the request was served successful. The error type in single quotation marks if the request failed.	'Success'
correlation-id	Text	GUID that is common between the corresponding client log and server log for a given request. This value can be useful to help troubleshoot client issues.	bbbb1111-cc22-3333-44dd-555555eeeeee
content-id	Text	GUID, enclosed in curly braces that identifies the encrypted content (for example, a document). This field has a value only if request-type is AcquireLicense and is blank for all other request types.	{bb4af47b-cfed-4719-831d-71b98191a4f2}
owner-email	String	Email address of the owner of the document. This field is blank if the request type is RevokeAccess.	alice@contoso.com
issuer	String	Email address of the document issuer. This field is blank if the request type is RevokeAccess.	alice@contoso.com (or) FederatedEmail.4c1f4d-93bf-00a95fa1e042@contoso.onmicrosoft.com'
template-id	String	ID of the rights management template used to encrypt the document. This field is	{6d9371a6-4e2d-4e97-9a38-202233fed26e}

Field name	W3C data type	Description	Example value
		blank if the request type is RevokeAccess.	
file-name	String	File name of an encrypted document that is tracked by using the Microsoft Purview Information Protection client. Currently, some files (such as Office documents) display as GUIDs rather than the actual file name. This field is blank if the request type is RevokeAccess.	TopSecretDocument.docx
date-published	Date	Date when the document was encrypted. This field is blank if the request type is RevokeAccess.	2015-10-15T21:37:00
c-info	String	Information about the client platform that is making the request. The specific string varies, depending on the application (for example, the operating system or the browser).	'MSIPC;version=1.0.623.47;AppName=WINWORD.EXE;AppVersion=15.0.4753.1000;AppArch=x86;O
c-ip	Address	IP address of the client that makes the request.	64.51.202.144
admin-action	Bool	Whether an administrator has accessed the document tracking site in Administrator mode.	True
acting-as-user	String	The email address of the user for whom an administrator is accessing	'joe@contoso.com'

Field name	W3C data type	Description	Example value
		the document tracking site.	

Exceptions for the user-id field

Although the user-id field usually indicates the user who made the request, there are two exceptions where the value doesn't map to a real user:

- The value '**microsoftrmsonline@<YourTenantID>.rms.<region>.aadrm.com**'.

  This indicates a Microsoft 365 service, such as Exchange or SharePoint, is making the request. In the string, *<YourTenantID>* is the GUID for your tenant and *<region>* is the region where your tenant is registered. For example, **na** represents North America, **eu** represents Europe, and **ap** represents Asia.

- If you're using the Rights Management connector.

  Requests from this connector are logged with the service principal name of **Aadrm_S-1-7-0**, which is automatically generated when you install the Rights Management connector.

Typical request types

There are many request types for the Azure Rights Management service but the following table identifies some of the most typically used request types.

Expand table

Request type	Description
AcquireLicense	A client from a Windows-based computer is requesting a use license for encrypted content.
AcquirePreLicense	A client, on behalf of the user, is requesting a use license for encrypted content.
AcquireTemplates	A call was made to acquire rights management templates based on template IDs
AcquireTemplateInformation	A call was made to get the IDs of the rights management template from the service.
AddTemplate	A call is made from an admin portal to add a rights management template.
AllDocsCsv	A call is made from the document tracking site to download the CSV file from the **All Documents** page.
BECreateEndUserLicenseV1	A call is made from a mobile device to create an end-user license.
BEGetAllTemplatesV1	A call is made from a mobile device (back-end) to get all the rights management templates.
Certify	The client is certifying the user for the consumption and creation of encrypted content.
FECreateEndUserLicenseV1	Similar to the AcquireLicense request but from mobile devices.
FECreatePublishingLicenseV1	The same as Certify and GetClientLicensorCert combined, from mobile clients.
FEGetAllTemplates	A call is made, from a mobile device (front-end) to get the rights management templates.
FindServiceLocationsForUser	A call is made to query for URLs, which is used to call Certify or AcquireLicense.
GetClientLicensorCert	The client is requesting a publishing certificate (that is later used to encrypt content) from a Windows-based computer.

Request type	Description
GetConfiguration	A PowerShell cmdlet is called to get the configuration of the tenant for the Azure Rights Management service.
GetConnectorAuthorizations	A call is made from the Rights Management connectors to get their configuration from the cloud.
GetRecipients	A call is made from the document tracking site to navigate to the list view for a single document.
GetTenantFunctionalState	An admin portal is checking whether the Azure Rights Management service is activated.
KeyVaultDecryptRequest	The client is attempting to decrypt the Rights Management encrypted content. Applicable only for a customer-managed tenant key (BYOK) in Azure Key Vault.
KeyVaultGetKeyInfoRequest	A call is made to verify that the key specified to be used in Azure Key Vault for the Azure Rights Management tenant key is accessible and not already used.
KeyVaultSignDigest	A call is made when a customer-managed key (BYOK) in Azure Key Vault is used for signing purposes. This is called typically once per AcquireLicence (or FECreateEndUserLicenseV1), Certify, and GetClientLicensorCert (or FECreatePublishingLicenseV1).
KMSPDecrypt	The client is attempting to decrypt the Rights Management encrypted content. Applicable only for a legacy customer-managed tenant key (BYOK).
KMSPSignDigest	A call is made when a legacy customer-managed key (BYOK) is used for signing purposes. This is called typically once per AcquireLicence (or FECreateEndUserLicenseV1), Certify, and GetClientLicensorCert (or FECreatePublishingLicenseV1).
ServerCertify	A call is made from a Rights Management-enabled client (such as SharePoint) to certify the server.
SetUsageLogFeatureState	A call is made to enable usage logging.
SetUsageLogStorageAccount	A call is made to specify the location of the Azure Rights Management service logs.
UpdateTemplate	A call is made from an admin portal to update an existing rights management template.

Azure Rights Management usage logs and Microsoft Purview auditing

File access and denied events don't include the file name and aren't accessible in the Microsoft Purview unified audit log.

PowerShell reference

After you connect to the Azure Rights Management service, the only PowerShell cmdlet that you need to access your Azure Rights Management usage logging is Get-AipServiceUserLog.

For more information about using PowerShell for the Azure Rights Management service, see Administering the Azure Rights Management service by using PowerShell.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

Last updated on 09/01/2026

Configure Azure Rights Management super users for discovery services or data recovery

> *Microsoft Purview service description*

The super user feature of the Azure Rights Management service from Microsoft Purview Information Protection ensures that authorized people and services can always read and inspect the data that Azure Rights Management encrypts for your organization. If necessary, the encryption protection can then be removed or changed.

A super user always has the Rights Management Full Control usage right for documents and emails that have been encrypted by your organization's tenant. This ability is sometimes referred to as "reasoning over data" and is a crucial element in maintaining control of your organization's data. For example, you would use this feature for any of the following scenarios:

- An employee leaves the organization and you need to read the files that they encrypted.
- An IT administrator needs to remove the current encryption settings that were configured for files and apply new encryption settings.
- Exchange Server needs to index mailboxes for search operations.
- You have existing IT services for data loss prevention (DLP) solutions, content encryption gateways (CEG), and anti-malware products that need to inspect files that are already encrypted.
- You need to bulk decrypt files for auditing, legal, or other compliance reasons.

Configuration for the super user feature

By default, the super user feature isn't enabled, and no users are assigned to this role. It's enabled for you automatically if you configure the Rights Management connector for Exchange, and it's not required for standard services that run Exchange Online, Microsoft SharePoint Server, or SharePoint in Microsoft 365.

If you need to manually enable the super user feature, use the PowerShell cmdlet Enable-AipServiceSuperUserFeature, and then assign users (or service accounts) as needed by using the

Add-AipServiceSuperUser cmdlet or the Set-AipServiceSuperUserGroup cmdlet and add users (or other groups) as needed to this group.

Although using a group for your super users is easier to manage, for performance reasons, the Azure Rights Management service caches the group membership. So if you need to assign a new user to be a super user to decrypt content immediately, add that user by using Add-AipServiceSuperUser, rather than adding the user to an existing group that you configured by using Set-AipServiceSuperUserGroup.

> **Note**
>
> - When adding a user with the Add-AipServiceSuperUser cmdlet, you must also add the primary mail address or user principal name to the group. Email aliases aren't evaluated.
> - If you haven't yet installed the PowerShell module for the Azure Rights Management service, see Install the AIPService PowerShell module for the Azure Right Management service.

It doesn't matter when you enable the super user feature or when you add users as super users. For example, if you enable the feature on Thursday and then add a user on Friday, that user can immediately open content that was protected at the very beginning of the week.

Security best practices for the super user feature

Follow these security best practices when using the super user feature:

- Restrict and monitor the administrators who are assigned a global administrator for your tenant, or who are assigned the GlobalAdministrator role by using the Add-AipServiceRoleBasedAdministrator cmdlet. These users can enable the super user feature and assign users (and themselves) as super users, and potentially decrypt all files that your organization encrypts.
- To see which users and service accounts are individually assigned as super users, use the Get-AipServiceSuperUser cmdlet.
- To see whether a super user group is configured, use the Get-AipServiceSuperUserGroup cmdlet and your standard user management tools to check which users are a member of this group.

- Like all administration actions, enabling or disabling the super feature, and adding or removing super users are logged and can be audited by using the Get-AipServiceAdminLog command. For example, see Example auditing for the super user feature.
- When super users decrypt files, file decryption is logged and can be audited with usage logging.

  > **Note**
  >
  > While the logs include details about the decryption, including the user who decrypted the file, they don't detail when the user is a super user.
  >
  > Use the logs together with Get-AipServiceSuperUser and Get-AipServiceSuperUserGroup to first collect a list of super users that you can identify in the logs.

- If you don't need the super user feature for everyday services, enable the feature only when you need it, and disable it again by using the Disable-AipServiceSuperUserFeature cmdlet.

Example auditing for the super user feature

The following log extract shows some example entries from using the Get-AipServiceAdminLog cmdlet.

In this example, the administrator for Contoso Ltd confirms that the super user feature is disabled, adds Richard Simone as a super user, checks that Richard is the only super user configured for the Azure Rights Management service, and then enables the super user feature so that Richard can now decrypt some files that were protected by an employee who has now left the company.

```
2015-08-01T18:58:20 admin@contoso.com GetSuperUserFeatureState Passed Disabled
```

```
2015-08-01T18:59:44 admin@contoso.com AddSuperUser -id rsimone@contoso.com Passed True
```

```
2015-08-01T19:00:51 admin@contoso.com GetSuperUser Passed rsimone@contoso.com
```

```
2015-08-01T19:01:45 admin@contoso.com SetSuperUserFeatureState -state Enabled Passed True
```

Scripting options for super users

Often, somebody who is assigned a super user for the Azure Rights Management service needs to remove encryption from multiple files, in multiple locations. While it's possible to do this task manually, it's more efficient (and often more reliable) to script this using the Set-FileLabel cmdlet.

You can also use this cmdlet to apply a new label that doesn't apply encryption, or remove the label that applied encryption.

For more information about these cmdlets, see Use PowerShell with the Microsoft Purview Information Protection client from the PurviewInformationProtection PowerShell documentation.

> ⓘ **Note**
>
> The PurviewInformationProtection module is different from and supplements the **AIPService PowerShell module** that manages the Azure Rights Management service for Microsoft Purview Information Protection.

Removing encryption from PST files

To remove encryption from PST files, we recommend that you use eDiscovery from Microsoft Purview to search and extract encrypted emails and encrypted attachments in emails.

The super user ability is automatically integrated with Exchange Online so that eDiscovery in the Microsoft Purview portal can search for encrypted items prior to export, or decrypt encrypted email on export.

If you can't use Microsoft Purview eDiscovery, you might have another eDiscovery solution that integrates with the Azure Rights Management service to similarly reason over data.

Or, if your eDiscovery solution can't automatically read and decrypt protected content, you can still use this solution in a multi-step process together with the Set-FileLabel cmdlet:

1. Export the email in question to a PST file from Exchange Online or Exchange Server, or from the workstation where the user stored their email.

2. Import the PST file into your eDiscovery tool. Because the tool can't read encrypted content, expect these items to generate errors.

3. From all the items that the tool couldn't open, generate a new PST file that this time, contains just encrypted items. This second PST file will likely be much smaller than the original PST file.

4. Run Set-FileLabel on this second PST file to decrypt the contents of this much smaller file. From the output, import the now-decrypted PST file into your discovery tool.

For more detailed information and guidance for performing eDiscovery across mailboxes and PST files, see the following blog post: Azure Information Protection and eDiscovery Processes.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 09/01/2026

Decommission and deactivate the Azure Rights Management service

> *Microsoft Purview service description*

You're always in control of whether your organization protects content by using the Azure Rights Management encryption service from Microsoft Purview Information Protection. If you decide you no longer want to use this encryption service, you have the assurance that you won't be locked out of content that was previously encrypted.

If you don't need continued access to previously encrypted content, deactivate the Azure Rights Management encryption services and let your supporting subscription for this service expire. For example, this would be appropriate for when you have completed testing Microsoft Purview Information Protection before you deploy it in a production environment.

However, if you have deployed Microsoft Purview Information Protection in production and encrypted items by using the Azure Rights Management service, make sure that you have a copy of your Azure Rights Management service root key and suitable trusted publishing domain (TPD) before you deactivate the Azure Rights Management service. Make sure that you have a copy of your key and the TPD before your subscription expires to ensure that you can retain access to content that was encrypted by Azure Rights Management after the service is deactivated.

If you used the bring your own key solution (BYOK) where you generate and manage your own key in an HSM, you already have your Azure Rights Management root key. You'll also have a suitable TPD if you followed the instructions that prepare for a future cloud exit. However, if your Azure Rights Management root key was managed by Microsoft (the default), see the instructions for exporting your root key in the Manage the root key for your tenant's Azure Rights Management service article.

> **Tip**
>
> Even after your subscription expires, your tenant remains available for consuming encrypted content for an extended period. However, you will no longer be able to export your root key from the Azure Rights Management service.

When you have your Azure Rights Management service root key and the TPD, you can deploy Rights Management on premises (AD RMS) and import your Azure Rights Management root key as a trusted publishing domain (TPD). You then have the following options for decommissioning your Microsoft Purview Information Protection deployment:

Expand table

If this applies to you ...	... do this:
You want all users to continue using Rights Management, but use an on-premises solution rather than using the Azure Rights Management service →	Redirect your clients to the on-premises deployment by using the **LicensingRedirection** registry key for Office apps. For instructions, see Configure service discovery in the RMS client deployment notes.
You want to stop using Rights Management technologies completely →	Grant a designated administrator super user rights for Microsoft Purview encryption and install the Microsoft Purview Information Protection client for this user. This administrator can then use the PowerShell module from this client to bulk-decrypt files in folders that were encrypted by the Azure Rights Management service. Files revert to being unencrypted and can therefore be read without a Rights Management technology such as Microsoft Purview Information Protection or AD RMS. Because this PowerShell module can be used with both Microsoft Purview Information Protection and AD RMS, you have the choice of decrypting files before or after you deactivate the Azure Rights Management service, or a combination.
You can't identify all the files that were encrypted by the Azure Rights Management service. Or, you want all users to be able to automatically read any encrypted items that were missed →	Deploy a registry setting on all client computers by using the **LicensingRedirection** registry key for Office apps, as described in Configure service discovery in the RMS client deployment notes.
You want a controlled, manual recovery service for any files that were missed →	Grant designated users in a data recovery group super user rights for Microsoft Purview encryption and install the Microsoft Purview Information Protection client for these users so that they can decrypt files when this action is requested by standard users. On all computers, deploy the registry setting to prevent users from encrypting new files by setting **DisableCreation** to **1**, as described in Office registry settings for the DisableCreation policy.

For more information about the decommissioning options for your Microsoft Purview Information Protection deployment, see the following resources:

- For information about AD RMS and deployment references, see Active Directory Rights Management Services Overview.
- For instructions to import your Azure Rights Management service root key as a TPD file, see Add a Trusted Publishing Domain.
- To use PowerShell with the Microsoft Purview Information Protection client, see Set up the information protection client using PowerShell.

When you're ready to deactivate the Azure Rights Management encryption service, use the following instructions.

Deactivate the Azure Rights Management service

You must use PowerShell to deactivate the Azure Rights Management service from Microsoft Purview Information Protection. You can no longer activate or deactivate this encryption service from admin portals.

Before you begin, ensure you have an account with permissions to manage the Azure Rights Management service, such as a Compliance Administrator or Compliance Data Administrator for your tenant.

1. Install the AIPService module, to configure and manage the Azure Rights Management service. For instructions, see Install the AIPService PowerShell module for the Azure Right Management service.
2. From a PowerShell session, run Connect-AipService, and when prompted, provide the credentials for an account that has permissions to manage the Azure Rights Management service, such as a Compliance Administrator or Compliance Data Administrator for your tenant.
3. Run Get-AipService to confirm the current status of the Azure Rights Management service. A status of **Enabled** confirms activation; **Disabled** indicates that the service is deactivated.
4. To deactivate the service, run Disable-AipService.
5. Run Get-AipService again to confirm the encryption service is now deactivated. This time, the status should display **Disabled**.
6. Run Disconnect-AipService to disconnect from the service, and close your PowerShell session.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Email encryption

This article compares encryption options in Microsoft 365 including Microsoft Purview Message Encryption, S/MIME, Information Rights Management (IRM), and introduces Transport Layer Security (TLS).

Microsoft 365 delivers multiple encryption options to help you meet your business needs for email security. This article presents three ways to encrypt email in Office 365. If you want to learn more about all security features in Office 365, visit the Office 365 Trust Center. This article introduces the three types of encryption available for Microsoft 365 administrators to help secure email in Office 365:

- Microsoft Purview Message Encryption.
- Secure/Multipurpose Internet Mail Extensions (S/MIME).
- Information Rights Management (IRM).

How Microsoft 365 uses email encryption

Encryption is the process by which information is encoded so that only an authorized recipient can decode and consume the information. Microsoft 365 uses encryption in two ways: in the service, and as a customer control. In the service, encryption is used in Microsoft 365 by default; you don't have to configure anything. For example, Microsoft 365 uses Transport Layer Security (TLS) to encrypt the connection, or session, between two servers.

Here's how email encryption typically works:

- A message is encrypted, or transformed from plain text into unreadable ciphertext, either on the sender's machine, or by a central server while the message is in transit.
- The message remains in ciphertext while it's in transit in order to protect it from being read in case the message is intercepted.
- Once the message is received by the recipient, the message is transformed back into readable plain text in one of two ways:
    - The recipient's machine uses a key to decrypt the message, or

- A central server decrypts the message on behalf of the recipient, after validating the recipient's identity.

For more information on how Microsoft 365 secures communication between servers, such as between organizations within Microsoft 365 or between Microsoft 365 and a trusted business partner outside of Microsoft 365, see How Exchange Online uses TLS to secure email connections in Office 365.

Comparing email encryption options available in Office 365

Expand table

Email encryption technology			
What is it?	Message encryption is a service built on the Azure Rights Management service that lets you send encrypted email to people inside or outside your organization, regardless of the destination email address (Gmail, Yahoo! Mail, Outlook.com, etc.). As an admin, you can set up transport rules that define the conditions for encryption. When a user sends a message that matches a rule, encryption is applied automatically. To view encrypted messages, recipients can either get a one-time passcode, sign in with a Microsoft account, or sign in with a work or school account associated with Office 365. Recipients can also send encrypted replies. They don't need a Microsoft 365 subscription to view	IRM is an encryption solution that also applies usage restrictions to email messages. It helps prevent sensitive information from being printed, forwarded, or copied by unauthorized people. IRM capabilities in Microsoft 365 use the Azure Rights Management service.	S/MIME is a certificate-based encryption solution that allows you to both encrypt and digitally sign a message. The message encryption helps ensure that only the intended recipient can open and read the message. A digital signature helps the recipient validate the identity of the sender. Both digital signatures and message encryption are made possible through the use of unique digital certificates that contain the keys for verifying digital signatures and encrypting or decrypting messages. To use S/MIME, you must have public keys on file for each recipient. Recipients have to maintain their own private keys, which must remain secure. If a recipient's private keys are compromised, the recipient

Email encryption technology	OME 	IRM 	S/MIME 
	encrypted messages or send encrypted replies.		needs to get a new private key and redistribute public keys to all potential senders.
What does it do?	OME: Encrypts messages sent to internal or external recipients. Allows users to send encrypted messages to any email address, including Outlook.com, Yahoo! Mail, and Gmail. Allows you, as an admin, to customize the email viewing portal to reflect your organization's brand. Microsoft securely manages and stores the keys, so you don't have to. No special client side software is needed as long as the encrypted message (sent as an HTML attachment) can be opened in a browser.	IRM: Uses encryption and usage restrictions to provide online and offline protection for email messages and attachments. Gives you, as an admin, the ability to set up transport rules or Outlook protection rules to automatically apply IRM to select messages. Lets users manually apply templates in Outlook or Outlook on the web (formerly known as Outlook Web App).	S/MIME addresses sender authentication with digital signatures, and message confidentiality with encryption.
What does it not do?	OME doesn't let you apply usage restrictions to messages. For example, you can't use it to stop a recipient from forwarding or printing an encrypted message.	Some applications may not support IRM emails on all devices.	S/MIME doesn't allow encrypted messages to be scanned for malware, spam, or policies.
Recommendations and example scenarios	We recommend using OME when you want to send sensitive business information to people outside your organization, whether they're consumers or other businesses. For example: A bank employee sending	We recommend using IRM when you want to apply usage restrictions as well as encryption. For example: A manager sending confidential details to her team about a new	We recommend using S/MIME when either your organization or the recipient's organization requires true peer-to-peer encryption. S/MIME is most commonly used in the following scenarios:

Email encryption technology	OME 	IRM 	S/MIME 
	credit card statements to customers A doctor's office sending medical records to a patient An attorney sending confidential legal information to another attorney	product applies the "Do Not Forward" option. An executive needs to share a bid proposal with another company, which includes an attachment from a partner who is using Office 365, and require both the email and the attachment to be protected.	Government agencies communicating with other government agencies A business communicating with a government agency

Don't apply multiple email encryption technologies to the same email message. Some email clients, such as Outlook for Mac, Outlook for iOS, and Outlook for Android, aren't able to open messages with multiple email encryption technologies applied.

What encryption options are available for my Microsoft 365 subscription?

For information about email encryption options for your Microsoft 365 subscription see the Exchange Online service description. Here, you can find information about the following encryption features:

- The Azure Rights Management service, including both IRM capabilities and Microsoft Purview Message Encryption
- S/MIME
- TLS
- Encryption of data at rest (through BitLocker, service encryption, and DKM)

You can also use third-party encryption tools with Microsoft 365, for example, PGP (Pretty Good Privacy). Microsoft 365 does not support PGP/MIME and you can only use PGP/Inline to send and receive PGP-encrypted emails.

What about encryption for data at rest?

"Data at rest" refers to data that isn't actively in transit. In Microsoft 365, email data at rest is encrypted using BitLocker Drive Encryption. BitLocker encrypts the hard drives in Microsoft data centers to provide enhanced protection against unauthorized access. To learn more, see BitLocker Overview.

More information about email encryption options

For more information about the email encryption options in this article as well as TLS, see these articles:

Microsoft Purview Message Encryption

Message encryption

IRM

Information Rights Management in Exchange Online

Learn about the Azure Rights Management encryption service

S/MIME

S/MIME for message signing and encryption

Understanding S/MIME

Understanding Public Key Cryptography

TLS

Configure custom mail flow by using connectors

How Exchange Online uses TLS to secure email connections

Disabling TLS 1.0 and 1.1 for Microsoft 365

Last updated on 05/04/2026

Message encryption

Article • 01/10/2025

People often use email to exchange sensitive information, such as financial data, legal contracts, confidential product information, sales reports and projections, patient health information, or customer and employee information. As a result, mailboxes can become repositories for large amounts of potentially sensitive information and information leakage can become a serious threat to your organization.

With Message encryption, your organization can send and receive encrypted email messages between people inside and outside your organization. Message encryption works with Outlook.com, Yahoo!, Gmail, and other email services. Email message encryption helps ensure that only intended recipients can view message content.

> **Tip**
>
> If you're not an E5 customer, use the 90-day Microsoft Purview solutions trial to explore how additional Purview capabilities can help your organization manage data security and compliance needs. Start now at the **Microsoft Purview trials hub**. Learn details about **signing up and trial terms**.

How Message Encryption works

The rest of this article applies to Microsoft Purview Message Encryption. Office 365 Message Encryption (OME) was deprecated.

Microsoft Purview Message Encryption is an online service that's built on Azure Rights Management (Azure RMS) which is part of Microsoft Purview Information Protection. This service includes encryption, identity, and authorization policies to help secure your email. You can encrypt messages by using rights management templates, the Do Not Forward option, and the encrypt-only option.

Users can then encrypt email messages and various attachments by using these options. For a full list of supported attachment types, see "File types covered by IRM policies when they're attached to messages" in Introduction to IRM for email messages.

As an administrator, you can also define mail flow rules to apply this protection. For example, you can create a rule that requires the encryption of all messages addressed to a specific recipient, or that contains specific words in the subject line, and also specify that recipients can't copy or print the contents of the message.

Message encryption provides a unified sender experience whether you're sending mail inside your organization or to recipients outside of Microsoft 365. In addition, recipients who receive a protected email message sent to a Microsoft 365 account in Outlook 2016 or Outlook on the web, don't have to take any other action to view the message. It works seamlessly. Recipients using other email clients and email service providers also have an improved experience. For information, see Learn about protected messages in Microsoft 365 and How do I open a protected message.

For a detailed list of the differences between the previous version of OME and Microsoft Purview Message Encryption, see Compare versions of message encryption.

When someone sends an email message that matches an encryption mail flow rule, the message is encrypted before it's sent. All Microsoft 365 end users that use Outlook clients to read mail receive native, first-class reading experiences for encrypted and rights-protected mail even if they're not in the same organization as the sender. Supported Outlook clients include Outlook desktop, Outlook Mac, Outlook mobile on iOS and Android, and Outlook on the web (formerly known as Outlook Web App).

Recipients of encrypted messages who receive encrypted or rights-protected mail sent to their Gmail and Yahoo accounts receive a wrapper mail that directs them to the encrypted message portal where they can easily authenticate using a Microsoft account, Gmail, or Yahoo credentials.

End users that read encrypted or rights-protected mail on clients other than Outlook also use the encrypted message portal to view encrypted and rights-protected messages that they receive.

If the sender of the protected mail is in GCC High and the recipient is outside of GCC High, including commercial users, Outlook.com users, and users of other email providers such as Gmail, the recipient receives a wrapper mail. The wrapper mail directs the recipient to the encrypted message portal where the recipient is able to read and reply to the message. Otherwise, if the sender and recipient are both in the GCC High environment, even if they're not in the same organization, then recipients that use Outlook clients to read mail receive native, first-class reading experiences for encrypted and rights-protected mail. For more information about the different experience in GCC High, see Compare versions of OME.

For more information about size limits for messages and attachments that you can encrypt using OME, see Exchange Online Limits.

How Microsoft Purview Advanced Message Encryption works on top of Microsoft Purview

Message Encryption

Microsoft Purview Advanced Message Encryption lets you create multiple branding templates so you can fine-tune control over recipient mail and create custom branding experiences to support a diverse organizational structure.

Advanced Message Encryption in Microsoft 365 helps you meet compliance obligations that require more flexible control over external recipient's access to encrypted emails. With Advanced Message Encryption, as an administrator, you can control sensitive emails shared outside the organization with automatic policies that detect sensitive information types (for example, personal data, Financial or Health IDs) or keywords to enhance protection by expiring access through a secure web portal to encrypted emails. As an admin you can further control encrypted emails accessed through a web portal by revoking access to an email anytime.

Message revocation and expiration only work for emails that your users send to recipients outside your organization. In addition, the recipients must access the email through the web portal. To ensure the recipient uses the portal to receive email, you set up a custom branding template that applies the wrapper. Then, you apply the branding template in a mail flow rule. For more information about Advanced Message Encryption, see Advanced Message Encryption.

Defining rules for Microsoft Purview Message Encryption

One way to enable Microsoft Purview Message Encryption is for Exchange Online and Exchange Online Protection administrators to define mail flow rules. These rules determine under what conditions email messages should be encrypted. When an encryption action is set for a rule, any messages that match the rule conditions are encrypted before they're sent.

Mail flow rules are flexible, letting you combine conditions so you can meet specific security requirements in a single rule. For example, you can create a rule to encrypt all messages that contain specified keywords and are addressed to external recipients. Microsoft Purview Message Encryption also encrypts replies from recipients of encrypted email.

For more information about how to create mail flow rules to take advantage of Microsoft Purview Message Encryption, see Define mail flow rules to encrypt email messages.

Get started with the Microsoft Purview Message Encryption

If you're ready to get started using Microsoft Purview Message Encryption within your organization, see Set up Microsoft Purview Message Encryption.

Sending, viewing, and replying to encrypted email messages

With Microsoft Purview Message Encryption, users can send encrypted email from Outlook and Outlook on the web. Additionally, admins can set up mail flow rules in Microsoft 365 to automatically encrypt emails based on keyword matching or other conditions.

Recipients of encrypted messages who are in organizations will be able to read those messages seamlessly in any version Outlook, including Outlook for PC, Outlook for Mac, Outlook on the web, Outlook for iOS, and Outlook for Android. Users that receive encrypted messages on other email clients can view the messages in the encrypted message portal.

For detailed guidance about how to send and view encrypted messages, take a look at these articles.

Expand table

Read this article...	If you are...
Learn about protected messages in Microsoft 365	An end user that wants to learn more about how encrypted messages work and what options are available to you.
How do I open a protected message?	An end user that wants to read a protected message that was sent to you. This article includes information about reading messages in several versions of Outlook and from different email accounts, including those accounts outside of Microsoft 365 such as gmail and Yahoo! accounts.
Send, view, and reply to encrypted messages in Outlook	An end user that wants to send, view, or reply to an encrypted message from Outlook. Even if you're not a member of an organization, you still receive notification of encrypted messages sent to you in Outlook. Use this article for instructions on how to view and reply to encrypted messages sent from Office 365.
Send a digitally signed or encrypted	An end user that wants to send, view, or reply to encrypted messages using Outlook for Mac. This article also covers using encryption methods

Read this article...	If you are...
message for Mac	other than OME, such as S/MIME.
View encrypted messages on your Android device	An end user who has received a message encrypted with Office 365 Message Encryption on your Android device, you can use the free OME Viewer app to view the message and send an encrypted reply. This article explains how.
View encrypted messages on your iPhone or iPad	An end user who has received a message encrypted with Office 365 Message Encryption on your iPhone or iPad, you can use the free OME Viewer app to view the message and send an encrypted reply. This article explains how.

Feedback

Was this page helpful? Yes No

Set up Message Encryption

Microsoft Purview Message Encryption allows organizations to share encrypted email with anyone on any device. Users can exchange encrypted messages with other Microsoft 365 organizations, as well as third-parties using Outlook.com, Gmail, and other email services.

Follow the steps below to ensure that Microsoft Purview Message Encryption is available in your organization.

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

Verify that Azure Rights Management is active

Microsoft Purview Message Encryption uses the protection features from the Azure Rights Management service, the encrytion technology used by Microsoft Purview Information Protection to protect emails and documents by using encryption and access controls.

The only prerequisite for using Microsoft Purview Message Encryption is that the Azure Rights Management service must be activated in your organization's tenant. If it is, Microsoft 365 activates message encryption automatically and you don't need to do anything.

The Azure Rights Management service is also activated automatically for most eligible plans, so you probably don't have to do anything in this regard either. For more information, see Activate the Azure Rights Management service.

> **Important**
>
> If you use Active Directory Rights Management service (AD RMS) with Exchange Online, you need to migrate to the Azure Rights Management service before you can use message encryption. Microsoft Purview Message Encryption is not compatible with AD RMS.

For more information, see:

- Message Encryption FAQ to check whether your subscription plan includes Microsoft Purview Information Protection (which includes the Azure Rights Management encryption service).
- Microsoft Purview Information Protection for information about purchasing an eligible subscription.

Manually activating Azure Rights Management

If you disabled the Azure Rights Management service, or if it was not automatically activated for any reason, you can activate it manually.

For instructions, see How to activate or confirm the status of the encryption service.

Configure management of your Azure Rights Management tenant key

This is an optional step. Allowing Microsoft to manage the root key for the Azure Rights Management service is the default setting and recommended best practice for most organizations. If this is the case, you don't need to do anything.

There are many reasons, for example compliance requirements, that may necessitate you generating and managing your own root key, also known as "bring your own key" (BYOK). If this is the case, we recommend that you complete the required steps before setting up Microsoft Purview Message Encryption. See Bring your own key for the Azure Rights Management service root key for more.

Verify Microsoft Purview Message Encryption configuration in Exchange Online PowerShell

You can verify that your Microsoft 365 tenant is properly configured to use Microsoft Purview Message Encryption in Exchange Online PowerShell.

1. Using a work or school account that has sufficient permissions in your organization, such as Compliance Administrator, start a Windows PowerShell session and connect to Exchange Online. For instructions, see Connect to Exchange Online PowerShell.

2. Run the Get-IRMConfiguration cmdlet.

   You should see a value of `$True` for the AzureRMSLicensingEnabled parameter, which indicates that Microsoft Purview Message Encryption is configured in your tenant. If it is not,

use Set-IRMConfiguration to set the value of AzureRMSLicensingEnabled to `$True` to enable Microsoft Purview Message Encryption.

3. Run the Test-IRMConfiguration cmdlet using the following syntax:

```
PowerShell
Test-IRMConfiguration [-Sender <email address> -Recipient <email address>]
```

Example:

```
PowerShell
Test-IRMConfiguration -Sender securityadmin@contoso.com -Recipient
securityadmin@contoso.com
```

- For sender and recipient, use the email address of any user in your Microsoft 365 tenant.

  Your results should be similar to:

  ```
  Console
  Results : Acquiring RMS Templates ...
            - PASS: RMS Templates acquired.  Templates available: Contoso  -
  Confidential View Only, Contoso  - Confidential, Do Not
         Forward.
         Verifying encryption ...
             - PASS: Encryption verified successfully.
         Verifying decryption ...
             - PASS: Decryption verified successfully.
         Verifying IRM is enabled ...
             - PASS: IRM verified successfully.

         OVERALL RESULT: PASS
  ```

- Your organization name will replace *Contoso*.
- The default template names may be different from those displayed above.

4. If the test fails with an error message **Failed to acquire RMS templates**, execute the following commands and run the Test-IRMConfiguration cmdlet to verify that it passes. Connect to the AIPService module to run the cmdlet.

```
PowerShell
```

```
$RMSConfig = Get-AipServiceConfiguration
$LicenseUri = $RMSConfig.LicensingIntranetDistributionPointUrl
Set-IRMConfiguration -LicensingLocation $LicenseUri
Set-IRMConfiguration -InternalLicensingEnabled $true
```

5. Run the Remove-PSSession cmdlet to disconnect from the Rights Management service.

PowerShell

```
Remove-PSSession $session
```

Next steps: Define mail flow rules to use Microsoft Purview Message Encryption

If there are previously configured mail flow rules to encrypt email in your organization, you need to update the existing rules to use Microsoft Purview Message Encryption. For new deployments, you need to create new mail flow rules.

> **Important**
>
> If you do not update existing mail flow rules, your users will continue to receive encrypted mail that uses the previous HTML attachment format, instead of the new seamless experience.

Mail flow rules determine under what conditions email messages should be encrypted, as well as conditions for removing that encryption. When you set an action for a rule, any messages that match the rule conditions are encrypted when they're sent.

For steps on creating mail flow rules for message encryption, see Define mail flow rules to encrypt email messages.

To update existing rules to use Microsoft Purview Message Encryption:

1. In the Microsoft 365 admin center, go to **Admin centers** > **Exchange**.
2. In the Exchange admin center, go to **Mail flow** > **Rules**.
3. For each rule, in **Do the following**:

   - Select **Modify the message security**.
   - Select **Apply Office 365 Message Encryption and rights protection**.
   - Select **Encrypt** from the RMS template list.

- Select **Save**.
- Select **OK**.

Last updated on 05/04/2026

Define mail flow rules to encrypt email messages

> *Microsoft Purview service description*

As an administrator that manages Exchange Online, you can create mail flow rules (also known as transport rules) to help protect email messages you send and receive. You can set up rules to encrypt any outgoing email messages and remove encryption from encrypted messages coming from inside your organization or from replies to encrypted messages sent from your organization. You can use the Exchange admin center (EAC) or Exchange Online PowerShell to create these rules. In addition to overall encryption rules, you can also choose to enable or disable individual message encryption options for end users.

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

You can't encrypt inbound mail from senders outside of your Exchange Online organization. If a mail flow rule is set up to encrypt mail from outside the organization, the inbound mail will be delivered without encryption.

If you recently migrated from Active Directory RMS to the Azure Rights Management service, you'll need to review your existing mail flow rules to ensure that they continue to work in your new environment. Also, to use Microsoft Purview Message Encryption with the Azure Rights Management service, you need to update your existing mail flow rules. Otherwise, your users will continue to receive encrypted mail that uses the previous HTML attachment format instead of the new, seamless experience. If you haven't set up message encryption yet, see Set up Microsoft Purview Message Encryption for information.

For information about the components that make up mail flow rules and how mail flow rules work, see Mail flow rules (transport rules) in Exchange Online.

> **Important**

For hybrid Exchange environments, on-premises users can send and receive encrypted mail using message encryption only if email is routed through Exchange Online. To configure message encryption in a hybrid Exchange environment, you need to first configure hybrid using the Hybrid Configuration wizard and then configure mail to flow from Office 365 to your email server and configure mail to flow from your email server to Office 365. Once you've configured mail to flow through Office 365, then you can configure mail flow rules for message encryption by using the procedures in this article.

Create mail flow rules to encrypt email messages with Microsoft Purview Message Encryption

You can define mail flow rules for triggering message encryption with by using the EAC.

Use the EAC to create a rule for encrypting email messages with Microsoft Purview Message Encryption

Perform the following steps to create a mail flow rule that encrypts email messages using Microsoft Purview Message Encryption.

1. In a web browser, using a work or school account that has been granted sufficient permissions, such as Compliance Administrator, sign in to Microsoft 365.
2. Choose the **Admin** tile.
3. In the Microsoft 365 admin center, choose **Admin centers** > **Exchange**.
4. In the EAC, go to **Mail flow** > **Rules** and select **New** + > **Create a new rule**. For more information about using the EAC, see Exchange admin center in Exchange Online.
5. In **Name**, type a name for the rule, such as Encrypt mail for DrToniRamos@hotmail.com.
6. In **Apply this rule if**, select where the mail originates from inside the Exchange Online organization. Add **The sender is located** > **Inside the organization** for sending mail.
7. In **Apply this rule if**, select a condition, and enter a value if necessary. For example, to encrypt messages going to DrToniRamos@hotmail.com:
    a. In **Apply this rule if**, select **the recipient is**.
    b. Select an existing name from the contact list or type a new email address in the **check names** box.

- To select an existing name, select it from the list and then click **OK**.
- To enter a new name, type an email address in the **check names** box and then select **check names** > **OK**.

8. To add more conditions, choose **More options** and then choose **add condition** and select from the list.

   For example, to apply the rule only if the recipient is outside your organization, select **add condition** and then select **The recipient is external/internal** > **Outside the organization** > **OK**.

9. To enable message encryption, from **Do the following**, select **Modify the message security** and then choose **Apply Office 365 Message Encryption and rights protection**. Select an RMS template from the list, choose **Save**, and then choose **OK**.

The list of templates includes all default templates and options as well as any custom templates you've created for use by Microsoft 365. If the list is empty, ensure that you have set up Microsoft Purview Message Encryption as described in Set up Microsoft Purview Message Encryption. For information about the **Do Not Forward** option, see Do Not Forward option for emails. For information about the **encrypt-only** option, see Encrypt-only option for emails.

You can choose **add action** if you want to specify another action.

Use the EAC to update an existing mail flow rule to use Microsoft Purview Message Encryption

Perform the following steps to update an existing mail flow rule to use Microsoft Purview Message Encryption.

1. In a web browser, using a work or school account that has been granted sufficient administrator permissions, such as Compliance Administrator, sign in to Microsoft 365.
2. Choose the **Admin** tile.
3. In the Microsoft 365 admin center, choose **Admin centers** > **Exchange**.
4. In the EAC, go to **Mail flow** > **Rules**.
5. In the list of mail flow rules, select the rule you want to modify to use with Microsoft Purview Message Encryption and then choose **Edit**.

6. To enable encryption using Microsoft Purview Message Encryption, from **Do the following**, choose **Modify the message security** and then choose **Apply Office 365 Message Encryption and rights protection**. Select an RMS template from the list, choose **Save** and then choose **OK**.

   The list of templates includes all default templates and options as well as any custom templates you've created for use by Microsoft 365. If the list is empty, ensure that you have set up Microsoft Purview Message Encryption as described in Set up Microsoft Purview Message Encryption. For information about the Do Not Forward option, see Do Not Forward option for emails. For information about the encrypt-only option, see Encrypt Only option for emails.

   You can choose **add action** if you want to specify another action.

7. From the **Do the following** list, remove any actions that are assigned to **Modify the message security** > **Apply the previous version of OME**.

8. Choose **Save**.

Create mail flow rules to remove encryption for email messages with Microsoft Purview Message Encryption

You can define mail flow rules to remove message encryption with Microsoft Purview Message Encryption by using the EAC.

Use the EAC to create a rule to remove encryption from email messages with Microsoft Purview Message Encryption

You can remove encryption from messages that was applied by your organization. You can also remove encryption from any encrypted attachments to ensure the whole email message is without any protection.

1. In a web browser, using a work or school account that has been granted sufficient permissions, such as Compliance Administrator, sign in to Microsoft 365.

2. Choose the **Admin** tile.

3. In the Microsoft 365 admin center, choose **Admin centers** > **Exchange**.

4. In the EAC, go to **Mail flow** > **Rules** and select **New +** > **Create a new rule**. For more information about using the EAC, see Exchange admin center in Exchange Online.

5. In **Name**, type a name for the rule, such as `Remove encryption from outgoing mail`.

6. In **Apply this rule if**, select the conditions where encryption should be removed from messages. Add **The sender is located** > **Inside the organization** for sending mail out to any recipients *or* add **The recipient is located** > **Inside the organization** for receiving mail replies from outside the organization.

7. In **Do the following**, select **Modify the message security** > **Remove Office 365 Message Encryption and rights protection applied by the organization**.

8. (Optional) In **Do the following**, select **Modify the message security** > **Remove attachment rights protection protection applied by the organization**.

9. Save the rule.

Create mail flow rules for Office 365 Message Encryption without Microsoft Purview Message Encryption

Office 365 Message Encryption is deprecated as of July 1, 2023. If you haven't yet moved your organization to Microsoft Purview Message Encryption, Microsoft recommends that you make a plan to move as soon as it is reasonable for your organization. For instructions, see Set up Microsoft Purview Message Encryption.

Related content

Encryption in Microsoft 365

Set up Microsoft Purview Message Encryption

Add branding to encrypted messages

Mail flow rules (transport rules) in Exchange Online

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/22/2026

Add your organization's brand to your Microsoft Purview Message Encryption encrypted messages

> *Microsoft Purview service description*

Apply your company branding to customize the look of your organization's email messages and the encryption portal. You need to apply sufficient permissions to your work or school account before you can get started, for example Compliance Administrator. You customize branding in one of two ways, using Exchange Online PowerShell or Microsoft Purview Data Loss Prevention (DLP) policies.

For more information about using Microsoft Purview Data Loss Prevention (DLP) policies to add customized branding to encrypted messages, see these resources.

- Supported action: Exchange for details on this action.
- Design a data loss prevention policy if you're new to DLP and want to learn more about what goes into preparing to create a DLP policy.
- Create and Deploy data loss prevention policies for examples on how to create and deploy a DLP policy.

The following sections describe how to use Exchange Online PowerShell to customize branding.

Use the Get-OMEConfiguration and Set-OMEConfiguration cmdlets in Exchange Online PowerShell to customize these parts of encrypted email messages:

- Introductory text
- Disclaimer text
- URL for Your organization's privacy statement
- Text in the encrypted message portal
- Logo that appears in the email message and encrypted message portal, or whether to use a logo at all
- Background color in the email message and encrypted message portal

You can also revert back to the default look and feel at any time.

If you'd like more control, use Microsoft Purview Advanced Message Encryption to create multiple templates for encrypted emails originating from your organization. Use these templates to control parts of the end-user experience. For example, specify whether recipients can use Google, Yahoo, and Microsoft Accounts to sign in to the encryption portal. Use templates to fulfill several use cases, such as:

- Individual departments, such as Finance, Sales, and so on.
- Different products
- Different geographical regions or countries
- Whether you want to allow emails to be revoked
- Whether you want emails sent to external recipients to expire after a specified number of days.

Once you've created the templates, apply them to encrypted emails sent from your online mailbox by using Exchange mail flow rules. If you have Microsoft Purview Advanced Message Encryption, you can revoke any email that you have branded.

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

Work with branding templates

You can modify several features within a branding template, and modify, but not remove, the default template. If you have Advanced Message Encryption, you can also create, modify, and remove custom templates. Use Exchange Online PowerShell to work with one branding template at a time.

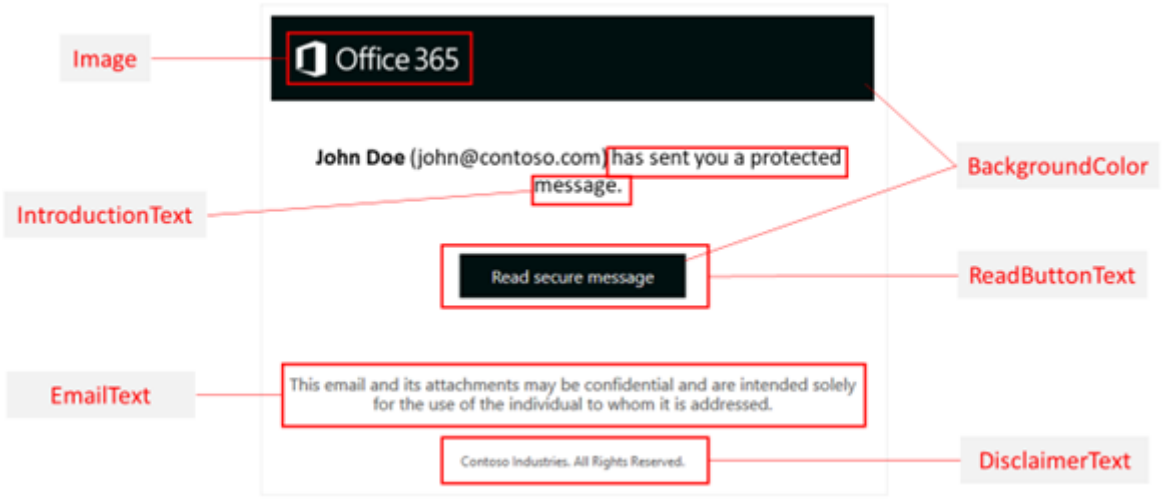
- Set-OMEConfiguration - Modify the default branding template or a custom branding template that you created.
- New-OMEConfiguration - Create a new branding template, Advanced Message Encryption only.
- Remove-OMEConfiguration - Remove a custom branding template, Advanced Message Encryption only. You can't delete the default branding template.

Modify a branding template

Use Exchange Online PowerShell to modify one branding template at a time. If you have Advanced Message Encryption, you can also create, modify, and remove custom templates.

1. Use a work or school account that has sufficient administrator permissions in your organization, such as Compliance Administrator, and connect to Exchange Online PowerShell. For instructions, see Connect to Exchange Online PowerShell.

2. Use the Set-OMEConfiguration cmdlet as described in Set-OMEConfiguration or use the following graphic and table for guidance.



Expand table

To customize this feature of the encryption experience	Use these commands

Background color	`Set-OMEConfiguration -Identity "<ConfigurationName>" -BackgroundColor "<#RRGGBB hexadecimal color code or name value>"`
	
	Example:
	
	`Set-OMEConfiguration -Identity "Branding Template 1" -BackgroundColor "#ffffff"`
	
	For more information about background colors, see Background colors.

Logo	`Set-OMEConfiguration -Identity "<ConfigurationName>" -Image <Byte[]>`
	
	Example:
	
	`Set-OMEConfiguration -Identity "Branding Template 1" -Image ([System.IO.File]::ReadAllBytes('C:\Temp\contosologo.png'))`
	
	Supported file formats: .png, .jpg, .bmp, or .tiff

To customize this feature of the encryption experience	Use these commands
	Optimal size of logo file: less than 40 KB Optimal size of logo image: 170x70 pixels. If your image exceeds these dimensions, the service resizes your logo for display in the portal. The service doesn't modify the graphic file itself. For best results, use the optimal size.

Text next to the sender's name and email address	`Set-OMEConfiguration -Identity "<ConfigurationName>" -IntroductionText "<String up to 1024 characters>"`
	
	Example:
	
	`Set-OMEConfiguration -Identity "Branding Template 1" -IntroductionText "has sent you a secure message."`

Text that appears on the "Read Message" button	`Set-OMEConfiguration -Identity "<ConfigurationName>" -ReadButtonText "<String up to 1024 characters>"`
	
	Example:
	
	`Set-OMEConfiguration -Identity "Message encryption configuration" -ReadButtonText "Read Secure Message."`

Text that appears below the "Read Message" button	`Set-OMEConfiguration -Identity "<ConfigurationName>" -EmailText "<String up to 1024 characters>"`
	
	Example:
	
	`Set-OMEConfiguration -Identity "Message encryption configuration" -EmailText "Encrypted message from ContosoPharma secure messaging system."`

URL for the Privacy Statement link	`Set-OMEConfiguration -Identity "<ConfigurationName>" -PrivacyStatementURL "<URL>"`
	
	Example:
	
	`Set-OMEConfiguration -Identity "Branding Template 1" -PrivacyStatementURL "https://contoso.com/privacystatement.html"`

Disclaimer statement in the email that contains the encrypted message	`Set-OMEConfiguration -Identity "<OMEConfigurationName>" -DisclaimerText "<Disclaimer statement. String of up to 1024 characters.>"`
	
	Example:
	
	`Set-OMEConfiguration -Identity "Branding Template 1" -DisclaimerText "This message is confidential for the use of the addressee only."`

Text that appears at the top of the encrypted	`Set-OMEConfiguration -Identity "<OMEConfigurationName>" -PortalText "<Text for your portal. String of up to 128 characters.>"`

To customize this feature of the encryption experience	Use these commands
mail viewing portal	**Example:** `Set-OMEConfiguration -Identity "Message encryption cfonfiguration" -PortalText "ContosoPharma secure email portal."`

To enable or disable authentication with a one-time pass code for this custom template	`Set-OMEConfiguration -Identity "<OMEConfigurationName>" -OTPEnabled <$true\|$false>`
	
	Examples:
	To enable one-time passcodes for this custom template
	
	`Set-OMEConfiguration -Identity "Branding Template 1" -OTPEnabled $true`
	
	To disable one-time passcodes for this custom template
	
	`Set-OMEConfiguration -Identity "Branding Template 1" -OTPEnabled $false`

To enable or disable authentication with Microsoft, Google, or Yahoo identities for this custom template	`Set-OMEConfiguration -Identity "<OMEConfigurationName>" -SocialIdSignIn <$true\|$false>`
	
	Examples:
	To enable social IDs for this custom template
	
	`Set-OMEConfiguration -Identity "Branding Template 1" -SocialIdSignIn $true`
	
	To disable social IDs for this custom template
	
	`Set-OMEConfiguration -Identity "Branding Template 1" -SocialIdSignIn $false`

Create an encrypted message branding template (Advanced Message Encryption)

If you have Microsoft Purview Advanced Message Encryption, you can create custom branding templates for your organization by using the New-OMEConfiguration cmdlet. Once you've created the template, you modify the template by using the Set-OMEConfiguration cmdlet as described in Modify a branding template. You can create multiple templates.

To create a new custom branding template:

1. Use a work or school account that has sufficient administrator permissions in your organization, such as Compliance Administrator, and connect to Exchange Online PowerShell. For instructions, see Connect to Exchange Online PowerShell.

2. Use the New-OMEConfiguration cmdlet to create a new template.

PowerShell

```
New-OMEConfiguration -Identity "<OMEConfigurationName>"
```

For example,

PowerShell

```
New-OMEConfiguration -Identity "Custom branding template"
```

Return the default branding template to its original values

To remove all modifications from the default template, including brand customizations, and so on, complete these steps:

1. Use a work or school account that has sufficient administrator permissions in your organization, such as Compliance Administrator, and connect to Exchange Online PowerShell. For instructions, see Connect to Exchange Online PowerShell.

2. Use the **Set-OMEConfiguration** cmdlet as described in Set-OMEConfiguration. To remove your organization's branded customizations from the DisclaimerText, EmailText, and PortalText values, set the value to an empty string, `""`. For all image values, such as Logo, set the value to `"$null"`.

   The following table describes the encryption customization option defaults.

Expand table

To revert this feature of the encryption experience back to the default text and image	Use these commands

Default text that comes with encrypted email messages. The default text appears above the instructions for viewing encrypted messages	`Set-OMEConfiguration -Identity "<OMEConfigurationName>" -EmailText "<empty string>"`
	
	Example:
	
	`Set-OMEConfiguration -Identity "Message encryption configuration" -EmailText ""`

Disclaimer statement in the email that contains the encrypted message	`Set-OMEConfiguration -Identity "<OMEConfigurationName>" DisclaimerText "<empty string>"`

To revert this feature of the encryption experience back to the default text and image	Use these commands
	Example: `Set-OMEConfiguration -Identity "Message encryption configuration" -DisclaimerText ""`

Text that appears at the top of the encrypted mail viewing portal	`Set-OMEConfiguration -Identity "<OMEConfigurationName>" -PortalText "<empty string>"`
	
	Example reverting back to default:
	
	`Set-OMEConfiguration -Identity "Message encryption configuration" -PortalText ""`

Logo	`Set-OMEConfiguration -Identity "<OMEConfigurationName>" -Image <"$null">`
	
	Example reverting back to default:
	
	`Set-OMEConfiguration -Identity "Message encryption configuration" -Image $null`

Background color	`Set-OMEConfiguration -Identity "<ConfigurationName>" -BackgroundColor "$null">`
	
	Example reverting back to default:
	
	`Set-OMEConfiguration -Identity "Message encryption configuration" -BackgroundColor $null`

Remove a custom branding template (Advanced Message Encryption)

You can only remove or delete branding templates that you've made. You can't remove the default branding template.

To remove a custom branding template:

1. Use a work or school account that has sufficient administrator permissions in your organization, such as Compliance Administrator, and connect to Exchange Online PowerShell. For instructions, see Connect to Exchange Online PowerShell.

2. Use the **Remove-OMEConfiguration** cmdlet as follows:

> ⚠ **Warning**
>
> The following command permanently removes the specified custom branding template. This action can't be undone, and you can't remove the default branding template.

PowerShell

```
Remove-OMEConfiguration -Identity "<OMEConfigurationName>"
```

For example,

PowerShell

```
Remove-OMEConfiguration -Identity "Branding template 1"
```

For more information, see Remove-OMEConfiguration.

Create an Exchange mail flow rule that applies your custom branding to encrypted emails sent from your online organization to external recipients

> ⓘ **Important**
>
> Third-party applications that scan and modify mail can prevent branding from being applied correctly.

After you've either modified the default template or created new branding templates, you can create Exchange mail flow rules to apply your custom branding based on certain conditions. Most importantly, the email must be encrypted. Such a rule applies custom branding to mail sent from your online mailbox in the following scenarios:

- If the email was manually encrypted by the end user using Outlook or Outlook on the web, formerly Outlook Web App
- If the email was automatically encrypted by an Exchange mail flow rule or Microsoft Purview Data Loss Prevention policy

To ensure Microsoft Purview Message Encryption applies your custom branding, set up a mail flow rule to encrypt your messages. The priority of the encryption rule should be higher than the branding rule so that the encryption rule is processed first. By default, if you create the encryption rule before the branding rule, then the encryption rule has a higher priority. For information, see Define mail flow rules to encrypt email messages. For information on setting the priority of a mail flow rule, see Manage mail flow rules.

1. In a web browser, using a work or school account that has sufficient permissions in your organization, such as Compliance Administrator, sign in to Microsoft 365.

2. Choose the **Admin** tile.

3. In the Microsoft 365 admin center, choose **Admin centers** > **Exchange**.

4. In the EAC, go to **Mail flow** > **Rules** and select **New +** > **Create a new rule**. For more information about using the EAC, see Exchange admin center in Exchange Online.

5. In **Name**, type a name for the rule, such as **Branding for sales department**.

6. In **Apply this rule if**, select the condition **The sender is located inside the organization** and other conditions you want from the list of available conditions. For example, you might want to apply a particular branding template to:

   - All encrypted emails sent from members of the finance department
   - Encrypted emails sent with a certain keyword such as "External" or "Partner"
   - Encrypted emails sent to a particular domain

7. If you've already defined a mail flow rule to apply encryption, skip this step. Otherwise, to configure the mail flow rule to apply encryption, from **Do the following**, select **Modify the message security**, and then select **Apply Office 365 Message Encryption and rights protection**. Select a Rights Management Service (RMS) template from the list and then select **add action**.

   The list of templates includes default templates and options and any custom templates you create. If the list is empty, ensure that you have set up Microsoft Purview Message Encryption. For instructions, see Set up Microsoft Purview Message Encryption. For information about the **Do Not Forward** option, see the Do Not Forward option for emails. For information about the **Encrypt Only** option, see Encrypt Only option for emails.

8. From **Do the following**, select **Modify the message security** > **Apply custom branding to OME messages**. Next, from the drop-down, select a branding template.

Select **add action** if you want to specify another action, or select **Save**, and then select **OK**.

Background color reference

The color names that you can use for the background color are limited. Instead of a color name, you can use a hex code value (`#RRGGBB`). You can use a hex code value that corresponds to a color name, or you can use a custom hex code value. Be sure to enclose the hex code value in quotation marks (for example, `"#f0f8ff"`).

The available background color names and their corresponding hex code values are described in the following table.

Expand table

Color name	Color code
`aliceblue`	#f0f8ff
`antiquewhite`	#faebd7
`aqua`	#00ffff
`aquamarine`	#7fffd4
`azure`	#f0ffff
`beige`	#f5f5dc
`bisque`	#ffe4c4
`black`	#000000
`blanchedalmond`	#ffebcd
`blue`	#0000ff
`blueviolet`	#8a2be2
`brown`	#a52a2a
`burlywood`	#deb887
`cadetblue`	#5f9ea0
`chartreuse`	#7fff00
`chocolate`	#d2691e
`coral`	#ff7f50

Color name	Color code
`cornflowerblue`	#6495ed
`cornsilk`	#fff8dc
`crimson`	#dc143c
`cyan`	#00ffff
`darkblue`	#00008b
`darkcyan`	#008b8b
`darkgoldenrod`	#b8860b
`darkgray`	#a9a9a9
`darkgreen`	#006400
`darkkhaki`	#bdb76b
`darkmagenta`	#8b008b
`darkolivegreen`	#556b2f
`darkorange`	#ff8c00
`darkorchid`	#9932cc
`darkred`	#8b0000
`darksalmon`	#e9967a
`darkseagreen`	#8fbc8f
`darkslateblue`	#483d8b
`darkslategray`	#2f4f4f
`darkturquoise`	#00ced1
`darkviolet`	#9400d3
`deeppink`	#ff1493
`deepskyblue`	#00bfff
`dimgray`	#696969
`dodgerblue`	#1e90ff
`firebrick`	#b22222

Color name	Color code
`floralwhite`	#fffaf0
`forestgreen`	#228b22
`fuchsia`	#ff00ff
`gainsboro`	#dcdcdc
`ghostwhite`	#f8f8ff
`gold`	#ffd700
`goldenrod`	#daa520
`gray`	#808080
`green`	#008000
`greenyellow`	#adff2f
`honeydew`	#f0fff0
`hotpink`	#ff69b4
`indianred`	#cd5c5c
`indigo`	#4b0082
`ivory`	#fffff0
`khaki`	#f0e68c
`lavender`	#e6e6fa
`lavenderblush`	#fff0f5
`lawngreen`	#7cfc00
`lemonchiffon`	#fffacd
`lightblue`	#add8e6
`lightcoral`	#f08080
`lightcyan`	#e0ffff
`lightgoldenrodyellow`	#fafad2
`lightgray`	#d3d3d3
`lightgrey`	#d3d3d3

Color name	Color code
`lightgreen`	#90ee90
`lightpink`	#ffb6c1
`lightsalmon`	#ffa07a
`lightseagreen`	#20b2aa
`lightskyblue`	#87cefa
`lightslategray`	#778899
`lightsteelblue`	#b0c4de
`lightyellow`	#ffffe0
`lime`	#00ff00
`limegreen`	#32cd32
`linen`	#faf0e6
`magenta`	#ff00ff
`maroon`	#800000
`mediumaquamarine`	#66cdaa
`mediumblue`	#0000cd
`mediumorchid`	#ba55d3
`mediumpurple`	#9370db
`mediumseagreen`	#3cb371
`mediumslateblue`	#7b68ee
`mediumspringgreen`	#00fa9a
`mediumturquoise`	#48d1cc
`mediumvioletred`	#c71585
`midnightblue`	#191970
`mintcream`	#f5fffa
`mistyrose`	#ffe4e1
`moccasin`	#ffe4b5

Color name	Color code
`navajowhite`	#ffdead
`navy`	#000080
`oldlace`	#fdf5e6
`olive`	#808000
`olivedrab`	#6b8e23
`orange`	#ffa500
`orangered`	#ff4500
`orchid`	#da70d6
`palegoldenrod`	#eee8aa
`palegreen`	#98fb98
`paleturquoise`	#afeeee
`palevioletred`	#db7093
`papayawhip`	#ffefd5
`peachpuff`	#ffdab9
`peru`	#cd853f
`pink`	#ffc0cb
`plum`	#dda0dd
`powderblue`	#b0e0e6
`purple`	#800080
`red`	#ff0000
`rosybrown`	#bc8f8f
`royalblue`	#4169e1
`saddlebrown`	#8b4513
`salmon`	#fa8072
`sandybrown`	#f4a460
`seagreen`	#00ff00

Color name	Color code
`seashell`	#fff5ee
`sienna`	#a0522d
`silver`	#c0c0c0
`skyblue`	#87ceeb
`slateblue`	#6a5acd
`slategray`	#708090
`snow`	#fffafa
`springgreen`	#00ff7f
`steelblue`	#4682b4
`tan`	#d2b48c
`teal`	#008080
`thistle`	#d8bfd8
`tomato`	#ff6347
`turquoise`	#40e0d0
`violet`	#ee82ee
`wheat`	#f5deb3
`white`	#ffffff
`whitesmoke`	#f5f5f5
`yellow`	#ffff00
`yellowgreen`	#9acd32

ⓘ**Note:** The author created this article with assistance from AI. Learn more

Last updated on 06/22/2026

Create a sensitive information type policy for your organization using Microsoft Purview Message Encryption

> *Microsoft Purview service description*

You can use either Exchange mail flow rules or Microsoft Purview Data Loss Prevention (DLP) to create a sensitive information type policy with Microsoft Purview Message Encryption. To create an Exchange mail flow rule, you can use either the Exchange admin center (EAC) or Exchange PowerShell.

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview **roles and permissions**.

To create the policy by using mail flow rules in the EAC

Sign in to the Exchange admin center and go to **Mail flow** > **Rules**. On the Rules page, create a rule that applies Message Encryption. You can create a rule based on conditions such as the presence of certain keywords or sensitive information types in the message or attachment.

To create the policy by using mail flow rules in Exchange PowerShell

Use a work or school account that has sufficient permissions in your organization, such as Compliance Administrator, connect to Exchange Online PowerShell. For instructions, see Connect to Exchange Online PowerShell. Use the `Set-IRMConfiguration` and `New-TransportRule` cmdlets to create the policy.

Example mail flow rule created with Exchange PowerShell

Run the following commands in PowerShell to create an Exchange mail flow rule that automatically encrypts emails sent outside your organization with the encrypt-only option if the emails or their attachments contain the following sensitive information types:

- ABA routing number
- Credit card number
- Drug Enforcement Agency (DEA) number
- U.S. or U.K. passport number
- U.S. bank account number
- U.S. Individual Taxpayer Identification Number (ITIN)
- U.S. Social Security Number (SSN)

PowerShell

```
Set-IRMConfiguration -DecryptAttachmentForEncryptOnly $true
New-TransportRule -Name "Encrypt outbound sensitive emails (out of box rule)" -
SentToScope  NotInOrganization  -ApplyRightsProtectionTemplate "Encrypt" -
MessageContainsDataClassifications @(@{Name="ABA Routing Number";
minCount="1"},@{Name="Credit Card Number"; minCount="1"},@{Name="Drug Enforcement
Agency (DEA) Number"; minCount="1"},@{Name="U.S. / U.K. Passport Number";
minCount="1"},@{Name="U.S. Bank Account Number"; minCount="1"},@{Name="U.S. Individual
Taxpayer Identification Number (ITIN)"; minCount="1"},@{Name="U.S. Social Security
Number (SSN)"; minCount="1"}) -SenderNotificationType "NotifyOnly"
```

For cmdlet syntax and parameter details, see Set-IRMConfiguration and New-TransportRule.

How recipients access attachments

After Microsoft encrypts a message, recipients have unrestricted access to attachments when they access and open their encrypted email.

To prepare for automatic encryption of outbound sensitive emails

You may want to update any applicable end-user documentation and training materials to prepare people in your organization for the automatic encryption of outbound messages containing sensitive information. Share these message encryption resources with your users as appropriate:

- Send, view, and reply to encrypted messages in Outlook for PC
- Microsoft 365 Essentials Video: Message Encryption

View these changes in the audit log

Microsoft 365 audits the creation of the mail flow rule and makes the audit record available to administrators. The operation is 'New-TransportRule' and a snippet of a sample audit entry from the Audit Log Search in the Microsoft Purview portal is as follows:

text


```
*{"CreationTime":"2018-11-28T23:35:01","Id":"a1b2c3d4-daa0-4c4f-a019-
03a1234a1b0c","Operation":"New-TransportRule","OrganizationId":"123456-221d-12345
","RecordType":1,"ResultStatus":"True","UserKey":"Microsoft
Operator","UserType":3,"Version":1,"Workload":"Exchange","ClientIP":"123.456.147.68:17
584","ObjectId":"","UserId":"Microsoft
Operator","ExternalAccess":true,"OrganizationName":"contoso.onmicrosoft.com","Originat
ingServer":"CY4PR13MBXXXX (15.20.1382.008)","Parameters":
{"Name":"Organization","Value":"123456-221d-12346"
{"Name":"ApplyRightsProtectionTemplate","Value":"Encrypt"},
{"Name":"Name","Value":"Encrypt outbound sensitive emails (out of box rule)"},
{"Name":"MessageContainsDataClassifications"...etc.*
```


To disable or customize the sensitive information types policy

Once you've created the Exchange mail flow rule, you can disable or edit the rule by going to **Mail flow** > **Rules** in the Exchange admin center and disabling the rule "*Encrypt outbound sensitive emails (out of box rule)*".

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Manage message encryption

> *Microsoft Purview service description*

Once you've finished setting up Purview Message Encryption, you can customize the configuration of your deployment in several ways. For example, you can configure whether to enable one-time pass codes, display the **Encrypt** button in Outlook on the web, and more. The tasks in this article describe how.

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

Manage whether Google, Yahoo, and Microsoft Account recipients can use these accounts to sign in to the encrypted message portal

When you set up the message encryption, users in your organization can send messages to recipients that are outside of your organization. If the recipient uses a *social ID* such as a Google account, Yahoo account, or Microsoft account, the recipient can sign in to the encrypted message portal with a social ID. If you want, you can choose not to allow recipients to use social IDs to sign in to the encrypted message portal.

To manage whether recipients can use social IDs to sign in to the encrypted message portal

1. Connect to Exchange Online PowerShell.
2. Run the Set-OMEConfiguration cmdlet with the SocialIdSignIn parameter as follows:

```
PowerShell
Set-OMEConfiguration -Identity <"OMEConfigurationIdParameter"> -SocialIdSignIn
<$true|$false>
```

For example, to disable social IDs:

```
PowerShell
Set-OMEConfiguration -Identity "OME Configuration" -SocialIdSignIn $false
```

To enable social IDs:

```
PowerShell
Set-OMEConfiguration -Identity "OME Configuration" -SocialIdSignIn $true
```

Manage the use of one-time pass codes for the encrypted message portal

If the recipient of a message encrypted by message encryption doesn't use Outlook, regardless of the account used by the recipient, the recipient receives a limited-time web-view link that lets them read the message. This link includes a one-time pass code. As an administrator, you can decide if recipients can use one-time pass codes to sign in to the encrypted message portal.

To manage whether OME generates one-time pass codes

Perform the following steps to manage whether OME generates one-time pass codes.

1. Use a work or school account that has sufficient administrator permissions in your organization, such as Compliance Administrator, and connect to Exchange Online PowerShell. For instructions, see Connect to Exchange Online PowerShell.

2. Run the Set-OMEConfiguration cmdlet with the OTPEnabled parameter:

```
PowerShell
Set-OMEConfiguration -Identity <"OMEConfigurationIdParameter"> -OTPEnabled
<$true|$false>
```

For example, to disable one-time pass codes:

```
PowerShell
Set-OMEConfiguration -Identity "OME Configuration" -OTPEnabled $false
```

To enable one-time pass codes:

PowerShell

```
Set-OMEConfiguration -Identity "OME Configuration" -OTPEnabled $true
```

Manage the display of the Encrypt button in Outlook on the web

As an administrator, you can manage whether to display the **Encrypt** button in Outlook on the web to end users.

To manage whether the Encrypt button appears in Outlook on the web

Perform the following steps to manage whether the **Encrypt** button appears in Outlook on the web.

1. Use a work or school account that has sufficient permissions in your organization, such as Compliance Administrator, and connect to Exchange Online PowerShell. For instructions, see Connect to Exchange Online PowerShell.

2. Run the Set-IRMConfiguration cmdlet with the -SimplifiedClientAccessEnabled parameter:

   PowerShell

   ```
   Set-IRMConfiguration -SimplifiedClientAccessEnabled <$true|$false>
   ```

   For example, to disable the **Encrypt** button:

   PowerShell

   ```
   Set-IRMConfiguration -SimplifiedClientAccessEnabled $false
   ```

   To enable the **Encrypt** button:

   PowerShell

   ```
   Set-IRMConfiguration -SimplifiedClientAccessEnabled $true
   ```

Enable service-side decryption of email messages for iOS mail app users

The iOS mail app can't decrypt messages protected with message encryption. As a Microsoft 365 administrator, you can apply service-side decryption for messages delivered to the iOS mail app. When you choose to do use service-side decryption, the service sends a decrypted copy of the message to the iOS device. The client device stores a decrypted copy of the message. The message also retains information about usage rights even though the iOS mail app doesn't apply client-side usage rights to the user. The user can copy or print the message even if they didn't originally have the rights to do so. However, if the user attempts to complete an action that requires the Microsoft 365 mail server, such as forwarding the message, the server won't permit the action if the user didn't originally have the usage right to do so. However, end users can work around "Do Not Forward" usage restriction by forwarding the message from a different account within the iOS mail app. Regardless of whether you set up service-side decryption of mail, attachments to encrypted and rights protected mail can't be viewed in the iOS mail app.

If you choose not to allow decrypted messages to be sent to iOS mail app users, users receive a message that states that they don't have the rights to view the message. By default, service-side decryption of email messages isn't enabled.

For more information, and for a view of the client experience, see View encrypted messages on your iPhone or iPad.

To manage whether iOS mail app users can view messages protected by message encryption

Perform the following steps to manage whether iOS mail app users can view messages protected by message encryption.

1. Use a work or school account that has sufficient permissions in your organization, such as Compliance Administrator, and connect to Exchange Online PowerShell. For instructions, see Connect to Exchange Online PowerShell.

2. Run the Set-ActiveSyncOrganizations cmdlet with the AllowRMSSupportForUnenlightenedApps parameter:

   PowerShell

   ```
   Set-ActiveSyncOrganizationSettings -AllowRMSSupportForUnenlightenedApps
   <$true|$false>
   ```

   For example, to configure the service to decrypt messages before they're sent to unenlightened apps like the iOS mail app:

PowerShell

```
Set-ActiveSyncOrganizationSettings -AllowRMSSupportForUnenlightenedApps $true
```

Or, to configure the service not to send decrypted messages to unenlightened apps:

PowerShell

```
Set-ActiveSyncOrganizationSettings -AllowRMSSupportForUnenlightenedApps $false
```

> **Note**
>
> Individual mailbox policies (OWA/ActiveSync) override the AllowRMSSupportForUnenlightenedApps setting (i.e. if -IRMEnabled is set to False within the respective OWA Mailbox policy or ActiveSync Mailbox policy, then the AllowRMSSupportForUnenlightenedApps setting doesn't apply).

Enable service-side decryption of email attachments for web browser mail clients

Normally, when you use message encryption, attachments are automatically encrypted. As an administrator, you can apply service-side decryption for email attachments that users download from a web browser.

When you use service-side decryption, the service sends a decrypted copy of the file to the device. The message is still encrypted. The email attachment also keeps information about usage rights even though the browser doesn't apply client-side usage rights to the user. The user can copy or print the email attachment even if they didn't originally have the rights to do so. However, if the user tries to complete an action that requires the Microsoft 365 mail server, such as forwarding the attachment, the server won't permit the action if the user didn't originally have the usage right to do so.

Regardless of whether you set up service-side decryption of attachments, users can't view any attachments to encrypted and rights protected mail in the iOS mail app.

If you choose not to allow decrypted email attachments, which is the default, users receive a message that states that they don't have the rights to view the attachment.

For more information about how Microsoft 365 implements encryption for emails and email attachments with the Encrypt-Only option, see Encrypt-Only option for emails.

To manage whether email attachments are decrypted on download from a web browser

Perform the following steps to manage whether email attachments are decrypted when downloaded from a web browser.

1. Use a work or school account that has sufficient permissions in your organization, such as Compliance Administrator, and connect to Exchange Online PowerShell. For instructions, see Connect to Exchange Online PowerShell.

2. Run the Set-IRMConfiguration cmdlet with the DecryptAttachmentForEncryptOnly parameter:

   PowerShell

   ```
   Set-IRMConfiguration -DecryptAttachmentForEncryptOnly <$true|$false>
   ```

   For example, to configure the service to decrypt email attachments when a user downloads them from a web browser:

   PowerShell

   ```
   Set-IRMConfiguration -DecryptAttachmentForEncryptOnly $true
   ```

   To configure the service to leave encrypted email attachments as they are upon download:

   PowerShell

   ```
   Set-IRMConfiguration -DecryptAttachmentForEncryptOnly $false
   ```

Ensure all external recipients use the encrypted message portal to read encrypted mail

You can use custom branding templates to force recipients to receive a wrapper mail that directs them to read encrypted email in the encrypted message portal instead of using Outlook or Outlook on the web. You might want to do force this experience if you use want greater control over how recipients use the mail they receive. For example, if external recipients view email in the web portal, you can set an expiration date for the email, and you can revoke the email. These features are only supported through the encrypted message portal. You can use the Encrypt option and the Do Not Forward option when creating the mail flow rules.

Use a custom template to force all external recipients to use the encrypted message portal and for encrypted email

Perform the following steps to create a custom template that forces all external recipients to use the encrypted message portal for encrypted email.

1. Use a work or school account that has sufficient permissions in your organization, such as Compliance Administrator, and connect to Exchange Online PowerShell. For instructions, see Connect to Exchange Online PowerShell.

2. Run the New-TransportRule cmdlet:

   PowerShell

   ```
   New-TransportRule -name "<mail flow rule name>" -FromScope "InOrganization" -
   ApplyRightsProtectionTemplate "<option name>" -
   ApplyRightsProtectionCustomizationTemplate "<template name>"
   ```

   where:

   - `mail flow rule name` is the name you want to use for the new mail flow rule.
   - `option name` is either `Encrypt` or `Do Not Forward`.
   - `template name` is the name you gave the custom branding template, for example `OME Configuration`.

   To encrypt all external email with the "OME Configuration" template and apply the Encrypt-Only option:

   PowerShell

   ```
   New-TransportRule -name "<All outgoing mail>" -FromScope "InOrganization" -
   ApplyRightsProtectionTemplate "Encrypt" -
   ApplyRightsProtectionCustomizationTemplate "OME Configuration"
   ```

   To encrypt all external email with the "OME Configuration" template and apply the Do Not Forward option:

   PowerShell

   ```
   New-TransportRule -name "<All outgoing mail>" -FromScope "InOrganization" -
   ApplyRightsProtectionTemplate "Do Not Forward" -
   ApplyRightsProtectionCustomizationTemplate "OME Configuration"
   ```

Customize the appearance of email messages and the encrypted message portal

For detailed information about how you can customize Microsoft Purview Message Encryption for your organization, see Add your organization's brand to your encrypted messages. To track and revoke encrypted messages, you must add your custom branding to the encrypted message portal.

Disable Microsoft Purview Message Encryption

We hope it doesn't come to it, but if you need to, disabling Microsoft Purview Message Encryption is straightforward. First, remove any mail flow rules you've created that use Microsoft Purview Message Encryption. For information about removing mail flow rules, see Manage mail flow rules. Then, complete these steps in Exchange Online PowerShell.

To disable Microsoft Purview Message Encryption

Perform the following steps to disable Microsoft Purview Message Encryption.

1. Using a work or school account that has sufficient permissions in your organization, such as Compliance Administrator, connect to Exchange Online PowerShell. For instructions, see Connect to Exchange Online PowerShell.

2. If you enabled the **Encrypt** button in Outlook on the web, disable it by running the Set-IRMConfiguration cmdlet with the SimplifiedClientAccessEnabled parameter. Otherwise, skip this step.

   PowerShell

   ```
   Set-IRMConfiguration -SimplifiedClientAccessEnabled $false
   ```

3. Disable the Microsoft Purview Message Encryption by running the Set-IRMConfiguration cmdlet with the AzureRMSLicensingEnabled parameter set to false:

   PowerShell

   ```
   Set-IRMConfiguration -AzureRMSLicensingEnabled $false
   ```

ⓘ**Note:** The author created this article with assistance from AI. Learn more

Last updated on 06/26/2026

Advanced Message Encryption

> *Microsoft Purview service description*

Advanced Message Encryption helps customers meet compliance obligations that require more flexible controls over external recipients and their access to encrypted emails. With Advanced Message Encryption in Office 365, you can control sensitive emails shared outside the organization with automatic policies and track those activities through the encrypted message portal access logs. You configure these policies to identify sensitive information types such as PII, Financial, or Health IDs, or you can use keywords to enhance protection. Once you've configured the policies, you pair policies with custom branded email templates and then add an expiration date for extra control of emails that fit the policy. Also, admins can further control encrypted emails accessed externally through a secure web portal by revoking access to the mail at any time.

You can only revoke and set an expiration date for emails sent to external recipients.

Get started with Microsoft Purview Advanced Message Encryption

The following articles describe how you set up and use Advanced Message Encryption.

If you do not have Office 365 Message Encryption set up already, see Set up new Office 365 Message Encryption capabilities.

With Advanced Message Encryption, you're not limited to a single branding template. Instead, you can create and use multiple branding templates. Adding custom branding also lets you enable tracking a revocation of encrypted messages. For information, see Add your organization's brand to your encrypted messages. When you use custom branding, external recipients receive a notification email that contains a link to the OME portal. The mail flow rule determines which branding template the notification email and OME Portal use. This way, your secure content isn't sent outside your organization.

You can only revoke messages and apply expiration dates to messages that users receive through the portal. In other words, email that has a custom branding template applied. For more information and an example, see the guidance in Ensure all external recipients use the encrypted message portal to read encrypted mail.

Set an expiration date for email encrypted by Microsoft Purview Advanced Message Encryption. Control sensitive emails shared outside the organization with automatic policies that enhance protection by expiring access through a secure web portal to encrypted emails.

Revoke email encrypted by Microsoft Purview Advanced Message Encryption. Control sensitive emails shared outside the organization and enhance protection by revoking access through a secure web portal to encrypted emails.

Encrypted message portal activity log by Microsoft Purview Advanced Message Encryption. Monitor sensitive emails shared outside the organization in the encrypted message portal.

Last updated on 05/14/2026

Encrypted message portal activity log by Microsoft Purview Advanced Message Encryption

You can access logs for encrypted messages through the encrypted message portal. With these logs, your organization can determine when external recipients read and forward messages. To ensure availability of logs for external recipients, apply a custom branding template to protected emails that your organization sends to external recipients. This template enforces a portal experience. For more information, see Add your organization's brand to your encrypted messages.

Enable message access audit logs in PowerShell

You can enable the access log by using Exchange Online PowerShell. The *EnablePortalTrackingLogs* parameter of the **Set-IrmConfiguration** cmdlet specifies whether to enable audit logs for accessing the encrypted message portal. Valid values are:

- $true: Turns on the audit feature.
- $false: Turns off the audit feature.

Example:

PowerShell

```
Set-IrmConfiguration -EnablePortalTrackingLogs $true
```

For more information, see Set-IRMConfiguration (ExchangePowerShell).

Message access audit information

The access log contains entries for messages sent through the encrypted message portal for the following types of activity:

- External user sign-in timestamp and authentication method
- External user reading messages or attachments
- Attachment download
- Mail replies and forwards

For more information on the message access log schema, see Audit log activities.

Search for events in the message access logs

Complete the following steps to get started with search:

1. Sign in to the Microsoft Purview portal.
2. Select the **Audit** solution card. If the **Audit** solution card isn't displayed, select **View all solutions** and then select **Audit** from the **Core** section.
3. Under **Search**, select the drop-down for **Activities** and type encrypted message portal activities.
4. Under encrypted message portal activities, select the event types to use in the search. Set the date range for the search (default is the previous week), you can also optionally add a particular user in your organization for the search. When ready, select **Search**.
5. Select an event from the list to view the audit properties.

Next steps

- Search the audit log: Search for broader activity related to the senders or recipients identified in encrypted message logs.
- Export, configure, and view audit log records: Export your encrypted message activity results for compliance reporting.

ⓘ **Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 04/14/2026

Set an expiration date for email encrypted by Microsoft Purview Advanced Message Encryption

> *Microsoft Purview service description*

You can use message expiration on emails that your users send to external recipients who use the OME Portal to access encrypted emails. You force recipients to use the OME portal to view and reply to encrypted emails sent by your organization by using a custom branded template that specifies an expiration date in PowerShell.

When you apply your company brand to customize the look of your organization's email messages, you can also specify an expiration for these email messages. With Microsoft Purview Advanced Message Encryption, you can create multiple templates for encrypted emails that originate from your organization. Using a template, you can control how long recipients have access to mail sent by your users.

When an end user receives mail that has an expiration date set, the user sees the expiration date in the wrapper email. If a user tries to open an expired mail, an error appears in the OME portal.

You can only set expiration dates for emails to external recipients.

With Microsoft Purview Advanced Message Encryption, anytime you apply custom branding, Microsoft 365 applies the wrapper to email that fits the mail flow rule to which you apply the template. You can only use expiration if you use custom branding.

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

Create a custom branding template to force mail expiration by using PowerShell

To create a custom branding template that forces mail expiration, complete the following steps.

1. Using a work or school account that has sufficient permissions in your organization, such as Compliance Administrator, start a Windows PowerShell session and connect to Exchange Online. For instructions, see Connect to Exchange Online PowerShell.

2. Run the New-OMEConfiguration cmdlet.

   PowerShell

   ```
   New-OMEConfiguration -Identity "Expire in 7 days" -ExternalMailExpiryInDays 7
   ```

Where:

- `Identity` is the name of the custom template.
- `ExternalMailExpiryInDays` identifies the number of days that recipients can keep mail before it expires. You can use any value between 1–730 days.

Related content

- Advanced Message Encryption
- Revoke email encrypted by Microsoft Purview Advanced Message Encryption
- Message policy and compliance service description

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Revoke email encrypted by Advanced Message Encryption

> *Microsoft Purview service description*

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

This article is part of a larger series of articles about Microsoft Purview Message Encryption.

If a message was encrypted using Microsoft Purview Advanced Message Encryption, and you are a Microsoft 365 admin or you are the sender of the message, you can revoke the message under certain conditions. Admins revoke messages using PowerShell. As a sender, you revoke a message that you sent directly from Outlook on the web. This article describes the circumstances under which revocation is possible and how to do it.

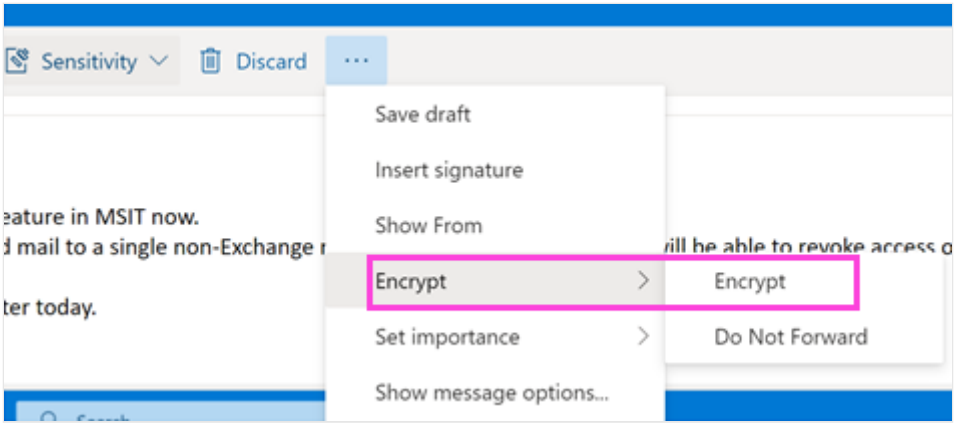
> **Note**
>
> To guarantee that the ability to track and revoke encrypted messages is available, you must add a custom branding template. See Add your organization's brand to your encrypted messages

Encrypted emails that you can revoke

Admins and message senders can revoke encrypted emails if the recipient received a link-based, branded encrypted email. If the recipient received a native inline experience in a supported Outlook client, then you can't revoke the message.

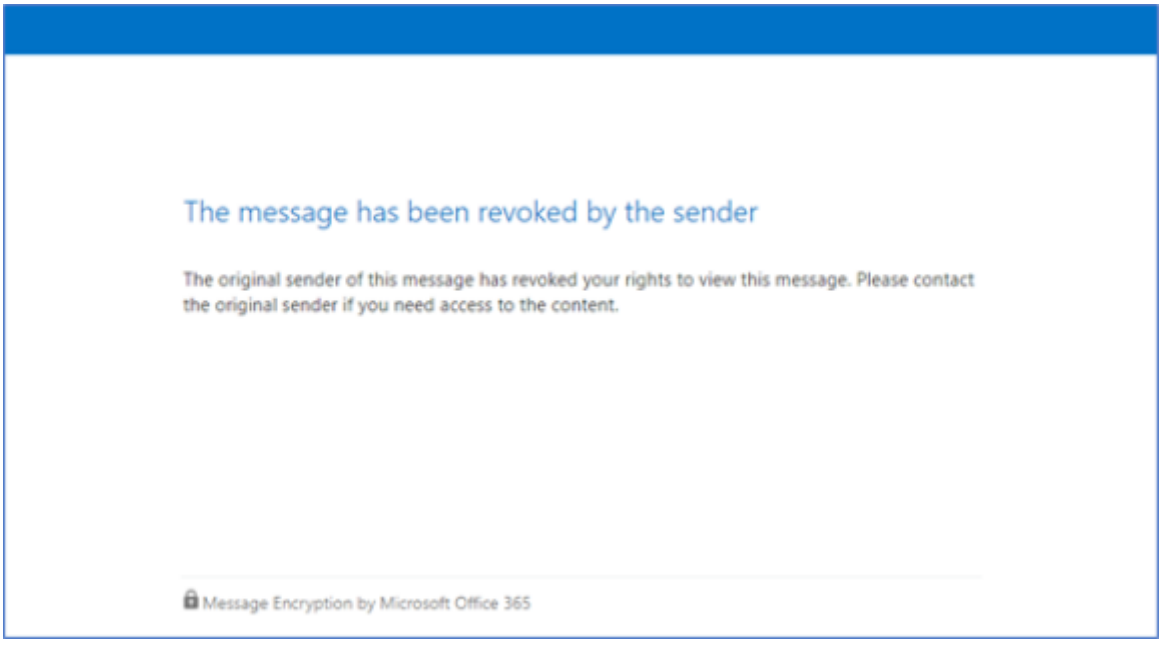
Whether a recipient receives a link-based experience or an inline experience depends on the recipient identity type: Microsoft 365 and Microsoft account recipients (for example, outlook.com users) get an inline experience in supported Outlook clients. All other recipient types, such as Gmail and Yahoo recipients, get a link-based experience.

Admins and message senders can revoke messages that are encrypted using encryption applied directly from Outlook on the web. For example, messages encrypted with the Encrypt Only option.



Recipient experience for revoked encrypted emails

Once an email has been revoked, the recipient receives an error when they access the encrypted email through the encryption portal: "The message has been revoked by the sender".



How to revoke an encrypted message that you sent

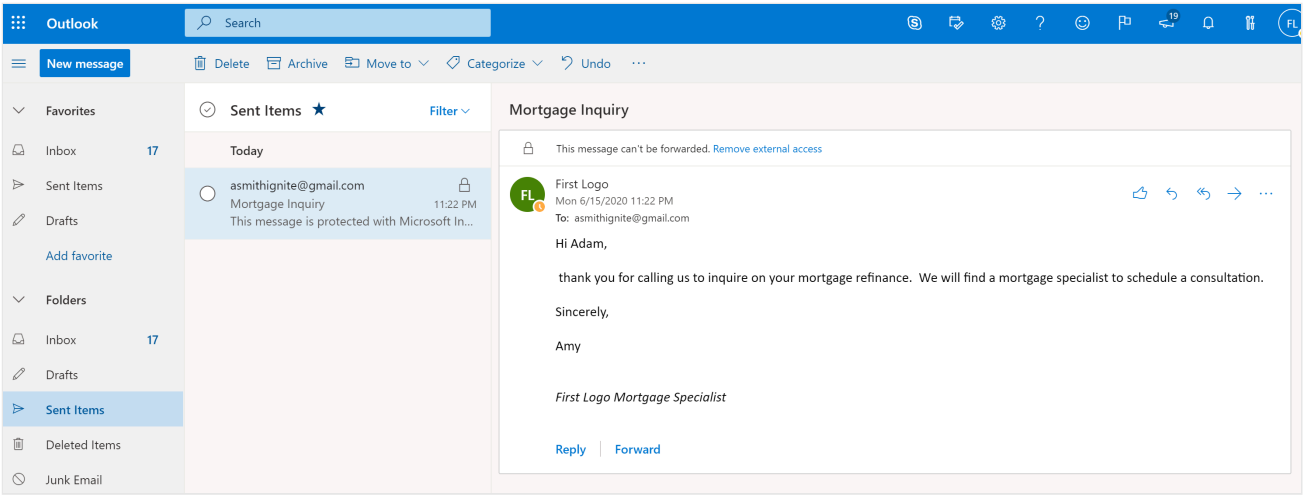
You can revoke a mail that you sent to a single recipient that uses a social account such as gmail.com or yahoo.com. In other words, you can revoke an email sent to a single recipient that received the link-based experience.

You cannot revoke a mail that you sent to a recipient that uses a work or school account from Microsoft 365 or a user that uses a Microsoft account, for example, an outlook.com account.

To revoke an encrypted message that you sent, complete these steps

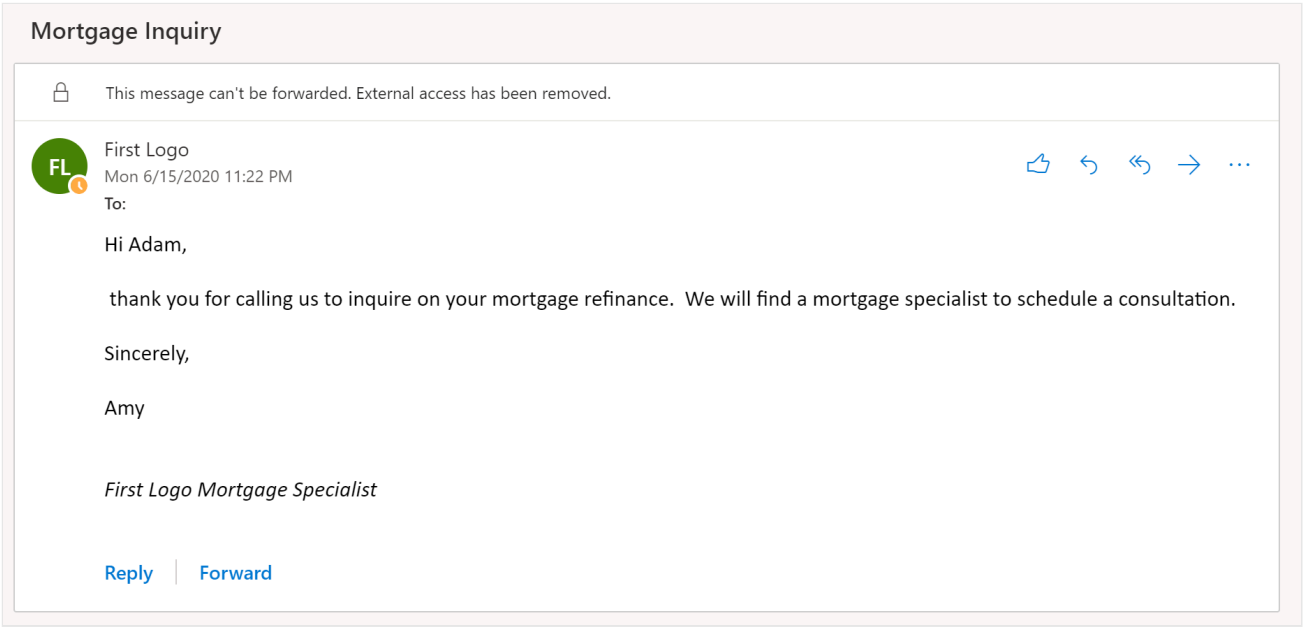
1. In Outlook on the web, in your **Sent** folder, browse to the message you want to revoke.

   If the mail is revocable, you'll see the "Remove external access" link at the top of the message.



2. Click **Remove external access** to revoke the message.

   The message shows that its status is revoked.



How to revoke an encrypted message as an administrator

Microsoft 365 administrators follow these general steps to revoke an eligible encrypted email:

- Get the Message ID of the email.
- Verify that you can revoke the message.

- Revoke the mail.

Step 1: Obtain the Message ID of the email

Before you can revoke an encrypted mail, gather the Message ID of the mail. The MessageId is usually of the format:

`<xxxxxxxxxxxxxxxxxxxxxxx@xxxxxx.xxxx.prod.outlook.com>`

There are multiple ways to find the Message ID of the email that you want to revoke. The following procedures describe two options, but you can use any method that provides the ID.

To identify the Message ID of the email you want to revoke by using Message Trace in the Microsoft Purview portal

Use Message Trace in the Microsoft Purview portal to find the Message ID for the encrypted email.

1. Search for the email by sender or recipient using New Message Trace in Microsoft Purview portal.
2. Once you've located the email, select it to bring up the **Message trace details** pane. Expand **More Information** to locate the Message ID.

To identify the Message ID of the email you want to revoke by using Message Encryption reports in the Microsoft Purview portal

You can also use Message Encryption reports in the Microsoft Purview portal to locate the Message ID.

1. In the Microsoft Purview portal, navigate to the **Message encryption report**. For information on this report, see View email security reports in the Security & Compliance Center.
2. Choose the **View details** table and identify the message that you want to revoke.
3. Double-click the message to view details that include the Message ID.

Step 2: Verify that the mail is revocable

To verify whether you can revoke a message, check whether the Revocation Status field is visible in the Encryption report, in the **Details** table in the Microsoft Purview portal.

To verify whether you can revoke a particular email message by using Windows PowerShell, complete these steps.

1. Using a work or school account that has sufficient permissions in your organization, such as Compliance Administrator, start a Windows PowerShell session and connect to Exchange Online. For instructions, see Connect to Exchange Online PowerShell.

2. Run the Get-OMEMessageStatus cmdlet as follows:

   PowerShell

   ```
   Get-OMEMessageStatus -MessageId "<message id>" | ft -a  Subject, IsRevocable
   ```

   This command returns the subject of the message and whether the message is revocable. For example,

   Console

   ```
   Subject        IsRevocable
   -------        -----------
   "Test message" True
   ```

Step 3: Revoke the mail

Once you know the Message ID of the email you want to revoke, and you have verified that the message is revocable, you can revoke the email using the Microsoft Purview portal or Windows PowerShell.

To revoke the message using the Microsoft Purview portal

1. Using a work or school account that has compliance administrator permissions in your organization, connect to the Microsoft Purview portal.

2. In the **Encryption report**, in the **Details** table for the message, choose **Revoke message**.

To revoke an email by using Windows PowerShell, use the Set-OMEMessageRevocation cmdlet.

1. Using a work or school account that has compliance administrator, or Exchange administrator permissions in your organization, connect to Exchange Online PowerShell.

2. Run the Set-OMEMessageRevocation cmdlet as follows:

   PowerShell

```
Set-OMEMessageRevocation -Revoke $true -MessageId "<messageId>"
```

3. To check whether the email was revoked, run the Get-OMEMessageStatus cmdlet as follows:

PowerShell

```
Get-OMEMessageStatus -MessageId "<messageId>" | ft -a  Subject, Revoked
```

If revocation was successful, the cmdlet returns the following result:

Console

```
Revoked: True
```

Advanced Message Encryption references and related guidance

- Microsoft Purview Advanced Message Encryption
- Microsoft Purview Advanced Message Encryption - email expiration
- Message policy and compliance service description

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Compare versions of message encryption

Article • 03/31/2025

> **Important**
>
> On February 28, 2021, Microsoft deprecated support for AD RMS in Exchange Online. If you've deployed a hybrid environment where your Exchange mailboxes are online and you're using IRM with Active Directory RMS on-premises, you'll need to migrate to Azure. Organizations that have deployed into the GCC Moderate environment are also affected. See "Overview of AD RMS deprecation in Exchange Online" in this article for information.

The rest of this article compares legacy Office 365 Message Encryption (OME) to Microsoft Purview Message Encryption and Microsoft Purview Advanced Message Encryption. Microsoft Purview Message Encryption is merger and newer version of both OME and Information Rights Management (IRM). Unique characteristics of deploying into GCC High are also outlined. The two can coexist in your organization. For information on how the new capabilities work, see Office 365 Message Encryption (OME).

This article is part of a larger series of articles about message encryption. This article is intended for administrators and ITPros. If you're just looking for information on sending or receiving an encrypted message, see the list of articles in Message encryption and locate the article that best fits your needs.

> **Tip**
>
> If you're not an E5 customer, use the 90-day Microsoft Purview solutions trial to explore how additional Purview capabilities can help your organization manage data security and compliance needs. Start now at the **Microsoft Purview trials hub**. Learn details about **signing up and trial terms**.

Overview of AD RMS deprecation in Exchange Online

Exchange Online includes Information Rights Management (IRM) functionality that provides online and offline protection of email messages and attachments. By default, Exchange Online uses Azure Information Protection. However, your organization may have configured Exchange Online IRM to use on-premises Active Directory Rights Management Service (AD RMS). AD RMS support in Exchange Online is retiring. Instead, Azure Information Protection will replace AD RMS entirely.

To assess whether this deprecation impacts your organization, see How to migrate AD RMS to Azure RMS in Exchange Online. This article provides recommendations on migration options.

Side-by-side comparison of message encryption features and capabilities

Expand table

Situation	Legacy OME	IRM in AD RMS	Microsoft Purview Message Encryption
Sending an encrypted mail	Through Exchange mail flow rules	End-user initiated from Outlook desktop or Outlook on the Web; or through Exchange mail flow rules	End-user initiated from Outlook desktop, Outlook for Mac, or Outlook on the Web; through Exchange mail flow rules (also known as transport rules) and data loss prevention (DLP)
Rights management template	N/A	Do Not Forward option and custom templates	Do Not Forward option, encrypt-only option, and custom templates
Recipient type	Internal and external recipients	Internal recipients only	Internal and external recipients
Experience for internal recipient	Recipients receive an HTML message that they download and open in a web browser or mobile app	Native inline experience in Outlook clients	Native inline experience for recipients in the same organization using Outlook clients. Recipients can read message from encrypted message portal using clients other than Outlook (no download or app required).
Experience for external recipient	Recipients receive an HTML message that they download and open in a web	N/A	Native inline experience for Microsoft 365 recipients. All other recipients can read

Situation	Legacy OME	IRM in AD RMS	Microsoft Purview Message Encryption
	browser or mobile app		message from OME portal (no download or app required).
Attachment permissions	No restrictions on attachments	Attachments are protected	Attachments are protected for the Do Not Forward option and custom templates. Admins can choose whether attachments for the encrypt-only option are protected or not.
Bring your own key (BYOK) support	None	None	BYOK supported

Advantages of Microsoft Purview Message Encryption over legacy OME

The new capabilities provide the following advantages:

- Ability to use the encrypt-only option (which enables secure collaboration), Do Not Forward option, and custom restrictions.
- Senders can send mail encrypted with the new capabilities manually from Outlook Desktop, Outlook for Mac and Outlook on the web clients.
- Microsoft 365 recipients get to use an inline experience in supported Outlook clients. Alternatively, admins can choose to show Microsoft 365 recipients a branded experience.
- Accounts outside of Microsoft 365, such as Gmail, Yahoo, and Microsoft accounts, are federated with the OME portal, which provides a better user experience for these recipients. All other identities use a one-time pass code to access encrypted messages.
- Admins can customize branding, and create multiple branding templates.
- Admins can revoke emails encrypted with the new capabilities.
- The new capabilities provide detailed usage reports through the Microsoft Purview portal.

Microsoft Purview Advanced Message Encryption capabilities

Microsoft Purview Advanced Message Encryption offers more capabilities on top of Microsoft Purview Message Encryption. You must have Microsoft Purview Message Encryption set up in your organization in order to use Advanced Message Encryption. Also, in order to use these capabilities, recipients must view and reply to secure mail through the Microsoft Purview Message Encryption Portal. The advanced capabilities include:

- Message revocation
- Message expiration
- Multiple branding templates
- Encrypted message portal activity logs

For information on using Advanced Message Encryption, see Microsoft Purview Advanced Message Encryption.

Unique characteristics of Microsoft Purview Message Encryption in a GCC High deployment

If you plan to use Microsoft Purview Message Encryption in a GCC High environment, there are some unique characteristics regarding the recipient experience.

Encrypted email between GCC High and GCC High recipients

Senders can manually encrypt emails in Outlook for PC and Mac and Outlook on the web, or organizations can set up a policy to encrypt emails using Exchange mail flow rules.

Recipients inside GCC High receive the same inline reading experience in Outlook for PC and Mac and Outlook on the web as all other users.

Encrypted email between GCC High and Non-GCC High recipients

Senders inside GCC High can send encrypted email outside of the GCC High boundary and vice versa.

All recipients outside GCC High, including commercial Microsoft 365 users, Outlook.com users, and other users of other email providers such as Gmail and Yahoo, receive a

wrapper mail. This wrapper mail redirects the recipient to the Microsoft Purview Message Encryption Portal where the recipient can read and reply to the message. This is also true for senders outside GCC High sending OME encrypted mail to GCC High.

Coexistence of legacy OME and Microsoft Purview Message Encryption in the same tenant

You can use both legacy OME and Microsoft Purview Message Encryption in the same tenant. As an administrator, you do this by choosing which version of message encryption you want to use when you create your mail flow rules.

- To specify the legacy version of OME, use the Exchange mail flow rule action **Apply the previous version of OME**.
- To specify Microsoft Purview Message Encryption, use the Exchange mail flow rule action **Apply Office 365 Message Encryption and rights protection**.

Users can manually send mail that is encrypted with Microsoft Purview Message Encryption from Outlook Desktop, Outlook for Mac, and Outlook on the web.

Migrate from legacy OME to Microsoft Purview Message Encryption

Even though both versions can coexist, we highly recommend that you edit your old mail flow rules that use the rule action **Apply the previous version of OME** to use Microsoft Purview Message Encryption. Update these rules to use the mail flow rule action **Apply Office 365 Message Encryption and rights protection**, select "Encrypt" in the RMS template list. For instructions, see Define mail flow rules to encrypt email messages.

Get started with OME

Typically, Microsoft Purview Message Encryption is automatically enabled for your organization. For more information about Microsoft Purview Message Encryption within your organization, see Set up Microsoft Purview Message Encryption.

The legacy version of OME is automatically enabled for your organization if you have enabled Azure Information Protection. In the past, legacy OME worked even if Azure Information Protection wasn't enabled. This is no longer the case.

To start using legacy OME, if you have enabled Azure Information Protection, configure mail flow rules that use the rule action **Apply the previous version of OME**. For instructions, see Define mail flow rules to encrypt email messages.

Feedback

Was this page helpful? Yes No

Message encryption FAQ

Have a question about how the new message protection capabilities work? Check for an answer here. Also, look at [Frequently asked questions]](azure-rights-management-learn-about.md#frequently-asked-questions) for answers to questions about the Azure Rights Management encryption service from Microsoft Purview Information Protection.

What is Microsoft Purview Message Encryption?

Microsoft Purview Message Encryption combines email encryption and rights management capabilities. Rights management capabilities are powered by Microsoft Purview Information Protection.

Who can use Microsoft Purview Message Encryption?

You can use Microsoft Purview Message Encryption under the following conditions:

- If you haven't set up Office 365 Message Encryption (OME) or Information Rights Management (IRM) for Exchange.
- If you set up OME and IRM, you can use these steps if you're also using the Azure Rights Management service from Microsoft Purview Information Protection.
- If you're using Exchange with Active Directory Rights Management service (AD RMS), you can't enable these new capabilities right away. Instead, you need to migrate AD RMS to Azure Information Protection first. When you finish the migration, you can successfully set up Microsoft Purview Message Encryption.

  If you choose to continue to use on-premises AD RMS with Exchange instead of migrating to the Azure Rights Management service, you can't use Microsoft Purview Message Encryption.

What subscriptions do I need to use Microsoft Purview Message Encryption?

To use Microsoft Purview Message Encryption, you need one of the following plans:

- Microsoft Purview Message Encryption is offered as part of Office 365 Enterprise E3 and E5, Microsoft 365 Enterprise E3 and E5, Microsoft 365 Business Premium, Office 365 A1, A3, and A5, and Office 365 Government G3 and G5. You don't need extra licenses to receive the new protection capabilities powered by Microsoft Purview Information Protection.
- You can also add Azure Information Protection Plan 1 to the following plans to receive Microsoft Purview Message Encryption: Exchange Plan 1, Exchange Plan 2, Office 365 F3, Microsoft 365 Business Basic, Microsoft 365 Business Standard, or Office 365 Enterprise E1.
- Each user benefiting from Microsoft Purview Message Encryption needs a license to use message encryption.
- For the full list, see the Exchange service descriptions for Microsoft Purview Message Encryption.

Can I use Exchange with bring your own key (BYOK) for the Azure Rights Management service?

Yes! Microsoft recommends that you complete the steps to set up BYOK before you set up Microsoft Purview Message Encryption.

For more information about BYOK, see Planning and implementing your Azure Information Protection tenant key.

Do Microsoft Purview Message Encryption and BYOK with the Azure Rights Management service change Microsoft's approach to non-Microsoft data requests such as subpoenas?

No. Microsoft Purview Message Encryption and the option to provide and control your own encryption keys, called BYOK for the Azure Rights Management service, weren't designed to respond to law enforcement subpoenas. OME, with BYOK for the Azure Rights Managment service, was designed for compliance-focused organizations. Microsoft takes requests for customer data seriously. As a cloud service provider, we always advocate for the privacy of your

data. In the event we get a subpoena, we always attempt to redirect the request directly to you to obtain the information. (Read Brad Smith's blog: Protecting customer data from government snooping). We periodically publish detailed information of the request we receive. For more information regarding non-Microsoft data requests, see Responding to government and law enforcement requests to access customer data on the Microsoft Trust Center. Also, see "Disclosure of Customer Data" in the Online Services Terms (OST) .

How is this feature related to legacy Office 365 Message Encryption (OME) and Information Rights Management (IRM) features?

Microsoft Purview Message Encryption is an evolution of the existing IRM and legacy OME solutions. The following table provides more details.

Comparison of legacy OME, IRM, and Microsoft Purview Message Encryption

Expand table

Capability	Previous versions of OME	IRM	Microsoft Purview Message Encryption
Sending an encrypted email	Only through Exchange mail flow rules	End-user initiated from Outlook for Windows, Outlook for Mac, or Outlook on the web; or through Exchange mail flow rules	End-user initiated from Outlook for Windows, Outlook for Mac, or Outlook on the web; or through mail flow rules
Rights management	-	Do Not Forward option and custom templates	Do Not Forward option, encrypt-only option, default and custom templates
Supported recipient type	External recipients only	Internal recipients only	Internal and external recipients

Capability	Previous versions of OME	IRM	Microsoft Purview Message Encryption
Experience for recipient	External recipients received an HTML message that they downloaded and opened in a browser or downloaded mobile app.	Internal recipients only received encrypted email in Outlook for Windows, Outlook for Mac, and Outlook on the web.	Internal and external recipients receive email in Outlook for Windows, Outlook for Mac, Outlook on the web, Outlook for Android, and Outlook for iOS, or through a web portal, regardless of whether they are in the same organization or in any organization. The encrypted message portal requires no separate download.
Bring Your Own Key support	Not available	Not available	BYOK supported

How do I enable Microsoft Purview Message Encryption for my organization?

See Set up Microsoft Purview Message Encryption.

Is Office 365 Message Encryption deprecated?

Office 365 Message Encryption (OME) was deprecated on July 1, 2023. It's being automatically replaced with Microsoft Purview Message Encryption. If you have an active sender mailbox, you can still view mail from OME.

My organization uses Active Directory Rights Management, can I use this functionality?

No. If you're using Exchange Online with Active Directory Rights Management service (AD RMS), you can't enable these new capabilities right away. Instead, you need to migrate AD RMS to Azure Information Protection first.

My organization has an Exchange Hybrid deployment. Can I use this feature?

On-premises users can send encrypted mail using Exchange Online mail flow rules. You need to route email through Exchange. For more information, see Part 2: Configure mail to flow from your email server to Microsoft 365.

What email client do I need to use to create an encrypted message? What applications are supported for sending protected messages?

You can create protected messages from Outlook 2016, Outlook 2013 for Windows and Mac, and from Outlook on the web. For more information on sending encrypted messages, see Send, view, and reply to encrypted messages in Outlook for PC.

What email clients are supported to read and reply to protected emails?

Microsoft 365 users can read and respond from Outlook for Windows and Mac (2013 and 2016), Outlook on the web, and Outlook mobile (Android and iOS). You can also use the iOS native mail client if your organization allows it. If you aren't a Microsoft 365 user, you can read and reply to encrypted messages on the web through your web browser.

What email clients support the encrypt-only protected emails?

Microsoft 365 users can use Outlook for PC versions 2019 and Microsoft 365 to create mail protected with the encrypt-only policy. Messages that have the encrypt-only policy applied can be read directly in Outlook on the web, in Outlook for iOS and Android, and Outlook for PC versions 2019 and Microsoft 365.

Is there a size limit for messages you can send with OME?

Yes. The maximum message size you can send with Microsoft Purview Message Encryption, including attachments, is 25 MB. For more information, see Message limits.

Is there a limit to the number of recipients you can send OME-encrypted messages to?

Yes. In some cases that have connectors configured such as an Exchange Hybrid deployment, when you include recipients on the **BCC** line, the BCC recipients are removed before mail gets encrypted. Best practice is to move to an Exchange Online, or put all recipients in the **To:** or **CC** fields.

What type of messages does the encrypted message portal support?

The encrypted message portal only supports mail. The portal doesn't support other message types such as calendar or voice mail.

What file types are supported as attachments in protected emails? Do attachments inherit the protection policies and permissions associated with protected emails? What about PDF?

You can attach any file type to a protected mail. Protection policies are applied only to a subset of the file formats mentioned in Supported file types. By default, Microsoft Purview Message Encryption encrypts the following Office files extensions:

- docx
- docm
- dotx
- dotm
- pptx
- pptm
- potx
- potm

- ppsx
- ppsm
- thmx
- xlsx
- xlsm
- xlsb
- xltx
- xltm
- xlam
- xps

Microsoft Purview Message Encryption doesn't support the 97-2003 versions of the following Office programs: Word (.doc), Excel (.xls), and PowerPoint (.ppt).

If enabled in Exchange Online, PDF encryption allows you to protect sensitive PDF documents attached to emails. When you send an email, the Office 365 service encrypts PDF file attachments for the newest versions of Outlook, including:

- Outlook (new) Desktop
- Outlook on the web
- Outlook for Mac
- Outlook for iOS
- Outlook for Android

To enable encryption for PDF attachments, using a work or school account that has the Information Rights Management role at a minimum in your tenant, run the following command in Exchange Online PowerShell:

PowerShell

```
Set-IRMConfiguration -EnablePdfEncryption $true
```

Protection is inherited from mail to unencrypted attachments only. If a file format is supported, such as a Word, Excel, or PowerPoint file, the file is always protected, even after the recipient downloads the attachment.

For example, say an attachment is protected by Do Not Forward. The original recipient downloads the file, creates a message to a new recipient, and attaches the file. When the new recipient receives the file, they can't open it.

Can I use mail flow rules to encrypt or decrypt meeting invites and their responses?

No, you must use sensitivity labels to apply or remove encryption for meeting invites and their responses. For more information, see Use sensitivity labels to protect calendar items, Teams meetings, and chat.

Are SharePoint or OneDrive attachments supported?

Not yet. SharePoint or OneDrive attachments aren't supported. You can encrypt a mail message, but not the cloud attachments.

What email clients support preview of encrypted attachments in protected emails?

When attachments are protected with a protected mail, you can preview documents directly using Outlook clients. Outlook supports preview of Office documents (docx, xlsx, pptx, doc, xls, ppt). Outlook on the web supports preview of Office documents (docx, xlsx, pptx) and PDF.

What email clients support revocation of protected emails?

Outlook on the web supports revocation of protected mail. See How to revoke an encrypted message that you sent for details.

Does the encrypted message portal support preview of encrypted attachments in protected emails?

The encrypted message portal supports preview of any encrypted attachment copies added to the encrypted mail. The support file types include Word, Excel, PowerPoint, and PDF files.

Can I automatically encrypt messages by setting up policies?

Yes. Use mail flow rules in Exchange Online to automatically encrypt a message based on certain conditions. For example, you can create policies that are based on recipient ID, recipient domain, or on the content in the body or subject of the message. See Define mail flow rules to encrypt email messages in Office 365.

Can I automatically remove encryption on incoming and outgoing mail?

Admins can set up a mail flow rule to remove encryption for outgoing mail. You can only set up a rule to remove encryption for incoming mail that originates from your Exchange Online organization.

Can I automatically remove encryption on journal mail?

For an Exchange Online mailbox, admins must enable journal decryption and set up an Exchange Online journaling rule to generate a decrypted copy of the mail into the journaling mailbox. The journaling rule takes any mail or attachment that has encryption and send the original plus a decrypted copy into the journaling mailbox. You can only set up a journaling rule that can decrypt mail or attachments when the encrypted item originates from your organization.
To enable Exchange Online journaling:

PowerShell

```
Set-IRMConfiguration -JournalReportDecryptionEnabled $true
```

Can I automatically encrypt messages by setting up policies in Data Loss Prevention (DLP) through the Microsoft Purview portal?

Yes! You can set up mail flow rules in Exchange Online or by using DLP in the Microsoft Purview portal.

Can I customize encrypted messages with my company branding?

Yes, for mail sent from an Exchange Online mailbox in your organization! For information on customizing email messages and the encrypted message portal, see Add your organization's brand to your encrypted messages.

On what types of recipients do the encrypted message portal activity logs work?

The encrypted message portal activity logs only capture events for external recipients by accessing the encrypted message portals. Any activities in email clients triggered by external recipients aren't recorded. For internal recipients, see the *MailItemsAccessed* mailbox-auditing action in Purview Audit (Premium) - Mail items accessed logs.

Are there any reporting capabilities or insights for encrypted emails?

There's an Encryption report in the Microsoft Purview portal. See Microsoft Purview Reports overview.

Can I use message encryption with compliance features such as eDiscovery?

Yes, most messages protected by Microsoft Purview Message Encryption are discoverable. Microsoft Purview Message Encryption protected mail that you receive from another Microsoft 365 organization that has custom branding applied through a mail flow rule is undiscoverable by your eDiscovery service. In other words, if the mail isn't accessible through the user's mailbox, but rather, surfaced only through a link to the encrypted message portal, the mail isn't searchable. See eDiscovery activities that support encrypted items for details.

Can I send as a shared mailbox and encrypt emails?

When an email message matches an encryption mail flow rule, Exchange encrypts the message sending it.

Can I open encrypted messages sent to a shared mailbox?

Yes! You can open encrypted messages for a shared mailbox. When the mail is sent from the same organization, you can open the mail when you're signed in to a supported Outlook client. If the mail is sent from an external organization, you need to use Outlook on the web.

- Users can open protected mails in a shared mailbox where the shared mailbox received a protected mail as part of a distribution group.
- Users can view attachments that inherit protection from email when they use Outlook for Windows, Outlook for Mac, Outlook for Android, Outlook for iOS, and Outlook on the web.

The following table lists the supported clients for shared mailboxes.

Expand table

Platform	Read mail	View email attachments
Outlook on the web	Yes	Yes
Outlook for Windows	Yes	Yes*
Outlook for Mac	Yes	Yes
Outlook for Android	Yes	Yes*
Outlook for iOS	Yes	Yes*

> **Note**
>
> Android and iOS use the Office mobile app to display encrypted attachments, and don't display attachments directly in Outlook mobile. Outlook for Windows can display encrypted

attachments when user is directly assigned to a shared mailbox. (See below on user assignment.)

There are currently two known limitations:

- You can't open attachments to emails that you receive on mobile devices by using Outlook mobile.
- There are two ways to allow a user to view the encrypted mail directly in Outlook for Windows:

    1. Directly assign a user to the shared mailbox, with full access permissions and automapping enabled. For Outlook 64-bit, users don't need to be directly assigned to the mailbox. Automapping is enabled by default for Exchange.
    2. Assign a mail-enabled security group to a shared mailbox. This method requires Outlook version 2402 and only supports mail generated after June 2024.

To assign a user to a shared mailbox

1. Connect to Exchange Online PowerShell.
2. Run the `Add-MailboxPermission` cmdlet with the `Automapping` parameter. This example gives Ayla full access permission to a support mailbox.

    PowerShell

    ```
    Add-MailboxPermission -Identity support@contoso.onmicrosoft.com -User
    ayla@contoso.com -AccessRights FullAccess -AutoMapping $true
    ```

To assign a mail enabled security group to the shared mailbox

To use this method, you must encrypt mail with the **Do Not Forward** or **Encrypt-only** protection options. Only mail initiated by the shared mailbox or mail sent within an organization can be opened.

1. Connect to Exchange Online PowerShell.
2. Run the `Add-MailboxPermission` cmdlet to assign the security group. The following example gives the Contoso front desk security group full access permission to a support mailbox.

    PowerShell

```
Add-MailboxPermission -Identity support@contoso.onmicrosoft.com -User
frontdesk@contoso.com -AccessRights FullAccess
```

Is delegated access supported with opening encrypted messages? Even if a delegate has full access to another user's mailbox?

When delegates are given full access permission to a user's mailbox, delegated access of encrypted mail is supported in Outlook on the web, Outlook for Mac, Outlook for iOS, and Outlook for Android. Outlook for Windows doesn't support delegated access.

How long do I have access to the mail in the encrypted message portal?

You can sign in to the encrypted message portal to retrieve mail as long as the sender's organization is active and the mail isn't configured to expire.

What do I do if I don't receive the one-time pass code after I requested it?

First, check the junk or spam folder in your email client. DKIM and DMARC settings for your organization might cause these emails to end up filtered as spam.

Next, check quarantine in the Microsoft Purview portal. Often, messages containing a one-time pass code, especially the first ones your organization receives, end up in quarantine.

Last updated on 11/19/2025

How Exchange Online secures your email secrets

> *Microsoft Purview service description*

This article describes how Microsoft secures your email secrets in its data centers.

How we secure secret information provided by you

In addition to the Office 365 Trust Center that provides Security, Privacy, and Compliance Information for Office 365, we use a technology called Distributed Key Manager (DKM).

Distributed Key Manager (DKM) is a client-side technology that uses a set of secret keys to encrypt and decrypt information. Only members of a specific security group in Active Directory Domain Services can access those keys in order to decrypt the data that is encrypted by DKM. In Exchange Online, only certain service accounts under which the Exchange processes run are part of that security group. No human is given credentials that are part of this security group and therefore no human has access to the keys that can decrypt these secrets.

For debugging, troubleshooting, or auditing purposes, a data center administrator must request elevated access to gain temporary credentials that are part of the security group. This process requires multiple levels of legal approval. If access is granted, all activity is logged and audited. Access is only granted for a set interval of time after which it automatically expires.

For extra protection, DKM technology includes automated key rollover and archiving. Automated rollover and archiving ensure that you can continue to access your older content without having to rely on the same key indefinitely.

Where Exchange Online uses DKM

Microsoft uses Distributed Key Manager to encrypt your secrets in Exchange Online data centers. For example:

- Email account credentials for connected accounts. Connected accounts are third-party accounts such as Hotmail, Gmail, and Yahoo! mail accounts.

- Customer Key. If you're using Customer Key, you'll use Azure Key Vault to safeguard your secrets.

Related articles

Encryption in Office 365

Technical reference details about encryption

Service assurance in the Microsoft Purview portal

Last updated on 05/14/2026

How Exchange Online uses TLS to secure email connections

Article • 11/27/2023

Learn how Exchange Online and Microsoft 365 use Transport Layer Security (TLS) and Forward Secrecy (FS) to secure email communications.

> **Tip**
>
> If you're not an E5 customer, use the 90-day Microsoft Purview solutions trial to explore how additional Purview capabilities can help your organization manage data security and compliance needs. Start now at the **Microsoft Purview trials hub**. Learn details about **signing up and trial terms**.

TLS basics for Microsoft 365 and Exchange Online

Transport Layer Security (TLS), and Secure Sockets Layer (SSL) that came before TLS, are cryptographic protocols. These protocols secure communication over a network by using security certificates to encrypt a connection between computers. TLS supersedes SSL and is often referred to as SSL 3.1. Exchange Online uses TLS to encrypt the connections between Exchange servers and the connections between Exchange servers and other servers. For example, TLS is used to encrypt the connection between Exchange Online and your on-premises Exchange servers or your recipients' mail servers. Once the connection is encrypted, all data sent through that connection is sent through the encrypted channel.

TLS doesn't encrypt the message, just the connection. So, if you forward a message that was sent through a TLS-encrypted connection to a recipient organization that doesn't support TLS encryption, that message isn't necessarily encrypted.

If you want to encrypt the message, use an encryption technology that encrypts the message contents. For example, you can use Microsoft Purview Message Encryption or S/MIME. See Email encryption in Office 365 and Message encryption for information on message encryption in Office 365.

Use TLS in situations where you want to set up a secure channel of correspondence between Microsoft and your on-premises organization or another organization, such as a partner. Exchange Online always attempts to use TLS first to secure your email but

can't if the other party doesn't offer TLS security. Keep reading to find out how you can secure all mail to your on-premises servers or important partners by using *connectors*.

To provide the best-in-class encryption to our customers, Microsoft deprecated Transport Layer Security (TLS) versions 1.0 and 1.1 in Office 365 and Office 365 GCC. However, you can continue to use an unencrypted SMTP connection without any TLS. We don't recommend email transmission without any encryption.

How Exchange Online uses TLS between Exchange Online customers

Exchange Online servers always encrypt connections to other Exchange Online servers in our data centers with TLS 1.2. When you send a message to a recipient that is within your organization, Exchange Online automatically sends the message over an encrypted connection using TLS. Exchange Online also sends email that you send to other customers over encrypted connections using TLS that are secured using Forward Secrecy.

How Microsoft 365 uses TLS between Microsoft 365 and external, trusted partners

By default, Exchange Online always uses *opportunistic TLS*. Opportunistic TLS means Exchange Online always tries to encrypt connections with the most secure version of TLS first, then works its way down the list of TLS ciphers until it finds one on which both parties can agree. Unless you configure Exchange Online to ensure that messages to that recipient must use a secure connection, then by default Exchange sends the message without encryption if the recipient's organization doesn't support TLS encryption. Opportunistic TLS is sufficient for most businesses. However, for businesses that have compliance requirements such as medical, banking, or government organizations, you can configure Exchange Online to require, or force, TLS. For instructions, see Configure mail flow using connectors in Office 365.

If you decide to configure TLS between your organization and a trusted partner organization, Exchange Online can use *forced TLS* to create trusted channels of communication. Forced TLS requires your partner organization to authenticate to Exchange Online with a security certificate to send mail to you. Your partner needs to manage their own certificates. Exchange Online uses connectors to protect messages that you send from unauthorized access before they arrive at the recipient's email provider. For information on using connectors to configure mail flow, see Configure mail flow using connectors in Office 365.

TLS and hybrid Exchange Server deployments

If you're managing a hybrid Exchange deployment, your on-premises Exchange server needs to authenticate to Microsoft 365 using a security certificate to send mail to recipients whose mailboxes are only in Office 365. As a result, you need to manage your own security certificates for your on-premises Exchange servers. You must also securely store and maintain these server certificates. For more information about managing certificates in hybrid deployments, see Certificate requirements for hybrid deployments.

How to set up forced TLS for Exchange Online in Office 365

For Exchange Online customers, in order for forced TLS to work to secure all of your sent and received email, you need to set up more than one connector that requires TLS. You need one connector for messages sent to user mailboxes and another connector for messages sent from user mailboxes. Create these connectors in the Exchange admin center in Office 365. For instructions, see Configure mail flow using connectors in Office 365.

TLS certificate information for Exchange Online

The certificate information used by Exchange Online is described in the following table. If your business partner is setting up forced TLS on their email server, you need to provide this information to them. For security reasons, our certificates do change from time to time. The current certificate is valid from September 24, 2020.

Current certificate information valid from September 24, 2020

Expand table

Attribute	Value
Certificate authority root issuer	DigiCert CA - 1
Certificate name	mail.protection.outlook.com
Organization	Microsoft Corporation
Organization unit	www.digicert.com

Attribute	Value
Certificate key strength	2048

Get more information about TLS, certificates, and Microsoft 365 and download certificates

Microsoft 365 encryption chains and certificate downloads

Microsoft 365 encryption chains and certificate downloads - DOD and GCC High

For a list of supported cipher suites, see Technical reference details about encryption.

Set up connectors for secure mail flow with a partner organization

Connectors with enhanced email security

Encryption in Microsoft 365

Feedback

Was this page helpful? Yes No

What is Microsoft 365 service encryption?

09/26/2025

In addition to using volume-level encryption, Exchange Online, Microsoft Teams, SharePoint Online, OneDrive for Business, and Windows 365 Cloud PCs also use service encryption to encrypt customer data.

Service encryption allows for two key management options, Microsoft-managed keys and customer-managed keys.

Service encryption provides multiple benefits:

- Provides a layer of protection for all Microsoft 365 services and Windows 365 Cloud PCs. For Microsoft 365 services, service encryption is an extra layer of protection is on top of BitLocker.
- Provides separation of Windows operating system administrators from access to application data stored or processed by the operating system.
- Includes a Customer Key option that enables multitenant services to provide per-tenant key management.
- Enhances the ability of Microsoft 365 and Windows 365 to meet your specific compliance requirements regarding encryption.

What are Microsoft-managed keys?

By default, Microsoft manages all cryptographic keys including the root keys for service encryption. This option, called *Microsoft-managed keys*, is enabled by default for Exchange Online, SharePoint Online, OneDrive for Business, and Windows 365 Cloud PCs. Microsoft-managed keys provide default service encryption unless you decide to onboard using Customer Key. If, at a later date, you decide to stop using Customer Key without following the data purge path, then your data stays encrypted using the Microsoft-managed keys. Your data is always encrypted at this default level at a minimum.

What is Customer Key?

Customer Key is a customer-managed key solution. With Customer Key, you supply root keys used with service encryption and you manage these keys using Azure Key Vault while Microsoft manages all other keys. Customer Key is available for Exchange, SharePoint and OneDrive for Business, and Windows 365 Cloud PCs (Enterprise and Frontline modes).

Windows 365 support for Microsoft Purview Customer Key is generally available, including Windows 365 Frontline (Dedicated and Shared).

Using Customer Key, you generate your own cryptographic keys. You can use an on-premises Hardware Service Module (HSM) or Azure Key Vault (AKV) to generate your keys. AKV lets you control and manage the cryptographic keys used by Microsoft 365. Customer Key uses the keys stored in the AKV as the root of one of the key chains that encrypts your mailbox data or files.

Customer Key provides you with more control over how Microsoft processes your data. For example, You can use Customer Key as a technical control if you want to terminate service with Microsoft or remove a portion of your data stored in the cloud. Removing data ensures no one, including Microsoft, can access or process the data. Customer Key is in addition and complementary to Customer Lockbox that you use to control access to your data by Microsoft personnel.

To learn how to set up Customer Key for Exchange Online, Microsoft Teams, SharePoint Online, including Team Sites, and OneDrive for Business, see these articles:

- Overview of Customer Key
- Set up Customer Key
- Manage Customer Key
- Roll or rotate a Customer Key or an availability key
- Understand the availability key

BitLocker and Distributed Key Manager (DKM) for Encryption

Article • 09/08/2023

Microsoft servers use BitLocker to encrypt the disk drives containing customer data at rest at the volume-level. BitLocker encryption is a data protection feature that is built into Windows. BitLocker is one of the technologies used to safeguard against threats in case there are lapses in other processes or controls (e.g., access control or recycling of hardware) that could lead to someone gaining physical access to disks containing customer data. In this case, BitLocker eliminates the potential for data theft or exposure because of lost, stolen, or inappropriately decommissioned computers and disks.

BitLocker is deployed with Advanced Encryption Standard (AES) 256-bit encryption on disks containing customer data in Exchange Online, SharePoint Online, and Skype for Business. Disk sectors are encrypted with a Full Volume Encryption Key (FVEK), which is encrypted with the Volume Master Key (VMK), which in turn is bound to the Trusted Platform Module (TPM) in the server. The VMK directly protects the FVEK and therefore, protecting the VMK becomes critical. The following figure illustrates an example of the BitLocker key protection chain for a given server (in this case, using an Exchange Online server).

The following table describes the BitLocker key protection chain for a given server (in this case, an Exchange Online server).

Expand table

KEY PROTECTOR	GRANULARITY	HOW GENERATED?	WHERE IS IT STORED?	PROTECTION
AES 256-bit External Key	Per Server	BitLocker APIs	TPM or Secret Safe	Lockbox / Access Control
			Mailbox Server Registry	TPM encrypted
48-digit Numerical Password	Per Disk	BitLocker APIs	Active Directory	Lockbox / Access Control

KEY PROTECTOR	GRANULARITY	HOW GENERATED?	WHERE IS IT STORED?	PROTECTION
X.509 Certificate as Data Recovery Agent (DRA) also called Public Key Protector	Environment (for example, Exchange Online multitenant)	Microsoft CA	Build System	No one user has the full password to the private key. The password is under physical protection.

BitLocker key management involves the management of recovery keys that are used to unlock/recover encrypted disks in a Microsoft data center. Microsoft 365 stores the master keys in a secured share, only accessible by individuals who have been screened and approved. The credentials for the keys are stored in a secured repository for access control data (what we call a "secret store"), which requires a high level of elevation and management approvals to access using a just-in-time access elevation tool.

BitLocker supports keys which fall into two management categories:

- BitLocker-managed keys, which are short-lived and tied to the lifetime of an operating system instance installed on a server or to a given disk. These keys are deleted and reset during server reinstallation or disk formatting.
- BitLocker recovery keys, which are managed outside of BitLocker but used for disk decryption. BitLocker uses recovery keys for the scenario in which an operating system is reinstalled, and encrypted data disks already exist. Recovery keys are also used by Managed Availability monitoring probes in Exchange Online where a responder may need to unlock a disk.

BitLocker-protected volumes are encrypted with a full volume encryption key, which in turn is encrypted with a volume master key. BitLocker uses FIPS-compliant algorithms to ensure that encryption keys are never stored or sent over the wire in the clear. The Microsoft 365 implementation of customer data-at-rest-protection doesn't deviate from the default BitLocker implementation.

Feedback

Was this page helpful? Yes No

Overview of Service encryption with Microsoft Purview Customer Key

Microsoft 365 provides baseline, volume-level encryption through BitLocker and Distributed Key Manager (DKM). Windows 365 Enterprise and Business Cloud PC disks are encrypted using Azure Storage server-side encryption (SSE).

To give you more control, Microsoft 365 also offers an added layer of encryption for your content through Customer Key. This content includes data from Microsoft Exchange, SharePoint, OneDrive, Teams, and Windows 365 Cloud PCs (Enterprise), including Windows 365 Frontline Dedicated and Shared modes.

BitLocker isn't supported as an encryption option for Windows 365 Cloud PCs. For details, see Using Windows 10 virtual machines in Intune.

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

How service encryption, BitLocker, SSE, and Customer Key work together

Your Microsoft 365 data is always encrypted at rest using BitLocker and Distributed Key Manager (DKM). For details, see How Exchange secures your email secrets.

Customer Key adds extra protection against unauthorized access to your data. It complements BitLocker disk encryption and server-side encryption (SSE) in Microsoft data centers. Customer Key helps you meet compliance or regulatory requirements by letting you control the root encryption keys at the application level.

You explicitly authorize Microsoft 365 to use your encryption keys to deliver value-added services like eDiscovery, anti-malware, anti-spam, and search indexing. Microsoft 365 uses these keys to encrypt your data at rest, as described in the Online Services Terms (OST).

Customer Key with hybrid deployments

Customer Key encrypts only data at rest in the cloud. It doesn't protect on-premises mailboxes or files. To protect on-premises data, use a separate method like BitLocker.

Learn about data encryption policies

A data encryption policy (DEP) defines the encryption hierarchy. Services use this hierarchy to encrypt data with both the keys you manage and the availability key that Microsoft protects. You create a DEP using PowerShell cmdlets, then assign it to encrypt application data.

Customer Key supports three types of DEPs. Each type uses different cmdlets and protects a different kind of data:

DEP for Multiple Microsoft 365 workloads

These DEPs encrypt data across several Microsoft 365 workloads for all users in the tenant:

- **Windows 365 Cloud PCs** (Enterprise), including Frontline Dedicated and Shared modes (details)
- **Teams**: Chat messages, media messages, calls/meeting recordings (in Teams storage), notifications, Cortana suggestions, status messages
- **Microsoft 365 Copilot** interactions
- **Exchange**: User/signal information and mailboxes not covered by a mailbox DEP
- **Microsoft Purview Information Protection**: EDM data (schemas, rule packages, salts), sensitivity label configurations

> ⓘ **Note**
>
> For EDM and Teams, the DEP encrypts new data from assignment. For Exchange, it encrypts all existing and new data.

Not encrypted by multi-workload DEPs (protected by other methods):

- SharePoint and OneDrive data (use SharePoint DEP)
- Teams files and recordings saved in SharePoint/OneDrive
- Teams Live Events, Viva Engage, Planner

You can create multiple DEPs per tenant but assign only one at a time. Encryption begins automatically after assignment.

DEPs for Exchange mailboxes

Mailbox DEPs give you more control over individual Exchange Online mailboxes. You can use them to encrypt data in UserMailbox, MailUser, Group, PublicFolder, and Shared mailboxes.

You can have up to 50 active mailbox DEPs per tenant. You can assign one DEP to multiple mailboxes, but only one DEP per mailbox.

By default, Exchange mailboxes are encrypted using Microsoft-managed keys. When you assign a Customer Key DEP, the service rewraps existing encrypted mailboxes on next access. Unencrypted mailboxes are marked for a move, with encryption occurring after the move completes. For details on move priorities, see Move requests in the Microsoft 365 service.

You can later refresh the DEP or assign a different one as described in Manage Customer Key for Office 365.

Each mailbox must meet licensing requirements to use Customer Key. For more info, see Prerequisites.

You can assign DEPs to shared, public folder, and group mailboxes as long as your tenant meets the licensing requirements for user mailboxes. You don't need separate licenses for non-user-specific mailboxes.

You can also request that Microsoft purge specific DEPs when leaving the service. Revoking access to your keys triggers deletion of the availability key, resulting in cryptographic deletion of your data. For details, see Revoke your keys and start the data purge path process.

DEP for SharePoint and OneDrive

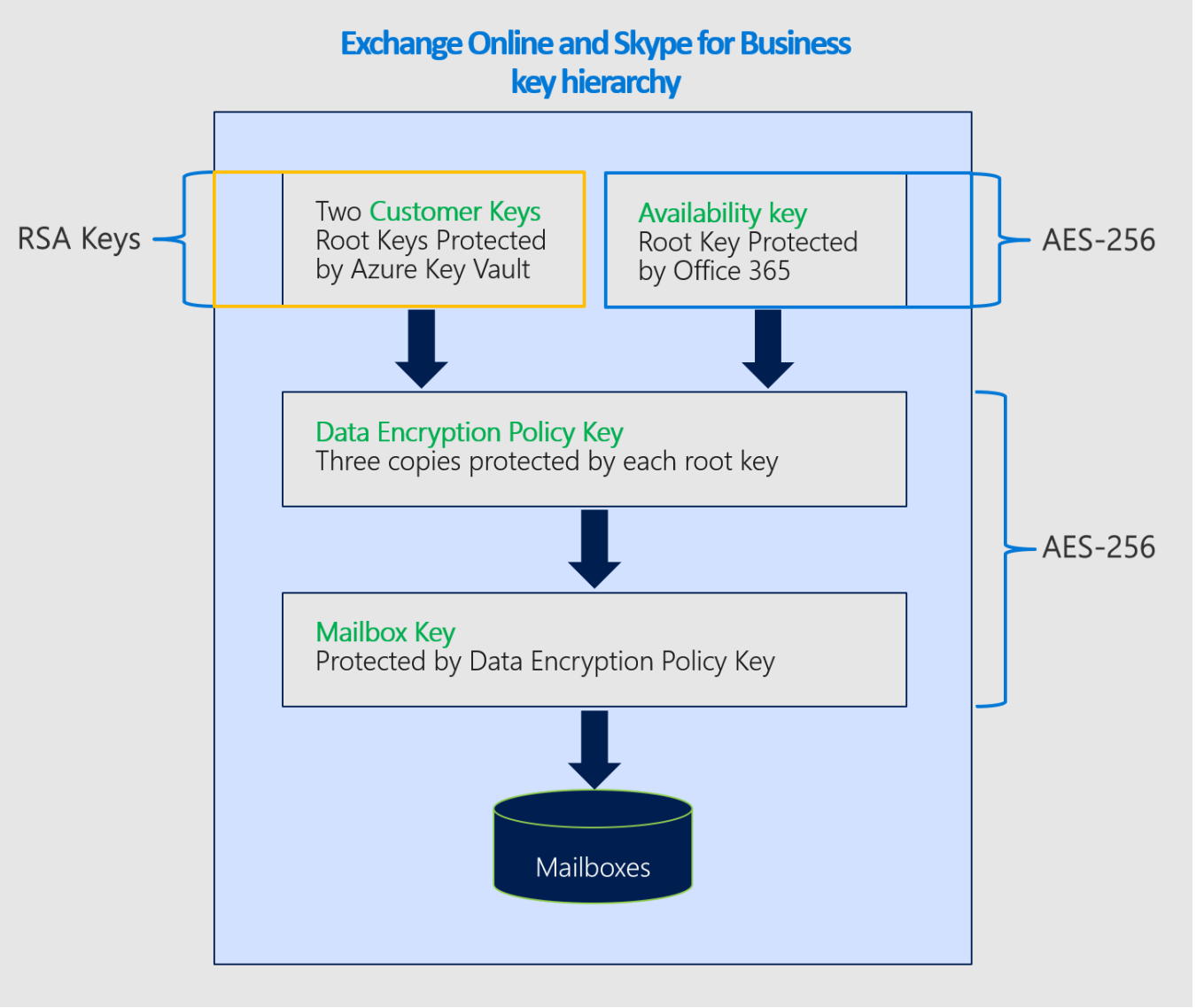
This DEP encrypts content stored in SharePoint and OneDrive, including Teams files stored in SharePoint. Multi-geo tenants can create one DEP per geo; single-geo tenants can create one DEP. For setup instructions, see Set up Customer Key.

Encryption ciphers used by Customer Key

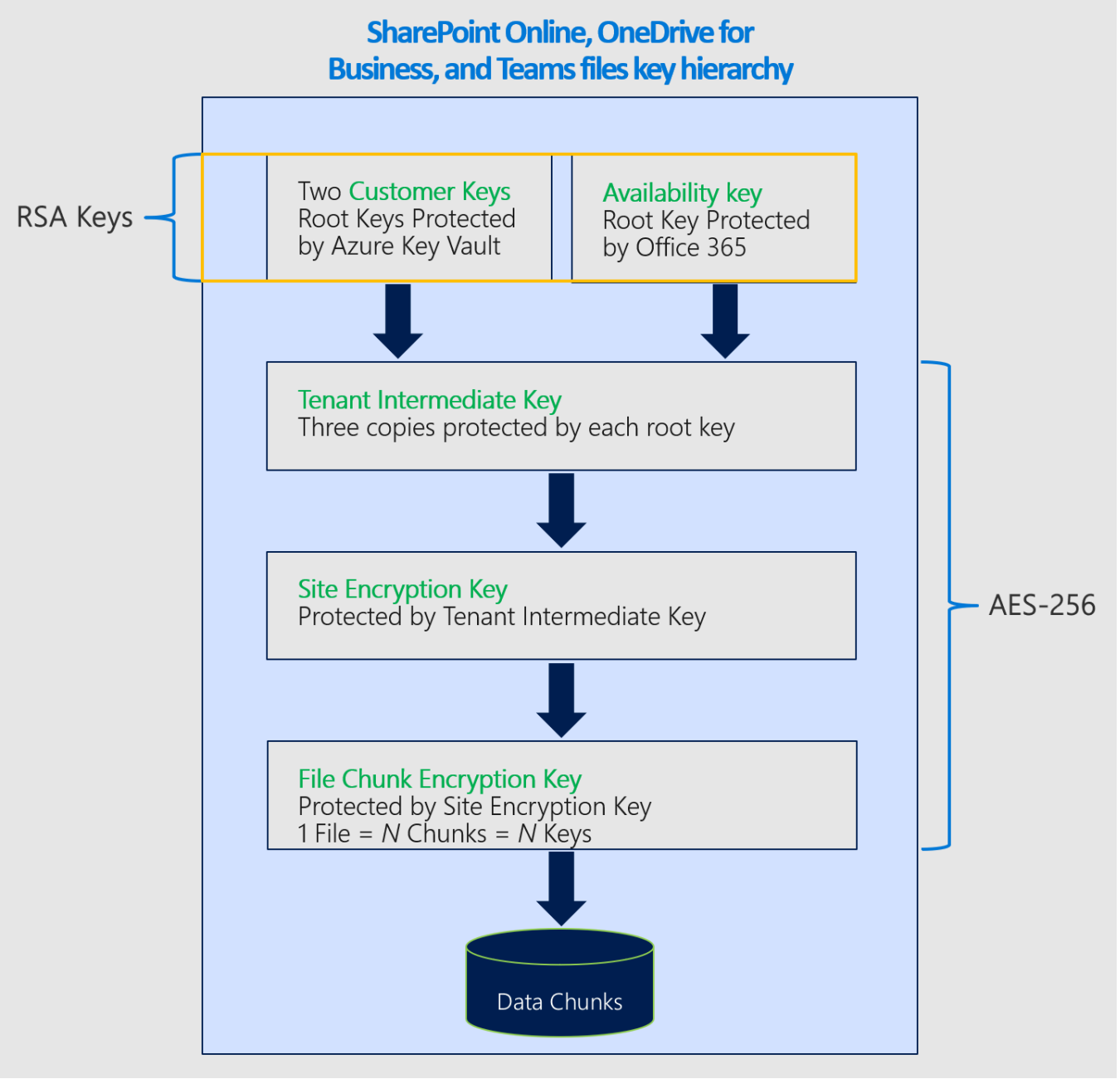
Customer Key uses different encryption ciphers to protect keys, as shown in the following diagrams.

The key hierarchy used for DEPs that encrypt data across multiple Microsoft 365 workloads is similar to the one used for individual Exchange mailboxes. The corresponding Microsoft 365 Workload Key replaces the Mailbox Key.

Encryption ciphers used to encrypt keys for Exchange



Encryption ciphers used to encrypt keys for SharePoint and OneDrive



Related articles

- Set up Customer Key
- Manage Customer Key
- Roll or rotate a Customer Key or an availability key
- Learn about the availability key
- Learn about Customer Lockbox
- Understand Service encryption and key management

Last updated on 01/15/2026

Set up Customer Key

Customer Key lets you control your organization's encryption keys and configure Microsoft 365 to use those keys to encrypt your data at rest in Microsoft's datacenters.

This article walks you through creating and configuring the required Azure resources, then enabling Customer Key. After setup, you assign policies that determine which keys encrypt data across Microsoft 365 workloads.

For a general overview and more information, see Overview of Customer Key.

> ⓘ **Important**
>
> Be sure to follow the best practices in this article marked as **TIP**, **IMPORTANT**, and **NOTE**. Customer Key gives you control over root encryption keys that can affect your entire organization. Mistakes with these keys can lead to service interruptions or permanent data loss.

Prerequisites

Before setting up Customer Key, ensure you have:

- **Licensing**: One of the following:
  - Office 365 E5
  - Microsoft 365 E5
  - Microsoft Purview Suite (formerly Microsoft 365 E5 Compliance)
  - Microsoft 365 E5 Information Protection & Governance
  - Microsoft 365 Security and Compliance for FLW
  - Existing Office 365 Advanced Compliance licenses are also supported
- **Azure subscriptions**: Two paid Azure subscriptions (Free, Trial, Sponsorship, MSDN, or Legacy Support aren't eligible)
- **Permissions**: Exchange admin permissions (all Customer Key operations run through Exchange Online PowerShell)
- **PowerShell modules**:
  - Azure PowerShell (v4.4.0 or higher recommended) for Key Vault/HSM operations
  - Exchange Online PowerShell for DEP operations
- **Azure knowledge**: Familiarity with Azure Key Vault and Microsoft Entra tenant concepts

> ⓘ **Important**

Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview **roles and permissions**.

If you need help beyond the documentation, contact Microsoft Support. To share feedback or suggestions about Customer Key or this documentation, visit the Microsoft 365 Community.

Overview of steps to set up Customer Key

To set up Customer Key, complete the following tasks in order. The rest of this article provides detailed instructions for each task, or links out to more information for each step in the process.

> **Important**
>
> The setup process differs depending on whether you're using **Azure Key Vault** or **Managed HSM**. The shared prerequisites appear first, followed by configuration-specific instructions.

Prerequisites for all configurations

- Create two new Azure subscriptions
- Register the required Service Principals

Configuration-specific setup

- Set up using Azure Key Vault
- Set up using Managed HSM

Onboarding options

Choose the onboarding method that applies to your scenario:

- Onboard using the Customer Key Onboarding Service (recommended for most tenants)
- Onboard to Customer Key for SharePoint and OneDrive
- Onboard to Customer Key for special cloud environments

Quick reference: Key setup commands

Use the following tables to quickly find the commands you need during setup.

Service principal registration

Expand table

Task	Command
Check if Customer Key Onboarding app is registered	`Get-AzADServicePrincipal -ServicePrincipalName 19f7f505-34aa-44a4-9dcc-6a768854d2ea`
Register Customer Key Onboarding app	`New-AzADServicePrincipal -ApplicationId 19f7f505-34aa-44a4-9dcc-6a768854d2ea`
Check if M365DataAtRestEncryption app is registered	`Get-AzADServicePrincipal -ServicePrincipalName c066d759-24ae-40e7-a56f-027002b5d3e4`
Register M365DataAtRestEncryption app	`New-AzADServicePrincipal -ApplicationId c066d759-24ae-40e7-a56f-027002b5d3e4`
Check if Exchange Online app is registered	`Get-AzADServicePrincipal -ServicePrincipalName 00000002-0000-0ff1-ce00-000000000000`
Register Exchange Online app	`New-AzADServicePrincipal -ApplicationId 00000002-0000-0ff1-ce00-000000000000`

Azure Key Vault operations

Expand table

Task	Command
Create a key	`Add-AzKeyVaultKey -VaultName <VaultName> -Name <KeyName> -Destination HSM`
Get key URI	`(Get-AzKeyVaultKey -VaultName <VaultName>).Id`
Check key expiration	`Get-AzKeyVaultKey -VaultName <VaultName>`
Update key expiration	`Update-AzKeyVaultKey -VaultName <VaultName> -Name <KeyName> -Expires (Get-Date -Date "12/31/9999")`
Back up a key	`Backup-AzKeyVaultKey -VaultName <VaultName> -Name <KeyName> -OutputFile <BackupFile>`

Managed HSM operations

Expand table

Task	Command
Create a key	`Add-AzKeyVaultKey -HsmName <HSMName> -Name <KeyName> -Destination HSM`
Get key URI	`(Get-AzKeyVaultKey -HsmName <HSMName>).Id`
Check key expiration	`Get-AzKeyVaultKey -HsmName <HSMName>`
Update key expiration	`Update-AzKeyVaultKey -HsmName <HSMName> -Name <KeyName> -Expires (Get-Date -Date "12/31/9999")`
Back up a key	`Backup-AzKeyVaultKey -HsmName <HSMName> -Name <KeyName> -OutputFile <BackupFile>`

Customer Key onboarding

Expand table

Task	Command
Install onboarding module	`Install-Module -Name M365CustomerKeyOnboarding`
Create onboarding request	`New-CustomerKeyOnboardingRequest -Organization <TenantID> -Scenario <MDEP|EXO> -Subscription1 <SubID1> -KeyIdentifier1 <KeyURI1> -Subscription2 <SubID2> -KeyIdentifier2 <KeyURI2> -OnboardingMode <Validate|Enable>`
View onboarding requests	`Get-CustomerKeyOnboardingRequest -OrganizationID <TenantID>`
View failed validations	`$request.FailedValidations`
View passed validations	`$request.PassedValidations`

Create two new Azure subscriptions

Customer Key requires two Azure subscriptions. As a best practice, Microsoft recommends creating new Azure subscriptions specifically for use with Customer Key.

Azure Key Vault keys can only be authorized for applications in the same Microsoft Entra tenant. To assign data encryption policies (DEPs), make sure to create both subscriptions under the same Microsoft Entra tenant used by your organization. For example, use your work or school account that has proper administrator privileges in your organization. For step-by-step guidance, see Sign up for Azure as an organization.

> **Important**
>
> Customer Key requires two keys for each DEP. To support this requirement, you must create two separate Azure subscriptions. As a best practice, have different members of your organization manage one key in each subscription. These subscriptions should be used only to administer encryption keys for Microsoft 365. This configuration helps protect your organization in case someone accidentally, intentionally, or maliciously deletes or mismanages a key they control.

Following these best practices reduces the risk of human error and simplifies Customer Key management.

Register the required Service Principals

To use the Customer Key, your tenant must have the required service principals registered. The following sections show you how to check whether the service principals are already registered in your tenant. If they aren't, run the *'New-AzADServicePrincipal'* cmdlet.

Register the Service Principal for the Customer Key Onboarding Application

To check if the Customer Key Onboarding application is already registered with the correct permissions, run the following command:

PowerShell

```
Get-AzADServicePrincipal -ServicePrincipalName 19f7f505-34aa-44a4-9dcc-6a768854d2ea
```

If it isn't registered, run:

PowerShell

```
New-AzADServicePrincipal -ApplicationId 19f7f505-34aa-44a4-9dcc-6a768854d2ea
```

Register the Service Principal for the M365DataAtRestEncryption Application

To check if the M365DataAtRestEncryption application is already registered with the correct permissions, run the following command:

```
PowerShell
Get-AzADServicePrincipal -ServicePrincipalName c066d759-24ae-40e7-a56f-027002b5d3e4
```

If it isn't registered, run:

```
PowerShell
New-AzADServicePrincipal -ApplicationId c066d759-24ae-40e7-a56f-027002b5d3e4
```

Register the Service Principal for the Office 365 Exchange Online Application

To check if the Office 365 Exchange Online application is already registered with the correct permissions, run the following command:

```
PowerShell
Get-AzADServicePrincipal -ServicePrincipalName 00000002-0000-0ff1-ce00-000000000000
```

If it isn't registered, run:

```
PowerShell
New-AzADServicePrincipal -ApplicationId 00000002-0000-0ff1-ce00-000000000000
```

Configuration-specific setup steps

> **Important**
>
> The following steps differ depending on whether you're using **Azure Key Vault** or **Managed HSM**. Use the tabs below to select your configuration.

Azure Key Vault

Create a premium Azure Key Vault in each subscription

Prior to creating a key vault, perform the steps in Getting Started with Azure Key Vault. These steps guide you through installing and launching Azure PowerShell, and connecting to your subscription. Next, create a resource group and a key vault.

> **Important**
>
> When creating your key vaults, you **must** enable both **soft delete** and **purge protection** during the initial vault creation process. Customer Key requires all vaults to have both features enabled with a retention period of **90** days. These settings can't be disabled once enabled, so it's critical to configure them correctly from the start. If you plan to enable firewall on your key vaults, ensure you select the option for **Allow Trusted Microsoft Services**.

When you create a key vault, you must choose an SKU: either Standard or Premium. The Standard SKU uses software-protected keys without a Hardware Security Module (HSM), while the premium SKU allows the use of HSMs to protect keys. Customer Key supports key vaults with either of the two SKUs, but Microsoft strongly recommends using the Premium SKU. The cost of operations is the same for both; so, the only price difference comes from the monthly cost of each HSM-protected key. For pricing details, see Key Vault pricing.

> **Important**
>
> Use the Premium SKU key vaults and HSM-protected keys for production data. Use Standard SKU key vaults and keys only for testing and validation.

Enabling soft delete and purge protection

Soft delete allows you to recover deleted keys or vaults within 90 days. Purge protection prevents permanent deletion until the retention period passes. Customer Key requires both features with a **90-day** retention period. Soft delete is enabled by default for new vaults; enable it manually for existing vaults.

To enable these features, see Azure Key Vault recovery management with soft delete and purge protection.

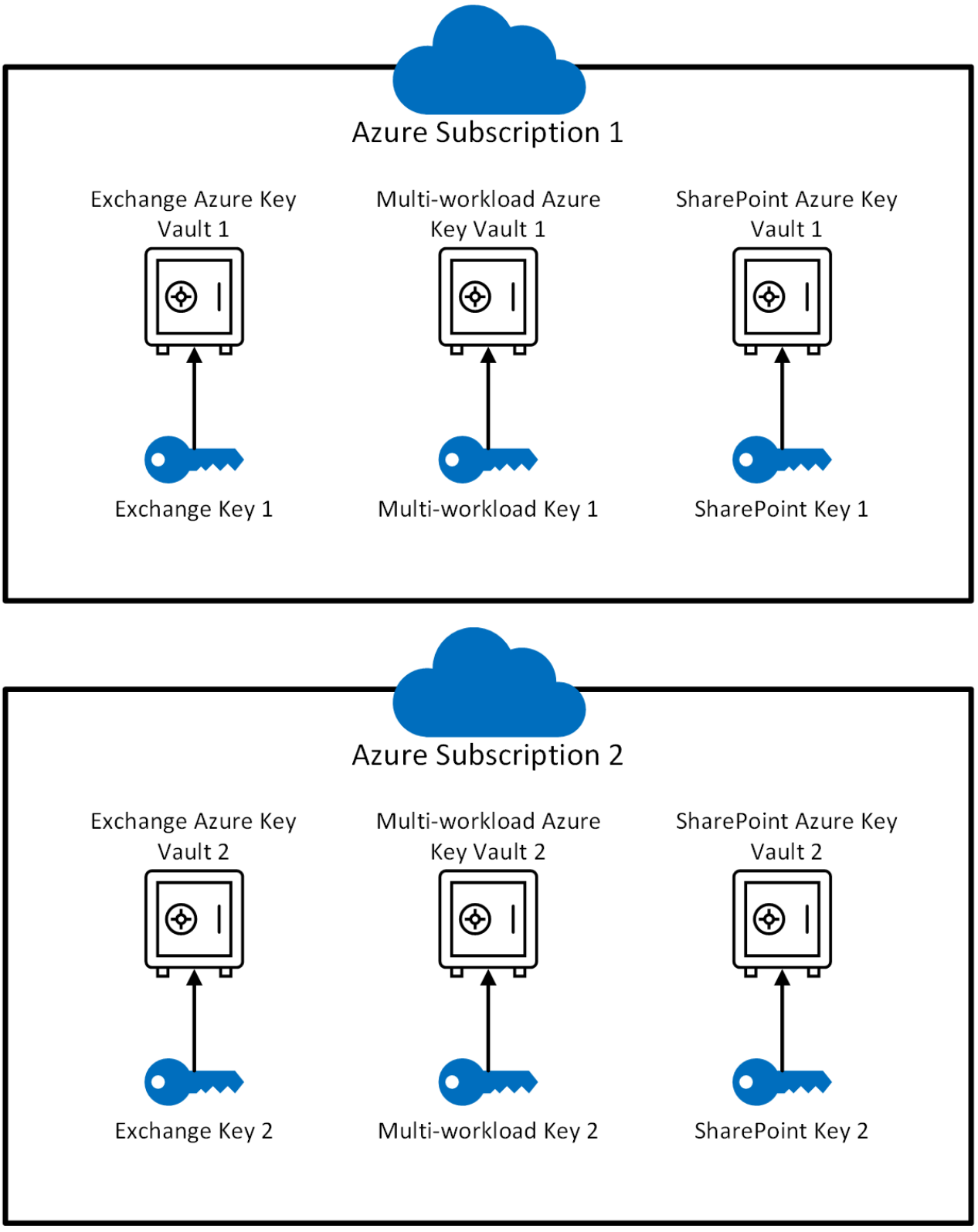
For each Microsoft 365 workload for which you use Customer Key, create a key vault in each of the two Azure subscriptions you set up earlier.

Key Vault configuration

For example, if you're using Customer Key for Multiple Workloads, Exchange and SharePoint scenarios, you need **three pairs** of key vaults, for a total of **six**. Use a clear naming convention that reflects the intended use of the DEP with which you associate the vaults. The following table shows how to map each Azure Key Vault to each workload.

Expand table

Key Vault Name	Permissions for Multiple Microsoft 365 Workloads (M365DataAtRestEncryption)	Permissions for Exchange	Permissions for SharePoint and OneDrive
ContosoM365AKV01	**Yes**	No	No
ContosoM365AKV02	**Yes**	No	No
ContosoEXOAKV01	No	**Yes**	No
ContosoEXOAKV02	No	**Yes**	No
ContosoSPOAKV01	No	No	**Yes**
ContosoSPOAKV02	No	No	**Yes**



Creating key vaults also requires setting up Azure resource groups. Key vaults require a small amount of storage, and logging (if enabled) also stores data. As a best practice, Microsoft recommends assigning different administrators to manage each resource group. These roles should align with the admins responsible for managing the associated Customer Key resources.

For Exchange, a DEP is scoped at the mailbox level. Each mailbox can have only one policy assigned, and you can create up to 50 policies. A SharePoint policy covers all data in an

organization's geographic location (or geo), while a multi-workload policy covers supported workloads across all users in the organization.

> **ⓘ Important**
>
> If you're using Customer Key for Multiple Workloads, Exchange, and SharePoint and OneDrive, make sure to create two Azure Key Vaults for each workload. This means you need a total of six key vaults.

Assign permissions to each Azure Key Vault

Assign the required permissions to each key vault using Azure role-based access control (Azure RBAC) in the Azure portal. This section explains how to apply the correct permissions using RBAC.

Assigning permissions using RBAC method

To assign (`wrapKey`, `unwrapKey`, and `get`) on your **Azure Key Vault**, assign the **Key Vault Crypto Service Encryption User** role to the appropriate Microsoft 365 app. See Grant permission to applications to access an Azure Key Vault using Azure RBAC.

When assigning the role, search for the following app names in your tenant:

- **Multiple Workloads**: `M365DataAtRestEncryption`
- **Exchange**: `Office 365 Exchange Online`
- **SharePoint and OneDrive**: `Office 365 SharePoint Online`

If you don't see the app you're looking for, make sure you register the app in the tenant.

For more information about assigning roles and permissions, see Use role-based access control to manage access to your Azure subscription resources.

Assigning user roles

Customer Key requires both **Key Vault Administrators** (for daily key management tasks) and **Key Vault Contributors** (for permission and role assignment). By design, Administrators can't delete keys to prevent data loss.

For detailed steps, see Azure built-in roles for Key Vault data plane operations.

Add a key to each Key Vault either by creating or importing a key

There are two ways to add keys to an Azure Key Vault: you can create a key directly in the vault or import an existing key. Creating a key in Azure is simpler, but importing a key gives you full control over how the key is generated. Use RSA keys. Customer Key supports RSA key lengths up to 4,096 bits. Azure Key Vault doesn't support wrapping and unwrapping with elliptical curve (EC) keys.

To add a key to each vault, see Add-AzKeyVaultKey.

If you prefer to generate a key on-premises and then import it into Azure, perform the steps in How to generate and transfer HSM-protected keys for Azure Key Vault. Use the Azure instructions to add a key to each Key Vault.

Verify expiration date of your keys

To check that your keys don't have an expiration date, run the Get-AzKeyVaultKey cmdlet.

For Azure Key Vault:

PowerShell

```
Get-AzKeyVaultKey -VaultName <vault name>
```

Customer Key can't use expired keys. Any operation using an expired key fails and can cause service outages.

Once set, an expiration date can't be removed, but you can change it. If you must use a key with an expiration date, update it to `12/31/9999` and use the legacy onboarding method.

> **⊗ Caution**
>
> Don't set expiration dates on encryption keys used with Customer Key. The Customer Key Onboarding Service only accepts keys without an expiration date.

To change the expiration date, use the Update-AzKeyVaultKey cmdlet:

PowerShell

```
Update-AzKeyVaultKey -VaultName <vault name> -Name <key name> -Expires (Get-Date
-Date "12/31/9999")
```

Back up Azure Key Vault Key

After you create or modify a key, make sure to immediately back it up. Store backup copies in both online and offline locations to help prevent data loss.

To back up a key in Azure Key Vault, use the Backup-AzKeyVaultKey cmdlet.

> **Important**
>
> If a key is deleted without a backup, it can't be recovered. Always create a backup after any key change or creation.

Obtain the URI for each Azure Key Vault key

After you set up your Key Vaults and add your keys, run the following command to get the URI for each key. You need these URIs when creating and assigning DEPs; so, be sure to save them somewhere secure.

Run the following command in Azure PowerShell—once for each Key Vault:

PowerShell

```
(Get-AzKeyVaultKey -VaultName <vault name>).Id
```

Onboard using the Customer Key Onboarding Service

The Microsoft 365 Customer Key Onboarding Service allows you to enable Customer Key in your tenant. This service automatically validates the required Customer Key resources. If you'd prefer, you can validate your resources separately before proceeding with enablement.

> **Important**
>
> This service isn't currently available for the following scenarios:
>
> - **Special clouds tenant**: See Onboard to Customer Key for special cloud environments.
> - **SharePoint and OneDrive**: See Onboard to Customer Key for SharePoint and OneDrive.

The account used for onboarding **must** have the following permissions:

1. Service principal registration permissions: To register the required service principals.
2. Reader role: On each Azure Key Vault used in the onboarding cmdlet.

Install the `M365CustomerKeyOnboarding` PowerShell module

1. Sign in to your Azure subscription with Azure PowerShell. For guidance, see Sign in with Azure PowerShell.

2. Install the **latest version** of the `M365CustomerKeyOnboarding` module from the PowerShell Gallery.

- To confirm you're using the latest version, check the **Version History** tab at the bottom of the module page.
- Copy and paste the install command into your session and run it.
- If you're prompted, select **Yes to All** to continue.

Using the two different onboarding modes

There are two different onboarding modes available when using the Customer Key Onboarding Service: **Validate**, and **Enable**. Each mode serves a different purpose in the onboarding process.

When running the cmdlet (as guided in Create an Onboarding request), specify the mode using the `-OnboardingMode` parameter.

Validate

Use the `Validate` mode to confirm that your Customer Key resource configuration is correct. This mode doesn't make any changes to your environment.

> **Important**
>
> You can run this mode as many times as needed, especially after making changes to your configuration.

Enable

Use the `Enable` mode when you're ready to onboard your tenant to Customer Key. This mode enables Customer Key for the workload you specify using the `-Scenario` parameter.

If you want to enable Customer Key for both **Multiple Workloads** and **Exchange**, run the cmdlet twice, once for each workload.

Before running in Enable mode, make sure your validation results show **Passed** under `ValidationResult`.

> ⓘ **Important**
>
> Your resources must pass all checks in `Validate` mode for the onboarding process to succeed in Enable mode.

Creating an onboarding request

The first step in the onboarding process is to create a new request. In PowerShell, you can store the results of a cmdlet in a variable using the `$` symbol followed by the variable name.

In this example, the onboarding request is stored in a variable named `$request`:

PowerShell

```
$request = New-CustomerKeyOnboardingRequest -Organization <tenantID> -Scenario
<Scenario> -Subscription1 <subscriptionID1> -KeyIdentifier1 <KeyURI1> -Subscription2
<subscriptionID2> -KeyIdentifier2 <KeyURI2> -OnboardingMode <OnboardingMode>
```

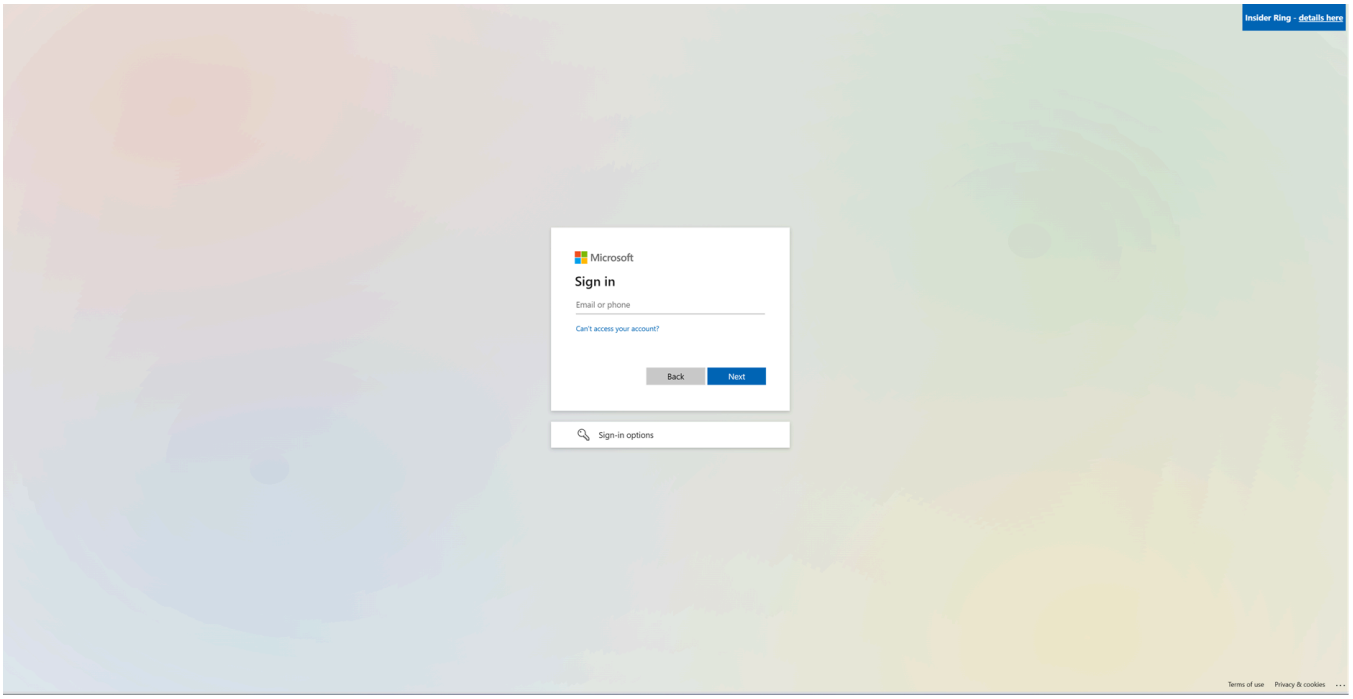
Expand table

Parameter	Description	Example
`-Organization`	Your tenant ID (GUID format)	`abcd1234-abcd-efgh-hijk-abcdef123456`
`-Scenario`	Workload: `MDEP` (Multiple Workloads) or `EXO` (Exchange)	`MDEP` or `EXO`
`-Subscription1`	First Azure subscription ID	`p121d534-1234-5678-9876-g3def67j9k12`
`-KeyIdentifier1`	Key URI for first AKV/HSM key	`https://exampleVault1.vault.azure.net/keys/customerKey1`
`-Subscription2`	Second Azure subscription ID	`21k9j76f-ed3g-6789-8765-43215dl21pbd`

Parameter	Description	Example
`-KeyIdentifier2`	Key URI for second AKV/HSM key	`https://exampleVault2.vault.azure.net/keys/customerKey2`
`-OnboardingMode`	`Validate` (test only) or `Enable` (test and enable)	`Validate` or `Enable`

Sign in with your tenant admin credentials

When prompted, a browser window opens. Sign in using your tenant admin account with the necessary privileges to complete onboarding.



View validation and enablement details

After successfully signing in, return to your PowerShell window. Run the variable you used when creating the onboarding request to view its output:

PowerShell

```
$request
```

```
PS C:\Windows\System32> $request

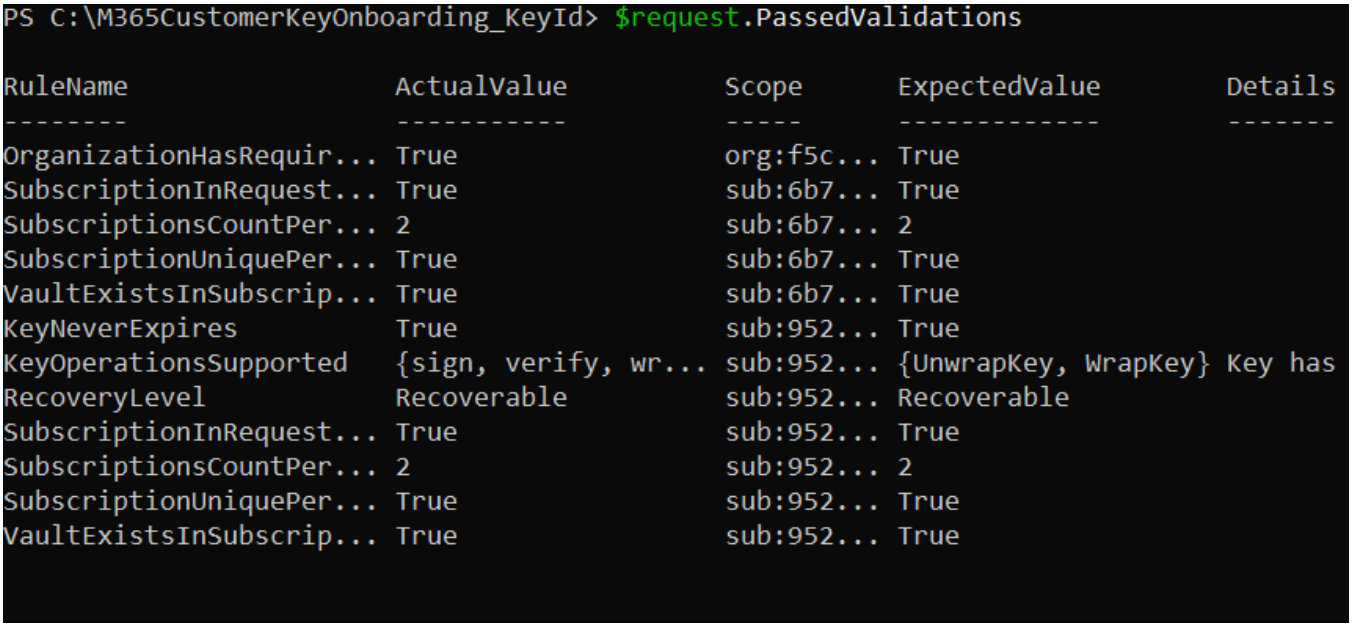
Id                                   CreatedDate           ValidationResult EnablementResult
--                                   -----------           ---------------- ----------------
94a83419-edb8-4bf1-8037-cc54ceb71594 11/22/2023 12:07:22 AM Passed           NotAttempted
```

You receive an output that includes `ID`, `CreatedDate`, `ValidationResult`, and `EnablementResult`.

Expand table

Output	Description
`ID`	ID associated with the created onboarding request.
`CreatedDate`	Date on which the request was created.
`ValidationResult`	Indicator of successful/unsuccessful validation.
`EnablementResult`	Indicator of successful/unsuccessful enablement.

A tenant that's ready to use Customer Key shows **Success** for both `ValidationResult` and `EnablementResult` as shown in the following screenshot:



If both values show **Success**, proceed to Next Steps.

Troubleshooting failed validations details

If validation fails during onboarding, perform the following steps to investigate the cause. This process helps identify which Customer Key resources are misconfigured.

1. List all onboarding requests for your tenant to find the `RequestID` of the one you want to investigate:

PowerShell

```
Get-CustomerKeyOnboardingRequest -OrganizationID <OrganizationID>
```

2. Store the specific onboarding request into a variable:

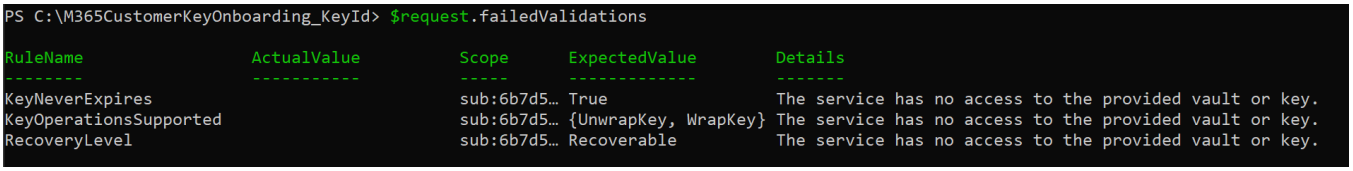
PowerShell

```
$request = Get-CustomerKeyOnboardingRequest -OrganizationID <OrganizationID> -
RequestID <RequestID>
```

3. View the failed validation details:

PowerShell

```
$request.FailedValidations
```



Each validation rule includes the following fields:

- **ExpectedValue**: What the resource configuration should be.
- **ActualValue**: What was detected in your environment.
- **Scope**: The first five characters of the subscription ID that helps you identify which subscription (and its associated Key Vault) has the issue.
- **Details**: Describes the root cause and offers guidance to resolve the problem.

The following table summarizes common validation rules and how to resolve failures:

Expand table

RuleName	Description	Solution
`OrganizationHasRequiredServicePlan`	Verifies whether your organization has the required licenses.	See Prerequisites.
`KeyNeverExpires`	Ensures that the key has no expiration date.	See the key expiration verification steps in Configuration-specific setup
`KeyOperationsSupported`	Verifies whether the key supports required operations for the workload.	See the permissions assignment steps in Configuration-specific setup
`RecoveryLevel`	Checks if the key recovery level is `Recoverable`.	Ensure both soft delete and purge protection are enabled. See Enabling Soft Delete and Purge Protection for Azure Key Vault or Create a resource group

RuleName	Description	Solution
		provision, and activate a Managed HSM for Managed HSM.
`SubscriptionInRequestOrganization`	Confirms that your Azure subscription belongs to your organization.	Make sure the subscription was created within the specified tenant.
`SubscriptionsCountPerScenario`	Confirms that two subscriptions were provided.	Make sure your onboarding request includes **two** subscriptions.
`SubscriptionUniquePerScenario`	Ensures you're using two unique subscriptions per scenario.	Double-check that both subscription IDs are different.
`VaultExistsinSubscription`	Confirms that the AKV/Managed HSM is part of the given subscription.	Verify whether the Key Vault/HSM was created in the correct subscription.

Checking passed validations

To view which validations were successful during the onboarding process, perform the following steps:

1. Store the specific onboarding request into the variable '$request'

PowerShell

```
$request = Get-CustomerKeyOnboardingRequest -OrganizationID <OrganizationID> -
RequestID <RequestID>
```

2. Run the following command to display all passed validations:

PowerShell

```
$request.PassedValidations
```

This command returns a list of all the validations that passed successfully for the selected onboarding request.

```
PS C:\M365CustomerKeyOnboarding_KeyId> $request.PassedValidations

RuleName                 ActualValue          Scope      ExpectedValue        Details
--------                 -----------          -----      -------------        -------
OrganizationHasRequired… True                 org:f5cfc… True
SubscriptionInRequestOr… True                 sub:6b7d5… True
SubscriptionsCountPerSc… 2                    sub:6b7d5… 2
SubscriptionUniquePerSc… True                 sub:6b7d5… True
VaultExistsInSubscripti… True                 sub:6b7d5… True
KeyNeverExpires          True                 sub:952b8… True
KeyOperationsSupported   {sign, verify, wrap… sub:952b8… {UnwrapKey, WrapKey} Key has more ope…
RecoveryLevel            Recoverable          sub:952b8… Recoverable
SubscriptionInRequestOr… True                 sub:952b8… True
SubscriptionsCountPerSc… 2                    sub:952b8… 2
SubscriptionUniquePerSc… True                 sub:952b8… True
VaultExistsInSubscripti… True                 sub:952b8… True
```

Onboard to Customer Key for special cloud environments

If your tenant is located in **GCC-H**, **DoD**, or **M365 operated by 21Vianet**, complete all required Customer Key setup steps first.

- For **GCC-H** and **DoD** tenants, contact Microsoft Support and request onboarding for government tenants.
- For customers in Microsoft 365 operated by 21Vianet (China), configure your Customer Key resources using Azure operated by 21Vianet portal. Then, contact Microsoft Support and request onboarding for your tenant.

Onboard to Customer Key for SharePoint and OneDrive

To onboard to Customer Key for SharePoint and OneDrive, your Azure Key Vaults must meet the following prerequisites:

1. Enable both soft delete and purge protection during the initial vault creation process. See Enabling Soft Delete and Purge Protection for Azure Key Vault.

If this prerequisite is satisfied, perform the steps in Create a DEP for use with SharePoint and OneDrive.

Next steps

After completing the setup steps in this article, you're ready to create and assign DEPs. For detailed instructions, see Manage Customer Key.

Related articles

- Customer Key Service encryption with Microsoft Purview Customer Key
- Manage Customer Key
- Roll or rotate a Customer Key or an availability key
- Learn about the availability key
- Service encryption and key management

Last updated on 01/15/2026

Manage Customer Key encryption policies

> *Microsoft Purview service description*

After setting up Customer Key, create and assign one or more **data encryption policies (DEPs)** to control how Microsoft 365 encrypts your organization's data at rest. After you assign them, you can manage your encryption keys and policies. For example, you can verify encryption status, view DEP details, disable a DEP, restore keys, or roll back to Microsoft-managed keys.

Microsoft Purview Customer Key also supports **Windows 365 Cloud PCs**. For more information, see Microsoft Purview Customer Key for Windows 365 Cloud PCs.

Prerequisites

Before managing Customer Key, make sure you have:

- **Completed setup**: Finish all steps in Set up Customer Key
- **Key Vault URIs**: Get the URIs for your Azure Key Vault keys. See Obtain the URI for each Azure Key Vault key
- **Permissions**: Exchange admin permissions (all DEP operations run through Exchange Online PowerShell)
- **PowerShell modules**:
  - Exchange Online PowerShell for DEP operations
  - Azure PowerShell for Key Vault operations

> **ⓘ Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

Quick reference: Commands by task

Use the following tables to quickly find the command you need.

Create and assign DEPs

Expand table

Task	Workload	Command
Create a DEP	Multi-workload	`New-M365DataAtRestEncryptionPolicy -Name <Name> -AzureKeyIDs <KeyURI1>,<KeyURI2>`
Create a DEP	Exchange	`New-DataEncryptionPolicy -Name <Name> -AzureKeyIDs <KeyURI1>, <KeyURI2>`
Create a DEP	SharePoint/OneDrive	`Register-SPODataEncryptionPolicy -PrimaryKeyVaultName <Name> -PrimaryKeyName <Name> -PrimaryKeyVersion <Version> -SecondaryKeyVaultName <Name> -SecondaryKeyName <Name> -SecondaryKeyVersion <Version>`
Assign a DEP	Multi-workload	`Set-M365DataAtRestEncryptionPolicyAssignment -DataEncryptionPolicy <PolicyName>`
Assign a DEP	Exchange mailbox	`Set-Mailbox -Identity <Mailbox> -DataEncryptionPolicy <PolicyName>`
Assign a DEP	Hybrid mailbox	`Set-MailUser -Identity <MailUser> -DataEncryptionPolicy <PolicyName>`

View and verify DEPs

Expand table

Task	Workload	Command
List all DEPs	Multi-workload	`Get-M365DataAtRestEncryptionPolicy`
List all DEPs	Exchange	`Get-DataEncryptionPolicy`
View DEP assignment	Multi-workload	`Get-M365DataAtRestEncryptionPolicyAssignment`
View mailbox DEP	Exchange	`Get-MailboxStatistics -Identity <Mailbox> | fl DataEncryptionPolicyID`
View encryption status	SharePoint/OneDrive	`Get-SPODataEncryptionPolicy`
Check mailbox encryption	Exchange	`Get-MailboxStatistics -Identity <Mailbox> | fl IsEncrypted`

Manage and disable DEPs

Expand table

Task	Workload	Command
Disable a DEP	Multi-workload	`Set-M365DataAtRestEncryptionPolicy -Identity <Policy> -Enabled $false`
Unassign a DEP	Exchange	`Set-Mailbox -Identity <Mailbox> -DataEncryptionPolicy $null`

Key Vault operations

Expand table

Task	Command
View Key Vault permissions	`Get-AzKeyVault -VaultName <VaultName>`
Remove user access	`Remove-AzKeyVaultAccessPolicy -VaultName <VaultName> -UserPrincipalName <UPN>`
Restore a key	`Restore-AzKeyVaultKey -VaultName <VaultName> -InputFile <BackupFile>`

Create a multi-workload DEP

To create a multi-workload DEP, follow these steps:

1. On your local computer, sign in using a **work or school account** with **Compliance Administrator** permissions, and connect to Exchange Online PowerShell.

2. Run the following cmdlet to create a multi-workload data encryption policy:

   PowerShell

   ```
   New-M365DataAtRestEncryptionPolicy -Name <PolicyName> -AzureKeyIDs <KeyVaultURI1, KeyVaultURI2> [-Description <String>]
   ```

   Parameter definitions:

   - `-Name`: The name you want to use for the policy. No spaces allowed.
   - `-AzureKeyIDs`: URIs of the two Azure Key Vault keys used in the policy, separated by commas.

Example: `"https://contosoCentralUSvault1.vault.azure.net/keys/Key_02"` `"https://contosoWestUSvault1.vault.azure.net/keys/Key_01"`

- `Description` *(optional)*: A human-readable description of the policy.

  Example: `"Policy for multiple workloads for all users in the tenant."`

PowerShell

```
New-M365DataAtRestEncryptionPolicy -Name "Contoso_Global" -AzureKeyIDs "https://contosoWestUSvault1.vault.azure.net/keys/Key_01","https://contosoCentralUSvault1.vault.azure.net/keys/Key_02" -Description "Policy for multiple workloads for all users in the tenant."
```

Assign multi-workload policy

After creating a multi-workload data encryption policy (DEP), assign it using the `Set-M365DataAtRestEncryptionPolicyAssignment` cmdlet. After assignment, Microsoft 365 encrypts your organization's data using the keys specified in the DEP.

PowerShell

```
Set-M365DataAtRestEncryptionPolicyAssignment -DataEncryptionPolicy <PolicyName or ID>
```

Replace `<PolicyName or ID>` with the name or GUID of the policy.

Example:

PowerShell

```
Set-M365DataAtRestEncryptionPolicyAssignment -DataEncryptionPolicy "Contoso_Global"
```

Windows 365 Cloud PCs:

After assigning the policy, allow 3–4 hours for the Intune admin center to reflect the update. After the update, follow the steps in the admin center to encrypt existing Cloud PCs.

For more information, see Set up Customer Keys for your Windows 365 Cloud PCs.

Create a DEP for Exchange mailboxes

Before you begin, complete the required setup steps. For details, see Set up Customer Key. To create the DEP, you'll need the Azure Key Vault URIs obtained during setup. For more

information, see Obtain the URI for each Azure Key Vault key.

To create a DEP for Exchange mailboxes, follow these steps:

1. On your local computer, using a work or school account that has Exchange admin permissions, connect to Exchange Online PowerShell.

2. Create a DEP using the `New-DataEncryptionPolicy` cmdlet:

PowerShell

```
New-DataEncryptionPolicy -Name <PolicyName> -Description "Policy Description" -
AzureKeyIDs <KeyVaultURI1>, <KeyVaultURI2>
```

Expand table

Parameter	Description	Example
`-Name`	Enter a name for the policy. Names can't contain spaces.	`USA_mailboxes`
`-AzureKeyIDs`	Enter the URIs for two keys in separate Azure Key Vaults. Separate them with a comma and a space.	`https://contosoEastUSvault01.vault.azure.net/keys/USA_key_01`, `https://contosoEastUS2vault01.vault.azure.net/keys/USA_key_02`
`-Description`	A user-friendly description that helps identify the policy's purpose.	`"Root key for mailboxes in USA and its territories"`

Example:

PowerShell

```
New-DataEncryptionPolicy -Name USA_mailboxes -Description "Root key for mailboxes
in USA and its territories" -AzureKeyIDs
https://contoso_EastUSvault02.vault.azure.net/keys/USA_key_01,
https://contoso_CentralUSvault02.vault.azure.net/keys/USA_Key_02
```

For syntax and parameter information, see New-DataEncryptionPolicy.

Assign a DEP to a mailbox

Assign the DEP with the `Set-Mailbox` cmdlet. Once assigned, Microsoft 365 encrypts the mailbox with the keys in the DEP.

PowerShell

```
Set-Mailbox -Identity <MailboxIdParameter> -DataEncryptionPolicy <PolicyName>
```

For more information, see Set-Mailbox.

Assign a DEP to hybrid mailboxes

In hybrid environments, assign a DEP to synchronized on-premises mailbox data using `Set-MailUser`:

PowerShell

```
Set-MailUser -Identity <MailUserIdParameter> -DataEncryptionPolicy <PolicyName>
```

For more information, see Set-MailUser and on-premises mailboxes using Outlook for iOS and Android with hybrid Modern Authentication.

Assign a DEP before you migrate a mailbox to the cloud

Assigning a DEP before migration encrypts your mailbox contents during the move, which is more efficient than assigning a DEP after migration.

1. Connect to Exchange Online PowerShell using a work or school account with Exchange Online cmdlet permissions.

2. Run the following command:

   PowerShell

   ```
   Set-MailUser -Identity <GeneralMailboxOrMailUserIdParameter> -
   DataEncryptionPolicy <DataEncryptionPolicyIdParameter>
   ```

View the DEPs you created for Exchange mailboxes

To list all DEPs in your organization, connect to Exchange Online PowerShell and run:

PowerShell

```
Get-DataEncryptionPolicy
```

For detailed cmdlet information, see Get-DataEncryptionPolicy.

Determine the DEP assigned to a mailbox

To find the DEP assigned to a mailbox, connect to Exchange Online PowerShell and run:

PowerShell

```
Get-Mailbox -Identity <MailboxIdParameter> | fl DataEncryptionPolicy
```

To get the friendly name from the returned GUID:

PowerShell

```
Get-DataEncryptionPolicy <GUID>
```

For more information, see Get-MailboxStatistics.

Create a DEP for use with SharePoint and OneDrive

Before you begin, make sure you complete the required setup steps. For more information, see Set up Customer Key.

To configure Customer Key for SharePoint and OneDrive, use SharePoint PowerShell.

A DEP uses two keys from separate Azure Key Vaults in different regions. Single-geo tenants need one DEP; multi-geo tenants need one DEP per geo with different keys. You need the Key Vault URIs for both keys (see Obtain the URI for each Azure Key Vault key).

1. On your local computer, using a work or school account that has Exchange Online cmdlet permissions in your organization, connect to SharePoint PowerShell.

2. Use the `Register-SPODataEncryptionPolicy` cmdlet to register the DEP.

   PowerShell

   ```
   Register-SPODataEncryptionPolicy -PrimaryKeyVaultName <PrimaryKeyVaultName> -
   PrimaryKeyName <PrimaryKeyName> -PrimaryKeyVersion <PrimaryKeyVersion> -
   ```

```
SecondaryKeyVaultName <SecondaryKeyVaultName> -SecondaryKeyName
<SecondaryKeyName> -SecondaryKeyVersion <SecondaryKeyVersion>
```

Example:

PowerShell

```
Register-SPODataEncryptionPolicy -PrimaryKeyVaultName 'stageRG3vault' -
PrimaryKeyName 'SPKey3' -PrimaryKeyVersion 'f635a23bd4a44b9996ff6aadd88d42ba' -
SecondaryKeyVaultName 'stageRG5vault' -SecondaryKeyName 'SPKey5' -
SecondaryKeyVersion '2b3e8f1d754f438dacdec1f0945f251a'
```

If using Managed HSM: Use the full URL of the key **including version** for each vault.

PowerShell

```
Register-SPODataEncryptionPolicy -PrimaryKeyVaultUri <PrimaryKeyVaultURL> -
SecondaryKeyVaultUri <SecondaryKeyVaultURL>
```

Example:

PowerShell

```
Register-SPODataEncryptionPolicy -PrimaryKeyVaultURL https://M365-
Test.managedhsm.azure.net/keys/Sharepoint-01/aaaa5513974f4780ea67b2f5d8c3dd -
SecondaryKeyVaultURL https://M365-Test-02.managedhsm.azure.net/keys/Sharepoint-
02/7d8f30343bed4e44a57225bae2012388
```

> ⓘ **Note**
>
> After the DEP is registered, encryption of the geo's data begins. This process might take some time.

For full cmdlet reference, see Register-SPODataEncryptionPolicy.

Verify Customer Key encryption status

After rolling a Customer Key, assigning a new data encryption policy (DEP), or migrating a mailbox, follow the steps in this section to confirm that encryption is complete.

Verify encryption completes for Exchange mailboxes

Encrypting a mailbox with Customer Key can take time. To verify that encryption is complete, determine the DEP assigned to the mailbox and confirm that the returned data encryption policy ID matches the policy you assigned.

Verify encryption completes for SharePoint and OneDrive

Check on the status of encryption by running the Get-SPODataEncryptionPolicy cmdlet as follows:

PowerShell

```
Get-SPODataEncryptionPolicy
```

The output includes primary/secondary key URIs, encryption status, and percentage of sites onboarded. Status values:

Expand table

Status	Description
Unregistered	Customer Key encryption not applied
Registering	Encryption in progress (shows percentage complete)
Registered	Encryption complete
Rolling	Key roll in progress (shows percentage complete)

Get details about DEPs you use with multiple workloads

To view multi-workload DEPs, connect to Exchange Online PowerShell with Compliance Administrator permissions and run:

PowerShell

```
Get-M365DataAtRestEncryptionPolicy
```

For details about a specific DEP:

PowerShell

```
Get-M365DataAtRestEncryptionPolicy -Identity "Contoso_Global"
```

Get multi-workload DEP assignment information

To identify which DEP is assigned to your tenant, connect to Exchange Online PowerShell and run:

PowerShell

```
Get-M365DataAtRestEncryptionPolicyAssignment
```

Disable a multi-workload DEP

Before disabling, unassign the DEP from workloads in your tenant. Then connect to Exchange Online PowerShell and run:

PowerShell

```
Set-M365DataAtRestEncryptionPolicy -Identity "PolicyName" -Enabled $false
```

Restore Azure Key Vault keys

Before restoring, try using soft delete to recover keys (available for up to 90 days). Full restoration is only needed if a key or vault is permanently lost.

To restore a key using Azure PowerShell, run the following command:

PowerShell

```
Restore-AzKeyVaultKey -VaultName <vault name> -InputFile <filename>
```

Example:

PowerShell

```
Restore-AzKeyVaultKey -VaultName Contoso-O365EX-NA-VaultA1 -InputFile Contoso-O365EX-NA-VaultA1-Key001-Backup-20170802.backup
```

> **Important**
>
> If the key vault already contains a key with the same name, the restore operation fails. The `Restore-AzKeyVaultKey` cmdlet restores all versions of the key, including its metadata and name.

Manage key vault permissions

Use Azure PowerShell to view or remove key vault permissions (for example, when a user leaves the team).

To view permissions:

```
PowerShell
Get-AzKeyVault -VaultName <vault name>
```

Example:

```
PowerShell
Get-AzKeyVault -VaultName Contoso-O365EX-NA-VaultA1
```

To remove a user's access:

```
PowerShell
Remove-AzKeyVaultAccessPolicy -VaultName <vault name> -UserPrincipalName <UPN of user>
```

Roll back from Customer Key to Microsoft-managed keys

You can revert to Microsoft-managed keys, which re-encrypts your data using the default encryption method for each workload.

> ⓘ **Important**
>
> Rolling back reverts encryption but doesn't delete data. A **data purge** permanently crypto-deletes data and can't be performed on multi-workload policies.

Roll back from Customer Key for multiple workloads

> ⓘ **Important**
>
> Rolling back from Customer Key for multiple workloads is **not** supported in Gallatin.

If you no longer want to use Customer Key with multi-workload data encryption policies (DEPs), submit a support request through Microsoft Support. Include the following information in your request:

- Tenant fully qualified domain name (FQDN)
- Tenant contact for the rollback request
- Reason for discontinuing Customer Key
- Incident number

> **Important**
>
> You must retain your Azure Key Vaults (AKVs) and encryption keys with the appropriate permissions. This action is required so that data can be rewrapped using Microsoft-managed keys.

Roll back from Customer Key for Exchange

> **Important**
>
> Rolling back from Customer Key for Exchange is **not** supported in Gallatin.

If you no longer want to encrypt individual mailboxes using mailbox-level data encryption policies (DEPs), you can unassign those DEPs from all mailboxes.

To unassign a mailbox-level DEP, follow these steps:

1. Use a work or school account that has the necessary Exchange Online PowerShell permissions, and connect to Exchange Online PowerShell.
2. Run the following cmdlet.

```
PowerShell
Set-Mailbox -Identity <mailbox> -DataEncryptionPolicy $null
```

This command unassigns the current DEP and re-encrypts the mailbox using the default Microsoft-managed keys.

> **Note**

You can't unassign the DEP associated with Microsoft-managed keys. If you don't want to use Microsoft-managed keys, assign a different DEP to the mailbox instead.

Roll back from Customer Key for SharePoint and OneDrive

> **Note**
>
> Roll back for Customer Key for SharePoint and OneDrive is currently in preview and is not available to tenants in GCC High, DoD, or M365 operated by 21Vianet.

If you no longer want to use Customer Key for SharePoint and OneDrive, submit a support request through Microsoft Support. Include the following information in your request:

- Tenant fully qualified domain name (FQDN)
- Tenant contact for the rollback request
- Reason for discontinuing Customer Key
- Incident number

> **Important**
>
> You must retain your Azure Key Vaults (AKVs) and encryption keys with the appropriate permissions. This action is required so that data can be rewrapped using Microsoft-managed keys.

Revoke your keys and start the data purge path

You control revocation of all root keys, including the availability key. When you revoke keys to purge data and exit the service, the availability key is deleted after the purge completes. This functionality is only supported for DEPs assigned to individual mailboxes.

Microsoft 365 audits and validates the data purge process. For more information, see the SSAE 18 SOC 2 Report, Risk Assessment and Compliance Guide for Financial Institutions, and Microsoft 365 Exit Planning Considerations.

> **Important**
>
> Purging a multi-workload DEP isn't supported. These policies encrypt data across multiple workloads and users within your tenant. Purging such a DEP would make data across all

workloads inaccessible.

If you're exiting Microsoft 365 services entirely, see how to delete a tenant in Microsoft Entra ID.

Revoke your Customer Keys and the availability key for Exchange

When you initiate the data purge path for Exchange, you place a permanent data purge request on a data encryption policy (DEP). This action permanently deletes encrypted data in all mailboxes assigned to that DEP.

Because the PowerShell cmdlet works on only one DEP at a time, consider reassigning a single DEP to all mailboxes before starting the data purge process.

> ⚠ **Warning**
>
> Don't use the data purge path to delete a subset of mailboxes. This process is intended only for organizations that are fully exiting the service.

> ⓘ **Important**
>
> Rolling back from Customer Key for Exchange, is **not** supported in Gallatin.

To initiate the data purge path, follow these steps:

1. Remove wrap and unwrap permissions for **O365 Exchange Online** from your Azure Key Vaults.
2. Use a work or school account with the appropriate Exchange Online PowerShell permissions, and connect to Exchange Online PowerShell.
3. For each DEP that includes mailboxes you want to purge, run the Set-DataEncryptionPolicy cmdlet:

   PowerShell

   ```
   Set-DataEncryptionPolicy <Policy ID> -PermanentDataPurgeRequested -
   PermanentDataPurgeReason <Reason> -PermanentDataPurgeContact <ContactName>
   ```

If the command fails, confirm that Exchange Online permissions were removed from both keys in the Azure Key Vault, as described in step 1. After you run the `Set-DataEncryptionPolicy` cmdlet with the `-PermanentDataPurgeRequested` switch, the DEP can't be assigned to mailboxes anymore.

4. Contact Microsoft Support and request the Data Purge eDocument.

   Microsoft sends a legal document to confirm and authorize the data purge. The person who signs this is usually an executive or other legally authorized representative.

5. After your representative signs the document, return it to Microsoft (usually via electronic signature).

   After Microsoft receives the signed eDocument, it verifies the authenticity of the signature and signee. After verification, they run the necessary cmdlets to complete the data purge. This process deletes the DEP, marks affected mailboxes for permanent deletion, and removes the availability key. After the process finishes, the data is purged, inaccessible to Exchange, and can't be recovered.

Revoke your Customer Keys and the availability key for SharePoint and OneDrive

Purging DEPs for SharePoint and OneDrive isn't supported in Customer Key. If you're exiting Microsoft 365 services entirely, you can follow the documented process to delete your tenant.

For details, see how to delete a tenant in Microsoft Entra ID.

Migrate your Key Vaults from the legacy access policy model to RBAC

If you onboarded to Customer Key using the legacy access policy model, follow these steps to migrate all Azure Key Vaults to use role-based access control (RBAC). To compare the two models and understand why Microsoft recommends RBAC, see Azure role-based access control (Azure RBAC) vs. access policies (legacy).

Remove legacy access policies

To remove the existing access policies from your Key Vaults, use the `Remove-AzKeyVaultAccessPolicy` cmdlet.

1. Sign in to your Azure subscription with Azure PowerShell. For guidance, see Sign in with Azure PowerShell.

2. Run the following command to remove access for the **Microsoft 365 Service** principal:

```
PowerShell

Remove-AzKeyVaultAccessPolicy -VaultName <VaultName> -ServicePrincipalName
c066d759-24ae-40e7-a56f-027002b5d3e4
```

3. Run the following command to remove access for the **Exchange Online** principal:

```
PowerShell

Remove-AzKeyVaultAccessPolicy -VaultName <VaultName> -ServicePrincipalName
00000002-0000-0ff1-ce00-000000000000
```

4. Run the following command to remove access for the **SharePoint and OneDrive for work or school service principal**:

```
PowerShell

Remove-AzKeyVaultAccessPolicy -VaultName <VaultName> -ServicePrincipalName
00000003-0000-0ff1-ce00-000000000000
```

Change the access configuration permission model

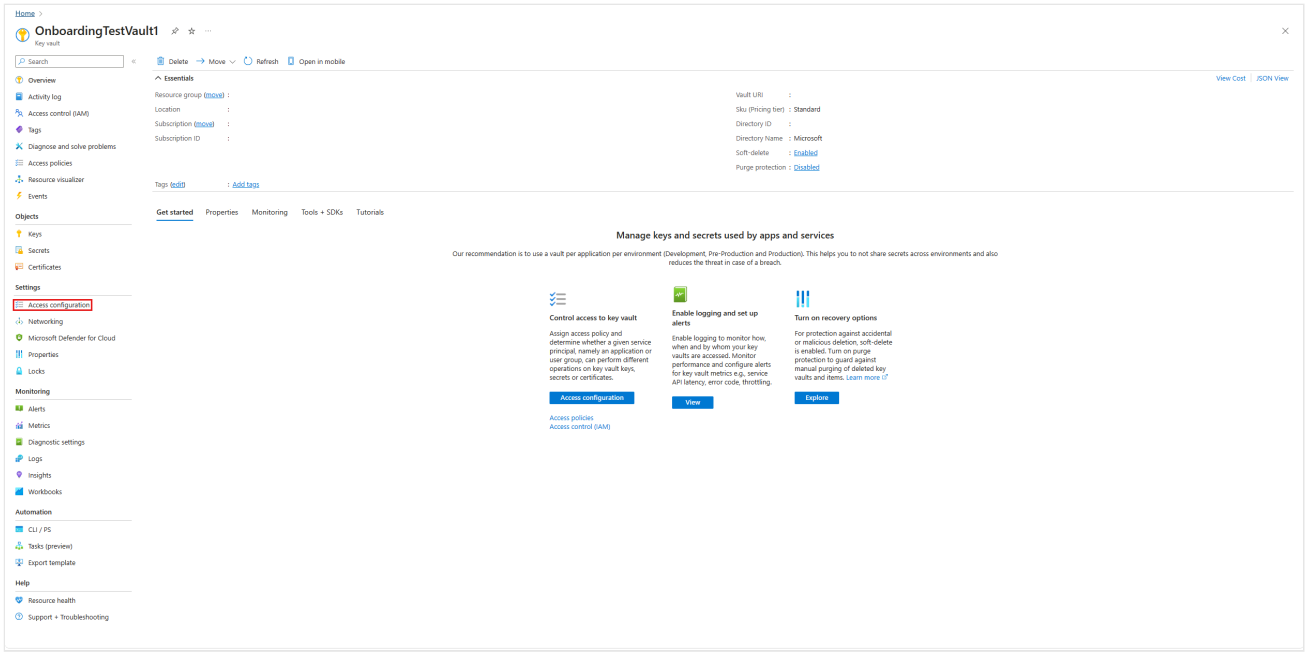
After removing the access policies, update the permission model for each Key Vault in the Azure portal:

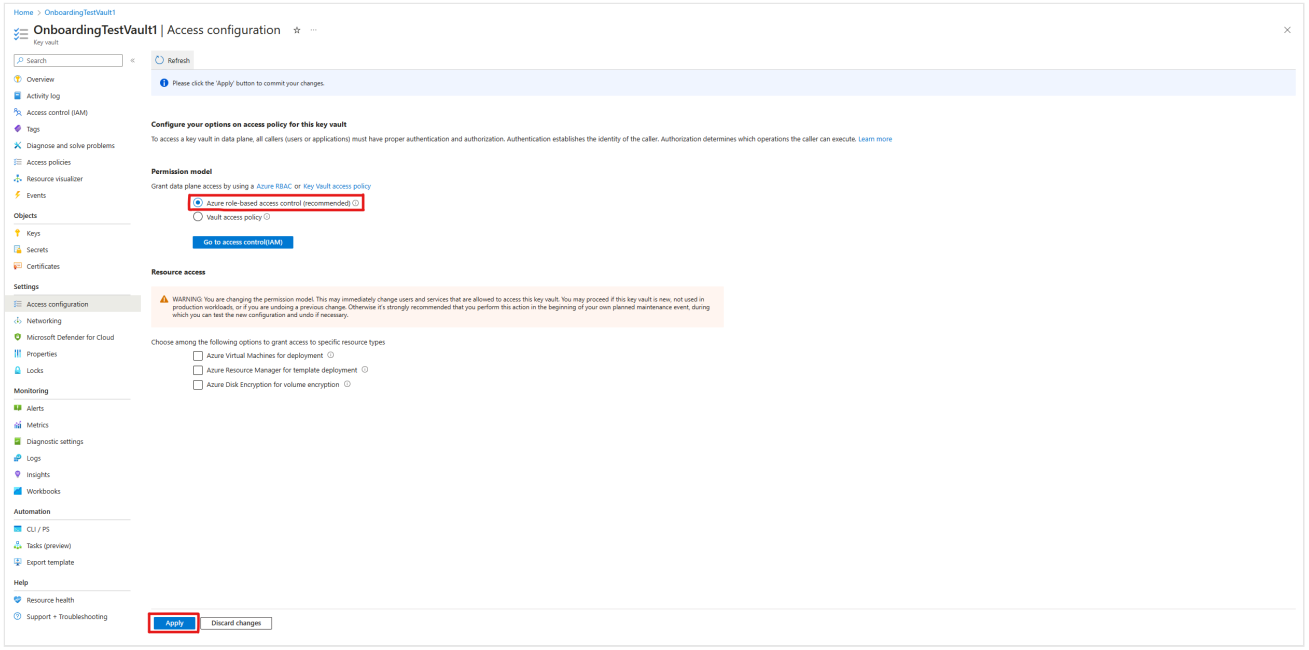
1. In the Azure portal, navigate to your Key Vault.

2. In the left-hand menu, under **Settings**, select **Access configuration**.

3. Under **Permission model**, choose **Azure role-based access control**.

4. Select **Apply** at the bottom of the screen.





Assign RBAC permissions

After you switch the permission model, follow the steps in Assign permissions to each Key Vault to grant the necessary roles.

Troubleshooting

The following table lists common Customer Key issues, possible causes, and recommended resolutions.

Expand table

Issue	Possible cause	Resolution
DEP creation fails	Invalid Key Vault URI	Verify the URI format and that the key exists in the vault
DEP assignment fails	Insufficient permissions	Ensure you have Exchange admin permissions
Mailbox not encrypted after one week	Database move not complete	Contact Microsoft Support
`Get-MailboxStatistics` returns no `DataEncryptionPolicyID`	DEP not assigned	Assign a DEP using `Set-Mailbox`
SharePoint encryption stuck at "Registering"	Large tenant or key access issue	Verify key permissions; for large tenants, encryption can take time
Key restore fails	Key with same name exists	Use a different vault or remove the existing key first
Can't connect to Exchange Online PowerShell	Module not installed or outdated	Install or update the Exchange Online PowerShell module

Related content

- Overview of Customer Key
- Learn about the availability key
- Set up Customer Key
- Roll or rotate a Customer Key or an availability key
- Customer Lockbox
- Service encryption and key management

ⓘ**Note:** The author created this article with assistance from AI. Learn more

Last updated on 06/22/2026

Roll or rotate a Customer Key or an availability key

> *Microsoft Purview service description*

> **⊗ Caution**
>
> Roll an encryption key used with Customer Key **only** if required by your organization's security or compliance policies.
>
> **Do not delete or disable any keys**, including older versions, that are or were associated with **Exchange**, **Multi-workload**, or **SharePoint** encryption policies.
>
> For example:
>
> - SharePoint retains backup content for restore and recovery, which might also rely on older keys.

Prerequisites

Before you roll or rotate keys, ensure you have:

- **Permissions**: Global Administrator or appropriate Exchange permissions
- **PowerShell modules installed**:
  - Azure PowerShell for Azure Key Vault operations
  - Exchange Online PowerShell to run the given cmdlets
- **Access to Azure Key Vault**: You must have permissions to create new key versions in your key vaults

> **ⓘ Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

Quick reference: Commands by workload

Use the following table to quickly find the command you need based on your workload and scenario.

Expand table

Workload	Scenario	Command
Multi-workload	Refresh DEP with new key version	`Set-M365DataAtRestEncryptionPolicy -Identity <Policy> -Refresh`
Multi-workload	Replace keys in DEP	`Set-M365DataAtRestEncryptionPolicy -Identity <Policy> -Replace -AzureKeyIDs <KeyURI1>,<KeyURI2>`
Exchange	Refresh DEP with new key version	`Set-DataEncryptionPolicy -Identity <Policy> -Refresh`
Exchange	Replace keys in DEP	`Set-DataEncryptionPolicy -Identity <Policy> -Replace -AzureKeyIDs <KeyURI1>,<KeyURI2>`
SharePoint/OneDrive	Roll/Rotate a key	`Update-SPODataEncryptionPolicy <SPOAdminSiteUrl> -KeyVaultName <Name> -KeyName <Name> -KeyVersion <Version> -KeyType <Primary | Secondary>`
SharePoint/OneDrive	Check roll progress	`Get-SPODataEncryptionPolicy <SPOAdminSiteUrl>`
Azure Key Vault	Create new key version	`Add-AzKeyVaultKey -VaultName <VaultName> -Name <KeyName> -Destination HSM -KeyOps @('wrapKey','unwrapKey')`

> **Note**
>
> For SharePoint and OneDrive, the same cmdlet, `Update-SPODataEncryptionPolicy`, is used to roll (or rotate) and replace keys.

Choose your key rolling method

There are two ways to roll customer-managed root keys. Use the following table to determine which method is right for your scenario.

Expand table

Method	When to use	What happens
Update existing key (create new version)	Routine key rotation; maintaining the same key name	Creates a new version of the existing key. The DEP references the same key name but uses the new version after refresh.
Replace with new key	Compromised key; changing key vault; organizational restructuring	Creates an entirely new key with a new name. The DEP is updated to reference the new key URIs.

About rolling the availability key

You can't directly roll the availability key. You can only roll keys you manage in Azure Key Vault.

Microsoft 365 rolls availability keys automatically on an internal schedule with no customer-facing SLA. Administrators don't have direct access to the key store. For more information, see Understand the availability key.

> **Important**
>
> For Exchange, creating a new DEP effectively rolls the availability key since each new data encryption policy (DEP) generates a unique availability key.
>
> For SharePoint and OneDrive, availability keys are created at the forest level and shared across DEPs. These keys roll only on Microsoft's internal schedule, but SharePoint, OneDrive, and Teams roll the tenant intermediate key (TIK) with each new DEP to mitigate this.

Roll customer-managed root keys

Use one of the following methods to roll customer-managed root keys stored in Azure Key Vault. Choose the method that matches your scenario based on the guidance in Choose your key rolling method.

Method 1: Update an existing key (create a new version)

Use this method to create a new version of an existing key. This approach is recommended for routine key rotations.

Step 1: Create a new key version in Azure Key Vault

To request a new version of an existing key, use the `Add-AzKeyVaultKey` cmdlet with the same key name you used when creating the original key.

1. Sign in to your Azure subscription with Azure PowerShell. For instructions, see Sign in with Azure PowerShell.

2. Run the `Add-AzKeyVaultKey` cmdlet:

   PowerShell

   ```
   Add-AzKeyVaultKey -VaultName <VaultName> -Name <KeyName> -Destination HSM -KeyOps @('wrapKey','unwrapKey') -NotBefore (Get-Date)
   ```

   Example:

   PowerShell

   ```
   Add-AzKeyVaultKey -VaultName Contoso-CK-EX-NA-VaultA1 -Name Contoso-CK-EX-NA-VaultA1-Key001 -Destination HSM -KeyOps @('wrapKey','unwrapKey') -NotBefore (Get-Date -Date "12/27/2016 12:01 AM")
   ```

   In this example, a key named Contoso-CK-EX-NA-VaultA1-Key001 already exists in the Contoso-CK-EX-NA-VaultA1 vault. The cmdlet creates a new version of the key. Previous versions are retained in the key's version history.

3. Repeat this step in each Azure Key Vault that contains keys you want to roll.

Step 2: Refresh the DEP to use the new key version

After creating a new key version, refresh the DEP to use it. Connect to Exchange Online PowerShell with appropriate Exchange cmdlet permissions, then run the appropriate cmdlet:

For Multi-workload DEPs:

PowerShell

```
Set-M365DataAtRestEncryptionPolicy -Identity <Policy> -Refresh
```

For Exchange DEPs:

PowerShell

```
Set-DataEncryptionPolicy -Identity <Policy> -Refresh
```

> ⓘ Note
>
> The `DataEncryptionPolicyID` property remains the same when refreshing with a new version of the same key. Refreshing the DEP with a new version of the same key doesn't rotate the availability key.

For SharePoint and OneDrive DEPs:

SharePoint supports rolling only one key at a time. Wait for the first operation to finish before rolling a second key.

PowerShell

```
Update-SPODataEncryptionPolicy <SPOAdminSiteUrl> -KeyVaultName <Name> -KeyName <Name>
-KeyVersion <Version> -KeyType <Primary | Secondary>
```

For Managed HSM keys, use `-KeyVaultURL` instead of `-KeyVaultName`.

To check progress: `Get-SPODataEncryptionPolicy <SPOAdminSiteUrl>`

Method 2: Replace keys with newly generated keys

Use this method to replace existing keys entirely with new keys.

> ⓘ Important
>
> Keep old keys enabled and accessible until the replacement process is fully completed.

Connect to Exchange Online PowerShell with the appropriate Exchange cmdlet permissions, then run the appropriate cmdlet:

For Exchange DEPs:

PowerShell

```
Set-DataEncryptionPolicy -Identity <Policy> -Replace -AzureKeyIDs <KeyURI1>,<KeyURI2>
```

For Multi-workload DEPs:

PowerShell

```
Set-M365DataAtRestEncryptionPolicy -Identity <Policy> -Replace -AzureKeyIDs <KeyURI1>,
<KeyURI2>
```

> **Note**
>
> The order of key URIs doesn't matter. To replace only one key, you must still provide both URIs. Wait **24 hours** before disabling old keys, and then an **additional 24-48 hours** before purging.

Verify key roll status

Use the following commands to verify the status of your key roll operations.

Expand table

Workload	Verification command
SharePoint/OneDrive	`Get-SPODataEncryptionPolicy <SPOAdminSiteUrl>`
Exchange	`Get-DataEncryptionPolicy -Identity <Policy>`
Multi-workload	`Get-M365DataAtRestEncryptionPolicy -Identity <Policy>`

Troubleshooting

The following table lists common issues you might encounter when rolling keys and how to resolve them.

Expand table

Issue	Possible cause	Resolution
Key roll fails with permission error	Insufficient Azure Key Vault permissions	Verify your account has `wrapKey` and `unwrapKey` permissions on the key vault
DEP refresh fails	New key version not yet propagated	Wait a few minutes and retry the refresh command
SharePoint roll shows "In Progress" for extended time	Large tenant with significant data	SharePoint key rolls can take time for large tenants. Continue monitoring with `Get-SPODataEncryptionPolicy`
Can't access encrypted	Old key was disabled	Re-enable the old key and wait 24-48 hours before

Issue	Possible cause	Resolution
content after key roll	too soon	disabling again

Related articles

For more information about Customer Key, see the following articles:

- Overview of Customer Key
- Set up Customer Key
- Manage Customer Key
- Learn about the availability key

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Learn about the availability key for Customer Key

The availability key is a root key automatically generated when you create a data encryption policy (DEP). Microsoft 365 stores and protects this key. Unlike the keys you manage in Azure Key Vault, you can't directly access it. Only Microsoft 365 service code can use it programmatically.

Its primary purpose is recovery from unexpected loss of your root keys. If you lose control of your keys, contact Microsoft Support to recover data using the availability key and migrate to a new DEP.

The availability key differs from Azure Key Vault keys in three ways:

- **Recovery**: Provides a break-glass option if both Azure Key Vault keys are lost
- **Defense in depth**: Logical separation of control and storage prevents total loss from a single failure
- **High availability**: Supports temporary Azure Key Vault outages for Exchange/Multi-Workload (SharePoint/OneDrive only uses it during explicit recovery requests)

You have sole authority to disable or destroy the availability key when leaving the service. For more information, see the Microsoft Trust Center.

Availability key uses

The availability key supports recovery if an attacker gains control of your key vault or mismanagement causes key loss. All services share these behaviors:

- **Recovery**: Decrypts data so you can re-encrypt with a new DEP
- **Programmatic only**: No direct admin access; only service code can use it
- **Not a substitute**: Your two Customer Keys remain the primary encryption roots

Service-specific behavior:

Exchange and Multi-workload

In addition to supporting recovery, the availability key ensures continued access during short Azure Key Vault outages for background operations such as anti-malware scans, eDiscovery, DLP, mailbox migrations, and indexing.

Fallback logic: If unwrap fails with one Customer Key, the service tries the second. If both fail with system errors, it falls back to the availability key. Access denied errors don't

trigger fallback for user actions.

Availability key security

Microsoft protects availability keys in access-controlled internal secret stores. Customers can't directly access the availability key. You can only roll keys you manage in Azure Key Vault. All operations occur through automated service code. For more information, see Roll or rotate a Customer Key or an availability key.

Availability key secret stores

Management operations require privilege escalation through Lockbox with justification, manager approval, and automatic time-bound revocation.

Expand table

Workload	Storage location
Exchange and Multi-workload	Exchange Active Directory secret store in tenant-specific containers
SharePoint and OneDrive	Internal secret store with SQL Database back end; keys wrapped with AES-256/HMAC and RSA-2048 certificates

Defense-in-depth

Microsoft uses defense-in-depth to protect availability keys:

- **Application isolation**: Only Microsoft 365 service code can use keys for encryption/decryption
- **Logical separation**: Customer Keys, availability keys, and customer data are stored in isolated locations
- **Access controls**: Engineers have no direct access to secret stores (see Administrative Access Controls)
- **Technical controls**: Block interactive sign-in to privileged service accounts
- **Monitoring**: Continuous intrusion detection, centralized logging, and alerts for unauthorized access attempts or baseline deviations

Use the availability key to recover from key loss

If you lose control of your Customer Keys, the availability key lets you decrypt affected data and re-encrypt it under a new DEP with new Customer Keys.

Exchange and Multi-workload

1. Create two new Customer Keys in separate Azure subscriptions.
2. Create a new DEP (Exchange or Multi-workload).
3. Assign the DEP to impacted mailboxes or tenant.
4. Allow background re-encryption (up to 72 hours). The availability key protects data during transition.

How the availability key is used

Exchange and Multi-workload

When you create a DEP, Microsoft 365 generates a DEP Key encrypted three times: once with each Customer Key and once with the availability key. Only encrypted versions are stored.

Decryption flow:

1. Decrypt the DEP Key using a Customer Key.
2. Use the DEP Key to decrypt the Mailbox/Workload Key.
3. Use that key to decrypt and access the data.

Availability key triggers

Exchange and Multi-workload

When Microsoft 365 needs a DEP Key:

1. Reads the DEP to identify both Customer Keys.
2. Randomly selects one key and requests Azure Key Vault to unwrap the DEP Key.
3. If that fails, tries the alternate key.
4. If both fail, evaluates failure type:

    - **System errors** (Key Vault unavailable, timeouts, network faults): Falls back to availability key.

- **Access denied** (key deleted, permissions removed): User request fails with error (no fallback).

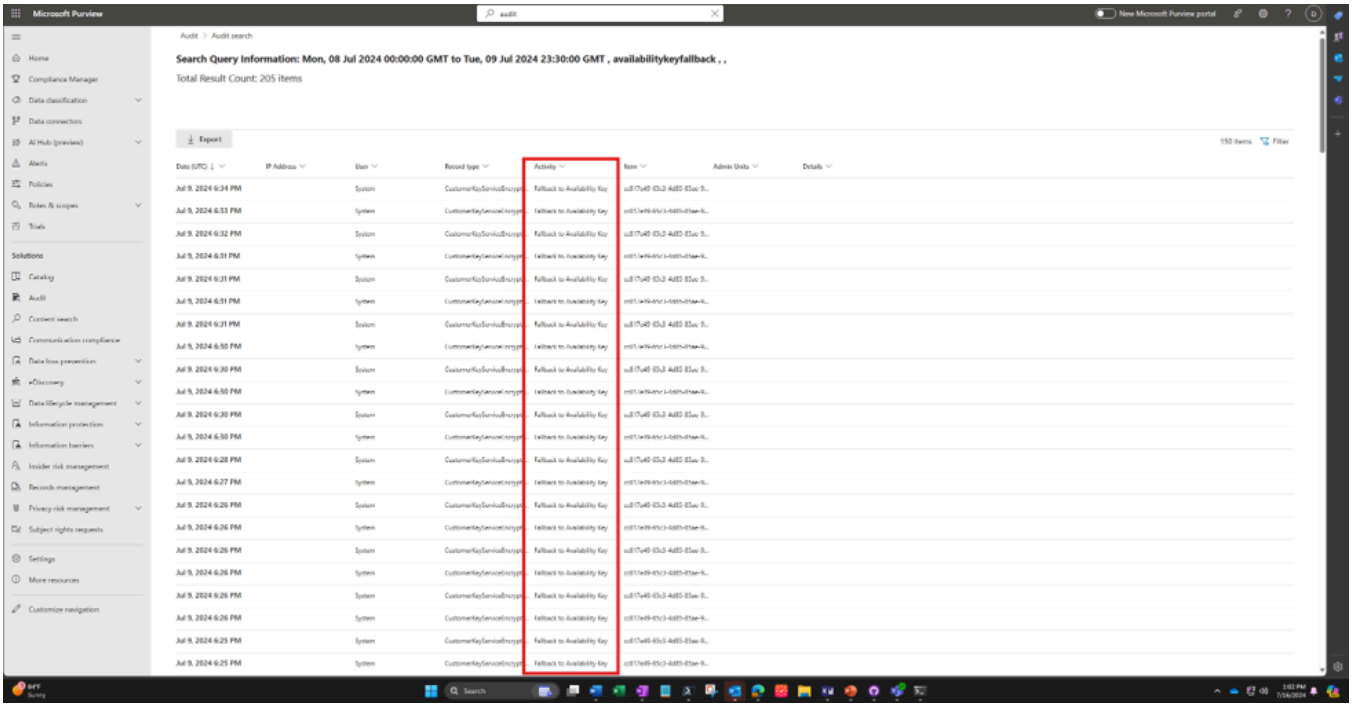
> **ⓘ Important**
>
> Internal operations (mailbox moves, indexing, anti-virus, eDiscovery, DLP) can still fall back to the availability key when both Customer Keys are unreachable, regardless of error type, until you delete the availability key.

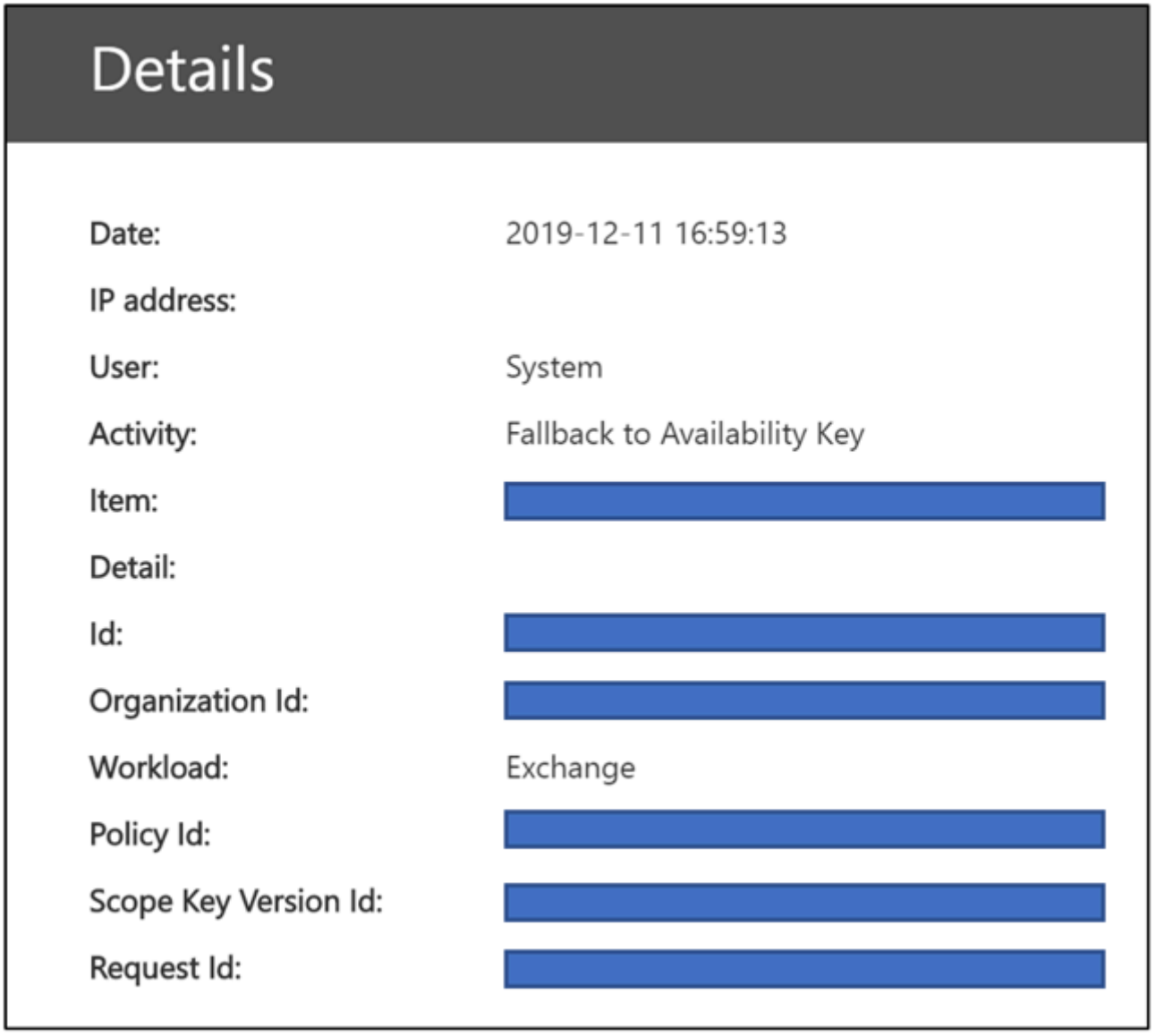
Audit logs and the availability key

Automated background processing doesn't generate customer-visible logs. When the availability key is used, fallback events create Unified Audit Log entries in the Microsoft Purview portal:

- **Record type**: CustomerKeyServiceEncryption
- **Activity**: Fallback to availability key
- **Workload field**: `Exchange` or `M365DataAtRestEncryption`

Fields include date, time, organization ID, DEP ID, Policy ID, and Request ID (schema details).





> **Note**
>
> SharePoint availability key usage isn't logged until the customer provides Lockbox approval.

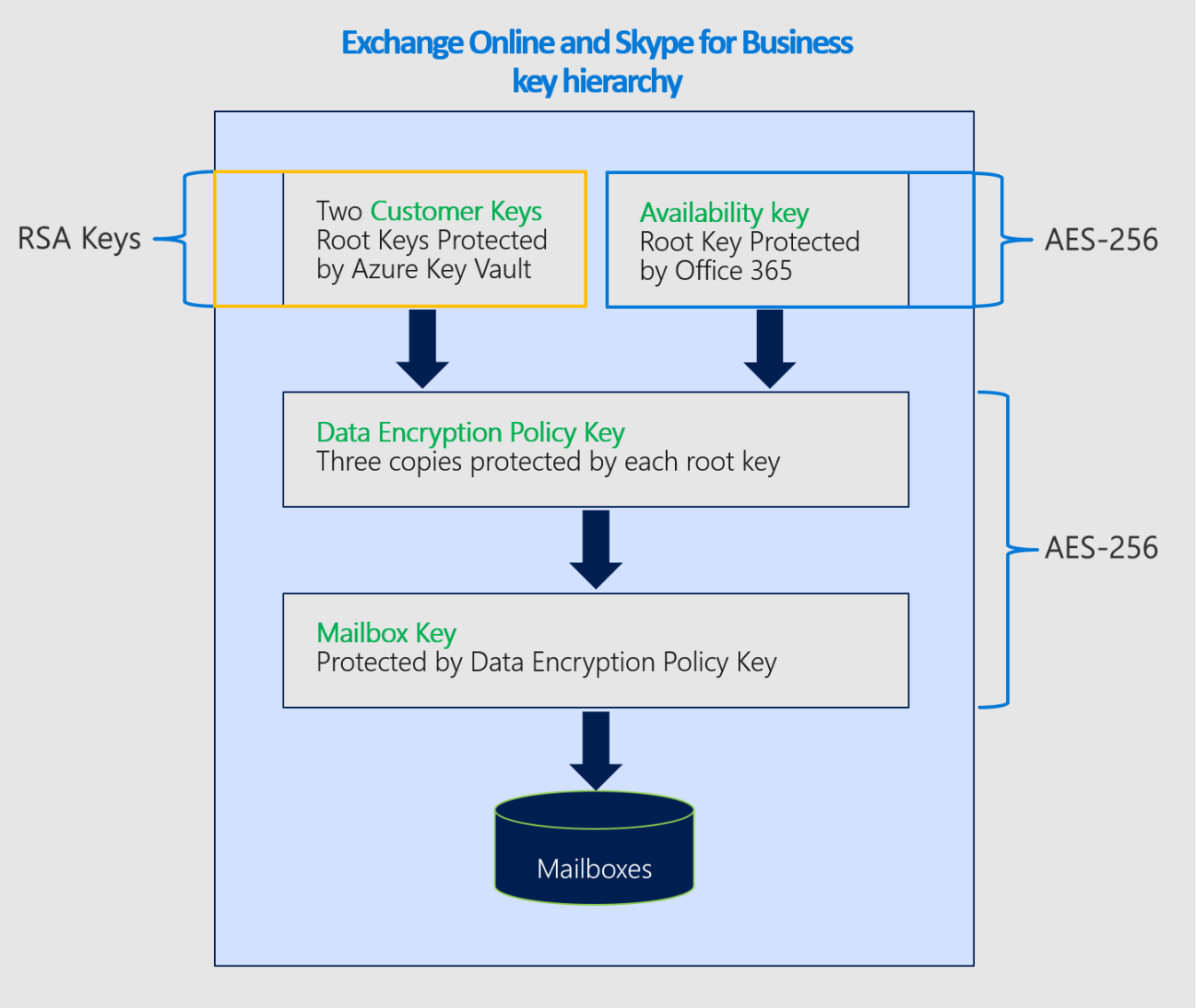
Availability key in the Customer Key hierarchy

The availability key wraps the tier of keys below it in the encryption hierarchy. Algorithms by service:

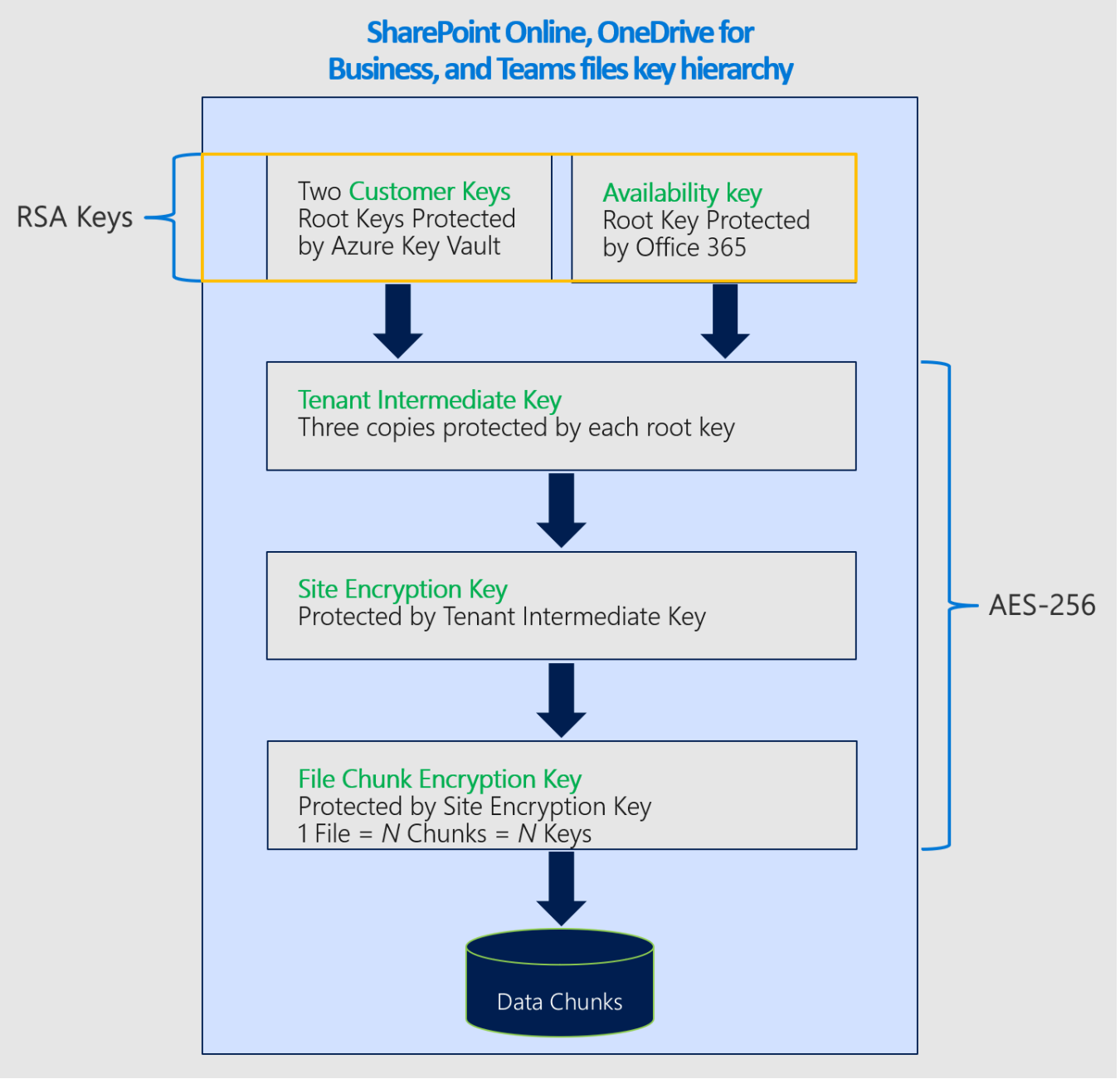
Expand table

Service	Algorithm
Exchange and Multi-workload	AES-256
SharePoint and OneDrive	RSA-2048

Encryption ciphers for Exchange



Encryption ciphers for SharePoint



Related articles

- Service encryption with Customer Key
- Set up Customer Key
- Manage Customer Key
- Roll or rotate a Customer Key or an availability key

Last updated on 01/15/2026

Microsoft Purview Customer Lockbox

Article • 02/03/2025

This article provides deployment and configuration guidance for Customer Lockbox. Customer Lockbox supports requests to access data in Exchange Online, SharePoint, OneDrive, Teams, and Windows 365. Additionally, all Microsoft 365 Copilot interactions are covered by Customer Lockbox through the support available for Exchange Online. To recommend support for other services, submit a request at Feedback Portal.

To see the options for licensing your users to benefit from Microsoft Purview offerings, see the Microsoft 365 licensing guidance for security & compliance.

Customer Lockbox ensures that Microsoft can't access your content to do service operations without your explicit approval. Customer Lockbox brings you into the approval workflow process that Microsoft uses to ensure only authorized requests allow access to your content. To learn more about Microsoft's workflow process, see Privileged access management.

Occasionally, Microsoft engineers help troubleshoot and fix issues that arise with the service. Usually, engineers fix issues using extensive telemetry and debugging tools Microsoft has in place for its services. However, some cases require a Microsoft engineer to access your content to determine the root cause and fix the issue. Customer Lockbox requires the engineer to request access from you as a final step in the approval workflow. This gives you the option to approve or deny the request for your organization, and provide direct-access control to your content.

> **Tip**
>
> If you're not an E5 customer, use the 90-day Microsoft Purview solutions trial to explore how additional Purview capabilities can help your organization manage data security and compliance needs. Start now at the **Microsoft Purview trials hub**. Learn details about **signing up and trial terms**.

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview **roles and permissions**.

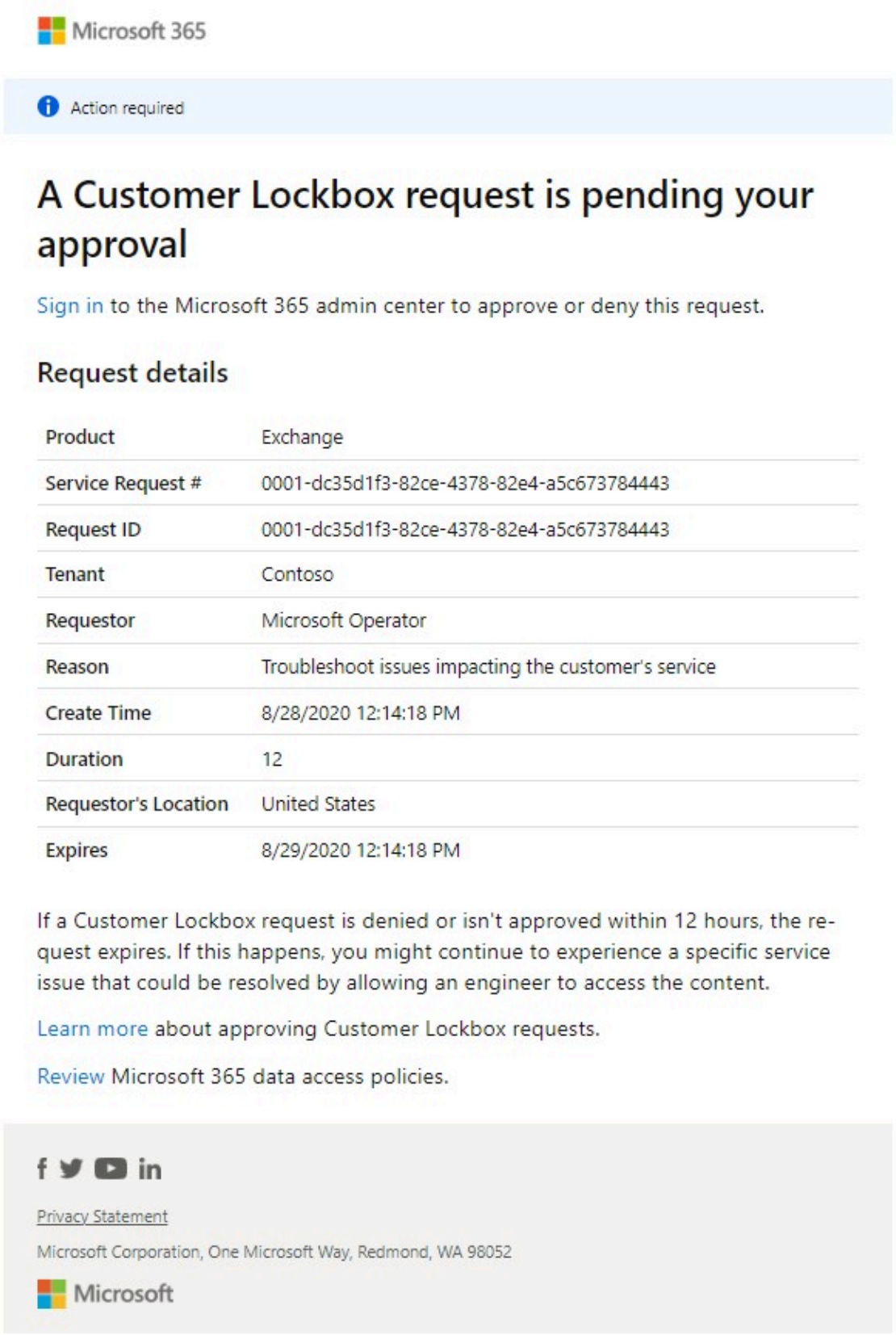
Customer Lockbox overview video

https://learn-video.azurefd.net/vod/player?id=016276d0-7e5d-43b5-8dee-8d4c6c542e23&locale=en-us&embedUrl=%2Fpurview%2Fcustomer-lockbox-requests

Customer Lockbox workflow

These steps outline the typical workflow when a Microsoft engineer starts a Customer Lockbox request:

1. Someone at an organization experiences an issue with their Microsoft 365 mailbox.
2. After the user troubleshoots the issue, but can't fix it, they open a support request with Microsoft Support.
3. A Microsoft support engineer reviews the service request and determines a need to access the organization's tenant to repair the issue.
4. The Microsoft support engineer logs into the Customer Lockbox request tool and makes a data access request that includes the organization's tenant name, service request number, expected start time of access (starts immediately post-approval if not specified), the estimated amount of time the engineer needs access to the data, and the service the request is for.
5. After a Microsoft Support manager approves the request, Customer Lockbox sends the designated approver at the organization an email notification about the pending access request from Microsoft.



Anyone who is assigned the Customer Lockbox access approver admin role in Microsoft 365 admin center can approve Customer Lockbox requests.

6. The approver signs in to the Microsoft 365 admin center and approves the request. This step triggers the creation of an audit record available by searching the audit log. For more information, see Auditing Customer Lockbox requests.

If the customer rejects the request or doesn't approve the request within 12 hours, the request expires and no access is granted to the Microsoft engineer.

> **Important**
>
> Microsoft does not include any links in Customer Lockbox email notifications requiring you to sign in to Office 365.

7. After the approver from the organization approves the request, the Microsoft engineer receives the approval message, logs into the tenant, and fixes the customer's issue. Microsoft engineers have the requested duration to fix the issue after which the access is automatically revoked.

> **Note**
>
> All actions performed by a Microsoft engineer are logged in the audit log. You can search for and review these audit records.

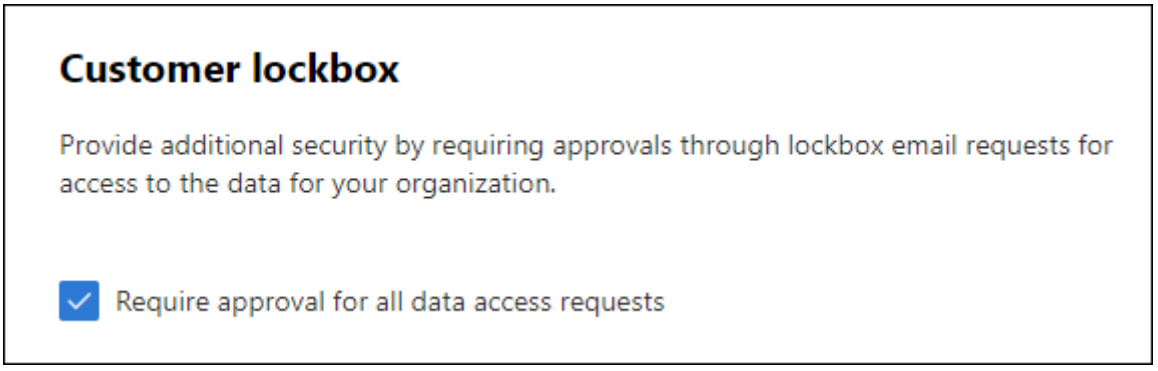
Turn Customer Lockbox requests on or off

You can turn on Customer Lockbox controls in the Microsoft 365 admin center. When you turn on Customer Lockbox, Microsoft must obtain your organization's approval before accessing any of your tenant's content.

1. Using a work or school account that has either the global administrator or the **Customer Lockbox access approver** role assigned, go to https://admin.microsoft.com and sign in.

2. Choose **Settings** > **Org Settings** > **Security & Privacy**.

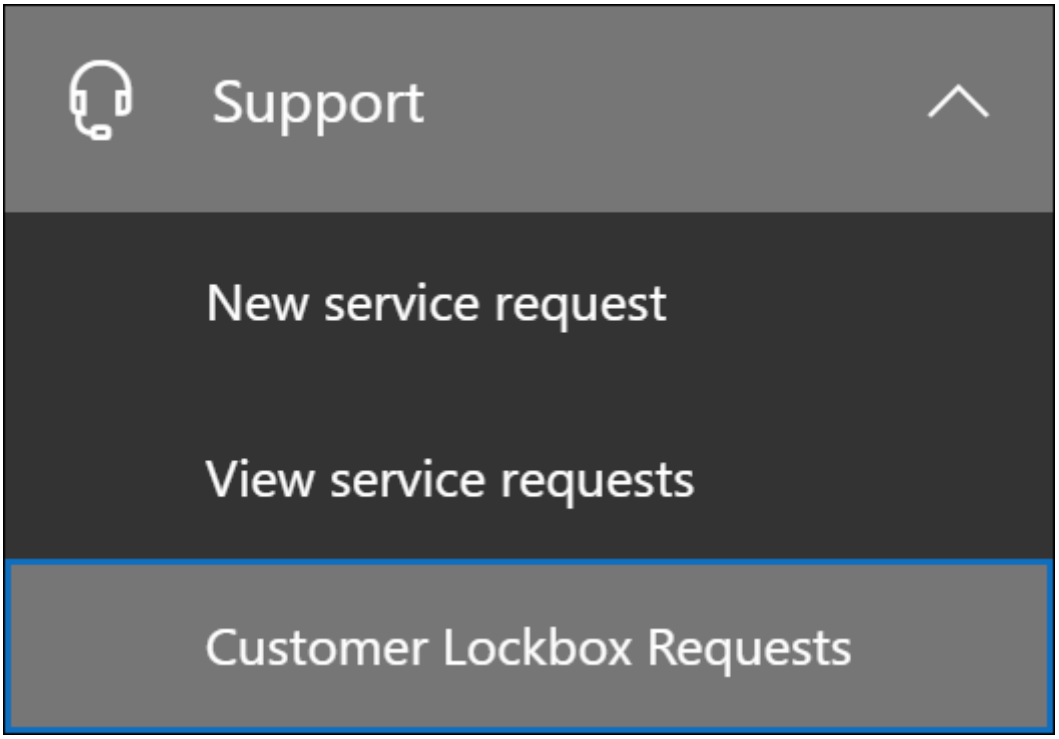
3. Select **Security & Privacy**, then select **Customer Lockbox** in the left column. Check the **Require approval for all data access requests** checkbox and save the changes to turn on the feature.



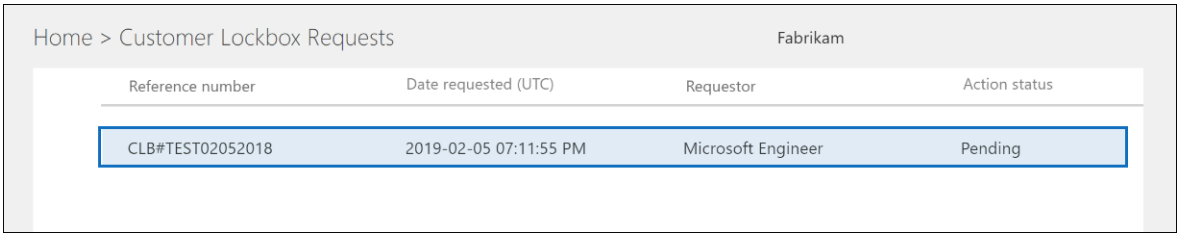
Approve or deny a Customer Lockbox request

1. Using a work or school account that has either the global administrator or the **Customer Lockbox access approver** role assigned, go to https://admin.microsoft.com and sign in.

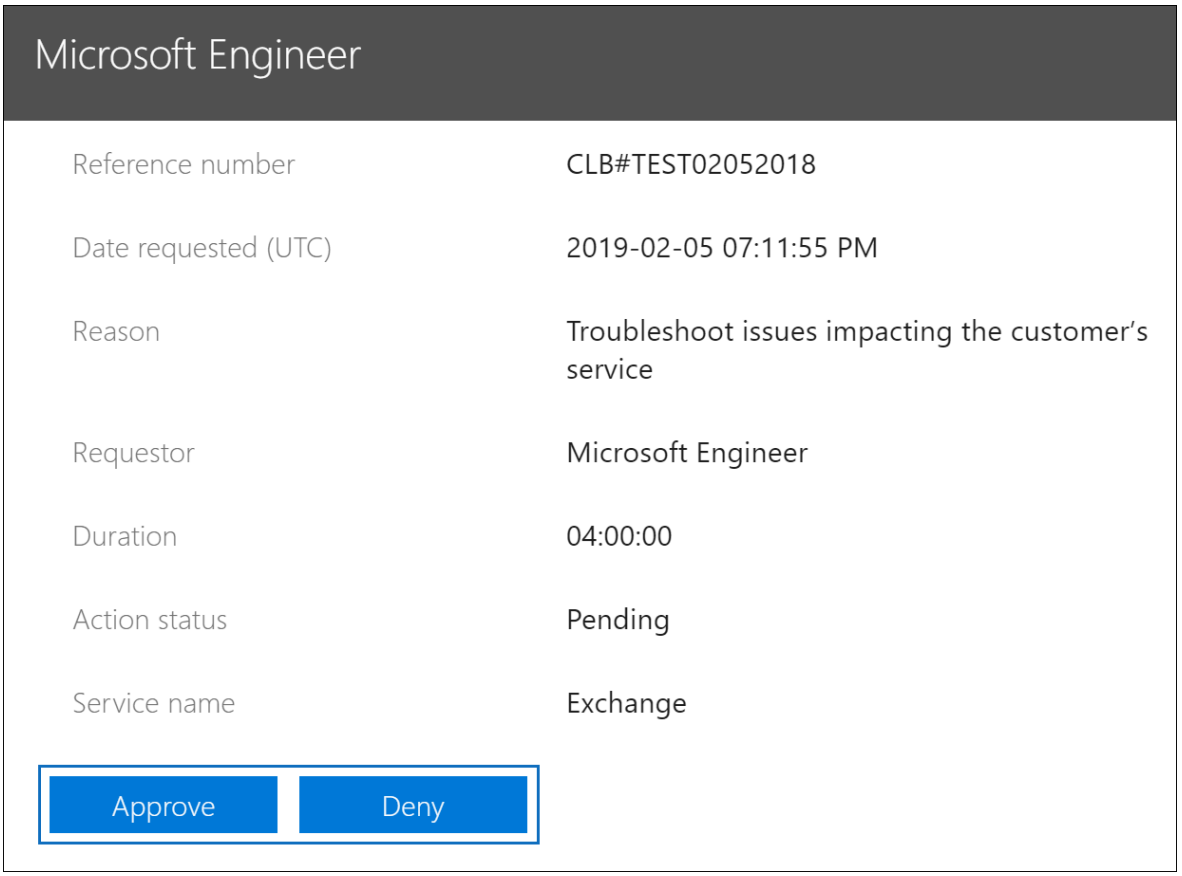
2. Choose **Support** > **Customer Lockbox Requests**.



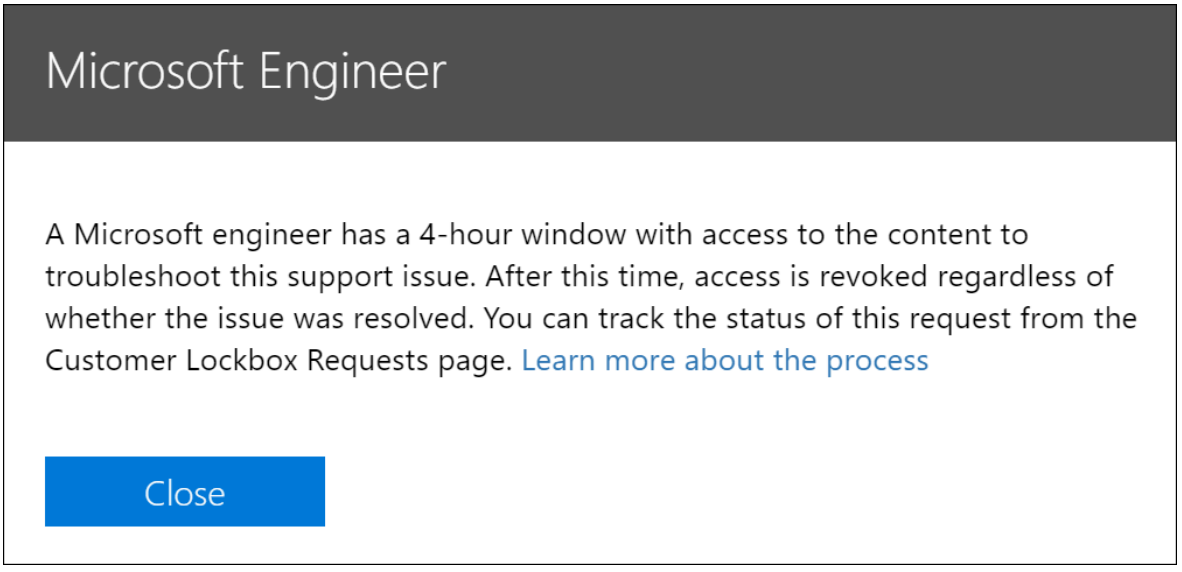
   A list of Customer Lockbox requests displays.



3. Select a Customer Lockbox request, and then choose **Approve** or **Deny**.



A confirmation message about the approval of the Customer Lockbox request displays.



> **Note**
>
> Use the Set-AccessToCustomerDataRequest cmdlet to approve, deny, or cancel Microsoft Purview Customer Lockbox requests that control access to your data by Microsoft support engineers. For more information, see Set-AccessToCustomerDataRequest.

Auditing Customer Lockbox requests

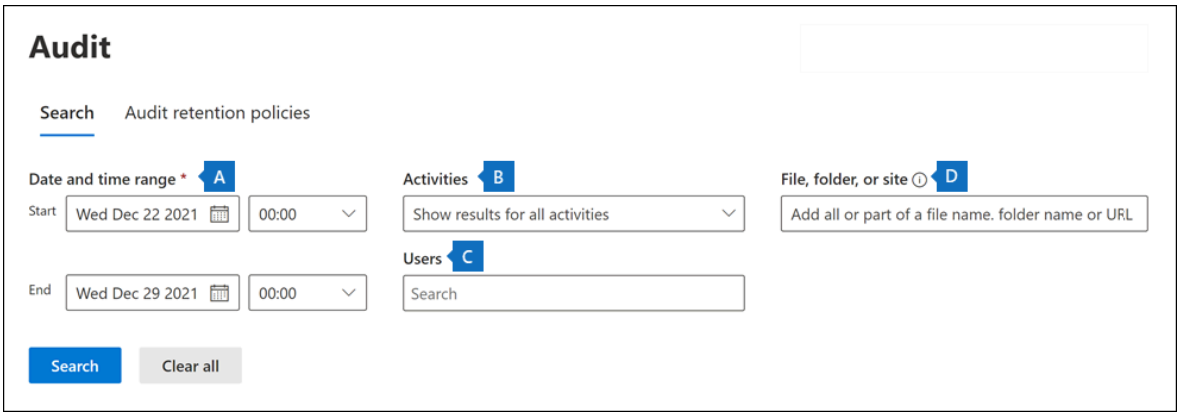
Audit records that correspond to Customer Lockbox requests are logged in the Microsoft 365 audit log. You can access these logs by using the audit log search tool in the Microsoft Purview portal. Actions related to accepting or denying a Customer Lockbox request and actions performed by Microsoft engineers (when access requests are approved) are also logged in the audit log. You can search for and review these audit records.

Search the audit log for activity related to Customer Lockbox requests

Before you can use the audit log to track requests for Customer Lockbox, there are some steps you need to take to set up audit logging, including assigning permissions to search the audit log. For more information, see Get started with auditing solutions. Once you've completed the setup, use these steps to create an audit log search query to return audit records related to Customer Lockbox:

1. Go to https://compliance.microsoft.com.
2. Sign in using an account that has been assigned the appropriate permissions to search the audit log.
3. In the left pane of the compliance center, choose **Audit**.

   The **Search** tab on the **Audit** page is displayed.



4. Configure the following search criteria:

   a. **Start date** and **End date**. Select a date and time range to display the events that occurred within that period.

   b. **Activities**. Leave this field blank so that the search returns audit records for all activities. This is necessary to return any audit records related to Customer

Lockbox requests and corresponding activity performed by Microsoft engineers.

c. **Users**. Leave this field blank.

d. **File, folder, or site**. Leave this field blank.

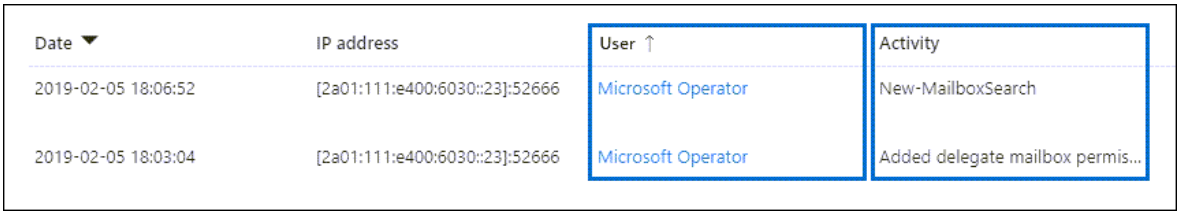
5. Click **Search** to run the search using your search criteria.

   The search results are displayed after a few moments. More search results will be added to the page until the search is complete.

6. Click the header in the **Activity** column to sort the results alphabetically based on the values in the **Activity** column.

7. Scroll down and look for audit records with an activity of **Set-AccessToCustomerDataRequest**. Records with this activity are related to an approver in your organization approving or denying a Customer Lockbox request.

8. Alternatively, click the header in the **User** column to sort the results alphabetically using the values in the **User** column. Look for the value of **Microsoft Operator**, which indicates activities performed by a Microsoft engineer in response to an approved Customer Lockbox request. The **Activity** column displays the action performed by the engineer.



9. In the list of results, click an audit record to display it.

Export the audit log search results

You can also export the audit log search results to a CSV file and then open the file in Excel to use the filtering and sorting capabilities to make it easier to find and view audit records related to a Customer Lockbox access request.

To export audit records, use the previous steps to search the audit log. When the search is complete, select **Export** > **Download all results** at the top of the search results page. When the export process is complete, you can download the CSV file to your local computer. For more detailed instructions, see Export, configure, and view audit log records.

After you download the file, you can open it in Excel and then filter on the **Operations** column to display audit records for **Set-AccessToCustomerDataRequest** activities. You can also filter on the **UserIds** column (using the value **Microsoft Operator**) to display audit records for activities performed by Microsoft engineers.

> **Note**
>
> When viewing audit records in the CSV file, additional information is contained in the **AuditData** column. The information in this column is contained in a JSON object, which contains multiple properties that are configured as *property:value* pairs separated by commas. You can use the JSON transform feature in the Power Query Editor in Excel to split each property in the JSON object in the **AuditData** column into multiple columns so that each property has its own column. This makes it easier to interpret this information. For detail instructions, see **Format the exported audit log using the Power Query Editor**.

Use PowerShell to search and export audit records

An alternative to using the audit search tool in the Microsoft Purview portal is to run the Search-UnifiedAuditLog cmdlet in Exchange Online PowerShell. One advantage of using PowerShell is that you can specifically search for **Set-AccessToCustomerDataRequest** activities or activities performed by Microsoft engineers related to a Customer Lockbox request.

After you connect to Exchange Online PowerShell, run one of the following commands. Replace the placeholders with a specific date range.

Search for `Set-AccessToCustomerDataRequest` activities

PowerShell

```
Search-UnifiedAuditLog -StartDate xx/xx/xxxx -EndDate xx/xx/xxxx -Operations
Set-AccessToCustomerDataRequest
```

Search for activities performed by Microsoft engineers

PowerShell

```
Search-UnifiedAuditLog -StartDate xx/xx/xxxx -EndDate xx/xx/xxxx -UserIds
"Microsoft Operator"
```

For more information and examples, see Use PowerShell to search and export audit log records.

We've also provided a PowerShell script that you can use to search the audit log and export the results to a CSV file. For more information, see Use a PowerShell script to search the audit log.

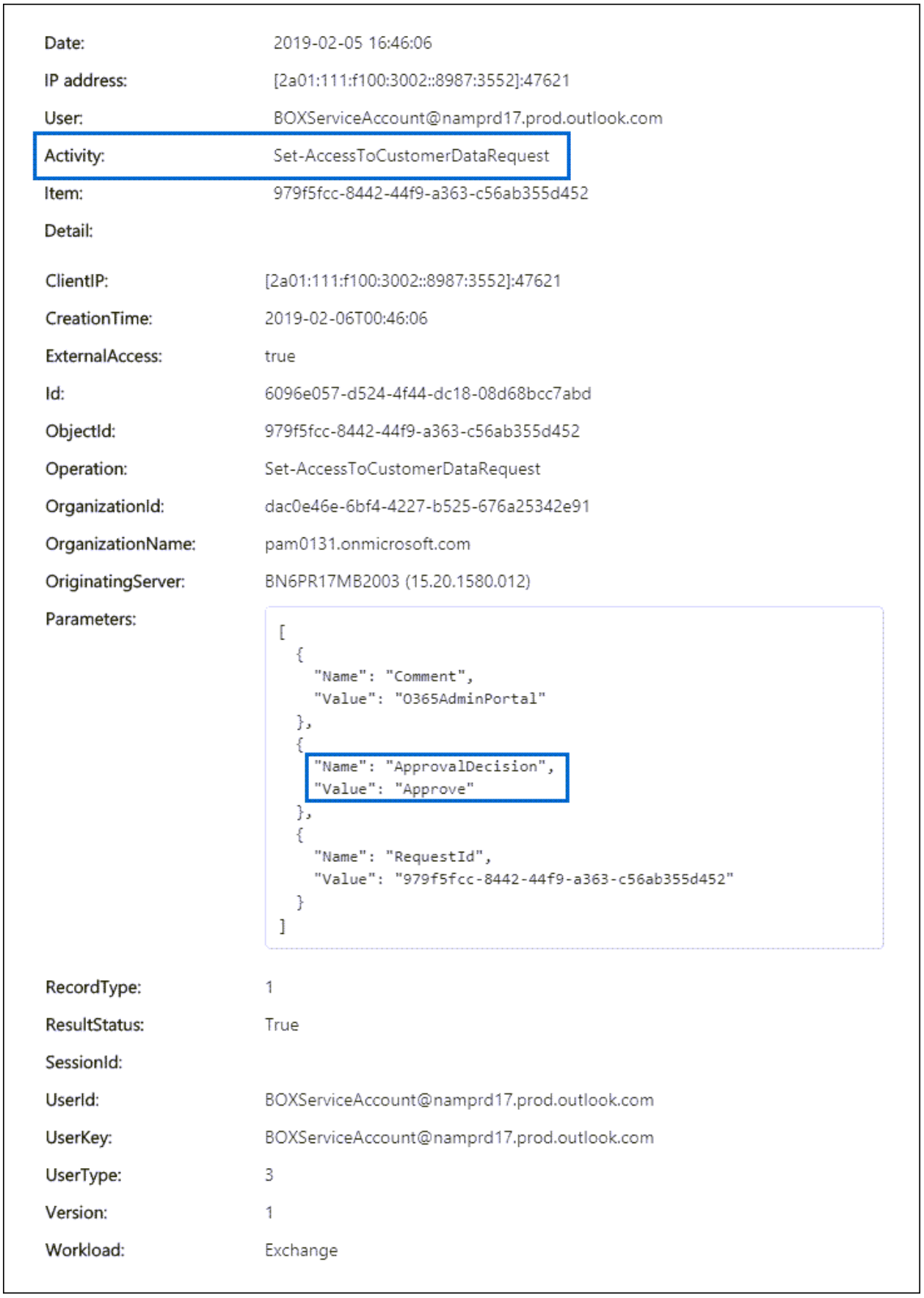
Audit record for a Customer Lockbox request

When a person in your organization approves or denies a Customer Lockbox request, the audit record is logged in the audit log contains the following information.

Expand table

Audit record property	Description
Date	The date and time when the Customer Lockbox request was approved or denied.
IP address	The IP address of the machine the approver used to approve or deny a request.
User	The service account BOXServiceAccount@[customerforest].prod.outlook.com.
Activity	Set-AccessToCustomerDataRequest; this is the auditing activity that is logged when you approve or deny a Customer Lockbox request.
Item	The Guid of the Customer Lockbox request

The following screenshot shows an example of an audit record that corresponds to an approved Customer Lockbox request. If a Customer Lockbox request was denied, then the value of `ApprovalDecision` parameter would be `Deny`.



Audit record for an action performed by a Microsoft engineer

The actions performed by a Microsoft engineer after a Customer Lockbox request is approved (and that may result in accessing customer content) are logged in the audit

log. These records contain the following information.

Expand table

Audit record property	Description
Date	Date time when the action was performed. The time that this action was performed will be within 4 hours of when the Customer Lockbox request was approved.
IP address	The IP Address of the machine Microsoft engineer used.
User	Microsoft Operator; this value indicates the record is related to a Customer Lockbox request.
Activity	Name of the activity performed by the Microsoft engineer.
Item	<empty>

Frequently asked questions

Which Microsoft 365 services does Customer Lockbox apply to?

Customer Lockbox is currently supported in Exchange Online, SharePoint Online, OneDrive for Business, Teams, and Windows 365.

Is Customer Lockbox available to all customers?

Customer Lockbox is included with the Microsoft 365 or Office 365 E5 subscriptions and can be added to other plans with an Information Protection and Compliance or an Advanced Compliance add-on subscription. For Government customers, Customer Lockbox is included with Microsoft 365 or Office 365 G5 subscription. See Plans and pricing for more information.

What is customer content?

Customer content is the data created by users of Microsoft 365 services and applications. Examples of customer content include:

- Email body or email attachments

- SharePoint site contents
- Information in the body of a SharePoint file
- Skype for Business presentation file body
- Instant messages (IM) or voice conversations
- Text entered in Teams chats and Teams channels, for example, 1:1 chats, group chats, shared channels, private channels, and meeting chat
- Other data pasted into Teams chat threads, such as code snippets, images, audio and video messages, and links
- App and bot data in Teams chats and Teams channels
- Teams activity feed
- Teams meeting recordings and transcripts
- Voicemail
- Files posted to Teams chats and Teams channels
- Microsoft 365 Copilot interactions
- Customer-generated blob or structured storage data (for example, SQL Containers)
- Customer-owned security information (for example, certificates, encryption keys, and passwords)
- Inferences, and all subsequent inferences, if customer content remains

For more information about customer content in Office 365, see the Office 365 Trust Center.

Who is notified when there is a request to access my content?

Global administrators and anyone assigned the Customer Lockbox access approver admin role are notified. These are also the same users who can approve for Customer Lockbox requests.

Who can approve or reject these requests in my organization?

Global administrators and anyone assigned the Customer Lockbox access approver admin role can approve Customer Lockbox requests. Customers control these role assignments in their organizations.

How do I opt in to Customer Lockbox?

A global administrator can enable and configure Customer Lockbox in the Microsoft 365 admin center.

If I approve a Customer Lockbox request, what can the engineer do and how will I know what the Microsoft engineer did?

After you approve a Customer Lockbox request, the Microsoft engineer granted these necessary privileges to access customer content by using pre-approved cmdlets. Actions taken by Microsoft engineers in response to Customer Lockbox requests are logged and accessible in the audit log in the Security & Compliance Center.

How do I know that Microsoft follows the approval process?

You can cross-reference the email approval notifications sent to admins and approvers in your organization with the Customer Lockbox request history in the Microsoft 365 admin center.

Customer Lockbox is included in the latest SOC 1 SSAE 16 audit report. For more details, you can find the latest reports in the Microsoft Service Trust Portal.

Can Microsoft modify the list of approvers for my tenant? If not, how is it prevented?

Only a global administrator in your organization can specify who can approve Customer Lockbox requests. That means only the members of the Global administrator group in Microsoft Entra ID can specify who can approve request. Membership of the Global administrator group in Microsoft Entra ID is managed only by your organization.

What if I need more information about a content access request to approve it?

Each Customer Lockbox request contains a Microsoft 365 service request number. You can contact Microsoft Support and reference this service number to get more information about the request.

When a Customer Lockbox request is approved, how long are the permissions valid?

Currently, the maximum period for the access permissions granted to the Microsoft engineer is 4 hours. The Microsoft engineer can also request a shorter period.

How can I get a history of all Customer Lockbox requests?

All Customer Lockbox requests are viewed in the Microsoft 365 admin center.

How do I correlate the content access requests with the related audit logs?

The Compliance Center Activity Feed contains log activities of Customer Lockbox. Customers can cross-reference the Customer Lockbox log activities from the activity feed against the email request they receive.

What happens when a customer doesn't respond to a Customer Lockbox request?

Customer Lockbox requests have a default duration of 12 hours. If you don't respond to a request within 12 hour, the request expires.

What does Microsoft do when a customer rejects a Customer Lockbox request?

If a customer rejects a Customer Lockbox request, no access to customer content occurs. If a user in your organization continues to experience a service issue requiring Microsoft to access customer content to resolve the issue, then the service issue might persist and Microsoft will inform the user about this.

How do I set up alerts whenever a request has been approved?

There is no built-in option to alert administrators. However, administrators can set up alerts using Microsoft Defender for Cloud Apps.

Does Customer Lockbox protect against data requests from law enforcement agencies or other third parties?

No. Microsoft takes third-party requests for customer data seriously. As a cloud service provider, Microsoft always advocates for the privacy of customer data. In the event we get a subpoena, Microsoft always attempts to redirect the third party to the customer to obtain the information. (Read Brad Smith's blog: Protecting customer data from government snooping). We periodically publish detailed information about the law enforcement requests that Microsoft receives.

See the Microsoft Trust Center regarding third-party data requests and the "Disclosure of Customer Data" section in the Online Services Terms for more information.

How does Microsoft ensure that a member of its staff doesn't have standing access to customer content in Office 365 applications?

Microsoft implements extensive preventive measures through access control systems, and detective measures to identify and address attempts to circumvent these access control systems. Microsoft 365 operates with the principles of least privilege and just-in-time access. Therefore, no Microsoft personnel have permission to access customer content on an ongoing basis. If permission is granted, it is for a limited duration.

Microsoft 365 uses an access control system called *Lockbox* to process requests for permissions that grant the ability to perform operational and administrative functions within the service. An operator must request access to customer content using Lockbox, which then requires a second person to take action on the request (for example, approve it) before access is granted. That second person can't be the requestor and must be designated to approve access to customer content. Only if the request is approved does the operator acquire temporary access to customer content. After the elevation period expires, Lockbox revokes access.

Refer to the Online Services Terms for more details about Microsoft general security practices.

Under what circumstances do Microsoft engineers need access to my content?

The most common scenario where Microsoft engineers need access customer content is when the customer makes a support request that requires access for troubleshooting. A foundational principle of Microsoft 365 is that the service operates without Microsoft access to customer content. Nearly all service operations performed by Microsoft are fully automated and human involvement is highly controlled and abstracted away from customer content. The goal for Microsoft 365 is access to customer content to support the service isn't needed until the customer approves a specific request for Microsoft access.

I already thought my data was secure with the Microsoft cloud, so why do I need Customer Lockbox?

Customer Lockbox provides an extra layer of control by offering customers the ability to give explicit access authorization for service operations. By demonstrating that procedures are in place for explicit data access authorization, Customer Lockbox also helps customers meet certain compliance obligations such as HIPAA and FEDRAMP.

Feedback

Was this page helpful? Yes No

What is Double Key Encryption (DKE)?

> *Microsoft Purview service description*

> *Applies to: Microsoft Purview Double Key Encryption, Microsoft Purview , Microsoft Purview Information Protection*
>
> *Service description for: Microsoft Purview*

Double Key Encryption (DKE) enables you to protect your highly sensitive data to meet specialized requirements. DKE lets you maintain control of your encryption keys. It uses two keys to protect data; one key in your control and a second key you store securely in Microsoft Azure. You maintain control of one of your keys using the Double Key Encryption service. Viewing data protected with Double Key Encryption requires access to both keys.

DKE helps you meet regulatory requirements across several regulations and standards such as the General Data Protection Regulation (GDPR), the Health Insurance Portability and Accountability Act (HIPAA), the Gramm-Leach-Bliley Act (GLBA), Russia's data localization law – Federal Law No. 242-FZ, Australia's Federal Privacy Act 1988, and New Zealand's Privacy Act 1993.

After you set up the DKE service and your keys, you apply protection to your highly sensitive content by using sensitivity labels.

Supported deployment scenarios

DKE supports several different configurations including both cloud and on-premises deployments. These deployments help to ensure that encrypted data remains opaque wherever you store it.

You can host the Double Key Encryption service used to request your key in a location of your choice (on-premises key management server or in the cloud). You maintain the service as you would any other application. Double Key Encryption lets you control access to the Double Key Encryption service. You can store your highly sensitive data on-premises or move it to the cloud. Double Key Encryption gives you the control to store your data and key in the same geographical location.

For more information about the default, cloud-based tenant root keys, see Managing the root key for your Azure Rights Management service.

When your organization should adopt DKE

DKE isn't for every organization, nor for all of your data. Let's say a typical organizational data landscape has the following structure:

- Nonsensitive data (about 80% of data): Most of an organization's data falls into this category. There are no issues or concerns with moving this data to the cloud today. Moving such data to the cloud can be beneficial and the organization can use the security built into the cloud.

- Sensitive (about 15% of data): Sensitive data needs to be protected. The organization expects the cloud service provider to provide security while enhancing productivity for this category of data so that they can meet compliance regulations. You want to ensure this data is labeled correctly using Microsoft Purview Information Protection and is protected with access control and retention and audit policies.

- Highly sensitive (about 5% of data): This set is the organization's crown jewels and needs to be heavily guarded. The organization doesn't want anyone to have access to such data. This category of data can also have regulatory requirements to have the keys in the same geographical region as the data. The keys might also need to be under the organization's strict custody. This content has the highest classification in your organization ("Top Secret") and access is restricted to just a few people. Highly sensitive data is what malicious users are after. Loss of this data can damage the organization's reputation and break trust with their customers.

As mentioned, Double Key Encryption is intended for your most sensitive data that is subject to the strictest protection requirements. You should do due diligence in identifying the right data to cover with this solution before you deploy. In some cases, you might need to narrow your scope and use other solutions. For example, for most of your data, consider Microsoft Purview Information Protection with Microsoft-managed keys or bring your own key (BYOK). These solutions are sufficient for documents that aren't subject to enhanced protections and regulatory requirements. Also, these solutions enable you to use the most powerful Microsoft 365 services; services that you can't use with DKE encrypted content. For example:

- Mail flow rules including anti-malware and spam that require visibility into the attachment
- Microsoft Delve

- eDiscovery
- Content search and indexing
- Office Web Apps including coauthoring functionality
- Copilot

DKE encrypted data isn't accessible at rest to Microsoft 365 services including Copilot. While you're using your DKE encrypted data in Office, the data still isn't accessible to Copilot, and you can't use Copilot in apps while you're using DKE encrypted data.

External applications or services that aren't integrated with DKE through the Microsoft Information Protection SDK are unable to perform actions on the encrypted data. The Microsoft Information Protection SDK 1.7+ supports Double Key Encryption. Applications that integrate with our SDK can reason over this data with sufficient permissions and integrations in place.

When your DKE encrypted data is in use in an Office app, it might be accessible to other Microsoft 365 services, depending on your Office version:

- In the latest versions of Office, DKE encrypted data in use is also not accessible to Microsoft 365 services. Because this change rolled out at the same time as the separate privacy control for sensitivity labels that prevents sending labeled content to connected experiences for analysis, you can identify the minimum Office versions by using the capabilities table and the row **Prevent connected experiences that analyze content**.

- In previous versions of Office, DKE encrypted data in use is accessible to Microsoft 365 services unless you use a policy setting to turn off connected experiences that analyze content. For more information, see Use policy settings to manage privacy controls for Microsoft 365 Apps for enterprise.

Use Microsoft Purview Information Protection capabilities (classification and labeling) to protect most of your sensitive data and only use DKE for your mission-critical data. Double Key Encryption is relevant for sensitive data in highly regulated industries such as financial services and healthcare.

Sensitivity label features that don't support DKE

Because DKE doesn't use a cloud-based key managed by the Azure Rights Management service, several sensitivity label features can't work with DKE-encrypted content:

- **Co-authoring and AutoSave**

    - Coauthoring and AutoSave aren't supported for Office documents encrypted with DKE. Labels with DKE still display in Office apps, but coauthoring and AutoSave are disabled when the label is applied. For more information, see Coauthoring for files encrypted with sensitivity labels.

- **Files in SharePoint and OneDrive**
    - SharePoint and OneDrive can't process DKE-encrypted files. These files can't be opened in Office for the web, and collaborative features such as coauthoring, eDiscovery, data loss prevention, and search don't work for DKE-encrypted files in these locations. For more information, see Enable sensitivity labels for files in SharePoint and OneDrive.

- **Default label for a SharePoint document library**
    - A sensitivity label that uses DKE can't be set as the default sensitivity label for a SharePoint document library. For more information, see Configure a default sensitivity label for a SharePoint document library.

- **Loop components and pages**
    - Sensitivity labels configured with DKE aren't supported for Loop components and pages. For more information, see Sensitivity labels for Loop components and pages.

- **Teams meetings and chat**
    - DKE labels can't be used to protect calendar items, Teams meetings, and chat. Meetings and chat require encryption that uses the Azure Rights Management service. For more information, see Use sensitivity labels to protect calendar items, Teams meetings, and chat.

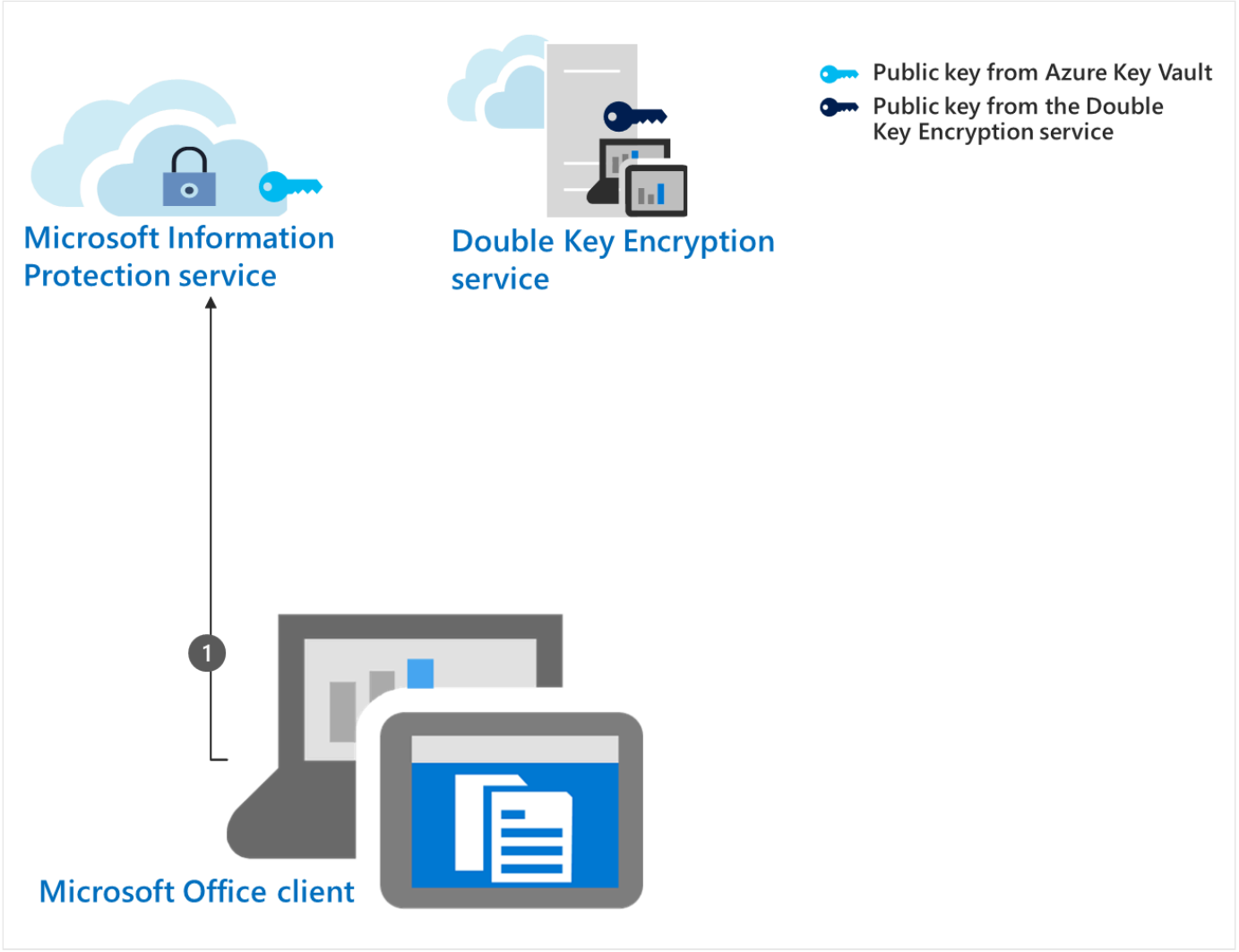
If your organizations have any of the following requirements, you can use DKE to help secure your content:

- You have regulatory requirements to hold keys within a geographical boundary.
- All of the keys that you hold for data encryption and decryption are maintained in your data center.

DKE encryption workflow

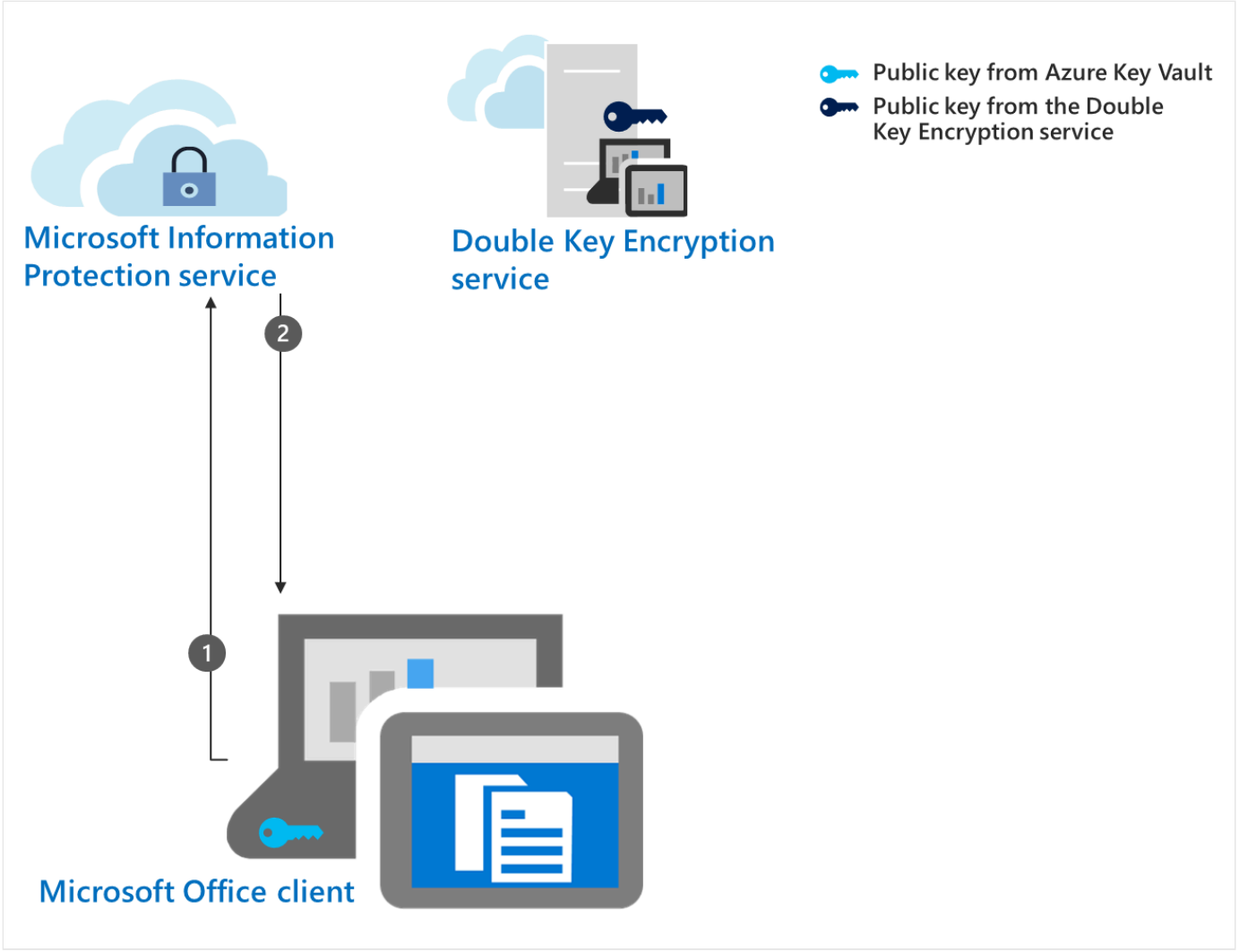
This section breaks apart the workflow into separate steps to illustrate how two keys are used to protect an Office document.

Step 1: Bootstrapping



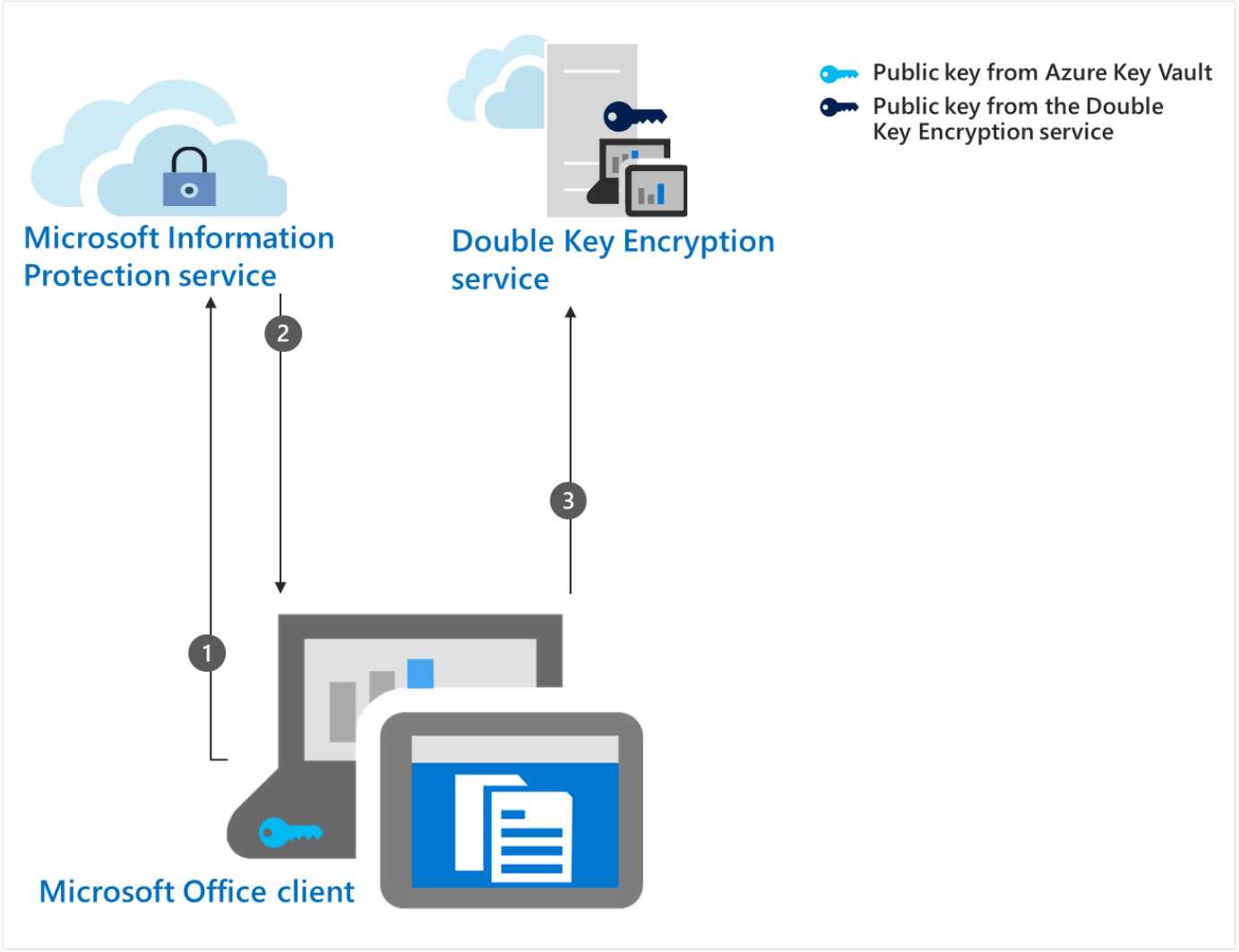
The Microsoft Office client performs startup configuration tasks and sends requests and information to the Azure Rights Management encryption service. This process is also called *bootstrapping*. Tasks include authorizing the user using Microsoft Entra ID, downloading certificates and templates, and so on. Bootstrapping tasks are first-time connection and startup tasks that give the user access to the Azure Rights Management encryption policies.

Step 2: Collect and cache the public key for the Azure Rights Management service



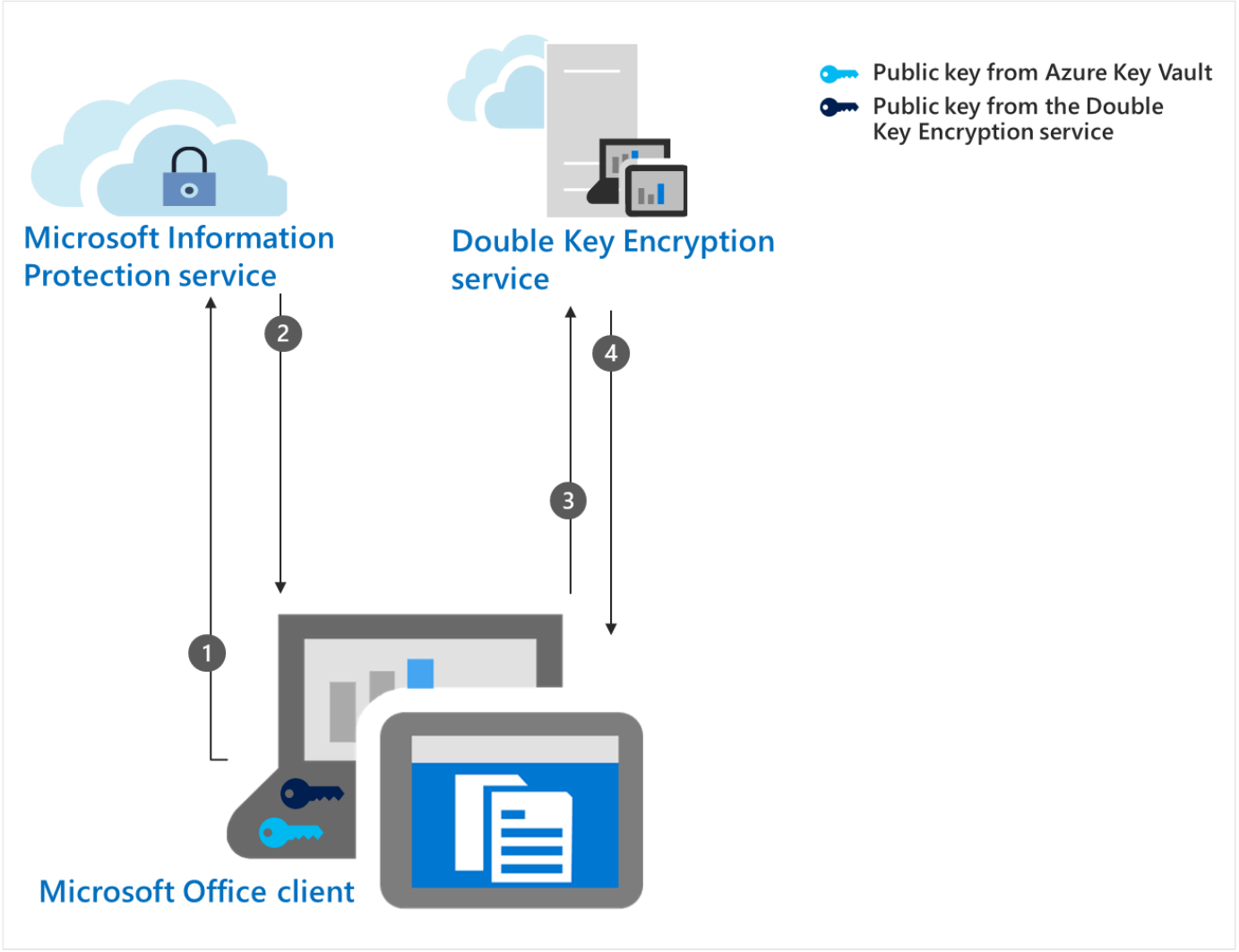
The Office app retrieves the public key from the Azure Key Vault in the encryption service based on the authorized user using Microsoft Entra ID. Once collected, the client caches the key for 30 days by default. Once cached, the client doesn't have to bootstrap to the Azure Rights Management service again until the key expires. As an administrator, you can configure a different caching period for the Azure Rights Management service. You need to set a cache period for this key or accept the 30 day default. Without a cache period, offline publishing doesn't work.

Step 3: Request the DKE public key



The Office client requests your other public key from the Double Key Encryption service based on the authorized user using Microsoft Entra ID.

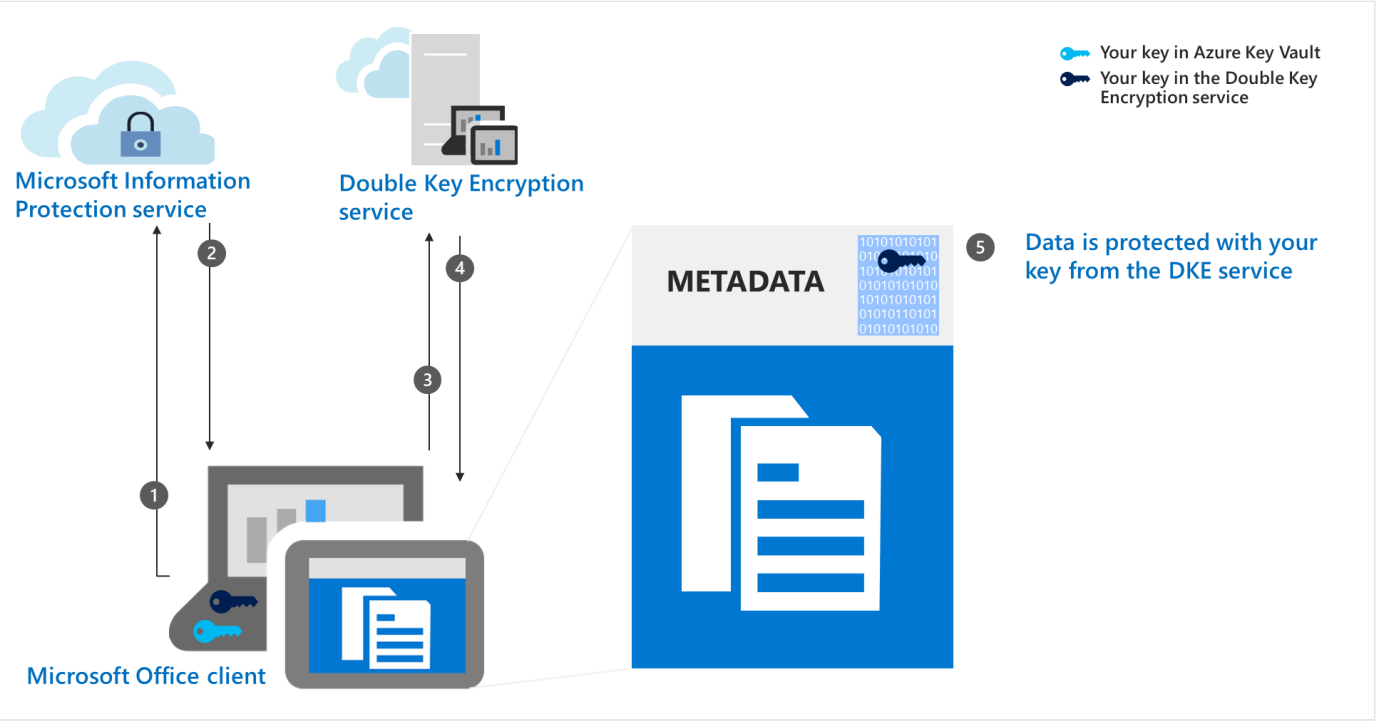
Step 4: Collect and cache the DKE key



The Double Key Encryption service sends this public key to the Office client. The client caches the key in the device for as long you configured it to. Unlike the key from Azure,

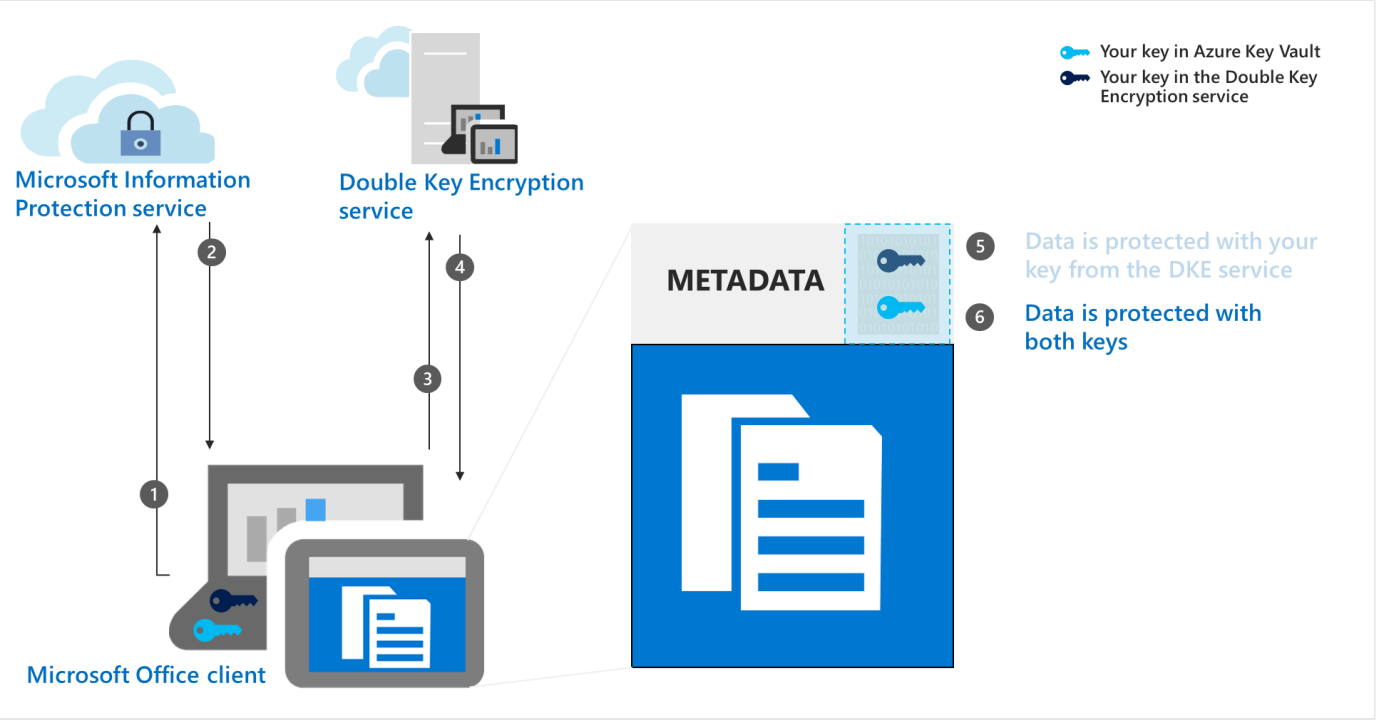
- You don't have to set up a cache period for the key hosted by the Double Key Encryption service.
- If you do want to set up a cache period, you can set it up when you deploy the Double Key Encryption service, or after you deploy.

Step 5: Protect the document with the DKE key



The Microsoft Office client encrypts the part of the metadata **that controls access to the content** using your public key that was retrieved from the Double Key Encryption service.

Step 6: Protect the document with the Azure key



The Microsoft Office client encrypts **the already encrypted part of the document metadata** with your public key from Azure.

The data is now protected with both keys.

System and licensing requirements for DKE

This section details server and client system and configuration requirements that must be met before you can successfully deploy DKE in your environment.

licensing requirements for DKE

Double Key Encryption comes with Microsoft 365 E5. If you don't have a Microsoft 365 E5 license, you can sign up for a trial. For more information about these licenses, see Microsoft 365 licensing guidance for security & compliance.

The Azure Rights Management service is required for DKE

DKE works with sensitivity labels and requires encryption with rights management from Microsoft Purview Information Protection.

DKE labeling requirements for Office apps

Use sensitivity labels built in to Office apps to support DKE in Word, Excel, PowerPoint, and Outlook. For supported versions, see the capabilities tables and the row **Double Key Encryption (DKE)**.

DKE on client computers

Users apply DKE sensitivity labels through these interfaces.

- Sensitivity labeling in Windows Office apps
- Microsoft Purview Information Protection file labeler in Windows File Explorer
- Microsoft Purview Information Protection PowerShell
- Microsoft Purview Information Protection scanner

Install prerequisites on each client computer where you want to protect and consume protected documents.

Supported environments for storing and viewing DKE-protected content

Supported applications. Microsoft 365 Apps for enterprise clients on Windows, including Word, Excel, PowerPoint, and Outlook.

Online content support. You can store documents and files that are protected with Double Key Encryption online in both SharePoint and OneDrive. You must label and protect documents and files with DKE by supported applications before you upload to these locations. You can share encrypted content by email, but you can't view encrypted documents and files online. Instead, you must view protected content using the supported desktop applications and clients on your local computer.

Labeling scenarios outside of Office apps. Apply DKE labels outside of Office apps using the Microsoft Purview Information Protection file labeler in File Explorer right-click options, Microsoft Purview Information Protection PowerShell labeling cmdlets, or the Microsoft Purview Information Protection scanner.

Viewing DKE-protected content outside of Office apps. You can use the Microsoft Purview Information Protection Viewer to consume DKE-protected emails and documents such as PDF.

Encrypt Only and Do Not Forward scenarios. Encrypt Only and Do Not Forward aren't supported with DKE in Outlook. All other Office applications support user-defined permission labels with DKE.

Related content

- Set up Double Key Encryption
- Double Key Encryption FAQ

Last updated on 05/12/2026

Set up Double Key Encryption

07/23/2025

> *Applies to: Microsoft Purview Double Key Encryption, Microsoft Purview, Azure Information Protection*

> *Service description for: Microsoft Purview*

Follow these general steps to set up DKE. Once you complete these steps, your end users can protect your highly sensitive data with Double Key Encryption.

1. Deploy the DKE service as described in this article. Ensure your environment meets the minimum system and licensing requirements. For more information, see System and licensing requirements for DKE.

2. Create a label with Double Key Encryption. In the Microsoft Purview portal, navigate to **Information protection** and create a new label with Double Key Encryption. See Restrict access to content by using sensitivity labels to apply encryption.

3. Configure the registry on client devices so you can use Double Key Encryption labels. Next, protect your data by selecting the Double Key Encrypted label from the Sensitivity ribbon in Microsoft Office.

There are several ways you can complete some of the steps to deploy Double Key Encryption. This article provides detailed instructions so that less experienced admins successfully deploy the service. If you're comfortable doing so, you can choose to use your own methods.

Deploy DKE

This article and the deployment video use Azure as the deployment destination for the DKE service. If you're deploying to another location, you need to provide your own values.

Follow these general steps to set up Double Key Encryption for your organization.

1. Install software prerequisites for the DKE service
2. Clone the Double Key Encryption GitHub repository
3. Modify application settings
4. Generate test keys
5. Build the project
6. Deploy the DKE service and publish the key store
7. Validate your deployment

8. Register your key store
9. Create sensitivity labels using DKE
10. Migrate protected files from HYOK labels to DKE labels

When you're done, you can encrypt documents and files using DKE. For information, see Apply sensitivity labels to your files and email in Office.

Install software prerequisites for the DKE service

Install these prerequisites on the computer where you want to install the DKE service.

.NET 8.0 SDK. Download and install the SDK from Download .NET 8.0.

Visual Studio Code. Download Visual Studio Code from https://code.visualstudio.com/. Once installed, run Visual Studio Code and select **View** > **Extensions**. Install these extensions.

- C# for Visual Studio Code
- NuGet Package Manager

Git resources. Download and install one of the following.

- Git
- GitHub Desktop
- GitHub Enterprise

OpenSSL. You must have OpenSSL installed to generate test keys after you deploy DKE. Make sure you're invoking it correctly from your environment variables path. For example, see "Add the installation directory to PATH" at https://www.osradar.com/install-openssl-windows/ for details.

Clone the DKE GitHub repository

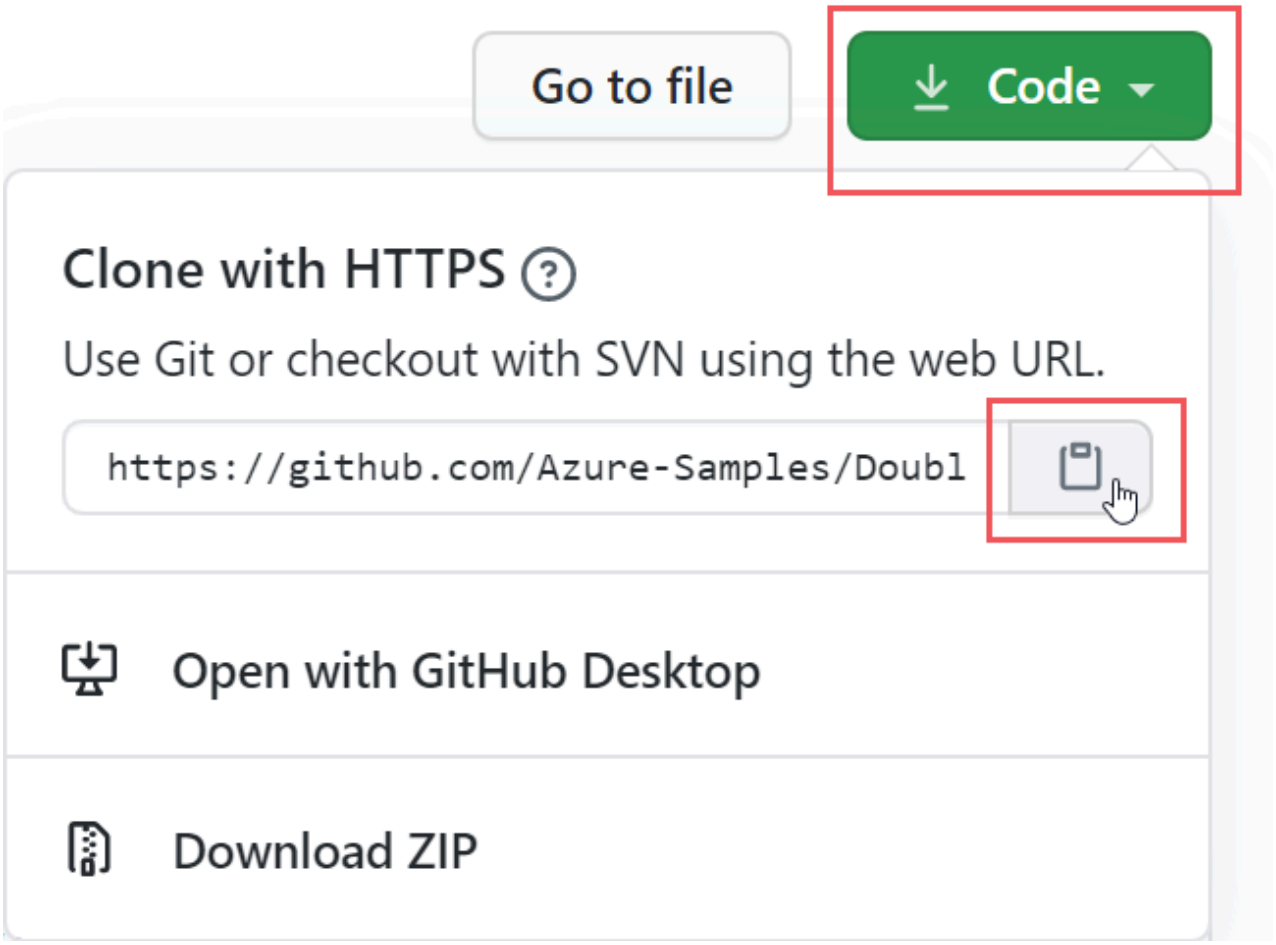
Microsoft supplies the DKE source files in a GitHub repository. You clone the repository to build the project locally for your organization's use. The DKE GitHub repository is located at https://github.com/Azure-Samples/DoubleKeyEncryptionService.

The following instructions are intended for inexperienced git or Visual Studio Code users:

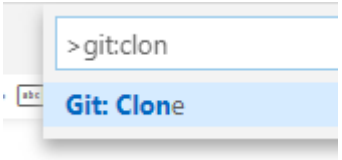
1. In your browser, go to: https://github.com/Azure-Samples/DoubleKeyEncryptionService.

2. Towards the right side of the screen, select **Code**. Your version of the UI might show a **Clone or download** button. Then, in the dropdown that appears, select the copy icon to copy the URL to your clipboard.

   For example:



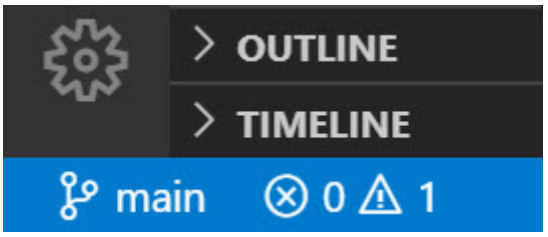
3. In Visual Studio Code, select **View** > **Command Palette** and select **Git: Clone**. To jump to the option in the list, start typing `git: clone` to filter the entries, and then select it from the drop-down. For example:



4. In the text box, paste the URL that you copied from Git and select **Clone from GitHub**.

5. In the **Select Folder** dialog that appears, browse to and select a location to store the repository. Select **Open**.

   The repository opens in Visual Studio Code, and displays the current Git branch at the bottom left. The branch should be **main**. For example:



6. If you're not on the main branch, select it. In Visual Studio Code, select the branch and choose **main** from the list of branches that displays.

   > **Important**
   >
   > Selecting the main branch ensures that you have the correct files to build the project. If you don't choose the correct branch your deployment will fail.

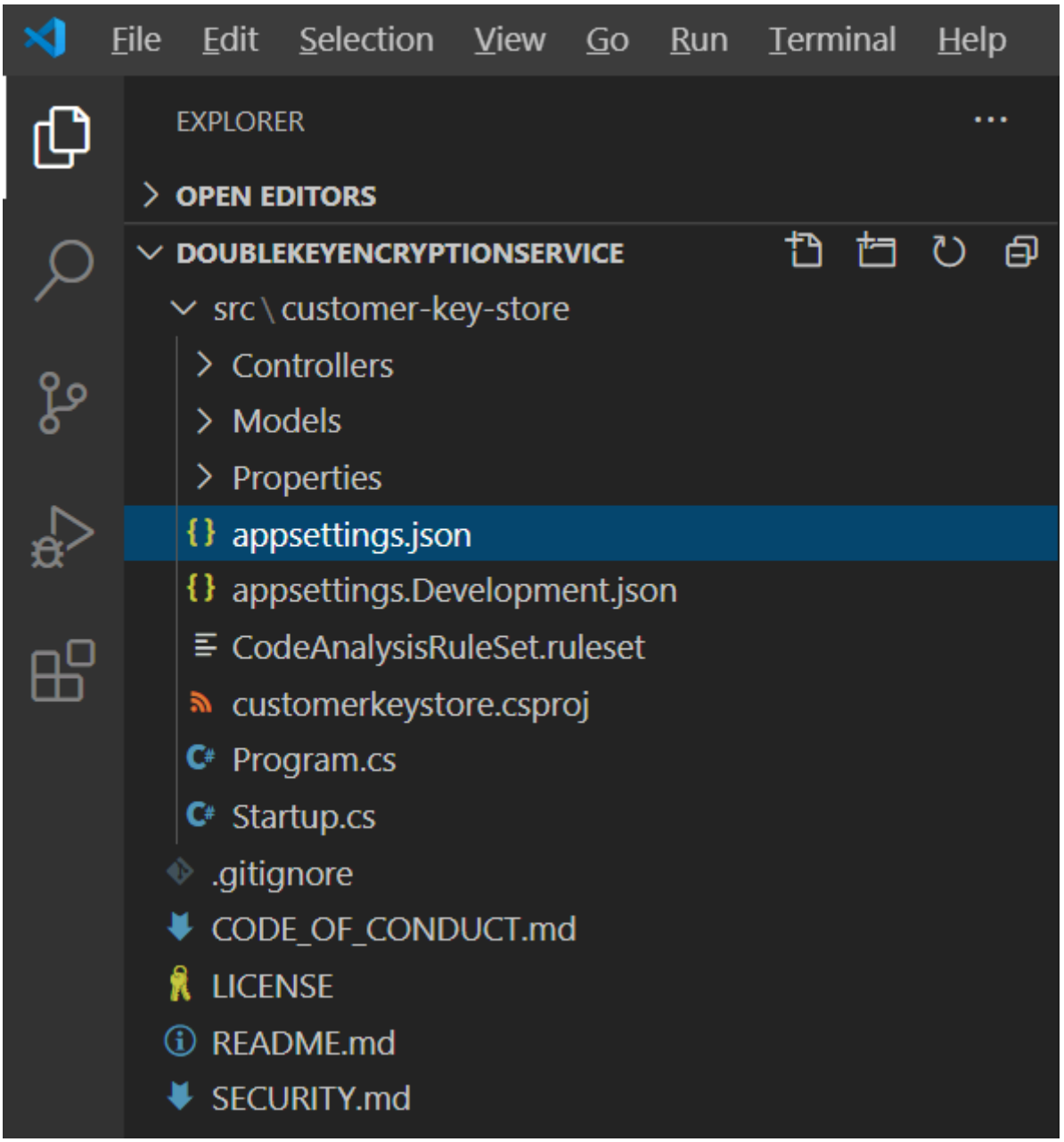
You now have your DKE source repository set up locally. Next, modify application settings for your organization.

Modify application settings

To deploy the DKE service, you must modify the following types of application settings:

- Key access settings
- Tenant and key settings

You modify application settings in the appsettings.json file. This file is located in the DoubleKeyEncryptionService repo you cloned locally under DoubleKeyEncryptionService\src\customer-key-store. For example, in Visual Studio Code, you can browse to the file as shown in the following picture.



Key access settings

Choose whether to use email or role authorization. DKE supports only one of these authentication methods at a time.

- **Email authorization**. Allows your organization to authorize access to keys based on email addresses only.
- **Role authorization**. Allows your organization to authorize access to keys based on Active Directory groups, and requires that the web service can query LDAP.

To set key access settings for DKE using email authorization

1. Open the **appsettings.json** file and locate the `AuthorizedEmailAddress` setting.

2. Add the email address or addresses that you want to authorize. Separate multiple email addresses with double quotes and commas. For example:

```
JSON

"AuthorizedEmailAddress": ["email1@company.com", "email2@company.com ",
"email3@company.com"]
```

3. Locate the `LDAPPath` setting and remove the text `If you use role authorization (AuthorizedRoles) then this is the LDAP path.` between the double quotes. Leave the double quotes in place. When you're finished, the setting should look like this.

```
JSON

"LDAPPath": ""
```

4. Locate the `AuthorizedRoles` setting and delete the entire line.

This image shows the **appsettings.json** file correctly formatted for email authorization.



To set key access settings for DKE using role authorization

1. Open the **appsettings.json** file and locate the `AuthorizedRoles` setting.

2. Add the Active Directory group names you want to authorize. Separate multiple group names with double quotes and commas. For example:

```
JSON

"AuthorizedRoles": ["group1", "group2", "group3"]
```

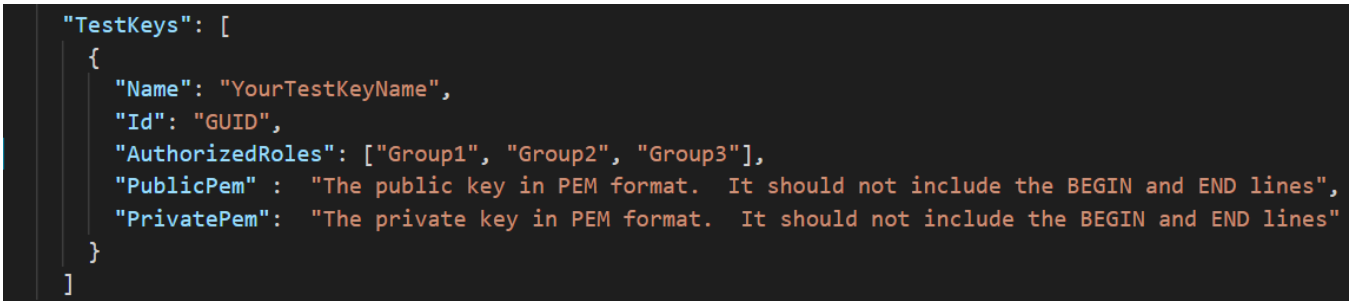
3. Locate the `LDAPPath` setting and add the Active Directory domain. For example:

```
JSON

"LDAPPath": "contoso.com"
```

4. Locate the `AuthorizedEmailAddress` setting and delete the entire line.

This image shows the **appsettings.json** file correctly formatted for role authorization.



```
"TestKeys": [
  {
    "Name": "YourTestKeyName",
    "Id": "GUID",
    "AuthorizedRoles": ["Group1", "Group2", "Group3"],
    "PublicPem" :  "The public key in PEM format.  It should not include the BEGIN and END lines",
    "PrivatePem":  "The private key in PEM format.  It should not include the BEGIN and END lines"
  }
]
```

Tenant and key settings

DKE tenant and key settings are located in the **appsettings.json** file.

To configure tenant and key settings for DKE

1. Open the **appsettings.json** file.
2. Locate the `ValidIssuers` setting and replace `<tenantid>` with your tenant ID, making sure you keep the slash at the end of the tenant ID. You can locate your tenant ID by going to the Azure portal and viewing the tenant properties. For example:

   JSON

   ```
   "ValidIssuers": [
   "https://sts.windows.net/9c99431e-b513-44be-a7d9-e7b500002d4b/"
   ]
   ```

> **Note**
>
> If you want to enable external B2B access to your key store, you will also need to include these external tenants as part of the valid issuers' list.

Locate the `JwtAudience`. Replace the placeholder value, `"https://dkeservice.contoso.com"` in the json with the hostname of the machine where you want to run the DKE service.

> **Important**
>
> The value for `JwtAudience` must match the name of your host *exactly*.

- `TestKeys:Name`. Enter a name for your key. For example: `TestKey1`
- `TestKeys:Id`. Create a GUID and enter it as the `TestKeys:ID` value. For example, `DCE1CC21-FF9B-4424-8FF4-9914BD19A1BE`. You can use a site like Online GUID Generator to randomly generate a GUID.
- `JwtAuthorization`. Replace `{tenant-id}` with your tenant id.

The following json shows the correct format for tenant and keys settings in **appsettings.json**. `LDAPPath` is configured for role authorization.

JSON

```
{
  "TokenValidationParameters": {
    "ValidIssuers": [
      "https://sts.windows.net/9c99431e-b513-44be-a7d9-e7b50002d4b/"
    ]
  },
  "Logging": {
    "LogLevel": {
      "Default": "Warning"
    }
  },
  "AllowedHosts": "*",
  "JwtAudience": "https://dkeservice.contoso.com/",
  "JwtAuthorization": "https://login.microsoftonline.com/{tenant-
id}/oauth2/v2.0/authorize",
  "RoleAuthorizer": {
    "LDAPPath": ""
  },
  "TestKeys": [
    {
      "Name": "TestKey1",
      "Id": "DCE1CC21-FF98-4424-8FF4-9914BD19A1BE"
    }
  ]
}
```

Generate test keys

Once you have your application settings defined, you're ready to generate public and private test keys.

To generate keys:

1. From the Windows Start menu, run the OpenSSL Command Prompt.

2. Change to the folder where you want to save the test keys. The files you create by completing the steps in this task are stored in the same folder.

3. Generate the new test key.

```
Console

openssl req -x509 -newkey rsa:2048 -keyout key.pem -out cert.pem -days 365
```

4. Generate the private key.

   If you installed OpenSSL version 3 or later, run the following command:

```
Console

openssl rsa -in key.pem -out privkeynopass.pem -outform PEM -traditional
```

   Otherwise, run the following command:

```
Console

openssl rsa -in key.pem -out privkeynopass.pem -outform PEM
```

5. Generate the public key.

```
Console

openssl rsa -in key.pem -pubout > pubkeyonly.pem
```

6. In a text editor, open **pubkeyonly.pem**. Copy all of the content in the **pubkeyonly.pem** file, except the first and last lines, into the `PublicPem` section of the **appsettings.json** file.

7. In a text editor, open **privkeynopass.pem**. Copy all of the content in the **privkeynopass.pem** file, except the first and last lines, into the `PrivatePem` section of the **appsettings.json** file.

8. Remove all blank spaces and newlines in both the `PublicPem` and `PrivatePem` sections.

   > **ⓘ Important**
   >
   > When you copy this content, do not delete any of the PEM data.

9. In Visual Studio Code, browse to the **Startup.cs** file. This file is located in the DoubleKeyEncryptionService repo you cloned locally under DoubleKeyEncryptionService\src\customer-key-store.

10. Locate the following lines:

```
C#

#if USE_TEST_KEYS
#error !!!!!!!!!!!!!!!!!!!! Use of test keys is only supported for testing,
DO NOT USE FOR PRODUCTION !!!!!!!!!!!!!!!!!!!!!!!!!!!!
services.AddSingleton<ippw.IKeyStore, ippw.TestKeyStore>();
#endif
```

11. Replace these lines with the following text:

```
C#

services.AddSingleton<ippw.IKeyStore, ippw.TestKeyStore>();
```

    The end results should look similar to the following.

```
services.AddSingleton<ippw.IKeyStore, ippw.TestKeyStore>();

services.AddTransient<ippw.KeyManager, ippw.KeyManager>();

services.AddMvc().SetCompatibilityVersion(CompatibilityVersion.Version_2_1);
```

Now you're ready to build your DKE project.

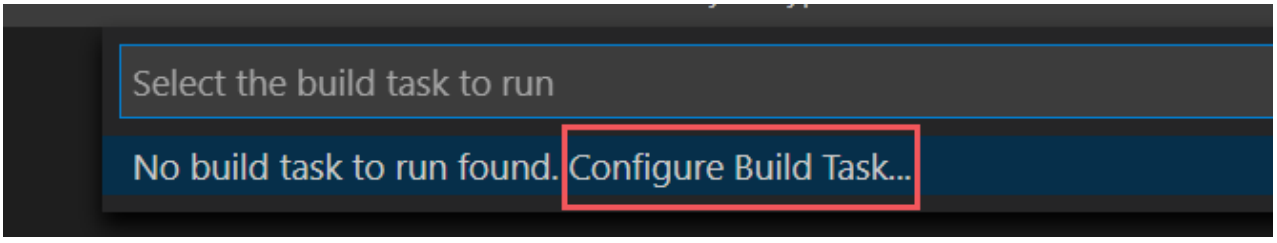
Build the project

Use the following instructions to build the DKE project locally:

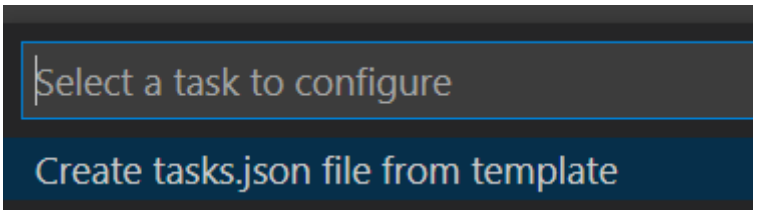
1. In Visual Studio Code, in the DKE service repository, select **View** > **Command Palette** and then type **build** at the prompt.

2. From the list, choose **Tasks: Run build task**.

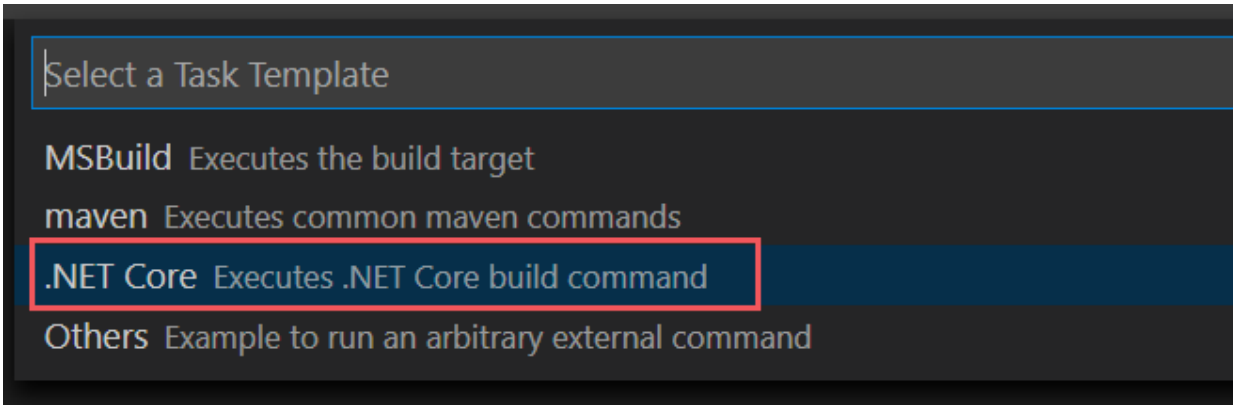
    If there are no build tasks found, select **Configure Build Task** and create one for .NET core as follows.



    a. Choose **Create tasks.json from template**.



b. From the list of template types, select **.NET Core**.



c. In the build section, locate the path to the **customerkeystore.csproj** file. If it's not there, add the following line:

```JSON
"${workspaceFolder}/src/customer-key-store/customerkeystore.csproj",
```

d. Run the build again.

3. Verify that there are no red errors in the output window.

   If there are red errors, check the console output. Ensure that you completed all the previous steps correctly and the correct build versions are present.

Your setup is now complete. Before you publish the keystore, in appsettings.json, for the JwtAudience setting, ensure the value for hostname exactly matches your App Service host name.

Deploy the DKE service and publish the key store

For production deployments, deploy the service either in a third-party cloud or publish to an on-premises system.

You may prefer other methods to deploy your keys. Select the method that works best for your organization.

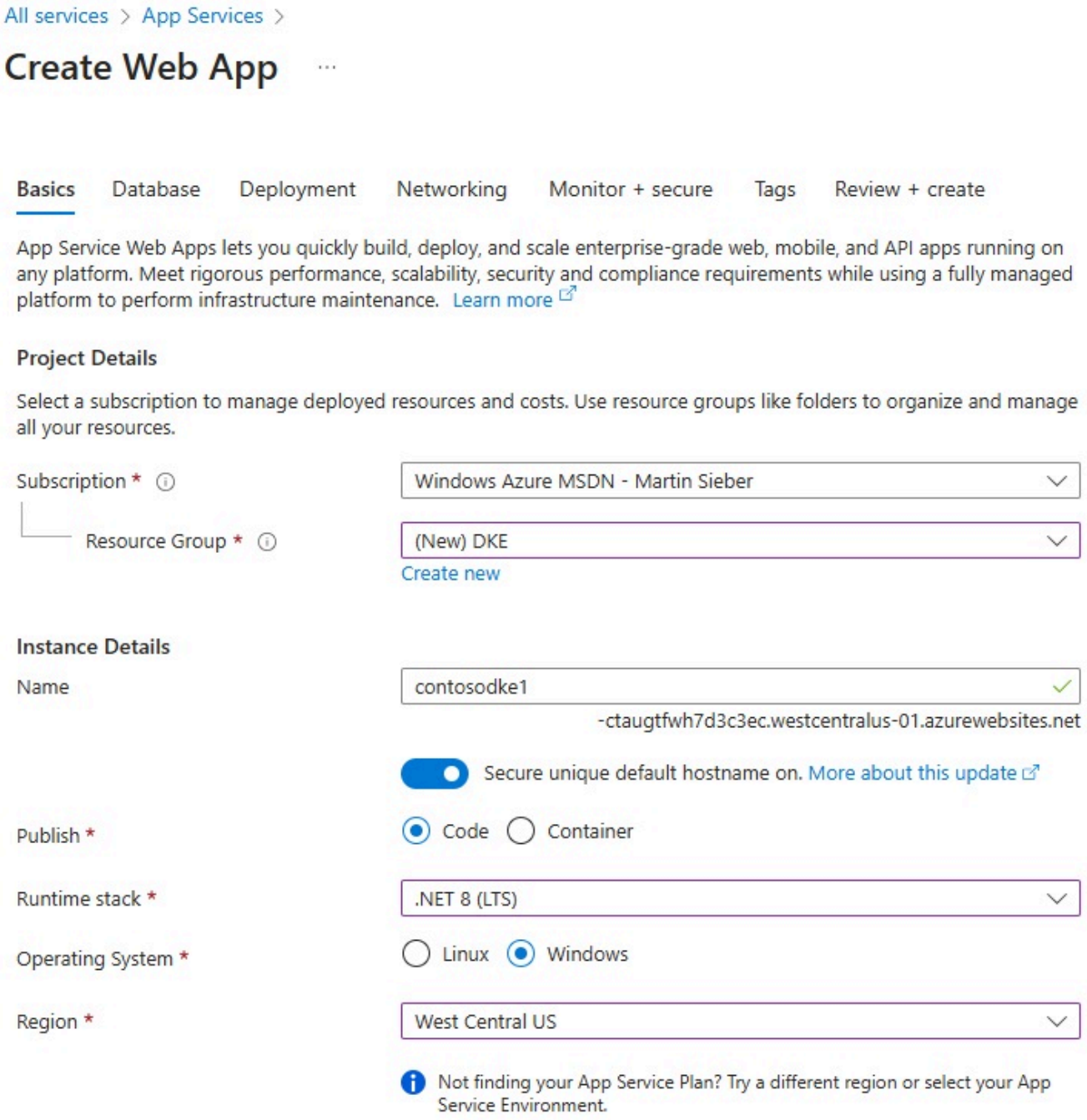
For pilot deployments, you can deploy in Azure and get started right away.

To create an Azure Web App instance to host your DKE deployment

To publish the key store, create an Azure App Service instance to host your DKE deployment. Next, publish your generated keys to Azure.

1. In your browser, sign in to the Microsoft Azure portal, and go to **App Services** > **Create** > **Web App**.
2. Create a new or select an existing resource group.
3. Provide a name for the instance, e.g. contosodke1. This name does not match the name in the config file. The value you provide for the name is also the WebAppInstanceName.
4. In pricing plan, select a size that allows custom domains, e.g. Basic B1.
5. For **Publish**, select **code**, and for **Runtime stack**, select **.NET 8 (LTS)**.
6. Select **Review + create**

For example:

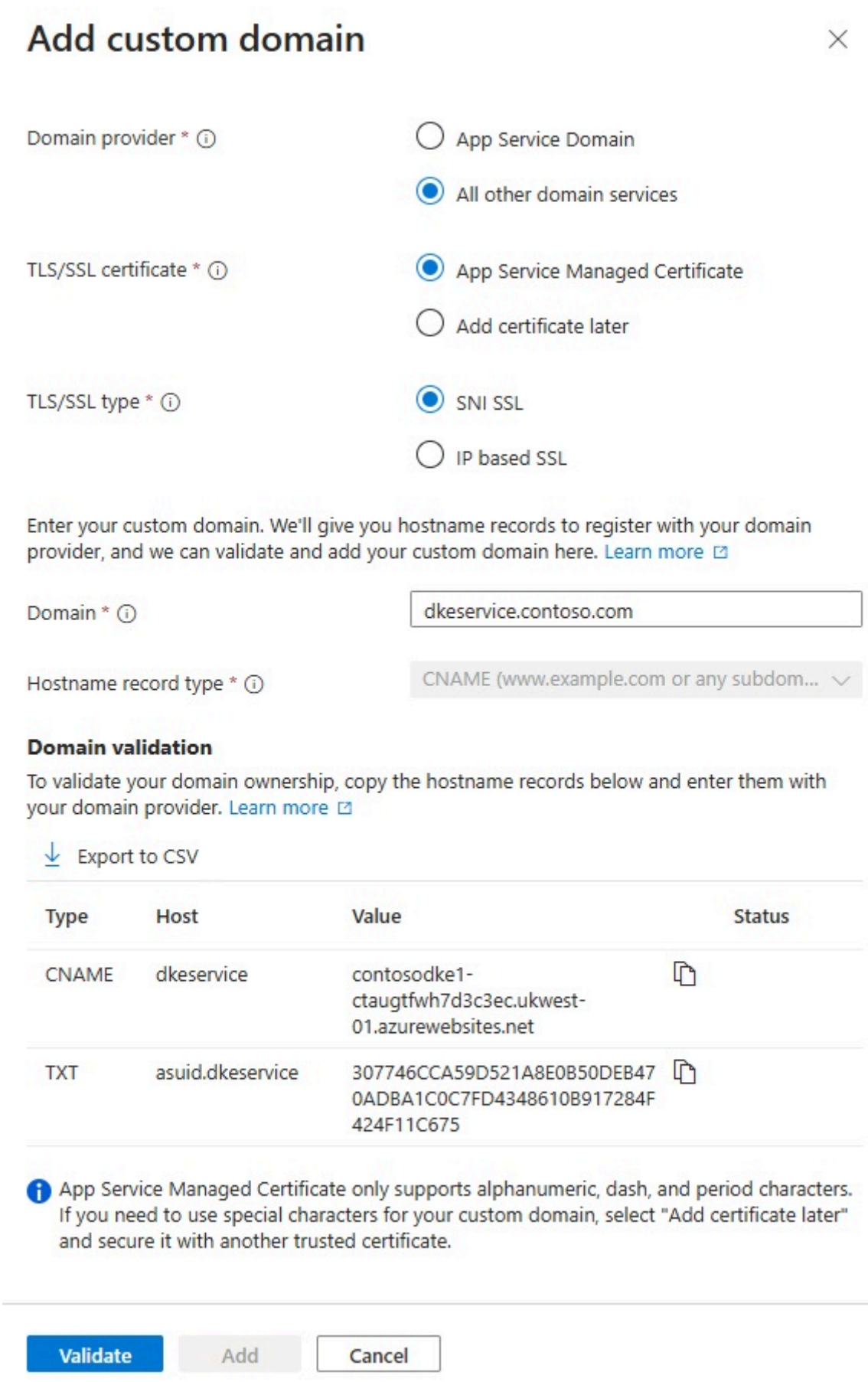


1. In the newly created web app, go to “Settings”, “Custom domains”, select **Add custom domain**
2. In **Domain provider,** select **All other domain services**
3. In **Domain**, enter the hostname of the computer where you want to install the DKE service.

> **ⓘ Important**
>
> Make sure the hostname is the same name as the one defined for the JwtAudience setting in the appsettings.json file.

10. In the public DNS, perform the suggested changes as indicated in the Domain validation section and select **Validate**.

For example:



1. Do the following to publish your generated keys:

    - Publish via ZipDeployUI
    - You can use other methods to upload your binaries to that app, such as FTP.

for example, ftp

Publish via ZipDeployUI

1. Go to `https://<WebAppInstanceName>.scm.<location>.azurewebsites.net/ZipDeployUI`.

    For example: `https://dkeservice-ctaugtfwh7d3c3c.scm.westcentralus-01.azurewebsites.net/ZipDeployUI`

2. Launch a PowerShell prompt.

3. In the codebase for the key store, go to the **customer-key-store\src\customer-key-store** folder, and verify that this folder contains the **customerkeystore.csproj** file.

4. Run: **dotnet publish**

    The output window displays the directory where the publish was deployed.

    For example: `customer-key-store\src\customer-key-store\bin\Debug\net8.0\publish\`

5. Send all files in the publish directory to a .zip file. When creating the .zip file, make sure that all files in the directory are at the root level of the .zip file.

6. Drag and drop the .zip file you created to the ZipDeployUI site you opened before.

DKE is deployed and you can browse to the test keys you created. Continue to Validate your deployment in this article.

Validate your deployment

After deploying DKE using one of the methods described in this article, validate the deployment and the key store settings.

Run:

```PowerShell
src\customer-key-store\scripts\key_store_tester.ps1 dkeserviceurl/mykey
```

For example:

PowerShell

```
key_store_tester.ps1 https://dkeservice.contoso.com/TestKey1
```

Ensure that no errors appear in the output. When you're ready, register your key store.

The key name is case-sensitive. Enter the key name as it appears in the appsettings.json file.

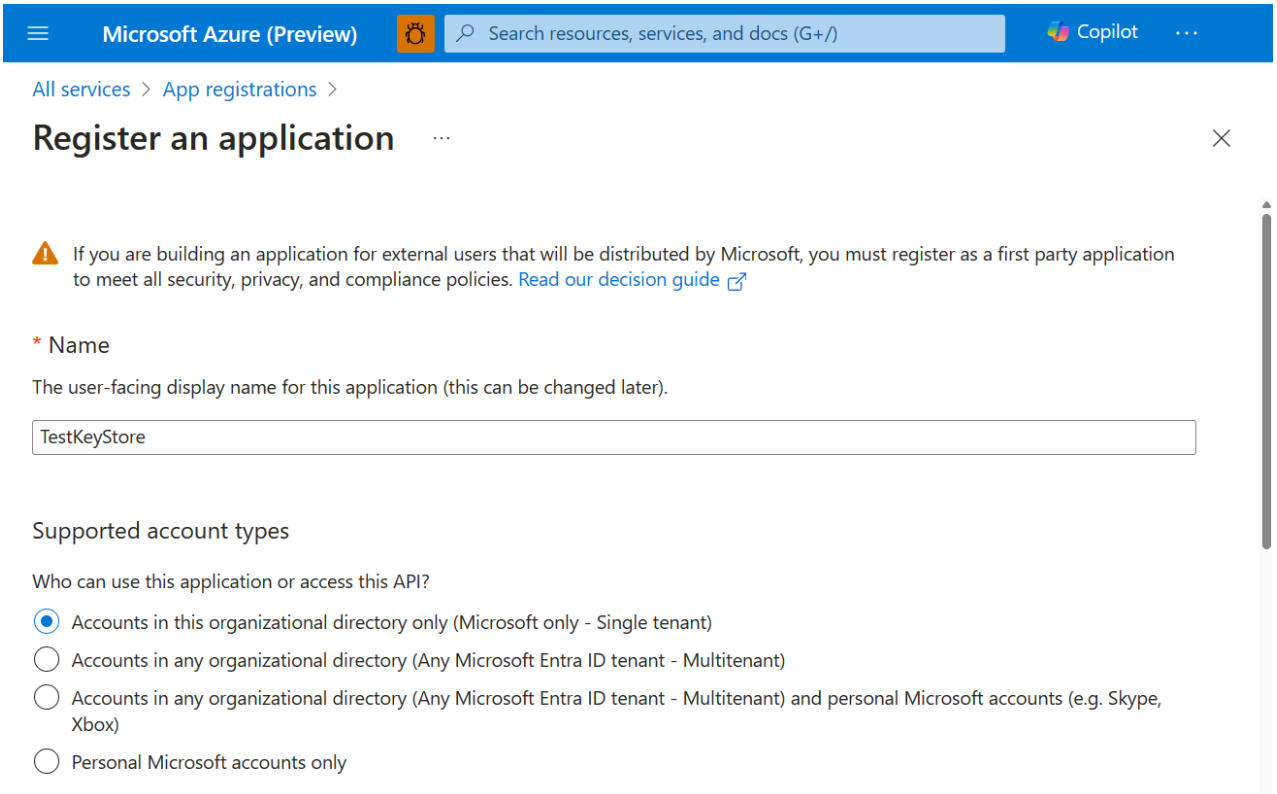
Register your key store

The following steps enable you to register your DKE service. Registering your DKE service is the last step in deploying DKE before you can start creating labels.

To register the DKE service:

1. In your browser, open the Microsoft Azure portal, and go to **All Services** > **Identity** > **App Registrations**.
2. Select **New registration**, and enter a meaningful name.
3. Select an account type from the options displayed.

   For example:



4. At the bottom of the page, select **Register** to create the new App Registration.
5. In your new App Registration, in the left pane, under **Manage**, select **Authentication**.

6. Select **Add a platform**.

7. On the **Configure platforms** popup, select **Web**.

8. Under **Redirect URIs**, enter the URI of your double key encryption service. Enter the App Service URL, including both the hostname and domain.

   For example: `https://mydkeservicetest.com`

   - The URL you enter must match the hostname where your DKE service is deployed.
   - The domain must be a verified domain.
   - In all cases, the scheme must be **https**.

   Ensure the hostname exactly matches your App Service hostname.

9. Under **Implicit grant**, select the **ID tokens** checkbox.

10. Select **Save** to save your changes.

11. On the left pane, select **Expose an API**, next to Application ID URI, enter your App Service URL, including both hostname and domain, and then select **Set**.

12. Still on the **Expose an API** page, in the **Scopes defined by this API** area, select **Add a scope**. In the new scope:

    a. Define the scope name as **user_impersonation**.

    b. Select the administrators and users who can consent.

    c. Define any remaining values required.

    d. Select **Add scope**.

    e. Select **Save** at the top to save your changes.

13. Still on the **Expose an API** page, in the **Authorized client applications** area, select **Add a client application**.

    In the new client application:

    a. Define the Client ID as `d3590ed6-52b3-4102-aeff-aad2292ab01c`. This value is the Microsoft Office client ID, and enables Office to obtain an access token for your key store.

    b. Under **Authorized scopes**, select the **user_impersonation** scope.

    c. Select **Add application**.

d. Select **Save** at the top to save your changes.

e. Repeat these steps, but this time, if you still use Azure Information Protection client, define the client ID as `c00e9d32-3c8d-4a7d-832b-029040e7db99`.

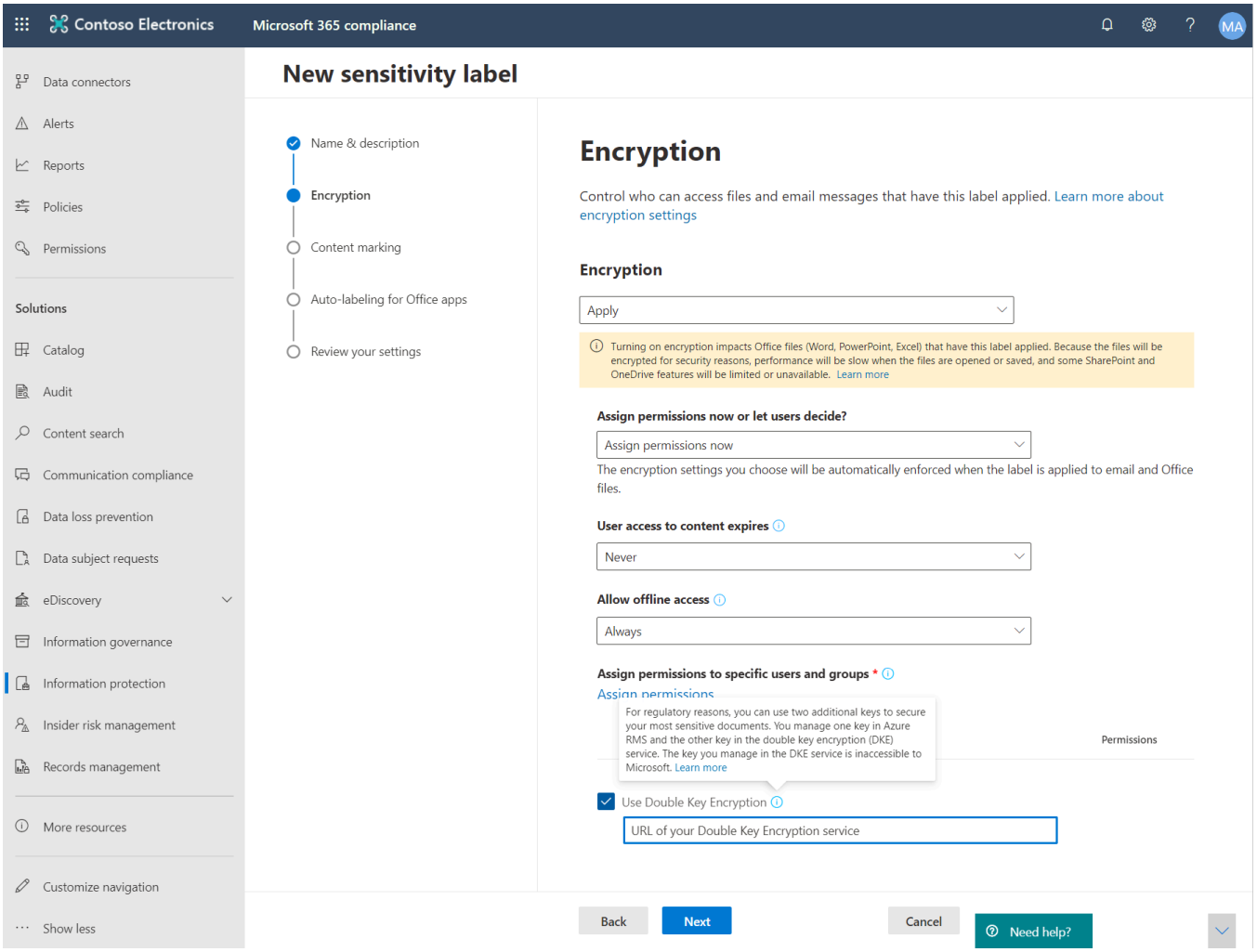
f. Repeat these steps, but this time, define the client ID as `0e25d998-b19a-4080-811c-d74d60d65e42`. This value is the Purview Information Protection client ID.

Your DKE service is now registered. Continue by creating labels using DKE.

Create sensitivity labels using DKE

In the Microsoft Purview portal, create a new sensitivity label and apply encryption as you would otherwise. Select **Use Double Key Encryption** and enter the endpoint URL for your key. You need to include the key name you provided in the "TestKeys" section of the appsettings.json file in the URL.

For example: `https://testingdke1.azurewebsites.net/KEYNAME`



Any DKE labels you add appear for users in the latest versions of Microsoft 365 Apps for enterprise.

> **Note**
>
> It may take up to 24 hours for the clients to refresh with the new labels.

Migrate protected files from HYOK labels to DKE labels

If you want, once you finish setting up DKE, you can migrate content that you protected using HYOK labels to DKE labels. To migrate, use the Microsoft Purview Information Protection scanner. To get started using the scanner, see Understand the information protection scanner.

If you don't migrate content, your HYOK protected content remains unaffected.

Other deployment options

We realize that for some customers in highly regulated industries, this standard reference implementation using software-based keys might not be sufficient to meet their enhanced compliance obligations and needs. We partner with third-party hardware security module (HSM) vendors to support enhanced key management options in the DKE service, including:

- Entrust
- Thales
- Utimaco

Reach out directly to these vendors for more information and guidance on their in-market DKE HSM solutions.

Set up clients to apply DKE sensitivity labels

On each client device, complete this task.

1. Ensure the following registry values are defined on each client. Create registry keys that aren't already there:

   Console

   ```
   [HKEY_LOCAL_MACHINE\SOFTWARE\WOW6432Node\Microsoft\MSIPC\flighting]
   "DoubleKeyProtection"=dword:00000001
   ```

```
[HKEY_LOCAL_MACHINE\SOFTWARE\Microsoft\MSIPC\flighting]
"DoubleKeyProtection"=dword:00000001
```

Double Key Encryption FAQ

"Have questions about how Double Key Encryption works that we didn't cover elsewhere? Check for an answer here."

What Microsoft 365 Apps can I use with DKE?

You can use DKE labels to protect documents using the desktop versions of Word, Excel, PowerPoint, and Outlook on Windows. To ensure that you're using a supported version of Office apps, see the capabilities tables and the row **Double Key Encryption (DKE)**.

Can I use Double Key Encryption with Microsoft Office built-in sensitivity labeling?

Yes! Use can use built-in sensitivity labeling with Office apps. For information, see the capabilities tables and the row **Double Key Encryption (DKE)**. While you can use the information protection client to protect documents with Double Key Encryption for now, this method will be deprecated in the future.

How is Double Key Encryption different from the existing hold your own key (HYOK) solution?

Double Key Encryption encrypts your data with two keys. Your encryption key is in your control and the second key is stored in Microsoft Azure, allowing you to move your encrypted data to the cloud. HYOK protects your content with only one key and the key is always on premises.

Can Double Key Encrypted documents be shared externally?

You can share Double Key Encrypted documents with users on a separate tenant as long as those users:

- Have the required permission to access your key in your Double Key Encryption service.

- Have the required permission to access your key in Microsoft Azure.

What happens to documents that are protected with HYOK?

Deploying Double Key Encryption doesn't affect your existing HYOK setup. However, we recommend that you start using Double Key Encryption in parallel with HYOK.

Can I run Double Key Encryption in my non-Microsoft air-gapped environment?

DKE doesn't support these environments because the service requires access to Microsoft Azure.

Where can I store Double Key Encrypted documents?

You can store Double Key Encrypted documents on-premises or in the cloud. In the cloud, you can move encrypted content to SharePoint Online and OneDrive for Business. You can't view the encrypted documents online in Office Web Apps.

What regions and languages is Double Key Encryption available in? Is Double Key Encryption available worldwide?

DKE labels are localized to the same languages as other sensitivity labels in Microsoft Purview Information Protection. Double Key Encryption is available worldwide.

Can I convert a non-DKE label to a DKE label?

No. You can’t add DKE to a label after you create it. Instead, you must choose **Use Double Key Encryption** and provide the URL of your Double Key Encryption service when you create the label.

How do I roll my DKE keys?

For instructions on rolling (also called rotating or rekeying) the key you store in Azure, see Operations for your tenant key. See Tenant and key settings for information on creating a new key for the DKE service. When you create a key, you set up a name and a GUID. Then, if you rotate a key, you keep the old record with the name and GUID but add a new record with the same name but different GUID. The new key gets set as active so that the public key API starts returning it for new encryption. Both keys are available for decryption so that new content and old content can be decrypted.

Why can't I access Double Key Encrypted documents?

The right to view DKE-protected content requires users to authenticate to endpoints managed by your organization. If users in your organization can't view DKE-protected content, review the key access settings configuration.

What kind of mutual transport layer security (mTLS) is supported with the DKE service on Office Desktop Apps on Windows?

- Before 2402: No mTLS support.
- After 2402: The apps don’t support sending the client certification. Instead the desktop apps send the request with `WINHTTP_NO_CLIENT_CERT_CONTEXT`. For more information, see SSL in WinHTTP - Win32 apps.

Last updated on 11/18/2025

Set up Information Rights Management (IRM) in SharePoint admin center

> *Microsoft Purview service description*

Within SharePoint in Microsoft 365, IRM protection is applied to files at the list and library level. Before your organization can use IRM protection, you must first set up Rights Management. IRM relies on the Azure Rights Management service from Azure Information Protection to encrypt and assign usage restrictions. Some Microsoft 365 plans include Azure Rights Management, but not all.

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

Turn on IRM service using SharePoint admin center

Before your organization can IRM-protect SharePoint lists and libraries, you must first activate the Rights Management service for your organization. To learn how, see Activating Azure Rights Management.

After you activate the Rights Management service, sign in to the SharePoint admin center to turn on IRM.

1. Sign in to SharePoint using a work or school account that has site owner or SharePoint administrator permissions.
2. Select the app launcher icon and choose **Admin** to open the Microsoft 365 admin center. (If you don't see the Admin tile, your work or school account doesn't have administrator permissions in your organization.)
3. In the left pane, choose **Admin centers** > SharePoint admin center.
4. In the left pane, choose **settings**, and then choose **classic settings page**.

5. In the **Information Rights Management (IRM)** section, choose **Use the IRM service specified in your configuration**, and then choose **Refresh IRM Settings**. After you refresh IRM settings, people in your organization can begin using IRM in their SharePoint lists and document libraries. However, the options to do so might take up to an hour to appear in Library Settings and List Settings.

IRM-enable SharePoint document libraries and lists

After IRM settings refresh, site owners can IRM-protect their SharePoint lists and document libraries. For more information, see Apply Information Rights Management to a list or library.

When site owners enable IRM for a list or library, they can protect any supported file types in that list or library. When IRM is enabled for a library, rights management applies to all of the files in that library. When you enable IRM for a list, rights management applies only to files that are attached to list items, not the actual list items.

When people download files in an IRM-enabled list or library, the files are encrypted so that only authorized people can view them. Each rights-managed file also contains an issuance license that imposes restrictions on the people who view the file. Typical restrictions include:

- Making a file read-only
- Disabling the copying of text
- Preventing people from saving a local copy
- Preventing people from printing the file

Client programs that can read IRM-supported file types use the issuance license within the rights-managed file to enforce these restrictions. So, a rights-managed file retains its protection even after download. To enable IRM on a list or library, see Apply Information Rights Management to a list or library.

You can't create or edit documents in an IRM-enabled library using Office in a browser. Instead, one person at a time can download and edit IRM-encrypted files. Use check-in and check-out to manage *coauthor*, or authoring across multiple users.

When you download a PDF file from an IRM-protected library, Microsoft 365 creates a protected PDF file. The file's extension doesn't change, but the file is protected. To view this file, you need the information protection viewer, the full information protection client, or another application that supports viewing protected PDF files.

SharePoint supports encryption of the following file types:

- PDF
- The 97-2003 file formats for the following Microsoft Office programs: Word, Excel, and PowerPoint
- The Office Open XML formats for the following Microsoft Office programs: Word, Excel, and PowerPoint
- The XML Paper Specification (XPS) format

> **Note**
>
> IRM protection can't be applied to protected documents (like digitally signed PDF files) as SharePoint needs to open the document on upload.

Next steps

Once you enabled IRM for SharePoint, you can start applying rights management to lists and libraries. For information, see Apply Information Rights Management to a list or library.

The OneDrive sync client for Windows supports synchronizing IRM-protected SharePoint document libraries and OneDrive locations (as long as the IRM setting for the library isn't set to expire document access rights). For more information, or to get started deploying the sync client, see Deploy the new OneDrive sync client for Windows.

Last updated on 05/14/2026

Technical reference details about encryption

09/26/2025

Refer to this article for information about certificates, technologies, and TLS cipher suites used for encryption in Microsoft 365. This article also provides details about planned deprecations.

- If you're looking for overview information, see Encryption in Microsoft 365.
- If you're looking for setup information, see Set up encryption in Microsoft 365 Enterprise.
- For specific information about TLS 1.1 and 1.0 deprecation, see Disabling TLS 1.0 and 1.1 for Microsoft 365.
- For information about cipher suites supported by specific versions of Windows, see Cipher Suites in TLS/SSL (Schannel SSP).
- For certificate chains, see Microsoft 365 encryption chains and Microsoft 365 encryption chains - DOD and GCC High.

Microsoft Office 365 certificate ownership and management

You don't need to purchase or maintain certificates for Office 365. Instead, Office 365 uses its own certificates.

Current encryption standards and planned deprecations

To provide best-in-class encryption, Office 365 regularly reviews supported encryption standards. Sometimes, old standards are deprecated as they become out of date and less secure. This article describes currently supported cipher suites and other standards and details about planned deprecations.

FIPS compliance for Microsoft 365

All cipher suites supported by Office 365 use algorithms acceptable under FIPS 140-2. Office 365 inherits FIPS validations from Windows (through Schannel). For information about Schannel, see Cipher Suites in TLS/SSL (Schannel SSP).

AES256-CBC support for Microsoft 365

In late August 2023, Microsoft Purview Information Protection will begin to use Advanced Encryption Standard (AES) with 256-bit key length in Cipher Block Chaining mode (AES256-CBC). By October 2023, AES256-CBC will be the default for encryption of Microsoft 365 Apps documents and emails. You may need to take action to support this change in your organization.

Who is impacted and what do I need to do?

Use this table to figure out if you have to take action:

Expand table

Client applications	Service applications	Action required?	What do I need to do?
Microsoft 365 Apps	Exchange Online, SharePoint Online	No	N/A
Office 2013, 2016, 2019, or 2021	Exchange Online, SharePoint Online	Yes (Optional)	See Set up Office 2013, 2016, 2019, or 2021 for AES256-CBC mode.
Microsoft 365 Apps	Exchange Server or hybrid	Yes (Mandatory)	See Set up Exchange Server for AES256-CBC support.
Office 2013, 2016, 2019, or 2021	Exchange Server or hybrid	Yes (Mandatory)	Complete Option 1 (required), and then see Set up Office 2013, 2016, 2019, or 2021 for AES256-CBC mode.
Microsoft 365 Apps	MIP SDK	Yes (Optional)	See Set up MIP SDK for AES256-CBC support.
Any	SharePoint Server	No	N/A

Set up Office 2013, 2016, 2019, or 2021 for AES256-CBC mode

You need to configure Office 2013, 2016, 2019, or 2021 to use AES256-CBC mode using Group Policy, or by using the Cloud Policy service for Microsoft 365. Starting with version 16.0.16227 of Microsoft 365 Apps, CBC mode is used by default. Use the `Encryption mode for Information Rights Management (IRM)` setting under `User Configuration/Administrative Templates/Microsoft Office 2016/Security Settings`.

For example, to force CBC mode, select the group policy setting as follows:

Encryption mode for Information Rights Management (IRM): [1, Cipher Block Chaining (CBC)]

Set up Exchange Server for AES256-CBC support

Exchange Server doesn't support decrypting content that uses AES256-CBC. To work around this problem, you have two options.

Option 1

Customers using Exchange Online with the Azure Rights Management Connector service deployed will be opted out of the AES256-CBC publishing change in both Exchange Online and SharePoint Online.

To move to AES256-CBC mode, complete these steps:

1. Install the hotfix on your Exchange Servers when it becomes available. For the most recent information about ship dates, see the Microsoft 365 product roadmap.
2. If you're using Exchange Server with the Azure Rights Management Connector Service, you'll need to run the GenConnectorConfig.ps1 script on each Exchange server. For more information, see Configure servers for the Rights Management connector.

Once your organization has installed the patch across all of your Exchange Servers, open a support case and request these services to be enabled for AES256-CBC publishing.

Option 2

This option gives you some extra time before you need to patch all of your Exchange servers. Use this option if you're unable to complete the steps in option 1 when the hotfix becomes available. Instead, deploy group policy or client settings that force Microsoft 365 clients to keep using AES128-ECB mode. Deploy this setting using Group Policy, or by using the Cloud Policy service for Microsoft 365. You can configure Office and Microsoft 365 Apps for Windows to use ECB or CBC mode with the `Encryption mode for Information Rights Management (IRM)` setting under `User Configuration/Administrative Templates/Microsoft Office 2016/Security Settings`. Starting with version 16.0.16327 of Microsoft 365 Apps, CBC mode is used by default.

For example, to force EBC mode for Windows clients, set the group policy setting as follows:

Encryption mode for Information Rights Management (IRM): [2, Electronic Codebook (ECB)]

To configure settings for Office for Mac clients, see Set suite-wide preferences for Office for Mac.

As soon as you can, complete the steps in option 1.

Set up MIP SDK for AES256-CBC support

Update to MIP SDK 1.13 or later. If you choose to update to MIP SDK 1.13, you'll need to configure a setting to force AES256-CBC. For more information, see the MIP SDK Version 1.13.158 Critical Update. Later versions of the MIP SDK will protect Microsoft 365 files and email with AES256-CBC by default.

Versions of TLS supported by Microsoft 365

TLS, and SSL that came before TLS, are cryptographic protocols that secure communication over a network by using security certificates to encrypt a connection between computers. All Microsoft 365 support TLS version 1.2 (TLS 1.2). Support for TLS version 1.3 (TLS 1.3) is being rolled out throughout the service by the various applications and services. Exchange Online now supports TLS 1.3 for all email submissions and server communications with third-parties on the internet for worldwide customers. There is no timeline for TLS 1.3 for Exchange Online operated by 21Vianet.

> **Important**
>
> Be aware that TLS versions deprecate, and that deprecated versions *should not be used* where newer versions are available. If your legacy services do not require TLS 1.0 or 1.1 you should disable them.

Deprecations

- Cipher suites: We are constantly reviewing our supported cipher suite list. We have stop supporting weak SHA1 message authentication, and the non-forward-secrecy RSA key exchange algorithm in the past. The current list of supported ciphers can be seen here: TLS cipher suites supported by Microsoft 365.
- TLS 1.0 and TLS 1.1 versions: Support for these TLS versions ended on October 31, 2018 with their removal from the service completed since 2022. All connections with Microsoft 365 now use at least TLS 1.2 to communicate. One exception is the SMTP AUTH client submission protocol in Exchange Online which offers an opt-in endpoint for customers with legacy devices that still need TLS 1.0 or TLS 1.1.
- 3DES algorithm: Support for this encryption algorithm ended on October 31, 2018. Microsoft 365 removed the cipher suite using it, TLS_RSA_WITH_3DES_EDE_CBC_SHA cipher suite, from the service from February 28, 2019. For a list of supported ciphers, see TLS cipher suites supported by Microsoft 365.
- SHA-1 certificates: Support in Microsoft 365 for communications where the other third-party provides SHA-1 certificates ended from June 2016. SHA-2 (Secure Hash Algorithm

2) or stronger hashing algorithms are now needed in the certificate chain.

TLS cipher suites supported by Microsoft 365

TLS uses *cipher suites*, collections of encryption algorithms, to establish secure connections. Microsoft 365 supports the cipher suites listed in the following table. The table lists the cipher suites in order of strength, with the strongest cipher suite listed first.

Microsoft 365 responds to a connection request by first attempting to connect using the most secure cipher suite. If the connection doesn't work, Microsoft 365 tries the second most secure cipher suite in the list, and so on. The service continues down the list until the connection is accepted. Likewise, when Microsoft 365 requests a connection, the receiving service chooses whether to use TLS and which cipher suite to use.

Expand table

Cipher suite name	TLS protocol version
TLS_AES_256_GCM_SHA384	1.3
TLS_AES_128_GCM_SHA256	1.3
TLS_ECDHE_RSA_WITH_AES_256_GCM_SHA384	1.2
TLS_ECDHE_RSA_WITH_AES_128_GCM_SHA256	1.2
TLS_ECDHE_RSA_WITH_AES_256_CBC_SHA384	1.2
TLS_ECDHE_RSA_WITH_AES_128_CBC_SHA256	1.2

Related articles

Cipher Suites in TLS/SSL (Schannel SSP)

Encryption in Office 365

Preparing for TLS 1.2 in Office 365 and Office 365 GCC

Azure Certificate Authority details

This article outlines the specific root and subordinate Certificate Authorities (CAs) that Azure service endpoints use. This CA list is distinct from the trust anchors on Azure VMs and hosted services, which use the trust anchors provided by the operating systems. The scope includes government and national clouds. The sections after the CA details tables provide the minimum requirements for public key encryption and signature algorithms, links to certificate downloads and revocation lists, and information about key concepts. These sections also provide the host names for the URIs that you should add to your firewall allow lists.

Certificate Authority details

Any entity trying to access Microsoft Entra identity services through the TLS/SSL protocols receives certificates from the CAs listed in this article. Different services might use different root or intermediate CAs. The following root and subordinate CAs are relevant to entities that use certificate pinning.

How to read the certificate details:

- The Serial Number (top string in the table) contains the hexadecimal value of the certificate serial number.
- The Thumbprint (bottom string in the table) is the SHA1 thumbprint.
- For CAs that you must download over HTTP instead of HTTPS, copy the link and paste it into the URL field in a new tab.

How to find certificate validity (expiry) dates:

To determine when a certificate expires, download the certificate file and view its properties. On Windows, double-click the downloaded `.crt` file and check the **Valid from** and **Valid to** fields under the **General** tab. On Linux or macOS, use `openssl x509 -in <certificate-file> -noout -dates` to display the validity period. Root CAs typically have validity periods of 15 to 25 years, while subordinate CAs have shorter validity periods of 3 to 10 years.

Root and Subordinate CAs list

Root Certificate Authorities

Expand table

Certificate Authority	Download URL	Serial Number Thumbprint
DigiCert Global Root CA	`https://cacerts.digicert.com/DigiCertGlobalRootCA.crt`	0x083be05690424 A8985D3A65E5E5(
DigiCert Global Root G2	`https://cacerts.digicert.com/DigiCertGlobalRootG2.crt`	0x033af1e6a711a9 DF3C24F9BFD666
DigiCert Global Root G3	`https://cacerts.digicert.com/DigiCertGlobalRootG3.crt`	0x055556bcf25ea4 7E04DE896A3E666
DigiCert TLS ECC P384 Root G5	`https://cacerts.digicert.com/DigiCertTLSECCP384RootG5.crt`	0x09e09365acf7d9 17F3DE5E9F0F19E
DigiCert TLS RSA	`https://cacerts.digicert.com/DigiCertTLSRSA4096RootG5.crt`	0x08f9b478a8fa7e A78849DC5D7C75

Certificate Authority	Download URL	Serial Number Thumbprint
4096 Root G5		
Microsoft ECC Root Certificate Authority 2017	`https://www.microsoft.com/pkiops/certs/Microsoft%20ECC%20Root%20Certificate%20Authority%202017.crt`	0x66f23daf87de8b 999A64C37FF47D9
Microsoft RSA Root Certificate Authority 2017	`https://www.microsoft.com/pkiops/certs/Microsoft%20RSA%20Root%20Certificate%20Authority%202017.crt`	0x1ed397095fd8b4 73A5E64A3BFF831

Subordinate Certificate Authorities

Expand table

Certificate Authority	Download URL	Serial Number Thumbprint
DigiCert Basic OV G2 TLS CN RSA406 SHA256 2022 CA1	`http://cacerts.digicert.com/DigiCertBasicOVG2TLSCNRSA4096SHA2562022CA1.crt`	0x02361857392757c7ce4 60707270F2100EE2B771I
DigiCert Basic OV G3 TLS CN ECC P-384 SHA384 2022 CA1	`http://cacerts.digicert.com/DigiCertBasicOVG3TLSCNECCP-384SHA3842022CA1.crt`	0x01484637b5aacf53c27 9355B7783FFF5DF69FE98
DigiCert Basic RSA CN CA G2	`https://cacerts.digicert.com/DigiCertBasicRSACNCAG2.crt`	0x02f7e1f982bad009aff4 4D1FA5D1FB1AC3917C0
DigiCert Cloud Services CA-1	`https://cacerts.digicert.com/DigiCertCloudServicesCA-1.crt`	0x019ec1c6bd3f597bb20 81B68D6CD2F221F8F534
DigiCert Global G2 TLS RSA SHA256 2020 CA1	`http://cacerts.digicert.com/DigiCertGlobalG2TLSRSASHA2562020CA1-1.crt`	0cf5bd062b5602f47ab85 1B511ABEAD59C6CE207(
DigiCert Global G3 TLS ECC SHA384 2020 CA1	`https://cacerts.digicert.com/DigiCertGlobalG3TLSECCSHA3842020CA1-2.crt`	0b00e92d4d6d731fca305 9577F91FE86C27D99121
DigiCert SHA2 Secure Server CA	`http://cacerts.digicert.com/DigiCertSHA2SecureServerCA-2.crt`	0x02742eaa17ca8e21c71 626D44E704D1CEABE3BI
DigiCert TLS Hybrid ECC SHA384 2020 CA1	`http://cacerts.digicert.com/DigiCertTLSHybridECCSHA3842020CA1-1.crt`	07f2f35c87a877af7aefe94 aec13cdd5ea6a3998aec1
DigiCert TLS RSA SHA256 2020 CA1	`https://cacerts.digicert.com/DigiCertTLSRSASHA2562020CA1-1.crt`	0x06d8d904d5584346f68 1C58A3A8518E8759BF07

Certificate Authority	Download URL	Serial Number Thumbprint
GeoTrust Global TLS RSA4096 SHA256 2022 CA1	`https://cacerts.digicert.com/GeoTrustGlobalTLSRSA4096SHA2562022CA1.crt`	0x0f622f6f21c2ff5d521f7 7E6DB7B7584D8CF2003E
Microsoft Azure ECC TLS Issuing CA 03	`https://www.microsoft.com/pkiops/certs/Microsoft%20Azure%20ECC%20TLS%20Issuing%20CA%2003%20-%20xsign.crt`	0x01529ee8368f0b5d72b 56D955C849887874AA17
Microsoft Azure ECC TLS Issuing CA 03	`https://www.microsoft.com/pkiops/certs/Microsoft%20Azure%20ECC%20TLS%20Issuing%20CA%2003.crt`	0x330000003322a2579b5 91503BE7BF74E2A10AA0
Microsoft Azure ECC TLS Issuing CA 04	`https://www.microsoft.com/pkiops/certs/Microsoft%20Azure%20ECC%20TLS%20Issuing%20CA%2004%20-%20xsign.crt`	0x02393d48d702425a7cb FB73FDC24F06998E070A
Microsoft Azure ECC TLS Issuing CA 04	`https://www.microsoft.com/pkiops/certs/Microsoft%20Azure%20ECC%20TLS%20Issuing%20CA%2004.crt`	0x33000000322164aedab 406E3B38EFF35A727F276
Microsoft Azure ECC TLS Issuing CA 07	`https://www.microsoft.com/pkiops/certs/Microsoft%20Azure%20ECC%20TLS%20Issuing%20CA%2007%20-%20xsign.crt`	0x0f1f157582cdcd33734b 3BE6CA5856E3B9709056
Microsoft Azure ECC TLS Issuing CA 07	`https://www.microsoft.com/pkiops/certs/Microsoft%20Azure%20ECC%20TLS%20Issuing%20CA%2007.crt`	0x3300000034c732435db AB3490B7E37B3A8A1E71
Microsoft Azure ECC TLS Issuing CA 08	`https://www.microsoft.com/pkiops/certs/Microsoft%20Azure%20ECC%20TLS%20Issuing%20CA%2008%20-%20xsign.crt`	0x0ef2e5d83681520255e 716DF84638AC8E6EEBE6
Microsoft Azure ECC TLS Issuing CA 08	`https://www.microsoft.com/pkiops/certs/Microsoft%20Azure%20ECC%20TLS%20Issuing%20CA%2008.crt`	0x330000031526979844 CF33D5A1C2F0355B207F
Microsoft Azure RSA TLS Issuing CA 03	`https://www.microsoft.com/pkiops/certs/Microsoft%20Azure%20RSA%20TLS%20Issuing%20CA%2003%20-%20xsign.crt`	0x05196526449a5e3d1a3 F9388EA2C9B7D632B66A
Microsoft Azure RSA TLS Issuing CA 03	`https://www.microsoft.com/pkiops/certs/Microsoft%20Azure%20RSA%20TLS%20Issuing%20CA%2003.crt`	0x330000003968ea517d8 37461AACFA5970F7F2D2
Microsoft Azure RSA TLS Issuing CA 04	`https://www.microsoft.com/pkiops/certs/Microsoft%20Azure%20RSA%20TLS%20Issuing%20CA%2004%20-%20xsign.crt`	0x09f96ec295555f24749e BE68D0ADAA2345B48E5
Microsoft Azure RSA TLS Issuing CA 04	`https://www.microsoft.com/pkiops/certs/Microsoft%20Azure%20RSA%20TLS%20Issuing%20CA%2004.crt`	0x330000003cd7cb44ee5 7304022CA8A9FF7E3E0C
Microsoft Azure RSA TLS Issuing CA 07	`https://www.microsoft.com/pkiops/certs/Microsoft%20Azure%20RSA%20TLS%20Issuing%20CA%2007%20-%20xsign.crt`	0x0a43a9509b01352f899 3382517058A0C20228D5
Microsoft Azure RSA	`https://www.microsoft.com/pkiops/certs/Microsoft%20Azure%20RSA%20TLS%20Issuing%20CA%2007.crt`	0x330000003bf980b0c83 0E5F41B697DAADD808B

Certificate Authority	Download URL	Serial Number Thumbprint
TLS Issuing CA 07		
Microsoft Azure RSA TLS Issuing CA 08	`https://www.microsoft.com/pkiops/certs/Microsoft%20Azure%20RSA%20TLS%20Issuing%20CA%2008%20-%20xsign.crt`	0x0efb7e547edf0ff1069a 31600991ED5FEC63D355
Microsoft Azure RSA TLS Issuing CA 08	`https://www.microsoft.com/pkiops/certs/Microsoft%20Azure%20RSA%20TLS%20Issuing%20CA%2008.crt`	0x330000003a5dc2ffc321 512C8F3FB71EDACF7AD/
Microsoft TLS ECC Root G2	`http://caissuers.microsoft.com/pkiops/certs/Microsoft%20TLS%20ECC%20Root%20G2%20-%20xsign.crt`	0x08d3c6d001f26cb5a23 7FA05ACF9B4AABB0DE0
Microsoft TLS G2 ECC CA OCSP 02	`https://www.microsoft.com/pkiops/certs/Microsoft%20TLS%20G2%20ECC%20CA%20OCSP%2002.crt`	0x3300000005516a21c89 336D3E5451315294F6E1
Microsoft TLS G2 ECC CA OCSP 04	`https://www.microsoft.com/pkiops/certs/Microsoft%20TLS%20G2%20ECC%20CA%20OCSP%2004.crt`	0x3300000008d690fff489 5BFF3CC51FB9C1CCE0D5
Microsoft TLS G2 ECC CA OCSP 06	`https://www.microsoft.com/pkiops/certs/Microsoft%20TLS%20G2%20ECC%20CA%20OCSP%2006.crt`	0x33000000066fa24c31c 9BC57546ACBDB29E388(
Microsoft TLS G2 ECC CA OCSP 08	`https://www.microsoft.com/pkiops/certs/Microsoft%20TLS%20G2%20ECC%20CA%20OCSP%2008.crt`	0x3300000007737deb49 F1C0A1828C10B43F9C65
Microsoft TLS G2 RSA CA OCSP 02	`https://www.microsoft.com/pkiops/certs/Microsoft%20TLS%20G2%20RSA%20CA%20OCSP%2002.crt`	0x330000000c4964a16f4 F8A408B8FBEEA077B757
Microsoft TLS G2 RSA CA OCSP 04	`https://www.microsoft.com/pkiops/certs/Microsoft%20TLS%20G2%20RSA%20CA%20OCSP%2004.crt`	0x330000000b13e5667d4 DA6D0400641B45AECC5
Microsoft TLS G2 RSA CA OCSP 06	`https://www.microsoft.com/pkiops/certs/Microsoft%20TLS%20G2%20RSA%20CA%20OCSP%2006.crt`	0x330000000def69cd2d7 88283C78D08489718F78
Microsoft TLS G2 RSA CA OCSP 08	`https://www.microsoft.com/pkiops/certs/Microsoft%20TLS%20G2%20RSA%20CA%20OCSP%2008.crt`	0x3300000011f100a7eb0 DDBD3DA5C09386CC360
Microsoft TLS G2 RSA CA OCSP 10	`https://www.microsoft.com/pkiops/certs/Microsoft%20TLS%20G2%20RSA%20CA%20OCSP%2010.crt`	0x330000000f33206537e B663D8413CA35F3F2264
Microsoft TLS G2 RSA CA OCSP 12	`https://www.microsoft.com/pkiops/certs/Microsoft%20TLS%20G2%20RSA%20CA%20OCSP%2012.crt`	0x33000000107d59860fc D3089DF9F474908870454
Microsoft TLS G2 RSA CA OCSP 14	`https://www.microsoft.com/pkiops/certs/Microsoft%20TLS%20G2%20RSA%20CA%20OCSP%2014.crt`	0x33000000127e739347e 55D9B33BDC16DDB3B66
Microsoft TLS G2	`https://www.microsoft.com/pkiops/certs/Microsoft%20TLS%20G2%20RSA%20CA%20OCSP%2016.crt`	0x330000000e42938f25d 27E6663299D340DFE545

Certificate Authority	Download URL	Serial Number Thumbprint
RSA CA OCSP 16		
Microsoft TLS RSA Root G2	`http://caissuers.microsoft.com/pkiops/certs/Microsoft%20TLS%20RSA%20Root%20G2%20-%20xsign.crt`	0x0b0c6b2c466917b047 B5EE89E77326AB2BF177

Client compatibility for public PKIs

The CAs used by Azure are compatible with the following OS versions:

Expand table

Windows	Firefox	iOS	macOS	Android	Java
Windows XP SP3+	Firefox 32+	iOS 7+	OS X Mavericks (10.9)+	Android SDK 5.x+	Java JRE 1.8.0_101+

Review the following action steps when CAs expire or change:

- Update to a supported version of the required OS.
- If you can't change the OS version, you might need to manually update the trusted root store to include the new CAs. See the documentation provided by the manufacturer.
- If your scenario includes disabling the trusted root store or running the Windows client in disconnected environments, ensure that all root CAs are included in the Trusted Root CA store and all sub CAs listed in this article are included in the Intermediate CA store.
- Many distributions of **Linux** require you to add CAs to `/etc/ssl/certs`. See the distribution documentation.
- Ensure that the **Java** key store contains the CAs listed in this article. For more information, see the Java applications section of this article.
- If your application explicitly specifies a list of acceptable CAs, check whether you need to update the pinned certificates when CAs change or expire. For more information, see Certificate pinning.

Public key encryption and signature algorithms

Support for the following algorithms, elliptical curves, and key sizes are required:

Signature algorithms:

- ES256
- ES384
- ES512
- RS256
- RS384
- RS512

Elliptical curves:

- P256
- P384
- P521

Key sizes:

- ECDSA 256
- ECDSA 384
- ECDSA 521
- RSA 2048

- RSA 3072
- RSA 4096

Certificate downloads and revocation lists

To optimize connectivity, add the following domains (HTTP on port 80) to your firewall allow list. These domains serve Authority Information Access (AIA), certificate revocation list (CRL), and Online Certificate Status Protocol (OCSP) requests. The certificate revocation and chain-building standards define these requests as HTTP rather than HTTPS, so port 80 is expected.

This article is the authoritative list of these domains. When Azure adds or removes a certificate authority, it updates the corresponding AIA, CRL, and OCSP domains here and records the change in the Article change log. If your environment restricts outbound connectivity by fully qualified domain name (FQDN), monitor this article and the change log so that you can update your firewall allow list.

AIA:

- `cacerts.digicert.com`
- `cacerts.digicert.cn`
- `cacerts.geotrust.com`
- `caissuers.microsoft.com`
- `www.microsoft.com`

CRL:

- `crl3.digicert.com`
- `crl4.digicert.com`
- `crl.digicert.cn`
- `www.microsoft.com`

OCSP:

- `ocsp.digicert.com`
- `ocsp.digicert.cn`
- `oneocsp.microsoft.com`

Certificate pinning

Certificate pinning is a security technique where you accept only authorized, or *pinned*, certificates when establishing a secure session. The process rejects any attempt to establish a secure session using a different certificate. For more information, including guidance on static pinning of publicly trusted certificates, see Certificate pinning.

How to address certificate pinning

If your application explicitly specifies a list of acceptable CAs, you might need to periodically update pinned certificates when Certificate Authorities change or expire.

To detect certificate pinning, take the following steps:

- If you're an application developer, search your source code for references to certificate thumbprints, Subject Distinguished Names, Common Names, serial numbers, public keys, and other certificate properties of any of the Sub CAs involved in this change.
  - If there's a match, update the application to include the missing CAs.
- If you have an application that integrates with Azure APIs or other Azure services and you're unsure whether it uses certificate pinning, check with the application vendor.

Java applications

To determine if your Java application trusts the **Microsoft ECC Root Certificate Authority 2017** and **Microsoft RSA Root Certificate Authority 2017** root certificates, check the list of trusted root certificates that the Java Virtual Machine (JVM) uses.

1. Open a terminal window on your system.
2. Run the following command:

   Bash

   ```
   keytool -list -keystore $JAVA_HOME/jre/lib/security/cacerts
   ```

   - `$JAVA_HOME` refers to the path to the Java home directory.
   - If you're unsure of the path, find it by running the following command:

   Bash

   ```
   readlink -f $(which java) | xargs dirname | xargs dirname
   ```

3. Look for the **Microsoft RSA Root Certificate Authority 2017** in the output. The output should look similar to the following sample:

   - If the **Microsoft ECC Root Certificate Authority 2017** and **Microsoft RSA Root Certificate Authority 2017** root certificates are trusted, they should appear in the list of trusted root certificates used by the JVM.
   - If the certificate isn't in the list, add it.

   Bash

   ```
   ...
   Microsoft ECC Root Certificate Authority 2017, 20-Aug-2022, Root CA,
   Microsoft RSA Root Certificate Authority 2017, 20-Aug-2022, Root CA,
   ...
   ```

4. To add a root certificate to the trusted root certificate store in Java, use the `keytool` utility. The following example adds the **Microsoft ECC Root Certificate Authority 2017** and **Microsoft RSA Root Certificate Authority 2017** root certificates:

   Bash

   ```
   keytool -import -file microsoft-ecc-root-ca.crt -alias microsoft-ecc-root-ca -
   keystore $JAVA_HOME/jre/lib/security/cacerts
   keytool -import -file microsoft-rsa-root-ca.crt -alias microsoft-rsa-root-ca -
   keystore $JAVA_HOME/jre/lib/security/cacerts
   ```

   > ⓘ Note
   >
   > In this example, `microsoft-ecc-root-ca.crt` and `microsoft-rsa-root-ca.crt` are the names of the files that contain the **Microsoft ECC Root Certificate Authority 2017** and **Microsoft RSA Root Certificate Authority 2017** root certificates, respectively.

Past changes

The CA/Browser Forum updated the Baseline Requirements to require all publicly trusted public key infrastructures (PKIs) to stop using the SHA-1 hash algorithms for Online Certificate Status Protocol (OCSP) on May 31, 2022. Microsoft updated all remaining OCSP responders that used the SHA-1 hash algorithm to use the SHA-256 hash algorithm.

Microsoft updated Azure services to use TLS certificates from a different set of Root Certificate Authorities (CAs) on February 15, 2021, to comply with changes set forth by the CA/Browser Forum Baseline Requirements. Some services finalized these updates in 2022. For the latest information regarding managed TLS and Azure, see Managed TLS changes.

Article change log

- September 09, 2025:
  - Added the following CAs:

Expand table

Certificate Authority	Serial Number Thumbprint
Microsoft TLS G2 ECC CA OCSP 04	0x3300000008d690fff489d9c4d7000000000008 5BFF3CC51FB9C1CCE0D52F754427347545E583C5
Microsoft TLS G2 ECC CA OCSP 08	0x3300000007737deb49d007dc64000000000007 F1C0A1828C10B43F9C651FDFC494832174BE8D07
Microsoft Azure ECC TLS Issuing CA 03	0x330000003322a2579b5e698bcc000000000033 91503BE7BF74E2A10AA078B48B71C3477175FEC3
Microsoft Azure ECC TLS Issuing CA 04	0x33000000322164aedab61f509d000000000032 406E3B38EFF35A727F276FE993590B70F8224AED
Microsoft Azure ECC TLS Issuing CA 07	0x3300000034c732435db22a0a2b000000000034 AB3490B7E37B3A8A1E715036522AB42652C3CFFE
Microsoft Azure ECC TLS Issuing CA 08	0x330000003152697984479 8bbb8000000000031 CF33D5A1C2F0355B207FCE940026E6C1580067FD

  - Removed the following CAs:

Expand table

Certificate Authority	Serial Number Thumbprint
Microsoft RSA TLS Issuing AOC CA 01	0x330000002ffaf06f6697e2469c00000000002f 4697FDBED95739B457B347056F8F16A975BAF8EE
Microsoft RSA TLS Issuing AOC CA 02	0x3300000030c756cc88f5c1e7eb000000000030 90ED2E9CB40D0CB49A20651033086B1EA2F76E0E
Microsoft RSA TLS Issuing EOC CA 01	0x33000000310c4914b18c8f339a000000000031 A04D3750DEBFCCF1259D553DBEC33162C6B42737
Microsoft RSA TLS Issuing EOC CA 02	0x3300000032444d7521341496a9000000000032 697C6404399CC4E7BB3C0D4A8328B71DD3205563
Microsoft ECC TLS Issuing AOC CA 01	0x33000000282bfd23e7d1add707000000000028 30AB5C33EB4B77D4CBFF00A11EE0A7507D9DD316
Microsoft ECC TLS Issuing AOC CA 02	0x33000000290f8a6222ef6a5695000000000029 3709CD92105D074349D00EA8327F7D5303D729C8
Microsoft ECC TLS Issuing EOC CA 01	0x330000002a2d006485fdacbfeb00000000002a 5FA13B879B2AD1B12E69D476E6CAD90D01013B46
Microsoft ECC TLS Issuing EOC CA 02	0x330000002be6902838672b667900000000002b 58A1D8B1056571D32BE6A7C77ED27F73081D6E7A
DigiCert TLS RSA SHA256 2020 CA1	0a3508d55c292b017df8ad65c00ff7e4 6938FD4D98BAB03FAADB97B34396831E3780AEA1

- September 5, 2025:
  - Added the following CAs:

Expand table

Certificate Authority	Serial Number Thumbprint
DigiCert Global G2 TLS RSA SHA256 2020 CA1	0cf5bd062b5602f47ab8502c23ccf066 1B511ABEAD59C6CE207077C0BF0E0043B1382612

Certificate Authority	Serial Number Thumbprint
DigiCert Global G3 TLS ECC SHA384 2020 CA1	0b00e92d4d6d731fca3059c7cb1e1886 9577F91FE86C27D9912129730E8166373FC2EEB8
DigiCert Basic OV G2 TLS CN RSA406 SHA256 2022 CA1	02361857392757c7ce46336cbc47cda2 60707270F2100EE2B771FEC9EFFAD8C9BFFE3358
DigiCert Basic OV G3 TLS CN ECC P-384 SHA384 2022 CA1	01484637b5aacf53c276a96852c2bf32 9355B7783FFF5DF69FE98FAF08F7B66CD4B94B99
Microsoft TLS G2 ECC CA OCSP 02	3300000005516a21c89814ccf7000000000005 336D3E5451315294F6E1E7CFDA84E51C8B1E4054
Microsoft TLS G2 ECC CA OCSP 06	33000000066fa24c31c3956fab000000000006 9BC57546ACBDB29E388CD1791B43DE4F659B844B
Microsoft TLS G2 RSA CA OCSP 02	330000000c4964a16f44203b2200000000000c F8A408B8FBEEA077B75738658641BE77154720EB
Microsoft TLS G2 RSA CA OCSP 04	330000000b13e5667d4a9b5580000000000000b DA6D0400641B45AECC595D24E5037AA6BC09C358
Microsoft TLS G2 RSA CA OCSP 06	330000000def69cd2d72237356000000000000d 88283C78D08489718F78CD2B797333765164E911
Microsoft TLS G2 RSA CA OCSP 08	3300000011f100a7eb05eba100000000000011 DDBD3DA5C09386CC36CFEC4680F17084BA7A5628
Microsoft TLS G2 RSA CA OCSP 10	330000000f33206537ee42ae4f00000000000f B663D8413CA35F3F22643CCBC2D3250700916B1A
Microsoft TLS G2 RSA CA OCSP 12	33000000107d59860fd64d7243000000000010 D3089DF9F47490887045B91D3FDEF6ED87F02F52
Microsoft TLS G2 RSA CA OCSP 14	33000000127e739347e75be1a1000000000012 55D9B33BDC16DDB3B663F1771F96D4324C118073
Microsoft TLS G2 RSA CA OCSP 16	330000000e42938f25dc51a99b00000000000e 27E6663299D340DFE54571CECCEC01BE74495DC6

- Removed the following CAs:

Expand table

Certificate Authority	Serial Number Thumbprint
Entrust Root Certification Authority G2	4a538c28 8CF427FD790C3AD166068DE81E57EFBB932272D4
Entrust Root Certification Authority G2	4a538c28 8CF427FD790C3AD166068DE81E57EFBB932272D4
└ Entrust Certification Authority - L1K	0ee94cc30000000051d37785 F21C12F46CDB6B2E16F09F9419CDFF328437B2D7
└ Entrust Certification Authority - L1M	61a1e7d20000000051d366a6 CC136695639065FAB47074D28C55314C66077E90
Entrust Certification Authority - L1K	0ee94cc30000000051d37785 F21C12F46CDB6B2E16F09F9419CDFF328437B2D7
Entrust Certification Authority - L1M	61a1e7d20000000051d366a6 CC136695639065FAB47074D28C55314C66077E90

- Updated AIA, CRL, and OCSP URLs to reflect recent CA changes.

- October 8, 2024: Removed the following CAs and CDP endpoints: crl.microsoft.com, mscrl.microsoft.com, and ocsp.msocsp.com.

Expand table

Certificate Authority	Serial Number Thumbprint
Baltimore CyberTrust Root	0x20000b9 D4DE20D05E66FC53FE1A50882C78DB2852CAE474
Microsoft RSA TLS CA 01	0x0f14965f202069994fd5c7ac788941e2 703D7A8F0EBF55AAA59F98EAF4A206004EB2516A
Microsoft RSA TLS CA 02	0x0fa74722c53d88c80f589efb1f9d4a3a B0C2D2D13CDD56CDAA6AB6E2C04440BE4A429C75
DigiCert Cloud Services CA-1	0f171a48c6f223809218cd2ed6ddc0e8 B3F6B64A07BB9611F47174407841F564FB991F29

- July 22, 2024: Added Entrust CAs from a parallel Microsoft 365 article to provide a comprehensive list.
- June 27, 2024: Removed the following CAs, which were superseded by both versions of Microsoft Azure ECC TLS Issuing CAs 03, 04, 07, 08.

Expand table

Certificate Authority	Serial Number Thumbprint
Microsoft Azure ECC TLS Issuing CA 01	0x09dc42a5f574ff3a389ee06d5d4de440 92503D0D74A7D3708197B6EE13082D52117A6AB0
Microsoft Azure ECC TLS Issuing CA 01	0x330000001aa9564f44321c54b900000000001a CDA57423EC5E7192901CA1BF6169DBE48E8D1268
Microsoft Azure ECC TLS Issuing CA 02	0x0e8dbe5ea610e6cbb569c736f6d7004b 1E981CCDDC69102A45C6693EE84389C3CF2329F1
Microsoft Azure ECC TLS Issuing CA 02	0x330000001b498d6736ed5612c200000000001b 489FF5765030EB28342477693EB183A4DED4D2A6
Microsoft Azure ECC TLS Issuing CA 05	x0ce59c30fd7a83532e2d0146b332f965 C6363570AF8303CDF31C1D5AD81E19DBFE172531
Microsoft Azure ECC TLS Issuing CA 05	0x330000001cc0d2a3cd78cf2c1000000000001c 4C15BC8D7AA5089A84F2AC4750F040D064040CD4
Microsoft Azure ECC TLS Issuing CA 06	0x066e79cd7624c63130c77abeb6a8bb94 7365ADAEDFEA4909C1BAADBAB68719AD0C381163
Microsoft Azure ECC TLS Issuing CA 06	0x330000001d0913c309da3f05a600000000001d DFEB65E575D03D0CC59FD60066C6D39421E65483
Microsoft Azure TLS Issuing CA 01	0x0aafa6c5ca63c45141ea3be1f7c75317 2F2877C5D778C31E0F29C7E371DF5471BD673173
Microsoft Azure TLS Issuing CA 01	0x1dbe9496f3db8b8de700000000001d B9ED88EB05C15C79639493016200FDAB08137AF3
Microsoft Azure TLS Issuing CA 02	0x0c6ae97cced599838690a00a9ea53214 E7EEA674CA718E3BEFD90858E09F8372AD0AE2AA
Microsoft Azure TLS Issuing CA 02	0x330000001ec6749f058517b4d000000000001e C5FB956A0E7672E9857B402008E7CCAD031F9B08
Microsoft Azure TLS Issuing CA 05	0x0d7bede97d8209967a52631b8bdd18bd 6C3AF02E7F269AA73AFD0EFF2A88A4A1F04ED1E5
Microsoft Azure TLS Issuing CA 05	0x330000001f9f1fa2043bc28db900000000001f 56F1CA470BB94E274B516A330494C792C419CF87
Microsoft Azure TLS Issuing CA 06	0x02e79171fb8021e93fe2d983834c50c0 30E01761AB97E59A06B41EF20AF6F2DE7EF4F7B0
Microsoft Azure TLS Issuing CA 06	0x3300000020a2f1491a37fbd31f000000000020 8F1FD57F27C828D7BE29743B4D02CD7E6E5F43E6

- July 17, 2023: Added 16 new subordinate Certificate Authorities.

- February 7, 2023: Added eight new subordinate Certificate Authorities.

Next steps

To learn more about Certificate Authorities and PKI, see:

- Microsoft PKI Repository
- Microsoft PKI Repository, including CRL and policy information
- Azure Firewall Premium certificates
- PKI certificates and Configuration Manager
- Securing PKI

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 09/10/2026

Disabling TLS 1.0 and 1.1 for Microsoft 365

09/26/2025

Applies to: Microsoft 365 Apps for enterprise, Office 365 Business, Office 365 Personal, Office Online Server, Office Web Apps

> **Important**
>
> We have already disabled TLS 1.0 and 1.1 for most Microsoft 365 services in the world wide environment. For Microsoft 365 operated by 21 Vianet, TLS 1.0/1.1 was disabled on June 30, 2023.

As of October 31, 2018, the Transport Layer Security (TLS) 1.0 and 1.1 protocols are deprecated for the Microsoft 365 service. The effect for end-users is minimal. This change has been publicized for over two years, with the first public announcement made in December 2017. This article is only intended to cover the Office 365 local client in relation to the Office 365 service but can also apply to on-premises TLS issues with Office and Office Online Server/Office Web Apps.

For SharePoint and OneDrive, you'll need to update and configure .NET to support TLS 1.2. For information, see How to enable TLS 1.2 on clients.

Office 365 and TLS overview

The Office client relies on the Windows web service (WINHTTP) to send and receive traffic over TLS protocols. The Office client can use TLS 1.2 if the web service of the local computer can use TLS 1.2. All Office clients can use TLS protocols, as TLS and SSL protocols are part of the operating system and not specific to the Office client.

On Windows 8 and later versions

By default, the TLS 1.2 and 1.1 protocols are available if no network devices are configured to reject TLS 1.2 traffic.

On Windows 7

TLS 1.1 and 1.2 protocols are not available without the KB 3140245 update. The update addresses this issue and adds the following registry sub key:

```
Console
```

```
HKEY_LOCAL_MACHINE\SOFTWARE\Microsoft\Windows\CurrentVersion\Internet
Settings\WinHttp
```

> **Note**
>
> Windows 7 users who do not have this update are affected as of October 31, 2018. KB 3140245 has details about how to change WINHTTP settings to enable TLS protocols.

More information

The value of the **DefaultSecureProtocols** registry key that the KB article describes determines which network protocols can be used:

Expand table

DefaultSecureProtocols Value	Protocol enabled
0x00000008	Enable SSL 2.0 by default
0x00000020	Enable SSL 3.0 by default
0x00000080	Enable TLS 1.0 by default
0x00000200	Enable TLS 1.1 by default
0x00000800	Enable TLS 1.2 by default
0x00002000	Enable TLS 1.3 by default

Office clients and TLS registry keys

You can refer to KB 4057306 Preparing for the mandatory use of TLS 1.2 in Office 365. This is a general article for IT administrators, and it's official documentation about the TLS 1.2 change.

The following table shows the appropriate registry key values in Office 365 clients after October 31, 2018.

Expand table

Enabled protocols for Office 365 service after October 31, 2018	Hexadecimal value
TLS 1.0 + 1.1 + 1.2	0x00000A80

Enabled protocols for Office 365 service after October 31, 2018	Hexadecimal value
TLS 1.1 + 1.2	0x00000A00
TLS 1.0 + 1.2	0x00000880
TLS 1.2	0x00000800

> **Important**
>
> Don't use the SSL 2.0 and 3.0 protocols, which can also be set by using the **DefaultSecureProtocols** key. SSL 2.0 and 3.0 are considered outdated and insecure protocols. The best practice is to end the use of SSL 2.0 and SSL 3.0, although the decision to do this ultimately depends on what best meets your product needs. For more information about SSL 3.0 vulnerabilities, refer to **KB 3009008**.

You can use the default Windows Calculator in Programmer mode to set up the same reference registry key values. For more information, see KB 3140245 Update to enable TLS 1.1 and TLS 1.2 as a default secure protocols in WinHTTP in Windows.

Regardless if the Windows 7 update (KB 3140245) is installed or not, the DefaultSecureProtocols registry sub key isn't present and must be added manually or through a group policy object (GPO). That is, unless you have to customize what secure protocols are enabled or restricted, this key is not required. You only need the Windows 7 SP1 (KB 3140245) update.

Update and configure the .NET Framework to support TLS 1.2

You'll need to update applications that call Microsoft 365 APIs over TLS 1.0 or TLS 1.1 to use TLS 1.2. .NET 4.5 defaults to TLS 1.1. To update your .NET configuration, see How to enable Transport Layer Security (TLS) 1.2 on clients.

More information

For more information, see Preparing for the mandatory use of TLS 1.2 in Office 365.

References

The following resources provide guidance to help make sure that your clients are using TLS 1.2 or a later version and to disable TLS 1.0 and 1.1:

- For Windows 7 clients that connect to Office 365, make sure that TLS 1.2 is the default secure protocol in WinHTTP in Windows. For more information, see KB 3140245 - Update to enable TLS 1.1 and TLS 1.2 as default secure protocols in WinHTTP in Windows.
- To address weak TLS usage by removing TLS 1.0 and 1.1 dependencies, see TLS 1.2 support at Microsoft.
- New IIS functionality makes it easier to find clients on Windows Server 2012 R2 and Windows Server 2016 that connect to the service by using weak security protocols.
- Get more information about how to solve the TLS 1.0 problem.
- For general information about our approach to security, go to the Office 365 Trust Center.
- Preparing for TLS 1.0/1.1 Deprecation - Office 365 Skype for Business
- Exchange Server TLS guidance, part 1: Getting Ready for TLS 1.2
- Exchange Server TLS guidance Part 2: Enabling TLS 1.2 and Identifying Clients Not Using It
- Exchange Server TLS guidance Part 3: Turning Off TLS 1.0/1.1
- Enable TLS 1.1 and TLS 1.2 support in Office Online Server

Disabling TLS 1.0 and 1.1 in Microsoft 365 GCC High and DoD

09/26/2025Applies to: Office 365 Business

Summary

In order to comply with the latest compliance standards for the Federal Risk and Authorization Management Program (FedRAMP), we are disabling Transport Layer Security (TLS) versions 1.1 and 1.0 in Microsoft 365 for GCC High and DoD environments. This change was previously announced through Microsoft Support in Preparing for the mandatory use of TLS 1.2 in Office 365.

The security of your data is important, and we are committed to transparency about changes that could affect your use of the service.

Although the Microsoft TLS 1.0 implementation has no known security vulnerabilities, we remain committed to the FedRAMP compliance standards. Therefore, we disabled TLS 1.1 and 1.0 in Microsoft 365 in GCC High and DoD environments on January 15, 2020.

More information

Starting on January 15, 2020, Microsoft 365 in the GCC High and DoD environments will disable TLS 1.1 and 1.0.

By January 15, 2020, all combinations of client servers and browser servers should use TLS version 1.2 (or a later version) to make sure that all connections can be made without issues to Microsoft 365. This may require updates to certain combinations of client servers and browser servers.

For SharePoint and OneDrive, you'll need to update and configure .NET to support TLS 1.2. For information, see How to enable TLS 1.2 on clients.

You must update your client computers to make sure that you maintain uninterrupted access to Office 365 GCC High and DoD.

You'll need to update applications that call Microsoft 365 APIs over TLS 1.0 or TLS 1.1 to use TLS 1.2. .NET 4.5 defaults to TLS 1.1. To update your .NET configuration, see How to enable Transport Layer Security (TLS) 1.2 on clients. For more information, see Preparing for the mandatory use of TLS 1.2 in Office 365.

We know that the following client applications cannot use TLS 1.2:

- Android 4.3 and earlier versions
- Firefox version 5.0 and earlier versions
- Internet Explorer 8-10 on Windows 7 and earlier versions
- Internet Explorer 10 on Windows Phone 8.0
- Safari 6.0.4/OS X 10.8.4 and earlier versions

Although current analysis of connections to Microsoft Online services shows that most services and endpoints see little TLS 1.1 and 1.0 usage, we're providing notice of this change so that you can update any affected clients or servers as necessary before support for TLS 1.1 and 1.0 ends. If you are using any on-premises infrastructure for hybrid scenarios or Active Directory Federation Services (AD FS), make sure that the infrastructure can support both inbound and outbound connections that use TLS 1.2 (or a later version).

In addition to the outages that you might experience if you use the listed clients that cannot use TLS 1.2, removing TLS 1.1 and 1.0 will prevent you from being able to use the following Microsoft product:

- Lync phone

References

For guidance and references to help make sure that your clients are using TLS 1.2, see Preparing for the mandatory use of TLS 1.2 in Office 365.

Preparing for TLS 1.2 in Office 365 and Office 365 GCC

09/26/2025Applies to: Office 365 Business

Summary

To provide the best-in-class encryption to our customers, Microsoft has deprecated Transport Layer Security (TLS) versions 1.0 and 1.1 in Office 365 and Office 365 GCC. We understand that the security of your data is important, and we're committed to transparency about changes that may affect your use of the TLS service.

The Microsoft TLS 1.0 implementation has no known security vulnerabilities. But because of the potential for future protocol downgrade attacks and other TLS vulnerabilities, we discontinued support for TLS 1.0 and 1.1 in Microsoft Office 365 and Office 365 GCC.

For information about how to remove TLS 1.0 and 1.1 dependencies, see the following white paper: Solving the TLS 1.0 problem.

After you upgrade to TLS 1.2, make sure that the cipher suites you're using are supported by Azure Front Door. Microsoft 365 and Azure Front Door have slight differences in cipher suite support. For details, see What are the current cipher suites supported by Azure Front Door?.

More information

We began deprecation of TLS 1.0 and 1.1 as of January 2020. Any clients, devices, or services that connect to Office 365 through TLS 1.0 or 1.1 in our DoD or GCC High instances are unsupported. For our commercial customers of Office 365, deprecation of TLS 1.0 and 1.1 began October 15, 2020 and rollout continued over the following weeks and months.

We recommend that all client-server and browser-server combinations use TLS 1.2 (or a later version) in order to maintain connection to Office 365 services. You might have to update certain client-server and browser-server combinations.

> **Note**
>
> For SMTP Inbound mail flow, after deprecation of TLS 1.0 and 1.1, we will accept only TLS 1.2 connection. However, we will continue accepting SMTP Connection which is unencrypted without any TLS. We do not recommend email transmission without any encryption.

You'll need to update applications that call Microsoft 365 APIs over TLS 1.0 or TLS 1.1 to use TLS 1.2. .NET 4.5 defaults to TLS 1.1. To update your .NET configuration, see How to enable Transport Layer Security (TLS) 1.2 on clients.

The following clients are known to be unable to use TLS 1.2. Update these clients to ensure uninterrupted access to the service.

- Android 4.3 and earlier versions
- Firefox version 5.0 and earlier versions
- Internet Explorer 8-10 on Windows 7 and earlier versions
- Internet Explorer 10 on Windows Phone 8
- Safari 6.0.4/OS X10.8.4 and earlier versions

TLS 1.2 for Microsoft Teams Rooms and Surface Hub

Microsoft Teams Rooms (previously known as Skype Room System V2 SRS V2) have supported TLS 1.2 since December 2018. We recommend that Rooms devices have Microsoft Teams Rooms app version 4.0.64.0 or later installed. For more information, see the Release notes. The changes are backward and forward compatible.

Surface Hub released TLS 1.2 support in May 2019.

TLS 1.2 support for Microsoft Teams Rooms and Surface Hub products also requires the following server-side code changes:

- Skype for Business Online server changes were made live in April 2019. Now, Skype for Business Online supports connecting Microsoft Teams Rooms and Surface Hub devices by using TLS 1.2.

- Skype for Business Server customers must install a cumulative update (CU) to use TLS 1.2 for Teams Rooms Systems and Surface Hub.
  - For Skype for Business Server 2015, CU9 is already released in May 2019.
  - For Skype for Business Server 2019, CU1 was previously planned for April 2019 but is delayed to June 2019.

  > ⓘ **Note**
  >
  > Skype for Business on-premises customers should not disable TLS 1.0/1.1 before installing specific CUs for Skype for Business Server.

If you are using any on-premises infrastructure for hybrid scenarios or Active Directory Federation Services, make sure that the infrastructure can support both inbound and outbound

connections that use TLS 1.2.

References

The following resources provide guidance to help make sure that your clients are using TLS 1.2 or a later version and to disable TLS 1.0 and 1.1.

- For Windows 7 clients that connect to Office 365, make sure that TLS 1.2 is the default secure protocol in WinHTTP in Windows. For more information see KB 3140245 - Update to enable TLS 1.1 and TLS 1.2 as a default secure protocols in WinHTTP in Windows.
- TLS cipher suites supported by Office 365
- To start addressing weak TLS use by removing TLS 1.0 and 1.1 dependencies, see TLS 1.2 support at Microsoft.
- New IIS functionality makes it easier to find clients on Windows Server 2012 R2 and Windows Server 2016 that connect to the service by using weak security protocols.
- Get more information about how to solve the TLS 1.0 problem.
- For general information about our approach to security, go to the Office 365 Trust Center.
- To identify the TLS version that is used by SMTP clients, see SMTP Auth clients report in the EAC.
- Preparing for TLS 1.0/1.1 Deprecation - Office 365 Skype for Business
- Exchange Server TLS guidance, part 1: Getting Ready for TLS 1.2
- Exchange Server TLS guidance Part 2: Enabling TLS 1.2 and Identifying Clients Not Using It
- Exchange Server TLS guidance Part 3: Turning Off TLS 1.0/1.1
- Enable TLS 1.1 and TLS 1.2 support in Office Online Server
- Enable TLS and SSL support in SharePoint 2013
- Enable TLS 1.1 and TLS 1.2 support in SharePoint Server 2016
- Enable TLS 1.1 and TLS 1.2 support in SharePoint Server 2019

Microsoft Purview data compliance solutions

Microsoft Purview data compliance solutions help you manage and monitor your data, protect information, minimize compliance risks, and meet regulatory requirements. This article helps you learn about Microsoft Purview data compliance solutions and quickly get started with deploying these solutions to meet specific compliance needs for your organization.

If you're new to Microsoft Purview, consider starting with the Microsoft Purview Suite trial to explore compliance capabilities at the E5 level for free.

Data compliance solutions

Microsoft Purview offers the following data compliance solutions to help your organization manage data risk, meet regulatory requirements, and respond to legal obligations:

- Monitor organizational activity with Microsoft Purview Audit
- Detect inappropriate messages with Microsoft Purview Communication Compliance
- Manage regulatory requirements with Microsoft Purview Compliance Manager
- Retain and delete content with Microsoft Purview Data Lifecycle Management
- Identify and collect data for legal cases with Microsoft Purview eDiscovery
- Meet record-keeping requirements with Microsoft Purview Records Management

Monitor organizational activity with Audit

Microsoft Purview auditing solutions provide an integrated solution to help organizations effectively respond to security events, forensic investigations, internal investigations, and compliance obligations. Your organization's unified audit log captures, records, and retains thousands of user and admin operations performed in dozens of Microsoft 365 services and solutions.

Use Audit when your organization needs to:

- Investigate possible security breaches and determine the scope of compromise.
- Support internal investigations by tracking user and admin activities across Microsoft 365 services.
- Meet compliance obligations through long-term audit record retention.

- Conduct forensic investigations by reviewing when mail items were accessed, forwarded, or replied to.

To get started with Audit or learn more, see the following articles:

- Get started with auditing solutions
- Search the audit log
- Audit logs for Copilot and AI applications
- Manage audit log retention policies

Detect inappropriate messages with Communication Compliance

Microsoft Purview Communication Compliance helps minimize risks by helping you quickly detect, capture, and take remediation actions for email and Microsoft Teams communications. These actions include inappropriate communications containing profanity, threats, and harassment and communications that share sensitive information inside and outside of your organization.

Use Communication Compliance when your organization needs to:

- Detect and act on workplace harassment, threats, or profanity in user communications.
- Monitor for inappropriate sharing of sensitive information inside and outside your organization.
- Meet regulatory requirements for messaging oversight in financial services, such as FINRA Rule 3110.
- Identify potential insider trading, collusion, or unauthorized communications about confidential projects.

To get started with Communication Compliance or learn more, see the following articles:

- Learn about communication compliance
- Plan for communication compliance
- Create and manage policies
- Investigate and remediate alerts

Manage regulatory requirements with Compliance Manager

Microsoft Purview Compliance Manager helps you manage your organization's compliance requirements with greater ease and convenience. Compliance Manager can help you throughout

your compliance journey, from taking inventory of your data protection risks to managing the complexities of implementing controls, staying current with regulations and certifications, and reporting to auditors. Microsoft also offers a comprehensive set of compliance offerings to help your organization comply with national, regional, and industry-specific requirements governing the collection and use of data.

Use Compliance Manager when your organization needs to:

- Assess and manage compliance posture across multicloud environments.
- Prepare for audits by tracking improvement actions and compliance scores.
- Stay current with evolving regulations by using prebuilt assessments for more than 360 industry and regional standards.
- Prioritize which actions to take first based on risk-based compliance scoring.

To get started with Compliance Manager or learn more, see the following articles:

- Get started with Compliance Manager
- Build and manage assessments
- Learn about regulations
- Working with improvement actions

Retain and delete content with Data Lifecycle Management

Microsoft Purview Data Lifecycle Management provides tools and capabilities to retain and delete content across Exchange, SharePoint, OneDrive, Microsoft 365 Groups, Teams, and Viva Engage. Retaining and deleting emails, documents, and messages often meets compliance and regulatory requirements. However, deleting content that no longer has business value also reduces your exposure to cyberattacks.

Use Data Lifecycle Management when your organization needs to:

- Retain content for a specified period to meet compliance and regulatory requirements.
- Delete content that no longer has business value to reduce your exposure to cyberattacks.
- Archive mailbox data for users who need more than 100 GB of storage.
- Retain mailbox content after users leave your organization through inactive mailboxes.

To get started with Data Lifecycle Management or learn more, see the following articles:

- Get started with data lifecycle management
- Learn about retention policies and retention labels
- Create retention policies

- Learn about archive mailboxes

Identify and collect data for legal cases with eDiscovery

Microsoft Purview eDiscovery is the process of identifying, collecting, and auditing electronic information for legal, regulatory, or business reasons. Use eDiscovery to search for data and content in Exchange Online, OneDrive, SharePoint, Microsoft Teams, Microsoft 365 Groups, and Viva Engage teams. You can search mailboxes and sites in the same eDiscovery search, and then export the search results for analysis and review.

Use eDiscovery when your organization needs to:

- Identify and collect electronically stored information as evidence for legal cases.
- Respond to regulatory investigation and litigation requests.
- Preserve content in place to prevent deletion during an investigation.
- Reduce review costs by using analytics like near-duplicate detection and email threading.

To get started with eDiscovery or learn more, see the following articles:

- Learn about the eDiscovery workflow
- Get started with eDiscovery
- Create and manage cases
- Search for content
- Manage review sets

Meet record-keeping requirements with Records Management

Microsoft Purview Records Management helps your organization manage its legal obligations, provides the ability to demonstrate compliance with regulations, and increases efficiency by regularly disposing of items that are no longer required to be retained, no longer of value, or no longer required for business purposes.

Use Records Management when your organization needs to:

- Declare items as records or regulatory records to ensure content remains immutable.
- Start retention periods based on specific events, such as contract expiration or employee termination.
- Conduct disposition reviews with proof of deletion for compliance auditing.
- Manage the lifecycle of business-critical content with strict control and audit trails.

To get started with Records Management or learn more, see the following articles:

- Get started with records management
- Declare records by using retention labels
- Create and manage retention labels with file plan
- Manage disposition of data

Related Microsoft Purview solutions

Microsoft Purview data compliance solutions work alongside other Microsoft Purview solution areas:

- Microsoft Purview data security solutions
- Microsoft Purview data governance solutions

To deploy integrated compliance solutions, see Deploy a data governance solution.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/04/2026

Learn about auditing solutions in Microsoft Purview

Microsoft Purview auditing solutions provide an integrated solution to help organizations effectively respond to security events, forensic investigations, internal investigations, and compliance obligations. Your organization's unified audit log captures, records, and retains thousands of user and admin operations performed in dozens of Microsoft services and solutions. Security ops, IT admins, insider risk teams, and compliance and legal investigators in your organization can search audit records for these events. This capability provides visibility into the activities performed across your organization.

Comparison of key capabilities

The following table compares the key capabilities available in Audit (Standard) and Audit (Premium). Audit (Premium) includes all Audit (Standard) functionality.

Expand table

Capability	Audit (Standard)	Audit (Premium)
Enabled by default	✔	✔
Thousands of searchable audit events	✔	✔
Audit search tool in the Microsoft Purview portal	✔	✔
Audit Search Graph API	✔	✔
Search-UnifiedAuditLog cmdlet	✔	✔
Export audit records to CSV file	✔	✔
Access to audit logs via Office 365 Management Activity API [1]	✔	✔
180-day audit log retention	✔	✔
Up to 1-year audit log retention		✔

Capability	Audit (Standard)	Audit (Premium)
10-year audit log retention [2]		✔
Audit log retention policies		✔
Intelligent insights		✔

> **Note**
>
> [1] Audit (Premium) includes higher bandwidth access to the Office 365 Management Activity API, which provides faster access to audit data.
>
> [2] In addition to the required licensing for Audit (Premium) (described in the next section), a user must be assigned a 10-Year Audit Log Retention add-on license to retain their audit records for 10 years. Audit records generated by non-user entities (such as service principal actions, system events, and application activities) are retained for a fixed period of one year. This retention period isn't configurable and custom audit log retention policies don't apply to these records.

Audit (Standard)

Microsoft Purview Audit (Standard) enables you to log and search for audited activities to support your forensic, IT, compliance, and legal investigations.

- **Enabled by default**. Audit (Standard) is enabled by default for all organizations with the appropriate subscription. That configuration captures and makes searchable records for audited activities. You only need to assign the necessary permissions to access the audit log search tool (and the corresponding cmdlet) and ensure that users have the right license for Microsoft Purview Audit (Premium) features.
- **Thousands of searchable audit events**. You can search for a wide range of audited activities that occur in most of the Microsoft services in your organization. For a list of the activities you can search for, see Audit log activities. For a list of the services and features that support audited activities, see Audit log record type.
- **Audit search tool in the Microsoft Purview portal**. Use the Audit log search tool in the portal to search for audit records. You can search for specific activities, for activities performed by specific users, and activities that occurred within a date range.

- **Audit Search Graph API**. Microsoft Graph offers a unified API endpoint for accessing data from multiple Microsoft cloud services in a single response. The Audit Search Graph API allows you to programmatically access the audit search experience through Microsoft Graph.

- **Search-UnifiedAuditLog cmdlet**. You can also use the **Search-UnifiedAuditLog** cmdlet in Exchange Online PowerShell (the underlying cmdlet for the search tool) to search for audit events or to use in a script. For more information, see:
  - Search-UnifiedAuditLog cmdlet reference
  - Use a PowerShell script to search the audit log

- **Export audit records to a CSV file**. After running the Audit log search tool in the Microsoft Purview portal, you can export the audit records returned by the search to a CSV file. This process lets you use Microsoft Excel to sort and filter on different audit record properties. You can also use Excel Power Query transform functionality to split each property in the AuditData JSON object into its own column. This process lets you effectively view and compare similar data for different events. For more information, see Export, configure, and view audit log records.

- **Access to audit logs via Office 365 Management Activity API**. A third method for accessing and retrieving audit records is to use the Office 365 Management Activity API. This method lets organizations retain auditing data for longer periods than the default 180 days and lets them import their auditing data to a SIEM solution. For more information, see Office 365 Management Activity API reference.

- **180-day audit log retention**. When a user or admin performs an audited activity, the system generates an audit record and stores it in the audit log for your organization. In Audit (Standard), the system retains records for 180 days, which means you can search for activities that occurred within the past six months.

> **Important**
>
> The default retention period for Audit (Standard) changed from 90 days to 180 days. Audit (Standard) logs generated before October 17, 2023, are retained for 90 days. Audit (Standard) logs generated on or after October 17, 2023, follow the new default retention of 180 days.

Audit (Premium)

> **Important**
>
> Classic Search retired on November 30, 2023. New Search includes enhancements such as faster search times, additional search options, ability to save searches, and more.

Audit (Premium) builds on the capabilities of Audit (Standard) by providing audit log retention policies, longer retention of audit records, high-value intelligent insights, and higher bandwidth access to the Office 365 Management Activity API.

- **Audit log retention policies**. Create customized audit log retention policies to retain audit records for longer periods, up to one year (and up to 10 years for users with the required add-on license). Create a policy to retain audit records based on the service where the audited activities occur, specific audited activities, or the user who performs an audited activity.
- **Longer retention of audit records**. Microsoft Entra ID, Exchange, OneDrive, and SharePoint audit records are retained for one year by default. Audit records for all other activities are retained for 180 days by default, or you can use audit log retention policies to configure longer retention periods.
- **Audit (Premium) intelligent insights**. Audit records for intelligent insights can help your organization conduct forensic and compliance investigations by providing visibility to events such as when mail items were accessed, or when mail items were replied to and forwarded, or when and what a user searched for in Exchange Online and SharePoint Online. These intelligent insights can help you investigate possible breaches and determine the scope of compromise.
- **Higher bandwidth to the Office 365 Management Activity API**. Audit (Premium) provides organizations with more bandwidth to access auditing logs through the Office 365 Management Activity API. Although all organizations (that have Audit (Standard) or Audit (Premium)) initially receive a baseline of 2,000 requests per minute, this limit dynamically increases depending on an organization's seat count and their licensing subscription. This change results in organizations with Audit (Premium) getting about twice the bandwidth as organizations with Audit (Standard).

Long-term retention of audit logs

Audit (Premium) retains all Exchange, SharePoint, and Microsoft Entra audit records for one year. This retention happens through a default audit log retention policy that retains any audit record that contains the value of **AzureActiveDirectory**, **Exchange**, **OneDrive**, or **SharePoint**, for the **Workload** property (which indicates the service in which the activity occurred) for one year.

Retaining audit records for longer periods can help with ongoing forensic or compliance investigations. For more information, see the "Default audit log retention policy" section in Manage audit log retention policies.

In addition to the one-year retention capabilities of Audit (Premium), Microsoft also released the capability to retain audit logs for 10 years. The 10-year retention of audit logs helps support long running investigations and respond to regulatory, legal, and internal obligations.

> **Note**
>
> Retaining audit logs for 10 years requires an additional per-user add-on license. After you assign this license to a user and set an appropriate 10-year audit log retention policy for that user, audit logs covered by that policy start to be retained for the 10-year period. This policy isn't retroactive and can't retain audit logs that were generated before the 10-year audit log retention policy was created.

Audit log retention policies

All audit records that other services generate and that the default audit log retention policy doesn't cover are retained for 180 days. You can create customized audit log retention policies to retain other audit records for longer periods, up to 10 years. You can create a policy to retain audit records based on one or more of the following criteria:

- The Microsoft service where the audited activities occur.
- Specific audited activities.
- The user who performs an audited activity.

> **Important**
>
> The default retention period for Audit (Standard) changed from 90 days to 180 days. Audit (Standard) logs generated before October 17, 2023, are retained for 90 days. Audit (Standard) logs generated on or after October 17, 2023, follow the new default retention of 180 days. You can also specify how long to retain audit records that match the policy and a priority level so that specific policies take priority over other policies. Any custom audit log retention policy takes precedence over the default audit retention policy if you need to retain Exchange, SharePoint, or Microsoft Entra ID audit records for less than a year or for 10

years for some or all users in your organization. For more information, see Manage audit log retention policies.

> **Important**
>
> The audit item lifetime for data is determined when the auditing pipeline adds the data and is based on the licensing defaults or applicable retention policies. Any changes to licensing or applicable retention policies change the expiration time of the audit data after updating. These changes don't affect any previously committed items.

Audit (Premium) activity properties

Audit (Premium) helps organizations conduct forensic and compliance investigations by providing access to important events, such as when users access mail items, reply to and forward mail items, and search in Exchange Online and SharePoint Online. These events can help you investigate possible breaches and determine the scope of compromise. In addition to these events in Exchange and SharePoint, other Microsoft services include important events that require assigning users the appropriate Audit (Premium) license. Assign an Audit (Premium) license to users so the system generates audit logs when they perform these events.

These activities require that you assign users the appropriate Audit (Premium) license. Assign an Audit (Premium) license to users so the system generates audit logs when they perform these activities and properties.

Audit (Premium) provides access to the following activity properties:

Exchange Online

Expand table

Activity	Property
MailItemsAccessed	SensitivityLabel

Microsoft Teams

Expand table

Activity	Property
ChatCreated	AppAccessContext
ChatRetrieved	AppAccessContext
ChatUpdated	AppAccessContext
MeetingParticipantDetail	IsJoinedFromLobby ArtifactShared
MessageCreatedNotification	AppAccessContext
MessageDeletedNotification	AppAccessContext
MessageHostedContentsListed	AppAccessContext
MessageHostedContentRead	AppAccessContext
MessagesListed	AppAccessContext
MessageRead	AppAccessContext
MessageSent	AppAccessContext ParticipatingDomainInformation ParticipantInfo
MessageUpdated	ParticipantInfo AppAccessContext
MessageUpdatedNotification	AppAccessContext
SubscribedToMessages	AppAccessContext

High-bandwidth access to the Office 365 Management Activity API

Organizations that access auditing logs through the Office 365 Management Activity API faced throttling limits at the publisher level. This throttling limit meant that if a publisher pulled data for multiple customers, all those customers shared the same limit.

With Audit (Premium), this limit changed from a publisher-level limit to a tenant-level limit. Each organization now gets its own fully allocated bandwidth quota to access its auditing data. The bandwidth isn't a static, predefined limit. Instead, it's modeled on a combination of factors, including the number of seats in the organization. E5, A5, and G5 organizations get more bandwidth than non-E5, non-A5, and non-G5 organizations.

All organizations initially get a baseline of 2,000 requests per minute. This limit dynamically increases based on an organization's seat count and licensing subscription. E5, A5, and G5 organizations get about twice as much bandwidth as non-E5, non-A5, and non-G5 organizations. A cap on the maximum bandwidth protects the health of the service.

For more information, see the **API throttling** section in Office 365 Management Activity API reference.

Licensing requirements

Before you get started, review the subscription requirements for Audit (Standard) and Audit (Premium).

Next steps

> **Tip**
>
> Training your security operations team, IT administrators, and compliance investigators in Audit (Standard) and Audit (Premium) can help your organization get started more quickly. For more information, see **Describe the eDiscovery and audit capabilities of Microsoft Purview**.

- Get started with auditing solutions: Walk through the setup steps to enable Audit (Standard) or Audit (Premium) in your organization.
- Search the audit log: Run your first audit log search to find user and admin activities.
- Manage audit log retention policies: Create custom retention policies to keep audit records beyond the default period.

ⓘ **Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 05/18/2026

Get started with auditing solutions

Microsoft Purview Audit (Standard) and Audit (Premium) enable you to search for audit records of activities that users and admins perform in different Microsoft services. Because Audit (Standard) is enabled by default for most Microsoft 365 organizations, you only need to complete a few steps before you and others in your organization can search the audit log. To use features available only in Audit (Premium), you need to complete a few more configuration steps.

For more information about Audit (Standard) and Audit (Premium) capabilities, see Microsoft Purview auditing solutions.

Step 1: Verify organization subscription and billing

Licensing for Audit (Standard) and Audit (Premium) requires the appropriate organization subscription that provides access to the audit log search tool and per-user licensing that's required to log and retain audit records.

When a user or admin performs an audited activity, the system generates an audit record and stores it in the audit log for your organization. In Audit (Standard) and Audit (Premium), you can retain and search audit records in the audit log for 180 days.

> **Important**
>
> The default retention period for Audit (Standard) changed from 90 days to 180 days. Audit (Standard) logs that were generated before October 17, 2023, are retained for 90 days. Audit (Standard) logs that are generated on or after October 17, 2023, follow the new default retention of 180 days.

For a list of subscription and licensing requirements for these auditing solutions, see the subscription requirements for Audit (Standard) and Audit (Premium).

To audit user interactions with non-Microsoft 365 AI data, which includes information from other generative AI applications from Microsoft and connected external AI applications, you must enable pay-as-you-go billing for your organization. This data type includes Copilot in Microsoft Fabric, Microsoft Security Copilot, Microsoft Copilot Studio, and any connected or cloud AI application.

Step 2: Assign permissions to search the audit log

Admins and members of investigation teams must be assigned the *View-Only Audit Logs* or *Audit Logs* role in the Microsoft Purview portal to search or export the audit log. By default, the

Microsoft Purview portal **Role groups** page assigns these roles to the *Audit Reader* and *Audit Manager* role groups.

> ⓘ **Note**
>
> Currently, Exchange admin center permissions are required to enable or disable auditing and to access audit cmdlets. Use the existing *Audit Logs* and *View-Only Audit Logs* roles in Exchange admin center to grant access to audit cmdlets. Use the existing *Audit Logs* role in Exchange admin center to grant access to enable or disable auditing.

You can also create custom role groups with the ability to search the audit log by adding the *View-Only Audit Logs* or *Audit Logs* roles to a custom role group. For more information, see Permissions in the Microsoft Purview portal.

> ⓘ **Note**
>
> Access to the **Audit Search Graph API** requires additional permissions to be configured in Microsoft Graph. For more information, see **Permissions in Audit Search Graph API**.

Assign permissions to scope audit logs

To search or export the audit log, administrators or members of investigation teams must be assigned to at least one of the following audit-related role groups in the Microsoft Purview portal:

- **Audit Manager**: A user assigned to the *Audit Manager* role group can search and export the audit log and manage audit settings for the tenant (like enabling or disabling audit logging). This role group grants the *View-Only Audit Logs* and *Audit Logs* roles to the user.
- **Audit Reader**: A user assigned to the **Audit Reader** role group can only search and export the audit log. They can't enable or disable audit logging. This role group grants the *View-Only Audit Logs* role to the user.

Step 3: Enable/Disable SearchQueryInitiatedExchange/SearchQueryInitiatedShare events

SearchQueryInitiatedExchange must be enabled explicitly for logging when users perform searches in Exchange Online.

SearchQueryInitiatedSharePoint is by default enabled for logging when users perform searches in SharePoint.

To enable Exchange Online and SharePoint events to be audited for users, run the following cmdlet (for each user) in Exchange Online PowerShell:

PowerShell

```
Set-Mailbox <user> -AuditOwner @{Add="SearchQueryInitiated"} // For Exchange Online

Set-Mailbox <user> -AuditOwner @{Add="SharepointSearchQueryInitiated"} // For
SharePoint. Enabled by default.
```

In a multi-geo environment, you must run the **Set-Mailbox** command in the forest where the user's mailbox is located. To identify the user's mailbox location, run the following cmdlet:

PowerShell

```
Get-Mailbox <user identity> | FL MailboxLocations
```

If you previously ran the cmdlet to enable auditing of search queries in a forest that's different than the forest where the user's mailbox is located, remove the *SearchQueryInitiated/SharepointSearchQueryInitiated* value from the user's mailbox. To remove the value, run `Set-Mailbox -AuditOwner @{Remove="SearchQueryInitiated"} for Exchange Online and Set-Mailbox -AuditOwner @{Remove="SharepointSearchQueryInitiated"}`. Then, add it to the user's mailbox in the forest where the user's mailbox is located.

> **Note**
>
> These PowerShell cmdlets are valid only for SearchQueryInitiated and SharepointSearchQueryInitiated events and don't apply to other events within the audit log.

Step 4: Set up Audit (Premium) for users

> **Tip**
>
> Organizations using Audit (Standard) can skip this step.

Audit (Premium) features, such as the ability to log intelligent insights, require an appropriate E5 license assigned to users. Additionally, you must enable the Advanced Auditing app/service plan for those users.

To enable the Advanced Auditing service plan for users, complete the following steps for each user:

1. In the Microsoft 365 admin center, go to **Users** > **Active users**, and select a user.
2. On the user properties flyout page, select **Licenses and apps**.

3. In the **Licenses** section, verify that the user is assigned an E5 license or is assigned an appropriate add-on license. For a list of licenses that support Audit (Premium), see Audit licensing requirements.
4. Expand the **Apps** section and select the **Microsoft 365 Advanced Auditing** checkbox.
5. Select **Save changes**. The logging of Audit (Premium) insights begins within 24 hours.

If you customize the mailbox actions that are logged on user mailboxes or shared mailboxes, new Audit (Premium) events released by Microsoft aren't automatically audited on those mailboxes. For information about changing the mailbox actions that are audited for each sign-in type, see the "Change or restore mailbox actions logged by default" section in Manage mailbox auditing.

Step 5: Set up audit retention policies in Audit (Premium)

> **Tip**
>
> Organizations using Audit (Standard) can skip this step.

In addition to the default policy that retains Microsoft Entra ID, Exchange, OneDrive, and SharePoint audit records for one year, organizations using Audit (Premium) can create audit log retention policies to meet the requirements of your organization's security operations, IT, and compliance teams.

For more information, see Manage audit log retention policies.

Step 6: Search for audited events

Now that you have Audit (Standard) or Audit (Premium) configured for your organization, you're ready to search the audit log in the Microsoft Purview portal. For detailed guidance, see Search the audit log.

Next steps

- Search the audit log: Run your first audit log search now that auditing is configured.
- Manage audit log retention policies: Create custom retention policies to keep audit records longer than the default period.
- Export, configure, and view audit log records: Export search results to CSV for offline analysis.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 09/08/2026

Turn auditing on or off

Audit logging is on by default for Microsoft 365 organizations. However, when you set up a new Microsoft 365 organization, verify the auditing status for your organization. For instructions, see Verify the auditing status for your organization.

When you turn on auditing in the Microsoft Purview portal, the audit log records user and admin activity from your organization and automatically retains it for 180 days. The retention (lifetime) for audit data starts when it's added to the auditing log and is retained based on audit log retention policies and the license assigned to users.

> **Important**
>
> Auditing isn't enabled by default for Small and Medium Business (SMB) licenses, including Microsoft 365 Business Basic, Business Standard, and Business Premium. Additionally, unmanaged tenants that use free trials of enterprise licenses don't have auditing enabled by default. In both these cases, you must manually enable auditing for your organization.

Changes to the user licensing or retention policies also change the expiration date of audit data.

Your organization might have reasons for not wanting to record and retain audit log data. In these cases, a global admin can turn off auditing in Microsoft 365 for your organization. For instructions, see Turn off auditing.

> **Important**
>
> If you turn off auditing in Microsoft 365, you can't use the Office 365 Management Activity API or Microsoft Sentinel to access auditing data or logs for your organization. When you turn off auditing, searches of the audit log in Microsoft Purview return no results. Running the **Search-UnifiedAuditLog** cmdlet in Exchange Online PowerShell also returns no results.

Before you turn auditing on or off

You must be assigned the *Audit Logs* role in Exchange Online to turn auditing on or off. By default, the *Compliance Management* and *Organization Management* role groups on the

Permissions page in the Exchange admin center have this role.

- For step-by-step instructions on searching the audit log, see Search the audit log.
- For more information about the Microsoft 365 Management Activity API, see Get started with Microsoft 365 Management APIs.

Verify the auditing status for your organization

To verify that auditing is turned on for your organization, run the following command in Exchange Online PowerShell:

```
PowerShell
Get-AdminAuditLogConfig | Format-List UnifiedAuditLogIngestionEnabled
```

A value of `True` for the *UnifiedAuditLogIngestionEnabled* property indicates that auditing is turned on. A value of `False` indicates that auditing isn't turned on.

> **Important**
>
> Be sure to run the previous command in Exchange Online PowerShell. Although the **Get-AdminAuditLogConfig** cmdlet is also available in Security & Compliance PowerShell, the *UnifiedAuditLogIngestionEnabled* property is always `False`, even when auditing is turned on.

Turn on auditing

If auditing isn't turned on for your organization, turn it on in the Microsoft Purview portal or by using Exchange Online PowerShell. Audit events might take several hours after you turn on auditing before they become searchable in the audit log.

Complete the following steps to turn on auditing:

1. Sign in to the Microsoft Purview portal.
2. Select the **Audit** solution card. If the **Audit** solution card isn't displayed, select **View all solutions** and then select **Audit** from the **Core** section.
3. If auditing isn't turned on for your organization, a banner is displayed prompting you to start recording user and admin activity.
4. Select the **Start recording user and admin activity** banner.

Enabling auditing might take up to 60 minutes to take effect.

Use PowerShell to turn on auditing

Complete the following steps to turn on auditing by using Exchange Online PowerShell:

1. Connect to Exchange Online PowerShell.
2. Run the following PowerShell command to turn on auditing.

   PowerShell

   ```
   Set-AdminAuditLogConfig -UnifiedAuditLogIngestionEnabled $true
   ```

   A message is displayed saying that enabling auditing might take up to 60 minutes to take effect.

Turn off auditing

Use Exchange Online PowerShell to turn off auditing.

1. Connect to Exchange Online PowerShell.
2. Run the following PowerShell command to turn off auditing.

   PowerShell

   ```
   Set-AdminAuditLogConfig -UnifiedAuditLogIngestionEnabled $false
   ```

3. After a while, verify that auditing is turned off (disabled). You can do this in two ways:

   - In Exchange Online PowerShell, run the following command:

     PowerShell

     ```
     Get-AdminAuditLogConfig | FL UnifiedAuditLogIngestionEnabled
     ```

     The value of `False` for the *UnifiedAuditLogIngestionEnabled* property indicates that auditing is turned off.

   - Go to the **Audit** page in the Microsoft Purview portal.

     If auditing isn't turned on for your organization, a banner prompts you to start recording user and admin activity.

Audit records when auditing status is changed

Your organization audits changes to the auditing status. Audit records are logged when auditing is turned on or turned off. You can search the Exchange admin audit log for these audit records.

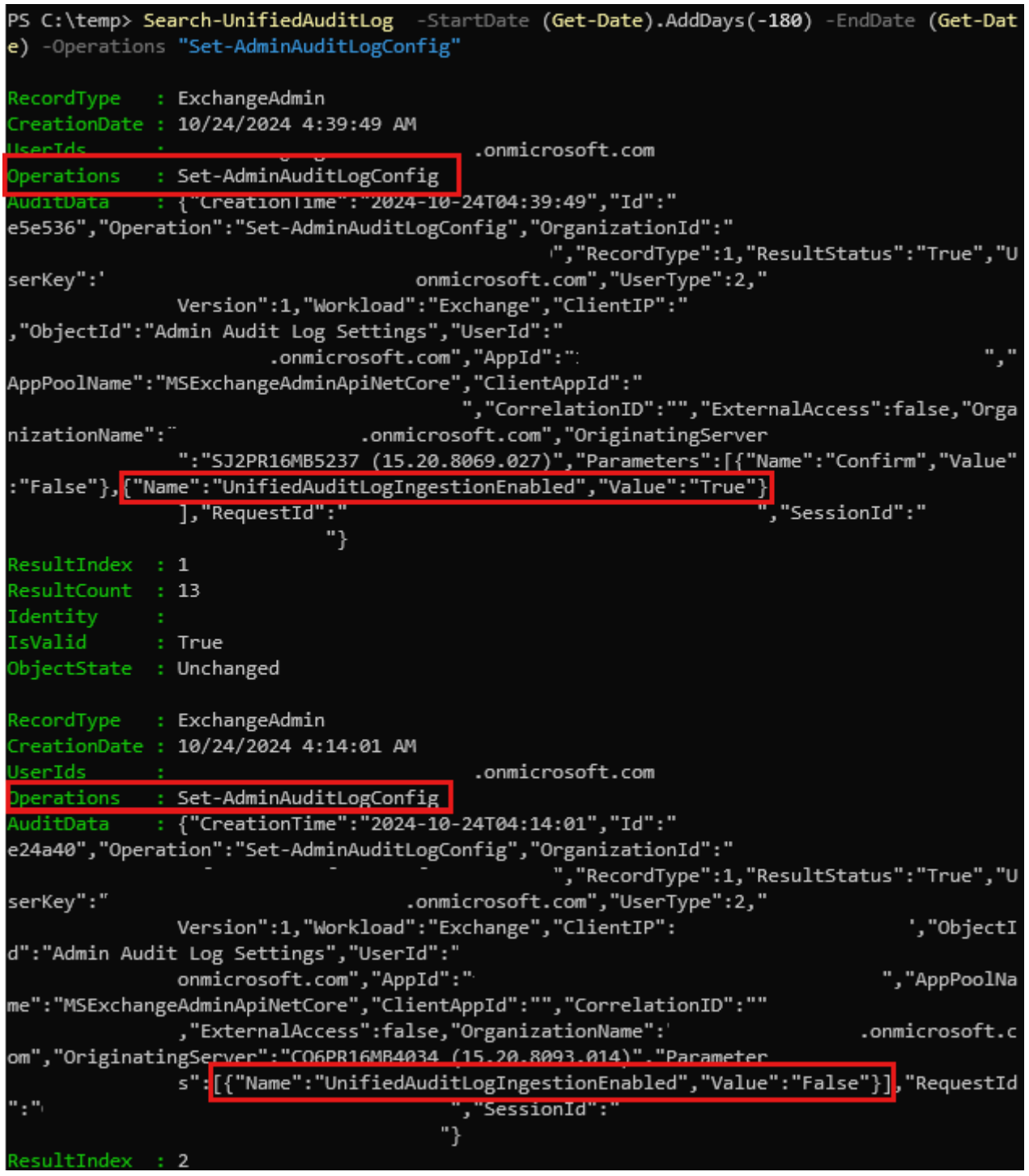
To search the Exchange admin audit log for audit records that are generated when turning auditing on or off, run the following command in Exchange Online PowerShell:

PowerShell

```
Search-UnifiedAuditLog -Operations Set-AdminAuditLogConfig
```

Audit records for turning auditing on or off contain information about when the auditing status was changed, the admin who changed it, and the IP address of the computer that was used to make the change. The following screenshots show audit records that correspond to changing the auditing status in your organization.

Audit record for Set-AdminAuditLogConfig



Look for the result value of `UnifiedAuditLogIngestionEnabled` in the *AuditData* property. The `UnifiedAuditLogIngestionEnabled` value indicates whether unified audit logging was turned on or off in the Microsoft Purview portal or by running the **Set-AdminAuditLogConfig -UnifiedAuditLogIngestionEnabled $true/false** cmdlet.

For more information about searching the Exchange admin audit log, see Search-UnifiedAuditLog.

Next steps

- Get started with auditing solutions: Complete the remaining setup steps after enabling auditing, including permissions and retention policies.
- Search the audit log: Start searching for user and admin activities now that auditing is turned on.
- Manage mailbox auditing: Verify and customize which mailbox actions are audited in your organization.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/19/2026

Search the audit log

Search in Microsoft Purview Audit (Standard) and Audit (Premium) gives your organization access to critical audit log event data so you can gain insight and further investigate user activities.

- Search jobs that you start through the Microsoft Purview portal don't need the web browser window to stay open to finish. These jobs keep running even after you close the browser window.
- The system now keeps completed search jobs for 30 days, so you can look back at past audit searches.
- Each admin Audit account user can have up to 10 search jobs running at the same time, with a limit of one unfiltered search job.

Before you search the audit log

Review the following items before you start searching the audit log.

- Audit log search is turned on by default for Microsoft 365 and Office 365 enterprise organizations. Verify the current unified audit log ingestion setting for your organization by running the following command in **Exchange Online PowerShell**:

  ```
  PowerShell
  Get-AdminAuditLogConfig | Format-List UnifiedAuditLogIngestionEnabled
  ```

  The value of `True` for the *UnifiedAuditLogIngestionEnabled* property indicates that audit log search is turned on. For more information, see Turn audit log search on or off.

  > **Important**
  >
  > Be sure to run the previous command in Exchange Online PowerShell. Although the **Get-AdminAuditLogConfig** cmdlet is also available in Security & Compliance PowerShell, the *UnifiedAuditLogIngestionEnabled* property is always `False`, even when audit log search is turned on.

- You must be assigned the *Audit Logs* or *View-Only Audit Logs* roles in the Microsoft Purview portal to search the audit log. For more information, see Get started with auditing solutions. To access audit cmdlets, you must be assigned the *Audit Logs* or *View-Only Audit Logs* roles

in the Exchange admin center. You can also create custom role groups with the ability to search the audit log by adding the *View-Only Audit Logs* or *Audit Logs* roles to a custom role group. For more information, see Permissions in the Microsoft Purview portal.

- When a user or admin performs an audited activity, the system generates an audit record and stores it in the audit log for your organization. The length of time that an audit record is retained (and searchable in the audit log) depends on your Office 365 or Microsoft 365 Enterprise subscription, and specifically the type of the license that is assigned to specific users.

  - For users assigned an Office 365 E5 or Microsoft 365 E5 license (or users with a Microsoft Purview Suite (formerly known as Microsoft 365 E5 Compliance) or Microsoft 365 E5 eDiscovery and Audit add-on license), the system retains audit records for Microsoft Entra ID, Exchange, and SharePoint activity for one year by default. Organizations can also create audit log retention policies to retain audit records for activities in other services for up to one year. For more information, see Manage audit log retention policies.

    > **Note**
    >
    > If your organization participated in the private preview program for the one-year retention of audit records, the retention duration for audit records that were generated before the general availability rollout date won't be reset.

  - For users assigned any other (non-E5) Office 365 or Microsoft 365 license, the system retains audit records for 180 days. For a list of Office 365 and Microsoft 365 subscriptions that support unified audit logging, see the subscription requirements for Audit (Standard) and Audit (Premium).

    > **Important**
    >
    > The default retention period for Audit (Standard) changed from 90 days to 180 days. Audit (Standard) logs generated before October 17, 2023 are retained for 90 days. Audit (Standard) logs generated on or after October 17, 2023 follow the new default retention of 180 days.

    > **Note**

> Even when mailbox auditing on by default is turned on, you might notice that mailbox audit events for some users aren't found in audit log searches in the Microsoft Purview portal or via the Office 365 Management Activity API. For more information, see **Mailbox audit logging**.

- To turn off audit log search for your organization, use the following command to disable unified audit log ingestion in Exchange Online PowerShell:

  ```
  PowerShell
  Set-AdminAuditLogConfig -UnifiedAuditLogIngestionEnabled $false
  ```

  To turn unified audit log ingestion back on for your organization, run the following command in Exchange Online PowerShell:

  ```
  PowerShell
  Set-AdminAuditLogConfig -UnifiedAuditLogIngestionEnabled $true
  ```

  For more information, see Turn off audit log search.

- The underlying cmdlet used to search the audit log is an Exchange Online cmdlet, which is **Search-UnifiedAuditLog**. That means you can use this cmdlet to search the audit log instead of using the search tool on the **Audit** page in the Microsoft Purview portal. You have to run this cmdlet in Exchange Online PowerShell. For more information, see Search-UnifiedAuditLog.

  For information about exporting the search results returned by the **Search-UnifiedAuditLog** cmdlet to a CSV file, see Tips for exporting and viewing the audit log.

- To programmatically download data from the audit log, use the Office 365 Management Activity API instead of a PowerShell script. The Office 365 Management Activity API is a REST web service that you can use to develop operations, security, and compliance monitoring solutions for your organization. For more information, see Office 365 Management Activity API reference.

- Microsoft Entra ID is the directory service for Microsoft 365. The unified audit log contains user, group, application, domain, and directory activities performed in the Microsoft 365 admin center or in the Azure management portal. For a complete list of Microsoft Entra events, see Microsoft Entra audit Report Events.

- Microsoft doesn't guarantee a specific time after an event occurs for the corresponding audit record to be returned in the results of an audit log search. For core services (such as Exchange, SharePoint, OneDrive, and Teams), audit record availability is typically 60 to 90 minutes after an event occurs. For other services, audit record availability might be longer. However, some issues that are unavoidable (such as a server outage) might occur outside of the audit service that delays the availability of audit records. For this reason, Microsoft doesn't commit to a specific time.
- To search for Power BI activities in the audit log, you have to enable auditing in the Power BI admin portal. For instructions, see the "Audit logs" section in Power BI admin portal.

Get started with search

> **Important**
>
> The execution time (search time) of a Search job depends on several factors, including the selected date range, the number of users included, the volume of activities, and overall tenant size. In large tenants with a high number of users, broadly-scoped Search jobs may take up to 48 hours to complete.

Complete the following steps to get started with search:

1. Sign in to the Microsoft Purview portal.
2. Select the **Audit** solution card. If the **Audit** solution card isn't displayed, select **View all solutions** and then select **Audit** from the **Core** section.
3. On the **Search** page, configure the following search criteria as applicable:
    a. **Date and time range (UTC)**: The last seven days are selected by default. Select a date and time range to display the events that occurred within that period. The date and time are presented in Coordinated Universal Time (UTC). The maximum date range that you can specify is 180 days. An error is displayed if the selected date range is greater than 180 days.

    > **Tip**
    >
    > If you use the maximum date range of 180 days, select the current time for the **Start date**. Otherwise, you receive an error saying that the start date is earlier than the

end date. If you turn on auditing within the last 180 days, the maximum date range can't start before the date that auditing was turned on.

b. **Keyword Search**: Enter a keyword or phrase to search for in the audit log. The keyword or phrase is searched for in the audit log or in the file, folder, or sites (if specified) for the search. To search for text that contains special characters, replace the special characters with an asterisk(*) in your keyword search. For example, to search for *test_search_document*, use *test*search*document*.

> **Important**
>
> Terms entered in the **Keyword Search** field are only searched within indexed content (content within the Audit *common schema*). Audit *data content* in the audit log isn't searched for these keywords.

c. **Admin Units**: Select the drop-down list to display the administrative units you want the audited activities scoped to for your search. You can select one or more administrative units to scope your search to. Leave this box blank to return entries for all administrative units in your organization.

d. **Activities - friendly names**: Select the drop-down list to display the friendly names for audited activities that you can search for. Friendly names for user and admin activities are organized into groups of related activities. Using friendly names, you can select specific audited activities or you can select the activity group name to select all activities in the group. You can also select a selected activity to clear the selection. To search for a friendly name in the **Activities - friendly names** drop-down list, use the search box above that list.

e. **Activities - operations names**: Enter the exact operation names to search for audited activities to include in your search results. You can enter one or more operation names, separated by commas. The **Activities - operations names** search criterion is similar to previous searches only available in PowerShell and provides greater flexibility helping you find the data that you need.

> **Important**

> Operation names must be entered exactly as they are named. If operation names are entered incorrectly, no results are returned.

For example, to search for all activities related to enabling and disabling information barriers for a SharePoint site in your organization, you would:

- Review the audit activities article to find the exact operation name for the information barriers activities you want to search for. In this information barriers activities example, the operation names are *SPOIBIsEnabled* and *SPOIBIsDisabled*.
- Enter *SPOIBIsEnabled,SPOIBIsDisabled* in the **Activities - operations names** field. Copy and paste the operation names directly from audit activities to the **Activities - operations names** field to ensure that they're entered correctly and without typos.

f. **Record types**: Select the drop-down list to display the record types for audited activities that you can search for. You can select one or more record types to search for. To search for a record type in the list, use the search box over the list.

   Specific record types are associated with specific Microsoft services and applications. For example, if you want to scope your search for specific record types associated with sensitivity labels in Microsoft Purview Information Protection (MIP), you could select the *MIPLabel*, *MipAutoLabelExchangeItem*, *MipAutoLabelSharePointItem*, and *MipAutoLabelSharePointPolicyLocation* record types from the list.

g. **Search name**: Enter in a custom name for your search job. This name is used to identify your search job in the search job history. If you don't enter a name, the search job is automatically named using a combination of the date and time defined for the search and other defined search criteria values.

h. **Users**: Select this field and select the names one or more users to display search results for. The audit log entries for the selected activity performed by the users you select in this box are displayed in the list of results. Leave this box blank to return entries for all users (and service accounts) in your organization.

i. **File, folder, or site**: Enter part or all of a file or folder name to find related activity. This returns results for matching files, folders, and sites. You can also enter a full URL or part of one, just avoid special characters or spaces. You can use * as a wildcard at the end of the URL. For example, `https://<tenantname>.sharepoint.com/sites/site123*`. Leave the field blank to see activity for all files and folders in your organization.

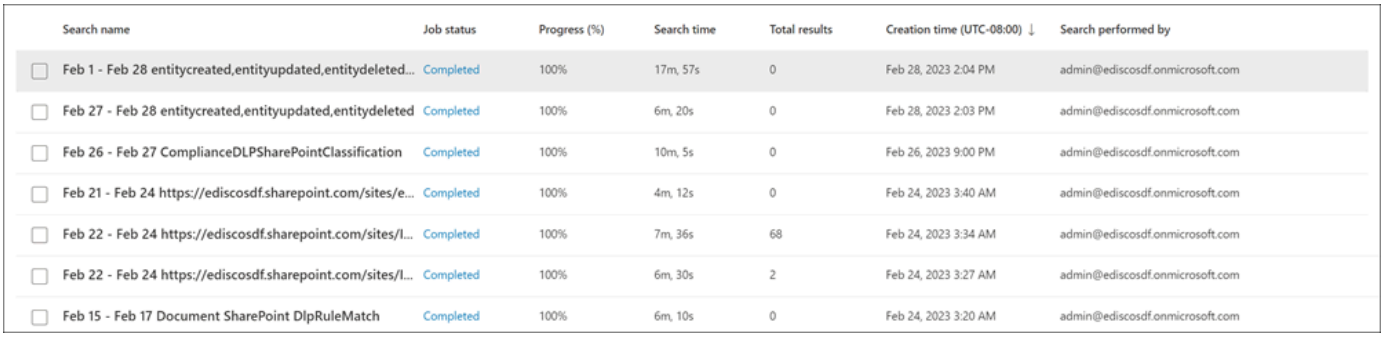
j. **Workloads**: Enter or search for workload services to search for activity related to the selected workloads. Enter the name of a workload to jump to the workload in the list or scroll to the workloads you'd like to select.

4. Select **Search** to start your search job. A maximum of 10 search jobs can be run in parallel for one user account. If a user requires more than 10 search jobs, they must wait for an *In progress* job to finish or delete a search job.

Search job dashboard

The search job dashboard displays active and completed search jobs. For each search job, the dashboard shows the following information:

- **Search name**: The name of the search job. You can see the full search name for a job by hovering the cursor over the search job name.
- **Job status**: The status of the search job. The status can be *Queued*, *In Progress*, or *Completed*.
- **Progress (%)**: The percentage of the search job that the job completes.
- **Search time**: The total running time that elapses to complete the search job.
- **Total results**: The total number of results the search job returns.
- **Creation time**: The date and time the search job created in UTC.
- **Search performed by**: The user account that creates the search job.



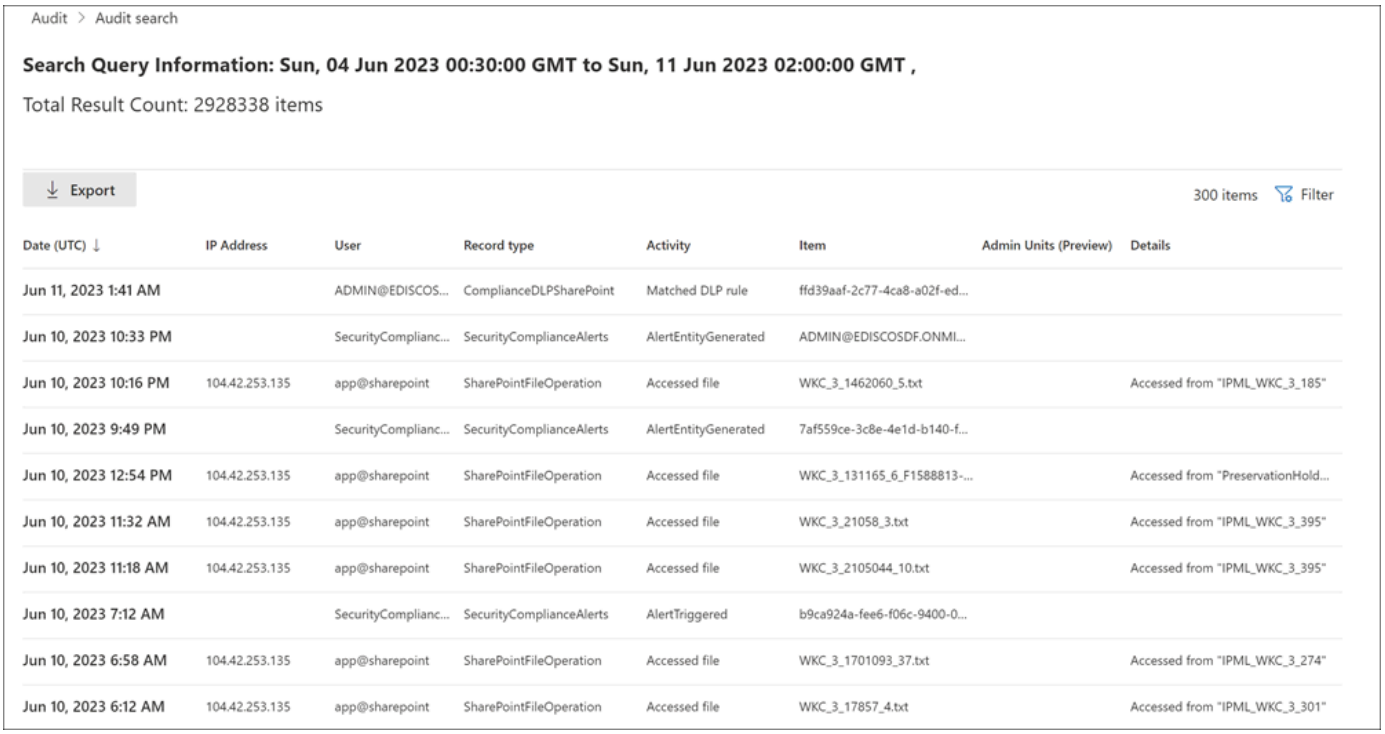
Search name	Job status	Progress (%)	Search time	Total results	Creation time (UTC-08:00) ↓	Search performed by
Feb 1 - Feb 28 entitycreated,entityupdated,entitydeleted...	Completed	100%	17m, 57s	0	Feb 28, 2023 2:04 PM	admin@ediscosdf.onmicrosoft.com
Feb 27 - Feb 28 entitycreated,entityupdated,entitydeleted	Completed	100%	6m, 20s	0	Feb 28, 2023 2:03 PM	admin@ediscosdf.onmicrosoft.com
Feb 26 - Feb 27 ComplianceDLPSharePointClassification	Completed	100%	10m, 5s	0	Feb 26, 2023 9:00 PM	admin@ediscosdf.onmicrosoft.com
Feb 21 - Feb 24 https://ediscosdf.sharepoint.com/sites/e...	Completed	100%	4m, 12s	0	Feb 24, 2023 3:40 AM	admin@ediscosdf.onmicrosoft.com
Feb 22 - Feb 24 https://ediscosdf.sharepoint.com/sites/l...	Completed	100%	7m, 36s	68	Feb 24, 2023 3:34 AM	admin@ediscosdf.onmicrosoft.com
Feb 22 - Feb 24 https://ediscosdf.sharepoint.com/sites/l...	Completed	100%	6m, 30s	2	Feb 24, 2023 3:27 AM	admin@ediscosdf.onmicrosoft.com
Feb 15 - Feb 17 Document SharePoint DlpRuleMatch	Completed	100%	6m, 10s	0	Feb 24, 2023 3:20 AM	admin@ediscosdf.onmicrosoft.com

Deleting a search job removes only the search job definition and the associated search result, not the backend data associated with search. To delete a search job, select the search job in the dashboard and then select **Delete** on the command bar.

To copy the search criteria for an existing search job, select the search job you want to copy and then select **Copy this search** on the command bar. The search criteria are copied to the search page and you can modify the search criteria as needed for a new search.

Search job details dashboard

To view details about a search job, select the search job in the dashboard. The dashboard shows the total number of items in the job at the top. The total result number removes duplicates, so it might be less than the number of items in the search job dashboard.



Date (UTC)	IP Address	User	Record type	Activity	Item	Admin Units (Preview)	Details
Jun 11, 2023 1:41 AM		ADMIN@EDISCOS...	ComplianceDLPSharePoint	Matched DLP rule	ffd39aaf-2c77-4ca8-a02f-ed...		
Jun 10, 2023 10:33 PM		SecurityComplianc...	SecurityComplianceAlerts	AlertEntityGenerated	ADMIN@EDISCOSDF.ONMI...		
Jun 10, 2023 10:16 PM	104.42.253.135	app@sharepoint	SharePointFileOperation	Accessed file	WKC_3_1462060_5.txt		Accessed from "IPML_WKC_3_185"
Jun 10, 2023 9:49 PM		SecurityComplianc...	SecurityComplianceAlerts	AlertEntityGenerated	7af559ce-3c8e-4e1d-b140-f...		
Jun 10, 2023 12:54 PM	104.42.253.135	app@sharepoint	SharePointFileOperation	Accessed file	WKC_3_131165_6_F1588813-...		Accessed from "PreservationHold...
Jun 10, 2023 11:32 AM	104.42.253.135	app@sharepoint	SharePointFileOperation	Accessed file	WKC_3_21058_3.txt		Accessed from "IPML_WKC_3_395"
Jun 10, 2023 11:18 AM	104.42.253.135	app@sharepoint	SharePointFileOperation	Accessed file	WKC_3_2105044_10.txt		Accessed from "IPML_WKC_3_395"
Jun 10, 2023 7:12 AM		SecurityComplianc...	SecurityComplianceAlerts	AlertTriggered	b9ca924a-fee6-f06c-9400-0...		
Jun 10, 2023 6:58 AM	104.42.253.135	app@sharepoint	SharePointFileOperation	Accessed file	WKC_3_1701093_37.txt		Accessed from "IPML_WKC_3_274"
Jun 10, 2023 6:12 AM	104.42.253.135	app@sharepoint	SharePointFileOperation	Accessed file	WKC_3_17857_4.txt		Accessed from "IPML_WKC_3_301"

The search job details dashboard displays the following information about the individual items gathered in the search job results:

- **Date (UTC)**: The date and time the activity occurred.
- **IP Address**: The IP address of the device that was used to perform the activity.
- **User**: The user account that performed the activity.
- **Record type**: The record type associated with the activity.
- **Activity**: The friendly name of the activity that was performed.
- **Item**: The name of the file, folder, or site that the activity was acted on.
- **Admin Units**: The admin unit that the user account that performed the activity belongs to.
- **Details**: Additional details about the activity.

You can sort the search job items by using the column headers or create a custom filter by using the filter pane. Use the filter to filter the search job items for specific values for any of the dashboard column criteria. To export all search job items to a .csv file, select **Export** on the command bar. Export supports results up to 50,000 rows for Audit (Standard) and up to 1,000,000 rows for Audit (Premium).

Select a specific activity to see more details about the activity in a fly-out window. The fly-out window displays the additional information about the activity.

Detail

Date (UTC)

2023-02-20T06:45:05

IP Address

Users

i:0i.t|00000003-0000-0ff1-ce00-000000000000|app@sharepoint

Activity

FileAccessed

Item

https://contoso.sharepoint.com/sites/IPML_WKC_3_164/IPML_WKC_3_164/WKC_3_1411322_1.txt

Detail

AppAccessContext


```
{
    "ClientAppId": "00001111-aaaa-2222-bbbb-3333cccc4444",
    "ClientAppName": "0a0a0a0a-1111-bbbb-2222-3c3c3c3c3c3c",
    "CorrelationId": "dddd3333-ee44-5555-66ff-777777aaaaaa ",
    "TokenIssuedAtTime": "2023-02-20T06:36:25",
    "UniqueTokenId": "HWLdylVJX0G6_QuqC0ULAA"
}
```


CreationTime

2023-02-20T06:45:05

Id

Operation

FileAccessed

OrganizationId

RecordType

6

UserKey

i:0i.t|00000003-0000-0ff1-ce00-000000000000|app@sharepoint

Close

Scoping access to audit logs using administrative units

Access to search the audit log is based on the administrative units assigned to the user in the Microsoft Purview portal. A restricted admin can only search and export user-generated audit logs within the scope of their assigned administrative units. An unrestricted admin has access to all audit logs, including logs generated by non-user and system accounts. To access scoped activity logs from any Microsoft service, including Exchange mailbox activity logs, use the Search-UnifiedAuditLog cmdlet.

Expand table

Admin units assigned to admins	Admin units available to perform scoped search on	Access to search and export audit logs
None (Default): Unrestricted admin	All administrative units are available	Access to all activity logs from any user, nonuser, or system account.
One or more administrative units: Restricted admin	Only those administrative units assigned to the admin are available	Access to activity logs from users with a matching administrative unit assignment.

Only unrestricted admins can access the following audit activities through search queries. These logs are not yet accessible when queried by a restricted admin. To view a complete list of audit logs for these activities, submit a search request by using an unrestricted admin account.

Expand table

Service	Operation
Azure Information Protection	Discover
Dynamics 365	CrmDefaultActivity
Endpoint data loss prevention	FileCreated FileCreatedOnNetworkShare FileCreatedOnRemovableMedia FileDeleted
Exchange	Set-Mailbox Set-MailboxPlan SupervisionBulkEmailExclusion
Microsoft Forms	ViewRuntimeForm

For more information about administrative units, see Permissions in the Microsoft Purview portal.

Next steps

- Export, configure, and view audit log records: Export your search results to CSV and format them in Excel for analysis.
- Use a PowerShell script to search the audit log: Automate large or recurring audit log searches using PowerShell.
- Search the audit log to investigate common support issues: Apply your search skills to common investigation scenarios like compromised accounts and email forwarding.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/19/2026

Export, configure, and view audit log records

When you search the audit log and download the search results to a comma-separated value (.csv) file, the file contains a column named **AuditData**. This column holds extra information about each event. The data in this column is formatted as a JSON object. This object contains multiple properties that appear as *property:value* pairs separated by commas. You can use the JSON transform feature in the Power Query Editor in Excel to split each property in the JSON object in the **AuditData** column into multiple columns so that each property has its own column. This feature lets you sort and filter on one or more of these properties, which can help you quickly locate the specific auditing data you're looking for.

Step 1: Export audit log search results

> **Important**
>
> For export processes to work correctly, make sure that your network firewall doesn't block the Azure Front Door domain (azurefd.net).

When you use Audit (Standard), you can export up to **50,000 records** to a .csv file from a single audit search query. For Audit (Premium), the export limit increases to **1,000,000 records**. If the number of results returned by your audit search query exceeds these limits, the exported .csv file doesn't include all results and might omit some audit logs. To ensure complete data export, consider the following recommendations:

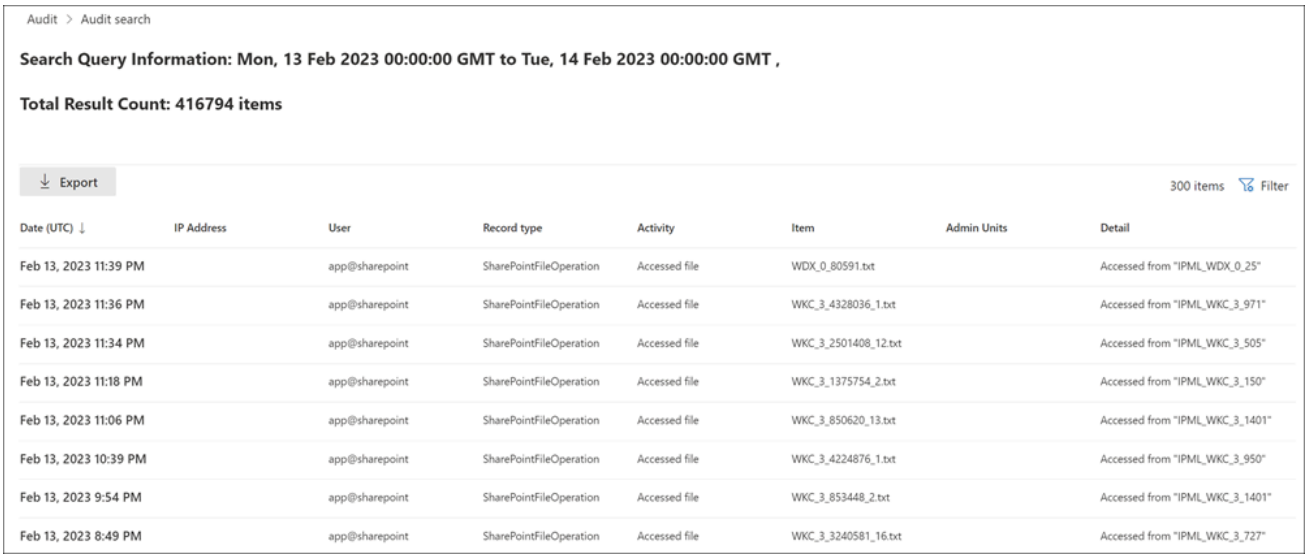
1. **Refine the scope of your search query** by narrowing the date range or applying other filters like UserId, Operation, and more.
2. **Perform multiple searches** with smaller, segmented date ranges to capture all relevant audit logs.

Narrowing the search scope and splitting large date ranges into smaller segments helps ensure comprehensive coverage of audit data within the export limits.

To search the audit log and export the results in a .csv file to your local computer, complete the following steps:

1. Run an audit log search and revise the search criteria if necessary until you have the desired results.

2. On the search results page, select **Export**.



   This option exports all the audit records from the audit log search you ran in step 1. It adds the raw data from the audit log to a .csv file. It takes a while to prepare the download file for a large search. Large files result when searching for all activities or using a wide date range.

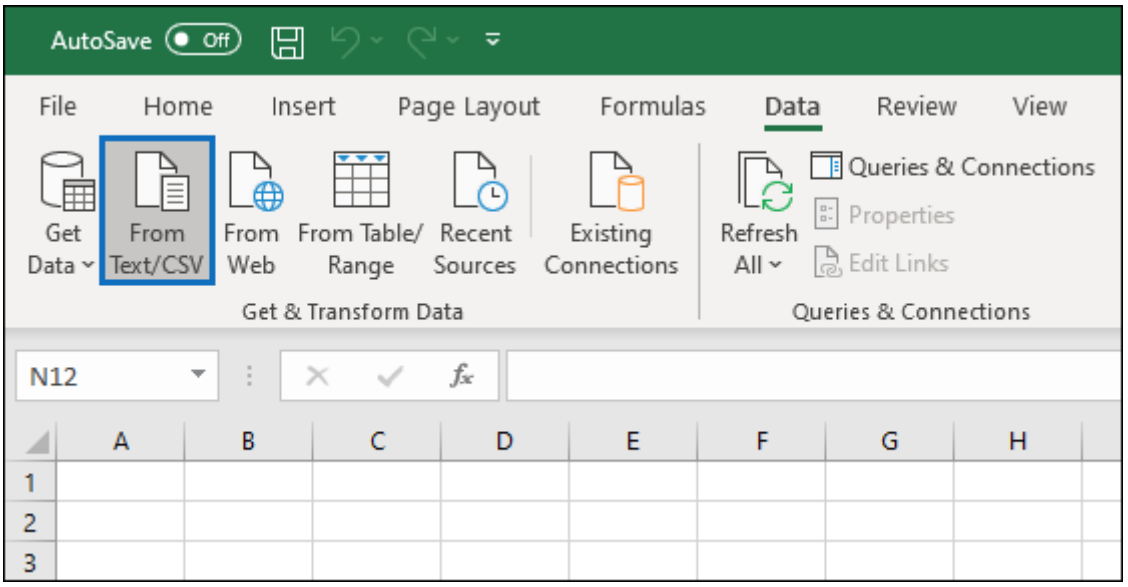
3. When the export process completes, a message appears at the top of the window that prompts you to open the .csv file and save it to your local computer. You might need to refresh the page to see this message. You can also access the .csv file in the **Downloads** folder.

Step 2: Format the exported audit log using the Power Query Editor

In this step, you use the JSON transform feature in the Power Query Editor in Excel to split each property in the JSON object in the **AuditData** column into its own column. Then, you filter columns to view records based on the values of specific properties. This process helps you quickly locate the specific auditing data you're looking for.

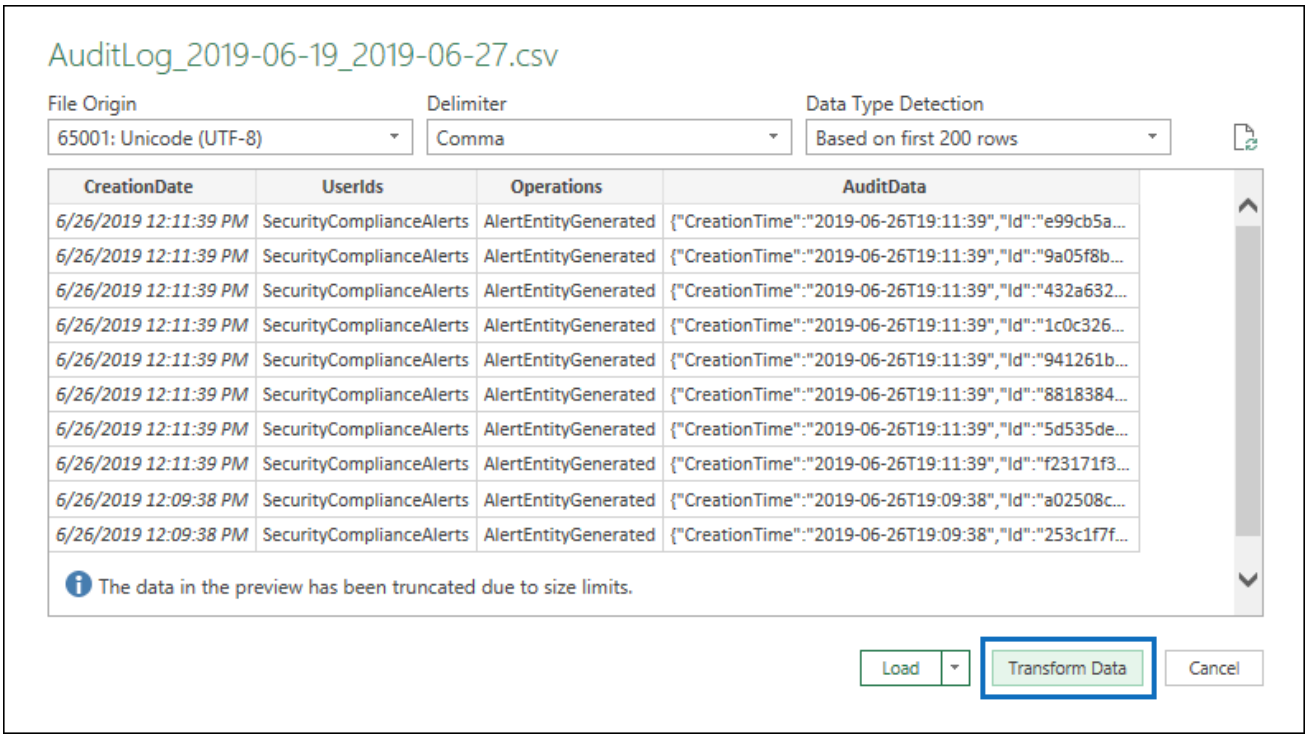
1. Open a blank workbook in Excel for Microsoft 365, Excel 2019, or Excel 2016.

2. On the **Data** tab, in the **Get & Transform Data** ribbon group, select **From Text/CSV**.



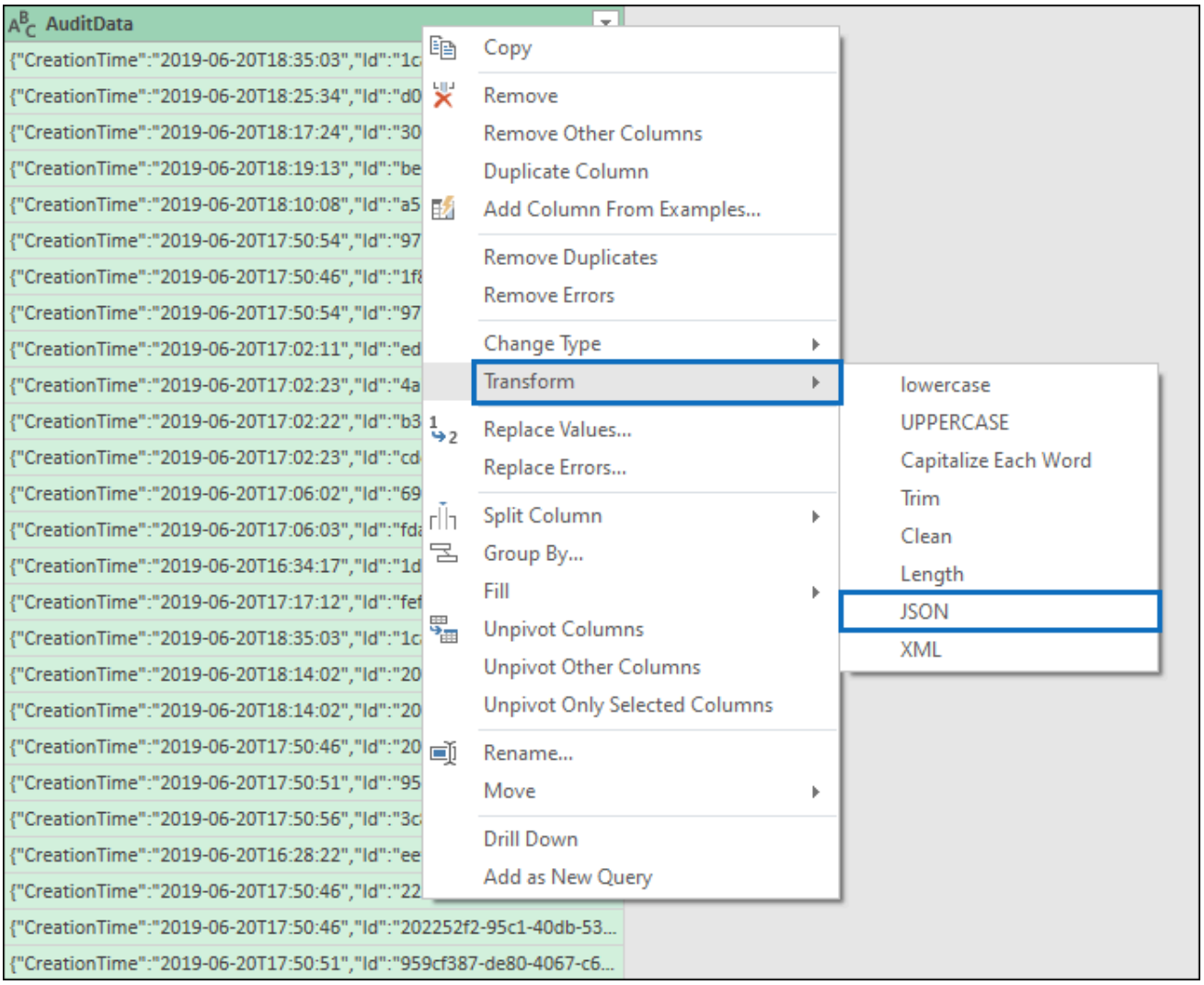
3. Open the .csv file that you downloaded in Step 1.

4. Select **Transform Data** in the window that's displayed.

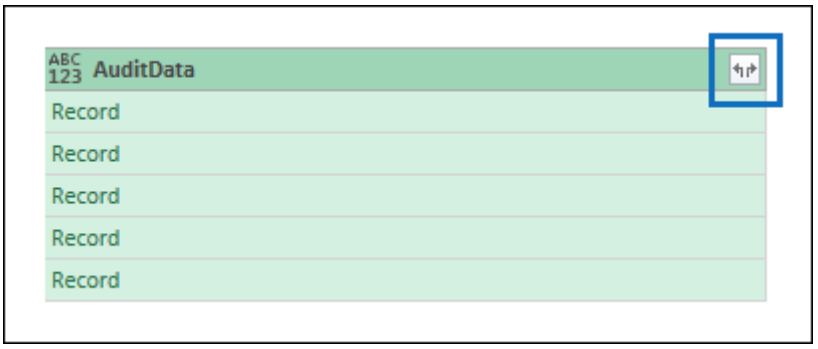


The **Query Editor** opens the .csv file. You see four columns: **CreationDate**, **UserIds**, **Operations**, and **AuditData**. The **AuditData** column is a JSON object that contains multiple properties. You need to create a column for each property in the JSON object.

5. Right-click the title in the **AuditData** column, select **Transform**, and then select **JSON**.

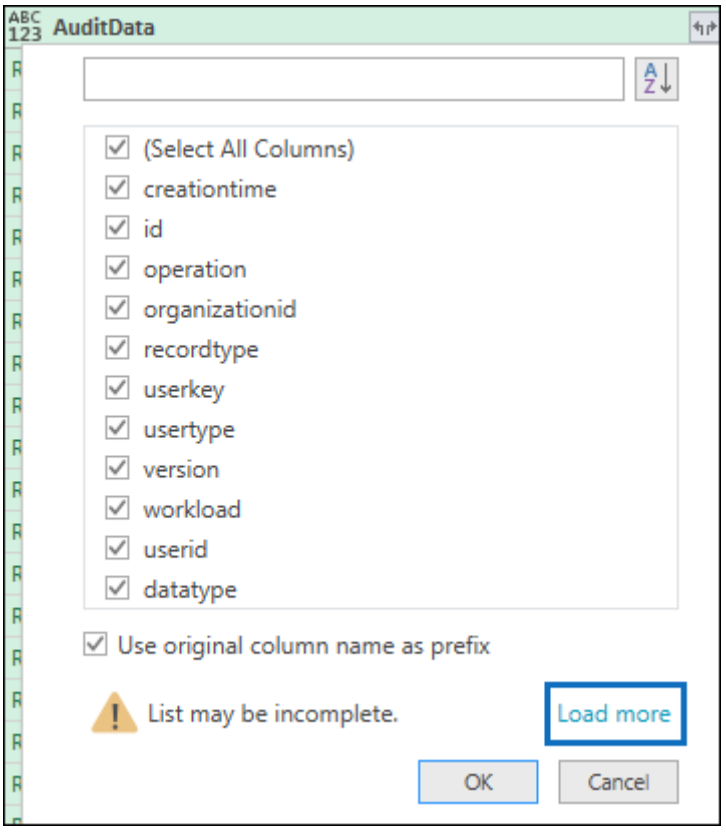


6. Select the expand icon in the upper-right corner of the **AuditData** column.



You see a partial list of the properties in the JSON objects in the **AuditData** column.

7. Select **Load more** to display all properties in the JSON objects in the **AuditData** column.



You can clear the check box next to any property that you don't want to include. Eliminating columns that aren't useful for your investigation is a good way to reduce the amount of data displayed in the audit log.

> **Note**
>
> The JSON properties displayed in the previous screenshot (after you select **Load more**) are based on the properties found in the **AuditData** column from the first 1,000 rows in the .csv file. If different JSON properties exist in records after the first 1,000 rows, these properties (and a corresponding column) don't appear when you split the **AuditData** column into multiple columns. To help prevent this issue, consider rerunning the audit log search and narrow the search criteria so that fewer records return. Another workaround is to filter items in the **Operations** column to reduce the number of rows (before you perform step 5) before transforming the JSON object in the **AuditData** column.

> **Tip**
>
> To view an attribute within a list such as AuditData.AffectedItems, select the **Expand** icon in the upper right corner of the column you want to pull an attribute from, and then select **Expand to New Row**. From there it's a record and you can select the

Expand icon in the upper right corner of the column, view the attributes, and select the one you want to view or extract.

8. Do one of the following things to format the title of the columns that you add for each JSON property that you select.

   - Clear the **Use original column name as prefix** check box to use the name of the JSON property as the column names; for example, **RecordType** or **SourceFileName**.
   - Leave the **Use original column name as prefix** check box selected to add the AuditData prefix to the column names; for example, **AuditData.RecordType** or **AuditData.SourceFileName**.

9. Select **OK**.

   You split the **AuditData** column into multiple columns. Each new column corresponds to a property in the AuditData JSON object. Each row in the column contains the value for the property. If the property doesn't contain a value, the *null* value appears. In Excel, cells with null values are empty.

10. On the **Home** tab, select **Close & Load** to close the Power Query Editor and open the transformed .csv file in an Excel workbook.

Use PowerShell to search and export audit log records

Instead of using the audit log search tool in the Microsoft Purview portal, you can use the Search-UnifiedAuditLog cmdlet in Exchange Online PowerShell to export the results of an audit log search to a .csv file. Then you can follow the same procedure described in Step 2 to format the audit log using the Power Query editor. One advantage of using the PowerShell cmdlet is that you can search for events from a specific service by using the *RecordType* parameter. The following examples show how to use PowerShell to export audit records to a .csv file so you can use the Power Query editor to transform the JSON object in the **AuditData** column as described in Step 2.

In the following example, run the commands to return all records related to SharePoint sharing operations.

PowerShell

```
$auditlog = Search-UnifiedAuditLog -StartDate 06/01/2019 -EndDate 06/30/2019 -
RecordType SharePointSharingOperation
```

PowerShell

```
$auditlog | Select-Object -Property CreationDate,UserIds,RecordType,AuditData |
Export-Csv -Path c:\AuditLogs\PowerShellAuditlog.csv -NoTypeInformation
```

The search results are exported to a .csv file named *PowerShellAuditlog* that contains four columns: CreationDate, UserIds, RecordType, AuditData.

You can use the name or enum value for the record type. For a list of record type names and their corresponding enum values, see the *AuditLogRecordType* table in Office 365 Management Activity API schema.

You can only include a single value for the *RecordType* parameter. To search for audit records for other record types, run the two previous commands again to specify a different record type and append those results to the original .csv file. For example, run the following two commands to add SharePoint file activities from the same date range to the PowerShellAuditlog.csv file.

PowerShell

```
$auditlog = Search-UnifiedAuditLog -StartDate 06/01/2019 -EndDate 06/30/2019 -
RecordType SharePointFileOperation
```

PowerShell

```
$auditlog | Select-Object -Property CreationDate,UserIds,RecordType,AuditData |
Export-Csv -Append -Path c:\AuditLogs\PowerShellAuditlog.csv -NoTypeInformation
```

Tips for exporting and viewing the audit log

Here are some tips and examples of exporting and viewing the audit log before and after you use the JSON transform feature to split the **AuditData** column into multiple columns.

- Filter the **RecordType** column to display only the records from a specific service or functional area. For example, to show events related to SharePoint sharing, select **14** (the enum value for records triggered by SharePoint sharing activities). For a list of the services that correspond to the enum values displayed in the **RecordType** column, see Detailed properties in the audit log.
- Filter the **Operations** column to display the records for specific activities. For a list of most operations that correspond to a searchable activity in the audit log search tool in the Microsoft Purview portal, see the "Audited activities" section in Search the audit log.

Run the diagnostic for audit log configuration check

> **Note**
>
> You must be a Microsoft 365 global administrator to run the diagnostic.

To run the diagnostic for audit log configuration check, complete the following steps:

1. Select the diagnostic in the Microsoft 365 admin center.
2. Enter the User principal name (UPN) or email address of the user missing Audit entries.

> **Note**
>
> If you don't enter user details, the diagnostic performs general audit configuration validation checks.

1. Select **Run Tests**.
2. Review the key findings from the diagnostic to check the user and audit configuration for your organization.

Next steps

- Use a PowerShell script to search the audit log: Automate recurring searches and export large result sets using PowerShell.
- Detailed activity properties in the audit log: Understand the properties in your exported data to interpret results accurately.
- Search the audit log to investigate common support issues: Use exported data to investigate scenarios like compromised accounts or unauthorized access.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 06/19/2026

Use a PowerShell script to search the audit log

Security, compliance, and auditing are top priorities for IT administrators. Microsoft 365 has several built-in capabilities to help organizations manage security, compliance, and auditing. In particular, unified audit logging can help you investigate security incidents and compliance issues. You can retrieve audit logs by using the following methods:

- The Office 365 Management Activity API
- The audit log search tool in the Microsoft Purview portal
- The Search-UnifiedAuditLog cmdlet in Exchange Online PowerShell

If you need to regularly retrieve audit logs, consider a solution that uses the Office 365 Management Activity API. It provides large organizations with the scalability and performance to retrieve millions of audit records on an ongoing basis. Using the audit log search tool in Microsoft Purview portal is a good way to quickly find audit records for specific operations that occur in shorter time range. Using longer time ranges in the audit log search tool, especially for large organizations, might return too many records to easily manage or export.

When you need to manually retrieve auditing data for a specific investigation or incident, particularly for longer date ranges in larger organizations, using the **Search-UnifiedAuditLog** cmdlet might be the best option. This article includes a PowerShell script that uses the cmdlet. The script can retrieve 50,000 audit records each time you run the cmdlet and then export them to a CSV file. You can format the file by using Power Query in Excel to help with your review. Using the script in this article also minimizes the chance that large audit log searches time out in the service.

Before you run the script

- Audit logging must be enabled for your organization to successfully use the script to return audit records. Audit logging is turned on by default for Microsoft 365 and Office 365 enterprise organizations. To verify that unified audit log ingestion is enabled for your organization, run the following command in Exchange Online PowerShell. This check confirms that audit records are being collected and available for search:

  ```
  PowerShell
  ```

```
Get-AdminAuditLogConfig | FL UnifiedAuditLogIngestionEnabled
```

The value of `True` for the **UnifiedAuditLogIngestionEnabled** property indicates that audit log search is turned on.

- You must be assigned the *View-Only Audit Logs* or *Audit Logs* role in Exchange Online to successfully run the script. By default, these roles are assigned to the *Compliance Management* and *Organization Management* role groups on the **Permissions** page in the Exchange admin center.

- It might take a long time for the script to complete. How long it takes to run depends on the date range and the size of the interval that you configure the script to retrieve audit records for. Larger date ranges and smaller intervals result in a long running time. See the variable description table later in this article for more information about the date range and intervals.

- The sample script provided in this article isn't supported under any Microsoft standard support program or service. The sample script is provided **as-is** without warranty of any kind. Microsoft further disclaims all implied warranties including, without limitation, any implied warranties of merchantability or of fitness for a particular purpose. The entire risk arising out of the use or performance of the sample script and documentation remains with you. In no event shall Microsoft, its authors, or anyone else involved in the creation, production, or delivery of the script be liable for any damages whatsoever (including, without limitation, damages for loss of business profits, business interruption, loss of business information, or other pecuniary loss) arising out of the use of or inability to use the sample script or documentation, even if Microsoft has been advised of the possibility of such damages.

Step 1: Connect to Exchange Online PowerShell

First, connect to Exchange Online PowerShell. You can connect by using modern authentication or multifactor authentication (MFA). For step-by-step instructions, see Connect to Exchange Online PowerShell.

Step 2: Modify and run the script to retrieve audit records

After you connect to Exchange Online PowerShell, create, modify, and run the script to retrieve the auditing data. The first seven lines in the audit log search script contain the following

variables that you can modify to configure your search. See the following variable description table for details about these variables.

1. Save the following text to a Windows PowerShell script by using a filename suffix of .ps1. For example, SearchAuditLog.ps1.

PowerShell

```
#Modify the values for the following variables to configure the audit log search.
$logFile = "d:\AuditLogSearch\AuditLogSearchLog.txt"
$outputFile = "d:\AuditLogSearch\AuditLogRecords.csv"
[DateTime]$start = [DateTime]::UtcNow.AddDays(-1)
[DateTime]$end = [DateTime]::UtcNow
$record = "AzureActiveDirectory"
$resultSize = 5000
$intervalMinutes = 60

#Start script
[DateTime]$currentStart = $start
[DateTime]$currentEnd = $end

Function Write-LogFile ([String]$Message)
{
    $final = [DateTime]::Now.ToUniversalTime().ToString("s") + ":" + $Message
    $final | Out-File $logFile -Append
}

Write-LogFile "BEGIN: Retrieving audit records between $($start) and $($end),
RecordType=$record, PageSize=$resultSize."
Write-Host "Retrieving audit records for the date range between $($start) and
$($end), RecordType=$record, ResultsSize=$resultSize"

$totalCount = 0
while ($true)
{
    $currentEnd = $currentStart.AddMinutes($intervalMinutes)
    if ($currentEnd -gt $end)
    {
        $currentEnd = $end
    }

    if ($currentStart -eq $currentEnd)
    {
        break
    }

    $sessionID = [Guid]::NewGuid().ToString() + "_" +  "ExtractLogs" + (Get-
Date).ToString("yyyyMMddHHmmssfff")
    Write-LogFile "INFO: Retrieving audit records for activities performed
between $($currentStart) and $($currentEnd)"
    Write-Host "Retrieving audit records for activities performed between
$($currentStart) and $($currentEnd)"
```

```
    $currentCount = 0

    $sw = [Diagnostics.StopWatch]::StartNew()
    do
    {
        $results = Search-UnifiedAuditLog -StartDate $currentStart -EndDate
$currentEnd -RecordType $record -SessionId $sessionID -SessionCommand
ReturnLargeSet -ResultSize $resultSize

        if (($results | Measure-Object).Count -ne 0)
        {
            $results | export-csv -Path $outputFile -Append -NoTypeInformation

            $currentTotal = $results[0].ResultCount
            $totalCount += $results.Count
            $currentCount += $results.Count
            Write-LogFile "INFO: Retrieved $($currentCount) audit records out of
the total $($currentTotal)"

            if ($currentTotal -eq $results[$results.Count - 1].ResultIndex)
            {
                $message = "INFO: Successfully retrieved $($currentTotal) audit
records for the current time range. Moving on!"
                Write-LogFile $message
                Write-Host "Successfully retrieved $($currentTotal) audit records
for the current time range. Moving on to the next interval." -foregroundColor
Yellow
                ""
                break
            }
        }
    }
    while (($results | Measure-Object).Count -ne 0)

    $currentStart = $currentEnd
}

Write-LogFile "END: Retrieving audit records between $($start) and $($end),
RecordType=$record, PageSize=$resultSize, total count: $totalCount."
Write-Host "Script complete! Finished retrieving audit records for the date range
between $($start) and $($end). Total count: $totalCount" -foregroundColor Green
```

Alternatively, you can use the moreRecordsAvailable parameter, as shown in the following example.

PowerShell

```
#Modify the values for the following variables to configure the audit log search.
$logFile = "d:\AuditLogSearch\AuditLogSearchLog.txt"
$outputFile = "d:\AuditLogSearch\AuditLogRecords.csv"
[DateTime]$start = [DateTime]::UtcNow.AddDays(-1)
[DateTime]$end = [DateTime]::UtcNow
$record = "AzureActiveDirectory"
```

```
$resultSize = 5000
$intervalMinutes = 60

#Start script
[DateTime]$currentStart = $start
[DateTime]$currentEnd = $end

Function Write-LogFile ([String]$Message)
{
    $final = [DateTime]::Now.ToUniversalTime().ToString("s") + ":" + $Message
    $final | Out-File $logFile -Append
}

Write-LogFile "BEGIN: Retrieving audit records between $($start) and $($end),
RecordType=$record, PageSize=$resultSize."
Write-Host "Retrieving audit records for the date range between $($start) and
$($end), RecordType=$record, ResultsSize=$resultSize"

$totalCount = 0
while ($true)
{
    $currentEnd = $currentStart.AddMinutes($intervalMinutes)
    if ($currentEnd -gt $end)
    {
        $currentEnd = $end
    }

    if ($currentStart -eq $currentEnd)
    {
        break
    }

    $sessionID = [Guid]::NewGuid().ToString() + "_" +  "ExtractLogs" + (Get-
Date).ToString("yyyyMMddHHmmssfff")
    Write-LogFile "INFO: Retrieving audit records for activities performed
between $($currentStart) and $($currentEnd)"
    Write-Host "Retrieving audit records for activities performed between
$($currentStart) and $($currentEnd)"
    $currentCount = 0

    $sw = [Diagnostics.StopWatch]::StartNew()
    do
    {
        $results = Search-UnifiedAuditLog -StartDate $currentStart -EndDate
$currentEnd -RecordType $record -SessionId $sessionID -SessionCommand
ReturnLargeSet -ResultSize $resultSize

        if (($results | Measure-Object).Count -ne 0)
        {
            $results | export-csv -Path $outputFile -Append -NoTypeInformation

            $totalCount += $results.Count
            $currentCount += $results.Count
            Write-LogFile "INFO: Retrieved $($currentCount) audit records so far
for the current time range (running ResultCount: $($results[$results.Count -
```

```
1].ResultCount))"

            if (-not $results[$results.Count -
1].AuditSearchRequestMetadata.moreRecordsAvailable)
            {
                $message = "INFO: Successfully retrieved $($currentCount) audit
records for the current time range. Moving on!"
                Write-LogFile $message
                Write-Host "Successfully retrieved $($currentCount) audit records
for the current time range. Moving on to the next interval." -foregroundColor
Yellow
                ""
                break
            }
    }
    while (($results | Measure-Object).Count -ne 0)

    $currentStart = $currentEnd
}

Write-LogFile "END: Retrieving audit records between $($start) and $($end),
RecordType=$record, PageSize=$resultSize, total count: $totalCount."
Write-Host "Script complete! Finished retrieving audit records for the date range
between $($start) and $($end). Total count: $totalCount" -foregroundColor Green
```

2. Modify the variables listed in the following table to configure the search criteria. The script includes sample values for these variables, but change them (unless stated otherwise) to meet your specific requirements.

Expand table

Variable	Sample value	Description
`$logFile`	"d:\temp\AuditSearchLog.txt"	Specifies the name and location for the log file that contains information about the progress of the audit log search performed by the script. The script writes UTC timestamps to the log file.
`$outputFile`	"d:\temp\AuditRecords.csv"	Specifies the name and location of the CSV file that contains the audit records returned by the script.
`[DateTime]$start` and `[DateTime]$end`	[DateTime]::UtcNow.AddDays(-1) [DateTime]::UtcNow	Specifies the date range for the audit log search. The script returns records for audit activities that occurred within the specified date range. For example, to return activities performed in January 2021, you can use a start date of `"2021-01-01"` and an end date of `"2021-01-31"` (be sure to surround the values in double-quotation marks) The sample

Variable	Sample value	Description
		value in the script returns records for activities performed in the previous 24 hours. If you don't include a timestamp in the value, the default timestamp is 12:00 AM (midnight) on the specified date.
`$record`	"AzureActiveDirectory"	Specifies the record type of the audit activities (also called *operations*) to search for. This property indicates the service or feature that an activity was triggered in. For a list of record types that you can use for this variable, see Audit log record type. You can use the record type name or ENUM value. **Tip:** To return audit records for all record types, use the value `$null` (without double-quotations marks).
`$resultSize`	5000	Specifies the number of results returned each time the **Search-UnifiedAuditLog** cmdlet is called by the script (called a *result set*). The value of 5,000 is the maximum value supported by the cmdlet. Leave this value as-is.
`$intervalMinutes`	60	To help overcome the limit of 5,000 records returned, this variable takes the data range you specified and slices it up into smaller time intervals. Now each interval, not the entire date range, is subject to the 5000 record output limit of the command. The default value of 5,000 records per 60-minute interval within the date range should be sufficient for most organizations. But, if the script returns an error that says, `maximum results limitation reached`, decrease the time interval (for example, to 30 minutes or even 15 minutes) and rerun the script.

Most of the variables listed in the previous table correspond to parameters for the **Search-UnifiedAuditLog** cmdlet. For more information about these parameters, see Search-UnifiedAuditLog.

1. On your local computer, open Windows PowerShell and go to the folder where you saved the modified script.

2. After saving and configuring the script, run it from the directory where the script file is stored to start the audit log search:

PowerShell

```
.\SearchAuditLog.ps1
```

The script displays progress messages while it's running. After the script finishes running, it creates the log file and the .csv file that contains the audit records and saves them to the folders defined by the `$logFile` and `$outputFile` variables.

> **Important**
>
> There's a 50,000 limit for the maximum number of audit records returned each time you run the cmdlet in the script. If you run this script and it returns 50,000 results, audit records for activities that occurred within the date range likely weren't included. If this happens, divide the date range into smaller durations and then rerun the script for each date range. For example, if a date range of 90 days returns 50,000 results then you can rerun the script twice, once for the first 45 days in the date range and then again for the next 45 days.

Step 3: Format and view the audit records

After running the script and exporting the audit records to a .csv file, you might want to format the .csv to make it easier to review and analyze the audit records. One way to do this is to use the Power Query JSON transform feature in Excel to split each property in the JSON object in the **AuditData** column into its own column. For step-by-step instructions, see "Step 2: Format the exported audit log using the Power Query Editor" in Export, configure, and view audit log records.

Next steps

- Export, configure, and view audit log records: Format and view your exported audit records using Power Query in Excel.
- Detailed activity properties in the audit log: Interpret the properties in your exported results to understand activity details.
- Manage audit log retention policies: Ensure the audit data your scripts query is retained for the required period.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 06/19/2026

Manage mailbox auditing

Mailbox audit logging is turned on by default in all organizations. This setting automatically logs certain actions that mailbox owners, delegates, and admins perform. Admins can search the mailbox audit log for the corresponding mailbox audit records.

Some benefits of mailbox auditing on by default include:

- Auditing is automatically turned on when you create a new mailbox. You don't need to manually enable mailbox auditing for new users.
- You don't need to manage the mailbox actions that are audited. A predefined set of mailbox actions are audited by default for each sign-in type (Admin, Delegate, and Owner).
- When Microsoft releases a new mailbox action, the action might be automatically added to the list of mailbox actions that are audited by default, subject to the user having the appropriate license. Because new actions are added automatically, you don't need to add new actions on mailboxes as they're released.
- You have a consistent mailbox auditing policy across your organization because you're auditing the same actions for all mailboxes.

Verify mailbox auditing on by default is turned on

To verify that mailbox auditing on by default is turned on for your organization, run the following command in Exchange Online PowerShell:

PowerShell

```
Get-OrganizationConfig | Format-List AuditDisabled
```

The value **False** indicates that mailbox auditing on by default is turned on for the organization. Mailbox auditing on by default in the organization overrides whether auditing is enabled or disabled on individual mailboxes. It doesn't override mailbox-level audit operation settings.

How organization-level and mailbox-level auditing work together:

- **If organization-level auditing is ON and mailbox-level auditing is OFF**: Organization-level auditing overrides the mailbox setting and enables auditing. The mailbox's configured audit operations apply.

- **If organization-level auditing is ON and mailbox-level auditing is ON**: Mailbox-level auditing remains enabled, and the mailbox's configured audit operations apply.

For example, if mailbox auditing on by default is turned on for the organization, the *AuditEnabled* property returned by the `Get-Mailbox` cmdlet always displays **True**. The mailbox audits mailbox activity by using the mailbox's configured mailbox actions.

To keep mailbox auditing disabled for specific mailboxes, configure *mailbox auditing bypass* for the mailbox owner and other users with delegated access to the mailbox. For more information, see Bypass mailbox audit logging.

Supported mailbox types

The following table describes mailbox types that mailbox auditing supports by default:

Expand table

Mailbox type	Auditing supported	On by default supported
Microsoft 365 Group mailboxes	✓	✓
Public folder mailboxes		
Resource mailboxes		
Shared mailboxes	✓	✓
User mailboxes	✓	✓

> **Important**
>
> In a multigeo environment, cross-geo mailbox auditing isn't supported. If you assign a user permissions to access a shared mailbox in a different geo location, mailbox actions that the user performs aren't logged in the mailbox audit log of the shared mailbox. Exchange admin audit events are available for all locations via Microsoft Purview Audit and the Search-UnifiedAuditLog cmdlet.

Sign-in types and mailbox actions

Sign-in types classify who's responsible for the audited actions on the mailbox. The following list describes the sign-in types that mailbox audit logging uses:

- **Owner**: The mailbox owner (the account associated with the mailbox).
- **Delegate**:
  - A user assigned the SendAs, SendOnBehalf, or FullAccess permission to another mailbox.
  - An admin assigned the FullAccess permission to a user's mailbox.
- **Admin**:
  - The mailbox is searched by using one of the following Microsoft eDiscovery tools:
    - eDiscovery in the Microsoft Purview portal.
    - In-Place eDiscovery in Exchange Online.
  - The mailbox is accessed by using the Microsoft Exchange Server MAPI Editor.
  - The mailbox is accessed by an account impersonating another user. This access happens when the *ApplicationImpersonation* role is assigned to an account, such as an application, which is now actively accessing the data. For more information, see Configure impersonation.

Mailbox actions for user mailboxes and shared mailboxes

The following table describes the mailbox actions that you can use in mailbox audit logging for user mailboxes and shared mailboxes.

- A check mark (✓) indicates the mailbox action you can log for the sign-in type (not all actions are available for all sign-in types).
- An asterisk (*) after the check mark indicates the mailbox action is logged by default for the sign-in type.
- An admin with Full Access permission to a mailbox is considered a delegate.

Expand table

Mailbox action	Description	Admin	Delegate	Owner
AddFolderPermissions	Although this value is accepted as a mailbox action, it's already included in the **UpdateFolderPermissions** action and isn't audited separately. In other words, don't use this value.			
ApplyRecord	An item is labeled as a record.	✓*	✓*	✓*
AttachmentAccess	A mail message attachment was accessed through the Microsoft Graph API. This operation is enabled by default and can't be configured with the other mailbox actions.	✓*	✓*	✓*
Copy	A message was copied to another folder.	✓		

Mailbox action	Description	Admin	Delegate	Owner
Create	An item was created in the Calendar, Contacts, Draft, Notes, or Tasks folder in the mailbox (for example, a new meeting request is created). Creating, sending, or receiving a message isn't audited. Also, creating a mailbox folder isn't audited.	✓*	✓*	✓
FolderBind	A mailbox folder was accessed. This action is also logged when the admin or delegate opens the mailbox. **Note**: Audit records for folder bind actions performed by delegates are consolidated. One audit record is generated for individual folder access within a 24-hour period.	✓	✓	
HardDelete	A message purged from the Recoverable Items folder.	✓*	✓*	✓*
MailboxLogin	The user signed into their mailbox.			✓
MailItemsAccessed	Occurs when mail data is accessed by mail protocols and clients.	✓*	✓*	✓*
MessageBind	**Note**: This value is available only for users *without* E5/A5/G5 licenses. A message was viewed in the preview pane or opened by an admin.	✓		
ModifyFolderPermissions	Although this value is accepted as a mailbox action, it's already included in the **UpdateFolderPermissions** action and isn't audited separately. In other words, don't use this value.			
Move	A message was moved to another folder.	✓	✓	✓
MoveToDeletedItems	A message deleted and moved to the Deleted Items folder.	✓*	✓*	✓*
RecordDelete	An item that's labeled as a record was soft-deleted (moved to the Recoverable Items folder). Items labeled as records can't be permanently deleted (purged from the Recoverable Items folder).	✓	✓	✓
RemoveFolderPermissions	Although this value is accepted as a mailbox action, it's already included in the			

Mailbox action	Description	Admin	Delegate	Owner
	UpdateFolderPermissions action and isn't audited separately. In other words, don't use this value.			
SearchQueryInitiated	A person uses Outlook (Windows, Mac, iOS, Android, or Outlook on the web) or the Mail app for Windows 10 to search for items in a mailbox.			✓
Send	The user sends an email message, replies to an email message, or forwards an email message.	✓*		✓*
SendAs	A message was sent using the SendAs permission. This permission allows another user to send the message as though it came from the mailbox owner.	✓*	✓*	
SendOnBehalf	A message was sent using the SendOnBehalf permission. This permission allows another user to send the message on behalf of the mailbox owner. The message indicates to the recipient who the message was sent on behalf of and who actually sent the message.	✓*	✓*	
SoftDelete	A message permanently deleted or deleted from the Deleted Items folder. Soft-deleted items are moved to the Recoverable Items folder.	✓*	✓*	✓*
Update	A message or any of its properties changed.	✓*	✓*	✓*
UpdateCalendarDelegation	A calendar delegation was assigned to a mailbox. Calendar delegation gives someone else in the same organization permissions to manage the mailbox owner's calendar.	✓*		✓*
UpdateFolderPermissions	A folder permission was changed. Folder permissions control which users in your organization can access folders in a mailbox and the messages located in those folders.	✓*	✓*	✓*
UpdateInboxRules	An inbox rule was added, removed, or changed. Inbox rules process messages in the user's Inbox based on conditions. Actions specify what to do to messages that match the conditions of the rule. For example, move the message to a specified folder or delete the message.	✓*	✓*	✓*

> **Important**
>
> If you customize the mailbox actions to audit *before* mailbox auditing is turned on by default in your organization, the customized mailbox auditing settings are preserved on the mailbox and aren't overwritten by the default mailbox actions described in this section. To revert the audit mailbox actions to their default values (which you can do at any time), see Restore the default mailbox actions.

Mailbox actions for Microsoft 365 Group mailboxes

When you turn on mailbox auditing, Microsoft 365 Group mailboxes start logging mailbox audit data. However, you can't customize the logged data or add or remove mailbox actions for any sign-in type.

The following table describes the mailbox actions that Microsoft 365 Group mailboxes log by default for each sign-in type.

An admin with Full Access permission to a Microsoft 365 Group mailbox is a delegate.

Expand table

Mailbox action	Description	Admin	Delegate	Owner
Create	Creation of a calendar item. Creating, sending, or receiving a message isn't audited.	✓*	✓*	
HardDelete	A message purged from the Recoverable Items folder.	✓*	✓*	✓*
MoveToDeletedItems	A message deleted and moved to the Deleted Items folder.	✓*	✓*	✓*
SendAs	A message sent by using the SendAs permission.	✓*	✓*	
SendOnBehalf	A message sent by using the SendOnBehalf permission.	✓*	✓*	
SoftDelete	A message permanently deleted or deleted from the Deleted Items folder. Soft-deleted items are moved to the Recoverable Items folder.	✓*	✓*	✓*
Update	A message or any of its properties changed.	✓*	✓*	✓*

Verify that default mailbox actions are logged for each sign-in type

The *DefaultAuditSet* property's value shows whether the default mailbox actions (managed by Microsoft) are audited on the mailbox.

To see which default audit action sets are currently applied to a user mailbox or shared mailbox, replace `<MailboxIdentity>` with the name, alias, email address, or user principal name (username) of the mailbox and run the following command in Exchange Online PowerShell:

PowerShell

```
Get-Mailbox -Identity <MailboxIdentity> | Format-List DefaultAuditSet
```

To view the default audit set for a Microsoft 365 Group mailbox, replace `<MailboxIdentity>` with the name, alias, or email address of the group mailbox and run the following command in Exchange Online PowerShell:

PowerShell

```
Get-Mailbox -Identity <MailboxIdentity> -GroupMailbox | Format-List DefaultAuditSet
```

The value `Admin, Delegate, Owner` means:

- The system audits the default mailbox actions for all three sign-in types. You see this value on Microsoft 365 Group mailboxes.
- An admin didn't change the audited mailbox actions for any sign-in type on a user mailbox or a shared mailbox.

If an admin changes the mailbox actions that the system audits for a sign-in type (by using the *AuditAdmin*, *AuditDelegate*, or *AuditOwner* parameters on the **Set-Mailbox** cmdlet), the property value is different.

For example, the value `Owner` for the *DefaultAuditSet* property on a user mailbox or shared mailbox means:

- The system audits the default mailbox actions for the mailbox owner.
- The audited mailbox actions for the `Delegate` and `Admin` sign-in types are changed from the default actions.

A blank value for the *DefaultAuditSet* property means the mailbox actions for all three sign-in types are changed on the user mailbox or a shared mailbox.

For more information, see Change or restore mailbox actions logged by default.

Display the mailbox actions that are logged on mailboxes

To see the mailbox actions that are currently logged on user mailboxes or shared mailboxes, replace `<MailboxIdentity>` with the name, alias, email address, or user principal name (username) of the mailbox. Then run one or more of the following commands in Exchange Online PowerShell.

> **Note**
>
> Although you can add the `-GroupMailbox` switch to the following **Get-Mailbox** commands for Microsoft 365 Group mailboxes, don't trust the values that are returned. The default and static mailbox actions that are audited for Microsoft 365 Group mailboxes are described in Mailbox actions for Microsoft 365 Group mailboxes.

Owner actions

To view the mailbox actions currently audited for the Owner sign-in type, run the following command:

PowerShell

```
Get-Mailbox -Identity <MailboxIdentity> | Select-Object -ExpandProperty AuditOwner
```

Delegate actions

To review the delegate actions currently being audited on the mailbox, run the following command:

PowerShell

```
Get-Mailbox -Identity <MailboxIdentity> | Select-Object -ExpandProperty AuditDelegate
```

Admin actions

To verify the admin actions currently being audited on the mailbox, run the following command:

PowerShell

```
Get-Mailbox -Identity <MailboxIdentity> | Select-Object -ExpandProperty AuditAdmin
```

Change or restore mailbox actions logged by default

One of the key benefits of using the default mailbox actions is that you don't need to manage the mailbox actions that are audited. Microsoft manages the actions for you and automatically adds new mailbox actions to audit by default as they're released.

However, your organization might be required to audit a different set of mailbox actions for user mailboxes and shared mailboxes. The procedures in this section show you how to change the mailbox actions that are audited for each sign-in type, and how to revert back to the Microsoft-managed default actions.

> **Important**
>
> If you use the following procedures to customize the mailbox actions that are logged on user mailboxes or shared mailboxes, new default mailbox actions released by Microsoft aren't automatically audited on those mailboxes. You need to manually add any new mailbox actions to your customized list.

Change the mailbox actions to audit

To change the mailbox actions that Microsoft audits for user mailboxes and shared mailboxes, use the *AuditAdmin*, *AuditDelegate*, or *AuditOwner* parameters on the **Set-Mailbox** cmdlet. You can't customize audited actions for Microsoft 365 group mailboxes.

Use two different methods to specify the mailbox actions:

- *Replace* (overwrite) the existing mailbox actions by using this syntax: `action1,action2,...actionN`.
- *Add or remove* mailbox actions without affecting other existing values by using this syntax: `@{Add="action1","action2",..."actionN"}` or `@{Remove="action1","action2",..."actionN"}`.

The following example changes the admin mailbox actions for the mailbox named "Gabriela Laureano" by overwriting the default actions with SoftDelete and HardDelete.

PowerShell

```
Set-Mailbox -Identity "Gabriela Laureano" -AuditAdmin HardDelete,SoftDelete
```

The following example adds the MailboxLogin action to the audited owner actions for the mailbox laura@contoso.onmicrosoft.com.

PowerShell

```
Set-Mailbox -Identity laura@contoso.onmicrosoft.com -AuditOwner @{Add="MailboxLogin"}
```

The following example removes the MoveToDeletedItems action from the audited delegate actions for the Team Discussion mailbox.

PowerShell

```
Set-Mailbox -Identity "Team Discussion" -AuditDelegate @{Remove="MoveToDeletedItems"}
```

Regardless of the method you use, customizing the audited mailbox actions on user mailboxes or shared mailboxes has the following results:

- Microsoft no longer manages the audited mailbox actions for the sign-in type that you customized.
- The sign-in type that you customized is removed from the *DefaultAuditSet* property value for the mailbox, indicating that Microsoft no longer manages the default actions for that sign-in type. For details on how to verify this value, see Verify that default mailbox actions are logged for each sign-in type.

Restore the default mailbox actions

> **Note**
>
> The following procedures don't apply to Microsoft 365 Group mailboxes. Those mailboxes are limited to the default actions described in Mailbox actions for Microsoft 365 Group mailboxes.

If you customize the mailbox actions that the system audits on a user mailbox or a shared mailbox, you can restore the default mailbox actions for one or all sign-in types by using this syntax:

PowerShell

```
Set-Mailbox -Identity <MailboxIdentity> -DefaultAuditSet <Admin | Delegate | Owner>
```

You can specify multiple *DefaultAuditSet* values separated by commas.

This example restores the default audited mailbox actions for all sign-in types on the mailbox **mark@contoso.onmicrosoft.com**.

PowerShell

```
Set-Mailbox -Identity mark@contoso.onmicrosoft.com -DefaultAuditSet
Admin,Delegate,Owner
```

This example restores the default audited mailbox actions for the Admin sign-in type on the mailbox **chris@contoso.onmicrosoft.com**, but keeps the customized audited mailbox actions for the Delegate and Owner sign-in types.

PowerShell

```
Set-Mailbox -Identity chris@contoso.onmicrosoft.com -DefaultAuditSet Admin
```

When you restore the default audited mailbox actions for a sign-in type, you get the following outcomes:

- The current list of mailbox actions is replaced with the default mailbox actions for the sign-in type.
- Any new mailbox actions that Microsoft releases are automatically added to the list of audited actions for the sign-in type.
- The *DefaultAuditSet* property value for the mailbox is updated to include the restored sign-in type.

Turn off mailbox auditing on by default for your organization

To turn off mailbox auditing on by default for your entire organization, run the following command in Exchange Online PowerShell:

PowerShell

```
Set-OrganizationConfig -AuditDisabled $true
```

When you turn off mailbox auditing on by default, you make the following changes:

- Mailbox auditing is turned off for your organization.
- From the time you turn off mailbox auditing on by default, no mailbox actions are audited, even if mailbox auditing is enabled on a mailbox (the *AuditEnabled* property on the mailbox is **True**).
- Mailbox auditing isn't turned on for new mailboxes and setting the *AuditEnabled* property on a new or existing mailbox to **True** is ignored.
- Any mailbox audit bypass association settings (configured by using the **Set-MailboxAuditBypassAssociation** cmdlet) are ignored.
- Existing mailbox audit records are retained until the audit log age limit for the record expires.

Turn on mailbox auditing on by default

To turn mailbox auditing back on for your organization, run the following command in Exchange Online PowerShell:

PowerShell

```
Set-OrganizationConfig -AuditDisabled $false
```

Bypass mailbox audit logging

Currently, you can't disable mailbox auditing for specific mailboxes when mailbox auditing on by default is turned on in your organization. For example, the system ignores setting the *AuditEnabled* mailbox property to **False**.

However, you can use the **Set-MailboxAuditBypassAssociation** cmdlet in Exchange Online PowerShell to prevent *all* mailbox actions by the specified users from being logged, regardless of where the actions occur. For example:

- Mailbox owner actions that the bypassed users perform aren't logged.
- Delegate actions that the bypassed users perform on other users' mailboxes (including shared mailboxes) aren't logged.
- Admin actions that the bypassed users perform aren't logged.

To bypass mailbox audit logging for a specific user, replace `<MailboxIdentity>` with the name, email address, alias, or user principal name (username) of the user and run the following command:

PowerShell

```
Set-MailboxAuditBypassAssociation -Identity <MailboxIdentity> -AuditByPassEnabled $true
```

To verify that auditing is bypassed for the specified user, run the following command:

PowerShell

```
Get-MailboxAuditBypassAssociation -Identity <MailboxIdentity> | Format-List AuditByPassEnabled
```

The value **True** indicates that mailbox audit logging is bypassed for the user.

> **Note**
>
> The `AuditLogAgeLimit` property is no longer applicable for managing mailbox audit log retention. Audit log retention is now configured and managed through Microsoft Purview. For more information, see Manage audit log retention policies.

Next steps

- Search the audit log for mailbox activities in specific mailboxes: Search for audited mailbox actions using the filters you configured.
- Use MailItemsAccessed to investigate compromised accounts: Use the MailItemsAccessed action to investigate potential email breaches.
- Investigate shared mailbox activities using audit logs: Monitor activity in shared mailboxes for unauthorized access or changes.

Note: The author created this article with assistance from AI. Learn more

Last updated on 06/19/2026

Manage audit log retention policies

You can create and manage audit log retention policies in the Microsoft Purview portal. Audit log retention policies are part of the new Microsoft Purview Audit (Premium) capabilities. An audit log retention policy lets you specify how long to retain audit logs in your organization. You can retain audit logs for up to 10 years. You can create policies based on the following criteria:

- All activities in one or more Microsoft services
- Specific activities in a Microsoft service performed by all users or by specific users
- A priority level that specifies which policy takes precedence if you have multiple policies in your organization

Default audit log retention policy in Audit (Premium)

Audit (Premium) in Microsoft Purview provides a default audit log retention policy for all organizations. You can't modify this policy. It retains all Exchange Online, SharePoint, OneDrive, and Microsoft Entra audit records for one year. This default policy retains audit records that contain the value of **AzureActiveDirectory**, **Exchange**, **OneDrive**, and **SharePoint** for the **Workload** property (which is the service in which the activity occurred). Audit records for all other activities are retained for 180 days by default or you can change the retention to a different duration using a custom retention policy.

> **Note**
>
> The default audit log retention policy only applies to audit records for activity performed by users who are assigned an Office 365 or Microsoft 365 E5 license or have a Microsoft Purview Suite (formerly known as Microsoft 365 E5 Compliance) or E5 eDiscovery and Audit add-on license. If you have non-E5 users or guest users in your organization, their corresponding audit records are retained for 180 days.

The default retention period for audit records varies by license type and when the records were generated.

> **Important**

The default retention period for Audit (Standard) changed from 90 days to 180 days. Audit (Standard) logs generated before October 17, 2023 are retained for 90 days. Audit (Standard) logs generated on or after October 17, 2023 follow the new default retention of 180 days.

Before you create an audit log retention policy

- You need the *Organization Configuration* role in the Microsoft Purview portal to create or modify an audit retention policy.
- Your organization can have up to 50 audit log retention policies.
- To retain an audit log for longer than 180 days (and up to 1 year), the user who generates the audit log (by performing an audited activity) must have an Office 365 E5 or Microsoft 365 E5 license or a Microsoft Purview Suite (formerly known as Microsoft 365 E5 Compliance) or E5 eDiscovery and Audit add-on license. To retain audit logs for 10 years, the user who generates the audit log must also have a 10-year audit log retention add-on license in addition to an E5 license.

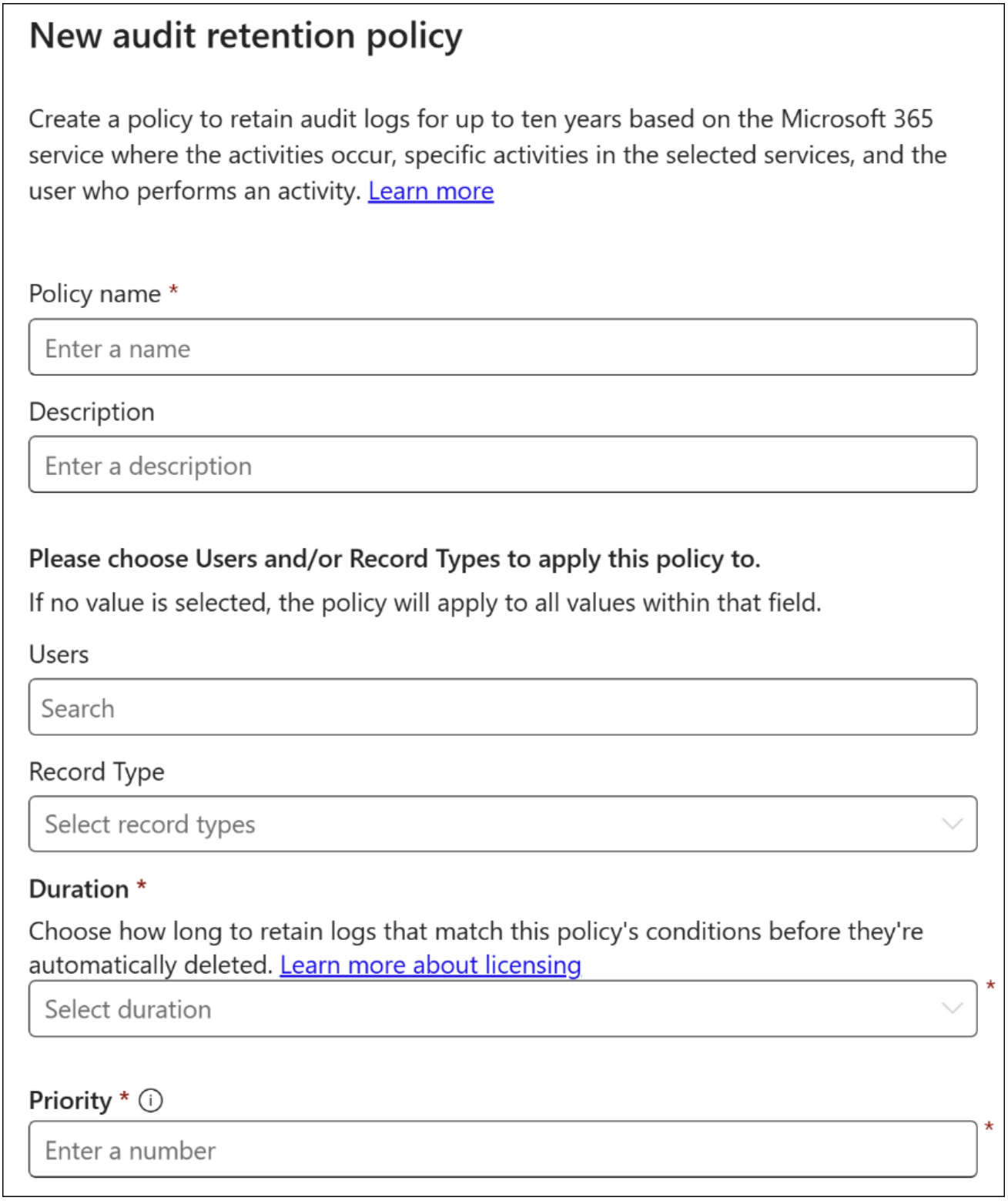
  > **Note**
  >
  > If the user generating the audit log doesn't have the licenses required for the selected retention duration, data is retained according to the highest priority retention policy. This retention might be either the default retention policy for the user's license or the highest priority policy that matches the user and its record type.

- All custom audit log retention policies (created by your organization) take priority over the default retention policy. For example, if you create an audit log retention policy for Exchange mailbox activity that has a retention period that's shorter than one year, audit records for Exchange mailbox activities are retained for the shorter duration specified by the custom policy.
- The audit item lifetime for data is determined when you add it to the auditing pipeline and is based on the licensing defaults or applicable retention policies. Any changes to licensing or applicable retention policies change the expiration time of the audit data after updating. These changes don't update any previously committed items.

Create an audit log retention policy

Complete the following steps to create an audit retention policy:

1. Sign in to the Microsoft Purview portal with a user account assigned the *Organization Configuration* role on the **Roles & scopes** page in the Microsoft Purview portal.
2. Select the **Audit** solution card. If the **Audit** solution card isn't displayed, select **View all solutions** and then select **Audit** from the **Core** section.
3. Select **Create audit retention policy**, and then complete the following fields on the flyout page:



- **Policy name**: The name of the audit log retention policy. This name must be unique in your organization, and you can't change it after creating the policy.
- **Description**: Optional, but helpful to provide information about the policy, such as the record type or workload, users specified in the policy, and the duration.
- **Users:** Select one or more users to apply the policy to. If you leave this box blank, the policy applies to all users.
- **Record type**: The audit record type the policy applies to. If you leave this property blank, the policy applies to all record types. You can select a single record type or multiple record types:
- If you select a single record type, the **Activities** field is dynamically displayed. Use the drop-down list to select activities from the selected record type to apply the policy to. If you don't select specific activities, the policy applies to all activities of the selected record type.
  - If you select multiple record types, you don't have the ability to select activities. The policy applies to all activities of the selected record types.
  - **Duration:** The amount of time to retain the audit logs that meet the criteria of the policy. The available options are *7 Days*, *30 Days*, *6 Months*, *9 Months*, *1 Year*, *3 Years*, *5 Years*, and *7 Years*. Users with the 10-year Audit Log Retention add-on license can select a *10 Years* option.

> **ⓘ Important**
>
> To retain audit logs for the 7 and 30 days duration options, you must have a Microsoft 365 Enterprise E5 subscription. To retain audit logs for the 3, 5, and 7 years duration options, you must be assigned to a 10-Year Audit Log Retention add-on license in addition to your Microsoft 365 Enterprise E5 subscription. For more information about Audit subscriptions and add-ons, see Auditing solutions in Microsoft Purview

- **Priority**: This value determines the order in which audit log retention policies in your organization are processed. A lower value indicates a higher priority. Valid priorities are numerical values between **1** and **10000**. A value of **1** is the highest priority, and a value of **10000** is the lowest priority. For example, a policy with a value of **5** takes priority over a policy with a value of **10**. Any custom audit log retention policy takes priority over the default policy for your organization.

4. Select **Save** to create the new audit log retention policy.

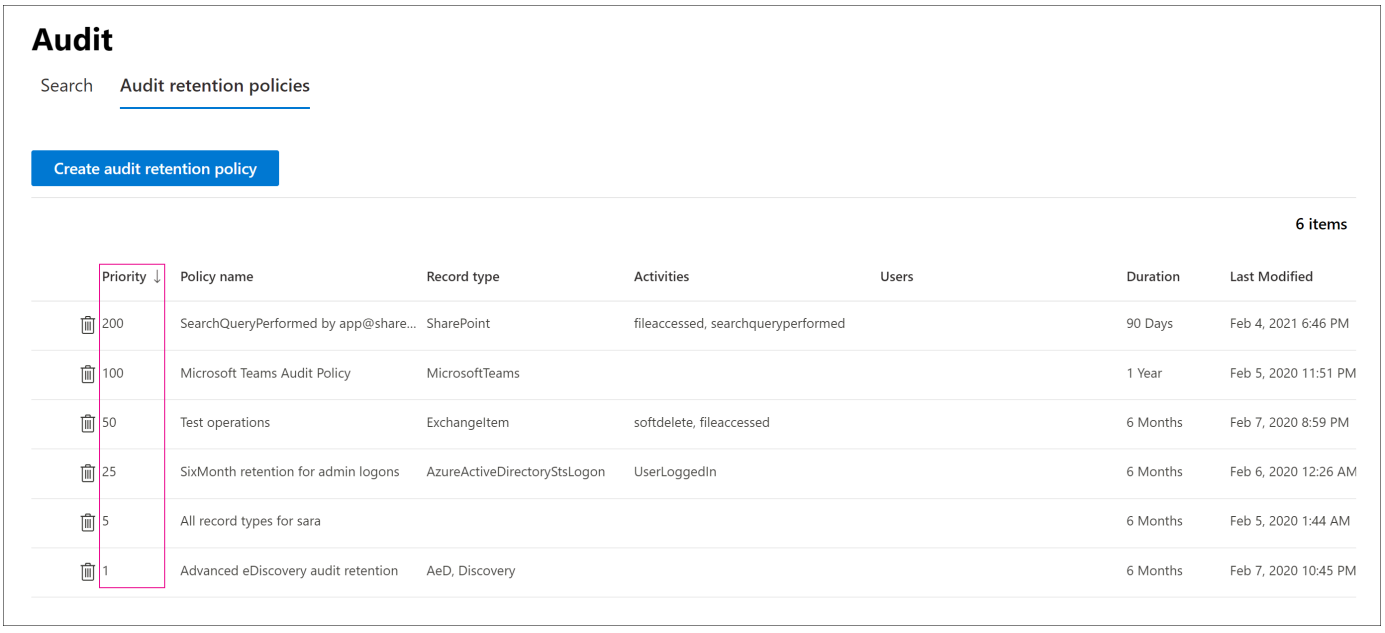
The new policy appears in the list on the **Policies** page.

Manage audit log retention policies in the Microsoft Purview portal

The **Audit retention policies** tab (also called the *dashboard*) lists audit log retention policies. You can use the dashboard to view, edit, and delete audit retention policies.

View policies in the dashboard

The dashboard lists audit log retention policies. One advantage of viewing policies in the dashboard is that you can select the **Priority** column to list the policies in the priority order in which they're applied. A lower value indicates a higher priority.



You can also select a policy to display its settings on the flyout page.

> **Note**
>
> The dashboard doesn't display the default audit log retention policy for your organization.

Edit policies in the dashboard

To edit a policy, select it to display the flyout page. You can modify one or more settings and then save your changes.

> **Important**
>
> If you use the *New-UnifiedAuditLogRetentionPolicy* cmdlet, you might create an audit log retention policy for record types or activities that aren't available in the **Create audit retention policy** tool in the dashboard. If the policy includes record types or activities that aren't available in the **Create audit retention policy** tool, you can't edit the policy (for example, change the retention duration or add and remove activities) from the **Audit retention policies** dashboard. You can only view and delete the policy in the Microsoft Purview portal. To edit the policy, you need to use the Set-UnifiedAuditLogRetentionPolicy cmdlet in Security & Compliance PowerShell. A message is displayed at the top of the flyout page for policies that you need to edit by using PowerShell.

Delete policies in the dashboard

To delete a policy, select the **Delete** icon and then confirm that you want to delete the policy. The policy is removed from the dashboard, but it might take up to 30 minutes for the policy to be removed from your organization.

Create and manage audit log retention policies in PowerShell

You can also use Security & Compliance PowerShell to create and manage audit log retention policies. One reason to use PowerShell is to create a policy for a record type or activity that isn't available in the UI.

Create an audit log retention policy in PowerShell

Complete the following steps to create an audit log retention policy in PowerShell:

1. Connect to Security & Compliance PowerShell.

2. Run the following command to create an audit log retention policy:

```
PowerShell
New-UnifiedAuditLogRetentionPolicy -Name "Microsoft Teams Audit Policy" -
Description "One year retention policy for all Microsoft Teams activities" -
RecordTypes MicrosoftTeams -RetentionDuration TenYears -Priority 100
```

This example creates an audit log retention policy named "Microsoft Teams Audit Policy" with these settings:

- A description of the policy.
- Retains all Microsoft Teams activities (as defined by the *RecordType* parameter).
- Retains Microsoft Teams audit logs for 10 years.
- A priority of 100.

The following example creates an audit log retention policy that retains audit logs for the "User logged in" activity for six months for the user admin@contoso.onmicrosoft.com.

PowerShell

```
New-UnifiedAuditLogRetentionPolicy -Name "SixMonth retention for admin logons" -
RecordTypes AzureActiveDirectoryStsLogon -Operations UserLoggedIn -UserIds
admin@contoso.onmicrosoft.com -RetentionDuration SixMonths -Priority 25
```

For more information, see New-UnifiedAuditLogRetentionPolicy.

View policies in PowerShell

Use the Get-UnifiedAuditLogRetentionPolicy cmdlet in Security & Compliance PowerShell to view audit log retention policies.

The following command displays the settings for all audit log retention policies in your organization. This command sorts the policies from the highest to lowest priority.

PowerShell

```
Get-UnifiedAuditLogRetentionPolicy | Sort-Object -Property Priority -Descending | FL
Priority,Name,Description,RecordTypes,Operations,UserIds,RetentionDuration
```

> ⓘ **Note**
>
> The *Get-UnifiedAuditLogRetentionPolicy* cmdlet doesn't return the default audit log retention policy for your organization.

Edit policies in PowerShell

Use the Set-UnifiedAuditLogRetentionPolicy cmdlet in Security & Compliance PowerShell to edit an existing audit log retention policy.

Delete policies in PowerShell

Use the Remove-UnifiedAuditLogRetentionPolicy cmdlet in Security & Compliance PowerShell to delete an audit log retention policy. It might take up to 30 minutes for the policy to be removed from your organization.

Next steps

- Search the audit log: Search for activities that are retained by your custom retention policies.
- Use a PowerShell script to search the audit log: Automate searches against your retained audit data for ongoing monitoring.
- Microsoft services that support auditing: Review which services generate the audit records your retention policies cover.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/19/2026

Audit logs for Copilot and AI applications

This article provides an overview of audit logs generated for user interactions and admin activities related to Microsoft Copilot and AI applications. The system automatically logs these activities as part of Audit (Standard). If your organization enables auditing, you don't need to take extra steps to configure auditing support for Copilot and AI applications.

Billing for auditing non-Microsoft AI applications

Audit logs for non-Microsoft AI applications use pay-as-you-go billing. This billing model provides user and admin interaction audit logs that the system retains for 180 days. These audit logs cover interactions with non-Microsoft AI applications.

Your enterprise subscription doesn't include audit logs for this type of user interaction. Instead, it falls under pay-as-you-go billing. The system logs these interactions under the *AIAppInteraction* recordType or *AIApp* workload. Some scenarios logged under the *ConnectedAiAppInteraction* recordType are also part of this pay-as-you-go billing model. You need to enable pay-as-you-go features to turn on these logs. When you enable them, the system retains these audit logs for 180 days. Consumption is charged based on the number of audit records ingested for user interactions with these non-Microsoft AI applications.

Pay-as-you-go billing doesn't apply to Microsoft applications. All Microsoft applications, including Microsoft Copilots like Microsoft Security Copilot, Copilot in Microsoft Fabric, and custom applications built using Microsoft Copilot Studio and Microsoft Foundry are included in Audit Standard.

Admin activities with Copilot and AI applications

The system generates audit logs when an administrator performs activities related to Copilot settings, plugins, promptbooks, or workspaces. For more information, see Microsoft 365 Copilot activities.

User activities with Copilot, Cowork, and AI applications

The system automatically generates audit logs when a user interacts with Copilot, Cowork, or an AI application. These audit records contain details about which user interacted with Copilot, when the interaction took place, and where it occurred. Audit records also include references to files, sites, or other resources Copilot, Cowork, and AI applications accessed to generate responses to user prompts.

Common properties in Copilot audit logs

The following table outlines some of the common properties included in audit logs.

Expand table

Attribute	Definition	Examples
AccessedResources	References to all resources (files, documents, emails, etc.) which Copilot accessed in response to the	For example: "AccessedResources": [{"Action":"Read","ID":"AAAAAEYE2GAACp1FlnN_CHXStUkHAGWJYgtgc 8706-4188-bd11-

Attribute	Definition	Examples
	user's request. - *ID* is the unique identifier for the resource. This could be a fileId on OneDrive, or a messageId in Teams, or email ID in Outlook, etc. - *SiteUrl* is the URL of the resource that was accessed. This could be the URL of a SharePoint site, full file path of a file, etc. - *ListItemUniqueId* is a unique identifier for an item in SharePoint. - *Type* refers to the type of resource that was accessed. It can contain values like the filetype extension (pptx, docx, etc.) or describe the type of resource (for non-SharePoint resources). - *Name* is the user-friendly readable name of the resource (for example, fileName). - *SensitivityLabelId* is the ID of the sensitivity label assigned to the resource. This is helpful in identifying whether Copilot accessed any sensitive information while generating its response. - *Action* refers to the nature of access which Copilot performed on the resource. Common values include *read*, *create*, *modify*. - *PolicyDetails* is used in scenarios where Copilot's access to a particular resource was blocked or restricted based on some policy. This property can include details like *PolicyId*, *PolicyName*, list of rules, etc. - *Status* is used to specify whether Copilot's action on a specific resource was a *success* or *failure*. - *XPIADetected* is a boolean that denotes whether there was an XPIA (Cross Prompt Injection Attack) detected from a particular resource which Copilot accessed.	ebb85995028c","SiteUrl":"`https://microsoft.sharepoint.com/teams/Off` `web=1`","Type":"docx","listItemUniqueId":"AAAAAEYE2GAACp1FlnN_CH)
AgentId	Unique identifier for an agent. The string can also include details about the category of agent involved in the interaction.	For example, when a user interacts with a Declarative Agent or a Custo *b424-4d54-8ddb-15dc19247088* or *CopilotStudio.CustomEngine.11fd28*
AgentName	A friendly readable name of the agent.	*JiraStatusAgent*, *SalesAgent*, *ReminderBot*, and others.
AgentVersion	The version number or version ID of the agent involved.	Values like *25.001*, *8076fbed-be52-4004-ac89-81181ecd7b33*, and othe
AISystemPlugin	Details of plugins or extensions enabled for the Copilot interaction. - *Name* is the name of the plugin that Copilot uses in generating the response. - *ID* is the unique identifier for the plugin. - *Version* refers to the version of plugin used.	
AppHost	The same Copilot application can be deployed within multiple host applications. This property helps identify the application that hosted	Some of the common AppHost scenarios are: - *BizChat*: The Copilot interaction was performed in the Microsoft 365 C - *Bing*: The Copilot interaction was performed through the Microsoft Ec

Attribute	Definition	Examples
	the interaction between a user and Copilot.	- *Office*: The Copilot interaction was performed through office.com or r - Other application-specific values: Values like *Word*, *Excel*, *PowerPoint*,
AppIdentity	A detailed string that you use to uniquely identify the specific Copilot or AI Application that the user interacted with. It typically follows the structure *workloadName.appGroup.appName*.	For example, interactions with first-party Copilot apps developed by M and others. Interactions with custom-built Copilots created through Copilot Studio Interactions with third-party AI apps deployed within your organizatior *ConnectedAIApp.AzureAI.AzureResourceName*. Interactions with third-p *AIApp.SaaS.AppName*.
CapacityId	Unique identifier for the Microsoft Fabric Capacity in a tenant.	You can find the *CapacityID* in the URL of the capacity management pa *CapacityID* is shown in the URL after */capacities/*. For example, `0000111 aaaa-2222-bbbb-3333cccc4444`.
ClientRegion	The user's region when they performed the operation.	
Contexts	Contains a collection of attributes to help describe where the user was during the Copilot interaction. - *ID* is the identifier of the resource that was being used during the Copilot interaction. - *Type* is the filetype/name of the app or service where the interaction occurred.	- *ID* contains values like *FileId* or *FilePath* (for SharePoint scenarios), or - *Type* contains values like docx, pptx, xlsx, TeamsMeeting, TeamsChanr
DLPEvaluationDeferred	An integer bitmask that indicates DLP evaluation of one or more content processing stages couldn't be completed, so it's deferred for later reevaluation.	The value is a bitmask where each bit represents a specific evaluation s - *1 – Prompt*: DLP evaluation of the user prompt was deferred. - *2 – Response*: DLP evaluation of the generated response was deferred - *4 – Grounding*: DLP evaluation of grounding content was deferred. - *8 – WebGrounding*: DLP evaluation of web-grounding content was de
DLPEvaluationDeferredReason	Provides the reason why one or more DLP evaluation stages identified by *DLPEvaluationDeferred* could not be completed during the original request processing and were deferred for later reevaluation.	This field is populated only when *DLPEvaluationDeferred* contains a no
Messages	Contains details about the prompt and response messages within the Copilot interaction. A single audit record typically contains a prompt-response pair but can also include a prompt with multiple response messages (that is, all Copilot responses associated with that prompt). - *ID* is the messageId of the prompt/response message in the Copilot interaction. - *IsPrompt* is a boolean flag to denote whether this message is a user prompt or Copilot response. - *JailbreakDetected* is a boolean flag to denote whether a jailbreak attempt was made using this prompt message. - *Size* is currently not used.	"Messages": [{"ID":"1715186983849", "isPrompt":true}, {"ID":"17151869
ModelTransparencyDetails	Details of the AI model used in the interaction. - *ModelProviderName* is the publisher of the model used. All Microsoft 365 Copilot scenarios include a ModelProviderName. - *ModelName* is the name of the model used. Not available in	

Attribute	Definition	Examples
	Microsoft 365 Copilot scenarios. - *ModelVersion* is the version of the model used. Not available in Microsoft 365 Copilot scenarios.	
Operation	Specifies the name of the activity that was audited.	For user interactions with Copilot or Cowork, this property contains *Co* as described in the RecordType row. Also includes Copilot admin operations like *UpdateTenantSettings*, *Crec*
RecordType	Identifies the category of Copilot or AI application that the user interacted with.	*CopilotInteraction* refers to scenarios where a user interacted with a Mi *ConnectedAIAppInteraction* refers to scenarios where a user interacted *AIAppInteraction* refers to interactions with third-party AI applications t
Workload	Identifies the app category, similar to RecordType.	*Copilot*, *ConnectedAIApp*, *AIApp*

Common AppHost scenarios in Copilot

The following table lists some of the commonly used values for AppHost and describes the scenarios in which they're used.

Expand table

AppHost	Copilot Scenario	Copilot product
Bing	**Business Chat via Bing/Windows interface.** Refers to Microsoft 365 Copilot's cross-app **Business Chat** access through the Bing Chat experience (for instance, in the Microsoft Edge browser sidebar, Windows Copilot, or the copilot.cloud.microsoft.com web portal). This scenario occurs when a user engages the Copilot in a general-purpose chat outside any specific Office app.	Microsoft 365 Copilot Chat
Bookings	Copilot in Microsoft Bookings.	Microsoft 365 Copilot
Copilot in Azure	**Copilot in Microsoft Azure (Cloud Management).** This refers to Copilot usage in the context of **Microsoft Azure**. Though labeled under Microsoft 365 Copilot, the scenario is a cloud admin or developer using Copilot to manage Azure resources. For example, in the Azure portal, ad admin might ask, "How do I set up an alert for CPU usage on my VMs?" and Copilot would generate an ARM template or Azure CLI steps. Or "list untagged resources and suggest a tagging scheme". Essentially, Copilot in Azure is an AI cloud assistant to query and control Azure services via natural language.	Microsoft 365 Copilot
Copilot in Defender	**Security Copilot in Microsoft Defender.** The user is using Copilot within the **Microsoft Defender** security portal. This assists security operations (SecOps) tasks. For example, an analyst could ask, "Investigate alert ID 23455 and summarize what happened," and Copilot analyzes Defender alerts and incidents to produce an explanation or even recommend next steps. It can also cross-reference threat intelligence. This scenario specifically focuses on **Defender for Endpoint/Office/Cloud, etc.** data via the Security Copilot interface.	Security Copilot
Copilot in Intune	**Security Copilot in Microsoft Intune (Endpoint Management).** Copilot is used in the **Microsoft Intune** admin center. It helps IT admins with device management and security posture. For instance, an admin might ask, "How many devices are noncompliant this week and why?". Copilot retrieves Intune data on device compliance and summarize reasons. It can also assist with troubleshooting by comparing device configuration or retrieving app deployment status. This is essentially Security Copilot tapping into Intune's information to answer questions and provide insights for IT and security teams.	Security Copilot

AppHost	Copilot Scenario	Copilot product
Designer	Copilot in Microsoft Designer.	Microsoft 365 Copilot
Edge	**Business Chat in Edge sidebar.** Represents Microsoft 365 Copilot chat accessed through the **Edge browser's Copilot (Bing Chat Enterprise) sidebar**. In this scenario, the user is likely using the Microsoft Edge sidebar Copilot to query organizational data (a BizChat experience within Microsoft Edge).	Microsoft 365 Copilot Chat
Excel	**Copilot in Microsoft Excel.** The user is using Copilot inside an Excel spreadsheet. For example, Copilot might be asked to analyze data, create formulas, or generate a summary of a table. It's the Excel-integrated Copilot helping with computations or insights in workbooks.	Microsoft 365 Copilot
Forms	**Copilot in Microsoft Forms.** Represents Copilot being used in the context of Forms. For example, Copilot could help create survey questions, quizzes, or analyze form responses. This would be the AI assistant helping content creation or summarization in Forms.	Microsoft 365 Copilot
Logic App	Copilot in Azure Logic Apps.	Microsoft 365 Copilot
Loop	**Copilot in Microsoft Loop.** This refers to Copilot assisting within the Microsoft Loop application or Loop components. The user might have Copilot generate or summarize content in a Loop workspace. For example, Copilot could help brainstorm in a Loop page, given Loop's collaborative canvas. This integration brings Copilot to the Loop app context.	Microsoft 365 Copilot
M365AdminCenter	**Copilot in Microsoft 365 Admin Center.** This refers to an AI assistant for IT administrators. In this scenario, Copilot could help an admin with tasks in the Microsoft 365 admin center. For example, answering questions about settings, generating PowerShell scripts, or summarizing user reports.	Microsoft 365 Copilot
M365App	**Business Chat via Microsoft 365 app (desktop or mobile).** Similar to the Office apphost, this denotes Business Chat launched from the **Microsoft 365 unified app** on Windows or mobile. It covers the scenario where a user uses the Office/Microsoft 365 app itself (outside a specific product like Word) to chat with Copilot across their data.	Microsoft 365 Copilot Chat
Microsoft Purview	**Security Copilot in Purview (Compliance).** The user is interacting with Copilot in the **Microsoft Purview** compliance portal. Copilot helps compliance officers and security teams triage and summarize issues related to data protection and governance. For example, Copilot can summarize a batch of Data Loss Prevention (DLP) alerts, highlight inside risk activities, or answer questions like "Have we had any policy violations in email this week?". It can work in both an embedded way (inside Microsoft Purview UI, summarizing whatever section you're on) and in a standalone Q&A way for Microsoft Purview data. This scenario brings AI to compliance workflows, making it faster to grasp risks and decide on actions.	Security Copilot
Office	**Business Chat via Office.com or Microsoft 365 app.** Indicates the Copilot Business Chat accessed through the Office.com or Microsoft 365 home app (web, desktop, or mobile). For example, when a user opens the "Copilot" chat on Office.com (microsoft365.com) or the Microsoft 365 mobile app to ask cross-domain questions.	Microsoft 365 Copilot Chat
OfficeCopilotNotebook	**Copilot Notebook (Microsoft 365).** Refers to the **Copilot Notebooks** feature in Microsoft 365 Copilot (a central AI-powered notebook). This scenario is when Copilot compiles or interacts with a **cross-application notebook** of content. Copilot Notebooks allow users to gather and generate information across multiple sources in a notebook interface. This AppHost appears when that notebook is used on the Office.com/Microsoft 365 side (outside of OneNote).	Microsoft 365 Copilot Chat

AppHost	Copilot Scenario	Copilot product
OfficeCopilotSearchAnswer	**Copilot answer in Microsoft 365 Search.** This value refers to Copilot generating an **AI-powered answer in the Office/Microsoft 365 search experience**. For instance, when a user searches on Office.com or SharePoint and Copilot provides a natural language answer (drawing from workplace data) instead of just search results. This is a Business Chat-like Q&A feature within the search context.	Microsoft 365 Copilot Chat
OneDrive	Also refers to the **Copilot in SharePoint** scenario, as described previously.	Microsoft 365 Copilot
OneNote	**Copilot in Microsoft OneNote.** The user is using Copilot inside OneNote notebooks. This in-app OneNote Copilot can summarize notes, generate plans or lists, or answer questions based on NoteNote content. Copilot "*supercharges your note-taking*" in OneNote, helping to create, recall, and organize information.	Microsoft 365 Copilot
OneNoteCopilotNotebook	**Copilot Notebook in OneNote.** Similar to the OfficeCopilotNotebook scenario, but specifically when the Copilot Notebook is accessed within OneNote. Microsoft introduced Copilot Notebooks integrated into OneNote, so this AppHost logs when a user uses the **AI-powered notebook inside OneNote** (bringing cross-data Copilot functionality into OneNote).	Microsoft 365 Copilot Chat
Outlook	**Copilot in Microsoft Outlook.** The user in engaging Copilot while using Outlook (desktop, web, or mobile). This typically means the Copilot is helping with email tasks. For example, drafting an email reply, summarizing a long email thread, or organizing an inbox.	Microsoft 365 Copilot Chat
OutlookOnCanvas	**Copilot inline in Outlook compose.** This refers to Copilot's assistance directly in the email canvas. For example, when composing an email, Copilot might autogenerate text right in the draft. It's the on-canvas helper in Outlook's compose window (as opposed to using a separate Copilot pane).	Microsoft 365 Copilot
OutlookSidepane	**Copilot in Outlook side pane.** Denotes the classic Outlook Copilot experience via the Copilot pane in Outlook. For example, a user opens the Copilot sidebar in Outlook to draft or summarize messages. This value explicitly captures that side pane interaction.	Microsoft 365 Copilot
Planner	**Copilot in Microsoft Planner.** Indicates a Copilot scenario in Planner, likely assisting with project plans or tasks. A user might ask Copilot to draft a plan, generate task checklists, or update task descriptions.	Microsoft 365 Copilot
Power BI	Copilot in Microsoft Power BI.	Microsoft 365 Copilot
PowerPoint	**Copilot in Microsoft PowerPoint.** Copilot is being used within a PowerPoint presentation. In this scenario, a user could ask Copilot to create slides, generate speaker notes, or redesign content in PowerPoint.	Microsoft 365 Copilot
PowerPointOnCanvas	**Copilot on PowerPoint slides.** Indicates an on-canvas Copilot experience in PowerPoint, where Copilot inserts or modifies content directly on slides. This could be the scenario of Copilot generating layouts or bulleted points straight into the presentation (without solely relying on the chat pane). It's a more embedded form of the PowerPoint Copilot.	Microsoft 365 Copilot
Security Copilot Standalone	**Security Copilot standalone experience.** The user is interacting with Security Copilot through the standalone experience, as opposed to interacting with Copilot through an embedded experience within Defender, Purview, etc.	Security Copilot
SharePoint	**Copilot in SharePoint.** Copilot is used within SharePoint (likely on a SharePoint site or page). For instance, a user might ask Copilot to summarize a SharePoint news post or draft content for a	Microsoft 365 Copilot

AppHost	Copilot Scenario	Copilot product
	SharePoint page. This scenario corresponds to a SharePoint-integrated Copilot helping with intranet content.	
Stream	Copilot in Microsoft Stream.	Microsoft 365 Copilot
Teams	**Copilot in Microsoft Teams.** Represents Copilot usage inside the Microsoft Teams app (Web/Desktop/Mobile). This covers interactions with Copilot in Teams chats or channels. For example, asking Copilot to summarize a Teams chat, answer a question in a channel, or assist in a meeting context. This is essentially the Copilot experience within Teams' interface. For example, the "Chat Copilot" in a Teams chat thread.	Microsoft 365 Copilot
TeamsAdminPortal	**Copilot in Microsoft Teams Admin Center.** Similar to the M365AdminCenter scenario, this indicated a Copilot scenario in the Teams Admin Portal. An admin might use Copilot to configure Teams settings or generate reports.	Microsoft 365 Copilot
VivaEngage	**Copilot in Viva Engage (Viva Engage).** The user is interaction with Copilot within Viva Engage. For example, drafting a post or summarizing conversation threads in a Viva Engage community. This scenario covers any AI assistance inside Viva Engage, such as helping craft announcements or answers.	Microsoft 365 Copilot
VivaGoals	**Copilot in Viva Goals.** Viva Goals manages OKRs (objectives and key results). A Copilot here could draft OKRs, update progress, or analyze goal attainment.	Microsoft 365 Copilot
VivaPulse	**Copilot in Viva Pulse.** Viva Pulse is a feedback survey tool; Copilot here might draft survey questions or summarize sentiment from responses.	Microsoft 365 Copilot
Whiteboard	**Copilot in Microsoft Whiteboard.** The user is using Copilot on a digital whiteboard. Copilot in Whiteboard helps brainstorm and organize ideas on the whiteboard canvas. For example, suggesting ideas, clustering sticky notes, or summarizing the board's content. This scenario covers using Copilot during a Whiteboard session (in Teams or the Whiteboard app) to enhance creativity and structure.	Microsoft 365 Copilot
Word	**Copilot in Microsoft Word.** The user is interacting with Copilot within Word. For example, asking it to draft or edit portions of a Word document. This is the in-app **Word Copilot** (side-pane chat and commands in Word). Copilot can generate content, summarize text, or adjust formatting the document context.	Microsoft 365 Copilot
WordOnCanvas	**Copilot inline in Word's document.** This refers to Copilot assistance directly *on the canvas* in Word. Instead of the side pane, Copilot acts within the document editing area. For example, the feature where Copilot writes directly into the document or provides inline suggestions. It's essentially Word Copilot's capabilities applies within the document body, rather than via the chat pane.	Microsoft 365 Copilot

Example Copilot scenarios for user activities

The following tables list some example scenarios and how they appear in the audit log. These example audit logs come from Copilot activities.

Microsoft Copilot

A user interacts with Microsoft Copilot through the Microsoft 365 Copilot Chat client.

Expand table

Operation	RecordType	AppIdentity	AppHost
CopilotInteraction	CopilotInteraction	Copilot.MicrosoftCopilot.BizChat	BizChat

Security Copilot

A user interacts with Security Copilot within Microsoft Defender.

Expand table

Operation	RecordType	AppIdentity	AppHost
CopilotInteraction	CopilotInteraction	Copilot.Security.SecurityCopilot	Defender

Copilot Studio applications

A user interacts with a custom-built Copilot Studio application (whose appId is the GUID contained in appIdentity). The interaction takes place within Microsoft Teams, where this custom-built application is deployed.

Expand table

Operation	RecordType	AppIdentity	AppHost
CopilotInteraction	CopilotInteraction	Copilot.Studio.f4d97b45-1deb-40ce-9004-b473b79eab85	Teams

Microsoft Facilitator

Microsoft Facilitator updates AI Notes, Live Notes, or Meeting Moderation in Microsoft Teams.

Expand table

Operation	RecordType	AppIdentity	AppHost
AINotesUpdate	TeamCopilotInteraction	Copilot.TeamCopilot.AINotes	Teams
LiveNotesUpdate	TeamCopilotInteraction	Copilot.TeamCopilot.LiveNotes	Teams
LiveNotesUpdate	TeamCopilotInteraction	Copilot.TeamCopilot.MeetingModerator	Teams
TeamCopilotMsgInteraction	TeamCopilotInteraction	Copilot.TeamCopilot.Message	Teams

Identifying if Copilot accessed the web

When you enable web search, Microsoft 365 Copilot and Microsoft 365 Copilot Chat parse user prompts and determine whether web search would improve the quality of the response. To identify if Copilot referenced the public web in a user interaction, review the *AISystemPlugin.Id* property in the *CopilotInteraction* audit record. *AISystemPlugin.Id* contains the value *BingWebSearch* when user Copilot requests use the public web via Microsoft Bing for additional data.

Accessing Copilot audit logs

Access Copilot audit logs by using the Microsoft Purview portal and selecting **Audit**.

To search for specific Copilot or AI application scenarios, use the **Activities – operation names** field in the Microsoft Purview portal to filter audit logs by properties like *Operation*, *RecordType*, and *Workload*.

If you need to search for audit logs containing a specific *AppIdentity* value or set of values, first search and export all relevant Copilot audit logs by filtering by operation name. From the exported search results, apply a filter on the *AppIdentity* property offline.

Audit logs for Microsoft Copilot and Microsoft Copilot Chat include the model provider and model name when a request is processed. This information can help support auditing, compliance, and governance requirements in organizations that enable multiple model providers.

For example, if a user selects a specific model instead of **Auto**, audit records might include entries similar to:

- `ModelProvider: OpenAI, ModelName: gpt-55-high`
- `ModelProvider: Anthropic, ModelName: claude-sonnet-4-6` And, if a user selects **Auto**, the model provider and model name might not be available for every request, such as when an internal model is used.

Currently, Cowork doesn't show provider details.

Next steps

- Search the audit log: Run targeted searches for Copilot and AI application activities in the audit log.
- Manage audit log retention policies: Create retention policies to preserve Copilot audit records for your required period.
- Export, configure, and view audit log records: Export Copilot activity records for detailed analysis or compliance reporting.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 08/26/2026

Search the audit log to investigate common support issues

This article describes how to use the audit log search tool to help you investigate common support issues. This guidance includes using the audit log to:

- Find the IP address of the computer used to access a compromised account
- Determine who set up email forwarding for a mailbox
- Determine if a user deleted email items in their mailbox
- Determine if a user created an inbox rule
- Investigate why there was a successful sign-in by a user outside your organization
- Search for mailbox activities performed by users with non-E5 licenses
- Search for mailbox activities performed by delegate users

Use the audit log search tool

Each investigation scenario described in this article is based on using the audit log search tool in the Microsoft Purview portal. This section lists the permissions required to search the audit log and describes the steps to access and run audit log searches. Each scenario section explains how to configure an audit log search query and what to look for in the detailed information in the audit records that match the search criteria.

Permissions required to use the audit log search tool

To search the audit log, you must be assigned the *Audit Logs* or *View-Only Audit Logs* roles in Purview. By default, the *Audit Reader* and *Audit Manager* role groups on the **Role groups** page in the Microsoft Purview portal have these roles assigned.

To access audit cmdlets, you must be assigned the *Audit Logs* and *View-Only Audit Logs* roles in the Exchange admin center. By default, the *Compliance Management* and *Organization Management* role groups on the **Permissions page** in the Exchange admin center have these roles assigned.

For more information, see Get started with auditing solutions.

Running audit log searches

For detailed guidance on how to run a search of the audit log, see Search the audit log.

Find the IP address of the computer used to access a compromised account

Most audit records include the IP address for any user activity. The audit record also includes information about the client used.

To configure an audit log search query for this scenario, use the following settings:

- **Activities:** If relevant to your case, select a specific activity to search for. For investigating compromised accounts, consider selecting the **User signed in to mailbox** activity under **Exchange mailbox activities**. This selection returns auditing records that show the IP address used when signing in to the mailbox. Otherwise, leave this field blank to return audit records for all activities.

> **Tip**
>
> If you leave this field blank, the search returns **UserLoggedIn** activities, which is a Microsoft Entra activity that indicates that someone signed in to a user account. Use filtering in the search results to display the **UserLoggedIn** audit records.

- **Start date** and **End date:** Select a date range that's applicable to your investigation.
- **Users:** If you're investigating a compromised account, select the user whose account was compromised. This selection returns audit records for activities performed by that user account.
- **File, folder, or site:** Leave this field blank.

After you run the search, the IP address for each activity appears in the **IP address** column in the search results. Select the record in the search results to view more detailed information on the detail page.

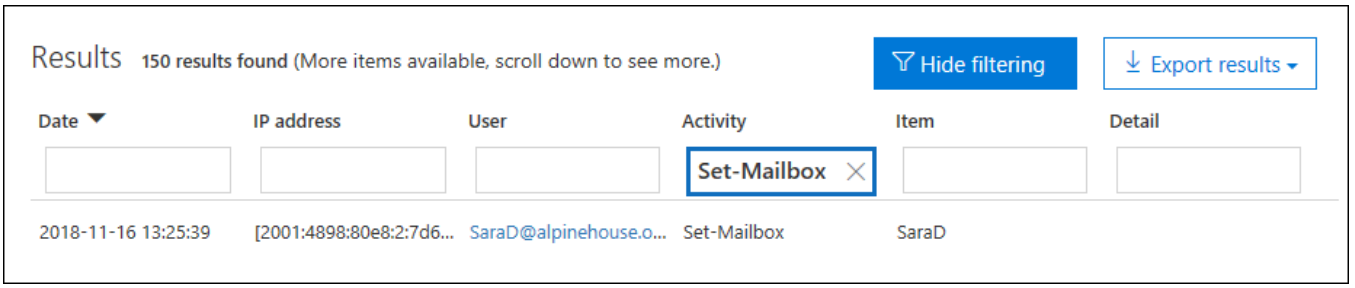
Determine who set up email forwarding for a mailbox

When you configure email forwarding for a mailbox, you forward email messages sent to that mailbox to another mailbox. You can forward messages to users inside or outside of your organization. When you set up email forwarding on a mailbox, use the underlying Exchange Online cmdlet **Set-Mailbox**.

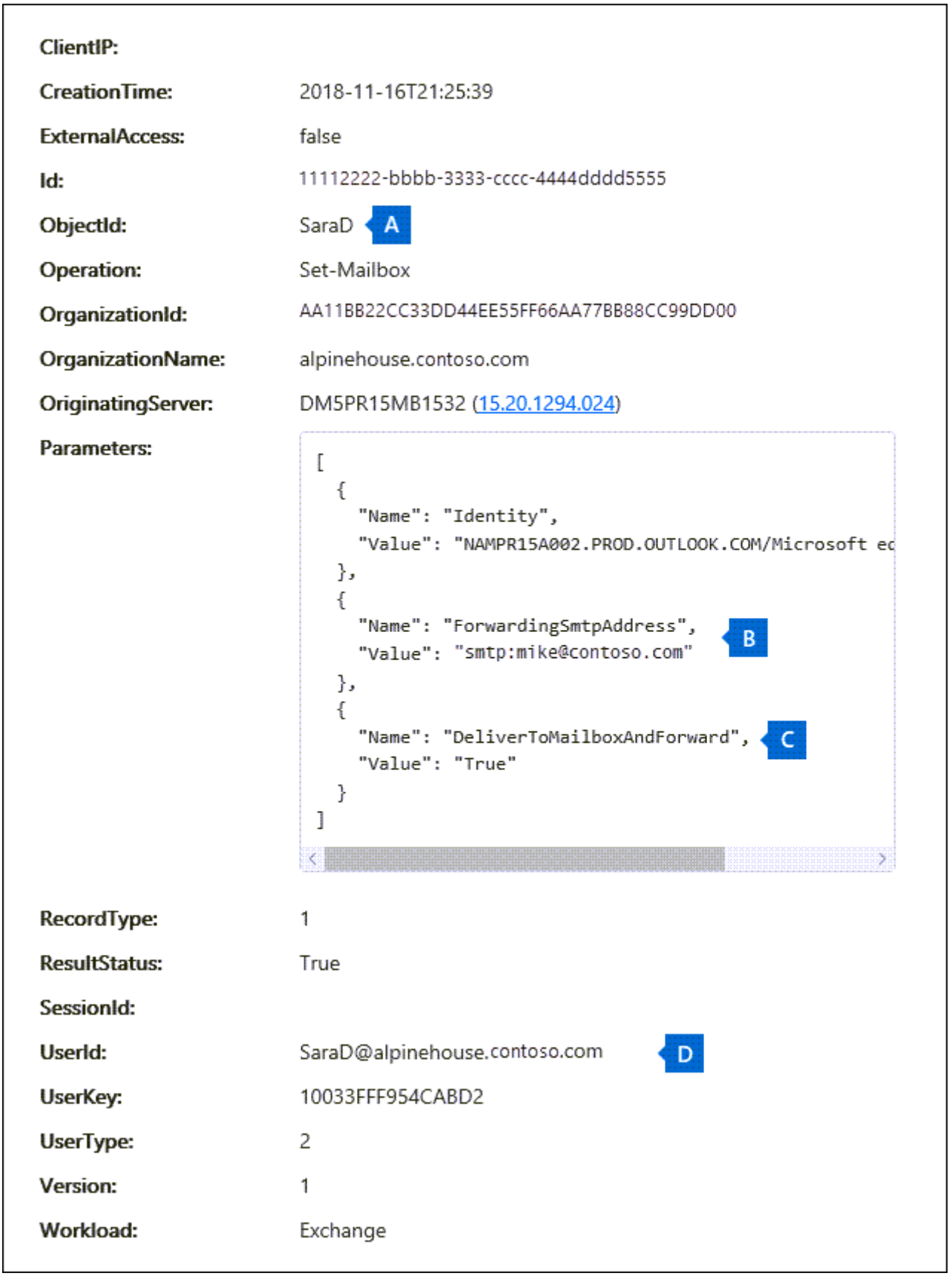
To configure an audit log search query for this scenario, use the following settings:

- **Activities:** Leave this field blank so the search returns audit records for all activities. This setting is necessary to return any audit records related to the **Set-Mailbox** cmdlet.
- **Start date** and **End date:** Select a date range that's applicable to your investigation.
- **Users:** Unless you're investigating an email forwarding issue for a specific user, leave this field blank. This setting helps you identify if email forwarding was set up for any user.
- **File, folder, or site:** Leave this field blank.

After you run the search, select **Filter results** on the search results page. In the box under the **Activity** column header, type **Set-Mailbox** so only audit records related to the **Set-Mailbox** cmdlet are displayed.



You need to look at the details of each audit record to determine if the activity is related to email forwarding. Select the audit record to display the **Details** flyout page, and then select **More information**. The following screenshot and descriptions highlight the information that indicates email forwarding was set on the mailbox.



a. In the **ObjectId** field, the alias of the mailbox that email forwarding was set on is displayed. This mailbox also appears in the **Item** column in the search results page.

b. In the **Parameters** field, the value *ForwardingSmtpAddress* indicates that email forwarding was set on the mailbox. In this example, mail is being forwarded to the email address mike@contoso.com, which is outside of the alpinehouse.onmicrosoft.com organization.

c. The *True* value for the *DeliverToMailboxAndForward* parameter indicates that a copy of the message is delivered to sarad@alpinehouse.onmicrosoft.com *and* is forwarded to the email address specified by the *ForwardingSmtpAddress* parameter, which in this example is mike@contoso.com. If the value for the *DeliverToMailboxAndForward* parameter is set to *False*, then email is only forwarded to the address specified by the *ForwardingSmtpAddress* parameter. It's not delivered to the mailbox specified in the **ObjectId** field.

d. The **UserId** field indicates the user who set email forwarding on the mailbox specified in the **ObjectId** field. This user also appears in the **User** column on the search results page. In this case, it seems that the owner of the mailbox set email forwarding on the mailbox.

If you determine that email forwarding shouldn't be set on the mailbox, remove it by running the following command in Exchange Online PowerShell:

PowerShell

```
Set-Mailbox <mailbox alias> -ForwardingSmtpAddress $null
```

For more information about the parameters related to email forwarding, see the Set-Mailbox article.

Determine if a user deleted email items

Starting in January 2019, Microsoft turned on mailbox audit logging by default for all Office 365 and Microsoft 365 organizations. This change means that the system automatically logs certain actions that mailbox owners perform, and you can access the corresponding mailbox audit records when you search the mailbox audit log. Before this change, you had to manually enable mailbox audit logging for every user mailbox in your organization.

By default, the system logs mailbox actions that include the SoftDelete and HardDelete mailbox actions that mailbox owners perform. By using this audit logging, you can search the audit log for events related to deleted email items. For more information about mailbox auditing by default, see Manage mailbox auditing.

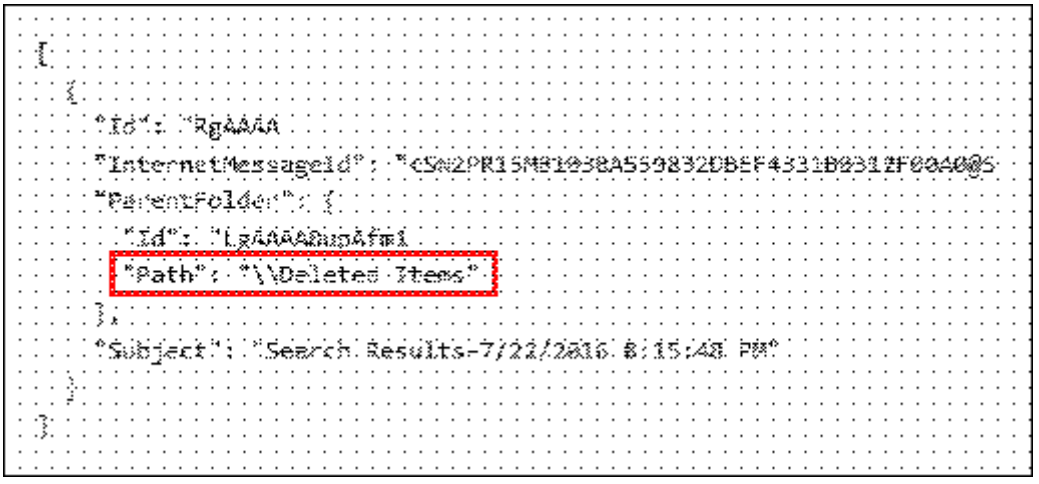
To configure an audit log search query for this scenario, use the following settings:

- **Activities:** Under **Exchange mailbox activities**, select one or both of the following activities:
- **Deleted messages from Deleted Items folder:** This activity corresponds to the **SoftDelete** mailbox auditing action. The audit log records this activity when a user deletes a message from the **Deleted Items** folder. It also records this activity when a user permanently deletes an item by selecting it and pressing **Shift+Delete**. After an item is permanently deleted, the user can recover it until the deleted item retention period expires.

- **Purged messages from mailbox:** This activity corresponds to the **HardDelete** mailbox auditing action. The audit log records this activity when a user purges an item from the **Recoverable Items** folder. Admins can use the search tool in the Microsoft Purview portal to search for and recover purged items until the deleted item retention period expires or longer if the user's mailbox is on hold.
- **Start date** and **End date:** Select a date range that's applicable to your investigation.
- **Users:** If you select a user in this field, the audit log search tool returns audit records for email items that the specified user deleted (SoftDeleted or HardDeleted). Sometimes the user who deletes an email isn't the mailbox owner.
- **File, folder, or site:** Leave this field blank.

After you run the search, you can filter the search results to display the audit records for soft-deleted items or for hard-deleted items. Select the audit record to display the **Details** flyout page, and then select **More information**. Additional information about the deleted item, such as the subject line and the location of the item when it was deleted, appears in the **AffectedItems** field. The following screenshots show an example of the **AffectedItems** field from a soft-deleted item and a hard-deleted item.

Example of AffectedItems field for soft-deleted item



Example of AffectedItems field for hard-deleted item



Recover deleted email items

Users can recover soft-deleted items if the deleted items retention period hasn't expired. In Exchange Online, the default deleted items retention period is 14 days, but admins can increase this setting to a maximum of 30 days. For instructions on recovering deleted items, see Recover deleted items or email in Outlook on the web.

As previously explained, admins might be able to recover hard-deleted items if the deleted item retention period hasn't expired or if the mailbox is on hold. In this case, items are kept until the hold duration expires. When you run a content search, soft-deleted and hard-deleted items in the Recoverable Items folder appear in the search results if they match the search query. For more information about running content searches, see Content Search in Office 365.

> **Tip**
>
> To search for deleted email items, search for all or part of the subject line that's displayed in the **AffectedItems** field in the audit record.

Determine if a user created an inbox rule

When users create an inbox rule for their Exchange Online mailbox, the system saves a corresponding audit record to the audit log. For more information about inbox rules, see:

- Use inbox rules in Outlook on the web
- Manage email messages in Outlook by using rules

To configure an audit log search query for this scenario, use the following settings:

- **Activities:** Under **Exchange mailbox activities**, select one or both of the following activities:
- **New-InboxRule Create new inbox rule from Outlook Web App**. This activity returns audit records when users create inbox rules by using Outlook web app or Exchange Online PowerShell.
- **Updated inbox rules from Outlook client**. This activity returns audit records when users create, modify, or remove inbox rules by using the Outlook desktop client.
- **Start date** and **End date:** Select a date range that's applicable to your investigation.
- **Users:** Unless you're investigating a specific user, leave this field blank. This approach helps you identify new inbox rules set up by any user.
- **File, folder, or site:** Leave this field blank.

After you run the search, the search results display any audit records for this activity. Select an audit record to display the **Details** flyout page, and then select **More information**. The

Parameters field displays information about the inbox rule settings. The following screenshot and descriptions highlight the information about inbox rules.

ClientIP:

CreationTime: 2018-11-27T18:50:21

ExternalAccess: false

Id:

ObjectId: A NAMPR15A002.PROD.OUTLOOK.COM/Microsoft Exchange Hosted Organizations/alpinehouse.onmicrosoft.com/SaraD\Move messages from admin

Operation: New-InboxRule

OrganizationId:

OrganizationName: alpinehouse.onmicrosoft.com

OriginatingServer: DM5PR15MB1532 (15.20.1361.017)

Parameters:


```
[
  {
    "Name": "AlwaysDeleteOutlookRulesBlob",
    "Value": "False"
  },
  {
    "Name": "Force",
    "Value": "False"
  },
  {
    "Name": "From",
    "Value": "admin@alpinehouse.onmicrosoft.com"
  },
  {
    "Name": "MoveToFolder",
    "Value": "adminsearch"
  },
  {
    "Name": "Name",
    "Value": "Move messages from admin"
  },
  {
    "Name": "StopProcessingRules",
    "Value": "False"
  }
]
```


B

C

RecordType: 1

ResultStatus: True

SessionId:

UserId: SaraD@alpinehouse.onmicrosoft.com D

UserKey: 10033FFF954CABD2

UserType: 2

Version: 1

Workload: Exchange

a. In the **ObjectId** field, the audit record displays the full name of the inbox rule. This name includes the alias of the user's mailbox (for example, SaraD) and the name of the inbox rule (for example, "Move messages from admin").

b. In the **Parameters** field, the audit record displays the condition of the inbox rule. In this example, the condition is specified by the *From* parameter. The value defined for the *From* parameter indicates that the inbox rule acts on email sent by admin@alpinehouse.onmicrosoft.com. For a complete list of the parameters that you can use to define conditions of inbox rules, see the New-InboxRule article.

c. The *MoveToFolder* parameter specifies the action for the inbox rule. In this example, messages received from admin@alpinehouse.onmicrosoft.com are moved to the folder named *AdminSearch*. For a complete list of parameters that you can use to define the action of an inbox rule, see the New-InboxRule article.

d. The **UserId** field indicates the user who created the inbox rule specified in the **ObjectId** field. The **User** column on the search results page also displays this user.

Investigate why a user outside your organization successfully signed in

When you review audit records in the audit log, you might see records that show Microsoft Entra ID authenticated an external user who successfully signed in to your organization. For example, an admin in contoso.onmicrosoft.com might see an audit record showing that a user from a different organization (for example, fabrikam.onmicrosoft.com) successfully signed in to contoso.onmicrosoft.com. Similarly, you might see audit records that show users with a Microsoft Account (MSA), such as an Outlook.com or Live.com, successfully signed in to your organization. In these situations, the audited activity is **User logged In**.

This behavior is by design. Microsoft Entra ID, the directory service, supports *pass-through authentication* when an external user tries to access a SharePoint site or a OneDrive location in your organization. When the external user tries to access the resource, they're prompted to enter their credentials. Microsoft Entra ID uses the credentials to authenticate the user, so only Microsoft Entra ID verifies that the user is who they say they are. The indication of the successful sign-in in the audit record comes from Microsoft Entra authenticating the user. The successful sign-in doesn't mean that the user accessed any resources or performed any other actions in your organization. It only shows that Microsoft Entra ID authenticated the user. For a pass-through user to access SharePoint or OneDrive resources, a user in your organization must explicitly share a resource with the external user by sending them a sharing invitation or anonymous sharing link.

> **Note**
>
> Microsoft Entra ID supports pass-through authentication only for *first-party applications*, such as SharePoint Online and OneDrive for Business. It doesn't support pass-through authentication for other third-party applications.

Here's an example and descriptions of relevant properties in an audit record for a **User logged In** event that results from pass-through authentication. Select the audit record to display the **Details** flyout page, and then select **More information**.

Details

Date: 2019-08-15 11:49:11

IP address: 131.107.160.12

User: pilarp@fabrikam.com

Activity: User logged in

Item: 00000003-0000-0ff1-ce00-000000000000

Detail:

More information

Actor:


```
[
  {
    "ID": "Unknown",
    "Type": 0
  },
  {
    "ID":"pilarp@fabrikam.com",
    "Type": 5
  }
]
```


A

B

ActorContextId:

ActorIpAddress: 131.107.160.12

ApplicationId: 11112222-bbbb-3333-cccc-4444dddd5555

C

AzureActiveDirectoryEve 1

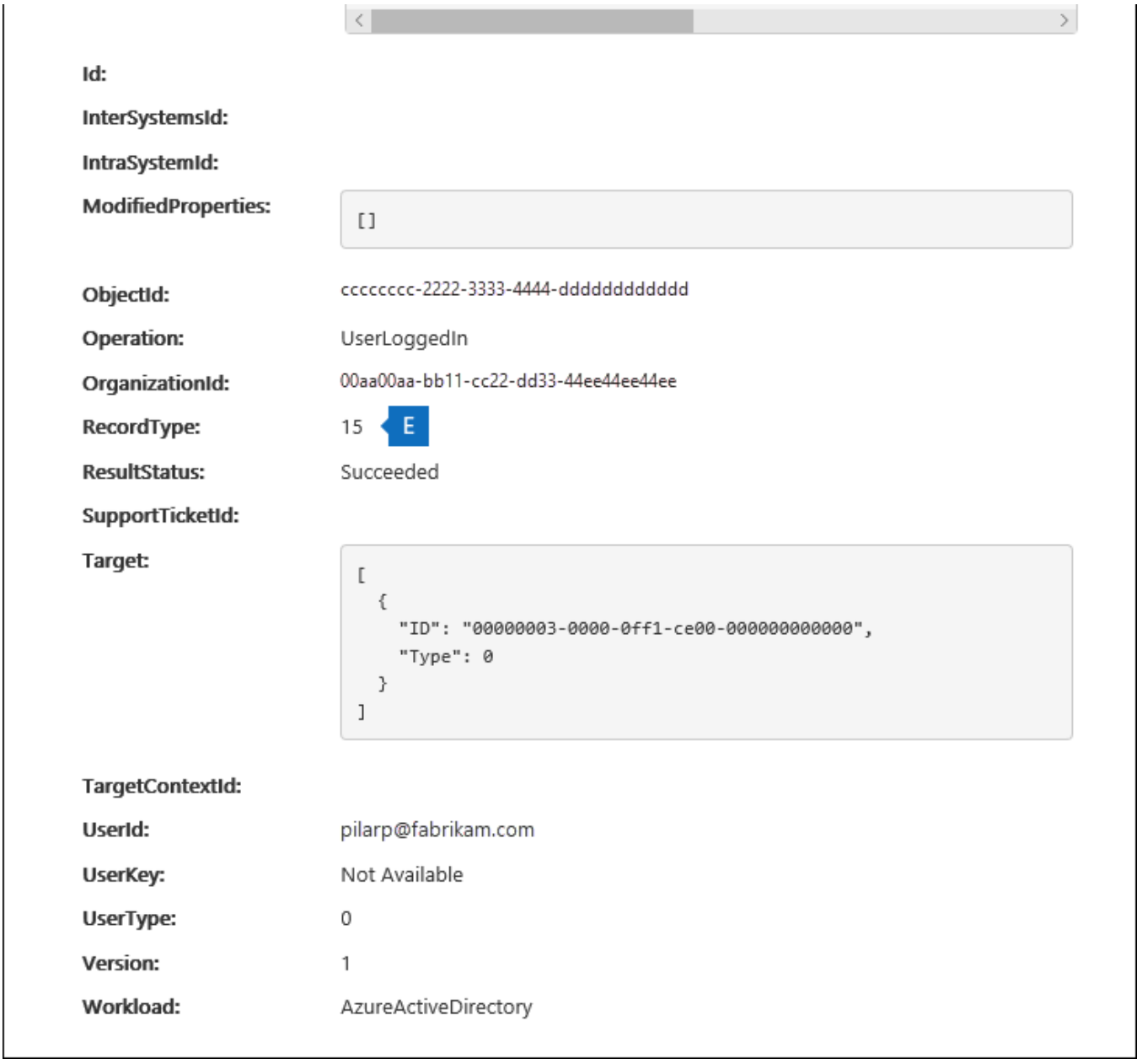
ClientIP: 131.107.160.12

CreationTime: 2019-08-15T18:49:11

ExtendedProperties:


```
[
  {
    "Name": "UserAgent",
    "Value": "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/
  },
  {
    "Name": "UserAuthenticationMethod",
    "Value": "1"
  },
  {
    "Name": "RequestType",
    "Value": "OAuth2:Authorize"
  },
  {
    "Name": "ResultStatusDetail",
    "Value": "Success"
  },
  {
    "Name": "KeepMeSignedIn",
    "Value": "True"
  }
]
```


D



a. This field shows that the user who tried to access a resource in your organization isn't found in your organization's Microsoft Entra ID.

b. This field shows the UPN of the external user that tried to access a resource in your organization. This user ID also appears in the **User** and **UserId** properties in the audit record.

c. The **ApplicationId** property shows the application that triggered the sign-in request. The value of `00000003-0000-0ff1-ce00-000000000000` displayed in the **ApplicationId** property in this audit record indicates SharePoint Online. OneDrive for Business also has this same **ApplicationId**.

d. This field shows that the pass-through authentication succeeded. In other words, Microsoft Entra ID successfully authenticated the user.

e. The **RecordType** value of **15** shows that the audited activity (UserLoggedIn) is a Secure Token Service (STS) sign-in event in Microsoft Entra ID.

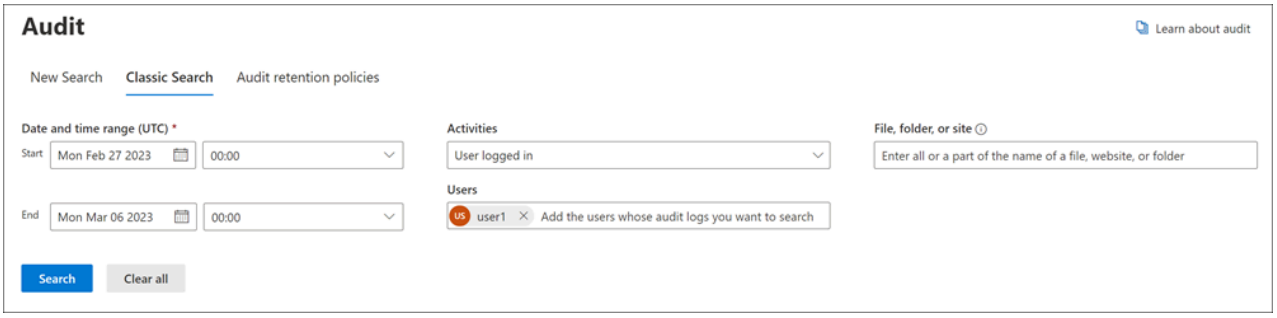
For more information about the other properties displayed in a UserLoggedIn audit record, see the Microsoft Entra related schema information in Office 365 Management Activity API schema.

Here are two example scenarios that result in a successful **User logged in** audit activity because of pass-through authentication:

- A user with a Microsoft Account (such as SaraD@outlook.com) tries to access a document in a OneDrive for Business account in fourthcoffee.onmicrosoft.com and there's no corresponding guest user account for SaraD@outlook.com in fourthcoffee.onmicrosoft.com.
- A user with a Work or School account in an organization (such as pilarp@fabrikam.onmicrosoft.com) tries to access a SharePoint site in contoso.onmicrosoft.com and there's no corresponding guest user account for pilarp@fabrikam.com in contoso.onmicrosoft.com.

Tips for investigating successful sign-ins resulting from pass-through authentication

- Search the audit log for activities performed by the external user identified in the **User logged in** audit record. Type the UPN for the external user in the **Users** box and use a date range if relevant to your scenario. For example, you can create a search using the following search criteria:

  In addition to the **User logged in** activities, you might see other audit records, such as records that indicate a user in your organization shared resources with the external user and whether the external user accessed, modified, or downloaded a document that was shared with them.

- Search for SharePoint sharing activities that indicate a file was shared with the external user identified by a **User logged in** audit record. For more information, see Use sharing auditing in the audit log.

- Export the audit log search results that contain records relevant to your investigation so that you can use Excel to search for other activities related to the external user. For more information, see Export, configure, and view audit log records.

Search for mailbox activities performed by users with non-E5 licenses

Even when mailbox auditing on by default is turned on for your organization, you might notice that mailbox audit events for some users don't appear in audit log searches by using the Microsoft Purview portal, the **Search-UnifiedAuditLog** cmdlet, or the Office 365 Management Activity API. The reason for this behavior is that mailbox audit events return only for users with E5 licenses when you use one of the previous methods to search the unified audit log.

To retrieve mailbox audit log records for non-E5 users, use one of the following workarounds:

- Manually enable mailbox auditing on individual mailboxes (run the `Set-Mailbox -Identity <MailboxIdentity> -AuditEnabled $true` command in Exchange Online PowerShell). After you enable mailbox auditing, search for mailbox audit activities by using the Microsoft Purview portal, the **Search-UnifiedAuditLog** cmdlet, or the Office 365 Management Activity API.

  > ⓘ **Note**
  >
  > If mailbox auditing already appears to be enabled on the mailbox, but your searches return no results, change the value of the *AuditEnabled* parameter to `$false` and then back to `$true`.

Search for mailbox activities performed in a specific mailbox (including shared mailboxes)

When you use the **Users** dropdown list in the audit log search tool in the Microsoft Purview portal or the **Search-UnifiedAuditLog -UserIds** command in Exchange Online PowerShell, you can search for activities performed by a specific user. For mailbox audit activities, this type of search finds activities performed by the specified user. It doesn't guarantee that the search results include all activities performed in the same mailbox. For example, an audit log search doesn't return audit records for activities performed by a delegate user because searching for mailbox activities performed by a specific user doesn't return activities performed by a delegate user who has been assigned permissions to access another user's mailbox. A delegate user is someone who is assigned the SendAs, SendOnBehalf, or FullAccess mailbox permission to another user's mailbox.

Also, using the **User** dropdown list in the audit log search tool or the **Search-UnifiedAuditLog -UserIds** doesn't return results for activities performed in a shared mailbox.

To search for the activities performed in a specific mailbox or to search for activities performed in a shared mailbox, use the following syntax when running the **Search-UnifiedAuditLog** cmdlet:

PowerShell

```
Search-UnifiedAuditLog -StartDate <date> -EndDate <date> -FreeText (Get-Mailbox <mailbox identity).ExchangeGuid
```

For example, the following command returns audit records for activities performed in the Contoso Compliance Team shared mailbox between August 2020 and October 2020:

PowerShell

```
Search-UnifiedAuditLog -StartDate 08/01/2020 -EndDate 10/31/2020 -FreeText (Get-Mailbox complianceteam@contoso.onmicrosoft.com).ExchangeGuid
```

Last updated on 03/30/2026

Use MailItemsAccessed to investigate compromised accounts

A compromised user account (also called an *account takeover*) is a type of cyberattack where an attacker gains access to a user account and operates as the user. When investigating compromised email accounts, assume that more mail data is compromised than might be indicated by tracing the attacker's actual presence. Depending on the type of data in email messages, you might face regulatory fines unless you can prove that sensitive information wasn't exposed. For example, HIPAA-regulated organizations face significant fines if there's evidence that patient health information (PHI) was exposed.

Use the *MailItemsAccessed* mailbox-auditing action to investigate compromised email accounts by auditing accesses of mail data by mail protocols and clients. This auditing action helps you better understand email data breaches and identify the scope of compromises to specific mail items. When a cyberattacker gains access to a specific piece of mail, Exchange Online audits the event even though there's no indication that the mail item was read.

The MailItemsAccessed mailbox-auditing action

The MailItemsAccessed action is part of Audit (Standard) functionality. It's part of Exchange mailbox auditing and is enabled by default for users that are assigned an Office 365 E3/E5 or Microsoft 365 E3/E5 license.

The MailItemsAccessed mailbox-auditing action covers all mail protocols: POP, IMAP, MAPI, EWS, Exchange ActiveSync, and REST. It also covers both types of accessing mail: *sync*, where clients download sets of messages from the server, and *bind*, where a user or client accesses an individual message.

Auditing sync access

Sync operations are only recorded when a mailbox is accessed by a desktop version of the Outlook client for Windows or Mac. During the sync operation, these clients typically download a large set of mail items from the cloud to a local computer. The audit volume for sync operations is huge. So, instead of generating an audit record for each mail item that's synced, the system generates an audit event for the mail folder containing items that were synced. All mail items in

the synced folder are assumed to be compromised. The access type is recorded in the OperationProperties field of the audit record.

For an example of displaying the sync access type in an audit record, see Use MailItemsAccessed audit records for forensic investigations, specifically the procedure for checking sync activities to determine context.

Auditing bind access

A bind operation is an individual access to an email message. For bind access, the InternetMessageId of individual messages is recorded in the audit record. The MailItemsAccessed audit action records bind operations and then aggregates them into a single audit record. All bind operations that occur within a 2-minute interval are aggregated in a single audit record in the Folders field within the AuditData property. Each message that was accessed is identified by its InternetMessageId. The number of bind operations that were aggregated in the record is displayed in the OperationCount field in the AuditData property.

For an example of displaying the bind access type in an audit record, see the "Check for bind activities" procedure in Use MailItemsAccessed audit records for forensic investigations.

Use MailItemsAccessed audit records for forensic investigations

Mailbox auditing generates audit records for access to email messages so that you can be confident that email messages aren't compromised. In circumstances where data access is uncertain, assume that the data is compromised by recording all mail access activity.

Use MailItemsAccessed audit records for forensic purposes after a data breach is resolved and the bad actor is evicted. Complete the following steps to begin your investigation:

1. Identify the set of mailboxes that the bad actor compromised.
2. Determine the time frame when the bad actor had access to mailboxes in your organization.
3. Use the **Search-UnifiedAuditLog** cmdlet in Exchange Online PowerShell to search audit records that correspond to the data breach.

Run one of the following commands to search for MailItemsAccessed audit records:

Unified audit log:

PowerShell

```
Search-UnifiedAuditLog -StartDate 01/06/2020 -EndDate 01/20/2020 -UserIds
<user1,user2> -Operations MailItemsAccessed -ResultSize 1000
```

Complete the following steps to use MailItemsAccessed audit records to investigate a compromised user account. Each step includes the command syntax for the **Search-UnifiedAuditLog** cmdlet.

1. Check for sync activities. If a bad actor uses an email client to download messages in a mailbox, they can disconnect the computer from the Internet and access the messages locally without interacting with the server. In this case, mailbox auditing can't audit these activities.

   To search for MailItemsAccessed records where the mail items are accessed by a sync operation, run the following command:

   Unified audit log:

   PowerShell

   ```
   Search-UnifiedAuditLog -StartDate 01/06/2020 -EndDate 02/20/2020 -UserIds
   <user1,user2> -Operations MailItemsAccessed -ResultSize 1000 | Where
   {$_.AuditData -like '*"MailAccessType","Value":"Sync"*'} | FL
   ```

2. Check sync activities to determine if any of them happen in the same context as the one used by the bad actor to access the mailbox. Context is identified and differentiated by the IP address of the client computer used to access the mailbox and the mail protocol.

   Use the properties listed in the following table to investigate. These properties are located in the **AuditData** or **OperationProperties** property. If any of the syncs occur in the same context as the bad actor activity, assume the bad actor synced all mail items to their client, which means the entire mailbox is compromised.

   Expand table

Property	Description
ClientInfoString	Describes protocol, client (includes version)
ClientIPAddress	IP address of the client machine.
SessionId	Session ID helps to differentiate bad actor actions vs day-to-day user activities on the same account (useful for compromised accounts)
UserId	UPN of the user reading the message.

3. Check for bind activities. After checking for sync activities and comparing their context to bad actor activity, you can be confident that all other access to email messages by the bad actor is captured in the MailItemsAccessed audit records that have a MailAccessType property with a value of "Bind".

   To search for MailItemsAccessed records where the mail items are accessed by a Bind operation, run the following command.

   Unified audit log:

   ```
   PowerShell
   Search-UnifiedAuditLog -StartDate 01/06/2020 -EndDate 01/20/2020 -UserIds
   <user1,user2> -Operations MailItemsAccessed -ResultSize 1000 | Where
   {$_.AuditData -like '*"MailAccessType","Value":"Bind"*'} | FL
   ```

   Email messages that are accessed are identified by their internet message ID. You can also check to see if any audit records have the same context as the ones for other bad actor activity.

   You can use the audit data for bind operations in two different ways:

   - Access or collect all email messages the bad actor accessed by using the InternetMessageId to find them and then checking to see if any of those messages contains sensitive information.
   - Use the InternetMessageId to search audit records related to a set of potentially sensitive email messages. This is useful if you're concerned only about a few messages.

Filter duplicate audit records

To remove auditing noise, the system filters out duplicate audit records for the same bind operations that occur within an hour of each other. The system also filters out sync operations at one-hour intervals. An exception to this deduplication process occurs if, for the same InternetMessageId, any of the properties described in the following table are different. If one of these properties is different in a duplicate operation, the system generates a new audit record. The following table describes each property in more detail.

Expand table

Property	Description
ClientIPAddress	IP address of the client computer.

Property	Description
ClientInfoString	The client protocol, client used to access the mailbox.
ParentFolder	The full folder path of the mail item that was accessed.
Logon_type	The logon type of the user who performed the action. The logon types (and their corresponding Enum value) are Owner (0), Admin (1), or Delegate (2).
MailAccessType	Whether the access is a bind or a sync operation.
MailboxUPN	The UPN of the mailbox where the message being read is located.
User	The UPN of the user reading the message.
SessionId	The Session ID helps to differentiate attacker actions and day-to-day user activities in the same mailbox (if an account is compromised). For more information about sessions, see Contextualizing attacker activity within sessions in Exchange Online.

Next steps

Learn more about related audit tasks:

- Identify who deleted an email message or why an email is missing: Check for deleted or moved messages.
- Identify who modified mailbox rules: Check for forwarding or auto-delete rules.
- Export, configure, and view audit log records: Export audit data for your report.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 06/24/2026

Identify who deleted an email message or why an email is missing

Use this article to investigate missing email messages in Exchange Online and identify who deleted them. It guides administrators through using the Microsoft Purview audit log and Exchange Online PowerShell to locate deletion events, analyze mailbox configurations, check retention and migration activity, and determine whether mailbox rules or shared-mailbox access caused the loss.

Use these methods to investigate:

- Emails deleted by users or administrators
- Missing emails after migration or sync issues
- Emails removed by mailbox rules or retention policies
- Deleted emails from shared mailboxes
- Emails not appearing in expected folders

Before you begin

To investigate deleted emails and missing messages, you need:

- The *Audit Logs* role assigned in Microsoft Purview
- To connect to Exchange Online PowerShell by using Connect-ExchangeOnline

How to identify deleted emails

Use the following methods to investigate missing emails and identify deletion activities. Choose the method based on the type of deletion you're investigating.

Search for deleted emails by operation type

Use this method to search for the following operations:

- **SoftDelete**: Items moved to Deleted Items folder.
- **HardDelete**: Items permanently removed from mailbox.
- **MoveToDeletedItems**: Items moved to Deleted Items folder by user action.

To search for audit records of email deletions by using specific deletion operations, run the following command:

PowerShell

```
Search-UnifiedAuditLog -StartDate 01/06/2020 -EndDate 01/20/2020 -UserIds
<user1,user2> -Operations SoftDelete,HardDelete,MoveToDeletedItems -ResultSize 1000
```

Replace `<user1,user2>` with user email addresses. Specify multiple users by separating user names with commas.

Search for shared mailbox deletions

To investigate deletions from shared mailboxes, run the following command:

PowerShell

```
Search-UnifiedAuditLog -StartDate 01/06/2020 -EndDate 01/20/2020 -UserIds
<user1,user2> -FreeText <sharedmailbox@domain.com> -Operations
SoftDelete,HardDelete,MoveToDeletedItems -ResultSize 1000
```

> ⓘ Note
>
> Shared mailbox auditing might not be enabled by default. If this search returns no results, see Shared mailbox auditing not configured to enable auditing.

Search for missing emails by using keywords

This method helps identify deletion records for emails with specific subjects or content.

To search for audit records related to specific missing emails, run the following command:

PowerShell

```
Search-UnifiedAuditLog -StartDate 01/06/2020 -EndDate 01/20/2020 -FreeText "<email
subject or keyword>" -Operations SoftDelete,HardDelete,MoveToDeletedItems -ResultSize
1000
```

Search for comprehensive email activities

This broader search includes moves and updates that might explain missing emails.

To search for all activities that affect email visibility, run the following command:

PowerShell

```
Search-UnifiedAuditLog -StartDate 01/06/2020 -EndDate 01/20/2020 -UserIds
<user1,user2> -Operations SoftDelete,HardDelete,MoveToDeletedItems,Move,Update,Create
-ResultSize 1000
```

Understanding search results

Getting no results from audit searches can provide important insights about what didn't cause the missing emails:

- **Deletion operations not found**: Rules out user-initiated deletions during the timeframe.
- **No retention policy actions**: Indicates automated policies didn't delete the emails.
- **No migration activities**: Shows migration processes didn't remove the emails.
- **No administrator actions**: Confirms admins didn't perform bulk operations.

> ⓘ Note
>
> Document which searches return no results. This information helps narrow down the root cause by eliminating potential causes.

Missing search results

If your audit log searches don't find deletion records for missing emails, try the following steps.

Missing emails with no deletion audit records

Use the following steps to investigate when emails are missing but no deletion audit records are found.

1. Check if auditing was enabled when the deletion occurred.

   PowerShell

   ```
   Get-Mailbox <mailbox> | Select AuditEnabled,AuditOwner,AuditDelegate,WhenCreated
   ```

2. Enable comprehensive auditing for future monitoring.

   PowerShell

```
Set-Mailbox <mailbox> -AuditEnabled $true -AuditOwner
@{Add="SoftDelete","HardDelete","MoveToDeletedItems","Move","Create","Update"}
```

3. Search for retention policy actions that might remove emails.

```
PowerShell
Search-UnifiedAuditLog -StartDate 01/06/2020 -EndDate 01/20/2020 -UserIds
<user1,user2> -Operations TaggedAsRecord,ApplyRetentionTag -ResultSize 1000
```

If this command returns no results, it indicates that retention policies didn't automatically delete emails during the specified timeframe. This information helps you rule out automated policy-based deletions as the cause of missing emails.

Shared mailbox auditing not configured

Use the following steps to enable auditing for shared mailboxes when deletion records aren't found.

1. Check the current auditing configuration for the shared mailbox.

```
PowerShell
Get-Mailbox <sharedmailbox> | Select
AuditEnabled,AuditOwner,AuditDelegate,AuditAdmin
```

2. Enable comprehensive auditing for the shared mailbox.

```
PowerShell
Set-Mailbox <sharedmailbox> -AuditEnabled $true -AuditOwner
@{Add="SoftDelete","HardDelete","MoveToDeletedItems"} -AuditDelegate
@{Add="SoftDelete","HardDelete","MoveToDeletedItems","SendAs","Move"}
```

3. Search for activities by users with shared mailbox access.

```
PowerShell
Get-MailboxPermission <sharedmailbox> | Where {$_.User -ne "NT AUTHORITY\SELF"} |
ForEach {
    Search-UnifiedAuditLog -StartDate 01/06/2020 -EndDate 01/20/2020 -UserIds
$_.User -Operations SoftDelete,HardDelete -ResultSize 100 }
```

Migration-related email loss investigation

The following steps show how to investigate emails lost during migration processes.

1. Search for migration-related activities during the migration timeframe.

```
PowerShell
Search-UnifiedAuditLog -StartDate <migration_start_date> -EndDate
<migration_end_date> -UserIds <user1,user2> -FreeText "<migration subject or
keyword>" -ResultSize 1000
```

Advanced investigation procedures

The following procedures show how to perform detailed analysis when standard searches don't reveal the cause of missing emails.

Analyze mailbox rules that might delete emails

If you suspect that mailbox rules are causing emails to be deleted or moved to unexpected folders, use the following dedicated mailbox rule investigation procedures.

> **Tip**
>
> For comprehensive mailbox rule investigation, see **Identify who modified mailbox rules** for detailed guidance on identifying who created, modified, or deleted mailbox rules that might affect email delivery.

Investigate retention and compliance policies

To check if compliance policies are causing email deletion, run the following commands:

1. Check retention policies applied to the mailbox.

```
PowerShell
Get-Mailbox <mailbox> | Select
RetentionPolicy,LitigationHoldEnabled,RetainDeletedItemsFor
```

2. Search for compliance-related deletions.

```
PowerShell
```

```
Search-UnifiedAuditLog -StartDate 01/06/2020 -EndDate 01/20/2020 -UserIds
<user1,user2> -Operations ComplianceSettingChanged,RetentionPolicyApplied -
ResultSize 1000
```

Check for administrator actions

To check if administrators performed actions that affected emails, run the following command:

PowerShell

```
Search-UnifiedAuditLog -StartDate 01/06/2020 -EndDate 01/20/2020 -UserIds
<user1,user2> -Operations Remove-MailboxFolderPermission,Set-Mailbox,New-
MailboxExportRequest,New-MailboxImportRequest -ResultSize 1000
```

Quick reference

Key operations for investigating deletions

The following table lists the key audit operations to use when investigating deleted or missing emails.

Expand table

Operation	Description	Use Case
ApplyRetentionTag	Retention policy applied to items	Automated policy actions
HardDelete	Items permanently removed from mailbox	Permanent deletions, second-level investigation
Move	Items moved between folders	Investigate folder changes
MoveToDeletedItems	Items moved to Deleted Items by user action	User-initiated moves to deleted items
SoftDelete	Items moved to Deleted Items folder	User deletions, first-level investigation
TaggedAsRecord	Items marked for retention	Compliance-related actions

Search parameters for investigating deletions

The following table describes the main `Search-UnifiedAuditLog` parameters used in these investigations.

Expand table

Parameter	Description	Example
-FreeText	Search for specific email identifiers	`<email subject or unique identifier>`
-Operations	Filter by deletion activity types	SoftDelete,HardDelete,MoveToDeletedItems
-ResultSize	Number of results to return	1,000 (standard), 5,000 (comprehensive)
-StartDate/-EndDate	Define investigation timeframe	Based on when emails went missing
-UserIds	Filter by who performed the action	`<user1@domain.com,user2@domain.com>`

Next steps

- Use MailItemsAccessed to investigate compromised accounts: Investigate whether deletions are part of a broader account compromise.
- Identify who modified mailbox rules: Check whether mailbox rules are auto-deleting or forwarding messages.
- Export, configure, and view audit log records: Export your investigation results for reporting or evidence preservation.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/19/2026

Search the audit log for specific email messages

Use this article to search the Microsoft Purview audit log for specific email messages that are missing, moved, or need investigation. It shows practical Exchange Online PowerShell searches and investigation steps you can use when standard audit searches don't return the records you expect.

Use these methods to investigate:

- Email migration issues
- Messages moved between folders
- Email access patterns
- Missing email items

Before you begin

To search audit logs for specific email items, you need:

- The *Audit Logs* role assigned in Microsoft Purview
- To connect to Exchange Online PowerShell by using Connect-ExchangeOnline

Search for specific email messages

Use these methods to find audit records for specific email messages. Select the method based on what information you have about the email you're looking for.

Search by email subject or keywords

Use the FreeText parameter to search for:

- Email subject lines
- Keywords from email content
- Sender names
- Any identifying text from the email

To search for audit records by using email subjects, keywords, or other identifying information, run the following command:

PowerShell

```
Search-UnifiedAuditLog -StartDate 01/06/2020 -EndDate 01/20/2020 -FreeText "<subject
or keyword>" -ResultSize 1000
```

Search for moved emails

To search for Move and MoveToDeletedItems operations in a mailbox, run the following command:

PowerShell

```
Search-UnifiedAuditLog -StartDate 01/06/2020 -EndDate 01/20/2020 -
UserIds <user1,user2> -Operations Move,MoveToDeletedItems -ResultSize 1000
```

> **Important**
>
> The audit feature for move operations isn't enabled by default.

Search all email activities

To search for all email-related activities in a mailbox, run the following command:

PowerShell

```
Search-UnifiedAuditLog -StartDate 01/06/2020 -EndDate 01/20/2020 -
UserIds <user1,user2> -
Operations Send,Create,Move,SoftDelete,HardDelete,MailItemsAccessed -ResultSize 1000
```

What to do when searches return no result

If your audit log searches don't find the email records you're looking for, try these investigation steps.

Missing emails after migration

Use the following steps to investigate missing emails after migration when searches return no results.

1. Check if auditing was enabled before the migration occurred.

PowerShell

```
Get-Mailbox <mailbox> | Select AuditEnabled,WhenCreated
```

1. Expand the search date range to cover a broader timeframe.

PowerShell

```
Search-UnifiedAuditLog -StartDate 01/01/2020 -EndDate 03/31/2020 -
ObjectIds <mailbox> -ResultSize 1000
```

Emails moved but no audit records found

Use the following steps to investigate when users move emails but you can't find audit records.

1. Check if auditing is configured for move operations.

PowerShell

```
Get-Mailbox <mailbox> | Select AuditOwner
```

1. Enable auditing for move operations if it's not configured.

PowerShell

```
Set-Mailbox <mailbox> -
AuditOwner @{Add="Move","MoveToDeletedItems","SoftDelete","HardDelete"}
```

1. Search using additional operation types that might capture the activity.

PowerShell

```
Search-UnifiedAuditLog -StartDate 01/06/2020 -EndDate 01/20/2020 -
ObjectIds <mailbox> -Operations Move,MoveToDeletedItems,Update -ResultSize 1000
```

Locating specific email messages

The following steps help you investigate when you can't find audit records for specific email messages.

1. Search by subject or keywords

PowerShell

```
Search-UnifiedAuditLog -StartDate 01/06/2020 -EndDate 01/20/2020 -FreeText "<subject or keyword>" -ResultSize 1000 | FL
```

1. Search by user

PowerShell

```
Search-UnifiedAuditLog -StartDate 01/06/2020 -EndDate 01/20/2020 -UserIds <sender> -ResultSize 1000 | FL
```

Advanced investigation procedures

The following procedures help you resolve issues when standard investigation steps don't resolve the issue.

Enable comprehensive mailbox auditing

To enable all relevant audit actions and ensure future activities are captured, run the following command:

PowerShell

```
Set-Mailbox <mailbox> -AuditEnabled $true -AuditOwner @{Add="Move","MoveToDeletedItems","SoftDelete","HardDelete","Create"} -AuditDelegate @{Add="SendAs","Move"}
```

Verify organization-wide audit configuration

To ensure that auditing is enabled at the organization level, run the following command:

PowerShell

```
Set-OrganizationConfig -AuditDisabled $false
```

Quick reference

Search parameters

Use these key parameters to search audit logs for specific email items:

Expand table

Parameter	Description	Example
-FreeText	Search for specific email identifiers	`<email subject or unique identifier>`
-Operations	Filter by deletion activity types	SoftDelete,HardDelete,MoveToDeletedItems
-ResultSize	Number of results to return	1,000 (standard), 5,000 (comprehensive)
-StartDate/-EndDate	Define investigation timeframe	Based on when emails went missing
-UserIds	Filter by who performed the action	`<user1@domain.com,user2@domain.com>`

Next steps

- Identify who deleted an email message or why an email is missing: Investigate deletion causes if your search reveals missing or deleted messages.
- Search the audit log for deleted mailbox items: Narrow your search to deleted items using deletion-specific operations and filters.
- Export, configure, and view audit log records: Export your email search results for detailed review or reporting.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 07/14/2026

Identify who modified mailbox rules

This guide explains how to investigate changes to mailbox rules in Exchange Online, including how to view a mailbox's current inbox and forwarding rules and how to search the Microsoft Purview audit log to identify who created, modified, or deleted those rules.

Use these methods to investigate:

- Changes to email forwarding rules
- Rules that cause emails to not appear in expected folders
- Unauthorized rule modifications

Before you begin

To investigate mailbox rule modifications, you need:

- The *Audit Logs* role assigned in Microsoft Purview
- To connect to Exchange Online PowerShell by using Connect-ExchangeOnline

How to identify mailbox rule modifications

Use Get-InboxRule to check current mailbox rules and Search-UnifiedAuditLog to find mailbox rule modifications.

Check current mailbox rules

The Get-InboxRule output shows:

- **Current rule configuration**: Rule configuration
- **Rule actions**: Move, delete, or forward
- **Rule status**: Enabled or disabled

To see what rules currently exist in a mailbox, run the following command:

PowerShell

```
Get-InboxRule -Mailbox <mailbox> | FL
Name,Description,DeleteMessage,MoveToFolder,Enabled
```

Search for rule modification audit records

The Search-UnifiedAuditLog command looks for:

- **New-InboxRule**: New rules created
- **Set-InboxRule**: Existing rules modified
- **Remove-InboxRule**: Rules deleted

To find out who created, modified, or deleted mailbox rules, run the following command:

PowerShell

```
Search-UnifiedAuditLog -StartDate 01/06/2020 -EndDate 01/20/2020 -UserIds
<user1,user2> -Operations New-InboxRule,Set-InboxRule,Remove-InboxRule -ResultSize
1000
```

What to do when searches return no results

If your audit searches don't find rule modification records:

1. **Expand the date range** to capture older changes:

PowerShell

```
Search-UnifiedAuditLog -StartDate 01/01/2020 -EndDate 03/31/2020 -UserIds
<user1,user2> -Operations New-InboxRule,Set-InboxRule,Remove-InboxRule -ResultSize
1000
```

1. **Enable auditing** for future rule changes:

PowerShell

```
Set-Mailbox <mailbox> -AuditEnabled $true -AuditOwner @{Add="Create","Update"}
```

Quick reference

Key operations for rule investigation

The following table lists the mailbox rule operations that appear in audit results.

Expand table

Operation	Description
New-InboxRule	New mailbox rule created.

Operation	Description
Remove-InboxRule	Mailbox rule deleted.
Set-InboxRule	Existing mailbox rule modified.

Essential commands

The following commands are the primary tools for investigating mailbox rule changes.

Expand table

Command	Purpose
`Get-InboxRule -Mailbox <mailbox>`	Check current rule configuration.
`Search-UnifiedAuditLog -Operations New-InboxRule,Set-InboxRule,Remove-InboxRule`	Find who made rule changes.

Next steps

- Use MailItemsAccessed to investigate compromised accounts: Determine whether unauthorized rule changes indicate a compromised account.
- Identify who deleted an email message or why an email is missing: Investigate whether modified rules caused email deletions or missing messages.
- Export, configure, and view audit log records: Export your rule modification findings for compliance documentation.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/24/2026

Use the audit log to identify Exchange transport rule activities

This article explains how to determine whether a user created, modified, or removed a transport rule in Exchange Online. Transport rules, also known as mail flow rules, can significantly impact email delivery and processing, so understanding their origin is important for security and compliance investigations.

Before you begin

Before you begin, make sure you have the following information:

- The email address of the sender or recipient targeted by the transport rule.
- The license type assigned to the affected user.
- The exact date and time when the transport rule change occurred.
- The date and time when the audit search was run. If you perform the search shortly after the transport rule update, the logs might not yet reflect the change.
- An Extended Message Trace (EMT) from the Exchange Online admin center for sample emails sent before and after the transport rule change.

You must sign in to Exchange Online PowerShell and Microsoft Purview portal to complete the audit investigation procedure.

Check mailbox auditing status

> **Note**
>
> Before searching for audit activity, verify that mailbox auditing is enabled for the affected user.

Complete the following steps to confirm that mailbox auditing is enabled:

1. In Exchange Online PowerShell, run the following command to check the organization-level audit configuration:

   PowerShell

```
Get-AdminAuditLogConfig | FL UnifiedAudit*
```

> **Note**
>
> Run this command in Exchange Online PowerShell. In Security & Compliance PowerShell, the `UnifiedAuditLogIngestionEnabled` property always returns `False`, even when auditing is enabled.

2. Check if mailbox audit logs are disabled at the tenant level:

PowerShell

```
Get-OrganizationConfig | FL AuditDisabled
```

3. At the mailbox level, `AuditEnabled` always returns `True`, even if auditing is disabled for a specific mailbox. To verify, use:

PowerShell

```
Get-Mailbox -Filter {AuditEnabled -eq "True"}
```

If you receive an error that the mailbox can't be found, auditing is disabled for that mailbox. For more information, see Verify the auditing status for your organization.

Verify administrator permissions

Ensure that you have the required permissions to search or export audit logs:

1. In the Microsoft Purview portal, go to **Settings** > **Roles and scopes** > **Role groups**.
2. Verify that your account is assigned to the **View-Only Audit Logs** or **Audit Logs** role.
3. In the Exchange admin center, use the **Audit Logs** role to enable or disable auditing and access audit cmdlets. Use the **View-Only Audit Logs** role to grant read-only access to audit logs.

For more information, see Assign permissions to search the audit log.

Check audit log retention policies

Audit log retention policies determine whether older activities are still available:

1. Determine when the activity occurred.
2. By default, audit logs are retained for 180 days.
3. If the affected user doesn't have an Audit (Premium) license and no audit retention policy is configured, audit logs older than 180 days aren't available.

For more information, see Get-UnifiedAuditLogRetentionPolicy.

Search the audit log for rule activities

When searching in Microsoft Purview Audit, use filters to narrow your results.

Recommended audit filters:

- **Date and time range (UTC)**: Specify the exact range when the activity occurred. Audit timestamps are always in UTC.
- **Activities – operation names**:
  - New-TransportRule: Logged when you create a new transport rule.
  - Set-TransportRule: Logged when you modify an existing transport rule.
  - Disable-TransportRule: Logged when you disable a transport rule.
  - Enable-TransportRule: Logged when you enable a transport rule.
  - Remove-TransportRule: Logged when you remove a transport rule.
- **User**: Enter the email address of the person who performed the operation. This field is optional if the user identity is unknown.

By applying these filters, you can identify whether a transport rule was created, modified, enabled, disabled, or removed, and by whom.

Next steps

Use the following resources to continue your investigation:

- Identify who modified mailbox rules: Investigate specific rule changes and determine who created, modified, or deleted mailbox rules.
- Use MailItemsAccessed to investigate compromised accounts: Investigate whether suspicious rule changes are linked to a compromised account.
- Export, configure, and view audit log records: Export your rule activity results for evidence collection or compliance review.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

Last updated on 06/24/2026

Investigate shared mailbox activities using audit logs

This article explains how to investigate shared mailbox activity by using the Microsoft Purview audit log and Exchange Online PowerShell. It describes practical search patterns and remediation steps to find deletions, sent-as activity, folder browsing, mailbox rule and forwarding changes, and other delegate actions.

Use these methods to investigate:

- Email deletions from shared mailboxes
- Who sent emails from shared mailboxes
- Delegate access activities
- Email moves between folders
- Forwarding and rule configurations
- Missing emails in shared mailboxes

Before you begin

To investigate shared mailbox activities, you need:

- The *Audit Logs* role assigned in Microsoft Purview
- To connect to Exchange Online PowerShell by using Connect-ExchangeOnline

How to investigate shared mailbox activities

Use these methods to investigate activities in shared mailboxes. Choose the method based on the type of activity you're investigating.

Search for deleted emails in shared mailboxes

To search for audit records of email deletions from a shared mailbox, run the following command:

PowerShell

```
Search-UnifiedAuditLog -StartDate 01/06/2020 -EndDate 01/20/2020 -FreeText <shared-mailbox@domain.com> -Operations SoftDelete,HardDelete,MoveToDeletedItems -ResultSize 500
```

This command searches for:

- **SoftDelete**: Items moved to Deleted Items folder.
- **HardDelete**: Items permanently removed from mailbox.
- **MoveToDeletedItems**: Items moved to Deleted Items by user action.

Search for emails sent from shared mailboxes

To identify who sent emails from a shared mailbox using delegate permissions, run the following command:

PowerShell

```
Search-UnifiedAuditLog -StartDate 01/06/2020 -EndDate 01/20/2020 -UserIds
<user1,user2> -FreeText <shared-mailbox@domain.com> -Operations SendAs -ResultSize
500
```

Search for email moves between folders

To search for move operations in a shared mailbox, run the following command:

PowerShell

```
Search-UnifiedAuditLog -StartDate 01/06/2020 -EndDate 01/20/2020 -UserIds
<user1,user2> -FreeText <shared-mailbox@domain.com> -Operations
Move,MoveToDeletedItems -ResultSize 300
```

Monitor delegate folder access (FolderBind)

To track when delegates browse specific folders in shared mailboxes, use the following steps:

Check FolderBind configuration:

PowerShell

```
Get-Mailbox <shared-mailbox@domain.com> | Select AuditDelegate | Where-Object
{$_.AuditDelegate -contains "FolderBind"}
```

Enable FolderBind auditing:

PowerShell

```
Set-Mailbox <shared-mailbox@domain.com> -AuditDelegate @{Add="FolderBind"}
```

Search for folder browsing activities:

PowerShell

```
Search-UnifiedAuditLog -StartDate 01/06/2020 -EndDate 01/20/2020 -FreeText <shared-mailbox@domain.com> -Operations FolderBind -ResultSize 500
```

> **Important**
>
> FolderBind tracks delegate and admin folder access only. It doesn't log shared mailbox owners browsing folders. Enable it proactively for compliance and security monitoring.

Common use cases:

- Monitor access to confidential folders in executive shared mailboxes
- Compliance audit trails for regulatory requirements
- Security investigations of unauthorized folder browsing
- Delegate behavior analysis for governance

Investigate delegate access activities

To identify users with delegate permissions to a shared mailbox, run the following command:

PowerShell

```
Get-MailboxPermission <shared-mailbox@domain.com> | Where-Object {$_.AccessRights -eq "FullAccess"}
```

To search for activities performed by a specific delegate, run the following command:

PowerShell

```
Search-UnifiedAuditLog -StartDate 01/06/2020 -EndDate 01/20/2020 -UserIds <delegate@domain.com> -ResultSize 500
```

> **Important**
>
> Use `-UserIds` to find activities performed **BY** a specific delegate user. Don't use `-UserIds` with the shared mailbox address, as this action doesn't return delegate activities performed **IN** the

shared mailbox. For activities performed in the shared mailbox by any user (including delegates), use the `-FreeText` parameter as shown in other sections.

Monitor email access activities

> **Note**
>
> This command requires Microsoft 365 E5 licensing to capture `MailItemsAccessed` operations.

To search for email access activities in shared mailboxes, run the following command:

PowerShell

```
Search-UnifiedAuditLog -StartDate 01/06/2020 -EndDate 01/20/2020 -UserIds
<user1,user2> -FreeText <shared-mailbox@domain.com> -Operations MailItemsAccessed -
ResultSize 200
```

Search for email forwarding configuration

To find who configured email forwarding on a shared mailbox, run the following command:

PowerShell

```
Search-UnifiedAuditLog -StartDate 01/06/2020 -EndDate 01/20/2020 -UserIds
<user1,user2> -FreeText <shared-mailbox@domain.com> -Operations Set-Mailbox -
ResultSize 200
```

Search for mailbox rule modifications

To search for inbox rule creation or modification activities, run the following command:

PowerShell

```
Search-UnifiedAuditLog -StartDate 01/06/2020 -EndDate 01/20/2020 -FreeText <shared-
mailbox@domain.com> -Operations New-InboxRule,Set-InboxRule -ResultSize 100
```

> **Tip**

For comprehensive mailbox rule investigation, see **Identify who modified mailbox rules** for detailed guidance on identifying who created, modified, or deleted mailbox rules that might be affecting email delivery.

What to do when searches return no results

If your audit log searches don't find shared mailbox activity records, try these steps to verify the auditing configuration.

Verify and enable mailbox auditing

1. Check if auditing is enabled at the organization level.

PowerShell

```
Get-OrganizationConfig | Select AuditDisabled
```

1. Check if auditing is enabled for the specific shared mailbox.

PowerShell

```
Get-Mailbox <shared-mailbox@domain.com> | Select AuditEnabled
```

1. Enable auditing if it's disabled.

PowerShell

```
Set-OrganizationConfig -AuditDisabled $false
Set-Mailbox <shared-mailbox@domain.com> -AuditEnabled $true
```

Advanced procedures

The following commands apply common shared mailbox fixes:

To enable sent items copying to shared mailbox (so sent emails appear in shared mailbox Sent Items), run the following command:

PowerShell

```
Set-Mailbox <shared-mailbox@domain.com> -MessageCopyForSentAsEnabled $true
```

To remove email forwarding from a shared mailbox, run the following command:

```
PowerShell
Set-Mailbox <shared-mailbox@domain.com> -ForwardingSmtpAddress $null
```

Quick reference

The following tables summarize common audit operations and key search parameters for shared mailbox investigations.

Common audit operations for shared mailboxes

The following table lists common audit operations you can use when investigating shared mailbox activity.

Expand table

Operation	Description	Investigation Focus
Create	Emails created or sent	Email creation activities
FolderBind	**Delegate folder access/browsing**	**Monitor who accessed specific folders**
HardDelete	Items permanently deleted	Permanent deletions from shared mailbox
MailItemsAccessed	Mailbox items accessed	Email access tracking (E5 required)
Move	Items moved between folders	Folder organization changes
New-InboxRule	Inbox rules created	Rule creation investigation
SendAs	Emails sent using delegate permissions	Identify who sent from shared mailbox
Set-Mailbox	Mailbox settings modified	Forwarding and configuration changes
SoftDelete	Items moved to Deleted Items folder	User deletions from shared mailbox

Key search parameters

The following table describes the main parameters used in the shared mailbox audit log search examples.

Expand table

Parameter	Description	Example
-FreeText	Activities performed in a specific mailbox	`<shared-mailbox@domain.com>`
-Operations	Filter by operation type	SoftDelete,HardDelete,SendAs
-ResultSize	Limit results (max 5000)	500 (standard), 1000 (comprehensive)
-StartDate/-EndDate	Define investigation timeframe	01/06/2020, 01/20/2020
-UserIds	Activities performed by a specific user	`<delegate@domain.com>`

> **Important**
>
> Replace `<shared-mailbox@domain.com>` and `<delegate@domain.com>` with actual email addresses before running commands.

Next steps

- Manage mailbox auditing: Adjust which shared mailbox actions are audited to capture the activities you need.
- Identify who deleted an email message or why an email is missing: Investigate specific email deletions discovered in shared mailbox activity.
- Export, configure, and view audit log records: Export your shared mailbox investigation results for evidence preservation.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 06/24/2026

Search the audit log for mailbox activities in specific mailboxes

This article explains how to search for mailbox activities in Microsoft Purview Audit for a specific mailbox, including shared mailboxes. It covers common investigation scenarios, such as when mailbox audit logs are missing or can't be found. It provides step-by-step instructions to verify auditing configuration, permissions, retention policies, and search methods.

Before you begin

To complete the steps in this article, you need:

- The email address of the affected mailbox.
- The license type assigned to the affected mailbox.
- Details about the operation you're searching for and the date and time it occurred.
- An understanding of how audit log searches are performed in Microsoft Purview Audit.

Check mailbox auditing status

> ⓘ **Note**
>
> Before searching for audit activity, verify that mailbox auditing is enabled for the affected user.

Complete the following steps to confirm that mailbox auditing is enabled:

1. In Exchange Online PowerShell, run the following command to check the organization-level audit configuration:

   PowerShell

   ```
   Get-AdminAuditLogConfig | FL UnifiedAudit*
   ```

   > ⓘ **Note**
   >
   > Run this command in Exchange Online PowerShell. In Security & Compliance PowerShell, the `UnifiedAuditLogIngestionEnabled` property always returns `False`,

even when auditing is enabled.

2. Check if mailbox audit logs are disabled at the tenant level:

```
PowerShell

Get-OrganizationConfig | FL AuditDisabled
```

3. At the mailbox level, `AuditEnabled` always returns `True`, even if auditing is disabled for a specific mailbox. To verify, use:

```
PowerShell

Get-Mailbox -Filter {AuditEnabled -eq "True"}
```

If you receive an error that the mailbox can't be found, auditing is disabled for that mailbox. For more information, see Verify the auditing status for your organization.

Verify administrator permissions

Ensure that you have the required permissions to search or export audit logs:

1. In the Microsoft Purview portal, go to **Settings** > **Roles and scopes** > **Role groups**.
2. Verify that your account is assigned to the **View-Only Audit Logs** or **Audit Logs** role.
3. In the Exchange admin center, use the **Audit Logs** role to enable or disable auditing and access audit cmdlets. Use the **View-Only Audit Logs** role to grant read-only access to audit logs.

For more information, see Assign permissions to search the audit log.

Check audit log retention policies

Audit log retention policies determine whether older activities are still available:

1. Determine when the activity occurred.
2. By default, audit logs are retained for 180 days.
3. If the affected user doesn't have an Audit (Premium) license and no audit retention policy is configured, audit logs older than 180 days aren't available.

For more information, see Get-UnifiedAuditLogRetentionPolicy.

Search for mailbox activities in Microsoft Purview Audit

When searching for activities, use the correct filters based on mailbox type.

User mailbox – recommended audit filters

- **Date and time range (UTC)**: Specify when the activity occurred. Audit timestamps are always in UTC.
- **Users**: Enter the affected mailbox.
- **Activities – operation names**: Select the specific operation you're looking for.
- **Activities – friendly names**: Alternatively, use the friendly activity names from the dropdown list.

Shared mailbox – recommended audit filters

- **Date and time range (UTC)**: Specify when the activity occurred.
- **Keywords**: Enter the primary SMTP address or Exchange GUID of the shared mailbox.

> ⓘ Note
>
> When searching for shared mailbox activities, enter the SMTP address under the **Keywords** field, not the **Users** field.

Shared mailbox auditing specifics

By default, the system doesn't audit all actions for shared mailboxes. The audited actions depend on the role of the person accessing the mailbox:

- **Owner**: The account associated with the mailbox.
- **Delegate**: A user with SendAs, SendOnBehalf, or FullAccess permissions to another mailbox. This role can include admins assigned FullAccess to a user's mailbox.
- **Admin**: Access via Microsoft eDiscovery tools in Microsoft Purview or Exchange Online, use of the Microsoft Exchange Server MAPI Editor, or access by an account impersonating another user (with the ApplicationImpersonation role).

Before searching, check the official documentation to confirm whether the action you're investigating is audited for the specific sign-in type. See Mailbox actions for user mailboxes and shared mailboxes.

Next steps

- Identify who deleted an email message or why an email is missing: Investigate specific email deletions found in your mailbox search results.
- Investigate shared mailbox activities using audit logs: Dig deeper into shared mailbox activities including delegate access and forwarding.
- Export, configure, and view audit log records: Export your mailbox activity results to CSV for offline analysis.

ⓘ **Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 04/14/2026

Search the audit log for events in Microsoft Teams

Applies to: Microsoft Teams

> **Important**
>
> The Microsoft Teams admin center is gradually replacing the Skype for Business admin center, and we're migrating Teams settings to it from the Microsoft 365 admin center. If a setting has been migrated, you'll see a notification and then be directed to the setting's location in the Teams admin center. For more information, see Manage Teams during the transition to the Teams admin center.

The audit log helps you investigate specific activities across Microsoft services. For Microsoft Teams, the audit log tracks activities such as:

- Team creation
- Team deletion
- Added channel
- Deleted channel
- Changed channel setting

For a complete list of Teams activities that are audited, see Teams activities in the audit log.

> **Note**
>
> Audit events from private channels are also logged as they are for teams and standard channels.

Turn on auditing in Teams

Before you can view audit data, you need to turn on auditing in the Microsoft Purview portal. For more information, see Turn auditing on or off.

> **Important**

Audit data is only available from the point at which you turn on auditing.

Retrieve Teams data from the audit log

To retrieve audit logs for Teams activities, use the Microsoft Purview portal. For step-by-step instructions, see Search the audit log.

> **Important**
>
> Audit data is only visible in the audit log if auditing is turned on.

The length of time that an audit record is retained and searchable in the audit log depends on your Microsoft 365 or Office 365 subscription, and specifically the type of license assigned to users. For more information, see the Security & Compliance Center service description.

Search the audit log effectively

Use the following tips when searching for Teams activities in the audit log.

- Select specific activities to search for by selecting the checkbox next to one or more activities. To cancel the selection, select the activity. Use the search box to display the activities that contain the keyword that you type.
- Select **Show results for all activities** in the **Activities** list to display events for activities run using cmdlets. If you know the name of the operation for these activities, type it in the search box to display the activity, then select it.
- Select **Clear all** to clear the current search criteria. The date range returns to the default of the last seven days.
- If 5,000 results are found, you can probably assume that there are more than 5,000 events that met the search criteria. Refine the search criteria and rerun the search to return fewer results, or export all the search results by selecting **Export** > **Download all results**. For step-by-step instructions to export audit logs, see Search the audit log.

For more information about using audit log search, see How to do an audit log search for Teams.

Teams activities

For a list of all events that are logged for user and admin activities in Teams in the Microsoft 365 audit log, see:

- Teams activities in the audit log
- Shifts app in Teams activities in the audit log
- Updates app in Teams activities in the audit log

Office 365 Management Activity API

You can use the Office 365 Management Activity API to get information about Teams events. For more information about the Management Activity API schema for Teams, see Teams schema.

Attribution in Teams audit logs

When you change membership in Teams (such as adding or deleting users) through Microsoft Entra ID, Microsoft 365 admin portal, or Microsoft 365 Groups Graph API, Teams audit messages and the General channel show an existing owner of the team as the initiator, not the actual initiator of the action. In these scenarios, check Microsoft Entra ID or Microsoft 365 Group audit logs to see the relevant information.

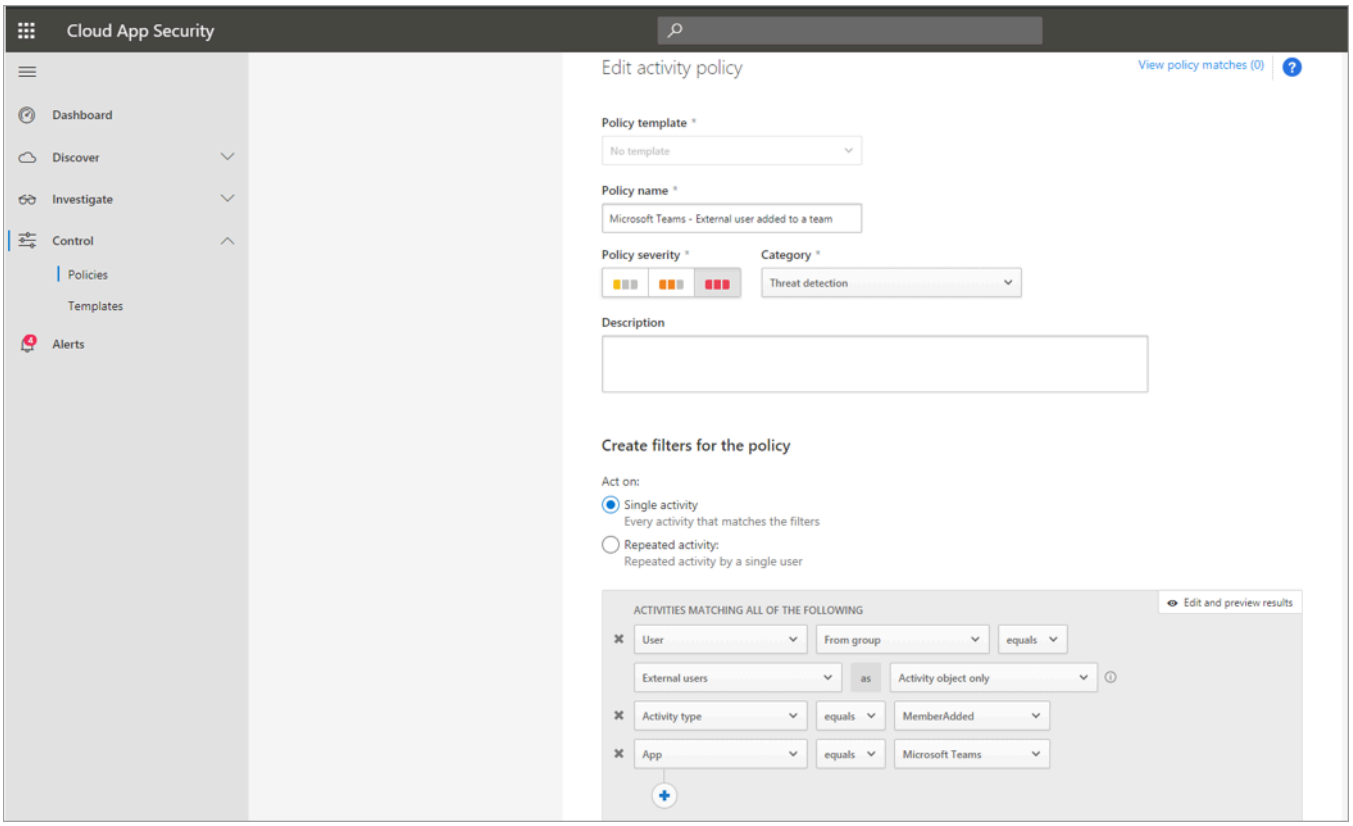
Use Defender for Cloud Apps to set activity policies

By using Microsoft Defender for Cloud Apps integration, you can set activity policies to enforce a wide range of automated processes by using the app provider's APIs. With these policies, you can monitor specific activities carried out by various users or follow unexpectedly high rates of one certain type of activity.

After you set an activity detection policy, it starts to generate alerts. Alerts only happen for activities that occur after you create the policy. Here are some example scenarios for how you can use activity policies in Defender for Cloud Apps to monitor Teams activities.

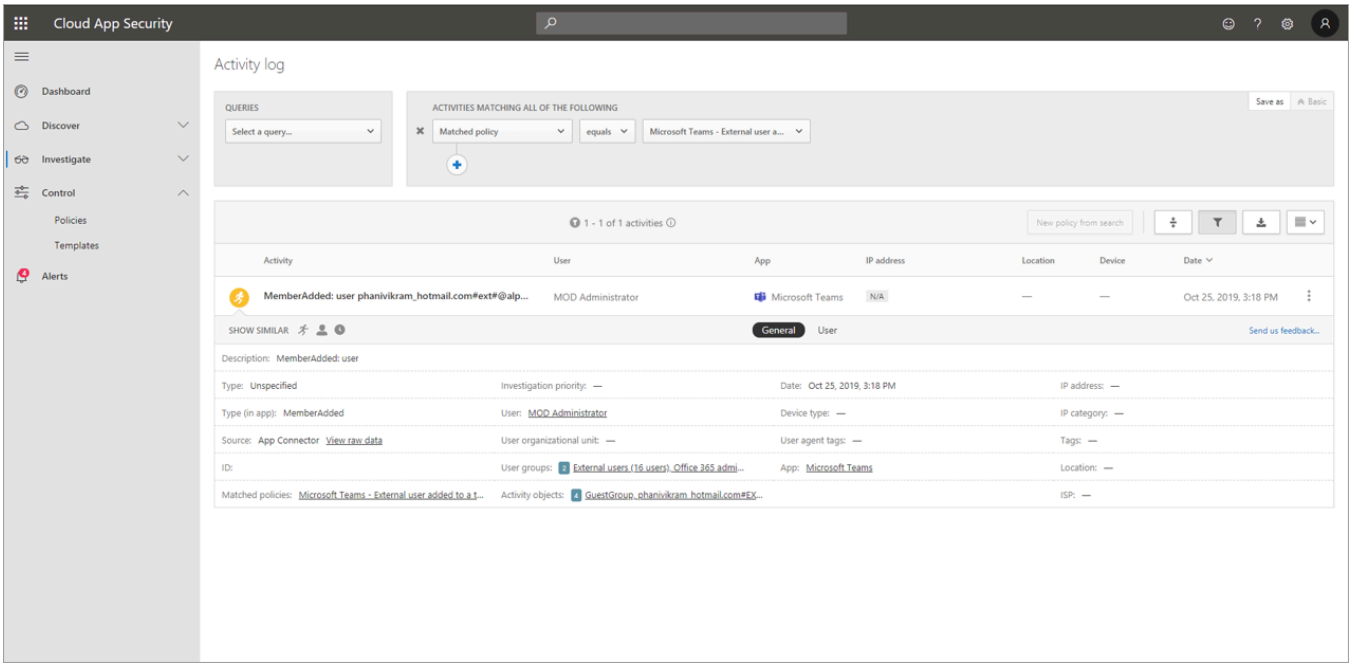
External user scenario

From a business perspective, you might want to keep an eye on the addition of external users to your Teams environment. If you enable external users, monitoring their presence is a good idea. You can use Defender for Cloud Apps to identify potential threats.



The screenshot of this policy to monitor adding external users shows you how to name the policy, set the severity according to your business needs, set it as (in this case) a single activity, and then establish the parameters that specifically monitor only the addition of noninternal users and limit this activity to Teams.

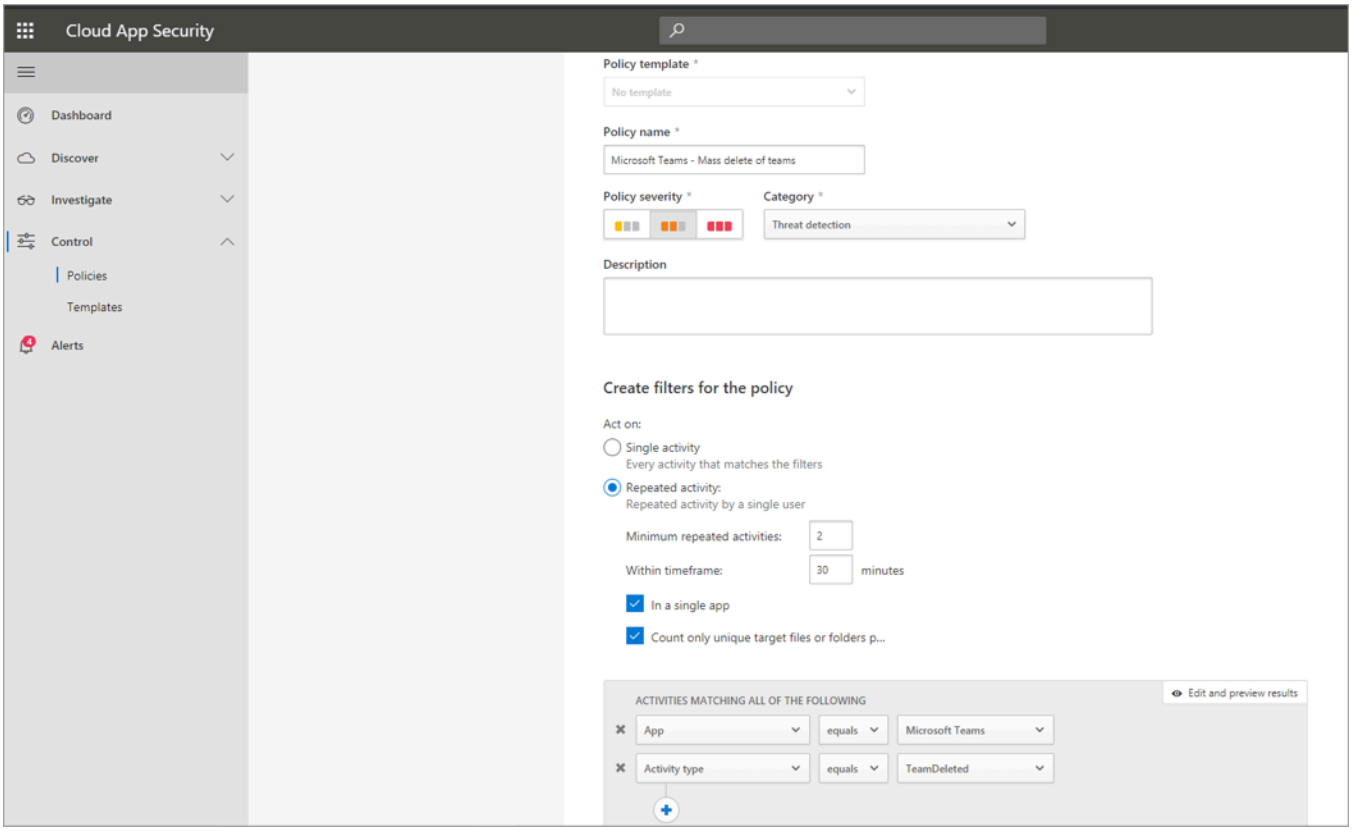
You can view the results from this policy in the activity log:



Here you can review matches to the policy you set, make any adjustments as needed, or export the results to use elsewhere.

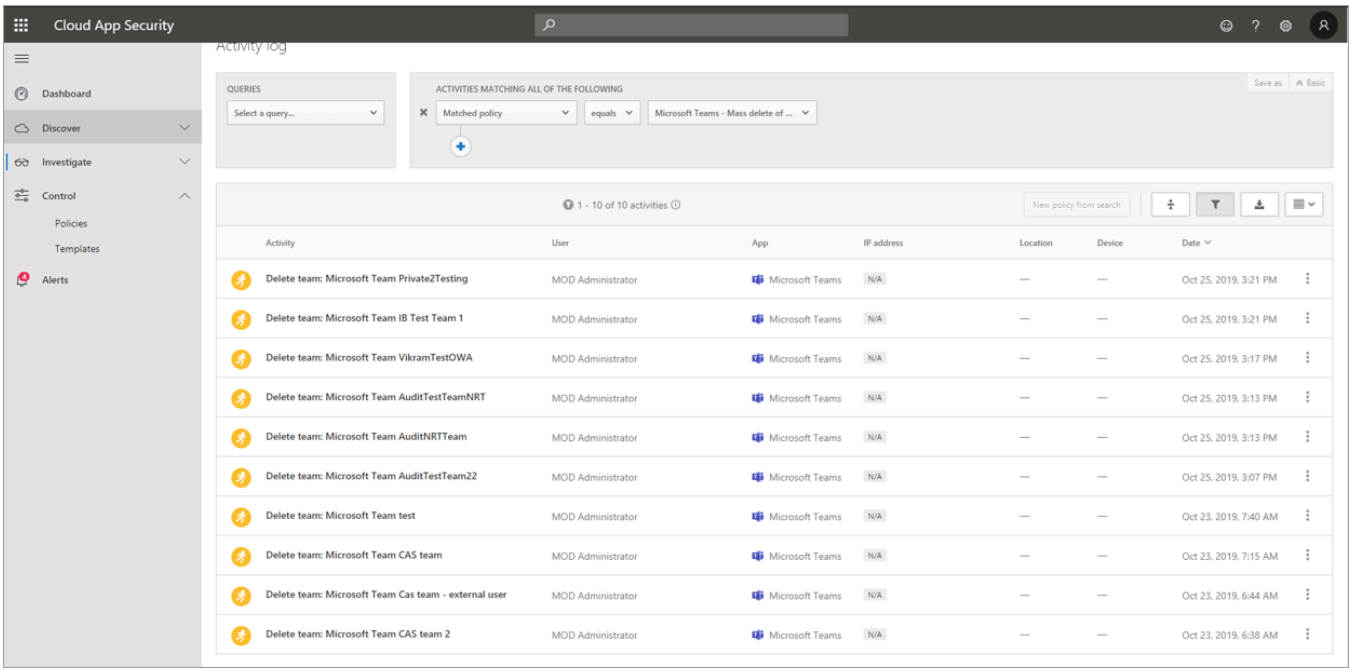
Mass delete scenario

As mentioned earlier, you can monitor deletion scenarios. You can create a policy that monitors mass deletion of Teams sites. In this example, you set up an alert-based policy to detect mass deletion of teams within 30 minutes.



As the screenshot shows, you can set many different parameters for this policy to monitor Teams deletions, including severity, single or repeated action, and parameters limiting this to Teams and site deletion. You can set these parameters independently of a template, or you can use a template, depending on your organizational needs.

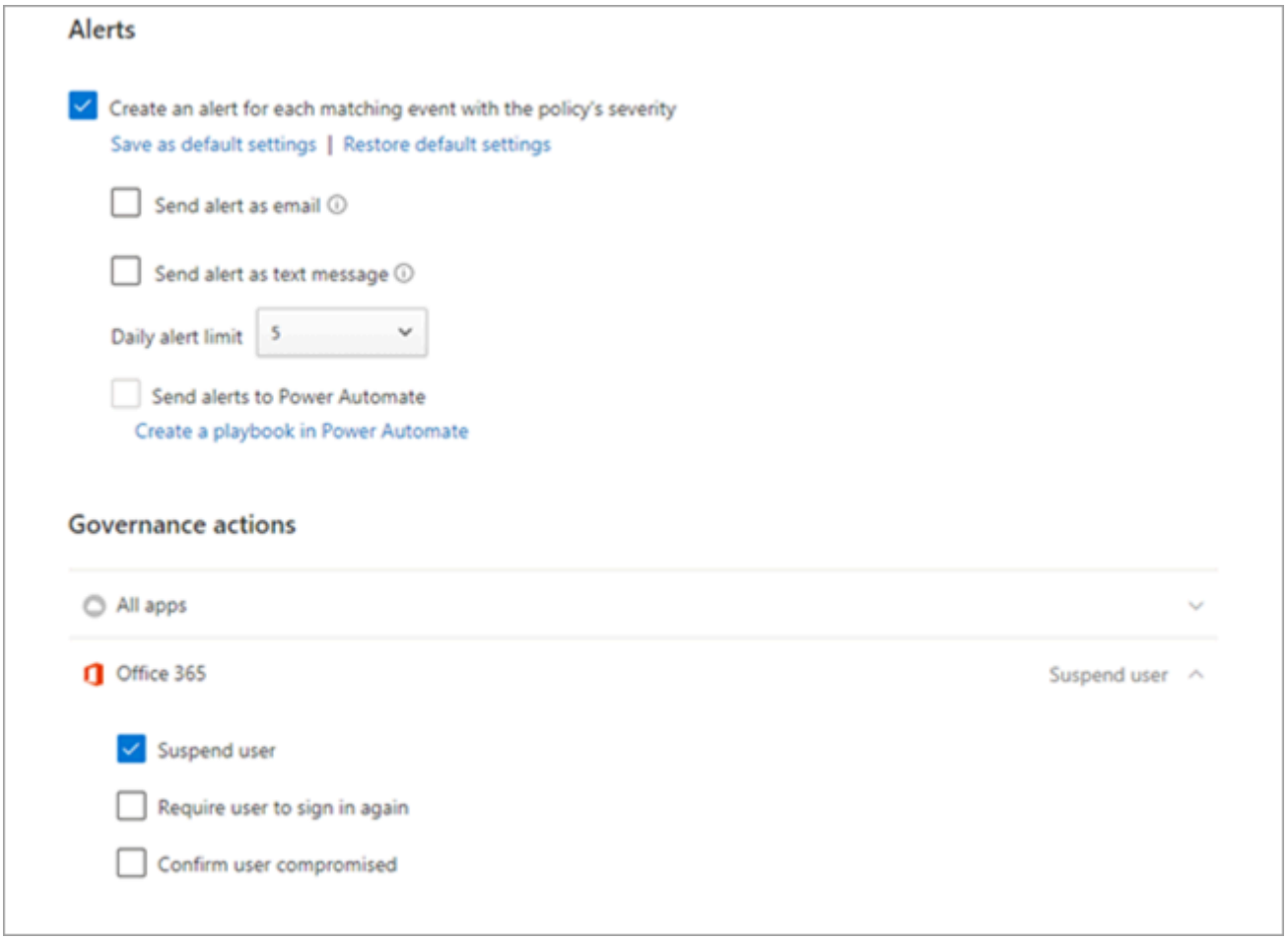
After you create a policy that works for your business, you can review the results in the activity log as events are triggered:



You can filter down to the policy you set to see the results. If the results you get in the activity log aren't satisfactory (maybe you see lots of results, or nothing at all), you can fine-tune the query to make it more relevant to what you need.

Alert and governance scenario

You can set alerts and send emails to admins and other users when an activity policy is triggered. You can set automated governance actions such as suspending a user or requiring a user to sign in again. This example shows how a user account can be suspended when an activity policy is triggered and determines a user deleted two or more teams in 30 minutes.



Use Defender for Cloud Apps to set anomaly detection policies

Anomaly detection policies in Defender for Cloud Apps provide out-of-the-box user and entity behavioral analytics (UEBA) and machine learning (ML) so that you can immediately run advanced threat detection across your cloud environment. Because they're automatically enabled, the new anomaly detection policies provide immediate results by providing immediate detections, targeting numerous behavioral anomalies across your users and the machines and devices connected to your network. Additionally, the new policies expose more data from the Defender for Cloud Apps detection engine, to help you speed up the investigation process and contain ongoing threats.

Teams events integration into anomaly detection policies is currently in development. You can set up anomaly detection policies for other Microsoft 365 products and take action on users who match those policies.

Next steps

- Export, configure, and view audit log records: Export your Teams audit results for analysis or compliance documentation.

- Use a PowerShell script to search the audit log: Automate recurring Teams activity searches using PowerShell.
- Manage audit log retention policies: Create retention policies to preserve Teams audit records beyond the default period.

ⓘ **Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 04/14/2026

Use sharing auditing in the audit log

Sharing is a key activity in SharePoint Online and OneDrive for Business, and organizations widely use it. Administrators can use sharing auditing in the audit log to determine how sharing is used in their organization.

SharePoint sharing schema reference

Sharing events (not including events related to sharing policy and sharing links) differ from file- and folder-related events in one primary way: one user performs an action that affects another user. For example, when a resource User A gives User B access to a file. In this example, User A is the *acting user* and User B is the *target user*. In the SharePoint File schema, the acting user's action only affects the file itself. When User A opens a file, the only information needed in the **FileAccessed** event is the acting user. To address this difference, there's a separate schema, called the *SharePoint Sharing schema* that captures more information about sharing events. This schema ensures that administrators have visibility into who shared a resource and the user the resource was shared with.

The Sharing schema provides two additional fields in an audit record related to sharing events:

- **TargetUserOrGroupType:** Identifies whether the target user or group is a Member, Guest, SharePointGroup, SecurityGroup, or Partner.
- **TargetUserOrGroupName:** Stores the UPN or name of the target user or group that a resource was shared with (User B in the previous example).

These two fields, in addition to other properties from the audit log schema such as User, Operation, and Date tell the full story about *which* user shared *what* resource with *whom* and *when*.

Another schema property is important to the sharing story. When you export audit log search results, the **AuditData** column in the exported CSV file stores information about sharing events. For example, when a user shares a site with another user, sharing the site adds the target user to a SharePoint group. The **AuditData** column captures this information to provide context for administrators. To parse the **AuditData** column into individual properties, use the JSON transform feature in the Power Query Editor in Excel as described in Use the PowerQuery Editor to format the exported audit log.

SharePoint sharing events

Sharing occurs when a user (the *acting* user) shares a resource with another user (the *target* user). Audit records related to sharing a resource with an external user (a user who is outside of your organization and doesn't have a guest account in your organization's Microsoft Entra ID) are identified by the following events, which the audit log records:

- **SharingInvitationCreated:** A user in your organization tries to share a resource (likely a site) with an external user. This event results in an external sharing invitation sent to the target user. The invitation grants no access to the resource at this point.
- **SharingInvitationAccepted:** The external user accepts the sharing invitation sent by the acting user and now has access to the resource.
- **AnonymousLinkCreated:** An anonymous link (also called an "Anyone" link) is created for a resource. Because an anonymous link can be created and then copied, it's reasonable to assume that any document that has an anonymous link is shared with a target user.
- **AnonymousLinkUsed:** This event is logged when an anonymous link is used to access a resource.
- **SecureLinkCreated:** A user creates a "specific people link" to share a resource with a specific person. This target user might be someone who is external to your organization. The person that the resource is shared with is identified in the audit record for the **AddedToSecureLink** event. The time stamps for these two events are nearly identical.
- **AddedToSecureLink:** A user is added to a specific people link. Use the **TargetUserOrGroupName** field in this event to identify the user added to the corresponding specific people link. This target user might be someone who is external to your organization.

How sharing auditing works

When a user (the acting user) wants to share a resource with another user (the target user), SharePoint (or OneDrive for Business) first checks if the email address of the target user is already associated with a user account in the organization's directory. If the target user is in the directory (and has a corresponding guest user account), SharePoint takes the following actions:

- Immediately assigns the target user permissions to access the resource by adding the target user to the appropriate SharePoint group, and logs an **AddedToGroup** event.
- Sends a sharing notification to the email address of the target user.
- Logs a **SharingSet** event. This event has a friendly name of "Shared file, folder, or site" and appears under **Sharing and access request activities** in the activities picker of the audit log

search tool.

If a user account for the target user isn't in the directory, SharePoint takes the following actions:

- Logs one of the following events, based on how the resource is shared:
  - **AnonymousLinkCreated**
  - **SecureLinkCreated**
  - **AddedToSecureLink**
  - **SharingInvitationCreated** (this event is logged only when the shared resource is a site)
- When the target user accepts the sharing invitation (by clicking the link in the invitation), SharePoint logs a **SharingInvitationAccepted** event and assigns the target user permissions to access the resource. If the target user is sent an anonymous link, the **AnonymousLinkUsed** event is logged after the target user uses the link to access the resource. For secure links, a **FileAccessed** event is logged when an external user uses the link to access the resource.

Additional information about the target user is also logged, such as the identity of the user the invitation is to and the user who accepts the invitation. In some cases, these users (or email addresses) can be different.

How to identify resources shared with external users

A common requirement for administrators is creating a list of all resources that administrators shared with users outside of the organization. By using sharing auditing in Office 365, administrators can generate this list. Complete the following steps.

Step 1: Search for sharing events and export the results to a CSV file

To search the audit log for sharing events, complete the following steps. For more information including the required permissions, see Search the audit log.

1. Sign in to the Microsoft Purview portal.
2. Select the **Audit** solution card. If the **Audit** solution card isn't displayed, select **View all solutions** and then select **Audit** from the **Core** section.
3. On the **Search** page and under **Activities - friendly names**, select **Sharing and access request activities** to search for sharing-related events.

4. Select a date and time range to find the sharing events that occurred within that period.
5. Select **Search** to run the search.
6. When the search finishes and displays the results, select **Export results** > **Download all results**.

   After you select the export option, a message at the bottom of the window prompts you to open or save the CSV file.
7. Select **Save** > **Save as** and save the CSV file to a folder on your local computer.

Step 2: Use the PowerQuery Editor to format the exported audit log

Use the JSON transform feature in the Power Query Editor in Excel to split each property in the **AuditData** column (which consists of a multi-property JSON object) into its own column. This feature lets you filter columns to view records related to sharing.

For step-by-step instructions, see "Step 2: Format the exported audit log using the Power Query Editor" in Export, configure, and view audit log records.

Step 3: Filter the CSV file for resources shared with external users

In this step, you filter the CSV file for the different sharing-related events that the SharePoint sharing events section describes. Alternatively, you can filter the **TargetUserOrGroupType** column to display all records where the value of this property is **Guest**.

After formatting the exported audit log with Power Query Editor in Excel so that each **AuditData** property has its own column, complete the following steps:

1. Open the formatted Excel file that contains the exported audit log with **AuditData** properties split into individual columns by the Power Query transformation.
2. On the **Home** tab, select **Sort & Filter**, then select **Filter**.
3. In the **Sort & Filter** dropdown list on the **Operations** column, clear all selections, then select one or more of the following sharing-related events and select **Ok**.

   - **SharingInvitationCreated**
   - **AnonymousLinkCreated**

- **SecureLinkCreated**
- **AddedToSecureLink**

Excel displays the rows for the events you selected.

4. Go to the column named **TargetUserOrGroupType** and select it.

5. In the **Sort & Filter** dropdown list, clear all selections, then select **TargetUserOrGroupType:Guest**, and select **Ok**.

   Now Excel displays the rows for sharing events **and** where the target user is outside of your organization, because external users are identified by the value **TargetUserOrGroupType:Guest**.

> **Tip**
>
> For the audit records that are displayed, the **ObjectId** column identifies the resource that you shared with the target user. For example, `ObjectId:https:\/\/contoso-my.sharepoint.com\/personal\/sarad_contoso_com\/Documents\/Southwater Proposal.docx`.

Next steps

- Export, configure, and view audit log records: Format and analyze your sharing audit exports using Power Query in Excel.
- Search the audit log: Broaden your investigation to search for other activity types related to the shared resources.
- Use a PowerShell script to search the audit log: Automate recurring sharing audits using PowerShell for ongoing monitoring.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/24/2026

Search the audit log for deleted mailbox items

Use this article to investigate whether a user deleted email items from a mailbox in Exchange Online. It provides step-by-step guidance on checking audit settings, verifying permissions, reviewing retention policies, and searching audit logs in Microsoft Purview. By following the procedure in this article, you can determine if deletion activity occurred and confirm whether the audit data is still available.

Before you start, review the prerequisites to ensure you have the required permissions, information, and access.

Before you begin

To complete this procedure, you need:

- The email address of the affected mailbox.
- The license type assigned to the affected mailbox.
- The approximate date and time when the operation occurred.

You must also have access to Exchange Online PowerShell and the Microsoft Purview portal with your administrator credentials.

Check mailbox auditing status

> **Note**
>
> Before searching for audit activity, verify that mailbox auditing is enabled for the affected user.

Complete the following steps to confirm that mailbox auditing is enabled:

1. In Exchange Online PowerShell, run the following command to check the organization-level audit configuration:

   ```
   PowerShell
   ```

```
Get-AdminAuditLogConfig | FL UnifiedAudit*
```

> **Note**
>
> Run this command in Exchange Online PowerShell. In Security & Compliance PowerShell, the `UnifiedAuditLogIngestionEnabled` property always returns `False`, even when auditing is enabled.

2. Check if mailbox audit logs are disabled at the tenant level:

PowerShell

```
Get-OrganizationConfig | FL AuditDisabled
```

3. At the mailbox level, `AuditEnabled` always returns `True`, even if auditing is disabled for a specific mailbox. To verify, use:

PowerShell

```
Get-Mailbox -Filter {AuditEnabled -eq "True"}
```

If you receive an error that the mailbox can't be found, auditing is disabled for that mailbox. For more information, see Verify the auditing status for your organization.

Verify administrator permissions

Ensure that you have the required permissions to search or export audit logs:

1. In the Microsoft Purview portal, go to **Settings** > **Roles and scopes** > **Role groups**.
2. Verify that your account is assigned to the **View-Only Audit Logs** or **Audit Logs** role.
3. In the Exchange admin center, use the **Audit Logs** role to enable or disable auditing and access audit cmdlets. Use the **View-Only Audit Logs** role to grant read-only access to audit logs.

For more information, see Assign permissions to search the audit log.

Check audit log retention policies

Audit log retention policies determine whether older activities are still available:

1. Determine when the activity occurred.
2. By default, audit logs are retained for 180 days.
3. If the affected user doesn't have an Audit (Premium) license and no audit retention policy is configured, audit logs older than 180 days aren't available.

For more information, see Get-UnifiedAuditLogRetentionPolicy.

Search for deletion activity in Microsoft Purview Audit

When searching for audit data, use filters that match the suspected deletion time and the mailbox deletion activity being investigated.

Option 1: Search by user and activities

Use this option to search for deletion activity by specifying the mailbox user and relevant audit operations.

- **Date and time range (UTC)**: Specify when the deletion occurred. Audit timestamps are always in UTC.
- **Users**: Enter the affected mailbox.
- **Activities – operation names**: Select Move, MoveToDeletedItems, Create, SoftDelete, and HardDelete.
- **Activities – friendly names**: If available, use the friendly names in the dropdown list.

Option 2: Search by activities and keywords

Use this option when you want to search by audit activity and identify the mailbox with a keyword such as its SMTP address or ExchangeGUID.

- **Date and time range (UTC)**: Specify when the deletion occurred.
- **Activities – operation names**: Select Move, MoveToDeletedItems, Create, SoftDelete, and HardDelete.
- **Keywords**: Enter the primary SMTP address or ExchangeGUID of the mailbox.

Option 3: Search by message ID prefix

Use this option if you know the message ID prefix and want to narrow the audit search to a specific message.

- **Date and time range (UTC)**: Specify when the deletion occurred.

- **Keywords**: Enter the prefix of the email messageID.

Next steps

After you complete your audit log search, use these resources to continue your investigation:

- Identify who deleted an email message or why an email is missing: Determine the root cause of deletions, including retention policies, rules, and admin actions.
- Export, configure, and view audit log records: Export your deletion search results for evidence collection or reporting.
- Use the MailItemsAccessed audit event to investigate compromised accounts: Check whether email deletions are linked to a compromised account.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/24/2026

Encrypted message portal activity log by Microsoft Purview Advanced Message Encryption

You can access logs for encrypted messages through the encrypted message portal. With these logs, your organization can determine when external recipients read and forward messages. To ensure availability of logs for external recipients, apply a custom branding template to protected emails that your organization sends to external recipients. This template enforces a portal experience. For more information, see Add your organization's brand to your encrypted messages.

Enable message access audit logs in PowerShell

You can enable the access log by using Exchange Online PowerShell. The *EnablePortalTrackingLogs* parameter of the **Set-IrmConfiguration** cmdlet specifies whether to enable audit logs for accessing the encrypted message portal. Valid values are:

- $true: Turns on the audit feature.
- $false: Turns off the audit feature.

Example:

PowerShell

```
Set-IrmConfiguration -EnablePortalTrackingLogs $true
```

For more information, see Set-IRMConfiguration (ExchangePowerShell).

Message access audit information

The access log contains entries for messages sent through the encrypted message portal for the following types of activity:

- External user sign-in timestamp and authentication method
- External user reading messages or attachments
- Attachment download
- Mail replies and forwards

For more information on the message access log schema, see Audit log activities.

Search for events in the message access logs

Complete the following steps to get started with search:

1. Sign in to the Microsoft Purview portal.
2. Select the **Audit** solution card. If the **Audit** solution card isn't displayed, select **View all solutions** and then select **Audit** from the **Core** section.
3. Under **Search**, select the drop-down for **Activities** and type encrypted message portal activities.
4. Under encrypted message portal activities, select the event types to use in the search. Set the date range for the search (default is the previous week), you can also optionally add a particular user in your organization for the search. When ready, select **Search**.
5. Select an event from the list to view the audit properties.

Next steps

- Search the audit log: Search for broader activity related to the senders or recipients identified in encrypted message logs.
- Export, configure, and view audit log records: Export your encrypted message activity results for compliance reporting.

ⓘ **Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 04/14/2026

Audit log activities

The tables in this article describe the activities recorded in the Microsoft 365 audit log. You can search for these activities by searching the audit log in the Microsoft Purview portal.

> **Tip**
>
> Select one of the links in the **In this article** list at the top of this article to go directly to a specific product table.

Audit logging is turned on by default for Microsoft 365 organizations. If auditing isn't turned on for your organization, a banner appears that prompts you to start recording user and admin activity. For instructions, see Turn on auditing.

These tables group related activities or the activities from a specific service. The tables include the friendly name that's displayed in the **Activities** drop-down list (or that are available in PowerShell) and the name of the corresponding operation that appears in the detailed information of an audit record and in the CSV file when you export the search results. For descriptions of the detailed information, see Audit log detailed properties.

> **Important**
>
> Some operation names listed in the **Operation** column in the following tables contain a period (`.`). You must include the period in the operation name if you specify the operation in a PowerShell command when searching the audit log, creating audit retention policies, creating alert policies, or creating activity alerts. Also be sure to use double quotation marks (`" "`) to contain the operation name.

Agent 365 activities

The following table lists Agent 365 activities that are logged when an agent is invoked or performs any actions or tool calls.

Expand table

Friendly name	Operation	Description
Executed AI tool	AIExecuteTool	Agent executed a tool call.
Invoked AI agent	AIInvokeAgent	AI agent is invoked by a user, agent, or event.
Made AI inference call	AIInferenceCall	AI agent leveraged an AI model to produce an answer or determine next steps.
Applied AI guardrail	AIGuardrail	A guardrail was applied to an AI agent's request or response to enforce safety or compliance policies.

Application administration activities

The following table lists application admin activities that Microsoft 365 audit logs when an admin adds or changes an application registered in Microsoft Entra ID. You must register any application that relies on Microsoft Entra ID for authentication in the directory.

Expand table

Friendly name	Operation	Description
Added credentials to a service principal	Add service principal credentials.	Credentials were added to a service principal in Microsoft Entra ID. A service principal represents an application in the directory.
Added delegation entry	Add delegation entry.	An authentication permission was created or granted to an application in Microsoft Entra ID.
Added service principal	Add service principal.	An application was registered in Microsoft Entra ID. A service principal represents an application in the directory.
Removed a service principal from the directory	Remove service principal.	An application was deleted or unregistered from Microsoft Entra ID. A service principal represents an application in the directory.
Removed credentials from a service principal	Remove service principal credentials.	Credentials were removed from a service principal in Microsoft Entra ID. A service principal represents an application in the directory.
Removed delegation entry	Remove delegation entry.	An authentication permission was removed from an application in Microsoft Entra ID.
Set delegation entry	Set delegation entry.	An authentication permission was updated for an application in Microsoft Entra ID.

Briefing email activities

The following table lists the activities in Briefing email that the Microsoft 365 audit log records. For more information about Briefing email, see:

- Overview of Briefing email
- Configure Briefing email

Expand table

Friendly name	Operation	Description
Updated organization privacy settings	UpdatedOrganizationBriefingSettings	Admin updates the organization privacy settings for Briefing email.
Updated user privacy settings	UpdatedUserBriefingSettings	Admin updates the user privacy settings for Briefing email.

Communication compliance activities

The following table lists communication compliance activities that the Microsoft 365 audit log records. For more information, see Learn about Microsoft Purview Communication Compliance.

> **Note**
>
> You can see these activities when you use the Search-UnifiedAuditLog PowerShell cmdlet. You can't see these activities in the **Activities** drop-down list.

Expand table

Friendly name	Operation	Description
Policy match	SupervisionRuleMatch	A user sent a message that matches a policy's condition.
Policy update	SupervisionPolicyCreated, SupervisionPolicyUpdated, SupervisionPolicyDeleted	A communication compliance administrator performed a policy update.
Tag applied to messages	SupervisoryReviewTag	Tags are applied to messages or messages are resolved.

Compliance Manager activities

The following table lists the operations and activities that the audit log records when an admin manages settings in Compliance Manager. For more information, see Learn about Compliance Manager.

Expand table

Friendly name	Operation	Description
Roles change	ComplianceManagerRolesChange	An admin changed the roles for users.
Tenant automation level change	ComplianceManagerAutomationLevelChange	An admin changed the automation level for the organization across all actions.
Testing source automation change	ComplianceManagerAutomationChange	An admin changed the testing source automation settings.

Content explorer activities

The following table lists the activities in Content explorer recorded in the Microsoft 365 audit log. You access content explorer on the Data classifications tool in the Microsoft Purview portal. For more information, see Using data classification content explorer.

Expand table

Friendly name	Operation	Description
Accessed item	LabelContentExplorerAccessedItem	An admin (or a user who's a member of the Content Explorer Content Viewer role group) used content explorer to view an email message or SharePoint and OneDrive document.
AccessedAggregates	AccessedAggregates	An admin used content explorer to navigate between different aggregates.
AccessedContentList	AccessedContentList	When admin Clicked on any site/user in Content Explorer to get list of items.
AccessedItemDownload	AccessedItemDownload	An admin used content explorer to download an item.
AccessedItemPreview	AccessedItemPreview	An admin used content explorer to preview an item.

Friendly name	Operation	Description
AccessedSiteList	AccessedSiteList	An admin used content explorer to navigate to SharePoint and OneDrive to get the site list.
AccessedUserList	AccessedUserList	An admin used content explorer to navigate to Exchange/Teams workload to get the user list.
SearchedContentList	SearchedContentList	When admin Searched for a file in Content Explorer.
SearchedSiteList	SearchedSiteList	An admin used content explorer to search for a site.
SearchedUserList	SearchedUserList	An admin used content explorer to search for a user.

Data Security Investigations activities

Data Security Investigations provides powerful investigation tools that might require the ability to audit activities to ensure secure and approved use of the solution. To meet these requirements, the unified audit log records Data Security Investigations activities.

The following table lists Data Security Investigations activities that the Microsoft 365 audit log records. For more information, see Learn about Microsoft Purview Data Security Investigations.

Expand table

Friendly name	Operation	Description
Canceled DSI sample view	DSISampleViewCancelled	A user canceled a samples view.
Canceled DSI statistics view	DSIStatisticsViewCancelled	A user canceled a statistics job.
Created DSI Investigation	CreatedDSIInvestigation	A user created an investigation.
Deleted Data Security Investigations capacity	DSICapacityDeleted	A user deleted the Data Security Investigation capacity.
Deleted DSI Investigation	DeletedDSIInvestigation	A user deleted an investigation.

Friendly name	Operation	Description
DSI AI feedback provided	DSIAIFeedbackProvided	A user provided AI feedback.
DSI investigation list viewed	DSIInvestigationListViewed	A user viewed the list of investigations.
DSI search added	DSIPurviewSearchAdded	A user added a search.
DSI search updated	DSIPurviewSearchUpdated	A user updated a search.
DSIProbeJobResultsViewed	DSIProbeJobResultsViewed	A user viewed the results of examination.
Get DSI Investigation	GetDSIInvestigation	A user viewed an investigation.
Get DSI search	GetSearchDSIInvestigation	A user viewed a search.
Investigation created from Defender XDR	DSIInvestigationCreatedFromXDR	A user created an investigation from Defender XDR.
Investigation created from Purview IRM	DSIInvestigationCreatedFromIRM	A user created an investigation from Microsoft Purview Insider Risk Management.
Item added to mitigation	DSIItemAddedToMitigation	A user added an item to an investigation's mitigation plan.
Item removed from mitigation	DSIItemRemovedFromMitigation	A user removed an item from an investigation's mitigation plan.
Mitigation plan list viewed	DSIMitigationPlanListViewed	A user viewed the mitigation plan list.
Started categorization job	DSIInvestigationCategorizationJobSubmitted	A user started a categorization job.
Started probing job	DSIProbeJobSubmitted	A user started an examination job.
Started purge job	DSIPurgeStarted	A user started a purge job.

Friendly name	Operation	Description
Started vectorization job	DSIInvestigationVectorizationJobSubmitted	A user stated a vectorization job.
Submitted DSI search add to evidence set job	DSIPurviewSearchAddToEvidenceSetJobSubmitted	A user added searched data to an investigation scope.
Submitted DSI search sample job	DSIPurviewSearchSampleJobSubmitted	A user started a samples job.
Submitted DSI search statistics job	DSIPurviewSearchStatisticsJobSubmitted	A user started a statistics job.
Updated Data Security Investigations capacity	DSICapacityUpdated	A user updated the Data Security Investigation capacity.
Updated DSI Investigation	UpdatedDSIInvestigation	A user updated an investigation.
Updated DSI Investigation members	DSIInvestigationMembersUpdated	A user updated an investigation's members.
Updated status on item from mitigation plan	DSIMitigationItemStatusUpdated	A user updated the status of an item on an investigation's mitigation plan.
Uploaded file for DSI search	DSIPurviewSearchUploadFile	A user uploaded a file to search.
Vector search started	DSIVectorSearchStarted	A user ran a vector search.
Viewed data processing errors for a DSI evidence set	DSIInvestigationSettingsUpdated	A user updated investigation settings.
Viewed Data Security Investigations capacity	DSICapacityViewed	A user viewed the Data Security Investigation capacity.
Viewed default DSI investigation settings	DSIInvestigationSettingsDefaultViewed	A user viewed investigation settings.
Viewed DSI evidence set document	DSIEvidenceSetDocumentViewed	A user viewed a document in an investigation.

Friendly name	Operation	Description
Viewed DSI investigation settings	DSIInvestigationSettingsViewed	A user viewed investigation settings.
Viewed DSI sample results	DSISampleResultsViewed	A user viewed sample results.
Viewed DSI sample status	DSISampleStatusViewed	A user viewed the status of a samples job.
Viewed DSI statistics results	DSIStatisticsResultsViewed	A user viewed search statistics.
Viewed individual activity	DSISingleActivityViewed	A user viewed an activity in an investigation.
Viewed investigation activities list	DSIActivitiesViewed	A user viewed the investigation activities.
Viewed investigation evidence set	DSIEvidenceSetViewed	A user viewed the list of items in the investigation.
Viewed item from mitigation plan	DSIItemViewedFromMitigation	A user viewed an item from an investigation's mitigation plan.

Data Security Posture Agent activities

The following table lists the activities for the Data Security Posture Agent in the Microsoft 365 audit log.

Expand table

Friendly name	Operation	Description
Agent Ran	AgentRan	User ran the Data Security Posture Agent.
Agent re-ran	Agent Reran	User reran the Data Security Posture Agent.
Agent result viewed	AgentResultViewed	User viewed the results of a Data Security Posture Agent run.
Agent run cancelled	AgentRunCancelled	User canceled a run for Data Security Posture Agent.
Agent run deleted	AgentRunDeleted	User deleted a run for Data Security Posture Agent.

Directory administration activities

The following table lists Microsoft Entra directory and domain-related activities that are logged when an administrator manages their organization in the Microsoft 365 admin center or in the Azure management portal.

Expand table

Friendly name	Operation	Description
Added a partner to the directory	Add partner to company.	Added a partner (delegated administrator) to your organization.
Added domain to company	Add domain to company.	Added a domain to your organization.
Removed a partner from the directory	Remove partner from company.	Removed a partner (delegated administrator) from your organization.
Removed domain from company	Remove domain from company.	Removed a domain from your organization.
Set company information	Set company information.	Updated the company information for your organization. Includes email addresses for subscription-related email that Microsoft 365 sends, and technical notifications about Microsoft 365 services.
Set domain authentication	Set domain authentication.	Changed the domain authentication setting for your organization.
Set password policy	Set password policy.	Changed the length and character constraints for user passwords in your organization.
Turned on Microsoft Entra ID Sync	Set DirSyncEnabled flag.	Set the property that enables a directory for Microsoft Entra ID Sync.
Updated domain	Update domain.	Updated the settings of a domain in your organization.
Updated the federation settings for a domain	Set federation settings on domain.	Changed the federation (external sharing) settings for your organization.
Verified domain	Verify domain.	Verified that your organization was the owner of a domain.
Verified email verified domain	Verify email verified domain.	Used email verification to verify that your organization is the owner of a domain.

Disposition review activities

The following table lists the activities a disposition reviewer took when an item reached the end of its configured retention period, or when an item automatically moved to the next disposition stage or was permanently deleted as a result of autoapproval.

Expand table

Friendly name	Operation	Description
Added reviewers	AddReviewer	A disposition reviewer added one or more other users to the current disposition review stage.
Approved disposal	ApproveDisposal	For manual approval: A disposition reviewer approved the disposition of the item to move it to the next disposition stage. If the item was in the only or final stage of disposition review, the disposition approval marked the item as eligible for permanent deletion. For autoapproval: No manual action was taken within the configured autoapproval time period so the item automatically moved to the next disposition stage. If the item was in the only or final stage of disposition review, the item automatically became eligible for permanent deletion.
Extended retention period	ExtendRetention	A disposition reviewer extended the retention period of the item.
Relabeled item	RelabelItem	A disposition reviewer relabeled the retention label.

Dragon Copilot admin activities

The following table lists administrative activities for Dragon Copilot recorded in the Microsoft 365 audit log. These events are logged when an administrator changes Dragon Copilot configuration. These events include changes in environments, products, Electronic Health Record (EHR) connectors, billing plans, groups, role assignments, settings, and extensions. If your organization uses Dragon Copilot, you can search the audit log for these activities by using the **Activities** picker list. For more information about Dragon Copilot, see the Dragon Copilot documentation.

You can also search for Dragon Copilot admin activities by running the `Search-UnifiedAuditLog -RecordType DragonCopilotAdmin` command in Exchange Online PowerShell.

Expand table

Friendly name	Operation	Description
Accepted terms	AcceptedTerms	Terms of service were accepted.
Activated subscription	ActivatedSubscription	A subscription was activated.
Added group member	AddedGroupMember	A user was added to a group.
Canceled deprovision	CanceledDeprovision	A previously queued deprovision was cancelled.
Created billing plan	CreatedBillingPlan	A billing plan was created.
Created child organization	CreatedChildOrganization	A child organization was created.
Created data export configuration	CreatedDataExportConfig	A data export configuration was created.
Created EHR connector	CreatedEhrConnector	An EHR connector was created.
Created EHR instance	CreatedEhrInstance	An EHR integration instance was created.
Created EHR user	CreatedEhrUser	An EHR user mapping was created.
Created environment	CreatedEnvironment	A new Dragon Copilot environment was created.
Created extension	CreatedExtension	An extension was installed.
Created group	CreatedGroup	A user group was created.
Created organization role assignment	CreatedOrganizationRoleAssignment	An organization role assignment was created.
Created role assignment	CreatedRoleAssignment	A role was assigned to a user.
Deleted billing plan	DeletedBillingPlan	A billing plan was deleted.
Deleted data export configuration	DeletedDataExportConfig	A data export configuration was deleted.
Deleted device connector	DeletedDeviceConnector	A device connector under an EHR connector was deleted.
Deleted EHR instance	DeletedEhrInstance	An EHR integration instance was removed.

Friendly name	Operation	Description
Deleted EHR user	DeletedEhrUser	An EHR user mapping was deleted.
Deleted environment	DeletedEnvironment	An environment was deleted.
Deleted group	DeletedGroup	A user group was deleted.
Deleted organization role assignment	DeletedOrganizationRoleAssignment	An organization role assignment was deleted.
Deleted role assignment	DeletedRoleAssignment	A role assignment was removed.
Deleted setting	DeletedSetting	A configuration setting was deleted.
Deprovisioned product	DeprovisionedProduct	A product was deprovisioned from an environment.
Installed device connector	InstalledDeviceConnector	A device connector was installed or upgraded under an EHR connector from a published package.
Installed device connector from package	InstalledDeviceConnectorFromPackage	A device connector was installed or upgraded under an EHR connector from a sideloaded device package.
Listed billing plans	ListedBillingPlans	Billing plans were listed.
Listed device connectors	ListedDeviceConnectors	Device connectors installed under an EHR connector were listed.
Listed EHR instances	ListedEhrInstances	EHR instances were listed.
Listed environments	ListedEnvironments	All environments were listed.
Listed extensions	ListedExtensions	Extensions were listed.
Listed groups	ListedGroups	User groups were listed.
Listed organizations	ListedOrganizations	Organizations were listed.
Listed role assignments	ListedRoleAssignments	Role assignments were listed.
Listed role definitions	ListedRoleDefinitions	An administrator listed the available role definitions.
Listed settings	ListedSettings	Configuration settings were listed.
Provisioned product	ProvisionedProduct	A product was provisioned for an environment.

Friendly name	Operation	Description
Queued deprovision	QueuedDeprovision	An environment was queued for deprovisioning.
Removed group member	RemovedGroupMember	A user was removed from a group.
Set device connector state	SetDeviceConnectorState	A device connector under an EHR connector was enabled or disabled.
Updated billing plan	UpdatedBillingPlan	Billing plan properties were modified.
Updated data export configuration	UpdatedDataExportConfig	A data export configuration was modified.
Updated device connector configuration	UpdatedDeviceConnectorConfiguration	The configuration of a device connector under an EHR connector was updated.
Updated EHR connector	UpdatedEhrConnector	An EHR connector was modified.
Updated EHR instance	UpdatedEhrInstance	An EHR integration instance was modified.
Updated environment	UpdatedEnvironment	Environment properties were modified.
Updated extension	UpdatedExtension	An extension was updated.
Updated group	UpdatedGroup	Group properties were modified.
Updated organization	UpdatedOrganization	Organization properties were modified.
Updated setting	UpdatedSetting	A configuration setting was changed.
Verified device connector package	VerifiedDeviceConnectorPackage	A sideloaded device package was verified against an EHR connector without being installed.
Viewed billing plan	ViewedBillingPlan	Billing plan details were viewed.
Viewed data export configuration	ViewedDataExportConfig	A data export configuration was viewed.
Viewed device connector	ViewedDeviceConnector	An installed device connector under an EHR connector was viewed.
Viewed EHR connector	ViewedEhrConnector	An EHR connector was viewed.
Viewed EHR instance	ViewedEhrInstance	An EHR instance was viewed.

Friendly name	Operation	Description
Viewed environment	ViewedEnvironment	Environment details were viewed.
Viewed extension	ViewedExtension	Extension details were viewed.
Viewed group	ViewedGroup	Group details were viewed.
Viewed organization	ViewedOrganization	Organization details were viewed.
Viewed role assignment	ViewedRoleAssignment	A role assignment was viewed.
Viewed role definition	ViewedRoleDefinition	An administrator viewed a role definition.
Viewed setting	ViewedSetting	A setting value was viewed.

Dragon Copilot Web activities

The following table lists user and application activities for Dragon Copilot Web recorded in the Microsoft 365 audit log. These events capture user access, clinical workflow actions, AI-assisted operations, and session lifecycle events performed within Dragon Copilot Web. Dragon Copilot Web supports clinical documentation workflows including ambient recording, transcript review, AI-assisted note generation, clinical chat, and EHR-integrated experiences.

If your organization uses Dragon Copilot Web, you can search for these activities in Microsoft Purview Audit by selecting the appropriate activity name or by filtering on the Dragon Copilot record types. Audit records are emitted under the following record types:

- DragonCopilotAccess
- DragonCopilotClinicalData
- DragonCopilotSession

You can also search for Dragon Copilot Web audit events by using `Search-UnifiedAuditLog` and filtering on the corresponding record type.

Dragon Copilot access activities

The following table lists user access and interaction activities for Dragon Copilot recorded in the Microsoft 365 audit log. These events are logged when users access Dragon Copilot or interact with its features and underlying services. These activities include launching Dragon Copilot from an EHR system, accessing Dragon Copilot, and updating patient session data. Events also capture failures related to these operations. If your organization uses Dragon Copilot, you can search the

audit log for these activities by using the **Activities** picker list. For more information about Dragon Copilot, see the Dragon Copilot documentation.

Expand table

Friendly name	Operation	Description
Application accessed	AppAccessed	A user opened the Dragon Copilot web application.
EHR launch completed	EhrLaunchCompleted	Dragon Copilot launch from EHR completed.
EHR launch failed	EhrLaunchFailed	Dragon Copilot launch from EHR failed.
Encounter accessed	EncounterAccessed	A user opened a patient encounter in Dragon Copilot.
License check failed	LicenseCheckFailed	User doesn't have Dragon Copilot license.
Session data updated	SessionDataUpdated	Session data updated.
Session data update failed	SessionDataUpdateFailed	Session data update failed.

Dragon Copilot clinical data activities

Expand table

Friendly name	Operation	Description
Transcript viewed	TranscriptViewed	A user viewed the transcript of an ambient or dictation session.
Clinical chat opened	ClinicalChatOpened	A user opened the clinical chat interface in Dragon Copilot.
Clinical chat episode viewed	ClinicalChatEpisodeViewed	A user viewed a previous clinical chat conversation episode.
Clinical chat prompt sent	ClinicalChatPromptSent	A user sent a message or prompt in the clinical chat interface.
Copilot prompt executed	CopilotPromptExecuted	A user triggered an AI action such as note summarization, generation, or rewrite.
Note generated	NoteGenerated	A clinical note was generated by the AI assistant and presented to the user.
Note updated	NoteUpdated	An AI-generated clinical note was updated with new or revised content and presented to the user.

Friendly name	Operation	Description
AI suggested edit presented	AiSuggestedEditPresented	An AI-suggested edit to a clinical note was presented to the user for review.
AI suggested edit accepted	AiSuggestedEditAccepted	A user accepted an AI-suggested edit to a clinical note.
AI suggested edit rejected	AiSuggestedEditRejected	A user rejected an AI-suggested edit to a clinical note.
Clinical chat feedback submitted	ClinicalChatFeedbackSubmitted	A user submitted feedback on a clinical chat response.

Dragon Copilot session activities

Expand table

Friendly name	Operation	Description
Session started	SessionStarted	A user started a new clinical documentation session in Dragon Copilot.
Session ended	SessionEnded	A user ended a clinical documentation session in Dragon Copilot.
Recording started	RecordingStarted	A user started audio recording during a clinical session.
Recording ended	RecordingEnded	A user stopped audio recording during a clinical session.

eDiscovery activities

The audit log records eDiscovery activities that you perform in the Microsoft Purview portal. It logs events when administrators or eDiscovery managers (or any user assigned eDiscovery permissions) perform the following tasks in the Microsoft Purview portal:

- Creating and managing eDiscovery cases.
- Creating and editing searches for eDiscovery cases.
- Performing search actions, such as generating statistics, creating a sample, and exporting from search.
- Creating, editing, and removing holds for eDiscovery cases.
- Creating review sets and performing review activities in eDiscovery cases.

For more information about searching the audit log, the permissions that are required, and exporting search results, see Search the audit log.

How to search for and view eDiscovery activities

The audit log records eDiscovery activities that you perform in the Microsoft Purview portal or in PowerShell. To search for eDiscovery activities, see Search for eDiscovery activities in the audit log.

eDiscovery activity reference

The following table lists the eDiscovery activities that are logged when an administrator or eDiscovery manager performs an eDiscovery-related activity by using the Microsoft Purview portal. Some activities performed in the classic eDiscovery user experience might be returned when you search for activities in this list.

> **Note**
>
> The client IP field is included for all activities performed in the new eDiscovery experience. To distinguish activities performed in the new experience from those in the classic experience, check for the presence of the client IP field in the audit entry. (In the following image, the first entry is eDiscovery case deleted in new modern ux and the other two entries without IP addresses are classic ux).

Expand table

Friendly name	Operation	Description
Added favorite	FavoriteSet	User set a case to favorite.
Added Purview search	PurviewSearchAdded	A new search was created. This audit activity includes the case name, case ID, and search name that was created.
Added review set	ReviewSetAdded	User created a review set.
Added saved search for a review set	ReviewSetSavedSearchAdded	User created saved filters in a review set. The audit activity includes the review set name, saved filters name, and the query used.
Added tag template	TagTemplateAdded	User created tag template in eDiscovery settings.
Cancelled sample view	SampleViewCancelled	User canceled sample view. The audit activity includes the canceled search's

Friendly name	Operation	Description
		name, ID, query, and associated settings.
Cancelled statistics view	StatisticsViewCancelled	User canceled statistics view. The audit activity includes the canceled search's name, ID, query, and associated settings.
Changed eDiscovery case	CaseUpdated	Updated an eDiscovery case.
Changed hold in eDiscovery case	HoldUpdated	Hold was modified or edited. This audit activity includes the hold name, ID, and the specific data sources and query values that were modified.
Closed eDiscovery case	CaseClosed	An eDiscovery case was closed.
Copied review set	ReviewSetCopied	User triggers the "Add to another review set" process.
Created eDiscovery case	CaseAdded	Added a new eDiscovery case.
Created hold in eDiscovery case	HoldCreated	A new hold was created. This audit activity includes the case name, case ID, and hold name that was created.
Deleted eDiscovery case	CaseRemoved	An eDiscovery case was deleted. You must remove any hold associated with the case before you can delete the case.
Deleted file for Purview search	PurviewSearchDeleteFile	User deleted a file from a search. This audit activity includes the case name, case ID, search name, search ID, and the ID of the file the user deleted.
Deleted hold in eDiscovery case	HoldRemoved	A hold that is associated with an eDiscovery case was deleted. Deleting a hold releases all of the content locations from the hold.
Deleted Purview search	PurviewSearchDeleted	A search was deleted. This activity includes the deleted case name, case ID, and search name.
Deleted Purview Search export job	PurviewSearchExportJobDeleted	An export from search was deleted. This activity includes the deleted export name, the case information in which the export is included, the user, and the deletion timestamp.

Friendly name	Operation	Description
Removed favorite	FavoriteRemoved	User removed a case from favorites.
Removed saved search for a review set	ReviewSetSavedSearchRemoved	User deleted saved filters in a review set.
Removed tag template	TagTemplateRemoved	User removed a tag template in eDiscovery settings. The audit activity includes the name and object ID of the tag template removed.
Reopened eDiscovery case	CaseReopened	An eDiscovery case was reopened.
Retried distribution sync for hold	HoldRetryDistributionSync	User triggered “retry policy” in a hold.
Retrieved actions for all review sets	GetActionsForAllReviewSets	User viewed a list of actions associated with all review sets in a case.
Retrieved all action for hold	GetActionsForAllHolds	User viewed a list of actions associated with all holds in a case. The audit activity includes case name, ID, hold name, ID, and the name of the list of actions.
Retrieved all action for search	GetActionsForSearch	User viewed a list of actions (such as exports) associated with a search. The audit activity includes case name, ID, search name, ID, and the name of the list of actions.
Retrieved all actions for all exports	GetActionsForAllExports	User viewed a list of exports in a case. The audit activity includes case name, ID, case settings, and the list of export names viewed.
Retrieved all actions for case	GetActionsForCase	User viewed a list of actions (such as exports) associated with an eDiscovery case.
Retrieved all actions for hold	GetActionsForHold	User viewed a list of actions associated with a hold.
Retrieved all actions for review set	GetActionsForReviewSet	User viewed a list of actions (such as exports) associated with a review set. The audit activity includes case name, ID,

Friendly name	Operation	Description
		review set name, ID, and the name of the list of actions.
Retrieved single action for case	GetSingleActionForCase	User viewed a single action associated with an eDiscovery case. For example, the user viewed the export process in process manager. The audit activity includes case name, ID, and settings and ID associated with the process.
Retrieved single action for hold	GetSingleActionForHold	User viewed a single action associated with a hold.
Retrieved single action for review set	GetSingleActionForReviewSet	User viewed a single action associated with a review set.
Retrieved single action for search	GetSingleActionForSearch	User viewed a single action associated with a search. For example, the user viewed the export process in the search's process manager. The audit activity includes case name, ID, and settings and ID associated with the process.
Submitted new algo job	AlgoJobSubmitted	User ran analytics on the documents in a review set.
Submitted new burn job	BurnJobSubmitted	User converted all the redacted documents in a review set to PDF files.
Submitted purview search add to review set job	PurviewSearchAddToReviewSetJobSubmitted	User triggered an "add to review set" process from a search. This audit activity includes the search name, ID, and the specific data sources and query values that were associated to the search. It also includes relevant add to review set process settings used.
Submitted purview search export job	PurviewSearchExportJobSubmitted	User triggered an "export" process from a search. This audit activity includes the search name, ID, and the specific data sources and query values that were associated to the search. It also includes relevant export process settings used and the export name.
Submitted purview search sample job	PurviewSearchSampleJobSubmitted	User triggered a "generate sample" process from a search. This audit activity includes the search name, ID, and the

Friendly name	Operation	Description
		specific data sources and query values that were associated to the search. It also includes relevant sample process settings used.
Submitted purview search statistics job	PurviewSearchStatisticsJobSubmitted	User triggered a “generate statistics” process from a search. This audit activity includes the search name, ID, and the specific data sources and query values that were associated to the search. It also includes relevant statistics process settings used.
Submitted review set export job	ReviewSetExportJobSubmitted	Exported documents from review set. This audit activity includes the review set name, list of documents IDs being exported and relevant export process settings.
Tagged documents by ID for a review set	ReviewSetDocumentsTaggedById	User applied tags to specific documents in a review set. The audit activity includes the case name, review set name, and the list of items being tagged.
Tagged documents by query for a review set	ReviewSetDocumentsTaggedByQuery	User applied tags to specific documents in a review set after filtering by query. The audit activity includes the case name, review set name, and the list of items being tagged.
Updated case members	CaseMembersUpdated	Case membership was updated. As a member of a case, a user can perform various case-related tasks depending on whether they're assigned the necessary permissions.
Updated case settings	CaseSettingsUpdated	User modified the settings for a case. Case settings include case information, premium feature access, case permissions, and settings that control search and analytics behavior.
Updated notes for a review set	ReviewSetNotesUpdated	User updated note for a review set document.
Updated Purview search	PurviewSearchUpdated	Search was modified or edited. This audit activity includes the search name, ID, and the specific data sources and query values that were modified.

Friendly name	Operation	Description
Updated review set	ReviewSetUpdated	User updated the review set such as review set name and description.
Updated review set annotation	ReviewSetAnnotationsUpdated	User annotated a document in a review set. This audit activity includes the review set name and other information such as the annotated document ID.
Updated saved search for a review set	ReviewSetSavedSearchUpdated	User modified saved filters in a review set. The audit activity includes the review set name, saved filters name, and the query used.
Updated tag template	TagTemplateUpdated	User updated a tag template in eDiscovery settings. The audit activity includes the name and object ID of the tag template updated.
Updated tags	TagsUpdated	User updated review set tags in a case.
Updated tags for a tag template	TagTemplateTagsUpdated	User updated tags for a tag template. The audit activity includes the name and ID of the tag template tags updated.
Updated tenant settings	TenantSettingsUpdated	User updated eDiscovery's organization settings.
Viewed analytics review set attachments report	AlgoGetReviewSetAttachmentsReport	User viewed analytics attachment report. This audit activity includes case and review set name and ID.
Viewed analytics review set documents report	AlgoGetReviewSetDocumentsReport	User viewed analytics document report.
Viewed analytics review set email report	AlgoGetReviewSetEmailsReport	User viewed analytics email report.
Viewed analytics review set report	AlgoGetReviewSetReport	User viewed analytics report in a review set. This audit activity includes case and review set name and ID, as well as the report type.
Viewed analytics review set	AlgoGetReviewSetReportByFileType	User viewed analytics report graph by file type.

Friendly name	Operation	Description
report by file type		
Viewed analytics review set report by source	AlgoGetReviewSetReportBySource	User viewed analytics document by source. This audit activity includes case and review set name and ID.
Viewed case metadata	CaseMetadataViewed	Metadata about an eDiscovery case was viewed.
Viewed case settings	CaseSettingsViewed	User viewed the settings for a case. Case settings include case information, premium feature access, case permissions, and settings that control search and analytics behavior.
Viewed data processing errors for a review set	ReviewSetDataProcessingErrorViewed	User viewed processing errors in a review set.
Viewed default case settings	CaseSettingsDefaultViewed	Default case setting viewed.
Viewed eDiscovery case	CaseViewed	A user viewed an eDiscovery case in the Microsoft Purview portal. The audit record for this event includes the name, case settings, and case ID of the case that was viewed.
Viewed eDiscovery case list	CaseListViewed	This activity is logged when a user viewed a list of eDiscovery cases. The audit record includes the name and ID of the cases that was viewed in the cases list.
Viewed favorite list	FavoriteListViewed	User viewed a list of favorite cases.
Viewed hold	HoldViewed	User viewed a hold inside a case. This audit activity includes the hold name and ID the user viewed. It also includes the specific data sources and query values inside the hold that was viewed by the user.
Viewed hold list	HoldListViewed	User viewed the holds list inside a case. This audit activity includes the case name, case ID, and the list of holds name and ID the user viewed.

Friendly name	Operation	Description
Viewed hold results	HoldResultsViewed	User viewed hold's results. The audit activity includes case name, case ID, hold name, and hold ID.
Viewed hold status	HoldStatusViewed	User viewed hold's status. The audit activity includes case name, case ID, hold name, and hold ID.
Viewed if analytics settings are changed	AlgoIsSettingChanged	This audit activity includes case and review set name and ID.
Viewed load count in case	LoadCountInCaseViewed	User viewed the count of load sets in a case.
Viewed load list	LoadListViewed	User viewed a list of load sets in review set.
Viewed Purview search	PurviewSearchViewed	User viewed a specific search. This audit activity includes the search name, ID, and the specific data sources and query values inside the search that was viewed by the user.
Viewed Purview search list	PurviewSearchListViewed	User viewed the searches list inside a case. This audit activity includes the case name, case ID, and the list of searches name and ID the user viewed.
Viewed query report for a review set	ReviewSetQueryReportViewed	User viewed query report inside a review set.
Viewed review set document	ReviewSetDocumentViewed	User viewed a document inside a review set. This audit activity includes the case name, case ID, review set name, review set ID and other information such as the ID of the document viewed.
Viewed review set list	ReviewSetListViewed	User viewed the list of review sets in a case.
Viewed review set summary	ReviewSetSummaryViewed	User viewed the Overview of a Review set.
Viewed sample results	SampleResultsViewed	User viewed search sample results view.

Friendly name	Operation	Description
Viewed sample status	SampleStatusViewed	User viewed search sample view's status. The audit activity includes search's name, ID, query, and associated settings.
Viewed saved search list for a review set	ReviewSetSavedSearchListViewed	User viewed a list of saved filters in a review set. The audit activity includes the review set name and the name of the list of saved filters viewed.
Viewed search count for a review set	ReviewSetSearchCountViewed	User ran a search within a review set and viewed the count returned. The audit activity includes the case name, review set name and the count of items returned in the search results.
Viewed search options for a review set	ReviewSetSearchOptionsViewed	User viewed the settings of a review set. The audit activity includes the case name and review set name.
Viewed search results for a review set	ReviewSetSearchRun	User ran a search within a review set. The audit activity includes the case name, review set name and the list of items returned in the search results.
Viewed statistics results	StatisticsResultsViewed	User viewed search statistics results.
Viewed statistics status	StatisticsStatusViewed	User viewed search statistics view's status. The audit activity includes search's name, ID, query, and associated settings.
Viewed tag template list	TagTemplateListViewed	User viewed the list of tag template in eDiscovery settings. The audit activity includes the list of tag templates viewed.
Viewed tags	TagsViewed	User viewed review set tags in a case.
Viewed tags for a tag template	TagTemplateTagsViewed	User viewed tags for a tag template. The audit activity includes the name and ID of the tag template tags viewed.
Viewed tenant settings	TenantSettingsViewed	User viewed eDiscovery's organization settings. The audit activity contains information on the setting values.

Detailed properties for eDiscovery activities

The following table lists the properties that the portal shows on the details page for an eDiscovery activity listed in the search results. These properties are also included in the CSV file when you export the audit log search results. An audit log record for an eDiscovery activity doesn't include every detailed property listed in the following table.

> **Tip**
>
> When you export the search results, the CSV file contains a column named **AuditData**, which contains the detailed properties described in the following table in a multivalue property. You can use the Power Query feature in Excel to split this column into multiple columns so that each property has its own column. This setup lets you sort and filter on one or more of these properties. For more information, see Search the audit log.

Property	Description
CaseId	The identity (GUID) of the eDiscovery case created, changed, or deleted.
CaseName	The name of the eDiscovery case created, changed, or deleted.
ClientApplication	eDiscovery cmdlet activities have a value of **EMC** for this property. This value indicates the activity was performed by using the Microsoft Purview portal GUI or running the cmdlet in PowerShell.
ClientIP	The IP address of the device that was used when the activity was logged. The IP address is displayed in either an IPv4 or IPv6 address format.
ClientRequestId	For eDiscovery activities, this property is typically blank.
CmdletVersion	The build number for the version of the Microsoft Purview portal running in your organization.
CreationTime	The creation date and time in Coordinated Universal Time (UTC) for the activity.
DataSources	A list of source ID, source name, and location details associated to the activity.
EffectiveOrganization	The name of the Microsoft 365 organization.
ExportName	Name of the eDiscovery export.
ExtendedProperties	Additional properties related to the eDiscovery activity. For example, when case members are updated, this field contains the updated case member information; or when the review set description is updated, this field contains the updated review set description updated by the user.
ID	The ID of the report entry. The ID uniquely identifies the audit log entry.
Item	The name of the item associated with the activity. For example, this field contains the name of the document viewed when a user views a review set document.

Property	Description
JobId	The GUID of the eDiscovery process.
ObjectId	The GUID of the object (for example, a search, hold, or review set) created, accessed, or changed, by the activity listed in the *Operation* property.
ObjectName	The name of the object (for example, a search, hold or review set) created, accessed, or changed, by the activity listed in the Operation property.
ObjectType	The type of eDiscovery object created, deleted, or modified. For example, a search action (generate sample results or trigger an export from search) are listed as *Search* in this field.
Operation	The name of the operation that corresponds to the eDiscovery activity that was performed.
OrganizationId	The GUID for your Microsoft 365 organization.
QueryId	The GUID of the query associated with the activity.
QueryText	The query text associated with the activity, such as a search statistic process or add to review process.
RecordType	The type of operation indicated by the record.
ResultStatus	If the action (specified in the Operation property) was successful or not.
Settings	Settings applied to the eDiscovery activity.
StartTime	The date and time in Coordinated Universal Time (UTC) when the eDiscovery activity was started.
UserCancelled	Whether the specific activity was canceled by the user.
UserId	The user who performed the activity (specified in the Operation property) that resulted in the record being logged.
UserKey	An alternative ID for the user identified in the UserId property. For eDiscovery activities, the value for this property is typically the same as the UserId property.
UserServicePlan	The subscription used by your organization. For eDiscovery activities, this property is typically blank.
UserType	The type of user that performed the operation. The following values indicate the user type. 0 A regular user. 2 An administrator in your organization. 3 A Microsoft datacenter administrator or datacenter system account. 4 A system account. 5 An application. 6 A service principal.
Version	The version number of the activity (identified by the Operation property) that's logged.

Property	Description
Workload	The service where the activity occurred.

Encrypted message portal activities

Your organization can access logs for encrypted messages through the encrypted message portal. These logs help you determine when your external recipients read and forward messages. For more information about enabling and using encrypted message portal activity logs, see Encrypted message portal activity log.

Each audit entry for a tracked message contains the following fields:

- **MessageID**: Contains the ID of the message being tracked. Use this key identifier to follow a message through the system.
- **Recipient**: List of all recipient email addresses.
- **Sender**: The originating email address.
- **AuthenticationMethod**: Describes the authenticating method for accessing the message, such as OTP, Yahoo, Gmail, or Microsoft.
- **AuthenticationStatus**: Contains a value indicating that the authentication succeeded or failed.
- **OperationStatus**: Whether the indicated operation succeeded or failed.
- **AttachmentName**: Name of the attachment.
- **OperationProperties**: A list of optional properties. For example, the number of OTP passcodes sent, or the email subject.

eSignature activities

The following table lists the eSignature activities that the Microsoft 365 audit log records. For more information, see eSignature for Microsoft 365.

Friendly name	Operation	Description
Cancelled by the requestor	ESignatureRequestCanceled	The signing request was cancelled by the requestor.
Declined by a signer	ESignatureRequestDeclined	A signer declined the request.
Document was downloaded	ESignatureRequestDownloaded	The document was downloaded by a member of the workflow.
Document was viewed	ESignatureDocumentViewed	The document was viewed by a member of the workflow.

Friendly name	Operation	Description
Link to the signed document expired	ESignatureRequestExpired	The link to the signed document expired.
Request was completed	ESignatureRequestCompleted	An electronic signature request was completed.
Request was created	ESignatureRequestCreated	An electronic signature request was created by a member of the organization.
Request was sent	ESignatureRequestSent	An electronic signature request was sent to recipients.
Request was signed by a user	ESignatureDocumentSigned	An electronic signature request was signed by a recipient.

Exchange admin activities

Exchange administrator audit logging (which Microsoft 365 enables by default) logs an event in the audit log when an administrator (or a user assigned administrative permissions) makes a change in your Exchange Online organization. The Exchange admin audit log records changes made by using the Exchange admin center or by running a cmdlet in Exchange Online PowerShell. The audit log doesn't record cmdlets that begin with the verbs **Get-**, **Search-**, or **Test-**. For more detailed information about admin audit logging in Exchange, see Administrator audit logging.

> **ⓘ Important**
>
> Some Exchange Online cmdlets aren't logged in the Exchange admin audit log or in the audit log. Many of these cmdlets relate to maintaining the Exchange Online service and Microsoft datacenter personnel or service accounts run them. These cmdlets aren't logged because they would result in a large number of "noisy" auditing events. If there's an Exchange Online cmdlet that isn't being audited, submit a design change request (DCR) to Microsoft Support.

Use these tips to search for Exchange admin activities when searching the audit log:

- Use the date range boxes and the **Users** list to narrow the search results for cmdlets run by a specific Exchange administrator within a specific date range.
- Select the **Activity** column to sort the cmdlet names in alphabetical order and display events from the Exchange admin audit log.

- Select the **Download all results** option to export the search results and get information about what cmdlet was run, which parameters and parameter values were used, and what objects were affected. For more information, see Export, configure, and view audit log records.
- Use the `Search-UnifiedAuditLog -RecordType ExchangeAdmin` command in Exchange Online PowerShell to return only audit records from the Exchange admin audit log. It might take up to 30 minutes after an Exchange cmdlet is run for the corresponding audit log entry to be returned in the search results. For more information, see Search-UnifiedAuditLog. For information about exporting the search results returned by the **Search-UnifiedAuditLog** cmdlet to a CSV file, see the "Tips for exporting and viewing the audit log" section in Export, configure, and view audit log records.

Exchange mailbox activities

The following table lists the activities recorded in the Microsoft 365 audit log. Mailbox audit logging automatically logs mailbox activities performed by the mailbox owner, a delegated user, or an administrator in the audit log for up to 180 days. An admin can turn off mailbox audit logging for all users in your organization. In this case, no mailbox actions for any user are logged. For more information, see Manage mailbox auditing.

> **Important**
>
> The default retention period for Audit (Standard) changed from 90 days to 180 days. Audit (Standard) logs generated before October 17, 2023, are retained for 90 days. Audit (Standard) logs generated on or after October 17, 2023, follow the new default retention of 180 days.

You can also search for mailbox activities by using the **Search-UnifiedAuditLog** cmdlet in Exchange Online PowerShell.

Friendly name	Operation	Description
Accessed mailbox attachments	AttachmentAccess	A message attachment was accessed.
Accessed Mailbox folder	FolderBind	A mailbox folder was accessed. This action is also logged when the admin or delegate opens the mailbox. Audit records for folder bind actions performed by delegates are consolidated. One audit record is

Friendly name	Operation	Description
		generated for individual folder access within a 24-hour period.
Accessed mailbox items	MailItemsAccessed	Messages were read or accessed in mailbox. Audit records for this activity are triggered in one of two ways: when a mail client (such as Outlook) performs a bind operation on messages or when mail protocols (such as Exchange ActiveSync or IMAP) sync items in a mail folder. Analyzing audit records for this activity is useful when investigating compromised email account.
Accessed message	MessageBind	A message was viewed in the preview pane or opened by an admin. This value is available only for users without E5/A5/G5 licenses.
Added delegate mailbox permissions	Add-MailboxPermission	An administrator assigned the FullAccess mailbox permission to a user (known as a delegate) to another person's mailbox. The FullAccess permission allows the delegate to open the other person's mailbox, and read and manage the contents of the mailbox. The audit record for this activity is also generated when a system account in the Microsoft 365 service periodically performs maintenance tasks in behalf of your organization. A common task performed by a system account is updating the permissions for system mailboxes. For more information, see System accounts in Exchange mailbox audit records.
Added or removed user with delegate access to calendar folder	UpdateCalendarDelegation	A user was added or removed as a delegate to the calendar of another user's mailbox. Calendar delegation gives someone else in the same organization permissions to manage the mailbox owner's calendar.
Added permissions to folder	AddFolderPermissions	A folder permission was added. Folder permissions control which users in your organization can access folders in a mailbox and the messages located in those folders.
Copied messages to another folder	Copy	A message was copied to another folder.
Created mailbox item	Create	An item is created in the Calendar, Contacts, Notes, or Tasks folder in the mailbox. For example, a new meeting request is created. Creating, sending, or receiving a message isn't audited. Also, creating a mailbox folder isn't audited.

Friendly name	Operation	Description
Created new inbox rule in Outlook web app	New-InboxRule	A mailbox owner or other user with access to the mailbox created an inbox rule in the Outlook web app.
Delete priority cleanup label item	PriorityCleanupDelete	The item with a label of Priority Cleanup type is getting deleted.
Deleted messages from Deleted Items folder	SoftDelete	A message was permanently deleted or deleted from the Deleted Items folder. These items are moved to the Recoverable Items folder. Messages are also moved to the Recoverable Items folder when a user selects it and presses **Shift+Delete**.
Label message as a record		A user applied a retention label to an email message and that label is configured to mark the item as a record.
Labeled item as a record	ApplyRecord	An item is labeled as a record.
Labeled message as a record	ApplyRecordLabel	A message was classified as a record. Occurs when a retention label that classifies content as a record is manually or automatically applied to a message.
Modified folder permission	ModifyFolderPermissions	A folder permission was changed. Folder permissions control which users in your organization can access mailbox folders and the messages in the folder.
Modified inbox rule from Outlook web app	Set-InboxRule	A mailbox owner or other user with access to the mailbox modified an inbox rule using the Outlook web app.
Moved messages to another folder	Move	A message was moved to another folder.
Moved messages to Deleted Items folder	MoveToDeletedItems	A message was deleted and moved to the Deleted Items folder.
Perform search query	SharepointSearchQueryInitiated	The user performs a search on SharePoint.
Preserve Mailbox item	PreservedMailItemProactively	The mail item is proactively preserved. Proactive preservation is a Microsoft Purview Data Lifecycle

Friendly name	Operation	Description
		Management (DLM) feature, where mails deleted by risky users in the tenant are preserved in dumpster.
Priority cleanup label applied	ApplyPriorityCleanup	The Priority Cleanup label is applied to an item.
Priority cleanup label applied on Exchange item	ApplyPriorityCleanup	A retention label for priority cleanup is applied to an item on Exchange.
Purged messages from the mailbox	HardDelete	A message was purged from the Recoverable Items folder (permanently deleted from the mailbox).
Removed delegate mailbox permissions	Remove-MailboxPermission	An administrator removed the FullAccess permission (that was assigned to a delegate) from a person's mailbox. After the FullAccess permission is removed, the delegate can't open the other person's mailbox or access any content in it.
Removed permissions from folder	RemoveFolderPermissions	A folder permission was removed. Folder permissions control which users in your organization can access folders in a mailbox and the messages located in those folders.
Search items in a mailbox	SearchQueryInitiated	A person uses Outlook (Windows, Mac, iOS, Android, or Outlook on the web) or the Mail app for Windows 10 to search for items in a mailbox.
Sent message	Send	A message was sent, replied to, or forwarded.
Sent message using Send As permissions	SendAs	A message was sent using the SendAs permission. This means that another user sent the message as though it came from the mailbox owner.
Sent message using Send On Behalf permissions	SendOnBehalf	A message was sent using the SendOnBehalf permission. This means that another user sent the message on behalf of the mailbox owner. The message the recipient to whom the message was sent on behalf of and who actually sent the message.
Update item label	UpdateComplianceTag	When the label is applied to an item.
Updated inbox rules from Outlook client	UpdateInboxRules	A mailbox owner or other user with access to the mailbox created, modified, or removed an inbox rule by using the Outlook client.

Friendly name	Operation	Description
Updated message	Update	A message or its properties was changed.
User signed in to mailbox	MailboxLogin	The user signed in to their mailbox.

System accounts in Exchange mailbox audit records

In audit records for some mailbox activities, especially **Add-MailboxPermissions**, you might see the user who performed the activity identified as NT AUTHORITY\SYSTEM or NT SERVICE\MSExchangeAdminApiNetCore(Microsoft.Exchange.AdminApi.NetCore). This system account in the Exchange service in the Microsoft cloud performs scheduled maintenance tasks on behalf of your organization. For example, a common audited activity performed by the NT SERVICE\MSExchangeAdminApiNetCore(Microsoft.Exchange.AdminApi.NetCore) account is updating the permissions on the DiscoverySearchMailbox, which is a system mailbox. This update verifies that the FullAccess permission (which is the default) is assigned to the Discovery Management role group for the DiscoverySearchMailbox. This update ensures that eDiscovery administrators can perform necessary tasks in their organization.

Another system user account that might appear in an audit record for **Add-MailboxPermission** is Administrator@apcprd03.prod.outlook.com. This service account also appears in mailbox audit records related to verifying and updating the FullAccess permission assigned to the Discovery Management role group for the DiscoverySearchMailbox system mailbox. Specifically, audit records that identify the Administrator@apcprd03.prod.outlook.com account typically trigger when Microsoft support personnel run a role-based access control diagnostic tool on behalf of your organization.

Fabric activities

The following table lists the Microsoft Fabric activities in Power BI that the Microsoft 365 audit log records. You can search the audit log for these activities in Power BI. For information about audit settings, see Audit and usage tenant settings.

Friendly name	Operation	Description
Applied a PostgreSQL database schema	ApplyPostgreSQLDatabaseSchema	Generated when a user applies (plans and executes) a PostgreSQL database schema through the custom PG schema service (pgschema-based ALM

Friendly name	Operation	Description
		flow). The audit log records caller identity, operation result, and the affected PostgreSQL database artifact.
Assign Warehouse Server Alias	AssignWarehouseServerAlias	A workspace administrator assigned a server alias to a Fabric Warehouse workspace.
Automatically bound user credentials to Git	AutoBoundGitCredentials	Automatically bound user credentials to Git.
Branch workspace admin profile configured	GitBranchWorkspaceAdminProfileConfigured	Branch workspace admin profile configured.
Branch workspace configured	GitBranchWorkspaceConfigured	Branch workspace configured.
Branched out to a workspace in Git	GitBranchedOut	Branched out to a workspace in Git.
Browsed PostgreSQL database objects	PgSQLDbObjectExplorer	Generated when a user browses PostgreSQL database schema objects through Object Explorer. The audit log records caller identity, operation result, and the affected PostgreSQL database artifact.
Cancelled a Digital Operations Ontology Agent conversation	DigitalOperationsOntologyAgentConversationCancelled	A user canceled a Digital Operations Ontology Agent conversation.
Checked out Git branch	GitCheckedOutBranch	Checked out Git branch.
Compute Items Sizes	ComputeItemsSize	Calculated current size of all items for the OneLake item-size report.

Friendly name	Operation	Description
Create the cross tenant auth mapping	CreateCrossTenantAuthMapping	Create a user mapping for the cross tenant auth feature.
Created Artifact Definition	CreatedArtifactDefinition	Created Artifact Definition is a Fabric External Sources activity which is generated when an External Sources artifact definition is created. Shared across item types: for MirroredStorage it carries SourceType and ConnectionId; for AWS Databricks Catalog it carries CatalogName, WorkspaceConnectionId, and MirroringMode.
Created AWS Databricks Catalog	CreatedAWSDatabricksCatalog	Created AWS Databricks Catalog is an AWSDatabricksCatalog activity which is generated when a Fabric AWS Databricks Catalog item is created, capturing the bound AWS Databricks workspace ConnectionId and the initial MirroredScope (catalogs/schemas/tables selected for mirroring).
Created a Digital Operations Ontology Agent conversation	DigitalOperationsOntologyAgentConversationCreated	A user created a Digital Operations Ontology Agent conversation.
Created a Fabric Copilot session	FabricCopilotSessionCreated	A user created a Fabric Copilot session.
Create a Microsoft Fabric free trial	CreatedFabricTrial	A new Microsoft Fabric free trial capacity was started.
Created MirroredStorage	CreatedMirroredStorage	Generated when a user or service principal creates a Fabric MirroredStorage item, linking an external storage source to a workspace as OneLake shortcuts.

Friendly name	Operation	Description
Created workspace relation	CreatedWorkspaceRelation	Created workspace relation.
Deleted a Fabric Copilot session	FabricCopilotSessionDeleted	A user deleted a Fabric Copilot session.
Deleted a task flow	DeletedTaskFlow	Deleted a task flow.
Deleted connection by tenant admin	DeletedGatewayClusterDatasourceAsAdmin	A tenant admin deleted the connection.
Deleted connection role assignment by tenant admin	DeletedGatewayDatasourceByAdmin	A tenant admin deleted the connection role assignment.
Deleted workspace relation	DeletedWorkspaceRelation	Deleted workspace relation.
Downloaded dataflow refresh logs	DownloadedDataflowRefreshLogs	Downloaded dataflow refresh logs.
Edited Power BI semantic model options	EditedSemanticModelOptions	A user made a change to their semantic model options. This occurs when changes are made in the model options dialog.
Executed a tenant relocation	TenantRelocationExecuted	Executed tenant relocation.
Export item definitions	ExportItemDefinitions	Export multiple item definitions from a workspace.
Exported a PostgreSQL database schema	ExportPostgreSQLDatabaseSchema	Generated when a user dumps a PostgreSQL database schema through the custom PG schema service (pgschema-based ALM flow). The audit log records caller identity, operation result, and the affected PostgreSQL database artifact.

Friendly name	Operation	Description
Get Items Sizes	GetItemsSize	Retrieved cached item sizes for the OneLake item-size report.
Get connection by tenant admin	GetGatewayClusterDatasourceAsAdmin	A tenant admin retrieved the connection details.
Get on-demand billing limits	GetOnDemandBillingLimits	Generated when a capacity admin opens the On-demand billing page and the current billing categories and their 24-hour compute limits are returned for a Fabric capacity.
Get on-demand billing quota	GetOnDemandBillingQuota	Generated when the On-demand billing experience loads and the subscription-level on-demand billing quota usage and limits are retrieved.
Get Capacity Overage Config. This can happen via Public API or using UX.	GetCapacityOverageBillingConfiguration	Get Capacity Overage Config. This can happen via Public API or using UX.
Git connection settings updated	GitConnectionSettingsUpdated	Git connection settings updated.
Granted consent to tenant relocation	TenantRelocationConsentGranted	Tenant relocation consent granted.
Import item definitions	ImportItemDefinitions	Import multiple item definitions into a workspace.
Imported OneLake lifecycle policy	ImportedLifecyclePolicy	Imported OneLake Lifecycle policy.
Imported PostgreSQL sample data	ImportSampleDataToPostgreSQLDatabase	Generated when a user imports supported sample data into a Fabric Native PostgreSQL database. The audit log records caller identity, operation result, and the affected PostgreSQL database artifact.

Friendly name	Operation	Description
Initialized connection to Git	GitConnectionInitialized	Initialized connection to Git.
Item definition exported	ItemDefinitionExported	A Fabric item definition has been exported.
List all of the cross tenant auth mappings	ListCrossTenantAuthMappings	List cross tenant auth mappings in a tenant.
Logged in to Git provider	ConfiguredGitProviderCredentials	Logged in to Git provider.
Logged out of Git provider	DeletedGitProviderCredentials	Logged out of Git provider.
Microsoft Fabric Audit activity to retrieve FMI bindings	GetFabricManagedIdentityBindings	Microsoft Fabric Audit activity to retrieve FMI bindings.
Migrate Gen1 dataflow	MigrateGen1Dataflow	A user upgraded a Power BI Gen1 dataflow to a Fabric Gen2.1 dataflow in place by using the Upgrade Wizard.
Modified OneLake default tier	ModifiedDefaultTier	Modified OneLake default tier.
Modified OneLake file soft-delete settings	ModifiedOneLakeFileSoftDeleteSettings	Generated when a workspace admin changes the OneLake file-level soft-delete retention settings (enable/disable, retention period).
Planning session upgraded	PlanningSessionUpgraded	Session type is upgraded in planning workload.
Recover an artifact	ArtifactRecovered	A user recovered a previously soft-deleted artifact, such as a gateway or gateway cluster member, restoring it to active state.
Refreshed AWS Databricks	RefreshedAWSDatabricksCatalogShortcuts	Refreshed AWS Databricks Catalog shortcuts is an

Friendly name	Operation	Description
Catalog Shortcuts		AWSDatabricksCatalog activity which is generated when the refresh-catalog-metadata operation re-syncs the already-mirrored shortcuts against the underlying AWS Databricks Unity Catalog.
Retrieved artifact's Logical Id	ArtifactLogicalIdRetrieved	Retrieved artifact's Logical Id.
Retrieved PostgreSQL SQL audit policy	GetSqlAuditPolicyOnDatabase	Generated when a user retrieves the SQL audit policy settings for a PostgreSQL database artifact. The audit log records caller identity, operation result, and the affected PostgreSQL database artifact.
Revoked consent to tenant relocation	TenantRelocationConsentRevoked	Tenant relocation consent revoked.
Sent a Digital Operations Ontology Agent message	DigitalOperationsOntologyAgentMessageSent	A user sent a message in a Digital Operations Ontology Agent conversation.
Sent a Fabric Copilot message	FabricCopilotSessionMessageSent	A user sent a message in a Fabric Copilot session.
Set PostgreSQL SQL audit policy	SetSqlAuditPolicyOnDatabase	Generated when a user updates SQL audit policy settings for a PostgreSQL database artifact. The audit log records caller identity, operation result, and the affected PostgreSQL database artifact.
Set Warehouse Connection Mode To Automatic	SetWarehouseConnectionModeToAutomatic	A workspace administrator set the Fabric Warehouse SQL connection mode to Automatic.

Friendly name	Operation	Description
Set Warehouse Connection Mode To Proxy	SetWarehouseConnectionModeToProxy	A workspace administrator set the Fabric Warehouse SQL connection mode to Proxy.
Set Warehouse Connection Mode To Redirect	SetWarehouseConnectionModeToRedirect	A workspace administrator set the Fabric Warehouse SQL connection mode to Redirect.
Set Warehouse Hardware Acceleration	SetWarehouseHardwarePlatform	Changed the current hardware acceleration settings for warehouses in a workspace.
Soft delete an artifact	ArtifactSoftDeleted	A user soft deleted an artifact, such as a gateway or gateway cluster member, marking it for deletion while allowing recovery within a retention period.
Switched Git branch	GitSwitchedBranch	Switched Git branch for workspace.
Update on-demand billing limits	UpdateOnDemandBillingLimits	Generated when a capacity admin enables or disables a billing category, or changes its 24-hour compute limit, on a Fabric capacity. One audit event is emitted per billing category.
Update Capacity Overage Config. This can happen via Public API or using UX.	UpdateCapacityOverageBillingConfiguration	Update Capacity Overage Config. This can happen via Public API or using UX.
Updated a Fabric Copilot session	FabricCopilotSessionUpdated	A user updated a Fabric Copilot session.
Updated a Fabric Copilot session state	FabricCopilotSessionStateUpdated	A user updated the state of a Fabric Copilot session.
Updated artifact's Logical Id	ArtifactLogicalIdUpdated	Updated artifact's Logical Id.

Friendly name	Operation	Description
Updated authorization setting in GraphQL	UpdatedAuthorizationSettingGraphQL	Updated authorization setting in GraphQL.
Updated Artifact Definition	UpdatedArtifactDefinition	Updated Artifact Definition is a Fabric External Sources activity which is generated when an External Sources artifact definition is updated via the update-item-definition API. Shared across item types (e.g. MirroredStorage, AWS Databricks Catalog); the properties carried depend on the item type.
Updated AWS Databricks Catalog Definition	UpdatedAWSDatabricksCatalogDefinition	Updated AWS Databricks Catalog definition is an AWSDatabricksCatalog activity which is generated when the item's Databricks workspace connection or mirrored scope (catalogs/schemas/tables) is updated via the update-item-definition API.
Updated connection role assignment by tenant admin	UpdatedGatewayDatasourceByAdmin	A tenant admin add or updates the connection role assignment.
Updated git items selection	GitItemsSelectionUpdated	Updated git items selection.
Updated MirroredStorage definition	UpdatedMirroredStorageDefinition	Generated when a user or service principal modifies the mirroring scope of an existing MirroredStorage item, changing which external paths are exposed in OneLake.
Unassign Warehouse Server Alias	UnassignWarehouseServerAlias	A workspace administrator assigned a server alias to a Fabric Warehouse workspace.
Workspace relations	WorkspaceRelationsRetrieved	Workspace relations retrieved.

Friendly name	Operation	Description
retrieved		

File and page activities

The following table lists the file and page activities in SharePoint and OneDrive that the Microsoft 365 audit log records.

Friendly name	Operation	Description
(none)	FileAccessedExtended	This activity is related to the "Accessed file" (FileAccessed) activity. A FileAccessedExtended event is logged when the same person continually accesses a file for an extended period (up to three hours). The purpose of logging FileAccessedExtended events is to reduce the number of FileAccessed events that are logged when a file is continually accessed. This approach helps reduce the noise of multiple FileAccessed records for what is essentially the same user activity, and lets you focus on the initial (and more important) FileAccessed event.
(none)	FileModifiedExtended	This activity is related to the "Modified file" (FileModified) activity. A FileModifiedExtended event is logged when the same person continually modifies a file for an extended period (up to three hours). The purpose of logging FileModifiedExtended events is to reduce the number of FileModified events that are logged when a file is continually modified. This approach helps reduce the noise of multiple FileModified records for what is essentially the same user activity, and lets you focus on the initial (and more important) FileModified event.

Friendly name	Operation	Description
(none)	FilePreviewed	User previews files on a SharePoint or OneDrive site. These events typically occur in high volumes based on a single activity, such as viewing an image gallery.
(none)	PagePrefetched	A user's client (such as website or mobile app) requests the indicated page to help improve performance if the user browses to it. This event is logged to indicate that the page content is served to the user's client. This event isn't a definitive indication that the user navigated to the page. When the client renders the page content (as per the user's request), it should generate a ClientViewSignaled event. Not all clients support indicating a prefetch, and therefore some prefetched activities might instead be logged as PageViewed events.
(none)	PageViewedExtended	This activity is related to the "Viewed page" (PageViewed) activity. A PageViewedExtended event is logged when the same person continually views a web page for an extended period (up to three hours). The purpose of logging PageViewedExtended events is to reduce the number of PageViewed events that are logged when a page is continually viewed. This approach helps reduce the noise of multiple PageViewed records for what is essentially the same user activity, and lets you focus on the initial (and more important) PageViewed event.
Accessed file	FileAccessed	User or system account accesses a file. After a user accesses a file, the system doesn't log the FileAccessed event again for the same user and file for the next five minutes.

Friendly name	Operation	Description
Archived file	FileArchived	A user or system account archived a file with Microsoft 365 Archive. When an automatic file archive policy archives the file, the event includes the policy ID, policy version, and the file's last access date.
Changed record status to locked	LockRecord	The record status of a retention label that classifies a document as a record is locked. This status means the document wasn't modified or deleted. Only users assigned at least the contributor permission for a site can change the record status of a document.
Changed record status to unlocked	UnlockRecord	The record status of a retention label that classifies a document as a record is unlocked. This status means that the document can be modified or deleted. Only users assigned at least the contributor permission for a site can change the record status of a document.
Changed retention label for a file	ComplianceSettingChanged	A retention label is applied to or removed from a document. This event is triggered when a retention label is manually or automatically applied to a message.
Checked in file	FileCheckedIn	User checks in a document that they checked out from a document library.
Checked out file	FileCheckedOut	User checks out a document located in a document library. Users can check out and make changes to documents that are shared with them.
Copied file	FileCopied	User copies a document from a site. The copied file can be saved to another folder on the site.
Deleted file	FileDeleted	User deletes a document from a site.
Deleted file from recycle	FileDeletedFirstStageRecycleBin	User deletes a file from the recycle bin of a site.

Friendly name	Operation	Description
bin		
Deleted file from second-stage recycle bin	FileDeletedSecondStageRecycleBin	User deletes a file from the second-stage recycle bin of a site.
Deleted file marked as a record	RecordDelete	A document or email that was marked as a record is deleted. An item is considered as a record when a retention label that marks items as a record is applied to content.
Detected document sensitivity mismatch	DocumentSensitivityMismatchDetected	User uploads a document to a site that's protected with a sensitivity label and the document has a higher priority sensitivity label than the sensitivity label applied to the site. For example, a document labeled Confidential is uploaded to a site labeled General. This event isn't triggered if the document has a lower priority sensitivity label than the sensitivity label applied to the site. For example, a document labeled General is uploaded to a site labeled Confidential. For more information about sensitivity label priority, see Label priority (order matters).
Detected malware in file	FileMalwareDetected	SharePoint anti-virus engine detects malware in a file.
Discarded file checkout	FileCheckOutDiscarded	User discards (or undoes) a checked out file. That means any changes they made to the file when it was checked out are discarded, and not saved to the version of the document in the document library.
Downloaded file	FileDownloaded	User downloads a document from a site.
Modified file	FileModified	User or system account modifies the content or the properties of a document on a site. The system waits five minutes before it logs another

Friendly name	Operation	Description
		FileModified event when the same user modifies the content or properties of the same document.
Moved file	FileMoved	User moves a document from its current location on a site to a new location.
Performed search query	SearchQueryPerformed	User or system account performs a search in SharePoint or OneDrive. Some common scenarios where a service account performs a search query include applying an eDiscovery holds and retention policy to sites and OneDrive accounts, and autoapplying retention or sensitivity labels to site content. To enable logging for this activity, see Get started with auditing solutions.
Priority cleanup move file to recycle bin	PriorityCleanupFileRecycled	An item is moved to Stage 2 recycle bin by a priority cleanup policy on OneDrive or SharePoint Online.
Priority cleanup permanent delete of file	PriorityCleanupFileDeleted	An item is permanently deleted by a priority cleanup policy on OneDrive or SharePoint Online.
Reactivated file	FileUnarchived	A file archived with Microsoft 365 Archive was reactivated and is available to access.
Recycled a file	FileRecycled	User moves a file into the SharePoint Recycle Bin.
Recycled a folder	FolderRecycled	User moves a folder into the SharePoint Recycle Bin.
Recycled all minor versions of file	FileVersionsAllMinorsRecycled	User deletes all minor versions from the version history of a file. The deleted versions are moved to the site's recycle bin.
Recycled all versions of file	FileVersionsAllRecycled	User deletes all versions from the version history of a file. The deleted

Friendly name	Operation	Description
		versions are moved to the site's recycle bin.
Recycled version of file	FileVersionRecycled	User deletes a version from the version history of a file. The deleted version is moved to the site's recycle bin.
Renamed file	FileRenamed	User renames a document.
Requested file reactivation	FileUnarchiveRequested	A user requested to reactivate a file archived with Microsoft 365 Archive.
Restored file	FileRestored	User restores a document from the recycle bin of a site.
Started bulk deleting file versions	FileVersionBulkDeletionStarted	Admin started a bulk deletion of file versions across a SharePoint site or document library to reclaim storage space by permanently removing old versions that exceed the version policy.
Started bulk expiring file versions by using a schedule file	FileVersionBulkExpirationByScheduleStarted	Admin starts a bulk expiration of file versions using a schedule file on a SharePoint site, applying version expiration according to that schedule.
Started generating file version expiration report	FileVersionExpirationReportStarted	Admin started generating a file version expiration report for a SharePoint site or document library that includes certain version metadata and expiration information.
Started tenant apply file version policy	TenantApplyFileVersionPolicyStarted	Admin started applying the tenant-level file version policy across all sites in the organization, which sets version limits or trims existing versions for all document libraries tenant-wide.
Started updating version settings on all document libraries in the site	DocumentLibraryVersionPolicyBulkUpdateStarted	Admin started a bulk update of version policy settings across all document libraries in a SharePoint site.
Uploaded file	FileUploaded	User uploads a document to a folder

Friendly name	Operation	Description
		on a site.
View signaled by client	ClientViewSignaled	A user's client (such as website or mobile app) signals that the indicated page is viewed by the user. This activity is often logged following a PagePrefetched event for a page. **NOTE**: Because ClientViewSignaled events are signaled by the client, rather than the server, it's possible the event might not be logged by the server and therefore might not appear in the audit log. It's also possible that information in the audit record might not be trustworthy. However, because the user's identity is validated by the token used to create the signal, the user's identity listed in the corresponding audit record is accurate. The system waits five minutes before it logs the same event when the same user's client signals that the page is viewed again by the user.
Viewed page	PageViewed	User views a page on a site. This activity doesn't include using a Web browser to view files located in a document library. After a user views a page, the system doesn't log the PageViewed event again for the same user and page for the next five minutes.

Frequently asked questions about FileAccessed and FilePreviewed events

Can non-user activities trigger FilePreviewed audit records that include a user agent like "OneDriveMpc-Transform_Thumbnail"?

There are no known scenarios where non-user actions generate events like these. User actions like opening a user profile card (by selecting their name or email address in a message in Outlook on the web) generate similar events.

Are calls to the OneDriveMpc-Transform_Thumbnail always intentionally triggered by the user?

No. But browser prefetch can result in similar events.

If I see a FilePreviewed event coming from a Microsoft-registered IP address, does that mean the preview shows on the screen of the user's device?

No. Browser prefetch might log the event.

Are there scenarios where a user previewing a document generates FileAccessed events?

Both the FilePreviewed and FileAccessed events indicate that a user's call led to a read of the file (or a read of a thumbnail rendering of the file). While these events are intended to align with preview versus access intention, the event distinction isn't a guarantee of the user's intent.

What causes FileAccessed audit events in Insider Risk Management scenarios?

Creating a case in Insider Risk Management and enabling Content Explorer generates and records a FileAccessed event in the audit log. In this scenario, the event is generated by the Insider Risk Management/Content Explorer workflow rather than by a direct user file-open action, and the same ApplicationId value, `00001111-aaaa-2222-bbbb-3333cccc4444`, is used for these events. This is important for administrators because it helps distinguish these system-generated FileAccessed events from other FileAccessed activity during an investigation. You can use the ApplicationId field as one of the properties to help identify or filter these specific events when reviewing audit log results.

The app@sharepoint user in audit records

In audit records for some file activities (and other SharePoint-related activities), you might notice the user who performed the activity (identified in the User and UserId fields) is app@sharepoint. This user means that an application performed the activity. In this case, the application was granted permissions in SharePoint to perform organization-wide actions (such as searching a SharePoint site or OneDrive account) on behalf of a user, admin, or service. This process of giving permissions to an application is called *SharePoint App-Only* access. This user means that an application, instead of a user, presented the authentication to SharePoint to perform an action. This is why the app@sharepoint user is identified in certain audit records. For more information, see Grant access using SharePoint App-Only.

For example, app@sharepoint is often identified as the user for "Performed search query" and "Accessed file" events. That's because an application with SharePoint App-Only access in your organization performs search queries and accesses files when applying retention policies to sites and OneDrive accounts.

Here are a few other scenarios where app@sharepoint might be identified in an audit record as the user who performed an activity:

- **Microsoft 365 Groups**. When a user or admin creates a new group, audit records are generated for creating a site collection, updating lists, and adding members to a SharePoint group. An application performs these tasks on behalf of the user who created the group.
- **Microsoft Teams**. Similar to Microsoft 365 Groups, audit records are generated for creating a site collection, updating lists, and adding members to a SharePoint group when a team is created.
- **Compliance features**. When an admin implements compliance features, such as retention policies, eDiscovery holds, and autoapplying sensitivity labels.

In these and other scenarios, you might also notice that multiple audit records with app@sharepoint as the specified user were created within a short time frame, often within a few seconds of each other. This pattern also means they were probably triggered by the same user-initiated task. Also, the ApplicationDisplayName and EventData fields in the audit record might help you identify the scenario or application that triggered the event.

Microsoft 365 Archive file-level policy activities

The following table lists Microsoft 365 Archive file-level policy activities in SharePoint and OneDrive that are recorded in the Microsoft 365 audit log. For audit activity generated when an individual file is archived, see File and page activities.

Friendly name	Operation	Description
Added site to file archive policy	FileArchivePolicySiteAdded	An admin added a site to an automatic file archive policy.
Changed file archive policy	FileArchivePolicyChanged	An admin changed an automatic file archive policy.
Completed file archive policy run	FileArchivePolicyRunCompleted	A run of an automatic file archive policy completed. The event includes the run outcome, number of files archived, amount of storage archived, number of sites processed, and whether the policy ran in what-if mode.
Created file archive policy	FileArchivePolicyCreated	An admin created an automatic file archive policy.
Deleted file archive policy	FileArchivePolicyDeleted	An admin deleted an automatic file archive policy.

Friendly name	Operation	Description
Removed site from file archive policy	FileArchivePolicySiteRemoved	An admin removed a site from an automatic file archive policy.
Started file archive policy	FileArchivePolicyStarted	An admin started an automatic file archive policy.

Folder activities

The following table lists the folder activities in SharePoint and OneDrive recorded in the Microsoft 365 audit log. Audit records for some SharePoint activities indicate the app@sharepoint user performed the activity on behalf of the user or admin who initiated the action. For more information, see The app@sharepoint user in audit records.

Friendly name	Operation	Description
Copied folder	FolderCopied	User copies a folder from a site to another location in SharePoint or OneDrive.
Created folder	FolderCreated	User creates a folder on a site.
Deleted folder	FolderDeleted	User deletes a folder from a site.
Deleted folder from recycle bin	FolderDeletedFirstStageRecycleBin	User deletes a folder from the recycle bin on a site.
Deleted folder from second-stage recycle bin	FolderDeletedSecondStageRecycleBin	User deletes a folder from the second-stage recycle bin on a site.
Modified folder	FolderModified	User modifies a folder on a site. This action includes changing the folder metadata, such as changing tags and properties.
Moved folder	FolderMoved	User moves a folder to a different location on a site.
Renamed folder	FolderRenamed	User renames a folder on a site.
Restored folder	FolderRestored	User restores a deleted folder from the recycle bin on a site.

Forms activities

The following table lists the user and admin activities in Microsoft Forms that the Microsoft 365 audit log records. Microsoft Forms is a forms, quiz, and survey tool you can use to collect data for

analysis. Some operations include additional activity parameters, as noted in the descriptions.

If a coauthor or an anonymous responder performs a Forms activity, the audit log records it slightly differently. For more information, see the Forms activities performed by coauthors and anonymous responders section.

Friendly name	Operation	Description
Added form coauthor	AddFormCoauthor	A user uses a collaboration link to help design the form or view responses. This event is logged when a user uses a collaboration URL (not when the collaboration URL is first generated).
Added specific responder	AddSpecificResponder	Form owner adds a new user or group to the specific responders list.
Allowed share form for copy	AllowShareFormForCopy	Form owner creates a template link to share the form with other users. This event is logged when the form owner selects to generate the template URL.
Connected to Excel workbook	ConnectToExcelWorkbook	Connected the form to an Excel workbook. Property ExcelWorkbookLink:string is the associated Excel workbook ID of the current form.
Created a collection	CollectionCreated	Form owner created a collection.
Created comment	CreateComment	Form owner adds comment or score to a quiz.
Created form	CreateForm	Form owner creates a new form. Property DataMode:string is the current form is set to sync with a new or existing Excel workbook if the property value equals DataSync. Property ExcelWorkbookLink:string is the associated Excel workbook ID of the current form.
Created response	CreateResponse	Similar to receiving a new response. A user submitted a response to a form. Property ResponseId:string and Property ResponderId:string is which result is being viewed. For an anonymous responder, the ResponderId property is null.
Created summary link	GetSummaryLink	Form owner creates summary results link to share results.

Friendly name	Operation	Description
Deleted all responses	DeleteAllResponses	Form owner deletes all response data.
Deleted collection from the Recycle Bin	CollectionHardDeleted	Form owner hard-deleted a collection from the Recycle Bin.
Deleted form	DeleteForm	Form owner deletes a form. This action includes SoftDelete (delete option used and form moved to recycle bin) and HardDelete (Recycle bin is emptied).
Deleted Response	DeleteResponse	Form owner deletes one response. Property ResponseId:string is the response being deleted.
Deleted summary link	DeleteSummaryLink	Form owner deletes summary results link.
Disabled anyone can respond setting	DisallowAnonymousResponse	Form owner turns off the setting allowing anyone to respond to the form.
Disabled collaboration	DisableCollaboration	Form owner turns off the setting of collaboration on the form.
Disabled specific people can respond setting	DisableSpecificResponse	Form owner turns off the setting allowing only specific people or specific groups in the current organization to respond to the form.
Disallowed share form for copy	DisallowShareFormForCopy	Form owner deletes template link.
Edited form	EditForm	Form owner edits a form such as creating, removing, or editing a question. The property *EditOperation:string* is the edit operation name. The possible operations are: - CreateQuestion - CreateQuestionChoice - DeleteQuestion - DeleteQuestionChoice - DeleteFormImage - DeleteQuestionImage - UpdateQuestion - UpdateQuestionChoice - UploadFormImage/Bing/Onedrive

Friendly name	Operation	Description
		- UploadQuestionImage - ChangeTheme FormImage includes any place within Forms where the user can upload an image, such as in a query or as a background theme.
Enabled anyone can respond setting	AllowAnonymousResponse	Form owner turns on the setting allowing anyone to respond to the form.
Enabled Office 365 work or school account collaboration	EnableWorkOrSchoolCollaboration	Form owner turns on the setting allowing users with a Microsoft 365 work or school account to view and edit the form.
Enabled people in my organization collaboration	EnableSameOrgCollaboration	Form owner turns on the setting allowing users in the current organization to view and edit the form.
Enabled specific people can respond setting	EnableSpecificResponse	Form owner turns on the setting allowing only specific people or specific groups in the current organization to respond to the form.
Enabled specific people collaboration	EnableSpecificCollaboaration	Form owner turns on the setting allowing only specific people or specific groups in the current organization to view and edit the form.
Exported form	ExportForm	Form owner exports results to Excel. Property ExportFormat:string is if the Excel file is Download or Online.
Listed forms	ListForms	Form owner is viewing a list of forms. Property ViewType:string is which view the form owner is looking at: All Forms, Shared with Me, or Group Forms.
Moved a form into collection	MovedFormIntoCollection	Form owner moved a form into a collection.
Moved a form out of collection	MovedFormOutofCollection	Form owner moved a form out of a collection.

Friendly name	Operation	Description
Moved collection to the Recycle Bin	CollectionSoftDeleted	Form owner moved a collection to the Recycle Bin.
Moved form	MoveForm	Form owner moves a form. Property DestinationUserId:string is the user ID of the person who moved the form. Property NewFormId:string is the new ID for the newly copied form. Property IsDelegateAccess:boolean is the current form move action is performed through the admin delegate page.
Previewed form	PreviewForm	Form owner previews a form by using the Preview function.
Removed form coauthor	RemoveFormCoauthor	Form owner deletes a collaboration link.
Removed specific responder	RemoveSpecificResponder	Form owner removes a user or group from the specific responders list.
Renamed a collection	CollectionRenamed	Form owner changed the name of a collection.
Sent Forms Pro invitation	ProInvitation	User selects to activate a Pro trial.
Submitted response	SubmitResponse	A user submits a response to a form. Property IsInternalForm:boolean is if the responder is within the same organization as the form owner.
Updated a collection	CollectionUpdated	Form owner updated a collection property.
Updated form phishing status	UpdatePhishingStatus	This event is logged whenever the detailed value for the internal security status changes, regardless of whether this change affects the final security state (for example, form is now Closed or Opened). This change might result in duplicate events without a final security state change. The possible status values for this event are: - Take Down - Take Down by Admin - Admin Unblocked - Auto Blocked - Auto Unblocked

Friendly name	Operation	Description
		- Customer Reported - Reset Customer Reported.
Updated form setting	UpdateFormSetting	Form owner updates one or multiple form settings. Property FormSettingName:string is updated sensitive settings' name. Property NewFormSettings:string is updated settings' name and new value. Property thankYouMessageContainsLink:boolean is updated thank-you message contains a URL link.
Updated response	UpdateResponse	Form owner updated a comment or score on a quiz. Property ResponseId:string and Property ResponderId:string is which result is being viewed. For an anonymous responder, the ResponderId property is null.
Updated user phishing status	UpdateUserPhishingStatus	This event is logged whenever the value for the user security status changes. The value of the user status in the audit record is **Confirmed as Phisher** when the user created a phishing form that the Microsoft Online safety team took down. If an admin unblocks the user, the value of the user's status is set to **Reset as Normal User**.
Updated user setting	UpdateUserSetting	Form owner updates a user setting. Property UserSettingName:string is the setting's name and new value.
Viewed form (design time)	ViewForm	Form owner opens an existing form for editing. Property AccessDenied:boolean is access of current form is denied due to permission check. Property FromSummaryLink:boolean is current request comes from the summary link page.
Viewed response	ViewResponse	Form owner views a particular response. Property ResponseId:string and Property ResponderId:string is which result is being viewed. For an anonymous responder, the ResponderId property is null.

Friendly name	Operation	Description
Viewed response page	ViewRuntimeForm	User opened a response page to view. This event is logged regardless of whether the user submits a response or not.
Viewed responses	ViewResponses	Form owner views the aggregated list of responses. Property ViewType:string is whether form owner is viewing Detail or Aggregate.

Forms activities performed by coauthors and anonymous responders

Forms supports collaboration when designing forms and analyzing responses. A form collaborator is known as a *coauthor*. Coauthors can do everything a form owner can do, except delete or move a form. Forms also allows you to create a form that can be responded to anonymously. This means the responder doesn't have to sign in to your organization to respond to a form.

The following table lists the auditing activities and information recorded in the Microsoft 365 audit log for activities performed by coauthors and anonymous responders.

Activity type	Internal or external user	User ID that's logged	Organization logged in to	Forms user type
Coauthoring activities	External	`urn:forms:coauthor#a0b1c2d3@forms.office.com` (The second part of the ID is a hash, which differs for different users)	Form owner's org	Coauthor
Coauthoring activities	External	UPN	Coauthor's org	Coauthor
Coauthoring activities	Internal	UPN	Form owner's org	Coauthor
Response activities	Anonymous	`urn:forms:anonymous#a0b1c2d3@forms.office.com` (The second part of the User ID is a hash, which differs for different users)	Form owner's org	Responder
Response activities	External	`urn:forms:external#a0b1c2d3@forms.office.com` (The second part of the User ID is a hash, which differs for different users)	Form owner's org	Responder
Response activities	External	UPN	Responder's org	Responder

Information barriers activities

The following table lists the activities in Information Barriers that the Microsoft 365 audit log records. For more information about Information Barriers, see Learn about Information Barriers in Microsoft 365.

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

Friendly name	Operation	Description
Applied information barrier mode to site	SiteIBModeSet	A SharePoint or global administrator applied a mode to the site.
Applied segments to site	SiteIBSegmentsSet	A SharePoint, global administrator, or site owner added one or more information barriers segments to a site.
Changed AppBypassInformationBarrier setting for the tenant	AppBypassInformationBarrier	A SharePoint or global administrator changed apps access for SharePoint sites.
Changed information barrier mode of site	SiteIBModeChanged	A SharePoint or global administrator updated the mode of the site.

Friendly name	Operation	Description
Changed segments of site	SiteIBSegmentsChanged	A SharePoint or global administrator changed one or more information barriers segments for a site.
Disabled information barriers for SharePoint and OneDrive	SPOIBIsDisabled	A SharePoint or global administrator disabled information barriers for SharePoint and OneDrive in the organization.
Enabled information barriers for SharePoint and OneDrive	SPOIBIsEnabled	A SharePoint or global administrator disabled information barriers for SharePoint and OneDrive in the organization.
Information barriers insights report completed	InformationBarriersInsightsReportCompleted	System completes build of the information barriers insights report.
Information barriers insights report OneDrive section queried	InformationBarriersInsightsReportOneDriveSectionQueried	An administrator queries the information barriers insights report for

Friendly name	Operation	Description
		OneDrive accounts.
Information barriers insights report scheduled	InformationBarriersInsightsReportSchedule	An administrator schedules the information barriers insights report.
Information barriers insights report SharePoint section queried	InformationBarriersInsightsReportSharePointSectionQueried	An administrator queries the information barriers insights report for SharePoint sites.
Removed segment from site	SiteIBSegmentsRemoved	A SharePoint or global administrator removes one or more information barriers segments from a site.

Microsoft 365 Admin Center Agent Management activities

The following table lists Agent Management activities that the Microsoft 365 audit log records. For more information, see Manage Microsoft 365 Copilot agents in the Microsoft 365 admin center.

Friendly name	Operation	Description
Blocked Agent	BlockedAgent	An admin blocked an agent. Users in the organization can't use the agent.
Deleted Agent	DeletedAgent	An admin deleted a shared agent that deletes the underlying files and the agent.

Friendly name	Operation	Description
Deployed Agent	DeployedAgent	An admin deployed an agent. The agent becomes active and usable for specific users, groups, or the entire organization.
Removed Agent	RemovedAgent	An admin removed a previously installed agent for the whole organization or for specific users and groups.
Unblocked Agent	UnblockedAgent	An admin unblocked a previously blocked agent.
Updated Agent	UpdatedAgent	An admin updated a custom agent by uploading the latest manifest file.
Updated Tenant-level Agent Settings	UpdatedTenantSettings	An admin updated tenant level settings that apply to agents.

Microsoft 365 Apps Admin Services cloud policy activities

The following table lists the activities for policy configuration changes in the Cloud Policy service that the Microsoft 365 audit log records.

Friendly name	Operation	Description
Created policy configuration	CreatedPolicyConfig	A new policy configuration was created.
Deleted policy configuration	DeletedPolicyConfig	A policy configuration was deleted.
Updated policy configuration	UpdatedPolicyConfig	An existing policy configuration was updated, for example by adding, changing, or removing policy settings, or by changing the policy configuration's name, description, or scope.
Updated policy configuration priority	UpdatedPolicyConfigPriority	The priority of a policy configuration was changed.

Microsoft 365 Apps Admin Services cloud update activities

The following table lists the activities for configuration changes and triggered actions in the Cloud Update service that the Microsoft 365 audit log records.

Friendly name	Operation	Description
Device configuration updated	updateddeviceconfiguration	An action for a device managed by Cloud Update was triggered. This action includes triggering a rollback, resuming a rolled back device, or changing the update channel.
Profile configuration updated	updatedprofileconfiguration	A Cloud Update profile's configuration changed. This change includes updates to the profile's state, set deadline, Update Validation enablement, and configured waves and wave delay.
Tenant configuration updated	updatedtenantconfiguration	The Cloud Update's organization-wide configuration changed. This change includes updates to exclusions, exclusion windows, and generating a new Tenant Association Key (TAK).

Microsoft 365 Backup activities

The following table lists the activities in Microsoft 365 Backup that the Microsoft 365 audit log records. Microsoft 365 Backup is designed to ensure your organization's data is always protected and easily recoverable. For more information about Microsoft 365 Backup, see Overview of Microsoft 365 Backup.

Friendly name	Operation	Description
Activated a Backup Policy	BackupPolicyActivated	A Microsoft 365 Backup policy is activated from an inactive state.
Activated Draft Restore Task	RestoreTaskActivated	A draft restore task for Microsoft 365 Backup is activated. This is a long running operation.
Backup Item Created	BackupItemAdded	One or more backup items are added to a Microsoft 365 Backup policy.
Backup Item Removed	BackupItemRemoved	One or more backup items are removed from a Microsoft 365 Backup policy.
Browsed Site for Restore	SiteBrowsedForRestore	A site was browsed in Microsoft 365 Backup to identify and select items for granular restore.
Cancel Offboard Backup Item	CancelOffboardBackupItem	Offboarding canceled for a backup item.
Completed Granular Restore for Backup Item	BackupItemGranularRestoreCompleted	A granular restore for one or more backup items is completed in Microsoft 365 Backup.

Friendly name	Operation	Description
Completed Granular Restore Session	GranularRestoreSessionCompleted	A granular restore session is completed in Microsoft 365 Backup.
Completed Restore Task	RestoreTaskCompleted	A restore task is completed in Microsoft 365 Backup.
Created Draft Restore Task	DraftRestoreTaskCreated	A restore task is created in Microsoft 365 Backup. By default, the restore task is created as draft.
Created Granular Browse Session	GranularBrowseSessionCreated	A granular browse session is created in Microsoft 365 Backup to enable browsing of backed-up content for item-level restore.
Created Granular Restore Session	GranularRestoreSessionCreated	A granular restore session is created in Microsoft 365 Backup.
Created new Backup policy	NewBackupPolicyCreated	A new Microsoft 365 Backup policy was created by a Global Admin. By default, a backup policy is created in an inactive state.
Deleted a Backup Policy	BackupPolicyDeleted	A Microsoft 365 Backup policy is deleted.
Deleted Draft Restore Task	DraftRestoreTaskDeleted	A draft restore task is deleted in Microsoft 365 Backup.
Edited a Backup Policy	BackupPolicyEdited	A Microsoft 365 Backup policy was updated. You can update a backup policy by performing any of the following actions: 1. Renaming the policy. 2. Add one or more protection units. 3. Removing one or more protection units.
Edited Draft Restore Task	DraftRestoreTaskEdited	A draft restore task is edited in Microsoft 365 Backup.
Offboard Backup Item	OffboardBackupItem	Backup item is offboarding.
Paused a Backup Policy	BackupPolicyPaused	A Microsoft 365 Backup policy is paused from active state.
Programmatically got Backup Item	GetBackupItem	User requests the protection unit backed up using Microsoft 365 Backup

Friendly name	Operation	Description
		programmatically.
Programmatically got Details of Backup Policy	ViewBackupPolicyDetails	Details of a Microsoft 365 Backup policy are accessed programmatically.
Programmatically got Details of Restore Task	GetRestoreTaskDetails	A user requests details of restore task by its identifier in Microsoft 365 Backup.
Programmatically got list of all Backup Policies	ListAllBackupPolicies	A user programmatically gets the list of all backup policies backed up using Microsoft 365 Backup.
Programmatically got list of Backup Items in Backup Policies	ListAllBackupItemsInPolicies	A list of all protection units present in a backup policy in Microsoft 365 Backup is requested programmatically.
Programmatically got list of Backup Items in Tenant	ListAllBackupItemsInTenant	A user programmatically gets list of all protection units in the organization backed up using Microsoft 365 Backup.
Programmatically got list of Backup Items in Workload	ListAllBackupItemsInWorkload	A user programmatically gets list of all protection units in workload (SharePoint, OneDrive, or Exchange) backed up using Microsoft 365 Backup.
Programmatically got list of Restore Items in Restore Task	GetAllRestoreArtifactsInTask	A user programmatically gets all artifacts in a restore task in Microsoft 365 Backup.
Programmatically got list of Restore Points	ListAllRestorePoints	A user programmatically gets all restore points in Microsoft 365 Backup. Restore Points represent the timestamp when an artifact is protected (per Protection Policy). Only Global Admins have access.
Programmatically got list of Restore Tasks	ListAllRestoreTasks	A user programmatically gets the list of existing restore sessions in Microsoft 365 Backup. This list includes the restore session created for every service type enrolled in the organization.
Restore item restoration complete	BackupItemRestoreCompleted	The restoration of an item backed up by Microsoft 365 Backup is completed.
Restore item restoration triggered	BackupItemRestoreTriggered	Restore is triggered for an item backed up with Microsoft 365 Backup.
Triggered Granular Restore for Backup	BackupItemGranularRestoreTriggered	A granular restore is triggered for a backup item in a granular restore session

Friendly name	Operation	Description
Item		in Microsoft 365 Backup.
Triggered Granular Restore Session	GranularRestoreSessionTriggered	A granular restore session is triggered in Microsoft 365 Backup. This is a long running operation.

Microsoft 365 Copilot admin activities

The following table lists admin activities related to Microsoft 365 Copilot and Microsoft 365 Copilot Chat that the Microsoft 365 audit log records. You can access Copilot across Microsoft services. The audit log records activities that show how and when users interact with Copilot. This information includes the Microsoft service where the activity took place and references to the files accessed during the interaction.

To access the text from the user's prompt during the interaction, see Content Search or view the **AI interaction** event from the activity explorer in Data Security Posture Management for AI.

For more information about auditing and other compliance management options for Microsoft 365 Copilot, see Use Microsoft Purview to manage data security & compliance for Microsoft 365 Copilot & Microsoft 365 Copilot Chat.

For more information about events related to user interactions with AI, see Audit logs for Copilot and AI activities.

Friendly name	Operation	Description
Created a new Copilot plugin	CreatePlugin	A user (or admin or system on behalf of a user) created a new Copilot plugin.
Created a new Copilot promptbook	CreatePromptBook	A user (or admin or system on behalf of a user) created a new promptbook in Copilot.
Deleted a Copilot plugin	DeletePlugin	A user (or admin or system on behalf of a user) deleted a Copilot plugin.
Deleted a Copilot promptbook	DeletePromptBook	A user (or admin or system on behalf of a user) deleted a Copilot promptbook.
Disabled a Copilot plugin	DisableCopilotPlugin	A user (or admin or system on behalf of a user) disabled a Copilot plugin.
Disabled a Copilot promptbook	DisablePromptBook	A user (or admin or system on behalf of a user) disabled a Copilot promptbook.

Friendly name	Operation	Description
Enabled a Copilot plugin	EnablePlugin	A user (or admin or system on behalf of a user) enabled a Copilot plugin.
Enabled a Copilot promptbook	EnablePromptBook	A user (or admin or system on behalf of a user) enabled a Copilot promptbook.
Exported AI Interactions	AIEnterpriseInteractionsExported	An admin exported interactions with Copilot.
Interacted with Copilot	CopilotInteraction	A user (or admin or system on behalf of a user) entered prompts in Copilot.
Sent change notification for AI interaction creation	AIInteractionCreatedNotification	A change notification was sent to notify a subscribed listener application of a new Copilot AI interaction.
Sent change notification for AI interaction deletion	AIInteractionDeletedNotification	A change notification was sent to notify a subscribed listener application of a deleted Copilot AI interaction.
Sent change notification for AI interaction update	AIInteractionUpdatedNotification	A change notification was sent to notify a subscribed listener application of an updated Copilot AI interaction.
Subscribed to AI interactions	SubscribedToAIInteractions	A subscription was created by a listener application to receive change notifications for Copilot AI interactions.
Updated a Copilot plugin setting	UpdatePlugin	A user (or admin or system on behalf of a user) updated a Copilot plugin setting.
Updated a Copilot promptbook setting	UpdatePromptBook	A user (or admin or system on behalf of a user) updated a Copilot promptbook setting.
Updated a Copilot setting	UpdateTenantSettings	An admin (or system on behalf of an admin) updated a Copilot setting.

Microsoft 365 Copilot scheduled prompts activities

By using Copilot Scheduled Prompts, users can automate Copilot prompts to run on a defined schedule in Microsoft 365 Copilot chat. The following table lists the activities recorded in the Microsoft 365 audit log. Admins can manage this feature in your organization. For more information, see Manage scheduled prompts for Microsoft 365 Copilot.

Friendly name	Operation	Description
Created a scheduled prompt	ScheduledPromptCreated	A user scheduled a Copilot prompt from Microsoft 365 Copilot Chat (Office.com chat).

Friendly name	Operation	Description
Deleted a scheduled prompt	ScheduledPromptDeleted	A user deletes a scheduled prompt from the management window. The prompt is removed from the list and no longer runs.
Executed a scheduled prompt	ScheduledPromptExecute	Logged at the start of each scheduled run. It doesn't confirm the prompt was completed.

Microsoft Defender for Cloud activities

The following table lists Microsoft Defender for Cloud activities that are logged in the Microsoft 365 audit log. For more information, see What is Microsoft Defender for Cloud?.

Friendly name	Operation	Description
Connector created	ConnectorCreated	A user created a cloud environment connector for AWS, Azure, Google Cloud Platform (GCP), GitHub, JFrog, or Azure DevOps.
Connector deleted	ConnectorDeleted	A user deleted a cloud environment connector.
Connector modified	ConnectorModified	A user updated the configuration of an existing cloud environment connector.
Enablement rule collection created	EnablementRuleCollectionCreated	A user created an enablement rule collection to configure Defender plans and extensions for an environment.
Enablement rule collection deleted	EnablementRuleCollectionDeleted	A user deleted an enablement rule collection.
Enablement rule collection modified	EnablementRuleCollectionModified	A user modified an enablement rule collection to change Defender plan or extension settings.
Global setting modified	GlobalSettingModified	A user modified a global setting.
Rule collection created	RuleCollectionCreated	A user created a rule collection.
Rule collection deleted	RuleCollectionDeleted	A user deleted a rule collection and its associated rules.
Rule collection disabled	RuleCollectionDisabled	A user disabled an active rule collection without deleting it.
Rule collection enabled	RuleCollectionEnabled	A user enabled a previously disabled rule collection.

Friendly name	Operation	Description
Rule collection modified	RuleCollectionModified	A user modified a rule collection.
Rule created	RuleCreated	A user created a rule.
Rule deleted	RuleDeleted	A user deleted a rule.

Microsoft Defender for Endpoint general settings activities

The following table lists the activities for Microsoft Defender for Endpoint general settings recorded in the Microsoft 365 audit log. For more information about the settings Microsoft Defender for Endpoint, see Configure general Defender for Endpoint settings.

Friendly name	Operation	Description
Changed data retention	ChangeDataRetention	Edited the data retention settings for Defender for Endpoint.
Downloaded offboarding package	DownloadOffboardingPkg	Downloaded the package used to remove devices from Defender for Endpoint.
Downloaded onboarding package	DownloadOnboardingPkg	Downloaded the package used to onboard devices to Defender for Endpoint.
Set advanced features	SetAdvancedFeatures	Changed advanced features in Defender for Endpoint, enabling more precise controls during an incident. Only settings for advanced features that are generally available are logged in the Microsoft 365 audit log.

Microsoft Defender for Endpoint indicator settings activities

The following table lists the activities for Microsoft Defender for Endpoint indicator settings that the Microsoft 365 audit log records. For more information about Microsoft Defender for Endpoint indicators, see Manage indicators.

Friendly name	Operation	Description
Added indicator	AddIndicator	Created a new indicator of compromise used to track attacks and events.

Friendly name	Operation	Description
Deleted indicator	DeleteIndicator	Removed an indicator of compromise.
Edited indicator	EditIndicator	Edited an indicator of compromise.

Microsoft Defender for Endpoint response actions activities

The following table lists the activities for Microsoft Defender for Endpoint response actions, including Live response, that the Microsoft 365 audit log records. For more information about Microsoft Defender for Endpoint response actions, see Take response actions on a device.

Friendly name	Operation	Description
Collected investigation package	CollectInvestigationPackage	Collected information about a device that's used to understand attack tools and techniques.
Collected Logs	LogsCollection	Collected log information about Defender for Endpoint.
Downloaded file	DownloadFile	Downloaded a suspicious file for further investigation.
Isolated device	IsolateDevice	Isolated a device from the network, helping prevent attacks from spreading.
Offboarded device	DeviceOffBoarding	Removed a device from Defender for Endpoint.
Ran Antivirus scan	RunAntiVirusScan	Ran Microsoft Defender Antivirus scan on a device.
Ran get Live response file link	LiveResponseGetFile	Opened a Live response session, opening a remote shell into a device.
Ran Live response API	RunLiveResponseApi	Opened a Live response session via an API call, opening a remote shell into a device.
Ran Live response session	RunLiveResponseSession	Ran a Live response session, opening a remote shell into a device.
Released from isolation	ReleaseFromIsolation	Added an isolated device back onto the network.
Removed app restrictions	RemoveAppRestrictions	Allow an app to run.
Restricted app execution	RestrictAppExecution	Prevent a malicious app from running.

Friendly name	Operation	Description
Stopped and quarantined file	StopAndQuarantineFile	Quarantined a suspicious file for further analysis.

Microsoft Defender for Endpoint roles settings activities

The following table lists the activities for Microsoft Defender for Endpoint role settings that the Microsoft 365 audit log records. For more information about roles in Microsoft Defender for Endpoint, see Create and manage roles for role-based access control.

Friendly name	Operation	Description
Added role	AddRole	Added new security role and permissions.
Deleted role	DeleteRole	Removed security roles and permissions.
Edited role	EditRole	Edited security roles and permissions.

Microsoft Defender for Experts activities

The following table lists the activities in Microsoft Defender Experts that the Microsoft 365 audit log records. For more information about Microsoft Defender Experts, see Learn about Microsoft Defender Experts for XDR and Learn about Microsoft Defender Experts for Hunting.

Friendly name	Operation	Description
Defender Experts analyst permission created	DefenderExpertsAnalystPermissionCreated	An administrator granted one or more role permissions to Defender Experts analysts to investigate incidents or remediate threats.
Defender Experts analyst permission modified	DefenderExpertsAnalystPermissionModified	An administrator modified role permissions for Defender Experts analysts to investigate incidents or remediate threats.

Microsoft Defender for Identity activities

The following table lists the activities for Microsoft Defender for Identity that the Microsoft 365 audit log records.

Friendly name	Operation	Description
Added alert notification recipient	AlertNotificationsRecipientAdded	A recipient was added to the security alerts notification configuration.
Added exclusions configuration	ExclusionConfigurationAdded	A global exclusion was added for an alert or for an entity.
Added health issues notification recipient	MonitoringAlertNotificationRecipientAdded	A recipient was added to the health issues notification configuration.
Added remediation action	RemediationActionAdded	A remediation action was added to the queue.
Added sensor	SensorCreated	A new sensor was added.
Created workspace	WorkspaceCreated	The workspace was created.
Deleted alert notification recipient	AlertNotificationsRecipientDeleted	A recipient was removed from the security alerts notification configuration.
Deleted exclusions configuration	ExclusionConfigurationDeleted	A global exclusion was deleted for an alert or for an entity.
Deleted health issues notification recipient	MonitoringAlertNotificationRecipientDeleted	A recipient was removed from the heath issues notification configuration.
Deleted workspace	WorkspaceDeleted	The workspace was deleted.
Downloaded Domain Controller coverage Excel	DomainControllerCoverageExcelDownloaded	The Domain Controller coverage Excel file was downloaded.
Downloaded report	ReportDownloaded	A report was downloaded.
Downloaded security alert Excel	AlertExcelDownloaded	The detailed Excel file of an alert was downloaded.
Regenerated sensor deployment key	SensorDeploymentAccessKeyUpdated	The sensors access key was regenerated.
Removed sensor	SensorDeleted	A sensor was deleted.
Retrieved sensor deployment key	SensorDeploymentAccessKeyReceived	The sensors access key was retrieved.
Updated alert threshold configuration	WorkspaceAlertThresholdLevelUpdated	The alerts thresholds configuration was updated.

Friendly name	Operation	Description
Updated directory services account configuration	DirectoryServicesAccountConfigurationUpdated	The Directory services accounts set was modified.
Updated entity remediation configuration	EntityRemediatorConfigurationUpdated	The Manage action accounts mode was modified.
Updated entity tags	TaggingConfigurationUpdated	A tag was added or removed from an entity.
Updated exclusions configuration	ExclusionConfigurationUpdated	A global exclusion was updated for an alert or for an entity.
Updated health issue	MonitoringAlertUpdated	A health issue was modified.
Updated remediation action	RemediationActionUpdated	A remediation action was completed.
Updated remediation action configuration	RemediationActionConfigurationUpdated	The Manage action accounts set was modified.
Updated reporter configuration	ReporterConfigurationUpdated	A report's scheduling was modified.
Updated security notification configuration	NotificationConfigurationUpdated	The security alerts or health issues notifications configuration was modified.
Updated sensor configuration	SensorConfigurationUpdated	A sensor's configuration was updated.
Updated sensors activation mode	SensorActivationMethodConfigurationUpdated	The sensors activation mode was modified.
Updated syslog forwarding configuration	SyslogServiceConfigurationUpdated	The syslog forwarding configuration was modified.
Updated Unified RBAC configuration	URbacAuthorizationStatusChanged	The Unified Role Based Access Control configuration was modified.
Updated VPN configuration	VpnConfigurationUpdated	The VPN (radius accounting) configuration was modified.

Microsoft Defender for Office 365 activities

The following table lists the activities for Microsoft Defender for Office 365 logged in the Microsoft 365 audit log. For more information about Microsoft Defender for Office 365, see Microsoft Defender for Office 365 overview.

> **Note**
>
> These events aren't listed as friendly names in audit search. This table is a complementary resource for the Office 365 Management Activity API schema.

Events	Operation	Description
Alert events	SecurityComplianceAlerts	Security and compliance alert signals for alerts generated based on Alert policies in the Microsoft Purview compliance portal.
Attack Simulation admin events	AttackSimAdmin	Events related to admin activities in Attack Simulation & Training in Microsoft Defender for Office 365.
Attack Simulation training events	UserTraining	Events related to user training in Attack Simulation & Training in Microsoft Defender for Office 365.
Attack Simulation user events	AttackSim	Events related to user activities in Attack Simulation & Training in Microsoft Defender for Office 365.
Automated incident response (AIR) events	AirInvestigation	Automated incident response (AIR) events and data.
Campaign events	Campaign	Email campaign events from Microsoft Defender for Office 365.
Content detection events	ThreatIntelligenceAtpContent	Phishing and malware events for files in SharePoint, OneDrive for Business, and Microsoft Teams from Microsoft Defender for Office 365.
Outbound spam events	HygieneEvent	Events related to outbound spam protection.
Phishing and malware events	ThreatIntelligence	Phishing and malware events from the built-in security add-on for on-premises mailboxes (formerly Exchange Online Protection) and Microsoft Defender for Office 365.
Quarantine events	Quarantine	Quarantine events from Microsoft Defender for Office 365.
Submission events	MailSubmission	Admin and user submission events from the built-in security add-on for on-premises mailboxes (formerly

Events	Operation	Description
		Exchange Online Protection) and Microsoft Defender for Office 365.
URL Click Events	ThreatIntelligenceUrl	Safe links time-of-block and block override events from Microsoft Defender for Office 365.

Microsoft Defender XDR automation activities

The following table lists automation activities for Microsoft Defender XDR recorded in the Microsoft 365 audit log. For more information, see Search the audit log for events in Microsoft Defender XDR.

Friendly name	Operation	Description
Created a generated playbook	GeneratedPlaybookCreated	A user created a generated playbook.
Created an automation rule	AutomationRuleCreated	A user created an automation rule.
Deactivated a generated playbook	GeneratedPlaybookInactive	A user deactivated a generated playbook.
Deactivated an automation rule	AutomationRuleInactive	A user deactivated an automation rule.
Edited a generated playbook	GeneratedPlaybookEdited	A user edited a generated playbook.
Edited an automation rule	AutomationRuleEdited	A user edited an automation rule.
Activated a generated playbook	GeneratedPlaybookActivated	A user activated a generated playbook.
Activated an automation rule	AutomationRuleActivated	A user activated an automation rule.
Removed a generated playbook	GeneratedPlaybookRemoved	A user removed a generated playbook.
Removed an automation rule	AutomationRuleRemoved	A user removed an automation rule.
Triggered a generated playbook	TriggerGeneratedPlaybook	A user manually triggered a generated playbook.

Microsoft Defender XDR case management activities

The following table lists the case management activities that are logged in the Microsoft 365 audit log when users create and manage cases in the Microsoft Defender portal. For more information, see Microsoft Defender XDR.

Friendly name	Operation	Description
A case was assigned to a different user or team member	ChangedCaseAssignment	Generated when a case is assigned or reassigned to a different user.
A comment in a case was edited	UpdatedCommentInCase	Generated when an existing comment on a case is edited.
A comment was added to a case	AddedCommentToCase	Generated when a comment is added to a case.
A comment was deleted from a case	DeletedCommentFromCase	Generated when a comment is deleted from a case.
A new case was created in the system	CreatedCase	Generated when a user creates a new case. Not logged for cases created by application or service accounts.
An existing case was deleted from the system	DeletedCase	Generated when a case is deleted.
An object was linked to a case	LinkedObject	Generated when an object, such as an alert or entity, is linked to a case.
An object was unlinked from a case	UnlinkedObject	Generated when an object is unlinked from a case.
Case content was changed	UpdatedCase	Generated when a case is updated with changes, such as custom fields, that aren't captured by a more specific operation.
One or more tags were added to a case	AddedTagsToCase	Generated when one or more tags are added to a case.
One or more tags were removed from a case	RemovedTagsFromCase	Generated when one or more tags are removed from a case.
The classification of a case was changed	ChangedCaseClassification	Generated when the classification of a case is changed.
The closing notes of a case were modified or updated	ChangedCaseClosingNotes	Generated when the closing notes of a case are changed.
The description of a case was modified or updated	ChangedCaseDescription	Generated when the description of a case is changed.
The name or title of a case was changed	ChangedCaseName	Generated when the name (title) of a case is changed.
The priority level of a case was modified	ChangedCasePriority	Generated when the priority of a case is changed.

Friendly name	Operation	Description
The severity of a case was changed	ChangedCaseSeverity	Generated when the severity of a case is changed.
The status of a case was updated	ChangedCaseStatus	Generated when the status of a case is changed.

Microsoft Defender XDR custom detection activities

The following table lists the custom detection activities for Microsoft Defender XDR recorded in the Microsoft 365 audit log. For more information about Microsoft Defender XDR custom detections, see Create and manage custom detections rules.

Friendly name	Operation	Description
Changed custom detection rule status	ChangeCustomDetectionRuleStatus	A custom detection rule was turned off or on.
Created custom detection rule	CreateCustomDetection	A new custom detection rule was created.
Deleted custom detection rule	DeleteCustomDetection	A custom detection rule was removed.
Edited custom detection rule	EditCustomDetection	A custom detection rule was modified.
Ran custom detection rule	RunCustomDetection	A custom detection rule was manually run.

Microsoft Defender XDR incident activities

The following table lists the incident activities for Microsoft Defender XDR recorded in the Microsoft 365 audit log. For more information about incidents in Microsoft Defender XDR, see Incidents overview.

Friendly name	Operation	Description
Added comment to incident	AddCommentToIncident.	Added comment to incident.
Added tags to incident	AddTagsToIncident	Added tags to incident.
Assigned user to incident	AssignUserToIncident	Assigned user to incident.
Edit incident classification	EditIncidentClassification	Edit incident classification.
Removed tags from incident	RemoveTagsFromIncident	Removed tags from the incident.

Friendly name	Operation	Description
Unassigned user to incident	UnAssignUserFromIncident	Unassigned user to incident.
Updated incidents status	UpdateIncidentStatus	Updated incidents status.

Microsoft Defender XDR suppression rule activities

The following table lists the activities for suppression rules for Microsoft Defender XDR recorded in the Microsoft 365 audit log. For more information about suppression rules in Microsoft Defender XDR, see Manage suppression rules.

Friendly name	Operation	Description
Created suppression rule	CreateSuppressionRule	Created alert suppression rule.
Deleted suppression rule	DeleteSuppressionRule	Deleted alert suppression rule.
Disabled suppression rule	DisableSuppressionRule	Disabled alert suppression rule.
Edited suppression rule	EditSuppressionRule	Deleted alert suppression rule.
Enabled suppression rule	EnableSuppressionRule	Enabled alert suppression rule.

Microsoft Edge WebContentFiltering activities

The following table lists the browsing activities in Microsoft Edge recorded in the Microsoft 365 audit log for the WebContentFiltering feature. The table includes the friendly name that displays in the **Activities** column and the name of the corresponding operation that appears in the detailed information of an audit record and in the .csv file when you export the search results.

Search the audit log describes how you can search the audit logs from the Microsoft Purview portal. Users need to be a global admin or have audit read permissions to access audit logs. Use the **Activities** filter to search for specific activities. To list all WebContentFiltering activities, select *WebContentFiltering* in the **Record type** filter. Use the date range boxes and the **Users** list to scope the search results further if needed.

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

Friendly name	Operation	Description
Allowed URL Navigation in Microsoft Edge	URLNavigationAllowed	User browsed a URL.
Blocked URL Navigation in Microsoft Edge	URLNavigationBlocked	Browsed URL is blocked by WebContentFiltering policy.

Microsoft Entra group administration activities

The following table lists group administration activities that the Microsoft 365 audit log records when an admin or a user creates or changes a Microsoft 365 group or when an admin creates a security group by using the Microsoft 365 admin center or the Azure management portal. For more information about groups in Microsoft 365, see View, create, and delete Groups in the Microsoft 365 admin center.

Friendly name	Operation	Description
Added group	AddGroup	A group was created.
Added member to group	AddMemberToGroup.	A member was added to a group.
Deleted group	Delete group.	A group was deleted.
Removed member from group	RemoveMemberFromGroup.	A member was removed from a group.
Updated group	UpdateGroup.	A property of a group was changed.

Microsoft Entra risk detection activities

The following table lists risk detection activities recorded in the Microsoft 365 audit log when using Microsoft Entra ID Protection. For more information about risk detection, see Learn about risk detections..

Friendly name	Operation	Description
Compromised sign-in detection	CompromisedSignIn	Sign-in risk detections represent the probability that a given authentication request isn't the authorized owner of the account.
Compromised user detection	CompromisedUser	User risk detections might flag a legitimate user account as at risk, when a potential threat actor gains access to an account by compromising their credentials or when they detect some type of anomalous user activity.

Microsoft Graph Data Connect

The following table lists the user and admin activities in Microsoft Graph Data Connect recorded in the Microsoft 365 audit log. The table includes the friendly name that displays in the **Activities** column and the name of the corresponding operation that appears in the detailed information of an audit record and in the CSV file when you export the search results.

Friendly name	Operation	Description
Approved or denied an app	ConsentModificationRequest	A user performed a consent modification request.
Extraction ran	DataAccessRequestOperation	A user performed an extraction. Partner datasets and ISV runs are excluded.

Microsoft Places Directory activities

The following table lists the admin activities in Microsoft Places Directory that the Microsoft 365 audit log records. The table includes the friendly name that displays in the **Activities** column and the name of the corresponding operation that appears in the detailed information of an audit record and in the CSV file when you export the search results.

Friendly name	Operation	Description
Created a place	CreatedPlace	An admin performed an operation for creating places.
Deleted a place	DeletedPlace	An admin performed an operation for deleting places.
Updated a place	UpdatedPlace	An admin performed an operation for updating places.

Microsoft Purview Audit solution activities

The following table lists activities associated with audit solutions in the Microsoft Purview portal and the Audit Search Graph API. The *SecurityComplianceCenter* workload audits these activities. For more information, see Search the audit log.

Audit search and export activities that use Search-UnifiedAuditLog aren't audited.

Friendly name	Operation	Description
Audit search canceled	AuditSearchCancelled	An audit search job was canceled.
Audit search completed	AuditSearchCompleted	An audit search job was completed.
Audit search created	AuditSearchCreated	A new audit search request was submitted.
Audit search deleted	AuditSearchDeleted	An audit search job was deleted.

Friendly name	Operation	Description
Audit search export job completed	AuditSearchExportJobCompleted	The export job to export the search results of an audit search query was completed.
Audit search export job created	AuditSearchExportJobCreated	An export job was created to export the search results of an audit search query.
Audit search export results downloaded	AuditSearchExportResultsDownloaded	The search results from an audit search query were downloaded.

Microsoft Purview governance activities

The following table lists Microsoft Purview governance activities recorded in the Microsoft 365 audit log. For more information, see Microsoft Purview governance solutions.

Friendly name	Operation	Description
Asset or entity created	EntityCreated	A new asset or entity was created or added to an existing asset.
Classification added	ClassificationAdded	Added classifications to asset entity.
Classification definition created	ClassificationDefinitionCreated	Created a classification type.
Classification definition deleted	ClassificationDefinitionDeleted	Deleted a classification type.
Classification definition updated	ClassificationDefinitionUpdated	Updated a classification type.
Classification deleted	ClassificationDeleted	Deleted classifications from asset entity.
Classification updated	ClassificationUpdated	Updated classifications on asset entity.
Entity deleted	EntityDeleted	An asset or entity was deleted from an existing asset.
Entity updated	EntityUpdated	**Includes**: attribute update, business attribute update, collection info (only include collection ID), contacts update, customAttributes, hierarchy, homeId, labels, sourceDetails.
Glossary term assigned	GlossaryTermAssigned	Assigned terms to asset entity.

Friendly name	Operation	Description
Glossary term created	GlossaryTermCreated	Created a term.
Glossary term deleted	GlossaryTermDeleted	Deleted a term.
Glossary term disassociated	GlossaryTermDisassociated	Disassociated terms from asset entity.
Glossary term updated	GlossaryTermUpdated	Updated a term.
Sensitivity label changed	SensitivityLabelChanged	Sensitivity label was added, updated, or deleted.

Microsoft Purview on-demand classification activities

The following table lists the activities from Microsoft Purview on-demand classification that the Microsoft 365 audit log records. For more information about on-demand classification, see Learn about on-demand classification in Microsoft Purview.

> **Note**
>
> These audit activities are only available in Audit (Premium). You must be assigned the appropriate license before these activities are logged in the audit log. For more information, see Audit (Premium). For Audit (Premium) licensing requirements, see Auditing solutions in Microsoft 365.

Friendly name	Operation	Description
Classified file as part of an on-demand classification scan	DataScanClassification	File was evaluated for sensitive content as part of an on-demand classification scan for SharePoint or OneDrive.
Classified file on device as part of an on-demand classification scan	SensitiveInfoDiscovered	File was evaluated for sensitive content as part of an on-demand classification scan for Endpoint devices.
Declassified file as part of an on-demand classification scan	DataScanDeClassification	File no longer matches the sensitive information types defined in the corresponding On-demand classification scan triggered for SharePoint or OneDrive locations.

Microsoft Purview permission activities

The following table lists Microsoft Purview administration activities that the Microsoft 365 audit log records when an administrator manages role-based access control (RBAC) role groups and user assignments. For more information, see Permissions in the Microsoft Purview portal. See Office 365 Management Activity API schema for the audit log schema.

Friendly name	Operation	Description
Create RBAC role group definition	CreateRBACRoleGroupDefinition	An RBAC role group definition is created.
Delete RBAC role group definition	DeleteRBACRoleGroupDefinition	An RBAC role group definition is deleted.
Update RBAC role group definition	UpdateRBACRoleGroupDefinition	An RBAC role group definition is updated.
Add user assignments to role group	GrantPermissionsAsync	User assignments to an RBAC role group are added.
Remove user assignments to role group	DeletePermissionAsync	User assignments to an RBAC role group are removed.
Updated user role assignments expiration in role group	UpdatePermission	A user assignment expiration date is updated in an RBAC role group.

Microsoft Security Copilot agent management

The following table lists the Agent management operations in Security Copilot that the Microsoft 365 audit log records. These activities characterize activities for agents and the components necessary for them to function such as triggers. For more information about Security Copilot agent management, see Microsoft Security Copilot agents overview.

Friendly name	Operation	Description
Created a new Security Copilot agent	CreateCopilotAgent	A user (or admin or system on behalf of a user) created a new Security Copilot agent.
Created a new Security Copilot agent trigger	CreateCopilotforSecurityAgentTrigger	A user (or admin or system on behalf of a user) created a new promptbook in Copilot.
Deleted a Security Copilot agent	DeleteCopilotAgent	A user (or admin or system on behalf of a user) deleted a Security Copilot agent.

Friendly name	Operation	Description
Deleted a Security Copilot agent trigger	DeleteCopilotForSecurityAgentTrigger	A user (or admin or system on behalf of a user) deleted a Security Copilot trigger.
Updated a Security Copilot agent setting	UpdateCopilotAgent	A user (or admin or system on behalf of a user) updated a Security Copilot agent setting.
Updated a Security Copilot agent trigger	UpdateCopilotforSecurityAgentTrigger	A user (or admin or system on behalf of a user) updated a Security Copilot agent trigger setting.

Microsoft Security Copilot interactions

The following table lists the types of user interactions with AI components in Security Copilot that the Microsoft 365 audit log records. These operations represent individual prompting, embedded experiences, or agentic workflows. For more information about Security Copilot interactions, see Prompting in Security Copilot and Microsoft Security Copilot agents overview.

Friendly name	Operation	Description
Interacted with Copilot	CopilotInteraction	A user (or admin or system on behalf of a user) entered prompts in Copilot or an agent.

Microsoft Security Copilot platform management

The following table lists management operations of a Security Copilot and its components that the Microsoft 365 audit log records. These operations represent tenant or workspace settings or AI components such as plugins and promptbooks. For more information about the management elements of Security Copilot, see What is Microsoft Security Copilot.

Friendly name	Operation	Description
Created a new Copilot plugin	CreatePlugin	A user (or admin or system on behalf of a user) created a new Copilot plugin.
Created a new Copilot promptbook	CreatePromptBook	A user (or admin or system on behalf of a user) created a new promptbook in Copilot.
Deleted a Copilot plugin	DeletePlugin	A user (or admin or system on behalf of a user) deleted a Copilot plugin.
Deleted a Copilot promptbook	DeletePromptBook	A user (or admin or system on behalf of a user) deleted a Copilot promptbook.

Friendly name	Operation	Description
Disabled a Copilot plugin	DisableCopilotPlugin	A user (or admin or system on behalf of a user) disabled a Copilot plugin.
Disabled a Copilot promptbook	DisablePromptBook	A user (or admin or system on behalf of a user) disabled a Copilot promptbook.
Enabled a Copilot plugin	EnablePlugin	A user (or admin or system on behalf of a user) enabled a Copilot plugin.
Enabled a Copilot promptbook	EnablePromptBook	A user (or admin or system on behalf of a user) enabled a Copilot promptbook.
Updated a Copilot plugin setting	UpdatePlugin	A user (or admin or system on behalf of a user) updated a Copilot plugin setting.
Updated a Copilot promptbook setting	UpdatePromptBook	A user (or admin or system on behalf of a user) updated a Copilot promptbook setting.
Updated a Copilot setting	UpdateTenantSettings	An admin (or system on behalf of an admin) updated a Copilot setting.

Microsoft Sentinel AI tool activities

The following table lists the AI tool related activities in Microsoft Sentinel data lake recorded in the Microsoft 365 audit log. For more information about MCP tools in Microsoft Sentinel data lake, see Tool collection in Microsoft Sentinel MCP server.

Friendly name	Operation	Description
Completed AI tool run	SentinelAIToolRunCompleted	Microsoft Sentinel AI tool run completed.
Created AI tool	SentinelAIToolCreated	User creates a Microsoft Sentinel AI tool.
Started AI tool run	SentinelAIToolRunStarted	Microsoft Sentinel AI tool run started.

Microsoft Sentinel data lake job activities

The following table lists the job activities in Microsoft Sentinel data lake that the Microsoft 365 audit log records. For more information about jobs in Microsoft Sentinel data lake, see Create and manage Jupyter notebook jobs (preview) and Create KQL jobs in the Microsoft Sentinel data lake (preview).

Friendly name	Operation	Description
Completed a job run adhoc	JobRunAdhocCompleted	Ad hoc Job execution completed.

Friendly name	Operation	Description
Completed a job run scheduled	JobRunScheduledCompleted	Scheduled Job execution completed.
Created job	JobCreated	A job is created.
Deleted a custom table	CustomTableDelete	As part of a job run, the custom table was deleted.
Deleted job	JobDeleted	A job is deleted.
Disabled job	JobDisabled	The job is disabled.
Enabled job	JobEnabled	A disabled job is reenabled.
Ran a job adhoc	JobRunAdhoc	Job is triggered manually and run started.
Ran a job on schedule	JobRunScheduled	Job run is triggered due to schedule.
Read from table	TableRead	As part of the job run, a table is read.
Stopped a job run	JobRunStopped	User manually cancels or stops an ongoing job run.
Updated job	JobUpdated	The job definition and/or configuration and schedule details of the job if updated.
Wrote to a custom table	CustomTableWrite	As part of the job run, data was written to a custom table.

Microsoft Sentinel data lake KQL activities

The following table lists the KQL activities in Microsoft Sentinel data lake that the Microsoft 365 audit log records. For more information about KQL in Microsoft Sentinel data lake, see KQL and the Microsoft Sentinel data lake (preview).

Friendly name	Operation	Description
Completed KQL query	KQLQueryCompleted	User runs interactive queries via KQL on data in their Microsoft Sentinel data lake.

Microsoft Sentinel data lake notebook activities

The following table lists the notebook activities in Microsoft Sentinel Data lake recorded in the Microsoft 365 audit log. For more information about notebooks in Microsoft Sentinel data lake, see Use notebooks in Microsoft Sentinel data lake.

Friendly name	Operation	Description
Deleted a custom table	CustomTableDelete	User deleted a table as part of their notebook execution.
Read from table	TableRead	User read a table as part of their notebook execution.
Started session	SessionStarted	User started a notebook session.
Stopped session	SessionStopped	User stopped a notebook session.
Wrote to a custom table	CustomTableWrite	User wrote to a table as part of their notebook execution.

Microsoft Sentinel data lake onboarding activities

The following table lists the Microsoft Sentinel data lake onboarding activities recorded in the Microsoft 365 audit log. For more information about onboarding to Microsoft Sentinel data lake, see Onboarding to Microsoft Sentinel data lake (preview).

Friendly name	Operation	Description
Changed subscription of Sentinel lake	SentinelLakeSubscriptionChanged	User modified the billing subscription ID or resource group associated with the data lake.
Onboarded data for Sentinel lake	SentinelLakeDefaultDataOnboarded	During onboarding, default data onboard is logged.
Set up Sentinel lake	SentinelLakeSetup	At the time of onboarding, Microsoft Sentinel data lake is set up and the details are logged.

Microsoft Sentinel graph activities

The following table lists the graph activities in Microsoft Sentinel recorded in the Microsoft 365 audit log. For more information about Microsoft Sentinel graph, see What is Microsoft Sentinel Graph? (preview).

Friendly name	Operation	Description
Created a graph scenario	GraphScenarioCreated	User created a graph instance for a predefined graph scenario.
Deleted a graph scenario	GraphScenarioDeleted	User deleted or disable a graph instance for a predefined graph scenario.
Ran a graph query	GraphQueryRun	User ran a graph query.

Microsoft Sentinel data lake customer managed keys activities

The following table lists the customer-managed key (CMK) activities in Microsoft Sentinel data lake that the Microsoft 365 audit log records. For more information about CMK in Microsoft Sentinel data lake, see Set up Microsoft Sentinel customer-managed key.

Friendly name	Operation	Description
Disabled CMK encryption	CMKDisabled	The customer-managed key was removed and encryption reverted to Microsoft-managed keys for the data lake.
Enabled CMK encryption	CMKEnabled	A customer-managed key was configured for the data lake successfully.
Received CMK encryption request	CMKRequestReceived	A request to enable customer-managed key encryption was received for the data lake, including the required key details.
Accepted CMK encryption request	CMKRequestAccepted	The request to enable customer-managed key encryption was accepted for the data lake.
Rejected CMK encryption request	CMKRequestRejected	A key wrap, unwrap, or validation operation failed.
Rolled back CMK operation	CMKRollbackExecuted	The customer-managed key configuration could not be completed, so the operation was rolled back for the data lake and reverted to Microsoft-managed keys.
Updated CMK encryption	CMKUpdated	The encryption key was rotated or changed for the data lake.

Microsoft To Do activities

The following table lists the activities in Microsoft To Do that the Microsoft 365 audit log records. For more information about Microsoft To Do, see Support for Microsoft To Do.

> **Note**
>
> To audit events for Microsoft To Do activities, you need a paid Project Plan 1 license (or higher) in addition to the relevant Microsoft 365 license that includes entitlements to Audit (Standard).

Friendly name	Operation	Description
Accepted sharing link on folder	AcceptedSharingLinkOnFolder	Accepted sharing link for a folder.
Attachment created	AttachmentCreated	An attachment was created for a task.
Attachment deleted	AttachmentDeleted	An attachment was deleted.
Attachment updated	AttachmentUpdated	An attachment was updated.
Folder sharing link shared	FolderSharingLinkShared	Created a sharing link for a folder.
Linked entity created	LinkedEntityCreated	A linked entity of task was created.
Linked entity deleted	LinkedEntityDeleted	A linked entity was deleted.
Linked entity updated	LinkedEntityUpdated	A linked entity was updated.
SubTask created	SubTaskCreated	A subtask was created.
SubTask deleted	SubTaskDeleted	A subtask was deleted.
SubTask updated	SubTaskUpdated	A subtask was updated.
Task created	TaskCreated	A task was created.
Task deleted	TaskDeleted	A task was deleted.
Task read	TaskRead	A task was read.
Task updated	TaskUpdated	A task was updated.
TaskList created	TaskListCreated	A task list was created.
TaskList read	TaskListRead	A task list was read.
TaskList updated	TaskListUpdated	A task list was updated.
User invited	UserInvited	Invited user to a folder.

Microsoft 365 Organizational Data Service activities

The following table lists the activities in Microsoft 365 Organizational Data Service that the Microsoft 365 audit log records. For more information about Microsoft 365 Organizational Data Service, see Microsoft 365 Organizational Data Service.

Friendly name	Operation	Description
Updated data sharing settings	DataSharingSettingsUpdated	Logged when a customer updates the list of apps allowed to share data, including which attributes each app can access.

Friendly name	Operation	Description
Updated access policies	AccessPoliciesUpdated	Logged when a customer updates access policies that control access to organizational data.
Updated access policy associations	AccessPolicyAssociationsUpdated	Logged when a customer updates attribute or app associations to access policies that control access to organizational data.

People Skills activities

The following table lists the activities in People Skills (Viva) that are logged in the Microsoft 365 audit log. For more information about this feature, see **People Skills documentation**.

Friendly name	Operation	Description
Added custom skills	CustomSkillsAdded	Generated when a tenant administrator adds custom skills to the tenant skills library.
Added graph skills	GraphSkillsAdded	Generated when a tenant administrator adds graph-based skills to the tenant library.
Added skill from inference	SkillAddedFromInference	Generated when the system adds an inferred skill to a user profile via TBA-triggered inference.
Added skill from learning	SkillAddedFromLearning	Generated when the system adds a skill inferred from learning activity.
Added user skill	UserSkillAdded	Generated when a user adds a skill to their own profile.
Applied external bulk tag	ExternalBulkTagApplied	Generated when a third-party application applies bulk content tags via the service-to-service (S2S) API.
Changed inferencing setting	InferencingToggleChanged	Generated when the E3/E5 inferencing toggle status changes.
Completed tenant offboarding	TenantOffboardCompleted	Generated when tenant offboarding completes and all tenant data has been deleted.
Deactivated tenant auto-on	TenantAutoOnDeactivated	Generated when the auto-on tenant setting is deactivated by the system.
Deleted custom skills	CustomSkillsDeleted	Generated when a tenant administrator deletes custom skills from the tenant library.
Deleted external skills	ExternalSkillsDeleted	Generated when a third-party application deletes previously ingested skills data.
Deleted graph skills	GraphSkillsDeleted	Generated when a tenant administrator removes graph-based skills from the tenant library.

Friendly name	Operation	Description
Deleted tenant skills library	TenantSkillsLibraryDeleted	Generated when the tenant skills library is removed as part of offboarding.
Deleted user data	UserDataDeleted	Generated when user data is deleted during tenant offboarding.
Deleted user skill	UserSkillDeleted	Generated when a user removes a skill from their profile.
Exported skills	SkillsExported	Generated when a tenant administrator exports skills data from the service.
Ingested external skills	ExternalSkillsIngested	Generated when a third-party application ingests skills data into People Skills via the service-to-service (S2S) API.
Provisioned skills library	SkillsLibraryProvisioned	Generated when the skills library is auto-provisioned for a tenant.
Provisioned tenant	TenantProvisioned	Generated when a tenant administrator provisions the People Skills skills library for their tenant.
Started tenant offboarding	TenantOffboardKickoff	Generated when tenant offboarding (data deletion) is initiated by the system.
Updated tenant settings	TenantSettingsUpdated	Generated when a tenant administrator modifies People Skills tenant-level configuration settings.
Updated user skill	UserSkillUpdated	Generated when a user updates metadata on an existing skill in their profile.
Validated custom skills	CustomSkillsValidated	Generated when a tenant administrator validates custom skills data (marks skills as validated).

Personal insights activities

The following table lists the activities in personal insights recorded in the Microsoft 365 audit log. For more information about personal insights, see Admin guide for personal insights.

Friendly name	Operation	Description
Updated organization MyAnalytics settings	UpdatedOrganizationMyAnalyticsSettings	Admin updates organization-level settings for personal insights.
Updated user MyAnalytics settings	UpdatedUserMyAnalyticsSettings	Admin updates user settings for personal insights.

Planner activities

The following table lists the user and admin activities in Microsoft Planner that the Microsoft 365 audit log records. The table includes the friendly name that displays in the **Activities** column and the name of the corresponding operation that appears in the detailed information of an audit record and in the CSV file when you export the search results.

> **Note**
>
> Portfolio activity in Planner is logged with Microsoft Project for web roadmap activity. For details, see the Project for the web activities section.

Friendly name	Operation	Description
Added a member to a roster	RosterMemberAdded	A member was added to a roster. If the add operation was a *ResultStatus.Failure* or *ResultStatus.AuthorizationFailure*, the *MemberIds* is a list of member IDs attempted.
Added a reaction to a chat message	ChatMessageReactionAdded	A reaction was added to a chat message by a user or an app.
Assigned a task	TaskAssigned	A task's assignee was modified by a user or an app. This might be an unassigned task getting assigned or an assigned task with a new assignee.
Completed a task	TaskCompleted	A task was marked completed by a user or an app.
Copied a plan	PlanCopied	A plan was copied by a user or an app. If the copy operation was a *ResultStatus.Failure* or *ResultStatus.Failure*, the *newPlanId* is null, the *newContainerType* is *ContainerType.Invalid*, and *newContainerId* is null.
Created a chat message	ChatMessageCreated	A chat message was created by a user or an app.
Created a goal	GoalCreated	A goal was created by a user or an app. If the create operation was a *ResultStatus.Failure* or *ResultStatus.AuthorizationFailure*, the *ObjectId* is null and *PlanId* is null.
Created a plan	PlanCreated	A plan was created by a user or an app. If the create operation was a *ResultStatus.Failure* or

Friendly name	Operation	Description
		ResultStatus.AuthorizationFailure, the *ObjectId* is null, *ContainerType* is *ContainerType.Invalid* and *ContainerId* is null.
Created a roster	RosterCreated	A roster was created by a user or an app. If the create operation was a *ResultStatus.Failure* or *ResultStatus.AuthorizationFailure*, the *ObjectId* is null, *MemberIds* is an empty string.
Created a task	TaskCreated	A task was created by a user or an app. If the create operation was a *ResultStatus.Failure* or *ResultStatus.AuthorizationFailure*, the *ObjectId* is null and *PlanId* is null.
Deleted a chat message	ChatMessageDeleted	A chat message was deleted from a task chat by a user or an app.
Deleted a goal	GoalDeleted	A goal was deleted by a user or an app.
Deleted a plan	PlanDeleted	A plan was deleted by a user or an app.
Deleted a roster	RosterDeleted	A roster was deleted by a user or an app.
Deleted a task	TaskDeleted	A task was deleted by a user or an app.
Modified a chat message	ChatMessageModified	A chat message was modified by a user or an app.
Modified a goal	GoalModified	A goal was modified by a user or an app.
Modified a plan	PlanModified	A plan was modified by a user or an app.
Modified a task	TaskModified	A task was modified by a user or an app.
Read a goal	GoalRead	A goal was read by a user or an app. If the read operation was a *ResultStatus.Failure* or *ResultStatus.AuthorizationFailure*, the *PlanId* is null.
Read a list of chat messages	ChatMessageListRead	A list of chat messages was queried by a user or an app. - if the query operation was a *ResultStatus.Failure* or *ResultStatus.AuthorizationFailure*, *MessageList* is an empty string.

Friendly name	Operation	Description
Read a list of goals	GoalListRead	A list of goals was queried by a user or an app. If the query operation was a *ResultStatus.Failure* or *ResultStatus.AuthorizationFailure*, *GoalList* is an empty string.
Read a list of plans	PlanListRead	A list of plans was queried by a user or an app. If the query operation was a *ResultStatus.Failure* or *ResultStatus.AuthorizationFailure*, *PlanList* is an empty string.
Read a list of tasks	TaskListRead	A list of tasks was queried by a user or an app. If the query operation was a *ResultStatus.Failure* or *ResultStatus.AuthorizationFailure*, *TaskList* is an empty string.
Read a plan	PlanRead	A plan was read by a user or an app. If the read operation was a *ResultStatus.Failure* or *ResultStatus.AuthorizationFailure*, *ContainerType* is *ContainerType.Invalid* and ContainerId is null.
Read a task	TaskRead	A task was read by a user or an app. If the read operation was a *ResultStatus.Failure* or *ResultStatus.AuthorizationFailure*, *PlanId* is null.
Removed a member to a roster	RosterMemberDeleted	A member was removed from a roster. If the remove operation was a *ResultStatus.Failure* or *ResultStatus.AuthorizationFailure*, *MemberIds* is the list of member IDs attempted.
Removed a reaction to a chat message	ChatMessageReactionRemoved	A reaction was removed from a chat message by a user or an app
Updated a roster's sensitivity label	RosterSensitivityLabelUpdated	A user or an app updates a roster's sensitivity label.
Updated tenant settings	TenantSettingsUpdated	Organization settings are updated by an organization admin. If the update operation was a *ResultStatus.Failure* or *ResultStatus.AuthorizationFailure*, the *ObjectId* is original settings, and *TenantSettings* is the attempted organization settings.

Power Apps activities

You can search the audit log for app-related activities in Power Apps. These activities include creating, launching, and publishing an app. Assigning permissions to apps is also audited. For a description of all Power Apps activities, see Activity logging for Power Apps.

> **Note**
>
> Alert Policies (Protection Alert) audit events (RecordType:45) aren't currently supported for PowerApps.

Power Automate activities

You can search the audit log for activities in Power Automate (formerly called Microsoft Flow). These activities include creating, editing, and deleting flows, and changing flow permissions.

Project for the web activities

You can search the audit log for activities in Microsoft Project for the web. Microsoft Project for the web is built on the Microsoft Dataverse and has an associated Project Power App. To enable auditing for scenarios where the user is using the Microsoft Dataverse or the Project Power App, see the System Settings Auditing tab guidance. For a list of entities related to Project for the web, see the Export user data from Project for the web guidance.

For information about Microsoft Project for the web, see Microsoft Project for the web.

> **Note**
>
> Auditing events for Microsoft Project for the web activities requires a paid Project Plan 1 license (or higher).

The following table lists the Microsoft Project for web activities recorded in the Microsoft 365 audit log:

Friendly name	Operation	Description
Created project	ProjectCreated	A project was created by a user or app.
Created roadmap	RoadmapCreated	A roadmap or portfolio was created by a user or app.
Created roadmap item	RoadmapItemCreated	A roadmap or portfolio item was created by a user or app.

Friendly name	Operation	Description
Created task	TaskCreated	A task was created by a user or app.
Deleted project	ProjectDeleted	A project was deleted by a user or app.
Deleted roadmap	RoadmapDeleted	A roadmap or portfolio was deleted by a user or app.
Deleted roadmap item	RoadmapItemDeleted	A roadmap or portfolio item was deleted by a user or app.
Deleted task	TaskDeleted	A task was deleted by a user or app.
Project accessed	ProjectAccessed	A project was read by a user or app.
Project home accessed	ProjectListAccessed	A user queried a list of projects and/or roadmaps.
Roadmap accessed	RoadmapAccessed	A roadmap or portfolio was read by a user or app.
Roadmap item accessed	RoadmapItemAccessed	A roadmap or portfolio item was read by a user or app.
Task accessed	TaskAccessed	A task was read by a user or app.
Updated project	ProjectUpdated	A project was modified by a user or app.
Updated project settings	ProjectForTheWebProjectSettings	An admin updated project settings.
Updated roadmap	RoadmapUpdated	A roadmap or portfolio was modified by a user or app.
Updated roadmap item	RoadmapItemUpdated	A roadmap or portfolio item was modified by a user or app.
Updated roadmap settings	ProjectForTheWebRoadmaptSettings	An admin updated roadmap or portfolio settings.
Updated task	TaskUpdated	A task was modified by a user or app.

Quarantine activities

The following table lists the quarantine activities recorded in the Microsoft 365 audit log. For more information about quarantine, see Quarantine email messages.

Friendly name	Operation	Description
Deleted quarantine message	QuarantineDeleteMessage	An admin or user deleted an email message that was harmful.
Denied quarantine message release request	QuarantineReleaseRequestDeny	An admin denial for a release request from a user for an email message that was harmful.
Exported quarantine message	QuarantineExport	An admin or user exported an email message that was harmful.
Previewed quarantine message	QuarantinePreview	An admin or user previewed an email message that was harmful.
Release request quarantine message	QuarantineReleaseRequest	A user requested the release of an email message that was harmful.
Released quarantine message	QuarantineRelease	An admin or user released an email message from quarantine that was harmful.
Viewed quarantine message's header	QuarantineViewHeader	An admin or user viewed the header of an email message that was harmful.

Report activities

The following table lists the activities for usage reports recorded in the Microsoft 365 audit log.

Friendly name	Operation	Description
Updated usage report privacy settings	UpdateUsageReportsPrivacySetting	Admin updated privacy settings for usage reports.

Retention policy and retention label activities

The following table lists the configuration activities recorded in the Microsoft 365 audit log for retention policies and retention labels when you create, reconfigure, or delete them.

Friendly name	Operation	Description
Changed adaptive scope membership	ApplicableAdaptiveScopeChange	Users, sites, or groups were added to or removed from the adaptive scope. These changes result from running the scope's query. Because the changes are system-initiated, the reported user displays as a GUID rather than a user account.
Configured settings for a retention policy	NewRetentionComplianceRule	Administrator configured the retention settings for a new retention policy. Retention settings include how long items are retained, and what happens to items when the retention period expires (such as

Friendly name	Operation	Description
		deleting items, retaining items, or retaining and then deleting them). This activity also corresponds to running the New-RetentionComplianceRule cmdlet.
Created adaptive scope	NewAdaptiveScope	Administrator created an adaptive scope.
Created retention label	NewComplianceTag	Administrator created a new retention label.
Created retention policy	NewRetentionCompliancePolicy	Administrator created a new retention policy.
Deleted adaptive scope	RemoveAdaptiveScope	Administrator deleted an adaptive scope.
Deleted retention label	RemoveComplianceTag	Administrator deleted a retention label.
Deleted retention policy	RemoveRetentionCompliancePolicy	Administrator deleted a retention policy.
Deleted settings from a retention policy	RemoveRetentionComplianceRule	Administrator deleted the configuration settings of a retention policy. Most likely, this activity is logged when an administrator deletes a retention policy or runs the Remove-RetentionComplianceRule cmdlet.
Enabled regulatory record option for retention labels	SetRestrictiveRetentionUI	Administrator ran the Set-RegulatoryComplianceUI cmdlet so that an administrator can then select the UI configuration option for a retention label to mark content as a regulatory record.
Priority cleanup label applied on file	PriorityCleanupTagApplied	A retention label for priority cleanup is applied to an item on OneDrive or SharePoint (includes label replacement).
Retained email item proactively	ExchangeDataProactivelyPreserved	Adaptive protection automatically applied a retention label to retain an item in Exchange.
Retained file proactively	SharePointDataProactivelyPreserved	Adaptive protection automatically applied a retention label to retain an item in SharePoint or OneDrive.
Updated adaptive scope	SetAdaptiveScope	Administrator changed the description or query for an existing adaptive scope.

Friendly name	Operation	Description
Updated retention label	SetComplianceTag	Administrator updated an existing retention label.
Updated retention policy	SetRetentionCompliancePolicy	Administrator updated an existing a retention policy. Updates that trigger this event include adding or excluding content locations that the retention policy is applied to.
Updated settings for a retention policy	SetRetentionComplianceRule	Administrator changed the retention settings for an existing retention policy. Retention settings include how long items are retained, and what happens to items when the retention period expires (such as deleting items, retaining items, or retaining and then deleting them). This activity also corresponds to running the Set-RetentionComplianceRule cmdlet.

Role administration activities

The following table lists Microsoft Entra role administration activities recorded in the Microsoft 365 audit log when an admin manages admin roles in the Microsoft 365 admin center or in the Azure management portal.

Friendly name	Operation	Description
Add member to Role	Add member to role.	Added a user to an admin role in Microsoft 365.
Removed a user from a directory role	Remove member from role.	Removed a user to from an admin role in Microsoft 365.
Set company contact information	Set company contact information.	Updated the company-level contact preferences for your organization. This includes email addresses for subscription-related email sent by Microsoft 365, and technical notifications about services.

Sensitive information types activities

The following table lists the audit events for activities recorded in the Microsoft 365 audit log involving creation and updating of sensitive information types.

Friendly name	Operation	Description
Created new sensitive information type	CreateRulePackage / EditRulePackage*	A new sensitive information type was created. This information includes any SITs created by copying an out of the box SIT. **Note**: This activity surfaces under the audit activities 'Created rule package' or 'Edited rule package'.
Deleted a sensitive information type	EditRulePackage / RemoveRulePackage	An existing sensitive information type was deleted. This activity surfaces under the audit activity "Edited rule package" or "Removed rule package."
Edited a sensitive information type	EditRulePackage	An existing sensitive information type was edited. This information can include operations like adding or removing a pattern and editing the regex or keyword associated with the sensitive information type. This activity surfaces under the audit activity "Edited rule package."

Sensitivity label activities

The following table lists events recorded in the Microsoft 365 audit log that result from using sensitivity labels with containers and items managed by Microsoft Purview. Containers include SharePoint sites, Teams sites, and Loop workspaces. Items include documents, emails, calendar events, Loop components and pages, and Copilot pages. For auto-labeling policies, items also include files and data assets in Microsoft Purview Data Map.

Friendly name	Operation	Description
Applied sensitivity label to file	FileSensitivityLabelApplied SensitivityLabelApplied	A sensitivity label was applied to an item by using Microsoft 365 apps, Office for the web, the details pane in a SharePoint document library or from the Files tab in Teams, or an auto-labeling policy. The operations for this activity are different depending on how the label was applied: - Office for the web, details pane in a SharePoint document library or the Files tab in Teams, or an auto-labeling policy (FileSensitivityLabelApplied) - Microsoft 365 apps (SensitivityLabelApplied).
Applied sensitivity label to site	SiteSensitivityLabelApplied	A sensitivity label was applied to a SharePoint site, Teams site that isn't group-connected, or a Loop workspace.
Changed sensitivity	FileSensitivityLabelChanged	A different sensitivity label was applied to an item.

Friendly name	Operation	Description
label applied to file	SensitivityLabelUpdated	The operations for this activity are different depending on how the label was changed: - Office for the web, details pane in a SharePoint document library or the Files tab in Teams, or an auto-labeling policy (FileSensitivityLabelChanged) - Microsoft 365 apps (SensitivityLabelUpdated).
Changed sensitivity label on a site	SiteSensitivityLabelChanged	A different sensitivity label was applied to a SharePoint site, Teams site that isn't group-connected, or a Loop workspace.
Failed to apply file sensitivity label	FileSensitivityLabelAppliedFailed	A sensitivity label failed to be applied to an item by using Office for the web, the details pane in a SharePoint document library or from the Files tab in Teams, or an auto-labeling policy.
Failed to change file sensitivity label	FileSensitivityLabelChangedFailed	A different sensitivity label failed to be applied to an item using Office for the web, the details pane in a SharePoint document library or from the Files tab in Teams, or an auto-labeling policy.
Failed to remove file sensitivity label	FileSensitivityLabelRemovedFailed	A sensitivity label failed to be removed from an item by using Office for the web, the details pane in a SharePoint document library or from the Files tab in Teams, or an auto-labeling policy.
Removed sensitivity label from file	FileSensitivityLabelRemoved SensitivityLabelRemoved	A sensitivity label was removed from an item by using Microsoft 365 apps, Office for the web, the details pane in a SharePoint document library or the Files tab in Teams, an auto-labeling policy, or the Unlock-SPOSensitivityLabelEncryptedFile cmdlet. The operations for this activity are different depending on how the label was removed: - Office for the web, details pane in a SharePoint document library or the Files tab in Teams, or an auto-labeling policy (FileSensitivityLabelRemoved) - Microsoft 365 apps (SensitivityLabelRemoved).
Removed sensitivity label from site	SiteSensitivityLabelRemoved	A sensitivity label was removed from a SharePoint site, Teams site that isn't group-connected, or a Loop workspace.

The following table lists descriptions for values recorded in the Microsoft 365 audit log included in the *FailureReason* property for audit activities that log failures to apply, change, or remove

sensitivity labels. These properties aren't available for other sensitivity label activities:

FailureReason property	Property description
CannotOverrideCurrentLabel	A sensitivity label wasn't assigned to the file because the currently applied sensitivity label has a higher priority.
DowngradeJustificationTextMissing	Sensitivity label wasn't assigned to the file because justification text is required for this operation.
FileEncrypted	A sensitivity label wasn't assigned to the file because it's encrypted.
FileExtensionNotSupported	A sensitivity label wasn't assigned to the file because the file type isn't supported for this operation.
FileLockedOrCheckedOut	A sensitivity label wasn't assigned to the file because it's locked or checked out.
FileNotFound	A sensitivity label wasn't assigned to the file because it isn't available.
FileNotSupported	A sensitivity label wasn't assigned to the file because it doesn't support this operation.
FileTooLarge	A sensitivity label wasn't assigned to the file because it's too large to support this operation.
InvalidArgument	A sensitivity label wasn't assigned to the file because some parameters provided to perform this operation aren't valid.
LabelIdNotFound	A sensitivity label wasn't assigned to the file because the sensitivity label provided is unavailable or isn't valid.
LabelNotSupported	A sensitivity label wasn't assigned to the file because the sensitivity label provided isn't supported.
LabelOperationsDisabled	A sensitivity label wasn't assigned to the file because the sensitivity label operation is disabled on the file.
MinorVersionLimitExceeded	A sensitivity label wasn't assigned to the file because the version limit is exceeded.
UnauthorizedAccessException	A sensitivity label wasn't assigned to the file because the current user isn't authorized to perform this operation.
Unknown	A sensitivity label wasn't assigned to the file because of an internal error.
ZeroByteFileError	A sensitivity label wasn't assigned to the file because the operation can't be performed on a zero-byte file.

Additional auditing information for sensitivity labels:

- When you use sensitivity labels for Microsoft 365 Groups, and therefore Teams sites that are group-connected, the labels are audited with group management in Microsoft Entra ID. For more information, see Audit logs in Microsoft Entra ID.
- When you use sensitivity labels for Teams meeting invites, and Teams meeting options and chat, see Search the audit log for events in Microsoft Teams.
- When you use sensitivity labels with Power BI, see Audit schema for sensitivity labels in Power BI.
- When you use sensitivity labels with Microsoft Defender for cloud apps, see Governing connected apps and the labeling information for file governance actions.

SharePoint list activities

The following table lists the activities recorded in the Microsoft 365 audit log related to when users interact with lists and list items in SharePoint. Audit records for some SharePoint activities show that the app@sharepoint user performed the activity on behalf of the user or admin who initiated the action. For more information, see The app@sharepoint user in audit records.

Friendly name	Operation	Description
Created list	ListCreated	A user created a SharePoint list.
Created list column	ListColumnCreated	A user created a SharePoint list column. A list column is a column that's attached to one or more SharePoint lists.
Created list content type	ListContentTypeCreated	A user created a list content type. A list content type is a content type that's attached to one or more SharePoint lists.
Created list item	ListItemCreated	A user created an item in an existing SharePoint list.
Created site column	SiteColumnCreated	A user created a SharePoint site column. A site column is a column that isn't attached to a list. A site column is also a metadata structure that can be used by any list in a given web.
Created site content type	Site ContentType Created	A user created a site content type. A site content type is a content type that's attached to the parent site.
Deleted list	ListDeleted	A user deleted a SharePoint list.
Deleted list column	List Column Deleted	A user deleted a SharePoint list column.
Deleted list content type	ListContentTypeDeleted	A user deleted a list content type.

Friendly name	Operation	Description
Deleted list item	List Item Deleted	A user deleted a SharePoint list item.
Deleted site column	SiteColumnDeleted	A user deleted a SharePoint site column.
Deleted site content type	SiteContentTypeDeleted	A user deleted a site content type.
Recycled list item	ListItemRecycled	A user moved a SharePoint list item to the Recycle Bin.
Restored list	ListRestored	A user restored a SharePoint list from the Recycle Bin.
Restored list item	ListItemRestored	A user restored a SharePoint list item from the Recycle Bin.
Updated list	ListUpdated	A user updated a SharePoint list by modifying one or more properties.
Updated list column	ListColumnUpdated	A user updated a SharePoint list column by modifying one or more properties.
Updated list content type	ListContentTypeUpdated	A user updated a list content type by modifying one or more properties.
Updated list item	ListItemUpdated	A user updated a SharePoint list item by modifying one or more properties.
Updated list view	ListViewUpdated	A user updated a SharePoint list view by modifying one or more properties.
Updated site column	SiteColumnUpdated	A user updated a SharePoint site column by modifying one or more properties.
Updated site content type	SiteContentTypeUpdated	A user updated a site content type by modifying one or more properties.

SharePoint Embedded container type activities

The following table lists SharePoint Embedded container type activities that the Microsoft 365 audit log records. For more information about SharePoint Embedded, see SharePoint Embedded overview.

Friendly name	Operation	Description
Created container type	ContainerTypeCreated	A SharePoint Embedded container type was created.
Deleted container type	ContainerTypeDeleted	A SharePoint Embedded container type was deleted.
Updated container type	ContainerTypeUpdated	A SharePoint Embedded container type was updated.
Updated container type owners	ContainerTypeOwnersUpdated	The owners of a SharePoint Embedded container type were updated.

Sharing and access request activities

The following table lists the user sharing and access request activities in SharePoint and OneDrive recorded in the Microsoft 365 audit log. For sharing events, the **Detail** column under **Results** identifies the name of the user or group the item was shared with and whether that user or group is a member or guest in your organization. For more information, see Use sharing auditing in the audit log.

> **Note**
>
> Users can be either *members* or *guests* based on the UserType property of the user object. A member is usually a user within the organization, and a guest is usually a collaborator outside of your organization. When a user accepts a sharing invitation (and isn't already part of your organization), a guest account is created for them in your organization's directory. After the guest user has an account in your directory, you might share resources directly with them (without requiring an invitation).

Friendly name	Operation	Description
Accepted access request	AccessRequestAccepted	An access request to a site, folder, or document was accepted and the requesting user was granted access.
Accepted sharing invitation	SharingInvitationAccepted	User (member or guest) accepted a sharing invitation and was granted access to a resource. This event includes information about the user who was invited and the email address that was used to accept the invitation (they could be different). This activity is often accompanied by a second event that describes how the user was granted access to the resource,

Friendly name	Operation	Description
		for example, adding the user to a group that has access to the resource.
Added permission level to site collection	PermissionLevelAdded	A permission level was added to a site collection.
Blocked sharing invitation	SharingInvitationBlocked	A sharing invitation sent by a user in your organization is blocked because of an external sharing policy that either allows or denies external sharing based on the domain of the target user. In this case, the sharing invitation was blocked because: The target user's domain isn't included in the list of allowed domains. Or The target user's domain is included in the list of blocked domains. For more information about allowing or blocking external sharing based on domains, see Restricted domains sharing in SharePoint and OneDrive.
Created a company shareable link	CompanyLinkCreated	User created a company-wide link to a resource. Company-wide links can only be used by members in your organization. They can't be used by guests.
Created access request	AccessRequestCreated	User requests access to a site, folder, or document they don't have permissions to access.
Created an anonymous link	AnonymousLinkCreated	User created an anonymous link to a resource. Anyone with this link can access the resource without having to be authenticated.
Created secure link	SecureLinkCreated	A secure sharing link was created to this item.
Created sharing invitation	SharingInvitationCreated	User shared a resource in SharePoint or OneDrive with a user who isn't in your organization's directory.
Deleted secure link	SecureLinkDeleted	A secure sharing link was deleted.
Denied access request	AccessRequestDenied	An access request to a site, folder, or document was denied.
Removed a company	CompanyLinkRemoved	User removed a company-wide link to a resource. The link can no longer be used to access the resource.

Friendly name	Operation	Description
shareable link		
Removed an anonymous link	AnonymousLinkRemoved	User removed an anonymous link to a resource. The link can no longer be used to access the resource.
Shared file, folder, or site	SharingSet	User (member or guest) shared a file, folder, or site in SharePoint or OneDrive with a user in your organization's directory. The value in the **Detail** column for this activity identifies the name of the user the resource was shared with and whether this user is a member or a guest. This activity is often accompanied by a second event that describes how the user was granted access to the resource. For example, adding the user to a group that has access to the resource.
Unshared file, folder, or site	SharingRevoked	User (member or guest) unshared a file, folder, or site that was previously shared with another user.
Updated access request	AccessRequestUpdated	An access request to an item was updated.
Updated an anonymous link	AnonymousLinkUpdated	User updated an anonymous link to a resource. The updated field is included in the EventData property when you export the search results.
Updated sharing invitation	SharingInvitationUpdated	An external sharing invitation was updated.
Used a company shareable link	CompanyLinkUsed	User accessed a resource by using a company-wide link.
Used an anonymous link	AnonymousLinkUsed	An anonymous user accessed a resource by using an anonymous link. The user's identity might be unknown, but you can get other details such as the user's IP address.
Used secure link	SecureLinkUsed	A user used a secure link.
User added to secure link	AddedToSecureLink	A user was added to the list of entities who can use a secure sharing link.
User removed from secure link	RemovedFromSecureLink	A user was removed from the list of entities who can use a secure sharing link.

Friendly name	Operation	Description
Withdrew sharing invitation	SharingInvitationRevoked	User withdrew a sharing invitation to a resource.

Site administration activities

The following table lists events that result from site administration tasks in SharePoint and are recorded in the Microsoft 365 audit log. Audit records for some SharePoint activities indicate the app@sharepoint user performed the activity on behalf of the user or admin who initiated the action. For more information, see The app@sharepoint user in audit records.

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

Friendly name	Operation	Description
Added allowed data location	AllowedDataLocationAdded	A SharePoint or global administrator added an allowed data location in a multigeo environment.
Added exempt user agent	ExemptUserAgentSet	A SharePoint or global administrator added a user agent to the list of exempt user agents in the SharePoint admin center.
Added geo location admin	GeoAdminAdded	A SharePoint or global administrator added a user as a geo admin of a location.
Allowed user to create groups	AllowGroupCreationSet	Site administrator or owner adds a permission level to a site that allows a user assigned that permission to create a group for that site.
Apply theme to a web	WebThemeApplied	A SharePoint, global administrator, or a site owner applied a web theme.
Canceled site geo move	SiteGeoMoveCancelled	A SharePoint or global administrator canceled a SharePoint or OneDrive site geo move. The multigeo capability lets an organization span multiple Microsoft datacenter geographies, which are called geos. For more information, see Multi-Geo Capabilities in OneDrive and SharePoint.

Friendly name	Operation	Description
Changed a sharing policy	SharingPolicyChanged	A SharePoint or global administrator changed a SharePoint sharing policy by using the Microsoft 365 admin center, SharePoint admin center, or SharePoint Management Shell. Any change to the settings in the sharing policy in your organization are logged. The policy that was changed is identified in the **ModifiedProperties** field in the detailed properties of the event record.
Changed device access policy	DeviceAccessPolicyChanged	A SharePoint or global administrator changed the unmanaged devices policy for your organization. This policy controls access to SharePoint, OneDrive, and Microsoft 365 from devices that aren't joined to your organization. Configuring this policy requires an Enterprise Mobility + Security subscription. For more information, see Control access from unmanaged devices.
Changed exempt user agents	CustomizeExemptUsers	A SharePoint or global administrator customized the list of exempt user agents in the SharePoint admin center. You can specify which user agents to exempt from receiving an entire web page to index. This means when a user agent you specified as exempt encounters an InfoPath form, the form is returned as an XML file, instead of an entire web page. This makes indexing InfoPath forms faster.
Changed network access policy	NetworkAccessPolicyChanged	A SharePoint or global administrator changed the location-based access policy (also called a trusted network boundary) in the SharePoint admin center or by using SharePoint PowerShell. This type of policy controls who can access SharePoint and OneDrive resources in your organization based on authorized IP address ranges that you specify. For more information, see Control access to SharePoint and OneDrive data based on network location.
Completed migration job	MigrationJobCompleted	A migration job completed.
Completed site geo move	SiteGeoMoveCompleted	A site geo move that a global administrator in your organization scheduled completed. The multigeo capability lets an organization span multiple Microsoft datacenter geographies, which are called geos. For more information, see Multi-Geo Capabilities in OneDrive and SharePoint.

Friendly name	Operation	Description
Create tenant-wide theme	TenantWideThemeCreated	A SharePoint administrator, global administrator, or brand manager created a custom theme applied across all SharePoint sites in your organization.
Created Sent To connection	SendToConnectionAdded	A SharePoint or global administrator creates a new Send To connection on the Records management page in the SharePoint admin center. A Send To connection specifies settings for a document repository or a records center. When you create a Send To connection, a Content Organizer can submit documents to the specified location.
Created site collection	SiteCollectionCreated	A SharePoint or global administrator creates a site collection in your SharePoint organization or a user provisions their OneDrive site.
Delete tenant-wide theme	TenantWideThemeDeleted	A SharePoint administrator, global administrator, or brand manager deleted a custom theme applied across all SharePoint sites in your organization.
Deleted orphaned hub site	HubSiteOrphanHubDeleted	A SharePoint or global administrator deleted an orphan hub site, which is a hub site that doesn't have any sites associated with it. An orphaned hub is likely caused by the deletion of the original hub site.
Deleted Sent To connection	SendToConnectionRemoved	A SharePoint or global administrator deletes a Send To connection on the Records management page in the SharePoint admin center.
Deleted site	SiteDeleted	Site administrator deletes a site.
Enabled document preview	PreviewModeEnabledSet	Site administrator enables document preview for a site.
Enabled legacy workflow	LegacyWorkflowEnabledSet	Site administrator or owner adds the SharePoint 2013 Workflow Task content type to the site. Global administrators can also enable workflows for the entire organization in the SharePoint admin center.
Enabled Office on Demand	OfficeOnDemandSet	Site administrator enables Office on Demand, which lets users access the latest version of Office desktop applications. Office on Demand is enabled in the SharePoint admin center and requires a Microsoft 365 subscription that includes full, installed Office applications.
Enabled result source for People Searches	PeopleResultsScopeSet	Site administrator creates the result source for People Searches for a site.

Friendly name	Operation	Description
Enabled RSS feeds	NewsFeedEnabledSet	Site administrator or owner enables RSS feeds for a site. Global administrators can enable RSS feeds for the entire organization in the SharePoint admin center.
Joined site to hub site	HubSiteJoined	A site owner associates their site with a hub site.
Modified site collection quota	SiteCollectionQuotaModified	Site administrator modifies the quota for a site collection.
Permanently delete site	SitePermanentlyDeleted	A SharePoint or global administrator permanently deletes a site from SharePoint Admin Center Deleted Sites.
Registered hub site	HubSiteRegistered	A SharePoint or global administrator creates a hub site. The results are that the site is registered to be a hub site.
Removed allowed data location	AllowedDataLocationDeleted	A SharePoint or global administrator removed an allowed data location in a multigeo environment.
Removed geo location admin	GeoAdminDeleted	A SharePoint or global administrator removed a user as a geo admin of a location.
Renamed site	SiteRenamed	Site administrator or owner renames a site.
Restore Site	SiteRestored	A SharePoint or global administrator restores a site from SharePoint Admin Center Deleted Sites.
Scheduled site geo move	SiteGeoMoveScheduled	A SharePoint or global administrator scheduled a SharePoint or OneDrive site geo move. The multigeo capability lets an organization span multiple Microsoft datacenter geographies, which are called geos. For more information, see Multi-Geo Capabilities in OneDrive and SharePoint.
Set host site	HostSiteSet	A SharePoint or global administrator changes the designated site to host personal or OneDrive sites.
Set storage quota for geo location	GeoQuotaAllocated	A SharePoint or global administrator configured the storage quota for a geo location in a multigeo environment.
Unjoined site from hub site	HubSiteUnjoined	A site owner disassociates their site from a hub site.
Unregistered hub site	HubSiteUnregistered	A SharePoint or global administrator unregisters a site as a hub site. When a hub site is unregistered,

Friendly name	Operation	Description
		it no longer functions as a hub site.
Update DisableSiteBranding config	UpdateDisableSiteBranding	A SharePoint or global administrator enables or disables site branding.
Update tenant-wide theme	TenantWideThemeUpdated	A SharePoint administrator, global administrator, or brand manager updated a custom theme applied across all SharePoint sites in your organization.

Site permissions activities

The following table lists events related to assigning permissions in SharePoint and using groups to give (and revoke) access to sites recorded in the Microsoft 365 audit log. Audit records for some SharePoint activities indicate the app@sharepoint user performed the activity on behalf of the user or admin who initiated the action. For more information, see The app@sharepoint user in audit records.

Friendly name	Operation	Description
Added site collection admin	SiteCollectionAdminAdded	Site collection administrator or owner adds a person as a site collection administrator for a site. Site collection administrators have full control permissions for the site collection and all subsites. This activity is also logged when an admin gives themselves access to a user's OneDrive account (by editing the user profile in the SharePoint admin center or by using the Microsoft 365 admin center).
Added user or group to SharePoint group	AddedToGroup	User added a member or guest to a SharePoint group. This action might be intentional or the result of another activity, such as a sharing event.
Broke permission level inheritance	PermissionLevelsInheritanceBroken	An item was changed so that it no longer inherits permission levels from its parent.
Broke sharing inheritance	SharingInheritanceBroken	An item was changed so that it no longer inherits sharing permissions from its parent.
Created group	GroupAdded	Site administrator or owner creates a group for a site, or performs a task that results in a group being created. For example, the first time a user creates a link to share a file, a system group is added to the user's OneDrive site. This event can also be a result

Friendly name	Operation	Description
		of a user creating a link with edit permissions to a shared file.
Deleted group	GroupRemoved	User deletes a group from a site.
Modified 'Members Can Share' setting	WebMembersCanShareModified	The **Members Can Share** setting was modified on a site.
Modified access request setting	WebRequestAccessModified	The access request settings were modified on a site.
Modified permission level on a site collection	PermissionLevelModified	A permission level was changed on a site collection.
Modified site permissions	SitePermissionsModified	Site administrator or owner (or system account) changes the permission level that is assigned to a group on a site. This activity is also logged if all permissions are removed from a group. **NOTE**: This operation is deprecated in SharePoint. To find related events, you can search for other permission-related activities such as **Added site collection admin**, **Added user or group to SharePoint group**, **Allowed user to create groups**, **Created group**, and **Deleted group.**
Removed permission level from site collection	PermissionLevelRemoved	A permission level was removed from a site collection.
Removed site collection admin	SiteCollectionAdminRemoved	Site collection administrator or owner removes a person as a site collection administrator for a site. This activity is also logged when an admin removes themselves from the list of site collection administrators for a user's OneDrive account (by editing the user profile in the SharePoint admin center). To return this activity in the audit log search results, you have to search for all activities.
Removed user or group from SharePoint group	RemovedFromGroup	User removed a member or guest from a SharePoint group. This action might be intentional or the result of another activity, such as an unsharing event.

Friendly name	Operation	Description
Requested site admin permissions	SiteAdminChangeRequest	User requests to be added as a site collection administrator for a site collection. Site collection administrators have full control permissions for the site collection and all subsites.
Restored sharing inheritance	SharingInheritanceReset	A change was made so that an item inherits sharing permissions from its parent.
Updated group	GroupUpdated	Site administrator or owner changes the settings of a group for a site. This change can include the group's name, who can view or edit the group membership, and how membership requests are handled.

Stream activities

You can search the audit log for activities in Microsoft Stream. These activities include video activities performed by users, group channel activities, and admin activities such as managing users, managing organization settings, and exporting reports. For a description of these activities, see the "Actions logged in Stream" section in Audit Logs in Microsoft Stream.

Synchronization activities

The following table lists file synchronization activities recorded in the Microsoft 365 audit log for SharePoint and OneDrive.

Friendly name	Operation	Description
Allowed computer to sync files	ManagedSyncClientAllowed	User established a sync relationship with a site. The sync relationship is successful because the user's computer is a member of a domain that's been added to the list of domains (called the *safe recipients list*) that can access document libraries in your organization. For more information about this feature, see Use PowerShell cmdlets to enable OneDrive sync for domains that are on the safe recipients list.
Blocked computer from syncing files	UnmanagedSyncClientBlocked	User tries to establish a sync relationship with a site from a computer that isn't a member of your organization's domain or is a member of a domain that hasn't been added to the list of domains (called the *safe recipients list)* that can access document libraries in your organization. The sync relationship isn't allowed, and the user's computer is blocked from syncing, downloading,

Friendly name	Operation	Description
		or uploading files on a document library. For information about this feature, see Use PowerShell cmdlets to enable OneDrive sync for domains that are on the safe recipients list.
Downloaded file changes to computer	FileSyncDownloadedPartial	This event is deprecated along with the old OneDrive sync app (Groove.exe).
Downloaded files to computer	FileSyncDownloadedFull	User downloaded a file to their computer from a SharePoint document library or OneDrive using OneDrive sync app (OneDrive.exe).
Uploaded file changes to document library	FileSyncUploadedPartial	This event is deprecated along with the old OneDrive sync app (Groove.exe).
Uploaded files to document library	FileSyncUploadedFull	User uploaded a new file or changes to a file in SharePoint document library or OneDrive by using OneDrive sync app (OneDrive.exe).

SystemSync activities

The following table lists the activities for SystemSync recorded in the Microsoft 365 audit log.

Friendly name	Operation	Description
Data Share Created	DataShareCreated	When the user created the data export.
Data Share Deleted	DataShareDeleted	When the user deleted the data export.
Download Copy Of Lake Data	DownloadCopyOfLakeData	When the copy of Lake Data is downloaded.
Generate Copy Of Lake Data	GenerateCopyOfLakeData	When the copy of Lake Data is generated.

Teams activities

The following table lists Microsoft Teams activities recorded in the Microsoft 365 audit log. You can search the audit log for user and admin activities in Microsoft Teams. Teams is a chat-centered workspace in Microsoft 365. It brings a team's conversations, meetings, files, and notes together into a single place. For more detailed search guidance, see Search the audit log for events in Microsoft Teams.

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

Friendly name	Operation	Description
Accept message	UserAccepted	Accept a new message from a person who never established contact with.
Added bot to team	BotAddedToTeam	A user adds a bot to a team.
Added channel	ChannelAdded	A user adds a channel to a team.
Added connector	ConnectorAdded	A user adds a connector to a channel.
Added details about Teams meeting [9]	MeetingDetail	Teams added information about a meeting, including the start time, the end time, the URL to join the meeting, and the recording start/stop time for the meeting.
Added information about call participants [10, 11, 12]	CallParticipantDetail	Teams added information about the participants of a call, including the user ID of each participant, the time each participant joined and left the call, and participant connection information.
Added information about meeting participants [9, 12, 14]	MeetingParticipantDetail	Teams added information about the participants of a meeting, including the user ID of each participant, the time a participant joined the meeting, the time a participant left the meeting, and the start/stop time of a meeting recording by the user. You can also see logs with timestamps and information about: - Who joined a meeting - Where screensharing occurred - Who started the screenshare - When screensharing occurred - When screensharing stopped - When take, give, or request control was activated - If these control requests were accepted

Friendly name	Operation	Description
		- Who accepted these control requests - Who the content was shared with.
Added members [6, 8]	MemberAdded	A team owner adds members to a team, channel, or group chat.
Added tab	TabAdded	A user adds a tab to a channel.
Applied sensitivity label	SensitivityLabelApplied	A user or meeting organizer applied a sensitivity label to a Teams meeting.
Approved request	ApprovedRequest	After the user views the details of the approval request, they approved it.
Block user	UserBlocked	Block a user from sending messages to you.
Canceled approval request	CanceledApprovalRequest	User canceled an approval request they made.
Cancelled approval asynchronously	CancelledApprovalAsync	User cancels an approval request asynchronously.
Changed channel setting	ChannelSettingChanged	The ChannelSettingChanged operation is logged when the following activities are performed by a team member. For each of these activities, a description of the setting that was changed (shown in parentheses) is displayed in the **Item** column in the audit log search results. • Changes name of a team channel (**Channel name**) • Changes description of a team channel (**Channel description**).
Changed organization setting	TeamsTenantSettingChanged	The TeamsTenantSettingChanged operation is logged when the following activities are performed by a global admin in the Microsoft 365 admin center. These activities affect org-wide Teams settings. For more information, see Manage Teams settings for your organization. For each of these activities, a description of the setting that was changed (shown in parentheses) is displayed in the **Item** column in the audit log search results. • Enables or disables Teams for the organization (**Microsoft Teams**).

Friendly name	Operation	Description
		• Enables or disables interoperability between Microsoft Teams and Skype for Business for the organization (**Skype for Business interoperability**). • Enables or disables the organizational chart view in Microsoft Teams clients (**Org chart view**). • Enables or disables the ability for team members to schedule private meetings (**Private meeting scheduling**). • Enables or disables the ability for team members to schedule channel meetings (**Channel meeting scheduling**). • Enables or disables video calling in Teams meetings (**Video for Skype meetings**). • Enables or disables screen sharing in Microsoft Teams meetups for the organization (**Screen sharing for Skype meetings**). • Enables or disables that ability to add animated images (called Giphys) to Teams conversations (**Animated images**). • Changes the content rating setting for the organization (**Content rating**). The content rating restricts the type of animated image that can be displayed in conversations. • Enables or disables the ability for team members to add customizable images (called custom memes) from the internet to team conversations (**Customizable images from the Internet**). • Enables or disables the ability for team members to add editable images (called stickers) to team conversations (**Editable images**). • Enables or disables that ability for team members to use bots in Microsoft Teams chats and channels (**Org-wide bots)**. • Enables specific bots for Microsoft Teams. This doesn't include the T-Bot, which is Teams help bot that's available when bots are enabled for the organization (**Individual bots**). • Enables or disables the ability for team members to add extensions or tabs (**Extensions or tabs**).

Friendly name	Operation	Description
		• Enables or disables the side-loading of proprietary bots for Microsoft Teams (**Side loading of Bots**). • Enables or disables the ability for users to send email messages to a Microsoft Teams channel (**Channel email**).
Changed role of members in team	MemberRoleChanged	A team owner changes the role of members in a team. The following values indicate the role type assigned to the user. **1** - The Member role. **2** - The Owner role. **3** - The Guest role. The Members property also includes the name of your organization and the member's email address.
Changed sensitivity label	SensitivityLabelChanged	A user changed a sensitivity label on a Teams meeting.
Changed team setting	TeamSettingChanged	The TeamSettingChanged operation is logged when the following activities are performed by a team owner. For each of these activities, a description of the setting that was changed (shown in parentheses) is displayed in the **Item** column in the audit log search results. • Changes the access type for a team. Teams can be set as private or public (**Team access type**). When a team is private (the default setting), users can access the team only by invitation. When a team is public, it's discoverable by anyone. • Changes the information classification of a team (**Team classification**). For example, team data can be classified as high business impact, medium business impact, or low business impact. • Changes the name of a team (**Team name**). • Changes the team description (**Team description**). • Changes made to team settings. To access these settings, a team owner can right-click a team, select **Manage team**, and

Friendly name	Operation	Description
		then select the **Settings** tab. For these activities, the name of the setting that was changed is displayed in the **Item** column in the audit log search results.
Create org scoped template instance	CreateOrgScopedTemplateInstance	User created an organization scoped template instance.
Create team scoped template instance	CreateTeamScopedTemplateInstance	User created a team scoped template instance.
Create update request	CreateUpdateRequest	User created a new update template.
Created a chat [1, 6]	ChatCreated	A Teams chat was created.
Created approval	CreatedApproval	User created a new approval request.
Created approval asynchronously	CreatedApprovalAsync	User creates a new approval request asynchronously.
Created export job	CreatedExportJob	User created an export job.
Created provisioning	CreatedProvisioning	User created provisioning of CDS (Common Data Service) instance.
Created team	TeamCreated	A user creates a team.
Delete update request	DeleteUpdateRequest	User deletes an update template.
Deleted a message [2]	MessageDeleted	A message in a chat or channel was deleted.
Deleted all organization apps	DeletedAllOrganizationApps	Deleted all organization apps from the catalog.
Deleted app	AppDeletedFromCatalog	An app was deleted from the catalog.
Deleted channel	ChannelDeleted	A user deletes a channel from a team.
Deleted team	TeamDeleted	A team owner deletes a team.
Detected impersonation	TeamsImpersonationDetected	An external message sender is detected to have potential impersonation activity.

Friendly name	Operation	Description
attempt		
Detected security risk in call	SecurityRiskInCallDetected	A security risk was detected during an incoming call in Microsoft Teams, such as brand, domain, or user impersonation.
Downloaded file	DownloadedFile	User downloaded file attached to an approval request.
Edit org scoped template instance	EditOrgScopedTemplateInstance	User edited an organization scoped template instance.
Edit team scoped template instance	EditTeamScopedTemplateInstance	User edited a team scoped template instance.
Edit update request	EditUpdateRequest	User opens template edit wizard and selects the save button to confirm any update or enable/disable templates.
Edited a message	GraphMessageUpdated	A user updates a chat or channel message.
Edited a message with a URL link in Teams	MessageEditedHasLink	A user edits a message and adds a URL link to it in Teams.
Edited approval request	EditedApprovalRequest	User took an edit action on an approval request.
Exported messages [1, 2]	MessagesExported	Chat or channel messages were exported.
Exported recordings [1]	RecordingsExported	Chat recordings were exported.
Exported transcripts [1]	TranscriptsExported	Chat transcripts were exported.
Failed to validate invitation to shared channel [3]	FailedValidation	A user responds to an invitation to a shared channel but the invitation failed validation.
Fetched all hosted content of a message[1]	MessageHostedContentsListed	All hosted content in a message, such as images or code snippets, was retrieved.
Fetched chat [1]	ChatRetrieved	A Microsoft Teams chat was retrieved.

Friendly name	Operation	Description
Get org scoped template instance	GetOrgScopedTemplateInstance	User fetched an organization scoped template instance.
Get team scoped template instance	GetTeamScopedTemplateInstance	User fetched a team scoped template instance.
Got async operation	GotAsyncOperation	User got async operation entity.
Installed app	AppInstalled	An app was installed.
Obtain list of meeting recordings [13]	GraphMeetingRecordingsListed	List recording details for an online meeting.
Obtain list of meeting transcripts [13]	GraphMeetingTranscriptsListed	List transcript details for an online meeting.
Obtain meeting transcript content [13]	GraphMeetingTranscriptContentRead	Get the transcript content for an online meeting or an ad hoc call.
Obtain meeting transcript instance [13]	GraphMeetingTranscriptRead	Get the transcript details for an online meeting or ad hoc call.
Obtained meeting recording content [13]	GraphMeetingRecordingContentRead	Get the recording content for an online meeting or an ad hoc call.
Obtained meeting recording instance [13]	GraphMeetingRecordingRead	Get the recording details for an online meeting or an ad hoc call.
Performed action on card	PerformedCardAction	A user took action on an adaptive card within a chat. Adaptive cards are typically used by bots to allow the rich display of information and interaction in chats. Only inline input actions on an adaptive card inside a chat are available in the audit log. For example, when a user submits a poll response in a channel conversation on an adaptive card generated by a Poll bot. User actions such as

Friendly name	Operation	Description
		"View result", which opens a dialog, or user actions inside dialogs aren't available in the audit log.
Populate AI-generated notes in chat	AINotesUpdate	AI-generated notes were populated for a group chat.
Populate AI-generated notes in live meetings	LiveNotesUpdate	AI-generated notes were populated for a live meeting.
Posted new message [3, 4, 6, 8]	MessageSent	A new message was posted to a chat or a channel.
Published app	AppPublishedToCatalog	An app was added to the catalog.
Read a message [1]	MessageRead	A message of a chat or channel was retrieved.
Read hosted content of a message [1]	MessageHostedContentRead	Hosted content in a message, such as an image or a code snippet, was retrieved.
Reassigned approval request	ReassignedApprovalRequest	User reassigned an approval request to another user.
Received Read Receipt on a message	MessageReadReceiptReceived	A read receipt was received indicating that a user has read one or more messages in a chat. The audit record includes the chat thread and the point up to which messages were read.
Rejected approval request	RejectedApprovalRequest	After the user views the details of the approval request, they rejected it.
Removed bot from team	BotRemovedFromTeam	A user removes a bot from a team.
Removed connector	ConnectorRemoved	A user removes a connector from a channel.
Removed members [6, 8]	MemberRemoved	A team owner removes members from a team, channel, or group chat.
Removed sensitivity label	SensitivityLabelRemoved	A user removed a sensitivity label from a Teams meeting.
Removed sharing of team channel [3]	TerminatedSharing	A team or channel owner disabled sharing for a shared channel.

Friendly name	Operation	Description
Removed tab	TabRemoved	A user removes a tab from a channel.
Responded to approval	RespondedToApproval	User took action on an approval request.
Responded to invitation for shared channel [3]	InviteeResponded	A user responded to a shared channel invitation.
Responded to invitee response to shared channel [3]	ChannelOwnerResponded	A channel owner responded to a response from a user who responded to a shared channel invitation.
Responded with custom response	RespondedWithCustomResponse	User responded with a custom response text to an approval request.
Restored sharing of team channel [3]	SharingRestored	A team or channel owner re-enabled sharing for a shared channel.
Retrieved messages [1]	MessagesListed	Messages from a chat or channel were retrieved.
Sensitive content shared	SensitiveContentShared	Logged when **Detect sensitive content during screen sharing** meeting option is enabled for a Teams meeting and within that meeting there's sensitive content (for example, credit card numbers, bank account numbers, social security numbers and similar identification numbers) displayed on screen during a screen sharing session. For more information, see Manage whether meetings in your organization can detect sensitive content during screen sharing.
Sent a message with a URL link in Teams	MessageCreatedHasLink	A user sends a message containing a URL link in Teams.
Sent change notification for message creation [1]	MessageCreatedNotification	A change notification was sent to notify a subscribed listener application of a new message.
Sent change notification for message deletion [1]	MessageDeletedNotification	A change notification was sent to notify a subscribed listener application of a deleted message.

Friendly name	Operation	Description
Sent change notification for message update [1]	MessageUpdatedNotification	A change notification was sent to notify a subscribed listener application of an updated message.
Sent invitation for shared channel [3]	InviteSent	A channel owner or member sends an invitation to a shared channel. Invitations to shared channels can be sent to people outside of your organization if the channel policy is configured to share the channel with external users.
Submit update	SubmitUpdate	User submitted a new update.
Subscribed to message change notifications [1]	SubscribedToMessages	A subscription was created by a listener application to receive change notifications for messages.
Security Risk detected in a Teams call	SecurityRiskInCallDetected	An external Teams call is detected to have potential security concern such as impersonation
Teams Admin Action	TeamsAdminAction	Logged when a Teams admin performs actions like updating, adding, or deleting Teams related settings or configurations using admin tools like PowerShell, Admin Center, API, etc.
Thread creation probed	CreateThreadProbe	A user probed whether they can create a chat with users in your organization. The following properties are populated for each of these activities: **Activity**: The name of this activity, CreateThreadProbe. **CorrelationId**: Unique request identifier for the audit log. **CreationTime**: The time at which the audit log was created.ExtraProperties**: Contains a single key-value pair that describes the environment where the probe request initiated. **ID**: The ID of the report entry. The ID uniquely identifies the audit log entry. **OrganizationId**: The GUID for your Microsoft 365 organization. **ParticipantInfo**: Contains a list of objectId and tenantId pairs describing the set of target users against which the initiating user (UserKey user) initiated the probe. In other words, these are the set of users with which the initiating user wants

Friendly name	Operation	Description
		to create a thread. Only contains users from your organization. **RecordType**: The type of operation indicated by the record. **UserID**: The user principal name of the user that initiated the probe. **UserKey**: GUID of the user that initiated the probe. **UserSIPDomain**: The SIP domain of the user that initiated the probe. **UserTenantId**:The tenant of the user that initiated the probe. **UserType**: The type of the user that initiated the probe. Only takes on the value 0, which represents a regular AAD user. **Version**: The version number of the activity (identified by the Operation property) that's logged. **Workload**: The service where the activity occurred.
Unblock User	UserUnblocked	Unblock a user that was previously blocked.
Undeleted a message	GraphMessageUndeleted	A message in a chat or channel was deleted.
Uninstalled app	AppUninstalled	An app was uninstalled.
Update Teams Message	GraphMessageUpdated	Occurs when a user edits an existing chat message.
Updated a chat [1]	ChatUpdated	A Teams chat was updated.
Updated a message [1]	MessageUpdated	A message of a chat or channel was updated.
Updated app	AppUpdatedInCatalog	An app was updated in the catalog.
Updated approval request asynchronously	UpdatedApprovalRequestAsync	User took action on an approval request asynchronously.
Updated connector	ConnectorUpdated	A user modified a connector in a channel.
Updated tab	TabUpdated	A user modified a tab in a channel.
Upgraded app	AppUpgraded	An app was upgraded to its latest version in the catalog.

Friendly name	Operation	Description
User signed in to Teams	TeamsSessionStarted	A user signs in to a Microsoft Teams client. This event doesn't capture token refresh activities.
View approval request	ViewApprovalRequest	User makes a request to view an approval request.
View update	ViewUpdate	User views the details of the update.
Viewed approval request	ViewedApprovalRequest	User views the details of the approval request from either the adaptive card or within the app.
Viewed approval with additional field	ViewedApprovalWithAdditionalField	User views the details of the approval request that contains at least one editable text field.
Viewed approval with group request	ViewedApprovalWithGroupRequest	User views the details of the approval request with a group included.

Note

[1] An audit record for this event is only logged when the operation is performed by calling a Microsoft Graph API. If the operation is performed in the Teams client, an audit record isn't logged.

[2] This event is only available in Audit (Premium). That means users must be assigned the appropriate license before these events are logged in the audit log. For more information about activities only available in Audit (Premium), see Audit (Premium) in Microsoft Purview. For Audit (Premium) licensing requirements, see Auditing solutions in Microsoft 365.

[3] This event is in public preview.

[4]This event is generated for chat only if there are guests, federated and/or anonymous users.

[5] This event is currently unavailable in Government Community Cloud (GCC), Government Community Cloud High (GCC-High), and Department of Defense (DoD) organizations.

[6] This event is included in all participating tenants for federated chats.

[7] This event has participating domain info for 1:1 federated chats. [8] This event is included in all chat conversations between external Teams users managed by an organization and external Teams users not managed by an organization.

[9] This event is currently available in Government Community Cloud (GCC), Government Community Cloud High (GCC-High), Department of Defense (DoD) organizations and air-gapped environments.

[10] This event is included in all participating tenants.

[11] An audit record isn't logged for activities with Agents or Bots. This capability is currently in development.

[12] Calls that are recorded or transcribed are included in MeetingParticipantDetail.

[13] This event is logged by Graph API access and not by Teams client.

[14] This event is now shared with all participating tenants, not just the organizer tenant. Participating tenants will receive Meeting Participant Detail audit records for their own users who attended a Teams meeting organized by another tenant.

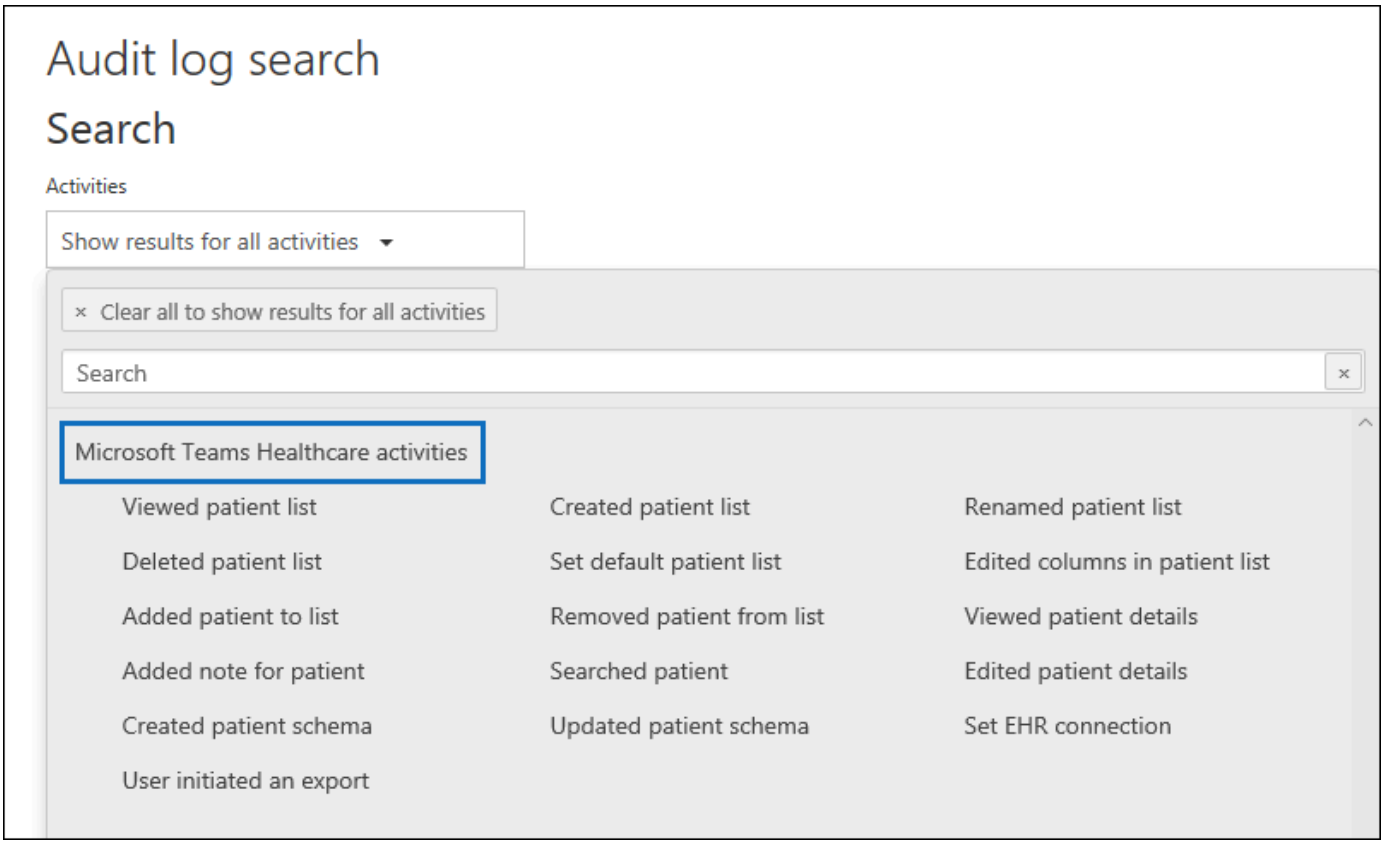
Teams Evaluation Data Hub activities

The following table lists activities for Teams Evaluation Data Hub recorded in the Microsoft 365 audit log. Teams Evaluation Data Hub provides storage and data-access capabilities for conversational data sets (for example, meetings, chats, and Q&A logs) used to evaluate Microsoft Teams experiences. The audit log records when users or services read or write data sets in the hub and when access permissions on a data set are granted or revoked.

Friendly name	Operation	Description
Granted access permission	PermissionGranted	A user or service was granted access permission to a Teams Evaluation Data Hub data set or data asset.
Read data from storage	DataRead	A user or service read, browsed, or downloaded data from Teams Evaluation Data Hub storage.
Revoked access permission	PermissionRevoked	A user or service had access permission revoked for a Teams Evaluation Data Hub data set or data asset.
Wrote data to storage	DataWrite	Data in Teams Evaluation Data Hub storage was created, updated, overwritten, or deleted by a user or service.

Teams Healthcare activities

If your organization uses the Patients application in Microsoft Teams, you can search the audit log for activities related to the Patients app. If your environment is configured to support the Patients app, an additional activity group for these activities is available in the **Activities** picker list.



For a description of the Patients app activities, see Audit logs for Patients app.

Teams Shifts activities

The following table lists Shift app in Microsoft Teams activities recorded in the Microsoft 365 audit log. If your organization uses the Shifts app in Microsoft Teams, you can search the audit log for activities related to the Shifts app. If your environment is configured to support Shifts apps, the **Activities** picker list includes an additional activity group for these activities.

Friendly name	Operation	Description
Accepted off shift message	OffShiftDialogAccepted	A user acknowledged the off-shift message to access Teams after shift hours.
Added open shift	OpenShiftAdded	A user added an open shift to a scheduling group.
Added scheduling group	ScheduleGroupAdded	A user added a new scheduling group to the schedule.
Added shift	ShiftAdded	A user added a shift.
Added shift request	RequestAdded	A user added a shift request.
Added Time clock entry	TimeClockEntryAdded	A user added a new manual Time clock entry on Time Sheet.
Added time off	TimeOffAdded	A user added time off on the schedule.

Friendly name	Operation	Description
Added workforce integration	WorkforceIntegrationAdded	The Shifts app is integrated with a third-party system.
Canceled shift request	RequestCancelled	A user canceled a shift request.
Changed schedule setting	ScheduleSettingChanged	A user changed a setting in Shifts settings.
Clocked in using Time clock	ClockedIn	A user clocked in using Time clock.
Clocked out using Time clock	ClockedOut	A user clocked out using Time clock.
Deleted open shift	OpenShiftDeleted	A user deleted an open shift from a scheduling group.
Deleted scheduling group	ScheduleGroupDeleted	A user deleted a scheduling group from the schedule.
Deleted shift	ShiftDeleted	A user deleted a shift.
Deleted Time clock entry	TimeClockEntryDeleted	A user deleted a Time clock entry on Time Sheet.
Deleted time off	TimeOffDeleted	A user deleted time off.
Edited open shift	OpenShiftEdited	A user edited an open shift in a scheduling group.
Edited scheduling group	ScheduleGroupEdited	A user edited a scheduling group.
Edited shift	ShiftEdited	A user edited a shift.
Edited Time clock entry	TimeClockEntryEdited	A user edited a Time clock entry on Time Sheet.
Edited time off	TimeOffEdited	A user edited time off.
Ended break using Time clock	BreakEnded	A user ended a break during an active Time clock session.
Responded to shift request	RequestRespondedTo	A user responded to a shift request.
Shared schedule	ScheduleShared	A user shared a team schedule for a date range.
Started break using Time clock	BreakStarted	A user started a break during an active Time clock session.

Friendly name	Operation	Description
Withdrew schedule	ScheduleWithdrawn	A user withdrew a published schedule.

Teams Updates activities

The following table lists activities for the Updates app in Microsoft Teams recorded in the Microsoft 365 audit log. If your organization uses the Updates app in Microsoft Teams, you can search the audit log for activities related to the Updates app. If your environment is configured to support Updates apps, the **Activities** picker list includes an additional activity group for these activities.

Friendly name	Operation	Description
Create an update request	CreateUpdateRequest	A user created an update request.
Edit an update request	EditUpdateRequest	A user opens the request editing workflow and selects **Save** to confirm and save any changes, or enables or disables the update request directly.
Submit an update	SubmitUpdate	A user submitted an update.
View the details of one update	ViewUpdate	A user views the details of the update.

Teams User Concern activities

The following table lists user concern activities in Microsoft Teams recorded in the Microsoft 365 audit log. These events are generated when a user reports a concern about a message, call, or another user in Microsoft Teams. For more information, see End user reporting for Security.

Friendly name	Operation	Description
Reported call in Teams	CallReported	A user reported a call as a concern in Microsoft Teams.
Reported message in Teams	MessageReported	A user reported a message as a concern in Microsoft Teams.
Reported user in Teams	UserReported	A user reported another user as a concern in Microsoft Teams.

Trainable Classifier activities

The following table lists the activities for trainable classifiers recorded in the Microsoft 365 audit log.

Friendly name	Operation	Description
Created Purview Data Classification model	ModelCreate	A user (or admin or system on behalf of a user) created a new trainable model in Purview Data Classification.
Deleted Purview Data Classification model	ModelUpdate	A user (or admin or system on behalf of a user) updated a new trainable model in Purview Data Classification.
Published Purview Data Classification model	ModelPublish	A user (or admin or system on behalf of a user) published a new trainable model in Purview Data Classification.
Updated Purview Data Classification model	ModelDelete	A user (or admin or system on behalf of a user) deleted a new trainable model in Purview Data Classification.

Universal Print management activities

The following table lists the activities recorded in the Microsoft 365 audit log when actions are performed on printers, connectors, printer shares, and tenant-level settings using the Universal Print service.

Friendly name	Operation	Description
Added group access to printer share	ShareAccessControlGroupAdded	A group is granted access to a printer share.
Added member to pull-print printer	PullPrintPrinterMemberAdded	A member is added to a pull-print printer group for access control.
Added user access to printer share	ShareAccessControlUserAdded	A user is granted access to a printer share.
Created printer share	ShareCreated	A printer share is created, allowing multiple users to access the printer.
Created pull-print printer	PullPrintPrinterCreated	A pull-print virtual printer is created for secure print release.
Deleted printer share	ShareDeleted	A printer share is deleted, removing shared access.
Deleted pull-print printer	PullPrintPrinterDeleted	A pull-print virtual printer is deleted from Universal Print.
Disabled allow-all access for printer share	ShareAccessControlAllowAllDisabled	Universal access to the printer share is disabled.

Friendly name	Operation	Description
Disabled include-all-printers	PullPrintPrinterIncludeAllPrintersDisabled	All printers are excluded from the pull-print configuration.
Downloaded report	ReportDownloaded	A report is downloaded from Universal Print for analysis or auditing.
Enabled allow-all access for printer share	ShareAccessControlAllowAllEnabled	All users are allowed access to the printer share.
Enabled include-all-printers	PullPrintPrinterIncludeAllPrintersEnabled	All printers are included in the pull-print configuration.
Modified printer settings	PrinterSettingsModified	Printer configuration settings are updated in Universal Print.
Modified printer share settings	ShareSettingsModified	Settings for a printer share are updated in Universal Print.
Modified pull-print printer settings	PullPrintPrinterSettingsModified	Settings for a pull-print printer are updated in Universal Print.
Modified tenant print settings	TenantPrintSettingsModified	Tenant-level print settings are updated in Universal Print.
Registered connector	ConnectorRegistered	The connector is registered with Universal Print, enabling communication between printers and the cloud service.
Registered printer	PrinterRegistered	A printer is registered with Universal Print, making it available for cloud-based printing.
Removed group access from printer share	ShareAccessControlGroupRemoved	A group's access to a printer share is removed.
Removed member from pull-print printer	PullPrintPrinterMemberRemoved	A member is removed from a pull-print printer group.
Removed user access from printer share	ShareAccessControlUserRemoved	A user's access to a printer share is removed.
Swapped printer	PrinterSwapped	A printer is replaced or swapped with another device in Universal Print.
Unregistered connector	ConnectorUnregistered	The connector is unregistered from Universal Print, removing its ability to

Friendly name	Operation	Description
		manage printers.
Unregistered printer	PrinterUnregistered	A printer is removed from Universal Print, making it unavailable for cloud-based printing.

Universal Print print job activities

The following table lists the activities recorded in the Microsoft 365 audit log when actions are performed on print jobs using the Universal Print service.

Friendly name	Operation	Description
Aborted print job	PrintJobAborted	The print job is terminated unexpectedly due to an error or interruption.
Acknowledged print job	PrintJobAcknowledged	The print job is acknowledged by the printer, indicating that the printer has received and is ready to process the print job.
Canceled print job	PrintJobCanceled	The print job is canceled by the user or system before completion.
Completed print job	PrintJobCompleted	The print job finishes successfully, and all pages are printed.
Fetched print job document	PrintJobDocumentFetched	A print job document is retrieved by the printer or the Universal Print connector.
Submitted print job	PrintJobCreated	The print job is submitted to the print queue, indicating the start of the printing process.
Uploaded print job document	PrintJobDocumentUploaded	A print job document is successfully uploaded to Universal Print.

User administration activities

The following table lists user administration activities recorded in the Microsoft 365 audit log when an admin adds or changes a user account by using the Microsoft 365 admin center or the Azure management portal.

Friendly name	Operation	Description
Added user	Add user.	A user account is created.
Changed user license	Change user license.	The license assigned to a user changed. To see what licenses were changed, see the corresponding **Updated user** activity.

Friendly name	Operation	Description
Changed user password	Change user password.	A user changes their password. Self-service password reset has to be enabled (for all or selected users) in your organization to allow users to reset their password. You can also track self-service password reset activity in Microsoft Entra ID. For more information, see Reporting options for Microsoft Entra password management.
Deleted user	Delete user.	A user account was deleted.
Reset user password	Reset user password.	Administrator resets the password for a user.
Set license properties	Set license properties.	Administrator modified the properties of a license assigned to a user.
Set property that forces user to change password	Set force change user password.	Administrator set the property that forces a user to change their password the next time the user signs in to Microsoft 365.
Updated user	Update user.	Administrator changes one or more properties of a user account. For a list of the user properties that can be updated, see the "Update user attributes" section in Microsoft Entra audit Report Events.

Viva Engage activities

The following table lists the user and admin activities in Viva Engage recorded in the Microsoft 365 audit log. To return Viva Engage-related activities from the audit log, select **Show results for all activities** in the **Activities** list. Use the date range boxes and the **Users** list to narrow the search results.

> **Note**
>
> Some Viva Engage audit activities are only available in Audit (Premium). That means users must be assigned the appropriate license before these activities are logged in the audit log. For more information, see Audit (Premium). For Audit (Premium) licensing requirements, see Auditing solutions in Microsoft 365.
>
> In the following table, Audit (Premium) activities are highlighted with an asterisk (*).

Friendly name	Operation	Description
Added corporate communicator	AddUserRole	A user is assigned as a corporate communicator.

Friendly name	Operation	Description
Changed custom usage policy	UsagePolicyUpdated	A tenant admin updates a custom usage policy.
Changed data retention policy	SoftDeleteSettingsUpdated	Verified admin updates the setting for the network data retention policy to either Hard Delete or Soft Delete. Only verified admins can perform this operation.
Changed network configuration	NetworkConfigurationUpdated	Network or verified admin changes the Viva Engage network's configuration. This change includes setting the interval for exporting data and enabling chat.
Changed network profile settings	ProcessProfileFields	Network or verified admin changes the information that appears on member profiles for network users network.
Changed private content mode	SupervisorAdminToggled	Verified admin turns *Private Content Mode* on or off. This mode lets an admin view the posts in private groups and view private messages between individual users (or groups of users). Only verified admins only can perform this operation.
Changed security configuration	NetworkSecurityConfigurationUpdated	Verified admin updates the Viva Engage network's security configuration. This update includes setting password expiration policies and restrictions on IP addresses. Only verified admins can perform this operation.
Conversation closed	CloseConversation	The Viva Engage thread was closed, preventing users from replying to it. This action is available to an admin, corporate communicator, or a user delegate.
Conversation opened	OpenConversation	The Viva Engage thread conversation was opened which allows users to reply to the thread. This action is available to an admin, corporate communications, or a user delegate.
Create a segment	SegmentCreated	Admin creates a segmentation.
Created file	FileCreated	User uploads a file.
Created group	GroupCreation	User creates a group.
Created message	MessageCreation	User creates a message.

Friendly name	Operation	Description
Delete a segment	SegmentDeleted	Admin deletes a segmentation.
Deleted group	GroupDeletion	A group is deleted from Viva Engage.
Deleted message	MessageDeleted	User deletes a message.
Download segmentation admin reports	SegmentationReportDownloaded	Admin reports are downloaded.
Downloaded file	FileDownloaded	User downloads a file.
Event content exported	EventContentExported	Event organizer exports event questions, replies, reaction, and upvote count to a CSV. This action is available to event organizer, co-organizers only. Network admin can toggle this feature off for the entire network in the Engage admin settings.
Event moderation enabled	EventModerationEnabled	Event organizer enabled moderation for an event. This action is available to event organizers only.
Exported data	DataExport	Verified admin exports Viva Engage network data. Only verified admins can perform this operation.
Failed to access file	FileAccessFailure	User failed to access a file.
Failed to access group	GroupAccessFailure	User failed to access a group.
Failed to access thread	ThreadAccessFailure	User failed to access a message thread.
Generate segmentation admin reports	SegmentationReportGenerated	Admin user triggers a report generation.
Reacted to message	MarkedMessageChanged	User reacted to a message.
Remove curated topic*	RemoveCuratedTopic	User removes a curated topic.
Removed Corporate Communicator	RemoveUserRole	A user is removed from the Corporate Communicator role.

Friendly name	Operation	Description
Shared file	FileShared	User shares a file with another user.
Suspended network user	NetworkUserSuspended	Network or verified admin suspends (deactivates) a user from Viva Engage.
Suspended user	UserSuspension	User account is suspended (deactivated).
Tenant Usage Policy Acceptance	UsagePolicyAcceptance	A user accepts an organization-specific usage policy.
Thread Muted	AdminThreadMuted	A thread was muted by admin or watch alert (system user). A watch alert occurs when a thread is flagged for a certain theme (set by the admin) and is automuted by the system.
Thread Unmuted	AdminThreadUnmuted	Admin unmuted a thread.
Updated file description	FileUpdateDescription	User changes the description of a file.
Updated file name	FileUpdateName	User changes the name of a file.
Updated message	MessageUpdated	User updates a message.
Updated role assignment for Network Admin	NetworkAdminUpdated	A user is either promoted or demoted to the Network Admin role.
Updated role assignment for Verified Admin	NetworkVerifiedAdminUpdated	A user is either promoted or demoted to the Verified Admin role.
Viewed file	FileVisited	User views a file.
Viewed thread	ThreadViewed	User views a message thread.

Viva Glint activities

The following table lists the user and admin activities in Viva Glint that the Microsoft 365 audit log records. The table includes the friendly name that's displayed in the **Activities** column and the name of the corresponding operation that appears in the detailed information of an audit record and in the .csv file when you export the search results.

Search the audit log explains how you can search for audit logs from the Microsoft Purview portal. To access audit logs, you need to be a global admin or have audit read permissions. Use the *Activities* filter to search for specific activities and to list all Viva Glint activities by choosing

VivaGlint in the *Record type* filter. Use the date range boxes and the *Users* list to narrow the search results further.

Friendly name	Operation	Description
Added agentic policy	PolicyAdded	An agentic policy was added.
Added Company Admin	AddCompanyAdmin	A user was added to the Company Admin role.
Added Support User	AddSupportUser	A user was added to the Support User role.
Assigned User Role	AssignRole	A user is assigned to a new role.
Changed Client Details	ClientDetailsChanged	Change activity in Viva Glint General Settings or Advanced Configuration: Details.
Company Admin Logged In	AdminLogin	A Company Admin user logged into Viva Glint.
Company Admin Logged Out	AdminLogout	A Company Admin user logged out of Viva Glint.
Completed agentic campaign	CampaignCompleted	An agentic campaign was completed.
Created agentic campaign	CampaignCreated	An agentic campaign was created.
Created User Role	RoleCreated	Glint User Role creation activity.
Deleted agentic campaign	CampaignDeleted	An agentic campaign was deleted.
Deleted agentic policy	PolicyDeleted	An agentic policy was deleted.
Deleted retained agentic campaign data	CampaignRetentionDeleted	Retained agentic campaign data was deleted.
Deleted Support Access	GlintSupportAccessDeleted	A user was removed from the Support User role.
Deleted System	GlintSystemDeleted	The Viva Glint system was deleted.
Deleted tenant data for a DSR request	DsrTenantDeleted	Tenant data was deleted for a data subject rights (DSR) request.
Deleted User	GlintUserDeleted	Viva Glint user deletion activity.
Deleted user data for a	DsrUserDeleted	User data was deleted for a data

Friendly name	Operation	Description
DSR request		subject rights (DSR) request.
Disabled Advanced Configuration Access	UserAdvConfigDisabled	A Company Admin user disabled Advanced Configuration access in Viva Glint.
Disabled Glint to Insights integration	GlintToInsightsIntegrationDisabled	A user disabled importing Viva Insights data in Viva Glint
Enabled Advanced Configuration Access	UserAdvConfigEnabled	A Company Admin user enabled Advanced Configuration access in Viva Glint.
Enabled Glint to Insights integration	GlintToInsightsIntegrationEnabled	A user enabled importing Viva Insights data in Viva Glint
Executed Data Apps Job	GlintDataAppsJobRun	Viva Glint Advanced Configuration Data app and Upload activity.
Exported agentic campaign policy violations	VivaGlintAgenticPolicyViolationsExported	Agentic campaign policy violations export activity.
Exported agentic campaign topic analysis	VivaGlintAgenticTopicAnalysisExported	Agentic campaign topic analysis export activity.
Exported all agentic campaign data	VivaGlintAgenticAllDataExported	Agentic campaign raw data export activity.
Exported Glint 360 Feedback	GlintFeedbackProgramExport	Viva Glint 360 feedback program data export activity.
Exported Glint Distribution Lists	GlintUserGroupExport	Distribution List export activity.
Exported Glint People	GlintUserExport	Glint People data export activity.
Exported Glint Pulse Program Response Rate	GlintPulseProgramRespondentRateExport	Glint pulse program response rate export activity.
Exported Glint Question Library	GlintQuestionExport	Glint question library export activity.
Exported Glint Survey Program	GlintPulseProgramExport	Glint survey program content export activity.
Exported Glint User Role	GlintRoleExport	Glint User Role data export activity.
Exported Raw Survey Data	RawSurveyReponseExport	Raw survey response data export activity.

Friendly name	Operation	Description
Exported User Data	UserDataExport	Employee data export activity.
Generated agentic campaign report	ReportGenerated	An agentic campaign report was generated.
Imported Glint 360 Feedback	GlintFeedbackProgramImport	360 Feedback program data import activity.
Imported Glint SFTP	GlintRubiconImport	Employee data SFTP import activity.
Imported Glint User Data	GlintUserImport	Employee data import activity.
Imported Glint User Role	GlintRoleImport	User Role employee data activity.
Launched agentic campaign	CampaignLaunched	An agentic campaign was launched.
Removed Company Admin	RemoveCompanyAdmin	A user was removed from the Company Admin role.
Reopened Survey	SurveyReopen	A closed Viva Glint survey was reopened.
Reordered agentic policy priorities	PolicyPrioritiesReordered	Agentic policy priorities were reordered.
Set Data DSR Control	DataDsrControlSet	User Data controls for survey data deletion in General Settings update activity.
Set Discarding Employee IDs	DiscardingEmployeeIdsSet	User Data controls for Employee ID reuse in General Settings update activity.
Shared agentic campaign report	ReportShared	An agentic campaign report was shared.
Unassigned User Role	UnassignRole	A user is removed from a role.
Updated agentic campaign	CampaignUpdated	An agentic campaign was updated.
Updated agentic policy	PolicyUpdated	An agentic policy was updated.
Updated agentic settings	AgenticSettingsUpdated	Agentic settings were updated.

Friendly name	Operation	Description
Updated Glint to Insights integration	GlintToInsightsIntegrationUpdated	A user updated the configuration for importing Viva Insights data in Viva Glint
Viewed As	ViewAs	Admins' "View As" another user activity.

Viva Goals activities

The following table lists the user and admin activities in Viva Goals that the Microsoft 365 audit log records. The table includes the friendly name that displays in the **Activities** column and the name of the corresponding operation that appears in the detailed information of an audit record and in the CSV file when you export the search results.

Search the audit log explains how you can search for audit logs from the Microsoft Purview portal. To access audit logs, you need to be a global admin or have audit read permissions. You can use the **Activities** filter to search for specific activities. To list all Viva Goals activities, select **VivaGoals** in the **Record type** filter. You can also use the date range boxes and the **Users** list to narrow the search results further.

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

Friendly name	Operation	Description
Dashboard created	Dashboard created	User created a new dashboard on Viva Goals.
Dashboard deleted	Dashboard deleted	User deleted a dashboard on Viva Goals.
Dashboard updated	Dashboard updated	User updated a dashboard on Viva Goals.
Data exported	Data exported	A user exported list of OKRs or list of users in an organization on Viva Goals.
Goals policy updated	Goals policy updated	The global admin modified the policy or settings at organization level on Viva Goals. For example, the global admin configured who can create organizations on Viva Goals.

Friendly name	Operation	Description
OKR or Project created	OKR or Project created	User created an OKR or Project on Viva Goals.
OKR or Project deleted	OKR or Project deleted	User deleted an OKR or Project.
OKR or Project updated	OKR or Project updated	An OKR/Project was modified or a check-in was made by the user or an integration on Viva Goals.
Organization created	Organization created	Admin or the user created a new organization on Viva Goals.
Organization integrations updated	Organization integrations updated	The user (typically Organization owners or admins) configured a third party integration or updated an existing third party integration for an organization on Viva Goals.
Organization settings updated	Organization settings updated	The user (typically Organization owners or admins) updated organization specific settings on Viva Goals.
Team added	Team added	A new team was created within an organization on Viva Goals.
Team deleted	Team deleted	A team within an organization on Viva Goals was deleted by the user.
Team updated	Team updated	A team within an organization on Viva Goals was modified or updated.
User added	User added	A new user was added to an organization on Viva Goals.
User deactivated	User deactivated	A user was deactivated in an organization.
User deleted	User deleted	A user was deleted from an organization on Viva Goals.
User logged in	User logged in	User signed in to Viva Goals.

Viva Insights activities

Viva Insights provides insight into how groups collaborate across your organization. The following table lists activities recorded in the Microsoft 365 audit log that users assigned the Administrator or Analyst roles in Viva Insights perform. Users assigned the Analyst role have full access to all service features and use the product to do analysis. Users assigned the Administrator role can configure privacy settings and system defaults, and can prepare, upload, and verify organizational data in Viva Insights. For more information, see Introducing Microsoft Viva Insights.

Friendly name	Operation	Description
Accessed OData link	AccessedOdataLink	Analyst accessed the OData link for a query.
Added delegate access	AddDelegates	A user added delegate access for organization insights or Copilot dashboard.
Canceled query	CanceledQuery	Analyst canceled a running query.
Created meeting exclusion	MeetingExclusionCreated	Analyst created a meeting exclusion rule.
Deleted result	DeletedResult	Analyst deleted a query result.
Downloaded report	DownloadedReport	Analyst downloaded a query result file.
Executed query	ExecutedQuery	Analyst ran a query.
Modified CXO Settings	ModifiedCXOSettings	This event logs the assignment or removal of users/groups to access the Microsoft 365 Copilot Dashboard.
Modified Exclusion List Using Entra Group	ModifiedExclusionListUsingEntraGroup	This event logs changes to the exclusion list using the Entra group in the Copilot Dashboard.
Modified Minimum Group Size	ModifiedMinimumGroupSize	This event logs modifications to the organization's minimum group size for generating insights in the Copilot Dashboard and Viva Insights advanced analysis.
Removed delegate access	RemoveDelegates	A user removed delegate access for organization insights or Copilot dashboard.
Updated data access setting	UpdatedDataAccessSetting	Admin updated data access settings.
Uploaded organization data	UploadedOrgData	Admin uploaded organizational data file.
User logged in*	UserLoggedIn	A user signed in to their Microsoft 365 user account.
Viewed Explore	ViewedExplore	Analyst viewed visualizations in one or more Explore page tabs.

Note

*A Microsoft Entra sign-in activity event is created when a user signs in. This activity is logged even if you don't turn on Viva Insights in your organization. For more information about user sign in activities, see Sign-in logs in Microsoft Entra ID.

Viva Pulse activities

The following table lists the user and admin activities in Viva Pulse recorded in the Microsoft 365 audit log. The table includes the friendly name that's displayed in the **Activities** column and the name of the corresponding operation that appears in the detailed information of an audit record and in the CSV file when you export the search results.

Search the audit log explains how you can search for audit logs from the Microsoft Purview portal. To access audit logs, you need to be a global admin or have audit read permissions. You can use the **Activities** filter to search for specific activities. To list all Viva Pulse activities, select *VivaPulse* in the **Record type** filter. You can also use the date range boxes and the **Users** list to narrow the search results further.

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

Friendly name	Operation	Description
Admin deleted a user's data	PulseDeleteUserData	Admin deleted a user's data.
Admin updated tenant's settings	PulseTenantSettingsUpdate	Admin updated an organization setting for Viva Pulse.
Confidential conversation started	PulseConfidentialConversationStarted	Author started a confidential conversation.
Conversation message deleted	PulseConfidentialConversationMessageDeleted	User deleted a conversation message.
Conversation response added	PulseConfidentialConversationResponded	Author or recipient responded on a conversation.
User canceled a Pulse survey	PulseCancel	User canceled a Pulse survey.

Friendly name	Operation	Description
User created a Pulse draft	PulseCreateDraft	User created a Pulse draft.
User created a Pulse survey	PulseCreate	New Viva Pulse feedback request is created.
User deleted a Pulse draft	PulseDeleteDraft	User deleted a Pulse draft.
User deleted a question for Library	PulseQuestionDeleted	User removed a question for pulse question library.
User extended their pulse survey deadline	PulseExtendDeadline	User extended the deadline for the existing Viva Pulse feedback request.
User invited additional users to Pulse survey	PulseInvite	User invited additional users to an existing Viva Pulse feedback request.
User request a pulse data export	PulseDataExport	Admin or Author request a pulse export operation.
User shared a pulse report	PulseShareResults	User shared Viva Pulse feedback results with other users.
User submitted response to a pulse survey	PulseSubmit	Admin or user provided feedback for a Viva Pulse feedback request.

Windows 365 Customer Lockbox activities

The following table lists the user and admin activities in Windows 365 Customer Lockbox that the Microsoft 365 audit log records. To return Windows 365 Customer Lockbox-related activities from the audit log, select **Windows365CustomerLockbox** under **Record types**. Use the date range boxes and the **Users** list to narrow the search results.

Friendly name	Operation	Description
Backfill CMD Agent	AddDevicesToBackfill Operation	Backfill CMD Agent on Cloud PC.
Bulk reinstall CMD Agent	TriggerClientAgentCheckBulkAction Operation	Bulk reinstall CMD agent on demand.
Check PowerShell Execution Policy	Check PowerShell Execution Policy	Check PowerShell execution policy.

Friendly name	Operation	Description
Create LogCollection Request	Create LogCollection Request	Create log collection request to Cloud PC.
Create new work items (Scheduler)	Create new work items (Scheduler)	Create new work items, such as Provision, Deprovision, Reprovision, and others.
Create Remote Action Operation in ActionModeratorService	Create Remote Action Operation in ActionModeratorService	Create remote action by tenantID, workspaceID, actionType, actionParameters.
Create VmExtension Request	Create VmExtention Request	Creates VM extension requests to execute VM extension to run custom scripts on the customer device.
Execute AppHealthPlugin	Execute AppHealthPlugin	Execute AppHealthPlugin.
Install RD agent	Install RD agent	Install RD agent on the device of the user.
OCE Run Commands on VM	OCE Run Commands on VM	Run commands by tenantID, deviceID list, and script.
Post Remote Action Operation	Post Remote Action Operation	Post Remote Action, plus polling the result by GetActions operation.
Reinstall CMD Agent	AddDevicesToReinstall Operation	Reinstall CMD Agent on demand.
Run hybrid AADJ extension	Run hybrid AADJ extension	Run hybrid AADJ extension on the device of the user.
Trigger CMD Agent Canary check.	Trigger CMD Agent Canary check.	Trigger CMD Agent Canary check on specific device.
Trigger device remediation	Trigger device remediation	Trigger device remediation.
Trigger generic action	Trigger generic action	Trigger device action for user.
Trigger generic action by SaaF	Trigger generic action by SaaF	Trigger device action (Retargeting, EnableRdpAccessForCitrix, DisableRdpAccessFprCitrix) for user.
Trigger generic action with options	Trigger generic action with options	Trigger device action with option parameters, more powerful than that of trigger generic action.
Trigger orchestrator	Trigger orchestrator	Trigger orchestrator for the user.

Friendly name	Operation	Description
Upload folder to blob	Upload folder to blob	Compress and upload customer device's folder to blob.

Fabric Workspace Inbound External Data Share setting

The following table lists the Fabric admin activities for enabling and disabling the Workspace Inbound External Data Share setting.

Friendly name	Operation	Description
Enabled external data shares bypass on the workspace	ExternalDataSharesBypassForWorkspaceEnabled	Admin allowed External Data Shares to access data in workspaces where inbound public network access is restricted.
Disabled external data shares bypass on the workspace	ExternalDataSharesBypassForWorkspaceDisabled	Admin blocked External Data Shares from accessing data in workspaces where inbound public network access is restricted.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 10/01/2026

Detailed activity properties in the audit log

Summarize this article for me

When you export the results of an audit log search from the Microsoft Purview portal, you can download all the results that meet your search criteria. You can export this information by selecting **Export results** > **Download all results** on the **Audit log search** page. For more information, see Search the audit log.

When you export all results for an audit log search, the raw data from the unified audit log is copied to a comma-separated value (CSV) file that you download to your local computer. This file contains extra property information from each audit activity record in a column named **AuditData**. This column contains a multi-value property for multiple properties from the audit log record. Each of the **property: value** pairs in this multi-value property are separated by a comma.

The following table describes the activity properties that are included (depending on the service in which an activity occurs) in the multi-property **AuditData** column. The **Microsoft service that has this property** column indicates the service and type of activity (user or admin) that includes the property. For more detailed information about these properties or about properties that might not be listed in this article, see Management Activity API Schema.

> **Tip**
>
> You can use the JSON transform feature in Power Query in Excel to split the **AuditData** column into multiple columns so that each property has its own column. This feature lets you sort and filter on one or more of these properties. To learn how to do this, see **Export, configure, and view audit log records**.

Expand table

Property	Description	Microsoft service that has this property
Actor	The user or service account that performed the action.	Microsoft Entra ID
AddOnName	The name of an add-on that was added, removed, or updated in a team. The type of add-ons in Microsoft Teams is a bot, a connector, or a tab.	Microsoft Teams

Property	Description	Microsoft service that has this property
AddOnType	The type of an add-on that was added, removed, or updated in a team. The following values indicate the type of add-on. **1** - Indicates a bot. **2** - Indicates a connector. **3** - Indicates a tab.	Microsoft Teams
AppAccessContext	The application context for the user or service principal that performed the action.	Microsoft Teams
ArtifactShared	Files or content shared by the user.	Microsoft Teams
AzureActiveDirectoryEventType	The type of Microsoft Entra ID activity. The following values indicate the type of activity. **0** - Indicates an account sign-in activity. **1** - Indicates an Azure application security activity.	Microsoft Entra ID
ChannelGuid	The ID of a Microsoft Teams channel. The team that the channel is located in is identified by the **TeamName** and **TeamGuid** properties.	Microsoft Teams
ChannelName	The name of a Microsoft Teams channel. The team that the channel is located in is identified by the **TeamName** and **TeamGuid** properties.	Microsoft Teams
Client	The client device, the device OS, and the device browser used for the sign-in activity (for example, Nokia Lumia 920; Windows Phone 8; IE Mobile 11).	Microsoft Entra ID
ClientInfoString	Information about the email client that was used to perform the operation, such as a browser version, Outlook version, and mobile device information	Exchange (mailbox activity)
ClientIP	The IP address of the device that was used when the activity was logged. The IP address is displayed in either an IPv4 or IPv6 address format. For some services, the value displayed in this property might be the IP address for a trusted application (for example, Office on the web apps) calling into the service on behalf of a user and	Microsoft Entra ID, Exchange, SharePoint

Property	Description	Microsoft service that has this property
	not the IP address of the device used by person who performed the activity. Also, for admin activity (or activity performed by a system account) for Microsoft Entra ID-related activities, the IP address isn't logged and the value for the ClientIP property is `null`.	
CreationTime	The date and time in Coordinated Universal Time (UTC) when the audit log record is generated.	All
CurrentProtectionType	A complex property type containing fields to describe the current protection status of a document. Includes the following: **ProtectionType**: Enumerates the type of protection applied to the document. These values and their meanings apply: *0* (no protection), *1* (template-based protection), *2* (don't forward, for email), *3* (encrypt only), and *4* (custom, user configured protection) **Owner**: The email address of the user that configured protection. **TemplateId**: When the *ProtectionType* is set to *1* (template), this field contains the GUID of the template applied to the document. When the value of *ProtectionType* doesn't equal *1*, this field is blank. **DocumentEncrypted**: Boolean flag indicating if any type of encryption is applied to the document. Values are *True* or *False*.	All
DestinationFileExtension	The file extension of a file that is copied or moved. This property is displayed only for the FileCopied and FileMoved user activities.	SharePoint
DestinationFileName	The name of the file is copied or moved. This property is displayed only for the FileCopied and FileMoved actions.	SharePoint
DestinationRelativeUrl	The URL of the destination folder where a file is copied or moved. The combination of the values for the **SiteURL**, the **DestinationRelativeURL**, and the **DestinationFileName** property is the same as the value for the **ObjectID** property, which is the full path name for the file that was copied. This	SharePoint

Property	Description	Microsoft service that has this property
	property is displayed only for the FileCopied and FileMoved user activities.	
EventSource	Identifies that an activity occurred in SharePoint. Possible values are **SharePoint** and **ObjectModel**.	SharePoint
ExtendedProperties	The extended properties for Microsoft Purview and Microsoft Entra ID activities.	Microsoft Entra ID Microsoft Purview
ExternalAccess	For Exchange admin activity, specifies whether the cmdlet was run by a user in your organization, by Microsoft datacenter personnel or a datacenter service account, or by a delegated administrator. The value **False** indicates that the cmdlet was run by someone in your organization. The value **True** indicates that the cmdlet was run by datacenter personnel, a datacenter service account, or a delegated administrator. For Exchange mailbox activity, specifies whether a mailbox was accessed by a user outside your organization.	Exchange
ID	The ID of the report entry. The ID uniquely identifies the report entry.	All
InternalLogonType	Reserved for internal use.	Exchange (mailbox activity)
IsJoinedFromLobby	Whether or not the user joined a Teams session from the lobby.	Microsoft Teams
ItemType	The type of object that was accessed or modified. Possible values include **File**, **Folder**, **Web**, **Site**, **Tenant**, and **DocumentLibrary**.	SharePoint
LoginStatus	Identifies sign-in failures that might have occurred.	Microsoft Entra ID
LogonType	The type of mailbox access. The following values indicate the type of user who accessed the mailbox.	Exchange (mailbox activity)

Property	Description	Microsoft service that has this property
	0 - Indicates a mailbox owner. **1** - Indicates an administrator. **2** - Indicates a delegate. **3** - Indicates the transport service in the Microsoft datacenter. **4** - Indicates a service account in the Microsoft datacenter. **6** - Indicates a delegated administrator.	
MailboxGuid	The Exchange GUID of the mailbox that was accessed.	Exchange (mailbox activity)
MailboxOwnerUPN	The email address of the person who owns the mailbox that was accessed.	Exchange (mailbox activity)
Members	Lists the users that have been added or removed from a team. The following values indicate the Role type assigned to the user. **1** - Indicates the Member role. **2** - Indicates the Owner role. **3** - Indicates the Guest role. The Members property also includes the name of your organization, and the member's email address.	Microsoft Teams
ModifiedProperties (Name, NewValue, OldValue)	The property is included for admin activities, such as adding a user as a member of a site or a site collection admin group. The property includes the name of the property that was modified (for example, the Site Admin group) the new value of the modified property (such the user who was added as a site admin, and the previous value of the modified object).	All (admin activity)
NewValue	The value after a change, includes all properties updated or deleted.	Microsoft Purview (governance)
ObjectFullyQualifiedName	The fully qualified name for an entity.	Microsoft Purview (governance)

Property	Description	Microsoft service that has this property
ObjectId	For Exchange admin audit logging, the name of the object that was modified by the cmdlet. For SharePoint activity, the full URL path name of the file or folder accessed by a user. For Azure AD activity, the name of the user account that was modified.	All
ObjectName	The main entity name.	Microsoft Purview (governance)
ObjectType	The entity type.	Microsoft Purview (governance)
OldValue	The value before a change, includes all properties updated or deleted.	Microsoft Purview (governance)
Operation	The name of the user or admin activity. The value of this property corresponds to the value that was selected in the **Activities** drop down list. If **Show results for all activities** was selected, the report will included entries for all user and admin activities for all services. For a description of the operations/activities that are logged in the audit log, see the **Audited activities** tab in Search the audit log in the Office 365. For Exchange admin activity, this property identifies the name of the cmdlet that was run.	All
OrganizationId	The GUID for your organization.	All
Parameters	For Exchange admin activity, the name and value for all parameters that were used with the cmdlet that is identified in the Operation property.	Exchange (admin activity)
ParticipantInfo	Additional properties about the participant identity.	Microsoft Teams
ParticipatingDomainInformation	Domain information about the participant.	Microsoft Teams
Path	The name of the mailbox folder where the message that was accessed is located. This	Exchange (mailbox activity)

Property	Description	Microsoft service that has this property
	property also identifies the folder a where a message is created in or copied/moved to.	
PreviousProtectionType	A complex property type containing fields to describe the previous protection status of a document. Includes the following: **ProtectionType**: Enumerates the type of protection applied to the document. These values and their meanings apply: *0* (no protection), *1* (template-based protection), *2* (don't forward, for email), *3* (encrypt only), *4* (custom, user configured protection) **Owner**: The email address of the user that configured protection. **TemplateId**: When the *ProtectionType* is set to *1* (template), this field contains the GUID of the template applied to the document. When the value of *ProtectionType* doesn't equal *1*, this field is blank. **DocumentEncrypted**: Boolean flag indicating if any type of encryption is applied to the document. Values are *True* or *False*.	All
ProtectionEventType	Enumerates how the protection was changed by the operation being audited. The following values and meanings apply: **0** - Indicates unchanged. **1** - Indicates added. **2** - Indicates changed. **3** - Indicates removed.	All
RecordType	The type of operation indicated by the record. This property indicates the service or feature that the operation was triggered in. For a list of record types and their corresponding ENUM value (which is the value displayed in the **RecordType** property in an audit record), see Audit log record type.	
ResultStatus	Indicates whether the action (specified in the **Operation** property) was successful or not. For Exchange admin activity, the value is either **True** (successful) or **False** (failed).	All

Property	Description	Microsoft service that has this property
SecurityComplianceCenterEventType	Indicates that the activity was a Microsoft Purview portal activity. All Microsoft Purview portal activities have a value of **0** for this property.	Microsoft Purview portal
SensitivityLabel	The sensitivity label assigned to a specific mail item.	Exchange
SharingType	The type of sharing permissions that you assigned to the user that you shared the resource with. This user is identified in the **UserSharedWith** property.	SharePoint
Site	The GUID of the site where the file or folder accessed by the user is located.	SharePoint
SiteUrl	The URL of the site where the file or folder accessed by the user is located.	SharePoint
SourceFileExtension	The file extension of the file that the user accessed. This property is blank if the object that was accessed is a folder.	SharePoint
SourceFileName	The name of the file or folder accessed by the user.	SharePoint
SourceRelativeUrl	The URL of the folder that contains the file the user accessed. The combination of the values for the **SiteURL**, the **SourceRelativeURL**, and the **SourceFileName** property is the same as the value for the **ObjectID** property, which is the full path name for the file the user accessed.	SharePoint
Subject	The subject line of the message that was accessed.	Exchange (mailbox activity)
TabType	The type of tab added, removed, or updated in a team. The possible values for this property are: **Excel pin** - An Excel tab. **Extension** - All first-party and third-party apps; such as Class Schedule, VSTS, and Forms. **Notes** - OneNote tab. **Pdfpin** - A PDF tab. **Powerbi** - A Power BI tab.	Microsoft Teams

Property	Description	Microsoft service that has this property
	Powerpointpin - A PowerPoint tab. **Sharepointfiles** - A SharePoint tab. **Webpage** - A pinned website tab. **Wiki-tab** - A wiki tab. **Wordpin** - A Word tab.	
Target	The user that the action (identified in the **Operation** property) was performed on. For example, if a guest is added to SharePoint or a Microsoft Team, that user is listed in this property.	Microsoft Entra ID
TeamGuid	The ID of a team in Microsoft Teams.	Microsoft Teams
TeamName	The name of a team in Microsoft Teams.	Microsoft Teams
UserAgent	Information about the user's browser. The browser provides this information.	SharePoint
UserDomain	Identity information about the tenant organization of the user (actor) who performed the action.	Microsoft Entra ID
UserId	The user who performed the action (specified in the **Operation** property) that resulted in the record being logged. Audit records for activity performed by system accounts (such as SHAREPOINT\system or NT AUTHORITY\SYSTEM) are also included in the audit log. Another common value for the UserId property is app@sharepoint. This value indicates that the "user" who performed the activity was an application that has the necessary permissions in SharePoint to perform organization-wide actions (such as search a SharePoint site or OneDrive account) on behalf of a user, admin, or service. For more information, see: The app@sharepoint user in audit records or System accounts in Exchange mailbox audit records.	All

Property	Description	Microsoft service that has this property
UserKey	Contains a valid Microsoft Entra ID Object ID in GUID format or hex format. For scenarios where the primary actor isn't a user, the *UserKey* is an empty string. See UserType and UserKey scenarios for details on various *UserKey* scenarios.	All
UserType	The type of user that performed the operation. See the UserType and UserKey scenarios for details on various *UserType* scenarios.	All
Version	Indicates the version number of the activity (identified by the **Operation** property) that's logged.	All
Workload	The Microsoft 365 service where the activity occurred.	All

UserType and UserKey scenarios

The following table provides details for *UserType* and *UserKey* scenarios:

Expand table

Value	UserType member name	Description	UserKey
0	Regular	A regular user without admin permissions.	Microsoft Entra Object ID in GUID format
2	Admin	An administrator in your Microsoft 365 organization.[1]	Microsoft Entra Object ID in GUID format
3	DCAdmin	A Microsoft datacenter administrator or datacenter system account.	Microsoft Entra Object ID in GUID format
4	System	An audit event triggered by server-side logic. For example, Windows services or background processes.	Guid.Empty.ToString() (or the value '00000000-0000-0000-0000-000000000000').

Value	UserType member name	Description	UserKey
5	Application	An audit event triggered by a Microsoft Entra application.	Microsoft Entra Application Name or Application ID (when available). Otherwise, an empty string.
6	ServicePrincipal	A service principal.	Guid.Empty.ToString() (or the value '00000000-0000-0000-0000-000000000000').
7	CustomPolicy	A customer created or managed policy.	Guid.Empty.ToString() (or the value '00000000-0000-0000-0000-000000000000').
8	SystemPolicy	A Microsoft-managed or system policy.	Guid.Empty.ToString() (or the value '00000000-0000-0000-0000-000000000000').
9	PartnerTechnician	A partner tenant's user working on behalf of the customer tenant (in GDAP scenarios).	Guid.Empty.ToString() (or the value '00000000-0000-0000-0000-000000000000').
10	Guest	A guest or anonymous user.	Guid.Empty.ToString() (or the value '00000000-0000-0000-0000-000000000000').
11	Agent	An AI agent	Microsoft Entra Object ID of the primary agent in GUID format.

Note

[1] For Microsoft Entra related events, the value for an administrator isn't used in an audit record. Audit records for activities performed by administrators indicate that a regular user (for example, **UserType: 0**) performed the activity. The **UserID** property identifies the person (regular user or administrator) who performed the activity.

Last updated on 02/18/2026

Learn about properties in Exchange audit records

This article explains the properties and fields you see in Exchange audit records returned by Search-UnifiedAuditLog. It's written for Exchange and Microsoft 365 admins who investigate mailbox activity or build scripts to parse audit output. You find a concise description of primary properties, representative AuditData JSON for common operations, PowerShell extraction examples, and a quick reference for common property values and investigation scenarios.

Use this reference to understand:

- Key properties in audit record results
- What each field represents and contains
- Common property values and their meanings
- How to extract basic information from audit data

Audit record structure reference

Exchange audit records contain several key properties that provide different types of information about the audited activity.

Primary audit record properties

When you run Search-UnifiedAuditLog, each result contains these main properties:

Expand table

Property	Data Type	Description	Example value
AuditData	JSON String	Detailed JSON data about the activity	Complex JSON structure
CreationDate	DateTime	When the activity occurred	11/17/2025 2:30:15 PM
Identity	GUID	Unique identifier for the audit record	00aa00aa-bb11-cc22-dd33-44ee44ee44ee
Operations	String	What type of activity occurred	SoftDelete, HardDelete, Move
ResultCount	Integer	Total results available	150

Property	Data Type	Description	Example value
ResultIndex	Integer	Position in search results	1, 2, 3...
UserIds	String	Who performed the activity	`<user@domain.com>`

AuditData JSON structure reference

The *AuditData* property contains detailed information in JSON format. Here are some typical JSON response structures for common operations:

- **Email deletion (SoftDelete/HardDelete)**

JSON


```
{
  "CreationTime": "2025-11-17T14:30:15",
  "Id": "00aa00aa-bb11-cc22-dd33-44ee44ee44ee",
  "Operation": "SoftDelete",
  "OrganizationId": "00aa00aa-bb11-cc22-dd33-44ee44ee44ee",
  "RecordType": 2,
  "ResultStatus": "Success",
  "UserKey": "user@domain.com",
  "UserType": 0,
  "Version": 1,
  "Workload": "Exchange",
  "ClientIP": "192.168.1.100",
  "ObjectId": "AAMkADM2NDIyMzUzLWE1ZjQtNGVkNS04OGE3LTg5NzY4ZDc1ZTZhNwBGAAAAAAC7...",
  "UserId": "user@domain.com",
  "ClientInfoString": "Client=OWA;Mozilla/5.0...",
  "ExternalAccess": false,
  "InternalLogonType": 0,
  "MailboxGuid": "87654321-4321-4321-4321-210987654321",
  "MailboxOwnerUPN": "mailboxowner@domain.com",
  "LogonType": 0,
  "Item": {
    "Id": "AAMkADM2NDIyMzUzLWE1ZjQtNGVkNS04OGE3LTg5NzY4ZDc1ZTZhNwBGAAAAAAC7...",
    "Subject": "Meeting Request - Q4 Planning",
    "ParentFolder": {
      "Id": "AAMkADM2NDIyMzUzLWE1ZjQtNGVkNS04OGE3LTg5NzY4ZDc1ZTZhNwAuAAAAAAC7...",
      "Name": "Inbox",
      "Path": "[\\Inbox](file:///\\inbox\)"
    }
  },
  "SessionId": "11111111-2222-3333-4444-555555555555"
}
```


- **Mailbox rule creation (New-InboxRule)**

JSON


```
{
  "CreationTime": "2025-11-17T14:45:22",
  "Id": "98765432-8765-8765-8765-876543218765",
  "Operation": "New-InboxRule",
  "OrganizationId": "00aa00aa-bb11-cc22-dd33-44ee44ee44ee",
  "RecordType": 1,
  "ResultStatus": "Success",
  "UserKey": "user@domain.com",
  "UserType": 0,
  "Version": 1,
  "Workload": "Exchange",
  "ClientIP": "10.0.0.50",
  "ObjectId": "InboxRule:aaaaaaaa-0000-1111-2222-bbbbbbbbbbbb",
  "UserId": "user@domain.com",
  "ClientInfoString": "Client=WebServices;ExchangeWebServices",
  "MailboxGuid": "87654321-4321-4321-4321-210987654321",
  "MailboxOwnerUPN": "mailboxowner@domain.com",
  "Parameters": \[
    {
      "Name": "Name",
      "Value": "Move Microsoft Security Emails"
    },
    {
      "Name": "MoveToFolder",
      "Value": "[\\Inbox\\Security](file:///\\inbox\Security)"
    },
    {
      "Name": "From",
      "Value": "security-noreply@microsoft.com"
    }
  \],
  "SessionId": "22222222-3333-4444-5555-666666666666"
}
```


- Mailbox access (MailItemsAccessed)

JSON


```
{
  "CreationTime": "2025-11-17T15:00:10",
  "Id": "13579246-1357-1357-1357-135792468135",
  "Operation": "MailItemsAccessed",
  "OrganizationId": "00aa00aa-bb11-cc22-dd33-44ee44ee44ee",
  "RecordType": 50,
  "ResultStatus": "Success",
  "UserKey": "user@domain.com",
  "UserType": 0,
  "Version": 1,
  "Workload": "Exchange",
  "ClientIP": "203.0.113.45",
  "UserId": "user@domain.com",
```

```
  "ClientInfoString": "Client=ActiveSync;Apple-iPhone/1309.63",
  "MailboxGuid": "87654321-4321-4321-4321-210987654321",
  "MailboxOwnerUPN": "mailboxowner@domain.com",
  "ClientAppId": "00000002-0000-0ff1-ce00-000000000000",
  "Folders": \[
    {
      "Id": "AAMkADM2NDIyMzUzLWE1ZjQtNGVkNS04OGE3LTg5NzY4ZDc1ZTZhNwAuAAAAAAC7...",
      "Path": "[\\Inbox](file:///\\inbox\)",
      "FolderItems": \[
        {
          "Id": "AAMkADM2NDIyMzUzLWE1ZjQtNGVkNS04OGE3LTg5NzY4ZDc1ZTZhNwBGAAAAAAC7...",

          "Subject": "Weekly Status Report"
        }
      \]
    }
  \],
  "OperationCount": 5,
  "SessionId": "33333333-4444-5555-6666-777777777777"
}
```


Here are the key fields organized by category:

User and session information

Expand table

Field	Data type	Description	Example value
ClientInfoString	String	Client application information	Client=OWA, Client=WebServices
ClientIP	String	IP address of the client	192.168.1.100
SessionId	GUID	Unique session identifier	00aa00aa-bb11-cc22-dd33-44ee44ee44ee
UserId	String	User who performed the action	user@domain.com
UserKey	String	User identifier key	user@domain.com
UserType	String	Type of user account	Regular, Admin, System, Application

Activity details

Expand table

Field	Data type	Description	Example value
CreationTime	DateTime	When the activity occurred (UTC)	2025-11-17T14:30:15Z

Field	Data type	Description	Example value
Operation	String	Specific operation performed	SoftDelete, New-InboxRule, MailItemsAccessed
OrganizationId	GUID	Organization identifier	00aa00aa-bb11-cc22-dd33-44ee44ee44ee
RecordType	String	Type of audit record	ExchangeItem, ExchangeAdmin
Workload	String	Microsoft 365 service	Exchange, SharePoint, OneDrive

Target information

Expand table

Field	Data type	Description	Example value
Item.Id	String	Exchange item identifier	AAMkADM2...
Item.ParentFolder.Name	String	Folder name where item was located	Inbox, Sent Items, Deleted Items
Item.ParentFolder.Path	String	Full folder path	\Inbox, \Deleted Items
Item.Subject	String	Email subject line	Meeting Request
MailboxGuid	GUID	Target mailbox identifier	00aa00aa-bb11-cc22-dd33-44ee44ee44ee
MailboxOwnerUPN	String	Mailbox owner's email address	mailboxowner@domain.com

Basic data extraction techniques

Use these techniques to extract and inspect key values from the AuditData JSON payload.

Convert JSON data

To extract information from the `AuditData` JSON, use the following commands:

```
\$AuditData = ConvertFrom-Json \$Results[0].AuditData
\$AuditData.UserId
\$AuditData.ClientIP
\$AuditData.MailboxOwnerUPN
```

Display key properties

To see the main properties in a readable format, use the following command:

```
\$Results \| Select CreationDate, UserIds, Operations, @{Name="MailboxOwner";Expression={(ConvertFrom-Json \$\_.AuditData).MailboxOwnerUPN}} \| Format-Table -AutoSize
```

Understanding common property values

The following tables help you understand common values found in audit record properties.

Operation types

The following table lists common Exchange audit operations and what they typically indicate during an investigation.

Expand table

Operation	Description	Investigation focus
HardDelete	Item permanently removed	Permanent deletions
MailItemsAccessed	Mailbox items were accessed	Unauthorized access investigation
Move	Item moved between folders	Folder organization changes
New-InboxRule	New mailbox rule created	Rule creation investigation
Send	Email was sent	Email communication tracking
Set-InboxRule	Mailbox rule modified	Rule modification tracking
SoftDelete	Item moved to Deleted Items	User-initiated deletions

RecordType values

The following table explains common RecordType values and the Exchange activities they represent.

Expand table

RecordType	Workload	Typical operations
ExchangeAdmin	Exchange	New-InboxRule, Set-InboxRule, Set-Mailbox
ExchangeAggregatedOperation	Exchange	MailItemsAccessed (bulk operations)
ExchangeItem	Exchange	SoftDelete, HardDelete, Move, Send

UserType values

UserType identifies what kind of account performed the audited action.

Expand table

UserType	Description	Investigation notes
Admin	Administrator account	Administrative actions
Application	Application service principal	App-based access
Regular	Standard user account	Normal user activity
System	System or service account	Automated processes

ClientInfoString examples

ClientInfoString values show which client, protocol, or access method was used.

Expand table

ClientInfoString	Description	Access method
Client=ActiveSync	Exchange ActiveSync	Mobile device sync
Client=IMAP4	IMAP protocol	IMAP client access
Client=OWA	Outlook Web App	Web browser access
Client=POP3	POP3 protocol	POP3 client access
Client=WebServices	Exchange Web Services (EWS)	Programmatic access

Quick reference

Use the following tables as a quick lookup for the most useful audit properties and PowerShell techniques.

Essential properties for investigations

The following table highlights which audit properties are most useful for common investigation scenarios.

Expand table

Investigation type	Key properties to focus on
Bulk Operations	SessionId, ItemCount, ClientIP, UserIds, Operations
Email Deletions	CreationDate, UserIds, Operations, Item.Subject, Item.ParentFolder.Name
Rule Changes	UserIds, Operations, Parameters, ClientIP, MailboxOwnerUPN
Unauthorized Access	UserIds, ClientIP, ClientInfoString, MailboxOwnerUPN, SessionId

Useful PowerShell techniques

These PowerShell techniques help you inspect, filter, and present Exchange audit records more effectively.

Expand table

Technique	Command example	Purpose
Convert JSON	`ConvertFrom-Json \$\_.AuditData`	Extract detailed audit information
Filter Data	`Where-Object {\$\_.ClientIP -like "192.168.\*"}`	Focus on specific criteria
Format Output	`Format-Table -Wrap -AutoSize`	Display results clearly
Group Results	`Group-Object ClientIP`	Identify patterns and anomalies
Select Properties	Select CreationDate, UserIds, Operations	Show only relevant information

Next steps

- Search the audit log for mailbox activities: Search for specific mailbox activities.
- Use a PowerShell script to search the audit log: Automate audit log searches with PowerShell.
- Use MailItemsAccessed to investigate compromised accounts: Investigate mailbox access patterns.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

Last updated on 06/24/2026

Microsoft services that support auditing

Summarize this article for me

You can search the unified audit log for activities performed in different Microsoft services. The following table lists the Microsoft services, apps, and features that the unified audit log supports.

Expand table

Microsoft service or feature	Record types
Microsoft Entra ID	AzureActiveDirectory, AzureActiveDirectoryAccountLogon, AzureActiveDirectoryStsLogon
Azure Information Protection (For the Microsoft Purview Information Protection client and scanner)	AipDiscover, AipSensitivityLabelAction, AipProtectionAction, AipFileDeleted, AipHeartBeat
Communication compliance	ComplianceSupervisionExchange
Content explorer	LabelContentExplorer
Data connectors	ComplianceConnector
Data loss prevention (DLP)	ComplianceDLPSharePoint, ComplianceDLPExchange, DLPEndpoint
Dynamics 365	CRM
eDiscovery (Standard + Premium)	Discovery, AeD
Encrypted message portal	OMEPortal
Exact Data Match	MipExactDataMatch
Exchange Online	ExchangeAdmin, ExchangeItem, ExchangeItemAggregated
Forms	MicrosoftForms

Microsoft service or feature	Record types
Information barriers	InformationBarrierPolicyApplication
Microsoft Defender XDR	AirInvestigation, AirManualInvestigation, AirAdminActionInvestigation, MS365DCustomDetection
Microsoft 365 Copilot and Microsoft 365 Copilot Chat	CopilotInteraction
Microsoft Defender Experts	DefenderExpertsforXDRAdmin
Microsoft Defender for Identity (MDI)	MicrosoftDefenderForIdentityAudit
Microsoft Planner	PlannerCopyPlan, PlannerPlan, PlannerPlanList, PlannerRoster, PlannerRosterSensitivityLabel, PlannerTask, PlannerTaskList, PlannerTenantSettings
Microsoft Project for the web	ProjectAccessed, ProjectCreated, ProjectDeleted, ProjectTenantSettingsUpdated, ProjectUpdated, RoadmapAccessed,RoadmapCreated, RoadmapDeleted, RoadmapItemAccessed,RoadmapItemCreated,RoadmapItemDeleted, RoadmapItemUpdated, RoadmapTenantSettingsUpdated, RoadmapUpdated, TaskAccessed, TaskCreated,TaskDeleted, TaskUpdated
Microsoft Purview Information Protection (MIP) labels	MIPLabel, MipAutoLabelExchangeItem, MipAutoLabelSharePointItem, MipAutoLabelSharePointPolicyLocation
Microsoft Teams	MicrosoftTeams
Microsoft To Do	MicrosoftToDo, MicrosoftToDoAudit
MyAnalytics	MyAnalyticsSettings
OneDrive for Business	OneDrive
Power Apps	PowerAppsApp, PowerAppsPlan
Power Automate	MicrosoftFlow
Power BI	PowerBIAudit
Quarantine	Quarantine

Microsoft service or feature	Record types
Sensitive information types	DlpSensitiveInformationType
Sensitivity labels	MIPLabel, SensitivityLabelAction, SensitivityLabeledFileAction, SensitivityLabelPolicyMatch
SharePoint Online	SharePoint, SharePointFileOperation,SharePointSharingOperation, SharePointListOperation, SharePointCommentOperation
Stream	MicrosoftStream
SystemSync	DataShareCreated, DataShareDeleted, GenerateCopyOfLakeData, DownloadCopyOfLakeData
Threat Intelligence	ThreatIntelligence, ThreatIntelligenceUrl, ThreatFinder, ThreatIntelligenceAtpContent
Viva Goals	VivaGoals
Viva Insights	VivaInsights
Viva Engage	Yammer

For more information about the operations that each service audits, see the Audit log activities article.

The previous table also identifies the record type value to use to search the audit log for activities in the corresponding service by using the **Search-UnifiedAuditLog** cmdlet in Exchange Online PowerShell or by using a PowerShell script. Some services have multiple record types for different types of activities within the same service. For a more complete list of auditing record types, see Office 365 Management Activity API schema.

For more information about using PowerShell to search the audit log, see:

- Search-UnifiedAuditLog
- Use a PowerShell script to search the audit log

Last updated on 02/18/2026

Use the Microsoft Graph security API

The Microsoft Graph security API provides a unified interface and schema to integrate with security solutions from Microsoft and ecosystem partners. This empowers customers to streamline security operations and better defend against increasing cyber threats. The Microsoft Graph security API federates queries to all onboarded security providers and aggregates responses. Use the Microsoft Graph security API to build applications that:

- Consolidate and correlate security alerts from multiple sources.
- Pull and investigate all incidents and alerts from services that are part of or integrated with Microsoft 365 Defender.
- Unlock contextual data to inform investigations.
- Automate security tasks, business processes, workflows, and reporting.
- Send threat indicators to Microsoft products for customized detections.
- Invoke actions to in response to new threats.
- Provide visibility into security data to enable proactive risk management.

The Microsoft Graph security API provides key features as described in the following sections.

Advanced hunting

Advanced hunting is a query-based threat-hunting tool that lets you explore up to 30 days of raw data. You can proactively inspect events in your network to locate threat indicators and entities. The flexible access to data enables unconstrained hunting for both known and potential threats.

Use runHuntingQuery to run a Kusto Query Language (KQL) query on data stored in Microsoft 365 Defender. Use the returned result set to enrich an existing investigation or to uncover undetected threats in your network.

Quotas and resource allocation

The following conditions relate to all queries.

1. Queries explore and return data from the past 30 days.
2. Results can return up to 100,000 rows.
3. You can make up to at least 45 calls per minute per tenant. The number of calls varies per tenant based on its size.
4. Each tenant is allocated CPU resources, based on the tenant size. Queries are blocked if the tenant reaches 100% of the allocated resources until after the next 15-minute cycle. To avoid blocked queries due to excess consumption, follow the guidance in Optimize your queries to avoid hitting CPU quotas.
5. If a single request runs for more than three minutes, it times out and returns an error.
6. A `429` HTTP response code indicates that you reached the allocated CPU resources, either by the number of requests sent or by allotted running time. Read the response body to understand the limit you reached.
7. Query results have an overall size limit of 50 MB. This limit doesn't just refer to the number of records; factors such as the number of columns, data types, and field lengths also contribute to the query result size.

Migrate from the older APIs

The advanced hunting APIs in Microsoft Graph replace the older version of the API that was available through the `https://api.security.microsoft.com/api/advancedhunting/run` and `https://api.security.microsoft.com/api/advancedqueries/run` endpoints. The older APIs are now retired and will stop returning data on February 1, 2027.

To migrate to the advanced hunting APIs in Microsoft Graph, update the following parameters in your application:

Expand table

Subject	Older parameters	Microsoft Graph
Endpoints	https://api.securitycenter.microsoft.com/api/advancedqueries/run https://api.security.microsoft.com/api/advancedhunting/run	https://graph.microsoft.com/beta/security/run
Resource URI	Microsoft Defender for Endpoint on `https://api.securitycenter.microsoft.com`	Microsoft Graph on `https://graph.microsoft.`
API permissions	*AdvancedQuery.Read* (delegated) and *AdvancedQuery.Read.All* (application) under Microsoft Defender for Endpoint (formerly Windows Defender Advanced Threat Protection) *AdvancedHunting.Read* (delegated) and *AdvancedHunting.Read.All* (application) under Microsoft Threat Protection	*ThreatHunting.Read.All* (delegated and applica
Request body	**Query** property. For example `{"Query":"DeviceProcessEvents |where InitiatingProcessFileName =~ 'powershell.exe' |where ProcessCommandLine contains 'appdata'|project Timestamp, FileName, InitiatingProcessFileName, DeviceId|limit 2"}`	**Query** and **Timespan** properties. For example, `"DeviceProcessEvents", "Timespan": "P90D"}`
Response	**QueryResponse** object consisting of **Stats**, **Schema**, and **Results**	huntingQueryResults resource type", consistin (instead of **Schema**) and **results** (instead of **Re**

For more information about how to authorize your app to call Microsoft Graph APIs, see Get access on behalf of a user and Get access without a user.

Power Platform flow migration (PowerApps / Power Automate / Logic Apps)

Microsoft Graph does't have a built-in Advanced Hunting action that was available in the Power Platform connector for Microsoft Defender ATP. To continue using Advanced Hunting in your Power Platform flows, create a custom connector. For more information, see Create a Microsoft Graph JSON Batch Custom Connector for Power Automate and use the Microsoft Graph parameters described in the preceding table.

Power BI flow migration

If you're using a custom Power BI report created with the older API, update your Power BI query to use the Microsoft Graph parameters described in the preceding table. For more information, see Create custom Microsoft Defender XDR reports using Microsoft Graph security API and Power BI.

Alerts

Alerts are detailed warnings about suspicious activities in a customer's tenant that Microsoft or partner security providers identified and flagged for action. Attacks typically employ various techniques against different types of entities, such as devices, users, and mailboxes. The result is alerts from multiple security providers for multiple entities in the tenant. Piecing the individual alerts together to gain insight into an attack can be challenging and time-consuming.

The security API offers two types of alerts that aggregate other alerts from security providers and make analyzing attacks and determining responses easier:

- Alerts and incidents - these are the latest generation of alerts in the Microsoft Graph security API. They're represented by the alert resource and its collection, incident resource, defined in the `microsoft.graph.security` namespace.
- Legacy alerts - these are the first generation of alerts in the Microsoft Graph security AI. They're represented by the alert resource defined in the `microsoft.graph` namespace.

> **Important**
>
> To view Sentinel alerts and incidents you must onboard Sentinel to the Defender Portal. For more information see Connect Microsoft Sentinel to the Microsoft Defender portal.

Alerts and incidents

These alert resources first pull alert data from security provider services, that are either part of or integrated with Microsoft 365 Defender. Then they consume the data to return rich, valuable clues about a completed or ongoing attack, the impacted assets, and associated evidence. In addition, they automatically correlate other alerts with the same attack techniques or the same attacker into an incident to provide a broader context of an attack. They recommend response and remediation actions, offering consistent actionability across all the different providers. The rich content makes it easier for analysts to collectively investigate and respond to threats.

Alerts from the following security providers are available via these rich alerts and incidents:

- Microsoft Entra ID Protection
- Microsoft 365 Defender
- Microsoft Defender for Cloud Apps
- Microsoft Defender for Endpoint
- Microsoft Defender for Identity
- Microsoft Defender for Office 365
- Microsoft Purview Data Loss Prevention
- Microsoft Purview Insider Risk Management

Legacy alerts

> **Note**

The legacy alerts API is deprecated. We recommend that you migrate to the new alerts and incidents API.

The legacy alert resources federate calling of supported Azure and Microsoft 365 Defender security providers. They aggregate common alert data among the different domains to allow applications to unify and streamline management of security issues across all integrated solutions. They enable applications to correlate alerts and context to improve threat protection and response.

The legacy version of the security API offers the alert resource that federates calling of supported Azure and Microsoft 365 Defender security providers. This **alert** resource aggregates alert data that's common among the different domains to allow applications to unify and streamline management of security issues across all integrated solutions. This enables applications to correlate alerts and context to improve threat protection and response.

With the alert update capability, you can sync the status of specific alerts across different security products and services that are integrated with the Microsoft Graph security API by updating your **alert** entity.

Alerts from the following providers are available via the **alert** resource. Support for GET alerts, PATCH alerts, and subscribe (via webhooks) is indicated in the following table.

Expand table

Security provider	GET alert	PATCH alert	Subscribe to alert
Microsoft Entra ID Protection	✓	File issue *	✓
Microsoft 365 • Default • Cloud App Security • Custom Alert	✓	File issue	File issue
Microsoft Defender for Cloud Apps	✓	File issue *	✓
Microsoft Defender for Endpoint **	✓	✓	File issue
Microsoft Defender for Identity ***	✓	File issue *	✓
Microsoft Sentinel (formerly Azure Sentinel)	✓	Not supported in Microsoft Sentinel	✓

> **Note:** New providers are continuously onboarding to the Microsoft Graph security ecosystem. To request new providers or for extended support from existing providers, file an issue in the Microsoft Graph security GitHub repo.

* File issue: Alert status gets updated across Microsoft Graph security API integrated applications but not reflected in the provider's management experience.

** Microsoft Defender for Endpoint requires additional user roles to those required by the Microsoft Graph security API. Only the users in both Microsoft Defender for Endpoint and Microsoft Graph security API roles can access the Microsoft Defender for Endpoint data. Because

application-only authentication isn't limited by this, we recommend that you use an application-only authentication token.

*** Microsoft Defender for Identity alerts are available via the Microsoft Defender for Cloud Apps integration. This means you get Microsoft Defender for Identity alerts only if you joined Unified SecOps and connected Microsoft Defender for Identity to Microsoft Defender for Cloud Apps. Learn more about how to integrate Microsoft Defender for Identity and Microsoft Defender for Cloud Apps.

Attack simulation and training

Attack simulation and training is part of Microsoft Defender for Office 365. This service lets users in a tenant experience a realistic benign phishing attack and learn from it. Social engineering simulation and training experiences for end users help reduce the risk of users being breached via those attack techniques. The attack simulation and training API enables tenant administrators to view launched simulation exercises and trainings, and get reports on derived insights into online behaviors of users in the phishing simulations.

eDiscovery

Microsoft Purview eDiscovery provides an end-to-end workflow to preserve, collect, analyze, review, and export content that's responsive to your organization's internal and external investigations.

Audit log query

Microsoft Purview Audit provides an integrated solution to help organizations effectively respond to security events, forensic investigations, internal investigations, and compliance obligations. Thousands of user and admin operations performed in dozens of Microsoft 365 services and solutions are captured, recorded, and retained in your organization's unified audit log. Audit records for these events are searchable by security ops, IT admins, insider risk teams, and compliance and legal investigators in your organization. This capability provides visibility into the activities performed across your Microsoft 365 organization.

Identities

Health issues

The Microsoft Defender for Identity health issues API allows you to monitor the health status of your sensors and agents across your hybrid identity infrastructure. You can use the health issues API to retrieve information about the current health issues of your sensors, such as the issue type, status, configuration, and severity. You can also use this API to identify and resolve any issues that might affect the functionality or security of your sensors and agents.

> **Note:** The Microsoft Defender for Identity health issues API is only available on the Defender for Identity plan or Microsoft 365 E5/A5/G5/F5 Security service plans.

Sensors

The Defender for Identity sensors management APIs allows you to:

- Create detailed reports of the sensors in your workspace, including information about the server name, sensor version, type, state, and health status.
- Manage sensor settings, such as adding descriptions, enabling or disabling delayed updates, and specifying the domain controller that the sensor connects to for querying Entra ID.
- Identify sensors that are ready to be activated.
- Define whether the sensors in your infrastructure are to be activated automatically or manually.
- Identify servers that are ready to be activated with the unified agent.
- Enable or disable the automatic activation of eligible servers for the unified agent.
- Activate or deactivate the unified agent on eligible servers.
- Enable or disable the automatic enabling of the required events auditing configuration during the sensor's activation.

identityAccounts

The identityAccounts resource and related APIs allows you to retrieve details of users that are flagged by Microsoft Defender for Identity alerts, and apply actions such as disabling accounts and resetting the user password for the compromised user.

Incidents

An incident is a collection of correlated alerts and associated data that make up the story of an attack. Incident management is part of Microsoft 365 Defender and is available in the Microsoft 365 Defender portal (https://security.microsoft.com/).

Microsoft 365 services and apps create alerts when they detect a suspicious or malicious event or activity. Individual alerts provide valuable clues about a completed or ongoing attack. However, attacks typically employ various techniques against different types of entities, such as devices, users, and mailboxes. The result is multiple alerts for multiple entities in your tenant.

Because piecing the individual alerts together to gain insight into an attack can be challenging and time-consuming, Microsoft 365 Defender automatically aggregates the alerts and their associated information into an incident.

Grouping related alerts into an incident gives you a comprehensive view of an attack. For example, you can see:

- Where the attack started.
- What tactics were used.
- How far the attack went into your tenant.
- The scope of the attack, such as how many devices, users, and mailboxes were impacted.
- All of the data associated with the attack.

The incident resource and its APIs allow you to sort through incidents to create an informed cybersecurity response. It exposes a collection of incidents, with their related alerts, that were flagged in your network, within the time range you specified in your environment retention policy.

Information protection

The Microsoft Graph threat assessment API helps organizations assess the threat received by any user in a tenant. This empowers customers to report spam emails, phishing URLs, or malware attachments they receive to Microsoft. The policy check result and rescan result can help tenant administrators understand the threat scanning verdict and adjust their organizational policy.

Records management

Most organizations need to manage data to proactively comply with industry regulations and internal policies, reduce risk in the event of litigation or a security breach, and let people effectively and agilely share knowledge that is current and relevant to them. You can use the records management APIs to systematically apply retention labels to different types of content that require different retention settings. For example, you can configure the start of the retention period from when the content was created, last modified, labeled, or when an event occurs for a particular event type. Further, you can use file plan descriptors to improve the manageability of these retention labels.

Secure Score

Microsoft Secure Score is a security analytics solution that gives you visibility into your security portfolio and how to improve it. With a single score, you can better understand what you did to reduce your risk in Microsoft solutions. You can also compare your score with other organizations and see how your score has been trending over time. The Microsoft Graph security secureScore and secureScoreControlProfile entities help you balance your organization's security and productivity needs while enabling the appropriate mix of security features. You can also project what your score would be after you adopt security features.

Threat intelligence

Microsoft Defender Threat Intelligence delivers world-class threat intelligence to help protect your organization from modern cyber threats. You can use Threat Intelligence to identify adversaries and their operations, accelerate detection and remediation, and enhance your security investments and workflows.

The threat intelligence APIs allow you to operationalize intelligence found within the user interface. This includes finished intelligence in the forms of articles and intel profiles, machine intelligence such as IoCs and reputation verdicts, and enrichment data such as passive DNS, cookies, components, and trackers.

Common use cases

The following are some of the most popular requests for working with the Microsoft Graph security API.

Expand table

Use cases	REST resources	Try it in Graph Explorer
Update secure score control profiles	Update secureScoreControlProfile	https://graph.microsoft.com/v1.0/security/secureScoreControlProfiles/{id}
Alerts and incidents		
List alerts	List alerts	https://graph.microsoft.com/v1.0/security/alerts_v2
Update alert	Update alert	https://graph.microsoft.com/v1.0/security/alerts/{id}
List incidents	List incidents	https://graph.microsoft.com/v1.0/security/incidents
List incidents with alerts	List incidents	https://graph.microsoft.com/v1.0/security/incidents?$expand=alerts
Update incident	Update incident	https://graph.microsoft.com/v1.0/security/incidents/{id}
eDiscovery		
List eDiscovery cases	List eDiscoveryCases	https://graph.microsoft.com/v1.0/security/cases/eDiscoveryCases
List eDiscovery case operations	List caseOperations	https://graph.microsoft.com/v1.0/security/cases/ediscoveryCases/{id}/operations
Identities		
List health issues	List health issues	https://graph.microsoft.com/v1.0/security/identities/healthIssues
List sensors	List sensors	https://graph.microsoft.com/v1.0/security/identities/sensors
Legacy alerts		
List alerts	List alerts	https://graph.microsoft.com/v1.0/security/alerts
Update alerts	Update alert	https://graph.microsoft.com/v1.0/security/alerts/{alert-id}
Secure scores		
List secure scores	List secureScores	https://graph.microsoft.com/v1.0/security/secureScores
Get secure score	Get secureScore	https://graph.microsoft.com/v1.0/security/secureScores/{id}
List secure score control profiles	List secureScoreControlProfiles	https://graph.microsoft.com/v1.0/security/secureScoreControlProfiles

Use cases	REST resources	Try it in Graph Explorer
Get secure score control profile	Get secureScoreControlProfile	https://graph.microsoft.com/v1.0/security/secureScoreControlProfiles/{id}

You can use Microsoft Graph change notifications to subscribe to and receive notifications about updates to Microsoft Graph security entities.

Resources

Code and contribute to these Microsoft Graph security API samples:

- ASP.NET (C#) sample
- Python sample
- Node.js (JavaScript) sample

Engage with the community:

- Join the tech community
- Discuss on StackOverflow

Next steps

The Microsoft Graph security API can open up new ways for you to engage with different security solutions from Microsoft and its partners. Follow these steps to get started:

- Drill down into alerts, secureScore, and secureScoreControlProfiles.
- Try the API in the Graph Explorer. Under **Sample Queries**, choose **show more samples** and set the Security category to **on**.
- Try subscribing to and receiving notifications on entity changes.

Related content

Code and contribute to this Microsoft Graph security API sample:

- PowerShell sample

Explore other options to connect with the Microsoft Graph security API:

- Microsoft Graph security connectors for Logic Apps, Flow and Power Apps
- Jupyter Notebook samples

Engage with the community:

- Join the tech community
- Discuss on StackOverflow

Last updated on 01/21/2026

Office 365 Management Activity API reference

Article • 12/03/2024

Use the Office 365 Management Activity API to retrieve information about user, admin, system, and policy actions and events from Office 365 and Microsoft Entra activity logs.

You can use the actions and events from the Office 365 and Microsoft Entra audit and activity logs to create solutions that provide monitoring, analysis, and data visualization. These solutions give organizations greater visibility into actions taken on their content. These actions and events are also available in the Office 365 Activity Reports. For more information, see Search the audit log in Microsoft 365.

> **Tip**
>
> If you're interested in creating custom reports from Audit Logs, you might find the following blogs helpful.
>
> - Microsoft Purview audit log activities via O365 Management API - Part 1
> - Microsoft Purview audit log activities via O365 Management API - Part 2

The Office 365 Management Activity API is a REST web service that you can use to develop solutions using any language and hosting environment that supports HTTPS and X.509 certificates. The API relies on Microsoft Entra ID and the OAuth2 protocol for authentication and authorization. To access the API from your application, you'll need to first register it in Microsoft Entra ID and configure it with appropriate permissions. This will enable your application to request the OAuth2 access tokens it needs to call the API. For more information, see Get started with Office 365 Management APIs.

For information about the data that the Office 365 Management Activity API returns, see Office 365 Management Activity API schema.

> **Important**
>
> Before you can access data through the Office 365 Management Activity API, you must enable unified audit logging for your Office 365 organization. You do this by turning on the Office 365 audit log. For instructions, see Turn Office 365 audit log search on or off.

Working with the Office 365 Management Activity API

The Office 365 Management Activity API aggregates actions and events into tenant-specific content blobs, which are classified by the type and source of the content they contain. Currently, these content types are supported:

- Audit.AzureActiveDirectory
- Audit.Exchange
- Audit.SharePoint
- Audit.General (includes all other workloads not included in the previous content types)
- DLP.All (DLP events only for all workloads)

For details about the events and properties associated with these content types, see Office 365 Management Activity API schema.

To begin retrieving content blobs for a tenant, you first create a subscription to the desired content types. If you are retrieving content blobs for multiple tenants, you create multiple subscriptions to each of the desired content types, one for each tenant.

After you create a subscription, you can poll regularly to discover new content blobs that are available for download, or you can register a webhook endpoint with the subscription and we will send notifications to this endpoint as new content blobs are available.

> ⓘ **Note**
>
> When a subscription is created, it can take up to 12 hours for the first content blobs to become available for that subscription. The content blobs are created by collecting and aggregating actions and events across multiple servers and datacenters. As a result of this distributed process, the actions and events contained in the content blobs will not necessarily appear in the order in which they occurred. One content blob can contain actions and events that occurred prior to the actions and events contained in an earlier content blob. We are working to decrease the latency between the occurrence of actions and events and their availability within a content blob, but we can't guarantee that they appear sequentially.

> **Note**
>
> DLP sensitive data is only available in the activity feed API to users that have been granted "Read DLP sensitive data" permissions. For more on Data Loss Prevention (DLP) see Overview of Data Loss Prevention Policies

Activity API operations

All API operations are scoped to a single tenant and the root URL of the API includes a tenant ID that specifies the tenant context. The tenant ID is a GUID. For information about how to get the GUID, see Get started with Office 365 Management APIs.

Because the notifications we send to your webhook include the tenant ID, you can use the same webhook to receive notifications for all tenants.

The URL for the API endpoint that you use is based on the type of Microsoft 365 or Office 365 subscription plan for your organization.

Enterprise plan

HTTP

```
https://manage.office.com/api/v1.0/{tenant_id}/activity/feed/{operation}
```

GCC government plan

HTTP

```
https://manage-gcc.office.com/api/v1.0/{tenant_id}/activity/feed/{operation}
```

GCC High government plan

HTTP

```
https://manage.office365.us/api/v1.0/{tenant_id}/activity/feed/{operation}
```

DoD government plan

HTTP

```
https://manage.protection.apps.mil/api/v1.0/{tenant_id}/activity/feed/{operation}
```

All API operations require an Authorization HTTP header with an access token obtained from Microsoft Entra ID. The tenant ID in the access token must match the tenant ID in the root URL of the API and the access token must contain the ActivityFeed.Read claim (this corresponds to the permission [Read activity data for an organization] that you configured for you application in Microsoft Entra ID).

JSON

```
Authorization: Bearer eyJ0e...Qa6wg
```

The Activity API supports the following operations:

- **Start a subscription** to begin receiving notifications and retrieving activity data for a tenant.
- **Stop a subscription** to discontinue retrieving data for a tenant.
- **List current subscriptions**
- **List available content** and the corresponding content URLs.
- **Receiving notifications** sent by a webhook when new content is available.
- **Retrieving content** by using the content URL.
- **List notifications** sent by a webhook.
- **Retrieve resource friendly names** for objects in the data feed identified by guids.

Start a subscription

This operation starts a subscription to the specified content type. If a subscription to the specified content type already exists, this operation is used to:

- Update the properties of an active webhook.
- Enable a webhook that was disabled because of excessive failed notifications.
- Re-enable an expired webhook by specifying a later or null expiration date.
- Remove a webhook.

Expand table

Subscription	Description
Path	`/subscriptions/start?contentType={ContentType}`
Parameters	contentType - Must be a valid content type.
PublisherIdentifier	The tenant GUID of the vendor coding against the API. This is **not** the application GUID or the GUID of the customer using the application, but the GUID of the company writing the code. This parameter is used for throttling the request rate. Make sure this parameter is specified in all issued requests to get a dedicated quota. All requests received without this parameter will share the same quota.
Body	webhook - Optional JSON object with three properties: • **address**: Required HTTPS endpoint that can receive notifications. A test message will be sent to the webhook to validate the webhook before creating the subscription. • **authId**: Optional string that will be included as the WebHook-AuthID header in notifications sent to the webhook as a means of identifying and authorizing the source of the request to the webhook. • **expiration**: Optional datetime that indicates a datetime after which notifications should no longer be sent to the webhook.
Response	contentType - The content type specified in the call.
status	The status of the subscription. If a subscription is disabled, you'll not be able to list or retrieve content.
webhook	The webhook properties specified in the call together with the status of the webhook. If the webhook is disabled, you'll not receive notification, but you'll still be able to list and retrieve content, provided the subscription is enabled.

Sample request

JSON

```
POST {root}/subscriptions/start?
contentType=Audit.SharePoint&PublisherIdentifier=46b472a7-c68e-4adf-8ade-
3db49497518e
Content-Type: application/json; utf-8
Authorization: Bearer eyJ0e...Qa6wg

{
    "webhook" : {
        "address": "https://webhook.myapp.com/o365/",
        "authId": "o365activityapinotification",
        "expiration": ""
    }
```

```
}
```

Sample response

JSON

```
HTTP/1.1 200 OK
Content-Type: application/json; charset=utf-8

{
    "contentType": "Audit.SharePoint",
    "status": "enabled",
    "webhook": {
        "status": "enabled",
        "address":  "https://webhook.myapp.com/o365/",
        "authId": "o365activityapinotification",
        "expiration": null
    }
}
```

Webhook validation

When the /start operation is called and a webhook is specified, we will send a validation notification to the specified webhook address to validate that an active listener can accept and process notifications. If we do not receive an HTTP 200 OK response, the subscription will not be created. Or, if /start is being called to add a webhook to an existing subscription and a response of HTTP 200 OK is not received, the webhook will not be added and the subscription will remain unchanged.

Sample request

JSON

```
POST {webhook address}
Content-Type: application/json; charset=utf-8
Webhook-AuthID: (webhook authId, if provided)
Webhook-ValidationCode: (random opaque string)

{
    "validationCode": (random opaque string, same as header)
}
```

Sample response

JSON

```
HTTP/1.1 200 OK
```

Stop a subscription

This operation stops a subscription to the specified content type.

When a subscription is stopped, you'll no longer receive notifications and you'll not be able to retrieve available content. If the subscription is later restarted, you'll have access to new content from that point forward. You will not be able to retrieve content that was available between the time the subscription was stopped and restarted.

Expand table

Subscription	Description
Path	`/subscriptions/stop?contentType={ContentType}`
Parameters	contentType - Must be a valid content type.
PublisherIdentifier	The tenant GUID of the vendor coding against the API. This is **not** the application GUID or the GUID of the customer using the application, but the GUID of the company writing the code. This parameter is used for throttling the request rate. Make sure this parameter is specified in all issued requests to get a dedicated quota. All requests received without this parameter will share the same quota.
Body	(empty)
Response	(empty)

Sample request

JSON

```
POST {root}/subscriptions/stop?
contentType=Audit.SharePoint&PublisherIdentifier=46b472a7-c68e-4adf-8ade-
3db49497518e
Authorization: Bearer eyJ0e...Qa6wg
```

Sample response

JSON

```
HTTP/1.1 200 OK
```

List current subscriptions

This operation returns a collection of the current subscriptions together with the associated webhooks.

Expand table

	Subscription	Description
Path	`/subscriptions/list`	
Parameters	PublisherIdentifier	The tenant GUID of the vendor coding against the API. This is **not** the application GUID or the GUID of the customer using the application, but the GUID of the company writing the code. This parameter is used for throttling the request rate. Make sure this parameter is specified in all issued requests to get a dedicated quota. All requests received without this parameter will share the same quota.
Body	(empty)	
Response	JSON array	Each subscription will be represented by a JSON object with three properties: • **contentType**: Indicates the content type. • **status**: Indicates the status of the subscription. • **webhook**: Indicates the configured webhook, together with the status (enabled, disabled, expired) of the webhook. If a subscription does not have a webhook, the webhook property will be present but with null value.

Sample request

JSON

```
GET {root}/subscriptions/list?PublisherIdentifier=46b472a7-c68e-4adf-8ade-
3db49497518e
```

```
Authorization: Bearer eyJ0e...Qa6wg
```

Sample response

JSON

```
HTTP/1.1 200 OK
Content-Type: application/json; charset=utf-8

[
    {
        "contentType" : "Audit.SharePoint",
        "status": "enabled",
        "webhook": {
            "status": "enabled",
            "address": "https://webhook.myapp.com/o365/",
            "authId": "o365activityapinotification",
            "expiration": null
        }
    },

    ...

    {
        "contentType": "Audit.Exchange",
        "webhook": null
    }
]
```

List available content

This operation lists the content currently available for retrieval for the specified content type. The content is an aggregation of actions and events harvested from multiple servers across multiple datacenters. The content will be listed in the order in which the aggregations become available, but the events and actions within the aggregations are not guaranteed to be sequential. An error is returned if the subscription status is disabled.

Expand table

Subscription	Description
Path	`/subscriptions/content?contentType={ContentType}&startTime={0}&endTime={1}`

Subscription	Description
Parameters	contentType - Must be a valid content type.
PublisherIdentifier	The tenant GUID of the vendor coding against the API. This is **not** the application GUID or the GUID of the customer using the application, but the GUID of the company writing the code. This parameter is used for throttling the request rate. Make sure this parameter is specified in all issued requests to get a dedicated quota. All requests received without this parameter will share the same quota.
startTime endTime	Optional datetimes (UTC) indicating the time range of content to return, based on when the content became available. The time range is inclusive with respect to startTime (startTime <= contentCreated) and exclusive with respect to endTime (contentCreated < endTime), so that non-overlapping, incrementing time intervals can used to page through available content. • YYYY-MM-DD • YYYY-MM-DDTHH:MM • YYYY-MM-DDTHH:MM:SS Both must be specified (or both omitted) and they must be no more than 24 hours apart, with the start time no more than 7 days in the past. By default, if startTime and endTime are omitted, then the content available in the last 24 hours is returned. **NOTE**: Even though it is possible to specify a startTime and endTime more than 24 hours apart, this is not recommended. Furthermore, if you do get any results in response to a request for more than 24 hours, these could be partial results and should not be taken into account. The request should be issued with an interval of no more than 24 hours between the startTime and endTime.
Response	JSON array - The available content will be represented by JSON objects with the following properties: • **contentType**: Indicates the content type. • **contentId**: An opaque string that uniquely identifies the content. • **contentUri**: The URL to use when retrieving the content. • **contentCreated**: The datetime when the content was made available. • **contentExpiration**: The datetime after which the content will no longer be available for retrieval.

Sample request

JSON

```
GET {root}/subscriptions/content?
contentType=Audit.SharePoint&PublisherIdentifier=46b472a7-c68e-4adf-8ade-
3db49497518e
```

```
Authorization: Bearer eyJ0e...Qa6wg
```

Sample response

JSON

```
HTTP/1.1 200 OK
Content-Type: application/json; charset=utf-8

[
    {
        "contentType": "Audit.SharePoint",
        "contentId":
"492638008028$492638008028$f28ab78ad40140608012736e373933ebspo2015043022$4a8
1a7c326fc4aed89c62e6039ab833b$04",
        "contentUri": "https://manage.office.com/api/v1.0/f28ab78a-d401-
4060-8012-
736e373933eb/activity/feed/audit/492638008028$492638008028$f28ab78ad40140608
012736e373933ebspo2015043022$4a81a7c326fc4aed89c62e6039ab833b$04",
        "contentCreated": "2015-05-23T17:35:00.000Z",
        "contentExpiration": "2015-05-30T17:35:00.000Z"
    },
    ...
]
```

Pagination

When listing available content for a time range, the number of results returned is limited to prevent response timeouts. If there are more results in the specified time range than can be returned in single response, the results will be truncated and a header will be added to the response indicating the URL to use to retrieve the next page of results. The URL will contain the same *startTime* and *endTime* parameters that were specified in the original request, together with a parameter indicating the internal ID of the next page. If *startTime* and *endTime* were not specified in the original request, they will be set to reflect the 24-hour interval that preceded the original request.

JSON

```
HTTP/1.1 200 OK
Content-Type: application/json; charset=utf-8
NextPageUri:
https://manage.office.com/api/v1/{tenant_id}/activity/feed/subscriptions/con
tent?contentType=Audit.SharePoint&startTime=2015-10-01&endTime=2015-10-
```

```
02&nextPage=2015101900R022885001761
```

To list all available content for a specified time range, you might need to retrieve multiple pages until a response without the **NextPageUri** header is received.

Receiving notifications

Notifications are sent to the configured webhook for a subscription as new content becomes available. Because the notification includes the tenant identifier, you can use the same webhook to receive notifications for all tenants for which you have subscriptions.

The notification is made as an HTTP POST over TLS (TLS 1.0 and later versions) to the specified webhook address. If the webhook configuration includes an auth ID, we will send it as an HTTP header: Webhook-AuthID. Any response other than HTTP 200 OK will be considered a failure and the notification will be retried. You can also configure your webhook to require client certificate-based authentication and we will authenticate using the manage.office.com certificate.

The body of the request will contain an array of one or more JSON objects that represent the available content blobs. The number of content blobs in each notification is limited to keep the size of the notification relatively small. Because this limit might change, your implementation should query for the length of the array instead of expecting a fixed size. Each object will include the same properties returned by the /content operation, together with the GUID of the tenant to which the data belongs and the GUID of your application that created the subscriptions. This allows the webhook to establish context when it is being used with multiple tenants and applications.

- **tenantId**: The GUID of the tenant to which the content belongs.
- **clientId**: The GUID of your application that created the subscription.
- **contentType**: Indicates the content type.
- **contentId**: An opaque string that uniquely identifies the content.
- **contentUri**: The URL to use when retrieving the content.
- **contentCreated**: The datetime when the content was made available.
- **contentExpiration**: The datetime after which the content will no longer be available for retrieval.

The following is an example of a notification.

JSON

```
POST https://webhook.myapp.com/o365/
Content-Type: application/json; utf-8
Webhook-AuthID: o365activityapinotification

[
    {
        "tenantId": "{GUID}",
        "clientId": "{GUID}",
        "contentType": "Audit.SharePoint",
        "contentId":
"492638008028$492638008028$f28ab78ad40140608012736e373933ebspo2015043022$4a8
1a7c326fc4aed89c62e6039ab833b$04",
        "contentUri": "https://manage.office.com/api/v1.0/f28ab78a-d401-
4060-8012-
736e373933eb/activity/feed/audit/492638008028$492638008028$f28ab78ad40140608
012736e373933ebspo2015043022$4a81a7c326fc4aed89c62e6039ab833b$04",
        "contentCreated": "2015-05-23T17:35:00.000Z",
        "contentExpiration": "2015-05-30T17:35:00.000Z"
    },
    ...
]
```

Notification failure and retry

The notification system sends notifications as new content becomes available. If we encounter excessive failures when sending notifications, our retry mechanism will exponentially increase the time between retries. If we continue to encounter failures, we reserve the right to disable the webhook and stop sending notifications to it altogether. The /start operation can be used to re-enable a disabled webhook.

Retrieve content

To retrieve a content blob, make a GET request against the corresponding content URI that is included in the list of available content and in the notifications sent to a webhook. The returned content will be a collection of one more actions or events in JSON format.

Sample request

JSON

```
GET https://manage.office.com/api/v1.0/41463f53-8812-40f4-890f-
865bf6e35190/activity/feed/audit/301299007231$301299007231$41463f53881240f48
90f865bf6e35190aad2015062920$e1c2ab19858a469fb1f1fd097effffc9$04 HTTP/1.1
Authorization: Bearer eyJ0e...Qa6wg
```

Sample response

JSON


```
HTTP/1.1 200 OK
Content-Type: application/json; charset=utf-8

[
    {
        "CreationTime": "2015-06-29T20:03:19",
        "Id": "80c76bd2-9d81-4c57-a97a-accfc3443dca",
        "Operation": "PasswordLogonInitialAuthUsingPassword",
        "OrganizationId": "41463f53-8812-40f4-890f-865bf6e35190",
        "RecordType": 9,
        "ResultStatus": "failed",
        "UserKey": "1153977025279851686@contoso.onmicrosoft.com",
        "UserType": 0,
        "Workload": "AzureActiveDirectory",
        "ClientIP": "134.170.188.221",
        "ObjectId": "admin@contoso.onmicrosoft.com",
        "UserId": "admin@contoso.onmicrosoft.com",
        "AzureActiveDirectoryEventType": 0,
        "ExtendedProperties": [
            {
                "Name": "LoginError",
                "Value": "-2147217390;PP_E_BAD_PASSWORD;The entered and
stored passwords do not match."
            }
        ],
        "Client": "Exchange",
        "LoginStatus": -2147217390,
        "UserDomain": "contoso.onmicrosoft.com"
    },
    {
        "CreationTime": "2015-06-29T20:03:34",
        "Id": "4e655d3f-35fa-42e0-b050-264b2d255c7a",
        "Operation": "PasswordLogonInitialAuthUsingPassword",
        "OrganizationId": "41463f53-8812-40f4-890f-865bf6e35190",
        "RecordType": 9,
        "ResultStatus": "success",
        "UserKey": "1153977025279851686@contoso.onmicrosoft.com",
        "UserType": 0,
        "Workload": "AzureActiveDirectory",
        "ClientIP": "134.170.188.221",
        "ObjectId": "admin@contoso.onmicrosoft.com",
        "UserId": "admin@contoso.onmicrosoft.com",
```

```
        "AzureActiveDirectoryEventType": 0,
        "Client": "Exchange",
        "LoginStatus": 0,
        "UserDomain": "contoso.onmicrosoft.com"
    },
    {
        "CreationTime": "2015-06-29T20:04:55",
        "Id": "b567caf0-088e-4c1c-a4ea-633a1e3d66c8",
        "Operation": "Add User.",
        "OrganizationId": "41463f53-8812-40f4-890f-865bf6e35190",
        "RecordType": 8,
        "ResultStatus": "success",
        "UserKey": "1003BFFD8EC47CA6@contoso.onmicrosoft.com",
        "UserType": 0,
        "Workload": "AzureActiveDirectory",
        "ObjectId": "user001@contoso.onmicrosoft.com",
        "UserId": "admin@contoso.onmicrosoft.com",
        "AzureActiveDirectoryEventType": 0,
        "Actor": [
            {
                "ID": "1cef1fdb-ff52-48c4-8e4e-dfb5ea83d357",
                "Type": 2
            },
            {
                "ID": "admin@contoso.onmicrosoft.com",
                "Type": 5
            },
            {
                "ID": "1003BFFD8EC47CA6",
                "Type": 3
            }
        ],
        "ActorContextId": "41463f53-8812-40f4-890f-865bf6e35190",
        "InterSystemsId": "c2ced078-ad57-4079-a743-5c37f5284790",
        "IntraSystemId": "d1497f7e-15b4-49aa-83ad-11a17ca4a2f4",
        "Target": [
            {
                "ID": "user001@contoso.onmicrosoft.com",
                "Type": 5
            },
            {
                "ID": "10037FFE91510806",
                "Type": 3
            }
        ],
        "TargetContextId": "41463f53-8812-40f4-890f-865bf6e35190"
    }
]
```


List notifications

This operation lists all notification attempts for the specified content type. If you did not include a webhook when starting the subscription to the content type, there will be no notifications to retrieve. Because we retry notifications in the event of failure, this operation can return multiple notifications for the same content, and the order in which the notifications are sent will not necessarily match the order in which the content became available (especially when there are failures and retries).

You can use this operation to help investigate issues related to webhooks and notifications, but you should not use it to determine what content is currently available for retrieval. Use the /content operation instead. We return an error if the subscription status is disabled.

Expand table

Subscription	**Description**
Path	`/subscriptions/notifications?contentType={ContentType}&startTime={0}&endTime={1}`
Parameters	contentType - Must be a valid content type.
PublisherIdentifier	The tenant GUID of the vendor coding against the API. This is **not** the application GUID or the GUID of the customer using the application, but the GUID of the company writing the code. This parameter is used for throttling the request rate. Make sure this parameter is specified in all issued requests to get a dedicated quota. All requests received without this parameter will share the same quota.
startTime endTime	Optional datetimes (UTC) that indicate the time range of content to return, based on when the content became available. The time range is inclusive with respect to *startTime* (*startTime* <= contentCreated) and exclusive with respect to *endTime* (*contentCreated* < endTime), so that non-overlapping, incrementing time intervals can used to page through available content. • YYYY-MM-DD • YYYY-MM-DDTHH:MM • YYYY-MM-DDTHH:MM:SS Both must be specified (or both omitted) and they must be no more than 24 hours apart, with the start time no more than 7 days in the past. By default, if *startTime* and *endTime* are omitted, the content available in the last 24 hours is returned.
Response	JSON array - The notifications will be represented by JSON objects with the following properties: • **contentType**: indicates the content type. • **contentId**: an opaque string that uniquely identifies the content. • **contentUri**: the URL to use when retrieving the content.

Subscription	Description
	• **contentCreated**: the datetime when the content was made available. • **contentExpiration**: the datetime after which the content will no longer be available for retrieval. • **notificationSent**: the datetime when the notification was sent. • **notificationStatus**: indicates the success or failure of the notification attempt.

Sample request

JSON

```
GET {root}/subscriptions/notifications?
contentType=Audit.SharePoint&PublisherIdentifier=46b472a7-c68e-4adf-8ade-
3db49497518e
Authorization: Bearer eyJ0e...Qa6wg
```

Sample response

JSON

```
HTTP/1.1 200 OK
Content-Type: application/json; charset=utf-8

[
    {
        "contentType": "Audit.SharePoint",
        "contentId":
"492638008028$492638008028$f28ab78ad40140608012736e373933ebspo2015043022$4a8
1a7c326fc4aed89c62e6039ab833b$04",
        "contentUri": "https://manage.office.com/api/v1.0/f28ab78a-d401-
4060-8012-
736e373933eb/activity/feed/audit/492638008028$492638008028$f28ab78ad40140608
012736e373933ebspo2015043022$4a81a7c326fc4aed89c62e6039ab833b$04",
        "contentCreated": "2015-05-23T17:35:00.000Z",
        "contentExpiration": "2015-05-30T17:35:00.000Z",
        "notificationSent": "2015-05-23T17:36:00.000Z",
        "notificationStatus": "success"

    },
    ...
]
```

Pagination

When listing notification history for a time range, the number of results returned is limited to prevent response timeouts. If there are more results in the specified time range than can be returned in a single response, the results are truncated and a header is added to the response indicating the URL to use to retrieve the next page of results. The URL will contain the same *startTime* and *endTime* parameters that were specified in the original request, together with a parameter indicating the internal ID of the next page. If *startTime* and *endTime* were not specified in the original request, they will be set to reflect the 24-hour interval that preceded the original request.

JSON

```
HTTP/1.1 200 OK
Content-Type: application/json; charset=utf-8
NextPageUri:
https://manage.office.com/api/v1/{tenant_id}/activity/feed/subscriptions/con
tent?contentType=Audit.SharePoint&startTime=2015-10-01&endTime=2015-10-
02&nextPage=2015101900R022885001761
```

To list all available content for a specified time range, you might need to retrieve multiple pages until a response without the **NextPageUri** header is received.

Retrieve resource friendly names

This operation retrieves friendly names for objects in the data feed identified by guids. Currently "DlpSensitiveType" is the only supported object.

Expand table

Object	Subscription	Description
Path	`/resources/dlpSensitiveTypes`	
Parameters	PublisherIdentifier	The tenant GUID of the vendor coding against the API. This is **not** the application GUID or the GUID of the customer using the application, but the GUID of the company writing the code. This parameter is used for throttling the request rate. Make sure this parameter is specified in all issued requests to get a dedicated quota. All requests received without this parameter will share the same quota.

Object	Subscription	Description
Headers	Accept-Language	Header to specify the desired language for localized names. For example, use "en-US" for English or "es" for Spanish. The default language (en-US) will be returned if this header is not present.
Body	(empty)	
Response	JSON array	The available content will be represented by JSON objects with the following properties: • **id**: Indicates the guid of the sensitive information type. • **name**: The friendly name of the sensitive information type.

Sample request

JSON

```
GET {root}/resources/dlpSensitiveTypes?PublisherIdentifier=46b472a7-c68e-
4adf-8ade-3db49497518e
Authorization: Bearer eyJ0e...Qa6wg
Accept-Language: {language code}
```

Sample response

JSON

```
HTTP/1.1 200 OK

[
    {
        "id": "50842eb7-edc8-4019-85dd-5a5c1f2bb085",
        "name": "CreditCardNumber"
    },
    {
        "id": "0e9b3178-9678-47dd-a509-37222ca96b42",
        "name": "EUDebitCardNumber"
    },
    ...
    {
    }
]
```

API throttling

Organizations that access auditing logs through the Office 365 Management Activity API were restricted by throttling limits at the publisher level. This means that for a publisher pulling data on behalf of multiple customers, the limit was shared by all those customers.

We're moving from a publisher-level limit to a tenant-level limit. The result is that each organization will get their own fully allocated bandwidth quota to access their auditing data. All organizations are initially allocated a baseline of 2,000 requests per minute. This is not a static, predefined limit but is modeled on a combination of factors including the number of seats in the organization and that Office 365 and Microsoft 365 E5 organizations will get approximately twice as much bandwidth as non-E5 organizations. There will also be cap on the maximum bandwidth to protect the health of the service.

For more information, see the "High-bandwidth access to the Office 365 Management Activity API" section in Advanced audit in Microsoft 365.

> **Note**
>
> Even though each tenant can initially submit up to 2,000 requests per minute, Microsoft cannot guarantee a response rate. The response rate depends on various factors, such as client system performance, network capacity, and network speed.

Errors

When the service encounters an error, it will report the error response code to the caller, using standard HTTP error-code syntax. . Additional information is included in the body of the failed call as a single JSON object. An example of a full JSON error body is shown below:

JSON

```
{
    "error":{
        "code":"AF50000",
        "message": "An internal server error occurred. Retry the request."
    }
}
```

Expand table

Code	Message
AF10001	The permission set ({0}) sent in the request did not include the expected permission **ActivityFeed.Read**. {0} = the permission set in the access token.
AF20001	Missing parameter: {0}. {0} = the name of the missing parameter.
AF20002	Invalid parameter type: {0}. Expected type: {1} {0} = the name of the invalid parameter. {1} = the expected type (int, datetime, guid).
AF20003	Expiration {0} provided is set to past date and time. {0} = the expiration passed in the API call.
AF20010	The tenant ID passed in the URL ({0}) does not match the tenant ID passed in the access token ({1}). {0} = tenant ID passed in the URL{1} = tenant ID passed in the access token
AF20011	Specified tenant ID ({0}) does not exist in the system or has been deleted. {0} = tenant ID passed in the URL
AF20012	Specified tenant ID ({0}) is incorrectly configured in the system. {0} = tenant ID passed in the URL
AF20013	The tenant ID passed in the URL ({0}) is not a valid GUID. {0} = tenant ID passed in the URL
AF20020	The specified content type is not valid.
AF20021	The webhook endpoint {{0}) could not be validated. {1} {0} = webhook address.

Code	Message
	{1} = "The endpoint did not return HTTP 200." or "The address must begin with HTTPS."
AF20022	No subscription found for the specified content type.
AF20023	The subscription was disabled by {0}. {0} = "a tenant admin" or "a service admin"
AF20030	Start time and end time must both be specified (or both omitted) and must be less than or equal to 24 hours apart, with the start time no more than 7 days in the past.
AF20031	Invalid nextPage Input: {0}. {0} = the next page indicator passed in the URL
AF20050	The specified content ({0}) does not exist. {0} = resource id or resource URL
AF20051	Content requested with the key {0} has already expired. Content older than 7 days cannot be retrieved. {0} = resource id or resource URL
AF20052	Content ID {0} in the URL is invalid. {0} = resource id or resource URL
AF20053	Only one language may be present in the Accept-Language header.
AF20054	Invalid syntax in Accept-Language header.
AF429	Too many requests. Method={0}, PublisherId={1} {0} = HTTP Method {1} = Tenant GUID used as PublisherIdentifier
AF50000	An internal error occurred. Retry the request.

Office 365 Management Activity API schema

The Office 365 Management Activity API schema is provided as a data service in two layers:

- **Common schema**. The interface to access core Office 365 auditing concepts such as Record Type, Creation Time, User Type, and Action as well as to provide core dimensions (such as User ID), location specifics (such as Client IP address), and service-specific properties (such as Object ID). It establishes consistent and uniform views for users to extract all Office 365 audit data in a few top level views with the appropriate parameters, and provides a fixed schema for all the data sources, which significantly reduces the cost of learning. Common schema is sourced from product data that is owned by each product team, such as Exchange, SharePoint, Azure Active Directory, Viva Engage, and OneDrive. The Object ID field can be extended by Microsoft 365 product teams to add service-specific properties.
- **Service-specific schema**. Built on top of the Common schema to provide a set of Microsoft 365 service-specific attributes; for example, SharePoint schema, OneDrive schema, and Exchange admin schema.

Office 365 Management API schemas

This article provides details on the Common schema as well as service-specific schemas. The following table describes the available schemas.

Expand table

Name of schema	Description
Common schema	The view to extract Record Type, User ID, Client IP, User type and Action along with core dimensions such as user properties (such as UserID), location properties (such as Client IP), and service-specific properties (such as Object Id).
Copilot schema	Events include how and when to interact with Copilot, in which Microsoft 365 service the activity took place, and references to the files stored in Microsoft 365 that were accessed during the interaction.
SharePoint Base schema	Extends the Common schema with the properties specific to all SharePoint audit data.
SharePoint File Operations	Extends the SharePoint Base schema with the properties specific to file access and manipulation in SharePoint.
SharePoint List Operations	Extends the SharePoint Base schema with the properties specific to interactions with lists and list items in SharePoint.

Name of schema	Description
SharePoint Sharing schema	Extends the SharePoint Base schema with the properties specific to file sharing.
SharePoint schema	Extends the SharePoint Base schema with the properties specific to SharePoint, but unrelated to file access and manipulation.
Project schema	Extends the SharePoint Base schema with the properties specific to Project.
Exchange Admin schema	Extends the Common schema with the properties specific to all Exchange admin audit data.
Exchange Mailbox schema	Extends the Common schema with the properties specific to all Exchange mailbox audit data.
OWA Auth schema	Extends the Common schema with the properties specific to OWA Auth data.
Microsoft Entra ID Base schema	Extends the Common schema with the properties specific to all Microsoft Entra audit data.
Microsoft Entra account Logon schema	Extends the Microsoft Entra ID Base schema with the properties specific to all Microsoft Entra logon events.
Microsoft Entra ID Secure STS Logon schema	Extends the Microsoft Entra ID Base schema with the properties specific to all Microsoft Entra ID Secure Token Service (STS) logon events.
Microsoft Entra schema	Extends the Common schema with the properties specific to all Microsoft Entra audit data.
DLP schema	Extends the Common schema with the properties specific to Data Loss Prevention events.
Security and Compliance Center schema	Extends the Common schema with the properties specific to all Security & Compliance Center events.
Security and Compliance Alerts schema	Extends the Common schema with the properties specific to all Office 365 security and compliance alerts.
Viva Engage schema	Extends the Common schema with the properties specific to all Viva Engage events.
Data Center Security Base schema	Extends the Common schema with the properties specific to all data center security audit data.
Data Center Security Cmdlet schema	Extends the Data Center Security Base schema with the properties specific to all data center security cmdlet audit data.
Microsoft Teams schema	Extends the Common schema with the properties specific to all Microsoft Teams events.
Microsoft Defender for Office 365 and Threat Investigation and Response schema	Extends the Common schema with the properties specific to Defender for Office 365 and threat investigation and response data.

Name of schema	Description
Submission schema	Extends the Common schema with the properties specific to user and admin submissions in Microsoft Defender for Office 365.
Automated investigation and response events in Microsoft Defender for Office 365 Plan 2	Extends the Common schema with the properties specific to Office 365 automated investigation and response (AIR) events. To see an example, see Tech Community blog: Improve the Effectiveness of your SOC with Microsoft Defender for Office 365 and the Office 365 Management API .
Hygiene events schema	Extends the Common schema with the properties specific to events in the built-in security features for all cloud mailboxes and Microsoft Defender for Office 365.
Power BI schema	Extends the Common schema with the properties specific to all Power BI events.
Dynamics 365 schema	Extends the Common schema with the properties specific to Dynamics 365 events.
Viva Insights schema	Extends the Common schema with the properties specific to all Microsoft Viva Insights events.
Quarantine schema	Extends the Common schema with the properties specific to all quarantine events.
Microsoft Forms schema	Extends the Common schema with the properties specific to all Microsoft Forms events.
MIP label schema	Extends the Common schema with the properties specific to sensitivity labels manually or automatically applied to email messages.
Encrypted message portal event schema	Extends the Common schema with the properties specific to encrypted message portal accessed by external recipients.
Communication compliance Exchange schema	Extends the Common schema with the properties specific to the Communication compliance offensive language model.
Reports schema	Extends the Common schema with the properties specific to all reports events.
Compliance connector schema	Extends the Common schema with the properties specific to importing non-Microsoft data by using data connectors.
SystemSync schema	Extends the Common schema with the properties specific to data ingested via SystemSync.
Viva Goals schema	Extends the Common schema with the properties specific to all Viva Goals events.
Microsoft Planner schema	Extends the Common schema with the properties specific to Microsoft Planner events.
Microsoft Project for the web schema	Extends the Common schema with the properties specific to Microsoft Project for the web events.

Name of schema	Description
Viva Pulse schema	Extends the Common schema with the properties specific to all Viva Pulse events.
Compliance Manager schema	Extends the Common schema with the properties specific to Compliance Manager events.
Backup Policy Schema	Extends the Common schema with the properties specific to Microsoft 365 Backup Policies.
Restore Task schema	Extends the Common schema with the properties specific to Microsoft 365 Backup Restore Tasks.
Backup Item schema	Extends the Common schema with the properties specific to Microsoft 365 Backup artifacts.
Restore Item schema	Extends the common schema with the properties specific to Microsoft 365 Backup Restore Items.
Granular Browse Task schema	Extends the Common schema with the properties specific to Microsoft 365 Backup Granular Browse Tasks.
Cloud Policy service schema	Extends the Common schema with the properties specific to all Cloud Policy service audit data.
Cloud Update profile configuration schema	Extends the Common schema with the properties specific to the Cloud Update profile configuration audit data.
Cloud Update tenant configuration schema	Extends the Common schema with the properties specific to the Cloud Update tenant configuration audit data.
Cloud Update device configuration schema	Extends the Common schema with the properties specific to the Cloud Update device configuration audit data.
Microsoft Entra Risk Detection schema	Extends the Common schema with the properties specific to Microsoft Entra Risk Detection events.
Microsoft Edge WebContentFiltering schema	Extends the Common schema with the properties specific to Microsoft Edge WebContentFiltering events.
Microsoft 365 Copilot scheduled prompt schema	Extends the Common schema with the properties specific to Microsoft 365 Copilot scheduled prompt audit data.
Microsoft Places Directory schema	Extends the Common schema with the properties specific to Microsoft Places Directory audit data.
Teams Evaluation Data Hub schema	Extends the Common schema with the properties specific to Teams Evaluation Data Hub data-access and permission-change events.
Security and Compliance RBAC	Extends the Common schema with the properties specific to Security and Compliance RBAC permission events.
Dragon Copilot Admin schema	Extends the Common schema with the properties specific to Dragon Copilot administrative audit data.
Microsoft Agent 365 schema	Events generated by Microsoft Agent 365, including AI guardrail

Name of schema	Description
	evaluations, decisions, findings, and policy enforcement actions.

Common schema

EntityType Name: AuditRecord

Expand table

Parameter	Type	Mandatory?	Description
Id	Combination GUIDEdm.Guid	Yes	Unique identifier of an audit record.
RecordType	Self.AuditLogRecordType	Yes	The type of operation indicated by the record. See the AuditLogRecordType table for details on the types of audit log records.
CreationTime	Edm.Date	Yes	The date and time in Coordinated Universal Time (UTC) when the audit log record was generated.
Operation	Edm.String	Yes	The name of the user or admin activity. For a description of the most common operations/activities, see Search the audit log. For Exchange admin activity, this property identifies the name of the cmdlet that was run. For Dlp events, this can be "DlpRuleMatch", "DlpRuleUndo" or "DlpInfo", which are described under "DLP schema" below.
OrganizationId	Edm.Guid	Yes	The GUID for your organization's Office 365 tenant. This value is always the same for your organization, regardless of the Office 365 service in which it occurs.
UserType	Self.UserType	Yes	The type of user that performed the operation. See the UserType table for details on the types of users.
UserKey	Edm.String	Yes	An alternative ID for the user identified in the UserId property. This property is populated with the passport unique ID (PUID) for events performed by users in SharePoint, OneDrive, and Exchange.

Parameter	Type	Mandatory?	Description
Workload	Edm.String	Yes	The Office 365 service where the activity occurred.
ResultStatus	Edm.String	No	Indicates whether the action (specified in the Operation property) was successful or not. Possible values are **Succeeded**, **PartiallySucceeded**, or **Failed**. For Exchange admin activity, the value is either **True** or **False**. **Important**: Different workloads might overwrite the value of the ResultStatus property. For example, for Microsoft Entra ID STS logon events, a value of **Succeeded** for ResultStatus indicates only that the HTTP operation was successful; it doesn't mean the logon was successful. To determine if the actual logon was successful or not, see the LogonError property in the Microsoft Entra ID STS Logon schema. If the logon failed, the value of this property contains the reason for the failed logon attempt.
ObjectId	Edm.string	No	For SharePoint and OneDrive activity, the full path name of the file or folder accessed by the user. For Exchange admin audit logging, the name of the object that was modified by the cmdlet. For Cloud Policy service, the object ID of the policy configuration. For TrainableClassifier activity, the id of the model created, published, updated, or deleted by the user.
UserId	Edm.string	Yes	The UPN (User Principal Name) of the user who performed the action (specified in the Operation property) that resulted in the record being logged; for example, `my_name@my_domain_name`. **Note**: Records for activity performed by system accounts (such as SHAREPOINT\system or NT AUTHORITY\SYSTEM) are also included. In SharePoint, another value display in the UserId property is `app@sharepoint`. This value indicates the "user" who performed

Parameter	Type	Mandatory?	Description
			the activity was an application that has the necessary permissions in SharePoint to perform organization-wide actions (such as search a SharePoint site or OneDrive account) on behalf of a user, admin, or service.
ClientIP	Edm.String	Yes	The IPv4 or IPv6 address of the device that was used when the activity was logged. For some services, the value displayed in this property might be the IP address for a trusted application (for example, Office on the web apps) calling into the service on behalf of a user and not the IP address of the device used by person who performed the activity. Also, for Microsoft Entra ID-related events, the IP address isn't logged and the value for the ClientIP property is `null`.
Scope	Self.AuditLogScope	No	Event created by one of the following sources: • **online**: A service in Microsoft 365. • **onprem**: A service in an on-premises organization. Currently, SharePoint is the only workload sending events from on-premises to Microsoft 365.
AppAccessContext	CollectionSelf.AppAccessContext	No	The application context for the user or service principal that performed the action.

AgentAdminActivity

Expand table

Parameters	Type	Mandatory?	Description
AgentID	Edm.String	Yes	Unique identifier of the Agent
AgentName	Edm.String	Yes	The name of the Agent

Parameters	Type	Mandatory?	Description
AgentType	Edm.Int32	Yes	The type of the Agent - Microsoft (1P), External (3P), Shared by users (Shared), Built by your org (External).
UserAssignments	Collection (Edm.Guid)	No	Collection of GUID of Users and Groups associated with the agent activity.
ForAllUsers	Edm.Boolean	Yes	Whether the activity associated with the agent was captured for the entire organization or not. True when it applies to all users in the organization. False when it applies to specific users/group.

AgentSettingsAdminActivity

Expand table

Parameters	Type	Mandatory?	Description
Property	Edm.String	Yes	The tenant wide setting name that was updated by the admin.
RemovedIdentities	Collection(Edm.Guid)	No	Collection of GUID of Users or Groups that were removed.
AddedIdentities	Collection(Edm.Guid)	No	Collection of GUID of Users or Groups that were added.
NewValue	Edm.String	Yes	The new value of the setting.
ForAllUsers	Edm.Boolean	Yes	Boolean value indicating if the tenant wide setting impacted all users.
OldValue	Edm.String	Yes	The old value or the new setting.

Enum: AuditLogRecordType - Type: Edm.Int32

AuditLogRecordType

Expand table

Value	Member name	Description
1	ExchangeAdmin	Events from the Exchange admin audit log.
2	ExchangeItem	Events from an Exchange mailbox audit log for actions that are performed on a single item, such as creating or receiving an email message.
3	ExchangeItemGroup	Events from an Exchange mailbox audit log for actions that can be performed on multiple items, such as moving or deleted one or more email messages.

Value	Member name	Description
4	SharePoint	SharePoint events.
6	SharePointFileOperation	SharePoint file operation events.
7	OneDrive	OneDrive events.
8	AzureActiveDirectory	Microsoft Entra ID events.
9	AzureActiveDirectoryAccountLogon	Microsoft Entra ID OrgId logon events (deprecated).
10	DataCenterSecurityCmdlet	Data Center security cmdlet events.
11	ComplianceDLPSharePoint	Data loss protection (DLP) events in SharePoint and OneDrive.
13	ComplianceDLPExchange	Data loss protection (DLP) events in Exchange, when configured via Unified DLP Policy. DLP events based on Exchange Transport Rules are not supported.
14	SharePointSharingOperation	SharePoint sharing events.
15	AzureActiveDirectoryStsLogon	Secure Token Service (STS) logon events in Microsoft Entra ID.
16	SkypeForBusinessPSTNUsage	Public Switched Telephone Network (PSTN) events from Skype for Business.
17	SkypeForBusinessUsersBlocked	Blocked user events from Skype for Business.
18	SecurityComplianceCenterEOPCmdlet	Admin actions in the built-in security features for all cloud mailboxes from the Microsoft Defender portal.
19	ExchangeAggregatedOperation	Aggregated Exchange mailbox auditing events.
20	PowerBIAudit	Power BI events.
21	CRM	Dynamics 365 events.
22	Viva Engage	Viva Engage events.
23	SkypeForBusinessCmdlets	Skype for Business events.
24	Discovery	Events for eDiscovery activities performed by running content searches and managing eDiscovery cases in the Microsoft Purview portal.
25	MicrosoftTeams	Events from Microsoft Teams.
28	ThreatIntelligence	Phishing and malware events from the built-in security features for all cloud mailboxes and Microsoft Defender for Office 365.
29	MailSubmission	Submission events from the built-in security features for all cloud mailboxes and Microsoft Defender for Office 365.

Value	Member name	Description
30	MicrosoftFlow	Microsoft Power Automate (formerly called Microsoft Flow) events.
31	AeD	Advanced eDiscovery events.
32	MicrosoftStream	Microsoft Stream events.
33	ComplianceDLPSharePointClassification	Events related to DLP classification in SharePoint.
34	ThreatFinder	Campaign-related events from Microsoft Defender for Office 365.
35	Project	Microsoft Project events.
36	SharePointListOperation	SharePoint List events.
37	SharePointCommentOperation	SharePoint comment events.
38	DataGovernance	Events related to retention policies and retention labels in the Microsoft Purview portal.
39	Kaizala	Kaizala events.
40	SecurityComplianceAlerts	Security and compliance alert signals.
41	ThreatIntelligenceUrl	Safe links time-of-block and block override events from Microsoft Defender for Office 365.
42	SecurityComplianceInsights	Events related to insights and reports in the Microsoft Defender portal.
43	MIPLabel	Events related to the detection in the Transport pipeline of email messages that have been tagged (manually or automatically) with sensitivity labels.
44	VivaInsights	Viva Insights events.
45	PowerAppsApp	Power Apps events.
46	PowerAppsPlan	Subscription plan events for Power Apps.
47	ThreatIntelligenceAtpContent	Phishing and malware events for files in SharePoint, OneDrive, and Microsoft Teams from Microsoft Defender for Office 365.
48	LabelContentExplorer	Events related to data classification content explorer.
49	TeamsHealthcare	Events related to the Patients app in Microsoft Teams for Healthcare.
50	ExchangeItemAggregated	Events related to the MailItemsAccessed mailbox auditing action.
51	HygieneEvent	Events related to outbound spam protection.
52	DataInsightsRestApiAudit	Data Insights REST API events.

Value	Member name	Description
53	InformationBarrierPolicyApplication	Events related to the application of information barrier policies.
54	SharePointListItemOperation	SharePoint list item events.
55	SharePointContentTypeOperation	SharePoint list content type events.
56	SharePointFieldOperation	SharePoint list field events.
57	MicrosoftTeamsAdmin	Teams admin events.
58	HRSignal	Events related to HR data signals that support the Insider risk management solution.
59	MicrosoftTeamsDevice	Teams device events.
60	MicrosoftTeamsAnalytics	Teams analytics events.
61	InformationWorkerProtection	Events related to compromised user alerts.
62	Campaign	Email campaign events from Microsoft Defender for Office 365.
63	DLPEndpoint	Endpoint DLP events.
64	AirInvestigation	Automated incident response (AIR) events.
65	Quarantine	Quarantine events.
66	MicrosoftForms	Microsoft Forms events.
67	ApplicationAudit	Application audit events.
68	ComplianceSupervisionExchange	Events tracked by the Communication compliance offensive language model.
69	CustomerKeyServiceEncryption	Events related to the customer key encryption service.
70	OfficeNative	Events related to sensitivity labels applied to Office documents.
71	MipAutoLabelSharePointItem	Auto-labeling events in SharePoint.
72	MipAutoLabelSharePointPolicyLocation	Auto-labeling policy events in SharePoint.
73	MicrosoftTeamsShifts	Teams Shifts events.
75	MipAutoLabelExchangeItem	Auto-labeling events in Exchange.
76	CortanaBriefing	Briefing email events.
78	WDATPAlerts	Events related to alerts generated by Windows Defender for Endpoint.
79	PowerAppsResource	Events related to Microsoft Power Platform

Value	Member name	Description
		Connectors (Preview).
82	SensitivityLabelPolicyMatch	Events generated when the file labeled with a sensitivity label is opened or renamed.
83	SensitivityLabelAction	Event generated when sensitivity labels are applied, updated, or removed from a file.
84	SensitivityLabeledFileAction	Events generated when a file labeled with a sensitivity label is opened or renamed.
85	AttackSim	Events related to user activities in Attack Simulation & Training in Microsoft Defender for Office 365.
86	AirManualInvestigation	Events related to manual investigations in Automated investigation and response (AIR).
87	SecurityComplianceRBAC	Security and compliance RBAC events.
88	UserTraining	Events related to user training in Attack Simulation & Training in Microsoft Defender for Office 365.
89	AirAdminActionInvestigation	Events related to admin actions in Automated investigation and response (AIR).
90	MSTIC	Threat intelligence events in Microsoft Defender for Office 365.
91	PhysicalBadgingSignal	Events related to physical badging signals that support the Insider risk management solution.
92	TeamsEasyApprovals	Events related to Teams Easy Approvals
93	AipDiscover	AIP scanner events
94	AipSensitivityLabelAction	AIP sensitivity label events
95	AipProtectionAction	AIP protection events
96	AipFileDeleted	AIP file deletion events
97	AipHeartBeat	AIP heartbeat events
98	MCASAlerts	Events corresponding to alerts triggered by Microsoft Cloud App Security.
99	OnPremisesFileShareScannerDlp	Events related to scanning for sensitive data on file shares.
100	OnPremisesSharePointScannerDlp	Events related to scanning for sensitive data in SharePoint.
101	ExchangeSearch	Events related to using Outlook on the web (OWA) to search for mailbox items.
102	SharePointSearch	Events related to searching an organization's

Value	Member name	Description
		SharePoint home site.
103	PrivacyInsights	Privacy insight events.
105	MyAnalyticsSettings	MyAnalytics events.
106	SecurityComplianceUserChange	Events related to modifying or deleting a user.
107	ComplianceDLPExchangeClassification	Exchange DLP classification events.
109	MipExactDataMatch	Exact Data Match (EDM) classification events.
113	MS365DCustomDetection	Events related to custom detection actions in Microsoft 365 Defender.
147	CoreReportingSettings	Reports settings events.
148	ComplianceConnector	Events related to importing non-Microsoft data using data connectors in the Microsoft Purview portal.
154	OMEPortal	Encrypted message portal event logs generated by external recipients.
157	MipLabelAnalyticsAuditRecord	MIP Label Analytics events.
164	ScorePlatformGenericAuditRecord	Generic Audit Record used for Data Connectors.
174	DataShareOperation	Events related to sharing of data ingested via SystemSync.
181	EduDataLakeDownloadOperation	Events related to the export of SystemSync ingested data from the lake.
183	MicrosoftGraphDataConnectOperation	Events related to extractions done by Microsoft Graph Data Connect.
186	PowerPagesSite	Activities related to Power Pages site.
187	PowerPlatformAdminDlp	Events related to Microsoft Power Platform DLP (Preview).
188	PlannerPlan	Microsoft Planner plan events.
189	PlannerCopyPlan	Microsoft Planner copy plan events.
190	PlannerTask	Microsoft Planner task events.
191	PlannerRoster	Microsoft Planner roster and roster membership events.
192	PlannerPlanList	Microsoft Planner plan list events.
193	PlannerTaskList	Microsoft Planner task list events.
194	PlannerTenantSettings	Microsoft Planner tenant settings events.
195	ProjectForThewebProject	Microsoft Project for the web project events.

Value	Member name	Description
196	ProjectForThewebTask	Microsoft Project for the web task events.
197	ProjectForThewebRoadmap	Microsoft Project for the web roadmap events.
198	ProjectForThewebRoadmapItem	Microsoft Project for the web roadmap item events.
199	ProjectForThewebProjectSettings	Microsoft Project for the web project tenant settings events.
200	ProjectForThewebRoadmapSettings	Microsoft Project for the web roadmap tenant settings events.
202	MicrosoftTodoAudit	Microsoft To Do Audit events.
206	MicrosoftTeamsSensitivityLabelAction	Microsoft Teams Sensitivity Label events.
216	Viva Goals	Viva Goals events.
217	MicrosoftGraphDataConnectConsent	Events for consent actions performed by tenant admins for Microsoft Graph Data Connect applications.
218	AttackSimAdmin	Events related to admin activities in Attack Simulation & Training in Microsoft Defender for Office 365.
230	TeamsUpdates	Teams Updates App Events.
231	PlannerRosterSensitivityLabel	Microsoft Planner roster sensitivity label events.
235	MicrosoftDefenderForIdentityAudit	Microsoft Defender for Identity audit events.
237	DefenderExpertsforXDRAdmin	Microsoft Defender Experts Administrator action events.
251	VfamCreatePolicy	Viva Access Management policy create events.
252	VfamUpdatePolicy	Viva Access Management policy update events.
253	VfamDeletePolicy	Viva Access Management policy delete events.
256	PowerPlatformAdministratorActivity	Power Platform Administrator activity events.
257	Windows365CustomerLockbox	Windows365 Customer Lockbox audit events.
261	CopilotInteraction	Copilot interaction events.
265	VivaLearning	User activities in Viva Learning.
266	VivaLearningAdmin	Admin activities in Viva Learning.
269	PeopleAdminSettings	Audit events for People Admin settings.
275	OWAAuth	Access Token for Resource issued successfully events.
277	SharePointESignature	SharePoint eSignature audit events.

Value	Member name	Description
278	Dynamics365BusinessCentral	Audit events for Dynamics 365 Business Central.
279	MeshWorlds	Audit events for Mesh.
280	VivaPulseResponse	Viva Pulse survey response events.
281	VivaPulseOrganizer	Viva Pulse survey organizer events.
282	VivaPulseAdmin	Viva Pulse admin events.
283	VivaPulseReport	Viva Pulse report related events.
284	AIAppInteraction	Audit events for user interactions with third-party (non-Microsoft and non-custom-built) AI applications.
285	ComplianceDLMExchange	ComplianceDLMExchange events.
286	ComplianceDLMSharePoint	ComplianceDLMSharePoint events.
287	ProjectForThewebAssignedToMeSettings	Microsoft Project for the web assigned to me tenant settings events.
288	CloudPolicyService	Events from the Cloud Policy service.
291	SensitiveInfoDiscovered	Events from Purview On-Demand classification for Endpoint devices.
292	InsiderRiskScopedUserInsights	Events related to scoped user insights in Insider Risk Management.
293	MicrosoftTeamsRetentionLabelAction	Audit events for retention labels in Microsoft Teams.
294	AadRiskDetection	Audit events for Risk Detection in Microsoft Entra (previously called Azure Active Directory).
295	AuditSearch	Audit events for an admin interacting with Search in Purview Audit.
296	AuditRetentionPolicy	Audit events for an admin interacting with Retention Policies in Purview Audit.
297	AuditConfig	Audit events for an admin interacting with configuration settings related to Purview Audit.
298	BackupPolicy	Events related to Microsoft 365 Backup Policies.
299	RestoreTask	Events related to Microsoft 365 Backup Restore Tasks.
300	RestoreItem	Events related to artifacts backed up with Microsoft 365 Backup.
301	BackupItem	Events related to items being restored using Microsoft 365 Backup.
302	URBACAssignment	Events related to Unified RBAC assignments.

Value	Member name	Description
303	URBACRole	Events related to Unified RBAC roles.
304	URBACEnableState	Events related to Unified RBAC state enablement.
306	PurviewInsiderRiskCases	Events related to Purview Insider Risk cases.
307	PurviewInsiderRiskAlerts	Events related to Purview Insider Risk alerts.
308	InsiderRiskScopedUsers	Events related to Purview Insider Risk scoped users.
310	CreateCopilotPlugin	Audit events for Copilot plugin creation.
311	UpdateCopilotPlugin	Audit events for updating a Copilot plugin.
312	DeleteCopilotPlugin	Audit events for deleting a Copilot plugin.
313	EnableCopilotPlugin	Audit events for enabling a Copilot plugin.
314	DisableCopilotPlugin	Audit events for disabling a Copilot plugin.
315	CreateCopilotWorkspace	Audit events for creating a Copilot workspace.
316	UpdateCopilotWorkspace	Audit events for updating a Copilot workspace.
317	DeleteCopilotWorkspace	Audit events for deleting a Copilot workspace.
318	EnableCopilotWorkspace	Audit events for enabling a Copilot workspace.
319	DisableCopilotWorkspace	Audit events for disabling a Copilot workspace.
320	CreateCopilotPromptBook	Audit events for creating a Copilot promptbook.
321	UpdateCopilotPromptBook	Audit events for updating a Copilot promptbook.
322	DeleteCopilotPromptBook	Audit events for deleting a Copilot promptbook.
323	EnableCopilotPromptBook	Audit events for enabling a Copilot promptbook.
324	DisableCopilotPromptBook	Audit events for disabling a Copilot promptbook.
325	UpdateCopilotSettings	Audit event for updating user or tenant settings related to Copilot.
328	ConnectedAIAppInteraction	Audit events related to user interactions with third-party (non-Microsoft developed) AI applications which are deployed within the organization's Microsoft tenant.
329	PrivaPrivacyConsentOperation	Audit events related to Consent in Microsoft Priva.
330	PrivaPrivacyAssessmentOperation	Audit events related to Assessments in Microsoft Priva.
331	DataCatalogAccessRequests	Audit events related to Data Catalog access requests.
332	ComplianceSettingsChange	Microsoft Purview Compliance settings change

Value	Member name	Description
		events.
333	DataSecurityInvestigation	Events from Data Security Investigations in Microsoft Purview.
334	TeamCopilotInteraction	Audit events for user interactions with Facilitator and activities performed by Facilitator in Microsoft Teams.
335	IRMActivityAuditTrail	Events from Purview Insider Risk Management.
336	SharePointContentSecurityPolicy	Events from Content Security Policy in SharePoint.
337	CloudUpdateProfileConfig	Events from the Cloud Update's profile configuration.
338	CloudUpdateTenantConfig	Events from the Cloud Update's tenant configuration.
339	CloudUpdateDeviceConfig	Events from Cloud Update's managed devices configuration.
341	DeviceDiscoverySettingsExclusion	Events related to Exclusions in Device Discovery settings.
342	DeviceDiscoverySettingsAuthenticatedScans	Events related to Authenticated Scans in Device Discovery settings.
344	DeviceDiscoverySettings	Events related to settings in Device Discovery.
345	USXWorkspaceOnboarding	Events related to workspace onboarding in Microsoft Defender USX.
346	VivaGlintAdvancedConfiguration	Events related to Advanced Configuration in Viva Glint.
347	VivaGlintPulseProgram	Events related to Pulse Program in Viva Glint.
348	VivaGlintPulseProgramRespondentRate	Events related to Pulse Program Respondent Rate in Viva Glint.
349	VivaGlintQuestion	Events related to Questions in Viva Glint.
350	VivaGlintRole	Events related to Roles in Viva Glint.
351	VivaGlintRubicon	Events related to Rubicon in Viva Glint.
352	VivaGlintSupportAccess	Events related to Support Access in Viva Glint.
353	VivaGlintSystem	Events related to system activities in Viva Glint.
354	VivaGlintUser	Events related to user activities in Viva Glint.
355	VivaGlintUserGroup	Events related to user group activities in Viva Glint.
356	VivaGlintFeedbackProgram	Events related to Feedback Programs in Viva Glint.
357	FabricAudit	Events related to Microsoft Fabric.
358	TrainableClassifier	Events from Purview Data Classification.

Value	Member name	Description
359	WebContentFiltering	Events from Microsoft Edge WebContentFiltering.
360	NoisyAlertPolicy	Events related to Noisy Alert Policies in Microsoft Defender.
361	DataScanClassification	Events from Purview On-Demand classification for SharePoint and OneDrive.
362	AIInteractionsExport	Events related to export of AI interactions.
363	Microsoft365CopilotScheduledPrompt	Events from Microsoft 365 Copilot scheduled prompt.
364	PlacesDirectory	Events from Microsoft Places Directory.
365	SentinelNotebookOnLake	Events from notebook execution on Sentinel data lake.
366	SentinelJob	Events from operations on jobs in Sentinel data lake.
367	SentinelKQLOnLake	Events from running KQL on Sentinel data lake.
368	SentinelLakeOnboarding	Events from onboarding to Sentinel data lake.
369	SentinelLakeDataOnboarding	Data loading events into Sentinel data lake.
370	SentinelAITool	Events from operations on AI tool in Microsoft Sentinel
371	SentinelGraph	Events related to Graph in Microsoft Sentinel.
372	CrossTenantAccessPolicy	Events from Cross Tenant Access Policies.
373	OutlookCopilotAutomation	Events related to back-end automation (without an explicit user interaction) in Microsoft Outlook driven by Agents, Copilot, or other AI scenarios.
374	VivaEngageNetworkAssociation	Events related to Network Association in Viva Engage.
375	AppAdminActivity	Events related to app admin activity.
376	AppSettingsAdminActivity	Events related to app settings admin activity.
377	UniversalPrintPrintJob	Audit events related to Print Jobs in Microsoft Universal Print.
378	VivaAmplifyOutlookSensitivityLabel	Events related to Outlook Sensitivity Labels in Viva Amplify.
379	AIInteractionsSubscription	Events related to AI interaction subscriptions.
380	AIInteractionsChangeNotification	Events related to AI interaction change notifications.
381	FilteringMailMetadataExtended	Events related to filtering mail metadata.

Value	Member name	Description
382	OfficeRestrictedModeAction	Audit events related to activities performed in Office Restricted Mode.
383	CopilotForSecurityTrigger	Events related to triggers for Security Copilot agents.
384	CopilotAgentManagement	Events related to admin activities for Microsoft Copilot agents.
385	P4AIAssessmentFabricScannerRecord	Events related to Purview for AI Assessment Fabric Scanner.
386	PlannerGoal	Microsoft Planner goal events.
387	PlannerGoalList	Microsoft Planner goal list events.
401	PlannerChatMessage	Microsoft Planner chat message events.
402	PlannerChatMessageList	Microsoft Planner chat message list events.
414	VivaEngageSegment	Viva Engage segmentation events.
422	VivaEngageEvents	Events related to Viva Engage hosted events.
427	UniversalPrintManagement	Audit events related to Management events in Microsoft Universal Print.
430	PurviewPostureAgent	Data Security Posture Agent events.
431	GranularBrowseTask	Events related to browsing backed up site's restore point using Microsoft 365 Backup.
444	TeamsEvalDataHubDataAccess	Teams Evaluation Data Hub data access events.
445	TeamsEvalDataHubPermissionChange	Teams Evaluation Data Hub permission change events.
454	DragonCopilotAdmin	Events from Dragon Copilot administrative operations.
462	MicrosoftTeamsUserConcern	Events related to user security concern in Microsoft Teams.
463	VivaGlintAgenticCampaign	Events related to agentic campaigns in Viva Glint.
479	VivaGlintVivaInsightsIntegration	Events related to the Viva Insights to Viva Glint integration.
480	PeopleSkillsAdmin	People Skills (Viva) events for tenant administrator and application (service-to-service) operations.
481	PeopleSkillsUser	People Skills (Viva) events for user and system operations on user skills.
489	AIGuardrail	Events generated when an AI guardrail evaluates content and applies a guardrail decision.

Enum: User Type - Type: Edm.Int32

User Type

Expand table

Value	Member name	Description
0	Regular	A regular user without admin permissions.
1	Reserved	A reserved value, not for use.
2	Admin	An administrator in your Microsoft 365 organization.
3	DCAdmin	A Microsoft datacenter administrator or datacenter system account.
4	System	An audit event triggered by server-side logic. For example, Windows services or background processes.
5	Application	An audit event triggered by a Microsoft Entra application.
6	ServicePrincipal	A service principal.
7	CustomPolicy	A customer created or managed policy.
8	SystemPolicy	A Microsoft-managed or system policy.
9	PartnerTechnician	A partner tenant's user working on behalf of the customer tenant (in GDAPscenarios).
10	Guest	A guest or anonymous user.

> **Note**
>
> ** For Microsoft Entra related events, the value for an administrator isn't used in an audit record. Audit records for activities performed by administrators indicate that a regular user (for example, **UserType: 0**) performed the activity. The **UserID** property identifies the person (regular user or administrator) who performed this activity.

Enum: AuditLogScope - Type: Edm.Int32

AuditLogScope

Expand table

Value	Member name	Description
0	Online	This event was created by a Microsoft 365 service.
1	Onprem	This event was created by an on-premises server.

Complex Type AppAccessContext

Expand table

Parameters	Type	Mandatory?	Description
AADSessionId	Edm.String	No	The Microsoft Entra SessionId of the Entra sign-in that was performed by the app on behalf of the user.
APIId	Edm.String	No	The Id for the API pathway that is used to access the resource; for example access via the Microsoft Graph API.
ClientAppId	Edm.String	No	The Id of the Microsoft Entra app that performed the access on behalf of the user.
ClientAppName	Edm.String	No	The name of the Microsoft Entra app that performed the access on behalf of the user.
CorrelationId	Edm.String	No	An identifier that can be used to correlate a specific user's actions across Microsoft 365 services.
UniqueTokenId	Edm.String	No	UniqueTokenId gets set if the Microsoft Entra token is available for the request. It's a unique, per-token identifier that is case-sensitive.
IssuedAtTime	Edm.Date	No	"Issued At" gets set if the Microsoft Entra token is available for the request and it indicates when the authentication for this Microsoft Entra token occurred.

SharePoint Base schema

Expand table

Parameter	Type	Mandatory?	Descripti
Site	Edm.Guid	No	The GUID the site where the or folder accessed the user is located.
ItemType	Edm.String String="Microsoft.Office.Audit.Schema.SharePoint.ItemType"	No	The type object tha was acces or modifie See the ItemType table for details on types of objects.

Parameter	Type	Mandatory?	Descripti
EventSource	Edm.String String="Microsoft.Office.Audit.Schema.SharePoint.EventSource"	No	Identifies an event occurred i SharePoin Possible values are **SharePoir ObjectMc**
SourceName	Edm.String	No	The entity that trigg the audite operation Possible values are SharePoin **ObjectMc**
UserAgent	Edm.String	No	Informatic about the user's clie or browse This informatic provided the client browser.
MachineDomainInfo	Edm.String Term="Microsoft.Office.Audit.Schema.PIIFlag" Bool="true"	No	Informatic about dev sync operation This informatic reported if it's pres in the request.
MachineId	Edm.String Term="Microsoft.Office.Audit.Schema.PIIFlag" Bool="true"	No	Informatic about dev sync operation This informatic reported if it's pres in the request.
ListItemUniqueId	Edm.Guid	No	The Guid uniquely a identifiabl

Parameter	Type	Mandatory?	Descripti
			item of lis This informatic present o if it is applicable
ListId	Edm.Guid	No	The Guid the list. Th informatic present o if it is applicable
ApplicationId	Edm.String	No	The ID of applicatio performin the operation
ApplicationDisplayName	Edm.String	No	The displa name of t applicatio performin the operation
IsWorkflow	Edm.Boolean	No	This is set `True` if SharePoin Workflow triggered audited event.

Enum: ItemType - Type: Edm.Int32

ItemType

Expand table

Value	Member name	Description
0	Invalid	The item is none of the other item types (that are listed in this table).
1	File	The item is a file.
5	Folder	The item is a folder.
6	web	The item is a web.
7	Site	The item is a site.

Value	Member name	Description
8	Tenant	The item is a tenant.
9	DocumentLibrary	The item is a document library.
11	Page	The item is a Page.

Enum: EventSource - Type: Edm.Int32

EventSource

Expand table

Value	Member name	Description
0	SharePoint	The event source is SharePoint.
1	ObjectModel	The event source is ObjectModel.

Enum: SharePointAuditOperation - Type: Edm.Int32

Expand table

Member name	Description
AccessInvitationAccepted	The recipient of an invitation to view or edit a shared file (or folder) has accessed the shared file by clicking on the link in the invitation.
AccessInvitationCreated	User sends an invitation to another person (inside or outside their organization) to view or edit a shared file or folder on a SharePoint or OneDrive site. The details of the event entry identifies the name of the file that was shared, the user the invitation was sent to, and the type of the sharing permission selected by the person who sent the invitation.
AccessInvitationExpired	An invitation sent to an external user expires. By default, an invitation sent to a user outside of your organization expires after 7 days if the invitation isn't accepted.
AccessInvitationRevoked	The site administrator or owner of a site or document in SharePoint or OneDrive withdraws an invitation that was sent to a user outside your organization. An invitation can be withdrawn only before it's accepted.
AccessInvitationUpdated	The user who created and sent an invitation to another person to view or edit a shared file or folder on a SharePoint or OneDrive site resent the invitation.
AccessRequestApproved	The site administrator or owner of a site or document in SharePoint or OneDrive approves a user request to access the site or document.

Member name	Description
AccessRequestCreated	User requests access to a site or document in SharePoint or OneDrive that they don't have permission to access.
AccessRequestRejected	The site administrator or owner of a site or document in SharePoint declines a user request to access the site or document.
ActivationEnabled	Users can browser-enable form templates that don't contain form code, require full trust, enable rendering on a mobile device, or use a data connection managed by a server administrator.
AdministratorAddedToTermStore	Term store administrator added.
AdministratorDeletedFromTermStore	Term store administrator deleted.
AllowGroupCreationSet	Site administrator or owner adds a permission level to a SharePoint or OneDrive site that allows a user assigned that permission to create a group for that site.
AppCatalogCreated	App catalog created to make custom business apps available for your SharePoint Environment.
AuditPolicyRemoved	Document LifeCycle Policy has been removed for a site collection.
AuditPolicyUpdate	Document LifeCycle Policy has been updated for a site collection.
AzureStreamingEnabledSet	A video portal owner has allowed video streaming from Azure.
CollaborationTypeModified	The type of collaboration allowed on sites (for example, intranet, extranet, or public) has been modified.
ConnectedSiteSettingModified	User has either created, modified or deleted the link between a project and a project site or the user modifies the synchronization setting on the link in Project web app.
ContainerTypeCreated	A SharePoint Embedded container type was created.
ContainerTypeDeleted	A SharePoint Embedded container type was deleted.
ContainerTypeOwnersUpdated	The owners of a SharePoint Embedded container type were updated.
ContainerTypeUpdated	A SharePoint Embedded container type was updated.
CreateSSOApplication	Target application created in Secure store service.
CustomFieldOrLookupTableCreated	User created a custom field or lookup table/item in Project web app.
CustomFieldOrLookupTableDeleted	User deleted a custom field or lookup table/item in Project web app.

Member name	Description
CustomFieldOrLookupTableModified	User modified a custom field or lookup table/item in Project web app.
CustomizeExemptUsers	Global administrator customized the list of exempt user agents in SharePoint admin center. You can specify which user agents to exempt from receiving an entire web page to index. This configuration means when an exempt user agent encounters an InfoPath form, the form is returned as an XML file instead of an entire web page. This result makes indexing InfoPath forms faster.
DefaultLanguageChangedInTermStore*	Language setting changed in the terminology store.
DelegateModified	User created or modified a security delegate in Project web app.
DelegateRemoved	User deleted a security delegate in Project web app.
DeleteSSOApplication	An SSO application was deleted.
eDiscoveryHoldApplied	An In-Place Hold was placed on a content source. In-Place Holds are managed by using an eDiscovery site collection (such as the eDiscovery Center) in SharePoint.
eDiscoveryHoldRemoved	An In-Place Hold was removed from a content source. In-Place Holds are managed by using an eDiscovery site collection (such as the eDiscovery Center) in SharePoint.
eDiscoverySearchPerformed	An eDiscovery search was performed using an eDiscovery site collection in SharePoint.
EngagementAccepted	User accepts a resource engagement in Project web app.
EngagementModified	User modifies a resource engagement in Project web app.
EngagementRejected	User rejects a resource engagement in Project web app.
EnterpriseCalendarModified	User copies, modifies or delete an enterprise calendar in Project web app.
EntityDeleted	User deletes a timesheet in Project web app.
EntityForceCheckedIn	User forces a check-in on a calendar, custom field or lookup table in Project web app.
ExemptUserAgentSet	Global administrator adds a user agent to the list of exempt user agents in the SharePoint admin center.
FeatureActivated	Site admin activates a site collection feature.
FeatureDeactivated	Site admin deactivates a site collection feature.
FileAccessed	User or system account accesses a file on a SharePoint or OneDrive site. System accounts can also generate FileAccessed events.

Member name	Description
FileCheckOutDiscarded	User discards a checked out file. That means any changes they made to the file when it was checked out are discarded, and not saved to the version of the document in the document library.
FileCheckedIn	User checks in a document that they checked out from a SharePoint or OneDrive document library.
FileCheckedOut	User checks out a document located in a SharePoint or OneDrive document library. Users can check out and make changes to documents that have been shared with them.
FileCopied	User copies a document from a SharePoint or OneDrive site. The copied file can be saved to another folder on the site.
FileDeleted	User deletes a document from a SharePoint or OneDrive site.
FileDeletedFirstStageRecycleBin	User deletes a file from the recycle bin on a SharePoint or OneDrive site.
FileDeletedSecondStageRecycleBin	User deletes a file from the second-stage recycle bin on a SharePoint or OneDrive site.
FileDownloaded	User downloads a document from a SharePoint or OneDrive site.
FileFetched	This event has been replaced by the FileAccessed event, and has been deprecated.
FileModified	User or system account modifies the content or the properties of a document located on a SharePoint or OneDrive site.
FileMoved	User moves a document from its current location on a SharePoint or OneDrive site to a new location.
FilePreviewed	User previews a document on a SharePoint or OneDrive site.
FileRecycled	User moves a document into the SharePoint or OneDrive Recycle Bin.
FileRenamed	User renames a document on a SharePoint or OneDrive site.
FileRestored	User restores a document from the recycle bin of a SharePoint or OneDrive site.
FileSyncDownloadedFull	User downloads a file to their computer from a SharePoint document library or OneDrive using OneDrive sync app (OneDrive.exe).
FileSyncDownloadedPartial	This event has been deprecated along with the old OneDrive sync app (Groove.exe).
FileSyncUploadedFull	User uploads a new file or changes to a file in SharePoint document library or OneDrive using OneDrive sync app (OneDrive.exe).

Member name	Description
FileSyncUploadedPartial	This event has been deprecated along with the old OneDrive sync app (Groove.exe).
FileUploaded	User uploads a document to a folder on a SharePoint or OneDrive site.
FileViewed	This event has been replaced by the FileAccessed event, and has been deprecated.
FolderCopied	User copies a folder from a SharePoint or OneDrive site to another location in SharePoint or OneDrive.
FolderCreated	User creates a folder on a SharePoint or OneDrive site.
FolderDeleted	User deletes a folder from a SharePoint or OneDrive site.
FolderDeletedFirstStageRecycleBin	User deletes a folder from the recycle bin on a SharePoint or OneDrive site .
FolderDeletedSecondStageRecycleBin	User deletes a folder from the second-stage recycle bin on a SharePoint or OneDrive site.
FolderModified	User modifies a folder on a SharePoint or OneDrive site. This event includes folder metadata changes, such as tags and properties.
FolderMoved	User moves a folder from a SharePoint or OneDrive site.
FolderRecycled	User moves a folder into the SharePoint or OneDrive Recycle Bin.
FolderRenamed	User renames a folder on a SharePoint or OneDrive site.
FolderRestored	User restores a folder from the Recycle Bin on a SharePoint or OneDrive site.
GroupAdded	Site administrator or owner creates a group for a SharePoint or OneDrive site, or performs a task that results in a group being created. For example, the first time a user creates a link to share a file, a system group is added to the user's OneDrive site. This event can also be a result of a user creating a link with edit permissions to a shared file.
GroupRemoved	User deletes a group from a SharePoint or OneDrive site.
GroupUpdated	Site administrator or owner changes the settings of a group for a SharePoint or OneDrive site. This can include changing the group's name, who can view or edit the group membership, and how membership requests are handled.
LanguageAddedToTermStore	Language added to the terminology store.
LanguageRemovedFromTermStore	Language removed from the terminology store.
LegacyWorkflowEnabledSet	Site administrator or owner adds the SharePoint Workflow Task content type to the site. Global administrators can also enable

Member name	Description
	work flows for the entire organization in the SharePoint admin center.
LookAndFeelModified	User modifies a quick launch, Gantt chart formats, or group formats. Or the user creates, modifies, or deletes a view in Project web app.
ManagedSyncClientAllowed	User successfully establishes a sync relationship with a SharePoint or OneDrive site. The sync relationship is successful because the user's computer is a member of a domain that's been added to the list of domains (called the safe recipients list) that can access document libraries in your organization. For more information, see Get started with SharePoint Online Management Shell to enable OneDrive sync for domains that are on the safe recipients list.
MaxQuotaModified	The maximum quota for a site has been modified.
MaxResourceUsageModified	The maximum allowable resource usage for a site has been modified.
MySitePublicEnabledSet	The flag enabling users to have public MySites has been set by the SharePoint administrator.
NewsFeedEnabledSet	Site administrator or owner enables RSS feeds for a SharePoint or OneDrive site. Global administrators can enable RSS feeds for the entire organization in the SharePoint admin center.
ODBNextUXSettings	New UI for OneDrive has been enabled.
OfficeOnDemandSet	Site administrator enables Office on Demand, which lets users access the latest version of Office desktop applications. Office on Demand is enabled in the SharePoint admin center and requires an Office 365 subscription that includes full, installed Office applications.
PageViewed	User views a page on a SharePoint site or OneDrive site. This does not include viewing document library files from a SharePoint site or One Drive for Business site on a browser.
PeopleResultsScopeSet	Site administrator creates or changes the result source for People Searches for a SharePoint site.
PermissionSyncSettingModified	User modifies the project permission sync settings in Project web app.
PermissionTemplateModified	User creates, modifies or deletes a permissions template in Project web app.
PortfolioDataAccessed	User accesses portfolio content (driver library, driver prioritization, portfolio analyses) in Project web app.
PortfolioDataModified	User creates, modifies, or deletes portfolio data (driver library, driver prioritization, portfolio analyses) in Project web app.

Member name	Description
PreviewModeEnabledSet	Site administrator enables document preview for a SharePoint site.
ProjectAccessed	User accesses project content in Project web app.
ProjectCheckedIn	User checks in a project that they checked out from a Project web app.
ProjectCheckedOut	User checks out a project located in a Project web app. Users can check out and make changes to projects that they have permission to open.
ProjectCreated	User creates a project in Project web app.
ProjectDeleted	User deletes a project in Project web app.
ProjectForceCheckedIn	User forces a check in on a project in Project web app.
ProjectModified	User modifies a project in Project web app.
ProjectPublished	User publishes a project in Project web app.
ProjectWorkflowRestarted	User restarts a workflow in Project web app.
PWASettingsAccessed	User access the Project web app settings via CSOM.
PWASettingsModified	User modifies the a Project web app configuration.
QueueJobStateModified	User cancels or restarts a queue job in Project web app.
QuotaWarningEnabledModified	Storage quota warning modified.
RenderingEnabled	Browser-enabled form templates are rendered by InfoPath forms services.
ReportingAccessed	User accessed the reporting endpoint in Project web app.
ReportingSettingModified	User modifies the reporting configuration in Project web app.
ResourceAccessed	User accesses an enterprise resource content in Project web app.
ResourceCheckedIn	User checks in an enterprise resource that they checked out from Project web app.
ResourceCheckedOut	User checks out an enterprise resource located in Project web app.
ResourceCreated	User creates an enterprise resource in Project web app.
ResourceDeleted	User deletes an enterprise resource in Project web app.
ResourceForceCheckedIn	User forces a check in of an enterprise resource in Project web app.
ResourceModified	User modifies an enterprise resource in Project web app.

Member name	Description
ResourcePlanCheckedInOrOut	User checks in or out a resource plan in Project web app.
ResourcePlanModified	User modifies a resource plan in Project web app.
ResourcePlanPublished	User publishes a resource plan in Project web app.
ResourceRedacted	User redacts an enterprise resource removing all personal information in Project web app.
ResourceWarningEnabledModified	Resource quota warning modified.
SSOGroupCredentialsSet	Group credentials set in Secure store service.
SSOUserCredentialsSet	User credentials set in Secure store service.
SearchCenterUrlSet	Search center URL set.
SecondaryMySiteOwnerSet	A user has added a secondary owner to their MySite.
SecurityCategoryModified	User creates, modifies or deletes a security category in Project web app.
SecurityGroupModified	User creates, modifies or deletes a security group in Project web app.
SendToConnectionAdded	Global administrator creates a new Send To connection on the Records management page in the SharePoint admin center. A Send To connection specifies settings for a document repository or a records center. When you create a Send To connection, a Content Organizer can submit documents to the specified location.
SendToConnectionRemoved	Global administrator deletes a Send To connection on the Records management page in the SharePoint admin center.
SharedLinkCreated	User creates a link to a shared file in SharePoint or OneDrive. This link can be sent to other people to give them access to the file. A user can create two types of links: a link that allows a user to view and edit the shared file, or a link that allows the user to just view the file.
SharedLinkDisabled	User disables (permanently) a link that was created to share a file.
SharingInvitationAccepted*	User accepts an invitation to share a file or folder. This event is logged when a user shares a file with other users.
SharingRevoked	User unshared a file or folder that was previously shared with other users. This event is logged when a user stops sharing a file with other users.
SharingSet	User shares a file or folder located in SharePoint or OneDrive with another user inside their organization.
SiteAdminChangeRequest	User requests to be added as a site collection administrator for a SharePoint site collection. Site collection administrators have

Member name	Description
	full control permissions for the site collection and all subsites.
SiteCollectionAdminAdded*	Site collection administrator or owner adds a person as a site collection administrator for a SharePoint or OneDrive site. Site collection administrators have full control permissions for the site collection and all subsites.
SiteCollectionCreated	Global administrator creates a new site collection in your SharePoint organization.
SitePermanentlyDeleted	A SharePoint or global administrator permanently deletes a site from SharePoint Admin Center Deleted Sites.
SiteRenamed	Site administrator or owner renames a SharePoint or OneDrive site
SiteRestored	A SharePoint or global administrator restores a site from SharePoint Admin Center Deleted Sites.
StatusReportModified	User creates, modifies or deletes a status report in Project web app.
SyncGetChanges	User clicks **Sync** in the action tray on in SharePoint or OneDrive to synchronize any changes to file in a document library to their computer.
SyntexBillingSubscriptionSettingsChanged	The Syntex Billing subscription settings have changed. This event is triggered when a Syntex trial expires.
TaskStatusAccessed	User accesses the status of one or more tasks in Project web app.
TaskStatusApproved	User approves a status update of one or more tasks in Project web app.
TaskStatusRejected	User rejects a status update of one or more tasks in Project web app.
TaskStatusSaved	User saves a status update of one or more tasks in Project web app.
TaskStatusSubmitted	User submits a status update of one or more tasks in Project web app.
TenantWideThemeCreated	A SharePoint administrator, global administrator, or brand manager creates a custom theme that applies across all SharePoint sites in your organization.
TenantWideThemeDeleted	A SharePoint administrator, global administrator, or brand manager deletes a custom theme that applies across all SharePoint sites in your organization.
TenantWideThemeUpdated	A SharePoint administrator, global administrator, or brand manager updates a custom theme that applies across all SharePoint sites in your organization.

Member name	Description
TimesheetAccessed	User accesses a timesheet in Project web app.
TimesheetApproved	User approves timesheet in Project web app.
TimesheetRejected	User rejects a timesheet in Project web app.
TimesheetSaved	User saves a timesheet in Project web app.
TimesheetSubmitted	User submits a status timesheet in Project web app.
UnmanagedSyncClientBlocked	User tries to establish a sync relationship with a SharePoint or OneDrive site from a computer that isn't a member of your organization's domain or is a member of a domain that hasn't been added to the list of domains (called the safe recipients list) that can access document libraries in your organization. The sync relationship is not allowed, and the user's computer is blocked from syncing, downloading, or uploading files on a document library.
UpdateDisableSiteBranding	A SharePoint or global administrator enables or disables site branding.
UpdateSSOApplication	Target application updated in Secure store service.
UserAddedToGroup	Site administrator or owner adds a person to a group on a SharePoint or OneDrive site. Adding a person to a group grants the user the permissions that were assigned to the group.
UserRemovedFromGroup	Site administrator or owner removes a person from a group on a SharePoint or OneDrive site. After the person is removed, they no longer are granted the permissions that were assigned to the group.
WebThemeApplied	A SharePoint or global administrator, or a site owner applies a web theme.
WorkflowModified	User creates, modifies, or deletes an Enterprise Project Type or Workflow phases or stages in Project web app.

SharePoint file operations

File and page events returned in audit log searches use this schema.

Expand table

Parameter	Type	Mandatory?	Description
SiteUrl	Edm.String	Yes	The URL of the site where the file or folder accessed by the user is located.
SourceRelativeUrl	Edm.String	No	The URL of the folder that contains the file accessed by the user. The combination of the values for the *SiteURL*, *SourceRelativeURL*, and

Parameter	Type	Mandatory?	Description
			SourceFileName parameters is the same as the value for the **ObjectID** property, which is the full path name for the file accessed by the user.
SourceFileName	Edm.String	Yes	The name of the file or folder accessed by the user.
SourceFileExtension	Edm.String	No	The file extension of the file that was accessed by the user. This property is blank if the object that was accessed is a folder.
DestinationRelativeUrl	Edm.String	No	The URL of the destination folder where a file is copied or moved. The combination of the values for *SiteURL*, *DestinationRelativeURL*, and *DestinationFileName* parameters is the same as the value for the **ObjectID** property, which is the full path name for the file that was copied. This property is displayed only for FileCopied and FileMoved events.
DestinationFileName	Edm.String	No	The name of the file that is copied or moved. This property is displayed only for FileCopied and FileMoved events.
DestinationFileExtension	Edm.String	No	The file extension of a file that is copied or moved. This property is displayed only for FileCopied and FileMoved events.
UserSharedWith	Edm.String	No	The user that a resource was shared with.
SharingType	Edm.String	No	The type of sharing permissions that were assigned to the user that the resource was shared with. This user is identified by the *UserSharedWith* parameter.
SourceLabel	Edm.String	No	The original label of the file before it's changed by a user action.
DestinationLabel	Edm.String	No	The final label of the file after it's changed by a user action.
SensitivityLabelOwnerEmail	Edm.String	No	The email address of the owner of the sensitivity label.
SensitivityLabelId	Edm.String	No	The current sensitivity label ID of the file.

SharePoint list operations

SharePoint list events returned in audit log searches use this schema.

Expand table

Parameter	Type	Mandatory?	Description
ListTitle	Edm.String	No	The title of the SharePoint list.
ListName	Edm.String	No	The name of the SharePoint list.
ListUrl	Edm.String	No	The URL of the list relative to the containing website.
ListBaseType	Edm.String	No	Specifies the base type for a list.
ListBaseTemplateType	Edm.String	No	The list definition type on which the list is based.
IsHiddenList	Edm.Boolean	No	This value is set to `True` if the SharePoint list is hidden.
IsDocLib	Edm.Boolean	No	This value is set to `True` if the SharePoint list is of the type Document Library.

SharePoint Sharing schema

Sharing and access request events returned in audit log searches use this schema.

The file share-related SharePoint events. They are different from file- and folder-related events in that a user is taking an action that has some effect on another user. For information about the SharePoint Sharing schema, see Use sharing auditing in the audit log.

Expand table

Parameter	Type	Mandatory?	Description
TargetUserOrGroupName	Edm.String	No	Stores the UPN or name of the target user or group that a resource was shared with.
TargetUserOrGroupType	Edm.String	No	Identifies whether the target user or group is a Member, Guest, Group, or Partner.
EventData	XML code	No	Conveys follow-up information about the sharing action that has occurred, such as adding a user to a group or granting edit permissions.
SiteUrl	Edm.String	No	The URL of the site where the file or folder accessed by the user is located.
SourceRelativeUrl	Edm.String	No	The URL of the folder that contains the file accessed by the user. The combination of the values for the *SiteURL*, *SourceRelativeURL*, and *SourceFileName* parameters is the same as the value for the **ObjectID** property, which is the full path name for the file accessed by the user.
SourceFileName	Edm.String	No	The name of the file or folder accessed by the user.
SourceFileExtension	Edm.String	No	The file extension of the file that was accessed by the user. This property is blank if the object that

Parameter	Type	Mandatory?	Description
			was accessed is a folder.
UniqueSharingId	Edm.String	No	The unique sharing ID associated with the sharing operation.

SharePoint schema

SharePoint events (excluding the file and folder events) returned in audit log searches use this schema.

Parameter	Type	Mandatory?	Description
CustomEvent	Edm.String	No	Optional string for custom events.
EventData	Edm.String	No	Optional payload for custom events.
ModifiedProperties	Collection(ModifiedProperty)	No	The property is included for admin events, such as adding a user as a member of a site or a site collection admin group. The property includes the name of the property that was modified (for example, the Site Admin group), the new value of the modified property (such the user who was added as a site admin), and the previous value of the modified object.

Project schema

Microsoft Project for the web events returned in audit log searches use these schemas.

Parameter	Type	Mandatory?	Description
Entity	Edm.String	Yes	ProjectEntity the audit was for.
Action	Edm.String	Yes	ProjectAction that was taken.
OnBehalfOfResId	Edm.Guid	No	The resource Id the action was taken on behalf of.

Enum: Project Action - Type: Edm.Int32

Project action

Member name	Description
Accepted	The user accepted an event or workflow.
Accessed	The user accessed an entity.
Activated	The user activated an entity, event or workflow.

Member name	Description
Cancelled	The user cancelled an event or workflow.
CheckedIn	The user check in an entity.
CheckedOut	The user checkout an entity.
Copied	The user copied an entity.
Created	The user created an entity.
Deactivated	The user deactivated an entity.
Deleted	The user deleted an entity.
Exported	The user exported an entity.
ForceCheckedIn	The user caused an entity to be force checked in.
Modified	The user modified an entity.
Published	The user published an entity.
Redacted	The user redacted an entity.
Rejected	The user rejected an entity.
Restarted	The user restarted an event or workflow.
Saved	The user saved an entity.
Sent	The user sent an entity.
Submitted	The user submitted an entity for review or workflow.

Enum: Project Entity - Type: Edm.Int32

Project entity

Member name	Description
CustomField	Represents an enterprise custom field.
Driver	Represents a portfolio driver.
DriverPrioritization	Represents a portfolio prioritization.
Engagement	Represents a resource engagement.
EnterpriseCalendar	Represents a enterprise resource calendar.
EnterpriseProjectType	Represents an enterprise project type.
FiscalPeriod	Represents a fiscal period.
GanttChartFormat	Represents a Gantt chart format.

Member name	Description
GroupingFormat	Represents a view grouping format.
LineClassification	Represents a timesheet line classification.
LookupTable	Represents a enterprise lookup table.
PermissionTemplate	Represents a security permission template.
PortfolioAnalysis	Represents a portfolio analysis.
Project	Represents a project.
QueueJob	Represents a queue job.
QuickLaunch	Represents a quick launch item.
Reporting	Represents the reporting endpoint.
Resource	Represents an enterprise resource.
ResourcePlan	Represents a resource plan associated with A project.
SecurityCategory	Represents a security category.
SecurityGroup	Represents a security group.
Setting	Represents a Project web app setting.
Statusing	Represents a task status update.
StatusReport	Represents a status report.
TimeReportingPeriod	Represents a period of time for a timesheet.
Timesheet	Represents a timesheet entity.
TimesheetAuditLog	Represents a timesheet audit log.
TimesheetManager	Represents the manager of a timesheet.
UserDelegate	Represents a user delegation for another user.
View	Represents a view definition.
WorkflowPhase	Represents a phase in a workflow.
WorkflowStage	Represents a stage in a workflow.

Exchange Admin schema

Parameters	Type	Mandatory	Description
ModifiedObjectResolvedName	Edm.String	No	This is the user friendly name of the object that was

Parameters	Type	Mandatory	Description
			modified by the cmdlet. This is logged only if the cmdlet modifies the object.
Parameters	Collection(Common.NameValuePair)	No	The name and value for all parameters that were used with the cmdlet that is identified in the Operations property.
ModifiedProperties	Collection(Common.ModifiedProperty)	No	The property is included for admin events. The property includes the name of the property that was modified, the new value of the modified property, and the previous value of the modified object.
ExternalAccess	Edm.Boolean	Yes	Specifies whether the cmdlet was run by a user in your organization, by Microsoft datacenter personnel or a datacenter service account, or by a delegated administrator. The value **False** indicates that the cmdlet was run by someone in your organization. The value **True** indicates that the cmdlet was run by datacenter personnel, a datacenter service account, or a delegated administrator.

Parameters	Type	Mandatory	Description
OriginatingServer	Edm.String	No	The name of the server from which the cmdlet was executed.
OrganizationName	Edm.String	No	The name of the tenant.
DeviceId	Edm.String	No	Available only for registered/domain joined device.
TokenObjectId	Edm.String	Yes	oid claim of the Microsoft Entra tokens.
TokenTenantId	Edm.String	Yes	tid claim of the Microsoft Entra tokens.

Exchange Mailbox schema

Exchange mailbox events returned in audit log searches use these schemas.

Parameters	Type	Mandatory	Description
LogonType	Self.LogonType	Yes	Indicates the type of user who accessed the mailbox and performed the operation that was logged.
InternalLogonType	Self.LogonType	Yes	Reserved for internal use.
MailboxGuid	Edm.String	No	The Exchange GUID of the mailbox that was accessed.
MailboxOwnerUPN	Edm.String	No	The email address of the person who owns the mailbox that was accessed.
MailboxOwnerSid	Edm.String	No	The SID of the mailbox owner.
MailboxOwnerMasterAccountSid	Edm.String	No	Mailbox owner account's master account SID.

Parameters	Type	Mandatory	Description
LogonUserSid	Edm.String	No	The SID of the user who performed the operation.
LogonUserDisplayName	Edm.String	No	The user-friendly name of the user who performed the operation.
ExternalAccess	Edm.Boolean	Yes	This is true if the logon user's domain is different from the mailbox owner's domain.
OriginatingServer	Edm.String	No	This is from where the operation originated.
OrganizationName	Edm.String	No	The name of the tenant.
ClientInfoString	Edm.String	No	Information about the email client that was used to perform the operation, such as a browser version, Outlook version, and mobile device information.
ClientIPAddress	Edm.String	No	The IP address of the device that was used when the operation was logged. The IP address is displayed in either an IPv4 or IPv6 address format.
ClientMachineName	Edm.String	No	The machine name that hosts the Outlook client.
ClientProcessName	Edm.String	No	The email client that was used to access the mailbox.
ClientVersion	Edm.String	No	The version of the email client.
SessionId	Edm.String	No	Unique identifier for the session. It helps to differentiate attacker actions vs day-to-day user activities on the

Parameters	Type	Mandatory	Description
			same account (useful for compromised accounts)
AppId	Edm.String	No	Application Identifier
ClientAppId	Edm.String	No	Original caller(service/protocol) identifier of the request.
HostAppId	Edm.String	No	The executing service/protocol identifier.
OperationProperties	Collection(Common.NameValuePair)	No	Collection of operation-specific properties.
ClientRequestId	Edm.String	No	Identifier for the client request.
ExternalUserTenantId	Edm.String	No	Tenant identifier for external users.
ActorIP	Edm.String	No	IP address of the actor performing the operation.
ActorInfoString	Edm.String	No	Name of the cloud instance (e.g., Public, GCC, GCC-H)
DeviceId	Edm.String	No	Available only for registered/domain joined device.
CloudInstanceName	Edm.String	No	Name of the cloud instance (e.g., Public, GCC, GCC-H)
TokenObjectId	Edm.String	No	oid claim of the Microsoft Entra tokens.
TokenTenantId	Edm.String	No	tid claim of the Microsoft Entra tokens.
AuthType	Edm.String	No	Specifies the authentication scheme used by the client to access the service.

Parameters	Type	Mandatory	Description
TokenType	Edm.String	No	Indicates the type of token presented during authentication, providing context on the token's role and scope.
recipientCount	Edm.Int64	No	Represents the total number of recipients, to whom email was sent.
recipientList	Collection(Self.EmailAddress)	No	Captures the list of mailbox recipients, maximum 1000.

Enum: LogonType - Type: Edm.Int32

LogonType

Value	Member name	Description
0	Owner	The mailbox owner.
1	Admin	A person with administrative privileges for someone's mailbox.
2	Delegated	A person with the delegate privileges for someone's mailbox.
3	Transport	A transport service in the Microsoft datacenter.
4	SystemService	A service account in the Microsoft datacenter
5	BestAccess	Reserved for internal use.
6	DelegatedAdmin	A delegated administrator.

ExchangeMailboxAuditGroupRecord schema

Parameters	Type	Mandatory?	Description
Folder	Self.ExchangeFolder	No	The folder where a group of items is located.
CrossMailboxOperations	Edm.Boolean	No	Indicates if the operation involved more than one mailbox.
DestMailboxId	Edm.Guid	No	Set only if the CrossMailboxOperations parameter is **True**.

Parameters	Type	Mandatory?	Description
			Specifies the target mailbox GUID.
DestMailboxOwnerUPN	Edm.String	No	Set only if the CrossMailboxOperations parameter is **True**. Specifies the UPN of the owner of the target mailbox.
DestMailboxOwnerSid	Edm.String	No	Set only if the CrossMailboxOperations parameter is **True**. Specifies the SID of the target mailbox.
DestMailboxOwnerMasterAccountSid	Edm.String	No	Set only if the CrossMailboxOperations parameter is **True**. Specifies the SID for the master account SID of the target mailbox owner.
DestFolder	Self.ExchangeFolder	No	The destination folder, for operations such as Move.
Folders	Collection(Self.ExchangeFolder)	No	Information about the source folders involved in an operation; for example, if folders are selected and then deleted.
AffectedItems	Collection(Self.ExchangeItem)	No	Information about each item in the group.
Teams	Self.TeamsData	No	Microsoft Teams-related data for group operations.

ExchangeMailboxAuditRecord schema

Parameters	Type	Mandatory?	Description
Item	Self.ExchangeItem	No	Represents the item upon which the operation was performed
ModifiedProperties	Collection(Edm.String)	No	Properties that are modified.

Parameters	Type	Mandatory?	Description
MbxLoginLocalQueueADLookupInfo	Self.LocalQueueADLookupInfo	No	Contains Active Directory lookup details for the mailbox login context.
SendAsUserSmtp	Edm.String	No	SMTP address of the user who is being impersonated.
SendAsUserMailboxGuid	Edm.Guid	No	The Exchange GUID of the mailbox that was accessed to send email as.
SendOnBehalfOfUserSmtp	Edm.String	No	SMTP address of the user on whose behalf the email is sent.
SendOnBehalfOfUserMailboxGuid	Edm.Guid	No	The Exchange GUID of the mailbox that was accessed to send mail on behalf of.
ContactEmail1EmailAddress	Edm.String	No	First email address for contact information.
ContactEmail1DisplayName	Edm.String	No	Display name for first contact email.
ContactEmail2EmailAddress	Edm.String	No	Second email address for contact information.
ContactEmail2DisplayName	Edm.String	No	Display name for second contact email.
ContactEmail3EmailAddress	Edm.String	No	Third email address for contact information.
ContactEmail3DisplayName	Edm.String	No	Display name for third contact email.
AttachmentId	Edm.String	No	Identifier for email attachment.
AttachmentSizeInBytes	Edm.Int64	No	Size of attachment in bytes.

Parameters	Type	Mandatory?	Description
SaveToSentItems	Edm.Boolean	No	Flag indicating whether item should be saved to sent items folder.
Teams	Self.TeamsData	No	Microsoft Teams-related data.

ExchangeAggregatedMailboxAuditRecord schema

Parameters	Type	Mandatory?	Description
OperationCount	Edm.Int32	No	The total number of operations performed on the aggregated set of items.
Folders	Collection(Self.ExchangeAggregatedFolder)	No	A collection of aggregated folder objects involved in the operation.
Messages	Collection(Self.ExchangeAggregatedMessage)	No	A collection of aggregated message objects involved in the operation.

ExchangeAggregatedOperationRecord schema

Parameters	Type	Mandatory?	Description
OperationCount	Edm.Int32	No	The total number of operations included aggregated in this record.
AggregateDurationInSeconds	Edm.Int32	No	Combined duration of all aggregated operations in seconds.

ExchangeItem complex type

Parameters	Type	Mandatory?	Description
Id	Edm.String	Yes	The store ID.
Subject	Edm.String	No	The subject line of the message that was accessed.
ParentFolder	Self.ExchangeFolder	No	The name of the folder where the item is located.
ParentMessage	Self.ExchangeMessage	No	Parent message if this item is nested.

Parameters	Type	Mandatory?	Description
Attachments	Edm.String	No	A list of the names and file size of all items that are attached to the message.
ImmutableId	Edm.String	No	A permanent, unchanging identifier for the item.
InternetMessageId	Edm.String	No	Internet message identifier.
ComplianceLabel	Edm.String	No	Compliance label applied to the item.
IsRecord	Edm.Boolean	No	Flag indicating if the item is a record.
IsPriorityCleanup	Edm.Boolean	No	Flag for priority cleanup processing
SizeInBytes	Edm.Int64	No	Size of the Exchange item in bytes.
Sensitivity	Edm.String	No	Indicates the sensitivity level of the item (e.g., Normal, Personal, Private, Confidential).

LocalQueueADLookupInfo complex type

Parameters	Type	Mandatory?	Description
Puid	Edm.String	No	Passport unique identifier (Puid) for a Microsoft account or user object.
OrgIdPuid	Edm.String	No	Organization identifier PUID.
Smtp	Edm.String	No	SMTP address for AD lookup.
SearchScope	Edm.String	No	Active Directory search scope
ConsumerMailbox	Edm.String	No	Flag indicating if this is a consumer mailbox.

EmailAddress complex type

Parameters	Type	Mandatory?	Description
Address	Edm.String	No	Email address associated with user.
Name	Edm.String	No	Display name for email address.

ExchangeFolder complex type

Parameters	Type	Mandatory?	Description
Id	Edm.String	Yes	The store ID of the folder object.
Path	Edm.String	No	The name of the mailbox folder where the message that was accessed is located.

Parameters	Type	Mandatory?	Description
Name	Edm.String	No	Folder Name
MemberSid	Edm.String	No	Member security identifier
MemberRights	Edm.String	No	Access Rights on the folder
MemberUpn	Edm.String	No	Member user principal name

ExchangeMessage complex type

Parameters	Type	Mandatory?	Description
Id	Edm.String	Yes	Unique identifier for the Message.
Path	Edm.String	No	The path where the message is located

ExchangeFolderItem complex type

Parameters	Type	Mandatory?	Description
Id	Edm.String	No	A guid that identifies the folder item within Exchange.
ImmutableId	Edm.String	No	Immutable identifier for each Exchange folder item.
InternetMessageId	Edm.String	No	Internet message identifier for each folder item.
CreationTime	Edm.Date	No	Time when Record/item was created.
Subject	Edm.String	No	Subject of the record/item.
SizeInBytes	Edm.Int64	No	Size of the record/item in bytes.
Sensitivity	Edm.String	No	Sensitivity labels of the record/item.
ClientRequestId	Edm.String	No	A unique identifier for the client request that triggered the operation.
Teams	Self.TeamsData	No	Microsoft Teams-related data

ExchangeMessageItem complex type

Parameters	Type	Mandatory?	Description
Id	Edm.String	No	A guid that identifies the Exchange Message Item.
SizeInBytes	Edm.Int64	No	The size of the message in bytes, including any attachments.

ExchangeAggregatedFolder complex type

Parameters	Type	Mandatory?	Description
Id	Edm.String	Yes	The guid of the aggregated folder.
Path	Edm.String	No	The path of the aggregated folder.
FolderItems	Collection(Self.ExchangeFolderItem)	No	The collection of Exchange folder items.

ExchangeAggregatedMessage complex type

Parameters	Type	Mandatory?	Description
Id	Edm.String	Yes	The guid of the aggregated message container.
Path	Edm.String	No	The path of the aggregated message.
MessageItems	Collection(Self.ExchangeMessageItem)	No	The collection of exchange message items.

OWA Auth schema

Parameters	Type	Mandatory?	Description
UniqueTokenIdentifier	Edm.String	No	A unique identifier for the resource.
ResourceURL	Edm.String	No	The resource URL.

Azure Active Directory Base schema

Parameters	Type	Mandatory?	Description
AzureActiveDirectoryEventType	Self.AzureActiveDirectoryEventType	Yes	The type of Microsoft Entra event.
ExtendedProperties	Collection(Common.NameValuePair)	No	The extended properties of the Microsoft Entra event.
ModifiedProperties	Collection(Common.ModifiedProperty)	No	This property is included for admin events. The property includes the name of the property that was modified, the new value of the modified

Parameters	Type	Mandatory?	Description
			property, and the previous value of the modified property.

Enum: AzureActiveDirectoryEventType - Type -Edm.Int32

AzureActiveDirectoryEventType

Member name	Description
AccountLogon	The account login event.
AzureApplicationAuditEvent	The Azure application security event.

Azure Active Directory Account Logon schema

Parameters	Type	Mandatory?	Description
Application	Edm.String	No	The application that triggers the account login event, such as Office 15.
Client	Edm.String	No	Details about the client device, device OS, and device browser that was used for the of the account login event.
LoginStatus	Edm.Int32	Yes	This property is from OrgIdLogon.LoginStatus directly. The mapping of various interesting logon failures could be done by alerting algorithms.
UserDomain	Edm.String	Yes	The Tenant Identity Information (TII).

Enum: CredentialType - Type: Edm.Int32

Value	Member name	Description
-1	Other	Other authentication.
0	Password	User credential is username and password.
1	MobilePhone	User credential is mobile phone.
2	SecretQuestion	User credential is secret question.
3	SecurePin	User credential is secure PIN.
4	SecurePinReset	User credential is secure PIN reset.
11	EasyID	User credential is EasyID.
14	PasswordIndexCredentialType	User credential is PasswordIndexCredentialType.

Value	Member name	Description
16	Device	User credential is a device.
17	ForeignRealmIndex	User credential is ForeignRealmIndex.

Enum: LoginType - Type: Edm.Int32

Value	Member name	Description
-1	Other	Other i type.
1	InitialAuth	Login with initial authentication
2	CookieCopy	Login with cookie.
3	SilentReAuth	Login with silent re-authentication.

Enum: AuthenticationMethod - Type: Edm.Int32

Value	Member name	Description
0	Min	The authentication method is a Min
1	Password	The authentication method is a password.
2	Digest	The authentication method is a digest.
3	ProxyAuth	The authentication method is a ProxyAuth.
4	InfoCard	The authentication method is an InfoCard
5	DAToken	The authentication method is a DAToken.
6	Sha1RememberMyPassword	The authentication method is a Sha1RememberMyPassword.
7	LMPasswordHash	The authentication method is an LMPasswordHash.
8	ADFSFederatedToken	The authentication method is an ADFSFederatedToken.
9	EID	The authentication method is an EID.
10	DeviceID	The authentication method is a DeviceID.
11	MD5	The authentication method is MD5.
12	EncProxyPasswordHash	The authentication method is a EncProxyPasswordHash.
13	LWAFederation	The authentication method is a LWAFederation.
14	Sha1HashedPassword	The authentication method is a Sha1HashedPassword.
15	SecurePin	The authentication method is a secure Pin.
16	SecurePinReset	The authentication method is a secure PIN reset.

Value	Member name	Description
17	SAML20PostSimpleSign	The authentication method is a SAML20PostSimpleSign.
18	SAML20Post	The authentication method is a SAML20Post.
19	OneTimeCode	The authentication method is a one-time code.

Azure Active Directory schema

Parameters	Type	Mandatory?	Description
Actor	Collection(Self.IdentityTypeValuePair)	No	The user or service principal that performed the action.
ActorContextId	Edm.String	No	The GUID of the organization that the actor belongs to.
ActorIpAddress	Edm.String	No	The actor's IP address in IPV4 or IPV6 address format.
InterSystemsId	Edm.String	No	The GUID that track the actions across components within the Office 365 service.
IntraSystemsId	Edm.String	No	The GUID that's generated by Azure Active Directory to track the action.
SupportTicketId	Edm.String	No	The customer support ticket ID for the action in "act-on-behalf-of" situations.
Target	Collection(Self.IdentityTypeValuePair)	No	The user that the action (identified by the Operation property) was performed on.
TargetContextId	Edm.String	No	The GUID of the organization that the targeted user belongs to.

Complex Type IdentityTypeValuePair

Parameters	Type	Mandatory?	Description
ID	Edm.String	Yes	The value of the identity given the type.
Type	Self.IdentityType	Yes	The type of the identity.

Enum: IdentityType - Type: Edm.Int32

IdentityType

Member name	Description
Claim	The identity is a claim for authorization purpose.
Name	The audit action actor or target identity display name.
Other	The identity of the actor is other type, such as the ObjectId in GUID generated by the Office 365 service.
PUID	The audit action actor or the target passport unique ID (PUID).
SPN	The identity of a service principal if the action is performed by the Office 365 service.
UPN	The user principal name.

Azure Active Directory Secure Token Service (STS) Logon schema

Parameters	Type	Mandatory?	Description
ApplicationId	Edm.String	No	The GUID that represents the application that is requesting the login. The display name can be looked up via the Azure Active Directory Graph API.
Client	Edm.String	No	Client device information, provided by the browser performing the login.
DeviceProperties	Collection(Common.NameValuePair)	No	This property includes various device details, including Id, Display name, OS, Browser, IsCompliant, IsCompliantAndManaged, SessionId, and DeviceTrustType. The DeviceTrustType property can have the following values: **0** - Microsoft Entra registered **1** - Microsoft Entra joined **2** - Hybrid Microsoft Entra joined
ErrorCode	Edm.String	No	For failed logins (where the value for the Operation property is UserLoginFailed), this property contains the Azure Active Directory STS (AADSTS) error code. For descriptions of these error codes, see Authentication and authorization error codes. A value of `0` indicates a successful login.

Parameters	Type	Mandatory?	Description
LogonError	Edm.String	No	For failed logins, this property contains a user-readable description of the reason for the failed login.

DLP schema

Data loss prevention (DLP) events in Microsoft Purview are available for Exchange Online, Endpoint(devices), and SharePoint and OneDrive.

DLP events always have `UserKey="DlpAgent"` in the common schema. There are three types of DlpEvents that are stored as the value of the Operation property of the common schema:

- **DlpRuleMatch**: Indicates a rule was matched. These events exist in all Exchange, Endpoint(devices) and SharePoint and OneDrive. For Exchange it includes false positive and override information. For SharePoint and OneDrive, false positive and overrides generate separate events.
- **DlpRuleUndo**: SharePoint and OneDrive only. Indicates a previously applied policy action has been "undone" due to:
  - AFalse positive/override designation by user.
  - The document is no longer subject to policy (either due to policy change or change to content in doc).
- **DlpInfo**: SharePoint and OneDrive only. Indicates a false positive designation but no action was "undone."

Parameters	Type	Mandatory	Description
SharePointMetaData	Self.SharePointMetadata	No	Describes metadata about the document in SharePoint or OneDrive that contained the sensitive information.
ExchangeMetaData	Self.ExchangeMetadata	No	Describes metadata about the email message that contained the sensitive information.
EndpointMetaData	Self.EndpointMetadata	No	Describes metadata about the document in endpoint that contained the sensitive information
ExceptionInfo	Edm.String	No	Identifies reasons why a policy no longer applies and/or any information about false positive and/or

Parameters	Type	Mandatory	Description
			override noted by the end user.
PolicyDetails	Collection(Self.PolicyDetails)	Yes	Information about 1 or more policies that triggered the DLP event.
SensitiveInfoDetectionIsIncluded	Boolean	Yes	Indicates whether the event contains the value of the sensitive data type and surrounding context from the source content. Accessing sensitive data requires the "Read DLP policy events including sensitive details" permission in Azure Active Directory.

eDiscovery schema

The eDiscovery audit schema is designed to capture and log activities related to eDiscovery processes within the organization.

Parameters	Type	Mandatory	Description
CaseId	Edm.Guid	No	The identity (GUID) of the eDiscovery case.
CaseName	Edm.String	No	The name of the eDiscovery case.
Object1Id	Edm.String	No	The ID of the object (for example, a search, hold, or review set) that was created, accessed, or changed.
Object1Name	Edm.String	No	The name of the object (for example, a search, hold or review set) that was created, accessed, or changed.
Object1Type	Edm.String	No	The type of the eDiscovery object that the user created, deleted, or modified.
Object2Id	Edm.String	No	The ID of the object (for example, a search, hold, or review set) that was created, accessed, or changed.
Object2Name	Edm.String	No	The name of the object (for example, a search, hold or review

Parameters	Type	Mandatory	Description
			set) that was created, accessed, or changed.
Object2Type	Edm.String	No	The type of the eDiscovery object that the user created, deleted, or modified.
StartTime	Edm.Date	No	The date and time in Coordinated Universal Time (UTC) when the eDiscovery activity was started.
EndTime	Edm.Date	No	The date and time in Coordinated Universal Time (UTC) when the eDiscovery activity was ended.
UserCancelled	Edm.Boolean	No	Whether the specific activity was cancelled by the user.
ItemIds	Collection(Edm.String)	No	Item IDs associated with the activity.
ItemNames	Collection(Edm.String)	No	Item names associated with the activity.
DataSources	Collection(Edm.String)	No	A list of source IDs, source names and locations associated to the activity.
QueryId	Edm.String	No	The ID of the query associated with the activity.
QueryText	Edm.String	No	The query text associated with the activity, such as a search statistic process or add to review process.
QueryFiles	Collection(Edm.String)	No	Input file names used to generate the query. This is specific to the search by file gesture.
Settings	Collection(Common.NameValuePair)	No	Settings applied to the eDiscovery activity.
ExtendedProperties	Collection(Common.NameValuePair)	No	Additional properties related to the eDiscovery activity.
ExportName	Edm.String	No	Name of the eDiscovery export.
JobId	Edm.String	No	The GUID of the eDiscovery process.

Parameters	Type	Mandatory	Description
RecordNumber	Edm.String	No	Used when an audit record is divided into multiple parts due to size. It indicates the sequence of each part within the total splits.

SharePointMetadata complex type

Parameters	Type	Mandatory?	Description
From	Edm.String	Yes	The user who triggered the event. Values are the FileOwner, LastModifier, or LastSharer.
itemCreationTime	Edm.Date	Yes	Date/time stamp in UTC of when event logged.
SiteCollectionGuid	Edm.Guid	Yes	The GUID of the site collection.
SiteCollectionUrl	Edm.String	Yes	Name of the SharePoint site.
FileName	Edm.String	Yes	Name of the path.
FileOwner	Edm.String	Yes	The document owner.
FilePathUrl	Edm.String	Yes	The URL of the document
DocumentLastModifier	Edm.String	Yes	The user who last modified the document.
DocumentSharer	Edm.String	Yes	The user who last modified sharing of the document.
UniqueId	Edm.String	Yes	A guid that identifies the file.
LastModifiedTime	Edm.DateTime	Yes	Timestamp in UTC for when doc was last modified.
IsViewableByExternalUsers	Edm.Boolean	Yes	Determines if the file is accessible to any external user.

ExchangeMetadata complex type

Parameters	Type	Mandatory?	Description
MessageID	Edm.String	Yes	The message ID of the email that triggered the event.
From	Edm.String	Yes	The user who sent the email.
To	Collection(Edm.String)	No	A collection of email addresses that were on the To line of the message.
CC	Collection(Edm.String)	No	A collection of email addresses that were on the CC line of the message.

Parameters	Type	Mandatory?	Description
BCC	Collection(Edm.String)	No	A collection of email addresses that were on the BCC line of the message.
Subject	Edm.String	Yes	Subject of the email message.
Sent	Edm.DateTime	Yes	The time in UTC of when the email was sent.
RecipientCount	Edm.Int32	Yes	The total number of all recipients on the TO, CC, and BCC lines of the message.

EndpointMetadata complex type

Parameters	Type	Mandatory?	Description
SensitiveInformation	Collection(Self.SensitiveInformation)	No	Information about the type of sensitive information detected.
EnforcementMode	Edm.String	Yes	Indicate whether the DLP Rule set to 1/2/3/4/5 depicting audit/warn(block with override)/warn and bypass/block/allow(audit without alerts) respectively.
FileExtension	Edm.String	No	The file extension of the document that contained the sensitive information.
FileType	Edm.String	No	The file type of the document that contained the sensitive information.
DeviceName	Edm.String	No	The name of the device on which DLP rule match was detected.

PolicyDetails complex type

Parameters	Type	Mandatory?	Description
PolicyId	Edm.Guid	Yes	The guid of the DLP policy for this event.
PolicyName	Edm.String	Yes	The friendly name of the DLP policy for this event.
Rules	Collection(Self.Rules)	Yes	Information about the rules within the policy that were matched for this event.

Rules complex type

Parameters	Type	Mandatory?	Description
RuleId	Edm.Guid	Yes	The guid of the DLP rule for this event.
RuleName	Edm.String	Yes	The friendly name of the DLP rule for this event.
Actions	Collection(Edm.String)	No	A list of actions taken as a result of a DLP RuleMatch event.
OverriddenActions	Collection(Edm.String)	No	A list of actions previously taken that were now undone as a result of a DLPRuleUndo event.
Severity	Edm.String	No	The severity (Low, Medium and High) of the rule match.
RuleMode	Edm.String	Yes	Indicate whether the DLP Rule was set to Enforce, Audit with Notify, or Audit only.
ConditionsMatched	Self.ConditionsMatched	No	Details about what conditions of the rule were matched for this event.

ConditionsMatched complex type

Parameters	Type	Mandatory?	Description
SensitiveInformation	Collection(Self.SensitiveInformation)	No	Information about the type of sensitive information detected.
DocumentProperties	Collection(NameValuePair)	No	Information about document properties that triggered a rule match.
OtherConditions	Collection(NameValuePair)	No	A list of key value pairs describing any other conditions that were matched.

SensitiveInformation complex type

Parameters	Type	Mandatory?	Description
Confidence	Edm.Int	Yes	The aggregated confidence of all pattern matches for the Sensitive Information Type.
Count	Edm.Int	Yes	The total number of sensitive

Parameters	Type	Mandatory?	Description
			instances detected.
Location	Edm.String	No	
SensitiveType	Edm.Guid	Yes	A guid that identifies the type of sensitive data detected.
SensitiveInformationDetections	Self.SensitiveInformationDetections	No	An array of objects that contain sensitive information data with the following details: matched value and context of matched value.
SensitiveInformationDetailed ClassificationAttributes	Collection(SensitiveInformationDetailed ConfidenceLevelResult)	Yes	Information about the count of sensitive information type detected for each of the three confidence levels (High, Medium and Low) and wether it matches the DLP rule or not.
SensitiveInformationTypeName	Edm.String	No	The name of the sensitive information type.
UniqueCount	Edm.Int32	Yes	The unique count of sensitive instances detected.

SensitiveInformationDetailedClassificationAttributes complex type

Parameters	Type	Mandatory?	Description
Confidence	Edm.int32	Yes	The confidence level of the pattern that was detected.

Parameters	Type	Mandatory?	Description
Count	Edm.Int32	Yes	The number of sensitive instances detected for a particular confidence level.
IsMatch	Edm.Boolean	Yes	Indicates if the given count and confidence level of the sensitive type detected results in a DLP rule match.

SensitiveInformationDetections complex type

DLP sensitive data is only available in the activity feed API to users that have been granted "Read DLP sensitive data" permissions.

Parameters	Type	Mandatory?	Description
DetectedValues	Collection(Common.NameValuePair)	Yes	An array of sensitive information that was detected. Information contains key value pairs with Value = matched value (eg. Value of credit card) and Context = an excerpt from source content that contains the matched value.
ResultsTruncated	Edm.Boolean	Yes	Indicates if the logs were truncated due to large number of results.

ExceptionInfo complex type

Parameters	Type	Mandatory?	Description
Reason	Edm.String	No	For a DLPRuleUndo event, this indicates why the rule no longer applies, which can be one of 3 reasons: Override, Document Change, or Policy Change
FalsePositive	Edm.Boolean	No	Indicates whether the user designated this event as a false positive.
Justification	Edm.String	No	If the user chose to override policy, any user-specified justification is captured here.
Rules	Collection(Edm.Guid)	No	A collection of GUIDs for each rule that was designated as a false positive or override, or for which an action was undone.

Security and Compliance Center schema

Parameters	Type	Mandatory	Description
StartTime	Edm.Date	No	The date and time when the cmdlet was run.

Parameters	Type	Mandatory	Description
ClientRequestId	Edm.String	No	A GUID that can be used to correlate this cmdlet with portal operations. This information is used only by Microsoft support.
CmdletVersion	Edm.String	No	The build version of the cmdlet when it was run.
EffectiveOrganization	Edm.String	No	The GUID for the organization impacted by the cmdlet. (Deprecated: This parameter will eventually stop appearing.)
UserServicePlan	Edm.String	No	The service plan assigned to the user who ran the cmdlet.
ClientApplication	Edm.String	No	If the cmdlet was run by an application, as opposed to remote PowerShell, this field contains the application name.
Parameters	Edm.String	No	The name and value for parameters that were used with the cmdlet that do not include personal data.
NonPiiParameters	Edm.String	No	The name and value for parameters that were used with the cmdlet that include personal data. (Deprecated: This field will eventually stop appearing and its content will be merged with the Parameters field.)

Security and Compliance Alerts schema

Alert signals include:

- All alerts generated by Alert policies in the Microsoft Defender portal.
- Microsoft 365 related alerts generated in Microsoft Defender for Cloud Apps.

The UserId and UserKey of these events are always SecurityComplianceAlerts. There are three types of alert events that are stored as the value of the Operation property of the common schema:

- **AlertTriggered**: A new alert is generated due to a policy match.
- **AlertEntityGenerated**: A new entity is added to an alert. This event is only applicable to alerts generated based on Alert policies in the Microsoft Defender portal. Each generated alert can be associated with one or multiple of these events. For example, an alert policy is defined to trigger an alert if any user deletes more than 100 files in 5 minutes. If two users exceed the threshold around the same time, there are two AlertEntityGenerated events, but only one AlertTriggered event.
- **AlertUpdated**: An update was made to the metadata of an alert. This event is logged when the status of an alert is changed (for example, from **Active** to **Resolved**) and when someone adds a comment to the alert.

Parameters	Type	Mandatory	Description
AlertId	Edm.Guid	Yes	The Guid of the alert.
AlertType	Self.String	Yes	Type of the alert. Alert types include: • System • Custom
Name	Edm.String	Yes	Name of the alert.
PolicyId	Edm.Guid	No	The Guid of the policy that triggered the alert.
Status	Edm.String	No	Status of the alert. Statuses include: • Active • Investigating • Resolved • Dismissed
Severity	Edm.String	No	Severity of the alert. Severity levels include: • Low • Medium • High
Category	Edm.String	No	Category of the alert. Categories include: • AccessGovernance • DataGovernance • DataLossPrevention • InsiderRiskManagement • MailFlow • ThreatManagement • Other
Source	Edm.String	No	Source of the alert. Sources include: • Office 365 Security & Compliance • Cloud App Security
Comments	Edm.String	No	Comments left by the users who have viewed the alert. By default, it's "New alert".
Data	Edm.String	No	The detailed data blob of the alert or alert entity.
AlertEntityId	Edm.String	No	The identifier for the alert entity. This parameter is only applicable to AlertEntityGenerated events.
EntityType	Edm.String	No	Type of the alert or alert entity. Entity types include: • User • Recipients • Sender • MalwareFamily This parameter applies only to AlertEntityGenerated events.

Viva Engage schema

Viva Engage events returned in audit log searches use this schema.

Parameters	Type	Mandatory	Description
ActorUserId	Edm.String	No	Email of user that performed the operation.
ActorYammerUserId	Edm.Int64	No	ID of user that performed the operation.
DataExportType	Edm.String	No	Returns "data" if data export includes messages, notes, files, topics, users and groups; returns "user" if data export includes users only.
FileId	Edm.Int64	No	ID of the file in the operation.
FileName	Edm.String	No	Name of the file in the operation. Appears blank if not relevant to the operation.
GroupName	Edm.String	No	Name of the group in the operation. Appears blank if not relevant to the operation.
IsSoftDelete	Edm.Boolean	No	Returns "true" if the network's data retention policy is set to Soft Delete; returns "false" if the network's data retention policy is set to Hard Delete.
MeetingId	Edm.String	No	The ID of the event/teams meeting in the operation.
MessageId	Edm.Int64	No	ID of the message in the operation.
ModifiedProperties	Collection(ModifiedProperty)	No	Includes the name of the property that was modified, the new value of the modified object and the previous value of the modified object.
YammerNetworkId	Edm.Int64	No	Network ID of the user that performed the operation.
TargetObjectId	Edm.String	No	Entra Id of the target user in the operation.
TargetUserId	Edm.String	No	Email of target user in the operation. Appears blank if not relevant to the operation.
TargetYammerUserId	Edm.Int64	No	ID of target user in the operation.
ThreadId	Edm.Int64	No	ID of the Message thread in the operation.

Parameters	Type	Mandatory	Description
VersionId	Edm.Int64	No	Version ID of the file in the operation.

Data Center Security Base schema

Parameters	Type	Mandatory?	Description
DataCenterSecurityEventType	Self.DataCenterSecurityEventType	Yes	The type of cmdlet event in lock box.

Enum: DataCenterSecurityEventType - Type: Edm.Int32

DataCenterSecurityEventType

Member name	Description
DataCenterSecurityCmdletAuditEvent	This is the enum value for cmdlet audit type event.

Data Center Security Cmdlet schema

Parameters	Type	Mandatory?	Description
StartTime	Edm.Date	Yes	The start time of the cmdlet execution.
EffectiveOrganization	Edm.String	Yes	The name of the tenant that the elevation/cmdlet was targeted at.
ElevationTime	Edm.Date	Yes	The start time of the elevation.
ElevationApprover	Edm.String	Yes	The name of a Microsoft manager.
ElevationApprovedTime	Edm.Date	No	The timestamp for when the elevation was approved.
ElevationRequestId	Edm.Guid	Yes	A unique identifier for the elevation request.
ElevationRole	Edm.String	No	The role the elevation was requested for.
ElevationDuration	Edm.Int32	Yes	The duration for which the elevation was active.
GenericInfo	Edm.String	No	Used for comments and other generic information.

Microsoft Teams schema

Microsoft Teams events returned in audit log searches use these schemas.

Parameters	Type	Mandatory?	Description
Action	Edm.String	No	For shared channel events, the action taken by the invitee or the channel

Parameters	Type	Mandatory?	Description
			owner for a share with team invite.
AddOnGuid	Edm.Guid	No	A unique identifier for the add-on that generated the event.
AddOnName	Edm.String	No	The name of the add-on that generated the event.
AddOnType	Self.AddOnType	No	The type of add-on that generated this event.
CallId	Edm.String	No	A unique identifier for the call.
ChannelGuid	Edm.Guid	No	A unique identifier for the channel being audited.
ChannelName	Edm.String	No	The name of the channel being audited.
ChannelType	Edm.String	No	The type of channel being audited (Standard/Private).
ExtraProperties	Collection(Self.KeyValuePair)	No	A list of extra properties.
HostedContents	Collection(Self.HostedContent)	No	A collection of chat or channel message hosted contents.
ImpersonationType	Edm.String	No	The type of impersonation detected, such as brand, domain, or user.
Invitee	Edm.String	No	For shared channel events, the UPN of the invitee team owner who accepts or declines the invite for a share with team invite.
Members	Collection(Self.MicrosoftTeamsMember)	No	A list of users within a Team.
MessageId	Edm.String	No	An identifier for a chat or channel message.
MessageURLs	Edm.String	No	Present for any URL sent in Teams messages.
Messages	Collection(Self.Message)	No	A collection of chat or channel messages.
MessageSizeInBytes	Edm.Int64	No	The size of a chat or channel message in bytes

Parameters	Type	Mandatory?	Description
			with UTF-16 encoding.
Name	Edm.String	No	Only present for settings events. Name of the setting that changed.
NewValue	Edm.String	No	Only present for settings events. New value of the setting.
OldValue	Edm.String	No	Only present for settings events. Old value of the setting.
Organizer	Self.MicrosoftTeamsMember	No	The organizer of the meeting or call.
SubscriptionId	Edm.String	No	A unique identifier of a Microsoft Graph change notification subscription.
TabType	Edm.String	No	Only present for tab events. The type of tab that generated the event.
TeamGuid	Edm.Guid	No	A unique identifier for the team being audited.
TeamName	Edm.String	No	The name of the team being audited.

MicrosoftTeamsMember complex type

Parameters	Type	Mandatory?	Description
UPN	Edm.String	No	The user principal name of the user.
Role	Self.MemberRoleType	No	The role of the user within the team.
DisplayName	Edm.String	No	The display name of the user.

Enum: MemberRoleType - Type: Edm.Int32

MemberRoleType

Value	Member name	Description
0	Member	A user who is a member of the team.
1	Owner	A user who is the owner of the team.
2	Guest	A user who is not a member of the team.

KeyValuePair complex type

Parameters	Type	Mandatory?	Description
Key	Edm.String	No	The key of the key-value pair.
Value	Edm.String	No	The value of the key-value pair.

Enum: AddOnType - Type: Edm.Int32

AddOnType

Value	Member name	Description
1	Bot	A Microsoft Teams bot.
2	Connector	A Microsoft Teams connector.
3	Tab	A Microsoft Teams tab.

HostedContent complex type

Parameters	Type	Mandatory?	Description
Id	Edm.String	Yes	A unique identifier of the message hosted content.
SizeInBytes	Edm.Int64	No	The message hosted content size in bytes.

Message complex type

Parameters	Type	Mandatory?	Description
AADGroupId	Edm.String	No	A unique identifier of the group in Azure Active Directory that the message belongs to.
Id	Edm.String	Yes	A unique identifier of the chat or channel message.
ChannelGuid	Edm.String	No	A unique identifier of the channel the message belongs to.
ChannelName	Edm.String	No	The name of the channel the message belongs to.
ChannelType	Edm.String	No	The type of the channel the message belongs to.
ChatName	Edm.String	No	The name of the chat the message belongs to.
ChatThreadId	Edm.String	No	A unique identifier of the chat the message belongs to.
ParentMessageId	Edm.String	No	A unique identifier of the parent chat or channel message.
SizeInBytes	Edm.Int64	No	The size of the message in bytes with UTF-16 encoding.
TeamGuid	Edm.String	No	A unique identifier of the team the message belongs to.

Parameters	Type	Mandatory?	Description
TeamName	Edm.String	No	The name of the team the message belongs to.
Version	Edm.String	No	The version of the chat or channel message.

Microsoft Teams user concern schema

Microsoft Teams user concern events returned in audit log searches use this schema. These events have `RecordType` set to `MicrosoftTeamsUserConcern`(462), and apply to the following operations: `MessageReported`, `CallReported`, and `UserReported`.

Parameters	Type	Mandatory?	Description
SubmissionId	Edm.String	Yes	A unique identifier of the user concern report submission.
ReportMetadata	Edm.String	Yes	A JSON-encoded string containing metadata about the report, including the report reason and details about the reported entity (message, call, or user).

Microsoft Defender for Office 365 and Threat Investigation and Response schema

Microsoft Defender for Office 365 events are available for Microsoft 365 customers who have Defender for Office 365, either included or as an add-on subscription. For example, Microsoft 365 Business Premium includes Defender for Office 365 Plan 1, and Microsoft 365 A5/E5/G5 includes Defender for Office 365 Plan 2.

Threat Investigation and Response events are available only to customers with Defender for Office 365 Plan 2.

Each event in the Defender for Office 365 feed corresponds to the following features:

- Delivered email messages detected and acted on by zero-hour auto purge (ZAP).
- URLs detected at time-of-click by Safe Links.
- Files detected by Safe Attachments for SharePoint, OneDrive, and Microsoft Teams.
- Alerts that triggered automated investigations (Defender for Office 365 Plan 2 only).
- Attack simulation training events (Defender for Office 365 Plan 2 only).

Email message events

Parameters	Type	Mandatory?	Description
AttachmentData	Collection(Self.AttachmentData)	No	Data about attachments in the email message that triggered the event.

Parameters	Type	Mandatory?	Description
DetectionType	Edm.String	Yes	The type of detection. For example • **Inline** • Detected at delivery. • **Delayed**: Detected after delivery • **ZAP**: Removed after delivery by zero-hour auto purge (ZAP). Events with ZAP detection type are typically preceded by a message with a **Delayed** detection type.
DetectionMethod	Edm.String	Yes	The detection method or technology used by Defender for Office 365.
InternetMessageId	Edm.String	Yes	The Internet Message Id.
NetworkMessageId	Edm.String	Yes	The Exchange Online Network Message Id.
P1Sender	Edm.String	Yes	The MAIL FROM address used in SMTP transmission of the email message between email servers (also known as the `5321.MailFrom` address, P1 sender, or envelope sender).
P2Sender	Edm.String	Yes	The From address (also known as the `5322.From` address or P2 sender) that's shown in email clients.
Policy	Self.Policy	Yes	The type of threat policy (for example, **Anti-spam** or **Anti-phish**) and related action type (for example, **High Confidence Spam**, **Spam**, or **Phish**) relevant to the email message.
Policy	Self.PolicyAction	Yes	The action configured in the threat policy (for example, **Move to Junk Mail folder** or **Quarantine**)

Parameters	Type	Mandatory?	Description
			relevant to the email message.
Recipients	Collection(Edm.String)	Yes	An array of recipients of the email message.
SenderIp	Edm.String	Yes	The IPv4 or IPv6 address that submitted the message.
Subject	Edm.String	Yes	The subject line of the message.
Verdict	Edm.String	Yes	The message verdict.
MessageTime	Edm.Date	Yes	Date and time in Coordinated Universal Time (UTC) the email message was received or sent.
EventDeepLink	Edm.String	Yes	Deep-link to the email event in Threat Explorer or Real-time detections.
Delivery Action	Edm.String	Yes	The original delivery action on the email message.
Original Delivery location	Edm.String	Yes	The original delivery location of the email message.
Latest Delivery location	Edm.String	Yes	The latest delivery location of the email message at the time of the event.
Directionality	Edm.String	Yes	Identifies whether an email message was inbound, outbound, or an intra-org message.
ThreatsAndDetectionTech	Edm.String	Yes	The threats and the corresponding detection technologies. This field exposes all the threats on an email message, including the latest addition on spam verdict. For example, `["Phish: [Spoof DMARC]","Spam: [URL malicious reputation]"]`. The different detection threat

Parameters	Type	Mandatory?	Description
			and detection technologies are described in Detection technologies.
AdditionalActionsAndResults	Collection(Edm.String)	No	The additional actions that were taken on the email, such as ZAP or Manual Remediation. Also includes the corresponding results.
Connectors	Edm.String	No	The names and GUIDs of the connectors associated with the email.
AuthDetails	Collection(Self.AuthDetails)	No	The authentication checks that are done for the email. Also includes the values for SPF, DKIM, DMARC, and CompAuth.
SystemOverrides	Collection(Self.SystemOverrides)	No	Overrides that are applicable to the email. These can be system or user overrides.
Phish Confidence Level	Edm.String	No	Indicates the confidence level associated with Phish verdict. It can be Normal or High.

> **Note**
>
> We recommend that you use the new ThreatsAndDetectionTech field because it shows multiple verdicts and the updated detection technologies. This field also aligns with the values you see within other experiences like Threat Explorer and Advanced Hunting.

Detection technologies

Name	Description
Advanced filter	Machine learning models to detect phishing and spam.
Antimalware protection	Detection from signature-based anti-malware.
Bulk	Detection for advertising/marketing and similar message types with their relative bulk complaint levels (BCL).
Campaign	Messages identified and grouped as part of a malware or phishing campaign.

Name	Description
Domain reputation	The message was sent from a domain that was identified as spam or phishing domain, based on internal or external signals.
File detonation	Safe Attachments detected a malicious attachment during detonation within a sandbox.
File detonation reputation	File attachments previously detected by Safe Attachments during detonation.
File reputation	The message contains a file that was previously identified as malicious by other sources.
Fingerprint matching	The message resembles a previously detected malicious or spam message.
General filter	Phishing or spam signals based on analyst heuristics.
Impersonation brand	Sender impersonation of well-known brands.
Impersonation domain	Impersonation of sender domains that you own or specified for protection in anti-phishing policies.
Impersonation user	Impersonation of protected senders that you specified in anti-phishing policies.
IP reputation	The message was sent from an IP that was identified as potentially malicious.
LLM content analysis	Analysis by Microsoft's purpose-built large language models to detect harmful email.
Mail bombing	A distributed denial of service (DDoS) attack that typically subscribes recipients to a large number of legitimate newsletters and services. The resulting volume of incoming email within minutes intends to overwhelm the recipient's mailbox and email security systems, and acts as a precursor to malware, ransomware, or data exfiltration.
Mailbox intelligence impersonation	Sender detected as impersonating an address in the user's personal sender map.
Mixed analysis detection	Multiple filters contributed to the verdict for this message.
Spoof DMARC	The message failed DMARC authentication.
Spoof external domain	Spoof intelligence detected email spoofing of a domain that is external to your organization.
Spoof intra-org	Spoof intelligence detected email spoofing of a user or domain that is internal to your organization.
URL detonation	Safe Links detected a malicious URL in the message during detonation within a sandbox.
URL detonation	URLs previously detected by Safe Links during detonation.

Name	Description
reputation	
URL malicious reputation	The message contains a URL that was previously identified as malicious or spam by other sources.

AttachmentData complex type

AttachmentData

Parameters	Type	Mandatory?	Description
FileName	Edm.String	Yes	The file name of the attachment.
FileType	Edm.String	Yes	The file type of the attachment.
FileVerdict	Self.FileVerdict	Yes	The file malware verdict.
MalwareFamily	Edm.String	No	The file malware family.
SHA256	Edm.String	Yes	The file SHA256 hash.

> **Note**
>
> Within the Malware family, you see the exact MalwareFamily name (for example, HTML/Phish.VS!MSR) or Malicious Payload as a static string. A Malicious Payload should still be treated as malicious email when a specific name isn't identified.

SystemOverrides complex type

SystemOverrides

Parameters	Type	Mandatory?	Description
Details	Edm.String	No	The details about the specific override (such as ETR or Safe Sender) that was applied.
FinalOverride	Edm.String	No	Indicates the override that impacted the delivery in the case of multiple overrides.
Result	Edm.String	No	Indicates whether the email was set to allowed or blocked based on the override.
Source	Edm.String	No	Indicates whether the override was user-configured or tenant-configured.

AuthDetails complex type

AuthDetails

Parameters	Type	Mandatory?	Description
Name	Edm.String	No	The name of the specific auth check, such as DKIM or DMARC.
Value	Edm.String	No	The value associated with the specific auth check, such as True or False.

Enum: FileVerdict - Type: Edm.Int32

FileVerdict

Value	Member name	Description
0	Good	No threats detected.
1	Bad	Malware found in attachment.
-1	Error	Scan / analysis error.
-2	Timeout	Scan / analysis timeout.
-3	Pending	Scan / analysis not complete.

Enum: Policy - Type: Edm.Int32

Policy type and action type

Value	Member name	Description
1	Anti-spam, HSPM	High Confidence Spam (HSPM) action in the anti-spam policy.
2	Anti-spam, SPM	Spam (SPM) action in the anti-spam policy.
3	Anti-spam, Bulk	Bulk action in the anti-spam policy.
4	Anti-spam, PHSH	Phish (PHSH) action in the anti-spam policy.
5	Anti-phish, DIMP	Domain Impersonation (DIMP) action in the anti-phishing policy.
6	Anti-phish, UIMP	User Impersonation (UIMP) action in the anti-phishing policy.
7	Anti-phish, SPOOF	Spoof action in the anti-phishing policy.
8	Anti-phish, GIMP	Mailbox intelligence action in the anti-phishing policy.
9	Anti-malware, AMP	Malware policy action in the anti-malware policy.
10	Safe attachment, SAP	Policy action in the Safe Attachments policy.
11	Exchange transport rule, ETR	Policy action in the Exchange mail flow rule (also known as transport rule).

Value	Member name	Description
12	Anti-malware, ZAPM	Malware policy action in the anti-malware policy applied to zero-hour auto purge (ZAP).
13	Anti-phish, ZAPP	Phishing policy action in the anti-phishing policy applied to ZAP.
14	Anti-phish, ZAPS	Spam policy action in the anti-spam policy applied to ZAP.
15	Anti-spam, High confidence phish email (HPHISH)	High confidence Phish policy action in anti-spam policy.
17	Anti-spam, Outbound spam policy (OSPM)	Policy action in the outbound spam policy.

Enum: PolicyAction - Type: Edm.Int32

Policy action

Value	Member name	Description
0	MoveToJMF	Policy action is to move the message to the Junk Email folder.
1	AddXHeader	Policy action is to add an X-header to the email message.
2	ModifySubject	Policy action is to modify subject in the email message with information specified by the threat policy.
3	Redirect	Policy action is to redirect email message to email address specified by the threat policy.
4	Delete	Policy action is to delete (drop) the email message.
5	Quarantine	Policy action is to quarantine the email message.
6	NoAction	Policy is configured to take no action on the email message.
7	BccMessage	Policy action is to Bcc the email message to email address specified by the threat policy.
8	ReplaceAttachment	Policy action is to replace the attachment in the email message as specified by the threat policy.

URL time-of-click events

Parameters	Type	Mandatory?	Description
UserId	Edm.String	Yes	Identifier (for example, email address) of the user who clicked on the URL.
AppName	Edm.String	Yes	Office 365 service from which the URL was clicked (for example, Mail).

Parameters	Type	Mandatory?	Description
URLClickAction	Self.URLClickAction	Yes	Click action for the URL based on the organization's policies for Safe Links.
SourceId	Edm.String	Yes	Identifier for the Office 365 service from which the URL was clicked (for example, for email, this value is the Exchange Online Network Message Id).
TimeOfClick	Edm.Date	Yes	The date and time in Coordinated Universal Time (UTC) when the user clicked the URL.
URL	Edm.String	Yes	URL clicked by the user.
UserIp	Edm.String	Yes	The IPv4 or IPv6 address of the user who clicked the URL.

Enum: URLClickAction - Type: Edm.Int32

URLClickAction

Value	Member name	Description
2	Blockpage	User blocked from navigating to the URL by Safe Links.
3	PendingDetonationPage	User presented with the detonation pending page by Safe Links.
4	BlockPageOverride	User blocked from navigating to the URL by Safe Links, but the user overrode the block to navigate to the URL.
5	PendingDetonationPageOverride	User presented with the detonation page by Safe Links, but the user overrode the page to navigate to the URL.

File events

Parameters	Type	Mandatory?	Description
FileData	Self.FileData	Yes	Data about the file that triggered the event.
SourceWorkload	Self.SourceWorkload	Yes	Workload or service where the file was found (for example, SharePoint, OneDrive, or Microsoft Teams)
DetectionMethod	Edm.String	Yes	The method or technology used by Microsoft Defender for Office 365 for the detection.
LastModifiedDate	Edm.Date	Yes	The date and time in Coordinated Universal Time (UTC) when the file was created or last modified.
LastModifiedBy	Edm.String	Yes	Identifier (for example, an email address) for the user who created or last modified the file.
EventDeepLink	Edm.String	Yes	Deep-link to the file event in Threat Explorer or Real-time detections.

FileData complex type

FileData

Parameters	Type	Mandatory?	Description
DocumentId	Edm.String	Yes	Unique identifier for the file in SharePoint, OneDrive, or Microsoft Teams.
FileName	Edm.String	Yes	Name of the file that triggered the event.
FilePath	Edm.String	Yes	Path (location) for the file in SharePoint, OneDrive, or Microsoft Teams.
FileVerdict	Self.FileVerdict	Yes	The file malware verdict.
MalwareFamily	Edm.String	No	The file malware family.
SHA256	Edm.String	Yes	The file SHA256 hash.
FileSize	Edm.String	Yes	Size for the file in bytes.

Enum: SourceWorkload - Type: Edm.Int32

SourceWorkload

Value	Member name
0	SharePoint
1	OneDrive
2	Microsoft Teams

Attack Sim schema in Microsoft Defender for Office 365 Plan 2

Attack simulation training events are available for Microsoft 365 customers who have Defender for Office 365 Plan 2, either included or as an add-on subscription. For example Microsoft 365 A5/E5/G5 includes Defender for Office 365 Plan 2.

Parameters	Type	Mandatory?	Description
Batch ID	Edm.String	Yes	A unique identifier for a group of records that are processed together.
Campaign ID	Edm.String	Yes	A unique identifier for the attack simulation training campaign.

Parameters	Type	Mandatory?	Description
UserDisplayName	Edm.String	No	The display name of the user involved in the attack simulation training campaign.
AttackTechnique	Edm.String	No	The attack technique used in the simulation.
CampaignType	Edm.String	No	The type of attack simulation campaign. The different types are simulation campaign, training campaign, simulation automation.
TimeData	Edm.Date	No	Campaign launch time.
EndTimeData	Edm.Date	No	Campaign end time.
AttackSimEvent	Self.AttackSimEventType	Yes	Type of event (user activity or message delivery state).
ExtendedProperties	Collection(Common.NameValuePair)	Yes	Additional properties associated with the audit record.

Enum: AttackSimEventType - Type: Edm.Int32

Value	Member Name	Description
1	MessageDelivered	Message delivered to recipient successfully.
2	MessageFailed	Message delivery failed.
6	CredentialEntered	User entered credentials in the phish simulation for techniques such as Credential Harvesting, Link in attachment.
7	PermissionGranted	User granted permission in the phish simulation for technique such as Oauth consent grant.
8	LinkClicked	User clicked on phish simulation URL.
10	AttachmentOpened	Attachment opened by user for techniques such as Malware attachment, Link in attachment.
12	MessageRead	Message opened and read by recipient.
13	MessageReplied	Recipient replied to the message.
14	MessageForwarded	Message forwarded to another user.
15	MessageReported	Message reported as phish by recipient.
16	MessageDeleted	Message deleted by recipient.
17	OutOfOffice	Automatic replies in Outlook enabled for recipient.

Value	Member Name	Description
18	PositiveReinforcementMessageDelivered	Positive reinforcement message delivered successfully to recipient.

Attack Sim Admin schema in Microsoft Defender for Office 365 Plan 2

Attack simulation training admin events are available for Microsoft 365 customers who have Defender for Office 365 Plan 2, either included or as an add-on subscription. For example Microsoft 365 A5/E5/G5 includes Defender for Office 365 Plan 2.

Parameters	Type	Mandatory?	Description
Batch ID	Edm.String	Yes	A unique identifier for a group of records that are processed together.
Campaign ID	Edm.String	Yes	A unique identifier for an attack simulation training campaign.
AttackTechnique	Edm.String	No	The attack technique used in a simulation.
CampaignType	Edm.String	No	The type of attack simulation campaign. The different types are simulation campaign, training campaign, simulation automation.
TimeData	Edm.Date	Yes	Campaign launch time.
EndTimeData	Edm.Date	Yes	Campaign end time.
AttackSimAdminEvent	Self.AttackSimAdminEventType	Yes	The type of attack sim admin event.
ExtendedProperties	Collection(Common.NameValuePair)	Yes	Additional properties associated with the audit record.

Enum: AttackSimAdminEventType - Type: Edm.Int32

Value	Member Name	Description
0	CampaignLaunched	When an attack simulation training campaign was launched. The different types of campaigns are simulation, training campaign.
1	CampaignCompleted	When an attack simulation training campaign was completed.
2	CampaignReportDownloaded	Admin downloaded a campaign report.

Value	Member Name	Description
3	CampaignReportViewed	Admin viewed a campaign report.
4	CampaignCancelled	When an attack simulation training campaign was cancelled.
5	CampaignScheduled	When an attack simulation training campaign was scheduled.
6	CampaignDeleted	When an attack simulation training campaign was deleted.
7	SimulationExcluded	When a simulation was excluded from reporting.
8	AutomationSubmitted	When an automation was submitted, includes simulation and payload automation.
9	AutomationTurnOn	When an automation was turned on, includes simulation and payload automation.
10	AutomationTurnOff	When an automation was turned off, includes simulation and payload automation.
11	AutomationCompleted	When a simulation automation was completed.
12	AutomationDeleted	When an automation was deleted, includes simulation and payload automation.

User Training schema

Parameters	Type	Mandatory?	Description
Batch ID	Edm.String	Yes	A unique identifier for a group of records that are processed together.
Campaign ID	Edm.String	Yes	A unique identifier for an attack simulation training campaign.
Course ID	Edm.String	Yes	A unique identifier for a training course.
UserDisplayName	Edm.String	No	The display name of the user involved in the attack simulation.
CampaignType	Edm.String	Yes	The type of attack simulation campaign. The different types are simulation campaign, training campaign, simulation automation.
TimeData	Edm.Date	Yes	Campaign launch time.
EndTimeData	Edm.Date	Yes	Campaign end time.
UserTrainingEvent	Self.UserTrainingEventType	Yes	The type of user training event.

Parameters	Type	Mandatory?	Description
ExtendedProperties	Collection(Common.NameValuePair)	Yes	Additional properties associated with the audit record.

Enum: UserTrainingEventType - Type: Edm.Int32

Value	Member Name	Description
1	TrainingAssigned	Training assigned to user.
2	TrainingUpdated	Training updated for user.
3	TrainingCompleted	User completed training.
4	TrainingPreviouslyAssigned	Previously assigned training to user.
5	TrainingNotCompleted	User did not complete assigned training.

Submission schema

Events for submitting false positives or false negatives to Microsoft for analysis are available in all organizations with mailboxes in Exchange Online. Each event in the submission feed corresponds to false positives or false negatives that were submitted by:

- Admins: Messages, files, or URLs submitted to Microsoft for analysis.
- Users: Messages reported by users to admins or Microsoft for review.

Submission events

Parameters	Type	Mandatory?	Description
AdminSubmissionRegistered	Edm.String	No	Admin submission is registered and is pending for processing.
AdminSubmissionDeliveryCheck	Edm.String	No	Admin submission system checked the email's policy.
AdminSubmissionSubmitting	Edm.String	No	Admin submission system is submitting the email.
AdminSubmissionSubmitted	Edm.String	No	Admin submission system submitted the email.
AdminSubmissionTriage	Edm.String	No	Admin submission triaged by grader.
AdminSubmissionTimeout	Edm.String	No	Admin submission timed out with no result.
UserSubmission	Edm.String	No	Submission was first reported by an end user.
UserSubmissionTriage	Edm.String	No	User submission is triaged by grader.

Parameters	Type	Mandatory?	Description
CustomSubmission	Edm.String	No	Message reported by a user was sent to the organization's custom mailbox as set in the user reported messages settings.
AttackSimUserSubmission	Edm.String	No	The user-reported message was actually a phish simulation training message.
AdminSubmissionTablAllow	Edm.String	No	An allow entry in the Tenant Allow/Block List was created at time of submission to immediately take action on similar messages while it is being rescanned.
SubmissionNotification	Edm.String	No	Admin feedback is sent to end user.

Automated investigation and response events in Microsoft Defender for Office 365 Plan 2

Automated investigation and response (AIR) events are available for Microsoft 365 customers who have Defender for Office 365 Plan 2, either included or as an add-on subscription. For example Microsoft 365 A5/E5/G5 includes Defender for Office 365 Plan 2.

Investigation events are logged based on a change in investigation status. For example, when an admin takes an action that changes the status of an investigation from Pending Actions to Completed, an event is logged.

Currently, only automated investigation are logged. The following status values are logged:

- Investigation Started
- No threats found
- Terminated by System
- Pending Action
- Threats Found
- Remediated
- Failed
- Terminated by throttling
- Terminated By User
- Running

Main investigation schema

Name	Type	Description
InvestigationId	Edm.String	Investigation ID/GUID.
InvestigationName	Edm.String	Name of the investigation.

Name	Type	Description
InvestigationType	Edm.String	Type of the investigation. Can take one of the following values: • User-Reported Messages • Zapped Malware • Zapped Phish • Url Verdict Change (Currently, manual investigations aren't available.)
LastUpdateTimeUtc	Edm.Date	UTC time of the last update for an investigation.
StartTimeUtc	Edm.Date	Start time for an investigation.
Status	Edm.String	State of investigation, Running, Pending Actions, etc.
DeeplinkURL	Edm.String	Deep link URL to an investigation in the Microsoft Defender portal.
Actions	Collection (Edm.String)	Collection of actions recommended by an investigation.
Data	Edm.String	Data string which contains more details about investigation entities, and information about alerts related to the investigation. Entities are available in a separate node within the data blob.

Actions

Field	Type	Description
ID	Edm.String	Action ID
ActionType	Edm.String	The type of the action, such as email remediation.
ActionStatus	Edm.String	Values include: • Pending • Running • Waiting on resource • Completed • Failed
ApprovedBy	Edm.String	Null if auto approved; otherwise, the username/id (this is coming soon)
TimestampUtc	Edm.DateTime	The timestamp of the action status change.
ActionId	Edm.String	Unique identifier for action.
InvestigationId	Edm.String	Unique identifier for investigation.
RelatedAlertIds	Collection(Edm.String)	Alerts related to an investigation.
StartTimeUtc	Edm.DateTime	Timestamp of action creation.
EndTimeUtc	Edm.DateTime	Action final status update timestamp.

Field	Type	Description
Resource Identifiers	Edm.String	Consists of the Azure Active Directory tenant ID.
Entities	Collection(Edm.String)	List of one or more affected entities by action.
Related Alert IDs	Edm.String	Alert related to an investigation.

Entities

MailMessage (email)

Field	Type	Description
Type	Edm.String	"mail-message"
Files	Collection (Self.File)	Details about the files of this message's attachments.
Recipient	Edm.String	The recipient of this mail message.
Urls	Collection(Self.URL)	The Urls contained in this mail message.
Sender	Edm.String	The sender's email address.
SenderIP	Edm.String	The sender's IP address.
ReceivedDate	Edm.DateTime	The received date of this message.
NetworkMessageId	Edm.Guid	The network message id of this mail message.
InternetMessageId	Edm.String	The internet message id of this mail message.
Subject	Edm.String	The subject of this mail message.

IP

Field	Type	Description
Type	Edm.String	"ip"
Address	Edm.String	The IP address as a string, such as `127.0.0.1`.

URL

Field	Type	Description
Type	Edm.String	"url"
Url	Edm.String	The full URL to which an entity points.

Mailbox (also equivalent to the user)

Field	Type	Description
Type	Edm.String	"mailbox"
MailboxPrimaryAddress	Edm.String	The mailbox's primary address.
DisplayName	Edm.String	The mailbox's display name.
Upn	Edm.String	The mailbox's UPN.

File

Field	Type	Description
Type	Edm.String	"file"
Name	Edm.String	The file name without path.
FileHashes	Collection (Edm.String)	The file hashes associated with the file.

FileHash

Field	Type	Description
Type	Edm.String	"filehash"
Algorithm	Edm.String	The hash algorithm type, which can be one of these values: • Unknown • MD5 • SHA1 • SHA256 • SHA256AC
Value	Edm.String	The hash value.

MailCluster

Field	Type	Description
Type	Edm.String	"MailCluster" Determines the type of entity being discussed.
NetworkMessageIds	Collection (Edm.String)	List of the mail message IDs that are part of the mail cluster.
CountByDeliveryStatus	Collections (Edm.String)	Count of mail messages by DeliveryStatus string representation.
CountByThreatType	Collections (Edm.String)	Count of mail messages by ThreatType string representation.
Threats	Collections	The threats of mail messages that are part of the mail cluster.

Field	Type	Description
	(Edm.String)	Threats include values like Phish and Malware.
Query	Edm.String	The query that was used to identify the messages of the mail cluster.
QueryTime	Edm.DateTime	The query time.
MailCount	Edm.int	The number of mail messages that are part of the mail cluster.
Source	String	The source of the mail cluster; the value of the cluster source.

Hygiene events schema

Hygiene events are related to outbound spam protection. These events are related to users who are restricted from sending email. For more information, see the following articles:

- Outbound spam protection
- Remove blocked users from the Restricted entities page

Parameters	Type	Mandatory?	Description
Audit	Edm.String	No	System information related to the hygiene event.
Event	Edm.String	No	The type of hygiene event. The values for this parameter are **Listed** or **Delisted**.
EventId	Edm.Int64	No	The ID of the hygiene event type.
EventValue	Edm.String	No	The user who was impacted.
Reason	Edm.String	No	Details about the hygiene event.

Power BI schema

The Power BI events listed in Search the audit log use this schema.

Parameters	Type	Mandatory?	Description
AppName	Edm.String Term="Microsoft.Office.Audit.Schema.PIIFlag" Bool="true"	No	The name of the app where the event occurred.
DashboardName	Edm.String Term="Microsoft.Office.Audit.Schema.PIIFlag" Bool="true"	No	The name of the dashboard where the event occurred.
DataClassification	Edm.String Term="Microsoft.Office.Audit.Schema.PIIFlag" Bool="true"	No	The data classification, if any, for the dashboard

Parameters	Type	Mandatory?	Description
			where the event occurred.
DatasetName	Edm.String Term="Microsoft.Office.Audit.Schema.PIIFlag" Bool="true"	No	The name of the dataset where the event occurred.
ActivityStatus	Edm.String Term="Microsoft.Office.Audit.Schema.PIIFlag" Bool="true"	No	The status of the Gen1 dataflow migration audit event. Possible values include `Started`, `Succeeded`, and `Failed`.
MigrationTriggerType	Edm.String Term="Microsoft.Office.Audit.Schema.PIIFlag" Bool="true"	No	The source that triggered the Gen1 dataflow migration, such as the Upgrade Wizard or another migration entry point.
MembershipInformation	Collection(MembershipInformationType) Term="Microsoft.Office.Audit.Schema.PIIFlag" Bool="true"	No	Membership information about the group.
OrgAppPermission	Edm.String Term="Microsoft.Office.Audit.Schema.PIIFlag" Bool="true"	No	Permissions list for an organizational app (entire organization, specific users, or specific groups).
ReportName	Edm.String Term="Microsoft.Office.Audit.Schema.PIIFlag" Bool="true"	No	The name of the report where the event occurred.
SharingInformation	Collection(SharingInformationType) Term="Microsoft.Office.Audit.Schema.PIIFlag" Bool="true"	No	Information about the person to whom a sharing invitation is sent.
SwitchState	Edm.String Term="Microsoft.Office.Audit.Schema.PIIFlag" Bool="true"	No	Information about the state of various

Parameters	Type	Mandatory?	Description
			tenant level switches.
WorkSpaceName	Edm.String Term="Microsoft.Office.Audit.Schema.PIIFlag" Bool="true"	No	The name of the workspace where the event occurred.

MembershipInformationType complex type

Parameters	Type	Mandatory?	Description
MemberEmail	Edm.String Term="Microsoft.Office.Audit.Schema.PIIFlag" Bool="true"	No	The email address of the group.
Status	Edm.String Term="Microsoft.Office.Audit.Schema.PIIFlag" Bool="true"	No	Not currently populated.

SharingInformationType complex type

Parameters	Type	Mandatory?	Description
RecipientEmail	Edm.String Term="Microsoft.Office.Audit.Schema.PIIFlag" Bool="true"	No	The email address of the recipient of a sharing invitation.
RecipientName	Edm.String Term="Microsoft.Office.Audit.Schema.PIIFlag" Bool="true"	No	The name of the recipient of a sharing invitation.
ResharePermission	Edm.String Term="Microsoft.Office.Audit.Schema.PIIFlag" Bool="true"	No	The permission being granted to the recipient.

Dynamics 365 schema

The audit records for events related to model-driven apps in Dynamics 365 events use both a base and an entity operation schema. For more information, see Enable and use Activity Logging.

Dynamics 365 base schema

Parameters	Type	Mandatory?	Description
CrmOrganizationUniqueName	Edm.String	Yes	The unique name of the organization.

Parameters	Type	Mandatory?	Description
InstanceUrl	Edm.String	Yes	The URL to the instance.
ItemUrl	Edm.String	No	The URL to the record emitting the log.
ItemType	Edm.String	No	The name of the entity.
UserAgent	Edm.String	No	The unique identifier of the user GUID in the organization.
Fields	Collection(Common.NameValuePair)	No	A JSON object that contains the property key-value pairs that were created or updated.

Dynamics 365 entity operation schema

Entity events from model-driven apps in Dynamics 365 use this schema to build on the Dynamics 365 base schema. This schema includes information about the entity operation that triggered the audited event.

Parameters	Type	Mandatory?	Description
EntityId	Edm.Guid	No	The unique identifier of the entity.
EntityName	Edm.String	Yes	The name of the entity in the organization. Example of entities include `contact` or `authentication`.
Message	Edm.String	Yes	This parameter contains the operation that was performed in related to the entity. For example, if a new contact was created, the value of the Message property is `Create` and the corresponding value of the EntityName property is `contact`.
Query	Edm.String	No	The parameters of the filter query that was used while executing the FetchXML operation.
PrimaryFieldValue	Edm.String	No	Indicates the value for the attribute that is the primary field for the entity.

eSignature schema

This schema includes eSignature for Microsoft 365 events returned in audit log searches.

Parameters	Type	Mandatory?	Description
AgreementId	Edm.Guid	Yes	ID of the electronic signature request.
AgreementSender	Edm.String	Yes	Sender of the electronic signature request.
AgreementCreationDateTime	Edm.Date	Yes	Date & time request was sent.
AgreementStatus	Self.AgreementStatusType	Yes	Status of the request.
AgreementRecipients	Collection(Self.AgreementRecipient)	Yes	Recipients of the request.

Viva Insights schema

Viva Insights events returned in audit log searches use this schema.

Parameters	Type	Mandatory?	Description
WpaUserRole	Edm.String	No	The Viva Insights role of the user who performed the action.
ModifiedProperties	Collection (Common.ModifiedProperty)	No	This property includes the name of the property that was modified, the new value of the modified property, and the previous value of the modified property.
OperationDetails	Collection (Common.NameValuePair)	No	A list of extended properties for the setting that was changed. Each property has a **Name** and **Value**.

Quarantine schema

Quarantine events returned in audit log searches use this schema. For more information about quarantine, see Quarantined email messages in cloud organizations.

Parameters	Type	Mandatory?	Description
RequestType	Self.RequestType	No	The type of quarantine request performed by a user.
RequestSource	Self.RequestSource	No	Quarantine request come from the Microsoft Defender portal, a cmdlet, or a URLlink.
NetworkMessageId	Edm.String	No	The network message id of quarantined email message.
ReleaseTo	Edm.String	No	The recipient of the email message.

Enum: RequestType - Type: Edm.Int32

Value	Member name	Description
0	Preview	User request to preview an email message identified as harmful.
1	Delete	User request to delete an email message identified as harmful.
2	Release	User request to release an email message identified as harmful.
3	Export	User request to export an email message identified as harmful.
4	ViewHeader	User request to view the header of an email message identified as harmful.
5	Release request	User request to release an email message identified as harmful.

Enum: RequestSource - Type: Edm.Int32

The source of a user request to preview, delete, release, export, or view the header of a potentially harmful email message.

Value	Member name	Description
0	SCC	The Microsoft Defender portal.
1	Cmdlet	A cmdlet.
2	URLlink	A link.

Microsoft Forms schema

Microsoft Forms events returned in audit log searches use this schema.

Parameters	Type	Mandatory?	Description
FormsUserTypes	Collection(Self.FormsUserTypes)	Yes	The role of the user who performed the action. The values for this parameter are Admin, Owner, Responder, or Coauthor.
SourceApp	Edm.String	Yes	Indicates if the action is from Forms website or from another App.
FormName	Edm.String	No	The name of the current form.
FormId	Edm.String	No	The Id of the target form.
FormTypes	Collection(Self.FormTypes)	No	Indicates whether this is a Form, Quiz, or Survey.
ActivityParameters	Edm.String	No	JSON string containing activity parameters. For more information, see Microsoft Forms activities.

Enum: FormsUserTypes - Type: Edm.Int32

FormsUserTypes

Value	Form User Type	Description
0	Admin	An administrator who has access to the form.
1	Owner	A user who is the owner of the form.
2	Responder	A user who has submitted a response to a form.
3	Coauthor	A user who has used a collaboration link provided by the form owner to login and edit a form.

Enum: FormTypes - Type: Edm.Int32

FormTypes

Value	Form Types	Description
0	Form	Forms that are created with the New Form option.
1	Quiz	Quizzes that are created with the New Quiz option. A quiz is a special type of form that includes additional features such as point values, auto and manual grading, and commenting.
2	Survey	Surveys that are created with the New Survey option. A survey is a special type of form that includes additional features such as CMS integration and support for Flow rules.

MIP label schema

Events in the Microsoft Purview Information Protection label schema are triggered when Microsoft 365 detects an email message processed by agents in the Transport pipeline that has a sensitivity label applied to it. The sensitivity label might have been applied manually or automatically, and it might have been applied within or outside of the Transport pipeline. Sensitivity labels can be automatically applied to email messages by auto-apply label policies.

The intent of this audit schema is to represent the sum of all email activity that involves sensitivity labels. In other words, there should be an recorded audit activity for each email message that is sent to or from users in the organization that has a sensitivity label applied to it, regardless of when or how the sensitivity label was applied. For more information about sensitivity labels, see:

- Learn about sensitivity labels
- Apply a sensitivity label to content automatically

Parameters	Type	Mandatory?	Description
Sender	Edm.String	No	The email address in the From field of the email message.
Receivers	Collection(Edm.String)	No	All email addresses in the To, CC, and Bcc fields of the email message.
ItemName	Edm.String	No	The string in the Subject field of the email message.
LabelId	Edm.Guid	No	The GUID of the sensitivity label applied to the email message.
LabelName	Edm.String	No	The name of the sensitivity label applied to the email message.
LabelAction	Edm.String	No	The actions specified by the sensitivity label that were applied to the email message before the message entered the mail transport pipeline.
LabelAppliedDateTime	Edm.Date	No	The date the sensitivity label was applied to the email message.
ApplicationMode	Edm.String	No	Specifies how the sensitivity label was applied to the email message. The **Privileged** value indicates the label was manually applied by a user. The **Standard** value indicates the label was auto-applied by a client-side or service-side labeling process.

Encrypted message portal events schema

Events for encrypted message portal schema are triggered when when Purview Message Encryption detects an encrypted email message is accessed through the portal by an external recipient. The mail might have been encrypted manually with a sensitivity label or an RMS template, or automatically by a transport rule, a Data Loss Prevention policy, or an auto-labeling policy.

The intent of this audit schema is to represent the sum of all portal activity that involves accessing the encrypted mail by external recipients. In other words, there should be a recorded audit activity for a recipient that attempts to sign in to the portal and any activities related to accessing the encrypted mail. This includes mail sent to or from users in the organization when the mail has encryption applied to it, regardless of when or how the encryption was applied. For more information, see, Learn about encrypted message portal logs.

Parameters	Type	Mandatory?	Description
MessageId	Edm.String	No	The Id of the message.

Parameters	Type	Mandatory?	Description
Recipient	Edm.String	No	Recipient email address.
Sender	Edm.String	No	Email address of sender.
AuthenticationMethod	Self.AuthenticationMethod	No	Authentication method when accessing the message, i.e. OTP, Yahoo, Gmail, Microsoft.
AuthenticationStatus	Self.AuthenticationStatus	No	0: Success, 1: Failure.
OperationStatus	Self.OperationStatus	No	0: Success, 1: Failure.
AttachmentName	Edm.String	No	Name of the attachment.
OperationProperties	Collection(Common.NameValuePair)	No	Extra properties, i.e. number of OTP passcode sent, email subject, etc.

Communication compliance Exchange schema

The communication compliance events listed in the Office 365 audit log use this schema. This includes audit records for the **SupervisoryReviewOLAudit** operation that's generated when email message content contains offensive language identified by anti-spam models with a match accuracy of >= 99.5%.

Parameters	Type	Mandatory?	Description
ExchangeDetails	ExchangeDetails	No	Properties of the email message that triggered the SupervisoryReviewOLAudit event.

Enum: ExchangeDetails - Type: ExchangeDetails

ExchangeDetails

Member name	Type	Description
NetworkMessageId	Edm.Guid	The network message ID of the email message.
InternetMessageId	Edm.String	The internet message ID of the email message.
AttachmentData	Collection(AttachmentDetails)	Information about files attached to the email message.
Recipients	Collection(Edm.String)	The email addresses in the To, Cc, and Bcc fields of the email message.
Subject	Edm.String	The text in the Subject field of the email message.
MessageTime	Edm.Date	The date and time the email message was sent.

Member name	Type	Description
From	Edm.String	The email address in the From field of the email message.
Directionality	Edm.String	The origination status of the email message.

Enum: AttachmentDetails - Type: Edm.Int32

AttachmentDetails

Member name	Type	Description
FileName	Edm.String	The name of the file attached to the email message.
FileType	Edm.String	The file extension of the file attached to the email message.
SHA256	Edm.String	The SHA-256 hash of the file attached to the email message.

Reports schema

Report events returned in audit log searches use this schema.

Parameters	Type	Mandatory?	Description
ModifiedProperties	Collection (Common.ModifiedProperty)	No	This property includes the name of the property that was modified, the new value of the modified property, and the previous value of the modified property.

Compliance connector schema

Events in the compliance connector schema are triggered when items that are imported by a data connector are skipped or failed to be import to user mailboxes. For more information about data connectors, see Learn about connectors for third-party data.

Parameters	Type	Mandatory?	Description
JobId	Edm.String	No	This is a unique identifier of the data connector.
TaskId	Edm.String	No	Unique identifier of the periodic data connector instance. Data connectors import data in periodic intervals.
JobType	Edm.String	No	The name of the data connector.
ItemId	Edm.String	No	Unique identifier of the item (for example, an email message) being imported.
ItemSize	Edm.Int64	No	The size of the item being imported.

Parameters	Type	Mandatory?	Description
SourceUserId	Edm.String	No	The unique identifier of the user from the third-party data source. For example, for a Slack data connector, this property specifies the user Id in Slack workspace.
FailureType	Self.FailureType	No	Indicates the type of data import failure. For example, the value **incorrectusermapping** indicates the item wasn't imported because no user mapping between the third-party data source and Microsoft 365 could be found.
ResultMessage	Edm.String	No	Indicates the type of failure, such as **Duplicate message**.
IsRetry	Edm.Boolean	No	Indicates whether the data connector retried to import the item.
Attachments	Collection.Attachment	No	A list of attachments received from the third-party data source.

Enum: FailureType - Type: Edm.Int32

Value	Member name
0	Default
1	MailboxWrite

Attachment complex type

Parameters	Type	Mandatory?	Description
FileName	Edm.String	No	The name of the attachment.
Details	Edm.String	No	Other details about the attachment.

SystemSync schema

Events in the SystemSync schema are triggered when the SystemSync ingested data is either exported via Data Lake or shared via other services.

DataLakeExportOperationAuditRecord

Parameters	Type	Mandatory?	Description
DataStoreType	DataStoreType	Yes	Indicates which data store the data was downloaded from. Refer DataStoreType for all possible values.

Parameters	Type	Mandatory?	Description
UserAction	DataLakeUserAction	Yes	Indicates what action user had performed on the data store. Refer DataLakeUserAction for all possible values.
ExportTriggeredAt	Edm.DateTimeOffset	Yes	Indicates when the data export was triggered.
NameOfDownloadedZipFile	Edm.String	No	The name of the compressed file the admin had downloaded from the Data Lake.

DataShareOperationAuditRecord

Parameters	Type	Mandatory?	Description
Invitation	DataShareInvitationType	No	Details of the invite sent to the recipient of the Data Share.

DataShareInvitationType complex type

Parameters	Type	Mandatory?	Description
ShareId	Edm.Guid	Yes	System assigned identifier for the Data Share.
Invitees	Collection(Edm.Guid)	Yes	List of admin users the invite was sent to.
InviteeTenantId	Edm.Guid	Yes	The target tenant whom the invite is intended to.
ShareName	Edm.String	Yes	System assigned name for the Data Share.
SyncFrequency	Self.SyncFrequency	Yes	Frequency at which the data is synced to the destination storage account once share is established. See SyncFrequency for possible values.
SyncStartTime	Edm.DateTimeOffset	Yes	Date and time of first sync.

Enum: SyncFrequency - Type: Edm.Int32:

Value	Member name	Description
0	Hourly	Indicates the data is synced every hour.
1	Daily	Indicates the data is synced once a day.

Enum: DataStoreType - Type: Edm.Int32:

Value	Member name	Description
0	CanonicalStore	Indicates data is downloaded from Canonical store.
1	StagingStore	Indicates data is downloaded from Staging store.

Enum: DataLakeUserAction - Type: Edm.Int32:

Value	Member name	Description
0	TriggerExport	The admin user triggered export from Data Lake.
1	DownloadZipFile	The admin user downloaded the exported data.

MicrosoftGraphDataConnectOperation complex type

Parameters	Type	Mandatory?	Description
ApplicationId	Edm.Guid	Yes	The application identification.
ApplicationName	Edm.String	Yes	The application name.
PipelineName	Edm.String	Yes	The pipeline name.
PipelineRunId	Edm.Guid	No	The identification of this pipeline run.
CopyActivityRunId	Edm.Guid	No	The identification of the ADF copy activity.
RunStartTime	Edm.Date	Yes	Date and time of the extraction.
RunEndTime	Edm.Date	Yes	Date and time of the extraction.
DatasetName	Edm.String	Yes	The dataset name being extracted.
DatasetColumns	Edm.String	Yes	The set of selected columns being extracted.
ScopeList	Edm.String	Yes	The scope of the extraction.
ScopeCountRequested	Edm.Int64	No	The requested scope count for this extraction.
ScopeCountDelivered	Edm.Int64	No	The delivered scope count for this extraction.
UndeliveredScope	Edm.String	No	The undelivered scope of the extraction.
RowCount	Edm.Int64	No	The number of rows extracted.
Status	Edm.String	Yes	The extraction status.
Reason	Edm.String	No	The error message in case of failure.

AipDiscover

The following table contains information related to Azure Information Protection (AIP) scanner events.

Event	Description
ApplicationId	The ID of the application performing the operation.
ApplicationName	Friendly name of the application performing the operation. Outlook (for email), OWA (for email), Word (for file), Excel (for file), PowerPoint (for file).

Event	Description
ClientIP	The IP address of the device that was used when the activity was logged. The IP address is displayed in either an IPv4 or IPv6 address format. For some services, the value displayed in this property might be the IP address for a trusted application (for example, Office on the web apps) calling into the service on behalf of a user and not the IP address of the device used by person who performed the activity. Also, for Azure Active Directory-related events, the IP address isn't logged and the value for the ClientIP property is null. The IP address is displayed in either an IPv4 or IPv6 address format.
CreationTime	The date and time in Coordinated Universal Time (UTC) in ISO8601 format when the audit log record was generated.
DataState	Describes the state of the data.
DeviceName	The device on which the activity happened.
Id	GUID of the current record.
IsProtected	Whether protected: True/False
Location	The location of the document with respect to the user's device. The possible values are **unknown**, **localMedia**, **removableMedia**, **fileshare**, and **cloud**.
ObjectId	File full path (URL). For SharePoint and OneDrive activity, the full path name of the file or folder accessed by the user.
Operation	Describes type of access.
OrganizationId	The GUID for your organization's Office 365 tenant. This value is always the same for your organization, regardless of the Office 365 service in which it occurs.
Platform	Device platform (Win, OSX, Android, iOS)
ProcessName	The relevant process name. For example, `Outlook`, `msip.app`, or `WinWord`.
ProductVersion	Version of the AIP client.
ProtectionOwner	Rights Management owner in UPN format.
ProtectionType	Protection type can be template or ad-hoc.
RecordType	The type of operation indicated by the record. See the AuditLogRecordType table for details on the types of audit log records. For a complete updated list and full description of the Log RecordType, see the Microsoft 365 Compliance audit log activities via Office 365 Management API blog post. Here we only list the relevant MIP Record types.
Scope	Event created by one of the following sources: • **online**: A service in Microsoft 365. • **onprem**: A service in an on-premises organization. Currently, SharePoint is the only workload sending events from on-premises to Microsoft 365.
SensitiveInfoTypeData	Stores the datatype of the Sensitive Info type data.

Event	Description
SensitivityLabelId	The current MIP sensitivity label GUID. Use the **Get-Label** to get the full values of the GUID.
TemplateId	TemplateID parameter to get a specific template. The Get-AipServiceTemplate cmdlet gets all existing or selected protection templates from Azure Information Protection.
UserId	The UPN (User Principal Name) of the user who performed the action (specified in the Operation property) that resulted in the record being logged; for example, `my_name@my_domain_name`. **Note**: Records for activity performed by system accounts (such as SHAREPOINT\system or NT AUTHORITY\SYSTEM) are also included. In SharePoint, another value display in the UserId property is `app@sharepoint`. This value indicates the "user" who performed the activity was an application that has the necessary permissions in SharePoint to perform organization-wide actions (such as search a SharePoint site or OneDrive account) on behalf of a user, admin, or service.
UserKey	An alternative ID for the user identified in the UserId property. This property is populated with the passport unique ID (PUID) for events performed by users in SharePoint, OneDrive, and Exchange.
UserType	The type of user that performed the operation. See the UserType table for details on the types of users. 0 = Regular 1 = Reserved 2 = Admin 3 = DcAdmin 4 = System 5 = Application 6 = ServicePrincipal 7 = CustomPolicy 8 = SystemPolicy
Version	Version ID of the file in the operation.
Workload	Stores the Office 365 service where the activity occurred.

AipSensitivityLabelAction

The following table contains information related to AIP sensitivity label events.

Event	Description
ApplicationId	Corresponds to the Microsoft Entra Application ID.
ApplicationName	Application friendly name of the application performing the operation.
CreationDate	The date and time in Coordinated Universal Time (UTC) in ISO8601 format when the user performed the activity.
DataState	Specifies the state of the data.
DeviceName	The name of the user's device.

Event	Description
Identity	The identity of the user or service to be authenticated.
IsProtected	Whether protected: True/False
IsProtectedBefore	Whether the content was protected before change: True/False
IsValid	Boolean
Location	The location of the document with respect to the user's device. The possible values are unknown, localMedia, removableMedia, fileshare, and cloud.
ObjectState	Specifies the state of the object.
Operation	The operation type for the audit log.The name of the user or admin activity. For a description of the most common operations/activities: SensitivityLabelApplied SensitivityLabelUpdated SensitivityLabelRemoved SensitivityLabelPolicyMatched SensitivityLabeledFileOpened.
Identity	The identity of the user or service to be authenticated.
PSComputerName	Computer Name
PSShowComputerName	The value is False for documented edited in Office 365.
Platform	Device platform (Win, OSX, Android, iOS).
ProcessName	Process that hosts MIP SDK.
ProductVersion	Version of the Azure Information Protection client that performed the audit action.
ProtectionType	Protection type can be template or ad-hoc.
RecordType	Shows the value of Label Action. The operation type indicated by the record. For more information, see the full list of record types.
RunspaceId	The Runspace is a specific instance of PowerShell which contains modifiable collections of commands, providers, variables, functions, and language elements that are available to the command line user.
SensitiveInfoTypeData	Stores the datatype of the Sensitive Info Type Data
TemplateId	TemplateID parameter to get a specific template. The Get-AipServiceTemplate cmdlet gets all existing or selected protection templates from Azure Information Protection.
UserId	The User Principal Name (UPN) of the user who performed the action (specified in the Operation property) that resulted in the record being logged; for example, `my_name@my_domain_name`. **Note**: Records for activity performed by system accounts (such as SHAREPOINT\system or NT AUTHORITY\SYSTEM) are also included. In SharePoint,

Event	Description
	another value display in the UserId property is `app@sharepoint`. This value indicates the "user" who performed the activity was an application that has the necessary permissions in SharePoint to perform organization-wide actions (such as search a SharePoint site or OneDrive account) on behalf of a user, admin, or service.

AipProtectionAction

Event	Description
PSComputerName	Computer name
RunspaceId	The Runspace is a specific instance of PowerShell which contains modifiable collections of commands, providers, variables, functions, and language elements that are available to the command line user.
PSShowComputerName	The value is false for a documented edited in Office 365.
RecordType	Shows the value of Label Action. The operation type indicated by the record. Here we are only listing the relevant MIP Record types. For more information, see the full list of record types.
CreationTime	The date and time in Coordinated Universal Time (UTC) in ISO8601 format when the audit log record was generated.
UserId	The User Principal Name (UPN) of the user who performed the action (specified in the Operation property) that resulted in the record being logged. For example, `my_name@my_domain_name`. **Note**: Records for activity performed by system accounts (such as SHAREPOINT\system or NT AUTHORITY\SYSTEM) are also included. In SharePoint, another value display in the UserId property is `app@sharepoint`. This value indicates the "user" who performed the activity was an application that has the necessary permissions in SharePoint to perform organization-wide actions (such as search a SharePoint site or OneDrive account) on behalf of a user, admin, or service.
Operation	The operation type for the audit log. The name of the user or admin activity. For a description of the most common operations/activities. SensitivityLabelApplied SensitivityLabelUpdated SensitivityLabelRemoved SensitivityLabelPolicyMatched SensitivityLabeledFileOpened.
Identity	The identity of the user or service to be authenticated.
ObjectState	State of the Object after the current event.
ApplicationId	The application where the activity happened and displayed in GUID.
ApplicationName	Application friendly name of the application performing the operation.Outlook (for email), OWA (for email), Word (for file), Excel (for file), PowerPoint (for file).

Event	Description
ProcessName	Process name of the Office application.
Platform	The platform on which the activity happened. For example, Windows.
DeviceName	Device the event was recorded on.
ProductVersion	Version of the Azure Information Protection client that performed the audit action.
UserId	The UPN of the user who performed the action (specified in the Operation property) that resulted in the record being logged; for example, `my_name@my_domain_name`. **Note**: Records for activity performed by system accounts (such as SHAREPOINT\system or NT AUTHORITY\SYSTEM) are also included. In SharePoint, another value display in the UserId property is `app@sharepoint`. This indicates the "user" who performed the activity was an application that has the necessary permissions in SharePoint to perform organization-wide actions (such as search a SharePoint site or OneDrive account) on behalf of a user, admin, or service.
ClientIP	The IP address of the device that was used when the activity was logged. The IP address is displayed in either an IPv4 or IPv6 address format. For some services, the value displayed in this property might be the IP address for a trusted application (for example, Office on the web apps) calling into the service on behalf of a user and not the IP address of the device used by person who performed the activity. Also, for Azure Active Directory-related events, the IP address isn't logged and the value for the ClientIP property is null. The IP address is displayed in either an IPv4 or IPv6 address format.
Id	GUID of the current record.
RecordType	Shows the value of Label Action. The operation type indicated by the record. Here we are only listing the relevant MIP Record types.
CreationTime	The date and time in Coordinated Universal Time (UTC) in ISO8601 format when the audit log record was generated.
Operation	The name of the user or admin activity. For a description of the most common operations/activities, see Search the audit log.
OrganizationId	The GUID for your organization's Office 365 tenant. This value is always the same for your organization, regardless of the Office 365 service in which it occurs.
UserType	The type of user that performed the operation. See the UserType table for details on the types of users. 0 = Regular 1 = Reserved 2 = Admin 3 = DcAdmin 4 = System 5 = Application 6 = ServicePrincipal 7 = CustomPolicy 8 = SystemPolicy

Event	Description
UserKey	An alternative ID for the user identified in the UserId property. This property is populated with the passport unique ID (PUID) for events performed by users in SharePoint, OneDrive, and Exchange.
Workload	Stores the Office 365 service where the activity occurred.
Version	Version of the Azure Information Protection client that performed the audit action
Scope	Specifies scope.

AipFileDeleted

Event	Description
PSComputerName	Computer name
RunspaceId	The Runspace is a specific instance of PowerShell which contains modifiable collections of commands, providers, variables, functions, and language elements that are available to the command line user.
PSShowComputerName	The value is false for a documented edited in Office 365.
RecordType	Shows the value of Label Action. The operation type indicated by the record. Here we are only listing the relevant MIP Record types. For more information, see the full list of record types.
CreationTime	The date and time in Coordinated Universal Time (UTC) in ISO8601 format when the audit log record was generated.
UserId	The User Principal Name (UPN) of the user who performed the action (specified in the Operation property) that resulted in the record being logged. For example, `my_name@my_domain_name`. **Note**: Records for activity performed by system accounts (such as SHAREPOINT\system or NT AUTHORITY\SYSTEM) are also included. In SharePoint, another value display in the UserId property is `app@sharepoint`. This value indicates the "user" who performed the activity was an application that has the necessary permissions in SharePoint to perform organization-wide actions (such as search a SharePoint site or OneDrive account) on behalf of a user, admin, or service.
Operation	The operation type for the audit log. The name of the user or admin activity. For a description of the most common operations/activities. SensitivityLabelApplied SensitivityLabelUpdated SensitivityLabelRemoved SensitivityLabelPolicyMatched SensitivityLabeledFileOpened.
Identity	The identity of the user or service to be authenticated.
ObjectState	State of the Object after the current event.

Event	Description
ApplicationId	The application where the activity happened and displayed in GUID.
ApplicationName	Application friendly name of the application performing the operation.Outlook (for email), OWA (for email), Word (for file), Excel (for file), PowerPoint (for file).
ProcessName	Process name of the Office application.
Platform	The platform on which the activity happened. For example, Windows.
DeviceName	Device the event was recorded on.
ProductVersion	Version of the Azure Information Protection client that performed the audit action.
UserId	The UPN of the user who performed the action (specified in the Operation property) that resulted in the record being logged; for example, `my_name@my_domain_name`. **Note**: Records for activity performed by system accounts (such as SHAREPOINT\system or NT AUTHORITY\SYSTEM) are also included. In SharePoint, another value display in the UserId property is `app@sharepoint`. This value indicates the "user" who performed the activity was an application that has the necessary permissions in SharePoint to perform organization-wide actions (such as search a SharePoint site or OneDrive account) on behalf of a user, admin, or service..
ClientIP	The IP address of the device that was used when the activity was logged. The IP address is displayed in either an IPv4 or IPv6 address format. For some services, the value displayed in this property might be the IP address for a trusted application (for example, Office on the web apps) calling into the service on behalf of a user and not the IP address of the device used by person who performed the activity. Also, for Azure Active Directory-related events, the IP address isn't logged and the value for the ClientIP property is null. The IP address is displayed in either an IPv4 or IPv6 address format.
Id	GUID of the current record.
RecordType	Shows the value of Label Action. The operation type indicated by the record. Here we are only listing the relevant MIP Record types.
CreationTime	The date and time in Coordinated Universal Time (UTC) in ISO8601 format when the audit log record was generated.
Operation	The name of the user or admin activity. For a description of the most common operations/activities, see Search the audit log.
OrganizationId	The GUID for your organization's Office 365 tenant. This value is always the same for your organization, regardless of the Office 365 service in which it occurs.
UserType	The type of user that performed the operation. See the UserType table for details on the types of users. 0 = Regular 1 = Reserved 2 = Admin 3 = DcAdmin

Event	Description
	4 = System 5 = Application 6 = ServicePrincipal 7 = CustomPolicy 8 = SystemPolicy
UserKey	An alternative ID for the user identified in the UserId property. This property is populated with the passport unique ID (PUID) for events performed by users in SharePoint, OneDrive, and Exchange.
Workload	Stores the Office 365 service where the activity occurred.
Version	Version of the Azure Information Protection client that performed the audit action
Scope	Specifies scope.

AipHeartBeat

The following table contain information related to AIP heartbeat events.

Event	Description
ApplicationId	Corresponds to the Microsoft Entra Application ID.
ApplicationName	Application friendly name of the application performing the operation.
CreationDate	The date and time in Coordinated Universal Time (UTC) in ISO8601 format when the user performed the activity.
DataState	Specifies the state of the data.
DeviceName	The name of the user's device.
Identity	The identity of the user or service to be authenticated.
IsProtected	Whether protected: True/False
IsProtectedBefore	Whether the content was protected before change: True/False
IsValid	Boolean
Location	The location of the document with respect to the user's device. The possible values are unknown, localMedia, removableMedia, fileshare, and cloud.
ObjectState	Specifies the state of the object.
Operation	The operation type for the audit log. The name of the user or admin activity.
PSComputerName	Computer Name
PSShowComputerName	The value is False for documented edited in Office 365.
Platform	Device platform (Win, OSX, Android, iOS).
ProcessName	Process that hosts MIP SDK.

Event	Description
ProductVersion	Version of the Azure Information Protection client that performed the audit action.
ProtectionType	Protection type can be template or ad-hoc.
RecordType	Shows the value of Label Action. The operation type indicated by the record. For more information, see the full list of record types.
RunspaceId	The Runspace is a specific instance of PowerShell which contains modifiable collections of commands, providers, variables, functions, and language elements that are available to the command line user.
SensitiveInfoTypeData	Stores the datatype of the Sensitive Info Type Data.
TemplateId	TemplateID parameter to get a specific template. The Get-AipServiceTemplate cmdlet gets all existing or selected protection templates from Azure Information Protection.
UserId	The UPN of the user who performed the action (specified in the Operation property) that resulted in the record being logged; for example, `my_name@my_domain_name`. **Note**: Records for activity performed by system accounts (such as SHAREPOINT\system or NT AUTHORITY\SYSTEM) are also included. In SharePoint, another value display in the UserId property is `app@sharepoint`. This value indicates the "user" who performed the activity was an application that has the necessary permissions in SharePoint to perform organization-wide actions (such as search a SharePoint site or OneDrive account) on behalf of a user, admin, or service.
UserType	The type of user that performed the operation. See the UserType table for details on the types of users. 0 = Regular 1 = Reserved 2 = Admin 3 = DcAdmin 4 = System 5 = Application 6 = ServicePrincipal 7 = CustomPolicy 8 = SystemPolicy
UserKey	An alternative ID for the user identified in the UserId property. This property is populated with the passport unique ID (PUID) for events performed by users in SharePoint, OneDrive, and Exchange.

MicrosoftGraphDataConnectConsent complex type

Parameters	Type	Mandatory?	Description
ApplicationId	Edm.Guid	Yes	The application identification.
ApplicationVersion	Edm.String	Yes	The application version.

Parameters	Type	Mandatory?	Description
AppRegistrationTenantId	Edm.Guid	Yes	The application registration tenant id.
Approver	Edm.String	Yes	The approver's user principal name.
ApprovalUpdatedDateInUTC	Edm.Date	Yes	The update date time in UTC.
ApprovalExpiryDateInUTC	Edm.Date	Yes	The expiry date time in UTC.
ApprovalValidDays	Edm.Int32	Yes	The number of days from update for which the approval is valid.
DestinationSinks	Edm.String	Yes	The destination sinks.
DestinationTenantId	Edm.Guid	Yes	The destination tenant id.
Reason	Edm.String	No	The reason provided by the admin who performed the operation.
State	Edm.String	Yes	The consent state.
Datasets	CollectionSelf.MGDCDataset	Yes	Details on the datasets which were consented to as part of this operation.

Complex Type MGDCDataset

Parameters	Type	Mandatory?	Description
DatasetName	Edm.String	Yes	The name of the dataset in the consent operation.
DatasetColumns	Edm.String	Yes	The list of columns for the dataset in the consent operation.
DenyGroups	Edm.String	No	The deny groups list for the dataset in the consent operation.
Scope	Edm.String	Yes	The scope types for the dataset in the consent operation. Possible values are **All**, **List**, and **FilterUri**.
ScopeFiltersUris	Edm.String	No	The scope filter URI for the dataset in the consent operation.
ScopeList	Edm.String	No	The scope group list for the dataset in the consent operation.
PrivacyPolicyType	Edm.String	Yes	The privacy policy types for the dataset in the consent operation. Possible values are None and DenyList.

Viva Glint schema

Viva Glint events returned in audit log searches use this schema (and also the Common schema).

Parameters	Type	Mandatory?	Description
ClientUUID	Edm.String	Y	The Client UUID of the Viva Glint instance.
CampaignUUID	Edm.String	N	The stable internal identifier of the agentic campaign.
DataSubjectUUID	Edm.String	N	The stable internal identifier of the data subject affected by a data subject rights deletion.
ExportFormat	Edm.String	N	The format of the exported agentic campaign data.
ModifiedProperties	Edm.String	N	The properties that were changed by the operation.
PolicyUUID	Edm.String	N	The stable internal identifier of the agentic policy.
ReportUUID	Edm.String	N	The stable internal identifier of the agentic campaign report.
RetentionCutoff	Edm.Date	N	The retention cutoff date used when agentic campaign data is deleted by the retention process.
SharedWith	Edm.String	N	The recipients or scope with which an agentic campaign report was shared.
ImportType	Edm.String	N	The data import source type.
DataDSRControl	Edm.String	N	This in-platform control determines whether survey data is deleted when a user is removed from the Viva Glint platform.
DiscardEmployeeIds	Edm.String	N	This in-platform control specifies whether the Employee IDs of previously deleted employees are disregarded or retained in the Viva Glint platform.
JobName	Edm.String	N	The name of a Glint data app job that was run.
ExtendedCompletionDate	Edm.Date	N	The new date to which a closed survey cycle has been extended.
FeedBackComponentName	Edm.String	N	The name of a specific component within the 360 feedback program.
ExportApiClientAppId	Edm.String	N	The application ID which made the data export request using APIs.

Viva Goals schema

Viva Goals events returned in audit log searches use this schema (and also the Common schema).

Parameters	Type	Mandatory?	Description
Detail	Edm.String	No	A description of the event or the activity that occurred in Viva Goals.
Username	Edm.String Term="Microsoft.Office.Audit.Schema.PIIFlag" Bool="true"	No	The name of the user who triggered the event.
UserRole	Edm.String	No	The role of the user who triggered this event in Viva Goals. This value mentions if the user is an organization admin or an owner.
OrganizationName	Edm.String Term="Microsoft.Office.Audit.Schema.PIIFlag" Bool="true"	No	The name of the organization in Viva Goals where the event was triggered.
OrganizationOwner	Edm.String Term="Microsoft.Office.Audit.Schema.PIIFlag" Bool="true"	No	The owner of the organization in Viva Goals where the event occurred.
OrganizationAdmins	Collection(Edm.String) Term="Microsoft.Office.Audit.Schema.PIIFlag" Bool="true"	No	The admin(s) of the organization in Viva Goals where the event occurred. There can be one or more admins in the organization.
UserAgent	Edm.String Term="Microsoft.Office.Audit.Schema.PIIFlag" Bool="true"	No	The user agent (browser details) of the user who triggered the event. UserAgent might not be present in case of a system generated event.
ModifiedFields	Collection(Common.NameValuePair)	No	A list of attributes that were modified along with its new and old values output as a JSON.
ItemDetails	Collection(Common.NameValuePair)	No	Additional properties about the object that was modified.

Microsoft Planner schema

Microsoft Planner events returned in audit log searches use this schema.

Microsoft Planner overwrites the definition of ObjectId and ResultStatus in the Common schema. Microsoft Planner's ObjectId definition is bound to each Microsoft Planner's record type and is illustrated individually.

Microsoft Planner's ResultStatus is defined as the following.

Enum: ResultStatus - Type: Edm.Int32

ResultStatus

Value	Member name	Description
1	Success	The user request succeeded.
2	Failure	The user request failed due to reasons other than authorization.
3	AuthorizationFailure	The user requested failed due to failed authorization.

Microsoft Planner extends the Common schema with the following record types.

PlannerGoal record type

Properties	Type	Description
ObjectId	Edm.String	Id of the goal requested.
PlanId	Edm.String	Id of the plan containing the goal.

PlannerPlan record type

Properties	Type	Description
ObjectId	Edm.String	Id of the plan requested.
ContainerType	Self.ContainerType	Type of the container associated with the plan.
ContainerId	Edm.String	Id of the container associated with the plan.
SharedWithContainerId	Edm.String	Id of the container with shared access to the plan.
SharedWithContainerType	Self.ContainerType	Type of the container with shared access to the plan.
SharedWithContainerAccessLevel	Self.PlanAccessLevel	Level of access given to container with shared access to the plan.

Enum: ContainerType - Type Edm.Int32

ContainerType

Value	Member name	Description
0	Invalid	Used when the requested plan is not found.
2	Group	The plan is associated with a Microsoft 365 Group.
3	TeamsConversation	The plan is associated with a Teams conversation.
4	OfficeDocument	The plan is associated with an Office document.
5	Roster	The plan is associated with a roster group.
6	Project	The plan originates from Microsoft Project.
7	User	The plan is associated with a user.
8	TeamsChannel	The plan is associated with a Teams channel.
10	PlannerTask	The plan is associated with a Planner task.

Enum: PlanAccessLevel - Type Edm.Int32

PlanAccessLevel

Value	Member name	Description
1	ReadAccess	Access to read Plan.
2	ReadWriteAccess	Access to read and write to Plan.
3	FullAccess	Access to read, write and configure Plan.

PlannerCopyPlan record type

Properties	Type	Description
ObjectId	Edm.String	Id of the plan being copied.
OriginalPlanId	Edm.String	Id of the plan being copied. Same as ObjectId.
OriginalContainerType	Self.ContainerType	Type of the container associated with the original plan.
OriginalContainerId	Edm.String	Id of the container associated with the original plan.
NewPlanId	Edm.String	Id of the new plan. Null when the operation failed.
NewContainerType	Self.ContainerType	Type of the container associated with the new plan.
NewContainerId	Edm.String	Id of the container associated with the new plan.

PlannerTask record type

Properties	Type	Description
ObjectId	Edm.String	Id of the task requested.
PlanId	Edm.String	Id of the plan containing the task.

PlannerRoster record type

Properties	Type	Description
ObjectId	Edm.String	Id of the roster requested.
MemberIds	Edm.String	A comma-separated string of member ids changed to the roster.

PlannerGoalList record type

Properties	Type	Description
ObjectId	Edm.String	A representation of the view query for a list of goals.
GoalList	Edm.String	A comma-separated string of goal ids queried.

PlannerPlanList record type

Properties	Type	Description
ObjectId	Edm.String	A representation of the view query for a list of plans.
PlanList	Edm.String	A comma-separated string of plan ids queried.

PlannerTaskList record type

Properties	Type	Description
ObjectId	Edm.String	A representation of the view query for a list of tasks.
TaskList	Edm.String	A comma-separated string of task ids queried.

PlannerTenantSettings record type

Properties	Type	Description
ObjectId	Edm.String	Original tenant settings in JSON.
TenantSettings	Edm.String	New tenant settings in JSON.

PlannerRosterSensitivityLabel record type

Properties	Type	Description
ObjectId	Edm.String	Id of the sensitivity label. Null when the sensitivity label is removed.
Roster	Edm.String	Id of the roster to which the sensitivity label is changed.
AssignmentMethod	Self.SensitivityLabelAssignmentMethod	The assignment method of the sensitivity label.

Enum: SensitivityLabelAssignmentMethod - Type Edm.Int32

SensitivityLabelAssignmentMethod

Value	Member name	Description
0	Standard	The sensitivity label is automatically applied but not allowed to override a privileged label assignment.
1	Privileged	The sensitivity label is applied manually by a user or by an admin.
2	Auto	The sensitivity label is automatically applied and is allowed to override a privileged label assignment.

PlannerChatMessage record type

Properties	Type	Description
ObjectId	Edm.String	Id of the chat message
ThreadType	Self.ThreadType	The type of chat. Only option is Task, for a chat scoped to a single Task.
ParentEntityId	Edm.String	Id of the element containing the chat, the Task id.

Enum: ThreadType - Type Edm.Int32

ThreadType

Value	Member name	Description
0	Task	Chat is contained inside of a Task.

PlannerChatMessageList record type

Properties	Type	Description
ObjectId	Edm.String	A representation of the view query for a list of chat messages.
MessageList	Edm.String	A comma-separated string of chat message ids queried.

Microsoft Project for the web schema

Microsoft Project For The web extends the Common schema with the following record types.

ProjectForThewebProject record type

Properties	Type	Mandatory?	Description
ProjectId	Edm.Guid	No	Id of the Project being audited.
AdditionalInfo	CollectionSelf.AdditionalInfo	No	Additional information.

ProjectForThewebTask record type

Properties	Type	Mandatory?	Description
ProjectId	Edm.Guid	Yes	Id of the Project being audited.
TaskId	Edm.Guid	Yes	Id of the Task being audited.
AdditionalInfo	CollectionSelf.AdditionalInfo	No	Additional information.

ProjectForThewebRoadmap record type

Properties	Type	Mandatory?	Description
RoadmapId	Edm.Guid	Yes	Id of the Roadmap being audited.
AdditionalInfo	CollectionSelf.AdditionalInfo	No	Additional information.

ProjectForThewebRoadmapItem record type

Properties	Type	Mandatory?	Description
RoadmapItemId	Edm.Guid	Yes	Id of the Roadmap Item being audited.
AdditionalInfo	CollectionSelf.AdditionalInfo	No	Additional information.

Complex Type AdditionalInfo

Parameters	Type	Mandatory?	Description
EnvironmentName	Edm.String	No	Id of the environment where action was performed.

ProjectForThewebProjectSetting record type

Properties	Type	Mandatory?	Description
ProjectEnabled	Edm.Boolean	Yes	The value that was set for Project for the web (1= enabled, 0 disabled).

ProjectForThewebRoadmapSetting record type

Properties	Type	Mandatory?	Description
RoadmapEnabled	Edm.Boolean	Yes	The value that was set for Roadmap (1= enabled, 0 disabled).

ProjectForThewebAssignedToMeSetting record type

Properties	Type	Mandatory?	Description
AssignedToMeEnabled	Edm.Boolean	Yes	The value that was set for AssignedToMe (1= enabled, 0 disabled).

Viva Pulse schema

Viva pulse events returned in audit log searches use this schema (and also the Common schema).

Parameters	Type	Mandatory?	Description
EventName	Edm.String	No	A description of the event or the activity that occurred in Viva Pulse.
PulseId	Edm.String	No	Id of the pulse survey.
EventDetails	Collection(Common.NameValuePair)	No	Additional properties about the event.

Some of the VivaPulse events record different properties in EventDetails. Each property recorded in EventDetail are described in the table.

EventName	PropertName	Description
PulseReportShare	Recipients	List of recipient ids with whom the pulse survey is shared.
PulseCreate	Recipients	List of user ids who are participants of the specified pulse survey.

EventName	PropertName	Description
PulseInvite	Recipients	List of user IDs who are additionally invited to the pulse.
PulseTenantSettingsUpdate	TenantSettingName	Changed settings name.
PulseSubmit	Not Applicable	This event does not record any property into EventDetails.
PulseExtendDeadline	Not Applicable	This event does not record any property into EventDetails.
PulseCancel	Not Applicable	This event does not record any property into EventDetails.
PulseCreateDraft	Not Applicable	This event does not record any property into EventDetails.
PulseDeleteDraft	Not Applicable	This event does not record any property into EventDetails.
PulseDeleteUserData	Not Applicable	This event does not record any property into EventDetails.
PulseQuestionDeleted	Not Applicable	This event does not record any property into EventDetails.
PulseDataExport	ExportType,ExportCompletedDate	ExportType indicates whether the data export is an Admin-level or Author-level export, ExportCompletedDate specifies when the data export was finalized.
PulseConfidentialConversationStarted	Not Applicable	This event does not record any property into EventDetails.
PulseConfidentialConversationResponded	Not Applicable	This event does not record any property into EventDetails.

EventName	PropertName	Description
PulseConfidentialConversationMessageDeleted	Not Applicable	This event does not record any property into EventDetails.

Compliance Manager schema

Compliance Manager events returned in audit log searches use this schema (and also the Common schema).

Parameters	Type	Mandatory?	Description
Details	Collection(SettingsChange)	No	A description of the event that occurred for a Microsoft Purview Compliance Manager.

Values taken by SettingsChange properties in Details for different operations are described in the table below.

Operation	PropertyName	Description
All Operations	Name	Name of setting involved in the Compliance Manager operation.
All Operations	NewValue	New value for the new settings.
All Operations	OriginalValue	Original value for new setting.

> **Note**
>
> - For role changes the name is the role type.
> - The audit record reflects the change in event, such as user is assigned a role or revoked role.
> - The original and new value have the emails of the user for which the role has changed.
> - In case there is no change in the role, that role type isn't present in the audit record.

Backup Policy schema

Parameters	Type	Mandatory?	Description
PolicyID	Edm.String	Yes	The ID of the policy.
EditMethodology	Edm.String	No	How the policy was created / edited.
CountOfArtifactsBeingAdded	Edm.Int32	No	Number of artifacts being added.
CountOfArtifactsBeingRemoved	Edm.Int32	No	Number of artifacts being removed.
ServiceType	Edm.String	No	Whether it is a SharePoint, Exchange, or OneDriveForBusiness policy.

Parameters	Type	Mandatory?	Description
BillingPolicyId	Edm.String	No	The ID of the associated Billing Policy.

Granular Browse Task schema

Parameters	Type	Mandatory?	Description
SessionId	Edm.String	Yes	The ID of the Browse Task.
RestoreTime	Edm.Date	No	Restore Point dateTime on which browse session was created.
BackupItemId	Edm.String	No	Artifact Id for which browse session was created.
ServiceType	Edm.String	Yes	Whether it is a SharePoint, Exchange, or OneDriveForBusiness policy.

Restore Task schema

Parameters	Type	Mandatory?	Description
TaskID	Edm.String	Yes	The ID of the Restore Task.
CreationMethodology	Edm.String	No	How the Restore Task was created / edited.
CountOfArtifactsBeingAdded	Edm.Int32	No	Number of artifacts being added.
CountOfArtifactsBeingRemoved	Edm.Int32	No	Number of artifacts being removed.
ServiceType	Edm.String	No	Whether it is a SharePoint, Exchange, or OneDriveForBusiness policy.

Restore Item schema

Parameters	Type	Mandatory?	Description
RestoreTime	Edm.DateTime	Yes	Time which the item is being restored to.
RestoreLocationType	Edm.String	Yes	Location type that the item is being restored to.
RestoreLocation	Edm.String	No	Location that the item is being restored to.
TaskID	Edm.String	Yes	The ID of the Restore task.
BackupItemID	Edm.String	Yes	ID of the Backup Item being restored.
ProtectionUnitID	Edm.String	Yes	Protection Unit ID of the item being restored.
SuccessStatus	Edm.String	No	Whether the restore operation was successful.
BackupItemType	Edm.String	Yes	Whether the Backup Item is a Site / Account / Mailbox.

Parameters	Type	Mandatory?	Description
ServiceType	Edm.String	No	Whether it is a SharePoint, Exchange, or OneDriveForBusiness policy.

Backup Item schema

Parameters	Type	Mandatory?	Description
PolicyID	Edm.String	Yes	Policy ID of the Policy the item is getting added to.
ItemID	Edm.String	Yes	ID of the Backup Item.
ProtectionUnitID	Edm.String	Yes	Protection Unit ID of the item being backed up.
ResultStatus	Edm.String	No	Whether the restore operation was successful.
BackupItemType	Edm.String	Yes	Whether the Backup Item is a Site / Account / Mailbox.
EditMethodology	Edm.String	No	How the backup item is to be added.
ServiceType	Edm.String	No	Whether it is a SharePoint, Exchange, or OneDriveForBusiness policy.

Cloud Policy service schema

The audit records for events related to the Cloud Policy service extend the Common schema as follows:

PolicyConfigChangeAuditRecord

Parameters	Type	Mandatory?	Description
ConfigId	Edm.String	No	ID of the policy configuration
ConfigName	Edm.String	No	Given name of policy configuration
Description	Edm.String	No	Description provided for policy configuration
ConfigScope	Self.CPSScope	No	Scope of policy configuration
Groups	Collection(Common.NameValuePair)	No	List of configured groups
Priority	Edm.Int32	No	Priority value of policy configuration
Policies	Collection(Self.Policy)	No	List of configured policy settings
Priorities	Collection(Self.PrioritySetting)	No	List of policy configurations and their priority values

Enum: CPSScope - Type: Edm.Int32

Value	Member name	Description
1	Tenant	The policy configuration is scoped to all users in the tenant.
2	Anonymous	The policy configuration is scoped to anonymous users.
3	User	The policy configuration is scoped to users in configured Microsoft Entra group(s).

Complex Type Policy

Parameters	Type	Mandatory?	Description
PolicyId	Edm.String	No	ID of the policy setting
PolicyName	Edm.String	No	Name of policy setting
Value	Edm.String	No	Configured value of policy setting
Settings	Collection(Self.Setting)	No	Configured setting

Complex Type Setting

Parameters	Type	Mandatory?	Description
SettingId	Edm.String	No	ID of the setting
SettingName	Edm.String	No	Name of setting
Value	Edm.String	No	Configured value of setting

Complex Type PrioritySetting

Parameters	Type	Mandatory?	Description
ConfigId	Edm.String	No	ID of the policy configuration
ConfigName	Edm.String	No	Given name of policy configuration
Value	Edm.String	No	Configured priority of policy configuration

Cloud update profile configuration schema

The Cloud Update profile configuration related events extend the Common schema with the following record types.

CloudUpdateProfileConfigAuditRecord

Parameters	Type	Mandatory?	Description
ProfileName	Edm.String	Yes	Name of profile
ProfileState	Self.ProfileState	No	State of profile
Deadline	Edm.Int32	No	Configured deadline
UpdateValidationState	Self.UpdateValidationState	No	State of Update Validation
Waves	Collection(Self.Wave)	No	Collection containing configured waves
WaveDelay	Edm.Int32	No	Set delay between waves

Enum: ProfileState - Type: Edm.Int32

Value	Member name	Description
1	Enabled	The profile is enabled and active.
2	Disabled	The profile is disabled.
3	Paused	The profile is enabled, but in paused state.

Enum: UpdateValidationState - Type: Edm.Int32

Value	Member name	Description
1	Enabled	Update Validation is enabled.
2	Disabled	Update Validation is disabled.

Complex Type Wave

Parameters	Type	Mandatory?	Description
Name	Edm.String	No	Name of wave
Type	Self.WaveType	No	Wave specified by administrator or automatic catch-all wave
Groups	Collection(Common.NameValuePair)	No	Collection of groups configured for this wave

Enum: WaveType - Type: Edm.Int32

Value	Member name	Description
1	Groups	Wave was configured by administrator using groups.

Value	Member name	Description
2	RemainingDevices	Automatically created wave which includes all devices in profile's scope which are not covered by previous waves.

Cloud Update tenant configuration schema

The Cloud Update tenant configuration related events extend the Common schema with the following record types.

CloudUpdateTenantConfigAuditRecord

Parameters	Type	Mandatory?	Description
ProfileExclusionWindows	Collection(Self.ExclusionWindow)	No	Collection of configured exclusion windows
ExclusionList	Collection(Common.NameValuePair)	No	Collection of configured exclusions
TAK	Edm.String	No	Tenant Association Key

Complex Type ExclusionWindow

Parameters	Type	Mandatory?	Description
Name	Edm.String	No	Given name of exclusion window
StartDate	Edm.Date	No	Start date of exclusion windows
EndDate	Edm.Date	No	End date of exclusion windows
Groups	Collection(Common.NameValuePair)	No	Collection of groups the exclusion window is scoped to

Cloud Update device schema

The Cloud Update device related events extend the Common schema with the following record types.

CloudUpdateDeviceConfigAuditRecord

Parameters	Type	Mandatory?	Description
RollbackDevices	Self.Rollback	No	Rollbacks triggered
ChannelChangeDevices	Self.ChannelChange	No	Channel changes triggered

Complex Type Rollback

Parameters	Type	Mandatory?	Description
RollbackType	Self.RollbackType	No	Type of rollback
RollbackBuildNumber	Edm.String	No	Targeted build number to roll back to
Devices	Collection(Edm.String)	No	Collection of devices targeted by rollback

Enum: RollbackType - Type: Edm.Int32

Value	Member name	Description
1	SpecificDevices	Rollback was targeted at a specific subset of devices.
2	AllDevices	Rollback was targeted at all devices within the profile's scope.

Complex Type ChannelChange

Parameters	Type	Mandatory?	Description
ChannelChange	Self.Channel	No	Targeted update channel
Devices	Collection(Edm.String)	No	Collection of targeted devices
Groups	Collection(Common.NameValuePair)	No	Collection of targeted groups

Enum: Channel - Type: Edm.Int32

Value	Member name	Description
1	MonthlyEnterpriseChannel	Monthly Enterprise Channel
2	CurrentChannel	Current Channel

Microsoft Entra Risk Detection schema

Microsoft Entra risk detection events returned in audit log searches use this schema (and also the Common schema).

Parameters	Type	Mandatory?	Description
Activity	Edm.String	No	The activity type for which risk was detected.
ActivityDateTime	Edm.Date	No	The date and time in Coordinated Universal Time (UTC) when the risky activity occurred.
AdditionalInfo	Edm.String	No	Additional Information in JSON format for the detected risk.

Parameters	Type	Mandatory?	Description
CorrelationId	Edm.String	No	An identifier that can used to correlate sign-in activities associated with a risk detection.
DetectedDateTime	Edm.Date	No	The date and time in Coordinated Universal Time (UTC) when the risk was detected.
DetectionTimingType	Edm.String	No	Indicates the timing type of the detected risk. Possible values are real-time/offline.
LastUpdatedDateTime	Edm.Date	No	The date and time in Coordinated Universal Time (UTC) when the risk detection was last updated.
Location	Self.LocationType	No	Information about the location from which the sign-in activity was detected.
RequestId	Edm.String	No	Indicates the Request Id of the sign-in activity for which risk was detected. If risk detection is not associated with sign-in then this property is null.
RiskDetail	Edm.String	No	The details of the detected risk.
RiskEventType	Edm.String	No	The type of event for which risk was detected.
RiskId	Edm.String	No	An unique identifier for the risk detection.
RiskLevel	Edm.String	No	The level of the risk detected.
RiskState	Edm.String	No	The risk state of a risky user or a sign-in linked to the risk detection.
Source	Edm.String	No	The source of the detected risk.
TokenIssuerType	Edm.String	No	The type of token issuer for the sign-in linked to the risk detection.
UserDisplayName	Edm.String	No	The user principal name (UPN) of the user for whom the risk was detected.

LocationType complex type

Parameters	Type	Mandatory?	Description
City	Edm.String	No	The city where the sign-in was performed.
CountryOrRegion	Edm.String	No	The country or region where the sign-in was performed.
GeoCoordinates	Self.GeoCoordinatesType	No	The geo co-ordinates of the location where the sign-in was performed.
State	Edm.String	No	The state where the sign-in was performed.

GeoCoordinatesType complex type

Parameters	Type	Mandatory?	Description
Altitude	Edm.String	No	The altitude of the location where the sign-in was performed.
Latitude	Edm.String	No	The latitude of the location where the sign-in was performed.
Longitude	Edm.String	No	The longitude of the location where the sign-in was performed.

Microsoft Edge WebContentFiltering schema

Microsoft Edge WebContentFiltering events returned in audit log searches use this schema (and also the Common schema).

Parameters	Type	Mandatory?	Description
URLPath	Edm.String	Yes	The URL that was browsed.
DomainURL	Edm.String	Yes	The domain URL that was browsed.
Category	Edm.String	Yes	Category of browsed URL.

Microsoft 365 Copilot scheduled prompt schema

Microsoft 365 Copilot scheduled prompt events returned in audit log searches use this schema (and also the Common schema).

Copilot scheduled prompts allow users to automate Copilot prompts, so they run on a defined schedule in Microsoft 365 Copilot Chat.

Parameters	Type	Mandatory?	Description
ScenarioType	Edm.String	Yes	The name of automation.
PromptText	Edm.String	No	The Copilot prompt that's run by the LLM.
AutomationId	Edm.String	Yes	A unique identifier that correlates all activities related to each run of the Copilot prompt.
TriggerMode	Edm.String	No	The schedule for when the Copilot prompt runs, shown in cron format.

Microsoft Places Directory schema

Microsoft Places Directory events returned in audit log searches use this schema (and also the Common schema).

Parameters	Type	Mandatory?	Description
PlaceType	Edm.String	No	The type for the place item that is

Parameters	Type	Mandatory?	Description
			created/updated/deleted by the request. For example, "Building", "Room", "Desk", and so on.
Parameters	Collection(Common.NameValuePair)	No	The name and value for all parameters that were used with the cmdlet that is identified in the Operations property.
ModifiedProperties	Collection(Common.ModifiedProperty)	No	The property includes the name of the property that was modified, the new value of the modified property, and the previous value of the modified object.

Teams Evaluation Data Hub schema

Teams Evaluation Data Hub events returned in audit log searches use the Common schema. Teams Evaluation Data Hub doesn't introduce additional service-specific properties beyond the Common schema. For these events, the following Common schema properties have service-specific meanings.

Parameters	Type	Mandatory?	Description
Workload	Edm.String	Yes	The Office 365 service where the activity occurred. Always `TeamsEvalDataHub`.
RecordType	Self.AuditLogRecordType	Yes	The type of operation indicated by the record. `TeamsEvalDataHubDataAccess` (444) for `DataRead` and `DataWrite` operations; `TeamsEvalDataHubPermissionChange` (445) for `PermissionGranted` and `PermissionRevoked` operations.
Operation	Edm.String	Yes	The name of the user or service activity. One of `DataRead`, `DataWrite`, `PermissionGranted`, or `PermissionRevoked`.
ObjectId	Edm.String	No	Identifies the data set or access-control object that the operation targets.

Microsoft Sentinel data lake and graph schema

The following Microsoft Sentinel events returned in audit log searches use these schemas (and also the Common schema).

SentinelNotebookOnLake

For audit log activities, see Microsoft Sentinel data lake notebook activities.

Parameter	Type	Mandatory?	Description
EventTime	Edm.Date	Yes	Timestamp when the Spark Notebook execution was submitted/started.
Compute	Edm.String	No	Compute selected.
DatabaseName	Edm.String	No	The workspace the command ran on.
TableName	Edm.String	No	Name of the table.
SessionDurationInSecs	Edm.String	Yes	Total duration of the session.
SessionStartTime	Edm.Date	No	Start time for the session.
SessionEndTime	Edm.Date	No	End time for the session.
KernalId	Edm.Guid	Yes	Jupyter KernelId, uniquely identifies the Notebook session.
SessionId	Edm.Guid	No	Correlation ID for the various code blocks executed with a Spark session.
Interface	Edm.String	Yes	Interface from where the notebook was run.

SentinelJob

For audit log activities, see Microsoft Sentinel data lake job activities.

Parameter	Type	Mandatory?	Description
EventTime	Edm.Date	Yes	Timestamp when the operation on job was initiated.
Operation	Edm.String	Yes	Job Operation name.
Job Type	Edm.String	Yes	Type of Job like Notebook, KQL.
Job ID	Edm.Guid	Yes	Unique identifier of a job.
Job Name	Edm.String	Yes	Name of the job.
Compute	Edm.String	No	Compute configuration of the job.
Schedule	Edm.String	No	Schedule details, if configured for the job.
Run ID	Edm.Guid	No	ID of a specific job run instance.
JobRunStatus	Edm.String	No	Status of the job execution.
Logs	Edm.String	No	Logs from the notebook run.

Parameter	Type	Mandatory?	Description
JobExecutionDurationInSecs	Edm.Int64	No	Time taken for job execution.
JobTotalDurationInSecs	Edm.Int64	No	Total Duration of the job operation including session.
JobStartTime	Edm.Date	No	Start time of the job.
JobEndTime	Edm.Date	No	End time of the job.
Interface	Edm.String	Yes	Interface from where the job operation was done.
DatabasesRead	Collection(Edm.String)	No	The list of workspaces read by the job.
DatabasesWrite	Collection(Edm.String)	No	The list of workspaces written to by the job
TablesRead	Collection(Edm.String)	No	The list of tables read by the job.
TablesWrite	Collection(Edm.String)	No	The list of tables written to by the job.
Query	Edm.String	No	KQL query or notebook executed in the job.

SentinelKQLOnLake

For audit log activities, see Microsoft Sentinel data lake KQL activities.

Parameter	Type	Mandatory?	Description
EventTime	Edm.Date	Yes	Timestamp of KQL query execution.
DatabaseName	Collection(Edm.String)	Yes	The workspaces the KQL query ran on.
ResultTableCount	Edm.Int64	No	Output Table Count.
QueryResponse	Edm.String	Yes	Response from executing the KQL query.
TotalRows	Collection(Edm.Int64)	Yes	Total Rows returned from query execution. List of values if multiple queries executed at once.
ComponentFault	Edm.String	No	The entity that caused the query to fail. For example, if the query result is too large, the ComponentFault is 'Client'. If an internal error occurred, the ComponentFault is 'Server'.
FailureReason	Edm.String	No	If KQL query failed, the reason for failure.
ExecutionDuration	Edm.Int64	Yes	Time taken for query execution, in milliseconds.
TotalCPU	Edm.Int64	Yes	Total CPU Duration of the run.

Parameter	Type	Mandatory?	Description
MemoryPeak	Edm.Int64	Yes	Memory Peak of the KQL query execution.
Interface	Edm.String	Yes	Interface from where the KQL query was executed.
TablesRead	Collection(Edm.String)	Yes	The list of tables read in the KQL query.
QueryText	Edm.String	Yes	The KQL query executed in scrubbed form.

SentinelLakeOnboarding

For audit log activities, see Microsoft Sentinel data lake onboarding activities.

Parameter	Type	Mandatory?	Description
BillingAzureSubscriptionId	Edm.String	No	Azure subscription chosen for Sentinel data lake billing.
BillingAzureResourceGroupName	Edm.String	No	Azure resource group chosen for Sentinel data lake billing.
TenantId	Edm.String	No	Tenant ID associated with the lake setup or update.
ProvisioningStatus	Edm.String	No	Status of provisioning the lake.

SentinelLakeDataOnboarding

Property Name	Type	Mandatory?	Description
DataOnboardingAtSetup	Edm.String	No	Data sets ingested during Sentinel data lake onboarding.
Tables	Collection(Edm.String)	No	List of table names ingested during Sentinel data lake onboarding.
SubscriptionsEnabled	Collection(Edm.String)	No	List of subscriptions enabled for ARG ingestion.
DataOnboardingStatus	Edm.String	No	Status of operation.

SentinelAITool

For audit log activities, see Microsoft Sentinel AI tool activities.

Parameter	Type	Mandatory?	Description
EventOccurenceTime	Edm.Date	Yes	Timestamp of the operation.
ToolID	Edm.Guid	Yes	Identifier of the AI Tool.
ToolName	Edm.String	Yes	Name of the AI Tool.

Parameter	Type	Mandatory?	Description
Interface	Edm.String	Yes	Interface from where the AI Tool was run.
InputParameters	Edm.String	No	Parameters given to the tool.
DatabasesRead	Collection(Edm.String)	No	The list of databases read from in the run.
TablesRead	Collection(Edm.String)	No	The list of tables read from in the run.
APIsCalled	Collection(Edm.String)	No	APIs called by the AI Tool.
FailureReason	Edm.String	No	If the run failed, the reason for failure.
TotalRows	Collection(Edm.Int64)	No	Total Rows returned.
DataScanned	Edm.Int64	No	Total GBs Scanned.
ExecutionDuration	Edm.Int64	No	Time taken for query execution, in milliseconds.
TotalCpuHours	Edm.Int64	No	Total CPU Duration of the run.
TotalSCUHours	Edm.Int64	No	Total SCUs used in the run.

SentinelGraph

For audit log activities, see Microsoft Sentinel graph activities.

Parameters	Type	Mandatory?	Description
EventTime	Edm.Date	Yes	Timestamp of the operation.
GraphName	Edm.String	Yes	Name of the graph instance.
Operation	Edm.string	Yes	Action taken on graph.
OperationInput	Edm.string	No	Parameters of graph action.
GraphQuery Stats	Edm.string	No	Collection of response metadata.
GraphQuery Status	Edm.String	No	Response of the query or action on graph.
Interface	Edm.String	Yes	Interface from where the graph action was taken.

Purview On-demand classification schema

Microsoft Purview on-demand classification events returned in audit log searches use this schema.

DataScanClassification schema

The DataScanClassification audit schema is designed to capture and log activities when a file was evaluated for sensitive content as part of an On-demand classification scan for SharePoint or

OneDrive.

Parameters	Type	Mandatory	Description
ClassificationInfo	Collection(Self.SensitiveInformation)	No	Details about sensitive information found in the file after the scan.
PolicyId	Edm.Guid	Yes	The guid of the On-demand classification scan.
ClassificationMode	Edm.Int32	Yes	Whether scan was executed for specific classifiers or all.
ClassificationPosture	Edm.Int32	Yes	Comparison of sensitive info discovered before and after the scan.
ClassificationResult	Edm.Int32	Yes	File classification status.
DocumentMetaData	Self.DocumentMetadata	No	Describes metadata about the document in SharePoint or OneDrive that contained the sensitive information.
PreviousClassificationInfo	Collection(Self.SensitiveInformation)	No	Details about sensitive information found in the file after the scan.

DocumentMetaData complex type

Parameters	Type	Mandatory?	Description
itemCreationTime	Edm.Date	Yes	Datetimestamp in UTC of when event logged.
SiteCollectionGuid	Edm.Guid	Yes	The GUID of the site collection.
SiteCollectionUrl	Edm.String	Yes	Name of the SharePoint site.
FileName	Edm.String	Yes	Name of the path.
FileOwner	Edm.String	Yes	The document owner.
FileOwnerEmail	Edm.String	Yes	Email address of document owner.
FilePathUrl	Edm.String	Yes	The URL of the document
DocumentLastModifier	Edm.String	Yes	The user who last modified the document.
UniqueId	Edm.String	Yes	A guid that identifies the file.
LastModifiedTime	Edm.DateTime	Yes	Timestamp in UTC for when doc was last modified.

Enum: ClassificationMode - Type: Edm.Int32

Value	Description
1	File evaluated for all classifiers configured in tenant.
2	File evaluated only for selected classifiers specified in scan.

Enum: ClassificationPosture - Type: Edm.Int32

Value	Description
1	File did not match any classifiers both before and after the scan.
2	No new classifier found after the scan.
3	File did not match any classifier before the scan. One or more classifiers found in file after the scan.
4	File matched some classifiers before the scan. One or more classifiers no longer match
5	File matched some classifiers before the scan. New classifier or change in existing classifier count or confidence level after the scan.

Enum: ClassificationResult - Type: Edm.Int32

Value	Description
1	File classification completed successfully.
2	File classification completed with error. One or more classifier evaluation failed.
3	File classification failed.

Security and Compliance RBAC schema

The Securityand Compliance RBAC audit schema is designed to capture and log activities when role and role groups management and assignments occur in the Microsoft Purview portal and Microsoft Defender portal. Some of the fields and values specific to the common audit schema are listed below.

Parameters	Type	Mandatory	Description
`Operation`	`Edm.String`	Yes	Name of the SecurityComplianceRBAC operation.
`RecordType`	`Self.AuditLogRecordType`	Yes	Audit record type. SecurityComplianceRBAC record value is 87.
`ResultStatus`	`Edm.String`	Yes	Result of the operation. Values emitted by this workload are `Success` and `Failed`.
`PreExecutionMessage`	`Edm.String`	No	Operation-specific input details. See the applicable operation schema.

Parameters	Type	Mandatory	Description
`PostExecutionMessage`	`Edm.String`	No	Operation-specific result details. See the applicable operation schema.
`AssociatedAdminUnits`	`Collection(Edm.Guid)`	No	Administrative units associated with the record.

CreateRBACRoleGroupDefinition Operation

Creates an RBAC role-group definition.

Parameters	Type	Mandatory	Description
`PreExecutionMessage`	`Edm.String`	Yes	Custom role group object information.
`PreExecutionMessage.RoleGroupId`	`Edm.Guid`	Yes	Identifier of the role group.
`PreExecutionMessage.RoleGroupName`	`Edm.String`	Yes	Internal name of the role group.
`PreExecutionMessage.RoleGroupDisplayName`	`Edm.String`	No	User-facing display name of the role group.
`PreExecutionMessage.RoleGroupDescription`	`Edm.String`	No	Description of the role group.
`PreExecutionMessage.RolesDetail`	`Collection(RoleDetail)`	No	Roles included in the role group.
`PreExecutionMessage.RolesDetail[].RoleId`	`Edm.Guid`	No	Identifier of an included role.
`PreExecutionMessage.RolesDetail[].RoleName`	`Edm.String`	No	Name of an included role.
`PostExecutionMessage`	`Edm.String`	No	New custom role group object information.
`PostExecutionMessage.RoleGroupId`	`Edm.Guid`	No	Identifier of the created role group.
`PostExecutionMessage.RoleGroupName`	`Edm.String`	No	Internal name of the created role group.
`PostExecutionMessage.RoleGroupDisplayName`	`Edm.String`	No	Display name of the created role group.

Parameters	Type	Mandatory	Description
`PostExecutionMessage.RoleGroupDescription`	`Edm.String`	No	Description of the created role group.
`PostExecutionMessage.RolesDetail`	`Collection(RoleDetail)`	No	Roles included in the created role group.
`PostExecutionMessage.RolesDetail[].RoleId`	`Edm.Guid`	No	Identifier of an included role.
`PostExecutionMessage.RolesDetail[].RoleName`	`Edm.String`	No	Name of an included role.

UpdateRBACRoleGroupDefinition Operation

Updates an existing RBAC role-group definition.

Parameters	Type	Mandatory	Description
`PreExecutionMessage`	`Edm.String`	Yes	Custom role group object information.
`PreExecutionMessage.RoleGroupId`	`Edm.Guid`	Yes	Identifier of the role group.
`PreExecutionMessage.RoleGroupName`	`Edm.String`	Yes	Internal name of the role group.
`PreExecutionMessage.RoleGroupDisplayName`	`Edm.String`	No	Updated display name of the role group.
`PreExecutionMessage.RoleGroupDescription`	`Edm.String`	No	Updated description of the role group.
`PreExecutionMessage.RolesDetail`	`Collection(RoleDetail)`	No	Updated collection of roles included in the role group.
`PreExecutionMessage.RolesDetail[].RoleId`	`Edm.Guid`	No	Identifier of an included role.
`PreExecutionMessage.RolesDetail[].RoleName`	`Edm.String`	No	Name of an included role.
`PostExecutionMessage`	`Edm.String`	No	Empty field.

DeleteRBACRoleGroupDefinition Operation

Deletes an RBAC role-group definition.

Parameters	Type	Mandatory	Description
`PreExecutionMessage`	`Edm.String`	Yes	Custom role group object information.
`PreExecutionMessage.RoleGroupId`	`Edm.Guid`	Yes	Identifier of the role group being deleted.
`PreExecutionMessage.RoleGroupName`	`Edm.String`	Yes	Internal name of the role group being deleted.
`PreExecutionMessage.RoleGroupDisplayName`	`Edm.String`	No	Display name of the role group being deleted.
`PreExecutionMessage.RoleGroupDescription`	`Edm.String`	No	Description of the role group being deleted.
`PreExecutionMessage.RolesDetail`	`Collection(RoleDetail)`	No	Roles included in the role group being deleted.
`PreExecutionMessage.RolesDetail[].RoleId`	`Edm.Guid`	No	Identifier of an included role.
`PreExecutionMessage.RolesDetail[].RoleName`	`Edm.String`	No	Name of an included role.
`PostExecutionMessage`	`Edm.String`	No	Empty field.

DeletePermissionAsync Operation

Remove assigned members from a role-group.

Parameters	Type	Mandatory	Description
`PreExecutionMessage`	`Edm.String`	Yes	Input container.
`PreExecutionMessage.MemberDetail`	`MemberDetail`	No	Member assigned to role group.
`PreExecutionMessage.MemberDetail.MemberId`	`Edm.Guid`	No	Identifier of the assigned member.
`PreExecutionMessage.MemberDetail.MemberType`	`Edm.String`	No	Member category, such as `User`, `SecurityGroup`, `ServicePrincipal`, `Application`.
`PreExecutionMessage.MemberDetail.MemberName`	`Edm.String`	No	Name of the assigned member.

Parameters	Type	Mandatory	Description
`PreExecutionMessage.MemberDetail.MemberDisplayName`	`Edm.String`	No	Display name of the assigned member.
`PreExecutionMessage.RoleGroupDetail`	`RoleGroupDetail`	No	Role group being removed from the member.
`PreExecutionMessage.RoleGroupDetail.RoleGroupId`	`Edm.Guid`	No	Identifier of the role group.
`PreExecutionMessage.RoleGroupDetail.RoleGroupName`	`Edm.String`	No	Name of the role group.
`PreExecutionMessage.ScopeDetail`	`ScopeDetail`	No	Scope of the role assignment. Null for global scope.
`PreExecutionMessage.ScopeDetail.ScopeId`	`Edm.Guid`	No	Identifier of the assigned scope.
`PreExecutionMessage.ScopeDetail.ScopeType`	`Edm.String`	No	Resource type of the assigned scope.
`PreExecutionMessage.ScopeDetail.ScopeFeature`	`Edm.String`	No	Feature associated with the assigned scope.
`PreExecutionMessage.ScopeDetail.ScopePathById`	`Edm.String`	No	Resource path identifying the assigned scope.
`PreExecutionMessage.ExceptionMessage`	`Edm.String`	No	Reserved field, currently `null`.
`PreExecutionMessage.ExpiresAt`	`Edm.String`	No	Expiration of the role assignment.
`PostExecutionMessage`	`Edm.String`	No	Empty.

GrantPermissionsAsync Operation

Assigns members to a role-group.

Parameters	Type	Mandatory	Description
`PreExecutionMessage`	`Edm.String`	Yes	Input container.
`PreExecutionMessage.MemberDetail`	`MemberDetail`	No	Member assigned to role group.

Parameters	Type	Mandatory	Description
`PreExecutionMessage.MemberDetail.MemberId`	`Edm.Guid`	No	Identifier of the assigned member.
`PreExecutionMessage.MemberDetail.MemberType`	`Edm.String`	No	Member category, such as `User`, `SecurityGroup`, `ServicePrincipal`, `Application`.
`PreExecutionMessage.MemberDetail.MemberName`	`Edm.String`	No	Name of the assigned member.
`PreExecutionMessage.MemberDetail.MemberDisplayName`	`Edm.String`	No	Display name of the assigned member.
`PreExecutionMessage.RoleGroupDetail`	`RoleGroupDetail`	No	Role group being assigned.
`PreExecutionMessage.RoleGroupDetail.RoleGroupId`	`Edm.Guid`	No	Identifier of the assigned role group.
`PreExecutionMessage.RoleGroupDetail.RoleGroupName`	`Edm.String`	No	Name of the assigned role group.
`PreExecutionMessage.ScopeDetail`	`ScopeDetail`	No	Scope of the role assignment. Null for global scope.
`PreExecutionMessage.ScopeDetail.ScopeId`	`Edm.Guid`	No	Identifier of the assigned scope.
`PreExecutionMessage.ScopeDetail.ScopeType`	`Edm.String`	No	Resource type of the assigned scope, such as `Case`.
`PreExecutionMessage.ScopeDetail.ScopeFeature`	`Edm.String`	No	Feature associated with the assigned scope, such as `InsiderRiskManagement`
`PreExecutionMessage.ScopeDetail.ScopePathById`	`Edm.String`	No	Resource path identifying the assigned scope.
`PreExecutionMessage.ExceptionMessage`	`Edm.String`	No	Reserved field, currently `null`.
`PreExecutionMessage.ExpiresAt`	`Edm.String`	No	Expiration of the assignment. The value is `Permanent` or an ISO 8601 UTC timestamp.
`PostExecutionMessage`	`Edm.String`	No	Result role assignment output container.

Parameters	Type	Mandatory	Description
			Empty when the operation fails before returning a result.
`PostExecutionMessage.MemberDetail`	`MemberDetail`	No	Member contained in the resulting assignment. Uses the same `MemberDetail` schema as the input.
`PostExecutionMessage.RoleGroupDetail`	`RoleGroupDetail`	No	Role group contained in the resulting assignment. Uses the same `RoleGroupDetail` schema as the input.
`PostExecutionMessage.ScopeDetail`	`ScopeDetail`	No	Scope contained in the resulting permission. Uses the same `ScopeDetail` schema as the input.
`PostExecutionMessage.ExceptionMessage`	`Edm.String`	No	Reserved field, currently `null`.
`PostExecutionMessage.ExpiresAt`	`Edm.String`	No	Expiration of the resulting assignment. The value is `Permanent` or an ISO 8601 UTC timestamp.

UpdatePermission Operation

Updates role assignment expiration in a role-group.

Parameters	Type	Mandatory	Description
`PreExecutionMessage`	`Edm.String`	Yes	Input container.
`PreExecutionMessage.MemberDetail`	`MemberDetail`	No	Member assigned to role group.
`PreExecutionMessage.MemberDetail.MemberId`	`Edm.Guid`	No	Identifier of the assigned member.
`PreExecutionMessage.MemberDetail.MemberType`	`Edm.String`	No	Member category, such as `User`, `SecurityGroup`, `ServicePrincipal`, `Application`.
`PreExecutionMessage.MemberDetail.MemberName`	`Edm.String`	No	Name of the assigned member.

Parameters	Type	Mandatory	Description
`PreExecutionMessage.MemberDetail.MemberDisplayName`	`Edm.String`	No	Display name of the assigned member.
`PreExecutionMessage.RoleGroupDetail`	`RoleGroupDetail`	No	Role group being assigned.
`PreExecutionMessage.RoleGroupDetail.RoleGroupId`	`Edm.Guid`	No	Identifier of the assigned role group.
`PreExecutionMessage.RoleGroupDetail.RoleGroupName`	`Edm.String`	No	Name of the assigned role group.
`PreExecutionMessage.ScopeDetail.ScopeId`	`Edm.Guid`	No	Identifier of the assigned scope.
`PreExecutionMessage.ScopeDetail.ScopeType`	`Edm.String`	No	Resource type of the assigned scope, such as `Case`.
`PreExecutionMessage.ScopeDetail.ScopeFeature`	`Edm.String`	No	Feature associated with the assigned scope, such as `InsiderRiskManagement`
`PreExecutionMessage.ScopeDetail.ScopePathById`	`Edm.String`	No	Resource path identifying the assigned scope.
`PreExecutionMessage.ExceptionMessage`	`Edm.String`	No	Reserved field, currently `null`.
`PreExecutionMessage.ExpiresAt`	`Edm.String`	No	Updated role assignment expiration. The value is `Permanent` or an ISO 8601 UTC timestamp.
`PostExecutionMessage`	`Edm.String`	No	Result role assignment output container.
`PostExecutionMessage.MemberDetail`	`MemberDetail`	No	Member contained in the resulting assignment. Uses the same `MemberDetail` schema as the input.
`PostExecutionMessage.RoleGroupDetail`	`RoleGroupDetail`	No	Role group contained in the resulting assignment. Uses the same `RoleGroupDetail` schema as the input.
`PostExecutionMessage.ScopeDetail`	`ScopeDetail`	No	Scope contained in the resulting permission.

Parameters	Type	Mandatory	Description
			Uses the same `ScopeDetail` schema as the input.
`PostExecutionMessage.ExceptionMessage`	`Edm.String`	No	Reserved field, currently `null`.
`PostExecutionMessage.ExpiresAt`	`Edm.String`	No	Expiration of the resulting assignment. The value is `Permanent` or an ISO 8601 UTC timestamp.

PurviewPostureAgent schema

Data Security Posture Agent events, returned in audit log searches use this schema (and also the Common schema).

PurviewPostureAgent schema

The PurviewPostureAgent audit schema is designed to capture and log activities related to the Data Security Posture Agent.

Parameters	Type	Mandatory?	Description
DiscoveryId	Edm.String	No	Unique identifier for the exploration event.
Scope	Collection(Self.PostureAgentSearchLocation)	No	Collection of search location information for the posture agent.
OldLabel	Edm.String	No	The previous label value before the change.
NewLabel	Edm.String	No	The new label value after the change.
FileName	Edm.String	No	The name of the file associated with the posture agent event.

PostureAgentSearchLocation complex type

Parameters	Type	Mandatory?	Description
Type	Edm.String	Yes	The type of the search location.
Id	Edm.String	Yes	The identifier of the search location.

Dragon Copilot Admin schema

Dragon Copilot admin events returned in audit log searches use this schema (and also the Common schema). These events have `RecordType` set to `DragonCopilotAdmin` (454).

For more information about Dragon Copilot, see the Dragon Copilot documentation.

Parameters	Type	Mandatory?	Description
EnvironmentId	Edm.String	No	The Dragon Copilot environment GUID. Enables customers to scope or filter audit logs by organizational unit. Only present for environment-scoped operations.
ProductType	Edm.String	No	The Dragon Copilot product the operation is scoped to. Possible values: `Physician`, `Nursing`, `Radiology`. Only present when the operation is product-scoped (for example, `ProvisionedProduct`, `DeprovisionedProduct`).
ModifiedProperties	Collection(Common.ModifiedProperty)	No	This property is included for certain admin events. The property includes the name of the property that was modified, the new value of the modified property, and the previous value of the modified property.

Microsoft Agent 365 schema

Microsoft Agent 365 events returned in audit log searches use the Common schema and the applicable event-specific schema described in this section.

AI Guardrail schema

AI guardrail events use this schema and the Common schema. These events have `RecordType` set to `AIGuardrail` (489) and `Operation` set to `AIGuardrailApplied`.

Parameters	Type	Mandatory?	Description
ContentInputId	Edm.String	No	The identifier of the content submitted to the guardrail for evaluation.
ContentModified	Edm.Boolean	No	Indicates whether the guardrail modified the evaluated content.

Parameters	Type	Mandatory?	Description
ContentOutputId	Edm.String	No	The identifier of the content produced after the guardrail evaluation.
ConversationId	Edm.String	No	The identifier of the conversation associated with the guardrail evaluation.
DecisionCode	Edm.String	No	The provider-specific code associated with the guardrail decision.
DecisionReason	Edm.String	No	The reason for the guardrail decision.
DecisionType	Self.DecisionType	Yes	The action selected by the guardrail after evaluating the content.
ExternalEventId	Edm.String	No	The identifier assigned to the event by the originating system or provider.
Findings	Collection(Self.AIGuardrailFinding)	No	The risks or policy findings identified during the guardrail evaluation.
GuardianId	Edm.String	Yes	The unique identifier of the guardrail that evaluated the content.
GuardianName	Edm.String	Yes	The name of the guardrail that evaluated the content.
GuardianVersion	Edm.String	No	The version of the guardrail that evaluated the content.
ParentSpanId	Edm.String	Yes	The identifier of the parent span in the distributed trace.
PlatformType	Edm.String	Yes	The platform on which the guardrail evaluation occurred.
PolicyId	Edm.String	No	The identifier of the policy applied during the evaluation.
PolicyName	Edm.String	No	The name of the policy applied during the evaluation.
PolicyVersion	Edm.String	No	The version of the policy applied during the evaluation.
ProviderName	Edm.String	Yes	The name of the provider that performed the guardrail evaluation.
SpanId	Edm.String	Yes	The identifier of the span representing the guardrail

Parameters	Type	Mandatory?	Description
			evaluation.
TargetId	Edm.String	No	The identifier of the resource or target evaluated by the guardrail.
TargetType	Edm.String	Yes	The type of resource or target evaluated by the guardrail.
TraceId	Edm.String	Yes	The identifier used to correlate the evaluation with other spans in the same distributed trace.
ContentId	Edm.String	No	A reference to the content associated with the guardrail event.

Complex Type AIGuardrailFinding

Parameters	Type	Mandatory?	Description
RiskCategory	Edm.String	Yes	The category of the risk identified by the guardrail.
RiskSeverity	Self.RiskSeverity	Yes	The severity assigned to the identified risk.
RiskScore	Edm.Double	No	The numeric score assigned to the identified risk.
RiskMetadata	Collection(Edm.String)	No	Additional metadata associated with the finding.
DecisionType	Self.DecisionType	Yes	The guardrail decision associated with the finding.
PolicyId	Edm.String	No	The identifier of the policy that generated the finding.
PolicyName	Edm.String	No	The name of the policy that generated the finding.
PolicyVersion	Edm.String	No	The version of the policy that generated the finding.

Enum: DecisionType - Type: Edm.Int32

DecisionType

Value	Member name	Description
0	block	The content or operation was blocked.
1	audit	The evaluation was recorded without blocking the content or operation.
2	allow	The content or operation was allowed.
3	warn	A warning was generated for the content or operation.
4	modify	The content was modified before it was allowed to continue.

Enum: RiskSeverity - Type: Edm.Int32

RiskSeverity

Value	Member name	Description
0	low	The finding has low severity.
1	medium	The finding has medium severity.
2	high	The finding has high severity.
3	critical	The finding has critical severity.
4	none	No risk severity was assigned.

People Skills schema

Extends the Common schema with the properties specific to People Skills (Viva) audit data. These properties apply to both the `PeopleSkillsAdmin` (476) and `PeopleSkillsUser` (477) record types. For more information about this feature, see People Skills documentation.

Parameters	Type	Mandatory?	Description
APIEndpoint	Edm.String	Yes	The originating API route or processor name for the operation.
RequestId	Edm.String	Yes	The per-request correlation id, present on every record.
SkillId	Edm.String	No	The identifier of the affected skill. Emitted for skill-scoped operations.
OperationCount	Edm.Int32	No	The count of affected items for the operation.
SourceApp	Edm.String	No	The originating client app or source of the operation.
ResponseTimeMs	Edm.Double	No	The server-side response time, in milliseconds.

Last updated on 09/23/2026

Office 365 Management Activity API FAQs and troubleshooting

Article • 04/09/2025

The Office 365 Management Activity API (also known as the *Unified Auditing API*) is a part of Office 365 security and compliance offerings, that:

- Allows programmatic access to multiple auditing pipeline workloads (such as SharePoint and Exchange)
- Is the main interface used by a variety of third-party vendor products to aggregate and index auditing data

The Management Activity API shouldn't be confused with the Office 365 Service Communications API. The Management Activity API is for auditing end user activities in the various workloads. The Service Communications API is for auditing status and messages that are sent by the services that are available in Office 365 (such as Dynamics CRM or Identity Service).

Frequently asked questions about the Office 365 Management Activity API

How do I onboard to the Management Activity API?

To get started with the Office 365 Management Activity API, see Get started with Office 365 Management APIs.

What happens if I disable auditing for my Office 365 organization? Will I still get events via the Management Activity API?

No. Office 365 unified auditing must be enabled for your organization to pull records via the Management Activity API. For instructions, see Turn audit log search on or off.

What events are audited for a specific Office 365 service?

Office 365 Management Activity API schema documentation has a comprehensive list of events. For details, see Office 365 Management Activity API schema. Also see the "Audited activities" section in Search the audit log in the Security & Compliance Center for a list of events for most of the Office 365 services that are audited.

Are there any differences in the records that are fetched by the Management Activity API versus the records that are returned by using the audit log search tool in the Microsoft

Purview compliance portal?

The data that is returned by both methods is the same. The only difference is that with the API, you can get data for the last 7 days only (more details in the questions below). When searching the audit log in the Security & Compliance Center (or by using the corresponding **Search-UnifiedAuditLog** cmdlet in Exchange Online PowerShell), you can get data for the retention period in effect when the data is generated (for example, 90 days or one year).

What is the maximum time I will have to wait before a notification is sent about a given Office 365 event?

There is no guaranteed maximum latency for notification delivery (in other words, there is no SLA). Typically, most notifications are sent within one hour of the event. Often the latency is much shorter, but this period might be longer since this varies from workload to workload.

How long before events show within the Office 365 service?

The Office 365 Audit service doesn't guarantee a specified time when events will be delivered. Audit tries to deliver data as quickly as possible. However, some issues may arise (such as server outages) upstream from the Audit service and are unavoidable. Because audit events are often used for forensic investigations, Microsoft prioritizes data completeness over latency. For core services (such as Exchange, SharePoint, OneDrive, and Teams) event availability is typically 60 to 90 minutes. For other services, event availability may be longer. While these are typical event delivery times, we acknowledge that anomalies may occur. For these reasons, Microsoft doesn't commit to a specific delivery time.

Aren't webhook notifications more immediate?

No. Recently, there have been longer wait times for notifications when using a webhook compared to querying the API directly with the `/content` operation.

How long will the content remain available to fetch via the API?

Content is available to fetch via the API for 7 days after the notification of the content availability. Even if the notification is delayed for an unusually long period (for example, in the case of a service interruption), you would still have 7 days after the first availability of the notification to download the content blob related to the originating event.

Can I query the Management Activity API for a particular event ID or RecordType or other properties in the content blob?

No. The Management Activity API provides all the event details for a particular log. It can be used to download the entire details, and then you can implement your own query logic on the downloaded data; for example, by using a custom application or a third-party tool.

How do I know the data coming from my existing auditing solution, which collects data from the Management Activity API, is accurate and complete?

The short answer is that Microsoft doesn't provide any kind of a log that allows you to cross-check any given application or third-party (ISV) application. There are other Microsoft security products that obtain their data from the same pipeline, but those products fall outside the scope of this discussion and can't be used to directly cross-check the Management Activity API. If you're concerned about discrepancies between what your existing solution is providing and what you expect, you should implement the operations illustrated above. But this can be difficult, depending on how your existing tool or solution lists and indexes data. If your existing solution only presents data sorted by the creation time of the actual event, there's no way to query the API by event creation time so that you can compare result sets. In this scenario, you'd have to collect the notified content blobs for several days, index or sort them manually, and then do a manual comparison.

What is the throttling limit for the Management Activity API?

All organizations are initially allocated a baseline of 2,000 requests per minute. The throttling limit is then adjusted based on a combination of factors, including the number of seats in the organization. Also, Office 365 E5 and Microsoft 365 E5 organizations will get approximately twice as much bandwidth as non-E5 organizations. There will also be cap on the maximum bandwidth to protect the health of the service.

> ⓘ **Note**
>
> Even though each tenant can initially submit up to 2,000 requests per minute, Microsoft cannot guarantee a response rate. The response rate depends on various factors, such as client system performance, network capacity, and network speed.

I'm encountering a throttling error in the Management Activity API. What should I do?

Open a ticket with Microsoft Support and request a new throttling limit, and include a business justification for increasing the limit. We will evaluate the request, and if accepted, we will increase the throttling limit.

Why are TargetUpdatedProperties no longer in ExtendedProperties in the audit logs for Microsoft Entra activities?

TargetUpdatedProperties were appearing in ExtendedProperties. However, they have been removed from ExtendedProperties and now appear in ModifiedProperties.

Why aren't audit logs with UserAccountNotFound "LogonError" for Microsoft Entra sign-in activities available via the Management Activity API?

Starting in November 2020, audit logs for Microsoft Entra sign-in activities are ingested into the unified audit log from Microsoft Entra Event Hubs. As a result of this change, populating the "LogonError" property with the UserAccountNotFound value is not possible. Starting the first week of February 2021, the ErrorCode property in the Microsoft Entra logon auditing schema now matches AADSTS error codes. Also, the UserId parameter will not be populated with the user name from the attempted login for UserAccountNotFound errors because that user name doesn't exist in the organization's Microsoft Entra directory.

I can see duplicate audit logs when using the Management Activity API. Why is this happening?

This duplication of events is an expected and designed behavior intended to prevent the loss of any audit events. When an error is encountered by an audit event, the audit service retries the process, which may result in duplicate audit events being generated and stored in the audit log storage(s). While exporting audit search results using the UI from the Microsoft Purview portal automatically removes duplicates, the Office 365 Management Activity API does not have this de-duplication feature. Consequently, duplicates may appear in the results returned by the Office 365 Management Activity API. It is the responsibility of the SIEM solution to implement logic to remove such duplicated events upon ingestion.

Troubleshooting the Office 365 Management Activity API

One thing that should be made clear for anyone who's getting started with the Office 365 Management Activity API is that there is no concept of querying by event specifics, such as date that the event occurred, which site collection an event might have been fired from, or the type of event. Instead, you create subscriptions to specific workloads (for example, SharePoint or Microsoft Entra ID) and each subscription is per tenant.

The following sections summarizes the most common questions that customers have in using the Office 365 Management Activity API:

- Questions about third-party tools and clients
- Enable unified audit logging in Office 365
- Connect to the API
- Check your subscriptions
- Create a new subscription
- Check content availability

- Use webhooks
- Request content blobs and throttling
- Service tags

We'll show a selection of simple PowerShell scripts that can help you answer the most common questions asked by customers or get you started implementing a custom solution by demonstrating the main operations. Not all the operations are explained in these sections, but they are all listed in Office 365 Management Activity API reference.

Questions about third-party tools and clients

The most common category of questions come from customers using third-party products to download and aggregate auditing data. Depending on the third-party product, customers may encounter difficulty with the setup or experience an interruption or an inconsistency in the data exposed in those products. Here it should be stated that the first action such customers should take is to contact their vendor's support organization. Typically, a tenant-specific vendor configuration or service problem is the cause of customer complaints.

Enable unified audit logging in Office 365

If you've just set up an app that's trying to use the Management Activity API and it's not working, be sure that you've enabled unified audit logging for your Office 365 organization. You do this by turning on the Office 365 audit log. For instructions, see Turn Office 365 audit log search on or off.

> **Note**
>
> The unified audit log configuration change can take up to 60 minutes to take effect.

If unified auditing isn't enabled, you'll typically receive an error that contains the following string: `Microsoft.Office.Compliance.Audit``.DataServiceException: Tenant <tenantID> does not exist.`

Connect to the API

Most applications connect to the API using a straightforward Client Credentials OAuth2 flow. Therefore, the first step is to create a Microsoft Entra application that has the permissions needed to access the Management Activity API data. It's outside the scope of this article to

explain the steps to create a Microsoft Entra App registration. For more information, see Register your application with your Microsoft Entra tenant.

Azure application permissions

The three permissions currently used for the Office 365 Management Activity API are:

1. Read activity data for your organization
2. Read service health information for your organization
3. Read Data Loss Prevention (DLP) policy events, including detected sensitive information

> **Note**
>
> You should enable both the Application Permissions and the Delegated Permissions for at least the first two of the above permission sets. Read DLP policy events will only be necessary if you are interested in the DLP workloads.

Get an access token

The following PowerShell script uses the App ID and a Client Secret to obtain the OAuth2 token from the Management Activity API authentication endpoint. It then places the access token into the `$headerParams` array variable, which you'll attach to your HTTP request. For the value for the API endpoint (in the $resource variable) use one of the following values based on your organization's Microsoft 365 or Office 365 subscription plan:

- Enterprise plan: `manage.office.com`
- GCC government plan: `manage-gcc.office.com`
- GCC High government plan: `manage.office365.us`
- DoD government plan: `manage.protection.apps.mil`

PowerShell

```
# Create app of type Web app / API in Azure AD, generate a Client Secret, and
update the client id and client secret here
$ClientID = "<YOUR_APPLICATION_ID"
$ClientSecret = "<YOUR_CLIENT_SECRET>"
$loginURL = "https://login.microsoftonline.com/"
$tenantdomain = "<YOUR_DOMAIN>.onmicrosoft.com"
# Get the tenant GUID from Properties | Directory ID under the Azure Active
Directory section. For $resource, use one of these endpoint values based on your
```

```
subscription plan: Enterprise - manage.office.com; GCC - manage-gcc.office.com;
GCC High: manage.office365.us; DoD: manage.protection.apps.mil
$TenantGUID = "<YOUR_TENANT_GUID>"
$resource = "https://<YOUR_API_ENDPOINT>"
# auth
$body =
@{grant_type="client_credentials";resource=$resource;client_id=$ClientID;client_se
cret=$ClientSecret}
$oauth = Invoke-RestMethod -Method Post -Uri $loginURL/$tenantdomain/oauth2/token?
api-version=1.0 -Body $body
$headerParams = @{'Authorization'="$($oauth.token_type) $($oauth.access_token)"}
```

The `$oauth` variable will contain the response object, which has several properties including the access token. Here's an example of a response:

JSON

```
token_type     : Bearer
expires_in     : 3599
ext_expires_in : 0
expires_on     : 1508285860
not_before     : 1508281960
resource       : https://manage.office.com
access_token   :
eyJ0eXAiOiJKV1QiLCJhbGciOiJSUzI1NiIsIng1dCI6IjJLVmN1enFBaWRPTHFXU2FvbDd3Z0ZSR0NZby
IsImtpZCI6IjJLVmN1enFBaWRPTHFXU2FvbDd3Z0ZSR0NZbyJ9.eyJhdWQiOiJodHRwczov…
```

Check your subscriptions

If you've experienced an interruption in data flowing to an existing Management Activity API client or solution, you might wonder if something happened to your subscription. To check your active subscriptions, add the following to the previous script:

PowerShell

```
Invoke-WebRequest -Headers $headerParams -Uri
"$resource/api/v1.0/$tenantGUID/activity/feed/subscriptions/list"
```

Sample response

JSON

```
StatusCode        : 200
Status Description : OK
Content           :
[{"contentType":"Audit.Exchange","status":"enabled","webhook":null},
{"contentType":"Audit.SharePoint","status":"enabled","webhook":
```

```
{"authId":"","address":"https://mvcwebapiwebhookreceiver.azurewebsite...
RawContent        : HTTP/1.1 200 OK
                    Pragma: no-cache
                    Content-Length: 266
                    Cache-Control: no-cache
                    Content-Type: application/json; charset=utf-8
                    Expires: -1
                    Server: Microsoft-IIS/8.5,Microsoft-IIS/8.5
                    X-AspNet-Versi...
Forms             : {}
Headers           : {[Pragma, no-cache], [Content-Length, 266], [Cache-Control,
no-cache], [Content-Type, application/json; charset=utf-8]...}
Images            : {}
InputFields       : {}
Links             : {}
ParsedHtml        : mshtml.HTMLDocumentClass
RawContentLength  : 266
```

This says that the tenant has both Audit.Exchange and Audit.SharePoint subscriptions enabled. The Exchange subscription has no webhook enabled (null) and the SharePoint subscription has a webhook enabled with the address of the registered endpoint shown.

Create a new subscription

To create a new subscription, you use the /start operation. For the API endpoint, use one of these values base on your subscription plan:

- Enterprise plan: `manage.office.com`
- GCC government plan: `manage-gcc.office.com`
- GCC High government plan: `manage.office365.us`
- DoD government plan: `manage.protection.apps.mil`

PowerShell

```
Invoke-WebRequest -Method Post -Headers $headerParams -Uri
"https://<YOUR_API_ENDPOINT>/api/v1.0/$tenantGUID/activity/feed/subscriptions/star
t?contentType=Audit.AzureActiveDirectory"
```

> **Note**
>
> Remember that `$headerParams` was populated in the first part of the script listed in the Connect to the API section in this article.

The previous code will create a new subscription to the Audit.AzureActiveDirectory content type, with a webhook that is null. You can then check your subscriptions using the code in the Check your subscriptions section in this article.

Check content availability

To check what content blobs were created during a certain period, you can add the following line to the script in the "Connecting to the API" section.

PowerShell

```
Invoke-WebRequest -Method GET -Headers $headerParams -Uri
"$resource/api/v1.0/$tenantGUID/activity/feed/subscriptions/content?
contentType=Audit.SharePoint"
```

The previous example will get all the content notifications that became available today, which means from 12:00 AM UTC to the current time. If you want to specify a different time period (keeping in mind that the maximum period for which you can query is 24 hours), add the *starttime* and *endtime* parameters to the URI; for example:

PowerShell

```
Invoke-WebRequest -Method GET -Headers $headerParams -Uri
"$resource/api/v1.0/$tenantGUID/activity/feed/subscriptions/content?
contentType=Audit.SharePoint&startTime=2017-10-13T00:00&endTime=2017-10-13T11:59"
```

> **Note**
>
> You must use either both *starttime* and *endtime* parameters or neither.

JSON

```
[{      "contentUri" :
"https://<your_API_endpoint>/api/v1.0/<your_tenant_guid>/activity/feed/audit/20171
014180051748005825$20171014180051748005825$audit_sharepoint$Audit_SharePoint",
        "contentId" :
"20171014180051748005825$20171014180051748005825$audit_sharepoint$Audit_SharePoint
",
        "contentType" : "Audit.SharePoint",
        "contentCreated" : "2017-10-13T18:00:51.748Z",
        "contentExpiration" : "2017-10-20T18:00:51.748Z",
}]
```

> **ⓘ Important**
>
> - The *contentUri* property is the URI from which you can retrieve the content blob. The blob itself is what contains the event details; it will contain details about 1 – N events. While there may be 30 JSON objects in the collection, there may be many more events detailed in those 30 content URIs.
> - The *contentCreated* property is not the date that the event being notified was created. This is the date the notification was created. The events detailed in that blob may have been created well before the content blob was created. Therefore, you can never query the API directly for events that occurred within any given period.

Page contents for busy tenants

Many larger Office 365 tenants have thousands of events being generated every hour. If this is the case with your organization, and you try to execute a query for a 24-hour period like in the example above, you may need to retrieve more notifications than can be returned in one response. In this case, you'll need to implement a logical loop of some kind, each time checking the response headers for the **NextPageUrl:** header value. See the Pagination section in the Office 365 Management Activity API reference for more details.

PowerShell

```
IF the NextPageUrl header is present in the response...

THEN issue a new request substituting the Uri parameter in the above Invoke-
WebRequest example for the value of the NextPageUrl header...

ELSE exit the loop
```

It will be difficult to test this looping code unless you have a very active tenant. In our testing, we tried to execute several thousand update operations in a script and was unable to generate a large enough number of notifications to require the **NextPageUrl** header to be sent.

Use webhooks

There two ways to get a notification that content blobs have been created. The *push* approach is implemented with a webhook endpoint, which is a web application that you create and host yourself or on a cloud platform. You register the webhook at the time you create a subscription to an audited content type. You may also add a webhook registration to an existing

subscription using the approach shown below. The *pull* approach requires you to query for a particular timespan (no more than 24 hours) using the /content operation. The response will tell you which content blobs were created during the period specified.

Before you add a webhook, be aware of the following two issues:

1. Webhooks are being de-emphasized by Microsoft because of the difficulty in debugging and troubleshooting. Typically, you would host a WebApi endpoint that responds to incoming requests. Many customers either use hosting environments (which they don't have full control over) or on-premises environments (that have difficulty allowing incoming HTTP requests). Support has seen many problems where the incoming requests from the Office 365 auditing pipeline have been blocked by a firewall or router. In this case, the API will implement a back-off algorithm, which can be confusing and may result in disabling the webhook in the subscription.

2. The webhook must be ready to immediately respond to a validation request after the start operation is executed. If there is a bug in the webhook application, the validation will fail, and the webhook will not be enabled. For information about the validation request schema, see the Webhook validation section in the Office 365 Management Activity API reference. You should seriously consider the challenges in producing a production-ready webhook to respond to Management Activity API notifications.

If you decide to implement a webhook, it should be tested to verify that the endpoint is prepared to respond with an HTTP 200 response to both the validation request and the normal notification requests before any /start operation that specifies a webhook endpoint is called for the first time. Typically, you would use a mocked request from the API to test this readiness. Carefully read and observe both the Webhook validation and Retrieving content sections in the Office 365 Management Activity API reference to understand the schema of these two types of requests.

To add a subscription with a webhook endpoint, add a body parameter to the POST to the /start endpoint; for example:

JSON

```
$body = '{
    "webhook" : {
        "address": "https://webhook.myapp.com/o365/",
        "authId": "o365activityapinotification",
        "expiration": ""
    }}'
$uri =
"https://manage.office.com/api/v1.0/$tenantGUID/activity/feed//subscriptions/start
?contentType=Audit.SharePoint"
Invoke-RestMethod -Method Post -uri $uri -Headers $headerParams -Body $body
```

Immediately following this call, a validation request will be sent out to `https://webhook.myapp.com/o365/ …` and there should be a listener ready to respond, as per the description in the Webhook validation section in the Office 365 Management Activity API reference. Your listener must respond with HTTP 200. If you immediately run the /list operation at this point, you'll not see the webhook shown as other than null until the validation has returned successfully.

Check notifications to webhooks

It's important to distinguish between the /notifications operation and the /content operation. Checking notifications is only relevant if you have subscribed with a webhook endpoint. This operation will let you know if attempts to send notifications to your webhook have been successful or not. This operation should not be used to list available content. To cross-check the availability of content against what you may have already received in your webhook, you would use the /content operation. But be sure to first check for failed notifications using the /notifications operation.

Request content blobs and throttling

After you've obtained a list of content URIs, you must request the blobs specified by the URIs. The following is an example of requesting a content blob (using the manage.office.com API endpoint for Enterprise organizations) using PowerShell. This example assumes you have already used the previous example in the Get an access token section in this article to get an access token and have populated the `$headerParams` variable appropriately.

PowerShell

```
# Get a content blob
$uri = 'https://manage.office.com/api/v1.0/<<your-tenant-
guid>>/activity/feed/audit/<<ContentID>$audit_sharepoint$Audit_SharePoint'
$contents = Invoke-WebRequest -Method GET -Headers $headerParams -Uri $uri
```

Note the following things about the *$uri* variable in the previous example:

- We used single quotes so that the *$* symbols aren't interpreted as variables in PowerShell
- The entire URI will be returned in the *contentUri* property of the response to your previous call to the /content endpoint. The placeholder tokens shown are just for illustration.

When attempting to retrieve the available content blobs, many customers (mostly working with busy tenants) have experienced errors that look like this:

JSON

```
Response Code 403: {'error':{'code':'AF429','message':'Too many requests.
Method=GetBlob, PublisherId=00000000-0000-0000-0000-000000000000'}}
```

This is likely due to throttling. Note that the value of the PublisherId parameter likely indicates that the client didn't specify the *PublisherIdentifier* in the request. Also, keep in mind that the correct parameter name is *PublisherIdentifier* even though you see *PublisherId* listed in the 403 error responses.

> **Note**
>
> In the API reference, the *PublisherIdentifier* parameter is listed in every operation of the API, but it should also be included in the GET request to the contentUri URL when retrieving the content blob.

If you're doing simple API calls to troubleshoot problems (for example, checking if a given subscription is active) you can safely omit the *PublisherIdentifier* parameter, but absolutely any code that is eventually meant for production use should include the *PublisherIdentifier* parameter on every call.

If you're implementing a client for your company's tenant, the *PublisherIdentifier* is the Tenant GUID. If you are creating an ISV application or add-in for multiple customers, the *PublisherIdentifier* should be the ISV's Tenant GUID, and not the tenant GUID of end user's company.

If you include the valid *PublisherIdentifier*, then you'll be in a pool that is allotted 60K requests per minute per tenant. This is an exceptionally large number of requests. However, if you don't include the *PublisherIdentifier* parameter, you'll be in the general pool allotted 60K requests per minute for all tenants. In this case, you'll more than likely find that your calls are getting throttled. To prevent this, here's how you would request a content blob using the *PublisherIdentifier*:

JSON

```
$contentUri = ($response.Content | ConvertFrom-Json).contentUri[0]
$uri = $contentUri + '?PublisherIdentifier=82b24b6d-0591-4604-827b-705d55d0992f'
$contents = Invoke-WebRequest -Method GET -Headers $headerParams -Uri $uri
```

The previous example assumes that the *$response* variable was populated with the response to a request to the /content endpoint and that the *$headerParams* variable includes a valid access token. The script grabs the first item in the array of content URIs from the response and then

invokes the GET to download that blob and put it in the *$contents* variable. Your code will likely loop through the contentUri collection, issuing the GET for each *contentUri*.

Service tags

The Office 365 Management Activity API and Office 365 Management Activity API Webhook now support service tags to find the required IP address prefixes that need to be allowed through the firewall. A service tag represents a pre-defined group of IP address prefixes that is managed and updated by Microsoft. Service tags help minimize the complexity for security rule creation.

The following service tags include the IP address prefixes that support the Office 365 Management Activity API and Office 365 Management Activity API Webhook:

- M365ManagementActivityApi
- M365ManagementActivityApiWebhook

For a list of the IP address prefixes that are encompassed by previous service tags, download one of the following files:

- Microsoft Azure IP Ranges and Service Tags – Public Cloud
- Microsoft Azure IP Ranges and Service Tags – US Government Cloud
- Microsoft Azure IP Ranges and Service Tags – China Cloud

After you download the appropriate file, search for the strings "M365ManagementActivityApi" and "M365ManagementActivityApiWebhook" to find the section that contains the IP address prefixes for each service tag.

Configure service tags for network security groups

You can use Az or AzureRM PowerShell cmdlets to set the network security group rule with service tag. You can configure security group rules with service tags, and then add those rules to a new network security group. A few resources, such as Azure Functions, also provide a service tag option to configure network security groups using the Azure portal

For information about using Az or AzureRM PowerShell to set up network security rules and network security groups, see:

- New-AzNetworkSecurityRuleConfig
- New-AzNetworkSecurityGroup

Here's an example of a AzureRM PowerShell script that restricts outbound traffic to all IP address prefixes except for the ones included in M365ManagementActivityApi service tag. The example script creates two outbound network security group rules. The first one allows outbound traffic to the IP address prefixes included in the M365ManagementActivityApi service tag. The second one denies outbound traffic to the Internet. The rules are then added to a network security group, which can then be configured to specific Azure services.

PowerShell

```
$allowMyServiceRule = New-AzureRmNetworkSecurityRuleConfig ` -Name
"AllowMyService"  -Access Allow  -Protocol Tcp  -Direction Outbound  -Priority 100
-DestinationAddressPrefix M365ManagementActivityApi  -SourcePortRange * -
SourceAddressPrefix * -DestinationPortRange *
$denyInternetRule = New-AzureRmNetworkSecurityRuleConfig -Name "denyInternetOut" -
Access Deny `
-Protocol Tcp -Direction Outbound  -Priority 4000 -DestinationAddressPrefix
Internet -SourcePortRange * -SourceAddressPrefix * -DestinationPortRange *
$nsg = New-AzureRmNetworkSecurityGroup `
-ResourceGroupName $resourceGroupName `
-Location $location `
-Name $nsgName `
-SecurityRules $denyInternetRule,$allowMyServiceRule
```

An inbound rule to accept traffic only from M365ManagementActivityApiWebhook can be configured in a similar way.

Communication Compliance

08/20/2025

> **Important**
>
> Microsoft Purview Communication Compliance provides the tools to help organizations detect regulatory compliance (for example, SEC or FINRA) and business conduct violations such as sensitive or confidential information, harassing or threatening language, and sharing of adult content. Communication Compliance is built with privacy by design. Usernames are pseudonymized by default, role-based access controls are built in, investigators are opted in by an admin, and audit logs are in place to help ensure user-level privacy.

Protecting sensitive information and detecting and acting on workplace harassment incidents is an important part of compliance with internal policies and standards. Microsoft Purview Communication Compliance helps minimize these risks by helping you quickly detect, capture, and take remediation actions for email, Microsoft Teams communications, Microsoft 365 Copilot, and Microsoft 365 Copilot Chat interactions. These communications can include potentially inappropriate communications containing profanity, threats, and harassment and communications that share sensitive information inside and outside of your organization.

> **Tip**
>
> If you're not an E5 customer, use the 90-day Microsoft Purview solutions trial to explore how additional Purview capabilities can help your organization manage data security and compliance needs. Start now at the Microsoft Purview trials hub. Learn details about signing up and trial terms.

Microsoft Purview Communication Compliance is an insider risk solution that helps you detect, capture, and act on inappropriate messages that can lead to potential data security or compliance incidents within your organization. Communication Compliance evaluates text and image-based messages in Microsoft and third-party apps (Teams, Microsoft 365 Copilot and Microsoft 365 Copilot Chat, Viva Engage, Outlook, WhatsApp, and others) for potential business policy violations including inappropriate sharing of sensitive information, threatening or harassing language, and potential regulatory violations such as stock and capital manipulations.

Communication Compliance's mission is to foster safe and compliant communications across customers' enterprise communication channels. With role-based access controls, human

investigators can take remediation actions such as removing a message from Teams or notifying senders of potentially inappropriate conduct.

System capabilities

> **Important**
>
> Microsoft is committed to making sure artificial intelligence (AI) systems are developed responsibly and in ways that warrant people's trust. As part of this commitment, Microsoft Purview engineering teams are operationalizing the six core principles of Microsoft's Responsible AI strategy to design, build, and manage AI solutions. As part of our effort to responsibly deploy AI, we provide documentation, role-based access, scenario attestation, and more to help organizations use AI systems responsibly.

Communication Compliance uses machine learning models and keyword matching to identify messages containing potential business conduct or regulatory policy violations that an investigator reviews. Communication Compliance cultivates user privacy with pseudonymization and responsible use of the product by providing role-based access controls.

Users get permissions that allow them to perform the tasks in their role, such as:

- **Administrator role**: Administrators configure policies to detect specific patterns in an organization. Admins choose which machine learning classifiers to include in policies and can use the value of machine learning classifiers alongside keyword and phrase matching to meet organizational needs.
- **Investigator role**: Investigators view and remediate potentially risky messages flagged by built-in classifiers. Investigators have a range of actions they can take to address the displayed message and no automatic actions are taken without manual investigator review. Investigators can share feedback on messages misclassified by Microsoft classifiers via the *Report as Misclassified* action, which aims to continually improve classifier performance through customer submitted feedback.
- **Communication Compliance role**: Communication Compliance role can create policies for the organization and review policy matches for every policy.

Intended uses of Communication Compliance

> **Important**
>
> Microsoft Purview Communication Compliance provides built-in machine learning tools for detecting messages that might violate regulatory, code of conduct, or organizational

requirements. Users can choose built-in classifiers during policy configuration to evaluate organization communications. Admins can edit policy settings at any time according to organizational needs. Built with privacy by design, Microsoft doesn't receive any customer data or evaluations unless your organization's investigators explicitly consent to the *Report as Misclassified* action.

Communication Compliance empowers organizations to detect, triage, and remediate communications with potential business conduct and regulatory compliance violations. Communication Compliance provides the following policy templates that use machine learning classifiers for users:

- **Business conduct**: Discrimination, Profanity, Threat, and Targeted harassment classifiers
- **Regulatory compliance**: Corporate sabotage, customer complaints, gifts & entertainment, money laundering, regulatory collusion, stock manipulation, unauthorized disclosure classifiers

Metrics used to evaluate and measure performance

Communication Compliance considers many metrics to holistically evaluate system accuracy and performance. False positive rates and true positive rates (also known as recall rates) from test data sets are two of the most important metrics that inform the classifier's performance. Quantitative metrics, such as the number of hits, and qualitative feedback from customers also help evaluate performance and improve the classifiers.

Operational factors and settings

Operational factors and settings allow for effective and responsible use of the system. The following table details these factors and settings:

Expand table

Operational factor	Benefits
Role-based access control	- No individual has too much power to influence usage of Communication Compliance in the organization: - Investigators can only view policies they have permission to view - Administrators can only create and edit policies - Communication Compliance role has the most widespread permissions to view and edit policies.
Using multiple conditions in a policy	You can fine-tune the value from machine learning classifiers with other capabilities such as key phrase matching, detection of sensitive

Operational factor	Benefits
	information, or scoping to specific domains.
End user reporting of inappropriate content in Teams	Message recipients can report messages they find inappropriate or deem a data security incident as a proactive approach to identifying communication risks.
Keyword highlighting of messages caught by built-in classifiers	Investigators gain clarity on which key phrases trigger which classifiers.
Pseudonymization	Sender and recipient aliases are pseudonymized to minimize investigation bias.
Users choose which communication channels to detect for patterns in	Users can choose which channels, such as Microsoft Teams, Exchange, Viva Engage, or third-party sources, to identify and detect content patterns.
Feedback loop to report misclassified items	Items misclassified by machine learning classifiers can be reported to Microsoft to improve future performance of classifiers. User privacy is honored as per the Microsoft Privacy Statement.

Limitations of Communication Compliance

Every artificial intelligence solution has limitations. You can minimize the impact of these limitations when using Communication Compliance. The following table describes these limitations.

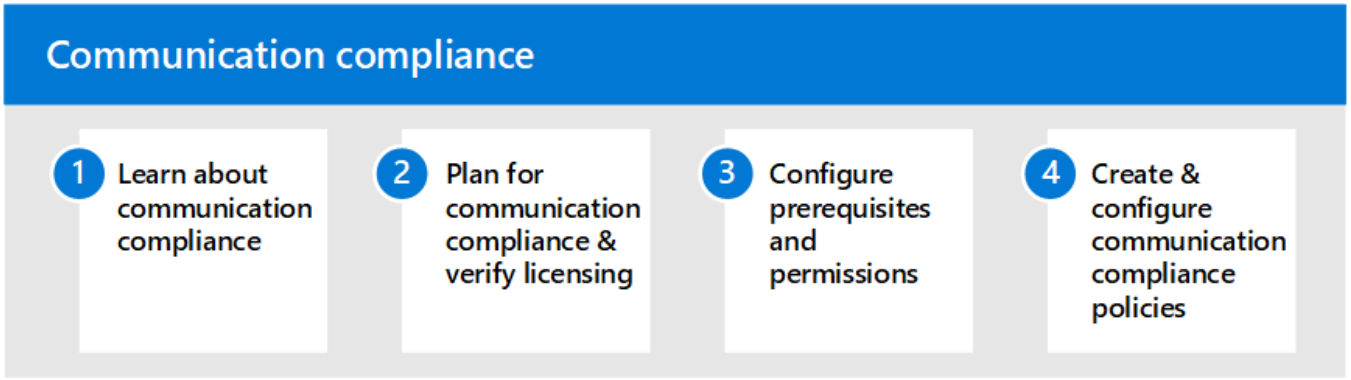
Expand table

Limitations	Impacted users	Ways to minimize impact
False positive messages	Message senders	- Communications compliance uses a feedback loop to report misclassified items to continually improve classifier performance. - Human reviewers must view message before action is taken.
False negative messages	Admins and message recipients	Machine learning classifiers can be combined with multiple conditions such as key phrase matching or user scoping to satisfy unique organizational needs.
Limited support for evasive typing and other adversarial inputs	Message recipients	Basic evasive typing coverage is provided to address swapping letters with numbers with improvements to come in future releases.
Not all languages covered	Message	12 languages supported today corresponding to highest

Limitations	Impacted users	Ways to minimize impact
	recipients	usage regions.

Configure Communication Compliance

Use the following steps to configure Communication Compliance for your organization:



1. Learn about Communication Compliance
2. Plan for Communication Compliance and verify licensing
3. Configure prerequisites and permissions
4. Create and configure Communication Compliance policies

More information about Communication Compliance

- Investigate and remediate alerts
- Case study: Contoso quickly configures an inappropriate content policy for Microsoft Teams, Exchange, and Viva Engage communications

Learn about Communication Compliance

07/18/2025

> **Important**
>
> Microsoft Purview Communication Compliance provides the tools to help organizations detect regulatory compliance (for example, SEC or FINRA) and business conduct violations such as sensitive or confidential information, harassing or threatening language, and sharing of adult content. Communication Compliance is built with privacy by design. Usernames are pseudonymized by default, role-based access controls are built in, investigators are opted in by an admin, and audit logs are in place to help ensure user-level privacy.

Microsoft Purview Communication Compliance is an insider risk solution that helps minimize communication risks by helping you detect, capture, and act on potentially inappropriate messages in your organization. Predefined and custom policies allow you to check internal and external communications for policy matches so they can be examined by designated reviewers. Reviewers can investigate email, Microsoft Teams, Microsoft 365 Copilot and Microsoft 365 Copilot Chat, Viva Engage, or third-party communications in your organization and take appropriate actions to make sure they're compliant with your organization's message standards.

Communication Compliance policies in Microsoft 365 help you overcome many modern challenges associated with compliance and internal and external communications, including:

- Checking increasing types of communication channels
- The increasing volume of message data
- Regulatory enforcement and the risk of fines

Additionally, there may be a separation of duties between your IT admins and your compliance management team. Communication Compliance supports the separation between configuration of policies and the investigation and review of messages. For example, the IT group for your organization may be responsible for setting up Communication Compliance role permissions, groups, and policies and investigators and reviewers may be responsible for message triage, review, and mitigation actions.

For more information and an overview of the planning process to address compliance and risky activities in your organization, see Starting an Insider Risk Management program.

Watch the following video to learn how to fulfill regulatory compliance requirements with Communication Compliance:

https://learn-video.azurefd.net/vod/player?id=bf50d9ad-7ecb-4d3c-b0ec-0d9b1f3f7dcd&locale=en-us&embedUrl=%2Fpurview%2Fcommunication-compliance

> **Important**
>
> Communication Compliance is currently available in tenants hosted in geographical regions and countries supported by Azure service dependencies. To verify that Communication Compliance is supported for your organization, see Azure dependency availability by country/region.

> **Tip**
>
> If you're not an E5 customer, use the 90-day Microsoft Purview solutions trial to explore how additional Purview capabilities can help your organization manage data security and compliance needs. Start now at the Microsoft Purview trials hub. Learn details about signing up and trial terms.

Scenarios for Communication Compliance

Communication Compliance policies can assist with reviewing messages in your organization in several important compliance areas:

- **Corporate policies**

  Users must comply with acceptable use, ethical standards, and other corporate policies in all their business-related communications. Communication Compliance policies can detect policy matches and help you take corrective actions to help mitigate these types of incidents. For example, you could check user communications in your organization for human resources concerns such as harassment or the use of potentially inappropriate or offensive language.

- **Risk management**

  Organizations are responsible to all communications distributed throughout their infrastructure and corporate network systems. Using Communication Compliance policies to help identify and manage potential legal exposure and risk can help minimize risks before they can damage corporate operations. For example, you could check messages in your organization for unauthorized communications and conflicts of interest about confidential projects such as upcoming acquisitions, mergers, earnings disclosures, reorganizations, or leadership team changes.

- **Regulatory compliance**

  Most organizations must comply with some type of regulatory compliance standards as part of their normal operating procedures. These regulations often require organizations to implement some type of scoping or oversight process for messaging that is appropriate for their industry. The Financial Industry Regulatory Authority (FINRA) Rule 3110 is a good example of a requirement for organizations to have scoping procedures in place to check user communications and the types of businesses in which it engages. Another example may be a need to review broker-dealer communications in your organization to safeguard against potential insider trading, collusion, or bribery activities. Communication Compliance policies can help your organization meet these requirements by providing a process to both analyze and report on corporate communications. For more information on support for financial organizations, see Key compliance and security considerations for US banking and capital markets.

Key feature areas

Communication Compliance offers several important features to help address compliance concerns on your messaging platforms:

- Intelligent customizable templates
- Flexible remediation workflows
- Actionable insights

Intelligent customizable templates

Intelligent customizable templates in Communication Compliance allow you to apply machine learning to intelligently detect communication violations in your organization.

- **Customizable pre-configured templates**: Policy templates help address the most common communications risks. Initial policy creation and follow-on updating are now quicker with predefined templates to analyze and mitigate potentially inappropriate content, sensitive information, conflict of interest, and regulatory compliance issues.
- **New machine learning support**: Built-in classifiers to analyze and mitigate discrimination, threats, harassment, profanity, and potentially inappropriate images and help reduce misclassified content in communication messages, saving reviewers time during the investigation and remediation process.
- **Improved condition builder**: Configure policy conditions that are now streamlined into a single, integrated experience in the policy workflow, reducing confusion in how conditions are applied for policies.

Flexible remediation workflows

Built-in remediation workflows allow you to quickly identify and take action on messages with policy matches in your organization. The following new features increase efficiency for investigation and remediation activities:

- **Flexible remediation workflow**: New remediation workflow helps you quickly take action on policy matches, including new options to escalate messages to other reviewers and to send email notifications to users with policy matches.
- **Conversation policy matching**: Messages in conversations are grouped by policy matches to give you more visibility about how conversations relate to your communication policies. For example, conversation policy matching in the **Pending** tab will automatically show all messages in a Teams channel that have matches for your communications policies for analyzing and mitigating potentially inappropriate messages. Other messages in conversations that don't match your communications policies wouldn't be displayed.
- **Keyword highlighting**: Terms matching policy conditions are highlighted in the message text view to help reviewers quickly analyze and remediate policy alerts.
- **Optical character recognition (OCR)**: Check, detect, and investigate printed and handwritten text within images embedded or attached to email or Microsoft Teams chat messages.
- **New filters**: Investigate and remediate policy alerts faster with message filters for several fields, including sender, recipient, date, domains, and many more.
- **Improved message views**: Investigation and remediation actions are now quicker with new message source and text views. Message attachments are now viewable to provide complete context when taking remediation actions.
- **User history**: Historical view of all user message remediation activities, such as past notifications and escalations for policy matches, now provides reviewers with more context during the remediation workflow process. First-time or repeat instances of policy matches for users are now archived and easily viewable.
- **Pattern detected notification**: Many harassing and bullying actions take place over time and involve reoccurring instances of the same behavior by a user. The pattern detected notification displayed in alert details helps raise attention to these alerts and this type of behavior.
- **Translation**: Quickly investigate message details in eight languages using translate support in the remediation workflow. Messages in other languages are automatically converted to the display language of the reviewer.
- **Attachment detection**: Check, detect, and investigate linked content (Modern attachments) from OneDrive and Microsoft Teams that match policy classifiers and conditions for Microsoft Teams messages. Attachment content is automatically extracted to a text file for detailed review and action.

- **Summarize message content for policy matches**: Save time for investigators by using Microsoft Copilot in Microsoft Purview to summarize lengthy Teams, email, or Viva Engage messages. Copilot in Microsoft Purview creates a summary of the conversation, including recordings, meeting transcripts, and attachments.

Actionable insights

New interactive dashboards for alerts, policy matches, actions, and trends help you quickly view the status of pending and resolved alerts in your organization.

- **Proactive intelligent alerts**: Alerts for policy matches requiring immediate attention include new dashboards for pending items sorted by severity and new automatic email notifications sent to designated reviewers.
- **Interactive dashboards**: New dashboards display policy matches, pending and resolved actions, and trends by users and policy.
- **Auditing support**: A full log of policy and review activities is easily exported from the Microsoft Purview portal to help support audit review requests.

Integration with Microsoft 365 services

Communication Compliance policies check, detect, and capture messages across several communication channels to help you quickly review and remediate compliance issues:

- **Generative AI**: Use Communication Compliance policies to analyze interactions (prompts and responses) entered into generative AI applications to help detect inappropriate or risky interactions or sharing of confidential information. Coverage is supported for Microsoft 365 Copilot, Copilots built using Microsoft Copilot Studio, AI applications connected by Microsoft Entra or Microsoft Purview Data Map connectors, and more.
- **Microsoft Teams**: Chat communications for public and private Microsoft Teams channels and individual chats are supported in Communication Compliance as a standalone channel source or with other Microsoft 365 services. You can also detect communications included in meetings transcripts (preview). You need to manually add individual users, distribution groups, or specific Microsoft Teams channels when you select users and groups to apply a Communication Compliance policy to. Teams users can also self-report potentially inappropriate messages in private and group channels and chats for review and remediation.
- **Exchange Online**: All mailboxes hosted on Exchange Online in your Microsoft 365 organization are eligible for analyses. Emails and attachments matching Communication Compliance policy conditions are instantly available for investigation and in compliance reports. Exchange Online is now an optional source channel and is no longer required in Communication Compliance policies.

- **Microsoft 365 Copilot and Microsoft 365 Copilot Chat**: Interactions (prompts and responses) entered by users into Copilot are detected by Communication Compliance policies.
- **Viva Engage**: Private messages and public community conversations in Viva Engage are supported in Communication Compliance policies. Viva Engage is an optional channel and must be in native mode to support checking of messages and attachments.
- **Third-party sources**: You can check messages from third-party sources for data imported into mailboxes in your Microsoft 365 organization. Communication Compliance supports connections to several popular platforms, including Instant Bloomberg and others.

To learn more about messaging channel support in Communication Compliance policies, see Detect channel signals with Communication Compliance.

Watch the video below to learn how to detect communication risks in Microsoft Teams with Communication Compliance:

https://learn-video.azurefd.net/vod/player?id=7857b91f-d0c1-4b9f-ac84-5c70a8453216&locale=en-us&embedUrl=%2Fpurview%2Fcommunication-compliance

Integration with Microsoft Purview Insider Risk Management

When users experience employment stressors, they may engage in risky activities. Workplace stress can lead to uncharacteristic or malicious behavior by some users that could surface as potentially inappropriate behavior on your organization's messaging systems. Counterproductive work behavior can be a pre-cursor to more serious violations, such as sabotaging company assets or leaking sensitive information. By integrating Communication Compliance with Microsoft Purview Insider Risk Management, you can detect for stressors that indicate an unhealthy workplace environment.

Learn more about integrating Communication Compliance with Insider Risk Management

Get started with recommended actions

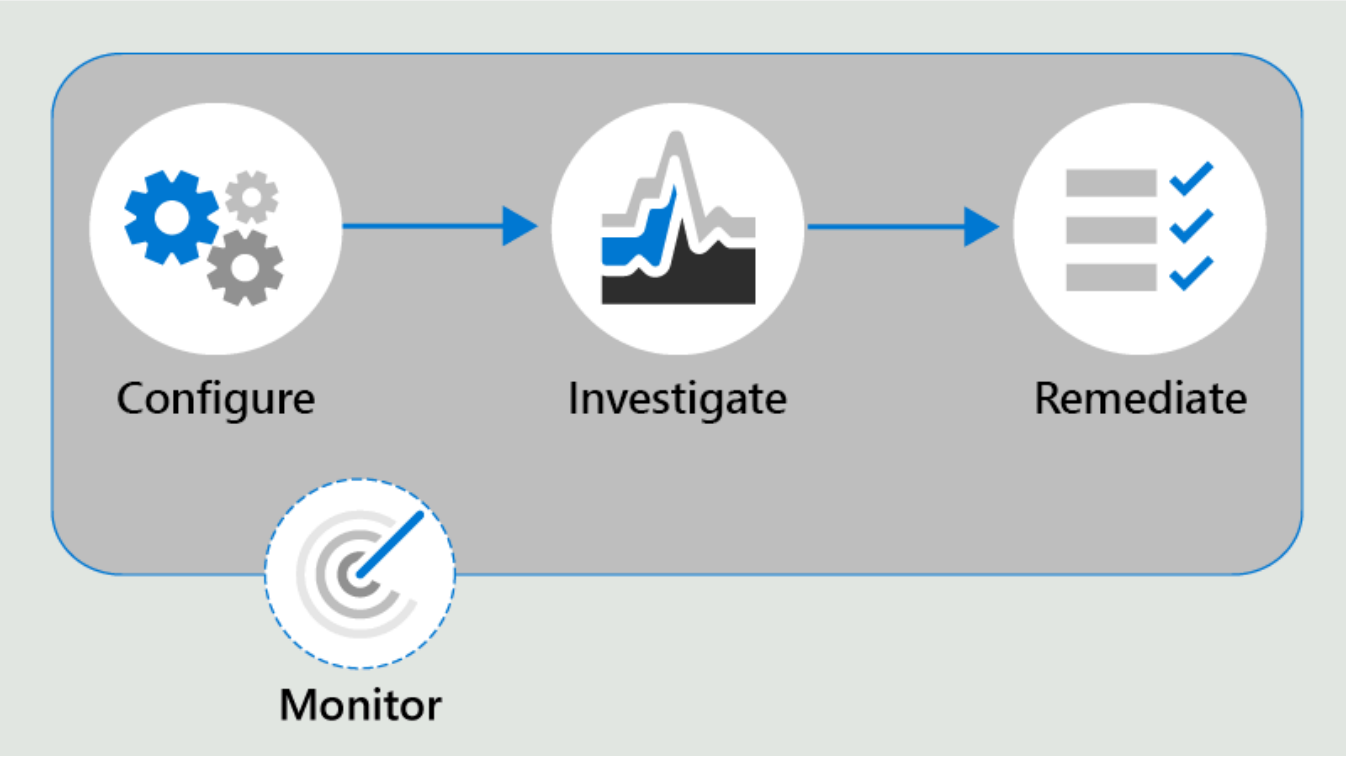
Whether you're setting up Communication Compliance for the first time or getting started with creating new policies, the new recommended actions experience can help you get the most out of Communication Compliance capabilities. Recommended actions include setting up permissions, creating distribution groups, creating policies, and more.

Workflow

Communication Compliance helps you address common pain points associated with complying with internal policies and regulatory compliance requirements. With focused policy templates and a flexible workflow, you can use actionable insights to quickly resolve detected compliance issues.

Before you create a policy, you can decide whether you want to apply an **adaptive scope**. For more information, see Adaptive policy scopes for compliance solutions. If you decide to create an adaptive policy, you must create one or more adaptive scopes before you create your policy, and then select them during the create new policy process. For instructions, see Configuration information for adaptive scopes.

Identifying and resolving compliance issues with Communication Compliance uses the following workflow:



Configure

In this workflow step, you identify your compliance requirements and configure applicable Communication Compliance policies. Policy templates are a great way to not only quickly configure a new compliance policy, but to also quickly modify and update policies as your requirements change. For example, you may want to quickly test a policy for potentially inappropriate content on communications for a small group of users before configuring a policy for all users in your organization.

> ⓘ **Important**

By default Global Administrators do not have access to Communication Compliance features. To enable permissions for Communication Compliance features, see Make Communication Compliance available in your organization.

You can choose from the following policy templates in the Microsoft Purview portal:

- **Detect inappropriate text**: Use this template to quickly create a policy that uses built-in classifiers to automatically detect text in messages that may be considered inappropriate, abusive, or offensive.
- **Detect inappropriate images**: Use this template to quickly create a policy that uses built-in classifiers to automatically detect content that contains adult and racy images that may be considered as inappropriate in your organization.
- **Detect sensitive info types**: Use this template to quickly create a policy to check communications containing defined sensitive information types or keywords to help make sure that important data isn't shared with people that shouldn't have access.
- **Detect financial regulatory compliance**: Use this template to quickly create a policy to check communications for references to standard financial terms associated with regulatory standards.
- **Detect conflict of interest**: Use this template to quickly create a policy to detect communications between two groups or two users to help avoid conflicts of interest.
- **Custom policy**: Use this template to configure specific communication channels, individual detection conditions, and the amount of content to detect and review in your organization.
- **User-reported messages policy**: This system policy supports user reported messages from channel, group, and private chat messages. Enabled by default in the Teams admin center.

Tip

Use recommended actions to help you determine if you need a sensitive information type policy or if you need to update existing inappropriate content policies.

Investigate

In this step, you can look deeper into the issues detected as matching your Communication Compliance policies. This step includes the following actions available in the Microsoft Purview portal:

- **Alerts**: When a group of messages matches a policy condition, an alert is automatically generated. For each alert, you can see the status, the severity, the time detected, and if an

eDiscovery (Premium) case is assigned and its status. New alerts are displayed on the Communication Compliance home page and the **Alerts** page and are listed in order of severity.

- **Issue management**: For each alert, you can take investigative actions to help remediate the issue detected in the message.
- **Document review**: During the investigation of an issue, you can use several views of the message to help properly evaluate the detected issue. The views include a conversation summary, text-only, and detail views of the communication conversation.
- **Reviewing user activity history**: View the history of user message activities and remediation actions, such as past notifications and escalations, for policy matches.
- **Filters**: Use filters such as sender, recipient, date, and subject to quickly narrow down the message alerts that you want to review.

Remediate

The next step is to remediate Communication Compliance issues you've investigated using the following options:

- **Resolve**: After reviewing an issue, you can remediate by resolving the alert. Resolving an alert removes it from the **Pending** tab, and the action is preserved as an entry on the **Resolved** tab for the matching policy. Alerts are automatically resolved after marking the alert as misclassified, sending a notice to a user about the alert, or opening a new case for the alert.
- **Tag a message**: As part of the resolution of an issue, you can tag a policy match as *Compliant*, *Non-compliant*, or *Questionable* as it relates to the policies and standards for your organization. You can also create a custom tag. Tagging can help you micro-filter policy alerts for escalations or as part of other internal review processes. You can filter on any tag value.
- **Notify the user**: Often, users accidentally or inadvertently violate a Communication Compliance policy. You can use the notify feature to provide a warning notice to the user and to resolve the issue.
- **Escalate to another reviewer**: Sometimes, the initial reviewer of an issue needs input from other reviewers to help resolve the incident. You can easily escalate message issues to reviewers in other areas of your organization as part of the resolution process.
- **Report as misclassified**: Messages incorrectly detected as matches of compliance policies will occasionally slip through to the review process. You can mark these types of alerts as misclassified, submit feedback to Microsoft about the misclassification to help improve global classifiers, and automatically resolve the issue.
- **Remove message in Teams**: Potentially inappropriate messages may be removed from displaying in Microsoft Teams channels or personal and group chat messages. Those

identified messages that are removed are replaced with a notification that the message has been removed for a policy violation.

- **Escalate for investigation**: In the most serious situations, you may need to share Communication Compliance information with other reviewers in your organization. Communication Compliance is tightly integrated with other Microsoft Purview features to help you with end-to-end risk resolution. Escalating a case for investigation allows you to transfer data and management of the case to Microsoft Purview eDiscovery (Premium). eDiscovery (Premium) provides an end-to-end workflow to preserve, collect, review, analyze, and export content that's responsive to your organization's internal and external investigations. It allows legal teams to manage the entire legal hold notification workflow. To learn more about eDiscovery (Premium) cases, see Overview of Microsoft Purview eDiscovery (Premium).

Maintain

Keeping track and mitigating compliance issues identified by Communication Compliance policies spans the entire workflow process. As alerts are generated and investigation and remediation actions are implemented, existing policies may need review and updates, and new policies may need to be created.

- **Review and report**: Use Communication Compliance dashboard widgets, export logs, and events recorded in the unified audit logs to continually evaluate and improve your compliance posture.
- **Get insights on policy health**: Communication Compliance provides warnings and recommendations to improve policy health. Learn more about policy health

Ready to get started?

- For planning information, see Plan for Communication Compliance.
- Check out the case study for Contoso and see how they quickly configured a Communication Compliance policy to detect potentially inappropriate content in Microsoft Teams, Exchange Online, and Viva Engage communications.
- To configure Communication Compliance for your Microsoft 365 organization, see Configure Communication Compliance.

More resources

For the latest Ignite videos for Communication Compliance, see the following:

- Foster a culture of safety and inclusion with Communication Compliance
- Learn how to reduce communication risks within your organization

- Better with Microsoft Teams - Learn more about the latest native Teams integrated features in Communication Compliance

For a quick overview of Communication Compliance, see the Detect workplace harassment and respond with Communication Compliance video on the Microsoft Mechanics channel.

Check out how TD Securities is using Communication Compliance to address their regulatory obligations and meet their security and stability needs.

Check out the Microsoft Mechanics video on how Insider Risk Management and Communication Compliance work together to help minimize data risks from users in your organization.

To keep up with the latest Communication Compliance updates, select **What's new** in the Communication Compliance solution for your organization.

Plan for Communication Compliance

08/20/2025

> **Important**
>
> Microsoft Purview Communication Compliance provides the tools to help organizations detect regulatory compliance (for example, SEC or FINRA) and business conduct violations such as sensitive or confidential information, harassing or threatening language, and sharing of adult content. Communication Compliance is built with privacy by design. Usernames are pseudonymized by default, role-based access controls are built in, investigators are opted in by an admin, and audit logs are in place to help ensure user-level privacy.

Before you get started with Communication Compliance in your organization, your information technology and compliance management teams should review important planning activities and considerations. Thoroughly understanding and planning for deployment in the following areas helps ensure that your implementation and use of Communication Compliance features goes smoothly and aligns with the best practices for the solution.

To learn how to fulfill regulatory compliance requirements with Communication Compliance, watch the following video:

https://learn-video.azurefd.net/vod/player?id=bf50d9ad-7ecb-4d3c-b0ec-0d9b1f3f7dcd&locale=en-us&embedUrl=%2Fpurview%2Fcommunication-compliance-plan

For more information and an overview of the planning process to address compliance and risky activities in your organization, see Starting an Insider Risk Management program.

To see how Insider Risk Management and Communication Compliance work together to help minimize data risks from users in your organization, watch the Microsoft Mechanics video.

> **Important**
>
> Communication Compliance is currently available in tenants hosted in geographical regions and countries supported by Azure service dependencies. To verify that Communication Compliance is supported for your organization, see Azure dependency availability by country/region.

> **Tip**

If you're not an E5 customer, use the 90-day Microsoft Purview solutions trial to explore how additional Purview capabilities can help your organization manage data security and compliance needs. Start now at the **Microsoft Purview trials hub**. Learn details about **signing up and trial terms**.

Work with stakeholders in your organization

Identify the appropriate stakeholders in your organization to collaborate for taking actions on Communication Compliance alerts. Consider including the following stakeholders in your initial planning and the end-to-end Communication Compliance workflow:

- Information technology
- Compliance
- Privacy
- Security
- Human resources
- Legal

Plan for the investigation and remediation workflow

Select dedicated stakeholders to investigate and review the alerts and cases on a regular cadence in the Microsoft Purview portal. Make sure you understand how you assign users and stakeholders to different Communication Compliance role groups in your organization.

> **Important**
>
> After configuring your role groups, it might take up to 30 minutes for the role group permissions to apply to assigned users across your organization.

Configure permissions

Six role groups configure initial permissions to manage Communication Compliance features. To make **Communication Compliance** available as a menu option in Microsoft Purview portal and to continue with these configuration steps, you must be assigned to one those groups. For more information, see Assign permissions in Communication Compliance.

Scoped users

Before you start using Communication Compliance, determine who needs their communications reviewed. In the policy, user email addresses identify individuals or groups of people to apply the policy to. Some examples of these groups are Microsoft 365 Groups, Exchange-based distribution lists, Viva Engage communities, and Microsoft Teams channels. You can also exclude specific users or groups from checking with a specific exclusion group or a list of groups. For more information about groups types supported in Communication Compliance policies, see Get started with Communication Compliance.

> **Important**
>
> Users covered by Communication Compliance policies must have either a Microsoft 365 E5 Compliance license, an Office 365 Enterprise E3 license with the Advanced Compliance add-on, or be included in an Office 365 Enterprise E5 subscription. If you don't have an existing Enterprise E5 plan and want to try Communication Compliance, you can **sign up for a trial of Office 365 Enterprise E5**.

Reviewers

When you create a Communication Compliance policy, you decide who reviews the messages from the scoped users. In the policy, user email addresses identify individuals or groups of people who review scoped communications. All reviewers must have mailboxes hosted on Exchange Online, be assigned to either the *Communication Compliance Analysts* or *Communication Compliance Investigators* role groups, and be assigned in the policy they need to investigate. When you add reviewers to a policy, they automatically receive an email message that notifies them of the assignment to the policy and provides links to information about the review process.

Groups for scoped users and reviewers

To simplify your setup, we recommend creating groups for people who need their communications reviewed and groups for people who review those communications. If you use groups, you might need several. For example, you might want to identify communications between two distinct groups of people, or you might want to specify a group that isn't in scope. When you assign a Distribution group in the policy, the policy detects all emails from each user in the Distribution group. When you assign a Microsoft 365 group in the policy, the policy detects all emails sent to that group, not the individual emails received by each group member.

> **Note**

Before you create a policy, decide whether you want to apply an **adaptive scope** for users or groups. For more information, see Adaptive policy scopes for compliance solutions.

Adding groups and distribution lists to Communication Compliance policies are part of the overall conditions and rules set, so the maximum number of groups and distribution lists that a policy supports varies depending on the number of conditions you also add to the policy. Each policy supports approximately 20 groups or distribution lists, depending on the number of additional conditions present in the policy.

The following chart can help you configure groups in your organization for Communication Compliance policies:

Expand table

Policy Member	Supported Groups	Unsupported Groups
Scoped users Excluded users	Distribution groups Microsoft 365 Groups	Dynamic distribution groups Nested distribution groups Mail-enabled security groups Microsoft 365 groups with dynamic membership
Reviewers	None	Distribution groups Dynamic distribution groups Nested distribution groups Mail-enabled security groups

Privacy

Protecting the privacy of users that have policy matches is important and can help promote objectivity in data investigation and analysis reviews for Communication Compliance alerts. This setting applies only to user names displayed the Communication Compliance solution. It doesn't affect how names are displayed in other compliance solutions or admin center.

For users with a Communication Compliance match, you can choose one of the following settings in **Communication Compliance settings**:

- **Show anonymized versions of usernames**: User names are anonymized to prevent users in *Communication Compliance Analysts* role group from seeing who is associated with policy alerts. Users in the *Communication Compliance Investigators* role group always see user names, not the anonymized versions. For example, a user 'Grace Taylor' appears with a randomized pseudonym such as 'AnonIS8-988' in all areas of the Communication Compliance experience. Choosing this setting anonymizes all users with current and past policy matches and applies to all policies. User profile information in the Communication

Compliance alert details isn't available when you choose this option. However, user names are displayed when adding new users to existing policies or when assigning users to new policies. If you turn off this setting, user names are displayed for all users that have current or past policy matches.

- **Do not show anonymized versions of usernames**: User names are displayed for all current and past policy matches for Communication Compliance alerts. User profile information (the name, title, alias, and organization or department) is displayed for the user for all Communication Compliance alerts.

Plan for Communication Compliance policies

Creating Communication Compliance policies is quick and easy with the predefined templates for analyzing potentially inappropriate content, sensitive information, and regulatory compliance issues. Custom Communication Compliance policies give you the flexibility to detect and investigate issues specific to your organization and requirements.

When planning for Communication Compliance policies, consider the following areas:

- Consider adding all users in your organization as in-scope for your Communication Compliance policies. Identifying specific users as in-scope for individual policies can be useful in some circumstances. However, most organizations should include all users in Communication Compliance policies optimized for harassment or discrimination detection.
- Decide whether you want to apply an adaptive scope to your Communication Compliance policy. For more information, see Adaptive policy scopes for retention. Creating multiple policies can result in higher administrative overhead.
- Configure the percentage of communications to review at 100% to ensure that policies catch all issues of concern in communications for your organization.
- You can analyze communications from non-Microsoft sources for data imported into mailboxes in your Microsoft 365 organization. To include review of communications in these platforms, you need to configure a third-party connector to these services before messages meeting policy conditions are detected by a communication policy.
- Policies can support detecting languages other than English in custom Communication Compliance policies. Build a custom keyword dictionary of offensive words in the language of your choice or build your own machine learning model by using trainable classifiers in Microsoft 365.
- All organizations have different communication standards and policy needs. Detect specific keywords by using Communication Compliance policy conditions or detect specific types of information with custom sensitive information types.

Migrating between Microsoft 365 US Government Cloud and the commercial cloud

If you migrate your organization from the Microsoft 365 US Government Cloud to the worldwide commercial cloud or from the worldwide commercial cloud to the Government Cloud, active cases and alerts aren't migrated. Close any alerts and cases before starting the migration.

Create a Communication Compliance policy walkthrough

Want to see an in-depth walkthrough of setting up a new Communication Compliance policy and remediating an alert? Check out the following 15-minute video to see a demonstration of how Communication Compliance policies can help you detect potentially inappropriate messages, investigate potential violations, and remediate compliance issues.

https://learn-video.azurefd.net/vod/player?id=643aa064-e8bd-4b0e-b3ad-f58ab1585626&locale=en-us&embedUrl=%2Fpurview%2Fcommunication-compliance-plan

Ready to get started?

To configure Communication Compliance for your Microsoft 365 organization, see Configure Communication Compliance or check out the case study for Contoso and how they quickly configured a Communication Compliance policy to detect potentially inappropriate content in Microsoft Teams, Exchange Online, and Viva Engage communications.

Get started with Communication Compliance

09/26/2025

> **Important**
>
> Microsoft Purview Communication Compliance provides the tools to help organizations detect regulatory compliance (for example, SEC or FINRA) and business conduct violations such as sensitive or confidential information, harassing or threatening language, and sharing of adult content. Communication Compliance is built with privacy by design. Usernames are pseudonymized by default, role-based access controls are built in, investigators are opted in by an admin, and audit logs are in place to help ensure user-level privacy.

Use Communication Compliance policies to identify user communications for analysis by internal or external reviewers. For more information about how Communication Compliance policies can help you detect communications in your organization, see Communication Compliance policies. To see how Contoso quickly configured a Communication Compliance policy to detect potentially inappropriate content in Microsoft Teams, Exchange Online, and Viva Engage communications, check out this case study.

Subscriptions and billing

Before getting started with Communication Compliance, confirm your Microsoft 365 subscription and any add-ons. To access and use Communication Compliance, administrators need to verify that their organization has a supported subscription and that users have the appropriate licenses assigned. For more information about subscriptions and licensing, see the subscription requirements for Communication Compliance.

Additionally, to detect inappropriate or risky interactions for non-Microsoft 365 AI data, you need to enable pay-as-you-go billing in your organization. Non-Microsoft 365 AI data includes information from other generative AI applications from Microsoft and connected external AI applications. This data type includes Copilot in Microsoft Fabric, Microsoft Security Copilot, Microsoft Copilot Studio, and any connected or cloud AI application. There aren't any pay-as-you-go billing requirements or charges for Microsoft 365 detecting inappropriate or risky interaction for Microsoft 365 Copilot data.

> **Important**

Communication Compliance is currently available for tenants hosted in geographical regions and countries supported by Azure service dependencies. To verify that Communication Compliance is supported for your organization, see Azure dependency availability by country/region.

Recommended actions

Recommended actions can help your organization quickly get started with Communication Compliance. Included on the **Overview** page, recommended actions help guide you through the steps to configure and deploy policies.

The following recommendations help you get started and maximize your Communication Compliance configuration:

- **Get to know Communication Compliance**: Before completing set up, review our official documentation to learn about, plan for, and deploy Communication Compliance in your organization.
- **Assign permissions to ensure your team can get their jobs done**: Ensure that only the appropriate stakeholders can access the solution by assigning team members responsible for managing Communication Compliance features and investigating and reviewing alerts.
- **Create distribution groups for users' whose communications you want to detect**: Create distribution groups containing users who are included in Communication Compliance policies.
- **Create your first policy to start detecting communications**: Detect and investigate potential regulatory compliance violations by first setting up a policy that identifies potential violations across your organization's internal and/or external communications.
- **Review alerts to investigate detected messages and take action**: Identify and analyze messages that match a policy's conditions to trigger alerts that provide context around a policy violation, so you can investigate and take action if needed.
- **Review reports for quick insights into how policies are performing**: Get quick insights into how your policies are performing, view detailed reports to drill down further, and export results for further analyses.

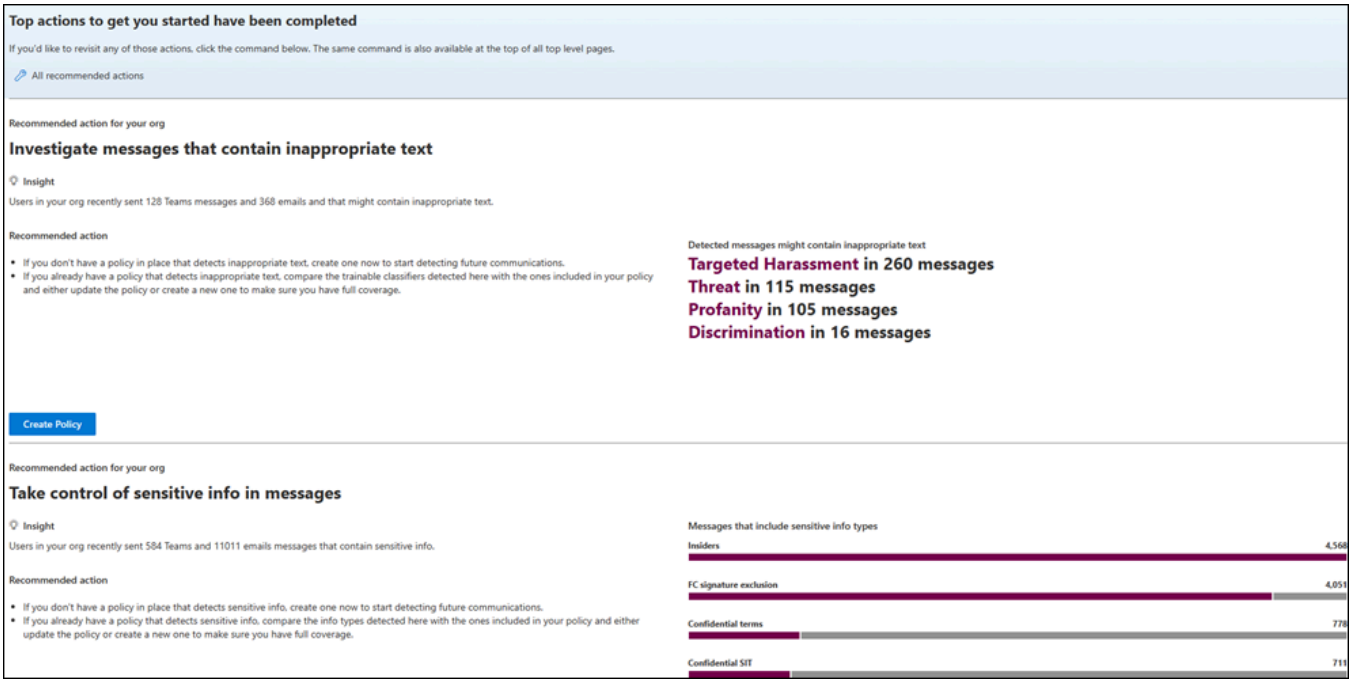
Each action in Communication Compliance has three attributes:

- **Action**: The name and description of the recommended action.
- **Recommended, required, or optional**: Whether the recommended action is highly recommended, required, or optional for Communication Compliance features to function as expected.

- **Estimated time to complete**: Estimated time to complete the recommended action in minutes.

Select recommendations from the list to get started with configuring Communication Compliance. Each recommended action guides you through the required activities for the recommendation, including any requirements, what to expect, and the impact of configuring the feature in your organization. Some recommended actions are automatically marked as complete when configured. If not, manually select the action as complete when configured.

Recommended actions insights on the **Policies** page help summarize current sensitive information types and potential regulatory compliance violations in communications in your organization. Data classification and the application of sensitivity labels, retention labels, and sensitive information type classification support these insights. These insights are aggregated and don't include any personal data for users in your organization.



Activity in messages aggregates by trainable classifier type from existing policies that use the *Detect inappropriate text* policy template or custom policies that use classifiers. Investigate alerts for these messages on the **Alert dashboard** for your policies.

Activity involving sensitive information types detects in messages covered in existing policies and for messages that aren't covered by existing policies. Insight messages that aren't covered by existing policies can't be investigated and remediated, a new policy must be created to detect and remediate similar activity in future messages. Insights aggregate for all sensitive information types, including ones that your organization didn't previously define in an existing Communication Compliance policy. Use these insights to create a new Communication Compliance policy or to update existing policies. After creating a new policy, message alerts for this policy might or might not match an equal number of messages identified in a similar

insight. Your policy might have different conditions, a different number of in-scope users, and only detects message activity that occurs after the policy is active.

> **Tip**
>
> To disable the display of recommended action insights, open a request with Microsoft Support.

Step 1 (required): Assign permissions for Communication Compliance

To use any of the Communication Compliance-related tools in the Microsoft Purview portal, users need the appropriate permissions. The easiest way to assign roles is to add the user to the appropriate role group on the **Role groups** page in the Microsoft Purview portal.

For step-by-step guidance, see Assign permissions in Communication Compliance.

Step 2 (required): Enable the audit log

Communication Compliance requires audit logs to show alerts and log remediation actions that reviewers take. The audit logs summarize all activities associated with a defined organizational policy or any time a Communication Compliance policy changes.

Auditing is enabled for Microsoft 365 organizations by default. Some organizations might disable auditing for specific reasons. If auditing is disabled for your organization, another administrator might have turned it off. We recommend confirming that it's OK to turn auditing back on when you complete this step.

For step-by-step instructions to turn on auditing, see Turn audit log search on or off. After you turn on auditing, a message is displayed that says the audit log is being prepared and that you can run a search in a couple of hours after the preparation is complete. You only need to perform this action once. For more information about using the audit log, see Search the audit log.

Step 3 (optional): Set up groups for Communication Compliance

When you create a Communication Compliance policy, you define who has their communications reviewed and who performs reviews. In the policy, use email addresses to

identify individuals or groups of people. To simplify your setup, you can create groups for people who have their communication reviewed and groups for people who review those communications. If you use groups, you might need several. For example, you might need multiple groups if you want to detect communications between two distinct groups of people or if you want to specify a group that isn't going to be scoped.

Use the following chart to help you configure groups in your organization for Communication Compliance policies:

Expand table

Policy Member	Supported Groups	Unsupported Groups
Scoped users Excluded users	Distribution groups Microsoft 365 Groups Mail-enabled security groups	Dynamic distribution groups Shared mailbox Nested distribution groups Microsoft 365 Groups with dynamic membership
Reviewers	None	Distribution groups Dynamic distribution groups Nested distribution groups Mail-enabled security groups

When you assign a *distribution group* in the policy, the policy detects all emails and Teams chats from each user in the *distribution group*. When you assign a *Microsoft 365 group* in the policy, the policy detects all emails and Teams chats sent to the *Microsoft 365 group*,* not the individual emails and chats received by each group member*. We recommend using distribution groups in Communication Compliance policies so that individual emails and Teams chats from each user are automatically detected.

> **Tip**
>
> For a more flexible configuration and to reduce administrative overhead, consider using an **adaptive scope** instead of a distribution group.

If you're an organization with an Exchange on-premises deployment or an external email provider and you want to detect Microsoft Teams chats for your users, you must create a distribution group for the users with on-premises or external mailboxes. Later in these steps, you assign this distribution group by using the **Choose users and groups** selection in the policy workflow. For more information about the requirements and limitations for enabling cloud-based storage and Teams support for on-premises users, see Search for and export Teams chat data for on-premises users.

To manage scoped users in large enterprise organizations, you might need to detect messages for all users across large groups. You can use PowerShell to configure a distribution group for a global Communication Compliance policy for the assigned group. This configuration enables you to detect messages for thousands of users with a single policy and keep the Communication Compliance policy updated as new employees join your organization.

1. Create a dedicated distribution group for your global Communication Compliance policy with the following properties: Make sure that you don't use this distribution group for other purposes or other Office 365 services.

   - **MemberDepartRestriction = Closed**. Ensures that users can't remove themselves from the distribution group.
   - **MemberJoinRestriction = Closed**. Ensures that users can't add themselves to the distribution group.
   - **ModerationEnabled = True**. Ensures that all messages sent to this group are subject to approval and that the group isn't being used to communicate outside of the Communication Compliance policy configuration.

   PowerShell

   ```
   New-DistributionGroup -Name <your group name> -Alias <your group alias> -
   MemberDepartRestriction 'Closed' -MemberJoinRestriction 'Closed' -
   ModerationEnabled $true
   ```

2. Select an unused Exchange custom attribute to log users added to the Communication Compliance policy in your organization.

3. Run the following PowerShell script on a recurring schedule to add users to the Communication Compliance policy:

   PowerShell

   ```
   $Mbx = (Get-Mailbox -RecipientTypeDetails UserMailbox -ResultSize Unlimited -
   Filter {CustomAttribute9 -eq $Null})
   $i = 0
   ForEach ($M in $Mbx)
   {
     Write-Host "Adding" $M.DisplayName
     Add-DistributionGroupMember -Identity <your group name> -Member
   $M.DistinguishedName -ErrorAction SilentlyContinue
     Set-Mailbox -Identity $M.Alias -<your custom attribute name> SRAdded
     $i++
   }
   Write-Host $i "Mailboxes added to supervisory review distribution group."
   ```

For more information about setting up groups, see:

- Create and manage distribution groups
- Overview of Microsoft 365 Groups

Step 4 (optional): Verify your Viva Engage tenant is in Native Mode

In Native Mode, all Viva Engage users are in Microsoft Entra ID, all groups are Microsoft 365 Groups, and all files are stored in SharePoint Online. Your Viva Engage tenant must be in Native Mode for Communication Compliance policies to check and identify risky conversations in private messages and community conversations in Viva Engage.

For more information about configuring Viva Engage in Native Mode, see:

- Overview of Viva Engage Native Mode in Microsoft 365
- Configure your Viva Engage network for Native Mode for Microsoft 365

Step 5 (required): Create a Communication Compliance policy

You can quickly create a Communication Compliance policy by choosing from several policy templates, or you can create a custom policy.

> **Important**
>
> If your role is scoped by one or more admin units, you can't create a policy based on a template. You can create a custom policy, however.

When you create a policy from a template, many settings are automatically chosen for you based on the template you choose. As the final step, you can customize the policy if you want to change any of the settings.

When you create a custom policy, you select all of the settings yourself.

Create a policy from a template

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.

2. Go to the **Communication Compliance** solution.

3. Select **Policies** in the left navigation.

4. Select **Create policy**, then select one of the policy templates. For example, select **Detect inappropriate text**.

5. In the pane on the right side of the screen:
   a. In the **Policy name** field, confirm or update the policy name. You can't change policy names after you create the policy.
   b. In the **Users or groups in scope** field, choose the users or groups to apply the policy to. If you select the *Detect conflict of interest* template, select two scoped groups or two scoped users to detect internal communications.
   c. In the **Reviewers** field, choose the reviewers for the policy. Reviewers that you add in this field are the reviewers that you can choose from when escalating an alert in the investigation and remediation workflow. When you add reviewers to a policy, they automatically receive an email message that notifies them of the assignment to the policy and links to information about the review process. **Reviewers are individual users and all reviewers must have mailboxes hosted on Exchange Online.**
   d. Review the list of settings that are automatically chosen based on the template you choose. To customize any settings, select **Customize policy**, then make the changes you want. If everything looks OK, select **Create policy**.

> ⓘ **Note**
>
> To enable optical character recognition (OCR) to identify embedded or attached images in messages for printed or handwritten text that match policy conditions, select **Customize policy**, then on the **Choose conditions and percentage** page, select the **Use OCR to extract text from images** checkbox.

Create a custom policy

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.

2. Go to the **Communication Compliance** solution.

3. Select **Policies** in the left navigation.

4. Select **Create policy**, then select **Custom policy**.

5. On the **Name and describe your policy** page, enter a name (required) and description (optional) for the policy. You can't change the policy name after you create it. Select **Next** when you're done.

6. If one or more admin units exist for your organization, you see the **Admin units (preview)** page. Otherwise, you see the **Choose users and reviewers** page (next step) and a banner with a link to learn more about admin units.

   To scope the policy to one or more admin units, select **Add admin units**, select the admin units to apply to the policy, then select **Save**.

   > ⓘ **Note**
   >
   > You can only see the admin units that are scoped to your role. If you're an unrestricted administrator, you can see all admin units for the organization. To view a summary of the role groups and admin units that you're assigned to, select **View my permissions**.

   Select **Next** when you're done with the **Admin units (preview)** page.

7. On the **Choose users and reviewers** page:
   a. In the **Choose users and groups** section, select one of the following options:

   - **All users**: This option is the most inclusive and recommended for the broadest protection.

     > ⓘ **Note**
     >
     > If the policy is scoped by one or more admin units, selecting this option selects all users in the admin units.

   - **Select users**: Start typing to find specific users or groups. The **Show insights and recommendations for users who match this policy's conditions but weren't included in the policy** checkbox appears when you select this option. Leave this checkbox selected if you want to reduce blind spots by receiving recommendations when users outside of the selected users match the policy conditions. This setting isn't available if you choose the **All users** option or the **Select adaptive scopes** option.

     > ⓘ **Note**
     >
     > If the policy is scoped by one or more admin units, you can only select users that are part of those admin units.

- **Select adaptive scopes**: An adaptive scope uses a query that you specify to define the membership of users or groups. To use an adaptive policy, create the adaptive scope before you create your policy, then select the scope when you choose this option. Site scopes aren't applicable to Communication Compliance so you won't see them when you add a scope. Learn more about the advantages of using an adaptive scope.

  > ⓘ Note
  >
  > If the policy is scoped by one or more admin units the **Select adaptive scopes** option doesn't appear. At this time, you can't use adaptive scopes together with admin units.

a. In the **Excluded users and groups** section, add any users or groups to exclude from the policy.

> ⓘ Note
>
> If the policy is scoped by one or more admin units, you can only exclude users and groups that are part of those admin units.

b. In the **Reviewers** section, choose the reviewers for the policy. Reviewers you add in this section are the reviewers you can choose when escalating an alert in the investigation and remediation workflow. When you add reviewers to a policy, they automatically receive an email message that notifies them of the assignment to the policy and links to information about the review process. **Reviewers are individual users and all reviewers must have mailboxes hosted on Exchange Online.**

c. Select **Next** when you're done.

8. On the **Choose locations to detect communications** page, choose the communication channels to check, including **Exchange**, **Teams**, **Viva Engage**, or a specific generative AI channel. Supported generative AI channels include **Microsoft Copilot experiences**, **Enterprise AI apps**, and **Other AI apps**. You can also choose to check third-party sources if you configured a connector in Microsoft 365. Select **Next** to move to the next page.

9. On the **Choose conditions and review percentage** page:
   a. In the **Communication direction** section, choose the communication direction to detect, including inbound, outbound, or internal communications.
   b. In the **Conditions** section, add the conditions you want to detect. You can choose from many different conditions for messages and message attachments, including

conditions that detect for:

- Default or custom sensitive information types. You can create sensitive info types before running the workflow or create them from the workflow.
- Default or custom keyword dictionaries.
- Microsoft provided trainable classifiers. Microsoft provided trainable classifiers can detect potentially inappropriate language and images sent or received in the body of email messages or other types of text. You can choose from the following built-in Microsoft provided trainable classifiers: *Targeted threat*, *Profanity*, *Targeted harassment*, *Adult images*, *Racy images*, and *Gory images*. Communication Compliance also includes content safety classifiers (preview) for Microsoft Teams that are based on large language models (LLMs). These classifiers include *Hate*, *Sexual*, *Violence*, and *Self-harm*. Learn more about content safety classifiers based on large language models.

a. In the **Optical character recognition (OCR)** section, to enable optical character recognition (OCR) to identify embedded or attached images in messages for printed or handwritten text that match policy conditions, select the **Use OCR to extract text from images** checkbox. One or more conditional settings associated with text, keywords, Microsoft provided trainable classifiers, or sensitive info types must be configured in the policy to enable the selection of the checkbox.
b. In the **Review percentage** section, move the slider to change the amount of content to review.
c. In the **Filter email blasts** section, the **Filter out messages from email blasting services** check box is selected by default. This setting causes messages sent from email blast services to be excluded. Messages that match specific conditions won't generate alerts. This exclusion includes bulk email (such as newsletters), spam, phishing, and malware. You can view a report containing the bulk email senders that are filtered out. The list of senders is filtered before the content is analyzed so there might be senders that don't match the content conditions (extra senders included in the report).
d. Select **Next** to move to the next page.

10. On the **Review and finish** page, review your policy selections, then select **Create policy**.

Notes and tips on creating Communication Compliance policies

- After configuring a policy, learn about best practices for managing the volume of alerts.
- You can't use PowerShell to create and manage Communication Compliance policies. You must use the Communication Compliance solution.

- For an in-depth walkthrough of setting up a new Communication Compliance policy and remediating an alert, see this 15-minute video:

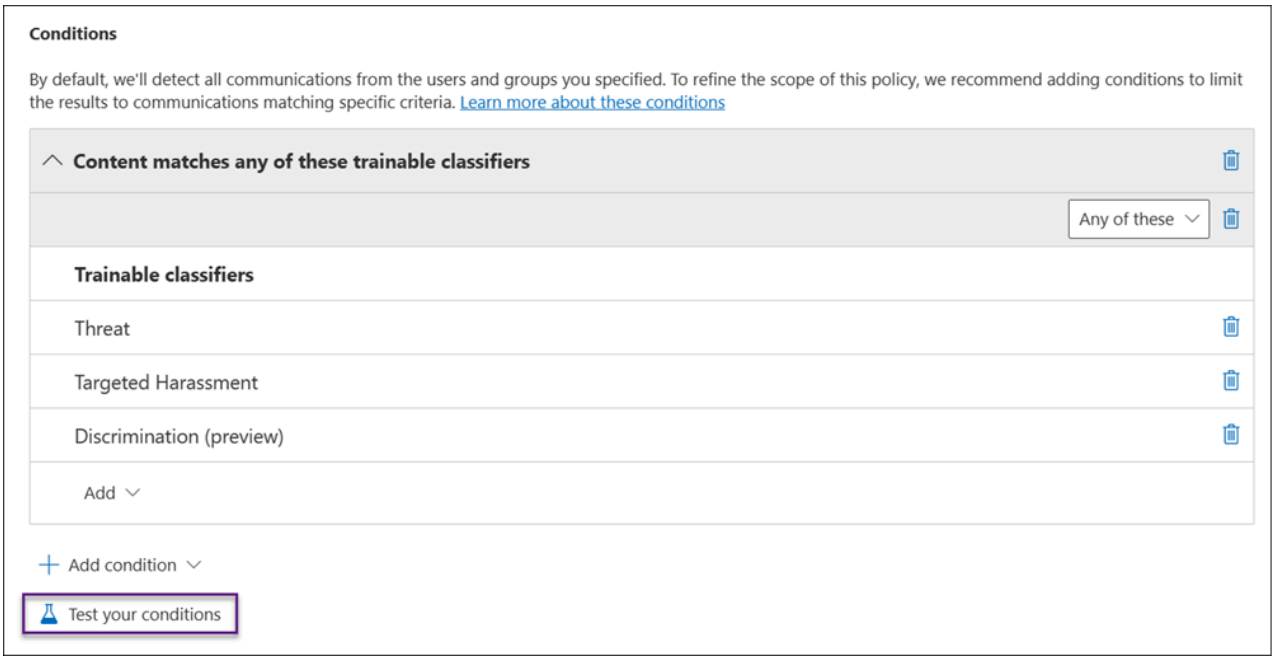
  https://learn-video.azurefd.net/vod/player?id=643aa064-e8bd-4b0e-b3ad-f58ab1585626&locale=en-us&embedUrl=%2Fpurview%2Fcommunication-compliance-configure

Test conditions (preview) when you create or edit a policy

If you create or edit a policy without testing conditions first, you might wait a long time before you can verify that the policy works as intended. Typically, you need to create a pilot policy and test it for several days before rolling the policy out to the wider organization. If you're a member of the Communication Compliance Admins role group or the Communication Compliance role group, you can save time by testing your conditions when you create or edit a policy. You can fine-tune the conditions to make sure the policy works as intended before rolling it out to the wider organization.

> **Note**
>
> Currently, you can only test the following conditions:
>
> - *Content matches any of these classifiers*
> - *Content contains any of these sensitive info types*

1. Make sure you're a member of the Communication Compliance Admins role group or the Communication Compliance role group. You must be a member of one of these role groups to use this feature.
2. Create a policy or edit an existing policy.
3. On the **Conditions and percentage** page, after entering your conditions, select **Test your conditions**.



> **Tip**
>
> You can also access the **Test your conditions** command from the list of policies on the **Policies** page or in the policy details panel if you choose to edit the policy.

4. In the **Test policy conditions** pane that appears on the right side of the page, do one of the following actions:

   - Select the **Enter messages to test** option, then enter some messages that you expect the policy to detect. Separate messages with a comma.
   - If you have a .txt file that includes a list of messages to detect, select the **Upload a file to test whether the trainable classifiers detects the matching elements you specified** option, then select **Upload file** to upload your text file.

5. Select **Test** to see a list of test results.

> **Tip**
>
> If you don't see the results you're looking for:
>
> - Check the definition of the sensitive info type or the trainable classifier to make sure it's designed to detect the type of info you expect.
> - Make sure that the message meets minimum word count requirements.

Step 6 (optional): Update compliance boundaries for Communication Compliance policies

Compliance boundaries create logical boundaries within an organization that control the user content locations (such as mailboxes, OneDrive accounts, and SharePoint sites) that eDiscovery managers can search.

If you configure compliance boundaries in your organization, update the compliance boundaries to allow certain users access to mailboxes that support Communication Compliance policies. You need to allow access to Communication Compliance administrators and Communication Compliance reviewers for your policy management and investigation and remediation actions to work properly.

To allow access for Communication Compliance admins and reviewers, run the following PowerShell commands. You only need to run these commands once, even if you add new Communication Compliance policies in the future:

PowerShell

```
Import-Module ExchangeOnlineManagement
$UserCredential = Get-Credential
Connect-IPPSSession -Credential $UserCredential
New-ComplianceSecurityFilter -FilterName "CC_mailbox" -Users <list your
Communication Compliance admins and reviewers user alias or email address> -
Filters "Mailbox_Name -like 'SupervisoryReview{*'" -Action All
```

For more information about cmdlet syntax, see New-ComplianceSecurityFilter.

Step 7 (optional): Create notice templates and configure user anonymization

If you want to respond to a policy alert by sending a reminder notice to the associated user, create at least one notice template in your organization. The notice template fields are editable before they're sent as part of the alert remediation process. We recommend creating a customized notice template for each Communication Compliance policy.

You can also choose to enable anonymization for displayed usernames when investigating policy matches and taking action on messages.

Create templates and configure user anonymization

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.
2. Go to the **Communication Compliance** solution.
3. Select **Settings** in the upper-right corner of the page, select **Communication Compliance**, select the **Privacy** tab, select **Show anonymized versions of usernames**, and then select **Save**.
4. Select the **Notice templates** tab, then select **Create notice template**.
5. In the right pane, complete the following fields:
    - Template name (required)
    - Send from (required)
    - Cc and Bcc (optional)
    - Subject (required)
    - Message body (required)
6. Select **Save** to create and save the notice template.

Step 8 (optional): Test your Communication Compliance policy

After you create a Communication Compliance policy, test it to make sure that the policy properly enforces the conditions you defined. If your Communication Compliance policies include sensitive information types, you might also want to test your Microsoft Purview Data Loss Prevention (DLP) policies. Make sure you give your policies time to activate so that the communications you want to test are captured.

> **Tip**
>
> To save time, you can test the following conditions before creating your policy:
>
> - *Content matches any of these classifiers*
> - *Content contains any of these sensitive information types*
>
> Learn more about testing conditions before creating a policy.

Follow these steps to test your Communication Compliance policy:

1. Open an email client, Microsoft Teams, or Viva Engage while signed in as a scoped user defined in the policy you want to test.

2. Send an email, Microsoft Teams chat, or Viva Engage message that meets the criteria you defined in the Communication Compliance policy. This test can be a keyword, attachment size, domain, and more. Make sure you determine if your configured conditional settings in the policy are too restrictive or too lenient.

   > **Note**
   >
   > Email messages can take approximately 24 hours to fully process in a policy. Communications in Microsoft Teams, Viva Engage, and third-party platforms can take approximately 48 hours to fully process in a policy.

3. Sign in to Microsoft 365 as a reviewer designated in the Communication Compliance policy. Go to **Communication Compliance** > **Alerts** to view the alerts for your policies.

4. Remediate the alert by using the remediation controls and verify that the alert is properly resolved.

Next steps

After you complete these steps to create your first Communication Compliance policy, you start to receive alerts from activity indicators after 24-48 hours. Configure additional policies as needed by using the guidance in Step 5 of this article.

To learn more about investigating Communication Compliance alerts, see Investigate and remediate Communication Compliance alerts.

To keep up with the latest Communication Compliance updates, select **What's new** in Communication Compliance for your organization.

Assign permissions in Communication Compliance

08/20/2025

Use six role groups to configure initial permissions for managing Communication Compliance features. To make **Communication Compliance** available as a menu option in Microsoft Purview portal and to continue with these configuration steps, you must be assigned to one of the following roles or role groups:

- Microsoft Entra ID *Global Administrator* role
- Microsoft Entra ID *Compliance Administrator* role
- Microsoft Purview portal *Organization Management* role group
- Microsoft Purview portal *Compliance Administrator* role group
- *Communication Compliance* role group
- *Communication Compliance Admins* role group

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview **roles and permissions**.

Members of the following roles have the same solution permissions included with the *Communication Compliance Admins* role group:

- Microsoft Entra ID *Global Administrator*
- Microsoft Entra ID *Compliance Administrator*
- Microsoft Purview portal *Organization Management*
- Microsoft Purview portal *Compliance Administrator*

> **Important**
>
> Make sure you always have at least one user in the *Communication Compliance* or *Communication Compliance Admins* role groups (depending on the option you choose) so that your Communication Compliance configuration doesn't get in to a 'zero administrator' scenario if specific users leave your organization.

Depending on how you want to manage Communication Compliance policies and alerts, assign users to specific role groups to manage different sets of Communication Compliance features. You can assign users with different compliance responsibilities to specific role groups to manage

different areas of Communication Compliance features. Or you might decide to assign all user accounts for designated administrators, analysts, investigators, and viewers to the *Communication Compliance* role group. Use a single role group or multiple role groups to best fit your compliance management requirements.

> **Important**
>
> After configuring your role groups, it might take up to 30 minutes for the role group permissions to apply to assigned users across your organization.

Choose from these solution role group options when configuring and managing Communication Compliance:

Expand table

Actions	Communication Compliance	Communication Compliance Admins	Communication Compliance Analysts	Communication Compliance Investigators	Communication Compliance Viewers
Access and investigate alerts	Yes	No	Yes	Yes	No
Access reports	Yes	No	No	No	Yes
Configure policies and settings	Yes	Yes	No	No	No
Create message details report	Yes	No	No	Yes	No
Manage privacy settings and notice templates	Yes	Yes	No	No	No
Take advanced remediation actions: - Escalate for investigation - Remove	Yes	No	No	Yes	No

Actions	Communication Compliance	Communication Compliance Admins	Communication Compliance Analysts	Communication Compliance Investigators	Communication Compliance Viewers
message in Teams - Download items and reports - Run Power Automate flows					
View **Conversation** and **Translation** tabs for a specific message	Yes	No	No	Yes	No
View and export policy updates	Yes	Yes	No	No	Yes

Consider admin units if you want to scope user permissions to a region or department

You can use administrative units in Communication Compliance to scope user permissions to a particular geography or department. For example, a global company that has subsidiaries throughout the world might want to create an admin unit that provides a German scope for investigators so that they only see user activity for German users.

To use admin units in Communication Compliance, you must first create the admin units (if they aren't already created), then assign the admin units to members of role groups. After you assign admin units to members of role groups, those members become *restricted administrators* and have limited access to Communication Compliance settings, policies, and user data in the organization. Members who aren't assigned administrative units are *unrestricted administrators* and have access to all settings, policies, and user data.

> **ⓘ Important**
>
> At this time, you can't use administrative units together with adaptive scopes in Communication Compliance. SharePoint sites and inactive mailboxes can only be segmented through adaptive scopes.

Effect of admin unit scoping on Communication Compliance roles

The following table shows how admin units, when enforced, affect each combination of Communication Compliance task and role.

> **Note**
>
> *Scoped*, in the following table, means that the admin actions for that role are limited by their assigned admin unit.

Expand table

Task	Scoped Communication Compliance	Scoped Communication Compliance Admins	Scoped Communication Compliance Analysts	Scoped Communication Compliance Investigators	Scoped Communication Compliance Viewers
Access and investigate alerts	Scoped	No	Scoped	Scoped	No
Access reports	No	No	No	No	No
Configure policies	Scoped	Scoped	No	No	No
Configure settings (including notice templates)	No	No	No	No	No
View and export audit logs	No	No	No	No	No

Option 1: Assign all compliance users to the Communication Compliance role group

1. Sign in to the Microsoft Purview portal with an admin account in your Microsoft 365 organization.
2. Select **Settings** in the upper-right corner of the page, select **Roles and groups**, then select **Role groups** in the left navigation pane.
3. Select the *Communication Compliance* role group, then select **Edit**.

4. Select **Choose users**, then select the checkboxes for all the users you want to add to the role group.
5. Select **Select**, then select **Next**.
6. Select **Save** to add the users to the role group, then select **Done**.

Option 2: Assign users to specific Communication Compliance role groups

Use this option to assign users to specific role groups to segment Communication Compliance access and responsibilities among different users in your organization.

1. Sign in to the Microsoft Purview portal with an admin account in your Microsoft 365 organization.
2. Select **Settings** in the upper-right corner of the page, then select **Role groups** in the left navigation pane.
3. Select one of the Communication Compliance role groups, then select **Edit**.
4. Select **Choose users**, then select the checkboxes for all the users you want to add to the role group.
5. Select **Select**, then select **Next**.
6. Select **Save** to add the users to the role group.
7. Select the next Communication Compliance role group, then repeat the previous steps for each required role group.
8. Select **Close** when you're done.

For more information about role groups and permissions, see Permissions in the Microsoft Purview portal.

Case study - Contoso configures a Communication Compliance policy to identify potentially inappropriate text

> **Important**
>
> Microsoft Purview Communication Compliance provides the tools to help organizations detect regulatory compliance (for example, SEC or FINRA) and business conduct violations such as sensitive or confidential information, harassing or threatening language, and sharing of adult content. Communication Compliance is built with privacy by design. Usernames are pseudonymized by default, role-based access controls are built in, investigators are opted in by an admin, and audit logs are in place to help ensure user-level privacy.

Microsoft Purview Communication Compliance helps you minimize communication risks by helping you detect, capture, and act on messages with potentially inappropriate text in your organization. Potentially inappropriate text might include profanity, threats, harassment, and adult content. Predefined and custom policies allow you to review internal and external communications for policy matches, so designated reviewers can examine them. Reviewers can investigate alerts for email, Microsoft Teams, Viva Engage, or third-party communications throughout your organization and take appropriate remediation actions to make sure they're compliant with your organization's message standards.

The Contoso Corporation is a fictional organization that needs to quickly configure a policy to detect potentially inappropriate text. They use Microsoft 365 primarily for email, Microsoft Teams, and Viva Engage support for their users, but have new requirements to enforce company policy around workplace harassment. Contoso IT administrators and compliance specialists have a basic understanding of the fundamentals of working with Microsoft 365 and are looking for end-to-end guidance for how to quickly get started with Communication Compliance.

This case study covers the basics for quickly configuring a Communication Compliance policy to detect potentially inappropriate text. This guidance includes:

- Step 1: Planning for Communication Compliance
- Step 2: Accessing Communication Compliance
- Step 3: Configuring prerequisites and creating a Communication Compliance policy
- Step 4: Investigate and remediate alerts

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Step 1: Planning for Communication Compliance

Contoso IT administrators and compliance specialists attend online webinars about compliance solutions in Microsoft Purview and decide that Communication Compliance policies can help them meet the updated corporate policy requirements for reducing workplace harassment. Working together, they develop a plan to create and enable a Communication Compliance policy to detect potentially inappropriate messages. This configuration includes detecting text for chats sent in Microsoft Teams, private messages and community conversations in Viva Engage, and in email messages sent in Exchange Online.

Their plan includes identifying the:

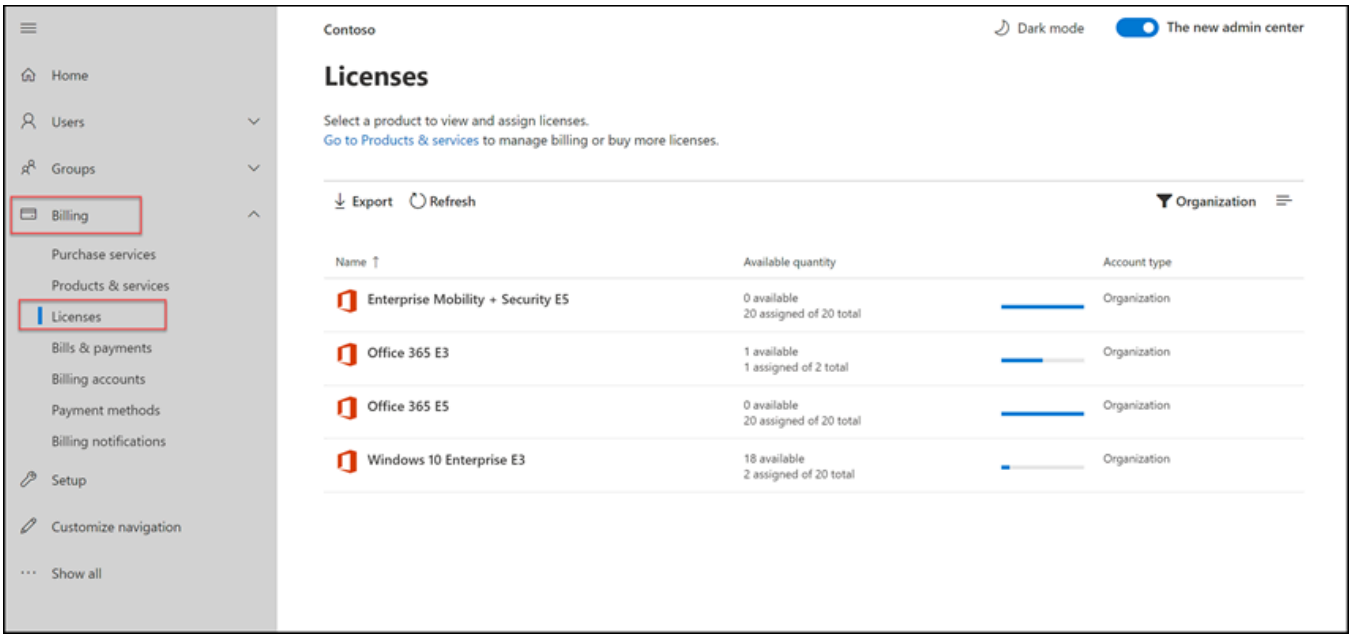
- IT administrators who need access to Communication Compliance features.
- Compliance specialists who need to create and manage Communication Compliance policies.
- Compliance specialists and other colleague in other departments (Human Resources, Legal, etc.) who need to investigate and remediate Communication Compliance alerts.
- Users who are in-scope for the Communication Compliance potentially inappropriate text policy.

Licensing

First, confirm whether Contoso's Microsoft 365 licensing includes support for the Communication Compliance solution. To access and use Communication Compliance, Contoso IT administrators need to verify Contoso has a supported subscription and assigned the appropriate licenses to users. For more information about subscriptions and licensing, see the subscription requirements for Communication Compliance.

Contoso IT administrators take the following steps to verify the licensing support for Contoso:

1. IT administrators sign in to the Microsoft 365 admin center and go to Microsoft 365 admin center > **Billing** > **Licenses**.
2. They confirm that they have one of the license options that includes support for Communication Compliance.



Permissions for Communication Compliance

Five role groups used to configure permissions manage Communication Compliance features. To make **Communication Compliance** available as a menu option in Microsoft Purview portal and to continue with these configuration steps, assign Contoso administrators the *Communication Compliance Admins* role.

Contoso decides to use the *Communication Compliance* role group and assign all the Communication Compliance administrators, analysts, investigators, and viewers to the group. This role group configuration makes it easier for Contoso to get started quickly and best fits their compliance management requirements.

Use the Microsoft Purview portal to select the *Communication Compliance* role group.

1. Contoso IT admins sign in to the Microsoft Purview portal with credentials for an admin account in their Microsoft 365 organization.
2. The administrators select **Settings** in the upper-right corner of the page, then select **Role groups** in the left navigation pane.
3. The administrators select the *Communication Compliance* role group, then select **Edit**.
4. The administrators select **Choose users**, then select the checkboxes for all the Contoso users who manage Communication Compliance and investigate and review alerts.
5. The administrators choose **Select**, then choose **Next** to review the role group changes.
6. They select **Save** to add Contoso users to the role group.

Step 2: Accessing Communication Compliance

After configuring the permissions for Communication Compliance, Contoso IT administrators and compliance specialists assigned to the Communication Compliance role group can access

the Communication Compliance solution in Microsoft Purview. Contoso IT administrators and compliance specialists have a couple of ways to access Communication Compliance and get started creating a new policy:

- Starting from the Microsoft Purview portal
- Starting directly from the Communication Compliance solution

Starting from the Microsoft Purview portal

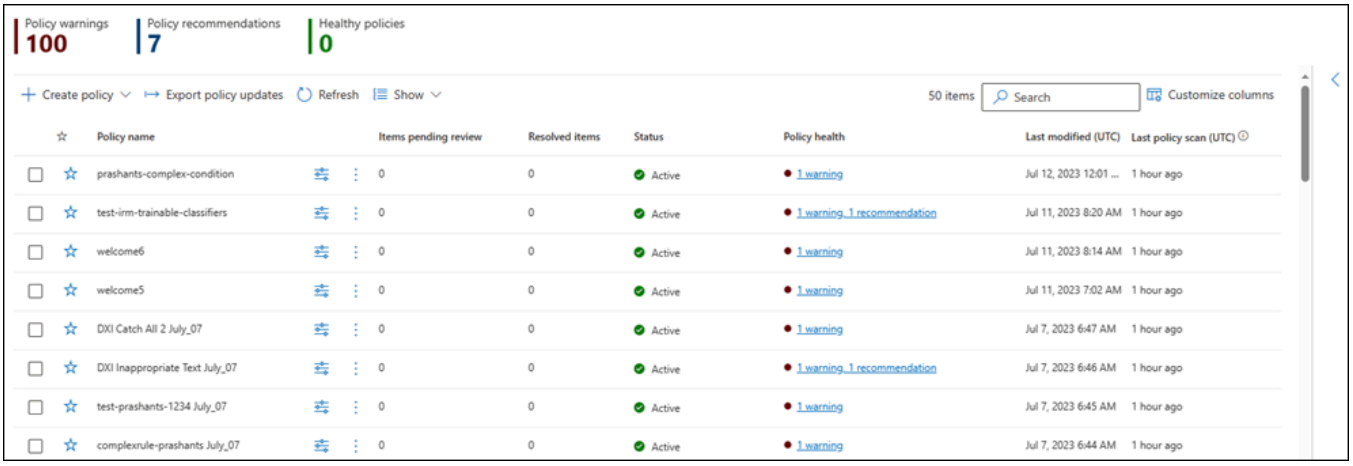
Admins and compliance specialists can sign in to the Microsoft Purview portal and then go to the Communication Compliance solution.

1. Sign in to the Microsoft Purview portal.
2. Select the **Communication Compliance** solution.

Starting directly from the Communication Compliance solution

To quickly access the solution, sign in directly. Sign directly into Communication Compliance in the Microsoft Purview portal.

When they select one of these links, Contoso IT administrators and compliance specialists go to the Communication Compliance page where they can quickly review the status of alerts and create new policies from the predefined templates.



Step 3: Configuring prerequisites and creating a Communication Compliance policy

To get started with a Communication Compliance policy, Contoso IT administrators need to configure several prerequisites before setting up the new policy to detect potentially inappropriate text. After completing these prerequisites, Contoso IT administrators and

compliance specialists can configure the new policy. Compliance specialists can start investigating and remediating any generated alerts.

Enabling auditing in Microsoft 365

Communication Compliance requires audit logs to show alerts and track remediation actions that reviewers take. The audit logs summarize all activities associated with a defined organizational policy or anytime there's a change to a Communication Compliance policy.

Contoso IT administrators review and complete the step-by-step instructions to turn on auditing. After they turn on auditing, a message displays that says the audit log is being prepared and that they can run a search in a couple of hours after the preparation is complete. The Contoso IT administrators only need to perform this action once.

Configuring Viva Engage tenant for Native Mode

Communication Compliance requires that the Viva Engage tenant for an organization is in Native Mode to detect potentially inappropriate text in private messages and public community conversations.

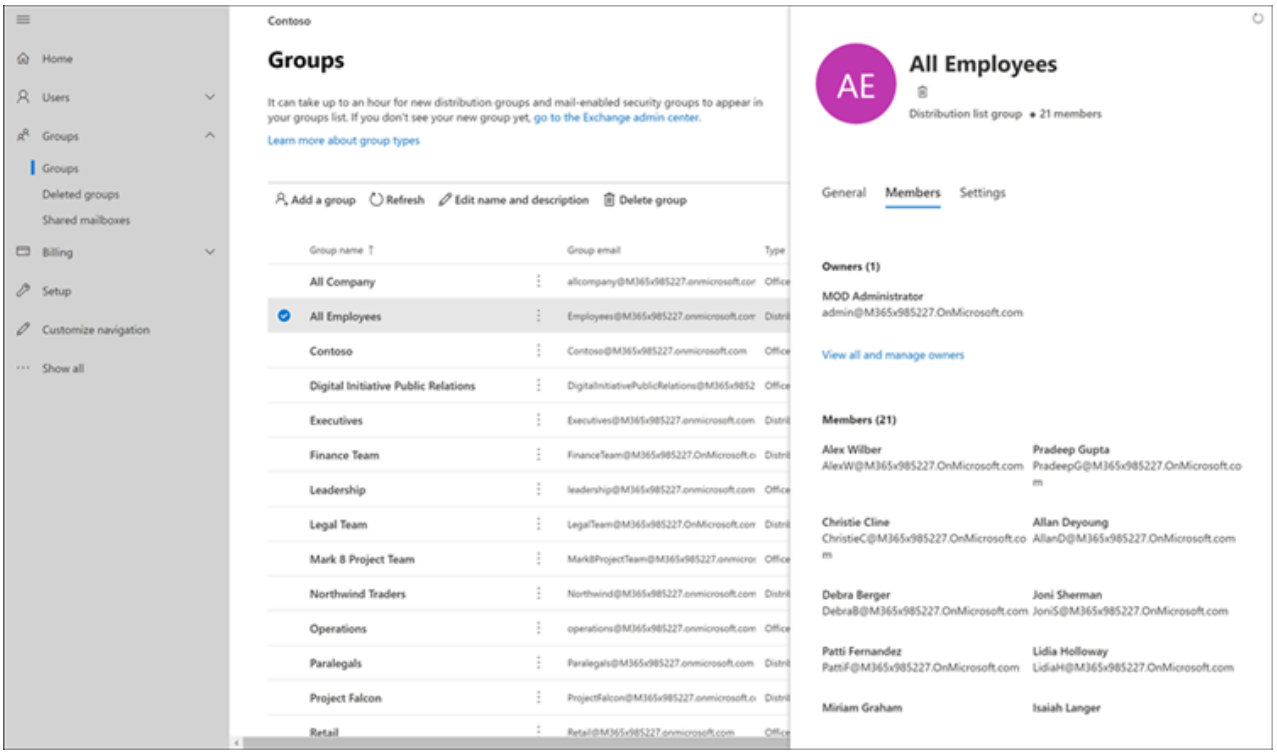
Contoso IT administrators make sure they review the information in the Overview of Viva Engage Native Mode in Microsoft 365 article and follow the steps for running the migration tool in the Configure your Viva Engage network for Native Mode for Microsoft 365 article.

Setting up a group to add users to the policy

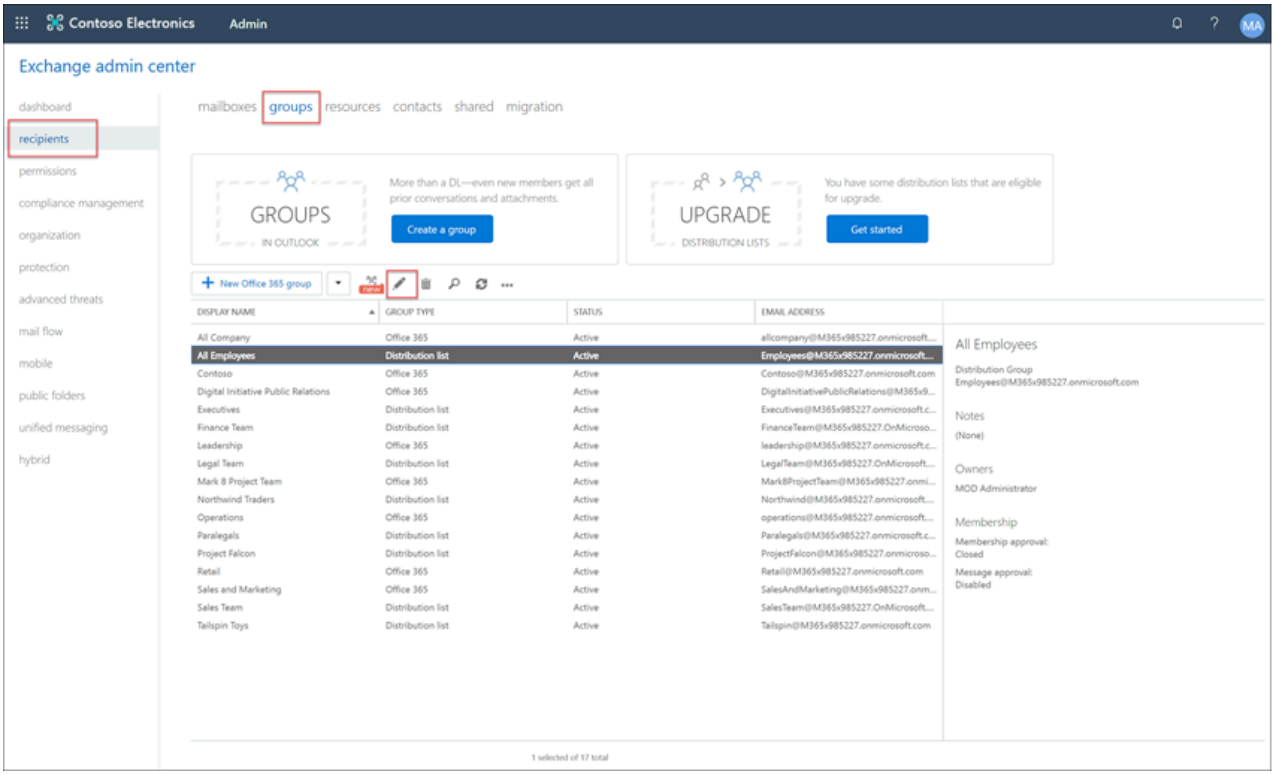
Contoso compliance specialists want to add all users to the communication policy that detects potentially inappropriate text. They could decide to add each user account to the policy separately, but they decide it's easier and saves time to use an **All Users** distribution group for the users for this policy.

They need to create a new group to include all Contoso users, so they take the following steps:

1. Contoso IT administrators sign in to the Microsoft 365 admin center and go to Microsoft 365 admin center > **Groups** > **Groups**.
2. They select **Add a group** and complete the workflow to create a new *Microsoft 365 group* or *Distribution group*.



3. After they create the new group, they add all Contoso users to the new group. They open the Exchange admin center and navigate to **Exchange admin center** > **recipients** > **Groups**. The Contoso IT administrators select the Membership area and the new *All Employees* group they created and select the **Edit** control to add all Contoso users to the new group in the workflow.



Creating the policy to detect potentially inappropriate text

After completing all the prerequisites, the IT administrators and compliance specialists for Contoso are ready to configure the Communication Compliance policy to detect potentially inappropriate text. By using the text policy template, you can quickly and easily configure this new policy.

1. Contoso IT admins sign in to the Microsoft Purview portal with credentials for an admin account in their Microsoft 365 organization.
2. They go to the **Communication Compliance** solution. This step opens a dashboard with quick links to Communication Compliance policy templates.
3. They choose **Policies** in the left navigation, select **Create policy**, and then select **Detect inappropriate text** to open the *Detect communications for inappropriate text* template.
4. In the policy template workflow, the Contoso IT administrators and compliance specialists work together to complete the three required fields: **Policy name**, **Users or groups in scope**, and **Reviewers**.
5. Since the policy workflow already suggests a name for the policy, the IT administrators and compliance specialists decide to keep the suggested name and focus on the remaining fields. They select the *All users* group for the **Users or groups in scope** field and select the compliance specialists who should investigate and remediate policy alerts for the **Reviewers** field. To finish configuring the policy and start gathering alert information, they select **Create policy**.

Step 4: Investigate and remediate alerts

After configuring the Communication Compliance policy to detect potentially inappropriate text, the Contoso compliance specialists investigate and remediate any alerts the policy generates. It can take up to an hour for the policy to fully process communications in all the communication source channels and for alerts to show up in the **Alert dashboard**.

When alerts are generated, Contoso compliance specialists follow the workflow instructions to investigate and remediate potentially inappropriate text issues.

Last updated on 12/17/2025

Microsoft Purview Insider Risk Management and Communication Compliance privacy guide

08/20/2025

Microsoft Purview insider risk solutions give organizations the tools to help detect and reduce potential risks and policy violations. The Microsoft Purview insider risk solutions include:

- **Microsoft Purview Insider Risk Management** uses different signals to find potential malicious or accidental insider risks, such as intellectual property theft, data leakage, and security violations. Insider risk management lets customers create policies to manage security and compliance.
- **Microsoft Purview Communication Compliance** provides tools to help organizations detect potential regulatory compliance issues (for example SEC or FINRA) and business conduct violations such as sensitive or confidential information, harassing or threatening language, and sharing of adult content.

Insider risk management and communication compliance are built with privacy by design and balance user privacy with tools that help detect and reduce organizational risks. We're committed to protecting user trust and maintaining user-level privacy through our core privacy principles:

- Pseudonymization
- Role-based access controls
- Admin explicit opt-in
- Audit logs

Pseudonymization

Pseudonymization helps protect end-user privacy by removing identifiable user details like user name or email address. Pseudonymization also helps prevent potential bias and conflicts of interest by removing identifiable user details (name, email) and personal data (title, department, or location) exposed in the solution. For example, a user named John Smith is pseudonymized into a nonpersonal identifier such as ANON2340. Pseudonyms are on by default for specific roles such as *Insider Risk Management Analysts* and *Insider Risk Management Investigators* (review alerts and take action respectively) and *Communication Compliance Analysts* (review policy alerts).

Role-based access controls

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

We recommend using stringent role-based access controls so only authorized insider risk management and communication compliance roles can use and access alerts and insights into potential policy violations. By default, global administrators don't have access to insider risk management and communication compliance features. This default setting helps ensure that only the appropriate stakeholders can access the solution and details specific to their role permissions. Organizations can assign users to specific role groups to manage different sets of features based on their responsibilities. For example, insider risk management and communication compliance admins can create, configure, and delete policies but can't access or investigate alerts or cases. On the other hand, insider risk management and communication compliance investigators can access and investigate alerts and cases but can't configure policies.

> **Note**
>
> Insider risk management admins can allow investigators and analysts to make edits to policy indicators and thresholds by using the inline alert customization setting.

Whether your organization chooses a single role group or multiple role groups to fit your organization's compliance and privacy requirements, both insider risk management and communication compliance allow admins to choose from predefined role group options within each solution.

Learn more about role group options for each solution:

- Get started with insider risk management
- Get started with communication compliance

Admin explicit opt-in

Insider risk management and communication compliance policies detect risky activities and communications and potential policy violations that could result in a security incident. An admin with the right permissions can explicitly scope employees into a policy.

Additionally, by default, insider risk management and communication compliance indicators that help detect risky activities and communications that might lead to potential data security incidents are disabled. For example, indicators like "downloading content from OneDrive", "sharing SharePoint files with people outside the organization", or "sending sensitive information or harassing messages" are off by default. Insider risk management and communication compliance don't detect those activities without an admin's explicit opt-in. Admins with the right permissions must explicitly select and opt in to one or more indicators in settings before a policy can detect those activities.

Admin explicit opt-in controls help safeguard end-user privacy by ensuring the solutions only flag alerts and policy violations for users and indicators specified in the policies.

Audit logs

Microsoft Purview insider risk solutions audit logs record all admin actions. With these logs, organizations can stay informed about all actions taken within Microsoft Purview insider risk solutions, including when an admin creates or edits a policy, adds a user, views user activity insights, or adds indicators.

Audit logs are enabled by default for all Microsoft 365 organizations. With these logs, organizations can audit privileged admins' actions and meet compliance and privacy requirements.

For more information about audit logs capabilities for each solution, see:

- Review activities with the insider risk management audit log
- Use communication compliance reports and audits

Protect user trust and build a holistic insider risk program

User privacy and trust are essential for organizations to establish a holistic insider risk program. The right set of tools can help you address risks in a way that meets security needs. Learn how to build a holistic insider risk management program with five elements that help companies have stronger data protection while ensuring user trust.

Create and manage Communication Compliance policies

> **Important**
>
> Microsoft Purview Communication Compliance provides the tools to help organizations detect regulatory compliance (for example, SEC or FINRA) and business conduct violations such as sensitive or confidential information, harassing or threatening language, and sharing of adult content. Communication Compliance is built with privacy by design. Usernames are pseudonymized by default, role-based access controls are built in, investigators are opted in by an admin, and audit logs are in place to help ensure user-level privacy.

Create and manage policies

Create Communication Compliance policies for Microsoft 365 organizations in the Microsoft Purview portal. Communication Compliance policies define which communications and users are subject to review in your organization, set custom conditions the communications must meet, and specify who should do reviews.

- Users assigned the *Communication Compliance Admins* role can set up policies and access the **Communication Compliance** page and global settings in Microsoft Purview.
- You can export the modification history to a .csv file that includes the status of alerts pending review, escalated items, and resolved items.
- You can't rename policies, but you can delete them when no longer needed.

> **Important**
>
> PowerShell isn't supported for creating and managing Communication Compliance policies. To create and manage these policies, use the policy management controls in the Communication Compliance solution.

Choose a policy template

Policy templates are predefined policy settings that you can use to quickly create policies to address common compliance scenarios. Each of these templates has differences in conditions

and scope. All templates use the same types of detection signals. You can choose from the following policy templates:

Expand table

Area	Policy Template	Details
Conflict of interest	Detect conflict of interest	- Locations: Exchange Online, Microsoft Teams, Viva Engage - Direction: Internal - Review Percentage: 100% - Conditions: None
Copilot interactions	Detect Microsoft 365 Copilot and Microsoft 365 Copilot Chat interactions	- Location: Microsoft 365 Copilot and Microsoft 365 Copilot Chat - Direction: Inbound, Outbound, Internal - Review Percentage: 100% - Conditions: Prompt Shields, Protected material classifiers
Inappropriate content	Detect inappropriate content	- Location: Microsoft Teams, Viva Engage - Direction: Inbound, Outbound, Internal - Review Percentage: 100% - Conditions: Hate, Violence, Sexual, Self-harm classifiers
Inappropriate images	Detect inappropriate images	- Locations: Exchange Online, Microsoft Teams - Direction: Inbound, Outbound, Internal - Review Percentage: 100% - Conditions: Adult and Racy image classifiers
Inappropriate text	Detect inappropriate text	- Locations: Exchange Online, Microsoft Teams, Viva Engage - Direction: Inbound, Outbound, Internal - Review Percentage: 100% - Conditions: Threat, Discrimination, and Targeted harassment classifiers
Regulatory compliance	Detect financial regulatory compliance	- Locations: Exchange Online, Microsoft Teams, Viva Engage - Direction: Inbound, Outbound - Review Percentage: 10% - Conditions: Customer complaints, Gifts & entertainment, Money laundering, Regulatory collusion, Stock manipulation, and Unauthorized disclosure classifiers
Sensitive information	Detect sensitive info types	- Locations: Exchange Online, Microsoft Teams, Viva Engage - Direction: Inbound, Outbound, Internal

Area	Policy Template	Details
		- Review Percentage: 10% - Conditions: Sensitive information, out-of-the-box content patterns, and types, custom dictionary option, attachments larger than 1 MB

User-reported messages policy

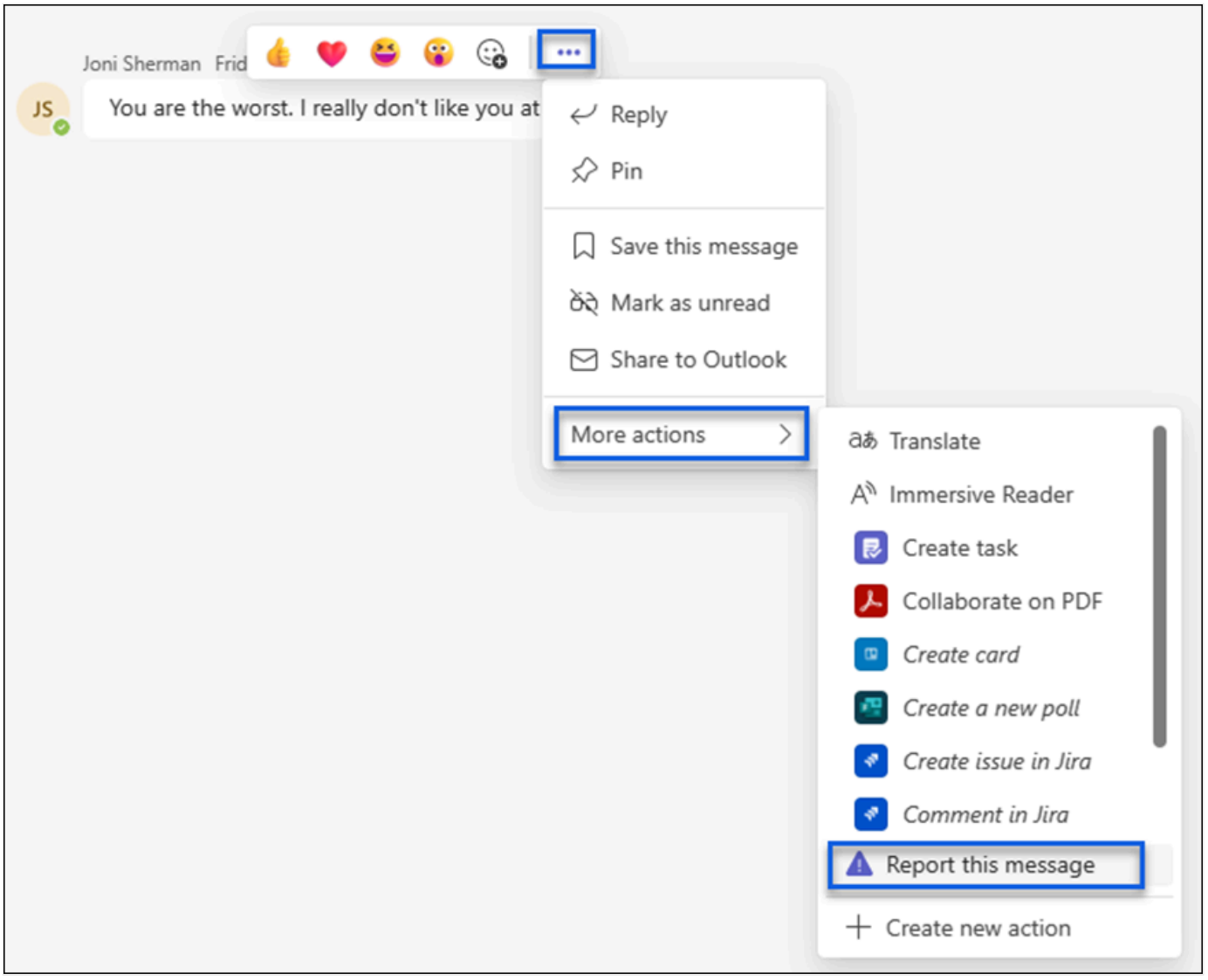
The *User-reported messages* policy is implemented for your organization when you purchase a license that includes Microsoft Purview Communication Compliance. However, this feature can take up to 30 days to become available after license purchase.

As part of a layered defense to detect and remediate inappropriate messages in your organization, you can supplement Communication Compliance policies with user-reported messages in Microsoft Teams and Viva Engage (preview). To support a safe and compliant work environment, this feature empowers users in your organization to self-report inappropriate internal Teams chat messages and Viva Engage conversations, such as harassing or threatening language, sharing of adult content, and sharing of sensitive or confidential information.

User-reported messages in Microsoft Teams

Enabled by default in the Teams admin center, the *Report inappropriate content* option in Teams messages lets users in your organization submit inappropriate internal personal and group chat messages for review by Communication Compliance reviewers for the User-reported messages policy. A default system policy supports these messages and supports reporting messages in Teams group and private chats.

Azure AI Content Safety's hate, self-harm, sexual, and violence classifiers evaluate user-reported messages. This evaluation helps investigators quickly assess potential risk in user-reported content and take action.

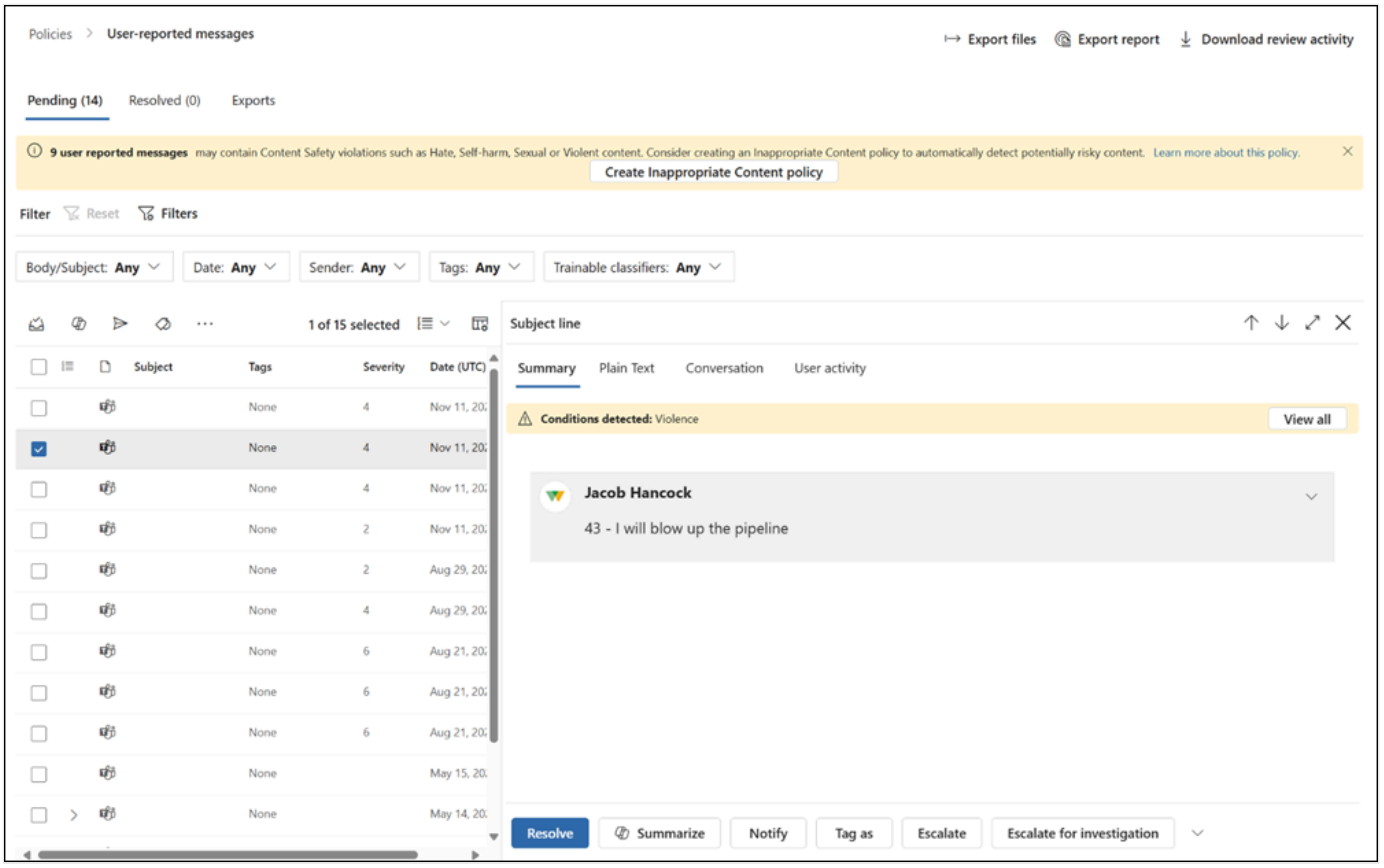


When a user submits a Teams chat message for review, the message is copied to the User-reported message policy. Reported messages initially remain visible to all chat members and there's no notification to chat members or the submitter that a message has been reported in channel, private, or group chats. A user can't report the same message more than once and the message remains visible to all users included in the chat session during the policy review process.

During the review process, Communication Compliance reviewers can perform all the standard remediation actions on the message, including removing the message from the Teams chat. Depending on how the messages are remediated, the message sender and recipients see different notification messages in Teams chats after the review. Any user-reported content that triggers the Content Safety classifiers shows the classifier name that flagged the message and a corresponding severity value is visible in the **Severity** column.

> **Important**

If a user reports a message that was sent before they were added to a chat, Teams message remediation isn't supported; the Teams message can't be removed from the chat.



User-reported messages from Teams chats are the only messages processed by the User-reported message policy and you can only modify the assigned reviewers for the policy. You can't edit all other policy properties. When you create the policy, all members of the *Communication Compliance Admins* role group (if populated with at least one user) or all members of your organization's *Global Admin* role group are assigned as initial reviewers. A randomly selected user from the *Communication Compliance Admins* role group (if populated with at least one user) or a randomly selected user from your organization's *Global Admin* role group is the policy creator.

> **ⓘ Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

Admins should immediately assign custom reviewers to this policy as appropriate for your organization. This assignment might include reviewers such as your Compliance Officer, Risk Officer, or members of your Human Resources department.

Customize the reviewers for chat messages submitted as user-reported messages

To customize the reviewers assigned to the User-reported messages policy, complete the following steps:

1. Sign in to the Microsoft Purview portal using credentials for an admin account in your Microsoft 365 organization.
2. Go to the **Communication Compliance** solution.
3. Select **Policies** in the left navigation.
4. Select the *User-reported messages* policy, then select **Edit**.
5. On the **Detect user-reported messages** pane, assign reviewers for the policy. Reviewers must have mailboxes hosted on Exchange Online. When you add reviewers to a policy, they automatically receive an email message that notifies them of the assignment to the policy and provides links to information about the review process.
6. Select **Save**.

The *Report inappropriate content* option is enabled by default and you can control it via Teams messaging policies in the Teams Admin Center. Users in your organization automatically receive the global policy, unless you create and assign a custom policy. Edit the settings in the global policy or create and assign one or more custom policies to turn on or turn off the *Report inappropriate content* option. For more information, see Manage messaging policies in Teams.

> **Important**
>
> If you use PowerShell to turn on or turn off the **End user reporting** option in the Teams Admin Center, use **Microsoft Teams cmdlets module version 4.2.0** or later.

User-reported messages in Viva Engage

The **Report Conversations** option is off by default in the Viva Engage admin center. When you turn on this option, you see different options depending on whether you have a license that includes Communication Compliance:

- **Licenses that don't include Communication Compliance**. If you don't have a license that includes Communication Compliance, when you turn on the option, the Viva Engage admin can specify an email address to receive reported conversations. The admin can also enter pre-submission instructions and post-submission confirmations for the user. Learn more

about enabling the **Report Conversations** option if you don't have a license that includes Communication Compliance

- **Licenses that do include Communication Compliance**. If you have a license that includes Communication Compliance, when you turn on the option, any reported conversations automatically route through Communication Compliance for investigation.

How upgrades and downgrades in licensing affect the reported conversations workflow

The following behaviors apply when your organization upgrades or downgrades between licenses that do and don't include Communication Compliance:

- If you upgrade from a license that doesn't include Communication Compliance to a license that includes Communication Compliance and turn on the **Report Conversations** option, reported conversations automatically route to Communication Compliance. The Viva Engage admin can continue to access earlier reported conversations and can view new reported conversations if they're added as an investigator in Communication Compliance.

- If you downgrade from a license that includes Communication Compliance to a license that doesn't include Communication Compliance and turn on the **Report Conversations** option, the workflow reverts to the process described in this article for customers that don't have a Communication Compliance license. To continue to use the feature, an admin must turn on the **Report Conversations** option in the Viva Engage admin center again and specify an email address to send reported conversations to.

User experience for customers that have a license that includes Communication Compliance

If you choose to route user-reported conversations through Communication Compliance, you can create pre-submission instructions for the user.

To report a conversation in Viva Engage, the user selects the ellipsis (three dots), then selects **Report conversation**.

The command name changes to **Report comment**, **Report question**, or **Report answer** depending on whether the user is viewing a comment, question, or answer in Viva Engage or Answers in Viva.

If the user submits the report successfully, they see a confirmation message.

When a conversation is reported, the conversation is copied to the Communication Compliance User-reported message policy. The Communication Compliance investigator can review reported conversations in the User Reported policy inbox and can do all the standard remediation actions for the conversation.

Turn on the Report Conversation feature in the Viva Engage admin center

To turn on the Report Conversation feature in the Viva Engage admin center, complete the following steps:

1. Go to the Viva Engage admin center.
2. On the **Settings** page, select **Report Conversations**.
3. Turn on the setting.
4. Under **Report recipient**, add an email address to route reports to. This email address is used to notify the admin in case there's any problem with the configuration.
5. In the **Pre-submission details or instructions for users** box, enter any instructions you want to display for the user.
6. In the **Post-submission instructions to user** box, enter any instructions (optional) that the user will see after reporting a conversation.

Integrate Communication Compliance with Microsoft Purview Insider Risk Management

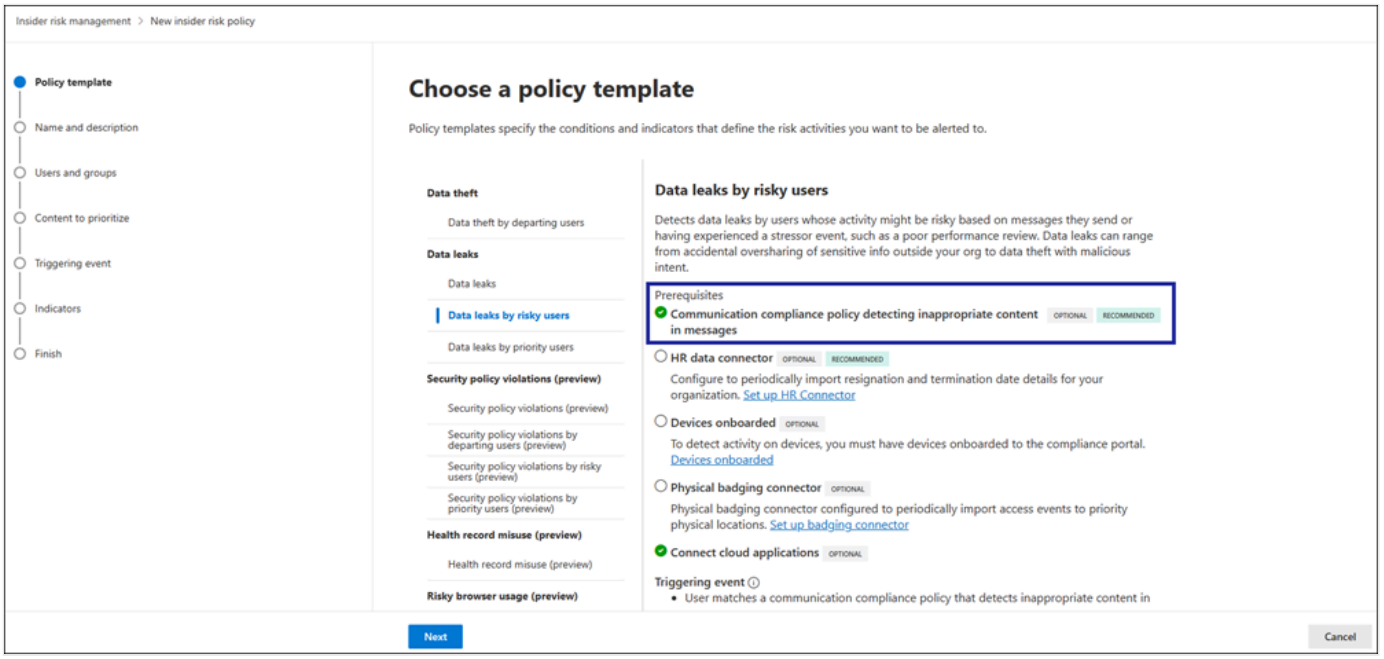
When users experience employment stressors, they might engage in risky activities. Workplace stress can lead to uncharacteristic or malicious behavior by some users that surfaces as potentially inappropriate behavior on your organization's messaging systems. Counterproductive work behavior can be a precursor to more serious violations, such as sabotaging company assets or leaking sensitive information. By integrating Communication Compliance with Microsoft Purview Insider Risk Management, you can detect stressors that indicate an unhealthy workplace environment.

You can integrate Communication Compliance with Insider Risk Management by automatically creating a Communication Compliance policy from Insider Risk Management. You can create this type of policy in two different ways:

- By creating an Insider Risk Management trigger for policy templates
- By selecting Insider Risk Management policy indicators for data-based policy templates
- By selecting generative AI policy indicators for policy templates

Create an Insider Risk Management trigger for policy templates

Communication Compliance can provide risk signals detected in applicable messages to Insider Risk Management risky user policies by using a dedicated Detect inappropriate text policy. This policy is automatically created if you select it as an option when creating a policy by using the Data leaks by risky users template or the Security policy violations by risky users template in Insider Risk Management. For example, the following screenshot shows the option selected in the *Data leaks by risky users* template.



When configured for an Insider Risk Management policy, Communication Compliance creates a dedicated policy named *Insider risk trigger - (date created)* and automatically includes all organization users in the policy. This policy starts detecting risky behavior in messages by using the Microsoft provided Threat, Harassment, and Discrimination trainable classifiers and automatically sends these signals to Insider Risk Management. If needed, you can edit this policy to update the scope of included users and the policy conditions and classifiers.

Users who send five or more messages classified as potentially risky within 24 hours are automatically brought in-scope for Insider Risk Management policies that include this option. Once in-scope, the Insider Risk Management policy detects potentially risky activities configured in the policy. The policy generates alerts as applicable. It can take up to 48 hours from the time risky messages are sent until the time a user is brought in-scope in an Insider Risk Management policy. If an alert is generated for a potentially risky activity detected by the Insider Risk Management policy, the triggering event for the alert is identified as being sourced from the Communication Compliance risky activity.

All users assigned to the Insider Risk Management Investigators role group are automatically assigned as reviewers in the dedicated Communication Compliance policy. Insider Risk Management investigators might need to review the associated risky user alert directly on the Communication Compliance alerts page (linked from the Insider Risk Management alert details). To enable this access, manually add them to the *Communication Compliance Investigators* role group.

Before integrating Communication Compliance with Insider Risk Management, consider the following guidance when detecting messages containing potentially inappropriate text:

- **For organizations without an existing *Detect inappropriate text* policy**. The Insider Risk Management policy workflow automatically creates the new *Risky user in messages - (date created)* policy. In most cases, you don't need to take any further action.
- **For organizations with an existing *Detect inappropriate text* policy**. The Insider Risk Management policy workflow automatically creates the new *Risky user in messages - (date created)* policy. Although you have two Communication Compliance policies for potentially inappropriate text in messages, investigators don't see duplicate alerts for the same activity. Insider Risk Management investigators only see alerts for the dedicated integration policy and Communication Compliance investigators only see the alerts for the existing policy. If needed, you can edit the dedicated policy to change the included users or individual policy conditions as applicable.

Select Insider Risk Management policy indicators for data-based policy templates

The **Policy indicators** setting in Insider Risk Management provides Communication Compliance indicators to choose from when using the *Data theft*, *Data leaks*, *Data leaks by risky users*, and *Data leaks by priority users* templates:

- Sending financial regulatory text that might be risky
- Sending inappropriate images
- Sending inappropriate content
- Sending messages that contain specific sensitive info types

For more information on creating a Communication Compliance policy this way, see Insider Risk Management policy indicators settings.

Select generative AI policy indicators for policy templates

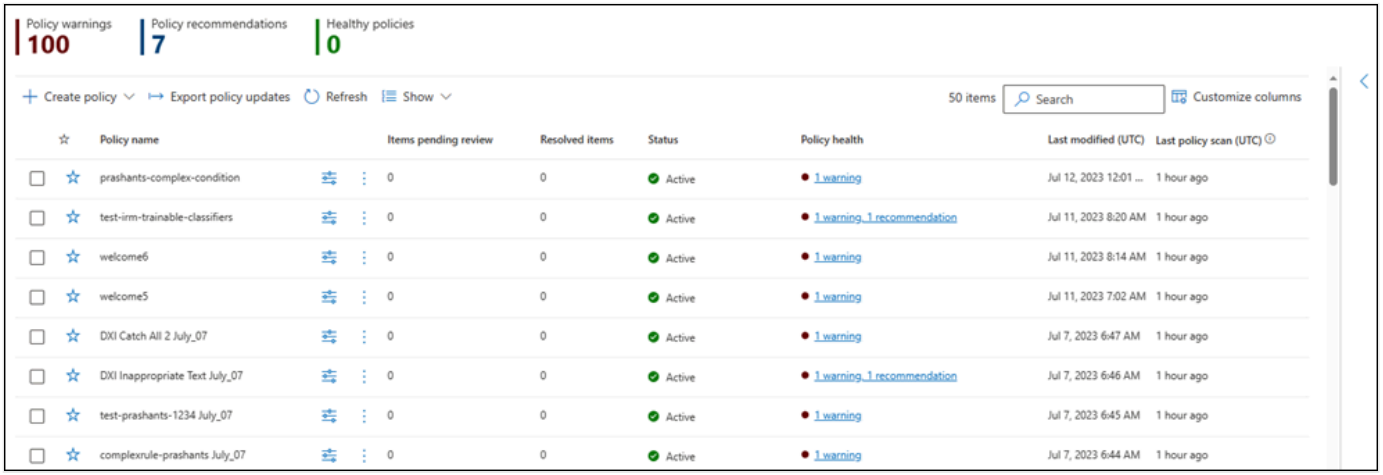
The **Policy indicators** setting in Insider Risk Management provides generative AI app indicators to choose from when using the *Data leaks*, *Data leaks by risky users*, *Data leaks by priority users*, or *Risky AI usage* policy templates. These Azure AI Content Safety indicators help detect when prompts and responses in AI apps match the following classifiers:

- Prompt Shields
- Protected material detection

For more information on creating a Communication Compliance policy this way, see Insider Risk Management policy indicators settings. For more information on Azure AI Content Safety, see Azure AI Content Safety.

Policy health (preview)

The policy health status gives you insights into potential issues or optimizations for your Communication Compliance policies. At the top of the **Policies** page, you see a summary that lists the total number of policy warnings and recommendations, and the total number of healthy policies. If a specific policy has a warning or recommendations, the **Policy health** column lists a link to the warning or recommendations. When you select the link, a details pane opens on the right side of the screen with the **Policy Health** tab selected, which makes it easy to quickly review the warning or recommendation and take action on it.



The **Policy Health** tab has two parts. The upper part of the tab shows general information about the policy, including whether the policy is active, whether it's updated, and general tips. The lower part of the tab shows the specific warning or recommendations. If you're a member of the *Communication Compliance* or *Communication Compliance Admins* group, you can take action directly from the warning or recommendation. If you're a member of the *Communication Compliance Analysts* or *Communication Compliance Investigators* group, you can see the warning or recommendation so you can prompt your admin to take action.

- **Warnings**: If you don't take action on a warning, the policy stops working. For the policy health preview, there's one warning, related to a policy's storage limit size.
- **Recommendations**: Recommendations provide suggestions for optimizing policies. If you ignore a policy recommendation, the policy continues to work. If you're a member of the *Communication Compliance* or *Communication Compliance Admins* group, you can act on a recommendation as long as the policy is active and if it's not updating. If you're a member of the *Communication Compliance Analysts* or *Communication Compliance Investigators* group, you can see the recommendation and prompt your admin to take action. For the preview, there are two recommendations:
  - **Potentially reduce false positives by filtering out bulk emails**: This recommendation prompts you to turn on the Filter email blasts setting in Communication Compliance to exclude bulk emails from policy evaluation. This setting helps reduce false positives by excluding bulk emails, such as newsletters, and spam, phishing, and malware, from getting flagged by Communication Compliance policies if the policy conditions are matched.
  - **Reduce user scoping blind spots**: This recommendation prompts you to reduce blind spots by turning on the **Show insights and recommendations for users who match this policy's conditions but weren't included in the policy** checkbox (used with the **Selected users** policy option).
- **Healthy policies**: If there are no warnings or recommendations for a specific policy, the policy is considered healthy.

Reduce user scoping blind spots

You can reduce user-scoping related blind spots by turning on the **Show insights and recommendations for users who match this policy's conditions but weren't included in the policy** checkbox. When you turn on this checkbox, you see insights and recommendations for users who are sending messages that match the policy condition but aren't included in the scope of the policy. In this case, Communication Compliance recommends that you include those users in the scope of the policy so that you can review the messages sent by those users and take actions. You can select the specific users that you want to add or you can extend the policy to all users.

Insights for these users are aggregated; you can't see the messages sent by them until you add them to the scope of the policy.

If an admin doesn't act on a recommendation, the recommendation continues to recur. To turn off recommendations for users outside the scope of the policy, turn off the checkbox.

Pause a policy

After you create a Communication Compliance policy, you can temporarily pause the policy. Use pausing for testing or troubleshooting policy matches, or for optimizing policy conditions. Instead of deleting a policy in these circumstances, pausing a policy also preserves existing policy alerts and messages for ongoing investigations and reviews. Pausing a policy prevents inspection and alert generation for all user message conditions defined in the policy for the time the policy is paused. To pause or restart a policy, you must be a member of the *Communication Compliance Admins* role group.

To pause a policy, go to the **Policy** page, select a policy, then select **Pause policy** from the actions toolbar. On the **Pause policy** pane, confirm you'd like to pause the policy by selecting **Pause**. In some cases, it can take up to 24 hours for a policy to be paused. Once the policy is paused, alerts aren't generated for messages matching the policy. However, messages associated with alerts that you created before pausing the policy remain available for investigation, review, and remediation.

The policy status for paused policies can show several states:

- **Active**: The policy is active.
- **Paused**: The policy is fully paused.
- **Pausing**: The policy is in the process of being paused.
- **Resuming**: The policy in the process of being resumed.
- **Error in resuming**: An error was encountered when resuming the policy. For the error stack trace, hover your mouse over the *Error in resuming* status in the Status column on the Policy page.
- **Error in pausing**: An error was encountered when pausing the policy. For the error stack trace, hover your mouse over the *Error in pausing* status in the Status column on the Policy page.

To resume a policy, go to the **Policy** page, select a policy, then select **Resume policy** from the actions toolbar. On the **Resume policy** pane, confirm you'd like to resume the policy by selecting **Resume**. In some cases, it can take up to 24 hours for a policy to be resumed. Once the policy is resumed, alerts are generated for messages matching the policy and made available for investigation, review, and remediation.

Copy a policy

If your organization has existing Communication Compliance policies, you might find it helpful to create a new policy from an existing policy. When you copy a policy, you create an exact duplicate of an existing policy, including all included users, all assigned reviewers, and all policy conditions. Some scenarios where copying a policy is helpful include:

- **Policy storage limit reached**: Active Communication Compliance policies have message storage limits. When the storage limit for a policy is reached, the policy is automatically deactivated. If your organization needs to continue detecting, capturing, and acting on inappropriate messages covered by the deactivated policy, you can quickly create a new policy with an identical configuration.
- **Detect and review inappropriate messages for different groups of users**: Some organizations prefer to create multiple policies with the same configuration but include different groups of users and different reviewers for each policy.
- **Similar policies with small changes**: For policies with complex configurations or conditions, it might save time to create a new policy from a similar policy.

To copy a policy, you must be a member of the *Communication Compliance* or *Communication Compliance Admins* role groups. After you create a new policy from an existing policy, it can take up to 24 hours to view messages that match the new policy configuration.

To copy a policy and create a new policy, complete the following steps:

1. Select the policy you want to copy.
2. Select **Copy policy** on the command bar or select **Copy policy** from the action menu for the policy.
3. In the **Copy policy** pane, accept the default name for the policy in the **Policy name** field or rename the policy. The policy name for the new policy can't be the same as an existing active or deactivated policy. Complete the **Description** field as needed.
4. If you don't need further customization of the policy, select **Copy policy** to complete the process. If you need to update the configuration of the new policy, select **Customize policy**. This selection starts the policy workflow to help you update and customize the new policy.

Mark a policy as a favorite

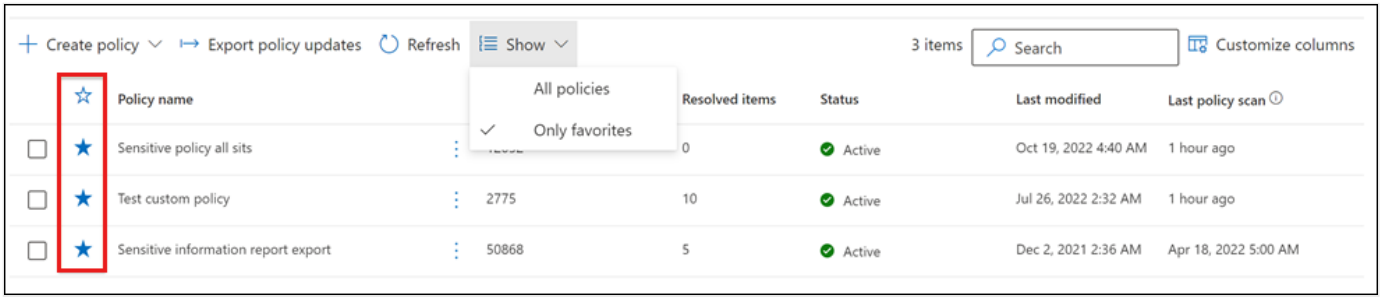
After you create a Communication Compliance policy, you can mark the policy as a favorite. When you mark a policy as a favorite, you can filter favorite policies to appear at the top of the policies list. By marking a policy as a favorite, you can also easily sort policies by favorites.

To mark a policy as a favorite, you have the following options:

- **Mark as favorite**: Enables you to mark selected policies as favorites, so you can easily find the policies that you're most interested in rather than having to search for them.
- **Sort favorites**: Sorts the policies by favorites, so your favorite policies appear at the top of the list.
- **Customize columns**: Choose to list the favorites that you want to see. You can also choose to sort favorite policies in ascending or descending order.

To sort policies by groups:

- **All policies**: This is the default view, displaying all the policies in the list.
- **Only favorites**: Groups policies by favorites at the top of the list.



Policy activity detection

Communications are scanned every hour from the time you create a policy. For example, if you create an inappropriate content policy at 11:00 AM, the policy gathers Communication Compliance signals every hour starting from when you created the policy. Editing a policy doesn't change this time. To view the last scan date and Coordinated Universal Time (UTC) for a policy, go to the *Last policy scan* column on the **Policy** page. After creating a new policy, you might need to wait up to an hour to view the first policy scan date and time.

The following table outlines the time to detection for supported content types:

Expand table

Content type	Time to detection
Email attachment	24 hours
Email body content	1 hour
Email metadata	1 hour
Email OCR	24 hours

Content type	Time to detection
Microsoft 365 Copilot and Microsoft 365 Copilot Chat body content (prompts and responses)	1 hour
Team attachment	24 hours
Teams body content	1 hour
Teams metadata	1 hour
Teams modern attachment	24 hours
Teams OCR	24 hours
Teams shared channels	24 hours
Teams transcripts	1 hour
Viva Engage attachment	Up to 24 hours
Viva Engage body content	1 hour

For existing policies created before July 31, 2022, it might take up to 24 hours to detect messages and review alerts that match these policies. To reduce the latency for these policies, copy the existing policy and create a new policy from the copy. If you don't need to retain any data from the older policy, you can pause or delete it.

To identify an older policy, review *Last policy scan* column on the **Policy** page. Older policies display a full date for the scan while policies created after July 31, 2022 display *1 hour ago* for the scan. To reduce latency, you can also wait until February 28, 2023 for your existing policies to be automatically migrated to the new detection criteria.

Storage limit notification

Each Communication Compliance policy has a storage limit size of 100 GB or 1 million messages, whichever limit you reach first. As the policy approaches these limits, Communication Compliance automatically sends notification emails to users assigned to the *Communication Compliance* or *Communication Compliance Admins* role groups. Notification messages are sent when the storage size or message count reaches 80, 90, and 95 percent of the limit. When the policy limit is reached, the policy is automatically deactivated and stops processing messages for alerts.

> **Important**

If a policy is deactivated because it reached the storage and message limits, evaluate how to manage the deactivated policy. If you delete the policy, you permanently delete all messages, associated attachments, and message alerts. If you need to maintain these items for future use, don't delete the deactivated policy.

To manage policies that approach the storage and message limits, consider making a copy of the policy to maintain coverage continuity or take the following actions to help minimize current policy storage size and message counts:

- Reduce the number of users assigned to the policy. Removing users from the policy or creating different policies for different groups of users can help slow the growth of policy size and total messages.
- Examine the policy for excessive false positive alerts. Add exceptions or make changes to the policy conditions to ignore common false positive alerts.
- Make a copy of the policy if a policy reaches the storage or message limits and deactivates it to continue detecting and taking action for the same conditions and users.

Policy settings

The following settings control how Communication Compliance policies are scoped, evaluated, and reviewed.

Choose which users to include

You can select **All users**, define specific users in a Communication Compliance policy, or select an adaptive scope.

- **All users**: When you select **All users**, you apply the policy to all users and all groups that any user is included in as a member.
- **Select users**: When you define specific users, you apply the policy to the defined users and any groups the defined users are included in as a member. If you choose the **Selected users** option, the **Show insights and recommendations for users who match this policy's conditions but weren't included in the policy** checkbox automatically appears. Leave this checkbox selected if you want to receive recommendations when users outside of the selected users match the policy conditions. This setting isn't available if you choose the **All users** option or the **Select adaptive scopes** option.
- **Select adaptive scope**: An adaptive scope uses a query that you specify to define the membership of users or groups. If you decide to create an adaptive policy, you must create

one or more adaptive scopes before you create your policy, and then select them when you choose this option. The scopes that you can select depend on the scope types that you add. For example, if you only added a scope type of **User**, you can select **Groups**. To learn more, see Advantages of using an adaptive scope.

Choose the communication direction to monitor

By default, the **Direction is** condition appears and you can't remove it. Choose communication direction settings in a policy individually or together:

- **Inbound**: Detects communications sent **to** scoped users from external and internal senders, including other scoped users in the policy.
- **Outbound**: Detects communications sent **from** scoped users to external and internal recipients, including other scoped users in the policy.
- **Internal**: Detects communications **between** the scoped users or groups in the policy.

Choose sensitive information types

You can include sensitive information types as part of your Communication Compliance policy. Sensitive information types are either predefined or custom data types that help identify and protect credit card numbers, bank account numbers, passport numbers, and more. As part of Learn about Microsoft Purview Data Loss Prevention, the sensitive information configuration can use patterns, character proximity, confidence levels, and even custom data types to help identify and flag content that might be sensitive. The default sensitive information types are:

- Financial
- Medical and health
- Privacy
- Custom information type

> **Important**
>
> Sensitive info types have two different ways of defining the max unique instance count parameters. To learn more, see **Create custom sensitive information types**.

The Communication Compliance solution supports default sensitive information types and bundled named-entity sensitive information types, which are collections of sensitive information types. For more information about sensitive information details and the patterns included in the

default types, see Sensitive information type entity definitions. For information on supported bundled named-entity sensitive information types, see the following articles:

- All credentials
- All full names
- All medical terms and conditions
- All Physical Addresses

For more information about sensitive information details and the patterns included in the default types, see Sensitive information type entity definitions.

Configure custom keyword dictionaries

Configure custom keyword dictionaries (or lexicons) to easily manage keywords specific to your organization or industry. Keyword dictionaries support up to 100 KB of terms (post-compression) in the dictionary and support any language. The tenant limit is also 100 KB after compression. If needed, you can apply multiple custom keyword dictionaries to a single policy or have a single keyword dictionary per policy. You can source these dictionaries from a file, such as a .CSV or .TXT list, or from an imported keyword list. Use custom dictionaries when you need to support terms or languages specific to your organization and policies.

Use Microsoft-provided trainable classifiers

Communication Compliance policies that use Microsoft provided trainable classifiers inspect and evaluate messages that meet a minimum word count requirement, depending on the language of the content. Custom trainable classifiers aren't supported. For a complete list of supported languages, word count requirements, and file types for Microsoft provided trainable classifiers, see Trainable classifier definitions.

To identify and take action on messages containing inappropriate language content that doesn't meet the word count requirement, create a sensitive information type or custom keyword dictionary for Communication Compliance policies detecting this type of content.

Expand table

Microsoft provided trainable classifier	Description
Corporate sabotage	Detects messages that might mention acts to damage or destroy corporate assets or property. This classifier can help you manage regulatory compliance obligations such as NERC Critical Infrastructure Protection standards or state by state regulations like Chapter 9.05 RCW in Washington state.
Customer complaints	Detects messages that might suggest customer complaints made on your organization's products or services, as required by law for regulated industries. This classifier can help you manage regulatory compliance obligations such as FINRA Rule 4530, FINRA 4513, FINRA 2111, Consumer Financial Protection Bureau, Code of Federal Regulations Title 21: Food and Drugs, and the Federal Trade Commission Act.
Discrimination	Detects potentially explicit discriminatory language and is particularly sensitive to discriminatory language against the African American/Black communities when compared to other communities.
Gifts & entertainment	Detects messages that might suggest exchanging gifts or entertainment in return for service, which violates regulations related to bribery. This classifier can help you manage regulatory compliance obligations such as Foreign Corrupt Practices Act (FCPA), UK Bribery Act, and FINRA Rule 2320.
Harassment	Detects potentially offensive content in multiple languages that targets people regarding race, color, religion, national origin.
Money laundering	Detects signs that might suggest money laundering or engagement in acts to conceal or disguise the origin or destination of proceeds. This classifier can help you manage regulatory compliance obligations such as the Bank Secrecy Act, the USA Patriot Act, FINRA Rule 3310, and the Anti-Money Laundering Act of 2020.
Profanity	Detects potentially profane content in multiple languages that would likely offend most people.
Regulatory collusion	Detects messages that might violate regulatory anti-collusion requirements such as an attempted concealment of sensitive information. This classifier can help you manage regulatory compliance obligations such as the Sherman Antitrust Act, Securities Exchange Act 1933, Securities Exchange Act of 1934, Investment Advisers Act of 1940, Federal Commission Act, and the Robinson-Patman Act.
Stock manipulation	Detects signs of possible stock manipulation, such as recommendations to buy, sell, or hold stocks that might suggest an attempt to manipulate the stock price. This classifier can help you manage regulatory compliance obligations such as the Securities Exchange Act of 1934, FINRA Rule 2372, and FINRA Rule 5270.
Threat	Detects potential threatening content in multiple languages aimed at committing violence or physical harm to a person or property.

Microsoft provided trainable classifier	Description
Unauthorized disclosure	Detects sharing of information containing content that is explicitly designated as confidential or internal to unauthorized individuals. This classifier can help you manage regulatory compliance obligations such as FINRA Rule 2010 and SEC Rule 10b-5.

> **Important**
>
> Microsoft provided trainable classifiers might detect a large volume of bulk sender/newsletter content due to a known issue. You can mitigate the detection of large volumes of bulk sender/newsletter content by selecting the **Filter email blasts check box** when you create the policy. You can also edit an existing policy to turn on this feature.

Content safety classifiers based on large language models

Communication Compliance includes a set of Azure AI classifiers for Microsoft 365 Copilot, Teams, and Viva Engage communications. These classifiers run on large language models and are highly accurate. They evaluate messages containing three or more words. If the evaluated severity is 4 or higher, the solution displays the message as an alert and adds a **Severity** column to the **Alerts** dashboard to make it easier to prioritize alerts. Investigators can sort and filter on the **Severity** column.

The **Severity** column only appears if the policy match comes from one of the classifiers listed in the following table. For all other classifiers, the column is empty.

The following table describes the classifiers:

Expand table

Classifier	Description
Hate	*Hate* refers to any content that attacks or uses pejorative or discriminatory language with reference to a person or identity group based on certain differentiating attributes of these groups, including but not limited to race, ethnicity, nationality, gender identity and expression, sexual orientation, religion, immigration status, ability status, personal appearance, and body size.
Sexual	*Sexual* describes language related to anatomical organs and genitals, romantic relationships, acts portrayed in erotic or affectionate terms, pregnancy, physical sexual acts, including those

Classifier	Description
	portrayed as an assault or a forced sexual violent act against one's will, prostitution, pornography, and abuse.
Violence	*Violence* describes language related to physical actions intended to hurt, injure, damage, or kill someone or something; describes weapons, guns and related entities, such as manufactures, associations, legislation, and so on.
Self-harm	*Self-harm* describes language related to physical actions intended to purposely hurt, injure, damage one's body, or kill oneself.

Learn more about Azure AI Content Safety

Considerations when using classifiers

Keep the following limitations and considerations in mind when using content safety classifiers:

- Only Microsoft 365 Copilot, Teams, and Viva Engage workloads are supported.
- The solution evaluates only messages containing five or more words. It doesn't evaluate attachments and OCR images.
- The solution doesn't evaluate Teams meeting transcripts.
- Feedback loop isn't yet supported to submit misclassified items to Microsoft.
- The solution supports up to 10,000 characters per message.
- Learn about supported languages

Detect risky generative AI interactions

Communication Compliance detects harmful user-generated and AI-generated content in applications and services. This detection includes evaluation of user prompts submitted to generative AI services and the inclusion of known text content that might be sensitive to your organization.

The following table describes the classifiers.

Expand table

Classifier	Description
Prompt Shield	Detects adversarial user input attacks such as user prompt injection attacks (jailbreak). User prompt injection attacks deliberately exploit system vulnerabilities to elicit unauthorized behavior from the large language model.
Protected material	Detects known text content that might be protected under copyright or branding laws. Detecting the display of protected material in generative AI responses ensures compliance

Classifier	Description
	with intellectual property laws and maintains content originality.

For more information, see Learn more about Azure AI Content Safety.

Configure optical character recognition (OCR) scanning

> **Note**
>
> Microsoft Purview includes **OCR (preview) settings** for Microsoft Purview Insider Risk Management, Microsoft Purview Data Loss Prevention, Microsoft Purview Data Loss Management, and auto-labeling. Use the OCR (preview) settings to provide image-scanning capabilities for those solutions and technologies. Communication Compliance has its own built-in OCR scanning functionality as described in this section and doesn't support the OCR (preview) settings at this time.

You can configure built-in or custom Communication Compliance policies to scan and identify printed or handwritten text from images that might be inappropriate in your organization. Integrated Azure AI services and optical scanning support for identifying text in images help analysts and investigators detect and act on instances where inappropriate conduct might be missed in communications that is primarily nontextual.

You can enable optical character recognition (OCR) in new policies from templates, custom policies, or update existing policies to expand support for processing embedded images and attachments. When you enable OCR in a policy created from a policy template, the solution automatically scans embedded or attached images in email and Microsoft Teams chat messages. For images embedded in document files, OCR scanning isn't supported. For custom policies, configure one or more conditional settings associated with keywords, Microsoft provided trainable classifiers, or sensitive info types in the policy to enable the selection of OCR scanning.

The solution scans and processes images from 100 KB to 4 MB in the following image formats:

- .jpg/.jpeg (joint photographic experts group)
- .png (portable network graphics)
- .bmp (bitmap)
- .tiff (tag image file format)

When you review pending policy matches for policies with OCR enabled, you see images identified and matched to policy conditions as child items for associated alerts. You can view the

original image to evaluate the identified text in context with the original message. It might take up to 48 hours for detected images to be available with alerts.

Choose conditions for your policies

The conditions you choose for a policy apply to communications from both email and third-party sources in your organization (Instant Bloomberg, for example).

At this time, the condition names that you choose depend on whether you're creating a new policy or an existing policy:

- For *new policies*, we recommend using the condition builder. The condition builder supports the *OR* operator, which enables you to combine multiple conditions in the same policy to create compound conditions with AND, OR, and NOT operators. This support helps you address complex scenarios for your unique compliance requirements. If you create a new policy with a template, Communication Compliance automatically uses the condition builder. Learn more about the condition builder.

  All exceptions in the condition builder are replaced with a NOT condition. You must include the NOT condition within a nested group.

- For *existing policies*, use the condition names listed in the first column in the following table.

> **Tip**
>
> To save time, you can **test certain conditions before creating your policy**.

Condition names for existing and new policies

The following table lists condition names to use for existing policies vs. new policies (using the condition builder). It also lists how to use each condition.

Expand table

Condition name for existing policies	Condition name for new policies	How to use this condition
Attachment contains any of these words	**Attachment contains**	To apply the policy when certain words or phrases are included or excluded in a message attachment (such as a Word document).

Condition name for existing policies	Condition name for new policies	How to use this condition
Attachment contains none of these words	**words or phrases**	Make sure to use the following syntax when entering conditional text: - Remove all leading and trailing spaces. - Add quotation marks before and after each keyword or key phrase. - Separate each keyword or key phrase with a comma. - Don't include spaces between items separated by a comma. **Example:** "banker","insider trading","confidential 123" Each word or phrase you enter is applied separately (only one word must apply for the policy to apply to the attachment). For more information about entering words or phrases, see the next section Matching words and phrases to emails or attachments.
Attachment is any of these file types **Attachment is none of these file types**	**Attachment file extension is**	To bring communications into scope that include or exclude specific types of attachments, enter the file extensions (such as .exe or .pdf). If you want to include or exclude multiple file extensions, enter file types separated by a comma (example *.exe,.pdf,.zip*). Don't include spaces between items separated by a comma. Only one attachment extension must match for the policy to apply.
Attachment is larger than **Attachment is not larger than**	**Attachment size equals or is larger than**	To review messages based on the size of their attachments, specify the maximum or minimum size an attachment can be before the message and its attachments are subject to review. For example, for an existing policy, if you specify **Attachment is larger than** > **2.0 MB**, all messages with attachments 2.01 MB and over are subject to review. You can choose bytes, kilobytes, megabytes, or gigabytes for this condition.
Content contains any of these sensitive info types	**Content contains sensitive info types**	Apply to the policy when any sensitive information types are included or excluded in a message. Each sensitive information type you choose is applied separately and only one of these sensitive information types must apply for the policy to apply to the message. For more information about custom sensitive information types, see Learn about sensitive information types.
Content matches any of these classifiers	**Content matches trainable classifiers**	Apply to the policy when any trainable classifiers are included or excluded in a message. Some classifiers are predefined in your organization, and custom classifiers must be configured separately before they're available for this condition. For existing policies, only one trainable classifier can be defined as a condition in a policy. If you're using the condition builder for a new policy, you're not limited to a single trainable classifier. For more information about configuring trainable classifiers, see Learn about trainable classifiers.

Condition name for existing policies	Condition name for new policies	How to use this condition
Message contains any of these words **Message contains none of these words**	**Message contains words or phrases**	To apply the policy when certain words or phrases are included or excluded in a message. Make sure to use the following syntax when entering conditional text: - Remove all leading and trailing spaces. - Add quotation marks before and after each keyword or key phrase. - Separate each keyword or key phrase with a comma. - Don't include spaces between items separated by a comma. **Example:** "banker","insider trading","confidential 123" Each word or phrase you enter is applied separately (only one word must apply for the policy to apply to the message). For more information about entering words and phrases, see the next section Matching words and phrases to emails or attachments.
Message is classified with any of these labels **Message is not classified with any of these labels**	**Message has retention labels applied**	To apply the policy when certain retention labels are included or excluded in a message. Retention labels must be configured separately and configured labels are chosen as part of this condition. Each label you choose is applied separately (only one of these labels must apply for the policy to apply to the message). For more information about retention labels, see Learn about retention policies and retention labels.
Message is received from any of these domains **Message is not received from any of these domains**	**Sender domain is**	Apply the policy to include or exclude specific domains in received messages. Make sure to use the following syntax when entering conditional text: -Enter each domain and separate multiple domains with a comma. -Don't include spaces between items separated by a comma. -Remove all leading and trailing spaces. Each domain entered is applied separately, only one domain must apply for the policy to apply to the message. If you want to use **Message is received from any of these domains** to look for messages from specific domains, you need to combine this with another condition like **Message contains any of these words**, or **Content matches any of these classifiers**, or you might get unexpected results. For existing policies, if you want to scan all emails but want to exclude messages from a specific domain that don't need review (newsletters,

Condition name for existing policies	Condition name for new policies	How to use this condition
		announcements, and so on), you must configure a **Message is not received from any of these domains** condition that excludes the domain (example 'contoso.com,wingtiptoys.com').
Message is received from any of these external email addresses **Message is not received from any of these external email addresses**	**Sender is**	Apply the policy to include or exclude messages received or not received from specific external email addresses (example someone@outlook.com). Use this condition to detect only messages that come from outside the organization (messages that cross the firewall). Make sure to use the following syntax when entering email addresses: - Enter each email address and separate multiple email addresses with a comma. - Don't include spaces between email addresses separated by a comma. - Remove all leading and trailing spaces. - Remove any single quote or double quotes
Message is sent to any of these domains **Message is not sent to any of these domains**	**Recipient domain is**	Apply the policy to include or exclude specific domains in sent messages. Make sure to use the following syntax when entering conditional text: -Enter each domain and separate multiple domains with a comma. -Don't include spaces between items separated by a comma. -Remove all leading and trailing spaces. Each domain is applied separately; only one domain must apply for the policy to apply to the message. For existing policies, if you want to exclude all emails sent to two specific domains, configure the **Message is not sent to any of these domains** condition with the two domains (example 'contoso.com,wingtiptoys.com').
Message is sent to any of these external email addresses **Message is not sent to any of these external email addresses**	**Recipient is**	Apply the policy to include or exclude messages sent or not sent to specific external email addresses (example someone@outlook.com). Use this condition to detect only messages that are sent outside the organization (messages that cross the firewall). Make sure to use the following syntax when entering email addresses: - Enter each email address and separate multiple email addresses with a

Condition name for existing policies	Condition name for new policies	How to use this condition
		comma. - Don't include spaces between email addresses separated by a comma. - Remove all leading and trailing spaces. - Remove any single quote or double quotes
Message size is larger than **Message size is not larger than**	**Message size equals or is larger than**	To review messages based on a certain size, use these conditions to specify the maximum or minimum size a message can be before it's subject to review. For example, for an existing policy, if you specify **Message size is larger than** > **1.0 MB**, all messages that are 1.01 MB and larger are subject to review. You can choose bytes, kilobytes, megabytes, or gigabytes for this condition.

> **Important**
>
> If a condition includes a list, don't include spaces between list items. For example, enter "bias,harassment" instead of "bias, harassment".

Matching words and phrases to emails or attachments

Each word you enter and separate with a comma is applied separately (only one word must apply for the policy condition to apply to the email or attachment). For example, let's use the condition **Message contains any of these words** with the keywords "banker", "confidential", and "insider trading" separated by a comma (banker,confidential,"insider trading"). The policy applies to any messages that includes the word "banker", "confidential", or the phrase "insider trading". Only one of these words or phrases must occur for this policy condition to apply. Words in the message or attachment must exactly match what you enter.

> **Important**
>
> When importing a custom dictionary file, each word or phrase must be separated with a carriage return and on a separate line. For example:
>
> *banker*
> *confidential*
> *insider trading*

To scan both email messages and attachments for the same keywords, create a custom keyword dictionary for the terms you want to scan. This policy configuration identifies defined keywords that appear in either the email message **OR** in the email attachment. Using the standard conditional policy settings (**Message contains any of these words** and **Attachment contains any of these words**) to identify terms in messages and in attachments requires the terms to be present in **BOTH** the message and the attachment.

Use the condition builder to create complex conditions (new policies)

If you want to create nested conditions or use the **OR** operator in addition to the **AND** operator in new policies, use the condition builder. To use multiple operators (both **AND** and **OR**) in the same condition, you must create a separate group for each operator. Select **Add group** to create a group.

To use the condition builder, you must first opt in. To do this, in the **Conditions** section, above the existing condition builder, turn the **Condition builder** option on. If you create a new policy by using a template, you don't have to opt in to the condition builder. Communication Compliance automatically uses the condition builder for new policies based on templates.

All exceptions in the condition builder are replaced with a NOT condition. For example, the existing condition builder includes the **Message is not received from any of these domains** condition. This condition is renamed to **Sender domain is** in the condition builder, so to build the **Message is not received from any of these domains** condition in the condition builder, use the **Sender domain is** condition together with the **NOT** condition. You must include the **NOT** condition within a nested group.

You can use conditions with trainable classifiers and sensitive info types as an exception by using the **NOT** condition in a nested group, but keyword highlighting isn't available in this case.

To make it easier to see a summary of a complex condition that includes multiple **AND** and **OR** operators or groups, the condition builder includes a simplified summary view of the condition. To see the summary for a condition you built, turn on the **Quick Summary** option.

For scenarios that show how to use conditions in policies, see Scenarios for using conditions in Communication Compliance policies.

Set the review percentage for policy matches

To reduce the amount of content to review, specify a percentage of all communications governed by a Communication Compliance policy. Communication Compliance selects a real-time, random

sample of content from the total percentage of content that matches chosen policy conditions. If you want reviewers to review all items, configure **100%** in a Communication Compliance policy.

Filter email blasts

Use the **Filter email blasts** setting to exclude messages sent from email blast services. Messages that match the conditions you specify don't generate alerts. This exclusion includes bulk email, such as newsletters, and spam, phishing, and malware. When you select this option, you can view a report that lists the bulk email senders that were filtered out. Reports are retained for 60 days. The list of senders is filtered before the content is analyzed, so there might be senders that don't match the content conditions.

This setting is on by default for new policies. If the **Filter email blasts setting** is off for existing policies, a recommendation is generated to turn on the setting.

Manage alert policies

After you configure a policy, Communication Compliance automatically creates a corresponding alert policy and generates alerts for messages that match conditions defined in the policy. It can take up to 24 hours after creating a policy to start receiving alerts from activity indicators. By default, all policy matches alert triggers are assigned a severity level of medium in the associated alert policy. Communication Compliance generates alerts for a policy once the aggregation trigger threshold level is met in the associated alert policy. A single email notification is sent once every 24 hours for any alerts, regardless of the number of individual messages that match policy conditions. For example, Contoso has an inappropriate content policy enabled and for January 1, there were 100 policy matches that generated six alerts. A single email notification for the six alerts is sent at end of January 1.

For Communication Compliance policies, the system configures the following alert policy values by default:

Expand table

Alert policy trigger	Default value
Aggregation	Simple aggregation
Threshold	Default: 4 activities Minimum: 3 activities Maximum: 2,147,483,647 activities

Alert policy trigger	Default value
Window	Default: 60 minutes Minimum: 60 minutes Maximum: 10,000 minutes

> **Note**
>
> The alert policy threshold trigger settings for activities support a minimum value of 3 or higher for Communication Compliance policies.

You can change the default settings for triggers on number of activities, period for the activities, and for specific users in alert policies on the **Alert policies** page in Microsoft Purview.

Change the severity level for an alert policy

To change the severity level for an alert policy, see Alert policies in the Microsoft Defender portal.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 06/22/2026

Configure conditions for key Communication Compliance scenarios

08/13/2025

> **Important**
>
> Microsoft Purview Communication Compliance provides the tools to help organizations detect regulatory compliance (for example, SEC or FINRA) and business conduct violations such as sensitive or confidential information, harassing or threatening language, and sharing of adult content. Communication Compliance is built with privacy by design. Usernames are pseudonymized by default, role-based access controls are built in, investigators are opted in by an admin, and audit logs are in place to help ensure user-level privacy.

This article provides multiple examples of creating conditions with the condition builder in Microsoft Purview Communication Compliance. Use the examples for your own scenarios and modify as needed for your organization.

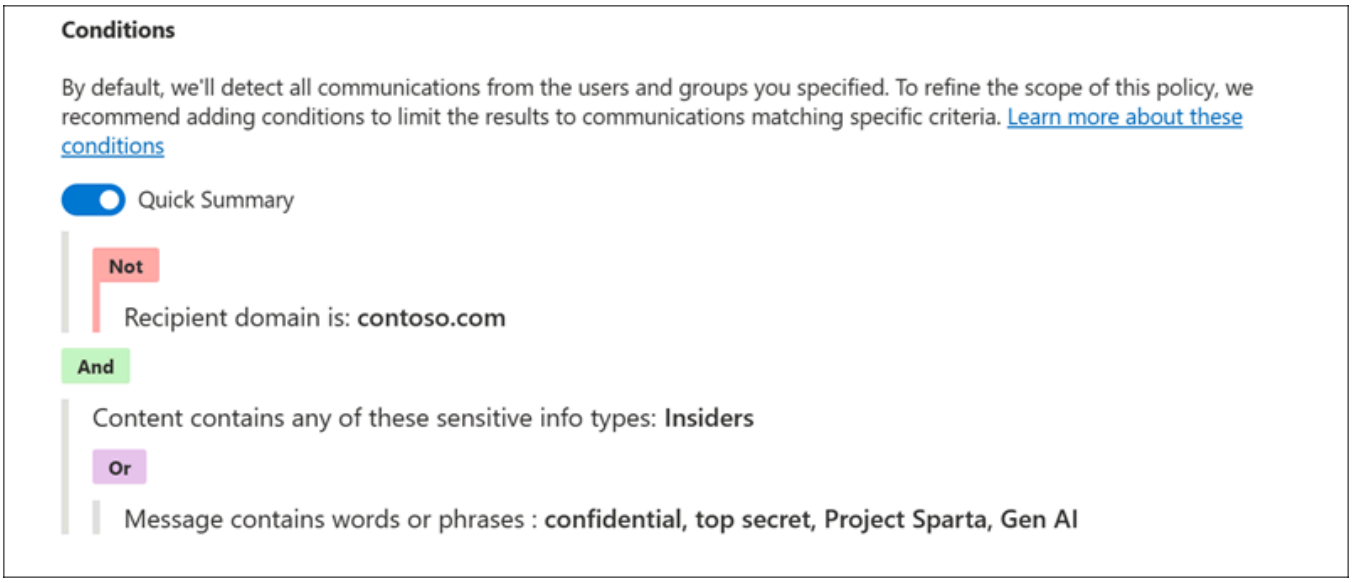
For general information on creating conditions for Communication Compliance policies, see Choose conditions for your Communication Compliance policies.

Data security examples

Example 1:

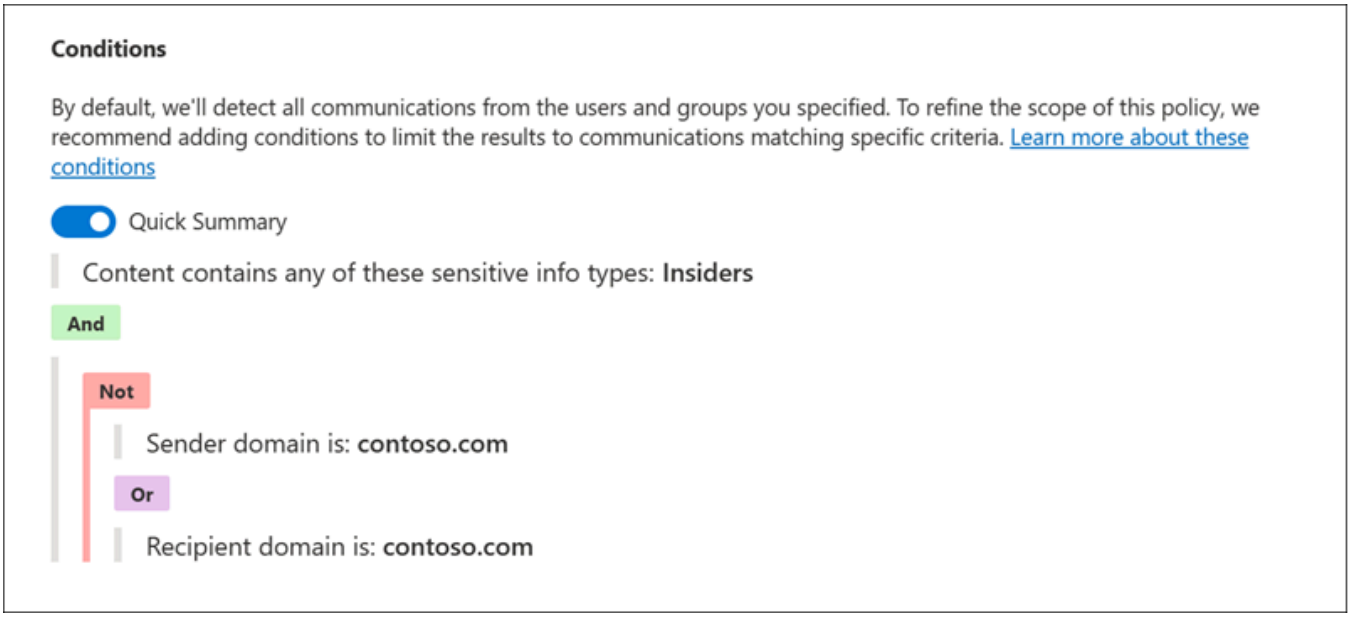
Detect all communications sent outside of Contoso.com that contain patterns detected by the *Insiders* sensitive info type OR that include any of these keywords:

- "confidential"
- "top secret"
- "Project Sparta"
- "GenAI"



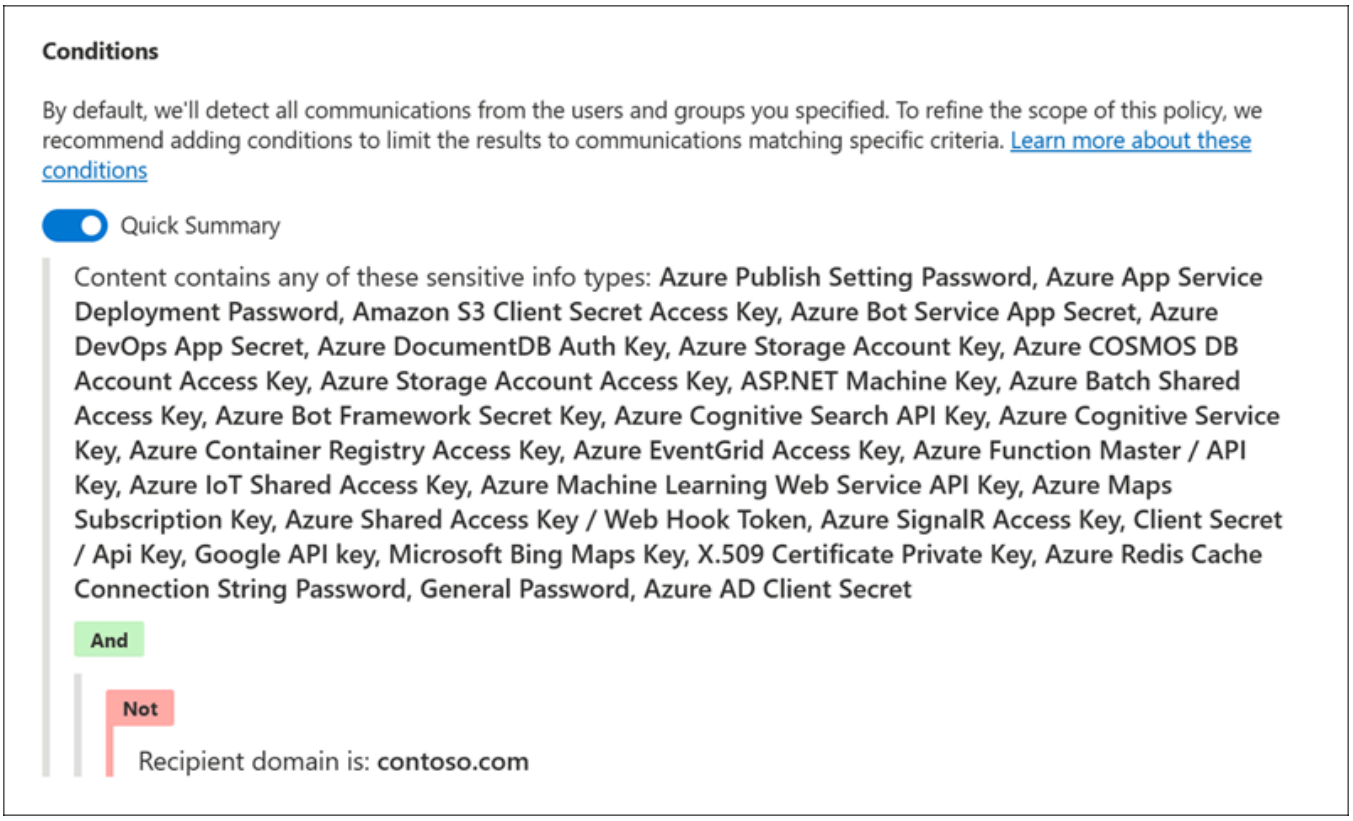
Example 2:

Detect all communications coming from outside or sent outside contoso.com that contains patterns detected by the *Insiders* sensitive info type.



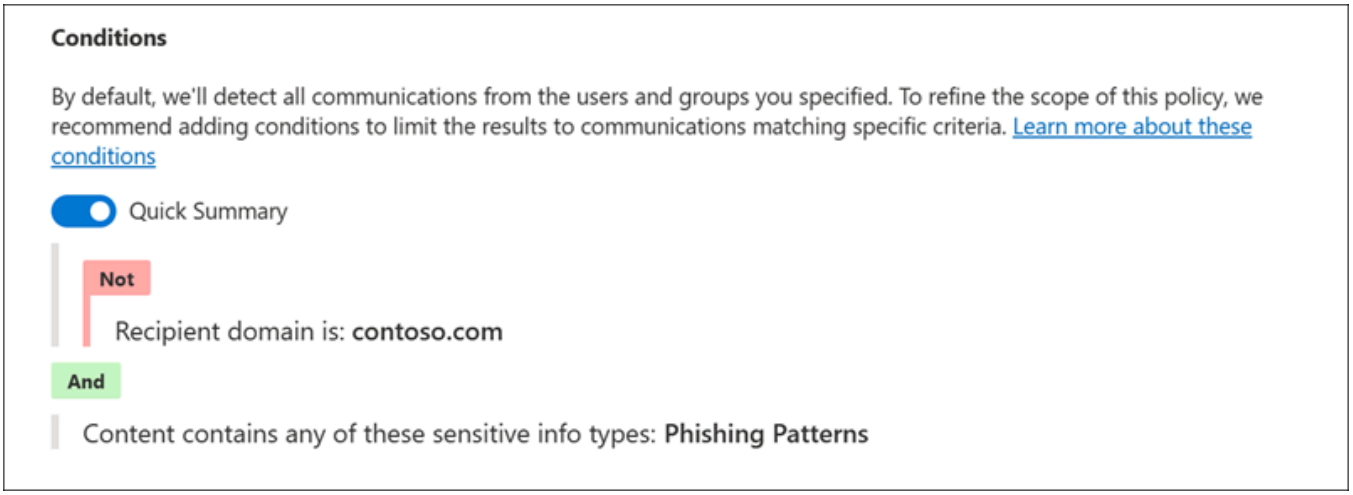
Example 3:

Proactively detect sharing of secret keys and passwords outside of contoso.com.



Example 4:

Detect Teams phishing requests sent as multifactor authentication requests by using the *Phishing Patterns* sensitive info type.

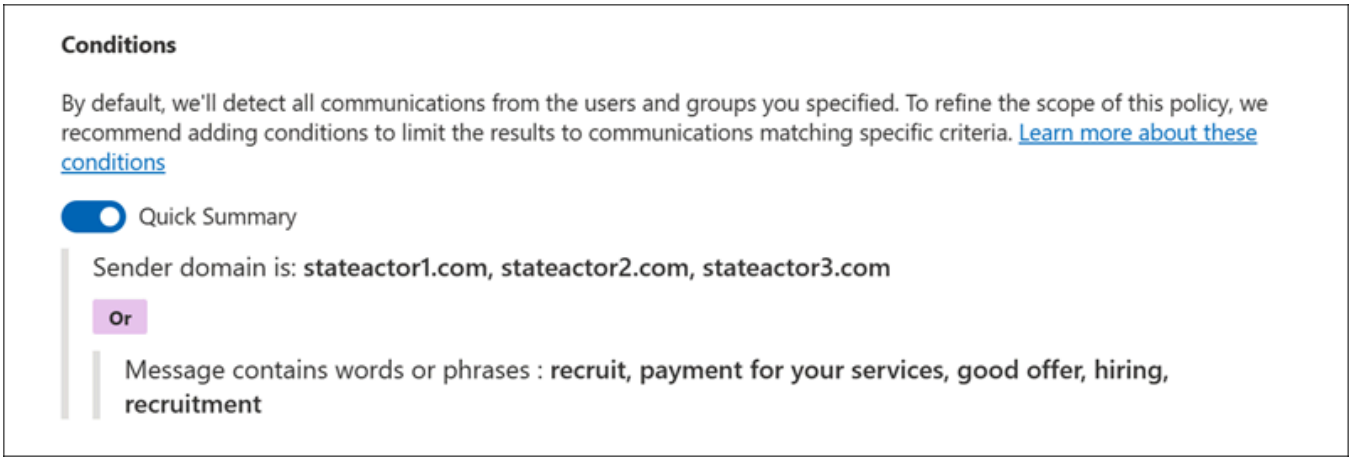


Example 5:

Analyze all messages coming from the stateactor1.com, stateactor2.com, and stateactor3.com domains OR messages containing these keywords:

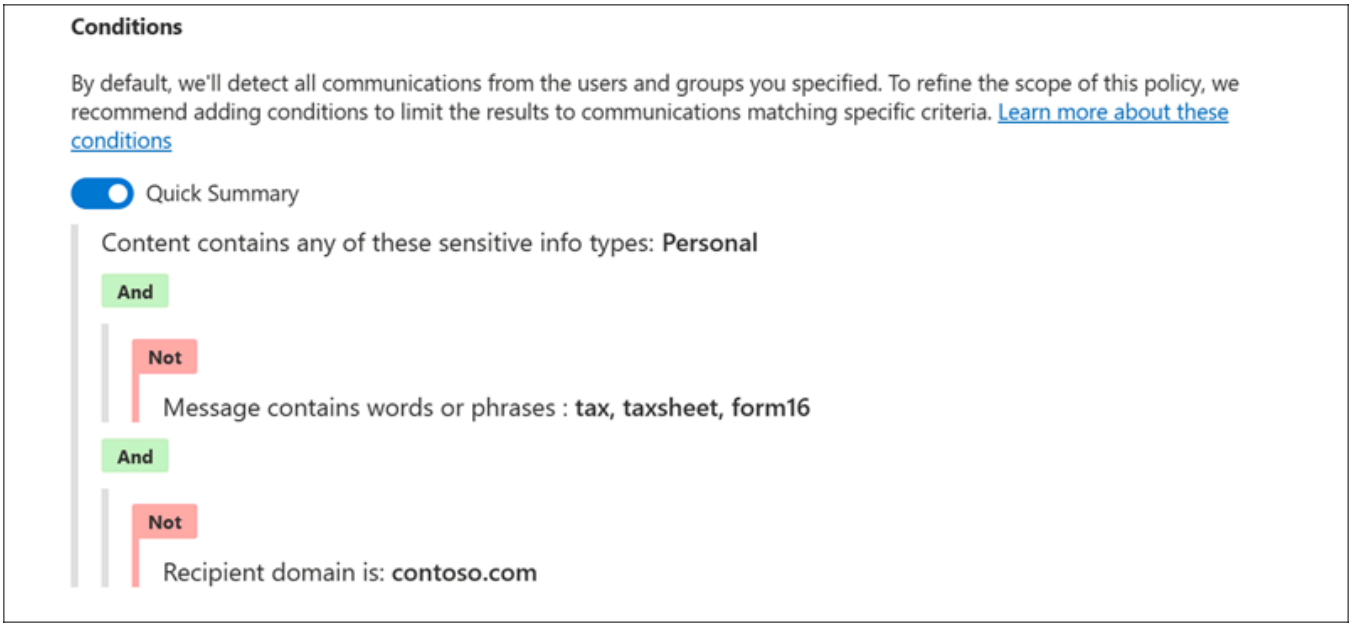
- recruit
- payment for your services
- good offer

- hiring, recruitment



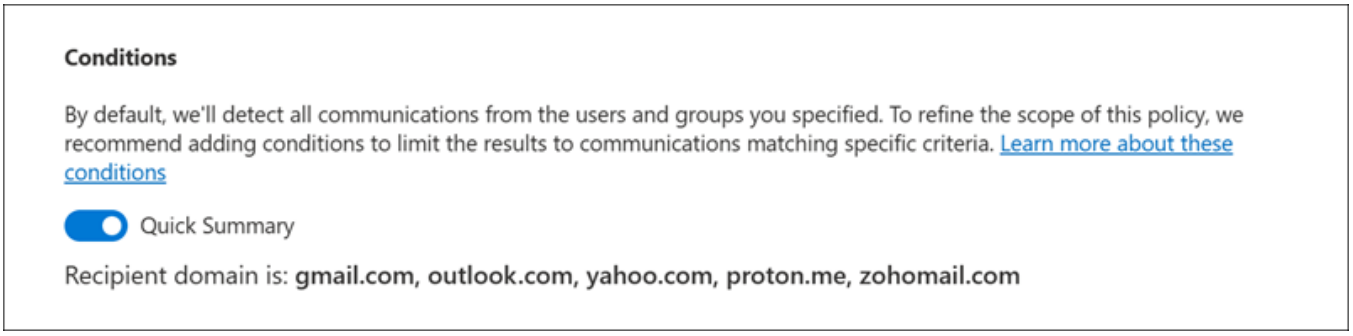
Example 6:

Analyze specific emails sent to external domains by detecting patterns using the *Personal* sensitive info type, but ignore specific keywords.



Example 7:

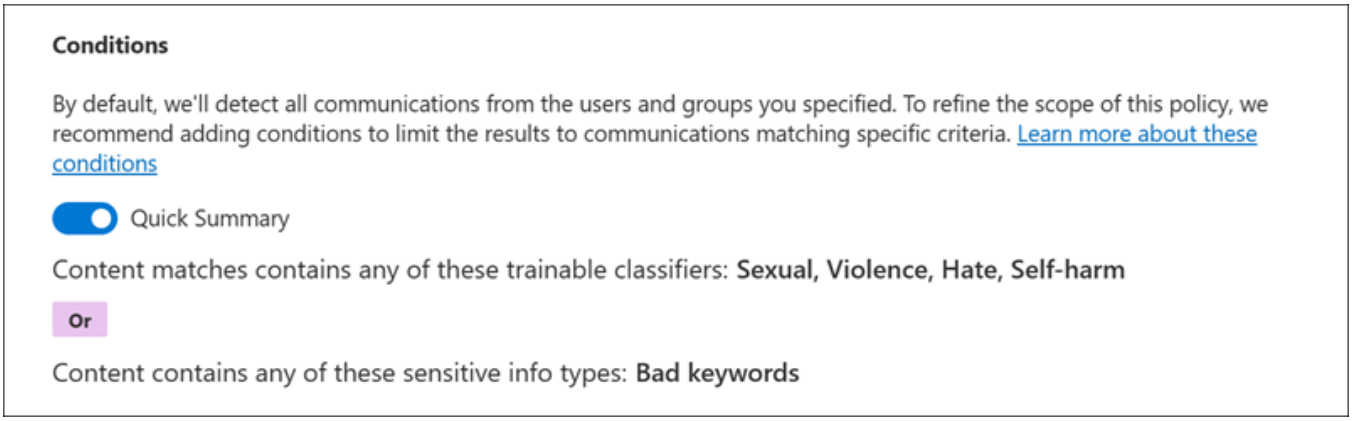
Analyze and view all emails sent to personal email accounts.



Business conduct examples

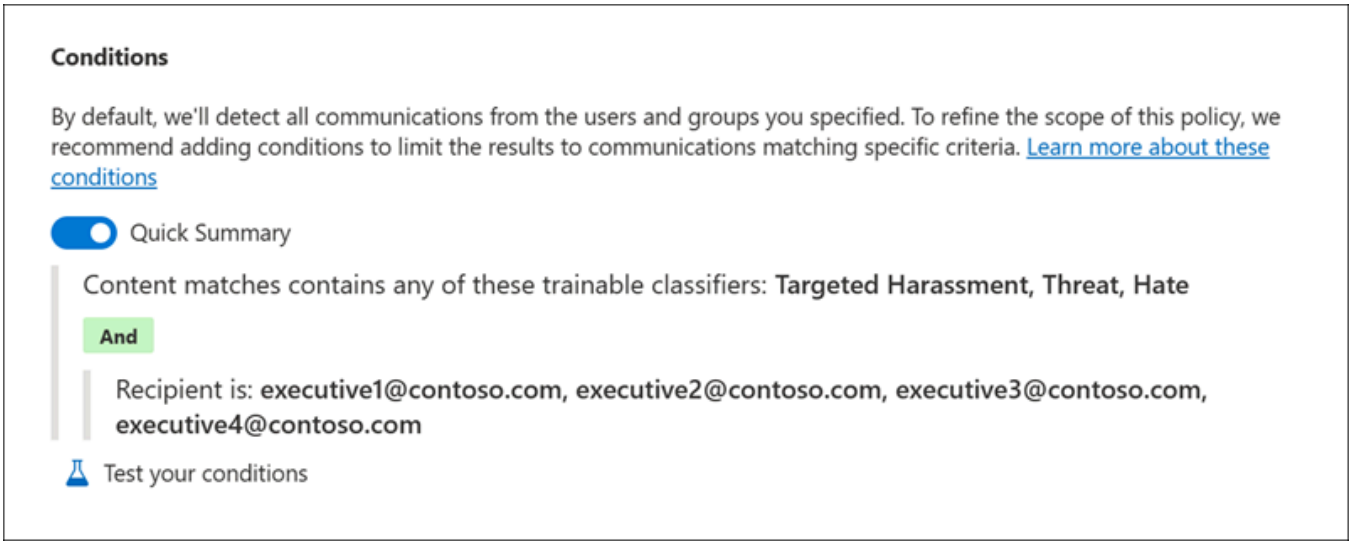
Example 1:

Detect inappropriate content sent in messages by using the *Sexual*, *Violence*, *Hate*, and *Self-harm* Microsoft provided trainable classifiers OR by using a curated keyword dictionary.



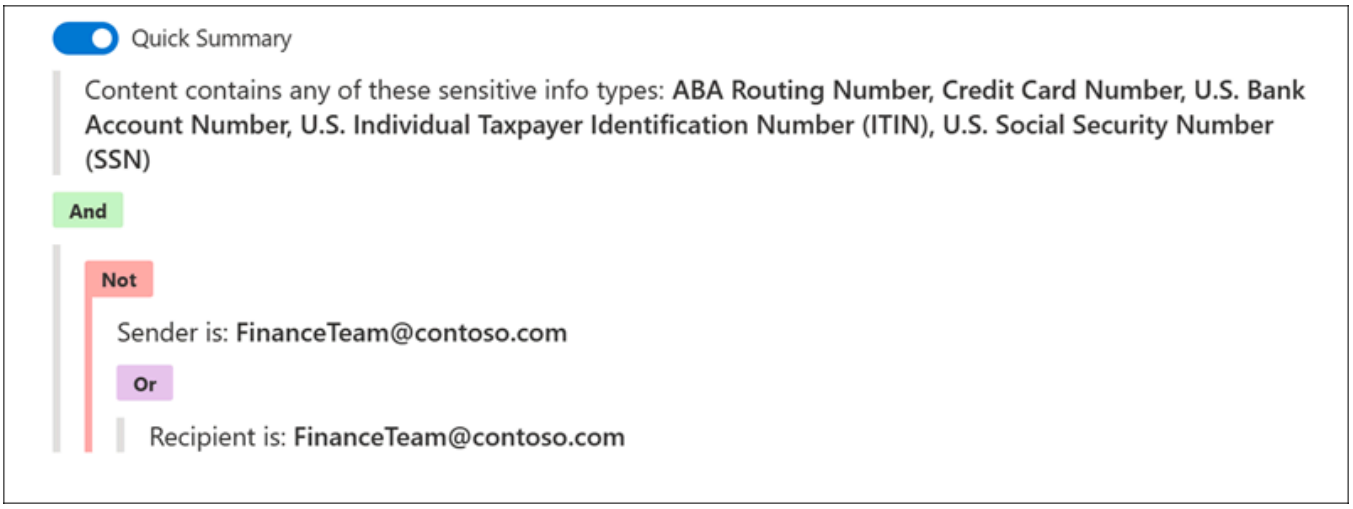
Example 2:

Detect threats and harassment addressed to Contoso's key executives by using the *Targeted Harassment*, *Threat*, and *Hate* Microsoft provided trainable classifiers.



Example 3:

Detect all messages that contain patterns detected by the *ABA Routing Number*, *Credit Card Number*, *U.S. Bank Account Number*, *U.S. Individual Taxpayer Identification Number (ITIN)*, and *U.S. Social Security Number (SSN)* sensitive information types sent by any user, but not sent to or sent by the Finance Team.



See also

- Get started with Communication Compliance policies
- Create and manage Communication Compliance policies

Use Communication Compliance reports and audits

> **Important**
>
> Microsoft Purview Communication Compliance provides the tools to help organizations detect regulatory compliance (for example, SEC or FINRA) and business conduct violations such as sensitive or confidential information, harassing or threatening language, and sharing of adult content. Communication Compliance is built with privacy by design. Usernames are pseudonymized by default, role-based access controls are built in, investigators are opted in by an admin, and audit logs are in place to help ensure user-level privacy.

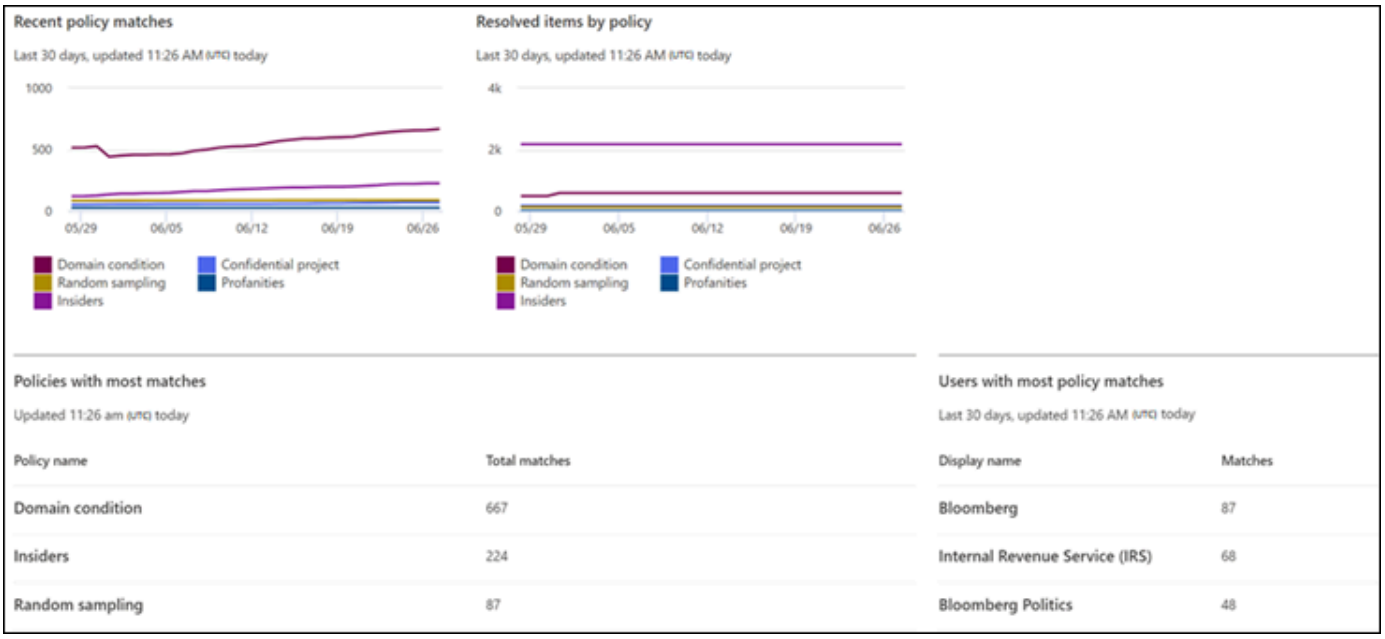
Use Communication Compliance reports

> **Note**
>
> Currently, if your role is scoped by one or more admin units in Microsoft Purview, the admin units don't apply to reports. You can see all reports in the organization.

The **Reports** dashboard is the central location for viewing all Communication Compliance reports. To view and manage reports, assign users to the *Communication Compliance Viewers* role group.

Report widgets provide a quick view of insights most commonly needed for an overall assessment of the status of Communication Compliance activities. You can't export the information contained in the report widgets. Detailed reports provide in-depth information related to specific Communication Compliance areas and offer the ability to filter, group, sort, and export information while reviewing.

For the date range filter, the date and time for events are listed in Coordinated Universal Time (UTC). When filtering messages, all filters for the report apply at the 00:00:00 UTC start date to the 23:59:59 UTC end date.



The **Reports dashboard** contains the following report widgets and detailed reports links:

Report widgets

The Reports dashboard includes the following widgets that provide at-a-glance summaries of policy activity:

- **Recent policy matches**: displays the number of matches by active policies over time.
- **Resolved items by policy**: displays the number of policy match alerts resolved by policies over time.
- **Policy with most matches**: displays the policies and the number of matches for a given period, ranked highest to lowest for matches.
- **Users with most policy match**: displays the users (or anonymized usernames) and number of policy matches for a given period.
- **Escalations by policy**: displays the number of escalations per policy over a given time.

Detailed reports

Use the *Export report* option to create a .CSV file containing the report details for any detailed report. The *Export report* option supports file size downloads up to 3 MB.

- **Policy settings and status**: provides a detailed look at policy configuration and settings, as well as the general status for each of the policy (matches and actions) on messages. Includes policy information and how policies are associated with users and groups, locations, review percentages, reviewers, status, and when the policy was last modified. Use the *Export* option to create a .CSV file containing the report details.

- **Items and actions per policy**: Review and export matching items and remediation actions per policy. Includes policy information and how policies are associated with:
  - Case created
  - Escalated items
  - Items matched
  - Items pending review
  - Resolved items
  - Tagged as compliant
  - Tagged as noncompliant
  - Tagged as questionable
  - User notified

- **Item and actions per location**: Review and export matching items and remediation actions per Microsoft 365 location. Includes information about how workload platforms are associated with:
  - Case created
  - Escalated items
  - Items matched
  - Items pending review
  - Resolved items
  - Tagged as compliant
  - Tagged as noncompliant
  - Tagged as questionable
  - User notified

- **Activity by user**: Review and export matching items and remediation actions per user. Includes information about how users are associated with:
  - Case created
  - Escalated items
  - Items matched
  - Items pending review
  - Resolved items
  - Tagged as compliant
  - Tagged as noncompliant
  - Tagged as questionable
  - User notified

> ⓘ **Note**

The items and actions displayed are only for the items and actions *matched* within the report's selected date range filter.

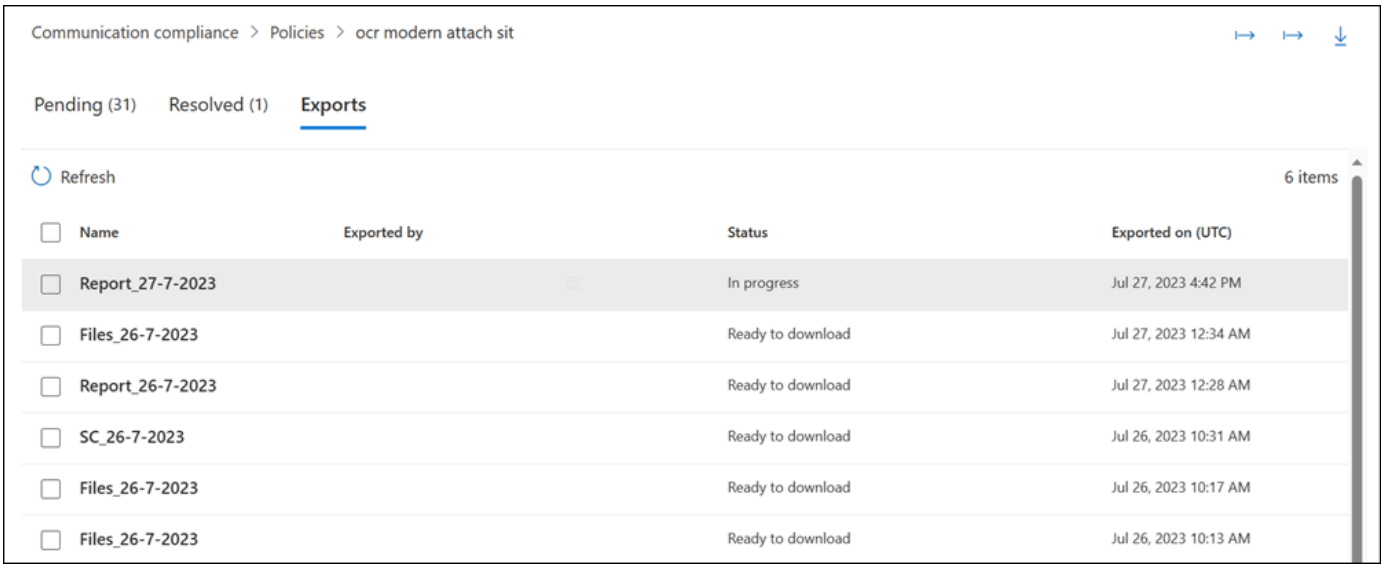
- **Sensitive information type per location** (preview): Review and export information about the detection of sensitive information types and the associated sources in Communication Compliance policies. Includes the overall total and the specific breakdown of sensitive information type instances in the sources configured in your organization. The values for each third-party source are displayed in separate columns in the .CSV file. Examples are:
  - **Email**: Sensitive information types detected in Exchange email messages.
  - **Teams**: Sensitive information types detected in Microsoft Teams channels and chat messages.
  - **Microsoft 365 Copilot and Microsoft 365 Copilot Chat**: Sensitive information types detected in Copilot interactions.
  - **Viva Engage**: Sensitive information types detected in Viva Engage inboxes, posts, chats, and replies.
  - **Third-party sources**: Sensitive information types detected for activities associated with third-party connectors configured in your organization. To view the breakdown of third-party sources for a specific sensitive information type in the report, hover your mouse over the value for the sensitive information type in the Third-party source column.
  - **Other**: Sensitive information types used for internal system processing. Selecting or deselecting this source for the report won't affect any values.

- **Email blast senders**: Review and export the list of senders of email blast messages that were filtered out from your Communication Compliance policies to reduce "noise." **Filter email blasts** is a Communication Compliance policy setting. The Email blast senders report includes the following fields:
  - Policy name
  - Policy last modified date
  - Sender
  - Number of mails filtered

Export a message details report

You can create custom reports and review details for messages brought into the scope of specific policies. Use these reports for all-up reviews of messages and for creating a report snapshot for a customizable time period. You can create a one-time report, or if you often need to create the same report, schedule report creation on a daily, weekly, or monthly basis. You can create up to five schedules per policy.

When you create a report, reviewers for the policy receive an email notification with a link to the **Exports** tab, where they can download the report as a CSV file.



Create the report

To create a message details report, complete the following steps:

1. Sign in to the Microsoft Purview portal using credentials for an admin account in your Microsoft 365 organization.
2. Go to the **Communication Compliance** solution.
3. Select **Policies** in the left navigation.
4. Select a policy, then select **Export report** in the upper-right corner of the page.
5. On the **Export options** pane, under **Include items from these users**, select one of the following options:
    - **All users**. Select this option to create a report with message information for all users in the policy.
    - **Select users**. Select this option, then select specific users from the list, to create a report with messages sent by those specific users. Only the users who sent a message that the system flagged are available in the list.
6. In the **Select frequency** list, select how often you want to create the report.
7. If it's a one-time report, enter a start date and end date for the report. If it's a recurring report, enter a date to start sending reports and a date to stop sending reports.

> **Note**
>
> If you select **Daily**, **Weekly**, or **Monthly** for a recurring report, the report starts collecting messages from the previous day, week, or month, depending on which option you select. For example, if you choose **Daily** as the frequency and choose **June 4, 2024** as the start date for the report, the report starts collecting messages on June 3.

8. In the **Send report to these reviewers** list, add or delete reviewers as needed. By default, all reviewers for the policy (added during policy creation) are listed.
9. In the **Report name** field, accept the suggested name for the report or change it as needed.
10. Select **Create report** (for one-time reports) or **Start scheduled export** (for recurring reports).

    The **Preparing scheduled report** dialog box appears.
11. Select **Manage scheduled exports** to view the schedule of reports.

> **Tip**
>
> To edit a report schedule, select **Manage scheduled reports** from the **Exports** tab.

12. To make changes to a scheduled report:
    a. Select the checkbox for the report, then select **Edit**.
    b. In the **Export options** dialog box, make the changes that you want.

> **Note**
>
> To change the frequency or start date of the report, delete the report from the **Manage scheduled exports** page, then create the report again.

> **Note**
>
> You can create up to five scheduled reports per policy. If you already have five scheduled reports for a policy and want to create a new one, delete one of the existing report schedules before you can create a new one.

Download a report when it's ready

The time it takes to create a report depends on the number of scheduled reports and the size of the reports. Reviewers on the policy receive an email notification with a link to the **Exports** tab when a report is ready. To download a report from the **Exports** tab, select a report in the *Ready to download* state, then select **Download export(s)**.

> ⓘ **Note**
>
> If your selected time period doesn't return any message results in the report, it means there were no messages for the selected time period. The report is blank.

What's included in a Message details report

Message details reports contain the following information for each message item in the policy:

- **Match ID**: Unique ID for a copy of the message in Communication Compliance.
- **Internet Message ID**: Unique ID for the message across platforms.
- **Conversation Family ID**: Thread ID for the message.
- **Sender column**: Sender of the message. If the policy match is a response from Microsoft 365 Copilot and Microsoft 365 Copilot Chat, the value is "Copilot".
- **Recipient column**: Recipients included in the message. If the policy match is a prompt to Copilot, the value is "Copilot".
- **Date**: The date when the message was sent.
- **Location**: Channel where the message was sent. This channel can be Exchange Online, Teams, Viva Engage, or any third-party channel that Communication Compliance supports.
- **Subject**: Subject of the message. If it's a Copilot interaction, the subject of the message is "Copilot" and the name of the Microsoft 365 app. For example: "Copilot in Excel".
- **Contains Attachments**: Status of any attachments for the message. Values are either *Yes* or *No*.
- **Policy Name**: Name of the policy associated with the message. This value is the same for all messages in the report.
- **Item Status**: Status of the message item in the policy. Values are *Pending* or *Resolved*.
- **Tags**: Tags assigned to the message. Values are *Questionable, Compliant*, or *Noncompliant*.
- **Keyword Matches**: Keyword matches for the message.
- **Trainable Classifier ID**: ID of the trainable classifier that was matched.
- **Trainable Classifier Name and Matched Keywords**: The name of the trainable classifier and the keywords that triggered the classifier match.
- **Reviewers**: Reviewers assigned to message.

- **Pending for (days)**: Number of days the message has been in a pending state. For resolved messages, the value is 0.
- **Comment for Resolved**: Comments for the message entered when resolved.
- **Resolved Date**: Date and Coordinated Universal Time (UTC) the message was resolved.
- **Last Updated By**: User name of the last updater.
- **Last Updated On**: Date and Coordinated Universal Time (UTC) the message was last updated.
- **History of Comments**: List of all comments for the message alert, including comment author and date and Coordinated Universal Time (UTC) of the comment.

> **Tip**
>
> You can also archive message details outside of Communication Compliance.

Review Communication Compliance audit data

In some instances, you must provide information to regulatory or compliance auditors to prove that user activities and communications are scoped. This information might be a summary of all activities associated with a defined organizational policy or anytime a Communication Compliance policy changes. Communication Compliance policies have built-in audit trails for complete readiness for internal or external audits. Your communication policies capture detailed audit histories of every create, edit, and delete action to provide proof of scoped procedures.

> **Important**
>
> You must enable auditing for your organization before Communication Compliance events are recorded. To enable auditing, see Step 2: Enable the audit log in Communication Compliance configuration. When activities trigger events that the Microsoft 365 Audit log captures, it can take up to 48 hours before you can view these events in Communication Compliance policies.

To view Communication Compliance policy update activities, select the **Export policy updates** control on the main page for any policy. You must be assigned the *Global Admin* or *Communication Compliance Admins* roles to export update activities. This action generates an audit file in the .CSV format that contains the following information:

> **Important**

Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

Expand table

Field	Details
AuditData	Main data source for all policy update activities. All update activities are recorded and separated by comma delimiters.
CreationDate	Date the update activity was performed in a policy.
Operations	Update operations performed on the policy.
UserIds	User that performed the update activity in a policy.

To view Communication Compliance review activities for a policy, select the **Export review activities** control on the **Overview** page for a specific policy. You must be assigned the *Global Admin* or *Communication Compliance Admins* roles to export review activities. This action generates an audit file in the .CSV format that contains the following information:

Important

Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

Expand table

Field	Details
AuditData	Main data source for all policy review activities. All review activities are recorded and separated by comma delimiters.
CreationDate	Date the review activity was performed in a policy.
Operations	Review operations performed on the policy.
UserIds	User that performed the review activity in a policy.

You can also view audit activities in the unified audit log or with the Search-UnifiedAuditLog PowerShell cmdlet. To learn more about audit log retention policies, see Manage audit log

retention policies.

For example, the following command searches the unified audit log for all scoped review activities (policies and rules):

PowerShell

```
Search-UnifiedAuditLog -StartDate $startDate -EndDate $endDate -RecordType AeD -
Operations SupervisoryReviewTag
```

To audit when Communication Compliance policies are created, updated, or deleted, run the following command:

PowerShell

```
Search-UnifiedAuditLog -StartDate $startDate -EndDate $endDate -RecordType Discovery
-Operations SupervisionPolicyCreated,SupervisionPolicyUpdated,SupervisionPolicyDeleted
```

To review events where content matched your current Communication Compliance policy rules, search the audit log with the following command:

PowerShell

```
Search-UnifiedAuditLog -StartDate $startDate -EndDate $endDate -Operations
SupervisionRuleMatch
```

Communication Compliance policy matches are stored in a scoped mailbox for each policy. In some cases, you might need to check the size of your scoped mailbox for a policy to make sure you aren't approaching the current 100-GB storage size or 1 million message limit. If you reach the mailbox limit, the system doesn't capture policy matches and you need to create a new policy (with the same settings) to continue to capture matches for the same activities.

To check the size of a scoped mailbox for a policy, complete the following steps:

1. Connect to Exchange Online PowerShell.
2. Run the following command:

   PowerShell

   ```
   ForEach ($p in Get-SupervisoryReviewPolicyV2 | Sort-Object Name)
   {
      "<Name of your Communication Compliance policy>: " + $p.Name
   ```

```
    Get-MailboxStatistics $p.ReviewMailbox | ft ItemCount,TotalItemSize
}
```

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/22/2026

Detect channel signals with Communication Compliance

08/28/2025

> **Important**
>
> Microsoft Purview Communication Compliance provides the tools to help organizations detect regulatory compliance (for example, SEC or FINRA) and business conduct violations such as sensitive or confidential information, harassing or threatening language, and sharing of adult content. Communication Compliance is built with privacy by design. Usernames are pseudonymized by default, role-based access controls are built in, investigators are opted in by an admin, and audit logs are in place to help ensure user-level privacy.

With Communication Compliance policies, you can choose to analyze messages in one or more of the following communication platforms as a group or as standalone sources. Your organization retains original messages captured across these platforms in the original platform location in accordance with your retention and hold policies. Communication Compliance policies retain copies of messages used for analysis and investigation for as long as a policy is in place, even if users leave your organization and their mailboxes are deleted. When you delete a communication policy, you also delete copies of messages associated with the policy.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Generative AI

> **Important**
>
> To detect inappropriate or risky interactions for non-Microsoft 365 AI data, you need to enable pay-as-you-go billing in your organization. Non-Microsoft 365 AI data includes information from other generative AI applications from Microsoft and other connected external AI applications. This data type includes Copilot in Microsoft Fabric, Microsoft Security Copilot, Microsoft Copilot Studio, and any connected or cloud AI application.

There aren't any pay-as-you-go billing requirements or charges for Microsoft 365 detecting inappropriate or risky interaction for Microsoft 365 Copilot data.

Communication Compliance supports numerous generative AI applications. You can use Communication Compliance policies to analyze interactions (prompts and responses) entered into these applications and help detect inappropriate or risky interactions or sharing of confidential information. Communication Compliance supports the following generative AI applications:

- **Microsoft Copilot experiences**: Support for user interactions in Microsoft 365 Copilot and other Copilots built using Microsoft Copilot Studio.
- **Enterprise AI apps**: Non-Copilot AI apps that you onboard or connect to your organization by using methods like Microsoft Entra registration and data connectors.
- **Other AI apps**: Other AI applications detected through browser activity and categorized as *Generative AI* in the Defender for Cloud Apps catalog.

Use the following links to learn more about creating a Communication Compliance policy that detects Microsoft 365 Copilot and Microsoft 365 Copilot Chat interactions and other Microsoft Purview data security and compliance protections for Microsoft 365 Copilot and Microsoft 365 Copilot Chat:

- Configure a Communication Compliance policy to detect for Microsoft 365 Copilot and Microsoft 365 Copilot Chat interactions
- Microsoft Purview data security and compliance protections for generative AI apps

Microsoft Teams

You can analyze chats in public and private Microsoft Teams channels and individual communications. When you assign users to a Communication Compliance policy with Microsoft Teams coverage selected, the solution automatically detects chat communications across all Microsoft Teams where users are a member.

To learn how to detect communication risks in Microsoft Teams with Communication Compliance, watch the following video:

https://learn-video.azurefd.net/vod/player?id=7857b91f-d0c1-4b9b-ac84-5c70a8453216&locale=en-us&embedUrl=%2Fpurview%2Fcommunication-compliance-channels

Microsoft Purview Communication Compliance automatically includes Microsoft Teams coverage for predefined policy templates and selects it as the default in the custom policy

template. Teams transcripts are also included. Teams chats matching Communication Compliance policy conditions might take up to 48 hours to process.

For Teams private chat and private channels, Communication Compliance policies support modern attachment analysis. Shared Channels support in Teams is handled automatically and doesn't require additional Communication Compliance configuration changes. The following table summarizes Communication Compliance behavior when sharing Teams channels with groups and users:

Expand table

Scenario	Communication Compliance behavior
Share a channel with an internal team	Communication Compliance policies apply to in-scope users and all messages in the shared channel
Share a channel with an external team	Communication Compliance policies apply to internal in-scope users and messages in the shared channel for the internal organization

Modern attachments are files sourced from OneDrive or SharePoint sites that are included in Teams messages. Text is automatically extracted from these attachments for automated processing and potential matches with active Communication Compliance policy conditions and classifiers. There's no extra configuration necessary for modern attachment detection and processing. Text is only extracted for attachments matching policy conditions at the time the message is sent. Text isn't extracted for attachments for messages with policy matches, even if the attachment also has a policy match.

Modern attachment analysis is supported for the following file types:

- Microsoft Word (.docx)
- Microsoft Excel (.xlsx, .xls, .csv)
- Microsoft PowerPoint (.pptx)
- Text (.txt)
- Portable Document Format (.pdf)

Modern attachments like files hosted on OneDrive or SharePoint sites are supported if shared as Teams messages. The extracted text from these attachments is saved as .txt files and included with the Teams message in Communication Compliance. Additionally, the URL to the original file is attached separately for review. Extracted text for modern attachments is included with the associated message on the **Pending** tab of the policy's dashboard. The extracted text for an attachment is named as the attachment file name (and format extension) and the .txt extension. For example, the extracted text for an attachment named *ContosoBusinessPlan.docx* appears as *ContosoBusinessPlan.docx.txt* in the **Pending** tab of the policy's dashboard.

Select the extracted attachment text to view the details in the *Source* and *Plain text* views. After reviewing, you can resolve or take action on the attachment text using the command bar controls. You can also download the attachment for review outside of the Communication Compliance review process.

Use the following group management configurations to bring individual user chats and channel communications in Teams in scope:

- **For Teams chats:** Assign individual users or assign a distribution group to the Communication Compliance policy. This setting is for one-to-one or one-to-many user/chat relationships.

- **For Teams channel communications:** If any member of the channel is a scoped user within a policy and the **Inbound** direction is configured as the **Communication direction** for the policy, all messages sent within the channel are subject to review by Communication Compliance, even for users in the channel that aren't explicitly scoped.

  For example, consider the following set of circumstances:

- User A is scoped in a Communication Compliance policy.
  - User A is a member of a Teams channel.
  - **Inbound** is selected as the direction for the policy.
  - User B and User C aren't scoped in a Communication Compliance policy.
  - User B and User C are members of the same Teams channel as User A.
  - User B and User C are posting in the channel, but aren't communicating directly with User A.

  In this scenario, a message from User B or User C can be flagged as a policy match even if they don't communicate directly with user A. This result is because User A is technically receiving the message from User B and User C since they're part of the same channel. In this case, if you want to exclude User B and User C from potentially being flagged for policy matches, select a different communication direction in the applicable Communication Compliance policy.

  On the other hand, if User B and User C have a conversation on another channel or in a chat that doesn't include User A, the conversation isn't reviewed by the policy even if the conversation would otherwise provide a policy match since User A, the scoped user, isn't included in the conversation.

- **For Teams chats with hybrid email environments**: Communication Compliance can detect chat messages for organizations with an Exchange on-premises deployment or an external email provider that enabled Microsoft Teams. You must create a distribution group for the users with on-premises or external mailboxes. When creating a Communication Compliance policy, assign this distribution group using **Choose users**

and groups in the policy workflow. For more information about the requirements and limitations for enabling cloud-based storage and Teams support for on-premises users, see Search for and export Teams chat data for on-premises users.

> **Tip**
>
> For a more flexible configuration and to reduce administrative overhead, consider using an **adaptive scope** instead of a distribution group.

Exchange

You can analyze all mailboxes hosted on Microsoft Exchange Online as part of your Microsoft 365 or Office 365 subscription. Exchange email messages and attachments that match Communication Compliance policy conditions might take about 24 hours to process. Supported attachment types for Communication Compliance are the same as the file types supported for Exchange mail flow rule content inspections.

Viva Engage

You can also analyze private messages, public conversations, and associated attachments in Microsoft Viva Engage communities. When you add users to a Communication Compliance policy that includes Viva Engage as a defined channel, the analysis includes communications across all Viva Engage communities that a user is a member of. Viva Engage chats without attachments that match Communication Compliance policy conditions can take up to one hour to process. Viva Engage chats with attachments can take up to 24 hours to process.

You must configure Viva Engage in Native Mode for Communication Compliance policies to detect Viva Engage communications and attachments. In Native Mode, all Viva Engage users are in Microsoft Entra ID, all groups are Microsoft 365 Groups, and all files are stored in SharePoint.

Non-Microsoft sources

You can analyze communications for data imported into all mailboxes in your Microsoft 365 organization from non-Microsoft sources like Instant Bloomberg, Slack, Zoom, SMS, and many others. For a full list of connectors supported in Communication Compliance, see Learn about connectors for third-party data.

You must configure a third-party connector for your Microsoft 365 organization before you can assign the connector to a Communication Compliance policy. The **Third-Party Sources** section

of the Communication Compliance policy workflow only displays currently configured third-party connectors.

> **Note**
>
> Content from all non-Microsoft channels (such as WhatsApp or Instant Bloomberg) takes 24 to 48 hours for detection by Communication Compliance, depending on how frequently the data is imported from those non-Microsoft channels.

Channel limits

Expand table

Channel	Current limit
All channels	Maximum of 25 attachments per policy match
Teams public channels	Maximum of 250 users for public channel messages for static scopes with users
Teams public channels	No support or coverage for public channel messages for adaptive scopes with users

Configure a Communication Compliance policy to detect generative AI interactions

> **Important**
>
> Microsoft Purview Communication Compliance provides the tools to help organizations detect regulatory compliance (for example, SEC or FINRA) and business conduct violations such as sensitive or confidential information, harassing or threatening language, and sharing of adult content. Communication Compliance is built with privacy by design. Usernames are pseudonymized by default, role-based access controls are built in, investigators are opted in by an admin, and audit logs are in place to help ensure user-level privacy.

Overview

You can use Communication Compliance to analyze interactions (prompts and responses) to detect inappropriate or risky interactions or sharing of confidential information entered into numerous generative AI applications.

These applications include Microsoft 365 Copilot, Copilots built with Microsoft Copilot Studio, AI applications connected by Microsoft Entra or Microsoft Purview Data Map connectors, and more.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Microsoft Copilot experience

> **Important**
>
> To detect inappropriate or risky interactions for non-Microsoft 365 AI data, you must enable pay-as-you-go billing in your organization. Non-Microsoft 365 AI data includes information from other generative AI applications from Microsoft and connected external AI applications. Examples of non-Microsoft 365 AI data include Copilot in Microsoft Fabric, Microsoft Security Copilot, Microsoft Copilot Studio, and any connected or cloud AI application.

There are no pay-as-you-go billing requirements or charges for Microsoft 365 detecting inappropriate or risky interactions for Microsoft 365 Copilot data.

Communication Compliance can detect interactions in any message with the IPM.SkypeTeams.Message.Copilot.*application* item class. These item classes include Copilot applications in Microsoft solutions like Teams, Outlook, and many more.

For example, Communication Compliance detects interactions in Copilot messages with the IPM.SkypeTeams.Message.Copilot.Teams, IPM.SkypeTeams.Message.Copilot.Outlook item classes, and others.

Connected generative AI application

> **Important**
>
> To detect inappropriate or risky interactions for connected generative AI applications, you must enable pay-as-you-go billing in your organization.

Communication Compliance can detect prompt and response interactions with non-Copilot AI applications. These applications are generative AI applications connected by using Microsoft Entra and Microsoft Purview Data Map connectors.

Other AI applications

> **Important**
>
> To detect inappropriate or risky interactions for other AI applications, your organization must enable pay-as-you-go billing.

Communication Compliance can also detect interactions with AI applications from browser and network activity by users in your organization. This capability helps you detect inappropriate or risky interactions or sharing of confidential information that users enter into AI applications outside of your organization.

Prerequisites

To investigate Copilot interactions in Communication Compliance, you must be assigned one of the following roles:

- *Communication Compliance*
- *Communication Compliance Investigators*
- *Communication Compliance Analysts*

You must also be assigned as a reviewer in the Reviewers field when creating a communication compliance policy.

How it works

> **Important**
>
> Microsoft is committed to making sure artificial intelligence (AI) systems are developed responsibly and in ways that warrant people's trust. As part of this commitment, Microsoft Purview engineering teams operationalize the six core principles of Microsoft's Responsible AI strategy to design, build, and manage AI solutions. To responsibly deploy AI, we provide documentation, role-based access, scenario attestation, and more to help organizations use AI systems responsibly.

You can take advantage of all Communication Compliance features when you create a Communication Compliance policy that detects Microsoft 365 Copilot and Microsoft 365 Copilot Chat interactions, including:

- Sensitive information types
- Keywords
- Microsoft provided trainable classifiers
- Content safety classifiers for Teams based on large language models
- Conditions

Any prompt or response that a user enters into a supported generative AI app and that matches a Communication Compliance policy appears as a policy match on the **Policies** page on the **Pending** tab, with separate entries for prompts and responses. If only the prompt or only the response matches a policy, the **Pending** tab shows an item just for that policy match. You can remediate policy matches for generative AI apps in the same way that you remediate any other policy match.

The following information appears for each item on the **Pending** tab for AI policy matches:

- **Icons**: The Copilot icon identifies the policy match as a generative AI interaction with a Microsoft-based Copilot For all other generative AI interactions, this icon is an email icon.
- **Subject column**: The *[Copilot]* value in this column identifies the policy match as a generative AI interaction with Microsoft-based Copilots. The *[AI app]* value in this column identifies the policy match for all other generative AI interactions.
- **Sender column**: Sender of the message. Depending on the source AI application, the sender appears as follows:
  - If the policy match is a response from a Copilot application, the value is *Copilot*.
  - If the policy match is a response from a connected AI application, the value is *Connected AI app*.
  - If the policy match is a response from a cloud AI application, the value is *Cloud AI app*.
- **Recipient column**: Recipients included in the message. This value is the user interacting with the AI application.
- **Message text**: The message text that the user entered (the text that caused the policy match) appears on the right side of the screen in its entirety.

Create a policy that detects Microsoft Copilot interactions

Follow these steps to create a policy that detects Microsoft Copilot interactions:

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.
2. Go to the **Communication Compliance** solution.
3. Select **Policies** in the left navigation.
4. Select **Create policy**, then select the **Detect Microsoft Copilot interactions** template.
5. Enter the policy name, select the users and groups to apply the policy to, and select the reviewers for the policy. Learn more about these options when creating a policy from a template
6. Review the list of settings chosen for you based on the template. Select **Create policy** to create the policy or select **Customize policy** if you want to make any changes before creating the policy.

Add a generative AI app as a location for an existing policy

Follow these steps to add a generative AI app as a location for an existing Communication Compliance policy:

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.
2. Go to the **Communication Compliance** solution.
3. Select **Policies** in the left navigation.
4. Select **More actions** (ellipsis) in the row for the policy you want to change, then select **Edit**.
5. Select **Next** two times in the policy creation workflow to go to the **Choose locations to detect communications** page.
6. Select one or more of the following checkboxes to add a generative AI application as a location:
    - **Microsoft Copilot experiences**
    - **Enterprise AI apps**
    - **Other AI apps**
7. Make any other changes to the policy, then on the **Review and finish** page, select **Save**.

Create a policy to review all generative AI interactions

When you first work with generative AI interactions, consider reviewing all AI interactions to understand how people in your organization use these applications. To create a policy that reviews all generative AI interactions, when you create or edit the policy:

- Set the location to include all generative AI applications:
    - **Microsoft Copilot experiences**
    - **Enterprise AI apps**
    - **Other AI apps**
- Set the **Review percentage** option on the **Choose conditions and review percentage** page to 100%.
- Don't set any conditions for the policy.

> **Note**
>
> Depending on the size of your organization, a policy that detects all generative AI interactions might result in a high volume of detected messages. This volume could cause

your organization to reach its **storage limit**. If your organization reaches its storage limit, adjust the policy to reduce the number of detections.

Remediate policy matches and alerts that contain generative AI interactions

Remediate policy matches and alerts that contain generative AI interactions the same way you remediate any policy match or alert in Communication Compliance. For example, you can tag a policy match, escalate it, resolve it, download it, or export it. Learn more about resolving policy matches and alerts in Communication Compliance.

Reports

AI interactions that you bring into the scope of a Communication Compliance policy appear in Communication Compliance reports and audit data. Learn more about Communication Compliance reports and audits.

See also

- Microsoft Purview data security and compliance protections for generative AI apps
- Investigate and remediate Communication Compliance alerts
- Use Communication Compliance reports and audits

ⓘ**Note:** The author created this article with assistance from AI. Learn more

Last updated on 06/22/2026

Microsoft Teams and Microsoft Purview Communication Compliance

08/20/2025Applies to: Microsoft Teams

> **Important**
>
> Microsoft Purview Communication Compliance provides the tools to help organizations detect regulatory compliance (for example, SEC or FINRA) and business conduct violations such as sensitive or confidential information, harassing or threatening language, and sharing of adult content. Communication Compliance is built with privacy by design. Usernames are pseudonymized by default, role-based access controls are built in, investigators are opted in by an admin, and audit logs are in place to help ensure user-level privacy.

Microsoft Purview Communication Compliance is an insider risk solution in Microsoft 365 that helps you minimize communication risks by helping you detect, capture, and act on inappropriate messages in your organization.

For Microsoft Teams, Communication Compliance helps you identify the following types of inappropriate content in Teams channels, private Teams channels, or in 1:1 and group chats:

- Offensive, profane, and harassing language
- Adult, racy, and gory images
- Sharing of sensitive information

Watch the following video to learn how to detect communication risks in Microsoft Teams with Communication Compliance:

https://learn-video.azurefd.net/vod/player?id=7857b91f-d0c1-4b9f-ac84-5c70a8453216&locale=en-us&embedUrl=%2Fpurview%2Fcommunication-compliance-teams

You can also detect offensive communications included in Microsoft Teams transcripts (preview).

For more information on Communication Compliance and how to configure policies for your organization, see Learn about Communication Compliance.

> **Tip**

Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

How to use Communication Compliance in Microsoft Teams

Communication Compliance and Microsoft Teams are tightly integrated and can help you minimize communication risks in your organization. After you configure your first Communication Compliance policies, you can actively manage inappropriate Microsoft Teams messages and content that the solution automatically flags in alerts.

Getting started

To get started with Communication Compliance in Microsoft Teams, plan and create predefined or custom policies to identify inappropriate user activities in Teams channels or in 1:1 and group chats. Remember that you need to configure some permissions and basic prerequisites as part of the configuration process.

Teams administrators can configure Communication Compliance policies at the following levels:

- **User level**: Apply these policies to an individual Teams user or to all Teams users in your organization. These policies cover messages that users send in 1:1 or group chats. The solution automatically detects chat communications for the users across all Microsoft Teams where the users are a member.
- **Teams level**: Apply these policies to a Microsoft Teams channel, including a Private channel. These policies cover messages sent in the Teams channel only.

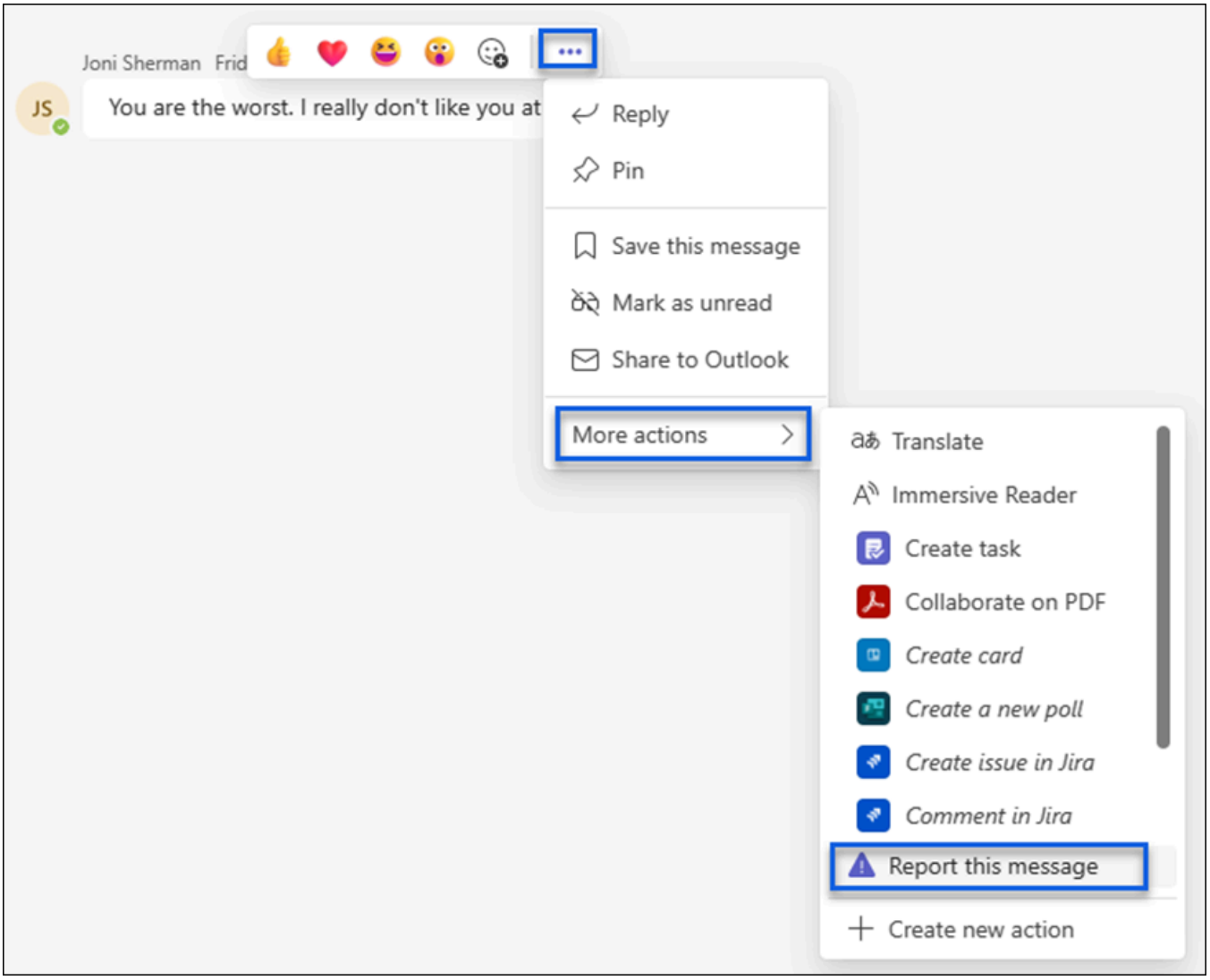
Report a message in Microsoft Teams

> **Note**
>
> After you create your first Communication Compliance policy, the organization implements the *User-reported messages* policy. It can take up to 30 days for this feature to become available after you create your first policy.

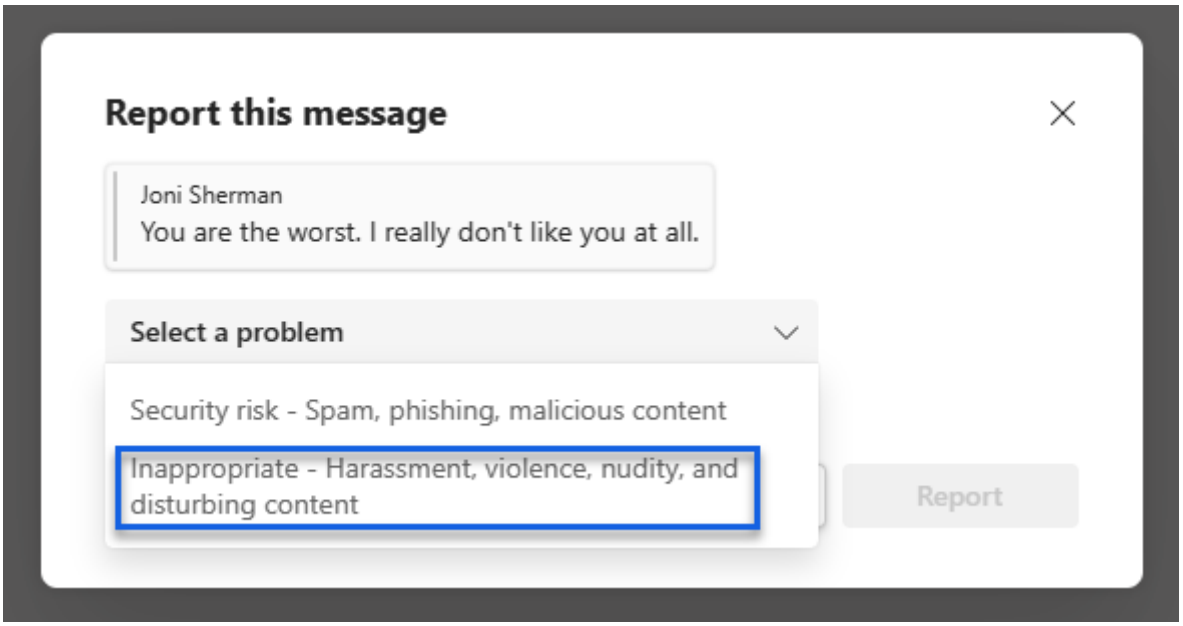
By default, Teams messaging policies in the Teams admin center enable the *Report inappropriate content* option for Teams personal and group chat messages. This option lets

users in your organization submit inappropriate internal chat messages for review by Communication Compliance reviewers for the policy. For more information about user-reported messages in Communication Compliance, see Communication Compliance policies.

To access the feature, a user selects **More options** (...) > **More actions** > **Report this message** from a Teams chat.



In the next dialog box, the user selects the **Inappropriate - Harassment, violence, nudity, and disturbing content** option from the **Select a problem** list.



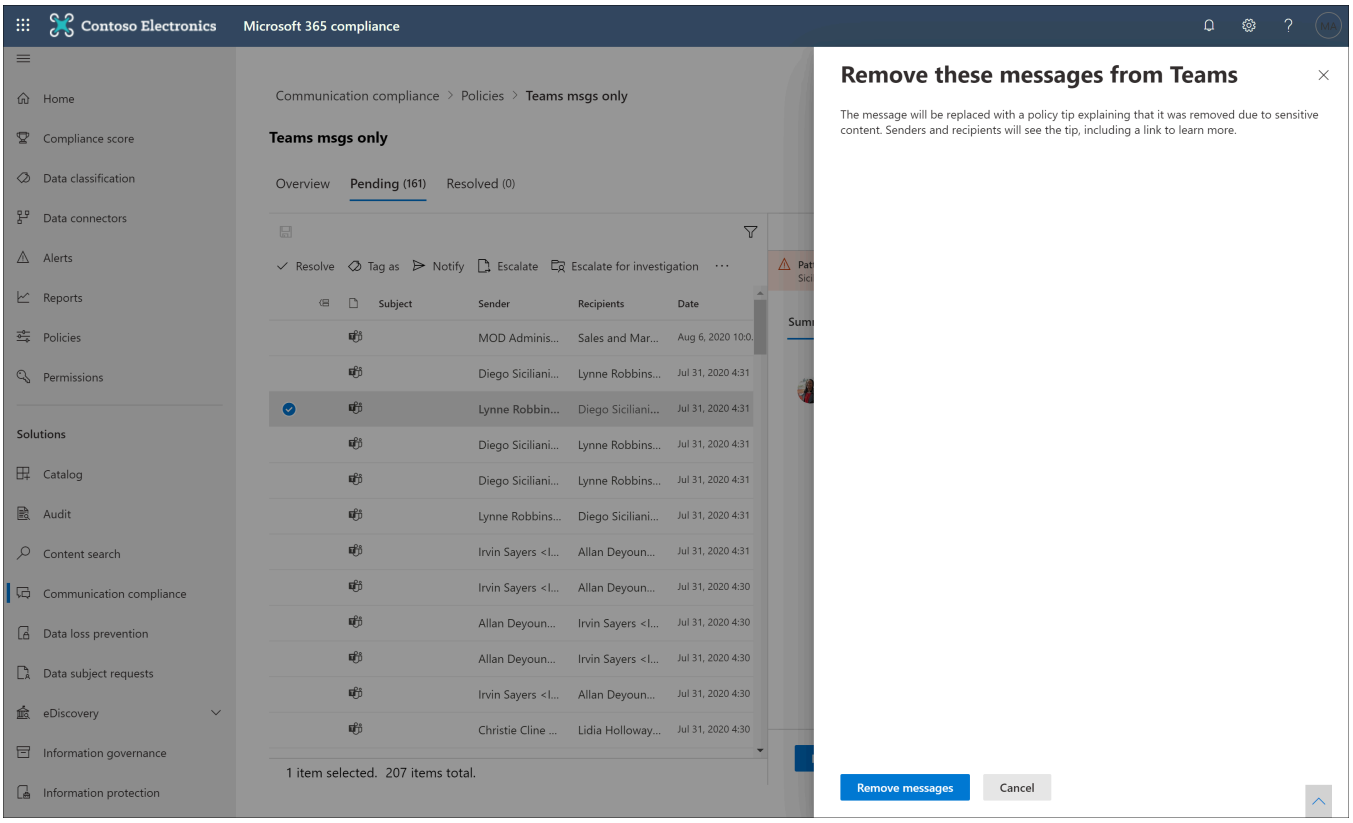
> **Note**
>
> Microsoft Defender for Office 365 manages the other choice in the list (**Security risk- Spam, phishing, malicious content**), if available. Depending on which policy options are turned on in the Microsoft Teams admin center, the user might also be presented with just the **Inappropriate - Harassment, violence, nudity, and disturbing content** option. Learn more about the Microsoft Defender for Office setting

After submitting the message for review, the user receives a confirmation of the submittal in Microsoft Teams. Other participants in the chat don't see this notification.

Users in your organization automatically get the global policy unless you create and assign a custom policy. Edit the settings in the global policy or create and assign one or more custom policies to turn on or turn off this feature. For more information, see Manage messaging policies in Teams.

Act on inappropriate messages in Microsoft Teams

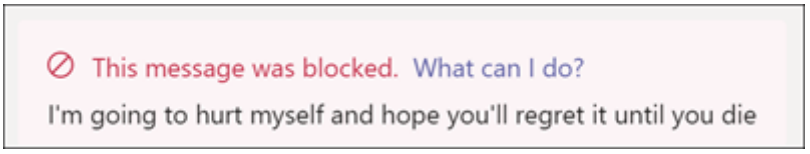
After you configure your policies and receive Communication Compliance alerts for Microsoft Teams messages, compliance reviewers in your organization can act on these messages. This process also includes user-reported messages if your organization enables that option. Reviewers help safeguard your organization by reviewing Communication Compliance alerts and removing flagged messages from view in Microsoft Teams.



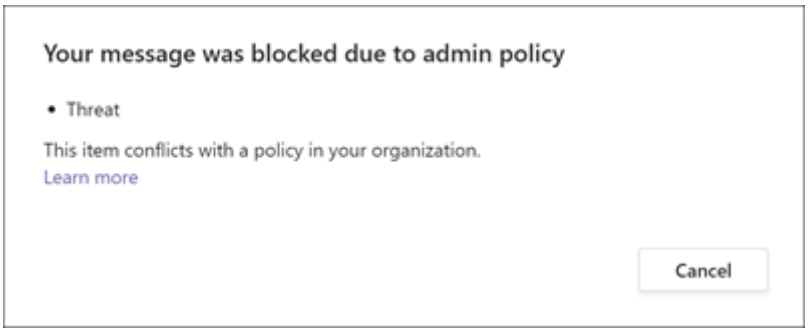
> **Important**
>
> If a user reports a message that was sent before they were added to a chat, Teams message remediation isn't supported; the Teams message can't be removed.

When reviewers remove messages and content, viewers see notifications that explain the message or content was removed and what policy is applicable to the removal. The sender of the removed message or content also receives a removal notification and gets the original message content for context relating to its removal. The sender can view the specific policy condition that applies to the message removal.

Example of policy tip seen by sender:



Example of policy notification seen by the sender:



Example of policy tip seen by recipient:



Use Communication Compliance with SIEM solutions

> **Important**
>
> Microsoft Purview Communication Compliance provides the tools to help organizations detect regulatory compliance (for example, SEC or FINRA) and business conduct violations such as sensitive or confidential information, harassing or threatening language, and sharing of adult content. Communication Compliance is built with privacy by design. Usernames are pseudonymized by default, role-based access controls are built in, investigators are opted in by an admin, and audit logs are in place to help ensure user-level privacy.

Microsoft Purview Communication Compliance is an insider risk solution that helps you minimize communication risks by helping you detect, capture, and act on potentially inappropriate messages in your organization. Security information and event management (SIEM) solutions such as Microsoft Sentinel or Splunk are commonly used to aggregate and track threats within an organization.

A common need for organizations is to integrate Communication Compliance alerts and their SIEM solutions. With this integration, organizations can view Communication Compliance alerts in their SIEM solution and then remediate alerts within the Communication Compliance workflow and user experience.

For example, an employee sends an offensive message to another employee and the offensive message is detected by a Communication Compliance policy for potentially inappropriate content. Communication Compliance solution logs events such as policy matches for inappropriate content in Microsoft 365 Audit (also known as "unified audit log") and imports them into the SIEM solution. You can associate SIEM alerts that were created from Microsoft 365 Audit events with the corresponding Communication Compliance alerts. Investigators receive notifications of these alerts in their SIEM solution, and they can investigate and remediate the corresponding alerts in the Communication Compliance dashboard.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Communication Compliance alerts in Microsoft 365 Audit

Microsoft 365 Audit captures all Communication Compliance policy matches. The following examples show the details available for selected Communication Compliance policy match activities:

Example of audit log entry for an Inappropriate Content policy template match:

XML


```
RunspaceId: 5c7bc9b0-7672-4091-a112-0635bd5f7732
RecordType: ComplianceSupervisionExchange
CreationDate: 7/7/2022 5:30:11 AM
UserIds: user1@contoso.onmicrosoft.com
Operations: SupervisionRuleMatch
AuditData: {"CreationTime":"2022-07-07T05:30:11","Id":"44e98a7e-57fd-4f89-79b8-
08d941084a35","Operation":"SupervisionRuleMatch","OrganizationId":"338397e6\-697e-
4dbe-a66b-2ea3497ef15c","RecordType":68,"ResultStatus":"
{\\"ItemClass\\":\\"IPM.Note\\",\\"CcsiResults\\":\\"\\"}","UserKey":"SupervisionStore
DeliveryAgent","UserType":0,"Version":1,"Workload":"Exchange","ObjectId":"\
<HE1P190MB04600526C0524C75E5750C5AC61A9@HE1P190MB0460.EURP190.PROD.OUTLOOK.COM\>","Use
rId":"user1@contoso.onmicrosoft.com","IsPolicyHit":true,"SRPolicyMatchDetails":
{"SRPolicyId":"00aa00aa-bb11-cc22-dd33-44ee44ee44ee","SRPolicyName":"Adult
images","SRRuleMatchDetails":\[\]}}
ResultIndex: 24
ResultCount: 48
Identity: 44e98a7e-57fd-4f89-79b8-08d941084a35
IsValid: True
ObjectState: Unchanged
```


Example of a Microsoft 365 audit log entry for a policy with custom keyword match (custom sensitive information type):

XML


```
RunspaceId: 5c7bc9b0-7672-4091-a112-0635bd5f7732
RecordType: ComplianceSupervisionExchange
CreationDate: 7/6/2022 9:50:12 PM
UserIds: user2@contoso.onmicrosoft.com
Operations: SupervisionRuleMatch
AuditData: {"CreationTime":"2022-07-06T21:50:12","Id":"5c61aae5-26fc-4c8e-0791-
08d940c8086f","Operation":"SupervisionRuleMatch","OrganizationId":"338397e6\-697e-
4dbe-a66b-2ea3497ef15c","RecordType":68,"ResultStatus":"
{\\"ItemClass\\":\\"IPM.Note\\",\\"CcsiResults\\":\\"public\\"}","UserKey":"Supervisio
nStoreDeliveryAgent","UserType":0,"Version":1,"Workload":"Exchange","ObjectId":"\
<20210706174831.24375086.807067@sailthru.com\>","UserId":"user2@contoso.onmicrosoft.co
m","IsPolicyHit":true,"SRPolicyMatchDetails":{"SRPolicyId":"11bb11bb-cc22-dd33-ee44-
55ff55ff55ff","SRPolicyName":"Insiders","SRRuleMatchDetails":\[{"SRCategoryName":"New
insiders lexicon"}\]}}
```

```
ResultIndex: 46
ResultCount: 48
Identity: 5c61aae5-26fc-4c8e-0791-08d940c8086f
IsValid: True
ObjectState: Unchanged
```

> **Note**
>
> Currently, there might be up to a 24-hour delay between the time Microsoft 365 Audit records a policy match and the time you can investigate policy matches in Communication Compliance.

Configure Communication Compliance and Microsoft Sentinel integration

When you use Microsoft Sentinel to aggregate Communication Compliance policy matches, Microsoft Sentinel uses Microsoft 365 Audit as the data source. To integrate Communication Compliance alerts with Microsoft Sentinel, complete the following steps:

1. Onboard to Microsoft Sentinel. As part of the onboarding process, you configure your data sources.

2. Configure the Microsoft Sentinel Microsoft Office 365 data connector and under connector configuration, select *Exchange*.

3. Configure search query to retrieve Communication Compliance alerts. For example:

   | *OfficeActivity* | *where OfficeWorkload == "Exchange" and Operation == "SupervisionRuleMatch"* | *sort by TimeGenerated*

   To filter for a specific user, use the following query format:

   | *OfficeActivity* | *where OfficeWorkload == "Exchange" and Operation == "SupervisionRuleMatch" and UserId == "User1@Contoso.com"* | *sort by TimeGenerated*

For more information about the Microsoft 365 Audit logs for Office 365 that Microsoft Sentinel collects, see Azure Monitor Logs reference.

Configure Communication Compliance and Splunk integration

To integrate Communication Compliance alerts with Splunk, complete the following steps:

1. Install the Splunk Add-on for Microsoft Office 365
2. Configure an integration application in Microsoft Entra ID for the Splunk Add-on for Microsoft Office 365
3. Configure search queries in your Splunk solution. Use the following search example to identify all Communication Compliance alerts:

   index= sourcetype="o365:management:activity" Workload=Exchange Operation=SupervisionRuleMatch*

To filter results for a specific Communication Compliance policy, use the *SRPolicyMatchDetails.SRPolicyName* parameter.

For example, the following search example returns alerts for matches to a Communication Compliance policy named *Inappropriate content*:

index= sourcetype='o365:management:activity' Workload=Exchange Operation=SupervisionRuleMatch SRPolicyMatchDetails.SRPolicyName=<Inappropriate content>*

The following table shows sample search results for different policy types:

Expand table

Policy types	Example search results

Policy detecting a custom sensitive information type keyword list	{
	CreationTime: 2022-09-17T16:29:57
	ID: 4b9ce23d-ee60-4f66-f38d-08d979f8631f
	IsPolicyHit: true
	ObjectId: CY1PR05MB27158B96AF7F3AFE62E1F762CFDD9@CY1PR05MB2715.namprd05.prod.outlook.com
	Operation: SupervisionRuleMatch
	OrganizationId: 22cc22cc-dd33-ee44-ff55-66aa66aa66aa
	RecordType: 68
	ResultStatus: {"ItemClass":"IPM.Note","CcsiResults":"leak"}
	SRPolicyMatchDetails: { [+] }
	UserId: user1@contoso.OnMicrosoft.com
	UserKey: SupervisionStoreDeliveryAgent
	UserType: 0
	Version: 1
	Workload: Exchange
	}

Policy detecting potentially	{ CreationTime: 2022-09-17T23:44:35 ID: e0ef6f54-9a52-4e4c-9584-08d97a351ad0 IsPolicyHit: true

Policy types	Example search results
inappropriate language	ObjectId: BN6PR05MB3571AD9FBB85C4E12C1F66B4CCDD9@BN6PR05MB3571.namprd05.prod.outlook.com Operation: SupervisionRuleMatch OrganizationId: 22cc22cc-dd33-ee44-ff55-66aa66aa66aa RecordType: 68 ResultStatus: {"ItemClass":"IPM.Yammer.Message","CcsiResults":""} SRPolicyMatchDetails: { [+] } UserId: user1@contoso.com UserKey: SupervisionStoreDeliveryAgent UserType: 0 Version: 1 }

Configure Communication Compliance with other SIEM solutions

To retrieve Communication Compliance policy matches from Microsoft 365 Audit, use PowerShell or the Office 365 Management API.

When you use PowerShell, you can use either of these parameters with the **Search-UnifiedAuditLog** cmdlet to filter audit log events for Communication Compliance activities.

Expand table

Audit log parameter	Communication Compliance parameter value
Operations	SupervisionRuleMatch
RecordType	ComplianceSupervisionExchange

The following examples show how to search for Communication Compliance matches.

PowerShell

```
Search-UnifiedAuditLog -StartDate $startDate -EndDate $endDate -Operations
SupervisionRuleMatch | ft CreationDate,UserIds,AuditData
```

The following is a sample search using the **RecordsType** parameter and the *ComplianceSupervisionExchange* value:

PowerShell

```
Search-UnifiedAuditLog -StartDate $startDate -EndDate $endDate -RecordType
ComplianceSuperVisionExchange | ft CreationDate,UserIds,AuditData
```

Additional SIEM integration resources

- Communication Compliance auditing
- Microsoft Purview Audit (Premium)
- Office 365 Management Activity API reference

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

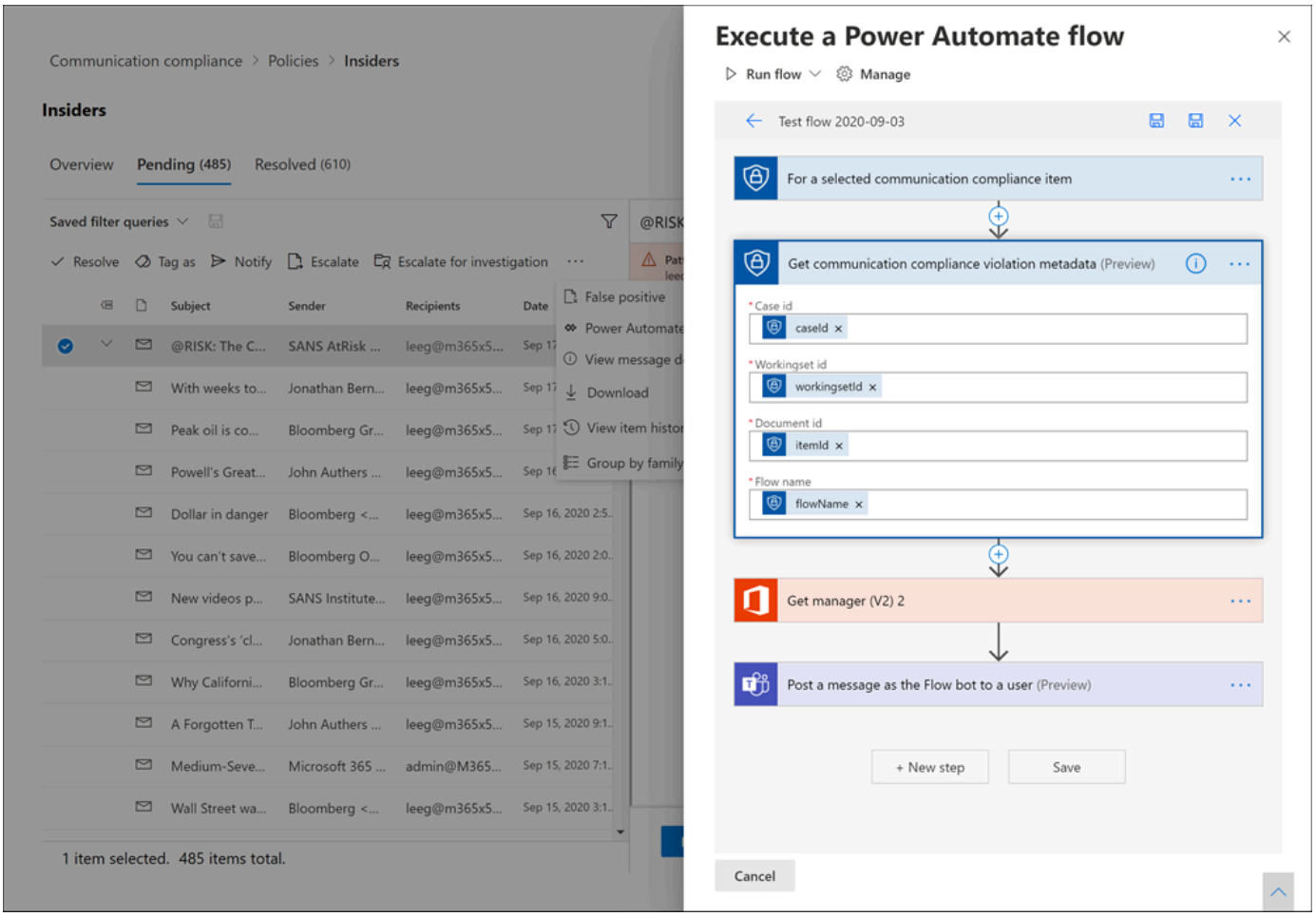
Last updated on 06/26/2026

Use Microsoft Power Automate with Communication Compliance

Microsoft Power Automate is a workflow service that automates actions across applications and services. By using flows from templates or creating them manually, you can automate common tasks associated with these applications and services. When you enable Power Automate flows for Communication Compliance, you can automate important tasks for alerts and users. You can configure Power Automate flows to notify managers when users have Communication Compliance alerts and other applications.

Customers with Microsoft 365 subscriptions that include Communication Compliance don't need extra Power Automate licenses to use the recommended default Communication Compliance Power Automate template. Customize the default template to support your organization and cover core Communication Compliance scenarios. If you choose to use premium Power Automate features in these templates, create a custom template by using the Microsoft Purview connector, or use Power Automate templates for other compliance areas in Microsoft Purview, you might need extra Power Automate licenses.

> **Important**
>
> Are you receiving requests for extra license validation when testing Power Automate flows? Your organization might not have received service updates for this preview feature yet. We're deploying updates, and all organizations with Microsoft 365 subscriptions that include Communication Compliance should have license support for flows created from the recommended Power Automate templates before October 30, 2020.



The following Power Automate template is provided to customers to support process automation for Communication Compliance alerts:

- **Notify manager when a user has a Communication Compliance alert**: Some organizations might need immediate management notification when a user has a Communication Compliance alert. When you configure and select this flow, the manager for the case user receives an email message with the following information about all alerts:
  - Applicable policy for the alert
  - Date and time of the alert
  - Severity level of the alert

Automate tasks for alerts and users with Microsoft Power Automate

To create a Power Automate flow from a recommended default template, use the **Manage Power Automate flows** option from **Automate** when working directly in an alert. To create a Power Automate flow with **Manage Power Automate flows**, you must be a member of at least one Communication Compliance role group.

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.
2. Go to the **Communication Compliance** solution.
3. Select **Policies** in the left navigation, then select the policy that includes the policy matches that you want to review.
4. Select the **Pending** tab, and select a policy match.
5. Select **Power Automate** from the alert action menu.
6. On the **Power Automate** page, select a default template from the **Communication Compliance templates you may like** section on the page.
7. Review the embedded connections needed for the flow and check the connection statuses. If needed, update any connections that aren't displayed as available. Select **Continue**.
8. By default, the recommended flows are preconfigured with the recommended Communication Compliance and Microsoft 365 service data fields required to complete the assigned task for the flow. If needed, customize the flow components by using **Show advanced options** and configuring the available properties for the flow component.
9. If needed, add any additional steps to the flow by selecting **New step**. In most cases, this change shouldn't be needed for the recommended default templates.
10. Select **Save draft** to save the flow for further configuration later, or select **Save** to complete the configuration for the flow.
11. Select **Close** to return to the **Power Automate flows** page. The new template is listed as a flow on the **My flows** tab and is automatically available for the user that created the flow when working with Communication Compliance alerts.

Share a Power Automate flow

By default, only the creator can access a Power Automate flow. To give other Communication Compliance users access to a flow, the creator must share it. Use **Power Automate** to share a flow when working directly in an alert.

To share a Power Automate flow, you must be a member of at least one Communication Compliance role group.

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.
2. Go to the **Communication Compliance** solution.
3. Select **Policies** in the left navigation, then select the policy that contains the policy matches you want to review.
4. Select the **Pending** tab, then select a pending policy match.

5. Select **Power Automate** from the alert action menu.
6. On the **Power Automate flows** page, select the **My flows** or **Team flows** tab.
7. Select the flow to share, then select **Share** from the **Flow Options** menu.
8. On the flow sharing page, enter the name of the user or group you want to add as an owner for the flow.
9. On the **Connection Used** dialog, select **OK** to acknowledge that the added user or group has full access to the flow.

Edit a Power Automate flow

Use **Power Automate** to edit a flow when working directly in an alert. To edit a Power Automate flow, you must be a member of at least one Communication Compliance role group.

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.
2. Go to the **Communication Compliance** solution.
3. Select **Policies** in the left navigation, then select the policy that contains the policy matches you want to review.
4. Select the **Pending** tab, then select a pending policy match.
5. Select **Power Automate** from the alert action menu.
6. On the **Power Automate flows** page, select the flow to edit. Select **Edit** from the flow control menu.
7. Select **ellipsis** > **Settings** to change a flow component setting or **ellipsis** > **Delete** to delete a flow component.
8. Select **Save**, then select **Close** when finished editing the flow.

Delete a Power Automate flow

If you need to delete a flow, use **Power Automate** when working directly in an alert. To delete a Power Automate flow, you must be a member of at least one Communication Compliance role group.

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.
2. Go to the **Communication Compliance** solution.
3. Select **Policies** in the left navigation, then select the policy that contains the policy matches you want to review.
4. Select the **Pending** tab, then select a pending policy match.

5. Select **Power Automate** from the alert action menu.
6. On the **Power Automate flows** page, select the flow that you want to delete. Select **Delete** from the flow control menu.
7. On the deletion confirmation dialog, select **Delete** to remove the flow or select **Cancel**.

Last updated on 06/25/2026

Use notification templates in Communication Compliance

08/20/2025

Create notification templates when you want to send users an email reminder for policy matches as part of the issue resolution process. You can only send notifications to the user email address associated with the policy match that generates the specific alert for remediation. When selecting a notification template to apply to a policy violation as part of the remediation workflow, you can choose to accept the field values defined in the template or overwrite the fields as needed.

Notice templates are custom email templates where you define the following message fields in the **Communication Compliance settings** area:

Expand table

Field	Required	Details
Template name	Yes	Friendly name for the notice template that you select in the notify workflow during remediation, supports text characters.
Sender address	Yes	Address of one or more users or groups that send the message to the user with a policy match, selected from the Active Directory for your subscription.
CC and BCC addresses	No	Optional users or groups to be notified of the policy match, selected from the Active Directory for your subscription.
Subject	Yes	Information that appears in the subject line of the message, supports text characters.
Message body	Yes	Information that appears in the message body, supports text or HTML values.

HTML for notices

If you want to create more than a simple text-based email message for notifications, you can create a more detailed message by using HTML in the message body field of a notice template. The following example provides the message body format for a basic HTML-based email notification template:

```
HTML
```

```
<!DOCTYPE html>
<html>
    <body>
        <h2>Action Required: Contoso Employee Code of Conduct Policy Training</h2>
        <p>A recent message you've sent has generated a policy alert for the
Contoso Employee <a href='https://www.contoso.com'>Code of Conduct Policy</a>.</p>
        <p>You are required to attend the Contoso Employee Code of Conduct <a
href='https://www.contoso.com'>training</a> within the next 14 days. Please
contact <a href='mailto:hr@contoso.com'>Human Resources</a> with any questions
about this training request.</p>
        <p>Thank you,</p>
        <p><em>Human Resources</em></p>
    </body>
</html>
```

> **Note**
>
> In Communication Compliance notification templates, the HTML href attribute implementation currently supports only single quotation marks instead of double quotation marks for URL references.

Investigate and remediate Communication Compliance alerts

> **Important**
>
> Microsoft Purview Communication Compliance provides the tools to help organizations detect regulatory compliance (for example, SEC or FINRA) and business conduct violations such as sensitive or confidential information, harassing or threatening language, and sharing of adult content. Communication Compliance is built with privacy by design. Usernames are pseudonymized by default, role-based access controls are built in, investigators are opted in by an admin, and audit logs are in place to help ensure user-level privacy.

After you configure Communication Compliance policies, you receive alerts for message issues that match your policy conditions. To view and act on alerts, assign users the following permissions:

- The *Communication Compliance Analysts* or the *Communication Compliance Investigators* role group
- Reviewer in the policy that is associated with the alert

After you establish required permissions, use the following working instructions to investigate and remediate issues.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Investigate policy matches and alerts

> **Important**
>
> On June 1, 2025, the new **Policy Match Preservation** setting goes into effect. On that date, existing policy matches might be permanently deleted based on the time period you select in the **Policy Match Preservation** setting. New policy matches are preserved for that time

period. You have until **May 19** to change the time period if you don't want to use the default setting of **1 year**. For details, see Policy Match Preservation setting.

To investigate issues detected by your policies, review policy matches and alerts. The Communication Compliance area provides several features to help you quickly investigate policy matches and alerts:

Policies page

When you sign in to the Microsoft Purview portal with an admin account in your Microsoft 365 organization, select the **Communication Compliance** solution, then select the **Policies** page. This page shows Communication Compliance policies configured for your Microsoft 365 organization and links to recommended policy templates.

> **Note**
>
> If your role group is scoped by one or more admin units, you see a message at the top of the page that you can only view and manage the policies that you're scoped for. To learn more about what you can access, select **View role groups** in the banner.

Each policy listed includes the following columns:

- **Messages scanned today**: The total number of messages that the policy scanned on the current day for users and locations that are in scope of the policy. If there's a policy match, the item is added to the **New pending today** column. The number refreshes automatically once per hour, it doesn't update if you refresh the page. At the end of the UTC day, the count resets to zero.

  > **Tip**
  >
  > If the value in the column is 0 or a low number, it might be because the policy is too strict. For example, the policy might be focused on just one user or on just one location that the in-scope user doesn't use. You might see a wide variety of aggregate numbers in this column if your policies have different focuses. For example, one policy might be focused on an entire department of users sending messages from multiple locations while another policy might be focused on just one user or just one location.

- **New pending today**: Shows the number of policy matches for the current day. This value updates whenever you open the page. You can also select the **Refresh** button to get the latest count. At the end of the UTC day, the count resets to zero.
- **Total pending**: The count of policy matches that need review. This value updates whenever you open the page. You can select the **Refresh** button to get the latest count. After refreshing, this number matches the number on the **Pending** tab.
- **Total resolved**: The total number of resolved policy matches. This value updates whenever you open the page. You can select the **Refresh** button to get the latest count. After refreshing, this number matches the number on the **Resolved** tab.
- **Status**: The status of the policy (Active or Deactivated).
- **Last modified**: The date and Coordinated Universal Time (UTC) of the last policy modification.
- **Last policy scan**: The UTC date for the last policy scan.

To start remediation actions, select the policy associated with the alert to launch the **Policy details** page. From the **Policy details** page, you can review a summary of the activities, review, and act on policy matches on the **Pending** tab, summarize a lengthy message by using Microsoft Copilot in Microsoft Purview, or review the history of closed policy matches on the **Resolved** tab. For details, see Review and remediate policy matches and alerts.

Alerts page

Go to **Communication Compliance** > **Alerts** to display the last 30 days of alerts grouped by policy matches. This view lets you quickly see which Communication Compliance policies generate the most alerts, ordered by severity. An alert isn't the same thing as a policy match. An alert generally consists of multiple policy matches, not just one policy match. After the required number of policy matches is met for a particular alert, the alert is created and email is sent to the alert recipient.

- **Reports page**: Go to **Communication Compliance** > **Reports** to display Communication Compliance report widgets. Each widget provides an overview of Communication Compliance activities and statuses, including access to deeper insights about policy matches and remediation actions.

> **Tip**
>
> Analyze interactions entered into generative AI applications.

Tips for quickly reviewing policy matches on the Pending or Resolved tab

- When you select a message to review on the **Pending** tab or the **Resolved** tab, the condition that caused the policy match appears in an alert message bar (yellow banner) at the top of the **Source** tab. This alert message bar is a quick way to determine the condition or conditions that caused the policy match. If there are multiple conditions, select **View all** in the banner to see all the conditions that caused the policy match. At this time, only Microsoft provided trainable classifiers and sensitive information types are highlighted as conditions in the yellow banner.

- Sometimes it's useful to quickly review policy settings without opening a policy. For example, if you're testing multiple policies with different conditions, you might want to save time by reviewing conditions for each policy to determine risk before opening the policy. You can do this by selecting **Policy settings**, which opens a panel where you can view the policy settings. If you're a member of the Communication Compliance or Communication Compliance Admins role group, you can view and change settings from the panel. If you're a member of the Communication Compliance Investigators or Communication Compliance Analysts role group, you can view settings but you can't change them.

Policy Match Preservation setting

When a Communication Compliance policy finds a message that matches the policy, the solution stores a *copy* of the message (not the original message). Starting June 1, 2025, you can use the **Policy Match Preservation** setting to specify how long to save policy matches in Communication Compliance. You can choose **1 month**, **6 months**, **1 year**, or **7 years**. The default value is **1 year**.

> ⓘ **Note**
>
> You can also specify a preservation period when you create or edit a policy. The preservation period you select within a policy takes precedence over the global **Policy Match Preservation** setting.

When the global **Policy Match Preservation** setting goes into effect on April 1, 2025:

- The value in the **Policy Match Preservation** setting applies to all existing policy matches. The system automatically deletes policy matches that are older than the selected time period. For example, if the **Policy Match Preservation** setting is **1 year** (the default), the

system permanently deletes any policy matches older than 1 year as of September 1. Only the policy matches in Communication Compliance are deleted because they're copies of the original messages. This process doesn't delete original copies from user shards. You can use Microsoft Purview Data Lifecycle Management to set up a retention policy for original messages.

- New policy matches as of September 1 receive a timestamp with the preservation period specified in the setting.
- If you change the time period in the setting after September 1, the preservation period applies going forward. For example, if you select **1 year** as the time period on September 1, then change the setting to **7 years** on September 3, policy matches from September 1 and September 2 are preserved for 1 year, and policy matches starting on September 3 are preserved for 7 years.

What you need to do

The **Policy Match Preservation** setting becomes available for your organization on April 1, 2025. Between April 1 and May 30, select the time period for preserving policy matches for your organization. The selection isn't applied until June 1, 2025.

Change the time period for the global Policy Match Preservation setting

To change the global policy match preservation period, complete the following steps:

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.
2. Select the **Settings** button in the upper-right corner of the page, then select **Communication Compliance** to go to the Communication Compliance global settings.
3. Select the **Policy Match Preservation** setting, then select a new time period.
4. Select **Confirm** in the confirmation dialog box.

Change the policy match preservation time period for a custom policy

You can also select a time period for preserving policy matches in the policy workflow for a custom policy. When you set a time period in a custom policy, it takes precedence over the selection in the global **Policy Match Preservation** setting. This option is useful for policies that are the exception to the norm.

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.
2. Go to the **Communication Compliance** solution, then select **Policies** in the left navigation.

3. Select the **More actions** (ellipsis) button in the row for the policy you want to change, then select **Edit**.
4. On the **Name and describe your policy** page in the policy workflow, under **Preserve policy matches**, make a selection. If you leave the **Global Setting** selection, the policy uses the time period selected in the global **Policy Match Preservation** setting. If you choose any other time period, it takes precedence over the selection in the global setting.

Using filters

Sort the messages so it's easier for you to investigate. From the **Policy details** page, Communication Compliance supports multilevel filtering for several message fields to help you quickly investigate and review messages with policy matches. You can filter pending and resolved items for each configured policy. You can configure filter queries for a policy or configure and save custom and default filter queries for use in each specific policy. After configuring fields for a filter, you see the filter fields displayed on the top of the message queue that you can configure for specific filter values.

Key filters (the **Body/Subject**, **Date**, **Sender**, and **Tags** filters) always display on the **Pending** and **Resolved** tabs to make it easy to access those filters.

For the **Date** filter, the date and time for events are listed in Coordinated Universal Time (UTC). When filtering messages for views, the requesting user's local date/time determines the results based on the conversion of the user's local date/time to UTC. For example, if a user in U.S. Pacific Daylight Time (PDT) filters a report from 8/30/2021 to 8/31/2021 at 00:00, the report includes messages from 8/30/2021 07:00 UTC to 8/31/2021 07:00 UTC. If the same user was in U.S. Eastern Daylight Time (EDT) when filtering at 00:00, the report includes messages from 8/30/2021 04:00 UTC to 8/31/2021 04:00 UTC.

Filter details

Communication Compliance filters help you filter and sort messages for quicker investigation and remediation actions. You can filter on the **Pending** and **Resolved** tabs for each policy. To save a filter or filter set as a saved filter query, you must configure one or more values as filter selections.

The following table outlines filter details:

Expand table

Filter	Details
Body/Subject	The message body or subject. Use this filter to search for keywords or a keyword phrase in the body or subject of the message. The subject appears in the **Subject** column for email messages. For Teams messages, nothing appears in the **Subject** column.
Classifiers	The name of built-in and custom classifiers that apply to the message. Some examples include *Targeted Harassment*, *Profanity*, *Threat*, and more.
Date	The date the message was sent or received by a user in your organization. To filter for a single day, select a date range that starts with the day you want results for and end with the following day. For example, to filter results for 9/20/2020, choose a filter date range of 9/20/2020-9/21/2020.
Escalated To	The user name of the person included as part of a message escalation action.
File class	The class of the message based on the message type, either *message* or *attachment*.
Has attachment	The attachment presence in the message.
Item class	The source of the message based on the message type, email, Microsoft Teams chat, Bloomberg, and more. For more information, see Item Types and Message Classes.
Language	The detected language of text in the message. The message is classified according to the language of most the message text. For example, for a message containing both German and Italian text, but most text is German, the message is classified as German (DE). For a list of supported languages, see Trainable classifiers. You can also filter by more than one language. For example, to filter messages classified as German and Italian, enter 'DE,IT' (the two-digit language codes) in the Language filter search box. To view the detected language classification for a message, select a message, select View message details, and scroll to the *EmailDetectedLanguage* field.
Recipient domains	The domain to which the message was sent; this domain is typically your Microsoft 365 subscription domain by default.
Recipient	The user to which the message was sent. **Note**: The **Recipient** field includes recipients in the **To** and **CC** fields. **BCC** fields aren't supported.
Sender domain	The domain that sent the message.
Sender	The person who sent the message.
Size	The size of the message in KB.
Tags	The tags assigned to a message, either *Questionable*, *Compliant*, or *Noncompliant*.

Configure a filter

To configure a filter for policy matches, complete the following steps:

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.
2. Go to the **Communication Compliance** solution.
3. Select **Policies** in the left navigation, then select a policy to see policy matches (if any) for that policy.
4. On the **Policy** page, select either the **Pending** or **Resolved** tab to display the items for filtering.
5. Select **Filters**.
6. Select one or more filter checkboxes, then select **Apply**.
7. To save the selected filters as a filter query, select **Save the query** after you configure at least one filter value. Enter a name for the filter query, then select **Save**. This filter is available to use for only this policy and is listed in the **Saved filter queries** section of the **Filters** page.

Review and remediate policy matches and alerts

After you review policy matches or alerts and configure filtering, take remediation action. Start your remediation process with the following workflow on the **Policy** or **Alerts** pages.

> **Note**
>
> If you see a policy prefaced with "DSPM for AI" (or "AI hub" from the preview name), the policy was created in Data Security Posture Management for AI, not in Communication Compliance. For more information about this solution, see Data Security Posture Management (DSPM) for AI.

Examine the message basics

Sometimes it's obvious from the source or subject that you can immediately remediate a message. The message might be spurious or incorrectly matched to a policy and you should resolve it as misclassified. To mark a message as misclassified, select **Resolve** and **Item was misclassified** to leave a note in the item's history that it's misclassified and remove it from the **Pending** queue. To share misclassified content with Microsoft, select **Send message content, attachments, and subject** to remove it from the **Pending** queue and report the items to Microsoft.

> **Note**
>
> Misclassified options in the **Resolve** pane don't appear if the condition that caused the policy match isn't based on a trainable classifier. Reporting misclassified items to Microsoft isn't supported for Content Safety classifiers yet.

From the source or sender information, you might already know how the message should be routed or handled in these circumstances. Consider using **Tag as** or **Escalate** to assign a tag to applicable messages or to send messages to a designated reviewer.

Examine the message details

After reviewing the message basics, open the message to examine the details and determine further remediation actions.

> **Tip**
>
> For lengthy messages, you can save time by summarizing the message (including attachments, transcripts, and recordings) with Copilot in Microsoft Purview.

Select a message to view the complete message header and body information. Several different options and views are available to help you decide the proper course of action:

- **Sentiment**: Messages include a sentiment evaluation (powered by Azure Cognitive Service for Language) to help investigators quickly prioritize potentially riskier messages to address first. The **Sentiment** column displays the message sentiment and is enabled in the default view. Messages are flagged with one of the following values:

Expand table

Value	Description
Positive	Messages with *Positive* sentiment indicate a lower priority for triaging.
Negative	Messages with *Negative* sentiment indicate messages to prioritize.
Neutral	Messages with *Neutral* sentiment are determined to be neither positive nor negative.
Not available	*Not available* appears in the column for file formats that aren't supported. For example, sentiment analysis isn't available for any image that's attached to a Teams or email

Value	Description
	message, text extracted from an OCR image, or text or recordings from a Teams transcript. *Not available* also appears if the message includes more than 5,120 words.
Scanning	*Scanning* appears in the the column when Communication Compliance is trying to determine the appropriate sentiment value for the message.

- **Attachments**: This option allows you to examine Modern attachments that match policy conditions. Modern attachments content is extracted as text and is viewable on the **Pending** tab. For more information, see the Communication Compliance feature reference.

- **Source**: This view is the standard message view commonly seen in most web-based messaging platforms. The header information is formatted in the normal style and the message body supports embedded graphic files and word-wrapped text. If optical character recognition (OCR) is enabled for the policy, images containing printed or handwritten text that match policy conditional are viewed as a child item for the associated message in this view.

- **Plain text**: Text view that displays a line-numbered text-only view of the message and includes keyword highlighting in messages and attachments for sensitive info type terms, terms identified by built-in classifiers assigned to a policy, or for terms included in a dedicated keyword dictionary assigned to a policy. Keyword highlighting, which is currently available for English language only, can help direct you to the area of interest in long messages and attachments. In some cases, highlighted text might be only in attachments for messages matching policy conditions. Embedded files aren't displayed and the line numbering in this view is helpful for referencing pertinent details among multiple reviewers.

- **Conversation**: This view, which is available for Teams chat messages, displays up to five messages before and after a message to help reviewers view the activity in the conversational context. Select **Load more** to load up to 20 messages before and after a message. To download messages, select **Download conversation**. This action downloads an image file of everything you see in the user interface and also a .csv file of all the message metadata (UserId, UserName, and so on).

  The Conversation view context helps reviewers quickly evaluate messages and make more informed message resolution decisions. Real-time message additions to conversations are displayed, including all inline images, emojis, and stickers available in Teams. Image or text file attachments to messages aren't displayed. Notifications are automatically displayed for messages that have been edited or for messages deleted from the Conversation window.

When a message is resolved, the associated conversational messages aren't retained with the resolved message.

- **Pattern detected** notification: Many harassing and bullying actions over time involve recurring instances of the same behavior by a user. The *Pattern detected* notification is displayed in the message details and raises attention to the message. Detection of patterns is on a per-policy basis and evaluates behavior over the last 30 days when at least two messages are sent to the same recipient by a sender. Investigators and reviewers can use this notification to identify repeated behavior to evaluate the message as appropriate.

- **Translation**: This view automatically converts message text to the language configured in the *Displayed language* setting in the Microsoft 365 subscription for each reviewer. This conversion includes the text for the policy match and everything included in the conversation view (up to five messages before and five messages after the policy match). The *Translation* view helps broaden investigative support for organizations with multilingual users and eliminates the need for additional translation services outside of the Communication Compliance review process. Using Microsoft translation services, Communication Compliance automatically detects if the text is in a different language than the user's current system setting and displays alert message text accordingly. For a complete list of supported languages, see Microsoft Translator Languages. Languages listed in the *Translator Language List* are supported in the *Translation* view.

- **User activity**: This view provides risk profile, policy matches, and user activities captured by Insider Risk Management and Communication Compliance. This integration helps Communication Compliance investigators quickly see risk severity and the associated activities for the user while investigating and triaging pending policy matches.

  The **User activity** view displays the *Insider risk severity* level for the user and has two sections, one for Communication Compliance policy matches and activities and one for Insider Risk Management risk activities. View the risk severity of the user from Insider Risk Management in the *Source* tab if data sharing is enabled. For more information, see Share Insider Risk Management data with other solutions.

  Communication Compliance policy matches: This section displays the total number of communication *Policy matches* and the total number of *Remediation actions* for the user included in this policy match.

  Select **View details** in this section to display a timeline of Communication Compliance activities for the user. This timeline includes:
  - Totals for *Policy matches* and *Remediation actions* for the user

- Details for each activity associated with the user for the past 30 days.

Insider risk activity: This section displays the total number of *Risk activities* and the total number of *Unusual activities* for activities associated with this user for Insider Risk Management policies. *Unusual activities* are activities for the user that are considered potentially risky, as they're unusual and a departure from their typical activities.

Select **View details** in this section to display a timeline of insider risk activity for the user. This timeline includes:

- The insider risk severity level.
- Totals for *All activities* and *Unusual activities*.
- Details for each activity associated with the user from the Activity explorer in Insider Risk Management.

To view the insider risk activities in insider risk management, select **Open in insider risk management**.

Timeline: This section displays a history of all communications for user activities. This timeline allows investigators to review all communications for a user, helping to identify any inappropriate handling of confidential information. Only previous policy matches for the user associated with the alert are displayed.

For each Copilot interaction, the details for the user prompt and the Copilot response are available for review. The date and time for each interaction are provided in UTC.

Summarize a message by using Copilot in Microsoft Purview

You can use Copilot in Microsoft Purview to get a contextual summary of a Teams, email, or Viva Engage message included in a policy match. The summary is in the context of one or more Microsoft provided trainable classifiers that flag the message (for example, stock manipulation). This feature saves time for investigators if the message content is lengthy.

Copilot in Microsoft Purview summarizes the entire message, including any video recordings, meetings transcripts, or attachments (docx, pdf, or txt files).

Limitations

Be aware of the following limitations when using Copilot in Microsoft Purview to summarize messages:

- The message content must be at least 100 words and less than 15,000 words. If the message content is fewer than 100 words or more than 15,000 words, Copilot in Microsoft Purview doesn't summarize it.
- Summarize items in various languages supported by Security Copilot. For more information, see Supported languages in Microsoft Security Copilot.
- At this time, only individual Teams messages can be summarized. If you summarize an individual Teams message, Copilot doesn't pull in the surrounding conversation.

Prerequisites for using Copilot in Microsoft Purview

To summarize a message in Communication Compliance by using Copilot in Microsoft Purview, you must have specific licenses and onboard your organization to Copilot in Microsoft Purview. For details, see Copilot in Microsoft Purview licensing requirements and onboarding. You must also be a member of the *Communication Compliance*, *Communication Compliance Analysts*, or *Communication Compliance Investigators* role.

Summarize a message

1. In Communication Compliance, go to the **Policies** page, and then open any policy to view policy matches for that policy.

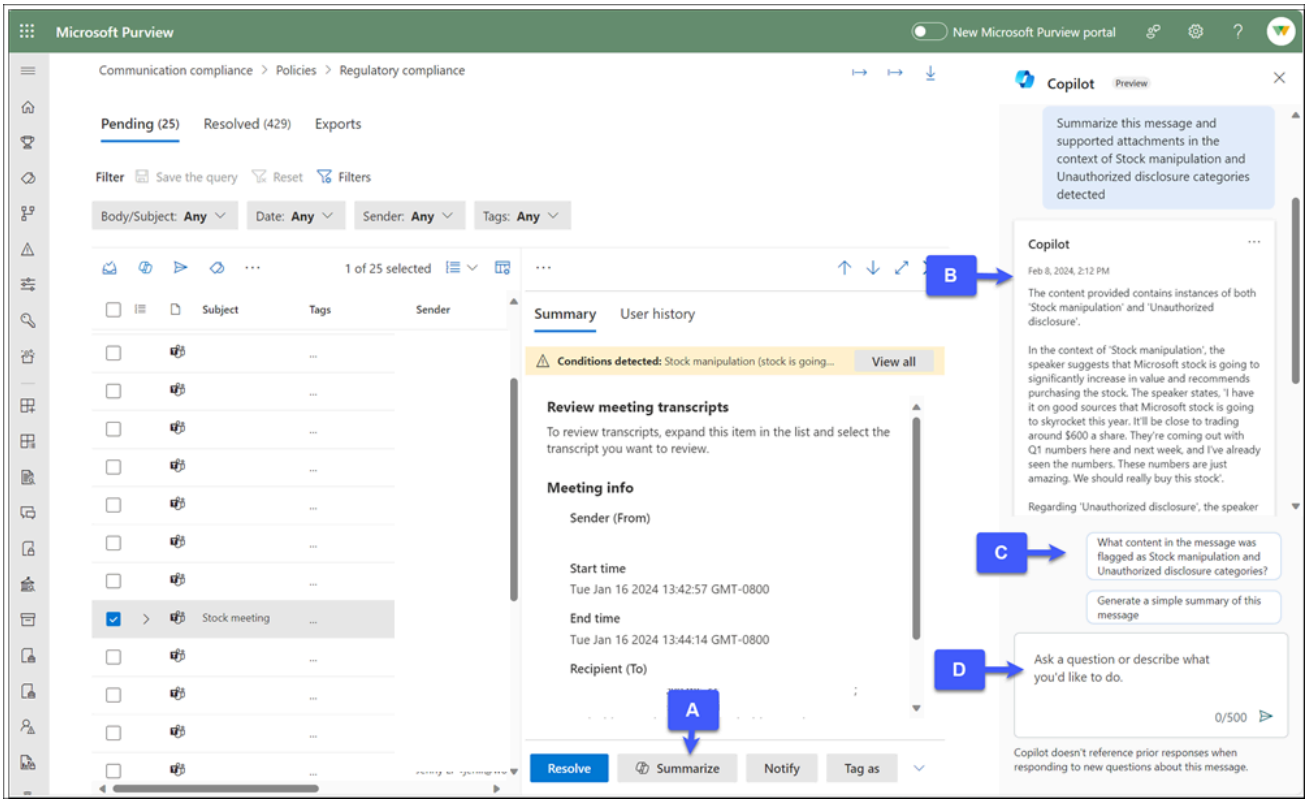
2. Select a message in the list. You can select a parent item or a child item (but it's easier to select the parent item).

3. Select **Summarize**. This option appears below the message details. It opens the **Copilot** panel on the right side of the screen and displays a summary of the message.

   > **ⓘ Important**
   >
   > You might see a generic error message that says "An error occurred" after selecting **Summarize**. This error message can appear for any of the following reasons:
   >
   > - You don't have the required license for Copilot in Microsoft Purview.
   > - You don't have the required Communication Compliance role and/or the Security Copilot contributor role. To use Copilot in Microsoft Purview, you must have one of the following Communication Compliance roles: *Communication Compliance*, *Communication Compliance Analysts*, *Communication Compliance Investigators*. You must also have the Security Copilot contributor role, which should be turned on by default for all users in a Microsoft Entra organization.

- There's an internal error. Microsoft provides more information for these errors.

4. Select one of the suggested prompts at the bottom of the panel to use Copilot to create the summary. For example, if the classifier that flagged a message is Stock manipulation, select *Summarize this message and supported attachments in the context of Stock manipulation category detected*. You can also enter an open-ended question in the "Ask a question" box.



a. **Summarize**.
b. Summary generated by **Summarize**.
c. Suggested Copilot prompts.
d. "Ask a question" box for open-ended Copilot questions.

> **Note**
>
> Only message content is summarized. If you ask a question relevant to the message content, Copilot provides a result. If the question isn't relevant to message content, Copilot instructs you to ask a different question.

5. If you want to provide feedback on the summarized content, select the drop-down arrow in the lower-right corner below the summary, and then enter your feedback.

6. After reviewing the summary, you can remediate the policy match like any other policy match. For remediation options, see Decide on a remediation action.

> ⓘ Note
>
> Copilot in Microsoft Purview responses are also recorded in Microsoft Security Copilot.

Understanding hidden matches

When a Communication Compliance policy triggers, it highlights the exact matches in the **Plain text** tab. However, policies can also flag messages where the keywords are hidden and embedded in nonvisible metadata, such as HTML tags or encoded strings. By default, the plain text view doesn't show this content.

The following scenarios show where hidden content matches might occur:

- **HTML tags**: Hyperlinks embedded in HTML tags, for example:

  ```HTML
  <a href="https://example.com">Click here</a>.
  ```

- **Alternative text for images**: Descriptive text for images, for example:

  ```HTML
  <img src="sunset.png" alt="Sunset over the lake">
  ```

- **Encoded strings**: Encoded strings that represent different content when decoded, for example:

  `Base64: U3Vuc2V0IG92ZXIgdGhlIGxha2U=` decodes to *Sunset over the lake*

- **Filename of attachment**: The file name of the attachment might lead to a policy match but isn't clearly highlighted in the text view.

- **OCR:** Text extracted from images using Optical Character Recognition (OCR).

If a message doesn't appear to match the conditions you set for the policy, complete the following steps:

1. Select and download the message with the policy match.
2. Extract the contents of the zipped folder.
3. Locate the .eml file.

4. Open the .eml file with any text editor.
5. Use the *Find* function to search for the individual keywords in the policy conditions to identify where the match occurs.

Review Microsoft Teams meetings transcripts

If you deploy Microsoft Teams in your organization, you can review Teams meetings transcripts for actionable alerts. Teams transcripts are automatically included if you choose Teams as a Microsoft 365 location when you create a custom policy or when you create a policy based on a template.

- **Scheduled meetings**: Communication Compliance ignores communication direction for Teams transcripts. If an individual is an invitee or is present in a scheduled (nonrecurring) meeting, the policy includes all of the meeting content, regardless of who says what in the meeting. This approach also helps in situations where a user is attending the meeting through a hub device in a conference room, since it's not always possible to tell whether an offending communication comes from an in-scope user. Since all meeting content is included, an investigator can review the recording of the meeting to determine if the offending communication is by an in-scope user.

- **Recurring meetings**: For recurring meetings, the policy evaluates only the following users:
  - Users who were invited to the meeting
  - Users identified by the transcript as having spoken during the meeting

- **Unscheduled meetings**: For unscheduled meetings (Meet now meetings), the policy evaluates only users who are identified by the transcript as having spoken during the meeting.

Requirements

To review Teams meeting transcripts, you must turn on meeting transcripts for your organization, since meeting transcripts aren't turned on by default. Learn more about turning on meeting transcripts for your organization.

Limitations

The following limitations apply when reviewing Microsoft Teams meeting transcripts in Communication Compliance:

- You can set one or more of the following policy conditions:

- Content matches any of these classifiers.
- Content contains any of these sensitive info types.
- Message contains any of these words.
- Message contains none of these words.

Any other policy conditions are ignored. Choose conditions for your policies.

> ⓘ **Note**
>
> If you don't set any conditions, the policy captures transcripts for all meetings.

- Only sensitive info types, keyword lists, and regulatory Microsoft provided trainable classifiers are detected. Regulatory trainable classifiers include:
  - Corporate sabotage
  - Gifts and entertainment
  - Money laundering
  - Regulatory collusion
  - Stock manipulation
  - Unauthorized disclosure

  > ⓘ **Note**
  >
  > All other classifiers, including business conduct classifiers, aren't detected.

- Ad-hoc (unscheduled) meetings aren't captured.
- External meetings aren't captured if the meeting organizer is outside your organization.
- Meeting recordings started by an uninvited user aren't captured.

Resolve an alert for a communication in a meeting transcript

After you create a policy, an alert triggers when the policy detects a transcript that contains offending content. The alert brings the offending content and background context to the attention of an investigator.

To resolve an alert related to a meeting transcript:

1. Select the **Source** tab and then review the transcript for the offending content. The **Source** tab shows the entire transcript. When you select the **Source** tab, the transcript automatically

scrolls to the line that contains the policy match. The offensive keyword or phrase is highlighted.

2. Use the **Plain text** tab to do a line-by-line review of the text, including start and stop times in relation to the overall meeting time. Text is captured 30 seconds before and after the offending communication.
3. Select the **Translation** tab to review translations in up to eight languages for the **Plain text** tab. Messages in other languages are automatically converted to the display language of the reviewer.
4. Select the **User history** tab to see a historical view of all user message remediation activities, such as past notifications and escalations for policy matches.
5. Use **Resolve**, **Notify**, **Tag as**, **Escalate**, and **Escalate for investigation** options to resolve the alert. To learn more about these options, see Decide on a remediation action.

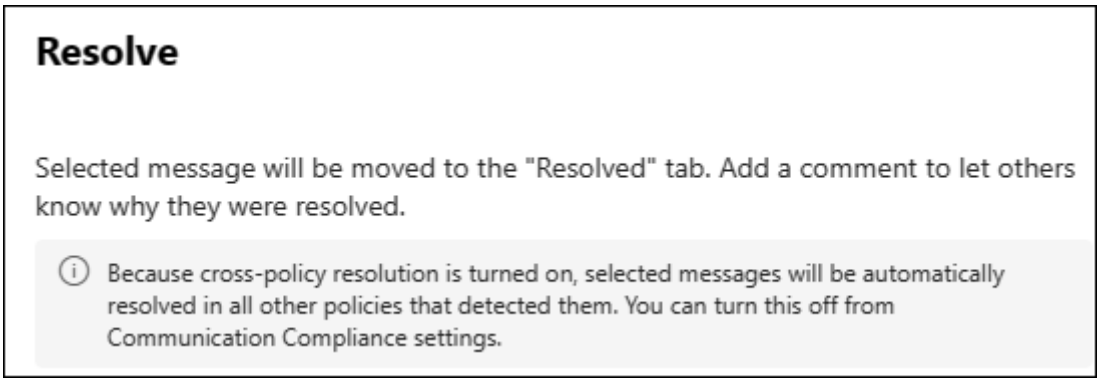
Decide on a remediation action

After reviewing message details, choose from several remediation actions:

- **Resolve**: Selecting **Resolve** immediately removes the message from the **Pending** queue and prevents any further action on the message. When you select **Resolve**, you close the message without further classification. You can also mark the message as misclassified if the alerting process or any Microsoft provided trainable classifiers incorrectly generated it. The **Resolved** tab displays all resolved messages.

  To save investigators time when duplicate policy matches appear across multiple policies, the solution turns on the **Cross-policy resolution** setting (preview) by default. When you resolve a policy match, Communication Compliance automatically resolves all instances of the same policy match in any policy where it's detected, regardless of the reviewer's scope. The item history for each policy notes cross-policy resolution actions, and the number of **Resolved** items increments in the *Items and actions per policy* and *Items and actions per location* reports for all related policies.

  When the **Cross-policy resolution** setting is turned on, this message appears in the **Resolve** pane.



If you don't want to automatically resolve all instances of the same policy match when an investigator resolves a policy match, turn off the **Cross-policy resolution** setting. Anyone who can resolve messages (anyone with the *Communication Compliance*, *Communication Compliance Investigators*, or *Communication Compliance Analyst* role) knows if the setting is turned off because the message in the **Resolve** pane disappears. To turn off the setting, in the Microsoft Purview portal, select **Settings** in the upper-right corner of the page, select **Communication Compliance**, then select the **Cross-policy resolution** setting.

- **Power Automate**: Use a Power Automate flow to automate process tasks for a message. By default, Communication Compliance includes the *Notify manager when a user has a Communication Compliance alert* flow template that reviewers can use to automate the notification process for users with message alerts. For more information about creating and managing Power Automate flows in Communication Compliance, see the Consider Power Automate flows.

- **Tag as**: Tag the message as *Compliant*, *Noncompliant*, or *Questionable* as it relates to the policies and standards for your organization. You can also create a custom tag. Adding tags and tagging comments helps you micro-filter messages for escalations or as part of other internal review processes. After tagging is complete, you can also choose to resolve the message to move it out of the pending queue. You can filter on any tag value.

- **Notify**: Use **Notify** to assign a custom notice template to the message and send a warning notice to the user. Choose the appropriate notice template configured in the **Communication Compliance settings** area and select **Send** to email a reminder to the user that sent the message and to resolve the issue.

- **Escalate**: Use **Escalate** to choose other people in your organization who should review the message. Choose from a list of reviewers configured in the Communication Compliance policy to send an email notification requesting additional review of the message. The selected reviewer can use a link in the email notification to go directly to items escalated to them for review.

- **Escalate for investigation**: Use **Escalate for investigation** to create a new eDiscovery (Premium) case for single or multiple messages. Provide a name and notes for the new case. The custodian is automatically filled in for you. You don't need any additional permissions to manage the case. Creating a case doesn't resolve or create a new tag for the message. You can select a total of 100 messages when creating an eDiscovery (Premium) case during the remediation process. Messages in all communication channels included in Communication Compliance are supported. For example, you could select 50 Microsoft Teams chats, 25 Exchange Online email messages, and 25 Viva Engage messages when you open a new eDiscovery (Premium) case for a user.

- **Remove message in Teams**: Use **Remove message in Teams** to block potentially inappropriate messages and content identified in messages from Microsoft Teams channels and 1:1 and group chats. This action includes Teams chat messages reported by users and chat messages detected using machine-learning and classifier-based Communication Compliance policies. Removed messages and content are replaced with a policy tip that explains that it's blocked and the policy that applies to its removal from view. Recipients are provided a link in the policy tip to learn more about the applicable policy and the review process. The sender receives a policy tip for the blocked message and content but can review the details of the blocked message and content for context regarding the removal.

Create a custom tag

You can tag a policy match as *Compliant*, *Non-compliant*, or *Questionable* as it relates to the policies and standards for your organization. If you need more flexibility than the standard tags provide, create a custom tag. For example, you might want to create an *Escalated* tag so investigators can let other team members know that a policy match is already escalated. Since you can apply multiple tags to any policy match, using the same example, investigators could apply that tag and also apply the *Non-compliant* tag for a policy match.

Keep the following points in mind for custom tags:

- You can apply tags to policy matches that appear on the **Pending** tab.
- You can apply multiple tags to a policy match.
- You can create up to 10 custom tags per policy.
- You create custom tags at the policy level so if you want to use the same custom tag for multiple policies, you must recreate the tag for each policy.
- When you apply a custom tag to a policy match:
  - The tag appears in the **Tags** column in the list of policy matches. If you apply more than one tag, each tag appears in the **Tags** column.

    - The tag appears in the **Tags** filter. You can then select the checkbox for that tag to filter all policy matches by that tag.
    - A **Custom tag** column appears in the following reports: *Items and actions per policy*, *Items and actions per location*, and *Activity by user*.
- If you delete a custom tag for a policy:
    - It's removed from any previous items that you tagged with that custom tag.
    - You can no longer filter by that tag.

> **Tip**
>
> Any action you take with custom tags is tracked in the **History of remediation actions** pane. Select **View item history** to display this pane.

Create and apply a custom tag for a policy match

1. On **Policies**, select a policy to view the policy matches for that policy.
2. Select the checkbox for the policy match that you want to create a custom tag for. Select the checkboxes for multiple policy matches if you want to apply a custom tag in bulk.
3. Select **Tag as** on the command bar.
4. In the **Tag item** pane, select **Add a new tag**.
5. Enter the name of the new tag, then press Enter to add the new tag.

   > **Note**
   >
   > A custom tag must be between 3 and 64 characters and can't include special characters. Spaces and hyphens are allowed.

6. Select the checkbox for the new tag to apply it to the policy match.
7. In the **Comment** box, add a comment to describe the purpose of the tag (optional).
8. Select **Save**.

Edit or delete a custom tag

1. On **Policies**, select a policy to view the policy matches for that policy.
2. Select the checkbox for the policy match that has the custom tag you want to edit.
3. Select **Tag as** on the command bar.
4. In the **Tag item** pane, select the ellipsis next to the custom tag.
5. Select **Edit**, make your changes, or select **Delete** to delete the tag.

> **Important**
>
> If you delete a custom tag, it's removed from any previous items that you tagged with that custom tag. You can no longer filter by that tag.

6. Select **Save**.

Unresolve messages

When you resolve messages, the messages are removed from the **Pending** tab and displayed in the **Resolved** tab. Investigation and remediation actions aren't available for messages in the *Resolved* tab. However, you might need to take extra action on a message that you mistakenly resolved or that needs further investigation after initial resolution. Use the **Unresolve** command to move one or more messages from the **Resolved** tab back to the **Pending** tab.

Unresolve a message

1. Sign in to the Microsoft Purview portal with credentials for an admin account in your Microsoft 365 organization.
2. Go to the **Communication Compliance** solution.
3. Select **Policies** in the left navigation, then select a policy that contains the resolved message to view the policy matches.
4. Select the **Resolved** tab.
5. On the **Resolved** tab, select one or more messages.
6. On the command bar, select **Unresolve**.
7. On the **Unresolve item** pane, add any comments, then select **Save**.
8. Select the **Pending** tab to verify that the selected items are displayed.

Archive message details outside of Communication Compliance (optional)

Export or download message details if you need to archive the messages in a separate storage solution. If you select one or more policy matches and then select **Download** in the bar over the list, you automatically add the selected messages to a .zip file that you can save in storage outside of Microsoft 365. The size limit for files downloaded by using **Download** is 3 MB.

Export a collection of message details that exceeds 3 MB in size

To download messages that cumulatively exceed 3 MB in size, use **Export files** (displayed in the upper-right corner of the **Policies** page). For example, you might want to download policy matches each week for archiving purposes. The download size limit for the **Export files** command is 3 GB. The maximum number of items depends on whether you select documents (limit of 1,000 items) or users (limit of 50,000). Files are exported to a .zip file that you can then download. The .zip file contains all the message files as well as a summary .TXT file that includes data for the following fields: Document, Doc ID, Custodian, Subject/title, File name, File type, Error, and Message.

On the **Exports** tab, the export has a status of *In progress* or *Ready to download*. To download a file that has a *Ready to download* status, select the files batch in the **Name** list, then select **Download export(s)** over the list.

You can either select the policy matches that you want to export, select **Export files**, and then select the list of documents or users you want to export, or you can select **Export files** and then select the documents, users, and date ranges that you want to export. After you select **Export**, it takes a few minutes for the job to complete. You receive an email notification with a link to the **Exports** tab when the export is ready to download. To download a file that has a *Ready to download* status, select the files batch in the **Name** list, then select **Download export(s)** over the list.

> **Note**
>
> To include attachments in exported files, you must select them manually in the list of policy matches. They're not automatically included with their parent message file.

> **Tip**
>
> You can also **create and download a Message details report**.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 06/22/2026

Best practices for managing the volume of alerts in Communication Compliance

> **Important**
>
> Microsoft Purview Communication Compliance provides the tools to help organizations detect regulatory compliance (for example, SEC or FINRA) and business conduct violations such as sensitive or confidential information, harassing or threatening language, and sharing of adult content. Communication Compliance is built with privacy by design. Usernames are pseudonymized by default, role-based access controls are built in, investigators are opted in by an admin, and audit logs are in place to help ensure user-level privacy.

After you configure Microsoft Purview Communication Compliance, make some adjustments to manage the volume of alerts that you receive. Use the list of best practices in this article to help create policies that cover as many users as possible while reducing the number of nonactionable alerts.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Understand keyword list volumes

Many customers use custom keyword lists for compliance scenarios. Understanding the volume of policy matches for each keyword can help you tune your policies. Use the Sensitive information type per location report to analyze keyword lists to see which keywords trigger most matches. You can investigate further to see if those keywords have high false-positive rates. You can also use the Message details reports to get data on keyword matches for a specific policy.

Use Microsoft Purview Reports and the data explorers

It's important to understand the volume of items classified by Microsoft provided trainable classifiers and sensitive information types. Microsoft Purview Reports is the primary tool you

should use. You can also use these tools to help you understand the volume that you can expect for your organization.

- Get started with content explorer (classic)
- Get started with data explorer
- Get started with activity explorer

When you first start using Microsoft provided trainable classifiers, you might not get enough matches, or you might get too many matches. The following table shows the volume level to expect for different types of Microsoft provided trainable classifiers.

Expand table

Microsoft trainable classifier	Volume
Adult images	Low
Customer complaints	Medium
Discrimination	Low
Gifts & entertainment	Medium
Gory images	Medium
Money laundering	Medium
Profanity	Medium
Racy images	Medium
Regulatory collusion	Medium
Stock manipulation	Medium
Targeted harassment	Low
Threat	Low
Unauthorized disclosure	High

Consider using the *Adult images* classifier instead of the *Racy images* classifier since the *Adult images* classifier detects a more explicit image. You can use the Content explorer to help you understand the volume that you can expect for your organization for each of the Microsoft provided trainable classifiers.

Filter email blasts

You can filter out email messages that are generic and intended for mass communication. For example, you can filter out spam, newsletters, and similar messages. For more information, see Learn about the Email blast senders report.

Filter out email signatures and disclaimers

Sensitive information types can be triggered from footers in emails, such as disclaimers. If many of your nonactionable alerts come from a specific set of sentences or phrases in an email signature or disclaimer, you can filter out the email signature or disclaimer.

Use sentiment evaluation

Messages in alerts include sentiment evaluation to help you quickly prioritize potentially riskier messages to address first. Using sentiment evaluation doesn't reduce your detection volumes but makes it easier to prioritize detections. Messages are flagged as *Positive*, *Negative*, or *Neutral* sentiment. For some organizations, messages with *Positive* sentiment might be a lower priority, allowing you to spend more time on other message alerts.

Report messages as misclassified

Reporting false positives as misclassified helps improve Microsoft's models and reduces the number of false positives that you see in the future.

Filter out specific senders by using a condition

If you have senders that consistently trigger detections, you can filter out these particular senders by using the following conditional setting:

- Message isn't received from any of these domains.

Some examples of consistently triggering detections can occur through newsletters, automated mails, and so on. For more scenario information, see Scenarios for creating conditions in Communication Compliance policies.

Use communication direction to target a particular set of users

If you're detecting standards of business conduct scenarios and only care about communications from your users (not from guests), consider using a policy that detects only outbound

communications. If you make the entire organization in scope, you can ensure that all of the users in your organization are covered but exclude users from outside your organization.

Combine trainable classifiers

Consider combining two or more Microsoft provided trainable classifiers together. For example, combine the Threat and Profanity classifiers or the Targeted harassment and Profanity classifiers to raise the threshold for messages captured.

Lower the percentage of reviewed communications

To sample a subset of all the messages that trigger alerts, specify a percentage of communications to review.

Last updated on 06/26/2026

Microsoft Purview Compliance Manager

What is Compliance Manager?

Microsoft Purview Compliance Manager is a solution that helps you automatically assess and manage compliance across your multicloud environment. Compliance Manager can help you throughout your compliance journey, from taking inventory of your data protection risks to managing the complexities of implementing controls, staying current with regulations and certifications, and reporting to auditors.

Watch the video below to learn how Compliance Manager can help simplify how your organization manages compliance:

https://learn-video.azurefd.net/vod/player?id=d9134cf3-a01d-4629-900e-8d32188377ee&locale=en-us&embedUrl=%2Fpurview%2Fcompliance-manager

Compliance Manager helps simplify compliance and reduce risk by providing:

- Pre-built assessments for common industry and regional standards and regulations, or custom assessments to meet your unique compliance needs (available assessments depend on your licensing agreement; learn more).
- Workflow capabilities to help you efficiently complete your risk assessments through a single tool.
- Detailed step-by-step guidance on suggested improvement actions to help you comply with the standards and regulations that are most relevant for your organization. For actions that are managed by Microsoft, you'll see implementation details and audit results.
- A risk-based compliance score to help you understand your compliance posture by measuring your progress in completing improvement actions.

The Compliance Manager overview page shows your current compliance score, helps you see what needs attention, and guides you to key improvement actions.

Understanding your compliance score

Compliance Manager awards you points for completing improvement actions taken to comply with a regulation, standard, or policy, and combines those points into an overall compliance score. Each action has a different impact on your score depending on the potential risks involved. Your compliance score can help prioritize which action to focus on to improve your

overall compliance posture. Compliance Manager gives you an initial score based on the Microsoft 365 data protection baseline. This baseline is a set of controls that includes key regulations and standards for data protection and general data governance.

Learn more

- Understand scoring in Compliance Manager.
- Learn how to work with improvement actions.

Key elements: controls, assessments, regulations, improvement actions

Compliance Manager uses several data elements to help you manage your compliance activities. As you use Compliance Manager to assign, test, and monitor compliance activities, it's helpful to have a basic understanding of the key elements: controls, assessments, regulations, and improvement actions.

Be sure to check out the Compliance Manager glossary of terms.

Controls

A control is a requirement of a regulation, standard, or policy. It defines how you assess and manage system configuration, organizational process, and people responsible for meeting a specific requirement of a regulation, standard, or policy. Compliance Manager tracks the following types of controls:

1. **Microsoft managed controls**: controls for Microsoft cloud services, which Microsoft is responsible for implementing
2. **Your controls**: sometimes referred to as customer managed controls, these are controls implemented and managed by your organization
3. **Shared controls**: these are controls that both your organization and Microsoft share responsibility for implementing

Learn more about monitoring control progress.

Assessments

An assessment is grouping of controls from a specific regulation, standard, or policy. Completing the actions within an assessment help you meet the requirements of a standard, regulation, or law. For example, you may have an assessment that, when you complete all

actions within it, helps to bring your Microsoft 365 settings in line with ISO 27001 requirements. Assessments have several components:

- **In-scope services**: the specific set of Microsoft services applicable to the assessment
- **Microsoft managed controls**: controls for Microsoft cloud services, which Microsoft implements on your behalf
- **Your controls**: sometimes referred to as customer managed controls, these are controls implemented and managed by your organization
- **Shared controls**: these are controls that both your organization and Microsoft share responsibility for implementing
- **Assessment score**: shows your progress in achieving total possible points from actions within the assessment that are managed by your organization and by Microsoft

Learn more about creating and managing assessments.

Regulations

Compliance Manager provides over 360 regulatory templates to help you quickly create assessments. For organizations with unique compliance needs, you can also create custom regulation templates. Learn more about working with regulations in Compliance Manager and view the full list of regulations. Learn more about creating regulation templates.

Improvement actions

Improvement actions help centralize your compliance activities. Each improvement action provides recommended guidance that's intended to help you align with data protection regulations and standards. Improvement actions can be assigned to users in your organization to perform implementation and testing work. You can also store evidence, notes, and record status updates within the improvement action. Learn more about working with improvement actions.

Supported languages

Compliance Manager is available in the following languages:

- English
- Bahasa Indonesian
- Bahasa Malay
- Chinese (Simplified)
- Chinese (Traditional)
- Czech

- Danish
- Dutch
- Finnish
- French
- German
- Hebrew
- Hungarian
- Italian
- Japanese
- Korean
- Norwegian
- Polish
- Portuguese (Brazilian)
- Russian
- Spanish
- Swedish
- Thai
- Turkish

Next steps

- Sign in, assign permissions and roles, configure settings, and personalize your dashboard view.
- Learn about and set up multicloud support.
- Create assessments to help you comply with industry standards that matter most to your organization.
- Get detailed scenario-based guidance on using Compliance Manager and other Purview solutions to help you manage data privacy and data protection.

Last updated on 12/03/2025

Compliance Manager glossary

Article • 02/04/2025

This glossary provides a brief description of important terms and concepts in the context of Microsoft Purview Compliance Manager. This glossary can help you learn and use the solution tools and features quickly and effectively.

Expand table

Term	Description
Assessment	A grouping of controls from a specific regulation, standard, or policy. Completing the actions within an assessment helps you meet the requirements of a standard, regulation, or law.
Control	Generally, a specific measure or action that an organization implements to mitigate or manage risks associated with a particular requirement or objective of a regulation, standard, or policy. As used in Compliance Manager: A control is a requirement of a regulation that defines how you assess and manage system configuration, organizational process, and people responsible for meeting a specific requirement of a regulation, standard, or policy.
Improvement action	A compliance activity with recommended implementation instructions, intended to help towards completion of a control.
License	In the context of Compliance Manager regulations: A single Compliance Manager license allows you to create an unlimited number of assessments for multiple versions of a regulation.
Regulation	A rule or requirement imposed by a governing authority, such as a government agency, to achieve a specific purpose. Also commonly understood as a standard or framework. Compliance Manager supports several industry regulations, providing over 360 regulatory templates for building assessments.
Service	A data source, such as Microsoft Azure or Amazon Web Services (AWS), or more broadly, the digital entity that's being assessed and that benefits from the actions taken. For an assessment, you designate the service that it should evaluate. Completing an improvement action in the assessment benefits the service.
Service instance	For Compliance Manager connectors, each service instance represents an account with a non-Microsoft service provider. For example, an organization may have multiple accounts in Salesforce, such as one for development and testing, one for production, etc. Connectors are set up for each service instance using one email address and password. So an organization may have several connectors for one service, which enables the organization to monitor assessment progress across all instances of a service.

Term	Description
Solution	A feature or capability used to complete an improvement action. For example, a Microsoft product, such as Microsoft Data Loss Prevention, or a setting in a service like Azure or AWS.
Subscription	A type of account to create, assess, and manage a service covered by Microsoft Defender for Cloud, such as Azure, Google Cloud Platform, or Amazon Web Services. Examples: an Azure account for development and testing purposes, an Azure account for production, etc.
Virtual resources	A cloud computing-based resource that is managed virtually, such as VMs and virtual storage disks.

Feedback

Was this page helpful? Yes No

Compliance Manager frequently asked questions

Are there licensing requirements for using Compliance Manager?

Yes. Organizations with Office 365 and Microsoft 365 licenses, and US Government Community (GCC) Moderate, GCC High, and Department of Defense (DoD) customers, have access to Compliance Manager. However, the assessments available to your organization and how you manage assessment templates depends on your licensing agreement. Visit the licensing guidance for details.

Can I use Compliance Manager for non-Microsoft products?

Yes. Learn more about multicloud support in Compliance Manager.

What changed with template licensing in December 2022?

Organizations at the E5/A5/G5 levels have greater flexibility in which regulatory templates they can use for free as part of their licensing agreement. These organizations can choose up to three premium regulatory templates for free instead of having a predetermined set of included templates. The Microsoft Data Protection Baseline is also included by default as part of your subscription.

For customers using any of the templates that were included before the December 2022 change (NIST 800-53, ISO 27001, GDPR), continued use of those templates will count against the allotted three free premium templates. Customers who have already purchased one or more premium templates can use three more premium templates for free.

Also new in December 2022: Templates that belong to the same regulation family now count as one template. For example, the templates for CMMC Levels 1 through 5, are now considered a

single template so that you don't need to purchase multiple versions. When you purchase a template license for a regulation, the license applies for all levels and versions of that regulation. Review the licensing details.

If I have a high score, does it mean I'm fully compliant?

No. Your compliance score measures your progress in completing recommended actions that help reduce risks around data protection and regulatory standards. It doesn't express an absolute measure of organizational compliance with regard to a particular standard or regulation. Compliance Manager, and your compliance score, shouldn't be interpreted as a guarantee in any way.

Last updated on 05/21/2026

Get started with Compliance Manager

Article • 11/27/2024

> **Tip**
>
> If you're not an E5 customer, use the 90-day Microsoft Purview solutions trial to explore how additional Purview capabilities can help your organization manage data security and compliance needs. Start now at the **Microsoft Purview trials hub**. Learn details about **signing up and trial terms**.

Who can access Compliance Manager

Compliance Manager is available to organizations with Office 365 and Microsoft 365 licenses, and to US Government Community Cloud (GCC) Moderate, GCC High, and Department of Defense (DoD) customers. Assessment availability and management capabilities depend on your licensing agreement. View service description details.

Sign in

1. Sign in to the Microsoft Purview portal using credentials for an admin account in your Microsoft 365 organization:
2. Select **Compliance Manager** on the left navigation pane. You arrive at your Compliance Manager dashboard.

Set user permissions and assign roles

Compliance Manager uses a role-based access control (RBAC) permission model. Only users who are assigned a role can access Compliance Manager, and the actions allowed by each user are restricted by role type. Our RBAC model also allows you to grant user access to individual assessments. See role-based access to assessments below to learn more.

The person holding the Global Administrator role for your organization can set user permissions for Compliance Manager. Permissions can be set in one of the following places:

- Permissions in the Microsoft Purview portal
- Permissions in Microsoft Entra ID

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

> **Note**
>
> Users with Microsoft Entra identities who don't have Office 365 or Microsoft 365 subscriptions won't be able to access Compliance Manager in the Microsoft Purview portal. To seek assistance in accessing Compliance Manager, contact cmresearch@microsoft.com.

> **Note**
>
> Customers in US Government Community (GCC) High and Department of Defense (DoD) environments can only set user permissions and roles for Compliance Manager in Microsoft Entra ID. See the Microsoft Entra permissions instructions and role type definitions.

Role types

The table below shows the functions allowed by each role in Compliance Manager. The table also shows how each Microsoft Entra role maps to Compliance Manager roles. Users need at least the Compliance Manager reader role, or Microsoft Entra global reader role, to access Compliance Manager.

A user can only hold one role at a time. Any change in a user's role overrides their previous role.

Expand table

User can:	Compliance Manager role	Microsoft Entra role
Read but not edit data	Compliance Manager Reader	Microsoft Entra Global reader, Security reader

User can:	Compliance Manager role	Microsoft Entra role
Edit data and create assessments - for example, can edit improvement action status, enter notes, upload evidence; can create assessments	Compliance Manager Contribution	Compliance Administrator
Edit data only - can't create assessments	Compliance Manager Assessor	Compliance Administrator
Manage assessments, regulatory templates, and tenant data; assign improvement actions	Compliance Manager Administration	Compliance Administrator, Compliance Data Administrator, Security Administrator

Role-based access to assessments and regulations

You can assign roles to users in order to grant access to specific assessments, or all assessments based on the regulation. Granting user access in these ways is useful when you need to ensure that only the people working on certain regulatory requirements have access to that data.

These four roles provide access to assessments:

- Compliance Manager Reader
- Compliance Manager Contribution
- Compliance Manager Assessor
- Compliance Manager Administration

What you can do with each assessment remains restricted based on which activities the role allows.

To grant users access to an assessment or all assessments for a regulation, open its details page and select **Manage users access** to add users by role. If a user has a role assigned to them in the Microsoft Purview portal for overall access to Compliance Manager, any role you assign them for a specific assessment applies only to that assessment.

More information:

- See Grant user access to assessments for more detailed instructions.
- See Grant user access to regulations for more detailed instructions.
- Learn more about managing all users' access to assessments in Compliance Manager settings.

Start a premium assessments trial

The Compliance Manager premium assessments trial is a great way to quickly create assessments that are most relevant to your organization. Our library of over 360 regulatory templates correspond to governmental regulations and industry standards around the world. Learn more about the premium assessments trial.

Compliance Manager settings

To locate Compliance Manager settings in the Microsoft Purview portal, select the **Settings** icon at the top of the page, then select **Compliance Manager**.

The types of settings for Compliance Manager include:

- Testing source: Allows you to turn off or on the automatic testing of improvement actions.
- Manage user history: Allows you to manage the data of users associated to improvement actions, including the ability to reassign improvement actions to a different user.
- User access: Allows you to view and manage user roles for access to assessments or assessment templates.
- Connectors: Allows you to activate and manage connectors for non-Microsoft services. Get details about accessing connectors in the portal.

Compliance Manager settings are accessible by users who hold a Compliance Manager Administrator role or Global Administrator role.

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

> **Note**
>
> The automated testing feature is not available to customers in GCC High and DoD environments because Secure Score isn't available in these environments. GCC High and DoD customers will need to manually implement and test their improvement actions.

Testing source for automated testing

Compliance Manager detects various signals to provide automated testing and monitoring of improvement actions. This automation derives from three primary sources: built-in, Microsoft Secure Score, and Microsoft Defender for Cloud (get details about automated testing sources). Compliance Manager also detects signals from Microsoft Priva (this capability is in preview; learn more). When an improvement action is successfully tested and implemented, you receive the maximum possible points for that action, which gets credited to your overall compliance score.

Initial settings

While automated testing helps maximize efficiency in your compliance activities, you have full control over whether to apply automated testing. Here's what's initially set up by Compliance Manager and how you can make changes:

- The first time you use Compliance Manager, automatic testing is turned on by default for all actions that can be automatically tested. It takes approximately seven days to fully collect data and factor that data into your compliance score.
- You can turn off automatic testing for all actions, which disables all automated activity on improvement actions by Compliance Manager. You can also select individual improvement actions for automatic testing. See Manage automated testing settings for instructions.

How to tell which actions are tested automatically

On an improvement action's details page, the **Testing type** status in the top information bar shows how the action is tested. If the value is listed as **Automatic**, then the action is automatically tested. If the value is **Manual**, then the action is tested by your organization. Get details about improvement action testing type.

When actions are added or updated

When automation becomes available for existing improvement actions, or when new automated improvement actions are added, the default behavior depends on your current settings and whether or not you've brought data into an improvement action. When you add your own testing data or evidence into an improvement action, automatic testing is turned off for that action to ensure that Compliance Manager doesn't overwrite any of your data.

If you've added data to an existing improvement action:

- Automated testing remains **off** by default for the action. You can choose to turn it back on.

If you haven't added data to an improvement action, or when new automated actions are added to Compliance Manager, the testing behavior follows your current settings:

- Automated testing is **on** if your setting is **on** for **all actions** or **per action**.
- Automated testing is **off** if your setting is **off for all actions**.

Manage automated testing settings

The Compliance Manager Administrator and the Global Administrator for your organization can change the settings for automated testing at any time. You can turn off automated testing for common improvement actions, or turn it on for individual actions. Follow the instructions below to change your automated testing settings. When you change a setting, it will be reflected in the audit log. Learn more about Compliance Manager activities in the audit log.

> **ⓘ Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

1. In Compliance Manager, select **Compliance Manager settings** in top right corner of the page.
2. Select **Testing source** from the left navigation.
3. Select your desired option for testing improvement actions:
   a. Turn on automatic testing for all actions.
   b. Turn off automatic testing for all actions.
   c. Turn on automatic testing per action.
4. If you select **Turn on per improvement action**, a list displays all the improvement actions that are eligible for testing. All actions are checked by default, so you need to uncheck the actions you **don't** want to be automatically tested.
5. Select **Save** to save your settings. You receive a confirmation message at the top of your screen that your selection was saved. If you receive a failure notice, try again.

Manage user history

The **Manage user history** settings help you quickly identify which users have worked with improvement actions in Compliance Manager. The identifiable user data associated with improvement actions includes the status of the improvement actions and documents they uploaded. Understanding and retrieving this type of data might be necessary for your organization's own compliance needs.

The user history settings also allow you to reassign all improvement actions from one user to another.

To find the user history settings:

1. In Compliance Manager, select **Compliance Manager settings** in the top right corner of the page.
2. Select **Manage user history** from the left navigation.

The **Manage user history** page shows a list of all users by email address who are assigned to an improvement action. Use the **Search** button to quickly find a specific user by typing in their email address.

To the right of each user's email address, the **Select** drop-down menu provides options to export a report, reassign improvement actions, or delete history. See each section below for details about each option.

Export a report of user history data

You can export an Excel file containing a list of improvement actions currently assigned to a user. The report also lists any evidence files uploaded by that user. This information can help you reassign open improvement actions.

The report reflects the improvement action's status as of its creation date. It's not a historical report of all previous changes to its status or assignment (learn how to export a report from your improvement actions page).

Follow the steps below to export a report by user:

1. In Compliance Manager, select **Compliance Manager settings** in top right part of the page.
2. Select **Manage user history** from the left navigation.

3. Find your intended user by searching the list email addresses, or by selecting **Search** and entering the user's email address.
4. From the **Select** drop-down menu, choose **Export report**.
5. Once the Excel file of your report is generated, you can open it and save it to your local machine.

Reassign improvement actions to another user

You can reassign ownership of improvement actions from one user to another. When you reassign an action, the evidence upload history doesn't change, but the name of the user who originally uploaded evidence no longer appears within the improvement action.

Follow the steps below to reassign improvement actions to another user:

1. In Compliance Manager, select **Compliance Manager settings** in top right corner of the page.
2. Select **Manage user history** from the left navigation.
3. Find a user by searching the list email addresses, or by selecting **Search** and entering that user's email address.
4. From the **Select** drop-down menu, choose **Reassign improvement actions**. The **Reassign improvement actions** flyout pane appears.
5. In the **Search users** field, enter the name or email address of the user to whom you're assigning the improvement actions.
6. When you see the name of your intended user under **Improvement actions will be assigned to**, select the user, then select **Assign actions**.
7. When the reassignment is complete, you see a confirmation message in the flyout pane confirming that all improvement actions from the previous user have been reassigned to the new user. If you receive a reassignment failure notice, close the window and try again. To close the flyout pane, select **Done**.

The new owner receives an email that they've been assigned to an improvement action. The email contains a direct link into the improvement action's details page.

> ⓘ **Note**

If you reassign an action that has a pending update, the direct link to the action in the reassignment email will break if the update is accepted after reassignment. You can fix this by re-assigning the action to the user after the update is accepted. Learn more about updates to improvement actions.

Delete user history

Deleting a user's history removes them as an owner of improvement actions and removes their name from all other fields in Compliance Manager. When you delete a user's history, the improvement actions they owned won't display an **Assigned to** value until a new user is assigned. Evidence uploaded to the improvement action shows **User removed** in place of the deleted user's name. Deleting user history is permanent.

To delete a user's history, follow the steps below:

1. In Compliance Manager, select **Compliance Manager settings** in top right corner of the page.
2. Select **Manage user history** from the left navigation.
3. Find a user by searching the list email addresses on the page, or by selecting **Search** and entering that user's email address.
4. From the **Select** drop-down menu, choose **Delete history**.
5. A window asks you to confirm the permanent deletion of the user's history. To continue with deletion, select **Delete history**. To leave without deleting the history, select **Cancel**.

You arrive back at the **Manage user history** page with a confirmation message at the top that the history for the user was deleted.

User access

The **User access** section of **Settings** displays a list of all users who have a role that allows access to one or more assessments. From this page, you can make changes to role assignments. When you change a setting, it will be reflected in the audit log. Learn more about Compliance Manager activities in the audit log.

- When you grant a user access to an **assessment**: The user has access to just that one assessment.

- When you grant a user access to a **regulation**: The user will have access to any assessment created with that regulation, including existing assessments and any assessments created in the future.

To add or remove user access roles for assessments and regulations, follow the steps below:

1. In Compliance Manager, select **Compliance Manager settings** in top right part of the page.
2. Select **User access** from the left navigation.
3. Select the checkbox next to the name of one or more users whose role you want to edit.
4. Depending on whether you're editing roles for assessments or regulations: From the **Edit assessment roles** or **Edit regulation roles** dropdown menu above the list of names, select **Add assessment/regulation permissions** or **Remove assessment/regulation permissions**.
5. **For adding a role**: From the flyout pane, go to the tab that corresponds to the role you want to add (Reader, Assessor, or Contributor), then select **Add assessments/regulations**. On the next flyout pane, choose the checkbox next to the assessments/regulations and select **Apply**, then select **Save**.
6. **For removing a role**: From the flyout pane, go to the tab that corresponds to the role you want to remove (Reader, Assessor, or Contributor). Select the button next to the assessments/regulations for which you want to remove access, and select the X mark in the **Remove** column.
    a. A **Remove access?** confirmation box appears. Select **Confirm** to remove the user's role, or select **Cancel** to cancel. The name of the assessments will now be removed from the role tab.
    b. Select **Save** on the flyout pane. The role removal won't be completed until you select the **Save** button. Selecting **Close** will cancel out of the process without saving the role removal.

The user list on the **User access** page will now reflect the changes you made.

> ⓘ **Note**
>
> Admins whose permissions for Compliance Manager were set in Microsoft Entra ID won't appear on the **User access** page. This means that if a user has access to one

or more assessments, and their role is Global Administrator, Compliance Administrator, Compliance Data Administrator, or Security Administrator, they won't appear on this page. Learn more about **setting Compliance Manager permissions**.

More information:

- Visit these instructions for assigning user roles from an assessment's details page.

Understand the Compliance Manager dashboard

The Compliance Manager dashboard is designed to provide you an at-a-glance view of your current compliance posture.

Overview page

Your compliance score

Your compliance score is featured prominently at the top of the **Overview** page. The score is a percentage based on points achievable for completing improvement actions that address key data protection standards and regulations. Points from Microsoft actions, which are managed my Microsoft, also count toward your compliance score.

When you come to Compliance Manager for the first time, your initial score is based on the Microsoft 365 data protection baseline. This baseline assessment, which is available to all organizations, is a set of controls that includes common industry regulations and standards. Compliance Manager checks your existing Microsoft 365 solutions and gives you an initial assessment based on your current privacy and security settings. As you add assessments that are relevant to your organization, your score becomes more meaningful for you.

Learn more: Understand how your compliance score is calculated.

Key improvement actions

This section lists the top improvement actions you can take right now to make the largest positive impact on your overall compliance score. Select **View all improvement actions** to go to your improvement actions page.

Solutions that affect your score

This section highlights solutions containing improvement actions that can positively impact your score, and the number of outstanding improvement actions in those solutions. Select **View all solutions** to visit your solutions page.

Compliance score breakdown

This section gives you a more detailed view of your score in two different ways:

- **Categories**: shows the percentage of your overall score within data protection categories, such as "protect information" or "manage devices."
- **Assessments**: shows the percentage of your progress in managing assessments for particular compliance and data protection standards, regulations, or laws, such as GDPR or NIST 800-53.

Filtering your dashboard view

You can filter your dashboard view to see only the items related to particular regulations and standards, solutions, type of action, assessment groups, or data protection categories. Filtering your view in this way will also filter the score on your dashboard, showing how many points you've achieved out of total possible points based on your filter criteria.

To apply filters:

1. Select **Filter** on the upper-right side of the dashboard.
2. Select your filter criteria from the **Filters** flyout pane, then select **Apply**.

After you apply a filter, you'll see your score adjusted in real time. The compliance score percentage and breakdown information, and the improvement actions and solutions, now only pertain to data covered by your filter criteria. If you sign out of Compliance Manager, your filtered view remains when you sign back in.

To remove filters:

- At the **Applied filters** heading above your compliance score, select the **X** next to the individual filter you want to remove; or
- Select **Filter** on the upper-right side of your dashboard, then on the **Filters** flyout pane, select **Clear filters**.

Improvement actions page

Improvement actions are recommended actions that can help to centralize your compliance activities and align with data protection regulations and standards. Each improvement action gives detailed implementation guidance and a link to launch you into the appropriate solution. Improvement actions can be assigned to users in your organization to perform implementation and testing work. You can also store evidence, notes, and record status updates within the improvement action.

Solutions page

The **Solutions** page shows the share of earned and potential points as organized by solution. Viewing your remaining points and improvement actions from this view helps you understand which solutions need more immediate attention.

Find the solutions page by selecting the **Solutions** tab on your Compliance Manager dashboard. You can also select **View all solutions** underneath **Solutions that affect your score** in the upper-right section of your dashboard. To filter your view of solutions:

1. Select **Filter** at the top-left corner of your assessments list.
2. On the **Filters** flyout pane, place a check next to the desired criteria (regulations, solutions, action types, groups, categories).
3. Select the **Apply** button. The filter pane closes and you see your filtered view.

You can also modify your view to see assessments by group, product, or regulation by selecting the type of grouping from the **Group** drop-down menu above your assessments list.

Taking action from the solution page

The **Solutions** page displays your organization's solutions that are connected to improvement actions. The table lists each solution's contribution to your overall score, the points achieved and possible within that solution, and the remaining number of improvement actions grouped in that solution that can increase your score.

There are two ways you can take action from this screen:

1. On the row of your intended solution, under the **Remaining actions** column, select the hyperlinked number. You see a filtered view of the improvement actions screen showing untested improvement actions for that solution.

2. On the row of your intended solution, under the **Open solution** column, select **Open**. You arrive at the solution's location in the Microsoft Purview portal,

Microsoft Defender portal, or its admin center, where you can take the recommended action.

Assessments page

The **Assessments** page lists all the assessments you set up for your organization. Your compliance score denominator is determined by all your tracked assessments. As you add more assessments, you see more improvement actions listed on your improvement actions page, and your compliance score denominator increases. See Build and manage assessments to learn about assessment status and how to create new assessments.

Regulations page

A regulatory template is a framework for creating an assessment in Compliance Manager. The **Regulations** page displays a list of regulatory templates and key details. The **Free regulation licenses used/Purchased regulation licenses used** counter near the top of the page shows the number of active regulations currently in use out of the total number available for your organization to use. See Regulation availability and licensing for more information.

Policies page

The **Policies** page lists policies you set up to alert you to key changes within Compliance Manager that affect your compliance score. Visit Compliance Manager alerts and alert policies to learn more.

Alerts page

The **Alerts** page lists alerts generated by policies you set up to track changes within Compliance Manager that affect your compliance score. Visit Compliance Manager alerts and alert policies to learn more.

Reports page

The **Reports** page displays a history report of activities in Compliance Manager within the previous seven days. You can look up data from the previous six months by adjusting filters.

The **Reports** page displays your compliance score and a chart representing your organization's overall compliance posture. Select **View detailed report** to bring up a detailed history report page. A table lists improvement actions that have had a score change within the last month. It takes about 24 hours for changes in an action to be reflected in the report.

Select an improvement action from its row on the table to see a details pane that explains the action, the related regulation, associated assessments, and score update. You can launch into the action by selecting **Go to action**. Select the **History** tab on the pane to view the score history for the action.

Filter by date range, regulation, and service

On the history report page, you can adjust the filters to broaden or refine the view of actions that appear in the table. Select the **Date range** filter to select a range that can retrieve data up to six months prior. Select **Add filter** to apply other filters that apply to actions based on regulation or by cloud service.

Next step

Begin setting up your assessments.

Feedback

Was this page helpful? Yes No

Multicloud support in Compliance Manager

09/26/2025

In this article: Learn how Compliance Manager helps you automatically assess and manage compliance across your multicloud environment.

Overview

Compliance Manager now integrates with Microsoft Defender for Cloud, which allows you to assess your compliance posture across Microsoft 365, Microsoft Azure, Google Cloud Platform (GCP), and Amazon Web Services (AWS) with resource-level testing and cloud-specific guidance. This new integration provides customers with a single interface in Compliance Manager to help make it easier to manage compliance across the organization's digital estate.

This integration allows Compliance Manger to track configurations in your organization's Microsoft Azure environment and detect signals from other services like GCP and AWS, so that you can assess your progress in meeting controls for the regulations you need to comply with. Compliance Manager provides guidance for implementing improvement actions in non-Microsoft services for meeting controls.

The integration with Defender for Cloud surfaces in two contexts within Compliance Manager:

1. Assessments: When you create an assessment in Compliance Manager, you'll select a supported regulation and then choose one or more services to assess. Compliance Manger then provides automatic monitoring of configurations in your selected services to determine whether controls are passing or failing.

2. Improvement actions: Using signals from Defender for Cloud, Compliance Manager automatically detects the test status and test result of improvement actions that pertain to Azure and to your non-Microsoft services. With these signals, Compliance Manager automatically tracks the status of actions and resource-level testing details from cloud services like AWS and GCP.

Supported services

The services listed below can be assessed by Compliance Manager:

- Microsoft 365
- Microsoft Azure cloud services

- Google Cloud Platform
- Amazon Web Services

Add a new service

If you want to assess a non-Microsoft service that isn't listed above, you can add it as a new service when you're creating a new assessment. When you add a new service, the universal version of the underlying regulation is used, and you perform manual implementation and testing work. For instructions on adding a new service, see step 5 of the assessment creation process.

> **Tip**
>
> Compliance Manager has a dedicated selection of connectors to support other non-Microsoft services such as Salesforce and Zoom. Visit **Working with connectors in Compliance Manager**.

Service subscriptions

When creating assessments, you can select a subscription if the service you choose for the assessment is monitored by Defender for Cloud. Your choice of subscription will affect the evaluation of improvement actions for that service. Learn more about monitoring assessment progress by service.

If you choose subscriptions that are in scope within Defender for Cloud for a matching regulation, automated test results are pulled from Defender for Cloud and shown in the assessment.

Supported regulations

View the list of regulations supported by both Compliance Manager and Defender for Cloud.

Get started

There are setup steps required before you can start building assessments for your cloud services. Visit Configure cloud settings to get started.

Resources

- Quickstart: Set up Microsoft Defender for Cloud

Configure cloud settings for use with Compliance Manager

Setting up for multicloud support

Compliance Manager integrates with Microsoft Defender for Cloud to provide multicloud support. Organizations must have at least one subscription within Microsoft Azure and then enable Defender for Cloud so that Compliance Manager can receive the necessary signals to monitor your cloud services. Once you have Defender for Cloud, you need to assign the relevant industry and regulatory standards to your subscriptions.

Depending on what your organization has already set up, use one of the following sections to get started:

- **You don't have Azure**: Activate Azure and create a subscription
- **You have Azure but don't have Defender for Cloud**: Enable Defender for Cloud on your Azure subscription
- **You have Defender for Cloud but haven't assigned standards**: Assign standards to your cloud service subscriptions

Standards supported by Compliance Manager and Defender for Cloud

The standards or regulations listed below are supported across Defender for Cloud and Compliance Manager. Each standard is available to support Microsoft 365 in addition to the other cloud services listed in parentheses.

> **Tip**
>
> Defender for Cloud refers to “standards,” while Compliance Manager uses “regulations” to refer to the same thing.

- AWS Foundational Security Best Practices
- CIS 1.1.0 (GCP)
- CIS Microsoft Azure Foundations Benchmark v1.1.0 (Azure)
- CIS 1.2.0 (AWS, GCP)

- CIS Microsoft Azure Foundations Benchmark v1.3.0 (Azure)
- CIS Microsoft Azure Foundations Benchmark v1.4.0 (Azure)
- FedRAMP High (Azure)
- FedRAMP Moderate (Azure)
- ISO 27001 (Azure, GCP)
- NIST SP 800-171 Rev.2 (Azure)
- NIST SP 800-53 Rev.4 (Azure)
- NIST SP 800 53 Rev.5 (Azure, AWS, GCP)
- PCI DSS 3.2.1 (AWS, GCP)
- PCI DSS v4.0 (Azure)
- SOC 2 Type 2 (Azure)
- SWIFT CSP-CDCF v2022 (Azure)

Activate Azure and create a subscription

Setting up a subscription within Microsoft Azure is a prerequisite for getting started with Defender for Cloud. If you don't have a subscription, you can sign up for an Azure free account.

Enable Defender for Cloud

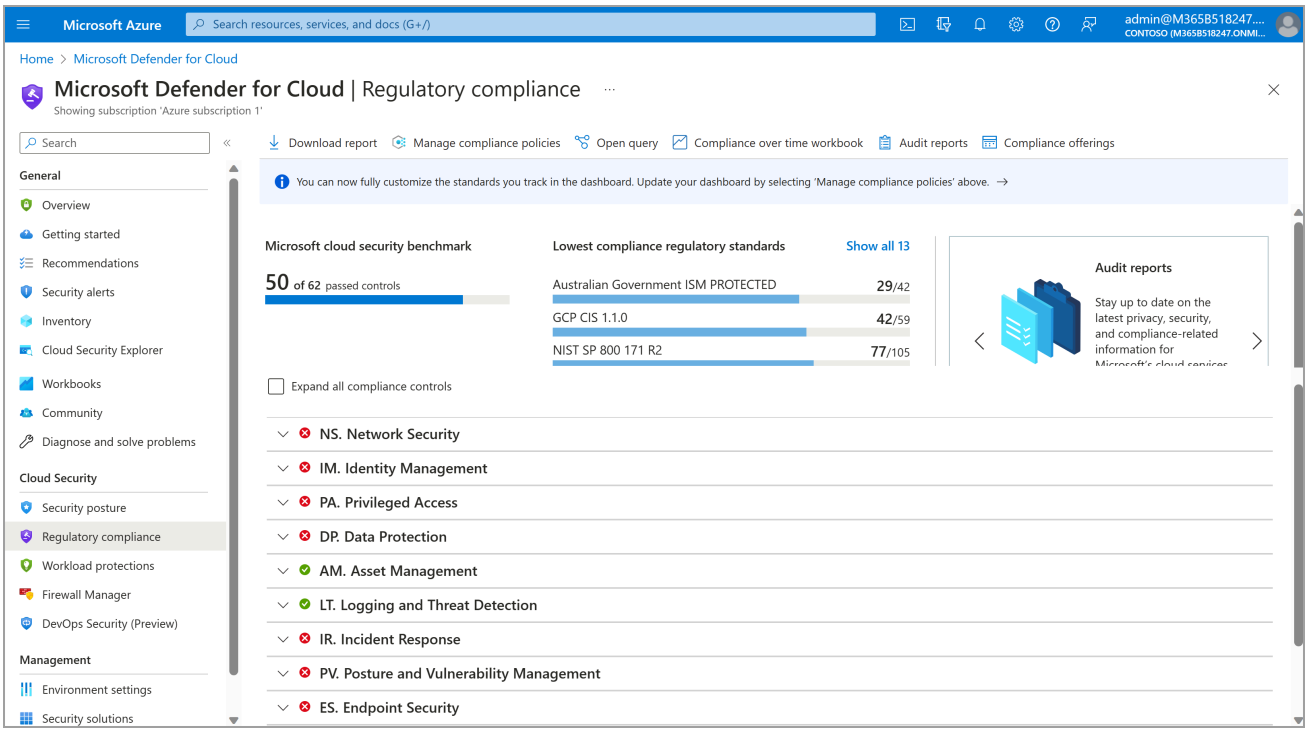
Visit Quickstart: Set up Microsoft Defender for Cloud. Follow the steps to enable Defender for Cloud on your Azure subscription and become familiar with the Defender for Cloud Overview page. Once you've enabled Defender for Cloud, follow the steps below to make sure you're set up for Compliance Manager integration.

Most setup functions require the user to hold the Owner role in Azure. Get more details about User roles and permissions for Defender for Cloud.

Confirm access to Defender for Cloud Regulatory compliance

1. Go to Microsoft Defender for Cloud | Regulatory compliance.

2. Validate that you see a dashboard like this:



3. If you don't see the dashboard above and instead see a notice about insufficient licensing, follow the prompts to activate an applicable Defender for Cloud plan. We recommend enabling one of the two plans: **Foundational CSPM** or **Defender CSPM** (cloud security posture management plans). You can manually select the plans by completing the following steps:
   a. In Defender for Cloud, select **Environment settings** on the left navigation.
   b. Select **Azure** from your list of environments. Expand the item underneath **Azure** to view the subscription, then select the subscription. You arrive at the **Defender plans** page.
   c. In the **Plan** column, find the rows for **Foundational CSPM** and **Defender CSPM**. In the **Status** row, select the **On** button for both plans.

View available environments

1. In Defender for Cloud, select **Environment settings** on the left navigation.

2. View the available environments and subscriptions currently visible to MDC for your tenant. You may need to expand your management groups to view subscriptions, which you can do by selecting **Expand all** below the search bar. In addition to your Azure subscriptions, you'll also see any Google Cloud Platform (GCP) projects or Amazon Web Services (AWS) accounts connected to Defender for Cloud.

3. If you don't see an expected subscription and have already confirmed your Defender for Cloud licensing in the previous steps, check your current directory and subscription filters in your Azure Portal settings. In this view, you can adjust any subscription filters or switch to

a different directory if one is available, and then return to the **Environment settings** view to check the results.

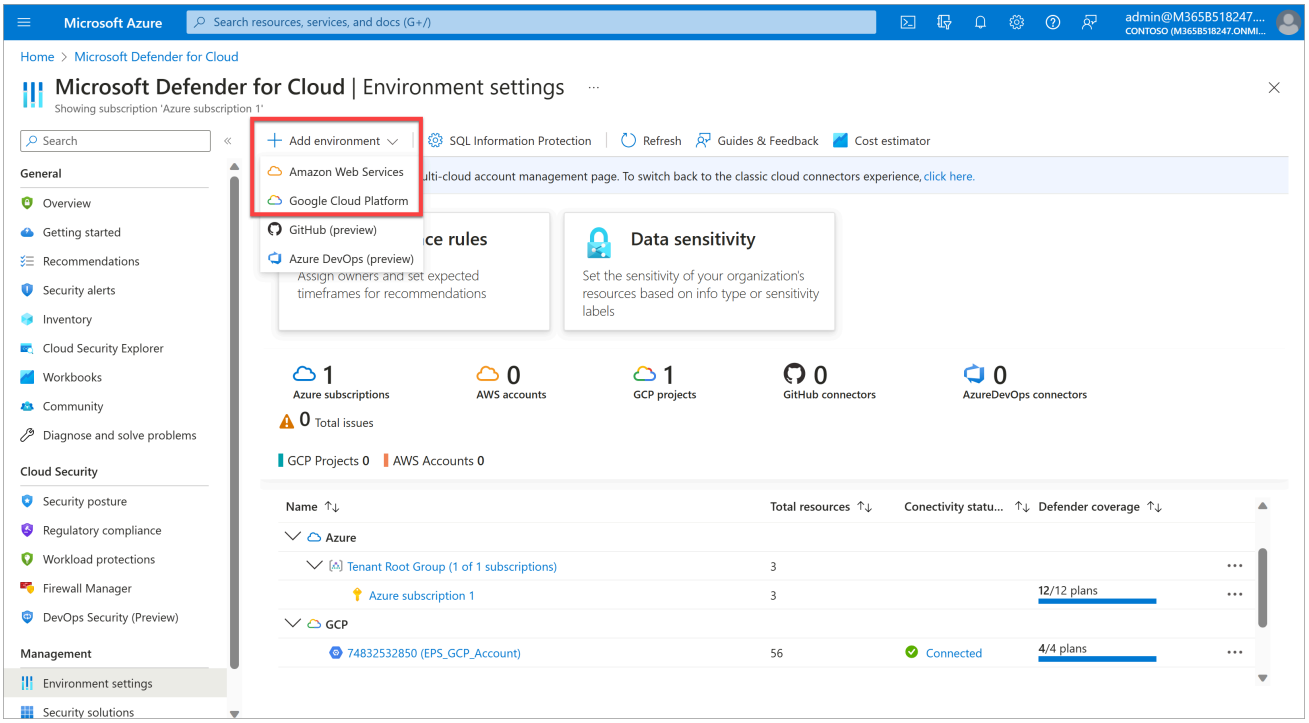
4. If you don't see an expected AWS or GCP environment, account, or project, proceed to the next step to set up the necessary connectors.

Connect to your Amazon Web Services or Google Cloud Provider accounts (optional)

Follow these instructions if you have an Amazon Web Services (AWS) account or Google Cloud Platform (GCP) project that you want Compliance Manager to assess compliance posture, and you don't already see those accounts or projects in your Azure Environment settings. When you complete this process, you can begin assigning standards to your connected AWS or GCP subscriptions within about an hour, though full data can take up to 24 hours to appear.

1. In Defender for Cloud, select **Environment settings** on the left navigation.

2. Select **Add environment** and choose either **Amazon Web Services** or **Google Cloud Platform**.



3. Follow the wizard steps to complete the account setup. Connecting to the accounts requires admin permissions in the AWS or GCP accounts being used, and some configuration steps within AWS or GCP. These steps are detailed in the wizard.
   a. For a simple setup option, consider starting with just one account such as GCP. In the first step of **Account details**, at **Onboard**, select **Single account**. This option requires the least amount of configuration effort.

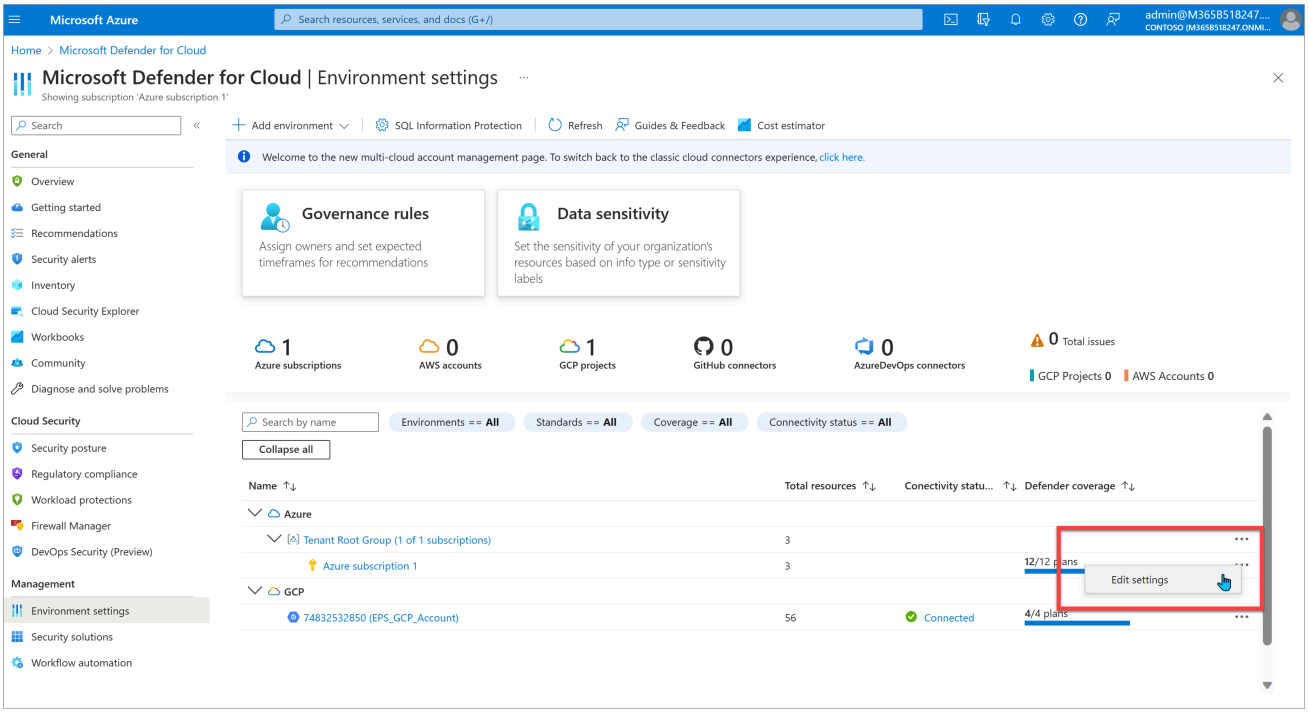
Add standards to your subscriptions

Check the list of standards supported by Defender for Cloud and Compliance Manager to ensure your desired standard is supported. Then follow the steps below.

1. In Defender for Cloud, select **Environment settings** on the left navigation.

2. Your available environments and subscriptions are listed on the page. You may need to expand your management groups to view subscriptions, which you can do by selecting **Expand all** below the search bar. Find the subscription to which you want to add a standard.

3. On the row for the subscription, select the ellipses on the far right and select **Edit settings**.



4. On the left navigation, under **Policy settings**, select **Security policy**.

5. Browse the list of available standards under **Industry & regulatory standards**. You can view more standards by selecting the **Add more standards** button at the bottom of the list. Assign at least one of the supported standards listed below to your subscription by selecting **Enable** on the standard's row.

Related content

For more information about setting up and managing Defender for Cloud, see the following articles:

- Quickstart: Set up Microsoft Defender for Cloud
- User roles and permissions for Defender for Cloud

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 07/01/2026

Working with connectors in Compliance Manager

Article • 03/31/2025

Compliance Manager offers a set of connectors designed to help you gain a clear understanding of your compliance obligations across the services used in your organization. The connectors provide a seamless link to non-Microsoft services so that you can include them in your assessments and take advantage of automatic monitoring and testing of controls. Connecting your services to Compliance Manager requires a few simple setup steps. Once your connector is activated, you can select it as an in-scope service when creating an assessment.

> **Tip**
>
> If you're not an E5 customer, use the 90-day Microsoft Purview solutions trial to explore how additional Purview capabilities can help your organization manage data security and compliance needs. Start now at the Microsoft Purview trials hub. Learn details about signing up and trial terms.

Available connectors

The connectors currently available for Compliance Manager are listed below:

- Corporate Sustainability
- Salesforce
- Zoom

> **Important**
>
> Compliance Manager connectors are unique to the Compliance Manager solution and are set up within Compliance Manager. The data connectors that are accessed from the left navigation of the Microsoft Purview portal aren't compatible with Compliance Manager. We recommend interacting with connectors by following these instructions below to ensure that any connectors you activate are the ones built specifically for use with Compliance Manager.

Roles for setting up connectors

Only users with a Data Connector Admininistrator role or a Global Administrator role can set up Compliance Manager connectors.

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

Connector setup steps

1. Configure settings in your service so that it can integrate with the connector.
2. Activate the connector in Compliance Manager.
3. Select the connector when building an assessment.

Step 1: Configure service settings to enable the connector

For each service, you need to do some configurations and get basic information about your accounts. This information is used to create the connection between the service and the Compliance Manager connector so that the connector can retrieve the necessary signals. Visit the pages listed below and follow the setup instructions before you activate its connector in Compliance Manager:

- Corporate Sustainability
- Salesforce setup
- Zoom setup

Step 2: Activate the connector in Compliance Manager

The next step is to "activate a connector" in Compliance Manager by connecting to your desired service. After you activate a connector, you can select it as a service to be monitored when creating an assessment. The connector allows Compliance Manager to receive signals from your designated non-Microsoft service and evaluate configurations in order to determine whether controls are passing or failing.

If you have more than one account for your product, such as accounts for production, for development and testing, or for backup disaster recovery, you need to activate a connector for each separate account. We refer to each account within a service as a **service instance**.

Follow the steps below to activate a connector:

1. Sign in to the Microsoft Purview portal. Select the **Settings** icon at the top of the page, then select **Compliance Manager**.
2. On the **Compliance Manager settings** page, select **Connectors**.
3. The page displays a list of available connectors to choose from. Select the checkbox next to the name for the product whose connector you want to activate, then select **Activate connector**.
4. You see an information page with details about the connector. Select **Add connector**, which takes you into the connector setup wizard.
5. Review the Microsoft Terms of Service, then select **Accept**.
6. On the **Connector name** page, enter a unique name. You might want a name that helps identify which account the connector is for; for example, "Salesforce Prod" for your Salesforce production account. Then select **Next**.
7. On the **Authentication** page, enter the credentials for the account that you want to link to the connector. The specific fields depend on the service you're connecting to, but are usually a variation of email or account ID, password, and token. See the setup instructions for each service to learn how to get this information. When done entering authentication information, select **Validate connection**. It takes a few seconds to validate the connection.
8. If the connection to your account is validated, you see a **Validation succeeded** message. If validation fails, enter your credentials again until the connection is validated. Select **Next**.
9. On the **Review and finish** page, review the connector details for accuracy. Select **Back** if you need to make changes. Select **Finish** to complete the process.

You arrive back at the Compliance Manager data connectors page. Go to **My activated connectors** tab to view your connector, along with any other connectors created by your organization. If you don't see the newly created connector, refresh your browser.

Step 3: Add a connector to assessments

Now you're ready to build an assessment with a connector so that it tracks progress related to the desired service.

1. From the **Assessments** page in Compliance Manager, select **Add assessment**.
2. From here, begin following the general assessment creations instructions.

   > **Tip**
   >
   > When selecting a regulation for the assessment, the **Supported services** column on the **Select regulation** flyout pane shows which services are supported by the regulation template. If the service related to your connector isn't listed, the connector won't show up as an option in step 3, below.

3. When you arrive at the **Select services** page and choose **Select services**, you should see the name of the product related to your activated connector; for example, Salesforce. Check the box next to the service name, select **Add**, then select **Next**.
4. At the **Service instances** page, you choose one or more connectors you activated for the service. On the row listing your service, select **Manage service instances**.
5. A flyout pane lists all the service instances for your connector. Each instance represents a connector activated by your organization. All service instances are checked by default. If you want the assessment to cover all instances, which means the assessment uses all connectors you activated for that service, leave them all checked and select **Cancel**. Or, uncheck any instances that you don't want to cover in the assessment and select **Update**.
6. Back at the **Service instances** page, select **Next**.
7. On the **Review and finish** page, confirm all your selections are correct, then select **Create assessment**.

The next screen confirms the assessment was created. When you select **Done**, you're taken to your new assessment's details page.

Viewing service data in an assessment

The assessment details page shows a detailed view of the assessment's rate of progress. When you add a connector as a service for the assessment, you see the connector name as one of the services listed on the assessment's **Progress** tab. You can view more details about service progress by selecting **View service details**.

Automated monitoring for connectors

Every 24 hours, connector signals are refreshed and any updated status is reflected in your assessment. This means that Compliance Manager provides daily automated refreshed status of configurations in the products for which you activated connectors.

All improvement actions that are monitored through connectors are automatically monitored and tested so that Compliance Manager can determine whether controls are passing or failing. Improvement action test status is reflected in your assessments so that you can take the necessary actions in your non-Microsoft products to satisfy a control requirement.

Editing a connector

If you need to edit a connector's name or account credentials, select the connector from your **My activated connectors** page and select the **Edit** command at the top of your connectors list.

Deleting a connector

To delete a connector, you need to first locate your connector in the **Data connectors** area of the Microsoft Purview portal. Follow the instructions below to delete a connector:

1. In the Microsoft Purview portal, select **Data connectors** in the left nav, then select **My connectors**.
2. Locate the connector that you want to delete on the list and select the checkbox next to its name. Select the **Delete** command at the top of the list.
3. In the delete confirmation dialog box, select **Delete**.

Your connector is deleted and removed from the list of your connectors.

Feedback

Was this page helpful? Yes No

Salesforce setup for Compliance Manager connector

09/26/2025

Follow the instructions on this page to enable the connection between your Salesforce account and the Compliance Manager connector for Salesforce.

This process involves obtaining a token for a given Salesforce account. Therefore, if you activate multiple connectors for multiple Salesforce accounts, you'll need to repeat this process for each account in order to get the token.

Setup steps

1. Sign in to your Salesforce account with your credentials.
2. Go to your account **Settings** and select **Reset My Security Token**. This can be done by any user.
3. Select **Reset Security** to get a new security token sent to your email.

   > ⓘ **Important**
   >
   > If the security token for your account is used anywhere else, resetting the token may cause issues in those locations.

4. Activate connector in Compliance Manager. When you activate a Salesforce connector in Compliance Manager, you'll provide your **Username**, **Password**, and **Security token**. Once you validate the connection during the connector activation process, you're all set up and ready to begin using the connector. Get details at Working with connectors in Compliance Manager.

Zoom setup for Compliance Manager connector

09/26/2025

Follow the instructions on this page to enable the connection between your Zoom account and the Compliance Manager connector for Zoom.

Zoom must authenticate each HTTP request made to the Zoom API. Zoom supports different authentication methods. Compliance Manager connectors use the **Server-to-Server OAuth app**, which enables you to securely integrate with Zoom APIs and get your account owner access token without user interaction. This is different from the OAuth app type, which requires user authentication. This app type is added and managed across an account by account admins. This app type also enables you to utilize event subscriptions using Webhooks. A paid Zoom account is required to access Zoom APIs.

Setup steps

1. Enable permissions

The user who handles the connector activation process for Zoom needs view and edit permissions so that they can set up a Server-to-Server OAuth app in Zoom. The administrator must enable the Server-to-Server OAuth app role by going to **User Management** > **Roles** > **Role Settings** > **Advanced features** and selecting the **View** and **Edit** checkboxes for **Server-to-Server OAuth** app. See Zoom's Using role management for details.

2. Create a Server-to-Server OAuth app

Follow the steps to create a Server-to-Server OAuth app to use with account credentials.

1. Sign in to your Zoom account and go to the Zoom App Marketplace: https://marketplace.zoom.com/develop/create.
2. From the **Develop** drop-down menu, select **Build application**.
3. Find the **Server-to-Server OAuth** app type card, then select **Create**.
4. Enter a name for your app, then select **Create**.
5. On the **App credentials** page: View and copy your Account ID, Client ID, and Client secret.

> **Tip**
>
> You'll enter these credentials when activating the Zoom connector in Compliance Manager. The client secret expires one hour after you generate it, so you may want to activate the Zoom connector in Compliance Manager soon after creating the Server-to-Server OAuth app.

6. On the **Information** page: Add information about your app, such as a short description and developer contact information (name and email address is required for activation).

7. On the **Feature** page: Set the toggle to your preferred setting for event subscriptions. If enabled, choose the event subscriptions you'd like to use.

8. On the **Scopes** page: If you have the role permission to add scopes, add any scopes that you'd like to enable. Choose **Add Scopes** to search for and add scopes. Select the required scopes listed below, which are required for Microsoft APIs:

    - account:read:lock_settings:admin
    - account:read:managed_domains:admin
    - account:read:settings:admin
    - user:read:user:admin
    - user:read:list_user:admin
    - scheduler:read:list_schedule:admin

9. On the **Activation** page: Select **Activate your app**. Your app should be activated. If you see errors preventing activation, address them and retry the activation. You won't be able to generate an access token to make API calls unless your app is activated.

If your app is deactivated, existing tokens will no longer work. You can also choose to deactivate your app in this section.

3. Activate connector in Compliance Manager

When you activate a Zoom connector in Compliance Manager, you'll provide your Account ID, Client ID, and Client secret. Once you validate the connection during the connector activation process, you're all set up and ready to begin using the connector. Get details at Working with connectors in Compliance Manager.

You can always retrieve your app credentials by following these steps:

1. From the Zoom App Marketplace, select **Manage** in the upper right corner.
2. Select the app you created in step 2 above.
3. On the **App credentials** page, select **Copy** next to the credentials you need to copy.

Build and manage assessments in Compliance Manager

Compliance Manager assessments help your organization evaluate its compliance with industry and regional regulations. Setting up the most relevant assessments for your organization can help you implement policies and operational procedures to limit your compliance risk. Ready-to-use regulatory templates for over 360 regulations contain the necessary controls and improvement actions for completing the assessment.

> **Tip**
>
> Get a comprehensive compliance overview before you deploy Microsoft services in your organization. Learn more about **predeployment compliance with Compliance Manager (preview)**.

Assessments page

All of your assessments are listed on the **Assessments** page Compliance Manager. You can create one assessment that covers multiple services. For example, you can create a single EU GDPR assessment that covers Microsoft 365, Microsoft Azure, Amazon Web Services (AWS), and Google Cloud Platform (GCP). The assessment details page shows a breakdown of control progress by service to help you evaluate how you're doing across all your services. Learn more about monitoring assessment progress from the assessment details page.

> **Important**
>
> The regulations that are available for your organization's use by default depend on your licensing agreement. **Review licensing details**.

The **Free regulation licenses used/Purchased regulation licenses used** counter near the top of the page shows the number of regulations currently in use out of the total number available for your organization to use. Learn more about regulation availability.

Assessment status and details

The assessments page summarizes key information about each assessment:

- **Assessment**: Name of the assessment.
- **Status**: See status types below.
  - **Complete**: All controls have a status of “Passed,” or at least one is passed and the rest are “Out of scope.”
  - **Incomplete**: At least one control has a status of “Failed." Review the failed controls and, within those controls, review both your improvement actions and Microsoft actions to see which have a "Failed" status.
  - **None**: Not all controls have been tested.
  - **In progress**: Improvement actions have a status of “In progress,” “Partial credit,” or “Undetected."
- **Progress**: The percentage of the work done toward completion, as measured by the number of controls successfully tested.
- **Your improvement actions**: The number of completed actions to satisfy implementation of your controls.
- **Microsoft actions**: The number of completed actions to satisfy implementation of Microsoft controls.
- **Group**: The name of the group to which the assessment belongs.
- **Service**: The services covered by the assessment, such as Microsoft 365, Microsoft Azure, or other cloud services.
- **Regulation**: The regulatory template serving as the basis for the assessment.

To filter your view of assessments, follow these steps:

1. Select **Filter** at the top-left corner of your assessments list.
2. On the **Filters** flyout pane, check your desired criteria.
3. Select the **Apply** button. The filter pane closes and you see your filtered view.

You can also modify your view to see assessments by group, product, or regulation by selecting the type of grouping from the **Group** drop-down menu above your assessments list.

Data protection baseline default assessment

To get you started, Microsoft provides a default **Data Protection Baseline** assessment that's included at all subscription levels. This baseline assessment has a set of controls for key regulations and standards for data protection and general data governance. The Data Protection Baseline assessment draws elements primarily from NIST CSF (National Institute of Standards and Technology Cybersecurity Framework) and ISO (International Organization for Standardization),

as well as from FedRAMP (Federal Risk and Authorization Management Program) and GDPR (General Data Protection Regulation of the European Union).

The default Data Protection Baseline assessment is used to calculate your initial compliance score the first time you come to Compliance Manager, before you configure any other assessments. Compliance Manager collects initial signals from your Microsoft 365 solutions. You see at a glance how your organization is performing relative to key data protection standards and regulations, and see suggested improvement actions to take. Compliance Manager becomes more helpful as you build and manage your own assessments to meet your organization's particular needs.

Assessments for AI regulations

Compliance Manager provides four premium regulatory templates to help your organization assess, implement, and strengthen its compliance against AI regulations. These templates are applicable to all generative AI apps that Microsoft Purview supports for AI interactions, such as Microsoft 365 Copilot, Security Copilot, ChatGPT Enterprise, Microsoft Foundry, Gemini and DeepSeek.

The AI regulations listed below align with compliance requirements such as monitoring AI interactions and preventing data loss in AI applications:

- EU Artificial Intelligence Act
- ISO/IEC 23894:2023
- ISO/IEC 42001:2023
- NIST AI Risk Management Framework (RMF) 1.0

Where to find AI regulation templates

On the **Regulations** page in Compliance Manager, the AI regulations are listed under the **Premium AI templates** header. All other premium regulations are listed under the **Premium templates** header. Using an AI regulation counts toward your purchased premium licenses. Learn more about regulation availability and licensing.

How to use AI regulation templates

The **Recommendations** section from Data Security Posture Management for AI provides guided assistance on working with AI regulations. This solution displays recent interactions with sensitive data and recommends actions to take to help you stay compliant with AI regulations.

Automatic Assessments for AI Apps and Agents in Compliance Manager

Compliance Manager integrates with Azure AI Foundry to automate compliance evaluations for AI models and agents. The Compliance Manager integration with Azure AI Foundry syncs evaluation results directly into Compliance Manager, reducing manual effort and improving regulatory alignment. Organizations benefit from built-in assessments for key AI regulations—including the EU AI Act, NIST AI RMF, and ISO/IEC standards—provided free for six months with Copilot or Agent licenses. The automated compliance evaluation capability enables quick compliance checks without additional purchases, streamlining governance for AI deployments.

Prerequisites

Before using automatic assessments for AI apps and agents, make sure the following prerequisites are met:

- The admin account must have either the Azure AI Project Manager or Azure AI User RBAC role assigned to the relevant AI Foundry accounts.
- The Azure AI Project Manager or Azure AI User RBAC role access is required to add Agents during assessment creation for the AI Foundry service in Compliance Manager.

Key Features

- **Scoped Assessments** – Users define the assessment scope directly in Compliance Manager.
- **Comprehensive Action Set** – AI Foundry provides 75 actions, including 15 automated evaluation actions for metrics like reliability, BLEU score, coherence, and fluency.
- **Automated Sync** – Compliance Manager regularly syncs these 15 evaluation actions from AI Foundry, displaying real-time pass/fail status and detailed metrics.
- **Flexible Action Management** – Users can convert automated actions to manual actions through Compliance Manager settings.
- **Manual Update Support** – Allows manual updates for organizations without AI Foundry access.
- **Effort Reduction** – Eliminates repetitive tasks by automatically syncing evaluation results.
- **Detailed Insights** – Provides granular evaluation metrics alongside compliance status for transparency.

Baseline Assessment for Microsoft 365 Copilot and Copilot Chat

Microsoft Purview Compliance Manager provides a baseline assessment for Microsoft Copilot for Microsoft 365 and Copilot Chat. The baseline assessment for Microsoft Copilot for Microsoft 365 and Copilot Chat is derived from the global AI regulations and includes recommended controls to help organizations maintain compliance and reduce risk.

When a Microsoft Copilot for Microsoft 365 or Copilot Chat license is purchased, the baseline assessment is automatically provisioned in the Microsoft 365 admin center. Administrators can view completed actions and pending tasks within the portal to track compliance progress and implement required controls.

Initial steps before creating assessments

Listed below are details about steps and information that will help you prepare for creating an assessment:

- Plan a grouping strategy for your assessments.
- Understand regulatory templates, which contain the controls and action recommendations for assessments.
- Set up connectors if you're assessing non-Microsoft services.

Groups for assessments

When you create an assessment, you must assign it to a group. Groups are containers that allow you to organize assessments in a way that is logical to you, such as by year or regulation, or based on your organization's divisions or geographies. Because groups determine how assessments are organized, we recommend planning a grouping strategy before you create assessments. Below are examples of two groups and their underlying assessments:

- **FFIEC IS assessment 2020**
  - FFIEC IS
- **Data security and privacy assessments**
  - ISO 27001:2013
  - ISO 27018:2014

Different assessments within a group or groups can share improvement actions. Improvement actions can be changes you make within technical solutions mapped to your tenant, like turning on two-factor authentication, or to nontechnical actions you perform outside the system, like instituting a new workplace policy. Any updates in details or status that you make to a technical improvement action will be picked up by assessments across all groups. Nontechnical

improvement action updates will be recognized by assessments within the group where you apply them. Shared improvement actions across assessments and groups allow you to implement one improvement action and meet several requirements simultaneously.

What to know when working with groups

- You can create a group during the process of creating an assessment.
- Groups can't be standalone entities. A group must contain at least one assessment.
- Group names must be unique within your organization.
- Groups don't have security properties. All permissions are associated with assessments.
- Once you add an assessment to a group, the grouping can't be changed.
- If you add a new assessment to an existing group, common information from assessments in that group are copied to the new assessment.
- Related assessment controls in different assessments within the same group automatically update when completed.
- Groups can contain assessments for the same certification or regulation, but each group can only contain one assessment for a specific product-certification pair. For example, a group can't contain two assessments for Office 365 and NIST CSF. A group can contain multiple assessments for the same product only if the corresponding certification or regulation for each one is different.
- Deleting an assessment breaks the relationship between that assessment and the group.
- Groups can't be deleted.

Set up connectors

Compliance Manager has an integrated set of connectors to build assessments that cover non-Microsoft services like Salesforce and Zoom. Visit Working with connectors to learn more and start the setup process.

Create assessments

To create and modify an assessment, a user must hold a role of Compliance Manager Administration, Compliance Manager Assessor, or Global Administrator. Learn more about roles and permissions.

> **Important**

Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview **roles and permissions**.

Before starting to create an assessment, be sure you know which group you'll assign it to, or be prepared to create a new group for this assessment. Read details in the Groups for assessments section. To create an assessment, you use a guided process to select a regulation and designate services.

Build a custom assessment

You can modify a regulatory template by adding controls and improvement actions to create a customized assessment. Visit Build custom assessments (preview) for instructions.

Create an assessment using a guided process

1. From your **Assessments** page, select **Add assessment** to begin the assessment creation wizard.

2. On the **Base your assessment on a regulation** page, select **Select regulation** to choose the regulatory template for the assessment. The **Select regulation** flyout page opens.

3. Use the search box to find your desired regulation, then select the check bubble to the left of the regulation name. Select **Save**, confirm your selection, then select **Next**.

4. On the **Add name and group** page, enter values in the following fields:

   - **Assessment name**: Assessment names must be unique. If the name matches another assessment in any group, you receive an error asking you to create a different name.
   - **Assessment group**: Assign your assessment to a group in one of two ways:
     - **Use existing group** to assign it to a group you created; or
     - **Create new group** to which you'll assign the assessment. Enter a name for this group. You can also **Copy data from an existing group**, such as implementation and testing details and documents, by selecting the appropriate boxes.

   When finished, select **Next**.

5. On the **Select services** page, designate which services this assessment applies to (learn more about multicloud support) using the **Select services** command. The flyout pane shows which services are available for your chosen regulation. Place a check next to your desired services, then select **Add**. Then select **Next**.

- If your desired service isn't listed, you can add it as a new service. When you add a new service, the universal version of the underlying regulation is used, and you perform manual implementation and testing work. To add a new service:
  - On the **Select services** page, select **Add new service**.
  - Enter a name and description for the service.
  - Select **Add**. The service is listed on the **Service** section of the assessment's details page.

6. If you selected a service that has more than one subscription covered by Microsoft Defender for Cloud, you arrive at a substep for **Select service subscriptions**. Select **Manage subscriptions**. On the flyout pane, a tab for each service displays a list of all subscriptions within that service. All subscriptions are selected by default, but you can remove any by selecting the **X** next to the name. On the **Select services** page, select **Next**.

7. **Review and finish:** Review all your selections and make any necessary edits. When you're satisfied with the settings, select **Create assessment**.

The next screen confirms the assessment was created. When you select **Done**, you're taken to your new assessment's details page. If you see an **Assessment failed** screen after selecting **Create assessment**, select **Try again** to re-create your assessment.

Edit an assessment

After creating an assessment, you can edit it to update its name and add or remove services and subscriptions. To update an assessment:

1. From the assessment details page, select the ellipses in the upper right corner and select **Edit assessment**. The assessment update wizard opens.

2. You can update the assessment name on the **Update assessment name** page, or leave it as-is, then select **Next**.

3. On the **Select services** page, add or remove services, then select **Next**.

4. On the **Select service subscriptions** page, select **Manage subscriptions** to make any changes to your subscriptions. Then select **Next**.

5. Review your updates, then select **Modify assessment** to save your changes.

Monitor assessment progress and controls

Each assessment has a details page that gives an at-a-glance view of your progress in completing the assessment. The page shows how your services are performing, and the status of controls and improvement actions. Expand the **Overview** section at the left side of the page to see basic details about the assessment, including its group, regulation, associated services, completion status, and a description.

The **Progress** tab shows the percentage of progress toward assessment completion. The progress bar displays a breakdown showing the number of points achieved within each service covered by the assessment. Get details on each service in the Assessment progress by service section. See all controls within the assessment and their current status on the Controls tab. Quickly access the status of all your improvement actions for the assessment the Your improvement actions tab. The actions handled by Microsoft for the assessment are listed on the Microsoft actions tab.

Assessment progress by service

The **Service** section on the assessment's **Progress** tab helps you understand how you're doing with respect to a regulation with each of your services individually, even at the subscription level, and collectively across your organization. The assessment gets its data on available subscriptions and improvement action status from Microsoft Defender for Cloud. Any errors associated with subscription accessibility should be addressed in your Defender for Cloud. See Configure cloud settings for more information.

Select the **View service details** command, located next to or under the **Assessment progress** bar graph or in the upper-right command bar, to view a flyout pane with more details. The **View service details** flyout pane lists each service and its progress toward completing the assessment. Selecting **View** next to a service name displays another pane that lists each subscription within the service and its status.

On a service's details panel, you see the list of subscriptions within the service that are covered by the assessment. The **Service progress** counter indicates the number of points achieved so far by improvement actions pertaining to the service for the assessment out of the total number of achievable points.

You can add more subscriptions to the service that you want the assessment to cover by editing the assessment.

Controls tab

The **Controls** tab displays detailed information for each control in the assessment. The **Control status breakdown** chart shows the status of controls by family (for example, Configuration

Management and Incident Response) so you can see at a glance which groupings of controls need attention. The table underneath the breakdown chart lists all controls. You can filter the list by control family, status, and service. The table shows the following details about each control:

- **Control title**
- **Status**: The test status of the improvement actions within the control:
  - **Passed**: All improvement actions have a test status of "passed," or at least one is passed and the rest are "out of scope."
  - **Failed** At least one improvement action has a test status of "failed."
  - **None**: All improvement actions haven't been tested.
  - **Out of scope**: All improvement actions are out of scope for this assessment.
  - **In progress**: Improvement actions have a status other than the ones listed above, which could include "in progress," "partial credit," or "undetected."
- **Control ID**: The control's identification number, assigned by its corresponding regulation, standard, or policy.
- **Points achieved**: The number of points earned by completing actions, out of the total number achievable.
- **Your improvement actions**: The number of your actions completed out of the total number to be done.
- **Microsoft actions**: The number of actions completed by Microsoft.

Select a control from the list to view its details page. A graph indicates the test status of the improvement actions within the control. A table below the graph lists the improvement actions for that control. Select an improvement action from the list to drill into the improvement action's details page, from where you can manage implementation and testing. Get details about working with improvement actions.

Your improvement actions tab

The **Improvement actions** tab on the assessment details page lists all your improvement actions for the control. The status bar chart details the aggregated test status of your improvement actions in the assessment so you can quickly gauge what has been tested and what still needs to be done. Hover over or select a test status label to highlight only that status on the bar.

Beneath the bar, a table lists all the actions and key details, including: service, test status, the number of potential and earned points, associated regulations and standards, applicable solution, action type, and control family.

Filter by **Service** to view actions related to a service and their progress. From the table, select an improvement action to go to its details page, from where you can manage implementation and testing. Get details about working with improvement actions.

Microsoft actions tab

The Microsoft actions tab appears for assessments based on templates that support Microsoft products. It lists all the actions in the assessment that are managed by Microsoft. The list shows key action details, including: service, test status, points that contribute to your overall compliance score, associated regulations and standards, applicable solution, action type, and control family. Select an improvement action to view its details page.

Grant user access to individual assessments

When you assign users a Compliance Manager role in the Microsoft Purview portal, they can view or edit data within all assessments by default (review the Compliance Manager role types). You can restrict user access to only certain assessments by managing user roles from within an assessment. Restricting access in this way can help ensure that users who play a role in overseeing compliance with particular regulations or standards have access only to the data and information they need to perform their duties. (You can also set user access for regulations, which allows users to access all assessments created for that regulation.)

External users who need access for auditing or other purposes can also be assigned a role for viewing assessments and editing test data. You provide access to external individual by assigning them a Microsoft Entra role. Learn more about assigning roles.

Steps for granting access

Follow the steps to grant user access to an assessment.

1. From your **Assessments** page, find the assessment you want to grant access to. Select it to open its details page.

2. In the upper-right corner, select **Manage user access**.

3. A **Manage user access** flyout pane appears. It has three tabs, one for each role of Readers, Assessors, and Contributors. Navigate to the tab for the role you want your user to hold for this assessment. Users who currently have access to the assessment will have a blue box with a check mark to the left of their name.

4. Select the + **Add** command for the role tab you're on: **Add reader**, or **Add assessor** or **Add contributor**.

5. Another flyout pane appears which lists all the users in your organization. You can select the checkbox next to the username you want to add, or you can enter their name in the search bar and select the user from there. You can select multiple users at once.

6. After making all your selections, select **Add**.

   > ⓘ **Note**
   >
   > If you assign a role to someone who already has an existing role, the new role assignment you choose will override their existing role. In this case, you'll see a confirmation box asking you to confirm the change in role.

7. The flyout pane closes and you arrive back at your assessment details page. A confirmation message at the top confirms the new role assignment for that assessment.

Steps for removing access

You can remove a user's access to individual assessments by following the steps below:

1. On the assessment's details page, select **Manage user access**.

2. On the **Manage user access** flyout pane, go the tab corresponding to the user's role you want to remove.

3. Find the user whose role you want to remove. Check the circle to the left of their name, then select the **Remove** command just below the role tab. To remove all users at once, select the **Remove all** command without checking the circle next to every user's name.

4. A **Remove access?** dialog appears, asking you to confirm the removal. Select **Remove access** to confirm the role removal.

5. Select **Save** on the flyout pane. The users' roles will now be removed from the assessment.

Learn how to get a broad view of all users with access to assessments.

Note about multiple roles

- A user can have one role that applies to an assessment, while also holding another role that applies broadly to overall Compliance Manager access.

    - For example, if you assigned a user a **Compliance Manager Reader** role, you can also assign that user a **Compliance Manager Assessor** role for a specific assessment. In effect, the user holds the two roles at the same time, but their ability to edit data is limited to the assessment to which they've been assigned the **Assessor** role.
    - Removing an assessment-based role won't remove the user's overall Compliance Manager role if they have one. If you want to change a user's overall role, you have to change it the Microsoft Purview portal. Learn more about assigning roles and permissions.

- For an individual assessment, one user can only hold one assessment-based role at a time.
    - For example, if a user holds a reader role for a GDPR assessment and you want to change them to a contributor role, you'll first need to remove their reader role, and then reassign them the reader role.

> **Note**
>
> Admins whose permissions for Compliance Manager were set in Microsoft Entra ID won't appear on the **Manage user access** flyout pane. This means that if a user has access to one or more assessments, and their role is Global Administrator, Compliance Administrator, Compliance Data Administrator, or Security Administrator, they won't appear on this pane. Learn more about setting Compliance Manager permissions and roles.

Accept updates to assessments

When an update is available for an assessment, you see a notification and have the option to accept the update or defer it for a later time. Updates are available for assessments based on the regulatory templates provided in Compliance Manager. If your organization is using universal templates for assessing other products, inheritance might not be supported.

What causes an update

An assessment update occurs when there are underlying template changes that impact scoring. Changes might involve adjusting control mapping or other guidance based on regulatory changes or product changes. Assessment updates can originate from your organization and from Microsoft.

If Microsoft updates a Compliance Manager template that you extended, your assessment inherits those updates once you accept them. Your assessment retains the other attributes you

applied to the assessment when you extended it.

Custom assessments that you create don't receive any template updates from Microsoft. Custom assessments can receive improvement action updates, but any Microsoft updates to control mapping between assessments and improvement actions don't apply to custom templates.

> **Note**
>
> Updates to assessments apply only at the group level. If you have two assessments built from the same template that exist in two different groups, each assessment will have a pending update notification, and you'll need to accept the update to each assessment in its respective group individually.

Where you see assessment update notifications

The assessment details page also shows a **Pending update** label next to the assessment with an update. Select that assessment to get to its details page.

A message near the top of the assessment details page shows that an update is available for that assessment. Select the **Review update** button in the banner to review the specific changes and accept or defer the update.

The assessment details page might also list improvement actions that have a **Pending update** label next to them. Those updates are for specific changes to the improvement actions themselves and need to be accepted separately. Visit Accepting updates to improvement actions to learn more.

Review update to accept or defer

When you select **Review update** from the assessment details page, a flyout pane appears on the right side of your screen. The flyout pane provides the key details below about the pending update:

- The template title
- Source of the update (Microsoft, your organization, or a specific user)
- The date the update was created
- An overview explaining the update
- Specific details about the changes, including the impact to your compliance score, the amount of progress toward completion of the assessment, and the specific number of changes to improvement actions and controls.

Selecting the **Updated template** command downloads an Excel file containing control data for the version of the template with the pending updates. Selecting the **Current template** command downloads a file of the existing template without the updates.

To accept the update and make the changes to your assessment, select **Accept update**. Accepted changes are permanent.

If you select **Cancel**, the update won't be applied to the assessment. However, you continue to see the **Pending update** notification until you accept the update.

- **Why we recommend accepting updates**: Accepting updates helps ensure you have the most updated guidance on using solutions and taking appropriate improvement actions to help you meet the requirements of the certification at hand.
- **Why you might want to defer an update**: If you're in the middle of completing an assessment, you might want to ensure you finished work on it before you accept an update to the assessment that could disrupt control mapping. You can defer the update for a later time by selecting **Cancel** on the review update flyout pane.

Export an assessment report

You can export an assessment to an Excel file for compliance stakeholders in your organization or for external auditors and regulators. On the assessment details page, select the **Export actions** in the top right corner of the page, which creates an Excel file you can save and share. The report is a snapshot of the assessment as of the date and time of the export. It contains the details for controls managed by both you and Microsoft, including implementation status, test date, and test results.

Delete an assessment

Deleting an assessment removes it from the list on your assessments page. Note these important points about deleting assessments:

- **Deleting an assessment is permanent; you cannot get it back.** If you want to use the same assessment again, you need to re-create it.
- If the improvement actions in the assessment don't appear in any other assessment, they're deleted when the assessment is deleted.
- We recommend exporting a report of the assessment before you permanently delete it.

To delete an assessment, follow the steps below:

1. From the **Assessments** page, select the assessment you wish to delete.
2. On the assessment's details page, select **Delete assessment** in the upper-right corner of your screen. If you don't see this option, select the ellipsis (...) in the upper-right corner, then select **Delete assessment** from the list.
3. A window appears asking you to confirm that you want to permanently delete the assessment. Select **Delete assessment** to close the window. You get a confirmation window that your assessment was deleted from Compliance Manager.

> **Note**
>
> You can't delete all of your assessments. Organizations need at least one assessment for Compliance Manager to function properly. If the assessment you want to delete is the only one, add another assessment before deleting the other assessment.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 07/01/2026

Build custom assessments (preview) in Compliance Manager

Article • 11/27/2024

> **Tip**
>
> If you're not an E5 customer, use the 90-day Microsoft Purview solutions trial to explore how additional Purview capabilities can help your organization manage data security and compliance needs. Start now at the Microsoft Purview trials hub. Learn details about signing up and trial terms.

Compliance Manager provides the option to modify a regulatory template by adding controls and improvement actions to suit your needs for building a custom assessment. First you copy a regulation, then add controls and improvement actions and services, then publish it so that you can start using it for assessments. This article explains how to copy and modify a regulation in order to create a customized assessment.

Customize a regulation

You can make changes to an existing Microsoft regulatory template in order to customize it for your needs. When you customize a regulation:

- It inherits all the controls and actions of the parent regulation.
- It inherits automatic updates to the Microsoft managed controls and automatically tested actions, except for any actions you modify.
- You can add more actions into a control that already exist in other controls.
- You can create your own actions and controls.
- You can add services.

> **Note**
>
> Copying a premium regulation activates the regulation for your organization and consumes a license for it. The license allows you to create multiple copies of a regulation, and build multiple assessments from a copied regulation. Learn more about regulation licensing.

Updates and versioning

When you customize a regulation, it inherits all of the details, controls, and actions from the parent regulation. The customized regulation receives updates from Microsoft when there are changes to the related regulation or product. However, if you modify an action that is automatically tested, your organization takes over testing and implementation, and the action doesn't receive updates by Microsoft.

> **Tip**
>
> We recommend not making changes to a copied and published regulation for now, as any udpates will not be reflected in versioning history. A versioning exprience that reflects all updates for customized regulations is in our product roadmap for the near future.

Steps for customizing a regulation

1. In Compliance Manager, go to the **Regulations** page.
2. Select the checkbox next to the name of the regulation you want to customize.
3. Select **Customize regulation** above the list of regulations.
4. In the **Create a copy of...** dialog, select **Copy**.

The copy of the assessment is created and the assessment's details page opens.

The name of the regulation is automatically created, starting with "Extension," followed by the name of the regulation, followed by "Copy 1." For example, if you copy the ISO 27005:2018 regulation, the name of the copied regulation will be "Extension ISO 27005:2018 Copy 1."

> **Tip**
>
> After you **publish** the regulation, you can **edit the copied regulation** to change the title or modify the description.

The **Controls** tab lists all of the controls inherited from the original, or parent, regulation. The **Service** lists the services to which the assessment applies, and is where you can add more services to the assessment.

The copied regulation has a status of **Draft** and can't be used to build an assessment until you publish it. To search for your newly copied regulation on the **Regulations** page, set the **Status** filter to **Draft** to see it and any other copied yet unpublished regulations.

A copied regulation can't be deleted.

Add controls

The **Controls** tab of the copied regulation's details page lists all the controls of the underlying regulation. You can add your own controls by following the steps below:

1. On the regulation's **Controls** tab, select **Create control**. The **Create control** flyout pane appears.
2. Enter a **Title** for the control.
3. Enter a **Control ID**. This must be a unique number, up to 150 characters. The ID can't belong to any existing control in this assessment or in any other regulation.
4. Select a control family from the dropdown menu. If the control family name you want isn't listed as an option, you can create a new family name as follows:
    - Select **Create new** from the **Control family** dropdown menu.
    - Enter a name for the new family in the **Control family name** field. This control family name will be added as a dropdown option after you create the control.
5. Enter an optional **Description** for the control.
6. Select **Create**.

After creating the control, you arrive at the details page for the control you created. The next step is to link improvement actions to the control.

Add improvement actions to controls

On the **Your improvement actions** tab of the control details page, you can import actions from the parent regulation, or you can create your own actions.

Import actions

You can import any of the other improvement actions that exist in other controls within the parent regulation into a control. Select **Import actions**. On the **Select actions to import** flyout pane, you can select all the actions by checking the box next to the **Improvement actions** column header. You can also check individual actions by selecting the checkbox next to the names.

After selecting your desired actions, select **Import**.

Create actions

You can create your own improvement actions to add to your copied regulation by following the steps below.

1. Enter a **Title** for the improvement action.
2. Select a **Score** value from the dropdown menu. Score values represent the maximum number of points awarded for action completion. Values are assigned based on whether the action is preventative, detective, or corrective actions, and whether they're mandatory or discretionary. Refer to this explanation of score values to see which number is assigned to which type of action.
3. At **Services**, select a service from the dropdown menu that the action applies to.
4. If the action applies to the Microsoft 365 service, then at **Solution**, select the appropriate solution from the dropdown menu.
5. Enter an optional **Description** for the action.
6. Select **Create**.

After the action is created, you arrive back at the control details page, and the improvement action you just created is listed on the **Your improvement actions** tab.

Add services

The **Services** tab lists the services that apply to the parent regulation. You can add a service that is available for that regulation by selecting **Add available services** and selecting one or more services. If there aren't any services listed, you can create a service and add it to the regulation.

To create your own service, select **Create service**. On the flyout pane, enter a name and description for the service, then select **Save**. After you publish regulation, the added service will be an available option if you make more copies of the same regulation.

Finalize changes by publishing

After adding controls and connecting improvement actions to the regulation you copied, you need to publish the changes in order to begin using the extended regulation for assessments.

1. Go to the details page of the extended regulation.

2. Select **Publish template** in the upper right corner.

After you publish the template, it will be listed on the **Regulations** page and it will be ready for assessment creation. The status of the regulation, which is a filter option, is now **Ready to use** instead of **Draft**.

Simple ways to search for a copied regulation on your **Regulations** page include:

- Set the **Created by** filter to your name.
- Select the **Includes custom content** column to sort by the **True** value. True indicates that the regulation has been customized.

> ⓘ **Note**
>
> You can't create an assessment from a copied regulation until you publish it by selecting **Publish template**.

Edit a copied regulation

After you publish a copied regulation, you can change its name, modify its description, and continue to add and edit controls and actions. To edit a regulation, select it from the list on the **Regulations** page. On the regulation details page, select the ellipsis (...) in the top right corner and select **Edit regulation**.

At this point, a working copy of the regulation is made so that you can make edits and save your work in progress without overriding the published copy. When you're done making edits, you need to again publish it by selecting **Publish template** in the upper right corner. When you select **Publish template**, a **Review and publish** flyout pane appears, where you can review your changes and add publishing notes. Select **Publish template** on the flyout pane to save the changes and republish your regulation.

Creating multiple copies of the same regulation

You can create multiple copies of a regulation so that you can customize it in different ways. You might want to have multiple versions of a regulation if you want to create different control configurations or include only certain services for the same regulation, and then create assessments based off of those extensions.

Multiple copies are automatically named with numbers appended to the name. For example, if you create two copies of the ISO 27005:2018 regulation, they're named as follows: "Extension ISO 27005:2018 Copy 1" and "Extension ISO 27005:2018 Copy 2."

We recommend editing the descriptions of extended regulations so that you can clearly identify how they differ or what they're intended to assess. To edit the description:

1. Go to the regulation's details page.
2. In the **Overview** section, under **Details**, select **Edit details**. The **Edit template details** flyout pane appears.
3. In the **Description** field, enter a detailed description.
4. Select **Save**.

Feedback

Was this page helpful? Yes No

Learn about regulations in Compliance Manager

09/26/2025

> **Important**
>
> The regulations that are available for your organization's use by default depend on your licensing agreement. Review licensing details.

Regulations overview

The **Regulations** page in Compliance Manager displays the list of regulations and certifications for which Compliance Manager provides control-mapping templates. When you build an assessment, you choose the underlying regulation by selecting from among our set of regulatory templates, then select the services you want to assess for that regulation. Setting up Compliance Manager for multicloud support provides you with greater automation in testing and monitoring of controls.

Each regulatory template also comes in a universal version, which provides general control mapping that can broadly apply to services. Universal templates provide the most general type of implementation guidance and require manual implementation and testing by the organization. Note that US Government Community (GCC) Moderate, GCC High, and Department of Defense (DoD) customers can't currently use universal templates.

List of regulations

Visit Compliance Manager regulations list to see the list of all regulatory templates available for building assessments.

Regulation availability and licensing

The Microsoft **Data Protection Baseline** regulatory template is available for organizations at all subscription levels to use. The regulations designated as **premium** require purchase of a license to use them.

On the **Regulations** page in Compliance Manager, the regulations are grouped under four headings:

- **Sub-service compliance readiness**: Use these to help assess your predeployment compliance posture (preview) (does not involve use of a license).
- **Included templates**: The Microsoft data protection baseline.
- **Premium AI templates**: To use for assessments for AI regulations.
- **Premium templates**: All other global, regional, and industry regulations.

When you purchase a license for a regulatory template, you can create as many assessments for that regulation as you wish. Depending on your licensing agreement, your organization might be able to use up to three more **premium** regulatory templates for free. When you begin creating assessments, Compliance Manager tracks how many templates are active so you can monitor your usage.

Resources

- Learn more about included regulations by subscription level.
- Go to the Compliance Manager service description.
- Understand your regulation usage through active and inactive regulation status.

Purchasing premium regulations

Licenses for premium regulatory templates can be obtained in various ways, depending on your Compliance Manager licensing agreement. Once your purchase has been finalized, the templates should become available in your tenant within 48 hours. Licenses for activated regulations are good for one year.

- **Commercial and GCC Moderate**: Purchase licenses in the admin center (learn more about subscriptions, licenses, and billing). Select the quantity of licenses you wish to purchase and your payment plan.
  - Commercial
  - GCC Moderate
  - You can also acquire licenses through your participation in the Cloud Solution Provider program or volume licensing.
- **GCC High and DOD accounts**: Purchase through volume licensing.

Staring a premium trial

You can try out premium regulation templates by acquiring trial versions of the licenses. Trial licenses are good for up to 25 templates for 90 days. Once you obtain your trial license, the templates should become available in your tenant within 48 hours. If your organization has a commercial license for Compliance Manager, you can learn how to start your trial at About the free trial for Microsoft Purview Compliance Manager premium assessments. If your

organization is under a GCC or DOD license, choose the appropriate trial link for your organization:

- GCC Moderate
- GCC High
- DOD

Active and inactive regulations

Regulations display a status as either active or inactive:

- **Active** indicates use in at least one assessment.
- **Inactive** indicates it's not being used for an assessment.

When you use a premium regulation to create an assessment, that regulation's availability status changes to **Active** and the purchased license is active for one year. Your purchase renews automatically unless you cancel.

Regulation licenses counter

The **Regulation licenses used** counter near the top of the **Regulations** page represents the number of regulatory templates in use out of the number you're eligible to use according to your licensing agreement and any purchased licenses. For example, if your counter shows 2/5, it means your organization has activated two regulations out of the five that are available to use. If your counter shows 5/2, it indicates that your organization exceeds its limits and needs to purchase three of the premium regulations in use.

Select **View details** the counter to view a detailed list of all regulations in use and their corresponding assessments.

Regulations details page

Select a regulation from the list on the **Regulations** page to open a regulation's details page. Regulatory templates that belong to the same regulation family are counted as one template. The regulation family is shown in the **Overarching regulation** column on the **Assessment templates** page. When you purchase a template license for a regulation and activate the template, it counts as one activated template even if you create assessments for different levels or versions of that regulation. For example, if you use a template for CMMC Level 1 and a template for CMMC Level 2, your activated templates counter increases by only one.

For more information, see Compliance Manager licensing guidance.

Grant user access to regulations

When you assign users a Compliance Manager role in the Microsoft Purview portal, their role extends by default to all regulations (review the Compliance Manager role types). This access means that, depending on the user's role type, they can view or interact with any existing or future assessments created with that regulation.

You can give users a specific role that applies to certain regulations by managing user roles from a regulation's details page. When you set a user's role for a regulation, the user will be able to interact at that role level for all assessments (existing and future) created with that regulation. (You can also restrict access to individual assessments following these instructions.)

External users who need access for auditing or other purposes can also be assigned an access role for regulations. You provide access to external individual by assigning them a Microsoft Entra role. Learn more about assigning roles.

Steps for granting access

Follow the steps to grant user access to a regulation.

1. From the **Regulations** page, find the regulation you want to grant access to. Select it to open its details page.
2. In the upper-right corner, select **Manage user access**.
3. A **Manage user access** flyout pane appears. It has three tabs, one for each role of Reader, Assessor, and Contributor. Navigate to the tab for the role you want your user to hold for this regulation.
4. Select the **+ Add** command for the role tab you're on: **Add readers**, or **Add assessors** or **Add contributors**.
5. Another flyout pane appears which lists all the users in your organization. You can select the checkbox next to the username you want to add, or you can enter their name in the search bar and select the user from there. You can select multiple users at once.
6. After making all your selections, select **Add**.

> **Note**
>
> If you assign a role to someone who already has an existing role, the new role assignment you choose will override their existing role. In this case, you'll see a confirmation box asking you to confirm the change in role.

7. The flyout pane closes and you arrive back at the regulation details page. Select **Save**. A confirmation message at the top confirms the new role assignment for that regulation.

Steps for removing access

You can remove a user's access to individual regulations by following these steps:

1. On the regulation's details page, select **Manage user access**.
2. On the **Manage user access** flyout pane, go the tab corresponding to the user's role you want to remove.
3. Find the user whose role you want to remove. Check the circle to the left of their name, then select the **Remove** command just below the role tab. To remove all users at once, select the **Remove all** command without checking the circle next to every user's name.
4. A **Remove access?** dialog appears, asking you to confirm the removal. Select **Remove access** to confirm the role removal.
5. Select **Save** on the flyout pane. The users' roles will now be removed from the assessment.

Note about multiple roles

- A user can have one role that applies to a regulation, while also holding another role that applies broadly to overall Compliance Manager access.
  - For example, if you've assigned a user a **Compliance Manager Reader** role, you can also assign that user a **Compliance Manager Assessor** role for a specific regulation. In effect, the user holds the two roles at the same time, but their ability to edit data will be limited to the assessment to which they've been assigned the **Assessor** role.
  - Removing a regulation-based role won't remove the user's overall Compliance Manager role if they have one. If you want to change a user's overall role, you have to change it in the Microsoft Purview portal. Learn more about assigning roles and permissions.
- For an individual regulation, one user can only hold one assessment-based role at a time.
  - For example, if a user holds a reader role for the GDPR regulation and you want to change them to a contributor role, you'll first need to remove their reader role, and then reassign them the reader role.

> **Note**

Admins whose permissions for Compliance Manager were set in Microsoft Entra ID won't appear on the **Manage user access** flyout pane. This means that if a user has access to one or more regulations, and their role is Global Administrator, Compliance Administrator, Compliance Data Administrator, or Security Administrator, they won't appear on this pane. Learn more about **setting Compliance Manager permissions and roles**.

Compliance Manager regulations list

Article • 03/06/2025

In this article: View the comprehensive list of **regulations** available to build assessments in Compliance Manager.

> **Important**
>
> The regulations that are available for your organization's use by default depend on your licensing agreement. Review licensing details.

> **Tip**
>
> If you're not an E5 customer, use the 90-day Microsoft Purview solutions trial to explore how additional Purview capabilities can help your organization manage data security and compliance needs. Start now at the Microsoft Purview trials hub. Learn details about signing up and trial terms.

Overview

Compliance Manager provides a comprehensive set of regulatory templates for creating assessments. These **regulations**, as they're referred to in Compliance Manager, can help your organization comply with national, regional, and industry-specific requirements governing the collection and use of data. Regulations are added to Compliance Manager as new laws and regulations are enacted. Compliance Manager also updates its regulations when the underlying laws or regulations change. Learn more about how to review and accept updates.

List of regulations and where to find them

Below is the complete list of regulations in Compliance Manager. In Compliance Manager, go to the **Regulations** page and select a regulation name to view its description, properties, controls, and associated improvement actions. Jump to a section below to view regulations by area or industry:

- AI
- Global
- Industry

- US Government
- US States and Territories
- Asia-Pacific Countries/Regions
- Europe, Middle East, and Africa (EMEA)
- Latin America
- North America
- South America

Included regulations

Some regulations are included in Compliance Manager by default, depending on subscription level:

- **Customers at all subscription levels**: The Microsoft Data Protection Baseline is included for all organizations as part of their subscription.
- **Customers at the A5/E5/G5 subscription levels**: In addition to the Microsoft Data Protection baseline, you can choose any three premium regulations to use for free.
- **US Government Community (GCC) Moderate, GCC High, and Department of Defense (DoD) customers**: Cybersecurity Maturity Model Certification (CMMC), levels 1 through 5, is included in addition to the Microsoft Data Protection Baseline.

Premium regulations

The regulatory templates listed below may be purchased by your organization. Certain licensing agreements allow for the use of three premium regulations for free. Review licensing details.

AI

- EU Artificial Intelligence Act
- ISO/IEC 23894:2023
- ISO/IEC 42001:2023
- NIST AI Risk Management Framework (RMF) 1.0

Global

- Guidelines and Functional Requirements for Electronic Records Management Systems (ICA Module 2)
- ISO 15489-1:2016

- ISO 16175-1:2020
- ISO 19791 - Information technology — Security techniques — Security assessment of operational systems
- ISO 22301:2019
- ISO 23081-1:2017
- ISO 27005:2018
- ISO 27017:2015
- ISO 27034-1 Information technology — Security techniques — Application security
- ISO 27799: 2016, Health informatics — Information security management in health
- ISO 28000 – Specifications for Security Management Systems for the Supply Chain
- ISO 31000:2018
- ISO 37301
- ISO 55001 – Asset management -- Management systems--Requirements
- ISO IEC 80001-1:2010
- ISO/IEC 27001:2013
- ISO/IEC 27001:2022
- ISO/IEC 27018:2019
- ISO/IEC 27033-1:2015
- ISO/IEC 27701:2019
- NIST 800-207 - Zero Trust Architecture
- SIG 2022
- System and Organization Controls (SOC) 1
- System and Organization Controls (SOC) 2

Industry

- AICPA/CICA Generally Accepted Privacy Principles (GAPP)
- ARMA - Implementing the Generally Accepted Record Keeping Principles (GARP)
- CDSA Content Protection & Security Standard
- CIS Implementation Group 1, Group 2, Group 3
- CIS Microsoft 365 Foundation Level 1 and 2
- CIS Microsoft 365 Foundations Benchmark Level 1 V4.0.0
- CIS Microsoft 365 Foundations Benchmark Level 2 V4.0.0
- Cloud Security Alliance (CSA) Cloud Controls Matrix (CCM)
- COBIT 5
- Digital Operational Resilience Act
- FINRA Cybersecurity Checklist
- ITU X.1052 Information Security Management Framework
- Joint Commission Information Management Standard
- Microsoft Cloud Security Benchmark v1

- Motion Picture Association (MPA) Content Security Best Practices
- NERC CIP
- SWIFT Customer Security Controls
- OWASP ProActive Controls for Developers 2018 v3.0
- (NAIC) Standards for Safeguarding Customer Information Model Regulation MDL-673
- PCI DSS v3.2.1
- PCI DSS v4.0
- Privacy of Consumer Financial and Health Information Regulation, NAIC MDL-672, Q2 2017
- Revisions to the principles for the sound management of operational risk (Basel III Ops Risks)
- Standardized Information Gathering (SIG) Questionnaire
- Trusted Information Security Assessment Exchange (TISAX) 5.1

US Government

- Appendix III to OMB Circular No. A-130 - Security of Federal Automated Information Resources
- CFR - Code of Federal Regulations Title 21, Part 11, Electronic Records, Electronic Signatures
- Children's Online Privacy Protection Rule (COPPA)
- CMMC Level 1, Level 2, Level 3, Level 4, Level 5
- CMMC v2 Level 1
- CMMC v2 Level 2
- CMS Information Systems Security and Privacy Policy (IS2P2)
- Computer Fraud and Abuse Act (CFAA)
- Content of Premarket Submissions for Management of Cybersecurity in Medical Devices
- Criminal Justice Information Services (CJIS) Security Policy
- Cybersecurity for Networked Medical Devices Containing Off-the-Shelf (OTS) Software - FDA
- Cybersecurity Maturity Model Certification (CMMC) Levels 1 through 5
- DFARS
- e-CFR - Identity Theft Rules
- Electronic Code of Federal Regulations - Part 748.0 and Appendix A
- FDIC Privacy Rules
- Federal Financial Institutions Examination Council (FFIEC) Information Security Booklet
- FedRAMP Moderate

- FedRAMP SSP High Baseline
- Freedom of Information Act (FOIA)
- FTC Privacy of Consumer Financial Information
- Gramm-Leach-Bliley Act, Title V, Subtitle A, Financial Privacy
- HIPAA/HITECH
- HITRUST
- Homeland Security Presidential Directive 7: Critical Infrastructure Identification, Prioritization, and Protection
- IRS - Revenue Procedure 98-25 Automated Records
- IRS-P1075
- Minimum Acceptable Risk Standards for Exchanges (MARS-E) 2.0
- National Archives Universal Electronic Records Management (ERM) Requirements
- NIST 800-37
- NIST 800-53 rev.4
- NIST 800-53 rev.5
- NIST 800-63 Digital Identity Guidelines
- NIST 800-78-4: Cryptographic Algorithms and Key Sizes for Personal Identity Verification
- NIST 800-137A -- Assessing Information Security Continuous Monitoring (ISCM) Programs
- NIST 800-171
- NIST 800-184: Guide for Cybersecurity Event Recovery
- NIST CSF
- NIST CSF 2.0
- NIST Privacy Framework
- NIST SP 1800-5 IT Asset Management
- NIST Special Publication 1800-1 Securing Electronic Health Records on Mobile Devices
- NIST Special Publication 800-128
- NIST Special Publication 800-210: General Access Control Guidance for Cloud Systems
- Sarbanes-Oxley Act
- SEC 17-4(a)
- United States of America Privacy Act
- US - Clarifying Lawful Overseas Use of Data (CLOUD) Act
- US - Commission Statement and Guidance on Public Company Cybersecurity Disclosures
- US - Department of Energy (DOE) Assistance to Foreign Atomic Energy Activities
- US - Family Educational Rights and Privacy Act (FERPA)
- US - Federal Information Security Modernization Act of 2014 (FISMA)
- US - Protecting and Securing Chemical Facilities From Terrorist Attacks Act

US States and Territories

- Alabama - Policy 621: Data Breach Notification - DRAFT
- Alaska - Chapter 48 - Personal Information Protection Act
- Arizona - Notification of Breaches in Security Systems
- Arkansas Code Title 4, Subtitle 7, Chapter 110, Personal Information Protection Act
- California - Civil Code Section 1798
- California - Database Breach Act (California SB 1386)
- California - Education Code-EDC, Title 3, Division 14, Part 65, Chapter 2.5- Social Media Privacy
- California - Privacy Rights Act (CPRA)
- California - SB-327 Information Privacy: Connected Devices
- California Consumer Credit Reporting Agencies Act
- Colorado Privacy Act (CPA)
- California Consumer Privacy Act (CCPA)
- Colorado Protections for Consumer Data Privacy
- Colorado Revised Statutes, Section 6-1-716, Notice of Security Breach
- Connecticut - Display and Use of Social Security Numbers and Personal Information
- Connecticut General Statutes - General Provisions for state contractors who receive confidential information
- Connecticut Information Security Program to Safeguard Personal Information
- Connecticut State Law - Breach of security re computerized data containing personal information
- D.C. Law 16-237 - Consumer Personal Information Security Breach Notification Act
- Delaware - Student Data Privacy Protection Act
- Delaware Computer Security Breaches- Commerce and Trade Subtitle II - 12B-100 to 12B-104
- Florida Title XXXII, Chapter 501, Section 501.171, Security of confidential personal information
- Georgia (US) Personal Identity Protection Act
- Guam's Notification of Breaches of Personal Information
- Hawaii - Security Breach of Personal Information Chapter 487N
- Idaho Identity Theft
- Illinois (740 ILCS 14/1) Biometric Information Privacy Act
- Illinois Personal Information Protection Act
- Indiana Disclosure of Security Breach
- Iowa - Student Personal Information Protection Act
- Iowa Code. Title XVI. Chapter 715C. Personal Information Security Breach Protection
- Kansas Consumer Information, Security Breach Statute

- Kentucky Data Breach Notification
- Louisiana Database Security Breach Notification Law (Act No. 382)
- Maine - Act to Protect the Privacy of Online Consumer Information
- Maine - Notice of Risk to Personal Data
- Code of Maryland State Government - Protection of Information by Government Agencies
- Maryland Personal Information Protection Act - Security Breach Notification Requirements, HB 1154
- Maryland's Student Data Privacy Act
- Massachusetts - 201 CMR 17.00: Standards For The Protection Of Personal Information Of Residents Of The Commonwealth
- Massachusetts Data Breach Notification Law 93H section 1-6
- Michigan Identity Theft Protection Act
- Mississippi Security Breach Notification
- Montana - Impediment of Identity Theft
- Nebraska's Data Protection and Consumer Notification of Data Security Breach Act
- Nevada Chapter 603A - Security and Privacy of Personal Information
- Nevada Senate Bill 220 Online Privacy Law
- New Hampshire Right to Privacy Act
- New Jersey Security Breach Disclosure
- New Mexico Chapter 57 - Privacy Protection (Article 57-12B-1 through 4)
- New Mexico Consumer Information Privacy Act
- New Mexico's Data Breach Notification Act
- New York - 23 NYCRR Part 500
- New York City Administrative Code - Security Breach Notification
- New York General Business Law - Data Security Breach Notification and Data Security Protections
- New York Privacy Act
- North Carolina - Identity Theft Protection Act
- North Dakota Chapter 51-30 Notice of Security Breach for Personal Information
- Ohio - Security Breach Notification
- Ohio Data Protection Act 2018
- Oklahoma Security Breach Notification Act
- Oregon Consumer Identity Theft Information Protection Act
- Pennsylvania Breach of Personal Information Notification Act
- Puerto Rico - Citizen Information on Data Banks Security Act
- Rhode Island - Identity Theft Protection Act
- South Carolina - Breach Notification
- South Dakota - Notice of Breach
- Tennessee 47-18-2107 Release of Personal Consumer Information
- Texas - Identity Theft Enforcement and Protection Act

- Texas Privacy Policy to Protect Social Security Numbers
- Utah Consumer Credit Protection Act
- Utah Electronic Information or Data Privacy
- Vermont - Act on Data Privacy and Consumer Protection
- Virginia Breach of Personal Information Act
- Virginia Consumer Data Privacy Act (CDPA)
- Washington DC - Consumer Security Breach Notification Standard
- West Virginia - Breach of Security of Consumer Information
- Wisconsin Security Breach Notification

Regional

Asia-Pacific Countries/Regions

- Asia Pacific Economic Cooperation (APEC) Privacy Framework
- Australia - ASD Essential 8
- Australia - ASD Essential 8 Maturity Level 1
- Australia - ASD Essential 8 Maturity Level 2
- Australia - ASD Essential 8 Maturity Level 3
- Australia - National Archives Act
- Australia - Public Records Office Victoria Recordkeeping Standards
- Australia - Spam Act 2003
- Australia Privacy (Credit Reporting) Code
- Australia Privacy Act
- Australia Public Record Act
- Australian Energy Sector Cyber Security Framework (AESCSF)
- Australian Information Security Registered Assessor Program (IRAP) with ISM Version 3.5 - Official
- Australian Information Security Registered Assessor Program (IRAP) with ISM Version 3.5 - Protected
- Australian Prudential Regulation Authority CPS
- Victorian Protective Data Security Standards V2.0 (VPDSS 2.0)
- Information Management Standard for Australian Government - National Archives of Australia (NAA)
- China - Personal Information Security Specification
- Cybersecurity Law of the People's Republic of China
- Hong Kong - Code of Banking Practice and Payment Card
- Hong Kong - Personal Data (Privacy) Ordinance
- India Digital Personal Data Protection Act

- India Information Technology (Reasonable Security Practices and Procedures and Sensitive Personal Data or Information) Rules
- India - Information Technology Act
- Reserve Bank of India Cyber Security Framework
- Indonesia - Law 11/2008
- Indonesia - Law 27/2022 (Personal Data Protection)
- Japan - Act on Prohibition of Unauthorized Computer Access
- Japan - Common Model of Information Security Measures for Government Agencies and Related Agencies
- Japan - Common Standards for Information Security Measures for Government Agencies and Related Agencies
- Japan Privacy Mark - JIS Q 15001: 2017
- Japanese Act on the Protection of Personal Information (Law No. 57 of 2003)
- Korea - Credit Information Use And Protection Act
- Korea - The Act on Promotion of Information and Communications Network Utilization and Data Protection
- Korea Personal Information Protection Act
- Malaysia - Personal Data Protection Act (PDPA)
- Malaysia Risk Management in Technology (RMiT)
- Myanmar - Law Protecting the Privacy and Security of Citizens
- Nepal - Right to Information Act
- New Zealand - Privacy Act / 2020
- New Zealand - Public Records Act
- New Zealand - Reserve Bank BS11 Outsourcing Policy
- New Zealand - Telecommunications Information Privacy Code
- New Zealand Health Data Retention Policy
- New Zealand Health Information Privacy Code
- New Zealand Health Information Security Framework (HISF)
- New Zealand Information Security Manual (NZISM)
- Pakistan - Electronic Data Protection Act - DRAFT
- Philippines BSP Information Security Management Guidelines
- Philippines Data Privacy Act of 2012
- Singapore - ABS Guidelines on Control Objectives and Procedures for Outsourced Service Providers
- Singapore - Banking Act (Cap.19)
- Singapore - Cybersecurity 2018
- Singapore - IMDA IoT Cyber Security Guide
- Singapore - Monetary Authority of Singapore Technology Risk Management Framework
- Singapore - Multi-Tier Cloud Security (MTCS) Standard
- Singapore - Personal Data Protection Act / 2012

- Singapore Spam Control Act
- Taiwan - Implementation Rules for the Internal Audit and Internal Control System of Electronic Payment Institutions - 2015
- Taiwan - Implementation Rules of Internal Audit and Internal Control System of Financial Holding Companies and Banking
- Taiwan - Regulations Governing Approval and Administration of Financial Information Service Enterprises Engaging in Interbank Funds Transfer and Settlement
- Taiwan - Regulations Governing the Standards for Information System and Security Management of Electronic Payment Institutions
- Taiwan – Trade Secrets Act
- Taiwan Personal Data Protection Act (PDPA)
- Thailand PDPA
- Taiwan – Trade Secrets Act
- Uzbekistan - Law of The Republic of Uzbekistan on Personal Data
- Vietnam - Consumer Rights Protection Law
- Vietnam - Law of Cybersecurity
- Vietnam - Law of Network Information Security
- Vietnam - Law on Information Technology

Europe, Middle East, and Africa (EMEA)

- Albania - The Law on the Protection of Personal Data No. 9887
- Austrian Telecommunications Act 2003
- Armenia - Law of the Republic of Armenia on the Protection of Personal Data
- Belarus Law On Information, Informatization and Protection of information
- Belgium - Act on the Protection of Natural Persons with Regard to the Processing of Personal Data
- Belgium NBB Dec 2015
- Bosnia and Herzegovina Law on the Protection of Personal Data
- Botswana - Data Protection Act
- Bulgaria Law for Protection of Personal Data 2002
- Central Bank of Kuwait Cybersecurity Framework
- Corporate Sustainability Reporting Directive (CSRD)
- Cyprus The Processing of Personal Data Law
- Czech - Act No. 110/2019 Coll. on Personal Data Processing - 2019
- Czech - On Cyber Security and Change of Related Acts (Act on Cyber Security) - Act No. 181
- Denmark - The Data Protection Act

- Denmark - Executive Order on Information and Consent Required in Case of Storing and Accessing Information in End-User Terminal Equipment
- Directive 2013/40/EU Of The European Parliament And Of The Council
- Dubai - Health Data Protection Regulation
- Dubai Consumer Protection Regulations (Telecommunications Regulatory Authority)
- Dubai ISR
- Egypt - Data Protection Law
- Estonia - Personal Data Protection Act
- Estonia - The system of security measures for information systems
- EU - Directive 2006/24/EC
- EU - ePrivacy Directive 2002 58 EC
- EU GDPR (General Data Protection Regulation)
- EudraLex - The Rules Governing Medicinal Products in the European Union
- European Network and Information Security Agency (ENISA) - Cloud Computing Information Assurance Framework
- Finland - Data Protection Act
- Finnish Criteria for Assessment of Information Security of Cloud Services
- France - The Data Protection Act
- Georgia Law on Personal Data Protection
- Germany - Annotated text of the Minimum Requirements for Risk Management
- Germany - Cloud Computing Compliance Controls Catalog (C5)
- Germany - Federal Data Protection Act
- Germany - Supervisory Requirements for IT in Financial Institutions (BAIT)
- Ghana - Data Protection Act
- Ireland Data Protection Act
- Israel - Privacy Protection (Transfer of Data to Databases Abroad) Regulations
- Israel Privacy Law
- Jordan Cloud Platforms & Services Policy
- Kenya Data Protection Act
- Kingdom of Saudi Arabia Personal Data Protection Law (KSA PDPL)
- Luxembourg Act
- Malta - Data Protection Act
- Mauritius Data Protection Act 2004
- Montenegro - Law on Personal Data Protection
- NATO Directive AC/322-D(2021)0032
- Nigeria Data Protection Regulation
- NIS2 Directive (EU) 2022/2555 of the European Parliament and of the Council
- Oman - Electronic Transactions Law
- Qatar Cloud Computing Regulation
- Qatar Cloud Security Policy

- Qatar National Information Assurance (NIA)
- Republic of Moldova Law on Personal Data Protection
- Romania - Data Protection Law 190/2018
- Russia - Federal Law 149-FZ On Information, Information Technology and Information Security
- Russian Federation Federal Law Regarding Personal Data
- Saudi Arabia - Saudi Arabia Monetary Authority (SAMA)
- Saudi Arabia - National Cybersecurity Authority (NCA)
- South Africa Consumer Protection ACT 68 2008
- South Africa Electronic Communications and Transactions Act, 2002
- South Africa - Promotion of Access to Information Act
- South African POPIA
- Slovakia Act on the Protection of Personal Data
- Spain - Nation Security Framework
- Switzerland - Federal Act on Data Protection (FADP)
- Turkey - Information and Communication Security Guide
- Turkey - KVKK Protection of Personal Data 6698
- UAE - Federal Decree Law on Combating Cyber Crimes
- UAE - Federal Law Concerning Electronic Transactions and Commerce
- UAE - Federal Law No 2 of 2019 On the Use of the Information and Communication Technology (ICT) in Health Fields
- UAE - NESA Information Assurance Standards
- UAE Data Privacy Law
- UAE Regulatory Policy TRA - Internet of Things
- UAE's Federal Decree Law Regulating the Telecommunications Sector
- Uganda - The Data Protection and Privacy Act
- UK - Cyber Security for Defence Suppliers Standard 05-138
- UK - The Offshore Petroleum Activities Regulations / 2011
- UK Cyber Essentials
- UK Data Protection Act
- UK Data Retention Act
- UK Privacy and Electronic Communications
- Ukraine - Protection of Personal Data Law
- Yemen - Yemen Law of the Right of Access to Information

Latin America

- Antigua and Barbuda - Data Protection Act /2013
- Bahamas - Data Protection Act
- Barbados - Data Protection Bill 2019

- Barbados - Electronic Transactions Act
- Bermuda - Electronic Transaction Act
- Saint Lucia Data Protection Act
- Trinidad and Tobago Data Protection (Act 13 of 2011)

North America

- Canada - Breach of Security Safeguards Regulations
- Canada - British Columbia - Information Privacy & Security - FOIPPA
- Canada - Office of the Superintendent of Financial Institutions Cyber Security Self-Assessment Guide
- Canada - Personal Health Information Protection Act (PHIPA) 2020
- Canada - Personal Information Protection and Electronic Documents Act (PIPEDA)
- Canada - Protected B
- Canada Cybersecure - Baseline Cyber Security Controls for Small and Medium Organizations
- CAN-SPAM Act
- Information Security Management Act - Province of British Columbia, CA
- Mexico - Federal Consumer Protection Law
- Mexico - Federal Law on Protection of Personal Data Held by Private Parties

South America

- Argentina - Personal Data Protection Act 25.326
- Brazil - Consumer Protection Code Law No. 8078 (Office 365)
- Brazil - General Data Protection Law (LGPD)
- Colombia - Decree No. 1377/2013
- Colombia - External Circular Letter 007 of 2018
- Colombia - Law 1266/2008- Habeas Data Act
- Peruvian Legislation Law 29733 Law of Data Privacy Protection

---

Feedback

Was this page helpful? Yes No

Creating regulation templates in Compliance Manager

Compliance Manager helps organizations manage regulatory requirements by providing ready-to-use templates. For organizations with unique compliance needs, you can create custom regulation templates using one of two methods:

- Upload a regulation in PDF format
- Start from a blank template

These options allow flexibility for organizations that need to comply with official standards or define custom internal regulations.

To learn more about regulations, see Learn about regulations in Compliance Manager. To see a list of available regulations, see Compliance Manager regulations list.

How It Works

The process starts when the admin uploads a regulatory PDF file. The system analyzes the document for malware, extracts controls and maps controls to existing improvement actions. Any detected errors are displayed for correction. The admin can then update existing controls, add new ones, and manually map actions to controls. Once all adjustments are complete, the draft template is published for use in Compliance Manager.

Requirements

- A user must hold a role of Compliance Manager Administration or Compliance Manager Assessor. Learn more about roles and permissions.
- For PDF upload:
  - A regulation document published by an official regulator or standards body.
  - PDF format only.

Create a new template based on regulation document (PDF)

1. From your **Regulations** page, select **Create new template**.
2. Select **Upload PDF document**.
3. Modify template name as needed.

4. Optionally, enter overarching regulation name if you want to group multiple templates.
5. Select applicable services.
6. Select **Scan and extract**.

   The system scans, extracts, and maps controls. This process may take several minutes to complete. When the process is complete, you see the results of the scan including malware found, controls extracted, and controls mapped.
7. When the scan is complete, select **Review draft template**.
8. Review the draft template. Add, remove, or edit the controls as needed. When you're satisfied, select **Publish template**.
9. Confirm template details and optionally provide notes, and then select **Publish template**.

The template is published and you can now create an assessment based on your new template.

Create a blank template

1. From your **Regulations** page, select **Create new template**.
2. Select **Create blank template**.
3. Enter a template name.
4. Enter overarching regulation name.
5. Select applicable services.
6. Select **Create draft template**.
7. Manually add controls to your template. Select **Create control**.
8. In the **Create control** dialog, complete each field and then select **Create**.
9. Manually assign improvement actions to your control. Select one of the following:

   - **Import actions** - select from a list of available actions.
   - **Create action** - define your action.
10. Add as many controls and actions as needed. When you're finish, select **Publish template**.
11. Confirm template details and optionally provide notes, and then select **Publish template**.

The template is published and you can now create an assessment based on your new template.

Last updated on 12/04/2025

Learn about the Microsoft Purview Compliance Manager premium assessments trial

07/24/2025

Compliance Manager helps you automatically assess and manage compliance across your multicloud environment. The assessments you build are based on templates that correspond to governmental regulations and industry standards around the world.

Your licensing agreement to use Compliance Manager may include one or more built-in regulatory templates for building assessments. An extensive library of premium regulations is also available for building assessments specific to your organization's needs. To try out these regulatory templates before purchasing licenses, sign up for the free Compliance Manager premium assessments trial.

> **Tip**
>
> If you're not an E5 customer, use the 90-day Microsoft Purview solutions trial to explore how additional Purview capabilities can help your organization manage data security and compliance needs. Start now at the **Microsoft Purview trials hub**. Learn details about **signing up and trial terms**.

Trial setup instructions

- Sign in to the Microsoft Purview portal.
- On the home page, go to the [**Trials and recommendations**](purview-portal.md#trials-and-recommendations section and select **View all trials and recommendations**.
- In the **Purview and Priva trials** section, find the **Compliance assessments** card and select **Try now**.

The premium assessment trial is available to organizations using Compliance Manager under a commercial license. For GCC and DOD information and trial options, see Learn about regulations in Compliance Manager.

Full access to Compliance Manager features depends on your licensing agreement for Compliance Manager. To learn more about the free compliance trial that includes Compliance Manager, see Learn about the Microsoft Purview risk and compliance solutions trial.

> **Note**
>
> Check your Microsoft 365 admin center settings to facilitate enrollment in the trial. In the Microsoft 365 admin center, go to the **Settings** > **Org settings** > **Services** page, and select **User owned apps and services**. Then check the box to let users start trials on behalf of your organization. When setup finishes in about two minutes, you can go back and uncheck the box to disable this setting. You can also start the trial without adjusting your org settings by going to the **Billing** > **Purchase services** page and searching for the Compliance Manager premium assessment trial.

Regulations available to use during the trial

When you start your one-time trial for premium assessments, you can use 25 premium regulatory templates of your choice for free for 90 days. You can create as many assessments from the 25 regulations as you wish. View the full list of Compliance Manager regulations.

After starting your trial, go to your **Assessments** tab in Compliance Manager and view the **Purchased regulation licenses used** counter to see how many regulations are available and are in use. This counter shows your 25 licensed regulatory templates. As you create assessments for these regulations, your activated number updates. Multiple assessments based on the same regulation count as one regulatory template use. For more information, see Learn about regulations in Compliance Manager.

Roles for starting the trial

A user with one of the roles listed below can start or end a trial:

- Billing Administrator
- Compliance Administrator
- Compliance Data Administrator
- Global Administrator

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

Get more details about roles for starting trials.

During your trial

Throughout your trial, you can review the **Trial summary** section of the **Overview** page in Compliance Manager to see how many assessments you created, the number of improvement actions taken, how your compliance score has changed, and how much time remains in your trial. From here, you can also review the list of available regulations and learn more about purchasing them for ongoing use.

After the trial

After your trial ends, the regulatory templates you used won't receive automatic updates and might fall out of date with any regulatory changes. We recommend that you delete assessments built from regulatory templates that you don't intend to purchase.

Ending the trial

Admins can disable the trial anytime by selecting **End trial** on the **Compliance assessments trial** card on the **Microsoft Purview trials and recommendations** page in the Microsoft Purview portal.

Learn more and explore other trials

To learn more about the features of Compliance Manager and how it can help your organization meet compliance and regulatory requirements, see Microsoft Purview Compliance Manager.

You can also learn about other Microsoft Purview solutions and how to start trials at Learn about the Microsoft Purview risk and compliance solutions trial.

Terms and conditions

See the terms and conditions for Microsoft 365 trials.

Working with improvement actions in Compliance Manager

Article • 02/24/2025

> **Tip**
>
> If you're not an E5 customer, use the 90-day Microsoft Purview solutions trial to explore how additional Purview capabilities can help your organization manage data security and compliance needs. Start now at the Microsoft Purview trials hub. Learn details about signing up and trial terms.

Overview

Improvement actions help centralize your compliance activities. Each improvement action recommends an action to take, with detailed guidance intended to help you align with data protection regulations and standards. Improvement actions can be assigned to users in your organization to perform implementation and testing work. You can also store evidence, add notes, and record status updates within the action. Many improvement actions come with automatic testing and monitoring.

Compliance Manager works within a shared responsibility model where some actions are completed by you and others are completed automatically by Microsoft in order to meet the requirements of a regulation, standard, or certification. The actions listed on the Improvement actions tab in Compliance Manager are the ones that your organization is responsible for completing.

> **Tip**
>
> Visit the Compliance Manager glossary for helpful definitions of many of the terms on this page.

Automated testing and monitoring

While some improvement actions must be manually tested by your organization, many actions can be automatically tested and monitored for you. Compliance Manager automatically identifies settings in your Microsoft 365 environment and your multicloud environment that help determine when certain configurations meet improvement action

implementation requirements. Compliance Manager utilizes four types of automation, explained below.

Built-in automation

Compliance Manager has built-in functionality to receive signals from other Microsoft solutions and non-Microsoft services. Compliance Manager detects signals from other Microsoft Purview solutions that your organization may subscribe to, including Data Lifecycle Management, Information Protection, Data Loss Prevention, Communication Compliance, and Insider Risk Management. Compliance Manager also detects signals from Microsoft Priva (this capability is in preview). The automation applies specifically to the solution and isn't scoped to cloud services. Learn more about automatic testing settings.

Microsoft Secure Score automation

Compliance Manager detects signals from complementary improvement actions that are monitored by Microsoft Secure Score. Through these signals, Compliance Manager can automatically test certain improvement actions in order to provide continuous control assessment. When an improvement action is successfully tested and implemented, you receive the maximum possible points for that action, which gets credited to your overall compliance score.

Microsoft Defender for Cloud automation

Integration with Defender for Cloud allows Compliance Manager to facilitate improvement actions and provide continuous monitoring across multiple Microsoft and non-Microsoft cloud services, such as Microsoft Azure, Amazon Web Services (AWS), and Google Cloud Platform (GCP). The cloud infrastructure of this monitoring means that action status can be evaluated and graded at the subscription level of the intended service. You can see specific implementation and testing results for each improvement action within each subscription of your service. The overall score you receive for the improvement action is an aggregate of the individual scores of each subscription within that service. Learn more about multicloud support and scoring.

Connectors

We're rolling out a selection of connectors built specifically for Compliance Manager to support other non-Microsoft services. Connectors for Salesforce and Zoom are available

now, with more connectors releasing soon. Learn more at Working with connectors in Compliance Manager.

> **Note**
>
> An improvement action that is set to automatic testing can be changed to manual testing, unless that action is implemented through Microsoft Defender for Cloud.

Testing type

Compliance Manager provides options for how to test improvement actions. On the improvement action's details page, the **Testing type** status in the top information bar shows how the action is tested. From here you can choose how you want the action to be tested: **Manual** or **Automatic**.

> **Note**
>
> Testing source can't be changed on actions for services supported by Defender for Cloud. If you don't agree with an automated testing result, you can go to the related assessment in Defender for Cloud to alter the testing logic and scope.

Manual

Improvement actions set for manual testing are actions that you manually test and implement. You set the necessary implementation and test status states, and upload any evidence files on the **Evidence** tab. For some actions, this is the only available method for testing improvement actions.

Automatic

Certain improvement actions can be automatically tested by Compliance Manager. Get details on which improvement actions can and can't be tested automatically.

For those improvement actions that can be automatically tested, you'll see the **Automatic** option for **Testing type**. Compliance Manager detects signals from other compliance solutions and cloud services. When signals indicate that an improvement action has been successfully implemented, you automatically receive the eligible points for that action, which factor into scores for related controls and assessments. Learn more about scoring.

Automatic testing is on by default for all eligible improvement actions. You can adjust these settings to automatically test only certain improvement actions, or you can turn off automatic testing for all actions. Learn more about how automated testing works and how to adjust your settings at Set up automated testing.

Improvement actions page

All of the improvement actions managed by your organization are listed on the **Improvement actions** page in Compliance Manager. The list of improvement actions displays basic information related to each action, such as the regulations they correspond to, testing status and source, owner, and action type. See below for descriptions of some of the columns. Filter the columns to customize your view.

Points achieved

Each improvement action is worth a number of points that contribute to your overall compliance score. This column shows how many points have been earned out of the total available to achieve.

Service

The service is the data source, such as Microsoft Azure or Amazon Web Services, that the improvement action assesses and impacts.

Group

This column shows the assessment groups that the improvement action is associated with. Learn more about assessment groups.

Solutions

The solutions are the location where you perform the recommended action; for example, in Microsoft Communication Compliance, Microsoft Priva, or Microsoft Defender for Cloud.

Categories

Compliance Manager groups improvement actions into general categories to help you understand which areas need attention in order to bolster your overall compliance posture. The categories are: Control Access, Discover and Respond, Govern Information,

Infrastructure Cloud, Manage Compliance, Manage Devices, Manage Internal Risks, Privacy Management, Protect Against Threats, and Protect Information.

Action type

Improvement actions can be technical or nontechnical in nature:

- **Technical actions** are implemented by interacting with the technology of a solution (for example, changing a configuration).
- **Nontechnical actions** are managed by your organization and implemented in ways other than working with the technology of a solution. There are two types of nontechnical actions: **documentation** and **operational**.

Learn how action type affects scoring for improvement actions.

Improvement action details page

Select an improvement action from the list on the **Improvement actions** page to view the action's details page. The details page provides detailed implementation guidance and a link to launch you into the appropriate solution or service. Here's what you can do from an improvement actions details page:

- Get basic status quickly from the top information bar.
- Assign an owner, who is responsible for performing the work.
- Review implementation guidance on the **Details** tab. Follow a direct link into the recommended solution to start implementation work and add implementation notes.
- Review the action's status within your cloud services' subscriptions from the **Details** tab and drill into subscription-level details.
- Upload and store evidence files and links on the **Evidence** tab.
- Check the **Related controls** tab to see all the controls that include the improvement action. You can filter the control list by regulation.

Implementation and testing

Assign an owner

To begin implementation work on an improvement action, you can do the work yourself or assign it to another user as an owner. The owner could be a business policy owner, an IT implementer, or another employee with responsibility to perform the task. Once you

identify the appropriate assignee, be sure they hold a sufficient Compliance Manager role to perform the work.

To assign an owner to an improvement action, open the improvement action's details page. In the **Owner** box in the information bar at the top of the page, begin typing a user's name and select the name when it appears. The assigned user receives an email explaining that the improvement action has been assigned to them, with a direct link to the improvement action.

> **Note**
>
> US Government Community (GCC) High and Department of Defense (DoD) customers won't receive an email when improvement actions are assigned to them.

Assign multiple improvement actions to a single user

You can assign multiple improvement actions to one user by following these steps:

1. In Compliance Manager, go to the **Improvement actions** page.
2. Select the checkbox to the left of the improvement actions you want to assign.
3. Select **Assign to user** at the top of the improvement actions table. The **Assign improvement actions** flyout pane appears.
4. Enter a user's name in the search bar to search for the user, or select a name from the list of suggested people. After choosing the user, select **Assign**.

You arrive back at your **Improvement actions** page with the new owner listed in the **Assigned to** column for the actions you assigned.

Implementation work

Implementation guidance varies depending on whether you go to Microsoft Defender for Cloud to perform the work to complete the action. Learn more about multicloud support. Instructions for each action type's implementation are below.

Actions that you implement in Defender for Cloud

On an improvement action's details page, view the **Testing source** on the top information bar. If the testing source is **Defender for Cloud**, then the action pertains to cloud services such as Microsoft Azure, Amazon Web Services (AWS), and Google Cloud Platform (GCP), and the action is implemented and monitored using Compliance Manager's integration with Defender for Cloud. These actions belong to the category of

infrastructure cloud. Whether they can be automatically tested or not depends on the action type:

- Technical actions are automatically tested and monitored.
- Non-technical actions must be manually tested.

View the implementation instructions for each action type below.

Subscription monitoring

For actions implemented in Defender for Cloud, the **Subscription monitoring** section on the **Details** tab shows a list of all related subscriptions (view glossary). This view provides a more granular view into the action's status across each service's subscriptions. Select a subscription name to view a flyout pane with implementation details, virtual resources status, and the related assessments.

Implementation for manual actions:

- On the improvement action details page, review the **How to implement** guidance and perform the necessary steps. This may involve non-technical work that takes place offline.
- Each subscription listed under **Subscription monitoring** needs to have its status updated. Each subscription contains a single virtual resource, which represents the subscription itself. Select the subscription name to bring up a flyout pane for that subscription. Select the appropriate status, date, and enter notes.
- When you're done, select **Save**.

> ⓘ **Note**
>
> Manual actions don't synchronize status between Compliance Manager and Defender for Cloud. You can update the status in either location, however the statuses won't synch.

Implementation for automatic actions:

- Select each subscription, one at a time, listed under **Subscription monitoring**.
- If a subscription shows that there are virtual resources that aren't complete, select the subscription and on the flyout pane, select the **Virtual resources** tab.
- Inspect the status of each resource to determine which require ones require remediation.
- For the resources needing remediation, review the **How to implement** guidance. Then select **Go to Defender for Cloud** to make the necessary changes.

Updates to the improvement action's status show within 24 hours.

Actions that you don't implement in Defender for Cloud

The implementation guidance on the **Details** tab provides instructions and a link into the solution where you go to perform the implementation work. To perform and record your implementation work:

1. Review the **How to implement** guidance and follow the link into the related solution to perform the necessary work. This may involve non-technical work that takes place offline.

2. After performing implementation work, select one of the following status options from the **Implementation status** dropdown menu in the top information bar:

    - **Not implemented**
    - **Implemented**
    - **Alternative implementation** - select this option if you used other non-Microsoft tools or took other actions not included in Microsoft recommendations
    - **Planned**
    - **Out of scope** - not relevant to your organization and doesn't contribute to your score

3. You then need to record the date for the implementation status change. Select **Edit notes** above the action's information bar. The **Edit action details** flyout pane opens to the **Implementation** tab. You can also edit the implementation status here.

4. At the **Implementation date** field, select an implementation date.

5. In the notes field, enter optional implementation notes (learn more about notes).

6. Select **Save** on the flyout pane to close the pane.

7. Select **Save** or **Save and close** above the action's information bar to save the changes you made.

Common actions synchronize across groups. When two different assessments in the same group share improvement actions that are managed by you, any updates you make to an action's implementation details or status will automatically synchronize to the same action in any other assessment in the group. This synchronization allows you to implement one improvement action and meet several requirements across multiple regulations.

Testing work

The test status of an improvement action is displayed in the **Test status** field on the action's details page. After an improvement action has been implemented, you can update the test status for manually tested actions by following the instructions below. A user with editing permissions can adjust the action status.

1. On the improvement action details page, you can select one of the following status options from the **Test status** dropdown menu in the top information bar:

    - **None** - no work has started on the action
    - **Not assessed** - action hasn't been tested
    - **Passed** - implementation has been verified by an assessor
    - **Failed low risk**
    - **Failed medium risk**
    - **Failed high risk**
    - **Out of scope** - the action is out of scope for the assessment and doesn't contribute to your score
    - **To be determined**
    - **Could not be determined**
    - **Partially tested**
    - **In progress**

2. You then need to record the test date. Select **Edit notes** above the action's information bar. The **Edit action details** flyout pane appears.

3. On the **Test and verification** tab, select a test date. You can also edit the test status here.

4. In the notes field, enter optional testing notes (learn more about notes).

5. Select **Save** on the flyout pane to close the pane.

6. Select **Save** or **Save and close** above the action's information bar to save the changes you made.

For manually tested actions that are implemented in Defender for Cloud, refer to the implementation instructions for how to set test status within subscriptions.

> **Note**
>
> If the action is manually tested, you can edit test status, test date, and notes. You can't edit test status and notes for automatically tested actions.

Automatically tested actions may show a status of **Out of scope** when automatic monitoring through Defender for Cloud is first set up. This is because it can take up to 24 hours to process the signals from Defender for Cloud. Improvement action statuses refresh every 24 hours.

Automatically tested actions may also show one of the following test statuses:

- **To be detected** - awaiting signals that indicate test status
- **Could not be detected** - couldn't detect a test status; will be automatically checked again
- **Partially tested** - action has been partially tested; neither passes nor fails

Exporting testing history

You can export a report that shows you a history of all changes in test status for an improvement action. These reports are especially helpful for monitoring progress on actions that are automatically tested, since such actions are regularly or frequently updated based on your tenant's data.

On an improvement action's details page, select **Export testing history** above the top information bar. The report downloads as an Excel file.

Enter and view notes

You can enter notes for internal reference about your implementation or testing work. These notes can be added or edited by any user with editing permissions, not just by the action's assigned owner.

The notes don't display on the improvement action's details page. To view notes, or to add notes, you need to select the **Edit notes** command in the upper right of the improvement action's details page. When you select **Edit notes**, the **Edit action details** flyout pane appears. A text field for notes appears on both the **Implementation** and **Test and verification** tabs of the flyout pane (this is also where you can select status and dates for implementation and testing work.)

After adding notes, select **Save** on the flyout pane. The pane closes. Then select **Save** or **Save and close** above the action's information bar to save your notes.

> **Note**

Testing notes can't be manually edited for automatically tested improvement actions. Compliance Manager updates notes for you.

Storing evidence

You can upload evidence related to implementation and testing work, in the form of files or links, directly to the **Evidence** tab of an improvement action's details page. This environment is a secure, centralized repository to help you demonstrate satisfaction of controls to meet compliance standards and regulations. Any user with read-only access can read content in this section. Only users with editing rights can upload, download, and delete evidence.

Upload evidence

- On the improvement action's details page, go to the **Evidence** tab.
- Select **Add evidence** and select either **Document** or **Link**.
- For a document, browse to select the file you want to upload, then select **Done**. Accepted file types are:
  - Documents (.doc, .xls, .ppt, .txt, .pdf)
  - Images (.jpg, .png)
  - Video (.mkv)
  - Compressed files (.zip, .rar)
- For a link, enter a name for the link and its URL, then select **Done**.

Delete evidence

You can delete evidence when the improvement action hasn't yet reached a pass or fail status. To delete evidence files or links, select the action menu (the three dots) to the right of the item's name and select **Delete**. Confirm the deletion when prompted.

When an improvement action has a test status of **Passed, Failed**, or **Out of scope**, evidence files and links can no longer be deleted.

Assign improvement action to assessor for completion

After you complete the work, conduct testing, and upload evidence, the next step is to assign the improvement action to an assessor for validation. The assessor validates the work, examines the evidence, and selects the appropriate test status.

- **If action status is set to "Pass"**: The action is complete and receives the total number of points.
- **If action status is set to "Fail"**: The action doesn't meet the requirements, and the assessor can assign it back to the appropriate user for more work.

A **Compliance Manager Assessor** role allows a user to perform this work, without the full rights to create assessments. Learn how to set permissions and how to grant role-based access to assessments and regulations.

Accepting updates to improvement actions

When an update is available for an improvement action, you see a notification next to its name. You can either accept the update or defer it for a later time.

What causes an update

An update occurs when there are changes related to scoring, automation, or scope. Changes may involve new guidance for improvement actions based on regulatory changes, or could be because of product changes. Only the improvement actions managed by your organizations receive update notifications.

Where you see assessment update notifications

When an improvement action is updated, you see a **Pending update** label next to its name on the improvement actions page, and on the details page of its related assessments.

Go to the improvement action's details page, and select the **Review update** button in the top banner to review details about the changes and accept or defer the update.

Review update to accept or defer

When you select **Review update** from the improvement action details page, a flyout pane appears on the right side of your screen. The flyout pane provides key details about the update, such as the assessments impacted and changes in score and scope.

Select **Accept update** to accept all the changes to the improvement action. **Accepted changes are permanent**.

> ⓘ Note

When you accept an update to an action, you're also accepting updates to any other versions or instances of this action. Updates will propagate tenant-wide for technical actions, and will propagate group-wide for non-technical actions.

If you select **Cancel**, the update won't be applied to the improvement action. However, you continue to see the **Pending update** notification until you accept the update.

- **Why we recommend accepting updates**: Accepting updates helps ensure you have the most updated guidance on using solutions and taking appropriate improvement actions to help you meet the requirements of the certification at hand.
- **Why you might want to defer an update**: If you're in the middle of completing an assessment that includes the improvement action, you may want to ensure you've finished work on it before you accept the update. You can defer the update for a later time by selecting **Cancel** on the review update flyout pane.

Accept all updates at once

If you have multiple updates and want to accept them all at one time, select the **Accept all updates** link at the top of your improvement actions table. A flyout pane appears which lists the number of actions to be updated. Select the **Accept updates** button to apply all updates.

When you return to your improvement actions page, you may see a message across the top of the page asking you to refresh the page for the updates to be completed.

Set up alerts for improvement action changes

You can set up alerts to notify you immediately when certain changes to improvement actions occur, such as a change in implementation or test status, or an increase or decrease in score. Getting quick notifications of such changes can help you stay on top of possible compliance risks. Visit Compliance Manager alerts and alert policies to learn how to set up alerts.

Export a report of all improvement actions

On the improvement actions page, select **Export actions** above the list of actions. An Excel file downloads. The file contains all your improvement actions and the filter categories shown on the improvement actions page. This is the same file to use for

updating multiple improvement actions at once. Get details about how to work with the file in order to update multiple improvement actions.

Feedback

Was this page helpful? Yes No

Update improvement actions and bring compliance data into Compliance Manager

Overview

Compliance Manager enables organizations to bring their existing compliance activity data and evidence into the Compliance Manager solution. Uploading a specially formatted Excel file allows organizations who are new to Compliance Manager to migrate compliance activities completed in other systems into Compliance Manager and quickly start increasing their overall compliance score.

This upload process also gives new and existing Compliance Manager users greater flexibility and ability to update improvement actions on a larger scale. For example, you can:

- Add test results and evidence to multiple improvement actions that were tested in a system other than Compliance Manager.
- Assign improvement actions to various users based on the actions' score potential.
- Update the implementation status or testing status of multiple improvement actions all at one time.
- Change improvement actions' testing type from automatic to manual.

Get started with updating improvement actions

To migrate existing data into Compliance Manager or to perform a bulk update of improvement actions, you can start from one of two places in Compliance Manager:

- From **Assessments** page: Helps you update specific assessments with information from elsewhere, such as test results or evidence for actions that were tested by a separate system.
- From **Improvement actions** page: Facilitates updating multiple actions at once, such as assigning them to users, changing implementation or test status, and adding notes and evidence.

To begin the process of migrating data or updating actions, follow the steps in Steps for updating actions.

> **Important**

- Only the improvement actions managed by your organization, not Microsoft managed actions, can be updated by this process. (Learn more about types of improvement actions.)
- Improvement actions must already be associated to an assessment before you can update them through this process. (Learn more about building and managing assessments.)

Migrating your existing work into Compliance Manager

If you're new to Compliance Manager, the steps below illustrate the basic workflow for bringing your existing compliance activities into Compliance Manager:

1. **Create an assessment**: Compliance Manager can recommend assessments that may be most relevant to your organization, or you can create one through a guided process. Visit Create assessments for instructions.

2. **Export improvement actions**: You'll export an Excel file containing the action data that you want to update. It may make more sense to start the export from your **Assessments** page, but you can also export from the **Improvement actions** page. See Steps for updating actions.

3. **Update the improvement action Excel file**: Use the instructions on the **How to update actions** tab of the Excel file to add your information.

4. **Upload the Excel file**: Upload your edited Excel file by selecting the **Upload actions** command on the **Assessments** or **Improvement actions** page.

Updating multiple improvement actions at once

To update the status, evidence, notes, or other data in multiple improvement actions at one time, you'll follow the basic flow outlined below:

1. **Export improvement actions.** Export the improvement actions you want to update, staring either from your **Improvement actions** page or your **Assessments** page. The export is a downloaded Excel file containing improvement action data. See Steps for updating actions.

2. **Update the improvement action Excel file.** Use the instructions on the **How to update actions** tab of the Excel file to add or update the information in the specially formatted Excel file.

3. **Upload the Excel file.** Upload your edited Excel file by selecting the **Upload actions** command on the **Assessments** or **Improvement actions** page.

> ⓘ **Note**
>
> If you export the improvement actions from an assessment, the exported Excel file will include control-mapping data for that assessment. However, you won't be able to change the control-mapping data when you edit your Excel file.

Steps for updating actions

The following steps outline the process for bringing compliance activity data into your assessments and updating improvement actions.

1. In Compliance Manager, begin from either from the **Assessments** page or the **Improvement actions** page.

2. Export the improvement actions that you want to update:

   a. From the **Improvement actions** page: Find the improvement actions you want to update, and select the checkbox to the left of their names. Then select the **Export actions** command above the list of actions.

   b. From the **Assessments** page: Find the assessment or assessments you want to update, and select the checkbox to the left of their names. Then select the **Export actions** command above the list of assessments.

3. An Excel file will download, which contains all the data related to the actions. Open the file and refer to the formatting instructions on the tab labeled **How to update actions.**

4. Edit the information on the **Action Update** tab of the spreadsheet according to the formatting instructions. Then save your updated version of the Excel file to your computer.

5. On your **Assessments** page or **Improvement actions** page, select the **Update actions** command, which will open the improvement actions update wizard.

6. The first page in the wizard lists the main prerequisites: that all improvement actions must be associated to at least one assessment, and that your data needs to be formatted in an Excel file for upload. Check the boxes next to the reminders, then continue the process by selecting **Next**.

7. On the **Import updated improvement actions** page, select **Browse** to locate your updated Excel file from its saved location, then select **Next**. If there are any problems with the format of your file, an error message will give instructions for fixing the problem. Upload your corrected file again, then select **Next**.

8. On the **Review and finish** page, review the summary showing the number of actions that will be updated, their associated assessments, and how they affect your compliance score. From here you can upload a different file, or continue with the upload by selecting **Update actions**.

9. When your file has been successfully uploaded, you'll see a confirmation screen. Select **Finish** to exit the wizard and arrive back at the page where you began the update actions process.

Most of the updates will take effect right away, but it may take up to a day for all the updated information to be fully reflected in Compliance Manager.

> **Note**
>
> Control mapping won't be included in the Excel file that's downloaded when you generate a report by selecting *Export actions*.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 07/01/2026

Microsoft Purview Compliance Manager alerts and alert policies

Article • 10/15/2024

> **Tip**
>
> If you're not an E5 customer, use the 90-day Microsoft Purview solutions trial to explore how additional Purview capabilities can help your organization manage data security and compliance needs. Start now at the Microsoft Purview trials hub. Learn details about signing up and trial terms.

Overview

Compliance Manager can alert you to changes as soon as they happen so that you can stay on track with your compliance goals. For example, you can set up alerts to inform you when an improvement action's score value has increased or decreased due to a configuration change in your tenant, or when an improvement action has been assigned to a user to perform implementation or testing work. View the types of events for which you can create alerts.

To create alerts, you first set up an alert policy to outline the conditions that trigger an alert and the frequency of notifications. When we detect a match to your policy conditions, you'll receive an email notification with details so you can determine whether to investigate or take further action.

All alerts are listed on the **Alerts** page in Compliance Manager, and all alert policies are listed on the **Policies** page. All organizations have a default score change policy already set up for them.

Understanding the Policies and Alerts pages

> **Important**
>
> Users must hold the **Security reader** role in Microsoft Entra ID in order to access the **Policies** and **Alerts** pages. Additional security and Compliance Manager roles are needed to work with alerts and alert policies. Get details below in Alert policy permissions.

Policies page

Select **Policies** (or **Alert policies** if using the classic portal) in Compliance Manager to view and manage your alert policies. The **Policies** page contains a table listing all the policies created by your organization. From this page, you can create new policies, edit existing policies, change activation status, and delete policies.

In the **Status column**, **Active** means the policy is in effect and triggering alerts when conditions are met. **Inactive** means the policy exists but isn't generating alerts. The policies table also shows you the severity of the policy and the date the policy was last modified.

To view an individual policy's details, select its row in the table. A flyout pane will appear that shows all details. Select the **Action** button at the bottom of the pane and select from options to edit the policy, view its alerts, or delete it. The commands to add, edit, delete, activate, and disable are also available near the top of the table, above the filters.

To get started creating an alert policy, see Create an alert policy.

Alerts page

Select **Alerts** in Compliance Manager to view and manage your alerts. The **Alerts** page contains a table listing each alert generated by an alert policy, along with its severity and the triggering event (for example, an action's score change) and date of the alert.

To view an individual alert, select its row in the table. A flyout pane will appear that shows all details on the **Overview** tab of the pane. The **Events log** tab displays actions taken by users that triggered the alert.

The **Actions** button at the bottom of the pane provides options to assign the alert to a user for follow-up, email the user whose actions generated the alert, or view the details of the policy that generate the alert. You can also take the same actions by selecting the round button that appears to the left of the alert name when you hover over its row, then selecting one of the buttons near the top of the table, above the filters.

To start working with alerts, see Viewing and managing alerts.

Alert policy permissions

The table below outlines which users can create and edit alerts and alert policies based on their role type. In addition to holding a Compliance Manager role, users also need a Microsoft Entra role as follows:

- To view alerts and alert policies: the **Security reader** role in Microsoft Entra ID.
- To create or update alert policies: the **Compliance Administrator**, **Compliance Data Administrator**, **Security Administrator**, or **Security Operator** role in Microsoft Entra ID.

Learn more about Azure roles in the Microsoft Purview portal.

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

Expand table

Role	Can create and edit policies	Can edit alerts
Compliance Manager Administration	Yes	Yes
Compliance Manager Assessor	Yes	Yes
Compliance Manager Contributor	Yes	Yes
Compliance Manager Reader	No	No
Global Administrator	Yes	Yes

Learn how to set user permissions and assign roles for Compliance Manager.

Create an alert policy

You can create policies to alert you when certain changes or events related to improvement actions happen. Event types are listed below.

Alert event types

- **Score change**: an increase or decrease in points awarded for an improvement action due to configuration changes made by someone in your organization. For example, if your organization creates an insider risk managing policy, that could increase your points for a certain action by a certain amount.
- **Assignment change**: an improvement action has been assigned to a user, reassigned to a different user, or unassigned from a user.

- **Implementation status change**: a user has changed an improvement action's implementation status.
- **Test status change**: a user has changed the testing status of an improvement action.
- **Evidence change**: a user has uploaded or deleted an evidence document in the **Documents** tab of the improvement action.

Default score change policy

Compliance Manager sets up a default alert policy to monitor for score changes in improvement actions. The default policy generates an alert when an improvement action's score changes. Most settings for the default policy can't be edited. But you can add other users as recipients of alert notifications, which we recommend.

Here are the settings for the default policy:

- All matches that are detected within a span of 60 minutes will be grouped into one single alert to reduce excessive notifications. For example, if five improvement actions experience a score change within one hour, one alert is generated.
- The severity level for these alerts is **medium**.
- The Global Admin for your organization is the default recipient of alert notifications.
- You can add more alert recipients by following these steps:
  - On the **Policies** page, find the **Compliance Manager default alert policy**.
  - Check the box to thee left of its name and select the **Edit** button near the top, above the filters.
  - Select the **Next** button until you come to the **Alert recipients** page.
  - Select **+Select recipients** and check the boxes next to each user name on the flyout pane whom you want to receive the email notification. When done, select **Add recipient**, then select **Next**.
  - On the **Review and finish** page, select **Update** to save your changes.
- The default policy can't be deleted, but you can disable it by following the steps outlined below.

Policy creation steps

To create a policy to generate alerts based on one or more events, follow the steps below:

1. In **Compliance Manager**, go to the **Policies** page and select **Add** to start the policy creation wizard.

   > **Tip**
   >
   > In the new Microsoft Purview portal, this page is called **Policies** and is found on the left navigation instead of as a tab at the top.

2. On the **Name and description** page, enter a name for the policy and an optional description, then select **Next**.

3. On the **Conditions** page, select one or more events that will trigger an alert. Under the **Improvement action activity** header, select **Add sub-conditions** and check the box that appears when hovering to the left of each condition name. You can choose one or more conditions for a policy: assignment change, evidence change, implementation status change, score change, test status change. When you're finished, select **Next**.

4. On the **Outcomes** page, choose what happens when a policy match is detected:

   - Select a severity level for the alert when a match is detected: low, medium, or high.
   - Select how often you want to be notified by email when a match is detected. You can choose to be notified with each match, or choose a threshold of a certain number of matches above three.
   - If you choose to be notified after three or more matches, you'll then designate the number of minutes within which that threshold must be reached (for example, 4 matches within 90 minutes).

   When you're done, select **Next**.

5. On the **Alert recipient** page, select additional users in your organization to receive an email when the policy conditions are met. The user who creates the policy is the default recipient. Select **+Select recipients** and check the boxes next to each user name on the flyout pane whom you want to receive the email notification. When done, select **Add recipients**, then select **Next**.

6. Review all selections, and make any changes to each section by selecting **Edit**. When finished reviewing, select **Create policy**.

7. When your policy is created, select **Done**. You arrive at your **Policies** page with the flyout pane for the policy you just created already open.

Your policy is active once you create it, which means it starts detecting matches and generating alerts. See the **Managing policies** section below for how to inactivate or delete policies.

It can take up to 24 hours after creating or updating a policy before alerts are generated by that policy. See View alert details below to learn about triggering events and alert aggregation.

Managing policies

The **Policies** page contains a table listing of all your policies. See Policies page to further understand this page. Any user in your organization can view policies, but certain actions are restricted to certain roles; see Alert policy permissions.

> **Tip**
>
> In the new Microsoft Purview portal, this page is called **Policies** and is found on the left navigation instead of as a tab at the top.

View policy details

Select a policy from its row on the **Policies** page to bring up a flyout panel showing the policy's details, including its match conditions, whether and when alert notifications are sent and to whom, and severity level.

The **Actions** button at the bottom of the panel gives you options to edit the policy, delete the policy, or view alerts.

View a policy's alerts

From the policy's flyout panel, select **Actions** and then **View alerts**. You are taken directly to the Alerts page with a filtered view of all the alerts generated by that policy. Learn how to work with alerts.

Edit a policy

You can edit any aspect of a policy except for its name. If you want to change its name, you need to create a new policy with a new name.

To edit a policy, select the round button that appears to the left of its name when you hover over its row on the **Policies** page and select the **Edit** button near the top, above the filters.

You are taken to the policy creation wizard where you can make and save changes to your policy. You can also select the policy to bring up its details panel, and from the **Actions** button, select **Edit policy**. After working your way through the wizard again, review your selections and in the final step, select **Update** to save your changes.

It can take up to 24 hours before alerts are generated by the updated policy.

Activate or inactivate a policy

Policies are activated by default as soon as they're created. When active, a policy will create an alert (shown on the **Alerts** page) when the conditions are met, and will send a notification email to the designated recipients.

To change a policy to an **inactive** state, which means it won't generate alerts, select the round button that appears to the left of the policy name when you hover over its row. Then select the **Disable** command above the table. The status of your policy will now read Inactive. To reactivate the policy, follow the same process and select the **Activate** button above the filters.

Delete a policy

To delete a policy, you can select the button next to its name on the **Policies** page and select **Delete** near the top of the page. You can also select the policy to bring up its details panel, and from the **Actions** button, select **Delete policy**.

Deleting is permanent. Once you delete a policy, it will no longer generate alerts or email notifications. Learn more about alerts connected to deleted policies.

Viewing and managing alerts

The **Alerts** page shows a table with all the alerts generated by all your policies. Alerts are generated almost immediately after an event matching the policy's conditions occurs. The alert name is the same name as the policy that generated the alert.

An alert can only be generated from an active policy. Once an alert is generated, it remains listed on the **Alerts** page regardless of whether the policy is active or inactive.

Filter your view of alerts

You can filter your view of alerts by selecting the **Filter** command above the table on your **Alerts** page. From the **Filter** flyout pane, select among these filter options:

- Event type
- Severity
- Status
- User assigned to
- Detection date
- Policy name

After making your selections, select **Apply**. The flyout pane closes and your updated **Alerts** page shows your filtered view. Your filters are displayed at the top of the table, though not all filter columns may show in the table.

View alert details

To view all the details about the alert, including the events which triggered it, select its row on the table. A flyout pane shows the details of the alert on the **Overview** tab of the panel.

The **Events log** tab of the flyout panel lists the activities that generated the alert, such as a score change or an assignment change, along with the name of the user associated to each action and the date detected.

Alert events

The **Event** column on the **Alerts** page indicates the conditions of a policy that were detected; in other words, the activity that generated the alert. The **Events log** tab on the alert's details panel lists each instance of an event, the associated user, and the date detected. Event values are listed below:

- **Score change**: shows the number of increase or decrease in points
- **Assignment change**: an improvement action has been assigned to a user, reassigned to a different user, or unassigned from a user
- **Implementation status change**: a user has changed an improvement action's implementation status
- **Test status change**: a user has changed the testing status of an improvement action.
- **Evidence change**: a user has uploaded or deleted an evidence document in the Documents tab of the improvement action
- **Multi-event**: multiple instances of the same type of event have been detected; for example, a single improvement action that has been reassigned multiple times

- **Multi-condition**: multiple conditions within a single policy were detected

Alert aggregation for multiple events within one minute

When multiple events that match the conditions of an alert policy occur with one minute, they're added to an existing alert by a process called alert aggregation.

For example, when one event occurs which matches a policy, an alert is generated and displayed on the **Alerts** page and a notification is sent. If another event matching the same policy occurs within one minute of the first event, then Compliance Manager adds details about the other event on the **Events log** tab of the existing alert instead of triggering a new alert. The goal of alert aggregation is to help reduce alert "fatigue" and let you focus and take action on fewer alerts.

Taking action on alerts

When one of your policies generates an alert, you can view the events that caused the alert and determine whether you need to verify or further investigate the events.

To take an action on an alert, select its row on the **Alerts** page to bring up the flyout panel with its details, select the **Actions** button, and chose among options listed below. You can also take actions by selecting the round button that appears to the left of the alert name when you hover over its row, and selecting one of the action buttons near the top of the page, above the filters.

Assign alert: You may want to assign the alert to a user to investigate or verify the events that caused the alert. When you choose this option, a panel opens where you can select a user in your organization and assign the alert to them. You can filter your alerts view by selecting **Filters** on the **Alerts** page, and entering the user's name in the **Assigned to** field.

Email alert: You may want to send an email to the user associated to the alert's activity to confirm that they took the action. When you chose this option, it opens an email template with basic information about the alert, which you can customize with further instructions and sent to the user.

View policy details: You may want to review the settings for the policy that triggered the alert. When you select this option, you are taken directly to the **Policies** page with the policy details panel already open. You'll no longer be on your **Alerts** page when you close the policy details panel.

Change status: You can update status for your alert based on your review of its impact and whether it needs investigating. Learn more about alert statuses in the next section.

Alert status

When an alert is created, its status is **Active**. As you review the details of each alert, you can update its status to any of the states listed below:

- **Active**: default state of the alert until its status is changed
- **Investigating**: alert is under investigation
- **Resolved**: the alert doesn't require further investigation or follow-up
- **Dismissed**: the alert isn't relevant or doesn't need investigation

To assign or change an alert's status, select an alert from its row on the table, select **Change status** near the top of the page, above the filters. From the Update alert status flyout pane, select a status from the drop-down menu, then select **Update alert**.

Once an alert is generated, its status is independent of the status of the policy that generated the alert. For example, it's possible to have an **active** alert associated to an **inactive** policy, and it's possible to have an **investigating** status on an alert that was generated by a policy that was later inactivated or deleted.

When policies are deleted

When a policy is deleted, any alerts that were generated by that policy remain on your **Alerts** page, but no new alerts are generated.

Email notifications of alerts

When you create a policy, an email is sent to the user who created the policy alerting them that a match was detected. You can choose to send these email notifications to more users in your organization. Alerts occur in near real-time, and the email notifications are sent out as soon as an alert is generated. The email contains the event name, severity, time detected, and a link to view the alert in Compliance Manager.

Remove users from receiving alerts

If you designate alert recipients and then later decide to remove them, follow the steps below. The policy creator will still receive email notifications when policy matches are detected.

1. Begin the steps to edit your policy.
2. When you arrive at the **Alert recipients** screen, select **+Select recipients**.
3. In the **Select recipients** flyout panel, find the user you want to remove from notifications and uncheck the box to the left of their name, then select the **Add**

recipients button (which has the effect of saving your selection).

4. Continue through the wizard and confirm that the user doesn't appear under **Recipients** on the Review and finish page. Select **Update** to save your settings and finish.

Feedback

Was this page helpful? Yes No

Compliance Manager scoring

09/26/2025

> **Important**
>
> Recommendations from Compliance Manager should not be interpreted as a guarantee of compliance. It is up to you to evaluate and validate the effectiveness of customer controls per your regulatory environment. These services are subject to the terms and conditions in Product Terms. See also Microsoft 365 licensing guidance for security and compliance.

Understanding your compliance score

The Compliance Manager dashboard displays your overall compliance score. This score measures your progress in completing recommended improvement actions within controls. Your score can help you understand your current compliance posture. It can also help you prioritize actions based on their potential to reduce risk.

A score value is assigned at these levels:

1. **Improvement action**: Each action has a different impact on your score depending on the potential risk involved. See Action types and scoring below for details.
2. **Assessment**: This score is calculated using improvement action scores. Each Microsoft action and each improvement action managed by your organization is counted once, regardless of how often it's referenced in a control.

The overall compliance score is calculated using improvement action scores, where each Microsoft action is counted once, each technical action you manage is counted once, and each nontechnical action you manage is counted once per group. This logic is designed to provide the most accurate accounting of how actions are implemented and tested in your organization. You may notice that this can cause your overall compliance score to differ from the average of your assessment scores. Read more below about how actions are scored.

Initial score based on Microsoft 365 data protection baseline

Compliance Manager gives you an initial score based on the Microsoft 365 data protection baseline. This baseline is a set of controls that includes key regulations and standards for data protection and general data governance. This baseline draws elements primarily from NIST CSF (National Institute of Standards and Technology Cybersecurity Framework) and ISO

(International Organization for Standardization), as well as from FedRAMP (Federal Risk and Authorization Management Program) and GDPR (General Data Protection Regulation of the European Union).

Your initial score is calculated according to the default Data Protection Baseline assessment provided to all organizations. Upon your first visit, Compliance Manager is already collecting signals from your Microsoft 365 solutions. You see at a glance how your organization is performing relative to key data protection standards and regulations, and see suggested improvement actions to take.

Because every organization has specific needs, Compliance Manager relies on you to set up and manage assessments to help minimize and mitigate risk as comprehensively as possible.

Action types and scoring

Improvement actions have points that are awarded when you complete the requirements for implementation. Your action status is updated on your dashboard within 24 hours of a change being made. Once you follow a recommendation to implement a control, you'll typically see the control status updated the next day.

Points are awarded per action per assessment. For example, if an action is worth 10 points but it appears in two assessments, the action is worth 20 points overall for your tenant. An exception is for technical actions that are scoped to your tenant; points for these actions are granted once per action, regardless of how many groups the action belongs to.

Actions for services supported by Microsoft Defender for Cloud

An improvement action's overall score is based on the average of scores received by its subscriptions. Each subscription is scored based on the status of the relevant virtual resources.

For example, consider an action with two subscriptions, A and B. Subscription A has 0 out of 1 resource completed, and subscription B has 1 out of 2 resources completed. The subscription scores are: A is 0%, B is 50%. The two subscription scores are averaged to get the overall action score of 25%.

How score values are determined

Actions are assigned a score value based on whether they're mandatory or discretionary, and whether they're preventative, detective, or corrective.

Mandatory and discretionary actions

- **Mandatory actions** can't be bypassed, either intentionally or accidentally. An example of a mandatory action is a centrally managed password policy that sets requirements for password length, complexity, and expiration. Users must follow these requirements to access the system.
- **Discretionary actions** rely upon users to understand and adhere to a policy. For example, a policy requiring users to lock their computer when unattended is a discretionary action because it relies on the user.

Preventative, detective, and corrective actions

- **Preventative actions** address specific risks. For example, protecting information at rest using encryption is a preventative action against attacks and breaches. Separation of duties is a preventative action to manage conflict of interest and guard against fraud.
- **Detective actions** actively monitor systems to identify irregular conditions or behaviors that represent risk, or that can be used to detect intrusions or breaches. Examples include system access auditing and privileged administrative actions. Regulatory compliance audits are a type of detective action used to find process issues.
- **Corrective actions** try to keep the adverse effects of a security incident to a minimum, take corrective action to reduce the immediate effect, and reverse the damage if possible. Privacy incident response is a corrective action to limit damage and restore systems to an operational state after a breach.

Each action has an assigned value in Compliance Manager based on the risk it represents:

Expand table

Type	Assigned score
Preventative mandatory	27
Preventative discretionary	9
Detective mandatory	3
Detective discretionary	1
Corrective mandatory	3
Corrective discretionary	1

How is your compliance score calculated?







	Preventative	Detective	Corrective
Mandatory	+27 points	+3 points	+3 points
Discretionary	+9 points	+1 points	+1 points

Assess predeployment compliance with Compliance Manager (preview)

09/26/2025

Compliance Manager provides a way for organizations to understand Microsoft's compliance posture with Microsoft services before deployment. By reviewing the data provided in the **Sub-service compliance readiness** group of regulations, you can analyze in detail how using Azure services can contribute to your organization's overall compliance posture in regard to these four regulations:

- ISO/IEC 27001:2013
- NIST 800-53 rev.4
- PCI DSS v4.0
- System and Organization Controls (SOC) 2

Get started on the Regulations page

In Compliance Manager, go to the **Regulations** page. The first grouping of regulations you see is **Sub-service compliance readiness**. You can filter your view to only these regulations by selecting **Pre-deployment** on the **Availability** filter. A deeper subservice filter is available once you select a regulation to start viewing its actions and controls.

The subservice compliance readiness regulations aren't for creating assessments, and interacting with these regulations doesn't activate them or require a license. When you open them, they display detailed information on the controls and improvement actions managed by you and Microsoft that can contribute to an improved compliance posture using Microsoft services.

To view an individual regulation, select its name from the list to open its details page.

Explore predeployment regulation details

When you select a regulation from the **Regulations** page, a details page opens. Each details page has expandable sections for **All actions** and **All controls.** When you open the regulation's page, the **All actions** section is automatically expanded. Minimize the **All actions** section to more easily get to the **All controls** section underneath.

All actions section

The **All actions** section of the regulation details page provides detailed information about the features supported by the service, and the relevant actions to be taken by your organization and actions managed by Microsoft (learn more about the shared responsibility model).

Your actions tab

The **Your actions** tab has a table with the features supported by the service. Select a name from the **Feature | Action** column to display a details panel along the right side, which you can maximize to full screen. The details screen includes a control description, feature description, and the related Cloud Security Benchmark ID. It also provides prescriptive guidance and the automated technical policies, with deep links to reference content, that you can enable for monitoring compliance. The **Related controls** tab displays related controls from various regulations to help you examine all potentially affected controls simultaneously.

An array of filters allows you to customize your view based on service, subservice, control family, and control ID.

Use the **Sub-service** filter to delve deeper into your service review. This detailed view provides specific information tailored to your selected subservice to enable a more customized compliance review process.

Microsoft actions tab

The **Microsoft actions** tab displays the actions that Microsoft takes care of as the service provider for meeting regulatory requirements. Select a name from the **Feature | Action** column to display a details panel along the right side. You can maximize the panel to full screen, which displays a control description, the Azure control framework ID, and, on the **Implementation** tab, details on the steps taken by Microsoft to fulfill the action. The **Related controls** tab displays related controls from various regulations to help you examine all potentially affected controls simultaneously.

All controls section

The **All controls** section displays a table listing all the controls for the regulation. Under the **Control title** column, select a control to display a details panel along the right side. You can maximize the panel to full screen, which displays details of the regulatory requirements, and tabs that list your actions and Microsoft actions.

Export action and control data

In both the **All actions** and **All controls** sections, there are commands at the top of each table that enable you to export all or selected actions and controls into a CSV or JSON file format.

Select **Export all** to export all items, and select your desired format from the dropdown. To export only selected items, check the box next to the action or control names, then select **Export selected**, then select the desired format.

Get started with Microsoft Service Trust Portal

The Microsoft Service Trust Portal provides a variety of content, tools, and other resources about how Microsoft cloud services protect your data, and how you can manage cloud data security and compliance for your organization.

Accessing the Service Trust Portal

The Service Trust Portal is Microsoft's public site for publishing audit reports and other compliance-related information associated with Microsoft's cloud services. STP users can download audit reports produced by external auditors and gain insight from Microsoft-authored whitepapers that provide details on how Microsoft cloud services protect your data, and how you can manage cloud data security and compliance for your organization. To access some of the resources on the Service Trust Portal, you must sign in as an authenticated user with your Microsoft cloud services account (Microsoft Entra organization account) and review and accept the Microsoft Non-Disclosure Agreement for Compliance Materials.

Existing customers

Existing customers can access the Service Trust Portal at https://aka.ms/STP with one of the following online subscriptions (trial or paid):

- Microsoft 365
- Dynamics 365
- Azure

> **Note**
>
> Microsoft Entra accounts associated with organizations have access to the full range of documents and resources like Compliance Manager.

New customers and customers evaluating Microsoft online services

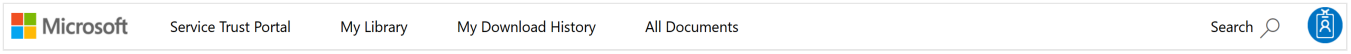
To create a new account or to create a trial account, use one of the following sign-up forms (also used for trial accounts) to get access to the STP.

- Sign up for a new Microsoft 365 Apps for business trial account or a new Office 365 Enterprise trial account
- Sign up for a new Dynamics 365 trial account
- Sign up for a new Azure trial account.

When you sign up for either a free trial, or a subscription, you must enable Microsoft Entra ID to support your access to the STP.

Using the Service Trust Portal

The Service Trust Portal features and content are accessible from the main menu. The following sections describe each item in the main menu.



Service Trust Portal

The **Service Trust Portal** link displays the home page. It provides a quick way to get back to the home page.

Certifications, Regulations, and Standards

Provides a wealth of security implementation and design information with the goal of making it easier for you to meet regulatory compliance objectives by understanding how Microsoft Cloud services keep your data secure. To review content, select one of the following tiles.

- **ISO/IEC** - International Organization for Standardization (ISO) / International Electrotechnical Commission (IEC)
- **SOC** - System and Organization Controls (SOC) 1, 2, and 3 Reports
- **GDPR** - General Data Protection Regulation
- **FedRAMP** - Federal Risk and Authorization Management Program
- **PCI** - Payment Card Industry (PCI) Data Security Standards (DSS)
- **CSA Star** - Cloud Security Alliance (CSA) Security, Trust and Assurance Registry (STAR)
- **Australia IRAP** - Australia Information Security Registered Assessors Program (IRAP)
- **Singapore MTCS** - Multi-Tier Cloud Security (MTCS) Singapore Standard
- **Spain ENS** - Spain Esquema Nacional de Seguridad (ENS)

Reports, Whitepapers, and Artifacts

General documents relating to the following categories:

- **AI Resources**: Resources describing the approach to Compliance, Security, and Privacy in the AI solutions such as Microsoft Copilot and Azure OpenAI.
- **BCP and DR**: Business Continuity and Disaster Recovery.
- **Pen Test and Security Assessments**: Attestation of penetration tests and security assessments conducted by third parties.
- **Privacy and Data Protection**: Privacy and data protection resources.
- **FAQ and Whitepapers**: Whitepapers and answers to frequently asked questions.

Industry and Regional Resources

Documents the apply to the following industries and regions:

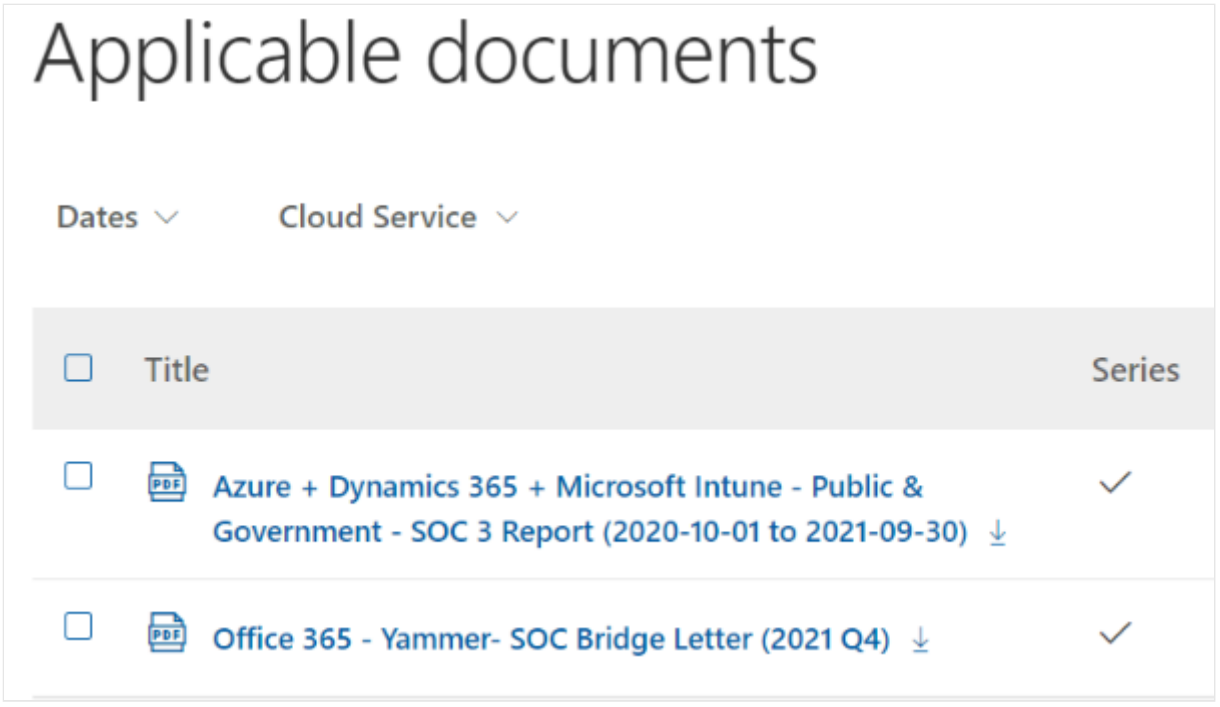
- **Financial Services** - Resources elaborating regulatory compliance guidance for FSI (by country/region)
- **Healthcare and Life Sciences** - Capabilities offered by Microsoft for Healthcare Industry
- **Media and Entertainment** - Media and Entertainment Industry Resources
- **United States Government** - Resources exclusively for US Government customers
- **Regional Resources** - Documents describing compliance of Microsoft's online services with various regional policies and regulations

Resources for your Organization

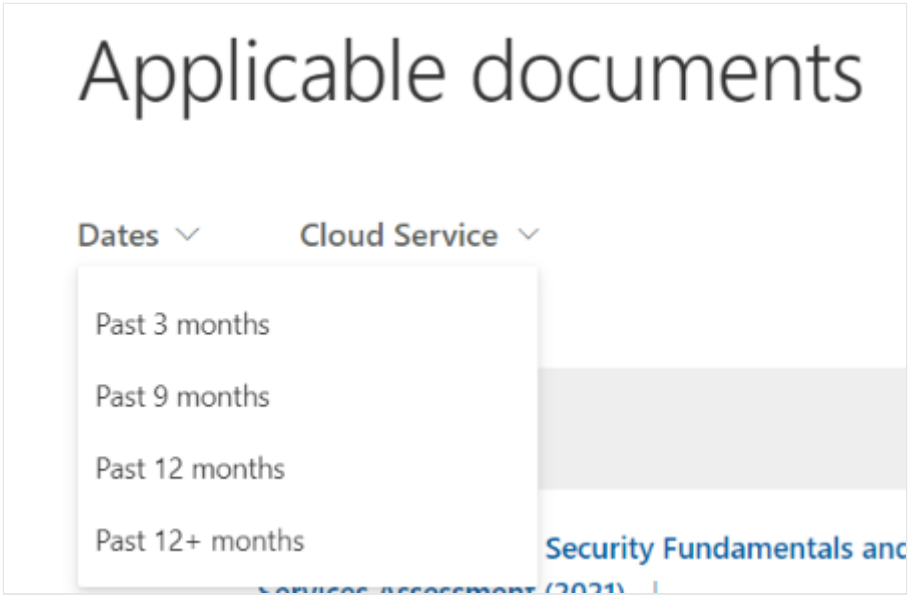
Documents applying to your organization (restricted by tenant).

- **Resources for your Organization** - Documents based on your organization's subscription and permissions

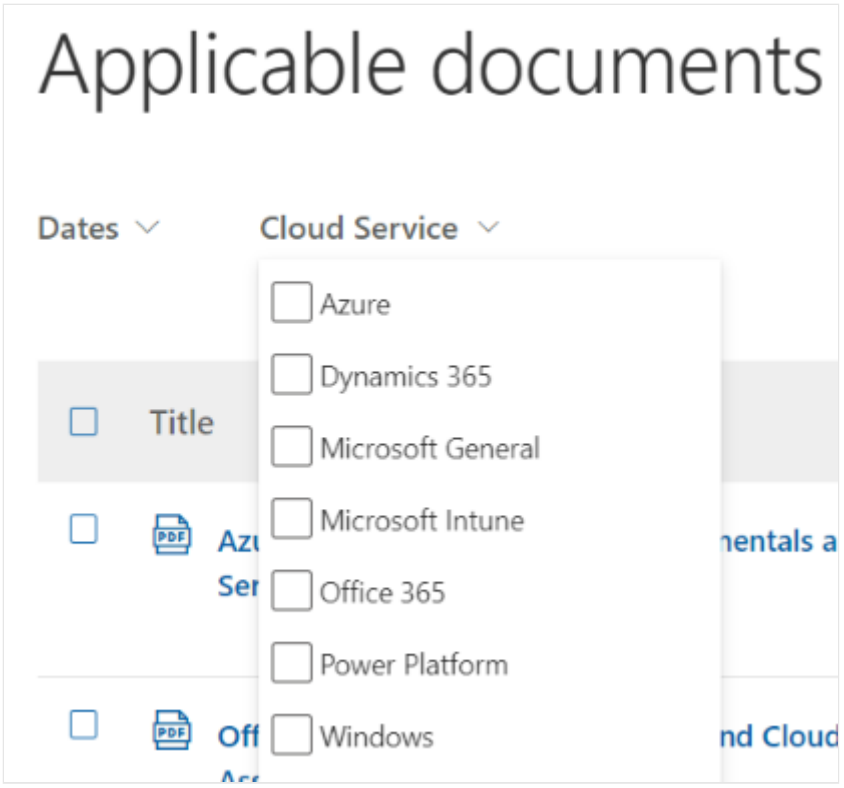
Resources with the series check mark indicate that the document has multiple versions, which can be viewed once you select the document and select **View all versions** on the download page.



Filter by date and cloud service - When viewing the available documents, you can filter the results by date range by selecting **Dates** and then selecting the range you want to use.



Document download view - When viewing the available documents, you can filter the results by the applicable **Cloud Service**.



> **Note**
>
> Many of the files on the STP require acceptance of a license agreement. Some browser-based PDF viewers don't allow JavaScript to run, which prevents the license agreement from being displayed and the file from opening.

All Documents

This section displays all available documents. Select the documents to save into your My Library section. Documents are sorted under the same categories shown under Certifications, Standards, Regulations, and Industry Resources. To view all resources for a particular cloud service use the **Cloud Service** filter.

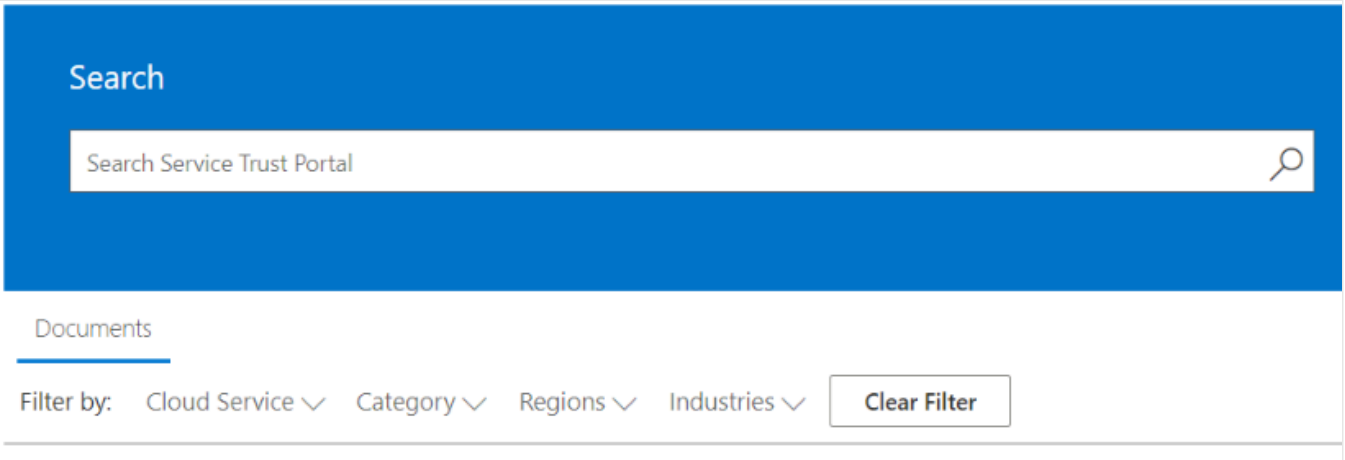
Restricted Documents

The Service Trust Portal has documents that, given the nature of their content, are available for users with specific permissions. You need to be assigned one of following roles to view restricted documents:

- Tenant Admin
- Compliance Administrator
- Security Administrator
- Security Reader

Search

Select the magnifying glass in the upper right-hand corner of the Service Trust Portal page to expand the box, enter your search terms, and press **Enter**. The **Search** page is displayed, with the search term displayed in the search box and the search results listed below.



By default, the search returns document results. You can filter the results by using the dropdown lists to refine the list of documents displayed. You can use multiple filters to narrow the list of documents. Filters include the specific cloud services, and regions. Select the document name link to download the document.

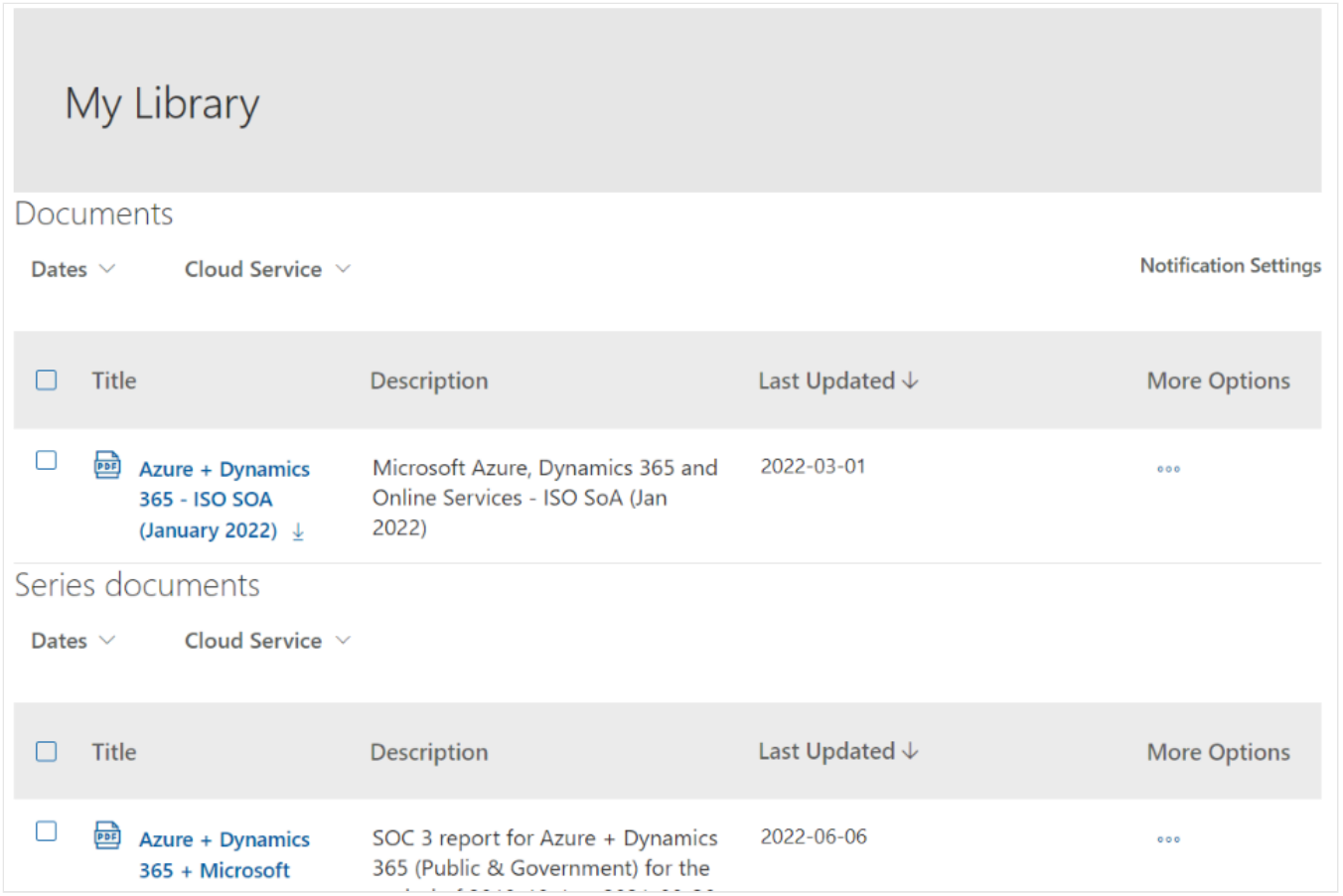
> **Note**
>
> Service Trust Portal reports and documents are available to download for at least 12 months after publishing or until a new version of document becomes available.

My Library

Use the My Library feature to add documents and resources on the Service Trust Portal to your My Library page. This lets you access documents that are relevant to you in a single place. To add a document to your My Library, select the ellipsis (...) menu to the right of a document and then select **Save to library**. You can add multiple documents to your My Library by clicking the checkbox next to one or more documents, and then clicking **Save to library** at the top of the page.

Additionally, the notifications feature lets you configure your My Library so that an email message is sent to you whenever Microsoft updates a document that you've added to your My Library. To set up notifications, go to your My Library and select **Notification Settings**. You can choose the frequency of notifications and specify an email address in your organization to send notifications to. Email notifications include links to the documents that have been updated and a brief description of the update.

If a document is part of a series, you'll be subscribed to the series and receive notifications when there's an update to that series. You can view the individual documents and Series documents that you have subscribed to, in two sections as shown below:



My Download History

On the My Download History tab, you can view and export a download history of documents downloaded from the Service Trust Portal within the last 18 months. The history includes the document title and download date, and the document status, such as whether it's live, has a newer version, or has been deleted. The full download history can be exported to a CSV file.

Localization support

The Service Trust Portal enables you to view the page content in different languages. To change the page language, select the globe icon in the lower left corner of the page and select the language of your choice.

Give feedback

We can help with questions about the Service Trust Portal, or errors you experience when you use the portal. You can also contact us with questions and feedback about Service Trust Portal

compliance reports and trust resources by using the Feedback link on the bottom of the STP pages.

Your feedback is important to us. Select **Feedback** at the bottom of the page to send us comments about what you did or didn't like, or suggestions you have for improving our products or product features.



Last updated on 04/06/2026

Data lifecycle and records management in Microsoft Purview

Summarize this article for me

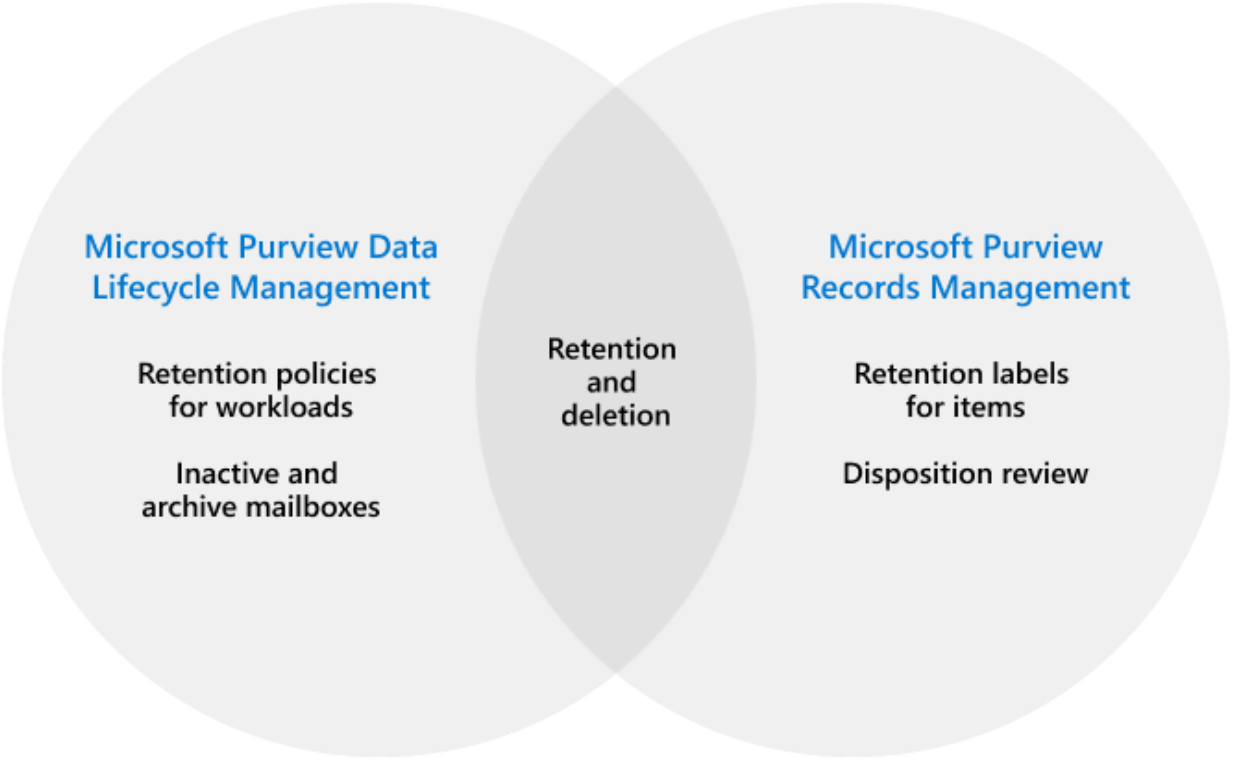
> *Microsoft Purview service description*

Use the **Data Lifecycle Management** and **Records Management** solutions in Microsoft Purview to manage your organization's compliance and regulatory requirements for data & records across their lifecycle.

> **Tip**
>
> Looking to map and manage your data across your entire data estate, including multi-cloud, and software-as-a-service (SaaS)? Use **Microsoft Purview Data Map and Microsoft Purview Unified Catalog**.

From a licensing perspective, there can be considerable overlap between data lifecycle management and records management. Both solutions support retention and deletion of data for Microsoft 365 apps and services.

Use the following graphic to help you identify the main configurable components for these solutions that each have their own configuration area in the Microsoft Purview portal:



The following sections detail the main capabilities for each solution, with links to understand more. However, if you're looking for a guided deployment, see Deploy a data governance solution with Microsoft Purview.

Looking for complementary capabilities to protect your data? See Protect your data with Microsoft Purview.

Microsoft Purview Data Lifecycle Management

Keep what you need and delete what you don't:

Expand table

Capability	What problems does it solve?
Retention policies for Microsoft 365 workloads, with retention labels for exceptions	Lets you retain or delete content with policy management for email, documents, Teams and Viva Engage messages.
Inactive mailboxes	Lets you retain mailbox content after employees leave the organization so that this content remains accessible to administrators, compliance officers, and records managers.
Archive mailboxes	Provides additional mailbox storage space for users.

Capability	What problems does it solve?
Import service for PST files	Supports bulk-importing PST files to Exchange Online mailboxes to retain and search email messages for compliance or regulatory requirements.

Want to learn more? See Learn about data lifecycle management.

Ready to start using some or all of these capabilities? See Get started with data lifecycle management.

Microsoft Purview Records Management

Manage high-value items for business, legal, or regulatory record-keeping requirements:

Expand table

Capability	What problems does it solve?
File plan	Lets you create retention labels interactively or import in bulk, and export for analysis. Labels support additional administrative information (optional) to help you identify and track business or regulatory requirements.
Retention labels for individual items, retention policies if needed for baseline retention	Labels support flexible retention and deletion schedules that can be applied manually or automatically, with records declaration when needed.
Disposition review and proof of disposition	Manual review of content before it's permanently deleted, with proof of disposition of records.

Want to learn more? See Learn about records management.

Ready to start using some or all of these capabilities? See Get started with records management.

Licensing requirements

To understand your licensing requirements and options, see the information from the Microsoft Purview Service Description, Microsoft Purview Data Lifecycle Management & Records Management section for feature-level licensing requirements.

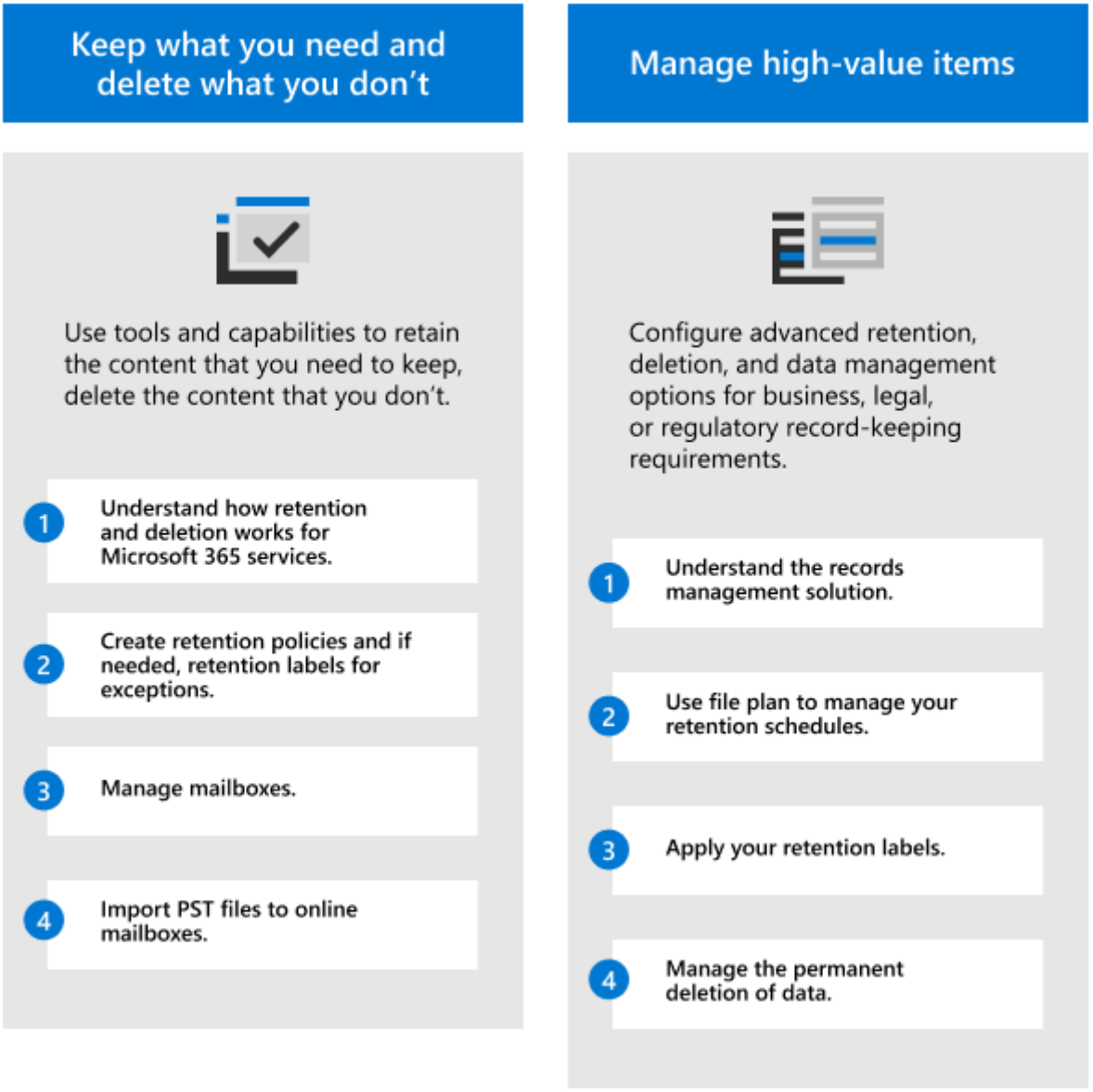
Last updated on 02/27/2026

Deploy Microsoft Purview data lifecycle and records management solutions

Summarize this article for me

> *Microsoft Purview service description*

Use **Microsoft Purview Data Lifecycle Management** and **Microsoft Purview Records Management** to manage compliance or regulatory requirements throughout the lifecycle of your Microsoft 365 data & records.



For data governance that maps and manages data across your data estate, including multi-cloud, and software-as-a-service (SaaS), use Microsoft Purview Data Map and Microsoft Purview Unified Catalog.

For a data protection solution, see Deploy an information protection solution with Microsoft Purview.

Licensing

To understand your licensing requirements and options, see the information from the Microsoft Purview Service Description, Microsoft Purview Data Lifecycle Management & Records Management section for feature-level licensing requirements.

Keep what you need and delete what you don't



Use **Microsoft Purview Data Lifecycle Management** to keep what you need and delete what you don't.

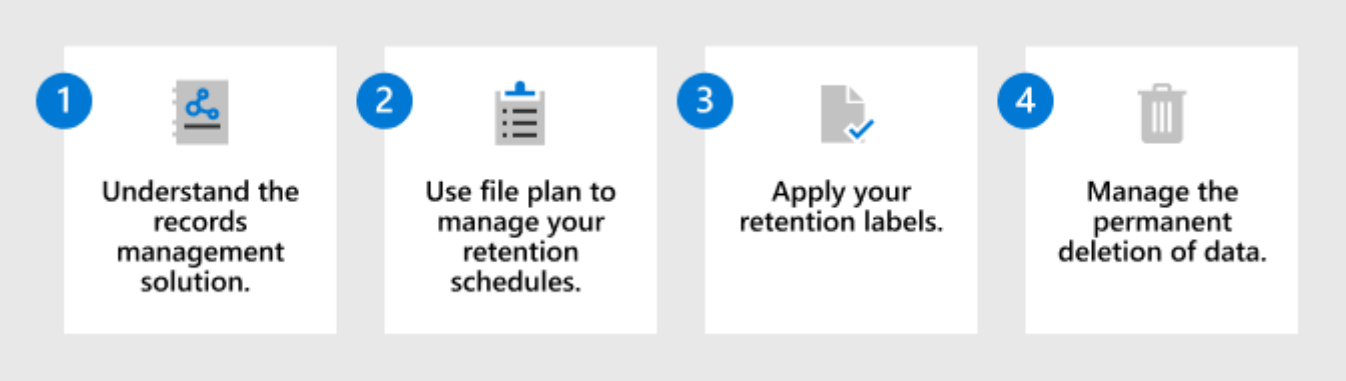
Expand table

Step	Description	More information
1	Understand how retention and deletion works for Microsoft 365 services. After you understand how you can use retention policies and retention labels, identify your workloads that need a retention policy and whether you need to create retention labels for exceptions.	Learn about retention policies and retention labels
2	Create retention policies and if needed, retention labels for exceptions. The most commonly used retention policies are for Exchange, SharePoint, Teams, Microsoft 365 Groups, and Viva Engage. You can configure exceptions for documents and emails, and dynamically apply retention labels for risky users.	Create retention policies Create and apply retention labels for your exceptions Dynamically mitigate the risk of accidental or malicious deletes

Step	Description	More information
	For exceptional circumstances, you might need to expedite the permanent deletion of sensitive information.	Expedite the permanent deletion of sensitive information from mailboxes
3	Manage mailboxes. Enable mailboxes for archiving and auto-expanding archiving, consider whether you need to customize when emails are moved to the archive mailbox, and make mailboxes inactive when users leave the organization.	Enable archive mailboxes Enable auto-expanding archiving Create and manage inactive mailboxes
4	Import PST files to online mailboxes. If you have PST files that contain data you want to oversee, you can import them by using network upload or drive shipping.	Use network upload to import your organization's PST files Use drive shipping to import your organization's PST files

To learn more about the capabilities from this solution, see Learn about data lifecycle management.

Manage high-value items



Use **Microsoft Purview Records Management** to manage your organization's high-value items for business, legal, or regulatory record-keeping requirements.

Expand table

Step	Description	More information
1	Understand the records management solution. Use retention labels with more flexible configuration options and when required, declare items as records.	Learn about records management

Step	Description	More information
2	Use file plan to manage your retention schedules. File plan lets you create retention labels interactively or import in bulk, and export for analysis. Labels that you create with file plan support additional administrative information to help you identify and track business or regulatory requirements.	Use file plan to create and manage retention labels
3	Apply your retention labels. Your retention labels can be published and applied manually or automatically in apps, or automatically applied based on sensitive information, keywords or searchable properties, trainable classifiers, or cloud attachments.	Publish retention labels and apply them in apps Apply a retention label to content automatically
4	Manage the permanent deletion of data. Known as the disposition of data, you can require a manual review of content before it's permanently deleted, and provide proof of disposition for records.	Manage content disposition

> **Tip**
>
> Check the list of **common scenarios** for additional configurations supported by records management.

To learn more about the capabilities from this solution, see Learn about records management.

Training resources

Interactive guides: Microsoft Purview Purview Data Lifecycle Management and Records Management

Learning modules for consultants and admins:

- Introduction to information protection and data lifecycle management in Microsoft Purview
- Manage the data lifecycle in Microsoft Purview
- Manage records in Microsoft Purview

For documentation to support users when these solutions are deployed, see the end user documentation sections for data lifecycle management and records management.

Last updated on 02/27/2026

Learn about retention policies and retention labels

> *Microsoft Purview service description*

> **Note**
>
> If you're seeing messages about retention policies in Teams or have questions about retention labels in your apps, contact your IT department for information about how they have been configured for you. In the meantime, you might find the following articles helpful:
>
> - Teams messages about retention policies
> - Apply retention labels to files in SharePoint or OneDrive
>
> The information on this page is for IT administrators who can create retention policies and retention labels for compliance reasons.

For most organizations, the volume and complexity of their data is increasing daily—email, documents, instant messages, and more. Effectively managing or governing this information is important because you need to:

- **Comply proactively with industry regulations and internal policies** that require you to retain content for a minimum period of time—for example, the Sarbanes-Oxley Act might require you to retain certain types of content for seven years.
- **Reduce your risk in the event of litigation or a security breach** by permanently deleting old content that you're no longer required to keep.
- **Help your organization to share knowledge effectively and be more agile** by ensuring that your users work only with content that's current and relevant to them.

Retention settings that you configure can help you achieve these goals. Managing content commonly requires two actions:

Expand table

Action	Purpose
Retain content	Prevent permanent deletion and remain available for eDiscovery
Delete content	Permanently delete content from your organization

With these two retention actions, you can configure retention settings for the following outcomes:

- Retain-only: Retain content forever or for a specified period of time.
- Delete-only: Permanently delete content after a specified period of time.
- Retain and then delete: Retain content for a specified period of time and then permanently delete it.

These retention settings work with content in place that saves you the additional overheads of creating and configuring additional storage when you need to retain content for compliance reasons. In addition, you don't need to implement customized processes to copy and synchronize this data.

Use the following sections to learn more about how retention policies and retention labels work, when to use them, and how they supplement each other. But if you're ready to get started and deploy retention settings for some common scenarios, see Get started with data lifecycle management.

How retention settings work with content in place

When content has retention settings assigned to it, that content remains in its original location. Most of the time, people continue to work with their documents or mail as if nothing's changed. But if they edit or delete content that's included in the retention policy, a copy of the content is automatically retained.

- For SharePoint and OneDrive sites: The copy is retained in the **Preservation Hold** library.
- For Exchange mailboxes: The copy is retained in the **Recoverable Items** folder.
- For Teams, Viva Engage messages, Copilot and AI apps: The copy is retained in a hidden folder named **SubstrateHolds** as a subfolder in the Exchange **Recoverable Items** folder.

> **Note**

Because the Preservation Hold library is included in the site's storage quota, you might need to increase your storage when you use retention settings for SharePoint, OneDrive, and Microsoft 365 groups.

These secure locations and the retained content aren't visible to most people. In most cases, people don't even need to know that their content is subject to retention settings.

For more detailed information about how retention settings work for different workloads, see the following articles:

- Learn about retention for SharePoint and OneDrive
- Learn about retention for Teams
- Learn about retention for Viva Engage
- Learn about retention for Exchange
- Learn about retention for Copilot

Retention policies and retention labels

To assign your retention settings to content, use **retention policies** and **retention labels with label policies**. You can use just one of these methods, or combine them.

Use a retention policy to assign the same retention settings for content at a site or mailbox level, and use a retention label to assign retention settings at an item level (folder, document, email).

For example, if all documents in a SharePoint site should be retained for five years, it's more efficient to do this with a retention policy than apply the same retention label to all documents in that site. However, if some documents in that site should be retained for five years and others retained for 10 years, a retention policy wouldn't be able to do this. When you need to specify retention settings at the item level, use retention labels.

Unlike retention policies, retention settings from retention labels travel with the content if it's moved to a different location within your Microsoft 365 tenant. In addition, retention labels have the following capabilities that retention policies don't support:

- Options to start the retention period from when the content was labeled or based on an event, in addition to the age of the content or when it was last modified.
- Use trainable classifiers to identify content to label.
- Apply a default label for SharePoint items or Exchange messages.

- Supported actions at the end retention period:
  - Disposition review to review the content before it's permanently deleted.
  - Automatically apply another retention label
- Mark the content as a record as part of the label settings, and always have proof of disposition when content is deleted at the end of its retention period.

Retention policies

Retention policies can be applied to the following locations:

- Exchange mailboxes
- SharePoint classic and communication sites
- OneDrive accounts
- Microsoft 365 Group mailboxes & sites
- Skype for Business
- Exchange public folders
- Teams channel messages (standard channels, shared channels, and private channels post-migration)
- Teams chats
- Teams private channel messages (pre-migration only)
- Teams call logs
- Microsoft Copilot experiences
- Enterprise AI apps
- Other AI apps
- Viva Engage community messages
- Viva Engage user messages

> **Note**
>
> If you have existing retention policies for Teams chats and Copilot interactions, they continue to be supported, although they can't be edited when your tenant supports the separate locations. At this point, any new retention policies must use the new locations.

You can efficiently apply a single policy to multiple locations, or to specific locations or users.

For the start of the retention period, you can choose when the content was created or, supported only for files and the SharePoint, OneDrive, and Microsoft 365 Groups locations, when the content was last modified.

Items inherit the retention settings from their container specified in the retention policy. If they are, then moved outside that container when the policy is configured to retain content, a copy of that item is retained in the workload's secured location. However, the retention settings don't travel with the content in its new location. If that's required, use retention labels instead of retention policies.

Retention labels

Use retention labels for different types of content that require different retention settings. For example:

- Tax forms that need to be retained for a minimum period of time.
- Press materials that need to be permanently deleted when they reach a specific age.
- Competitive research that needs to be retained for a specific period and then permanently deleted.
- Work visas that must be marked as a record so that they can't be edited or deleted.

In all these cases, retention labels let you apply retention settings for governance control at the item level (document or email).

With retention labels, you can:

- **Enable people in your organization to apply a retention label manually** to content in Outlook and Outlook on the web, OneDrive, SharePoint, and Microsoft 365 groups. Users often know best what type of content they're working with, so they can classify it and have the appropriate retention settings applied.
- **Apply retention labels to content automatically** if it matches specific conditions, that includes cloud attachments that are shared in email or Teams, or when the content contains:
  - Specific types of sensitive information.
  - Specific keywords that match a query you create.
  - Pattern matches for a trainable classifier.
- **Start the retention period from when the content was labeled** for documents in SharePoint sites and OneDrive accounts, and for email items.
- **Start the retention period when an event occurs**, such as employees leave the organization, or contracts expire.

- **Apply a default retention label to a document library, folder, or document set** in SharePoint, so that all documents that are stored in that location inherit the default retention label.

- **Mark items as a record** as part of your records management strategy. When this labeled content remains in Microsoft 365, further restrictions are placed on the content that might be needed for regulatory reasons. For more information, see Compare restrictions for what actions are allowed or blocked.

Retention labels, unlike sensitivity labels, don't persist if the content is moved outside Microsoft 365.

Dynamically mitigate the risk of accidental or malicious deletes

In preview, you can use this solution with Insider Risk Management so that retention labels are automatically applied with Adaptive Protection.

When you enable Adaptive Protection for your tenant, retention labels are automatically applied to unlabeled content if it's deleted by users who have been identified as an elevated risk. If these users delete content from SharePoint, OneDrive, or Exchange, a retention label is automatically applied to that content to retain it for 120 days. As a result, that content remains accessible for search and eDiscovery from the secured locations used by the workload.

> **Note**
>
> If you enabled and configured Adaptive Protection before this integration with data lifecycle management released, you'll need to opt-in to create this auto-labeling policy. See the instructions at the end of this section.

When these items are retained with Adaptive Protection, the following auditing events are generated and identify the user and item:

- For SharePoint and OneDrive: **Retained file proactively**
- For Exchange: **Retained email item proactively**

After the 120 days expire, the items then become eligible for permanent deletion. To learn more about when permanent deletion occurs, see How retention works for SharePoint and OneDrive and How retention works for Exchange.

Unlike other labeling scenarios, users don't see the retention label, and you don't need to create or manage the retention label or auto-labeling retention policy. At this time, you can't change the retention period or assign different policies based on the different risk levels, or for different locations. The single retention label and auto-labeling retention policy for your tenant aren't visible in the Microsoft Purview portal.

If you're using Adaptive Protection but don't want to automatically retain content in this way, you can turn off the auto-labeling policy without affecting other Adaptive Protection policies. Use the same control if you need to turn on the auto-labeling retention policy for Adaptive Protection, and confirm the status.

1. Sign in to the Microsoft Purview portal > **Solutions** > **Settings** > **Solution settings**> **Data lifecycle management** > **Adaptive protection**.

2. For **Adaptive protection in Data Lifecycle Management**, turn the setting off, confirm your choice, and select **Save**.

   Any retention labels that were applied as a result of Adaptive Protection are removed so that the items can then become eligible for permanent deletion.

You won't be able to turn on this setting unless Adaptive Protection is turned on for your tenant. If your account has the required permissions, you'll see an option to take you to the insider risk management solution where you can turn on and configure Adaptive Protection.

Override holds to reclaim disk space or permanently delete sensitive information

Also like adaptive protection, priority cleanup for files or mailbox items similarly applies retention labels under the covers. The labels are automatically configured and applied when you create a priority cleanup policy that identifies items with a query that you specify. This policy can override existing holds for retention and eDiscovery.

Priority cleanup is specific to data lifecycle management and isn't supported for labeled items that are marked as records. Typical uses depend on the workload:

- For SharePoint and OneDrive, you can use priority cleanup to periodically delete large files, such as Teams meeting recordings and transcripts. These files are regularly created if you're using the popular Copilot in Teams recap feature, and might be automatically retained with a retention policy that retains all items for many years. However, these specific files typically have little business value after 1-3 months, and you want to regularly delete them to save

disk space. You can also use priroty cleanup to delete files in the Preservation Hold library that can prevent OneDrive accounts from being deleted after a user leaves the organization.

- For Exchange, you can use priority cleanup to delete items when this action is required for security or privacy. For example, after a data spillage incident.

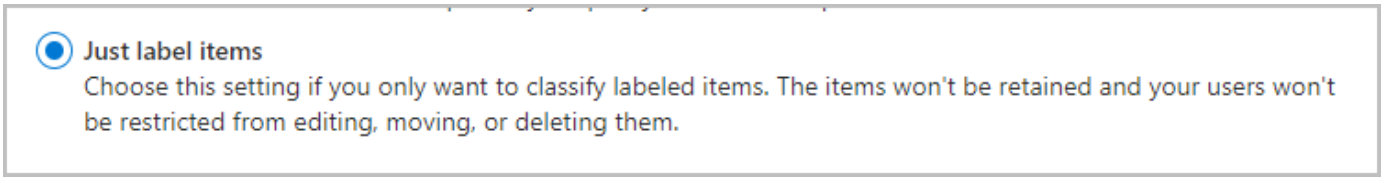
For more information and configuration instructions, see the following articles:

- Override holds to clean up files for Copilot and reclaim storage
- Expedite the permanent deletion of sensitive information from mailboxes.

Classifying content without applying any actions

Although the main purpose of retention labels is to retain or delete content, you can also use retention labels without turning on any retention or other actions. In this case, you can use a retention label simply as a text label, without enforcing any actions.

For example, you can create and apply a retention label named "Review later" with no actions, and then use that label to find that content later.



Using a retention label as a condition in a DLP policy

You can specify a retention label as a condition in a Microsoft Purview Data Loss Prevention (DLP) policy for documents in SharePoint. For example, configure a DLP policy to prevent documents from being shared outside the organization if they have a specified retention label applied to it.

For more information, see Create and Deploy data loss prevention policies.

Retention labels and policies that apply them

When you publish retention labels, they're included in a **retention label policy** that makes them available for admins and users to apply to content. As the following diagram shows:

1. A single retention label can be included in multiple retention label policies.

2. Retention label policies specify the locations to publish the retention labels. A single label retention policy can include multiple locations.



You can also create one or more **auto-apply retention label policies**, each with a single retention label. With this policy, a retention label is automatically applied when conditions that you specify in the policy are met.

Retention label policies and locations

Retention labels can be published to different locations, depending on what the retention label does.

Expand table

If the retention label is...	Then the label policy can be applied to...
Published to admins and end users	Exchange, SharePoint, OneDrive, Microsoft 365 Groups
Auto-applied based on sensitive information types, keywords or a query, or trainable classifiers	Exchange, SharePoint, OneDrive, Microsoft 365 Groups
Auto-applied to cloud attachments	SharePoint, OneDrive, Microsoft 365 Groups

Exchange public folders, Skype, Teams and Viva Engage messages don't support retention labels. To retain and delete content from these locations, use retention policies instead.

Only one retention label at a time

Unlike sensitivity labels, you can't configure priorities for retention labels. Use the following information to understand label behavior for retention labels.

As with sensitivity labels, an item such as an email or document can have only a single retention label applied to it at a time. A retention label can be applied manually by an end user or admin, or automatically by using any of the following methods:

- Auto-apply retention label policy
- A Microsoft Syntex model
- Default retention label for SharePoint or Outlook
- Outlook rules
- Power Automate compliance action of **Apply a retention label on the item**

If there are multiple auto-apply retention label policies that could apply a retention label, and the content meets the conditions of more than one of these policies, you can't control which retention label will be selected. However, in some cases, the retention label for the oldest auto-apply retention label policy (by date created) is selected. This happens only when the matching policies don't include multiple instances of the same type of condition (sensitive information types, specific keywords or searchable properties, or trainable classifiers).

For standard retention labels (they don't mark items as a record or regulatory record):

- Admins and end users can manually change or remove an existing retention label that's applied on content.
- When items already have a retention label applied, the existing label won't be automatically removed or replaced by another retention label with the following exceptions:
    - At the end of the retention period, the existing label is configured to automatically apply a different retention label, or the existing label is configured to run a Power Automate flow with the compliance action of **Relabel an item at the end of retention**.
    - You use the Power Automate compliance action of **Apply a retention label on the item**. If the item already has a retention label applied, it will be replaced.
    - The existing label was applied as a default label. When you use a default label, there are some scenarios when it can be replaced by another default label, or automatically removed. For more information, see Default labels for SharePoint and Outlook.

For retention labels that mark items as a record or a regulatory record:

- These retention labels are never automatically changed during their configured retention period, even if the existing label was applied as a default label, or identified for priority cleanup.
- Only admins for the container can manually change or remove retention labels that mark items as a record, but can't manually change or remove retention labels that mark items as a regulatory record. For more information, see Compare restrictions for what actions are allowed or blocked.
- At the end of the retention period, an existing label can be replaced if it's configured to mark items as a record and automatically apply a different retention label or to run a Power Automate flow with the compliance action of **Relabel an item at the end of retention**. You can't use these relabeling methods if the existing label is configured to mark items as a regulatory record.

Will an existing label be overridden or removed?

Use the following tables to help you quickly identify whether an existing retention label on items can be overridden by another retention label, or removed so that it's no longer labeled.

> **Note**
>
> Unless a labeled item is marked as a record or regulatory record, it can always be overridden by priority cleanup that under the covers, applies a retention label to delete the item.

A standard retention label refers to a retention label that isn't configured to mark items as records or regulatory records.

Will a label be overridden?

Expand table

New label application method	Standard retention label	Marks items as records	Marks items as regulatory records
Manually applied	Yes	Yes [1] if admin for the container	No
Applied with Power Automate actions	Yes	Yes [1]	Not applicable

New label application method	Standard retention label	Marks items as records	Marks items as regulatory records
Applied with the **Change label** label setting	Yes	Yes	Not applicable
Applied with the **Relabel** disposition review action	Yes	Yes	No
Applied with auto-apply retention label policy	No	No	Not applicable
Applied with Microsoft Syntex model	No	No	No
Outlook rules	No	No	No
Inherited from default label for SharePoint	Yes if originally applied by another default label [2]	No	No
Inherited from default label for Outlook	Yes if originally applied by another default label	No	No

Footnotes:

[1] The record must be locked.

[2] An exception is if you move the item to another location with a different default label, then the original label isn't overwritten. Only if you then change the default label for this new location will the original default label be overwritten.

Monitoring retention labels

Use the Microsoft Purview portal to monitor how retention labels are being used in your tenant, and identify where your labeled items are located:

- Sign in to the Microsoft Purview portal > **Solutions** > **Data Lifecycle Management** > **Overview**

  You can then drill down to see details by selecting **View details** that loads content explorer and activity explorer.

For more information, including important prerequisites, see Learn about data classification.

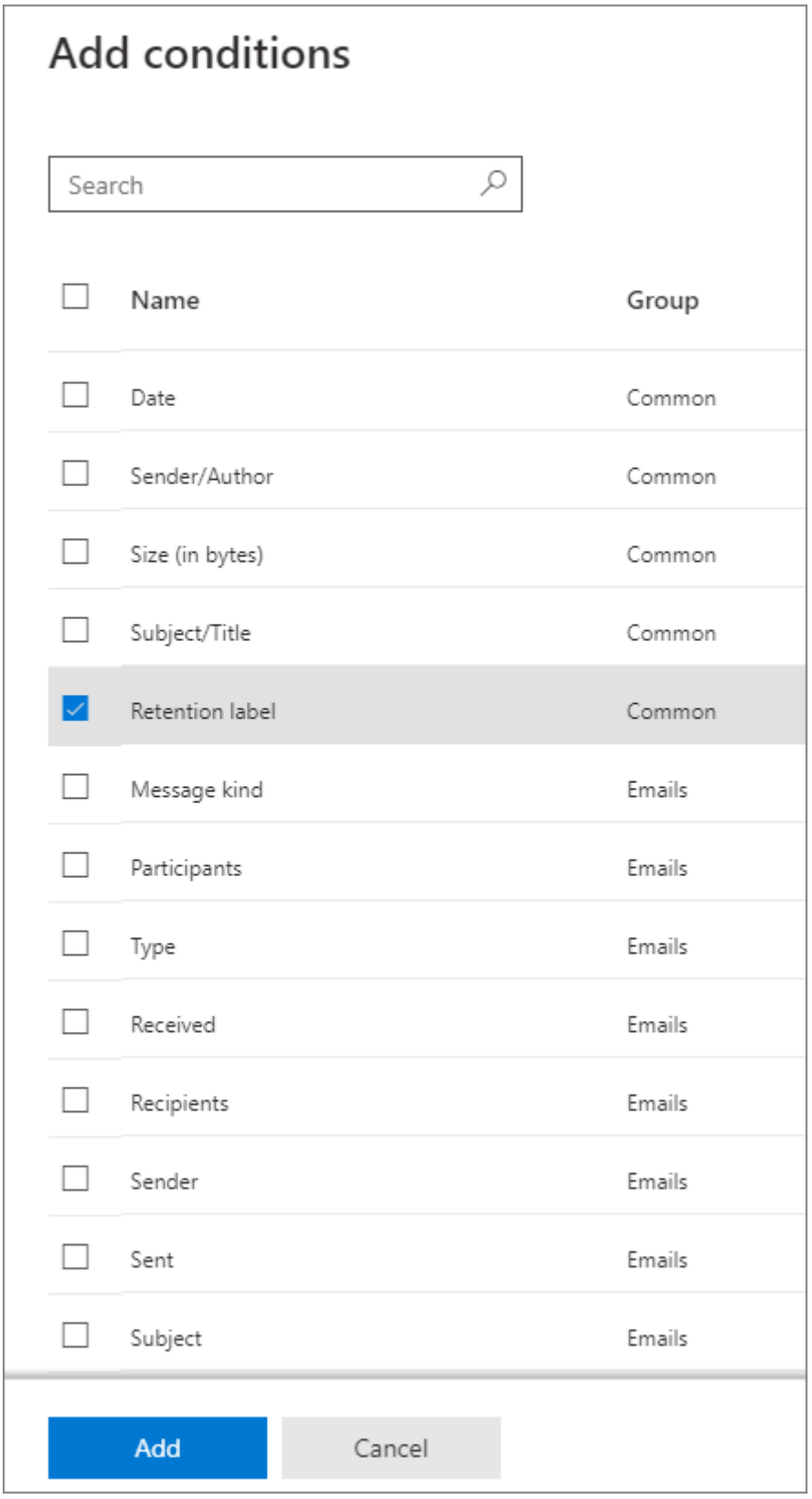
> **Tip**

Consider using some of the other data classification insights, such as trainable classifiers and sensitive info types, to help you identify content that you might need to retain or delete, or manage as records.

Using Content Search to find all content with a specific retention label

After retention labels are applied to content, either by users or auto-applied, you can use content search to find all items that have a specific retention label applied.

When you create a content search, choose the **Retention label** condition, and then enter the complete retention label name or part of the label name and use a wildcard. For more information, see Keyword queries and search conditions for Content Search.



When the retention period starts

The start of the retention period depends on how you configure the retention policy or retention label:

- **When the content was created.** The default start of the retention period.

- **When the content was last modified.** Supported only for files in the SharePoint, OneDrive, and Microsoft 365 Groups locations.
- **When the content was labeled.** Available with retention labels, for documents in SharePoint and OneDrive, and for email items.
- **When an event occurs.** Available with retention labels that are configured for event-based retention, such as when employees leave the organization or contracts expire.

Compare capabilities for retention policies and retention labels

Use the following table to help you identify whether to use a retention policy or retention label, based on capabilities.

Expand table

Capability	Retention policy	Retention label
Retention settings that can retain and then delete, retain-only, or delete-only	Yes	Yes
Workloads supported: - Exchange - SharePoint - OneDrive - Microsoft 365 groups - Skype for Business - Teams - Copilot and AI apps - Viva Engage	 Yes Yes Yes Yes Yes Yes Yes	 Yes, except public folders Yes Yes Yes No No No
Retention applied automatically	Yes	Yes
Automatically apply different retention settings at the end of the retention period	No	Yes
Retention applied based on conditions - sensitive info types, KeyQL queries and keywords, trainable classifiers, cloud attachments	No	Yes
Retention settings applied manually	No	Yes
End-user interaction	No	Yes
Persists if the content is moved	No	Yes, within your Microsoft 365 tenant

Capability	Retention policy	Retention label
Declare item as a record	No	Yes
Start the retention period when labeled or based on an event	No	Yes
Run a Power Automate flow at the end of the retention period	No	Yes
Disposition review	No	Yes
Proof of disposition for up to seven years	No	Yes, when you use disposition review or item is marked as a record
Audit admin activities	Yes	Yes
Audit retention actions	No	Yes *
Identify items subject to retention: - Content Search - Data classification page, content explorer, activity explorer	 No No	 Yes Yes

Footnote:

* For retention labels that don't mark the content as a record or regulatory record, auditing events are limited to when an item in SharePoint or OneDrive has a label applied, changed, or removed. Or, when a retention label is used with priority-cleanup. For auditing details for retention labels, see the Auditing retention actions section on this page.

Combining retention policies and retention labels

You don't have to choose between using retention policies only or retention labels only. Both methods can be used together and in fact, complementary each other for a more comprehensive solution.

The following examples are just some of the ways in which you can combine retention policies and retention labels for the same location.

For more information about how retention policies and retention labels work together and how to determine their combined outcome, see the section on this page that explains the principles of retention and what takes precedence.

Example for users to override automatic deletion

Scenario: By default, content in users' OneDrive accounts is automatically deleted after five years but users must have the option to override this for specific documents.

1. You create and configure a retention policy that automatically deletes content five years after it's last modified, and apply the policy to all OneDrive accounts.

2. You create and configure a retention label that keeps content forever and add this to a label policy that you publish to all OneDrive accounts. You explain to users how to manually apply this label to specific documents that should be excluded from automatic deletion if not modified after five years.

Example to retain items for longer

Scenario: By default, SharePoint items are automatically retained and then deleted after five years, but documents in specific libraries must be retained for 10 years.

1. You create and configure a retention policy that automatically retains and then deletes content after five years, and apply the policy to all SharePoint and Microsoft 365 Groups instances.

2. You create and configure a retention label that automatically retains content for 10 years. You add this label to a label policy that you publish to all SharePoint and Microsoft 365 Groups instances so that SharePoint admins can then apply it as a default label to be inherited by all items in specific document libraries.

Example to delete items in a shorter time period

Scenario: By default, emails aren't retained but are automatically deleted after 10 years. However, emails related to a specific project that has a prerelease code name must be automatically deleted after one year.

1. You create and configure a retention policy that automatically deletes content after 10 years, and apply the policy to all Exchange recipients.

2. You create and configure a retention label that automatically deletes content after one year. Options for applying this label to relevant emails include:

    - You create an auto-labeling policy that identifies content by using the project code name as the keyword, and apply the policy to all Exchange recipients
    - You publish the label and instruct users involved in the project how to create an automatic rule in Outlook that applies this label

- You publish the label and instruct users to create a folder in Outlook for all emails related to the project and they apply the published label to the folder, and then create an Outlook rule to move all project-related emails to this folder

How long it takes for retention settings to apply

When you submit retention policies for workloads and label policies to automatically apply a retention label, allow up to seven days for the retention settings to be applied to content:

- How long it takes for retention policies to take effect
- How long it takes for retention labels to take effect

Similarly, allow up to seven days for retention labels to be visible in apps after you publish the labels:

- When retention labels become available to apply

Often, the policies take effect and labels are visible quicker than seven days. But with many potential variables that can impact this process, it's best to plan for the maximum of seven days.

Adaptive or static policy scopes for retention

When you create a retention policy or retention label policy, you must choose between adaptive and static to define the scope of the policy.

- An **adaptive scope** uses a query that you specify, so the membership isn't static but dynamic by running daily against the attributes or properties that you specify for the selected locations. You can use multiple adaptive scopes with a single policy.

  Example: Emails and OneDrive documents for executives require a longer retention period than standard users. You create a retention policy with an adaptive scope that uses the Microsoft Entra attribute job title of "Executive," and then select the Exchange email and OneDrive accounts locations for the policy. There's no need to specify email addresses or OneDrive URLs for these users because the adaptive scope automatically retrieves these values. For new executives, there's no need to reconfigure the retention policy because these new users with their corresponding values for email and OneDrive are automatically picked up.

- A **static scope** doesn't use queries and is limited in configuration in that it can apply to all instances for a specified location, or use inclusion and exclusions for specific instances for

that location. These three choices are sometimes referred to as "org-wide," "includes," and "excludes" respectively.

Example: Emails and OneDrive documents for executives require a longer retention period than standard users. You create a retention policy with a static scope that selects the Exchange email and OneDrive accounts locations for the policy. For the Exchange email location, you're able to identify a group that contains just the executives, so you specify this group for the retention policy, and the group membership with the respective email addresses is retrieved when the policy is created. For the OneDrive accounts location, you must identify and then specify individual OneDrive URLs for each executive. For new executives, you must reconfigure the retention policy to add the new email addresses and OneDrive URLs. You must also update the OneDrive URLs anytime there is a change in an executive's UPN.

OneDrive URLs are particularly challenging to reliably specify because by default, these URLs aren't created until the user accesses their OneDrive for the first time. And if a user's UPN changes, which you might not know about, their OneDrive URL automatically changes.

Advantages of using adaptive scopes over static scopes:

- No limits on the number of items per policy. Although adaptive policies are still subject to the maximum number of policies per tenant limitations, the more flexible configuration will likely result in far fewer policies.
- You can apply specific retention settings to just inactive mailboxes. This configuration isn't possible with a static scope because at the time the policy is assigned, static scopes don't support the specific inclusion of recipients with inactive mailboxes.

For more advantages of using adaptive scopes, see Adaptive scopes.

Advantages of using static scopes over adaptive scopes:

- Simpler configuration if you want all instances automatically selected for a workload.

  For "includes" and "excludes," this choice can be a simpler configuration initially if the numbers of instances that you have to specify are low and don't change. However, when these number of instances start to increase and you have frequent changes in your organization that require you to reconfigure your policies, adaptive scopes can be simpler to configure and easier to maintain.

- The **Skype for Business** and **Exchange public folders** locations don't support adaptive scopes. For those locations, you must use a static scope.

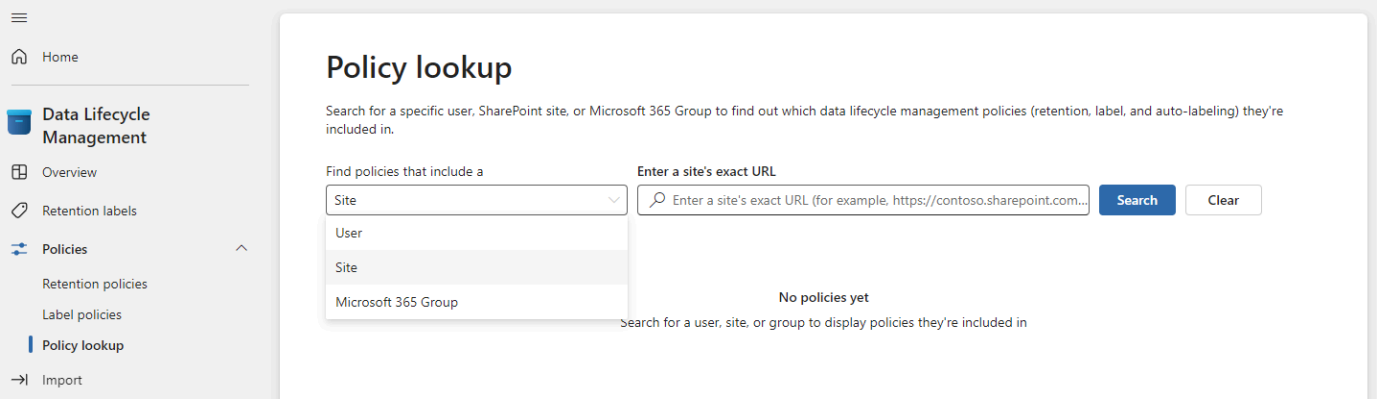
For configuration information, see Configuring adaptive scopes.

Currently, adaptive scopes don't support Preservation Lock to restrict changes to retention policies and retention label policies.

Policy lookup

You can configure multiple retention policies for Microsoft 365 locations, as well as multiple retention label policies that you publish or auto-apply. To find the policies for retention that are assigned to specific users, sites, and Microsoft 365 groups, use **Policy lookup** from the **Data lifecycle management** or **Records management** solutions in the Microsoft Purview portal.

For example, from the Microsoft Purview portal:



You must specify the exact email address for a user, exact URL for a site, or exact email address for a Microsoft 365 group. You can't use wildcards, or partial matches, for example.

The option for sites includes OneDrive accounts. For information how to specify the URL for a user's OneDrive account, see Get a list of all user OneDrive URLs in your organization.

The principles of retention, or what takes precedence?

Unlike retention labels, you can apply more than one retention policy to the same content. Each retention policy can result in a retain action and a delete action. Additionally, that item could also be subject to these actions from a retention label.

In this scenario, when items can be subject to multiple retention settings that could conflict with one another, what takes precedence to determine the outcome?

The outcome isn't which single retention policy or single retention label wins, but how long an item is retained (if applicable) and when an item is deleted (if applicable). These two actions are calculated independently from each other, from all the retention settings applied to an item.

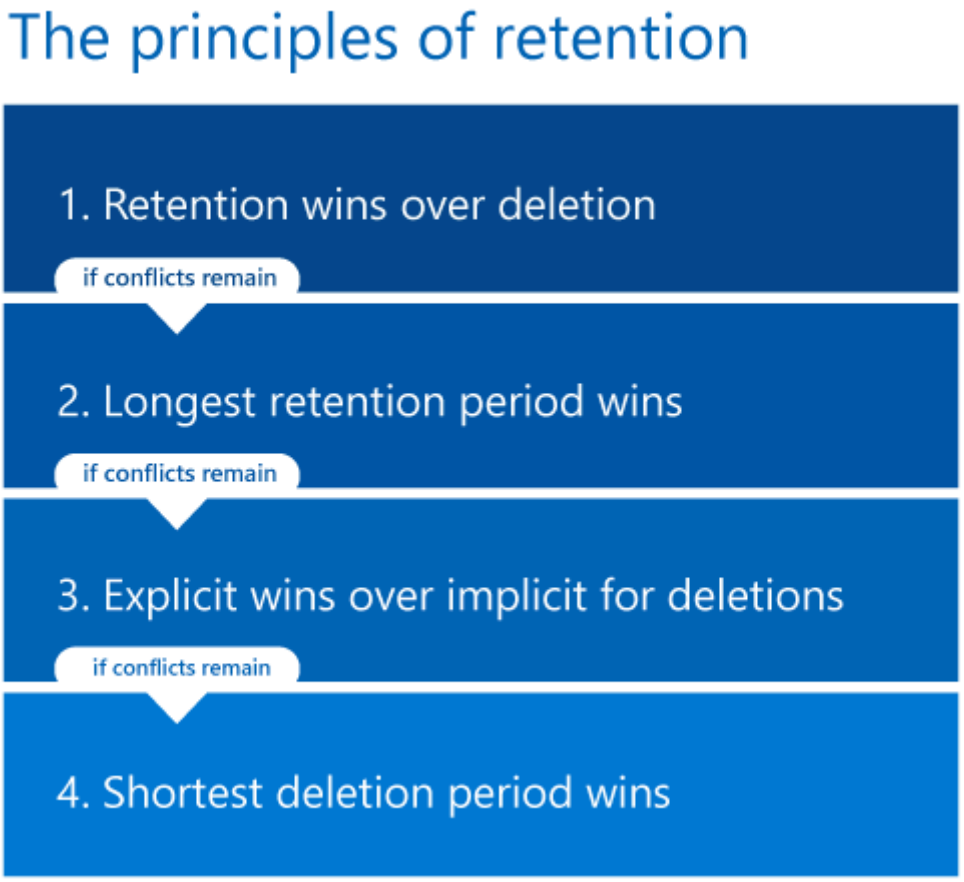
For example, an item might be subject to one retention policy that is configured for a delete-only action, and another retention policy that is configured to retain and then delete. Consequently, this item has just one retain action but two delete actions. The retention and deletion actions could be in conflict with one another and the two deletion actions might have a conflicting date. The principles of retention explain the outcome.

By default, retention always takes precedence over permanent deletion, and the longest retention period wins. These two simple rules always decide how long an item is retained unless it's subject to priority cleanup that might be needed to expedite permanent deletion for exceptional circumstances.

If you're not using priority cleanup, there are a few more factors that determine when an item is permanently deleted, which include the delete action from a retention label always takes precedence over the delete action from a retention policy.

Use the following flow to understand the retention and deletion outcomes for a single item that isn't subject to priority cleanup, where each level acts as a tie-breaker for conflicts, from top to bottom. If the outcome is determined by the first level because there are no further conflicts, there's no need to progress to the next level, and so on.

> **Important**
>
> If you are using retention labels: Before applying the principles to determine the outcome of multiple retention settings on the same item, make sure you know **which retention label is applied**.



Before explaining each principle in more detail, it's important to understand the difference between the retention period for the item vs. the specified retention period in the retention policy or retention label. That's because although the default configuration is to start the retention period when an item is created, so that the end of the retention period is fixed for the item, files also support the configuration to start the retention period from when the file is last modified. With this alternative configuration, every time the file is modified, the start of the retention period is reset, which extends the end of the retention period for the item. Retention labels also support starting the retention period when labeled and at the start of an event.

To apply the principles in action with a series of Yes and No questions, you can also use the retention flowchart.

Explanation for the four different principles:

1. **Retention wins over deletion.** Content won't be permanently deleted when it also has retention settings to retain it. While this principle ensures that content is preserved for compliance reasons, the delete process can still be initiated (user-initiated or system-initiated) and consequently, might remove the content from users' main view. However, permanent deletion is suspended. For more information about how and where content is retained, use the following links for each workload:

- How retention works for SharePoint and OneDrive
- How retention works with Microsoft Teams
- How retention works with Viva Engage
- How retention works for Exchange
- How retention works with Copilot & AI apps

Example for this first principle: An email message is subject to a retention policy for Exchange that is configured to delete items three years after they are created, and it also has a retention label applied that is configured to retain items five years after they are created.

The email message is retained for five years because this retention action takes precedence over deletion. The email message is permanently deleted at the end of the five years because of the delete action that was suspended while the retention action was in effect.

2. **The longest retention period wins.** If content is subject to multiple retention settings that retain content for different periods of time, the content is retained until the end of the longest retention period for the item.

   > ⓘ **Note**
   >
   > It's possible for a retention period of five years in a retention policy or label wins over a retention period of seven years in a retention policy or label, because the 5-year period is configured to start based on when the file is last modified, and the seven year period is configured to start from when the file is created.

   Example for this second principle: Documents in the Marketing SharePoint site are subject to two retention policies. The first retention policy is configured for all SharePoint sites to retain items for five years after they are created. The second retention policy is configured for specific SharePoint sites to retain items for 10 years after they are created.

   Documents in this Marketing SharePoint site are retained for 10 years because that's the longest retention period for the item.

3. **Explicit wins over implicit for deletions.** With conflicts now resolved for retention, only conflicts for deletions remain:

   a. A retention label (however it was applied) provides explicit retention in comparison with retention policies, because the retention settings are applied to an individual item rather

than implicitly assigned from a container. This means that a delete action from a retention label always takes precedence over a delete action from any retention policy.

Example 1 for this third principle (label): A document is subject to two retention policies that have a delete action of five years and 10 years respectively, and also a retention label that has a delete action of seven years.

The document is permanently deleted after seven years because the delete action from the retention label takes precedence.

Example 2 for this third principle (label): A document in the Marketing SharePoint site is subject to both a retention policy and a retention label. The retention policy is configured to retain items in all SharePoint sites for 10 years after creation, while the retention label specifies a delete action after seven years.

In this case, the document is retained for 10 years, which is the longest retention period. Because the retention label's delete action takes precedence, the document is permanently deleted after 10 years. The item is not moved to the Preservation Hold Library (PHL).

Example 3 for this third principle (label): An email message is subject to one retention policy and a retention label. The retention policy is configured for all Exchange mailboxes to retain items for 10 years after they are creation, and the retention label specifies a delete action after seven years.

The email message is retained for 10 years, which is the longest retention period. After seven years, the retention label causes the item to be moved to the Recoverable Items folder. The item remains there until the 10 year retention period expires, after which it is permanently deleted.

b. When you have retention policies only: If a retention policy for a location uses an adaptive scope or a static scope that includes specific instances (such as specific users for Exchange email) that retention policy takes precedence over a static scope that is configured for all instances for the same location.

   A static scope that is configured for all instances for a location is sometimes referred to as an "org-wide policy". For example, **Exchange mailboxes** and the default setting of **All mailboxes**. Or, **SharePoint classic and communication sites** and the default setting of **All sites**. When retention policies aren't org-wide but have been configured with an adaptive scope or a static scope that includes specific instances, they have equal precedence at this level.

Example 1 for this third principle (policies): An email message is subject to two retention policies. The first retention policy is unscoped and deletes items after 10 years. The second retention policy is scoped to specific mailboxes and deletes items after five years.

The email message is permanently deleted after five years because the deletion action from the scoped retention policy takes precedence over the org-wide retention policy.

Example 2 for this third principle (policies): A document in a user's OneDrive account is subject to two retention policies. The first retention policy is scoped to include this user's OneDrive account and has a delete action after 10 years. The second retention policy is scoped to include this user's OneDrive account and has a delete action after seven years.

When this document will be permanently deleted can't be determined at this level because both retention policies are scoped to include specific instances.

4. **The shortest deletion period wins.** Applicable to determine when items will be deleted from retention policies and the outcome couldn't be resolved from the previous level: Content is permanently deleted at the end of the shortest retention period for the item.

   > ⓘ **Note**
   >
   > It's possible that a retention policy that has a retention period of seven years wins over a retention policy of five years because the first policy is configured to start the retention period based on when the file is created, and the second retention policy from when the file is last modified.

   Example for this fourth principle: A document in a user's OneDrive account is subject to two retention policies. The first retention policy is scoped to include this user's OneDrive account and has a delete action of 10 years after the file is created. The second retention policy is scoped to include this user's OneDrive account and has a delete action of seven years after the file is created.

   This document will be permanently deleted after seven years because that's the shortest retention period for the item from these two scoped retention policies.

Items subject to eDiscovery hold also fall under the first principle of retention; they cannot be permanently deleted by any retention policy or retention label. When that hold is released, the principles of retention continue to apply to them. For example, they could then be subject to an unexpired retention period or a delete action.

Principles of retention examples that combine retain and delete actions

The following examples are more complex to illustrate the principles of retention when different retain and delete actions are combined. To make the examples easier to follow, all retention policies and labels use the default setting of starting the retention period when the item is created so the end of the retention period is the same for the item.

1. An item has the following retention settings applied to it:

   - A retention policy for delete-only after five years
   - A retention policy that retains for three years and then deletes
   - A retention label that retains-only for seven years

   Outcome: The item is retained for seven years because retention takes precedence over deletion and seven years is the longest retention period for the item. At the end of this retention period, the item is permanently deleted because of the delete action from the retention policies.

   Although the two retention policies have different dates for the delete actions, the earliest that the item can be permanently deleted is at the end of the longest retention period, which is longer than both deletion dates.

2. An item has the following retention settings applied to it:

   - An org-wide retention policy that deletes-only after 10 years
   - A retention policy scoped with specific instances that retains for five years and then deletes
   - A retention label that retains for three years and then deletes

   Outcome: The item is retained for five years because that's the longest retention period for the item. At the end of that retention period, the item is permanently deleted because of the delete action of three years from the retention label. Deletion from retention labels takes precedence over deletion from all retention policies. In this example, all conflicts are resolved by the third level.

Use Preservation Lock to restrict changes to policies

Some organizations might need to comply with rules defined by regulatory bodies such as the Securities and Exchange Commission (SEC) Rule 17a-4, which requires that after a policy for

retention is turned on, it cannot be turned off or made less restrictive.

Preservation Lock ensures your organization can meet such regulatory requirements because it locks a retention policy or retention label policy so that no one—including an administrator—can turn off the policy, delete the policy, or make it less restrictive.

You apply Preservation Lock after the retention policy or retention label policy is created. For more information and instructions, see Use Preservation Lock to restrict changes to retention policies and retention label policies.

Releasing a policy for retention

Providing your policies for retention don't have a Preservation Lock, you can delete your policies at any time, which effectively turns off the retention settings for a retention policy, and retention labels can no longer be applied from retention label policies. Any previously applied retention labels remain with their configured retention settings and for these labels, you can still update the retention period when it's not based on when items were labeled.

You can also keep a policy, but change the location status to off, or disable the policy. Another option is to reconfigure the policy so it no longer includes specific users, sites, groups, and so on.

Additional information for specific locations:

- **SharePoint sites and OneDrive accounts:**

  When you release a retention policy for SharePoint sites and OneDrive accounts, any content that's subject to retention from the policy continues to be retained for 30 days to prevent inadvertent data loss. During this 30-day grace period deleted files are still retained (files continue to be added to the Preservation Hold library), but the timer job that periodically cleans up the Preservation Hold library is suspended for these files so you can restore them if necessary.

  An exception to this 30-day grace period is when you update the policy to exclude one or more sites for SharePoint or accounts for OneDrive; in this case, the timer job deletes files for these locations in the Preservation Hold library without the 30-day delay.

  For more information about the Preservation Hold library, see How retention works for SharePoint and OneDrive.

  Because of the behavior during the grace period, if you re-enable the policy or change the location status back to on within 30 days, the policy resumes without any permanent data

loss during this time.

- **Exchange email and Microsoft 365 Groups**

  When you release a retention policy for mailboxes that are inactive at the time the policy is released:

  - If the retention policy is explicitly applied to a mailbox, the retention settings no longer apply. With no retention settings applied, an inactive mailbox becomes eligible for automatic deletion in the usual way.

    An explicit retention policy requires either an adaptive policy scope, or a static policy scope with an include configuration that specified an active mailbox at the time the policy was applied and later became inactive

  - If the retention policy is implicitly applied to a mailbox and the configured retention action is to retain, the retention policy continues to apply and an inactive mailbox never becomes eligible for automatic deletion. When the retain action no longer applies because the retention period has expired, the Exchange admin can now manually delete the inactive mailbox

    An implicit retention policy requires a static policy scope with the **All mailboxes** (for Exchange email) or **All groups** (for Microsoft 365 Groups) configuration.

    For more information about inactive mailboxes that have retention policies applied, see Inactive mailboxes and Microsoft 365 retention.

Auditing retention configuration and actions

When auditing is enabled, auditing events for retention are supported for both administration configuration (retention policies and retention labels) and retention actions (retention labels only).

Auditing retention configuration

Administrator configuration for retention policies and retention labels is logged as auditing events when a retention policy or label is created, reconfigured, or deleted.

For the full list of auditing events, see Retention policy and retention label activities.

Auditing retention actions

Retention actions that are logged as auditing events are available only for retention labels and not for retention policies:

- Specific to priority cleanup, use the **Cleanup ID** assigned to the priority cleanup policy as a keyword search term in the auditing log to find related activities.
- Specific to Adaptive Protection when a retention label is applied to an item:
  - For SharePoint and OneDrive, from **Retention policy and retention label activities**, select **Retained file proactively**
  - For Exchange, from **Retention policy and retention label activities**, select **Retained email item proactively**
- When a retention label is applied, changed, or removed from an item in SharePoint or OneDrive:
  - From **File and page activities**, select **Changed retention label for a file**
- When a labeled item in SharePoint is marked as a record, and it is unlocked or locked by a user:
  - From **File and page activities**, select **Changed record status to unlocked** and **Changed record status to locked**
- When a retention label that marks content as a record or regulatory record is applied to an item in Exchange:
  - From **Exchange mailbox activities**, select **Labeled message as a record**
- When a labeled item in SharePoint, OneDrive, or Exchange is marked as a record or regulatory record, and it is permanently deleted:
  - From **File and page activities**, select **Deleted file marked as a record**
- When a disposition reviewer takes action for an item that's reached the end of its retention period:
  - From **Disposition review activities**, select **Approved disposal**, **Extended retention period**, **Relabeled item**, or **Added reviewers**

PowerShell cmdlets for retention policies and retention labels

Use Security & Compliance PowerShell for Purview retention cmdlets that support configuration at scale, scripting for automation, or might be necessary for advanced configuration scenarios.

For a list of available cmdlets, and to identify which ones are supported for the different locations, see PowerShell cmdlets for retention policies and retention labels.

When to use retention policies and retention labels or eDiscovery holds

Although retention settings and holds that you create with an eDiscovery case can both prevent data from being permanently deleted, they are designed for different scenarios. To help you understand the differences and decide which to use, use the following guidance:

- Retention settings that you specify in retention policies and retention labels are designed for a long-term data lifecycle management strategy to retain or delete data for compliance requirements. The scope is usually broad with the main focus being the location and content rather than individual users. The start and end of the retention period is configurable, with the option to automatically delete content without additional administrator intervention.
- Holds for eDiscovery cases are designed for a limited duration to preserve data for a legal investigation. The scope is specific with the focus being content owned by identified users. The start and end of the preservation period isn't configurable but dependent on individual administrator actions, without an option to automatically delete content when the hold is released.

Summary to compare retention with holds:

Expand table

Consideration	Retention	eDiscovery holds
Business need:	Compliance	Legal
Time scope:	Long-term	Short-term
Focus:	Broad, content-based	Specific, user-based
Start and end date configurable:	Yes	No
Content deletion:	Yes (optional)	No
Administrative overheads:	Low	High

If content is subject to both retention settings and an eDiscovery hold, preserving content for the eDiscovery hold always takes precedence. In this way, the principles of retention expand to

eDiscovery holds because they preserve data until an administrator manually releases the hold. However, despite this precedence, don't use eDiscovery holds for long-term data lifecycle management. If you are concerned about automatic deletion of data, you can configure retention settings to retain items forever, or use disposition review with retention labels.

If you are using older eDiscovery tools to preserve data, see the following resources:

- Exchange:
  - In-Place Hold and Litigation Hold
  - How to identify the type of hold placed on an Exchange Online mailbox
- SharePoint and OneDrive:
  - Add content to a case and place sources on hold in the eDiscovery Center
- Retirement of legacy eDiscovery tools

Use retention policies and retention labels instead of older features

If you need to retain or delete content in Microsoft 365 for data lifecycle management, we recommend you use Microsoft 365 retention policies and retention labels instead of the following older features.

If you currently use these older features, they'll usually work side by side with Microsoft 365 retention policies and retention labels. Check their specific documentation for any restrictions. However, we recommend that going forward, you use Microsoft 365 retention policies and retention labels to benefit from a single solution to manage both retention and deletion of content across multiple workloads in Microsoft 365.

> **Tip**
>
> For SharePoint, see the following resources:
>
> - Use Microsoft Purview risk and compliance solutions instead of the older information management and records management features in SharePoint for Microsoft 365
> - Migration strategies for moving to Microsoft Purview risk and compliance solutions

Older features from Exchange Online:

- Retention tags and retention policies, also known as messaging records management (MRM) (deletion only)

  However, if you use the following MRM features, be aware that they aren't currently supported by Microsoft 365 retention policies:

  - An archive policy for archive mailboxes to automatically move emails from a user's primary mailbox to their archive mailbox after a specified period of time. An archive policy (with any settings) can be used in conjunction with a Microsoft 365 retention policy that applies to a user's primary and archive mailbox.
  - Retention policies applied by an admin to specific folders within a mailbox. A Microsoft 365 retention policy applies to all folders in the mailbox. However, an admin can configure different retention settings by using retention labels that a user can apply to folders in Outlook as a default retention label.

- Journaling (retention and archive)

  Might be required to integrate with third-party solutions and copies of email messages and their data communication are stored outside Exchange Online. Because you're moving data outside Microsoft 365, you must take extra precautions to secure it and also resolve any duplications that might result from this solution. It is your responsibility to monitor and follow up on any nondelivery receipts to the journaling mailbox that can occur because of external and dependent services. You don't have these additional administrative overheads when you use Microsoft 365 retention and other Microsoft Purview compliance solutions that also aren't limited to just email messages.

- Litigation hold (retention only)

  Although Litigation holds are still supported, we recommend you use Microsoft 365 retention or eDiscovery holds, as appropriate.

Older features from SharePoint and OneDrive:

- Document deletion policies (deletion only)
- Configuring in place records management (retention only)
- Use policies for site closure and deletion (deletion only)
- Information management policies (deletion only)

If you have configured SharePoint sites for content type policies or information management policies to retain content for a list or library, those policies are ignored while a retention policy or retention label policy is in effect.

Related information

- SharePoint Online Limits
- Limits and specifications for Microsoft Teams
- Resources to help you meet regulatory requirements for data lifecycle management and records management

Configuration guidance

See Get started with data lifecycle management. This article has information about subscriptions, permissions, and links to end-to-end configuration guidance for retention scenarios.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 07/22/2026

Flowchart to determine when an item will be retained or permanently deleted

Summarize this article for me

> *Microsoft Purview service description*

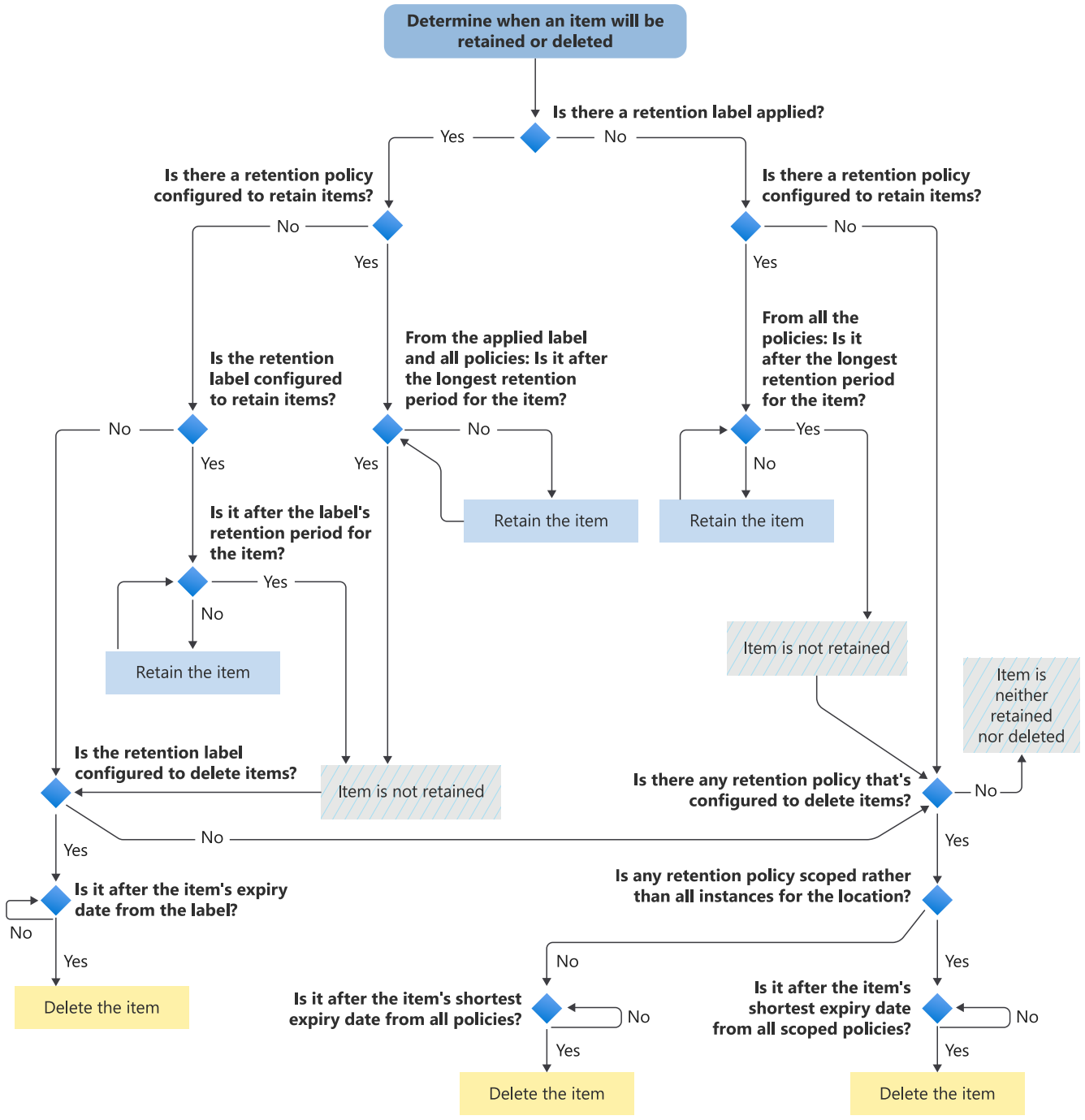
Use the following flowchart to apply the principles of retention to an item to determine if the system will retain it or permanently delete it as a result of a retention label or retention policy.

This logic flow is used for an item when any of the following conditions apply:

- There's more than one retention policy applied
- There's a retention label and one or more retention policies
- The item isn't subject to priority cleanup

When an item is subject to an eDiscovery hold (or the older technologies of Litigation hold or In-Place Hold), unless it is also subject to priority cleanup, it is always retained before the decision flows for retention policies and a retention label.

If any of the terms used in this flowchart are unfamiliar to you, see Learn about retention policies and retention labels.



> **Note**
>
> It's important to distinguish between the longest retention period for the item vs. the longest specified period in a retention policy or label. And similarly, between the shortest expiry date for the item vs. the shortest specified period in a retention policy.
>
> For more information, see the explanation after the graphic in the principles of retention section.

Last updated on 02/27/2026

Learn about retention for SharePoint and OneDrive

09/22/2025

> *Microsoft 365 licensing guidance for security & compliance*.

The information in this article supplements Learn about retention because it has information that's specific to SharePoint and OneDrive.

For other workloads, see:

- Learn about retention for Microsoft Teams
- Learn about retention for Viva Engage
- Learn about retention for Exchange
- Learn about retention for Copilot

What's included for retention and deletion

> **Note**
>
> In preview, Microsoft Facilitator AI-generated notes in meetings are supported by retention policies and retention labels.

All files stored in SharePoint (and SharePoint Embedded) or OneDrive sites can be retained by applying a retention policy or retention label. For SharePoint, archived sites are supported in addition to active sites.

The following files can be deleted:

- When you use a retention policy: All files in document libraries, which include any automatically created SharePoint document libraries, such as **Site Assets**.
- When you use retention labels: All files in all document libraries, and all files at the root level that aren't in a folder.

> **Tip**

When you use a **query with an auto-apply policy for a retention label**, you can exclude specific document libraries by using the following entry: `NOT(DocumentLink:"<URL to document library>")`

Files that can be retained and deleted include those used by Microsoft Loop and Copilot Pages. For these apps, content can be stored in SharePoint Embedded containers. Although the containers aren't SharePoint sites, for the purposes of retention and deletion, they behave the same as SharePoint sites. For more information about where the files are stored, see Overview of Loop storage.

List items aren't supported by retention policies but are supported by retention labels with the exception of items in system lists. These are hidden lists used by SharePoint to manage the system and include the master page catalog, solution catalog, and data sources. When retention labels are applied to supported list items, they will always be retained according to the retention settings, but not deleted if they are hidden from search.

When you apply a retention label to a supported list item that has a document attachment:

- For a standard retention label (doesn't declare the item to be a record):
  - The document attachment doesn't automatically inherit the retention settings of the label, but can be labeled independently.
- For a retention label that declares the item a record:
  - The document attachment automatically inherits the retention settings from the label if the document isn't already labeled.

Retention settings from both retention policies and retention labels don't apply to organizing structures that include libraries, lists, folders, and Loop workspaces.

For retention policies and label policies: SharePoint sites must be indexed for the retention settings to be applied. However, if items in SharePoint document libraries are configured to not appear in search results, this configuration doesn't exclude files from the retention settings.

If a site is configured with the Set-SPOSite parameter *LockState* that's set to NoAccess or ReadOnly, items in that site can't be deleted with this site configuration.

How retention works for SharePoint and OneDrive

To store content that needs to be retained, SharePoint and OneDrive create a Preservation Hold library if one doesn't exist for the site. The Preservation Hold library is a hidden system location that isn't designed to be used interactively but instead, automatically stores files when this is needed for compliance reasons. It's not supported to edit, delete, or move these automatically retained files yourself. Or, change or remove retention labels and sensitivity

labels for these files. Instead, use compliance tools, such as those supported by eDiscovery to access the content in these files.

The Preservation Hold library works in the following way to support retention policies and retention labels:

When a user changes an item that's subject to retention from a retention policy or a retention label that marks items as a record, or deletes any item subject to retention, the original content is copied to the Preservation Hold library. This behavior lets the user change or delete the content in their app, while keeping a copy of the original for compliance reasons.

Expand table

Retention item	User edit or delete action	Copy created in the Preservation Hold library
Retention policy	Edit item	Yes *
	Delete item	Yes
Standard retention label (doesn't mark items as records or regulatory records)	Edit item	No
	Delete item	Yes
Retention label that marks items as records	Edit unlocked item	Yes
	Edit locked item	Not applicable – action blocked
	Delete item	Not applicable – action blocked
Retention label that marks items as regulatory records	Edit item	Not applicable – action blocked
	Delete item	Not applicable – action blocked

Footnote:

* New content isn't copied to the Preservation Hold library the first time it's edited. To retain all versions of a file, versioning must be turned on for the site.

This behavior for copying files into the Preservation Hold library applies to content that exists when the retention settings were applied. In addition, for retention policies, any new content that's created or added to the site after it was included in the policy will be retained in the Preservation Hold library.

A timer job periodically runs on the Preservation Hold library. For content that has been in the Preservation Hold library for more than 30 days, this job compares the content to all queries

used by the retention settings for that content. Content that is older than their configured retention period and isn't awaiting disposition review is then deleted from the Preservation Hold library, and from the original location if it's still there. This timer job runs every seven days, which means that together with the minimal 30 days, it can take up to 37 days for content to be deleted from the Preservation Hold library.

Users see an error message if they try to delete a library, list, or site that's subject to retention.

Users also see an error message if they try to delete a labeled item in any of the following circumstances. The item isn't copied to the Preservation Hold library but remains in the original location:

- The records management setting that allows users to delete labeled items is turned off.

  To check or change this setting, go to the records management settings in the Microsoft Purview portal: **Settings** > **Solution settings **> Records Management > **Retention Labels** > **Deleting content labeled for retention**. There are separate settings for OneDrive and SharePoint.

  Alternatively, and if you don't have access these settings in the portals, you can use *AllowFilesWithKeepLabelToBeDeletedSPO* and *AllowFilesWithKeepLabelToBeDeletedODB* from Get-PnPTenant and Set-PnPTenant.

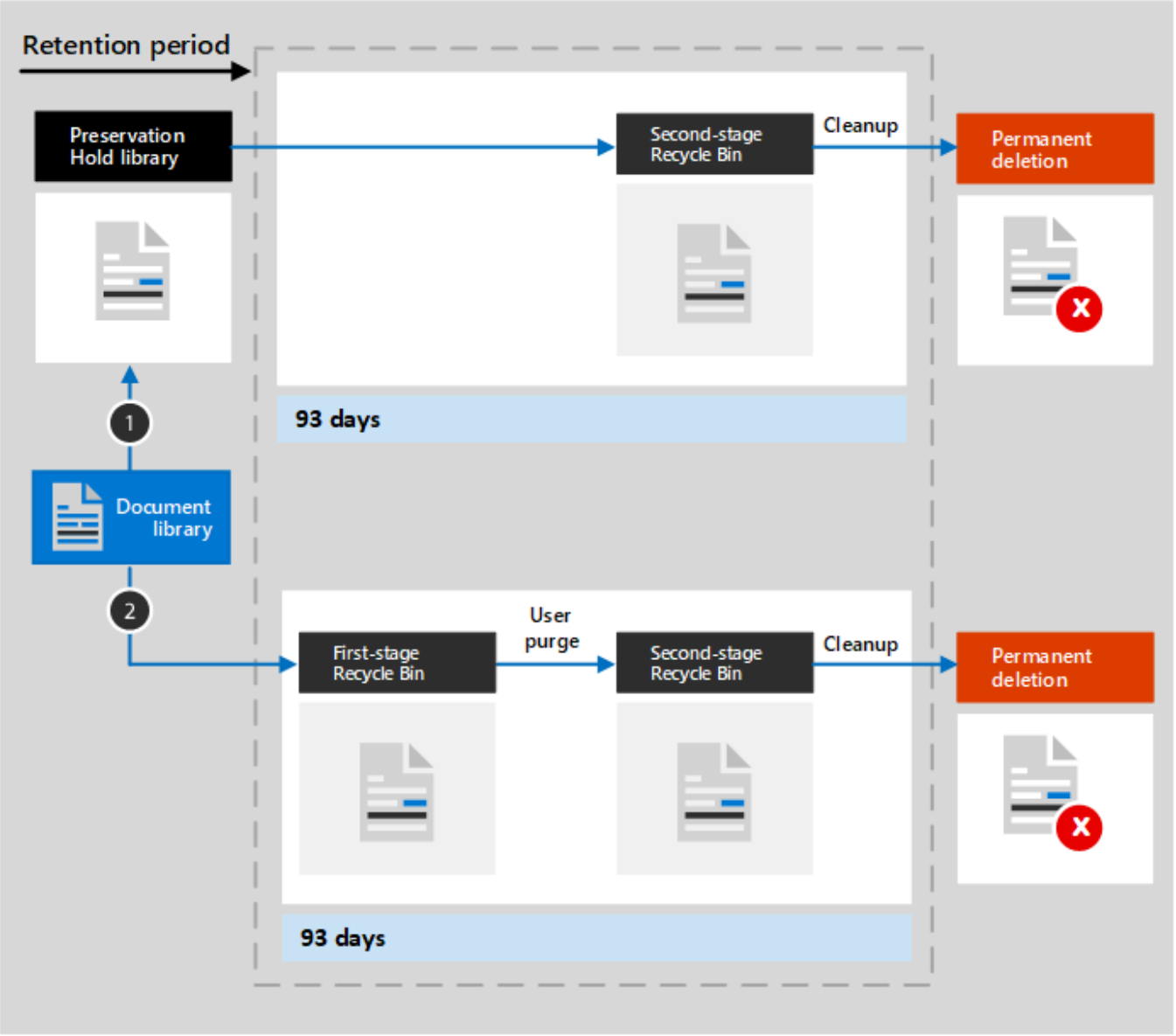
- The retention label marks items as a record and it's locked.

  Only when the record is unlocked, does a copy of the last version get stored in the Preservation Hold library.

- The retention label marks items as a regulatory record, which always prevents the item from being edited or deleted.

After retention settings are assigned to content in a OneDrive account, SharePoint site, or SharePoint Embedded container for a Loop workspace, the paths the content takes depend on whether the retention settings are to retain and delete, to retain only, or delete only. In the explanations that follow, modified content is moved to the Preservation Hold library for retention policies, and retention labels that mark items as records (and the content is unlocked). Items that are modified with retention labels that don't mark items as records don't create copies in the Preservation Hold library, but do when items are deleted.

When the retention settings are to retain and delete:



1. **If the content is modified or deleted** during the retention period, a copy of the original content as it existed when the retention settings were assigned is created in the Preservation Hold library. There, the timer job identifies items whose retention period has expired. Those items are moved to the second-stage Recycle Bin, where they're permanently deleted at the end of 93 days. The second-stage Recycle Bin isn't visible to end users (only the first-stage Recycle Bin is), but site collection admins can view and restore content from there.

> ⓘ Note
>
> To help prevent inadvertent data loss, we no longer permanently delete content from the Preservation Hold library. Instead, we permanently delete content only from the Recycle Bin, so all content from the Preservation Hold library now goes through the second-stage Recycle Bin.

2. **If the content is not modified or deleted** during the retention period, the timer job moves this content to the first-stage Recycle Bin at the end of the retention period. If a user deletes the content from there or empties this Recycle Bin (also known as purging), the document is moved to the second-stage Recycle Bin. A 93-day retention period spans both the first- and second-stage recycle bins. At the end of 93 days, the document is permanently deleted from wherever it resides, in either the first-stage or second-stage Recycle Bin. The Recycle Bin isn't indexed and therefore unavailable for searching. As a result, an eDiscovery search can't find any Recycle Bin content on which to place a hold.

> ⓘ **Note**
>
> Because of the first principle of retention, permanent deletion is always suspended if the same item must be retained because of another retention policy or retention label, or it's under eDiscovery holds for legal or investigative reasons.

When the retention settings are retain-only, or delete-only, the contents paths are variations of retain and delete:

Content paths for retain-only retention settings

1. **If the content is modified or deleted** during the retention period: A copy of the original document is created in the Preservation Hold library and retained until the end of the retention period, when the copy in the Preservation Hold library is moved to the second-stage Recycle Bin and is permanently deleted after 93 days.

2. **If the content is not modified or deleted** during the retention period: Nothing happens before and after the retention period; the document remains in its original location.

Content paths for delete-only retention settings

1. **If the content is deleted** during the configured period: The document is moved to first-stage Recycle Bin. If a user deletes the document from there or empties this Recycle Bin, the document is moved to the second-stage Recycle Bin. A 93-day retention period spans both the first-stage and second-stage recycle bins. At the end of 93 days, the document is permanently deleted from wherever it resides, in either the first-stage or second-stage Recycle Bin. If the content is modified during the configured period, it follows the same deletion path after the configured period.

2. **If the content is not deleted** during the configured period: At the end of the configured period in the retention policy, the document is moved to the first-stage Recycle Bin. If a user deletes the document from there or empties this Recycle Bin (also known as

purging), the document is moved to the second-stage Recycle Bin. A 93-day retention period spans both the first-stage and second-stage recycle bins. At the end of 93 days, the document is permanently deleted from wherever it resides, in either the first-stage or second-stage Recycle Bin. The Recycle Bin isn't indexed and therefore unavailable for searching. As a result, an eDiscovery search can't find any Recycle Bin content on which to place a hold.

How retention works with cloud attachments

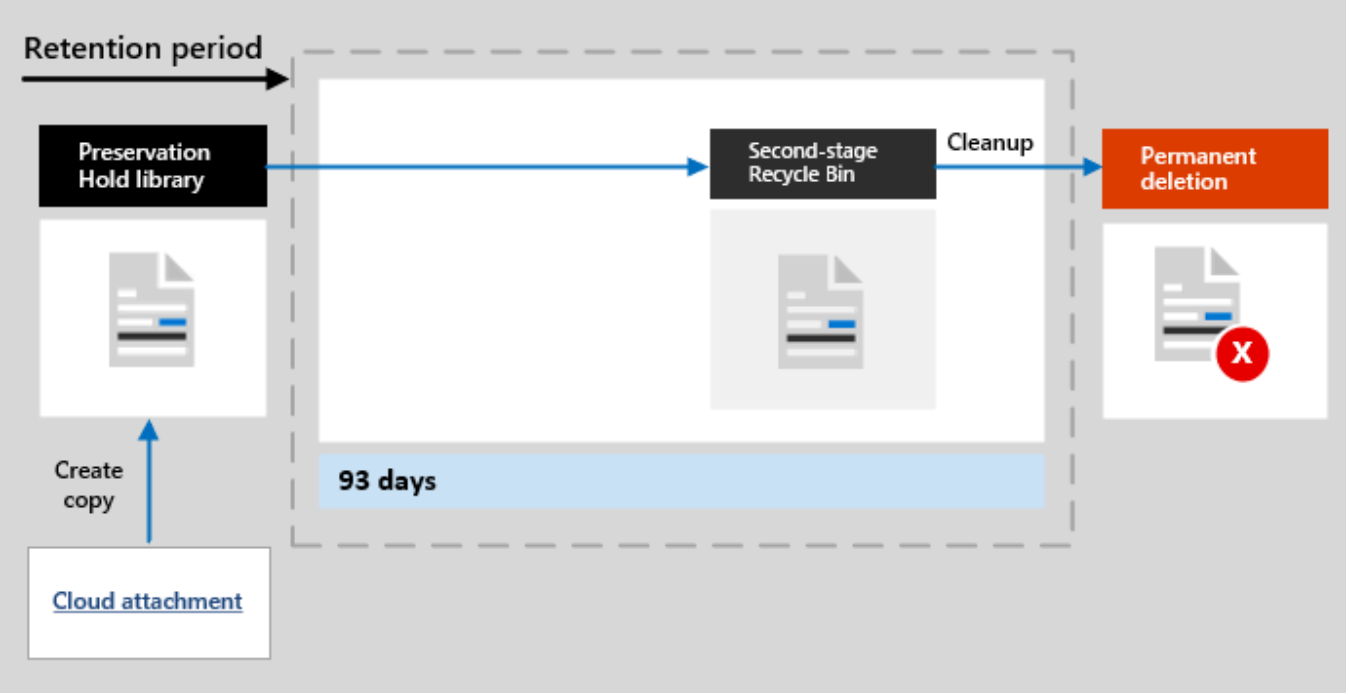
Cloud attachments are embedded links to files that users share, or referenced in interactions for Microsoft 365 Copilot and Microsoft 365 Copilot Chat. They can be retained and deleted when your users share them in Outlook emails and Teams or Viva Engage messages, and they are referenced in interactions with Copilot. When you automatically apply a retention label to cloud attachments, the retention label is applied to a copy of the shared file, which is stored in the Preservation Hold library.

For this scenario, we recommend you configure the label setting to start the retention period based on when the item is labeled. If you do configure the retention period based on when the item is created or last modified, this date is taken from the original file at the time of sharing. If you configure the start of retention to be when last modified, this setting has no effect for this copy in the Preservation Hold library.

However, if the original file is modified and then shared again, a new copy of the file as a new version is saved and labeled in the Preservation Hold library.

If the original file is shared again but not modified, the labeled date of the copy in the Preservation Hold library is updated. This action resets the start of the retention period and is why we recommend you configure the start of the retention period to be based on when the item is labeled.

Because the retention label isn't applied to the original file, the labeled file is never modified or deleted by a user. The labeled file remains in the Preservation Hold library until the timer job identifies that its retention period has expired. If the retention settings are configured to delete items, the file is then moved to the second-stage Recycle Bin, where it's permanently deleted at the end of 93 days:



The copy that's stored in the Preservation Hold library is typically created within an hour from the cloud attachment being shared.

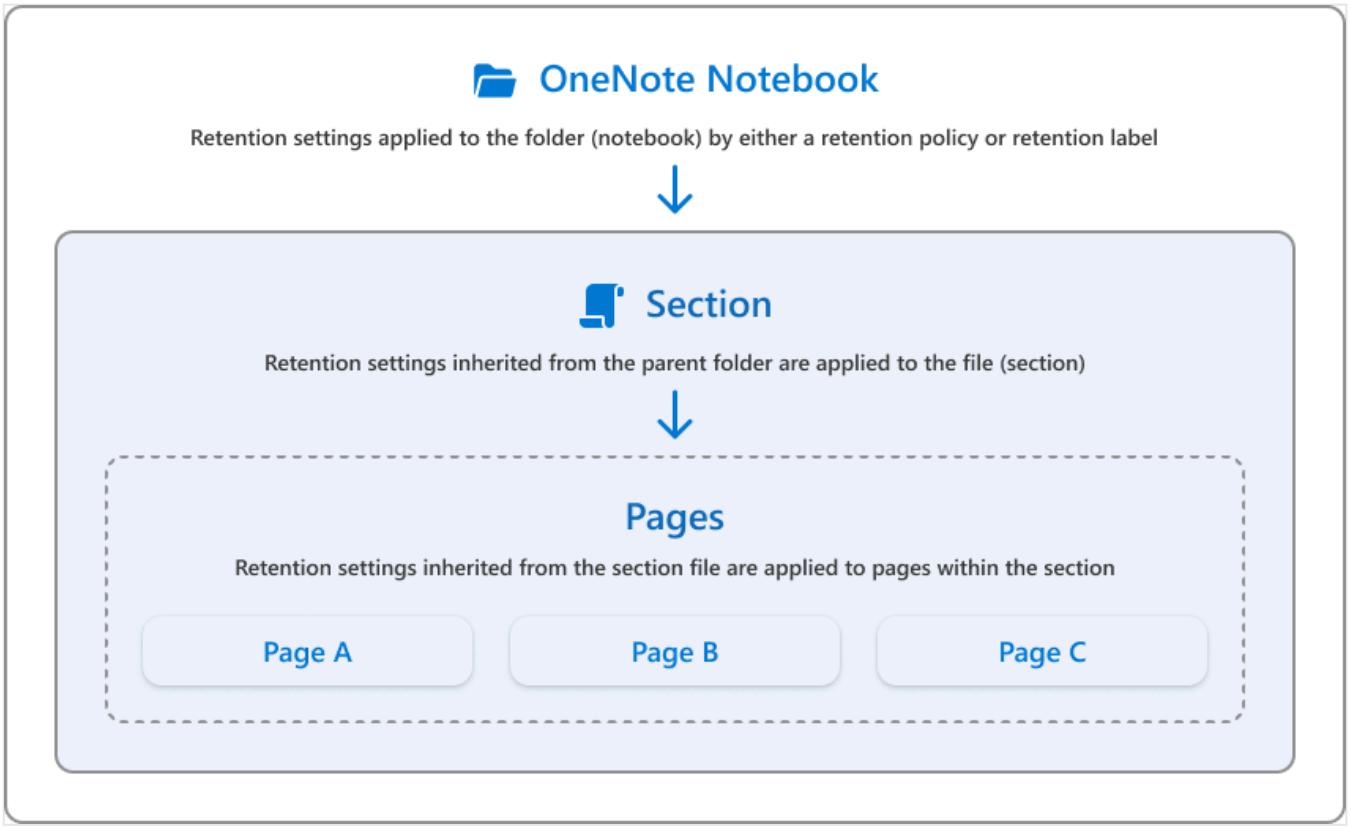
To safeguard against the original file being deleted or moved by users before the copy can be created and labeled, files in locations included in the auto-labeling policy are automatically copied into the Preservation Hold library if they are deleted or moved. These files have a temporary retention period of one day and then follow the standard cleanup process described on this page. When the original file has been deleted or moved, the copy for retaining cloud attachments uses this version of the file. The automatic and temporary retention of deleted or moved files in the Preservation Hold library is unique to auto-labeling policies for cloud attachments.

How retention works with OneNote content

When you apply a retention policy to a location that includes OneNote content, or a retention label to a OneNote folder, the different OneNote sections inherit the retention settings as individual files. Pages from each section are contained within the file and inherit the retention settings from their parent section.

Because of this structure, each section will be individually retained and deleted (with all its pages), according to the retention settings you specify.

Only sections are impacted by the retention settings that you specify. For example, although you see a **Modified** date for each individual notebook, this date isn't used by Microsoft 365 retention.



How retention works with document versions

Versioning is a feature of all document lists and libraries in SharePoint and OneDrive. By default, versioning retains a minimum of 500 major versions, although you can change this limit. For more information, see Enable and configure versioning for a list or library and How versioning works in lists and libraries.

When a document with versions is subject to retention settings to retain that content, and it's not marked as a record, how the versions are stored in the Preservation Hold library changed in July 2022 to improve performance. Now, all versions of that file are retained in a single file in the Preservation Hold library. Before the change, versions were copied to the Preservation Hold library as separate files, and after the change, remain as separate files.

> **Note**
>
> Versions that are from a record continue to be copied to the Preservation Hold library as separate files, which means that they can expire independently from each other and the current version.

If the label doesn't mark the item as a record and retention settings are configured to delete the item at the end of the retention period:

- If the retention period is based on when the content was created, when labeled, or when an event starts, each version has the same expiration date as the original document. The original document and its versions all expire at the same time.

- If the retention period is based on when the content was last modified:
  - **After the change where all versions of the file are retained in a single file in the Preservation Hold library**: Each version has the same expiration date as the last version of the document. The last version of the document and its versions all expire at the same time.
  - **Before the change where versions were copied to the Preservation Hold library as separate files**: Each version has its own expiration date based on when the original document was modified to create that version. The original document and its versions expire independently of each other.

When the retention action is to delete the document, all versions not in the Preservation Hold library are deleted at the same time according to the current version.

For items that are subject to a retention policy (or an eDiscovery hold), the versioning limits for the document library are ignored until the retention period of the document is reached (or the eDiscovery hold is released). In this scenario, old versions aren't automatically purged and users are prevented from deleting versions.

That's not the case for retention labels when the content isn't subject to a retention policy (or an eDiscovery hold). Instead, the versioning limits are honored so that older versions are automatically deleted to accommodate new versions, but users are still prevented from deleting versions.

How retention works with Microsoft 365 Archive

For administrators, there's very little change to how retention policies and retention labels work and are managed for sites that use Microsoft 365 Archive. For example, the default policy configuration of all sites automatically includes archived sites as well as active sites. An active site that's included in a retention policy and then changed to be an archive site will continue to be subject to the configuration settings in the retention policy. The same applies to labeled items in a site that becomes archived. You can create a new retention policy for an archived site, and auto-apply retention labels for archived sites. Items still support disposition review, Power Automate actions, simulation mode is supported, policy lookup, adaptive scopes, and Microsoft Graph API to programmatically apply and manage retention labels are all supported for archived sites.

The one exception is for cloud attachments, where an item that's currently in an archived site won't be retained with an auto-apply retention label policy. Cloud attachments that were

retained from an active site continue to to be subject to the configuration settings in the retention label.

Because users can't view and interact with items in archived sites, the user actions usually supported for retention labels won't be possible. For example, manually applying or removing retention labels, locking and unlocking records, editing of record properties that include the name and description. Similarly, although the Microsoft Purview portal supports disposition review, the contents of an item under disposition review can't be displayed and the URL link to the item won't work.

When a user leaves the organization

SharePoint:

When a user leaves your organization, any content created by that user isn't affected because SharePoint is considered a collaborative environment, unlike a user's mailbox or OneDrive account.

OneDrive:

If a user leaves your organization, any files that are subject to a retention policy or has a retention label will remain subject to the retention settings for the duration of the retention period specified in the policy or label. During that time, all sharing access continues to work and the content continues to be discoverable by Content Search and eDiscovery.

When the retention period expires and the retention settings included a delete action, content moves into the Site Collection Recycle Bin and isn't accessible to anyone except the admin.

Configuration guidance

If you're new to configuring retention in Microsoft 365, see Get started with data lifecycle management.

If you're ready to configure a retention policy or retention label for Exchange, see the following instructions:

- Create and configure retention policies
- Publish retention labels and apply them in apps
- Apply a retention label to content automatically

Learn about retention for Microsoft Teams

> *Microsoft Purview service description*

> **Note**
>
> If you're seeing a message in Teams that your chats or messages have been deleted by a retention policy, see **Teams messages about retention policies**.
>
> The information on this page is for IT administrators who manage these retention policies.

The information in this article supplements Learn about retention because it has information that's specific to Microsoft Teams messages and call logs.

For other workloads, see:

- Learn about retention for SharePoint and OneDrive
- Learn about retention for Viva Engage
- Learn about retention for Exchange
- Learn about retention for Copilot

What's included for retention and deletion

> **Note**
>
> Retention policies support **shared channels**. Any shared channels inherit retention settings from the parent team.

Teams chats messages and Facilitator interactions in meetings, channel messages, and private channel messages can be deleted by using retention policies for Teams. In addition to the text in the messages, the following items from these interactions can be retained for compliance reasons: Video clips, embedded images, tables, hypertext links, links to other Teams messages and files, and card content.

These chat messages and private channel messages include all the names of the people in the conversation, and channel messages include the team name and the message title (if supplied).

Code snippets, recorded voice memos from the Teams mobile client, thumbnails, announcement images, and reactions from others in the form of emoticons aren't retained when you use retention policies for Teams messages.

Some control messages are supported when you use retention policies for Teams messages. Control messages are system generated messages that appear on the Teams client. The types of control messages supported for retention policies are the same types of control messages available from Microsoft Teams Export APIs.

Call logs are also supported. Microsoft Teams call logs represent the collection of call-related data generated by Teams, including call data records (CDRs, sometimes also known as call detail records or just call records) and other call metadata. This data can be stored across multiple underlying locations used by Teams, but can be targeted as a single workload for retention purposes. For more information about Teams call logs, see the following resources:

- Personal call logs - View your call history in Microsoft Teams
- Shared call queue logs - View call history in the Queues app in Microsoft Teams

Emails and files that you use with Teams aren't included in retention policies for Teams. These items have their own retention policies that use the **Microsoft 365 Group mailboxes & sites** or **SharePoint classic and communication sites** locations. Teams meeting transcripts and Teams meeting recordings and transcripts from channel meetings are included with these files. However, for files and Teams meeting recordings and transcripts from user chats, you need a retention policy that includes the organizer's **OneDrive account** as the location.

> **Tip**
>
> If you want to retain or delete just Teams meeting transcripts or recordings with their accompanying transcripts, use an **auto-apply retention label policy** that identifies these files.

Migration of private channel messages in 2025

Towards the end of 2025, Teams is changing where private channel messages are stored. Before the migration, these messages are stored in user mailboxes and require their own location in a Teams retention policy: **Teams private channel messages**

After the migration, they're stored in group mailboxes, in the same way as messages for standard channels and shared channels. The **Teams channel messages** location in a retention policy will

therefore apply to all channel messages. Post migration, if a team has any private channels, they inherit retention settings from their parent team in the same way as shared channels do.

> **Important**
>
> If messages from private channels need different retention settings, during and post migration, create new retention policies for the **Teams channel messages** location that apply just to the parent teams that have private channels. Because of the **principles of retention**, you might also need to edit your existing retention policy for Teams channel messages, to exclude these parent teams that are now specified in your new policy.

Any previously configured retention policies that use the **Teams private channel messages** location continues to enforce the retention settings.

For more information, see the Tech Community blog post: Enhancing Private Channels in Microsoft Teams to Unlock Their Full Potential

How retention works with Microsoft Teams messages

Use this section to understand how your compliance requirements are met by backend storage and processes, and should be verified by eDiscovery tools rather than by messages that are currently visible in the Teams app.

You can use a retention policy to retain data from chats and channel messages in Teams, and delete these chats and messages. Behind the scenes, Exchange mailboxes are used to store data copied from these messages. Data from Teams chats is stored in a hidden folder in the mailbox of each user included in the chat, and a similar hidden folder in a group mailbox is used for Teams channel messages. These hidden folders aren't designed to be directly accessible to users or administrators, but instead, store data that compliance administrators can search with eDiscovery tools.

These mailboxes are, listed by their RecipientTypeDetails attribute:

- **UserMailbox**: These mailboxes store message data for cloud-based Teams users, and Teams private channels and before the migration.
- **MailUser**: These mailboxes store message data for on-premises Teams users.
- **GroupMailbox**: These mailboxes store message data for Teams standard channels, and for Teams private channels post migration.
- **SubstrateGroup**: These mailboxes store message data for Teams shared channels.

Other mailbox types, such as RoomMailbox that's used for Teams conference rooms, aren't supported for Teams retention policies.

Teams uses an Azure-powered chat service as its primary storage for all messages (chats and channel messages). If you need to delete Teams messages for compliance reasons, retention policies for Teams can delete messages after a specified period, based on when they were created. Messages are then permanently deleted from both the Exchange mailboxes where they stored for compliance operations, and from the primary storage used by the underlying Azure-powered chat service. For more information about the underlying architecture, see Security and compliance in Microsoft Teams and specifically, the Information Protection Architecture section.

Although this data from Teams chats and channel messages are stored in mailboxes, you must configure a retention policy for the **Teams channel messages** and **Teams chats** locations. Teams chats and channel messages aren't included in retention policies that are configured for Exchange user or group mailboxes. Similarly, retention policies for Teams don't affect other email items stored in mailboxes.

If a user is added to a chat, a copy of all messages shared with them are ingested into their mailbox. The created date of those messages doesn't change for the new user and remains the same for all users.

> ⓘ **Note**
>
> If a user is included in an active retention policy that retains Teams messages and you delete a mailbox of a user who is included in this policy, the mailbox is converted into an **inactive mailbox** to retain the Teams data. If you don't need to retain this Teams data for the user, exclude the user account from the retention policy and **wait for this change to take effect** before you delete their mailbox.

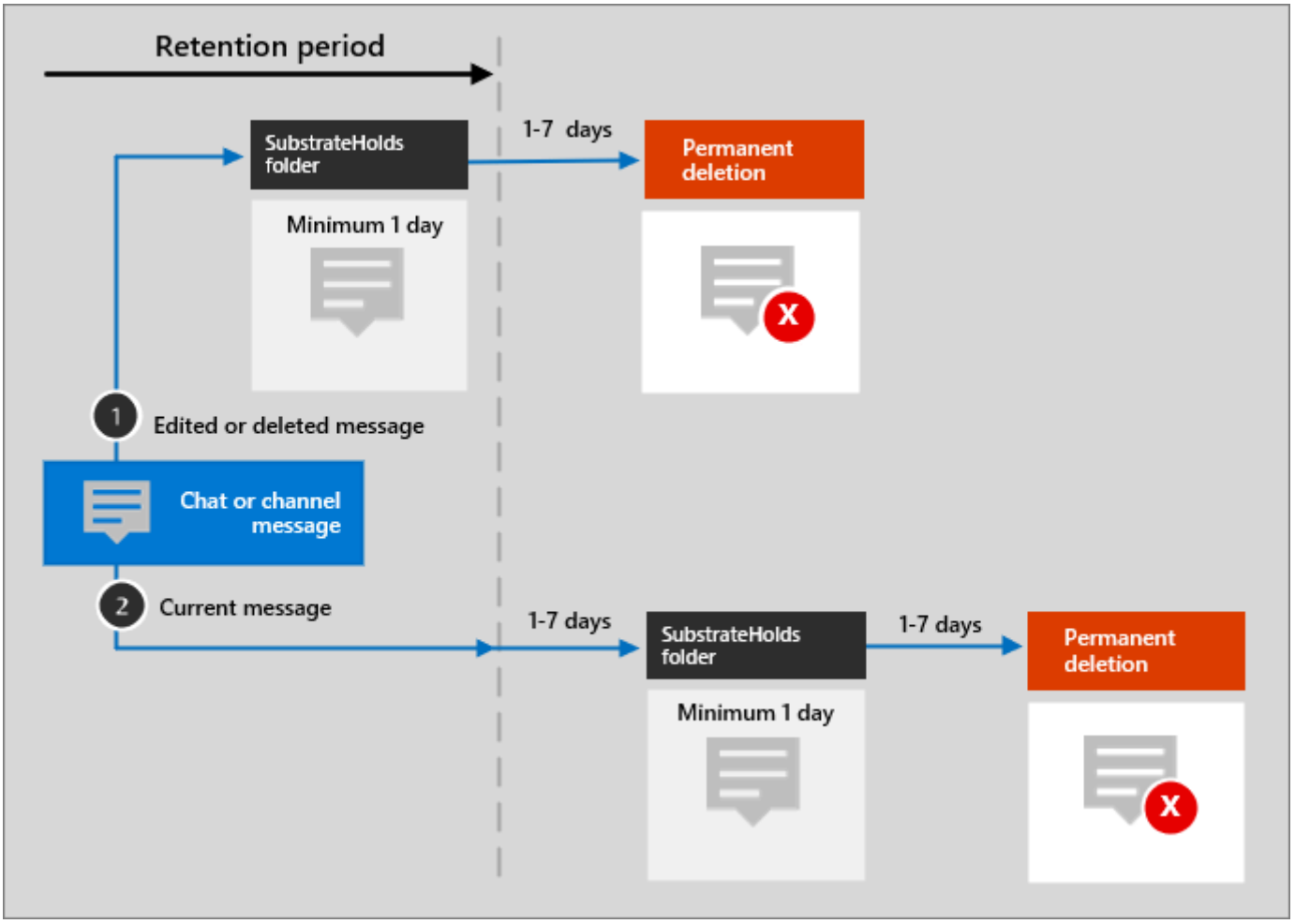
After a retention policy is configured for chat and channel messages, a timer job from the Exchange service periodically evaluates items in the hidden mailbox folder where these Teams messages are stored. The timer job typically takes 1-7 days to run. When these items have expired their retention period, they're moved to the SubstrateHolds folder—another hidden folder that's in every user or group mailbox to store "soft-deleted" items before they're permanently deleted.

Messages remain in the SubstrateHolds folder for at least 1 day, and then if they're eligible for deletion, the timer job permanently deletes them the next time it runs.

> **Important**
>
> Because of the [first principle of retention] and since Teams chat and channel messages are stored in Exchange Online mailboxes, permanent deletion from the SubstrateHolds folder is always suspended if the mailbox is affected by another Teams retention policy for the same location, Litigation Hold, delay hold, or if an eDiscovery hold is applied to the mailbox for legal or investigative reasons.
>
> While the mailbox is included in an applicable hold, Teams chat and channel messages that have been deleted will no longer be visible in the Teams app but will continue to be discoverable with eDiscovery.

After a retention policy is configured for chat and channel messages, the paths the content takes depend on whether the retention policy is to retain and then delete, to retain only, or delete only.

When the retention policy is to retain and then delete:



For the two paths in the diagram:

1. **If a chat or channel message is edited or deleted** by a user during the retention period, the original message is copied (if edited) or moved (if deleted) to the SubstrateHolds folder. When a user deletes a Teams message, although the message disappears from the Teams app, the message doesn't go into the SubstrateHolds folder for 21 days. The message is stored in the SubstrateHolds folder for at least 1 day. When the retention period expires, the message is permanently deleted the next time the timer job runs (typically between 1-7 days).

2. **If a chat or channel message is not deleted** by a user and for current messages after editing, the message is moved to the SubstrateHolds folder after the retention period expires. This action typically takes between 1-7 days from the expiry date. When the message is in the SubstrateHolds folder, it's stored there for at least 1 day, and then the message is permanently deleted the next time the timer job runs (typically between 1-7 days).

> ⓘ **Note**
>
> Messages stored in mailboxes, including the hidden folders, are searchable by eDiscovery tools. Until messages are permanently deleted from the SubstrateHolds folder, they remain searchable by eDiscovery tools.

When the retention period expires and moves a message to the SubstrateHolds folder, a delete operation is communicated to the backend Azure chat service, that then relays the same operation to the Teams client app. Delays in this communication or caching can explain why, for a short period of time, users continue to see these messages in their Teams app.

In this scenario where the Azure chat service receives a delete command because of a retention policy, the corresponding message in the Teams client app is deleted for all users in the conversation. Sometimes, this behavior might seem unexpected because some of these users can be from another organization, have a retention policy with a longer retention period, or no retention policy assigned to them. For these users, copies of the messages are still stored in their mailboxes and remain searchable for users in Teams and admins using eDiscovery until the messages are permanently deleted by another retention policy. Users won't be able to see the message in Teams.

> ⓘ **Important**

Messages visible in the Teams app are not an accurate reflection of whether they're retained or permanently deleted for compliance requirements.

When the retention policy is retain-only, or delete-only, the content's paths are variations of retain and delete.

Content paths for retain-only retention policy

1. **If a chat or channel message is edited or deleted** by a user during the retention period: The original message is copied (if edited) or moved (if deleted) to the SubstrateHolds folder. When a user deletes a Teams message, although the message disappears from the Teams app, the message doesn't go into the SubstrateHolds folder for 21 days. The message is stored in the SubstrateHolds folder for at least 1 day. If the retention policy is configured to retain forever, the item remains there. If the retention policy has an end date for the retention period and it expires, the message is permanently deleted the next time the timer job runs (typically between 1-7 days).

2. **If the chat or channel message is not modified or deleted** by a user and for current messages after editing during the retention period: Nothing happens before and after the retention period; the message remains in its original location.

Content paths for delete-only retention policy

1. **If the chat or channel message is edited or deleted** by a user during the retention period: The original message is copied (if edited) or moved (if deleted) to the SubstrateHolds folder. When a user deletes a Teams message, although the message disappears from the Teams app, the message doesn't go into the SubstrateHolds folder for 21 days. The message is stored in the SubstrateHolds folder for at least 1 day and permanently deleted the next time the timer job runs (typically between 1-7 days).

2. **If a chat or channel message is not deleted** by a user during the retention period: At the end of the retention period, the message is moved to the SubstrateHolds folder. This action typically takes between 1-7 days from the expiry date. The message is retained there for at least 1 day and then permanently deleted the next time the timer job runs (typically between 1-7 days).

Example flows and timings for retention policies

Use the following examples to see how the processes and timings explained in the previous sections apply to retention policies that have the following configurations:

- Example 1: Retain-only for 7 years
- Example 2: Retain for 30 days and then delete
- Example 3: Delete-only after 1 day

For all examples that refer to permanent deletion, because of the principles of retention, this action is suspended if the message is subject to another retention policy to retain the item or it's subject to an eDiscovery hold.

Example 1: Retain-only for 7 years

On day 1, a user creates a chat or channel message.

On day 5, the user edits that message.

On day 30, the user deletes the current message.

Retention outcomes:

- For the original message:
  - On day 5, the message is copied to the SubstrateHolds folder where it can still be searched with eDiscovery tools for a minimum of 7 years from day 1 (the retention period).
- For the current (edited) message:
  - On day 30, the message is no longer displayed in the Teams app and moves to the SubstrateHolds folder after 21 days where it continues to be searchable with eDiscovery tools for a minimum of 7 years from day 1 (the retention period).

If the user had deleted the current message after the specified retention period, instead of within the retention period, the message would still be moved to the SubstrateHolds folder after 21 days. However, now the retention period has expired, the message would be permanently deleted there after the minimum of 1 day and then typically within 1-7 days.

Example 2: Retain for 30 days and then delete

On day 1, a user creates a chat or channel message.

On day 10, the user edits that message.

The user doesn't make further edits and doesn't delete the message.

Retention outcomes:

- For the original message:
  - On day 10, the message is copied to the SubstrateHolds folder, where it can still be searched with eDiscovery tools.
  - At the end of the retention period (30 days from day 1), the message is permanently deleted typically within 1-7 days after the minimum of 1 day, and then won't be returned with eDiscovery searches.

- For the current (edited) message:
  - At the end of the retention period (30 days from day 1), the message moves to the SubstrateHolds folder typically within 1-7 days, where it can still be searched with eDiscovery tools.
  - The message is then permanently deleted typically within 1-7 days after the minimum of 1 day, and then won't be returned with eDiscovery searches.

Example 3: Delete-only after 1 day

> **Note**
>
> Because of the short one-day duration of this configuration and retention processes that operate within a time period of 1-7 days, this section shows example timings that are within the typical time ranges.

On day 1, a user creates a chat or channel message.

Example retention outcome if the user doesn't edit or delete the message:

- Day 5 (typically 1-7 days after the start of the retention period on day 2):
  - The message moves to the SubstrateHolds folder and remains there for at least 1 day where it can still be searched with eDiscovery tools.

- Day 9 (typically 1-7 days after a minimum of 1 day in the SubstrateHolds folder):
  - The message is permanently deleted and then won't be returned with eDiscovery searches.

As this example shows, although you can configure a retention policy to delete messages after just one day, the service undergoes multiple processes to ensure a compliant deletion. As a result, a delete action after 1 day could take 16 days before the message is permanently deleted so that it's no longer returned in eDiscovery searches.

How retention works with Microsoft Teams call logs

By default, Teams call logs are retained indefinitely. Use a retention policy for call logs if you need to delete this data for compliance reasons.

The deletion flow and process is the same as for Teams messages, only simplified because the call-related records in call logs can't be edited or deleted by users, and because deletion is the only applicable action. In the previous diagram and explanations, Teams call logs follow the "Content paths for delete-only retention policy" and the same behavior as "If a chat or channel message is not deleted".

As a result:

- At the end of the retention period, the call-related record is moved to the SubstrateHolds folder. This action typically takes between 1-7 days from the expiry date. The record is retained there for at least 1 day and then permanently deleted the next time the timer job runs (typically between 1-7 days).

From the previous section, only Example 3: Delete-only after 1 day is applicable for Teams call logs. Substitute "call-related record" for the word "message" and the behavior and timings remain the same.

Skype for Business and Teams interop chats

When a Skype for Business chat comes into Teams, it becomes a message in a Teams chat thread and is ingested into the appropriate mailbox. Teams retention policies apply to these messages from the Teams thread.

However, if conversation history is turned on for Skype for Business and from the Skype for Business client side that history is being saved into a mailbox, that chat data isn't handled by a Teams retention policy. For this content, use a retention policy that's configured for Skype for Business.

Messages and external users

When external users are included in a meeting or chat that your organization hosts:

- If an external user joins by using a guest account in your tenant, any Teams messages are stored in both your users' mailbox and a shadow mailbox that's granted to the guest account. However, retention policies aren't supported for shadow mailboxes, even though

they can be reported as included in a retention policy for the entire location (sometimes known as an "organization-wide policy").

- If an external user joins by using an account from another Microsoft 365 organization, your retention policies can't delete messages for this user because they're stored in that user's mailbox in another tenant. For the same meeting or chat however, your retention policies can delete messages for your users.

When a user leaves the organization

If a user who has a mailbox in Exchange Online leaves your organization and their Microsoft 365 account is deleted, their chat messages that are subject to retention are stored in an inactive mailbox. The chat messages remain subject to any retention policy that was placed on the user before their mailbox was made inactive, and the contents are available to an eDiscovery search. For more information, see Learn about inactive mailboxes.

If the user stored any files in Teams, see the equivalent section for SharePoint and OneDrive.

Configuration guidance

If you're new to configuring retention in Microsoft 365, see Get started with data lifecycle management.

If you're ready to configure a retention policy for Teams, see Create and configure retention policies.

Last updated on 06/03/2026

Learn about retention for Viva Engage

> *Microsoft Purview service description*

The information in this article supplements Learn about retention because it has information that's specific to Viva Engage.

For other workloads, see:

- Learn about retention for SharePoint and OneDrive
- Learn about retention for Microsoft Teams
- Learn about retention for Exchange
- Learn about retention for Copilot

What's included for retention and deletion

Viva Engage user messages and community messages can be deleted by using retention policies for Viva Engage, and in addition to the text in these messages, the following items can be retained for compliance reasons: Hypertext links and links to other Viva Engage messages.

> **Note**
>
> As explained in the following section, user messages include private messages for an individual user, and any community messages associated with that user.
>
> User messages also include **storyline posts**, which are supported by retention policies.

User messages include all the names of the people in the conversation, and community messages include the community name and the message title (if supplied).

Reactions from others in the form of emoticons aren't retained when you use retention policies for Viva Engage.

Files that you use with Viva Engage aren't included in retention policies for Viva Engage. These items have their own retention policies.

How retention works with Viva Engage

Use this section to understand how your compliance requirements are met by backend storage and processes, and should be verified by eDiscovery tools rather than by messages that are currently visible in the Viva Engage app.

You can use a retention policy to retain data from community messages and user messages in Viva Engage, and delete these messages. Behind the scenes, Exchange mailboxes are used to store data copied from these messages. Data from Viva Engage user messages is stored in a hidden folder in the mailbox of each user included in the user message, and a similar hidden folder in a group mailbox is used for community messages.

Copies of community messages can also be stored in the hidden folder of user mailboxes when they @ mention users or notify the user of a reply. Although these messages originate as a community message, a retention policy for Viva Engage user messages will often include copies of community messages. As a result, user messages aren't restricted to private messages.

These hidden folders aren't designed to be directly accessible to users or administrators, but instead, store data that compliance administrators can search with eDiscovery tools.

Even though they're stored in Exchange, Viva Engage messages are only included in a retention policy that's configured for the **Viva Engage community messages** or **Viva Engage user messages** locations.

> ⓘ **Note**
>
> If a user is included in an active retention policy that retains Viva Engage data and you delete a mailbox of a user who is included in this policy, to retain the Viva Engage data, the mailbox is converted into an inactive mailbox. If you don't need to retain this Viva Engage data for the user, exclude the user account from the retention policy before you delete their mailbox.

After a retention policy is configured for Viva Engage messages, a timer job from the Exchange service periodically evaluates items in the hidden folder where these Viva Engage messages are stored. The timer job takes up to seven days to run. When these items have expired their retention period, they're moved to the SubstrateHolds folder—a hidden folder that's in every user or group mailbox to store "soft-deleted" items before they're permanently deleted.

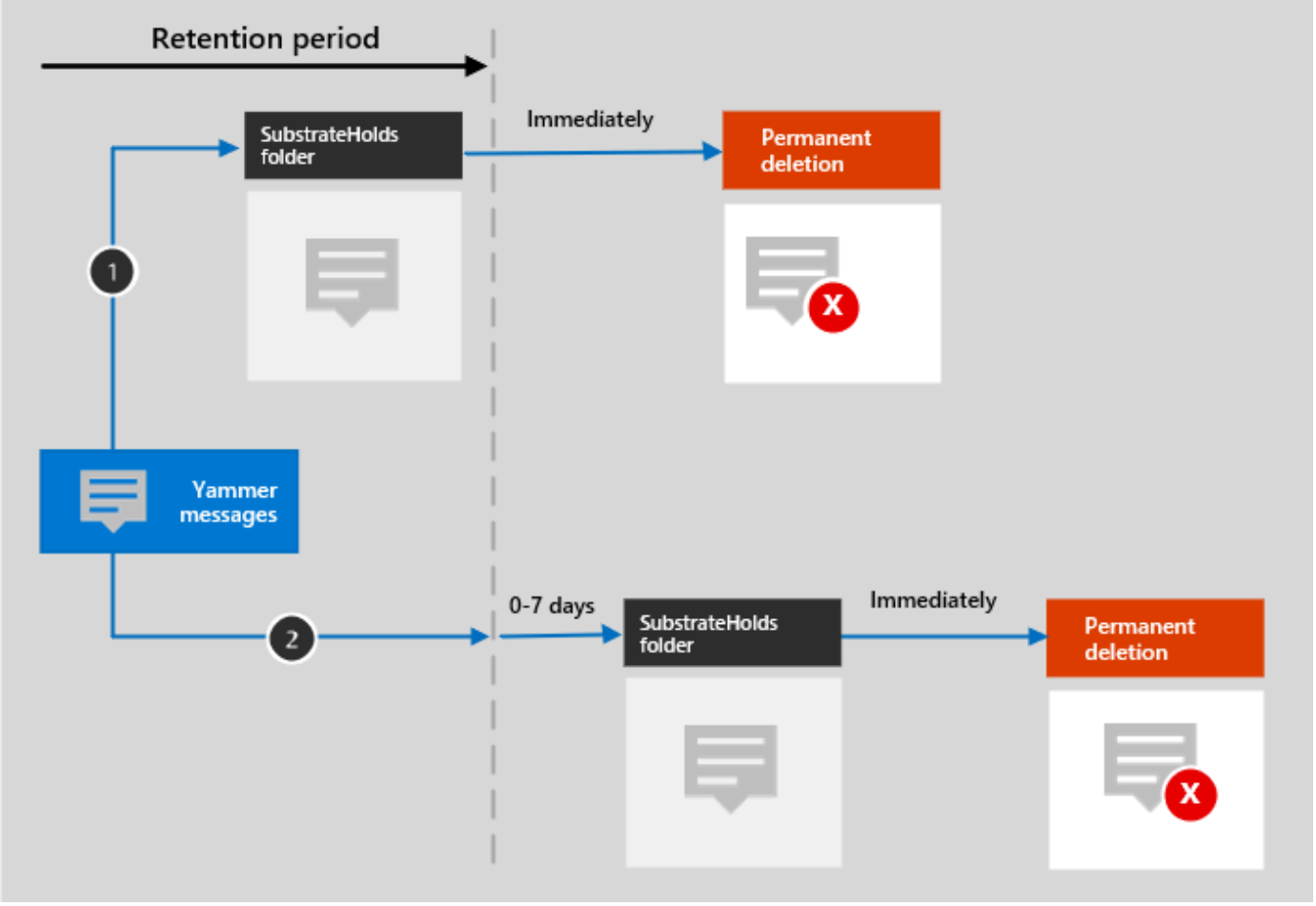
> ⓘ **Important**

Because of the first principle of retention and since Viva Engage messages are stored in Exchange Online mailboxes, permanent deletion from the SubstrateHolds folder is always suspended if the mailbox is affected by another Viva Engage retention policy for the same location, Litigation Hold, delay hold, or if an eDiscovery hold is applied to the mailbox for legal or investigative reasons.

While the mailbox is included in an applicable hold, Viva Engage messages that are deleted will no longer be visible in Viva Engage but continue to be discoverable with eDiscovery.

After a retention policy is configured for Viva Engage messages, the paths the content takes depend on whether the retention policy is to retain and then delete, to retain only, or delete only.

When the retention policy is to retain and then delete:



For the two paths in the diagram:

1. **If a Viva Engage message is edited or deleted** by the user during the retention period, the original message is immediately copied (if edited) or moved (if deleted) to the SubstrateHolds folder. The message is stored there until the retention period expires and then the message is immediately permanently deleted.

2. **If a Viva Engage message is not deleted** and for current messages after editing, the message is moved to the SubstrateHolds folder after the retention period expires. This action takes up to seven days from the expiry date. When the message is in the SubstrateHolds folder, it's then immediately permanently deleted.

> **Note**
>
> Messages in the SubstrateHolds folder are searchable by eDiscovery tools. Until messages are permanently deleted from the SubstrateHolds folder, they remain searchable by eDiscovery tools.

When the retention period expires and moves a message to the SubstrateHolds folder, a delete operation is communicated to the Viva Engage service, that then relays the same operation to the Viva Engage client app. Delays in this communication or caching can explain why, for a short period of time, users continue to see these messages in their Viva Engage app.

In this scenario where the Viva Engage service receives a delete command because of a retention policy, the corresponding message in the Viva Engage app is deleted for all users in the conversation. Some of these users might be from another organization, have a retention policy with a longer retention period, or no retention policy assigned to them. For these users, copies of the messages are still stored in their mailboxes and remain searchable for eDiscovery until the messages are permanently deleted by another retention policy.

> **Important**
>
> Messages visible in the Viva Engage app aren't an accurate reflection of whether they're retained or permanently deleted for compliance requirements.

When the retention policy is retain-only, or delete-only, the content's paths are variations of retain and delete.

Content paths for retain-only retention policy

1. **If a Viva Engage message is edited or deleted**: A copy of the original message is immediately created in the SubstrateHolds folder and retained there until the retention period expires. Then the message is immediately permanently deleted from the SubstrateHolds folder.

2. **If the Viva Engage message is not modified or deleted** and for current messages after editing during the retention period: Nothing happens before and after the retention period; the message remains in its original location.

Content paths for delete-only retention policy

1. **If the Viva Engage message is not deleted** during the retention period: At the end of the retention period, the message is moved to the SubstrateHolds folder. This action takes up to seven days from the expiry date. Then the message is immediately permanently deleted from the SubstrateHolds folder.

2. **If the Viva Engage message is deleted by the user** during the period, the item is immediately moved to the SubstrateHolds folder where it's immediately permanently deleted.

Example flows and timings for retention policies

Use the following examples to see how the processes and timings explained in the previous sections apply to retention policies that have the following configurations:

- Example 1: Retain-only for 7 years
- Example 2: Retain for 30 days and then delete
- Example 3: Delete-only after 1 day

For all examples that refer to permanent deletion, because of the principles of retention, this action is suspended if the message is subject to another retention policy to retain the item or it's subject to an eDiscovery hold.

Example 1: Retain-only for 7 years

On day 1, a user posts a new Viva Engage message.

On day 5, the user edits that message.

On day 30, the user deletes the current message.

Retention outcomes:

- For the original message:
  - On day 5, the message is copied to the SubstrateHolds folder where it can still be searched with eDiscovery tools for a minimum of 7 years from day 1 (the retention period).

- For the current (edited) message:
  - On day 30, the message moves to the SubstrateHolds folder where it can still be searched with eDiscovery tools for a minimum of 7 years from day 1 (the retention period).

If the user had deleted the current message after the specified retention period, instead of within the retention period, the message would still be moved to the SubstrateHolds folder. However, now the retention period has expired, the message would be permanently deleted after the minimum of 1 day and then typically within 1-7 days.

Example 2: Retain for 30 days and then delete

On day 1, a user posts a new Viva Engage message.

On day 10, the user edits that message.

The user doesn't make further edits and doesn't delete the message.

Retention outcomes:

- For the original message:
  - On day 10, the message is copied to the SubstrateHolds folder, where it can still be searched with eDiscovery tools.
  - At the end of the retention period (30 days from day 1), the message is permanently deleted typically within 1-7 days after the minimum of 1 day, and then won't be returned with eDiscovery searches.
- For the current (edited) message:
  - At the end of the retention period (30 days from day 1), the message moves to the SubstrateHolds folder typically within 1-7 days, where it can still be searched with eDiscovery tools.
  - The message is then permanently deleted typically within 1-7 days after the minimum of 1 day, and then won't be returned with eDiscovery searches.

Example 3: Delete-only after 1 day

> **Note**

Because of the short one-day duration of this configuration and retention processes that operate within a time period of 1-7 days, this section shows example timings that are within the typical time ranges.

On day 1, a user posts a new Viva Engage message.

Example retention outcome if the user doesn't edit or delete the message:

- Day 5 (typically 1-7 days after the start of the retention period on day 2):
  - The message moves to the SubstrateHolds folder and remains there for at least 1 day where it can still be searched with eDiscovery tools.
- Day 9 (typically 1-7 days after a minimum of 1 day in the SubstrateHolds folder):
  - The message is permanently deleted and then won't be returned with eDiscovery searches.

As this example shows, although you can configure a retention policy to delete messages after just one day, the service undergoes multiple processes to ensure a compliant deletion. As a result, a delete action after 1 day could take 16 days before the message is permanently deleted so that it's no longer returned in eDiscovery searches.

Messages and external users

By default, a retention policy for Viva Engage user messages applies to all users in your organization, but not external users. You can apply a retention policy to external users if you use the **Edit** option for users included, and specify their account.

At this time, Azure B2B guest users aren't supported.

When a user leaves the organization

If a user leaves your organization and their Microsoft 365 account is deleted, their Viva Engage user messages that are subject to retention are stored in an inactive mailbox. These messages remain subject to any retention policy that was placed on the user before their mailbox was made inactive, and the contents are available to an eDiscovery search. For more information, see Learn about inactive mailboxes.

If the user stored any files in Viva Engage, see the equivalent section for SharePoint and OneDrive.

Limitations

Be aware of the following limitation when you use retention for Viva Engage community messages and user messages:

- When you select **Edit** for the **Viva Engage user messages** location, you might see guests and nonmailbox users. Retention policies aren't designed for these users, so don't select them.

Configuration guidance

If you're new to configuring retention in Microsoft 365, see Get started with data lifecycle management.

If you're ready to configure a retention policy for Viva Engage, see Create and configure retention policies.

Last updated on 09/15/2024

Learn about retention for Exchange

09/26/2025

> *Microsoft 365 licensing guidance for security & compliance*.

The information in this article supplements Learn about retention because it has information that's specific to Exchange. For other workloads, see:

- Learn about retention for SharePoint and OneDrive
- Learn about retention for Microsoft Teams
- Learn about retention for Viva Engage
- Learn about retention for Copilot

What's included for retention and deletion

The following Exchange items from user mailboxes and shared mailboxes can be retained and deleted by using retention policies and retention labels: Mail messages (includes received messages, drafts, sent messages) with any attachments, tasks when they have an end date, and notes.

Calendar items that have an end date are supported for retention policies but aren't supported for retention labels.

Public folders are supported for retention policies but aren't supported for retention labels.

Contacts, and any tasks and calendar items that don't have an end date are not supported.

Other items stored in a mailbox, such as Skype and Teams messages, aren't included in retention policies or labels for Exchange. These items have their own retention policies.

Mailboxes must have at least 10 MB of data before retention settings will apply to them, and retention labels can be published to them.

How retention works for Exchange

Both a mailbox and a public folder use the Recoverable Items folder to retain items. Only people who have been assigned eDiscovery permissions can view items in another user's Recoverable Items folder.

When a user deletes a message in a folder other than the Deleted Items folder, by default, the message moves to the Deleted Items folder. However, a user can soft delete an item

(Shift+Delete) in any folder, which bypasses the Deleted Items folder and moves the item directly to the Recoverable Items folder.

When you apply retention settings to Exchange data, a timer job periodically evaluates items in the Recoverable Items folder. If an item doesn't match the rules of at least one retention policy or retention label to retain the item, it is permanently deleted (also called hard deleted) from the Recoverable Items folder. Items configured for disposition review are never permanently deleted from the Recoverable Items folder until the disposition is confirmed.

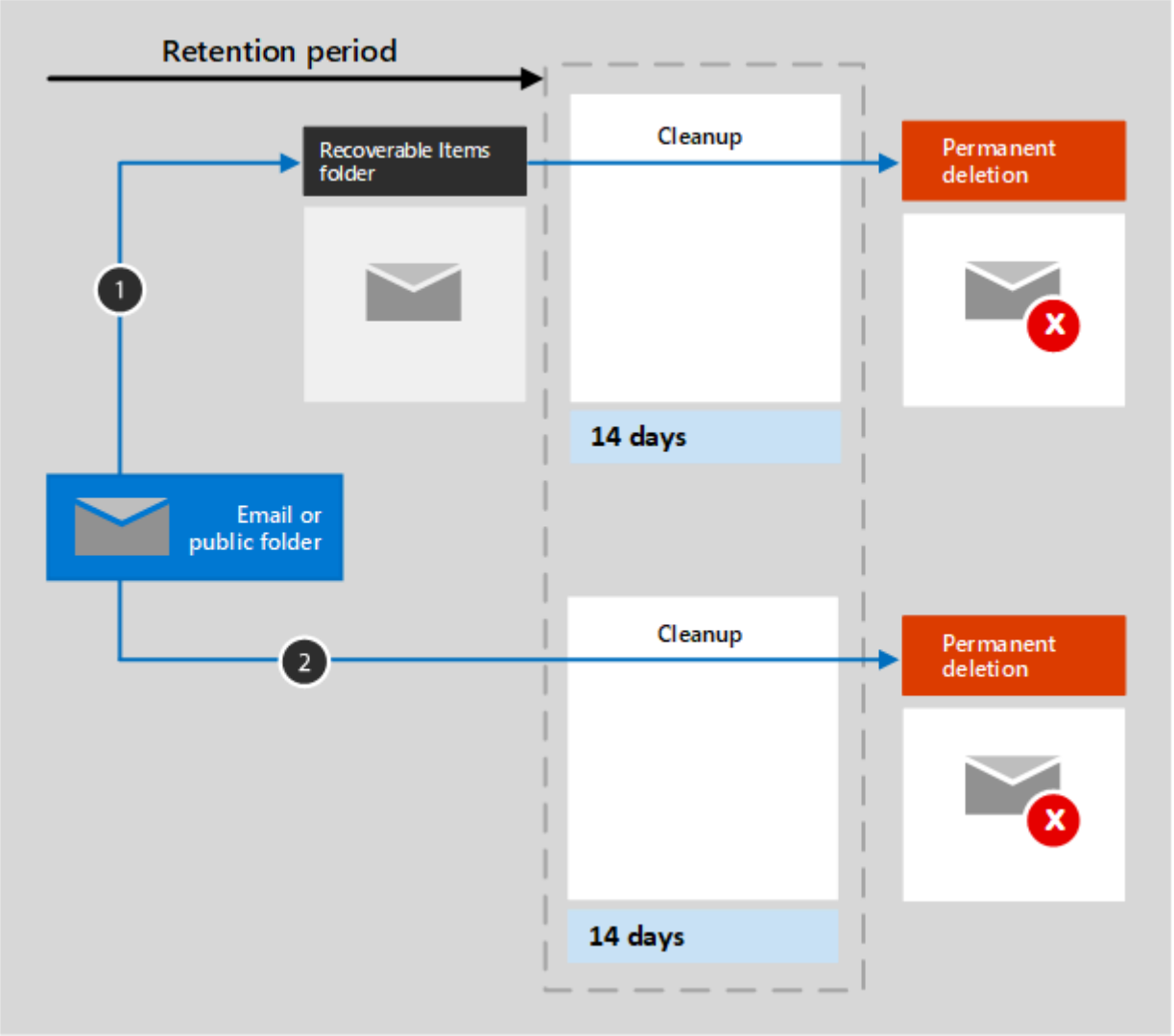
> **Note**
>
> Because of the first principle of retention, permanent deletion is always suspended if the same item must be retained because of another retention policy or retention label, or it is under eDiscovery holds for legal or investigative reasons.

The timer job can take up to seven days to run and the Exchange location must contain at least 10 MB.

When a user attempts to change properties of a mailbox item—such as the subject, body, attachments, senders and recipients, or date sent or received for a message—a copy of the original item is saved to the Recoverable Items folder before the change is committed. This action happens for each subsequent change. At the end of the retention period, copies in the Recoverable Items folder are permanently deleted.

After retention settings are applied to Exchange content, the paths the content takes depend on whether the retention settings are to retain and delete, to retain only, or delete only.

When the retention settings are to retain and delete:



1. **If the item is modified or permanently deleted** by the user (either SHIFT+DELETE or deleted from Deleted Items) during the retention period: The item is moved (or copied, in the case of edit) to the Recoverable Items folder. There, a timer job runs periodically and identifies items whose retention period has expired, and these items are permanently deleted within 14 days of the end of the retention period. Note that 14 days is the default setting, but it can be configured up to 30 days.

2. **If the item is not modified or deleted** during the retention period: The same process runs periodically on all folders in the mailbox and identifies items whose retention period has expired, and these items are permanently deleted within 14 days of the end of the retention period. Note that 14 days is the default setting, but it can be configured up to 30 days.

When the retention settings are retain-only, or delete-only, the contents paths are variations of retain and delete:

Content paths for retain-only retention settings

1. **If the item is modified or deleted** during the retention period: A copy of the original item is created in the Recoverable Items folder and retained until the end of the retention period, when the copy in the Recoverable Items folder is permanently deleted within 14 days after the item expires.

2. **If the item is not modified or deleted** during the retention period: Nothing happens before and after the retention period; the item remains in its original location.

Content paths for delete-only retention settings

1. **If the item is not deleted** during the configured period: At the end of the configured period in the retention policy, the item is moved to the Recoverable Items folder.

2. **If the item is deleted** during the configured period: The item is immediately moved to the Recoverable Items folder. If a user deletes the item from there or empties the Recoverable Items folder, the item is permanently deleted. Otherwise, the item is permanently deleted after being in the Recoverable Items folder for 14 days.

User notification of expiry date

Retention policies for Exchange, unlike retention policies for the other Microsoft 365 workloads, have a user presence by displaying at the top of each email message the name of the retention policy that has the shortest expiry date for the item, and the calculated expiry date for that item. Users don't see this notification if the retention policy doesn't delete items (retain-only).

If a retention label is applied to an email message, the name of that label and corresponding expiry date is always displayed, and will replace the name and date from any retention policy applied to the mailbox.

Remember, in this context, the expiry date for when an email will be deleted is when users can expect the email message to automatically move to the Recoverable Items folder (if not already there). Emails in the Recoverable Items folder will not be permanently deleted but remain there for compliance purposes if they are subject to any retention settings to retain it, or they are under an eDiscovery hold for legal or investigative reasons.

When a user leaves the organization

If a user leaves your organization and the user's mailbox is included in a policy for retention, the mailbox becomes an inactive mailbox when the user's Microsoft 365 account is deleted. The contents of an inactive mailbox are still subject to any retention policy that was placed on

the mailbox before it was made inactive, and the contents are available to an eDiscovery search. For more information, see Inactive mailboxes in Exchange Online.

When the retention settings no longer apply because the data is permanently deleted or the retention period has expired, the Exchange admin can now delete the inactive mailbox. In this scenario, the inactive mailbox isn't automatically deleted.

Configuration guidance

If you're new to configuring retention in Microsoft 365, see Get started with data lifecycle management.

If you're ready to configure a retention policy or retention label for Exchange, see the following instructions:

- Create and configure retention policies
- Publish retention labels and apply them in apps
- Apply a retention label to content automatically

Learn about retention for Copilot & AI apps

09/26/2025

> *Microsoft 365 licensing guidance for security & compliance*.

> **Note**
>
> Previously, messages from Microsoft 365 Copilot and Microsoft 365 Copilot Chat were automatically included in the retention policy location named **Teams chats and Copilot interactions** because they were retained and deleted by using the same mechanisms. Users didn't have to be using Teams for the retention policy to apply to Copilot.
>
> Retention policies for Microsoft 365 Copilot and Microsoft 365 Copilot Chat are now separate from Teams chats, and newly created retention policies support the following locations:
>
> - **Microsoft Copilot experiences**
> - Microsoft 365 Copilot
> - Security Copilot
> - Copilot in Fabric
> - Copilot Studio
> - **Enterprise AI apps**
> - Entra-registered AI apps
> - ChatGPT Enterprise
> - Azure AI services
> - **Other AI apps**
> - ChatGPT
> - Google Gemini
> - Microsoft Copilot (consumer version)
> - DeepSeek

The information in this article supplements Learn about retention because it has information that's specific to interactions with Copilot & AI apps.

For other workloads, see:

- Learn about retention for SharePoint and OneDrive
- Learn about retention for Viva Engage
- Learn about retention for Exchange

- Learn about retention for Teams

For more information about Microsoft Purview integration with Copilot & AI apps, see Microsoft Purview data security and compliance protections for generative AI apps.

What's included for retention and deletion

Retention policies for AI apps include user prompts and responses for Microsoft 365 Copilot and Copilot Studio, and also user prompts and responses for other Microsoft Copilots and generative AI apps when they have a collection policy with the setting to capture content. These messages can then be retained and deleted for compliance reasons.

User prompts include text that users type, and selecting AI app prompts that are captured as a prepopulated message. AI app responses include text, links, and references. For Copilot in Teams chat and channels, responses include threaded summaries. Because messages to indicate that a response is in progress don't have business value, these messages aren't captured.

How retention works with AI apps

Use this section to understand how your compliance requirements are met by backend storage and processes, and should be verified by eDiscovery tools rather than by messages that are currently visible in your AI apps, such as Copilot.

You can use a retention policy to retain data from messages in Copilot or another AI app, and delete those messages. Behind the scenes, Exchange mailboxes are used to store data copied from these messages. Data from generative AI messages is stored in a hidden folder in the mailbox of the user who runs the AI app. This hidden folder isn't designed to be directly accessible to users or administrators, but instead, store data that compliance administrators can search with eDiscovery tools.

The Exchange mailbox for retaining these messages has the RecipientTypeDetails attribute of **UserMailbox**, which also stores message data for Teams private channels and cloud-based Teams users.

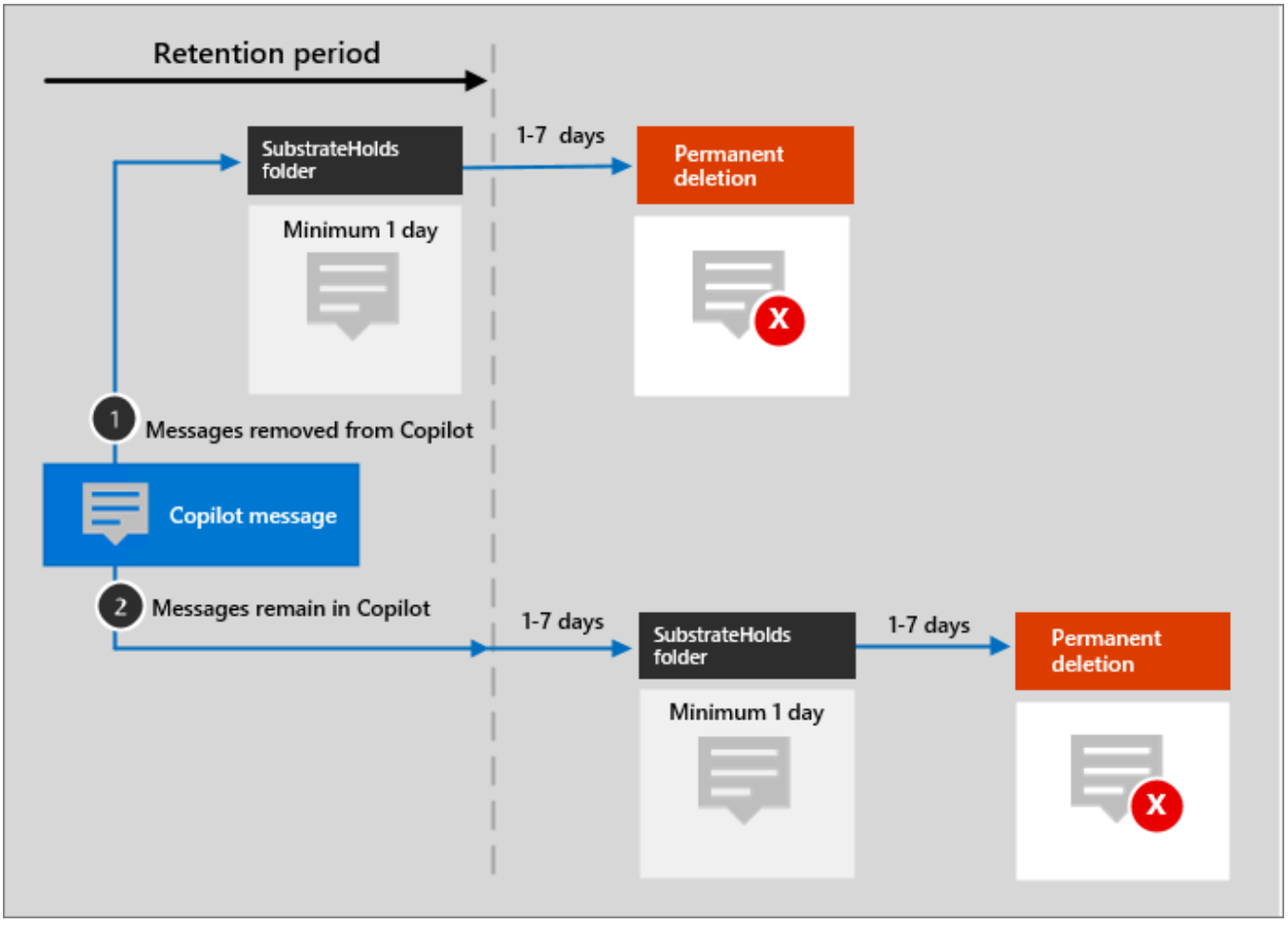
After a retention policy is configured for AI app interactions, a timer job from the Exchange service periodically evaluates items in the hidden mailbox folder where these messages are stored. The timer job typically takes 1-7 days to run. When these items have expired their retention period, they're moved to the SubstrateHolds folder—another hidden folder that's in every user mailbox to store "soft-deleted" items before they're permanently deleted.

Messages remain in the SubstrateHolds folder for at least 1 day, and then if they're eligible for deletion, the timer job permanently deletes them the next time it runs.

> **ⓘ Important**
>
> Because of the [first principle of retention], and because the AI app messages are stored in Exchange Online mailboxes, permanent deletion from the SubstrateHolds folder is always suspended if the mailbox is affected by another retention policy for the same location, Litigation Hold, delay hold, or if an eDiscovery hold is applied to the mailbox for legal or investigative reasons.

After a retention policy is configured, the paths the content takes depend on whether the retention policy is to retain and then delete, to retain only, or delete only.

When the retention policy is to retain and then delete, using Microsoft 365 Copilot as the example:



In most scenarios, these messages aren't removed. For example, they remain but are hidden when users close a chat window or close the app. However, Copilot messages are removed in the following scenarios:

- Users delete (when this option is available) the associated chat in Microsoft 365 Copilot Chat.
- A request is submitted to delete a user's history of all interactions with Microsoft 365 Copilot.

For the two paths in the diagram:

1. **If messages are removed from the AI app**, the message is moved to the SubstrateHolds folder where it remains for at least 1 day. When the retention period expires, the message is permanently deleted the next time the timer job runs (typically between 1-7 days).

2. **If messages remain in the AI app** after the retention period expires, the message is copied to the SubstrateHolds folder. This action typically takes between 1-7 days from the expiry date. When the message is in the SubstrateHolds folder, it's stored there for at least 1 day, and then the message is permanently deleted the next time the timer job runs (typically between 1-7 days).

> **Note**
>
> Messages stored in mailboxes, including the hidden folders, are searchable by eDiscovery tools. Until messages are permanently deleted from the SubstrateHolds folder, they remain searchable by eDiscovery tools.

When the retention period expires and copies a message to the SubstrateHolds folder, a delete operation is communicated to the backend service for Copilot, that then relays the same operation to the user app with Copilot. Delays in this communication or caching can explain why, for a short period of time, users continue to see these messages in Copilot.

> **Important**
>
> Messages visible in your AI apps are not an accurate reflection of whether they are retained or permanently deleted for compliance requirements.

When the retention policy is retain-only, or delete-only, the content's paths are variations of retain and delete.

Content paths for retain-only retention policy

1. **If messages are removed from your AI app** the message is moved to the SubstrateHolds folder after the retention period expires. This action typically takes between 1-7 days from the expiry date. If the retention policy is configured to retain forever, the item remains

there. If the retention policy has an end date for the retention period and it expires, the message is permanently deleted the next time the timer job runs (typically between 1-7 days).

2. **If messages remain in your AI app** after the retention period expires, nothing happens before and after the retention period; the message remains in its original location.

Content paths for delete-only retention policy

1. **If messages are removed from your AI app** during the retention period, the message is moved to the SubstrateHolds folder. The message is stored in the SubstrateHolds folder for at least 1 day and permanently deleted the next time the timer job runs (typically between 1-7 days).

2. **If messages remain in your AI app** after the retention period expires, the message is copied to the SubstrateHolds folder. This action typically takes between 1-7 days from the expiry date. The message is retained there for at least 1 day and then permanently deleted the next time the timer job runs (typically between 1-7 days).

Example flows and timings for retention policies

Use the following examples to see how the processes and timings explained in the previous sections apply to retention policies that have the following configurations:

- Example 1: Retain for 30 days and then delete
- Example 2: Delete-only after 1 day

For all examples that refer to permanent deletion, because of the principles of retention, this action is suspended if the message is subject to another retention policy to retain the item or it's subject to an eDiscovery hold.

Example 1: Retain for 30 days and then delete

On day 1, a user sends a prompt to Copilot and the prompt is removed after 10 days.

Retention outcome:

- After day 10, the message is moved to the SubstrateHolds folder, where it can still be searched with eDiscovery tools.
- At the end of the retention period (30 days from day 1), the message is permanently deleted typically within 1-7 days after the minimum of 1 day, and then won't be returned with eDiscovery searches.

Example 2: Delete-only after 1 day

> **Note**
>
> Because of the short one-day duration of this configuration and retention processes that operate within a time period of 1-7 days, this section shows example timings that are within the typical time ranges.

On day 1, a user sends a prompt to Copilot and this prompt isn't removed from Copilot.

Example retention outcome if the user's prompt isn't removed:

- Day 5 (typically 1-7 days after the start of the retention period on day 2):
  - The message is copied to the SubstrateHolds folder and remains there for at least 1 day.
- Day 9 (typically 1-7 days after a minimum of 1 day in the SubstrateHolds folder):
  - The message is permanently deleted and then won't be returned with eDiscovery searches.

As this example shows, although you can configure a retention policy to delete messages after just one day, the service undergoes multiple processes to ensure a compliant deletion. As a result, a delete action after 1 day could take 16 days before the message is permanently deleted so that it's no longer returned in eDiscovery searches.

When a user leaves the organization

If a user leaves your organization and their Microsoft 365 account is deleted, their Copilot and other AI app messages that are subject to retention are stored in an inactive mailbox. The messages remain subject to any retention policy that was placed on the user before their mailbox was made inactive, and the contents are available to an eDiscovery search. For more information, see Learn about inactive mailboxes.

Configuration guidance

If you're new to configuring retention in Microsoft 365, see Get started with data lifecycle management.

If you're ready to configure a retention policy for Copilot interactions, see Create and configure retention policies.

Common settings for retention policies and retention label policies

> *Microsoft 365 licensing guidance for security & compliance*.

Many settings for retention are common to both retention policies and retention label policies. Use the following information to help you configure these settings to proactively retain content, delete content, or both—retain and then delete the content.

For the scenarios that support these policies for retention, see:

- Create and configure retention policies.
- Publish retention labels and apply them in apps
- Apply a retention label to content automatically

Settings that are specific to each scenario are explained in their respective documentation.

For overview information about policies for retention and how retention works in Microsoft 365, see Learn about retention policies and retention labels.

Scopes - adaptive and static

If you're unfamiliar with adaptive and static scopes, and to help you choose which one to use when you configure a policy for retention, see Adaptive or static policy scopes for retention.

When you've decided whether to use an adaptive or static scope, use the following information to help you configure it:

- Configuration information for adaptive scopes
- Configuration information for static scopes

> **Tip**
>
> If you have policies that use static scopes and you want to convert them to adaptive scopes, leave your existing policies in place while you create new policies that use adaptive scopes with the same retention settings. Validate these new policies are targeting the correct users, sites, and groups before you disable or delete the old policies with static scopes.

Configuration information for static scopes

When you choose to use static scopes, you must then decide whether to apply the policy to all instances for the selected location (the entire location) or to include or exclude specific instances (specific inclusions or exclusions).

A policy that applies to entire locations

Except Skype for Business, the default is that all instances for the selected locations are automatically included in the policy without you having to specify them as included.

For example, **All mailboxes** for the **Exchange email** location. With this default setting, all existing user mailboxes will be included in the policy, and any new mailboxes created after the policy is applied will automatically inherit the policy.

For locations that include SharePoint sites, all instances include archived sites as well as active sites, and all containers in SharePoint Embedded for Loop workspaces.

A policy with specific inclusions or exclusions

Be aware that if you use the optional configuration to scope your retention settings to specific users, specific Microsoft 365 groups, or specific sites, there are some limits per policy to be aware of. For more information, see Limits for retention policies and retention label policies.

To use the optional configuration to scope your retention settings, make sure the **Status** of that location is **On**, and then use the links to include or exclude specific users, Microsoft 365 groups, or sites.

> ⚠ **Warning**
>
> If you configure instances to include and then remove the last one, the configuration reverts to **All** for the location. Make sure this is the configuration that you intend before you save the policy.
>
> For example, if you specify one SharePoint site to include in your retention policy that's configured to delete data, and then remove the single site, by default all SharePoint sites will then be subject to the retention policy that permanently deletes data. The same applies to includes for Exchange mailboxes, OneDrive accounts, Teams chat users, and so on.
>
> In this scenario, toggle the location off if you don't want the **All** setting for the location to be subject to the retention policy. Alternatively, specify exclude instances to be exempt from

the policy.

Locations

Locations in policies for retention identify specific Microsoft 365 services that support retention settings, such as Exchange email and SharePoint sites. Use the following section for the locations that have configuration details and possible exceptions that you need to be aware of when you select them for your policy.

Configuration information for Exchange mailboxes and Exchange public folders

Both the **Exchange mailboxes** location and the **Exchange public folders** location require mailboxes to have at least 10 MB of data before retention settings will apply to them.

The **Exchange mailboxes** location supports retention for users' email, calendar, and other mailbox items, by applying retention settings at the level of a mailbox. Shared mailboxes and resource mailboxes for equipment and rooms are also supported.

Email contacts and Microsoft 365 group mailboxes aren't supported for Exchange email. For Microsoft 365 group mailboxes, select the **Microsoft 365 Group mailboxes & sites** location instead. Although the Exchange location initially allows a group mailbox to be selected for a static scope, when you try to save the retention policy, you receive an error that "RemoteGroupMailbox" isn't a valid selection for this location.

Depending on your policy configuration, inactive mailboxes might be included or not:

- Static policy scopes include inactive mailboxes when you use the default **All mailboxes** configuration but aren't supported for specific inclusions or exclusions. However, if you include or exclude a recipient that has an active mailbox at the time the policy is applied and the mailbox later goes inactive, the retention settings continue to be applied or excluded.
- Adaptive policy scopes, by default, include inactive mailboxes when they meet the scope's query. You can exclude them by using the advanced query builder and the OPATH property *IsInactiveMailbox*:

  Console

  ```
  (IsInactiveMailbox -eq "False")
  ```

If you use a static policy scope and choose recipients to include or exclude, you can select distribution groups and email-enabled security groups as an efficient way to select multiple recipients instead of selecting them one-by-one. When you use this option, behind the scenes, these groups are automatically expanded at the time of configuration to select the mailboxes of the users in the group. If the membership of those groups later change, your existing retention policy isn't automatically updated, unlike adaptive policy scopes.

For detailed information about which mailbox items are included and excluded when you configure retention settings for Exchange, see What's included for retention and deletion.

The **Exchange public folders** location applies retention settings to all public folders and can't be applied at the folder or mailbox level.

Configuration information for SharePoint sites and OneDrive accounts

The location name for the SharePoint sites location changes, depending on whether you're using adaptive policy scopes or static policy scopes. The name change reflects what type of sites are included:

- Static policy scopes: **SharePoint classic and communication sites**
  - Can retain and delete documents in SharePoint communication sites, team sites that aren't connected by Microsoft 365 groups, and classic sites. This includes sites associated with Teams shared channels and Teams private channels. Team sites connected by Microsoft 365 groups aren't supported with this option and instead, use the **Microsoft 365 Group mailboxes & sites** location that applies to content in the group's mailbox, site, and files.

- Adaptive policy scopes: **SharePoint sites**
  - Can retain and delete documents in OneDrive sites and Microsoft 365 group-connected sites in addition to SharePoint communication sites, team sites that aren't connected by Microsoft 365 groups, and classic sites.

> **Tip**
>
> You can use a **filter in the SharePoint admin center** or a **SharePoint PowerShell command** to confirm whether a site is group-connected.

Microsoft Loop files can be stored in OneDrive, SharePoint, and SharePoint Embedded. For more information, see Overview of Loop storage. When you use the default policy setting of all sites for SharePoint, this configuration automatically includes all Loop workspaces but you can specify individual workspaces by their container URL.

For detailed information about what's included and excluded when you configure retention settings for SharePoint and OneDrive, see What's included for retention and deletion.

When you specify your locations for SharePoint sites or OneDrive accounts, you don't need permissions to access the sites. For static scopes, no validation is done at the time you specify the URL on the **Edit locations** page. However, the SharePoint sites that you specify are checked that they exist on the final page of the configuration. If this check fails, you see a message that validation failed for the URL you entered, and the retention policy can't be created until the validation check passes. If you see this message, go back in the configuration process to change the URL or remove the site from the retention policy.

To specify individual OneDrive accounts, see Get a list of all user OneDrive URLs in your organization.

> **Note**
>
> When you specify individual OneDrive accounts, be aware that unless OneDrive accounts are pre-provisioned, the URL isn't created until a user accesses their OneDrive for the first time.
>
> Also, the OneDrive URL will automatically change if there is a change in the user's UPN. For example, a name-changing event such as marriage, or a domain name change to support an organization's rename or business restructuring. If the UPN changes, you will need to update the OneDrive URLs you specify for retention settings.
>
> Because of the challenges of reliably specifying URLs for individual users to include or exclude for static scopes, adaptive scopes with the **User** scope type are better suited for this purpose.

Configuration information for Microsoft 365 Group mailboxes & sites

To retain or delete content for a Microsoft 365 group (formerly Office 365 group), use the **Microsoft 365 Group mailboxes & sites** location. For retention policies, this location includes the

group mailbox and SharePoint teams site. For retention labels, this location includes the SharePoint teams site only.

For detailed information about which items are included and excluded for Microsoft 365 Groups:

- For group mailboxes, see What's included for retention and deletion for Exchange retention.
- For SharePoint teams sites, see What's included for retention and deletion for SharePoint retention.

Mailboxes that you target with this policy location require at least 10 MB of data before retention settings will apply to them.

> **Note**
>
> Even though a Microsoft 365 group has an Exchange mailbox, a retention policy for the **Exchange mailboxes** location won't include content in Microsoft 365 group mailboxes.

If you use static scopes: Although the **Exchange mailboxes** location for a static scope initially allows you to specify a group mailbox to be included or excluded, when you try to save the retention policy, you'll see an error that "RemoteGroupMailbox" isn't a valid selection for the Exchange location.

By default, a retention policy applied to a Microsoft 365 group includes the group mailbox and SharePoint teams site. Files stored in the SharePoint teams site are covered with this location, but not Teams chats or Teams channel messages that have their own retention policy locations.

To change the default because you want the retention policy to apply to either just the Microsoft 365 mailboxes, or just the connected SharePoint teams sites, use the Set-RetentionCompliancePolicy PowerShell cmdlet and the *Applications* parameter with one of the following values:

- `Group:Exchange` for just Microsoft 365 mailboxes that are connected to the group.
- `Group:SharePoint` for just SharePoint sites that are connected to the group.

To return to the default value of both the mailbox and SharePoint site for the selected Microsoft 365 groups, specify `Group:Exchange,SharePoint`.

What happens if a Microsoft 365 group is deleted after a policy is applied

When a policy for retention (static policy scope or adaptive) is applied to a Microsoft 365 group, and that group is then deleted from Microsoft Entra ID:

- The group-connected SharePoint site is preserved and continues to be managed by the retention policy with the **Microsoft 365 Group mailboxes & sites** location. Although users who were granted access to the group-connected SharePoint site via the SharePoint admin center continue to be able to access the site, users who were granted access to this site via the deleted group must now be granted permissions from the SharePoint admin center. Now the group is deleted, any new permissions must be granted from the SharePoint admin center.

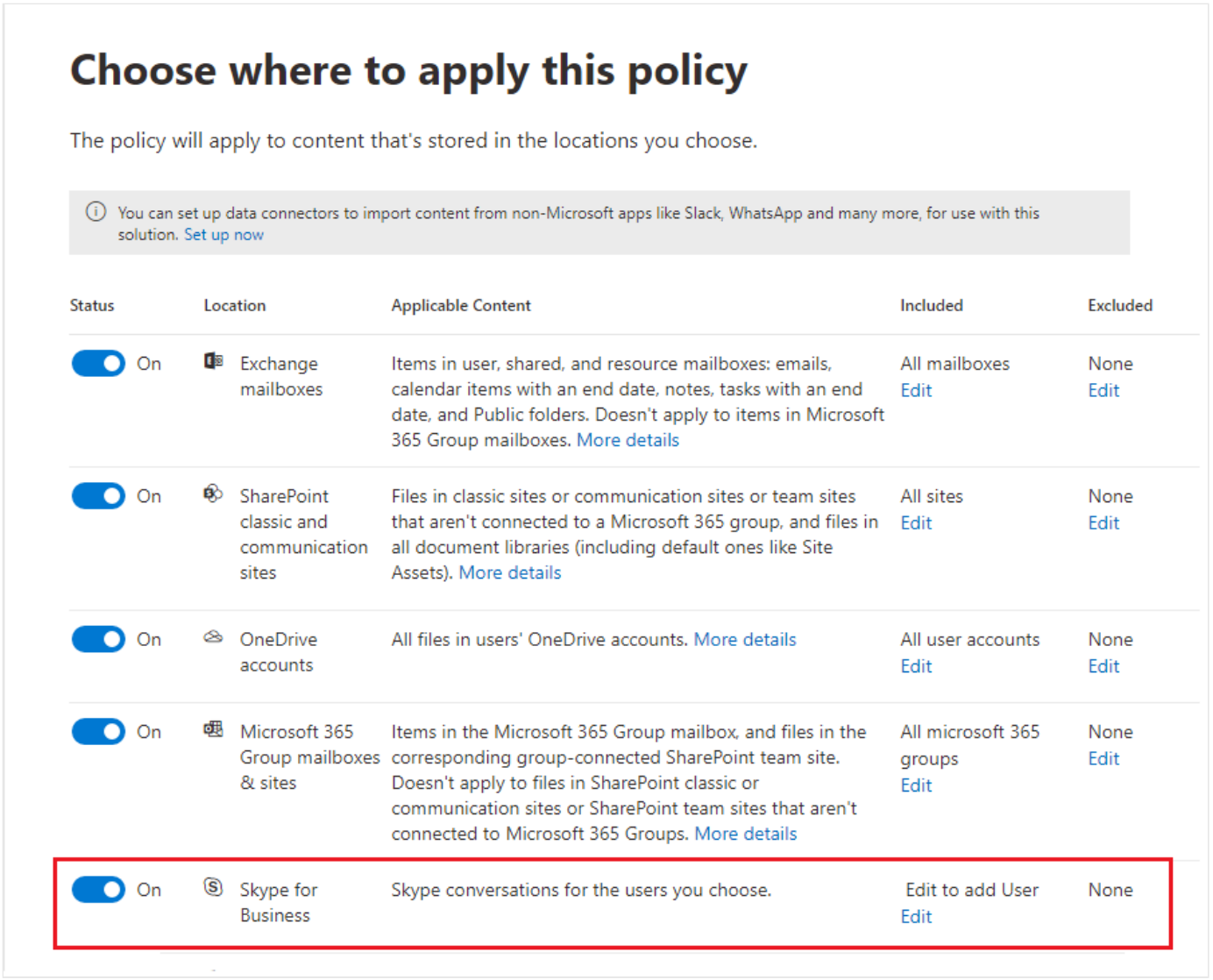
- To release the retention policy from this site, use the *DeletedResources* parameter with the Set-RetentionCompliancePolicy PowerShell cmdlet.

- The mailbox for the deleted group becomes inactive and like the SharePoint site, remains subject to retention settings. For more information, see Inactive mailboxes in Exchange Online.

Configuration information for Skype for Business

> **Note**
>
> Skype for Business was retired July 31, 2021 and we encourage customers to migrate to Microsoft Teams. However, retention policies for Skype for Business continues to be supported for existing customers.

Unlike Exchange email, you can't toggle the status of the Skype location on to automatically include all users, but when you turn on that location, you must then manually choose the users whose conversations you want to retain:



After you select this **Edit** option, in the **Skype for Business** pane you can quickly include all users by selecting the hidden box before the **Name** column. However, it's important to understand that each user counts as a specific inclusion in the policy. So if you include 1,000 users by selecting this box, it's the same as if you manually selected 1,000 users to include, which is the maximum supported for Skype for Business.

Be aware that **Conversation History**, a folder in Outlook, is a feature that has nothing to do with Skype archiving. **Conversation History** can be turned off by the end user, but archiving for Skype is done by storing a copy of Skype conversations in a hidden folder that is inaccessible to the user but available to eDiscovery.

Settings for retaining and deleting content

By choosing the settings for retaining and deleting content, your policy for retention will have one of the following configurations for a specified period of time:

- Retain-only

For this configuration, choose the following options:

- For retention policies: On the **Decide if you want to retain content, delete it, or both** page, select **Retain items for a specific period**, specify the retention period and then for **At end of the retention period** select **Do nothing** for the retention settings to be removed. Or to retain without an end date, select **Retain items forever** on this page.
- For retention labels: On the **Define label settings page**, select **Retain items indefinitely or for a specific period**, and then:
  - For the retention settings to no longer be in effect on the labeled content after a specific time: On the **Define the retention period** page, for **Retain items for**, specify the time period. Then on the **Choose what happens after the retention period** page, select **Deactivate retention settings**. The label remains on the content but with no retention restrictions, as if it's a label that just classifies. However, if the label marks items as a record, the item still supports lock and unlock that has its own set of restrictions.
  - To retain without an end date: On the **Define the retention period** page, for **Retain items for**, select **An indefinite period**. The label remains on the content with any existing restrictions.

- Retain and then delete

  For this configuration, choose the following options:

  - For retention policies: On the **Decide if you want to retain content, delete it, or both** page, select **Retain items for a specific period**, specify the retention period and then for **At end of the retention period** select **Delete items automatically**.
  - For retention labels: On the **Define label settings** page, select **Retain items indefinitely or for a specific period**, specify the retention period and then for **Choose what happens after the retention period**, select either **Delete items automatically** or **Start a disposition review**. For information about disposition reviews, see Disposition review.

- Delete-only

  For this configuration, choose the following options:

  - For retention policies: On the **Decide if you want to retain content, delete it, or both** page, select **Only delete items when they reach a certain age**, and specify the time period.

- For retention labels: On the **Define label settings** page, select **Enforce actions after a specific period** and specify the time period, still referred to as the retention period. The option **Choose what happens after the period** is automatically set to **Delete items automatically**.

Archiving retained SharePoint files

For SharePoint, you can add an archive action to retention policies and retention labels. This action moves supported inactive files to Microsoft 365 Archive while the files remain subject to retention and legal holds. Items that don't support the archive action remain active. Archived files remain available for Microsoft Purview eDiscovery and content search but are excluded from Microsoft 365 Copilot indexing.

The archive action isn't enabled by default. Existing retention policies and labels don't archive files unless an administrator updates them to include the archive action. Event-based retention isn't supported for file-level archive.

To configure the archive action, select **Archive retained items in secure and compliant Microsoft 365 Archive to reduce cost** in the retention settings. Then specify when the archive period starts and how long files remain active before they're archived. The archive period can start when items were created or last modified.

File-level archive requires a Microsoft 365 E5 license. No additional setup is required to archive files by using retention policies or labels. To use the cost-saving benefits of Microsoft 365 Archive, first set up pay-as-you-go billing for Microsoft 365 Archive.

Users with read access can reactivate an archived file when they need it. Retention and hold requirements continue to apply after reactivation or deletion. After a file is reactivated, the policy archives it again after a 120-day cooldown period.

For more information about managing archived files, see Manage Microsoft 365 Archive.

Retaining content for a specific period of time

When you configure a retention label or policy to retain content in the Microsoft Purview portal, you choose to retain items for a specific number of days, months (assumes 30 days), or years (assumes 365 days). Or alternatively, retain the items forever. The retention period isn't calculated from the time the policy was assigned, but according to the start of the retention period specified.

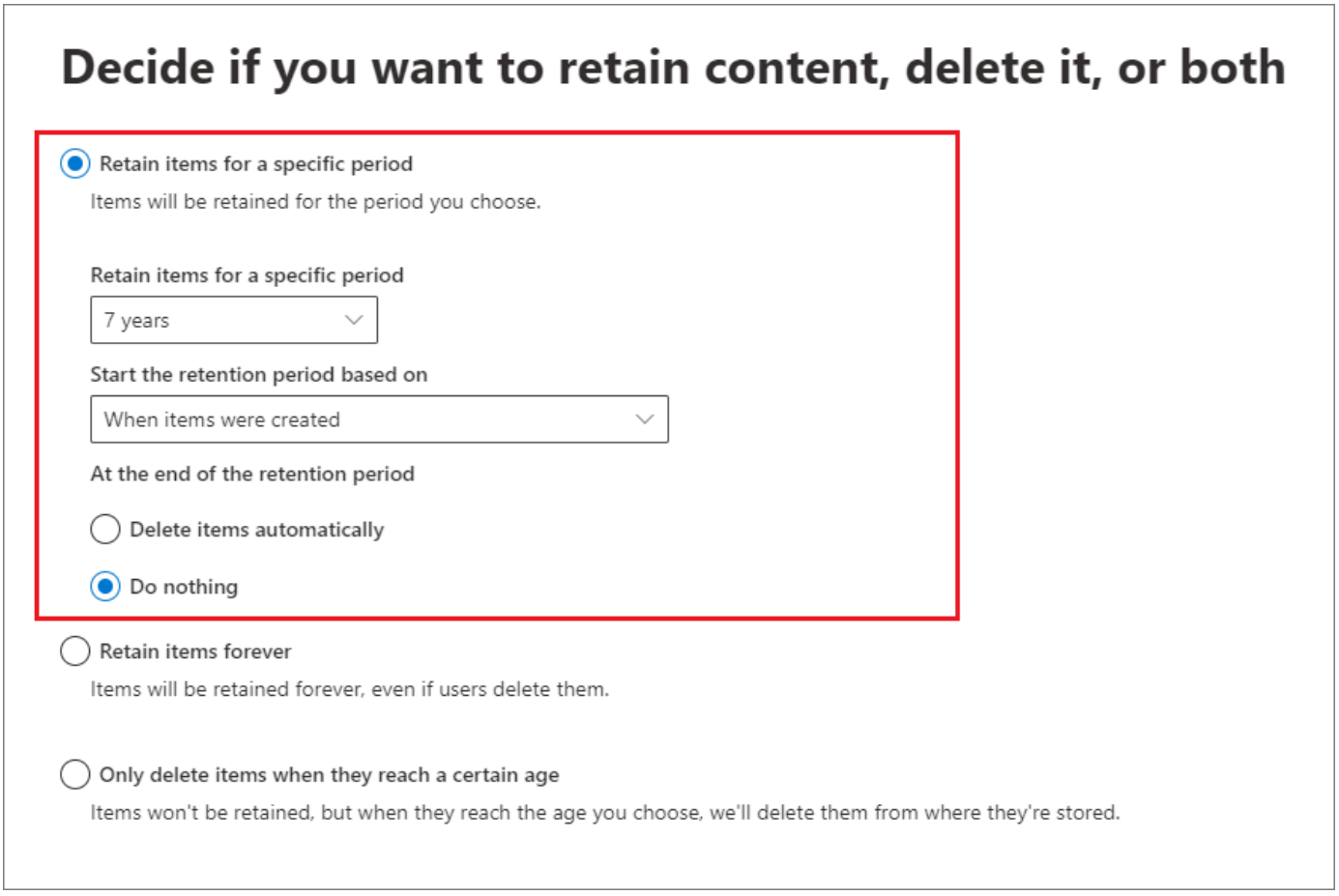
For the start of the retention period, you can choose when the content was created or, supported only for files and the SharePoint, OneDrive, and Microsoft 365 Groups, when the content was last modified. For retention labels, you can start the retention period from the content was labeled, and when an event occurs.

Examples:

- SharePoint: If you want to retain items in a site collection for seven years after this content is last modified, and a document in that site collection hasn't been modified in six years, the document will be retained for only another year if it's not modified. If the document is edited again, the age of the document is calculated from the new last modified date, and it will be retained for another seven years.

- Exchange: If you want to retain items in a mailbox for seven years, and a message was sent six years ago, the message will be retained for only one year. For Exchange items, the age is based on the date received for incoming email, or the date sent for outgoing email. Retaining items based on when it was last modified applies only to site content in OneDrive and SharePoint.

At the end of the retention period, you choose whether you want the content to be permanently deleted. For example, for retention policies:



Retention labels have two more options. As described in the next section, they can apply another retention label with its own retention period. Or, they can trigger a Power Automate flow for custom actions.

Before you configure retention, first familiarize yourself with capacity and storage limits for the respective workloads:

- For SharePoint and OneDrive, retained items are stored in the site's Preservation Hold library, which is included in the site's storage quota. For more information, see Manage site storage limits from the SharePoint documentation.
- For Exchange, Teams, AI apps, and Viva Engage, where retained messages are stored in mailboxes, see Exchange Online limits and enable auto-expanding archiving.

  In extreme cases where a high volume of email is deleted in a short time period, either by users or automatically from policy settings, you might also need to configure Exchange to more frequently move items from the Recoverable Items folder in the user's primary mailbox to the Recoverable Items folder in their archive mailbox. For step-by-step instructions, see Increase the Recoverable Items quota for mailboxes on hold.

Relabeling at the end of the retention period

When you configure a retention label to automatically apply a different retention label at the end of the retention period, the item is then subject to the retention settings of the newly selected retention label. This option lets you automatically change the retention settings for the item.

You can change the replacement label after you've created and saved the primary retention label. For items that already have the primary retention label applied and within the configured retention period, the change of replacement label will synchronize to these items. As with other label changes, allow up to 7 days for this synchronization period.

For the replacement label, you'll typically choose a label that has a longer retention period than the primary retention label. However, that isn't necessarily the case because of the label setting when to start the retention period. For example, the primary retention label is configured to start the retention period when the item is created, and the replacement label starts the retention period when labeled, or when an event occurs.

If there's also a change in whether the label marks the item as a record or a regularly record, the replacement retention label can also change the restrictions for what action are allowed or blocked for that item.

Relabeling example configuration

You create and configure a retention label for an industry-compliance requirement to retain content for three years after it's created, and mark the item as a record. When this label is applied, users won't be able to delete the item from their app, because that's one of the restrictions of a record.

At the end of the three years, you want to automatically retain the content for two more years because of internal compliance policies, but there's no need to mark it as a record with the restrictions that this configuration applies.

To complete the configuration, you select the label setting to change the label at the end of the retention period, and choose a label that retains content for five years after the content was created, and doesn't mark the item as a record.

With these concatenated settings, users will be able to delete the item from their app after three years but it remains accessible for eDiscovery searches for five years.

Considerations for the relabeling option

- You can't relabel a regulatory record but the replacement label can be configured to mark the content as a regulatory record.
- You won't be able to delete a retention label that's selected as a replacement label.
- You can choose a replacement label that's configured to apply another replacement label. There's no limit to the number of replacement labels an item can have.
- If the replacement label marks the item as a record or regulatory record but can't be applied because the file is currently checked out, the relabel process is retried when the file is checked back in again, or checkout is discarded.
- A replacement label is visible to users in Outlook only when that label is included in a published label policy for the same location, or it's configured for delete-only.

Configuration paths for relabeling

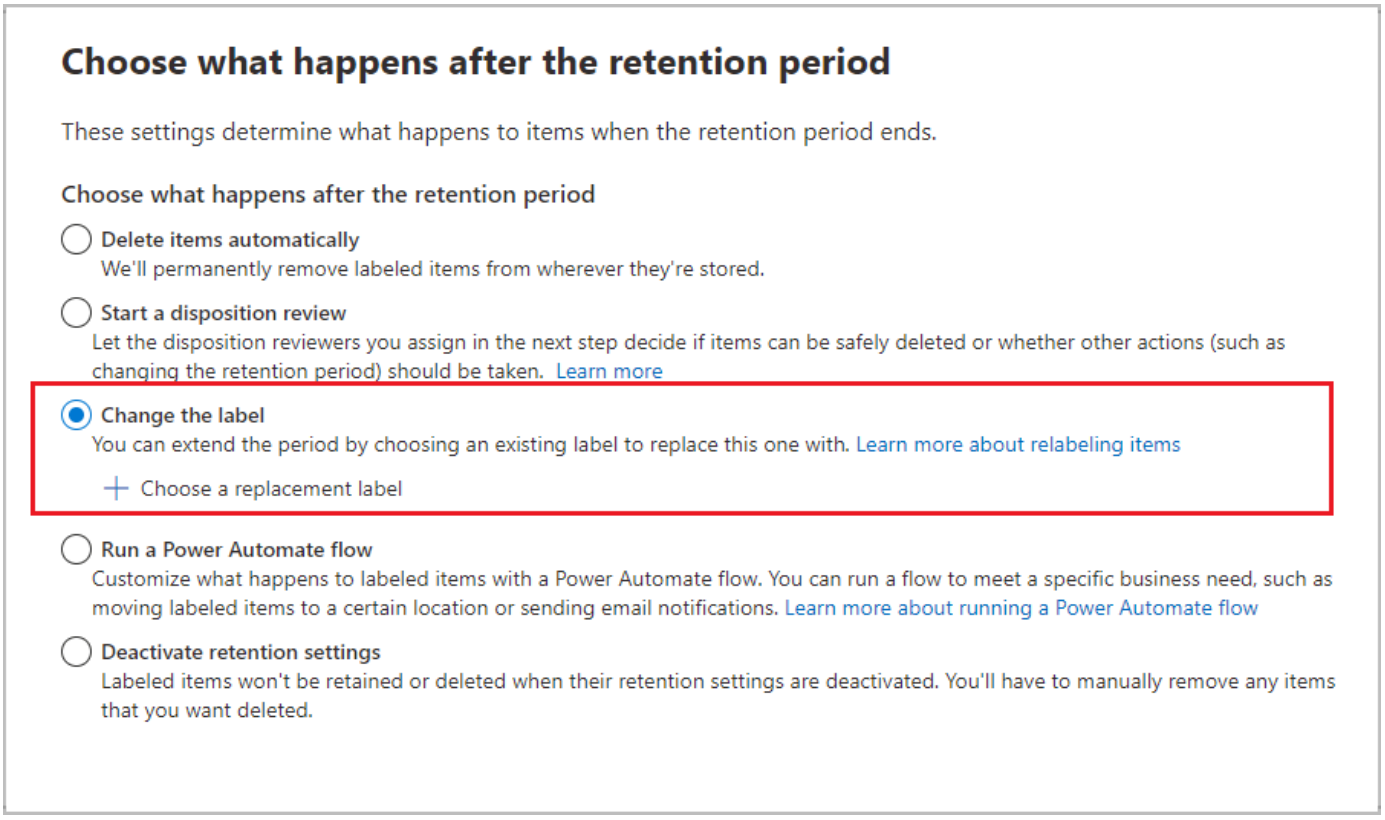
The option to relabel at the end of the retention period has two configuration paths when you create a retention label:

- If you need to initially retain content with the primary label (most typical): On the **Define label settings** page, select **Retain items indefinitely or for a specific period** and specify the

retention period. Then on the **Choose what happens after the retention period** page, select **Change the label** > **Choose a replacement label**.

- If you don't need to initially retain content with the primary label: On the **Define label settings** page, select **Enforce actions after a specific period**, specify the retention period, and then select **Change the label** > **Choose a replacement label**.

In both cases, the replacement label must already be created but doesn't need to be included in an existing label policy.



Alternatively, disposition reviewers can manually select a replacement label as part of the disposition review process if the label setting **Start a disposition review** is selected on this **Choose what happens after the retention period** page.

Deleting content that's older than a specific age

Retention settings can retain and then delete items, or delete old items without retaining them.

In both cases, if your retention settings delete items, it's important to understand that the time period you specify isn't calculated from the time the policy was assigned, but according to the start of the retention period specified. For example, from the time when the item was created or modified, or labeled.

For this reason, first consider the age of the existing content and how the settings might impact that content. Consider communicating your chosen settings to your users and help desk before the settings are applied to content, which gives them time to assess the possible impact.

A policy that applies to entire locations

When you choose locations, except for Skype for Business, the default setting is **All** when the status of the location is **On**.

When a retention policy applies to any combination of entire locations, there is no limit to the number of recipients, sites, accounts, groups, etc., that the policy can include.

For example, if a policy includes all Exchange email and all SharePoint sites, all sites and recipients will be included, no matter how many. And for Exchange, any new mailbox created after the policy is applied will automatically inherit the policy.

A policy with specific inclusions or exclusions

Be aware that if you use the optional configuration to scope your retention settings to specific users, specific Microsoft 365 groups, or specific sites, there are some limits per policy to be aware of. For more information, see Limits for retention policies and retention label policies.

To use the optional configuration to scope your retention settings, make sure the **Status** of that location is **On**, and then use the links to include or exclude specific users, Microsoft 365 groups, or sites.

> ⚠ **Warning**
>
> If you configure includes and then remove the last one, the configuration reverts to **All** for the location. Make sure this is the configuration that you intend before you save the policy.
>
> For example, if you specify one SharePoint site to include in your retention policy that's configured to delete data, and then remove the single site, by default all SharePoint sites will then be subject to the retention policy that permanently deletes data. The same applies to includes for Exchange recipients, OneDrive accounts, Teams chat users etc.
>
> In this scenario, toggle the location off if you don't want the **All** setting for the location to be subject to the retention policy. Alternatively, specify excludes to be exempt from the policy.

Updating policies for retention

Some settings can't be changed after a policy for retention is created and saved, which include:

- The policy name and the retention settings except the retention period and when to start the retention period.

If you edit a retention policy and items are already subject to the original settings in your retention policy, your updated settings will be automatically applied to these items in addition to items that are newly identified.

Usually this update is fairly quick but can take several days. When the policy replication across your Microsoft 365 locations is complete, you'll see the status of the retention policy in the Microsoft Purview portal change from **On (Pending)** to **On (Success)**.

Locking the policy to prevent changes

If you need to ensure that no one can turn off the policy, delete the policy, or make it less restrictive, see Use Preservation Lock to restrict changes to retention policies and retention label policies.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 09/10/2026

PowerShell cmdlets for retention policies and retention labels

> *Microsoft Purview service description*

Use the following sections to identify the main PowerShell cmdlets that are available for retention policies and retention labels that you might need for configuration at-scale, automated scripts, or advanced configuration scenarios. For the full list of cmdlets, see the policy-and-compliance-retention list from the PowerShell documentation.

Before you use these cmdlets, you must first connect to Security & Compliance PowerShell.

In the descriptions that follow, a policy for retention can refer to a retention policy (no labels), or a retention label policy. Each policy defines whether it's static or adaptive and the locations for the policy to be applied.

When you use retention policies, the policy requires one rule to complete the configuration, for example, a rule that defines the retention settings, such as retain for five years and then delete.

When you use retention labels, these labels contain the retention settings and their policies need different rules:

- A retention label policy that you publish needs a rule that defines which labels should be displayed in apps.
- An auto-apply retention label policy needs a rule that defines the label to apply and the conditions for applying the label.

Retention cmdlets for older locations

Use the cmdlets in the following table when the locations are **Exchange mailboxes**, **SharePoint sites** or **SharePoint classic and communication sites**, **OneDrive accounts**, **Microsoft 365 Group mailboxes & sites**, **Skype for Business**, **Exchange public folders**, **Teams chats and Copilot interactions**, or **Teams channel messages**.

Don't use these cmdlets when the locations are for Teams private channel messages (applicable premigration only), Viva Engage user messages, Viva Engage community messages, Teams chats

(without Copilot interactions), Microsoft Copilot experiences, Enterprise AI apps, or Other AI apps. These locations have alternative cmdlets that are identified in the next section.

Expand table

Cmdlet	Description	Applicable locations
Enable-ComplianceTagStorage Get-ComplianceTagStorage	A one-time operation to create storage, or view that storage for retention labels	Exchange mailboxes SharePoint sites (or SharePoint classic and communication sites) OneDrive accounts Microsoft 365 Group mailboxes & sites
Get-ComplianceTag New-ComplianceTag Remove-ComplianceTag Set-ComplianceTag	View, create, delete, configure retention labels	Exchange mailboxes SharePoint sites (or SharePoint classic and communication sites) OneDrive accounts Microsoft 365 Group mailboxes & sites
Get-RecordReviewNotificationTemplateConfig Set-RecordReviewNotificationTemplateConfig	View or configure the disposition review notification and reminder settings	Exchange mailboxes SharePoint sites (or SharePoint classic and communication sites) OneDrive accounts Microsoft 365 Group mailboxes & sites
Get-RetentionCompliancePolicy New-RetentionCompliancePolicy Remove-RetentionCompliancePolicy Set-RetentionCompliancePolicy	View, create, delete, configure retention policies and retention label policies	Exchange email SharePoint sites (or SharePoint classic and communication sites) OneDrive accounts Microsoft 365 Group

Cmdlet	Description	Applicable locations
		mailboxes & sites Skype for Business Exchange public folders Teams chats and Copilot interactions Teams channel messages
Get-RetentionComplianceRule New-RetentionComplianceRule Set-RetentionComplianceRule Remove-RetentionComplianceRule	View, create, configure, delete settings (rules) for retention policies and retention labels	Exchange email SharePoint sites (or SharePoint classic and communication sites) OneDrive accounts Microsoft 365 Group mailboxes & sites Skype for Business Exchange public folders Teams chats and Copilot interactions Teams channel messages

Retention cmdlets for newer locations

Use the following cmdlets when the locations are for **Teams private channel messages** (applicable premigration only), **Viva Engage user messages**, **Viva Engage community messages**, **Teams chats**, **Microsoft Copilot experiences**, **Enterprise AI apps**, **Other AI apps**, and for Teams call logs.

When the locations are for Teams chats and Copilot interactions, Teams channel messages, Exchange email, SharePoint sites (or SharePoint classic and communication sites), OneDrive accounts, Microsoft 365 Group mailboxes & sites, Skype for Business, or Exchange public folders, use the cmdlets listed in the previous section.

Expand table

Cmdlet	Description	Applicable locations
Get-AppRetentionCompliancePolicy New-AppRetentionCompliancePolicy Remove-AppRetentionCompliancePolicy Set-AppRetentionCompliancePolicy	View, create, delete, configure retention policies	Teams private channel messages Viva Engage user messages Viva Engage community messages Teams chats Teams call logs Microsoft Copilot experiences Enterprise AI apps Other AI apps
Get-AppRetentionComplianceRule New-AppRetentionComplianceRule Remove-AppRetentionComplianceRule Set-AppRetentionComplianceRule	View, create, configure, delete settings (rules) for retention policies	Teams private channel messages Viva Engage user messages Viva Engage community messages Teams chats Teams call logs Microsoft Copilot experiences Enterprise AI apps Other AI apps

Configuration guidance

Use the help pages associated with each cmdlet for detailed information and examples.

For guided help to create and then publish retention labels, see Create and publish retention labels by using PowerShell.

Last updated on 06/03/2026

Limits for retention policies and retention label policies

> *Microsoft Purview service description*

When you use retention policies and retention label policies to automatically retain or delete data for your organization, there are some maximum numbers to be aware of.

Maximum number of retention labels per tenant

A maximum of 1,000 retention labels are supported per tenant.

Maximum number of policies per tenant

A single tenant can have a maximum of 10,000 policies (any configuration). This maximum number includes the different policies for retention, and other policies for compliance such as policies for DLP, information barriers, eDiscovery holds, In-Place Holds, and sensitivity labels. However, this maximum excludes:

- Auto-labeling policies for SharePoint and OneDrive, unless they are for cloud attachments.
- Published label policies for SharePoint and OneDrive that delete-only, rather than retain-only, or retain and then delete.
- Exchange retention policies from messaging records management (MRM).
- Litigation holds.

Within this 10,000 policies limit, there are also some limits on the maximum number of policies for retention per workload:

- Exchange (any configuration): 1,800
  - Per mailbox: 25 is the recommended maximum before performance might be impacted; 50 is the supported limit.
- SharePoint or OneDrive: (all sites automatically included): 13
- SharePoint or OneDrive (specific locations included or excluded): 2,600

> ⓘ **Note**

These maximum numbers for Exchange and SharePoint are not exclusive to retention but are shared with other types of hold policies that include eDiscovery holds and In-Place Holds.

Although retention policies for Microsoft Teams and Viva Engage use mailboxes to store data for retention purposes, the maximum number of policies for Exchange Online exclude retention policies for Teams and Viva Engage.

Maximum number of items per policy

> **Important**
>
> Applicable only if you use static policy scopes rather than adaptive policy scopes.

For maximum limits for adaptive policy scopes, see Adaptive policy scopes.

If you use static scopes and the optional configuration to include or exclude specific users, specific Microsoft 365 groups, or specific sites, there are some limits per policy to be aware of.

Maximum numbers of items per policy for retention for static scopes:

- Exchange mailboxes: 1,000
- Microsoft 365 Groups: 500
- Teams channel messages: 500
- Teams chats messages: 1,000
- Copilot and other AI interactions: 1,000
- Viva Engage community messages: 1,000
- Viva Engage user messages: 1,000
- SharePoint sites: 100
- OneDrive accounts: 100

Skype for Business has to be scoped to specific users and the maximum number supported per policy is 1,000.

Because these limitations are per policy, if you need to use specific inclusions or exclusions that result in going over these numbers, you can create additional policies that have the same retention settings. See the next section for some example scenarios and solutions that use multiple retention policies for this reason.

However, multiple policies result in higher administrative overheads. Consider using adaptive scopes rather than creating and maintaining multiple policies with includes and excludes.

Examples of using multiple policies to avoid exceeding maximum numbers

The following examples are for static scopes and provide some design solutions for when you can't specify just the location for a retention policy, and must take into account the maximum number of items documented in the previous section.

Exchange example:

- **Requirement**: In an organization that has over 40,000 user mailboxes, most users must have their email retained for seven years but a subset of identified users (425) must have their email retained for only five years.
- **Solution**: Create one retention policy for Exchange email with a retention period of 7 years and exclude the subset of users. Then create a second retention policy for Exchange email with a retention period of 5 years and include the subset of users.

  The number included and excluded (425 in both cases) is below the maximum number (1,000) of specified mailboxes for a single policy. The subset of users must be explicitly excluded from the first policy because it has a longer retention period than the second policy. If the subset of users required a longer retention policy, you wouldn't need to exclude them from the first policy.

  With this solution, if anybody new joins the organization, their mailbox is automatically included in the first policy for 7 years and there is no impact to the maximum numbers supported. However, new users that require the 5-year retention period add to the include and exclude numbers, and this limit would be reached at 1,000.

SharePoint example:

- **Requirement**: An organization has several thousand SharePoint sites but only 2,000 sites require a retention period of 10 years, and 8,000 sites require a retention period of four years.

- **Solution**: Create 20 retention policies for SharePoint with a retention period of 10 years that includes 100 specific sites, and create 80 retention policies for SharePoint with a retention period of 4 years that includes 100 specific sites.

  Because you don't need to retain all SharePoint sites, you must create retention policies that specify the specific sites. Because a retention policy doesn't support more than 100 specified sites, you must create multiple policies for the two retention periods. These retention policies have the maximum number of included sites, so the next new site that needs retaining would require a new retention policy, irrespective of the retention period.

Maximum numbers for disposition review

For the disposition of content, there are some limits to be aware of:

- Maximum numbers per tenant:
  - 16,000,000 items in either of the following disposition review states: pending disposition or approved disposition
  - 16,000,000 items marked as records automatically disposed (no disposition review)
- Maximum number of disposition reviewers:
  - 10 reviewers per disposition stage in any combination of individual users and mail-enabled security groups
  - 200 reviewers per tenant in any combination of individual users and mail-enabled security groups. Because of this limit, we recommend using mail-enabled security groups whenever possible for larger organizations.
- Proof of disposition for up to 7 years after the item was disposed.

Last updated on 05/19/2026

Learn about data lifecycle management

09/26/2025

> *Microsoft 365 licensing guidance for security & compliance*.

Microsoft Purview Data Lifecycle Management (formerly Microsoft Information Governance) provides you with tools and capabilities to retain the content that you need to keep, and delete the content that you don't.

Retaining and deleting content is often needed for compliance and regulatory requirement, but deleting content that no longer has business value also helps you manage risk and liability. For example, it reduces your attack surface.

Microsoft 365 features

Retention policies are the cornerstone for data lifecycle management. Use these policies for Microsoft 365 workloads that include Exchange, SharePoint, OneDrive, Teams, and Viva Engage. Configure whether content for these services needs to be retained indefinitely, or for a specific period if users edit or delete it. Or you can configure the policy to automatically permanently delete the content after a specified period if it's not already deleted. You can also combine these two actions for retain and then delete, which is a very typical configuration. For example, retain email for three years and then delete it.

When you configure a retention policy, you can target all instances in your organization (such as all mailboxes and all SharePoint sites), or individual instances (such as only the mailboxes for specific departments or regions, or just selected SharePoint sites).

If you need exceptions for individual emails or documents, such as a longer retention period for legal documents, you do this with **retention labels** that you publish to apps so that users can apply them, or automatically apply them by inspecting the content.

Retention labels are also used with **Adaptive Protection**, if you're using this solution with insider risk management. In this case, the retention label and auto-apply policy is automatically created for you. For details, see Dynamically mitigate the risk of accidental or malicious deletes.

Under the covers, retention labels are also used with priority cleanup for scenarios where you need to expedite the permanent deletion of sensitive information from mailboxes, even if they have existing holds for retention or eDiscovery.

For more information about retention policies and retention labels, and how retention works in Microsoft 365, see Learn about retention policies and retention labels.

> **Note**
>
> If you need to manage high-value items for business, legal, or regulatory record-keeping requirements, use retention labels with **records management** rather than retention labels with data lifecycle management.

Other data lifecycle management capabilities to help you keep what you need and delete what you don't:

- **Mailbox archiving** to provide users with additional mailbox storage space, and auto-expanding archiving for mailboxes that need more than 100 GB storage. A default archiving policy automatically moves email to the archive mailbox, and if required, you can customize this policy. For more information about mailbox archiving, see Learn about archive mailboxes.
- **Inactive mailboxes** that retain mailbox content after employees leave the organization. For more information about inactive mailboxes, see Learn about inactive mailboxes.
- **Import service for PST files** by using network upload or drive shipping. For more information, see Learn about importing your organization's PST files.

Exchange (legacy) features

Retention policies and retention tags from messaging records management (MRM), and **journaling rules** are older compliance features from Exchange that were originally configurable from the Classic Exchange admin center. They haven't been brought forward to the new Exchange admin center.

If you're not already using these features, or have a specific business requirement to use them instead of the Microsoft 365 features for data lifecycle management, we don't recommend you use these older compliance features. Instead, use the newer Microsoft 365 features that retain data in place and support policies across other Microsoft 365 services.

For more information, see Use retention policies and retention labels instead of older features.

Deployment guidance

For deployment guidance for data lifecycle management that includes a recommended deployment roadmap, licensing information, permissions, a list of supported scenarios, and end-user documentation, see Get started with data lifecycle management.

Looking for deployment guidance to protect your data? See Deploy an information protection solution with Microsoft Purview.

Get started with data lifecycle management

Summarize this article for me

> *Microsoft Purview service description*

Ready to start managing the lifecycle of your organization's data by retaining the content that you need to keep, and deleting the content that you don't? To get started, use the following guidance for Microsoft Purview Data Lifecycle Management:

1. **Understand how retention and deletion works** in Microsoft 365, and then identify the workloads that need a retention policy and whether you need to create retention labels for exceptions: Learn about retention

   > **Note**
   >
   > If you need to manage high-value items for business, legal, or regulatory record-keeping requirements: Use retention labels with **records management** rather than data lifecycle management.

2. **Create retention policies** for the workloads you identified, specifying retention settings and actions that are required by your organization policies or industry regulations: Create retention policies

   If needed, create and apply retention labels for your exceptions.

3. **Enable mailbox archiving** to provide users with additional mailbox storage space: Enable archive mailboxes in Microsoft 365

   If required to support archive mailboxes:

   - Enable auto-expanding archiving for mailboxes that need more than 100 GB storage.
   - Use retention tags with a retention policy from messaging records management (MRM) if you need to customize how emails are automatically moved from a user's primary mailbox to their archive mailbox, or if you need to specify retention and deletion settings for specific folders rather than the whole mailbox.

4. **Understand and manage inactive mailboxes** that retain mailbox content after employees leave the organization: Learn about inactive mailboxes

5. If you have PST files that contain data you want to govern: **Import PST files to online mailboxes** by using network upload or drive shipping: Learn about importing your organization's PST files

Licensing and billing requirements

A number of different subscriptions support data lifecycle management capabilities.

To see the options for licensing your users to benefit from Microsoft Purview features, see the Microsoft Purview service description. For the features listed on this page, see the Microsoft Purview Data Lifecycle Management & Records Management section for feature-level licensing requirements.

You must set up pay-as-you-go billing to create retention policies that include the locations Microsoft Copilot experiences, Enterprise AI apps, and Other AI apps. There's no charge for Microsoft 365 Copilot interactions.

Permissions

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview **roles and permissions**.

See the following section for information about roles and role groups to manage Microsoft 365 retention.

For permissions to manage mailboxes for archiving, inactive mailboxes, and import, these typically require Exchange permissions, such as the Mail Recipients role. By default, this role is assigned to the Recipient Management and Organization Management role groups. For the exact permissions requirement for each management task, see the documentation that accompanies the admin instructions.

Permissions for retention policies and retention labels

Members of your compliance team who will create and manage retention policies and retention labels need permissions to the Microsoft Purview portal. To grant permissions for this limited administration, we recommend that you add users to the **Compliance Administrator** admin role group.

Alternatively to using this default role, you can create a new role group and add the **Retention Management** role to this group. For a read-only role, use **View-Only Retention Management**.

Specific to priority cleanup that can override retention labels and policies when it's necessary to expedite the permanent deletion of sensitive information, admins must have the **Priority Cleanup Admin** role. This role isn't included by default in the **Compliance Administrator** admin role group.

For instructions to add users to the default roles or create your own role groups, use the following guidance, depending on the portal you're using:

- Permissions in the Microsoft Purview portal

These permissions are required only to create, configure, and apply retention policies and retention labels. The person configuring these policies and labels doesn't require access to the content.

Support for administrative units

Data lifecycle management supports administrative units that have been configured in Microsoft Entra ID:

- You can assign administrative units to members of custom role groups and any others that support administrative units. For example, role groups used with Microsoft Purview Records Management. Edit these role groups and select individual members, and then the **Assign admin units** option to select administrative units from Microsoft Entra ID. These administrators are now restricted to managing just the users in those administrative units.
- You can define the initial scope of retention policies and retention label policies when you create or edit these policies. When you select administrative units, only the users in those administrative units will be eligible for the policy.

> **Important**
>
> Don't select administrative units for a policy that you want to apply to SharePoint sites or to Exchange public folders. Because administrative units support only users and groups, if you configure policy for retention to use administrative units, you won't be able to select the locations for SharePoint sites or Exchange public folders.

- Both adaptive scopes and static scopes support administrative units.
- Additional impact for restricted administrators

    - Policy lookup: Restricted administrators will see policies only from users within their assigned administrative units
    - Import PST files: Restricted administrators won't be able to use the network upload feature to bulk-import PST files to Microsoft 365 mailboxes
    - Exchange legacy features: Restricted administrators won't be able to configure the Exchange legacy features of retention policies and retention tags from messaging records management (MRM), and journaling rules

- Currently, retention labels don't support administrative units.
- Currently, a restricted administrator can create and view adaptive scopes for all administrative units when they use PowerShell cmdlets.
- Currently, inactive mailboxes aren't supported in a policy when you select one or more administrative units. To include inactive mailboxes in the policy, you must be an unrestricted administrator and select **Full directory**.

For more information about how Microsoft Purview supports administrative units, see Administrative units.

Common scenarios

Use the following table to help you map your business requirements to the most common scenarios for data lifecycle management.

Expand table

I want to ...	Documentation
Efficiently retain or delete data for Microsoft 365 services and Copilot interactions: - Exchange - SharePoint - OneDrive - Microsoft 365 Groups - Teams - Copilot and AI apps - Viva Engage - Skype for Business	Create and configure retention policies
Automatically retain content that's deleted by users who have been identified as having an elevated risk level	Dynamically mitigate the risk of accidental or malicious deletes
Provide users with additional mailbox storage	Enable archive mailboxes in Microsoft 365

I want to ...	Documentation
Retain mailbox data after employees leave the organization	Create and manage inactive mailboxes
Upload mailbox data from PST files	Use network upload to import PST files

For scenarios that require data management of individual items, see the common scenarios for records management.

End-user documentation

See the following section for information about end-user documentation to support Microsoft 365 retention.

The data lifecycle management capabilities for inactive mailboxes and import of PST files don't require end-user documentation because these are admin operations only. To help users understand and interact with their archive mailboxes in Outlook after you've enabled this capability, see Manage email storage with online archive mailboxes.

End-user documentation for retention and deletion

Most retention policies work unobtrusively in the background without user interaction, and so need little documentation for users. Retention policies for Teams inform users when their messages have been deleted with a link to Teams messages about retention policies.

However, if you supplement retention policies with retention labels, these labels do have a UI presence in Microsoft 365 apps. Before you deploy these labels to your production network, make sure you provide information and instructions for end users and your help desk. To help users apply retention labels in SharePoint and OneDrive, see Apply retention labels to files in SharePoint or OneDrive.

The most effective end-user documentation will always be customized guidance and instructions you provide for the retention label names and configurations you choose.

Last updated on 02/27/2026

Create and configure retention policies

> *Microsoft Purview service description*

Use a retention policy to manage the data for your organization by deciding proactively whether to retain content, delete content, or retain and then delete the content.

A retention policy lets you do this very efficiently by assigning the same retention settings at the container level to be automatically inherited by content in that container. For example, all items in SharePoint sites, all email messages in users' Exchange mailboxes, all channel messages for teams that are used with Microsoft Teams. If you're not sure whether to use a retention policy at the container level or a retention label at the item level, see Retention policies and retention labels.

For more information about retention policies and how retention works in Microsoft 365, see Learn about retention policies and retention labels.

> **Note**
>
> The information on this page is for compliance administrators. If you are not an administrator and want to understand how retention policies have been configured for the apps that you use, contact your help desk, IT department, or administrator. If you're seeing messages about retention policies in Teams chats and channel messages, you might find it helpful to review **Teams messages about retention policies**.

Before you begin

To make sure you have permissions to create and edit retention policies, see the permissions information for data lifecycle management.

Decide before you create your retention policy whether it will be **adaptive** or **static**. For more information, see Adaptive or static policy scopes for retention. If you decide to use an adaptive policy, you must create one or more adaptive scopes before you create your retention policy, and then select them during the create retention policy process. For instructions, see Configuration information for adaptive scopes.

To retain prompts and responses for AI apps other than Microsoft 365 Copilot and Copilot Studio, you must first have a collection policy for these AI apps, and that policy includes the setting to capture content. For more information, see Collection Policies solution overview.

If you're creating a retention policy for Teams and use private channels, make sure you're aware of the migration of private channel messages in 2025 so you know whether to select **Teams private channel messages** or **Teams channel messages** as the location.

Create and configure a retention policy

Although a retention policy can support multiple services that are identified as "locations" in the retention policy, you can't create a single retention policy that includes all the supported locations:

- **Exchange mailboxes**
- **SharePoint sites** or **SharePoint classic and communication sites**
- **OneDrive accounts**
- **Microsoft 365 Group mailboxes & sites**
- **Skype for Business**
- **Exchange public folders**
- **Teams channel messages**
- **Teams chats**
- **Teams private channel messages** (applicable pre-migration only)
- **Teams call logs**
- **Microsoft Copilot experiences**
- **Enterprise AI apps**
- **Other AI apps**
- **Viva Engage community messages**
- **Viva Engage user messages**

If you select the Teams or Viva Engage locations when you create a retention policy, the other locations are automatically excluded. This means that the instructions to follow depend on whether you need to include the Teams or Viva Engage locations.

> ⓘ **Note**

When you use adaptive policies instead of static policies, you can configure a single retention policy to include both Teams and Viva Engage locations. This isn't the case for static policies where Teams and Viva Engage locations require their own retention policy.

When you've more than one retention policy, and when you also use retention labels, see The principles of retention, or what takes precedence? to understand the outcome when multiple retention settings apply to the same content.

Select the tab for instructions to create a retention policy for Teams, Copilots, AI apps, Viva Engage, or the other supported services (Exchange, SharePoint, OneDrive, Microsoft 365 Groups, Skype for Business):

Retention policy for Teams & AI apps

> **Note**
>
> For AI apps other than Microsoft 365 Copilot and Copilot Studio, see **Before you begin** about the prerequisite of collection policies.

Retention policies for Teams support shared channels. When you configure retention settings for the **Teams channel message** location, if a team has any shared channels, they inherit retention settings from their parent team.

From late April 2026, retention policies also support newly created Teams call logs when you create the retention policies with PowerShell. For more information, see Retention policy for Teams call logs.

1. Sign in to the Microsoft Purview portal > **Solutions** > **Data Lifecycle Management** > **Policies** > **Retention policies**.
2. Select **New retention policy** to start the **Create retention policy** configuration, and name your new retention policy.
3. For the **Assign admin units** page, keep the default of **Full directory**. Currently, admin units aren't supported for this policy.
4. For the **Choose the type of retention policy to create** page, select **Adaptive** or **Static**, depending on the choice you made from the Before you begin instructions. If you haven't already created adaptive scopes, you can select **Adaptive** but because there

won't be any adaptive scopes to select, you won't be able to finish the configuration with this option.

5. Depending on your selected scope:

   - If you chose **Adaptive**: On the **Choose adaptive policy scopes and locations** page, select **Add scopes** and select one or more adaptive scopes that have been created. Then, select one or more locations. The locations that you can select depend on the scope types added. For example, if you only added a scope type of **User**, you'll be able to select **Teams chats** but not **Teams channel messages**.

   - If you chose **Static**: On the **Choose locations to apply the policy** page, select one or more locations:
     - **Teams channel message**: Messages from standard and shared channel chats, and standard and shared channel meetings. Includes messages from private channel chats and private channel meetings post-migration only.
     - **Teams chats**: For Teams, messages from private 1:1 chats, group chats, meeting chats, and chat with yourself.
     - **Teams private channel messages**: Applicable pre-migration only. Messages from private channel chats and private channel meetings. If you select this option, you can't select the other Teams locations in the same retention policy.
     - **Microsoft Copilot experiences**: For built-in and custom Copilot experiences, all user prompts responses. Includes Microsoft 365 Copilot, Security Copilot, Copilot in Fabric, Copilot Studio.
     - **Enterprise AI apps**: For non-Copilot Enterprise AI apps onboarded or connected to your org using methods like Entra registration or data connectors, all user prompts and responses. Includes Entra-registered AI apps, ChatGPT Enterprise, Microsoft Foundry.
     - **Other AI Apps**: For other supported AI apps, all user prompts and responses. includes ChatGPT, Google Gemini, Microsoft Bing, DeepSeek.

     By default, all teams and all users are selected, but you can refine this by selecting the **Choose** and **Exclude** options.

6. For **Decide if you want to retain content, delete it, or both** page, specify the configuration options for retaining and deleting content.

   You can create a retention policy that just retains content without deleting, retains and then deletes after a specified period of time, or just deletes content after a specified period of time. For more information, see Settings for retaining and deleting content.

7. Complete the configuration and save your settings.

For guidance when to use retention policies for Teams and understand the end user experience, see Manage retention policies for Microsoft Teams from the Teams documentation.

For technical details about how retention works for Teams and Copilot data, including what elements of messages are supported for retention and timing information with example walkthroughs, see Learn about retention for Microsoft Teams and Learn about retention for Copilot.

Known configuration issues for Teams retention policies

- Although you can select the option to start the retention period when items were last modified, the value of **When items were created** is always used. For messages that are edited, a copy of the original message is saved with its original timestamp to identify when this pre-edited message was created, and the post-edited message has a newer timestamp.
- When you select **Edit** for the Teams chats location, you might see guests and non-mailbox users. Retention policies aren't designed for these users, so don't select them.
- Before late April 2026: Teams call data records (CDRs) for Teams chat and channel messages were included in Teams chat retention policies. Now, CDRs are included only in a retention policy for Teams call logs, as documented in the next section.

Retention policy for Teams call logs

Teams call logs represent the collection of call-related data generated by Teams, including call data records (CDRs) and other call metadata. CDRs are also sometimes referred to as call detail records, or just call records.

Prior to supporting the retention of Teams call logs in late April 2026, CDRs for Teams chat and Teams channels were included in retention policies for the Teams chat location. Going forward, new CDRs are supported only when you create a retention policy for Teams call logs. CDRs included in previous Teams chat retention policies continue to be managed by those same policies.

This separate retention policy for call logs can be created and modified only by using PowerShell. It has the following considerations:

- The policy includes call logs for both Teams chat and Teams channels.
- The policy applies only to new call logs that are created after the policy is configured and active.

After you create the retention policy for Teams call logs, it's displayed in **Data Lifecycle Management** > **Policies** in the Microsoft Purview portal, where it's visible but read-only.

To create a retention policy for Teams call logs, first use the New-AppRetentionCompliancePolicy cmdlet. As an example, the following command creates a Teams call log retention policy for all mailboxes:

PowerShell

```
New-AppRetentionCompliancePolicy -Name "<PolicyName>" -Applications
"User:MicrosoftTeamsCallLog" -ExchangeLocation "All"
```

Then use the New-AppRetentionComplianceRule cmdlet to create a retention rule that specifies the retention duration and action. For example, the following command creates a delete-only rule that permanently deletes Teams call log data after one year:

PowerShell

```
New-AppRetentionComplianceRule -Name "<RuleName>" -Policy "<PolicyName>" -
RetentionDuration 365 -RetentionDurationDisplayHint Days -
RetentionComplianceAction Delete
```

For more information about using PowerShell cmdlets to create and manage retention policies, see PowerShell cmdlets for retention policies and retention labels.

Additional retention policies needed to support Teams

Teams is more than just messages and call logs from chats and channels. If you have teams that were created from a Microsoft 365 group (formerly Office 365 group), you should additionally configure a retention policy that includes that Microsoft 365 group by using the **Microsoft 365 Group mailboxes & sites** location. This retention policy applies to content in the group's mailbox, site, and files. Files include Teams meeting recordings and transcripts from channel meetings.

To retain or delete Teams meeting recordings with their transcripts from user chats, you'll need a retention policy that includes the organizer's OneDrive account as the location.

If you have team sites that aren't connected to a Microsoft 365 group, which includes sites for Teams shared channels and Teams private channels, you need a retention policy that includes the **SharePoint classic and communication sites** or **OneDrive accounts** locations to retain and delete files in Teams:

- Files that are shared in chat are stored in the OneDrive account of the user who shared the file.
- Files that are uploaded to channels are stored in the SharePoint site for the team.

> **Tip**
>
> You can apply a retention policy to the files of just a specific team when it's not connected to a Microsoft 365 group by selecting the SharePoint site for the team, and the OneDrive accounts of users in the Team.

It's possible that a retention policy that's applied to Microsoft 365 groups, SharePoint sites, or OneDrive accounts could delete a file that's referenced in a Teams chat or channel message before those messages get deleted. In this scenario, the file still displays in the Teams message, but when users select the file, they get a "File not found" error. This behavior isn't specific to retention policies and could also happen if a user manually deletes a file from SharePoint or OneDrive.

Separate an existing 'Teams chats and Copilot interactions' policy

Previously, retention policies used the location **Teams chats and Copilot interactions** that combined Teams chat and Copilot interactions. There are now separate retention locations for Teams chat and Copilot interactions.

- You can separate Teams chats from Copilot interactions from an existing retention policy.
- You can create new retention policies for just Teams chats, or for just Copilot interactions.

Use the following PowerShell commands to separate an existing retention policy for **Teams chats and Copilot interactions**:

- To make your existing retention policy for the older locations a Teams chat only policy:

```
PowerShell
```

```
Set-RetentionCompliancePolicy -Identity "<policy name>" -Applications
"User:TeamsChatUserInteractions"
```

Or, if you have an existing retention policy for newer locations, such as Teams private channel messages and Viva Engage, you can add Teams chat to it:

PowerShell

```
Set-AppRetentionCompliancePolicy -Identity "<policy name>" -Applications
"User:TeamsChatUserInteractions"
```

- To add Microsoft 365 Copilot interactions to an existing retention policy for newer locations, such as Teams private channel messages and Viva Engage:

PowerShell

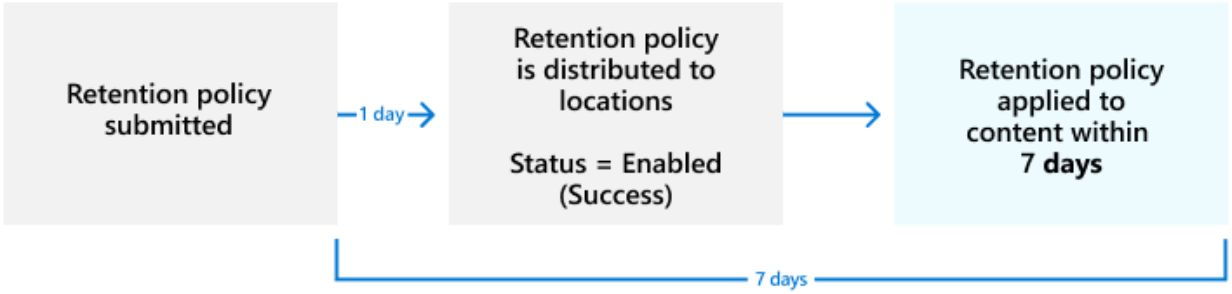
```
Set-AppRetentionCompliancePolicy -Identity “<policy name>” -Applications
"User:M365Copilot"
```

For new retention policies, select the new locations, such as **Teams chat** or **Microsoft Copilot Experiences**.

> ⓘ **Note**
>
> If you have existing retention policies for **Teams chats and Copilot interactions**, they continue to be supported, although they can't be edited when your tenant supports the separate locations. At this point, any new retention policies must use the new locations.

How long it takes for retention policies to take effect

When you create and submit a retention policy, it can take up to seven days for the retention policy to be applied:



First, the retention policy needs to be distributed to the locations that you selected, and then applied to content. You can always check the distribution status of the retention policy by selecting it from the **Retention policies** page in the Microsoft Purview portal. From the flyout pane, if you see **(Error)** included in the status, and in the details for the locations see a message that it's taking longer than expected to deploy the policy or to try redeploying the policy, try running the Set-AppRetentionCompliancePolicy or Set-RetentionCompliancePolicy PowerShell command to retry the policy distribution:

1. Connect to Security & Compliance PowerShell.
2. Run one of the following commands:

   - For the policy locations **Teams private channel messages**, **Viva Engage user messages** and **Viva Engage community messages**:

     PowerShell

     ```
     Set-AppRetentionCompliancePolicy -Identity <policy name> -RetryDistribution
     ```

   - For all other policy locations, such as **Exchange mailboxes**, **SharePoint classic and communication sites**, and **Teams channel messages**:

     PowerShell

     ```
     Set-RetentionCompliancePolicy -Identity <policy name> -RetryDistribution
     ```

Updating retention policies

When settings from the retention policy are already applied to content, a change in configuration to the policy will be automatically applied to this content in addition to content that's newly identified.

Some settings can't be changed after the policy is created and saved, which include the name of the retention policy, the scope type (adaptive or static), and the retention settings except the retention period.

If you no longer need the retention settings that you've configured, see Releasing a policy for retention.

Troubleshooting retention policies

If your retention policies aren't working as expected or you see errors related to your retention policies, use the following troubleshooting resources:

- Identify errors in Microsoft 365 retention and retention label policies
- Resolve errors in Microsoft 365 retention and retention label policies

Next steps

If some items for Exchange, SharePoint, OneDrive, or Microsoft 365 Groups need different retention settings from the retention policy settings you've configured, create retention labels for these exceptions.

However, if you're looking to manage high-value items for business, legal, or regulatory record-keeping requirements, use file plan to create and manage retention labels.

Last updated on 06/09/2026

Create retention labels for exceptions to your retention policies

> *Microsoft Purview service description*

As part of your data governance strategy to retain what you need and delete what you don't, you might need to create a few retention labels for items that need exceptions to your retention policies.

Whereas retention policies automatically apply to all items at the container level (such as SharePoint sites, user mailboxes, and so on), retention labels apply to individual items, such as a SharePoint document or an email message.

Make sure you understand the principles of retention before you use retention labels to supplement a retention policy for specific SharePoint, OneDrive, or Exchange items. Typically, you'll use retention labels to retain specific items longer than an applied retention policy, but they can also be used to override automatic deletion at the end of the retention period, or apply a different deletion period.

As a typical example: The majority of content on your SharePoint sites need to be retained for three years, which is covered with a retention policy. But you have some contract documents that must be retained for seven years. These exceptions can be addressed with retention labels. After assigning the retention policy to all SharePoint sites, you apply the retention labels to the contract documents. All SharePoint items will be retained for three years, and just the contract documents will be retained for seven years.

For more examples of how retention labels can be used as exceptions to retention policies, see Combining retention policies and retention labels.

Retention labels also support more capabilities than retention policies. For more information, see Compare capabilities for retention policies and retention labels.

Use the following information to help you create retention labels to supplement retention policies as part of your data lifecycle management strategy.

> **Note**

Create retention labels from the **Records management** solution rather than **Data lifecycle management** if you need to use retention labels to manage high-value items for business, legal, or regulatory record-keeping requirements. For example, you want to use event-based retention or disposition review. For instructions, see Use file plan to create and manage retention labels.

Before you begin

To make sure you have permissions to create and edit retention labels and their policies, see Permissions for retention policies and retention labels.

How to create retention labels for data lifecycle management

1. Sign in to the Microsoft Purview portal > **Solutions** > **Data Lifecycle Management** > **Retention labels**.

2. Select **Create a label** and follow the prompts to create the retention label. Be careful what name you choose, because this can't be changed after the label is saved.

   > **Note**
   >
   > Double quotation marks ("") aren't supported in retention label names. For example, ""Permanent Label"" isn't supported.

   For more information about the retention settings, see Settings for retaining and deleting content.

3. After you have created the label and you see the options to publish the label, auto-apply the label, or just save the label: Select **Just save the label for now**, and then select **Done**.

4. Repeat these steps to create any more retention labels that you need for different retention settings.

To edit an existing label, select it, and then select the **Edit label** option to start the Edit retention label configuration that lets you change the label descriptions and any eligible settings.

Most settings can't be changed after the label is created and saved, which include the label name, and the retention settings except the retention period.

Next steps

Now you've created retention labels, they are ready to be added to items by publishing the labels, or automatically applying them:

- Publish retention labels and apply them in apps
- Apply a retention label to content automatically

Last updated on 06/22/2026

Expedite the permanent deletion of sensitive information from mailboxes

> *Microsoft 365 licensing guidance for security & compliance.*

> **Note**
>
> Priority cleanup is rolling out in preview and subject to change.
>
> Although this feature requires additional permissions and multiple approvals, highly regulated organizations that use Preservation Lock might want the additional safeguard of turning off priority cleanup at the tenant level to ensure that their policies remain in compliance with required regulations.

Use the **Priority cleanup** feature under **Data Lifecycle Management** in Microsoft Purview when you need to permanently delete sensitive content from Exchange mailboxes, despite retention settings that retain the item or eDiscovery holds. Items still go through the same deletion process with the same timings to ensure a compliant permanent deletion, but there's no need to wait for the retention period to expire or the hold to be released. Priority cleanup might be implemented for security or privacy in response to an incident, or for compliance with regulatory requirements. For example:

- **Privacy request**: You receive a request to remove personal information about an employee who has left the organization and all mailboxes have a retention policy to retain items for two years.
- **Data spillage**: An employee accidentally emailed sensitive information about a future acquisition and you need to delete this information, despite any eDiscovery holds in progress in the organization.

Because the deletion is irreversible and can override existing holds and even Preservation Lock, the process requires multiple approvals, specific roles, and is audited. After considering these safeguards, if your organization still has concerns about this capability, you can just continue to use retention policies and retention labels to ensure a compliant deletion of content instead of using priority cleanup.

Under the covers, priority cleanup uses retention labels with auto-apply policies. However, you don't interact manually with these labels and policies, and they supersede the principles of retention to achieve the required expedited deletion.

> **Note**
>
> If an item is subject to multiple priority cleanups, the newest takes priority.

Important exceptions for priority cleanup:

- You can't use priority cleanup for items that are marked as a record or regulatory record.
- Although priority cleanup can override eDiscovery holds, the data might already be copied to an eDiscovery review set where it can't be deleted by priority cleanup. This data is automatically deleted when the entire eDiscovery case is deleted by an eDiscovery admin.

Similar to auto-apply retention labels, priority cleanup supports simulation, so you can check the returned samples in case the policy configuration needs any fine-tuning.

Comparing priority cleanup for different workloads

Although you can use the priority cleanup feature for mailboxes and for files in SharePoint and OneDrive, the typical use cases are different. There are some other differences in behavior and configuration that are summarized in the following table. As a result of these differences, you can't create a cleanup policy for all three locations in same policy.

Expand table

Behavior or configuration	Exchange	SharePoint/OneDrive
Typical use case:	Data spillage Delete sensitive data for compliance requirements	Delete stale Teams meeting recordings and transcripts Delete retained files in the Preservation Hold library
Typical policy use:	Rare	For Teams meeting recordings and transcripts: Continual For the Preservation Hold library: After a user leaves the organization
Two-person rule:	Another priority cleanup admin is assigned in the policy to review items after the policy is turned on	Another priority cleanup admin is included in the policy configuration flow before the policy is turned on

Behavior or configuration	Exchange	SharePoint/OneDrive
Assign policy approvers:	- Priority cleanup admin - eDiscovery admin - Retention manager	- eDiscovery admin
Simulation required before running the policy:	No (but recommended)	Yes
Override Preservation Lock:	Yes	Only if the policy is configured for delete-only

Additionally, the required permissions for Exchange are different from the required permissions for SharePoint and OneDrive.

Prerequisites for priority cleanup

Make sure you can meet the prerequisites that must be in place before you can use priority cleanup to expedite the permanent deletion of sensitive data. These requirements include permissions and approvers.

Because of the built-in safeguards, the feature itself is enabled by default at the tenant level. However, priority cleanup can be turned off on the priority cleanup settings page. If you can't create new priority cleanup policies, see the instructions for turning off the feature to check the status and reverse the configuration.

> **Note**
>
> A mailbox must have at least 10-MB data to support priority cleanup.

Permissions for priority cleanup

To successfully access and manage **Priority cleanup** in the Microsoft Purview portal, users must have the **Priority Cleanup Admin** role. This role is required to create and manage priority cleanup policies, enable or disable the feature, or approve items within the initial approval stage. This role is automatically added to the Organization Management role group but must be manually added to any other role group.

Alternatively, the **Priority Cleanup Viewer** role allows only the visibility of priority cleanup policies and settings without the ability to make changes or create new policies.

Content Explorer List Viewer and **Content Explorer Content Viewer** roles are required to view item content and details in simulation mode and approval stages.

Similar to records management disposition, each person that accesses the **Priority cleanup** > **Pending cleanups** page sees only items that they're assigned to approve. To monitor the end-to-end process of a priority cleanup, use auditing and the priority cleanup ID as a search term.

Policy approvers

As a safeguard against accidental or malicious deletions, each item subject to priority cleanup always requires at least one other person to approve the permanent deletion in addition to the person who created the priority cleanup policy. Approvers must be individual users. Mail-enabled security groups are not currently supported.

- Priority cleanup admin approval is required for all policies.
- If an item is also subject to retention settings from a retention label or retention policies, retention manager approval is also required.
- If an item is also subject to one or more eDiscovery holds, eDiscovery admin approval is also required.

Approver role requirements

The approver at each stage of the review process must have the correct roles assigned before the policy can be created.

Expand table

Reviewer	Required roles
Priority cleanup admins	- Priority Cleanup admin - Data Classification Content Viewer - Data Classification List Viewer - Disposition Management
Retention managers	- Retention Management - Data Classification Content Viewer - Data Classification List Viewer - Disposition Management
eDiscovery admins	- Search And Purge - Hold - Review - Data Classification Content Viewer - Data Classification List Viewer - Disposition Management

> **Note**
>
> If the reviewer at any stage doesn't have the correct roles assigned in advance, policy creation fails with an error.

For instructions to add users to the default roles or create your own role groups, use the following guidance:

- Permissions in the Microsoft Purview portal

Although you can specify multiple approvers for each stage, just one person from each stage is required to approve for their stage.

Enable auditing

Make sure that auditing is enabled at least one day before you create and run the first priority cleanup policy. Auditing is also required to view simulation results if the policy is enabled in simulation mode. For more information, see Search the audit log.

Limitations of priority cleanup

- Group mailboxes are supported via adaptive scopes only.
- For the KeyQL query, some properties and conditions supported by eDiscovery aren't supported by priority cleanup. These include SenderAuthor, SubjectTitle, (c:c), and (c:s).
- In simulation mode, the priority cleanup policy may incorrectly show email items marked as records and regulatory records. These items are not actually in scope for priority cleanup outside simulation mode.
- Unlike disposition review for retention labels:
  - You can't customize the email notification
  - Approvers can't nominate additional approvers
  - There's no automatic approval after a specified period of time
- If an approver doesn't agree to permanently delete an identified item, they must assign an existing retention label (any configuration) to the item. Make sure your approvers know which retention labels are suitable for this action.
- Although you can delete a priority cleanup policy, if the approval process for it is complete, items might still be permanently deleted.

Create a priority cleanup policy

Decide before you create your priority cleanup policy whether to make it **adaptive** or **static**. For more information, see Adaptive or static policy scopes for retention. If you decide to use an adaptive policy, you must create one or more adaptive scopes before you create your priority cleanup policy, and then select them during the policy configuration. For instructions, see Configuration information for adaptive scopes.

1. Sign in to the Microsoft Purview portal > **Solutions** > **Data Lifecycle Management** > **Priority cleanup**, and then select **+ Create a priority cleanup**.

   If you don't see the **Priority cleanup** option, check your permissions.

2. Enter a name and description for this priority cleanup policy, and then select **Next**. The name is visible to end users, but the optional description is visible only to priority cleanup admins and the policy's specified approvers. This restriction means that any details you enter can be informative and specific, without worrying about unauthorized people seeing these details.

3. For **Choose the type of priority cleanup**, select **Adaptive** or **Static** then select **Next**.
   a. For Adaptive scope, add one or more adaptive scopes. For a user mailbox, add a user scope. For a group mailbox, add a Microsoft 365 Group scope. Then select one or more of the mailbox locations to match your scopes.
   b. For Static scope, select **Policy for Exchange Online** and then select Exchange Online location. Optionally, you can edit what mailboxes to include or exclude.

   To help you decide:

   - **All mailboxes are included**: Best when you're unsure where the content resides. It may take longer to apply the policy, but it's worth it if content might be in unknown mailboxes.
   - **Specific mailboxes**: Fastest option for a small, known set of mailboxes. Use to include or exclude up to 100 mailboxes confidently.
   - **User mailboxes by attributes**: You can't use a static scope if you want to use attributes to target user mailboxes by region, department, etc. Instead, go back in the policy configuration and choose to use adaptive scopes instead. However, avoid if the scope includes over 1 million mailboxes.

4. Select **Next**.

5. For the **Tell us what you're looking for** page, enter text into the KeyQL editor box to construct a query using Exchange email properties. You can refine your query by using search operators such as AND, OR, and NOT.

For example, to find all content sent after February 2, 2024, with an attachment named ContosoEmployeeSalaries.xlsx: **AttachmentNames:ContosoEmployeeSalaries.xlsx AND sent>=2024-02-02**

For more information about the query syntax that uses Keyword Query Language (KeyQL), see Keyword Query Language (KeyQL) syntax reference.

This query-based policy uses the same search index as eDiscovery content search to identify content. For more information about the searchable properties that you can use for email, see Finding content in Exchange Online.

6. For the **Choose when content should be deleted** page, choose whether to permanently delete the matched items as soon as possible or retain them for a specific period and then delete them. Most of the time, you'll select the first option so that you can delete the item as soon as possible. Use the alternative option only if the items should be retained for compliance reasons and you can't use a retention label for this purpose. For example, the item already has a retention label applied with a longer retention period.

   > ⓘ **Note**
   >
   > A priority cleanup policy overrides the principles of retention that normally determine when an item should be retained or permanently deleted.

7. For the **Assign who'll approve what gets deleted** page, this is where you need to specify another priority cleanup approver, an approver for when an identified item has retention settings applied (such as a retention policy, retention label, or litigation hold policy), and an approver for when an identified item has one or more eDiscovery holds applied.

   - **Priority cleanup admins**: Must be assigned the Priority Cleanup Admin role and is the first-stage approver for all priority cleanups for this policy. This should be a different person to the user who created the priority cleanup policy, but isn't enforced.
   - **Retention managers**: Must be assigned the Retention Management role. Approval from specified users is required if the identified content is subject to one or more retention policies or Litigation holds.
   - **eDiscovery admins**: Must be assigned the eDiscovery Administrator role. Approval from specified users is required if the identified content is subject to one or more eDiscovery holds.

8. For the **Choose the policy mode** page, choose whether to run the policy first in simulation mode or turn it on, or neither for the time being.

Running the policy in simulation mode delays permanent deletion but allows you to verify sample matches and refine the query. It also lets others review the results before approval, even if they're not designated approvers.

9. Specific to priority cleanup, you must acknowledge by selecting a checkbox that you understand how this policy can override eDiscovery holds and other applied retention settings.

10. On the **Your priority cleanup policy has been created** page, you see the **Cleanup ID** that's used to track and monitor this policy. Use the Copy function, or copy it later from the policy details so you can monitor the progress of this policy from auditing details.

If you chose to run the policy in simulation mode:

- You might need to wait a couple of hours for results, depending on the number of mailboxes to search.
- You can turn on the policy for up to seven days. After seven days, the simulation must be restarted.

If you turn on the policy, as with auto-apply retention label policies, it can take up to seven days to apply the policy to items and trigger the approval process.

Approval process for a priority cleanup policy

When the priority cleanup policy is turned on and items are identified, approvers for the policy are notified by email, with a reminder once a week. They can select the link in the notification and reminder emails to go directly to the **Data lifecycle management** > **Priority cleanup** > **Pending cleanups** page in the portal to review the content to approve. Alternately, the approvers can manually navigate to this page in the portal.

To implement the security control that uses the two-person rule, each priority cleanup always requires another priority cleanup admin to approve the permanent deletion of identified items. Then if the items have retention settings applied, they require a next stage approval from retention admins. And finally, if the items are included in an eDiscovery hold, they also require another approval from an eDiscovery admin. When all the required approvals are complete, items are permanently deleted and cannot be restored by users, by admins, or by Microsoft.

On the **Pending cleanups** page, items identified by a priority cleanup policy are listed with a status of **Pending disposition** and an estimated count of how many items are identified. These might be different items, or the same item in multiple mailboxes.

When the approver selects one of the list items, the next page shows them the individual items with the item names, locations, and senders. When an item is selected, the preview pane

displays the item's subject, source, details, and history. The history displays any priority cleanup approvals to date for that item, with approver comments if available.

After reviewing all the items, the approver can individually or multi-select them, and select **Approve disposal**. They must then acknowledge the action with an optional comment, and select **Apply**.

Alternatively, if the item shouldn't be permanently deleted as soon as possible, the approver must select **Relabel**, and select an existing retention label.

Items that are approved or relabeled are then moved to the **Disposed items** tab. Allow up to seven days for items to be permanently deleted.

Export the views

An approver can use the **Export** option from the **Pending cleanups** and **Disposed items** pages to export information about the items in either view as a .csv file that they can then sort and manage with Excel.

How to monitor priority cleanup

You can monitor the status of priority cleanups for each policy from **Data lifecycle management** > **Priority cleanup**. For example, the status displays **In simulation**, or **Enabled (Pending)** that changes to **Enabled (Success)**.

Use the details of a policy to identify its cleanup ID, and paste this number as a keyword search string from the auditing solution. To use the date range, remember to specify the dates in UTC.

There are two auditing events specific to priority cleanup for mailboxes:

- Operation name of **PriorityCleanupTagApplied**: When an item is identified for priority cleanup, and if this resulted in removing an existing retention label.
- Operation name of **PriorityCleanupDelete**: The deletion of a mailbox item by priority cleanup.

Currently, these events don't have friendly names to select from the Microsoft Purview portal.

Other auditing events are the same as those used for creating and configuring a retention label and auto-labeling policy, and events for disposition review.

The end user experience for priority cleanup

Because priority cleanup doesn't use a soft-delete process, users see a **Retention:** message bar on their emails in Outlook when they're identified for priority cleanup. They also see the name of the priority cleanup policy, then **(-1 days)** to indicate that it should be deleted as soon as possible, and an estimated expiry day and time based on that -1 days.

For example, if your priority cleanup policy is named "Cleanup policy test":

Retention: Cleanup policy test (-1 days) Expires: Thu 2/6/2024 AM

> **Tip**
>
> If you prefer end users to not see the retention message, you can achieve this by first using **eDiscovery search and purge** that soft-deletes items. When that completes, then apply the priority cleanup policy to permanently delete the soft deleted items.

After the final priority cleanup approval, the item silently disappears from Outlook.

Turn off priority cleanup for the tenant

After considering the safeguards of additional permissions and multiple approvals, if your organization still has concerns about this capability, you can turn off the ability to create priority cleanup policies. When you turn off priority cleanup, you're turning it off for SharePoint and OneDrive as well as for Exchange.

1. Sign in to the Microsoft Purview portal > **Solutions** > **Data Lifecycle Management**.
2. From the top right, select **Priority cleanup settings**.
3. From the **Configuration** page, turn off the control for priority cleanup, and select **Save**.

New priority cleanup policies can't be created until you turn on the control and select **Save** again.

If priority cleanup policies are already created when you turn off the control:

- Existing priority cleanup policies continue to function
- Existing priority cleanup policies can be deleted
- Existing priority cleanup policies can't be modified

See Also

Override holds to clean up files for Copilot and reclaim storage

Last updated on 12/03/2025

Override holds to clean up files for Copilot and reclaim storage

> *Microsoft 365 licensing guidance for security & compliance.*

Use the **Priority cleanup** feature under **Data Lifecycle Management** in Microsoft Purview when you need to override existing retention settings or eDiscovery holds to delete content from OneDrive and SharePoint. Despite the existing holds, items are moved to the second-stage Recycle Bin and from there, go through the same processes with the same time periods to ensure a compliant deletion process.

Although priority cleanup might be needed for security or privacy in response to an incident, or for compliance with regulatory requirements, typical use cases for SharePoint and OneDrive include the following;

- **Delete stale Teams meeting recordings and transcripts, turned on for Copilot recap**: The Copilot in Teams recap feature uses Teams meeting recordings and transcripts that might be covered by a retention policy. These files are typically large and have little business value after 1-3 months. To reclaim the space from these files, or any other similarly large files that don't need to be retained as long as your retention policies, you can take action to remove them.

  > **Tip**
  >
  > Most meeting recordings and transcripts are saved to OneDrive. However, recordings from channel meetings are stored in SharePoint. For more information, see **Use OneDrive and SharePoint to store meeting recordings**.

  Priority cleanup lets you delete these files by overriding existing retention policies so you can better manage your storage requirements. Because these files are continually created, leave the priority cleanup policy running to automatically find and delete the stale files.

- **Delete labeled items in OneDrive's Preservation Hold library when a user leaves the organization**: When one or more items in a Preservation Hold library have an unexpired retention period, the OneDrive site for the user can't be deleted. As a result, when users leave the organization, you might not be able to delete their OneDrive site until the longest retention period expires. Priority cleanup can delete these files instead of waiting for the expiration period to end, so you can then delete the site for the OneDrive account and reclaim that OneDrive storage space. Create a priority cleanup policy after a user has left the organization and their OneDrive account is blocked from deletion.

Because the deletion from priority cleanup can override existing holds and even Preservation Lock when the retention settings are for delete-only, the process requires multiple approvals, specific roles, and is audited. After considering these safeguards, if your organization still has concerns about this capability, you can just continue to use retention policies and retention labels to ensure a compliant deletion of content instead of using priority cleanup.

Under the covers, priority cleanup uses retention labels with auto-apply policies. However, you don't interact manually with these labels and policies, and they supersede the principles of retention so you don't need to wait for the longest retention period to expire.

> **Note**
>
> If an item is subject to multiple priority cleanups, the newest takes priority.

Important exceptions for priority cleanup:

- You can't use priority cleanup for items that are marked as a record or regulatory record.
- Although priority cleanup can override eDiscovery holds, the data might already be copied to an eDiscovery review set where it can't be deleted by priority cleanup. This data is automatically deleted when the entire eDiscovery case is deleted by an eDiscovery admin.

Similar to auto-apply retention labels, priority cleanup supports simulation, to check the returned samples in case the policy configuration needs any fine-tuning. For SharePoint and OneDrive, you must run simulation at least one time, and for any change other than the policy description.

Comparing priority cleanup for different workloads

Although you can use the priority cleanup feature for mailboxes and for files in SharePoint and OneDrive, the typical use cases are different. There are some other differences in behavior and configuration that are summarized in the following table. As a result of these differences, you can't create a cleanup policy for all three locations in same policy.

Expand table

Behavior or configuration	Exchange	SharePoint/OneDrive
Typical use case:	Data spillage	Delete stale Teams meeting recordings and transcripts

Behavior or configuration	Exchange	SharePoint/OneDrive
	Delete sensitive data for compliance requirements	Delete retained files in the Preservation Hold library
Typical policy use:	Rare	For Teams meeting recordings and transcripts: Continual For the Preservation Hold library: After a user leaves the organization
Two-person rule:	Another priority cleanup admin is assigned in the policy to review items after the policy is turned on	Another priority cleanup admin is included in the policy configuration flow before the policy is turned on
Assign policy approvers:	- Priority cleanup admin - eDiscovery admin - Retention manager	- eDiscovery admin
Simulation required before running the policy:	No (but recommended)	Yes
Override Preservation Lock:	Yes	Only if the policy is configured for delete-only

Additionally, the required permissions for Exchange are different from the required permissions for SharePoint and OneDrive.

Prerequisites for priority cleanup

Make sure you can meet the prerequisites that must be in place before you can use priority cleanup to override existing holds for SharePoint and OneDrive. These requirements include permissions and approvers.

Because of the built-in safeguards, the feature itself is enabled by default at the tenant level. However, priority cleanup can be turned off on the priority cleanup settings page. If you can't create new priority cleanup policies, see the instructions for turning off the feature to check the status and reverse the configuration.

Permissions for priority cleanup

To successfully access and manage **Priority cleanup** in the Microsoft Purview portal, you must have the **Priority Cleanup Admin** role. For SharePoint and OneDrive, this role is required to create and manage priority cleanup policies, and enable or disable the feature. This role is

automatically added to the Organization Management role group but must be manually added to any other role group.

Alternatively, the **Priority Cleanup Viewer** role allows only the visibility of priority cleanup policies and settings without the ability to make changes or create new policies.

Content Explorer List Viewer and **Content Explorer Content Viewer** roles are required to view item content and details in simulation mode, and approval stages.

Similar to records management disposition, each person that accesses the **Priority cleanup** > **Pending cleanups** page sees only items that they're assigned to approve. To monitor the end-to-end process of a priority cleanup, use auditing and the priority cleanup ID as a search term.

Policy approvers

As a safeguard against accidental or malicious deletions, each item subject to priority cleanup always requires at least one other person to approve the deletion in addition to the person who created the priority cleanup policy. Approvers must be individual users. Mail-enabled security groups aren't currently supported.

- Priority cleanup admin approval is required for all policies.
- If an item is also subject to one or more eDiscovery holds, eDiscovery admin approval is also required.
- Separate approval from a retention admin isn't required to override retention settings.

Approver role requirements

The approver at each stage of the review process must have the correct roles assigned.

Expand table

Reviewer	Required roles
Priority cleanup admins	- Priority Cleanup Admin - Data Classification Content Viewer - Data Classification List Viewer - Retention Management
eDiscovery admins	- Search And Purge - Hold - Review - Data Classification Content Viewer - Data Classification List Viewer - Disposition Management

> ⓘ Note
>
> If the reviewer for eDiscovery admin doesn't have the correct roles assigned in advance, policy creation fails with an error.

For instructions to add users to the default roles or create your own role groups, use the following guidance:

- Permissions in the Microsoft Purview portal

Although you can specify multiple approvers for eDiscovery, just one person is required to approve overriding eDiscovery holds.

Enable auditing

Make sure that auditing is enabled at least one day before you create and run the first priority cleanup policy. Auditing is also required to view simulation results. For more information, see Search the audit log.

Limitations of priority cleanup

- In simulation mode, items under preservation lock where the retention settings are configured for delete-only may show up. These items are not actually in scope for priority cleanup outside simulation mode.
- Unlike disposition review for retention labels:
  - You can't customize the email notification
  - Approvers can't nominate additional approvers
  - There's no automatic approval after a specified period of time
- If an approver doesn't agree to delete an identified item, they must assign an existing retention label (any configuration) to the item. Make sure your approvers know which retention labels are suitable for this action.
- Although you can delete a priority cleanup policy, if the approval process for it is complete, items might still be deleted.

Create a priority cleanup policy

Decide before you create your priority cleanup policy whether to make it **adaptive** or **static**. For more information, see Adaptive or static policy scopes for retention. If you decide to use an

adaptive policy, you must have one or more adaptive scopes before you create your priority cleanup policy so that you can select them during the policy configuration. For instructions, see Configuration information for adaptive scopes.

1. Sign in to the Microsoft Purview portal > **Solutions** > **Data Lifecycle Management** > **Priority cleanup**, and then select **+ Create a priority cleanup**.

   If you don't see the **Priority cleanup** option, check your permissions.

2. Enter a name and description for this priority cleanup policy, and then select **Next**. The name is visible to end users, but the optional description is visible only to priority cleanup admins and the policy's specified approvers. This restriction means that any details you enter can be informative and specific, without worrying about unauthorized people seeing these details.

3. For **Choose the type of priority cleanup**, select **Adaptive** or **Static** then select **Next**.
   a. For Adaptive scope, add one or more adaptive scopes. For OneDrive accounts, add a scope for users. For sites, add a scope for SharePoint sites or Microsoft 365 Groups, as needed. Then select one or more of the locations to match your scopes.
   b. For Static scope, select **Policy for SharePoint/OneDrive** and then select one or more of the locations for sites and OneDrive accounts. Optionally, you can edit what sites and user accounts to include or exclude.

4. Select **Next**.

5. For the **Tell us what you're looking for** page, enter text into the KeyQL editor box to construct a query to target specific files. You can refine your query by using search operators such as AND, OR, and NOT.

   For example, to find Teams meeting recordings and transcripts, use **ProgID:Media AND ProgID:Meeting**. For files in a Preservation Hold library, use **ParentLink:PreservationHoldLibrary**.

   For more information about the query syntax that uses Keyword Query Language (KeyQL), see Keyword Query Language (KeyQL) syntax reference and Automatically apply a retention label to retain or delete content.

   This query-based policy uses the same search index as eDiscovery content search to identify content.

6. For the **Choose when content should be deleted** page, choose whether to delete the matched items as soon as possible or retain them for a specific period and then delete them. Most of the time, you'll select the first option so that you can delete the item as soon as possible. Use the alternative option only if the items should be retained for

compliance reasons and you can't use a retention label for this purpose. For example, the item already has a retention label applied with a longer retention period.

> **Note**
>
> A priority cleanup policy overrides the principles of retention that normally determine when an item should be retained or deleted.

7. For the **Assign who'll approve what gets deleted** page, this is where you need to specify an approver for when an identified item has one or more eDiscovery holds applied:

   - **eDiscovery admins**: Must be assigned the eDiscovery Administrator role. Approval from specified users is required if the identified content is subject to one or more eDiscovery holds.

8. For the **Choose the policy mode** page, choose whether to run the policy first in simulation mode or leave it turned off.

   Running the policy in simulation mode is required to initially set up priority cleanup for SharePoint and OneDrive locations. Simulation mode lets you verify that sample results match your query, so you can fine-tune it before approval. It also facilitates review with another priority cleanup admin.

9. Specific to priority cleanup, you must acknowledge by selecting a checkbox that you understand how this policy can override eDiscovery holds and other applied retention settings. Select **Submit**

10. On the **Your priority cleanup policy has been created** page, you see the **Cleanup ID** that's used to track and monitor this policy. Use the Copy function, or copy it later from the policy details so you can monitor the progress of this policy from auditing details.

11. Once simulation results are available, ask a second priority cleanup admin to review the results from the **Priority Cleanup Policies** page. They should select the policy from the list, and then select **View simulation details**.

12. If the results are satisfactory, this second priority cleanup admin can select **Turn on policy**. However, if the policy needs reconfiguration to add or remove items, they should make those edits and rerun simulation, and view results again. They should continue to make edits and rerun simulation each time. When the results are satisfactory, you, as the first priority cleanup admin, can turn on the policy.

> **Important**

> The last person to edit the policy can't also turn it on. We recommend that only one priority admin edits the policy after it's first created. So as a variation to step 12, the original priority cleanup admin could edit the policy and rerun simulation, and the second priority cleanup admin turns it on.

The Locations box shows scanned areas, and sample counts may differ from match counts. After the policy is turned on, wait for the "Enabled (Success)" status, which may take up to seven days.

Approval process for a priority cleanup policy

When the priority cleanup policy is turned on and items are identified that are subject to eDiscovery, your assigned eDiscovery approvers for the policy are notified by email, with a reminder once a week. They can select the link in the notification and reminder emails to go directly to the **Data lifecycle management** > **Priority cleanup** > **Pending cleanups** page in the portal to review the content to approve. Alternately, the approvers can manually navigate to this page in the portal.

When any required approvals are complete, items are moved to the second-stage Recycle Bin.

On the **Pending cleanups** page, items identified by a priority cleanup policy are listed with a status of **Pending disposition** and an estimated count of how many items are identified. These might be different items, or the same item in multiple locations.

When the approver selects one of the list items, the next page shows them the individual items with the item names, locations, and senders. When an item is selected, the preview pane displays the item's subject, source, details, and history. The history displays any priority cleanup approvals to date for that item, with approver comments if available.

After reviewing all the items, the approver can individually or multi-select them, and select **Approve disposal**. They must then acknowledge the action with an optional comment, and select **Apply**.

Alternatively, if the item shouldn't be deleted, the approver must select **Relabel**, and select an existing retention label.

Items that are approved or relabeled are then moved to the **Disposed items** tab. Allow up to seven days for items to be deleted.

Export the views

An approver can use the **Export** option from the **Pending cleanups** and **Disposed items** pages to export information about the items in either view as a .csv file that they can then sort and manage with Excel.

How to monitor priority cleanup

You can monitor the status of priority cleanups for each policy from **Data lifecycle management** > **Priority cleanup**. For example, the status displays **In simulation**, or **Enabled (Pending)** that changes to **Enabled (Success)**.

Use the details of a policy to identify its cleanup ID, and paste this number as a keyword search string from the auditing solution. To use the date range, remember to specify the dates in UTC.

There are two auditing events specific to priority cleanup for SharePoint and OneDrive:

- Operation name of **PriorityCleanupTagApplied**: When an item is identified for priority cleanup, and if this resulted in removing an existing retention label.
- Operation name of **PriorityCleanupFileRecycled**: The deletion of an item in SharePoint or OneDrive by priority cleanup.

Currently, these events don't have friendly names to select from the Microsoft Purview portal.

Other auditing events are the same as those used for creating and configuring a retention label and auto-labeling policy, and events for disposition review.

The end user experience for priority cleanup

Files that are identified for priority cleanup display to users as an applied retention label that has the same name as the priority cleanup policy.

If users remove or replace this retention label, it will be automatically reapplied by the policy.

Turn off priority cleanup for the tenant

After considering the safeguards of additional permissions and multiple approvals, if your organization still has concerns about this capability, you can turn off the ability to create priority cleanup policies. When you turn off priority cleanup, you're turning it off for Exchange as well as for SharePoint and OneDrive.

1. Sign in to the Microsoft Purview portal > **Solutions** > **Data Lifecycle Management**.
2. From the top right, select **Priority cleanup settings**.
3. From the **Configuration** page, turn off the control for priority cleanup, and select **Save**.

New priority cleanup policies can't be created until you turn on the control and select **Save** again.

If priority cleanup policies are already created when you turn off the control:

- Existing priority cleanup policies continue to function
- Existing priority cleanup policies can be deleted
- Existing priority cleanup policies can't be modified

See Also

Use priority cleanup to expedite the permanent deletion of sensitive information from mailboxes

---

Last updated on 11/17/2025

Permanently delete files with Microsoft Purview Priority Cleanup

> *Microsoft Purview service description*

Use Microsoft Purview Priority Cleanup to permanently delete targeted files from SharePoint and OneDrive after the required review and approval. This option bypasses the SharePoint and OneDrive recycle bins. After deletion, the content is no longer discoverable in SharePoint search, Microsoft 365 Copilot, or eDiscovery.

Permanent deletion is an additional option for a Priority Cleanup policy. Use the standard Priority Cleanup workflow when you want identified items to move to the second-stage Recycle Bin.

Before you begin, review the required roles, approvers, auditing, and simulation requirements.

Prerequisites

Before you configure permanent deletion:

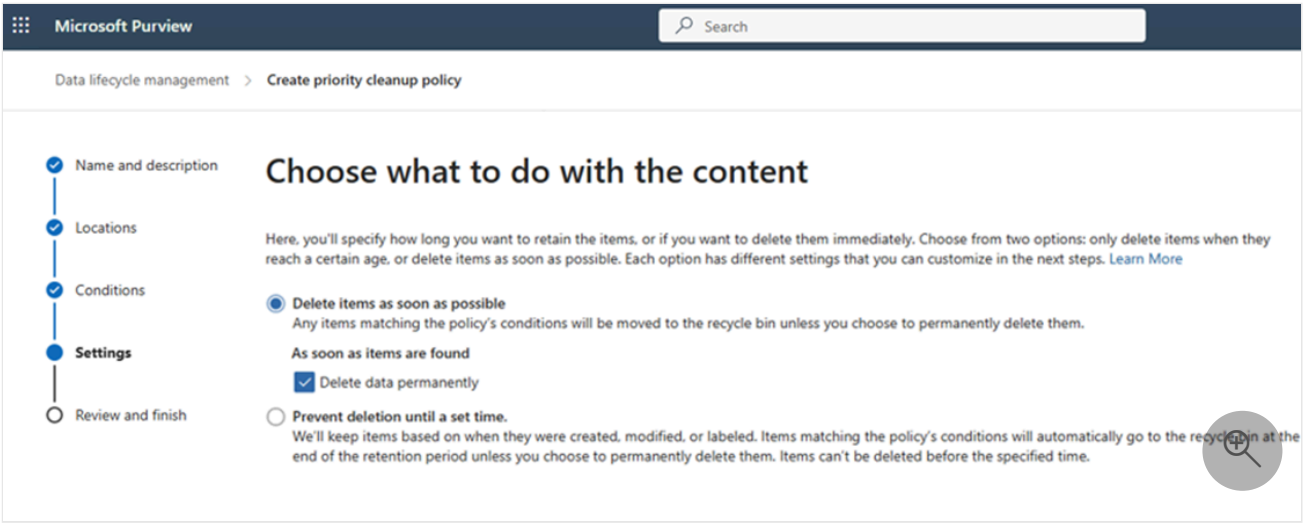
- Review the prerequisites, permissions, approver requirements, and simulation process in Override holds to clean up files for Copilot and reclaim storage.
- Assign the **Priority Cleanup Admin** role to the administrator who creates and manages the policy.
- Assign the required roles to each approver before policy creation.
- Enable auditing at least one day before you create and run the first Priority Cleanup policy.
- Confirm that the permanent deletion option is available in your tenant. Public preview rollout begins on August 24, 2026.

> ⓘ **Important**
>
> Permanent deletion can override retention settings or eDiscovery holds. Items already copied to an eDiscovery review set aren't deleted by Priority Cleanup.

Configure permanent deletion

Create a Priority Cleanup policy and explicitly select the permanent deletion option.

1. Sign in to the Microsoft Purview portal and go to **Solutions** > **Data Lifecycle Management** > **Priority cleanup**.
2. Select + **Create a priority cleanup**.
3. Enter the policy name and description, choose an adaptive or static scope, select the SharePoint or OneDrive locations, and define the query for the targeted files.
4. On the **Choose what to do with the content** page, select **Delete data permanently**.



5. Choose whether to delete matched items as soon as possible or prevent deletion until a set time.
6. Assign the required approvers.
7. Run the policy in simulation mode and review the returned samples.
8. After the simulation results are satisfactory, have a second Priority Cleanup admin review the results and turn on the policy.

> **Important**
>
> The last person to edit the policy can't also turn it on.

Review and approve permanent deletion

Required approvers review identified items before permanent deletion occurs.

1. In the Microsoft Purview portal, go to **Data Lifecycle Management** > **Priority cleanup** > **Pending cleanups**.

2. Select the cleanup and review the identified items, their locations, and their history.
3. Select the items to delete, and then select **Approve disposal**.
4. Acknowledge the action, optionally add a comment, and then select **Apply**.

After all required approvals are complete, the selected items bypass the recycle bins and are permanently deleted. The deleted content is no longer discoverable in SharePoint search, Microsoft 365 Copilot, or eDiscovery.

If an item shouldn't be deleted, select **Relabel**, and then assign an existing retention label.

Monitor permanent deletion

Use the cleanup ID and auditing to monitor the policy and deletion activity.

1. Go to **Data Lifecycle Management** > **Priority cleanup** and open the policy details.
2. Copy the cleanup ID.
3. In the auditing solution, use the cleanup ID as the keyword search string.
4. Check for the **PriorityCleanupFileDeleted** operation, which indicates that Priority Cleanup permanently deleted an item.

Related content

- Override holds to clean up files for Copilot and reclaim storage
- Search the audit log
- Use priority cleanup to expedite the permanent deletion of sensitive information from mailboxes

ⓘ**Note:** This article contains content created with AI. Learn more

---

Last updated on 08/24/2026

Remove a retention hold from an unlicensed OneDrive site

> *Microsoft Purview service description*

This article is for Microsoft Purview administrators who manage unlicensed OneDrive for Business sites. Use the `Fix-PurviewConfig` cmdlet in Security & Compliance PowerShell to force-exclude a specified retention hold from an unlicensed, archived site. Removing the blocking hold allows the site to be deleted when no other non-force-excluded holds apply.

Before you begin, review the account, site, and policy requirements in Prerequisites.

Prerequisites

Before you begin, make sure you have:

- An account that can connect to Security & Compliance PowerShell.
- The URL of the unlicensed OneDrive for Business site.
- The GUID of the retention or hold policy to remove. The policy must exist in EOP and apply to the specified site.
- Confirmation that the site owner is unlicensed and the OneDrive site is archived. For more information, see Manage unlicensed OneDrive user accounts.

Remove the retention hold

The `RemoveHoldFromUnlicencedODBSite` scenario removes the specified hold from the site. Review your organization's retention and legal requirements before you run the cmdlet because removing the blocking hold can allow site deletion.

1. Connect to Security & Compliance PowerShell.
2. Confirm the values that you will use for `SiteUrl` and `PolicyId`.
3. Run the `Fix-PurviewConfig` cmdlet with the following syntax:

```text
Fix-PurviewConfig -Scenario RemoveHoldFromUnlicencedODBSite
    -SiteUrl <String>
```

```
    -PolicyId <String>
```

For example:

PowerShell

```
Fix-PurviewConfig -Scenario RemoveHoldFromUnlicencedODBSite `
  -SiteUrl "https://<tenant>-my.sharepoint.com/personal/<user>" `
  -PolicyId "430b7b03-88b7-4ec5-9ae1-736f27bac9c3"
```

4. Review the result. A successful operation returns the following fields:

text

```
ScenarioName  : DLM_ODB_Remove_Hold_For_Unlicenced_Site
Description   : Remove hold from unlicenced ODB site.
Status        : Success
Comments      : {Successfully removed hold(s) from unlicenced ODB site.}
CorrelationId : 00000000-0000-0000-0000-000000000000
```

After successful execution, the specified hold is force-excluded. If no other non-force-excluded holds apply, site deletion is allowed.

You can run the command again without changing the result. If the hold is already excluded, the cmdlet reports that the hold has already been removed.

Resolve validation errors

Use the following table to resolve validation errors returned by the cmdlet.

Expand table

Validation error	Action
`SiteUrl` isn't provided.	Provide the URL of the unlicensed OneDrive for Business site.
`SiteUrl` isn't a valid OneDrive for Business URL.	Correct the URL format and run the command again.
The specified site isn't present in OneDrive for Business.	Confirm that the site exists and that you entered the correct URL.
`PolicyId` isn't provided.	Provide the retention or hold policy GUID.
`PolicyId` doesn't exist in EOP.	Confirm the policy ID and run the command again.

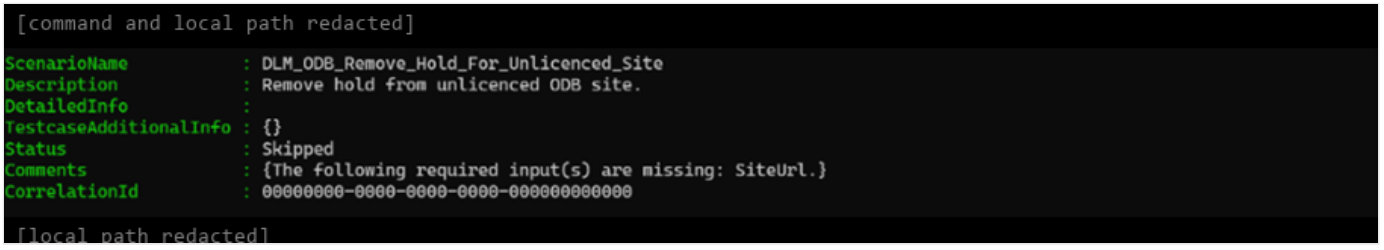
Validation error	Action
`PolicyId` isn't a valid GUID.	Provide a valid policy GUID.
The hold doesn't apply to the site.	Confirm that the policy applies to the specified site.
The site owner is licensed.	Remove the hold only after the site owner is unlicensed.

Review validation screenshots

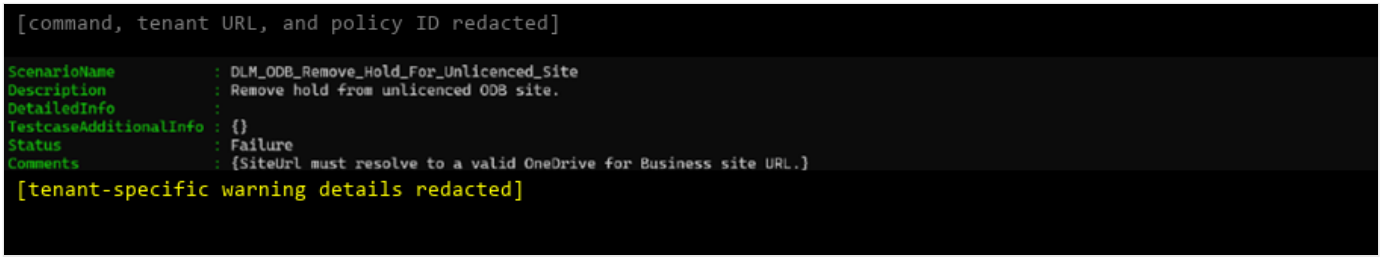
The following screenshots show the validation outcomes from the approved test run. Tenant URLs, user aliases, local file paths, policy identifiers, browser property values, and exception internals are redacted. The cmdlet results and status messages are unchanged.

Validate the site URL

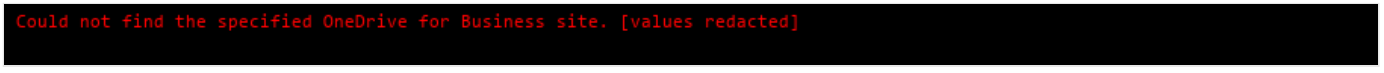
Site URL isn't provided



Site URL isn't in a valid OneDrive for Business format



Site URL isn't present in OneDrive for Business



Validate the policy ID

Policy ID isn't provided

```
[command, tenant URL, and local path redacted]

ScenarioName           : DLM_ODB_Remove_Hold_For_Unlicenced_Site
Description            : Remove hold from unlicenced ODB site.
DetailedInfo           :
TestcaseAdditionalInfo : {}
Status                 : Skipped
Comments               : {The following required input(s) are missing: PolicyId.}
CorrelationId          : 00000000-0000-0000-0000-000000000000
```

Policy ID doesn't exist in EOP

```
[command, tenant URL, and policy ID redacted]

ScenarioName           : DLM_ODB_Remove_Hold_For_Unlicenced_Site
Description            : Remove hold from unlicenced ODB site.
DetailedInfo           :
TestcaseAdditionalInfo : {}
Status                 : Failure
Comments               : {The provided PolicyId does not exist in EOP. The policy may have been deleted. To remove an orphan hold, please use the

[exception stack and tenant-specific identifiers redacted]
```

Policy ID isn't a valid GUID

```
[command, tenant URL, and policy ID redacted]

ScenarioName           : DLM_ODB_Remove_Hold_For_Unlicenced_Site
Description            : Remove hold from unlicenced ODB site.
DetailedInfo           :
TestcaseAdditionalInfo : {}
Status                 : Failure
Comments               : {PolicyId must contain one or more valid GUID values separated by ';' or ','. }

[tenant-specific warning details redacted]
```

Validate hold applicability and licensing

Hold isn't applicable to the site

```
[command, tenant URL, and policy ID redacted]

ScenarioName           : DLM_ODB_Remove_Hold_For_Unlicenced_Site
Description            : Remove hold from unlicenced ODB site.
DetailedInfo           :
TestcaseAdditionalInfo : {}
Status                 : Failure
Comments               : {The hold associated with the provided PolicyId is not applicable to this site.}
CorrelationId          : 00000000-0000-0000-0000-000000000000

[tenant-specific warning details redacted]
```

Site owner is licensed

```
[command, tenant URL, and policy ID redacted]

ScenarioName           : DLM_ODB_Remove_Hold_For_Unlicenced_Site
Description            : Remove hold from unlicenced ODB site.
DetailedInfo           :
TestcaseAdditionalInfo : {}
Status                 : Failure
Comments               : {Hold removal skipped because the ODB user is still licensed.}
CorrelationId          : 00000000-0000-0000-0000-000000000000
[tenant-specific warning details redacted]
```

Verify site-level hold removal

Before cmdlet execution

```
Before: site-level hold present. [browser URL and property values redacted]
```

Successful cmdlet execution

```
[command, tenant URL, and policy ID redacted]

ScenarioName           : DLM_ODB_Remove_Hold_For_Unlicenced_Site
Description            : Remove hold from unlicenced ODB site.
DetailedInfo           :
TestcaseAdditionalInfo : {}
Status                 : Success
Comments               : {Successfully removed hold(s) from unlicenced ODB site.}
CorrelationId          : 00000000-0000-0000-0000-000000000000
```

After cmdlet execution

```
After: site-level hold excluded. [browser URL and property values redacted]
```

Verify tenant-level hold removal

Before cmdlet execution

```
Before: tenant-level hold isn't stamped in AllWebHolds. [values redacted]
```

Successful cmdlet execution

```
[command, tenant URL, and policy ID redacted]

ScenarioName           : DLM_ODB_Remove_Hold_For_Unlicenced_Site
Description            : Remove hold from unlicenced ODB site.
DetailedInfo           :
TestcaseAdditionalInfo : {}
Status                 : Success
Comments               : {Successfully removed hold(s) from unlicenced ODB site.}
CorrelationId          : 00000000-0000-0000-0000-000000000000
```

After cmdlet execution

```
After: tenant-level hold excluded. [browser URL and property values redacted]
```

Verify repeated execution

Hold is already excluded

```
[first command, tenant URL, and policy ID redacted]

ScenarioName           : DLM_ODB_Remove_Hold_For_Unlicenced_Site
Description            : Remove hold from unlicenced ODB site.
DetailedInfo           :
TestcaseAdditionalInfo : {}
Status                 : Success
Comments               : {Successfully removed hold(s) from unlicenced ODB site.}
CorrelationId          : 00000000-0000-0000-0000-000000000000
[second command, tenant URL, and policy ID redacted]

ScenarioName           : DLM_ODB_Remove_Hold_For_Unlicenced_Site
Description            : Remove hold from unlicenced ODB site.
DetailedInfo           :
TestcaseAdditionalInfo : {}
Status                 : Failure
Comments               : {The hold associated with the provided PolicyId is not applicable to this site.}
CorrelationId          : 00000000-0000-0000-0000-000000000000
[tenant-specific warning details redacted]
```

Related content

- Learn about retention policies and retention labels
- Manage unlicensed OneDrive user accounts
- Security & Compliance PowerShell

ⓘ**Note:** This article contains content created with AI. Learn more

---

Last updated on 08/17/2026

Learn about archive mailboxes

> *Microsoft Purview service description*

Use this article if you're an Exchange or compliance administrator and want to learn about mailbox archives that provide more storage space for users' online mailboxes. This additional mailbox storage might be needed to support your organization's message retention, eDiscovery, and hold requirements.

> **Note**
>
> Mailbox users can learn more about their archive mailbox storage by using the following article: Manage email storage with online archive mailboxes

After you turn on archive mailboxes, sometimes also known as *In-Place Archiving*, a user's current mailbox becomes their *primary mailbox* and an additional and associated mailbox is created, called the *archive mailbox*.

Both mailboxes are considered a user's mailbox for compliance features such as Content search from the Microsoft Purview portal, Microsoft 365 retention, and Litigation Hold.

Users can access and store messages in their archive mailboxes by using Outlook and Outlook on the web. Users can also move or copy messages between their primary mailbox and their archive mailbox. They can also recover deleted items from the Recoverable Items folder in their archive mailbox by using the Recover Deleted Items tool.

Managing archive mailboxes with messaging records management (MRM)

Messages can also be moved automatically to the archive mailbox by the default Exchange retention policy from the messaging records management (MRM) feature. This default policy is automatically assigned to every mailbox and does the following actions:

- Moves items that are two years or older from a user's primary mailbox to their archive mailbox.

- Moves items that are 14 days or older from the Recoverable Items folder in the user's primary mailbox to the Recoverable Items folder in their archive mailbox.

You can customize your organization's Exchange MRM policy with retention tags. For an example configuration, see Customize an archive and deletion policy for mailboxes in your organization.

> ⓘ **Note**
>
> MRM, like Microsoft 365 retention policies and retention labels, can also automatically delete emails after a specified period. As an older technology than Microsoft 365 retention, MRM continues to work side-by-side with retention policies and retention labels from Microsoft Purview. For more information, see Use retention policies and retention labels instead of older features.

Auto-expanding archiving

After a user's archive mailbox is enabled, up to 100 GB of extra storage becomes available. If users need more storage space than this 100 GB, enable auto-expanding archiving to provide up to 1.5 TB of additional storage in archive mailboxes. For more information, see Learn about auto-expanding archiving.

> ⓘ **Important**
>
> Microsoft is increasing the maximum auto-expanding archive capacity for eligible Exchange Online mailboxes from 1.5 TB to 3 TB. Rollout begins in mid-October 2026 and is expected to complete by late October.
>
> The increase applies to mailboxes assigned the following eligible licenses:
>
> - Office 365 E5, A5, or G5
> - Microsoft 365 E5, A5, or G5
> - Microsoft 365 Purview Suite
> - Microsoft 365 Purview Suite FLW, EDU, or GOV
> - Microsoft Defender + Purview Suite FLW
> - Microsoft 365 F5 Compliance
> - Microsoft 365 F5 Security + Compliance
> - The combination of the eDiscovery & Audit, Insider Risk Management, and Information Protection & Governance mini suites

Qualifying offerings that include Microsoft Purview E5 capabilities, such as E7, also support the increase. **For all other licenses, the existing 1.5 TB limit remains unchanged.**

For E5 licenses, verify that **Records Management** is selected and remains enabled in the user's license settings.

If auto-expanding archiving is already enabled, no additional configuration or per-GB charge applies. Existing archive content, retention policies, and holds are unchanged. Verify license eligibility and monitor archive growth.

Auto-archiving

If a user's primary mailbox fills up, auto-archiving helps prevent mail flow disruptions by automatically moving the oldest items to the archive mailbox when mailbox usage exceeds a safe threshold, which is 96% of the mailbox quota by default. Unlike the time-based MRM archive policy, auto-archiving is *threshold-based*: it's triggered by mailbox size rather than by item age, and it complements your existing archive and deletion policies without changing them.

Auto-archiving requires the archive mailbox to be provisioned and to have remaining storage capacity. Items tagged **Never Move to Archive** and recently received items are excluded, and the threshold is configurable for the organization. Setting the threshold to 100 turns off auto-archiving for the tenant. For more information, see Auto-archiving in Exchange Online.

Next steps

See Enable archive mailboxes in Microsoft 365.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 10/14/2026

Enable archive mailboxes for Microsoft 365

> *Microsoft Purview service description*

Email archiving in Microsoft 365 (also called *In-Place Archiving*) provides users with more mailbox storage space. For more information, see Learn about archive mailboxes.

Use the information in this article to enable or disable an archive mailbox by using the Exchange admin portal or by using PowerShell. Also learn how to run an automated diagnostic check on a user's archive mailbox to identify any problems and suggested resolutions.

Configuration to enable or disable archive mailboxes has moved to the new Exchange admin center (EAC).

Get the necessary permissions

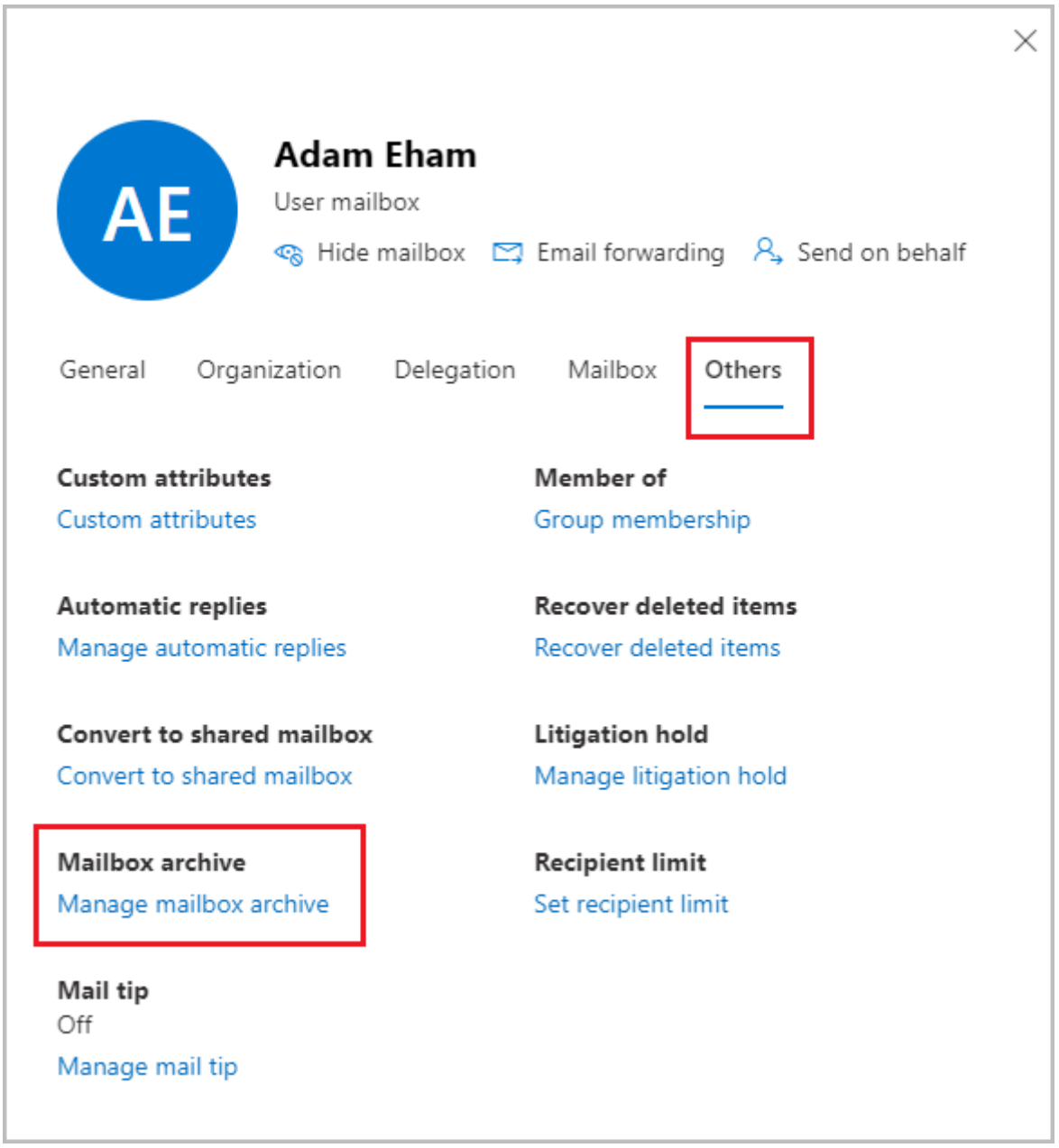
You must be assigned the Mail Recipients role in Exchange Online to enable or disable archive mailboxes. By default, this role is assigned to the Recipient Management and Organization Management role groups on the **Permissions** page in the Exchange admin center.

How to enable an archive mailbox

> **Note**
>
> When you enable the archive mailbox, items in the user's mailbox that are older than the archiving policy assigned to the mailbox will be moved to the new archive mailbox.
>
> The default archive policy that is part of the retention policy assigned to Exchange Online mailboxes moves items to the archive mailbox two years after the date the item was delivered to the mailbox or created by the user. For more information, see Learn about archive mailboxes.

1. Sign in to the Exchange admin center (EAC) and navigate to **Recipients** > **Mailboxes**.
2. In the list of mailboxes, select the user to enable their mailbox for archive.

3. In the flyout pane, select **Others**, and under **Mailbox archive**, select **Manage mailbox archive**:



4. On the **Manage mailbox archive** pane, turn on **Mailbox archive**, and then **Save**.

It might take a few moments to create the archive mailbox. When it's created, **Active** is displayed in the **Archive status** column for the selected user, although you might need to refresh the page to see the change of status.

How to auto-enable archive mailboxes

You can use the Set-OrganizationConfig PowerShell command to automatically provision an archive mailbox when a primary Exchange Online mailbox reaches 90% of its storage quota. An archive mailbox is created automatically only if all of the following conditions are met:

- The mailbox is a User Mailbox or Shared Mailbox in Exchange Online.
- Archiving is included in the Microsoft 365 assigned license.
- The mailbox is not synchronized from an on-premises environment (the IsDirSynced property is set to False).

PowerShell

```
Set-OrganizationConfig -AutoEnableArchiveMailbox $true
```

How to disable an archive mailbox

Similarly to how you enable an archive mailbox, you can use the same configuration in the Exchange admin center to disable a user's archive mailbox. This time, turn **Mailbox archive** off in the Exchange admin center.

After you disable an archive mailbox, you can reconnect it to the user's primary mailbox within 30 days of disabling it. In this case, the original contents of the archive mailbox are restored. After 30 days, the contents of the original archive mailbox are permanently deleted and can't be recovered. So if you re-enable the archive more than 30 days after disabling it, a new archive mailbox is created.

The default archive policy assigned to users' mailboxes moves items to the archive mailbox two years after the date the item is delivered. If you disable a user's archive mailbox, no action will be taken on mailbox items and they'll remain in the user's primary mailbox.

Use Exchange Online PowerShell to enable or disable archive mailboxes

You can also use Exchange Online PowerShell to enable archive mailboxes. The primary reason to use PowerShell is that you can quickly enable the archive mailbox for all users in your organization.

The first step is to connect to Exchange Online PowerShell. For instructions, see Connect to Exchange Online PowerShell.

After you're connected to Exchange Online, you can run the commands in the following sections to enable or disable archive mailboxes.

Enable archive mailboxes

Run the following command to enable the archive mailbox for a single user.

PowerShell

```
Enable-Mailbox -Identity <username> -Archive
```

Run the following command to enable the archive mailbox for all users in your organization whose archive mailbox is currently not enabled and who do not already have a disabled archive:

PowerShell

```
Get-Mailbox -Filter {ArchiveGuid -Eq "00000000-0000-0000-0000-000000000000" -AND
DisabledArchiveGuid -Eq "00000000-0000-0000-0000-000000000000" -AND
RecipientTypeDetails -Eq "UserMailbox"} | Enable-Mailbox -Archive
```

> ⓘ **Note**
>
> If archive for a mailbox was disabled more than 30 days ago, the command fails with the following error:
>
> "Write-ErrorMessage : |Microsoft.Exchange.Data.Directory.MissingDisconnectReceiptsException|AD modification is not allowed - disconnect of shard with ExchangeGuid *GUID* is not authorized".
>
> To resolve this error, run the Set-Mailbox cmdlet for the mailbox with -RemoveDisabledArchive. Then run the command again to enable the archive mailbox.

When enabling archiving for a specific user whose mailbox is on hold or assigned to a retention policy, the following configuration changes are implemented:

- The storage quota for the user's primary mailbox's Recoverable Items folder is expanded by 5 GB, raising it from 100 GB to 105 GB. This adjustment provides additional space to accommodate deleted items during the archiving process, even if auto-expanding archive isn't enabled. The Recoverable Items warning quota is also augmented by 5 GB, increasing it from 90 GB to 95 GB.

Disable archive mailboxes

Run the following command to disable the archive mailbox for a single user.

PowerShell

```
Disable-Mailbox -Identity <username> -Archive
```

Run the following command to disable the archive mailbox for all users in your organization (whose archive mailbox is currently enabled).

PowerShell

```
Get-Mailbox -Filter {ArchiveGuid -Ne "00000000-0000-0000-0000-000000000000" -AND
RecipientTypeDetails -Eq "UserMailbox"} | Disable-Mailbox -Archive
```

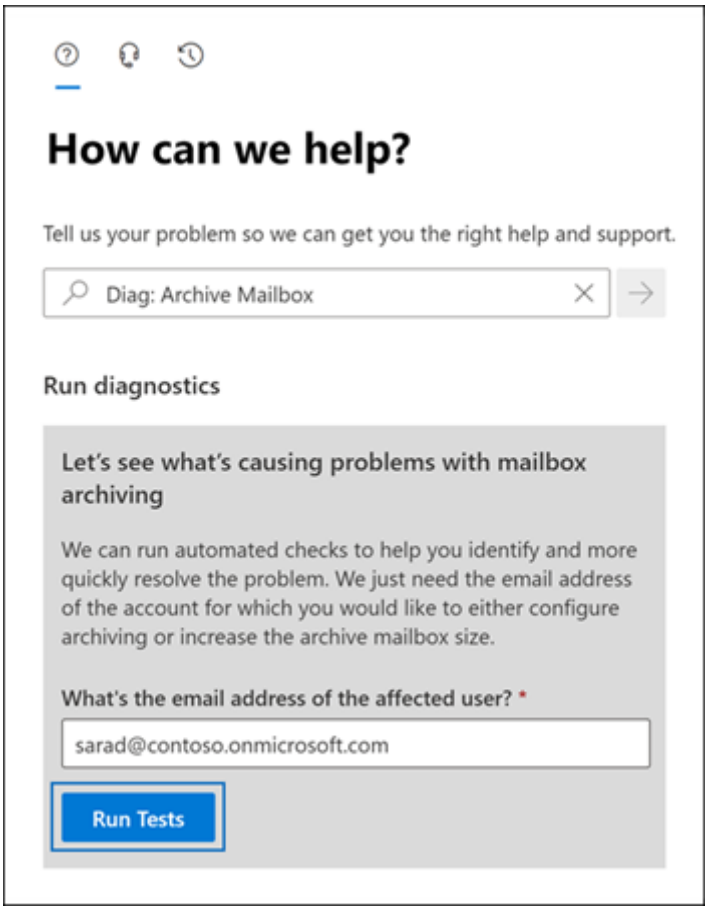
Run diagnostics on archive mailboxes

> ⓘ **Important**
>
> You must be a Microsoft 365 global admin to use the archive mailbox diagnostic check.
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

You can run an automated diagnostic check on a user's archive mailbox to identify any problems and suggested resolutions.

To run the diagnostic check, click the button below.

Run Tests: Archive Mailbox



A flyout page opens in the Microsoft 365 admin center. Enter the email address of the mailbox you want to check and click **Run Tests**.

> **Note**
>
> this feature isn't available in Microsoft 365 Government clouds, Microsoft 365 operated by 21Vianet, or Microsoft 365 Germany.

Instructions for end users

Explain to users how their archive mailbox works, and how they can interact with it in Outlook on Windows, macOS, and the web. The most effective documentation will be customized for your organization. But for basic instructions, see Manage email storage with online archive mailboxes.

Next steps

Consider enabling auto-expanding archiving for additional storage space. For instructions, see Enable auto-expanding archiving.

Last updated on 04/06/2026

Learn about auto-expanding archiving

> *Microsoft Purview service description*

Archive mailboxes in Microsoft Purview provide user mailboxes and shared mailboxes with additional mailbox storage space. After an archive mailbox is enabled, up to 100 GB of additional storage becomes available.

This archiving feature in Microsoft Purview (called *auto-expanding archiving*) provides up to 1.5 TB of additional storage in archive mailboxes. When the storage quota in the archive mailbox is reached, Microsoft Purview automatically (and incrementally) increases the size of the archive until the archive mailbox reaches 1.5 TB.

For step-by-step instructions for turning on auto-expanding archiving, see Enable auto-expanding archiving.

How auto-expanding archiving works

When auto-expanding archiving is enabled for a user's mailbox archive or a shared mailbox archive, Microsoft Purview periodically checks the size of the archive mailbox. When the archive mailbox gets close to its storage limit, additional storage space is automatically created and this process continues until the mailbox archive reaches 1.5 TB. The additional storage creation happens automatically, which means administrators don't have to request additional archive storage or manage auto-expanding archiving.

ⓘ Important

Microsoft is increasing the maximum auto-expanding archive capacity for eligible Exchange Online mailboxes from 1.5 TB to 3 TB. Rollout begins in mid-October 2026 and is expected to complete by late October.

The increase applies to mailboxes assigned the following eligible licenses:

- Office 365 E5, A5, or G5
- Microsoft 365 E5, A5, or G5
- Microsoft 365 Purview Suite
- Microsoft 365 Purview Suite FLW, EDU, or GOV

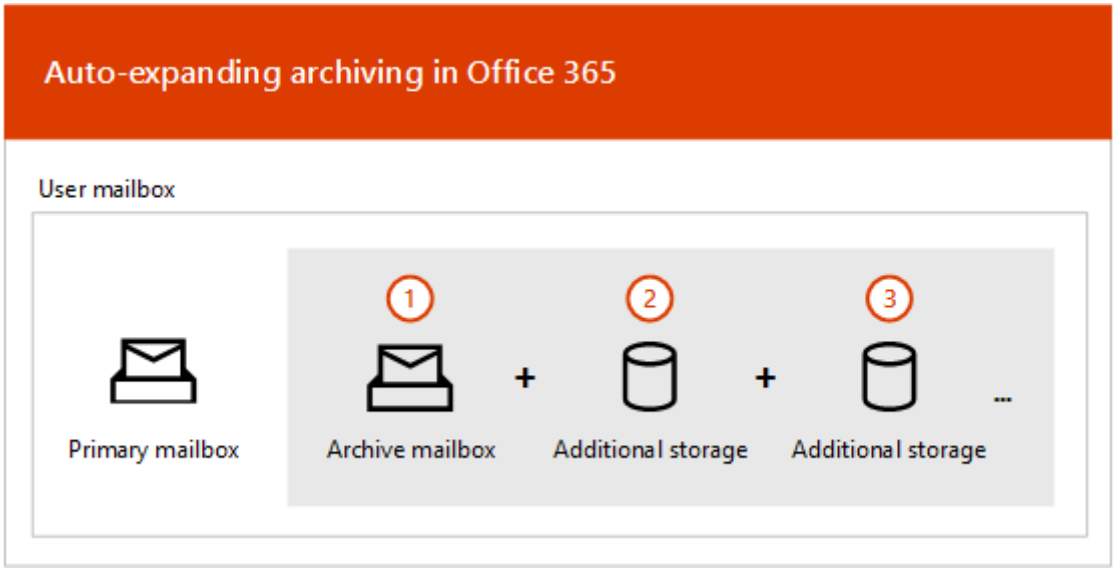
- Microsoft Defender + Purview Suite FLW
- Microsoft 365 F5 Compliance
- Microsoft 365 F5 Security + Compliance
- The combination of the eDiscovery & Audit, Insider Risk Management, and Information Protection & Governance mini suites

Qualifying offerings that include Microsoft Purview E5 capabilities, such as E7, also support the increase. **For all other licenses, the existing 1.5 TB limit remains unchanged.**

For E5 licenses, verify that **Records Management** is selected and remains enabled in the user's license settings.

If auto-expanding archiving is already enabled, no additional configuration or per-GB charge applies. Existing archive content, retention policies, and holds are unchanged. Verify license eligibility and monitor archive growth.

Here's a quick overview of the process.



1. Archiving is enabled for a user mailbox or a shared mailbox. An archive mailbox with 100 GB of storage space is created, and the warning quota for the archive mailbox is set to 90 GB.

2. An administrator enables auto-expanding archiving for the mailbox. If the mailbox has a hold or retention policy applied to it, the storage quota for the archive mailbox is increased to 110 GB and the archive warning quota is increased to 100 GB.

   Then, when the archive mailbox (including the Recoverable Items folder) reaches its storage quota, the archive mailbox is converted to an auto-expanding archive. Extra storage space is

added until it reaches a maximum size of 1.5 TB. It can take up to 30 days for the additional storage space to be provisioned.

3. Microsoft Purview automatically adds more storage space when necessary.

For information what happens when the maximum 1.5 TB is reached, and other size-related limitations for email storage, see Important considerations and limitations for auto-expanding archiving on this page.

What gets moved to the additional archive storage space?

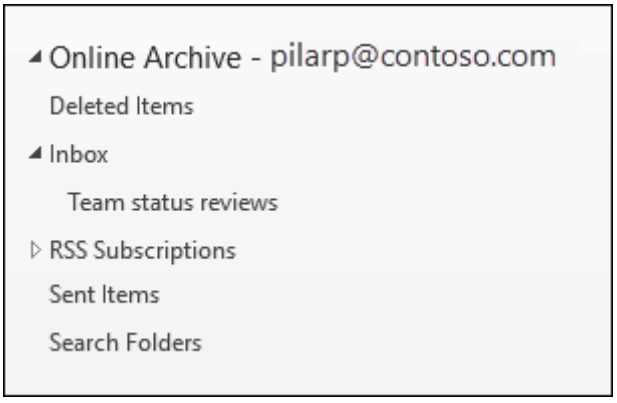
To make efficient use of auto-expanding archive storage, folders may get moved. Microsoft Purview determines which folders get moved when additional storage is added to the archive. Sometimes when a folder is moved, one or more subfolders are automatically created and items from the original folder are distributed to these folders to facilitate the moving process. You might need to communicate this behavior to end users after enabling their mailbox for auto-expanding archives, to help set expectations.

When viewing the archive portion of the folder list in Outlook, these subfolders are displayed under the original folder. The naming convention that Microsoft 365 uses to name these subfolders is **<folder name>_yyyy (Created on mmm dd, yyyy h_mm)**, where:

- **yyyy** is the year the messages in the folder were received.
- **mmm dd, yyyy h_m** is the date and time that the subfolder was created by Office 365, in UTC format, based on the user's time zone and regional settings in Outlook.

The following screenshots show a folder list before and after messages are moved to an auto-expanded archive.

Before additional storage is added



After additional storage is added



> **Note**
>
> In some scenarios, a small number of archive mailboxes might display "Inbox_Archive" and/or "Sent Items_Archive" folders. These folders, if present, are part of organizational enhancements to maintain the archive folder structure. Users can continue to access their archived emails without any issues, and these folders don't affect the functionality or performance of the mailbox.

As previously described, Microsoft Purview moves items to subfolders (and names them using the naming convention described above) to help distribute content to an auxiliary archive. But moving items to subfolders might not always be the case. Sometimes an entire folder may be moved to an auxiliary archive. In this case, the folder will retain its original name. It won't be apparent in the folder list in Outlook that the folder was moved to an auxiliary archive.

Important considerations and limitations for auto-expanding archiving

Auto-expanded archiving is supported by new Outlook for Windows, classic Outlook for Windows, and Outlook on the web.

If a mailbox with an auto-expanding archive reaches a total of 1.5 TB of data stored in the auto-expanding archive, no more data will be stored in the auto-expanding archive or moved from the main archive to the auto-expanding archive.

> **Note**

The size that counts toward the 1.5 TB limit is the total of all items in the archive, including items in the Recoverable Items folder. An archive can therefore approach its limit even when the visible item size looks smaller. Additional storage is provisioned automatically and can take up to 30 days, and auto-expanding archiving supports a growth rate of no more than 1 GB per day.

Should the main archive also reaches its quota, no more data will be stored there, which can result in the primary mailbox also reaching its quota. Then, processing can stop for sending or receiving email and other items. To reduce this risk and minimize liability for your organization, we recommend using Microsoft 365 retention policies to delete content that no longer has business value. Configure these policies to delete items after a specified period. If you're not yet familiar with retention policies that can automatically delete email and other Microsoft 365 items, see Learn about retention policies and retention labels.

Auto-expanding archiving is supported only for mailboxes used for individual users (or shared mailboxes) with a growth rate that doesn't exceed 1 GB per day. A user's archive mailbox is intended for just that user. Using journaling, transport rules, or auto-forwarding rules to copy messages to an archive mailbox isn't permitted. Microsoft reserves the right to deny additional archiving in instances where a user's archive mailbox is used to store archive data for other users or in other cases of the inappropriate use.

Here are some things to consider before you enable a mailbox for auto-expanding archives:

- Users can access any folder in their archive mailbox, including ones that were moved to the auto-expanded storage area.
- If an archive mailbox has at least one auto-expanded storage area, users can't delete a folder from the archive mailbox or from the auxiliary archive. In other words, after an auto-expanded storage area has been provisioned, they can't delete any folders in the archive.
- Users can delete items in an auto-expanded storage area. However, they can't use the Recover Deleted Items feature to recover an item after auto-expanding archiving is enabled for their mailbox.
- Users can search for items in an auto-expanded archive with the following limitations:
  - For classic Outlook on Windows, search is restricted to the current Outlook search scope. For example, search from the primary mailbox won’t return items from auto-expanded archives.
  - Not supported in a cloud-only archive situation (primary mailbox still on-premises).

- Item counts in Outlook and Read/Unread counts (in Outlook and Outlook on the web) in an auto-expanded archive might not be accurate.

Auto-expanding archiving and other compliance features

This section explains the functionality between auto-expanding archiving and other compliance features.

- **eDiscovery:** When you use an eDiscovery tool, such as Content Search, the additional storage areas in an auto-expanded archive are also searched.
- **Retention:** When you put a mailbox on hold by using methods such as Microsoft 365 retention, eDiscovery case holds, and Litigation Hold in Exchange Online, content located in an auto-expanded archive is also placed on hold.
- **Messaging records management (MRM):** If you use MRM deletion policies in Exchange Online to permanently delete expired mailbox items, expired items located in the auto-expanded archive will also be deleted.

Next steps

For more technical details about auto-expanding archiving, see Microsoft 365: Auto-Expanding Archives FAQ.

If you're ready to enable auto-expanding archiving, see Enable auto-expanding archiving.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 10/14/2026

Enable auto-expanding archiving

> *Microsoft Purview service description*

You can use the Exchange Online auto-expanding archiving feature to enable additional storage space for archive mailboxes. When auto-expanding archiving is turned on, additional storage space is automatically added to a user's archive mailbox until it reaches the storage limit of 1.5 TB. You can turn on auto-expanding archiving for everyone in your organization or just for specific users. For more information about auto-expanding archiving, see Learn about auto-expanding archiving.

ⓘ Important

Microsoft is increasing the maximum auto-expanding archive capacity for eligible Exchange Online mailboxes from 1.5 TB to 3 TB. Rollout begins in mid-October 2026 and is expected to complete by late October.

The increase applies to mailboxes assigned the following eligible licenses:

- Office 365 E5, A5, or G5
- Microsoft 365 E5, A5, or G5
- Microsoft 365 Purview Suite
- Microsoft 365 Purview Suite FLW, EDU, or GOV
- Microsoft Defender + Purview Suite FLW
- Microsoft 365 F5 Compliance
- Microsoft 365 F5 Security + Compliance
- The combination of the eDiscovery & Audit, Insider Risk Management, and Information Protection & Governance mini suites

Qualifying offerings that include Microsoft Purview E5 capabilities, such as E7, also support the increase. **For all other licenses, the existing 1.5 TB limit remains unchanged.**

For E5 licenses, verify that **Records Management** is selected and remains enabled in the user's license settings.

If auto-expanding archiving is already enabled, no additional configuration or per-GB charge applies. Existing archive content, retention policies, and holds are unchanged. Verify license

eligibility and monitor archive growth.

Before you enable auto-expanding archiving

- Understand the following restrictions:
  - After you turn on auto-expanding archiving for your organization or for a specific user, it can't be turned off.
  - Administrators can't adjust the storage quota for auto-expanding archiving.
  - Auto-expanding archiving prevents you from recovering or restoring an inactive mailbox. This means if you enable auto-expanding archiving for a mailbox, and the mailbox becomes inactive in the future, you won't be able to recover the inactive mailbox or restore it.

    After you enable auto-expanding archiving for an inactive mailbox, the only way to recover data is by using the content search tool in the Microsoft Purview portal to export the data from the mailbox and import to another mailbox. For more information, see the Inactive mailboxes and auto-expanding archives.
- To enable auto-expanding archiving, you must be a global administrator in your organization or a member of the Organization Management role group in your Exchange Online organization. Alternately, to enable auto-expanding archiving for specific users, you must be a member of a role group that includes the Mail Recipients role.
- A user's mailbox must already be enabled for archive before you can enable auto-expanding archiving.
- After you turn on auto-expanding archiving, an archive mailbox is converted to an auto-expanding archive when the archive mailbox (including the Recoverable Items folder) reaches 90 GB. It can take up to 30 days for the additional storage space to be provisioned.
- Auto-expanding archiving also supports shared mailboxes.
- You can't use the Exchange admin center or the Microsoft Purview portal to enable auto-expanding archiving. You must use Exchange Online PowerShell.

Enable auto-expanding archiving for your entire organization

You can enable auto-expanding archiving for your entire organization. After you turn it on, auto-expanding archiving will be enabled for existing user mailboxes and for new user mailboxes that are created. When you create user mailboxes, be sure to enable the user's main archive mailbox so the auto-expanding archiving feature works for the new user mailbox.

1. Connect to Exchange Online PowerShell

2. To enable auto-expanding archiving for your entire organization, run the following command:

```
PowerShell
Set-OrganizationConfig -AutoExpandingArchive
```

Enable auto-expanding archiving for specific users

Instead of enabling auto-expanding archiving for every user in your organization, you can enable it only for specific users. You might decide on this strategy because only some users have a need for a large archive storage capacity.

When you enable auto-expanding archiving for a specific user and the user's mailbox in on hold or assigned to a retention policy, the following two configurations changes are made:

- The storage quota for the user's primary archive mailbox is increased by 10 GB (from 100 GB to 110 GB). The archive warning quota is also increased by 10 GB (from 90 GB to 100 GB).

- The storage quota for the Recoverable Items folder in the user's primary mailbox is increased by 10 GB (also from 100 GB to 110 GB). The Recoverable Items warning quota is also increased by 10 GB (from 90 GB to 100 GB). These changes are applicable only if the mailbox in on hold or assigned to a retention policy.

This additional space is added to prevent any storage issues that may occur before the auto-expanding archive is provisioned. Additional storage space isn't added when you enable auto-expanding archiving for your entire organization, as described in the previous section.

1. Connect to Exchange Online PowerShell

2. To enable auto-expanding archiving for a specific user, run the following command. As previously explained, the user's archive mailbox (main archive) must be enabled before you can turn on auto-expanding archiving for that user.

PowerShell

```
Enable-Mailbox <user mailbox> -AutoExpandingArchive
```

> **Important**
>
> In an Exchange hybrid deployment, you can't use the **Enable-Mailbox -AutoExpandingArchive** command to enable auto-expanding archiving for a specific user whose primary mailbox is on-premises and whose archive mailbox is cloud-based. To enable auto-expanding archiving for cloud-based archive mailboxes in an Exchange hybrid deployment, you have to run the **Set-OrganizationConfig -AutoExpandingArchive** command in Exchange Online PowerShell to enable auto-expanding archiving for the entire organization. If a user's primary and archive mailboxes are both cloud-based, then you can use the **Enable-Mailbox -AutoExpandingArchive** command to enable auto-expanding archiving for that specific user.

Verify that auto-expanding archiving is enabled

To verify that auto-expanding archiving is enabled for your organization, run the following command in Exchange Online PowerShell.

PowerShell

```
Get-OrganizationConfig | FL AutoExpandingArchiveEnabled
```

A value of `True` indicates that auto-expanding archiving is enabled for the organization.

To verify that auto-expanding archiving is enabled for a specific user, run the following command in Exchange Online PowerShell.

PowerShell

```
Get-Mailbox <user mailbox> | FL AutoExpandingArchiveEnabled
```

A value of `True` indicates that auto-expanding archiving is enabled for the user.

To determine if auto-expanding archiving is enabled for inactive mailboxes, run the following command in Exchange Online PowerShell.

PowerShell

```
Get-Mailbox -InactiveMailboxOnly | FL UserPrincipalName,AutoExpandingArchiveEnabled
```

A value of `True` indicates that auto-expanding archiving is enabled for the inactive mailbox. A value of `False` indicates that auto-expanding archiving isn't enabled.

Keep the following things in mind after you enable auto-expanding archiving:

- If you run the **Set-OrganizationConfig -AutoExpandingArchive** command to enable auto-expanding archiving for your organization, you don't have to run the **Enable-Mailbox -AutoExpandingArchive** on individual mailboxes. Running the **Set-OrganizationConfig** cmdlet to enable auto-expanding archiving for your organization doesn't change the *AutoExpandingArchiveEnabled* property on user mailboxes to `True`.

- Similarly, the values for the *ArchiveQuota* and *ArchiveWarningQuota* mailbox properties aren't changed when you enable auto-expanding archiving. In fact, when you enable auto-expanding archiving for a user mailbox and the *AutoExpandingArchiveEnabled* property is set to `True`, the *ArchiveQuota* and *ArchiveWarningQuota* properties are ignored. Here's an example of these mailbox properties after auto-expanding archiving is enabled for a user's mailbox.

  ```
  PS C:\> Get-Mailbox sarad | FL Archive*Quota,AutoExpandingArchiveEnabled


  ArchiveQuota                : 100 GB (107,374,182,400 bytes)
  ArchiveWarningQuota         : 90 GB (96,636,764,160 bytes)
  AutoExpandingArchiveEnabled : True
  ```

More information

- You can also use PowerShell to enable archive mailboxes. For example, you can run the following command in Exchange Online PowerShell to enable archive mailboxes for all users whose archive mailbox isn't already enabled.

  PowerShell

  ```
  Get-Mailbox -Filter {ArchiveStatus -Eq "None" -AND RecipientTypeDetails -eq
  "UserMailbox"} | Enable-Mailbox -Archive
  ```

- Auto-expanding archiving is supported for cloud-based archive mailboxes in an Exchange hybrid deployment for users who have an on-premises primary mailbox. However, after auto-expanding archiving is enabled for a cloud-based archive mailbox, you can't off-board

that archive mailbox back to the on-premises Exchange organization. Auto-expanding archiving isn't supported for on-premises mailboxes in any version of Exchange Server.

- As previously explained, 10 GB is added to the storage quota of the user's primary archive mailbox (and to the Recoverable Items folder if the mailbox is on hold) when you run the **Enable-Mailbox -AutoExpandingArchive** command. This increase provides additional storage until the auto-expanded storage space is provisioned (which can take up to 30 days). This additional storage space isn't added when you run the **Set-OrganizationConfig -AutoExpandingArchive** to enable auto-expanding archiving for all mailboxes in your organization. If you enabled auto-expanding archiving for the entire organization, but need to add the additional 10 GB of storage space for a specific user, you can run the **Enable-Mailbox -AutoExpandingArchive** command on that mailbox. You see an error saying that auto-expanding archiving is already enabled, but the additional storage space is added to the mailbox.

> **Important**
>
> Auto-expanding archiving is supported only for mailboxes used by individual users or for shared mailboxes with a growth rate that doesn't exceed 1 GB per day. Using journaling, transport rules, or auto-forwarding rules to copy messages to an archive mailbox for the purposes of archiving is not permitted. A user's archive mailbox is intended for just that user. Microsoft reserves the right to deny additional archiving in instances where a user's archive mailbox is used to store archive data for other users or in other cases of inappropriate use.

Next steps

Using the information from this article and Learn about auto-expanding archiving, make sure you prepare end users for the changes they can expect after enabling their mailbox for auto-expanding archives.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 10/14/2026

Customize an archive and deletion policy for mailboxes in your organization

To retain and delete emails, we recommend you use Microsoft 365 retention policies and retention labels rather than the older messaging records management (MRM) from Exchange Online. However, a valid reason to still use this older feature is to automatically move emails from a user's primary mailbox to their archive mailbox. You might also need to use MRM to apply retention and deletion settings to specific folders in the mailbox, rather than the entire mailbox.

Use this article as an example scenario to better understand these two valid reasons to use MRM retention policies and retention tags. For all other retention and deletion scenarios, use Microsoft 365 retention policies and retention labels.

The configuration steps in our example scenario require you to create an MRM retention policy that you then assign to mailboxes. This policy moves items to a user's archive mailbox after a specified period of time and also deletes items from the Deleted items folder after they reach a specific age limit.

> **Note**
>
> Instructions in this article use the Microsoft Purview portal and the Exchange admin center.

The actual rules that determine what items are moved or deleted and when that happens are called **retention tags**. Retention tags are linked to an **MRM retention policy**, that in turn is assigned to a user's mailbox. A retention tag applies retention settings to individual messages and folders in a user's mailbox. It defines how long a message remains in the mailbox and what action is taken when the message reaches the specified retention age. When a message reaches its retention age, it's either moved to the user's archive mailbox or it's deleted.

Example scenario for the instructions

The steps in this article set up an archiving and retention policy for a fictitious organization named Alpine House. This example deployment requires the following tasks:

- Step 1: Enable an archive mailbox for every user in the organization. This procedure gives users more mailbox storage, and is required so that a retention policy can automatically

move items to the archive mailbox. A user can also manually move items to their archive mailbox for archival storage.

- Step 2: Create two custom retention tags to do the following actions:
  - Automatically move items that are 3 years old to the user's archive mailbox. Moving items to the archive mailbox frees up space in a user's primary mailbox.
  - Automatically delete items that are 5 years old from the Deleted Items folder. This also frees up space in the user's primary mailbox. Users will have the opportunity to recover these items if necessary. For more information, see the final bullet in the More information section.
- Step 3: Create a new retention policy and add the new custom retention tags to it. Additionally, you'll add a built-in retention tag that can't be achieved with a recommended Microsoft 365 retention label because it also moves items to the archive mailbox. It's a personal tag for archiving after 1 year that users can assign to items in their mailbox when they want a shorter archive period than their default of 3 years.
- Step 4: Replace the Default MRM policy with the new retention policy you've created.

Optionally, you can also run the Managed Folder Assistant to apply the new settings, and make the new retention policy the default for your organization.

You can follow some or all of the steps in this article to help understand how to set up your own archive and deletion policy for mailboxes in your organization. We recommend that you test this process on a few mailboxes before implementing it on all mailboxes in your organization.

Before you set up the archive and deletion policy

- To follow this example configuration, use an isolated, test environment that won't impact users on your production network. Then take your learnings to your production network.
- In your isolated test environment, consider using a global administrator account so you can perform all the steps for this example scenario. In a production environment, always use administrative roles that grant least privilege.
- When you create a new user account and assign the user an Exchange Online license, a mailbox is automatically created for the user. When the mailbox is created, it's automatically assigned a default retention policy, named Default MRM Policy. In this example deployment, you'll create a new MRM retention policy and then assign it to user mailboxes, replacing the

Default MRM policy. A mailbox can have only one MRM retention policy assigned to it at any one time.

- To learn more about retention tags and MRM retention policies in Exchange Online, see Retention tags and retention policies and the associated instructions. For example:
  - Create a retention policy for Exchange Online
  - Add retention tags to or remove retention tags from a retention policy in Exchange Online

Step 1: Enable archive mailboxes for users

The first step is to ensure each user in your organization has an archive mailbox. A user's archive mailbox must be enabled so that for our example scenario, a retention tag with a "Move to Archive" retention action can move the item after the retention age expires.

For instructions to enable archive mailboxes, see Enable archive mailboxes in Microsoft 365.

> **Note**
>
> You can enable archive mailboxes any time during this process, just as long as they're enabled at some point before you complete the process. If an archive mailbox isn't enabled, no action is taken on any items that have an archive or deletion policy assigned to it.

Step 2: Create new retention tags for the archive and deletion policies

In this step, you'll create the two custom retention tags that were previously described for our example scenario.

- A custom archive default policy tag, which we'll name "Alpine House 3 Year Move to Archive".
- A custom retention policy tag configured for the Deleted Items folder, which we'll name "Alpine House Deleted Items 5 Years Delete and Allow Recovery".

For this configuration, you use the Microsoft Purview portal.

1. Sign in to the Microsoft Purview portal > **Solutions** > **Data Lifecycle Management** > **Exchange (legacy)** > **MRM Retention tags**.

   A list of the retention tags for your organization is displayed.

Follow the step-by-steps instructions how to create the example retention tags for our scenario.

Create the custom archive default policy tag

First for our example scenario, you'll create a custom archive default policy tag (DPT) that will move items to the archive mailbox after 3 years.

1. On the **MRM Retention tags** page, select **+ New tag**, and then on the **Name your tag** page, type a name for the new retention tag, and an optional description that explains the purpose of the custom retention tag.

   For our example scenario, we'll name this tag "Alpine House 3 Year Move to Archive".

2. On the **Define how the tag will be applied** page, select **Automatically to entire mailbox (default)**.

3. On the **Define retention settings** page, complete the following fields:

   a. **When items reaches the following age (in days)** Enter the duration of the retention period. For this scenario, items will be moved to the archive mailbox after 1095 days (3 years).

   b. For the **Retention Action** Select **Move item to archive** to move items to the archive mailbox when the retention period expires.

4. Select **Next**, and then review and submit to create the custom archive DPT.

The new archive DPT is displayed in the list of retention tags.

Create the custom retention policy tag for the Deleted Items folder

The second retention tag to create for our example scenario is a custom retention policy tag (RPT) for the Deleted Items folder. This retention tag will delete items in the Deleted Items folder after 5 years, and provides a recovery period when users can use the Recover Deleted Items tool to recover an item.

1. Back on the **MRM Retention tags** page, select **+ New tag**, and on the **Name your tag** page, type a name for the new retention tag, and an optional description that explains the purpose of the custom retention tag.

For our example scenario, we'll name this tag "Alpine House Deleted Items 5 Years Delete and Allow Recovery".

2. On the **Define how the tag will be applied** page, select **Automatically to default folder** and then for the **Apply tag to this folder** option, select **Deleted items**.

3. On the **Define retention settings** page, complete the following fields:

   a. **When items reaches the following age (in days)** Enter the duration of the retention period. For this scenario, items will be deleted after 1825 days (5 years).

   b. For the **Retention Action** Select **Delete and allow recovery** to delete items when the retention period expires, but allow users to recover a deleted item within the deleted item retention period (which by default is 14 days).

4. Select **Next**, and then review and submit to create the custom deletion DPT.

The new RPT is displayed in the list of retention tags.

Step 3: Create a new retention policy

After you create the custom retention tags, the next step for our example scenario is to create a new retention policy and add the retention tags. You'll add the two custom retention tags that you created in Step 2, and a built-in tag that's preconfigured for you. In Step 4, you'll assign this new retention policy to user mailboxes.

1. Sign in to the Microsoft Purview portal > **Solutions** > **Data Lifecycle Management** card > **Exchange (legacy)** > **MRM Retention policies**.

2. On the **MRM Retention policies** page, select **New policy**.

3. In the **Name** box, type a name for the new retention policy; for example, **Alpine House Archive and Deletion Policy**.

4. Select + **Add tag**.

   A list of the retention tags in your organization is displayed, which includes the custom tags that you created in Step 2.

5. Add the 3 retention tags that are described in more detail in the More information section:

   - **Alpine House 3 Year Move to Archive** - the custom archive default policy tag created in step 2 of these instructions

- **Alpine House Deleted Items 5 Years Delete and Allow Recovery** - the custom tag for the Deleted Items folder created in step 2 of these instructions
- **Personal 1 year move to archive** - a built-in tag that's preconfigured

To add these retention tags, select them, and then select **Add**.

6. Back on the **Configure your policy** page, select **Next** to review and submit the new policy.

The new retention policy is displayed in the list. Select it to display the retention tags linked to it in the details pane.

Step 4: Assign the new retention policy to user mailboxes

When a new mailbox is created, a retention policy named Default MRM policy is assigned to it by default. In this step for our example scenario, you'll replace this retention policy by assigning the new retention policy that you created in Step 3 to the user mailboxes in your organization.

Replacement is required because a mailbox can have only one MRM retention policy assigned to it at a time. This step assumes that you'll assign the new policy to all mailboxes in your organization.

To follow these steps, make sure you use the new Exchange admin center, rather than the classic version.

1. Sign in to the Exchange admin center (EAC), and go to **Recipients** > **Mailboxes**.

   A list of all user mailboxes in your organization is displayed.

2. Select all the mailboxes by selecting the box for **Display name**.

3. Select the **Mailbox policies** option.

4. In the **Mailbox policies** flyout pane, under **Retention Policy**, select the retention policy that you created in Step 3; for example, **Alpine House Archive and Retention Policy**.

5. Select **Save** to save the new retention policy assignment.

6. To verify that the new retention policy was assigned to mailboxes:

   a. Select a mailbox on the **Mailboxes** page.

   b. On the mailbox properties page for the selected user, select **Mailbox**.

The name of the new policy assigned to the mailbox is displayed for the **Retention policy**.

(Optional) Step 5: Run the Managed Folder Assistant to apply the new settings

After you apply the new retention policy to mailboxes in Step 4, it can take up to 7 days in Exchange Online for the new retention settings to be applied to the mailboxes. This is because a process called the *Managed Folder Assistant* processes mailboxes at least once every 7 days. Instead of waiting for the Managed Folder Assistant to run, you can force this to happen by running the **Start-ManagedFolderAssistant** cmdlet in Exchange Online PowerShell.

What happens when you run the Managed Folder Assistant? It applies the settings in the retention policy by inspecting items in the mailbox and determining whether they're subject to retention. It then stamps items subject to retention with the appropriate retention tag, and then takes the specified retention action on items past their retention age.

Here are the steps to connect to Exchange Online PowerShell, and then run the Managed Folder Assistant on every mailbox in your organization.

1. Connect to Exchange Online PowerShell.
2. Run the following two commands to start the Managed Folder Assistant for all user mailboxes in your organization.

PowerShell

```
$Mailboxes = Get-Mailbox -ResultSize Unlimited -Filter {RecipientTypeDetails -eq "UserMailbox"}
```

PowerShell

```
$Mailboxes.Identity | Start-ManagedFolderAssistant
```

That's it! You've set up an archive and deletion policy for the Alpine House organization.

More information about the Managed Folder Assistant

As previously stated, the Managed Folder Assistant processes mailboxes at least once every 7 days. So it's possible that a mailbox can be processed by the Managed Folder Assistant more frequently. Also, admins can't predict the next time a mailbox is processed by the Managed Folder Assistant, which is one reason why you might want to run it manually.

However, if you want to temporarily prevent the Managed Folder Assistant from applying the new retention settings to a mailbox, you can run the `Set-Mailbox -ElcProcessingDisabled $true` command to temporarily disable the Managed Folder Assistant from processing a mailbox.

To re-enable the Managed Folder Assistant for a mailbox, run the `Set-Mailbox -ElcProcessingDisabled $false` command.

Finally, if a mailbox user has a disabled account, items aren't moved to the archive mailbox for that mailbox.

(Optional) Step 6: Make the new retention policy the default for your organization

In Step 4, you have to assign the new retention policy to existing mailboxes. But you can configure Exchange Online so that the new retention policy is assigned to new mailboxes that are created in the future.

You do this by using Exchange Online PowerShell to update your organization's default mailbox plan. A *mailbox plan* is a template that automatically configures properties on new mailboxes. In this optional step, you can replace the current retention policy that's assigned to the mailbox plan (by default, the Default MRM Policy) with the MRM retention policy that you created in Step 3. After you update the mailbox plan, the new MRM retention policy will be assigned to new mailboxes.

1. Connect to Exchange Online PowerShell.

2. Run the following command to display information about the mailbox plans in your organization.

   PowerShell

   ```
   Get-MailboxPlan | Format-Table DisplayName,RetentionPolicy,IsDefault
   ```

   Note the mailbox plan that's set as the default.

3. Run the following command to assign the new MRM retention policy that you created in Step 3 (for example, **Alpine House Archive and Retention Policy**) to the default mailbox plan. This example assumes the name of the default mailbox plan is **ExchangeOnlineEnterprise**.

   PowerShell

```
Set-MailboxPlan "ExchangeOnlineEnterprise" -RetentionPolicy "Alpine House Archive and Retention Policy"
```

4. You can rerun the command in step 2 to verify that the MRM retention policy assigned to the default mailbox plan was changed.

More information

- The retention age of mailbox items is calculated from the date of delivery. Or from the date of creation for items such as draft messages that aren't sent but are created by the user. When the Managed Folder Assistant processes items in a mailbox, it stamps a start date and an expiration date for all items that have retention tags with the Delete and Allow Recovery or Permanently Delete retention action. Items that have an archive tag are stamped with a move date.

- The following table provides more information about each retention tag for the custom MRM retention policy in this article.

Expand table

Retention tag	What this tag does	Built-in or custom?	Type
Alpine House 3 Year Move to Archive	Moves items that are 1095 days (3 years) old to the archive mailbox.	Custom (See Step 2: Create new retention tags for the archive and deletion policies)	Default Policy Tag (archive); this tag is automatically applied to the entire mailbox.
Alpine House Deleted Items 5 Years Delete and Allow Recovery	Deletes items from the Deleted Items folder that are 5 years old. Users can recover these items for up 14 days after they're deleted. See the next list entry for more information.	Custom (See Step 2: Create new retention tags for the archive and deletion policies)	Retention Policy Tag (Deleted Items); this tag is automatically applied to items in the Deleted items folder.
Personal 1 year move to archive	Moves items to the archive mailbox after 1 year.	Built-in	Personal; this tag can be applied by users.

- Users can use the Recover Deleted Items tool in Outlook and Outlook on the web to recover a deleted item within the deleted item retention period, which by default is 14 days in Exchange Online. An administrator can use Exchange Online PowerShell to increase the deleted item retention period to a maximum of 30 days. For more information, see: Recover

deleted items in Outlook for Windows and Change the deleted item retention period for a mailbox in Exchange Online.

Last updated on 12/03/2024

Learn about inactive mailboxes

> *Microsoft Purview service description*

Your organization might need to retain former employees' email after they leave the organization. Depending on your organization's retention requirements, you might need to retain mailbox content for a few months or years after employment ends, or you might need to retain mailbox content indefinitely. Regardless of how long you need to retain email, you can create inactive mailboxes to retain the mailbox of former employees.

What are inactive mailboxes?

An inactive mailbox is a type of soft-deleted mailbox used to retain a former employee's email after they leave your organization.

When an employee leaves your organization (or goes on an extended leave of absence), you can remove their Microsoft 365 account. The employee's mailbox data is retained for 30 days after the account is removed. During this period, you can still recover the mailbox data by undeleting the account. After 30 days, the data is permanently removed.

But if a hold is applied to the mailbox prior to deleting the Microsoft 365 account, the mailbox will be converted into an inactive mailbox. The following sections contain information about holds that can be applied with Microsoft 365 retention and eDiscovery holds.

Inactive mailboxes are useful when your organization needs to retain mailbox content of former employees for regulatory or other reasons. While any type of hold listed in this document will force a mailbox to be made inactive when a user object is deleted, we recommend doing this by applying a Microsoft 365 retention policy or retention labels, confirm the hold is applied, and then remove the corresponding Microsoft 365 account. At that point, the contents of the inactive mailbox are retained for the duration of the retention period specified before the user account was deleted. You can still recover the corresponding user account for a 30-day period, however after 30 days, the mailbox is retained in Microsoft 365 as an inactive mailbox until the retention policy or retention labels are removed.

Important

As we mentioned previously, we recommend you use Microsoft 365 retention to create an inactive mailbox:

- In-Place Holds in the Exchange admin center are now retired. As of July 1, 2020, new In-Place Holds were unable to be created in Exchange Online. As of October 1, 2020, the hold duration of in-place holds could no longer be changed. Any inactive mailbox that has an In-Place Hold applied can only be deleted by removing the In-Place Hold. Existing inactive mailboxes that are on In-Place Hold will continue to be preserved until the hold is removed. For more information about In-Place Holds retirement, see Retirement of legacy eDiscovery tools.
- Litigation hold remains supported as an alternative method to retain content in a mailbox and make it inactive after a user account is deleted. However, as an older technology, we recommend you use Microsoft 365 retention instead.

When there are multiple holds on the same content, the principle of retention apply and the content is retained for the longest period.

Confirming a hold is applied to a mailbox

Whether you apply a Microsoft 365 retention policy, retention labels, eDiscovery hold, Litigation hold, or have an existing In-Place Hold, you can confirm the hold is successfully applied to the mailbox by using PowerShell. If you've recently configured the hold, you might need to wait until it's applied to the mailbox.

To prevent accidental or unintentional deletion, we recommend you confirm the hold before you delete the user account. If the hold isn't applied, the mailbox won't be converted into an inactive mailbox.

For instructions, see How to identify the type of hold placed on an Exchange Online mailbox.

A hold can also be applied automatically: after you remove a hold from a mailbox, Microsoft 365 applies a *delay hold* so that content isn't purged immediately. A mailbox with a delay hold is still treated as being on hold, so it stays an inactive mailbox until the delay hold is removed or expires. For more information, see Managing mailboxes on delay hold and Remove a delay hold from an inactive mailbox.

Inactive mailboxes and Microsoft 365 retention

If a Microsoft 365 retention policy is applied to a mailbox, or one or more email items in a mailbox have a retention label applied, and then the Microsoft 365 user account is deleted, the mailbox will be converted into an inactive mailbox. For the inactive mailbox to be created:

- The retention settings must be configured to retain content, or retain and then delete content. If the retention action is configured to only delete content, the mailbox won't become inactive when the user account is deleted. For inactive mailboxes, we recommend using the retain and then delete option.

- The retention settings must be applied to a retention location that is associated with an Exchange mailbox:
  - Exchange email
  - Microsoft 365 Groups
  - Skype for Business
  - Exchange public folders
  - Teams channel messages
  - Teams chats and Copilot interactions
  - Teams private channel messages
  - Viva Engage community messages
  - Viva Engage user messages

For more information about Microsoft retention, see Learn about retention policies and retention labels.

If Microsoft 365 retention is used to create an inactive mailbox, the retention settings from the retention policy or retention labels continue to apply to the inactive mailbox. This means if the retention settings are configured to retain and then delete content, items will be moved to the Recoverable Items folder when the retention duration expires, and then eventually purged from the inactive mailbox. If the retention settings aren't configured to delete items, then items that haven't been permanently deleted by the user (before the mailbox was made inactive) will not be moved to the Recoverable Items folder and will be retained indefinitely after the mailbox becomes inactive.

> **Tip**
>
> You can use the *ComplianceTagHoldApplied* property to identify whether a mailbox has items that have one or more retention labels applied to retain, or retain and then delete content.

For more information, see Identifying mailboxes on hold because a retention label has been applied to a folder or item.

Using adaptive policy scopes to manage retention of inactive mailboxes

With adaptive policy scopes, you can apply retention settings specifically to inactive mailboxes. Benefits of this configuration include:

- You can meet your organization's regulations or policies that require different retention periods for active employees and former employees.
- You can configure the retention settings to retain mailbox content for only as long as necessary to meet your organization's requirement for former employees.
- You can quickly identify the policy for retention that's assigned to inactive mailboxes in your organization, which makes it easier to change the retention settings if necessary.
- It's easier to permanently delete an inactive mailbox because you can remove it from the policy by configuring the adaptive scope to exclude it, based on attributes or properties of the inactive mailbox. Otherwise, you must use Exchange Online PowerShell before deleting the mailbox.

> **Note**
>
> Depending on the configuration of your adaptive policy scope, inactive mailboxes might or might not be included. To specifically target or to exclude inactive mailboxes from an adaptive policy scope, see configuration information for Exchange email and Exchange public folders.

Using static policy scopes and inactive mailboxes

If you don't use adaptive policy scopes and instead use a static scope with Microsoft 365 retention, consider the following:

- Static policy scopes include inactive mailboxes when you use the default **All mailboxes** configuration but aren't supported for specific inclusions or exclusions. However, if you include or exclude a recipient that has an active mailbox at the time the policy is applied

and the mailbox later goes inactive, the retention settings continue to be applied or excluded. In this scenario, specific inclusion and exclusion limits still apply.

> **Note**
>
> This also means that any new Microsoft 365 retention settings using a static scope that is applied to the default selection of **All mailboxes** will automatically include all existing inactive mailboxes.

- If you change the default selection of **All mailboxes** to include specific recipients, retention settings for the policy will no longer apply to any inactive mailboxes, which now become eligible for automatic deletion.
- If you want to release a retention policy that's applied to an inactive mailbox, see Releasing a policy for retention.

> **Caution**
>
> When you use Microsoft 365 retention to make a mailbox inactive, do not change or remove the user principal name (UPN) for the mailbox before you delete the corresponding user account. Additionally, do not change the primary SMTP address (that's derived from the UPN) or remove this email address from the list of secondary SMTP addresses associated with the mailbox before making the mailbox inactive.
>
> If you change the UPN or email addresses (that were assigned to the mailbox at the time the retention settings were applied) and then delete the user account to make the mailbox inactive, you won't be able to delete the inactive mailbox if you no longer need to retain it. That's because you can't remove the inactive mailbox from the policy by using a UPN or email address (to identify the inactive mailbox) that's different from the ones that existed when the retention settings were initially applied to the mailbox. For more information about deleting inactive mailboxes, see **Delete an inactive mailbox in Office 365**.

Inactive mailboxes and eDiscovery case holds

If a hold that's associated with an eDiscovery case is placed on a mailbox and then the mailbox or the user's account is deleted, the mailbox becomes an inactive mailbox. However, don't use eDiscovery case holds to make a mailbox inactive. That's because eDiscovery cases are intended for specific, time-bound cases related to a legal issue. At some point, a legal case probably ends,

the holds associated with the case are removed, and the eDiscovery case is closed. In fact, if a hold that's placed on an inactive mailbox is associated with an eDiscovery case, and then the hold is released or the eDiscovery case is closed (or deleted), the inactive mailbox is permanently deleted. Also, you can't create a time-based eDiscovery hold. This condition means that content in an inactive mailbox is retained forever or until the hold is removed and the inactive mailbox is deleted. Therefore, use Microsoft 365 retention for inactive mailboxes.

> **Important**
>
> When a mailbox with an **active eDiscovery hold** is deleted, the mailbox content is preserved as an inactive mailbox. The content remains searchable and exportable in eDiscovery for as long as the hold or associated case remains active.

For more information about the differences between eDiscovery holds and Microsoft 365 retention, see When to use retention policies and retention labels or eDiscovery holds.

Inactive mailboxes and auto-expanding archives

An inactive mailbox that's configured with an auto-expanding archive can't be recovered or restored. If, for compliance reasons, you need to recover data from an inactive mailbox with an auto-expanding archive, use content search to export the data from the mailbox. This action is supported for eDiscovery purposes only, and can't be used as a backup solution. For instructions to use search for the recovery of data for eDiscovery, see Create a search for a case in eDiscovery

Inactive mailboxes and Exchange MRM retention policies

Applying an Exchange retention policy (the messaging records management, or MRM, feature in Exchange Online) does not create an inactive mailbox when the user account is deleted.

However, if this MRM retention policy was applied to a mailbox before it went inactive, any deletion policies (MRM retention tags configured with a **Delete** action) will continue to be processed on the inactive mailbox. That means items that are tagged with an MRM deletion policy will be moved to the Recoverable Items folder when the retention period expires. Those items are purged from the inactive mailbox when the hold duration expires. If a hold duration isn't specified for the inactive mailbox, items in the Recover Items folder will be retained indefinitely.

Conversely, any archive policies (MRM retention tags configured with a **MoveToArchive** action) that are included in the MRM retention policy assigned to an inactive mailbox are ignored. That means items in an inactive mailbox that are tagged with an archive policy remain in the primary mailbox when the retention period expires. They're not moved to the archive mailbox or to the Recoverable Items folder in the archive mailbox. They will be retained indefinitely.

Next steps

To make a mailbox inactive and manage it, such as recovering, restoring, and deleting it, see Create and manage inactive mailboxes.

Last updated on 08/05/2026

Create and manage inactive mailboxes

> *Microsoft Purview service description*

Inactive mailboxes let you retain former employees' email after they leave your organization and can be accessed by authorized people who have been granted eDiscovery permissions for compliance or legal reasons. For example, administrators, compliance officers, and records managers who can then use Content Search to search and export the contents of an inactive mailbox. Inactive mailboxes can't receive email and aren't displayed in your organization's shared address book or other lists.

For more information about inactive mailboxes, see Learn about inactive mailboxes.

Create an inactive mailbox

Making a mailbox inactive requires a hold on the mailbox, and then deleting the mailbox or corresponding user account.

To make a mailbox inactive, it must be licensed correctly so that a hold can be applied to the mailbox before it's deleted. After the user account is deleted, any Exchange Online license associated with the user account will be available to assign to a new user.

We recommend you use Microsoft 365 retention to apply the hold on the mailbox. Other methods are covered in Learn about inactive mailboxes.

The best way to delete a mailbox is to delete the corresponding user account in the Microsoft 365 admin center. For information about deleting user accounts, see Delete a user from your organization. However, you can also delete the mailbox by using the **Remove-Mailbox** cmdlet in Exchange Online PowerShell. For more information, see Delete or restore user mailboxes in Exchange Online.

The following table summarizes the process of making an inactive mailbox for different retention scenarios.

Expand table

To...	Do this...	Result
Retain mailbox content indefinitely after an employee leaves the organization	1. Apply Microsoft 365 retention settings with retain actions for the mailbox (a retention policy) or specific email items (one or more retention labels). 2. Wait for the retention settings to be applied. 3. Remove the user's Microsoft 365 account.	All content in the inactive mailbox that has retention settings applied, including items in the Recoverable Items folder, is retained indefinitely.
Retain all mailbox content for a specific period after an employee leaves the organization and then delete the mailbox	1. Apply a Microsoft 365 retention policy to the mailbox with retention settings that retain and then delete items when the retention period expires. 2. Wait for the retention settings to be applied. 3. Remove the user's Microsoft 365 account.	When the retention period for a mailbox item expires, the item is moved to the Recoverable Items folder and then it's permanently deleted (purged) from the inactive mailbox when the deleted item retention period (for Exchange mailboxes) expires. The retention period of the Microsoft 365 retention policy is always based on the original date a mailbox item was received or created.

> **Note**
>
> If Microsoft 365 retention settings that are configured to retain, or retain and then delete content, is already applied to the mailbox or mailbox items, or a Litigation hold is already placed on a mailbox, or then all you have to do to create an inactive mailbox is delete the corresponding user account.

View a list of inactive mailboxes

You can use the Microsoft Purview portal to view a list of the inactive mailboxes in your organization.

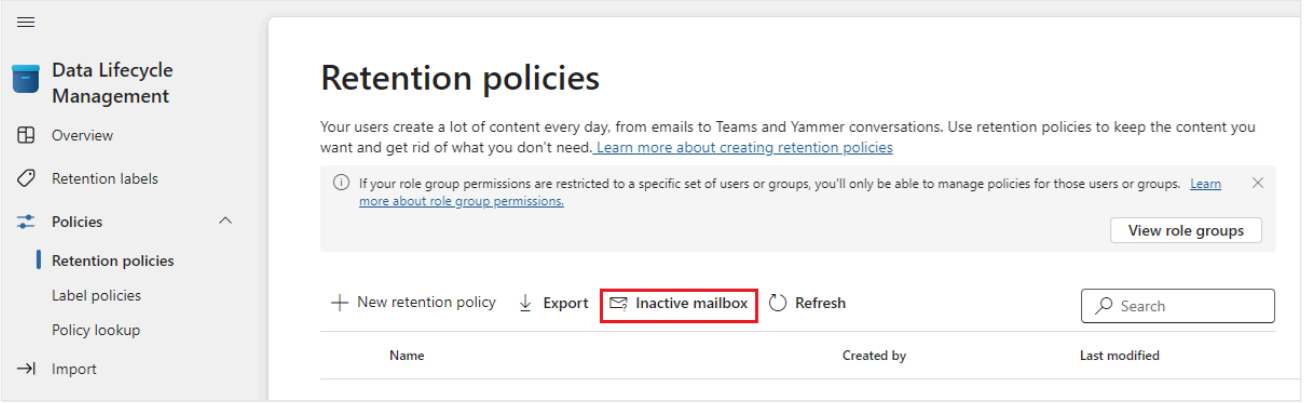
> **Note**

The portals are limited to displaying up to 5,000 inactive mailboxes. To view more than 5,000 inactive mailboxes, you must use Exchange Online PowerShell as described after the step instructions.

For this configuration, use the Microsoft Purview portal and a Compliance administrator account for your organization.

1. Sign in to the Microsoft Purview portal > **Solutions** > **Data Lifecycle Management** > **Policies** > **Retention policies**.

2. Select the **Inactive mailbox** option. For example, from the Microsoft Purview portal:



3. The **Inactive mailboxes** page displays a list of inactive mailboxes. Select one to see details about that inactive mailbox. Details include the Exchange identifier for the mailbox and whether it's on Litigation Hold.

   You won't see other types of holds in the details pane, for example a Microsoft 365 retention policy or eDiscovery hold. To find this information, see How to identify the type of hold placed on an Exchange Online mailbox.

If you have many inactive mailboxes, you might find it easier to search and sort a CSV file for the details you see in the list: On the **Inactive mailboxes** page, select **Export**.

Alternatively, you can run the following command in Exchange Online PowerShell to display the list of inactive mailboxes:

PowerShell

```
 Get-Mailbox -InactiveMailboxOnly -ResultSize Unlimited | FT
DisplayName,PrimarySMTPAddress,WhenSoftDeleted
```

You can also run the following command to export the list of inactive mailboxes and other information to a CSV file. In this example, the CSV file is created in the current directory.

PowerShell

```
Get-Mailbox -InactiveMailboxOnly -ResultSize Unlimited | Select
Displayname,PrimarySMTPAddress,DistinguishedName,ExchangeGuid,WhenSoftDeleted |
Export-Csv InactiveMailboxes.csv -NoTypeInformation
```

> **Note**
>
> It's possible that an inactive mailbox might have the same SMTP address as an active user mailbox. In this case, the value of the **DistinguishedName** or **ExchangeGuid** property can be used to uniquely identify an inactive mailbox.

Search and export the contents of an inactive mailbox

You can access the contents of the inactive mailbox by using the Content Search tool in the Microsoft Purview portal. For more information, see the Feature reference for Content search article.

Change the hold duration for an inactive mailbox

After a mailbox is made inactive, you might be able to change the duration of the hold applied to the inactive mailbox.

For step-by-step procedures, see Change the hold duration for an inactive mailbox.

Recover an inactive mailbox

If a former employee returns to your organization, or if a new employee is hired to take on the job responsibilities of the former employee, you can recover the contents of the inactive mailbox.

When you recover an inactive mailbox, the mailbox is converted to a new mailbox, the contents and folder structure of the inactive mailbox are retained, and the mailbox is linked to a new user account. After it's recovered, the inactive mailbox no longer exists.

For step-by-step procedures and more information about happens when you recover an inactive mailbox, see Recover an inactive mailbox.

Restore the contents of an inactive mailbox to another mailbox

If another employee takes on the job responsibilities of a former employee, or if another person needs access to the contents of the inactive mailbox, you can restore (or merge) the contents of the inactive mailbox to an existing mailbox.

When you restore an inactive mailbox, the contents are copied to another mailbox. The inactive mailbox is retained and remains an inactive mailbox. The inactive mailbox can still be searched using eDiscovery, its contents can be restored to another mailbox, or it can be recovered or deleted at a later date.

For step-by-step procedures, see Restore an inactive mailbox.

Delete an inactive mailbox

If you no longer need to retain the contents of an inactive mailbox, you can permanently delete the inactive mailbox by removing the hold applied to the inactive mailbox.

For more details and step-by-step procedures, see Delete an inactive mailbox.

> **Note**
>
> If a mailbox is still listed as inactive after you remove the last hold, check whether a delay hold is applied. A delay hold is applied automatically after another hold is removed, and it keeps the mailbox inactive until the delay hold is removed or expires. For instructions, see Remove a delay hold from an inactive mailbox.

Last updated on 08/05/2026

Change the hold duration for an inactive mailbox

Article • 03/26/2024

> *Microsoft 365 licensing guidance for security & compliance*.

An inactive mailbox is mailbox state that is used to retain a former employee's email after they leave your organization. A mailbox becomes inactive when an applicable hold is applied to it before the Microsoft 365 user object is deleted. The following types of holds initiate the creation of an inactive mailbox upon user account deletion:

- Microsoft 365 retention policies and labels with retain-only, or retain and delete settings.
- A hold associated with an eDiscovery case.
- Litigation Hold.
- An existing In-Place Hold.

ⓘ Important

While any of the above holds will force a mailbox to become inactive upon Microsoft 365 user account deletion, it is strongly recommended that you use Microsoft 365 retention when proactively planning to utilize inactive mailboxes.

- eDiscovery holds are intended for specific, time-bound cases related to a legal issue. At some point, a legal case will probably end and the holds associated with the case will be removed and the eDiscovery case will be closed (or deleted). If a hold that's placed on an inactive mailbox is associated with an eDiscovery case, and the hold is released or the eDiscovery case is closed or deleted, the inactive mailbox will be permanently deleted.
- In-Place Holds in the Exchange admin center are now retired. As of July 1, 2020, new In-Place Holds were unable to be created in Exchange Online. As of October 1, 2020, the hold duration of in-place holds could no longer be changed. Any inactive mailbox that has an In-Place Hold applied can only be deleted by removing the In-Place Hold. Existing inactive mailboxes that are on In-Place Hold will continue to be preserved until the hold is removed. For

more information about In-Place Holds retirement, see [Retirement of legacy eDiscovery tools](#).

- [Litigation hold](#) remains supported as an alternative method to retain content in a mailbox and make it inactive after a user account is deleted. However, as an older technology, we recommend you use Microsoft 365 retention instead.

Once made inactive, the contents of the mailbox including the Recoverable Items folder are retained until the hold that was placed on the mailbox before it was made inactive no longer applies.

If the applicable hold is not time-based, such as a hold associated with an indefinite retain-only Microsoft 365 retention policy or label, an eDiscovery case or Litigation Hold (without a `LitigationHoldDuration` configured), the mailbox content is retained indefinitely until the hold is removed.

However, if the hold is time-based, the mailbox content is retained until the hold duration expires. At this time, any items in the Recoverable Items folder are permanently deleted (purged) from the inactive mailbox.

> **Note**
>
> For inactive mailboxes, we recommend using a retain and delete setting for your Microsoft 365 retention policy or labels. If you choose a retain only setting, the Recoverable Items folder will purge at the end of the hold duration, however any other non-deleted items will remain within the inactive mailbox indefinitely.

As regulations and policies evolve, there may be some situations in which you need to change the duration of the hold assigned to the inactive mailbox. The following steps outline how to do this.

> **Tip**
>
> If you're not an E5 customer, use the 90-day Microsoft Purview solutions trial to explore how additional Purview capabilities can help your organization manage data security and compliance needs. Start now at the [Microsoft Purview trials hub](#). Learn details about [signing up and trial terms](#).

Connect to PowerShell

As we mentioned before, many different types of holds can trigger the creation of an inactive mailbox. For this reason, in order to change the hold duration applied to the inactive mailbox, you must first identify what type of holds are affecting it. To identify the type of holds, you must use Exchange Online PowerShell to identify the types of holds. If the inactive mailbox is affected by Microsoft 365 retention policies or labels, you must also use Security & Compliance PowerShell to identify the specific policies.

- To connect to Exchange Online PowerShell or Security & Compliance PowerShell, see one of the following topics:
  - Connect to Exchange Online PowerShell
  - Connect to Security & Compliance PowerShell

Step 1: Identify the holds on an inactive mailbox

Because different types of holds or one or more Microsoft 365 retention policies might be placed on an inactive mailbox, the first step is to identify the holds on an inactive mailbox.

Run the following command in Exchange Online PowerShell to display the hold information for a specific inactive mailbox in your organization.

PowerShell

```
Get-Mailbox -Identity <identity of inactive mailbox> -InactiveMailboxOnly |
FL
DisplayName,Name,DistinguishedName,ExchangeGuid,IsInactiveMailbox,Litigation
HoldEnabled,LitigationHoldDuration,LitigationHoldDate,LitigationHoldOwner,In
PlaceHolds,ComplianceTagHoldApplied
```

If you need to identify the type of hold for multiple inactive mailboxes and your organization has a large number of them, it may become unmanageable to view using PowerShell. In this case, you can export all of the applicable information to a CSV file using the following command and modifying the `Path` as needed for your environment:

PowerShell

```
Get-Mailbox -InactiveMailboxOnly -ResultSize Unlimited | Select
DisplayName,Name,DistinguishedName,ExchangeGuid,IsInactiveMailbox,Litigation
HoldEnabled,LitigationHoldDuration,LitigationHoldDate,LitigationHoldOwner,In
PlaceHolds,ComplianceTagHoldApplied | Export-Csv -NoTypeInformation -Path
"C:\Temp\InactiveMailboxHoldTypes.csv"
```

For the purposes of this example, the following shows results for six inactive mailboxes with different possible hold types.

> ⓘ **Note**
>
> Multiple holds including multiple types of holds can apply to a single inactive mailbox.

text

```
DisplayName               : Ann Beebe
Name                      : annb
DistinguishedName         : CN=annb,OU=Soft Deleted
                            Objects,OU=contoso.onmicrosoft.com,OU=Microsoft
Exchange
                            Hosted
Organizations,DC=NAMPR06A007,DC=PROD,DC=OUTLOOK,DC=COM
ExchangeGuid              : 664ef44e-c1a0-47b0-9553-2ecdfc6ef840
IsInactiveMailbox         : True
LitigationHoldEnabled     : True
LitigationHoldDuration    : 365.00:00:00
LitigationHoldDate        : 10/25/2021 6:54:57 PM
LitigationHoldOwner       : admin@contoso.com
InPlaceHolds              : {}
ComplianceTagHoldApplied  : False
...
DisplayName               : Carol Olson
Name                      : carolo
DistinguishedName         : CN=carolo,OU=Soft Deleted
                            Objects,OU=contoso.onmicrosoft.com,OU=Microsoft
Exchange
                            Hosted
Organizations,DC=NAMPR06A007,DC=PROD,DC=OUTLOOK,DC=COM
ExchangeGuid              : e646a369-00bf-43d3-837a-8eae8b301d44
IsInactiveMailbox         : True
LitigationHoldEnabled     : False
LitigationHoldDuration    : Unlimited
LitigationHoldDate        :
LitigationHoldOwner       :
InPlaceHolds              : {mbxcdbbb86ce60342489bff371876e7f224:3}
ComplianceTagHoldApplied  : False
...
DisplayName               : Megan Bowen
Name                      : meganb
DistinguishedName         : CN=meganb,OU=Soft Deleted
                            Objects,OU=contoso.onmicrosoft.com,OU=Microsoft
Exchange
                            Hosted
Organizations,DC=NAMPR06A007,DC=PROD,DC=OUTLOOK,DC=COM
ExchangeGuid              : 36ec77cb-c524-468a-a8e8-bfb75e01a176
IsInactiveMailbox         : True
```

```
LitigationHoldEnabled    : False
LitigationHoldDuration   : Unlimited
LitigationHoldDate       :
LitigationHoldOwner      :
InPlaceHolds             : {mbx6fe063689d404a5bb9940eed0f0bf5d2:1}
ComplianceTagHoldApplied : True
...
DisplayName              : Mario Necaise
Name                     : marion
DistinguishedName        : CN=marion,OU=Soft Deleted
                           Objects,OU=contoso.onmicrosoft.com,OU=Microsoft
Exchange
                           Hosted
Organizations,DC=NAMPR06A007,DC=PROD,DC=OUTLOOK,DC=COM
ExchangeGuid             : 0579e039-a695-40d5-8f0a-0dc04f4b4c8f
IsInactiveMailbox        : True
LitigationHoldEnabled    : False
LitigationHoldDuration   : Unlimited
LitigationHoldDate       :
LitigationHoldOwner      :
InPlaceHolds             : {}
ComplianceTagHoldApplied : False
...
DisplayName              : Abraham McMahon
Name                     : abrahamm
DistinguishedName        : CN=abrahamm,OU=Soft Deleted
                           Objects,OU=contoso.onmicrosoft.com,OU=Microsoft
Exchange
                           Hosted
Organizations,DC=NAMPR06A007,DC=PROD,DC=OUTLOOK,DC=COM
ExchangeGuid             : 55ad8905-4e68-4c8d-940d-e068ec6b51fc
IsInactiveMailbox        : True
LitigationHoldEnabled    : False
LitigationHoldDuration   : Unlimited
LitigationHoldDate       :
LitigationHoldOwner      :
InPlaceHolds             : {UniH7d895d48-7e23-4a8d-8346-533c3beac15d}
ComplianceTagHoldApplied : False
...
DisplayName              : Pilar Pinilla
Name                     : pilarp
DistinguishedName        : CN=pilarp,OU=Soft Deleted
                           Objects,OU=contoso.onmicrosoft.com,OU=Microsoft
Exchange
                           Hosted
Organizations,DC=NAMPR06A007,DC=PROD,DC=OUTLOOK,DC=COM
ExchangeGuid             : 8d7867d6-bb6d-4cd8-a51f-09d208f97fcc
IsInactiveMailbox        : True
LitigationHoldEnabled    : False
LitigationHoldDuration   : Unlimited
LitigationHoldDate       :
LitigationHoldOwner      :
InPlaceHolds             : {c0ba3ce811b6432a8751430937152491}
ComplianceTagHoldApplied : False
```

The following table identifies the six different hold types that were used to make each mailbox inactive from the above example.

Expand table

Inactive mailbox	Hold type	How to identify the hold on the inactive mailbox
Ann Beebe	Litigation Hold	The `LitigationHoldEnabled` property is set to `True` indicating the mailbox is on Litigation Hold. Additionally, the `LitigationHoldDuration` is set to `365.00:00:00` indicating that mailbox items will no longer be subject to litigation hold 365 days after their creation date (sent/received). The `LitigationHoldDate` indicates the date LitigationHold was enabled and `LitigationHoldOwner` identifies the person who initiated the litigation hold.
Carol Olson	Microsoft 365 retention policy from the Microsoft Purview portal that is applied to specific mailboxes	The `InPlaceHolds` property contains the GUID of the Microsoft 365 retention policy that's applied to the inactive mailbox. You can identify the hold type is a retention policy that's applied to specific mailboxes because the GUID starts with the `mbx` prefix and ends in a `:2` or `:3`. For more information, see Understanding the format of the InPlaceHolds value for retention policies.
Megan Bowen	Microsoft 365 retention label with a retain-only, or retain and delete action is applied to at least one item in the mailbox	The `ComplianceTagHoldApplied` property is `True` indicating an item is labeled with a retain-only, or retain and delete label. Additionally, the `InPlaceHolds` property contains the GUID of the Microsoft 365 retention label policy that's applied to the inactive mailbox. For more information, see Identifying mailboxes on hold because a retention label has been applied to a folder or item
Mario Necaise	Organization-wide Microsoft 365 retention policy from the Microsoft Purview portal	The `InPlaceHolds` property is empty, `LitigationHoldEnabled` is `False` and `ComplianceTagHoldApplied` is `False`. This result indicates that one or more entire (Exchange) location Microsoft 365 retention policies applied to the organization that the inactive mailbox is inheriting.

Inactive mailbox	Hold type	How to identify the hold on the inactive mailbox
		For more information, see How to confirm that an organization-wide retention policy is applied to a mailbox
Abraham McMahon	eDiscovery case hold in the Microsoft Purview portal	The `InPlaceHolds` property contains the GUID of the eDiscovery case hold that's placed on the inactive mailbox. You can tell this result is an eDiscovery case hold because the GUID starts with the `UniH` prefix. For more information, see eDiscovery holds.
Pilar Pinilla	In-Place Hold	The `InPlaceHolds` property contains the GUID of the In-Place Hold that's placed on the inactive mailbox. You can tell this result is an In-Place Hold because the GUID doesn't start with a prefix. **NOTE**: As of October 1, 2020, the hold duration of in-place holds can no longer be changed. You can only remove an In-Place Hold that results in the deletion of the inactive mailbox. For more information, see Retirement of legacy eDiscovery tools.

Step 2: Change the hold duration for an inactive mailbox

After you identify what type of hold is placed on the inactive mailbox (and whether there are multiple holds), the next step is to change the duration for the hold. The process varies depending on the type of hold applied.

- Change the duration for a Microsoft 365 retention policy
- Change the duration for a Microsoft 365 retention label
- Change the duration for an eDiscovery Hold
- Change the duration for a Litigation Hold
- Change the duration for an In-Place Hold

Change the duration for a Microsoft 365 retention policy

In order to modify the hold duration for a Microsoft 365 retention policy, you must first identify the policy affecting the inactive mailbox by running `Get-RetentionCompliancePolicy` with the associated GUID from the `InPlaceHolds` property on the mailbox in Security & Compliance PowerShell.

Be sure to remove the prefix and suffix from the GUID when running this command. For example, using our previous sample information, take the `InPlaceHolds` value of `mbxcdbbb86ce60342489bff371876e7f224:3` then remove `mbx` and `:3` resulting in a policy GUID of `cdbbb86ce60342489bff371876e7f224`. In this example, run:

```
PowerShell

Get-RetentionCompliancePolicy cdbbb86ce60342489bff371876e7f224 | FL Name
```

Once you know the name of the policy, you can simply modify the retention policy in the Microsoft Purview portal. Retention policies are typically applied to more than one location, so modifying the policy affects all applied locations - both inactive and active, which may also include locations other than Exchange. For more information, see Create and configure retention policies.

> **ⓘ Important**
>
> Retention policies with **preservation lock** enabled can have the retention period extended, but not decreased or removed.

If the intention is to modify the retention period for only inactive mailboxes, or only specific inactive mailboxes, you may consider deploying adaptive policy scopes, which can be used to individually target specific mailboxes - or mailbox types, such as inactive mailboxes - using Microsoft Entra ID and Exchange attributes and properties.

Change the duration for a Microsoft 365 retention label

As with retention policies, when you modify the hold duration of a Microsoft 365 retention label, you must first identify the policy that publishes the label affecting the content within the inactive mailbox. To identify that policy, run `Get-RetentionCompliancePolicy` with the associated GUID from the `InPlaceHolds` property on the mailbox in Security & Compliance PowerShell.

Be sure to remove the prefix and suffix from the GUID when running this command. For example, using the previous sample information, you take the `InPlaceHolds` value of

`mbx6fe063689d404a5bb9940eed0f0bf5d2:1` then remove `mbx` and `:1`, which results in a policy GUID of `6fe063689d404a5bb9940eed0f0bf5d2`. In this example, you run:

PowerShell

```
Get-RetentionCompliancePolicy 6fe063689d404a5bb9940eed0f0bf5d2 | FL Name
```

After you identify the policy, you'll know which labels have been published and their settings. Because labels apply to individual items, depending on the number of labels published with the policy and their settings, you may not be able to directly identify which label is affecting the content.

One method that you can use to identify the content each label applies to is to use Content Search. For example, using the previous sample information, assume the policy publishes several labels, one of which is named "HR-Content". With the correct permissions, you can run a Content Search with the New-ComplianceSearch PowerShell command, specify the inactive mailbox's primary SMTP address, pre-pended with a period (`.`), and the `-AllowNotFoundExchangeLocationsEnabled $true` parameter to skip validation:

PowerShell

```
New-ComplianceSearch -Name "MeganB Inactive Mailbox HR-Content Label Search"
-ExchangeLocation .meganb@contoso.onmicrosoft.com -
AllowNotFoundExchangeLocationsEnabled $true -ContentMatchQuery
"compliancetag=HR-Content"
```

After the search is created, start the search using the following command:

PowerShell

```
Start-ComplianceSearch "MeganB Inactive Mailbox HR-Content Label Search"
```

Using this method, you can then identify which labels from the identified label policy apply to content within the inactive mailbox so that you can modify their retention periods. Retention labels are typically applied to more than one location, so modifying a label affects all applied locations and labeled content, which may also include locations and content other than Exchange. For more information, see Publish retention labels and apply them in apps.

> ⓘ **Note**

Not all types of retention labels can be modified. For some labels, you may only be able to increase the time of retention, and for others you may not be able to modify the retention period at all.

Change the duration for an eDiscovery Hold

Holds associated with eDiscovery cases are indefinite holds, which means there's no hold duration that can be changed. Items are held forever or until the hold is removed or the case is closed.

Change the duration for a Litigation Hold

You must use Exchange Online PowerShell to change the hold duration for a Litigation Hold that is placed on an inactive mailbox. You can't use the Exchange admin center. Run the following command to change the hold duration. In this example, the hold duration is changed to an unlimited period of time:

PowerShell

```
Set-Mailbox -InactiveMailbox -Identity <identity of inactive mailbox> -
LitigationHoldDuration unlimited
```

The result is that items in the inactive mailbox are retained indefinitely or until the hold is removed or the hold duration is changed to a different value.

Tip

The best way to identify an inactive mailbox is by using its **Distinguished Name** or **Exchange GUID** value. Using one of these values helps prevent accidentally specifying the wrong mailbox.

Change the duration for an In-Place Hold

In-Place Holds are retired and can no longer be modified. If an inactive mailbox has an In-Place Hold applied to it, you can't change the hold duration. You can only remove the In-Place Hold, which will result in the deletion of the inactive mailbox. For more information, see Delete an inactive mailbox.

More information

- **How is the hold duration calculated for an item in an inactive mailbox?** The duration is calculated from the original date a mailbox item was received or created.

- **What happens when the hold duration expires?** When the hold duration expires for a mailbox item in the Recoverable Items folder, the item is permanently deleted (purged) from the inactive mailbox. If there's no duration specified for the hold placed on the inactive mailbox, items in the Recoverable Items folder are never purged (unless the hold duration for the inactive mailbox is changed).

- **Is an Exchange MRM policy still processed on inactive mailboxes?** If an MRM retention policy was applied to a mailbox before it went inactive, any deletion policies (MRM retention tags configured with a **Delete** action) continue to be processed on the inactive mailbox. That means items that are tagged with an MRM deletion policy are be moved to the Recoverable Items folder when the retention period expires. Those items are purged from the inactive mailbox when the hold duration expires. If a hold duration isn't specified for the inactive mailbox, items in the Recover Items folder are retained indefinitely.

  Conversely, any archive policies (MRM retention tags configured with a **MoveToArchive** action) that are included in the MRM retention policy assigned to an inactive mailbox are ignored. That means items in an inactive mailbox that are tagged with an archive policy remain in the primary mailbox when the retention period expires. They're not moved to the archive mailbox or to the Recoverable Items folder in the archive mailbox. They are retained indefinitely.

  > ⓘ **Note**
  >
  > Applying an Exchange retention policy (the messaging records management, or MRM, feature in Exchange Online) does not create an inactive mailbox when the user account is deleted.

- **As with regular mailboxes, the Managed Folder Assistant (MFA) also processes inactive mailboxes.** In Exchange Online, the MFA processes mailboxes approximately once every seven days. After you change the hold duration for an inactive mailbox, you can use the **Start-ManagedFolderAssistant** cmdlet to immediately start processing the new hold duration for the inactive mailbox. Run the following command.

  ```
  PowerShell
  ```

```
Start-ManagedFolderAssistant -InactiveMailbox <identity of inactive
mailbox>
```

- **If many `InPlaceHolds` are placed on an inactive mailbox, not all of the hold GUIDs will be displayed.** You can run the following command to display the GUIDs for all `InPlaceHolds` that are placed on an inactive mailbox.

PowerShell

```
Get-Mailbox -InactiveMailboxOnly -Identity <identity of inactive
mailbox> | Select-Object -ExpandProperty InPlaceHolds
```

Feedback

Was this page helpful? Yes No

Recover an inactive mailbox

09/26/2025

> *Microsoft 365 licensing guidance for security & compliance*.

An inactive mailbox (which is a type of soft-deleted mailbox) is used to preserve a former employee's email after they leave your organization. If that employee returns to your organization or if another employee takes on the job responsibilities of the former employee, there are two ways that you can make the contents of the inactive mailbox available to a user:

- **Recover an inactive mailbox.** If the former employee returns to your organization, or if a new employee is hired to take on the job responsibilities of the former employee, you can recover the contents of the inactive mailbox. This method converts the inactive mailbox to a new, active mailbox that contains the contents of the inactive mailbox. After it's recovered, the inactive mailbox no longer exists. The procedures in this article describe this method.
- **Restore an inactive mailbox.** If another employee takes on the job responsibilities of the former employee, or if another user needs access to the contents of the inactive mailbox, you can restore (or merge) the contents of the inactive mailbox to an existing mailbox. You can also restore the archive from an inactive mailbox. For the procedures for this method, see Restore an inactive mailbox in Office 365.

See the More information section for more details about the differences between recovering and restoring an inactive mailbox, and for a description of what happens when an inactive mailbox is recovered. Both recover and restore operations are designed for email data and don't support restoring or recovering application data that can be stored in mailbox hidden folders.

> **Note**
>
> You can't recover or restore an inactive mailbox that's configured with an auto-expanding archive. If, for compliance reasons, you need to recover data from an inactive mailbox with an auto-expanding archive, use content search to export the data from the mailbox. This action is supported for eDiscovery purposes only, and can't be used as a backup solution. For instructions to use content search for the recovery of data for eDiscovery, see following articles:
>
> - Content search
> - Export content search results

Requirements to recover an inactive mailbox

- You must use Exchange Online PowerShell to recover an inactive mailbox. You can't use the Exchange admin center (EAC), or the Microsoft Purview portal for this procedure. For step-by-step instructions to use Exchange Online PowerShell, see Connect to Exchange Online PowerShell.

- Run the following command to get identity information for the inactive mailboxes in your organization.

  PowerShell

  ```
  Get-Mailbox -InactiveMailboxOnly | Format-List
  Name,DistinguishedName,ExchangeGuid,PrimarySmtpAddress
  ```

  Use the information returned by this command to recover a specific inactive mailbox.

Recover inactive mailboxes

Use the **New-Mailbox** cmdlet with the *InactiveMailbox* parameter to recover an inactive mailbox.

1. Create a variable that contains the properties of the inactive mailbox.

   PowerShell

   ```
   $InactiveMailbox = Get-Mailbox -InactiveMailboxOnly -Identity <identity of
   inactive mailbox>
   ```

   > **Important**
   >
   > In the previous command, use the value of the **ExchangeGUID** property to identify the inactive mailbox.

2. This example uses the properties obtained in the previous command and recovers the inactive mailbox to an active mailbox for the user Ann Beebe. Be sure that the values specified for the *Name* and *MicrosoftOnlineServicesID* parameters are unique within your organization.

   PowerShell

   ```
   New-Mailbox -InactiveMailbox $InactiveMailbox.ExchangeGuid -Name annbeebe -
   FirstName Ann -LastName Beebe -DisplayName "Ann Beebe" -
   ```

```
MicrosoftOnlineServicesID Ann.Beebe@contoso.com -Password (ConvertTo-
SecureString -String 'P@ssw0rd' -AsPlainText -Force) -
ResetPasswordOnNextLogon $true
```

The primary SMTP address for the recovered inactive mailbox will have the same value as the one specified by the *MicrosoftOnlineServicesID* parameter.

After you recover an inactive mailbox, a new user account is also created. You need to activate this user account by assigning a license. To assign a license in the Microsoft 365 admin center, see Add users and assign licenses at the same time.

More information

- **What's the main difference between recovering and restoring an inactive mailbox?** When you recover an inactive mailbox, the mailbox is converted to a new mailbox, the contents and folder structure of the inactive mailbox are retained, and the mailbox is linked to a new user account. After it's recovered, the inactive mailbox no longer exists, and any changes made to the content in the new mailbox will affect the content that was originally on hold in the inactive mailbox. Conversely, when you restore an inactive mailbox, the contents are merely copied to another mailbox. The inactive mailbox is preserved and remains an inactive mailbox. Any changes made to the content in the target mailbox won't affect the original content held in the inactive mailbox. The inactive mailbox can still be searched by using In-Place eDiscovery, its contents can be restored to another mailbox, or it can be recovered or deleted at a later date.

- **What happens when you recover an inactive mailbox?** When you recover an inactive mailbox, the following things occur:

  - The hold that was applied to an inactive mailbox is changed or removed based on the type of hold that was applied to the inactive mailbox before it was recovered.

    - **Microsoft 365 retention policy with Preservation Lock.** If the inactive mailbox was included in a retention policy that has Preservation Lock, the recovered mailbox is assigned to the same retention policy.

    - **Microsoft 365 retention policy without Preservation Lock.** The inactive mailbox is removed from the Microsoft 365 retention policy. However, Litigation Hold is enabled on the recovered mailbox to prevent the deletion of mailbox content based on any organization-wide retention policies that delete content older than a specific age. You can keep the Litigation Hold or remove it. For more information, see Create a Litigation Hold.

    - **Litigation Hold.** If Litigation Hold was enabled for the inactive mailbox, it's removed from the recovered mailbox.

    - **In-Place Hold** In-Place Holds are removed from the recovered mailbox. This means the recovered mailbox is removed as a source mailbox from any In-Place Hold or In-Place eDiscovery search.

  - The single item recovery period (which is defined by the **RetainDeletedItemsFor** mailbox property) is set to 30 days. Typically, when a new mailbox is created in Exchange Online, this retention period is set to 14 days. Setting this to the maximum value of 30 days gives you more time to recover any data that's been permanently deleted (or purged) from the inactive mailbox. You can also disable single item recovery or set the single item recovery period back to the default of 14 days. For more information, see Enable or disable single item recovery for a mailbox.

  - Retention hold is enabled, and the retention hold duration is set to 30 days. This means that the default Exchange retention policy and any organization-wide or Exchange-wide Microsoft 365 retention policies that are assigned to the new mailbox won't be processed for 30 days. This gives the returning employee or the new owner of the recovered inactive mailbox time to manage the old messages. Otherwise, the Exchange or Microsoft 365 retention policy might delete old mailbox items (or move items to the archive mailbox, if it's enabled) that have expired based on the settings configured for the Exchange or Microsoft 365 retention policies. After 30 days, the retention hold expires, the **RetentionHoldEnabled** mailbox property is set to **False**, and the Managed Folder Assistant starts processing the policies assigned to the mailbox. If you don't need this additional time, you can just remove the retention hold. Alternatively, you can increase the duration of the retention hold by using the **Set-Mailbox -EndDateForRetentionHold** command. For more information, see Place a mailbox on retention hold.

- **Put a hold on the recovered mailbox if you need to preserve the original state of the inactive mailbox.** To prevent the new mailbox owner or retention policy from permanently deleting any messages from the recovered inactive mailbox, you can place the mailbox on Litigation Hold. For more information, see Create a Litigation Hold.

- **What user ID can you use when recovering an inactive mailbox?** When you recover an inactive mailbox, the value that you specify for the *MicrosoftOnlineServicesID* parameter can be different from the original one that was associated with the inactive mailbox. You can also use the original user ID. But as previously stated, make sure that the values used for *Name* and *MicrosoftOnlineServicesID* are unique within your organization when you recover the inactive mailbox.

- **What if the mailbox retention period for the inactive mailbox hasn't expired?** If an inactive mailbox was soft-deleted less than 30 days ago, you can't use the **New-Mailbox -InactiveMailbox** command to recover it. You need to recover it by restoring the corresponding user account. For more information, see Delete a user from your organization.

- **How do you know if the soft-deleted mailbox retention period for an inactive mailbox has expired?** Run the following command:

  PowerShell

  ```
  Get-Mailbox -InactiveMailboxOnly <identity of inactive mailbox> | Format-List ExternalDirectoryObjectId
  ```

  - If there's no value for the **ExternalDirectoryObjectId** property, the mailbox retention period has expired, and you can recover the inactive mailbox by running the **New-Mailbox -InactiveMailbox** command.
  - If there's a value for the **ExternalDirectoryObjectId** property, the soft-deleted mailbox retention period hasn't expired and you have to recover the mailbox by restoring the user account.

- **Consider enabling the archive mailbox after you recover an inactive mailbox.** This lets the returning user or new employee move old messages to the archive mailbox. And when the retention hold expires, the archive policy that is part of the default Exchange MRM retention policy assigned to Exchange Online mailboxes will move items that are two years or older to the archive mailbox. If you don't enable the archive mailbox, items older than two years will remain in the user's primary mailbox. For more information, see Enable archive mailboxes.

Restore an inactive mailbox

Article • 03/26/2024

> *Microsoft 365 licensing guidance for security & compliance*.

An inactive mailbox (which is a type of soft-deleted mailbox) is used to retain a former employee's email after they leave your organization. If another employee takes on the job responsibilities of the former employee or if that employee returns to your organization, there are two ways that you can make the contents of the inactive mailbox available to a user:

- **Restore an inactive mailbox** If another employee takes on the job responsibilities of the former employee, or if another user needs access to the contents of the inactive mailbox, you can restore (or merge) the contents of the inactive mailbox to an existing mailbox. You can also restore the archive from an inactive mailbox. After it's restored, the inactive mailbox is preserved and is retained as an inactive mailbox. This article describes the procedures for restoring an inactive mailbox.

- **Recover an inactive mailbox** If the former employee returns to your organization, or if a new employee is hired to take on the job responsibilities of the former employee, you can recover the contents of the inactive mailbox. This method converts the inactive mailbox to a new mailbox that contains the contents of the inactive mailbox. After it's recovered, the inactive mailbox no longer exists. For the step-by-step procedures, see Recover an inactive mailbox in Office 365.

See the More information section in this article for more details about the differences between restoring and recovering an inactive mailbox. Both operations are designed for email data and don't support restoring or recovering application data that can be stored in mailbox hidden folders.

> **Note**
>
> You can't recover or restore an inactive mailbox that's configured with an auto-expanding archive. If, for compliance reasons, you need to recover data from an inactive mailbox with an auto-expanding archive, use content search to export the data from the mailbox. This action is supported for eDiscovery purposes only, and can't be used as a backup solution. For instructions to use content search for the recovery of data for eDiscovery, see following articles:
>
> - Content search

- Export content search results

> **Tip**
>
> If you're not an E5 customer, use the 90-day Microsoft Purview solutions trial to explore how additional Purview capabilities can help your organization manage data security and compliance needs. Start now at the Microsoft Purview trials hub. Learn details about signing up and trial terms.

Requirements to restore an inactive mailbox

- You must use Exchange Online PowerShell to restore an inactive mailbox. You can't use the Exchange admin center (EAC), or the Microsoft Purview portal for this procedure. For step-by-step instructions to use Exchange Online PowerShell, see Connect to Exchange Online PowerShell.
- Run the following command in Exchange Online PowerShell to get identity information for the inactive mailboxes in your organization.

  ```
  PowerShell

  Get-Mailbox -InactiveMailboxOnly | Format-List
  Name,DistinguishedName,ExchangeGuid,PrimarySmtpAddress
  ```

  Use the information returned by this command to identify and restore a specific inactive mailbox.
- For more information about inactive mailboxes, see Inactive mailboxes in Office 365.

Restore inactive mailboxes

Use the **New-MailboxRestoreRequest** cmdlet with the *SourceMailbox* and *TargetMailbox* parameters to restore the contents of an inactive mailbox to an existing mailbox. For more information about using this cmdlet, see New-MailboxRestoreRequest.

Before you can restore an inactive mailbox, you have to add the LegacyExchangeDN of the inactive mailbox to the target mailbox, as an X500 proxy address of the target mailbox. This has to be done because the **New-MailboxRestoreRequest** cmdlet checks to make sure the value of the **LegacyExchangeDN** property on the source and target

mailboxes is the same. After you restore the inactive mailbox, you can optionally remove the LegacyExchangeDN of the inactive mailbox from source mailbox. Be sure to wait until the mailbox restore request is complete before removing the LegacyExchangeDN.

Follow these steps to restore an inactive mailbox to an existing mailbox:

1. Create a variable that contains the properties of the inactive mailbox.

   PowerShell

   ```
   $inactiveMailbox = Get-Mailbox -InactiveMailboxOnly -Identity <identity
   of inactive mailbox>
   ```

   > **Important**
   >
   > In the previous command, use the value of the **DistinguishedName** or **ExchangeGUID** property to identify the inactive mailbox. These properties are unique for each mailbox in your organization, whereas it's possible that an active and an inactive mailbox might have the same primary SMTP address.

2. Display the LegacyExchangeDN of the inactive mailbox so that you can add it as a proxy address to the target mailbox in the next step.

   PowerShell

   ```
   $inactiveMailbox.LegacyExchangeDN
   ```

3. Add the LegacyExchangeDN of the inactive mailbox as an X500 proxy address to the target mailbox.

   PowerShell

   ```
   Set-Mailbox <identity of target mailbox> -EmailAddresses @{Add="X500:
   <LegacyExchangeDN of inactive mailbox>"}
   ```

4. Restore the contents of the inactive mailbox to an existing mailbox. The contents of the inactive mailbox (source mailbox) will be merged into the corresponding folders in the existing mailbox (target mailbox).

   PowerShell

   ```
   New-MailboxRestoreRequest -SourceMailbox
   $inactiveMailbox.DistinguishedName -TargetMailbox <identity of target
   ```

```
mailbox>
```

Alternatively, you can specify a top-level folder in the target mailbox in which to restore the contents from the inactive mailbox. If the specified target folder or target folder structure doesn't already exist in the target mailbox, it is created during the restore process.

This example copies mailbox items and subfolders from an inactive mailbox to a folder named "Inactive Mailbox" in the top-level folder structure of the target mailbox.

PowerShell

```
New-MailboxRestoreRequest -SourceMailbox
$InactiveMailbox.DistinguishedName -TargetMailbox <identity of target
mailbox> -TargetRootFolder "Inactive Mailbox"
```

5. After the restore request is complete, you can optionally remove the LegacyExchangeDN of the inactive mailbox from the target mailbox. Leaving the LegacyExchangeDN from the inactive mailbox won't affect the target mailbox.

PowerShell

```
Set-Mailbox <identity of target mailbox> -EmailAddresses
@{Remove="X500:<LegacyExchangeDN of inactive mailbox>"}
```

Restore the archive from an inactive mailbox

If an inactive mailbox has an archive mailbox, you can also restore it to the archive mailbox of an existing mailbox. To restore the archive from an inactive mailbox, you have to add the *SourceIsArchive* and *TargetIsArchive* switches to the command used to restore an inactive mailbox.

1. Create a variable that contains the properties of the inactive mailbox.

PowerShell

```
$inactiveMailbox = Get-Mailbox -InactiveMailboxOnly -Identity <identity
of inactive mailbox>
```

> **Note**

In the previous command, use the value of the **DistinguishedName** or **ExchangeGUID** property to identify the inactive mailbox. These properties are unique for each mailbox in your organization, whereas it's possible that an active and an inactive mailbox might have the same primary SMTP address.

2. Display the LegacyExchangeDN of the inactive mailbox so that you can add it as a proxy address to the target mailbox in the next step.

```
PowerShell

$inactiveMailbox.LegacyExchangeDN
```

3. Add the LegacyExchangeDN of the inactive mailbox as an X500 proxy address to the target mailbox.

```
PowerShell

Set-Mailbox <identity of target mailbox> -EmailAddresses @{Add="X500:
<LegacyExchangeDN of inactive mailbox>"}
```

4. Restore the contents of the archive from the inactive mailbox (source archive) to the archive of an existing mailbox (target archive). In this example, the contents from the source archive are copied to a folder named "Inactive Mailbox Archive" in the archive of the target mailbox.

```
PowerShell

New-MailboxRestoreRequest -SourceMailbox
$InactiveMailbox.DistinguishedName -SourceIsArchive -TargetMailbox
<identity of target mailbox> -TargetIsArchive -TargetRootFolder
"Inactive Mailbox Archive"
```

5. After the restore request is complete, you can optionally remove the LegacyExchangeDN of the inactive mailbox from the target mailbox. Leaving the LegacyExchangeDN from the inactive mailbox won't affect the target mailbox.

```
PowerShell

Set-Mailbox <identity of target mailbox> -EmailAddresses
@{Remove="X500:<LegacyExchangeDN of inactive mailbox>"}
```

More information

- **What's the main difference between recovering and restoring an inactive mailbox?** When you recover an inactive mailbox, the mailbox is converted to a new mailbox. The contents and folder structure of the inactive mailbox are retained, and the mailbox is linked to a new user account. After it's recovered, the inactive mailbox no longer exists, and any changes made to the content in the new mailbox will affect the content that was originally on hold in the inactive mailbox. Conversely, when you restore an inactive mailbox, the contents are merely copied to another mailbox. The inactive mailbox is preserved and remains an inactive mailbox. Any changes made to the content in the target mailbox won't affect the original content held in the inactive mailbox. The inactive mailbox can still be searched by using the Content Search tool, its contents can be restored to another mailbox, or it can be recovered or deleted at a later date.

- **How do you find inactive mailboxes?** To get a list of the inactive mailboxes in your organization and display information that is useful for restoring an inactive mailbox, you can run this command.

  PowerShell

  ```
  Get-Mailbox -InactiveMailboxOnly | Format-List
  Name,PrimarySMTPAddress,DistinguishedName,ExchangeGUID,LegacyExchangeDN
  ,ArchiveStatus
  ```

- **Use a Microsoft 365 retention policy or Litigation Hold or to retain inactive mailbox content.** If you want to retain the state of an inactive mailbox after it's restored, you can apply a Microsoft 365 retention policy to the target mailbox or place the target mailbox on Litigation Hold before you restore the inactive mailbox. This will prevent the permanent deletion of any items from the inactive mailbox after they're restored to the target mailbox.

- **Enable retention hold on the target mailbox before you restore an inactive mailbox.** Because mailbox items from an inactive mailbox could be old, you might consider enabling retention hold on the target mailbox before you restore an inactive mailbox. When you put a mailbox on retention hold, the retention policy that's assigned to it won't be processed until the retention hold is removed or until the retention hold period expires. This gives the owner of the target mailbox time to manage old messages from the inactive mailbox. Otherwise, the retention policy might delete old items (or move items to the archive mailbox, if it's enabled) that have expired based on the retention settings configured for the target mailbox. For more information, see Place a mailbox on retention hold in Exchange Online.

- **You can use other parameters with the New-MailboxRestoreRequest cmdlet to implement different restore scenarios for inactive mailboxes.** For example, you

can run this command to restore the archive from the inactive mailbox into the primary mailbox of the target mailbox.

```
PowerShell

New-MailboxRestoreRequest -SourceMailbox <inactive mailbox> -
SourceIsArchive -TargetMailbox <target mailbox> -TargetRootFolder
"Inactive Mailbox Archive"
```

You can also restore the inactive primary mailbox into the archive of the target mailbox by running this command.

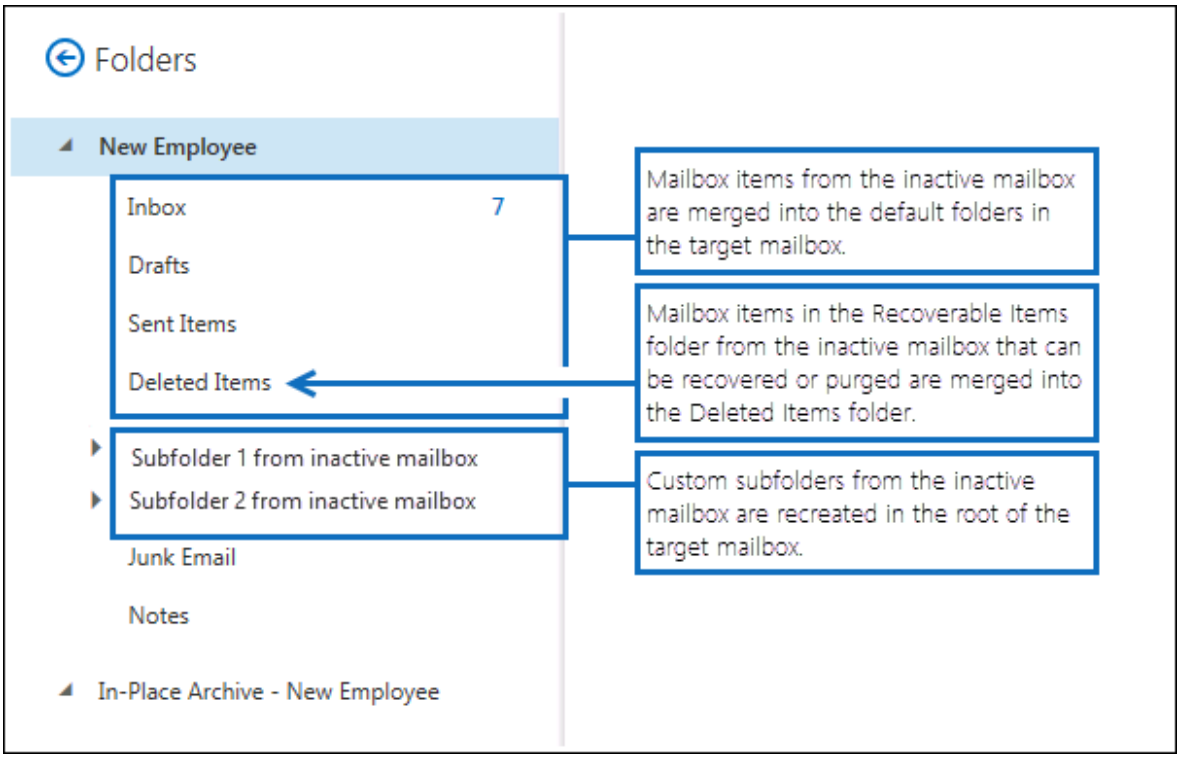
```
PowerShell

New-MailboxRestoreRequest -SourceMailbox <inactive mailbox> -
TargetMailbox <target mailbox> -TargetIsArchive -TargetRootFolder
"Inactive Mailbox"
```

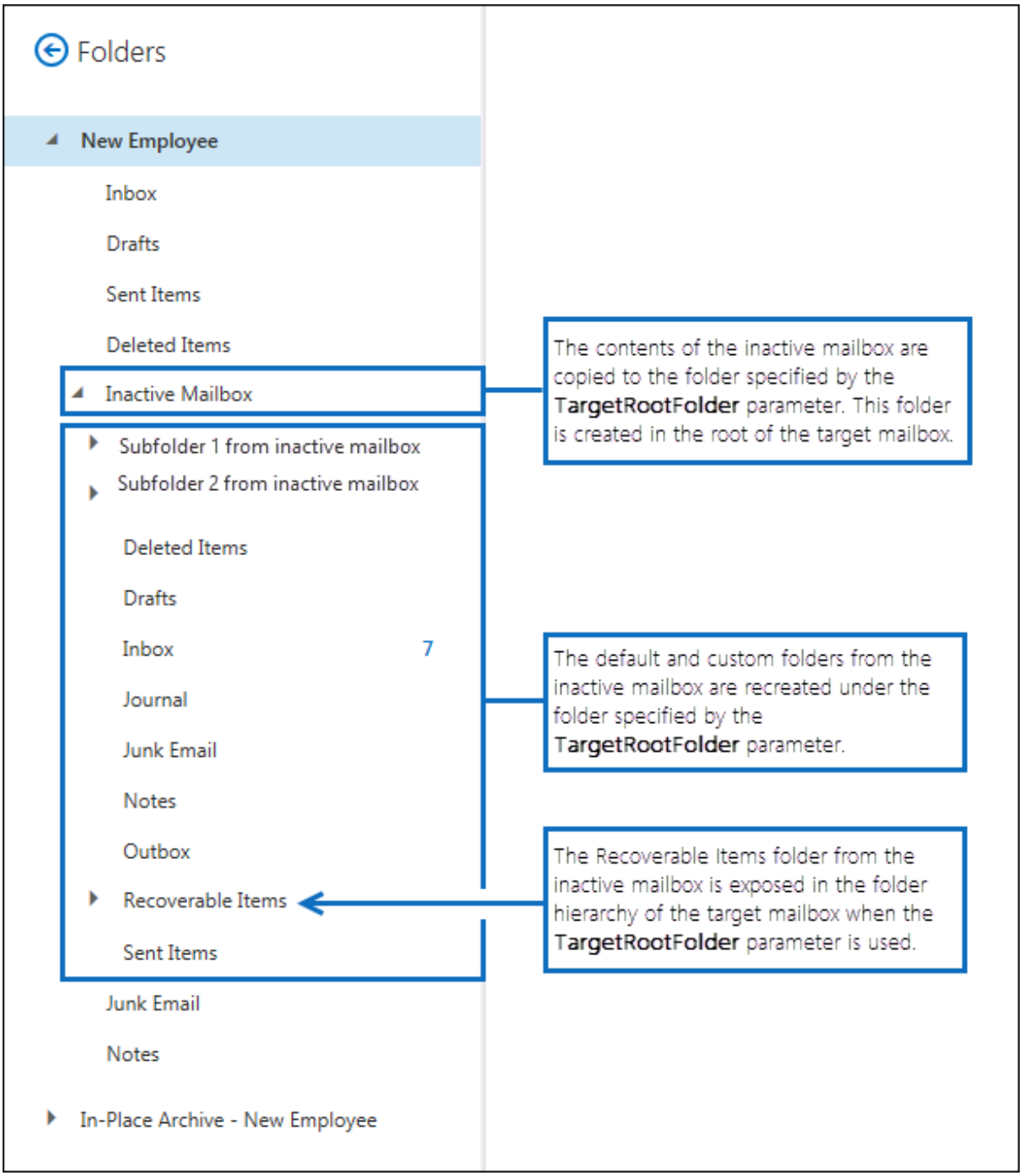
- **What does the TargetRootFolder parameter do?** As previously explained, you can use the **TargetRootFolder** parameter to specify a folder in the top of the folder structure (also called the root) in the target mailbox in which to restore the contents of the inactive mailbox. If you don't use this parameter, mailbox items from the inactive mailbox are merged into the corresponding default folders of the target mailbox, and custom folders are re-created in the root of the target mailbox. The following illustrations highlight these differences between not using and using the **TargetRootFolder** parameter.

  > **Note**
  >
  > To restore the contents of a primary mailbox to an archive mailbox, use the **TargetRootFolder** parameter to specify the archive mailbox folders to migrate the content to. This content will be visible after it's restored. If you don't use this parameter, the restored content is not visible because it's mapped to locations in the archive mailbox that aren't visible to users.

  Folder hierarchy in the target mailbox when the TargetRootFolder parameter isn't used



Folder hierarchy in the target mailbox when the TargetRootFolder parameter is used



Feedback

Was this page helpful? Yes No

Delete an inactive mailbox

> *Microsoft Purview service description*

When you no longer need to keep the contents of an inactive mailbox, you can permanently delete the mailbox by removing all holds on it. Then, the inactive mailbox is automatically marked for deletion and permanently deleted after it's processed.

Holds can be any of the following:

- Microsoft 365 retention policies and labels with retain or retain and delete settings
- A hold associated with an eDiscovery case
- Litigation hold
- An In-Place Hold
- A delay hold, which Microsoft 365 applies automatically after another hold is removed

Important

It is strongly recommended that you use Microsoft 365 retention when proactively planning to use inactive mailboxes.

- eDiscovery holds are intended for specific, time-bound cases related to a legal issue. At some point, a legal case will probably end and the holds associated with the case will be removed and the eDiscovery case will be closed (or deleted). If a hold that's placed on an inactive mailbox is associated with an eDiscovery case, and the hold is released or the eDiscovery case is closed or deleted, the inactive mailbox will be permanently deleted.

- In-Place Holds in the Exchange admin center are now retired. As of July 1, 2020, new In-Place Holds were unable to be created in Exchange Online. As of October 1, 2020, the hold duration of in-place holds could no longer be changed. Any inactive mailbox that has an In-Place Hold applied can only be deleted by removing the In-Place Hold. Existing inactive mailboxes that are on In-Place Hold will continue to be preserved until the hold is removed. For more information about In-Place Holds retirement, see Retirement of legacy eDiscovery tools.

- **Litigation hold** remains supported as an alternative method to retain content in a mailbox and make it inactive after a user account is deleted. However, as an older technology, we recommend you use Microsoft 365 retention instead.

See the More information section for a description of what happens after holds are removed from an inactive mailbox.

Before you delete an inactive mailbox

- You must use Exchange Online PowerShell to remove holds from an inactive mailbox. You can't use the Exchange admin center (EAC), or the Microsoft Purview portal for these procedures. For step-by-step instructions to use Exchange Online PowerShell, see Connect to Exchange Online PowerShell.
- Identify the holds on an inactive mailbox by using the instructions from Step 1: Identify the holds on an inactive mailbox. You'll need this information to know which holds to remove, and identifying information.
- Consider whether you want to copy the contents of an inactive mailbox to another mailbox before you remove the final hold that will result in the deletion of the inactive mailbox. For details, see Restore an inactive mailbox.
- Be aware that if you remove the final hold from an inactive mailbox the mailbox will be permanently deleted after the 30-day soft-deleted mailbox retention period expires. After the inactive mailbox is permanently deleted, it can't be recovered. Before you remove a hold, be sure that you no longer need the contents in the mailbox. If you want to reactivate an inactive mailbox, you can recover it. For details, see Recover an inactive mailbox.
- For more information about inactive mailboxes, see Learn about inactive mailboxes.

How to delete an inactive mailbox by removing holds

After you've identified what type of hold is placed on the inactive mailbox and whether there are multiple holds, you can remove the holds on the mailbox. An exception is if the inactive mailbox has one or more retention labels configured to retain items. In this case, you can't remove the hold.

When the last hold is removed, no further action is required from you because the inactive mailbox will be automatically deleted.

Use the instructions for the type of hold you need to remove:

- Remove an inactive mailbox from a retention policy
- Remove an inactive mailbox from an eDiscovery hold
- Remove Litigation hold from an inactive mailbox
- Remove an In-Place Hold from an inactive mailbox
- Remove a delay hold from an inactive mailbox

> **Note**
>
> A delay hold is applied automatically after you remove another hold, so it's a common reason a mailbox is still listed as inactive after you think you removed the last hold. Check for a delay hold before you conclude that hold removal didn't work.

> **Caution**
>
> When the instructions require you to specify an identity for the inactive mailbox, the best way to do this is by using its Distinguished Name or Exchange GUID value. Using one of these values instead of the SMTP address helps prevent accidentally specifying the wrong mailbox. However, for instructions that specifically tell you to use the primary SMTP address, you can't use the Distinguished Name or Exchange GUID value and must use the primary SMTP address.

Remove an inactive mailbox from a retention policy

The procedure to remove an inactive mailbox from a Microsoft 365 retention policy depends on whether the policy scope is adaptive or static.

You won't be able to remove an inactive mailbox if the retention policy has Preservation Lock applied.

How to remove an inactive mailbox from a retention policy with an adaptive scope

When you use the advanced query builder to define your adaptive scope for mailboxes, modify it to exclude inactive mailboxes by using the OPATH property *IsInactiveMailbox*:

Console

```
IsInactiveMailbox -eq "False"
```

You can't exclude inactive mailboxes by using the simple query builder. If the retention policy is configured to use an adaptive scope that used the simple query builder, create a new adaptive policy scope that uses the advanced query builder. Then modify the retention policy to use this new adaptive scope.

How to remove an inactive mailbox from a retention policy with a static scope

This procedure depends on whether the retention policy applied to the inactive mailbox is configured for the entire policy location (sometimes referred to as "organization-wide") or the policy is configured to include specific mailboxes (sometimes referred to as "explicit"). For more information, see Configuration information for static scopes.

Use the following PowerShell cmdlets to retrieve a policy GUID and identify the type of hold for a mailbox:

- If the policy is configured for the entire location: Use the **Get-OrganizationConfig** cmdlet in Exchange Online PowerShell to get information about these policies.
- If the policy is configured to include specific mailboxes: Use the **Get-Mailbox -IncludeInactiveMailbox** cmdlet and parameter in Exchange Online PowerShell to get information about these policies.

Static scope: How to remove an inactive mailbox from a retention policy that's configured for the entire location

Run the following PowerShell command for a specific retention policy:

PowerShell

```
Set-Mailbox <identity of inactive mailbox> -ExcludeFromOrgHolds <retention policy GUID without prefix or suffix>
```

Alternatively, you can run the following PowerShell commands to remove the inactive mailbox:

- For retention policies that use the older retention locations, such as **Exchange mailboxes**, **Teams chats and Copilot interactions**, and **Teams channel messages**:

  PowerShell

```
Set-Mailbox <identity of inactive mailbox> -ExcludeFromAllOrgHolds
```

- For retention policies that use the newer retention locations, such as **Teams private channel messages** (applicable pre-migration only), **Teams chats**, and **Microsoft Copilot experiences**:

  PowerShell

  ```
  Set-AppRetentionCompliancePolicy -Identity <retention policy GUID without prefix or suffix> -AddExchangeLocationException <identity of inactive mailbox> -DeletedResources "[{'EmailAddress':'<Primary SMTP address of inactive mailbox>'}]"
  ```

  You can also specify the identity of the inactive mailbox by using a variable and JSON, which is required if you need to specify more than one inactive mailbox with the same command. For more information, see the PowerShell examples.

  > **Important**
  >
  > Before you run the command, make sure you read the Caution information for the DeletedResources parameter about duplicate SMTP addresses.

Static scope: How to remove an inactive mailbox from a retention policy that's configured to include specific mailboxes

Run the following Security & Compliance PowerShell commands:

- For retention policies that use the older retention locations, such as **Exchange mailboxes**, **Teams chats and Copilot interactions**, and **Teams channel messages**:

  PowerShell

  ```
  Set-RetentionCompliancePolicy -Identity <retention policy GUID without prefix or suffix> -RemoveExchangeLocation <identity of inactive mailbox>
  ```

- For retention policies that use the newer retention locations, such as **Teams private channel messages**, **Teams chats**, and **Microsoft Copilot experiences**:

  PowerShell

  ```
  Set-AppRetentionCompliancePolicy -Identity <retention policy GUID without prefix or suffix> -RemoveExchangeLocation <identity of inactive mailbox>
  ```

Remove an inactive mailbox from an eDiscovery hold

See Removing content locations from an eDiscovery hold.

Remove a Litigation hold from an inactive mailbox

Run the following PowerShell command to remove a Litigation hold from a mailbox:

```
PowerShell
Set-Mailbox -InactiveMailbox -Identity <identity of inactive mailbox> -
LitigationHoldEnabled $false
```

Remove an In-Place Hold from an inactive mailbox

There are two ways to remove an In-Place Hold from an inactive mailbox:

- **Delete the In-Place Hold object**. If the inactive mailbox that you want to permanently delete is the only source mailbox for an In-Place Hold, you can just delete the In-Place Hold object.

  > ⓘ Note
  >
  > You have to disable the hold before you can delete an In-Place Hold object. If you try to delete an In-Place Hold object that has the hold enabled, you'll receive an error message.

- **Remove the inactive mailbox as a source mailbox of an In-Place Hold**. If you want to retain other source mailboxes for an In-Place Hold, you can remove the inactive mailbox from the list of source mailboxes and keep the In-Place Hold object.

How to remove an inactive mailbox from an eDiscovery In-Place Hold

If the eDiscovery hold is active, the user can be removed from eDiscovery location.

Use the Invoke-HoldRemovalAction) cmdlet to find and remove an eDiscovery in-place hold

1. When running the Invoke-HoldRemovalAction with GetHolds parameter, holds that apply on specified mailbox can be found:

   ```
   PowerShell
   ```

```
Invoke-HoldRemovalAction -Action GetHolds -ExchangeLocation <identity of inactive
mailbox>
```

If legacy eDiscovery hold is present, HoldId value will be only a GUID without any prefix or have "cld" one, see How to identify the type of hold placed on an Exchange Online mailbox.

1. The legacy eDiscovery holds or inactive case hold policy can be removed from the inactive mailbox by providing the HoldId:

PowerShell

```
Invoke-HoldRemovalAction -Action RemoveHold -ExchangeLocation <identity of
inactive mailbox> -HoldId <hold ID>
```

Remove a delay hold from an inactive mailbox

After you remove any other type of hold, Microsoft 365 automatically applies a *delay hold* so that content isn't purged immediately. While a delay hold is applied, the mailbox is still treated as being on hold, so it remains an inactive mailbox and isn't marked for deletion. This is a common reason a mailbox is still listed as inactive after you removed what you thought was the last hold.

A delay hold is applied by setting either of the following mailbox properties to **True**:

- *DelayHoldApplied* for Outlook content
- *DelayReleaseHoldApplied* for cloud content generated by non-Outlook apps such as Microsoft Teams

Run the following command to check whether either property is set on the inactive mailbox:

PowerShell

```
Get-Mailbox -InactiveMailbox -Identity <identity of inactive mailbox> | FL *Delay*
```

A delay hold expires automatically after 30 days. To remove it sooner, run one or both of the following commands, depending on which property is set to **True**:

PowerShell

```
Set-Mailbox -InactiveMailbox -Identity <identity of inactive mailbox> -
RemoveDelayHoldApplied
```

PowerShell

```
Set-Mailbox -InactiveMailbox -Identity <identity of inactive mailbox> -
RemoveDelayReleaseHoldApplied
```

> ⓘ Note
>
> You must be assigned the Legal Hold role in Exchange Online to use the *RemoveDelayHoldApplied* or *RemoveDelayReleaseHoldApplied* parameters.

Removing the delay hold also recalculates the hold status of the mailbox. If no other holds remain, the mailbox stops being an inactive mailbox and becomes a soft-deleted mailbox.

For more information about delay holds, including how they interact with each other, see Managing mailboxes on delay hold.

More information

- **An inactive mailbox is a type of soft-deleted mailbox.** In Exchange Online, a soft-deleted mailbox is a mailbox that's been deleted but can be recovered within a specific retention period. For soft-deleted mailboxes that aren't on hold, the mailbox is recoverable within 30 days. An inactive mailbox (a mailbox on hold before it was deleted) will remain in a soft-deleted with hold state until the hold is removed. After the last hold is removed from an inactive mailbox, the mailbox will no longer be in an inactive state. Instead it will become soft-deleted and remain in Exchange Online for 30 days from the day the hold was removed and recoverable during that time. After 30 days, a soft-deleted mailbox is marked for permanent deletion and can't be recovered.

- **What happens after you remove the last hold on an inactive mailbox?** The mailbox is treated like other soft-deleted mailboxes and is marked for permanent deletion after the 30-day soft-deleted mailbox retention period expires. This retention period starts on the date when the last hold is removed from the inactive mailbox. The *InactiveMailboxRetireTime* property is set when the mailbox transitions from being inactive (soft-deleted on hold) to no longer being inactive (soft-deleted with no holds). At that point, the *InactiveMailboxRetireTime* property is set to the current date when the transition occurred. There is an assistant that runs (called the *MailboxLifeCycle* assistant) that looks for mailboxes that have the *InactiveMailboxRetireTime* property set. If "InactiveMailboxRetireTime + 30 days" is less than the current date, then it will purge the mailbox.

If mailboxes still show as *Inactive* after you remove the last hold and *InactiveMailboxRetireTime* isn't set, first confirm that no hold is still applied — including a delay hold, which is applied automatically after another hold is removed:

PowerShell

```
Get-Mailbox -InactiveMailbox -Identity <identity of inactive mailbox> | FL
LitigationHoldEnabled,ComplianceTagHoldApplied,InPlaceHolds,*Delay*
```

If a hold is still applied, remove that hold. Recalculating the hold status doesn't remove a hold, so it doesn't change a mailbox that's still on hold.

If no holds remain, you can force the recalculation of the hold status for the inactive mailbox by using the following command:

PowerShell

```
Set-Mailbox -Identity <identity of inactive mailbox> -RecalculateInactiveMailbox
```

- **Is an inactive mailbox permanently deleted immediately after all holds are removed?** A formerly inactive mailbox will be available in the soft-deleted state for 30 days. After 30 days, the mailbox will be marked for permanent deletion.
- **How do you display information about an inactive mailbox after all holds are removed?** After all holds are removed and the inactive mailbox is reverted back to a soft-deleted mailbox, it won't be returned by using the *InactiveMailboxOnly* parameter with the **Get-Mailbox** cmdlet. But you can display information about the mailbox by using the **Get-Mailbox -SoftDeletedMailbox** command. For example:

PowerShell

```
Get-Mailbox -SoftDeletedMailbox -Identity pilarp | FL
Name,Identity,LitigationHoldEnabled,In
Placeholds,WhenSoftDeleted,IsInactiveMailbox,WasInactiveMailbox,InactiveMailboxRe
tireTime

Name                   : pilarp
Identity               : Soft Deleted Objects\pilarp
LitigationHoldEnabled  : False
InPlaceHolds           : {}
WhenSoftDeleted        : 6/16/2022 1:19:04 AM
IsInactiveMailbox      : False
WasInactiveMailbox     : True
InactiveMailboxRetireTime : 9/30/2022 11:16:23 PM
```

In the above example, the *WhenSoftDeleted* property identifies the soft-deleted date, which in this example is June 16, 2022. The *WasInactiveMailbox* property is listed as `True` because it was previously an inactive mailbox. The mailbox will be permanently deleted 30 days after September 30, 2022.

Last updated on 08/05/2026

Learn about importing your organization's PST files

> **Note**
>
> This article is for administrators. Are you trying to import PST files to your own mailbox? See Import email, contacts, and calendar from an Outlook .pst file.

Use the Import service in the Microsoft Purview portal to quickly bulk-import PST files to Exchange Online mailboxes in your organization. You can import PST files to Microsoft 365 in two ways:

- **Network upload**: Upload the PST files over the network to a temporary Azure Storage location in the Microsoft cloud. Then use the Microsoft 365 Import service to import the PST data to mailboxes in your organization.
- **Drive shipping**: Copy the PST files to a BitLocker-encrypted hard drive and then physically ship the drive to Microsoft. When Microsoft receives the hard drive, data center personnel upload the data to a temporary Azure Storage location in the Microsoft cloud. Then use the Microsoft 365 Import service to import the data to mailboxes in your organization.

Why import email data to Microsoft 365?

- Import your organization's archival messaging data to Microsoft 365.
- Use the Intelligent Import feature to filter the items in PST files that get imported to the target mailboxes. This feature lets you trim the imported data by setting filters.
- Importing email data to Microsoft 365 helps address your organization's compliance needs by letting you:
  - Enable archive mailboxes and auto-expanding archiving to give users additional mailbox storage space.
  - Place mailboxes on Litigation Hold to retain content.
  - Use the Content Search tool to search for mailbox content.
  - Use eDiscovery cases to manage your organization's legal investigations.
  - Use retention policies in the Microsoft Purview portal to control how long mailbox content is retained, and then delete content after the retention period expires.

- Importing data to Microsoft 365 helps protect against data loss. Email data that you import to Microsoft 365 inherits the high availability features of Exchange Online.
- Because you store email data in the cloud, users can access it from all their devices.

Select your import method

Use the following table to help decide which import method is best for your organization:

Expand table

Factor	Network upload	Drive shipping
Cost	Free	Costs apply - see Microsoft Purview service description for details
Region availability	United States, Canada, Brazil, United Kingdom, France, Germany, Switzerland, Norway, Europe, India, East Asia, Southeast Asia, Japan, Republic of Korea, Australia, UAE, Israel, Mexico, Taiwan, New Zealand	United States, Canada, Brazil, United Kingdom, Europe, India, East Asia, Southeast Asia, Japan, Republic of Korea, Australia, Mexico, Taiwan, New Zealand (not available in Germany or Switzerland)
Licensing	All commercial licensing agreements	Microsoft Enterprise Agreement (EA) only. Not available through Microsoft Products and Services Agreement (MPSA).
Max data volume	Unlimited (network dependent)	10 TB per drive, maximum 10 drives per import job
Upload time	Several hours per TB (depends on your network capacity)	7-10 business days after Microsoft receives the drive
Best for	Ongoing imports or smaller migrations	Large one-time data migrations

Requirements and limits

Before you import PST files, review the following requirements and limits that apply to both import methods.

Permissions

To create import jobs and import PST files, you need the following roles:

- **Mailbox Import Export** role in Exchange Online. By default, no role group has this role assigned. You can add it to the **Organization Management** role group, or create a new role group and assign the role.

- **Mail Recipients** role in Exchange Online. By default, the **Organization Management** and **Recipient Management** role groups have this role assigned.

  Or

- You're a global administrator in your organization.

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

> **Tip**
>
> Consider creating a new role group in Exchange Online specifically for importing PST files. Assign the **Mailbox Import Export** and **Mail Recipients** roles to the new role group, and then add members.

PST file format

Both ANSI and Unicode PST file formats are supported. Import files that use the Unicode PST file format. However, you can also import files that use the ANSI PST file format, such as those for languages that use a double-byte character set (DBCS). You can import PST files from Outlook 2007 and later versions to Microsoft 365.

For more information about ANSI PST files, see Step 4 in Use network upload to import PST files or Step 3 in Use drive shipping to import PST files.

Message size limit

If a PST file contains a mailbox item that's larger than 150 MB, the import process skips the item and doesn't import it. Items larger than 150 MB aren't imported because 150 MB is the message size limit in Exchange Online. For more information, see Message limits in Exchange Online.

Folder hierarchy limit

You can't import a PST file that has 300 or more levels of nested folders.

Supported mailbox scenarios

You can import PST files to the following mailbox types:

- **Inactive mailboxes**: You can import PST files to an inactive mailbox in Microsoft 365 by using either network upload or drive shipping. Specify the GUID of the inactive mailbox in the `Mailbox` parameter of the PST Import mapping file.
- **Online archive mailboxes in hybrid deployments**: You can import PST files to a cloud-based archive mailbox for a user whose primary mailbox is on-premises. Specify the email address for the user's on-premises mailbox in the `Mailbox` parameter, and set the `IsArchive` parameter to **TRUE** in the PST Import mapping file.
- **Public folders**: You can't import PST files to public folders.

PST file size

Upload or copy PST files that are no larger than 20 GB. Large PST files can affect the performance of the import process.

Archive mailbox import limit

When you import PST files to a user's archive mailbox (the `IsArchive` parameter set to **TRUE**), data is imported into the user's main archive mailbox only. The import doesn't target auto-expanding archive storage, because auto-expanding archive storage is provisioned on demand rather than in advance.

As a result, you can't import more than 100 GB of data into a user's archive in a single import. To import more than 100 GB, divide the data across multiple imports and allow time for the archive to auto-expand between imports. For more information, see Learn about auto-expanding archiving.

Plan your import

Before you start importing PST files, review the following planning considerations.

Import rate

A Microsoft 365 mailbox imports a PST file at a rate of about 24 GB per day. This rate isn't guaranteed and can vary based on server workload and transient performance issues. If this rate doesn't meet your needs, consider other methods for migrating email data to Microsoft 365. For more information, see Ways to migrate multiple email accounts to Microsoft 365.

If you import different PST files to different target mailboxes, the import process runs in parallel (each PST/mailbox pair is imported simultaneously). If you import multiple PST files to the same mailbox, the process imports them sequentially (one at a time).

Retention hold after import

After you import PST files to a Microsoft 365 mailbox, the retention hold setting for the mailbox is turned on indefinitely. This setting means that the retention policy assigned to the mailbox isn't processed until you turn off the retention hold or set a date to turn off the hold. If messages you import to a mailbox are old, the retention settings configured for the mailbox might cause those messages to be permanently deleted (purged). Placing the mailbox on retention hold gives the mailbox owner time to manage the newly imported messages, or gives you time to change the retention settings for the mailbox.

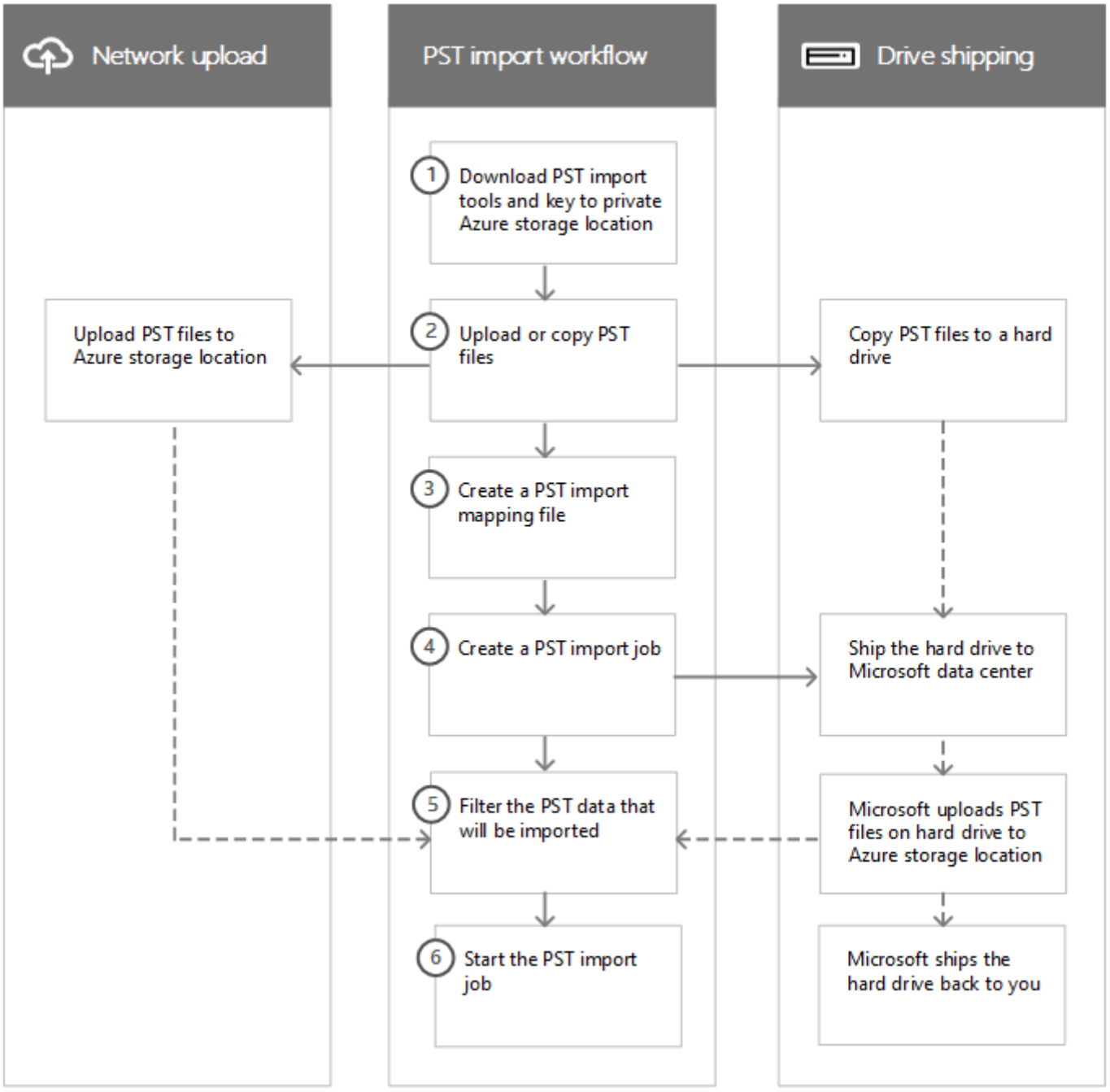
MaxReceiveSize property

By default, a Microsoft 365 mailbox can receive messages up to 35 MB. However, Microsoft 365 supports a maximum message receive size of 150 MB. If you import a PST file that contains an item larger than 35 MB, the Microsoft 365 Import service automatically sets the *MaxReceiveSize* property on the target mailbox to 150 MB.

> **Tip**
>
> To check the message receive size for a mailbox, run this command in Exchange Online PowerShell: `Get-Mailbox <user mailbox> | FL MaxReceiveSize`.

How the PST import process works

Here's an illustration and description of the complete PST import process. The illustration shows the primary workflow and highlights the differences between the network upload and drive shipping methods.



1. **Download the PST import tools and key to private Azure Storage location**: Download the tool and access key you use to upload the PST files or copy them to a hard drive. Get these items from the **Import** page in the Microsoft Purview portal. The key provides you (or Microsoft data center personnel in the case of drive shipping) with the necessary permissions to upload PST files to a private and secure Azure Storage location. This access key is unique to your organization and helps prevent unauthorized access to your PST files after you upload them to the Microsoft cloud. Importing PST files to Microsoft 365 doesn't require your organization to have a separate Azure subscription.

2. **Upload or copy the PST files**: The next step depends on whether you're using network upload or drive shipping to import PST files. In both cases, use the tool and secure storage key that you obtained in the previous step.

- **Network upload**: Use the AzCopy.exe tool (downloaded in step 1) to upload and store your PST files in an Azure Storage location in the Microsoft cloud. The Azure Storage location that you upload your PST files to is located in the same regional Microsoft datacenter as your organization.

  To upload them, the PST files that you want to import must be located in a file share or file server in your organization.

- **Drive shipping**: Use the WAImportExport.exe tool (downloaded in step 1) to copy your PST files to the hard drive. This tool encrypts the hard drive with BitLocker and then copies the PSTs to the hard drive. Like network upload, the PST files that you want to copy to the hard drive must be located in a file share or file server in your organization.

3. **Create a PST import mapping file**: After you upload the PST files to the Azure Storage location or copy them to a hard drive, create a comma-separated value (CSV) file that specifies which user mailboxes the PST files go to (and a PST file can go to a user's primary mailbox or their archive mailbox). Download a copy of the PST Import mapping file. The Microsoft 365 Import service uses the information to import the PST files.

4. **Create a PST import job**: Create a PST import job on the **Import PST files** page in the Microsoft Purview portal and submit the PST import mapping file you created in the previous step. For network upload (because the PST files are uploaded to Azure) Microsoft 365 analyzes the data in the PST files and then gives you an opportunity to set filters that control what data actually gets imported to the mailboxes specified in the PST import mapping file.

   For drive shipping, a few other things happen at this point in the process.

   - You physically ship the hard drive to a Microsoft data center (the shipping address for the Microsoft data center is displayed when the import job is created).

   - When Microsoft receives the hard drive, data center personnel upload the PST files on the hard drive to the Azure Storage location for your organization. As previously explained, your PST files are uploaded to an Azure Storage location that resides in the same regional Microsoft datacenter where your organization is located.

     > **Note**

> The PST files on the hard drive are uploaded to Azure within 7 to 10 business days after Microsoft receives the hard drive.

Like the network upload process, Microsoft 365 then analyzes the data in the PST files and gives you an opportunity to set filters that control what data actually gets imported to the mailboxes specified in the PST import mapping file.

- Microsoft ships the hard drive back to you.

5. **Filter the PST data that will be imported to mailboxes**: After you create the import job (and after the PST files from a drive shipping job are uploaded to the Azure Storage location) Microsoft 365 analyzes the data in the PST files (safely and securely) by identifying the age of the items and the different message types included in the PST files. When the analysis is completed and the data is ready to import, you have the option to import all the data contained in the PST files or you can trim the data that's imported by setting filters that control what data gets imported.

6. **Start the PST import job**: After you start the import job, Microsoft 365 uses the information in the PST import mapping file to import the PST files from the Azure Storage location to user mailboxes. Status information about the import job (including information about each PST file being imported) is displayed on the **Import PST files** page in the Microsoft Purview portal. When the import job finishes, the status for the job is set to **Complete**.

Step-by-step instructions

For detailed, step-by-step instructions for bulk-importing your organization's PST files to Microsoft 365, see one of the following articles:

- Use network upload to import PST files to Microsoft 365
- Use drive shipping to import PST files

Import SharePoint data to Microsoft 365

You can also import files and documents to SharePoint sites and OneDrive accounts in your organization. For more information, see the following articles:

- Migrate to SharePoint Online
- Introducing the SharePoint Migration Tool
- Migrate to SharePoint Online using PowerShell
- Migrate your file share content to SharePoint Online using the Azure Data Box

Frequently asked questions

For frequently asked questions about importing PST files, including duplicate handling, message properties, and common issues, see FAQ about importing PST files.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/11/2026

Use network upload to import your organization's PST files to Microsoft 365

> *Microsoft Purview service description*

> **Note**
>
> This article is for administrators. Are you trying to import PST files to your own mailbox? See Import email, contacts, and calendar from an Outlook .pst file.

Use the following step-by-step instructions to bulk-import multiple PST files to Microsoft 365 mailboxes by using network upload. For frequently asked questions about using network upload to bulk-import PST files to Microsoft 365 mailboxes, see FAQs for using network upload to import PST files. To import PST files by physically shipping a hard drive to Microsoft, see Use drive shipping to import PST files.

- Step 1: Copy the SAS URL and download AzCopy
- Step 2: Upload your PST files to Microsoft 365
- Step 3 (Optional): View a list of the uploaded PST files
- Step 4: Create the PST Import mapping file
- Step 5: Create a PST Import job
- Step 6: Filter data and start the PST Import job

You only need to perform Step 1 once to import PST files to Microsoft 365 mailboxes. After you perform Step 1, follow Step 2 through Step 6 each time you want to upload and import a batch of PST files.

Before you import PST files

Before you begin, review the requirements, limits, and planning guidance for importing PST files. The following prerequisites are specific to the network upload method:

- The only supported method for importing PST files to Microsoft 365 is to use the AzCopy tool, as described in this article. You can't use the Azure Storage Explorer to upload PST files directly to the Azure Storage area.
- This procedure involves copying and saving a copy of a URL that contains an access key. You use this information in Step 2 to upload your PST files, and in Step 3 if you want to view a list of the PST files uploaded to Microsoft 365. Take precautions to protect this URL like you would protect passwords or other security-related information. For example, you might save it to a password-protected Microsoft Word document or to an encrypted USB drive. See the More information section for an example of this combined URL and key.
- For a high-level overview of the PST Import process, see How the import process works section in this article.

Step 1: Copy the SAS URL and download AzCopy

First, download the AzCopy tool. Use this tool in Step 2 to upload PST files to Microsoft 365. Also, copy the SAS URL for your organization. This URL combines the network URL for the Azure Storage location in the Microsoft cloud for your organization and a Shared Access Signature (SAS) key. This key gives you the necessary permissions to upload PST files to an Azure Storage location. Protect the SAS URL. It's unique to your organization and you use it in Step 2.

> **Important**
>
> To import PST files by using the network upload method and command syntax documented in this article, use the version of AzCopy that you can download in step 6b in the following procedure. You can also download that same version of AzCopy from here. Using a different version of AzCopy isn't supported.

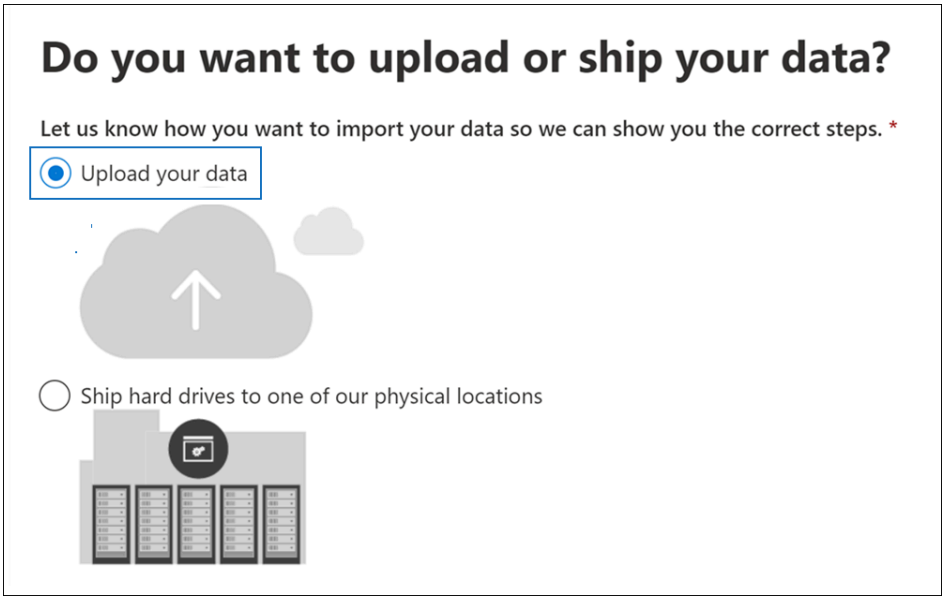
1. Sign in to the Microsoft Purview portal by using the credentials for an administrator account in your organization.
2. In the left navigation, select **Data lifecycle management** > **Microsoft 365** > **Import**.

   > **Note**
   >
   > You must be assigned the appropriate permissions to access the **Import** page in the Microsoft Purview portal. See the **Before you begin** section for more information.

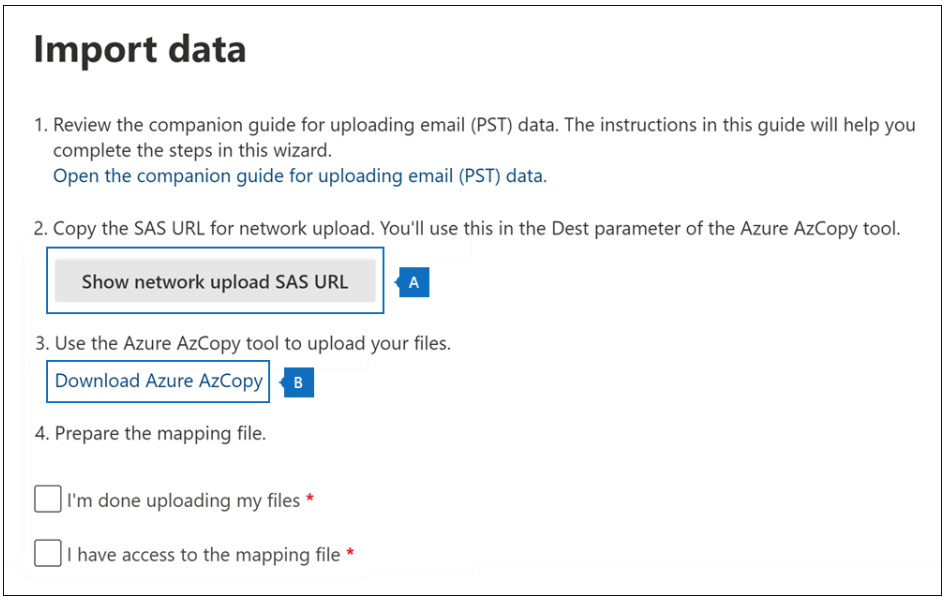
3. On the **Import** tab, select **+ New import job**.

   The import job wizard is displayed.

4. Type a name for the PST import job, and then select **Next**. Use lowercase letters, numbers, hyphens, and underscores. You can't use uppercase letters or include spaces in the name.
5. On the **Do you want to upload or ship data?** page, select **Upload your data** and then select **Next**.

6. On the **Import data** page, complete the following steps:



a. In step 2, select **Show network upload SAS URL**. After the SAS URL is displayed, select **Copy to clipboard** and then paste it and save it to a file.

b. In step 3, select **Download Azure AzCopy** to download the AzCopy tool to your local computer. This version of AzCopy is just an executable file, so there's nothing to install.

> ⓘ Note
>
> You can leave the **Import data** page open in case you need to copy the SAS URL again or select **Cancel** to close it.

Step 2: Upload your PST files to Microsoft 365

Now you're ready to use the AzCopy tool to upload PST files to Microsoft 365. This tool uploads and stores PST files in a Microsoft-provided Azure Storage location in the Microsoft cloud. As previously explained, the Azure Storage location that you upload your PST files to is located in the same regional Microsoft datacenter where your organization is located. To complete this step, the PST files must be located in a file share or file server in your organization or in an Azure Storage location managed by your organization. The PST storage location is known as the source location in this procedure. Each time you run the AzCopy tool, you can specify a different source location.

> ⓘ Note
>
> As previously stated, each PST file that you upload to the Azure Storage location should be no larger than 20 GB. PST files larger than 20 GB might impact the performance of the PST import process that you start in Step 6. Also, each PST file must have a unique name.

1. Open a Command Prompt on your local computer.
2. Go to the directory where you downloaded the azcopy.exe file in Step 1.
3. Run the following command to upload the PST files to Microsoft 365.

PowerShell

```
azcopy.exe copy "<Source location of PST files>" "<SAS URL>"
```

> **Important**
>
> You can specify a directory or an Azure Storage location as the source location in the previous command; you can't specify an individual PST file. All PST files in the source location are uploaded.

The following table describes the azcopy.exe fields and their required values. The information you obtained in the previous step is used in the values for these fields.

Expand table

Field	Description

Source	The first field specifies the source directory in your organization that contains the PST files that you want to upload Alternatively, you can specify an Azure Storage location as the source location of the PST files to upload.
	Be sure to surround the value of this field with double-quotation marks (" ").
	
	Examples:
	`"\\FILESERVER01\PSTs"`
	Or
	`"https://storageaccountid.blob.core.windows.net/PSTs?sp=racwdl&st=2021-09-21T07:25:53Z&se=2021-09-21T15:25:5 04&sr=c&sig=xxxxxx"`

Destination	Specifies the SAS URL that you obtained in Step 1.
	Be sure to surround the value of this parameter with double-quotation marks (" ").
	
	Note: If you use the SAS URL in a script or batch file, watch out for certain characters that need to be escaped. For e change `%` to `%%` and change `&` to `^&`.
	
	Tip: (Optional) You can specify a subfolder in the Azure Storage location to upload the PST files to. You do this by a location (after "ingestiondata") in the SAS URL. The first example doesn't specify a subfolder. That means the PST fil the root (named *ingestiondata*) of the Azure Storage location. The second example uploads the PST files to a subfol in the root of the Azure Storage location.
	
	Examples:
	`"https://3c3e5952a2764023ad14984.blob.core.windows.net/ingestiondata?sv=2012-02-12&se=9999-12- 31T23%3A59%3A59Z&sr=c&si=IngestionSasForAzCopy201601121920498117&sig=Vt5S4hVzlzMcBkuH8bH711atBffc`
	Or
	`"https://3c3e5952a2764023ad14984.blob.core.windows.net/ingestiondata/PSTFiles?sv=2012-02-12&se=9999-12- 31T23%3A59%3A59Z&sr=c&si=IngestionSasForAzCopy201601121920498117&sig=Vt5S4hVzlzMcBkuH8bH711atBffc`

`--recursive`	This optional flag specifies the recursive mode so that the AzCopy tool copies PST files that are located in subfolder directory that is specified by the source field. The default value for this flag is `true`. **Note:** If you include this flag, PST files in subfolders have a different file pathname in the Azure Storage location aft You have to specify the exact file pathname in the CSV file that you create in Step 4.
`--s2s-preserve-access-tier`	This optional flag is only required when the source location is a general-purpose v2 Azure Storage location that sup For the PST Import scenario, there's no need to preserve the access tier when you copy PST files from your Azure St Microsoft-provided Azure Storage location. In this case, you can include this flag and use a value of `false`. You don flag when copy PST files from a classic Azure Storage account, which doesn't support access tiers.

For more information about the **azcopy.exe copy** command, see azcopy copy.

Here are examples of the syntax for the AzCopy tool using actual values for each parameter.

Example 1

This example shows a source directory located on a file server or local computer.

PowerShell

```
azcopy.exe copy "\\FILESERVER1\PSTs"
"https://3c3e5952a2764023ad14984.blob.core.windows.net/ingestiondata?sv=2012-02-
```

```
12&amp;se=9999-12-
31T23%3A59%3A59Z&amp;sr=c&amp;si=IngestionSasForAzCopy2016011219204981l7&amp;sig=Vt5S4
hVzlzMcBkuH8bH711atBffdrOS72TlV1mNdORg%3D"
```

Example 2

This example shows a source directory located in a classic Azure Storage account with subdirectories.

PowerShell

```
azcopy.exe copy "https://storageaccountid.blob.core.windows.net/PSTs?
sp=racwdl&st=2021-09-21T07:25:53Z&se=2021-09-21T15:25:53Z&sv=2020-08-
04&sr=c&sig=xxxxxx"
"https://3c3e5952a2764023ad14984.blob.core.windows.net/ingestiondata?sv=2012-02-
12&amp;se=9999-12-
31T23%3A59%3A59Z&amp;sr=c&amp;si=IngestionSasForAzCopy2016011219204981l7&amp;sig=Vt5S4
hVzlzMcBkuH8bH711atBffdrOS72TlV1mNdORg%3D" --recursive
```

Example 3

This example shows a source directory located in a general-purpose v2 Azure Storage account. Access tiers aren't preserved when you upload the PST files.

PowerShell

```
azcopy.exe copy "https://storageaccountid.blob.core.windows.net/PSTs?
sp=racwdl&st=2021-09-21T07:25:53Z&se=2021-09-21T15:25:53Z&sv=2020-08-
04&sr=c&sig=xxxxxx"
"https://3c3e5952a2764023ad14984.blob.core.windows.net/ingestiondata?sv=2012-02-
12&amp;se=9999-12-
31T23%3A59%3A59Z&amp;sr=c&amp;si=IngestionSasForAzCopy2016011219204981l7&amp;sig=Vt5S4
hVzlzMcBkuH8bH711atBffdrOS72TlV1mNdORg%3D" --s2s-preserve-access-tier=false
```

When you run the command, status messages show the progress of uploading the PST files. A final status message shows the total number of files that you successfully uploaded.

> **Tip**
>
> After you successfully run the **azcopy.exe copy** command and verify that all the parameters are correct, save a copy of the command line syntax to the same secured file where you copied the information you obtained in Step 1. Then you can copy and paste this command in a Command Prompt each time that you want to run the AzCopy tool to upload PST files to Microsoft 365. The only value you might have to change is for the source field. This change depends on the source directory where the PST files are located.

(Optional) Step 3: View a list of the uploaded PST files

As an optional step, you can install and use the Microsoft Azure Storage Explorer (a free, open-source tool) to view the list of PST files that you uploaded to the Azure blob. Two good reasons to do this are:

- Verify that PST files from the shared folder or file server in your organization were successfully uploaded to the Azure blob.
- Verify the filename (and the subfolder pathname if you included one) for each PST file uploaded to the Azure blob. This verification is helpful when you're creating the PST mapping file in the next step because you need to specify both the folder pathname and

filename for each PST file. Verifying these names can help reduce potential errors in your PST mapping file.

The Azure Storage Explorer standalone application is generally available. You can download the latest version using the link in the following procedure.

> **Important**
>
> You can't use the Azure Storage Explorer to upload or modify PST files. The only supported method for importing PST files is to use AzCopy. Also, you can't delete PST files that you upload to the Azure blob. If you try to delete a PST file, you receive an error about not having the required permissions. All PST files are automatically deleted from your Azure storage area. If there are no import jobs in progress, all PST files in the **ingestiondata** container are deleted 30 days after the most recent import job was created.

To install the Azure Storage Explorer and connect to your Azure Storage area:

1. Download and install the Microsoft Azure Storage Explorer tool.
2. Start the Microsoft Azure Storage Explorer.
3. On the **Select Resource** page in the **Connect to Azure Storage** dialog, select **Blob container**.
4. On the **Select Authentication Method** page, select the **Shared access signature (SAS)** option, then select **Next**.
5. On the **Enter Connection Info** page, paste the SAS URL that you obtained in Step 1 into the box under **Blob container SAS URL**, then select **Next**. After you paste the SAS URL, the box under **Display name** is autopopulated with **ingestiondata**.
6. On the **Summary** page, review the connection information, then select **Connect**.

   The **ingestiondata** container opens. It contains the PST files that you uploaded in Step 2. The **ingestiondata** container is located under **Storage Accounts** > **(Attached Containers)** > **Blob Containers**.
7. When you're finished using the Microsoft Azure Storage Explorer, right-click **ingestiondata**, then select **Detach** to disconnect from your Azure Storage area. Otherwise, you receive an error the next time you try to attach.

Step 4: Create the PST Import mapping file

After you upload the PST files to the Azure Storage location for your organization, create a comma-separated value (CSV) file that specifies which user mailboxes the PST files go to. You'll submit this CSV file in the next step when you create a PST Import job.

1. Download a copy of the PST Import mapping file.
2. Open or save the CSV file to your local computer. The following example shows a completed PST Import mapping file (opened in NotePad). It's easier to use Microsoft Excel to edit the CSV file.

```
Console
Workload,FilePath,Name,Mailbox,IsArchive,TargetRootFolder,ContentCodePage,SPFileC
ontainer,SPManifestContainer,SPSiteUrl
Exchange,,annb.pst,annb@contoso.onmicrosoft.com,FALSE,/,,,,
```

```
Exchange,,annb_archive.pst,annb@contoso.onmicrosoft.com,TRUE,,,,,
Exchange,,donh.pst,donh@contoso.onmicrosoft.com,FALSE,/,,,,
Exchange,,donh_archive.pst,donh@contoso.onmicrosoft.com,TRUE,,,,,
Exchange,PSTFiles,pilarp.pst,pilarp@contoso.onmicrosoft.com,FALSE,/,,,,
Exchange,PSTFiles,pilarp_archive.pst,pilarp@contoso.onmicrosoft.com,TRUE,/Importe
dPst,,,,
Exchange,PSTFiles,tonyk.pst,tonyk@contoso.onmicrosoft.com,FALSE,,,,,
Exchange,PSTFiles,tonyk_archive.pst,tonyk@contoso.onmicrosoft.com,TRUE,/ImportedP
st,,,,
Exchange,PSTFiles,zrinkam.pst,zrinkam@contoso.onmicrosoft.com,FALSE,,,,,
Exchange,PSTFiles,zrinkam_archive.pst,zrinkam@contoso.onmicrosoft.com,TRUE,/Impor
tedPst,,,,
```

The first row, or header row, of the CSV file lists the parameters that the PST Import service uses to import the PST files to user mailboxes. Each parameter name is separated by a comma. Each row under the header row represents the parameter values for importing a PST file to a specific mailbox. You need a row for each PST file that you want to import to a user mailbox. You can have a maximum of 500 rows in the CSV mapping file. To import more than 500 PST files, create multiple mapping files and create multiple import jobs in Step 5.

> **Note**
>
> Don't change anything in the header row, including the SharePoint parameters; the PST Import process ignores them. Also, be sure to replace the placeholder data in the mapping file with your actual data.

3. Use the information in the following table to populate the CSV file with the required information.

Expand table

Parameter	Description	Example
`Workload`	Specifies the service that data is imported to. To import PST files to user mailboxes, use `Exchange`.	`Exchange`
`FilePath`	Specifies the folder location in the Azure Storage location that you uploaded the PST files to in Step 2. If you don't include an optional subfolder name in the SAS URL in the `/Dest:` parameter in Step 2, leave this parameter blank in the CSV file. If you include a subfolder name, specify it in this parameter (see the second example). The value for this parameter is case-sensitive. Either way, *don't* include "ingestiondata" in the value for the `FilePath` parameter. **Important:** The case for the file path name must be the same as the case you used if you include an optional subfolder name in the SAS URL in the destination field in Step 2. For example, if you use `PSTFiles` for the subfolder name and then use `pstfiles` in the `FilePath` parameter in CSV file, the import for the PST file fails. Be sure to use the same case in both instances.	(leave blank) Or `PSTFiles`

Parameter	Description	Example
`Name`	Specifies the name of the PST file that is imported to the user mailbox. The value for this parameter is case-sensitive. The file name of each PST file in the mapping file for an import job must be unique. **Important:** The case for the PST file name in the CSV file must be the same as the PST file that you uploaded to the Azure Storage location in Step 2. For example, if you use `annb.pst` in the `Name` parameter in the CSV file, but the name of the actual PST file is `AnnB.pst`, the import for that PST file fails. Be sure that the name of the PST in the CSV file uses the same case as the actual PST file.	`annb.pst`
`Mailbox`	Specifies the email address of the mailbox that the PST file is imported to. You can't specify a public folder or unified group because the PST Import Service doesn't support importing PST files to unified groups or to public folders. To import a PST file to an inactive mailbox, you have to specify the mailbox GUID for this parameter. To obtain this GUID, run the following PowerShell command in Exchange Online: `Get-Mailbox <identity of inactive mailbox> -InactiveMailboxOnly \FL ExchangeGuid`.	**Note:** Sometimes you have multiple mailboxes with the same email address, where one mailbox is an active mailbox and the other mailbox is in a soft-deleted (or inactive) state. In these situations, you have to specify the mailbox GUID to uniquely identify the mailbox to import the PST file to. To obtain this GUID for active mailboxes, run the following PowerShell command: `Get-Mailbox <identity of active mailbox> | FL ExchangeGuid`. To obtain the GUID for soft-deleted (or inactive) mailboxes, run this command: `Get-Mailbox <identity of soft-deleted or inactive mailbox> -SoftDeletedMailbox | FL ExchangeGuid`

`IsArchive`	Specifies whether to import the PST file to the user's archive mailbox. There are two options:
	
	FALSE: Imports the PST file to the user's primary mailbox.
	TRUE: Imports the PST file to the user's archive mailbox. This assumes that the user's archive mailbox is enabled.
	
	If you set this parameter to `TRUE` and the user's archive mailbox isn't enabled, the import for that user fails. If an import fails for one user (because their archive isn't enabled and this property is set to `TRUE`), the other users in the import job isn't affected.
	If you leave this parameter blank, the PST file is imported to the user's primary mailbox.
	
	Note: To import a PST file to a cloud-based archive mailbox for a user whose primary mailbox is on-premises, just specify `TRUE` for this parameter and specify the email address for the user's
		`FALSE`
		Or
		`TRUE`

Parameter	Description	Example
	on-premises mailbox for the `Mailbox` parameter. **Note:** When this parameter is `TRUE`, the PST file is imported into the user's main archive mailbox only, and a 100-GB-per-import limit applies. For more information, see Archive mailbox import limit.	

`TargetRootFolder`	Specifies the mailbox folder that the PST file is imported to.
	
	If you leave this parameter blank, the PST file is imported to a new folder named **Imported** at the root level of the mailbox (the same level as the Inbox folder and the other default mailbox folders).
	
	If you specify `/`, the folders and items in the PST file are imported to the top of the folder structure in the target mailbox or archive. If a folder exists in the target mailbox (for example, default folders such as Inbox, Sent Items, and Deleted Items), the items in that folder in the PST are merged into the existing folder in the target mailbox. For example, if the PST file contains an Inbox folder, items in that folder are imported to the Inbox folder in the target mailbox. New folders are created if they don't exist in the folder structure for the target mailbox.
	
	If you specify `/<foldername>`, items and folders in the PST file are imported to a folder named <*foldername*> . For example, if you use `/ImportedPst`, items would be imported to a folder named **ImportedPst**. This folder is located in the user's mailbox at the same level as the Inbox folder.
	
	Tip: Consider running a few test batches to experiment with this parameter so you can determine the best folder location to import PST files.
		(leave blank)
		Or
		`/`
		Or
		`/ImportedPst`

`ContentCodePage`	This optional parameter specifies a numeric value for the code page to use for importing PST files in the ANSI file format. This parameter is used for importing PST files from Chinese, Japanese, and Korean (CJK) organizations because these languages typically use a double byte character set (DBCS) for character encoding. If this parameter isn't used to import PST files for languages that use DBCS for mailbox folder names, the folder names are often garbled after they're imported. For a list of supported values to use for this parameter, see Code Page Identifiers.	(leave blank) Or `932` (which is the code page identifier for ANSI/OEM Japanese)

Parameter	Description	Example
	Note: As previously stated, this is an optional parameter and you don't have to include it in the CSV file. Or you can include it and leave the value blank for one or more rows.	
`SPFileContainer`	For PST Import, leave this parameter blank.	Not applicable
`SPManifestContainer`	For PST Import, leave this parameter blank.	Not applicable
`SPSiteUrl`	For PST Import, leave this parameter blank.	Not applicable

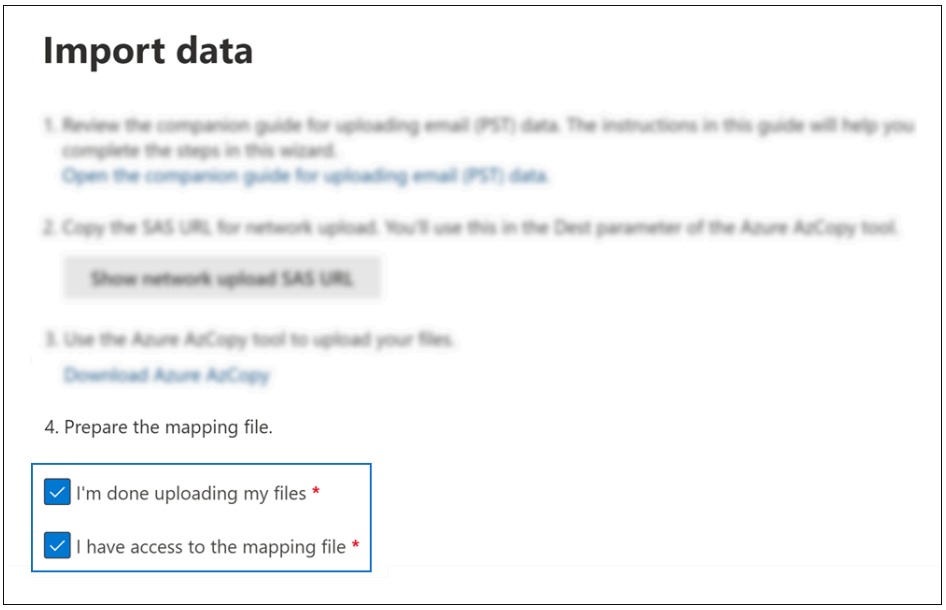
Step 5: Create a PST Import job

In this step, you create the PST import job in the Import service in Microsoft 365. As previously explained, you submit the PST import mapping file that you created in Step 4. After you create the job, Microsoft 365 analyzes the data in the PST files. Then, it gives you an opportunity to filter the data that actually gets imported to the mailboxes specified in the PST import mapping file (see Step 6).

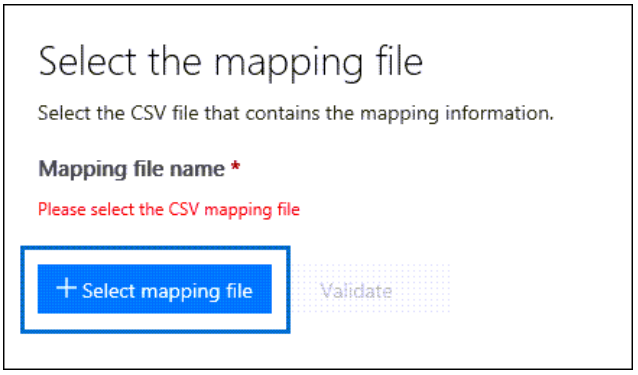
1. Go to https://purview.microsoft.com and sign in with the credentials for an administrator account in your organization.
2. In the left pane of the Microsoft Purview portal, select **Data lifecycle management** > **Microsoft 365** > **Import**.
3. On the **Import** tab, select **+ New import job**.

   > **Note**
   >
   > You need the appropriate permissions to access the **Import** page in the Microsoft Purview portal and create an import job. For more information, see the **Before you import PST files** section.

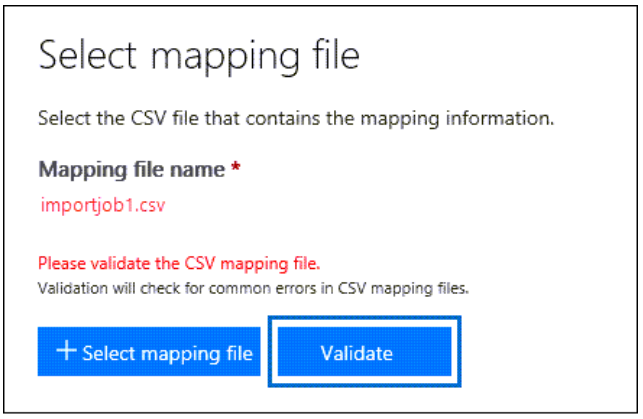
4. Type a name for the PST import job, and then select **Next**. Use lowercase letters, numbers, hyphens, and underscores. You can't use uppercase letters or include spaces in the name.
5. On the **Do you want to upload or ship data?** page, select **Upload your data** and then select **Next**.
6. In step 4 on the **Import data** page, select the **I'm done uploading my files** and **I have access to the mapping file** check boxes, and then select **Next**.



7. On the **Select the mapping file** page, select **Select mapping file** to submit the CSV mapping file that you created in Step 4.



8. After the name of the CSV file appears under **Mapping file name**, select **Validate** to check your CSV file for errors.



The CSV file must pass validation to create a PST import job. The file name turns green after successful validation. If validation fails, select the **View log** link. A validation error report opens, with an error message for each row in the file that failed.

> ⓘ Note
>
> As previously explained, a mapping file can have a maximum of 500 rows. Validation fails if the mapping file contains more than 500 rows. To import more than 500 PST files, create multiple mapping files and multiple import jobs.

9. After the mapping file is successfully validated, read the terms and conditions document, and then select the checkbox.

10. Select **Save** to submit the job, and then select **Close** after the job is successfully created.

    A status flyout page is displayed, with a status of **Analysis in progress**. The new import job is displayed in the list on the **Import PST files** page.

11. Select **Refresh** to update the status information displayed in the **Status** column. When the analysis completes and the data is ready to import, the status changes to **Analysis completed**.

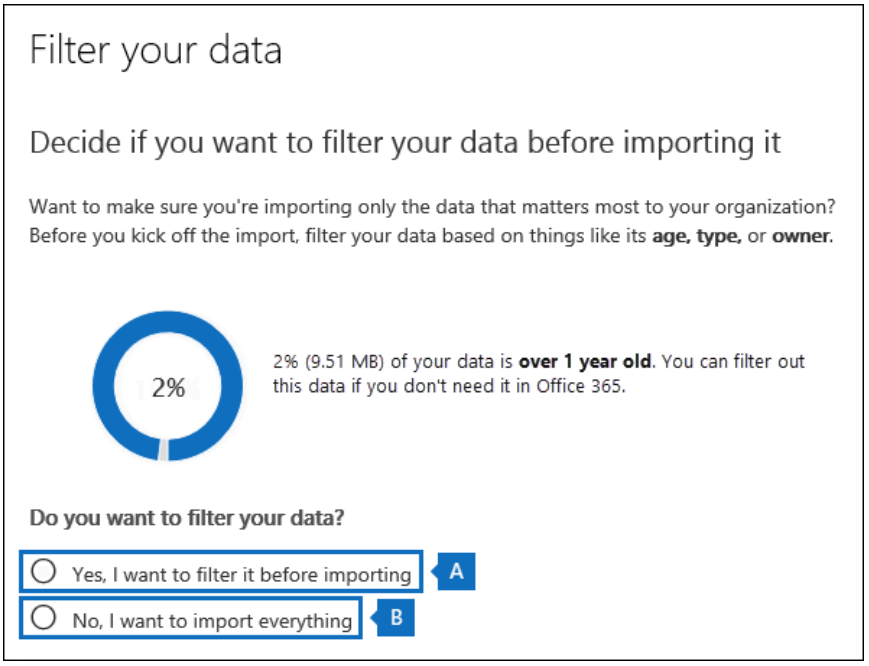
    You can select the import job to display the status flyout page, which contains more detailed information about the import job such as the status of each PST file listed in the mapping file.

Step 6: Filter data and start the PST Import job

After you create the import job in Step 5, Microsoft 365 safely and securely analyzes the data in the PST files. It identifies the age of the items and the different message types included in the PST files. When the analysis is complete and the data is ready to import, you can import all the data contained in the PST files or trim the data by setting filters that control what data gets imported.

1. On the **Import** tab in the Microsoft Purview portal, select the import jobs that you created in Step 5. Then select **Import to Microsoft 365**.

   The **Filter your data** page appears. It contains the data insights from the analysis of the PST files, including information about the age of the data. You can filter the data that you import or import all the data as is.



2. Do one of the following steps:

   a. To trim the data that you import, select **Yes, I want to filter it before importing**.

      For detailed step-by-step instructions about filtering the data in the PST files and then starting the import job, see Filter data when importing PST files to Microsoft 365.

      Or

b. To import all data in the PST files, select **No, I want to import everything,** and select **Next**.

3. If you chose to import all the data, select **Import data** to start the import job.

   The **Import PST files** page displays the status of the import job. Select **Refresh** to update the status information in the **Status** column. Select the import job to display the status flyout page, which shows status information about each PST file being imported.

More information

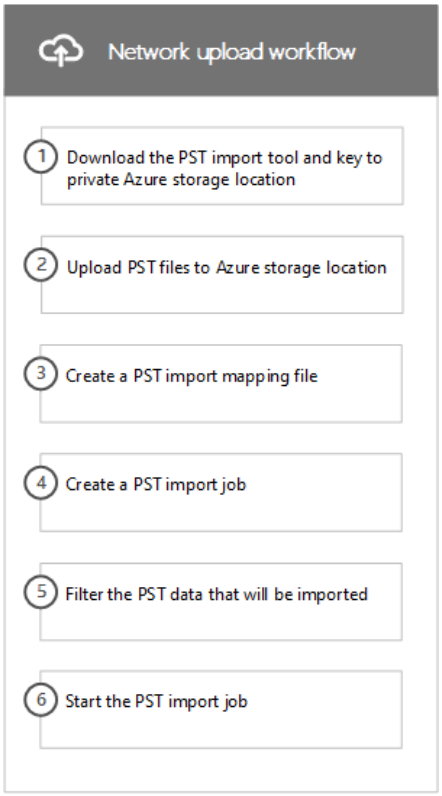
- Why import PST files to Microsoft 365?
  - It's a good way to import your organization's archival messaging data to Microsoft 365.
  - The data is available to the user from all devices because it's stored in the cloud.
  - It helps address compliance needs of your organization by letting you apply Microsoft Purview features to the data from the PST files that you imported. This benefit includes:
  - Enabling archive mailboxes and auto-expanding archiving to give users additional mailbox storage space to store the data that you imported.
  - Placing mailboxes on Litigation Hold to retain the data that you imported.
  - Using Microsoft eDiscovery tools to search the data that you imported.
  - Using Microsoft 365 retention policies to control how long the data that you imported will be retained, and what action to take after the retention period expires.
  - Searching the audit log for mailbox-related events that affect the data that you imported.
  - Importing data to inactive mailboxes to archive data for compliance purposes.
  - Using data loss prevention policies to prevent sensitive data from leaking outside your organization.
- As previously explained, the Microsoft 365 Import service turns on the retention hold setting (for an indefinite duration) after PST files are imported to a mailbox. This setting means the *RetentionHoldEnabled* property is set to **True** so that the retention policy assigned to the mailbox isn't processed. This setting gives the mailbox owner time to manage the newly imported messages by preventing a deletion or archive policy from deleting or archiving older messages. Here are some steps you can take to manage this retention hold:
  - After a certain time, turn off the retention hold by running the **Set-Mailbox -RetentionHoldEnabled $false** command. For instructions, see Place a mailbox on retention hold.
  - Configure the retention hold so that it's turned off on some date in the future. Run the **Set-Mailbox -EndDateForRetentionHold *date*** command. For example, assuming that today's date is June 1, 2016 and you want the retention hold turned off in 30 days, run the following command: **Set-Mailbox -EndDateForRetentionHold 7/1/2016**. In this scenario, leave the **RetentionHoldEnabled** property set to *True*. For more information, see Set-Mailbox.
  - Change the settings for the retention policy that's assigned to the mailbox so that older items that you imported aren't immediately deleted or moved to the user's archive

mailbox. For example, you could lengthen the retention age for a deletion or archive policy that's assigned to the mailbox. In this scenario, turn off the retention hold on the mailbox after you change the settings of the retention policy. For more information, see Set up an archive and deletion policy for mailboxes in your organization.

How the import process works

Use the network upload option and the Microsoft 365 Import service to bulk-import PST files to user mailboxes. Network upload means that you upload the PST files to a temporary storage area in the Microsoft cloud. Then the Microsoft 365 Import service copies the PST files from the storage area to the target user mailboxes.

Here's an illustration and description of the network upload process to import PST files to mailboxes in Microsoft 365.



1. **Download the PST import tool and key to private Azure Storage location:** Download the AzCopy command-line tool and an access key to upload the PST files to an Azure Storage location in the Microsoft cloud. Get these from the **Import** page in the Microsoft Purview portal. The key, called a secure access signature (SAS) key, gives you the necessary permissions to upload PST files to a private and secure Azure Storage location. This access key is unique to your organization and helps prevent unauthorized access to your PST files after you upload them to the Microsoft cloud. Importing PST files doesn't require your organization to have a separate Azure subscription.

2. **Upload the PST files to the Azure Storage location:** Use the azcopy.exe tool (downloaded in step 1) to upload and store your PST files in an Azure Storage location that resides in the same regional Microsoft datacenter where your organization is located. To upload them, the PST files that you want to import must be located in a file share or file server in your organization.

   You can optionally view the list of PST files after you upload them to the Azure Storage location.

3. **Create a PST import mapping file:** After you upload the PST files to the Azure Storage location, create a comma-separated value (CSV) file that specifies which user mailboxes the PST files go to. A PST file can go to a user's primary mailbox or their archive mailbox. The Microsoft 365 Import service uses the information in the CSV file to import the PST files.

4. **Create a PST import job:** Create a PST import job on the **Import PST files** page in the Microsoft Purview portal and submit the PST import mapping file you created. After you create the import job, Microsoft 365 analyzes the data in the PST files, then gives you an opportunity to set filters that control what data actually gets imported to the mailboxes specified in the PST import mapping file.

5. **Filter the PST data that you import to mailboxes:** After you create and start the import job, Microsoft 365 safely and securely analyzes the data in the PST files by identifying the age of the items and the different message types included in the PST files. When the analysis completes and the data is ready to import, you can import all the data contained in the PST files or trim the data by setting filters that control what data gets imported.

6. **Start the PST import job:** After you start the import job, Microsoft 365 uses the information in the PST import mapping file to import the PSTs files from the Azure Storage location to user mailboxes. The **Import PST files** page in the Microsoft Purview portal displays status information about the import job, including information about each PST file being imported. When the import job finishes, the status for the job is set to **Complete**.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/11/2026

Use drive shipping to import your organization's PST files

This article is for administrators. Are you trying to import PST files to your own mailbox? See Import email, contacts, and calendar from an Outlook .pst file.

Use the Microsoft 365 Import service and drive shipping to bulk-import PST files to user mailboxes. Drive shipping means that you copy the PST files to a hard disk drive and then physically ship the drive to Microsoft. When Microsoft receives your hard drive, data center personnel copy the data from the hard drive to a storage area in the Microsoft cloud. Then you can trim the PST data that's imported to the target mailboxes by setting filters that control what data gets imported. After you start the import job, the Import service imports the PST data from the storage area to user mailboxes. Using drive shipping to import PST files to user mailboxes is one way to migrate your organization's email to Microsoft 365.

Here are the steps required to use drive shipping to import PST files to Microsoft 365 mailboxes:

- Step 1: Download the PST Import tool
- Step 2: Copy the PST files to the hard drive
- Step 3: Create the PST Import mapping file
- Step 4: Create a PST Import job
- Step 5: Ship the hard drive to Microsoft
- Step 6: Filter data and start the PST Import job

> **Important**
>
> You perform Step 1 once to download the import tool. After you perform these steps, follow Step 2 through Step 6 each time you want to ship a hard drive to Microsoft.

For frequently asked questions about using drive shipping to import PST files to Microsoft 365, see FAQs for using drive shipping to import PST files.

Before you import PST files

Before you begin, review the requirements, limits, and planning guidance for importing PST files. The following prerequisites are specific to the drive shipping method:

- You need to store the PST files on a file server or shared folder in your organization. In Step 2, you run the Azure Import Export tool (WAImportExport.exe) to copy the PST files from this file server or shared folder to the hard drive.

- The Microsoft 365 Import service supports only 2.5-inch solid-state drives (SSDs) or 2.5-inch or 3.5-inch SATA II/III internal hard drives. You can use hard drives up to 10 TB. For import jobs, the service processes only the first data volume on the hard drive. You must format the data volume with NTFS. When you copy data to a hard drive, you can attach it directly by using a 2.5-inch SSD or 2.5-inch or 3.5-inch SATA II/III connector, or you can attach it externally by using an external 2.5-inch SSD or 2.5-inch or 3.5-inch SATA II/III USB adapter.

  > **ⓘ Important**
  >
  > The Microsoft 365 Import service doesn't support external hard drives that come with a built-in USB adapter. Additionally, the disk inside the casing of an external hard drive can't be used. Don't ship external hard drives.

- You must encrypt the hard drive that you copy the PST files to by using BitLocker. The WAImportExport.exe tool that you run in Step 2 helps you set up BitLocker. It also generates a BitLocker encryption key that Microsoft data center personnel use to access the drive to upload the PST files to the Azure Storage area in the Microsoft cloud.

- You or your organization must have an account with FedEx or DHL.

  - Organizations in the United States, Brazil, and Europe must have FedEx accounts.

  - Organizations in East Asia, Southeast Asia, Japan, Republic of Korea, and Australia must have DHL accounts.

    Microsoft uses (and charges) this account to return the hard drive back to you.

- The hard drive that you ship to Microsoft might cross international borders. In this case, you're responsible for ensuring that the hard drive and the data it contains are imported and exported in accordance with the applicable laws. Before shipping a hard drive, check with your advisors to verify that your drive and data can legally be shipped to the identified Microsoft data center. This check helps ensure that it reaches Microsoft in a timely manner.

- This procedure involves copying and saving a BitLocker encryption key. Take precautions to protect these keys like you protect passwords or other security-related information. For example, you might save them to a password-protected Microsoft Word document or save them to an encrypted USB drive. See the More information section for an example of these keys.

Step 1: Download the PST Import tool

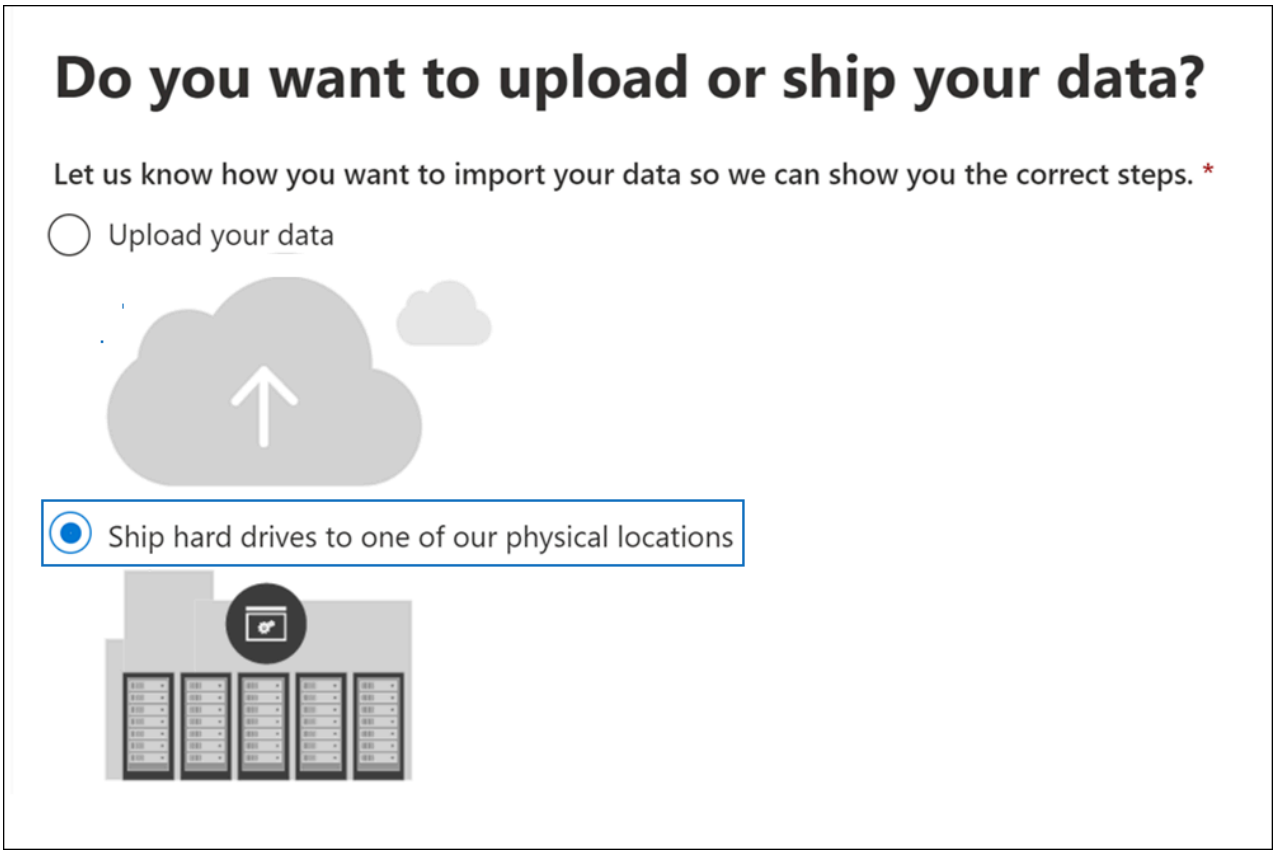
The first step is to download the tool. You use this tool in Step 2 to copy PST files to the hard drive.

> **Important**
>
> To successfully import PST files by using the drive shipping method, use Azure Import/Export tool version 1 (WAimportExportV1). Version 2 of the Azure Import/Export tool isn't supported. Using it results in incorrectly preparing the hard drive for the import job. Be sure to download the Azure Import/Export tool from the Microsoft Purview portal by following the procedures in this step.

1. Go to https://purview.microsoft.com and sign in by using the credentials for an administrator account in your organization.
2. In the left navigation pane of the Microsoft Purview portal, select **Data lifecycle management** > **Microsoft 365** > **Import**.

   > **Note**
   >
   > As previously stated, you need the appropriate permissions to access the **Import** page in the Microsoft Purview portal.

3. On the **Import** tab, select **+ New import job**.
4. In the import job wizard, type a name for the PST import job, and then select **Next**. Use lowercase letters, numbers, hyphens, and underscores. You can't use uppercase letters or include spaces in the name.
5. On the **Choose import job type** page, select **Ship hard drives to one of our physical locations** and then select **Next**.



6. On the **Import data** page, complete the following steps:

   Download the Azure Import/Export tool to download and install the Azure Import/Export (version 1) tool.

   - In the pop-up window, select **Save** > **Save as** to save the WaImportExportV1.zip file to a folder on your local computer.
   - Extract the WaImportExportV1.zip file.

7. Select **Cancel** to close the wizard.

   You return to the **Import** page in the Microsoft Purview portal when you create the import job in Step 4.

Step 2: Copy the PST files to the hard drive

In this step, use the WAImportExport.exe tool to copy PST files to the hard drive. This tool encrypts the hard drive by using BitLocker, copies the PSTs to the hard drive, and creates a journal file that stores information about the copy process. To complete this step, you need to locate the PST files in a file share or file server in your organization. This location is known as the source directory in the following procedure.

Each PST file that you copy to the hard drive should be no larger than 20 GB. PST files larger than 20 GB might affect the performance of the PST import process that you start in Step 6.

> **Important**
>
> After you run the WAImportExport.exe tool the first time for a hard drive, use a different syntax each time after that. Step 4 explains this syntax to copy PST files to the hard drive.

1. Open a Command Prompt on your local computer.

   > **Tip**
   >
   > If you run the command prompt as an administrator (by selecting **Run as administrator** when you open it), error messages display in the command prompt window. This information can help you resolve issues running the WAImportExport.exe tool.

2. Go to the directory where you installed the WAImportExport.exe tool in Step 1.

3. Run the following command the first time that you use the WAImportExport.exe to copy PST files to a hard drive.

   PowerShell

   ```
   WAImportExport.exe PrepImport /j:<Name of journal file> /t:<Drive letter> /id:
   <Name of session> /srcdir:<Location of PST files> /dstdir:<PST file path>
   /blobtype:BlockBlob /encrypt /logdir:<Log file location>
   ```

   The following table describes the parameters and their required values.

   Expand table

Parameter	Description	Example
`/j:`	Specifies the name of the journal file.The tool saves this file to the same folder where the WAImportExport.exe tool is located. Each hard drive you ship to Microsoft must have one journal file. Every time you run the WAImportTool.exe to copy PST files to a hard drive, the tool appends information to the journal file for that drive. Microsoft data center personnel use the information in the journal file to	`/j:PSTHDD1.jrn`

Parameter	Description	Example
	associate the hard drive with the import job that you create in Step 4, and to upload the PST files to the Azure Storage area in the Microsoft cloud.	
`/t:`	Specifies the drive letter of the hard drive when it's connected to your local computer.	`/t:h`
`/id:`	Specifies the name of the copy session. A session is defined as each time you run the WAImportExport.exe tool to copy files to the hard drive. The PST files are copied to a folder named with the session name specified by this parameter.	`/id:driveship1`
`/srcdir:`	Specifies the source directory in your organization that contains the PST files that the tool copies during the session. Surround the value of this parameter with double-quotation marks (" ").	`/srcdir:"\\FILESERVER01\PSTs"`
`/dstdir:`	Specifies the destination directory in the Azure Storage area in the Microsoft cloud where the PSTs are uploaded. Use the value `ingestiondata/`. Surround the value of this parameter with double-quotation marks (" "). Optionally, you can also add an extra file path to the value of this parameter. For example, you can use the file path of the source directory on the hard drive (converted to a URL format), which is specified in the `/srcdir:` parameter. For example, `\\FILESERVER01\PSTs` is changed to `FILESERVER01/PSTs`. In this case, you still must include `ingestiondata` in the file path. So in this example, the value for the `/dstdir:` parameter would be `"ingestiondata/FILESERVER01/PSTs"`.	`/dstdir:"ingestiondata/"` Or `/dstdir:"ingestiondata/FILESERVER01/PSTs"`

Parameter	Description	Example
	One reason to add the additional file path is if you have PST files with the same filename. > [!NOTE]> If you include the optional pathname, the namespace for a PST file after it's uploaded to the Azure Storage area includes the pathname and the name of the PST file; for example, `FILESERVER01/PSTs/annb.pst`. If you don't include a pathname, the namespace is only the PST filename; for example `annb.pst`.	
`/blobtype:`	Specifies the type of blobs in the Azure Storage area to import the PST files to. For importing PST files, use the value **BlockBlob**. This parameter is required.	`/blobtype:BlockBlob`
`/encrypt`	This switch turns on BitLocker for the hard drive. This parameter is required the first time you run the WAImportExport.exe tool. The BitLocker encryption key is copied to the journal file and the log file that is created if you use the `/logfile:` parameter. As previously explained, the journal file is saved to the same folder where the WAImportExport.exe tool is located.	`/encrypt`
`/logdir:`	This optional parameter specifies a folder to save log files to. If not specified, the log files are saved to the same folder where the WAImportExport.exe tool is located. Surround the value of this parameter with double-quotation marks (" ").	`/logdir:"c:\users\admin\desktop\PstImportLogs"`

Here's an example of the syntax for the WAImportExport.exe tool using actual values for each parameter:

PowerShell

```
WAImportExport.exe PrepImport /j:PSTHDD1.jrn /t:f /id:driveship1
/srcdir:"\\FILESERVER01\PSTs" /dstdir:"ingestiondata/" blobtype:BlockBlob
/encrypt /logdir:"c:\users\admin\desktop\PstImportLogs"
```

After you run the command, status messages display that show the progress of copying the PST files to the hard drive. A final status message shows the total number of files that you successfully copied.

4. Run this command each subsequent time you run the WAImportExport.ext tool to copy PST files to the same hard drive.

PowerShell

```
WAImportExport.exe PrepImport /j:<Name of journal file> /id:<Name of new
session> /srcdir:<Location of PST files> /dstdir:<PST file path>
/blobtype:BlockBlob
```

Here's an example of the syntax for running subsequent sessions to copy PST files to the same hard drive.

PowerShell

```
WAImportExport.exe PrepImport /j:PSTHDD1.jrn /id:driveship2
/srcdir:"\\FILESERVER01\PSTs\SecondBatch" /dstdir:"ingestiondata/"
/blobtype:BlockBlob
```

Step 3: Create the PST Import mapping file

After Microsoft data center personnel upload the PST files from the hard drive to the Azure Storage area, the Import service uses the information in the PST Import mapping file, which is a comma-separated value (CSV) file, to know which user mailboxes to import the PST files to. You submit this CSV file in the next step when you create a PST Import job.

1. Download a copy of the PST Import mapping file.

2. Open or save the CSV file to your local computer. The following example shows a completed PST Import mapping file (opened in NotePad). It's easier to use Microsoft Excel to edit the CSV file.

text

```
Workload,FilePath,Name,Mailbox,IsArchive,TargetRootFolder,ContentCodePage,SPFi
leContainer,SPManifestContainer,SPSiteUrl
Exchange,FILESERVER01/PSTs,annb.pst,annb@contoso.onmicrosoft.com,FALSE,/,,,,
```

```
Exchange,FILESERVER01/PSTs,annb_archive.pst,annb@contoso.onmicrosoft.com,TRUE,
/ImportedPst,,,,
Exchange,FILESERVER01/PSTs,donh.pst,donh@contoso.onmicrosoft.com,FALSE,/,,,,
Exchange,FILESERVER01/PSTs,donh_archive.pst,donh@contoso.onmicrosoft.com,TRUE,
/ImportedPst,,,,
Exchange,FILESERVER01/PSTs,pilarp.pst,pilarp@contoso.onmicrosoft.com,FALSE,/,,
,,
Exchange,FILESERVER01/PSTs,pilarp_archive.pst,pilarp@contoso.onmicrosoft.com,T
RUE,/ImportedPst,,,,
Exchange,,tonyk.pst,tonyk@contoso.onmicrosoft.com,FALSE,/,,,,
Exchange,,tonyk_archive.pst,tonyk@contoso.onmicrosoft.com,TRUE,,,,,
Exchange,,zrinkam.pst,zrinkam@contoso.onmicrosoft.com,FALSE,/,,,,
Exchange,,zrinkam_archive.pst,zrinkam@contoso.onmicrosoft.com,TRUE,,,,,
```

The first row, or header row, of the CSV file lists the parameters that the PST Import service uses to import the PST files to user mailboxes. Each parameter name is separated by a comma. Each row under the header row represents the parameter values for importing a PST file to a specific mailbox. You need a row for each PST file that you copied to the hard drive. Be sure to replace the placeholder data in the mapping file with your actual data.

> **Note**
>
> Don't change anything in the header row, including the SharePoint parameters; the PST Import process ignores them.

3. Use the information in the following table to populate the CSV file with the required information.

Expand table

Parameter	Description	Example
`Workload`	Specifies the service that data is imported to.To import PST files to user mailboxes, use `Exchange`.	`Exchange`
`FilePath`	Specifies the folder location in the Azure Storage area that PST files are copied to when the hard drive is shipped to Microsoft. What you add in this column in the CSV file depends on what you specified in for the `/dstdir:` parameter in the previous step. If you have subfolders on the source location, then the value in the `FilePath` parameter must contain the relative path for the subfolder; for example, /folder1/user1/.	(leave blank) Or `FILESERVER01/PSTs`

Parameter	Description	Example
	If you used `/dstdir:"ingestiondata/"`, then leave this parameter blank in the CSV file. If you included an optional pathname for the value of the `/dstdir:` parameter (for example, `/dstdir:"ingestiondata/FILESERVER01/PSTs"`, then use that pathname (not including "ingestiondata") for this parameter in the CSV file. The value for this parameter is case-sensitive. Either way, *don't* include "ingestiondata" in the value for the `FilePath` parameter. Leave this parameter blank or specify only the optional pathname. > [!IMPORTANT]> The case for the file path name must be the same case that you specified in the `/dstdir:` parameter in the previous step. For example, if you used `"ingestiondata/FILESERVER01/PSTs"` for the subfolder name in the previous step, but then used `fileserver01/psts` in the `FilePath` parameter in CSV file, the import for the PST file fails. Be sure to use the same case in both instances.	
`Name`	Specifies the name of the PST file that is imported to the user mailbox. The value for this parameter is case-sensitive. > [!IMPORTANT]> The case for the PST file name in the CSV file must be the same as the PST file that was uploaded to the Azure Storage location in Step 2. For example, if you use `annb.pst` in the `Name` parameter in the CSV file, but the name of the actual PST file is `AnnB.pst`, the import for that PST file fails. Be sure that the name of the PST in the CSV file uses the same case as the actual PST file.	`annb.pst`

`Mailbox`	Specifies the email address of the mailbox that the PST file is imported to. You can't specify a public folder because the PST Import Service doesn't support importing PST files to public folders.
	To import a PST file to an inactive mailbox, you have to specify the mailbox GUID for this parameter. To obtain this GUID, run the following PowerShell command in Exchange
		`annb@contoso.onmicrosoft.com`
		Or
		`2d7a87fe-d6a2-40cc-8aff-1ebea80d4ae7`

Parameter	Description	Example
	Online: `Get-Mailbox <identity of inactive mailbox> -InactiveMailboxOnly | FL Guid` > [!NOTE]> Sometimes, you have multiple mailboxes with the same email address, where one mailbox is an active mailbox and the other mailbox is in a soft-deleted (or inactive) state. In these situations, you have to specify the mailbox GUID to uniquely identify the mailbox to import the PST file to. To obtain this GUID for active mailboxes, run the following PowerShell command: `Get-Mailbox <identity of active mailbox> | FL Guid`. To obtain the GUID for soft-deleted (or inactive) mailboxes, run this command: `Get-Mailbox <identity of soft-deleted or inactive mailbox> -SoftDeletedMailbox | FL Guid`.	

`IsArchive`	Specifies whether to import the PST file to the user's archive mailbox. There are two options:
	FALSE Imports the PST file to the user's primary mailbox.
	TRUE Imports the PST file to the user's archive mailbox. This assumes that the user's archive mailbox is enabled. If you set this parameter to `TRUE` and the user's archive mailbox isn't enabled, the import for that user fails. If an import fails for one user (because their archive isn't enabled and this property is set to `TRUE`), the other users in the import job isn't affected.
	If you leave this parameter blank, the PST file is imported to the user's primary mailbox.
	Note: To import a PST file to a cloud-based archive mailbox for a user whose primary mailbox is on-premises, just specify `TRUE` for this parameter and specify the email address for the user's on-premises mailbox for the `Mailbox` parameter.
		`FALSE`
		Or
		`TRUE`

`TargetRootFolder`	Specifies the mailbox folder that the PST file is imported to.
	If you leave this parameter blank, the PST is imported to a new folder named **Imported** located at the root level of the mailbox (the same level as the Inbox folder and the other
		(leave blank)
		Or
		`/`
		Or
		`/ImportedPst`

Parameter	Description	Example
	default mailbox folders). If you specify `/`, items in the PST file are imported directly in to the user's Inbox folder. If you specify `/<foldername>`, items in the PST file are imported to a folder named *<foldername>*. For example, if you use `/ImportedPst`, items would be imported to a folder named **ImportedPst**. This folder is located in the user's mailbox at the same level as the Inbox folder.	
`ContentCodePage`	This optional parameter specifies a numeric value for the code page to use for importing PST files in the ANSI file format. This parameter is used for importing PST files from Chinese, Japanese, and Korean (CJK) organizations because these languages typically use a double byte character set (DBCS) for character encoding. If this parameter isn't used to import PST files for languages that use DBCS for mailbox folder names, the folder names are often garbled after they're imported. For a list of supported values to use for this parameter, see Code Page Identifiers. > [!NOTE]> As previously stated, this is an optional parameter and you don't have to include it in the CSV file. Or you can include it and leave the value blank for one or more rows.	(leave blank) Or `932` (which is the code page identifier for ANSI/OEM Japanese)
`SPFileContainer`	For PST Import, leave this parameter blank.	Not applicable
`SPManifestContainer`	For PST Import, leave this parameter blank.	Not applicable
`SPSiteUrl`	For PST Import, leave this parameter blank.	Not applicable

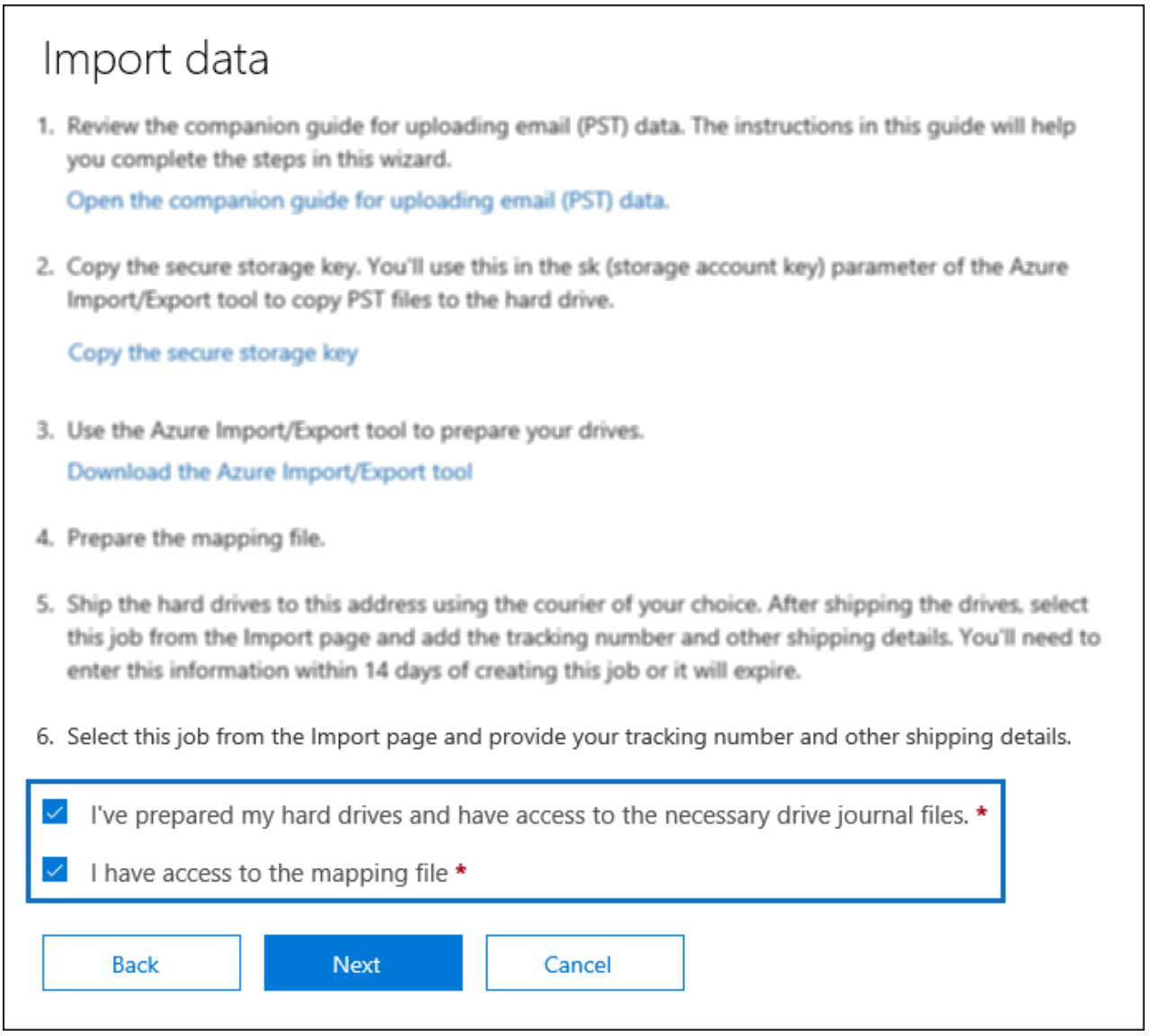
Step 4: Create a PST Import job

Next, create the PST Import job in the Import service in Microsoft 365. Submit the PST Import mapping file that you created in Step 3. After you create the job, the Import service uses the information in the mapping file to import the PST files to the specified user mailbox. The Import service imports the PST files after you copy the PST files from the hard drive to the Azure Storage area and you create and start the import job.

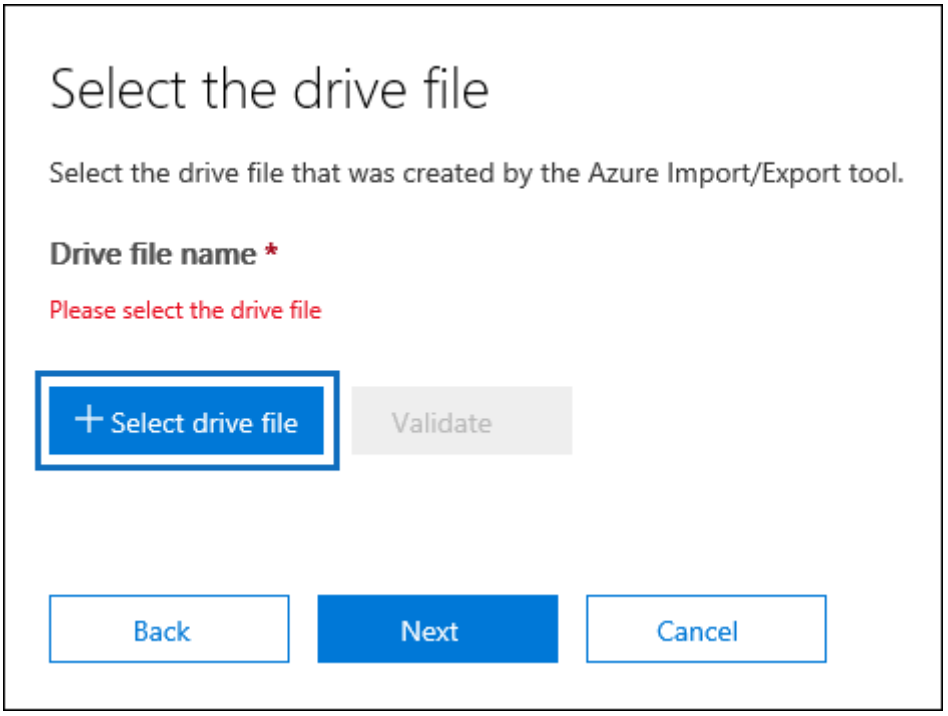
1. Go to https://purview.microsoft.com and sign in by using the credentials for an administrator account in your organization.
2. In the left navigation pane of the Microsoft Purview portal, select **Data lifecycle management** > **Microsoft 365** > **Import**.
3. On the **Import** tab, select **+ New import job**.

   > **Note**
   >
   > As previously stated, you need the appropriate permissions to access the **Import** page in the Microsoft Purview portal.

4. Type a name for the PST import job, then select **Next**. Use lowercase letters, numbers, hyphens, and underscores. You can't use uppercase letters or include spaces in the name.
5. On the **Choose import job type** page, select **Ship hard drives to one of our physical locations** and then select **Next**.
6. Select the **I've prepared my hard drives and have access to the necessary drive journal files** and **I have access to the mapping file** check boxes, then select **Next**.



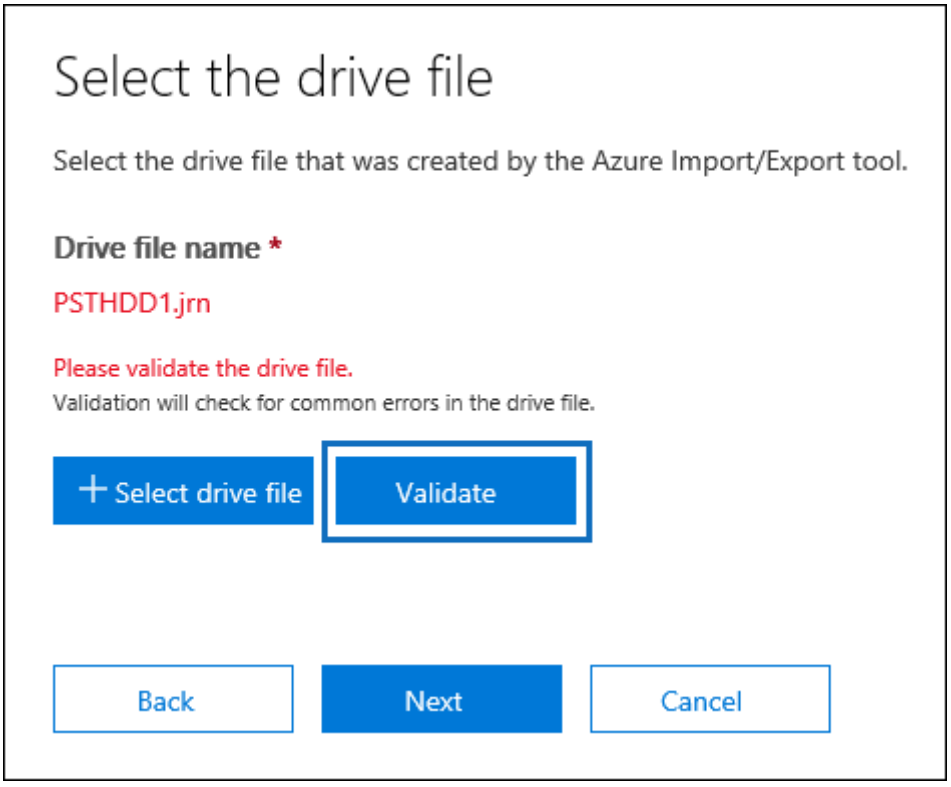
7. On the **Select the drive file** page, select **Select drive file**, then go to the same folder where the WAImportExport.exe tool is located. The journal file that you created in Step 2 is in this folder.



8. Select the journal file, such as `PSTHDD1.jrn`.

> **Tip**
>
> When you ran the WAImportExport.exe tool in Step 2, you specified the name of the journal file by using the `/j:` parameter.

9. After the name of the drive file appears under **Drive file name**, select **Validate** to check your drive file for errors.

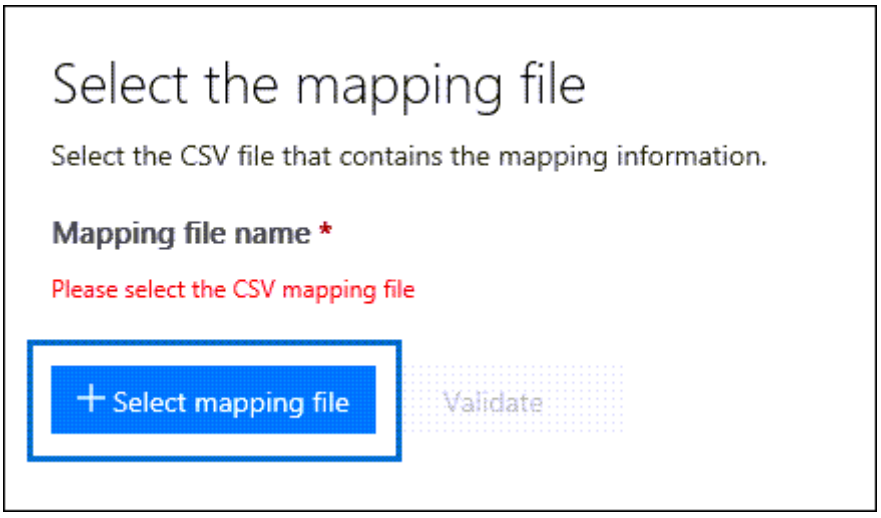


You must successfully validate the drive file to create a PST Import job. The file name changes to green after it's successfully validated. If the validation fails, select the **View log** link. A validation error report opens, with an error message that explains why the file failed.

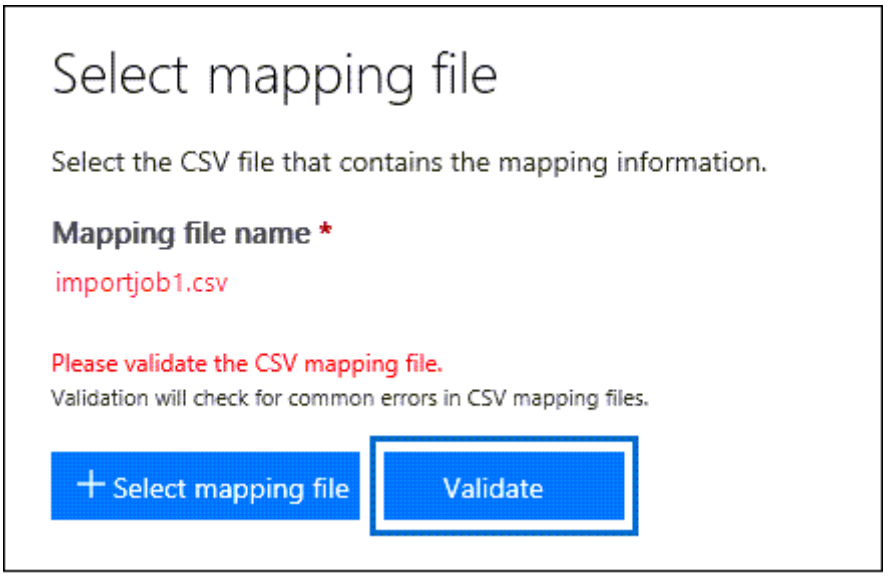
> **Note**
>
> You must add and validate a journal file for each hard drive you ship to Microsoft.

10. After adding and validating a journal file for each hard drive that you ship to Microsoft, select **Next**.

11. Select **+ Select mapping file** to submit the PST Import mapping file that you created in Step 3.



12. After the name of the CSV file appears under **Mapping file name**, select **Validate** to check your CSV file for errors.



    You must successfully validate the CSV file to create a PST Import job. The file name changes to green after it's successfully validated. If the validation fails, select the **View log** link. A validation error report opens, with an error message for each row in the file that failed.

13. After the PST mapping file is successfully validated, select **Next**.

14. On the **Provide contact information** page, type your contact information in the applicable boxes.

    The address for the Microsoft location that you ship your hard drives to is displayed. This address is autogenerated based on your Microsoft datacenter location. Copy this address to a file or take a screenshot.

15. Read the terms and conditions document, select the checkbox, then select **Save** to submit the import job.

When the import job is successfully created, a status page is displayed that explains the next steps of the drive shipping process.

16. On the **Import** tab, select **Refresh** to show the new drive shipping import job in the list of import jobs. The status is **Waiting for tracking number**. You can also select the import job to display the status flyout page, which contains more detailed information about the import job.

Step 5: Ship the hard drive to Microsoft

In this step, you ship the hard drive to Microsoft. You also provide the tracking number for the shipment and return shipment information for the drive shipping job. After Microsoft receives the drive, data center personnel take between 7 and 10 business days to upload your PST files to the Azure Storage area for your organization.

> **Note**
>
> If you don't provide the tracking number and return shipment information within 14 days of creating the import job, the import job expires. If this happens, you need to create a new drive shipping import job (see Step 4: Create a PST Import job) and re-submit the drive file and the PST import mapping file.

Ship the hard drive

Keep the following things in mind when you ship hard drives to Microsoft:

- Don't ship the SATA-to-USB adapter; you only need to ship the hard drive.
- Package the hard drive properly; for example, use an anti-static bag or bubble wrap.
- Use a delivery carrier of your choice to ship the hard drive to Microsoft.
- Ship the hard drive to the address for the Microsoft location that displays when you create the import job in Step 4. Be sure to include "Microsoft 365 Import Service" in the ship-to address.
- After you ship the hard drive, be sure to write down the name of the delivery carrier and the tracking number. You provide these in the next step.

Enter the tracking number and other shipping information

After you ship the hard drive to Microsoft, complete the following procedure on the Import service page.

1. Go to https://purview.microsoft.com and sign in by using the credentials for an administrator account in your organization.
2. In the left navigation pane of the Microsoft Purview portal, select **Data lifecycle management** > **Microsoft 365** > **Import**.
3. On the **Import** tab, select the job for the drive shipment that you want to enter the tracking number for.
4. On the status flyout page, select **Enter tracking number**.
5. Provide the following shipping information:
    a. **Delivery carrier** Type the name of the delivery carrier that you used to ship the hard drive to Microsoft.
    b. **Tracking number** Type the tracking number for the hard drive shipment.
    c. **Return carrier account number** Type your organization's account number for the carrier that listed under **Return carrier**. Microsoft uses (and charges) this account to ship your hard drive back to you. Organizations in the USA and Europe, must have an account with FedEx. Organizations in Asia and the rest of the world, must have an account with DHL.
6. Select **Save** to save this information for the import job.
    a. On the **Import** tab, select **Refresh** to update the information for your drive shipping import job. Notice that status is now set to **Drives in transit**.

Step 6: Filter data and start the PST Import job

After Microsoft receives your hard drive, the status for the import job on the **Import PST files** page changes to **Drives received**. Data center personnel use the information in the journal file to upload your PST files to the Azure Storage area for your organization. At this point, the status changes to **Import in-progress**. As previously stated, it takes between 7 and 10 business days after receiving your hard drive to upload the PST files.

After uploading the PST files to Azure, the status changes to **Analysis in progress**. This status indicates that Microsoft 365 is analyzing the data in the PST files (in a safe and secure manner) to identify the age of the items and the different message types included in the PST files. When the analysis completes and the data is ready to import, the status changes to **Analysis completed**. At this point, you can import all the data contained in the PST files or trim the data by setting filters that control what data gets imported.

1. Go to https://purview.microsoft.com and sign in by using the credentials for an administrator account in your organization.

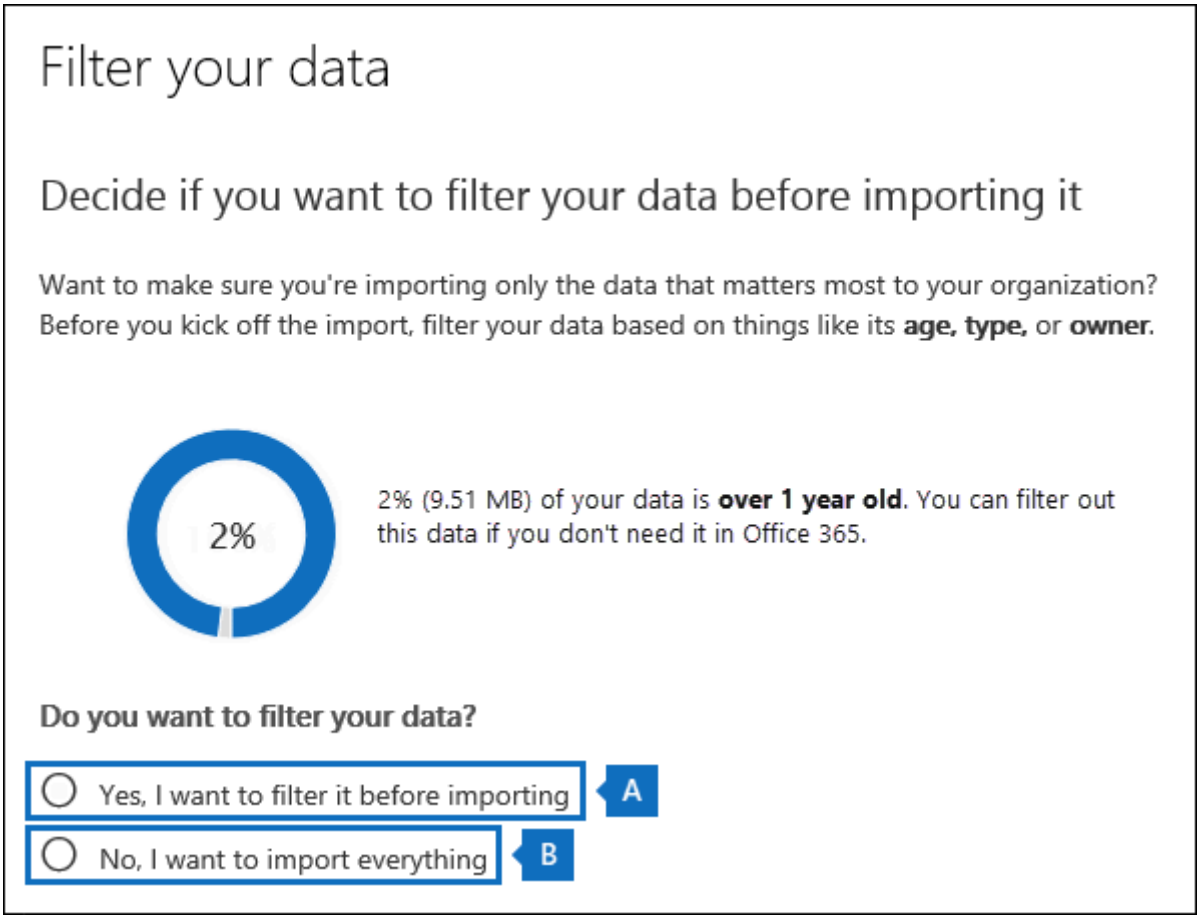
2. In the left navigation pane of the Microsoft Purview portal, select **Data lifecycle management** > **Microsoft 365** > **Import**.

3. On the **Import** tab, select the import job that you created in Step 4 and select **Import to Microsoft 365**.

   A flyout page appears with information about the PST files and other information about the import job.

4. Select **Import to Microsoft 365**.

5. The **Filter your data** page appears. It contains the data insights resulting from the analysis performed on the PST files by Microsoft 365, including information about the age of the data. At this point, you can filter the data that you want to import or import all the data as is.



6. Do one of the following steps:

   a. To trim the data that you import, select **Yes, I want to filter it before importing**.

      For detailed step-by-step instructions about filtering the data in the PST files and then starting the import job, see Filter data when importing PST files to Microsoft 365.

Or

b. To import all data in the PST files, select **No, I want to import everything,** and select **Next**.

7. If you chose to import all the data, select **Import data** to start the import job.

   The status of the import job appears on the **Import PST files** page. Select **Refresh** to update the status information that's displayed in the **Status** column. Select the import job to display the status flyout page, which displays status information about each PST file being imported. When the import completes and PST files are imported to user mailboxes, the status changes to **Completed**.

View a list of the PST files uploaded to Microsoft 365

Install and use Microsoft Azure Storage Explorer, a free, open-source tool, to view the list of PST files that Microsoft data center personnel uploaded to the Azure Storage area for your organization. Use this list to verify that PST files from the hard drives you sent to Microsoft were successfully uploaded to the Azure Storage area.

> **Important**
>
> You can't use Azure Storage Explorer to upload or modify PST files. The only supported method for importing PST files to Microsoft 365 is to use AzCopy. Also, you can't delete PST files that you upload to the Azure blob. If you try to delete a PST file, you receive an error about not having the required permissions. All PST files are automatically deleted from your Azure Storage area. If there are no import jobs in progress, the system deletes all PST files in the **ingestiondata** container 30 days after the most recent import job was created.

Perform the following steps to get the Shared Access Signature (SAS) URL for your organization. This URL combines the network URL for the Azure Storage location in the Microsoft cloud for your organization with an SAS key. This key provides you with the necessary permissions to access your organization's Azure Storage location.

To install Azure Storage Explorer and connect to your Azure Storage area:

1. Go to https://purview.microsoft.com and sign in by using the credentials for an administrator account in your organization.

2. In the left pane of the Microsoft Purview portal, select **Data lifecycle management** > **Microsoft 365** > **Import**.

3. On the **Import** tab, select **+ New import job**.

4. In the import job wizard, type a name for the PST import job, and then select **Next**. Use lowercase letters, numbers, hyphens, and underscores. You can't use uppercase letters or include spaces in the name.

5. On the **Choose import job type** page, select **Upload your data**, then select **Next**.

6. In step 2, select **Show network upload SAS URL**.

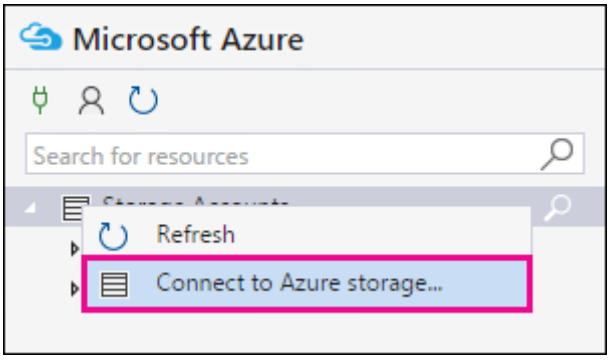
7. After the URL is displayed, copy it and save it to a file. Be sure to copy the entire URL.

   > **Important**
   >
   > Be sure to take precautions to protect the SAS URL. Anyone can use this URL to access the Azure storage area for your organization.

8. Select **Cancel** to close the import job wizard.

9. Download and install the Microsoft Azure Storage Explorer tool.

10. Start Microsoft Azure Storage Explorer, right-click **Storage Accounts** in the left pane, then select **Connect to Azure Storage**.



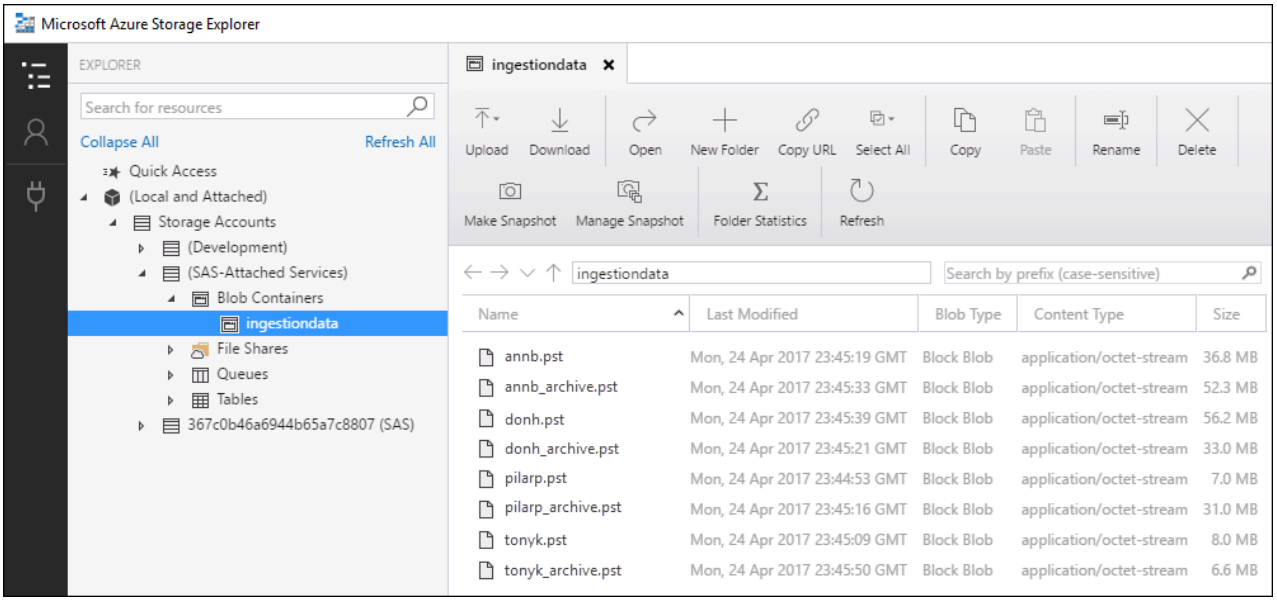
11. Select **Use a shared access signature (SAS) URI or connection string** and select **Next**.

12. Select **Use a SAS URI**, paste the SAS URL that you obtained in step 1 into the box under **URI**, then select **Next**.

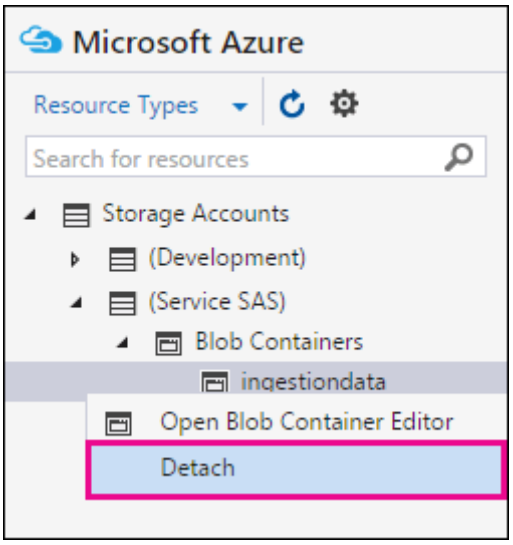
13. On the **Connection summary** page, review the connection information, then select **Connect**.

    The **ingestiondata** container opens. It contains the PST files from your hard drive. The **ingestiondata** container is located under **Storage Accounts** > **(SAS-Attached Services)** >

Blob Containers.



14. When you're finished using Microsoft Azure Storage Explorer, right-click **ingestiondata**, then select **Detach** to disconnect from your Azure Storage area. Otherwise, you receive an error the next time you try to attach.



Resolve common import errors

- **What happens if the import job fails because of errors in the PST Import CSV mapping file?** If an import job fails because of errors in the mapping file, you don't need to reship the hard drive to Microsoft to create an import job. That's because the PST files from the hard drive that you submitted for the drive shipping import job are already uploaded to the Azure Storage area for your organization. In this case, you only need to fix the errors in the PST Import CSV mapping file, then create a new "network upload" import job and submit the revised CSV mapping file. For more information, see Step 5: Create a PST Import job in Microsoft 365 and Step 6: Filter data and start the PST Import job.

> **Note**
>
> To help resolve errors in the PST Import CSV mapping file, use the Azure Storage Explorer tool to view the folder structure in the **ingestiondata** container for the PST files from your hard drive that you uploaded to the Azure storage area. Mapping file errors typically occur because of an incorrect value in the FilePath parameter. This parameter specifies the location of a PST file in the Azure storage area. See the description of the FilePath parameter in table in Step 3. As previously explained, the location of PST files in the Azure storage area is specified by the `/dstdir:` parameter when you run the WAImportExport.exe tool in Step 2.

More information

- Drive shipping is an effective way to import large amounts of archival messaging data to Microsoft 365 so you can take advantage of the compliance features that are available to your organization. After archival data is imported to user mailboxes, you can:
  - Enable archive mailboxes and auto-expanding archiving to give users more mailbox storage space for the data.
  - Place mailboxes on Litigation Hold to retain the data.
  - Use Microsoft eDiscovery tools to search the data.
  - Apply Microsoft 365 retention policies to control how long the data is retained, and what action to take after the retention period expires.
  - Search the audit log for events related to this data.
  - Import data to inactive mailboxes to archive data for compliance purposes.
  - Protect your organization against data loss of sensitive information.

- Here's an example of the secure storage account key and a BitLocker encryption key. This example also contains the syntax for the WAImportExport.exe command that you run to copy PST files to a hard drive. Be sure to take precautions to protect these just like you would protect passwords or other security-related information.

```text
Secure storage account key:

yaNIIs9Uy5g25Yoak+LlSHfqVBGOeNwjqtBEBGqRMoidq6/e5k/VPkjOXdDIXJHxHvNoNoFH5NcVUJ
XHwu9ZxQ==

BitLocker encryption key:

397386-221353-718905-535249-156728-127017-683716-083391
```

```
COMMAND SYNTAX

First time:

WAImportExport.exe PrepImport /j:<Name of journal file> /t:<Drive letter> /id:
<Name of session> /srcdir:<Location of PST files> /dstdir:<PST file path>
/blobtype:BlockBlob /encrypt /logdir:<Log file location>

Subsequent times:

WAImportExport.exe PrepImport /j:<Name of journal file> /id:<Name of new
session> /srcdir:<Location of PST files> /dstdir:<PST file path>
/blobtype:BlockBlob

EXAMPLES

First time:

WAImportExport.exe PrepImport /j:PSTHDD1.jrn /t:f /id:driveship1
/srcdir:"\\FILESERVER1\PSTs" /dstdir:"ingestiondata/"
/blobtype:BlockBlob /encrypt /logdir:"c:\users\admin\desktop\PstImportLogs"

Subsequent times:

WAImportExport.exe PrepImport /j:PSTHDD1.jrn /id:driveship2
/srcdir:"\\FILESERVER1\PSTs\SecondBatch" /dstdir:"ingestiondata/"
/blobtype:BlockBlob
```

- As previously explained, the Microsoft 365 Import service turns on the retention hold setting (for an indefinite duration) after PST files are imported to a mailbox. This setting means the *RetentionHoldEnabled* property is set to `True` so that the retention policy assigned to the mailbox isn't processed. This setting gives the mailbox owner time to manage the newly imported messages by preventing a deletion or archive policy from deleting or archiving older messages. Here are some steps you can take to manage this retention hold:
  - After a certain period of time, you can turn off the retention hold by running the `Set-Mailbox -RetentionHoldEnabled $false` command. For instructions, see Place a mailbox on retention hold.
  - You can configure the retention hold so that it's turned off on some date in the future. You do this by running the `Set-Mailbox -EndDateForRetentionHold <date>` command. For example, assuming that today's date is June 1, 2016 and you want the retention hold turned off in 30 days, you would run the following command: `Set-Mailbox -EndDateForRetentionHold 7/1/2016`. In this scenario, you would leave the *RetentionHoldEnabled* property set to *True*. For more information, see Set-Mailbox.
  - You can change the settings for the retention policy that's assigned to the mailbox so that older items that were imported aren't immediately deleted or moved to the user's archive mailbox. For example, you could lengthen the retention age for a deletion or

archive policy that's assigned to the mailbox. In this scenario, you turn off the retention hold on the mailbox after you change the settings of the retention policy. For more information, see Set up an archive and deletion policy for mailboxes in your organization.

Last updated on 04/06/2026

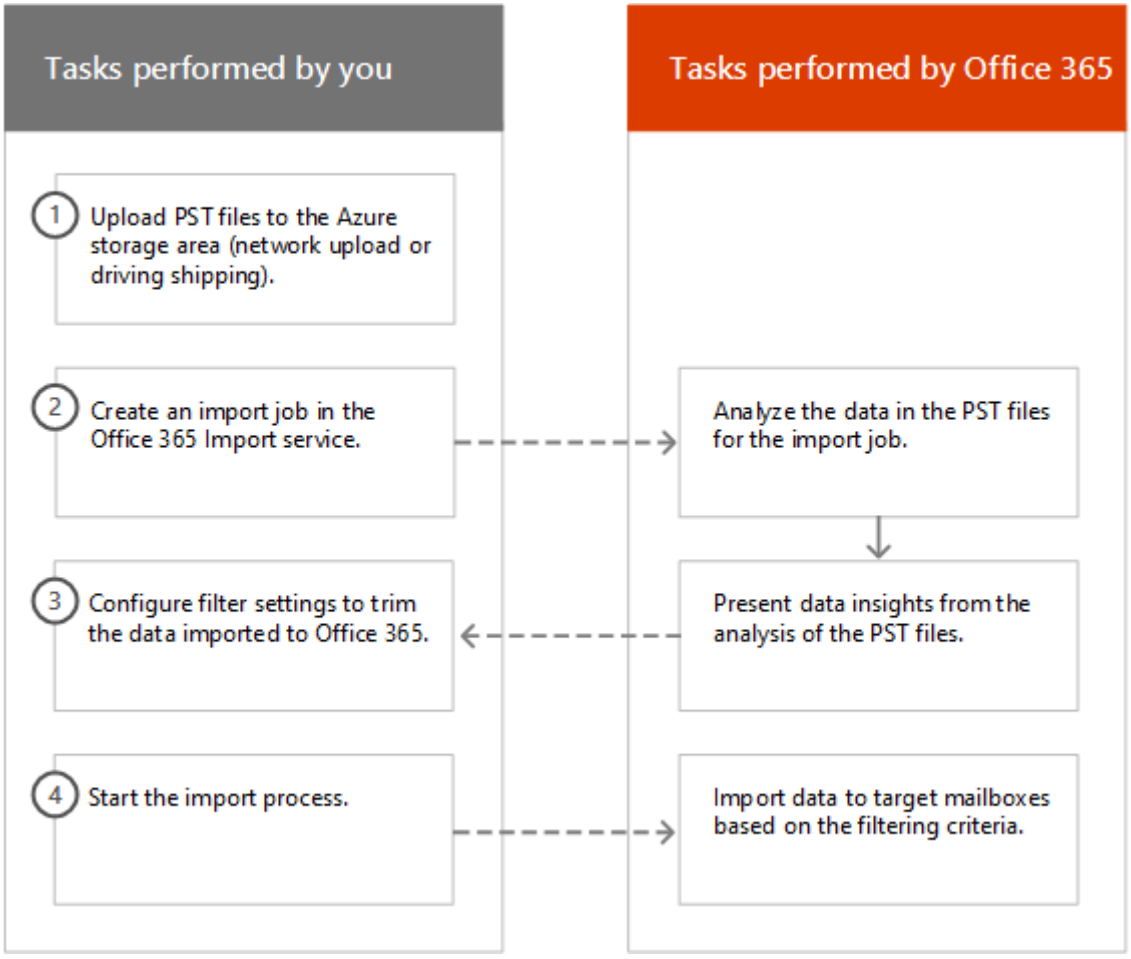
Filter data when importing PST files

> **Note**
>
> This article explains how to filter PST data during import. Before using this feature, create a PST import job by using network upload or drive shipping. For an overview of the PST import process, including requirements and limits, see Learn about importing PST files.

Use the new Intelligent Import feature in the Microsoft 365 Import service to filter the items in PST files that you import to the target mailboxes. Here's how it works:

- After you create and submit a PST import job, the process uploads PST files to an Azure storage area in the Microsoft cloud.
- Microsoft 365 analyzes the data in the PST files in a safe and secure manner. It identifies the age of the mailbox items and the different message types included in the PST files.
- When the analysis is complete and the data is ready to import, you can import all data in the PST files as is or trim the data by setting filters that control what data gets imported. For example, you can choose to:
  - Import only items of a certain age.
  - Import selected message types.
  - Exclude messages sent or received by specific people.
- After you configure the filter settings, Microsoft 365 imports only the data that meets the filtering criteria to the target mailboxes specified in the import job.

The following graphic shows the Intelligent Import process, and highlights the tasks you perform and the tasks performed by Microsoft 365.



Create a PST Import job

- The steps in this article assume that you created a PST import job in the Microsoft 365 Import service by using network upload or drive shipping. For step-by-step instructions, see one of the following articles:
  - Use network upload to import PST files to Microsoft 365
  - Use drive shipping to import PST files to Microsoft 365

- After you create an import job by using network upload, the status for the import job on the Import page in the Microsoft Purview portal is set to **Analysis in progress**, which means that Microsoft 365 is analyzing the data in the PST files that you uploaded. Select **Refresh** to update the status for the import job.

- For drive shipping import jobs, Microsoft 365 analyzes the data after Microsoft datacenter personnel receive your hard drive and upload the PST files to the Azure storage area for your organization.

- For answers to common questions about importing PST files, see FAQ about importing PST files.

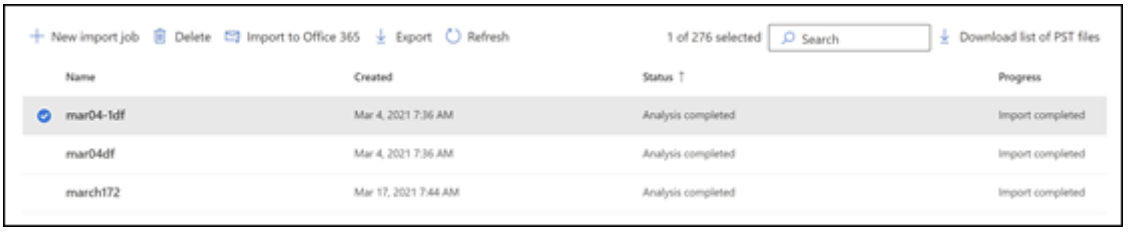
Filter data that gets imported to mailboxes

After you create a PST import job, follow these steps to filter the data before you import it to Microsoft 365.

1. Go to the Microsoft Purview portal and sign in by using the credentials for an administrator account in your organization.

2. In the left pane of the Microsoft Purview portal, select **Data lifecycle management** > **Microsoft 365** > **Import**.

   The **Import** tab lists the import jobs for your organization. The **Analysis completed** value in the **Status** column indicates the import jobs that Microsoft 365 analyzed and are ready for you to import.

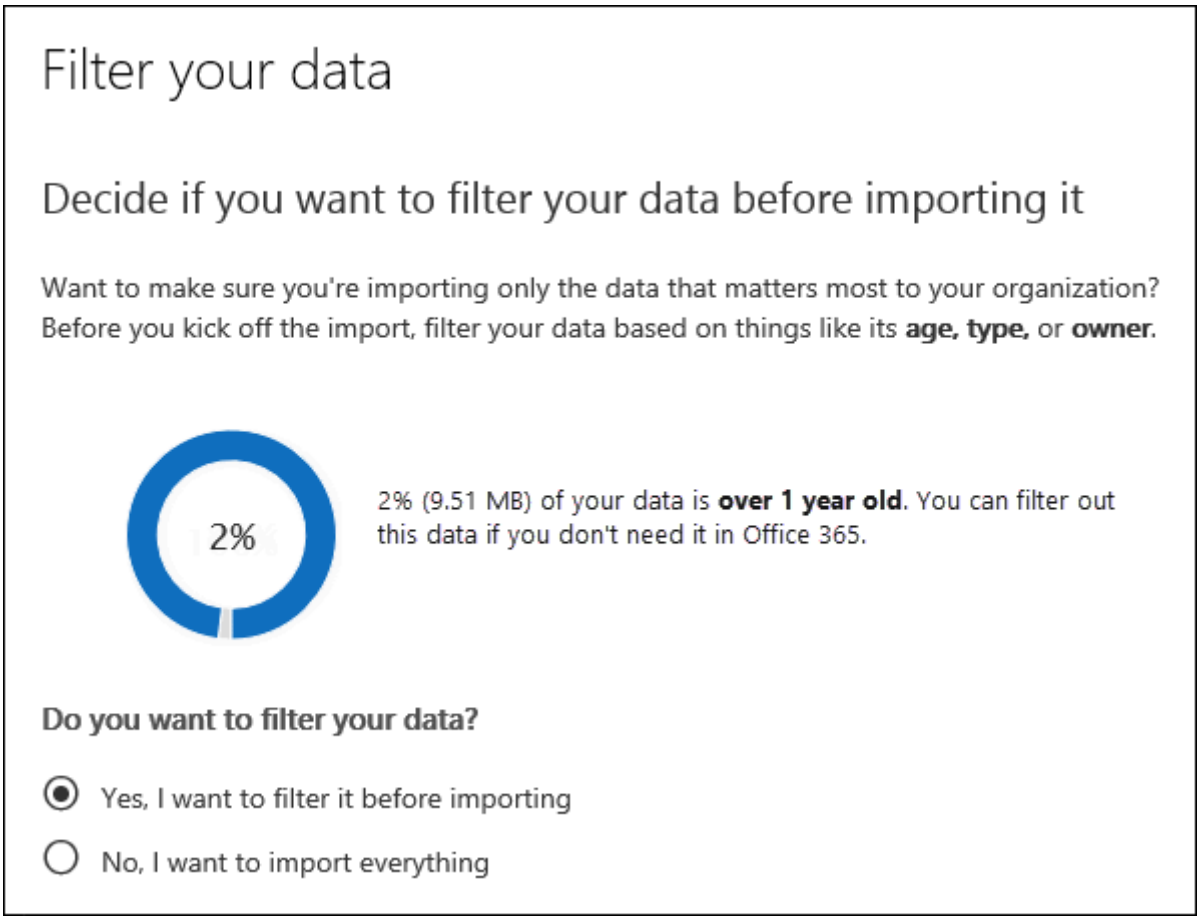


3. Select the import job that you want to complete and select **Import to Microsoft 365**.

   A flyout page is displayed with information about the PST files and other information about the import job.

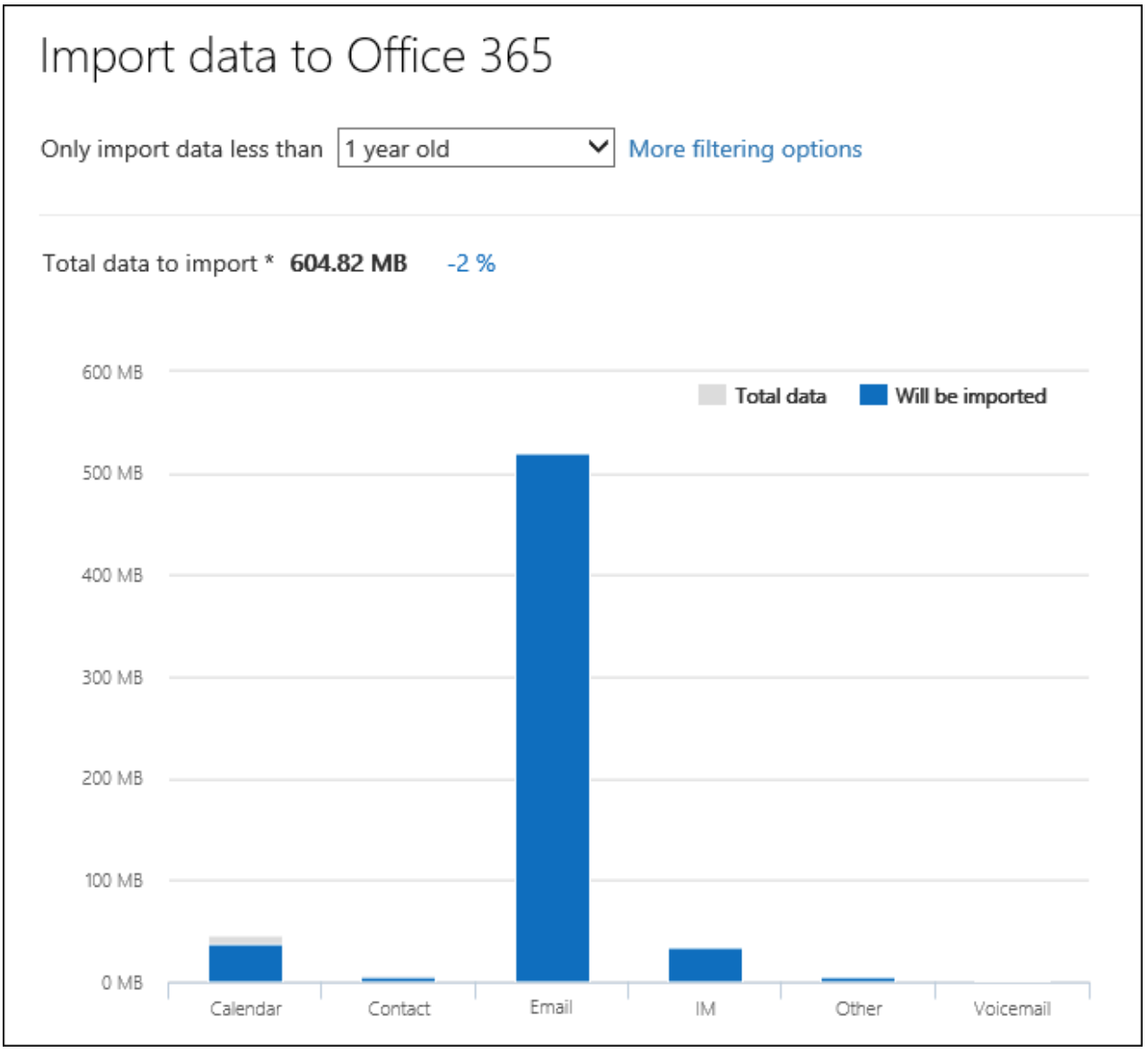
4. Select **Import to Microsoft 365**.

   The **Filter your data** page is displayed. It contains data insights about the data in the PST files for the import job, including information about the age of the data.



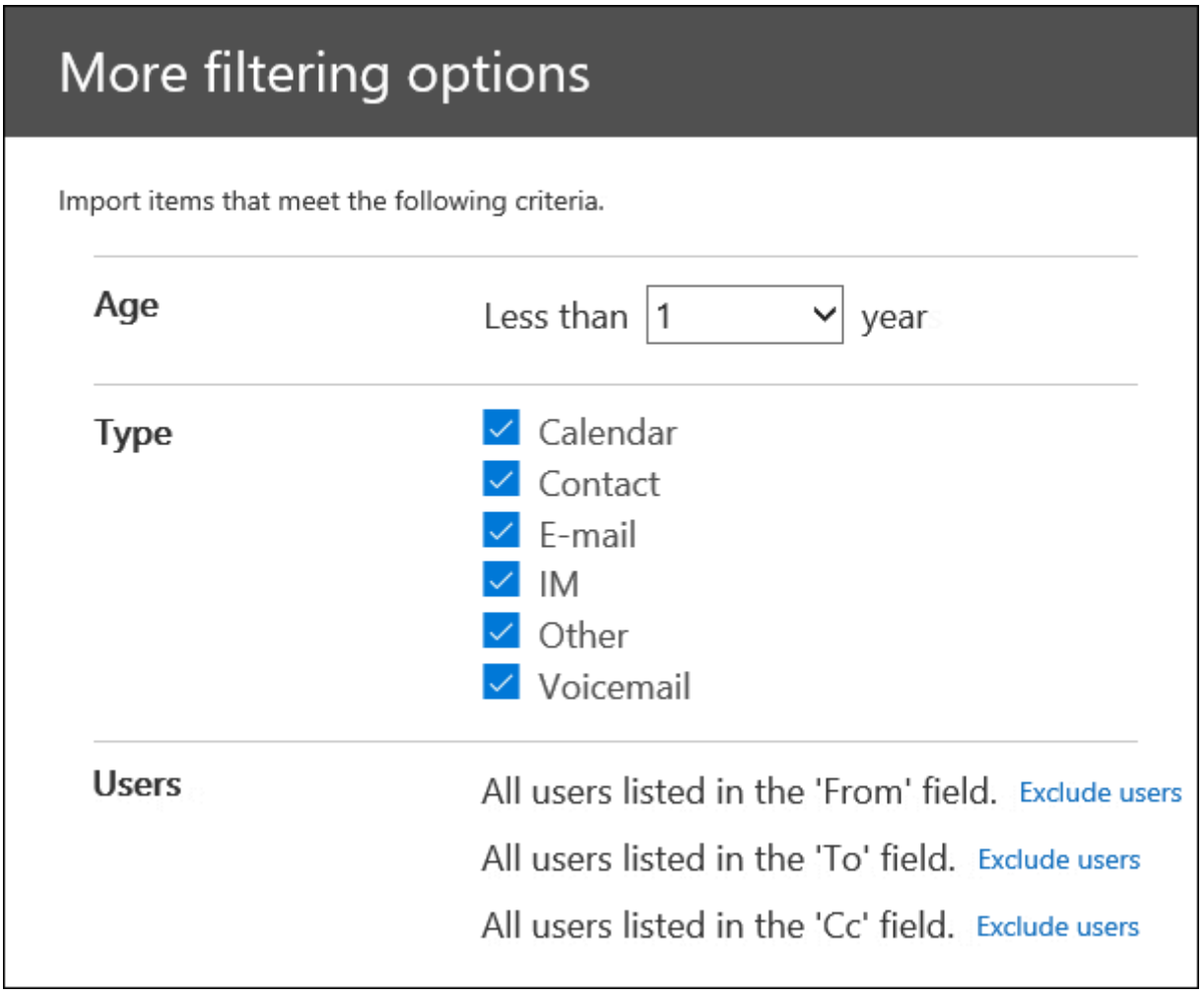
5. Based on whether you want to trim the data that's imported to Microsoft 365, under **Do you want to filter your data?**, do one of the following steps:
   a. Select **Yes, I want to filter it before importing** to trim the data that you import, and then select **Next**.

   The **Import data to Microsoft 365 page** is displayed with detailed data insights from the analysis that Microsoft 365 performed.



The graph on this page shows the amount of data that you import. It displays information about each message type found in the PST files. You can hover the cursor over each bar to display specific information about that message type. There's also a drop-down list with different age values based on the analysis of the PST files. When you select an age in the drop-down list, the graph is updated to show how much data is imported for the selected age.

a. To configure addition filters to reduce the amount of data that's imported, select **More filtering options**.



You can configure these filters:

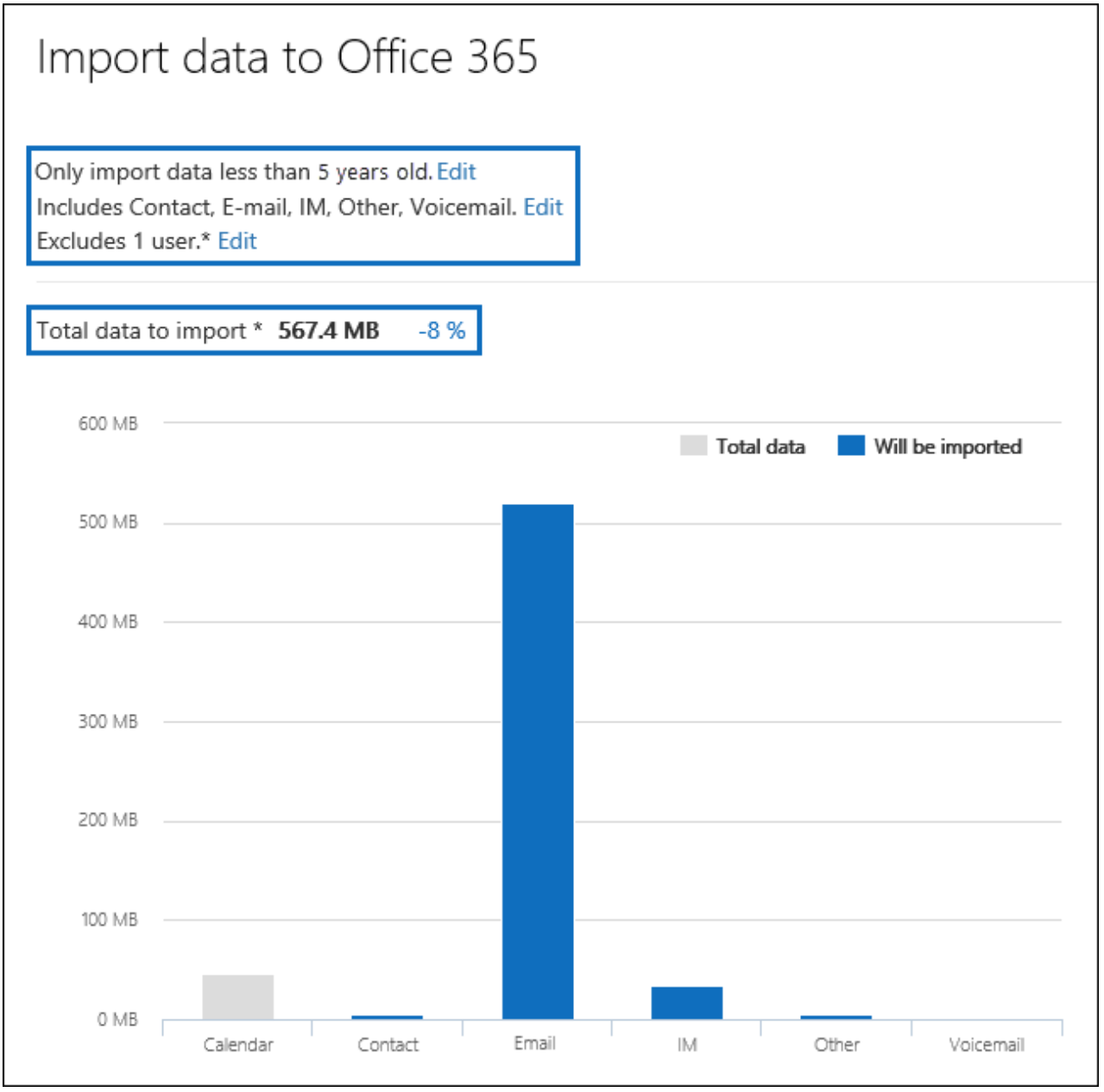
- **Age** - Select an age so only items that are newer than the specified age are imported. See the More information section for a description about how Microsoft 365 determines the age buckets for the **Age** filter.
- **Type** - This section shows all the message types that Microsoft 365 found in the PST files for the import job. You can clear a check box next to a message type that you want to exclude. You can't exclude the Other message type. See the More information section for a list of mailbox items that are included in the Other category.
- **Users** - You can exclude messages that are sent or received by specific people. To exclude people who appear in the From: field, To: field, or the Cc: field of messages, select **Exclude users** next to that recipient type. Type the email address (SMTP address) of the person, select **Add+** to add them to the list of excluded users for that recipient type, and then select **Save** to save the list of excluded users.

  > ⓘ Note

> Microsoft 365 doesn't show data insights that result from setting the **People** filter. However, if you set this filter to exclude messages sent or received by specific people, those messages are excluded during the actual import process.

a. Select **Apply** in the **More filtering options** flyout page to save your filter settings.

The data insights on the **Import data to Microsoft 365** page are updated based on your filter settings, including the total amount of data that you import based on the filter settings. A summary of the filter settings is also shown. You can select **Edit** next to a filter to change the setting if necessary.



a. Select **Next**.

A status page is displayed showing your filter settings. Again, you can edit any of the filter settings.

a. Select **Import data** to start the import. The total amount of data that you import is displayed.

Or

a. Select **No, I want to import everything** to import all data in the PST files to Microsoft 365, and then select **Next**.

b. On the **Import data to Microsoft 365** page, select **Import data** to start the import. The total amount of data that will be imported is displayed.

6. On the **Import** tab, select **Refresh**. The **Status** column displays the status for the import job.

7. Select the import job to display more detailed information, such as the status for each PST file and the filter settings that you configured.

More information

- How does Microsoft 365 determine the increments for the age filter? When Microsoft 365 analyzes a PST file, it looks at the sent or received timestamp of each item. If an item has both a sent and received timestamp, Microsoft 365 selects the oldest date. Then Microsoft 365 looks at the year value for that timestamp and compares it to the current date to determine the age of the item. The software uses these ages as the values in the drop-down list for the **Age** filter. For example, if a PST file has messages from 2016, 2015, and 2014, the values in the **Age** filter are **1 year**, **2 years**, and **3 years**.

- The following table lists the message types that the **Other** category includes in the **Type** filter on the **More options** flyout page (see Step 5b in the previous procedure). Currently, you can't exclude items in the **Other** category when you import PSTs to Microsoft 365.

Expand table

Message class ID	Mailbox items that use this message class
IPM.Activity	Journal entries
IPM.Document	Documents and files (not attached to an email message)
IPM.File	(same as IPM.Document)
IPM.Note.IMC.Notification	Reports sent by Internet Mail Connect, which is the Exchange Server gateway to the Internet
IPM.Note.Microsoft.Fax	Fax messages
IPM.Note.Rules.Oof.Template.Microsoft	Out-of-office autoreply messages
IPM.Note.Rules.ReplyTemplate.Microsoft	Replies sent by an inbox rule

Message class ID	Mailbox items that use this message class
IPM.OLE.Class	Exceptions for a recurring series
IPM.Recall.Report	Message recall reports
IPM.Remote	Remote mail messages
IPM.Report	Item status reports

Last updated on 04/06/2026

FAQ about importing PST files

This article is for administrators. Do you want to import PST files to your own mailbox? See Import email, contacts, and calendar from an Outlook .pst file

This article answers frequently asked questions about importing PST files to Microsoft 365 mailboxes. For requirements, limits, and planning guidance, see Learn about importing PST files.

For step-by-step instructions, see Use network upload to import PST files or Use drive shipping to import PST files. To learn about filtering imported data, see Filter data when importing PST files.

General

How does the PST import process handle duplicate email items?

The PST import process uses a SourceEntryId-based matching system to detect duplicate items. During import, each PST item's EntryId is stamped as a SourceEntryId property on the imported item in the target mailbox. The import process then checks for this property to determine whether an item was previously imported.

It's important to understand how this matching works in different scenarios:

- **Reimporting the same PST file to the same target folder**. Items imported during a previous import job already have a SourceEntryId stamp. The import process matches on this property and skips the item, preventing duplicates.
- **First-time import of a PST file**. Items that already exist natively in the target mailbox (items that weren't previously imported from a PST file) don't have a SourceEntryId property. Because no match is found, the import process copies the PST item to the mailbox, which can result in duplicate items.
- **Importing a different PST file that contains overlapping content**. SourceEntryId values differ between PST files. Even if a different PST file contains the same email content, the import process doesn't find a matching SourceEntryId and copies the item, which can result in duplicates.
- **Reimporting the same PST file to a different target folder**. If you specify a different target folder (using the TargetRootFolder property in the PST import mapping file) than the one

specified in a previous import job, all items in the PST file are reimported because the import process checks for matches only within the specified target folder.

The import process doesn't use content-based matching criteria such as message subject, date, body, or size to detect duplicates. The conflict resolution behavior (keeping the existing target item when a match is found) is built in to the import service and isn't configurable through the Microsoft Purview portal or the CSV mapping file.

Are message properties, such as when the message was sent or received, the list of recipients and other properties, preserved when PST files are imported to a Microsoft 365 mailbox?

Yes. The original message metadata isn't changed during the import process.

Are there any screen resolution requirements when I create an import job?

If the screen resolution on your local computer is set to 1366 x 768, you may have issues creating an import job. If so, we recommend that you change the screen resolution to a different setting and then create a new import job. We're working to fix this issue.

Network upload

After I upload my PST files to the Azure Storage area, how long are they kept in Azure before they're deleted?

When you use the network upload method to import PST files, you upload them to an Azure blob container named `ingestiondata`. If there are no import jobs in progress on the **Import PST files** page in the Microsoft Purview portal, then all PST files in the `ingestiondata` container in Azure are deleted 30 days after the most recent import job was created in the Microsoft Purview portal. That also means you have to create a new import job in the Microsoft Purview portal (described in Step 5 in the network upload instructions) within 30 days of uploading PST files to Azure.

This also means that after PST files are deleted from the Azure Storage area, they're no longer displayed in the list of files for a completed import job in the Microsoft Purview portal. Although

an import job might still be listed on the **Import PST files** page in the Microsoft Purview portal, the list of PST files might be empty when you view the details of older import jobs.

Drive shipping

After I ship my hard drive, how long does it take to get to the Microsoft data center?

That depends on a few things, such as your proximity to the Microsoft data center and what kind of shipping option you used to ship your hard drive (such as, next-day delivery, two-day delivery, or ground-delivery). With most shippers, you can use the tracking number to track the status of your delivery.

After Microsoft uploads my PST files to Azure, how long are they kept in Azure before they're deleted?

All PST files in the Azure Storage location for your organization (in blob container named `ingestiondata`), are deleted 30 days after the most recent import job was created on the **Import PST files** page in the Microsoft Purview portal.

This also means that after PST files are deleted from the Azure Storage area, they're no longer displayed in the list of files for a completed import job in the Microsoft Purview portal. Although an import job might still be listed on the **Import PST files** page in the Microsoft Purview portal, the list of PST files might be empty when you view the details of older import jobs.

Can Microsoft wipe my hard drive before they ship it back to me?

No, Microsoft can't wipe hard drives before shipping them back to customers. Hard drives are returned to you in the same state they were in when they were received by Microsoft.

Can Microsoft shred my hard drive instead of shipping it back to me?

No, Microsoft can't destroy your hard drive. Hard drives are returned to you in the same state they were in when they were received by Microsoft.

What courier services are supported for return shipping?

If you're a customer in the United States or Europe, Microsoft uses FedEx to return your hard drive. For all other regions, Microsoft uses DHL.

What are the return shipping costs?

Return shipping costs vary, depending on your proximity to the Microsoft data center that you shipped your hard drive to. Microsoft bills your FedEx or DHL account to return your hard drive. The cost of return shipping is your responsibility.

Can I use a custom courier shipping service, such as FedEx Custom Shipping, to ship my hard drive to Microsoft?

Yes.

If I have to ship my hard drive to another country/region, is there anything I need to do?

The hard drive that you ship to Microsoft might have to cross international borders. If so, you're responsible for ensuring that the hard drive and the data it contains are imported and/or exported in accordance with the applicable laws. Before shipping a hard drive, check with your advisors to verify that your drive and data can legally be shipped to the specified Microsoft data center. This helps to ensure that it reaches Microsoft in a timely manner.

Last updated on 06/11/2026

Learn about records management

09/26/2025

> *Microsoft 365 licensing guidance for security & compliance*

A records management system, also known as records and information management, is a solution for organizations to manage regulatory, legal, and business-critical records. Records management for Microsoft Purview helps you achieve your organization's legal obligations, provides the ability to demonstrate compliance with regulations, and increases efficiency with regular disposition of items that are no longer required to be retained, no longer of value, or no longer required for business purposes.

Use the following capabilities to support your records management solution for Microsoft 365 data:

- **Label items as a record**. Create and configure retention labels to mark items as a record that can then be applied by users or automatically applied by identifying sensitive information, keywords, or content types.
- **Migrate and manage your retention requirements with file plan**. By using a file plan, you can bring in an existing retention plan to Microsoft 365, or build a new one for enhanced management capabilities.
- **Configure retention and deletion settings with retention labels**. Configure retention labels with the retention periods and actions based on various factors that include the date last modified or created.
- **Start different retention periods when an event occurs** with event-based retention.
- **Review and validate disposition** with disposition reviews and proof of records deletion.
- **Export information about all disposed items** with the export option.
- **Set specific permissions** for records manager functions in your organization to have the right access.

Using these capabilities, you can incorporate your organization's retention schedules and requirements into a records management solution that manages retention, records declaration, and disposition, to support the full lifecycle of your content.

In addition to the online documentation, you might find it useful to download a deck with FAQs from a records management webinar. The recording of the actual webinar is no longer available.

Records

When an item is declared a record by using a retention label:

- Restrictions are placed on the item in terms of what actions are allowed or blocked.
- Additional activities about the item are logged.
- You have proof of disposition when the item is deleted at the end of their retention period.

You use retention labels to mark items as a **record**, or a **regulatory record**. The difference between these two are explained in the next section. You can either publish those labels so that users and administrators can manually apply them to items, or for labels that mark items as a record, you can auto-apply those labels.

By using retention labels to declare records, you can implement a single and consistent strategy for managing records across your Microsoft 365 environment.

Compare restrictions for what actions are allowed or blocked

Use the following table to identify what restrictions are placed on items as a result of applying a standard retention label, and retention labels that mark items as a record or regulatory record.

A standard retention label has retention settings and actions but doesn't mark items as a record or a regulatory record.

> **Note**
>
> For completeness, the table includes columns for a locked and unlocked record, which is applicable to SharePoint and OneDrive, but not Exchange. The ability to lock and unlock a record uses record versioning that isn't supported for Exchange items. So for all Exchange items that are marked as a record, the behavior maps to the **Record - locked** column, and the **Record - unlocked column** is not relevant.

Expand table

Action	Retention label	Record - locked	Record - unlocked	Regulatory record
Edit contents	Allowed	**Blocked**	Allowed	**Blocked**

Action	Retention label	Record - locked	Record - unlocked	Regulatory record
Edit properties, including rename	Allowed	Allowed [1]	Allowed	**Blocked**
Delete	Allowed [2]	**Blocked**	**Blocked**	**Blocked**
Copy	Allowed	Allowed	Allowed	Allowed
Move within container [3]	Allowed	Allowed	Allowed	Allowed
Move across containers [3]	Allowed	Allowed if never unlocked	**Blocked**	**Blocked**
Open/Read	Allowed	Allowed	Allowed	Allowed
Change label	Allowed	Allowed - container admin only	**Blocked**	**Blocked**
Remove label [4]	Allowed	Allowed - container admin only	**Blocked**	**Blocked**
Override with priority cleanup	Allowed	**Blocked**	**Blocked**	**Blocked**

Footnotes:

[1] Editing properties for a locked record is allowed by default but can be blocked by a tenant setting in the Microsoft Purview portal: Sign in to the Microsoft Purview portal > **Settings** > **Solutions settings** > **Records Management** > **Retention Labels** > **Allow editing of record properties**

[2] Deleting labeled items in SharePoint and OneDrive can be blocked as a tenant setting in the Microsoft Purview portal: Sign in to the Microsoft Purview portal > **Settings** > **Solutions settings** > **Records Management** > **Retention Labels** > **Deletion of items**.

When you apply a standard retention label to a list item that has a document attachment, that document doesn't inherit the retention settings and can be deleted from the list item. In comparison, if that retention label marked items as a record or regulatory record, the document attachment would inherit the retention settings and couldn't be deleted.

[3] Containers include SharePoint sites, OneDrive accounts, and Exchange mailboxes.

[4] Labels can be removed from items even if the labels are no longer published.

> **Important**

The most important difference for a regulatory record is that after it is applied to content, nobody, not even a global administrator, can remove the label.

Retention labels configured for regulatory records also have the following admin restrictions:

- The retention period can't be made shorter after the label is saved, only extended.
- These labels aren't supported by auto-labeling policies, and must be applied by using retention label policies.

In addition, a regulatory label can't be applied to a document that's checked out in SharePoint.

Because of the restrictions and irreversible actions, make sure you really do need to use regulatory records before you select this option for your retention labels. To help prevent accidental configuration, this option is not available by default but must first be enabled by using PowerShell. Instructions are included in Declare records by using retention labels.

Validating migrated records

If you're migrating files to SharePoint or OneDrive and your organization needs to manage these items as records, you might need to validate that the files haven't been altered and retain their immutability status. For example, you're using a migration solution and need to meet the chain of custody requirements. Typical file properties and methods often used for this type of validation, such as file size or file hash, might not be sufficient because SharePoint automatically updates the metadata for a file when it's uploaded.

Instead, to validate your migrated files, you can use the value of the `vti_writevalidationtoken` property, which is a base64-encoded XOR hash of the file before it is modified by SharePoint. Use the following steps:

1. Generate the XOR hash of the original file by using the QuickXorHash algorithm. For more information, see the QuickXorHash Algorithm code snippet.
2. Base64-encode the XOR hash. For more information, see the Base64Encode method documentation.
3. After the file is migrated, retrieve the value of the `vti_writevalidationtoken` property from the uploaded file.

4. Compare the value generated in step 2 with the value retrieved in step 3. These two values should match. If they do, you've validated that the file hasn't changed.

Configuration guidance

See Get started with records management. This article has information about subscriptions, permissions, and links to end-to-end configuration guidance for records management scenarios.

Regulatory requirements for data lifecycle management and records management

09/26/2025

> *Microsoft 365 licensing guidance for security & compliance*.

Use the resources on this page to help you meet specific regulatory requirements for data lifecycle management and records management in Microsoft 365. Each section of this document focuses on one or more related regulations and includes any existing guidance or third-party assessment of how to configure Microsoft 365 to help with the requirements outlined.

These resources are available to download from the Data Protection Resources, FAQ, and White Papers page of the Service Trust Portal.

New Zealand Public Records Act

Supporting New Zealand's Public Records Act compliance obligations with Microsoft 365 - Download assessment

Applicable workloads: SharePoint, OneDrive, Teams, and Exchange

Released January 2021, this report has been produced in partnership with Microsoft New Zealand to assess the capabilities of Microsoft 365 services for recording, storing, and managing requirements for electronic records, as specified by:

- New Zealand Public Records Act 2005, which sets guidelines for preservation of public archives and local authority archives in New Zealand.

This report helps you understand how the system aspects of the New Zealand Public Records Act 2005 (PRA) are achievable when using Microsoft 365.

SEC 17a-4(f), FINRA 4511(c), and CFTC 1.31(c)-(d)

Cohasset Assessment - Microsoft 365 - SEC Rule 17a-4(f) - Immutable Storage for SharePoint, OneDrive, Exchange, Teams, and Viva Engage - Download assessment

Applicable workloads: SharePoint, OneDrive, Teams, Exchange, and Viva Engage

Latest version released July 2022, this report has been produced in partnership with Cohasset Associates, Inc. (Cohasset) to assess the capabilities of Microsoft 365 services for recording,

storing, and managing requirements for electronic records, as specified by:

- Securities and Exchange Commission (SEC) in 17 CFR § 240.17a-4(f), which regulates exchange members, brokers or dealers.
- Financial Industry Regulatory Authority (FINRA) Rule 4511(c), which defers to the format and media requirements of SEC Rule 17a-4(f).
- The principles-based electronic records requirements of the Commodity Futures Trading Commission (CFTC) in 17 CFR § 1.31(c)-(d).

The opinion from Cohasset is that when compliance features are properly configured and carefully applied and managed as described in their report, the assessed Microsoft 365 services meet the five requirements related to the recording and non-rewriteable, non-erasable storage of electronic records.

Declare records by using retention labels

09/26/2025

> *Microsoft 365 licensing guidance for security & compliance*.

To declare documents and emails as records, you use retention labels that mark the content as a **record** or a **regulatory record**.

If you're not sure whether to use a record or a regulatory record, see Compare restrictions for what actions are allowed or blocked. If you need to use regulatory records, you must first run a PowerShell command, as described in the next section.

You can then either publish those labels in a retention label policy so that users and administrators can apply them to content, or for labels that mark items as records (but not regulatory records), auto-apply those labels to content that you want to declare a record.

How to display the option to mark content as a regulatory record

> **Note**
>
> The following procedure is an auditable action, logging **Enabled regulatory record option for retention labels** in the Retention policy and retention label activities section of the audit log.

By default, the retention label option to mark content as a regulatory record isn't displayed in the retention label wizard. To display this option, you must first run a PowerShell command:

1. Connect to the Office 365 Security & Compliance PowerShell.
2. Run the following cmdlet:

   ```
   PowerShell
   Set-RegulatoryComplianceUI -Enabled $true
   ```

   There is no prompt to confirm and the setting takes effect immediately.

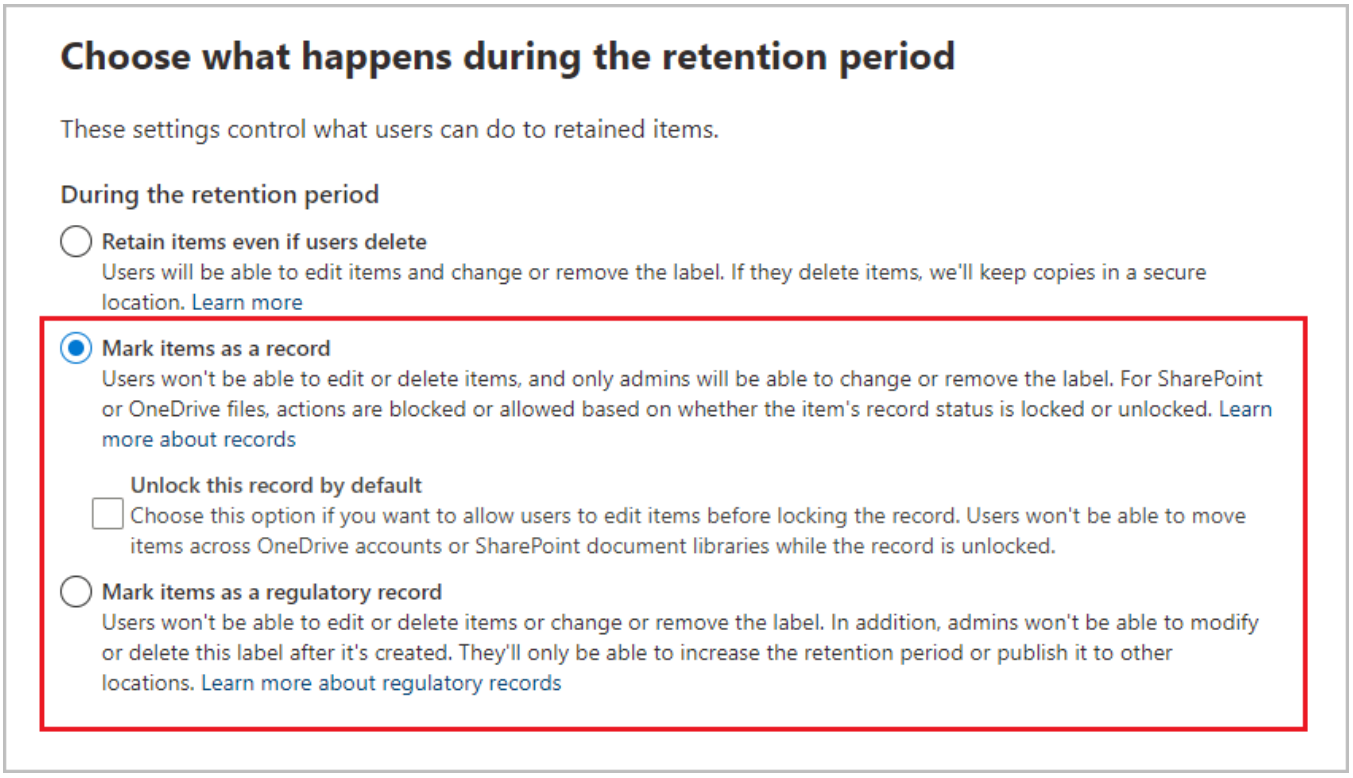
If you change your mind about seeing this option in the retention label wizard, you can hide it again by running the same cmdlet with the **false** value: `Set-RegulatoryComplianceUI -Enabled`

`$false`

Configuring retention labels to declare records

When you create a retention label from the **Records Management** solution in the Microsoft Purview portal, you have the option to mark items as a record. If you ran the PowerShell command from the previous section, you can alternatively mark items as a regulatory record.

For example:



Using this retention label, you can now apply it to SharePoint or OneDrive documents and Exchange emails, as needed.

For full instructions:

- Publish retention labels and apply them in apps
- Apply a retention label to content automatically (not supported for regulatory records)

Tenant setting for editing record properties

If you'll use retention labels to declare items as records (rather than regulatory records) in SharePoint and OneDrive, consider whether you need to change the default tenant setting that allows users to edit the properties for a locked record when files are larger than 0 bytes.

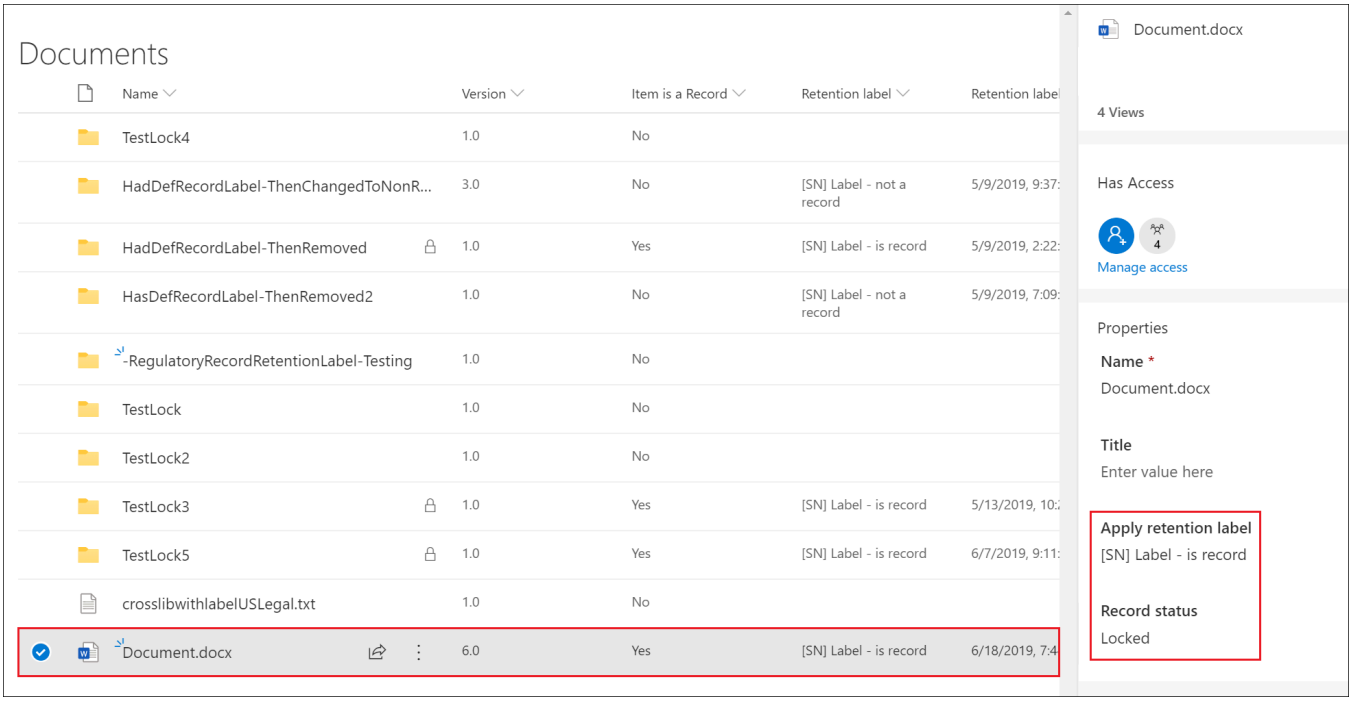
To change this default, turn off the setting **Allow users to edit record properties**: Sign in to the Microsoft Purview portal > **Settings** > **Solutions settings** > **Records Management** > **Retention Labels** > **Allow users to edit record properties**

Applying the configured retention label to content

When retention labels that mark items as a record or regulatory record are made available for users to apply them in apps:

- For Exchange, any user with write-access to the mailbox can apply these labels.
- For SharePoint and OneDrive, any user in the default Members group (the Contribute permission level) can apply these labels.

Example of a document marked as record by using a retention label:



Searching the audit log for labeled items that were declared records

The actions of labeling to declare items as records are logged in the audit log.

For SharePoint items:

- From **File and page activities**, select **Changed retention label for a file**. This audit event is for retention labels that mark items as records, regulatory records, or that are standard retention labels.

For Exchange items:

- From **Exchange mailbox activities**, select **Labeled message as a record**. This audit event is for retention labels that mark items as records or regulatory records.

For more information about these events, see Audit log activities.

Next steps

Understand how you can use record versioning to update records stored in SharePoint or OneDrive.

Use record versioning to update records stored in SharePoint or OneDrive

09/26/2025

> *Microsoft 365 licensing guidance for security & compliance*.

> **Note**
>
> Because regulatory records block editing, record versioning is not available for regulatory records.
>
> You can also prevent record versioning for your tenant, even if you're not using regulatory records. In the Microsoft Purview portal, go to **Settings**> **Records Management** > **Retention Labels** > **Configure record versioning** and then for **Configure record versioning**, turn off the setting for **Enable record versioning**.

The ability to mark a document as a record and restrict actions that can be performed on the record is an essential goal for any records management solution. However, collaboration might also be needed for people to create subsequent versions.

For example, you might mark a sales contract as a record, but then need to update the contract with new terms and mark the latest version as a new record while still retaining the previous record version. For these types of scenarios, SharePoint and OneDrive support *record versioning*. OneNote notebook folders don't support record versioning.

To use record versioning, you first label the document with a retention label that's configured to mark items as a record. At this point, a document property, called *Record status* is displayed next to the retention label. Depending on whether the label is configured to unlock the record by default, the initial record status is either **Locked** or **Unlocked**.

You can now do the following things:

- **Continually edit and retain individual versions of the document as records, by unlocking and locking the Record status property.** Only when the **Record status** property is set to **Locked** is a new version of the record retained. This toggle of locked and unlocked reduces the risk of retaining unnecessary versions and copies of the document.

  > **Note**

If the label is configured to unlock the record by default, but versioning is not enabled by the admin, or prevented by the records management setting, users will not be able to unlock the document after they lock it.

- **Have the records automatically stored in an in-place records repository located with the site.** Each site in SharePoint and OneDrive preserves content in its Preservation Hold library. Record versions are stored in the Records folder in this library as individual files. For more information about how the Preservation Hold library works, see How retention works for SharePoint and OneDrive.

- **Maintain an evergreen document that contains all versions.** By default, each SharePoint and OneDrive document has a version history available on the item menu. In this version history, you can easily see which versions are records and view those documents.

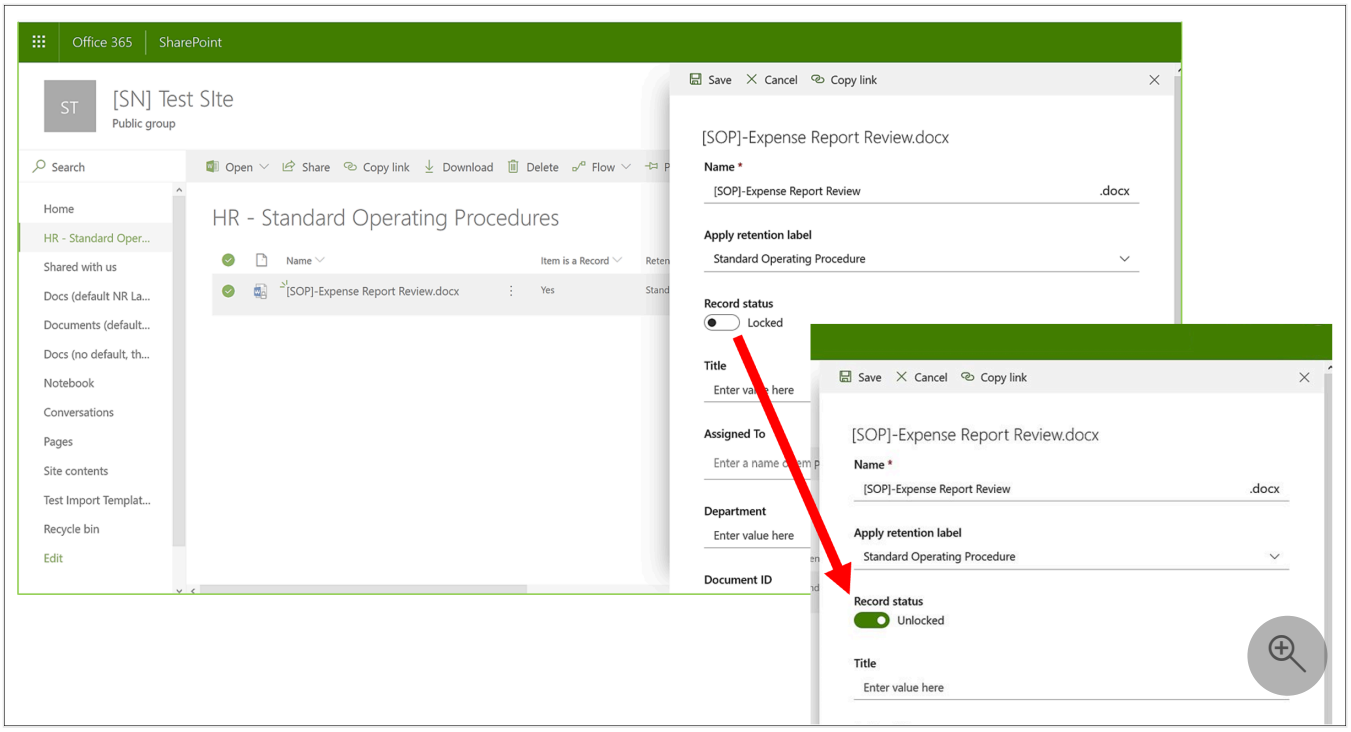
Tip

When you use record versioning with a retention label that has a delete action, consider configuring the retention setting **Start the retention period based on:** to be **When items were labeled**. With this label setting, the start of the retention period is reset for each new record version, which ensures that older versions will be deleted before newer versions.

If the retention label is configured for disposition review at the end of the retention period, each version undergoes its own disposition review.

By default, record versioning is automatically available for any document that has a retention label applied that marks the item as a record, and that label is published to the site. When a user views the document properties by using the details pane, they can toggle the **Record status** between **Locked** and **Unlocked**. A padlock icon might also be displayed for the document, to help identify when the file is locked.

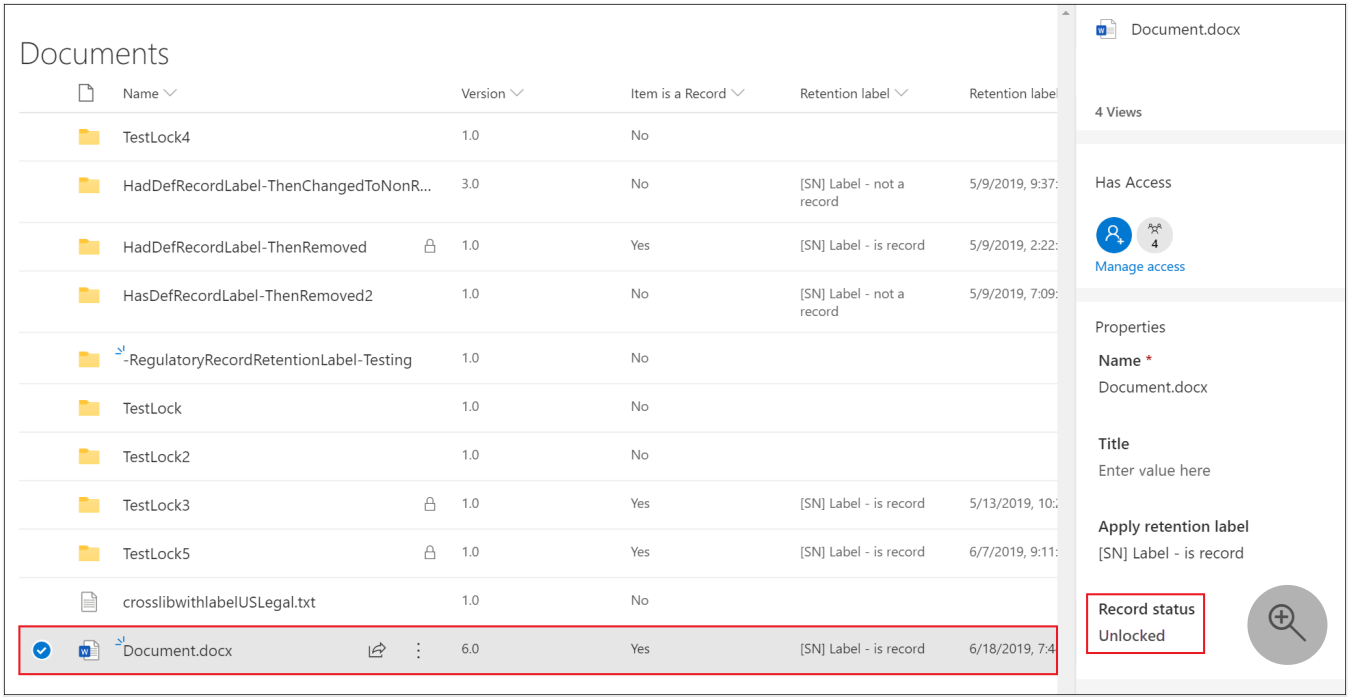
While the document is unlocked, any user with standard edit permissions can edit the file. However, users can't delete the file, because it's still a record. When editing is complete, a user can then toggle the **Record status** from **Unlocked** to **Locked**, which prevents further edits while in this status.



For more information about what user actions are allowed when a record is locked or unlocked, see Compare restrictions for what actions are allowed or blocked.

Locking and unlocking a record

After a retention label that marks content as a record is applied to a document, any user with Contribute permissions or a narrower permission level can unlock a record or lock an unlocked record.



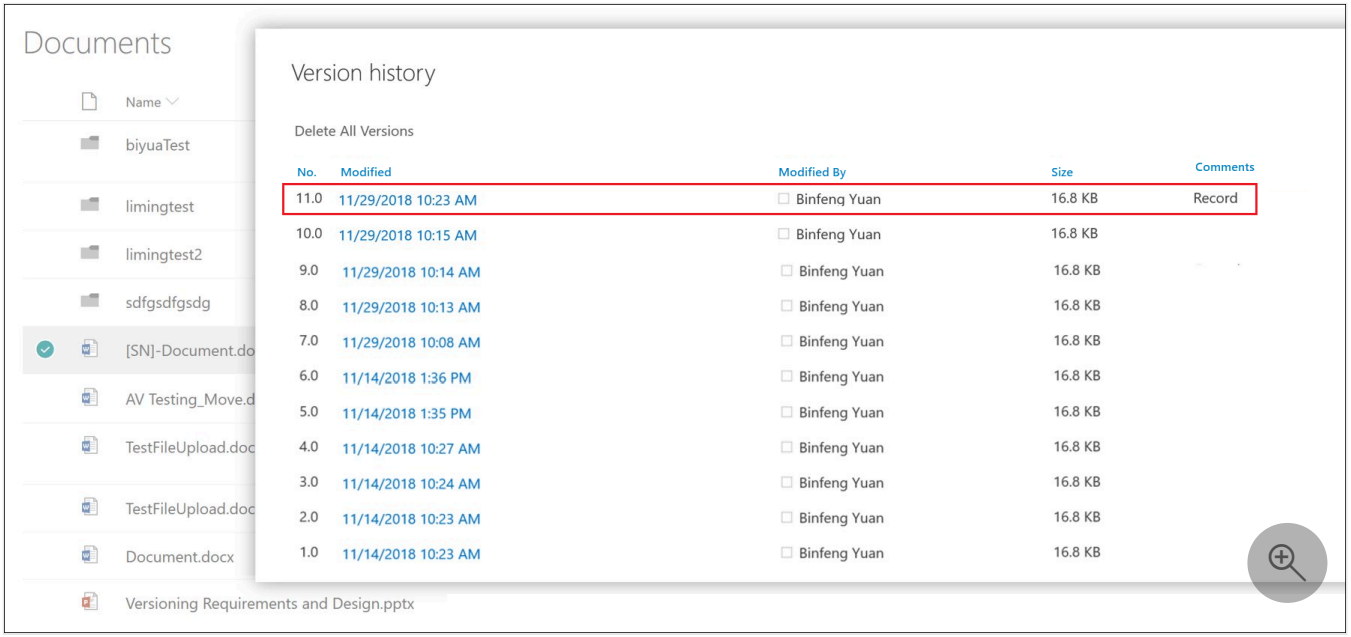
When a record is unlocked, the following actions occur:

1. If the current site doesn't have a Preservation Hold library, one is created.
2. If the Preservation Hold library doesn't have a Records folder, one is created.
3. A **Copy to** action copies the latest version of the document to the Records folder. The **Copy to** action includes only the latest version and no prior versions. This copied document is now considered a record version of the document, and its file name has the format: [Title GUID Version#]
4. The copy created in the Records folder is added to the version history of the original document, and this version shows the word **Record** in the comments field.
5. The original document is a new version that can be edited, but not deleted. The document library column **Item is a Record** still shows the **Yes** value because the document is still a record, even if it can now be edited.

When a user locks a record, the original document again can't be edited. But it is the action of unlocking a record that copies a version to the Records folder in the Preservation Hold library.

Record versions

Each time a record is unlocked, the latest version is copied to the Preservation Hold library, and that version contains the value of **Record** in the **Comments** field of the version history.



To view the version history, select a document in the document library and then click **Version history** in the item menu.

Searching the audit log for record versioning events

The actions of locking and unlocking records are logged in the audit log. From **File and page activities**, select **Changed record status to locked** and **Changed record status to unlocked**.

For more information about searching for these events, see Audit log activities.

Next steps

For other scenarios supported by records management, see Common scenarios for records management.

Get started with records management

09/26/2025

> *Microsoft 365 licensing guidance for security & compliance*.

Ready to start managing your organization's high-value content for legal, business, or regulatory obligations by using a records management solution in Microsoft 365? Use the following guidance to get started:

1. **Understand how retention and deletion works** in Microsoft 365, and identify whether you need to use retention policies to supplement retention labels that manage documents and emails at the item level: Learn about retention policies and retention labels

   If necessary, create retention policies for baseline governance of data across Microsoft 365 workloads.

2. **Understand the records management solution** and how retention labels can be used to allow or block actions when documents and emails are declared records: Learn about records management

3. **Create your file plan for retention and deletion settings and actions, and when items should be marked as records** by importing an existing plan if you have one, or create new retention labels: Use file plan to create and manage retention labels

4. **Publish and apply your retention labels**. Retention labels are reusable building blocks that can be used in multiple policies and can be incorporated into user workflows:

   - Publish retention labels and apply them in apps
   - Apply a retention label to content automatically

> **Tip**
>
> If you're migrating records to Microsoft 365 and need to validate that they haven't been altered, see **Validating migrated records**.

Licensing and billing requirements

A number of different subscriptions support records management and the licensing requirements for users depends on the features you use.

To see the options for licensing your users to benefit from Microsoft Purview features, see the Microsoft 365 licensing guidance for security & compliance. For Records Management, see the Microsoft Purview Data Lifecycle Management & Microsoft Purview Records Management section for feature-level licensing requirements.

You must set up pay-as-you-go billing to create retention policies that include the locations Microsoft Copilot experiences, Enterprise AI apps, and Other AI apps. There's no charge for Microsoft 365 Copilot interactions.

Permissions

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

Members of your compliance team who are responsible for records management need permissions to the Microsoft Purview portal. To grant permissions for this limited administration, we recommend that you add users to the **Records Management** admin role group. This role group grants permissions for all features related to records management, which include the permissions to create and manage adaptive policy scopes, and disposition review and verification.

For a read-only role, you can create a new role group and add the **View-Only Record Management** role to this group.

For instructions to add users to the default roles or create your own role groups, see Permissions in the Microsoft Purview portal.

These permissions are required only to create, configure, and apply retention labels that declare records, and manage disposition. The person configuring these labels doesn't require access to the content.

Support for administrative units

Records management supports administrative units that have been configured in Microsoft Entra ID:

- You can assign administrative units to members of role groups that are used with Microsoft Purview Records Management. Edit the Records Management role group or

other role groups that support administrative units. From these role groups, select individual members, and then the **Assign admin units** option to select administrative units from Microsoft Entra ID. These administrators are now restricted to managing just the users in those administrative units.

- You can define the initial scope of retention policies and retention label policies when you create or edit these policies. When you select administrative units, only the users in those administrative units will be eligible for the policy.

  > **Important**
  >
  > Don't select administrative units for a retention label policy that you want to apply to SharePoint sites. Because administrative units support only users and groups, if you configure a retention label policy to use administrative units, you won't be able to select the locations for SharePoint sites.

- Both adaptive scopes and static scopes support administrative units.
- Additional impact for restricted administrators
  - Policy lookup: Restricted administrators will see policies only from users within their assigned administrative units
  - Disposition review and verification: Restricted administrators will be able to add reviewers only from within their assigned administrative units, and see disposition reviews and items disposed only from users within their assigned administrative units
- Currently, retention labels and events don't support administrative units.
- Currently, retention policies for Copilots and AI apps don't support administrative units.
- Currently, a restricted administrator can create and view adaptive scopes for all administrative units when they use PowerShell cmdlets.
- Currently, inactive mailboxes aren't supported in a policy when you select one or more administrative units. To include inactive mailboxes in the policy, you must be an unrestricted administrator and select **Full directory**.

For more information about how Microsoft Purview supports administrative units, see Administrative units.

Common scenarios

Use the following table to help you map your business requirements to the scenarios that are supported by records management.

> **Tip**
>
> Need to comply with a specific industry regulation? Check **Regulatory requirements for data lifecycle management and records management** for regulation-specific guidance.

Expand table

I want to ...	Documentation
Declare a record	Declare records by using retention labels
Update a record	Use record versioning to update records stored in SharePoint or OneDrive
Let admins and users manually apply retain and delete actions for documents and emails: - SharePoint - OneDrive - Outlook and Outlook on the web	Publish retention labels and apply them in apps
Let site admins set default retain and delete actions for all content in a SharePoint library, folder, or document set	Publish retention labels and apply them in apps
Let users automatically apply retain and delete actions to emails by using Outlook rules	Publish retention labels and apply them in apps
Let admins apply retain and delete actions to a Microsoft Syntex model, so that these settings are automatically applied to identified files in a SharePoint library	Publish retention labels and apply them in apps
Automatically apply retain and delete actions to documents and emails	Apply a retention label to content automatically
Automatically retain and delete Teams meeting transcripts or recordings and their accompanying transcripts	Apply a retention label to content automatically
Automatically retain and delete documents and emails with a specific sensitivity label	Apply a retention label to content automatically
Start the retention period when an event occurs, such as: - Employees leave the organization - Contracts expire - End of product lifetime	Start retention when an event occurs
Do custom actions or integrate with other solutions at the end of the retention period	Customize what happens at the end of the retention period

I want to ...	Documentation
Restrict changes to policies to help meet regulatory requirements or safeguard against rogue administrators	Use Preservation Lock to restrict changes to retention policies and retention label policies
Manage the lifecycle of different document types in SharePoint	Use retention labels to manage the lifecycle of documents stored in SharePoint
Apply a retention label to a file when I receive an alert that content containing personal data is being stored or remains untouched for too long	Investigate and remediate alerts in Privacy Risk Management
Make sure somebody reviews and approves before content is deleted at the end of its retention period	Disposition reviews
Have proof of disposition for content that is permanently deleted at the end of its retention period	Disposition of records
Monitor how and where retain and delete settings are applied to items	Monitoring retention labels
Programamtically create retention labels, and manage events and event type retention triggers	Microsoft Graph API for records management
Programmatically apply and manage retention labels in SharePoint and OneDrive	Microsoft Graph API for retention labels in SharePoint and OneDrive: - lock and unlock records - set retention labels - remove retention labels - get metadata for retention labels

End-user documentation

If you're using retention policies for baseline data governance, they typically work unobtrusively in the background without user interaction. As a result, they need little documentation for users. Retention policies for Teams inform users when their messages have been deleted with a link to Teams messages about retention policies.

In comparison, retention labels have a UI presence in Microsoft 365 apps, so make sure you provide guidance for end users and your help desk before these labels are deployed to your production network. To help users apply retention labels in SharePoint and OneDrive, and information about unlocking records for editing, see Apply retention labels to files in SharePoint or OneDrive.

However, the most effective end-user documentation will be customized guidance and instructions you provide for the retention label names and configurations you choose.

Use file plan to create and manage retention labels

Summarize this article for me

> *Microsoft Purview service description*

Although you can create and manage retention labels from **Data lifecycle management** in the Microsoft Purview portal, file plan from **Records management** has additional management capabilities:

- You can bulk-create retention labels by importing the relevant information from a spreadsheet.
- You can export the information from existing retention labels for analysis and offline collaboration.
- More information about the retention labels is displayed to make it easier to see into and across the settings of all your retention labels from one view.
- File plan descriptors support additional and optional information for each label.

File plan can be used for all retention labels, even if they don't mark content as a record.

For information about what retention labels are and how to use them, see Learn about retention policies and retention labels.

Accessing file plan

To access file plan, you must have one of the following admin roles:

- Retention Manager
- View-only Retention Manager

Then sign in to the Microsoft Purview portal > **Solutions** > **Records Management** > **File plan**.

Navigating your file plan

If you've already created retention labels from **Data lifecycle management** in the Microsoft Purview portal, these labels automatically display in your file plan.

Similarly, if you now create retention labels in file plan, they're also available from **Data lifecycle management** if the labels aren't configured to mark content as a record.

On the **File plan** page, you see all your labels with their status and settings, optional file plan descriptors, an export option to analyze or enable offline reviews of your labels, and an import option to create retention labels.

Label settings columns

All columns except the label **Name** can be displayed or hidden by selecting the **Customize columns** option. But by default, the first few columns display information about the label status and its settings:

- **Status** identifies whether the label is included in a label policy or auto-apply policy (**Active**) or not (**Inactive**).
- **Based on** identifies how or when the retention period begins. Valid values:
  - Event
  - When created
  - Last modified
  - When labeled
- **Is record** identifies if the item is marked as a record when the label is applied. Valid values:
  - No
  - Yes
  - Yes(Regulatory)
- **Is unlocked by default** identifies if the item marked as a record is unlocked when the label is applied. Valid values:
  - No
  - Yes
- **Relabel to** identifies if the label is configured to apply another label at the end of the retention period. Valid values:
  - Blank or the selected label name
- **Retention duration** identifies the retention period. Valid values:
  - Days
  - Months
  - Years
  - Forever
  - None

- **Disposition type** identifies what happens to the content at the end of the retention period. Valid values:
  - No action
  - Auto-delete
  - Review required

File plan descriptors columns

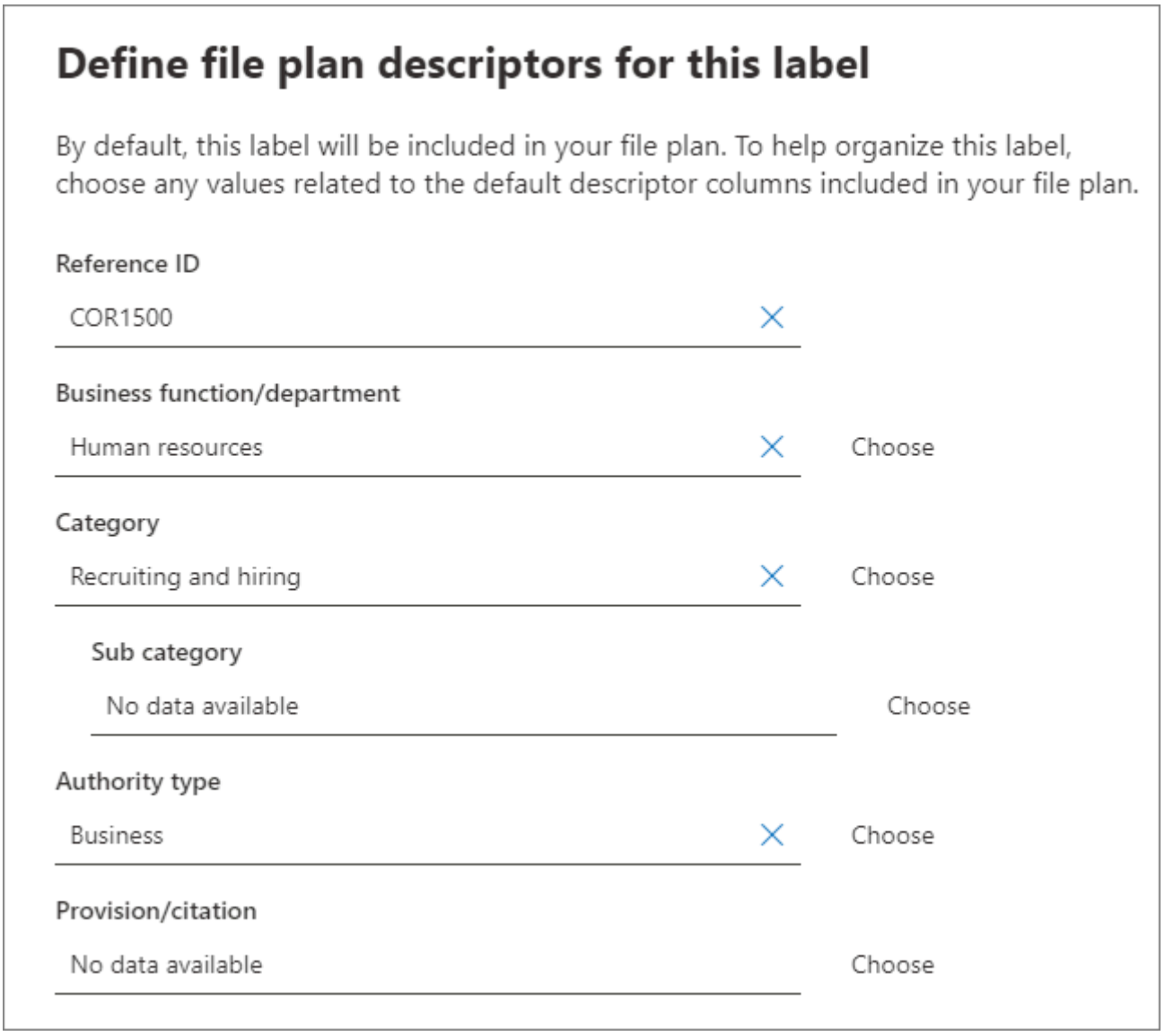
File plan lets you include more information as part of your retention labels. These file plan descriptors provide more options to improve the manageability and organization of the content you need to label.

By default, starting with **Reference ID**, the next few columns display these optional file plan descriptors that you can specify when you create a retention label, or edit an existing label.

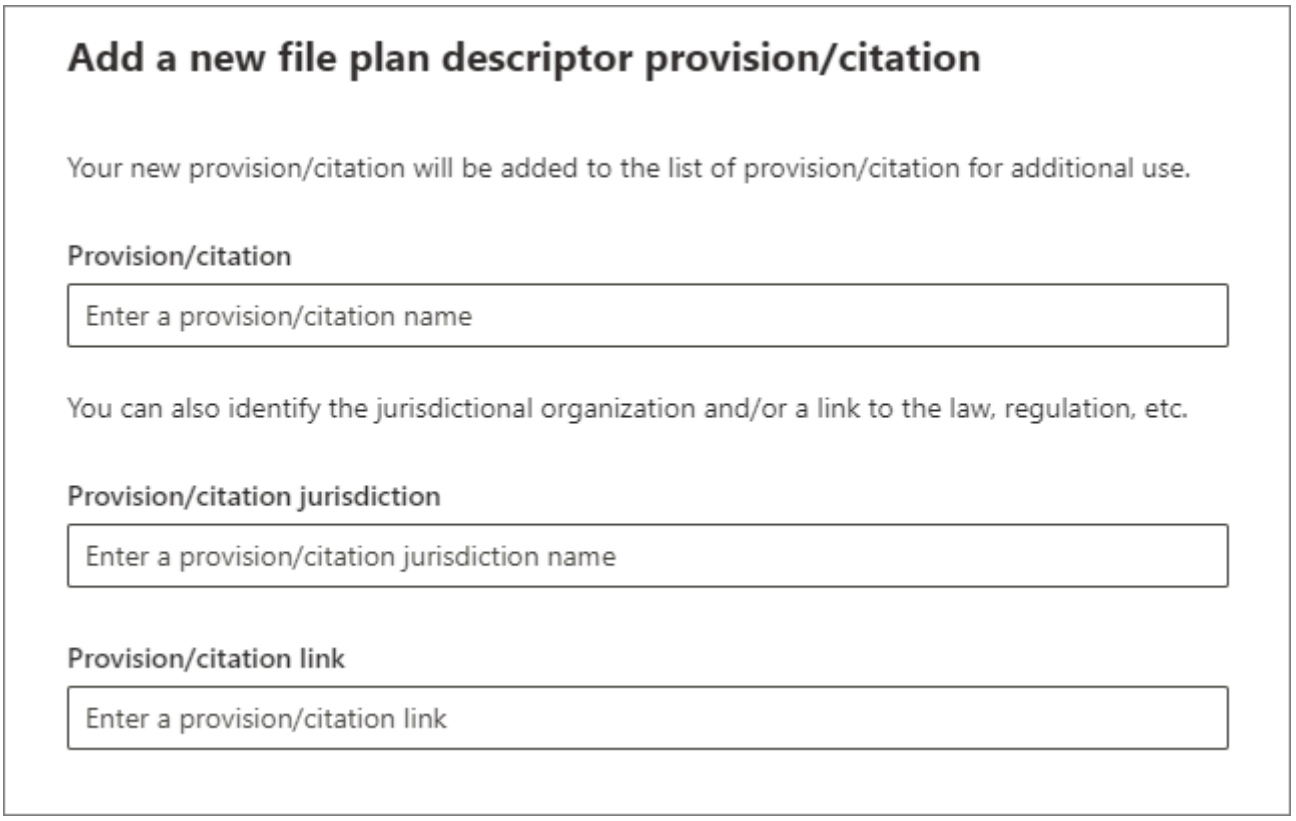
To get you started, there are some out-of-box values for the following file plan descriptors:

- Business function/department
- Category
- Authority type
- Provision/citation

Example of file plan descriptors when you create or edit a retention label:



When you select **Choose** for each of these optional descriptors, you can select one of the out-of-box values, or create your own and then select it. For example:



Create retention labels

1. From the **File plan** page, select **+ Create a label** > **Retention label**

2. Follow the prompts for the configuration process. Be careful what name you choose, because this can't be changed after the label is saved.

   For more information about the retention settings, see Settings for retaining and deleting content.

   To use the retention label to declare records, select **Mark items as records**, or **Mark items as regulatory records**. For more information, see Configuring retention labels to declare records.

3. After you've created the label and you see the options to publish the label, auto-apply the label, or just save the label: Select **Just save the label for now**, and then select **Done**.

4. Repeat these steps to create more labels.

Edit retention labels

To edit an existing retention label, select it from the **File Plan** page, and then select the **Edit label** option to start the edit retention process that lets you change the label description and any eligible settings.

Some settings can't be changed after the label is created and saved, which include:

- The retention label name and the retention settings except the retention period.
- The option to mark items as a record.

Delete retention labels

You can delete retention labels that aren't currently included in any published or auto-apply retention label policies, that aren't configured for event-based retention, or that mark items as regulatory records.

For retention labels that you can delete, if they've been applied to items, the deletion fails and you see a link to content explorer to identify the labeled items.

However, it can take up to two days for content explorer to show the items that are labeled. In this scenario, the retention label might be deleted without showing you the link to content explorer.

Export all retention labels to analyze or enable offline reviews

From your file plan, you can export the details of all retention labels into a .csv file to help you facilitate periodic compliance reviews with data governance stakeholders in your organization.

To export all retention labels to *.csv file, navigate to the **File plan** page, and then select **Export**.

Import retention labels into your file plan

In file plan, you can bulk-import new retention labels by using a .csv file with a specific format:

1. Navigate to the **File plan** page, and then select **Import**.
2. On the **Fill out and import your file plan** pane, select **Download a blank template**.
3. After the template downloads, add one row for each label, and save the file. See the next section of this document for information about the properties and values for each.
4. Select **Upload a file** to upload and validate the completed template.
    - If validation fails: Note the row number and column name to correct in the import file. Correct the errors in the file and save it, and then retry.

- If validation passes: A success message displays and the entries are converted into retention labels.

You can now publish your new retention labels, or auto-apply them. You can do both from the **Label policies** tab by selecting **Publish labels**, or **Auto-apply a label**.

Information about the label properties for import

Use the following information to help you fill out the downloaded template to import new retention labels. Some values have a maximum length for import:

- **LabelName**: Maximum length of 64 characters
- **Comment** and **Notes**: Maximum length of 1024 characters
- All other values: Unlimited length

Expand table

Property	Type	Required	Valid values
LabelName	String	Yes	This property specifies the name of the retention label and must be unique in your tenant. Supported characters for import: a-z, A-Z, 0-9, hyphen (-) and the space character.
Comment	String	No	Use this property to add a description about the retention label for admins. This description appears only to admins who manage the retention label in the Microsoft Purview portal.
Notes	String	No	Use this property to add a description about the retention label for users. This description appears when users hover over the label in apps like Outlook, SharePoint, and OneDrive. If you leave this property blank, a default description is displayed, which explains the label's retention settings.
IsRecordLabel	String	No, unless **Regulatory** is **TRUE**	This property specifies whether the label marks the content as a record. Valid values are: **TRUE**: The label marks the item as a record and as a result, the item

Property	Type	Required	Valid values
			can't be deleted. **FALSE**: The label doesn't mark the content as a record. This is the default value. Group dependencies: When this property is specified, RetentionAction, RetentionDuration, and RetentionType must also be specified.
RetentionAction	String	No, unless **RetentionDuration**, **RetentionType**, or **ReviewerEmail** are specified	This property specifies what action to take after the value specified by the RetentionDuration property (if specified) expires. Valid values are: **Delete**: Items older than the value specified by the RetentionDuration property are deleted. **Keep**: Retain items for the duration specified by the RetentionDuration property and then do nothing when the duration period expires. **KeepAndDelete**: Retain items for the duration specified by the RetentionDuration property and then delete them when the duration period expires. Group dependencies: When this property is specified, RetentionDuration and RetentionType must also be specified.
RetentionDuration	String	No, unless **RetentionAction** or **RetentionType** are specified	This property specifies the number of days to retain the content. Valid values are: **Unlimited**: Items will be retained indefinitely. ***n**: A positive integer in days; for example, **365**. The maximum number supported is 36,525, which is 100 years. If you need longer than this maximum, use Unlimited instead. Group dependencies: When this

Property	Type	Required	Valid values
			property is specified, RetentionAction and RetentionType must also be specified.
RetentionType	String	No, unless **RetentionAction** or **RetentionDuration** are specified	This property specifies whether the retention duration (if specified) is calculated from the content creation date, event date, when labeled date, or last modified date. Valid values are: **CreationAgeInDays** **EventAgeInDays** **TaggedAgeInDays** **ModificationAgeInDays** Group dependencies: When this property is specified, RetentionAction and RetentionDuration must also be specified.
ReviewerEmail	SmtpAddress	No	When this property is specified, a disposition review will be triggered when the retention duration expires. This property specifies the email address of a reviewer in your tenant for the **KeepAndDelete** retention action. You can include the email address of individual users, distribution groups, or security groups in your tenant. Specify multiple email addresses by separating them with semicolons. Group dependencies: When this property is specified, **RetentionAction** (must be **KeepAndDelete**), **RetentionDuration**, and **RetentionType** must also be specified.
ReferenceId	String	No	This property specifies the value that's displayed in the **Reference Id** file plan descriptor, which you can

Property	Type	Required	Valid values
			use as a unique value to your organization.
DepartmentName	String	No	This property specifies the value that's displayed in the **Function/department** file plan descriptor.
Category	String	No	This property specifies the value that's displayed in the **Category** file plan descriptor.
SubCategory	String	No	This property specifies the value that's displayed in the **Sub category** file plan descriptor.
AuthorityType	String	No	This property specifies the value that's displayed in the **Authority type** file plan descriptor.
CitationName	String	No	This property specifies the name of the citation displayed in the **Provision/citation** file plan descriptor. For example, "Sarbanes-Oxley Act of 2002".
CitationUrl	String	No	This property specifies the URL that's displayed in the **Provision/citation** file plan descriptor.
CitationJurisdiction	String	No	This property specifies the jurisdiction or agency that's displayed in the **Provision/citation** file plan descriptor. For example, "U.S. Securities and Exchange Commission (SEC)".
Regulatory	String	No	This property specifies whether the label marks the content as a regulatory record, which is more restrictive than a record. To use this label configuration, your tenant must be configured to display the option to mark content as a regulatory record, or the import validation will fail. Valid values are: **TRUE**: The label marks the item as a regulatory record. You must also set

Property	Type	Required	Valid values
			the **IsRecordLabel** property to TRUE. **FALSE**: The label doesn't mark the content as a regulatory record. This is the default value.
EventType	String	No, unless **RetentionType** is **EventAgeInDays**	This property specifies an event type used for event-based retention. Specify an existing event type that's displayed in **Records management** > **Events** > **Manage event types**. Alternatively, use the Get-ComplianceRetentionEventType cmdlet to view the available event types. Although there are some built-in event types, such as **Employee activity** and **Product lifetime**, you can also create your own event types. If you specify your own event type, it must exist before the import because the name is validated as part of the import process.
IsRecordUnlockedAsDefault	String	No	This property specifies whether the item marked as a record is unlocked when the label is applied. Valid values are: **TRUE**: The item marked as a record is unlocked when the label is applied. The **IsRecordLabel** property must be set to TRUE and the property **Regulatory** can't be set to TRUE. **FALSE**: The item isn't marked as a record or is marked as record but locked when the label is applied. This is the default value.
ComplianceTagForNextStage	String	No	This property specifies the name of a replacement label to be applied at the end of the retention period. Do not specify this property if **Regulatory** is **TRUE**.

Label settings not currently supported for import:

- Multi-stage disposition review: Although you can configure the settings for a single disposition review stage when you import retention labels with a template, you can't specify additional review stages. Instead, configure these in the Microsoft Purview portal after the import succeeds.

Next steps

Now you've created retention labels, they're ready to be added to items by publishing the labels, or automatically applying them:

- Publish retention labels and apply them in apps
- Apply a retention label to content automatically

Last updated on 02/27/2026

Start retention when an event occurs

Summarize this article for me

> *Microsoft Purview service description*

When you retain content, the retention period is often based on the age of the content. For example, you might retain documents for seven years after they're created and then delete them. But when you configure retention labels, you can also base a retention period on when a specific type of event occurs. The event triggers the start of the retention period, and all content with a retention label applied for that type of event get the label's retention actions enforced on them.

Examples for using event-based retention:

- **Employees leaving the organization** Suppose that employee records must be retained for 10 years from the time an employee leaves the organization. After 10 years elapse, all documents related to the hiring, performance, and termination of that employee must be disposed. The event that triggers the 10-year retention period is the employee leaving the organization.
- **Contract expiration** Suppose that all records related to contracts must be retained for five years from the time the contract expires. The event that triggers the five-year retention period is the expiration of the contract.
- **Product lifetime** Your organization might have retention requirements related to the last manufacturing date of products for content such as technical specifications. In this case, the last manufacturing date is the event that triggers the retention period.

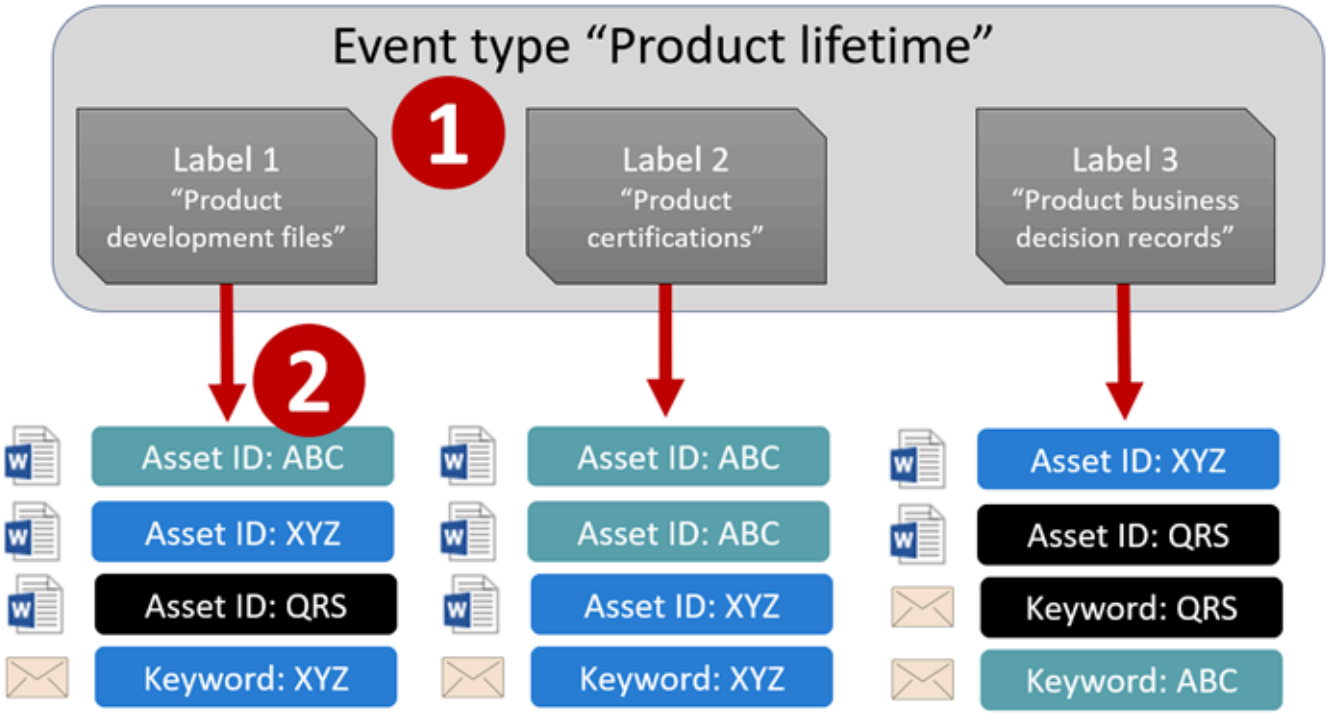
Event-based retention, sometimes also referred to as "event-driven retention", is typically used with records management processes. This means that:

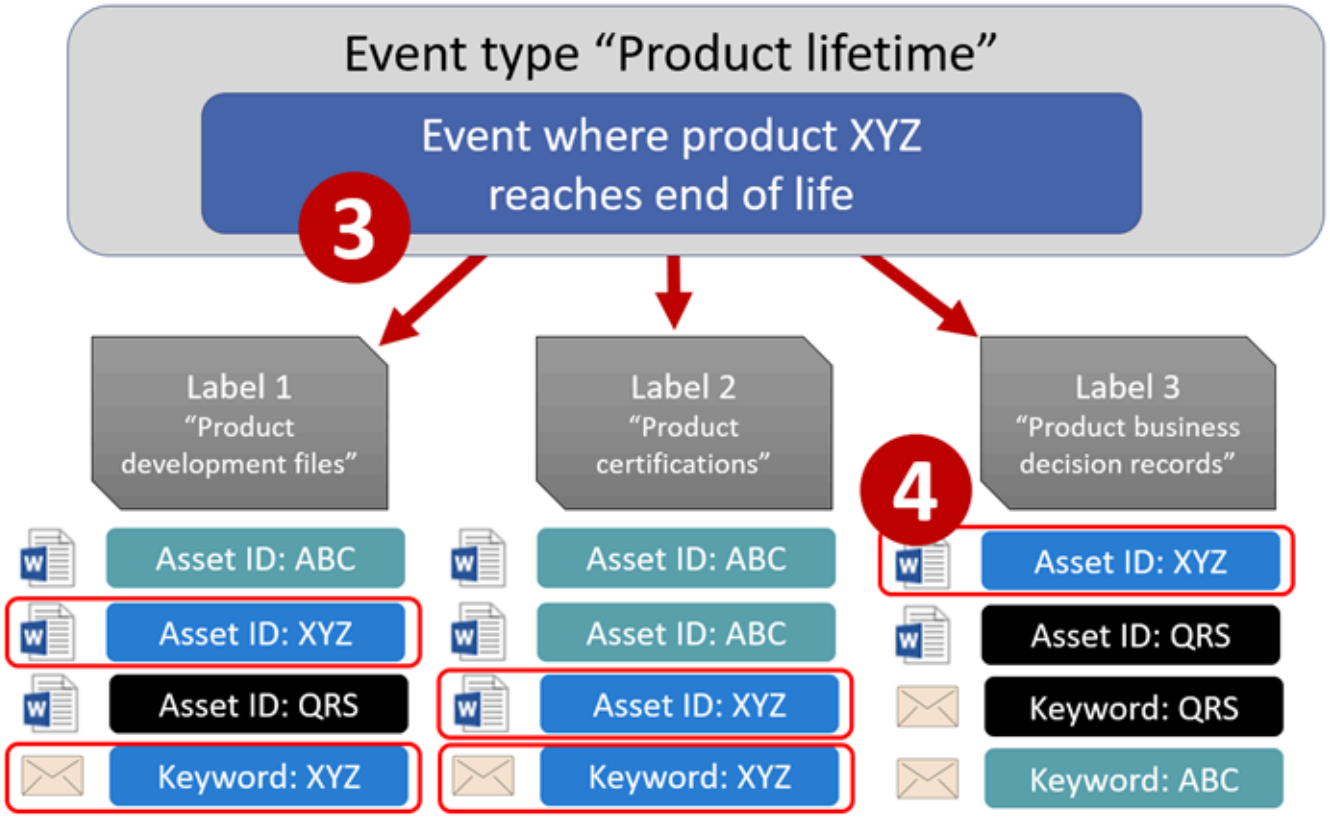
- Retention labels based on events also usually mark items as a record, as a part of a records management solution. For more information, see Learn about records management.
- A document that's been declared a record but whose event trigger hasn't yet happened is retained indefinitely. Until an event triggers that document's retention period, which then expires, records can't be permanently deleted.
- Retention labels based on events often trigger a disposition review at the end of the retention period, so that a records manager can manually review and dispose of the content. For more information, see Disposition of content.

A retention label based on an event has the same capabilities as any retention label in Microsoft Purview. For more information, see Learn about retention policies and retention labels.

Understanding the relationship between event types, labels, events, and asset IDs

To successfully use event-based retention, it's important to understand the relationship between event types, retention labels, events, and asset IDs as illustrated in the diagrams and the explanation that follows:





1. You create retention labels for different types of content and then associate them with a type of event. For example, retention labels for different types of product files and records are associated with an event type named Product Lifetime because those records must be retained for 10 years from the time the product reaches its end of life.

2. Users that are typically records managers apply those retention labels to content and (for documents in SharePoint and OneDrive) enter an asset ID for each item. In this example, the asset ID is a product name or code used by the organization. Then, each product's records are assigned a retention label, and each record has a property that contains an asset ID. The diagram represents **all the content** for all product records in an organization, and each item bears the asset ID of the product whose record it is.

3. Product Lifetime is the event type; a specific product reaching end of life is an event. When an event of that event type occurs—in this case, when a product reaches its end of life—you create an event that specifies:

    - An asset ID (for SharePoint and OneDrive documents)
    - Keywords (for Exchange items). In this example, the organization uses a product code in messages containing product records, so the keyword for Exchange items is functionally the same as the asset ID for SharePoint and OneDrive documents.
    - The date when the event occurred. This date is used as the start of the retention period. This date can be the current, a past, or a future date.

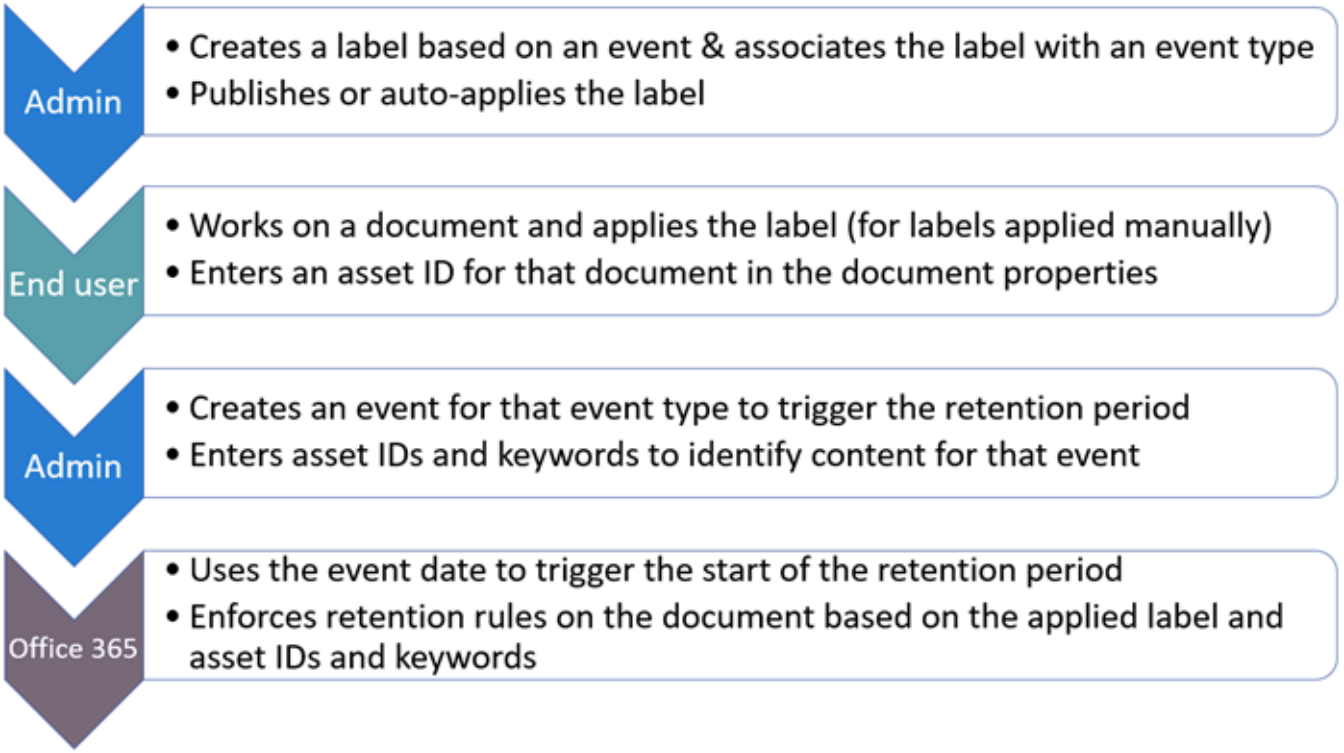
4. After you create an event, that event date is synchronized to all the content that has a retention label of that event type and that contains the specified asset ID or keyword. Like any retention label, this synchronization can take up to seven days. In the previous diagram, all the items circled in red have their retention period triggered by this event. In other words, when this product reaches its end of life, that event triggers the retention period for that product's records.

It's important to understand that if you don't specify an asset ID or keywords for an event, **all content** with a retention label of that event type will have its retention period triggered by the event. This means that in the previous diagram, all content would start being retained. It's unlikely this outcome is what you intended.

Finally, remember that each retention label has its own retention settings. In this example, they all specify 10 years, but it's possible for an event to trigger retention labels where each label has a different retention period.

How to set up event-based retention

High-level workflow for event-based retention:



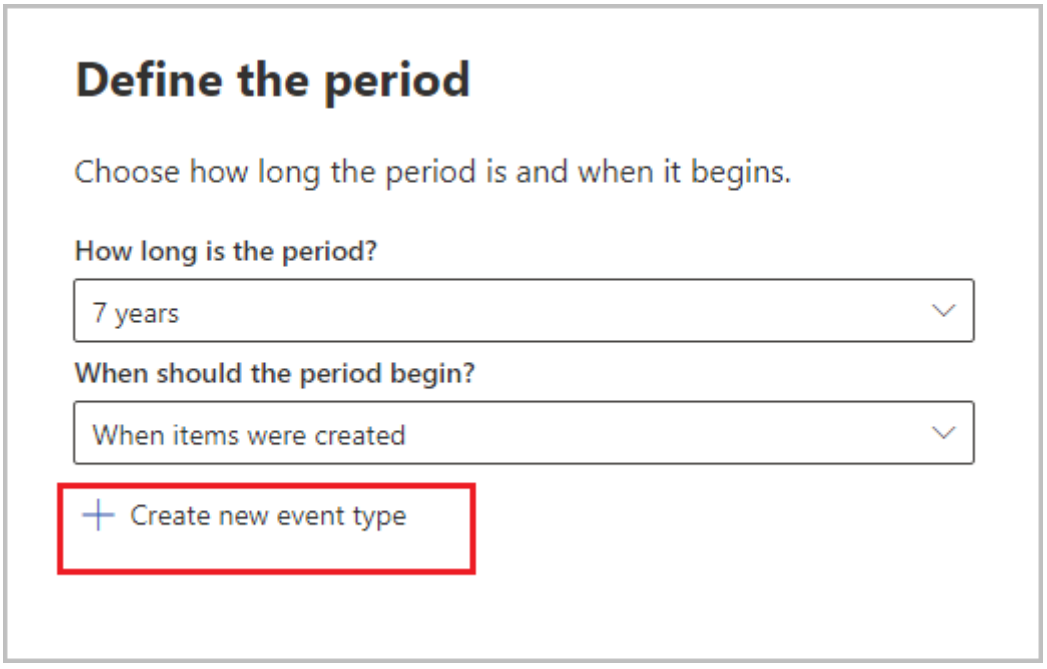
> **Tip**
>
> See Use retention labels to manage the lifecycle of documents stored in SharePoint for a detailed scenario about using managed properties in SharePoint to auto-apply retention labels and implement event-based retention.

Step 1: Create a label whose retention period is based on an event

To create and configure your retention label, see the instructions for Create retention labels for records management. But specific to event-based retention:

- On the **Define label settings** page when you create the retention label, make sure you select **Retain items forever or for a specific period**. Then:

  On the **Define the period** page, after you've set the time period, select one of the default event types from the dropdown list for **When should the period begin?**. Or, create your own event type by selecting **Create new event type** and follow the configuration prompts:



An event type is simply a general description of an event that you want to associate with a retention label.

The default event types have **(event type)** after their name in the dropdown list for easier identification, and you can also see and create event type from the **Records management** > **Events** tab > **Manage event types**.

Event-based retention requires retention settings that:

- Retain the content.
- Delete the content automatically or trigger a disposition review at the end of the retention period.

Event-based retention is typically used for content that's declared a record, so now is a good time to check whether you also need to select the option that marks content as a record.

If you're using an existing event type rather than creating a new event type, skip to step 3.

> ⓘ **Note**
>
> After you choose an event type and save the retention label, the event type can't be changed.

Step 2: Create a new event type for your label

For the retention settings, if you selected **Create new event type**, enter a name and description for your event type. Then select **Next**, **Submit**, and **Done**.

Back on the **Define the period** page, for **When should the period begin?**, use the dropdown list to select the event type that you created.

Step 3: Publish or auto-apply the event-based retention labels

Just like any retention label, you need to publish or auto-apply an event-based label, for it to be manually or automatically applied to content:

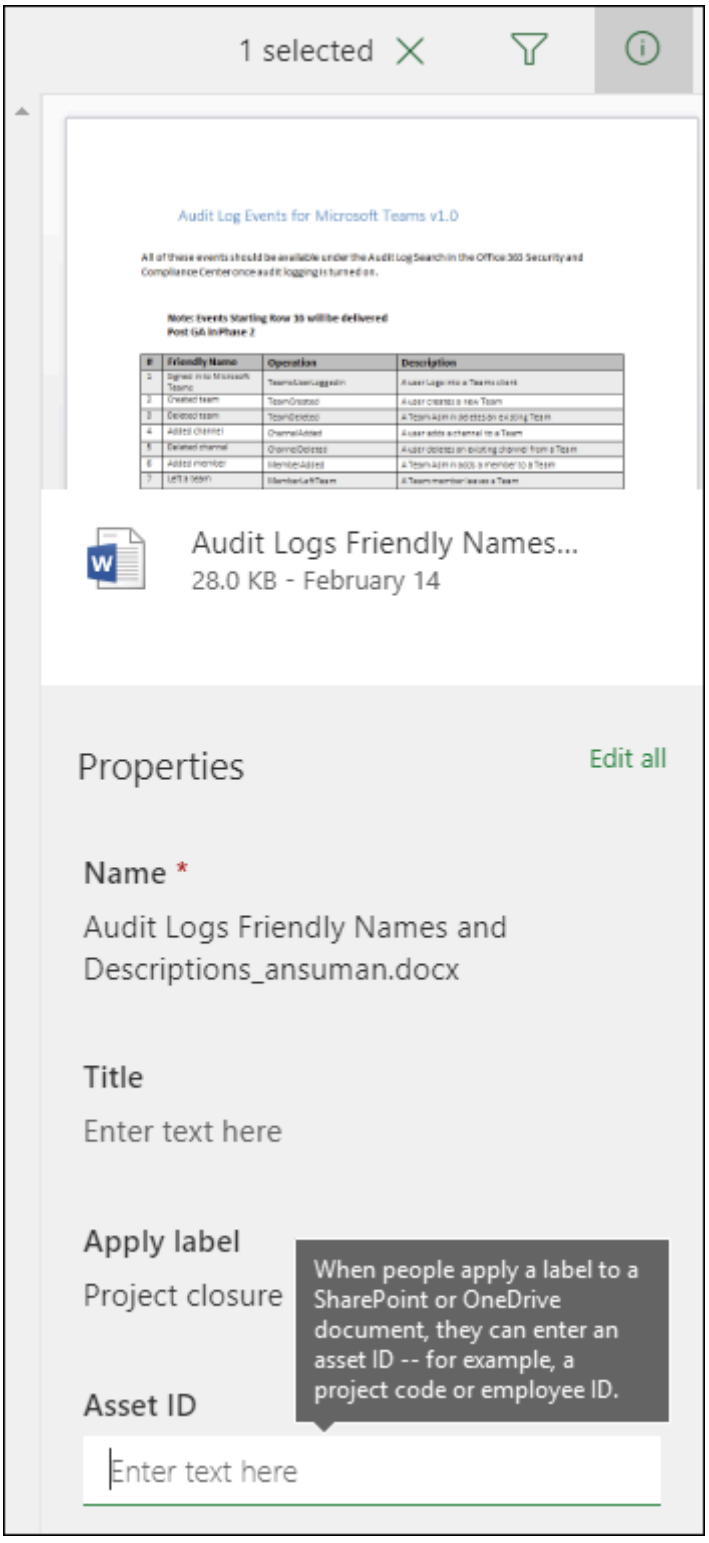
- Publish retention labels and apply them in apps
- Apply a retention label to content automatically

Step 4: Enter an asset ID

After an event-based label is applied to content, you can enter an asset ID for each item. For example, your organization might use:

- Product codes that you can use to retain content for only a specific product.
- Project codes that you can use to retain content for only a specific project.
- Employee IDs that you can use to retain content for only a specific person.

Asset ID is simply another document property that's available in SharePoint and OneDrive. Your organization might already use other document properties and IDs to classify content. If so, you can also use those properties and values when you create an event—see step 6 that follows. The important point is that you must use some *property:value* combination in the document properties to associate that item with an event type.

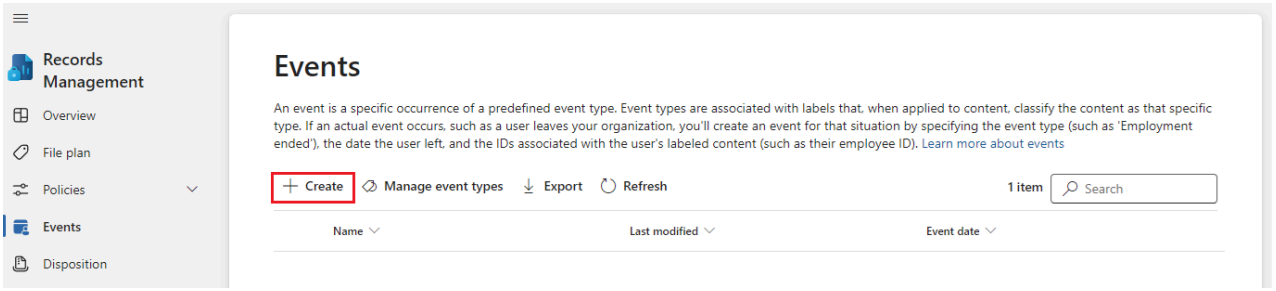


Step 5: Create an event

When a particular instance of that event type occurs, such as a product reaches its end of life, create an event in the Microsoft Purview portal:

1. Sign in to the Microsoft Purview portal > **Solutions** > **Records Management** > **Events**. **Events** tab.

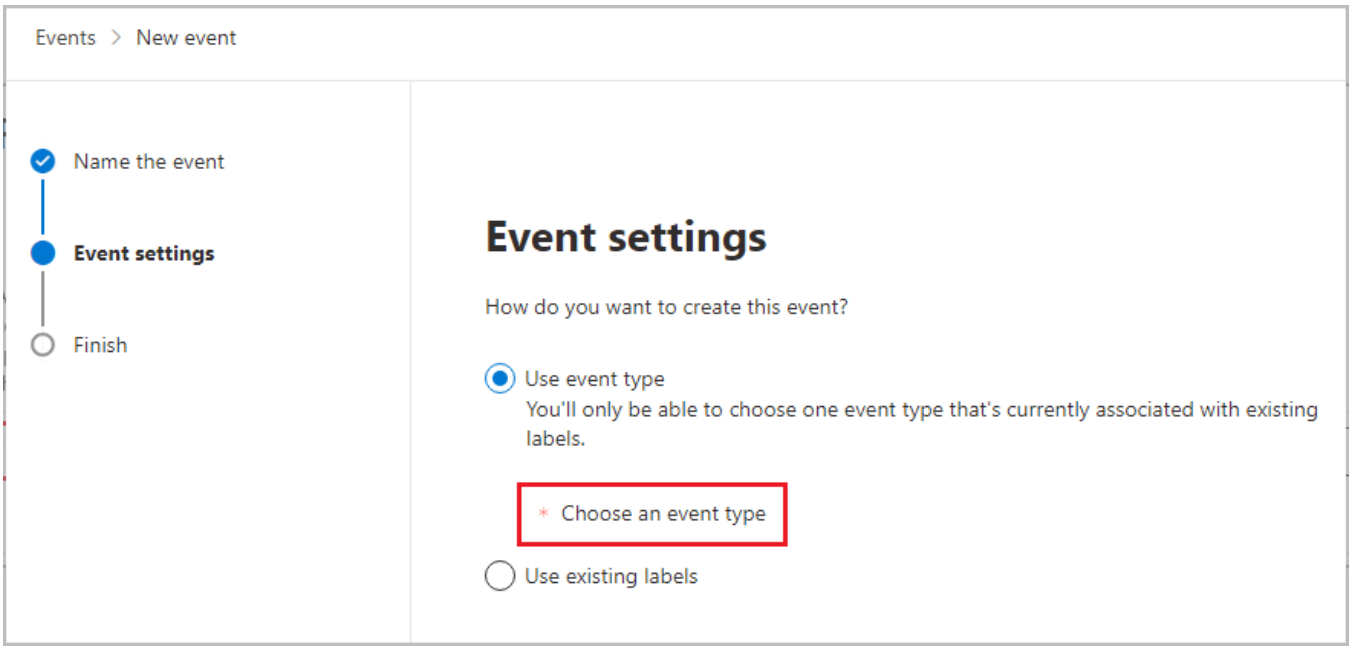
2. Select + **Create** to create an event. You trigger the event by creating it, here. For example:



Up to 1,000,000 events are supported per tenant.

Step 6: Choose the same event type used by the label in step 2

When you create the event, choose the same event type specified in the retention label settings in step 2. For example, if you selected **Product Lifetime** as your event type for the label settings, select **Product Lifetime** when you create the event. Only content with retention labels applied to it of that event type will have its retention period triggered.



Alternatively, if you need to create an event for multiple retention labels that have different event types, select the **Choose Existing Labels** option. Then, select the labels that are configured for the event types you want to associate with this event.

Step 7: Enter keywords or a query, or asset ID for SharePoint and OneDrive

Now you narrow the scope of the content. You do this by specifying keywords or a query. For SharePoint and OneDrive content, you can also do this by specifying asset IDs.

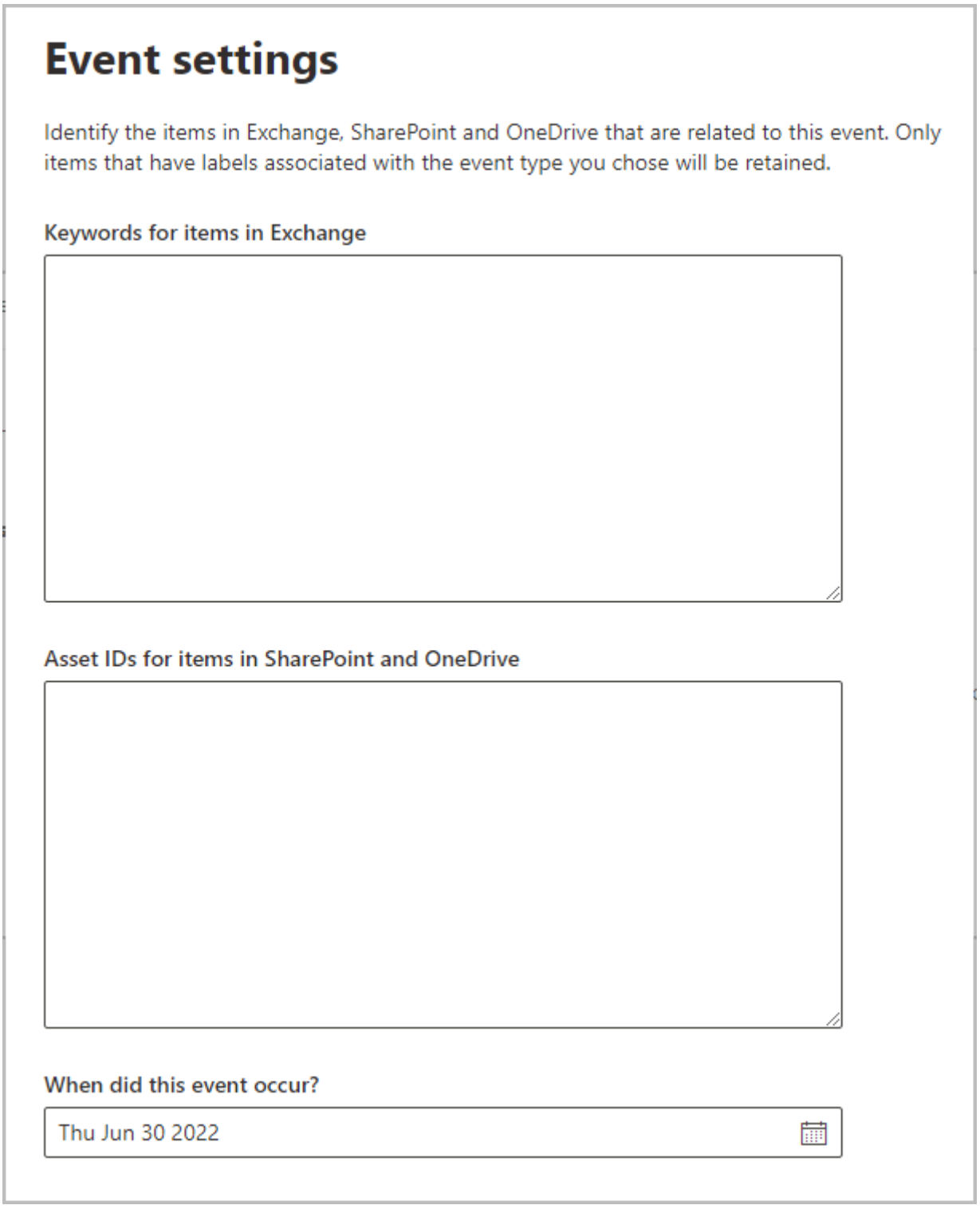
Queries use Keyword Query Language (KeyQL). For more information about the query syntax, see Keyword Query Language (KeyQL) syntax reference. For more information about the searchable properties that you can use, see Keyword queries and search conditions for Content Search.

For asset IDs, retention will be enforced only on content with the specified *property:value* pair. For example, if you're using the Asset ID property, enter `ComplianceAssetID:<value>` in the box for asset IDs shown in the following picture.

If an asset ID isn't entered, all content with labels of that event type get the same retention date applied to them.

Your organization might have applied other properties and IDs to the documents related to this event type. For example, if you need to detect a specific product's records, the ID might be a combination of your custom property ProductID and the value "XYZ". In this case, you'd enter `ProductID:XYZ` in the box for asset IDs shown in the following picture.

Finally, choose the date when the event occurred; this date is used as the start of the retention period. After you create an event, that event date is synchronized to all the content with a retention label of that event type, asset ID, and keywords or queries. As with any retention label, this synchronization can take up to seven days.



After creating an event, the retention settings take effect for the content that's already labeled and indexed. If the retention label is added to new content after the event is created, you must create a new event with the same details.

Deleting an event doesn't cancel the retention settings that are now in effect for the content that's already labeled. Currently, you can't cancel events after they're triggered.

Use Content Search to find all content with a specific label or asset ID

After retention labels are assigned to content, you can use content search to find all content that has a specific retention label or that contains a specific asset ID:

- To find all content with a specific retention label, choose the **Retention label** condition, and then enter the complete label name or part of the label name and use a wildcard.
- To find all content with a specific asset ID, enter the **ComplianceAssetID** property and a value, using the format `ComplianceAssetID:<value>`.

For more information, see Keyword queries and search conditions for Content Search.

Automate events by using PowerShell

You can use a PowerShell script to automate event-based retention from your business applications. The PowerShell cmdlets available for event-based retention:

- Get-ComplianceRetentionEventType
- New-ComplianceRetentionEventType
- Remove-ComplianceRetentionEventType
- Set-ComplianceRetentionEventType
- Get-ComplianceRetentionEvent
- New-ComplianceRetentionEvent

to help identify other cmdlets to create retention labels and their policies, see PowerShell cmdlets for retention policies and retention labels.

Automate events by using Graph API

You can use Microsoft Graph Records Management APIs to automatically create the events and event types that trigger the start of the retention time.

> **Note**
>
> In the past, you could use REST APIs to automate the events. This method is now deprecated and no longer works. Instead, use Microsoft Graph API.

Last updated on 02/27/2026

Customize what happens at the end of the retention period

Summarize this article for me

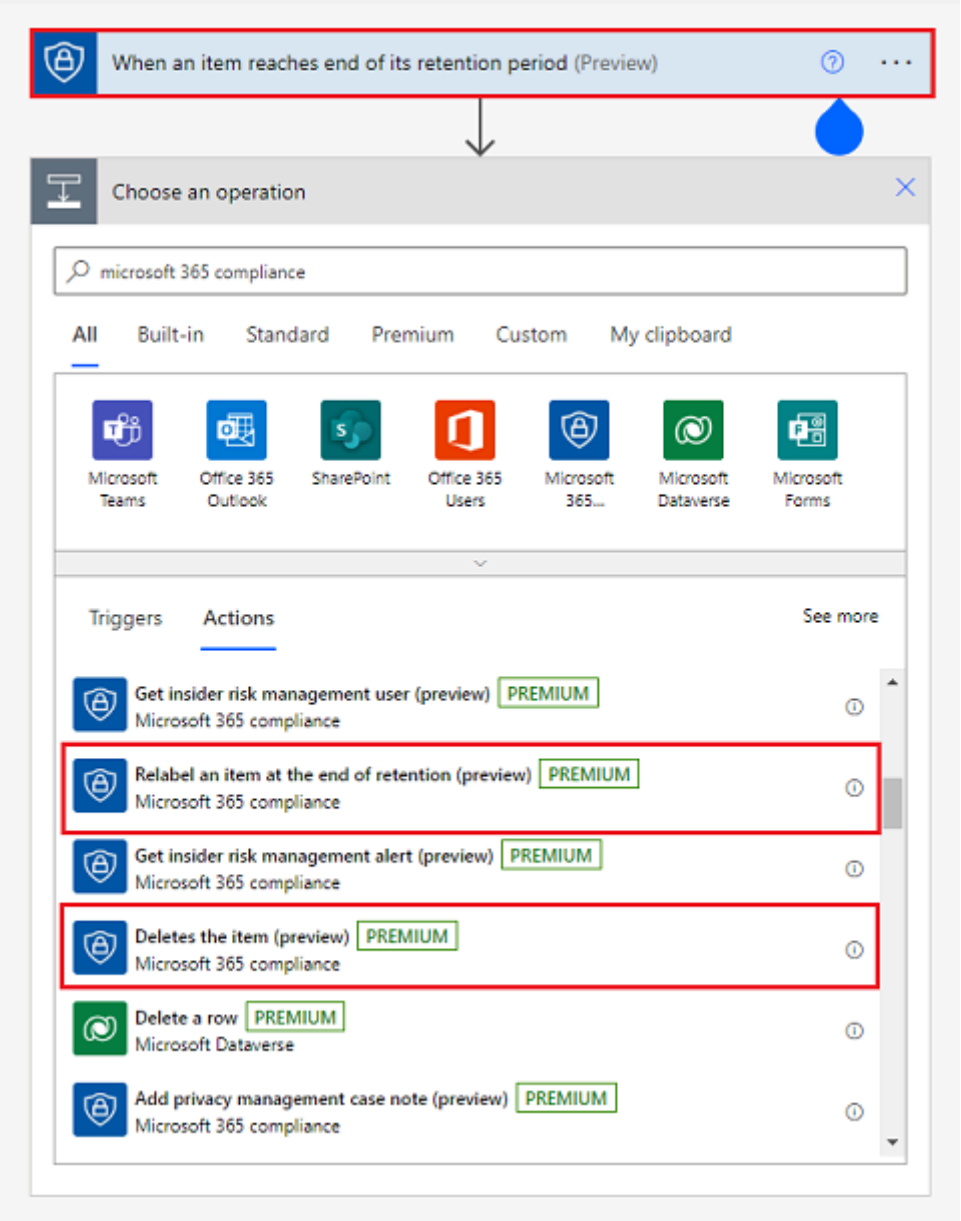
> *Microsoft 365 licensing guidance for security & compliance*.

When you configure a retention label to retain items for a specific period, you specify what action to take at the end of that retention period. You can choose from the built-in actions of permanently deleting the item, relabeling the item, deactivating the label, and starting a disposition review.

However, if you choose the option to run a Power Automate flow, this label configuration supports custom actions such as:

- Send an email notification to a records manager, or to the person who created or last modified a document
- Move a document item to a SharePoint location where you centrally store records

You can also select the Microsoft 365 compliance label-specific actions of relabel the item or permanently delete it:



For example, a flow can move the item, relabel it, and then send an email notification.

> **Note**
>
> You can also use the relabel action to remove the existing retention label when you select this action but don't specify a replacement label (leave the label name blank). If the existing retention label marked the item as a record, the record must be locked for the label to be removed.

The actions also allow integration with other apps and services. For example, you might have another solution that starts your own customized disposition review process when an email notification is received. Or your flow adds a new entry into a table and this change automatically starts an existing records management procedure.

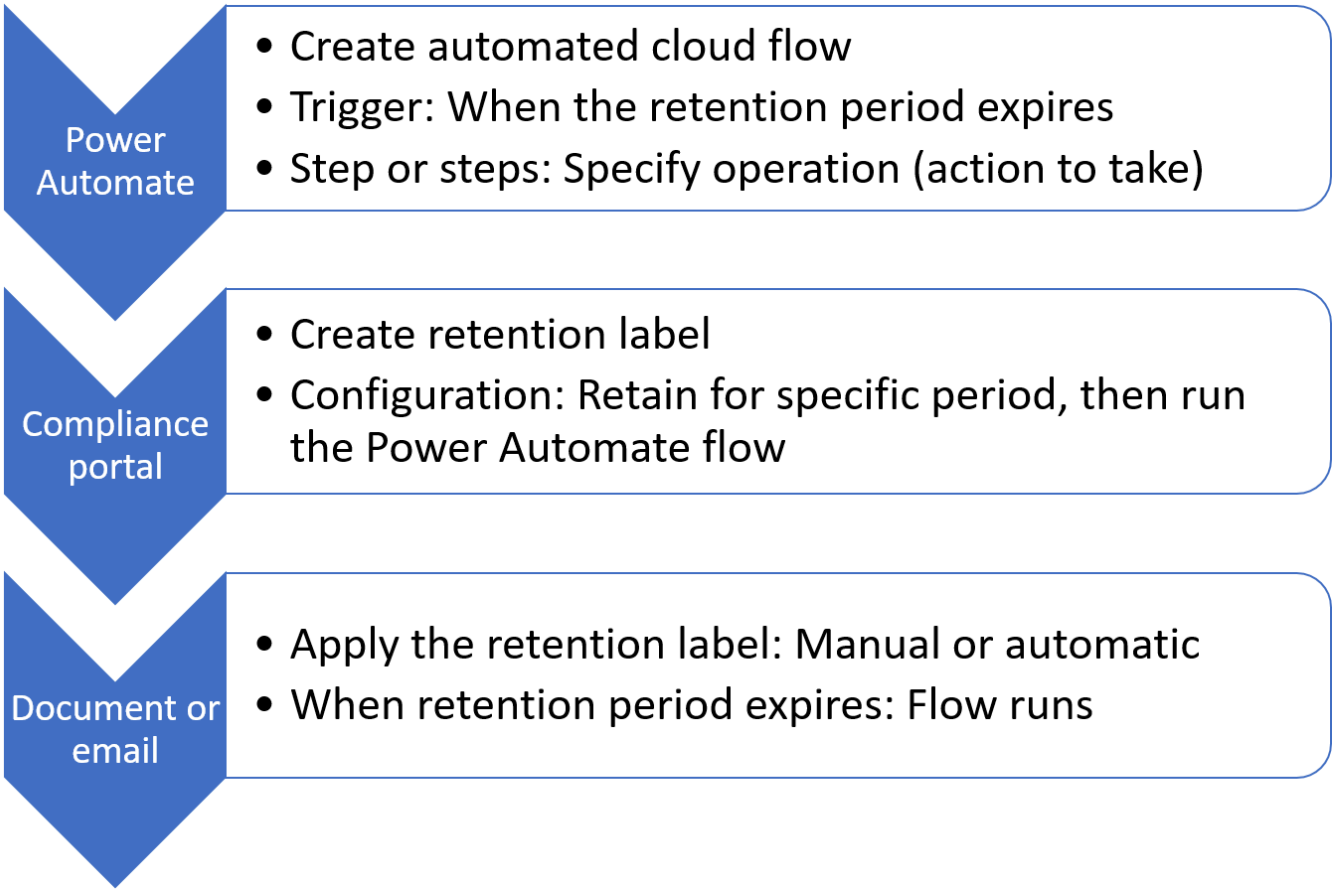
It's outside the scope of this article to document all the ways in which you can configure a flow and integrate with other solutions. However, to learn more about Power Automate, see Get

started with Power Automate and Overview of cloud flows. In this article, you learn about:

- Overview of using retention labels with a Power Automate flow
- Limitations
- Power Automate dependency
- How to configure a retention label to run a Power Automate flow
- When the flow is triggered

Overview of using retention labels with a Power Automate flow

The following processes show the high-level overview of how retention labels work with a Power Automate flow:



In Power Automate, create an automated cloud flow from blank that uses the trigger **When an item reaches the end of its retention period**. Then, add one or more steps and specify the operations that run your choice of actions.

Next, in the Microsoft Purview portal, create a new retention label with the following configuration:

- Retain the item for a specified period of time
- At the end of the retention period, run the flow that you just created

Apply the label by using either of the following methods:

- Publish retention labels and apply them in apps
- Apply a retention label to content automatically

When the retention period expires, your configured flow runs.

> **Note**
>
> In a Power Automate flow, you can also select the action **Apply a retention label on the item**. This action is independent from this scenario that applies actions at the end of the retention period. Instead, it's the equivalent to manually applying a retention label. The label is applied immediately.

Add resilience to your flow

Although optional, we recommend you always add the **Trigger a new instance of the flow if an item can't be located** action before **Relabel an item at the end of retention** and **Deletes an item at the end of retention**.

This additional action is required for files that you move or rename after an instance of the flow is triggered but your specified compliance actions aren't completed. Without it, a renamed or moved file results in the flow failing because it can't find the file.

Although renaming and moving files after an instance of the flow is triggered are an unusual occurrence, adding this recommended action improves the reliability of the compliance actions succeeding because the current path and name of the file is retrieved automatically just before the selected actions run.

Limitations for this scenario

- You must create a new retention label to run a flow at the end of the retention period. You can't edit an existing retention label for this configuration.
- By default, the account that creates the Power Automate flow must be the same account that creates the retention label. If you share the flow in Power Automate, the people you share it with can also select the flow when they create a retention label.
- In Power Automate, you must use the option **My flows** to create the flow. Currently, creating a flow under **Solutions** isn't supported.

- If you select the relabel action in the flow, the retention label that you select as the replacement can't be configured to mark the item as a regulatory record. It can be configured to mark the item as a record, or not mark the item as a record.
- If your chosen flow uses an environment type other than Default, you must use PowerShell to create the retention label, and use the *FlowId* parameter.

Power Automate dependency

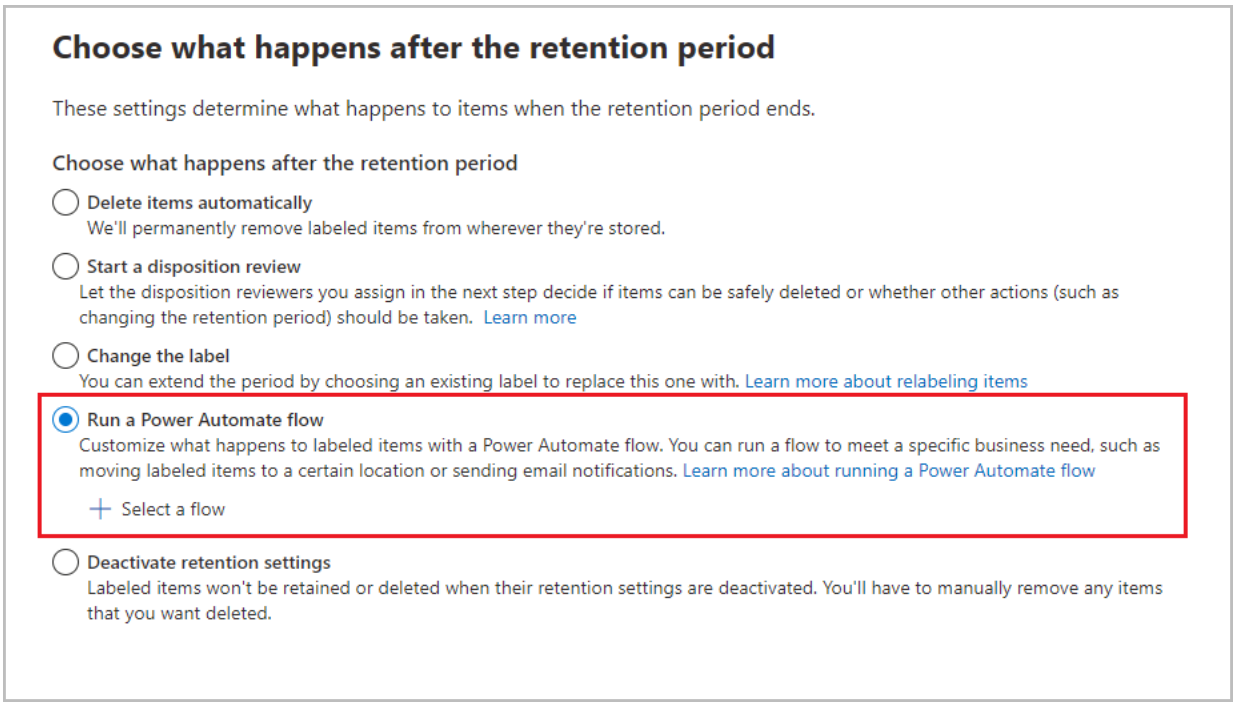
Power Automate is a workflow service that automates actions across applications and services. Specific to running a Power Automate flow at the end of the retention period:

- You must have a Power Automate plan that includes premium connectors, separate from your Microsoft 365 compliance plan.
- The Power Automate flow must be an automated cloud flow that you create from blank, and configure to use the trigger **When an item reaches the end of its retention period**.

For more information about how to use and configure the dependent automated cloud flows, see the Microsoft Power Automate Documentation.

How to configure a retention label to run a Power Automate flow

1. In the Microsoft Purview portal, create a new retention label with your preferred name and settings. For more information, see Use file plan to create and manage retention labels.
2. For this scenario, complete the following steps:
   a. On the **Define label settings** page, select **Retain items indefinitely or for a specific period** and specify the retention period.
   b. On the **Choose what happens after the retention period** page, select **Run a Power Automate flow**, and then **Select a flow**:



c. From the **Select a Power Automate flow** flyout pane, you'll see any automated cloud flows that you created (or somebody has shared with you) and that have the **When an item reaches the end of its retention period** trigger.

   Select the flow you want, and complete the configuration to create the retention label.

A retention label can be configured for only one Power Automate flow, but the same flow can be used by more than one retention label.

When the flow is triggered

Just as it can take up to seven days to automatically apply a retention label or display it so you can apply it, so you should allow up to seven days after the expiration period expires before the Power Automate flow is triggered. The initiation of the flow can then take a few hours.

For the Microsoft 365 compliance actions of delete and relabel, then allow up to seven days for these actions to complete. Specific for Exchange and the Microsoft 365 compliance actions of delete, allow up to 14 days.

Both the flow itself and the Microsoft 365 compliance actions in the flow are queued so that if there are any network or service delays, the flow and actions are automatically retried.

The principles of retention still apply to retention labels that are configured to run a flow at the end of the retention period, in that an item is always retained for the longest period if there is a conflict with one or more retention policies. However, the flow is always triggered at the end of the label's retention period.

Next steps

To help you track the retention labels with this configuration:

- Monitoring retention labels
- Using Content Search to find all content with a specific retention label
- Auditing retention actions

To confirm there aren't any failures or errors with your flow, go to the Power Automate dashboard, select **My flows**, and then select your flow to review the 28-day run history.

Last updated on 02/10/2026

Publish retention labels and apply them in apps

> *Microsoft 365 licensing guidance for security & compliance.*

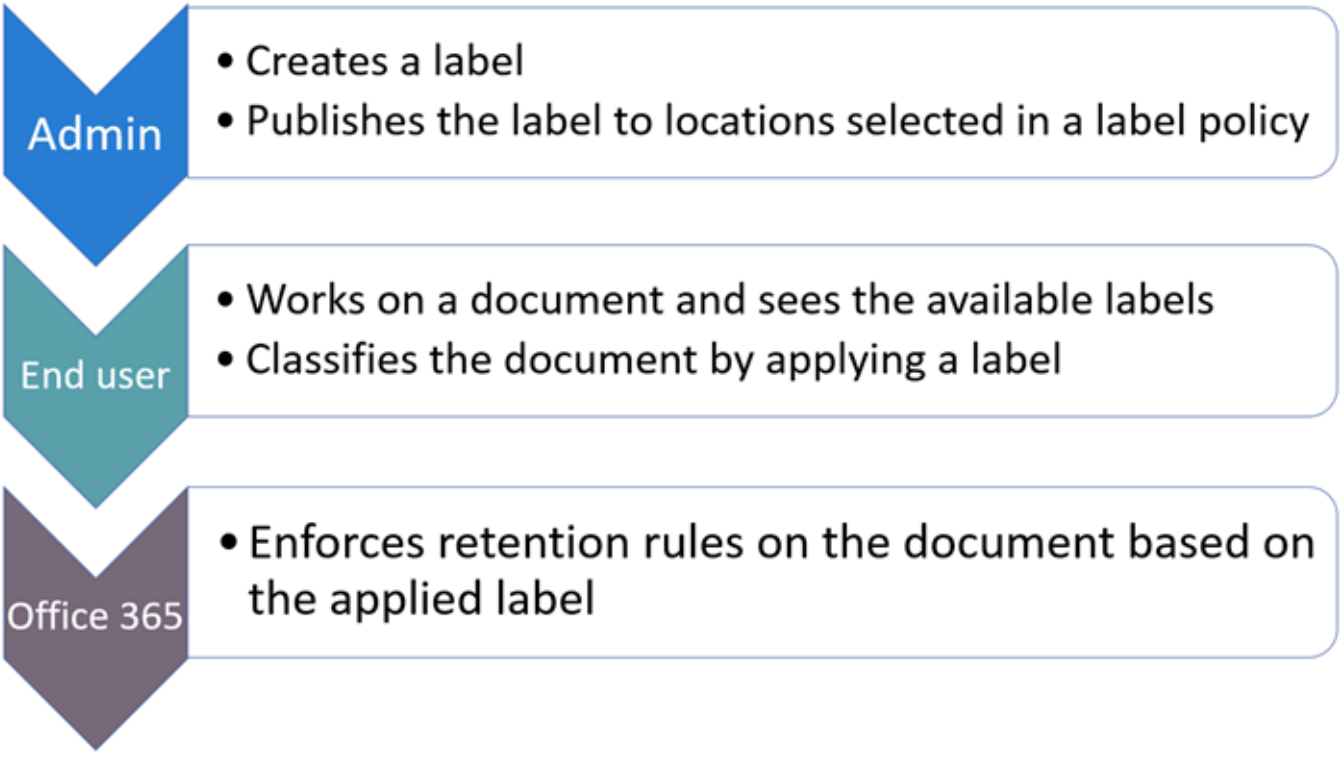
> **Note**
>
> This scenario is supported for all retention label configurations, including regulatory records.

Use the following information to help you publish retention labels, and then apply them to documents and emails.

Retention labels help you retain what you need and delete what you don't at the item level (document or email). They're also used to declare an item as a record as part of a records management solution for your Microsoft 365 data.

Making retention labels available to people in your organization so that they can classify content is a two-step process:

1. Create the retention labels.
2. Publish the retention labels by using a retention label policy.



Use the following instructions for the two admin steps.

Before you begin

To make sure you have permissions to create and edit retention labels and their policies, see the permissions information for records management or data lifecycle management, depending on the solution you're using.

Make sure you have created the retention labels you want to apply to items.

How to publish retention labels

Decide before you create your retention label policy whether it will be **adaptive** or **static**. For more information, see Adaptive or static policy scopes for retention. If you decide to use an adaptive policy, you must create one or more adaptive scopes before you create your retention label policy, and then select them during the create retention label policy process. For instructions, see Configuration information for adaptive scopes.

1. Sign in to the Microsoft Purview portal and then:

   If you're using records management: **Solutions** > **Records Management** > **Policies** > **Label policies**

   If you're using data lifecycle management: **Solutions** > **Data Lifecycle Management** > **Policies** > **Label policies**

2. Select **Publish labels** and follow the prompts to create the retention label policy. Be careful what name you choose for the policy, because this can't be changed after the policy is saved.

3. Use the link to select the retention labels to publish, and then select **Next**.

4. For the **Assign admin units** page, keep the default of **Full directory**. Currently, admin units aren't supported for this policy.

5. For the **Choose the type of retention policy to create** page, select **Adaptive** or **Static**, depending on the choice you made from the Before you begin instructions. If you haven't already created adaptive scopes, you can select **Adaptive** but because there won't be any adaptive scopes to select, you won't be able to finish the configuration with this option.

6. Depending on your selected scope:

   - If you chose **Adaptive**: On the **Choose adaptive policy scopes and locations** page, select **Add scopes** and select one or more adaptive scopes that have been created. Then, select one or more locations. The locations that you can select depend on the

scope types added. For example, if you only added a scope type of **User**, you'll be able to select **Exchange email** but not **SharePoint sites**.

- If you chose **Static**: On the **Choose locations** page, toggle on or off any of the locations. For each location, you can leave it at the default to apply the policy to the entire location, or specify includes and excludes

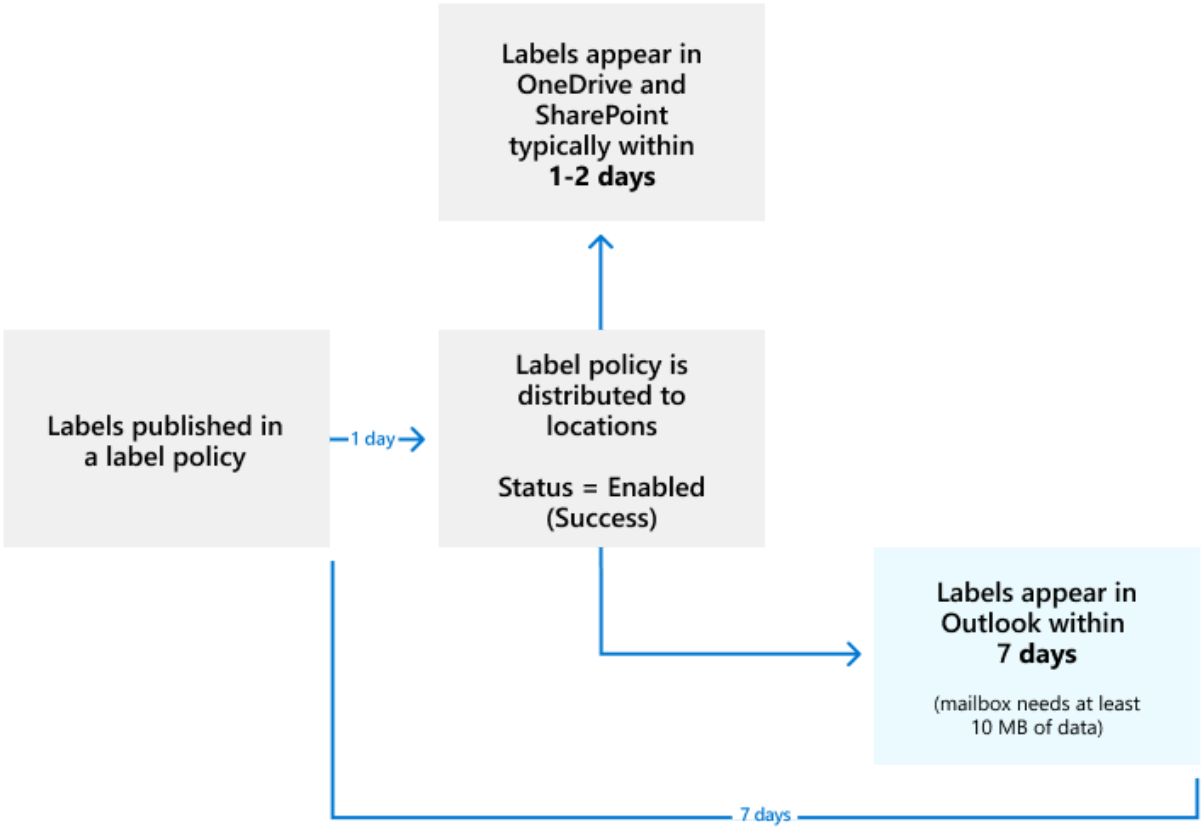
For information about the location choices, see Locations.

To edit an existing retention label policy (the policy type is **Publish**), select it, and then select the **Edit** option to start the **Edit retention policy** configuration.

When retention labels become available to apply

If you publish retention labels to SharePoint or OneDrive, those labels typically appear for users to select within one day. However, allow up to seven days.

If you publish retention labels to Exchange, it can take up to seven days for those retention labels to appear for users. As with all retention settings for Exchange, the mailbox must contain at least 10 MB of data.



If the labels don't appear after seven days, check the **Status** of the label policy by selecting it from the **Label policies** page in the Microsoft Purview portal. If you see **(Error)** included in the status and in the details for the locations see a message that it's taking longer than expected to deploy the policy or to try redeploying the policy, try running the Set-AppRetentionCompliancePolicy or Set-RetentionCompliancePolicy PowerShell command to retry the policy distribution:

1. Connect to Security & Compliance PowerShell.
2. Run one of the following commands:

   - For the policy locations **Teams private channel messages** (applicable pre-migration only), **Viva Engage user messages** and **Viva Engage community messages**:

     PowerShell

     ```
     Set-AppRetentionCompliancePolicy -Identity <policy name> -RetryDistribution
     ```

   - For all other policy locations, such as **Exchange mailboxes**, **SharePoint classic and communication sites**, **Teams channel messages** etc.:

     PowerShell

     ```
     Set-RetentionCompliancePolicy -Identity <policy name> -RetryDistribution
     ```

How to check on the status of retention labels published to Exchange

In Exchange Online, retention labels are made available to end users by a process that runs every seven days. By using PowerShell, you can see when this process last ran and therefore identify when it will run again.

1. Connect to Exchange Online PowerShell.
2. Run these commands.

   PowerShell

   ```
   $logProps = Export-MailboxDiagnosticLogs <user> -ExtendedProperties
   ```

   PowerShell

   ```
   $xmlprops = [xml]($logProps.MailboxLog)
   ```

PowerShell

```
$xmlprops.Properties.MailboxTable.Property | ? {$_.Name -like "ELC*"}
```

In the results, the `ELCLastSuccessTimeStamp` (UTC) property shows when the system last processed your mailbox. If it hasn't happened since the time you created the policy, the labels aren't going to appear. To force processing, run `Start-ManagedFolderAssistant -Identity <user>`.

If labels aren't appearing in Outlook on the web and you think they should be, make sure to clear the cache in your browser (CTRL+F5).

How to apply published retention labels

Use the following sections to learn how published retention labels can be applied in apps:

- Manually apply retention labels
- Default labels for SharePoint and Outlook
- Automatically apply a retention label to email by using Outlook rules

In addition, when you use Microsoft Syntex and publish retention labels to SharePoint locations, you can apply a retention label to a model so that identified files are automatically labeled.

After content is labeled, see the following information to understand when the applied label can be removed or changed: Only one retention label at a time.

Manually apply retention labels

Users can manually apply retention labels from the following locations:

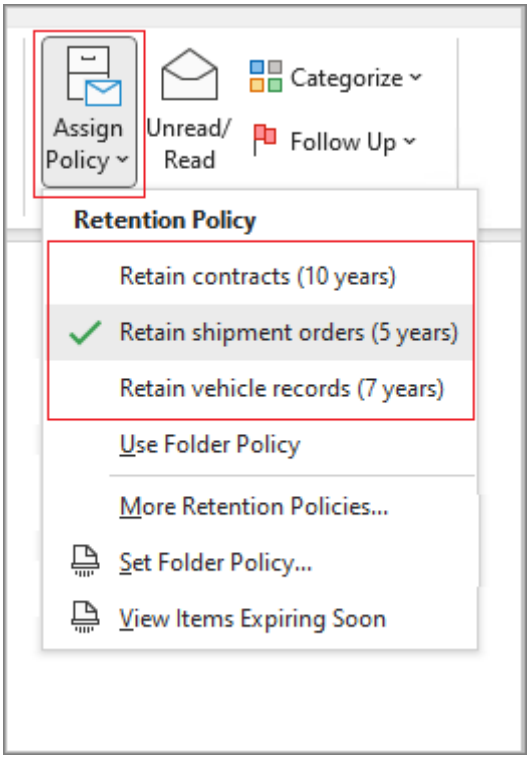
- Outlook
- OneDrive and SharePoint
- Teams site that's group-connected
- Loop and Copilot Pages

Use the following tabs to understand how to manually apply retention labels for each location:

Outlook

Your published retention labels display in Outlook alongside any legacy MRM retention tags that are assigned to the mailbox. Although you see references to retention policies, users can't select your Microsoft 365 retention policies in the Outlook apps.

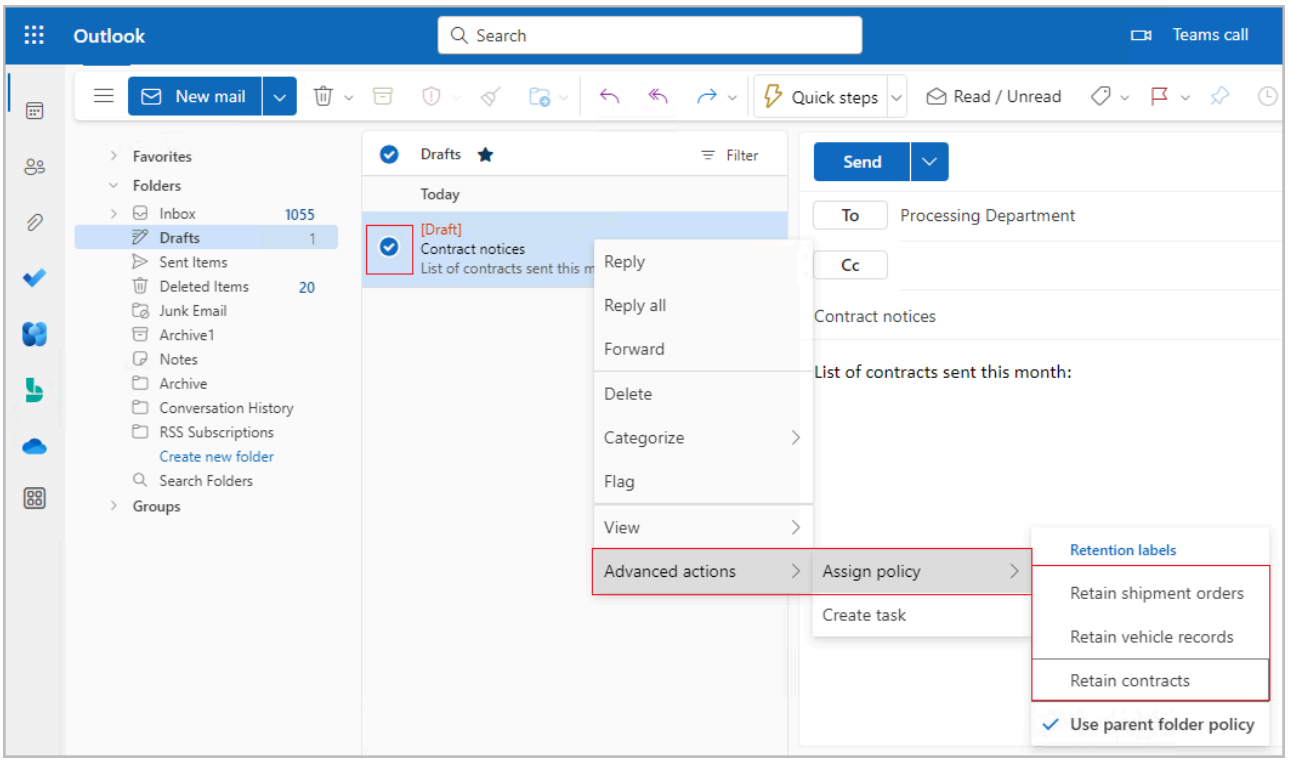
To label an item in the Outlook desktop client, select the item. On the **Home** tab on the ribbon, select **Assign Policy**, and then choose the retention label. For example:



If you don't immediately see the option to assign policy, look for the **Tags** group on the ribbon.

You can also right-click an item from a folder, such as the **Inbox** folder for emails received or the **Drafts** folder for an email to be sent. Then, select **Assign Policy** in the context menu, and then choose the retention label. When you select multiple items, you can use this method to apply the same retention label to multiple items at once.

To label an item in Outlook on the web, first right-click the item from a folder. Then, select **Advanced actions** > **Assign policy**, and then choose the retention label. For example:



As with Outlook desktop, you can also use this labeling method if you multi-select items.

After the retention label is applied, you can view that retention label at the top of the item, and the calculated expired date. For example:

Label assigned to email in Outlook on the web.

For more information about the expiry date displayed to users, see User notification of expiry date.

Default labels for SharePoint and Outlook

After a retention label is published to SharePoint and Outlook, users can apply it as a default retention label so that it's inherited by all unlabeled items. Although the same label is applied, each item will be retained and deleted separately, according to the start of the retention period setting in the label.

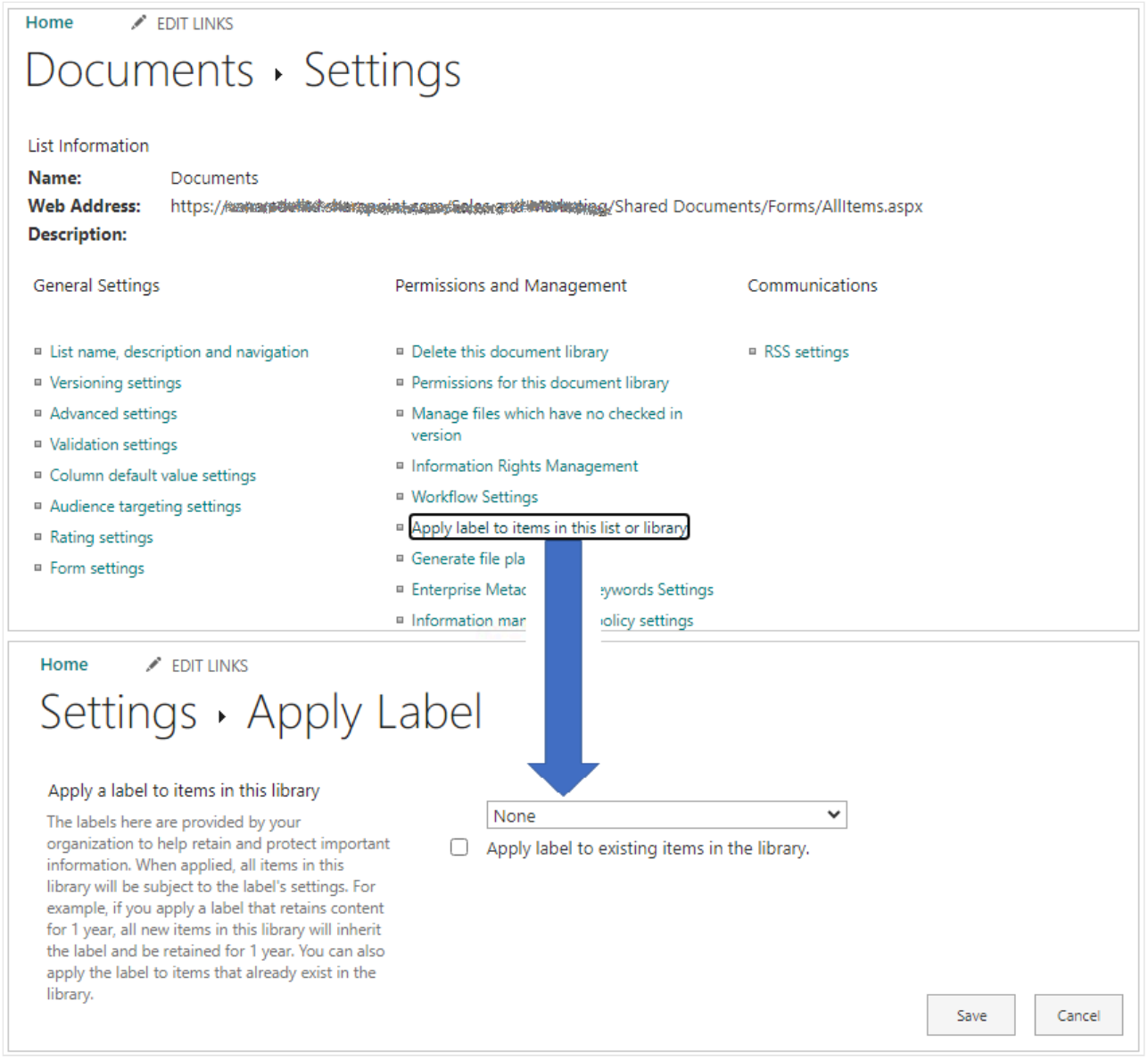
When you use default retention labels, there are some scenarios that can result in the inherited label being replaced or removed. See each tab for details about the label behavior.

Default label for SharePoint

You can apply a default retention label to all content in a SharePoint library, folder, or document set. Documents in that location then inherit your selected default retention label.

For a document library, the default label configuration is done on the **Library settings** page for a document library. When you choose the default retention label, you can also choose to apply it to existing items in the library.

For example, if you have a retention label for marketing materials, and you know a specific document library contains only that type of content, you can make the **Marketing Materials** retention label the default label for all documents in that library.



Label behavior when you use a default label for SharePoint

For standard retention labels that you apply as a default retention label to a library, folder, or document set:

- The default label is applied to all new unlabeled items in the container. However, this does not apply to subsequently migrated files.
- You can re-apply a default label to a location and select **Apply to existing items** to apply a default label to subsequently migrated items.

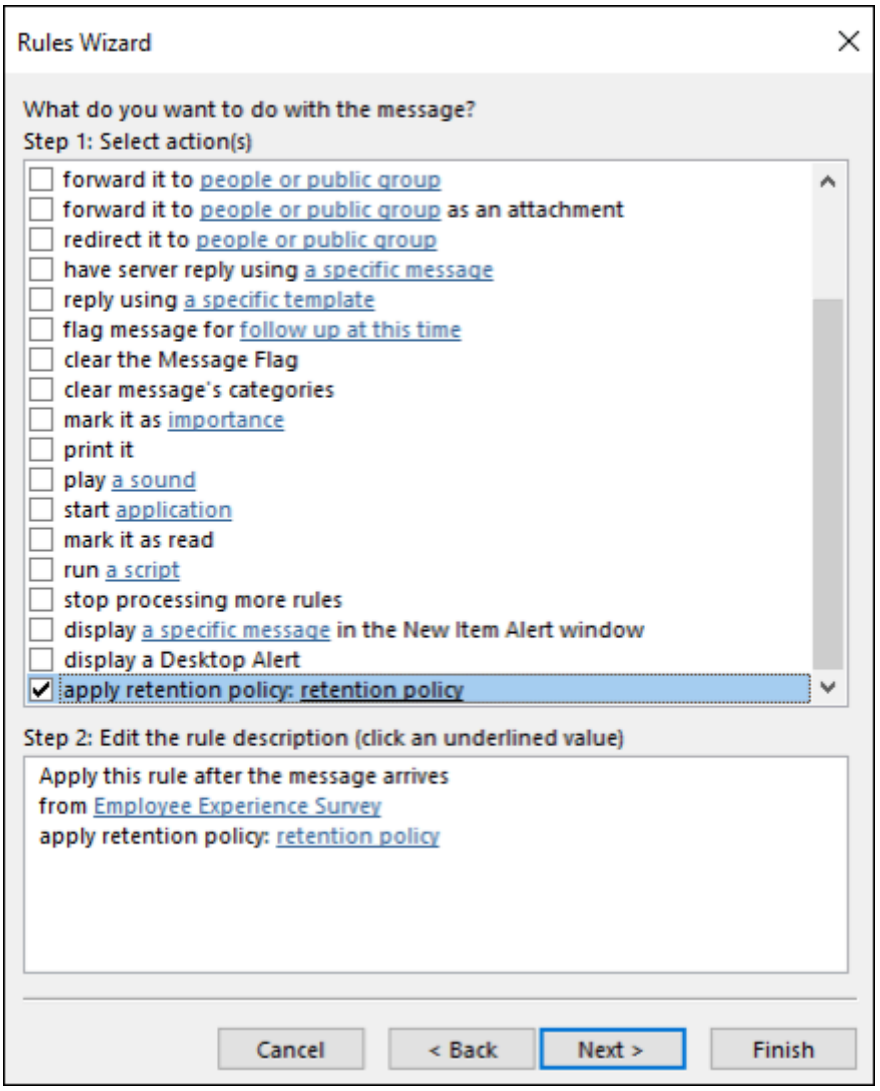
- For folders, the inheritance flows to any child folders and items inherit the label from their nearest folder.
- If you selected the option to apply the default label to existing items: Items that are already labeled retain their retention label, unless it was applied by a different default label.
- If you change the default retention label for the container: Existing retention labels applied to items in that container are changed only if you selected the option to apply the default label to existing items and those labels were applied by a default label.
- If you remove the default retention label for the container: Items retain their labels.
- If you move an item with a default retention label applied from one container to another container: The item keeps its existing default retention label, even if the new location has a different default retention label. Only if you then change the default label for this new location can the moved item inherit the default label from its current location.

When labels are applied that aren't standard retention labels but mark items as records (or regulatory records), these labels can only be manually changed or removed.

Automatically apply a retention label to email by using Outlook rules

In Outlook, you can create rules to apply a retention label. For example, you can create a rule that applies a specific retention label to all messages sent to or from a specific distribution group.

To create a rule, right-click an item, select **Rules** > **Create Rule** > **Advanced Options** > **Rules Wizard**. Specify the condition or conditions for the first step, and then **apply retention policy** for the action:



Although you see a reference to retention policies, retention labels are displayed and can be selected.

Updating retention labels and their policies

If you edit a retention label or a retention label policy, and the retention label or policy is already applied to content, your updated settings will automatically be applied to this content in addition to content that's newly identified.

Some settings can't be changed after the label or policy is created and saved, which include:

- Names for retention labels and their policies, the scope type (adaptive or static), and the retention settings except the retention period.
- The option to mark items as a record.

Deleting retention labels

To delete a retention label, all three conditions must apply:

- The label isn't included in any retention label policy
- The label isn't configured for event-based retention
- The label isn't configured to mark items as regulatory records

When all these conditions are met:

- You can always delete a retention label that doesn't mark items as records (sometimes referred to as a "standard retention label"). The deletion succeeds even if the label is applied to items, and the retention label is then removed from these items.

- You can delete a retention label that marks items as records only if the label isn't applied to items. If the label has been applied to items, the deletion fails and you see a link to content explorer to identify the labeled items. It can take up to two days for content explorer to show the items that are labeled. In this scenario, the retention label might be deleted without showing you the link to content explorer.

Locking the policy to prevent changes

If you need to ensure that no one can turn off the policy, delete the policy, or make it less restrictive, see Use Preservation Lock to restrict changes to retention policies and retention label policies.

Troubleshooting retention label policies

If your retention label policies aren't working as expected or you see errors related to these policies, use the following troubleshooting resources:

- Identify errors in Microsoft 365 retention and retention label policies
- Resolve errors in Microsoft 365 retention and retention label policies

Next steps

To help you track the labels applied from your published retention labeling policies:

- Monitoring retention labels
- Using Content Search to find all content with a specific retention label
- Auditing retention actions

Event-based retention is another supported scenario for retention labels. For more information, see Start retention when an event occurs.

Last updated on 10/20/2025

Automatically apply a retention label to retain or delete content

Summarize this article for me

> *Microsoft Purview service description*

> **Note**
>
> This scenario isn't supported for regulatory records or default labels for an organizing structure such as a document set or library in SharePoint, or a folder in Exchange. These scenarios require a published retention label policy.

One of the most powerful features of retention labels is the ability to apply them automatically to content that matches specified conditions. In this case, people in your organization don't need to apply the retention labels. Microsoft 365 does the work for them.

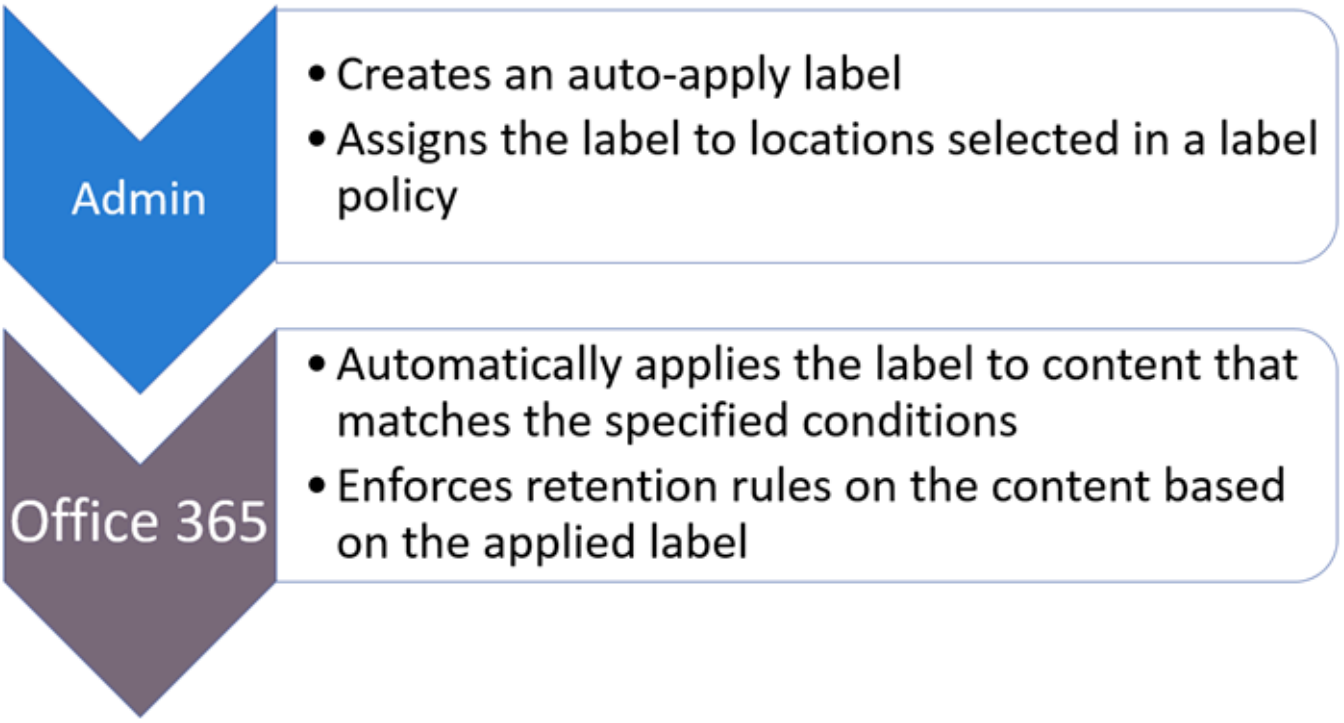
Auto-applying retention labels are powerful because:

- You don't need to train your users on all of your classifications.
- You don't need to rely on users to classify all content correctly.
- Users no longer need to know about data governance policies - they can focus on their work.

You can apply retention labels to content automatically when that content doesn't already have a retention label applied and contains sensitive information, keywords or searchable properties, or a match for trainable classifiers. You can also automatically apply a retention label to cloud attachments that are stored in SharePoint or OneDrive.

> **Tip**
>
> Use searchable properties to identify Teams meeting transcripts or recordings with accompanying transcripts and items that have a sensitivity label applied.

The processes to automatically apply a retention label based on these conditions:



Use the following instructions for the two admin steps.

> ⓘ **Note**
>
> Auto-policies use service-side labeling with conditions to automatically apply retention labels to items. You can also automatically apply a retention label with a label policy when you do the following:
>
> - Apply a retention label to a Microsoft Syntex model
> - Apply a default retention label for SharePoint and Outlook
> - Apply a retention label to email by using Outlook rules
>
> For these scenarios, see Publish retention labels and apply them in apps.

Learn about simulation mode

You can run an auto-labeling policy in simulation mode when it's configured for either of the following conditions:

- Specific types of sensitive information
- Specific keywords or searchable properties that match a query you create

Very similar to simulation mode for auto-labeling policies for sensitivity labels, this addition lets you run the auto-labeling deployment like the WhatIf parameter for PowerShell. You see results reported as if the auto-labeling policy applied your selected label, using the conditions that

you defined. You can then refine your conditions for accuracy if needed, and rerun the simulation.

Simulation mode also lets you gradually increase the scope of your auto-labeling policy before deployment. For example, you might start with a single location, such as a SharePoint site, with a single document library. Then, with iterative changes, increase the scope to multiple sites, and then to another location, such as OneDrive.

Typical workflow for an auto-labeling policy:

1. Create and configure an auto-labeling retention policy.
2. Run the policy in simulation mode, and wait for it to complete.
3. Review the results, and if necessary, refine your policy and rerun simulation. Wait for it to complete again.
4. Repeat step 3 as needed.
5. Deploy in production by turning on the policy.

Unlike simulation mode for automatically applying sensitivity labels:

- Simulation mode is optional, and not required to complete before you turn on the policy. You can even turn on the policy while simulation is still running.
- When simulation completes, the results automatically expire within seven days. Then, to view samples for your policy, you must restart the simulation.

Other considerations for simulation mode for auto-apply retention label policies:

- A maximum of 30 simulation jobs can be active in a 12-hour time period.
- A maximum of 100 item samples can be collected per mailbox.
- If you use adaptive scopes for your policy:
  - A maximum of 20,000 locations (any combination of sites and mailboxes) is supported.
  - Because these scopes use dynamic queries that run daily and can take a few days to fully populate, wait and confirm their membership before you start simulation.
  - For the **Microsoft 365 Group mailboxes & sites** location, items stored in AuxPrimary mailboxes aren't supported.
- You might need to be assigned additional permissions to see the simulation results. For information about the required roles, see the next section, Before you begin.
- Simulation counts all items matching the policy criteria at time of simulation. However, when the policy is turned on, only content that isn't already labeled is eligible for auto-applying retention labels.
- Although auto-labeling for sensitive information types applies to emails sent and received rather than emails stored in mailboxes, simulation for Exchange locations runs against

emails stored in mailboxes. Using historical data lets you more quickly assess the effectiveness of your chosen sensitive information types and configuration.

- For the **Microsoft 365 Group mailboxes & sites** and **OneDrive accounts** locations: Items that are stored in arbitration mailboxes aren't supported for simulation.
- Because simulation results are based on items available in the specified locations at the time the simulation job runs, remember to take the following considerations into account when you turn on the policy:
  - Items that are no longer within the specified location won't be labeled.
  - Items that no longer match the policy criteria won't be labeled.

On the **Label policies** page, the **Status** column displays **In simulation** for auto-labeling policies that are running in simulation, or configured for simulation and complete.

Simulation typically completes within one or two days, depending on the amount of data to analyze. The completed simulation triggers an email notification that's sent to the user configured to receive activity alerts.

To view the simulation results, select the policy from the **Label policies** page, and from the flyout pane, select **View simulation**. You can then view any samples, review the number of matching items and the locations, edit the policy, turn on the policy, or restart the simulation.

Before you begin

To make sure you have permissions to create and edit retention labels and their policies, see the permissions information for records management or data lifecycle management, depending on the solution you're using.

To run the policy in simulation mode:

- Auditing for Microsoft 365 must be turned on. If you need to turn on auditing or you're not sure whether auditing is already on, see Turn audit log search on or off.
- To view the list of samples, you must have the **Data Classification List Viewer** role, and to view file or email contents in the source view, you must have the **Data Classification Content Viewer** role.

Make sure you have created the retention labels you want to apply to items.

> **Note**
>
> You can extend your auto-labeling of retention labels to images. For more information, see Learn about optical character recognition in Microsoft Purview.

How to create an auto-apply retention label policy

Decide before you create your retention label policy whether to make it **adaptive** or **static**. For more information, see Adaptive or static policy scopes for retention. If you decide to use an adaptive policy, you must create one or more adaptive scopes before you create your retention label policy, and then select them during the create retention label policy process. For instructions, see Configuration information for adaptive scopes.

When you create an auto-apply policy, you select a retention label to automatically apply to content, based on the conditions that you specify:

1. Sign in to the Microsoft Purview portal and then:

   If you're using records management: **Solutions** > **Records Management** > **Policies** > **Label policies** > **Auto-apply a label**

   If you're using data lifecycle management: **Solutions** > **Data Lifecycle Management** > **Policies** > **Label policies** > **Auto-apply a label**

2. Enter a name and description for this auto-labeling policy, and then select **Next**.

3. For **Choose the type of content you want to apply this label to**, select one of the available conditions. For more information about the choices, see the Configuring conditions for auto-apply retention labels section on this page.

4. For the **Assign admin units** page, keep the default of **Full directory**. Currently, admin units aren't supported for this policy.

5. For the **Choose the type of retention policy to create** page, select **Adaptive** or **Static**, depending on the choice you made from the Before you begin instructions. If you haven't already created adaptive scopes, you can select **Adaptive** but because there won't be any adaptive scopes to select, you won't be able to finish the policy configuration with this option.

6. Depending on your selected scope:

   - If you chose **Adaptive**: On the **Choose adaptive policy scopes and locations** page, select **Add scopes** and select one or more adaptive scopes that have been created. Then, select one or more locations. The locations that you can select depend on the scope types added. For example, if you only added a scope type of **User**, you'll be able to select **Exchange mailboxes** but not **SharePoint sites**.

   - If you chose **Static**: On the **Choose locations** page, toggle on or off any of the locations. For each location, you can leave it at the default to apply the policy to the entire location, or specify includes and excludes

For information about the location choices, see Locations.

7. Follow the prompts to select a retention label, whether to run the policy in simulation mode or turn it on (if applicable for your chosen condition), and then review and submit your configuration choices.

To edit an existing retention label policy (the policy type is **Auto-apply**), select it, and then select the **Edit** option to start the **Edit retention policy** configuration.

After content is labeled by using an auto-apply label policy, the applied label can't be automatically removed or changed by changing the content or the policy, or by a new auto-apply label policy. For more information, see Only one retention label at a time.

> **Note**
>
> An auto-apply retention label policy will never replace an existing retention label that's applied to content. If you want to relabel content by using the conditions you configure, you'll need to manually remove the current retention label from existing content.

Configuring conditions for auto-apply retention labels

You can apply retention labels to content automatically when that content contains:

- Specific types of sensitive information
- Specific keywords or searchable properties that match a query you create
- A match for trainable classifiers

Or, you can automatically apply retention labels to newly shared cloud attachments.

When you configure retention labels to auto-apply based on sensitive information, keywords or searchable properties, or trainable classifiers, use the following table to identify when retention labels can be automatically applied.

SharePoint items that are in draft or that have never been published aren't supported for this scenario.

Exchange:

Expand table

Condition	Items in transit (sent or received)	Existing items (data at rest)
Sensitive info types	Yes	No
Specific keywords or searchable properties	Yes	Yes
Trainable classifiers	Yes	Yes (last six months only)

SharePoint and OneDrive:

Expand table

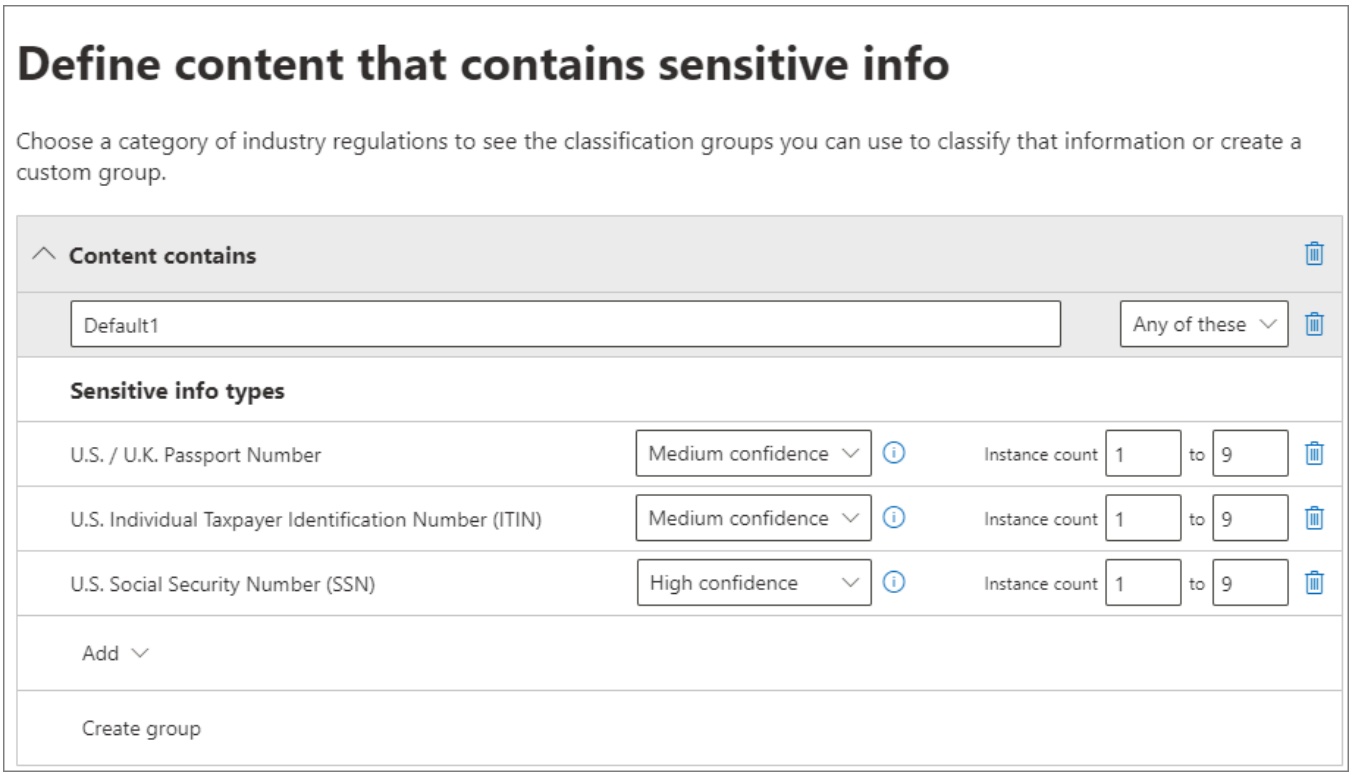
Condition	New or modified items	Existing items
Sensitive info types	Yes	Yes *
Specific keywords or searchable properties	Yes	Yes
Trainable classifiers	Yes	Yes (last six months only)

Footnote:

* Applies only to content that's already classified, which you can determine by using content explorer.

Auto-apply labels to content with specific types of sensitive information

When you create auto-apply retention label policies for sensitive information, you see the same list of policy templates as when you create a Microsoft Purview Data Loss Prevention (DLP) policy. Each template is preconfigured to look for specific types of sensitive information. In the following example, the sensitive info types are from the **Privacy** category, and **U.S Personally Identifiable Information (PII) Data** template:



To learn more about the sensitive information types, see Learn about sensitive information types. Currently, exact data match based sensitive information types and document fingerprinting aren't supported for this scenario.

After you select a policy template, you can add or remove any types of sensitive information, and you can change the confidence level and instance count. In the previous example screenshot, these options have been changed so that a retention label will be auto-applied only when:

- The type of sensitive information that's detected has a match accuracy (or confidence level) of at least **Medium confidence** for two of the sensitive info types, and **High confidence** for one. Many sensitive information types are defined with multiple patterns, where a pattern with a higher match accuracy requires more evidence to be found (such as keywords, dates, or addresses), while a pattern with a lower match accuracy requires less evidence. The lower the confidence level, the easier it is for content to match the condition but with the potential for more false positives.
- The content contains between 1 and 9 instances of any of these three sensitive info types. The default for the **to** value is **Any**.

For more information about these options, see the following guidance from the DLP documentation Test your Data Loss Prevention policies.

> ⓘ **Important**

Sensitive information types have two different ways of defining the max unique instance count parameters. To learn more, see Instance count supported values for SIT.

To consider when using sensitive information types to auto-apply retention labels:

- If you use custom sensitive information types, these can't auto-label existing items in SharePoint and OneDrive.

Auto-apply labels to content with keywords or searchable properties

You can auto-apply labels to content by using a query that contains specific words, phrases, or values of searchable properties. You can refine your query by using search operators such as AND, OR, and NOT.



For more information about the query syntax that uses Keyword Query Language (KeyQL), see Keyword Query Language (KeyQL) syntax reference.

Query-based auto-apply policies use the same search index as eDiscovery content search to identify content. For more information about the searchable properties that you can use, see Keyword queries and search conditions for Content Search.

Some things to consider when using keywords or searchable properties to auto-apply retention labels:

- For SharePoint, crawled properties and custom properties aren't supported for these KeyQL queries and you must use only predefined managed properties for documents. However, you can use mappings at the tenant level with the predefined managed

properties that are enabled as refiners by default (RefinableDate00-19, RefinableString00-99, RefinableInt00-49, RefinableDecimals00-09, and RefinableDouble00-09). For more information, see Overview of crawled and managed properties in SharePoint Server, and for instructions, see Create a new managed property.

- If you map a custom property to one of the refiner properties, wait 24 hours before you use it in your KeyQL query for a retention label.
- Although SharePoint managed properties can be renamed by using aliases, don't use these for KeyQL queries in your labels. Always specify the actual name of the managed property, for example, "RefinableString01".
- To search for values that contain spaces or special characters, use double quotation marks (`" "`) to contain the phrase; for example, `subject:"Financial Statements"`.
- Suffix wildcard searches (such as `*cat`) or substring wildcard searches (such as `*cat*`) aren't supported. However, prefix wildcard searches (such as `cat*`) are supported.
- Use the *DocumentLink* property instead of *Path* to match an item based on its URL. To include or exclude content inside a specific document library, use the *ParentLink* property.
- Be aware that partially indexed items can be responsible for not labeling items that you're expecting, or labeling items that you're expecting to be excluded from labeling when you use the NOT operator. For more information, see Partially indexed items in Content Search.
- We recommend that you don't use spaces between words in RefinableStrings values on documents. RefinableString isn't a word-break property.

Examples queries:

Expand table

Workload	Example	Description
Exchange	`subject:"Financial Statements"`	Finds emails with the exact phrase "Financial Statements" in the subject line, ignoring case.
Exchange	`recipients:garthf@contoso.com`	Finds emails sent to `garthf@contoso.com`, whether directly or copied.

Workload	Example	Description
SharePoint	`contenttype:document`	Finds all items classified as documents in SharePoint.
SharePoint	`site:https://contoso.sharepoint.com/sites/teams/procurement AND contenttype:document`	Finds all documents in the Procurement team site on SharePoint.
SharePoint	`ParentLink:SiteAssets`	Finds items stored in the Site Assets library of SharePoint sites.
SharePoint	`ParentLink:"https://contoso.sharepoint.com/sites/Legal/Shared Documents"`	Finds only items stored in the Shared Documents library of the Legal site.
SharePoint	`NOT ParentLink:"https://contoso.sharepoint.com/sites/Legal/Shared Documents"`	Finds all content except items stored in that document library.
OneDrive	`ParentLink:"https://contoso-my.sharepoint.com/personal/john_doe_contoso_com/Documents"`	Includes only content stored in John Doe's OneDrive Documents library.
OneDrive	`NOT ParentLink:"https://contoso-my.sharepoint.com/personal/john_doe_contoso_com/Documents"`	Excludes all content from that specific OneDrive library.
Exchange or SharePoint	`"customer information" OR "private"`	Finds content that includes either "customer information" or "private", in emails (Exchange) or documents (SharePoint).

More complex examples:

The following query for SharePoint identifies Word documents or Excel spreadsheets when those files contain the keywords **password**, **passwords**, or **pw**:

KeyQL

```
(password OR passwords OR pw) AND (filetype:doc* OR filetype:xls*)
```

The following query for Exchange identifies any Word document or PDF that contains the word **nda** or the phrase **non disclosure agreement** when those documents are attached to an email:

KeyQL

```
(nda OR "non disclosure agreement") AND (attachmentnames:.doc* OR
attachmentnames:.pdf)
```

The following query for SharePoint identifies documents that contain a credit card number:

KeyQL

```
sensitivetype:"credit card number"
```

The following query contains some typical keywords to help identify documents or emails that contain legal content:

KeyQL

```
ACP OR (Attorney Client Privilege*) OR (AC Privilege)
```

The following query contains typical keywords to help identify documents or emails for human resources:

KeyQL

```
(resume AND staff AND employee AND salary AND recruitment AND candidate)
```

Note that this final example uses the best practice of always including operators between keywords. A space between keywords (or two `property:value` expressions) is the same as using AND. By always adding operators, it's easier to see that this example query identifies only content that contains all these keywords, instead of content that contains any of the keywords. If your intention is to identify content that contains any of the keywords, specify OR instead of AND. As this example shows, when you always specify the operators, it's easier to correctly interpret the query.

Microsoft Teams meeting recordings and transcripts

To identify Microsoft meeting recordings and transcripts that are stored in users' OneDrive accounts or in SharePoint, specify the following for the **Keyword query editor**:

KeyQL

```
ProgID:Media AND ProgID:Meeting
```

Most of the time, meeting recordings and transcripts are saved to OneDrive. But for channel meetings, they're saved in SharePoint. For more information, see Use OneDrive and SharePoint to store meeting recordings.

The query will identify:

- A Teams meeting transcript, when there isn't an accompanying recording
- A Teams meeting recording and accompanying transcript

Because video transcripts are stored with the video in OneDrive or SharePoint, they can't be targeted for retention separately when there's a Teams meeting recordings video file.

Identify files and emails that have a sensitivity label

To identify files in SharePoint or OneDrive and Exchange emails that have a specific sensitivity label applied, specify the following for the **Keyword query editor**:

KeyQL

```
InformationProtectionLabelId:<GUID>
```

To find the GUID, use the Get-Label cmdlet from Security & Compliance PowerShell:

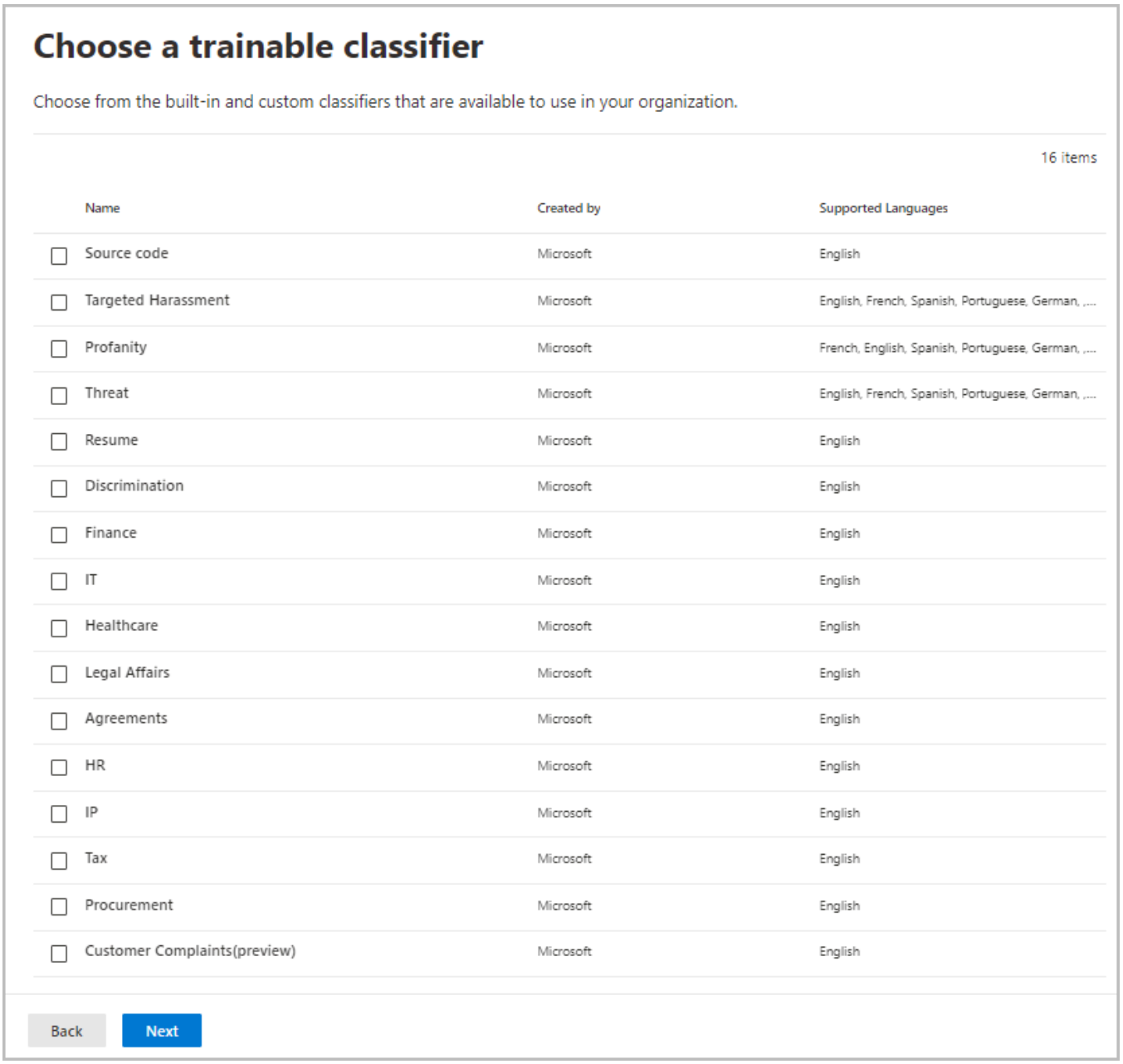
PowerShell

```
Get-Label | Format-Table -Property DisplayName, Name, Guid
```

Auto-apply labels to content by using trainable classifiers

> **Important**
>
> Currently, trainable classifiers for auto-labeling can't be used with adaptive scopes. Use a static scope instead.

When you choose the option for a trainable classifier, you can select one or more of the pre-trained or custom trainable classifiers:



The available pre-trained classifiers are often updated, so there might be more entries to select than the ones displayed in this screenshot.

For more information about trainable classifiers, see Learn about trainable classifiers.

To automatically apply a label by using this option, SharePoint sites, as well as mailboxes, must have at least 10 MB of data.

To consider when using trainable classifiers to auto-apply retention labels:

- You can't auto-label SharePoint and OneDrive items that are older than six months.

Auto-apply labels to cloud attachments

> **Note**

Support for cloud attachments that are shared in Viva Engage is in preview.

You might need to use this option if you're required to capture and retain all copies of files in your tenant that are sent over communications by users, or files that are referenced in Microsoft 365 Copilot or Microsoft 365 Copilot Chat. You use this option in conjunction with retention policies for the communication services themselves; Exchange, Teams, Viva Engage, and Copilot.

> **Important**
>
> When you select a label to use for auto-applying retention labels for cloud attachments, ensure that the label retention setting **Start the retention period based on** is **When items were labeled**.

Cloud attachments, sometimes also known as modern attachments, are a sharing mechanism that uses embedded links to files that are stored in the cloud. They support centralized storage for shared content with collaborative benefits, such as version control. Cloud attachments are not attached copies of a file or a URL text link to a file. However, support for URL text links is also supported. You might find it helpful to refer to the visual checklists for supported cloud attachments in Outlook, Teams, and Viva Engage.

When you choose the option to apply a retention label to cloud attachments, for compliance purposes, a copy of that file is created at the time of sharing. Your selected retention label is then applied to the copy that can then be identified using eDiscovery. Users aren't aware of the copy that is stored in the Preservation Hold library. The retention label isn't applied to the message itself, or to the original file.

If the file is modified and shared again, a new copy of the file as a new version is saved in the Preservation Hold library. For more information, including why you should use the **When items were labeled** label setting, see How retention works with cloud attachments.

The cloud attachments supported for this option are files such as documents, videos, and images that are stored in SharePoint and OneDrive. For Teams, cloud attachments shared in chat messages, and standard and private channels are supported. For Viva Engage, cloud attachments shared with users in storylines, community posts, and Inbox messages are supported. For Copilot, files referenced to users in the chat conversation and referenced by users are supported.

Cloud attachments shared over meeting invites and apps other than Teams, Outlook, Viva Engage, and Copilot aren't supported. The cloud attachments must be shared by users or Copilot; cloud attachments sent via bots aren't supported.

Although not required for this option, we recommend that you ensure versioning is enabled for your SharePoint sites and OneDrive accounts so that the version shared can be accurately captured. If versioning isn't enabled, the last available version is retained. Documents in draft or that have never been published aren't supported.

When you select a label to use for auto-applying retention labels for cloud attachments, make sure the label retention setting **Start the retention period based on** is **When items were labeled**.

When you configure the locations for this option, you can select:

- **SharePoint classic and communication sites** for shared files stored in SharePoint communication sites, team sites that aren't connected by Microsoft 365 groups, and classic sites.
- **Microsoft 365 Groups** for shared files that are stored in team sites connected by Microsoft 365 groups.
- **OneDrive accounts** for shared files stored in users' OneDrive.

You'll need to create separate retention policies if you want to retain or delete the original files, email messages, or messages from Teams and Copilot, and Viva Engage.

> ⓘ **Note**
>
> If you want retained cloud attachments to expire at the same time as the messages that contained them, configure the retention label to have the same retain and then delete actions and timings as your retention policies for Exchange, Teams and Copilot, and Viva Engage.

To consider when auto-applying retention labels to cloud attachments:

- If cloud attachments and links in a Teams or Viva Engage message are changed after the message is sent by editing the message, those changed cloud attachments and links aren't supported for retention.

- When a user is added to a Teams conversation and given access to the full history of the conversation, that history can include cloud attachments and URL text links. If these attachments were shared within 48 hours of the user added to the conversation, current copies of the attachments are auto-labeled for retention. Attachments shared before this time period aren't supported for newly added users.

- Attachments and links shared outside Teams, Outlook, Viva Engage, and Copilot aren't supported, and the attachments and links must be content stored in SharePoint or OneDrive.

- Cloud attachments and links in encrypted emails or encrypted messages aren't supported.
- Sharing an existing Viva Engage message with an attachment isn't supported.
- Specific to shared documents from URL text links:
  - Supported in the message body but not in the email subject or Teams channel subject, announcement, or subheadings.
  - Not supported for previous responses in the same thread, only the current message
  - Total limit of 25 attachments in a single message, where this maximum can be any combination of cloud attachments and shared documents from URL text links
  - Not supported beyond 5,000 characters in the initial email body or in Teams, Copilot, and Viva Engage messages
- The following items aren't supported as attachments that can be retained:
  - SharePoint sites, pages, lists, forms, folders, document sets, and OneNote pages.
  - Files shared by users who don't have access to those files at the time of sharing.
  - Files that are deleted or moved before the cloud attachment is sent. For example, a user copies and pastes a previously shared attachment from another message, without first confirming that the file is still available. Or, somebody forwards an old message when the file is now deleted.
  - Files that are shared by guests or users outside your organization.
  - Files in draft emails and messages that aren't sent.
  - Empty files.

How long it takes for retention labels to take effect

When you auto-apply retention labels based on sensitive information, keywords or searchable properties, or trainable classifiers, it can take up to seven days for the retention labels to be applied:



If the expected labels don't appear after seven days, check the **Status** of the auto-apply policy by selecting it from the **Label policies** page in the Microsoft Purview portal. If you see the status of **Off (Error)** and in the details for the locations see a message that it's taking longer than expected to deploy the policy (for SharePoint) or to try redeploying the policy (for OneDrive), try running the Set-RetentionCompliancePolicy PowerShell command to retry the policy distribution:

1. Connect to Security & Compliance PowerShell.
2. Run the following command:

```
PowerShell
Set-RetentionCompliancePolicy -Identity <policy name> -RetryDistribution
```

Updating retention labels and their policies

For auto-apply retention label policies that are configured for sensitive information, keywords or searchable properties, or a match for trainable classifiers: When a retention label from the policy is already applied to content, a change in configuration to the selected label and policy will be automatically applied to this content in addition to content that's newly identified.

For auto-apply retention label policies that are configured for cloud attachments: Because this policy applies to newly shared files rather than existing files, a change in configuration to the selected label and policy will be automatically applied to newly shared content only.

Some settings can't be changed after the label or policy is created and saved, which include:

- Names for retention labels and their policies, the scope type (adaptive or static), and the retention settings except the retention period.
- The option to mark items as a record.

Deleting retention labels

To delete a retention label, all three conditions must apply:

- The label isn't included in any retention label policy
- The label isn't configured for event-based retention
- The label isn't configured to mark items as regulatory records

When all these conditions are met:

- You can always delete a retention label that doesn't mark items as records (sometimes referred to as a "standard retention label"). The deletion succeeds even if the label is applied to items, and the retention label is then removed from these items.
- You can delete a retention label that marks items as records only if the label isn't applied to items. If the label has been applied to items, the deletion fails and you see a link to content explorer to identify the labeled items. It can take up to two days for content explorer to show the items that are labeled. In this scenario, the retention label might be deleted without showing you the link to content explorer.

Locking the policy to prevent changes

If you need to ensure that no one can turn off the policy, delete the policy, or make it less restrictive, see Use Preservation Lock to restrict changes to retention policies and retention label policies.

Troubleshooting retention label policies

If your auto-apply retention label policies aren't working as expected or you see errors related to these policies, use the following troubleshooting resources:

- Identify errors in Microsoft 365 retention and retention label policies
- Resolve errors in Microsoft 365 retention and retention label policies

Additionally, if you're also using the older feature, MRM retention tags and retention policies, see Auto-apply retention label doesn't apply to items in a mailbox.

Next steps

To help you track the labels applied from your auto-labeling policies:

- Monitoring retention labels
- Using Content Search to find all content with a specific retention label
- Auditing retention actions

See Use retention labels to manage the lifecycle of documents stored in SharePoint for an example scenario that uses an auto-apply retention label policy with managed properties in SharePoint, and event-based retention to start the retention period.

Last updated on 02/26/2026

Use Preservation Lock to restrict changes to retention policies and retention label policies

09/26/2025

> *Microsoft 365 licensing guidance for security & compliance*.

> **ⓘ Important**
>
> Currently, adaptive policy scopes don't support Preservation Lock.

Preservation Lock locks a retention policy or retention label policy so that no one—including a global admin—can turn off the policy, delete the policy, or make it less restrictive. This configuration might be needed for regulatory requirements and can help safeguard against rogue administrators.

When a retention policy is locked:

- No one can disable the policy or delete it
- Locations can be added but not removed
- You can extend the retention period but not decrease it

When a retention label policy is locked:

- No one can disable the policy or delete it
- Locations can be added but not removed
- Labels can be added but not removed

In summary, a locked policy can be increased or extended, but it can't be reduced or turned off.

> **ⓘ Important**
>
> Before you lock a retention policy or retention label policy, it's critical that you understand the impact and confirm whether it's required for your organization. For example, it might be needed to meet regulatory requirements. Administrators won't be able to disable or delete these policies after the preservation lock is applied.

Configure Preservation Lock after you've created a retention policy, or a retention label policy that you publish and contains only labels that mark items as regulatory records.

How to lock a retention policy or retention label policy

You must use PowerShell if you need to use Preservation Lock. Because administrators can't disable or delete a policy for retention after this lock is applied, enabling this feature is not available in the UI to safeguard against accidental configuration.

All retention policies with any configuration support Preservation Lock. To apply Preservation Lock on a retention label policy, it must contain only labels that mark items as regulatory records.

1. Connect to Security & Compliance PowerShell.

2. Find the name of the policy that you want to lock by running Get-RetentionCompliancePolicy. For example:

```
PS C:\WINDOWS\system32> Get-RetentionCompliancePolicy
Name                 Workload                                                Enabled Mode
----                 --------                                                ------- ----
PII Retention Policy Exchange, SharePoint, OneDriveForBusiness, Skype, ModernGroup True    Enforce
Employee Records     Exchange, SharePoint, OneDriveForBusiness, Skype, ModernGroup True    Enforce
Personal Financial PII Exchange, SharePoint, OneDriveForBusiness, Skype, ModernGroup True    Enforce
```

3. To place a Preservation Lock on your policy, run the Set-RetentionCompliancePolicy cmdlet with the name of the policy, and the *RestrictiveRetention* parameter set to true:

```
PowerShell

Set-RetentionCompliancePolicy -Identity "<Name of Policy>" -
RestrictiveRetention $true
```

   For example:

```
PS C:\WINDOWS\system32> Set-RetentionCompliancePolicy -Identity "Employee Records" -RestrictiveRetention $true
```

   When prompted, read and acknowledge the restrictions that come with this configuration by entering **Y**:

```
Confirm
You chose to make this a restrictive preservation policy. After you turn this policy on, you'll only be able to
add new locations to search or extend the time frame to preserve the content. You won't be able to turn off the
policy, delete it, or make any other changes.
[Y] Yes  [A] Yes to All  [N] No  [L] No to All  [?] Help (default is "Y"):
```

A Preservation Lock is now placed on the policy. To confirm, run `Get-RetentionCompliancePolicy` again, but specify the policy name and display the policy parameters:

PowerShell

```
Get-RetentionCompliancePolicy -Identity "<Name of Policy>" |Fl
```

You should see **RestrictiveRetention** is set to **True**. For example:

```
PS C:\WINDOWS\system32> Get-RetentionCompliancePolicy -Identity "Employee Records" |Fl


RunspaceId                    : d2b39144-2e5d-4dea-ab59-2786944f9ff3
Type                          : Hold
TeamsPolicy                   : False
SharePointLocation            : {}
SharePointLocationException   : {}
RetentionRuleTypes            : {}
ExchangeLocation              : {}
ExchangeLocationException     : {}
PublicFolderLocation          : {}
SkypeLocation                 : {}
SkypeLocationException        : {}
ModernGroupLocation           : {}
ModernGroupLocationException  : {}
OneDriveLocation              : {}
OneDriveLocationException     : {}
TeamsChatLocation             : {}
TeamsChatLocationException    : {}
TeamsChannelLocation          : {}
TeamsChannelLocationException : {}
DynamicScopeLocation          : {}
RestrictiveRetention          : True
Workload                      : Exchange, SharePoint, OneDriveForBusiness, Skype, ModernGroup
Priority                      : 1
ObjectVersion                 : 533c3975-eae4-4520-aa83-08d83ee404f2
```

See also

Resources to help you meet regulatory requirements for data lifecycle management and records management

Create and publish retention labels by using PowerShell

09/26/2025

> *Microsoft 365 licensing guidance for security & compliance*.

After you've decided to use retention labels to help you keep or delete documents and emails in Microsoft 365, you might have realized that you have many and possibly hundreds of retention labels to create and publish. The recommended method to create retention labels at scale is by using file plan from the Microsoft Purview portal. However, you can also use PowerShell.

Use the information, template files and examples, and script in this article to help you bulk-create retention labels and publish them in retention label policies. Then, the retention labels can be applied by administrators and users.

The supplied instructions don't support retention labels that are auto-applied.

Overview:

1. In Excel, create a list of your retention labels and a list of their retention label policies.
2. Use PowerShell to create the retention labels and retention label policies in those lists.

Disclaimer

The sample scripts provided in this article aren't supported under any Microsoft standard support program or service. The sample scripts are provided AS IS without warranty of any kind. Microsoft further disclaims all implied warranties including, without limitation, any implied warranties of merchantability or of fitness for a particular purpose. The entire risk arising out of the use or performance of the sample scripts and documentation remains with you. In no event shall Microsoft, its authors, or anyone else involved in the creation, production, or delivery of the scripts be liable for any damages whatsoever (including, without limitation, damages for loss of business profits, business interruption, loss of business information, or other pecuniary loss) arising out of the use of or inability to use the sample scripts or documentation, even if Microsoft has been advised of the possibility of such damages.

Step 1: Create a .csv file for the retention labels

1. Copy the following sample .csv file for a template and example entries for four different retention labels, and paste them into Excel.
2. Convert the text to columns: **Data** tab > **Text to Columns** > **Delimited** > **Comma** > **General**
3. Replace the examples with entries for your own retention labels and settings. For more information about the parameter values, see New-ComplianceTag.
4. Save the worksheet as a .csv file in a location that's easy to find for a later step. For example: C:>Scripts\Labels.csv

Notes:

- If the .csv file contains a retention label with the same name as one that already exists, the script skips creating that retention label. No duplicate retention labels are created.
- Don't change or rename the column headers from the sample .csv file provided, or the script will fail.

Sample .csv file for retention labels

text

```
Name (Required),Comment (Optional),IsRecordLabel (Required),RetentionAction
(Optional),RetentionDuration (Optional),RetentionType (Optional),ReviewerEmail
(Optional)
LabelName_t_1,Record - keep and delete - 2
years,$true,KeepAndDelete,730,CreationAgeInDays,
LabelName_t_2,Keep and delete tag - 7
years,$false,KeepAndDelete,2555,ModificationAgeInDays,
LabelName_t_3,5 year delete,$false,Delete,1825,TaggedAgeInDays,
LabelName_t_4,Record label tag - financial,$true,Keep,730,CreationAgeInDays,
```

Step 2: Create a .csv file for the retention label policies

1. Copy the following sample .csv file for a template and example entries for three different retention label policies, and paste them into Excel.
2. Convert the text to columns: **Data** tab > **Text to Columns** > **Delimited** > **Comma** > **General**

3. Replace the examples with entries for your own retention label policies and their settings. For more information about the parameter values for this cmdlet, see New-RetentionCompliancePolicy.

4. Save the worksheet as a .csv file in a location that's easy to find for a later step. For example: `<path>Policies.csv`

Notes:

- If the .csv file contains a retention label policy with the same name as one that already exists, the script skips creating that retention label policy. No duplicate retention label policies are created.
- Don't change or rename the column headers from the sample .csv file provided, or the script will fail.

Sample .csv file for retention policies

text

```
Policy Name (Required),PublishComplianceTag (Required),Comment (Optional),Enabled
(Required),ExchangeLocation (Optional),ExchangeLocationException
(Optional),ModernGroupLocation (Optional),ModernGroupLocationException
(Optional),OneDriveLocation (Optional),OneDriveLocationException
(Optional),PublicFolderLocation (Optional),SharePointLocation
(Optional),SharePointLocationException (Optional),SkypeLocation
(Optional),SkypeLocationException (Optional)
Publishing Policy Red1,"LabelName_t_1, LabelName_t_2, LabelName_t_3,
LabelName_t_4",N/A,$true,All,,All,,All,,,All,,,
Publishing Policy Orange1,"LabelName_t_1, LabelName_t_2",N/A,$true,All,,,,,,,,,
Publishing Policy Yellow1,"LabelName_t_3, LabelName_t_4",N/A,$false,All,,,,,,,,,
```

Step 3: Create the PowerShell script

1. Copy and paste the following PowerShell script into Notepad.

2. Save the file by using a file name extension of **.ps1** in a location that's easy to find. For example: `<path>CreateRetentionSchedule.ps1`

Notes:

- The script prompts you to provide the two source files that you created in the previous two steps:
  - If you don't specify the source file to create the retention labels, the script moves on to create the retention label policies.

- If you don't specify the source file to create the retention label policies, the script creates the retention labels only.

- The script generates a log file that records each action it took and whether the action succeeded or failed. See the final step for instructions how to locate this log file.

PowerShell script

Powershell

```
<#
. Steps: Import and publish retention labels
    - Load retention labels csv file
    - Validate csv file input
    - Create retention labels
    - Create retention policies
    - Publish retention labels for the policies
    - Generate the log for retention labels and policies creation
    - Generate the csv result for the labels and policies created
. Syntax
    .\Publish-ComplianceTag.ps1 [-LabelListCSV <string>] [-PolicyListCSV <string>]
. Detailed Description
    1) [-LabelListCSV <string>]
    -LabelListCSV ".\SampleInputFile_LabelList.csv"
    Load compliance tag for creation.
    2) [-PolicyListCSV <string>]
    -PolicyListCSV ".\SampleInputFile_PolicyList.csv"
    Load compliance tag for creation.
#>
param (
    [Parameter(Mandatory = $true)]
    [string]$LabelListCSV = "",
    [Parameter(Mandatory = $true)]
    [string]$PolicyListCSV = "",
    [Switch]$ResultCSV
)
# -------------------
# File operation
# -------------------
Function FileExist
{
    Param(
        # File path needed to check
        [Parameter(Mandatory = $true)]
        [String]$FilePath,
        [Switch]$Warning
    )
    $inputFileExist = Test-Path $FilePath
    if (!$inputFileExist)
    {
        if ($Warning -eq $false)
        {
```

```
            WriteToLog -Type "Failed" -Message "[File: $FilePath] The file doesn't
exist"
            throw
        }
        else
        {
            WriteToLog -Type "Warning" -Message "[File: $FilePath] The file
doesn't exist"
        }
    }
    else
    {
        WriteToLog -Type "Succeed" -Message "[File: $FilePath] The file is found"
    }
}
# -------------------
# Log operation
# -------------------
Function WriteToLog
{
    Param(
        # Message want to write to log file
        [Parameter(Mandatory = $true)]
        [String]$Message,
        # "Succeed" or "Faild"
        [String]$Type = "Message"
    )
    $date = Get-Date -Format 'HH:mm:ss'
    $logInfo = $date + " - [$Type] " + $Message
    $logInfo | Out-File -FilePath $logfilePath -Append
    if ($Type -eq "Succeed") { Write-Host $logInfo -ForegroundColor Green }
    elseif ($Type -eq "Failed") { Write-Host $logInfo -ForegroundColor Red }
    elseif ($Type -eq "Warning") { Write-Host $logInfo -ForegroundColor Yellow }
    elseif ($Type -eq "Start") { Write-Host $logInfo -ForegroundColor Cyan }
    else { Write-Verbose $logInfo }
}
Function Create-Log
{
    Param(
        # Log folder Root
        [Parameter(Mandatory = $true)]
        [String]$LogFolderRoot,
        # The function Log file for
        [Parameter(Mandatory = $true)]
        [String]$LogFunction
    )
    $logFolderPath = "$LogFolderRoot\logfiles"
    $folderExist = Test-Path "$logFolderPath"
    if (!$folderExist)
    {
        $folder = New-Item "$logFolderPath" -type directory
    }
    $date = Get-Date -Format 'MMddyyyy_HHmmss'
    $logfilePath = "$logFolderPath\Log_{0}_{1}.txt" -f $LogFunction, $date
    Write-Verbose "Log file is written to: $logfilePath"
```

```
    $logfile = New-Item $logfilePath  -type file
    return $logfilePath
}
Function Create-ResultCSV
{
    Param(
        # Result folder Root
        [Parameter(Mandatory = $true)]
        [String]$ResultFolderRoot,
        # The function Result file for
        [Parameter(Mandatory = $true)]
        [String]$ResultFunction
    )
    $retFolderPath = "$ResultFolderRoot\logfiles"
    $folderExist = Test-Path "$retFolderPath"
    if (!$folderExist)
    {
        $folder = New-Item "$retFolderPath" -type directory
    }
    $date = Get-Date -Format 'MMddyyyy_HHmmss'
    $retfilePath = "$retFolderPath\Result_{0}_{1}.csv" -f $ResultFunction, $date
    Write-Verbose "Result file is written to: $retfilePath"
    $retfile = New-Item $retfilePath  -type file
    return $retfilePath
}
# -------------------
# Prepare Log File
# -------------------
$scriptPath = '.\'
$logfilePath = Create-Log -LogFolderRoot $scriptPath -LogFunction
"Publish_Compliance_Tag"
if ($ResultCSV)
{
    $tagRetFile = Create-ResultCSV -ResultFolderRoot $scriptPath -ResultFunction
"Tag_Creation"
    $tagPubRetFile = Create-ResultCSV -ResultFolderRoot $scriptPath -
ResultFunction "Tag_Publish"
}
# -------------------
# Invoke Powershell cmdlet
# -------------------
Function InvokePowerShellCmdlet
{
    Param(
        [Parameter(Mandatory = $true)]
        [String]$CmdLet
    )
    try
    {
        WriteToLog -Type "Start" -Message "Execute Cmdlet : '$CmdLet'"
        return Invoke-Expression $CmdLet -ErrorAction SilentlyContinue
    }
    catch
    {
        WriteToLog -Type "Failed" "Failed to execute cmdlet!"
```

```
        WriteToLog -Type "Failed" $error[0]
        return $null
    }
}
# -------------------
# Create Compliance Tag
# -------------------
Function CreateComplianceTag
{
    Param(
        # File path needed to check
        [Parameter(Mandatory = $true)]
        [String]$FilePath
    )

    WriteToLog -Type "Start" "Start to create Compliance Tag"
    FileExist $FilePath

    # TODO Validate CSV file for the Header
    try
    {
        # Import csv
        $labels = Import-Csv $FilePath
        # Retrieve existing compliance tags
        $tags = InvokePowerShellCmdlet "Get-ComplianceTag"
        foreach($lab in $labels)
        {
            # Cmdlet parameters
            $para = [String]::Empty;
            $name = [String]::Empty;
            $cmdlet = 'New-ComplianceTag'
            if ([String]::IsNullOrEmpty($lab.'Name (Required)'))
            {
                WriteToLog -Type "Failed" -Message "Could not acquire table for
writing."
                throw;
            }
            else
            {
                $name = $lab.'Name (Required)'
                $cmdlet += " -Name '" + $name + "'"
            }
            if (![String]::IsNullOrEmpty($lab.'Comment (Optional)'))
            {
                $para = $lab.'Comment (Optional)'
                $cmdlet += " -Comment '" + $para + "'"
            }
            if (![String]::IsNullOrEmpty($lab.'IsRecordLabel (Required)'))
            {
                $para = $lab.'IsRecordLabel (Required)'
                $cmdlet += " -IsRecordLabel " + $para
            }
            if (![String]::IsNullOrEmpty($lab.'RetentionAction (Optional)'))
            {
                $para = $lab.'RetentionAction (Optional)'
```

```
                $cmdlet += " -RetentionAction " + $para
            }
            if (![String]::IsNullOrEmpty($lab.'RetentionDuration (Optional)'))
            {
                $para = $lab.'RetentionDuration (Optional)'
                $cmdlet += " -RetentionDuration " + $para
            }
            if (![String]::IsNullOrEmpty($lab.'RetentionType (Optional)'))
            {
                $para = $lab.'RetentionType (Optional)'
                $cmdlet += " -RetentionType " + $para
            }
            if (![String]::IsNullOrEmpty($lab.'ReviewerEmail (Optional)'))
            {
                $emails = $lab.'ReviewerEmail (Optional)'.Split(",") | ForEach-
Object { $_.Trim() }
                if (($emails -ne $null) -and ($emails.Count -ne 0))
                {
                    $eml = '@('
                    foreach($email in $emails)
                    {
                        $eml += "'{0}'," -f $email
                    }
                    $eml = $eml.Substring(0, $eml.Length - 1) + ')'

                    $cmdlet += " -ReviewerEmail " + $eml
                }
            }
            # If the tag already exists, skip for creation
            if (($tags -eq $null) -or ($tags | ? { $_.Name.ToLower() -eq
$name.ToLower() }) -eq $null)
            {
                # Create compliance tag
                $msg = "Execute Cmdlet : {0}" -f $cmdlet

                $ret = InvokePowerShellCmdlet $cmdlet

                if ($ret -eq $null)
                {
                    WriteToLog -Type "Failed" $error[0]
                    break;
                }
            }
            else
            {
                WriteToLog -Type "Warning" -Message "The tag '$name' already
exists! Skip for creation!"
            }
        }
    }
    catch
    {
        WriteToLog -Type "Failed" "Error in input"
    }
}
```

```
# -------------------
# Create Retention Compliance Policy
# -------------------
Function CreateRetentionCompliancePolicy
{
    Param(
        # File path needed to check
        [Parameter(Mandatory = $true)]
        [String]$FilePath
    )

    WriteToLog -Type "Start" "Start to Create Retention Policy"
    FileExist $FilePath
    try
    {
        # Import csv
        $list = Import-Csv -Path $FilePath
        # Retrieve existing retention compliance policy
        $policies = InvokePowerShellCmdlet "Get-RetentionCompliancePolicy"
        foreach($rp in $list)
        {
            # Cmdlet parameters
            $para = [String]::Empty;
            $name = [String]::Empty;
            $rpid = [String]::Empty;
            $cmdlet = 'New-RetentionCompliancePolicy'
            if ([String]::IsNullOrEmpty($rp.'Policy Name (Required)'))
            {
                WriteToLog -Type "Failed" -Message "Could not acquire table for
writing."
                throw;
            }
            else
            {
               $name = $rp.'Policy Name (Required)'
               $cmdlet += " -Name '" + $name + "'"
            }
            if ([String]::IsNullOrEmpty($rp.'Enabled (Required)'))
            {
                WriteToLog -Type "Failed" -Message "Could not acquire table for
writing."
                throw;
            }
            else
            {
                $enabled = $rp.'Enabled (Required)'
                $cmdlet += " -Enabled " + $enabled
            }
            if (![String]::IsNullOrEmpty($rp.'ExchangeLocation (Optional)'))
            {
                $para = $rp.'ExchangeLocation (Optional)'
                $cmdlet += " -ExchangeLocation " + $para
            }

            if (![String]::IsNullOrEmpty($rp.'ExchangeLocationException
```

```
(Optional)'))
            {
                $para = $rp.'ExchangeLocationException (Optional)'
                $cmdlet += " -ExchangeLocationException " + $para
            }
            if (![String]::IsNullOrEmpty($rp.'ModernGroupLocation (Optional)'))
            {
                $para = $rp.'ModernGroupLocation (Optional)'
                $cmdlet += " -ModernGroupLocation " + $para
            }
            if (![String]::IsNullOrEmpty($rp.'ModernGroupLocationException
(Optional)'))
            {
                $para = $rp.'ModernGroupLocationException (Optional)'
                $cmdlet += " -ModernGroupLocationException " + $para
            }
            if (![String]::IsNullOrEmpty($rp.'OneDriveLocation (Optional)'))
            {
                $para = $rp.'OneDriveLocation (Optional)'
                $cmdlet += " -OneDriveLocation " + $para
            }
            if (![String]::IsNullOrEmpty($rp.'OneDriveLocationException
(Optional)'))
            {
                $para = $rp.'OneDriveLocationException (Optional)'
                $cmdlet += " -OneDriveLocationException " + $para
            }
            if (![String]::IsNullOrEmpty($rp.'SharePointLocation (Optional)'))
            {
                $para = $rp.'SharePointLocation (Optional)'
                $cmdlet += " -SharePointLocation " + $para
            }
            if (![String]::IsNullOrEmpty($rp.'SharePointLocationException
(Optional)'))
            {
                $para = $rp.'SharePointLocationException (Optional)'
                $cmdlet += " -SharePointLocationException " + $para
            }
            if (![String]::IsNullOrEmpty($rp.'PublicFolderLocation (Optional)'))
            {
                $para = $rp.'PublicFolderLocation (Optional)'
                $cmdlet += " -PublicFolderLocation " + $para
            }
            if (![String]::IsNullOrEmpty($rp.'SkypeLocation (Optional)'))
            {
                $para = $rp.'SkypeLocation (Optional)'
                $cmdlet += " -SkypeLocation " + $para
            }
            if (![String]::IsNullOrEmpty($rp.'SkypeLocationException (Optional)'))
            {
                $para = $rp.'SkypeLocationException (Optional)'
                $cmdlet += " -SkypeLocationException " + $para
            }
            # If the policy already exists, skip for creation
            if (($policies -eq $null) -or ($policies | ? { $_.Name.ToLower() -eq
```

```
$name.ToLower() }) -eq $null)
            {
                # Create retention compliance policy
                $msg = "Execute Cmdlet : {0}" -f $cmdlet

                $ret = invokepowershellcmdlet $cmdlet

                if ($ret -eq $null)
                {
                    WriteToLog -Type "Failed" $error[0]
                    break;
                }
                $rpid = $ret.Guid
            }
            else
            {
                WriteToLog -Type "Warning" -Message "The policy '$name' already
exists! Skip for creation!"
                $rpid = ($policies | ? { $_.Name.ToLower() -eq $name.ToLower()
}).Guid
            }

            # Retrieve tag name for publishing
            $ts = $rp.'PublishComplianceTag (Required)'
            $tagList = $ts.Split(",") | ForEach-Object { $_.Trim() }

            WriteToLog -Type "Message" -Message "Publish Tags : '$ts'"

            PublishComplianceTag -PolicyGuid $rpid -TagName $tagList
        }
    }
    catch
    {
        WriteToLog -Type "Failed" "Error in input"
    }
}
# -------------------
# Publish Compliance Tag
# -------------------
Function PublishComplianceTag
{
    Param(
        [Parameter(Mandatory = $true)]
        [String]$PolicyGuid,
        [Parameter(Mandatory = $true)]
        [String[]]$TagNames
    )

    WriteToLog -Type "Start" "Start to Publish Compliance Tag"
    try
    {
        # Retrieve existing rule related to the given compliance policy
        $rule = InvokePowerShellCmdlet ("Get-RetentionComplianceRule -Policy {0}"
-f $PolicyGuid)
        $tagGuids = New-Object System.Collections.ArrayList
```

```
        foreach ($tn in $TagNames)
        {
            $t = InvokePowerShellCmdlet ("Get-ComplianceTag {0}" -f $tn)
            $tagGuids.Add($t.Guid) | Out-Null
        }
        if ($rule -ne $null)
        {
            foreach ($r in $rule)
            {
                if ([String]::IsNullOrEmpty($r.PublishComplianceTag))
                {
                    continue;
                }
                else
                {
                    $tl = $r.PublishComplianceTag.Split(",")
                    if ($tagGuids.Contains([GUID]$tl[0]))
                    {
                        $tagGuids.Remove([GUID]$tl[0]);
                    }
                }
            }
        }

        foreach($t in $tagGuids)
        {
            # Publish compliance tag
            $cmdlet = "New-RetentionComplianceRule -Policy {0} -
PublishComplianceTag {1}" -f $PolicyGuid, $t
            $ret = InvokePowerShellCmdlet $cmdlet

            if ($ret -eq $null)
            {
                WriteToLog -Type "Failed" $error[0]
                break;
            }
        }
    }
    catch
    {
        WriteToLog -Type "Failed" "Error in input"
    }
}
# -------------------
# Export All Labels Created in The Process
# -------------------
Function ExportCreatedComplianceTag
{
    Param(
        [Parameter(Mandatory = $true)]
        [String]$LabelFilePath
    )

    WriteToLog -Type "Start" "Start to Export Compliance Tag Created"
```

```
    try
    {
        # Import input csv
        $labels = Import-Csv $LabelFilePath
        # Create result table
        $tabName = "ResultTable"
        $table = New-Object system.Data.DataTable "$tabName"
        $col1 = New-Object system.Data.DataColumn Name,([string])
        $col2 = New-Object system.Data.DataColumn Comment,([string])
        $col3 = New-Object system.Data.DataColumn IsRecordLabel,([string])
        $col4 = New-Object system.Data.DataColumn RetentionAction,([string])
        $col5 = New-Object system.Data.DataColumn RetentionDuration,([string])
        $col6 = New-Object system.Data.DataColumn RetentionType,([string])
        $col7 = New-Object system.Data.DataColumn ReviewerEmail,([string])

        # Add the Columns
        $table.columns.add($col1)
        $table.columns.add($col2)
        $table.columns.add($col3)
        $table.columns.add($col4)
        $table.columns.add($col5)
        $table.columns.add($col6)
        $table.columns.add($col7)
        foreach($lab in $labels)
        {
            $t = InvokePowerShellCmdlet ("Get-ComplianceTag '{0}' " -f $lab.'Name
(Required)')

            # Create a result row
            $row = $table.NewRow()
            $row['Name'] = $t.Name
            $row['Comment'] = $t.Comment
            $row['IsRecordLabel'] = $t.IsRecordLabel
            $row['RetentionAction'] = $t.RetentionAction
            $row['RetentionDuration'] = $t.RetentionDuration
            $row['RetentionType'] = $t.RetentionType
            $row['ReviewerEmail'] = $t.ReviewerEmail

            # Add the row to the table
            $table.Rows.Add($row)
        }
        $table | Export-Csv $tagRetFile -NoTypeInformation
    }
    catch
    {
        WriteToLog -Type "Failed" "Error in exporting results."
    }
}
# --------------------
# Export All Published Labels and Policies in The Process
# --------------------
Function ExportPublishedComplianceTagAndPolicy
{
    Param(
        [Parameter(Mandatory = $true)]
```

```
        [String[]]$PolicyFilePath
    )

    WriteToLog -Type "Start" "Start to Export Published Compliance Tag and Policy"
    try
    {
        # Import input csv
        $policies = Import-Csv $PolicyFilePath
        # Create result table
        $tabName = "ResultTable"
        $table = New-Object system.Data.DataTable "$tabName"
        $col1 = New-Object system.Data.DataColumn 'Policy Name',([string])
        $col2 = New-Object system.Data.DataColumn PublishComplianceTag,([string])
        $col3 = New-Object system.Data.DataColumn Comment,([string])
        $col4 = New-Object system.Data.DataColumn Enabled,([string])
        $col5 = New-Object system.Data.DataColumn ExchangeLocation,([string])
        $col6 = New-Object system.Data.DataColumn ExchangeLocationException,
([string])
        $col7 = New-Object system.Data.DataColumn ModernGroupLocation,([string])
        $col8 = New-Object system.Data.DataColumn ModernGroupLocationException,
([string])
        $col9 = New-Object system.Data.DataColumn OneDriveLocation,([string])
        $col10 = New-Object system.Data.DataColumn OneDriveLocationException,
([string])
        $col11 = New-Object system.Data.DataColumn PublicFolderLocation,([string])
        $col12 = New-Object system.Data.DataColumn SharePointLocation,([string])
        $col13 = New-Object system.Data.DataColumn SharePointLocationException,
([string])
        $col14 = New-Object system.Data.DataColumn SkypeLocation,([string])
        $col15 = New-Object system.Data.DataColumn SkypeLocationException,
([string])

        # Add the Columns
        $table.columns.add($col1)
        $table.columns.add($col2)
        $table.columns.add($col3)
        $table.columns.add($col4)
        $table.columns.add($col5)
        $table.columns.add($col6)
        $table.columns.add($col7)
        $table.columns.add($col8)
        $table.columns.add($col9)
        $table.columns.add($col10)
        $table.columns.add($col11)
        $table.columns.add($col12)
        $table.columns.add($col13)
        $table.columns.add($col14)
        $table.columns.add($col15)
        foreach($policy in $policies)
        {
            $t = InvokePowerShellCmdlet ("Get-RetentionCompliancePolicy '{0}' -
DistributionDetail" -f $policy.'Policy Name (Required)')

            # Create a result row
            $row = $table.NewRow()
```

```
            $row['Policy Name'] = $t.Name

            $rules = InvokePowerShellCmdlet ("Get-RetentionComplianceRule -Policy
{0}" -f $t.Guid)
            $tagList = [String]::Empty
            foreach($rule in $rules)
            {
                if ([String]::IsNullOrEmpty($rule.PublishComplianceTag) -eq
$False)
                {
                    $tName = $rule.PublishComplianceTag.Split(',')[1]
                    $tagList = [String]::Concat($tagList, $tName, ",")
                }
            }
            if (![String]::IsNullOrEmpty($tagList))
            {
                $tagList = $tagList.Substring(0, $tagList.LastIndexOf(','))
            }
            $row['PublishComplianceTag'] = $tagList
            $row['Comment'] = $t.Comment
            $row['Enabled'] = $t.Enabled
            $row['ExchangeLocation'] = $t.ExchangeLocation
            $row['ExchangeLocationException'] = $t.ExchangeLocationException
            $row['ModernGroupLocation'] = $t.ModernGroupLocation
            $row['ModernGroupLocationException'] = $t.ModernGroupLocationException
            $row['OneDriveLocation'] = $t.OneDriveLocation
            $row['OneDriveLocationException'] = $t.OneDriveLocationException
            $row['PublicFolderLocation'] = $t.PublicFolderLocation
            $row['SharePointLocation'] = $t.SharePointLocation
            $row['SharePointLocationException'] = $t.SharePointLocationException
            $row['SkypeLocation'] = $t.SkypeLocation
            $row['SkypeLocationException'] = $t.SkypeLocationException

            # Add the row to the table
            $table.Rows.Add($row)
        }
        $table | Export-Csv $tagPubRetFile -NoTypeInformation
    }
    catch
    {
        WriteToLog -Type "Failed" "Error in exporting results."
    }
}
# Create compliance tag
CreateComplianceTag -FilePath $LabelListCSV
# Create retention policy and publish compliance tag with the policy
CreateRetentionCompliancePolicy -FilePath $PolicyListCSV
# Export to result csv
if ($ResultCSV)
{
    ExportCreatedComplianceTag -LabelFilePath $LabelListCSV
    ExportPublishedComplianceTagAndPolicy -PolicyFilePath $PolicyListCSV
}
```

Step 4: Run the PowerShell script

First, Connect to Security & Compliance PowerShell.

Then, run the script that creates and publishes the retention labels:

1. In your Security & Compliance PowerShell session, enter the path, followed by the characters `.\` and the file name of the script, and then press ENTER to run the script. For example:

   ```
   PowerShell

   <path>.\CreateRetentionSchedule.ps1
   ```

2. The script prompts you for the locations of the .csv files that you created in the previous steps. Enter the path, followed by the characters `.\` and file name of the .csv file, and then press ENTER. For example, for the first prompt:

   ```
   PowerShell

   <path>.\Labels.csv
   ```

Step 5: View the log file with the results

Use the log file that the script created to check the results and identify any failures that need resolving.

You can find the log file at the following location, although the digits in the example file name vary.

```
DOS

<path>.\Log_Publish_Compliance_Tag_01112018_151239.txt
```

Use retention labels to manage the lifecycle of documents stored in SharePoint

09/26/2025

> *Microsoft 365 licensing guidance for security & compliance*.

This article describes how you can manage the lifecycle of documents that are stored in SharePoint by using automatically applied retention labels and event-based retention.

The auto-apply functionality can use SharePoint metadata for document classification. The example in this article is for product-related documents, but the same concepts can be used for other scenarios. For example, in the oil and gas industry, you could use it to manage the lifecycle of documents about physical assets such as oil platforms, well logs, or production licenses. In the financial services industry, you could manage bank account, mortgage, or insurance contract documents. In the public sector, you could manage construction permits or tax forms.

This article looks at the information architecture and definition of the retention labels, classifies documents by auto-applying the labels, then generates the events that start the retention period.

Information architecture

Our scenario is a manufacturing company that uses SharePoint to store all the documents about the products that the company develops. These documents include product specifications, agreements with suppliers, and user manuals. When these documents are stored in SharePoint through Enterprise Content Management policies, document metadata is defined to classify them. Each document has the following metadata properties:

- **Doc Type** (product specification, agreement, or user manual)
- **Product Name** (Spinning Widget, Flying Widget, Crawling Widget)
- **Status** (draft or final)

This metadata forms a base content type called *Production Document* for all the documents.

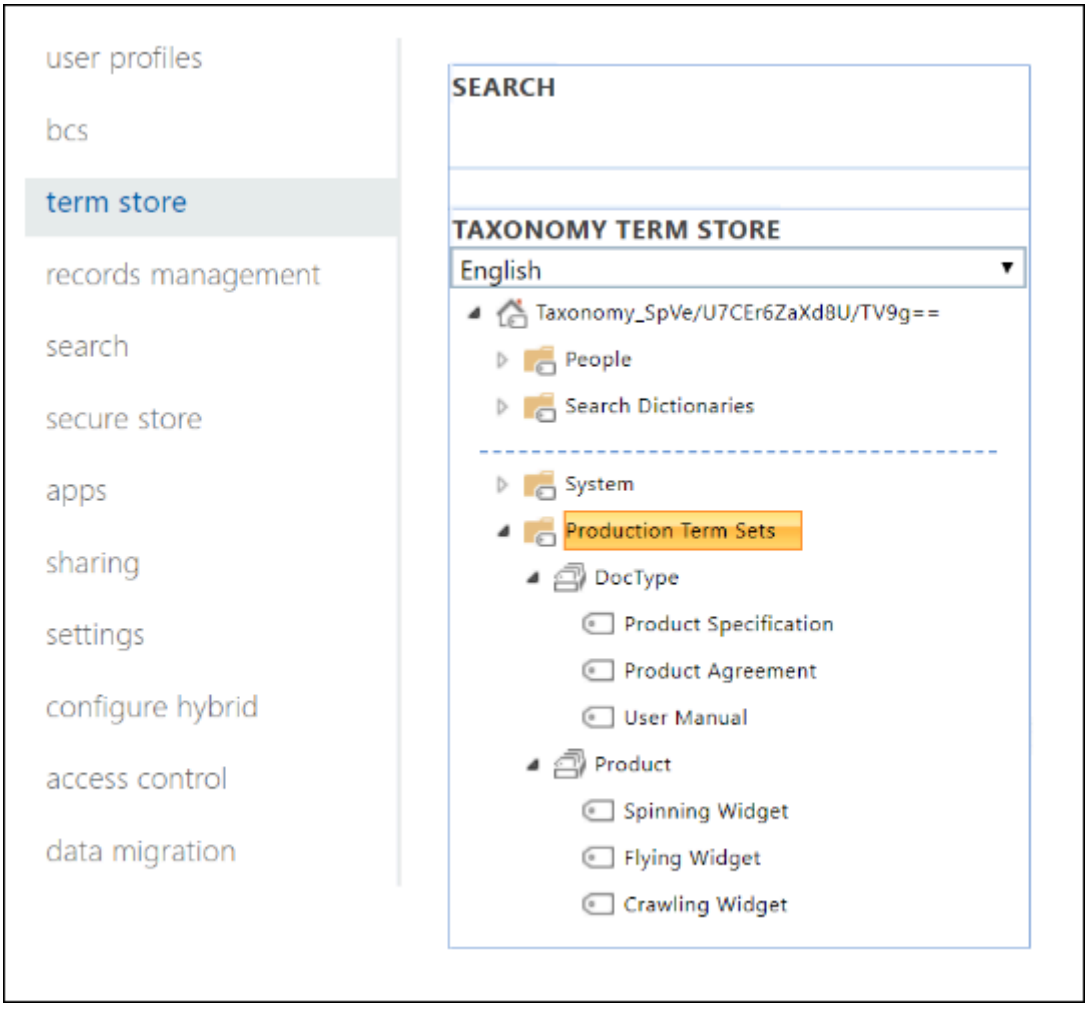
Columns

Name	Type	Status	Source
Name	File	Required	Document
Title	Single line of text	Optional	Item
Doc Type	Managed Metadata	Required	
Product Name	Managed Metadata	Required	
Status	Choice	Required	

> ⓘ **Note**
>
> The **Doc Type** and **Status** properties are used later in this scenario to classify and auto-apply retention labels.

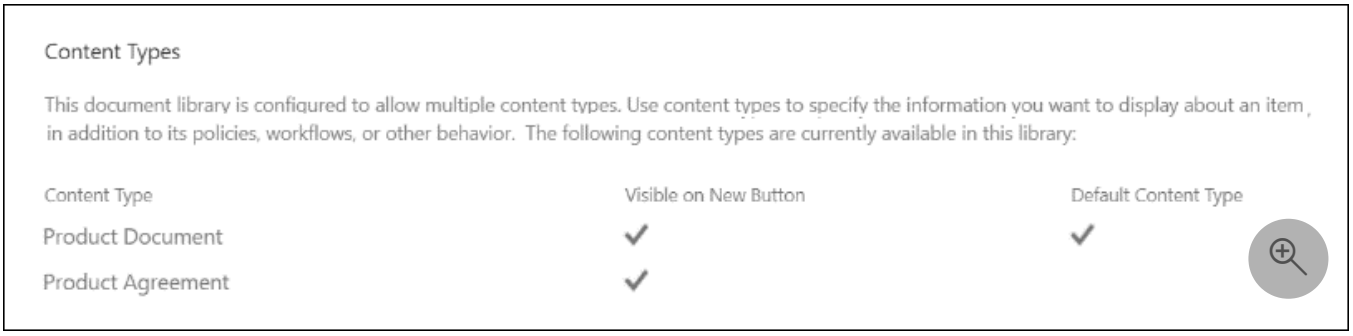
We might have several content types that represent different types of documents. This scenario focuses on the product documentation.

Here we use the Managed Metadata service and the Term Store to create a term set for *Doc Type* and for *Product Name*. For each term set, we create a term for each value. The outcome looks like this in Term Store for our SharePoint organization:



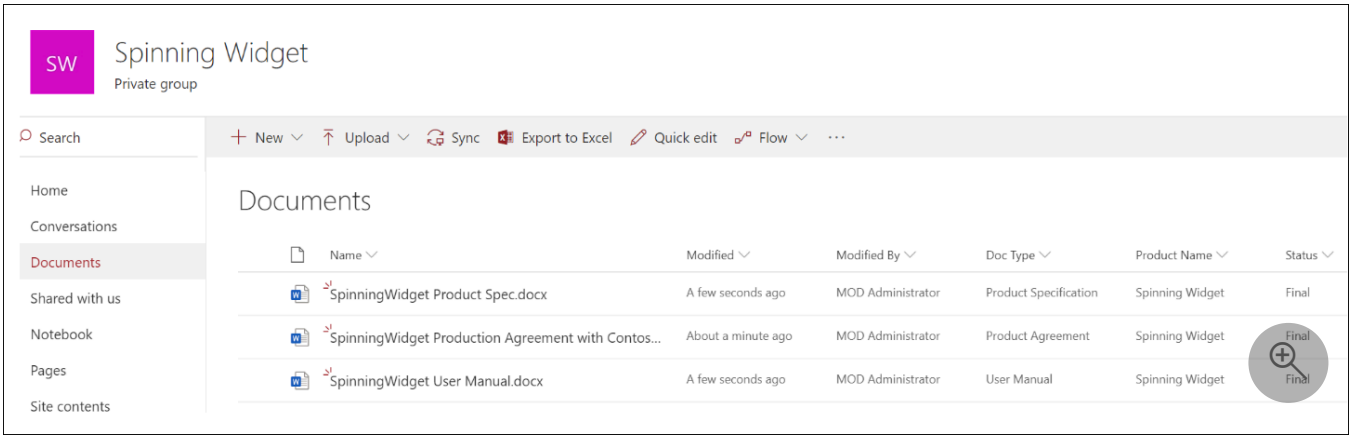
Content Type can be created and published by using the Content Type Hub. You can also create and publish a content type by using site provisioning tools, such as the PnP provisioning framework or site design JSON schema.

Each product has a dedicated SharePoint site that contains one document library that has the right content types enabled. All documents are stored in this document library.



> **Note**
>
> In this article we focus on documents and use a site, but the manufacturing company in this scenario could instead use a Microsoft Team per product to support collaboration activties such as persistent chat and using the **Files** tab in Teams for document management.

A view of the document library for the Spinning Widget product:



With the basic information architecture in place for document management, next look at the retention and disposal strategy for the documents that use the metadata and how they're classified.

Retention and disposition

The manufacturing company's compliance and data governance policies dictate how data is preserved and disposed of. Product-related documents must be kept for as long as the

product is manufactured and retained for a period of time afterwards. The retention time period differs for product specifications, agreements, and user manuals. The following table indicates the retention and disposition requirements:

Expand table

Document type	Retention	Disposition Requirements
Product specifications	5 years after production stops	Delete
Product agreements	10 years after production stops	Review
User manuals	5 years after production stops	Delete
All other types of documents	Don't actively retain	Delete when document is older than 3 years. A document is considered older than 3 years if not modified within the last 3 years.

The Microsoft Purview portal is used to create the following retention labels:

- Product Specification
- Product Agreement
- User Manual

In this article, we show how to create and auto-apply the Product Specification retention label. To implement the complete scenario, also create and auto-apply retention labels for the other two document types.

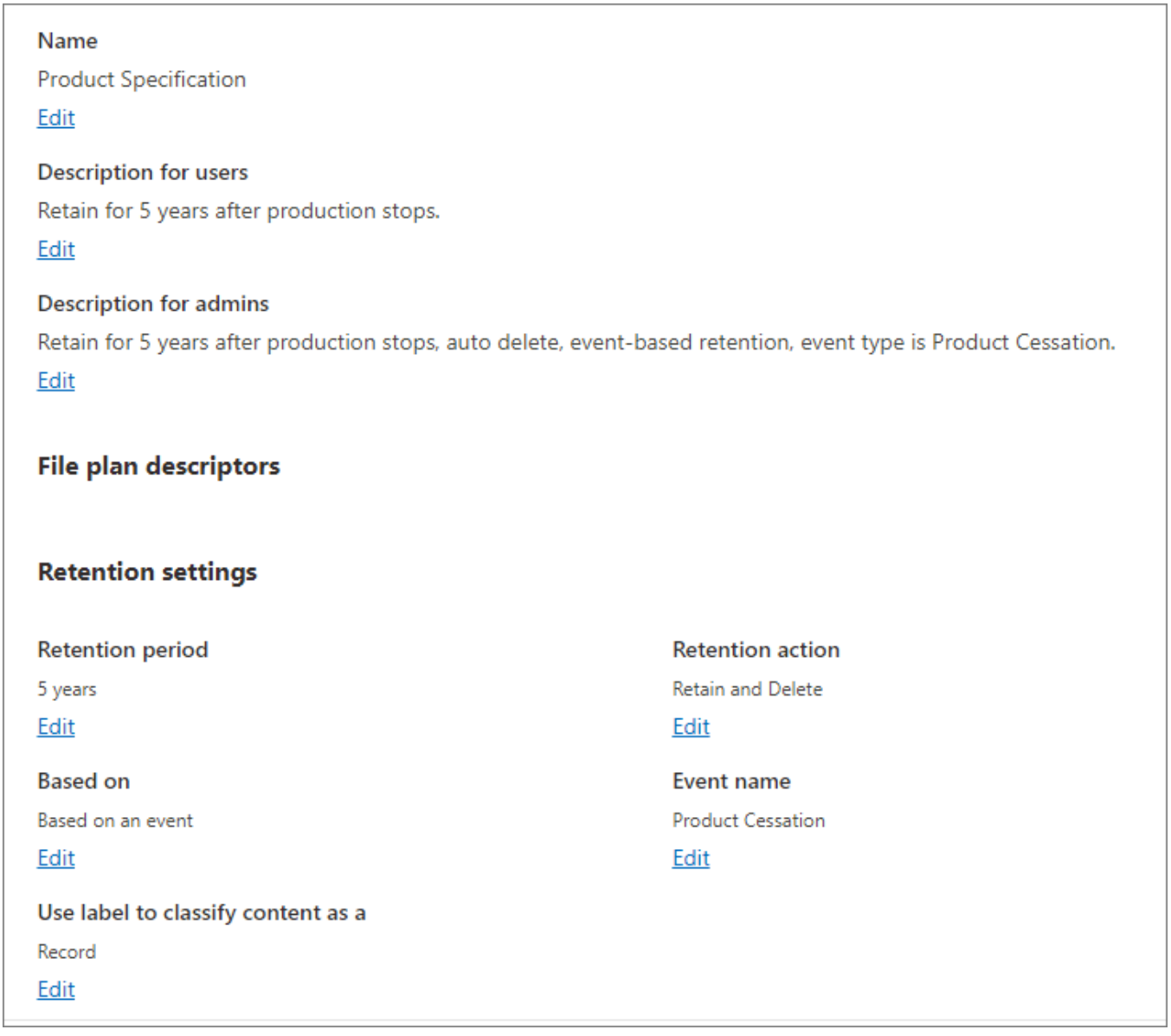
Settings for the Product Specification retention label

The file plan for the Product Specification retention label:

- **Name:** Product Specification
- **Description for users:** Retain for 5 years after production stops.
- **Description for admins:** Retain for 5 years after production stops, auto delete, event-based retention, event type is *Product Cessation*.
- **Retention action:** Retain and delete.

- **Retention duration:** 5 years (1,825 days).
- **Record label**: Configure the retention label to mark items as a record, which means users can't modify or delete the labeled documents.
- **File plan descriptors:** To simplify the scenario, no optional file descriptors are provided.

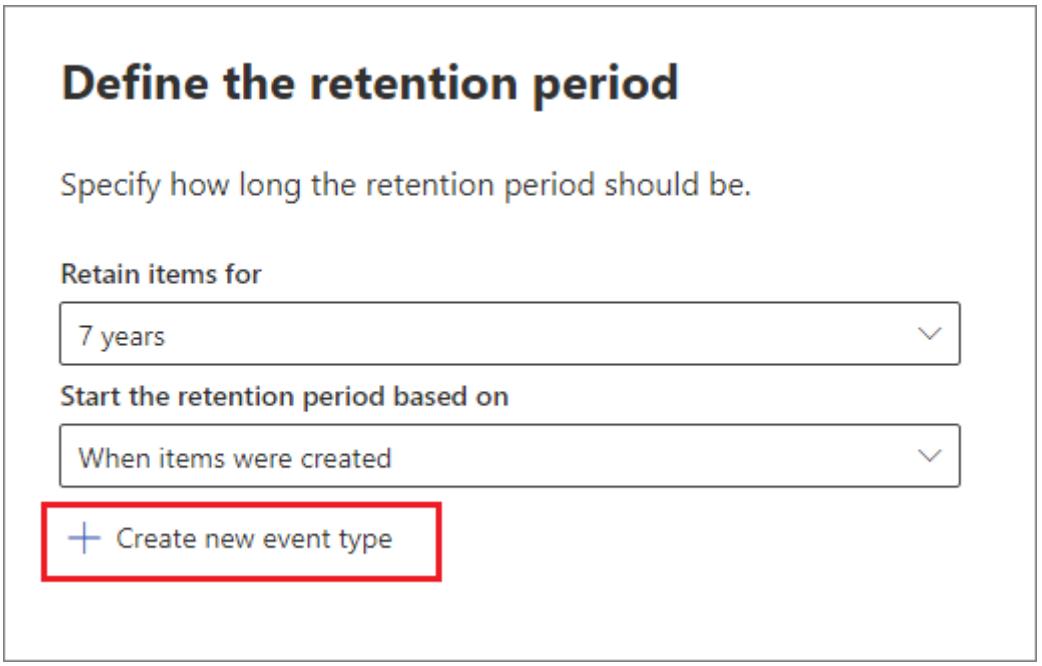
The following example shows a summary of the label settings when you create the Product Specification retention label in the Microsoft Purview portal. You can create the *Product Cessation* event type before you create the retention label, or during the retention label creation process.



> **Note**
>
> If you're recreating this scenario in a test enviornment, set the retention duration to a shorter period of time to avoid a five-year wait for document deletion.

Create an event type when you create a retention label

1. On the **Define the retention period** page of the retention label workflow, after **Start the retention period based on**, select **Create new event type**:



2. On the **Name your event type** page, enter **Product Cessation** and an optional description. Then select **Next**, **Submit**, and **Done**.

3. Back on the **Define the retention period** page, for **Start the retention period based on**, use the dropdown box to select the **Product Cessation** event type that you created.

> **Tip**
>
> For more detailed steps, see **Create a label whose retention period is based on an event**.

The next section shows how to auto-apply the retention label to product specification content.

Auto-apply retention labels to documents

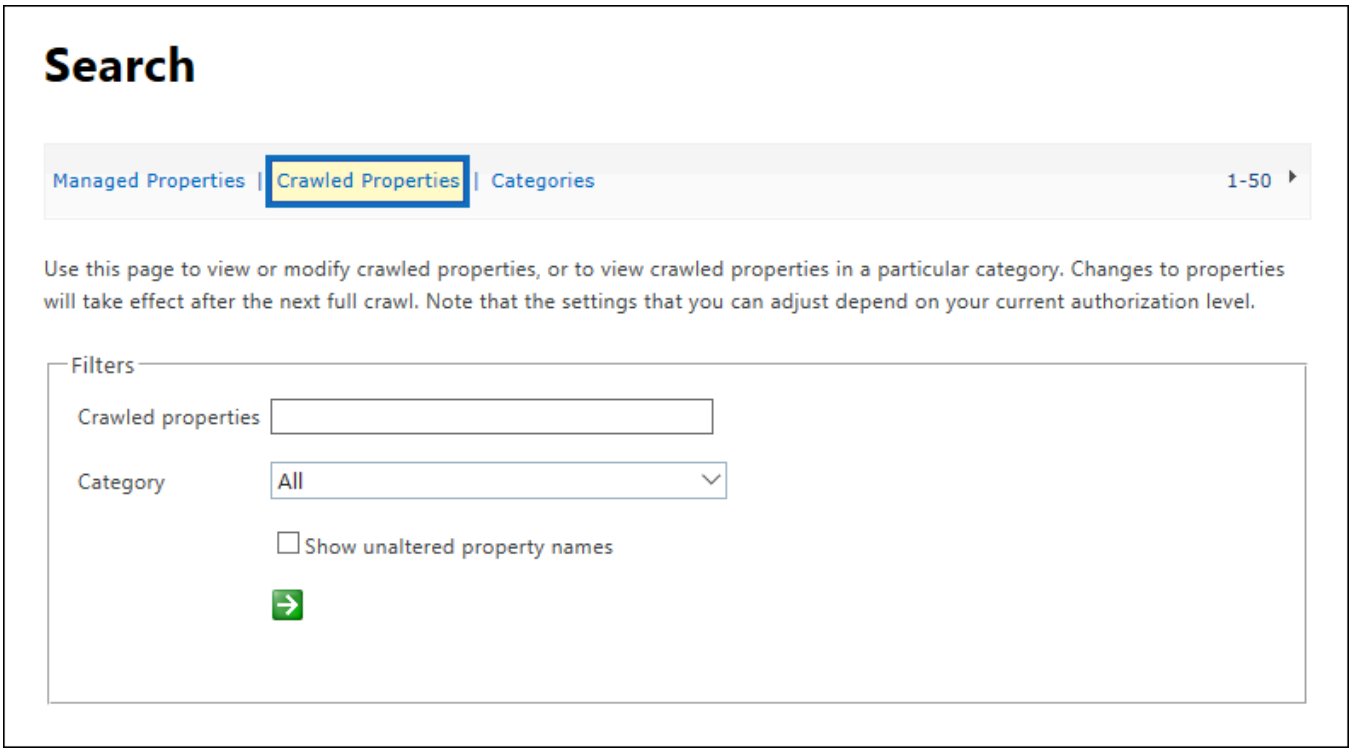
We're going to use Keyword Query Language (KeyQL) to auto-apply the retention labels that we created. KeyQL is the language used to build search queries. In KeyQL, you can search by using keywords or managed properties. For more information, see Keyword Query Language (KeyQL) syntax reference.

We need to configure the search schema to tell Microsoft 365 to "apply the **Product Specification** retention label to all documents that have a **Status** of **Final** and a **Doc Type** of

Product Specification." Recall that **Status** and **Doc Type** are the site columns that we defined for the Product Documentation content type in the Information architecture section.

When SharePoint indexes content, it automatically generates crawled properties for each site column. For this scenario, we're interested in the **Doc Type** and **Status** properties. We need documents in the library that are the right content type and have the site columns filled in for search to create the crawled properties.

In the SharePoint admin center, open the Search configuration, and select **Manage Search Schema** to view and configure the crawled properties.

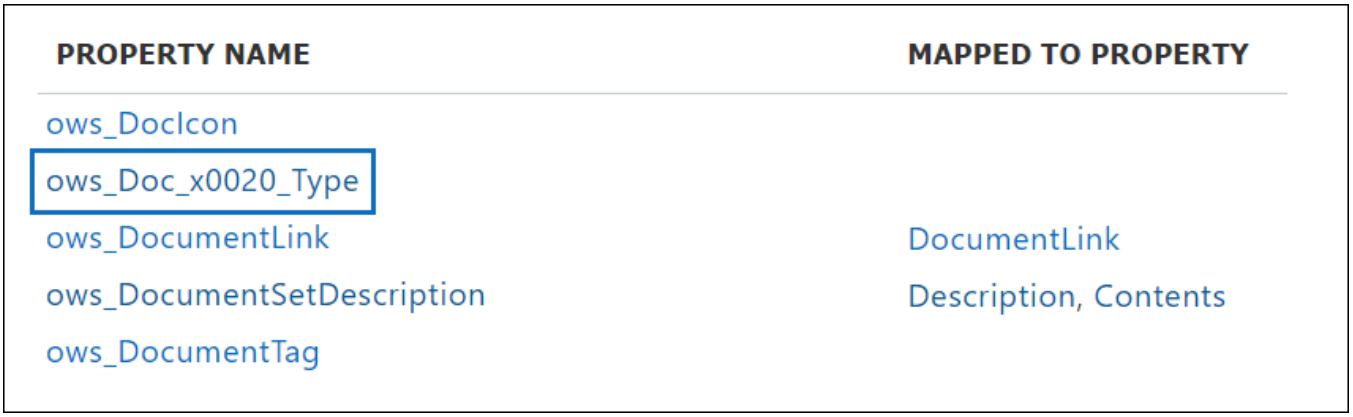


Run a search for ***status*** using the **Crawled properties** filter. This shows results that include "status" in the property name field:



The previous example shows the **ows__Status** property (notice the double underscore). This property maps to the **Status** property of the Production Document content type that is used in this scenario.

Run a search for ***ows_doc*** using the **Crawled properties** filter. This shows results that include "ows_doc" in the property name field:



PROPERTY NAME	MAPPED TO PROPERTY
ows_DocIcon	
ows_Doc_x0020_Type	
ows_DocumentLink	DocumentLink
ows_DocumentSetDescription	Description, Contents
ows_DocumentTag	

The above example shows the **ows_Doc_x0020_Type** property. This property maps to the **Doc Type** property of the Production Document content type that is used in this scenario.

> **Tip**
>
> To identify the name of a crawled property for this scenario:
>
> 1. Go to the document library that contains the production documents.
> 2. Go to the library settings.
> 3. For **Columns**, select the name of the column (for example, **Status** or **Doc Type**) to open the site column page.
> 4. The *Field* parameter in the URL for that page contains the name of the field. This field name, prefixed with "ows_", is the name of the crawled property. For example, the URL `https://tenantname.sharepoint.com/sites/SpinningWidget/_layouts/15/FldEdit.aspx?List=%7BC38C2F45-3BD6-4C3B-AA3B-EF5DF6B3D172%7D&Field=_Status` corresponds to the *ows__Status* crawled property.

If the crawled properties you're looking for don't appear in the Manage Search Schema section in SharePoint admin center:

- The documents might not be indexed yet. You can force a reindex of the library by going to **Document library settings** > **Advanced Settings**.
- If the document library is in a modern site, make sure that the SharePoint admin is also a site collection admin.

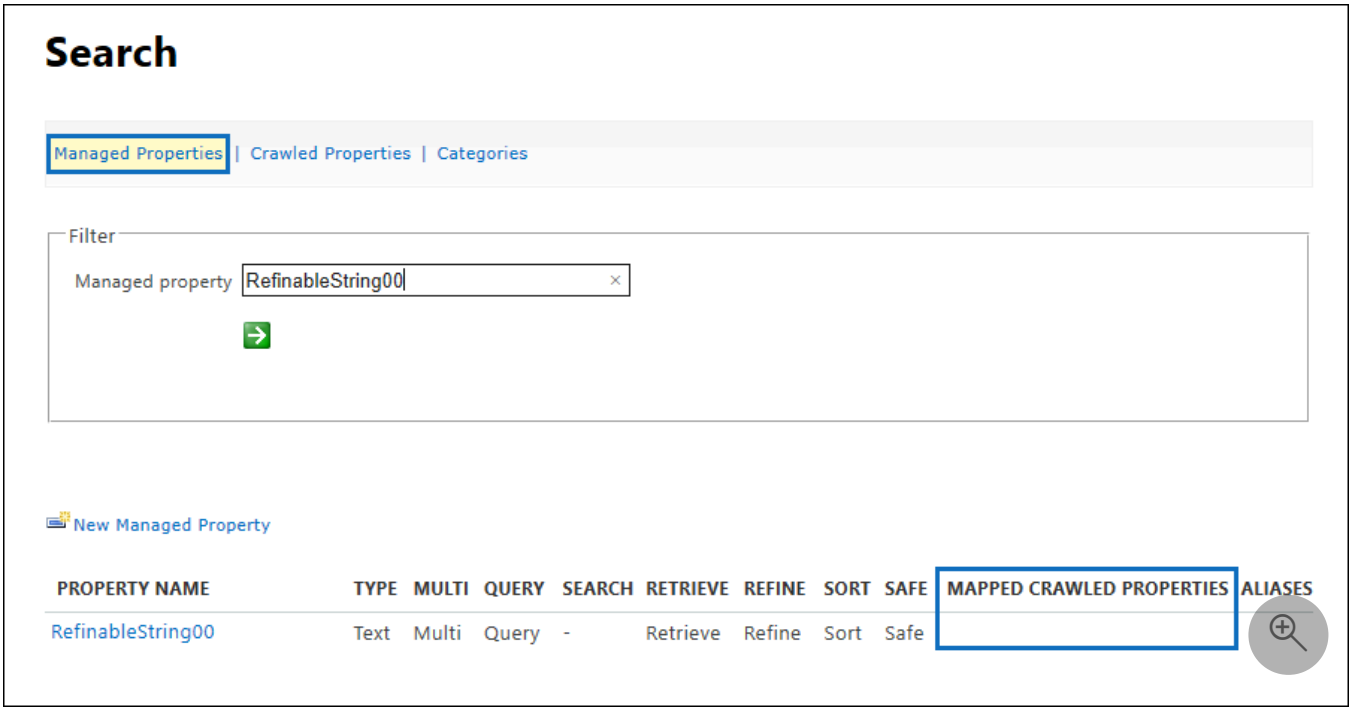
For more information about crawled and managed properties, see Automatically created managed properties in SharePoint Server.

Map crawled properties to predefined managed properties

KeyQL can't use crawled properties in search queries. It has to use a managed property. In a typical search scenario, create a managed property and map it to the crawled property. For auto-applying retention labels, you can only specify predefined managed properties in KeyQL (not custom managed properties). The system has a set of predefined (default) managed properties for string *RefinableString00* to *RefinableString199* that can be used. The default managed properties are typically used for defining search refiners. For a complete list, see Default unused managed properties.

For the KeyQL query to automatically apply the correct retention label to production document content, map the crawled properties *ows_Doc_x0020_Type and ows__Status* to two refinable managed properties.

In the scenario, **RefinableString00** and **RefinableString01** aren't being used. You can determine their use by looking at **Managed Properties** in **Manage Search Schema** in the SharePoint admin center:



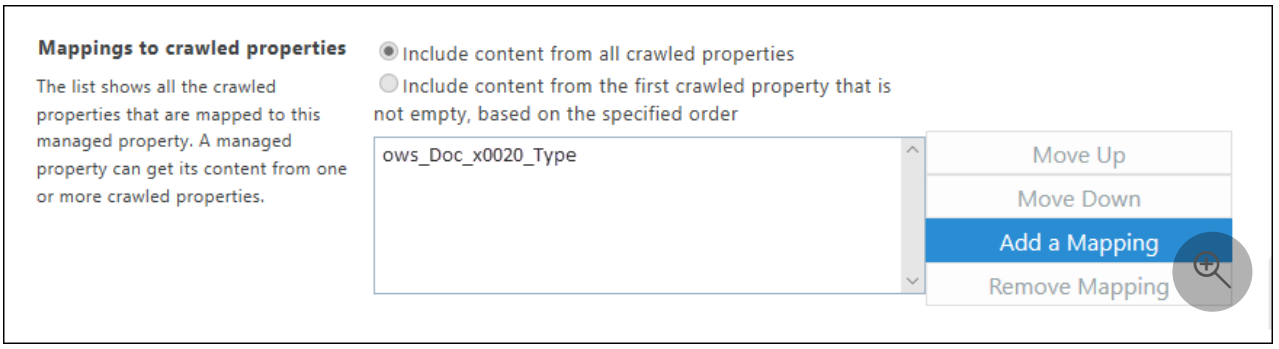
The **Mapped Crawled Properties** column in the previous example is empty for the managed property that isn't in use.

To map the **ows_Doc_x0020_Type** crawled property, follow these steps:

1. Run a search for ***RefinableString00*** using the **Managed property** filter.
2. In the results list, select the **RefinableString00** link, then navigate to the **Mappings to crawled properties** section.

3. Select **Add a Mapping**, and then run a search for ***ows_Doc_x0020_Type*** in the **Search for a crawled property name** box within the **Crawled property selection** window.

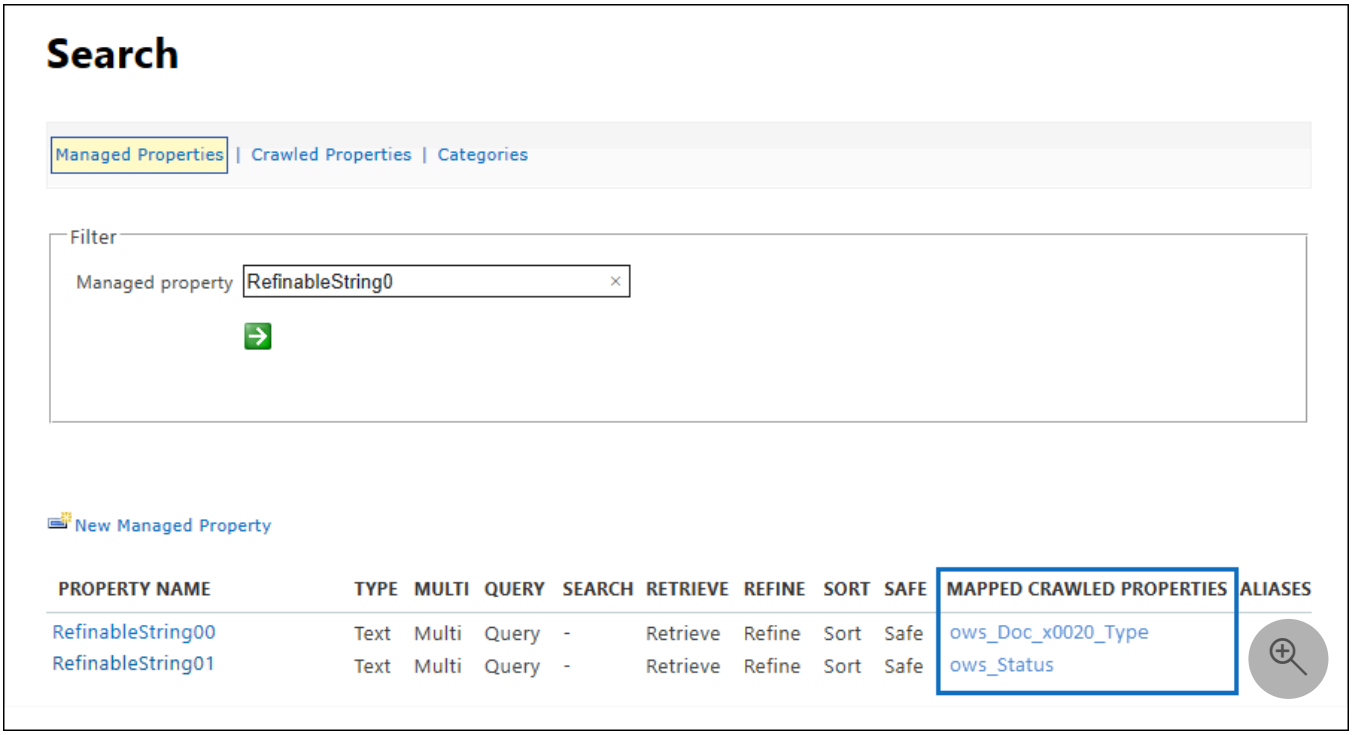
4. In the results list, select **ows_Doc_x0020_Type** and then select **OK**. The following example shows the new mapping in the **Mapped Crawled Properties** section:



5. Save the mapping.

Repeat these steps to map **RefinableString01** and **ows__Status**.

The following example shows the results of two managed properties mapped to the two crawled properties:



Setup can be verified by running an enterprise search. In a browser, go to *https://<your_tenant>.sharepoint.com/search*. Search for ***RefinableString00:"Product Specification"***. This search returns all documents that have a **Product Specification** of ***Doc Type***.

Next, search for **RefinableString00:"Product Specification" AND RefinableString01:Final**. This returns all documents that have **Product Specification** of ***Doc Type*** and a **Status** of ***Final***.

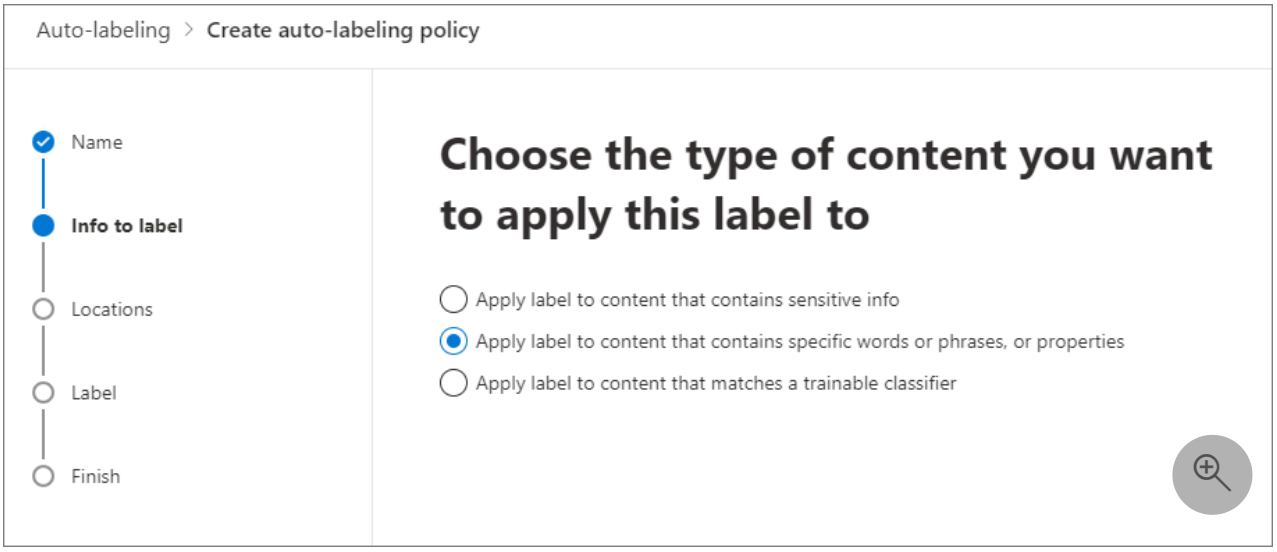
Create auto-apply label policies

After the KeyQL query is verified to return the expected documents, create an auto-apply label policy that uses a KeyQL query to automatically apply the Product Specification retention label to the appropriate documents.

1. Sign in to the Microsoft Purview portal > **Solutions** > **Records management** > **Policies** > **Label policies** > **Retention labels**

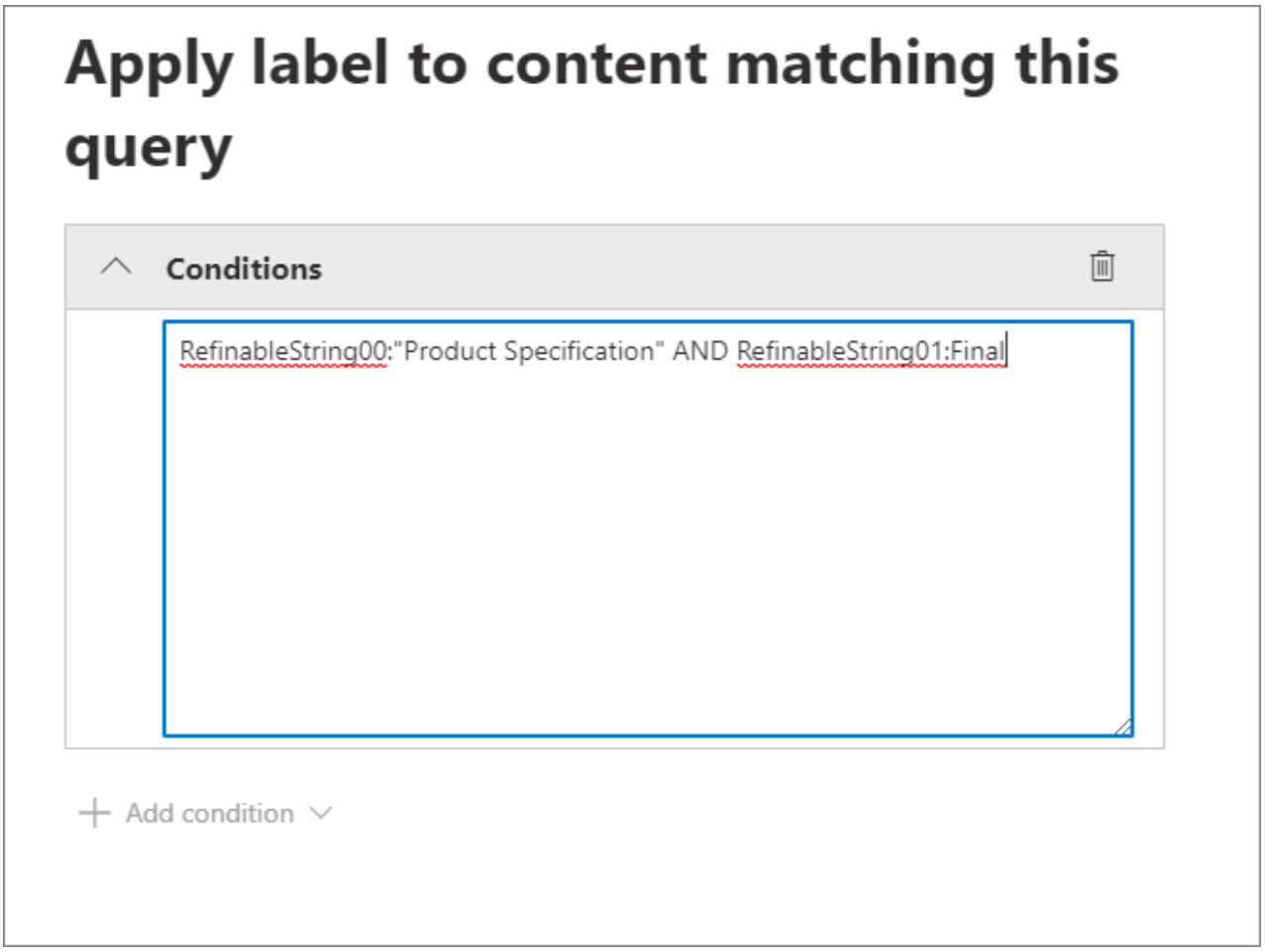
2. Select **Auto-apply a label**. In the auto-labeling policy configuration, on the **Name your auto-labeling policy** page, enter a name such as **Auto-apply Product Specification label**, and an optional description. Then select **Next**.

3. On the **Choose the type of content you want to apply this label to** page, select **Apply label to content that contains specific words or phrases, or properties**, then select **Next**.



   This option lets us provide the same KeyQL search query that we tested in the previous section. The query returns all Product Specification documents that have a status of *Final*. When the same query is used in the auto-apply label policy, the Product Specification retention label is automatically applied to documents that match it.

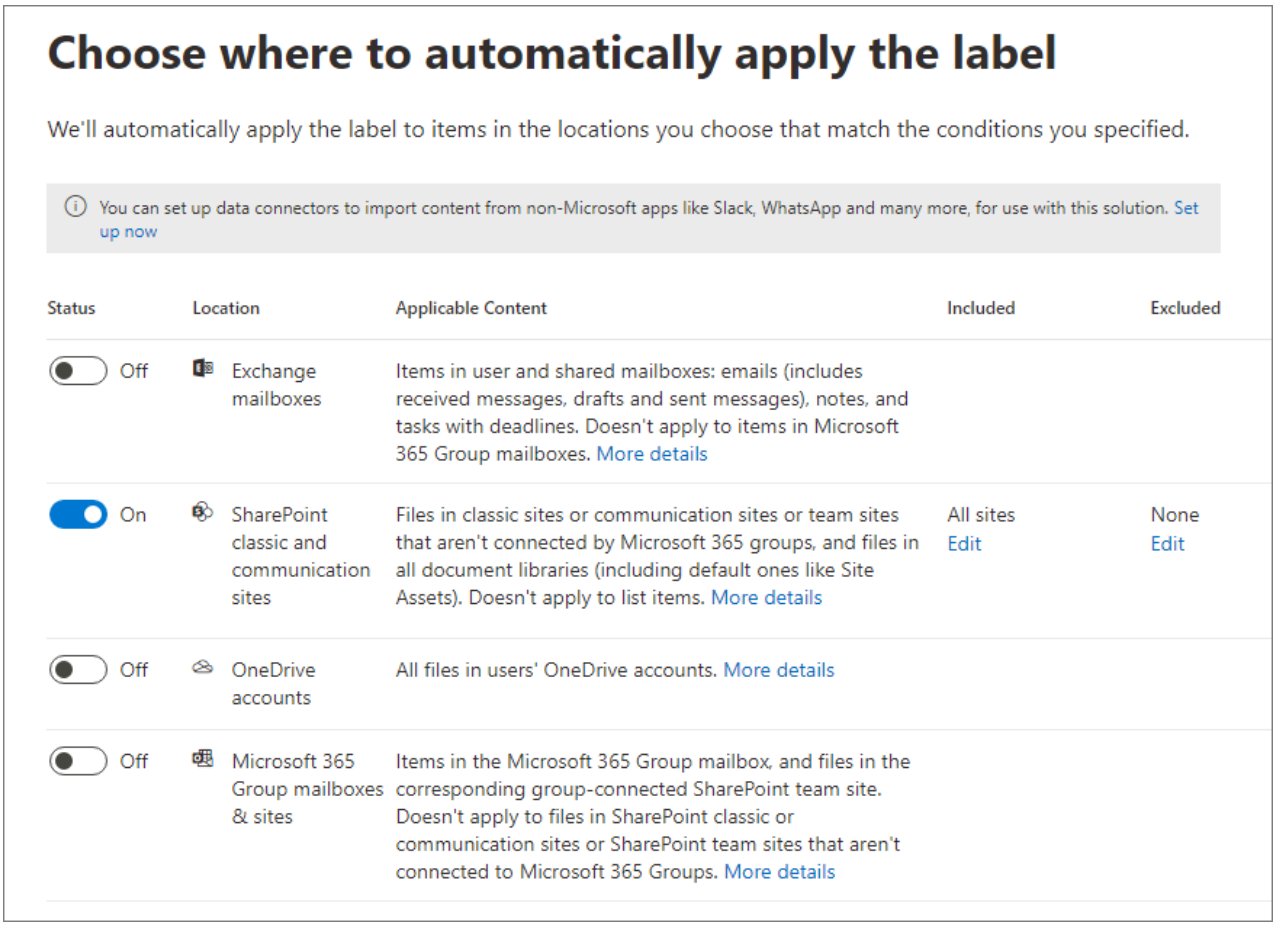
4. On the **Apply label to content matching this query** page, type **RefinableString00:"Product Specification" AND RefinableString01:Final**, then select **Next**.



5. For this scenario, choose **Static** on the **Choose the type of retention policy to create** page.

> **Note**
>
> For a production environment, we recommend adaptive scopes because they automatically update the locations included in the policy by using attributes and properties.

6. On the **Choose where to automatically apply the label** page, select the content locations that you want to apply the policy to. For this scenario, apply the policy only to **SharePoint classic and communication sites**. Toggle all other locations to **Off**:

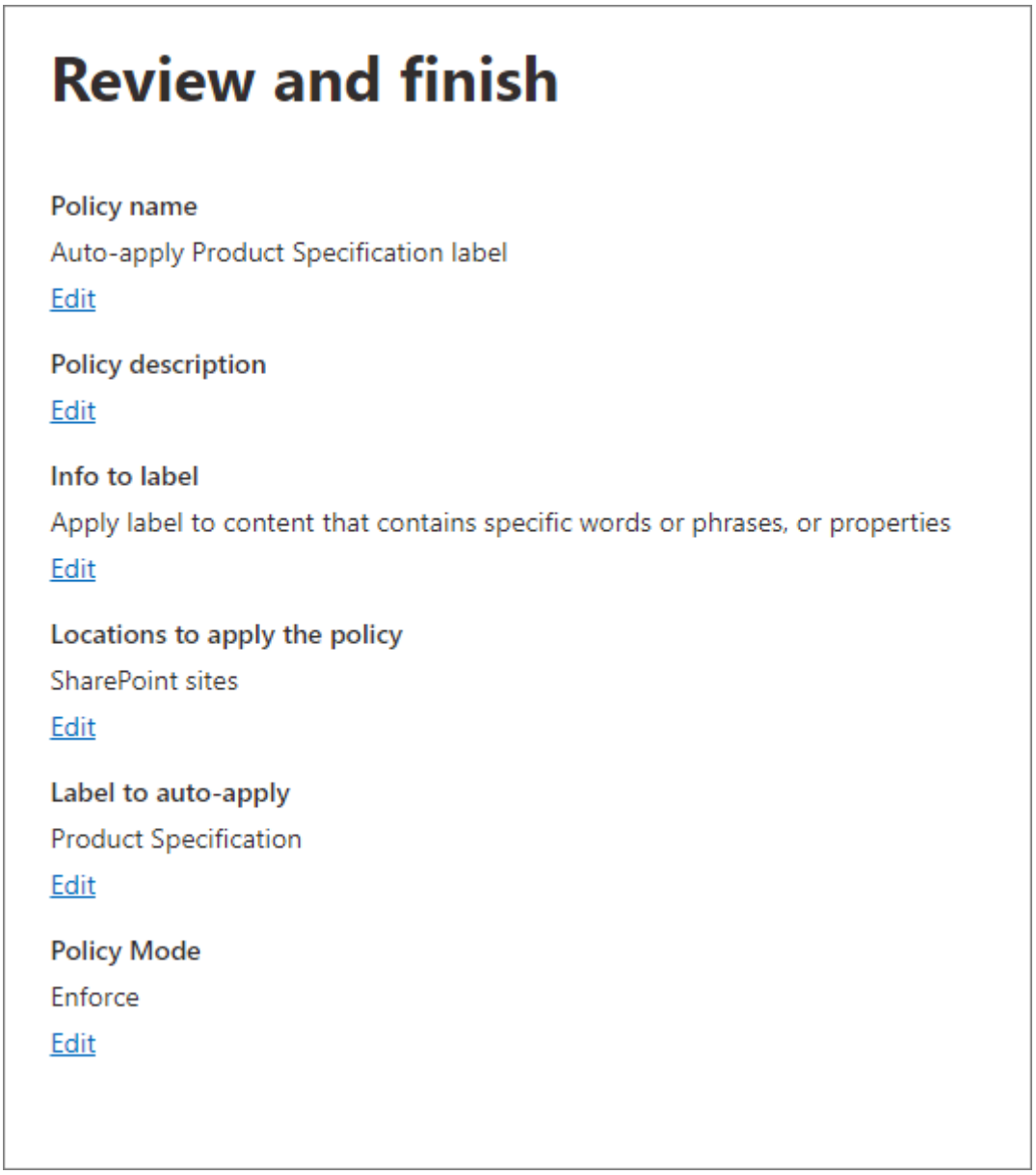


> **Tip**
>
> Instead of applying the policy to all SharePoint sites, you can apply the policy to individual sites by selecting **Edit** for the **Included** column and adding the site URLs.

7. On the **Choose a label to auto-apply** page, select the **Product Specification** label.

8. For this scenario, keep the default of **Turn on policy** on the **Decide whether to test or run your policy** page.

9. Review your settings to confirm accuracy before submitting:



10. Select **Submit** to create the auto-apply label policy.

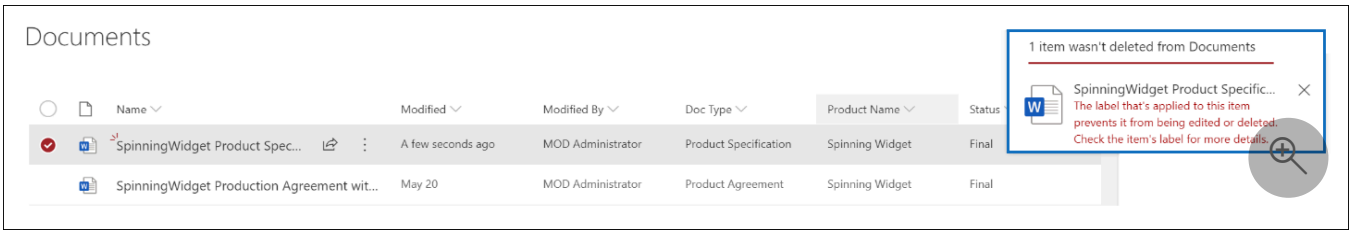
Verify that the retention label was automatically applied

After you've allowed sufficient time for the retention labels to be auto-applied, use activity explorer in the Microsoft Purview portal to verify that the auto-apply label policy applied retention labels to the product documents.

Within SharePoint, the properties of the documents in the Document Library show when a label is applied in the document details pane.



When retention labels mark items as a record, they prevent users from deleting those items. Users see an error message if they attempt to delete the document and the delete action is blocked:



Generate the event that triggers the retention period

In this scenario, an event defines when a product reaches the end of production. For product specification documents, the five-year retention period begins when the "end of production" event is triggered. This event triggers the start of the retention period defined in the retention label.

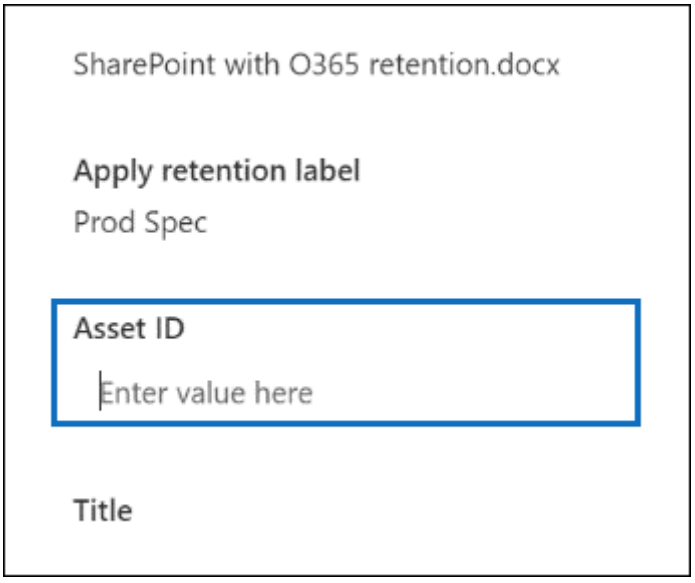
Events can be created in the Microsoft Purview portal. Choose the event type, set the correct asset IDs, and enter a date for the event. For more information, see Start retention when an event occurs.

You can also create and manage events using Microsoft Graph Records Management APIs.

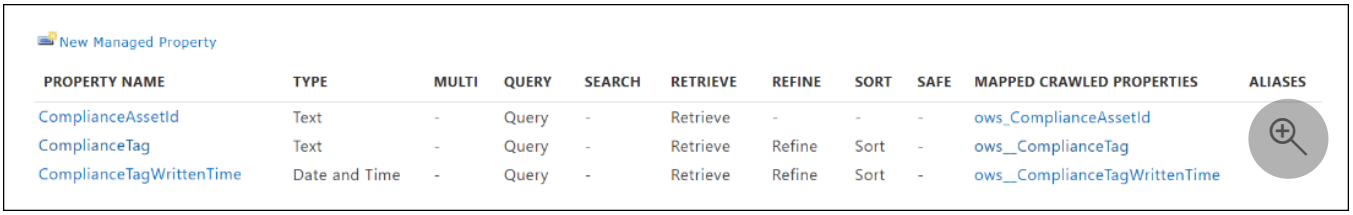
More about asset IDs

As explained in the Start retention when an event occurs article, it's important to understand the relationship between event types, retention labels, events, and asset IDs. The asset ID is a

document property in SharePoint and OneDrive. It helps identify the documents that have a retention period triggered by the event. By default, SharePoint has an **Asset ID** property that can be used for event-driven retention:



The following example shows the Asset ID managed property is **ComplianceAssetId**.



This scenario used the default **Asset ID** property. Other properties can also be used. If an Asset ID or keywords for an event aren't specified, the retention period is triggered for all the content that has a label of that event type.

Using advanced search in SharePoint

In the Asset Id example, there's another managed property related to retention labels called **ComplianceTag** that's mapped to a crawled property. The **ComplianceAssetId** managed property is also mapped to a crawled property. These managed properties can be used in advanced search to retrieve all documents that have a retention label applied.

Disposition of content

> *Microsoft Purview service description*

Use the **Disposition** page from **Records Management** in the Microsoft Purview portal to manage disposition reviews and view the metadata of items marked as records that have been automatically deleted at the end of their retention period.

Prerequisites for viewing content dispositions

To manage disposition reviews and confirm that items marked as records have been deleted, you must have sufficient permissions and auditing must be enabled. Also be aware of any limitations for disposition.

Permissions for disposition

To successfully access **Disposition** in the Microsoft Purview portal, users must have the **Disposition Management** role. This role is included in the **Records Management** default role group.

> ⓘ **Note**
>
> By default, a global admin isn't granted the **Disposition Management** role.

To grant users just the permissions they need for disposition reviews without granting them permissions to view and configure other features for retention and records management, create a custom role group (for example, named "Disposition Reviewers") and grant this group the **Disposition Management** role.

For instructions to add users to the default roles or create your own role groups, see Permissions in the Microsoft Purview portal.

Additionally:

- To view the contents of items during the disposition process, add users to the **Content Explorer Content Viewer** role group. If users don't have the permissions from this role group, they can still select a disposition review action to complete the disposition review,

but must do so without being able to view the item's contents from the mini-preview pane in the Microsoft Purview portal.

- By default, each person that accesses the **Disposition** page sees only items that they're assigned to review. For a records management administrator to see all items assigned to all users, and all retention labels that are configured for disposition review: From the Microsoft Purview portal, navigate to **Settings** > **Solution settings** > **Records Management ** > **Disposition** to select and then enable a mail-enabled security group that contains the administrator accounts. Permissions are then granted to the group members, but not the group owner.

  Microsoft 365 groups and security groups that aren't mail-enabled don't support this feature and don't display in the list to select. If you need to create a new mail-enabled security group, use the link to the Microsoft 365 admin center to create the new group.

  > **Important**
  >
  > After you've enabled the group, you can't change it in the Microsoft Purview portal. See the next section for how to enable a different group by using PowerShell.

- The Records Management settings are visible only to record management administrators.

Enabling another security group for disposition

After you've enabled a security group for disposition from the **Records Management settings** in the Microsoft Purview portal, you can't disable this permission for the group or replace the selected group in the portal. However, you can enable another mail-enabled security group by using the Enable-ComplianceTagStorage cmdlet.

For example:

PowerShell

```
Enable-ComplianceTagStorage -RecordsManagementSecurityGroupEmail
dispositionreviewers@contosoi.com
```

Enable auditing

Make sure that auditing is enabled at least one day before the first disposition action. For more information, see Search the audit log.

Disposition reviews

When content reaches the end of its retention period, there are several reasons why you might want to review that content and confirm whether it can be permanently deleted ("disposed"). For example, instead of deleting the content, you might need to:

- Suspend the deletion of relevant content for litigation or an audit.
- Assign a different retention period to the content, perhaps because the original retention settings were a temporary or provisional solution.
- Move the content from its existing location to an archive location, for example, if that content has research or historical value.

When a disposition review is triggered at the end of the retention period, your chosen reviewers receive an email notification that they have content to review. These reviewers can be individual users or members of a mail-enabled security group. When you use a mail-enabled security group, only group members and not the group owner receive the email notifications.

You can customize the notification email that reviewers receive, including instructions in different languages. For multi-language support, you must specify the translations yourself and this custom text is displayed to all reviewers irrespective of their locale.

Users receive an initial email notification per label at the end of the item's retention period, with a reminder per label once a week of all disposition reviews that they're assigned. They can click the link in the notification and reminder emails to go directly to the records management **Disposition** page in the Microsoft Purview portal to review the content and take an action. Alternately, the reviewers can navigate to this **Disposition** page in the Microsoft Purview portal. Then:

- Reviewers see only the disposition reviews that are assigned to them, whereas administrators who are added to the selected security group for records manager see all disposition reviews.
- Reviewers can add new users to the same disposition review. Note that this action does not automatically grant these added users the required permissions.
- For the disposition review process, a mini-review pane for each item shows a preview of the content if they have permissions to see it. If they don't have permissions, they can select the content link and request permissions. This mini-review pane also has tabs for additional information about the content:

- **Details** to display indexed properties, where it's located, who created it and when, and who last modified it and when.
- **History** that shows the history of any disposition review actions to date, with reviewer comments if available.

A disposition review can include content in Exchange mailboxes, SharePoint sites, and OneDrive accounts. Content pending a disposition review in those locations is permanently deleted only after a reviewer for the final stage of disposition chooses to permanently delete the content.

> **Note**
>
> A mailbox must have at least 10 MB data to support disposition reviews.

Administrators can see an overview of all pending dispositions in the **Overview** tab. Reviewers see only their items pending disposition.

Select the **View all pending dispositions**, to view them the **Disposition** page.

Disposition timelines

Disposition involves three separate timelines:

Expand table

Timeline	Length	Notes
Permanent deletion after approval	Within 15 days after the final (or only) reviewer selects **Approve disposal**	Applies when an item is marked as eligible for permanent deletion.
Auto-approval timeout (optional)	7-365 days; default 14 days	If reviewers take no manual action, the item advances to the next stage or, in the final stage, is disposed.
Configuration change propagation	Up to 7 days	Applies when you turn auto-approval on or off, or change the number of days, like all retention label changes.

Workflow summary for a disposition review

Basic workflow for a disposition review, single stage:

1. The admin configures a retention label to start a disposition review at the end of the retention period.
2. The admin publishes the label and a user applies the label to one or more items, or the admin automatically applies the retention label to items.
3. At the end of the retention period for each item, the designated disposition reviewers receive an email to review the item for disposition.
4. One of the reviewers selects the link to review the item in the disposition page from the Microsoft Purview portal, and confirms that the item should be permanently deleted.

Auto-approval for disposition

You can optionally specify a time period (7-365 days) for auto-approval. The default period if you select this option is 14 days.

If designated reviewers don't take manual action during this time period by using the standard disposition review process, the item automatically passes to the next review stage. If the item is in the final review stage, the item is automatically disposed with permanent deletion.

> **Important**
>
> If you configure this option and items are already pending disposition review, they automatically become auto-approved if they have already exceeded the number of days that you specified for auto-approval. The time period always starts from when the item is ready for disposition review and not from when you configure the option.

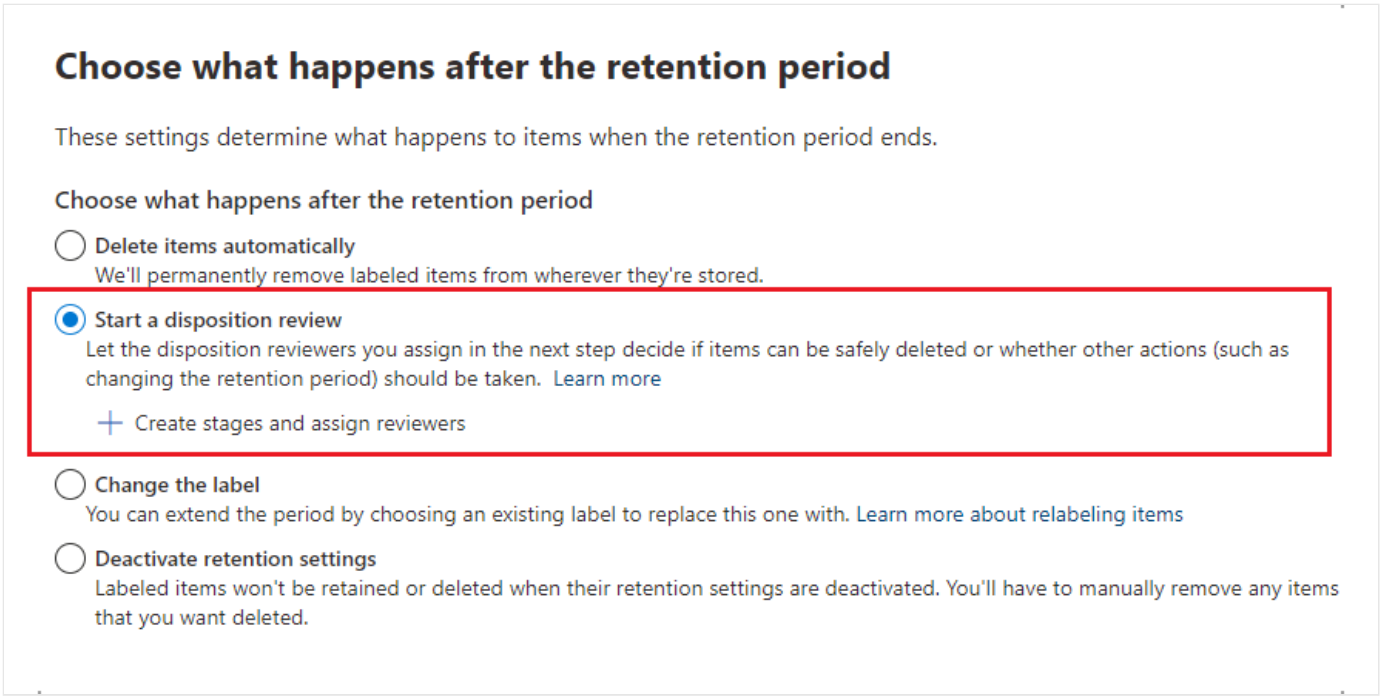
As with all retention label changes, allow up to 7 days if you turn on, turn off, or change the number of days for this option.

There's no new auditing event for auto-approval. Instead, use the details in the existing Approved disposal auditing event to identify whether the item was manually approved or automatically approved by using this option.

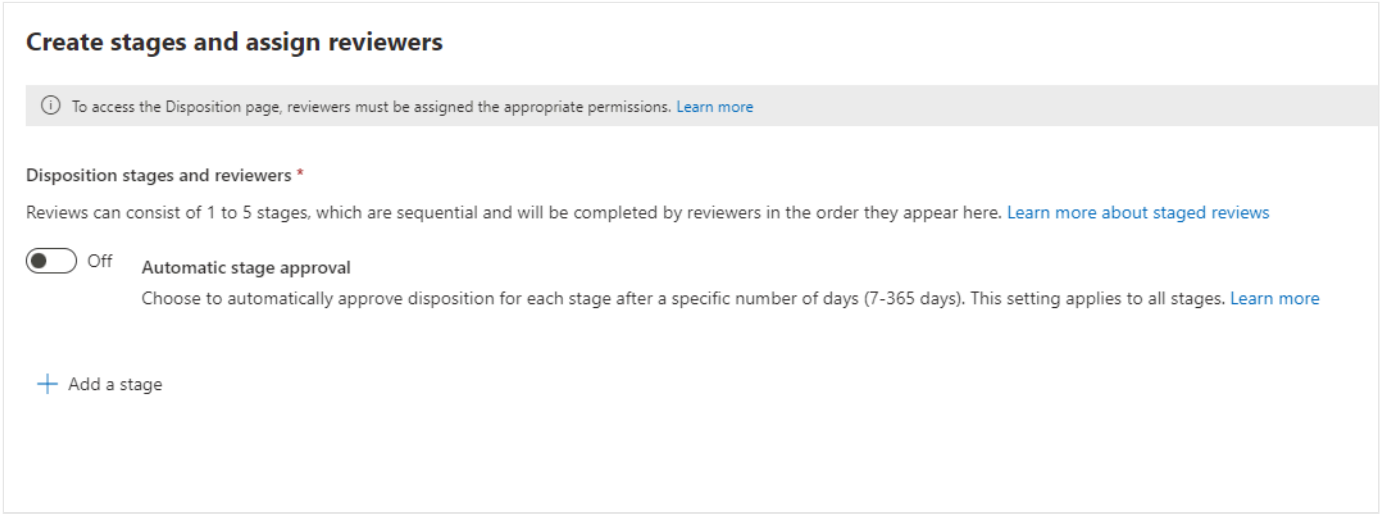
How to configure a retention label for disposition review

Triggering a disposition review at the end of the retention period is a configuration option that's available only with a retention label. Disposition review isn't available for a retention policy. For more information about these two retention solutions, see Learn about retention policies and retention labels.

From the **Choose what happens after the retention period** page for a retention label:



After you select the **Start a disposition review** option, select **+ Create stages and assign reviewers**. On the next page of the configuration, you'll specify how many consecutive stages of disposition you want and the disposition reviewers for each stage:



Optionally, select whether you want to use automatic-approval. If you use this option, specify the number of days reviewers have to take manual action before the item is automatically moved to the next disposition stage or automatically disposed.

Select **+ Add a stage**, and name your stage for identification purposes. Then specify the reviewers for that stage.

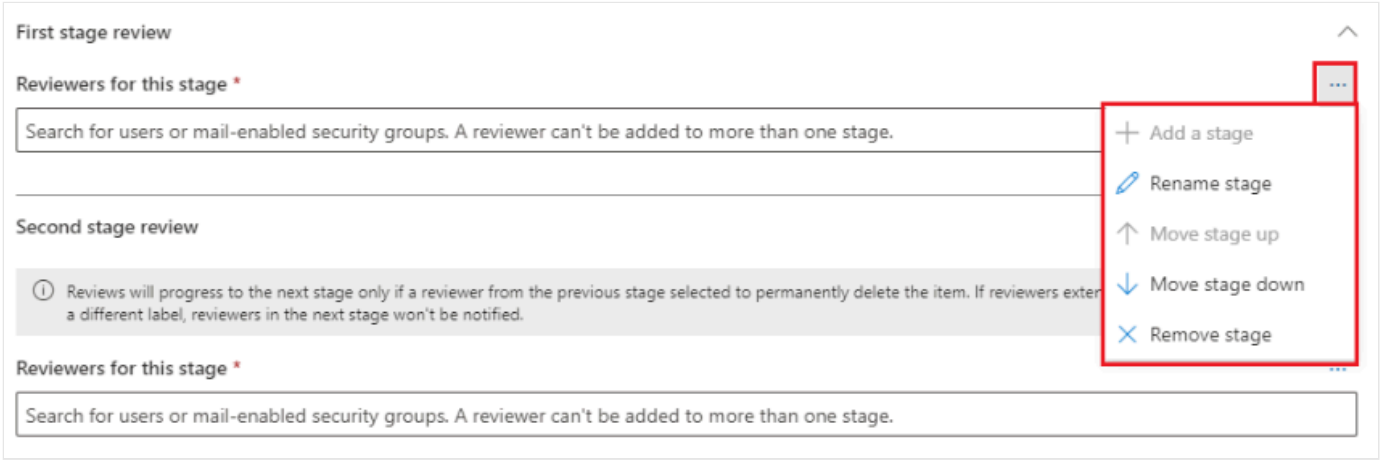
For the reviewers, specify up to 10 individual users or mail-enabled security groups. Microsoft 365 groups aren't supported for this option.

If you need more than one person to review an item at the end of its retention period, select **Add another stage** and repeat the configuration process for the number of stages that you need, with a maximum of five stages.

Within each individual stage of disposition, any of the users you specify for that stage are authorized to take the next action for the item at the end of its retention period. These users can also add other users to their disposition review stage.

> **Note**
>
> If you configured retention labels before multi-staged disposition review was available, you can upgrade your labels to support this feature: Edit the label and select **Edit stages and reviewers** on the **Choose what happens after the retention period** page.

During the configuration phase, for each stage specified, you can rename it, reorder it, or remove it by selecting **Edit stages and reviewers** that now displays for the **Start a disposition review** option. Then for each stage, you can select the Stage actions option (...):

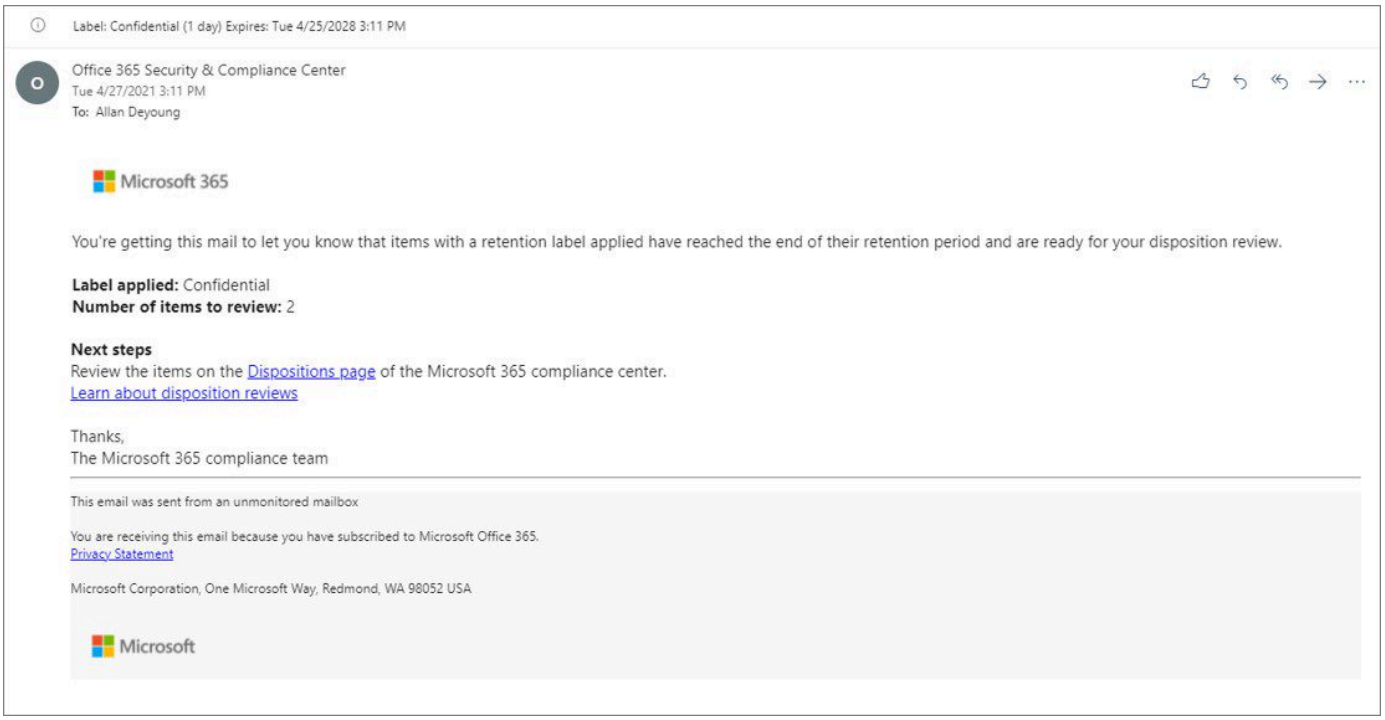


However, you can't reorder or remove a stage after you've created the retention label. You'll see only the **Add a stage** and **Rename a stage** options available. You can still edit the reviewers.

After you've specified your reviewers, remember to grant them the **Disposition Management** role permission. For more information, see the Permissions for disposition section on this page.

How to customize email messages for disposition review

Example default email notification sent to a reviewer:



You can customize the email messages that are sent to disposition reviewers for the initial notification and then reminders.

1. Sign in to the Microsoft Purview portal > **Settings** > **Solutions settings** > **Records Management** > **Disposition**.

2. From the **Email notifications for disposition reviews** section, select and specify whether you want to use just the default email message, or add your own text to the default message. Your custom text is added to the email instructions after the information about the retention label and before the next steps instructions.

   Text for all languages can be added, but formatting and images are unsupported. URLs and email addresses can be entered as text and depending on the email client, display as hyperlinks or unformatted text in the customized email.

   Example text to add:

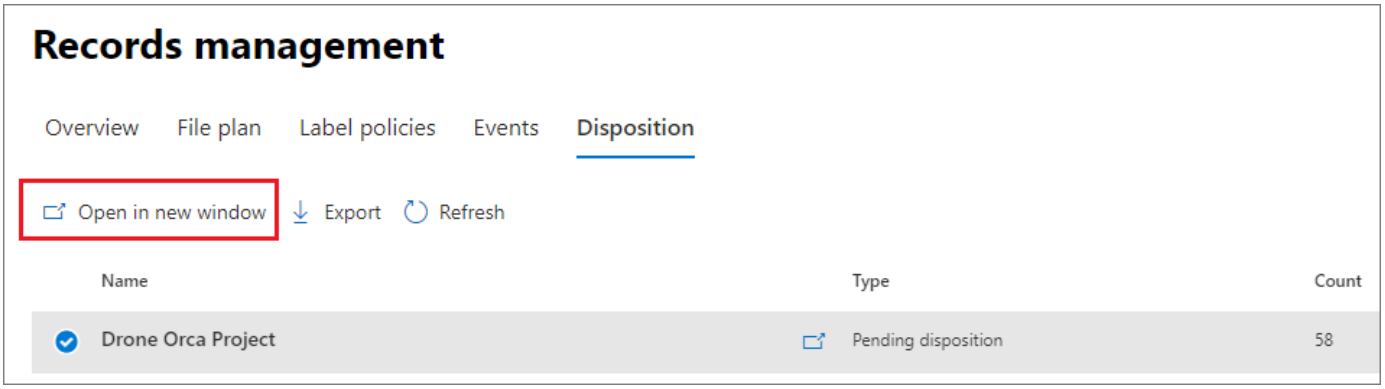
   Console

   ```
   If you need additional information, visit the helpdesk website
   (https://support.contoso.com) or send them an email (helpdesk@contoso.com).
   ```

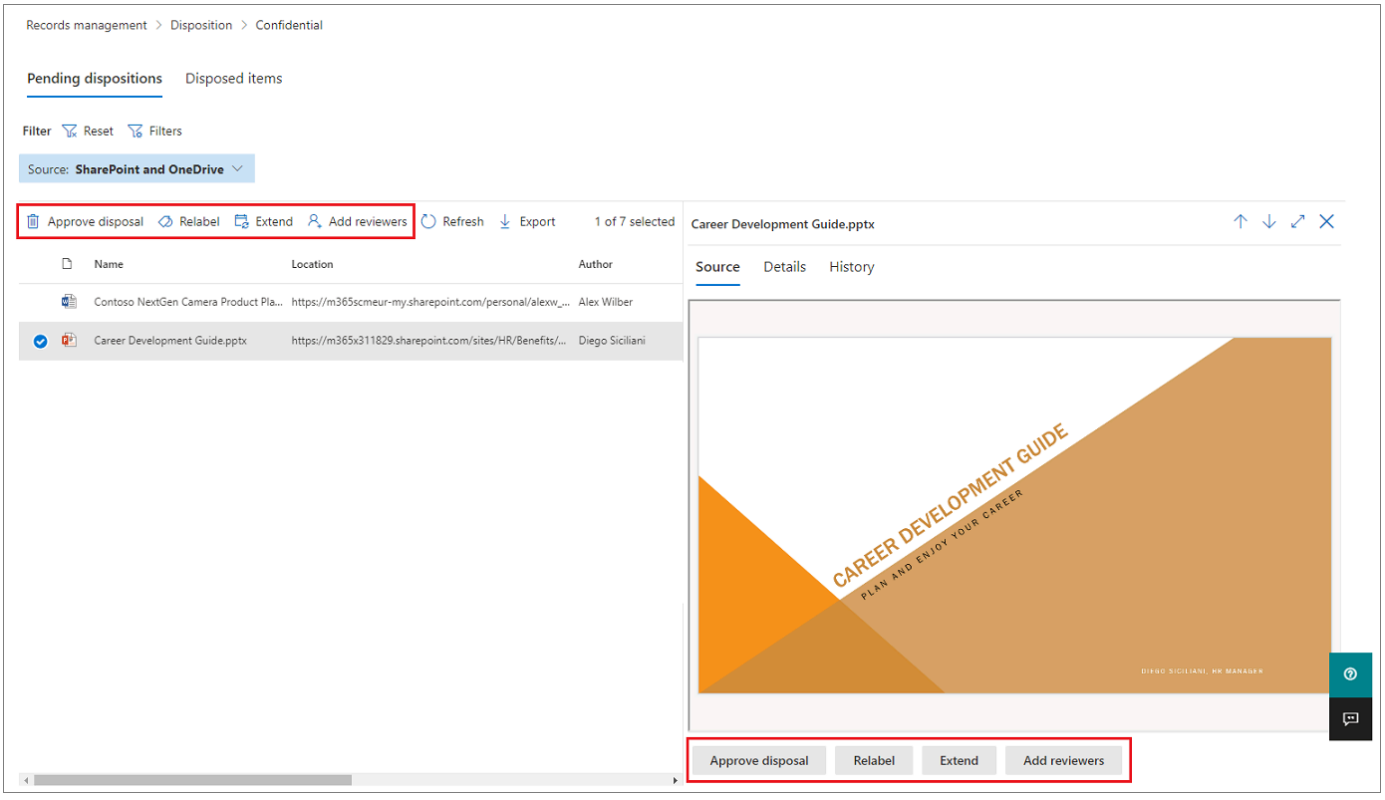
3. Select **Save** to save any changes.

Viewing and disposing of content

When a reviewer is notified by email that content is ready to review, they can click a link in the email that takes them directly to the **Disposition** page from **Records management** in the Microsoft Purview portal. There, the reviewers can see how many items for each retention label are waiting disposition with the **Type** displaying **Pending disposition**. They then select a retention label, and **Open in new window** to see all content with that label:



From the **Pending dispositions** page, they see all pending dispositions for that label. When one or more items are selected, they can use the mini-preview pane and the **Source**, **Details**, and **History** tab to inspect the content before taking action on it:



If you use the horizontal scroll bar, or close the min-review pane, you see more columns that include the expiry date and the name of the disposition review stage.

As you can see from the example shown, the actions supported are:

- **Approve disposal**:

    - When this action is selected for an interim stage of disposition review (you've configured multiple stages): The item moves to the next disposition stage.
    - When this action is selected for the final stage of disposition review, or there's only one stage of disposition: The item is marked as eligible for permanent deletion, which happens within 15 days.

- **Relabel**:
    - When this action is selected, the item exits the disposition review process for the original label. The item is then subject to the retention settings of the newly selected retention label.

- **Extend**:
    - When this action is selected, disposition review is effectively suspended until the end of the extended period and then disposition review is triggered again from the first stage.

- **Add reviewers**:
    - When this action is selected, the user is prompted to specify and add other users for review.

    > ⓘ **Note**
    >
    > This action doesn't automatically grant the required permissions to the users who are added. If they don't have these permissions, they can't participate in the disposition review.

Each action taken has a corresponding audit event in the Disposition review activities auditing activities group.

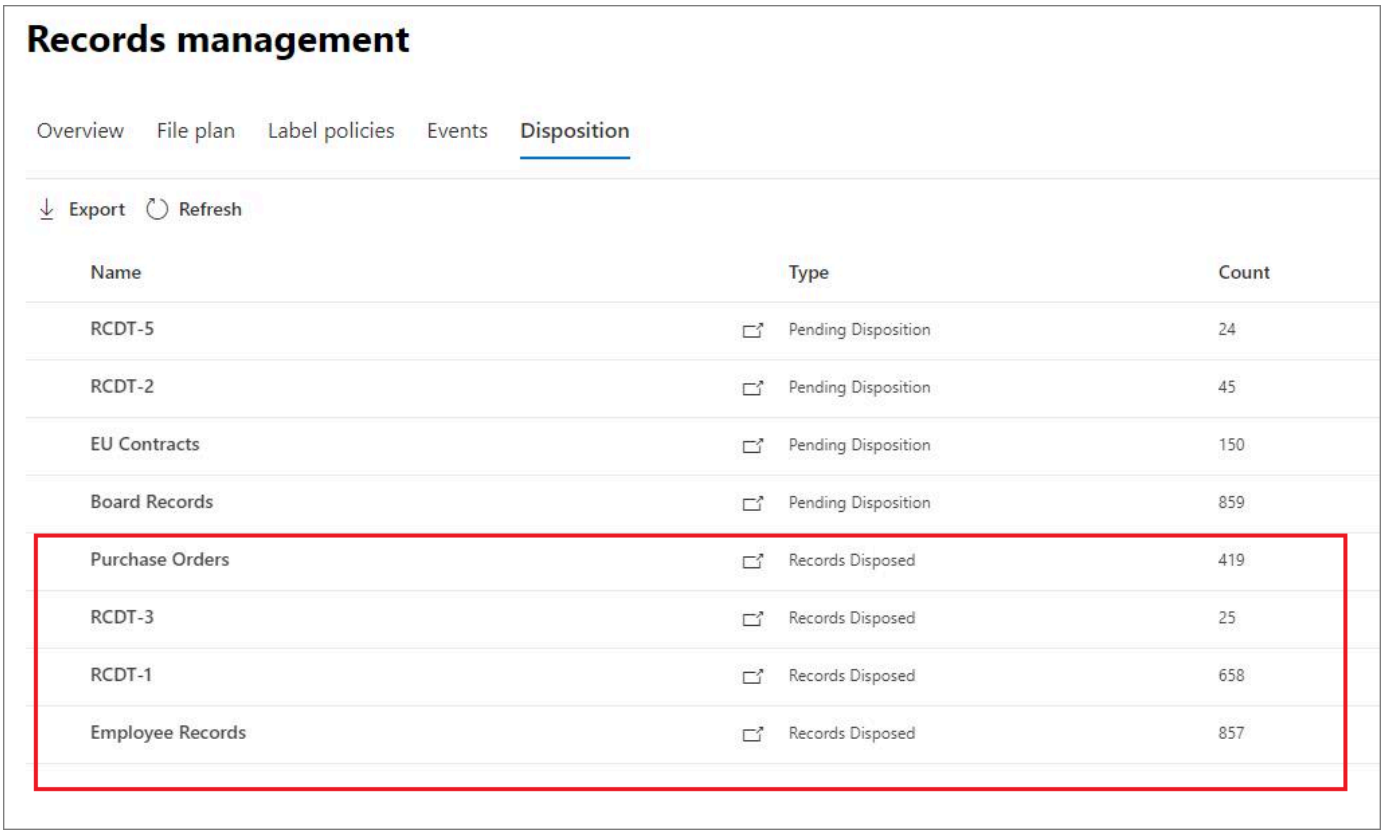
During the disposition review process, unless you're using the optional setting of an auto-approval timeout period, the content never moves from its original location, and it's not marked for permanent deletion until this action is selected by a reviewer for the final or only disposition stage.

Disposition of records

From the **Records management** main page > **Disposition** tab, you can identify:

- Items deleted as a result of a disposition review.
- Items marked as a record or regulatory record but not marked for disposition review and automatically deleted at the end of their retention period.

These items indicate **Records Disposed** in the **Type** column. For example:



> **Note**
>
> This functionality uses information from the unified audit log and therefore requires auditing to be enabled and searchable so the corresponding events are captured.

For auditing of deleted items that were marked as records or regulatory records, search for **Deleted file marked as a record** in the **File and page activities** category. This audit event is applicable to documents and emails.

Filter and export the views

When you select a retention label from the **Disposition** page, the **Pending disposition** tab (if applicable) and the **Disposed items** tab let you filter the views to help you more easily find items.

For pending dispositions, the time range is based on the expiration date. For disposed items, the time range is based on the deletion date.

You can export information about the items in either view as a .csv file that you can then sort and manage using Excel.

Note: The author created this article with assistance from AI. Learn more

Last updated on 07/22/2026

Learn about eDiscovery

Electronic discovery, or eDiscovery, is the process of identifying and delivering electronically stored information (ESI) that you can use as evidence in investigations and legal cases. Use Microsoft Purview eDiscovery to identify, review, and manage content in Microsoft 365 services to support your investigations. Supported Microsoft 365 services include:

- Exchange Online
- Microsoft Teams
- Microsoft 365 Groups
- OneDrive
- SharePoint
- Viva Engage

You can search mailboxes and sites in the same eDiscovery search, and then export the search results. Use eDiscovery cases to identify, hold, and export content found in mailboxes and sites. If your organization has an Office 365 E5 or Microsoft 365 E5 subscription (or related E5 add-on subscriptions), you can further manage cases and analyze content by using premium eDiscovery features.

eDiscovery in the Microsoft Purview portal

Depending on the licensing and subscriptions for your organization, you have access to specific eDiscovery or premium eDiscovery features in the Microsoft Purview portal. All Content Search features are now included within the search experience in eDiscovery, or you can use the Content Search case in eDiscovery to manage all new and existing content searches.

To get the latest information on planned eDiscovery features and capabilities and estimated release dates, see the Microsoft 365 Roadmap.

Notable changes in eDiscovery

If you're familiar with previous versions of eDiscovery, you might notice several differences when using eDiscovery in the Microsoft Purview portal:

- **Advanced indexing**: In previous versions of eDiscovery, when you added a custodian or noncustodial data source to a case, you needed to reindex any content that's partially indexed or has indexing errors. Reindexing determines if the contents are relevant to

defined search conditions. This reindexing process is called *advanced indexing*. As you add more partially and unindexed items to data sources (user's mailbox, OneDrive account, and so on), you need to separately update the index for specific custodian or noncustodial data sources.

In eDiscovery, advanced indexing runs automatically during each search that's scoped for statistics results and when you add results to a review set or export search results, depending on the indexing options you choose for the process. You no longer need to separately reindex data sources before the search process. This just in time indexing process helps avoid issues with stale indices that might result in indexing and search running sequentially and separately in the classic experience. Running (or rerunning) a search automatically updates all indexes.

- **Collections**: In previous versions of eDiscovery, *collections* provided managers with estimates of the content that might be relevant to cases. These estimates allowed managers to make quick, informed decisions about the size and scope of content relevant to cases. Once added to a review set, the collection is immutable.

  In eDiscovery, *Statistics* in searches replace *collections*. Statistics results in searches now allow managers to review important insights about the items included in the results and the relevance to the case. Searches aren't immutable in eDiscovery, even after the results are added to a review set. You can update searches at any time. Adding only a sample of the collection into review set and deleting a search is removed in eDiscovery.

- **Content Search**: Content search in the retired compliance portal was a separate solution from eDiscovery used for basic searches for content. Results from content search were estimated numbers of locations and search results that you could preview or export to a local computer.

  In the Microsoft Purview portal, all Content Search functionality is now included in a system generated eDiscovery case by default for all members of the *eDiscovery manager* and *Administrator* role groups. If you need to limit access to content searches, use *Case settings* to remove or add members to the case to manage access to these searches. You can also select **Content Search** in eDiscovery to create a Content Search case in your organization that contains all new and existing content searches.

  The Content Search case has the same capability as other user-created cases. You can create holds, review sets, and more in the content search case, depending on your subscription.

For more information about the changes to Content Search in eDiscovery, check out the following video:

https://learn-video.azurefd.net/vod/player?id=453b3db7-fd52-4c09-9e9a-03a3e1c4e213&locale=en-us&embedUrl=%2Fpurview%2Fedisc

- **Custodians**: In previous versions of eDiscovery, *custodians* (users) were the primary component of the eDiscovery workflow. Custodians were potential persons of interest in an investigation that you added to cases.

  In eDiscovery, cases are the primary component of the eDiscovery workflow. You still add people, groups, and data sources to cases, but the case is the central organizing unit.

- **Export updates**: The new export flow in eDiscovery supports a unified export structure across premium and nonpremium feature exports, faster export performance, detailed reporting, and flexible export options.

- **Jobs**: In previous versions of eDiscovery, tasks, activities, and reports associated with workflow components were called *jobs*. These events and reports are now referred to as *processes* in eDiscovery.

Features and capabilities

The following table compares key eDiscovery capabilities and features:

Expand table

Capability	eDiscovery feature support	Premium eDiscovery feature support
Search for content	✔	✔
Keyword queries and search conditions	✔	✔
Search statistics and samples	✔	✔
Export search results	✔	✔
Role-based permissions	✔	✔
Case management	✔	✔

Capability	eDiscovery feature support	Premium eDiscovery feature support
Place content locations on hold	✔	✔
Search and delete data	✔	✔
Advanced indexing		✔
Review sets		✔
Import external data to review sets		✔
Support for cloud attachments and SharePoint versions		✔
Optical character recognition		✔
Conversation threading		✔
Decryption		✔
Review set filtering		✔
Review set KQL queries (preview)		✔
Query Report (preview)		✔
Tagging		✔
Analytics		✔
Computed document metadata		✔
Guest user access (preview)		✔
Security Copilot		✔
Transparency of long-running processes		✔
Full reporting for all processes		✔
Enhanced data source mapping		✔

Capability	eDiscovery feature support	Premium eDiscovery feature support
Graph API delegated authentication	✔	✔
Graph API app-only authentication		✔

Here's a description of each eDiscovery capability.

- **Search for content**: Search for content that's stored in Exchange mailboxes, OneDrive accounts, SharePoint sites, Microsoft Teams, Microsoft 365 Groups, and Viva Engage Teams. Searches include content generated by other Microsoft 365 apps that store data in mailboxes and sites.
- **Keyword queries and search conditions**: Create Keyword Query Language (KeyQL) search queries to search for content keywords that match query criteria. You can also include conditions to narrow the scope of your search.
- **Search statistics and samples**: After you run a search, you can view statistics of the estimated search results, such as the number and total size of items matching your search criteria, the content locations that contain the most items, and the number of items returned by the search query. You can also view a representative sample of the items included in the search results.
- **Export search results**: Export search results to a local computer in your organization. When you export search results, items are copied from their original content location and packaged. Then you can download those items in the export package to a local computer.
- **Role-based permissions**: Use role-based access control (RBAC) permissions to control what eDiscovery-related tasks that different users can perform. You can use a built-in eDiscovery-related role group or create custom role groups that assign specific eDiscovery permissions.
- **Case management**: An eDiscovery case contains all searches, holds, and review sets related to a specific investigation. You can also assign members to a case to control who can access the case and view the contents of the case.
- **Place content locations on hold**: Preserve content relevant to your investigation by placing a hold on the content locations in a case. Holds let you secure electronically stored information from inadvertent (or intentional) deletion during your investigation.
- **Search and delete data**: Use eDiscovery to search for and delete email messages, Microsoft Teams chat messages, and Copilot and AI application data across your organization. This capability helps you find and remove content that's potentially harmful or high-risk.
- **Advanced indexing**: When a search, review set, or export process runs, the associated content locations where items are partially indexed are reindexed in a process called

Advanced indexing. Advanced indexing ensures any content deemed as partially indexed is reprocessed to make it fully searchable when you collect data for an investigation.

- **Review sets**: Add relevant data to a review set. A review set is a secure, Microsoft-provided Azure Storage location in the Microsoft cloud. When you add data to a review set, the collected items are copied from their original content location to the review set. Review sets provide a static, known set of content that you can search, filter, tag, analyze, and predict relevancy using predictive coding models. You can also track and report on what content gets added to the review set.
- **Import external data to review sets**: Not all documents that you need to analyze in eDiscovery are in Microsoft 365. By using the external data import feature, you can upload documents that aren't in Microsoft 365 to a review set.
- **Support for cloud attachments and SharePoint versions**: When you add content to a review set, you can include cloud attachments or linked files. The target file of a cloud attachment or linked file is added to the review set. You also can add all versions of a SharePoint document to a review set.
- **Optical character recognition (OCR)**: When content is added to a review set, OCR functionality extracts text from images, and includes the image text with the content that's added to a review set. This feature lets you search for image text when you query the content in the review set.
- **Conversation threading**: When chat messages from Teams and Viva Engage conversations are added to a review set, you can collect the entire conversation thread. The entire chat conversation that contains items that match the search criteria is added to the review set. This feature lets you review chat items in the context of the back-and-forth conversation.
- **Decryption**: eDiscovery supports decryption of content protected with Microsoft encryption technologies. Email messages and attachments encrypted with Microsoft Purview Message Encryption, and documents protected with sensitivity labels or Azure Rights Management, are automatically decrypted when added to search results or review sets.
- **Review set filtering**: After content is added to a review set, you can apply filters to display only the set of items that match your filtering criteria. Then you can save the filter sets as a query, which lets you quickly reapply the saved filters. Review set filtering and saved queries help you quickly select content items that are most relevant to your investigation.
- **Review set KQL queries (preview)**: Use the Advanced review set explorer (preview) to run Kusto Query Language (KQL) queries directly against review set data. Build complex queries with filtering, aggregation, pattern matching, regex, and data visualization to analyze and identify key information in the review set.
- **Query Report (preview)**: Generate and download a consolidated Query Report on multiple queries for a review set. This report lets you quickly see the total count and volume of

filtered items on a particular keyword search or multiple compound KeyQL queries.

- **Tagging**: Tags also help you omit nonrelevant content and identify the most relevant content. When experts, attorneys, or other users review content in a review set, their opinions related to the content can be captured by using tags. For example, if the intent is to exclude unnecessary content, a user can tag documents with a tag such as "nonresponsive". After content is reviewed and tagged, a review set query can be created to exclude any content tagged as "nonresponsive". This process eliminates the nonresponsive content from subsequent steps in the eDiscovery workflow.
- **Analytics**: eDiscovery allows you to analyze review set documents to help you organize the documents in a coherent manner and reduce the volume of documents to be reviewed. *Near duplicate detection* groups textually similar documents together to help you make your review process more efficient. *Email threading* identifies specific email messages that give a complete context of the conversation in an email thread. *Themes* functionality attempts to analyze themes in review set documents and assign a theme to documents so that you can review documents with related theme. These analytics capabilities help make your review process more efficient so that reviewers can review a fraction of collected documents.
- **Computed document metadata**: Many of the eDiscovery premium features, such as conversation threading and analytics, add metadata properties to review set documents. This metadata contains information related to the function performed by a specific feature. When reviewing documents, you can filter on metadata properties to display documents that match your filter criteria. This metadata can be imported into third-party review applications after review set documents are exported.
- **Guest user access (preview)**: Invite guest users to access eDiscovery cases for collaboration with external reviewers. Guest users are assigned the *Reviewer* role group permission and can view and work with case content shared with them.
- **Security Copilot**: Use Microsoft Security Copilot features to draft KeyQL search queries by using natural language and to generate contextual summaries of items in a review set.
- **Transparency of long-running processes**: Processes in eDiscovery typically involve long-running activities that user actions trigger, such as adding custodians to a case, adding content to a review set, running analytics, and creating search queries. You can track the status of these processes and get support information if you need to escalate issues to Microsoft Support.
- **Full reporting for all processes**: Use the Process report to view and manage processes in cases, searches, review sets, and holds.
- **Enhanced data source mapping**: Data sources streamline the process of identifying and managing data. You can search locations based on users, search one site for groups, and

map other locations with groups. You can also explore frequent collaborators as part of data sources.

- **Graph API delegated authentication**: Graph API calls are made on behalf of a signed-in user. The application acquires a token that combines the user's identity with the app's permissions (scopes), so Graph enforces both the user's privileges and the app's consented permissions.
- **Graph API app-only authentication**: Graph API calls are made by the application itself, without a user context. The app authenticates using its own credentials (client secret or certificate) and acquires a token tied solely to its consented application permissions.

Integration with other solutions

Insider risk management

In Microsoft Purview Insider Risk Management, you can quickly escalate cases to new cases in eDiscovery when you need additional legal review for potentially risky user activity. The tight integration between these solutions helps your risk and legal teams work more efficiently and provides a complete end-to-end view of user activities under review.

For more information, see get started with Insider Risk Management and escalate an Insider Risk Management case to an eDiscovery (Premium) case.

Microsoft Security Copilot

In eDiscovery, you can use Microsoft Security Copilot features to draft KeyQL search queries by using natural language. Copilot translates natural language to KeyQL without requiring you to learn how to construct a KeyQL query, know operators, or know supported search metadata fields.

Copilot can also provide a contextual summary of most items in a review set. The summary is in the context of text included in a selected item. This summary saves time for reviewers by quickly identifying information helpful when tagging or exporting items. Security Copilot summarizes the entire item, including any documents, meetings transcripts, or attachments. Most of the common document file types are supported.

For more information about using Security Copilot with review sets, see Summarize an item by using Security Copilot.

Ready to get started?

- Learn about the eDiscovery workflow
- Get started with eDiscovery

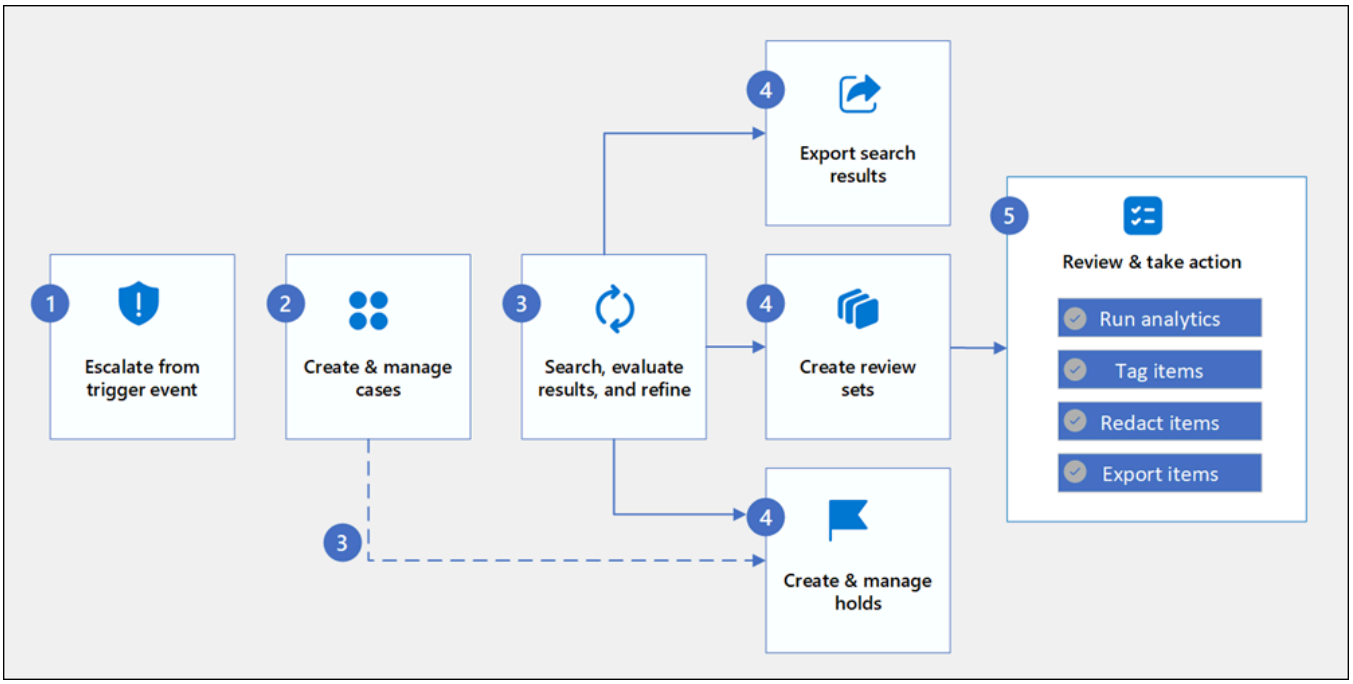
ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/29/2026

Learn about the eDiscovery workflow

The eDiscovery workflow helps you quickly identify, investigate, and take action on electronically stored information (ESI) in your organization. Identifying and taking action on ESI items by using eDiscovery uses the following improved workflow:



Step 1: Escalate from trigger event

Trigger events are activities that escalate in your organization and start the creation of a new case in eDiscovery. These events can be requests from internal or external partners, integrated events associated with alerts in other Microsoft Purview solutions (for example, Insider Risk Management cases), or any other activity that might benefit from the search, investigation, and mitigation actions included with eDiscovery.

Step 2: Create and manage cases

A case in eDiscovery contains all searches, holds, and review sets related to a specific investigation. Cases might include responding to regulatory, investigation, and litigation requests. You can also assign members to a case to control who can access the case and view the contents of the case. eDiscovery also supports new case creation integration with Microsoft Purview Insider Risk Management cases.

Step 3: Search, evaluate results, and refine

After you create a case, use the built-in search tools in eDiscovery to search the content locations in your organization. You can create and run different searches that are associated with the case. Use conditions (such as keywords) to build multiple search queries that return search results with the data that's most likely relevant to the case. You can also:

- View search statistics that might help you refine a search query to narrow the results.
- Preview the search results to quickly verify whether the relevant data is being found.
- Revise queries and rerun searches.

Step 4a: Actions from search results

- **Export search results**: After a search in an eDiscovery case successfully completes, you can export the search results. When you export search results, mailbox items are downloaded in PST files or as individual messages. When you export content from SharePoint and OneDrive sites, copies of native Office documents and other documents are exported. Search exports expire 14 days after creation, so download the export packages before they're automatically deleted.
- **Create review sets**: A review set is a secure, Microsoft-provided Azure Storage location in the Microsoft cloud. When you add data to a review set, the collected items are copied from their original content location to the review set. You can add all search results or configure sampling to add a representative subset. Review sets provide a static, known set of content that you can search, filter, tag, and analyze. You can also track and report on what content gets added to the review set.

Step 4b: Create holds

To preserve and protect data that's relevant to an investigation, place an eDiscovery hold on the data sources associated with a case. After creating a case, you can immediately place a hold on the content locations of the people of interest in your investigation. You can also create query-based holds if needed. Content locations include Exchange mailboxes, SharePoint sites, OneDrive accounts, and mailboxes and sites associated with Microsoft Teams and Microsoft 365 Groups. While placing a hold is optional, creating a hold preserves content that might be relevant to the case during the investigation.

When you create a hold, you can preserve all content in specific content locations or you can create a query-based hold to preserve only the content that matches a hold query. In addition to preserving content, another good reason to create holds is to quickly search the content locations on hold (instead of having to select each location to search) when you create and run

searches in the next step. After you complete your investigation, you can release any hold that you created. For more information, see Manage holds in eDiscovery.

Step 5: Review and take action from review sets

- **Search for content**: In most cases, it's useful to dig deeper into the content in a review set and organize it to facilitate a more efficient review. Using filters and queries in a review set helps you focus on a subset of documents that meet the criteria of your review.
- **Run KQL queries (preview)**: Use the Advanced review set explorer (preview) to run Kusto Query Language (KQL) queries directly against review set data for advanced filtering, aggregation, pattern matching, and data visualization.
- **Run analytics**: eDiscovery provides an integrated analytics tool that helps you further cull data from the review set that you determine isn't relevant to the investigation. In addition to reducing the volume of relevant data, eDiscovery also helps you save legal review costs by letting you organize content to make the review process easier and more efficient. For more information, see Analyze data in a review set in eDiscovery.
- **Tag items**: Organizing content in a review set is important to complete various workflows in the eDiscovery process. This organization often includes identifying relevant content, culling unnecessary content, and identifying content that needs review by an expert or attorney. When experts, attorneys, or other users review content in a review set, they can capture their opinions related to the content by using tags. Tags provide structure and organization to items included in an investigation. For more information, see Tag documents in a review set in eDiscovery.
- **Create a Query report (preview)**: Generate and download a consolidated report on multiple queries for a review set. This report lets you quickly see the total count and volume of filtered items on a particular keyword search or multiple compound KeyQL queries.
- **Add items from the review set to another review set**: In some cases, it might be necessary to select documents from one review set and work with them individually in another review set.
- **Export items**: After you search for and find data that's relevant to your investigation, you can export it out of your Microsoft 365 organization for review by people outside of the investigation team. In addition to the exported data files, the export package contains an export report, a summary report, and an error report. For more information, see Export documents from a review set in eDiscovery.

Throughout the workflow, eDiscovery administrators can use Process reports to track activities that count towards concurrency and daily limits, and Process managers to view process

information scoped to cases, searches, review sets, and holds.

Ready to get started?

- Learn more about eDiscovery features and components
- Get started with eDiscovery

ⓘ **Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 04/02/2026

Learn about eDiscovery features and components

This article provides a summary of the key features and components in Microsoft Purview eDiscovery. Use this article to learn about what each feature and component does and to find links to more detailed guidance. For information about the overall eDiscovery workflow, see Learn about the eDiscovery workflow.

Auditing

eDiscovery activities performed by admins, compliance officers, and other users are logged in the audit log. You can search for and review these activities to track actions taken across cases, searches, holds, and review sets. For more information, see Search for eDiscovery activities in the audit log. For a complete list of logged eDiscovery activities, see eDiscovery activities.

Cases

A case contains all searches, holds, and review sets related to a specific investigation. This investigation might include responding to regulatory and litigation requests. You can also assign members to a case to control who can access the case and view the contents of the case. eDiscovery also supports new case creation integration with Microsoft Purview Insider Risk Management cases.

Two general types of cases exist in eDiscovery:

- Cases you create in eDiscovery by using **Create case** on the **Cases** dashboard
- A single Content Search case you create in eDiscovery by using **Create case** on the **Content Search** dashboard.

Content Search case

All content searches are contained in a single eDiscovery case named *Content search*. The **Content Search** option in the eDiscovery (classic) experience in the Microsoft Purview portal is retired.

The new **Content Search** option in the new eDiscovery experience displays a set of all existing content searches in your organization and contains any new content searches that you create in

the Microsoft Purview portal.

Cloud attachments

Cloud attachments are links to documents and files shared in Microsoft Teams messages, Outlook emails, and other Microsoft 365 services. Instead of attaching a traditional copy of a file, users share a link to the file stored in SharePoint or OneDrive. eDiscovery can collect these cloud attachments and include them in search results and review sets, ensuring that shared content is preserved and available for review.

Data sources

In eDiscovery, the concept of *data sources* streamlines the process of identifying and managing data across Microsoft 365 platforms. eDiscovery users select a user, group, or bulk add list of SMTP addresses or URLs, which creates data sources and eDiscovery automatically identifies and organizes relevant data stored across platforms. The data source gathers locations related to the user or group (mailboxes, OneDrive sites, SharePoint sites) and adds the locations in the data source hierarchy. eDiscovery users refine the scope by selecting or excluding specific locations as needed.

SharePoint Embedded container support in eDiscovery (preview) lets data sources include supported user-owned and tenant-owned containers. You can discover relevant containers for a selected user, filter the containers by application, add a known container URL, or bulk import supported container URLs.

Group-owned and app-owned SharePoint Embedded containers aren't currently supported. For details, including the compliance boundary limitation, see Data sources in eDiscovery.

Decryption

eDiscovery supports decryption of content protected with Microsoft encryption technologies. When you add email messages and attachments encrypted with Microsoft Purview Message Encryption, and documents protected with sensitivity labels or Azure Rights Management, to search results or review sets, the system automatically decrypts them. This capability ensures that protected content is available for review and analysis without requiring manual decryption steps.

Exports and downloads

After a search associated with an eDiscovery case runs successfully, you can export the search results. When you export search results, mailbox items are downloaded in PST files or as individual messages. When you export content from SharePoint and OneDrive sites, you export copies of native Office documents and other documents. Search export packages expire 14 days after creation and must be downloaded before they're automatically deleted.

If you add search results to a review set from a case, you can also export review set content to a download package. This package is configurable and includes options to export selected documents only, all filtered documents, or all documents in the review set.

Holds and hold policies

Use an eDiscovery case to create hold policies that preserve content relevant to the investigation by applying an eDiscovery hold. Place a hold on the Exchange mailboxes and OneDrive accounts of people you're investigating in the case. You can also place a hold on the mailboxes and sites associated with Microsoft Teams, Microsoft 365 groups, and Viva Engage Groups. When you place content locations on hold, you preserve content until you remove the content location from the hold or delete the hold. Use the hold management dashboard to view hold status, modify holds, and track hold policies across cases. You can also use the hold report to view information about holds associated with all eDiscovery cases in your organization.

If needed, you can also place a mailbox on Litigation Hold to preserve all mailbox content, including deleted items and original versions of modified items. When you place a mailbox on Litigation Hold, the user's archive mailbox (if it's enabled) is also placed on hold.

Permissions and licensing

To use any of the eDiscovery-related features in the Microsoft Purview portal, assign the appropriate permissions. The easiest way to assign roles is to add the person to the appropriate role group on the **Role groups** page in the Microsoft Purview portal.

Both admins and users working with eDiscovery cases require a Microsoft 365 Enterprise E3 or E5 license. If an admin account doesn't also have a Microsoft SharePoint E3 license, some SharePoint capabilities might be limited. License both admins and users appropriately. Some eDiscovery capabilities also require pay-as-you-go billing configuration. Additionally, if the admin account has any conditional access Entra policy some operations might be restricted, see Troubleshoot sign-in problems with Conditional Access to troubleshoot policy issues.

> **Tip**

You can view your own permissions on the eDiscovery overview page in the Microsoft Purview portal. You must have at least one role assigned for your permissions to be displayed.

Processes

eDiscovery includes a Process report that lists all activities that count towards case concurrency and daily limits in eDiscovery for a defined time period. Processes in eDiscovery (preview) are activities associated with specific tasks that support cases, searches, and review sets. User actions trigger processes when using these components.

eDiscovery administrators and *eDiscovery Managers* (preview) can access this report. Process managers help you view information that is automatically scoped to cases, searches, review sets, and holds.

Review sets

A review set is a secure, Microsoft-provided Azure Storage location in the Microsoft cloud. When you add data to a review set, the collected items are copied from their original content location to the review set. Review sets provide a static, known set of content that you can search, filter, tag, analyze, and predict relevancy using predictive coding models. You can also track and report on what content gets added to the review set.

Key review set capabilities include:

- Search and filter content using the query condition builder to focus on specific subsets of documents.
- Use the Advanced review set explorer (preview) to run Kusto Query Language (KQL) queries directly against review set data for advanced filtering, aggregation, and visualization.
- Group and view items using different grouping options and specialized viewers to examine item details.
- Tag items to organize content, identify relevant documents, and track review progress using customizable tag structures.
- Run analytics to identify near-duplicate documents, email threads, and themes, helping reduce the volume of content to review.
- Export items from a review set to a download package for use outside of eDiscovery.

Searches

Use search to quickly find content relevant to a case. Searches find content across Exchange mailboxes, SharePoint sites and OneDrive locations, Microsoft Teams conversations, and Copilot and AI application data.

Create and run different searches that are associated with the case. Build search queries using the condition builder for a guided experience, or use Keyword Query Language (KQL) for advanced query syntax with autocompletion support.

You can also:

- View search statistics for insights on data volume, content locations, and query hit counts to help you refine your search.
- View search samples to examine a representative set of items returned by the search criteria.
- Preview the search results to quickly verify whether the relevant data is being found.
- Revise a query and rerun the search.
- Export the search results or add the search results to a review set.

Trigger events

Trigger events are activities that escalate in your organization and start the creation of a new case in eDiscovery. These events can be requests from internal or external partners, integrated events associated with alerts in other Microsoft Purview solutions (for example, Insider Risk Management cases), or any other activity that might benefit from the search, investigation, and mitigation actions included with eDiscovery. For more information, see Learn about the eDiscovery workflow.

Ready to get started?

- Get started with eDiscovery
- Learn about the eDiscovery workflow

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 09/19/2026

Get started with eDiscovery

Use eDiscovery to identify, preserve, collect, review, analyze, and export content that's responsive to your organization's internal and external investigations. Complete the following steps in the eDiscovery workflow to set up prerequisites and configure eDiscovery.

> **Tip**
>
> Do you prefer an interactive configuration guide experience? Check out the **Getting started with Microsoft Purview eDiscovery** guide.

For more information about how eDiscovery can help you with investigations in your organization, see Learn about eDiscovery.

Subscriptions and billing

Before getting started with eDiscovery, review the new billing model for how data storage charges are calculated and what is included with your enterprise subscription. Depending on the types of data included in your eDiscovery cases, different charges might apply.

For more information, see Billing in eDiscovery.

Both admins and users working with eDiscovery cases require a Microsoft 365 Enterprise E3 or E5 license. If an admin account doesn't also have a Microsoft SharePoint E3 license, some SharePoint capabilities might be limited. It is recommended that both admins and users are licensed appropriately. Additionally, if the admin account has any conditional access Entra policy some operations may be restricted, see Troubleshoot sign-in problems with Conditional Access to troubleshoot policy issues.

Step 1: Assign permissions

To use any of the eDiscovery-related tools in the Microsoft Purview portal, you must assign the appropriate permissions to users. The easiest way to assign roles is to add the user to the appropriate role group on the **Role groups** page in the Microsoft Purview portal.

For step-by-step guidance, see Assign permissions in eDiscovery.

> **Tip**
>
> Do you prefer an interactive configuration guide experience? Check out the **Configure permissions and global settings** guide.

Step 2: Verify that required eDiscovery apps are enabled, up to date, and configured correctly

eDiscovery requires the following Enterprise apps to be enabled in your Microsoft 365 or Microsoft Office 365 organization. If you don't enable these apps, you can't access all eDiscovery view, filter, and search features. These apps must be up to date and correctly configured. Make sure that *Enabled for users to sign-in* is selected.

Expand table

App	App ID
Built-in security add-on for on-premises mailboxes (formerly Exchange Online Protection)	00000007-0000-0ff1-ce00-000000000000
ComplianceWorkbenchApp	92876b03-76a3-4da8-ad6a-0511ffdf8647
MicrosoftPurviewEDiscovery	b26e684c-5068-4120-a679-64a5d2c909d9
Office365Zoom	0d38933a-0bbd-41ca-9ebd-28c4b5ba7cb7

To check your app status and configuration, complete the following steps:

1. Go to the Microsoft Azure portal.
2. Go to or search for *Enterprise Applications*.
3. Filter the results for the apps listed in the previous table to review the status and configuration for each app.
4. In each application overview blade, go to **Manage** > **Properties**.
5. Make sure that **Enabled for users to sign-in** property is set to *Yes*.

> **Note**

The *MicrosoftPurviewEDiscovery* app provides support for exporting data through scripts when premium eDiscovery features are enabled. For more information about exporting data through scripts, see Get caseOperation.

Step 3: Configure settings

Depending on the needs of your organization and individual cases, you might need to review and configure global eDiscovery settings:

- **General**: Turn on or off premium eDiscovery features for new cases.
- **Analytics**: Turn on Attorney-client privilege detection within your working sets. If you turn on attorney-client privilege detection, the attorney-client privilege model runs on your data and helps flag documents likely to be privileged. This flag is based on the content and by comparing participants against a user-provided attorney list. This feature doesn't replace the need for privilege review, it's meant to help you get started.
- **Guest users**: Manage guest access to cases in your organization. Invitations to guests are sent after your approval.
- **Tag templates**: Tags help organize content in a review set to complete various workflows. Tags can be reused across multiple review sets and cases with tag templates. You can group tags by sections and allow single or multiple choice tagging.

Step 4: Create a case

To start an investigation, you must create a case and configure case settings. The user who creates the case is automatically added as a member. Members of the case can access the case in the Microsoft Purview portal and perform eDiscovery tasks.

Step 5: Create a search

After creating a new case, you're automatically directed to the **Searches** tab in the case and you're ready to create a search for the case. Searches help you find the items you want to collect for the case.

Step 6: Review and refine search results

After the search query completes, you see the results displayed on the **Statistics** or **Sample** tabs (depending on the initial results type you chose) for the search. To view the results for the

result type not chosen initially, select the applicable tab and you can generate results for the view.

For example, if you choose **Statistics** as the search result type initially, but after the search is completed you'd like to view **Sample** results, you'd select the **Sample** tab and select **Generate sample results** and define sample view settings. Select **Run query** to generate the sample results.

For more information about working with search results, see Review, evaluate, and refine search results in eDiscovery.

Step 7: Take action on search results

After reviewing the items returned in the search query and refining your search as needed, you can take specific actions on selected items included in the search results.

Add search results to a review set

When you're satisfied with the results of a search and you're ready to review and analyze those search results, add them to a review set in the case. Review sets provide a static, known set of content that you can search, filter, tag, analyze, and take other actions on as part of your investigation. To add search result items to a review set, see Add search results to a review set in eDiscovery.

Export or download search results

Exporting and downloading search results helps internal and external partners stay current in your investigations. Download reports provide summary information for selected items in the search results in a .csv file. Exporting search results allows you to customize what items are included in the export, the types of items exported, the export format for item, and more. For more information about exporting search results, see Export search results in eDiscovery.

Download search results

To download the summary information for the search result items, complete the following steps:

1. On the **Statistics** or **Sample** tabs, select **Download reports**.
2. Save the .csv file to your local computer or network share.

Export search results

To export information for items in the search results, select **Export** and complete the details needed for the item export. For step-by-step guidance, see Export search results in eDiscovery.

Create holds

Create holds to preserve content that might be relevant to an eDiscovery case. You can create holds at any time and they're not dependent on searches. For step-by-step guidance, see Create holds in eDiscovery.

Last updated on 04/01/2026

Billing in eDiscovery

eDiscovery uses a billing model based on the combination of your enterprise subscription and pay-as-you-go billing for certain features. This new billing combination includes:

- Your Microsoft 365 subscription and any add-ons. Depending on the licensing and subscriptions for your organization, you have access to specific premium eDiscovery features in the Microsoft Purview portal. For more information about subscriptions and licensing, see the subscription requirements for eDiscovery.
- Pay-as-you-go billing for data storage for non-Microsoft 365 enterprise and cloud AI application data in eDiscovery cases.
- Pay-as-you-go billing for using Microsoft Graph APIs for tenants with E3 eDiscovery licenses.

Configure pay-as-you-go billing

Before you can use features associated with non-Microsoft 365 AI data in eDiscovery in your organization, or use Microsoft Graph APIs with an E3 eDiscovery license, you must configure an Azure subscription for pay-as-you-go billing. For step-by-step guidance, see Enable Microsoft Purview pay-as-you-go features.

Billing details for Microsoft Graph API usage for E3 eDiscovery licensed organizations

Microsoft Graph APIs support tenants with E3 eDiscovery licenses, enabling automation and integration of eDiscovery workflows. These APIs support case management, search, hold, and export operations, offering a secure and consistent experience aligned with the Microsoft Purview portal.

These APIs replicate the portal experience and include:

- **Case management:** Create, Get, List, Update, Close, Reopen, Delete
- **Search management:** Create, List, Get, Update, Delete, Estimate Statistics
- **Hold management:** List, Create, Get, Update, Delete
- **Data source management:** Add, List
- **Export:** Export results, Export report only

For more details on how to use the APIs, see Use the Microsoft Purview eDiscovery API v1.0

Billing calculation for API usage

To use any of the APIs, you need to activate the pay-as-you-go subscription.

Currently, usage of all APIs except Export APIs doesn't contribute towards billing. Usage of Export API is billed based on the volume of data exported. Each organization receives an included amount of free storage per month, with additional usage billed at a set price per GB. The monthly billing details are available in the billing section on the Azure portal.

For more information about API storage and pricing, see Microsoft Purview pricing.

For example, the following table outlines the Graph API billing for a sample month of eDiscovery activity and uses a hypothetical billing rate of $10 USD per GB:

Expand table

Day	Case	API used	Volume	Monthly cumulative volume	Charges	Cumulative charges
Day 1	Case 1	Create case, Create Search 1	N/A	0	0	0
Day 2	Case 1	Create Search 2, Create Hold, Add data sources	N/A	0	0	0
Day 3	Case 1	Export report	1 GB	1 GB	0	0
Day 6	Case 1	Export results	24 GB	25 GB	0	0
Day 10	Case 2	Create case, Create Search	0	25 GB	0	0
Day 11	Case 2	Export results and reports	20 GB	45 GB	0	0
Day 14	Case 1	Create Search 3	0	45 GB	0	0
Day 17	Case 1	Export results from Search 3	15 GB	60 GB	$100	$100
Day 21	Case 2	Create search and export results	12 GB	72 GB	$120	$220

Billing details for non-Microsoft 365 enterprise and cloud AI application data

Data types

Depending on the type of data included in eDiscovery cases in your organization, subscription or pay-as-you-go billing applies:

- **Microsoft 365 data**: Non-AI data from Microsoft 365 services and applications. This data type includes data from Microsoft services like Exchange, SharePoint, OneDrive, and others.

  This data storage is included in your enterprise subscription and isn't subject to pay-as-you-go billing.

- **Microsoft 365 AI data**: AI data from Microsoft 365 services and applications. This data type includes AI data from Microsoft services like Microsoft 365 Copilot and AI data from Microsoft Office applications like Microsoft Word, Microsoft Excel, Microsoft PowerPoint, and others.

  This data storage is included in your enterprise subscription and isn't subject to pay-as-you-go billing.

- **Non-Microsoft 365 AI data**: Data from other generative AI applications from Microsoft and other connected external AI applications. This data type includes Copilot in Microsoft Fabric, Microsoft Security Copilot, and any connected or cloud AI application.

  This data storage isn't included in your enterprise subscription and is subject to pay-as-you-go billing. Data for non-Microsoft 365 services and applications is identified by *ItemClass*.

  The following ItemClasses are applicable to pay-as-you-go billing:
  - IPM.SkypeTeams.Message.Copilot.Fabric.*
  - IPM.SkypeTeams.Message.Copilot.Security.SecurityCopilot
  - IPM.SkypeTeams.Message.Copilot.BAP*
  - IPM.Skypeteams.Message.Copilot.Sql.*
  - IPM.SkypeTeams.Message.ConnectedAIApp.*
  - IPM.SkypeTeams.Message.CloudAIApp.SaaS.*

Non-Microsoft 365 AI data scope

You can add and export non-Microsoft 365 AI data to an eDiscovery case in several ways. Your organization retains this data as long as a case is active or closed. Depending on the data type and the type of process taken in eDiscovery, the application of data storage costs varies.

The following eDiscovery processes use pay-as-you-go billing for non-Microsoft 365 AI data:

- Direct export of this data from cases by using the Microsoft Purview portal
- Adding this data to a review set
- Exporting this data by using Microsoft Graph API

> **Important**
>
> If you use the classic eDiscovery (Premium) experience in the Microsoft Purview portal, importing non-Microsoft 365 or non-Office 365 data to a review set is excluded from pay-as-you-go billing and isn't charged.

The following eDiscovery processes don't use pay-as-you-go billing for non-Microsoft 365 AI data and are included in your enterprise subscription:

- All search processes
- All hold processes
- All case management processes

> **Tip**
>
> A breakdown of the total size of Microsoft 365 data and non-Microsoft 365 AI data is included in the ***summary.csv report*** for exports from search and exports from review sets. This report includes details for metadata when exporting from a review set and can help you estimate data storage charges.

Billing calculation for non-Microsoft 365 AI data

Every 24 hours, the system calculates the total amount of data storage for non-Microsoft 365 AI data for all active and closed eDiscovery cases. Deleted cases don't incur data storage charges after they're deleted. The system multiplies the total amount of data by the daily storage rate to determine the total data storage cost for the day.

At the end of each month, the system adds together the daily charges for the total eDiscovery data storage costs. These costs are part of the all-up data meter storage costs in your Azure account. The system itemizes details for each eDiscovery process and the costs. Use the

Summary.csv report to review a breakdown of the volume of Microsoft 365 and non-Microsoft 365 AI data for processes in scope for pay-as-you-go billing.

For example, the following table outlines the non-Microsoft 365 AI data storage amounts for a sample month of eDiscovery activity:

Expand table

Day	Case	Action	Daily data volume	Running daily total
Day 1	Case 1	Direct export in active case	1 GB	1 GB
Day 2	All cases	No actions	0 GB	1 GB
Day 3	Case 2	Add to review set in active case	5 GB	6 GB
Day 4	All cases	No actions	0 GB	6 GB
Day 5	Case 3	Direct API export in active case	2 GB	8 GB
Day 6	All cases	No actions	0 GB	8 GB
Day 7	All cases	No actions	0 GB	8 GB
Day 8	All cases	No actions	0 GB	8 GB
Day 9	Case 1	Apply holds in active case (not billable)	2 GB	8 GB
Day 10	All cases	No actions	0 GB	8 GB
Day 11	Case 1	Closed case	0 GB	8 GB
Day 12	All cases	No actions	0 GB	8 GB
Day 13	Case 2	Add to review set in active case	2 GB	10 GB
Day 14	All cases	No actions	0 GB	10 GB
Day 15	All cases	No actions	0 GB	10 GB

Day	Case	Action	Daily data volume	Running daily total
Day 16	All cases	No actions	0 GB	10 GB
Day 17	Case 1	Reopened as active case and export	1 GB	11 GB
Day 18	All cases	No actions	0 GB	11 GB
Day 19	All cases	No actions	0 GB	11 GB
Day 20	All cases	No actions	0 GB	11 GB
Day 21	Case 1	Direct export in active case	1 GB	12 GB
Day 22	All cases	No actions	0 GB	12 GB
Day 23	All cases	No actions	0 GB	12 GB
Day 24	All cases	No actions	0 GB	12 GB
Day 25	Case 1	Deleted case (removes data storage)	3 GB	9 GB
Day 26	All cases	No actions	0 GB	9 GB
Day 27	Case 2	Export from review set in active case (not billable)	2 GB	9 GB
Day 28	All cases	No actions	0 GB	9 GB

In this example, the total amount of non-Microsoft 365 AI data storage that is subject to pay-as-you-go-billing rates for the month is 246 GB. Data associated with processes that aren't applicable to data storage billing and data that you remove from your organization don't appear on your monthly bill. You can view billing details in the Azure portal.

Excluding non-Microsoft 365 AI data

To exclude non-Microsoft 365 AI data from processes subject to pay-as-you-go billing, use search conditions and filters to exclude this data type.

KeyQL condition

Use the following KeyQL statement with the *Equals* operator to exclude non-Microsoft 365 AI data from a search query:

```
(NOT((ItemClass:IPM.SkypeTeams.Message.Copilot.Fabric.*) OR
(ItemClass:IPM.SkypeTeams.Message.Copilot.Security.SecurityCopilot) OR
(ItemClass:IPM.SkypeTeams.Message.Copilot.BAP*) OR
(ItemClass:IPM.SkypeTeams.Message.ConnectedAIApp.*) OR
(ItemClass:IPM.SkypeTeams.Message.CloudAIApp.SaaS.*) OR
(ItemClass:ipm.skypeteams.message.copilot.sql.*)))
```

Filter

Use the *Contains none of* operator for the **Non M365 AI activity** type filter to exclude non-Microsoft 365 AI data.

Last updated on 12/01/2025

Assign permissions in eDiscovery

To use eDiscovery capabilities and features in the Microsoft Purview portal, assign the appropriate permissions to users. The easiest way to assign roles is to add users to the appropriate role group on the **Role groups** page in the Microsoft Purview portal. This article describes the permissions required to perform eDiscovery tasks.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

eDiscovery roles and role groups

The primary eDiscovery-related role group in Microsoft Purview portal is called **eDiscovery Manager**. This role group contains two subgroups:

- **eDiscovery Manager**: An *eDiscovery Manager* can search content locations in the organization and perform various search-related actions such as preview and export search results in eDiscovery. Members can also create and manage cases, add, and remove users for a case, create case holds, run searches associated with a case, and access case data. eDiscovery Managers can only access and manage the cases they create. They can't access or manage cases created by other eDiscovery Managers.
  - You can add a mail-enabled security group as a member of the eDiscovery Managers subgroup in the *eDiscovery Manager* role group by using the **Add-RoleGroupMember** cmdlet in Security & Compliance PowerShell. For example, you can run the following command to add a mail-enabled security group to the *eDiscovery Manager* role group.

    PowerShell

    ```
    Add-RoleGroupMember "eDiscoveryManager" -Member <name of security group>
    ```

  - Exchange distribution groups and Microsoft 365 Groups aren't supported. You must use a mail-enabled security group, which you can create in Exchange Online PowerShell by running `New-DistributionGroup -Type Security`. You can also create a mail-enabled security group (and add members) in the Exchange admin center or in the Microsoft

365 admin center. It might take up to 60 minutes after you create it for a new mail-enabled security group to be available to add to the eDiscovery Managers role group.

  - You can't make a mail-enabled security group an eDiscovery Administrator by using the **Add-eDiscoveryCaseAdmin** cmdlet in Security & Compliance PowerShell. You can only add individual users as eDiscovery Administrators.

  - You also can't add a mail-enabled security group as a member of a case.

- **eDiscovery Administrator**: An *eDiscovery Administrator* is a member of the *eDiscovery Manager* role group, and can perform the same content search and case management-related tasks that an eDiscovery Manager can perform. Additionally, an eDiscovery Administrator can:
  - Access all cases that are listed on the **eDiscovery** area in the Microsoft Purview portal.
  - Configure eDiscovery solution settings.
  - Access process and hold reports scoped to all cases.
  - Access case data for any case in the organization.
  - Manage any eDiscovery case after they add themselves as a member of the case.
  - Remove members from an eDiscovery case. Only an eDiscovery Administrator can remove members from a case. Users who are members of the eDiscovery Manager subgroup can't remove members from a case, even if the user created the case.
  - If a person who is the only member of an eDiscovery case leaves your organization, no one (including members of the *Organization Management* role group or another member of the *eDiscovery Manager* role group) can access that eDiscovery case because they aren't a member of a case. In this situation, there's no way to access the data in the case. But because an eDiscovery Administrator can access all eDiscovery cases in the organization, they can view the case and add themselves or another eDiscovery manager as a member of the case.
  - An eDiscovery Administrator can view and access all eDiscovery cases, they can audit and oversee all cases and associated compliance searches. This functionality can help to prevent any misuse of compliance searches or eDiscovery cases. And because eDiscovery Administrators can access potentially sensitive information in the results of a compliance search, you should limit the number of people who are eDiscovery Administrators.

    > **Note**
    >
    > To analyze a user's data when premium eDiscovery features are enabled, the user must be assigned an Office 365 E5 or Microsoft 365 E5 license. Alternatively, users with an Office 365 E1 or an Office 365 or Microsoft 365 E3 license can be assigned a Microsoft

Purview Suite (formerly known as Microsoft 365 E5 Compliance) or Microsoft 365 eDiscovery and Audit add-on license. Administrators, compliance officers, or legal personnel who are assigned to cases as members and use premium eDiscovery features to collect, view, and analyze data don't need an E5 license. For more information about subscriptions and licensing, see the subscription requirements for eDiscovery.

Before you assign permissions

- You need to be a member of the *Organization Management* role group or be assigned the *Role Management* role to assign eDiscovery permissions in the Microsoft Purview portal.
- You can use the Add-RoleGroupMember cmdlet in Security & Compliance PowerShell to add a mail-enabled security group as a member of the *eDiscovery Managers* subgroup in the *eDiscovery Manager* role group. However, you can't add a mail-enabled security group to the *eDiscovery Administrators* subgroup.

Configure app-only authentication for eDiscovery PowerShell

Important

App-only authentication for eDiscovery cmdlets in Security & Compliance PowerShell is unsupported. Transition automations to Microsoft Graph APIs where available. The following steps are best-effort guidance for existing automations that continue to use this unsupported configuration and don't change its support status.

Certificate-based authentication (CBA), also known as app-only authentication, requires both Microsoft Entra authorization and eDiscovery role-based access control (RBAC). The API permission allows the application to authenticate, while service principal registration and role group membership authorize eDiscovery cmdlets.

1. Follow App-only authentication for unattended scripts in Exchange Online PowerShell and Security & Compliance PowerShell to create the Microsoft Entra application and configure its certificate.

2. On the application, select **Add a permission** > **APIs my organization uses** > **Microsoft Exchange Online Protection** > **Application permissions** > **Exchange** >

Exchange.ManageAsApp, and then grant tenant-wide admin consent. The similarly named **Office 365 Exchange Online** permission isn't required for this Security & Compliance PowerShell connection. If the application also runs other Exchange Online automations, keep any permissions that those separate workflows require.

3. Connect to Security & Compliance PowerShell with an account that can manage role groups. Register the enterprise application's service principal in Exchange and Microsoft Purview:

   PowerShell

   ```
   New-ServicePrincipal -AppId <application-client-id> -ObjectId <enterprise-application-object-id> -DisplayName "<display-name>"
   ```

   The *AppId* value is the application (client) ID. The *ObjectId* value is the Object ID from **Enterprise applications**, not the Object ID from **App registrations**. Use a descriptive *DisplayName* value to make the service principal easy to identify.

4. In the Microsoft Purview portal, the *eDiscovery Manager* role group appears under **Settings** > **Role groups**. Assign the registered service principal to this role group by using Security & Compliance PowerShell and the same enterprise application Object ID:

   PowerShell

   ```
   Add-RoleGroupMember -Identity "eDiscoveryManager" -Member <enterprise-application-object-id>
   ```

   The *eDiscovery Manager* role group is the validated least-privileged assignment for app-only eDiscovery. Don't use *eDiscovery Administrator* as an interchangeable service principal assignment.

5. Verify the role group membership:

   PowerShell

   ```
   Get-RoleGroupMember -Identity "eDiscoveryManager"
   ```

6. Use ExchangeOnlineManagement 3.10.1 or later, and include the *EnableSearchOnlySession* switch when you connect:

   PowerShell

```
Connect-IPPSSession -CertificateThumbprint "<certificate-thumbprint>" -AppId "
<application-client-id>" -Organization "<tenant>.onmicrosoft.com" -
EnableSearchOnlySession
```

If the connection returns HTTP 401, verify that **Microsoft Exchange Online Protection** > **Exchange.ManageAsApp** is configured as an Application permission and has tenant-wide admin consent. If the connection succeeds but eDiscovery cmdlets aren't recognized, verify the service principal registration and its membership in the *eDiscovery Manager* role group.

Assign eDiscovery permissions

1. Go to the Microsoft Purview portal and sign in with an account that can assign permissions.
2. Go to **Settings** > **Role groups**.
3. On **Role groups for Microsoft Purview solutions**, select **eDiscovery Manager**.
4. On the **eDiscovery Manager** pane, do one of the following actions based on the eDiscovery permissions that you want to assign.
    - Select **Edit**.
    - On **Manage eDiscovery Manager**, select **Choose users** or **Choose groups**.
    - Search and select the users you want to add as an *eDiscovery Manager*, then select **Select**.
    - Select **Next**.
    - To assign users to the *eDiscovery Administrator* role group, select **Choose users** or **Choose groups**.
    - Search and select the users you want to add as an *eDiscovery Administrator*, then select **Select**.
    - Select **Next**.
5. If the selected users or groups need organization-wide access as part of this role group assignment, go to Step 8.
6. If the selected users or groups need to be assigned to administrative units, select the users or groups and select **Assign admin units**.
7. On **Assign admin units**, select the checkbox for all the administrative units you want to assign to the users or groups. Select **Select**.

8. Select **Next** and **Save** to add the users or groups to the role group. Select **Done** to complete the steps.

> **Note**
>
> You can also use the **Add-eDiscoveryCaseAdmin** cmdlet to make a user an eDiscovery Administrator. However, you must assign the *Case Management* role to the user before you can use this cmdlet to make them an eDiscovery Administrator. For more information, see Add-eDiscoveryCaseAdmin.

On the **Role groups** page in the Microsoft Purview portal, you can also assign users eDiscovery-related permissions by adding them to the *Compliance Administrator*, *Organization Management*, and *Reviewer* role groups. For a description of the eDiscovery-related role-based access control roles assigned to each of these role groups, see Role-based access control roles related to eDiscovery.

RBAC roles related to eDiscovery

The following table lists the eDiscovery-related role-based access control roles in the Microsoft Purview portal, and shows the built-in role groups that each role belongs to by default.

Expand table

Role	Compliance Administrator	eDiscovery Manager & Administrator	Organization Management	Reviewer
Case Management	✓	✓	✓	
Communication		✓		
Compliance Search	✓	✓	✓	
Custodian (Data Sources)		✓		
Export		✓		
Hold	✓	✓	✓	
Manage review set tags		✓		
Preview		✓		

Role	Compliance Administrator	eDiscovery Manager & Administrator	Organization Management	Reviewer
Review		✓		✓
RMS Decrypt		✓		
Search And Purge			✓	

> **Note**
>
> To view the list of eDiscovery cases, a user must have at least one of the roles listed in the preceding table. For users who aren't eDiscovery Administrators, the cases shown are limited to those where the user is a member. The Custodian role provides access to the **Data Sources** tab in eDiscovery to manage data sources related to specific cases.

Run the following diagnostic test to check if the *Export*, *Preview*, or *Search* roles are assigned to the designated admin account.

1. Select the **Help** control in the top right of the Microsoft Purview portal. Enter *Diag:edisRBACdiag* in search (or select this link) to run the **eDiscovery RBAC Check** test.
2. In the **Run diagnostics** section, enter the UPN or email address of the user trying to run an export, preview, or search task.
3. Select **Run Tests**. If the user doesn't have the necessary eDiscovery roles, assign the roles to perform the desired task.

Custom role combinations

If you need to provide custom access to specific feature components of eDiscovery for specific users, use custom role combinations for specific users. For example, you might need to allow a user to manage data sources, but without access to search features. For another user, you might need to allow to manage searches, but without access to data sources.

Consider using the following role combinations as needed:

- *Case Management* and *Search* roles for access to only search features.
- *Case Management* and *Custodian* roles for access to only data source features.
- *Case Management*, *Review*, and *Manage Tags* roles for access to only review features.
- *Case Management* and *Hold* roles for access to only hold features.
- *Case Management* and *Export* roles for access to only export features.

- *Case Management*, *Search*, and *Purge* roles for access to only purge features.

eDiscovery roles

The following sections describe each of the eDiscovery-related RBAC roles listed in the RBAC roles related to eDiscovery table.

Case Management

This role lets users create, edit, delete, and control access to eDiscovery cases in the Microsoft Purview portal. You must assign the *Case Management* role before you can use the **Add-eDiscoveryCaseAdmin** cmdlet to make a user an eDiscovery Administrator. For more information, see Get started with eDiscovery.

Compliance Search

This role lets users search mailboxes and public folders, SharePoint sites, OneDrive sites, Skype for Business conversations, Microsoft 365 groups, Microsoft Teams, and Viva Engage groups. This role allows a user to get an estimate of the search results and create export reports, but other roles are needed to initiate search actions such as previewing, exporting, or deleting search results.

In eDiscovery, users who are assigned the *Compliance Search* role but don't have the *Preview* role can preview the results of a search when a user assigned the *Preview* role initiates the preview action. The user without the *Preview* role can preview results for up to two weeks after the initial preview action was created.

Similarly, users in eDiscovery who are assigned the *Compliance Search* role but don't have the *Export* role can download the results of a search when a user assigned the *Export* role initiates the export action. The user without the *Export* role can download the results of a search for up to two weeks after the initial export action was created. After that, they can't download the results unless someone with the *Export* role restarts the export.

The two-week grace period for previewing and exporting search results (without the corresponding search and export roles) doesn't apply when premium features are enabled in eDiscovery. Users must be assigned the *Preview* and *Export* roles to preview and export content when premium eDiscovery features are enabled.

Custodian

This role lets users identify and manage custodians for eDiscovery cases managed in the retired Microsoft Purview compliance portal and use the information from Microsoft Entra ID and other sources to find data sources associated with custodians. The user can associate other data sources such as mailboxes, SharePoint sites, and Teams with custodians in a case. The user can also place a legal hold on the data sources associated with custodians to preserve content in the context of a case.

Export

This role lets users export search results to a local computer. It also lets them prepare search results for analysis when premium eDiscovery features are enabled. For more information about exporting search results, see Export search results in eDiscovery.

Hold

This role lets users place content on hold in mailboxes, public folders, sites, Skype for Business conversations, and Microsoft 365 groups. When content is on hold, content owners can still modify or delete the original content, but the content is preserved until the hold is removed or until the hold duration expires. For more information about holds, see Create a hold in eDiscovery.

Manage review set tags

This role lets users create, edit, and delete review set tags for cases they can access. Users need to at least have the *Review* role and this role to manage tags during reviews.

> **Note**
>
> eDiscovery guest users are limited to the *Reviewer* role group and can't be assigned the *Manage review set tags* role. Therefore, guest users can't perform review set tagging. For more information about guest-user permissions, see **Configure guest user settings in eDiscovery**.

Preview

This role lets users view a list of items that a search returns. They can also open and view each item from the list to view its contents.

Review

This role lets users access review sets in eDiscovery. Users who are assigned this role can see and open the list of cases that they're members of. After the user accesses an eDiscovery case, they can select **Review sets** to access case data. This role doesn't allow the user to preview the results of a search that's associated with the case or do other search or case management tasks. Users with this role can only access the data in a review set.

RMS Decrypt

This role lets users view rights-protected email messages when previewing search results and export decrypted rights-protected email messages. This role also lets users view (and export) a file that's encrypted with a Microsoft encryption technology when the encrypted file is attached to an email message that's included in the results of an eDiscovery search. Additionally, this role lets users review and query encrypted email attachments that are added to a review set in eDiscovery. For more information about decryption in eDiscovery, see Decryption in Microsoft 365 eDiscovery tools.

Search And Purge

This role lets users perform bulk removal of data matching the criteria of a search. For more information, see Find and delete email messages in eDiscovery.

Add role groups as members of eDiscovery cases

You can add role groups as members of eDiscovery cases so that members of the role groups can access and perform tasks in the assigned cases. The roles assigned to the role group define what members of the role group can do. When you add a role group as a member of the case, members can access and perform those tasks in a specific case.

With this requirement in mind, if you add or remove a role from a role group, the role group is automatically removed as a member of any case it belongs to. This behavior protects your organization from inadvertently providing extra permissions to members of a case. Similarly, if you delete a role group, you remove it from all cases it was a member of.

Before you add or remove roles to a role group that might be a member of an eDiscovery case, run the following commands in Security & Compliance PowerShell to get a list of cases the role group is a member of. After you update the role group, add the role group back as a member of those cases.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

Last updated on 09/18/2026

Configure general settings in eDiscovery

General settings in eDiscovery let you set the default license level for all cases you create in eDiscovery.

Only users assigned to the eDiscovery *Administrators* role can access this setting. When enabled, the Microsoft 365 Enterprise E5 subscription is the default level for all new eDiscovery cases. You can override this setting on a per-case basis after creating a new case in the settings for a case. This setting isn't available for organizations with Microsoft 365 Enterprise E3 subscription.

Enable or disable premium eDiscovery features

> **Note**
>
> Enabling premium eDiscovery features only affects new cases and doesn't affect the license state of any existing cases.

To enable or disable premium features for your organization:

1. Go to the Microsoft Purview portal and sign in with the credentials for a user account assigned eDiscovery *Administrator* permissions.
2. Select the **eDiscovery** solution card, then go to **Settings** > **eDiscovery** > **General**.
3. Select or clear the toggle for **eDiscovery (Premium)**.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 06/19/2026

Configure analytics settings in eDiscovery

This article explains how to configure organization-level analytics settings for eDiscovery in the Microsoft Purview portal. eDiscovery administrators can use these settings to enable machine learning-based detection of privileged content across all eDiscovery cases in the organization.

> **Note**
>
> The **Attorney-client privilege detection** feature is the only analytics setting currently available in eDiscovery organization settings. You configure additional case level search and analytics settings in the case settings for individual cases.

Configure attorney-client privilege detection

A major and costly aspect of the review phase of any eDiscovery process is reviewing documents for privileged content. eDiscovery provides machine learning-based detection of privileged content to make this process more efficient. This feature is called *attorney-client privilege detection*.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

How attorney-client privilege detection works

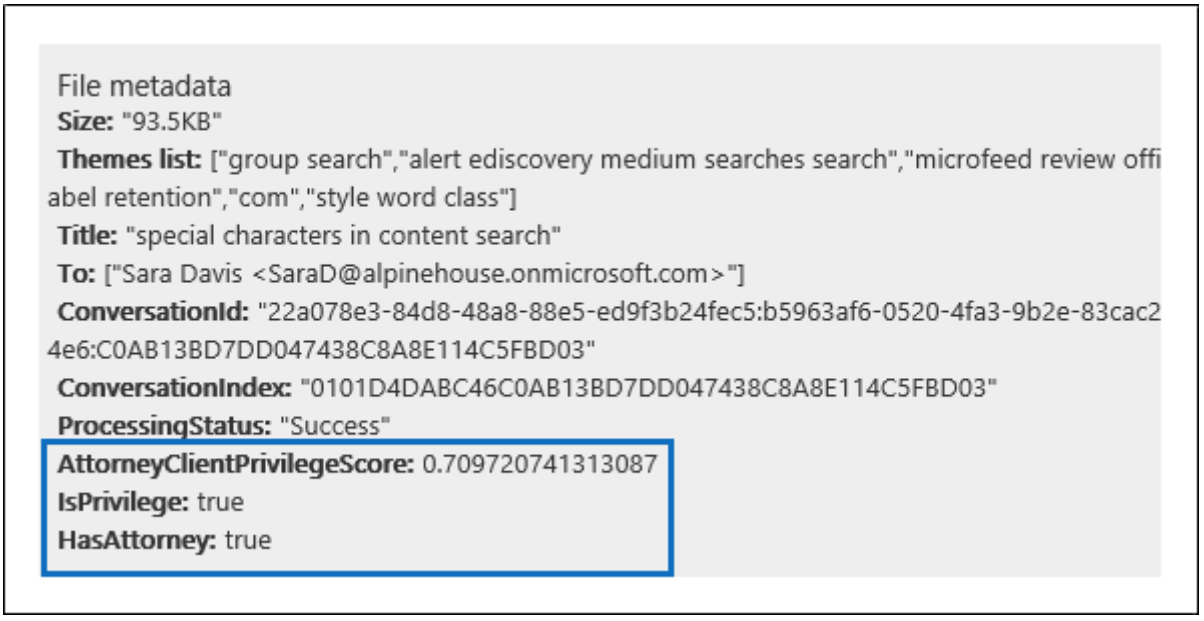
When you enable attorney-client privilege detection, the model processes all documents in a review set when you analyze the data in the review set. The model looks for two things:

- **Privileged content**: The model uses machine learning to determine the likelihood that the document contains content that is legal in nature.
- **Participants**: As part of setting up attorney-client privilege detection, you submit a list of attorneys for your organization. The model then compares the participants of the document with the attorney list to determine if a document has at least one attorney participant.

The model produces the following three properties for every document:

- **AttorneyClientPrivilegeScore**: The likelihood the document is legal in nature; the values for the score are between *0* and *1*.
- **HasAttorney**: This property is set to *true* if one of the document participants is listed in the attorney list; otherwise the value is *false*. The value is also set to *false* if your organization didn't upload an attorney list.
- **IsPrivilege**: This property is set to *true* if the value for **AttorneyClientPrivilegeScore** is over the threshold *or* if the document has an attorney participant; otherwise the value is set to *false*.

These properties (and their corresponding values) are added to the file metadata of the documents in a review set and are searchable within a review set:



Set up the attorney-client privilege detection model

To enable the attorney-client privilege detection model, your organization needs to turn it on and (optionally) upload an attorney list.

Step 1: Turn on attorney-client privilege detection

An eDiscovery Administrator in your organization (a member of the eDiscovery Administrator subgroup in the eDiscovery Manager role group) must make the model available in your eDiscovery cases.

1. In the Microsoft Purview portal, go to **Settings** > **eDiscovery** > **Analytics**.
2. Toggle **Attorney-client privilege detection** to *On*.
3. Select **Save** to save the change.

Step 2: Upload a list of attorneys (optional)

To take full advantage of the attorney-client privilege detection model and use the results of the **Has Attorney** or **Potentially Privileged** detection, upload a list of email addresses for the lawyers and legal personnel who work for your organization.

To upload an attorney list for the attorney-client privilege detection model:

1. Create a .csv file (without a header row) and add the email address for each appropriate person on a separate line. Save this file to your local computer.
2. In the Microsoft Purview portal, go to **Settings** > **eDiscovery** > **Analytics**.
3. Select **Choose file**, then find and select the .csv file that you created in step 1.
4. Select **Save** to upload the attorney list.

Use the attorney-client privilege detection model

Complete the following steps to use attorney-client privilege detection for documents in a review set.

Step 1: Create a smart tag group with attorney-client privilege detection model

One of the primary ways to see the results of attorney-client privilege detection in your review process is by using a smart tag group. A smart tag group shows the results of the attorney-client privilege detection in-line next to the tags. This setup lets you quickly identify potentially privileged documents during document review. You can also use the tags in the smart tag group to tag documents as privileged or nonprivileged. For more information about smart tags, see Set up smart tags.

1. In the review set that contains documents you previously analyzed, select **Manage review set** and then select **Manage tags**.

2. Under **Tags**, select the pull-down next to **Add group** and then select **Add smart tag group**.

3. On the **Choose a model for your smart tag** page, select **Select** next to **Attorney-client privilege**.

   A tag group named **Attorney-client privilege** appears. It contains two child tags named **Positive** and **Negative**, which correspond to the possible results produced by the model.

4. Rename the tag group and tags as appropriate for your review. For example, you can rename **Positive** to **Privileged** and **Negative** to **Not privileged**.

Step 2: Analyze a review set

When you analyze the documents in a review set, the attorney-client privilege detection model also runs and adds the **AttorneyClientPrivilegeScore**, **HasAttorney**, and **IsPrivilege** properties (described in How attorney-client privilege detection works) to every document in the review set. For more information about analyzing data in review set, see Analyze data in a review set in eDiscovery.

Step 3: Use the smart tag group for review of privileged content

After analyzing the review set and setting up smart tags, the next step is to review the documents. If the model determines the document is potentially privileged, the corresponding smart tag in the **Tagging panel** indicates the following results produced by the attorney-client privilege detection:

- If the document has content that might be legal in nature, the label **Legal content** appears next to the corresponding smart tag (which in this case is the default **Positive** tag).
- If the document has a participant who is found in your organization's attorney list, the label **Attorney** appears next to the corresponding smart tag (which in this case is also the default **Positive** tag).
- If the document has content that might be legal in nature *and* has a participant found in the attorney list, both the **Legal content** and **Attorney** labels appear.

If the model determines that a document doesn't contain content that is legal in nature or doesn't contain a participant from the attorney list, then neither label appears in the tagging panel.

After you review a document to see if it contains privileged content, you can tag the document with the appropriate tag.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/29/2026

Configure guest user settings in eDiscovery (preview)

With guest access, you can provide access to eDiscovery cases to people outside your organization. You can invite guest users to eDiscovery cases just like you can invite guests into your Teams environment.

A guest is someone who doesn't have a work or school account with your organization. For example, guests can include external partners, vendors, suppliers, or consultants. Anyone who isn't part of your organization can be added as a guest for an eDiscovery case. This access means that anyone with a Microsoft Entra account or personal email account (Outlook, Gmail, or others) can participate as a guest in eDiscovery and have access to specific eDiscovery cases.

Guest access in eDiscovery uses Azure B2B technology, which depends on an organization-wide Azure collaboration setting. When you invite a guest to eDiscovery, a guest account is created for them in Microsoft Entra ID. They're covered by the same compliance and auditing protection as other Microsoft 365 users. Guest access is subject to Microsoft Entra ID and Microsoft 365 service limits.

> **Note**
>
> eDiscovery guests are currently limited to *Reviewer* role group permission only. The *Reviewer* role allows guests to access review sets and review case data, but it doesn't include the *Manage review set tags* role. As a result, guest users can't create, edit, delete, or apply review set tags. Review set tagging requires both the *Review* and *Manage review set tags* roles. For more information about the reviewer role, see **Assign eDiscovery permissions in the Microsoft Purview portal**.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about **Microsoft Security Copilot in Microsoft Purview**.

Set up guest access

Setting up guest access in eDiscovery requires configuring other settings in Microsoft Entra ID. To enable access to invite guests to eDiscovery cases, complete the following steps:

1. Check the B2B external collaboration settings to ensure that collaboration with guests isn't blocked. If the **Guest invite settings** is set to *No one in the organization can invite guests including admins* then inviting guests into eDiscovery is disabled. Contact your system administrator to ensure the B2B external collaboration settings are properly configured. You can view and update the collaboration setting using the Microsoft Graph authorizationPolicy resource.

2. Depending on the specific configuration in Step 1, eDiscovery administrators might or might not need additional roles to approve a guest access request submitted from the eDiscovery user.

    - If the *Anyone in the organization can invite guest including guests and nonadmins (most inclusive)* or *Member users and users assigned to specific admin roles can invite guests including guests with member permission* options are selected, then eDiscovery administrators can approve invitation requests without extra roles.
    - If *Only users assigned to specific admin roles can invite guests* is selected, then the eDiscovery administrator who approves the invitation request requires the *Invite guest* role to approve and trigger the invitation email to the guest. See the Assign Microsoft Entra roles to users article for guidance on how to add the *Guest Inviter* role to users. If you prefer to use PowerShell to add this role, see Assign the Guest Inviter role to a user.

3. To enable collaboration with guest reviewers from other Microsoft 365 cloud environments, including if you're a Microsoft Azure Government organization, both your organization and the guest's organization must enable specific connection settings. For detailed configuration guidance, see Collaborate with guests from other Microsoft 365 cloud environments.

Enable guest access in eDiscovery

eDiscovery admins must enable guest access in eDiscovery before they can invite users as guests.

To enable guest access for guests, complete the following steps:

1. In the Microsoft Purview portal, go to **Settings** > **eDiscovery** > **Guest users**.
2. Set **Guest access** to *On*.

Now eDiscovery managers and admins can start the guest invitation process within any eDiscovery case. To configure guest access to specific cases, see Learn about access and permission settings in eDiscovery cases.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 09/18/2026

Configure tag templates in eDiscovery

Use tag templates to define and standardize tags and the review workflow across all cases and review sets in your organization. Each tag template contains tag groups where you can create a grouping structure for individual tags.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Create tag templates

To create a new tag template, complete the following steps:

1. In the Microsoft Purview portal, select **Settings** > **eDiscovery** and then select **Tag templates**.
2. Select **Create**.
3. On **New tag template**, enter a name and description for the tag template.
4. In the initial tag group, enter a tag group name and use the dropdown to define this group as a **Multi-select** or **Single-select** tag group.
5. Enter at least one tag name in the tag group. Select **Add tag** to add more tags to the tag group.
6. If needed, select **Add tag group** to add more tag groups and tags to the tag template.
7. Preview the tag template group and tag structure in the **Preview** pane on the right side of **New tag template**.
8. When complete, select **Save**.

The tag template is now available for use in all cases and review sets in your organization.

Manage tag templates

After you create a tag template, you can edit or delete it. You can also view information about the tag template, including the dates when you created or last modified it.

1. In the Microsoft Purview portal, select **Settings** > **eDiscovery** and then select **Tag templates**.

    - To edit a tag template, select the tag template name on **Manage tag templates** and select **Edit**. On **Edit tag template**, you can edit the tag template name, description, and tag groups and tags. When done, select **Save**.
    - To delete a tag template, select the tag template name on **Manage tag templates** and select **Delete**. Select **Yes** to confirm the deletion or **No** to cancel the deletion.

Apply tag templates to a case and review sets

After you create a tag template, you can apply it to a review set. You can apply a tag template when you create a new review set or when you edit an existing review set.

To apply a tag template to a new review set, complete the following steps:

1. In the Microsoft Purview portal, select **eDiscovery** > **Cases** and select a case.
2. In the case, select **Review sets**, select a review set, then select the **All items** tab.
3. On the review set command bar, select **Manage** > **Tags**.
4. On **Manage tags**, delete the blank tag group, then select **Import tag templates**.
5. On **Import tag templates**, select the tag template that you want to apply to the review set, then select **Import**.
6. If needed, edit the tag template as it applies to the case and review set.
7. Select **Save** to apply the tag templates to the case and review set.

The applied tag template is now available for use in the case and all included review sets.

ⓘ **Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 04/02/2026

Configure export settings in eDiscovery

> **Note**
>
> This setting isn't available in DoD cloud organizations.

Export settings let you manage download links for all exports in eDiscovery. When exporting large eDiscovery results, network interruptions and browser limitations might affect the download experience. Whether you enable the **Enable export download via pre-authorized link** option depends on your environment, security requirements, network bandwidth, and the tools you plan to use.

> **Note**
>
> For export download best practices, see Best practices for downloading export packages.

Use a pre-authorized link when downloading large export packages or when your network connection might be limited. Large exports take longer to download, which naturally increases the chances of a network interruption during the process. Pre-authorized links are designed for this scenario. By using a pre-authorized link, the download can resume when interrupted. Using a pre-authorized link works with tools that support parallel or chunked downloads, such as third-party download accelerators or browser extensions. When you initiate the download, the pre-authorized download link is generated. The link acts like an access token, so tools can download data without repeated sign-ins.

These links grant direct access to your export. If shared, anyone with the link can access the data. To reduce risk, set the shortest possible expiration time based on the size of your export. You can configure export package download link availability for the following durations:

- 1 hour
- 6 hours
- 12 hours
- 24 hours (default)
- 48 hours
- 72 hours
- 168 hours

Disable pre-authorized link for high bandwidth environments. The standard signed-in download method without pre-authorized links is ideal when you have an ExpressRoute connection with high bandwidth to Microsoft. The standard signed-in download method doesn't generate a pre-authorized token, reducing the risk of link misuse. Each session requires sign-in. If the download is interrupted, you need to re-authenticate and restart. Browser resume behavior varies and isn't guaranteed.

Enable preauthorized export download links

> **Important**
>
> Don't share export download links with anyone. Treat these links as secrets.

Changing this setting to turn preauthorized links on or off only applies to future export downloads. It doesn't affect existing links. For example, if you enable preauthorized links with a one-day expiration, any links you created before you disable the setting remain valid until they expire. Turning off the pre-authorized export download link setting doesn't retroactively change the behavior of existing links.

To enable preauthorized export download links for your organization:

1. Go to the Microsoft Purview portal and sign in with the credentials for a user account assigned eDiscovery *Administrator* permissions.
2. Select the **eDiscovery** solution card, then go to **Settings** > **eDiscovery** > **Export**.
3. Select the toggle for **Enable export download via pre-authorized link**.
4. Select a duration in **Export Link Settings**.
5. Select **Save**.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 06/29/2026

Create and manage cases in eDiscovery

After setting up Microsoft Purview eDiscovery and assigning permissions to eDiscovery managers in your organization who manage cases, create and manage a case for escalation activities or investigation requests.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Cases dashboard

The **Cases** dashboard in the Microsoft Purview portal displays all eDiscovery cases in your organization. You can quickly see the cases in your organization, the status of each case, and important information associated with each case.

- **Name**: The name of the case. The case name must be unique in your organization.
- **Case status**: The current status of the case.
- **Case type**: The type of case.
- **Created on**: The date and time in Coordinated Universal Time (UTC) when the case was created.
- **Last modified**: The date and time in Coordinated Universal Time (UTC) when the case was last modified.
- **Last modified by**: The name of the user that last modified the case.
- **Number**: The optional docket number or other custom numeric identifier for the case.
- **Description**: The description of the case entered when the case was created.

Select **Customize columns** to customize the displayed columns and the order of the columns in the **Cases** dashboard. Choose the columns to display or drag and drop columns to reorder.

Filter cases

To filter cases on the **Cases** dashboard, enter a keyword to filter by case names that contain the keyword or select **Filter** on the command bar. Select one or more of the following filters and select **Apply** to help you organize cases for review on the dashboard.

Case status options:

- **Active**: The case is active.
- **Pending delete**: The case likely has a hold in the pending deletion state. Check the hold policy and redistribute to resolve the error. Afterward, cases without any holds should be deleted successfully.
- **Closing**: The case is being closed. This status is seen only for case with active holds. Cases without holds are immediately closed.
- **Closed**: The case is closed without errors and the case can be reopened.
- **Closed with errors**: The case is closed, but there are errors. Errors are most likely associated with holds that weren't released during closing. To resolve, check policies with errors, resolve the errors, and redistribute the policies.

Case type options:

- **Premium**: eDiscovery (Premium) cases.
- **Standard**: eDiscovery (Standard) cases.
- **User data search**: Legacy user data search cases.
- **Content search**: Content searches in your organization.

> **Important**
>
> Case type is used only to map cases to the Microsoft Purview compliance portal during the preview of eDiscovery in the Microsoft Purview portal. This mapping helps you find cases in the compliance portal (if needed). All cases created in the Microsoft Purview portal are the *Premium* type, regardless of your license or subscription. Access to premium features isn't governed by the case settings toggle.

To clear **Case** dashboard filters, select **Filter** > **Clear filters** > **Apply**.

Grouping cases

To group cases on the **Cases** dashboard, select **Group** on the command bar. Select one of the following groupings to help you organize cases for review on the dashboard.

- **None**: Clears grouping for cases.
- **Status**: Groups cases by the status of the case.
- **Last modified by**: Groups cases by the user that last modified the case.
- **Case type**: Groups cases by the case type.

Download a list of cases

To download a list of all cases displayed in the **Cases** dashboard, select **Download list**. A .csv file is generated that contains the case name and all of the **Cases** dashboard column properties for each case.

Create a case

> **Tip**
>
> Do you prefer an interactive configuration guide experience? Check out the **Create and configure a case** guide.

Complete the following steps to create a case and configure case properties. The user who creates the case is automatically added as a member. Members of the case can access the case in the Microsoft Purview portal and perform eDiscovery tasks.

1. Go to the Microsoft Purview portal and sign in using the credentials for a user account assigned eDiscovery permissions. Members of the *Organization Management* role group can also create eDiscovery cases.
2. Select the **eDiscovery** solution card and then select **Cases** in the left navigation.
3. Select **Create case**.
4. On **Enter details to get started**, complete the following fields:
    - **Case name**: give the case a name (required). The case name must be unique in your organization.
    - **Case description**: Add an optional description to help others understand this case.
5. Select **Create** to create the new case and start your investigation.

Add data sources to a case

> **Tip**
>
> Want to try premium eDiscovery features? See the **subscription requirements** for Microsoft 365 Enterprise E5 licensing.

Data sources define the locations that you can quickly include in all case searches and holds. Add data sources to a case with premium eDiscovery features enabled to quickly use data sources across multiple case activities. When you add a data source at the case level, you include all associated resources (mailboxes, sites, and more) for the data source.

SharePoint Embedded container support in eDiscovery (preview) lets you add supported container URLs individually or through bulk import as case data sources.

To discover supported containers relevant to a selected user or filter containers by application, manage the data source in an individual search or hold. Group-owned and app-owned SharePoint Embedded containers aren't currently supported.

For step-by-step guidance on adding data sources to a case, see Data sources in eDiscovery.

Summarize a case with Copilot (preview)

When you summarize cases with Copilot, you can quickly access a comprehensive summary of eDiscovery cases. Copilot consolidates case information into a single, easy-to-understand summary. Copilot analyzes content within the case and generates natural language summaries, giving you a clear, actionable overview of key details in real-time. It eliminates the need to navigate through multiple case tabs by consolidating status updates, completed actions, pending tasks, and ongoing jobs into a single, easy-to-read summary. To use Copilot to summarize a case, Security Copilot must be enabled and the user must be assigned the *eDiscovery Administrator* or *eDiscovery Manager* role.

Whenever you view a case, Copilot automatically reviews the case and displays summary information in the Copilot pane. You can also select **Summarize this case** to run Copilot and display or refresh this summary and pane. When Copilot completes the summary, it displays details about the case in the following areas:

- **Case information**: A high level summary information about the case, including information about the case type, when the case was created, when the case was modified, who the last user was that modified the case, and more.
- **Data sources**: A summary of data sources associated with the case.
- **Exports**: A summary of search and review set exports associated with the case, including details about export items, file sizes, and item counts.
- **Holds**: A summary of hold policies associated with the case, including hold policy, query, and location details.
- **Jobs**: A summary of all the processes associated with the case.

- **Review sets**: A summary of the review sets associated with the case, when they were created, modified, and who edited.
- **Searches**: A summary of searches associated with the case, including query conditions, sources, options, and settings.
- **Settings**: A summary of the case settings, such as OCR, near-duplicate detection, search, document processing, and review set settings.

If the case doesn't have activity for an area, the Copilot summary indicates that the area doesn't have any activities to summarize. For example, if the case doesn't have associated exports, the Copilot summary in **Exports** section confirms that the case doesn't have any exports.

You can also enter customized prompts directly in Copilot to further refine and narrow the scope to specific areas about the case.

To change the data source used for the Copilot summary, select **Microsoft Learn documentation** at the end of the summary. This scopes the Copilot summary to support guidance that can help you learn more about working with cases and can provide step by step guidance or details about eDiscovery features and capabilities.

Mark a case as a favorite

You can mark an eDiscovery case as a favorite for quicker access to cases you want to prioritize. You can quickly access cases marked as favorites through the **Cases** dashboard. To mark a case as a favorite, select the star icon to the left of a case in the case list on the **Cases** dashboard.

Limited format cases

Some cases you previously created in the eDiscovery classic experience might have the *Limited format* status applied. If applicable to a case, this tag appears next to the case title when on any of the action tabs within a case (**Searches**, **Hold policies**, etc.). These cases have some limitations from the features and functionality available in cases created in the Microsoft Purview portal and with the new eDiscovery experience. Advanced indexing isn't available for searches, exporting items from a search, and when adding a search results to a review set.

Manage case settings

Case settings include case information, case license level control, access permissions, and settings that control search and analytics behavior for the case. You can access settings for a specific case by selecting **Case settings** after selecting a case. You can define specific parameters for feature

capabilities based on eDiscovery licenses in your organization, which users, role groups, and guests have access and permissions for the case, and which additional locations are included during organization-wide searches for the case.

For more information about case settings, see:

- Learn about case settings in eDiscovery cases
- Learn about access and permission settings in eDiscovery cases
- Learn about data sources settings in eDiscovery cases
- Learn about search and analytics settings in eDiscovery cases
- Learn about review set settings in eDiscovery cases

Manage case status

> ⚠ **Warning**
>
> Closing or deleting a case turns off all holds. Any content that was on hold is released, which might result in data loss.

To close or delete a case, select **Actions** for an individual case and choose **Close case** or **Delete case**.

For more information about managing case actions, see:

- Closing a case
- Reopening a case
- Deleting a case

Related content

For more information about tracking eDiscovery case activities, see the following resource:

- Search for eDiscovery activities in the audit log

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 09/19/2026

Learn about case settings in eDiscovery

Case settings in eDiscovery let you view and update case information and take action on specific cases in eDiscovery.

The following information displays for the selected case:

- **Premium features**: Shows the status of premium eDiscovery features enabled in the case. To use premium eDiscovery features for the selected case, turn on the **eDiscovery (Premium)** toggle.

> **Note**
>
> Premium features are enabled by default for tenants with premium feature access. To turn off, toggle the setting after a case is created.

- **License**: Shows the current licensing subscription for the case. Values are *eDiscovery (Premium)* or *eDiscovery (Standard)*.
- **ID**: Shows the case identification number. You can review this number, but you can't change it once the case is created.
- **Case name**: Shows the case name. This field is required. To change the case name, enter the new case name and select **Actions** > **Save**.
- **Case number**: Shows an optional docket number or other numeric identifier for the case. To change the case number, enter the new case number and select **Actions** > **Save**.
- **Case description**: Shows an optional description to help others understand this case. To change the case description, enter the new case description and select **Actions** > **Save**.
- **Case status**: Shows the current case status.
- **Case created**: Shows the date and time when the case was created.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Close a case

When you complete the legal case or investigation supported by eDiscovery, you can close the case. Here's what happens when you close an eDiscovery case:

> **Note**
>
> You need to either disable or delete all active hold policies in the case before you can close the case.

- Delete any active hold policies. After you delete the hold, a 30-day grace period (called a *delay hold*) is applied to content locations that were on hold. This delay hold helps prevent content from being immediately deleted and gives you an opportunity to search for or recover content that is permanently deleted after the delay hold period expires.
- The case is still listed on the **Cases** page in the Microsoft Purview portal. The details, holds, searches, and members of a closed case are retained.
- You can edit a closed case. For example, you can add or remove members, create searches, export search results, and prepare search results for analysis in eDiscovery.

To close a case, complete the following steps:

1. Go to the Microsoft Purview portal and sign in using the credentials for a user account assigned eDiscovery permissions.
2. Select the **eDiscovery** solution card and then select **Cases**.
3. Select a case, then navigate to the *Hold policies* tab.
4. Delete all hold policies in the case.
5. Once the hold policies are deleted, select **Case settings**.
6. On **Case settings**, select **Actions**, then select **Close case**. It might take up to 60 minutes for the closing process to complete.

When you close a case, you see the status of the case changing as the case is processed for closing.

Reopen a closed case

When you reopen an eDiscovery case, the process doesn't automatically reinstate any holds that were in place when the case was closed. After you reopen the case, go to the **Holds** tab and turn on the previous holds. To turn on a hold, select it to display the flyout page, then set the **Status** toggle to **On**.

To reopen a closed case, complete the following steps:

1. Go to the Microsoft Purview portal and sign in using the credentials for a user account assigned eDiscovery permissions.
2. Select the **eDiscovery** solution card, then select **Cases** in the left nav.
3. Select a case, then select **Case settings**.
4. On **Case settings**, select **Actions**, then select **Reopen case**. It might take up to 60 minutes for the reopening process to complete.

Delete a case

You can delete both active and closed eDiscovery cases. When you delete a case, you also delete all components associated with the case, such as the list of communications, searches, review sets, and exports. You can't recover or reopen a deleted case.

> **Note**
>
> In data spillage scenarios, the only way to remove items in a review set is to delete the eDiscovery case. Other search and purge methods don't remove items from a review set.

Before you can delete a case, you must delete *all* hold policies listed on the **Hold policies** tab of the case. This requirement includes deleting holds with a status of **Off**.

To delete hold policies associated with a case, complete the following steps:

1. Go the **Hold policies** tab in the eDiscovery case that you want to delete.
2. Select the hold policy that you want to delete.
3. On the hold policy page, select **Policy actions** > **Delete policy**.

To delete a case, complete the following steps:

1. Go to the Microsoft Purview portal and sign in using the credentials for a user account assigned eDiscovery permissions.
2. Select the **eDiscovery** solution card, then select **Cases** in the left nav.
3. Select a case, then select **Case settings**.
4. On the **Case settings** page, select **Actions**, then select **Delete case**.

Last updated on 04/02/2026

Learn about access and permission settings in eDiscovery cases

Access and permission settings in eDiscovery let you add or remove users to a case, manage role group membership for a case, and add people outside your organization as guest users for a case.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Add or remove users from an eDiscovery case

You can add or remove users to manage who can access the case. However, before a user can access an eDiscovery case (and perform tasks in the case), you must add the user to the eDiscovery Manager role group in the Microsoft Purview portal. For more information, see Assign eDiscovery permissions.

Add users to a case

Complete the following steps to add users to a case:

1. Go to the Microsoft Purview portal and sign in with the credentials for a user account assigned eDiscovery permissions.
2. Select the **eDiscovery** solution card and then select **Cases** in the left navigation.
3. Select a case, then select **Case settings**.
4. On **Case settings**, select **Access and permissions**.
5. Under **Users**, select **Add** to add users to the case. You can also choose to add a role group to the case by selecting **Add** under **Role groups**.
6. In the list of users or role groups that you can add to the case, select the check box next to the names of the users or role groups that you want to add.

> **Note**
>
> When adding a role group to a case, you can only add the role groups that you're a member of.

7. After you select the people or role groups to add as members of the case, select **Add**. The selected users are added to the case.

> **Important**
>
> If you add or remove a role from a role group that you added as a member of a case, the role group is automatically removed as a member of the case (or any case the role group is a member of). This behavior protects your organization from inadvertently providing extra permissions to members of a case. Similarly, if you delete a role group, it's removed from all cases it was a member of. For more information, see Assign eDiscovery permissions.

Remove users from a case

Only an eDiscovery Administrator can remove users from a case. Even if you're assigned to the *eDiscovery Manager* role group or initially created the case, you can't remove yourself or other members from a case unless you're also an eDiscovery Administrator. To remove yourself or other members from a case, contact an eDiscovery Administrator in your organization.

Complete the following steps to remove users from a case:

1. Go to the Microsoft Purview portal and sign in with the credentials for a user account assigned eDiscovery permissions.
2. Select the **eDiscovery** solution card and then select **Cases** in the left navigation.
3. Select a case, then select **Case settings**.
4. On **Case settings**, select **Access and permissions**.
5. Under **Users**, select **Remove** to remove users from the case. You can also remove users from a role group for the case by selecting **Remove** under **Role groups**.
6. In the list of users or role groups that you can remove from the case, select the check box next to the names of the users or role groups that you want to remove.
7. After you select the people or role groups to add as members of the case, select **Remove**. The selected users are removed from the case.

Guest access in eDiscovery

With guest access, you can provide access to an eDiscovery case to people outside your organization. You can invite guest users to eDiscovery cases just like you can invite guests into your Teams environment. eDiscovery admins must enable guest access in eDiscovery before they can invite users as guests.

Invite guests to an eDiscovery case

After you enable guest access for eDiscovery, complete the following steps to invite guests to an eDiscovery case:

1. In the Microsoft Purview portal, go to **eDiscovery** > **Cases** and select the case you want to invite guests to.
2. In the selected case, select the **Case settings** tab.
3. Select **Access and permissions**.
4. In the **Guest users** section on the **Access and permissions** page, select **Invite**.
5. On the **Invite guest** flyout pane, complete the following fields:
    - **Full name**: Enter the full name of the guest you want to invite.
    - **Email address**: Enter the email address of the guest you want to invite.
    - **Organization**: Enter the organization of the guest you want to invite.
    - **Guest role**: Select the role you want to assign to the guest. The guest is currently limited to *Reviewer* role group permission only. The *Reviewer* role doesn't include the *Manage review set tags* role, so guest users can't perform review set tagging.
    - **Justification**: Explain why this person needs access to the case.
6. Select **Invite** to submit the invitation request for review and approval by an eDiscovery admin.

Approve guest access requests

After submitting guest invitations, eDiscovery admins must approve the requests before the guest can access the case. eDiscovery managers receive notifications of pending invitation requests on the **Guest users** page in eDiscovery settings in the Microsoft Purview portal. To approve guest access requests, eDiscovery administrators must have the *Role Management* role assigned.

To review and approve guest access requests, eDiscovery managers complete the following steps:

1. Go to the Microsoft Purview portal and sign in with the credentials for a user account assigned eDiscovery permissions.
2. Go to **Settings** > **eDiscovery** > **Guest users**.
3. Select the guest request and select **Approve** or **Deny** from the flyout pane.

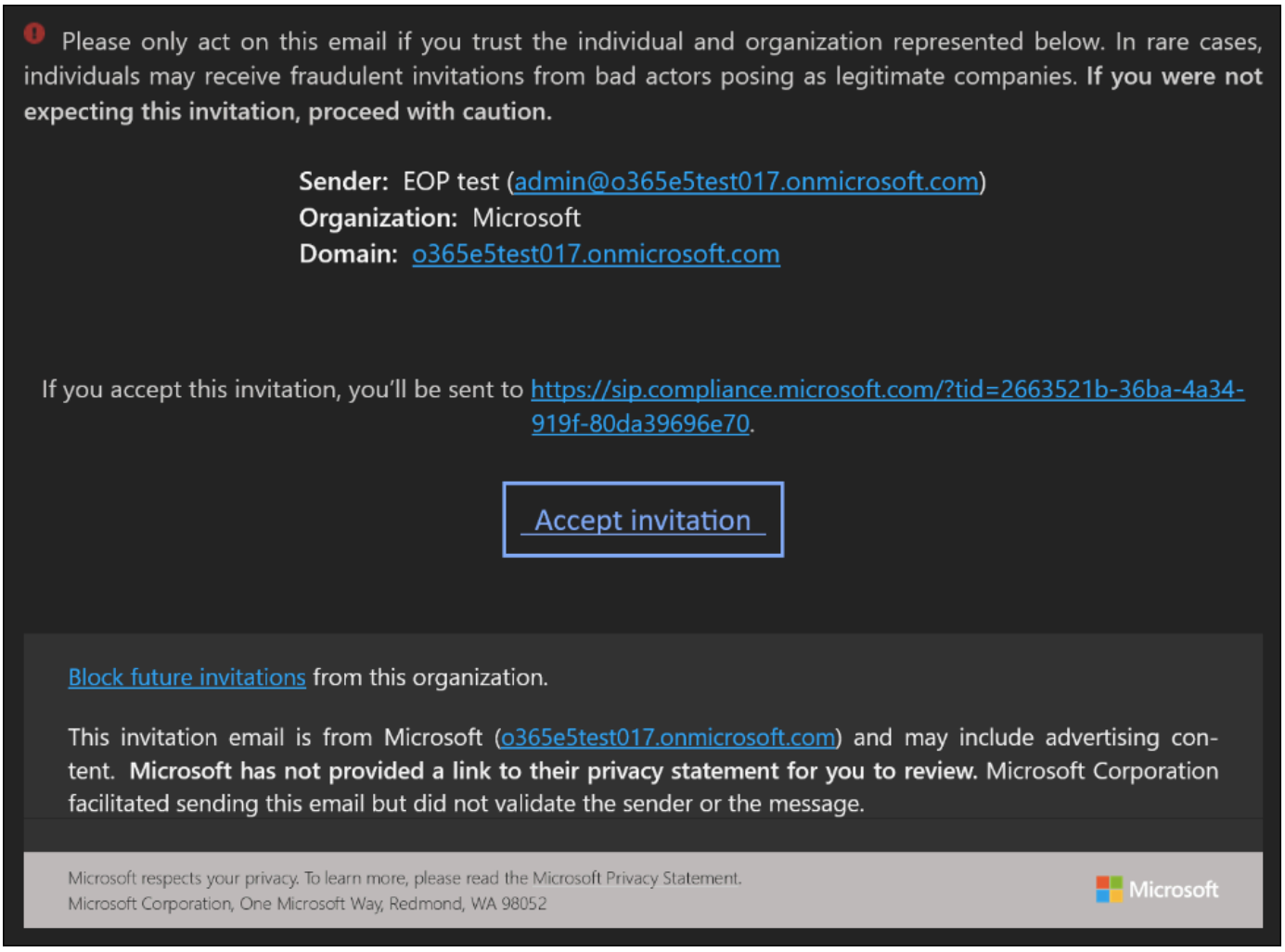
If the guest is being added to the tenant for the first time, eDiscovery admins need to perform a two-step approval process. They first need to approve creation of an account, which might take up to 24 hours. After 24 hours, eDiscovery admins can approve guest access to the case. For future access requests for access to other cases for the same guest, only one-step approval is required.

If you encounter an error when approving the guest access to a case, check the following points:

- The guest's email isn't disabled, blocked, or deleted by the tenant admin.
- If there are any members in the *Reviewer* role group whose Microsoft Entra account is deleted. Currently, an error occurs when adding a new member to a role group that has any deleted users. If the *Reviewer* role group has any deleted users, remove them from the role group and retry the guest approval.

Accepting guest invitations

After the eDiscovery admin approves the request, the guest receives a welcome email from *Microsoft Invitations*. This email contains information about the eDiscovery invitation and a link where the guest can accept the invite and sign in. The guest must select the unique link provided in the email or copy and paste the link in a browser to sign in to the specific eDiscovery organization.



Guests with a work or school account in Microsoft Entra ID can accept the invitation and authenticate immediately. Other users receive a one-time passcode to validate their identity. You must ensure that the one-time passcode setting is enabled in your organization.

> **Note**
>
> If the guest is signed in to another eDiscovery organization in their browser, they need to sign out from the organization and clear their browser cache before selecting the link in the invitation email.

After accepting the invitation and authenticating, the guest can participate in the eDiscovery case like an internal reviewer of the case and perform review activities such as querying, filtering, redaction, downloading single items for review, and more. Guest users can't perform review set tagging because tagging requires the *Manage review set tags* role, which isn't included in the *Reviewer* role group available to guest users.

Remove guest access

An eDiscovery admin can remove a guest's access to a case from the eDiscovery settings page. To remove guest access for a user, eDiscovery admins must complete the following steps:

1. Go to the Microsoft Purview portal and sign in with the credentials for a user account assigned eDiscovery permissions.
2. Go to **Settings** > **eDiscovery** > **Guest users**.
3. Select the guest and select **Remove** from the flyout pane.

The guest keeps access to other cases they were added to. When you remove a guest from the last case they have access to, you remove them from the *Reviewer* role group and they lose access to all cases in eDiscovery.

The guest keeps their Microsoft Entra guest account in the organization. Regularly review and clean up guest accounts from Microsoft Entra ID.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 09/18/2026

Learn about search and analytics settings in eDiscovery cases

You can configure settings for each eDiscovery case to control the following functionality:

- Near duplicates and email threading
- Themes
- Autogenerated review set query
- Ignore text
- Optical character recognition

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Configure analytics settings for a case

To configure search and analytics settings for a case:

1. Go to the Microsoft Purview portal and sign in with the credentials for a user account assigned eDiscovery permissions.
2. Select the **eDiscovery** solution card, then select **Cases** in the left nav.
3. Select a case, then select **Case settings**.
4. On **Case settings**, select **Search & analytics**.
5. The case **Search & analytics** page appears. These settings apply to all review sets in a case.
6. After selecting the applicable search and analytics options, select **Save**.

The following sections describe the analytics settings that you can configure for a case.

Near duplicates and email threading

In this section, set parameters for duplicate detection, near duplicate detection, and email threading.

- **Near duplicates/email threading:** When you turn on this setting, the workflow includes duplicate detection, near duplicate detection, and email threading when you run analytics on the data in a review set.
- **Document and email similarity threshold:** If the similarity level for two documents is over the threshold, both documents are in the same near duplicate set.
- **Minimum/maximum number of words:** These settings specify that near duplicates and email threading analysis are performed only on documents that have at least the minimum number of words and at most the maximum number of words.

Near duplicate detection

Consider a set of documents to review where a subset of documents uses the same template and mostly the same boilerplate language, with a few differences. If a reviewer can identify this subset, review one of the documents thoroughly, and review the differences for the rest, they don't miss any unique information while taking only a fraction of the time it would take to read all documents cover to cover. Near duplicate detection groups textually similar documents together to help you make your review process more efficient.

When you run near duplicate detection, the system parses every document with text. Then, it compares every document against each other to determine whether their similarity is greater than the set threshold. If it is, the system groups the documents together. Once all documents are compared and grouped, the system marks a document from each group as the "pivot." In reviewing your documents, you can review a pivot first and review the other documents in the same near duplicate set, focusing on the difference between the pivot and the document that is in review.

Email threading

Consider an email conversation that goes on for a while. In most cases, the last message in the email thread includes the contents of all the preceding messages. Therefore, reviewing the last message gives a complete context of the conversation that happened in the thread. Email threading identifies such messages so that reviewers can review a fraction of collected documents without losing any context.

Email threading in eDiscovery is the process of organizing a sequence of related emails that are part of the same conversation. This sequence includes the initial email and all subsequent replies and forwards linked to the original email. By grouping these emails into threads, reviewers see the entire context of a conversation, making it easier to understand the flow of communication. This approach helps reviewers identify relevant information more efficiently

and eliminates the need to review each email individually. Email messages included in the analytics process have the following metadata populated:

- **Is Inclusive**: This field identifies whether an email contains all the unique content from a thread, including all previous replies. It ensures that only the most comprehensive email in a thread is reviewed, which is essential for understanding the full context of the conversation without having to review each individual reply.
- **Has Unique Attachments**: This field marks emails that contain attachments not found in other emails within the same thread. Even if the email content is duplicated, unique attachments are flagged to ensure that all relevant documents are reviewed. This aspect is important in the legal review process to ensure that no unique evidence is overlooked, even if the email body itself is not unique.

How is it different from conversations in Outlook?

At a glance, this process sounds similar to conversation groupings in Outlook. However, there are some important distinctions. Consider an email conversation that forks into two conversations. For instance, someone responds to an email that isn't the latest in the conversation so the last two emails in the conversation both have unique content.

Outlook still groups the emails into a single conversation. Reading only the last email might miss the context of the second-to-last email, which also contains unique content. Because email threading parses out each email into individual components and compares them, email threading marks both of the last two emails as inclusive, ensuring that you don't miss any context as long as you read all emails marked as inclusive.

Let's also consider an email thread with multiple replies, where some replies include inline responses that modify the quoted content. If an inline reply alters part of the previous email, the latest reply doesn't fully encompass the content of the earlier email. Both the latest reply and the earlier email with unique content are marked as inclusive. This approach ensures that any unique information from the inline reply is preserved and not overlooked.

Themes

In this section, set the following parameters for themes:

- **Themes:** When turned on, the workflow performs themes clustering when you run analytics on the data in a review set.
- **Maximum number of themes:** Specifies the maximum number of themes that the workflow can generate when you run analytics on the data in a review set.

- **Include numbers in themes:** When turned on, the workflow includes numbers that identify a theme when generating themes.
- **Adjust maximum number of themes dynamically:** In certain situations, there might not be enough documents in a review set to produce the desired number of themes. When you enable this setting, eDiscovery adjusts the maximum number of themes dynamically rather than attempting to enforce the maximum number of themes.

When you create a new document, you generally start with one or more ideas that you want to convey in the document. Then you compose the document using words that align with these ideas. The more prevalent an idea is, the more frequent the words that are related to that idea tend to be. This method also aligns with how readers consume documents. The important things to understand from reading a document are the main ideas that the document is trying to convey. This understanding also includes which ideas appear where and what the relationships between the ideas are.

You can extend this process to how an eDiscovery reviewer wants to consume a set of documents in a case. They want to see which ideas are present in the review sets and which documents discuss those ideas. If they find a particular document of interest, they want to be able to see documents that discuss similar ideas.

The **Themes** functionality in eDiscovery attempts to mimic how humans reason about documents. It analyzes the *themes* that are discussed in a review set and assigns a theme to documents in the review set. In eDiscovery, Themes goes one step further and identifies the *dominant theme* in each review set and document. The *dominant theme* is the one that appears the most often in a document.

How do themes work?

The **Themes** functionality analyzes documents with text in a review set to parse out common themes that appear across all the documents in the review set. eDiscovery assigns those themes to the documents in which they appear. It also labels each theme with the words used in the documents that are representative of the theme. Because a document can contain various types of subject matter, eDiscovery often assigns multiple themes to review sets and documents. This assignment is referred to as the *Themes list*. The theme that appears most prominently in a review set or document is designated as its *dominant theme*.

Configuring Themes

Cases support themes, and they apply to all the review sets within them. You can configure the settings for themes when you create a new case or you can update the theme settings for an

existing case.

To configure themes in a case, complete the following steps:

1. Go to the Microsoft Purview portal and sign in with the credentials for a user account assigned eDiscovery permissions.
2. Select the **eDiscovery** solution card and then select **Cases** in the left nav.
3. Select a case, then select **Case settings**.
4. On **Case settings**, select **Search & analytics**.
5. Select the following theme options as applicable:

    - **Max number of themes:** Specifies the maximum number of themes that the workflow can generate when you run analytics on the data in review sets included in a case. For more information on limits, see Limits in eDiscovery.
    - **Include numbers in themes:** Numbers that identify a theme are included when generating themes.
    - **Adjust maximum number of themes dynamically:** In certain situations, there might not be enough documents in a review set to produce the desired number of themes for the case. When you enable this setting, the maximum number of themes is adjusted dynamically rather than attempting to enforce the maximum number of themes.

6. If you need to exclude keywords associated with themes, enter the text or regular expression needed in the **Ignore text** field. In the **Apply to** field, select **Themes** to apply the text or regular expression to all themes.
7. Select **Save**.

After you create a new case, the workflow automatically runs analytics on the data when you add the review sets to the case. The workflow generates themes for the review sets as part of the analytics processing.

Review set query

If you select the **Automatically create a For Review saved search after analytics** checkbox, eDiscovery autogenerates a review set query named **For Review.**

This query filters out duplicate items from the review set, so you can quickly review the unique items in the review set. You create this query only when you run analytics for a review set in the case. For more information about review set queries, see Query the data in a review set.

Ignore text

Certain text can diminish the quality of analytics, such as lengthy disclaimers that get added to email messages regardless of the content of the email. If you know of text that should be ignored, you can exclude it from analytics by specifying the text string and the analytics functionality (near-duplicates, email threading, themes, and relevance) that the text should be excluded for. The system also supports using regular expressions (RegEx) for ignored text.

Optical Character Recognition (OCR)

When you turn on this setting, OCR processing runs on image files when they're advanced indexed and added to a review set. When OCR is applied to image files, text in these files is available in review sets. This extracted text is then searchable within the review set. OCR runs only on items processed during Advanced indexing (if you select this option in the search query).

For example, if a large PDF file that is partially indexed or had other indexing errors is processed during Advanced indexing, OCR is applied. OCR processing only occurs on files that are reindexed during the Advanced indexing process. This condition means there might be situations where content is added to a review set, but some email attachments aren't processed for OCR because these files aren't processed during Advanced indexing.

After you add data to a review set, you can review, search, tag, and analyze image text. You can view the extracted text in the Text viewer of the selected image file in the review set. For more information, see:

- Add search results to a review set
- Supported image file types

The Purview eDiscovery OCR doesn't make images searchable during initial search. The case level setting for OCR is distinct from the tenant level OCR setting and doesn't incur the charges associated to that setting. To learn more about tenant level OCR, see https://learn.microsoft.com/en-us/purview/ocr-learn-about.

Last updated on 04/09/2026

Learn about review set settings in eDiscovery cases

You can configure grouping settings for each Microsoft Purview eDiscovery case to control how the data in a review set is grouped and displayed. You can turn on the **Enable group** setting only for review sets in cases created in your organization after March 15, 2023.

Depending on when the case is created, items in a review set can be grouped in two ways:

- **Group review set items by Group ID and Thread ID**: When enabled, review set items are grouped by *Group ID* and *Thread ID*. This setting is the default for all cases created after March 15, 2023.
- **Group review set items by Family ID and Conversation ID**: When disabled, review set items are grouped by *Family ID* and *Conversation ID*. This setting is the default for all cases created before March 15, 2023.

For more information about review set grouping, see Group and view documents in a review set.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Enable review set grouping settings for a case

To enable review set grouping settings for a case:

1. Go to the Microsoft Purview portal and sign in with the credentials for a user account assigned eDiscovery permissions.
2. Select the **eDiscovery** solution card, then select **Cases** in the left nav.
3. Select a case, then select **Case settings**.
4. On **Case settings**, select **Review sets**.
5. Select the **Enable group** toggle to enable grouping review set items by *Group ID* and *Thread ID*.

> **Note**
>
> For cases created before March 15, 2023, the **Enable group** toggle is disabled. Item grouping in review sets for these cases is based on *Family ID* and *Conversation ID*.

Disable review set grouping settings for a case

To disable review set grouping settings for a case:

1. Go to the Microsoft Purview portal and sign in with the credentials for a user account assigned eDiscovery permissions.
2. Select the **eDiscovery** solution card, then select **Cases** in the left nav.
3. Select a case, then select **Case settings**.
4. On **Case settings**, select **Review sets**.
5. Select the **Enable group** toggle to disable grouping review set items by *Group ID* and *Thread ID*. Item grouping in review sets for cases is now based on *Family ID* and *Conversation ID*.

Last updated on 04/01/2026

Create a search for a case in eDiscovery

Use search in eDiscovery to find in-place content such as email, documents, and instant messaging conversations in your organization that are relevant to a case. Use search to find content in these cloud-based Microsoft 365 data sources:

- Exchange Online mailboxes
- SharePoint sites
- OneDrive accounts
- Microsoft Teams
- Microsoft 365 Groups
- Viva Engage

You can create and run different searches that are associated with the case. Use conditions, such as keywords, to build search queries that return search results with the data that's most likely relevant to the case. You can also:

- View search statistics that might help you refine a search query to narrow the results.
- Preview the search results to quickly verify whether the relevant data is being found.
- Revise a query and rerun the search.

After you search for and find data that's relevant to your investigation, you can send the results to a review set for further investigation or export it for review by people outside of the investigation team.

> **Note**
>
> For organizations that have EU General Data Protection Regulation (GDPR) requirements to protect and enable individuals' privacy rights inside the European Union (EU), you can also manage investigations in response to Data Subject Requests (DSRs) submitted by a person in your organization. The User Data Search case tool is retired, and its functionality is merged with eDiscovery. You can now use search to find content to support DSRs in all locations supported by eDiscovery searches.

> **Tip**

Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Search query tips for eDiscovery cases

Use the following tips when creating and running searches in eDiscovery.

- The time zone for all searches is Coordinated Universal Time (UTC). You can't change the time zone for your organization. Time zone display settings in the search view only apply to values in the *Data* column and don't affect time stamps on collected items.

- Keyword searches are **always case-insensitive**, with no per-query or per-term opt-out. For example, **cat** and **CAT** return the same results, and `subject:"Apple"` matches both **Apple** and **apple**. This rule applies to keyword text and `property:value` syntax across all eDiscovery search surfaces (search, query-based holds, and review set search). The only case-sensitive tokens in KeyQL are the Boolean operators themselves - **AND**, **OR**, **NOT**, and **NEAR** must be uppercase to be parsed as operators.

  > ⓘ Note
  >
  > **Case-sensitive matching isn't supported in eDiscovery search.** If your scenario requires distinguishing **Apple** (the company) from **apple** (the fruit), or differentiating identifiers like **ID** from **id**, use the following pattern: (1) run a deliberately broad case-insensitive collection, (2) add the items to a review set, (3) export the review set or its text and metadata, and (4) apply case-sensitive filtering post-export by using external tooling (for example, PowerShell `Select-String -CaseSensitive`, regular expressions, or a third-party review platform).

- **An empty keyword condition returns all content in the selected data sources, not no content.** If you don't enter any keywords or conditions, the search returns every item in every selected mailbox, site, and group across the entire date range.

  > ⓘ Important
  >
  > For broad-scope cases, an empty keyword condition can return millions of items, exceed result limits, and significantly delay the search. To preserve a defined search scope, enter at least one keyword or condition before you run the search. To

intentionally collect all content in a scope (for example, for a custodial preservation snapshot), confirm the scope size by running **Generate statistics** first.

- The Boolean operators **AND**, **OR**, **NOT**, and **NEAR** must be uppercase.

  > **Important**
  >
  > Use these operators in the *KeyQL* condition field. The search doesn't support these operators in the *Keyword* condition field. The search treats these operators as literal keywords.

- Using quotes stops wild cards and any operations inside the quotes.
- A space between two keywords or two `property:value` expressions is the same as using **OR**. For example, `from:"Sara Davis" subject:reorganization` returns all messages sent by Sara Davis or messages that contain the word reorganization in the subject line. However, using a mix of spaces and **OR** conditionals in a single query might lead to unexpected results. Use either spaces or **OR** in a single query.
- Use syntax that matches the `property:value` format. Values aren't case-sensitive, and they can't have a space after the operator. If there's a space, your intended value is a full-text search. For example `to: pilarp` searches for "pilarp" as a keyword, rather than for messages sent to pilarp.
- When searching a recipient property, such as To, From, Cc, or Recipients, you can use an SMTP address, alias, or display name to denote a recipient. For example, you can use pilarp@contoso.com, pilarp, or "Pilar Pinilla."
- You can use only prefix searches; for example, **cat*** or **set***. Suffix searches (***cat**), infix searches (**c*t**), and substring searches (***cat***) aren't supported.
- When searching a property, use double quotation marks (" ") if the search value consists of multiple words. For example, `subject:budget Q1` returns messages that contain **budget** in the subject line and that contain **Q1** anywhere in the message or in any of the message properties. Using `subject:"budget Q1"` returns all messages that contain **budget Q1** anywhere in the subject line.
- To exclude content marked with a certain property value from your search results, place a minus sign (-) before the name of the property. For example, `-from:"Sara Davis"` excludes

any messages sent by Sara Davis.

- You can export items based on message type. For example, to export Skype conversations and chats in Microsoft Teams, use the syntax `kind:im`. To return only email messages, use `kind:email`. To return chats, meetings, and calls in Microsoft Teams, use `kind:microsoftteams`.

- When searching sites, you have to add the trailing `/` to the end of the URL when using the `path` property to return only items in a specified site. If you don't include the trailing `/`, items from a site with a similar path name are also returned. For example, if you use `path:sites/HelloWorld` then items from sites named `sites/HelloWorld_East` or `sites/HelloWorld_West` are also returned. To return items only from the HelloWorld site, you have to use `path:sites/HelloWorld/`.

- You can define the **Query language-country/region** in your search query before collecting content. Query language that is Chinese, Japanese, or Korean might require special treatment such as breaking a phrase or sentence into individual words or characters for more accurate test results. For more information, see Double-byte character language support in eDiscovery.

- When searching the **Sent** folders for emails, the search doesn't support using the SMTP address for the sender. Items in the **Sent** folder contain only display names.

- When searching data sources for content, optical character recognition (OCR) capabilities use tenant-level OCR scanning. Once data is collected in an eDiscovery review set, the OCR capability uses eDiscovery OCR scanning.

- The search-time *Include partially indexed items* choice affects statistics and preview only. Whether partially indexed items appear in a later export is set independently at export time. If you need them in the export, select the partial-index option on the export dialog as well. For details, see Search-time and export-time settings are independent.

Create a search query

> **Tip**
>
> Do you prefer an interactive configuration guide experience? Check out the **Design a search** guide.

When you create a new case, you're automatically directed to the **Searches** tab in the case and you're ready to create a search for the case. You can also create a new search in the Content Search case in your organization. Searches help you find the items you want to collect for the case.

SharePoint Embedded container support in eDiscovery (preview) includes container discovery and application filtering when you add a user as a data source. You can also add a supported container directly by using its URL.

1. Select **Create a search** in a case or Content Search case. If this new case doesn't have any previous searches, you can also select **Create a search** in the main pane under **Start searching for relevant data**.

2. On the **Enter details to get started** page, complete the following fields:

   - **Search name**: give the search a name (required). The search name must be unique in your organization.
   - **Search description**: Add an optional description to help others understand this search.

3. Select **Create** to create the new search and start your queries to find relevant data for the case.

4. On the **Query** tab, add data sources for the search. For complete guidance and limitations, see Data sources in eDiscovery.

   For SharePoint Embedded container support in eDiscovery (preview), select a user to discover relevant supported user-owned and tenant-owned containers, or enter a known supported container URL as a site location.

5. Select **Manage** to review the mailbox and site locations associated with the selected sources. Otherwise, select **Save and close**.

   If you added a user as a data source, use **Filter by app** as needed, and then select or clear the supported SharePoint Embedded containers relevant to the investigation.

6. Review your selections and confirm the included resources for each data source. Select **Save**. You scoped the data sources that your search queries examine.

7. In the **Data sources** section, select the ellipse menu for any data source for management options for the data source.

   For management options, select from the following options:

- **Manage data source**: Manage data sources for the selected user or site.
- **Disable mailboxes**: Disable mailboxes for the selected user or site. Select this option again to enable the mailbox for the user or site.
- **Disable sites**: Disable the site for the selected user or site. Select this option again to enable the site for the user or site.
- **Frequent collaborators**: Select associated mailboxes and sites for users that frequently collaborate with the selected user.
- **Manager**: Select associated mailboxes and sites of the user's manager.
- **Direct reports**: Select associated mailboxes and sites of the user's direct reports.
- **Groups the user owns**: Select associated mailboxes and sites of groups the user is an owner of.
- **Groups the user is in**: Select associated mailboxes and sites of groups the user is a member of.

For group sources, select from the following options:

- **Members**: Select associated mailboxes and sites of people who are members of the group.

8. Use the **Data sources** command bar controls to add, update, sync, and search for other data sources for the search as needed.

   - **Search and add**: Select the + icon to add data sources.
   - **Manage**: Select the pencil icon to manage assigned data sources.
   - **Sync**: Select the sync icon to synchronize data sources and update the data sources with the most recent data sources in your organization.
   - **Search**: Select the search icon to search the data sources currently included in the search query.

9. To define the parameters of your search query, choose from the following options on the **Query** tab:

   - **Condition builder**: The condition builder option in search provides an easy-to-user search experience when you build search queries in eDiscovery. Use keywords or custom conditions to focus the scope of your search queries. Additionally, you can use the Keyword Query Language (KeyQL) query condition option in search that provides guidance and lets you quickly paste long, complex queries directly into the editor. It also helps you build search queries from scratch and identifies potential errors and displays hints about how to resolve issues.

You can also quickly build KeyQL queries for your search using Microsoft Security Copilot. For guidance, see Create a KeyQL search query with Microsoft Copilot (preview).

10. Select **Run query**. If you want to save the query parameters you defined and run the query later, select **Save as draft**.

Create a KeyQL search query with Microsoft Copilot (preview)

The Natural Language Query (preview) KeyQL builder option in search lets you use natural language and Microsoft Security Copilot to quickly generate a Keyword Query Language (KeyQL) statement. Use the builder to construct complex queries with extra functionality, including AND, OR, and grouping of conditions, all while using natural language prompts.

This feature helps you build queries more easily by using predefined prompts for example scenarios. It also lets you refine and enhance custom prompts for more accurate search queries. You can also choose to use prompt suggestions as a starting point to create and refine KeyQL queries for common or custom search scenarios.

To create a search query with Copilot, complete the following steps:

1. After you select data sources for your query, select **Draft a query with Copilot**.
2. In the **Natural language prompt** pane, choose one of the following options:

    - Enter your search query question. You can include user, data source, and other content details as applicable.
    - Select **View prompts** to select one of the following prompt suggestions:
        - *Find all emails containing the words budget and finance and have attachments*
        - *Search all chats in the month of January 2020 that contain the word 'financial year'*
        - *Search for files of type .docx that contain the words confidential and budget*

3. Review the natural language prompt. To refine the prompt with Copilot, select **Refine**.
4. When the prompt is finalized, select **Generate KeyQL**.
5. Review the KeyQL query in the **Keyword Query Language (KeyQL) result** pane. If you need to refine the KeyQL query results, you can update the prompt in the **Natural language prompt** pane and select **Generate KeyQL** again. 1. When the KeyQL results are finalized, select **Copy KeyQL**.
6. Paste the KeyQL results in the query field on the **Keyword query language (KeyQL)** tab. You can close **Draft a query with Copilot**.

7. Select **Run query**. If you want to save the query parameters you defined and run the query later, select **Save as draft**.

Run a search query

After you create a search query manually or by using Security Copilot, run the query to generate search results.

To run a search query, complete the following steps:

1. Go to the Microsoft Purview portal and sign in with the credentials for a user account assigned eDiscovery permissions.
2. Select the **eDiscovery** solution card, and then select **Cases** in the left navigation.
3. Select a case. On the **Searches** tab, select a saved search.
4. Select **Run query**.
5. After you select **Run query**, you see the **Choose search results** flyout pane. Choose the view you want to generate for the query and its settings. You can choose either the **Statistics** or **Sample** view:

    - **Statistics**: This view generates a summary of collected data estimates arranged by top indicators. Choose one or more of the following options:
    - **Include categories**: Refine your view to include people, sensitive information types, item types, and errors.
    - **Include query keywords report**: Assess keyword relevance for different parts of your search query.
    - **Investigate partially indexed items**: Partially indexed items typically account for approximately 1% of content in data sources by count. Selecting this option (and only this option) generates summary information (item count and location) about partially indexed items included in the selected data sources for the search. No partially indexed items are reindexed or processed. For more information about how to calculate the ratio of partially indexed items in your organization, see Assess partially indexed items in your organization. To further process partially indexed items in scoped data sources, consider the following advanced indexing options:

- **Exclude partially indexed items in locations without search hits**: Choosing this option reduces the scope of partially indexed items (or advanced indexing if that option is selected) by limiting the inclusion of partially indexed items to only the data sources that include items relevant to your search. This exclusion removes partially indexed items from data sources that don't include any items relevant to your search.

  For example, you select multiple mailboxes, SharePoint sites, and OneDrive sites as data sources for your search. When the search runs, only a few of the mailboxes and sites have indexed items relevant to the search conditions. The rest of the mailboxes and sites don't contain any indexed items relevant to your search conditions. If you select this option, the partially indexed items in the mailboxes and sites that contain indexed items relevant to the search are included in the search statistics results. The partially indexed items in the mailboxes and sites that don't contain any indexed items relevant to the search are ignored and not reported in the search statistics results.

- **Perform advanced indexing on partially indexed items**: The scope of where advanced indexing runs depends on if you select the **Exclude partially indexed items in locations without search hits** option. The advanced indexing process runs a statistical sample of partially indexed items in scope and determines if these items match the query or not:
  - Selected *with* the **Exclude partially indexed items in locations without search hits** option: This setting applies to partially indexed items only to data sources with fully indexed items matching the search query. These items are sampled, indexed, and the items matching the search query are displayed in search statistics (as applicable).
  - Selected *without* the **Exclude partially indexed items in locations without search hits** option: This setting applies to all partially indexed items in all data sources included in the search. All items are sampled, indexed, and the items matching the search query are displayed in search statistics (as applicable).

  > **Important**
  >
  > This feature isn't available for limited format cases.

- **Sample**: This view generates a representative selection of the full search results. Define the parameters for the following options:

- **Select the number of sample items to generate per location**: Choose either *1*, *10*, or *100*.
- **Select the number of locations to get samples from**: Choose *10*, *100*, *1000*, or *10000*.

6. **Tenant-wide source configuration** options in eDiscovery let you include additional data sources during organization-wide searches. Adding more data sources might cause searches to take longer than usual.

   Configure the organization-wide source "All people and groups" to include one or more user or content types:

   - **"All people and groups" should include unlicensed and on-premises users**
   - **"All people and groups" should include guest users**
   - **"All people and groups" should include shared Teams channels**
   - **"All people and groups" should include departed users**

   > **ⓘ Important**
   >
   > For large organizations, tenant-wide searches across all Exchange mailboxes and SharePoint sites are long-running processes. Service throttling can affect the number of mailboxes and sites to search. Before running these tenant-wide searches, review scale limits and throughput guidance in **Limits in eDiscovery** and plan to conduct searches in batches as applicable. To improve performance and operability, break the search into scoped runs using **compliance boundaries** (for example, by region, business unit, or department) so each search targets a subset of mailboxes and sites.

7. Select **Run query** to immediately run the query.

Depending on the query view options you selected, you're automatically directed to the **Statistics** or **Sample** tab. The search query assessment starts and the time remaining to process the query is calculated. For more information about evaluating and fine-tuning your search results, see Review and evaluate search results.

Create a new search from an existing search

In some scenarios, it might be helpful to create a new search based on an existing search. This approach maintains the original data sources and search while making modifications to get new search statistics without changing the original search. When you create a duplicate search, you

keep all the data sources and the search from the original search in the new search. This approach preserves the integrity of the original search and enables you to explore new insights.

To create a new search from an existing search, complete the following steps:

1. Select a case. On the **Searches** tab, open a search.
2. Select **Duplicate**.
3. A new search is automatically created in the same case.

The new search name includes the source search title with *Copy of* prefixed to the title and the timestamp appended. For example, if the source search title is *Search for budget emails*, the new search title is *Copy of Search for budget emails + the timestamp*. To edit the new search title, select the edit control next to the title text.

Create a new search from an existing hold

In some scenarios, it might be helpful to create a new search based on an existing hold. This approach allows you to use the original data sources and search used for the hold to run a new search. When you create a search from an existing hold, you keep all the data sources and the search from the hold in the new search.

> ⓘ **Note**
>
> To create a search from an existing hold, you must be assigned the *Compliance Search* role. For information about adding users to a role group, see Assign eDiscovery permissions.

To create a new search from an existing hold, complete the following steps:

1. Select a case. On the **Hold policies** tab, select and open a hold policy.
2. Select **Create search**.
3. A new search is automatically created and you're automatically redirected to the new search.

The new search name includes the source hold policy name with *Copy of* prefixed to the title and timestamp. For example, if the source hold policy title is *Hold to preserve Contoso project*, the new search title is *Copy of Hold to preserve Contoso project + the timestamp*. To edit the new search title, select the edit control next to the title text.

Delete a search

To reduce the number of previous searches associated with a case, delete specific searches in eDiscovery. If the search has associated exports and review sets, the history of these exports and review sets remains available in the Process Manager and audit logs.

To delete an existing search in a case, complete the following steps:

1. Select a case and select the search to delete.
2. Select the ellipsis (...) in the upper right of the search page, and then select **Delete search**.
3. Select **Yes** on the **Delete search** dialog to delete the search from the case.

> **Note**
>
> You can't delete a search while processes are in progress. Wait for all processes to finish before deleting a search.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 09/19/2026

Use the condition builder to create search queries in eDiscovery

The condition builder provides an easy-to-use search experience when you build search queries in eDiscovery. Use the condition builder in search and review sets to construct simple and complex keyword queries, queries with operators (AND, OR), or both to help identify items in your organization.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Use the condition builder

To create a query and custom conditional filtering for your search, use the following controls:

- The **Keywords** condition is always available as the first condition in your query, making it easy to get started with search tasks. This condition only supports the **Equal** operator and is removed from your query using **Delete**. In the **Value** field of the Keywords condition, apply logical operators such as **AND**, **OR**, and **NEAR**. For example, entering `red AND blue` treats **AND** as a logical operator rather than a literal keyword. To search for an exact phrase like `"red and blue"`, enclose the text in quotation marks. This tells the search engine to treat the text as a single string. You can also select **Enter** on your keyboard to create a new grid for the next keyword. If you paste a list of keywords separated by line breaks (carriage returns), the system automatically generates a separate grid for each keyword.

  When multiple keyword segments are entered using the condition builder, the query can be converted to KeyQL and each keyword segment is separated by c:s. If your query includes multiple keyword conditions, the resulting KeyQL includes c:s to ensure keyword segmenting so you see hits per segment. This behavior is automatic and doesn't require manual entry of c:s in the KeyQL builder.

  > **Note**

To generate a keyword report in the **statistics** view, you must populate at least two or more keyword grids. If you enter only a single keyword, the total hit count shown reflects the results for that one keyword, and a keyword report isn't generated.

- **Add conditions**: Add a condition for the specific data sources for the search. To add extra conditions to your query, select **Add conditions** to display the list of available conditions. Each condition value selection adds a new condition to your query. Select the **AND/OR** operator as appropriate.
- **AND/OR**: These conditional logical operators let you select the query operation that applies to a specific condition. These operators let you use multiple conditions connected to your query. The operator following the *keyword* condition must be **AND**.
- **Selecting an operator**: Depending on the selected condition, the operators compatible for the condition are available to select. For example, if you select the *Date* condition, the available operators are *Before*, *After*, and *Between*. If you select the *Size (in bytes)* condition, the available operators are *Greater than*, *Greater or equal*, *Less than*, *Less or equal*, *Between*, and *Equal*.
- **Value**: Depending on the selected condition, the values compatible for the condition appear in the value details pane or you can add inline. Depending on the condition type associated with the value, you see options to define, filter, or search for values associated with the selected condition. For example, if you select *Sender* as the condition, you can search for and add specific users in your organization or external users. If you select *Size (in bytes)* as the condition, you see the option to enter a number for the size as the value. If the value is blank, the value field border appears in red to help notify you that a value is needed.
- **Delete a filter condition**: To remove an individual condition, select the delete icon to the right of each filter line.
- **Clear a filter value:** To clear the value for a specific condition, select the remove icon to the right of each filter line.
- **Save as draft**: To save the current set of conditions as a draft, select **Save as draft** above the condition builder area.
- **Discard**: To discard any changes made to the search, including conditions, and data source, select **Discard** above the condition builder area.

Guidelines for using conditions

Keep the following guidelines in mind when using search conditions:

- A condition connects logically to the keyword query (specified in the keyword condition) by the **AND** operator. That connection means that items need to satisfy both the keyword query and the condition to be included in the results.
- All conditions that follow the *keyword* condition are grouped together.
- If you add two or more unique conditions to a search query (conditions that specify different properties), those conditions connect logically by the **AND** and **OR** operators. That connection means the search returns only items that satisfy all the conditions (in addition to any keyword query).
- If you add multiple values (separated by commas or semicolons) to a single condition, those values connect by the **OR** operator. That connection means the search returns items if they contain any of the specified values for the property in the condition.
- Any condition that uses an operator with **Contains** and **Equals** logic returns similar search results for simple string searches. A simple string search is a string in the condition that doesn't include a wildcard. For example, a condition that uses **Equals any of** returns the same items as a condition that uses **Contains any of**.
- The search query that you create by using the keywords box and conditions appears on the **Search** page, in the details pane for the selected search. In a query, everything to the right of the notation `(c:c)` indicates conditions that you add to the query. Don't use `(c:c)` in manually entered queries. `(c:c)` isn't equal to **AND** or **OR**.
- If you are searching using *Compliance Tag* or *Retention Label* and the name of the label includes special characters that aren't supported (listed in the Special characters section, they are ignored as a condition of the search.

Find and select conditions

When you select **Add conditions** in the condition builder, you get quick access to some commonly used conditions such as *Date*, *Subject/Title*, and *Participants* along with conditions you used most recently. Select **Show more** and the **Choose which conditions to add** flyout pane appears to help you refine your search query with specific conditions. Use the options in the following sections to help you choose applicable conditions.

Filter conditions by area

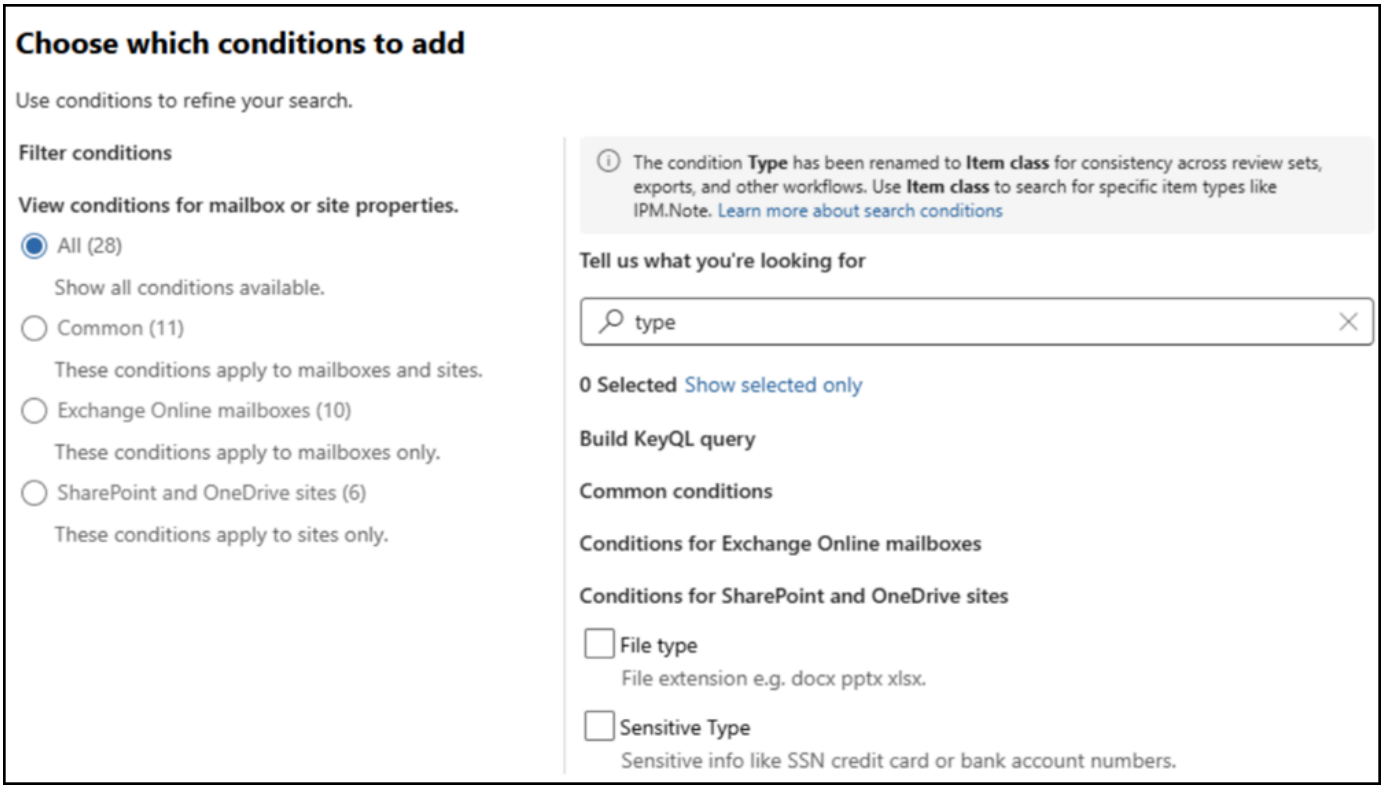
Quickly filter the condition view for mailboxes and site properties to help locate a specific condition for your search query. Filter available conditions in the following global groups:

- **All**: Shows all conditions and condition groups.

- **Common**: Filters and displays only the conditions that apply to both mailboxes and sites.
- **Exchange mailboxes**: Filters and displays only the conditions that apply to mailboxes.
- **SharePoint and OneDrive sites**: Filters and displays only the conditions that apply to SharePoint and OneDrive sites.

Condition picker

To quickly search for a specific condition, use the **Tell us what you're looking for** field to enter the name of the condition. The results automatically scope to the filter for global groups. For example, to search for any condition named *Type* (or one that contains the term *type* in the condition name), select **All** as the global filter, then enter *type* in the **Tell us what you're looking for** field. The condition view returns all conditions in all condition groups that contain the term *type*. Select the applicable condition to add to your search query.



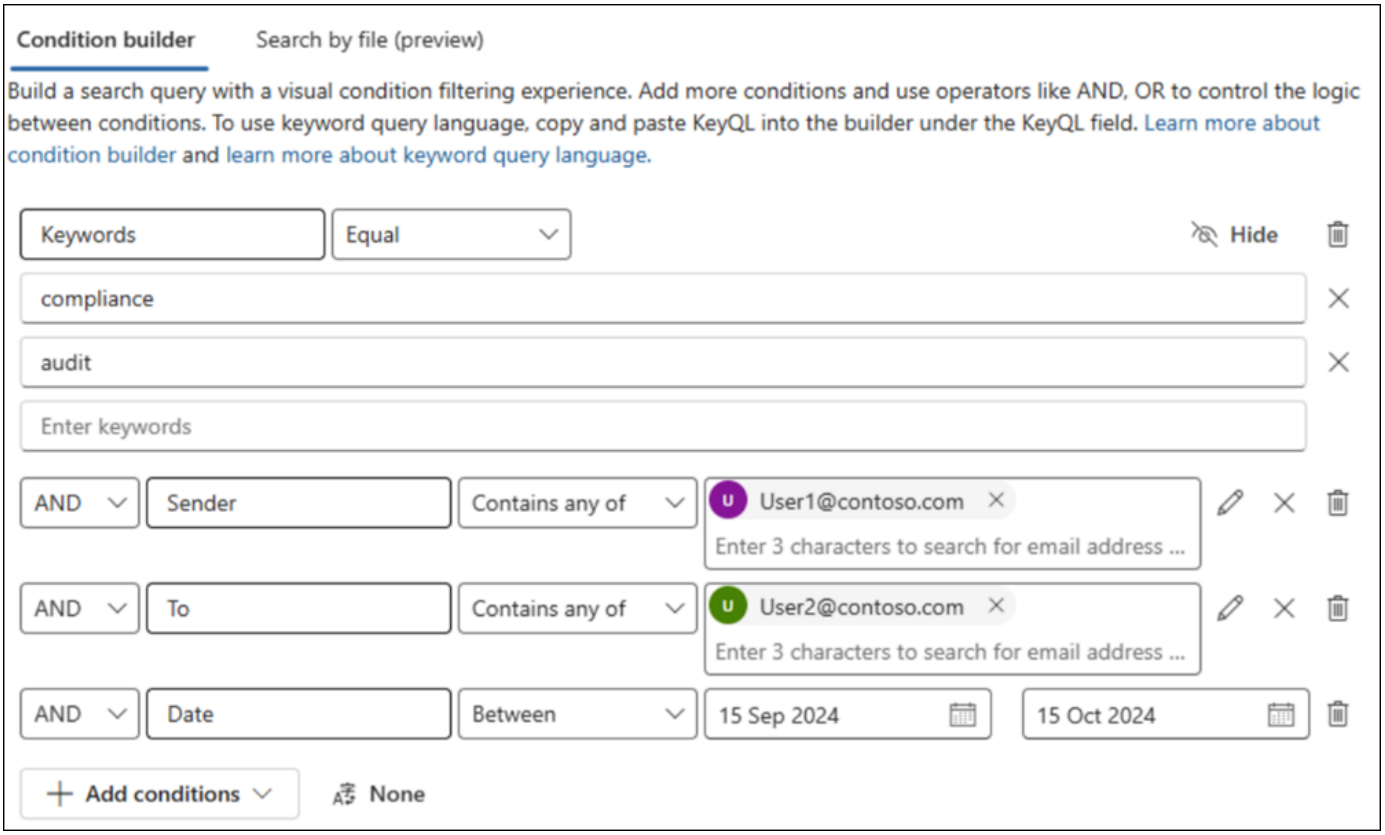
> **Important**
>
> You can only use the *Keyword* and *KeyQL* conditions once each in the condition builder.

Scenario example

The eDiscovery administrator needs to create a query to find emails sent from User1 to User2 that were sent between September 15, 2024 and October 15, 2024 that contains the keywords

compliance and *audit*. For this example, the administrator creates the following query using the new query builder:

1. For the first filter, the administrator uses the *Keywords* condition, the *Equal* operator, and *compliance, audit* as the keyword *Value*.
2. Next, the administrator selects **Add conditions**, selects *Sender*, then selects the *Contains any of* operator, then enter *User1@contoso.com* in the **Value** field. This can include external users.
3. Next, the administrator selects **Add conditions**, selects the *To* filter, then selects the *Contains any of* operator, then selects *User2@contoso.com* in the **Value** field. This can include external users.
4. To define the date range, the administrator selects **Add conditions**, selects *Date*, then selects the *Between* operator, and then selects the starting and ending dates for the *Value*.
5. Finally, the administrator selects **Run query** to return results matching the conditions and their values.



Use search conditions

You can add conditions to a search query to narrow a search and return a more refined set of results. Each condition adds a clause to the KeyQL search query that is created and run when you start the search.

- Special characters

- Conditions for common properties
- Conditions for mail properties
- Conditions for document properties
- Operators used with conditions

Special characters

Some special characters aren't included in the search index and therefore aren't searchable. This limitation also applies to the special characters that represent search operators in the search query. Here's a list of special characters that are either replaced by a blank space in the actual search query or cause a search error.

```
+ - = : ! @ # % ^ & , ; _ / ? ( ) [ ] { }
```

Conditions for common properties

Create a condition using common properties when searching mailboxes and sites in the same search. The following table lists the available properties to use when adding a condition.

Expand table

Condition	Description
Content kind[1]	Applied to both Exchange and SharePoint items, it refers to the type or category of the content. For example, *ContentKind:SharePointDocument*, *ContentKind:Copilot*, etc.
Content source application[1]	Identifies the application or service where the content originated. For example, *ContentSourceApplication:OneDriveForBusiness*, *ContentSourceApplication:SharePoint*, etc.
Date	For email, the date a message was created or imported from a PST file. For documents, the date a document was last modified. For Teams messages, the original message date. Item reports and review set metadata might display different timestamps (such as ingestion or modification dates) for the same Teams message. The date condition filters on the original event date. If you're searching for email messages for a specific time period, use the message *Received* and *Sent* conditions if you're unsure if the email messages might be imported instead of natively created in Exchange.
Identifier[1]	For email, the ID for a specific message. *Message IDs* are included in the audit record, data loss prevention (DLP) alerts, or review set metadata and allow you build a specific search for an individual message.

Condition	Description
	For Microsoft Teams messages, the ID of the chat or reaction. The *ChatThreadID* is included in the audit record, data loss prevention (DLP) alerts, or review set metadata and allow you build a specific search for an individual chat or reaction.
Sender/Author	For email, the person who sent a message. For documents, the person cited in the author field from Office documents. You can type more than one name, separated by commas. Two or more values are logically connected by the **OR** operator. (See Recipient expansion.)
Size (in bytes)	For both email and documents, the size of the item (in bytes).
Subject/Title	For email, the text in the subject line of a message. For documents, the title of the document. The Title property is metadata specified in Microsoft Office documents. You can type the name of more than one subject/title values, separated by commas. Two or more values are logically connected by the **OR** operator. **Note**: Don't include double quotation marks to the values for this condition because quotation marks are automatically added when using this search condition. If you add quotation marks to the value, two pairs of double quotations are added to the condition value, and the search query returns an error.
Retention label	For both email and documents, retention labels applied to messages and documents. Retention labels can be used to declare records and help you manage the data lifecycle of content by enforcing retention and deletion rules specified by the label. **Note**: If using Retention Label names that include unsupported special characters, they will be ignored as a condition for the search. For more information about retention labels, see Retention policies and retention labels.

Conditions for mail properties

Create a condition using mail properties when searching mailboxes or public folders in Exchange Online. The following table lists the email properties that you can use for a condition. These properties are a subset of the email properties listed in Conditions for common properties. These descriptions are repeated for your convenience.

Expand table

Condition	Description
Message kind	The message type to search. This property is the same as the Kind email property. Possible values: • contacts • docs • email • externaldata

Condition	Description
	• fax • im • journals • meetings • microsoftteams • notes • posts • rssfeeds • tasks • voicemail
Participants	All the people fields in an email message. These fields are From, To, Cc, and Bcc. (See Recipient expansion.)
Received	The date that an email message was received by a recipient. This property is the same as the Received email property.
Recipients	All recipient fields in an email message. These fields are To, Cc, and Bcc. For more information, see Recipient expansion in mailbox searches.
Sender	The sender of an email message.
Sent	The date that an email message was sent by the sender. This property is the same as the Sent email property.
Subject	The text in the subject line of an email message. **Note**: Don't include double quotation marks to the values for this condition because quotation marks are automatically added when using this search condition. If you add quotation marks to the value, two pairs of double quotations are added to the condition value, and the search query returns an error.
To	The recipient of an email message in the To field.
Topic[1]	Summary of the main subject or theme discussed in an email thread or conversation.
Item class	The message class property for an email item. It's a multivalue condition. To select multiple message classes, hold the **CTRL** key and select two or more message classes in the drop-down list that you want to add to the condition. Each message class selected in the list are logically connected by the **OR** operator in the corresponding search query. For a list of the message classes (and their corresponding message class ID) used by Exchange and that you can select in the **Message class** list, see Item Types and Message Classes.

Conditions for document properties

Create a condition using document properties when searching for documents on SharePoint and OneDrive sites. The following table lists the document properties that you can use for a condition. These properties are a subset of the site properties listed in Conditions for common properties. These descriptions are repeated for your convenience.

Expand table

Condition	Description
Author	The author field from Office documents.
Created	The date that a document is created.
File type	The extension of a file; for example, docx, one, pptx, or xlsx. This is the same property as the FileExtension site property. **Note:** If you include a File type condition using the **Equals** or **Equals any of** operator in a search query, you can't use a prefix search (by including the wildcard character (*) at the end of the file type) to return all versions of a file type. If you do, the wildcard is ignored. For example if you include the condition `Equals any of doc*`, only files with an extension of `.doc` are returned. Files with an extension of `.docx` aren't returned. To return all versions of a file type, use the *property:value* pair in a keyword query; for example, `filetype:doc*`.
Last modified	The date that a document was last changed.
Path[1]	The URL or location of a file or folder within a SharePoint site.
Sensitive information type (SIT)[1]	Sensitive information types included in documents. SITs are pattern-based classifiers and they detect sensitive information like social security, credit card, or bank account numbers to identify sensitive items. For more information about SITs, see Sensitive information types.
Sensitivity label[1]	Sensitivity labels applied to documents. Sensitivity labels let you classify and protect your organization's data, while making sure that user productivity and their ability to collaborate isn't hindered. For more information about sensitivity labels, see Sensitivity labels.
Title	The title of the document. The Title property is metadata that's specified in Office documents. It's different from the file name of the document.

Operators used with conditions

When you add a condition, select an operator that matches the type of property for the condition. The following table describes the operators that work with conditions and lists the equivalent operator used in the search query.

Expand table

Operator	Query equivalent	Description
After	`property>date`	Use with date conditions. Returns items that were sent, received, or modified after the specified date.
Before	`property<date`	Use with date conditions. Returns items that were sent, received, or modified before the specified date.
Between	`date..date`	Use with date and size conditions. When used with a date condition, returns items that were sent, received, or modified within the specified date range. When used with a size condition, returns items whose size is within the specified range.
Contains any of	`(property:value) OR (property:value)`	Use with conditions for properties that specify a string value. Returns items that contain any part of one or more specified string values.
Doesn't contain any of	`-property:value` `NOT property:value`	Use with conditions for properties that specify a string value. Returns items that don't contain any part of the specified string value.
Doesn't equal any of	`-property=value` `NOT property=value`	Use with conditions for properties that specify a string value. Returns items that don't contain the specific string.
Equals[2]	`size=value`	Returns items that are equal to the specified size.
Equals any of	`(property=value) OR (property=value)`	Use with conditions for properties that specify a string value. Returns items that match one or more specified string values.
Greater[2]	`size>value`	Returns items where the specified property is greater than the specified value.
Greater or equal[2]	`size>=value`	Returns items where the specified property is greater than or equal to the specified value.
Less[2]	`size<value`	Returns items that are less than the specified value.
Less or equal[2]	`size<=value`	Returns items that are less than or equal to the specified value.
Not equal[2]	`size<>value`	Returns items that don't equal the specified size.

Note

[1] This operator is an eDiscovery premium feature condition.

[2] This operator is available only for conditions that use the *Size* property.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/19/2026

Use Keyword Query Language to create search queries in eDiscovery

09/04/2025

The Keyword Query Language (KeyQL) query condition option provides feedback and guidance when you build search and hold policy queries in eDiscovery. When you enter queries in the editor, it autocompletes supported searchable properties and conditions. It also provides lists of supported values for standard properties and conditions. For example, if you specify the `kind` email property in your query, the editor presents a list of supported values that you can select.

The KeyQL editor also displays potential query errors in real time that you can fix before you run the search. Paste complex queries directly into the editor without manually building queries using the conditions cards in the condition builder. The KeyQL editor is also available when you create query-based holds in eDiscovery.

Here are the key benefits to using the KeyQL editor:

- Provides guidance and helps you build search queries from scratch.
- Lets you quickly paste long, complex queries directly into the editor. For example, if you receive a complex query from opposing counsel, you can paste that into the KeyQL editor instead of having to use the condition builder.
- Quickly identifies potential errors and displays hints about how to resolve issues.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

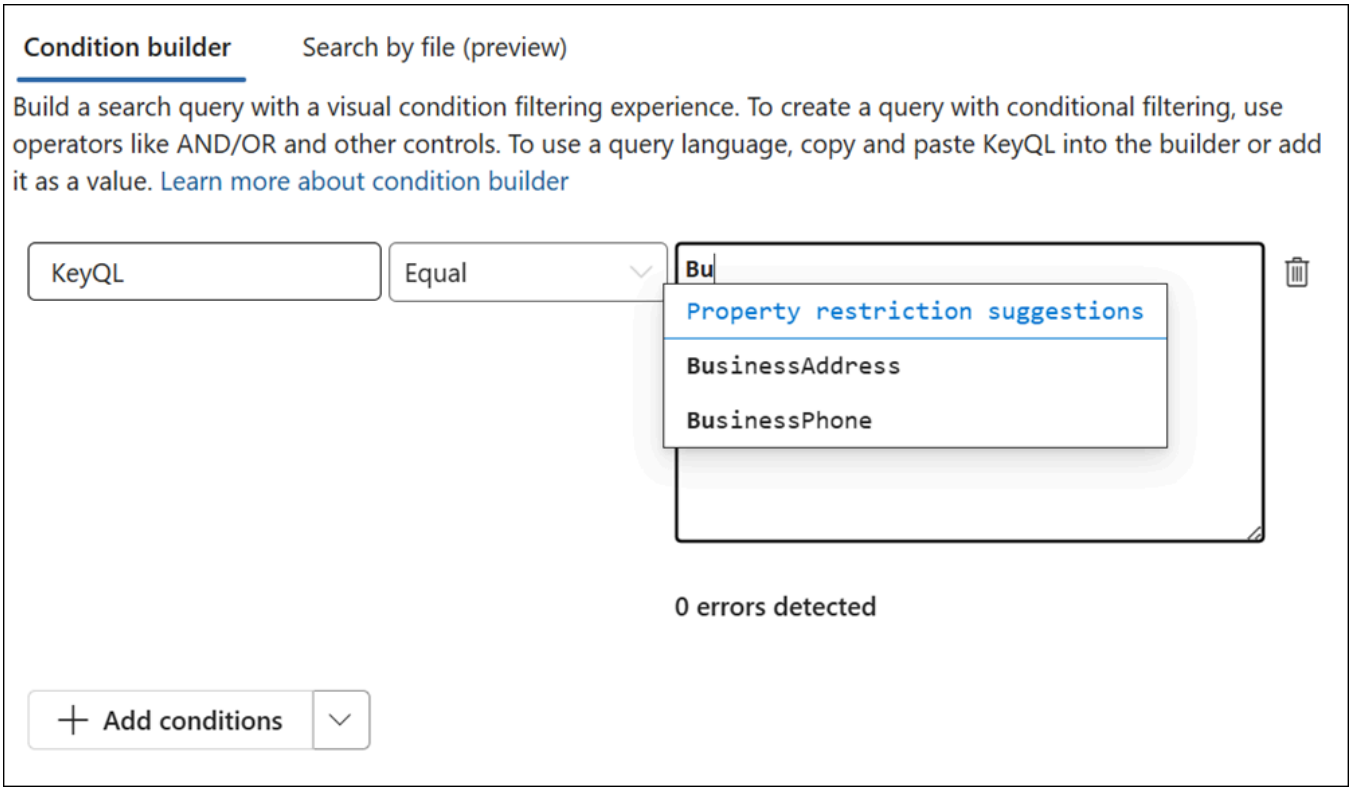
Using the KeyQL editor

Use the KeyQL editor by selecting **Add conditions** and selecting *KeyQL*. For conditions and values already populated, these conditions and values are automatically converted into KeyQL. KeyQL query conditions can't be used with other conditions and need to be used as standalone conditions in your query.

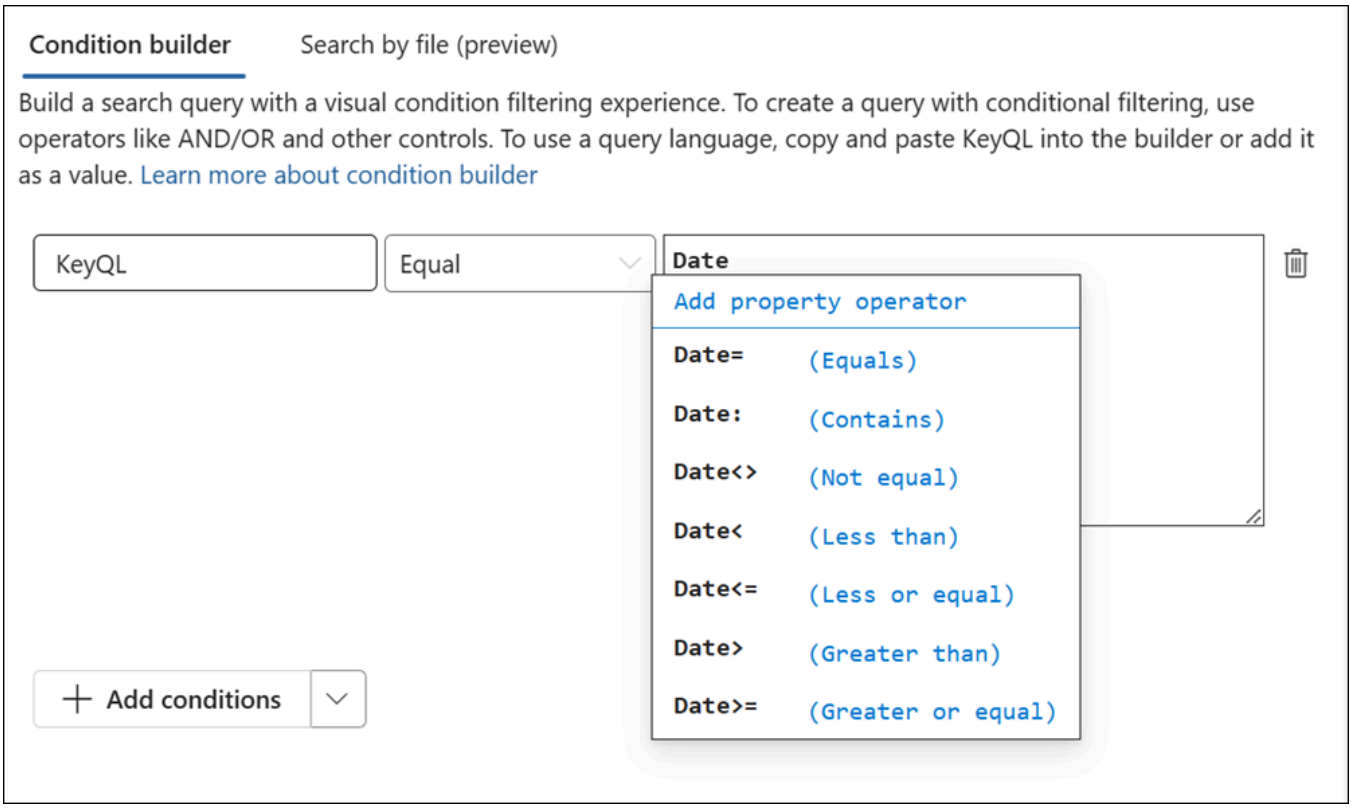
The following sections show examples of how the KeyQL editor provides suggestions and detects potential errors.

Autocompletion of search properties and operators

When you start to enter a search query in the KeyQL editor, the editor displays suggested autocompletion of supported search properties (also called *property restrictions*) that you can select. You need to type at least two characters to display a list of supported properties that begin with those two characters. For example, the following screenshot shows the suggested search properties that begin with `Bu`.



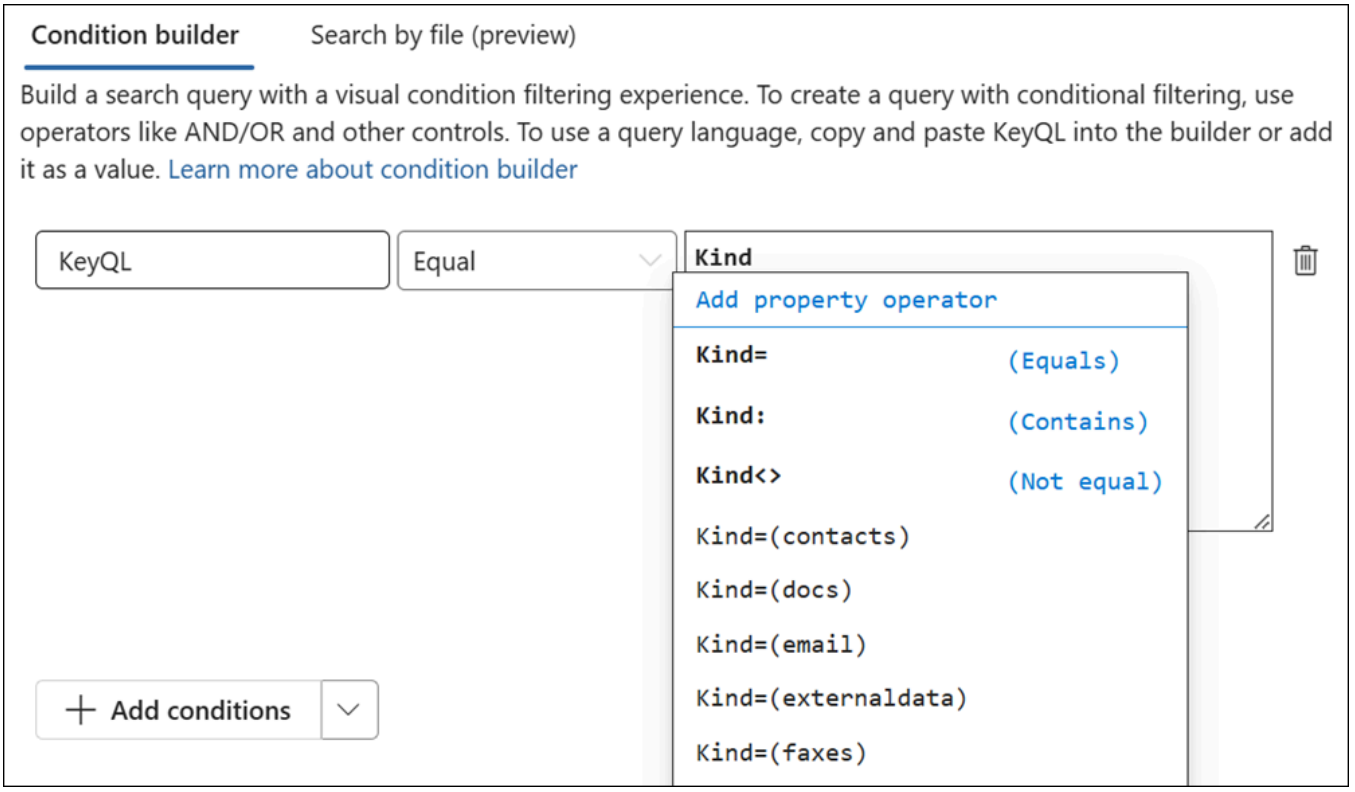
Additionally, the editor suggests a list of supported operators (such as `:`, `=`, and `<>`) when you type in a complete property name. For example, the following screenshot shows the suggested operators for the `Date` property.



For more information about the supported search conditions, see Use the condition builder to create search queries in eDiscovery.

Property value suggestions

The KeyQL editor provides suggestions for possible values of some properties. For example, the following screenshot shows the suggested values for the `Kind` property.



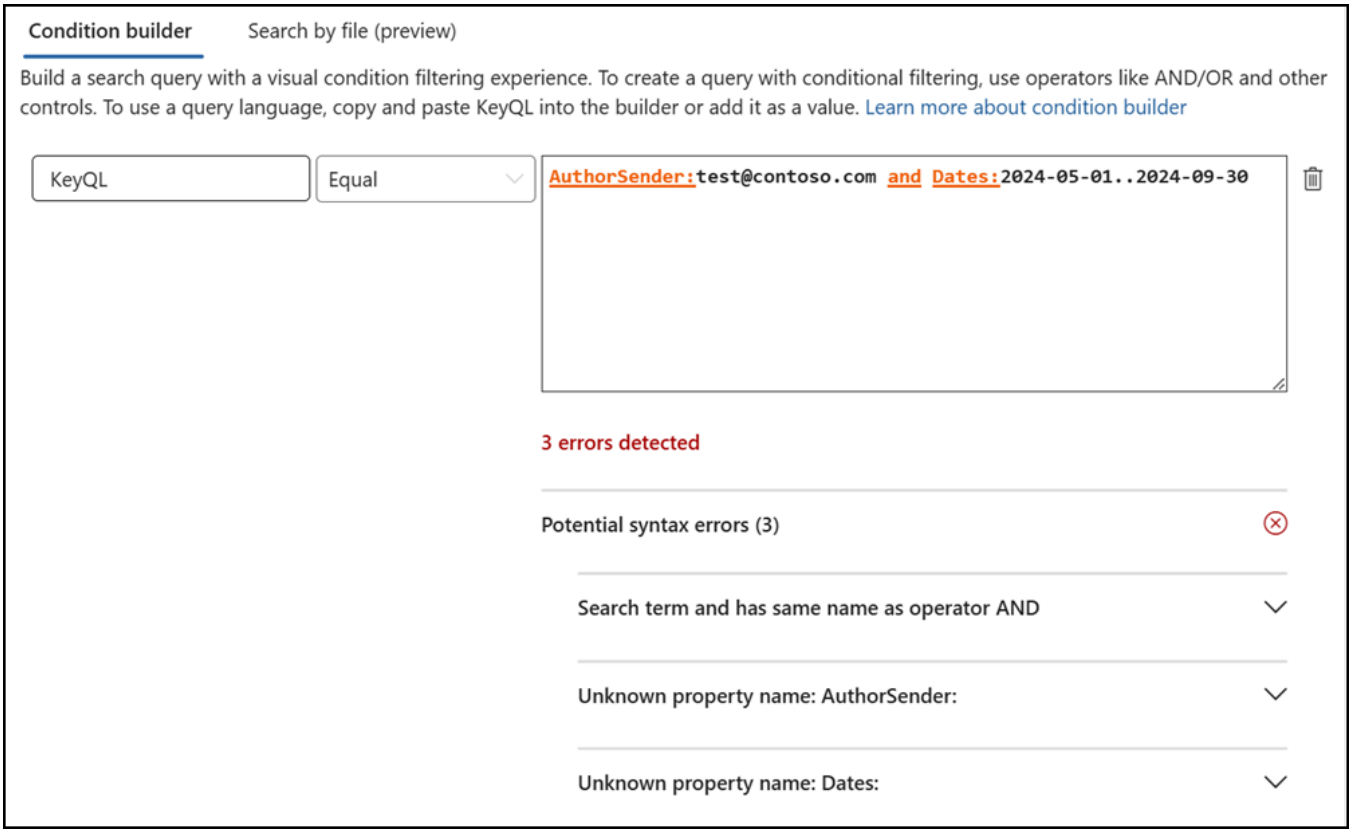
The editor also suggests a list of users (in User Principal Name (UPN) format) when you type email recipient properties, such as `From`, `To`, `Recipients`, and `Participants`.

Detection of potential errors

The KeyQL editor detects potential errors in search queries, and provides a hint of what is causing the error to help you resolve the error. The editor also indicates a potential error when a property doesn't have a corresponding operation or value. Potential errors in the query are highlighted in red text, and explanations and possible fixes for the error are displayed in the **Potential errors** drop-down section.

> **Important**
>
> Nested quotation marks aren't supported in the KeyQL editor.

For example, if you paste the following query into the KeyQL editor, the editor detects three potential errors. Use the potential error hints to help troubleshoot and fix the query.



Data sources in eDiscovery

In Microsoft 365, data is stored across three platforms: Exchange, Microsoft Teams, and SharePoint. These platforms organize and manage data within Microsoft 365 applications. Most Microsoft 365 apps store data in one or more of the following containers:

- **Users**: Data associated with individual users, such as their mail, 1:1 Teams messages, and OneDrive files.
- **Groups**: Data owned by the organization or a group of users within an organization. These groups are often referred to as Unified Groups or Teams.

In Microsoft Purview eDiscovery, the concept of *data source* streamlines the process of identifying and managing data across Microsoft 365 platforms. eDiscovery users select a user or group, which creates a data source. eDiscovery automatically identifies and organizes relevant data stored across platforms. The data source gathers locations related to the user or group (mailboxes, OneDrive sites, SharePoint sites) and adds the locations in the data source hierarchy. eDiscovery users refine the scope by selecting or excluding specific locations as needed.

A user data source typically includes:

- User mailbox
- OneDrive site

SharePoint Embedded application content

SharePoint Embedded container support in eDiscovery (preview) might change before general availability. Review the limitations in this section before using the feature in an investigation.

Some Microsoft 365 applications store content in SharePoint Embedded containers. Examples include Microsoft Loop, Microsoft Copilot Pages, Microsoft Copilot Notebooks, Microsoft Designer, and Outlook Newsletters. The applications represented in the Manage sources experience depend on the supported containers returned for the selected user.

When you add a user as an eDiscovery data source, eDiscovery identifies supported SharePoint Embedded containers relevant to that user and displays them with the user's other data locations. You can review and select the individual containers that are relevant to the investigation.

Add SharePoint Embedded containers as data sources

You can add supported SharePoint Embedded containers as data sources in the following ways:

- **Discover containers for a user:** Add a user as a data source to display supported user-owned and tenant-owned SharePoint Embedded containers relevant to that user. You can select or clear individual containers before saving the data source.
- **Filter containers by application:** When reviewing containers for a selected user, use **Filter by app** to narrow the displayed containers to a specific supported application. The available filters depend on the applications represented in the returned container list.
- **Add a container by URL:** If you know the SharePoint Embedded container URL, enter it when adding a site location. eDiscovery resolves the container URL so you can select the container as a data source.
- **Add container URLs in bulk:** Include supported SharePoint Embedded container URLs when you bulk import data sources. eDiscovery resolves the URLs as data-source locations.

For each supported SharePoint Embedded container, the **Manage sources** pane displays:

- The full container URL
- The container display name
- The associated application label
- The `ContainerTypeId`

Use these details to distinguish between containers and select the locations relevant to the investigation.

> ⓘ **Note**
>
> Microsoft Loop My workspace, Copilot Pages, and Copilot Notebooks use the same user-owned SharePoint Embedded container. In SharePoint Embedded administration and compliance experiences, this shared container is identified as **Loop**. As a result, Copilot Pages and Copilot Notebooks content can appear in a container labeled **Loop** and might not appear as separate application filters. For more information, see Overview of Copilot Pages and Copilot Notebooks storage.

> ⓘ **Important**
>
> **SharePoint Embedded limitations:**
>
> - **Container ownership:** Supported user-owned and tenant-owned SharePoint Embedded containers relevant to a selected user are surfaced in eDiscovery. Group-

owned and app-owned SharePoint Embedded containers aren't currently supported.
- **Compliance boundaries:** eDiscovery compliance boundary enforcement isn't currently supported for SharePoint Embedded containers.

Group and organization-wide data sources

Unified groups are classified into three types, each covering specific data locations:

- **Teams**: Includes Exchange mailbox storage for Teams chats and emails, as well as all associated SharePoint sites for channels and the team.
- **Viva Engage**: Includes Exchange mailbox storage for Viva Engage messages.
- **Classic**: Includes only SharePoint sites.

eDiscovery users can also use organization-wide sources to perform searches across your organization. Organization-wide sources include:

- **All people and groups**: Includes all users and all groups in your organization.
- **All public folders**: Includes all content in Exchange public folders mailboxes.

> **Note**
>
> Organization-wide searches include content from supported SharePoint Embedded containers. However, selecting an organization-wide source doesn't display a tenant-wide list of every SharePoint Embedded container. To review and select individual supported containers, use the user-based **Manage sources** experience.

> **Important**
>
> For large organizations, tenant-wide searches across all Exchange mailboxes and SharePoint sites are long-running processes. Service throttling can affect these searches because of the number of mailboxes and sites to search. Before running these tenant-wide searches, review scale limits and throughput guidance in Limits in eDiscovery. Plan to conduct searches in batches as applicable. To improve performance and operability, break the search into scoped runs by using compliance boundaries (for example, by region, business unit, or department) so each search targets a subset of mailboxes and sites.

It's common to search a set of user mailboxes when you manage the list as a distribution list. Adding distribution lists is supported, and the search includes only the Exchange mailboxes listed

in the group.

You can search for specific data sources or data locations by using user or group's names, mailbox Simple Mail Transfer Protocol (SMTP) addresses, and OneDrive or SharePoint site URLs. When you create a search by using specific data sources, you search only the locations specified in the data source.

If you use the organization-wide source **All people and groups**, the search covers all Exchange mailboxes, OneDrive, and SharePoint sites. To search all Exchange mailboxes only, select mailboxes under **All people and groups** and deselect the sites. To search all SharePoint and OneDrive sites only, select sites under **All people and groups** and deselect the mailboxes.

Real-time data source sync helps ensure that you're always informed about the latest changes in data locations associated with users and groups. You can query if any specific data sources are added to a search, if a hold includes newly provisioned data locations, or if data locations are removed.

For example, if you create a private channel for a Teams group, the sync feature on the data source panel alerts you of the new location. You can quickly and easily include it in searches or holds. This sync ensures that new data doesn't go unnoticed and is included in your investigations. This sync also helps prevent potential data loss from location changes.

For more information about using data sources effectively, check out the following video:

https://learn-video.azurefd.net/vod/player?id=c8081650-7272-4e22-8c7d-52c03ca69bcb&locale=en-us&embedUrl=%2Fpurview%2Fedisc-data-sources

Supported group types

The following table summarizes which group types can be added as data sources in eDiscovery, and how membership is handled for hold policies vs. searches.

Expand table

Group type	Supported as data source	Hold behavior	Search behavior
Distribution list (static)	Yes	Members snapshotted at hold creation. Re-add the group to reflect membership changes.	Members re-resolved each time the search runs.

Group type	Supported as data source	Hold behavior	Search behavior
Exchange Dynamic Distribution Group (DDG)	No	Not supported as a hold data source.	Not supported as a search data source.
Mail-enabled security group	Yes	Same as distribution list.	Same as distribution list.
Microsoft 365 group	Yes	Group mailbox and associated SharePoint site snapshotted at hold creation.	Group mailbox and site re-resolved each time the search runs.
Microsoft Entra ID dynamic Microsoft 365 group	No	Same limitation as Entra dynamic security groups.	Same limitation as Entra dynamic security groups.
Microsoft Entra ID dynamic security group	No	Entra dynamic security groups (groups with a `membershipRule`) aren't supported.	Entra dynamic security groups aren't supported.
Microsoft Teams group	Yes	Group mailbox plus the associated SharePoint site snapshotted at hold creation. Channel sites resolved at hold creation.	Group mailbox and site re-resolved each time the search runs.
Viva Engage group	Yes	Group mailbox snapshotted at hold creation.	Group mailbox re-resolved each time the search runs.

> **Note**
>
> If you need a scope that tracks dynamic membership criteria, create a static mail-enabled security group, synchronize its membership from the dynamic group on a schedule (for example, by using a scheduled Microsoft Graph and Exchange Online PowerShell job), and add the static shadow group as the data source. For holds, re-add the shadow group periodically to reflect updates. For searches, membership is re-resolved at each run.

Add data sources to cases

> **Important**
>
> Data sources in the legacy eDiscovery experience don't synchronize to the same case in the new eDiscovery experience. You must add the data sources in the case in the new eDiscovery

experience in the Microsoft Purview portal again.

After you enable premium eDiscovery features for a case, use the following options to add data sources to the case:

> **Tip**
>
> Want to try premium eDiscovery features? See the **subscription requirements** for Microsoft 365 Enterprise E5 licensing.

Bulk import data sources to a case

In an eDiscovery case, you can add multiple data sources by importing a list of SMTP addresses, SharePoint site URLs, or supported SharePoint Embedded container URLs. The import feature validates the entries before adding the sources, so you can correct any potential problems, such as typos. This streamlined process helps you save time and add many data sources to your cases more efficiently.

Before you import data sources, review the following considerations:

- You can import up to 500 data sources at a time. To minimize the processing time for searching and adding locations, split the import into multiple imports instead of importing the maximum number of data sources at one time.
- You must use the `;` delimiter in your list of SMTP addresses or URLs. Your list should follow the formatting of the following example:

  ```
  abh784@contoso.ms;https://contosodemos2.sharepoint.com/sites/Logistics;https://contosodemos2.sharepoint.com/sites/Mark8Project-UETConference;https://contosodemos2.sharepoint.com/sites/WG-ProjectObsidian;maib31@contoso.ms;rma230@contoso.ms;sva561@contoso.ms;mmz776@contoso.ms;LogisticsDSIGoldenSetDataToloka1@contoso.ms;jka132@contoso.ms;shafar929@contoso.ms;ika636@contoso.ms;brobet@contoso.ms;koka81@contoso.ms;rash95@contoso.ms;mvga66@contoso.ms;mra715@contoso.ms;abe543@contoso.ms;kto888@contoso.ms;moba80@contoso.ms;mian41@contoso.ms;inch48@contoso.ms;Mario.Smith_0803@contoso.ms;yape61@contoso.ms;mwo365@contoso.ms;fasi19@contoso.ms;HumanResourcesDugoPlannerTest@contoso.ms;kamo78@contoso.ms;bga844@contoso.ms;salesteam@contoso.ms;cko475@contoso.ms;ProjectZ@contoso.ms;TalentSourcingDanieltest5@contoso.ms;las450@contoso.ms;elfl69@contoso.ms;jch952@contoso.ms;asch81@contoso.ms;gmu955@contoso.ms;okbo68@contoso.ms
  ```

- Distribution groups aren't expanded to individual members during the import process.
- To import an inactive mailbox as a data source, add a `.` prefix to the email address of the inactive mailbox. For example, `.sarad@contoso.onmicrosoft.com`.
- Unverified SMTP addresses aren't added, but unverified URLs are supported. An unverified URL means the data source can't be verified against a valid SharePoint site that is in a locked state.

To bulk import data sources to a case, complete the following steps:

1. Go to the Microsoft Purview portal and sign in by using the credentials for a user account assigned eDiscovery permissions.
2. Select the **eDiscovery** solution card, and then select **Cases** in the left navigation.
3. Select a case, and then select the **Data sources** tab to add data sources to the case. When you add sources to this tab, you make these data sources available to choose from in all searches and holds in the case.
4. Select **Bulk import**.
5. On **Add data sources**, enter or copy a list of SMTP addresses or URLs separated by a semicolon.
6. Select **Next**.
7. On **Confirm data sources**, view the list of data sources and the verification status.
8. After verification of the data sources finishes, unverified sources are highlighted in red and marked as *Unverified* for you to review.
    - To update unverified sources, select the check box next to the data source and select **Edit** to update the data source information. You can correct any typo, path, or format errors. After editing the selected source, select **Save** and the updated data sources are automatically reverified.
    - To export the list of data sources, select **Export list**. This action creates a .csv file that contains user inputs, matched sources, and the status of the sources.
    - Any input with multiple matches is shown with *(Duplicate)* in front. You can include a subset of duplicated matched sources based on your needs.
    - To remove any data sources from the import, select the data source and select **Delete**.

9. When you're ready to add any of the listed data sources, select the check box for each item you want to add and then select **Add**.

> ⓘ **Note**
>
> Depending on the number of sources selected, it might take some time to process and add the data sources to the case.

Add one or more data sources to a case

1. Go to the Microsoft Purview portal and sign in by using the credentials for a user account assigned eDiscovery permissions.

2. Select the **eDiscovery** solution card, and then select **Cases** in the left navigation.

3. Select a case, and then select the **Data sources** tab to add data sources to the case. When you add sources to this tab, all searches and holds in the case can choose these data sources.

4. Select **Add** and pick choices from the following data source areas:

> ⓘ **Note**
>
> When searching for sources, some sources show up with a *?* in front of the source and labeled as *Unverified*. You can't verify these sources (sites or mailboxes) as valid sources when selected. These mailboxes or sites are external or the site is private or locked. You can add unverified sources to searches or holds and the system tries verification for these sources again when the search runs. Final search and hold results are included in the *location* .csv file in the process report.

   a. The left side of the pane displays the **Filter** options for sources.

   b. In the **Scope items by** section, only **All sources in the tenant (default)** is available. Select any data source available in your organization.

   c. Use one of the following options in the **Show for** filter to help scope your sources in the **Search** section:

      - *All people and groups (default)*
      - *People only*

- *Groups only*

d. If applicable, select **Exclude inactive users** to reduce the scope of sources to only currently active users in your organization.

e. After you filter data sources, use the search control and selectors in the **Search** section to add specific data sources, users, and groups to the search query. Enter the specific users, groups, or organization locations you want to add in the search field and select **Search**. You can use ";" as delimiter to trigger multiple searches at once. For example, input *john@contoso.com;david@contoso.com* allows you to search and find matches to both mailboxes at once. The `;` delimiter doesn't support site URLs.

   To add a supported SharePoint Embedded container directly, enter its full container URL as a site location.

   Search for people using the following values:

   - First and family name of the user display name (for example, John Smith)
   - First name only
   - User SMTP address
   - User alias
   - Exchange GUID
   - URL of the user's OneDrive site

   Search for groups using the following values:

   - Group mailbox SMTP address
   - URL of group site. The URL of a Teams channel site resolves the Teams group as a data source.

     If you add a distribution group, the list of group members isn't listed and the group is added to the search as a data source mailbox. When the search runs, the distribution group member mailboxes are expanded and fully searched.

     To confirm the mailboxes searched for the distribution group members, use the Locations_*the date/time of the report* information in the Process report after the search is completed. To confirm group membership before running the search, select the ellipse menu for the group and *Members*.

5. Select **Add** to add the data source to the selected case. This data source is now available when adding data sources for searches and holds in the case.

> **Important**
>
> You can manage individual data resources for a data source only when you add the case-level data source to an individual search and hold in the case. For example, to include or exclude mailboxes or sites in a search or hold for a data source configured at the case level, add the case-level data source to the search or hold, and then select **Manage**.

Add data sources to searches and holds

Add data sources to a specific search or hold by following these steps:

1. In a new search or hold, select **Add sources**. Or, select + and choose **Add data sources** from the dropdown. You can search for and select specific data sources to search or hold against.

   If you need to perform the search against organizational-wide mailboxes or sites, select **Add tenant-side sources in the empty search** or select the + button and choose an organizational-wide source from the dropdown.

   > **Note**
   >
   > Organizational-wide sources are only available in searches, not holds.

2. The left side of the pane displays the **Filter** options for sources. Use filters to scope the data sources by:

   - **All sources in the tenant (default):** Use this option to search from data sources available in your organization.
   - **All sources in this case**: Use this option to choose from data sources added at the case level. By using this option, you can quickly use data sources added to the case in a search or hold without having to search across your entire organization.
   - Use one of the following options in the **Show for** filter to help scope your sources in the **Search** section:
     - *All people and groups (default)*
     - *People only*
     - *Groups only*

- If applicable, select **Exclude inactive users** to reduce the scope of sources to only currently active users.

- Use **Locations to include control** to specify if the selected data sources added to the search or hold include:
  - *Mailboxes and sites (default)*: Selected people or group sources include the mailbox and site to the search or hold. This selection means selecting the user includes both the mailbox and OneDrive for the user. Selecting a Microsoft 365 group includes the group mailbox and all associated group sites.
  - *Mailboxes only*: Only mailbox associated with the select user or group is included. OneDrive and SharePoint sites aren't included.
  - *Sites only*: Only sites associated with the select user or group are included. Mailboxes aren't included.

3. After you filter the data sources, use the search control and selectors in the **Search** section to add specific data sources, users, and groups to the search query. Enter the specific users, groups, or organization locations you want to add in the search field and select **Search**. Use a semicolon as delimiter for multiple searches at once. For example, *john@contoso.com;david@contoso.com* searches and finds matches for both mailboxes at once. The semicolon delimiter doesn't support site URLs.

   To add a supported SharePoint Embedded container directly, enter its full container URL as a site location.

4. Select **Manage** to review the locations associated with the selected sources.

   If you added a user as a data source, use **Filter by app** as needed, and then select or clear the supported SharePoint Embedded containers relevant to the investigation.

5. Select **Save and close** to add the data source to the current search or hold.

Other considerations

- In the data source picker, you might not find some users or groups. In the **Manage** sources view, some selected sources might not display associated sites. This problem can happen for several reasons:
  - **OneDrive not provisioned, provisioning delay, or sync issues**: If a user doesn't have OneDrive provisioned, their source includes only a mailbox and not a OneDrive site. The OneDrive site exists, but it isn't synced in the compliance directory due to provisioning latency.

- **Deleted or departed user**: If you remove a user from the directory or soft-delete or deactivate their account, their OneDrive site might still exist but isn't linked to their user object. It doesn't appear under their source in the **Manage** view. For email content, consider using an inactive mailbox if the mailbox was placed on hold before deletion. For OneDrive content from a deleted user, add the OneDrive site URL directly as a SharePoint location in the search or hold.
- **Unlicensed user**: The user's Microsoft 365 license was removed, which breaks the link to their OneDrive site and can deactivate their mailbox. Reassign the required license and wait for the service to provision the data source before adding it again.
- **Mailbox converted**: The user's mailbox was converted to a shared mailbox or a mail user. This change might affect how the system associates OneDrive.
- **Multiple site collection admins**: If the OneDrive site has multiple admins and ownership is unclear, the system might not map it to the original user.
- **Retention or hold-only state**: The user has an inactive mailbox or hold-only state, and OneDrive might be cleaned up per retention policies.
- **Unverified source**: When searching for data sources, some mailboxes or sites might appear with a **?** icon and the label *Unverified*. This status means the system couldn't confirm the source as valid at selection time – often because the mailbox or site is external, private, or locked. You can still add unverified sources to searches or holds. Verification is attempted again when the search runs, and the final status is reflected in the **location.csv** file in the **process report**.
- **Ambiguous identity**: Duplicate entries exist in the directory, such as two users with similar names or an old account with the same alias. Work with your Microsoft Entra ID or Exchange admin to identify and remove duplicate objects so each email address corresponds to a single unique mailbox.
- **Unresolved recipient**: The user can't be found in the directory. For newly created accounts, allow up to 24 hours for directory synchronization to complete. Verify in the Exchange admin center that the mailbox exists and is active.

If a user shows *Not available* after adding, remove and re-add the user after the underlying account issue is resolved. You can verify data source availability by checking the Microsoft 365 admin center for the user's account status. To check for duplicate objects, run `Get-Recipient -Filter "EmailAddresses -eq 'user@domain.com'"` in Exchange Online PowerShell.

> **Note**

When sites show as **Not available**, manually add the URLs to the search or hold data sources.

- When you add a distribution group, the list of group members isn't displayed. The distribution group is added to the search as a data source mailbox. When you run the search, the system expands and fully searches the distribution group member mailboxes. To confirm the mailboxes searched for the distribution group members, use the `Locations_*the date/time of the report*` information in the **Process report** after the search completes. To confirm group membership before running the search, select the ellipsis menu for the group and members.

  If the membership of the distribution group changes after adding it as a data source in the search, run the search again to include items for the current members of the group. Each time you run the search, current members are included and former members are excluded.

Inactive mailboxes

You can search, review, and export inactive mailboxes stored in your organization in eDiscovery. To hold inactive mailbox content, first apply the hold policy to the active mailbox, verify success, and then change the mailbox to inactive.

When using inactive mailboxes in eDiscovery searches and holds, consider the following points:

- Search inactive mailboxes in an eDiscovery search. For a list of the inactive mailboxes in your organization, run `Get-Mailbox -InactiveMailboxOnly` in Exchange Online PowerShell. Alternatively, go to **Data lifecycle management** > **Microsoft 365** > **Retention** in the Microsoft Purview portal and select **More** (the navigation bar ellipses).

- If an existing search includes a user mailbox and you update that mailbox to inactive, the search continues to search the inactive mailbox when you rerun the search.

- Some users might have an active mailbox and an inactive mailbox that have different SMTP addresses. In this case, only the specific mailbox you select as a location for a search is searched. If you add a user's mailbox to a search, you can't assume that both active and inactive mailboxes are searched. Only the mailbox you explicitly add to the search is searched.

- Use Security & Compliance PowerShell to create a search for an inactive mailbox. You must append a period to the email address of the inactive mailbox. For example, the following

command creates a content search that searches an inactive mailbox with the email address pavelb@contoso.onmicrosoft.com:

```
New-ComplianceSearch -Name InactiveMailboxSearch -ExchangeLocation
.pavelb@contoso.onmicrosoft.com -AllowNotFoundExchangeLocationsEnabled $true
```

- Avoid having an active mailbox and inactive mailbox with the same SMTP address. If you need to reuse the SMTP address assigned to an inactive mailbox, recover the inactive mailbox or restore the contents of an inactive mailbox to an active mailbox (or the archive of an active mailbox), and then delete the inactive mailbox. For more information, see the following articles:
  - Recover an inactive mailbox in Office 365
  - Restore an inactive mailbox in Office 365
  - Delete an inactive mailbox in Office 365

- Hold policies aren't supported for inactive mailboxes. You must place holds on active mailboxes before the mailbox is deleted. For more information, see Create and manage inactive mailboxes.

Data source options

After you add sources to a search or hold, you can edit the sources or explore related sources as needed. Select the *ellipsis* next to the source you want to edit or remove.

The following options are available in **User or group options**:

- Manage data source
- Include or remove a mailbox
- Include or remove sites

Explore related sources

You can explore and add related sources from existing sources in a search or hold. These options give you the ability to investigate potentially relevant sources and bring in more sources to a specific search or hold.

For users you add in the **Data Sources** pane, you can explore the following connections:

- Manager
- Direct reports
- Frequent collaborators

- Groups the user owns
- Groups the user is a member of

By exploring these relationships, you can quickly identify and include relevant individuals and data sources in a search scope. Use *Manage* to fine-tune whether to include only mailboxes, only sites, or both related sources in a search or hold.

Manager

Use this option to explore upward in the reporting chain. When you select the user's manager, you include supervisory or decision-making context in your investigation.

Direct reports

Use this option to explore downward in the organization hierarchy for the user. This option is useful when the selected user is a manager or team lead, and you want to include their team members' communications or content.

Frequent collaborators

Use this option to find other users that frequently collaborate with the selected user. *Frequent collaborators* are the top 10 users who are most relevant to the selected user. You can select the mailboxes and sites for these users as data sources for searches.

Groups the user owns

This option shows groups that the selected user is listed as an owner. These groups might contain shared content or communications relevant to the investigation.

Groups the user is a member of

This option includes all groups the user is a member of, even if they're not the owner. These groups might provide additional context or shared data sources.

Group members

When you add a group in the **Data Sources** pane, you can explore and expand the group to view its members. By using this feature, you can:

- Identify individual users within the group that might hold relevant data.
- Select specific members as additional data sources for targeted searches.

This feature helps when investigating shared content or communications originating from collaborative groups such as Microsoft 365 Groups, Teams, or distribution lists.

> **Note**
>
> The list of group members and groups that a user is in (or owns) is capped at 100. If a user is included in more than 100 groups, owns more than 100 groups, or if a group has more than 100 members, the portal displays only the first 100 members.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 09/22/2026

Finding content in mailboxes in eDiscovery

Admins often need to find out who knew what and when. They need to respond to requests about ongoing or potential litigation, internal investigations, and other scenarios in the most efficient and effective way possible. These requests are often urgent, involve multiple stakeholder teams, and have significant impact if not completed in a timely manner. Knowing how to find the right information is critical for admins to complete searches successfully and help their organizations manage the risk and cost associated with eDiscovery requirements.

For more information about finding content in mailboxes in eDiscovery, watch the following video:

https://learn-video.azurefd.net/vod/player?id=6976ca23-3877-4cba-b052-d509d842f896&locale=en-us&embedUrl=%2Fpurview%2Fedisc-search-mailboxes

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about **Microsoft Security Copilot in Microsoft Purview**.

When an eDiscovery request is submitted, admins often receive only partial information to start collecting content that might be related to a particular investigation. The request might include user names, project titles, rough date ranges when the project was active, and not much more. From this information, admins need to create queries to find relevant content across Microsoft 365 services to determine the information needed for a particular project or subject. Understanding how information is stored and managed for these services helps admins more efficiently find what they need quickly and in an effective manner.

Exchange Online stores email, chat, meeting, and Microsoft 365 Copilot and Microsoft 365 Copilot Chat activity data (user prompts and Copilot responses). Many communication properties are available for searching items included in Exchange Online. Some properties such as *From*, *Sent*, *Subject*, and *To* are unique to certain items and aren't relevant when searching for files or documents in SharePoint and OneDrive. Including these types of properties when searching across workloads can sometimes lead to unexpected results.

For example, to find content related to specific users (*User 1* and *User 2*), associated with a project called *Tradewinds*, and during January 2020 through January 2022, you might use a query with the following properties:

- Add User 1 and User 2's Exchange Online locations as data sources to the case
- Select User 1 and User 2's Exchange Online locations as data source.
- For **Keyword**, use *Tradewinds*
- For **Date Range**, use the *January 1, 2020* to *January 31, 2022* range

When searching for emails or other mailbox content that might include cloud attachments, the cloud attachment link to a storage location in SharePoint or OneDrive should be considered as attachments of the message file rather than separate documents.

This behavior is similar to how a search returns emails with embedded local attachments, these attachments are retrieved as if they're part of the emails. Any compliance boundaries for SharePoint or OneDrive aren't reflected in the collection of cloud attachments included in responsive messages.

> ⓘ **Important**
>
> For emails, when you use a keyword, the search includes the subject, body, and many properties related to the participants. However, due to recipient expansion, search might not return expected results when using the alias or part of the alias. Therefore, use the full UPN.

Searchable email properties in KeyQL

> ⓘ **Important**
>
> While email messages might have other properties supported in other Microsoft 365 services, eDiscovery search tools support only the email properties listed in this table. You can't include other email message properties in searches.

The following table lists the email message properties supported in search using the eDiscovery KeyQL editor in the Microsoft Purview portal or by using the New-ComplianceSearch or Set-ComplianceSearch cmdlets. For a list of supported condition builder properties, see Use the condition builder to create search queries in eDiscovery.

The table includes an example of the *property:value* syntax for each property and a description of the search results returned by the examples. You can enter these `property:value` pairs in the keywords box for an eDiscovery search.

> ⓘ **Note**

When searching email properties, you can't search for message headers. Header information isn't indexed for searches. Additionally, you can't search for items in which the specified property is empty or blank. For example, using the *property:value* pair of **subject:""** to search for email messages with an empty subject line returns zero results. This limitation also applies when searching site and contact properties.

Expand table

Property	Property description	Examples	Search results returned by the examples
AttachmentNames	The names of files attached to an email message.	`attachmentnames:annualreport.ppt` `attachmentnames:annual*`	Messages that have an attached file named *annualreport.ppt*. In the second example, using the wildcard character (*) returns messages with the word *annual* in the file name of an attachment.[1]

Bcc	The Bcc field of an email message.[1]	`bcc:pilarp@contoso.com`
		`bcc:pilarp`
		`bcc:"Pilar Pinilla"`
			All examples return messages with *Pilar Pinilla* included in the Bcc field. (See Recipient Expansion)

Category	The categories to search. Users can define categories by using Outlook or Outlook on the web (formerly known as Outlook Web App). The possible values are: • blue • green • orange • purple • red • yellow	`category:"Red Category"`	Messages that are assigned the *red* category in the source mailboxes.
Cc	The Cc field of an email message.[1]	`cc:pilarp@contoso.com` `cc:"Pilar Pinilla"`	In both examples, messages with *Pilar Pinilla* specified in the Cc field.

Property	Property description	Examples	Search results returned by the examples
			(See Recipient Expansion)
From	The sender of an email message.[1]	`from:pilarp@contoso.com`	Messages sent by the specified user. (See Recipient Expansion)
HasAttachment	Indicates whether a message has an attachment. Use the values **true** or **false**.	`from:pilar@contoso.com AND hasattachment:true`	Messages sent by the specified user that have attachments.

Importance	The importance of an email message, which a sender can specify when sending a message. By default, messages are sent with normal importance, unless the sender sets the importance as **high** or **low**.	`importance:high`
		
		`importance:medium`
		
		`importance:low`
			Messages that are marked as high importance, medium importance, or low importance.

IsRead	Indicates whether messages are read. Use the values **true** or **false**.	`isread:true`
		
		`isread:false`
			The first example returns messages with the IsRead property set to **True**. The second example returns messages with the IsRead property set to **False**.

ItemClass	Use this property to search specific third-party data types that your organization imported to Microsoft 365. Use the following syntax for this property: `itemclass:ipm.externaldata.<third-party data type>*`	`itemclass:ipm.externaldata.Facebook* AND subject:contoso`
		
		`itemclass:ipm.externaldata.Twitter* AND from:"Ann Beebe" AND "Northwind Traders"`
			The first example returns Facebook items that contain the word "contoso" in the Subject property. The second example returns Twitter items that were posted by Ann Beebe and that contain the keyword phrase "Northwind Traders".

Kind	The type of email message to search for. Possible values:
	contacts
		`kind:email`
		
		`kind:email OR kind:im OR kind:voicemail`
			The first example returns email messages that meet the search criteria. The second example

Property	Property description	Examples	Search results returned by the examples
	docs email externaldata faxes im journals meetings microsoftteams (returns items from chats, meetings, and calls in Microsoft Teams) notes posts rssfeeds tasks voicemail	`kind:externaldata`	returns email messages, instant messaging conversations (including Skype for Business conversations and chats in Microsoft Teams), and voice messages that meet the search criteria. The third example returns items that were imported to mailboxes in Microsoft 365 from third-party data sources, such as Twitter, Facebook, and Cisco Jabber tha meet the search criteria. For more information, see Archiving third-party data.
Participants	All the people fields in an email message. These fields are From, To, Cc, and Bcc.[1]	`participants:garthf@contoso.com` `participants:contoso.com`	Messages sent by or sent to garthf@contoso.com The second example returns all messages sent by or sent to a user in the contoso.com domain. (See Recipient Expansion)
Received	The date that an email message is received by a recipient.	`received:2021-04-15` `received>=2021-01-01 AND received<=2021-03-31`	Messages that are received on April 15, 2021. The second example returns all messages received between January 1, 2021 and March 31, 2021.

Property	Property description	Examples	Search results returned by the examples

Recipients	All recipient fields in an email message. These fields are To, Cc, and Bcc.[1]	`recipients:garthf@contoso.com`
		
		`recipients:contoso.com`
			Messages sent to garthf@contoso.com The second example returns messages sent to any recipient in the contoso.com domain. (See Recipient Expansion)

Sent	The date that an email message is sent by the sender.	`sent:2021-07-01`
		
		`sent>=2021-06-01 AND sent<=2021-07-01`
			Messages that are sent on the specified date or sent within the specified date range.

Size	The size of an item, in bytes.	`size>26214400`
		
		`size:1..1048567`
			Messages larger than 25 MB. The second example returns messages from 1 through 1,048,567 bytes (1 MB) in size.

Subject	The text in the subject line of an email message. **Note:** When you use the Subject property in a query, the search returns all messages in which the subject line contains the text you're searching for. In other words, the query doesn't return only those messages that have an exact match. For example, if you search for `subject:"Quarterly Financials"`, your results include messages with the subject "Quarterly Financials 2018".	`subject:"Quarterly Financials"`
		
		`subject:northwind`
			Messages that contain the phrase "Quarterly Financials" anywhere in the text of the subject line. The second example returns all messages that contain the word northwind in the subject line.

To	The To field of an email message.[1]	`to:annb@contoso.com`
		
		`to:annb`
		`to:"Ann Beebe"`
			All examples return messages where Ann Beebe is specified in the To: line.

> **Note**
>
> [1] For the value of a recipient property, you can use email address (also called user principal name (UPN)), display name, or alias to specify a user. For example, you can use annb@contoso.com, annb, or "Ann Beebe" to specify the user Ann Beebe.

Searchable sensitive data types

You can use eDiscovery search tools in the Microsoft Purview portal to search for sensitive data, such as credit card numbers or social security numbers, stored in documents in mailboxes. To search for this data, use the `SensitiveType` property and the name or ID of a sensitive information type in a keyword query. For example, the query `SensitiveType:"Credit Card Number"` returns documents that contain a credit card number. The query `SensitiveType:"U.S. Social Security Number (SSN)"` returns documents that contain a U.S. social security number.

To see a list of sensitive information types that you can search for, go to **Data classifications** > **Sensitive info types** in the Microsoft Purview portal. Or, use the **Get-DlpSensitiveInformationType** cmdlet in Security & Compliance PowerShell to display a list of sensitive information types.

`SensitiveType` behavior depends on the workload and how the content is indexed. Use the following matrix to set expectations before running a query:

Expand table

Workload	`SensitiveType` supported?	When the match is evaluated	Common caveats
Exchange Online (mail, calendar, contacts)	Yes	At indexing time, against the message body and indexed attachments	Sensitive types added or modified *after* an item was indexed don't match that item until it's reindexed. Inline images, scanned PDFs, and other non-text content aren't evaluated.
Items that are partially indexed	Not evaluated	N/A	`SensitiveType` queries don't return partially indexed items. To include them, also export partially indexed items separately.
Microsoft Teams (chat and channel messages)	Limited	At indexing time, against message text only	Match accuracy is lower for short messages because most sensitive info types require corroborating context. Attachments in Teams messages follow the SharePoint/OneDrive rules.

Workload	`SensitiveType` supported?	When the match is evaluated	Common caveats
SharePoint Online / OneDrive for Business (documents)	Yes	At indexing time, against the document content	Same reindexing caveat as Exchange. Files larger than the indexing size limit are partially indexed and might not match.

If you expect a hit but don't get one, confirm that (1) the workload is in the supported column, (2) the item was indexed *after* the sensitive type was last published, and (3) the item isn't in the partially indexed set.

Recipient expansion

Mailboxes are flexible storage, and clients connecting to a mailbox control some aspects of recipient information, especially for the sender. Clients can choose the *SMTP*, *Name*, or *LegacyDN* properties as the sender address. To compensate for variations in client behavior and how data is stored, eDiscovery search's recipient expansion is a useful feature.

Often, sender information created by some clients only stores the name of the sender, such as *John Doe*, without including the SMTP address like *johndoe@contoso.com*. Additionally, federated items with legacy systems might only store the LegacyExchangeDN, which is a unique identifier used in older versions of Exchange to represent a mailbox or a distribution list. Federated systems refer to messages or data that are integrated from different systems or organizations, often involving legacy systems that use older formats or identifiers.

The recipient expansion function solves the issue of not searching broadly enough by expanding the search to capture content stored with these variations. Querying Microsoft Entra ID expands any values specified in the participant filter. This expansion includes the user's email address, UPN, alias, display name, and LegacyExchangeDN. This expansion ensures that searches cast a wider net, capturing all relevant content regardless of how participant information is stored and improves the accuracy and comprehensiveness of eDiscovery searches.

> **Note**
>
> Recipient expansion doesn't resolve cases where the user's Microsoft Entra ID object is deleted - typically a departed user whose mailbox was deleted without first being placed on hold. Departed users whose mailboxes are placed on hold remain in Microsoft Entra ID as inactive mailboxes, and recipient expansion still works for them (subject to the SMTP and display name conditions described later in this section). When the Microsoft Entra ID object is gone, you must manually include all known identifiers (for example, previous SMTP

addresses, aliases, or display names) in your query. To preserve recipient expansion for departing employees, place the mailbox on hold before deleting the account.

When you search any of the recipient properties (From, To, Cc, Bcc, Participants, and Recipients), Microsoft 365 attempts to expand the identity of each user by looking them up in Microsoft Entra ID. If the user is found in Microsoft Entra ID, the query is expanded to include the user's email address (or UPN), alias, display name, and LegacyExchangeDN. For example, a query such as `participants:ronnie@contoso.com` expands to `participants:ronnie@contoso.com OR participants:ronnie OR participants:"Ronald Nelson" OR participants:"<LegacyExchangeDN>"`.

To prevent recipient expansion, add a wildcard character (asterisk) to the end of the email address and use a reduced domain name; for example, `participants:"ronnie@contoco*"` Be sure to surround the email address with double quotation marks.

Preventing recipient expansion in the search query might result in relevant items not being returned in the search results. Email messages in Exchange can be saved with different text formats in the recipient fields. Recipient expansion is intended to help mitigate this fact by returning messages that might contain different text formats. So preventing recipient expansion might result in the search query not returning all items that might be relevant to your investigation.

Recipient expansion isn't designed to support scenarios involving user name and alias changes. If the SMTP/UPN for a user changes, Microsoft Entra ID might not find the user, leading to incomplete search results. Also, LegacyExchangeDN, which rarely changes, might not be present in the substrate for all email items. Recipient expansion only handles cases where the client uses LegacyExchangeDN instead of SMTP. If the SMTP address changes but the LegacyExchangeDN doesn't, then recipient expansion doesn't help and you need to manually find and use all variations of these addresses. This limitation can lead to users mistakenly believing that recipient expansion catches all variations, including name and SMTP changes.

When you use the *From* condition with the *equal* condition (=), only the display name is queryable in the index. This means that if you use an SMTP address in the *From* or *Sender* conditions, search returns a result if both of the following conditions are met:

- The SMTP address used in the condition matches an active or inactive user in Microsoft Entra.
- The display name of the active or inactive user hasn't changed since the targeted item was sent or received.

When you use the *Sender* or *From* conditions in an eDiscovery search, the system attempts recipient expansion and expands the query to include *Display Name*. If successful, the query

includes all email addresses assigned to the user's SMTP address and legacy Exchange display name. If recipient expansion isn't successful, the search uses only the value supplied in the *From* and *Sender* conditions to search the index. Conditions that might result in recipient expansion failing are:

- Departed users without a retained inactive mailbox.
- Active or inactive users where the SMTP address is no longer present on an object.
- Active or inactive users where the display name has been changed (when display name is used with the *Sender* or *From* conditions).

In some scenarios, recipient expansion might also cause hits on additional items when organization searches are used. For example, one user's display name might form a part of another user's display name. For example, a user might have a display name of *John Doe* and another user might have a display name of *John Doe Jr*. An organization-wide search might return result hits from both mailboxes. Consider suppressing recipient expansion in this scenario by adding a period at the end of the SMTP address.

Additional considerations include, the recipient expansion only supports sending from your own mailbox and no delegated or send-as activity. Make sure you review relevant limits, including the accuracy of the maximum number of distribution group members and the maximum level of nesting for distribution groups. Security groups are supported, and only distribution groups and the distribution group should be valid when the email is sent.

> ⓘ **Note**
>
> If you need to review or reduce the items returned by a search query due to recipient expansion, consider using premium eDiscovery features. You can search for messages (taking advantage of recipient expansion), add them to a review set, and then use review set queries or filters to review or narrow the results.

Content stored in Exchange Online mailboxes for eDiscovery

You primarily use a mailbox in Exchange Online to store email-related items such as messages, calendar items, tasks, and notes. But that's changing as more cloud-based apps also store their data in a user's mailbox. One advantage of storing data in a mailbox is that you can use the search tools in eDiscovery to find, view, and export the data from these cloud-based apps.

The data from some of these apps is stored in hidden folders located in a non-interpersonal message (non-IPM) subtree in the mailbox. Data from other cloud-based apps might not be

stored *in* the mailbox, but it's *associated with* the mailbox, and is returned in searches if that data matches the search query. Regardless of whether cloud-based data is stored in or associated with a user mailbox, the data is typically not visible in an email client when a user opens their mailbox.

The following table lists the apps that either store or associate data with a cloud-based mailbox. The table also describes the type of content that each app produces.

Expand table

Microsoft 365 app	Description
Class Schedule	Plans you create in Class Schedule are stored in the mailbox of the corresponding Microsoft 365 Group that's provisioned when you create a new plan. The alias for the group mailbox is the name of the plan.
Forms*	Forms and responses to a form are stored in files that are attached to email messages and stored in a hidden folder in the mailbox of the user who created the form. Forms created before April 2020 are stored as a PDF file. Forms created after 2020 are stored as a JSON file. Responses to a form are stored in a CSV file. When you export content from Forms in a PST file, this data is located in the **ApplicationDataRoot** folder in a subfolder named with the following globally unique identified (GUID): **c9a559d2-7aab-4f13-a6ed-e7e9c52aec87**.
Microsoft 365 Copilot and Microsoft 365 Copilot Chat	All Copilot activity data (user prompts and Copilot responses) generated in supported Microsoft 365 apps and services is stored in custodian mailboxes.
Microsoft 365 Groups	Email messages, calendar items, contacts (People), notes, and tasks are stored in the mailbox that's associated with a Microsoft 365 group.
Outlook/Exchange Online	Email messages, calendar items, contacts (People), notes, and tasks are stored in a user's mailbox.
People	Contacts in the People app (which are the same contacts as the ones accessible in Outlook) are stored in a user's mailbox.
Skype for Business	Conversations in Skype for Business are stored in the Conversation History folder in a user's mailbox. If the mailbox of a participant of a Skype meeting is placed on Litigation Hold or assigned to a retention policy, files attached to a meeting are retained in the participants mailbox.
Sway*	Sways are stored as an HTML file that is attached to an email message and stored in a hidden folder in the mailbox of the user who created the sway. When you export content from Sway in a PST file, this data is located in the **ApplicationDataRoot** folder in a subfolder named with the following GUID: **905fcf26-4eb7-48a0-9ff0-8dcc7194b5ba**.
Tasks	Tasks in the Tasks app (which are the same tasks as the ones accessible in Outlook) are stored in a user's mailbox.

Microsoft 365 app	Description
Teams	Conversations that are part of a Teams channel are associated with the Teams mailbox. Conversations that are part of the Chat list in Teams (also called *1 x N chats*) are associated with the mailbox of the users who participate in the chat. Also, summary information for meetings and calls in a Teams channel are associated with mailboxes of users who dialed into the meeting or call. So when searching for Teams content, you search the Teams mailbox for content in channel conversations and search user mailboxes for content in 1 x N chats.
To-Do	Tasks (called *to-dos*, which are saved in to-do lists) in the To-Do app are stored in a user's mailbox.
Viva Engage	Conversations and comments within a Viva Engage community are associated with the Microsoft 365 group mailbox, as well as the user mailbox of the author and any named recipients (@ mentioned or Cc'ed users). Private messages sent outside of a Viva Engage community are stored in the mailbox of the users who participate in the private message.

> **Note**
>
> * At this time, if you place a hold on a mailbox using holds in eDiscovery cases, the hold doesn't preserve content from this app.

Note: The author created this article with assistance from AI. Learn more

Last updated on 06/11/2026

Find and delete email messages in eDiscovery

> **Tip**
>
> You can also use the purge feature in Data Security Investigations to search for and permanently delete Exchange mailbox and Microsoft Teams items in your organization. Data Security Investigations provides a built-in search and purge workflow with a purge queue dashboard to help you manage and track purge queries. For more information, see **Purge in Data Security Investigations**.

You can use the search feature to search for and delete email messages from all mailboxes in your organization. This process can help you find and remove potentially harmful or high-risk email, such as:

- Messages that contain dangerous attachments or viruses
- Phishing messages
- Messages that contain sensitive data

However, the eDiscovery purge process isn't intended for general mailbox management tasks such as:

- Reducing mailbox quota
- Clearing out mailboxes
- Enforcing data retention policies

The purge process described in this article is specifically designed to address **data spillage incidents**, such as when a small message is accidentally shared with a large group of users. For routine mailbox management and compliance, use the following tools instead:

- **Retention Policies**: To manage the lifecycle of data and ensure compliance with organizational standards.
- **Archive Policies**: To offload older content and manage mailbox storage effectively.
- **Mailbox Ownership and Manual Deletion**: For precise control over specific mailboxes, especially when content needs to be reviewed or removed by the mailbox owner.

> **Tip**
>
> If your organization has a Defender for Office 365 Plan 2 subscription, use the procedure detailed in Remediate malicious email delivered in Office 365, rather than following the procedure described in this article.

Before you begin

> **Important**
>
> Carefully review the following guidance before starting the search and purge steps for email messages.

- Depending on the eDiscovery subscription for your organization, cases can either be enabled for eDiscovery premium features, if the organization has E5 licenses, or not. Verify the level of feature support for the case to determine if you should use PowerShell or Microsoft Graph to search and purge.
- For cases not configured for eDiscovery premium features, you can only use PowerShell to search and delete email messages. For E3 customers, cases created by cmdlets or the Microsoft Purview portal can only purge email by using PowerShell. You can purge up to 10 items per location.
- For cases configured for eDiscovery premium features, you can use PowerShell or Microsoft Graph to search and delete email messages. Cases configured by using Graph or the Microsoft Purview portal with premium features can only purge email and Teams messages by using Graph. You can purge up to 100 items per location.

  > **Important**
  >
  > Don't use a combination of PowerShell and Microsoft Graph to purge email messages.

- After data is permanently deleted, it can't be recovered. Follow the guidance in this article carefully and validate the scope of the search before issuing the purge command. After you run the purge command, it can't be undone and email messages can't be restored.
- Search conditions such as identifier (including message ID), sensitivity label, and sensitive information type aren't supported for search and delete for non-premium enabled cases in

eDiscovery. Using these conditions results in unexpected data loss.

- Run an export report to review all items that match your search criteria prior to the purge. By using the search and export experience in the Microsoft Purview portal and exporting the results in report-only format, you can examine detailed metadata before deletion. This approach helps refine your search scope and ensures a more targeted and accurate purge.

- Don't use the search and purge workflow described in this article to delete chat messages or other content from Microsoft Teams. If the search that you create in Step 2 returns items from Microsoft Teams, follow the steps outlined in Find and delete Microsoft Teams chat messages in eDiscovery to delete them.

- To create and run a search, you must be a member of the *eDiscovery Manager* role group or be assigned the *Compliance Search* role in the Microsoft Purview portal. To delete messages, you must be a member of the *Organization Management* role group or be assigned the *Search And Purge* role in the Microsoft Purview portal. For information about adding users to a role group, see Assign eDiscovery permissions.

  > **Note**
  >
  > The *Organization Management* role group exists in both Exchange Online and in the Microsoft Purview portal. These are separate role groups that give different permissions. Being a member of *Organization Management* in Exchange Online doesn't grant the required permissions to delete email messages. If you aren't assigned the *Search And Purge* role in the Microsoft Purview portal (either directly or through a role group such as *Organization Management*), you receive an error in Step 3 when you run the *New-ComplianceSearchAction* cmdlet with the message "A parameter can't be found that matches parameter name 'Purge'".

- You must use Security & Compliance PowerShell to delete messages. See Step 1: Connect to Security & Compliance PowerShell for instructions about how to connect.

- You can remove a maximum of 10 items per mailbox at one time. Because the capability to search for and remove messages is intended to be an event-response tool, this limit helps ensure that messages are quickly removed from mailboxes. This feature isn't intended to clean up user mailboxes.

> **Note**

Search and purge isn't supported if a mailbox has a litigation hold. Only 10 items are removed from view of the user. These 10 items aren't permanently deleted, so these 10 items are the only ones processed.

- The maximum number of mailboxes in a content search that you can use to delete items by doing a search and purge action is 50,000. If the search (that you create in Create a search query to find the message to delete) searches more than 50,000 mailboxes, the purge action (that you create in Step 3) fails. Searching more than 50,000 mailboxes in a single search might typically happen when you configure the search to include all mailboxes in your organization. This restriction still applies even when fewer than 50,000 mailboxes contain items that match the search query. See the Frequently asked questions section for guidance about using search permissions filters to search for and purge items from more than 50,000 mailboxes.
- You can only use the procedure in this article to delete items in Exchange Online mailboxes and public folders. You can't use it to delete content from SharePoint or OneDrive sites.
- You can't delete email items in a review set in an eDiscovery case by using the procedures in this article. That's because items in a review set are stored in an Azure Storage location, and not in the live service. This means they're not returned by the content search that you create in Step 1. To delete items in a review set, you have to delete the eDiscovery case that contains the review set.

Step 1: Connect to Security & Compliance PowerShell

First, connect to Security & Compliance PowerShell for your organization. For step-by-step instructions, see Connect to Security & Compliance PowerShell.

ⓘ Important

To use this cmdlet for purge actions in Security & Compliance PowerShell, you need to meet the following requirements:

- Use the Exchange Online PowerShell version 3.9.0 or later (August 2025). For more information, see What's new in the Exchange Online PowerShell module. Installing this version of the module might require you to uninstall previous versions of the module.
- Run the **Connect-IPPSSession** cmdlet with the EnableSearchOnlySession switch.

Step 2: Create a search query to find the message to delete

Next, create and run a search to find the message that you want to remove from mailboxes in your organization. You can create the search by using the Microsoft Purview portal or by running the **New-ComplianceSearch** and **Start-ComplianceSearch** cmdlets in Security & Compliance PowerShell. The messages that match the query for this search are deleted by running the **New-ComplianceSearchAction -Purge** command in Delete the message. For information about creating and configuring search queries, see the following articles:

- Create a search query
- Use the condition builder
- New-ComplianceSearch
- Start-ComplianceSearch

> **Note**
>
> The content locations that you include in the search query can't include SharePoint or OneDrive sites. You can only include mailboxes and public folders in a search that is used for email messages. If the search includes sites, you receive an error when you run the **New-ComplianceSearchAction** cmdlet in Delete the message.

Tips for finding messages to remove

The goal of the search query is to narrow the results to only the messages that you want to remove. Here are some tips:

- If you know the exact text or phrase used in the subject line of the message, use the **Subject** property in the search query.
- If you know the exact date or date range of the message, include the **Received** property in the search query.
- If you know who sent the message, include the **From** property in the search query.
- Preview the search results to verify that the search returns only the messages that you want to delete.
- Use the search estimate statistics, displayed in the details pane of the search in the Microsoft Purview portal or by using the Get-ComplianceSearch cmdlet, to get a count of the total number of results.

Here are two examples of queries to find suspicious email messages.

- This query returns messages that users received between April 13, 2024 and April 14, 2024 and that contain the words "action" and "required" in the subject line.

  PowerShell

  ```
  (Received:4/13/2024..4/14/2024) AND (Subject:'Action required')
  ```

- This query returns messages sent by user@contoso.com and that contain the exact phrase "Update your account information" in the subject line.

  PowerShell

  ```
  (From:user@contoso.com) AND (Subject:"Update your account information")
  ```

Here's an example of using a query to create and start a search by running the **New-ComplianceSearch** and **Start-ComplianceSearch** cmdlets to search all mailboxes in the organization:

PowerShell

```
$Search=New-ComplianceSearch -Name "Remove Phishing Message" -ExchangeLocation All -
ContentMatchQuery '(Received:4/13/2016..4/14/2016) AND (Subject:"Action required")'
Start-ComplianceSearch -Identity $Search.Identity
```

Step 3: Delete the message

After you create and refine a search query to return the messages that you want to remove, run the **New-ComplianceSearchAction -Purge** command in Security & Compliance PowerShell to delete the message.

> **Important**
>
> If you use search in eDiscovery to validate the result and scope what you want to delete, rerun the search cmdlet to validate the search in eDiscovery. The search in eDiscovery isn't run in PowerShell, so the ItemCount is 0 and the JobEndTime isn't set. The search needs to be rerun using PowerShell and you should wait for the search to complete. Verify the results in the cmdlet return the same results before performing the purge.

You can soft-delete or hard-delete the message. A soft-deleted message is moved to a user's Recoverable Items folder and retained until the deleted item retention period expires. Hard-deleted messages are marked for permanent removal from the mailbox and are permanently removed the next time the mailbox is processed by the Managed Folder Assistant. If single item recovery is enabled for the mailbox, hard-deleted items are permanently removed after the deleted item retention period expires. If a mailbox is placed on hold, deleted messages are preserved until the hold duration for the item expires or until the hold is removed from the mailbox.

> ⓘ **Note**
>
> Items from Microsoft Teams that are returned by the search query aren't deleted when you run the **New-ComplianceSearchAction -Purge** command.

To run the following commands to delete messages, make sure you're connected to Security & Compliance PowerShell.

Soft-delete messages

In the following example, the command soft-deletes the search results returned by a search query named "Remove Phishing Message".

PowerShell

```
New-ComplianceSearchAction -SearchName "Remove Phishing Message" -Purge -PurgeType SoftDelete
```

Hard-delete messages

> ⚠ **Warning**
>
> Hard-deleted messages are marked for permanent removal and can't be recovered by the user. If single item recovery is enabled, hard-deleted items are permanently removed after the deleted item retention period expires. Verify your search scope before running this command.

To hard-delete the items returned by the "Remove Phishing Message" content search, run the following command:

PowerShell

```
New-ComplianceSearchAction -SearchName "Remove Phishing Message" -Purge -PurgeType
HardDelete
```

When you run the **New-ComplianceSearchAction -Purge -PurgeType SoftDelete** or **New-ComplianceSearchAction -Purge -PurgeType HardDelete** command, the search specified by the *SearchName* parameter is the search query that you created in Create a search query to find the message to delete.

For more information, see New-ComplianceSearchAction.

Frequently asked questions

The following questions address common issues with search and purge operations.

- **I followed the steps in this article for search, verify, and purge but it still doesn't delete items from a mailbox.**

  Sometimes the items matching the search query are more than 10 items (or 100 items if using Microsoft Graph) for a mailbox. If you run the purge by using the PowerShell cmdlet and only delete 10 items, or run the purge by using Microsoft Graph and only delete 100 items, this behavior is normal. Repeat the process multiple times if you want to delete more than 10 (or 100) items per mailbox. We strongly recommend against working around these limits by continually running the purge command. If you need to delete more data from a single location, see Remediate malicious email delivered in Office 365 or Zero-hour auto purge (ZAP) in Microsoft Defender for Office 365.

  A mailbox on litigation hold can also prevent items from being deleted. If the export report associated to the search run in Step 2 (or done in eDiscovery) indicates items are in the *Recoverable Items* folder, this result means there are holds but the items are moved out of view. If you want to remove items from hold, we strongly recommend using Priority Cleanup. Alternatively, you can release the hold and try purging again. Check the status of the hold by running the following cmdlet in Exchange PowerShell:

  PowerShell

  ```
  Get-Mailbox (or MailUser) | fl *Hold*,*Recovery*
  ```

Verify the following values in the cmdlet results:

- InPlaceHolds should be empty or only have values prefixed with –
  - DelayHold should be false
  - SingleItemRecovery should be false

  Check for tenant level policies by running the following cmdlet:

  ```
  PowerShell
  Get-OrganizationConfig | fl *Hold*
  ```

  If there are any tenant level policies, you must exclude each affected mailbox from the purge:

- Mailbox is on delay hold.
  - Items of interest are still found after purge because they're moved to the *Recoverable items* folder.

- **How do I verify if items are under recoverable items folder?**

  To verify items are in the *Recoverable Items* folder, run an export report for the search and review the file path of the items in the CSV report to see if the items are in the *Recoverable items* folder. If yes, the *New-ComplianceSearch* cmdlet doesn't attempt to delete the item. If the CSV report indicates that the items are in the *Recoverable items* folder and the item's mailbox is archived, the mailbox is being resized. These backup items can't be deleted and eventually are removed automatically.

- **How do you get status on the search and remove operation?**

  Run the **Get-ComplianceSearchAction** to get the status on the delete operation. The object that is created when you run the **New-ComplianceSearchAction** cmdlet is named using this format: `<name of Content Search>_Purge`.

- **What happens after you delete a message?**

  A message that's deleted with the `New-ComplianceSearchAction -Purge -PurgeType HardDelete` command is moved to the Purges folder and can't be accessed by the user. After the message is moved to the Purges folder, the message is retained for the deleted item retention period if single item recovery is enabled for the mailbox. (In Microsoft 365, single item recovery is enabled by default when a new mailbox is created.) After the deleted item retention period expires, the message is marked for permanent deletion and is purged from Microsoft 365 the next time the mailbox is processed by the Managed Folder assistant.

If you use the `New-ComplianceSearchAction -Purge -PurgeType SoftDelete` command, messages are moved to the Deletions folder in the user's Recoverable Items folder. It isn't immediately purged from Microsoft 365. The user can recover messages in the Deleted Items folder for the duration based on the deleted item retention period configured for the mailbox. After this retention period expires (or if user purges the message before it expires), the message is moved to the Purges folder and can no longer be accessed by the user. After the message reaches the Purges folder, the message is retained for the duration based on the deleted item retention period configured for the mailbox if single items recovery is enabled for the mailbox. (In Microsoft 365, single item recovery is enabled by default when a new mailbox is created.) After the deleted item retention period expires, the message is marked for permanent deletion and is purged from Microsoft 365 the next time that the mailbox is processed by the Managed Folder assistant.

- **What if I have to delete a message from more than 50,000 mailboxes?**

  You can perform a search and purge operation on a maximum of 50,000 mailboxes (even if less than 50,000 contain items that match the search query). If you need to do a search and purge operation on more than 50,000 mailboxes, consider creating temporary search permissions filters that reduce the number of mailboxes that would be searched to less than 50,000 mailboxes. For example, if your organization contains mailboxes in different departments, states, or a different country/region, you can create a mailbox search permissions filter based on one of those mailbox properties to search a subset of mailboxes in your organization. After you create the search permissions filter, you would create the search (described in Step 1) and then delete the message (described in Step 3). Then you can edit the search permissions filter to search for and purge messages in a different set of mailboxes. For more information about creating search permissions filters, see Configure search permissions filtering in eDiscovery.

> **Note**
>
> Each mailbox location including primary, archive, and any others are counted as individual locations for this limit.

- **Will unindexed items included in the search results be deleted?**

  No, the `New-ComplianceSearchAction -Purge` command doesn't delete unindexed items.

- **What happens if a message is deleted from a mailbox that is placed on Litigation Hold or assigned to a Microsoft 365 retention policy?**

After the message is purged and moved to the Purges folder, the message is retained until the hold duration expires. If the hold duration is unlimited, then items are retained until the hold is removed or the hold duration is changed.

- **Why is the search and remove workflow divided among different Microsoft Purview portal role groups?**

  A person has to be a member of the *eDiscovery Manager* role group or be assigned the *Compliance Search management* role to search mailboxes. To delete messages, a person has to be a member of the *Organization Management* role group or be assigned the *Search And Purge management* role. This division makes it possible to control who can search mailboxes in the organization and who can delete messages.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/19/2026

Finding content in sites in eDiscovery

When you search for documents and files in SharePoint or OneDrive sites, adjust your query approach based on the metadata for the documents and files you're interested in. Files and documents have relevant properties like *Author*, *Created*, *CreatedBy*, *FileName*, *LastModifiedTime*, and *Title*. Most of these properties aren't relevant when searching for communications content in Exchange Online. Using these properties might lead to unexpected results if you use them across both documents and communications. Additionally, the *FileName* and *Title* of a document might not be the same. Using one or the other to try to find a file with specific content might lead to different or inaccurate results. Keep these properties in mind when searching for specific document and file content in SharePoint and OneDrive.

To learn more about finding content in sites in eDiscovery, watch the following video:

https://learn-video.azurefd.net/vod/player?id=4605a9a1-081f-4018-9bd3-77fb6da223e8&locale=en-us&embedUrl=%2Fpurview%2Fedisc-search-sites

For example, to find content related to documents created by User 1, for a project called *Tradewinds*, for specific files named *Financials*, and from January 2020 to January 2022, use a query with the following properties:

- Add User 1 as a data source to the search.
- Select User 1's OneDrive site as the location of interest.
- Add additional groups and their associated SharePoint sites related to the project as data sources.
- For **FileName**, use *Financials*
- For **Keyword**, use *Tradewinds*
- For **Date Range**, use the *January 1, 2020* to *January 31, 2022* range

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Searchable site properties in KeyQL

The following table lists the SharePoint and OneDrive properties supported in search using the eDiscovery KeyQL editor in the Microsoft Purview portal or by using the New-ComplianceSearch or Set-ComplianceSearch cmdlets. For a list of supported condition builder properties, see Use the condition builder to create search queries in eDiscovery.

> **Important**
>
> While documents and files stored on SharePoint and OneDrive might have other properties supported in other Microsoft 365 services, only the document and file properties listed in this table are supported in eDiscovery search tools. Attempting to include other document or file properties in searches isn't supported.

If you set a site status to *Locked* and the site to *NoAccess*, eDiscovery doesn't return search results. For more information about managing site lock statuses, see Lock and unlock sites. In addition, if a document is restricted in the content store, the eDiscovery manager must be a site admin, site owner, or the owner or last modifier of the restricted document to access it in eDiscovery. For more information about restricted content, see Data Loss Prevention policy reference.

> **Important**
>
> If a OneDrive site is archived and then searched within eDiscovery, the search triggers reindexing and the number of items increases over time as the site fully indexes. This effect persists until the site reindexing is complete. To learn more and manually trigger re-indexing, see Manually request crawling and reindexing of a site, a library, or a list.

The table includes an example of the *property:value* syntax for each property and a description of the search results returned by the examples.

Expand table

Property	Property description	Example	Search results returned by the examples
Author	The author field from Office documents, which persists if a document is copied. For example, if a user creates a document and the emails it to someone else who then uploads it to SharePoint, the document still retains the original author. Be sure to use the user's display name for this property.	`author:"Garth Fort"`	All documents authored by Garth Fort.
ContentType	The SharePoint content type of an item, such as Item, Document, or Video.	`contenttype:document`	All documents.
Created	The date that an item is created.	`created>=2021-06-01`	All items created on or after June 1, 2021.
CreatedBy	The person that created or uploaded an item. Be sure to use the user's display name for this property.	`createdby:"Garth Fort"`	All items created or uploaded by Garth Fort.
DetectedLanguage	The language of an item.	`detectedlanguage:english`	All items in English.

DocumentLink	The path (URL) of a specific folder on a SharePoint or OneDrive site. If you use this property, be sure to search the site that the specified folder is located in. Use this property instead of the *Site* and *Path* properties.
	
	To return items located in subfolders of the folder that you specify for the documentlink property, you have to add /* to the URL of the specified folder; for example, `documentlink:"https://contoso.sharepoint.com/Shared Documents/*"`
		`documentlink:"https://contoso-my.sharepoint.com/personal/garthf_contoso_com/Documents/Private/*"`
		
		`documentlink:"https://contoso-my.sharepoint.com/personal/garthf_contoso_com/Documents/Shared with Everyone/*" AND filename:confidential`
			The first example returns all items in the specified OneDrive folder. The second example returns documents in the specified site folder (and all subfolders) that contain the word "confidential" in the file name.

FileExtension	The extension of a file; for example, docx, one, pptx, or xlsx.	`fileextension:xlsx`	All Excel files (Excel 2007 and later)

FileName	The name of a file.	`filename:"marketing plan"`
		
		`filename:estimate`
			The first example returns files with the exact phrase "marketing plan" in the title. The second example returns files with the word "estimate" in the file name.

LastModifiedTime	The date that an item was last changed.	`lastmodifiedtime>=2021-05-01`
		
		`lastmodifiedtime>=2021-05-01 AND lastmodifiedtime<=2021-06-01`
			The first example returns items that were changed on or after May 1, 2021. The

Property	Property description	Example	Search results returned by the examples
			second example returns items changed between May 1, 2021 and June 1, 2021.
MipSensitiveLabel	Search for sensitivity labels applied to items in your organization's data. By using the Condition Builder, users can select the sensitivity label name (for example, Confidential, Highly Confidential, etc.) from the dropdown and then identify the GUID used in KeyQL for the sensitivity label.	`MipSensitiveLabel=f0c0c7ef-3d45-4c95-87bb-6dfc7a081f46`	All items that match the GUID of the sensitivity label.
ModifiedBy	The person who last changed an item. Be sure to use the user's display name for this property.	`modifiedby:"Garth Fort"`	All items that were last changed by Garth Fort.
SensitiveType	Search for identified and classified sensitive information types in your organization's data. By using the Condition Builder, users can select the sensitive information type name (for example, SSN, credit card, bank account numbers, etc.) from the dropdown and then identify the GUID used in KeyQL for the sensitive information type.	`SensitiveType="ba147552-2ed1-4a8f-a441-981e94ab1895"`	All items that match the GUID of the sensitive information type.

SharedWithUsersOWSUser	Documents that are shared with the specified user and displayed on the **Shared with me** page in the user's OneDrive site. These documents are explicitly shared with the specified user by other people in your organization. When you export documents that match a search query that uses the SharedWithUsersOWSUser property, the documents are exported from the original content location of the person who shared the document with the specified user. For more information, see Searching for site content shared within your organization.	`sharedwithusersowsuser:garthf`
		
		`sharedwithusersowsuser:"garthf@contoso.com"`
			Both examples return all internal documents that are explicitly shared with Garth Fort and that appear on the **Shared with me** page in Garth Fort's OneDrive account.

Size	The size of an item, in bytes.	`size>=1`
		
		`size:1..10000`
			The first example returns items larger than 1 byte. The second example returns items from 1 through 10,000 bytes in size.

Title	The title of the document. The Title property is metadata that's specified in Microsoft Office documents. It's different from the file name of the document.	`title:"communication plan"`	Any document that contains the phrase "communication plan" in the Title metadata property of an Office document.

Searchable sensitive data types

You can use eDiscovery search tools in the Microsoft Purview portal to search for sensitive data, such as credit card numbers or social security numbers, stored in documents on SharePoint and OneDrive sites. You can do this by using the `SensitiveType` property and the name (or ID) of a sensitive information type in a keyword query. For example, the query `SensitiveType:"Credit Card Number"` returns documents that contain a credit card number. The query `SensitiveType:"U.S. Social Security Number (SSN)"` returns documents that contain a U.S. social security number.

To see a list of the sensitive information types that you can search for, go to **Data classifications** > **Sensitive info types** in the Microsoft Purview portal. Or you can use the **Get-DlpSensitiveInformationType** cmdlet in Security & Compliance PowerShell to display a list of sensitive information types.

Limitations for searching sensitive data types

- To search for custom sensitive information types, you need to specify the ID of the sensitive information type in the `SensitiveType` property. Using the name of a custom sensitive information type (as shown in the example for built-in sensitive information types in the previous section) returns no results. Use the **Publisher** column on the **Sensitive info types** page in the Microsoft Purview portal (or the **Publisher** property in PowerShell) to differentiate between built-in and custom sensitive information types. Built-in sensitive data types have a value of `Microsoft Corporation` for the **Publisher** property.

  To display the name and ID for the custom sensitive data types in your organization, run the following command in Security & Compliance PowerShell:

  PowerShell

  ```
  Get-DlpSensitiveInformationType | Where-Object {$_.Publisher -ne "Microsoft Corporation"} | FT Name,Id
  ```

  Then you can use the ID in the `SensitiveType` search property to return documents that contain the custom sensitive data type; for example, `SensitiveType:7e13277e-6b04-3b68-94ed-1aeb9d47de37`

- You can't use sensitive information types and the `SensitiveType` search property to search for sensitive data at-rest in Exchange Online mailboxes. This limitation includes 1:1 chat messages, 1:N group chat messages, and team channel conversations in Microsoft Teams because all of this content is stored in mailboxes. However, you can use data loss prevention (DLP) policies to protect sensitive email data in transit. For more information, see Learn about data loss prevention and Search for and find personal data.

Forming a basic query

A basic query has three parts: SensitiveType, count range, and confidence range. For example, **SensitiveType:"<type>"** is required, and both **|<count range>** and **|<confidence range>** are optional.

Sensitive type - required

Queries typically begin with the property `SensitiveType:"` and an information type name from the sensitive information types inventory, and end with a `"`. You can also use the name of a custom sensitive information type that you created for your organization. For example, you might be looking for documents that contain credit card numbers.

In such an instance, use the following format: `SensitiveType:"Credit Card Number"`. Because you didn't include count range or confidence range, the query returns every document in which a credit card number is detected. This is the simplest query that you can run, and it returns the most results. Keep in mind that the spelling and spacing of the sensitive type matters.

Ranges - optional

Both of the next two parts are ranges, so let's quickly examine what a range looks like. In SharePoint queries, a basic range is represented by two numbers separated by two periods, which looks like this: `[number]..[number]`. For instance, if `10..20` is used, that range captures numbers from 10 through 20. This article covers many different range combinations.

Let's add a count range to the query. You can use count range to define the number of occurrences of sensitive information a document needs to contain before it's included in the query results. For example, if you want your query to return only documents that contain exactly five credit card numbers, use this: `SensitiveType:"Credit Card Number|5"`. Count range can also help you identify documents that pose high degrees of risk. For example, your organization might consider documents with five or more credit card numbers a high risk. To find

documents fitting this criterion, use this query: `SensitiveType:"Credit Card Number|5.."`. Alternatively, you can find documents with five or fewer credit card numbers by using this query: `SensitiveType:"Credit Card Number|..5"`.

Confidence range

Finally, confidence range is the level of confidence that the detected sensitive type is actually a match. The values for confidence range work similarly to count range. You can form a query without including a count range. For example, to search for documents with any number of credit card numbers—as long as the confidence range is 85 percent or higher—use this query: `SensitiveType:"Credit Card Number|*|85.."`.

> **Important**
>
> The asterisk (`*`) is a wildcard character that means any value works. You can use the wildcard character (`*`) either in the count range or in the confidence range, but not in a sensitive type.

Additional query properties and search operators

Queries in SharePoint also contain the `LastSensitiveContentScan` property, which can help you search for files scanned within a specific timeframe. For query examples with the `LastSensitiveContentScan` property, see the Examples of complex queries in the next section.

You can use SharePoint eDiscovery search properties such as `Author` or `FileExtension`. You can use operators to build complex queries.

Examples of complex queries

The following examples use different sensitive types, properties, and operators to show how you can refine your queries to find exactly what you're looking for.

Expand table

Query	Explanation
`SensitiveType:"Credit Card Number | 1..4294967295 | 1..100"`	This query returns documents with at least one match to the sensitive type "Credit Card Number." The values for each range are the respective minimum and maximum values. A simpler way to write this query is `SensitiveType:"Credit Card Number"`.
`SensitiveType:"Credit Card Number | 5..25" AND LastSensitiveContentScan:"8/11/2018..8/13/2018" NOT FileExtension:XLSX`	This query returns documents with 5-25 credit card numbers that were scanned from August 11, 2018 through August 13, 2018. Files with an XLSX extension aren't included in the query results. `FileExtension` is one of many properties that you can include in a query. For more information, see Using Search Properties and Operators with eDiscovery.
`SensitiveType:"Credit Card Number | 5..25" AND LastSensitiveContentScan:"8/11/2018..8/13/2018"`	This query returns documents with 5-25 credit card numbers that were scanned from August 11, 2018 through August 13, 2018.
`SensitiveType:"Credit Card Number" OR SensitiveType:"U.S. Social Security Number (SSN)"`	This query returns documents that contain either a credit card number or a social security number.
`SensitiveType:"International Banking Account Number (IBAN)"`	The name might seem strange because it's so long, but it's the correct name for that sensitive type. Make sure to use exact names from the sensitive information types inventory. You can also use the name of a custom sensitive information type that you created for your organization.

Examples of queries to avoid

Not all queries work in SharePoint. The following table gives examples of queries that don't work and describes why.

Expand table

Unsupported query	Reason
`SensitiveType:"Credit Card Number .."`	You must add at least one number.
`SensitiveType:"Credit Card Number 0"`	Zero isn't valid as either the minimum value or the maximum value in a range.

Unsupported query	Reason
`SensitiveType:"Credit Card Number 1. .3"`	The two-period portion shouldn't be separated by a space.
`SensitiveType:"Credit Card Number 1.. 80.."`	There are too many pipe delimiters `|`. Follow this format instead: `SensitiveType: "Credit Card Number 1.. 80.."`
`SensitiveType:"Credit Card Number 1.. 80..101"`	Because confidence values represent a percentage, they can't exceed 100. Choose a number from 1 through 100 instead.
`SensitiveType:"Credit Card Number"`	It might be difficult to see, but there's extra white space between "Credit" and "Card" that makes the query invalid. Use exact sensitive type names from the sensitive information types inventory.
`SensitiveType:"NotARule"`	"NotARule" isn't a valid sensitive type name. Only names in the sensitive information types inventory work in eDiscovery queries.

Searching for site content shared with external users

You can use eDiscovery search tools in the Microsoft Purview portal to search for documents stored on SharePoint and OneDrive sites shared with people outside of your organization. This search can help you identify sensitive or proprietary information that's being shared outside your organization. Use the `ViewableByExternalUsers` property in a keyword query. This property returns documents or sites that you shared with external users by using one of the following sharing methods:

- A sharing invitation that requires users to sign in to your organization as an authenticated user.
- An anonymous guest link, which allows anyone with this link to access the resource without having to be authenticated.

Here are some examples:

- The query `ViewableByExternalUsers:true AND SensitiveType:"Credit Card Number"` returns all items shared with people outside your organization and contains a credit card number.
- The query `ViewableByExternalUsers:true AND ContentType:document AND site:"https://contoso.sharepoint.com/Sites/Teams"` returns a list of documents on all team sites in the organization that you shared with external users.

> **Tip**
>
> A search query such as `ViewableByExternalUsers:true AND ContentType:document` might return numerous .aspx files in the search results. To eliminate these or other types of files, use the `FileExtension` property to exclude specific file types; for example `ViewableByExternalUsers:true AND ContentType:document NOT FileExtension:aspx`.

What content counts as shared with people outside your organization? Documents in your organization's SharePoint and OneDrive sites that you shared by sending a sharing invitation or that you shared in public locations. For example, the following user activities result in content that external users can view:

- A user shares a file or folder with a person outside your organization.
- A user creates and sends a link to a shared file to a person outside your organization. This link allows the external user to view (or edit) the file.
- A user sends a sharing invitation or a guest link to a person outside your organization to view (or edit) a shared file.

Issues using the ViewableByExternalUsers property

While the `ViewableByExternalUsers` property represents the status of whether a document or site is shared with external users, it has some caveats. In the following scenarios, the value of the `ViewableByExternalUsers` property isn't updated, and the results of a search query that uses this property might be inaccurate.

- Changes to sharing policy, such as turning off external sharing for a site or for the organization. The property still shows previously shared documents as being externally accessible even though external access might be revoked.
- Changes to group membership, such as adding or removing external users to Microsoft 365 Groups or Microsoft 365 security groups. The property isn't automatically updated for items the group has access to.

- Sending sharing invitations to external users where the recipient hasn't accepted the invitation, and therefore doesn't yet have access to the content.

In these scenarios, the `ViewableByExternalUsers` property doesn't reflect the current sharing status until the site or document library is recrawled and reindexed.

Searching for site content shared within your organization

You can use the `SharedWithUsersOWSUser` property to search for documents shared between people in your organization. When a person shares a file or folder with another user inside your organization, a link to the shared file appears on the **Shared with me** page in the OneDrive account of the person who the file was shared with. For example, to search for the documents shared with Sara Davis, you can use the query `SharedWithUsersOWSUser:"sarad@contoso.com"`. If you export the results of this search, you download the original documents located in the content location of the person who shared the documents with Sara.

Documents must be explicitly shared with a specific user to be returned in search results when using the `SharedWithUsersOWSUser` property. For example, when a person shares a document in their OneDrive account, they can share it with anyone (inside or outside the organization), share it only with people inside the organization, or share it with a specific person.

Only documents that are shared by using the third option (shared with **Specific people**) are returned by a search query that uses the `SharedWithUsersOWSUser` property.

Character limits for searches

For more information about character limits, see eDiscovery search limits.

Last updated on 04/02/2026

Finding content in Microsoft Teams in eDiscovery

This article provides a comprehensive set of procedures, guidelines, and best practices for using eDiscovery to preserve, collect, review, and export content from Microsoft Teams. The goal of this article is to help you optimize your eDiscovery search for Teams content.

To learn more about finding content in Microsoft Teams with eDiscovery, check out the following video:

https://learn-video.azurefd.net/vod/player?id=c077f1d9-33f0-43c3-a44e-eafe3095cf0f&locale=en-us&embedUrl=%2Fpurview%2Fedisc-search-teams

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about **Microsoft Security Copilot in Microsoft Purview**.

Where Teams content is stored

To manage Teams content in eDiscovery, you need to understand the type of Teams content you can collect, process, and review in eDiscovery and where Microsoft 365 stores that content. To configure Teams content locations in a case, see Configure data sources. Teams data is stored in Azure Cosmos DB. Exchange Online stores Teams compliance records that the substrate captures, and eDiscovery can access these records. The data stored in Exchange Online is hidden from clients. eDiscovery never operates against the real Teams message data, which remains in Azure Cosmos DB.

The following table lists Teams content types and where Microsoft 365 stores each type for compliance purposes.

Expand table

Teams category	Description	Chat messages/posts location	Files/attachments location	Meeting recordings location
Events	All data from webinars and town halls is stored in Cosmos DB. Substrate (SCD) serves as secondary storage for redundancy and eDiscovery purposes.	N/A	N/A	N/A
Private channels	Message posts, replies, and attachments shared in a private Teams channel.	Messages sent in a private channel are stored in the Exchange Online mailboxes of all members of the private channel.	Files shared in a private channel are stored in a dedicated SharePoint site associated with the private channel.	N/A
Shared channels	Message posts, replies, and attachments shared in a shared Teams channel.	Messages sent in a shared channel are stored in a system mailbox associated with the shared channel.[2]	Files shared in a shared channel are stored in a dedicated SharePoint site associated with the shared channel.	N/A
Teams 1:1 chats	Chat messages, posts, and attachments shared in a Teams conversation between two people. Teams 1:1 chats are also called *conversations*.	Messages in 1:1 chats are stored in the Exchange Online mailbox of all chat participants.	Files shared in a 1:1 chat are stored in the OneDrive account of the person who shared the file.	N/A
Teams channels	Chat messages, posts, replies, and attachments shared in a standard Teams channel.	All channel messages and posts are stored in the Exchange Online mailbox associated with the team.	Files shared in a channel are stored in the SharePoint site associated with the team.	N/A
Teams group chats	Chat messages, posts, and attachments shared in a Teams conversation between three or more people. Also	Messages in group chats are stored in the Exchange Online mailbox of all chat participants.	Files shared in group chats are stored in the OneDrive account of the person who shared the file.	N/A

Teams category	Description	Chat messages/posts location	Files/attachments location	Meeting recordings location
	called *1:N* chats or *group conversations*.			
Teams meetings	Audio and transcripts from recorded Teams meetings.	Chats in recorded meetings are stored in an Exchange Online mailbox. For standard meetings, it's stored in the mailboxes of all participants. For channel meetings, it's stored in the mailbox associated with the team.	Files and attachments shared in recorded meetings are stored in the OneDrive account for the user recording the Teams meeting.	Meeting recordings are stored in the OneDrive account for the user recording the Teams meeting.[1]
Teams reactions	Reactions applied to chat messages, posts, and attachments in a Teams conversation.	Messages in group chats are stored in the Exchange Online mailbox of all chat participants.	Files shared in group chats are stored in the OneDrive account of the person who shared the file.	N/A

> **Note**
>
> [1] To locate Teams meeting audio and transcript content, create an eDiscovery search targeting the organizer's OneDrive. Include the date range and file extension in the query. When the search completes, refine the results in a review set and select only the pertinent data. For example, a date range might be from *March 17, 2025 to March 18, 2025* and the file extension might be *mp4*.
>
> [2] To search for (and preserve) messages sent in a shared channel, you must search or specify the Exchange Online mailbox for the parent Team.

Teams content types

All Microsoft Teams 1:1 or group chats are journaled through to the respective users' mailboxes. All standard channel messages are journaled through to the group mailbox representing the team. Files uploaded in standard channels are covered under the eDiscovery functionality for SharePoint and OneDrive.

eDiscovery of messages and files in private channels works differently than in standard channels. To learn more, see eDiscovery of private channels.

Recorded Teams meetings are stored in the OneDrive account of the meeting organizer or initiator, depending on the *MeetingRecordingOwnership* configuration. Attendee identification information when using the **Hide Attendee Names** functionality for Teams meetings is stored in the user mailbox of the meeting organizer.

Not all Teams content is eDiscoverable. The following table shows the Teams content types that you can search for using Microsoft eDiscovery tools:

Expand table

Content type	Notes
Audio recordings	Audio calls between Teams user and external contacts
Card content	See Search for card content for more information.
Chat links	
Chat messages	This includes content in standard Teams channels, 1:1 chats, 1:N group chats, chats with yourself, and chats with guests.
Code snippets	
Edited messages	If the user is on hold, previous versions of edited messages are also preserved.
Emojis, GIFs, and stickers	
Inline images	
Loop components	Content in a loop component is saved in a .fluid file that's stored in the OneDrive account of the user who sends the loop component. That means you have to include OneDrive as a data source when searching for content in loop components.
Meeting IM conversations	
Meeting metadata[1]	
Meeting recordings and transcripts	Transcripts of the meeting audio are extracted and provided as a separate file. Maximum supported recorded meeting .mp4 file size is 350 MB. If the recorded meeting file size is greater than 350 MB, a processing error occurs and the file is available for download.
Name of channel	

Content type	Notes
Quotes	Quoted content is searchable. However, search results don't indicate that the content was quoted.
Reactions (such as likes, hearts, and other reactions)	Reactions are supported for all commercial customers after June 1, 2022. Reactions before this date aren't available for eDiscovery. Expanded reactions are now supported. To understand reaction history, the content must be on legal hold.
Subject	
Tables	
Teams Video Clip (TVC)	Search TVC with "Video-Clip" keyword and "save as" a .mp4 file for each TVC attachment by right-clicking the preview. TVCs are collected as Teams conversation attachments (if smaller than 200 MB) and separate .mp4 files. TVC file data is discoverable in eDiscovery review sets and can be exported. Preview of video clips isn't currently supported.

[1] Meeting (and call) metadata includes the following:

- Meeting start and end time, and duration
- Meeting join and leave events for each participant
- VOIP joins/calls
- Federated user joins
- Guest joins

> **Important**
>
> Anonymous users joining meetings and calls aren't currently supported in eDiscovery search queries.

Microsoft Teams data appears as IM or Conversations in the Excel eDiscovery export output. You can open the `.pst` file in Outlook to view those messages after you export them.

When viewing the .pst file for the team, all conversations are located in the Team Chat folder under Conversation History. The title of the message contains the team name and channel name.

Private chats in a user's mailbox are stored in the Team Chat folder under Conversation History.

eDiscovery of private and shared channels

Compliance copies of messages in private and shared channels go to different mailboxes depending on the channel type. That difference means you need to search different mailbox locations based on the type of channel a user belongs to.

- **Private channels**: Compliance copies go to the dedicated private channel mailbox. You need to search private channel mailboxes when searching for content in private channel messages.
- **Shared channels**: Compliance copies go to a system mailbox that's associated with the parent team. Because Teams doesn't support an eDiscovery search of a single system mailbox for a shared channel, you need to search the mailbox for the parent team (by selecting the name of the Team mailbox) when searching for message content in shared channels.

Each private and shared channel has its own SharePoint site that's separate from the parent team site. Files in private and shared channels are stored in their own site and managed independently of the parent team. You must identify and search the specific site associated with a channel when searching for content in files and channel message attachments.

Use the following sections to help identify the private or shared channel to include in your eDiscovery search.

Identifying the members of a private channel

Use the following procedure to identify members of a private channel so that you can use eDiscovery tools to search the member's mailbox for content in private channel messages.

Before you perform these steps, make sure you have the latest version of the Teams PowerShell module installed.

1. Run the following command to get the group ID of the team that contains the shared channels you want to search.

   PowerShell

   ```
   Get-Team -DisplayName <display name of the parent team>
   ```

   > **Tip**
   >
   > Run the **Get-Team** cmdlet without any parameters to display a list of all Teams in your organization. The list contains the group ID and DisplayName for every team.

2. Run the following command to get a list of private channels in the parent team. Use the group ID for the team that you obtained in step 1.

```
PowerShell
Get-TeamChannel -GroupId <parent team GroupId> -MembershipType Private
```

3. Run the following command to get a list of private channel owners and members for a specific private channel.

```
PowerShell
Get-TeamChannelUser -GroupId <parent team GroupId> -DisplayName "Partner Shared Channel"
```

4. Include the mailboxes of owners and members of a private channel as part of your eDiscovery search query.

Search for content for guests

You can use eDiscovery tools to search for Teams content related to guests in your organization. Teams chat content that's associated with a guest is preserved in a cloud-based storage location, and you can search for it by using eDiscovery. This content includes 1:1 and 1:N chat conversations where a guest participates with other users in your organization. You can also search for private channel messages where a guest participates and for content in *guest:guest* chat conversations where the only participants are guests.

To search for content for guests:

1. Connect to Microsoft Graph PowerShell. For more information, see the Microsoft Graph PowerShell overview. Be sure to complete Step 1 and Step 2 in the previous article.

2. After you successfully connect to Microsoft Graph PowerShell, run the following command to display the user principal name (UPN) for all guests in your organization. You need to use the UPN of the guest when you create the search in step 4.

```
PowerShell
Get-MgUser -Filter "userType eq 'Guest'" -All $true | FL UserPrincipalName
```

> **Tip**

> Instead of displaying a list of user principal names on the computer screen, you can redirect the output of the command to a text file. You can do this by appending `> filename.txt` to the previous command. The text file with the user principal names is saved to the current folder.

3. In a different Windows PowerShell window, connect to Security & Compliance PowerShell. For instructions, see Connect to Security & Compliance PowerShell. You can connect with or without using multifactor authentication.

4. Create a search query that searches for all content (such as chat messages and email messages) where the specified guest was a participant by running the following command.

   PowerShell

   ```
   New-ComplianceSearch <search name> -ExchangeLocation <guest UPN>  -
   AllowNotFoundExchangeLocationsEnabled $true -IncludeUserAppContent $true
   ```

   For example, to search for content associated with the guest Sara Davis, run the following command.

   PowerShell

   ```
   New-ComplianceSearch "Sara Davis Guest" -ExchangeLocation
   "sara.davis_hotmail.com#EXT#@contoso.onmicrosoft.com" -
   AllowNotFoundExchangeLocationsEnabled $true -IncludeUserAppContent $true
   ```

   For more information about using PowerShell to create searches, see New-ComplianceSearch.

5. Run the following command to start the search that you created in step 4:

   PowerShell

   ```
   Start-ComplianceSearch <search name>
   ```

6. Go to the Microsoft Purview portal and sign in by using the credentials for a user account assigned eDiscovery permissions.

7. Select the **eDiscovery** solution card and then select **Cases** in the left nav.

8. Select a case.

9. In the list of searches on the **Searches** tab, select the search that you created in step 4 to display the flyout page.

10. On the flyout page, you can do the following things:

    - Select **Sample** to view the search results and preview the content.
    - Next to the **Query** field, select **Edit** to edit and then rerun the search. For example, you can add a search query to narrow the results.
    - Select **Export** to export and download the search results.

Search for card content

Apps generate card content in Teams channels, 1:1 chats, and 1xN chats. The content is stored in mailboxes and you can search it. A *card* is a UI container for short pieces of content. Cards can have multiple properties and attachments, and can include items that trigger card actions. For more information, see Cards.

Like other Teams content, where you store card content depends on where you used the card. The Teams group mailbox stores content for cards used in a Teams channel. The mailboxes of the chat participants store card content for 1:1 and 1xN chats.

To search for card content, use the `kind:microsoftteams` or `itemclass:IPM.SkypeTeams.Message` search conditions. When you review search results, you see that bots generate card content in a Teams channel and the **Sender/Author** email property is `<appname>@teams.microsoft.com`, where `appname` is the name of the app that generated the card content. If a user generates card content, the **Sender/Author** value identifies the user.

When you view card content in search results, the content appears as an attachment to the message. The attachment is named `appname.html`, where `appname` is the name of the app that generated the card content.

> **Note**
>
> To display images from card content in search results (such as the checkmarks in the previous screenshot), you need to be signed into Teams (at https://teams.microsoft.com) in a different tab in the same browser session that you use to view the search results. Otherwise, image placeholders are displayed.

Search for Teams meetings by date

Admins can search for Teams meeting content based on the meeting start or end dates. To filter review set items by specific Teams meeting dates, use the *Meeting start date* and *Meeting end date* properties in eDiscovery search tools.

Search for hidden attendees in Teams meetings

The **Hide Attendee Names** feature in Microsoft Teams hides the name of attendees from other attendees so only organizers can see attendees names. Use this feature for meetings or events with external businesses, vendors, confidential meetings, or other meetings where personal privacy between attendees is important. When you enable this feature, the names of attendees are hidden in the meeting roster, meeting chat, and meeting recordings.

To search for the names of attendees in Teams meetings where the **Hide Attendee Names** feature was enabled, you must include the organizer's mailbox in the scope for the search. Hidden attendee names are only available in the organizer's mailbox.

Search for Teams Events data

All data from webinars and town halls is stored in Cosmos DB. Substrate (SCD) serves as secondary storage for redundancy and eDiscovery purposes. For webinars created before September 2024, this data remains in the webinar organizer's SharePoint storage to ensure continuity for V1 webinars.

eDiscovery applies all event-related information, including *Event titles, Descriptions, Attendee registration data, Presenters,* and *Themes*. Event registration details, such as attendee lists and registrant email addresses, are captured in eDiscovery logs. Live attendance data isn't included.

For legal cases, use the **Microsoft Purview portal** to include "*Microsoft Teams Virtual Events*" as an external data source in the processes. After creating a case and selecting data sources, create a search query to collect relevant Teams Event data for legal and compliance needs.

For future updates, see the **Message Center** for feature availability and functionality.

eDiscovery in external access and guest environments

Admins can use eDiscovery to search for content in chat messages in a Teams meeting in external access and guest access environments based on the following restrictions:

- **External access**: In a Teams meeting with users from your organization and users from an external organization where external attendees are using external access, admins in both organizations can search for content in chat messages from the meeting.
- **Guest**: In a Teams meeting with users from your organization and guests, only admins in the organization who hosts the Teams meeting can search for content in chat messages from the meeting.

Searching for Skype for Business conversations

You can use the following keyword query to specifically search for content in Skype for Business conversations:

PowerShell

```
kind:im
```

The previous search query also returns chats from Microsoft Teams. To prevent this issue, you can narrow the search results to include only Skype for Business conversations by using the following keyword query:

PowerShell

```
kind:im AND subject:conversation
```

The previous keyword query excludes chats in Microsoft Teams because Skype for Business conversations are saved as email messages with a Subject line that starts with the word "Conversation".

To search for Skype for Business conversations that occurred within a specific date range, use the following keyword query:

PowerShell

```
kind:im AND subject:conversation AND (received=startdate..enddate)
```

Related articles

- Teams PowerShell Overview

Last updated on 04/02/2026

Find and delete Microsoft Teams chat messages in eDiscovery

> **Tip**
>
> You can also use the purge feature in Data Security Investigations to search for and permanently delete Microsoft Teams items in your organization. Data Security Investigations provides a built-in search and purge workflow with a purge queue dashboard to help you manage and track purge queries. For more information, see Purge in Data Security Investigations.

You can use eDiscovery and the Microsoft Graph Explorer to search for and delete chat messages in Microsoft Teams. This feature helps you find and remove sensitive information or inappropriate content. This search and delete workflow helps you respond to a data spillage event, when content containing confidential or malicious information is released through Teams chat messages.

Microsoft Teams chat data is stored in separate copies, one for compliance purposes (used by tools like eDiscovery) and one for user access.

- **Soft delete or hard delete operations** remove the user copy permanently and can't be recovered for user use.
- **Cmdlet-based purges** of Teams messages can remove the compliance copy of Teams messages, but the user copy remains visible. After the compliance copy is deleted, eDiscovery can no longer remove the user-visible message. Verify Teams messages carefully before purging and avoid using cmdlets to purge Teams messages (the user copy isn't deleted).
  - If you use cmdlets, review the export report and exclude Teams items. Deletes are only performed on the compliance copy and don't remove copies from user viewing. Deleting only the compliance copy also prevents any future deletion when using Microsoft Graph APIs.
  - If you use Microsoft Graph to review the report and ensure Teams items being deleted are intentional, there's no recovery of the items after the deletion is complete.

To ensure complete removal of items, plan your purge strategy carefully and understand which copy each deletion method targets.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Before you search and delete chat messages

> **Important**
>
> Carefully review the following guidance before starting the search and purge steps for Teams chat messages.

- Depending on the eDiscovery subscription for your organization, cases can either be enabled for premium features or not. Verify the level of feature support for the case to determine if the search and purge should be conducted using PowerShell or Microsoft Graph. Cases not configured for premium features can only use PowerShell to search and delete email messages and cases configured for premium features can use PowerShell or Microsoft Graph to search and delete email messages. Don't use a combination of PowerShell and Microsoft Graph to purge email messages and chats.
- After data is permanently deleted, it isn't recoverable. Follow the guidance in this article carefully and validate the scope of the search before issuing the purge command. After you run the purge command, you can't undo it and email messages and chats can't be restored.
- Run an export report to review all items that match your search criteria before the purge. Using the search and export experience in the Microsoft Purview portal and exporting the results in report-only format allows you to examine detailed metadata before deletion. This approach helps refine your search scope and ensures a more targeted and accurate purge.
- To create an eDiscovery case and search for chat messages, you need to be a member of the *eDiscovery Manager* role group in the Microsoft Purview portal. To delete chat messages, you need to be assigned the *Search And Purge* role. This role is assigned to the *Data Investigator* and *Organization Management* role groups by default. For more information, see Assign eDiscovery permissions.

- Search and purge are supported for most conversations within your organization. Search and purge for Teams Connect Chat (External Access or Federation) conversations isn't supported.

  > **Important**
  >
  > Chats with yourself (or chats by users with themselves) aren't supported for search and delete.

- You can remove up to 100 items per mailbox at one time. Because the capability to search for and remove chat messages is intended to be an event-response tool, this limit helps ensure that chat messages are quickly removed.

Step 1: Create a case in eDiscovery

The first step is to create a case in eDiscovery to manage the search and delete process.

Step 2: Create a search query

After you create a case, you need to search for the Teams chat messages that you want to delete. The delete process deletes all items that are found in the search (within the 10 item per location limit).

Data sources for chat messages

Use the following table to determine which data sources to search depending on the type of chat message you need to delete.

Expand table

For this type of chat...	Search this data source...
Teams 1:1 chats	The mailbox of chat participants.
Teams group chats	The mailboxes of chat participants.
Teams channels (standard and shared)	The mailbox associated with the parent team.
Teams private channels	A dedicated mailbox for each private channel.

> **Note**

Before deleting messages, you also need to identify and remove any holds and retention policies assigned to the mailbox that contains the type of chat messages you want to delete.

Tips for searching for chat messages

To help ensure the most comprehensive identification of Teams chat conversations (including 1:1 and group chats, and chats from standard, shared, and private chats), use the **Type** condition and select the **Instant messages** option when you build the search query. Include a date range or several keywords to narrow the scope of the search to items relevant to your investigation.

Step 3: Review and verify chat messages to delete

The deletion process deletes all items returned by the search. It's important that you review the search statistics to ensure that the search only returns the items that you want to delete. Additionally, you can use the search statistics (specifically the *Top Locations* statistics) to generate a list of the data sources that contain items returned by the search. Use this list when removing hold and retention policies from the data sources that contain search results.

Step 4: Remove all holds and retention policies from data sources

Before you can delete chat messages from a mailbox, you need to remove all organization-wide holds, site holds, or retention policy holds that are assigned to a target mailbox. If you don't remove these holds, the chat messages you're trying to delete are retained.

Use the list of mailboxes that contain the chat messages you want to delete to determine if there's a hold or retention policy assigned to those mailboxes, and then remove the hold or retention policy. Be sure to identify the hold or retention policy that you remove so that you can reassign it to the mailboxes after deletion is complete.

For instructions about how to identify and remove holds and retention policies, see "Step 3: Remove all holds from the mailbox" in Delete items in the Recoverable Items folder of cloud-based mailboxes on hold.

Step 5: Delete chat messages from Teams

ⓘ Note

Because Microsoft Graph Explorer isn't available in some US Government clouds (GCC High and DOD), you must use PowerShell to accomplish these tasks. See **Delete chat messages with PowerShell** for details.

Now you're ready to delete chat messages from Teams. Use the Microsoft Graph Explorer to perform the following tasks:

1. Get the ID of the eDiscovery case that you created in Step 1. This case contains the search results created in Step 2.
2. Get the ID of the search that you created in Step 2 and verified the search results in Step 3. The search query returns the chat messages that you want to delete.
3. Delete the chat messages returned by the search.

For information about using Graph Explorer, see Use Graph Explorer to try Microsoft Graph APIs.

> **Important**
>
> To perform these three tasks in Graph Explorer, you might have to consent to the eDiscovery.Read.All and eDiscovery.ReadWrite.All permissions. For more information, see the "Consent to permissions" section in **Working with Graph Explorer**.

Get the case ID

Use Graph Explorer to retrieve the ID of the eDiscovery case you created earlier.

1. Go to https://developer.microsoft.com/graph/graph-explorer and sign in to the Graph Explorer with an account that's assigned the *Search And Purge* role in the Microsoft Purview portal.

2. Run the following GET request to retrieve the ID for the eDiscovery case. Use the value `https://graph.microsoft.com/v1.0/security/cases/ediscoveryCases` in the address bar of the request query. Be sure to select **v1.0** in the API version dropdown list.

   This request returns information about all cases in your organization on the **Response preview** tab.

3. Scroll through the response to locate the eDiscovery case. Use the **displayName** property to identify the case.

4. Copy the corresponding ID (or copy and paste it to a text file). Use this ID in the next task to get the search ID.

> **Tip**
>
> Instead of using the previous procedure to obtain the case ID, you can open the case in the Microsoft Purview portal and copy the case ID from the URL.

Get the eDiscoverySearchID

Next, retrieve the ID of the search that contains the chat messages you want to purge.

1. In Graph Explorer, run the following GET request to retrieve the ID for the search that you created in Step 2, and contains the items you want to delete. Use the value `https://graph.microsoft.com/v1.0/security/cases/ediscoveryCases/{ediscoveryCaseID}/searches` in the address bar of the request query, where *{ediscoveryCaseID}* is the CaseID that you obtained in the previous procedure.

2. Scroll through the response to locate the search that contains the items you want to delete. Use the *displayName* property to identify the search that you created and verified earlier.

   In the response, the search query from the search is displayed in the *contentQuery* property. Items returned by this query are deleted in the next task.

3. Copy the corresponding ID (or copy and paste it to a text file). Use this ID in the next task to delete the chat messages.

Delete the chat messages

After you confirm the correct case and search IDs, use the following request to purge the matching chat messages.

1. In Graph Explorer, run the following POST request to delete the items returned by the search that you created in Step 2. Use the value `https://graph.microsoft.com/v1.0/security/cases/ediscoveryCases/{ediscoveryCaseID}/searches/{ediscoverySearchID}/purgeData` in the address bar of the request query, where *{ediscoveryCaseID}* and *{ediscoverySearchID}* are the IDs that you obtained in the previous procedures.

If the POST request is successful, an HTTP response code is displayed in a green banner stating that the request was accepted.

For more information about purgeData, see sourceCollection: purgeData.

Delete chat messages with PowerShell

You can also delete chat messages by using PowerShell. For example, to delete messages in the US Government cloud, you can use a command similar to the following example:

PowerShell

```
Connect-MgGraph -Scopes "ediscovery.ReadWrite.All" -Environment USGov
```

PowerShell

```
Invoke-MgGraphRequest  -Method POST -Uri
'/v1.0/security/cases/ediscoveryCases/<ediscoverySearchID>/searches/<search
ID>/purgeData'
```

For more information about using PowerShell to delete chat messages, see ediscoverySearch: purgeData.

Step 6: Verify chat messages are deleted

When you run the POST request to delete chat messages, the Teams client removes these messages and replaces them with an automatically generated message stating that an admin removed the message. For an example of this message, see the User experience section in this article.

If you need to confirm a chat message is removed but don't have access to the user message in Teams, rerun the search and check if any matching messages appear. If there aren't any results for the message, the message is removed.

When you issue a purge:

- The **user copy** of the message is deleted immediately and can't be recovered by the user.
- The **compliance copy** (stored in the mailbox's *SubstrateHolds* folder) is retained for at least 24 hours before permanent deletion by a background timer job (typically within 1–7 days). If a hold is removed and then re-applied within that 24-hour window, the compliance copy might be preserved by the new hold.

For more information, see Learn about retention for Microsoft Teams.

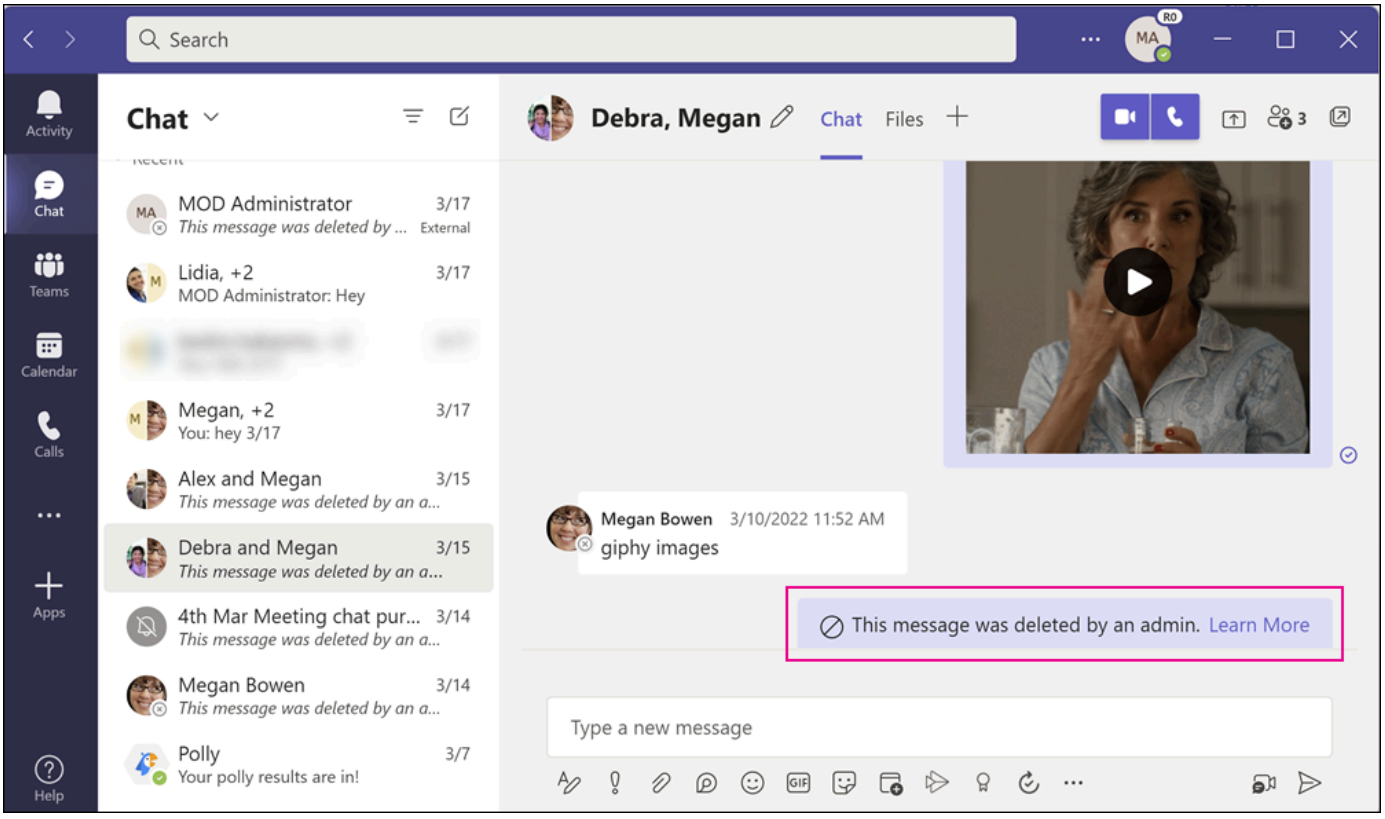
> **Note**
>
> Because Microsoft Graph Explorer isn't available in the US Government cloud (GCC, GCC High, and DOD), you must use PowerShell to accomplish these tasks.

Step 7: Reapply holds and retention policies to data sources

After you verify that chat messages are deleted and removed from the Teams client, reapply the holds and retention policies that you removed in Step 4.

User experience after Teams chat messages are deleted

For deleted chat messages, users see an automatically generated message stating *This message was deleted by an admin*.



The message in the previous screenshot replaces the chat message that was deleted.

> **Note**

If you're a user and a chat message was deleted, contact your admin for more information.

Related content

For frequently asked questions about search and purge, see Find and delete email messages in eDiscovery.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/29/2026

Search for and delete AI application data in eDiscovery

You can use eDiscovery and the Microsoft Graph Explorer to search for and delete user prompts and Microsoft Copilot and other AI solutions in supported applications and services. This feature can help you find and remove sensitive information or inappropriate content included in Copilot and other AI application activities. The eDiscovery search-and-deletion workflow described in this article can also help you respond to a data spillage event, when content containing confidential or malicious information is released through AI-related activity.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about **Microsoft Security Copilot in Microsoft Purview**.

Before you search and delete AI data

- To create an eDiscovery case and search for AI activity data, you need to be a member of the *eDiscovery Manager* role group. To delete AI activity data, you need to be assigned the *Search And Purge* role. The *Data Investigator* and *Organization Management* role groups have this role by default. For more information, see Assign eDiscovery permissions.
- You can remove up to 10 items per mailbox at one time. Because the search-and-removal capability is intended to be an event-response tool, the 10-item-per-mailbox limit helps ensure that AI activity data is removed quickly.

Step 1: Create a case in eDiscovery

To begin the search-and-deletion workflow, create an eDiscovery case to manage the process.

Step 2: Create a search in eDiscovery

After you create a case, create an eDiscovery search for the AI application data that you want to delete. The deletion process you perform in Step 5 deletes all AI-related items that are found in the search (within the 10 item per location limit).

Data sources for AI data

The following table lists the application and service sources for Copilot and other AI data. All user prompts and responses from AI applications are stored in a user's mailbox.

Expand table

For this type of Microsoft Copilot or other AI data...	Search this item class...
Microsoft Foundry	`IPM.SkypeTeams.Message.ConnectedAIApp.AzureAI.<AzureResourceName>`
ChatGPT Enterprise	- Local machine interactions: `IPM.SkypeTeams.Message.ConnectedAIApp.Connector.<ChatGPTEnterprise>` - Browser-based interactions: `IPM.SkypeTeams.Message.CloudAIApp.SaaS.<AppID>`
Copilot in Fabric	`IPM.SkypeTeams.Message.Copilot.Fabric.*`
Copilot personalization and memory	`IPM.Contact` For more information, see Copilot personalization and memory.
Copilot Studio	`IPM.SkypeTeams.Message.Copilot.Studio.*`
Excel	`IPM.SkypeTeams.Message.Copilot.Excel`
Facilitator	`IPM.SkypeTeams.Message`
Facilitator AI notes	`IPM.SkypeTeams.Message.TeamCopilot.AiNotes.Teams`
Loop	`IPM.SkypeTeams.Message.Copilot.Loop`
Microsoft 365 App	`IPM.SkypeTeams.Message.Copilot.M365App`
Microsoft 365 Copilot	`IPM.SkypeTeams.Message.Copilot.*`
Microsoft Copilot for Bing (Microsoft 365 Copilot Chat)	`IPM.SkypeTeams.Message.Copilot.BizChat`
Microsoft Entra apps	`IPM.SkypeTeams.Message.ConnectedAIApp.Entra.<AppID>`
Microsoft Forms	`IPM.SkypeTeams.Message.Copilot.Forms`
Microsoft Security Copilot	`IPM.SkypeTeams.Message.Copilot.Security.SecurityCopilot`
OneNote	`IPM.SkypeTeams.Message.Copilot.OneNote`
Other AI apps	- Local machine interactions: `IPM.SkypeTeams.Message.ConnectedAIApp.Connector.<AppName>` - Browser-based interactions: `IPM.SkypeTeams.Message.CloudAIApp.SaaS.<AppID>`
Outlook	`IPM.SkypeTeams.Message.Copilot.Outlook`

For this type of Microsoft Copilot or other AI data...	Search this item class...
PowerPoint	`IPM.SkypeTeams.Message.Copilot.Powerpoint`
SharePoint	`IPM.SkypeTeams.Message.Copilot.SharePoint`
Teams AI notes in Chat	`IPM.SkypeTeams.Message.TeamCopilot.AiNotes.Teams`
Teams Channel	`IPM.SkypeTeams.Message.Copilot.Teams`
Teams Chat	`IPM.SkypeTeams.Message.Copilot.Teams`
Teams Copilot Chat (Microsoft 365 Copilot Chat)	`IPM.SkypeTeams.Message.Copilot.BizChat`
Teams Meeting	`IPM.SkypeTeams.Message.Copilot.Teams`
Teams Microsoft 365 Chat (BF)	`IPM.SkypeTeams.Message`
WebChat	`IPM.SkypeTeams.Message.Copilot.WebChat`
Whiteboard	`IPM.SkypeTeams.Message.Copilot.Whiteboard`
Word	`IPM.SkypeTeams.Message.Copilot.Word`

> **Note**
>
> In step 4, you also need to identify and remove any holds and retention policies assigned to the mailbox that contains the type of Copilot or AI application data that you want to delete.

Tips for searching for Copilot and other AI data

To collect the most comprehensive set of AI data, use the *Item class* condition and select the **Copilot activity** option when you build the eDiscovery search query for this deletion workflow. This condition includes all Copilot and other AI application data. Add a date range or several keywords to narrow the scope of the search to items relevant to your search and delete investigation.

For Copilot memory data, consider the following when working in eDiscovery:

- Inferred memories can be deleted along with other mailbox items using Microsoft Graph.
- Copilot memories are stored as the *IPM.Contact* item class and they match as contacts in the eDiscovery condition builder.

- Deleting a conversation or message from Microsoft Purview or eDiscovery doesn't delete the associated Copilot memory.

How Copilot interactions appear in eDiscovery results

Each Microsoft 365 Copilot interaction is stored in the user's Exchange Online mailbox as an individual message-class item (see the Data sources for AI data table). When these items appear in eDiscovery search results, review sets, or exports (PST files, individual messages, or Microsoft Graph), the email-style headers reflect the two participants of the interaction:

Expand table

Interaction direction	`From`	`To`
User prompt to Copilot	The user's primary SMTP address	The Copilot application identity (for example, `Microsoft 365 Chat`, `Microsoft 365 Copilot Chat`, `WebChat`)
Copilot response to the user	The Copilot application identity	The user's primary SMTP address

The Copilot application identity is the service-principal display name of the Copilot app that handled the interaction. For example:

- **Microsoft 365 Chat** for Microsoft 365 Copilot Chat (BizChat) interactions.
- The **host app name** for in-app Copilot in Word, Excel, PowerPoint, Outlook, OneNote, Loop, and Teams.
- **WebChat** for browser-based Copilot WebChat interactions.

This identity isn't a routable mailbox and can't receive mail; it identifies the Copilot application that participated in the interaction. This representation is by design and is consistent across all tenants and across commercial and US Government clouds.

Step 3: Review and verify Copilot and other AI data to delete in eDiscovery

The deletion process in Step 5 deletes the items returned by the search. Review the search results to ensure that the search only returns the items that you want to delete.

You can use the search statistics (specifically the *Top Locations* statistics) to generate a list of the data sources that contain items returned by the search. Use the list of mailboxes from the Top

Locations statistics to remove hold and retention policies from the user mailboxes that contain search results.

Step 4: Remove holds and retention policies from data sources

Before you can delete Copilot or other AI application data from a mailbox, you have to remove any hold or retention policy that is assigned to a target mailbox. If you don't remove the hold or retention policy, the data you're trying to delete is retained.

Use the list of mailboxes that contain the Copilot and AI data that you want to delete to determine if there's a hold or retention policy assigned to those mailboxes, and then remove the hold or retention policy. Record each hold or retention policy that you remove so that you can reapply that policy to the affected mailboxes after the deletion is complete.

Step 5: Delete Copilot and other AI data in Microsoft Graph Explorer

> **Note**
>
> Because Microsoft Graph Explorer isn't available in some US Government clouds (GCC High and DOD), you must use PowerShell to accomplish these tasks. For more information, see Delete Copilot and other AI data with PowerShell.

Now you're ready to delete Copilot and other AI data from user mailboxes. Use Microsoft Graph Explorer to perform the following three tasks:

1. Get the ID of the eDiscovery case that you created in Step 1. This case contains the search created in Step 2.
2. Get the ID of the search that you created in Step 2 and verified the search results in Step 3. The query in the selected eDiscovery search returns the AI activity data to delete.
3. Delete the data returned by the search.

For information about using Graph Explorer, see Use Graph Explorer to try Microsoft Graph APIs.

> **Important**

To perform these three tasks in Graph Explorer, you might have to consent to the eDiscovery.Read.All and eDiscovery.ReadWrite.All permissions. For more information, see the "Consent to permissions" section in Working with Graph Explorer.

Get the case ID in Microsoft Graph Explorer

Perform the following steps to retrieve the eDiscovery case ID in Graph Explorer:

1. Go to https://developer.microsoft.com/graph/graph-explorer and sign in to Graph Explorer with an account that's assigned the *Search And Purge* role in the Microsoft Purview portal.
2. Run the following GET request to retrieve the ID for the eDiscovery case. Use the value `https://graph.microsoft.com/v1.0/security/cases/ediscoveryCases` in the address bar of the request query. Be sure to select **v1.0** in the API version dropdown list. This request returns information about all cases in your organization on the **Response preview** tab.
3. Scroll through the response to locate the eDiscovery case. Use the **displayName** property to identify the case.
4. Copy the corresponding ID (or copy and paste it to a text file). Save the eDiscovery case ID; you need it in the procedure for retrieving the search ID.

> **Tip**
>
> Instead of using Graph Explorer to obtain the case ID, you can open the case in the Microsoft Purview portal and copy the case ID from the URL.

Get the eDiscoverySearchID in Microsoft Graph Explorer

Perform the following steps to retrieve the eDiscovery search ID in Graph Explorer:

1. In Graph Explorer, run the following GET request to retrieve the ID for the search that you created in Step 2, and contains the items you want to delete. Use the value `https://graph.microsoft.com/v1.0/security/cases/ediscoveryCases/{ediscoveryCaseID}/searches` in the address bar of the request query, where *{ediscoveryCaseID}* is the CaseID that you obtained in the previous procedure.
2. Scroll through the response to locate the search that contains the items that you want to delete. Use the *displayName* property to identify the search that you created in Step 3. In the response, the search query from the search is displayed in the *contentQuery* property. The items returned by the selected search query are deleted in the purge request described in the following procedure.

3. Copy the corresponding ID (or copy and paste it to a text file). Save the eDiscovery search ID; you need it in the purge request to delete Copilot and other AI application data.

> **Tip**
>
> Instead of using Graph Explorer to obtain the search ID, you can open the case in the Microsoft Purview portal. Open the case and go to the Jobs tab. Select the relevant search and under Support information, find the job ID (the job ID displayed here's the same as the search ID).

Delete Copilot and other AI data in Microsoft Graph Explorer

Perform the following step to purge the items returned by the search:

1. In Graph Explorer, run the following POST request to delete the items returned by the search that you created in Step 2. Use the value `https://graph.microsoft.com/v1.0/security/cases/ediscoveryCases/{ediscoveryCaseID}/searches/{ediscoverySearchID}/purgeData` in the address bar of the request query, where *{ediscoveryCaseID}* and *{ediscoverySearchID}* are the IDs that you obtained in the previous procedures.

   If the purgeData POST request is successful, Graph Explorer displays an HTTP response code in a green banner stating that the request was accepted.

For more information about purgeData, see sourceCollection: purgeData.

Delete Copilot and other AI data with PowerShell

> **Note**
>
> Because Microsoft Graph Explorer isn't available in the US Government cloud (GCC, GCC High, and DOD), you must use PowerShell to accomplish these tasks.

You can also delete Copilot and other AI data with PowerShell. For example, to delete Copilot data in the US Government cloud, use a command similar to the following example:

PowerShell

```
Connect-MgGraph -Scopes "ediscovery.ReadWrite.All" -Environment USGov
```

PowerShell

```
Invoke-MgGraphRequest  -Method POST -Uri
'/v1.0/security/cases/ediscoveryCases/<ediscoverySearchID>/searches/<search
ID>/purgeData'
```

For more information about using PowerShell to delete Copilot and other AI data, see ediscoverySearch: purgeData.

Step 6: Verify Copilot and other AI data is deleted

After you run the purgeData POST request, Microsoft Purview removes the Copilot and other AI data returned by the search from the user's mailbox. The user doesn't receive any visible notification or confirmation that the data is deleted.

Deleted Copilot and other AI data moves to the *SubstrateHolds* folder, which is a hidden mailbox folder. Deleted data stays in the *SubstrateHolds* folder for at least one day. The next time the timer job runs (typically between 1-7 days), the timer job permanently deletes the data.

Step 7: Reapply holds and retention policies to user mailboxes

After you verify that the Copilot and other AI data is deleted, reapply the holds and retention policies to user mailboxes that you removed in Step 4.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/29/2026

Evaluate and refine search results in eDiscovery

Evaluating and refining your search results is an important step in your eDiscovery investigation. The search query you configure and the results that return help you determine if you discover items and information applicable to your investigation or if you need to modify your search to try to discover additional pertinent items. This initial search of items and initial review of information helps you determine what actions are required after you finalize your search parameters.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Evaluate search results

After you create and run a search, view the search statistics to help you verify whether relevant content is found and the content locations with the most hits. You can also review a sample of the search results to further help you determine if the content is within scope of your investigation.

Statistics dashboard

If you select **Statistics** as the initial result type for your search, the search automatically redirects you to this dashboard when the search results complete. If you're already familiar with previous versions of eDiscovery, the information on the **Statistics** tab is similar to collection estimates. The search results for the **Statistics** dashboard are included in the following sections:

- **Summary**: This section shows the number of search hits, locations, data sources, and the total file size of partially indexed items.
  - **Search hits**: Displays the total search hit count and volume from all items matching the query criteria from locations searched.
  - **Locations**: Displays the fraction of locations with hits out of all locations searched. The numerator shows the locations with hits and denominator shows the number of locations searched. Locations with errors are shown in red. To view full details on all the locations and associated hits and errors, select **Download report** to download the full .csv report.

Note that the displayed location count in this card might be different from the number of rows in locations CSV report. If multiple mailboxes share the same SMTP address, the location report might show several rows. This situation occurs because a single SMTP address might include different mailbox subtypes, such as primary, system, or archive. The row count in the CSV location report might exceed the count shown on the location card in the **Statistics** view. The **Statistics** view aggregates all mailboxes under the same SMTP address without distinguishing subtypes. For a more detailed breakdown of hits by sublocation, use the CSV report instead of the **Statistics** view. SharePoint and OneDrive locations can also have archive sites and are counted as separate locations in the report.

- **Retry failed locations**: This action is available when there are failed locations. *Retry failed locations* operation triggers a new search process targeting those locations with error and then aggregate the retried hit count with the original search process' hit count. After the retry process completes, the statistics view shows the new aggregated hit count.
- **Data sources**: Displays the fraction of data sources with hits out of all data sources searched. The numerator shows the data sources with hits and denominator shows the number of data sources included in the search. This data source is consistent with the data source in the search design flow and should match the number of people or groups included in the search. A tenant-wide data source of *All people and all groups* counts as a single data source.
- **Partially indexed items** or "Advanced indexed items hits": Displays the count and volume of partially and unindexed items returned as part of the search. This card displays partially indexed items information if you choose to include partially or unindexed items as part of the search configuration. If you choose to include partially and unindexed items and enable advanced indexing options, this card displays additional hits you get from advanced indexed items. The advanced indexed hit count is from a statistic sample on the partially indexed items, actual hits might be more and should be confirmed using the add to a review set and export search results actions. For more information, see Assess partially indexed items in your organization.

- **Search hit trends**: This section shows the following search result cards. The charts are interactive, hover to display section names, percentages, and item numbers. Select **View top 100** for more information about items included in each trend and to download the results to a .csv file:

  - **Top data sources**: Displays the top five data sources that make up the most search hits matching your query. The name of these data sources (names of users, groups, or organization-wide locations) are listed with the hit count. These data sources should match what you selected in the data sources workflow when building the search query.

- **Top sensitive information types (SITs)**: Displays the top five sensitive information types (SITs) in SharePoint files that are most often included in the search hits matching your query. Adding each SIT's count doesn't necessarily equate to the total count hits because a single item or document might contain more than one SIT type. For example, a document contains both a password and social security number (SSN). In this example, it's counted twice. Select **View top 100** to get a deeper understanding of the locations of these SIT counts to verify if they overlap or not.

- **Top keywords**: Query keywords, which resulted in the most search hits matching your query.

  > ⓘ Note
  >
  > To generate a keyword report in the statistics view, you must populate at least two or more keyword grids. If you enter only a single keyword, the total hit count shown reflects the results for that one keyword, and a keyword report isn't generated.

- **Top items types**: Most frequent item types within search hits matching your query. This count is determined by *itemClass* for Exchange content and *ContentType* for SharePoint content.

- **Indexing status**: Breakdown of unindexed (including partially indexed) and fully indexed data items.

- **Top communication participants**: Senders or recipients for emails, Microsoft Teams chats, and calendar invites in Exchange locations.

- **Top location type**: Hit count by location type (mailbox versus site).

Select **Regenerate view** to rerun the query and to review the most current results. Select **Download report** to combine all **Statistics** results into a single .csv file. When viewing the top 100 results for any trend area, select **Download report** for a .csv file of the top 100 results of the selected hit trend.

> ⓘ Note
>
> Search statistics expire after 14 days. Re-run search statistics for any searches older than 14 days to view current statistics.

Running the **Generate statistics** process is also required to enable sampling options when adding search results to a review set.

Understanding statistics and search results

> **Tip**
>
> For more information about how to use process report CSVs (Summary.csv, Locations.csv, Settings.csv) to investigate count differences, see **Process reports in eDiscovery**.

Depending on when you run a search in eDiscovery, the statistics for the search can show different results. For example, if you run two searches with the exact same conditions but at different times, you likely see different statistics results. These differences might occur for the following reasons:

- **Your organization is active**: Because you have active users in a production environment, data in your organization is constantly moved, added, deleted, and retired. The same search conditions run against the same locations likely return different search results because the data in those locations changed between the time you ran the searches.
- **Transient errors**: When you run a search (or export or add to a review set), transient processing errors might occur, especially for large sets of data. These errors often happen because of processing timeouts and can be mitigated by breaking up searches into smaller date ranges and exporting the data in parallel. Always try to break up searches into smaller sizes with more specific search conditions and more targeted with selected locations. This approach helps the process run more efficiently with less chance of errors.
- **Location access**: Some scenarios cause locations included in a search to be invalid, not accessible, or time out during processing. When you compare the results between two searches with the same conditions, make sure the locations you searched successfully match. For example, a search against 1,000 locations might have one failed location in the first run and no failed locations in the second run. This example means the first run searched only 999 locations successfully and the second run searched 1,000 locations. The difference of one location is the reason why search results between two runs are different. Use the *locations.csv* report for search, export, and add to review set processes to view a comprehensive report on what locations were successful and what locations failed. Rerun searches for any failed locations.
- **User running the search**: Depending on the user starting the search process, the user might or might not have the compliance boundary or compliance search filter applied. This filter either filters locations based on mailbox properties or filters content based on content path

(SharePoint sites). The results for the user might be limited if a compliance boundary or search permission filter is applied. For example, one user doesn't have a compliance boundary applied but a second user has a compliance boundary applied that restricts this user to user mailboxes and OneDrive sites to a specific region. A search by the first user returns all mailbox and OneDrive matches for the search conditions for all regions and a search for the second user returns only matches for mailboxes and OneDrive sites only for the allowed region.

- **Search results counts might vary across searches because of legal holds**: If you run the same search query at different times, the number of items in the search or export might differ. This difference can happen when items are edited or deleted between searches. For example, items under legal hold retain previous versions and might appear in later exports, while items not on hold might change or be removed if they no longer meet the hold criteria.
- **Search statistics are estimates**: Don't use these estimates to compare to OneDrive and SharePoint site storage. Estimates use index-based approximations, so the estimated eDiscovery content size can differ. Site storage often includes data not reflected in eDiscovery estimates, such as file versions and items in the Recycle Bin. For a view of site content, use the export process rather than the statistics dashboard size estimation.

Sample dashboard

If you select **Sample** as the initial result type for your search, you're automatically redirected to this dashboard when the search results complete. The search results for the **Sample** dashboard columns contain the following information for each item:

- **Subject/Title**: The subject or title of the items included in the sample.
- **Date**: The date the item was created or sent.
- **Sender/Author**: The sender or author of the item.

Samples let you inspect a representative subset of individual items and details for each item returned for the search. The number of samples per location and the number of sample locations defined in the search determine the number of sample items and location representation in the sample items.

Select a sample item to view the Source information for the item. If available for the item, this view displays a rich view of a selected item so that you can evaluate the relevancy of the item as it relates to the defined search data source and conditions.

> ⓘ **Note**

Sample items you generate are valid for 24 hours. If you generated the view more than 24 hours ago, regenerate the view to retrieve the latest samples matching your search query.

Select **Regenerate view** to rerun the query and review the most current results. Select **Download reports** to combine all **Sample** results into a single .csv file. Select **View settings** to view the settings applied to the sample view generation.

Refine search results

Based on the estimates and statistics that the search returns, you can edit and refine the search. Change the data sources that the search includes and change the search query to expand or narrow the search. You can update and run the search again until you're confident that the search results contain the content that's most relevant to your case.

After you're satisfied with the search results, you can take the following actions:

- Add search results to a review set
- Export the search results
- Create a hold

Before you open a support ticket about unexpected search results

If your search returns **zero results, fewer results than expected, or items you weren't expecting**, work through this checklist before opening a ticket. Most `RFC` (request for clarification) tickets filed against eDiscovery search are resolved by one of the items in the following list.

Expand table

#	Check	Where to verify
1	**Hold vs. search confusion**: A hold preserves content but doesn't surface it in search results until you run a search across the held locations.	Create holds in eDiscovery
2	**Indexing lag**: Newly ingested or recently modified items can take time to appear in the index. Re-run the search after the lag window.	Limits in eDiscovery — Indexing limits
3	**Partially indexed items**: Encrypted, oversized, or unsupported-format items aren't returned by content queries. You must export them separately as partially indexed items.	Partially indexed items in eDiscovery
4	**Location scope**: Confirm you add every relevant mailbox, site, group, and Teams location as a data source to the search. Add inactive	Data sources in eDiscovery

#	Check	Where to verify
	mailboxes and group members explicitly.	
5	**RBAC and permission filters**: Search permission filters and case-level RBAC silently scope results to what the searcher is entitled to see - even when running the search as an admin.	Search permission filtering and Assign permissions in eDiscovery
6	**Region or sovereign cloud**: Confirm the data source is in the same Microsoft 365 cloud and region as the case. You can't search cross-boundary data from the wrong cloud.	Compliance boundaries in eDiscovery
7	**Recipient expansion behavior**: Departed users whose Microsoft Entra ID object was deleted aren't expanded. You must add their historical SMTP addresses, aliases, and display names manually.	Recipient expansion
8	**Recent service changes**: Check the Microsoft 365 Message Center for service changes (for example, security-filter or Teams-private-channel changes) that affect search scope or results.	Microsoft 365 admin center → Message center

If you complete all eight checks and the result is still unexpected, capture the exact KeyQL of the search, the case ID, the data sources selected, and the run timestamp before opening a ticket. Support's first request is always for these four items.

Differences between statistics and export results

> **Tip**
>
> For guidance on interpreting the Summary.csv and Locations.csv reports that explain count differences for a specific process, see **Interpret count differences between search and export**.

When you run an eDiscovery search, the statistics return an estimate of the number of items (and their total size) that match the search criteria. However, the size and number of actual exported search results that you download differ from the estimated size and number of search results.

Several potential reasons explain these differences:

- **The way results are estimated**: The estimate provides an estimate (and not an actual count) of the items that meet the search query criteria. To compile the estimate of Exchange items, eDiscovery requests from the Exchange database a list of the message IDs that meet the search criteria. But when you export the search results, the search reruns and the actual

messages are retrieved from the Exchange database. Differences might result because of how the estimated number of items and the actual number of items are determined.

- **The way size of results is estimated**: During the estimate, size is approximated. The system gathers large numbers of items and sums the sizes up using approximations. You should consider the size estimate as an order of magnitude, not a specific measure of size. For example, a size estimate of 10 MB indicates the data is expected to be between 1 MB and 100 MB. The larger the number, the more variance there is in the estimate.
  - For Exchange-based content, file size is the size of the text in the message and attachment bytes. When exported, the format is converted to .msg and added to .pst or .zip files. Both these operations can significantly impact size.
  - For SharePoint-based content, the file size is approximate bytes of the file. In many cases for SharePoint-based data, file size can't be estimated during search.

- **Changes that happen between the time when estimating and exporting search results**: When you export search results, the search restarts to collect the most recent items in the search index that meet the search criteria. It's possible additional items were created, sent, or received that meet the search criteria in the time between when the estimated search results were collected and when the search results were exported. It's also possible that items that were in the search index when the search results were estimated are no longer there because they're purged from the content location before the search results are exported. To mitigate this problem, specify a date range for an eDiscovery search or place a hold on content locations so that items are preserved and can't be purged.

  Other issues that can result in differences between estimated and exported search results include:

  - An increase in items when using a date query. This issue is typically caused by the following two things:
    - **Hold versioning in SharePoint**: If a document is deleted from a site that's on hold and document versioning is enabled, all versions of the deleted document are preserved.
    - **Calendar items**: Accept and reject messages and recurring meetings automatically continue creating new items in the background with old dates.

  - With holds, there can be cases where the same item is preserved in a user's primary mailbox and in their archive mailbox. This situation can happen when a user manually moves an item to their archive.

  - Although rare, even in the case when a hold is applied, maintenance of built-in calendar items (which aren't editable by the user, but are included in many search results) might

be removed from time to time. This periodic removal of calendar items results in fewer items that are exported.

- **Unindexed items**: Items that are unindexed for search can cause differences between estimated and actual search results. You can include unindexed items when you export the search results. If you include unindexed items when exporting search results, there might be more items that are exported. This difference causes a difference between the estimated and exported search results.

  When using search, you can include unindexed items when you export search results. The number of unindexed items returned by the search is listed on the statistics page. When you export search results, you can choose to include or not include unindexed items. How you configure these options might result in differences between estimated and the actual results exported.

- **Document versions in SharePoint and OneDrive**: When searching SharePoint sites and OneDrive accounts, multiple versions of a document aren't included in the count of estimated search results. But you have the option to include document versions when you export the search results. If you include document versions when exporting search results, the actual number (and total size) of the exported items increases.

- **SharePoint folders**: If folders in SharePoint match a search query, for example, searching by date, the search estimate includes a count of those folders with the last modified date range (but not the items in those folders). When you export the search results, you have the option to choose export items inside subfolders of a matched folder or only include items that match search query. This option can affect the number of exported items. If a folder is empty, then the number of actual search results exported is reduced by one item, because the actual folder isn't exported.

- **SharePoint lists**: If the name of a SharePoint list matches a search query, the search estimate includes a count of all the items in the list. When you export the search results, the list (and the list items) is exported as a single CSV file. You can choose export setting that include list attachments, the attachments are exported as separate documents, which might increase the number of items exported.

- **Raw file formats versus exported file formats**: For Exchange items, the estimated size of the search results is calculated by using the raw Exchange message sizes. However, email messages are exported in a PST file or as individual messages. Both of these export options use a different file format than raw Exchange messages, which results in the total exported file size being different than the estimated file size.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/29/2026

Export search results in eDiscovery

After you run a search associated with a Microsoft Purview eDiscovery (preview) case, you can export the search results. When you export search results, you download mailbox items in PST files or as individual messages. When you export content from SharePoint and OneDrive sites, you export copies of native Office documents and other documents. When you export items, you prepare the items for download depending on the export options that you define. You can monitor the export's progress and download the results from the **Process manager** page. Larger exports might take longer to complete.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

For more information about new eDiscovery export enhancements and the retirement of the Microsoft Office 365 eDiscovery Export Tool, check out the following video:

https://learn-video.azurefd.net/vod/player?id=0d7dc0af-30dc-410b-a405-66f919e92cd1&locale=en-us&embedUrl=%2Fpurview%2Fediscsearch-export

Export considerations

This section describes important considerations when exporting content from eDiscovery. These considerations include requirements, limitations, and behaviors that might affect how content is accessed, exported, or logged.

- Search export packages expire 14 days after you create the export. After 14 days, the export and all data contained in the export package are automatically deleted and can't be recovered. To retain the exported data, download the export packages before the 14-day expiry. This expiry applies only to search exports, not to review set exports.
- To export content in eDiscovery, assign the eDiscovery user a Microsoft 365 Enterprise E3 or E5 license. When exporting SharePoint content, the account must also have a Microsoft SharePoint E3 license; otherwise, SharePoint items might not export as expected.
- You can't export files that are restricted by DLP action in eDiscovery. These files can appear in search statistics results but are excluded from the export.

- You can't export content stored in SharePoint sites that are in a **Locked** state with **NoAccess**. When a site is locked, eDiscovery can't retrieve items from that site, and the content is skipped during export.
- You can't export individual files if they're restricted by SharePoint access controls. To export restricted files, the collecting user must have direct access to the content (for example, be a site admin, site owner, or the owner or last modifier of the document). Files that don't meet these access requirements might appear in search results but are excluded from export.
- When you add SharePoint items to a review set or export them from eDiscovery, a **FileAccessed** audit event might be recorded in the Microsoft 365 audit log. This event reflects that the eDiscovery service accessed the content to perform the requested operation.
- For large exports - especially tenant-wide search exports that target a large number of locations - running a **Generate statistics** process before exporting can significantly improve performance. When search statistics are available, eDiscovery uses those results to optimize the subsequent export process by limiting processing to only the locations that contain search hits, instead of reprocessing all locations in scope. This location optimization helps reduce process duration and improves reliability for large tenants. If the statistics process is outdated or the search query or data sources changed, eDiscovery processes all locations again to ensure accuracy. To confirm if your export has location optimization applied, refer to settings.csv of the export process report.

> **Note**
>
> Even if the **Generate statistics** process has errored locations, those locations are still included in the subsequent export process. If you retry the **Generate statistics** process, the subsequent export process doesn't have location optimization applied.

Enable CMK (Customer Managed Key) for direct exports (preview)

Organizations that configure Customer Key through the Microsoft 365 Data-at-Rest Encryption Platform (MDEP) with a Data Encryption Policy (DEP) in policy assigned status can extend CMK protections to eDiscovery direct export packages. When enabled, exported investigation data is encrypted at rest using tenant-specific encryption scopes backed by customer-managed keys.

The following prerequisites are required to enable CMK for direct exports:

- The CMK direct export feature is enabled for your organization. To enable this feature, contact Microsoft Support or your Customer Success Account Manager (CSAM) and request CMK for eDiscovery direct exports for your tenant.
- Your organization configured Azure Key Vault following the standard Customer Key setup process. For more information, see Set up Customer Key.
- A Data Encryption Policy (DEP) is in *PolicyAssigned* state via the Microsoft 365 Data-at-Rest Encryption Platform (MDEP) onboarding cmdlets.

> **Note**
>
> This encryption applies to direct exports only. Data at rest in an eDiscovery review set isn't included at this time. Your tenant must remain CMK enabled at the time of the export.

Export search results

> **Tip**
>
> Want an interactive configuration guide experience? Check out the **Export data** guide.

To export search results, complete the following steps:

> **Tip**
>
> Your download speeds might be slower if you don't configure *Enable export download via pre-authorized link* in **eDiscovery export settings**.

1. Go to the Microsoft Purview portal and sign in with the credentials for a user account assigned eDiscovery permissions.
2. Select the **eDiscovery** solution card, and then select **Cases**.
3. Select a case, select the **Searches** tab, and then select a search query to export.
4. In the search query, select **Export**.
5. On the **Export** pane, configure export options.
6. Enter a name for the export in **Export name**. This field is required.

7. In **Export description**, enter an optional description for the export.

8. In **Tenant-wide source configuration** options, configure the organization-wide source "All people and groups" to include one or more user or content types:

    - **"All people and groups" should include unlicensed and on-premises users**
    - **"All people and groups" should include guest users**
    - **"All people and groups" should include shared Teams channels**
    - **"All people and groups" should include departed users**

9. In the **Select items to include in your export**, select one of the following options:

    - *Indexed items that match your search query*: This option exports only the indexed items that matched the query.

    - *Indexed items that match your search query and partially indexed items that might not match query*: This option exports only the indexed items that matched the query and the partially indexed or unindexed items that might not completely match the query.

    - *Partially indexed items that might not match query*: This option exports only the partially indexed or unindexed items that might not completely match the query.

      For these options including partially indexed items, you can scope the partially indexed items further by selecting *Exclude partially indexed items in locations without search hits*. When you select this option, the export returns partially indexed items only from mailboxes and sites where the indexed search query already returned at least one match.

      > **ⓘ Important**
      >
      > This option scopes by **location** (a mailbox or a SharePoint or OneDrive site), not by Kusto Query Language (KQL) conditions. Partially indexed items aren't evaluated against your search query. If a location is responsive (has any indexed hit), every partially indexed item in that location is included, regardless of any path, URL, or property restriction in your query. If your scope is a single location, this option has no effect. For details and examples, see Locations and the "in locations without search hits" option.

      > **ⓘ Note**

Partially indexed item settings made on the original search are for statistics and preview only. The authoritative include/exclude decision is made at export time using the options on this dialog. See **Search-time and export-time settings are independent**.

To apply your KQL conditions to partially indexed items, also select *Perform advanced indexing on partially indexed items*. Advanced indexing re-indexes the items so your query restrictions evaluate against them. Learn more about advanced indexing in eDiscovery.

10. In the **OneDrive and SharePoint items** section, select from the following document options:

    - **Select document versions**: Specify how many versions of SharePoint documents to export. Values are *Latest version only*, *Recent 10 versions*, *Recent 100 versions*, *All versions*.
    - **Select folder items**: Export items inside subfolders of a matched folder. Choose *Only include items that match search query* or *All items in folder (even if they don't match search query)*.
    - **Select items in lists and attachments (list expansion)**: Export files attached to SharePoint lists and their child items. Choose *All items in SharePoint list if any item matches search query*. If applicable, select *Include list attachments*.

11. In the **Messages and related items from mailboxes and Exchange Online** section, select one or more of the following options:

    - **Organize conversation into HTML transcript**: Contextual chat messages are threaded into HTML transcripts for ease of review and handling.
    - **Include Teams and Viva Engage conversations**: Include messages posted within a 12-hour window before and after each responsive item.
    - **Access links (cloud attachments) in messages**: Collect items from links to SharePoint or OneDrive.

      Important

      Including **cloud attachments** increases the size and quantity of your exported results. Items that are encrypted don't have embedded entities extracted including attachments. To collect cloud attachments from Shared Channels in

Teams, the Shared Channel SharePoint site must be included in the search to include cloud attachments.

To improve export performance, reduce the size of the export by using date ranges or other types of filters to decrease the total size of the search results. If an export runs for more than seven days without completing, it automatically cancels. To prevent cancellation, divide exports into smaller batches.

When you search for emails or other mailbox content that might include cloud attachments, the cloud attachment link to a storage location in SharePoint or OneDrive is considered as attachments of the message file rather than separate documents.

This behavior is similar to how a search returns emails with embedded local attachments. These attachments are retrieved as if they're part of the emails. Any compliance boundaries for SharePoint or OneDrive aren't reflected in the collection of cloud attachments included in responsive messages.

Select a version for access links. Values are *Latest version only*, *Recent 10 versions*, *Recent 100 versions*, *All versions*.

> **Important**
>
> If you have versions as shared feature enabled, this option retrieves the version of the document when it was shared.

12. In the **Export type** section, choose one of the following options:

    - *Export items report only*: Only the summary and item report are created. The various options for organizing data, folder and path structure, condensing paths, and other structures are hidden.

> **Note**
>
> When you use the *Exporting items report only* option, all string fields in the CSV report are limited to the first 255 characters. This limitation might truncate some items with long compound paths. If you need the full details, use the *Export items with items report* option.

- *Export items with items report*: This option exports items with the item report. Other export format options are available with this option in the **Export format** section.

13. In the **Export format** section, choose one of the following options:

    - *Create PSTs for messages*: This option creates .pst files for messages.
    - *Create .msg files for messages*: This option creates .msg files for messages.

14. In the **Package size settings** section, configure the maximum export package size the export process creates.
    a. If you select *Create PSTs for messages* in the previous step, set the maximum PST package size to 1 GB, 2 GB, 5 GB (default), or 10 GB.
    b. Set the maximum .zip package size to 2 GB, 10 GB (default), 20 GB, or 40 GB. The maximum .zip package size needs to be the same or greater than the maximum PST package size to contain the PST files in .zip packages when downloaded.

15. Select one or more of the following output package options. For details about each option, see Export package options.

    - *Organize data from different locations into separate folders or PSTs*: This option organizes data into separate folders for each data location. When exporting mailbox content, even if you select this export option, items from the primary mailbox and any associated archive mailboxes (including multiple archives) merge into a single PST file.
    - *Include folder and path of the source*: This option includes the original folder and folder path structure for items.
    - *Condense paths to fit within 256 characters*: This option condenses the folder path for each item to 256 characters or less.
    - *Give each item a friendly name*: This option creates a friendly name for each item.

16. Select **Export**.

17. When the export process starts, a confirmation dialog appears. Select **OK** to continue.

For more information about export file types and item.csv report fields, see Document metadata fields in eDiscovery.

Export package options

The export package options control how the system packages content in downloadable export packages. Depending on your needs and the post-export process an eDiscovery user needs to perform, adjust these options.

Organize data from different locations into separate folders or PSTs

This setting controls whether the system groups exported content by its original source. This grouping includes specific mailboxes or specific OneDrive or SharePoint sites, or it combines these items into a single container.

> ⓘ **Note**
>
> Exchange mailboxes store data from many applications, not only email. Some of the data from these applications might be converted to HTML and exported into separate `Items.zip` files. For example, the system might convert data from Microsoft Forms, Sway, and other apps. The list of applications in this category is evolving. These applications generally store data in the `/ApplicationDataRoot/` folder within a mailbox.

For Exchange Mailbox items

- **Selected**: The system exports each mailbox's content into a separate PST file named after the mailbox (for example, *john.doe@contoso.com.pst*).
- **Not selected**: The system merges all mailbox content across the selected data sources into a single PST file (for example, *exchange.pst*), subject to the maximum PST size limit. If the combined data exceeds this limit, the system creates extra PST files.

For OneDrive and SharePoint items

- **Selected**: The system exports content into separate folders per site. Each folder is named after the site and the system places all files from that site in each folder to preserve site-level organization.
- **Not selected**: The system exports all content into a single folder regardless of the originating site. Use the items.csv to understand the original source location and other relevant metadata for the item.

Include folder and path of the source

This setting determines whether the export preserves the original folder structure from the source. The source is the folder where items reside in an Exchange mailbox or OneDrive and SharePoint site.

For Exchange Mailbox items

- **Selected**: The exported PST retains the full folder hierarchy from the mailbox. For example, if a message is located in *Inbox* > *Legal* > *Q2*, this exact path is preserved inside the PST. This structure helps reviewers understand the original context and organization of the content.
- **Not selected**: All messages go to the root of the PST file, with no folder structure retained. While the messages are still exported, the original location within the mailbox is lost in the content itself. To find the original path of the item, review the items.csv in process report.

In both cases, the export process creates one PST file per mailbox if you select the **Organize data from different locations into separate folders or PSTs** option. The **Include folder and path of the source** option only affects the internal folder structure within each PST.

For OneDrive and SharePoint items

- **Selected**: Files are exported with the full folder path preserved. For example, if a file is stored in *Department* > *Legal* > *Contracts* > *Proposal.docx*, the entire path is recreated in the export package for this file. This structure mirrors the original site structure in which the file resides.
- **Not selected**: All files from a site go directly into a single folder for that site, without any subfolders. For example, if a file is stored in *Department* > *Legal* > *Contracts* > *Proposal.docx*, the path is *SharePoint* > *Items* > *Proposal.docx*. This structure flattens the folder hierarchy and might mix files from different folders. To find the original folder structure of the item, review the items.csv in process report.

Condense paths to fit within 259 characters

This setting helps prevent unzipping errors caused by long file paths during export. This problem can happen when you export content from OneDrive and SharePoint sites where deeply nested items are inside many layers of folders.

OneDrive and SharePoint sites might have deeply nested folder structures and long file names. When you export content and select the option to preserve folder and path structure, the resulting file paths can exceed the Windows file system limit of 259 characters. This limit leads to the following error message when unzipping the export package:

```
The file name(s) would be too long for the destination folder. You can shorten the file name and try again.
```

To help avoid this issue, use this option as needed:

- **Selected**: The export process automatically truncates folder paths - by shortening folder names or collapsing intermediate folders - so that the total path length for each file stays within the 259-character limit. This setting helps ensure that you can unzip and access the export package without errors.
- **Not selected**: The export process preserves the full original folder path. While this setting maintains the most accurate representation of the source structure, it increases the risk of hitting the path length limit - especially for deeply nested OneDrive and SharePoint items.

> ⓘ **Note**
>
> Different unzip utilities handle long file names and folder paths differently. The built-in extraction tool in Windows File Explorer might fail if the file name or path exceeds its character limit. Other non-Microsoft tools support longer paths and can extract these files successfully.

Give each item a friendly name

This setting controls whether the export item uses its original friendly name or the unique GUID generated by eDiscovery.

For Exchange Mailbox items

> ⓘ **Note**
>
> This setting doesn't apply when exporting Exchange items to PST format. Regardless of whether you select this option, email messages in PST exports always retain the original subject lines for the message title within the PST file.

When exporting as a .msg file:

- **Selected**: Exported .msg file uses the subject of the email.
- **Not selected**: Exported .msg file uses a system-generated GUID.

For OneDrive and SharePoint items

- **Selected**: Files are exported using the original file names. For example, a file named *Q2_Budget_Plan.docx* in SharePoint is exported with the exact same name in the export package.

    - If multiple files have the same name and the folder structure isn't preserved (the Include folder and path of the source option isn't selected), this situation might lead to file name collisions in the export folder. If this happens, the suffix is appended to the collided file names to differentiate the files.

    - When exporting versions, friendly names include version suffixes to distinguish between versions of the file. For example, *Project Plan.docx*, *Project Plan_v1.docx*, and *Project Plan_v2.docx*.

- **Not selected**: Files are exported using system-generated identifiers. This approach ensures uniqueness and avoids naming conflicts. For example, *Project Plan.docx* might be named b7f3a9c2-4e1e-4c9a-9f3e-2a1d8e7f9c1a.docx in the export package. Use the items.csv in process report to correlate the original title and other relevant metadata for the file.

View the export process status

To view the progress of the export in the Process manager, select **Process manager** and the **Export** process. You can track the process status and download export packages.

The **Overview** tab shows a progress bar that displays the current status, number of locations, and number of items for the export. An estimated completion time and current elapsed time are displayed in hours, minutes, and seconds. The **Overview** tab also contains the following export information:

- **Name**: The name of the export.
- **Completed**: The UTC date and time the export process completed.
- **Duration**: The duration of the export process.
- **Status**: The current status of the export process.
- **Process type**: The *Export* process type value.
- **Created**: The UTC date and time the export was created.
- **Created by**: The user who created the export.

The **Settings** tab displays a summary of the export options you selected for the export process.

View completed exports from a search

To view a list and details about completed exports from a search, select the **Exports** tab in a case. Search exports expire 14 days after creation. You can't download expired exports. On the **Exports** tab, you see the following details for each export process:

- **Export name**: The name of the export.
- **Status**: The current status of the export process.
- **Review set name**: The name of the review set from which the export was created (when you create the export from a review set).
- **Search name**: The name of the search from which the export was created (when you create the export from a search).
- **Created**: The UTC date and time the export was created.
- **Completed**: The UTC date and time the export was completed.
- **Duration**: The duration of the export process.
- **Created by**: The user that created the export process.

To delete an export that you created from a search, complete the following steps:

> ⓘ **Important**
>
> Deleting an export removes the export and all data contained in the export package. This action doesn't delete data in the original location.

1. In a case, select the **Exports** tab.
2. Select a completed export from a previous search. You can't delete in progress exports.
3. Select the ellipsis by the export name, and then select **Delete**.
4. On the **Delete export** dialog, select **Yes**.

Export packages

> ⓘ **Important**
>
> The export package links that the eDiscovery export process generates are for individual use only. Don't share these links. The behavior of these links can change. They might become invalid or inaccessible if others use them or if they're accessed outside the context of the current user. If you need to share data from the portal, use guest accounts.

After the export process finishes, you can download export packages. You can download search export packages for 14 days from the export creation time. After 14 days, the export packages are automatically deleted. The **Export packages** section on the **Overview** tab lists the export packages and shows the package name and size. To download a package, select the package and select **Download packages**.

> **Tip**
>
> You can configure the **Microsoft Purview eDiscovery API** to enable the programmatic download of export packages and the reports from an export process in an eDiscovery case.

Best practices for downloading export packages

To ensure a smooth download experience, consider the following best practices:

- **Use a supported, up-to-date browser**: Use the latest version of Microsoft Edge for the best export download experience. Older browser versions can have issues with the download mechanism.
- **Allow pop-ups from Microsoft Purview**: The download process triggers a pop-up window. Make sure your browser is configured to allow pop-ups from the Microsoft Purview site. For example, see instructions on how to configure pop-ups on Microsoft Edge.
- **Check your browser's automatic download settings**: If the **Automatic Downloads** setting is configured to *Ask* or *Block*, it might prevent multiple files from downloading at the same time. For export package downloads, configure this setting to *Allow* or adjust permissions for the Microsoft Purview portal site. For the Microsoft Edge browser, go to **Settings** > **Privacy, search, and services** > **Site permissions** > **All permissions** > **Automatic downloads** and confirm that the site is **Allowed to automatically download multiple files**.
- **Set your preferred download location**: By default, export packages save to your browser's default download folder. If you want to choose where files are saved, update your browser settings to prompt for a download location each time. See instructions for changing the downloads folder location in Microsoft Edge.
- **Check your organization's network configuration**: Ensure that your organization allows downloads from Microsoft Purview by verifying that all required endpoints and IP addresses are allowlisted. This verification includes enabling access to the Microsoft Purview export service and associated Azure storage endpoints. If your organization uses network security tools such as Zscaler, Microsoft Defender for Endpoint (MDE), or other proxy or firewall solutions, make sure these tools are configured to allow traffic to the export endpoints. In some cases, you might need to isolate specific endpoints from third-party security scanning to prevent download interference. For optimal performance and to avoid download interruptions, consider using a local drive (not a network share), and test your connection

throughput to Azure endpoints. For detailed allowlist guidance, see Microsoft 365 URLs and IP address ranges.

- **Test on a different network if downloads are slow**: If downloads are slow or stalling, try downloading on a different network (for example, a home network or mobile hotspot). If performance improves, the corporate network is likely the bottleneck due to firewall scanning, VPN overhead, or proxy throttling.

- **Enable parallel downloading in your browser settings**: Use parallel downloading to improve download speeds for large eDiscovery export packages. For example, use `edge://flags` in Microsoft Edge. This feature can significantly reduce wait times by allowing simultaneous file chunk downloads. This setting is also available for other chromium-based browsers.

- **Use a download manager**: A download manager can improve the reliability and speed of large file downloads. A download manager allows you to resume interrupted downloads and manage network fluctuations, which is especially helpful when downloading large export packages.

- **Split very large exports**: For exports of tens of gigabytes or more, consider breaking the export into smaller parts by splitting searches by date range or data source. Smaller export packages download faster and are less likely to stall due to network interruptions.

- **Allow first-party and third-party cookies** for the download sites. These sites include the following depending on the global location of the tenant:
  - **APC**: apc.proxyservice.ediscovery.svc.cloud.microsoft
  - **NAM**: nam.proxyservice.ediscovery.svc.cloud.microsoft
  - **EUR**: eur.proxyservice.ediscovery.svc.cloud.microsoft
  - **GCC**: gccm.proxyservice.ediscovery.gcc.svc.cloud.microsoft
  - **GCCH**: gcch.proxyservice.ediscovery.svc.usgovcloud.microsoft
  - **DOD**: dod.proxyservice.ediscovery.svc.usgovcloud.microsoft

- **Unzipping export files**: Data is compressed into a .zip file for improved performance and reliability. Avoid using the built-in Windows .zip utility. Use 7-Zip, WinZip, or similar tools for the best experience.

- **Verify timestamps after extraction**: Microsoft Purview eDiscovery doesn't modify or rewrite timestamps. The export preserves the authoritative date and time metadata from source Microsoft 365 workloads such as Exchange, SharePoint, OneDrive, and Microsoft Teams. When exporting content, these original timestamps are preserved in the export metadata

and within the native file formats. However, local ZIP extraction tools may modify certain file properties. Use the Items.csv report or internal file metadata to confirm original timestamp values.

- Disable **Tracking prevention for this site**: If enabled, it blocks downloads and should be turned off.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/23/2026

Add search results to a review set in eDiscovery

When you're satisfied with the results of a search and ready to review and analyze the results, add them to a review set in the case. You can add all search results or configure sampling to add a statistically representative subset based on confidence or percentage. Copying the original data to the review set also facilitates the review and analysis process by providing you with advanced analytics tools such as themes detection, near-duplicate detection, and email thread identification.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Create or add results to a review set

> **Tip**
>
> Want an interactive configuration guide experience? Check out the Create and manage review sets guide.

When you add the results of a search to a review set (the review sets in a case are listed on the **Review sets** tab), the following things occur:

- The search runs again. This step means the actual search results you copy to the review set might be different from the estimated results that the search returned.
- All items in the search results are copied from the original data source in the live services to a secure Azure Storage location in the Microsoft cloud.
- All items (including the content and metadata) are reindexed so that all data in the review set is fully searchable during the review of the case data. Reindexing the data results in thorough and fast searches when you search the data in the review set during the case investigation.
- A file encrypted with a Microsoft encryption technology and attached to an email message that's returned in the search results is decrypted when the email message and attached file

are added to the review set. You can review and query the decrypted file in the review set. You must be assigned the *RMS Decrypt* role to add decrypted email attachments to a review set. For more information, see Decryption in Microsoft Purview eDiscovery tools.

When you add content to a review set, the items and their metadata are copied to a Microsoft-managed Azure Storage location dedicated to the case.

Expand table

Question	Answer
Is review set content stored in Microsoft-managed Azure Storage?	Yes. The storage is Microsoft-owned and Microsoft-managed. Customers don't see, manage, or pay for these storage resources directly.
Is storage provisioned per review set?	No. Storage is provisioned **per case**. Multiple review sets within the same case share the underlying case storage, with logical separation so each review set's content remains independently queryable, taggable, and exportable.
How are different tenants and cases isolated?	Each case is isolated from other cases and other tenants at the storage boundary. Tenant separation is enforced by the service.
What region hosts the storage?	The storage is provisioned in the region that matches the case's geo, consistent with Microsoft 365 data residency commitments.
Does Customer Key apply to review set content?	Yes. If your tenant is configured with Customer Key, review set content at rest is encrypted with the customer-managed key.
Can customers access the storage directly?	No. There's no customer-accessible SAS token, account key, or storage URL for review set content. Access to review set content is only through the Microsoft Purview portal and approved export workflows.

If the search you're adding to review set is tenant wide or contains large number of data sources, running a **Generate statistics** process before adding to review set can significantly improve performance. When search statistics are available, eDiscovery uses the search statistics to optimize subsequent Add to review set process by limiting processing to only the locations that contain search hits, instead of reprocessing all locations in scope. This location optimization helps reduce process duration and improves reliability for large tenants. If the **Generate statistics** process is outdated or the search query or data sources have changed, eDiscovery processes all locations again to ensure accuracy. To confirm if your Add to review set process has location optimization applied, refer to settings.csv of the Add to review set process report. Additionally, completing the **Generate statistics** process is required to enable sampling options when adding search results to a review set.

> **Note**
>
> Even if the **Generate statistics** process has errored locations, those locations are still included in subsequent Add to review set process. If the **Generate statistics** process is retried, the subsequent Add to review set process doesn't have location optimization applied.

To add data to a review set, complete the following steps:

1. Select a search on the **Searches** tab, and then select **Add to review set**.
2. On the **Add items to review set** flyout, select **Add to a new review set** or **Add to an existing review set**.

   For a new review set, enter a name for the review set. To add data to an existing review set, select a review set from the drop-down choices.

   Adding data to a review set is a long-running process that includes gathering items from the original data sources in Microsoft 365 (for example, from mailboxes and sites), copying them to the Azure Storage location (this copying process is also called *ingestion*), and then reindexing the items. You can track the progress on the **Process manager** page. After the review set processing completes, select the **Review sets** tab in the case, then select the review set to start the process of filtering, reviewing, tagging, and exporting data in the review set.
3. In the **Select items to add to review set** section, choose one of the following options:

   - *Indexed items that match your search query*: This option exports only the indexed items that matched the query.
   - *Indexed items that match your search query and partially indexed items that might not match query*: This option exports only the indexed items that matched the query and the partially indexed or unindexed items that might not completely match the query.
   - *Partially indexed items that might not match query*: This option exports only the partially indexed or unindexed items that might not completely match the query.

   Scope the partially indexed items further by selecting *Exclude partially indexed items in locations without search hits*. With this option selected, only partially indexed items from location where there's already an indexed hit are returned. Additionally, user also has the option to *perform advanced indexing on partially indexed items*. For more information about advanced indexing, see Learn more about advanced indexing in eDiscovery.

4. In the **Configure sampling** section, choose one of the following options to control how many items are added to the review set. Sampling options are only available after you complete the **Generate statistics** process on the search. If statistics haven't been generated, are outdated, or the query or data sources changed since statistics were last run, sampling options are disabled and only the **Add all search results** option is available.

    - **Add all search results to review set** (default): Commits every item that matches the search query to the review set.
    - **Add samples to review set based on confidence**: Commits a statistically derived sample based on a **Confidence level** and **Confidence interval**. The confidence level (default: 95%) is the probability that your sample reflects the characteristics of the entire search result population. The confidence interval (default: 5%) is the range within which the true value is expected to fall. Higher confidence and tighter intervals produce larger samples.
    - **Add samples to review set based on percentage**: Commits a specified percentage of the total estimated items. Enter a value between 1 and 99 in the **Sample size** field (default: 50%). For example, if the search estimates 2,000 items and you select 50%, approximately 1,000 items are added to the review set.

    The actual number of items added to a review set may differ slightly from the calculated sample size. During the add-to-review-set process, items undergo expansion and extraction, which can result in minor variances between the expected and actual item counts.

5. In the **What to include for files from ODSP sites?** section, complete the following options:

    - **Select document versions**: Specify how many versions of SharePoint documents to collect. Choose from *Latest version only*, *Recent 10 versions*, *Recent 100 versions*, or *All versions*.
    - **Select folder items**: Choose one of following options to collect items inside subfolders of a matched folder:
        - **Only include items that matches query**
        - **Include all items in folder even if they don't match query**
    - **Select items in lists and attachments (list expansion)**: Choose of the following to collect files attached to SharePoint lists and their child items.
        - **Include all items in SharePoint if any item matches the query**
        - If applicable, select the **Include attachments of lists** checkbox.
    - **Convert supported file formats to HTML**: Select this option to convert supported cloud-native file formats, such as `.loop` and `.page` files (used by Copilot Notebook and Page content), to HTML when adding items to the review set. Converted content is

indexed and available for keyword search in the review set, making it easier to review and analyze. This option isn't selected by default.

6. In the **What to include for messages and related items from Mailboxes?** section, select the following options as applicable:

   - **Organize conversation into HTML transcript**: Contextual chat messages are threaded into HTML transcript of ease of review/handling.
   - **Include Teams and Viva Engage conversations**: Collect up to 12 hours of related conversations when a message matches a search.
   - **Access links (cloud attachments) in messages**: Collect items from links to SharePoint or OneDrive. Choose from *Latest version only*, *Recent 10 versions*, *Recent 100 versions*, or *All versions*.

     When you search for emails or other mailbox content that might include cloud attachments, the cloud attachment link to a storage location in SharePoint or OneDrive is considered as an attachment of the message file rather than separate documents.

     Treating cloud attachment links as message attachments is similar to how a search returns emails with embedded local attachments. These attachments are retrieved as if they're part of the emails. Any compliance boundaries for SharePoint or OneDrive aren't reflected in the collection of cloud attachments included in responsive messages.

     > **ⓘ Important**
     >
     > Items that are encrypted aren't supported for embedded entity extractions for example attachments. To collect cloud attachments from Shared Channels in Teams,the Shared Channel SharePoint site must be included in the search scope.

7. In **Tenant-wide source configuration** options, configure the organization-wide source **All people and groups** to include one or more user or content types:

   - **All people and groups** should include unlicensed and on-premises users
   - **All people and groups** should include guest users
   - **All people and groups** should include shared Teams channels
   - **All people and groups** should include departed users

8. Review your selections for the review set, and then select **Add to review set**. To cancel, select **Cancel**.

Optical character recognition

When you add search results to a review set, eDiscovery's optical character recognition (OCR) functionality automatically extracts text from images. It includes the image text with the data that it adds to a review set. You can view the extracted text in the Text viewer of the selected image file in the review set. This extracted text is then searchable within the review set. This feature lets you conduct further review and analysis on text in images. OCR supports loose files, email attachments, and embedded images. For a list of image file formats that support OCR, see Supported file types in eDiscovery.

> **Important**
>
> Microsoft Purview eDiscovery OCR extracts text from images only when content is added to a review set, it doesn't make images searchable during the search phase. To find image-based content, first search broadly then use OCR text in the review set. OCR in eDiscovery is separate from the tenant-level Microsoft Purview OCR setting (used by DLP, Insider Risk Management, and Retention). Enabling tenant-level OCR doesn't **enable eDiscovery OCR**. For more information about tenant-level OCR, see **Learn about OCR in Microsoft Purview**.

You need to enable OCR functionality for each case that you create in eDiscovery. For more information, see Configure search and analytics settings.

Conversation threading

Instant messaging is a convenient way to ask questions, share ideas, or quickly communicate across large audiences. As instant messaging platforms, like Microsoft Teams and Viva Engage groups, become core to enterprise collaboration, organizations must evaluate how their eDiscovery workflow addresses these new forms of communication and collaboration.

The conversation reconstruction feature in eDiscovery helps you identify contextual content and produce distinct conversation views. This capability allows you to efficiently and rapidly review complete instant message conversations (also called *threaded conversations*) that platforms like Microsoft Teams generate.

By using conversation reconstruction, you can reconstruct, review, and export threaded conversations. Use eDiscovery conversation reconstruction to:

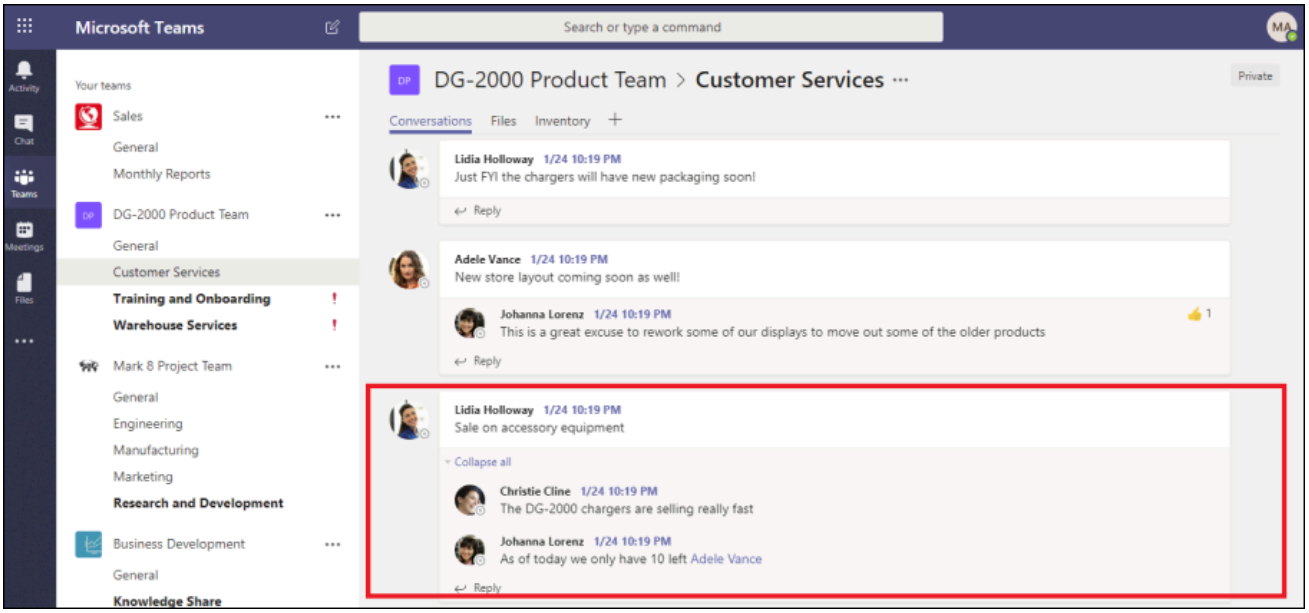
- Preserve unique message-level metadata across all messages within a conversation.
- Collect contextual messages around your search results.
- Review, annotate, and redact threaded conversations.
- Export individual messages or threaded conversations.

Conversation threading terminology

Here are a few definitions to help you get started using conversation reconstruction.

- **Messages:** Represent the smallest unit of a conversation. Messages might vary in size, structure, and metadata.

- **Conversation:** Represents a grouping of one or more messages. Across different applications, conversations might be represented in different ways. In some applications, an explicit action results from replying to an existing message. Conversations form explicitly as a result of this user action. For example, here's a screenshot of a channel conversation in Microsoft Teams.

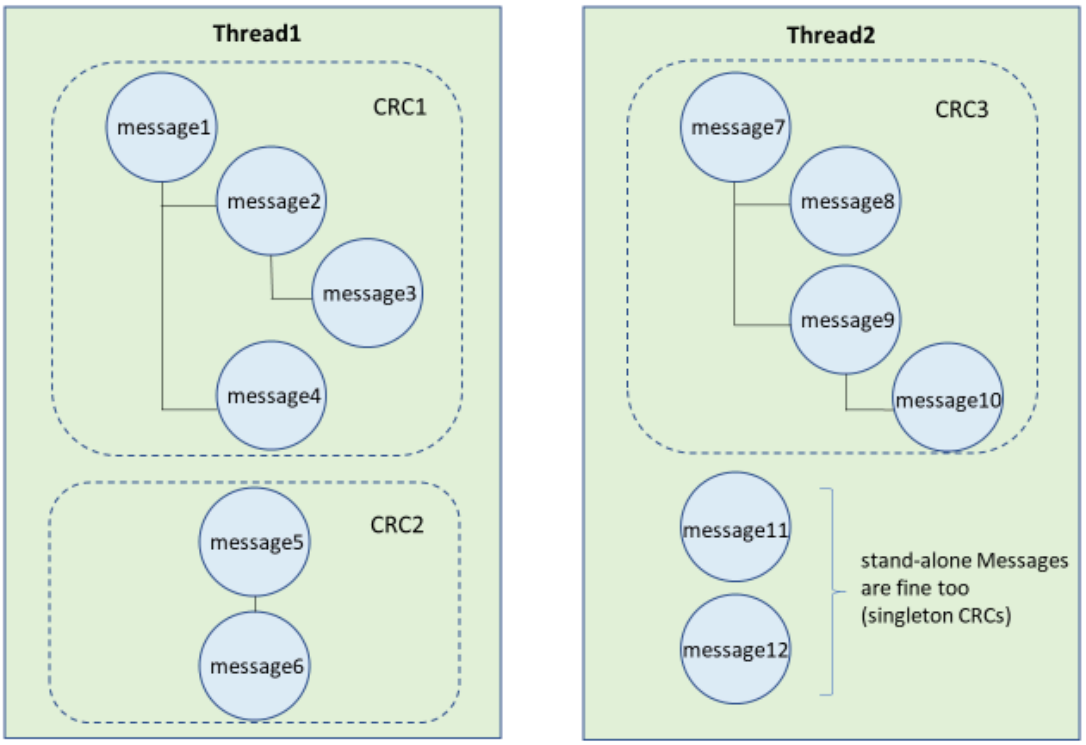


In other apps (such as group chat messages in Teams), there's no formal reply chain and instead messages appear as a "flat river of messages" within a single thread. In these types of apps, conversations are inferred from a group of messages that occur within a certain time. This "soft-grouping" of messages (as opposed to a reply chain) represent the "back and forth" conversation about a specific subject of interest.

Add conversations to a review set

After you review and finalize the search query, add the search results to a review set. Use the threaded conversations option to collect contextual messages from conversations that contain

items that match the criteria of the search. When you select the thread conversations option, the following things can happen:



1. Using a keyword and date range query, the search returns a hit on *Message 3*. This message is part of a larger conversation, illustrated by *CRC1*.
2. When you add the data into a review set and enable the conversation retrieval options, eDiscovery goes back and collects other items in *CRC1*.
3. After you add the items to the review set, you can review all the individual messages from *CRC1*.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/29/2026

Manage review sets in eDiscovery

Review sets are a static set of documents where you can analyze, query, view, tag, and export data in a case. To view all review sets in a case, select the **Review sets** tab. You can't delete items from a review set, and you can't delete review sets from a case. To delete a review set and its data, you must delete the eDiscovery case that contains the review set.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Review set summary page

The **Review sets** dashboard lists all the review sets in a case. This dashboard lets you create review sets and displays information about review sets in the case. You can filter and group the review sets. The **Review sets** dashboard contains the following information and controls:

- **Review set**: The name of the review set.
- **Size**: The total file size of the items contained in the review set.
- **Created by**: The user that created the review set.
- **Last modified**: The date and time the review set was last modified. Select **Time zone** to switch between using local time and Coordinated Universal Time (UTC).
- **Modified by**: The user that modified the review set.
- **Searches (sources)**: The number of search sources contained in the review set.
- **Exports**: The number of exports from the review set.
- **Description**: The description information for the review set.

To customize the columns display on the **Review sets** dashboard, select **Customize columns** to choose the columns to display or drag and drop the columns in the list to reorder. To search for a specific review set, enter a keyword in the **Search** field.

To download the list of review sets and the column information, select **Download list** to create a .csv file containing this information.

Grouping review sets

When you view a large list of review sets in a case, group the review sets by user information. Select **Group** to group review sets by the following parameters:

- **None**: Review sets aren't grouped on the **Review sets** dashboard.
- **Modified by**: Review sets are grouped by the *Modified by* values.
- **Created by**: Review sets are grouped by the *Created by* values.

Review set management

To open a review set and start working with review set items, select the review set. The review set opens and contains the following tabs and information to help you manage the review set.

All items

> **Important**
>
> Review sets for limited format cases use the legacy eDiscovery experience. Some features aren't available for these cases in the new eDiscovery experience.

After selecting a review set, you're ready to start working with the items contained in the review set. On the **All items** tab, use the controls on the command bar to take the following actions in the review set:

- Search content with the condition builder in a review set. Use the condition builder for visual, no-code filtering with standard conditions like keywords, dates, senders, and file types.
- Search content with the Advanced review set explorer (preview). Use the Advanced review set explorer (preview) for KQL-based queries that support complex filtering, aggregation, pattern matching, and data visualization.
- Group and view items in a review set
- Tag documents in a review set
- Run analytics on a review set
- Export items from a review set
- Add items from a review set to another review set
- Use the Query Report to create keyword reports for KeyQL queries (preview)

> **Tip**

Do you prefer an interactive configuration guide experience? Check out the **Create and manage review sets** guide.

Review set overview

Select **Overview** to display summary information about the review set:

- **Total extracted documents**: The total number of items added to the review set. This number indicates the sum of parent items and child items added to the review set.
- **Total processed items**: The total number of processed items added to the review set.
- **Total number of failed extracted documents**: The total number of documents that failed extraction to the review set.

Pagination

Pagination for review set items is enabled by default in review sets. When pagination is enabled, review set items are displayed as multiple pages of 50 items each. To view specific pages, use the forward and back controls to navigate pages or enter a specific page number. To disable automatic pagination, select **Manage** > **Turn pagination off**.

Advanced review set explorer (preview) tab

The **Advanced review set explorer (preview)** tab provides a Kusto Query Language (KQL) editor for running advanced queries directly against review set data. Use this tab for complex filtering, aggregation, pattern matching, and data visualization.

For detailed guidance, see Search content in a review set with the Advanced review set explorer (preview).

Dashboard (preview) tab

The **Dashboard tab** provides a visual experience to help you quickly understand the composition and trends within the review set. You don't need to create custom queries because the dashboard automatically creates interactive charts and summaries.

> **Note**
>
> The dashboard doesn't render if the review set is empty. Ensure the review set contains data before using the dashboard.

Summary cards

Summary cards provide at-a-glance metrics for:

- Total items in the review set
- Items by type (conversation vs documents)
- Items breakdown by analytics results (items to review excluding duplicates, or immaterial items) if analytics has been run.

Interactive charts

Visualize review set data by using the following charts:

- **Top Communication Participants**: Senders or recipients for emails, Teams chats, and calendar invites in Exchange locations.
- **Top Data Sources**: Data sources that make up the most search hits matching your query.
- **Top item classes**: Top item classes that make up search hits matching your query.
- **Top file extension**: Top file extensions that make up search hits matching your query.

Process manager

Select **Process manager** to display information about processes performed on the review set. For example, if you export data from the review set, an *Export* process appears as an item in the **Process manager** list.

- **Process type**: The type of process.
- **Status**: The status of the process.
- **Export name**: The name of the export. For non-export process types, this value is blank.
- **Created**: The date and time the process was created.
- **Completed**: The date and time the process was completed.
- **Duration**: The duration of the process.
- **Created by**: The user that created the process.

To download the list of processes and the column information, select **Download list** to create a .csv file containing this information.

To view information about all eDiscovery processes, see Use the Process report in eDiscovery.

ⓘ **Note:** The author created this article with assistance from AI. Learn more

Last updated on 04/02/2026

Search content in a review set with the query condition builder

In most cases, it's useful to dig deeper into the content in a review set and organize it to facilitate a more efficient review. By using filters and queries in a review set, you can focus on a subset of documents that meet the criteria for your review. This article explains how to use the query condition builder to create and run filter queries, work with condition types and load sets, use KeyQL query language, and save queries for reuse.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Use the query condition builder

To use the query condition builder in a review set, complete the following steps:

1. Go to the Microsoft Purview portal and sign in with the credentials for a user account assigned eDiscovery permissions.
2. Select the **eDiscovery** solution card, then select **Cases** in the left navigation.
3. Select a case, then select the **Review sets** tab.
4. Select a review set, then select the **All items** tab.
5. On the **All items** tab, select **Query**.
6. Add conditions as needed. For more information about using conditions, see Use the condition builder to create search queries in eDiscovery.

Query content in a review set

The eDiscovery review set provides a querying experience that you can use to build flexible queries for review sets. The querying experience enables you to:

- Quickly search for review set content matching certain conditions.
- Create complex queries by using *AND* or *OR* conditions in KeyQL.
- Manage saved filters without going to another area and easily change your queries by loading saved queries.

> **ⓘ Important**
>
> Review sets have pagination turned on by default and can display a maximum of 1,000 items per page. Use default or custom filters to adjust the displayed items as needed. Matched item counts are estimates. You can turn off pagination through **Manage** > **Turn pagination off**.

To query content in your review set, use the following controls:

- Select **Query** in the toolbar or the blue bubble above the toolbar. Any applied query also appears in the query bubble above the toolbar.
- **Keyword** condition automatically populates for any new query.
- **AND/OR**: These conditional logical operators enable you to select the query condition that applies to specific filters.
- **Add conditions**: Enables you to add multiple conditions to your query.
- **Select an operator**: Depending on the selected filter, the operators compatible with the filter are available to select. For example, if you select the *Date* filter, you see the available operators *Before*, *After*, and *Between*. If you select the *Size* filter, you see the available operators *Greater than*, *Greater or equal*, *Less than*, *Less or equal*, *Between*, and *Equal*.
- **Value**: Depending on the selected filter, the values compatible with the filter are available. Additionally, some filters support multiple values and some filters support one specific value. For example, if you select the *Date* filter, you select date values. If you select the *Size (in bytes)* filter, you select a value for bytes.
- **Clear a condition value:** To clear the value inside a condition, select the X to the right of each condition line.
- **Remove a filter condition**: To remove an individual filter or subgroup, select the **Delete** icon to the right of each condition line.
- **Clear all**: To clear the entire query of all conditions and their values, select **Clear all**.
- **Save**: Saves the current conditions into a saved query by giving it a name.
- **Load saved filters:** Saved filters allow you to load any saved queries and overwrite existing query that is built.
- **Run query:** For any manually built conditions or opened saved filters, select **Run query** to apply the query to the review set. The last run query is retained and is available the next time you open the review set, even after you exit.
- **Hide and show query**: Use the hide and show button to expand and collapse the query section.

Understand condition types

Every searchable field in a review set has a corresponding condition that you can use to filter items based on a specific property.

There are multiple types of filters:

- **Free-text**: A free-text condition is applied to text fields such as *Subject*. You can list multiple search terms by separating them with a comma.
- **Date**: A date condition is used for date fields such as *Last modified date*.
- **Search options**: A search options condition provides a list of possible values (each value is displayed with a checkbox that you can select) for particular fields in the review. This condition is used for fields, such as *Sender*, where there's a finite number of possible values in the review set.

> ⓘ **Note**
>
> Some search options let you enter free-text to match multiple values. For example, when using the **Sender** condition with **contains any**, entering *john* matches items where the sender is *johndoe@contoso.com*, *johnsmith@contoso.com*, or *johnwei@gmail.com*. Make sure the entered value appears as a pill for the free-text input to apply.

- **Keyword**: A keyword condition is a specific instance of free-text condition that you can use to search for terms. You can also use Kusto-like query language in this type of condition. For example user can type *red AND blue* in a keyword condition grid, run the query and get all items in the review set that has both the word "red" and the word "blue". The matched keywords are also highlighted in the **plain text** view in the preview pane.

Use load sets to filter by collection

A load set is an instance of adding data to a review set. For example, if you add the results of two different searches to the same review set, each one represents a load set. To filter review set items by a specific load, use the **Load Id** condition in the query builder:

1. In the review set, select **+ Add conditions** in the query area.
2. Select **Load Id** as the condition.
3. Set the operator to **Equals any of**.
4. Start typing in the search field. Suggested load IDs from the review set appear automatically.

5. Select one or more load IDs to filter by, then run the query.

This filter is useful when you want to isolate items from a specific collection or compare data across different loads added to the same review set.

To download the list of load sets and the column information, select **Download list** to create a .csv file containing this information.

Save and manage filter queries

After you build your conditions, save the condition combination as a filter query. This saved filter query lets you apply the same query in future review sessions.

To save, select **Save** and give it a name. You or other reviewers can run previously saved filter queries by selecting the **Saved filter** dropdown and selecting a filter query to open and select **Run query** to apply review set documents.

To edit or delete a saved filter query, select **Saved filter** and mouse over the saved filter and select **Edit** or **Delete** options for the saved filter query.

Use query language support for KeyQL and Keyword filters

When you use the *KeyQL* or *Keyword* filters, you can use a Kusto-like query language to build your review set search query. The query language for the KeyQL and Keyword filters supports standard Boolean operators, such as **AND**, **OR**, **NOT**, and **NEAR**. It also supports a single-character wildcard (?) and a multi-character wildcard (*).

> **Note**
>
> Review filters only support wildcards (? or *) on a single term. Using wildcards in searches on phrases that consist of multiple terms isn't supported.

Scenario examples

The following examples show common review set filtering scenarios using the query condition builder.

Filter for untagged items in a review set

An eDiscovery administrator needs to create a query to find all items in the review set without any tagging applied. For this example, the administrator creates the following review set query:

1. For the first condition, the administrator selects *Add conditions* and searches for the *tag* condition. The *Tags* condition is displayed as a matching option. The administrator selects it and selects *Apply*.
2. The administrator then selects *Is empty* operator for the *Tags* condition.
3. The administrator selects *Run query*.

The review set list is updated to show all items that don't have any tags applied.

Filter for native file type items in a review set

An eDiscovery administrator needs to create a query to find all items in the review set that are a certain type, such as .csv, .msg, or .pdf. For this example, the administrator creates the following review set query:

1. For the first condition, the administrator selects *Add conditions* and searches for the *file* in condition search. The condition *Native file extension* is one of the options displayed in the search results. The administrator selects it and selects *Apply*.
2. The administrator then selects the *Equals any of* operator.
3. The administrator selects the value box and chooses the file types in the dropdown such as csv to include in the query.
4. The administrator selects *Run query*.

The review set is updated and only the items that match the selected file types are displayed.

Filter partially indexed items

If you select the option to include partially indexed items when you add the search to a review set, you probably want to identify and view those items. You can determine if an item might be relevant to your investigation and whether you need to fix the error that resulted in the item being partially indexed. Use the KQL (Keyword Query Language) condition to find partially indexed items.

1. Select **Add conditions** in the query building space.
2. Find the condition *KQL* and select **Apply**.
3. Keep the operator as *Equal* and type *AddedBy* in the KQL editor box. You can filter out items added to the review set by applying the KQL of `AddedBy=UnindexedQuery`.
4. Select **Run query**.

The review set is updated and only the items that match the partially indexed query are displayed.

Filter documents by theme

Filtering documents by theme can save time when reviewing documents. For example, if you're looking for documents that discuss a particular subject, you can filter the documents by the dominant theme that relates to that subject. You can also filter documents by other themes in the theme list to find documents that are similar to a document that you're interested in. To display the themes for a document as a column in the document list for the review set, select **Customize columns** and select **Dominant theme** and **Themes list**.

To filter documents by theme, complete the following steps:

1. In a review set, add **Dominant theme** condition.
2. Select an operator to use with the **Dominant theme** and define the value to use with the operator.
3. Use an additional **Themes list** condition and the operator and values to apply to this filter. You can configure the **AND** and **OR** operators to filter documents by a combination of the **Dominant theme** and **Themes list** values.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/29/2026

Search content in a review set with the Advanced review set explorer (preview)

The **Advanced review set explorer (preview)** in Microsoft Purview eDiscovery lets reviewers run real-time big data analytics on review set content by using a defined subset of Kusto Query Language (KQL). With the Advanced review set explorer (preview), you can query, filter, and visualize review set data directly in the Microsoft Purview portal without modifying the underlying content. Queries run against the full set of indexed items and return either tabular results or chart visualizations.

Use the Advanced review set explorer (preview) to:

- Identify top item types, spot patterns, and find trends within the review set.
- Build powerful queries with complex filtering, pattern-based text extraction, and data format parsing.
- Analyze and find key information specific to your case or organization.
- Visualize query results by using charts for a comprehensive understanding of review set data.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Open the Advanced review set explorer (preview)

To open the Advanced review set explorer (preview), complete the following steps:

1. Go to the Microsoft Purview portal and sign in with the credentials for a user account assigned eDiscovery permissions.
2. Select the **eDiscovery** solution card, and then select **Cases** in the left navigation.
3. Select a case, and then select the **Review sets** tab.
4. Select a review set, and then select the **Advanced review set explorer (preview)** tab.

From here, you can write and run KQL queries directly against the review set data.

> **Tip**
>
> If you prefer a visual, no-code approach to filtering review set content, see **Search content in a review set with the query condition builder**.

Navigate the Advanced review set explorer (preview)

The Advanced review set explorer (preview) includes a left navigation pane and a results area to help you build, discover, and run KQL queries.

> **Tip**
>
> If you're new to KQL, start with the **Getting started** tab to explore basic query patterns, then use the **Sample queries** in the left pane for more targeted analysis.

Left navigation pane

The left navigation pane provides three collapsible sections that let you explore the review set schema and quickly insert query elements into the editor:

- **Schema**: Displays the available table (`ReviewSetTable`) for your review set. Select the table name to insert it into the query editor.
- **Operators**: Lists all supported KQL operators, including count, distinct, extend, limit, mv-expand, order, project, project-away, project-keep, render, sort, summarize, take, top, and where. Select any operator to add it to your query.
- **Sample queries**: Provides ready-to-use query templates for common review set analysis tasks, such as **FileClass count as piechart**, **FileType count**, **ItemClass count**, **Participants count**, **SensitivityLabel count**, and **Source count**. Select a sample query to load it directly into the query editor.

Getting started tab

The **Getting started** tab below the query results area provides a quick reference for building queries. It includes selectable examples organized into two categories.

Select any query name to load the example directly into the query editor:

Expand table

Category	Queries	Description
Basic queries	Limit	Shows 10 results from the specified table.
	Where	Filters the result set by date and file names ending with ".jpg".
	Count	Counts the number of rows that match the specified filter.
	Top	Arranges results by date and shows the first 10 rows.
	Project	Shows only the file name, file type, and date columns in the results.
Advanced queries	Summarize	Shows the top data sources and counts the number of items for each source.
	Extend	Adds a column to combine title and author values.
	Mv-expand	Expands the Participants list and counts occurrences per participant.
	Render	Counts items by file class and renders the results as a pie chart.

How the Advanced review set explorer (preview) works in a review set

Each query runs against the full review set's data and returns:

- A **tabular list** of matching items, or
- A **chart visualization**, depending on the query.

Queries follow standard KQL pipeline syntax, where each line builds on the results of the previous step. The portal automatically adds the `ReviewSetTable` table name as the first line. Each subsequent query line must start with `\|`.

For example:

Kusto

```
ReviewSetTable
| where SubjectTitle has "Case"
| summarize count() by FileClass
```

> ⓘ Note
>
> All queries in the Advanced review set explorer (preview) run against the current review set's indexed items. You don't need to specify a table name. The review set acts as the default

data source.

Best practices

Use the following best practices when writing queries in the Advanced review set explorer:

- Start with `where` filters to narrow data before using `summarize` or `render`.
- Use `take` to validate query logic before running full aggregations.
- Avoid overly broad regex patterns on large review sets.
- Use `where isnotempty(`*ColumnName*`)` to exclude items with empty values for a property, which helps avoid unexpected blank rows in your results.

Other considerations

The Advanced review set explorer (preview) doesn't support all Kusto operators and properties. Keep the following considerations in mind when you create queries:

- Queries are read-only and don't modify review set content. For more information about working with review sets, see Manage review sets in eDiscovery.
- Results reflect the current indexed state of the review set. If you add more data to the review set, the query might return different results.
- Not all file types are queryable. For more information, see Supported file types in eDiscovery.
- If the query returns more than 10,000 results, it returns only the first 10,000 results. Modify the query to narrow down the results.

For Kusto syntax and references, see the following articles:

- Syntax conventions for reference documentation
- Best practices for Kusto Query Language queries

Supported basic operators

> **Tip**
>
> For a complete list of available properties you can use in queries, see **Document metadata fields in eDiscovery**.

Where operator

Use the `where` operator to filter items based on metadata such as subject, keywords, item class, dates, or participants.

The following example filters the review set to return only items whose **Subject** contains the text "Day".

Kusto

```
ReviewSetTable
| where SubjectTitle has "Day"
```

This query filters the review set to include all calendar-related items whose **ItemClass** begins with `IPM.Appointment`, which covers standard appointments and extended classes like `IPM.AppointmentSnapshot.SkypeTeams.Call`.

Kusto

```
ReviewSetTable
| where ItemClass startswith "IPM.Appointment"
```

Project operator

Use `project` to return only the properties you care about. For example, the following query performs two actions:

- **Filters items by subject text**: The `where SubjectTitle has "Day"` clause limits the results to only those items whose **SubjectTitle** contains the word "Day".
- **Projects a single column**: The `project SubjectTitle` clause then returns only the **SubjectTitle** property for those filtered items, removing all other columns from the output.

Kusto

```
ReviewSetTable
| where SubjectTitle has "Day"
| project SubjectTitle
```

Project-away operator

The `project-away` operator removes specified columns from the query results while keeping all other columns. Use it when you want to exclude a few properties rather than listing every column you want to keep.

The following example returns all properties for items in the review set except **CompoundPath** and **InternetMessageId**, which are excluded from the output.

Kusto

```
ReviewSetTable
| where SubjectTitle has "Day"
| project-away CompoundPath, InternetMessageId
```

Project-keep operator

The `project-keep` operator keeps only the columns that match a specified pattern or list, removing everything else. Unlike `project`, `project-keep` preserves the original column order and supports wildcard patterns.

The following example keeps only columns whose names start with "Sender" or "Recipient", making it easy to focus on communication participants without listing each property individually.

Kusto

```
ReviewSetTable
| where SubjectTitle has "Day"
| project-keep Sender*, Recipient*
```

Mv-expand operator

The `mv-expand` operator expands multi-value properties (such as lists of participants or recipients) into separate rows, with one row per value. This is useful for analyzing individual entries within fields that contain multiple values.

The following example expands the **Participants** field so each participant appears on its own row, then counts the number of items associated with each participant. This helps identify the most active participants in the review set.

Kusto

```
ReviewSetTable
| mv-expand Participants
| summarize count() by tostring(Participants)
```

Extend operator

The `extend` operator adds one or more calculated columns to your query results by creating new values derived from existing fields without altering the underlying data.

The following example classifies review set items into four size categories - **Small (<1MB)**, **Medium (1–5MB)**, **Medium-Large (5–10MB)**, and **Large (>10MB)** - by creating a new column `SizeCategory` using the `case()` function, and then aggregates the count of items in each category by using `summarize`. It helps investigators quickly understand the distribution of file sizes in the review set for risk assessment or prioritization.

Kusto

```
ReviewSetTable
| extend SizeCategory = case(Size > 10485760, "Large (>10MB)", Size >= 5242880 and
Size <= 10485760, "Medium-Large (5–10MB)", Size >= 1048576 and Size < 5242880,
"Medium (1–5MB)", "Small (<1MB)")
| summarize count() by SizeCategory
```

Summarize operator

The `summarize` operator aggregates data by applying functions like `count()`, `sum()`, or `avg()`, and groups results based on specified columns.

The following example counts items grouped by **SubjectTitle** for items containing the word "Day".

Kusto

```
ReviewSetTable
| where SubjectTitle has "Day"
| summarize count() by SubjectTitle
```

Render barchart, render columnchart, render piechart

The `render barchart` operator takes the tabular results of your query and displays them as a bar chart. This makes it easier to visualize the distribution or comparison of aggregated values.

The following example renders subject title counts as a bar chart.

Kusto

```
ReviewSetTable
| where SubjectTitle has "Day"
```

```
| summarize count() by SubjectTitle
| render barchart
```

Take (limit) operator

The `take` operator limits the query output to a specified number of rows. It returns only the first *n* results from the dataset. The `limit` operator is a synonym for `take` and you can use them interchangeably.

The following example returns just the first 10 matching items, so you can validate query logic or spot-check content before running a full query. This approach helps when dealing with large review sets or complex filters, where results can take longer to compute and load without `take` limiting the output.

Kusto

```
ReviewSetTable
| where SubjectTitle has "day"
| project SubjectTitle, SenderDomain, Custodian, Size
| take 10
```

Sort operator

The `sort` operator orders the query results based on one or more specified columns in ascending or descending order.

The following example query calculates the total number of items for each **NativeFileExtension** in the review set by using `summarize count()`. Then, it sorts those extension groups in descending order of item count so you can see which file types are most prevalent.

Kusto

```
ReviewSetTable
| summarize ItemCount = count() by NativeFileExtension
| sort by ItemCount desc
```

Distinct operator

The `distinct` operator returns unique values for the specified columns by removing duplicate rows from the query results.

The following example uses `distinct` to remove duplicates so you only see unique custodians, and then uses `sort by Custodian asc` to order them alphabetically, making it easy to scan or export for reporting.

Kusto

```
ReviewSetTable
| distinct Custodian
| sort by Custodian asc
```

Count operator

The `count` operator returns the total number of rows in the query result set, so you can quickly see how many items are in the set.

The following example counts the number of unique subject titles that contain the word "Day".

Kusto

```
ReviewSetTable
| where SubjectTitle has "Day"
| summarize by SubjectTitle
| count
```

Top operator

The `top` operator returns the first *n* rows from the query results, ordered by one or more specified columns. Use it to get the highest or lowest values based on your sort criteria.

The following example gets the top five largest items in your review set.

Kusto

```
ReviewSetTable
| top 5 by Size desc
```

Bin operator

The `bin()` function rounds a datetime or numeric value down to the nearest multiple of a specified bin size. Use this function to group data into time intervals or numeric ranges, such as hourly buckets or size ranges.

The following example groups items in the review set into size buckets of **1 MB** using `bin(Size, 1048576)`, counts how many items fall into each bucket with `summarize count()`, and then sorts the buckets in ascending order so you can see the distribution of file sizes from smallest to largest.

Kusto

```
ReviewSetTable
| summarize ItemCount = count() by bin(Size, 1048576)
| sort by bin(Size, 1048576) asc
```

Common scenarios

The following scenarios demonstrate how Kusto queries in the Advanced review set explorer (preview) can help you analyze and better understand your review set data.

Scenario 1: Communications between two users

Find items exchanged exclusively between two specific users - John and David - with no other participants. This query isolates their direct one-to-one communications and shows only three properties: subject title, sender, and recipients.

Kusto

```
ReviewSetTable
| where isnotempty(Participants)
| extend ParticipantEmails = extract_all(@"\<(\[^\>\]+)\>", tostring(Participants))
| extend ParticipantCount = array_length(ParticipantEmails)
| where ParticipantCount == 2
| where ParticipantEmails has "John@contoso.com" and ParticipantEmails has
"David@contoso.com"
| project SubjectTitle, SenderAuthor, Recipients
```

Scenario 2: Emails sent externally

Find all review set items that include at least one external recipient (outside your organization contoso.com), and list those external email addresses for each item.

Kusto

```
ReviewSetTable
| where isnotempty(Recipients)
| extend Emails = extract_all(@"\<(\[^\>\]+)\>", tostring(Recipients))
| mv-expand Email = Emails
```

```
| extend Email = tostring(Email)
| where Email !contains "@contoso.com"
| summarize ExternalRecipients = make_set(Email) by ImmutableId, SubjectTitle
```

Scenario 3: Identify date and time patterns in the review set

Visualizing data over time helps investigators identify spikes or anomalies in communication or document creation, such as a sudden surge in emails that might correlate with a known incident. Quarterly views align with business cycles, making it easier to correlate findings with financial reporting periods, internal audits, or policy changes.

The following example analyzes how review set items are distributed over time by grouping items by year and quarter and visualizing the results as a column chart. It filters items by date range, extracts the year and quarter from each item's **Date** field, counts the items in each year-quarter combination, creates a readable label (for example, *Q2 2024*), and renders the results as a column chart.

Kusto

```
ReviewSetTable
| where Date > datetime(2015-01-01) and Date < datetime(2025-08-31)
| extend Year = datetime_part("year", Date), Quarter = datetime_part("quarter", Date)
| summarize EventCount = count() by Year, Quarter
| sort by Year asc, Quarter asc
| extend QuarterLabel = strcat("Q", tostring(Quarter), " ", tostring(Year))
| project QuarterLabel, EventCount
| render columnchart
```

Scenario 4: Visualize Microsoft Teams message types in the review set

The following example categorizes and visualizes Microsoft Teams message types in the review set by item class, helping you understand the distribution of different message formats.

Kusto

```
ReviewSetTable
| where ItemClass startswith "IPM.SkypeTeams.Message."
| project SubjectTitle, Date, SenderAuthor, FileClass, ItemClass, InternetMessageId
| summarize ItemCount = count() by ItemClass
| sort by ItemCount desc
| render piechart
```

Use regex in the Advanced review set explorer (preview)

The Advanced review set explorer (preview) supports **regular expressions (regex)** for powerful pattern matching in text fields. Regex support enables you to search for complex sequences of characters, such as alphanumeric combinations, special characters, or structured patterns, beyond simple keyword matching. Use regex to:

- Identify items with specific formats, such as IDs, codes, or structured tokens.
- Detect patterns that standard operators like `has` or `startswith` can't capture.
- Perform advanced filtering for compliance or forensic investigations.

Tips for writing regex in KQL

- Escape backslashes properly (`\\`) when needed.
- Use anchors (`^` for start, `$` for end) for exact positioning.
- Test patterns incrementally to avoid overly broad matches.
- Combine with other operators (for example, `and`, `or`) for precise queries.
- Reference: Regex syntax.

Syntax

Use the `matches regex` operator in your `where` clause:

Kusto

```
where <PropertyName> matches regex "<pattern>"
```

Example: Match a complex pattern in Keywords

The following query finds items whose **Keywords** field matches a pattern of 10 word characters, 10 digits, and 10 non-word characters:

Kusto

```
ReviewSetTable
| where Keywords matches regex "\\w{10}\\d{10}\\W{10}"
| project SubjectTitle, CompoundPath, Date, SenderAuthor, FileClass, ItemClass,
InternetMessageId
```

- Filters items where the **Keywords** property contains:

    - 10 word characters (`\w`)
    - followed by 10 digits (`\d`)
    - followed by 10 non-word characters (`\W`)

- Returns key metadata fields as specified for review.

Fix common KQL syntax errors

When you create queries in the Advanced review set explorer (preview), you might encounter syntax errors.

The following table lists common errors and how to fix them:

Expand table

Error	Cause	Fix
`sort` operator requires a column name	Using `sort` without specifying what to sort by.	Add a column name after `sort by`. For example, `sort by Date desc`.
Expected `)` or `,`	Missing closing parenthesis or incorrect function arguments.	Check that all opening parentheses have matching closing parentheses and that function arguments are separated by commas.
Unknown function	Using a function name that isn't supported or is misspelled.	Verify the function name against the KQL documentation. Common examples include `count()`, `tostring()`, and `datetime()`.
Token expected: `\ |`	A query line is missing the pipe operator at the start.	Ensure every line after `ReviewSetTable` begins with `\ |`.
Column not found	Referencing a property name that doesn't exist in the review set schema.	Check the available column names in your review set. Property names are case-sensitive.
`summarize` requires a `by` clause	Using `summarize` with an aggregation but no grouping column.	Add a `by` clause with one or more columns. For example, `summarize count() by FileClass`.
Unexpected end of query	Query is incomplete or has a trailing pipe with no operator.	Remove any trailing `\ |` at the end of the query, or add the intended operator after it.

> **Note**

Some older review sets might not support certain KQL properties if those properties weren't available when the review set was created and data was generated. Querying an unsupported property in an older review set results in a column not found error. To resolve this error, create a new review set and use the **Add to review set** process to regenerate the data with the latest supported properties.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/29/2026

Review Microsoft Teams content in a review set in eDiscovery

06/21/2025

After you add Microsoft Teams content to a review set, the next step is to review the content for its relevance to your investigation and cull it if necessary. An important prerequisite to reviewing Teams content is understanding how eDiscovery processes Teams meeting, chat conversations, and attachments when adding them to a review set. The following processing of Teams content is taken:

- **Grouping**: Messages, posts, and replies Teams conversations are grouped together and presented in the review set. This processing also includes attachments in chat conversations are extracted and group within the conversation.
- **Conversation transcript threading**: eDiscovery determines what additional content from a conversation to collect to provide context around items that matched the collection criteria.
- **Deduplication**: eDiscovery processes and manages duplicate Teams content.
- **Metadata**: eDiscovery adds metadata content for Teams content after the content is added to a review set.

> **Note**
>
> The date and time for all conversations in Teams are shown in Coordinated Universal Time (UTC).

Understand grouping, conversation threading, deduplication, and Teams metadata to help you optimize the review and analysis of Teams content. This section also has tips for viewing Teams content in a review set.

Grouping

When content from Teams chat conversations is added to a review set, messages, posts, and replies from conversations are aggregated in HTML transcript files when the setting of "Organize conversations into HTML transcript" is selected. A single chat conversation can have multiple transcript files. An important function of these transcript files is to present Teams content as continuous conversations and not as individual (or separate) messages. This helps provides context for items that matched the search criteria of your collections in the previous step and reduce the number of items collected into the review set. Group transcripts and associated items by *family* or *conversation*. Items in the same family have the same value for

the **FamilyId** or **GroupId" metadata property. Items in the same conversation have the same value for the **ConversationId** or **ThreadId" metadata property depending on the grouping configuration selected.

The following table describes how the different types of Teams content are grouped by family and conversation.

Expand table

Teams content type	Group by family	Group by conversation
Teams 1:1 and group chats	A transcript and all of its attachments and extracted items share the same **FamilyId**. Each transcript has a unique **FamilyId**.	All transcript files and their family items within the same conversation share the same **ConversationId**. This includes the following items: All extracted items and attachments of all transcripts that share the same **ConversationId**. • All transcripts for the same chat conversation • All custodian copies of each transcript • Transcripts from subsequent collections from the same chat conversation For Teams 1:1 and group chat conversations, you might have multiple transcript files, each one corresponding to a different time frame within the conversation. Because these transcript files are from the same conversation with the same participants, they share the same **ConversationId**.
Standard, private, and shared channel chats	Each post and all replies and attachments are saved to its own transcript. This transcript and all of its attachments and extracted items share the same **FamilyId**.	Each post and its attachments and extracted items have a unique **ConversationId**. If there are subsequent collections or new replies from the same post, the delta transcripts resulting from those collections will also have the same **ConversationId**.
Teams meetings	Each meeting	

Use the **Group** control in the command bar of a review set to view Teams content grouped by family or conversation.

- Select **Group family attachments** to view Teams content grouped by family. Each transcript file is displayed on a line in the list of review set items. Attachments are nested under the item.

- Select **Group Teams or Viva Engage conversations** to view Teams content grouped by conversation. Each conversation is displayed on a line in the list of review set items. Transcript files and attachments are nested under the top-level conversation.

> ⓘ **Note**
>
> Cloud attachments are grouped with the conversations they appear in. This grouping is accomplished by assigning the same **FamilyId** as the transcript file of the message the file was attached to and the same **ConversationId** as the conversation the message appeared in. This means multiple copies of cloud attachments might be added to the review set if they were attached to different conversations.

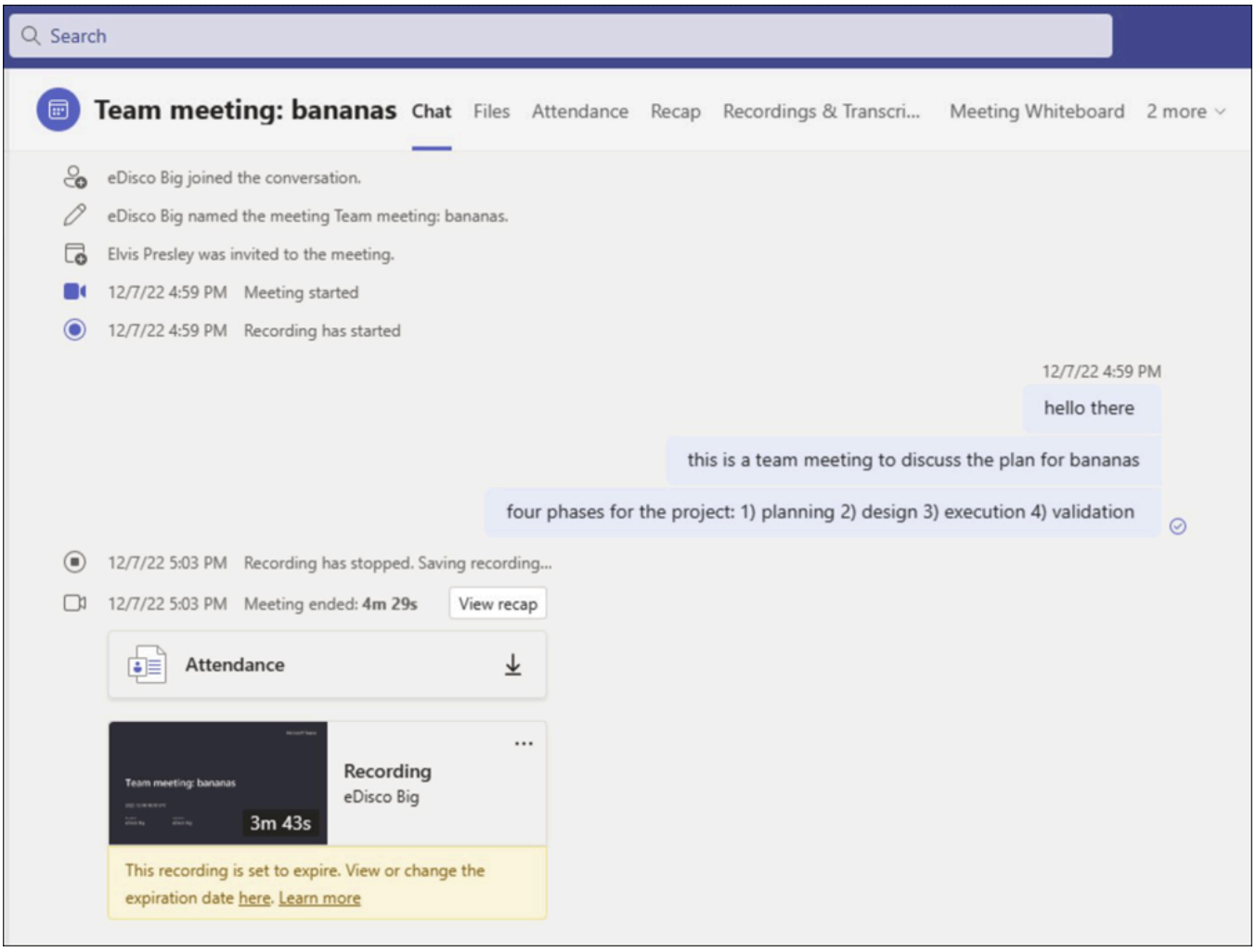
Viewing recorded meeting transcripts

The transcript of audio of the recorded meeting is captured as a separate file and indexed automatically for search. Recorded meetings in a review set are stored as a .zip file that contains the following files:

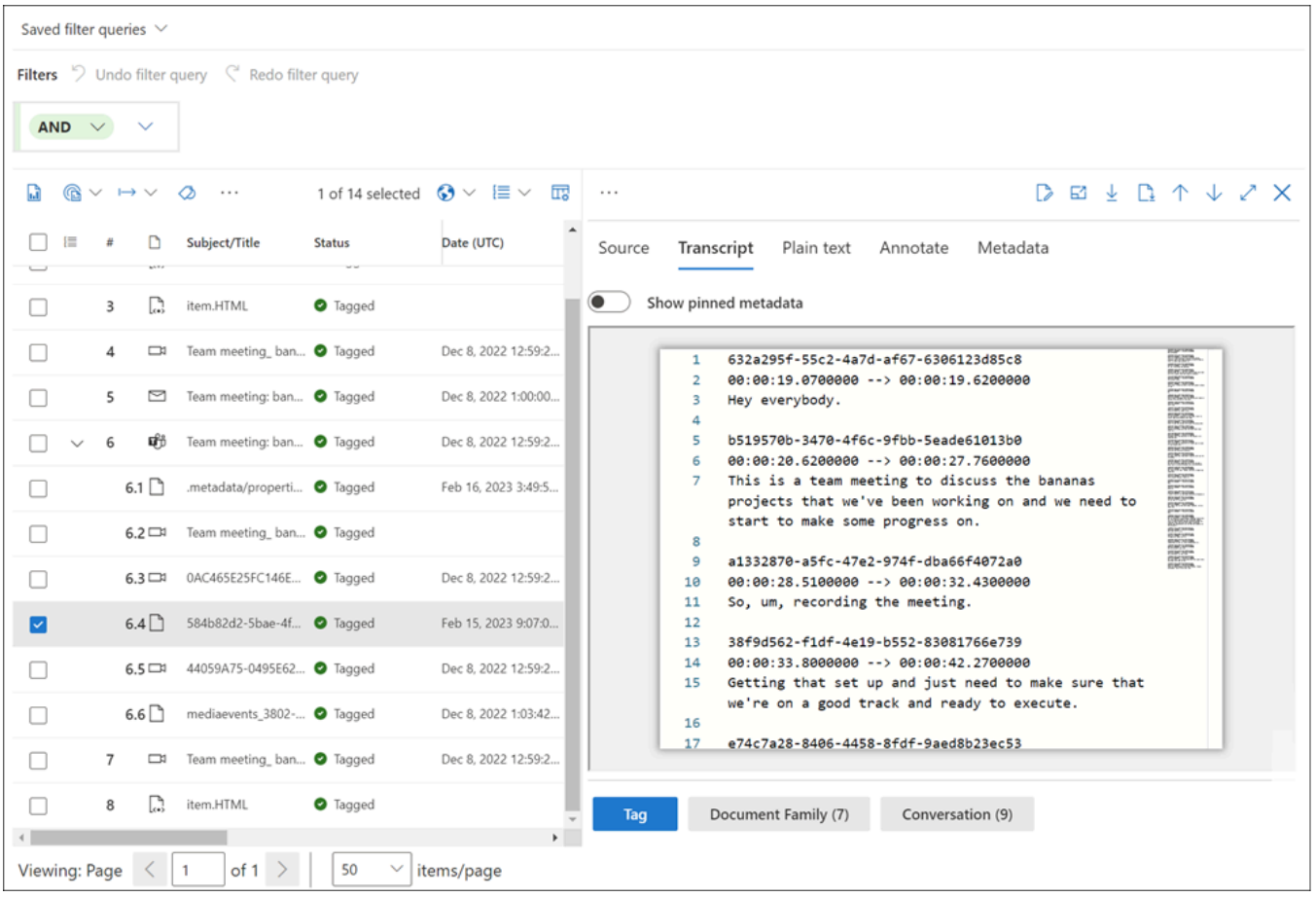
- The transcript of the meeting audio in .txt format
- The video recording of the meeting in .mp4 format
- The thumbnail image of the meeting in .jpg format
- Meeting metadata and meeting chapters (as applicable) in .json format

To view meeting audio transcript files in a review set, select the meeting and the **Transcript** viewer on the meeting details pane. The following screenshots show an example of a meeting in the Teams client and the meeting transcript file of the same meeting in the review set.

Meeting in Teams client



Meeting transcript file in a review set

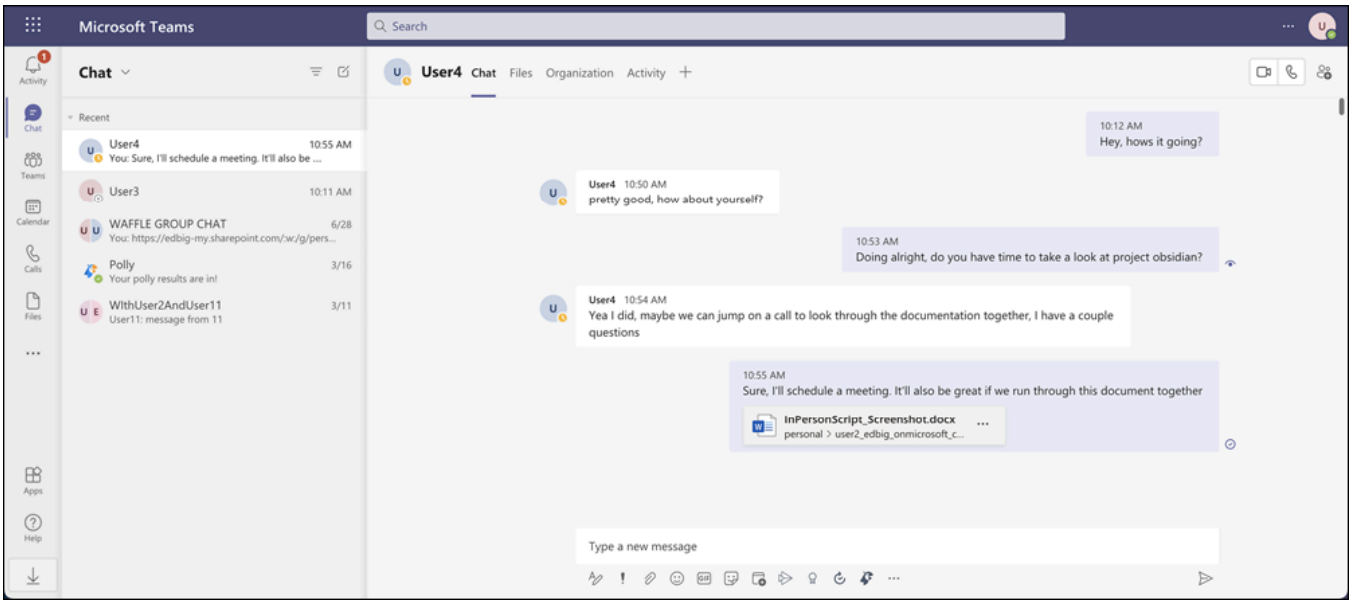


Viewing conversation transcript files

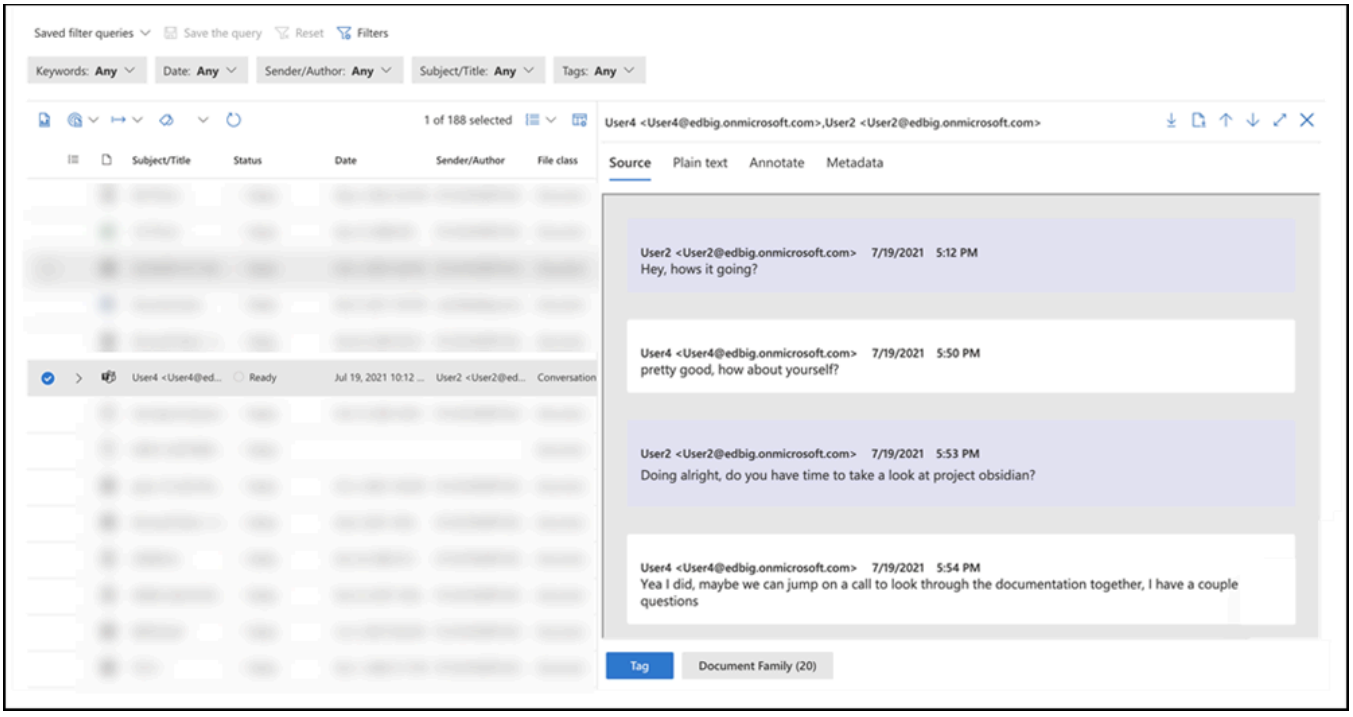
When viewing transcript files in a review set, some of the messages are highlighted in purple. The messages that are highlighted depend on which custodian copy of the transcript you're viewing. For example, in a 1:1 chat between User4 and User2, the messages posted by User4 are highlighted in purple when you view the transcript collected from User4's mailbox. When you view User2's transcript of the same conversation, messages posted by User2 are highlighted in purple. This highlighting behavior is based on the same Teams client experience, where a user's posts are highlighted in purple in the Teams client.

The following screenshots show an example of conversation in the Teams client and the transcript file of the same conversation in the review set. The purple highlighting in the transcript file indicates that the transcript was collected from User2's mailbox.

Conversation in Teams client



Conversation in a transcript file



Include contextual conversations for Teams chat messages

If you select "Include Teams and Viva Engage conversations" setting during the add to review set process, eDiscovery helps you identify contextual content related to items that might be relevant to your investigation. This feature produces distinct conversation views that include chat messages that precede and follow the items match the search query during collection. This capability allows you to review complete chat conversations in Microsoft Teams. This capability is independent from HTML threading and only affects whether contextual conversations are included in the process of adding items to a review set.

Here's the logic used by eDiscovery to include additional messages and replies that provide context around the items match the search query (called *responsive items*) you used when adding Teams content to a review set. Different behaviors are based on the types of chats and the search query used to collect the responsive items. There are two common search scenarios:

- Queries that use search parameters, such as keywords and property:value pairs
- Queries that only use date ranges

Expand table

Teams content type	Queries with search parameters	Queries with date ranges
Teams 1:1 and group chats	Messages that were posted 12 hours before and 12 hours after responsive items are included	Messages that fall in between the data range and any contextual messages within a 24-hour window are included.
Standard, private, and shared Teams channel chats	Each post that contains responsive items and all corresponding replies are included.	Each post that contains responsive items and all corresponding replies are included.

Deduplication of Teams content

The following list describes the deduplication (and duplication) behavior when collecting Teams content into a review set when user chooses to organize conversations in an HTML transcript.

- Each transcript file added to a review set should be a one-to-one mapping to content stored in data locations. That means eDiscovery doesn't collect any Teams content that have already been added to the review set. If a chat message is already collected in a review set, eDiscovery doesn't add the same message from the same data location to the review set in subsequent collections.

- For 1:1 and group chats, copies of messages are stored in the mailbox of each conversation participant. Copies of the same conversation that exist in different participants' mailboxes are collected with different metadata. As a result, each instance of the conversation is treated as unique and brought into the review set in separate transcript files. So if all participants of a 1:1 or group chat are added as custodians in a case and included in the scope of a collection, then copies of each transcript (for the same conservation) are added to the review set and are grouped together with the same **ConversationId** or **ThreadId**. Each of these copies is associated with a corresponding data source.

- In subsequent "Add to review set" processes where there are items from the same conversation, only the delta content that wasn't previously added previously is added to

the review set and grouped (by sharing the same **ConversationId** or **ThreadId**) with the previously added items from the same conversation. Here's an example of this behavior:

1. Process A adds messages in a conversation between User1 and User2 to the review set.
2. Process B adds messages from the same conversation, but there are new messages between User1 and User2 since process A was run.
3. Only the new messages in Process B are added to the review set. These messages are added to a separate transcript file, but the new transcript is grouped with the transcripts from Process A by the same **ConversationId**.

This behavior applies to all the types of Teams chats.

Metadata for Teams content

In large review sets with thousands or millions of items, it can be difficult to narrow the scope of your review to Teams content. To help you focus your review on Teams content, there are metadata properties that are specific to Teams content. You can use these properties to organize the columns in the review list and configure filters and queries to optimize the review of Teams content. These metadata properties are also included when you export Teams content from eDiscovery, to help you organize and view content post-export or in non-Microsoft eDiscovery tools.

The following table describes metadata properties for Teams content.

Expand table

Metadata property	Description
ContainsDeletedMessage	Indicates whether a transcript file contains a deleted message. Deleted messages are identified when viewing the transcript file.
ContainsEditedMessage	Indicates whether a transcript file contains an edited message. Edited messages are identified when viewing the transcript file.
ConversationId	A GUID that identifies the conversation that the item is associated with. Transcript files and attachments from the same conversation have the same value for this property.
Conversation name	The name of the conversation the transcript file or attachment is associated with. For Teams 1:1 and group chats, the value of this property is the User Principal Name (UPN) of all participants of the conversation are concatenated. For example, `User3 <User3@contoso.onmicrosoft.com>,User4 <User4@contoso.onmicrosoft.com>,User2 <User2@contoso.onmicrosoft.com>`. Teams channel (standard, private, and shared) chats use the following format

Metadata property	Description
	for conversation name: `<Team name>,<Channel name>`. For example, `eDiscovery vNext, General`.
ConversationType	Indicates the type of Team chat. For Teams 1:1 and group chats, the value for this property is `Group`. For standard, private, and shared channel chats, the value is `Channel`.
Date	The time stamp of the first message in the transcript file in UTC.
FamilyId	A GUID that identifies the transcript file for a chat conversation. Attachments have the same value for this property as the transcript file that contains the message the file was attached to.
FileClass	Indicates that type of content. Items from Teams chats have the value `Conversation`. In contrast, Exchange email messages have the value `Email`.
MessageKind	The message kind property. Teams content has the value `microsoftteams , im`.
Participants	List of all people who are involved in the chat, includes both senders and recipients in a conversation.
Recipients	A list of all users who received a message within the transcript conversation.
TeamsChannelName	The Teams channel name of the transcript.

For descriptions of other eDiscovery metadata properties, see Document metadata fields in eDiscovery.

Export Teams content

After you review and cull Teams content in a review set, you can export all items that are applicable to your investigation. There aren't any specific export settings for Teams content. If the content was added to review set as a transcript file, then each transcript is exported as an HTML message file. The metadata properties discussed in the previous section are included when Teams content is exported.

Each export item (transcript or individual messages) is referenced in the items.csv report and can be located using the target path field in the report. Transcript files are located in the Conversations folder in the root export folder.

Tips for viewing Teams content in a review set

Here are some tips and best practices for viewing Teams content in a review set.

- Use the **Customize columns** control in the command bar to add and organize columns to optimize the review of Teams content. You can add and remove columns that are useful for Teams content. You can also sequence the order of columns by dragging and dropping them in the **Edit column** flyout page. You can also sort on columns to group Teams content with similar values for the column you sort on.
- Useful columns that to help you review Teams content include **Participants**, **Recipients**, and **File type** or **Message kind**.
- Use filters for Teams-related properties to quickly display Teams content. There are filters for most of the metadata properties described in the previous section.

Deleting Teams chat messages

You can use eDiscovery and the Microsoft Graph Explorer to respond to data spillage incidents, when content containing confidential or malicious information is released through Teams chat messages. Admins in your organization can search for and delete chat messages in Microsoft Teams. This feature can help you remove sensitive information or inappropriate content in Teams chat messages. For more information, see Search and purge chat messages in Teams.

Group and view items in a review set in eDiscovery

eDiscovery review sets display content using different grouping options and include specialized viewers that you can use to examine details about individual items.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Grouping

Use the **Group** control in the command bar of a review set to view review content grouped by the following options:

Option 1: Group items by Group ID and Thread ID

If you're new to eDiscovery or are an existing eDiscovery customer with a steady stream of new cases, group review set items by using Group IDs and Thread IDs. This grouping option makes it easier to find and review related items in a review set by grouping them together in the list view. When the **Enable group** option is enabled in the review set case settings, you see the following grouping options on the **Group** control in the command bar of a review set:

- **Group by families**: All items related to a specific file are grouped together by using the same Group ID. For example, if you have a PowerPoint file in the review set that includes embedded images or .zip files, these images and files are grouped with the PowerPoint file and shown as nested items with the file in the item list view. Microsoft 365 Copilot and Microsoft 365 Copilot Chat user prompts and responses are also grouped together by using the same Group ID when there are multiple prompts and responses associated with a file.
- **Group by conversations**: All email messages, Teams conversations, and Viva Engage conversations are grouped by using the same Thread ID and appear as nested items. Additionally, all associated content for these messages and conversations is also grouped together. For example, if you have an email conversation that includes several email messages, some of which include attachments and some that include embedded images, all

of the email messages, attachments, and images are grouped together in the review set list view under an applicable item.

> ⓘ Note
>
> For cases created before March 15, 2023, the **Enable group** toggle is disabled and the **Group by families** and **Group by conversations** options aren't available. Item grouping in review sets for these cases is based on *Family ID* and *Conversation ID* described in Option 2: Group items by Family ID and Conversation ID.

Option 2: Group items by Family ID and Conversation ID

If you're an existing eDiscovery customer with a large number of existing cases or use existing internal or third-party automation to help process review set items, you might want to continue to group review set items by using Family IDs and Conversation IDs. When the **Enable group** option is disabled in the review set case settings, you see the following grouping options on the **Group** control in the command bar of a review set:

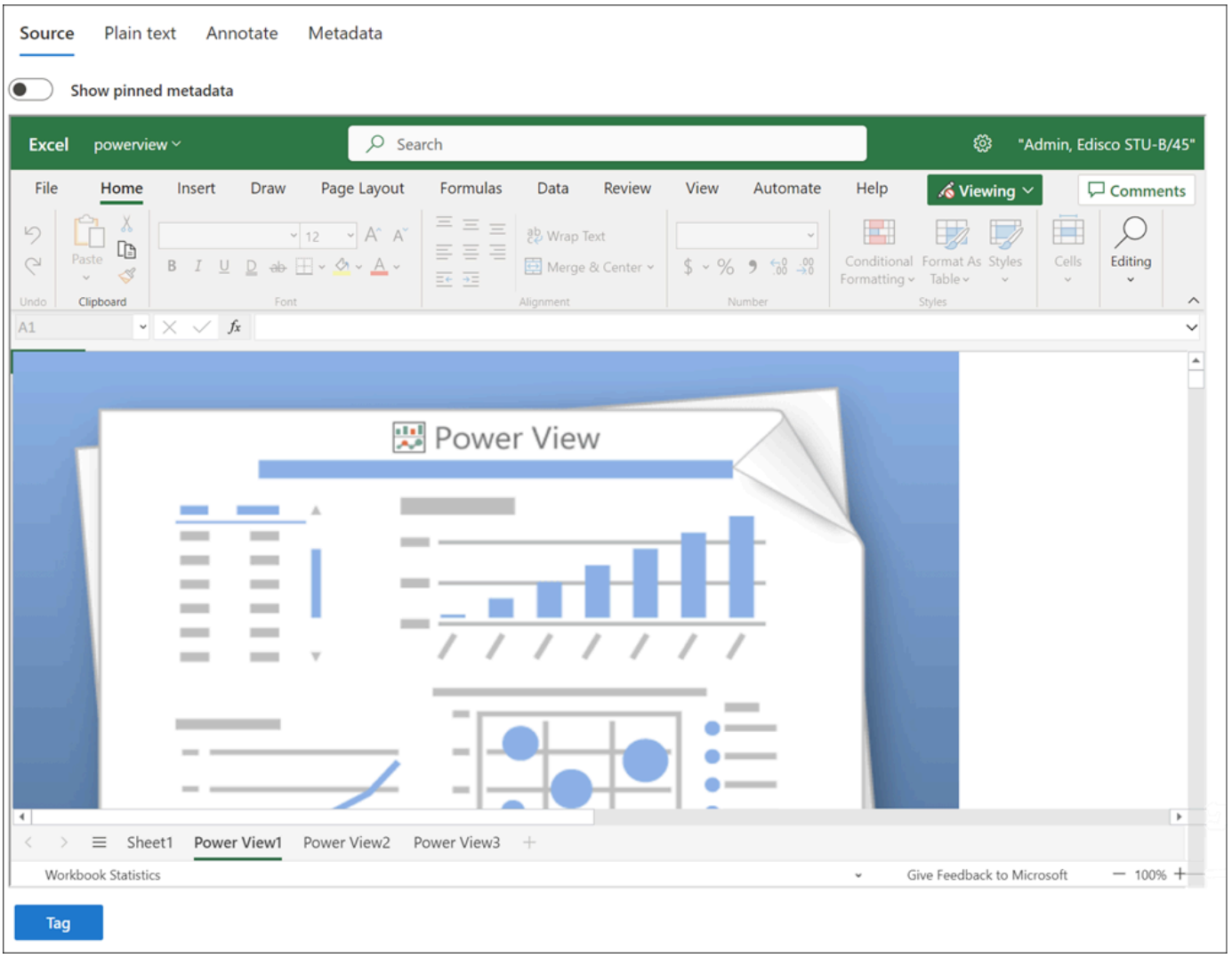
- **Group family attachments**: View review set content grouped by family. Each transcript file is displayed on a line in the list of review set items. Attachments are nested under the item.
- **Group Teams or Viva Engage conversations**: View Teams and Viva Engage content grouped by conversation. Each conversation is displayed on a line in the list of review set items. Transcript files and attachments are nested under the top-level conversation.

> ⓘ Note
>
> Cloud attachments are grouped with the conversations they appear in. Grouping cloud attachments with their conversations is accomplished by assigning the same **GroupID** as the transcript file of the message the file was attached to and the same **ConversationId** as the conversation the message appeared in. This means multiple copies of cloud attachments might be added to the review set if they were attached to different conversations.

Use source view

The **Source** viewer displays the richest view of a selected document. It supports hundreds of file types and is designed to display the truest to native experience possible. For Microsoft Office files, the viewer uses the web version of Office apps to display content such as document

comments, Microsoft Teams chats, Excel formulas, hidden rows and columns, and PowerPoint notes.

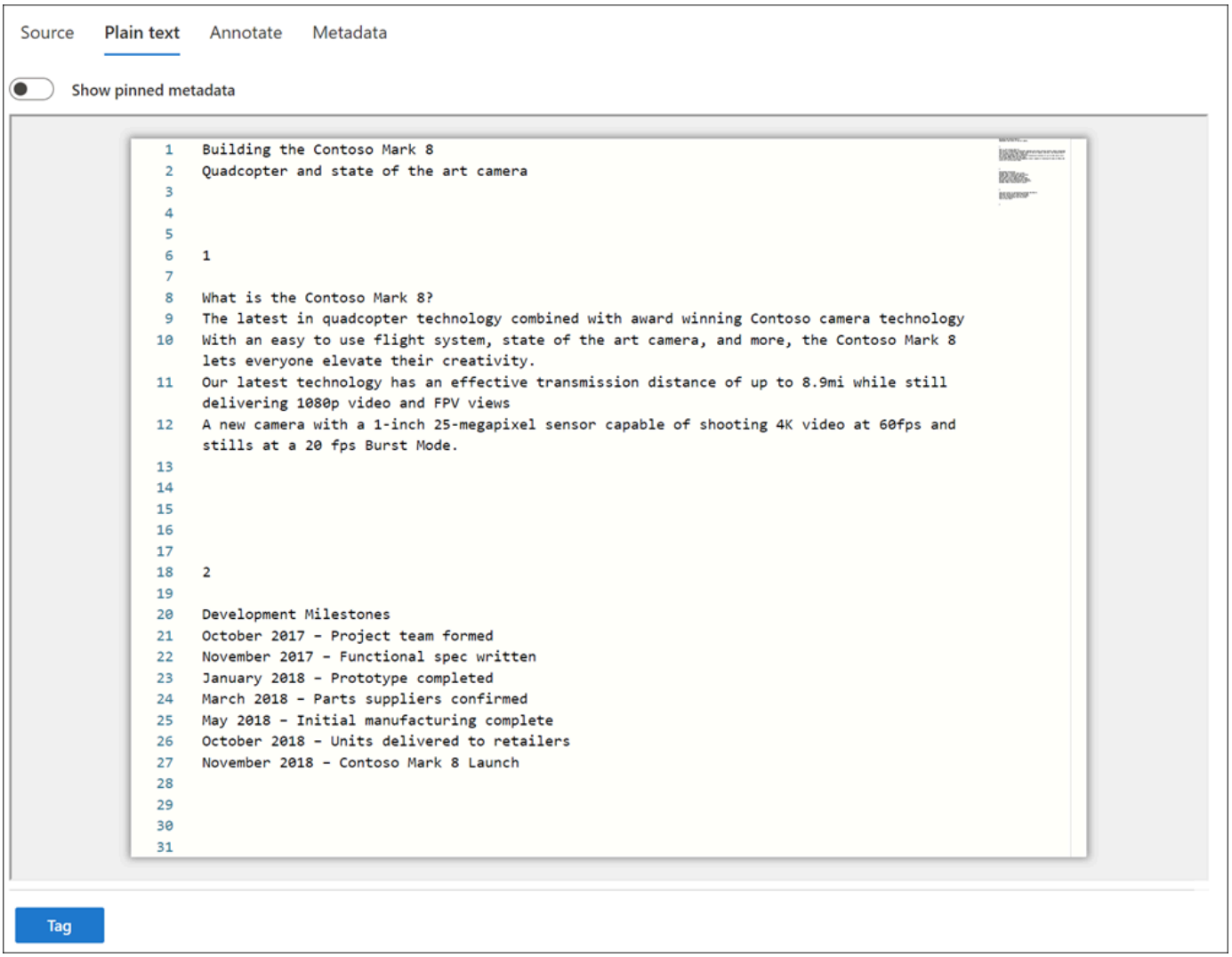


The **Source** view doesn't support previewing video and audio files. To view these files, export them from the review set as described in Export from a review set.

Use plain text view

The **Plain text** viewer provides a view of the extracted text of a selected file. It ignores any embedded images and formatting but is useful if you want to understand the content quickly. The text view also includes these features:

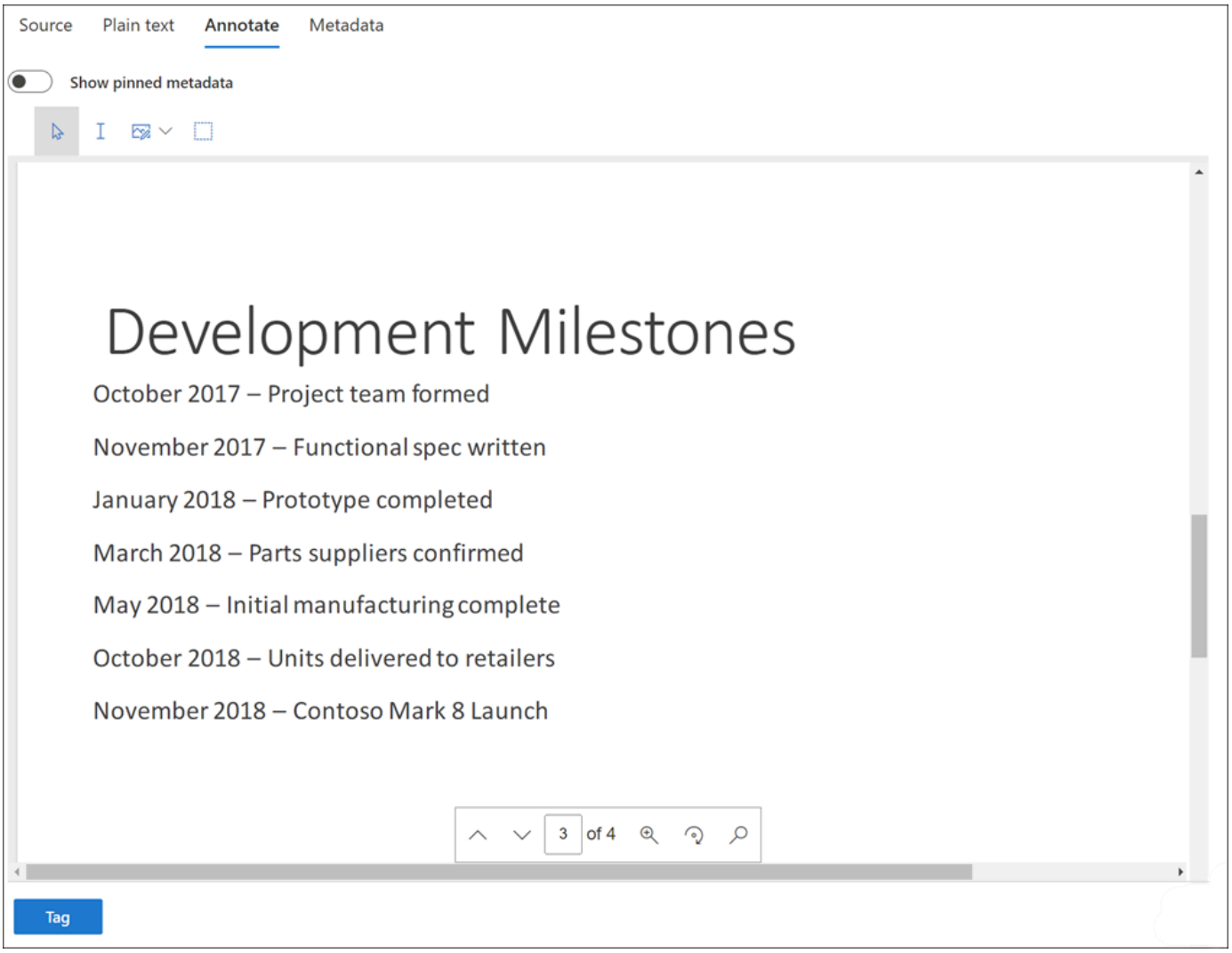
- Line counter makes it easier to reference specific portions of a document
- Search hit highlighting that highlights terms within the document and in the scrollbar



Use annotate view

The **Annotate** view provides features that let you apply markup on a selected document, including the ability to:

- **Select annotations**: Select annotations on a document to delete.
- **Select text**: Select text on the document to delete.
- **Area redactions**: Draw a box on the document to hide sensitive content.
- **Pencil**: Freehand draw in selectable colors on a document to bring attention to certain portions of a document.
- **Toggle annotation transparency**: Make annotations semi-transparent to view the content behind the annotation.
- **Previous page**: Go to the previous page.
- **Next page**: Go to the next page.
- **Go to page**: Enter a specific page number to go to.
- **Zoom**: Set the zoom level for annotate view.
- **Rotate**: Rotate the document clockwise.



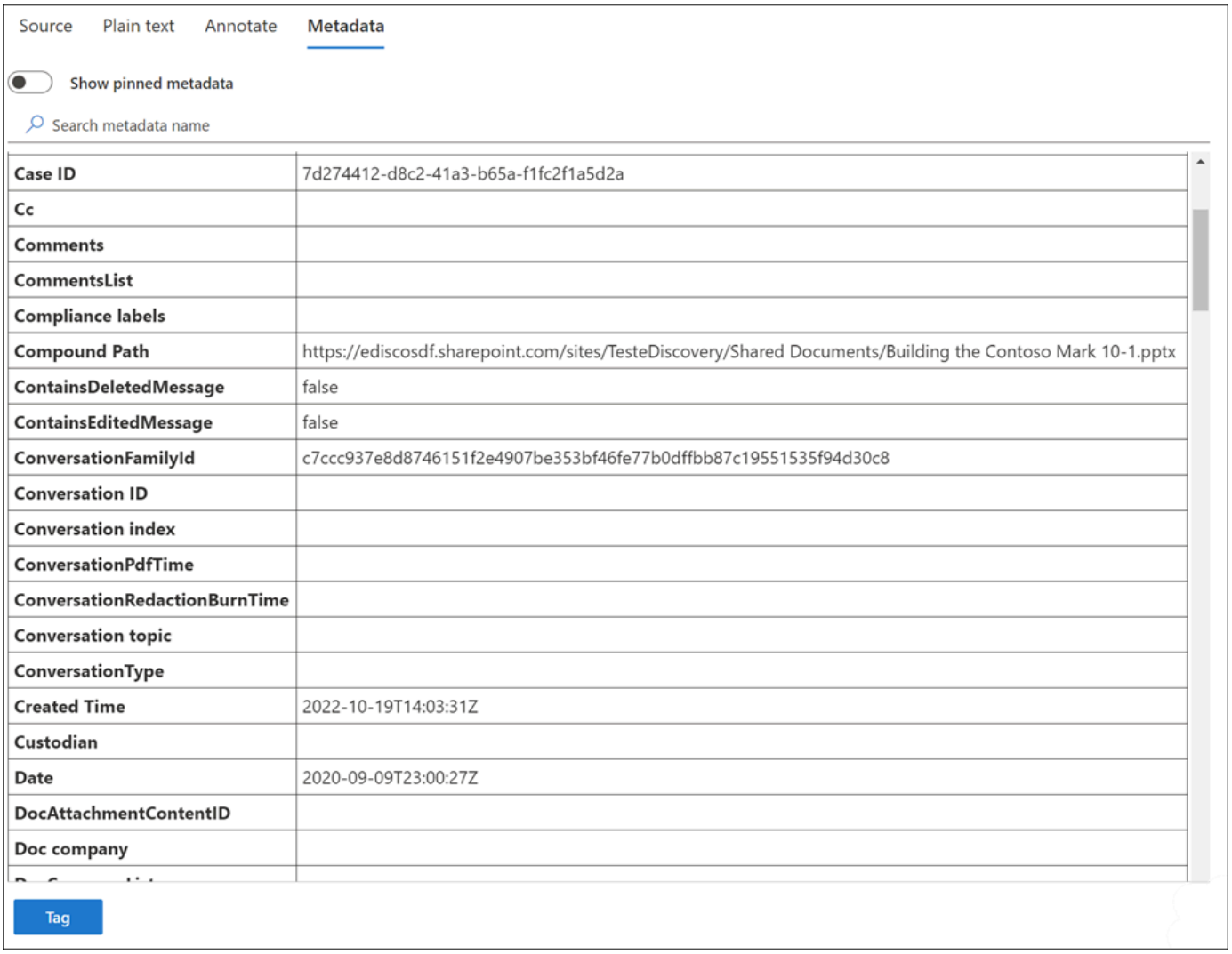
Commit redactions

When you apply annotations in the **Annotate** view (including area redactions, text deletion, and pencil markup), the annotations save with the document but don't change how the document appears in the review set, in search results, or in exports until you **commit** them. When you commit a redaction, it generates a redacted PDF (`{itemId}_burn.pdf`) that's persisted alongside the original item in the review set.

To include redacted PDFs in an export, select *Export redaction* when configuring the export. For details on how redacted items appear in the export package – including how individual documents, attachments, and conversations are handled – see Export from a review set.

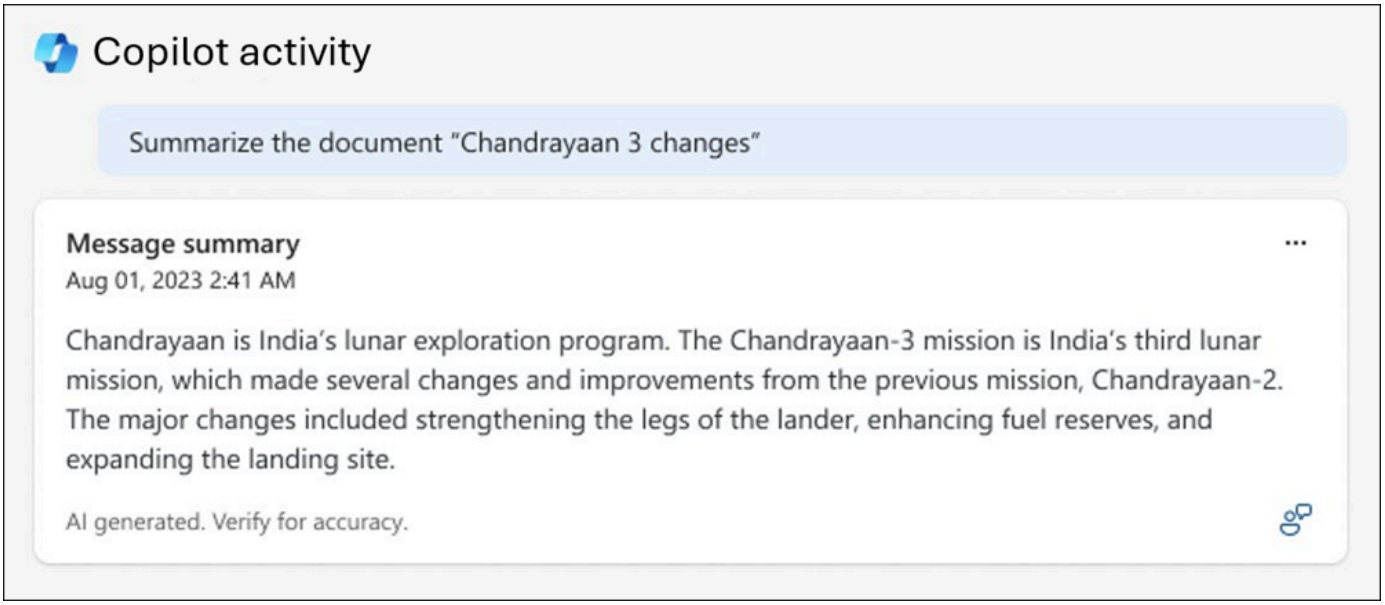
Use metadata view

You can toggle the panel in the **Metadata** view on or off to display various metadata associated with the selected document. Although you can customize the search results grid to display specific metadata, there are instances where scrolling horizontally can be difficult while reviewing data. The File metadata panel allows you to toggle on a view within the viewer.



Use Copilot activity view

The **Copilot activity** view displays the Microsoft 365 Copilot and Microsoft 365 Copilot Chat activity associated with the selected item. User prompts and Copilot responses display in the order they're generated.



Summarize an item by using Microsoft Security Copilot

You can use Security Copilot in Microsoft Purview to provide a contextual summary of most items in a review set. The summary is based on the text included in a selected item. The contextual summary saves time for reviewers by quickly identifying information helpful when tagging or exporting items.

Security Copilot summarizes the entire item, including any documents, meetings transcripts, or attachments. Most common document file types are supported.

Limitations

Security Copilot summarization in review sets has the following limitations:

- The *Plain text* view must be available for an item for Security Copilot availability. For more information about which file types support plain text views, see Supported file types in eDiscovery.
- The item content must be at least 100 words and less than 15,000 words. If the item content is fewer than 100 words or more than 15,000 words, Security Copilot doesn't summarize it.
- Summarize items in various languages supported by Security Copilot. For more information, see Supported languages in Microsoft Security Copilot.
- Currently, only individual Teams messages can be summarized. For example. if you summarize an individual Teams message, Copilot doesn't pull in the surrounding conversation (family or conversation groups).

For more information about Microsoft Security Copilot and Microsoft's artificial intelligence (AI) responsibilities, see Responsible AI FAQ.

Prerequisites for using Security Copilot

To summarize an item in a review set in eDiscovery by using Security Copilot, you must:

- Have specific licenses and onboard your organization to Security Copilot. For more information, see Learn more about Security Copilot licensing requirements and onboarding.
- Have an admin assign you at least one of the required roles to access the Security Copilot portal. For more information about these required roles, see Understand authentication in Security Copilot.

> ⓘ **Note**

Currently, eDiscovery only supports Security Copilot in the embedded experience (not the Security Copilot stand-alone experience).

Summarize an item in a review set

1. In eDiscovery, open a review set and filter the items as needed.
2. Select an item in the list. You can select a parent item or a child item.
3. In any item view, select **Summarize**. Selecting **Summarize** opens the **Copilot** panel on the right side of the screen and displays a summary of the item.

   > ⓘ **Note**
   >
   > If the plain text view isn't available for an item, **Summarize** isn't available to select.

4. To use Copilot to create the summary, select one of the suggested prompts at the bottom of the panel. For example, if the item is a legal contract in the Microsoft Word format, select the *Identify key participants*, *Identify key topics*, or *Identify action items* prompt.

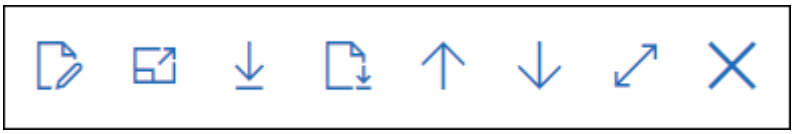
   You can also enter an open-ended question in the *Ask questions and get insights about this item* field. Copilot responses are generated without using information from prior messages. For more information about prompts and promptbooks, see Microsoft Copilot in Microsoft Purview prompts and promptbooks.

   > ⓘ **Note**
   >
   > Only item content is summarized. If you ask a question relevant to the item content, Copilot provides a result. If the question isn't relevant to the item content, Copilot instructs you to ask a different question.

5. To provide feedback on the summarized content, select the drop-down arrow in the lower-right corner below the summary, and then enter your feedback.
6. After reviewing the summary, you can take other eDiscovery actions connected with review set items.

Use viewer and management tools

For selected content, you get extra view and management tools that help you work with documents.



You can select the following actions for selected content:

- **View or update notes**: Opens the **Notes** pane where you can view or add notes to the selected document. You can display the notes for the document in a column in the dashboard list by customizing your displayed columns.
- **Show in a new window**: Displays the selected document in a new browser window or tab.
- **Download PDF**: Downloads the selected document as a .pdf file.
- **Move up**: Selects the previous review set item and displays the item in the viewer.
- **Move down**: Selects the next review set item and displays the item in the viewer.
- **Maximize/Minimize**: Maximizes the viewer pane to full screen and minimizes the viewer pane.
- **Close**: Closes the selected item from the viewers.

> **ⓘ Important**
>
> To download a PDF containing items from a review set, you must be assigned the *Export* role or be assigned to a role group that includes the *Export* role. For more information about eDiscovery permissions, see [Assign eDiscovery permissions](#).

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/29/2026

Tag items in a review set in eDiscovery

Organizing content in a review set is important to complete various workflows in the eDiscovery process. The review-set organization workflow includes:

- Culling unnecessary content
- Identifying relevant content
- Identifying content that must be reviewed by an expert or attorney

When experts, attorneys, or other users review content in a review set, they can capture their opinions related to the content by using tags. For example, if the intent is to cull unnecessary content, a user can tag documents with a tag such as "nonresponsive". After content is reviewed and tagged, you can create a review set search to exclude any content tagged as "nonresponsive". Excluding tagged nonresponsive content from the review set search eliminates that content from the next steps in the eDiscovery workflow.

You can use *tag templates* to define tags and review workflow across all cases and review sets in your organization. Using tag templates is the recommended approach and helps provide structure and standardization for tagging to ensure searches identify similar content across your organization. You can also use the tagging panel in an individual review set to customize tags for each case when tag templates don't meet your needs, but tags created here are specifically for the individual case and review set.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Tag types

eDiscovery has two types of tags:

- **Single-select tags**: Restricts reviewers to selecting a single tag within a tag group. These types of tags can be useful to ensure that reviewers don't select conflicting tags such as "responsive" and "nonresponsive". Single choice tags appear as radio buttons.
- **Multi-select tags**: Allow reviewers to select multiple tags within a tag group. These types of tags appear as checkboxes.

Tag groups

In addition to the tag types, the structure of how tag groups are organized in the tag panel can make tagging documents more intuitive. Tags are organized in tag groups. Review set search supports the ability to search by tag and by tag groups. Because review set search supports both levels, you can create a review set search to retrieve documents tagged with any tag in a group or the group itself.

For example, if you need to identify and tag privileged content, grouping can be used to make it clear that a reviewer can tag a document as *Privileged* (the tag group) and select *Attorney work product* (specific tag) as the type of privilege within the tag group.



Tag templates (preview)

Use tag templates to define and standardize tags and the review workflow across all cases and review sets in your organization. Each tag template contains tag groups where you can create a grouping structure for individual tags. You can create tag templates in the **Tag templates** tab in **eDiscovery settings**.

Create tag templates

To create a new tag template, complete the following steps:

1. Go to the Microsoft Purview portal and sign in with the credentials for a user account assigned eDiscovery permissions.
2. Select the **Settings** solution card, then select **eDiscovery**.
3. Select **Tag templates**.

4. On **Manage tag templates**, select **Create**.
5. On **New tag template**, enter a name and description for the tag template.
6. In the initial tag group, enter a tag group name and use the dropdown to define this as a **Multi-select** or **Single-select** tag group.
7. Enter at least one tag name in the tag group. Select **Add tag** to add more tags to the tag group.
8. If needed, select **Add tag group** to add more tag groups and tags to the tag template.
9. Preview the tag template group and tag structure in the **Preview** pane on the right side of **New tag template**.
10. When complete, select **Save**.

The tag template is now available for use in all cases and review sets in your organization.

Manage tag templates

After you create a tag template, you can edit or delete it. You can also view information about the tag template, including the dates when you created or last modified it.

To edit a tag template, select the tag template name on **Manage tag templates** and select **Edit**. On **Edit tag template**, you can edit the tag template name, description, and tag groups and tags. When you're done, select **Save**.

To delete a tag template, select the tag template name on **Manage tag templates** and select **Delete**. Select **Yes** to confirm the deletion or **No** to cancel the deletion.

Apply tag templates to a case and review sets

After you create a tag template, you can apply it to a review set. You can apply a tag template when you create a new review set or when you edit an existing review set.

To apply a tag template to a new review set, complete the following steps:

1. Go to the Microsoft Purview portal and sign in with the credentials for a user account assigned eDiscovery permissions.
2. Select the **eDiscovery** solution card, then select **Cases** in the left navigation.
3. Select a case, then select the **Review sets** tab.
4. Select a review set, then select the **All items** tab.
5. On the review set command bar, select **Manage** > **Tags**.
6. On **Manage tags**, delete the blank tag group, then select **Import tag templates**.

7. On **Import tag templates**, select the tag template that you want to apply to the review set and then select **Import**.
8. If needed, edit the tag template as it applies to the case and review set.
9. Select **Save** to apply the tag templates to the case and review set.

The applied tag template is now available for use in the case and all included review sets.

Create tags for a specific case or review set

If you don't want to use tag templates for a specific case or review set, you can create tags directly in the review set. This approach is useful when you need tags that are specific to a case or review set. However, these tags are unique to the case or review set and can't be used in other cases or review sets.

Before you apply tags to items in a review set, you need to create a tag structure.

1. Open a review set. On the review set command bar, select **Manage** > **Tags**.
2. On **Manage tags**, enter a *Tag group name* and select either **Multi-select** or **Single-select** for the tag group.
3. Enter one or more tags in the tag group. Select **Add tag** to add more tags to the tag group.
4. Select **Add tag group** if you need more tag groups and tags.
5. Preview the tag template group and tag structure in the **Preview** pane on the right side of **New tag template**.
6. Select **Save** to apply the tag group and tags to the case and review set.

Apply tags

With the tag structure from tag templates or custom tagging in place, reviewers can apply tags to items in a review set by configuring tagging settings.

1. In the review set command bar, select **Tag files** to display the **Tag files** flyout page (also called the *tagging panel*).

2. On the **Tag files** flyout page, set the following options to configure how to tag items displayed in the review set. The filters or filter queries currently applied to the review set determine which items are displayed and therefore the items that you can apply tags to. For more information, see Search content in a review set in eDiscovery.

    - **Choose selection**. Choose one the following options to determine the scope of items to apply tags to.

- **Tag selected items**: Applies tags to the items that you select. You can select items before or after launching the tagging panel. The **Tag selected items** option displays (in real time) the number of selected items that are tagged.
- **Tag all items in list**: Applies tags to all items displayed in the review set. The **Tag all items in list** option displays the total number of items that are tagged.

- **Expand selection**: Use the following options to tag additional items that are related to tagged items in the review set. Depending on the **Enable group** option enabled in the review set settings, you see the following options:

  With *Enable group* in settings disabled:
  - **None**: The **None** option doesn't apply tags to associated family items or associated conversation items. It only applies tags to the items that are selected or to all items in the review set list.
  - **Include family groups**: The **Include family groups** option applies the same tag to the associated family items of items that are tagged. *Family items* are items that share the same **FamilyId** metadata property value. For example, a document that's attached to an email message shares the same **FamilyId** as the email message. So if this option is selected for this example, the email message and the document are tagged, even though the document might not be included in the list of review set items.
  - **Include conversation groups**: The **Include conversation groups** option applies the same tag to all items that are in the same Teams or Viva Engage conversation as the items that are tagged. *Conversation items* are items that share the same **ConversationId** metadata property value. All messages, posts, and corresponding transcript file of a conversation share the same **ConversationId**. If this option is selected, then all items in the same conversation (and transcript file) are tagged, even though some of those conversation items might not be included in the list of review set items.

  > ⓘ Note
  >
  > Including family or conversation groups doesn't change the count of items shown in the **Tag selected items** or **Tag all items in list** options. In other words, the number of grouped items that are tagged isn't displayed.

  With *Enable group* in settings enabled:

    - **None**: The **None** option doesn't apply tags to family group items or conversation group items. It only applies tags to the items that are selected or to all items in the review set list.
    - **Include family groups**: The **Include family groups** option applies the same tag to the associated family items of items that are tagged. *Family group* items are items that share the same **GroupId** metadata property value. For example, a document that's attached to a PowerPoint file shares the same **GroupId** as the document. So if this option is selected for this example, the document and the PowerPoint file are both tagged. Both are grouped together in the list of review set items.
    - **Include conversation groups**: The **Include conversation groups** option applies the same tag to all items that are in the same email, Teams, or Viva Engage conversation as the items that are tagged. *Conversation group* items are items that share the same **ThreadId** metadata property value. All messages, posts, and corresponding transcript files of a conversation share the same **ThreadId**. If this option is selected, then all items in the same conversation (and transcript file) are tagged. All items are grouped together in the list of review set items.

- **Assign tags**: This section displays the tags (organized by tag groups) that you can apply to documents. These tags are either the tag groups and tags assigned using tag templates or custom tag groups and tags that you create for this case. You can only apply one single-choice tag (identified by a radio button) per tag group. However, you can apply multiple multichoice tags (identified by a checkbox).

3. Select **Apply tags** to apply the tags based on your settings.

   The **Applying tags** status message is displayed for each tag group on the tagging panel to indicate a tagging process has started. Tags for each tag group in the **Assign tags** section are greyed out until the process completes.

> **Tip**
>
> If you're in the process of configuring the settings on the tagging panel, but want to start over, select **Reset tag assignment** to clear the current setting. This control doesn't apply to items that are already tagged, and it doesn't change or remove tags from previously tagged items.

Set up smart tags

Machine learning (ML) capabilities in eDiscovery can help you make the decision process more efficient when reviewing case documents in a review set. Smart tags bring the ML capabilities to where you record decisions: when tagging documents during review. When you create a smart tag group, the decisions that result from the ML model you associate with the smart tag group display in-line with the tags in the tag group. Displaying model decisions inline with the tag group helps you see the ML results information in-line when you're reviewing specific documents.

How to set up a smart tag group

To create a smart tag group, complete the following steps:

1. In a review set, select **Manage review set** and then select **Manage tags**.
2. Select **Add tag group** and then select **Add smart tag group**.
3. Select the ML model that you want to associate to the tag group.

   Selecting an ML model creates a tag group and *N* child tags, where *N* is the number of possible outputs of the model. For example, the attorney-client privilege detection model has two possible outputs:

   - **Positive**: Use to tag documents that contain attorney-client privileged content.
   - **Negative**: Use to tag documents that don't contain attorney-client privileged content.

   If you select this model, you create a tag group with two child tags (one child tag named **Positive** and the other named **Negative**) for the review set. In this example, each child tag corresponds to one of the possible outputs from the attorney-client privilege detection model.
4. Optionally, rename the tag group and the child tags. For example, you could rename the **Positive** tag to **Privileged** and the **Negative** tag to **Not privileged**.

How to use smart tags

When reviewing a document, the model's results display next to the appropriate child tag. For example, if you have a smart tag group for attorney-client privilege detection and you review a document that is potentially privileged, the reason for that conclusion displays next to the appropriate tag. The tag isn't automatically applied to the document. The reviewer makes the decision about how to tag the document.

Monitor tagging jobs

When you tag a large number of items or select the **Tag all items in list** option, a **Tagging documents** job is created. You can view the status of this job on the **Jobs** tab in the case. The **Jobs** tab helps you track large tagging jobs that might take a long time to complete. In some cases, a tagging job might be complete, but the **Applying tags** status message in the tagging panel still displays. To update the status of tagging jobs, select **Refresh** in the review set command bar.

Remove tags

You can remove tags from items in a review set. However, you can't remove a single-choice tag that you applied to a review set item. You can only change a single-choice tag to another single-choice tag within the same tag group.

To remove a tag:

1. Select the items that you want to remove the tag from.
2. Select **Tag files** to display the tagging panel.
3. Under **Assign tags**, clear the tag, then select **Apply tags**.

You can also use the previous procedure to change the tag applied to selected items. After clearing the current tag, you can select a different one.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/19/2026

Analyze data in a review set in eDiscovery

When you collect a large number of documents, it can be difficult to review them all. eDiscovery provides tools to analyze the documents so you can reduce the number of documents to review without losing any information. eDiscovery analysis tools also help you organize the documents in a coherent manner. For more information about these capabilities, see:

- Near duplicate detection
- Email threading
- Themes

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Run analytics for a review set

To analyze data in a review set, complete the following steps:

1. Go to the Microsoft Purview portal and sign in with the credentials for a user account assigned eDiscovery permissions.
2. Select the **eDiscovery** solution card, and then select **Cases** in the left navigation.
3. Select a case.
   a. (Optional) select **Case settings**.
   b. On **Case settings**, select **Search & analytics**.
   c. On **Search & analytics**, configure analytics settings for your case.
   d. After selecting the applicable search and analytics options, select **Save**. Close **Case settings**.
4. Select the **Review sets** tab.
5. Select a review set, then select the **All items** tab.
6. Select **Analytics** > **Run document & email analytics**.
7. A notification dialog appears about the time it takes to complete the analytics process. Select **Yes** to continue with running the analytics process or select **Cancel**.
8. Select **OK** when notified that the analyze process is created.

You can check the progress of the analysis process by selecting **Process manager** in the review set. The process type is *Run analytics*.

After the analysis process completes, you can view the analytics report, run queries within your review set on outputs of the analysis, and see related documents of a given document.

Use the For Review (preview) filter query

After running analytics for the review set, use an automatically generated filter query (called *For Review*) to filter your review. This query excludes immaterial, duplicate, or noninclusive items. The For Review filter query leaves you with only the items that are representative, unique, and inclusive in the review set.

To apply the **For Review - Unique items only (preview)** filter query to a review set, select the **Saved filter queries** dropdown list, and then select **[AutoGen] For Review - Unique items only (preview)**. Minimize the query details to view the results.

Here's the syntax for the **For Review - Unique items only (preview)** filter query in Keyword Query Language (KeyQL) format:

```
(((FileClass="Email") AND (IsInclusive="True") AND (MarkAsRepresentative="Unique")) OR ((FileClass="Attachment") AND (MarkAsRepresentative="Unique")) OR ((FileClass="Document") AND (MarkAsRepresentative="Unique")) OR ((FileClass="Conversations") AND (MarkAsRepresentative="Unique")) OR ((FileClass="Attachment" OR FileClass="Conversations" OR FileClass="Document" OR FileClass="Email") AND  isempty (MarkAsRepresentative)))
```

The following list describes the result of the For Review - Unique items only (preview) filter query in terms of what content is displayed after you apply it to the review set.

- **Email**: Displays items that are marked as both **Inclusive** and **Representative**. An email is marked as IsInclusive True when it contains all the unique content from a thread, including all previous replies. It ensures that only the most comprehensive email in a thread is reviewed, which is essential for understanding the full context of the conversation without having to review each individual reply. An email is MarkedAsRepresentative Unique when it's the single representative copy among duplicates.
- **Attachments**: Filters out duplicate attachments in the review set. Only attachments that are unique in the review set are displayed.
- **Documents**: Filters out duplicate documents. Only documents that are unique in the review set are displayed.

- **Teams conversations**: Filter out duplicate Teams (and Viva Engage) conversations. Only conversations that are unique in the review set are displayed.
- **Other**: Any emails, attachments, documents, and conversations in the review set where the *MarkedAsRepresentative* field doesn't contain a value. The empty value means that analytics wasn't able to generate a value, so you should review and assess these items.

For more information about inclusive types and document uniqueness, see Email threading in eDiscovery.

Use the Potentially Immaterial Items filter query

The potentially immaterial query is designed to filter out specific types of files that could be immaterial in a review set. It targets two main categories:

1. Compressed files. It filters out files with extensions commonly associated with compressed or archived formats, such as .zip.
2. Small image files. It filters out image files that are smaller than 3,072 bytes (3 KB). These small image files are typically thumbnails, icons, or low-resolution images.

To apply the **Potentially Immaterial Items** filter query to a review set, select the **Saved filter queries** dropdown list, and then select **[AutoGen] Potentially Immaterial Items**. Minimize the query details to view the results.

View the analytics report

To view the analytics report for a review set, complete the following steps:

1. Go to the Microsoft Purview portal and sign in with the credentials for a user account assigned eDiscovery permissions.
2. Select the **eDiscovery** solution card, and then select **Cases** in the left navigation.
3. Select a case, then select the **Review sets** tab.
4. Select a review set, select the **All items** tab.
5. Select **Analytics** > **Show reports**.

The **Analytics** report has seven components from the analysis. Hover over the results to view counts for each component:

- **Target population:** The number of email messages, attachments, and loose documents found in the review set.

- **Documents (excluding attachments):** The number of loose documents that are pivots, unique near duplicates of a pivot, or an exact duplicate of another document.
- **Emails:** The number of email messages that are marked as *inclusive*.
- **Attachments:** The number of email attachments that are unique or duplicates of another email attachment in the review set.
- **Number documents by file type:** The number of files, identified by file extension.
- **Documents by source:** A summary of content by its original data source.
- **Documents aggregated by process:** A summary of content by review set processes.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/29/2026

Export items from a review set in eDiscovery

The export feature lets you customize the content that's included in the download package when you export items from a review set in eDiscovery. Export processes are retained for the life of the case. However, you must download the content from an export process within 30 days after the export process completes.

> **Tip**
>
> Do you prefer an interactive configuration guide experience? Check out the Export data guide.

To export documents from a review set, complete the following steps:

1. Go to the Microsoft Purview portal and sign in with the credentials for a user account assigned eDiscovery permissions.
2. Select the **eDiscovery** solution card, then select **Cases** in the left navigation.
3. Select a case, then select the **Review sets** tab.
4. Select the review set that you want to export and select the **All items** tab.
5. In the review set, select the items to export.
6. Select **Action** > **Export** on the command bar.
7. The **Export** flyout page appears with the settings to configure the export. Some options are selected by default, but you can change these options. See the export options section for descriptions of the options that you can configure.
8. After you configure the export options, select **Export** to start the export process. Go to the **Process manager** to track the progress and to download exported content.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Choose export options for a review set

Use the following options to configure the export. Some output options don't allow all options. For example, when exporting to the PST format, you can't export text files or redacted PDFs.

- **Export name**: Name of the export job. This value names the downloaded ZIP files.
- **Description for the export**: Free-text field for you to add a description.
- **Select items to include in your export**
  - **Selected documents only**: This option exports only the documents that you currently selected. This option is only available when you select items in a review set.
  - **All filtered documents**: This option exports the documents in an active filter. This option is only available when you apply a filter to the review set.
  - **All documents in the review set**: This option exports all documents in the review set.
- **Expand selected documents to include**: Use the following options to export additional items that are related to filtered or selected items in the review set:
  - **Associated family items**: This option includes associated family items. Family items are items that share the same *FamilyId* metadata property value. For example, a document that's attached to an email message shares the same *FamilyId* as the email message. If you select this option, the email message and the document are exported in this example, even though the document might not be included in the list of review set items.
  - **Associated conversation items**: This option includes associated items that are in the same Teams or Viva Engage conversation. Conversation items are items that share the same *ConversationId* metadata property value. All messages, posts, and corresponding transcript file of a conversation share the same *ConversationId*.

    If you select **Associated family items** or **Associated conversation items**, all items that share the same association ID are exported, even though some of those items might not be in the results of the current review set query.
- **Export type**: Exported content is either available for download directly through a web browser or can be sent to an Azure Storage account. The **Export item report only** and **Export items with item report** options enable direct download.
  - **Export item report only**: Only the summary and item report are created. When you select “Export item report only,” the various options for organizing data, folder and path structure, condensing paths, and other options are hidden to avoid any confusion.
    - *Export tags*: Includes tags you created and applied during review in the Item report.
  - **Export items with item report**: Items are exported with the item report.

- *Export redaction*: When selected, items in the export package that have committed redactions are delivered as the redacted PDF (`{itemId}_burn.pdf`) instead of the original native file. Specifically:
  - **Individual documents and email attachments** with committed redactions: the native file is **replaced** by the redacted PDF in the export package. The original native isn't included alongside.
  - **Conversations** (for example, Microsoft Teams or Viva Engage threads): both the native conversation file and the redacted PDF are included.
  - **Extracted text** files for redacted items contain the OCR'd text from the redacted PDF, not the original text.
  - **Family expansion** applies to redaction the same way it applies to the rest of the export: a redacted child attachment of a responsive parent is included as the redacted PDF.

  If you need both the original native file and the redacted PDF for the same item, run two exports — one with *Export redaction* selected and one without — and combine the packages.

  If you don't select this option, the export contains original native files only, even for items that have committed redactions.

- *Export text files*: This option exports the extracted text of the document in a separate file.

- *Export tags*: This option includes tags you created and applied during review in the Item report.

- *Convert supported file formats to HTML*: This option converts supported cloud-native file formats, such as `.loop` and `.page` files (used by Copilot Notebook and Page content), to HTML during export. This option is useful when your organization's post-export workflows support HTML but lack tooling for native cloud file formats. This option isn't selected by default.

- **Export format**: Exported content options for Exchange items.
  - **Create PSTs for messages where possible**: This option creates .pst files for messages.
  - **Create .msg files for messages**: This option creates .msg files for messages.

  The following settings control whether the output package contains splits based on location of the item and whether to include folder structure and naming convention.

- In the **Package size settings** section, configure the maximum export package size the export process creates.
  - If you chose *Create PSTs for messages*, you can configure the Maximum PST package size to 1 GB, 2 GB, 5 GB (default), or 10 GB.
  - You can configure the maximum .zip package size to 2 GB, 10 GB (default), 20 GB, or 40 GB. The maximum .zip package size needs to be the same or greater than the maximum PST package size to contain the PST files in .zip packages when downloaded.
  - Select one or more of the following options:
    - **Organize data from different locations into separate folders or PSTs**: This option organizes data into separate folders for each data location. When exporting mailbox content, items from the primary mailbox and any associated archive mailboxes (including multiple archives) are merged into a single PST file, even if this export option is selected.
    - **Include folder and path structure of the source**: This option includes the original folder and folder path structure for items.
    - **Condense paths to fit within 259 characters**: This option condenses the folder path for each item to 259 characters or less.
    - **Give each item a friendly name**: This option creates a friendly name for each item.

    The option for a condensed directory structure was previously available in the legacy eDiscovery experience. For a similar experience, deselect the following options in the **Export format** section:
    - *Organize data from different locations into separate folders or PSTs*
    - *Include folder and path structure of the source*
    - *Give each item a friendly name*

For more information about export file types and `item.csv` report fields, see Document metadata fields in eDiscovery.

Monitor your export process and view export settings

After you submit your export process, you can monitor the progress in the Process manager. You can also view the settings you chose for the export and download information for this export process in a consolidated report. If you're working with Microsoft support, you can copy process job information to share with your support representative.

To improve export performance, reduce the size of the exported package using date ranges or other types of filters to decrease the total size of the search results. If an export runs for more than seven days without completing, it's automatically canceled. To prevent cancellation, divide exports into smaller batches.

View completed exports from a review set

To view a list and details about completed exports from a review set, select the **Exports** tab in a case. On the **Exports** tab, you see the following details for each export process:

- **Export name**: The name of the export.
- **Status**: The current status of the export process.
- **Review set name**: The name of the review set from which the export was created (when the export is created from a review set).
- **Search name**: The name of the search from which the export was created (when the export is created from a search).
- **Created**: The UTC date and time the export was created.
- **Completed**: The UTC date and time the export was completed.
- **Duration**: The duration of the export process.
- **Created by**: The user that created the export process.

> **Tip**
>
> You can configure the **Microsoft Purview eDiscovery API** to enable the programmatic download of export packages and the reports from an export process in an eDiscovery case.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

Last updated on 06/29/2026

Add items from a review set to another review set in eDiscovery

In some cases, you might need to select documents from one review set and work with them individually in another review set. This approach is especially useful if you culled content in a review set and want to run analytics on the subset of data.

Follow the workflow in this article to add content from one review set to another.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Create a review set

Before you start, you need to create a review set to add the data to. You can add a new review set on the **Review sets** tab of the case. For more information, see Create a review set.

Step 1: Identify content to add to another review set

You can add content from one review set to another review set by selecting specific documents in the source review set or by selecting all items returned by a review set query, which filters items based on search criteria you define. If you're adding selected items, select the items, select **Action**, and then select **Add to another review set**.

Step 2: Specify options for adding to another review set

In the **Add to another review set** flyout, complete the following steps:

1. Select the review set you want to add the items to.

2. Choose from the following content options:

   - **Selected documents only**
   - **All filtered documents**
   - **All documents in the review set**

3. In the **Additional information** section, select options to include all metadata (by selecting the **Metadata** check box) from the items and whether to include the tags (by selecting the **Tags** check box) from the source review set when the documents are added to the new review set.

4. After you select **OK**, a new process named **Adding data to another review set** is created to add the content to another review set. You can go to the **Process manager** in the new review set to monitor the progress of this process.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/29/2026

Import external data into a review set in eDiscovery

Not all documents that you need to analyze in Microsoft Purview eDiscovery are in Microsoft 365. With the import external data feature in eDiscovery with premium features, you can upload documents that aren't in Microsoft 365 to a review set. This article shows you how to bring your external data into eDiscovery for analysis.

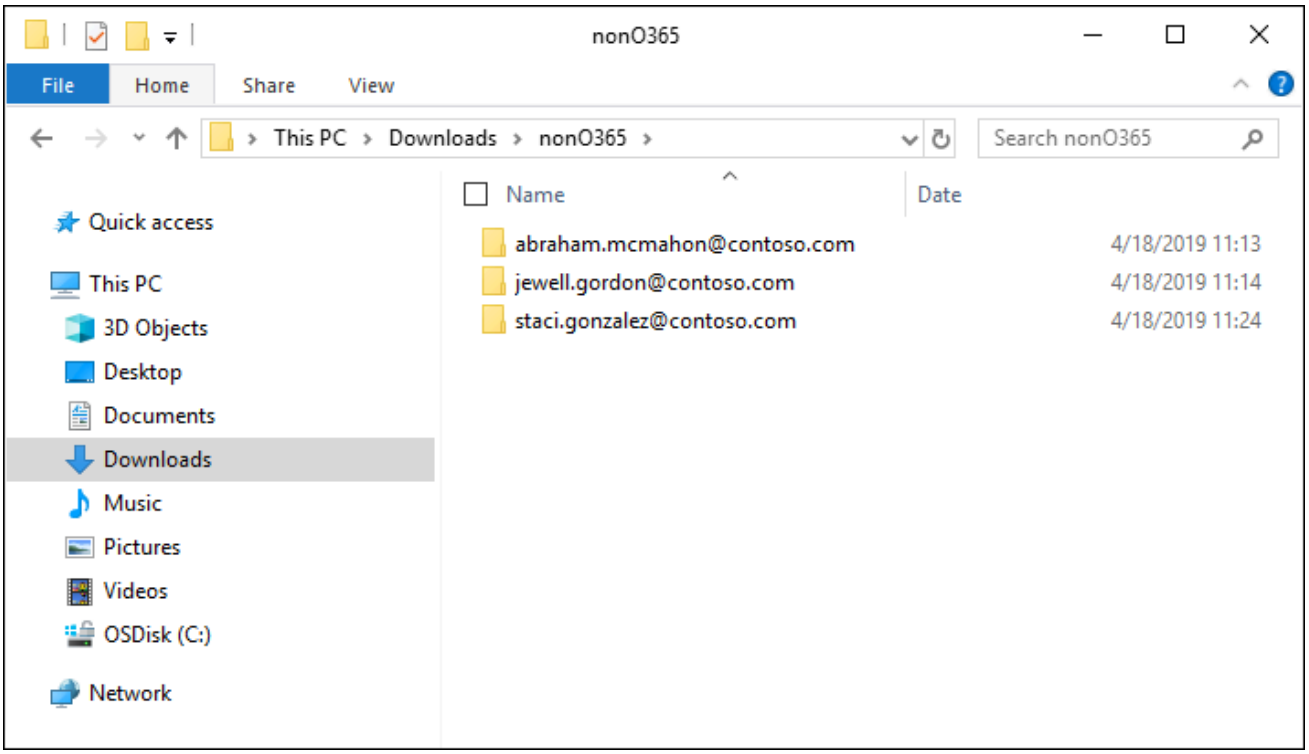
Requirements to import external data

To use the import external data feature in eDiscovery, you need the following requirements:

- eDiscovery with eDiscovery premium features enabled in your organization.
- All custodians that you want to associate external data to must have the appropriate license. For more information, see Get started with eDiscovery.
- An existing eDiscovery case.
- External data must be a file type supported by eDiscovery. For more information, see Supported file types in eDiscovery.
- All files that you upload to a review set must be located in folders, where each folder is associated with a specific custodian. The names for these folders must use the following naming format: *alias@domainname*. The alias@domainname must be the user's Microsoft 365 alias and domain. You can collect all the alias@domainname folders in a root folder. The root folder can only contain the alias@domainname folders. Loose files in the root folder aren't supported.

  The folder structure for the external data that you want to upload is similar to the following example:
  - c:\nonO365\abraham.mcmahon@contoso.com
  - c:\nonO365\jewell.gordon@contoso.com
  - c:\nonO365\staci.gonzalez@contoso.com

  Where abraham.mcmahon@contoso.com, jewell.gordon@contoso.com, and staci.gonzalez@contoso.com are the SMTP addresses of custodians in the case.



- An account that's assigned to the eDiscovery Manager role group (and added as eDiscovery Administrator). For more information about role groups, see Assign permissions in eDiscovery.

- The AzCopy tool installed on a computer that has access to the external data folder structure. To install AzCopy, see Get started with AzCopy. Be sure to install AzCopy in the default location, which is **%ProgramFiles(x86)%\Microsoft SDKs\Azure\AzCopy**. Use the latest version of AzCopy. Other versions of AzCopy might not work when loading external data in eDiscovery.

Import external data into eDiscovery

To import external data content into eDiscovery, complete the following steps:

1. Go to the Microsoft Purview portal and sign in with the credentials for a user account assigned eDiscovery *Administrator* permissions.

2. Select the **eDiscovery** solution card and then select **Cases** in the left navigation.

3. Select a case, then select the **Review sets** tab.

4. Select the review set that you want to import.

5. In the review set, select **Manage**, and then select **Import external data**.

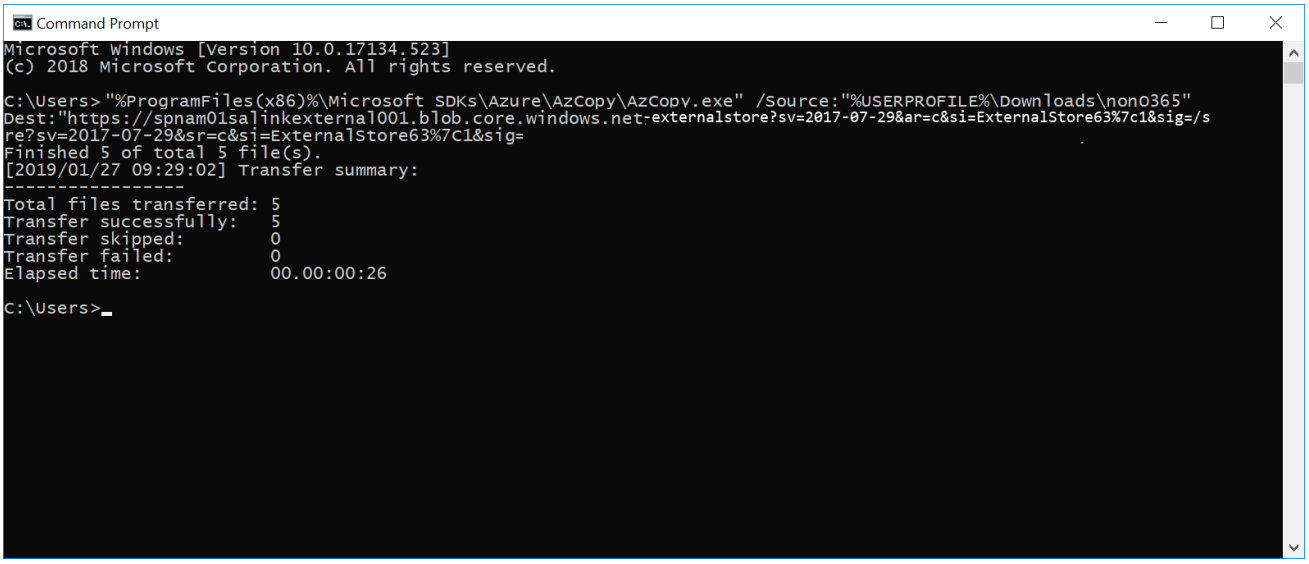
6. Select **Upload files** to start the data import workflow. The first step in the workflow prepares a secure Microsoft-provided Azure Storage location to upload the files to. When the preparation is completed, **Next: Upload files** becomes active.

7. Select **Next: Upload files**.

8. On the **Upload files** page, complete the following steps:

    - In the **Path to location of files** box, verify or type the location of the root folder where you stored the external data you want to upload. For example, for the location of the example files shown in the Requirements to import external data section, type **%USERPROFILE\Downloads\nonO365**. Providing the correct location ensures the AzCopy command displayed in the box under the path is properly updated.
    - Select **Copy to clipboard** to copy the command that is displayed in the box.

9. Start a Windows command prompt, paste the command that you copied in the previous step, and then select **Enter** to start the AzCopy command. After you start the command, the external data files are uploaded to the Azure Storage location that was prepared in step 6.



> **ⓘ Important**
>
> Use the latest version of AzCopy to successfully use the command that's provided on the **Upload files** page.

10. Go back to the Microsoft Purview portal, and select **Next: Process files** in the workflow. Selecting **Next: Process files** initiates processing, text extraction, and indexing of the external data files that you uploaded to the Azure Storage location.

11. Track the progress of processing the files on the **Process files** page or on the Process manager by viewing a job named **Adding external data to a review set**. After the job finishes, the new files are available in the review set.

12. After the processing finishes, you can close the workflow.

Review the upload process report

Use Process Manager to track the upload process. When the process finishes, download a process report that provides information about the imported external data.

The upload Non-Office 365 process report includes both a Summary CSV and an Items CSV report. These reports provide details about the upload process and a list of items added to the review set and associated metadata.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/19/2026

Use the Query Report to create keyword reports for KeyQL queries

When you filter in a review set, it can be difficult to compare filtered results when using multiple saved queries. The Query Report helps you generate and download a consolidated report on multiple queries for a review set. With the consolidated query report, you can quickly see the total count and volume of filtered items on a particular keyword search or multiple compound KeyQL queries.

Select an individual query report item to review specific filtered items. You can select individual items and view the source, plain text, and metadata information for an item.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Create a query report

To create a query report in a review set, complete the following steps:

1. Go to the Microsoft Purview portal and sign in with the credentials for a user account assigned eDiscovery permissions.
2. Select the **eDiscovery** solution card, then select **Cases**.
3. Select a case, then select the **Review sets** tab.
4. Select a review set, then select the **All items** tab.
5. Select **Actions** > **Query Report**.
6. On the **Query Report** page, enter your search queries and separate each with a line break (pressing enter on your keyboard). A report is generated for each search query line.

   For example, if you want to filter by:

   - a keyword *confidential* in items for one query,
   - a filter for two keywords *confidential* and *Project Infinity* in items for another query,

- and filter using a complex KeyQL statement *(((FileClass="Email") AND (InclusiveType="InclusiveMinus" OR InclusiveType="Inclusive")) OR ((FileClass="Attachment") AND (UniqueInEmailSet="true")) OR ((FileClass="Document") AND (MarkAsRepresentative="Unique")) OR (FileClass="Conversations"))* for a third query,

You would use the following values on separate lines in the **Query Report (preview)** search query field:

```
confidential
confidential AND Project Infinity
(((FileClass="Email") AND (InclusiveType="InclusiveMinus" OR InclusiveType="Inclusive")) OR ((FileClass="Attachment") AND (UniqueInEmailSet="true")) OR ((FileClass="Document") AND (MarkAsRepresentative="Unique")) OR (FileClass="Conversations"))
```

7. Select **Generate report**.

The results of the queries are summarized with the following information displayed for each query:

- **KeyQL**: The Keyword Query Language (KeyQL) statement used for the filter query.
- **Locations with hits**: The number of locations with filter query hits.
- **Count**: The number of items in the filter result.
- **Percentage**: The percentage of hits in the total count of items.
- **Size**: The file size of the items with hits.

Use the Query Report

To view the individual items that the filter queries return, double-select the filter query line in the report. Select an individual item to view item details. The item information is read-only, so you can't annotate, modify, or export these items.

To download the report summary information for the filter queries, select **Download report** to create a .csv file containing the report summary information. When you exit the Query Report, the filter queries aren't saved.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/29/2026

Create holds in eDiscovery

Create holds to preserve content that might be relevant to an eDiscovery case. Place a hold on the Exchange mailboxes and OneDrive accounts of people you're investigating in the case. You can also place a hold on the mailboxes and sites that are associated with Microsoft Teams, Microsoft 365 groups, and Viva Engage Groups. When you place content locations on hold, you preserve content until you remove the content location from the hold or until you delete or release the hold.

> **Important**
>
> After you create an eDiscovery hold, it might take up to 24 hours for the hold to take effect. For long term data retention not related to eDiscovery investigations, use retention policies and retention labels. For more information, see Learn about retention policies and retention labels.

eDiscovery holds **require an eligible license on every user whose content you preserve**. Frontline subscriptions such as Office 365 F1 and F3 don't include the rights required to be the *target* of an eDiscovery hold. Before you add a user to a hold, confirm the user's assigned subscription includes eDiscovery hold rights.

The eDiscovery (Premium) hold capabilities required for query-based, time-bound, and custodian-scoped holds are included only in higher tiers (for example, Microsoft 365 E5 / A5 / G5 or the Microsoft 365 E5 Compliance and E5 eDiscovery & Audit add-ons).

For the authoritative list of which hold capabilities are included in each subscription, see Microsoft 365 guidance for security & compliance. If the user's current subscription doesn't include the required rights, assign an add-on (such as Microsoft 365 E5 Compliance) **before** creating the hold. Placing a user on hold without proper licensing isn't supported.

When you create a hold, use the following options to scope the content that's preserved in the specified content locations:

- Create an infinite hold where all content in the specified locations is placed on hold. Alternatively, create a query-based hold where only the content in the specified locations that matches a search query is placed on hold.

- Specify a date range to preserve only the content that was sent, received, or created within that range. Alternatively, hold all content in specified locations regardless of when it was sent, received, or created.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Create a hold

> **Tip**
>
> Do you prefer an interactive configuration guide experience? Check out the Apply a hold guide.

To create an eDiscovery hold associated with an eDiscovery case, complete the following steps:

1. Go to the Microsoft Purview portal and sign in with credentials for a user account assigned eDiscovery permissions.
2. Select the **eDiscovery** solution card and then select **Cases** in the left navigation.
3. Select a case, and then select the **Hold policies** tab.
4. On the **Hold policies** dashboard, select **Create policy**.
5. On the **Enter details to get started** page, complete the following fields:
   - **Policy name**: Give the hold policy a name (required). The hold policy name must be unique in your organization.
   - **Policy description**: Add an optional description to help others understand this hold policy.
6. Select **Create** to create the new hold policy and start your hold on relevant data for the case.
7. To add data sources to a hold policy, see Data sources in eDiscovery for step-by-step guidance.

8. On the **Manage data sources** flyout pane, add or remove data sources for your hold policy. You can choose one or more users, groups, or organization locations.

- You must select at least one data source to create a hold policy.
- Enter the specific users, groups, or organization locations you want to add to the hold policy.

  Distribution groups: When you create a hold and select a distribution group as a data source, all current members of the distribution group are listed. You can select all mailboxes or sites for each user to include in the hold. For more information about preserving Bcc and expanded distribution group recipients, see Preserve Bcc and expanded distribution group recipients.

  > ⓘ **Important**
  >
  > **Group membership is captured as a point-in-time snapshot when the hold policy is created.** This applies to every supported group type – distribution list, mail-enabled security group, Microsoft 365 group, Microsoft Teams group, and Viva Engage group. When you add any group as a data source, members are expanded and selected at the moment the policy is created. Members added to the group later are **not** automatically placed on hold; members removed from the group remain on hold until you update or release the policy. To reflect current membership, add the group again as a data source. Search behavior differs, search re-resolves group membership each time the search runs. For more information, see **Data sources in eDiscovery**.
  >
  > Distribution list expansion is limited to a maximum of 100 members. If a distribution list contains more than 100 members, the expansion might fail and the distribution list isn't added successfully. In this case, split the distribution list into smaller groups before adding them as data sources, or use **bulk import data sources**.

  Exchange mailboxes: Applicable checkboxes are selected for the mailboxes placed on hold. Use **Edit** to find user mailboxes and distribution groups (the mailboxes of group members) to place on hold. You can also place a hold on the associated mailbox for a Microsoft Team, Microsoft 365 group, and Viva Engage Group. For more information about the application data that is preserved when a mailbox is placed on hold, see Content stored in mailboxes for eDiscovery.

SharePoint sites: Applicable checkboxes are selected for SharePoint sites and OneDrive accounts placed on hold. Use **Edit** to enter the URL for additional sites that you want to place on hold. You can also add the URL for the SharePoint site for a Microsoft Team, Microsoft 365 group, or Viva Engage group. To avoid potential URL truncation, ensure SharePoint site names stay under 70 characters. Hold policies might not be able to locate sites for truncated site URLs and might have access errors. Placing a hold on a subsite isn't allowed even if the subsite is displayed in the location picker. If you select a subsite, the hold is placed on the parent SharePoint site instead.

> **ⓘ Important**
>
> The Recycle Bin in SharePoint sites isn't indexed and therefore unavailable for searching. As a result, eDiscovery searches can't find any Recycle Bin content to place holds.
>
> eDiscovery doesn't continuously monitor location-level changes after a hold is applied. If mailbox or site identities change later (for example, SharePoint or OneDrive URL changes), or if the hold state changes outside the policy workflow, edit the hold policy data sources and reapply or retry the policy so hold status and scope stay accurate.

1. Select **Save**. You've now scoped the data sources for your hold policy.

2. To define the parameters of your hold policy, choose from the following options on the **Hold policy** tab:

   - **Condition builder**: The condition builder option in search provides a visual filtering experience when you build hold policies. Each condition adds a clause that is created and run when you create the hold. For example, you can specify a date range so that email or site documents that were created within the date range are preserved. For detailed information about using the condition builder option, see Use the condition builder to create search queries in eDiscovery.

   - **Keyword Query Language (KeyQL)**: The Keyword Query Language (KeyQL) query option in search provides guidance and lets you quickly paste long, complex queries directly into the editor. For detailed information about building search queries with the KeyQL option, see Use Keyword Query Language to create search queries in eDiscovery.

You can also quickly build KeyQL queries for your search using Microsoft Security Copilot. For more information, see Create a KeyQL search query with Microsoft Copilot (preview).

> **Important**
>
> For the best results when dealing with encrypted or partially indexed items, limit conditions to *Date*, *Participants*, and *Type* in query-based holds. Queries aren't effective on other conditions within encrypted or partially indexed items and holds might not be applied to these items.

3. Select **Apply hold**.

> **Important**
>
> To apply a hold to a SharePoint site, the site must have a title.

After creating a hold, check that the hold is applied successfully by navigating to the **Details** tab for the hold policy. You can review the following information for the hold policy:

- **Name**: The name of the data source.
- **Location**: The specific location (such as the SMTP address of the mailbox or the URL of the site) of the data sources included in the hold policy.
- **Hold status**: The hold status of the location. Indicates whether the location is on hold, not on hold, or if there's an error with the hold.
- **Source type**: Shows whether the data source type is *People* or *Group*.
- **Location type**: Shows whether the location is *Mailbox* or *Site*.

> **Note**
>
> When you create a query-based hold, you initially place all content from the selected locations on hold. Later, the system clears any content that doesn't match the specified query from the hold every seven to 14 days. However, a query-based hold doesn't clear content if you apply more than five holds of any type to a content location. The hold always includes partially or unindexed items from a location with a query-based hold.

Create a hold from a search

You can create an eDiscovery hold directly from search. However, creating a hold directly from a search has the following restrictions:

- The hold removes tenant-wide data sources contained in the search. Holds don't support tenant-wide data sources.
- The hold expands distribution lists contained in the search. You can choose to place all members' mailboxes and sites on hold or a subset or mix of these data sources on hold. Subsequent changes in distribution list membership don't change or update holds or the policy. If the distribution list is empty, it's removed from the hold because there's nothing to expand.
- The query used in the search isn't incorporated into the hold policy. The data sources in scope for the search are included in the hold policy.

To create an eDiscovery hold from a search, complete the following steps:

1. Complete the steps for creating a search query.
2. Select **Create a hold** on the command bar for the search.
3. View the new hold policy and update as needed.

Monitor hold processing in Process manager

The **Process manager** displays information about processes performed on the hold policy.

- **Process type**: The type of process.
- **Status**: The status of the process.
- **Created**: The date and time the process was created.
- **Completed**: The date and time the process was completed.
- **Duration**: The duration of the process.
- **Created by**: The user that created the process.

To download the list of processes and the column information, select **Download list** to create a .csv file containing this information.

To view information about all eDiscovery processes, see Use the Process report in eDiscovery.

Query-based holds placed on sites

Keep the following points in mind when you place a query-based eDiscovery hold on documents located in SharePoint sites:

- A query-based hold initially preserves all documents in a site for a short period of time after you delete them. When you delete a document, you move it to the Preservation Hold library even if it doesn't match the criteria of the query-based hold. However, a timer job removes deleted documents that don't match a query-based hold. The timer job runs periodically and compares all documents in the Preservation Hold library to your query-based eDiscovery holds (and other types of holds and retention policies). The timer job deletes the documents that don't match a query-based hold and preserves the documents that do.
- Use query-based holds to perform targeted preservation such as keywords, date ranges, or other document properties to preserve site documents.

Preserve content in Microsoft Teams

Conversations that are part of a Microsoft Teams channel are stored in the mailbox that's associated with the Microsoft Team. Similarly, files that team members share in a channel are stored on the team's SharePoint site. Therefore, you have to place the Team mailbox and SharePoint site on eDiscovery hold to preserve conversations and files in a channel.

Alternatively, conversations that are part of the Chat list in Teams (called *1:1 chats* or *1:N group chats*) are stored in the mailboxes of the users who participate in the chat. Files that users share in chat conversations are stored in the OneDrive account of the user who shares the file. Therefore, you have to add the individual user mailboxes and OneDrive accounts to an eDiscovery hold to preserve conversations and files in the chat list. It's a good idea to place a hold on the mailboxes of members of a Microsoft Team in addition to placing the team mailbox and site on hold.

> **Note**
>
> If your organization has an Exchange hybrid deployment (or your organization synchronizes an on-premises Exchange organization with Microsoft 365) and has enabled Microsoft Teams, on-premises users can use the Teams chat application and participate in 1:1 chats and 1:N group chats. These conversations are stored in cloud-based storage that's associated with an on-premises user. If you place an on-premises user on an eDiscovery hold, the Teams chat content in the cloud-based storage is preserved. For more information, see Find content in mailboxes.

Preserve card content

Card content generated by apps in Teams channels, 1:1 chats, and 1:N group chats are stored in mailboxes and are preserved when you place a mailbox on an eDiscovery hold. A *card* is a UI

container for short pieces of content. Cards can have multiple properties and attachments, and can include items that trigger card actions. For more information, see Cards. Like other Teams content, where card content is stored is based on where the card was used. Content for cards used in a Teams channel is stored in the Teams group mailbox. Card content for 1:1 and 1xN chats are stored in the mailboxes of the chat participants.

Preserve meeting and call information

Summary information for meetings and calls in a Teams channel is also stored in the mailboxes of users who dialed into the meeting or call. This content is also preserved when you place an eDiscovery hold on user mailboxes.

Preserve content in private channels

You can also preserve content in private channels. Private channel chats are stored in a dedicated private channel mailbox. Private channel message data from each user's mailbox is moved to the private channel's mailbox and message edits and deleted messages aren't copied into the mailbox. For users on hold, copies of message edits and deleted messages remain in the user's preserved library folder until the hold expires, while the latest version of private channel message data from each user's mailbox is copied to the private channel's mailbox.

Preserve wiki content

Every team or team channel also contains a Wiki for note taking and collaboration. The Wiki content automatically saves to a file in the .mht format. The Teams Wiki Data document library on the team's SharePoint site stores this file. You can preserve the wiki content by adding the team's SharePoint site to an eDiscovery hold.

> **Note**
>
> The capability to preserve Wiki content for a Team or team channel (when you place the team's SharePoint site on hold) was released on June 22, 2017. If a team site is on hold, the Wiki content is retained starting on that date. However, if a team site is on hold and the Wiki content was deleted before June 22, 2017, the Wiki content isn't preserved.

Microsoft 365 groups

Teams is built on Microsoft 365 groups. Therefore, placing Microsoft 365 groups on eDiscovery hold is similar to placing Teams content on hold.

Keep the following things in mind when placing both Teams and Microsoft 365 groups on an eDiscovery hold:

- To place content located in Teams and Microsoft 365 groups on hold, you must specify the mailbox and SharePoint site that are associated with a group or team.
- Run the **Get-UnifiedGroup** cmdlet in Exchange Online PowerShell to view properties for Teams and Microsoft 365 groups. This method is a good way to get the URL for the site that's associated with a Team or Microsoft 365 group. For example, the following command displays selected properties for a Microsoft 365 group named Senior Leadership Team:

  text

  ```
  Get-UnifiedGroup "Senior Leadership Team" | FL
  DisplayName,Alias,PrimarySmtpAddress,SharePointSiteUrl

  DisplayName            : Senior Leadership Team
  Alias                  : seniorleadershipteam
  PrimarySmtpAddress     : seniorleadershipteam@contoso.onmicrosoft.com
  SharePointSiteUrl      :
  https://contoso.sharepoint.com/sites/seniorleadershipteam
  ```

  > **Note**
  >
  > To run the **Get-UnifiedGroup** cmdlet, you need the View-Only Recipients role in Exchange Online or membership in a role group that's assigned the View-Only Recipients role.

- When you search a user's mailbox, the search doesn't include any Team or Microsoft 365 group that the user is a member of. Similarly, when you place a Team or Microsoft 365 group on eDiscovery hold, the hold applies only to the group mailbox and group site. The mailboxes and OneDrive sites of group members aren't placed on hold unless you explicitly add them to the eDiscovery hold. So if you have to place a Team or Microsoft 365 group on hold for a legal reason, consider adding the mailboxes and OneDrive accounts of team or group members on the same hold.
- To get a list of the members of a Team or Microsoft 365 group, you can view the properties on the **Groups** page in the Microsoft 365 admin center. Alternatively, you can run the following command in Exchange Online PowerShell:

  PowerShell

```
Get-UnifiedGroupLinks <group or team name> -LinkType Members | FL
DisplayName,PrimarySmtpAddress
```

> **Note**
>
> To run the **Get-UnifiedGroupLinks** cmdlet, you need the View-Only Recipients role in Exchange Online or membership in a role group that's assigned the View-Only Recipients role.

Preserve content in OneDrive accounts

> **Important**
>
> The retention period for deleted OneDrive accounts is different from the retention period for mailboxes. For more information about the retention period for deleted OneDrive accounts, see OneDrive retention and deletion.

To collect a list of the URLs for the OneDrive sites in your organization so you can add them to a hold or search associated with an eDiscovery case, see Create a list of all OneDrive locations in your organization.

The script in this article creates a text file that contains a list of all OneDrive sites in your organization. To run this script, you need to install and use the SharePoint Online Management Shell. Be sure to append the URL for your organization's MySite domain to each OneDrive site that you want to search. This domain contains all your OneDrive sites; for example, `https://contoso-my.sharepoint.com`. Here's an example of a URL for a user's OneDrive site: `https://contoso-my.sharepoint.com/personal/sarad_contoso_onmicrosoft.com`.

> **Important**
>
> The URL for a user's OneDrive account includes their user principal name (UPN) (for example, `https://alpinehouse-my.sharepoint.com/personal/sarad_alpinehouse_onmicrosoft_com`). In the rare case that a person's UPN is changed, their OneDrive URL also changes to incorporate the new UPN. If a user's OneDrive account is part of an eDiscovery hold, and their UPN is changed, you need to update the hold by adding the user's new OneDrive URL and removing the old one. If the URL for the OneDrive site changes, previously placed holds on the site remain effective and

content is preserved. For more information, see [How UPN changes affect the OneDrive URL](#).

Remove content locations from an eDiscovery hold

After you remove a mailbox, SharePoint site, or OneDrive account from an eDiscovery hold, a *delay hold* is applied. This hold delays the actual removal of the hold for 30 days to prevent data from being permanently deleted (purged) from a content location. This delay gives admins an opportunity to search for or recover content that is purged after an eDiscovery hold is removed. The details of how the delay hold works for mailboxes and sites are different.

- **Mailboxes:** A delay hold is placed on a mailbox the next time the Managed Folder Assistant processes the mailbox and detects that an eDiscovery hold was removed. Specifically, a delay hold is applied to a mailbox when the Managed Folder Assistant sets one of the following mailbox properties to **True**:
  - **DelayHoldApplied:** This property applies to email-related content (generated by people using Outlook and Outlook on the web) that's stored in a user's mailbox.
  - **DelayReleaseHoldApplied:** This property applies to cloud-based content (generated by non-Outlook apps such as Microsoft Teams, Microsoft Forms, and Microsoft Viva Engage) that's stored in a user's mailbox. Cloud data generated by a Microsoft app is typically stored in a hidden folder in a user's mailbox.

  When a delay hold is placed on the mailbox (when either of the previous properties is set to **True**), the mailbox is still considered to be on hold for an unlimited hold duration, as if the mailbox was on Litigation Hold. After 30 days, the delay hold expires, and Microsoft 365 automatically attempts to remove the delay hold (by setting the DelayHoldApplied or DelayReleaseHoldApplied property to **False**) so that the hold is removed. After either of these properties are set to **False**, the corresponding items that are marked for removal are purged the next time the mailbox is processed by the Managed Folder Assistant.

- **SharePoint and OneDrive sites:** Any SharePoint or OneDrive content that's being retained in the Preservation Hold library isn't deleted during the 30-day delay hold period after a site is removed from an eDiscovery hold. Retention of content in the Preservation Hold library during the 30-day delay hold period is similar to what happens when a site is released from a retention policy. Additionally, you can't manually delete this content in the Preservation Hold library during the 30-day delay hold period. To release a site from the 30-day delay hold/grace period hold, see the Can't delete a site because of an invalid retention policy or eDiscovery hold troubleshooting article.

For more information, see Releasing a policy for retention.

A delay hold is also applied to content locations on hold when you close an eDiscovery case because holds are turned off when a case is closed. For more information about closing a case, see Learn about case settings in eDiscovery.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/19/2026

Manage holds in eDiscovery

> **Note**
>
> Active hold policies that you create in cases in the legacy eDiscovery experience are automatically migrated and available in the new Microsoft Purview eDiscovery experience. Review the eDiscovery hold policy status report to check policy status and retry any policies that need attention.

Use eDiscovery to create hold policies that preserve content relevant to your cases. When you place content locations on hold, the content is held until you release the hold in the case, remove a specific data location, or delete the hold policy. You can place holds on data sources such as:

- User mailboxes and OneDrive sites.
- Microsoft Teams mailboxes and SharePoint and OneDrive sites.
- Microsoft 365 group mailboxes and SharePoint and OneDrive sites.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Hold policy dashboard

The **Hold policies** dashboard lists all the holds associated with a case. This dashboard lets you create hold policies, displays information about hold policies in the case, and lets you filter and group the hold policies. The **Hold policies** dashboard contains the following information and controls:

- **Name**: The name of the hold policy.
- **Created by**: The user that created the hold policy.
- **Last modified**: The date and time the hold policy was last modified. Select **Time zone** to switch between using local time and Coordinated Universal Time (UTC).
- **Hold policy Status**: The current status of the hold policy.

Select a hold policy to view the details, data sources, and condition and KeyQL filters.

To customize the columns display on the **Hold policies** dashboard, select **Customize columns** to choose the columns to display or drag and drop the columns in the list to reorder. To search for a specific hold policy, enter a keyword in the **Search** field. To download the list of hold policies and the column information, select **Download list** to create a .csv file containing this information.

Hold policy states

The hold policy page shows the state of a hold next to the hold name. Hold policies have the following states:

- **Draft**: Displayed when you create a new policy but don't apply it. If you navigate away from the policy draft, you cancel the draft and lose all policy changes.
- **On**: The policy is applied and all locations in the policy are on hold. Select the **Details** tab to view location information.
- **Off**: The policy is off for a previously applied hold. All included locations aren't on hold.
- **In progress**: The hold policy is in the process of being applied or updated.
- **Pending deletion**: The hold policy is in the process of being deleted.

Check for holds on a mailbox or site

Before you delete content, release a hold, or diagnose why a purge, site deletion, or storage cleanup isn't working as expected, identify every hold source that applies to the location. eDiscovery hold policies are only one of several hold sources. Microsoft Purview retention policies, retention labels, Litigation Hold, delay holds, and Single Item Recovery can independently preserve content, and any one of them is enough to keep items in the Recoverable Items folder or in the Preservation Hold Library.

Symptoms that warrant a hold check

Expand table

Symptom	Most likely hold source
`New-ComplianceSearchAction -Purge -PurgeType HardDelete` succeeds but items remain in `\Recoverable Items\Purges`	Any hold or retention policy on the mailbox; delay hold; Single Item Recovery
You can't delete a SharePoint site or OneDrive site	eDiscovery hold; tenant-wide, site-, or group-scoped retention policy; delay hold
You can't delete a folder or document library	Site-scoped retention policy; Preservation Hold Library content

Symptom	Most likely hold source
The Preservation Hold Library is consuming OneDrive storage and items aren't expiring	Retention policy or eDiscovery hold still applied to the OneDrive site
Mailbox shows as on hold even after you remove all known holds	Delay hold (`DelayHoldApplied` or `DelayReleaseHoldApplied`); retention label hold (`ComplianceTagHoldApplied`)
You can't permanently delete an inactive mailbox	A hold or retention policy is keeping the mailbox alive

Check for holds on a mailbox

Run the following cmdlets in Exchange Online PowerShell, in order. If any cmdlet returns a hold source, clear it before you attempt to delete content or remove the user from the location.

Expand table

#	Cmdlet	What it surfaces
1	`Invoke-HoldRemovalAssistant -Identity <user>`	All applicable hold sources in a single call. Recommended first step.
2	`Get-OrganizationConfig | FL InPlaceHolds`	Tenant-wide and organization-scoped retention policies. Entries with the `mbx`, `skp`, or `grp` prefix are mailbox-, Teams-, or group-scoped policies that surface at the organization level.
3	`Get-Mailbox <user> | FL LitigationHoldEnabled,InPlaceHolds,ComplianceTagHoldApplied,SingleItemRecoveryEnabled`	Mailbox-scoped holds, retention label hold, and

#	Cmdlet	What it surfaces
		Single Item Recovery.
4	`Get-Mailbox <user> | FL DelayHoldApplied,DelayReleaseHoldApplied`	Delay holds. A delay hold is automatically applied after a hold is removed and preserves items for up to 30 days.

Use the following table to identify the hold source for each `InPlaceHolds` prefix that you see, and where to clear it.

Expand table

Prefix	Hold source	Where to clear it
`mbx`	Microsoft Purview retention policy (mailbox-scoped)	Microsoft Purview portal: **Data lifecycle management** > **Retention policies** > exclude the mailbox
`skp`	Teams retention policy	Microsoft Purview portal: **Data lifecycle management** > **Retention policies** (Teams)
`grp`	Microsoft 365 group retention policy	Microsoft Purview portal: **Data lifecycle management** > **Retention policies**
`UniH`	eDiscovery case hold	Microsoft Purview portal: **eDiscovery** > open the case > **Hold policies**
`cld`	Cloud attachment hold	Microsoft Purview portal: **eDiscovery** > open the case > **Hold policies**

For more detail on each hold type and how to identify the associated retention policy or case, see Identify Exchange mailbox hold types in eDiscovery.

Check for holds on a SharePoint or OneDrive site

Run the following cmdlets, in order. If any cmdlet returns a hold source, clear it before you attempt to delete the site, the library, or content that's preserved in the Preservation Hold Library.

Expand table

#	Cmdlet	What it surfaces
1	`Invoke-HoldRemovalAssistant -Identity <siteUrl-or-userUPN>`	All applicable hold sources in a single call. Recommended first step.
2	`Get-OrganizationConfig | FL InPlaceHolds`	Tenant-wide retention policies. Entries with the `grp` prefix are group- or site-scoped policies.
3	`Get-RetentionCompliancePolicy | Where-Object { $_.SharePointLocation -or $_.OneDriveLocation }`	Retention policies that target SharePoint sites or OneDrive accounts.
4	Delay hold on the site	Surfaced in the `Invoke-HoldRemovalAssistant` output.

If the Preservation Hold Library (PHL) on a OneDrive site is consuming storage and items aren't draining, the most common cause is a retention policy or eDiscovery hold that still applies to the site. After you clear every hold source, the Managed Folder Assistant removes expired items from the PHL over time.

Common misconceptions

- **"Setting MRM to True deletes items from the Purges folder."** The Managed Folder Assistant (MRM/MFA) enforces retention. It isn't a delete mechanism. The supported way to permanently delete items from the Recoverable Items folder is the HardDelete purge described in Delete items in the Recoverable Items folder for mailboxes on hold in eDiscovery, and that purge respects every hold source listed earlier in this section.
- **"I removed the Litigation Hold so I can delete the items now."** Wait up to 240 minutes for the change to apply, then check for a delay hold (`DelayHoldApplied` or `DelayReleaseHoldApplied`) before you attempt to delete content. A delay hold is automatically applied after a hold is removed.
- **"`LitigationHoldEnabled` is False, so the mailbox isn't on hold."** This check alone is insufficient. Also check `InPlaceHolds`, `ComplianceTagHoldApplied`, `DelayHoldApplied`, `DelayReleaseHoldApplied`, and `Get-OrganizationConfig | FL InPlaceHolds`.
- **"`Get-OrganizationConfig` returned an `mbx` entry, so the policy applies to the whole organization."** An `mbx`-prefixed entry that surfaces under `Get-OrganizationConfig` is a mailbox-scoped Microsoft Purview retention policy. To clear it, exclude the specific mailbox from the policy in the Microsoft Purview portal. It might take up to 24 hours for the exclusion to replicate.

Create a hold policy

To create a new hold policy, see Create holds in eDiscovery.

Edit a hold policy

You can edit the hold policy name, description, or the policy details (data sources, condition filters, and KeyQL filters) as applicable.

To edit the hold policy name or description, complete the following steps:

1. Go to the Microsoft Purview portal and sign in with the credentials for a user account assigned eDiscovery permissions.
2. Select the **eDiscovery** solution card, then select **Cases** in the left navigation.
3. Select a case, then select the **Hold policies** tab.
4. On the **Hold policies** dashboard, select the hold policy you want to edit.
5. Select the edit (pencil) icon next to the policy name.
6. Update the policy name or description, then select **Continue**.

To edit hold policy details, complete the following steps:

1. Go to the Microsoft Purview portal and sign in with the credentials for a user account assigned eDiscovery permissions.
2. Select the **eDiscovery** solution card, then select **Cases** in the left navigation.
3. Select a case, then select the **Hold policies** tab.
4. On the **Hold policies** dashboard, select the hold policy you want to edit.
5. On the **Hold policy** page for the selected policy, select the **Hold policy** tab.
6. Update data sources, condition filters, and KeyQL filters as applicable.
7. Select **Apply hold**.

> **Important**
>
> eDiscovery doesn't continuously monitor workload or location-level changes after a hold operation completes. If mailbox or site identities change (for example, UPN, SharePoint URL, or OneDrive URL changes), or if a hold is changed outside this policy workflow, update the hold policy data sources and then select **Apply hold** to restamp the policy.

Retry a hold policy

A retry hold policy triggers the hold process to restamp all mailboxes and sites in the policy to enforce the hold. You might also encounter errors while placing a hold on data sources. For a list of possible errors, see Manage hold status errors.

Use **Retry policy** after location identity changes or external hold operations so that eDiscovery refreshes the policy state and per-location status.

To retry a hold policy, complete the following steps:

1. Go to the Microsoft Purview portal and sign in with the credentials for a user account assigned eDiscovery permissions.
2. Select the **eDiscovery** solution card, then select **Cases** in the left navigation.
3. Select a case, then select the **Hold policies** tab.
4. On the **Hold policies** dashboard, select the hold policy you want to retry.
5. On the **Hold policy** page for the selected policy, select **Policy actions** > **Retry policy**.

Turn off a hold policy

Turning off a hold policy might result in the permanent deletion of any content currently being preserved. It doesn't affect content preserved by other hold policies.

To turn off a hold policy, complete the following steps:

1. Go to the Microsoft Purview portal and sign in with the credentials for a user account assigned eDiscovery permissions.
2. Select the **eDiscovery** solution card, then select **Cases** in the left navigation.
3. Select a case, then select the **Hold policies** tab.
4. On the **Hold policies** dashboard, select the hold policy you want to turn off.
5. On the **Hold policy** page for the selected policy, select **Policy actions** > **Turn off**.

Turn on a hold policy

When you edit a policy, it doesn't affect content preserved by other hold policies.

To turn on a hold policy, complete the following steps:

1. Go to the Microsoft Purview portal and sign in with the credentials for a user account assigned eDiscovery permissions.
2. Select the **eDiscovery** solution card, then select **Cases** in the left navigation.
3. Select a case, then select the **Hold policies** tab.
4. On the **Hold policies** dashboard, select the hold policy you want to turn on.
5. On the **Hold policy** page for the selected policy, select **Policy actions** > **Turn on**.

Delete a hold policy

When you delete a hold policy, you remove all associated holds and release all sites and mailboxes. This action might result in permanent deletion of any content currently being preserved.

To delete a hold policy, complete the following steps:

1. Go to the Microsoft Purview portal and sign in with the credentials for a user account assigned eDiscovery permissions.
2. Select the **eDiscovery** solution card, then select **Cases** in the left navigation.
3. Select a case, then select the **Hold policies** tab.
4. On the **Hold policies** dashboard, select the hold policy you want to delete.
5. On the **Hold policy** page for the selected policy, select **Policy actions** > **Delete policy**.
6. On the **Delete policy?** dialog, select **Yes, delete**.

Manage hold status errors

You might encounter errors while placing a hold on data sources. The following table lists the errors that you might encounter and the recommended resolution.

Expand table

Hold error types	Description	Resolution
Distribution group has too many members	The distribution group associated with the requested hold has more than 1,000 email addresses. Currently, a distribution group with more than 1,000 email addresses can't be expanded and placed on hold.	Add the individual email addresses as data sources or split the distribution group into groups with fewer than 1,000 email addresses. Then select **Policy actions** > **Retry policy** in the hold policy to retry the hold application.
Invalid email address or URL	The location associated with the requested hold has an invalid email address or site URL.	Specify a valid email address or URL that exists within your organization.
Hold changed outside eDiscovery	The hold state for a location was changed outside this case policy workflow (for example, using `Invoke-HoldRemovalAction` in Security & Compliance PowerShell).	In eDiscovery, edit the hold policy so the intended locations are correct, then select **Apply hold** or **Policy actions** > **Retry policy** to restamp and refresh status.
Mailbox not found	The mailbox associated with the requested hold isn't a valid mailbox.	Verify the email address and check that it's a valid Exchange Online mailbox. After you confirm the mailbox, edit the data source for the mailbox and then select **Policy**

Hold error types	Description	Resolution
		actions > **Retry policy** in the hold policy to retry the hold application.
Policy deployment interrupted	A system error indicating a problem was encountered while applying the hold.	Select **Policy actions** > **Retry policy** in the hold policy to retry the hold application.
Site inaccessible	The SharePoint location associated with the requested hold request isn't accessible and might be read only.	Contact your SharePoint site administrator to configure the site as writable. Then select **Policy actions** > **Retry policy** in the hold policy to retry the hold application.
Site failed to apply hold	If the site URL changed for any reason (due to user UPN or tenant domain change), you must update the policy to include the site with the new name and URL.	Update the policy by removing the original site, then add the new site URL to the hold policy and reapply the hold.
Site not found	The SharePoint location associated with the requested hold might have been moved, deleted, or the site URL might not exist.	Check the site URL and confirm if the SharePoint site exists. After you confirm the site exists, edit the data source for the site and then select **Policy actions** > **Retry policy** in the hold policy to retry the hold application.

Hold policies have a *Site not found* error if the policy includes a location that had a UPN change due to change in status, departure policies, name change, or other circumstances. This error occurs when the hold is reapplied after these changes. The hold for site or mailbox can remain stamped on the original location identity, but the policy must be updated to include the new UPN or site URL for the policy to reflect the correct status and to support removal of holds.

After you add the new UPN or site URL to the policy, you can remove the prior SMTP UPN or site URL from the policy.

> **Important**
>
> To apply a hold to a SharePoint site, the site must have a title.

Place a hold on Microsoft Teams and Microsoft 365 groups

Microsoft Teams is built on Microsoft 365 groups. Therefore, placing them on hold in eDiscovery is similar. Keep the following things in mind when placing Microsoft 365 groups and Microsoft

Teams on hold:

- To place content located in Microsoft 365 groups and Microsoft Teams on hold, you need to specify the mailbox and SharePoint site that are associated with a group or team.
- Run the **Get-UnifiedGroup** cmdlet in Exchange Online to view properties for a Microsoft 365 group or Microsoft Team. This cmdlet is a good way to get the URL for the site that's associated with a Microsoft 365 group or a Microsoft Team. For example, the following command displays selected properties for a Microsoft 365 group named Senior Leadership Team:

  Console

  ```
  Get-UnifiedGroup "Senior Leadership Team" | FL
  DisplayName,Alias,PrimarySmtpAddress,SharePointSiteUrl
  DisplayName            : Senior Leadership Team
  Alias                  : seniorleadershipteam
  PrimarySmtpAddress     : seniorleadershipteam@contoso.onmicrosoft.com
  SharePointSiteUrl      :
  https://contoso.sharepoint.com/sites/seniorleadershipteam
  ```

  > **Note**
  >
  > To run the Get-UnifiedGroup cmdlet, you need to be assigned the View-Only Recipients role in Exchange Online or be a member of a role group that's assigned the View-Only Recipients role.

- When you search a user's mailbox, the search doesn't include any Microsoft 365 group or Microsoft Team that the user is a member of. Similarly, when you place a Microsoft 365 group or Microsoft Team hold, only the group mailbox and group site are placed on hold; the mailboxes and OneDrive sites of group members aren't placed on hold unless you explicitly add them to a case or place their data sources hold. Therefore, if you need to place a Microsoft 365 group or Microsoft Team on hold for a specific user, consider mapping the group site and group mailbox to the user. If the Microsoft 365 group or Microsoft Team isn't attributable to a single user, consider adding the source to a hold.
- To get a list of the members of a Microsoft 365 group or Microsoft Team, you can view the properties on the **Home** > **Groups page in the Microsoft 365 admin center**. Alternatively, you can run the following command in Exchange Online PowerShell:

  PowerShell

```
Get-UnifiedGroupLinks <group or team name> -LinkType Members | FL
DisplayName,PrimarySmtpAddress
```

> **Note**
>
> To run the **Get-UnifiedGroupLinks** cmdlet, you need to be assigned the View-Only Recipients role in Exchange Online or be a member of a role group that's assigned the View-Only Recipients role.

- Channel conversations that are part of a Microsoft Teams channel are stored in the mailbox that's associated with the Team. Similarly, files that team members share in a channel are stored on the team's SharePoint site. Therefore, you need to place the Microsoft Team mailbox and SharePoint site on hold to retain conversations and files in a channel.

- Alternatively, conversations that are part of the Chat list in Microsoft Teams are stored in the mailbox of the users who participate in the chat. Files that a user shares in Chat conversations are stored in the OneDrive site of the user who shares the file. Therefore, you need to place the individual user mailboxes and OneDrive sites on hold to retain conversations and files in the Chat list.

- Every Microsoft Team or team channel contains a Wiki for note-taking and collaboration. The Wiki content is automatically saved to a file with a .mht format. This file is stored in the Teams Wiki Data document library on the team's SharePoint site. You can place the content in the Wiki on hold by placing the team's SharePoint site on hold.

  > **Note**
  >
  > The capability to retain Wiki content for a Microsoft Team or team channel (when you place the team's SharePoint site on hold) was released on June 22, 2017. If a team site is on hold, the Wiki content is retained starting on that date. However, if a team site is on hold and the Wiki content was deleted before June 22, 2017, the Wiki content wasn't retained.

Related content

- eDiscovery hold limits

Note: The author created this article with assistance from AI. Learn more

Last updated on 06/23/2026

Use the hold report (preview) in eDiscovery

The hold report (preview) in eDiscovery lets users with *eDiscovery Administrator* or *eDiscovery Manager (preview)* permissions access a built-in report with information on all holds associated with eDiscovery cases in the Microsoft Purview portal. This report includes eDiscovery holds associated with eDiscovery cases. The hold report shows the data locations and reports on all hold policies within the tenant. It lists all locations that are part of those hold policies, whether enabled or disabled.

> **Important**
>
> The hold report is in public preview and provides a tenant-wide summary of hold policies for all eDiscovery cases. The report syncs policy and location status no more than once every seven days, so the information shown might not reflect recent changes to your hold policies. For the current status of a specific hold policy and its locations, go to the hold policy details page in the Microsoft Purview eDiscovery case. For more information, see Manage holds in eDiscovery.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Step 1: Verify your subscription and permissions

- Your organization must have an organization subscription that supports premium eDiscovery features to generate and access the hold report.
- Users accessing the hold report must be assigned the *eDiscovery Administrator* or *eDiscovery Manager (preview)* role. All eDiscovery Administrators and eDiscovery Managers in your organization can view the Hold Report. eDiscovery Managers can only view hold reports related to the cases they are a member of. For more information about eDiscovery permissions and the eDiscovery Administrator and eDiscovery Manager roles, see Assign eDiscovery permissions.

Step 2: View the hold report

1. Go to **eDiscovery** > **Dashboards** > **Hold report (preview)** to view your hold report.
2. Sort, filter, group by, and customize columns to customize your hold report.
3. Download the hold report to a .csv file for further analysis. The entire report is downloaded regardless of any filtering or grouping option you apply to the report. You can also use a PowerShell script to create a hold report for automated reporting.

The item count at the top of the hold report is the total count of locations within the hold policies in the tenant. If the same location is placed on hold under different policies, the same location appears multiple times in the report under the respective case and hold policy and contributes to multiple item counts.

When you search the report, only the **Case name** and **Location** columns are searched. For example, if you search for the keyword *Test*, the search only returns items that are associated with a case or location that contains the word Test.

Default hold report columns

- **Case Name**: Name of the case that the hold policy is associated with. If the case name changes and you update the hold policy associated with the case, the report immediately reflects the case name change for the associated row. When a full sync of the report occurs (currently every three to seven days), the case name also updates.
- **Hold name**: Name of the Hold Policy that contains the associated locations.
- **Last modified**: The date when the eDiscovery hold policy was last updated.
- **Location**: The Exchange mailbox or SharePoint site in the hold policy. If this field displays *None*, it indicates that the hold policy has no active locations within it. Any previous locations that were part of the policy are released from hold and removed from the policy.
- **Location error**: Displays the error message associated with the location within the hold policy. If this field displays *None*, it indicates that the associated location doesn't have errors, even if it's part of a *Partially Applied/Released* policy.
- **Location type**: Indicates if the location is a SharePoint site or Exchange mailbox.
- **Policy status**: The following policy status values are available:
  - *Applied successfully*: Indicates the location of the hold policy is turned on (enabled), a rule is enabled for the policy, and the policy is applied successfully. This status means the location is placed on hold.

- *Applying*: Indicates that the hold policy that the location is part of is set to be turned on (enabled) and is in progress to apply holds to the locations in it. This job is an asynchronous job, which means some locations might be placed on hold while others might still be in progress. Wait for the hold job to complete checking the final status.
- *Applied Partially*: – Indicates that the hold policy that the location is part of is set to be turned on (enabled) and has some locations where the hold might not be applied. Check the 'Location error' column. If the associated result is 'None' then the location is placed on hold. If there's an associated error, then the location isn't placed on hold and the error message is displayed here. Check the hold policy for more details to remediate and retry the hold.
- *Released successfully*: Indicates that the hold policy that the location is part of is set to be turned off (disabled) and has completed successfully. This status means the associated location is successfully released from hold.
- *Releasing*: Indicates that the hold policy that the location is part of is set to be turned off (disabled) and is in progress to release holds from the locations in it. This job is an asynchronous job, which means some locations might be removed from hold while others might still be in progress. Wait for the hold job to complete checking the final status.
- *Released partially*: Indicates that the hold policy that the location is part of is set to be turned off (disabled) and has some locations where the hold might not be released. Check the 'Location error' column. If the associated result is 'None' then the location is successfully released from hold. If there's an associated error, then the location isn't removed from hold and the error message is displayed here. Check the hold policy for more details to remediate and retry the action.
- *Policy failed*: Indicates that the policy is set up incorrectly. The policy doesn't have a rule or was set up with rule set as disabled. These policies aren't enforced even if the policy status is successful. To ensure these policies are enforced, create a rule and resync the policy. For more information, see New-CaseHoldRule.

> **Note**
>
> If you delete a user account while an eDiscovery hold is active on the mailbox, the mailbox becomes an inactive mailbox. The hold report continues to show the location and hold status. If the location shows an error after you delete the account, this error might indicate that the hold isn't yet in place because you delete the account before the system applies it, or that the mailbox transition is still processing. To confirm that the hold is in place before you delete the account, see **Identify Exchange mailbox hold types in eDiscovery**.

Other available hold report columns

- **Case type**: Indicates if the case is an eDiscovery or eDiscovery case with premium features.
- **Case status**: Shows the status of case. Values are *Active*, *Closing*, *Closed with error*, *Pending delete*, or *Closed*. If the case status changes and you update a hold policy associated with the case, the report immediately reflects the case status change for the associated row. When a full sync of the report occurs (currently every three to seven days), all case status changes are also updated.
- **Policy created by**: Shows the user that created the initial hold policy.
- **Policy last modified by**: Shows the user that last modified the hold policy.
- **Policy last modified**: Shows the last date that the policy was updated on. This date is the policy update date and doesn't necessarily match the date the location was placed on or released from hold.
- **Report last fetched**: Displays the date and time stamp of the latest fetch of policy status and location information by the report for the specific policy.
- **Policy created date**: Shows the first date that the policy was created on. This date is the policy creation date and doesn't necessarily match the date the location was placed on or released from hold.
- **Hold ID**: The unique ID of the hold policy.
- **Case ID**: The unique ID of the case associated with the hold policy.
- **Rule query**: Displays any content match query that is applied to the policy. If there's no content match query applied to the policy, this field is empty.
- **Rule ID**: The unique ID for the rule associated with the hold policy. This status is useful for sending information to Microsoft when reporting any hold errors. If there's no rule associated with the policy, *Rule not available* is displayed in this field. Policies where a rule isn't available aren't set up correctly and not enforced. Create rules and resync the policy. For more information, see New-CaseHoldRule.

ⓘ **Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 04/27/2026

Use a script to create a hold report in eDiscovery

This PowerShell script lets eDiscovery administrators and eDiscovery managers generate a report that contains information about all holds that are associated with eDiscovery cases in the Microsoft Purview portal. The report contains information such as the name of the case a hold is associated with, the content locations that are placed on hold, and whether the hold is query-based. If there are cases that don't have any holds, the script creates an additional report with a list of cases without holds.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

What the hold report includes

The case holds report that the script in this article creates contains the following information about each hold:

- The name of the hold and the name of the eDiscovery case that the hold is associated with.
- Whether the hold is associated with an eDiscovery case.
- Whether the eDiscovery case is active or closed.
- Whether the hold is enabled or disabled.
- The members of the eDiscovery case that the hold is associated with. Case members can view or manage a case, depending on assigned eDiscovery permissions.
- The time and date the case was created.
- If a case is closed, the person who closed it and the time and date it was closed.
- The Exchange mailboxes and SharePoint sites locations that are on hold.
- If the hold is query-based, the query syntax.
- The time and date the hold was created and the person who created it.
- The time and date the hold was last changed and the person who changed it.

Admin requirements and script information

- To generate a report on all eDiscovery cases in your organization, you need to be an eDiscovery Administrator in your organization. If you're an eDiscovery Manager, the report only includes information about the cases that you can access. For more information about eDiscovery permissions, see Assign eDiscovery permissions.
- The CaseHoldsReport.ps1 script has minimal error handling. Its primary purpose is to quickly create a report about the holds that are associated with the eDiscovery cases in your organization.

> **Important**
>
> The sample scripts provided in this article aren't supported under any Microsoft standard support program or service. The sample scripts are provided AS IS without warranty of any kind. Microsoft further disclaims all implied warranties including, without limitation, any implied warranties of merchantability or of fitness for a particular purpose. The entire risk arising out of the use or performance of the sample scripts and documentation remains with you. In no event shall Microsoft, its authors, or anyone else involved in the creation, production, or delivery of the scripts be liable for any damages whatsoever (including, without limitation, damages for loss of business profits, business interruption, loss of business information, or other pecuniary loss) arising out of the use of or inability to use the sample scripts or documentation, even if Microsoft has been advised of the possibility of such damages.

Step 1: Connect to Security & Compliance PowerShell

First, connect to Security & Compliance PowerShell for your organization. For step-by-step instructions, see Connect to Security & Compliance PowerShell.

Step 2: Run the script to report on holds associated with eDiscovery cases

After you connect to Security & Compliance PowerShell, create and run the script that collects information about the eDiscovery cases in your organization.

1. Save the following text to a Windows PowerShell script file by using a filename suffix of `.ps1`; for example, `CaseHoldsReport.ps1`.

   PowerShell

```
 #script begin
 " "
 write-host "********************************************"
 write-host "Security & Compliance Center   " -foregroundColor yellow -
backgroundcolor darkgreen
 write-host "eDiscovery cases - Holds report         " -foregroundColor yellow -
backgroundcolor darkgreen
 write-host "********************************************"
 " "

 #prompt users to specify a path to store the output files
 $time = get-date -Format dd-MM-yyyy_hh.mm
 $Path = Read-Host 'Enter a folder path to save the report to a .csv file
(filename is created automatically)'
 $outputpath = $Path + '\' + 'CaseHoldsReport' + ' ' + $time + '.csv'
 $noholdsfilepath = $Path + '\' + 'CaseswithNoHolds' + $time + '.csv'

 #add case details to the csv file
 function add-tocasereport {
     Param([string]$casename,
         [String]$casetype,
         [String]$casestatus,
         [datetime]$casecreatedtime,
         [string]$casemembers,
         [datetime]$caseClosedDateTime,
         [string]$caseclosedby,
         [string]$holdname,
         [String]$Holdenabled,
         [string]$holdcreatedby,
         [string]$holdlastmodifiedby,
         [string]$ExchangeLocation,
         [string]$sharePointlocation,
         [string]$ContentMatchQuery,
         [datetime]$holdcreatedtime,
         [datetime]$holdchangedtime,
         [string]$holdstatus,
         [string]$holderror
     )

     $addRow = New-Object PSObject
     Add-Member -InputObject $addRow -MemberType NoteProperty -Name "Case name" -
Value $casename
     Add-Member -InputObject $addRow -MemberType NoteProperty -Name "Case type" -
Value $casetype
     Add-Member -InputObject $addRow -MemberType NoteProperty -Name "Case status"
-Value $casestatus
     Add-Member -InputObject $addRow -MemberType NoteProperty -Name "Case
members" -Value $casemembers
     Add-Member -InputObject $addRow -MemberType NoteProperty -Name "Case created
time" -Value $casecreatedtime
     Add-Member -InputObject $addRow -MemberType NoteProperty -Name "Case closed
time" -Value $caseClosedDateTime
     Add-Member -InputObject $addRow -MemberType NoteProperty -Name "Case closed
by" -Value $caseclosedby
```

```
     Add-Member -InputObject $addRow -MemberType NoteProperty -Name "Hold name" -
Value $holdname
     Add-Member -InputObject $addRow -MemberType NoteProperty -Name "Hold
enabled" -Value $Holdenabled
     Add-Member -InputObject $addRow -MemberType NoteProperty -Name "Hold created
by" -Value $holdcreatedby
     Add-Member -InputObject $addRow -MemberType NoteProperty -Name "Hold last
changed by" -Value $holdlastmodifiedby
     Add-Member -InputObject $addRow -MemberType NoteProperty -Name "Exchange
locations" -Value  $ExchangeLocation
     Add-Member -InputObject $addRow -MemberType NoteProperty -Name "SharePoint
locations" -Value $sharePointlocation
     Add-Member -InputObject $addRow -MemberType NoteProperty -Name "Hold query"
-Value $ContentMatchQuery
     Add-Member -InputObject $addRow -MemberType NoteProperty -Name "Hold created
time (UTC)" -Value $holdcreatedtime
     Add-Member -InputObject $addRow -MemberType NoteProperty -Name "Hold changed
time (UTC)" -Value $holdchangedtime
     Add-Member -InputObject $addrow -MemberType NoteProperty -Name "Hold Status"
-Value $holdstatus
     Add-Member -InputObject $addrow -MemberType NoteProperty -Name "Hold Error"
-Value $holderror

     $allholdreport = $addRow | Select-Object "Case name", "Case type", "Case
status", "Hold name", "Hold enabled", "Case members", "Case created time", "Case
closed time", "Case closed by", "Exchange locations", "SharePoint locations",
"Hold query", "Hold created by", "Hold created time (UTC)", "Hold last changed
by", "Hold changed time (UTC)", "Hold Status", "Hold Error"
     $allholdreport | export-csv -path $outputPath -notypeinfo -append -Encoding
ascii
 }

 #get information on the cases and pass values to the case report function
 " "
 write-host "Gathering a list of eDiscovery cases and holds..."
 " "
 $edc = Get-ComplianceCase -ErrorAction SilentlyContinue
 foreach ($cc in $edc) {
     write-host "Working on case :" $cc.name
     if ($cc.status -eq 'Closed') {
         $cmembers = ((Get-ComplianceCaseMember -Case $cc.name).windowsLiveID) -
join ';'
         add-tocasereport -casename $cc.name -casetype $cc.casetype -casestatus
$cc.Status -caseclosedby $cc.closedby -caseClosedDateTime $cc.ClosedDateTime -
casemembers $cmembers
     }
     else {
         $cmembers = ((Get-ComplianceCaseMember -Case $cc.name).windowsLiveID) -
join ';'
         $policies = Get-CaseHoldPolicy -Case $cc.Name | % { Get-CaseHoldPolicy
$_.Name -Case $_.CaseId -DistributionDetail }
         if ($policies -ne $NULL) {
             foreach ($policy in $policies) {
                 $rule = Get-CaseHoldRule -Policy $policy.name
                 add-tocasereport -casename $cc.name -casetype $cc.casetype -
```

```
casemembers $cmembers -casestatus $cc.Status -casecreatedtime $cc.CreatedDateTime
-holdname $policy.name -holdenabled $policy.enabled -holdcreatedby
$policy.CreatedBy -holdlastmodifiedby $policy.LastModifiedBy -ExchangeLocation
(($policy.exchangelocation.name) -join ';') -SharePointLocation
(($policy.sharePointlocation.name) -join ';') -ContentMatchQuery
$rule.ContentMatchQuery -holdcreatedtime $policy.WhenCreatedUTC -holdchangedtime
$policy.WhenChangedUTC -holdstatus $policy.DistributionStatus -holderror
$policy.DistributionResults
            }
        }
        else {
            Write-Host "No hold policies found in case:" $cc.name -
foregroundColor 'Yellow'
            " "
            [string]$cc.name | out-file -filepath $noholdsfilepath -append
        }
    }
 }

 #get information on the cases and pass values to the case report function
 " "
 write-host "Gathering a list of eDiscovery cases and holds..."
 " "
 $edc = Get-ComplianceCase -CaseType Advanced -ErrorAction SilentlyContinue
 foreach ($cc in $edc) {
     write-host "Working on case :" $cc.name
     if ($cc.status -eq 'Closed') {
         $cmembers = ((Get-ComplianceCaseMember -Case $cc.name).windowsLiveID) -
join ';'
         add-tocasereport -casename $cc.name -casestatus $cc.Status -casetype
$cc.casetype -caseclosedby $cc.closedby -caseClosedDateTime $cc.ClosedDateTime -
casemembers $cmembers
     }
     else {
         $cmembers = ((Get-ComplianceCaseMember -Case $cc.name).windowsLiveID) -
join ';'
         $policies = Get-CaseHoldPolicy -Case $cc.Name | % { Get-CaseHoldPolicy
$_.Name -Case $_.CaseId -DistributionDetail }
         if ($policies -ne $NULL) {
             foreach ($policy in $policies) {
                 $rule = Get-CaseHoldRule -Policy $policy.name
                 add-tocasereport -casename $cc.name -casetype $cc.casetype -
casemembers $cmembers -casestatus $cc.Status -casecreatedtime $cc.CreatedDateTime
-holdname $policy.name -holdenabled $policy.enabled -holdcreatedby
$policy.CreatedBy -holdlastmodifiedby $policy.LastModifiedBy -ExchangeLocation
(($policy.exchangelocation.name) -join ';') -SharePointLocation
(($policy.sharePointlocation.name) -join ';') -ContentMatchQuery
$rule.ContentMatchQuery -holdcreatedtime $policy.WhenCreatedUTC -holdchangedtime
$policy.WhenChangedUTC -holdstatus $policy.DistributionStatus -holderror
$policy.DistributionResults

             }
         }
         else {
             write-host "No hold policies found in case:" $cc.name -
```

```
foregroundColor 'Yellow'
            " "
            [string]$cc.name | out-file -filepath $noholdsfilepath -append
        }
    }
}

" "
Write-host "Script complete! Report files saved to this folder: '$Path'"
" "
#script end
```

2. In the Windows PowerShell session that opened in Step 1, go to the folder where you saved the script.

3. Run the script; for example:

   PowerShell

   ```
   .\CaseHoldsReport.ps1
   ```

   The script prompts for a target folder to save the report to.

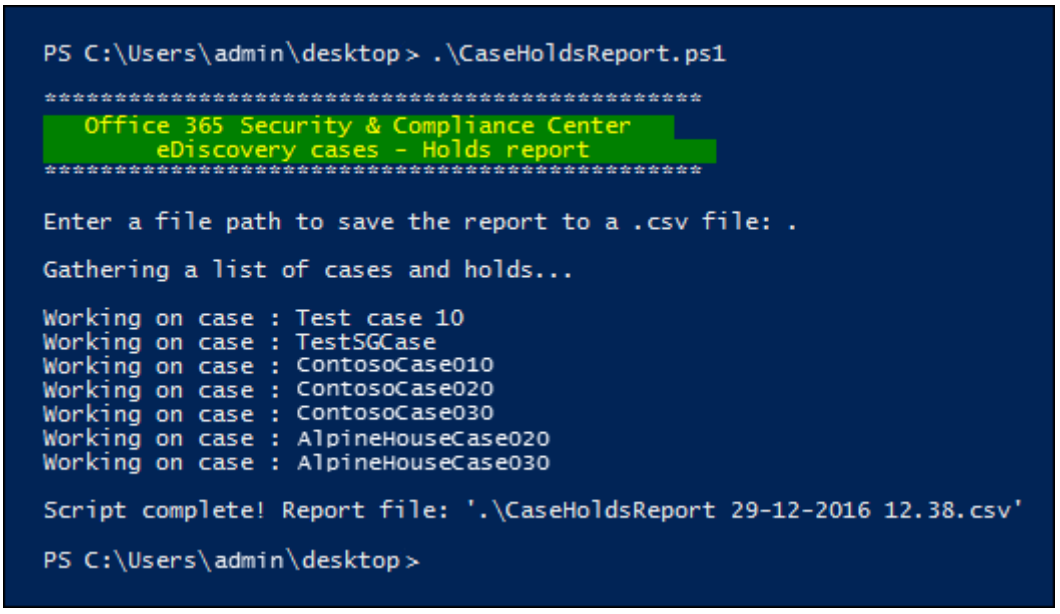
4. Type the full path name of the folder to save the report to, and then press **Enter**.

   > **Tip**
   >
   > To save the report in the same folder that the script is located in, type a period (".") when prompted for a target folder. To save the report in a subfolder in the folder where the script is located, just type the name of the subfolder.

   The script starts to collect information about all the eDiscovery cases in your organization. Don't access the report file while the script is running. After the script completes, a confirmation message is displayed in the Windows PowerShell session. After this message is displayed, you can access the report in the folder that you specified when prompted for a target folder. The file name for the report is `CaseHoldsReport<DateTimeStamp>.csv`.

   Additionally, the script creates a report with a list of cases that don't have any holds. The file name for this report is `CaseswithNoHolds<DateTimeStamp>.csv`.

   Here's an example of running the `CaseHoldsReport.ps1` script.



ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/25/2026

Identify Exchange mailbox hold types in eDiscovery

Microsoft Purview offers several ways that your organization can prevent mailbox content from being permanently deleted. This protection allows your organization to retain content to meet compliance regulations or during legal and other types of investigations. This article explains how to identify the different types of holds placed on an Exchange Online mailbox and how to use PowerShell cmdlets to get the GUID for each hold.

> **Tip**
>
> To determine *whether* a mailbox or site is on hold (for example, because a content purge or site deletion isn't working), start with the diagnostic checklist in **Check for holds on a mailbox or site**. This article explains the hold types themselves in more depth.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about **Microsoft Security Copilot in Microsoft Purview**.

Overview of Exchange mailbox hold types

Here's a list of the retention features (also called *holds*) in Microsoft Purview and Microsoft 365:

- **Litigation Hold**: Holds that you apply to user mailboxes in Exchange Online.
- **eDiscovery hold**: Holds that you associate with an eDiscovery case in the Microsoft Purview portal. You can place a hold on the Exchange mailboxes and OneDrive accounts of people you're investigating in the case. You can also place a hold on the mailboxes and sites that are associated with Microsoft Teams, Microsoft 365 groups, and Viva Engage Groups.
- **Microsoft Purview retention policies**: Can be configured to retain (or retain and then delete) content in user mailboxes in Exchange Online and in the corresponding mailbox for Microsoft 365 Groups and Microsoft Teams. You can also create a retention policy to retain Skype for Business Conversations, which are stored in user mailboxes.

There are two types of Microsoft Purview retention policies that you can assign to mailboxes.

- **Specific location retention policies**: These policies are assigned to the content locations of specific users. Use the **Get-Mailbox** cmdlet in Exchange Online PowerShell to get information about retention policies assigned to specific mailboxes. For more information about this type of retention policy, see the section A policy with specific inclusions or exclusions from the retention policy documentation.
- **Organization-wide retention policies**: These policies are assigned to all content locations in your organization. Use the **Get-OrganizationConfig** cmdlet in Exchange Online PowerShell to get information about organization-wide retention policies. For more information about this type of retention policy, see the section A policy that applies to entire locations from the retention policy documentation.

- **Microsoft Purview retention labels**: If a user applies a Microsoft Purview retention label (one that's configured to retain content or retain and then delete content) to *any* folder or item in their mailbox, a hold is placed on the mailbox as if you placed the mailbox on Litigation Hold or assigned it to a Microsoft Purview retention policy. For more information, see the Identifying mailboxes on hold because a retention label was applied to a folder or item section in this article.

To manage mailboxes on hold, you might need to identify the type of hold that's placed on a mailbox so that you can perform tasks such as changing the hold duration, temporarily or permanently removing the hold, or excluding a mailbox from a Microsoft Purview retention policy. In these cases, the first step is to identify the type of hold placed on the mailbox. Because multiple holds (and different types of holds) can be placed on a single mailbox, you need to identify all holds placed on a mailbox if you want to remove or change a hold.

Step 1: Get the GUID for holds placed on a mailbox

To connect to Exchange Online PowerShell, see Connect to Exchange Online PowerShell. You must connect before you run the cmdlets in this step.

Run the following two cmdlets in Exchange Online PowerShell to get the GUID of the holds that are placed on a mailbox. After you get a GUID, use it to identify the specific hold in Step 2. A Litigation Hold doesn't use a GUID. Litigation Holds are either enabled or disabled for a mailbox.

- **Get-Mailbox**: Use this cmdlet to check if Litigation Hold is enabled for a mailbox and to get the GUIDs for eDiscovery holds and Microsoft Purview retention policies that are assigned

to a mailbox. The output of this cmdlet also shows if a mailbox is explicitly excluded from an organization-wide retention policy.

- **Get-OrganizationConfig**: Use this cmdlet to get the GUIDs for organization-wide retention policies.

Use Get-Mailbox to view holds on a mailbox

Run the following command to get information about the holds and Microsoft Purview retention policies applied to a mailbox.

PowerShell

```
Get-Mailbox <username> | FL LitigationHoldEnabled,InPlaceHolds
```

> **Tip**
>
> If there are too many values in the InPlaceHolds property and not all of them show up, run the `Get-Mailbox <username> | Select-Object -ExpandProperty InPlaceHolds` command to display each GUID on a separate line.

The following table describes how to identify different types of holds based on the values in the *InPlaceHolds* property when you run the **Get-Mailbox** cmdlet.

Expand table

Hold type	Example value	How to identify the hold
Litigation Hold	`True`	Litigation Hold is enabled for a mailbox when the *LitigationHoldEnabled* property is set to `True`.
eDiscovery hold	`UniH7d895d48-7e23-4a8d-8346-533c3beac15d`	The *InPlaceHolds property* contains the GUID of any hold associated with an eDiscovery case in the Microsoft Purview portal. You can tell this is an eDiscovery hold because the GUID starts with the `UniH` prefix (which denotes a Unified Hold).
Microsoft Purview retention policy applied to the mailbox	`mbxcdbbb86ce60342489bff371876e7f224:1` or `skp127d7cf1076947929bf136b7a2a8c36f:3`	The *InPlaceHolds* property contains GUIDs of any specific location retention policy that's applied to the mailbox. You can identify retention policies because the GUID starts with the `mbx` or the `skp` prefix. The `skp`

Hold type	Example value	How to identify the hold
		prefix indicates that the retention policy is applied to Skype for Business conversations in the user's mailbox.
Excluded from an organization-wide Microsoft Purview retention policy	`-mbxe9b52bf7ab3b46a286308ecb29624696`	If a mailbox is excluded from an organization-wide Microsoft Purview retention policy, the GUID for the retention policy that the mailbox is excluded from shows up in the *InPlaceHolds* property and uses the `-mbx` prefix.

Use Get-OrganizationConfig to identify organization-wide holds

If the *InPlaceHolds* property is empty when you run the **Get-Mailbox** cmdlet, one or more organization-wide Microsoft Purview retention policies might still apply to the mailbox. Run the following command in Exchange Online PowerShell to get a list of GUIDs for organization-wide Microsoft Purview retention policies.

PowerShell

```
Get-OrganizationConfig | FL InPlaceHolds
```

> **Tip**
>
> If the `InPlaceHolds` property contains too many values and not all of them display, run the `Get-OrganizationConfig | Select-Object -ExpandProperty InPlaceHolds` command to display each GUID on a separate line.

The following table describes the different types of organization-wide holds and how to identify each type based on the GUIDs contained in the *InPlaceHolds* property when you run the **Get-OrganizationConfig** cmdlet.

Expand table

Hold type	Example value	Description
Microsoft Purview retention policies applied to Exchange mailboxes, Exchange	`mbx7cfb30345d454ac0a989ab3041051209:2`	Organization-wide retention policies applied to Exchange mailboxes, Exchange public folders, and 1xN chats in Microsoft Teams are identified by GUIDs that start

Hold type	Example value	Description
public folders, and Teams chats		with the `mbx` prefix. Note 1xN chats are stored in the mailbox of the individual chat participants.
Microsoft Purview retention policy applied to Microsoft 365 Groups and Teams channel messages	`grp1a0a132ee8944501a4bb6a452ec31171:3`	Organization-wide retention policies applied to Microsoft 365 groups and channel messages in Microsoft Teams are identified by GUIDs that start with the `grp` prefix. Note channel messages are stored in the group mailbox that is associated with a Microsoft Team.

For more information about retention policies applied to Microsoft Teams, see Learn about retention policies for Microsoft Teams.

Understand the format of the InPlaceHolds value for retention policies

In addition to the prefix (`mbx`, `skp`, or `grp`) that identifies an item in the `InPlaceHolds` property as a Microsoft Purview retention policy, the value also contains a suffix that identifies the type of retention action that's configured for the policy. For example, the action suffix is highlighted in bold type in the following examples:

`skp127d7cf1076947929bf136b7a2a8c36f`:1

`mbx7cfb30345d454ac0a989ab3041051209`:2

`grp1a0a132ee8944501a4bb6a452ec31171`:3

The following table defines the three possible retention actions:

Expand table

Value	Description
1	The retention policy is configured to delete items or the retention policy is used to publish retention labels. If the policy is publishing retention labels, check the retention configuration for each published label.
2	The retention policy is configured to hold items. The policy doesn't delete items after the retention period expires.
3	The retention policy is configured to hold items and then delete them after the retention period expires.

> **Note**
>
> Because retention label policies publish or autoapply labels that apply item-level actions, they always show an action value of 1 within the `InPlaceHolds` property of the mailbox.
>
> To identify whether a hold is applied to a folder or item within the mailbox, see **Identifying mailboxes on hold because a retention label has been applied to a folder or item**.

For more information about retention actions, see the Retaining content for a specific period of time section.

Step 2: Use the GUID to identify the hold

After you get the GUID for a hold that's applied to a mailbox, use that GUID to identify the hold. The following sections show how to identify the name of the hold and other information by using the hold GUID.

eDiscovery holds

Run the following commands in Security & Compliance PowerShell to find an eDiscovery hold that's applied to the mailbox. Use the GUID (without the UniH prefix) for the eDiscovery hold that you found in Step 1.

To connect to Security & Compliance PowerShell, see Connect to Security & Compliance PowerShell.

The first command creates a variable that holds information about the eDiscovery hold. Use this variable in the other commands. The second command shows the name of the eDiscovery case that the hold is associated with. The third command shows the name of the hold and lists the mailboxes that the hold applies to.

PowerShell

```
$CaseHold = Get-CaseHoldPolicy <hold GUID without prefix>
```

PowerShell

```
Get-ComplianceCase $CaseHold.CaseId | FL Name
```

PowerShell

```
$CaseHold | FL Name,ExchangeLocation
```

Identify Microsoft Purview retention policies by GUID

Connect to Security & Compliance PowerShell and run the following command to find the Microsoft Purview retention policy (organization-wide or specific location) that's applied to the mailbox. Use the GUID (without the mbx, skp, or grp prefix or the action suffix) that you found in Step 1.

PowerShell

```
Get-RetentionCompliancePolicy <hold GUID without prefix or suffix> -DistributionDetail
| FL Name,*Location
```

New retention policies that apply to the following locations don't stamp directly on organization configuration or Exchange Online objects:

- Enterprise AI apps
- Microsoft Copilot experiences
- Teams chats
- Teams private channel messages
- Viva Engage community messages
- Viva Engage user messages
- Other AI apps

To determine whether these retention compliance policies apply to a mailbox during an investigation, use Policy Lookup or Retention cmdlets for newer locations.

PowerShell

```
Get-AppRetentionCompliancePolicy <hold GUID without prefix or suffix> -
DistributionDetail  | FL Name, Applications, *Location
```

Identify mailboxes on hold because a retention label is applied to a folder or item

When a user applies a retention label that's configured to *retain* or *retain and then delete* content to any folder or item in their mailbox, the *ComplianceTagHoldApplied* mailbox property is set to **True**. When this property is **True**, treat the mailbox as if it's on hold, such as when you assign it to

a Microsoft Purview retention policy or place it on Litigation Hold, but with some caveats. When the *ComplianceTagHoldApplied* property is **True**, the following things occur:

- If you delete the mailbox or the user's Microsoft 365 account, the mailbox becomes an inactive mailbox.
- You can't disable the mailbox (either the primary mailbox or the archive mailbox, if it's enabled).
- Items that you delete from the mailbox follow one of two paths depending on whether they're labeled or not:
  - **Unlabeled items** follow the same path deleted items take when no holds apply to the mailbox. The time that it takes for these items to be permanently deleted is determined by the deleted item retention configuration and whether single item recovery is enabled for the mailbox or not.
  - **Labeled items** are retained within the recoverable items folder in the same way they would be if a Microsoft Purview retention policy applied, but at the individual item level. If multiple items have different labels that are configured to *retain* or *retain and then delete* content at different intervals, each item is retained based on the configuration of the applied label.
- Other holds, such as Microsoft Purview retention policies, eDiscovery holds, or litigation hold can extend how long labeled items are retained based on the principles of retention.

To view the value of the *ComplianceTagHoldApplied* property for a single mailbox, run the following command in Exchange Online PowerShell:

PowerShell

```
Get-Mailbox <username> | FL ComplianceTagHoldApplied
```

For more information about retention labels, see retention labels.

A *True* value for **ComplianceTagHoldApplied** doesn't mean the mailbox currently has items with a retention label. When you apply a retention label to a mailbox, **ComplianceTagHoldApplied** changes to *True*. After it changes to *True*, it doesn't automatically change back to *False*, even when the mailbox no longer has labeled items.

To change the *ComplianceTagHoldApplied* value to *False*, run the following command in Exchange Online PowerShell:

PowerShell

```
Set-Mailbox -Identity <username> -RemoveComplianceTagHoldApplied -ProvideConsent
```

For more information, see Set-Mailbox. For the full procedure to delete items from the Recoverable Items folder, see Delete items in the Recoverable Items folder.

Manage mailboxes on delay hold

When you remove any type of hold from a mailbox, Exchange Online applies a *delay hold*. This hold delays the actual removal of the hold for 30 days to prevent data from being permanently deleted (purged) from the mailbox. This delay gives admins an opportunity to search for or recover mailbox items that are purged after a hold is removed. The Managed Folder Assistant places a delay hold on a mailbox the next time it processes the mailbox and detects that a hold was removed. Specifically, the Managed Folder Assistant applies a delay hold to a mailbox when it sets one of the following mailbox properties to **True**:

- **DelayHoldApplied**: This property applies to email-related content (generated by people using Outlook and Outlook on the web) that's stored in a user's mailbox.
- **DelayReleaseHoldApplied**: This property applies to cloud-based content (generated by non-Outlook apps such as Microsoft Teams, Microsoft Forms, and Microsoft Viva Engage) that's stored in a user's mailbox. Cloud data generated by a Microsoft app is typically stored in a hidden folder in a user's mailbox.

When the Managed Folder Assistant places a delay hold on the mailbox (by setting either of the previous properties to **True**), the mailbox is still considered to be on hold for an unlimited hold duration, as if the mailbox was on Litigation Hold. After 30 days, the delay hold expires, and Microsoft 365 automatically attempts to remove the delay hold (by setting the *DelayHoldApplied* or *DelayReleaseHoldApplied* property to **False**) so that the hold is removed. After either of these properties are set to **False**, the corresponding items that are marked for removal are purged the next time the Managed Folder Assistant processes the mailbox.

> ⓘ **Note**
>
> If the user account for the mailbox is disabled, the Managed Folder Assistant doesn't process the mailbox and the delay hold remains after the 30 days expire. For more information, see Delay hold considerations.

To view the values for the DelayHoldApplied and DelayReleaseHoldApplied properties for a mailbox, run the following command in Exchange Online PowerShell.

PowerShell

```
Get-Mailbox <username> | FL *HoldApplied*
```

To remove the delay hold before it expires, run one (or both) of the following commands in Exchange Online PowerShell, depending on which property you want to change:

PowerShell

```
Set-Mailbox <username> -RemoveDelayHoldApplied
```

Or

PowerShell

```
Set-Mailbox <username> -RemoveDelayReleaseHoldApplied
```

You must be assigned the Legal Hold role in Exchange Online to use the *RemoveDelayHoldApplied* or *RemoveDelayReleaseHoldApplied* parameters.

To remove the delay hold on an inactive mailbox, run one of the following commands in Exchange Online PowerShell:

PowerShell

```
Set-Mailbox <DN or Exchange GUID> -InactiveMailbox -RemoveDelayHoldApplied
```

Or

PowerShell

```
Set-Mailbox <DN or Exchange GUID> -InactiveMailbox -RemoveDelayReleaseHoldApplied
```

> **Tip**
>
> The best way to specify an inactive mailbox in the previous command is to use its Distinguished Name or Exchange GUID value. Using one of these values helps prevent accidentally specifying the wrong mailbox.

For more information about using these parameters for managing delay holds, see Set-Mailbox.

Delay hold considerations

Keep the following points in mind when managing a mailbox on delay hold:

- If you set either the *DelayHoldApplied* or *DelayReleaseHoldApplied* property to **True** and you delete a mailbox (or the corresponding user account), the mailbox becomes an inactive mailbox. A mailbox is considered to be on hold if either property is set to **True**. Deleting a mailbox on hold results in an inactive mailbox. To delete a mailbox and not make it an inactive mailbox, set both properties to **False**.
- A mailbox is on hold for an unlimited hold duration if either the *DelayHoldApplied* or *DelayReleaseHoldApplied* property is set to **True**. However, that condition doesn't mean that *all* content in the mailbox is preserved. The value of each property determines what content is preserved. For example, let's say both properties are set to **True** because you remove holds from the mailbox. Then, you remove only the delay hold that applies to non-Outlook cloud data by using the *RemoveDelayReleaseHoldApplied* parameter. The next time the Managed Folder Assistant processes the mailbox, it purges the non-Outlook items marked for removal. It doesn't purge any Outlook items marked for removal because the *DelayHoldApplied* property is still set to **True**. The opposite condition is also true: if *DelayHoldApplied* is set to **False** and *DelayReleaseHoldApplied* is set to **True**, then only Outlook items marked for removal are purged.

How to confirm that an organization-wide retention policy is applied to a mailbox

When you apply or remove an organization-wide retention policy to a mailbox, exporting the mailbox diagnostics logs helps you confirm that Exchange Online applied or removed the retention policy to the mailbox. To view this information, you first need to validate a few things by using Exchange Online PowerShell.

Obtain the GUIDs for any retention policies explicitly applied to a mailbox

Run the following command to list the retention policy GUIDs that are explicitly applied to the mailbox.

PowerShell

```
Get-Mailbox <username> | Select-Object -ExpandProperty InPlaceHolds
```

Obtain the GUIDs for any organization-wide retention policies applied to mailboxes

Run the following command to list the GUIDs for organization-wide retention policies.

PowerShell

```
Get-OrganizationConfig | Select-Object -ExpandProperty InPlaceHolds
```

Get the Mailbox Diagnostics for HoldTracking

The Hold Tracking Mailbox Diagnostics logs maintain a history of the holds applied to a user mailbox. If the object is inactive, include `-IncludeInactiveMailboxes` parameter.

PowerShell

```
$ht = Export-MailboxDiagnosticLogs <username> -ComponentName HoldTracking
$ht.MailboxLog | Convertfrom-Json
```

Get the Mailbox Diagnostics for Substrate HoldTracking

The Substrate Hold Tracking Mailbox Diagnostics logs maintain a history of the holds applied to a user mailbox. If the object is inactive, include `-IncludeInactiveMailboxes` parameter.

PowerShell

```
$sht = Export-MailboxDiagnosticLogs <username> -ComponentName SubstrateHoldTracking
$sht.MailboxLog | Convertfrom-Json
```

Review the results of the Mailbox Diagnostics logs

After you run `Export-MailboxDiagnosticLogs` for HoldTracking or SubstrateHoldTracking, the resulting data might look something like this example:

> **ed** `: 0001-01-01T00:00:00.0000000` **hid** `: mbx7cfb30345d454ac0a989ab3041051209:1` **ht** `: 4`
> **lsd** `: 2020-03-23T18:24:37.1884606Z` **osd** `: 2020-03-23T18:24:37.1884606Z`

Use the following table to help you understand each of the values listed in the diagnostics log.

Expand table

Value	Description
ed	Indicates the end date, which is the date the retention policy was disabled. MinValue means the policy is still assigned to the mailbox.
hid	Indicates the GUID for the retention policy. This value correlates to the GUIDs that you collected for the explicit or organization-wide retention policies assigned to the mailbox.
ht	Indicates the hold type. Values are *0* for LitigationHold, *1* for InPlaceHold, *2* for ComplianceTagHold, *3* for DelayReleaseHold, *4* for OrganizationRetention, *5* for CompliancePolicy, *6* for SubstrateAppPolicy, and *7* for SharepointPolicy.
lsd	Indicates the last start date, which is the date the retention policy was assigned to the mailbox.
osd	Indicates the original start date, which is the date that Exchange first recorded information about the retention policy.

When you remove a retention policy from a mailbox, Exchange places a temporary delay hold on the user to prevent purging content. You can disable a delay hold by running the `Set-Mailbox -RemoveDelayHoldApplied` command.

Next steps

After you identify the holds that are applied to a mailbox, you can perform tasks such as changing the duration of the hold, temporarily or permanently removing the hold, or excluding an inactive mailbox from a Microsoft Purview retention policy. For more information about performing tasks related to holds, see one of the following articles:

- Run the Set-RetentionCompliancePolicy -Identity <Policy Name> -AddExchangeLocationException <user mailbox> command in Security & Compliance PowerShell to exclude a mailbox from an organization-wide Microsoft Purview retention policy. You can only use this command for retention policies where the value for the *ExchangeLocation* property equals `All`.
- Run the Set-AppRetentionCompliancePolicy -Identity <Policy Name> -AddExchangeLocationException <user mailbox> command in Security & Compliance PowerShell to exclude a mailbox from a Teams private channel messages, Viva Engage user messages, Viva Engage community messages, Teams chats, Microsoft Copilot experiences, Enterprise AI apps, and other AI apps. You can only use this command for retention policies where the value for the *ExchangeLocation* property equals `All`.
- Run the Set-Mailbox <\DN or Exchange GUID> -RemoveDelayHoldApplied if needed to remove DelayHoldApplied if that's True. If the object is inactive, include the `-InactiveMailbox` parameter.

- Run the Set-Mailbox <\DN or Exchange GUID> -RemoveDelayReleaseHoldApplied if needed to remove DelayReleaseHoldApplied if that's True. If the object is inactive, include the `-InactiveMailbox` parameter.
- Run the Set-MailUser <\DN or Exchange GUID> -RemoveDelayHoldApplied if needed to remove DelayReleaseHoldApplied if that's True.
- Run the Set-MailUser <\DN or Exchange GUID> -RemoveDelayReleaseHoldApplied if needed to remove DelayReleaseHoldApplied if that's True.
- Run the Set-Mailbox <\DN or Exchange GUID> -LitigationHoldEnabled $False if needed to remove LitigationHold. If the object is inactive, include the `-InactiveMailbox` parameter.
- Run the Set-MailUser <\DN or Exchange GUID> -RemoveLitigationHoldEnabled if needed to remove LitigationHold.
- Run the Set-Mailbox <\DN or Exchange GUID> -InactiveMailbox -RemoveComplianceTagHoldApplied if needed to remove ComplianceTagHoldApplied from inactive mailbox. If the object is active, include the `-ProvideConsent` parameter Set-Mailbox <\DN or Exchange GUID> -RemoveComplianceTagHoldApplied -ProvideConsent
- Run the Set-MailUser <\DN or Exchange GUID> -RemoveComplianceTagHoldApplied if needed to remove ComplianceTagHoldApplied from inactive mailuser.
- Run the Set-Mailbox <\DN or Exchange GUID> -ExcludeFromAllHolds if needed to remove holds from inactive mailbox. Check the Set-Mailbox | ExcludeFromAllHolds to understand which holds this parameter can remove and which it cannot.
- Run the Set-MailUser <\DN or Exchange GUID> -ExcludeFromAllHolds if needed to remove holds from inactive mailuser. Check the Set-MailUser | ExcludeFromAllHolds to understand which holds this parameter can remove and which it cannot.
- Change the hold duration for an inactive mailbox
- Delete an inactive mailbox
- Delete items in the Recoverable Items folder of cloud-based mailboxes on hold

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/23/2026

Delete items in the Recoverable Items folder for mailboxes on hold in eDiscovery

The **Recoverable Items** folder for an Exchange Online mailbox protects against accidental or malicious deletions. Exchange Online also uses this folder to store items that compliance features retain and access, such as holds and eDiscovery searches. However, in some situations, organizations might need to delete data unintentionally retained in the Recoverable Items folder.

For example, a user might unknowingly send or forward an email message that contains sensitive information or information that might have serious business consequences. Even if the user permanently deletes the message, it might be retained indefinitely because a legal hold was placed on the mailbox. This scenario is known as *data spillage* because data was unintentionally *spilled* into Microsoft 365. In these situations, you can delete items in a user's Recoverable Items folder for an Exchange Online mailbox, even if you placed the mailbox on hold by using one of the different hold features in Microsoft 365. These types of holds include Litigation Holds, eDiscovery holds, and retention policies created in Microsoft Purview.

The following procedure explains how to delete items from the Recoverable Items folder for cloud-based Exchange Online mailboxes that are on hold. This procedure involves disabling access to the mailbox and disabling single item recovery, disabling the Managed Folder Assistant from processing the mailbox, temporarily removing the hold, deleting items from the Recoverable Items folder, and then reverting the mailbox to its previous configuration.

Here's a summary of the process:

- Step 1: Collect information about the mailbox
- Step 2: Prepare the mailbox
- Step 3: Remove all holds from the mailbox
- Step 4: Remove the delay hold from the mailbox
- Step 5: Delete items in the Recoverable Items folder
- Step 6: Revert the mailbox to its previous state

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about **Microsoft Security Copilot in Microsoft Purview**.

Before you delete items

> **Important**
>
> Before you start, confirm that no hold source still applies to the mailbox. If any hold (eDiscovery hold, Litigation Hold, Microsoft Purview retention policy, retention label hold, or delay hold) is in effect, the HardDelete purge in Step 5 silently leaves items in the Recoverable Items folder by design. For the full diagnostic checklist and the `InPlaceHolds` prefix decoder, see Check for holds on a mailbox or site.

> **Caution**
>
> Following the procedures outlined in this article results in data being permanently deleted (purged) from an Exchange Online mailbox. That means messages that you delete from the Recoverable Items folder can't be recovered and won't be available for legal discovery or other compliance purposes. If you want to delete messages from a mailbox that's placed on hold as part of a Litigation Hold, eDiscovery hold, or retention policy created in the Microsoft Purview portal, check with your records management or legal departments before removing the hold. Your organization might have a policy that defines whether a mailbox on hold or a data spillage activity takes priority.

- To create and run a search, you need to be a member of the eDiscovery Manager role group or be assigned the *Compliance Search* management role. To delete messages, you need to be a member of the *Organization Management* role group or be assigned the Search And Purge management role. For information about adding users to a role group, see Assign eDiscovery permissions.
- If you assign a mailbox to an organization-wide retention policy, you need to exclude the mailbox from the policy before you can delete items from the Recoverable Items folder. It might take up to 24 hours to synchronize the policy change and remove the mailbox from the policy. For more information, see "Remove organization-wide retention policies" in the Remove all holds from the mailbox section in this article.
- You can't perform this procedure for a mailbox that has retention settings assigned with a retention policy locked by using Preservation Lock. This lock prevents you from removing or excluding the mailbox from the policy and from disabling the Managed Folder Assistant on the mailbox. For more information about locking policies for retention, see Use Preservation Lock to restrict changes to retention policies and retention label policies.

- The procedure described in this article isn't supported for inactive mailboxes. That's because you can't reapply a hold (or retention policy) to an inactive mailbox after you remove it. When you remove a hold from an inactive mailbox, the mailbox changes to a normal soft-deleted mailbox and is permanently deleted from your organization after it's processed by the Managed Folder Assistant.
- If a mailbox isn't on hold (or doesn't have single item recovery enabled), you can delete the items from the Recoverable Items folder.

Step 1: Collect information about the mailbox

In this step, collect selected properties from the target mailbox that affect this procedure. Write down these settings or save them to a text file because you change some of these properties and then revert back to the original values in Step 6, after you delete items from the Recoverable Items folder. Here's a list of the mailbox properties you need to collect.

- *SingleItemRecoveryEnabled* and *RetainDeletedItemsFor*. If necessary, you disable single recovery and increase the deleted items retention period in Step 3.
- *LitigationHoldEnabled* and *InPlaceHolds*. You need to identify all the holds placed on the mailbox so that you can temporarily remove them in Step 3. For tips about how to identify the type of hold that might be placed on a mailbox, see Identify hold types by InPlaceHolds property values.

Additionally, you need to get the mailbox client access settings so you can temporarily disable them so the owner (or other users) can't access the mailbox during this procedure. Finally, you get the current size and number of items in the Recoverable Items folder. After you delete items in the Recoverable Items folder in Step 5, you use this information to verify that items were removed.

1. Connect to Exchange Online PowerShell. Be sure to use a user name and password for an administrator account that's been assigned the appropriate management roles in Exchange Online.

2. Run the following command to get information about single item recovery and the deleted item retention period.

   PowerShell

   ```
   Get-Mailbox <username> | FL SingleItemRecoveryEnabled,RetainDeletedItemsFor
   ```

If single item recovery is enabled, you have to disable it in Step 2. If the deleted item retention period isn't set for 30 days (the maximum value in Exchange Online), then you can increase it in Step 2.

3. Run the following command to get the mailbox access settings for the mailbox.

```
PowerShell
Get-CASMailbox <username> | FL
EwsEnabled,ActiveSyncEnabled,MAPIEnabled,OWAEnabled,ImapEnabled,PopEnabled
```

You disable all of these access methods in Step 2.

4. Run the following command to get information about the holds and retention policies applied to the mailbox.

```
PowerShell
Get-Mailbox <username> | FL LitigationHoldEnabled,InPlaceHolds
```

> **Tip**
>
> If there are too many values in the *InPlaceHolds* property and not all of them are displayed, run the `Get-Mailbox <username> | Select-Object -ExpandProperty InPlaceHolds` command to display each value on a separate line.

5. Run the following command to get information about any organization-wide retention policies.

```
PowerShell
Get-OrganizationConfig | FL InPlaceHolds
```

If your organization has any organization-wide retention policies, you have to exclude the mailbox from these policies in Step 3. It might take up to 24 hours to replicate the change.

> **Tip**

If there are too many values in the *InPlaceHolds* property and not all of them are displayed, run the `Get-OrganizationConfig | Select-Object -ExpandProperty InPlaceHolds` command to display each value on a separate line.

6. Run the following command to determine if a delay hold is applied to the mailbox.

```
PowerShell
Get-Mailbox <username> | FL DelayHoldApplied,DelayReleaseHoldApplied
```

If the value of the *DelayHoldApplied* or *DelayReleaseHoldApplied* property is set to **True**, a delay hold is applied to the mailbox and must be removed. For more information about delay holds, see Step 4: Remove the delay hold from the mailbox.

If the value of either the *DelayHoldApplied* or *DelayReleaseHoldApplied* property is set to **False**, a delay hold isn't applied to the mailbox, and you can skip Step 4.

7. Run the following command to get the current size and total number of items in folders and subfolders in the Recoverable Items folder in the user's primary mailbox.

```
PowerShell
Get-MailboxFolderStatistics <username> -FolderScope RecoverableItems | FL
Name,FolderAndSubfolderSize,ItemsInFolderAndSubfolders
```

If the user's archive mailbox is enabled, run the following command to get the size and total number of items in folders and subfolders in the Recoverable Items folder in their archive mailbox.

```
PowerShell
Get-MailboxFolderStatistics <username> -FolderScope RecoverableItems -Archive |
FL Name,FolderAndSubfolderSize,ItemsInFolderAndSubfolders
```

When you delete items in Step 5, you can choose to delete or not delete items in the Recoverable Items folder in the user's primary archive mailbox. If auto-expanding archiving is enabled for the mailbox, items in an auxiliary archive mailbox won't be deleted.

Step 2: Prepare the mailbox

ⓘ Note

The actions in this step (disabling the Managed Folder Assistant and Single Item Recovery) are preparation. They aren't the delete mechanism. The Managed Folder Assistant (also called messaging records management, or MRM) enforces retention; it doesn't delete items from the Recoverable Items folder. The actual delete happens in Step 5 with the `New-ComplianceSearchAction -Purge -PurgeType HardDelete` command.

After you collect and save information about the mailbox, prepare the mailbox by performing the following tasks:

- **Disable client access to mailbox** so that the mailbox owner can't access their mailbox or make any changes to the mailbox data during this procedure.
- **Increase the deleted item retention period** to 30 days (the maximum value in Exchange Online) so that items aren't purged from the Recoverable Items folder before you can delete them in Step 5.
- **Disable single Item recovery** so that items aren't retained (during the deleted item retention period) after you delete them from the Recoverable Items folder in Step 5.
- **Disable the Managed Folder Assistant** so that it doesn't process the mailbox and retain the items that you delete in Step 5.

Perform the following steps in Exchange Online PowerShell.

1. Run the following command to disable all client access to the mailbox. The command syntax assumes that all client access methods are enabled on the mailbox.

   PowerShell

   ```
   Set-CASMailbox <username> -EwsEnabled $false -ActiveSyncEnabled $false -
   MAPIEnabled $false -OWAEnabled $false -ImapEnabled $false -PopEnabled $false
   ```

   > **Note**
   >
   > It might take up to 60 minutes to disable all client access methods to the mailbox. Disabling these access methods doesn't disconnect the mailbox owner if they're currently signed in. If the owner isn't signed in, they can't access their mailbox after these access methods are disabled.

2. Run the following command to increase the deleted item retention period to the maximum of 30 days. This step assumes that the current setting is less than 30 days.

PowerShell

```
Set-Mailbox <username> -RetainDeletedItemsFor 30
```

3. Run the following command to disable single item recovery.

PowerShell

```
Set-Mailbox <username> -SingleItemRecoveryEnabled $false
```

> ⓘ Note
>
> It might take up to 240 minutes to disable single item recovery. Don't delete items in the Recoverable Items folder until this period has elapsed.

4. Run the following command to prevent the Managed Folder Assistant from processing the mailbox. You can disable the Managed Folder Assistant only if a retention policy with a Preservation Lock isn't applied to the mailbox.

PowerShell

```
Set-Mailbox <username> -ElcProcessingDisabled $true
```

> ⓘ Note
>
> Setting the *ElcProcessingDisabled* parameter to *True* doesn't prevent all messages within the Recoverable Items folder from processing. If holds are removed from mailboxes, the substrate hold folder is deleted over time.

Step 3: Remove all holds from the mailbox

Before you can delete items from the Recoverable Items folder, remove all holds (identified in Step 1) that you placed on the mailbox. Remove all holds so that the system doesn't retain items after you delete them from the Recoverable Items folder. The following sections explain how to remove different types of holds on a mailbox. For tips about how to identify the type of hold that might be placed on a mailbox, see Identify hold types by InPlaceHolds property values. For more information, see How to identify the type of hold placed on an Exchange Online mailbox.

> ⊗ Caution

Check with your records management or legal departments before removing a hold from a mailbox.

Remove Litigation Hold from the mailbox

Run the following command in Exchange Online PowerShell to remove a Litigation Hold from the mailbox.

PowerShell

```
Set-Mailbox <username> -LitigationHoldEnabled $false
```

> **Note**
>
> Similar to disabling single item recovery, it might take up to 240 minutes to remove the Litigation Hold. Don't delete items from the Recoverable Items folder until this period has elapsed.

Remove retention policies applied to specific mailboxes

Run the following command in Security & Compliance PowerShell to identify the retention policy that is applied to the mailbox. This command also returns any Teams conversation retention policies applied to a mailbox. Use the GUID (not including the `mbx` or `skp` prefix) for the retention policy that you identified in Step 1.

PowerShell

```
Get-RetentionCompliancePolicy <retention policy GUID without prefix> | FL Name
```

After you identify the retention policy, go to **Data lifecycle management** > **Retention policies** in the Microsoft Purview portal, edit the retention policy that you identified in the previous step, and remove the mailbox from the list of recipients that are included in the retention policy.

Remove organization-wide retention policies

Organization-wide, Exchange-wide, and Teams-wide retention policies apply to every mailbox in the organization. You apply these policies at the organization level (not the mailbox level). When you run the **Get-OrganizationConfig** cmdlet in Step 1, it returns these policies. Run the following command in Security & Compliance PowerShell to identify the organization-wide retention

policies. Use the GUID (not including the `mbx` prefix) for the organization-wide retention policies that you identified in Step 1.

PowerShell

```
Get-RetentionCompliancePolicy <retention policy GUID without prefix> | FL Name
```

After you identify the organization-wide retention policies, go to the **Data lifecycle management** > **Retention policies** page in the Microsoft Purview portal, edit each organization-wide retention policy that you identified in the previous step, and add the mailbox to the list of excluded recipients. This action removes the user's mailbox from the retention policy.

> ⓘ **Important**
>
> After you exclude a mailbox from an organization-wide retention policy, it might take up to 24 hours to synchronize this change and remove the mailbox from the policy.

Remove retention labels that preserve mailbox items

When a user applies a label that's configured to retain content or retain and then delete content to any folder or item in their mailbox, the *ComplianceTagHoldApplied* mailbox property is set to **True**. When this property is set, the mailbox is considered to be on hold, as if it was placed on Litigation Hold or assigned to a retention policy.

To view the value of the *ComplianceTagHoldApplied* property, run the following command in Exchange Online PowerShell:

PowerShell

```
Get-Mailbox <username> |FL ComplianceTagHoldApplied
```

After you identify that a mailbox is on hold because a retention label is applied to a folder or item, you can use search in the Microsoft Purview portal to search for labeled items by using the **Retention label** condition. To learn how to search for content with specific retention labels, see:

- The "Using Content Search to find all content with a specific retention label" section in Learn about retention policies and retention labels
- The "Search conditions" section in Use the condition builder to create search queries in eDiscovery.

For more information about labels, see Learn about retention policies and retention labels.

> **Tip**
>
> The *ComplianceTagHoldApplied* value doesn't automatically revert to *False* after all labels are removed. For information about how to reset this property, see **Identify mailboxes on hold because a retention label is applied to a folder or item**.

Remove eDiscovery holds from the mailbox

Run the following commands in Security & Compliance PowerShell to identify the hold associated with an eDiscovery case (called *eDiscovery holds*) that you applied to the mailbox. Use the GUID (not including the `UniH` prefix) for the eDiscovery hold that you identified in Step 1. The second command displays the name of the eDiscovery case the hold is associated with; the third command displays the name of the hold.

PowerShell

```
$CaseHold = Get-CaseHoldPolicy <hold GUID without prefix>
```

PowerShell

```
Get-ComplianceCase $CaseHold.CaseId | FL Name
```

PowerShell

```
$CaseHold.Name
```

After you identify the name of the eDiscovery case and the hold, go to **eDiscovery** in the Microsoft Purview portal, open the case, and remove the mailbox from the hold. For more information about identifying eDiscovery holds, see the "eDiscovery holds" section in How to identify the type of hold placed on an Exchange Online mailbox.

Step 4: Remove the delay hold from the mailbox

After you remove any type of hold from a mailbox, the next time the Managed Folder Assistant processes the mailbox and detects that a hold is removed, it sets the value of the *DelayHoldApplied* or *DelayReleaseHoldApplied* mailbox property to **True**. This value is called a *delay hold* and means the actual removal of the hold is delayed for 30 days to prevent data from

being permanently deleted from the mailbox. The purpose of a delay hold is to give admins an opportunity to search for or recover mailbox items purged after a hold is removed. When a delay hold is placed on the mailbox, the mailbox is still considered to be on hold for an unlimited duration, as if the mailbox was on Litigation Hold. After 30 days, the delay hold expires, and Microsoft 365 automatically attempts to remove the delay hold (by setting the *DelayHoldApplied* or *DelayReleaseHoldApplied* property to **False**) so that the hold is removed. For more information about a delay hold, see the "Managing mailboxes on delay hold" section in How to identify the type of hold placed on an Exchange Online mailbox.

You must be assigned the Legal Hold role in Exchange Online to use the *RemoveDelayHoldApplied* or *RemoveDelayReleaseHoldApplied* parameter.

If the value of the *DelayHoldApplied* or *DelayReleaseHoldApplied* property is set to **True**, run one of the following commands to remove the delay hold:

PowerShell

```
Set-Mailbox <username> -RemoveDelayHoldApplied
```

Or

PowerShell

```
Set-Mailbox <username> -RemoveDelayReleaseHoldApplied
```

Step 5: Delete items in the Recoverable Items folder

Now you're ready to delete items in the Recoverable Items folder by using the New-ComplianceSearch and New-ComplianceSearchAction cmdlets in Security & Compliance PowerShell.

> **Note**
>
> If the mailbox quota is met and the user mailbox is declining emails, you might receive a 554 5.2.0 error when deleting recoverable items. For more information, see "554 5.2.0 STOREDRV.Deliver.Exception" when sending emails in Exchange Online. For information about increasing the Recoverable Items quota, see Increase the Recoverable Items quota for mailboxes on hold.

Here's an overview of the process to search for and delete items in a user's Recoverable Items folder:

1. **Identify recoverable items in the mailbox**: Use the Get-RecoverableItems cmdlet to retrieve information about items stored in the Recoverable Items folder of the target mailbox. This step helps you understand the volume and characteristics of recoverable items before attempting deletion. For example, you can run the cmdlet to review recoverable items associated with a specific mailbox and determine appropriate criteria for scoping a deletion.

   Review the output to identify common attributes, such as subject, sender, or date range, that you can use to precisely target the items you want to delete.

2. **Create a content search using searchable item properties**. After identifying the recoverable items, use the New-ComplianceSearch cmdlet (in Security & Compliance PowerShell) or use eDiscovery search tools in the Microsoft Purview portal to create a content search that returns items from the target items in the recoverable items folder. Use searchable item properties to target the specific items you want to delete. Use conditions that include properties such as *subject*, *sender*, *recipients*, *message type*, and *date range*.

   Define the search criteria carefully to ensure that only the intended items are included in the search results. After the search is created, run it and validate that the results match the items you identified in the previous step.

3. After you create a content search and validate that it returns the items that you want to delete, use the `New-ComplianceSearchAction -Purge -PurgeType HardDelete` command (in Security & Compliance PowerShell) to permanently delete the items returned by the content search. For example, you can run a command similar to the following command:

   PowerShell

   ```
   New-ComplianceSearchAction -SearchName "RecoverableItems" -Purge -PurgeType
   HardDelete
   ```

4. When you run the previous command, it deletes up to 10 items per mailbox. That limit means you might have to run the `New-ComplianceSearchAction -Purge` command multiple times to delete all the items that you want to delete in the Recoverable Items folder. To delete more items, you first have to remove the previous compliance search purge action. You do this by running the `Remove-ComplianceSearchAction` cmdlet. For example, to delete the purge action that you ran in the previous step, run the following command:

   ```
   ```

PowerShell

```
Remove-ComplianceSearchAction "RecoverableItems_Purge"
```

After you complete this step, you can create a new compliance search purge action to delete more items. You must delete each purge action before creating a new one.

To get a list of the compliance search actions, you can run the `Get-ComplianceSearchAction` cmdlet. Purge actions are identified by `_Purge` appended to the search name.

Verify that you deleted items

To verify that you deleted items from the Recoverable Items folder of a mailbox, use the **Get-MailboxFolderStatistics** cmdlet in Exchange Online PowerShell to check the size and number of items in the Recoverable Items folder. You can compare these statistics with the ones you collected in Step 1.

Run the following command to get the current size and total number of items in folders and subfolders in the Recoverable Items folder in the user's primary mailbox.

PowerShell

```
Get-MailboxFolderStatistics <username> -FolderScope RecoverableItems | FL
Name,FolderAndSubfolderSize,ItemsInFolderAndSubfolders
```

Run the following command to get the size and total number of items in folders and subfolders in the Recoverable Items folder in the user's archive mailbox.

PowerShell

```
Get-MailboxFolderStatistics <username> -FolderScope RecoverableItems -Archive | FL
Name,FolderAndSubfolderSize,ItemsInFolderAndSubfolders
```

Step 6: Revert the mailbox to its previous state

The final step is to revert the mailbox to its previous configuration. This step resets the properties that you changed in Step 2 and reapplies the holds that you removed in Step 3. This step includes:

- Changing the deleted item retention period back to its previous value. Alternatively, you can set this value to 30 days, the maximum value in Exchange Online.
- Re-enabling single item recovery.

- Re-enabling the client access methods so that the owner can access their mailbox.
- Reapplying the holds and retention policies that you removed.
- Re-enabling the Managed Folder Assistant to process the mailbox.

> **Important**
>
> Wait 24 hours after reapplying a hold or retention policy (and verifying that it's in place) before you re-enable the Managed Folder Assistant to process the mailbox.

Perform the following steps (in the specified sequence) in Exchange Online PowerShell.

1. Run the following command to change the deleted item retention period back to its original value. This example assumes that the previous setting is less than 30 days, such as 14 days.

   PowerShell

   ```
   Set-Mailbox <username> -RetainDeletedItemsFor 14
   ```

2. Run the following command to re-enable single item recovery.

   PowerShell

   ```
   Set-Mailbox <username> -SingleItemRecoveryEnabled $true
   ```

3. Run the following command to re-enable all client access methods to the mailbox.

   PowerShell

   ```
   Set-CASMailbox <username> -EwsEnabled $true -ActiveSyncEnabled $true -MAPIEnabled $true -OWAEnabled $true -ImapEnabled $true -PopEnabled $true
   ```

4. Reapply the holds that you removed in Step 3. Depending on the type of hold, use one of the following procedures.

   Litigation Hold

   Run the following command to re-enable a Litigation Hold for the mailbox.

   PowerShell

   ```
   Set-Mailbox <username> -LitigationHoldEnabled $true
   ```

Retention policies applied to specific mailboxes

Use the Microsoft Purview portal to add the mailbox back to the retention policy. Go to the **Data lifecycle management** > **Retention policies** page in the Microsoft Purview portal, edit the retention policy, and add the mailbox back to the list of recipients that the retention policy is applied to.

Organization-wide retention policies

If you removed an organization-wide or Exchange-wide retention policy by excluding it from the policy, use the Microsoft Purview portal to remove the mailbox from the list of excluded users. Go to the **Data lifecycle management** > **Retention policies** page in the Microsoft Purview portal, edit the organization-wide retention policy, and remove the mailbox from the list of excluded recipients. Doing this reapplies the retention policy to the user's mailbox.

eDiscovery case holds

Use the Microsoft Purview portal to add the mailbox back the hold that's associated with an eDiscovery case. Go to **eDiscovery**, open the case, and add the mailbox back to the hold.

5. Run the following command to allow the Managed Folder Assistant to process the mailbox again.

```
PowerShell
Set-Mailbox <username> -ElcProcessingDisabled $false
```

6. To verify that the mailbox is reverted back to its previous configuration, run the following commands and then compare the settings to the ones that you collected in Step 1.

```
PowerShell
Get-Mailbox <username> | FL
ElcProcessingDisabled,InPlaceHolds,LitigationHoldEnabled,RetainDeletedItemsFor,Si
ngleItemRecoveryEnabled
```

```
PowerShell
Get-CASMailbox <username> | FL
EwsEnabled,ActiveSyncEnabled,MAPIEnabled,OWAEnabled,ImapEnabled,PopEnabled
```

Identify hold types by InPlaceHolds property values

The following table describes how to identify different types of holds based on the values in the *InPlaceHolds* property when you run the **Get-Mailbox** or **Get-OrganizationConfig** cmdlets. For more detailed information, see How to identify the type of hold placed on an Exchange Online mailbox.

You need to remove all holds and retention policies from a mailbox before you can successfully delete items in the Recoverable Items folder.

Expand table

Hold type	Example value	How to identify the hold
Litigation Hold	`True`	The *LitigationHoldEnabled* property is set to `True`.
Retention policies in the Microsoft Purview portal applied to specific mailboxes	`mbxcdbbb86ce60342489bff371876e7f224` or `skp127d7cf1076947929bf136b7a2a8c36f`	When you run the **Get-Mailbox** cmdlet, the *InPlaceHolds* property also contains GUIDs of retention policies applied to the mailbox. You can identify retention policies because the GUID starts with the `mbx` prefix. If the GUID of the retention policy starts with the `skp` prefix, that indicates that the retention policy is applied to Skype for Business conversations. To identity the retention policy that's applied to the mailbox, run the following command in Security & Compliance PowerShell: `Get-RetentionCompliancePolicy <retention policy GUID without prefix> | FL Name` Be sure to remove the `mbx` or `skp` prefix when you run this command.
Organization-wide retention policies in the Microsoft Purview portal	No value or `-mbxe9b52bf7ab3b46a286308ecb29624696` (indicates that the mailbox is excluded from an organization-wide policy)	Even if the *InPlaceHolds* property is empty when you run the **Get-Mailbox** cmdlet, one or more organization-wide retention policies might still apply to the mailbox. To verify this, run the `Get-OrganizationConfig | FL InPlaceHolds` command in Exchange Online PowerShell to get a list of the GUIDs for organization-wide retention policies. The GUID for organization-wide retention policies applied to Exchange mailboxes starts with the `mbx` prefix. For example, `mbxa3056bb15562480fadb46ce523ff7b02`.

Hold type	Example value	How to identify the hold
		To identity the organization-wide retention policy that's applied to the mailbox, run the following command in Security & Compliance PowerShell: `Get-RetentionCompliancePolicy <retention policy GUID without prefix> | FL Name` If a mailbox is excluded from an organization-wide retention policy, the GUID for the retention policy appears in the *InPlaceHolds* property of the user's mailbox when you run the **Get-Mailbox** cmdlet. It's identified by the prefix `-mbx`. For example, `-mbxe9b52bf7ab3b46a286308ecb29624696`.

eDiscovery case hold in the Microsoft Purview portal	`UniH7d895d48-7e23-4a8d-8346-533c3beac15d`	The *InPlaceHolds* property also contains the GUID of any hold associated with an eDiscovery case in the Microsoft Purview portal that might be placed on the mailbox. You can tell this is an eDiscovery case hold because the GUID starts with the `UniH` prefix.
		You can use the `Get-CaseHoldPolicy` cmdlet in Security & Compliance PowerShell to get information about the eDiscovery case that the hold on the mailbox is associated with. For example, you can run the command `Get-CaseHoldPolicy <hold GUID without prefix> | FL Name` to display the name of the case hold that's on the mailbox. Be sure to remove the `UniH` prefix when you run this command.
		
		To identity the eDiscovery case that the hold on the mailbox is associated with, run the following commands:
		
		`$CaseHold = Get-CaseHoldPolicy <hold GUID without prefix>`
		
		`Get-ComplianceCase $CaseHold.CaseId | FL Name`

ⓘ**Note:** The author created this article with assistance from AI. Learn more

Last updated on 06/23/2026

Increase the Recoverable Items quota for mailboxes on hold in eDiscovery

The default Exchange retention policy—named *Default MRM Policy*—that automatically applied to new mailboxes in Exchange Online contains a retention tag named Recoverable Items 14 days move to archive. This retention tag moves items from the Recoverable Items folder in the user's primary mailbox to the Recoverable Items folder in the user's archive mailbox after the 14-day retention period expires for an item.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Emails in the Deletions folder retain based on the **RetainDeletedItemsFor** parameter and move to other folders in recoverable deleted items and then to the archive mailbox. For this process to work, the user's archive mailbox must be enabled. If the archive mailbox isn't enabled, no action is taken, which means that items in the Recoverable Items folder for a mailbox on hold aren't moved to the archive mailbox after the 14-day retention period expires. Because nothing is deleted from a mailbox on hold, it's possible that the storage quota for the Recoverable Items folder might be exceeded, especially if the user's archive mailbox isn't enabled. For information about deleting items from the Recoverable Items folder, see Delete Recoverable Items from mailboxes on hold.

To help reduce the chance of exceeding the Recoverable Items folder storage quota, the storage quota for the Recoverable Items folder automatically increases from 30 GB to 100 GB when you create an eDiscovery hold on a mailbox in Exchange Online. If the archive mailbox is enabled, the storage quota for the Recoverable Items folder in the archive mailbox also increases from 30 GB to 100 GB. If archive is enabled on a mailbox that is on hold with auto-expanding archive disabled, the storage quota of the Recoverable Items folder quota increases from 100 GB to 105 GB. If the auto-expanding archiving feature in Exchange Online is enabled, the total storage quota for the user's archive mailbox, including the Recoverable Items folder, is 1.5 TB.

The following table summarizes the storage quota for the Recoverable Items folder.

Expand table

Location of Recoverable Items folder	Mailboxes not on hold	Mailboxes on hold
Primary mailbox	30 GB	100 GB
Archive mailbox, including Recoverable Items folder *	1.5 TB	1.5 TB

> **Note**
>
> * The initial storage quota for the archive mailbox is 100 GB for users with an Exchange Online (Plan 2) license. However, when auto-expanding archiving is turned on for mailboxes on hold, the storage quota for both the archive mailbox and the Recoverable Items folder increases to 110 GB. Additional archive storage space (which includes the Recoverable Items folder) up to 1.5 TB is provisioned when necessary. For more information about auto-expanding archiving, see Learn about auto-expanding archiving.

> **Note**
>
> When archiving is enabled for a specific user whose mailbox is on hold or assigned to a retention policy, the storage quota for the user's primary mailbox's Recoverable Items folder raises from 100 GB to 105 GB and Recoverable items warning quota raises from 90 GB to 95 GB. This extra space supports an efficient archiving experience for users, even if auto-expanding archive isn't enabled.

When the storage quota for the Recoverable Items folder in the primary mailbox of a mailbox on hold is close to reaching its limit, you can do the following things:

- **Enable the archive mailbox and turn on auto-expanding archiving.** You can enable an extra storage capacity for the Recoverable Items folder by enabling the archive mailbox and then turning on the auto-expanding archiving feature in Exchange Online. This change results in 110 GB for the Recoverable Items folder in the primary mailbox and up to 1.5 TB of combined storage capacity for the archive and Recoverable Items folder. For more information, see Enable archive mailboxes and Enable auto-expanding archiving.

  > **Note**
  >
  > After you enable the archive for a mailbox that's close to exceeding the storage quota for the Recoverable Items folder, you might want to run the Managed Folder Assistant to manually trigger the assistant to process the mailbox so that expired items move to

the Recoverable Items folder in the archive mailbox. See Run the Managed Folder Assistant to apply the new retention settings for instructions. Other items in the user's mailbox might move to the new archive mailbox. Consider telling the user that this change might happen after you enable the archive mailbox.

- **Create a custom Exchange retention policy for mailboxes on hold.** In addition to enabling the archive mailbox and auto-expanding archiving for mailboxes on Litigation Hold, you might also want to create a custom Exchange retention policy for mailboxes on hold. This policy lets you apply a retention policy to mailboxes on hold that's different from the Default MRM Policy that's applied to mailboxes that aren't on hold, and lets you apply retention tags that are designed for mailboxes on hold. This policy includes creating a new retention tag for the Recoverable Items folder.

The rest of this article describes the step-by-step procedures to create a custom Exchange retention policy for mailboxes on hold.

- Step 1: Create a custom retention tag for the Recoverable Items folder
- Step 2: Create a new Exchange retention policy for mailboxes on hold
- Step 3: Apply the new Exchange retention policy to mailboxes on hold
- (Optional) Step 4: Run the Managed Folder Assistant to apply the new retention settings

Step 1: Create a custom retention tag for the Recoverable Items folder

First, create a custom retention tag (called a retention policy tag or RPT) for the Recoverable Items folder. This RPT moves items from the Recoverable Items folder in the user's primary mailbox to the Recoverable Items folder in the user's archive mailbox. Use PowerShell to create an RPT for the Recoverable Items folder. You can't use the Exchange admin center (EAC).

1. Connect to Exchange Online PowerShell

2. Run the following command to create a new RPT for the Recoverable Items folder:

   PowerShell

   ```
   New-RetentionPolicyTag -Name <Name of RPT> -Type RecoverableItems -
   AgeLimitForRetention <Number of days> -RetentionAction MoveToArchive
   ```

   For example, the following command creates an RPT for the Recoverable Items folder named "Recoverable Items 30 days for mailboxes on hold", with a retention period of 30

days. This means that after an item has been in the Recoverable Items folder for 30 days, it's moved to the Recoverable Items folder in the user's archive mailbox.

PowerShell

```
New-RetentionPolicyTag -Name "Recoverable Items 30 days for mailboxes on hold" -
Type RecoverableItems -AgeLimitForRetention 30 -RetentionAction MoveToArchive
```

Tip

Set the retention period (defined by the *AgeLimitForRetention* parameter) for the Recoverable Items RPT to match the deleted item retention period for the mailboxes that the RPT applies to. This setup gives users the entire deleted item retention period to recover deleted items before they're moved to the archive mailbox. In the previous example, the retention period was set to 30 days based on the assumption that the deleted item retention period for mailboxes is also 30 days. An Exchange Online mailbox is configured to retain deleted items for 14 days, by default. But you can change this setting to a maximum of 30 days. For more information, see **Change the deleted item retention period for a mailbox in Exchange Online**.

Step 2: Create a new Exchange retention policy for mailboxes on hold

Next, create a new retention policy and add retention tags to it, including the Recoverable Items RPT that you created in Step 1. In Step 3, you apply this new policy to mailboxes on hold.

Before you create the new retention policy, determine the additional retention tags that you want to add. For a list of the retention tags that are added to the Default MRM Policy and for information about creating new retention tags, see the following:

- Default Retention Policy in Exchange Online
- Default folders that support Retention Policy Tags
- The "Create a retention tag" section in the Create a Retention Policy article.

You can use the EAC or Exchange Online PowerShell to create a retention policy.

Use the EAC to create a retention policy

Perform the following steps in the EAC to create the retention policy:

1. In the EAC, go to **Compliance management** > **Retention policies**, and then select **Add** +.
2. On **New retention policy**, under **Name**, type a name that describes the purpose of the retention policy; for example, **MRM Policy for Mailboxes on Hold**.
3. Under **Retention tags**, select **Add** +.
4. In the list of retention tags, select the Recoverable Items RPT that you created in Step 1, and then select **Add**.
5. Select additional retention tags to add to the retention policy. For example, you might want to add the same tags that are included in the Default MRM Policy.
6. When you finish adding retention tags, select **OK**.
7. Select **Save** to create the new retention policy. The details pane displays the retention tags linked to the retention policy.

Use Exchange Online PowerShell to create a retention policy

Run the following command to create new retention policy for mailboxes on hold.

PowerShell

```
New-RetentionPolicy <Name of retention policy>  -RetentionPolicyTagLinks <list of retention tags>
```

For example, the following command creates the retention policy and links the retention tags listed in the EAC procedure earlier in this section.

PowerShell

```
New-RetentionPolicy "MRM Policy for Mailboxes on Hold"  -RetentionPolicyTagLinks "Recoverable Items 30 days for mailboxes on hold","1 Month Delete","1 Week Delete","1 Year Delete","5 Year Delete","6 Month Delete","Default 2 year move to archive","Junk Email","Never Delete","Personal 1 year move to archive","Personal 5 year move to archive"
```

Step 3: Apply the new Exchange retention policy to mailboxes on hold

In the last step, apply the new retention policy that you created in Step 2 to mailboxes on hold in your organization. You can use the EAC or Exchange Online PowerShell to apply the retention policy to a single mailbox or to multiple mailboxes.

Use the EAC to apply the new retention policy

Use the following steps in the EAC to apply the new retention policy to one or more mailboxes:

1. Go to **Recipients** > **Mailboxes**.
2. In the list view, select the mailbox you want to apply the retention policy to, then select **Edit**.
3. On the **User Mailbox** page, select **Mailbox features**.
4. Under **Retention policy**, select the retention policy that you created in Step 2, then select **Save**.

You can also use the EAC to apply the retention policy to multiple mailboxes.

1. Go to **Recipients** > **Mailboxes**.
2. In the list view, use the Shift or Ctrl keys to select multiple mailboxes.
3. In the details pane, select **More options**.
4. Under **Retention Policy**, select **Update**.
5. On the **Bulk assign retention policy** page, select the retention policy that you created in Step 2, then select **Save**.

Use Exchange Online PowerShell to apply the new retention policy

You can use Exchange Online PowerShell to apply a new retention policy to a single mailbox. But the real power of PowerShell is that you can use it to quickly identify all the mailboxes in your organization that are on Litigation Hold, and then apply the new retention policy to all mailboxes on hold in a single command. Here are some examples of using Exchange PowerShell to apply a retention policy to one or more mailboxes. All of the examples apply the retention policy that you created in Step 2.

This example applies the new retention policy to Pilar Pinilla's mailbox.

PowerShell

```
Set-Mailbox "Pilar Pinilla" -RetentionPolicy "MRM Policy for Mailboxes on Hold"
```

This example applies the new retention policy to all mailboxes in the organization that are on Litigation Hold.

PowerShell

```
$LitigationHolds = Get-Mailbox -ResultSize unlimited | Where-Object
{$_.LitigationHoldEnabled -eq 'True'}
```

PowerShell

```
$LitigationHolds.DistinguishedName | Set-Mailbox -RetentionPolicy "MRM Policy for Mailboxes on Hold"
```

You can use the **Get-Mailbox** cmdlet to verify that you applied the new retention policy.

Here are some examples to verify that the commands in the previous examples applied the "MRM Policy for Mailboxes on Hold" retention policy to mailboxes on Litigation Hold.

PowerShell

```
Get-Mailbox "Pilar Pinilla" | Select RetentionPolicy
```

PowerShell

```
Get-Mailbox -ResultSize unlimited | Where-Object {$_.LitigationHoldEnabled -eq 'True'} | FT DisplayName,RetentionPolicy -Auto
```

(Optional) Step 4: Run the Managed Folder Assistant to apply the new retention settings

After you apply the new Exchange retention policy to mailboxes on hold, it can take up to seven days in Exchange Online for the Managed Folder Assistant to process these mailboxes by using the settings in the new retention policy. Instead of waiting for the Managed Folder Assistant to run, you can use the **Start-ManagedFolderAssistant** cmdlet to manually trigger the assistant to process the mailboxes that you applied the new retention policy to.

Run the following command to start the Managed Folder Assistant for Pilar Pinilla's mailbox.

PowerShell

```
Start-ManagedFolderAssistant "Pilar Pinilla"
```

Run the following commands to start the Managed Folder Assistant for all mailboxes on hold.

PowerShell

```
$MailboxesOnHold = Get-Mailbox -ResultSize unlimited | Where-Object {($_.InPlaceHolds -ne $null) -or ($_.LitigationHoldEnabled -eq "True")}
```

PowerShell

```
$MailboxesOnHold.DistinguishedName | Start-ManagedFolderAssistant
```

Additional considerations for archive mailboxes and retention policies

User communication considerations

- When you enable a user's archive mailbox, consider telling the user that other items in their mailbox (not just items in the Recoverable Items folder) might move to the archive mailbox. This behavior occurs because the Default MRM Policy that's assigned to Exchange Online mailboxes contains a retention tag (named Default 2 years move to archive) that moves items to the archive mailbox two years after the date the item is delivered to the mailbox or created by the user. For more information, see Default Retention Policy in Exchange Online.
- When you enable a user's archive mailbox, consider telling the user that they can recover deleted items in the Recoverable Items folder in their archive mailbox. They can do this in Outlook by selecting the **Deleted Items** folder in the archive mailbox, then selecting **Recover Deleted Items from Server** on the **Home** tab. For more information about recovering deleted items, see Recover deleted items in Outlook for Windows.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/19/2026

Preserve Bcc and expanded distribution group recipients for eDiscovery

By creating litigation holds, eDiscovery holds, and retention policies in Microsoft Purview, you can preserve mailbox content to meet regulatory compliance and eDiscovery requirements. All messages include information about recipients directly addressed in the To and Cc fields by default. But your organization might require the ability to search for and reproduce details about all recipients of a message. This requirement includes:

- **Recipients addressed using the Bcc field of a message:** Bcc recipients are stored in the message in the sender's mailbox, but not included in headers of the message delivered to recipients.
- **Expanded distribution group recipients:** Recipients who receive the message because they're members of a distribution group to which the message was addressed, either in the To, Cc, or Bcc fields.

Exchange Online and Exchange Server 2013 (Cumulative Update 7 and later versions) retain information about Bcc and expanded distribution group recipients. You can search for this information by using an eDiscovery tool in the Microsoft Purview portal.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

How Bcc recipients and expanded distribution group recipients are preserved

> **Important**
>
> Only the sender's mailbox's Sent Items contains distribution group expansion.

Information about Bcc recipients is stored with the message in the sender's mailbox. This information is indexed and available to eDiscovery searches and holds.

Information about expanded distribution group recipients is stored with the message after you place a mailbox on Litigation Hold. In Microsoft 365, this information is also stored when a Microsoft Purview retention policy is applied to a mailbox. Distribution group membership is determined at the time the message is sent. The expanded recipients list stored with the message isn't impacted by changes to membership of the group after the message is sent.

Expand table

Information about...	Is stored in...	Is stored by default?	Is accessible to...
To: and Cc: recipients	Message properties in the sender and recipients' mailboxes.	Yes	Sender, recipients, and compliance officers
Bcc: recipients	Message property in the sender's mailbox.	Yes	Sender and compliance officers
Expanded distribution group recipients	Message properties in the sender's mailbox.	No. Expanded distribution group recipient information is stored after a mailbox is placed on Litigation Hold or assigned to a Microsoft Purview retention policy.	Sender and Compliance officers

Searching for messages with expanded distribution group recipients

When you search for messages sent to a recipient, eDiscovery search results now include messages sent to a distribution group that the recipient is a member of, but only in the sender mailbox and if the sender is on hold.

> **Important**
>
> Distribution group expansion details are only available in the sender's mailbox under Sent Items. If you search the recipient's mailbox, messages sent to distribution groups don't show the expanded list of group members.

Scenario: John is a member of the US-Sales distribution group. This table shows eDiscovery search results when Bob sends a message to John directly or indirectly via a distribution group.

Expand table

When you search Bob's mailbox for messages sent...	And the message is sent with...	Results include message?
To: John	John on To:	Yes
To: John	US-Sales on To:	Yes
To: US-Sales	US-Sales on To:	Yes
Cc: John	John on Cc:	Yes
Cc: John	US-Sales on Cc:	Yes
Cc: US-Sales	US-Sales on Cc:	Yes

Expand table

When you search John's mailbox for message...	And the message is sent with...	Results include message?
To/Cc	US-Sales on To/Cc:	No. Only sender mailbox can be searched.

Searching for messages with Bcc

Scenario: Bob sends an email to John (To/Cc) and Jack (Bcc directly, or indirectly via a distribution group). The following tables show the eDiscovery search results.

> **Important**
>
> Distribution group expansion detail is only available if the sender is on hold.

Expand table

When you search Bob's mailbox...	For messages sent...	Results include message?	Notes
Bob's mailbox	To:/Cc: John	Yes	Shows an indication that Jack was Bcc'ed.
Bob's mailbox	Bcc: Jack	Yes	Shows an indication that Jack was Bcc'ed.
Bob's mailbox	Bcc: Jack (via distribution	Yes	List of members of the Bcc'ed distribution group, expanded when the message was sent, is

When you search Bob's mailbox...	For messages sent...	Results include message?	Notes
	group)		visible in eDiscovery search preview, export, and logs.

Expand table

When you search John's mailbox...	For messages sent...	Results include message?	Notes
John's mailbox	To:/Cc: John	Yes	No indication of Bcc recipients.
John's mailbox	Bcc: Jack (directly or via distribution group)	No	Bcc: information isn't stored in the message delivered to recipients. You must search the **sender's mailbox**.

Expand table

When you search Jack's mailbox...	For messages sent...	Results include message?	Notes
Jack's mailbox	To:/Cc: John (directly or via distribution group)	Yes	To:/Cc: information is included in message delivered to all recipients.
Jack's mailbox	Bcc: Jack (directly or via distribution group)	No	Bcc: information isn't stored in the message delivered to recipients. You must search the **sender's mailbox.**

Frequently asked questions

When and where is Bcc recipient information stored?

By default, the original message in the sender's mailbox stores Bcc recipient information. If the Bcc recipient is a distribution group, the system expands distribution group membership only if the sender's mailbox is on hold or assigned to a Microsoft 365 retention policy.

When and where is the list of expanded distribution group recipients stored?

The system expands group membership when the sender sends the message. The original message in the sender's mailbox stores the list of expanded distribution group members. The sender's mailbox must be on Litigation Hold or assigned to a Microsoft 365 retention policy.

Can the To/Cc recipients see which recipients were Bcc'ed?

No. This information isn't included in message headers, and isn't visible to To/Cc recipients. The sender can see the Bcc field stored in the original message stored in their mailbox. Compliance officers can see this information when searching the sender's mailbox.

How can I ensure that expanded distribution group recipients are always preserved?

To ensure that expanded distribution group members are always preserved with a message, Place all mailboxes on hold or create an organization-wide Microsoft 365 retention policy.

Which types of groups are supported?

Distribution groups, mail-enabled security groups, and dynamic distribution groups are supported.

Is there a limit on the number of distribution group recipients that are expanded and stored in the message?

The system preserves up to 10,000 members of a distribution group.

Are nested distribution groups supported?

Yes, the system expands 25 levels of nested distribution groups.

Where is the Bcc and expanded distribution group recipient information visible?

Compliance officers can see Bcc and expanded distribution group recipient information when performing an eDiscovery search. Bcc and expanded distribution group recipients are included in search results copied to a Discovery mailbox or exported to a PST file and in the eDiscovery log included in search results. Bcc recipient information is also available in search preview.

What happens if a member of a distribution group is hidden from the organization's global address list (GAL)?

There's no impact. If recipients are hidden from the GAL, they're still included in the list of recipients for the expanded distribution group.

Last updated on 04/01/2026

Process reports in eDiscovery

Every process in Microsoft Purview eDiscovery generates a downloadable process report that contains detailed information about what happened during the process. This article provides scenario-based guidance for interpreting report data in depth. Use it to understand unexpected item counts, identify location errors, see how settings affected results, and know what to look for in reports when investigating specific outcomes.

For more information about the structure and contents of each report, see Use process reports in eDiscovery.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Best practices for using process reports

- **Download reports for every process**: Even when a process completes successfully, download the report package to keep a record of what was collected, from where, and with what settings. This information is valuable for transparency and audit purposes.
- **Check the Summary.csv first**: Start with the Summary CSV to get an overview of the process outcome, including item counts, errors, and the factors that affected the count.
- **Use the Items.csv for item-level investigation**: When you need to understand why a specific item was or wasn't included, look at the Items.csv for detailed per-item status and error information.
- **Use the Locations.csv for scope validation**: Verify that all intended locations were included and successfully processed. Identify any locations with errors that need attention.
- **Compare Settings.csv across processes**: When comparing the results of two processes, compare their Settings CSV reports to confirm the same configuration was used.
- **Use the Statistics.csv for search analysis**: After a *Generate statistics* process, use this report to understand the distribution of search hits across keywords, data types, sensitive information types, and communication participants.

Understanding search result counts

Search result counts can vary between runs and differ from export or collection counts. For a detailed explanation of the behavioral reasons behind these differences, see Understanding statistics and search results and Differences between statistics and export results. The sections below explain how to use the process reports to investigate these differences.

Investigating missing items in search results

If expected content doesn't appear in search results, use the process reports from your *Generate statistics* or *Generate sample* process to identify the cause.

What to look at in the process report:

- **Settings.csv — Partially indexed items setting**: Check whether the process was configured to include partially indexed items. By default, search results exclude content that isn't fully indexed, such as image-only files, unsupported file types, or items with processing errors. If the setting shows that the process didn't include partially indexed items, rerun the search with this option enabled to determine if the missing items are unindexed.
- **Settings.csv — Security filter applied**: If this value is *Yes*, a compliance boundary (search permissions filter) limits which locations the search can access. The search silently excludes items in locations outside the filter scope. The **Location restriction** field shows the filter criteria. If you suspect the filter is excluding expected content, ask an eDiscovery administrator or another user without search permission filter applied to run the same search for comparison.
- **Locations.csv — Status column**: Check for locations with errors. If a location shows a failure status, the **ErrorWarning** field explains why. For more information about common reasons that can affect location access, see Export considerations.
- **Locations.csv — Count column**: If a location you expected to have results shows a count of zero, the search query didn't match any content in that location. Verify the query is correct and that the content exists at the source. Items you recently added to a mailbox or site might not yet be indexed. Rerun the search after a short wait to resolve this condition.
- **Locations.csv — Missing locations**: If an expected location doesn't appear in the report at all, it wasn't included in the search scope. Verify the data sources configured for the search and add any missing mailboxes or sites.

> **Tip**
>
> When investigating missing items, start with the **Locations.csv** to confirm all expected locations were searched successfully, then check the **Settings.csv** to verify the right options were enabled. If both look correct and items are still missing, the content might be deleted or not yet indexed at the time of the search.

Search result differences with tenant-wide searches

When you use a tenant-wide search scope such as **All people and groups**, the results can vary depending on additional source configuration settings that control which user and content types are included. These settings expand or narrow the scope of the search and directly affect item counts.

What to look at in the process report:

- **Settings.csv — Tenant-wide process settings**: When a search uses **All people and groups**, the Settings CSV captures which additional source types were enabled at the time of the search export or **Add to Review** set. Check these settings to understand why certain content was included or excluded:
  - **Include departed users**: When enabled, the search includes mailboxes and OneDrive sites of users who left the organization (soft-deleted or inactive accounts). If this setting wasn't enabled and you expected to find content from former users, it explains why those items are missing from results.
  - **Include shared Teams channels**: When enabled, the search includes content from shared Teams channels. Shared channel data is stored in separate SharePoint sites, so it's only included when this setting is active. If you're missing shared channel content, verify this setting was enabled.
  - **Include unlicensed and on-premises users**: When enabled, the search includes mailboxes for users who don't have a license or are synced from on-premises environments.
  - **Include guest users**: When enabled, the search includes content associated with guest user accounts in your tenant.

> **Note**

You configure tenant-wide source settings at the process level. Including these settings increases the scope of the process and can affect performance.

Interpreting count differences between search and export

When you export search results or add them to a review set, the final item count often differs from the original search estimate. For the full list of reasons, see Differences between statistics and export results. The Summary CSV report shows the specific factors that affected the count for your process.

What to look at in the process report:

- **Information section of Summary.csv**: This section lists every factor that increased or decreased the item count compared to the original estimate, such as cloud attachments, contextual conversations, consolidated transcripts, SharePoint file versions, and duplicates skipped in a review set. For a complete list of these factors and what each one means, see Information section.
- **Search results section of Summary.csv**: Shows the breakdown of indexed items, partially indexed items, and advanced indexed items. If the count is lower than expected, check whether you included partially indexed items in the process settings. For more information, see Assess partially indexed items in your organization.
- **Error section of Summary.csv**: Shows retrieval exceptions — items that couldn't be exported or added to the review set. For the specific exception types tracked in this section, see Error.
- **Sampling configured**: If you configured sampling (confidence-based or percentage-based) when adding search results to a review set, the item count in the review set is intentionally lower than the total search estimate. This reduction reflects the configured sample size, not a processing discrepancy. Check the Settings CSV report to confirm the sampling mode and parameters used.

Understanding the impact of search settings on results

The Settings CSV report captures the exact configuration used when you submitted a process. Compare this report against your intended configuration to identify discrepancies.

Key settings that affect result scope include:

- **Include document versions**: Determines how many versions of each document are collected. Selecting *All versions* can significantly increase the item count and data volume.

- **Organize conversations into HTML transcripts**: When enabled, individual Teams chat messages consolidate into transcript files, which changes the item count.
- **Include contextual Teams and Viva Engage conversations**: When enabled, the search includes messages outside your search date range to provide conversation context. For 1:1 and group chats, this setting adds messages 12 hours before and after each responsive item. For Teams channel posts and Viva Engage posts, the search includes entire conversation threads regardless of date range. If you see items outside your expected date range, check whether this setting is enabled.
- **Access links in messages**: Controls whether the search retrieves cloud attachments from links in emails and Teams conversations.

Understanding export reports

Why certain items weren't exported

When an export finishes, the Items CSV report and the **Error** section of the Summary CSV report explain what happened to each item.

What to look at in the export report:

- **Items.csv — Status column**: Each row represents an item. Items with a status other than success include an **ErrorWarning** field with the failure reason. Common reasons include:
  - **Item deleted or moved**: The search found the item, but someone removed or relocated it at the source between the time of the search and the export. These items count in the search estimate but can't be retrieved at export time.
  - **Access timeout**: The system couldn't retrieve the item within the allowed time, often due to a large file size or transient service issue.
  - **Permission or access issue**: The export process couldn't access the item, possibly because the source location is locked or the user account lacks access.

- **Items.csv — AddedBy column**: Shows how each item was included in the results. For definitions of each value, see Items CSV report.

- **Summary.csv — Retrieval exceptions**: Shows the total count of items that couldn't be exported. This count matches the number of error entries in the Items.csv file.

Understanding file timestamps in exports

When you extract files from an export package, file timestamps might show the current date instead of the original modified date. This behavior occurs because the Windows built-in ZIP

extractor resets file dates during extraction. The original timestamps are preserved in the **Items.csv** report and inside the file formats themselves.

To preserve timestamps during extraction, use a tool like 7-Zip, WinZip, or WinRAR. For emails in PST format, opening the PST in Outlook shows the original sent and received dates.

Understanding export settings and traceability

The Settings CSV report serves as the authoritative record of the configuration used for an export. This report is essential for traceability and transparency, especially when responding to compliance inquiries.

Use this report to:

- Confirm which items were included (indexed, partially indexed, or both).
- Verify the export format (PST or .msg) and package size settings.
- Check whether decryption was enabled for Exchange or SharePoint content.
- Validate that the correct document version settings were applied.

Understanding location reports

Identifying which locations have results

The Locations CSV report provides a breakdown of every content location targeted during a search, export, or add to review set process. Each row represents a location and includes the item count and size.

What to look at in the location report:

- **Count column**: Shows the number of responsive items found in each location. Locations with a count of zero have no search hits for the query used.
- **Size column**: Shows the total size of responsive items from the location, in bytes.
- **Location Subtype column**: Distinguishes between different mailbox and site subtypes. A single user's SMTP address might appear in multiple rows if they have different mailbox subtypes. For the full list of subtype values, see Locations CSV report.
- **Status column**: Shows whether the location was successfully processed. If a location failed, the **ErrorWarning** field includes the reason.

Understanding location-level errors

Location-level errors in the Locations CSV report indicate that the process couldn't access a specific mailbox or site. Common errors include:

Expand table

Error	What it means	What to do
Permission or access issue	The process can't access the mailbox or site. The user might be unlicensed, the site might be locked, or the searching account lacks permissions.	Verify the user's license status and that the site isn't locked. Confirm that the eDiscovery admin or manager role has the required permissions, and then rerun the process.
Timeout	The location took too long to process. This error can happen with very large mailboxes or sites.	Rerun the process. If the timeout persists, narrow the search query or split the search to target fewer locations per run.
Location not found	The mailbox or site doesn't exist. The user might be deleted or the site URL might be incorrect.	Verify the user account or site URL in the Microsoft 365 admin center. If the user was deleted, use an inactive mailbox or add the OneDrive URL directly as a SharePoint location.

If a location has errors, it doesn't mean the entire process failed. Other locations might be processed successfully. Review the Locations CSV to identify which specific locations need attention.

Understanding hold reports

Interpreting hold distribution status

When you create or update a hold, the process report for the *Applying or updating hold* process includes a Locations CSV report that shows the status of each location in the hold.

What to look at in the hold report:

- **Status column**: Shows whether the hold was successfully applied to each location.
- **ErrorWarning column**: If a location failed, this field contains the specific error. Common hold distribution errors include:

Expand table

Error	What it means	What to do
PolicySyncTimeout / PolicyNotifyError	Transient errors. The hold took too long to apply.	Rerun the hold distribution. These errors usually resolve on retry.
FailedToOpenContainer	The mailbox or site can't be accessed. The user might not have a mailbox, or the site doesn't exist.	Verify the user account status in the Microsoft 365 admin center and confirm the mailbox or site is active.
SiteInReadonlyOrNotAccessible / SiteOutOfQuota	The SharePoint site or OneDrive is in a read-only state or has no storage quota left.	Contact your SharePoint admin to unlock the site or increase the storage quota, then rerun the hold.
RecipientTypeNotAllowed	The hold was attempted on an ineligible object, such as a distribution group. Only user mailboxes, group mailboxes, and SharePoint or OneDrive sites can be held.	Remove the ineligible object and add the individual user mailboxes or sites instead.

Understanding the consumption summary

The **Consumption summary** section of the Summary CSV report breaks down the data volume by content type. This information helps you understand storage and billing implications for processes that add data to a review set or export data.

> **Note**
>
> If you don't have an active pay-as-you-go subscription, data isn't exported or added to a review set. For more information, see Enable Microsoft Purview pay-as-you-go features.

Last updated on 04/01/2026

Use Process reports in eDiscovery

Process report for the organization

eDiscovery includes Process reports that list all processes in your organization that count towards case process concurrency and daily limits in eDiscovery. *eDiscovery administrators* and *eDiscovery Managers* can access this report on the **Dashboards** > **Process report**. Use this report to see the process activities across all activities and eDiscovery components. *eDiscovery Managers* can only view processes for the cases to which they're assigned.

> **Important**
>
> You can't delete or remove processes listed on the **Process manager** page. Long-running processes automatically time out after seven days.

Process report for items

eDiscovery also includes a **Process manager** that lists all processes scoped to the current eDiscovery item and page. For example, if you're on the **Search** page, selecting **Process manager** lists all processes relevant to the current search. If you're in a case, selecting **Process manager** lists all processes relevant to the current case. Each process in the Process manager list contains a detailed report .zip file that contains detailed information about the process once the process completes. You can download the process report .zip file for more information about when the process ran, the applied settings, the query used, and if the process had any item-level or location-level errors.

> **Important**
>
> You can't delete or remove processes listed on the **Process manager** page. Long-running processes automatically time out after seven days.

To learn more about Process reports in eDiscovery, watch the following video:

https://learn-video.azurefd.net/vod/player?id=72ab2c79-cbf6-44fe-b365-8a41cda94216&locale=en-us&embedUrl=%2Fpurview%2Fediscprocess-report-use

Process types and descriptions

User actions trigger these processes when managing searches, review sets, and holds.

Expand table

Process type	Area	Description
Add items to another review set	Review set	A user adds documents from one review set to a different review set in the same case. For more information, see Add data to a review set from another review set.
Add to review	Review set	A user adds search results to a review set or adds items into the review set from another review set. The process copies the items to an Azure Storage location and then reindexes those items in the Azure Storage location. This new index is used when querying and analyzing items in the data set.
Applying or updating hold	Hold	A user places data sources on hold or updates a hold.
Export	Search Review set	A user exports documents from a search or a review set. When the export process completes, they can download the exported data to a local computer.
Generate sample	Search	After a user creates and runs a new search estimate (or reruns an existing search estimate), the search tool prepares a sample subset of items (that match the search query) that you can preview. Previewing search results helps you determine the effectiveness of the search.
Generate statistics	Search	After a user creates and runs or reruns a search estimate, the search tool searches the index for items that match the search query and prepares an estimate that includes the number and total size of all items by the search, the number of data sources searched, and other relevant pattern to the search hits like keywords, sensitive information types, and more.
Import external data	Review set	A user uploads external data to a review set. The data is also indexed during this process. For more information, see Import external data into a review set.
Redact	Review set	After a user annotates a document in a review set and redacts a portion of it, they can choose to convert the redacted document to a PDF file. This ensures that the redacted portion won't be visible if the document is exported for presentation. For more information, see View documents in a review set.
Run analytics	Search	A user analyzes data in a review set by running eDiscovery analytics tools such as near duplicate detection, email threading analysis, and themes analysis.
Tag	Review set	This process triggers when a user selects **Start tagging process** in the **Tagging panel** when reviewing documents in a review set. A user can start this process

Process type	Area	Description
		after tagging documents in a review set and then bulk-selecting them in the view document panel.

Process status

The following table describes the different status states for processes:

> **Note**
>
> You can cancel only **Generate statistics** and **generate sample** processes.

Expand table

Status	Description
Abandoned	The process was automatically stopped. This stoppage might be due to excessive processing time or other reasons.
Canceled	The user canceled the process.
Complete	The process completed successfully.
Failed	The process encountered an error and didn't complete. This error might be caused by a wrongly formatted query.
In progress	The process is in progress. You can monitor the progress of the process in the **Process manager** for the area.

Download report

All processes support the ability to download a packaged report. The packaged report name is `Reports-*caseName-EntityName-ProcessName-timestamp*.zip`, with `EntityName` being the user-given name to the search or hold. Depending on the process, the report contains different `.csv` files.

- **Summary**: Tracks the time the process started, when it ended, the number of total items or locations, and the user who submitted the process request. It also contains the search query used and if a compliance boundary is applied. For *Add to review set* and *Export* from search processes, the summary report contains the summary of factors affecting the total item count added to review set or exported. The report name is `Summary_*the date/time of the report*.csv`.

- **Settings**: Contains the enumerated settings parameter for the process and values. The report name is `Settings_*the date/time of the report*`.
- **Statistics**: Contains all statistics details for the process, including all categories (if the setting was selected during process submission) such as sensitive information types, data type, and communication participants. The report name is `Statistics_*the date/time of the report*`.
- **Locations**: Tracks all data sources and associated locations scoped for the process. Includes the user/group entity name, location (mailbox/site URL), and count returned for that location. Also includes the status of the location (success/error and error detail). The report name is `Locations_*the date/time of the report*`.
- **Items**: Track all items processed, including information such as item ID, location of the item, subject/title of the item, item class/type, and success/error status. The report name is `Items_*the date/time of the report*`.

The following table shows the process types and available reports:

Expand table

Process type	Summary	Settings	Statistics	Location	Items
Add to review	✔	✔		✔	✔
Apply hold/rerun policy	✔			✔	
Export (review set)	✔	✔		✔	✔
Export (search)	✔	✔		✔	✔
Generate sample	✔	✔		✔	✔
Generate statistics	✔	✔	✔	✔	
Redact	✔	✔			✔
Tag	✔	✔			

> **Note**
>
> When running an export process using the *report only* option, all string fields in the report are limited to the first 255 characters. This means some items with long compound path

may be truncated. If full details are needed, use the *Export items with items report* option.

Summary CSV report

All reporting packages contain a `Summary-*the date/time of the report*.csv` file. This file contains the following information:

Overview

This section of the report includes the following details:

- **Job ID**: A unique identifier assigned to each eDiscovery process for support reference purposes.
- **Case name**: The name assigned to the eDiscovery case.
- **Search name**: The name given to a specific search within an eDiscovery case when the process runs. This detail only applies to processes like *Generate statistics*, *Generated sample*, *Add to review set*, and *Export* from a search.
- **Review set name**: The name of the review set where search results are added when the process runs. This detail only applies to the *Add to review set* and *Export from a review set* processes.
- **Export name**: The name given to the export. This detail only applies to the *Add to review set* and *Export from a review set* processes.
- **IsRetry**: Indicates if the current process is triggered by the *Retry failed locations* operation. If the value is *Yes*, it contains the aggregated hit count and the location reports detail the failed locations. This detail only applies to the *Generate statistics* process.
- **Query**: The query conditions for the applicable process types.
- **StartTime**: The date and time when the report generation started.
- **EndTime**: The date and time when the report generation completed.
- **SubmittedBySmtp**: The SMTP address for the user creating the report.
- **Security filter applied**: Indicates if security filters were applied during the process run. *Yes* indicates that filters were used to limit results based on specified criteria set by an admin in PowerShell.
- **Location restriction**: The filters applied to limit the search to specific locations. To review which mailboxes are filtered, use the provided query with the Get-Mailbox or Get-Recipient cmdlets and the *-Filter* attribute. For example, if the filter attribute is *Department -eq 'marketing'*, retrieve the relevant mailboxes with *Get-Recipient -Filter "Department -eq 'marketing'*. The output lists all mailboxes allowed to be searched.

- **Content restriction**: The query used to filter content during the search. Content matching this query is included in the search results. The provided query is available to perform KeyQL searches.

Search results

Summary of number of items that matched your search query, including partially indexed items or items where advanced indexing was performed (depending on what setting you selected). This section of the report includes the following details:

- **Indexed items**: Number of items matching the query that were natively indexed by Exchange and SharePoint.
- **Partially indexed items**: Number of partially indexed items that might not match the query. Included if you chose to include partially indexed items.
- **Advanced Indexed Items**: Number of items matching the query if you chose to perform advanced indexing on partially indexed items.

Consumption summary

This section of the report includes the following details:

- **Microsoft 365 content data**: Total size of data for all Microsoft 365 data.
- **Microsoft 365 metadata**: Total size of metadata for all Microsoft 365 data and Microsoft 365 Copilot data.
- **Non-Microsoft 365 AI content data**: Total size of data for all non-Microsoft 365 AI data. Includes data from Microsoft Fabric Copilot, Microsoft Security Copilot, and other AI data from enterprise and cloud applications.
- **Non-Microsoft 365 AI metadata**: Total size of metadata for all non-Microsoft 365 AI data.

> ⓘ **Note**
>
> For more information about the consumption summary and billing implications, see Understanding the consumption summary.

Error

This section of the report includes the following details:

- **Retrieval exceptions**: Number of items that weren't exported or added to review set due to exceptions such as empty files, access time-outs, and similar issues.

Warning

- **Items with processing errors**: Number of items that experienced processing errors but were still exported or added to the review set. These errors might be caused by unsupported file types, decryption problems, and similar issues.

Totals

This section of the report includes the following details:

- **Total items collected**: Total number of items exported or added to the review set. This number factors in items from estimated results, the settings that might increase or decrease the number of items retrieved, and items that weren't collected due to errors.

Information

> **Note**
>
> The **Information** section of this report shows how the item count can increase or decrease based on the settings you choose. The calculation process uses these settings and updates the search estimate count, so the counts don't add up exactly in most cases. This information gives you a sense of the changes and the reasons behind them.

This section of the report includes the following details:

- **Duplicates in review set (skipped)**: Items that are skipped because they already exist in the review set.
- **Cloud attachments**: Number of cloud attachments in email messages and Teams conversations that come from links shared through SharePoint or OneDrive. Maps to the setting *Access links(cloud attachments) in messages*.
- **Cloud attachments versions**: Number of cloud attachment versions that are collected depending on whether you choose to include more than just the latest version in your query.
- **Cloud attachments at time of sharing**: Number of cloud attachment versions that represent the original version shared from SharePoint or OneDrive. This count applies only to cloud attachments that have a retention label applied, which retains a copy of the file at the time of sharing.
- **Contextual conversation**: Number of contextual chat messages that are collected along with the message that matches your query. This count shows the number of additional

messages collected before and after the message with hit. This count corresponds to the setting "Include Teams and Viva Engage conversations".

- **Teams conversations consolidated into transcripts**: Teams chat conversations that are converted to HTML transcript files. This conversion results in many Teams chat messages being consolidated into a single HTML transcript. This count corresponds to the setting "Organize conversations into HTML transcripts".
- **SharePoint file versions**: Number of SharePoint file versions that are collected depending on whether you choose to include more than just the latest version.
- **SharePoint list items collected as .csv files**: Matches from the same SharePoint list are exported or added to review set as a single SharePoint list csv item. This count corresponds to the setting "Collect files attached to SharePoint lists and their child items."
- **List attachments**: Number of list attachments collected. This count corresponds to the setting *Include list attachments* when collecting files attached to SharePoint lists.
- **Items in SharePoint folders**: Number of items expanded from SharePoint folders that match your query. This count corresponds to the setting *Collect all items (even if they don't match search query)*.
- **Items extracted from parent items**: Number of embedded or attached items that the process extracts from parent items. For example, this count might include attachments or embedded images from an email that matches your query.

Settings CSV report

The settings report helps you validate and interpret the configuration of your submitted processes. The downloadable report package includes this report. It captures the specific options you select during the submission of a process, such as **Generate statistics**, **Add to review set**, and **Export**.

These settings directly influence the scope and format of the results. Some examples of these settings include:

- **Include document versions**: Determines the number of versions of a document to include.
- **Organize conversations into HTML transcripts**: Affects how Teams conversations are threaded and presented.
- **Access links in message**: Controls whether cloud attachments from message links are included.
- **Various export formats and options**: Influences how exported content is organized and packaged.

- **Decryption settings**: Indicates whether the specific process has associated Exchange or SharePoint decryption capability enabled.
- **Location optimization**: Indicates if location optimization is applied for the search *Export* or *Add to review set* process. If optimization is applied, the relevant *Generate statistics* process ID used.
- **Sampling mode**: Indicates the sampling method used for the *Add to review set* process. Values include *All items*, *Percentage*, or *Confidence*.
- **Sample percentage**: The percentage value configured when using percentage-based sampling. Present only for the *Percentage* sampling mode.
- **Confidence level**: The confidence level configured when using confidence-based sampling. Present only for the *Confidence* sampling mode.
- **Confidence interval**: The confidence interval configured when using confidence-based sampling. Present only for the *Confidence* sampling mode.

For more information about how these settings affect search result counts, see Understanding the impact of search settings on results. For guidance on using the Settings CSV for traceability and compliance, see Understanding export settings and traceability.

Statistics CSV report

This report provides detailed information for all the patterns derived from search hits when you complete the *Generate statistics* process. The report contains information on the top communication participants, the top item type, the top sensitive information type, and more.

The report includes the following details:

- **DataSourceName**: A list of data source names associated with the process. This information indicates the name of the people or groups.
- **DateSourceType:** Type of data source. For example, *Group* or *TeamsGroup*.
- **Location**: The location information. For example, the URL of the site or SMTP address of the mailbox.
- **LocationType**: Type of location. For example, a mailbox or site.
- **LocationSubType**: The subtype of the location type. If the location type is *Mailbox*, the subtype might have *PrimaryMailbox*, *SystemMailbox*, and other subtypes. If the location type is *Site*, it might have *OneDriveSite* and other subtypes.
- **OriginalStateName**: Native version of StatName.
- **StatName**: Name of the StatType. For example, for StatType *Keywords*, possible values are the keywords you entered in your query.

- **StatType:** The type of statistics information. For example, *Keywords*, *Communication participants*, and more.
- **Value**: The value or count associated with the statistics in the current data source.

Items CSV report

This report provides a detailed inventory of all items processed during an eDiscovery process, such as using the **Add to review set** or **Export** actions. The system automatically generates this report and includes it with the output to support auditing, traceability, and downstream analysis. This report is useful for users who need to validate the scope and accuracy of collected or exported content.

Each row in the file corresponds to a single item and includes the following metadata and more:

- **Date source**: Identifies the originating location of the item (mailbox, SharePoint site, Teams chat).
- **Compound path**: The full path to the item in its source container.
- **Subject/Title**: The subject line (for emails) or title (for documents) of the item.
- **Date**: The timestamp associated with the item, such as sent or last modified date.
- **Added by**: Shows how the item was included in the job:
  - *IndexedQuery*: The item is indexed and matches the search query.
  - *UnindexedQuery*: The item is included based on its partially indexed content and user specified setting.
- **Status**: Shows whether the item was successfully retrieved and processed. If the item failed, the **ErrorWarning** field includes the failure reason. For common item-level errors, see Why certain items weren't exported.
- **Other metadata**: Might include the file extension, content type, workload, and other relevant attributes depending on the source and process type.

Locations CSV report

This report provides an overview of the locations targeted during an eDiscovery process, such as **Search**, **Add to review set**, or **Export**. This report helps you understand the scope of the process and assess the distribution and relevance of search results across different locations.

This report can help with:

- **Heat mapping**: Identifying which locations yielded the most relevant content and which yielded none, helping you visualize the "hot spots" of data relevance.
- **Query refinement**: Understanding which sources you might need to exclude or more precisely target.
- **Audit and validation**: Verifying that the intended locations were included and assessing the effectiveness of the search scope.

By analyzing the `location.csv` report, you can make informed decisions about refining your search strategy and optimizing the scope of future eDiscovery workflows.

> ⓘ **Note**
>
> If multiple mailboxes share the same SMTP address, the location report might show several rows. This situation occurs because a single SMTP address might include different mailbox subtypes, such as primary, system, or archive. The row count in the CSV report might exceed the count shown on the location card in the **Statistics** view. The **Statistics** view aggregates all mailboxes under the same SMTP address without distinguishing subtypes. For a more detailed breakdown of hits by sublocation, use the CSV report instead of the **Statistics** view. SharePoint and OneDrive locations can also have archive sites and are counted as separate locations in the report.

Each row in the report represents a unique content location and includes the following key details:

- **Count**: The number of items in the location that matched the search criteria. This doesn't apply to the *Applying or updating hold* process.
- **Location ID**: Site ID for SharePoint or mailbox ID for mailboxes.
- **Location Subtype**: Specifies the type of location, such as *PrimaryMailbox*, *SystemMailbox*, *ArchiveMailbox*, or *OneDriveSite*.
- **Location**: The full URL or SMTP address of the location. For organization-wide searches, locations are empty for sites without hits. If needed, use the *LocationID* to identify the site URL.
- **Size**: The total size of responsive items from the location in bytes. This doesn't apply to *Applying or updating hold* process.
- **Status**: Indicates whether the location was successfully processed. If a location failed to be searched or exported, the **ErrorWarning** field includes the failure reason. For common

location errors and remediation steps, see Understanding location-level errors.

Last updated on 04/01/2026

Use Process managers in eDiscovery

Each primary area in eDiscovery includes a process manager. These managers let you view processes and related information for a specific area. You can access process managers by selecting **Process manager** in each of the following areas:

- Cases
- Searches
- Review sets
- Holds

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Process manager information

Each area's process manager shows information about processes performed in that area. For example, if you export data from the review set, an *Export* process item appears in the process manager list for the review set, but not in the process manager lists for searches and holds. Each process manager contains the following information for its area:

- **Process type**: The type of process.
- **Status**: The status of the process.
- **Export name**: The name of the export. For nonexport process types, this value is blank.
- **Created**: The date and time the process was created.
- **Completed**: The date and time the process was completed.
- **Duration**: The duration of the process.
- **Created by**: The user that created the process.

To customize the columns display for the process manager, select **Customize columns** to choose the columns to display or drag and drop the columns in the list to reorder. To download the list of processes and the column information, select **Download list** to create a .csv file containing this information.

Grouping processes

When you view a large list of processes for an area, it's often helpful to group processes. Select **Group** to group review sets by the following parameters:

- **None**: Processes are ungrouped in the Process manager.
- **Process type**: Processes grouped by the *Process type* values.
- **Created by**: Processes grouped by the *Created by* values.

Process details

Select a process listed in the process manager for more information in the following areas:

- **Overview** tab: Lists a summary of information for the process. Includes a calculation for the time remaining to completion for in-progress processes. Select **Download report** to combine all **Overview** information into a single .zip file. This file contains the files included in the Download report guidance.
- **Settings** tab: Summarizes the **Search & analytics** setting options selected for the case.

Progress bar

When you select a process in the Process manager, a progress bar appears that shows details about the current state of the process. For long-running processes, there are three phases:

- **Assessment phase**: In this phase, the progress is submitted and backend services assess the scope of the process. This assessment includes examining the input conditions based on data sources, query, and relevant process settings. The service analyzes the scope of the work and identifies the number of locations and number of items for processing. During this phase, a completion estimate isn't available and a flashing progress bar is displayed. The progress bar displays the status of the assessment while the scope is calculated. When all locations and items are identified, the process transitions to the next phase.
- **In progress phase**: In this phase, the progress bar displays the current status, number of locations, and number of items for the process. An estimated completion and current elapsed time are displayed in hours, minutes, and seconds.
- **Catch-up phase (optional)**: If there are locations or items with errors, processing is retried. The progress of the retries is displayed.

Last updated on 04/02/2026

Limits in eDiscovery

This article describes the limits in Microsoft Purview eDiscovery.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Case limits

The following table lists the limits for cases in eDiscovery.

Expand table

Description of limit	eDiscovery feature support	Premium eDiscovery feature support
Maximum number of cases for an organization.	10,000	50,000

Case and review set limits

The following table lists the limits for cases and review sets in eDiscovery with premium feature support.

> **Tip**
>
> Want to try premium eDiscovery features? See the subscription requirements for Microsoft 365 Enterprise E5 licensing.

Expand table

Description of limit	Premium eDiscovery feature support
Total number of documents that you can add to a case (for all review sets in a case).	40 million

Description of limit	Premium eDiscovery feature support
Total file size per load set. This limit includes loading non-Office 365 content into a review set.	1 TB
Maximum size per review set.	1 TB
Maximum number of load sets per case.	200
Maximum number of review sets per case.	100
Maximum number of tag groups per case.	1,000
Maximum number of unique tags per case.	1,000
Maximum number of tagged documents per case.	Unlimited

Export limits

> **Note**
>
> When exporting large individual items (5 GB or greater), configure the export job to include only one large item per export. This configuration reduces the likelihood of performance-related issues during the export.

The following table lists the limits when exporting the results of a search.

Expand table

Description of limit	eDiscovery feature support	Premium eDiscovery feature support
Export report total items for one report	1 million	1 million
Export report total locations for one report	500,000	500,000
Export results maximum number locations for one export	100,000	100,000
Export results maximum number of items for one export	500,000	500,000
Maximum amount of exportable data from a single search.	2 TB[1]	5 TB
Maximum an organization can export in a single day.[2]	2 TB	5 TB

Description of limit	eDiscovery feature support	Premium eDiscovery feature support
Maximum number of exports or reports displayed in Content Search or eDiscovery cases.	1,000	2,000
Maximum size of individual file supported for export	5 GB	5 GB
Maximum size of PST file that can be exported.	10 GB[3]	10 GB
Maximum size per export package.	40 GB	40 GB
Rate at which content from mailboxes are exported	2 GB per hour per mailbox[4]	2 GB per hour per mailbox[4]

> **Note**
>
> [1]If the search results are larger than 200 GB, consider using date ranges or other types of filters to decrease the total size of the search results.
>
> [2] The system resets this limit daily at 12:00 AM UTC.
>
> [3] The export process automatically splits search results larger than 10 GB into multiple PST files, even if you select the option to export results into a single PST.
>
> [4] If an export runs for seven days without completing, the system automatically cancels it. To prevent this issue, divide your export into smaller batches.

Hold limits

> **Tip**
>
> For better performance, limit each case to no more than 1,000 searches, holds, or exports. To spread this activity across more cases, consider updates to your organization processes and workflows.

The following table lists the limits for holds associated with an eDiscovery case.

Expand table

Description of limit	eDiscovery feature support	Premium eDiscovery feature support
Maximum number of hold policies for an organization.	10,000	20,000
Maximum number of hold policies within a single case	100	200
Maximum number of mailboxes expanded from a single distribution list when used as a hold data source.[1]	100	100
Maximum number of mailboxes in a single case hold. This limit includes the combined total of user mailboxes, and the mailboxes associated with Microsoft 365 Groups, Microsoft Teams, and Viva Engage Groups.	1,000	2,000
Maximum number of policies applied to a specific location	100	100
Maximum number of sites in a single case hold. This limit includes the combined total of OneDrive sites, SharePoint sites, and the sites associated with Microsoft 365 Groups, Microsoft Teams, and Viva Engage Groups.[2]	1,000	2,000

> **Note**
>
> [1] You can't add distribution lists with more than 100 members directly as a hold data source. Expand the distribution list membership (for example, in Outlook) and then use the **bulk import data sources** feature to import the individual user mailboxes as data sources for the hold. You can also split the distribution list into smaller groups of 100 or fewer members before adding them.
>
> [2] Higher premium limits (for example, 2,000 sites) apply only when the case has premium eDiscovery features enabled. Check that premium features are enabled in **case settings**. Use separate hold policies for Exchange mailboxes and for SharePoint and OneDrive sites instead of combining them in one policy.

To view a list of more than 1,000 cases, holds, searches, or exports, use the corresponding Security & Compliance PowerShell cmdlet:

- Get-ComplianceCase
- Get-CaseHoldPolicy
- Get-ComplianceSearch
- Get-ComplianceSearchAction

Indexing limits

The following table lists the indexing limits in eDiscovery with premium feature support.

> **Tip**
>
> Want to try premium eDiscovery features? See the subscription requirements for Microsoft 365 Enterprise E5 licensing.

Expand table

Description of limit	Premium eDiscovery feature support
Maximum number of characters extracted from a single file.[1]	10 million
Maximum size of a single file.[1]	150 MB
Maximum depth of embedded items in a document.[1]	25
Maximum size of files processed by Optical Character Recognition (OCR).[1]	24 MB
Maximum Advanced indexing throughput	2 GB per hour

> **Note**
>
> [1] If an item exceeds a single file limit, the process returns an error.

Indexing limits for email messages

The following table describes the indexing limits that might cause an email message to appear as an unindexed item or a partially indexed item in the results of a search.

Expand table

Indexing limit	Maximum value	Description
Maximum attachment size [1]	150 MB	The maximum size of an email attachment that the indexing service parses. The indexing service doesn't parse any attachment that's larger than this limit. It marks the message with the attachment as partially indexed.

Indexing limit	Maximum value	Description
Maximum number of attachments	250	The maximum number of files attached to an email message that the indexing service parses. If a message has more than 250 attachments, the indexing service parses and indexes the first 250 attachments. It marks the message as partially indexed because it didn't parse the extra attachments.
Maximum attachment depth	30	The maximum number of nested attachments that the indexing service parses. For example, if an email message has another message attached to it and the attached message has an attached Word document, the indexing service indexes the Word document and the attached message. This behavior continues for up to 30 nested attachments.
Maximum number of attached images	0	The parser skips an image that's attached to an email message and doesn't index it.
Maximum time spent parsing an item	30 seconds	The indexing service spends a maximum of 30 seconds parsing an item for indexing. If the parsing time exceeds 30 seconds, the indexing service marks the item as partially indexed.
Maximum parser output	2 million characters	The maximum amount of text output from the parser that's indexed. For example, if the parser extracted 8 million characters from a document, the indexing service indexes only the first 2 million characters.
Maximum annotation tokens	2 million	When the indexing service indexes an email message, it annotates each word with different processing instructions that specify how to index that word. Each set of processing instructions is called an annotation token. To maintain the quality of service in Microsoft 365, there's a limit of 2 million annotation tokens for an email message.
Maximum body size in index	67 million characters	The total number of characters in the body of an email message and all its attachments. When the indexing service indexes an email message, it concatenates all text in the body of the message and in all attachments into a single string. The maximum size of this string that is indexed is 67 million characters.
Maximum unique tokens in body	1 million	Tokens result from extracting text from content, removing punctuation and spaces, and then dividing it into words (called tokens) that the indexing service stores in the index. For example, the phrase `"cat, mouse, bird, dog, dog"` contains five tokens, but only four of these are unique tokens. There's a limit of 1 million unique tokens per email message, which helps prevent the index from getting too large with random tokens.

Note

[1] Parsing is the process where the indexing service extracts text from the attachment, removes unnecessary characters like punctuation and spaces, and then divides the text into words (in a process called tokenization) that it stores in the index.

Process limits

The following table lists the process limits in eDiscovery.

Expand table

Description of limit	eDiscovery feature support	Premium eDiscovery feature support
Maximum number of concurrent processes in your organization.	50	200
Maximum number of concurrent processes a single user can start at one time.	25	50
Maximum number of concurrent tenant-wide processes (for example, tenant-wide searches) in your organization.	5	50
Maximum number of concurrent tenant-wide processes (for example, tenant-wide searches) that a single user can start at one time.	5	25
Maximum number of jobs per day in your organization. [1]	500	1,000
Maximum duration the process runs before it's canceled.	7 days	7 days

Note

[1]This limit resets daily at 00:00 UTC.

Purge limits

The following table lists the purge limits in eDiscovery.

Expand table

Description of limit	eDiscovery feature support	Premium eDiscovery feature support
Items per location in Exchange	10	100

Description of limit	eDiscovery feature support	Premium eDiscovery feature support
Items per location in Teams	N/A	100
Total locations per purge	50,000	100,000

Review set download limits

The following table lists the review set download limits in eDiscovery with premium feature support.

> **Tip**
>
> Want to try premium eDiscovery features? See the subscription requirements for Microsoft 365 Enterprise E5 licensing.

Expand table

Description of limit	Premium eDiscovery feature support
Total file size or maximum number of documents downloaded from a review set.[1]	3 MB or 50 documents

> **Note**
>
> [1] This limit applies to downloading selected documents from a review set. It doesn't apply to exported documents from a review set. For more information about exporting documents from a review set, see Export items from a review set in eDiscovery.

Review set export limits

The following table lists the limits when exporting documents from a review set in eDiscovery with premium feature support.

> **Tip**
>
> Want to try premium eDiscovery features? See the subscription requirements for Microsoft 365 Enterprise E5 licensing.

Expand table

Description of limit	Premium eDiscovery feature support
Maximum size of a single export.	5 million documents or 500 GB, whichever is smaller.

Review set limits

The following table lists the review set limits in eDiscovery with premium feature support.

> **Tip**
>
> Want to try premium eDiscovery features? See the subscription requirements for Microsoft 365 Enterprise E5 licensing.

Expand table

Description of limit	Premium eDiscovery feature support
Maximum number of items displayed per page in a review set.[1]	1,000
Add to review set processing	To optimize performance during the *Add to review set* process, the system allocates up to one minute to process each item. If processing isn't completed within one minute, the item might be added to the review set without fully extracting its child items.[2]

> **Note**
>
> [1]Use default or custom filters to adjust the displayed items in a review set as needed.
>
> [2]Use the items.csv process report to review any items that encountered this processing limit after the *Add to review set* process is completed.

Search limits

The following table lists the search limits for Exchange mailboxes and SharePoint sites in eDiscovery with premium feature support.

Expand table

Description of limit	eDiscovery feature support	Premium eDiscovery feature support
Maximum recommended number of locations specified for a single targeted search.[1]	1,000	1,000
Maximum number of locations for a single tenant-wide search, export from search, or add to review set process (organization-wide "All people and groups" source).[1]	500,000	500,000
Maximum number of Sensitive Information Types per search query.[2]	20	20
Maximum number of participants per search query.[2]	25	50
Maximum number of characters for a search query (including operators and conditions).[3, 5]	10,000	10,000
Minimum number of alpha characters for prefix wildcards; for example, `one*` or `set*`.	3	3
Maximum variants returned when using prefix wildcard to search for an exact phrase or when using a prefix wildcard and the NEAR Boolean operator.[4]	10,000	10,000
Maximum number of operators in a query.[5]	100	100
Maximum number of keyword segments in a query for tenant wide search. [6]	5	5
Maximum number of keyword segments in a query for targeted search. [6]	20	100

Note

[1] The 1,000 location limit applies only to targeted searches you manually specify individual mailboxes or sites. It includes the combined total of user mailboxes, and the mailboxes associated with Microsoft 365 Groups, Microsoft Teams, and Viva Engage Groups. Tenant-wide processes that use the organization-wide "All people and groups" source support up to 500,000 locations and aren't subject to the 1,000 targeted search limit. The 500,000 tenant-wide location limit applies to tenant-wide search, tenant-wide export from search, and tenant-wide add to review set processes. [2] Including more than 20 Sensitive Information Types or 25 participants might result in search query timeouts. [3] When searching SharePoint and OneDrive locations, characters in the URLs of searched sites count against this limit. The total number of characters consists of:

- All characters in both the *Users* and *Filters* fields.
- All search permissions filters that apply to the user.
- The characters from any location properties in the search, including *ExchangeLocation*, *PublicFolderLocation*, *SharePointLocation*, *ExchangeLocationExclusion*, *PublicFolderLocationExclusion*, *SharePointLocationExclusion*, and *OneDriveLocationExclusion*. For example, including all SharePoint sites and OneDrive accounts in the search counts as six characters, as the word `ALL` appears for both the *SharePointLocation* and *OneDriveLocation* field.

[4] For non-phrase queries (a keyword value that doesn't use double quotation marks) the system uses a special prefix index. This index tells the system that a word occurs in a document, but not where it occurs in the document. To do a phrase query (a keyword value with double quotation marks), the system needs to compare the position within the document for the words in the phrase. This comparison means that the system can't use the prefix index for phrase queries. In this case, the system internally expands the query with all possible words that the prefix expands to. For example, `time*` can expand to "time OR timer OR times OR timex OR time boxed OR ...". The limit of 10,000 is the maximum number of variants the word can expand to, not the number of documents matching the query. There's no upper limit for non-phrase terms. [5] A search query can include a maximum of 100 operators and up to 10,000 characters. An operator is Boolean operator such as AND, OR, or NEAR. The limit of operators and characters apply to the combined total of operator and characters in the user input search query and any search permission filters that are applied to the user running the search. As a result, users with multiple or complex permission filters, especially filters applied to the same users or groups, are more likely to reach this limit. To reduce the risk of exceeding the maximum number of operator, keep search permission filters to the minimum required to meet your organization's business needs. For more information, see **Configure search permission filtering in eDiscovery**. [6] Keyword segment refers to each individual keyword or keyword phrase entered as a separate row in the Keyword condition of the condition builder. It doesn't refer to the number of words or characters within a keyword or phrase. Each keyword segment is executed as a separate search operation against the selected locations. As a result, using a large number of keyword segments can increase search duration or cause the search to time out or fail. If your scenario requires more keyword segments than supported in a single search, split the keywords across multiple searches and run them separately.

For more information about specific limits, see the following articles:

- Exchange
- SharePoint

Search sample limit

The search sample is a best effort and depends on the following factors. If you need essential data, see export the data or add data to review set.

Expand table

Description of limit	Limit
Maximum item size for sample	10 MB
Time spent per location during sampling	Less than 30 seconds

Export search and export as report limit

The following table lists limits for the **Export as report** and the **export search** results:

Expand table

Description of limit	Limit
Export items report only CSV report string field character limit	255
Exported results (items)	500,000
Exported results (locations)	100,000
Search export expiry	14 days from creation
Total items	1 million
Total locations	500,000

Search times

Microsoft collects performance information for searches run by all organizations. While the complexity of the search query can impact search times, the biggest factor that affects how long searches take is the number of mailboxes searched.

Although Microsoft doesn't provide a service level agreement for search times, the following table lists average search times for collection searches based on the number of mailboxes

included in the search.

Expand table

Number of mailboxes	Average search time
100	30 seconds
1,000	45 seconds
10,000	4 minutes
25,000	10 minutes
50,000	20 minutes
100,000	25 minutes

Themes limits

The limits described in this section relate to manually defining the maximum number of themes in settings for a case in eDiscovery with premium feature support.

> **Tip**
>
> Want to try premium eDiscovery features? See the subscription requirements for Microsoft 365 Enterprise E5 licensing.

Expand table

Description of limit	Premium eDiscovery feature support
Minimum number for the maximum number of defined themes.	3
Maximum number of defined themes.	10,000

Viewer limits

Expand table

Description of limit	Limit
Maximum size of an Excel file you can view in the native viewer.	4 MB

Tenant region or environment change limits

The same Microsoft 365 region and service environment store eDiscovery case data (cases, searches, holds, review sets, and exports) where you originally created it. This data doesn't move with your tenant if you later relocate the tenant to a different region or to a different service environment.

> **Important**
>
> If you move your Microsoft 365 tenant between regions (for example, through Advanced Data Residency), or migrate it across service environments (for example, from Microsoft 365 commercial to a Microsoft 365 Government environment such as GCC, GCC High, or DoD), you can't access eDiscovery cases and their associated data that you created **before** the move from the destination tenant. Cases created **after** the move work normally in the new location.

There's no in-place migration path for eDiscovery case data across regions or service environments. To avoid losing access to in-flight investigations, complete the following actions before the move:

- Complete or close any active investigations where possible.
- Export the results of any searches you want to retain.
- Export any review sets you want to retain.
- Document the membership of any case holds you need to recreate in the destination tenant.

After the move, recreate the holds, searches, and cases you still need in the destination tenant.

Note: The author created this article with assistance from AI. Learn more

---

Last updated on 06/08/2026

Supported file types in eDiscovery

Microsoft Purview eDiscovery supports many file types at many different levels. The following tables list the currently supported file types.

These tables indicate if a file type is supported for text extraction and Optical Character Recognition (OCR) text extraction for image files, if it's viewable in the native viewer, and if it's supported in the Annotate viewer in eDiscovery.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

Archive / Container

Expand table

Mime type	File identification	Metadata extraction	Container extraction	Possible Extensions
application/x-7z-compressed	Yes	Yes	Yes	.7z
application/x-rar-compressed	Yes	Yes	Yes	.rar[1]
application/x-tar	Yes	Yes	Yes	.tar
application/zip	Yes	Yes	Yes	.zip

> **Note**
>
> [1] Version 5 .rar files aren't supported.

Audio / Video

Expand table

Mime type	File identification	Metadata extraction	Text extraction	Native viewer	Annotate viewer	Possible Extensions
application/mp4	Yes	Yes	No	Yes	No	.f4v; .m4a; .m4v; .mp4; .mp4v; .mpeg; .mpeg4
audio/mpeg	Yes	Yes	No	Yes	No	.mpeg
video/3gpp	Yes	Yes	No	Yes	No	.3gp
video/3gpp2	Yes	Yes	No	Yes	No	.3g2; .3gp2
video/quicktime	Yes	Yes	No	Yes	No	.moov; .mov; .qt
video/x-m4v	Yes	Yes	No	Yes	No	.m4v

Database

Expand table

Mime type	File identification	Metadata extraction	Text extraction	Native viewer	Annotate viewer	Possible Extensions
application/x-msaccess	Yes	Yes	Yes	No	No	.mdb

Email

Expand table

Mime type	File identification	Metadata extraction	Text extraction	Native viewer	Annotate viewer	Possible Extensions
application/vnd.ms-outlook	Yes	Yes	Yes	Yes	Yes	.msg
message/rfc822	Yes	Yes	Yes	Yes	Yes	.eml
text/vcard-contact	Yes	Yes	Yes	Yes	Yes	.vcf

Email Container

Expand table

Mime type	File identification	Metadata extraction	Container extraction	Possible Extensions
application/mbox	Yes	Yes	Yes	.mbox
application/vnd.ms-outlook-pst	Yes	Yes	Yes	.pst

HTML

Expand table

Mime type	File identification	Metadata extraction	Text extraction	Native viewer	Annotate viewer	Possible Extensions
application/xhtml+xml	Yes	Yes	Yes	Yes	Yes	.xhtml
application/xml	Yes	Yes	Yes	Yes	Yes	.xml
text/html	Yes	Yes	Yes	Yes	Yes	.htm; .html; .shtml

Image

Expand table

Mime type	File identification	Metadata extraction	OCR text extraction	Native viewer	Annotate viewer	Possible Extensions
image/bmp	Yes	Yes	Yes	Yes	Yes	.bmp
image/emf	Yes	Yes	Yes	Yes	Yes	.emf
image/gif	Yes	Yes	Yes	Yes	Yes	.gif
image/jpeg	Yes	Yes	Yes	Yes	Yes	.jpeg; .jpg
image/png	Yes	Yes	Yes	Yes	Yes	.png
image/svg+xml	Yes	Yes	Yes	Yes	No	.svg
image/tiff	Yes	Yes	Yes	Yes	Yes	.tif
image/vnd.dwg	Yes	Yes	Yes	Yes	Yes	.dwg; .dxf
image/wmf	Yes	Yes	Yes	Yes	Yes	.wmf

Note

The **OCR text extraction** column indicates that you can extract text from these image formats when you enable eDiscovery OCR for the case and the item is processed during advanced indexing. OCR text extraction occurs when you add the items to the review set. It doesn't make these image types searchable by text content during search.

Microsoft Excel

Expand table

Mime type	File identification	Metadata extraction	Text extraction	Native viewer	Annotate viewer	Possible Extensions
application/vnd.ms-excel	Yes	Yes	Yes	Yes	Yes	.dat; .xls
application/vnd.ms-excel.sheet.binary.macroenabled.12	Yes	Yes	Yes	Yes	No	.xlsb
application/vnd.ms-excel.sheet.macroenabled.12	Yes	Yes	Yes	Yes	Yes	.xlsm
application/vnd.ms-excel.template.macroenabled.12	Yes	Yes	Yes	No	No	.xltm
application/vnd.openxmlformats-officedocument.spreadsheetml.sheet	Yes	Yes	Yes	Yes	Yes	.xlsx
application/vnd.openxmlformats-officedocument.spreadsheetml.template	Yes	Yes	Yes	Yes	Yes	.xltx

Microsoft OneNote

Expand table

Mime type	File identification	Metadata extraction	Text extraction	Native viewer	Annotate viewer	Possible Extensions
application/onenote	Yes	Yes	Yes	No	No	.one

Note

OneNote section files (.one) stored directly in SharePoint or OneDrive are supported for search when the site is included as a search location. OneNote notebooks are represented as folders in SharePoint, not individual files. For more information about collecting OneNote content as cloud attachments, see Cloud attachments.

Microsoft PowerPoint

Expand table

Mime type	File identification	Metadata extraction	Text extraction	Native viewer	Annotate viewer	Possi Extens
application/vnd.ms-powerpoint	Yes	Yes	Yes	Yes	Yes	.pot; .p .pp
application/vnd.openxmlformats-officedocument.presentationml.presentation	Yes	Yes	Yes	Yes	Yes	.ppt
application/vnd.openxmlformats-officedocument.presentationml.slideshow	Yes	Yes	Yes	Yes	Yes	.pps
application/vnd.openxmlformats-officedocument.presentationml.template	Yes	Yes	Yes	Yes	Yes	.pot

Microsoft Project

Expand table

Mime type	File identification	Metadata extraction	Text extraction	Native viewer	Annotate viewer	Possible Extensions
application/vnd.ms-project	Yes	Yes	Yes	No	Yes	.mpp

Microsoft Publisher

Expand table

Mime type	File identification	Metadata extraction	Text extraction	Native viewer	Annotate viewer	Possible Extensions
application/x-mspublisher	Yes	Yes	Yes	Yes	Yes	.pub

Microsoft Visio

Expand table

Mime type	File identification	Metadata extraction	Text extraction	Native viewer	Annotate viewer	Possible Extensions
application/vnd.ms-visio.drawing	Yes	Yes	Yes	Yes	No	
application/vnd.visio	Yes	Yes	Yes	Yes	Yes	.vsd

Microsoft Word

Expand table

Mime type	File identification	Metadata extraction	Text extraction	Native viewer	Annotate viewer	Poss Exten
application/msword	Yes	Yes	Yes	Yes	Yes	.dat;
application/rtf	Yes	Yes	Yes	Yes	Yes	.doc
application/vnd.ms-word.document.macroenabled.12	Yes	Yes	Yes	Yes	Yes	.do
application/vnd.ms-word.template.macroenabled.12	Yes	Yes	Yes	Yes	Yes	.do
application/vnd.openxmlformats-officedocument.wordprocessingml.document	Yes	Yes	Yes	Yes	Yes	.dc
application/vnd.openxmlformats-officedocument.wordprocessingml.template	Yes	Yes	Yes	Yes	Yes	.dc

Microsoft Works

Expand table

Mime type	File identification	Metadata extraction	Text extraction	Native viewer	Annotate viewer	Possible Extensions
application/vnd.ms-works-ss	Yes	Yes	No	No	No	.wps
application/vnd.ms-works-wp	Yes	Yes	No	No	No	.wps

Open Document Format

Expand table

Mime type	File identification	Metadata extraction	Text extraction	Native viewer	Annotate viewer	Possible Extension
application/vnd.oasis.opendocument.text	Yes	Yes	Yes	Yes	Yes	.odt

Other

Expand table

Mime type	File identification	Metadata extraction	Text extraction	Native viewer	Annotate viewer	Possible Extensions
application/json	Yes	Yes	Yes	Yes	Yes	n/a
application/octet-stream	Yes	No	No	No	No	.fluid
application/vnd.ms-graph	Yes	Yes	No	No	No	
application/winhlp	Yes	Yes	No	No	No	.hlp
application/x-tnef	Yes	Yes	No	No	No	

Plain Text

Expand table

Mime type	File identification	Metadata extraction	Text extraction	Native viewer	Annotate viewer	Possible Extensions
text/csv	Yes	Yes	Yes	Yes	Yes	.csv
text/plain	Yes	Yes	Yes	Yes	Yes	.con; .css; .csv; .dat; .pl; .txt

Portable Document Format

Expand table

Mime type	File identification	Metadata extraction	Text extraction	Native viewer	Annotate viewer	Possible Extensions
application/pdf	Yes	Yes	Yes	Yes	Yes	.pdf

Word Perfect

Expand table

Mime type	File identification	Metadata extraction	Text extraction	Native viewer	Annotate viewer	Possible Extensions
application/vnd.wordperfect; version=5.0	Yes	Yes	Yes	No	No	.wpd
application/vnd.wordperfect; version=5.1	Yes	Yes	Yes	No	No	.wpd
application/vnd.wordperfect; version=6.x	Yes	Yes	Yes	No	No	.wpd

Word Pro

Expand table

Mime type	File identification	Metadata extraction	Text extraction	Native viewer	Annotate viewer	Possible Extensions
application/vnd.lotus-wordpro	Yes	Yes	No	No	No	.lwp

ⓘ**Note:** The author created this article with assistance from AI. Learn more

Last updated on 06/29/2026

Document metadata fields in eDiscovery

The following table lists the metadata fields for documents in Microsoft Purview eDiscovery. For more information about searchable properties when searching Microsoft 365 content locations when you're collecting data for an eDiscovery case, see Create a search query for a case in eDiscovery.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

This table provides the following information:

- **Field name**: The name of the metadata field included in the `items.csv` report. Some metadata field values don't appear until analytics runs in the review set. These fields are highlighted with an asterisk (*).
- **Description**: A description of the metadata field.
- **Direct export**: If the metadata field value is populated for direct exports from search.
- **Add to review set**: If the metadata field value is populated when adding search results to a review set.
- **Review set export**: If the metadata field value is populated when exporting items from a review set.

> **Note**
>
> When you export search results by using the **report only** option (without including content), some metadata fields might not be fully populated. Report-only exports rely on indexed properties, which might not reflect the most current or complete metadata for all items. To ensure all metadata fields are accurately populated, export search results **with content**. Export with data fully inspects each item and returns complete metadata values as documented in the following table.

Expand table

Field name	Description	Direct export	Add to review set	Review set export
Added by	Indicates how the item was included in this process, whether it was a match from *IndexQuery*, *AdvancedIndexed*, or *PartiallyIndexed*.	Yes	Yes	Yes
All custodians	Union of the `Custodian` and the `Deduped custodians` fields.	No	Yes	Yes
Attachment names	List of attachment names of the email item.	No	Yes	Yes
Author	Author from the document metadata.	Yes	Yes	Yes
BCC	Bcc field for message types. The format is `DisplayName <SMTPAddress>`. This field is only available when the email is collected from the sender mailbox.	Yes	Yes	Yes
BCC expanded	The value in Bcc field of an email expanded. This field is only available when the email is collected from the sender mailbox. For example, if Bcc is a distribution list, this field shows the Bcc and all of its members.	Yes	Yes	Yes
Category	Category label applied in Outlook to the email item. This field is only available for items residing in the mailbox where the category label is applied.	Yes	Yes	Yes
CC	Cc field for message types. The format is `DisplayName <SMTPAddress>`.	Yes	Yes	Yes
CC expanded	Cc field of the email or message expanded. If the CC field is a distribution list, the Cc expanded field contains a list of its members. This expansion is only available if the item is collected from the sender's mailbox.	Yes	Yes	Yes
Channel Name	Teams channels and private channels: *Channel*.	Yes	Yes	Yes
Client conversation ID	The Exchange email or Teams conversation ID.	No	Yes	Yes
Compound path	Path that shows the full location of an item across its source systems (such as mailbox or SharePoint), including the source name, to describe where the item originated.	Yes	Yes	Yes
Contains deleted message	Boolean field indicating if the item contains a deleted message.	Yes	Yes	Yes

Field name	Description	Direct export	Add to review set	Review set export
Contains edited message	Boolean field indicating if the item contains an edited message.	Yes	Yes	Yes
Content source application	Application from which the content originates. For example, *SharePoint* for documents and list items stored in SharePoint; *MicrosoftTeams* for Teams messages, chats, and channel content.	No	Yes	Yes
Conversation index	Index of each item within the same conversation.	No	Yes	Yes
Conversation name	Sub-header of Teams announcement posts or the title of Teams channel post.	Yes	Yes	Yes
Conversation topic	This field depends on content type. Emails: typically the subject of the email without the forward or reply. Teams 1:1 chat: the first 40 characters of the first message. Teams 1:N chat: Name of the group chat; if not available, the first 40 characters of the first message. Teams Channel Post: Post title or announcement subheader; if not available, the first 40 characters of the first message.	No	Yes	Yes
Conversation type	The type of chat conversation. Values are: Teams 1:1 and group chats and all Viva Engage conversations: *Group*.	Yes	Yes	Yes
Created	Date and time of the item's creation time.	Yes	Yes	Yes
Created by	*Created by* value from SharePoint.	Yes	Yes	Yes
Custodian	SMTP address of the custodian the item was associated with.	No	Yes	Yes
Data source	The data source name of the person or group from which the item is collected.	Yes	Yes	No
Date[1]	Date is a computed field that depends on the file type: - **Email**: Sent date - **Email attachments**: Last modified date of the document; if not available, the parent's sent date. - **Embedded documents**: Last modified date of the document; if not available, the parent's last modified date. - **SharePoint documents (includes modern attachments)**: Last modified date of the document; if not available, SharePoint last modified date. - **Non-Office 365 documents**: Last modified date.	Yes	Yes	Yes

Field name	Description	Direct export	Add to review set	Review set export
	- **Meetings**: Meeting start date. - **VoiceMail**: Sent date. - **IM**: Sent date. - **Teams**: Sent date			
Deduped compound path*	List of compound paths of documents that are exact duplicates (email: based on content, documents: based on hash).	N/A	N/A	Yes
Deduped custodians*	List of custodians of documents that are exact duplicates (for email, based on content; for documents, based on hash).	N/A	N/A	Yes
Deduped file IDs*	List of file IDs of documents that are exact duplicates (for email, based on content; for documents, based on hash).	N/A	N/A	Yes
Deduped group file IDs*	List of Group file IDs of documents that are exact duplicates (for email, based on content; for documents, based on hash).	N/A	N/A	Yes
Deduped thread file IDs*	List of thread file IDs of documents that are exact duplicates (for email, based on content; for documents, based on hash).	N/A	N/A	Yes
Detected language	The language identified in the item, typically represented by a locale code (for example, en-us), as determined by SharePoint.	No	Yes	Yes
DG expansion result	Indicates whether the expansion of distribution group membership was successful. When emails are sent to distribution groups, additional fields such as *To Expanded*, *Cc Expanded*, or *Bcc Expanded* capture the expanded membership. These values are saved in the sender's mailbox copy if the sender is on hold. This field records failures in that expansion process.	Yes	Yes	Yes
Doc authors	Author from the document metadata.	Yes	Yes	Yes
Doc comments	Comments from the document metadata.	No	Yes	Yes
Doc company	Company from the document metadata.	No	No	Yes
Doc date created	Create date from document metadata.	No	Yes	Yes
Doc index	Position of a document within its hierarchical family, where a value of 0 represents the root (top-level item)	No	Yes	Yes

Field name	Description	Direct export	Add to review set	Review set export
	and higher numbers denote increasing levels of embedding within another document.			
Doc keywords	Keywords from the document metadata.	No	No	No
Doc subject	Subject from the document metadata.	No	Yes	Yes
Doc template	Template from the document metadata.	No	No	Yes
Document ID index	An internal index identifier for an item that's unique only within a single mailbox; items in different mailboxes (including different archive mailboxes for the same user) can share the same value.	Yes	Yes	Yes
Document ID path	Represents the hierarchical sequence of item IDs, showing the relationship when an item is embedded within another. A single ID indicates a top-level item with no parent.	No	Yes	Yes
Dominant theme*	Dominant theme as calculated for analytics.	N/A	N/A	Yes
Email action	Values are *None*, *Reply*, or *Forward* and are based on the subject line of a message.	Yes	Yes	Yes
Email date sent	Sent date of the message. For chats, the beginning date from the transcript.	Yes	Yes	Yes
Email delivery receipt	Boolean field indicating whether the sender requested a delivery receipt for the email message.	No	No	Yes
Email importance	Importance of the message: Low; Normal; High	Yes	Yes	Yes
Email internet headers	The full set of email headers from the email message	No	Yes	Yes
Email level	Indicates a message's level within the email thread it belongs to; attachments inherit its parent message's value.	No	No	No
Email participant domains	List of domains of all participants.	Yes	Yes	Yes
Email read receipt	True or false value if a read receipt was received at sender end.	No	Yes	Yes

Field name	Description	Direct export	Add to review set	Review set export
Email recipient domains	List of all domains of recipients of a message.	Yes	Yes	Yes
Email recipients	List of all recipients of a message (To, Cc, Bcc).	Yes	Yes	Yes
Email sender domain	Domain of the sender.	Yes	Yes	Yes
Email set	**Retired field**.	No	No	No
Email thread	**Retired field**	No	No	No
Error detail	If Status isn't successful, this field shows all the warnings and errors associated with the item.	Yes	Yes	Yes
Extracted content type	MIME media type of the item. For example, image/jpeg, application/vnd.ms-outlook, text/plain; charset=windows-1252, video/mp4. This type identifies the format of the extracted content.	No	Yes	Yes
Extracted text length	Number of characters in the extracted text.	No	Yes	Yes
Extracted text path	The path to the extracted text file in the export. This path applies only to review set export and only when user selects *export text file* in the settings.	No	No	Yes
Family duplicate set*	Numeric identifier for families that are exact duplicates of each other (same content and all the same attachments).	N/A	N/A	Yes
Family ID	Groups together attachments and extracted items from email and chats with its parent item. This grouping includes the chat or email and all attachments and extracted items.	No	Yes	Yes
Family size	Number of documents in the family.	No	Yes	Yes
File class	For content from SharePoint and OneDrive: Document. For content from Exchange: Email or Attachment. For content from Teams or Viva Engage: Conversations.	No	Yes	Yes
File extension	Extension of the file, such as .csv, .pdf, .html, .txt, .jpg, and more.	Yes	Yes	Yes
File ID	Document identifier unique within a case.	No	Yes	Yes
File name	Native file name of the item.	Yes	Yes	Yes

Field name	Description	Direct export	Add to review set	Review set export
Group ID	Groups together all items for email and documents. For email, this grouping includes the message and all attachments and extracted items. For documents, this grouping includes the document and any embedded items.	No	Yes	Yes
Group representative ID*	Numeric identifier of each set of exact duplicates within the same group. The Group Representative ID corresponds to the document ID of the chosen representative in the group.	N/A	N/A	Yes
Has attachment	Indicates whether the message has attachments. Some items might indicate they have an attachment when they are actually embeddings in the email (.png or .html). When viewed in the review set, the email item won't have any attachments associated with it.	Yes	Yes	Yes
Has text	A Boolean field indicating whether the item contains extractable text content; possible values are True (text is present) or False (no text available).	Yes	Yes	Yes
Has unique attachment*	This field marks emails that contain attachments not found in other emails within the same thread. Even if the email content is duplicated, unique attachments are flagged to ensure that all relevant documents are reviewed. This field is important in the legal review process to ensure that no unique evidence is overlooked, even if the email body isn't unique.	N/A	N/A	Yes
Identifier	The Internet Message ID for Exchange items and the SitePropertyPath for SharePoint items.	No	Yes	Yes
Immutable ID	An unchangeable identifier assigned to each item. This value is unique for parent items across mailboxes, but attachments and extracted items inherit the Immutable ID from their parent, so it isn't unique across all items. To uniquely identify individual items within a case, use the **File ID** field instead.	Yes	Yes	Yes
In reply to ID	In reply to ID from the message. This field identifies the unique message ID of the email or message that the current item is replying to. It's used to track conversational threads and message relationships, especially in threaded discussions or email chains. This field helps reconstruct the context of a reply by linking it	No	Yes	Yes

Field name	Description	Direct export	Add to review set	Review set export
	back to the original message, but it doesn't contain the full content of the original message itself.			
Inclusive type	**Retired field**.	N/A	N/A	N/A
Input file ID	The file ID of the top level item in the review set. For an attachment, this ID is the ID of the parent. Use this ID to group families together.	No	Yes	Yes
Internet message ID	Internet message ID from the message.	Yes	Yes	Yes
Is attachment from transcript	Is attachment part of teams transcript	Yes	Yes	Yes
Is bcc to me	Boolean metadata field that indicates whether the user was included in the Bcc line of the email.	No	Yes	Yes
Is doc from conversation	Boolean metadata field that indicates whether the document originated from a Microsoft Teams or Viva Engage conversation.	Yes	Yes	Yes
Is draft	Indicate if the email is a draft.	No	Yes	Yes
Is encrypted	Indicate if the item is encrypted.	Yes	Yes	Yes
Is external	Indicate if the email is from an external organization.	No	Yes	Yes
Is group representative*	This field indicates whether an email is marked as the single representative copy within a specific group. Only one email with the same hash from the same group is designated as the group representative.	N/A	N/A	Yes
Is inclusive*	This field identifies whether an email contains all the unique content from a thread, including all previous replies. It ensures that only the most comprehensive email in a thread is reviewed, which is essential for understanding the full context of the conversation without having to review each individual reply. This field helps in efficiently managing long email threads by focusing on the most informative emails.	N/A	N/A	Yes
Is modern attachment	Boolean field indicating if this file is a modern attachment or linked file.	Yes	Yes	Yes
Is read	Boolean field indicates whether an email is opened or viewed by the recipient.	Yes	Yes	No

Field name	Description	Direct export	Add to review set	Review set export
Is representative*	One document in every set of exact duplicates is marked as representative.	N/A	N/A	Yes
Is thread representative*	This field indicates whether an email is marked as the single representative copy within a specific thread. Only one email with the same hash from the same thread is designated as the thread representative.	N/A	N/A	Yes
Is top for group	Boolean field indicating if this item is the parent of the group.	No	Yes	Yes
Item class	Field that categorizes the type and structure of a message item such as *IPM.Note* for email, *IPM.File* for documents, or *IPM.SkypeTeams.Message* for Teams chats.	Yes	Yes	Yes
Item source	Mailbox or site from which the item originates.	Yes	Yes	Yes
Last modified by	The user who last modified the document.	No	Yes	Yes
Last modified time	Last modified date from document metadata.	No	Yes	Yes
Last modified time for retention	SharePoint document upload time for retention purposes.	No	No	No
Load ID	The ID of the load set in which the item was added to a review set. This ID corresponds to the job ID of the specific Add To Review Set job.	No	Yes	Yes
Location ID	Site ID for SharePoint locations and mailbox ID for exchange locations.	Yes	Yes	Yes
Location sub type	Specifies the type of location, such as *PrimaryMailbox*, *SystemMailbox*, *ArchiveMailbox*, or *OneDriveSite*.	Yes	Yes	Yes
Marked as pivot*	Boolean field indicating if This file is the pivot in a near duplicate set.	N/A	N/A	Yes
Meeting end date	End time of the meeting for items of item class. Values are *IPM.Schedule.Meeting.Request*, *IPM.Appointment*, or *IPM.AppointmentSnapshot.SkypeTeams.Meeting*.	Yes	No	Yes
Meeting name	Name of Teams meeting.	No	No	No
Meeting start date	Start time of the meeting for items such as item class. For example,	No	No	Yes

Field name	Description	Direct export	Add to review set	Review set export
	IPM.AppointmentSnapshot.SkypeTeams.Meeting.			
Message kind	The type of message to search for. Possible values: - *contacts* - *docs* - *email* - *externaldata* - *faxes* - *im* - *journals* - *meetings* - *microsoftteams* (returns items from chats, meetings, and calls in Microsoft Teams) - *notes* - *posts* - *rssfeeds* - *tasks* - *voicemail*.	Yes	Yes	Yes
Modern attachment embedded URLs	A modern attachment URL is embedded into item.	No	No	No
Modern attachment parent ID	The Immutable ID of the document's parent.	No	No	Yes
Native copy from	Field that contains the blob URL of the original document.	No	Yes	Yes
Native MD5	MD5 hash (128-bit hash value) of the file stream.	No	Yes	Yes
Native SHA 256	SHA256 hash (256-bit hash value) of the file stream.	No	Yes	Yes
ND ET sort excl attach	**Retired field**.	N/A	N/A	N/A
ND ET sort incl attach	**Retired field**.	N/A	N/A	N/A
ND set	**Retired field**.	N/A	N/A	N/A
Organizer	Meeting organizer's display name.	No	Yes	Yes

Field name	Description	Direct export	Add to review set	Review set export
Original file extension	The original file extension of the file.	No	Yes	Yes
Original path	Original path of item in mailbox or path in SharePoint.	Yes	Yes	No
Parent ID	ID of the item's parent.	No	Yes	Yes
Parent node	The closest preceding email message in the email thread.	No	No	No
Participant expansion	Lists all individual members of a group participant such as a distribution list by expanding it into its constituent participants.	No	Yes	No
Participants	List of all participants of a message; for example, sender, to, Cc, Bcc.	Yes	Yes	Yes
Pivot ID*	Analytics field value generated after running analytics in review set. The ID of a pivot.	N/A	N/A	Yes
Potentially privileged*	Analytics field value generated after running analytics in review set - True if attorney-client privilege detection model considers the document potentially privileged.	N/A	N/A	Yes
Preservation original URL	This field that captures the original location and version of an item before it was moved to the Preservation Hold Library (PHL).	No	Yes	No
Received	The date and time the email was received in UTC.	Yes	Yes	Yes
Recipient count	Number of recipients in the message.	Yes	Yes	Yes
Redacted file path	The path of the redacted replacement file in the export.	No	No	No
Redacted text path	The path of the redacted text file replacement in the export. For internal Microsoft use only.	No	No	No
Representative ID*	Numeric identifier of each set of exact duplicates.	N/A	N/A	Yes
Retention label	Retention labels applied to content in Office 365.	Yes	Yes	Yes
Retention URL	Captures the SharePoint or OneDrive path of an item under retention, typically pointing to its preserved version in the Preservation Hold Library (PHL).	Yes	No	No

Field name	Description	Direct export	Add to review set	Review set export
Sender	Sender (From) field for message types. The format is `DisplayName <SMTPAddress>`.	Yes	Yes	Yes
Sender/Author	Calculated field comprised of the sender or author of the item.	No	Yes	Yes
Sensitive type	List if SIT GUIDs associated with the item.	Yes	Yes	Yes
Sensitivity label	GUID of the Sensitivity label applied to the item.	Yes	Yes	Yes
Set ID	**Retired field**.	N/A	N/A	N/A
Set order inclusives first	**Retired field**.	N/A	N/A	N/A
Shared with	Lists the users or groups that have access to the SharePoint item.	No	No	No
SharePoint Item Content Type ID	Hexadecimal strings used as SharePoint content type identifiers are stored in the *SharepointItemContentTypeID* field, which helps identify the item's content type and map SharePoint content into the substrate's flexible schema.	Yes	Yes	No
Similarity percent*	Indicates how similar a document is to the pivot of the near duplicate set.	N/A	N/A	Yes
Size	Number of bytes of the item.	Yes	Yes	Yes
Source ID	SMTP for mailbox or URL for site of the item source.	Yes	Yes	No
SPO document link	SharePoint document path associated with an item. It references the location of the document within SharePoint.	No	Yes	Yes
SPO preservation original document unique ID	A unique identifier assigned to a document that's preserved in SharePoint under a preservation policy.	No	Yes	Yes
SPO unique ID	Unique ID of the SharePoint from where the document is retrieved.	No	Yes	Yes
Status	Overall status of the item in the process; indication whether item is successful or it encounters issues. Values are *Successful*, *Retrieval error*, *Skipped*, and more.	Yes	Yes	Yes

Field name	Description	Direct export	Add to review set	Review set export
Subject/Title	Calculated field comprised of the subject or title of the item.	Yes	Yes	Yes
Tags	Tags applied to the item, only apply to Review set export items.	No	No	Yes
Target path	Target path of the item in the export package .zip files.	Yes	No	Yes
Team name	Name of the Teams channel.	Yes	Yes	No
Teams announcement title	Title of the Teams announcement post.	Yes	Yes	No
Teams channel	Teams channel name. Only applies to Microsoft Teams content.	No	Yes	Yes
Themes list*	Themes list as calculated for analytics.	N/A	N/A	Yes
Thread participant domains*	Domains of all participants in a thread.	N/A	N/A	Yes
Thread participants*	Participants in a thread.	N/A	N/A	Yes
Thread representative ID*	Numeric identifier of each set of exact duplicates within the same thread. The thread Representative ID corresponds to the document ID of the chosen representative in the group.	N/A	N/A	Yes
Title	Title of the document.	No	Yes	Yes
To	To field value of email and messages.	Yes	Yes	Yes
To expanded	The *To* field of the email or message expanded, that is, if the To field is a distribution list, the To expanded field contains a list of its members. This expansion is only available if the item is collected from the sender's mailbox.	Yes	Yes	Yes
Type	Type of the content for the item. For example, *Message*.	Yes	Yes	Yes
Version group ID	Groups together the different versions of the same document.	No	Yes	Yes
Version number	The version number of a document collected from SharePoint or OneDrive. This version number matches	No	Yes	Yes

Field name	Description	Direct export	Add to review set	Review set export
	the one displayed in the version history in the SharePoint and OneDrive user experience.			
Was remediated	Indicate whether this item was remediated by the error remediation process. Only applies to review set exports.	No	No	Yes
Word count	The word count of the extracted text.	No	Yes	Yes
Workload	Indicates which workload the item comes from, Exchange or SharePoint.	Yes	Yes	Yes

Note

[1] When you use the *Date* condition in an eDiscovery search, you might see discrepancies in the results because Microsoft 365 content types store and interpret dates differently. For example, emails, documents, and Teams messages might have several date-related metadata fields, such as *Created*, *Last Modified*, *Received*, and *Sent*. The *Date* condition might apply differently across these items, depending on the content source and query. To improve accuracy and get the results you want, use date ranges instead of specific dates. Combine date conditions with keywords or other metadata filters to narrow your results. Review search statistics and preview results to validate search results before finalizing exports.

Last updated on 04/02/2026

Advanced indexing in eDiscovery

09/04/2025

When you add a data source to an eDiscovery case, the system reindexes any content it deems as partially indexed or that has indexing errors. This reindexing process is called *advanced indexing*. Many reasons can cause content to be partially indexed or have indexing errors. These errors might include image files, the presence of images in a file, unsupported file types, or file size indexing limits.

Advanced indexing is available in eDiscovery with premium feature support. The system applies advanced indexing when it runs processes like generating statistics, adding items to review set, and direct export. You don't need to wait for advanced indexing to complete before conducting a search or adding items to a review set like previous versions of eDiscovery. Also, you don't need to manually trigger any update on advanced indexing to ensure newly added content to the target locations need to be reindexed.

> **Tip**
>
> Want to try premium eDiscovery features? See the subscription requirements for Microsoft 365 Enterprise E5 licensing.

For SharePoint files, advanced indexing only runs on items marked as partially indexed or items with indexing errors. In Exchange, email messages with image attachments aren't marked as partially indexed or with indexing errors. This means that those files aren't reindexed by the advanced indexing process.

> **Note**
>
> Optical Character Recognition (OCR) automatically runs during advanced indexing. For more information about how OCR works and to configure OCR related settings, see Learn about search and analytics settings in eDiscovery cases.

View advanced indexing results

On the statistics dashboard, you can review advanced indexing hits and choose where to apply advanced indexing. You can view estimated hits on how many additional items from partially indexed items that match the search query after they're reindexed with advanced indexing. To ensure full advanced indexing runs on all partially indexed items in scope without sampling, use add to a review set and export with advanced indexing.

You can also review advanced indexed items in search results in the process report. This report provides details on natively indexed items, partially indexed items, and items identified by advanced indexing.

For more granular insights, you can use the CSV reports. This might include which specific items were exported or which locations were searched. These reports indicate whether an item was included due to a match with the indexed query or as a result of advanced indexing. The location report provides hit counts and data volume per location, broken down by matches to the standard index versus the advanced index.

Scope advanced indexing

You can scope advanced indexing to partially indexed items in locations where there are already indexed hits, partially indexed items in locations where there aren't any indexed hits, or both.

- **Partially indexed items from locations with indexed hits**: This option is more targeted and limits the scope to only those locations (mailboxes, sites) where some content was successfully indexed and matched the search criteria.

  Choose this option when:
  - You want to reduce noise and focus only on locations already known to be relevant.
  - You're conducting a narrow investigation and want to avoid reviewing irrelevant content.
  - You're optimizing for review efficiency and cost control (for example, in large-scale eDiscovery cases).
  - You want to prioritize the locations and the partially indexed items in locations with high relevancy to the matter first, then later come back to do a more comprehensive analysis that include all locations (including those without indexed hits).

- **Partially indexed items from all locations searched**: This option is more comprehensive and includes partially indexed items from every location searched, regardless of whether any indexed content matched.

  Choose this option when:
  - You need maximum completeness for legal defensibility or regulatory compliance.
  - You're investigating potential data concealment or malicious behavior where relevant content might exist only in unindexed formats (for example, image-based attachments or encrypted files).
  - You're dealing with high-risk scenarios where missing even one critical item could have legal or business consequences.

Partially indexed items in eDiscovery

When you select **Include partially indexed items**, eDiscovery searches automatically include partially indexed items in the estimated search results. Partially indexed items are Exchange mailbox items and documents on SharePoint and OneDrive sites that weren't completely indexed for search. In Exchange, a partially indexed item typically contains a file (of a file type that can't be indexed) attached to an email message.

The following reasons explain why items can't be indexed for search and are returned as partially indexed items:

- The file type is unrecognized or unsupported for indexing.
- Messages have an attached file that can't be opened; this issue is the most common cause of partially indexed email items.
- The file type is supported for indexing but an indexing error occurred for a specific file.
- Too many files attached to an email message.
- A file attached to an email message is too large.
- A file is encrypted with non-Microsoft technologies.
- A file attached to an email message is encrypted with a Microsoft Purview sensitivity label that applies encryption. Exchange doesn't index the attachment content, even though the attachment can be decrypted during export.
- A file is password-protected.

> **Note**
>
> Most organizations have less than 1% of content by volume and less than 12% by size that is partially indexed. The reason for the difference between volume and size is that larger files have a higher probability of containing content that can't be completely indexed.

For legal investigations, your organization might be required to review partially indexed items. You can specify whether to include partially indexed items when you export search results or when you add search results to a review set.

File types not indexed for search

Certain types of files, such as Bitmap (.bmp) or MP3 (.mp3) files, don't contain content that can be indexed. The search indexing servers in Exchange and SharePoint don't perform full-text indexing on these types of files. These file types are considered unsupported file types. There are also file types for which full-text indexing is disabled, either by default or by an administrator. Unsupported and disabled file types are labeled as unindexed items in searches.

For a list of supported and disabled file formats, see the following articles:

- **Exchange**: File formats indexed by Exchange Search
- **Exchange**: Get-SearchDocumentFormat
- **SharePoint**: Default crawled file name extensions and parsed file types in SharePoint

Messages and documents with partially indexed file types returned in search results

Not every email message with a partially indexed file attachment or every partially indexed SharePoint document is automatically returned as a partially indexed item. That's because other message or document properties, such as the **Subject** property in email messages and the **Title** or **Author** properties for documents, are indexed and available to be searched. For example, a keyword search for *financial* returns items with a partially indexed file attachment if that keyword appears in the subject of an email message or in the file name or title of a document. However, if the keyword appears only in the body of the file, the message or document is returned as a partially indexed item.

Similarly, messages with partially indexed file attachments and documents of a partially indexed file type are included in search results when other message or document properties, which are indexed and searchable, match the search criteria. Message properties that are indexed for search include sent and received dates, sender and recipient, the file name of an attachment, and text in the message body. Document properties indexed for search include created and modified dates. So even though a message attachment might be a partially indexed item, the message is included in the regular search results if the value of other message or document properties matches the search criteria.

For a list of email and document properties that you can search for by using eDiscovery tools in the Microsoft Purview portal, see Keyword queries and search conditions for eDiscovery.

> ⓘ **Note**

If you move a mailbox item from a folder that is indexed to a folder that isn't indexed, a flag is set to unindex the item. The item is removed from the index and isn't searchable. Later, if you move that same item back to a folder that is indexed, the flag isn't reset. That means the item remains unindexed, and not searchable.

Partially indexed items included in search results

Your organization might need to identify and analyze partially indexed items to determine what they are, what they contain, and whether they're relevant to a specific investigation. If you select **Include partially indexed items**, the partially indexed items in the content locations that you search are automatically included with the estimated search results. Depending on the specific setting you select, you can control whether to include partially indexed items in locations with indexed search hits, partially indexed items in locations without indexed search hits, or both. You can also include these partially indexed items when you export search results or add items to review sets.

Keep the following points in mind about partially indexed items:

- When you run an eDiscovery search, the total number and size of partially indexed Exchange items (returned by the search query) appear in the search statistics view, and they're labeled as **partially indexed items**. Statistics about partially indexed items don't include partially indexed items in SharePoint sites or OneDrive accounts.

- If you export results from a search of specific content locations or all content locations in your organization, only the unindexed items from content locations that contain items that match the search criteria are exported. In other words, if no search results are found in a mailbox or site, then any unindexed items in that mailbox or site aren't exported. The reason for this behavior is that exporting partially indexed items from many locations in the organization might increase the likelihood of export errors and increase the time it takes to export and download the search results.

  To export partially indexed items from all content locations for a search, configure the search to return all items (by removing any keywords from the search query) and then export only partially indexed items when you export the search results (by selecting **partially indexed items** under Select items to include in your export options).

- If you choose to include all mailbox items in the search results, or if a search query doesn't specify any keywords or only specifies a date range, partially indexed items might not be copied to the PST file that contains the partially indexed items. This behavior occurs because

all items, including any partially indexed items, are automatically included in the regular search results.

- You can't preview partially indexed items. To view partially indexed items returned by the search, you need to export the search results.

  Additionally, when you export search results and include partially indexed items in the export, partially indexed items from SharePoint items are exported to a folder named **Uncrawlable**. When you export partially indexed Exchange items, they're exported differently depending on whether the partially indexed items matched the search query and the configuration of the export settings.

- The following table shows the export behavior of indexed and partially indexed items and whether each is included for the different export configuration settings.

Expand table

Export configuration	Indexed items that match search query	Partially indexed items that match search query	Partially indexed items that don't match search query
Export only indexed items	Exported	Exported (included with the indexed items exported)	Not exported
Export only partially indexed items	Not exported	Exported (as partially indexed items)	Exported (as partially indexed items)
Export indexed and partially indexed items	Exported	Exported (included with the indexed items exported)	Exported (as partially indexed items)

Locations and the "in locations without search hits" option

When you include partially indexed items in an export (or add them to a review set), you can also choose to **exclude partially indexed items in locations without search hits**. Understanding what *location* means in this option determines what your export contains.

A **location** in eDiscovery is one of:

- A single **Exchange Online mailbox**, or
- A single **SharePoint Online site** or **OneDrive for Business site**.

A location is **not** a folder, a URL prefix, a SharePoint document library, or any subset of a mailbox or site defined by a Kusto Query Language (KQL) condition. The smallest unit the partial-index location filter operates on is the entire mailbox or the entire site.

What the option does

Expand table

Setting	Behavior
Exclude partially indexed items in locations without search hits (selected)	Partially indexed items are returned only from mailboxes and sites where at least one item also matched your search query. Mailboxes and sites with zero indexed matches contribute zero partially indexed items.
Exclude partially indexed items in locations without search hits (cleared)	Partially indexed items are returned from every selected mailbox and site, regardless of whether the indexed query matched anything in that mailbox or site.

KQL conditions don't constrain partially indexed items

Partially indexed items are returned based on the **location's index state**, not on whether the item matches your KQL query. If your query uses a path, URL, file-name pattern, or other property restriction (for example, `DocumentLink:"https://contoso.sharepoint.com/sites/Finance/Subfolder/*"`), partially indexed items from **anywhere** in the matched site can still appear in your export. They aren't evaluated against the restriction.

If you need to filter partially indexed items by KQL, select **Perform advanced indexing on partially indexed items** at export time. Advanced indexing re-indexes the items so your KQL conditions apply to them.

Single-location searches

If your scope is a single mailbox or a single site, the *Exclude partially indexed items in locations without search hits* option has **no effect**:

- If the location returns at least one indexed hit, it's considered responsive, and all of its partially indexed items are included.
- If the location returns no indexed hits, no items of any kind are exported from that location.

For single-location scopes, expect the location's full set of partially indexed items whenever you include partially indexed items at export, or use advanced indexing to apply your KQL conditions.

Example

A SharePoint search scoped to a single site uses the query `DocumentLink:"https://contoso.sharepoint.com/sites/Project/Subfolder/*"` and includes partially indexed items at export. Two outcomes are guaranteed:

1. **Indexed items** in the export are limited to the `Subfolder/*` path (the KQL condition applies).
2. **Partially indexed items** in the export include every partially indexed item across the entire `sites/Project` site, not just `Subfolder/*` (KQL doesn't apply to partially indexed items, and the location is the whole site).

To restrict partially indexed items to the same subfolder, enable **Perform advanced indexing on partially indexed items**.

Search-time and export-time settings are independent

eDiscovery search and export are separate jobs. The partially indexed item choices you make on each are independent:

Expand table

Where you set it	What it controls
Search query (search-time *Include partially indexed items*)	The statistics and preview returned by the search. It doesn't carry into a subsequent export.
Export dialog (export-time *Include partially indexed items*, *Exclude … in locations without search hits*, *Perform advanced indexing*)	The contents of the export package. These are the authoritative settings for what is exported.

If you set partial-index inclusion at search but clear it at export, the export contains no partially indexed items, regardless of what your search statistics showed. If you clear it at search but set it at export, the export contains partially indexed items according to the export-dialog settings. Always confirm partial-index behavior at the export step.

Date ranges and excluding partially indexed items

In eDiscovery searches, you can't use a date range to exclude partially indexed items from results in a search query. The search statistics include partially indexed items that fall outside of a date range, and you include them when you export partially indexed items. In eDiscovery with premium feature support, you can add partially indexed items to a review set and then filter them

in a review set before export. Alternatively, use the advanced indexing capability (a premium eDiscovery feature) to ensure these partially indexed items are reindexed to compare to the date range specified and to avoid a large volume of data getting exported.

For more information about indexing limits for email messages, see Limits in eDiscovery.

Assess partially indexed items in your organization

After you run an eDiscovery search, you can see the total number and size of partially indexed items in the locations that you searched in the search result statistics. The detailed statistics for the search show this information.

A few factors affect the number of partially indexed items that appear in the search results:

- If an item is partially indexed and matches the search query, the item is included in both the count and size of search result items and partially indexed items. However, when you export the results of that same search, the item is included only with set of search results and isn't included as a partially indexed item.
- The detailed statistics for the search *don't* include partially indexed items located in SharePoint and OneDrive sites in the estimate of partially indexed items. However, when you export the results of an eDiscovery search, you can include partially indexed items in the export. For example, if you only search sites, the estimated number of partially indexed items is zero.

To understand your organization's exposure to partially indexed items, run a search for all content in all mailboxes or a representative subset of mailboxes by using a blank keyword query.

In the following example, there are 1,629,904 (146.46 GB) fully indexed items and 10,025 (10.27 GB) partially indexed items.

You can determine the percentage of partially indexed items by using the following calculations.

To calculate the ratio of partially indexed items in your organization:

```
(Total number of partially indexed items/Total number of items) x 100
```

```
(10025/1629904) x 100 = 0.62%
```

By using the search results from the previous example, 0.62% of all mailbox items are partially indexed.

To calculate the percentage of the size of partially indexed items in your organization:

`(Size of all partially indexed items/Size of all items) x 100`

`(10.27 GB/146.46 GB) x 100 = 7.0%`

In the previous example, 7% of the total size of mailbox items comes from partially indexed items. Most customers have less than 1% of content by volume and less than 12% of content by size that is partially indexed.

For additional assessment of what types of content make up partially indexed items for your organization, select **Partially indexed items** to include in your export item report or overall export for further investigation.

More information about partially indexed items

- Because message and document properties and their metadata are indexed, a keyword search might return results if that keyword appears in the indexed metadata. However, that same keyword search might not return the same item if the keyword only appears in the content of an item with an unsupported file type. In this case, the item is returned as a partially indexed item.

- If a partially indexed item matches the search query criteria, it's included in the search results. However, it's not included with partially indexed items when you export search results.

- Although a file type is supported for indexing and is indexed, indexing or search errors can cause a file to be returned as a partially indexed item. For example, searching a large Excel file might be partially successful (because the first 4 MB are indexed), but then fails because the file size limit is exceeded. In this case, it's possible that the same file is returned with the search results and as a partially indexed item.

- Files that are encrypted with Microsoft encryption technologies, including Microsoft Purview sensitivity labels that apply encryption, behave differently depending on location:
  - Email attachments: Exchange doesn't index encrypted attachments. Keyword searches don't match content inside these attachments, and they're returned as partially indexed items. However, if the email message itself matches the search criteria (for example, the subject or sender), the message is included in the search results. You can decrypt encrypted attachments during preview and export.
  - SharePoint folders and OneDrive files: Files encrypted with Microsoft encryption technologies are partially indexed.

    - To ensure encrypted attachments are captured, include partially indexed items when exporting search results or adding results to a review set. Use review sets to analyze the decrypted content after collection.

- Email messages encrypted with S/MIME are partially indexed. This limitation includes encrypted messages with or without file attachments.

- Email messages protected using Azure Rights Management are indexed and are included in the search results if they match the search query. Rights-protected email messages are decrypted and can be previewed and exported. This functionality requires that you're assigned the RMS Decrypt role, which is assigned by default to the eDiscovery Manager role group.

- Don't use query-based holds to address encrypted or partially indexed items. Query-based hold with queries that include conditions beyond dates, participants, or item types - such as keywords or paths - might not be applicable for these items. There's a risk that the hold won't be applied as intended. To ensure coverage, limit query-based hold conditions for date ranges, participants, and item types only or apply a location-based hold. For more information, see Create an eDiscovery hold.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/23/2026

Search for eDiscovery activities in the audit log

The audit log records eDiscovery activities that you perform in Microsoft Purview portal or in PowerShell. Events are logged when administrators, eDiscovery managers, or any user assigned eDiscovery permissions perform the following tasks:

- Creating and managing eDiscovery cases.
- Creating and editing eDiscovery searches in cases.
- Performing search actions, such as generating statistics, sample, and exporting from search.
- Creating, editing, and removing eDiscovery holds from cases.
- Creating eDiscovery review sets and performing review activities in cases.

For more information about searching the audit log, the permissions that are required, and exporting search results, see Search the audit log.

How to search for and view eDiscovery activities

Perform the following steps to search for and view eDiscovery activities in the audit log.

1. Go to the Microsoft Purview portal and sign in with the credentials for a user account assigned eDiscovery permissions.

2. Select the **Audit** solution card.

3. On the **Search** page, complete the following steps:
   a. In the **Activities - friendly names** drop-down list, under **eDiscovery activities**, select one or more eDiscovery activities to search for. For a complete list of all eDiscovery activities, see Audit log activities.

   > **Note**
   >
   > The **Activities - friendly names** drop-down list also includes a group of activities named **eDiscovery cmdlet activities** that returns records from the cmdlet audit log.

   a. Select a date and time range to display eDiscovery events that occurred within that period.

b. In the **Users** box, select one or more users to display search results for. Leave this box blank to return entries for all users.

c. Select **Search** to run the search using your search criteria.

After the search results are displayed, you can select **Filter results** to filter or sort the resulting activity records. To view details about an activity, select the activity record in the list of search results. A **Details** flyout page is displayed that contains the detailed properties from the event record.

To display additional details, select **More information**. You can export the audit log search results to a .csv file, and then use the Excel Power Query feature to format and filter these records. For more information, see Export, configure, and view audit log records.

A client **IP Address** field is available for all activities performed in eDiscovery. If the **IP Address** column is blank for an activity, the activity was performed in the legacy eDiscovery experience. All eDiscovery activities performed in eDiscovery include an IP address in the **IP Address** column.

Audit activities logged for eDiscovery

For a complete listing of all eDiscovery detailed properties and activities included in the audit log, see Audit log activities.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/25/2026

Decryption in eDiscovery

Encryption is an important part of your file protection and information protection strategy. Organizations of all types use encryption technology to protect sensitive content within their organization and ensure that only the right people have access to that content.

To run common eDiscovery tasks on encrypted content, eDiscovery managers need to decrypt email message content when they export it from eDiscovery cases. Content encrypted with Microsoft encryption technologies isn't available for review until after export.

To make it easier to manage encrypted content in the eDiscovery workflow, eDiscovery tools now incorporate the decryption of encrypted files attached to email messages and sent in Exchange Online.[1] Additionally, encrypted documents stored in SharePoint and OneDrive are decrypted when premium eDiscovery features are enabled.[2]

Before this capability, only the content of an email message protected by rights management (and not attached files) was decrypted. Encrypted documents in SharePoint and OneDrive couldn't be decrypted during the eDiscovery workflow. Now, files that are encrypted with a Microsoft encryption technology and located on a SharePoint or OneDrive account are searchable and decrypted when the search results are prepared for preview, added to a review set, and exported. Additionally, encrypted documents in SharePoint and OneDrive that are attached to an email message (as a copy) are searchable. This decryption capability allows eDiscovery managers to view the content of encrypted email attachments and site documents when previewing search results, and review them after they're added to a review set when premium eDiscovery features are enabled.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about **Microsoft Security Copilot in Microsoft Purview**.

Requirements for decryption in eDiscovery

- **Permissions**: You need to be assigned the *RMS Decrypt* role to preview, review, and export files encrypted with Microsoft encryption technologies. You also need this role to review and query encrypted files that are added to a review set in eDiscovery. The Microsoft

Purview portal assigns this role by default to the eDiscovery Manager role group. For more information about the RMS Decrypt role, see Assign eDiscovery permissions.

- **Run the Inbox Repair tool on exported PST files**: After you export PST files, run the Inbox Repair tool (ScanPST.exe) to diagnose and repair any errors in the PST files.

Supported encryption technologies

For Exchange, eDiscovery tools support items encrypted with Microsoft encryption technologies. These technologies are Azure Rights Management (Azure RMS)[1] and Microsoft Purview Information Protection (specifically sensitivity labels). For more information about Microsoft encryption technologies, see Encryption and the various email encryption options available. Content encrypted by S/MIME or third-party encryption technologies isn't supported. For example, previewing or exporting content encrypted with non-Microsoft technologies isn't supported.

> **Note**
>
> Microsoft eDiscovery tools don't support the decryption of email messages sent with a Microsoft Purview Message Encryption custom branding template. When you use an OME custom branding template, email messages are delivered to the encrypted message portal instead of the recipient's mailbox. Therefore, you can't use eDiscovery tools to search for encrypted messages because those messages are never received by the recipient's mailbox.

For SharePoint, content labeled with *SharePoint online service* is decrypted. Items labeled or encrypted in the client before uploading to SharePoint, legacy document library RMS templates or settings, and S/MIME or other standards aren't supported.[2].

eDiscovery activities that support encrypted items

The following table identifies the supported tasks that you can perform in eDiscovery tools on encrypted files attached to email messages and encrypted documents in SharePoint and OneDrive. You can perform these supported tasks on encrypted files that match the criteria of a search. A value of `N/A` indicates the functionality isn't available in eDiscovery.

Expand table

eDiscovery task	eDiscovery	Premium features enabled
Search for content in encrypted files in sites and email attachments[1]	No	Yes
Preview encrypted files attached to email	No	Yes[2]
Preview encrypted documents in SharePoint and OneDrive	No	Yes
Review encrypted files in a review set	N/A	Yes
Export encrypted files attached to email	Yes	Yes
Export encrypted documents in SharePoint and OneDrive	No	Yes

Supported decryption

The following table describes the decryption that eDiscovery supports for email, email with attachments, and files hosted by SharePoint.

Expand table

Item type	Task	eDiscovery	Premium features enabled
Encrypted email	Search	Yes	Yes
Encrypted email	Decryption to .pst	No	Yes
Encrypted email	Decryption to file	No	Yes
Encrypted mail and attachment	Search	No	Yes (with Advanced indexing)[3]
Encrypted mail and attachment	Decryption to .pst	No	Yes
Encrypted mail and attachment	Decryption to file	No	Yes
File in SharePoint with MIP label	Search	No	Yes
File in SharePoint with MIP label	Decryption	No	Yes
File in SharePoint with other encryption[4]	Search, Decryption	No	No

Important

eDiscovery doesn't support legacy encryption protocols.

Decryption limitations with email and attachments

> **Important**
>
> Decryption and searchability are distinct. Email attachments encrypted with sensitivity labels are decrypted during preview and export, but the content in those attachments isn't indexed for search. Keyword searches don't match content inside these attachments. To ensure these attachments are captured, include partially indexed items in your collection and use review sets to analyze decrypted content. For more information, see Partially indexed items.

eDiscovery support for decryption of email messages and attachments is subject to the following limitations:

- Decryption isn't supported when the email or attachment encryption is applied in an external organization. eDiscovery only supports decryption for email and attachments that your organization encrypts.
- When decrypting emails or attachments, the owner of the mailbox must have access to view the encrypted content. Decryption isn't supported if the owner sends or forwards the emails or attachments to other recipients who can't view the encrypted content. Changes in the owner's groups or other organization permissions might also affect decryption support.
- PDFs generated from MIP-labeled Word documents (for example, a .docx file encrypted with a sensitivity label and then saved as .pdf) can only be decrypted when PDF encryption is enabled. PDF encryption and decryption control of the tenant is configured by running the following cmdlet:

  ```
  Set-IRMConfiguration -EnablePdfEncryption $true
  ```

If you don't set the configuration to *True*, eDiscovery doesn't decrypt the PDF generated from MIP labeled document.

Decryption limitations with sensitivity labels in SharePoint and OneDrive

eDiscovery doesn't support encrypted files in SharePoint and OneDrive when a sensitivity label that applies encryption is configured with the following setting:

- User access to the document has an expiration setting that is set to a value other than **Never**.

For more information about these settings, see the "Configure encryption settings" section in Restrict access to content by using sensitivity labels to apply encryption.

An eDiscovery search might return documents encrypted with the previous settings. This result might happen when a document property (such as the title, author, or modified date) matches the search criteria. Although these documents might be included in search results, you can't preview or review them. These documents remain encrypted when they're exported with premium eDiscovery features enabled.

Files in SharePoint with sensitivity labels configured for user-defined permissions are supported for eDiscovery decryption. For more information, see Learn more about support for labels configured for user-defined permissions.

> **Important**
>
> Decryption is supported for files that are locally encrypted and then uploaded to SharePoint or OneDrive as long as the user is signed in to a Microsoft 365 client when encrypting the file. For example, local files that are encrypted by the Microsoft Purview Information Protection client and then uploaded to Microsoft 365. If the encryption applied locally isn't by a user signed in to the Microsoft 365 client, the file isn't supported for decryption.

Decryption scope for nested and container files

eDiscovery decryption applies to root-level items only, such as files stored directly in SharePoint or OneDrive, direct email attachments, and the email messages themselves. If an encrypted file is nested inside a container format, eDiscovery doesn't recursively decrypt it.

For example:

- A MIP-labeled encrypted .pptx file stored directly in a SharePoint document library is decrypted as expected.
- The same MIP-labeled encrypted .pptx file inside a .zip archive stored in SharePoint isn't decrypted. The .zip archive is collected and processed, but encrypted files within the archive aren't decrypted.
- An encrypted attachment on an email message is decrypted. However, if that email message is attached as a nested .msg file inside another email, encrypted attachments

within the nested .msg file aren't decrypted.

This behavior applies to all container formats, including .zip archives and .msg files that contain encrypted items as nested content. Sensitivity label metadata for nested items might also not be available because label information is read from source-level properties of the root item.

Decrypting RMS-protected email messages and encrypted file attachments by using premium eDiscovery features

When you export rights-protected (RMS-protected) email messages that are included in the results of a search, the export process decrypts these messages. Members of the eDiscovery Manager role group have this decryption capability by default because the RMS Decrypt management role is assigned to this role group by default.

Keep the following things in mind when exporting encrypted email messages and attachments:

- To enable decryption of RMS-protected messages when you export them, you must export the search results as individual messages.
- Attachments encrypted separately from an email can be decrypted if the attachments are Microsoft 365 documents. For example, if a user encrypts a Word document and then attaches it to an email message that isn't encrypted, this attachment is decrypted.
- Attachments encrypted as part of the encryption of the associated email message are decrypted. For example, if a user creates an email message, attaches an unencrypted Word document, and then encrypts the message (including the attachment), this attachment is decrypted.
- The **ResultsLog** report identifies messages that are decrypted. This report contains a column named **Decode Status**, and a value of **Decoded** identifies the messages that are decrypted.
- In addition to decrypting file attachments when exporting search results, you can also preview the decrypted file when previewing search results. You can only view the rights-protected email message after you export it.
- To prevent someone from decrypting RMS-protected messages and encrypted file attachments, create a custom role group by copying the built-in eDiscovery Manager role group, then remove the RMS Decrypt management role from the custom role group. Add the person you don't want to decrypt messages as a member of the custom role group.

Notes

[1] When you enable premium eDiscovery features, you can only transparently decrypt content encrypted with RMS keys hosted in Microsoft 365. Double Key Encryption (DKE), Hold Your Own Key (HYOK), on-premises RMS, and similar options aren't supported. For more information, see Key types for the service.

[2] Email attachment decryption is supported only for Microsoft 365 documents (Microsoft Word, Excel, PowerPoint, etc.) and PDF files if PDF decryption is enabled. Encrypted email messages attached as attachments aren't supported.

[3] eDiscovery can decrypt and index encrypted files located on a local computer that a user copies to an email message as long as the encryption was applied when the user was signed in to a Microsoft 365 client. When you enable premium eDiscovery features, you need to advanced index encrypted email and attachments in recipient mailbox to decrypt them.

[4] You can only decrypt items labeled in SharePoint or uploaded to SharePoint after integration with sensitivity labels are enabled. These items must have labels with admin-defined permissions and have no expiration. You can't decrypt all other encrypted files in SharePoint. For more information, see Enable sensitivity labels for files in SharePoint and OneDrive.

You can't decrypt other documents, including:

- Files encrypted in the client and uploaded before sensitivity labels were integrated with SharePoint.
- Documents encrypted with legacy RMS templates and not labeled.
- Documents with user-defined permissions or with expiration settings (SMIME or other standards).

ⓘ **Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 04/28/2026

Set up compliance boundaries in eDiscovery

Compliance boundaries create logical boundaries within an organization that control the user content locations, such as mailboxes, OneDrive accounts, and SharePoint sites, that eDiscovery managers can search. Compliance boundaries also apply to searches in Data Security Investigations. Compliance boundaries control who can access eDiscovery cases used to manage the legal, human resources, or other investigations within your organization. Compliance boundaries control search scope within eDiscovery and DSI only. They don't restrict access in other Microsoft Purview or Defender solutions.

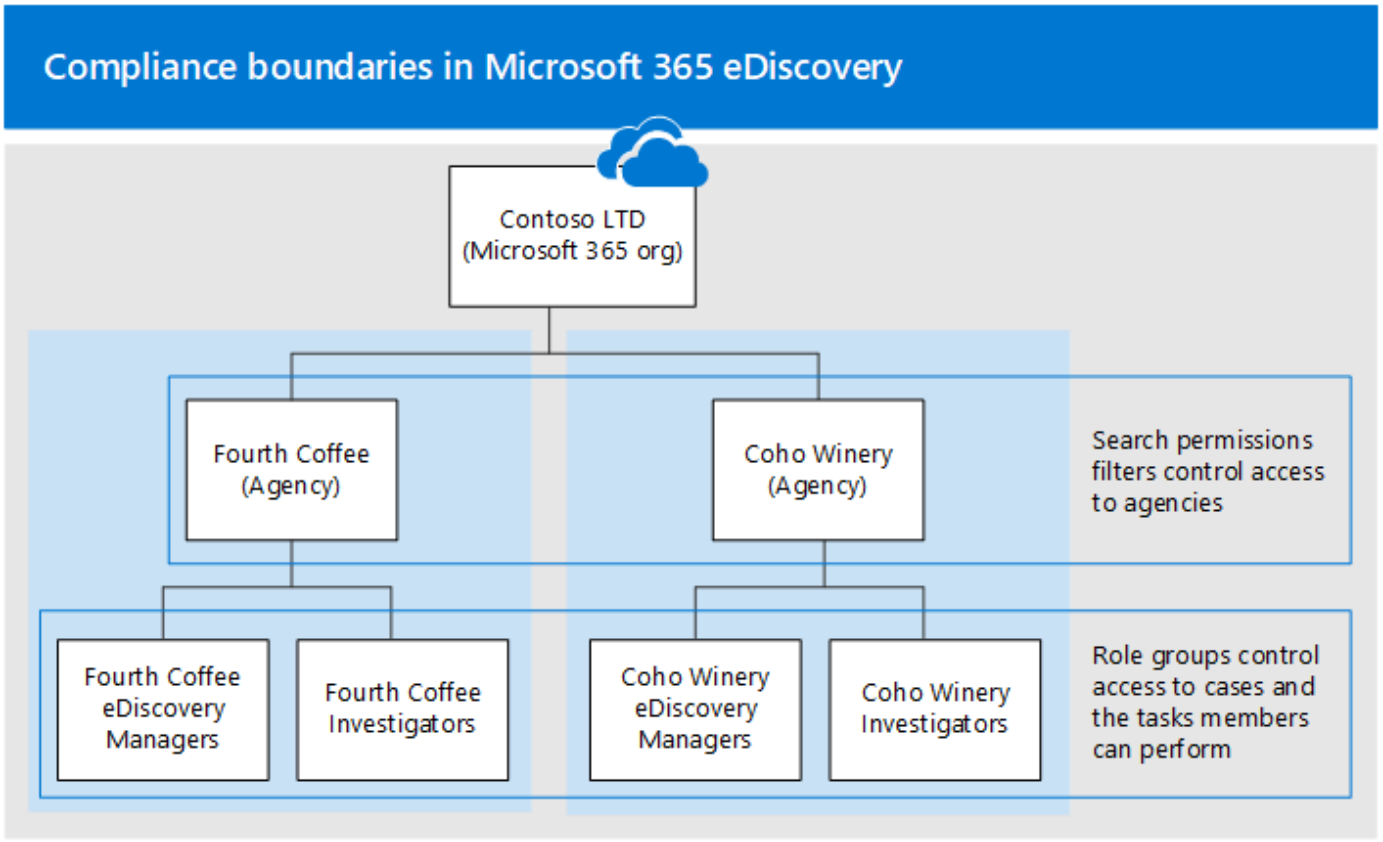
You often need compliance boundaries for multinational corporations that have to respect geographical borders and regulations, and for governments that are divided into different agencies. In Microsoft 365, compliance boundaries help you meet these requirements when performing content searches and managing investigations by using eDiscovery cases.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about Microsoft Security Copilot in Microsoft Purview.

> **Note**
>
> **SharePoint Embedded container support in eDiscovery (preview):** Compliance boundary enforcement isn't currently supported for SharePoint Embedded containers. For supported container scenarios and other limitations, see Data sources in eDiscovery.

Scenario example

The following illustration shows how compliance boundaries work.



In this example, Contoso LTD is an organization that consists of two subsidiaries, Fourth Coffee and Coho Winery. The business requires that eDiscovery managers and investigators can only search the Exchange mailboxes, OneDrive accounts, and SharePoint sites in their agency. Also, eDiscovery managers and investigators can only see eDiscovery cases in their agency, and they can only access the cases that they're a member of. In this scenario, investigators can't place content locations on hold or export content from a case. Here's how compliance boundaries meet these requirements.

- The search permissions filtering functionality for eDiscovery controls the content locations that eDiscovery managers and investigators can search. This control means eDiscovery managers and investigators in the Fourth Coffee agency can only search content locations in the Fourth Coffee subsidiary. The same restriction applies to the Coho Winery subsidiary.

- Role groups provide the following functions for compliance boundaries:
  - Control who can see the eDiscovery cases in the Microsoft Purview portal. This control means that eDiscovery managers and investigators can only see the eDiscovery cases in their agency.
  - Control who can assign members to an eDiscovery case. This control means eDiscovery managers and investigators can only assign members to cases that they themselves are a member of.
  - Control the eDiscovery-related tasks that members can perform by adding or removing roles that assign specific permissions.

- When a search permissions filter is applied to a role group, members of the role group can perform the following search-related actions as long as the permissions to perform an action are assigned to the role group:
  - Search for content
  - Preview search results
  - Export search results
  - Purge items returned by a search

Before you set up compliance boundaries

- Assign users an Exchange Online license. To verify assignments, use the Get-User cmdlet in Exchange Online PowerShell.

Set up compliance boundaries

Complete the following steps to set up compliance boundaries:

- Step 1: Identify a user attribute to define your agencies
- Step 2: Create a role group for each agency
- Step 3: Create a search permissions filter to enforce the compliance boundary
- Step 4: Create an eDiscovery case for intra-agency investigations

Step 1: Identify a user attribute to define your agencies

Choose an attribute to define your agencies. Use this attribute to create the search permissions filter that limits an eDiscovery manager to search only the content locations of users who are assigned a specific value for this attribute. For example, Contoso uses the **Department** attribute. The value for this attribute for users in the Fourth Coffee subsidiary is `FourthCoffee` and the value for users in Coho Winery subsidiary is `CohoWinery`. In Step 3, use this `attribute:value` pair (for example, *Department:FourthCoffee*) to limit the user content locations that eDiscovery managers can search.

Here are some examples of user attributes that you can use for compliance boundaries:

- Company
- CustomAttribute1 - CustomAttribute15
- Department
- Office
- CountryOrRegion (Two-letter country/region code)

> **Note**
>
> When selecting the attributes for the security filter, consider the attributes supported by different mailbox types in Exchange Online. Use CustomAttribute1 through CustomAttribute15 because all mailbox types in Exchange support these attributes.

Step 2: Create a role group for each agency

Create role groups in the Microsoft Purview portal that align with your agencies.

To create the role groups, go to **Settings** > **Roles and scopes** in the Microsoft Purview portal. Create a role group for each team in each agency that uses compliance boundaries and eDiscovery cases to manage investigations. You can use a single role group to assign both the eDiscovery roles and the security filter to its members. Membership in this role group manages both the roles assigned to the investigator and the content locations they can search.

> **Note**
>
> If you previously configured compliance boundaries using role groups with no roles attached (sometimes called 'empty role groups'), this configuration is no longer supported. Role groups used with search permissions filters must have at least one role assigned.

Use separate role groups. Use one role group to assign the security filter for the compliance boundary and another role group to assign the eDiscovery roles. By using two separate role groups, you manage the roles assigned to an investigator independently from the content locations they search. Use the role group for the compliance boundary only to apply the security filter to members. Use separate built-in (eDiscovery Manager) or custom role groups to assign roles to members. For more information about eDiscovery-related roles, see Assign eDiscovery permissions.

> **Important**
>
> Role groups with a security filter assigned must have at least one role added for it to take effect.

Using the Contoso compliance boundaries scenario, create four role groups and add the appropriate members to each one.

- Fourth Coffee eDiscovery Managers
- Fourth Coffee Investigators
- Coho Winery eDiscovery Managers
- Coho Winery Investigators

> **Important**
>
> If you add or remove a role from a role group that you added as a member of a case, the role group is automatically removed as a member of the case (or any case the role group is a member of). This behavior protects your organization from inadvertently providing extra permissions to members of a case. If you delete a role group, you remove it from all cases it was a member of.

Step 3: Create a search permissions filter to enforce the compliance boundary

After you create role groups for each agency, create search permissions filters that associate each role group with its specific agency and define the compliance boundary. You need to create one search permissions filter for each agency. For more information about creating security permissions filters, see Configure permissions filtering for eDiscovery.

Here's the syntax that's used to create a search permissions filter for compliance boundaries for the scenario in this article.

PowerShell

```
New-ComplianceSecurityFilter -FilterName <name of filter> -Users <role groups> -
Filters "Mailbox_<MailboxPropertyName>  -eq '<Value> '", "SiteContent_Path -like
'<SharePointURL>' -or SiteContent_Path -like '<OneDriveURL>'"
```

Here's a description of each parameter in the command:

- `FilterName`: Specifies the name of the filter. Use a name that describes or identifies the agency that the filter is used in.
- `Users`: Specifies the users or groups who get this filter applied to the search actions they perform. For compliance boundaries, this parameter specifies the role groups (that you created in Step 2) in the agency that you're creating the filter for. This parameter is a multivalue parameter so you can include one or more role groups, separated by commas.

- `Filters`: Specifies the search criteria for the filter. For compliance boundaries, you define the following filters. Each one applies to different content locations.

  - `Mailbox`: Specifies the mailboxes or OneDrive accounts that the role groups defined in the `Users` parameter can search. This filter allows members of the role group to search only the mailboxes or OneDrive accounts in a specific agency; for example, `"Mailbox_Department -eq 'FourthCoffee'"`.

    > **Note**
    >
    > Mailbox filters also apply to Microsoft 365 group mailboxes and connected SharePoint sites. To return results from a SharePoint site connected to a Microsoft 365 group, both the *Mailbox* and *SiteContent* filters must match to both the Microsoft 365 group mailbox and the associated SharePoint site.

- `SiteContent`: This filter includes two separate filters. The first `SiteContent_Path` specifies the SharePoint sites in the agency that the role groups defined in the `Users` parameter can search. For example, `SiteContent_Path -like 'https://contoso.sharepoint.com/sites/FourthCoffee*'`. The second `SiteContent_Path` filter (connected to the first `SiteContent_Path` filter by the `or` operator) specifies the agency's OneDrive domain (also called the *MySite* domain). For example, `SiteContent_Path -like 'https://contoso-my.sharepoint.com/personal*'`. You can also use the `Site_Path` filter in place of the `SiteContent` filter. The `Site` and `SiteContent` filters are interchangeable, and don't affect search permissions filters described in this article.

  > **Important**
  >
  > Why is the `SiteContent` filter for OneDrive included in the previous search permissions filter? Although the `Mailbox` filter applies to *both* mailboxes and OneDrive accounts, the inclusion of the SharePoint filter would exclude OneDrive accounts if you didn't also include the OneDrive `Site` filter. If the search permissions filter didn't include a SharePoint filter, then you wouldn't have to include a separate OneDrive filter because the Mailbox filter would include OneDrive accounts in the scope of the compliance boundary. In other words, a search permissions filter with only the `Mailbox` filter would include both mailboxes and OneDrive accounts.

Here are examples of the two search permissions filters that you create to support the Contoso compliance boundaries scenario. Both of these examples include a comma-separated filters list,

in which the mailbox and site filters are included in the same search permissions filter and are separated by a comma.

Fourth Coffee

PowerShell

```
New-ComplianceSecurityFilter -FilterName "Fourth Coffee Security Filter" -Users
"Fourth Coffee eDiscovery Managers", "Fourth Coffee Investigators" -Filters
"Mailbox_Department -eq 'FourthCoffee'", "SiteContent_Path -like
'https://contoso.sharepoint.com/sites/FourthCoffee*' -or SiteContent_Path -like
'https://contoso-my.sharepoint.com/personal*'"
```

Coho Winery

PowerShell

```
New-ComplianceSecurityFilter -FilterName "Coho Winery Security Filter" -Users "Coho
Winery eDiscovery Managers", "Coho Winery Investigators" -Filters "Mailbox_Department
-eq 'CohoWinery'", "SiteContent_Path -like
'https://contoso.sharepoint.com/sites/CohoWinery*' -or SiteContent_Path -like
'https://contoso-my.sharepoint.com/personal*'"
```

> **Note**
>
> The syntax for the `Filters` parameters in the previous examples includes a *filters list*. A filters list is a filter that includes a mailbox filter and a site path filter separated by a comma. In the previous example, notice that a comma separates `Mailbox` and `SiteContent` filters: `-Filters "Mailbox_<MailboxPropertyName> -eq '<Value> '", "SiteContent_Path -like '<SharePointURL>' -or SiteContent_Path -like '<OneDriveURL>'"`. When this filter is processed during the running of an eDiscovery search, two search permissions filters are created from the filters list: one mailbox filter and one SharePoint/OneDrive filter. An alternative to using a filters list would be to create two separate search permissions filters for each agency: one search permissions filter for the mailbox attribute and one filter for the SharePoint and OneDrive site attributes. In either case, the results are the same. Using a filters list or creating separate search permissions filters is a matter of preference.

How do the search permissions filters work in this scenario?

Here's how the search permission filters are applied for each agency in this scenario.

1. The `Mailbox` filter is first applied to define the content locations that eDiscovery managers can search. In this case, Coho Winery eDiscovery managers can only search the mailboxes and OneDrive accounts of users whose *Department* mailbox property has a value of **FourthCoffee**; Coho Winery eDiscovery managers can only search the mailboxes and OneDrive accounts of users whose *Department* mailbox property has a value of **CohoWinery**. The `Mailbox` filter is a *content location filter*, because it specifies the content locations that eDiscovery managers can search. In both filters, eDiscovery managers can only search content locations with a specific mailbox property value.

2. After the content locations that can be searched are defined, the next part of the filter defines the content that eDiscovery managers can search. The first `SiteContent` filter lets Fourth Coffee eDiscovery managers only search for documents that have a site path property that contains (or starts with) `https://contoso.sharepoint.com/sites/FourthCoffee`; Coho Winery eDiscovery managers can only search documents that have a site path property that contains (or starts with) `https://contoso.sharepoint.com/sites/CohoWinery`. Therefore, the two `SiteContent` filters are *content filters* because they define the content that can be searched for. In both filters, eDiscovery managers can only search for documents with a specific document property value. All SharePoint-related filters are content filters because searchable site properties are stamped on all documents. For more information, see Configure permissions filtering for eDiscovery.

   > **Note**
   >
   > Although the scenario in this article doesn't use them, you can also use mailbox content filters to specify the content that eDiscovery managers can search for. The syntax for mailbox content filters is `"MailboxContent_<property> -<comparison operator> '<value>'"`. You can create content filters based on date ranges, recipients, and domains or any searchable email property. For example, this filter would allow eDiscovery managers to only search for mail items sent or received by users in the contoso.com domain: `"MailboxContent_Participants -like 'contoso.com'"`. For more information about mailbox content filters, see Configure search permissions filtering.

3. The search permissions filter is joined to the search query by the **AND** Boolean operator. That means when an eDiscovery manager in one of the agencies runs an eDiscovery search, the items returned by the search must match the search query and the conditions defined in the search permissions filter.

Step 4: Create an eDiscovery case for intra-agency investigations

In the final step, create an eDiscovery case in the Microsoft Purview portal. Then, add the role group that you created in Step 2 as a member of the case. This approach provides two important benefits when using compliance boundaries:

- Only members of the role group added to the case can see and access the case in the Microsoft Purview portal. For example, if the Fourth Coffee Investigators role group is the only member of a case, then members of the Fourth Coffee eDiscovery Managers role group (or members of any other role group) can't see or access the case.
- When a member of the role group assigned to a case runs a search associated with the case, they can only search the content locations within their agency. This restriction is defined by the search permissions filter that you created in Step 3.

To create a case and assign members:

1. Go to **eDiscovery** in the Microsoft Purview portal and create a case.
2. In the list of cases, select the name of the case you created.
3. Add role groups as members to the case.

> ⓘ **Note**
>
> When adding a role group to a case, you can only add the role groups that you're a member of.

Search and export content in multigeo environments

Search permissions filters also let you control where content is routed for export and which datacenter can be searched when searching content locations in a SharePoint multigeo environment.

- **Export search results:** You can export the search results from Exchange mailboxes, SharePoint sites, and OneDrive accounts from a specific datacenter. This feature means that you can specify the datacenter location that search results are exported from.

  Use the *Region* parameter for **New-ComplianceSecurityFilter** or **Set-ComplianceSecurityFilter** cmdlets to create or change which datacenter the export is routed through.

Expand table

Parameter value	Datacenter location
NAM	North American (datacenters are in the US)
EUR	Europe
APC	Asia Pacific
CAN	Canada

- **Route content searches:** You can route the content searches of SharePoint sites and OneDrive accounts to a satellite datacenter. This feature means you can specify the datacenter location where searches are run.

  > **Note**
  >
  > If you're using premium eDiscovery features, the *Region* parameter doesn't control the region that data is exported from. Data is exported from the organization's central location. Also, searching for content in SharePoint and OneDrive isn't bound by the geographic location of datacenters. All datacenters are searched.

  Use one of the following values for the *Region* parameter to control the datacenter location that searches run in when searching SharePoint sites and OneDrive accounts.

Expand table

Parameter value	Datacenter routing locations for SharePoint
NAM	US
EUR	Europe
APC	Asia Pacific
CAN	US
AUS	Asia Pacific
KOR	The organization's default datacenter
GBR	Europe
JPN	Asia Pacific

Parameter value	Datacenter routing locations for SharePoint
IND	Asia Pacific
LAM	US
NOR	Europe
BRA	North American datacenters

If you don't specify the *Region* parameter for a search permissions filter, the organization's primary SharePoint region is searched. Search results are exported to the closest datacenter.

To simplify the concept, the *Region* parameter controls the datacenter that is used to search for content in SharePoint and OneDrive. This parameter doesn't apply to searching for content in Exchange because Exchange content searches aren't bound by the geographic location of datacenters. Also, the same *Region* parameter value might also dictate the datacenter that exports are routed through. This routing is often necessary to control the movement of data across geographic borders.

> **ⓘ Important**
>
> When creating security compliance filter and setting a specific GEO region: ESP, ITA, NZL and SWE aren't valid regions and generate an error redirecting to valid region values. This error includes the following valid regions: ARE, APC, AUS, AUT, BRA, CAN, CHE, CHL, DEU, EUR, FRA, GBR, IDN, IND, JPN, KOR, LAM, MYS, NAM, NOR, POL, QAT, ZAF.

Here are examples of using the *Region* parameter when creating search permission filters for compliance boundaries. This example assumes that the Fourth Coffee subsidiary is located in North America and that Coho Winery is in Europe.

PowerShell

```
New-ComplianceSecurityFilter -FilterName "Fourth Coffee Security Filter" -Users
"Fourth Coffee eDiscovery Managers", "Fourth Coffee Investigators" -Filters
"Mailbox_Department -eq 'FourthCoffee'", "SiteContent_Path -like
'https://contoso.sharepoint.com/sites/FourthCoffee*' -or SiteContent_Path -like
'https://contoso-my.sharepoint.com/personal*'" -Region NAM
```

PowerShell

```
New-ComplianceSecurityFilter -FilterName "Coho Winery Security Filter" -Users "Coho Winery eDiscovery Managers", "Coho Winery Investigators" -Filters "Mailbox_Department -eq 'CohoWinery'", "SiteContent_Path -like 'https://contoso.sharepoint.com/sites/CohoWinery*' -or SiteContent_Path -like 'https://contoso-my.sharepoint.com/personal*'" -Region EUR
```

Keep the following things in mind when searching and exporting content in multigeo environments.

- The *Region* parameter doesn't control searches of Exchange mailboxes. All datacenters are searched when you search mailboxes. To limit the scope of which Exchange mailboxes are searched, use the *Filters* parameter when creating or changing a search permissions filter.

- If it's necessary for an eDiscovery Manager to search across multiple SharePoint regions, you need to create a different user account for that eDiscovery manager to use in the search permissions filter to specify the region where the SharePoint sites or OneDrive accounts are located. For more information about setting this up, see Configure search permissions filtering in eDiscovery.

- When searching for content in SharePoint and OneDrive, the *Region* parameter directs searches to either the primary or satellite location where the eDiscovery manager conducts eDiscovery investigations. If an eDiscovery manager searches SharePoint and OneDrive sites outside of the region that's specified in the search permissions filter, no search results are returned.

- When exporting search results from eDiscovery, content from all content locations (including Exchange, Skype for Business, SharePoint, OneDrive, and other services that you can search by using the Content Search tool) are uploaded to the Azure Storage location in the datacenter that's specified by the *Region* parameter. This behavior helps organizations stay within compliance by not allowing content to be exported across controlled borders. If you don't specify a region in the search permissions filter, content is uploaded to the organization's primary datacenter.

  When exporting content with premium eDiscovery features enabled, you can't control where content is uploaded by using the *Region* parameter. Content is uploaded to an Azure Storage location in a datacenter in your organization's central location. For a list of geo locations based on your central location, see Microsoft 365 Multi-Geo eDiscovery configuration.

- You can edit an existing search permissions filter to add or change the region by running the following command:

PowerShell

```
Set-ComplianceSecurityFilter -FilterName <Filter name>  -Region <Region>
```

Use compliance boundaries for SharePoint hub sites

SharePoint hub sites often align with the same geographical or agency boundaries that eDiscovery compliance boundaries follow. You can use the site ID property of the hub site to create a compliance boundary. To do this, use the Get-SPOHubSite cmdlet in SharePoint Online PowerShell to get the SiteId for the hub site, then use this value for the department ID property to create a search permissions filter.

Use the following syntax to create a search permissions filter for a SharePoint hub site:

PowerShell

```
New-ComplianceSecurityFilter -FilterName <Filter Name> -Users <User or Group> -Filters
"Site_Departmentid -eq '{SiteId of hub site}'"
```

Here's an example of creating a search permissions filter for a hub site for the Coho Winery agency:

PowerShell

```
New-ComplianceSecurityFilter -FilterName "Coho Winery Hub Site Security Filter" -Users
"Coho Winery eDiscovery Managers", "Coho Winery Investigators" -Filters
"Site_Departmentid -eq '44252d09-62c4-4913-9eb0-a2a8b8d7f863'"
```

Compliance boundary considerations

Keep the following considerations in mind when managing eDiscovery cases and investigations that use compliance boundaries:

- When you create and run a search, you can select content locations that are outside of your agency. However, because of the search permissions filter, content from those locations isn't included in the search results.
- Compliance boundaries don't apply to holds in eDiscovery cases. That condition means an eDiscovery manager in one agency can place a user in a different agency on hold. However, the compliance boundary is enforced if the eDiscovery manager searches the content locations of the user who was placed on hold. That condition means the eDiscovery

manager can't search the user's content locations, even though they were able to place the user on hold.

- If you're assigned a search permissions filter - either mailbox or site filters - and you attempt to export partially indexed items from a search that includes all SharePoint sites in your organization, by using the scope option to include partially indexed items even from locations without search hits (see Export search results), the export contains partially indexed items from all sites, including those outside your assigned compliance boundary.

- When retrieving partially indexed items from sites that contain search hits, eDiscovery uses a prefix-based approach. For example, if a site `http://contoso.com/sales` is identified as a location matching the search query, the results might also include items from a similarly prefixed site such as `http://contoso.com/salesdepartment`, even if the `http://contoso.com/salesdepartment` site falls outside the configured compliance boundary. This behavior can occur when the filter isn't based on the site path. For instance, if the compliance boundary is defined as `Property -eq 'abc'`, a site that doesn't match the property value `'abc'` might be included if its path shares a prefix with a site that does match the property.

- Search permissions filters aren't applied to Exchange public folders.

- Supported search filters include `-and`, `-or`, `-like`, `-not`, `-eq`, and `-ne`. Don't use other exclusion filters, such as `-notlike` or `inotlike`, in a search permissions filter for a content-based compliance boundary. Using these exclusion filters can have unexpected results if content with recently updated attributes hasn't been indexed.

- Respect for compliance boundaries for OneDrive sites in searches might be impacted when:
  - The sites or associated mailboxes are moved or rehomed.
  - OneDrive ownership is reassigned, or more than one owner is added.
  - The OneDrive site is renamed or re-issued.
  - Moving a OneDrive site can change the ownership of the content and might include creating an arbitration mailbox. When you move these resources, there's a chance that content search results might be available across compliance boundaries. When a OneDrive site is reassigned, renamed, or assigned to more than one owner, it's possible that the content from these sites might be available across compliance boundaries.

- Compliance boundaries for mailboxes might be impacted when:
  - The mailbox isn't associated with a licensed user, which includes disabled or deleted users.
  - A mailbox isn't managed or synced from Microsoft Entra ID.

- Additionally, several types of mailboxes might produce content during search, regardless of their compliance boundaries. These mailbox types include:
  - EquipmentMailbox
  - GuestMailUser
  - LinkedMailbox
  - RoomList
  - RoomMailbox
  - SchedulingMailbox
  - SharedMailbox
  - SystemMailbox
  - TeamMailbox

- To make sure that the search respects the compliance boundaries as expected for these mailboxes, set attributes on shared mailboxes, equipment mailboxes, and other mailboxes by using the Set-Mailbox cmdlet or by using Invoke-ComplianceSecurityFilterAction to ensure the attribute is set correctly. If you properly configure the attribute and align it with the compliance boundary attribute, the mailbox respects the compliance boundary.

More information

- If you delicense or soft-delete a mailbox, the user is no longer within the compliance boundary. If you place a hold on the mailbox when you delete it, the content preserved in the mailbox is still subject to a compliance boundary or search permissions filter.
- If you implement compliance boundaries and search permissions filters for a user, don't delete the user's mailbox or their OneDrive account. If you delete a user's mailbox, also remove the user's OneDrive account since `mailbox_RecipientFilter` enforces the search permission filter for OneDrive.
- Compliance boundaries and search permissions filters depend on attributes stamped on content in Exchange, OneDrive, and SharePoint and the subsequent indexing of this stamped content.

Frequently asked questions

Who can create and manage search permissions filters by using the New-ComplianceSecurityFilter and Set-ComplianceSecurityFilter cmdlets?

To create, view, and modify search permissions filters in the Microsoft Purview portal, you need to have the *Organization Management* and *eDiscovery Administrator* roles assigned.

If an eDiscovery manager is assigned to more than one role group that spans multiple agencies, how do they search for content in one agency or the other?

The eDiscovery manager can add parameters to their search query that restrict the search to a specific agency. For example, if an organization specifies the **CustomAttribute10** property to differentiate agencies, they can append the following to their search query to search mailboxes and OneDrive accounts in a specific agency: `CustomAttribute10:<value>`.

What happens if I change the value of the attribute that's used as the compliance attribute in a search permissions filter?

It takes up to three days for a search permission filter to enforce the compliance boundary if you change the value of the attribute used in the filter. For example, let's say that a user in the Fourth Coffee agency is transferred to the Coho Winery agency. As a result, the value of the **Department** attribute on the user object changes from *FourthCoffee* to *CohoWinery*. In this situation, Fourth Coffee eDiscovery and investors get search results for that user for up to three days after the attribute is changed. Similarly, it takes up to three days before Coho Winery eDiscovery managers and investigators get search results for the user.

Can an eDiscovery manager see content from two separate compliance boundaries?

Yes, this capability is available when searching Exchange mailboxes by adding the eDiscovery manager to role groups that have visibility to both agencies. However, when searching SharePoint sites and OneDrive accounts, an eDiscovery manager can search for content in different compliance boundaries only if the agencies are in the same region or geo location.

> **Note**
>
> This limitation for sites doesn't apply when premium eDiscovery features are enabled because searching for content in SharePoint and OneDrive isn't bound by geographic location.

What happens to existing eDiscovery cases when a tenant moves between regions or across security boundaries?

eDiscovery cases, searches, holds, and review sets are created against the tenant's content location at the time of creation. When a tenant moves to a different region or to a different security boundary (for example, from commercial to GCC, GCC High, or DoD), existing artifacts remain bound to the origin location's data plane and might not be reusable from the destination.

Expand table

Move type	Impact on existing eDiscovery artifacts
Tenant relocation within the same security boundary (for example, a region move within commercial — NAM, EUR, APC)	Cases, searches, holds, and review sets created before the move might not be accessible from the destination region. Plan to export any required evidence before initiating the move, or expect to recreate artifacts after the move.
Tenant migration **across** security boundaries (commercial ↔ GCC, GCC High, or DoD)	Cases, searches, holds, and review sets created in the origin cloud **can't** be migrated and aren't accessible from the destination cloud. The artifacts continue to exist in the origin cloud's data plane but can't be opened, exported, or modified from the destination cloud. Plan to complete or export any active matters in the origin cloud before initiating the migration.

Symptoms after a move can include: existing cases appear in the portal but new case or search creation fails (often with a 500 error). If you encounter this error, contact Microsoft Support. There's a routing remediation, but for cross-cloud moves it requires accepting the loss of access to the origin-cloud artifacts in exchange for unblocking new artifact creation in the destination cloud.

For broader planning guidance on tenant migrations to the U.S. Government clouds, see Plan for Microsoft Purview DoD deployments, Plan for Microsoft Purview GCC High deployments, and Plan for Microsoft Purview GCC deployments. For tenant relocation and data residency planning, see Advanced Data Residency in Microsoft 365 and Plan for Microsoft 365 Multi-Geo.

Do search permissions filters work for eDiscovery case holds, Microsoft 365 retention policies, or DLP?

No, not at this time.

If I specify a region to control where content is exported, but I don't have a SharePoint organization in that region, can I still search SharePoint?

If the region specified in the search permissions filter doesn't exist in your organization, the default region is searched.

What is the maximum number of search permissions filters that I can create in an organization?

There's no limit to the number of search permissions filters that you can create in an organization. However, a search query can have a maximum of 100 conditions. In this case, a condition is defined as something that's connected to the query by a Boolean operator (such as

AND, **OR**, and **NEAR**). The limit of the number of conditions includes the search query itself plus all search permissions filters that are applied to the user who runs the search. Therefore, the more search permissions filters you have (especially if these filters are applied to the same user or group of users), the better the chance of exceeding the maximum number of conditions for a search.

To understand how this limit works, you need to understand that a search permissions filter is appended to the search query when a search is run. A search permissions filter is joined to the search query by the **AND** Boolean operator. The query logic for the search query and a single search permissions filter looks like this:

text

```
<SearchQuery> AND <PermissionsFilter>
```

Multiple search permissions filters are combined together by the **OR** Boolean operator, and then those conditions are connected to the search query by the **AND** operator.

The query logic for the search query and multiple search permissions filters looks like this:

text

```
<SearchQuery> AND (<PermissionsFilter1> OR <PermissionsFilter2> OR
<PermissionsFilter3>...)
```

It's possible the search query itself might consist of multiple conditions connected by Boolean operators. Each condition in the search query also counts against the 100-condition limit.

Also, the number of search permissions filters appended to a query depends on the user who is running the search. When a specific user runs a search, the search permissions filters that are applied to the user (which is defined by the *Users* parameter in the filter) are appended to the query. Your organization could have hundreds of search permissions filters, but if you apply more than 100 filters to the same users, then it's likely the 100-condition limit is exceeded when those users run searches.

There's one more thing to keep in mind about the condition limit. The number of specific SharePoint sites that are included in the search query or search permissions filters also count against this limit.

To prevent your organization from reaching the conditions limit, keep the number of search permissions filters in your organization as few as possible to meet your business requirements.

How do mailbox filters impact searches to OneDrive and M365 Group connected SharePoint sites?

OneDrive sites and Microsoft 365 Group connected SharePoint sites are linked to the associated user or Microsoft 365 Group mailbox in Exchange. For security filters that include both a Mailbox and SiteContent filter, to return results from the OneDrive or Microsoft 365 Group connected SharePoint sites both must be a match. The Mailbox filter is evaluated against the connected user or Microsoft 365 Group Mailbox first before evaluating the SharePoint and OneDrive items against the SiteContent filter.

What is an arbitration mailbox in the context of OneDrive sites?

For active and licensed users that have both an Exchange Mailbox and OneDrive account, there's a link between the mailbox and their OneDrive site. This means that the Mailbox filter assigned to a compliance boundary is evaluated against the user's mailbox to check for a match before results are returned.

When this link is lost between a mailbox and a OneDrive site, the OneDrive is associated with an arbitration mailbox instead. This arbitration mailbox doesn't have any of the mailbox properties from the original mailbox applied to it.

How do I check the status of a OneDrive site to determine if it is associated with an arbitration mailbox?

To review a specific OneDrive site and check if it's in arbitration or linked to a user's mailbox, use the Invoke-ComplianceSecurityFilterAction cmdlet.

Use the following syntax to check the status of a OneDrive site:

PowerShell

```
Invoke-ComplianceSecurityFilterAction -Action GetStatus -PropertyName <Boundary
Property> -SiteUrl <SharePoint/OneDrive URL>
```

For a OneDrive site that's not in arbitration, the commands return the *BoundaryType* value as UserMailbox, as shown in the following example:

PowerShell

```
PS C:\> Invoke-ComplianceSecurityFilterAction -action GetStatus -PropertyName
"CustomAttribute10" -siteurl "https://contoso-
my.sharepoint.com/personal/lila_contoso_onmicrosoft_com"
```

```
SiteUrlOrEmailAddress : https://contoso-
my.sharepoint.com/personal/lila_contoso_onmicrosoft_com
BoundaryType          : UserMailbox
BoundaryInstruction   : Set via Set-Mailbox
BoundaryObjectId      : SPO_4531d99e-21df-4903-a8c1-76cc3ec82685@SPO_525ef3f5-fbf7-
4f91-bb4a-e3c24116c8a9
BoundaryStatus        : Success
BoundaryProperty      : CustomAttribute10
BoundaryPropertyValue : Dept-CH
```

For a OneDrive site that's in arbitration, the commands return an empty or null value for the *BoundaryType* value, as shown in the following example.

PowerShell

```
PS C:\> Invoke-ComplianceSecurityFilterAction -action GetStatus -PropertyName
"CustomAttribute10" -siteurl "https://contoso-
my.sharepoint.com/personal/lila_contoso_onmicrosoft_com"

SiteUrlOrEmailAddress : https://contoso-
my.sharepoint.com/personal/lila_contoso_onmicrosoft_com
BoundaryType          :
BoundaryInstruction   :
BoundaryObjectId      : SPO_4531d99e-21df-4903-a8c1-76cc3ec82685@SPO_525ef3f5-fbf7-
4f91-bb4a-e3c24116c8a9
BoundaryStatus        : Success
BoundaryProperty      : CustomAttribute10
BoundaryPropertyValue :
```

How does a OneDrive site in arbitration affect results returned by an eDiscovery Manager with compliance boundaries applied to them?

This behavior depends on how the security filter is configured for the compliance boundary. When a OneDrive site is in arbitration, a Mailbox filter doesn't include it in the returned results.

For security filters that use only a Mailbox filter, items from the OneDrive are returned within the results of a search because the mailbox filter doesn't apply in this scenario.

For security filters that use both a mailbox filter and a SiteContent filter, results are returned if the *SiteContent* filter matches the items in the OneDrive instance that's in arbitration. If there's no match, the results aren't returned. For example, if the SiteContent_Path filter is used as separate agencies have a unique OneDrive URL due to the deployment of Microsoft 365 Multi-Geo, then the OneDrive items' path filters in or out results from the OneDrive in arbitration.

How can I update a OneDrive site in arbitration so that Mailbox filters continue to apply?

Use the Invoke-ComplianceSecurityFilterAction cmdlet to set the CustomAttribute1 through to CustomAttribute15 values on OneDrive instances that link to an arbitration mailbox.

Use the following syntax to set the status of a OneDrive site:

PowerShell

```
Invoke-ComplianceSecurityFilterAction -Action Set -PropertyName <Boundary Property> -SiteUrl <SharePoint/OneDrive URL>
```

After you set it on the OneDrive, it honors the Mailbox filters that use the CustomAttribute property set on the OneDrive.

What scenarios cause a OneDrive site to link to an arbitration mailbox?

Many scenarios can cause a OneDrive site to lose its link to the user's mailbox and link to an arbitration mailbox. See the Compliance Boundary Limitations section for more information. As a result, when you move, change, or reassign a OneDrive, use the Invoke-ComplianceSecurityFilterAction cmdlet to validate that the relevant Mailbox filter behaves as expected.

As part of user deprovisioning processes, run the Invoke-ComplianceSecurityFilterAction cmdlet against the user's OneDrive to set the required CustomAttribute value.

Why should I use CustomAttribute1 through CustomAttribute15 for Mailbox security filters?

Although other Mailbox attributes work with security filters, use the CustomAttributes when you define the Mailbox security filter for the following reasons:

- All Exchange Online mailboxes support these attributes.
- The attributes stamp directly on the mailbox, not the associated user object, so they remain set in scenarios where there's no associated user object.
- The Invoke-ComplianceSecurityFilterAction cmdlet only supports setting CustomAttribute on OneDrive instances linked to an arbitration mailbox. **My organization previously configured compliance boundaries using role groups with no roles attached (empty role groups), how should I update this congifuration?**

First identify roles used in your New-ComplianceSecurityFilter commands by running Get-ComplianceSecurityFilter | FL FIlterName, Users

Then, add a minimal role (such as View-Only Audit Logs) to each role group, or restructure to use the two-role group structure recommended where you set up one role group for the security

filter (with at least one role) and one role group for eDiscovery permissions.

Finally, verify your compliance boundaries are functioning as expected by running a test search.

Note: adding or removing roles from a role group that is a member of a case will automatically remove that role group from the case. Re-add the role group to the affected cases after makeing role changes.

My search returns results outside the expected compliance boundary, what should I do?

Verify the role group used in your security filter has at least one role assigned. Role groups with no roles assigned don't activate the security filter. Confirm the mailbox attribute is set on all target mailboxes, including shared mailboxes and group mailboxes. Check for OneDrive sites in arbitration status as sites linked to arbitration mailboxes don't inherit mailbox attributes. Allow up to three days for attribute changes to take effect in search permissions filters.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 09/19/2026

Configure search permissions filtering in eDiscovery

Use search permissions filtering to let an eDiscovery manager search only a subset of mailboxes and sites in your organization. You can also use permissions filtering to let that same eDiscovery manager search only for mailbox or site content that meets a specific search criteria.

For example, you might want an eDiscovery manager to search only the mailboxes of users in a specific location or department. Create a filter that uses a supported recipient filter to limit which mailboxes a specific user or group of users can search. You can also create a filter that specifies what mailbox content a user can search for. This filter uses a searchable message property. Similarly, you can let an eDiscovery manager search only specific SharePoint sites in your organization. Create a filter that limits which site can be searched. You can also create a filter that specifies what site content can be searched. This filter uses a searchable site property.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about **Microsoft Security Copilot in Microsoft Purview**.

Understand search permissions filters

The Microsoft Purview portal applies search permissions filters when you search for content. When you apply a search permissions filter to a specific user, that user can perform the following search-related actions:

- Search for content
- Preview search results
- Export search results
- Purge items returned by a search

Use search permissions filtering to create logical boundaries (called *compliance boundaries*) within an organization that control the user content locations (such as mailboxes, SharePoint sites, and OneDrive accounts) that specific eDiscovery managers can search. For more information, see Set up compliance boundaries for eDiscovery investigations.

The following four cmdlets in Security & Compliance PowerShell let you configure and manage search permissions filters:

- New-ComplianceSecurityFilter
- Get-ComplianceSecurityFilter
- Set-ComplianceSecurityFilter
- Remove-ComplianceSecurityFilter

Requirements to configure permissions filtering

Before you configure permissions filtering, make sure the following requirements are met:

- To run the compliance security filter cmdlets, you need to be a member of the *Organization Management* role group and have the *eDiscovery administrator* role assigned in the Microsoft Purview portal. For more information, see Permissions in the Security & Compliance Center.
- You need to connect to both Exchange Online and Security & Compliance PowerShell to use the compliance security filter cmdlets. These cmdlets require access to mailbox properties, which is why you need to connect to Exchange Online PowerShell. See the steps in Connect to Exchange Online and Security & Compliance PowerShell in a single session.
- See the Additional considerations for search permissions filtering section for additional information about search permissions filters.
- Search permissions filtering applies to inactive mailboxes, which means you can use mailbox and mailbox content filtering to limit who can search an inactive mailbox. See the Additional considerations for search permissions filtering section for additional information about permissions filtering and inactive mailboxes.
- You can't use search permissions filtering to limit who can search public folders in Exchange.

> **Note**
>
> You can create unlimited search permission filters in an organization. However, a search query can have a maximum of 100 conditions. In this case, a condition is something that's connected to the query by a Boolean operator, such as **AND**, **OR**, and **NEAR**. The limit for the number of conditions includes the search query itself plus all search permissions filters that the user who runs the search applies. Therefore, the more search permissions filters you have - especially if these filters are applied to the same user or group of users - the better the chance of exceeding the maximum number of conditions for a search. To prevent your organization from reaching the conditions limit, keep the number of search permissions

filters in your organization as few as possible to meet your business requirements. For more information, see Set up compliance boundaries for eDiscovery investigations.

Connect to Exchange Online and Security & Compliance PowerShell in a single session

Before you can successfully run the script in this section, you need to download and install the Exchange Online PowerShell module. For information, see Install and maintain the Exchange Online PowerShell module.

1. Save the following text to a Windows PowerShell script file by using a filename suffix of **.ps1**. For example, you could save it to a file named **ConnectEXO-SCC.ps1**.

   PowerShell

   ```
   Import-Module ExchangeOnlineManagement
   $UserCredential = Get-Credential
   Connect-ExchangeOnline -Credential $UserCredential -ShowBanner:$false
   Connect-IPPSSession -Credential $UserCredential
   $Host.UI.RawUI.WindowTitle = $UserCredential.UserName + " (Exchange Online + Security & Compliance)"
   ```

2. On your local computer, open Windows PowerShell, go to the folder where the script that you created in the previous step is located, and then run the script. For example:

   PowerShell

   ```
   .\ConnectEXO-SCC.ps1
   ```

How do you know if this worked? After you run the script, cmdlets from Exchange Online PowerShell and Security & Compliance PowerShell are available. If you don't receive any errors, you connected successfully. A quick test is to run Exchange Online PowerShell and Security & Compliance PowerShell cmdlets. For example, you can run **Get-Mailbox** and **Get-ComplianceSearch**.

For troubleshooting PowerShell connection errors, see:

- Connect to Exchange Online PowerShell
- Connect to Security & Compliance PowerShell

Create a search permissions filter with New-ComplianceSecurityFilter

Use the **New-ComplianceSecurityFilter** cmdlet to create a search permissions filter. Here's the basic syntax for this cmdlet:

PowerShell

```
New-ComplianceSecurityFilter -FilterName <name of filter> -Users <user or role group>
-Filters <filter>
```

The following sections describe the parameters for this cmdlet. You need all parameters to create a search permissions filter.

Specify the FilterName parameter

Use the *FilterName* parameter to specify the name of the permissions filter. Use this name to identify a filter when using the **Get-ComplianceSecurityFilter**, **Set-ComplianceSecurityFilter**, and **Remove-ComplianceSecurityFilter** cmdlets.

Specify the Filters parameter

Use the *Filters* parameter to specify the search criteria for the compliance security filter. You can create three different types of filters:

- **Mailbox or OneDrive filtering:** This filter specifies the mailboxes and OneDrive accounts that assigned users (specified by the *Users* parameter) can search. This filter is called a *content location* filter because it defines the content locations that a user can search. The syntax for this filter is **Mailbox_***MailboxPropertyName*, where *MailboxPropertyName* specifies a mailbox property used to scope the mailboxes and OneDrive accounts that can be searched. For example, the mailbox filter `"Mailbox_CustomAttribute10 -eq 'OttawaUsers'"` allows the user assigned this filter to search only the mailboxes and OneDrive accounts that have the value "OttawaUsers" in the CustomAttribute10 property.

  You can use any supported filterable recipient property for the *MailboxPropertyName* property in a mailbox or OneDrive filter. The following table lists four commonly used recipient properties used to create a mailbox or OneDrive filter. The table also includes an example of using the property in a filter.

Expand table

Property name	Example
Alias	`"Mailbox_Alias -like 'v-'"`
Company	`"Mailbox_Company -eq 'Contoso'"`
CountryOrRegion	`"Mailbox_CountryOrRegion -eq 'United States'"`
Department	`"Mailbox_Department -eq 'Finance'"`

> **Note**
>
> The Mailbox filter applies to user and group mailboxes. The filter also includes the associated OneDrive site for the user mailbox and SharePoint sites associated to the group mailbox.

- **Mailbox content filtering:** This filter applies to the content that can be searched. This filter is called a *content filter* because it specifies the mailbox content or searchable email properties the assigned users can search for. The syntax for this filter is **MailboxContent_**_SearchablePropertyName_, where *SearchablePropertyName* specifies a Keyword Query Language (KeyQL) property that can be specified in a search. For example, the mailbox content filter `"MailboxContent_Recipients -like 'contoso.com'"` allows the user assigned this filter to only search for messages sent to recipients in the contoso.com domain. For a list of searchable email properties, see Keyword queries and search conditions for eDiscovery.

  > **Important**
  >
  > Within a *single component* of the `Filters` parameter, you can't mix a mailbox filter and a mailbox content filter (for example, you can't combine `Mailbox_*` and `MailboxContent_*` together with `-and` inside one quoted string). To combine these types in a single security filter, put each in its own component of a **filters list that combines filter types** — separate them with a comma. A single component can contain a more complex query of the *same* type — for example, `"Mailbox_CustomAttribute10 -eq 'FTE' -and Mailbox_MemberOfGroup -eq '$($DG.DistinguishedName)'"`.

- **Site and site content filtering:** You can use two SharePoint- and OneDrive-related filters to specify what site or site content the assigned users can search.
  - **Site**_*QueryableSiteProperty*
  - **SiteContent**_*QueryableSiteProperty*

  These two filters are interchangeable. For example, `"Site_Path -like 'https://contoso.sharepoint.com/sites/doctors'"` and `"SiteContent_Path -like 'https://contoso.sharepoint.com/sites/doctors'"` return the same results. For a list of searchable site properties, see Keyword queries and search conditions for eDiscovery. For a more complete list, see Overview of crawled and managed properties in SharePoint. Properties marked with a **Yes** in the **Queryable** column can be used to create a site or site content filter.

  > **ⓘ Important**
  >
  > Setting up a site filter with one of the supported properties doesn't mean the site property in the filter propagates to all documents on that site. This condition means the user is still responsible for populating the specific property fields associated with the documents on that site in order for the site filter to work and capture the right content. For example, if the user has a security filter "Site_RefineableString00 -eq 'abc'" applied and then the user runs a search using keyword query "xyz". The security filter gets appended to the query and the actual query running would be "xyz **AND RefineableString0:'abc'**". The user needs to ensure that documents on the site have values in the RefineableString00 field as "abc". If not, the search query won't return any results.

Keep the following considerations in mind when configuring the *Filters* parameter for search permissions filters:

- Unlike mailboxes, there isn't a content location filter for sites even though the *Site* filter looks like a location filter. All filters for SharePoint and OneDrive are content filters (which is also why *Site_* and *SiteContent_* filters are interchangeable) because site-related properties like *Path* are stamped directly on the documents. This design is a result of the way that SharePoint is designed. In SharePoint, there isn't a "site object" with properties, like there is with Exchange mailboxes. Therefore, the *Path* property is stamped on the document and contains the URL of the site where the document is located. This design is why a *Site* filter is considered a content filter and not a content location filter.

- You have to create a search permissions filter to explicitly prevent users from searching content locations in a specific service (such as preventing a user from searching any Exchange mailbox or any SharePoint site). In other words, creating a search permissions filter that allows a user to search all SharePoint sites in the organization doesn't prevent that user from searching mailboxes. For example, to allow SharePoint admins to only search SharePoint sites, you have to create a filter that prevents them from searching mailboxes. Similarly, to allow Exchange admins to only search mailboxes, you have to create a filter that prevents them from searching sites.

Specify the Users parameter

Use the *Users* parameter to specify the users who get the filter applied to their searches. Identify users by their alias or primary SMTP address. You can specify multiple values separated by commas, or you can assign the filter to all users by using the value **All**.

Use the *User* parameter to specify a Microsoft Purview portal role group. This approach lets you create a custom role group and then assign that role group a search permissions filter. For example, suppose you have a custom role group for eDiscovery managers for the U.S. subsidiary of a multinational corporation. You can use the *Users* parameter to specify this role group (by using the Name property of the role group) and then use the *Filter* parameter to allow only mailboxes in the U.S. to be searched. You can't specify distribution groups with this parameter.

Use a filters list to combine filter types

A *filters list* is a `Filters` value that contains more than one filter component, with the components separated by commas. A single security filter can mix two different content-location filters (`Mailbox_*` and `Site_*`), or it can mix a content-location filter with a content filter on the same workload (`Mailbox_*` plus `MailboxContent_*`, or `Site_*` plus `SiteContent_*`). Each component is a separate quoted string in the `Filters` parameter.

PowerShell

```
-Filters "Mailbox_CustomAttribute10 -eq 'OttawaUsers'", "SiteContent_Path -like
'https://contoso.sharepoint.com/sites/doctors'"
```

How a filters list composes depends on the **classes** of the components. There are two filter classes:

- **Content-location filters** — `Mailbox_*` and `Site_*`. These filters define *which* content locations the assigned user can search.

- **Content filters** — `MailboxContent_*` and `SiteContent_*`. These filters define *what* content can be returned from the locations allowed by the corresponding location filter.

The composition rules are:

Expand table

Components in the filters list	Logical composition
Two content-location filters for different workloads (`Mailbox_*` and `Site_*`)	**OR** across workloads. The user is granted access to mailboxes matching the `Mailbox_*` predicate *or* sites matching the `Site_*` predicate. This composition is the original "comma = OR" behavior described in earlier versions of this article.
A content-location filter and a content filter for the *same* workload (`Mailbox_*` plus `MailboxContent_*`, or `Site_*` plus `SiteContent_*`)	**Layered AND**. The location filter narrows which content locations are searchable, and the content filter further restricts which items can be returned from those locations. Both conditions must be satisfied for an item to be returned.
Multiple components of the *same* class (for example, two `Mailbox_*` components, or two `MailboxContent_*` components)	**OR** within the class. Multiple mailbox rules combine to expand the allowed location set; multiple mailbox-content rules combine to expand the allowed content set. The combined content set is then AND'd with the user's search query (see How does search permissions filtering work?).

The preceding example contains a `Mailbox_*` filter and a `SiteContent_*` filter - two location filters for different workloads - so the comma is OR (mailboxes matching the first predicate *or* SharePoint content matching the second).

Compare that composition to the following example, which contains a `Mailbox_*` location filter and a `MailboxContent_*` content filter for the same workload - the comma here is a layered AND (mailboxes that are members of the group *and*, within those mailboxes, only items that aren't Microsoft 365 Copilot interactions):

PowerShell

```
-Filters "Mailbox_MemberOfGroup -eq 'CN=Litigation Hold
Group,OU=Groups,DC=contoso,DC=com'", "MailboxContent_ItemClass -notlike
'IPM.SkypeTeams.Message.Copilot*'"
```

An alternative to using a filters list is to create two separate search permissions filters and assign them to the same user - the behavior described in How does search permissions filtering work? applies (multiple security filters are OR'd together at evaluation time). When the two components fall into different *classes* (location vs. content), the filters list and the two-separate-filters

approaches are *not* equivalent - prefer the filters list to keep the layered AND semantics within a single filter.

Keep the following things in mind about using a filters list:

- You must use a filters list to create a filter that includes a `Mailbox` filter and a `MailboxContent` filter (or a `Site` filter and a `SiteContent` filter). Within a single quoted component, you can't combine the two classes with `-and`.
- Each component of a filters list can contain a complex filter syntax. For example, the mailbox and site filters can contain multiple filters separated by an **-or** operator:

  PowerShell

  ```
  -Filters "Mailbox_Department -eq 'CohoWinery' -or Mailbox_CustomAttribute10 -eq 'CohoUsers'", "SiteContent_Path -like 'https://contoso.sharepoint.com/sites/CohoWinery*'"
  ```

Examples of creating search permissions filters

The following examples show how to use the **New-ComplianceSecurityFilter** cmdlet to create a search permissions filter.

The first example enables members of the "US Discovery Managers" role group to search only the mailboxes and OneDrive accounts in the United States.

PowerShell

```
New-ComplianceSecurityFilter -FilterName USDiscoveryManagers  -Users "US Discovery Managers" -Filters "Mailbox_CountryOrRegion  -eq 'United States'"
```

The second example enables the user annb@contoso.com to perform search actions only for mailboxes and OneDrive accounts in Canada. This filter contains the three-digit numeric country code for Canada from ISO 3166-1.

PowerShell

```
New-ComplianceSecurityFilter -FilterName CountryFilter  -Users annb@contoso.com -Filters "Mailbox_CountryCode  -eq '124'"
```

The third example enables the users *donh* and *suzanf* to search only the mailboxes and OneDrive accounts that have the value 'Marketing' for the CustomAttribute1 mailbox property.

PowerShell

```
New-ComplianceSecurityFilter -FilterName MarketingFilter  -Users donh,suzanf -Filters
"Mailbox_CustomAttribute1  -eq 'Marketing'"
```

The fourth example enables members of the "Fourth Coffee eDiscovery Managers" role group to search only the mailboxes and OneDrive accounts that have the value 'FourthCoffee' for the Department mailbox property. The filter also enables the role group members to search for documents in the Fourth Coffee SharePoint site.

PowerShell

```
New-ComplianceSecurityFilter -FilterName "Fourth Coffee Security Filter" -Users
"Fourth Coffee eDiscovery Managers", "Fourth Coffee Investigators" -Filters
"Mailbox_Department -eq 'FourthCoffee'", "SiteContent_Path -like
'https://contoso.sharepoint.com/sites/FourthCoffee*' -or SiteContent_Path -like
'https://contoso-my.sharepoint.com/personal*'"
```

> **Note**
>
> In the previous example, you need to include an additional site content filter (`SiteContent_Path -like 'https://contoso-my.sharepoint.com/personal*'`) so that role group members can search for documents in OneDrive accounts. If you don't include this filter, the filter only enables role group members to search for documents located in `https://contoso.sharepoint.com/sites/FourthCoffee`.

The fifth example enables members of the eDiscovery Manager role group to search only the mailboxes and OneDrive accounts of members of the Ottawa Users distribution group. The `Get-DistributionGroup` cmdlet in Exchange Online PowerShell is used to find the members of the Ottawa Users group.

PowerShell

```
$DG = Get-DistributionGroup "Ottawa Users"
```

PowerShell

```
New-ComplianceSecurityFilter -FilterName DGFilter  -Users eDiscoveryManager -Filters
"Mailbox_MemberOfGroup -eq '$($DG.DistinguishedName)'"
```

The sixth example prevents any user from performing search actions on the mailboxes and OneDrive accounts of members of the Executive Team distribution group. That means users can delete content from these mailboxes. The `Get-DistributionGroup` cmdlet in Exchange Online PowerShell is used to find the members of the Executive Team group.

PowerShell

```
$DG = Get-DistributionGroup "Executive Team"
```

PowerShell

```
New-ComplianceSecurityFilter -FilterName NoExecutivesPreview  -Users All -Filters
"Mailbox_MemberOfGroup -ne '$($DG.DistinguishedName)'"
```

The seventh example enables members of the OneDrive eDiscovery Managers custom role group to only search for content in OneDrive accounts in the organization.

PowerShell

```
New-ComplianceSecurityFilter -FilterName OneDriveOnly  -Users "OneDrive eDiscovery
Managers" -Filters "SiteContent_Path -like 'https://contoso-
my.sharepoint.com/personal*'"
```

The eighth example restricts the user to performing search actions only on email messages sent during the calendar year 2015.

PowerShell

```
New-ComplianceSecurityFilter -FilterName EmailDateRestrictionFilter -Users
donh@contoso.com -Filters "MailboxContent_Received -ge '01-01-2015' -and
MailboxContent_Received -le '12-31-2015'"
```

Similar to the previous example, the ninth example restricts the user to performing search actions only on documents that were last changed sometime in the calendar year 2015.

PowerShell

```
New-ComplianceSecurityFilter -FilterName DocumentDateRestrictionFilter -Users
donh@contoso.com -Filters "SiteContent_LastModifiedTime -ge '01-01-2015' -and
SiteContent_LastModifiedTime -le '12-31-2015'"
```

The tenth example prevents members of the "OneDrive Discovery Managers" role group from performing search actions on any mailbox in the organization.

PowerShell

```
New-ComplianceSecurityFilter -FilterName NoEXO -Users "OneDrive Discovery Managers" -
Filters "Mailbox_Alias -notlike '*'"
```

The eleventh example prevents anyone in the organization from performing search actions on email messages that were sent or received by janets or sarad.

PowerShell

```
New-ComplianceSecurityFilter -FilterName NoSaraJanet -Users All -Filters
"MailboxContent_Participants -notlike 'janets@contoso.onmicrosoft.com' -and
MailboxContent_Participants -notlike 'sarad@contoso.onmicrosoft.com'"
```

The twelfth example uses a filters list to combine mailbox and site filters. In this example, the mailbox filter is a content location filter and the site filter is a content filter.

PowerShell

```
New-ComplianceSecurityFilter -FilterName "Coho Winery Security Filter" -Users "Coho
Winery eDiscovery Managers", "Coho Winery Investigators" -Filters "Mailbox_Department
-eq 'CohoWinery'", "SiteContent_Path -like
'https://contoso.sharepoint.com/sites/CohoWinery*'"
```

View search permissions filters with Get-ComplianceSecurityFilter

Use **Get-ComplianceSecurityFilter** to return a list of search permissions filters. Use the *FilterName* parameter to return information for a specific search filter.

Modify a search permissions filter with Set-ComplianceSecurityFilter

Use **Set-ComplianceSecurityFilter** to modify an existing search permissions filter. The following sections describe the parameters for this cmdlet. The only required parameter is *FilterName*.

FilterName

Use the *FilterName* parameter to specify the name of the permissions filter.

Users

Use the *Users* parameter to specify the users who get this filter applied to their searches. Because this parameter is a multivalue property, specifying a user or group of users overwrites the existing list of users. See the following examples for the syntax to add and remove selected users.

You can also use the *Users* parameter to specify a Microsoft Purview portal role group. This role group lets you create a custom role group and then assign that role group a search permissions filter. For example, suppose you have a custom role group for eDiscovery managers for the U.S. subsidiary of a multinational corporation. Use the *Users* parameter to specify this role group (by using the Name property of the role group) and then use the *Filters* parameter to allow only mailboxes in the U.S. to be searched. You can't specify distribution groups with this parameter.

Filters

Use the *Filters* parameter to specify the search criteria for the compliance security filter. You can create three different types of filters:

- **Mailbox and OneDrive filtering:** This filter specifies the mailboxes and OneDrive accounts that assigned users (specified by the *Users* parameter) can search. The syntax for this filter is **Mailbox**_ *MailboxPropertyName*. The *MailboxPropertyName* specifies a mailbox property used to scope the mailboxes that can be searched. For example, the mailbox filter `"Mailbox_CustomAttribute10 -eq 'OttawaUsers'"` allows the user assigned this filter to search only the mailboxes that have the value "OttawaUsers" in the CustomAttribute10 property. For the *MailboxPropertyName* property, you can use any supported filterable recipient property. For a list of supported properties, see Filterable properties for the -RecipientFilter parameter.

- **Mailbox content filtering:** This filter applies to the content that can be searched. It specifies the mailbox content that assigned users can search. The syntax for this filter is **MailboxContent**_*SearchablePropertyName*. The *SearchablePropertyName* specifies a Keyword Query Language (KQL) property that you can specify in a search. For example, the mailbox content filter `"MailboxContent_Recipients -like 'contoso.com'"` allows the user assigned this filter to only search for messages sent to recipients in the contoso.com domain. For a list of searchable email properties, see Use the condition builder to create search queries in eDiscovery.

- **Site and site content filtering:** Use two SharePoint and OneDrive for Business site-related filters to specify what site or site content the assigned users can search:
  - **Site**_ *SearchableSiteProperty*
  - **SiteContent**_*SearchableSiteProperty*

These two filters are interchangeable. For example, `"Site_Path -like 'https://contoso.spoppe.com/sites/doctors*'"` and `"SiteContent_Path -like 'https://contoso.spoppe.com/sites/doctors*'"` return the same results. For a list of searchable site properties, see Overview of crawled and managed properties in SharePoint. Properties marked with a **Yes** in the **Queryable** column can be used to create a site or site content filter.

Examples of changing search permissions filters

These examples show how to use the **Get-ComplianceSecurityFilter** and **Set-ComplianceSecurityFilter** cmdlets to add or remove a user from the existing list of users that the filter is assigned to. When you add or remove users from a filter, specify the user by using their SMTP address.

This example adds a user to the filter.

PowerShell

```
$filterusers = Get-ComplianceSecurityFilter -FilterName OttawaUsersFilter
```

PowerShell

```
$filterusers.users.add("pilarp@contoso.com")
```

PowerShell

```
Set-ComplianceSecurityFilter -FilterName OttawaUsersFilter -Users $filterusers.users
```

This example removes a user from the filter.

PowerShell

```
$filterusers = Get-ComplianceSecurityFilter -FilterName OttawaUsersFilter
```

PowerShell

```
$filterusers.users.remove("annb@contoso.com")
```

PowerShell

```
Set-ComplianceSecurityFilter -FilterName OttawaUsersFilter -Users $filterusers.users
```

Remove a search permissions filter with Remove-ComplianceSecurityFilter

Use the **Remove-ComplianceSecurityFilter** cmdlet to delete a search filter. Use the *FilterName* parameter to specify the filter you want to delete.

Additional considerations for search permissions filtering

How does search permissions filtering work? When you run a search, the permissions filter is added to the search query. The **AND** Boolean operator connects the permissions filter to the search query. The query logic for the search query and the permissions filter looks like this:

```
text
<SearchQuery> AND <PermissionsFilter>
```

For example, you have a permissions filter that allows Bob to perform all search actions on the mailboxes of members of the Workers distribution group. Then Bob runs a search on all mailboxes in the organization with the search query `sender:jerry@adatum.com`. Because the permissions filter and the search query are logically combined by an **AND** operator, the search returns any message sent by jerry@adatum.com to any member of the Workers distribution group.

What happens if you have multiple search permissions filters? In a search query, multiple permissions filters are combined by **OR** Boolean operators. So results are returned if any of the filters are true. In a search, all filters (combined by **OR** operators) are then combined with the search query by the **AND** operator.

```
text
<SearchQuery> AND  (<PermissionsFilter1> OR <PermissionsFilter2> OR
<PermissionsFilter3>)
```

Let's take the previous example, where a search filter allows Bob to search only the mailboxes of the members of the Workers distribution group. Then we create another filter that prevents Bob from searching Phil's mailbox ("Mailbox_Alias -ne 'Phil'"). And let's also assume that Phil is a member of the Workers group. When Bob runs a search (from the previous example) on all mailboxes in the organization, search results are returned for Phil's mailbox even though you

applied filter to prevent Bob from searching Phil's mailbox. This is because the first filter, which allows Bob to search the Workers group, is true. And because Phil is a member of the Workers group, Bob can search Phil's mailbox.

How are filter components within a single security filter combined? Filters within a single security filter are evaluated in two layers:

1. **Mailbox filters first.** All `Mailbox_*` components in the filter are evaluated and combined with **OR** to produce the set of allowed mailbox locations. If multiple `Mailbox_*` rules are present (whether across components of a filters list or within a single component using `-or`), they combine as OR to expand the location set.
2. **Mailbox content filters next, scoped to those locations.** All `MailboxContent_*` components are then applied to content within the allowed mailbox locations from step 1. If multiple `MailboxContent_*` rules are present, they combine with **OR**. The combined `MailboxContent_*` predicate is then **AND**'d to the user's search query.

The same two-layer model applies to SharePoint and OneDrive: `Site_*` location filters compose with OR to define the allowed sites, and `SiteContent_*` filters compose with OR and are then AND'd to the user's search query, scoped to those sites.

```text
(<MailboxRule1> OR <MailboxRule2> ...)           // Layer 1: location scope
  AND
((<MailboxContentRule1> OR <MailboxContentRule2> ...) AND <SearchQuery>)   // Layer 2:
content
```

For example, the filter `-Filters "Mailbox_MemberOfGroup -eq 'CN=Litigation Hold Group,...'", "MailboxContent_ItemClass -notlike 'IPM.SkypeTeams.Message.Copilot*'"` means: only mailboxes that are members of the group are searchable, and within those mailboxes, the user can return any item *except* Microsoft 365 Copilot interactions. Items that match the user's search query, are in an allowed mailbox, and pass the content filter are returned.

Does search permissions filtering work for inactive mailboxes? Yes, you can use mailbox and mailbox content filters to limit who can search inactive mailboxes in your organization. Like a regular mailbox, an inactive mailbox must be configured with the recipient property that's used to create a permissions filter. If necessary, you can use the **Get-Mailbox -InactiveMailboxOnly** command to display the properties of inactive mailboxes. For more information, see Create and manage inactive mailboxes.

Does search permissions filtering work for public folders? No. Search permissions filtering can't be used to limit who can search public folders in Exchange. For example, items in public folder locations can't be excluded from the search results by a permissions filter.

Does allowing a user to search all content locations in a specific service also prevent them from searching content locations in a different service? No. You have to create a search permissions filter to explicitly prevent users from searching content locations in a specific service (such as preventing a user from searching any Exchange mailbox or any SharePoint site). In other words, creating a search permissions filter that allows a user to search all SharePoint sites in the organization doesn't prevent that user from searching mailboxes. For example, to allow SharePoint admins to only search SharePoint sites, you have to create a filter that prevents them from searching mailboxes. Similarly, to allow Exchange admins to only search mailboxes, you have to create a filter that prevents them from searching sites.

Do search permissions filters count against search query character limits? Yes. Search permissions filters count against the character limit for search queries. For more information, see Limits in eDiscovery.

What is the maximum number of search permissions filters that can be created in an organization?

There's no limit to the number of search permissions filters that can be created in an organization. However, a search query can have a maximum of 100 conditions. In this case, a condition is defined as something that's connected to the query by a Boolean operator (such as **AND**, **OR**, and **NEAR**). The limit of the number of conditions includes the search query itself plus all search permissions filters that are applied to the user who runs the search. Therefore, the more search permissions filters you have (especially if these filters are applied to the same user or group of users), the better the chance of exceeding the maximum number of conditions for a search.

To understand how this limit works, you need to understand that a search permissions filter is appended to the search query when a search is run. A search permissions filter is joined to the search query by the **AND** Boolean operator. The query logic for the search query and a single search permissions filter looks like this:

```text
<SearchQuery> AND <PermissionsFilter>
```

Multiple search permissions filters are combined together by the **OR** Boolean operator, and then those conditions are connected to the search query by the **AND** operator.

The query logic for the search query and multiple search permissions filters looks like this:

text

```
<SearchQuery> AND (<PermissionsFilter1> OR <PermissionsFilter2> OR
<PermissionsFilter3>...)
```

It's possible the search query itself might consist of multiple conditions connected by Boolean operators. Each condition in the search query also counts against the 100-condition limit.

Also, the number of search permissions filters appended to a query depends on the user who is running the search. When a specific user runs a search, the search permissions filters that are applied to the user (which is defined by the *Users* parameter in the filter) are appended to the query. Your organization could have hundreds of search permissions filters, but if more than 100 filters are applied to the same users, then it's likely the 100-condition limit is exceeded when those users run searches.

There's one more thing to keep in mind about the condition limit. The number of specific SharePoint sites that are included in the search query or search permissions filters also count against this limit.

To prevent your organization from reaching the conditions limit, keep the number of search permissions filters in your organization as few as possible to meet your business requirements.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/29/2026

Cloud attachments in eDiscovery

Summarize this article for me

This article explains how eDiscovery handles cloud (modern or linked) attachments versus traditional (classic) attachments. Cloud attachments are links to files stored in SharePoint or OneDrive rather than embedded copies. Because a cloud attachment points to a live file stored in Microsoft 365 services that you can modify after sharing, eDiscovery offers options to collect the live file, the version that you shared, or multiple versions to preserve context. You can collect cloud attachments as part of the export or add to review set workflow.

Collect cloud attachments

eDiscovery supports the following options for the collection of modern attachments:

- Live version of the linked document at the moment of collection.
- Version at the time of sharing.
- Other version options (collect all versions or specific shared versions).

Regardless of the location of the modern attachment, the process attributes the *Custodian* property to the owner of the message. For example, if User A shares a linked attachment with User B that resides in User C's OneDrive and you collect the message from User B's mailbox, the linked attachment belongs to User B (custodian) and not User C.

To enable the collection of cloud attachments, select the **Access links (cloud attachments) in messages** option so linked content is included in Add to review set and export workflows.

> **Important**
>
> Including attachments increases the size and quantity of export results.

Once the workflow processes completes, you can group the modern attachment and parent message or email together by using **Group by conversations and related items** filter and the **Is modern attachment** column value.

Considerations and unsupported scenarios

Before collecting cloud attachments, review the following considerations and scenarios:

- Plain text emails aren't supported. Only HTML emails are processed.
- Encrypted emails and messages don't support modern attachment extraction.

- Non-clickable links aren't supported. Links that aren't clickable aren't processed. Most links in emails are rewritten to clickable safelinks by the Exchange backends for opt-in tenants.
- Malformed URLs aren't supported. URLs that don't parse correctly aren't stamped.
- Maximum message body length. If the body length is greater than 100,000 characters, any content (modern attachment link) beyond this limit isn't considered.
- Maximum URL length. URLs longer than 2,048 characters are skipped.
- Maximum total property size. If the total property size exceeds 1 MB, it's dropped.
- Maximum number of links per message. Only the first 50 links in a message are processed. Other links beyond 50 are ignored.
- If a message is forwarded and the modern attachment link is expired or stale, the extraction doesn't work. Similarly, if the modern attachment file is renamed or moved, the link isn't extracted.
- Cloud attachment links in forwarded or replied emails aren't extracted from the quoted portion of the original message. Only the original email retains the cloud attachment association. If you need to collect a cloud attachment that was shared in the original email, search for and collect the original message rather than a forward or reply.
- Teams shared channel messages and posts don't support modern attachment link extraction. Users need to include the shared channel site and include the file manually as part of their collection.
- Cloud attachments are treated as part of the parent's message, not separate entities.
  - If an email or Teams message matches the condition of a search, its linked cloud attachments are retrieved - even if the file doesn't meet the condition of the search.
  - If an email or Teams message matches the condition of a search, its linked cloud attachments are retrieved - even if the file resides outside the compliance boundary.
- Advanced indexing isn't applied to cloud attachment files referenced within emails or messages. This limitation means the content of the cloud attachment file itself isn't processed by advanced indexing or OCR.

Collecting the version shared in cloud attachments

By default, the eDiscovery workflow for collecting cloud attachments only includes adding the most current version of a cloud attachment to a review set or for export. This workflow means the version you collect and add to a review set or export might be different from the version that was originally shared in the cloud attachment. So, content that was present in the cloud attachment at the time it was shared might be removed and doesn't exist in the current version that's added to the review set or exported.

Use retention labels to preserve the version of a document at the time when it was shared as a cloud attachment. To do this, create a retention label, choose the option to apply the label to cloud attachments, and then automatically apply the label to documents stored in SharePoint

and OneDrive. After you set up this configuration, a copy of a document is created at the time when the file is shared. Also, if the document is modified and shared again as a cloud attachment, the modified version is also preserved. If the file is modified and shared again, a new copy of the file as a new version is preserved.

Preserving the shared versions of cloud attachments helps you scope the preservation and collection of potentially relevant content to the specific version of the document that you shared rather than the current live version. After you implement this retention solution, both the current live version of a cloud attachment and the version that you shared in the cloud attachment are collected and added to a review set or exported.

For instructions on setting up a retention label and automatically applying it to cloud attachments, see Auto-apply labels to cloud attachments.

Additionally, the current live version and the version that you shared have the same **GroupID** property value, which is the same as the **GroupID** for the parent object (such as an email message or a Teams chat conversation). This property value lets you group cloud attachments with the item in which you shared them.

After you implement the retention label and auto-apply the label to SharePoint documents, select the option to collect cloud attachments when adding items to a review set. When you collect the cloud attachments, both the current live version and the version that was originally shared are added to the review set or exported.

View in eDiscovery reports

eDiscovery includes a Process manager that lists all processes scoped for each area. Each process in the Process manager list contains a detailed report .zip file that contains detailed information about the process once the process completes.

Summary CSV report

All reporting packages include a `Summary-*the date/time of the report*.csv` file. This file includes information related to cloud attachments. This report includes information about the following items:

- Cloud attachments
- Cloud attachments versions
- Cloud attachment versions at time of sharing

For more information about these areas, see Summary CSV: Information.

Items CSV report

Reporting packages also include an Items CSV report. In the `items.csv` file, you can:

- Filter on all modern attachments by using the property **Is modern attachment** = `True`.
- Correlate modern attachment with its parent email or message through group ID and the property **Has attachment** = `True`.

Last updated on 02/27/2026

Double-byte character language support in eDiscovery

eDiscovery supports double-byte character set languages (these include Simplified Chinese, Traditional Chinese, Japanese, and Korean, which are collectively known as *CJK* languages) for the following scenarios in a review set:

- When you search for data.
- When you tag documents.
- When you analyze case data using near duplicate detection, email threading, and themes analytics.

> **Tip**
>
> Get started with Microsoft Security Copilot to explore new ways to work smarter and faster using the power of AI. Learn more about **Microsoft Security Copilot in Microsoft Purview**.

Limitations for CJK characters

The following features for CJK characters aren't currently supported in eDiscovery:

- Detecting CJK characters in file metadata.
- Filtering, sorting, or searching by document languages in a review set.
- Optical Character Recognition (OCR) for image files.
- Email files (such as *.eml and *.msg) in Annotate view.
- Search hit highlighting in Text view.
- Query-based holds for cases.
- The Relevance module used to analyze data.

Search for items containing CJK characters

You can use CJK characters when searching for content in eDiscovery. CJK characters are supported for all search operators and search conditions, including the Boolean operators **AND**, **OR**, **NOT**, and **NEAR**.

If you're certain that content locations or items contain CJK characters, but searches don't return any results, select the query language-country/region icon and select the corresponding language-country/region culture code value for the search. The default language/region is neutral.

When searching for phrases in CJK languages, break the phrase into individual words or characters and combine them using the AND operator. CJK phrases aren't natively recognized as single searchable units, so this approach helps to ensure more accurate search results. For example, instead of searching for `(SubjectTitle:Teamsのチャットメッセージ)`, use `(SubjectTitle:Teams) AND (SubjectTitle:の) AND (SubjectTitle:チャット) AND (SubjectTitle:メッセージ)`.

Search multiple languages at the same time

How you search multiple languages at the same time depends on your search scenario.

- When you query data in a review set with premium eDiscovery features enabled, you can search for multiple languages. For review set queries, you don't need to specify a document language. eDiscovery automatically detects document languages when you add content to a review set. This detection helps you optimize your query results in a review set.
- When you create a search to collect data, create separate searches for each language you're targeting. For example, if you're searching for a document that contains both Chinese and Korean, select Chinese for your first search and select Korean for your second search.

Changing your display language to Chinese, Japanese, or Korean

For information about how to change display language and time zone, see How to set language and region settings for Office 365.

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/29/2026

Use Microsoft Purview APIs for eDiscovery

The Microsoft Purview Application Programming Interface (API) for eDiscovery in Microsoft Graph enables your organization to automate repetitive tasks and integrate with your existing eDiscovery tools to build repeatable workflows that industry regulations might require. This article provides guidance on how to configure the required prerequisites to enable access to the Microsoft Purview APIs for eDiscovery. This guidance is based on using app-only access to the APIs, with either a client secret or a self-signed certificate to authenticate the requests.

Microsoft Purview APIs

The Microsoft Purview APIs for eDiscovery include two separate APIs:

- **Microsoft Graph**: Part of the `Microsoft.Graph.Security` namespace and used for working with eDiscovery cases.
- **Microsoft Purview eDiscovery API**: Used exclusively to programmatically download packages created when exporting from searches and review sets in eDiscovery.

eDiscovery APIs in Microsoft Graph support eDiscovery cases with and without premium features enabled. Delegated authentication supports core eDiscovery capabilities such as cases, holds, searches, and exports. App-only authentication is available for cases with premium features enabled and supports advanced operations such as review sets, tagging, and analytics.

For a list of supported API calls within the Microsoft Graph calls, see Use the Microsoft Purview eDiscovery API.

Application access to data

Before you can make any calls to the Microsoft Purview APIs for eDiscovery, you must first register an app in the Microsoft Identity Platform, Entra ID.

An application can access data in two ways:

- **Delegated access**: An app acting on behalf of a signed-in user.
- **App-only access**: An app acting with its own identity.

For more information about access scenarios, see Authentication and authorization basics.

> **Important**
>
> App-only authentication requires an eDiscovery case with premium features enabled. Delegated authentication is available for **core eDiscovery capabilities** including cases, legal holds, searches, and exports. Premium API operations (review sets, tagging, analytics) require premium features enabled regardless of authentication method. For more information about subscription requirements, see **subscription requirements for eDiscovery**.

Microsoft Graph API

Prerequisites for Microsoft Graph API

Implementing app-only accessinvolves registering an app in Azure portal, creating client secret/certificates, assigning API permissions, setting up a service principal, and then using app-only access to call Microsoft Graph APIs. To register an app, create client secret/certificates and assign API permissions the account must be a *Cloud Application Administrator*.

For more information about registering an app in the Azure portal, see Register an application with the Microsoft identity platform.

Granting tenant-wide admin consent for Microsoft Purview eDiscovery API application permissions requires you to sign in as a user that is authorized to consent on behalf of your organization. For more information, see Grant tenant-wide admin consent to an application.

Setting up a service principal requires the following prerequisites:

- A machine with the ExchangeOnlineManagement module installed.
- An account that has the *Role Management* role assigned in Microsoft Purview.

For detailed steps on implementing app-only access for eDiscovery, see Set up app-only access for Microsoft Purview eDiscovery.

Connecting to Microsoft Graph API using app-only access

Use the Connect-MgGraph cmdlet in PowerShell to authenticate and connect to Microsoft Graph using the app-only access method. This cmdlet enables your app to interact with Microsoft Graph securely and enables you to explore the Microsoft Purview eDiscovery APIs.

Connecting via client secret

To connect using a client secret, update and run the following example PowerShell code.

PowerShell

```
$clientSecret = "<client secret>" ## Update with client secret added to the registered
app
$appID = "<APP ID>" ## Update with Application ID of registered/Enterprise app
$tenantId = "<Tenant ID>" ## Update with tenant ID
$ClientSecretPW = ConvertTo-SecureString "$clientSecret" -AsPlainText -Force
$clientSecretCred = New-Object System.Management.Automation.PSCredential -ArgumentList
("$appID", $clientSecretPW)
Connect-MgGraph -TenantId "$tenantId" -ClientSecretCredential $clientSecretCred
```

Connecting via certificate

To connect using a certificate, update and run the following example PowerShell code.

PowerShell

```
$certPath = "Cert:\currentuser\my\<xxxxxxxxxx>" ## Update with the cert thumbnail
$appID = "<APP ID>" ## Update with Application ID of registered/Enterprise app
$tenantId = "<Tenant ID>" ## Update with tenant ID
$ClientCert = Get-ChildItem $certPath
Connect-MgGraph -TenantId $TenantId -ClientId $appId -Certificate $ClientCert
```

Invoke Microsoft Graph API calls

After you connect, you can start making calls to the Microsoft Graph API.

For example, you can list the eDiscovery cases within the tenant by using the ediscoveryCases API. The guidance for each operation lists the following information:

- Permissions required to make the API call
- HTTP request and method
- Request header and body information
- Response
- Examples (HTTP, C#, CLI, Go, Java, PHP, PowerShell, Python)

Because you're connected through the Microsoft Graph PowerShell module, you can use either the HTTP or PowerShell method.

First, let's look at the PowerShell example.

```
PS C:\Users        Get-MgSecurityCaseEdiscoveryCase

CreatedDateTime     Description                                DisplayName                          Id
---------------     -----------                                -----------                          --
14/02/2025 10:27:47                                            CE008 - Compliance Boundaries        2386be8...
17/01/2025 10:29:57                                            CE004 - Keyword Testing              dcede9f...
14/01/2025 08:59:24                                            CE003 - Message Headers              9d0aa11...
03/03/2025 17:25:40                                            CE012 - APAC Discovery               21ea867...
14/01/2025 13:29:16                                            HK001 - Identifier                   ef98a4b...
16/01/2025 13:29:30 Testing adding custodians                  MG001 - Custodian COllection         3bbb619...
23/01/2025 14:07:28                                            CE005 - SPO Export Testing           0d2f5c9...
23/01/2025 14:10:31                                            HK005 - SPO Export Testing           1f92354...
03/03/2025 17:24:08                                            CE011 - EMEA Request                 b5df12f...
15/01/2025 09:20:18                                            HK003 - Team Meeting Collection      8b61d90...
03/03/2025 17:40:28                                            CE016 - Sales                        3939a9e...
14/02/2025 10:30:13                                            CE009 - Compliance Boundary - No Roles eb44b12...
13/01/2025 17:13:09                                            CE002 - Encryption                   b1bad4d...
17/02/2025 16:33:57                                            HK008 - Compliance Boundary          e2edafe...
03/03/2025 17:30:20                                            CE015 - Marketing                    9d9da8c...
26/02/2025 16:17:07                                            CE010 - Communication Compliance     c6570da...
03/03/2025 16:41:26                                            Test                                 f145669...
14/01/2025 14:20:47                                            HK002 - Sensitivity Labels           5fa5d87...
03/03/2025 17:22:22                                            HK009 - Bates vs Contoso             69f4758...
10/02/2025 13:01:47                                            HK007 - Demo Case                    7af7146...
04/02/2025 10:40:07                                            CE007 - 1st Party Apps               09bfb77...
04/02/2025 10:43:01                                            HK006 - 1st Party Apps               91b08b2...
20/01/2025 11:38:10                                            HK004 - Polls Collection             5b4d3e7...
03/03/2025 17:27:42                                            CE013 - Australia Data Disclosure    a976b8e...
03/03/2025 17:28:50                                            CE014 - Engineering                  f8d6290...
```

As you can see, it returns a list of all the cases within the tenant. When delving deeper into a case, it's important to record the case *ID*. You need this ID for future API calls.

Now, let's look at an HTTP example. You use the Invoke-MgGraphRequest cmdlet to make the call by using PowerShell.

First, store the URL in a variable:

```
$uri = "https://graph.microsoft.com/v1.0/security/cases/ediscoveryCases"
```

Then use the Invoke-MgGraphRequest cmdlet to make the API call.

PowerShell

```
Invoke-MgGraphRequest -Method Get -Uri $uri
```

As you can see from the following output, you need to extract the values from the returned response.

```
PS C:\Users        $uri = "https://graph.microsoft.com/v1.0/security/cases/ediscoveryCases"
PS C:\Users        Invoke-MgGraphRequest -Method Get -Uri $uri

Name                           Value
----                           -----
@odata.context                 https://graph.microsoft.com/v1.0/$metadata#security/cases/ediscoveryCases
@odata.count                   27
value                          {System.Collections.Hashtable, System.Collections.Hashtable, System.Collections.Hashtable, System.Collections.Hashtable...}
```

You can save the *Value* elements of the response to a new variable by using the following command.

PowerShell

```
$cases = (Invoke-MgGraphRequest -Method Get -Uri $uri).value
```

```
PS C:\Users [redacted]  $cases = (Invoke-MgGraphRequest -Method Get -Uri $uri).value
PS C:\Users [redacted]  $cases.count
27
PS C:\Users [redacted]  $cases | Get-Member


   TypeName: System.Collections.Hashtable

Name                MemberType            Definition
----                ----------            ----------
Add                 Method                void Add(System.Object key, System.Object val
Clear               Method                void Clear(), void IDictionary.Clear()
Clone               Method                System.Object Clone(), System.Object ICloneab
Contains            Method                bool Contains(System.Object key), bool IDicti
ContainsKey         Method                bool ContainsKey(System.Object key)
ContainsValue       Method                bool ContainsValue(System.Object value)
CopyTo              Method                void CopyTo(array array, int arrayIndex), voi
Equals              Method                bool Equals(System.Object obj)
GetEnumerator       Method                System.Collections.IDictionaryEnumerator GetE
GetHashCode         Method                int GetHashCode()
GetObjectData       Method                void GetObjectData(System.Runtime.Serializati
GetType             Method                type GetType()
OnDeserialization   Method                void OnDeserialization(System.Object sender),
Remove              Method                void Remove(System.Object key), void IDictior
ToString            Method                string ToString()
Item                ParameterizedProperty System.Object Item(System.Object key) {get;se
Count               Property              int Count {get;}
IsFixedSize         Property              bool IsFixedSize {get;}
IsReadOnly          Property              bool IsReadOnly {get;}
IsSynchronized      Property              bool IsSynchronized {get;}
Keys                Property              System.Collections.ICollection Keys {get;}
SyncRoot            Property              System.Object SyncRoot {get;}
Values              Property              System.Collections.ICollection Values {get;}
```

```
PS C:\Users [redacted]  $cases

Name                           Value
----                           -----
externalId
id                             2386be89-6ae9-404f-9a33-0013314cc801
displayName                    CE008 - Compliance Boundaries
closedDateTime
status                         active
description
closedBy                       {user, application}
lastModifiedBy                 {user, application}
lastModifiedDateTime           14/02/2025 10:27:47
createdDateTime                14/02/2025 10:27:47
externalId
id                             dcede9fc-e4a4-4a66-b5a2-070736800764
displayName                    CE004 - Keyword Testing
closedDateTime
status                         active
description
```

This command returns a collection of hash tables. Optionally, you can run a small bit of PowerShell code to convert the hash tables into PowerShell objects for easier use with cmdlet

parameters such as *format-table* and *format-list*.

PowerShell

```
$CasesAsObjects = @()
foreach($i in $cases) {$CasesAsObjects += [pscustomobject]$i}
$CasesAsObjects | ft displayname,id,status
```

```
PS C:\Users      $CasesAsObjects = @()
PS C:\Users      foreach($i in $cases) {$CasesAsObjects += [pscustomobject]$i}
PS C:\Users      $CasesAsObjects | ft displayname,id,status

displayName                          id                                   status
-----------                          --                                   ------
CE008 - Compliance Boundaries        2386be89-6ae9-404f-9a33-0013314cc801 active
CE004 - Keyword Testing              dcede9fc-e4a4-4a66-b5a2-070736800764 active
CE003 - Message Headers              9d0aa119-76dc-4afc-9ae6-0d96fcfe8c9d closed
CE012 - APAC Discovery               21ea8671-ba19-4421-bb36-13ed3b117e31 active
HK001 - Identifier                   ef98a4b5-c2dc-4ac5-97e5-157c59a174a9 active
MG001 - Custodian COllection         3bbb6199-0cc8-4641-8a34-21236e5331a3 active
CE005 - SPO Export Testing           0d2f5c9b-b8de-4b0c-ab24-21c54e2c2665 active
HK005 - SPO Export Testing           1f92354f-e6a5-41ac-ac66-25233154c17e active
CE011 - EMEA Request                 b5df12f9-059c-4664-8d61-3abfb2efd84e active
HK003 - Team Meeting Collection      8b61d90a-45f8-4634-b557-3edf7eaf07ff active
CE016 - Sales                        3939a9ea-ba24-45e3-bd37-400060053213 active
CE009 - Compliance Boundary - No Roles eb44b12b-5944-4a24-b2b4-539dfcc5d967 active
CE002 - Encryption                   b1bad4db-ac47-4e47-b868-54a107cbcbf0 active
HK008 - Compliance Boundary          e2edafee-84f8-4ba7-bf85-58df63d2fcf8 active
CE015 - Marketing                    9d9da8cf-629a-4e4e-aad6-6a16f485f159 active
CE010 - Communication Compliance     c6570dab-aa84-4196-ae57-7287aa4cf224 active
```

Microsoft Purview eDiscovery API

You can configure the Microsoft Purview eDiscovery API to enable the programmatic download of export packages and the reports from an export process in an eDiscovery case.

Prerequisites for Microsoft Purview eDiscovery API

Before you execute the configuration steps in this section, complete and validate the configuration detailed in the Microsoft Graph API section. Extend the previously registered app in Microsoft Entra ID to include the required permissions to achieve programmatic download of the export package.

This configuration already provides the following prerequisites:

- Registered app in Azure portal configured with the appropriate client secret or certificate.
- Service principal in Microsoft Purview assigned the relevant eDiscovery roles.
- Microsoft eDiscovery API permissions configured for the Microsoft Graph.

To extend the existing registered app's API permissions to enable programmatic download, complete the following steps:

- Register a new Microsoft Application and service principal in the tenant.
- Assign additional API permissions to the previously registered app in the Azure portal.

To grant tenant-wide admin consent for Microsoft Purview eDiscovery APIs application permissions, sign in as a user that is authorized to consent on behalf of the organization. For more information, see Grant tenant-wide admin consent to an application.

Configuration steps

Step 1: Register the MicrosoftPurviewEDiscovery app in Microsoft Entra ID

Complete the following steps:

1. Validate that the **MicrosoftPurviewEDiscovery** app isn't already registered. Sign in to the Azure portal and go to **Microsoft Entra ID** > **Enterprise Applications**.

2. Change the **Application type** filter to show *Microsoft Applications*.

3. In the search box, enter *MicrosoftPurviewEDiscovery*. The **MicrosoftPurviewEDiscovery** app should be displayed. If the **MicrosoftPurviewEDiscovery** app isn't listed, register the app in Microsoft Entra ID.

   To register the app, complete the following steps:

   - Use the Microsoft.Graph PowerShell Module to register the **MicrosoftPurviewEDiscovery** app in Microsoft Entra ID. For more information, see Install the Microsoft Graph PowerShell SDK.
   - After the module is installed on a machine, run the following cmdlet to connect to Microsoft Graph using PowerShell:

   PowerShell

   ```
   Connect-MgGraph -scopes "Application.ReadWrite.All"
   ```

   If this is the first time using Microsoft Graph PowerShell cmdlets, you might be prompted to consent to required permissions.

   To register the **MicrosoftPurviewEDiscovery** app, run the following PowerShell commands:

PowerShell

```
$spId = @{"AppId" = "b26e684c-5068-4120-a679-64a5d2c909d9" }
```

PowerShell

```
New-MgServicePrincipal -BodyParameter $spId;
```

> **Note**
>
> Use the PowerShell script to register a new application in Microsoft Entra ID and assign the Microsoft Purview eDiscovery API permissions for application authentication if applicable. After you register the application, you need to configure the client secret or certificate and grant admin consent through the portal.

Step 2: Assign additional MicrosoftPurviewEDiscovery permissions to the registered app

Now that the service principal is added, update the permissions on your previously registered app created in the Microsoft Graph API section of this article. Sign in to the Azure portal and go to **Microsoft Entra ID** > **App Registrations**.

1. Find and select the app you created in the Microsoft Graph API section of this article.
2. Select **API Permissions** from the navigation menu.
3. Select **Add a permission** and then **APIs my organization uses**.
4. Search for MicrosoftPurviewEDiscovery and select it.
5. Select **Application Permissions**.
6. Select the check box for **eDiscovery.Download.Read**.
7. Select **Add Permissions**.
8. On **API permissions**, select **Grant Admin Consent (your organization)** to approve the added permissions.

After admin consent is granted, the status of the added permissions is updated for your organization.

Downloading the export packages and reports

Retrieving the case ID and export job ID

To download the export packages and reports of an export process in an eDiscovery case, you need the case ID and the operation or job ID for the export job.

To gather this information by using the Microsoft Purview portal:

- Open an eDiscovery case.
- Locate the export process.
- Select **Copy support information**.
- Add this information into a text editor (like Notepad).

Alternatively, access this information programmatically by using the following Graph API calls to locate the case ID and the job ID you want to export:

1. Connect to Microsoft Graph by following the steps in the Connecting to Microsoft Graph API using app-only access section of this article.

2. Use the eDiscovery Graph PowerShell cmdlets with the following command if you know the case name:

   PowerShell

   ```
   Get-MgSecurityCaseEdiscoveryCase | where {$_.displayname -eq "<Name of case>"}
   ```

3. After you have the case ID, look up the operations in the case to identify the job ID for the export by using the following command:

   PowerShell

   ```
   Get-MgSecurityCaseEdiscoveryCaseOperation -EdiscoveryCaseId "<case ID>"
   ```

Export jobs are logged under an action of *exportResult* for a direct export from search or *ContentExport* for an export from a review set. The name of the export jobs isn't returned by this API call. To find the name of the export process, you must query the specific operation ID. Use the following command to find the name of the export process:

PowerShell

```
Get-MgSecurityCaseEdiscoveryCaseOperation -EdiscoveryCaseId "<case ID>" -
CaseOperationId “<operation ID>”
```

The name of the export operation is included in the **AdditionalProperties** field.

To make the HTTP API calls directly to list cases in your organization, see List ediscoveryCases.

To make the HTTP API calls directly to list the operations for a case, see List caseOperations.

Use the case ID in the API call to indicate which case to list the operations from. For example:

```
https://graph.microsoft.com/v1.0/security/cases/ediscoveryCases/<CaseID>/operations/
```

The name of the export jobs isn't returned with this API call. To find the name of the export process, you must query the specific job ID. For example:

```
https://graph.microsoft.com/v1.0/security/cases/ediscoveryCases/<CaseID>/operations/<OperationID>
```

Downloading an export package

Retrieving the download URLs for export packages

The *exportFileMetaData* property contains the URL you need to download the export packages and reports. To get the URL, you need the case *ID* of the eDiscovery case where you ran the export process and the operation *ID* for the export process.

Use the following eDiscovery Graph PowerShell cmdlets to find this information:

PowerShell

```
$operation = Get-MgSecurityCaseEdiscoveryCaseOperation -EdiscoveryCaseId "<case ID>" -CaseOperationId “<operation ID>”
```

PowerShell

```
$Operation.AdditionalProperties.exportFileMetadata
```

To make the HTTP API calls directly to return the *exportFileMetaData* information for an operation, see List caseOperations.

Each export package in the Microsoft Purview portal has an entry in the *exportFileMetaData* property. Each entry lists the following information:

- The export package file name
- The downloadUrl to retrieve the export package
- The size of the export package

Example scripts to download the export package

Because the Microsoft Purview eDiscovery API is separate from Microsoft Graph API, you need a separate authentication token to authorize a download request. Use the **MSAL.PS PowerShell Module** and the **Get-MSALToken** cmdlet to get a separate token. You also need to connect to the Microsoft Graph APIs by using the Connect-MgGraph cmdlet.

The following example scripts can be used as a reference when developing your own scripts to enable the programmatic download of the export packages.

Connecting with a client secret

If you configured your app to use a client secret, use the following example script as a reference to download the export package and reports programmatically. Copy the contents into Notepad and save it as `DownloadExportUsingApp.ps1`.

PowerShell

```
[CmdletBinding()]
param (
    [Parameter(Mandatory = $true)]
    [string]$tenantId,
    [Parameter(Mandatory = $true)]
    [string]$appId,
    [Parameter(Mandatory = $true)]
    [string]$appSecret,
    [Parameter(Mandatory = $true)]
    [string]$caseId,
    [Parameter(Mandatory = $true)]
    [string]$exportId,
    [Parameter(Mandatory = $true)]
    [string]$path = "D:\Temp",
    [ValidateSet($null, 'USGov', 'USGovDoD')]
    [string]$environment = $null
)

if (-not(Get-Module -Name Microsoft.Graph -ListAvailable)) {
    Write-Host "Installing Microsoft.Graph module"
    Install-Module Microsoft.Graph -Scope CurrentUser
}

if (-not(Get-Module -Name MSAL.PS -ListAvailable)) {
    Write-Host "Installing MSAL.PS module"
    Install-Module MSAL.PS -Scope CurrentUser
}

$password = ConvertTo-SecureString $appSecret -AsPlainText -Force
$clientSecretCred = New-Object System.Management.Automation.PSCredential -ArgumentList
($appId, $password)
```

```
if (-not(Get-MgContext)) {
    Write-Host "Connect with credentials of a ediscovery admin (token for graph)"
    if (-not($environment)) {
        Connect-MgGraph -TenantId $TenantId -ClientSecretCredential $clientSecretCred
    }
    else {
        Connect-MgGraph -TenantId $TenantId -ClientSecretCredential $clientSecretCred
-Environment $environment
    }
}

Write-Host "Connect with credentials of a ediscovery admin (token for export)"
$exportToken = Get-MsalToken -ClientId $appId -Scopes "00001111-aaaa-2222-bbbb-
3333cccc4444/.default" -TenantId $tenantId -RedirectUri "http://localhost" -
ClientSecret $password

$uri = "/v1.0/security/cases/ediscoveryCases/$($caseId)/operations/$($exportId)"

$export = Invoke-MgGraphRequest -Uri $uri;
if (-not($export)){
    Write-Host "Export not found"
    exit
}
else{
    $export.exportFileMetadata | % {
        Write-Host "Downloading $($_.fileName)"
        Invoke-WebRequest -Uri $_.downloadUrl -OutFile "$($path)\$($_.fileName)" -
Headers @{"Authorization" = "Bearer $($exportToken.AccessToken)"; "X-
AllowWithAADToken" = "true" }
    }
}
```

Save the script and open a new PowerShell window with the following PowerShell modules installed:

- Microsoft.Graph
- MSAL.PS

Browse to the directory where you saved the script and run the following command:

PowerShell

```
.\DownloadExportUsingApp.ps1 -tenantId “<tenant ID>” -appId “<App ID>” -appSecret
“<Client Secret>” -caseId “<CaseID>” -exportId “<ExportID>” -path “<Output Path>”
```

Review the folder you specified as the path to view the downloaded files.

Connecting with a certificate

If you configured your app to use a certificate, use the following example script as a reference to download the export package and reports programmatically. Copy the contents into a text editor and save it as `DownloadExportUsingAppCert.ps1`.

PowerShell

```
[CmdletBinding()]
param (
    [Parameter(Mandatory = $true)]
    [string]$tenantId,
    [Parameter(Mandatory = $true)]
    [string]$appId,
    [Parameter(Mandatory = $true)]
    [String]$certPath,
    [Parameter(Mandatory = $true)]
    [string]$caseId,
    [Parameter(Mandatory = $true)]
    [string]$exportId,
    [Parameter(Mandatory = $true)]
    [string]$path = "D:\Temp",
    [ValidateSet($null, 'USGov', 'USGovDoD')]
    [string]$environment = $null
)

if (-not(Get-Module -Name Microsoft.Graph -ListAvailable)) {
    Write-Host "Installing Microsoft.Graph module"
    Install-Module Microsoft.Graph -Scope CurrentUser
}

if (-not(Get-Module -Name MSAL.PS -ListAvailable)) {
    Write-Host "Installing MSAL.PS module"
    Install-Module MSAL.PS -Scope CurrentUser
}

##$password = ConvertTo-SecureString $appSecret -AsPlainText -Force
##$clientSecretCred = New-Object System.Management.Automation.PSCredential -
ArgumentList ($appId, $password)

$ClientCert = Get-ChildItem $certPath

if (-not(Get-MgContext)) {
    Write-Host "Connect with credentials of a ediscovery admin (token for graph)"
    if (-not($environment)) {
        Connect-MgGraph -TenantId $TenantId -ClientId $appId -Certificate $ClientCert
    }
    else {
        Connect-MgGraph -TenantId $TenantId -ClientId $appId -Certificate $ClientCert
-Environment $environment
    }
}

Write-Host "Connect with credentials of a ediscovery admin (token for export)"
```

```
$connectionDetails = @{
    'TenantId'          = $tenantId
    'ClientId'          = $appID
    'ClientCertificate' = $ClientCert
    'Scope' = "00001111-aaaa-2222-bbbb-3333cccc4444/.default"
}

$exportToken = Get-MsalToken @connectionDetails

$uri = "/v1.0/security/cases/ediscoveryCases/$($caseId)/operations/$($exportId)"

$export = Invoke-MgGraphRequest -Uri $uri;
if (-not($export)){
    Write-Host "Export not found"
    exit
}
else{
    $export.exportFileMetadata | % {
        Write-Host "Downloading $($_.fileName)"
        Invoke-WebRequest -Uri $_.downloadUrl -OutFile "$($path)\$($_.fileName)" -
Headers @{"Authorization" = "Bearer $($exportToken.AccessToken)"; "X-
AllowWithAADToken" = "true" }
    }
}
```

When you save the script, open a new PowerShell window with the following PowerShell modules installed:

- Microsoft.Graph
- MSAL.PS

Browse to the directory where you saved the script and run the following command.

PowerShell

```
.\DownloadExportUsingAppCert.ps1 -tenantId “<tenant ID>” -appId “<App ID>” -certPath
“<Certificate Path>” -caseId “<CaseID>” -exportId “<ExportID>” -path “<Output Path>”
```

Review the folder you specified as the path to view the downloaded files.

ⓘ **Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 05/06/2026

Microsoft Purview eDiscovery solutions

Article • 05/28/2025

> **Important**
>
> The classic eDiscovery (Premium) experience will be **retired in August 2025** and won't be available as an experience option in the Microsoft Purview portal after retirement.
>
> We recommend that you start planning for this transition early and start using the new eDiscovery experience in the Microsoft Purview portal.

Electronic discovery, or eDiscovery, is the process of identifying and delivering electronic information that can be used as evidence in legal cases. You can use eDiscovery tools in Microsoft Purview to search for content in Exchange Online, OneDrive for Business, SharePoint Online, Microsoft Teams, Microsoft 365 Groups, and Viva Engage teams. You can search mailboxes and sites in the same eDiscovery search, and then export the search results. You can use Microsoft Purview eDiscovery (Standard) cases to identify, hold, and export content found in mailboxes and sites. If your organization has an Office 365 E5 or Microsoft 365 E5 subscription (or related E5 add-on subscriptions), you can further manage custodians and analyze content by using the feature-rich Microsoft Purview eDiscovery (Premium) solution in Microsoft 365.

> **Tip**
>
> If you're not an E5 customer, use the 90-day Microsoft Purview solutions trial to explore how additional Purview capabilities can help your organization manage data security and compliance needs. Start now at the Microsoft Purview trials hub. Learn details about signing up and trial terms.

eDiscovery solutions

> **Important**
>
> Content Search and eDiscovery (Standard) in the eDiscovery classic experience were retired on May 26, 2025. Unless you're working directly with Microsoft when using these legacy features for specific short-term scenarios, use the guidance for the new Content Search case and eDiscovery cases in the new eDiscovery experience in the Microsoft Purview portal.

Microsoft Purview provides three eDiscovery solutions: Content search, eDiscovery (Standard), and eDiscovery (Premium).

Expand table

Content Search	eDiscovery (Standard)	eDiscovery (Premium)
- Search for content - Keyword queries and search conditions - Export search results - Role-based permissions	- Search and export - Case management - Legal hold	- Custodian management - Legal hold notifications - Advanced indexing - Review set filtering - Tagging - Analytics - Predictive coding models And more...

- **Content search**. Use the Content search tool to search for content across Microsoft 365 data sources and then export the search results to a local computer.
- **eDiscovery (Standard)**. eDiscovery (Standard) builds on the basic search and export functionality of Content search by enabling you to create eDiscovery cases and assign eDiscovery managers to specific cases. eDiscovery managers can only access the cases of which they're members. eDiscovery (Standard) also lets you associate searches and exports with a case and lets you place an eDiscovery hold on content locations relevant to the case.
- **eDiscovery (Premium)**. The eDiscovery (Premium) tool builds on the existing case management, preservation, search, and export capabilities in eDiscovery (Standard). eDiscovery (Premium) provides an end-to-end workflow to identify, preserve, collect, review, analyze, and export content that's responsive to your organization's internal and external investigations. It lets legal teams manage custodians and the legal hold notification workflow to communicate with custodians involved in a case. It allows you to collect and copy data from the live service into review sets, when you can filter, search, and tag content to cull non-relevant content from further review so your workflow can identify and focus on content that's most relevant. eDiscovery (Premium) provides analytics and machine learning-based predictive coding models to further narrow to scope of your investigation to the most relevant content.

Comparison of key capabilities

> **Important**
>
> Content Search and eDiscovery (Standard) in the eDiscovery classic experience were retired on May 26, 2025. Unless you're working directly with Microsoft when using these

legacy features for specific short-term scenarios, use the guidance for the new Content Search case and eDiscovery cases in the new eDiscovery experience in the Microsoft Purview portal.

The following table compares the key capabilities available in Content search, eDiscovery (Standard), and eDiscovery (Premium).

Expand table

Capability	Content search	eDiscovery (Standard)	eDiscovery (Premium)
Search for content	✔	✔	✔
Keyword queries and search conditions	✔	✔	✔
Search statistics	✔	✔	✔
Export search results	✔	✔	✔
Role-based permissions	✔	✔	✔
Case management		✔	✔
Place content locations on legal hold		✔	✔
Custodian management			✔
Legal hold notifications			✔
Advanced indexing			✔
Error remediation			✔
Review sets			✔
Support for cloud attachments and SharePoint versions			✔
Optical character recognition			✔
Conversation threading			✔

Capability	Content search	eDiscovery (Standard)	eDiscovery (Premium)
Collection statistics and reports			✔
Review set filtering			✔
Tagging			✔
Analytics			✔
Predictive coding models			✔
Computed document metadata			✔
Transparency of long-running jobs			✔
Export to customer-owned Azure Storage location			✔

Here's a description of each eDiscovery capability.

- **Search for content**. Search for content that's stored in Exchange mailboxes, OneDrive for Business accounts, SharePoint sites, Microsoft Teams, Microsoft 365 Groups, and Viva Engage Teams. This includes content generated by other Microsoft 365 apps that store data in mailboxes and sites.
- **Keyword queries and search conditions**. Create Keyword Query Language (KQL) search queries to search for content keywords that match query criteria. You can also include conditions to narrow the scope of your search.
- **Search statistics**. After you run a search, you can view statistics of the estimated search results, such as the number and total size of items matching your search criteria. Other statistics include the top content locations that contain search results and the number of items that match different parts of the search query.
- **Export search results**. Export search results to a local computer in your organization in a two-step process. When you export search results, items are copied from their original content location in Microsoft 365 to a Microsoft-provided Azure Storage location. Then you can download those items to a local computer.
- **Role-based permissions**. Use role-based access control (RBAC) permissions to control what eDiscovery-related tasks that different users can perform. You can use a built-in eDiscovery-related role group or create custom role groups that assign specific eDiscovery permissions.

- **Case management**. eDiscovery cases in eDiscovery (Standard) and eDiscovery (Premium) let you associate specific searches and exports with a specific investigation. You can also assign members to a case to control who can access the case and view the contents of the case. eDiscovery (Premium) also supports new case creation integration with Microsoft Purview Insider Risk Management cases.
- **Place content locations on legal hold**. Preserve content relevant to your investigation by placing a legal hold on the content locations in a case. This lets you secure electronically stored information from inadvertent (or intentional) deletion during your investigation.
- **Custodian management**. Manage the people that you've identified as people of interest in the case (called *custodians*) and other data sources that may not be associated with a custodian. When you add custodians and non-custodial data sources to a case, you can place a legal hold on these data sources, communicate with custodians by using the legal hold notification process, and search custodian and non-custodial data sources to collect content relevant to the case.
- **Legal hold notifications**. Manage the process of communicating with case custodians. A legal hold notification instructs custodians to preserve content that's relevant to the case. You can track the notices that were received, read, and acknowledged by custodians. The communications workflow in eDiscovery (Premium) allows you to create and send initial notifications, reminders, and escalations if custodians fail to acknowledge a hold notification.
- **Advanced indexing**. When you add custodial and non-custodian data sources to a case, the associated content locations are reindexed in a process called *Advanced indexing*. Advanced indexing ensures any content deemed as partially indexed is reprocessed to make it fully searchable when you collect data for an investigation.
- **Error remediation**. Fix processing errors using a process called *error remediation*. Error remediation allows you to rectify data issues that prevent eDiscovery (Premium) from properly processing the content during Advanced indexing. For example, files that are password protected can't be processed since the files are locked or encrypted. Using error remediation, you can download files with errors, remove the password protection, and then upload the remediated files.
- **Review sets**. Add relevant data to a review set. A review set is a secure, Microsoft-provided Azure Storage location in the Microsoft cloud. When you add data to a review set, the collected items are copied from their original content location to the review set. Review sets provide a static, known set of content that you can search, filter, tag, analyze, and predict relevancy using predictive coding models. You can also track and report on what content gets added to the review set.
- **Support for cloud attachments and SharePoint versions**. When you add content to a review set, you have the option to include cloud attachments or linked files. This means that the target file of a cloud attachment or linked file is added to the review set. You also have the option to add all versions of a SharePoint document to a review set.

- **Optical character recognition (OCR)**. When content is added to a review set, OCR functionality extracts text from images, and includes the image text with the content that's added to a review set. This lets you search for image text when you query the content in the review set.
- **Conversation threading**. When chat messages from Teams and Viva Engage conversations are added to a review set, you can collect the entire conversation thread. This means that the entire chat conversation that contains items that match the collection criteria is added to the review set. This lets you review chat items in the context of the back-and-forth conversation.
- **Collection statistics and reports**. After you create a collection estimate or commit a collection to a review set, you can view a rich set of statistics on the retrieved items, such as the content locations that contain the most items that matched the search criteria and the number of items returned by the search query. You can also preview a subset of the results.
- **Review set filtering**. After content is added to a review set, you can apply filters to display only the set of items that match your filtering criteria. Then you can save the filter sets as a query, which lets you quickly reapply the saved filters. Review set filtering and saved queries help you quickly select content items that are most relevant to your investigation.
- **Tagging**. Tags also help you omit non-relevant content and identify the most relevant content. When experts, attorneys, or other users review content in a review set, their opinions related to the content can be captured by using tags. For example, if the intent is to exclude unnecessary content, a user can tag documents with a tag such as "non-responsive". After content has been reviewed and tagged, a review set query can be created to exclude any content tagged as "non-responsive". This process eliminates the non-responsive content from subsequent steps in the eDiscovery workflow.
- **Analytics**. eDiscovery (Premium) provides tools to analyze review set documents to help you organize the documents in a coherent manner and reduce the volume of documents to be reviewed. *Near duplicate detection* groups textually similar documents together to help you make your review process more efficient. *Email threading* identifies specific email messages that give a complete context of the conversation in an email thread. *Themes* functionality attempts to analyze themes in review set documents and assign a theme to documents so that you can review documents with related theme. These analytics capabilities help make your review process more efficient so that reviewers can review a fraction of collected documents.
- **Predictive coding models**. Use predictive coding models to reduce large volumes of case content to a relevant set of items that you can prioritize for review. This is accomplished by creating and training your own predictive coding models that help you prioritize the review of the most relevant items in a review set. The system uses the training to apply prediction scores to every item in the review set. This lets you filter items based on the prediction score, which allows you to review the most relevant (or non-relevant) items first.

- **Computed document metadata**. Many of the eDiscovery (Premium) features, such as Advanced indexing, conversation threading, analytics, and predictive coding add metadata properties to review set documents. This metadata contains information related to the function performed by a specific feature. When reviewing documents, you can filter on metadata properties to display documents that match your filter criteria. This metadata can be imported into third-party review applications after review set documents are exported.
- **Transparency of long-running jobs**. Jobs in eDiscovery (Premium) are typically long-running processes that are triggered by user actions, such as the adding custodians to a case, adding content to a review set, running analytics, and training predictive coding models. You can track the status of these jobs and get support information if you need to escalate issues to Microsoft Support.
- **Export to customer-owned Azure Storage location**. When you export documents from a review set, you have the option to export them to an Azure Storage account managed by your organization. Additionally, eDiscovery (Premium) lets you customize what data is exported. This includes exporting file metadata, native files, text files, tags, and redacted documents saved to a PDF file.

eDiscovery subscription comparison

Before you get started, review the subscription requirements for Content search, eDiscovery (Standard), and eDiscovery (Premium). Generally, subscriptions that support eDiscovery (Standard) also support Content search and subscriptions that support eDiscovery (Premium) also support Content search and eDiscovery (Standard).

Get started with eDiscovery

See the following articles to help you learn more and get started using Microsoft Purview eDiscovery solutions.

Get started with eDiscovery (Premium)

- Overview of eDiscovery (Premium)
- Get started with eDiscovery (Premium)
- Create and manage an eDiscovery (Premium) case

Integration with Insider Risk Management

Cases in Microsoft Purview Insider Risk Management can be quickly escalated to new cases in Microsoft Purview eDiscovery (Premium) when additional legal review is needed for potentially risky user activity. The tight integration between these solutions can help your risk and legal teams work more efficiently and can help provide a complete end-to-end view of user activities under review. Check out how to get started with Insider Risk Management and how to easily escalate an Insider Risk Management case to an eDiscovery (Premium) case.

eDiscovery roadmap

To see what eDiscovery features have been launched, are rolling out, or in development, see the Microsoft 365 Roadmap.

Training

Training your IT administrators, eDiscovery managers, and compliance investigation teams in the basics for Content search, eDiscovery (Standard), and eDiscovery (Premium) can help your organization get started more quickly using Microsoft Purview eDiscovery tools. To help these users in your organization getting started with eDiscovery, see Describe the eDiscovery and audit capabilities of Microsoft Purview.

Overview of Content search

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

Use the Content search tool in the Microsoft Purview portal to quickly find email in Exchange mailboxes, documents in SharePoint sites and OneDrive locations, and instant messaging conversations in Skype for Business. You can use the content search tool to search for email, documents, and instant messaging conversations in collaboration tools such as Microsoft Teams and Microsoft 365 Groups.

Search for content

The first step is to starting using the Content search tool to choose content locations to search and configure a keyword query to search for specific items. Or, you can just leave the query blank and return all items in the target locations.

- Create and run a Content search.
- Build search queries and use conditions to narrow your search.
- Feature reference for Content search.
- Configure search permissions filtering so that an eDiscovery manager can only search subset of mailboxes or sites in your organization.
- Search cloud-based mailboxes for on-premises users in Microsoft 365.
- View keyword statistics for the results of a search and then refine the query if necessary.
- Search for third-party data that your organization has imported to Microsoft 365.
- Preserve Bcc recipients so you can search for them.
- Search for Microsoft 365 Copilot and Microsoft 365 Copilot Chat data (user prompts and Copilot responses) stored in custodian mailboxes.

Perform actions on content you find

After you run a search and refine it as necessary, the next step is to do something with the results returned by the search. You can export and download the results to your local computer or in the case of an email attack on your organization, you can delete the results of a search from user mailboxes.

- Export the results of a content search and download them to your local computer..
- Search for and delete email messages, such as messages that content a virus, dangerous attachments, or phishing messages.

Learn more about content search

Content search is easy to use, but it's also a powerful tool. Behind-the-scenes, there's a lot going on. The more you know about it and understand its behavior and its limitations, the more successful you'll be using it for your organization's search and investigation needs.

- Estimated and actual search results and the reasons why there might be differences between them when you export and download search results.
- Partially indexed items in Exchange and SharePoint and how to include or exclude them when you export and download search results.
- Investigate partially indexed items and determine your organization's exposure to them.

Use scripts for advanced scenarios

Sometimes you have to perform more advanced, complex, and repetitive content search tasks. In these cases, it's easier and faster to use commands in Security & Compliance PowerShell.

To help make this easier, we've created several Security & Compliance PowerShell scripts to help you complete complex content search-related tasks.

- Search specific mailbox and site folders (called a *targeted* collection) when you're confident that items responsive to a case are located in that folder.
- Search the mailbox and OneDrive location for a list of users.
- Create, report on, and delete multiple searches to quickly and efficiently identify and cull search data.
- Clone a content search and quickly compare the results of different keyword search queries run on the same content locations; or use the script to save time by not having to re-enter a large number of content locations when you create a new search.

Last updated on 11/21/2025

Get started with Content search

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

You can use the Content search eDiscovery tool in the Microsoft Purview portal to search for in-place content such as email, documents, and instant messaging conversations in your organization. Use this tool to search for content in these cloud-based Microsoft 365 data sources:

- Exchange Online mailboxes
- SharePoint Online sites and OneDrive for Business accounts
- Microsoft Teams
- Microsoft 365 Groups
- Viva Engage

After you run a search, the number of content locations and an estimated number of search results are displayed on the search flyout page. You can quickly view statistics, such as the content locations that have the most items that match the search query. After you run a search, you can preview the results or export them to a local computer.

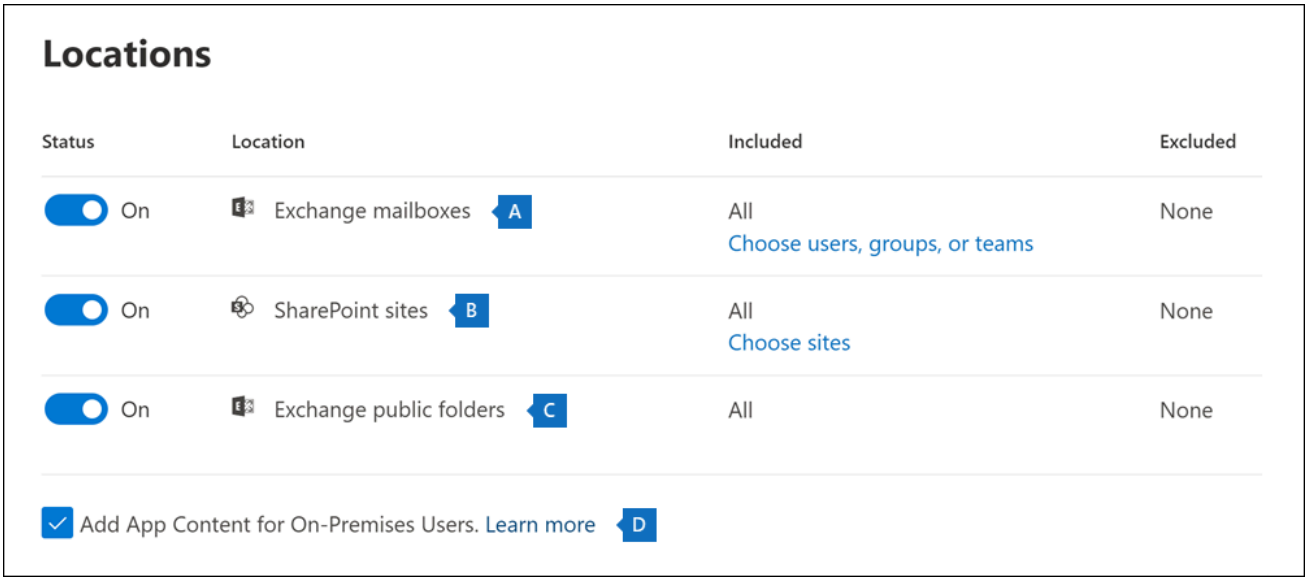
Before you run a search

- To access to the Content search tool in the Microsoft Purview portal (to run searches and preview results and export results), an administrator, compliance officer, or eDiscovery manager must be a member of the eDiscovery Manager role group in the Microsoft Purview portal. For more information, see Assign eDiscovery permissions.
- In an Exchange hybrid deployment, you can't use the Content search tool to search emails in on-premises mailboxes. You can only use the tool to search cloud-based mailboxes.
- In an Exchange hybrid deployment, you can search for Teams chat data in on-premises mailboxes. For more information, see Search for and export Teams chat data for on-

premises users.

Create and run a search

1. Go to the Microsoft Purview portal and sign in using the credentials of an account that's been assigned the appropriate permissions.
2. In the left navigation pane, select **Content search**.
3. On the **Content search** page, select **New search**.
4. On the **Name and description** page, enter a name for the search, an optional description that helps identify the search. The name of the search must be unique in your organization.
5. On the **Locations** page, choose the content locations that you want to search. You can search mailboxes, sites, and public folders.



a. **Exchange mailboxes**: Set the toggle to **On**. The option to search all Exchange mailboxes is automatically selected. If needed, select **Choose users, groups, or teams** to specify the mailboxes to search. Use the search box to find user mailboxes and distribution groups. You can also search the mailbox associated with a Microsoft Team (for channel messages), Microsoft 365 Group, and Viva Engage community. All Microsoft 365 Copilot and Microsoft 365 Copilot Chat activity data (user prompts and Copilot responses) generated in supported Microsoft 365 apps and services is stored in custodian mailboxes. For more information about the application data stored in mailboxes, see Content stored in mailboxes for eDiscovery.

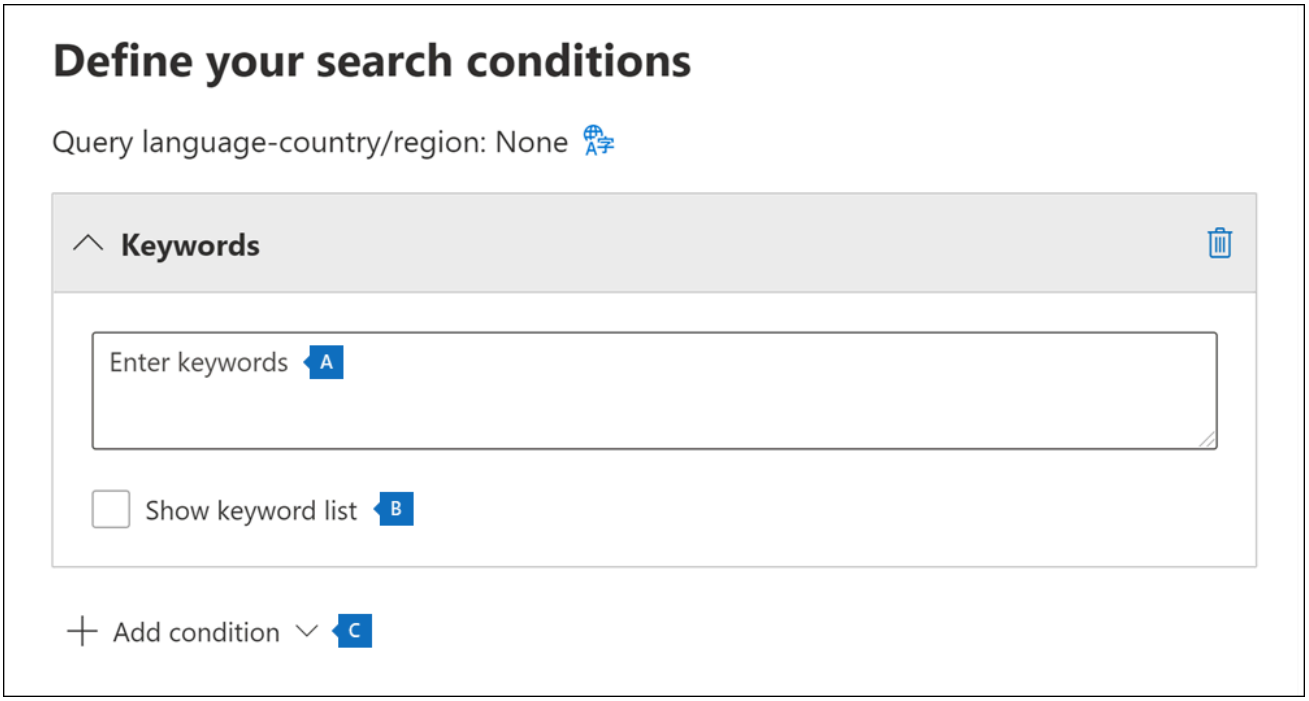
b. **SharePoint sites**: Set the toggle to **On**. The option to search all SharePoint sites is automatically selected. Select **Choose sites** to specify SharePoint sites and OneDrive sites

to search. Enter the URL for each site that you want to search. You can also add the URL for the SharePoint site for a Microsoft Team, Microsoft 365 Group, or Viva Engage community. The Recycle Bin in SharePoint sites is not indexed and therefore unavailable for searching. As a result, eDiscovery searches can't find any Recycle Bin content to place holds.

c. **Exchange public folders**: Set the toggle to **On**. The option to search all Exchange public folders is automatically selected to search all public folders in your Exchange Online organization. You can't choose specific public folders to search. Leave the toggle switch off if you don't want search all public folders.

d. Keep this checkbox selected to search for Teams content for on-premises users. For example, if you search all Exchange mailboxes in the organization and this checkbox is selected, the cloud-based storage used to store Teams chat data for on-premises users will be included in the scope of the search. For more information, see Search for Teams chat data for on-premises users.

6. On the **Conditions** page, enter a keyword query and add conditions to the search query if necessary.



a. Specify keywords, message properties such as sent and received dates, or document properties such as file names or the date that a document was last changed. You can use more complex queries that use a Boolean operator, such as **AND**, **OR**, **NOT**, and **NEAR**. If you leave the keyword box empty, all content located in the specified content locations is

included in the search results. For more information, see Keyword queries and search conditions for eDiscovery.

b. Alternatively, you can select the **Show keyword list** checkbox and then enter a keyword in each row. If you do this, the keywords on each row are connected by a logical operator (**c:s**) that is similar in functionality to the **OR** operator in the search query that's created.

   Why use the keyword list? You can get statistics that show how many items match each keyword. This can help you quickly identify which keywords are the most (and least) effective. You can also use a keyword phrase (surrounded by parentheses) in a row. For more information about the keyword list and search statistics, see Get keyword statistics for searches.

   > **Note**
   >
   > To help reduce issues caused by large keyword lists, you're limited to a maximum of 20 rows in the keyword list.

c. You can add search conditions to narrow a search and return a more refined set of results. Each condition adds a clause to the search query that is created and run when you start the search. A condition is logically connected to the keyword query (specified in the keyword box) by a logical operator (**c:c**) that is similar in functionality to the **AND** operator. That means that items have to satisfy both the keyword query and one or more conditions to be included in the results. This is how conditions help to narrow your results. For a list and description of conditions that you can use in a search query, see Search conditions.

7. Review the search settings (and edit if necessary), and then select **Submit** to start the search.

To access this content search again or access other content searches listed on the **Content search** page, select a search to view the search summary and search statistics.

For more information about Content search, such as searching for content in different Microsoft 365 services, see Feature reference for Content search.

Next steps

Here's a list of next steps to perform after you create and run a Content search.

- Preview search results
- View statistics for search results
- Export search results

Last updated on 11/21/2025

Use Content search in Microsoft Teams

Applies to: Microsoft Teams

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

> **Note**
>
> Content search of messages and files in private channels work differently than in standard channels. To learn more, see Content search of private channels.

Content search provides a way to query Microsoft Teams information spanning Exchange, SharePoint Online, and OneDrive for Business. For example, using **Content search** against your *Manufacturing Specs* mailbox and *Manufacturing Specs* SharePoint site, you can search against Teams standard channel conversations from Exchange, file uploads and modifications from SharePoint Online, and OneNote changes.

You can also add query criteria to the **Content Search** to narrow the results returned. In the previous example, you can look for content where the keywords "**New Factory Specs**" were used.

> **Tip**
>
> After adding search conditions, you can export a report or the actual content to your computer for analysis.

Content search of private channels

Records for messages sent in a private channel are delivered to the mailbox of all private channel members, rather than to a group mailbox. The titles of the records are formatted to indicate which private channel they were sent from.

Because each private channel has its own SharePoint site collection that's separate from the parent team site, files in a private channel are managed independently of the parent team.

Teams doesn't support content search of a single channel, so the whole team must be searched. To perform a content search of a private channel, search across the team, the site collection associated with the private channel (to include files), and mailboxes of private channel members (to include messages).

Use the following steps to identify files and messages in a private channel to include in your content search.

Include private channel files in a content search

Before you perform these steps, install the SharePoint Online Management Shell and connect to SharePoint Online.

1. Run the following to get a list of all SharePoint site collections associated with private channels in the team.

   PowerShell

   ```
   Get-SPOSite
   ```

2. Run the following PowerShell script to get a list of all SharePoint site collection URLs associated with private channels in the team and the parent team group ID.

   PowerShell

   ```
   $sites = get-sposite -template "teamchannel#0"
   foreach ($site in $sites) {$x= get-sposite -identity $site.url -detail;
   $x.relatedgroupID; $x.url}
   ```

3. For each team or group ID, run the following PowerShell script to identify all relevant private channel sites.

   PowerShell

   ```
   $sites = get-sposite -template "teamchannel#0"
   $groupID = "e8195240-4a70-4830-9106-80193cf717cb"
   ```

```
foreach ($site in $sites) {$x= Get-SpoSite -Identity $site.url -Detail; if
($x.RelatedGroupId -eq $groupID) {$x.RelatedGroupId;$x.url}}
```

Include private channel messages in a content search

Before you perform these steps, make sure you have the latest version of the Teams PowerShell module installed.

1. Run the following to get a list of private channels in the team.

   PowerShell

   ```
   Get-TeamChannel -GroupId <GroupID> -MembershipType Private
   ```

2. Run the following to get a list of private channel members.

   PowerShell

   ```
   Get-TeamChannelUser -GroupId <GroupID> -DisplayName "Engineering" -Role Member
   ```

3. Include the mailboxes of all members from each private channel in the team as part of your content search query.

Related articles

- eDiscovery cases in the Microsoft Purview portal

Last updated on 11/21/2025

Search for and export Teams chat data for on-premises users

> **⊗ Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic **Content Search**, classic **eDiscovery (Standard)**, and classic **eDiscovery (Premium)**.
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the **new eDiscovery experience** in the **Microsoft Purview portal**.

If your organization has an Exchange hybrid deployment, you can use eDiscovery tools to search for and export Teams chat data for on-premises users. This option is also available for organizations that synchronize on-premises Exchange organization with Microsoft 365.

For a cloud-based users using Microsoft Teams enabled in your organization, chat data (also called *1x1 or 1xN chats*) is saved to their primary cloud-based mailbox. When an on-premises user uses the Teams chat application, their chat messages can't be stored in their primary mailbox, which is located on-premises. To get around this limitation, a feature where a cloud-based storage area is created so that you use eDiscovery tools to search for and export Teams chat data for on-premises users.

Here are the requirements and limitations for enabling cloud-based storage for on-premises users:

- The user accounts in your on-premises directory service (such as Active Directory) must be synchronized with Microsoft Entra ID, the directory service in Microsoft 365. This means that a mail user account is created in Microsoft 365 and is associated with a user whose primary mailbox is located in the on-premises organization.
- The user whose primary mailbox is located in the on-premises organization must be assigned a Microsoft Teams license and a minimum of an Exchange Online Plan 1 license.
- If your organization doesn't have an Exchange hybrid deployment, you must synchronize your on-premises Exchange schema to Microsoft Entra ID. If you don't do this, you might

risk creating duplicate cloud-based mailboxes in Exchange Online for users that have a mailbox in your on-premises Exchange organization.

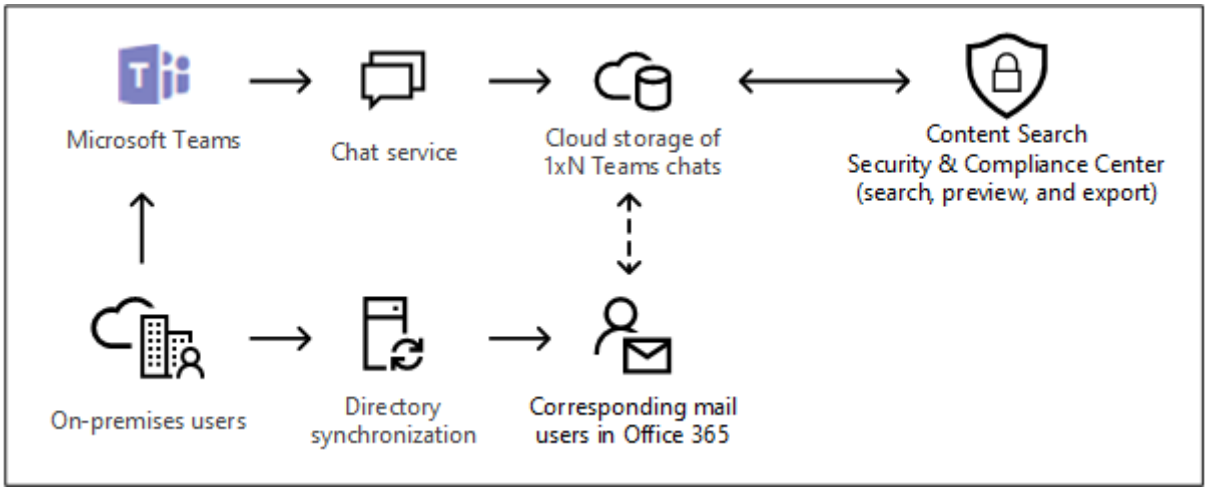
- Only the Teams chat data associated with an on-premises user is stored in the cloud-based storage area. An on-premises user can't access this storage area.

> ⓘ **Note**
>
> Teams channel conversations are always stored in the cloud-based mailbox that's associated with the Team, which means you can search for channel conversations. For more information about searching Teams channel conversations, see Searching Microsoft Teams and Microsoft 365 Groups.

How it works

If a Microsoft Teams-enabled user has an on-premises mailbox and their user account/identity has been synched to the cloud, Microsoft creates cloud-based storage to associate the on-premises user's 1xN Teams chat data with. Teams chat data for on-premises users is indexed for search. This lets you use Content search (and searches associated with Microsoft Purview eDiscovery (Standard) and Microsoft Purview eDiscovery (Premium) cases) to search, preview, and export Teams chat data for on-premises users. You can also use ***ComplianceSearch** cmdlets in Security & Compliance PowerShell to search for Teams chat data for on-premises users.

The following graphic shows the workflow of how Teams chat data for on-premises users is available to search, preview, and export.

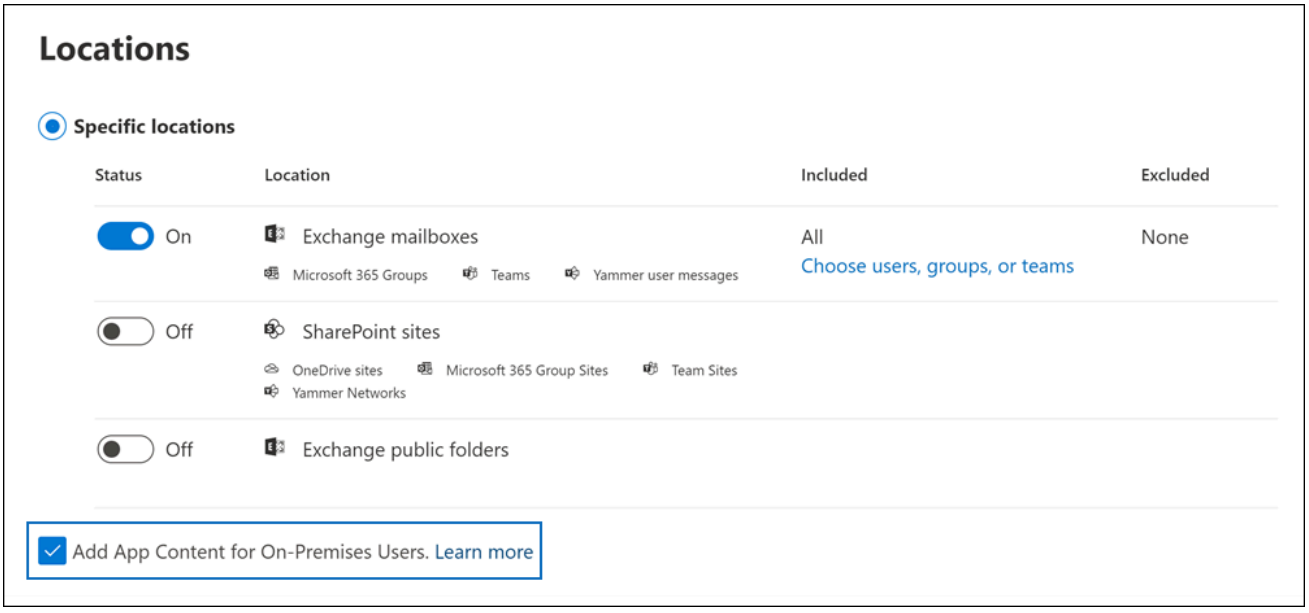


In addition to this capability, you can also use eDiscovery tools to search, preview, and export Teams content in the cloud-based SharePoint site and Exchange mailbox associated with each Microsoft Team and 1xN Teams chat data in the Exchange Online mailbox for cloud-based users.

Searching for and exporting Teams chat content for on-premises users

Here's how to use Content search in the Microsoft Purview portal to search for Teams chat data for on-premises users. You can also use the search tool in eDiscovery (Standard) to search for chat data for on-premises users.

1. In the Microsoft Purview portal, go to **Content search**.
2. On the **Searches** tab, select **New search**, and name the new search.
3. On the **Locations** page, set the toggle to **On** for Exchange mailboxes.
4. To search for Teams content for specific users (including on-premises users), select **Choose user, groups, or teams** and choose specific users to include in the search. If you don't list specific users, the search will include all users, including on-premises users.
5. Make sure the **Add app content for on-premises users** checkbox is selected. This ensures that the cloud-bases storage for on-premises users will be searched.



6. On the **Define your search conditions** page, create a keyword query and add conditions to the search query if necessary. To only search for Team chats data, you can add the following query in the **Keywords** box:

```text
kind:im AND kind:microsoftteams
```

7. Submit and run the search. Any search results for on-premises users can be previewed like any other search results. You can also export the search results (including any Teams chat data) to a PST file. For more information, see:

   - Create a search
   - Preview search results
   - Export search results

Using PowerShell to search for Teams chat data for on-premises users

You can use the **New-ComplianceSearch** cmdlets in Security & Compliance PowerShell to search for Teams chat data for on-premises users. As previously explained, you don't have to submit a support request to use PowerShell to search for Teams chat data for on-premises users.

1. Connect to Security & Compliance PowerShell.

2. Run the following PowerShell command to create a content search that searches for Teams chat data for on-premises users.

   PowerShell

   ```
   New-ComplianceSearch <name of new search> -ContentMatchQuery <search query> -
   ExchangeLocation <on-premises user> -IncludeUserAppContent $true -
   AllowNotFoundExchangeLocationsEnabled $true
   ```

   The *IncludeUserAppContent* parameter is used to specify the cloud-based storage for the user or users who are specified by the *ExchangeLocation* parameter. The *AllowNotFoundExchangeLocationsEnabled* allows you to search the cloud-based storage for on-premises users. When you use the `$true` value for this parameter, the search doesn't try to validate the existence of the mailbox before it runs. This is required to search the cloud-based storage for on-premises users because this cloud-based storage doesn't resolve as a regular cloud-based mailbox.

   The following example searches for Teams chats that contain keyword "redstone" in the cloud-based storage for Sara Davis, who is an on-premises user in the Contoso organization.

   PowerShell

```
New-ComplianceSearch "Redstone_Search" -ContentMatchQuery "redstone AND (kind:im
AND kind:microsoftteams)" -ExchangeLocation sarad@contoso.com -
IncludeUserAppContent $true -AllowNotFoundExchangeLocationsEnabled $true
```

After you create a search, be sure to use the **Start-ComplianceSearch** cmdlet to run the search.

For more information using these cmdlets, see:

- New-ComplianceSearch
- Start-ComplianceSearch

Limits

- Currently, you can search, preview, and export Teams chat data for on-premises users. You can also place the Teams chat data for an on-premises user on a hold associated with an eDiscovery (Standard) or eDiscovery (Premium) case, and apply a retention policy for Teams chats or channel messages for on-premises users. However, you can't apply a retention policy for other content locations (such as Exchange mailboxes and SharePoint sites) for on-premises users.

Frequently asked questions

Do I have to submit a support request to search for chat messages for on-premises users?

No. This feature is enabled by default for all organizations. At one point, you did have to contact Microsoft Support but that is no longer the case.

Can eDiscovery tools find older Teams chat data for on-premises users before the time that this feature was enabled by default for all organizations?

Microsoft started storing the Teams chat data for on-premises users on January 31, 2018. If the identity of an on-premises Teams user has been synched between your on-premises Active Directory and Microsoft Entra ID in Microsoft 365 since this date, then their Teams chat data is stored in the cloud and is searchable using eDiscovery tools.

Do on-premises users need a license to store their Teams chat data in the cloud?

Yes. To store Teams chat data for an on-premises user in a cloud-based storage, the user must be assigned a Microsoft Teams license and an Exchange Online Plan license in Office 365 (or Microsoft 365).

Where is the cloud-based storage for on-premises users located?

Teams chat data is stored in the Preferred Data Location (PDL) for an on-premises user. The PDL is honored in both Single-Geo and Multi-Geo environments. For more information, see Microsoft 365 Multi-Geo.

Is there a risk of losing the Teams chat data if the user's on-premises mailbox is migrated to the cloud?

No. When you migrate the primary mailbox of an on-premises user to the cloud, the Teams chat data for that user will be migrated to their new cloud-based primary mailbox.

Can I apply an eDiscovery hold or retention policies to on-premises users?

Yes. You can apply eDiscovery holds or retention policies for Teams chats and channel messages of on-premises users. But to preserve or retain Teams content for on-premises users, an on-premises user must be assigned an Exchange Online Plan 2 license.

Last updated on 09/02/2025

Preview eDiscovery search results

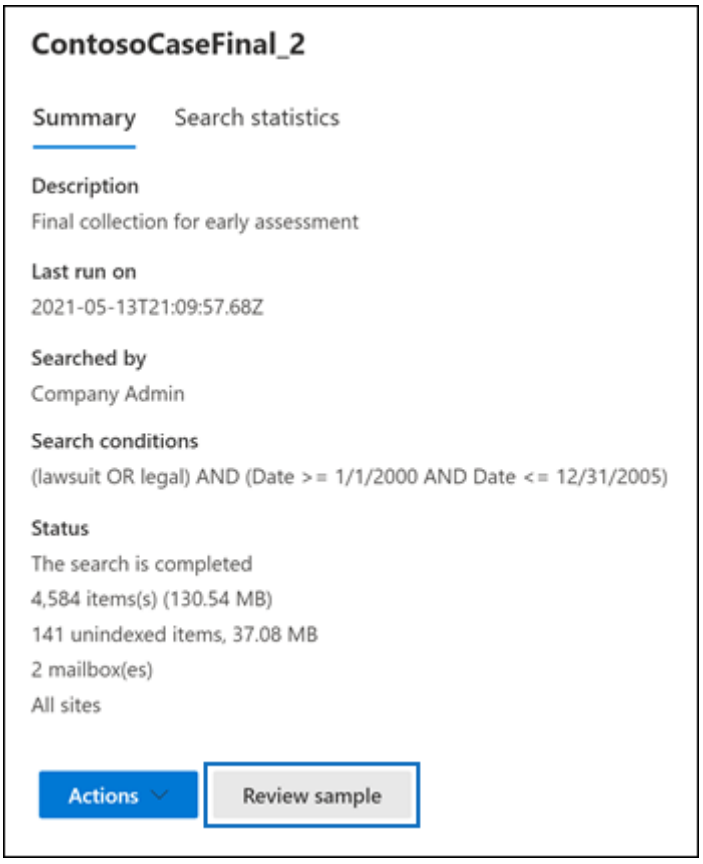
> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

After you run a Content search, you can preview a sample of the results returned by the search. Previewing items returned by the search query can help you determine if the search is returning the results you hope for or if you need to change the search query and rerun the search.

> **Note**
>
> There may be scenarios when the preview sample doesn't contain all available results of searches, and sampling is best effort. In some cases (such as in environments operated by 21Vianet in China), the review sample might not function reliably. For the most accurate and complete search results, use export to obtain the full set of search results.

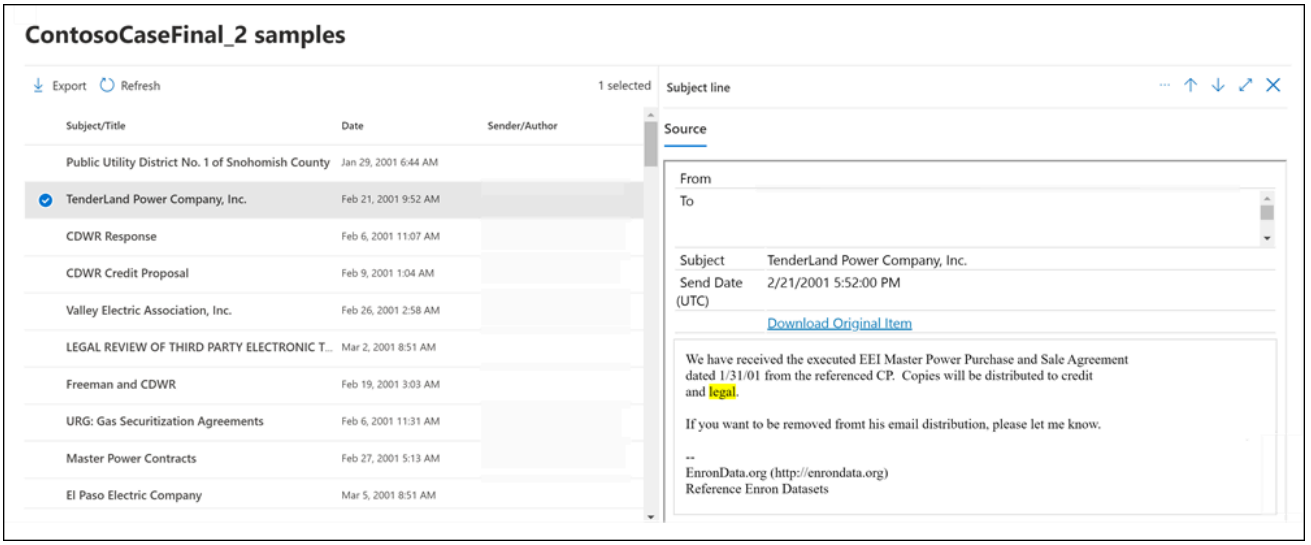
To preview a sample of results returned by a search:

1. In the Microsoft Purview portal, go to the **Content search**.
2. Select search to display the flyout page.
3. On the bottom of the flyout page, select **Review sample**.



A page is displayed containing a sample of the search results.

4. Select an item to view its contents in the reading pane.



In the previous screenshot, the keywords from the search query are highlighted when you preview items.

How the search result samples are selected

A maximum of 1,000 randomly selected items are available to preview. In addition to being randomly selected, items available for preview must also meet the following criteria:

- A maximum of 100 items from a single content location (a mailbox or a site) can be previewed. This maximum means that it's possible that less than 1,000 items might be available for preview. For example, if you search four mailboxes and the search returns 1,500 estimated items, only 400 are available for preview because only 100 items from each mailbox can be previewed.
- For mailbox items, only email messages are available to preview. Preview for items like tasks, calendar items, and contacts aren't available.
- For site items, only documents are available to preview. Preview for items like folders, lists, or list attachments aren't available.

File types supported when previewing search results

You can preview supported file types in the preview pane. If a file type isn't supported, you have to download a copy of the file to your local computer (by selecting **Download original item**). For .aspx Web pages, the URL for the page is included though you may not have permissions to access the page. Unindexed items aren't available for previewing.

The following file types are supported and can be previewed in the search results pane.

- .txt, .html, .mhtml
- .eml
- .doc, .docx, .docm
- .pptm, .pptx
- .pdf

Also, the following file container types are supported. You can view the list of files in the container in the preview pane.

- .zip
- .gzip

Last updated on 01/08/2026

View statistics for eDiscovery search results

> **⊗ Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

After you create and run a Content search or a search associated with a Microsoft Purview eDiscovery (Standard) case, you can view statistics about the estimated search results. This includes a summary of the search results (similar to the summary of the estimated search results displayed on the search flyout page), the query statistics such as the number of content locations with items that match the search query, and the identity of content locations that have the most matching items.

Additionally, you can use the keywords list to configure a search to return statistics for each keyword in a search query. This lets you compare the number of results returned by each keyword in a query.

You can also download search statistics to a comma-separated value (CSV) file. This lets you use the filtering and sorting features in Excel to compare results, and prepare reports for your search results.

Get statistics for searches

To display statistics for a Content search or a search associated with an eDiscovery (Standard) case.:

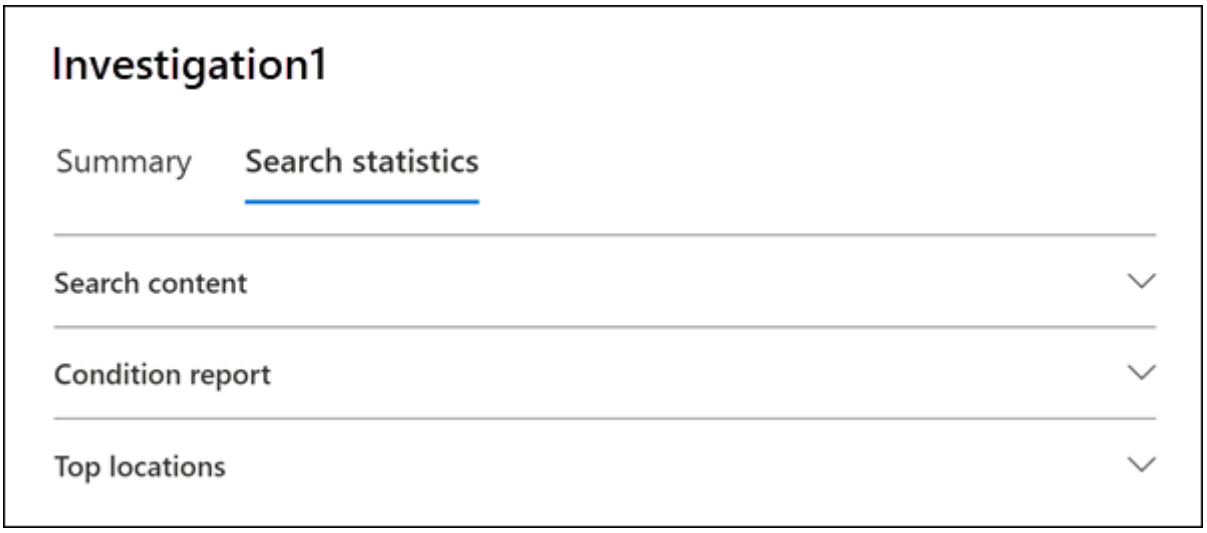
1. In the Microsoft Purview portal, select **Show all**, and then do one of the following:

   - Select **Content search** and then select a search to display the flyout page.

     OR

- Select **eDiscovery** > **eDiscovery (Standard)**, select a case, and then select a search on the **Searches** tab to display the flyout page.

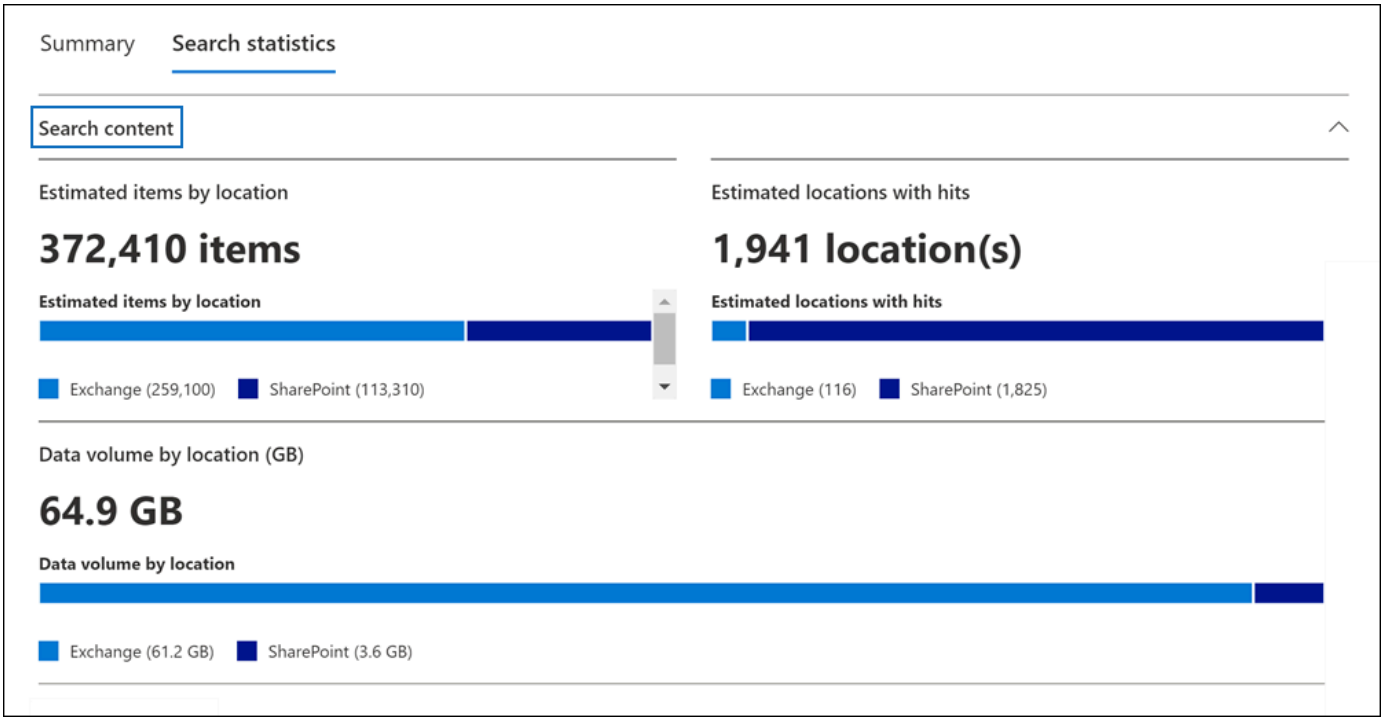
2. On the flyout page of the selected search, select the **Search statistics** tab.



The **Search statistics** tab contains for following sections that contain different types of statistics about the search.

Search content

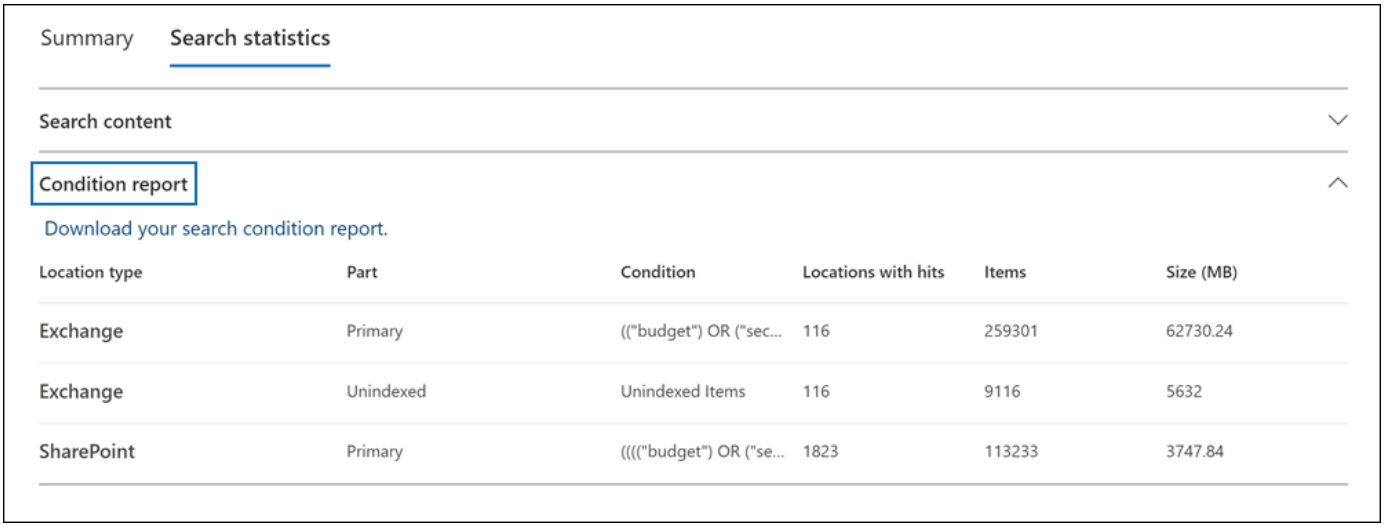
This section displays a graphical summary of the estimated items returned by the search. This indicates the number of items that match the search criteria. This information gives you an idea about the estimated number of items returned by the search.



- **Estimated items by locations**: The total number of estimated items returned by the search. The specific number of items located in mailboxes and located in sites is also displayed.
- **Estimated locations with hits**: The total number of content locations that contain items returned by the search. The specific number of mailbox and site locations is also displayed.
- **Data volume by location (in MB)**: The total size of all estimated items returned by the search. The specific size of mailbox items and site items is also displayed.

Condition report

This section displays statistics about the search query and the number of estimated items that matched different parts of the search query. You can use these statistics to analyze the number of items that match each component of search query. This can help you refine the search criteria and if necessary narrow the scope of the scope. You can also download a copy of this report in CSV format.



- **Location type**: The type of content location that the query statistics are applicable to. The value of **Exchange** indicates a mailbox location; a value of **SharePoint** indicates a site location.
- **Part**: The part of the search query the statistics are applicable to. **Primary** indicates the entire search query. **Keyword** indicates the statistics in the row are for a specific keyword. If you use a keyword list for search query, statistics for each component of the query are included in this table. For more information, see Get keyword statistics for searches.
- **Condition**: The actual component (keyword or condition) of the search query that returned the statistics displayed in the corresponding row.
- **Locations with hits**: The number of the content locations (specified by the **Location type** column) that contain items that match the primary or keyword query listed in the **Condition** column.

- **Items**: The number of items (from the specified content location) that match the query listed in the **Condition** column. As previously explained, if an item contains multiple instances of a keyword that is being searched for, it's only counted once in this column.
- **Size (MB)**: The total size of all items that were found (in the specified content location) that match the search query in the **Condition** column.

Top locations

This section displays statistics about the specific content locations with the most items returned by the search. The top 1,000 locations are displayed. You can also download a copy of this report in CSV format.

- The name of the location name (the email address of mailboxes and the URL for sites).
- Location type (a mailbox or site).
- Estimated number of items in the content location returned by the search.
- The total size of estimated items in each content location.

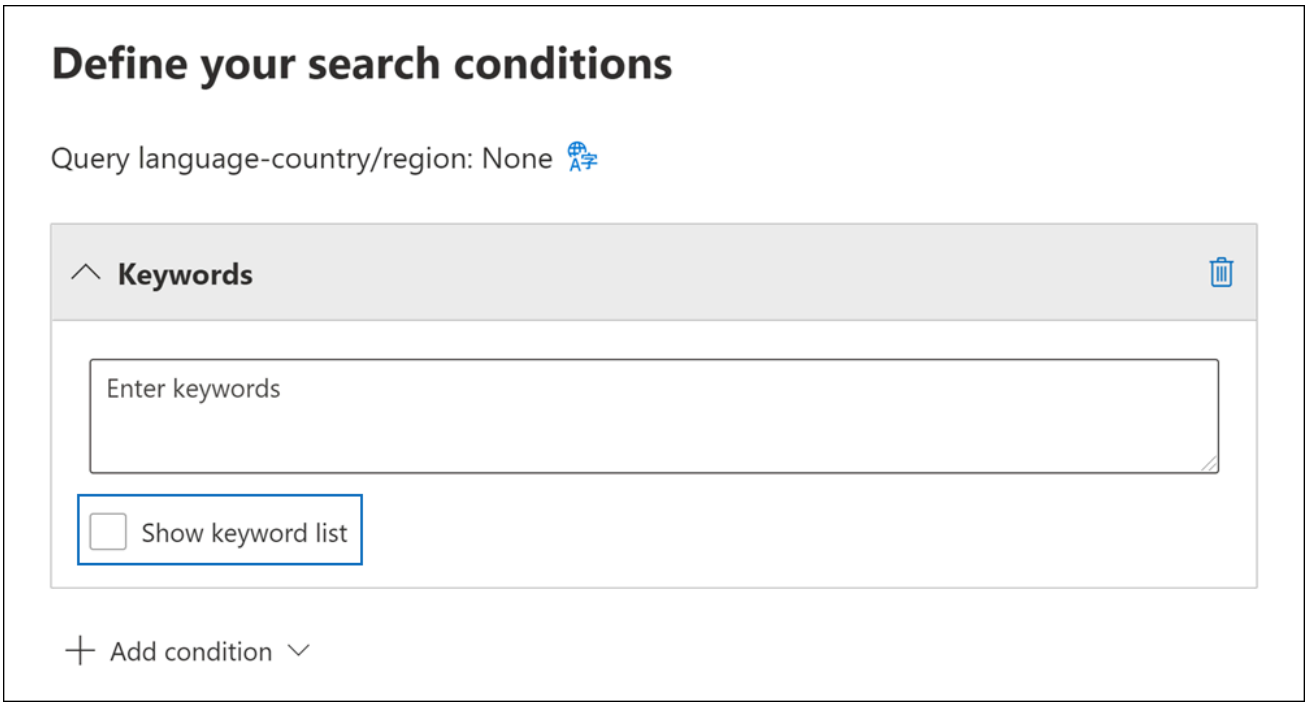
Get keyword statistics for searches

The **Condition report** section shows the search query and the number (and size) of items that match the query. If you use a keyword list when you create or edit a search query, you can get enhanced statistics that show how many items match each keyword or keyword phrase. This can help you quickly identify which parts of the query are the most (and least) effective. For example, if a keyword returns a large number of items, you might choose to refine the keyword query to narrow the search results.

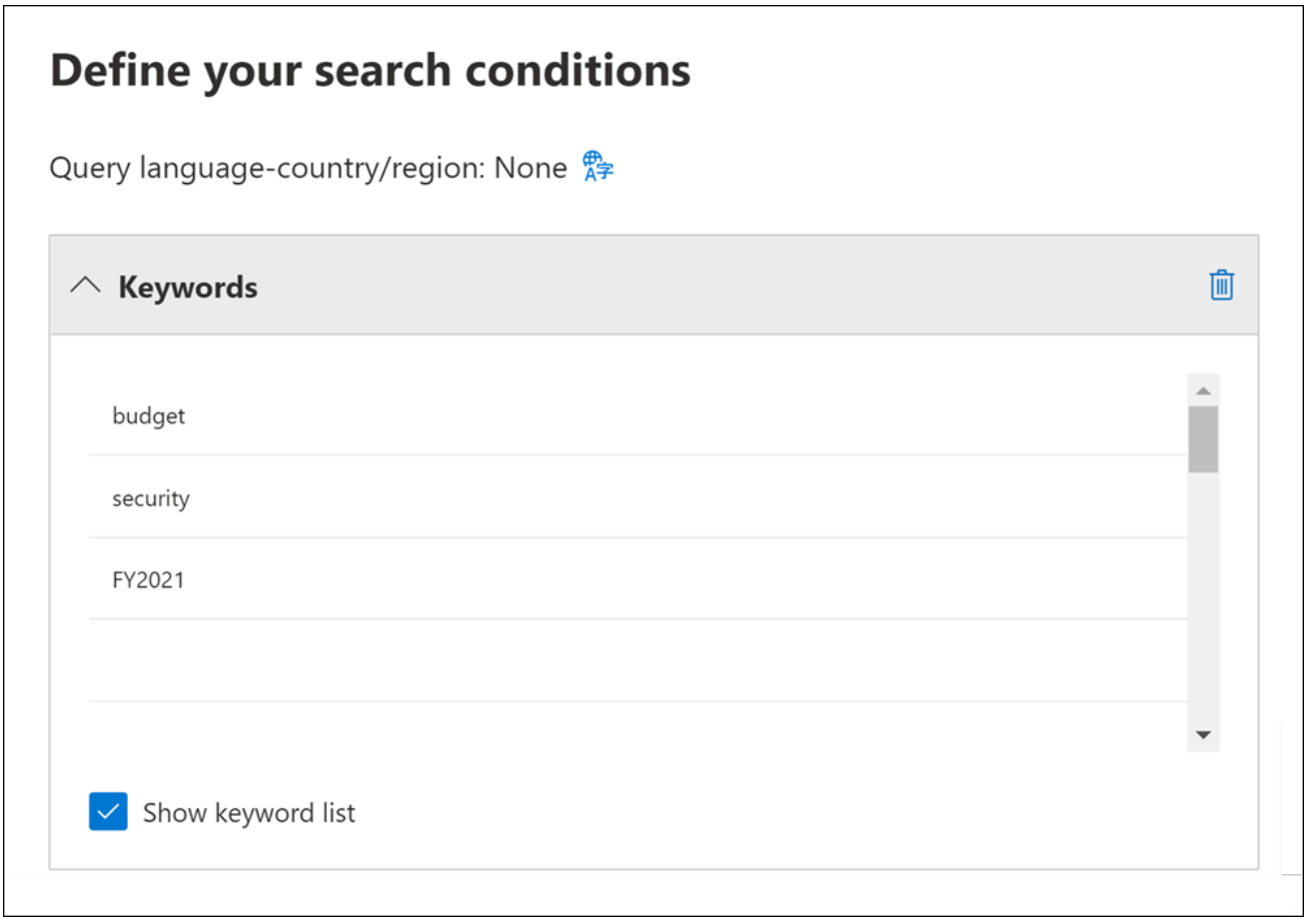
To create a keyword list and view keyword statistics for a search:

1. In the Microsoft Purview portal, create a new Content search or a search associated with an eDiscovery (Standard) case.

2. On the **Conditions** page of the search workflow. select the **Show keyword list** checkbox.



3. Type a keyword or keyword phase in a row in the keywords table. For example, type **budget** in the first row, type **security** in the second row, and type **FY2021** in the third row.



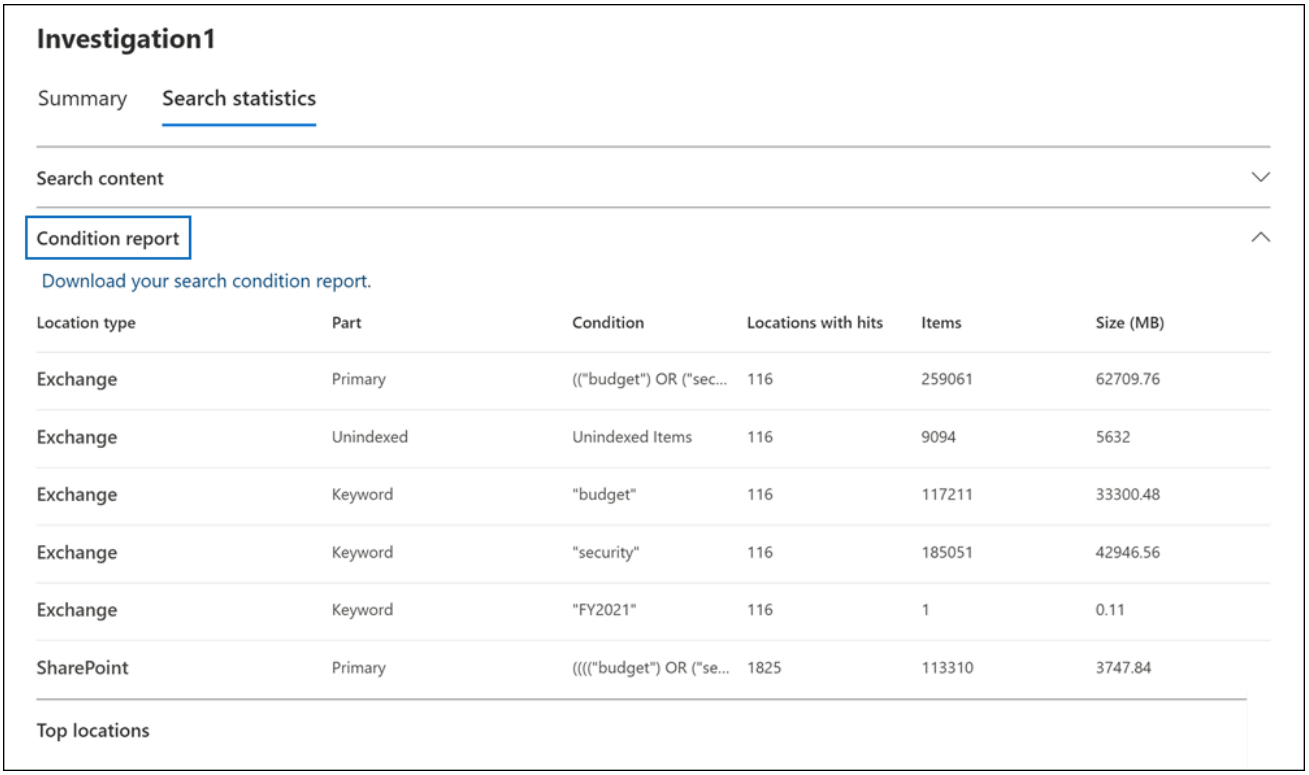
> **Note**

To help reduce issues caused by large keyword lists, you're limited to a maximum of 20 rows in the keyword list of a search query.

4. After adding the keywords to the list that you want to search and get statistics for, run the search.

5. When the search is completed, select it to display the flyout page.

6. On the **Search statistics** tab, select the **Condition report** to display the keyword statistics for the search.



As shown in the previous screenshot, the statistics for each keyword are displayed; this includes:

- The keyword statistics for each type of content location included in the search.
- The number of unindexed mailbox items.
- The actual search query and results for each keyword (identified as **Keyword** in the **Part** column), which includes any conditions from the search query.
- The complete search query (identified as **Primary** in the **Part** column) and the statistics for the complete query for each location type. Note these are the same statistics displayed on the **Summary** tab.

Last updated on 11/21/2025

Export Content search results

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

After a Content search is successfully run, you can export the search results to a local computer. When you export email results, they're downloaded to your computer as PST files. When you export content from SharePoint and OneDrive for Business sites, copies of native Office documents are exported. There are other documents and reports included with the exported search results.

Exporting the results of a Content search involves preparing the results, and then downloading them to a local computer. These steps for exporting search results also apply to exporting the results of a search that's associated with Microsoft Purview eDiscovery (Standard) cases.

Before you export search results

> **Important**
>
> The eDiscovery Export tool was permanently **retired on August 31, 2025**, this functionality isn't available in the new eDiscovery experience.

- To export search results, you have to be assigned the Export management role in Microsoft Purview portal. This role is assigned to the built-in eDiscovery Manager role group. It isn't assigned by default to the Organization Management role group. For more information, see Assign eDiscovery permissions.
- The computer you use to export the search results has to meet the following system requirements:

    - Latest version of Windows (32-bit or 64-bit)
    - Microsoft .NET Framework 4.7 or higher

- You have to use Microsoft Edge[1] to run the eDiscovery Export Tool. Using Internet Explorer 11 to export search results is no longer supported[2].

  > **Note**
  >
  > [1] As a result of recent changes to Microsoft Edge, SelectOnce support is no longer enabled by default. Also, Microsoft doesn't manufacture third-party extensions or add-ons for SelectOnce applications. Exporting search results using an unsupported browser with third-party extensions or add-ons isn't supported.
  >
  > [2] Beginning August 2021, Microsoft 365 apps and services will no longer support Internet Explorer 11 (IE11) and users may have a degraded experience or be unable to connect to those apps and services. These apps and services will phase out over the upcoming weeks and months to ensure a smooth end of support. Each app and service are being phased out on independent schedules. For more information, see this **blog post**.

- The eDiscovery Export Tool that you use in Step 2 to download search results doesn't support automation (by using a script or running cmdlets). We highly recommended that you don't automate the download process in Step 2. If you automate this processes, Microsoft Support won't provide assistance if you run into issues.

- We recommend downloading search results to a local computer. To eliminate your company's firewall or proxy infrastructure from causing issues when downloading search results, you might consider downloading search results to a virtual desktop outside of your network. This may decrease timeouts that occur in Azure data connections when exporting a large number of files. For more information about virtual desktops, see Windows Virtual Desktop.

- To improve performance when downloading search results, consider dividing searches that return a large set of results into smaller searches. For example, you can use date ranges in search queries to return a smaller set of results that can be downloaded faster.

- When you export search results, the data is temporarily stored in a Microsoft-provided Azure Storage location in the Microsoft cloud before it's downloaded to your local computer. Be sure that your organization can connect to the endpoint in Azure, which is

***.blob.core.windows.net** (the wildcard represents a unique identifier for your export). The search results data is deleted from the Azure Storage location two weeks after it's created.

- If your organization uses a proxy server to communicate with the Internet, you need to define the proxy server settings on the computer that you use to export the search results (so the export tool can be authenticated by your proxy server). To do this, open the *machine.config* file in the location that matches your version of Windows.

  - **32-bit:** `%windir%\Microsoft.NET\Framework\[version]\Config\machine.config`
  - **64-bit:** `%windir%\Microsoft.NET\Framework64\[version]\Config\machine.config`

    Add the following lines to the *machine.config* file somewhere between the `<configuration>` and `</configuration>` tags. Be sure to replace `ProxyServer` and `Port` with the correct values for your organization; for example, `proxy01.contoso.com:80`.

    XML

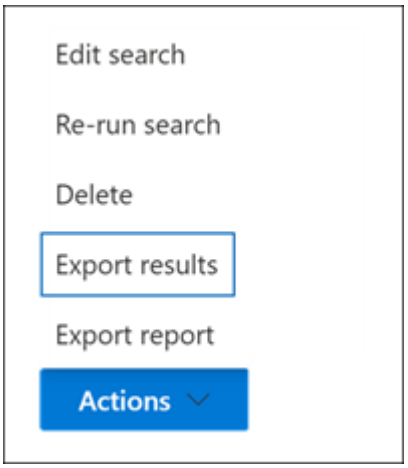
    ```
    <system.net>
       <defaultProxy enabled="true" useDefaultCredentials="true">
         <proxy proxyaddress="https://ProxyServer :Port "
                usesystemdefault="False"
                bypassonlocal="True"
                autoDetect="False" />
       </defaultProxy>
    </system.net>
    ```

- If the results of a search are older than 7 days and you submit an export job, an error message is displayed prompting you to rerun the search to update the search results. If this happens, cancel the export, rerun the search, and then start the export again.

Step 1: Prepare search results for export

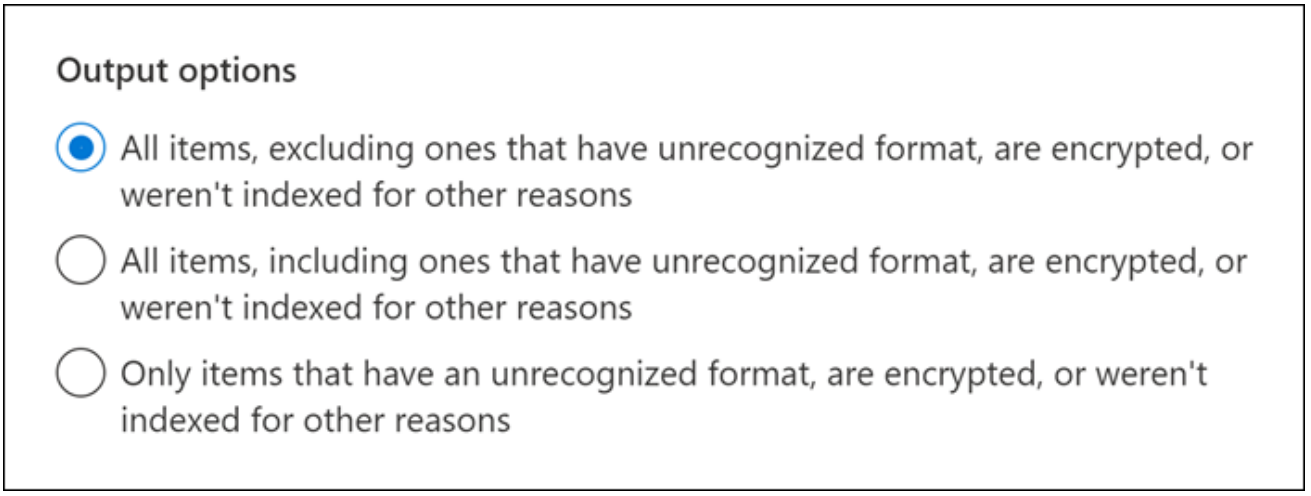
The first step is to prepare the search results for exporting. When you prepare results, they're uploaded to a Microsoft-provided Azure Storage location in the Microsoft cloud. Content from mailboxes and sites is uploaded at a maximum rate of 2 GB per hour.

1. In the Microsoft Purview portal, select the content search that you want to export results from.
2. On the **Actions** menu at the bottom of the flyout page, select **Export results**.



The **Export results** flyout page is displayed. The export options available to export content depend on whether search results are located in mailboxes or sites or a combination of both.

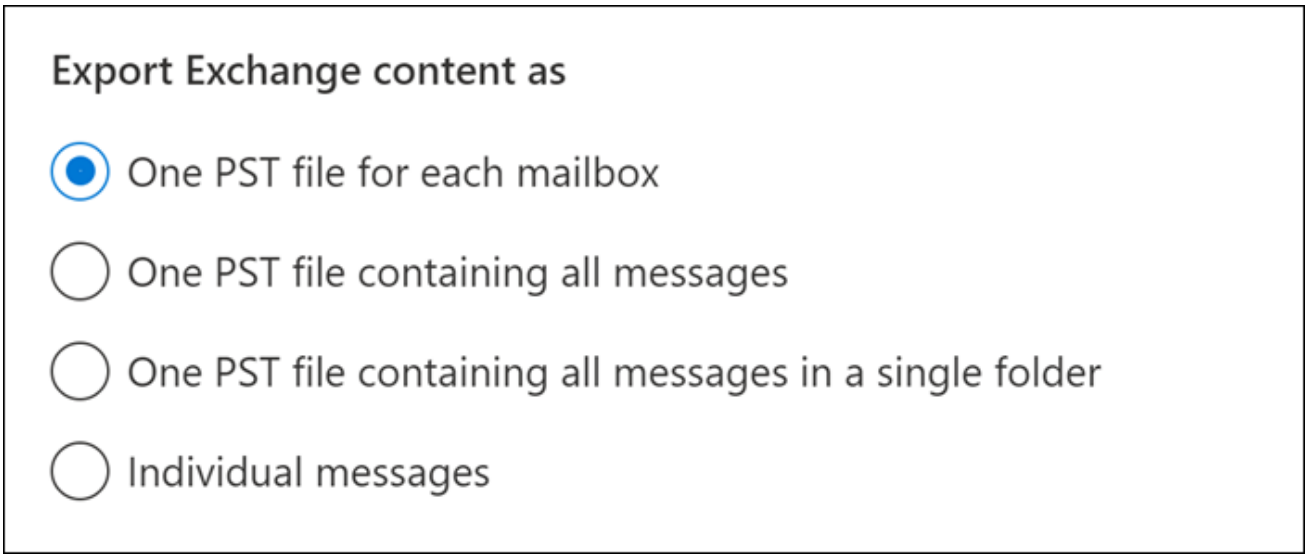
3. Under **Output options**, choose one of the following options:



- **All items, excluding ones that have unrecognized format, are encrypted, or weren't indexed for other reasons**. This option exports only indexed items.
- **All items, including ones that have unrecognized format, are encrypted, or weren't indexed for other reasons**. This option exports indexed and unindexed items.
- **Only items that have an unrecognized format, are encrypted, or weren't indexed for other reasons**. This option exports only unindexed items.

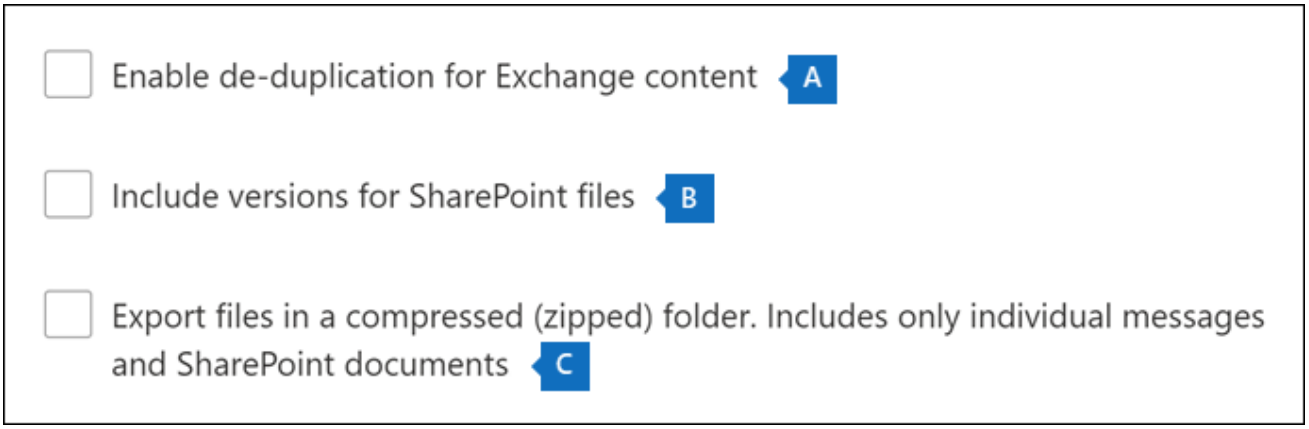
  See the More information section for a description about how partially indexed items are exported. For more information about partially indexed items, see Partially indexed items in Content search.

4. Under **Export Exchange content as**, choose one of the following options:



- **One PST file for each mailbox**: Exports one PST file for each user mailbox that contains search results. Any results from the user's archive mailbox are included in the same PST file. This option reproduces the mailbox folder structure from the source mailbox.
- **One PST file containing all messages**: Exports a single PST file (named *Exchange.pst*) that contains the search results from all source mailboxes included in the search. This option reproduces the mailbox folder structure for each message.
- **One PST file containing all messages in a single folder**: Exports search results to a single PST file where all messages are located in a single, top-level folder. This option lets reviewers review items in chronological order (items are sorted by sent date) without having to navigate the original mailbox folder structure for each item.
- **Individual messages**: Exports search results as individual email messages, using the .msg format. If you select this option, email search results are exported to a folder in the file system. The folder path for individual messages is the same as the one used if you exported the results to a PST file.

5. Configure the following additional options:



a. Select the **Enable de-duplication for Exchange content** checkbox to exclude duplicate messages.

   If you select this option, only one copy of a message will be exported even if multiple copies of the same message are found in the mailboxes that were searched. The export results report (which is a file named Results.csv) will contain a row for every copy of a duplicate message so that you can identify the mailboxes (or public folders) that contain a copy of the duplicate message.

b. Select the **Include versions for SharePoint files** checkbox to export all versions of SharePoint documents. This option appears only if the content sources of the search include SharePoint or OneDrive for Business sites.

c. Select the **Export files in a compressed (zipped) folder. Includes only individual messages and SharePoint documents** checkbox to export search results to compressed folders. This option appears only when you choose to export Exchange items as individual messages and when the search results include SharePoint or OneDrive documents. This option is primarily used to work around the 260 character limit in Windows file path names when items are exported. See the "Filenames of exported items" in the More information section.

> ⓘ **Important**
>
> Exporting files in a compressed (zipped) folder will increase export times.

6. Select **Export** to start the export process. The search results are prepared for downloading, which means they're collected from the original content locations and then uploaded to an Azure Storage location in the Microsoft cloud. This may take several minutes.

See the next section for instructions to download the exported search results.

Step 2: Download the search results

The next step is to download the search results from the Azure Storage location to your local computer.

> ⓘ **Important**

The exported search results must be downloaded within 14 days after you created the export job in Step 1.

1. On the **Content search** page in the Microsoft Purview portal, select the **Exports** tab

   You may have to select **Refresh** to update the list of export jobs so that it shows the export job you created. Export jobs have the same name as the corresponding search with **_Export** appended to the search name.

2. Select the export job that you created in Step 1.

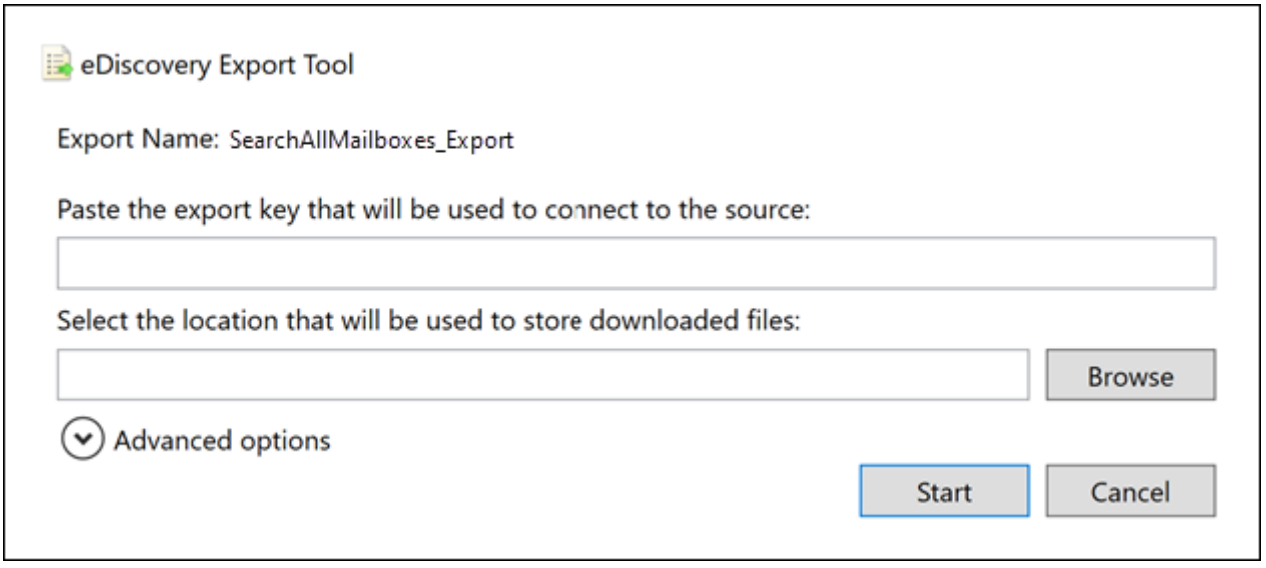
3. On the flyout page under **Export key**, select **Copy to clipboard**. You use this key in step 6 to download the search results.

   > **ⓘ Important**
   >
   > Because anyone can install and start the eDiscovery Export tool, and then use this key to download the search results, be sure to take precautions to protect this key just like you would protect passwords or other security-related information.

4. At the top of the flyout page, select **Download results**.

5. If you're prompted to install the **eDiscovery Export Tool**, select **Install**.

6. In the **eDiscovery Export Tool**, do the following:



   a. Paste the export key that you copied in Step 3 in the appropriate box.

   b. Select **Browse** to specify the location where you want to download the search result files.

> **ⓘ Important**
>
> Due to high network activity during download, you should download search results only to a location on an internal drive on your local computer. For the best download experience, follow these guidelines:
>
> - Don't download search results to a UNC path, a mapped network drive, an external USB drive, or a synched OneDrive for Business account.
> - Disable anti-virus scanning for the folder that you download the search result to.
> - Download search results to different folders for concurrent download jobs.

7. Select **Start** to download the search results to your computer.

   The **eDiscovery Export Tool** displays status information about the export process, including an estimate of the number (and size) of the remaining items to be downloaded. When the export process is complete, you can access the files in the location where they were downloaded.

More information

Here's more information about exporting search results.

- Export reports
- Exporting partially indexed items
- Exporting individual messages or PST files
- Decrypting RMS-protected email messages and encrypted file attachments
- Filenames of exported items
- Miscellaneous

Export reports

- When you export search results, the following reports are included in addition to the search results.
  - **Export Summary** An Excel document that contains a summary of the export. This includes information such as the number of content sources that were searched, the estimated and downloaded sizes of the search results, and the estimated and downloaded number of items that were exported.

- **Manifest** A manifest file (in XML format) that contains information about each item included in the search results.
- **Results** An Excel document that contains information about each item that is download as a search result. For email, the result log contains information about each message, including:
  - The location of the message in the source mailbox (including whether the message is in the primary or archive mailbox).
  - The date the message was sent or received.
  - The Subject line from the message.
  - The sender and recipients of the message.
  - Whether the message is a duplicate message if you enabled the de-duplication option when exporting the search results. Duplicate messages have a value in the **Duplicate to Item** column that identifies the message as a duplicate. The value in the **Duplicate to Item** column contains the item identity of the message that was exported.

    For documents from SharePoint and OneDrive for Business sites, the result log contains information about each document, including:
    - The URL for the document.
    - The URL for the site collection where the document is located.
    - The date that the document was last modified.
    - The name of the document (which is located in the Subject column in the result log).
- **Unindexed Items** An Excel document that contains information about any partially indexed items that would be included in the search results. If you don't include partially indexed items when you generate the search results report, this report will still be downloaded, but will be empty.
- **Errors and Warnings** Contains errors and warnings for files encountered during export. See the Error Details column for information specific to each individual error or warning.
- **Skipped Items** When you export search results from SharePoint and OneDrive for Business sites, the export will usually include a skipped items report (SkippedItems.csv). The items cited in this report are typically items that won't be downloaded, such as a folder or a document set. Not exporting these types of items is by design. For other

items that were skipped, the 'Error Type' and 'Error Details' field in the skipped items report show the reason the item was skipped and wasn't downloaded with the other search results.

- **Trace.log** Contains detailed logging information about the export process and can help uncover issues during export. If you open a ticket with Microsoft Support about an issue related to exporting search results, you may be asked to provide this trace log.

  > ⓘ **Note**
  >
  > You can just export these documents without having to export the actual search results.

Exporting partially indexed items

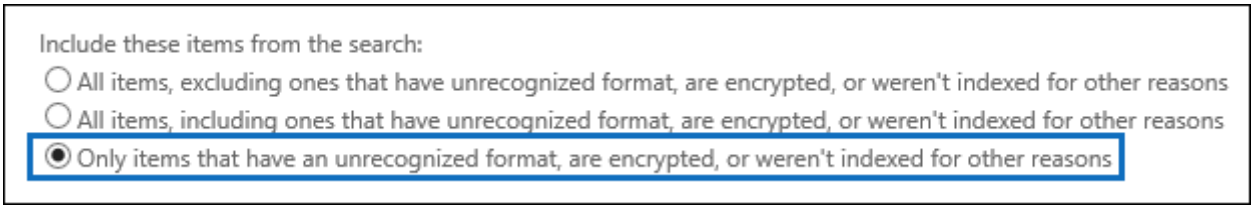
- If you're exporting mailbox items from a content search that returns all mailbox items in the search results (because no keywords where included in the search query), partially indexed items won't be copied to the PST file that contains the unindexed items. This is because all items, including any partially indexed items, are automatically included in the regular search results. This means that partially indexed items will be included in a PST file (or as individual messages) that contains the other, indexed items.

  If you export both the indexed and partially indexed items or if you export only the indexed items from a content search that returns all items, the same number of items will be downloaded. This happens even though the estimated search results for the content search (displayed in the search statistics in the Microsoft Purview portal) will still include a separate estimate for the number of partially indexed items. For example, let's say that the estimate for a search that includes all items (no keywords in the search query) shows that 1,000 items were found and that 200 partially indexed items were also found. In this case, the 1,000 items include the partially indexed items because the search returns all items. In other words, there are 1,000 total items returned by the search, and not 1,200 items (as you might expect). If you export the results of this search and choose to export indexed and partially indexed items (or export only partially indexed items), then 1,000 items are downloaded. Again, that's because partially indexed items are included with the regular (indexed) results when you use a blank search query to return all items. In this same example, if you choose to export only partially indexed items, then only the 200 unindexed items would be downloaded.

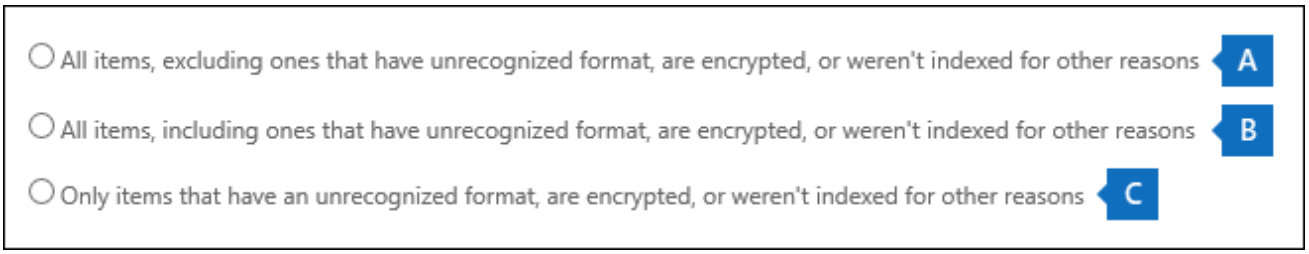
Also note that in the previous example (when you export indexed and partially indexed items or you export only indexed items), the **Export Summary** report included with the exported search results would list 1,000 items estimated items and 1,000 downloaded items for the same reasons as previously described.

- If the search that you're exporting results from was a search of specific content locations or all content locations in your organization, only the partial items from content locations that contain items that match the search criteria will be exported. In other words, if no search results are found in a mailbox or site, then any partially indexed items in that mailbox or site won't be exported. The reason for this is that exporting partially indexed items from lots of locations in the organization might increase the likelihood of export errors and increase the time it takes to export and download the search results.

  To export partially indexed items from all content locations for a search, configure the search to return all items (by removing any keywords from the search query) and then export only partially indexed items when you export the search results.



- When exporting search results from SharePoint or OneDrive for Business sites, the ability to export unindexed items also depends on the export option that you select and whether a site that was searched contains an indexed item that matches the search criteria. For example, if you search specific SharePoint or OneDrive for Business sites and no search results are found, then no unindexed items from those sites will be exported if you choose the second export option to export both indexed and unindexed items. If an indexed item from a site does match the search criteria, then all unindexed items from that site will be exported when exporting both indexed and unindexed items. The following illustration describes the export options based on whether a site contains an indexed item that matches the search criteria.



a. Only indexed items that match the search criteria are exported. No partially indexed items are exported.

b. If no indexed items from a site match the search criteria, then partially indexed items from that same site aren't exported. If indexed items from a site are returned in the search results, then the partially indexed items from that site are exported. In other words, only the partially indexed items from sites that contain items that match the search criteria are exported.

c. All partially indexed items from all sites in the search are exported, regardless of whether a site contains items that match the search criteria.

If you choose to export partially indexed items, partially indexed mailbox items are exported in a separate PST file regardless of the option that you choose under **Export Exchange content as**.

- If partially indexed items are returned in the search results (because other properties of partially indexed items matched the search criteria), then those partially indexed are exported with the regular search results. So, if you choose to export both indexed items and partially indexed items (by selecting the **All items, including ones that have unrecognized format, are encrypted, or weren't indexed for other reasons** export option), the partially indexed items exported with the regular results will be listed in the Results.csv report. They won't be listed in the Unindexed items.csv report.

Exporting individual messages or PST files

- If the file path name of a message exceeds the maximum character limit for Windows, the file path name is truncated. But the original file path name will be listed in the Manifest and ResultsLog.
- As previously explained, email search results are exported to a folder in the file system. The folder path for individual messages would replicate the folder path in the user's mailbox. For example, for a search named "ContosoCase101" messages in a user's inbox would be located in the folder path `~ContosoCase101\\<date of export\Exchange\user@contoso.com (Primary)\Top of Information Store\Inbox`.
- If you choose to export email messages in one PST file containing all messages in a single folder, a **Deleted Items** folder and a **Search Folders** folder are included in the top level of the PST folder. These folders are empty.
- As previously stated, you must export email search results as individual messages to decrypt RMS-protected messages when they're exported. Encrypted messages remain encrypted if

you export email search results as a PST file.

Decrypting RMS-protected email messages and encrypted file attachments

Any rights-protected (RMS-protected) email messages included in the results of a Content search will be decrypted when you export them. Additionally, any file that's encrypted with a Microsoft encryption technology and is attached to an email message that's included in the search results will also be decrypted when it's exported. This decryption capability is enabled by default for members of the eDiscovery Manager role group. This is because the RMS Decrypt management role is assigned to this role group by default. Keep the following things in mind when exporting encrypted email messages and attachments:

- As previously explained, to decrypt RMS-protected messages when you export them, you have to export the search results as individual messages. If you export search results to a PST file, RMS-protected messages remain encrypted.
- Messages that are decrypted are identified in the **ResultsLog** report. This report contains a column named **Decode Status**, and a value of **Decoded** in this column identifies the messages that were decrypted.
- In addition to decrypting file attachments when exporting search results, you can also preview the decrypted file when previewing search results. You can only view the rights-protected email message after you export it.
- At this time, the decryption capability when exporting search results doesn't include encrypted content from SharePoint and OneDrive for Business sites. However, support is coming soon for documents encrypted with Microsoft encryption technologies and stored in SharePoint Online and OneDrive for Business.
- If you need to prevent someone from decrypting RMS-protect messages and encrypted file attachments, you have to create a custom role group (by copying the built-in eDiscovery Manager role group) and then remove the RMS Decrypt management role from the custom role group. Then add the person who you don't want to decrypt messages as a member of the custom role group.

Filenames of exported items

- There's a 260-character limit (imposed by the operating system) for the full path name for email messages and site documents exported to your local computer. The full path name for exported items includes the item's original location and the folder location on the local computer where the search results are downloaded to. For example, if you specify to

download the search results to `C:\Users\Admin\Desktop\SearchResults` in the eDiscovery Export tool, then the full pathname for a downloaded email item would be `C:\Users\Admin\Desktop\SearchResults\ContentSearch1\03.15.2017-1242PM\Exchange\sarad@contoso.com (Primary)\Top of Information Store\Inbox\Insider trading investigation.msg`.

- If the 260-character limit is exceeded, the full path name for an item will be truncated, based on the following:

  - If the full path name is longer than 260 characters, the file name is shortened to get under the limit; note that the truncated filename (excluding the file extension) won't be fewer than eight characters.

  - If the full path name is still too long after shortening the file name, the item is moved from its current location to the parent folder. If the pathname is still too long, then the process is repeated: shorten the filename, and if necessary move again to the parent folder. This process is repeated until the full pathname is under the 260-character limit.

  - If a truncated full path name already exists, a version number is added to the end of the filename; for example, `statusmessage(2).msg`.

    To help mitigate this issue, consider downloading search results to a location with a short path name; for example, downloading search results to a folder named `C:\Results` would add fewer characters to the path names of exported items than downloading them to a folder named `C:\Users\Admin\Desktop\Results`.

- When you export site documents, it's also possible that the original file name of a document will be modified. This happens specifically for documents that have been deleted from a SharePoint or OneDrive for Business site that's been placed on hold. After a document that's on a site that's on hold is deleted, the deleted document is automatically moved to the Preservation Hold library for the site (which was created when the site was placed on hold). When the deleted document is moved to the Preservation Hold library, a randomly generated and unique ID is appended to the original filename of the document. For example, if the filename for a document is `FY2017Budget.xlsx` and that document is later deleted and moved to the Preservation Hold library, the filename of the document that is moved to the Preservation Hold library is modified to something like `FY2017Budget_DEAF727D-0478-4A7F-87DE-5487F033C81A2000-07-05T10-37-55.xlsx`. If a document in the Preservation Hold library matches the query of a Content search and you export the results of that search, the exported file has the modified filename; in this

example, the filename of the exported document would be `FY2017Budget_DEAF727D-0478-4A7F-87DE-5487F033C81A2000-07-05T10-37-55.xlsx`.

When a document on a site that's on hold is modified (and versioning for the document library in the site has been enabled), a copy of the file is automatically created in the Preservation Hold library. In this case, a randomly generated and unique ID is also appended to the filename of the document that's copied to the Preservation Hold library.

The reason why filenames of documents that are moved or copied to the Preservation Hold library is to prevent conflicting filenames. For more information about placing a hold on sites and the Preservation Hold library, see Overview of in-place hold in SharePoint Server 2016.

Miscellaneous

- When downloading search results using the eDiscovery Export Tool, it's possible you might receive the following error: `System.Net.WebException: The remote server returned an error: (412) The condition specified using HTTP conditional header(s) is not met.` This is transient error, which typically occurs in the Azure Storage location. To resolve this issue, retry downloading the search results, which will restart the eDiscovery Export Tool.
- All search results and the export reports are included in a folder that has the same name as the Content search. The email messages that were exported are located in a folder named **Exchange**. Documents are located in a folder named **SharePoint**.
- The file system metadata for documents on SharePoint and OneDrive for Business sites is maintained when documents are exported to your local computer. That means document properties, such as created and last modified dates, aren't changed when documents are exported.
- If your search results include a list item from SharePoint that matches the search query, all rows in the list will be exported in addition to the item that matches the search query and any attachments in the list. The reason for this behavior is to provide a context for list items that are returned in the search results. The other list items and attachments may cause the count of exported items to be different than the original estimate of search results.

Last updated on 10/31/2024

Search for and delete email messages

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

> **Tip**
>
> You can also use the purge feature in Data Security Investigations to search for and permanently delete Exchange mailbox and Microsoft Teams items in your organization. Data Security Investigations provides a built-in search and purge workflow with a purge queue dashboard to help you manage and track purge queries. For more information, see Purge in Data Security Investigations.

You can use the Content search feature to search for and delete email messages from all mailboxes in your organization. This can help you find and remove potentially harmful or high-risk email, such as:

- Messages that contain dangerous attachments or viruses
- Phishing messages
- Messages that contain sensitive data

> **Tip**
>
> If your organization has a Defender for Office 365 Plan 2 subscription, we recommend using the procedure detailed in Remediate malicious email delivered in Office 365, rather than following the procedure described in this article.

Before you begin

- The search and purge workflow described in this article doesn't delete chat messages or other content from Microsoft Teams. If the Content search that you create in Step 2 returns items from Microsoft Teams, those items won't be deleted when you purge items in Step 3. To search for and delete chat messages, see Search and purge chat messages in Teams.

- To create and run a Content search, you have to be a member of the *eDiscovery Manager* role group or be assigned the *Compliance Search* role in the Microsoft Purview portal. To delete messages, you have to be a member of the *Organization Management* role group or be assigned the *Search And Purge* role in the Microsoft Purview portal For information about adding users to a role group, see Assign eDiscovery permissions.

  > ⓘ **Note**
  >
  > The *Organization Management* role group exists in both Exchange Online and in the Microsoft Purview portal. These are separate role groups that give different permissions. Being a member of *Organization Management* in Exchange Online does not grant the required permissions to delete email messages. If you aren't assigned the *Search And Purge* role in the Microsoft Purview portal (either directly or through a role group such as *Organization Management*), you'll receive an error in Step 3 when you run the *New-ComplianceSearchAction* cmdlet with the message "A parameter cannot be found that matches parameter name 'Purge'".

- You have to use Security & Compliance PowerShell to delete messages. See Step 1: Connect to Security & Compliance PowerShell for instructions about how to connect.

- A maximum of 10 items per mailbox can be removed at one time. Because the capability to search for and remove messages is intended to be an incident-response tool, this limit helps ensure that messages are quickly removed from mailboxes. This feature isn't intended to clean up user mailboxes.

  > ⓘ **Note**
  >
  > This option isn't supported if a mailbox has a litigation hold or in-place hold. Only 10 items are removed from view of the user. These 10 items aren't permanently deleted, so these 10 items are the only ones processed.

- The maximum number of mailboxes in a content search that you can use to delete items by doing a search and purge action is 50,000. If the search (that you create in Step 2 searches

more than 50,000 mailboxes, the purge action (that you create in Step 3) will fail. Searching more than 50,000 mailbox in a single search might typically happen when you configure the search to include all mailboxes in your organization. This restriction still applies even when less than 50,000 mailboxes contain items that match the search query. See the More information section for guidance about using search permissions filters to search for and purge items from more than 50,000 mailboxes.

- The procedure in this article can only be used to delete items in Exchange Online mailboxes and public folders. You can't use it to delete content from SharePoint or OneDrive for Business sites.

- Email items in a review set in an eDiscovery (Premium) case can't be deleted by using the procedures in this article. That's because items in a review set are stored in an Azure Storage location, and not in the live service. This means they won't be returned by the content search that you create in Step 1. To delete items in a review set, you have to delete the eDiscovery (Premium) case that contains the review set. For more information, see Close or delete an eDiscovery (Premium) case.

Step 1: Connect to Security & Compliance PowerShell

The first step is to connect to Security & Compliance PowerShell for your organization. For step-by-step instructions, see Connect to Security & Compliance PowerShell.

Step 2: Create a Content Search to find the message to delete

The second step is to create and run a Content search to find the message that you want to remove from mailboxes in your organization. You can create the search by using the Microsoft Purview portal or by running the **New-ComplianceSearch** and **Start-ComplianceSearch** cmdlets in Security & Compliance PowerShell. The messages that match the query for this search will be deleted by running the **New-ComplianceSearchAction -Purge** command in Step 3. For information about creating a Content search and configuring search queries, see the following articles:

- Content search in Office 365
- Keyword queries for Content search
- New-ComplianceSearch
- Start-ComplianceSearch

> **Note**
>
> The content locations that are searched in the Content search that you create in this step can't include SharePoint or OneDrive for Business sites. You can include only mailboxes and public folders in a Content search that will be used to email messages. If the Content search includes sites, you'll receive an error in Step 3 when you run the **New-ComplianceSearchAction** cmdlet.

Tips for finding messages to remove

The goal of the search query is to narrow the results of the search to only the message or messages that you want to remove. Here are some tips:

- If you know the exact text or phrase used in the subject line of the message, use the **Subject** property in the search query.
- If you know that exact date (or date range) of the message, include the **Received** property in the search query.
- If you know who sent the message, include the **From** property in the search query.
- Preview the search results to verify that the search returned only the message (or messages) that you want to delete.
- Use the search estimate statistics (displayed in the details pane of the search in the Microsoft Purview portal or by using the Get-ComplianceSearch cmdlet) to get a count of the total number of results.

Here are two examples of queries to find suspicious email messages.

- This query returns messages that were received by users between April 13, 2016 and April 14, 2016 and that contain the words "action" and "required" in the subject line.

  PowerShell

  ```
  (Received:4/13/2016..4/14/2016) AND (Subject:'Action required')
  ```

- This query returns messages that were sent by user@contoso.com and that contain the exact phrase "Update your account information" in the subject line.

  PowerShell

  ```
  (From:user@contoso.com) AND (Subject:"Update your account information")
  ```

Here's an example of using a query to create and start a search by running the **New-ComplianceSearch** and **Start-ComplianceSearch** cmdlets to search all mailboxes in the organization:

PowerShell

```
$Search=New-ComplianceSearch -Name "Remove Phishing Message" -ExchangeLocation All -ContentMatchQuery '(Received:4/13/2016..4/14/2016) AND (Subject:"Action required")'
Start-ComplianceSearch -Identity $Search.Identity
```

Step 3: Delete the message

After you've created and refined a Content search to return the messages that you want to remove, the final step is to run the **New-ComplianceSearchAction -Purge** command in Security & Compliance PowerShell to delete the message.

You can soft- or hard-delete the message. A soft-deleted message is moved to a user's Recoverable Items folder and retained until the deleted item retention period expires. Hard-deleted messages are marked for permanent removal from the mailbox and will be permanently removed the next time the mailbox is processed by the Managed Folder Assistant. If single item recovery is enabled for the mailbox, hard-deleted items will be permanently removed after the deleted item retention period expires. If a mailbox is placed on hold, deleted messages are preserved until the hold duration for the item expires or until the hold is removed from the mailbox.

> ⓘ **Note**
>
> As previously stated, items from Microsoft Teams that are returned by Content search are not deleted when you run the **New-ComplianceSearchAction -Purge** command.

To run the following commands to delete messages, be sure that you're connected to Security & Compliance PowerShell.

Soft-delete messages

In the following example, the command soft-deletes the search results returned by a Content search named "Remove Phishing Message".

PowerShell

```
New-ComplianceSearchAction -SearchName "Remove Phishing Message" -Purge -PurgeType
SoftDelete
```

Hard-delete messages

To hard-delete the items returned by the "Remove Phishing Message" content search, you would run this command:

PowerShell

```
New-ComplianceSearchAction -SearchName "Remove Phishing Message" -Purge -PurgeType
HardDelete
```

When you run the previous commands to soft- or hard-delete messages, the search specified by the *SearchName* parameter is the Content search that you created in Step 1.

For more information, see New-ComplianceSearchAction.

More information

- **How do you get status on the search and remove operation?**

  Run the **Get-ComplianceSearchAction** to get the status on the delete operation. The object that is created when you run the **New-ComplianceSearchAction** cmdlet is named using this format: `<name of Content Search>_Purge`.

- **What happens after you delete a message?**

  A message that's deleted with the `New-ComplianceSearchAction -Purge -PurgeType HardDelete` command is moved to the Purges folder and can't be accessed by the user. After the message is moved to the Purges folder, the message is retained for the duration of the deleted item retention period if single item recovery is enabled for the mailbox. (In Microsoft 365, single item recovery is enabled by default when a new mailbox is created.) After the deleted item retention period expires, the message is marked for permanent deletion and will be purged from Microsoft 365 the next time the mailbox is processed by the Managed Folder assistant.

  If you use the `New-ComplianceSearchAction -Purge -PurgeType SoftDelete` command, messages are moved to the Deletions folder in the user's Recoverable Items folder. It isn't immediately purged from Microsoft 365. The user can recover messages in the Deleted Items folder for the duration based on the deleted item retention period configured for the

mailbox. After this retention period expires (or if user purges the message before it expires), the message is moved to the Purges folder and can no longer be accessed by the user. Once in the Purges folder, the message is retained for the duration based on the deleted item retention period configured for the mailbox if single items recovery is enabled for the mailbox. (In Microsoft 365, single item recovery is enabled by default when a new mailbox is created.) After the deleted item retention period expires, the message is marked for permanent deletion and will be purged from Microsoft 365 the next time that the mailbox is processed by the Managed Folder assistant.

- **What if you have to delete a message from more than 50,000 mailboxes?**

  As previously stated, you can perform a search and purge operation on a maximum of 50,000 mailboxes (even if less than 50,000 contain items that match the search query). If you have to do a search and purge operation on more than 50,000 mailboxes, consider creating temporary search permissions filters that reduce the number of mailboxes that would be searched to less than 50,000 mailboxes. For example, if your organization contains mailboxes in different departments, states, or countries/regions, you can create a mailbox search permissions filter based on one of those mailbox properties to search a subset of mailboxes in your organization. After you create the search permissions filter, you would create the search (described in Step 1) and then delete the message (described in Step 3). Then you can edit the filter to search for and purge messages in a different set of mailboxes. For more information about creating search permissions filters, see Configure permissions filtering for Content Search.

- **Will unindexed items included in the search results be deleted?**

  No, the `New-ComplianceSearchAction -Purge command doesn't delete unindexed items.

- **What happens if a message is deleted from a mailbox that has been placed on In-Place Hold or Litigation Hold or is assigned to an Microsoft 365 retention policy?**

  After the message is purged and moved to the Purges folder, the message is retained until the hold duration expires. If the hold duration is unlimited, then items are retained until the hold is removed or the hold duration is changed.

- **Why is the search and remove workflow divided among different Microsoft Purview portal role groups?**

  As previously explained, a person has to be a member of the eDiscovery Manager role group or be assigned the Compliance Search management role to search mailboxes. To

delete messages, a person has to be a member of the Organization Management role group or be assigned the Search And Purge management role. This makes it possible to control who can search mailboxes in the organization and who can delete messages.

Last updated on 12/01/2025

Retry a Content Search to resolve a content location error

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

When you use Content Search in the Microsoft Purview portal to search a large number of mailboxes, you may get search errors that are similar to the error:

```text
Error

The search on the following locations failed:
User1@contoso.com: Problem in processing the request. Please try again later. If you keep getting this error, contact your admin. (CS008-009)
User2@contoso.com: Application error occurred. Please try again later. (CS012-002)
```

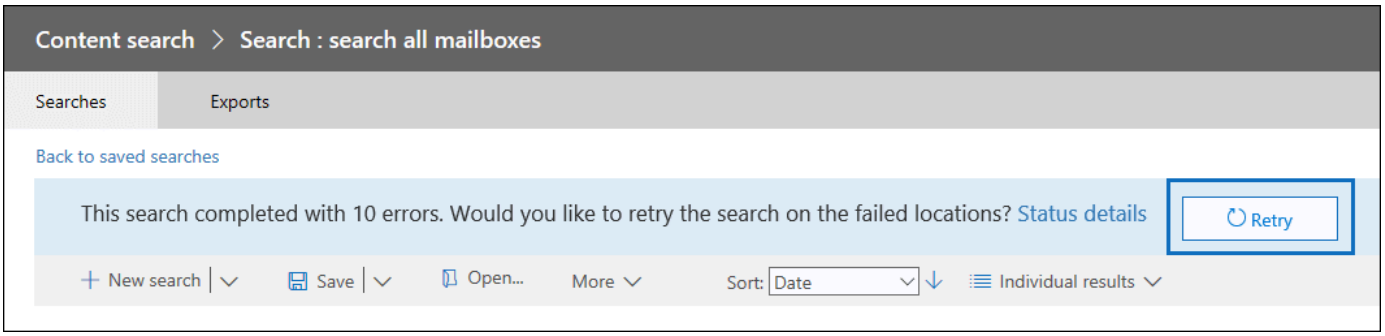
These errors (with error codes of CS001-002, CS003-002, CS008-009, CS012-002, and other errors of the form CS0XX-0XX) indicate that Content Search failed to search specific content locations. In this example, two mailboxes weren't searched. These errors are displayed on the status details flyout page of the Content Search.

Cause of content location errors

When searching a large number of mailboxes, the search is distributed across thousands of servers in a Microsoft datacenter. At any one time, specific servers could be in reboot state or in the process of failing over to redundant copies. In either of these cases, the Content Search's request to retrieve data will time out. In the previous example, the errors for the mailboxes that failed were the result of the search timing out.

Resolving content location errors

Restarting the search will often result in similar errors on different servers. Instead of restarting the search, select **Retry** at the top of the search results page.



This will result in the retrying the search only for the mailboxes that failed. When you retry the search, the other results that were successfully returned are retained.

Tips to avoid content location errors

Here are some additional causes of content location errors and some tips to help you avoid them when searching large numbers of mailboxes.

- The mailbox being searched might be busy due to user activity. In this case, the search service might throttle itself to prevent the mailbox from becoming unavailable. To avoid this, try running searches during non-business hours.
- The search query might be retrieving too much content from the mailbox. If possible, try to narrow the scope of the search by using keywords, date ranges, and search conditions.
- Too many keywords or keyword phrases when you create a search query using the keywords list. When you run a search query that uses the keywords list, the service essentially runs a separate search for each row in the keyword list so that statistics can be generated. If you're using the keywords list in search queries, minimize the number of rows in the keyword list or divide the number keywords into smaller lists and create a different search for each keyword list.

  > **Note**
  >
  > To help reduce issues caused by large keyword lists, you're now limited to a maximum of 20 rows in the keyword list of a search query.

- Too many searches are being performed on the same mailbox at the same time. If possible, try to run one search at a time on any one mailbox.
- Searching too many mailboxes in a single search. The probability of content location errors increases when searching a large number of mailboxes. If possible, try to run multiple searches so that each search includes a subset of mailboxes in your organization.
- Required maintenance is being performed on the mailbox. Though this cause probably occurs infrequently, wait a little while after receiving the content location error and then retry the search.

Last updated on 11/21/2025

Feature reference for Content search

> **⊗ Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

This article describes features and functionality of Content search. To use the Content search eDiscovery tool in the Microsoft Purview portal to search for in-place content, see Get started with Content search.

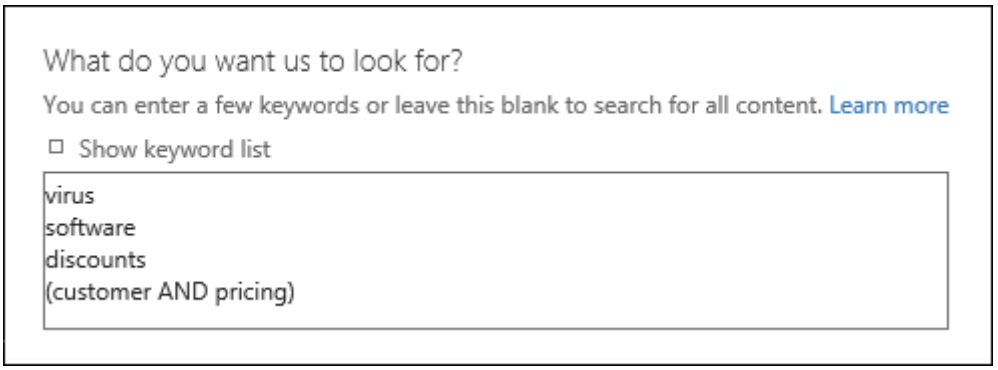
Building a search query

For detailed information about creating a search query, using Boolean search operators and search conditions, and searching for sensitive information types and content shared with users outside your organization, see Keyword queries and search conditions for Content Search.

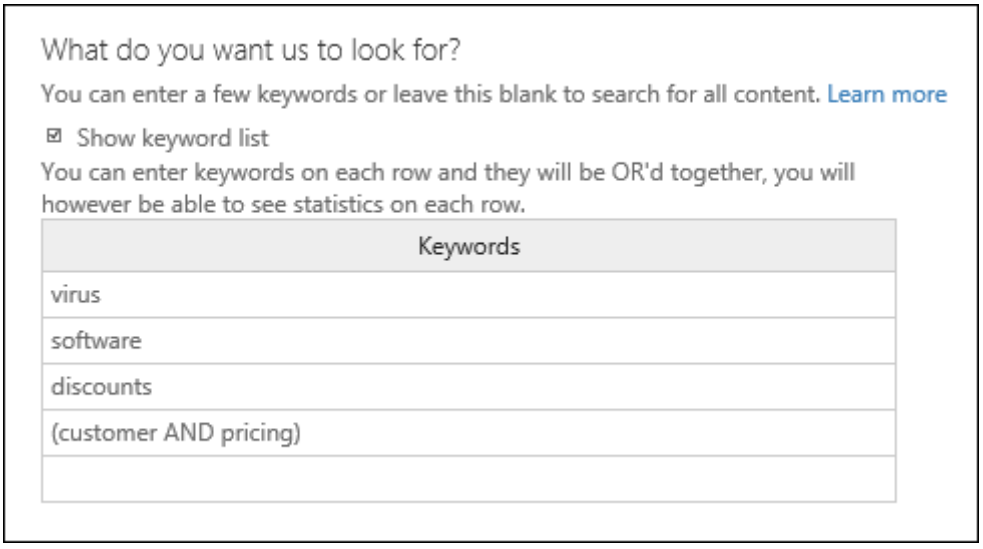
Keep the following things in mind when using the keyword list to create a search query.

- You have to select the **Show keyword list** checkbox and then type each keyword in a separate row to create a search query where the keywords (or keyword phrases) in each row are connected by the **OR** operator. If you paste a list of keywords in the keyword box or press the **Enter** key after typing a keyword, they won't be connected by the **OR** operator. Here are incorrect and correct examples of how to add a list of keywords.

  Incorrect

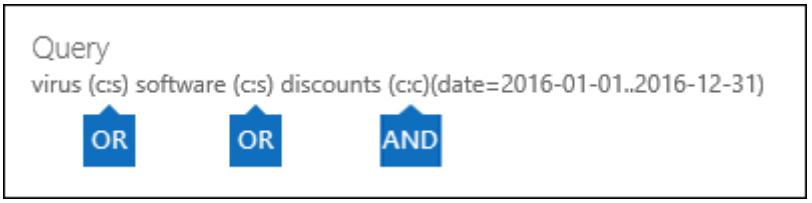


Correct



- You can also prepare a list of keywords or keyword phrases in an Excel file or a plain text file, and then copy and paste your list into the keyword list. To do this, you have to select the **Show keyword list** check box. Then, select the first row in the keyword list and paste your list. Each line from the Excel or text file is pasted into separate row in the keyword list.

- After you create a query using the keyword list, it's a good idea to verify the search query syntax to make the search query is what you intended. In the search query that's displayed under **Query** in the details pane, the keywords are separated by the text **(c:s)**. This indicates that the keywords are connected by a logical operator similar in functionality to the **OR** operator. Similarly, if your search query includes conditions, the keywords and the conditions are separated by the text **(c:c)**. This indicates that the keywords are connected to the conditions with a logical operator similar in functionality to the **AND** operator. Here's an example of the search query (displayed in the Details pane) that results when using the keyword list and a condition.



- When you run a content search, Microsoft 365 automatically checks your search query for unsupported characters and for Boolean operators that may not be capitalized. Unsupported characters are often hidden and typically cause a search error or return unintended results. For more information about the unsupported characters that are checked, see Check your Content Search query for errors.

- If you have a search query that contains keywords for non-English characters (such as Chinese characters), you can select **Query language-country/region** and select a language-country/region culture code value for the search. The default language/region is neutral. How can you tell if you need to change the language setting for a content search? If you're certain content locations contain the non-English characters you're searching for, but the search returns no results, the language setting may be the cause.

Partially indexed items

- Partially indexed items in mailboxes are included in the estimated search results. Partially indexed items from SharePoint and OneDrive aren't included in the estimated search results. For more information, see Partially indexed items in eDiscovery.

Searching OneDrive accounts

- To collect a list of the URLs for the OneDrive sites in your organization, see Create a list of all OneDrive locations in your organization. This script in this article creates a text file that contains a list of all OneDrive sites. To run this script, you have to install and use the SharePoint Online Management Shell. Be sure to append the URL for your organization's MySite domain to each OneDrive site that you want to search. This is the domain that contains all your OneDrive; for example, `https://contoso-my.sharepoint.com`. Here's an example of a URL for a user's OneDrive site: `https://contoso-my.sharepoint.com/personal/sarad_contoso_onmicrosoft.com`.

  In the rare case of a person's user principal name (UPN) being changed, the URL for their OneDrive location is changed to incorporate the new UPN. If this happens, you have to modify a content search by adding the user's new OneDrive URL and removing the old one. For more information, see How UPN changes affect the OneDrive URL.

Searching Microsoft Teams and Microsoft 365 Groups

You can search the mailbox that's associated with a Microsoft Team or Microsoft 365 Group. Because Microsoft Teams is built on Microsoft 365 Groups, searching them is similar. In both

cases, only the group or team mailbox is searched. The mailboxes of the group or team members aren't searched. To search them, you've to specifically add them to the search.

Keep the following things in mind when searching for content in Microsoft Teams and Microsoft 365 Groups.

- To search for content located in Teams and Microsoft 365 Groups, you have to specify the mailbox and SharePoint site that are associated with a team or group.
- Content from private channels is stored in each user's mailbox, not the team mailbox. To search for content in private channels, see eDiscovery of private and shared channels.
- Run the **Get-UnifiedGroup** cmdlet in Exchange Online to view properties for a team or a Microsoft 365 Group. This is a good way to get the URL for the site that's associated with a team or a group. For example, the following command displays selected properties for a Microsoft 365 Group named Senior Leadership Team:

  ```text
  Get-UnifiedGroup "Senior Leadership Team" | FL
  DisplayName,Alias,PrimarySmtpAddress,SharePointSiteUrl
  DisplayName            : Senior Leadership Team
  Alias                  : seniorleadershipteam
  PrimarySmtpAddress     : seniorleadershipteam@contoso.onmicrosoft.com
  SharePointSiteUrl      :
  https://contoso.sharepoint.com/sites/seniorleadershipteam
  ```

  > **Note**
  >
  > To run the **Get-UnifiedGroup** cmdlet, you have to be assigned the View-Only Recipients role in Exchange Online or be a member of a role group that's assigned the View-Only Recipients role.

- When a user's mailbox is searched, any team or Microsoft 365 Group that the user is a member of won't be searched. Similarly, when you search a team or a Microsoft 365 Group, only the group mailbox and group site that you specify is searched. The mailboxes and OneDrive for Business accounts of group members aren't searched unless you explicitly add them to the search.
- To get a list of the members of a team or a Microsoft 365 Group, you can view the properties on the **Home** > **Groups** page in the Microsoft 365 admin center. Alternatively, you can run the following command in Exchange Online PowerShell:

PowerShell

```
Get-UnifiedGroupLinks <group or team name> -LinkType Members | FL
DisplayName,PrimarySmtpAddress
```

> ⓘ Note
>
> To run the **Get-UnifiedGroupLinks** cmdlet, you have to be assigned the View-Only Recipients role in Exchange Online or be a member of a role group that's assigned the View-Only Recipients role.

- Conversations that are part of a Teams channel are stored in the mailbox that's associated with the team. Similarly, files that team members share in a channel are stored on the team's SharePoint site. Therefore, you have to add the team mailbox and SharePoint site as a content location to search conversations and files in a channel.

- Alternatively, conversations that are part of the Chat list in Teams are stored in the Exchange Online mailbox of the users who participate in the chat. And files that a user shares in Chat conversations are stored in the OneDrive for Business account of the user who shares the file. Therefore, you have to add the individual user mailboxes and OneDrive for Business accounts as content locations to search conversations and files in the Chat list.

  > ⓘ Note
  >
  > In an Exchange hybrid deployment, users with an on-premises mailbox might participate in conversations that are part of the Chat list in Teams. In this case, content from these conversations is also searchable because it's saved to a cloud-based storage area (called a *cloud-based mailbox for on-premises users*) for users who have an on-premises mailbox. For more information, see Search for Teams chat data for on-premises users.

- Every team or team channel contains a Wiki for note-taking and collaboration. The Wiki content is automatically saved to a file with a .mht format. This file is stored in the Teams Wiki Data document library on the team's SharePoint site. You can use the Content Search tool to search the Wiki by specifying the team's SharePoint site as the content location to search.

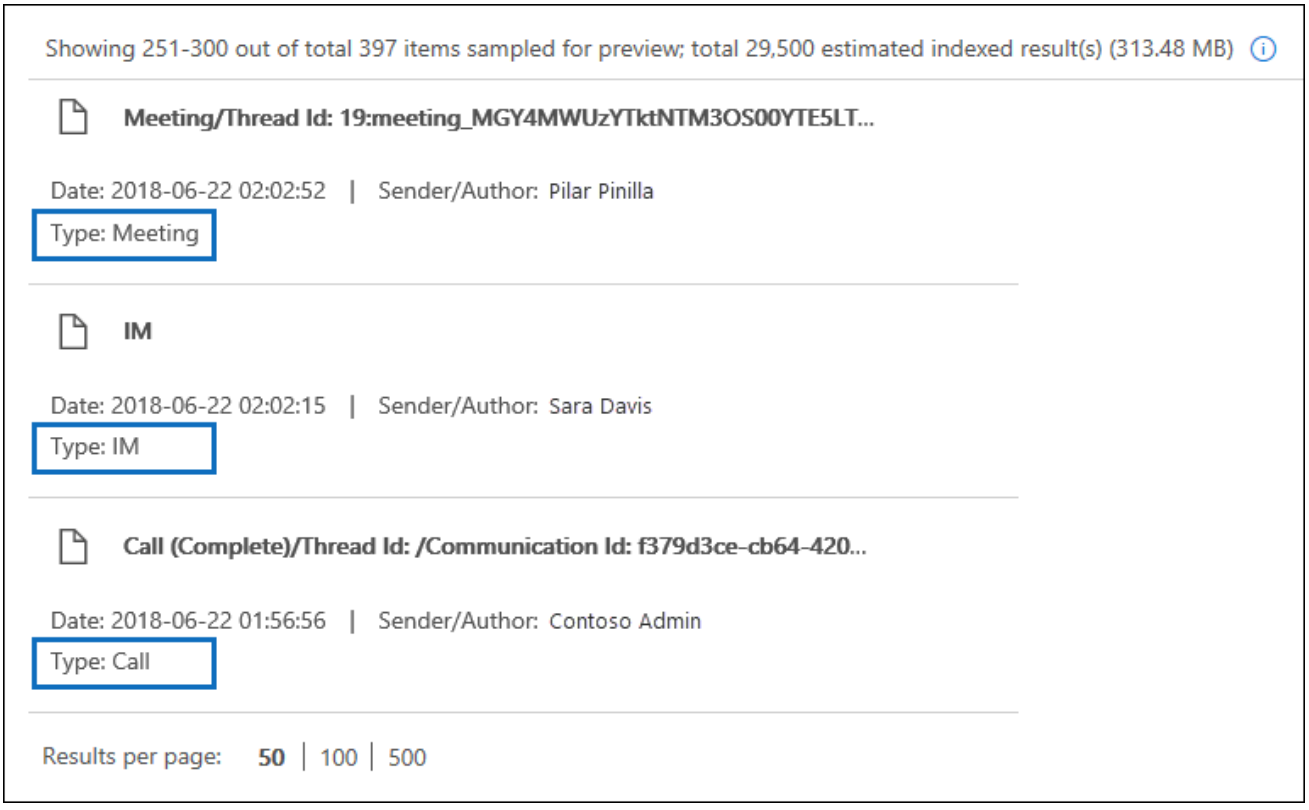
  > ⓘ Note

> The capability to search the Wiki for a team or channel (when you search the team's SharePoint site) was released on June 22, 2017. Wiki pages that were saved or updated on that date or after are available to be searched. Wiki pages last saved or updated before that date aren't available for search.

- Summary information for meetings and calls in a Teams channel are also stored in the mailboxes of users who dialed into the meeting or call. This means you can use Content Search to search these summary records. Summary information includes:
  - Date, start time, end time, and duration of a meeting or call
  - The date and time when each participant joined or left the meeting or call
  - Calls sent to voice mail
  - Missed or unanswered calls
  - Call transfers, which are represented as two separate calls

  It can take up to 8 hours for meeting and call summary records to be available to be searched.

  In the search results, meeting summaries are identified as **Meeting** in the **Type field**, and call summaries are identified as **Call**. Also, conversations that are part of a Teams channel and 1xN chats are identified as **IM** in the **Type** field.



For more information, see Microsoft Teams launches eDiscovery for calls and meetings.

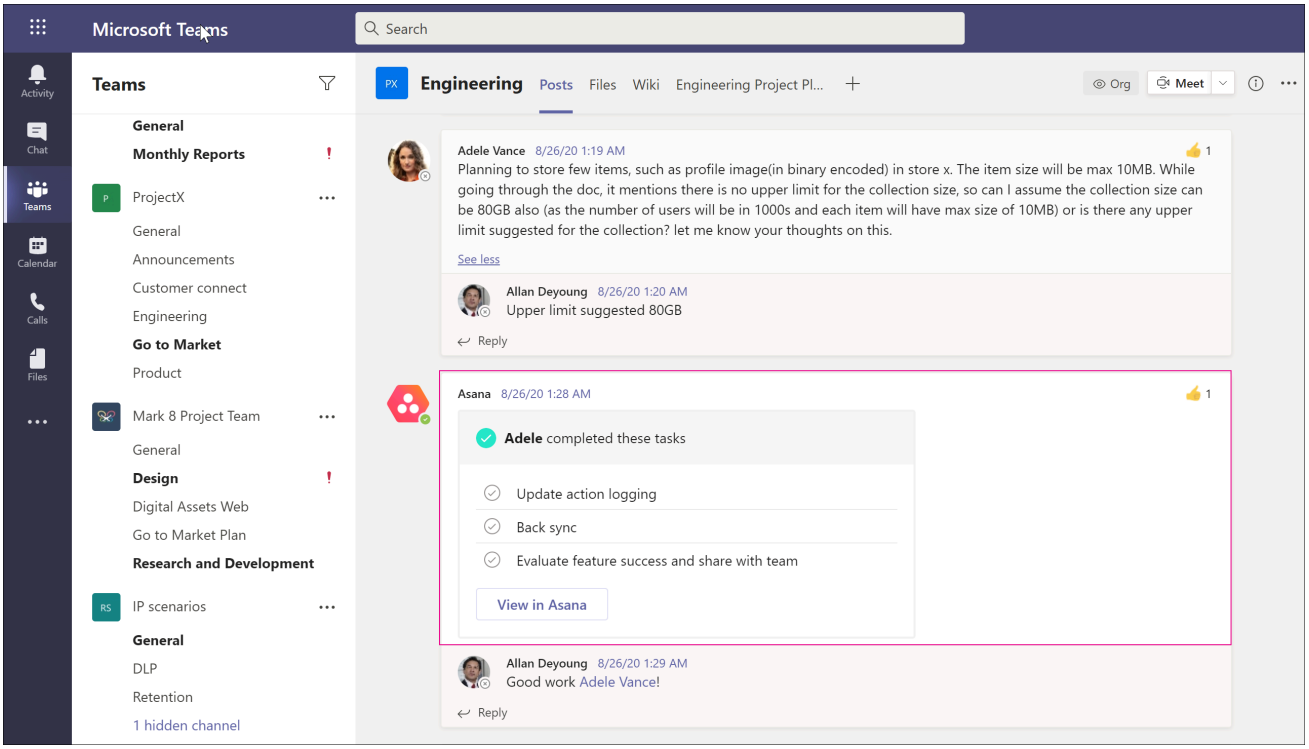
- Card content generated by apps in Teams channels, 1:1 chats, and 1xN chats are stored in mailboxes and can be searched. A *card* is a UI container for short pieces of content. Cards can have multiple properties and attachments, and can include items that trigger card actions. For more information, see Cards

  Like other Teams content, where card content is stored is based on where the card was used. Content for cards used in a Teams channel is stored in the Teams group mailbox. Card content for 1:1 and 1xN chats are stored in the mailboxes of the chat participants.

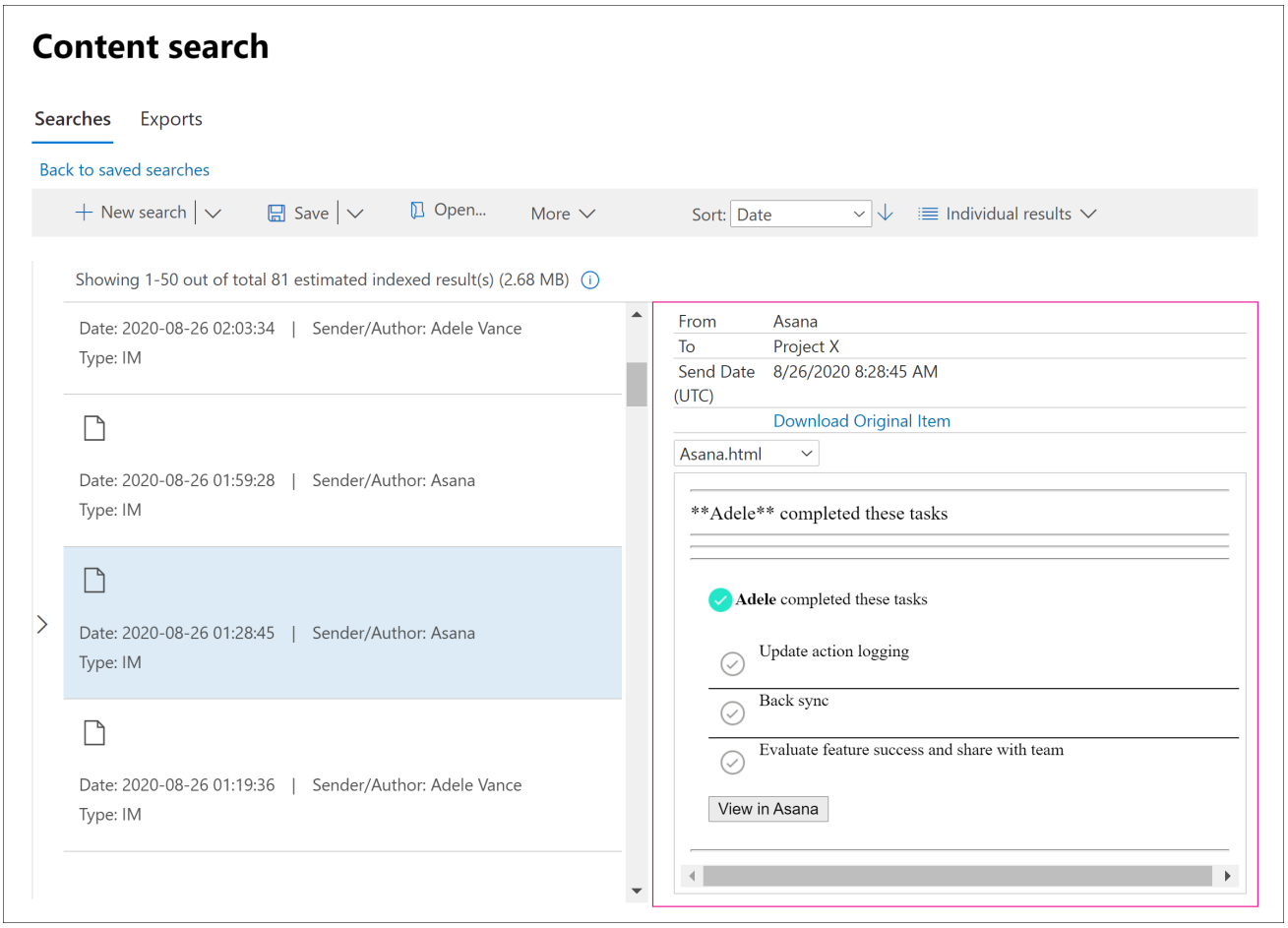
  To search for card content, you can use the `kind:microsoftteams` or `itemclass:IPM.SkypeTeams.Message` search conditions. When reviewing search results, card content generated by bots in a Teams channel has the **Sender/Author** email property as `<appname>@teams.microsoft.com`, where `appname` is the name of the app that generated the card content. If card content was generated by a user, the value of **Sender/Author** identifies the user.

  When viewing card content in Content search results, the content appears as an attachment to the message. The attachment is named `appname.html`, where `appname` is the name of the app that generated the card content. The following screenshots show how card content (for an app named Asana) appears in Teams and in the results of a search.

Card content in Teams

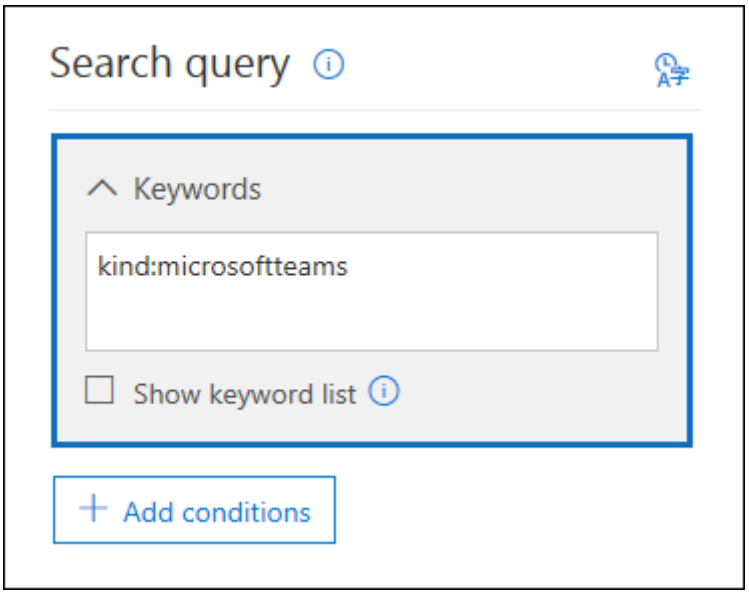


Card content in search results

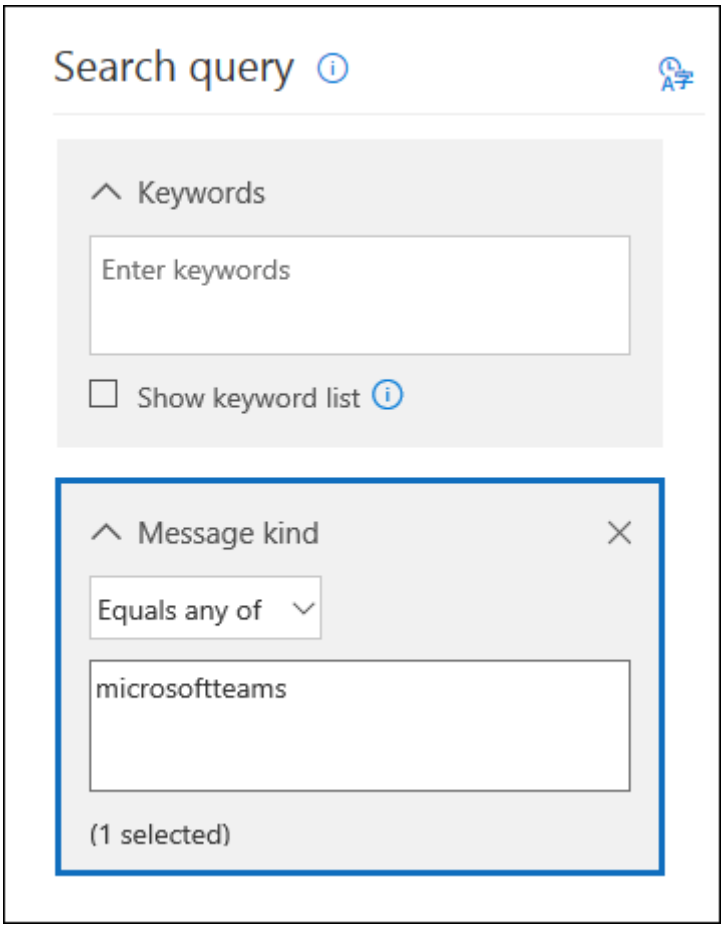


> **Note**
>
> To display images from card content in search results at this time (such as the checkmarks in the previous screenshot), you have to be signed into Teams (at https://teams.microsoft.com) in a different tab in the same browser session that you use to view the search results. Otherwise, image placeholders are displayed.

- You can use the **Kind** email property or the **Message kind** search condition to search specifically for content in Teams.
    - To use the **Kind** property as part of the keyword search query, in the **Keywords** box of a search query, type `kind:microsoftteams`.



- To use a search condition, add the **Message kind** condition and use the value `microsoftteams`.



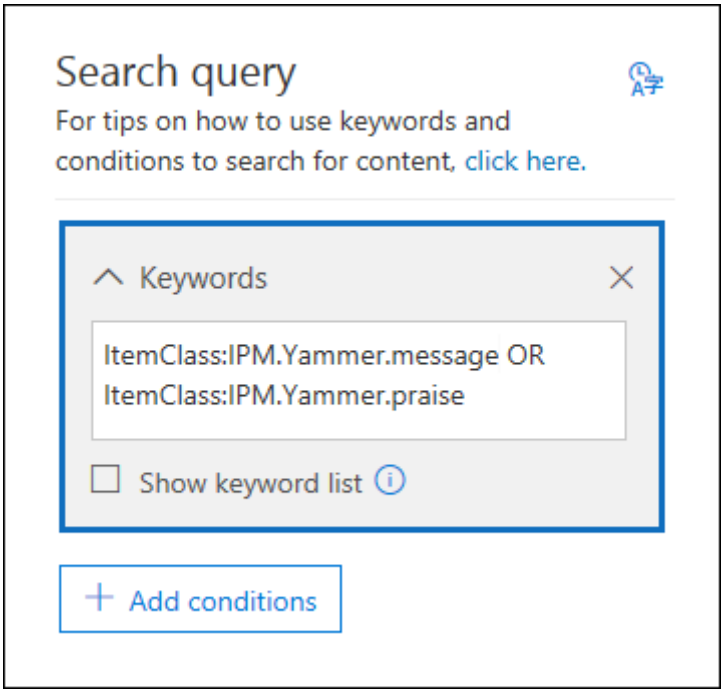
Conditions are logically connected to the keyword query by the **AND** operator. That means an item must match both the keyword query and the search condition to be returned in the search results. For more information, see the "Guidelines for using conditions" section in Keyword queries and search conditions for Content Search.

Searching Viva Engage Groups

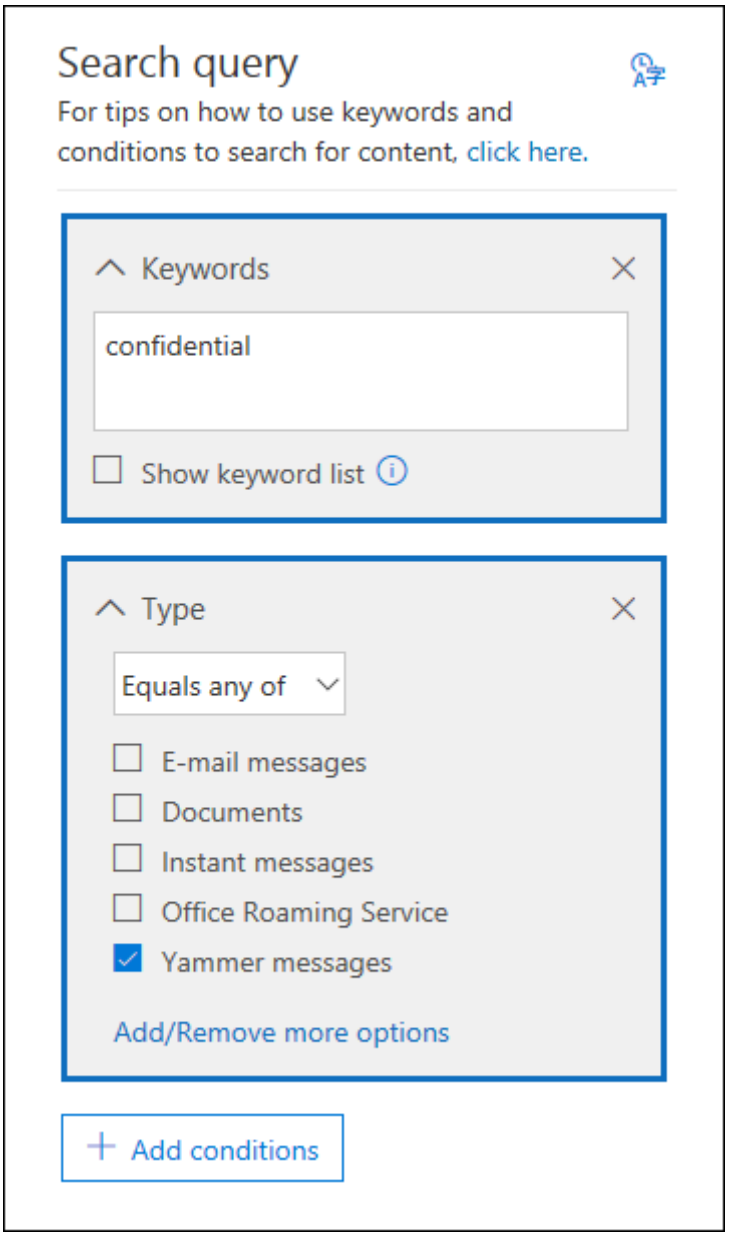
You can use the **ItemClass** email property or the **Type** search condition to search specifically for conversation items in Viva Engage Groups.

- To use the **ItemClass** property as part of the keyword search query, in the **Keywords** box of a search query, you can type one (or all) of the following property:value pairs:
  - ItemClass:IPM.Yammer.message
  - ItemClass:IPM.Yammer.poll
  - ItemClass:IPM.Yammer.praise
  - ItemClass:IPM.Yammer.question

  For example, you can use the following search query to return Viva Engage messages and Viva Engage praise items:



- Alternatively, you can use the **Type** email condition and select **Viva Engage messages** to return Viva Engage items. For example, the following search query returns all Viva Engage conversation items that contain the keyword "confidential".



Searching inactive mailboxes

You can search inactive mailboxes in a content search. To get a list of the inactive mailboxes in your organization, run the command `Get-Mailbox -InactiveMailboxOnly` in Exchange Online PowerShell. Alternatively, you can go to **Data lifecycle management** > **Microsoft 365** > **Retention** in the Microsoft Purview portal, and then select **More**••• > **Inactive mailboxes**.

Here are a few things to keep in mind when searching inactive mailboxes.

- If an existing content search includes a user mailbox and that mailbox is made inactive, the content search will continue to search the inactive mailbox when you rerun the search after it becomes inactive.
- Sometimes a user may have an active mailbox and an inactive mailbox that have different SMTP addresses. In this case, only the specific mailbox that you select as a location for a

content search is searched. In other words, if you add a user's mailbox to a search, you can't assume that both their active and inactive mailboxes are searched. Only the mailbox that you explicitly add to the search is searched.

- You can use Security & Compliance PowerShell to create a content search to search an inactive mailbox. To do this, you have to pre-append a period (.) to the email address of the inactive mailbox. For example, the following command creates a content search that searches an inactive mailbox with the email address pavelb@contoso.onmicrosoft.com:

  PowerShell

  ```
  New-ComplianceSearch -Name InactiveMailboxSearch -ExchangeLocation
  .pavelb@contoso.onmicrosoft.com -AllowNotFoundExchangeLocationsEnabled $true
  ```

- We strongly recommend that you avoid having an active mailbox and inactive mailbox with the same SMTP address. If you need to reuse the SMTP address that is assigned to an inactive mailbox, we recommend that you recover the inactive mailbox or restore the contents of an inactive mailbox to an active mailbox (or the archive of an active mailbox), and then delete the inactive mailbox. For more information, see one of the following articles:
  - Recover an inactive mailbox in Office 365
  - Restore an inactive mailbox in Office 365
  - Delete an inactive mailbox in Office 365

Searching disconnected or de-licensed mailboxes

If the Exchange Online license (or the entire Microsoft 365 license) is removed from a user account or in Microsoft Entra ID, the user's mailbox becomes a *disconnected* mailbox. This means that the mailbox is no longer associated with the user account. Here's what happens when searching disconnected mailboxes:

- If the license is removed from a mailbox, the mailbox is no longer searchable.
- If an existing content search includes a mailbox in which the license is removed, no search results from the disconnected mailbox will be returned if you rerun the content search.
- If you use the **New-ComplianceSearch** cmdlet to create a content search and specify a disconnected mailbox as the Exchange content location to search, the content search won't return any search results from the disconnected mailbox.

If you need to preserve the data in a disconnected mailbox so that it's searchable, you must place a hold on the mailbox before removing the license. This preserves the data and keeps the

disconnected mailbox searchable until the hold is removed. For more information about holds, see How to identify the type of hold placed on an Exchange Online mailbox.

Searching for content in a SharePoint multi-geo environment

If it's necessary for an eDiscovery manager to search for content in SharePoint and OneDrive in different regions in a SharePoint multi-geo environment, then you need to do the following things to make that happen:

1. Create a separate user account for each satellite geo location that the eDiscovery manager needs to search. To search for content in sites in that geo location, the eDiscovery manager must sign in to the account you created for that location, and then run a content search.

2. Create a search permissions filter for each satellite geo location (and corresponding user account) the eDiscovery manager needs to search. Each of these search permissions filters limits the scope of the content search to a specific geo location when the eDiscovery manager is signed in to the user account associated with that location.

> **Tip**
>
> You don't have to use this strategy when using the search tool in **eDiscovery (Premium)**. That's because all datacenters are searched when you search SharePoint sites and OneDrive accounts in Microsoft Purview eDiscovery (Premium). You have to use this strategy of region-specific user accounts and search permissions filters only when using the Content Search tool and running searches associated with **eDiscovery cases**.

For example, let's say that an eDiscovery manager needs to search for SharePoint and OneDrive content in satellite locations in North American, Europe, and Asia Pacific. The first step is to create three users accounts, one for each location. The next step is to create three search permissions filters, one for each location *and* corresponding user account. Here are examples of the three search permissions filters for this scenario. In each of these examples, the **Region** specifies the SharePoint datacenter location for that geo and the **Users** parameter specifies the corresponding user account.

North America

PowerShell

```
New-ComplianceSecurityFilter -FilterName "SPMultiGeo-NAM" -Users ediscovery-
nam@contoso.com -Region NAM -Action ALL
```

Europe

PowerShell

```
New-ComplianceSecurityFilter -FilterName "SPMultiGeo-EUR" -Users ediscovery-
eur@contoso.com -Region EUR -Action ALL
```

Asia Pacific

PowerShell

```
New-ComplianceSecurityFilter -FilterName "SPMultiGeo-APC" -Users ediscovery-
apc@contoso.com -Region APC -Action ALL
```

Keep the following things in mind when using search permissions filters to search for content in multi-geo environments:

- The **Region** parameter directs searches to the specified satellite location. If an eDiscovery manager only searches SharePoint and OneDrive sites outside of the region specified in the search permissions filter, no search results are returned.
- The **Region** parameter doesn't control searches of Exchange mailboxes. All datacenters are searched when you search mailboxes.

For more information about using search permissions filters in a multi-geo environment, see the "Searching and exporting content in Multi-Geo environments" section in Set up compliance boundaries for eDiscovery investigations.

Need help with eDiscovery issues?

Check out the Resolve common eDiscovery issues article for basic troubleshooting steps that you can take to identify and resolve issues that you might encounter during an eDiscovery search or elsewhere in the eDiscovery process.

Last updated on 10/31/2024

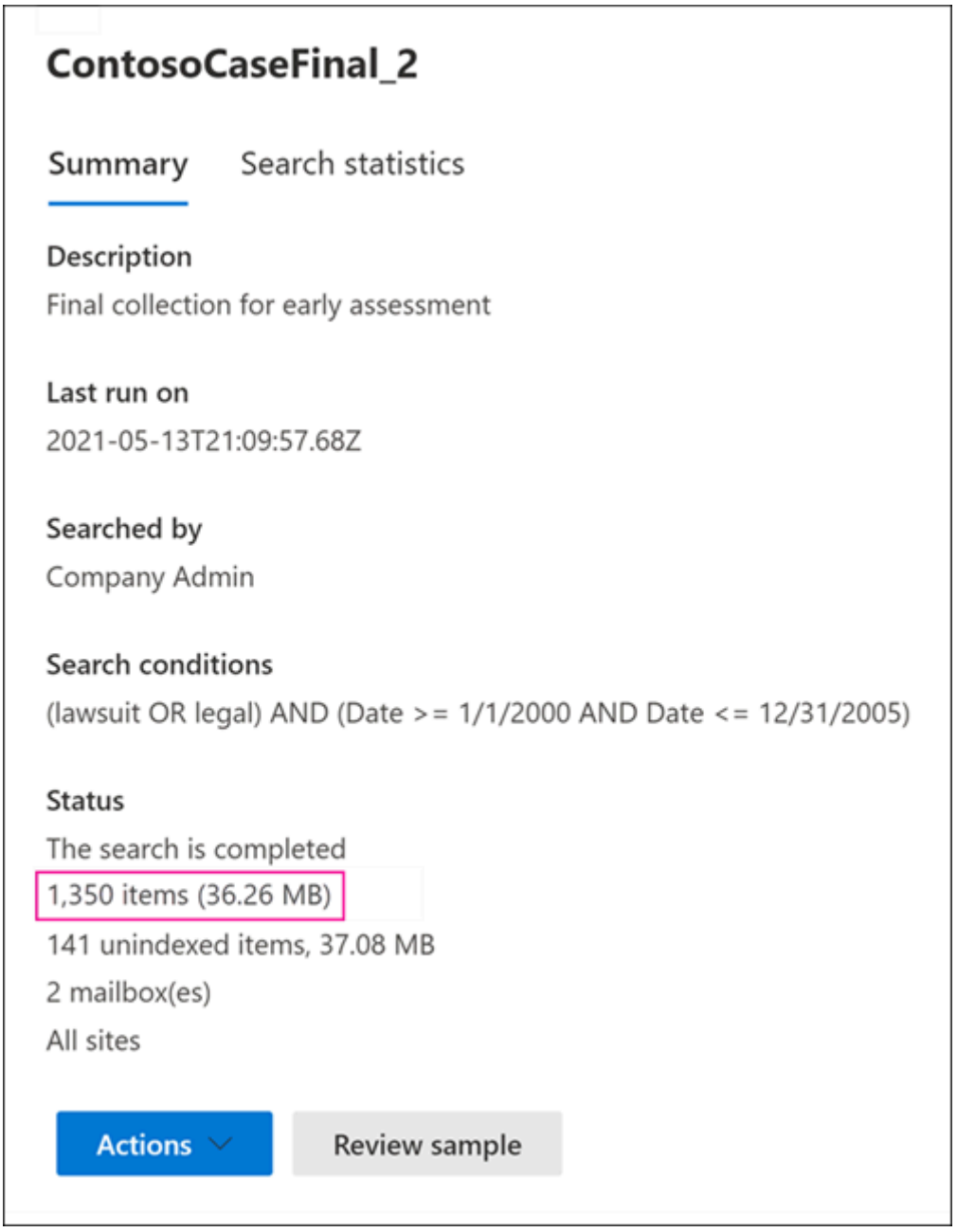
Differences between estimated and actual eDiscovery search results

> **⊗ Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic **Content Search**, classic **eDiscovery (Standard)**, and classic **eDiscovery (Premium)**.
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the **new eDiscovery experience** in the **Microsoft Purview portal**.

This article applies to searches that you can run using one of the following Microsoft Purview eDiscovery tools:

- Content search
- eDiscovery (Standard)

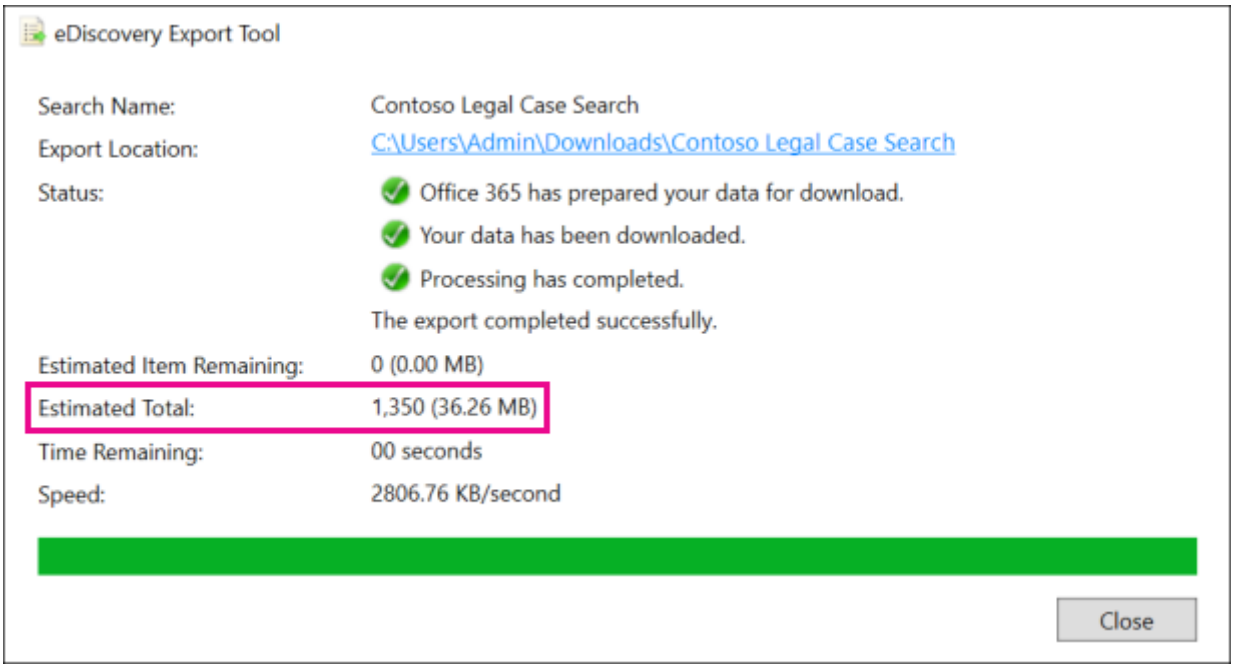
When you run an eDiscovery search, the tool you're using will return an estimate of the number of items (and their total size) that match the search criteria. For example, when you run a search in the Microsoft Purview portal, the estimated search results are displayed on the flyout page for the selected search.



This is the same estimate of total size and number of items that is displayed in the eDiscovery Export Tool when you export results to a local computer and in the Export Summary report that's downloaded with the search results.

> **Important**
>
> The eDiscovery Export tool was permanently **retired on August 31, 2025**, this functionality isn't available in the new eDiscovery experience.

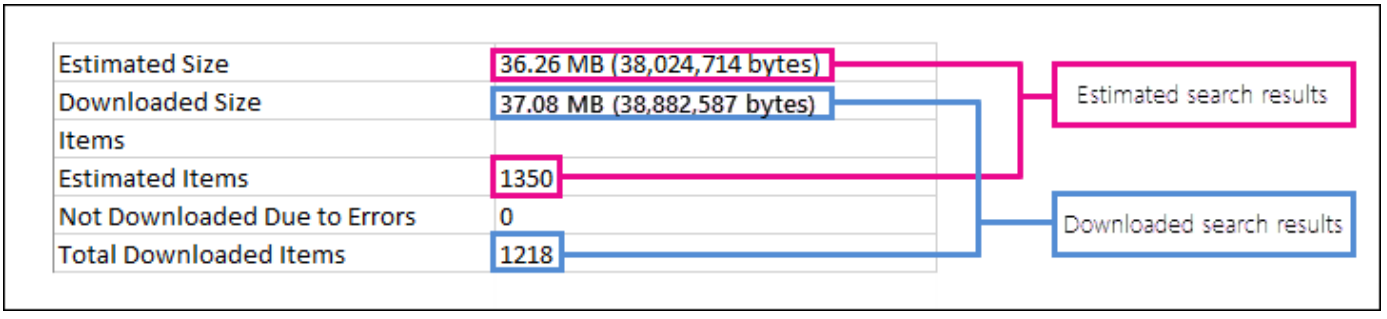
Estimated results in the eDiscovery Export tool



Estimated results in Export Summary report

1	Application launched from	
2	Export Name	Contoso Legal Case Search
3	Download Started	1/8/2016 14:43
4	Download Completed	1/8/2016 14:46
5	Status	The export completed successfully.
6	Locations	
7	Total Locations	22
8	Succeeded Locations	22
9	Failed Locations	0
10	Size	
11	Estimated Size	36.26 MB (38,024,714 bytes)
12	Downloaded Size	37.08 MB (38,882,587 bytes)
13	Items	
14	Estimated Items	1350
15	Not Downloaded Due to Errors	0
16	Total Downloaded Items	1218

However, as you'll notice in the previous screenshot of the Export Summary report, the size and number of actual search results that are downloaded are different than the size and number of estimated search results.



Here are some reasons for these differences:

- **The way results are estimated**. An estimate of the search results is just that, an estimate (and not an actual count) of the items that meet the search query criteria. To compile the estimate of Exchange items, a list of the message IDs that meet the search criteria is requested from the Exchange database by the eDiscovery tool you're using. But when you export the search results, the search is rerun and the actual messages are retrieved from the Exchange database. So these differences might result because of how the estimated number of items and the actual number of items are determined.

- **Changes that happen between the time when estimating and exporting search results**. When you export search results, the search is restarted to collect that most recent items in the search index that meet the search criteria. It's possible there are additional items were created, sent, or received that meet the search criteria in the time between when the estimated search results were collected and when the search results were exported. It's also possible that items that were in the search index when the search results were estimated are no longer there because they were purged from the content location before the search results are exported. One way to mitigate this issue is to specify a date range for an eDiscovery search. Another way is to place a hold on content locations so that items are preserved and can't be purged.

  Here are other issues that can result is differences between estimated and exported search results:
  - In increase in items when using a date query. This is typically cause by the following two things:
    - Hold versioning in SharePoint. If a document is deleted from a site that's on hold and document versioning is enabled, all versions of the deleted document will be preserved.
    - Calendar items. Accept and reject messages and recurring meetings will automatically continue creating new items in the background with old dates.
  - With holds, there can be cases where the same item is preserved in a user's primary mailbox and in their archive mailbox. This can happen when a user manually moves an item to their archive.
  - Although rare, even in the case when a hold is applied, maintenance of built-in calendar items (which aren't editable by the user, but are included in many search results) may be removed from time to time. This periodic removal of calendar items will result in fewer items that are exported.

- **Unindexed items**. Items that are unindexed for search can cause differences between estimated and actual search results. You can include unindexed items when you export the

search results. If you include unindexed items when exporting search results, there might be more items that are exported. This will cause a difference between the estimated and exported search results.

When using the Content search tool, you have the option to include unindexed items when you export search results. The number of unindexed items returned by the search is listed on the flyout page together with the other estimated search results. Any unindexed items would also be included in the total size of the estimated search results. When you export search results, you have the option to include or not include unindexed items. How you configure these options might result in differences between estimated and the actual search results that are downloaded.

- **Exporting the results of a Content search that includes all content locations**. If the search that you're exporting results from was a search of all content locations in your organization, then only the unindexed items from content locations that contain items that match the search criteria will be exported. In other words, if no search results are found in a mailbox or site, then any unindexed items in that mailbox or site won't be exported. However, unindexed items from all content locations (even those that don't contain items that match the search query) will be included in the estimated search results.

  Alternatively, if the search that you're exporting results from included specific content locations, then unindexed items (that aren't excluded by the search criteria) from all the content locations specified in the search will be exported. In this case, the estimated number of unindexed items and the number of unindexed items that are exported should be the same.

  The reason for not exporting unindexed items from every location in the organization is because it might increase the likelihood of export errors and increase the time it takes to export and download the search results.

- **Unindexed items in SharePoint and OneDrive not included in the search estimates**. Unindexed items from SharePoint sites and OneDrive for Business accounts aren't included in the estimated search results. This is because the SharePoint index doesn't contain data for unindexed items. Only unindexed items from mailboxes are included in the search estimates. However, if you include unindexed items when exporting search results, unindexed items in SharePoint and OneDrive are included, which will increase the number of items that are actually exported. This will result in differences between the estimated results (which don't include unindexed items in SharePoint and OneDrive sites) and the actual items that are downloaded. The rule about exporting unindexed items only from

content locations that contain items that match the search criteria still applies in this situation.

- **Document versions in SharePoint and OneDrive**. When searching SharePoint sites and OneDrive accounts, multiple versions of a document aren't included in the count of estimated search results. But you have the option to include all document versions when you export the search results. If you include document versions when exporting search results, the actual number (and total size) of the exported items will be increased.

- **SharePoint folders**. If folders in SharePoint match a search query, for example, searching by date, the search estimate will include a count of those folders with the last modified date range (but not the items in those folders). When you export the search results, the items in folder are exported but the actual folder isn't exported. The result is that the number of exported items will be more than the number of estimated search results. If a folder is empty, then the number of actual search results exported will be reduced by one item, because the actual folder isn't exported.

  > ⓘ **Note**
  >
  > When running a query-based search, you can exclude SharePoint folders by adding the following condition to the query: `NOT(ContentType:folder)`.

- **SharePoint lists**. If the name of a SharePoint list matches a search query, the search estimate will include a count of all the items in the list. When you export the search results, the list (and the list items) is exported as a single CSV file. This will reduce the actual number of items actually exported. If the list contains attachments, the attachments will be exported as separate documents, which will also increase the number of items exported.

  > ⓘ **Note**
  >
  > When running a query-based search, you can exclude SharePoint lists by adding the following condition to the query: `NOT(ContentType:list)`.

- **Raw file formats versus exported file formats**. For Exchange items, the estimated size of the search results is calculated by using the raw Exchange message sizes. However, email messages are exported in a PST file or as individual messages (which are formatted as EML files). Both of these export options use a different file format than raw Exchange messages, which results in the total exported file size being different than the estimated file size.

- **De-duplication of Exchange items during export**. For Exchange items, de-duplication reduces the number of items that are exported. You have the option to de-duplicate the search results when you export them. For Exchange messages, this means that only a single instance of a message is exported, even though that message might be found in multiple mailboxes. The estimated search results include every instance of a message. So if you choose the de-duplication option when exporting search results, the actual number of items that are exported might be considerably less than the estimated number of items.

The search results report (Results.csv file) contains an entry for each duplicate message and identifies the source mailbox where a duplicate message is located. This helps you identify all mailboxes that contain a duplicate message.

> **Note**
>
> If you don't select the **Include items that are encrypted or have an unrecognized format** option when you export search results or just download the reports, the index error reports are downloaded but they don't have any entries. This doesn't mean there aren't any indexing errors. It just means that unindexed items weren't included in the export.

Last updated on 11/21/2025

Partially indexed items in Content search

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

A Content search that you run from the Microsoft Purview portal automatically includes partially indexed items in the estimated search results when you run a search. Partially indexed items are Exchange mailbox items and documents on SharePoint and OneDrive for Business sites that for some reason weren't completely indexed for search. In Exchange, a partially indexed item typically contains a file (of a file type that can't be indexed) that is attached to an email message. Here are some other reasons why items can't be indexed for search and are returned as partially indexed items when you run an eDiscovery search:

- The file type is unrecognized or unsupported for indexing.
- Messages have an attached file that can't be opened; this is the most common cause of partially indexed email items.
- The file type is supported for indexing but an indexing error occurred for a specific file.
- Too many files attached to an email message.
- A file attached to an email message is too large.
- A file is encrypted with non-Microsoft technologies.
- A file is password-protected.

> **Note**
>
> Most organizations have less than 1% of content by volume and less than 12% by size that is partially indexed. The reason for the difference between volume and size is that larger files have a higher probability of containing content that can't be completely indexed.

For legal investigations, your organization may be required to review partially indexed items. You can also specify whether to include partially indexed items when you export search results to a local computer or when you prepare the results for analysis with eDiscovery (Premium). For more information, see Investigating partially indexed items in eDiscovery.

File types not indexed for search

Certain types of files, such as Bitmap (.bmp) or MP3 (.mp3) files, don't contain content that can be indexed. As a result, the search indexing servers in Exchange and SharePoint don't perform full-text indexing on these types of files. These types of files are considered to be unsupported file types. There are also file types for which full-text indexing has been disabled, either by default or by an administrator. Unsupported and disabled file types are labeled as unindexed items in Content Searches. As previously stated, partially indexed items can be included in the set of search results when you run a search, export the search results to a local computer, or prepare search results for eDiscovery (Premium).

For a list of supported and disabled file formats, see the following articles:

- **Exchange** - File formats indexed by Exchange Search
- **Exchange** - Get-SearchDocumentFormat
- **SharePoint** - Default crawled file name extensions and parsed file types in SharePoint

Messages and documents with partially indexed file types can be returned in search results

Not every email message with a partially indexed file attachment or every partially indexed SharePoint document is automatically returned as a partially indexed item. That's because other message or document properties, such as the **Subject** property in email messages and the **Title** or **Author** properties for documents are indexed and available to be searched. For example, a keyword search for "financial" will return items with a partially indexed file attachment if that keyword appears in the subject of an email message or in the file name or title of a document. However, if the keyword appears only in the body of the file, the message or document would be returned as a partially indexed item.

Similarly, messages with partially indexed file attachments and documents of a partially indexed file type are included in search results when other message or document properties, which are indexed and searchable, match the search criteria. Message properties that are indexed for search include sent and received dates, sender and recipient, the file name of an attachment, and text in the message body. Document properties indexed for search include created and modified dates.

So even though a message attachment may be a partially indexed item, the message will be included in the regular search results if the value of other message or document properties matches the search criteria.

For a list of email and document properties that you can search for by using eDiscovery tools in the Microsoft Purview portal, see Keyword queries and search conditions for eDiscovery.

> **Note**
>
> If a mailbox item is moved from a folder that is indexed to a folder that is not indexed, a flag is set to unindex the item and the item is removed from the index and will not be searchable. Later, if that same item is moved back to a folder that is indexed, the flag is not reset. That means the item will remain unindexed, and not searchable.

Partially indexed items included in the search results

Your organization might be required to identify and perform additional analysis on partially indexed items to determine what they are, what they contain, and whether they're relevant to a specific investigation. As previously explained, the partially indexed items in the content locations that are searched are automatically included with the estimated search results. You have the option to include these partially indexed items when you export search results or prepare the search results for eDiscovery (Premium).

Keep the following in mind about partially indexed items:

- When you run an eDiscovery search, the total number and size of partially indexed Exchange items (returned by the search query) are displayed in the search statistics on the flyout page, and labeled as **unindexed items**. Statistics about partially indexed items displayed on the flyout page don't include partially indexed items in SharePoint sites or OneDrive accounts.

- If the search that you're exporting results from was a search of specific content locations or all content locations in your organization, only the unindexed items from content locations that contain items that match the search criteria will be exported. In other words, if no search results are found in a mailbox or site, then any unindexed items in that mailbox or site won't be exported. The reason for this is that exporting partially indexed items from lots of locations in the organization might increase the likelihood of export errors and increase the time it takes to export and download the search results.

To export partially indexed items from all content locations for a search, configure the search to return all items (by removing any keywords from the search query) and then export only partially indexed items when you export the search results (by selecting **Only items that have an unrecognized format, are encrypted, or weren't indexed for other reasons** under **Output options**).

- If you choose to include all mailbox items in the search results, or if a search query doesn't specify any keywords or only specifies a date range, partially indexed items might not be copied to the PST file that contains the partially indexed items. This is because all items, including any partially indexed items, will be automatically included in the regular search results.

- Partially indexed items aren't available to be previewed. You have to export the search results to view partially indexed items returned by the search.

  Additionally, when you export search results and include partially indexed items in the export, partially indexed items from SharePoint items are exported to a folder named **Uncrawlable**. When you export partially indexed Exchange items, they're exported differently depending if the partially indexed items matched the search query and the configuration of the export settings.

- The following table shows the export behavior of indexed and partially indexed items and whether or not each is included for the different export configuration settings.

Expand table

Export configuration	Indexed items that match search query	Partially indexed items that match search query	Partially indexed items that don't match search query
Export only indexed items	Exported	Exported (included with the indexed items exported)	Not exported
Export only partially indexed items	Not exported	Exported (as partially indexed items)	Exported (as partially indexed items)
Export indexed and partially indexed items	Exported	Exported (included with the indexed items exported)	Exported (as partially indexed items)

Date ranges and excluding partially indexed items

In Content search and Microsoft Purview eDiscovery, you can't use a date range to exclude partially indexed items from being returned by a search query. In other words, partially indexed items that fall outside of a date range are still included as partially indexed items in the search statistics and when you export partially indexed items. In eDiscovery (Premium), partially indexed items can be collected and then filtered in a review set prior to export.

Indexing limits for messages

The following table describes the indexing limits that might result in an email message being returned as a partially indexed item in an eDiscovery search in Microsoft 365.

For a list of indexing limits for SharePoint documents, see Search limits for SharePoint Online.

Expand table

Indexing limit	Maximum value	Description
Maximum attachment size (excluding Excel files)	150 MB	The maximum size of an email attachment that will parse for indexing. Any attachment that's larger than this limit won't be parsed for indexing, and the message with the attachment will be marked as partially indexed. **Note:** Parsing is the process where the indexing service extracts text from the attachment, removes unnecessary characters like punctuation and spaces, and then divides the text into words (in a process called tokenization), that are then stored in the index.
Maximum size of Excel files	4 MB	The maximum size of an Excel file located on a site or attached to an email message that will be parsed for indexing. Any Excel file that's larger than this limit won't be parsed, and the file or the email the message with the file attachment will be marked as unindexed.
Maximum number of attachments	250	The maximum number of files attached to an email message that will be parsed for indexing. If a message has more than 250 attachments, the first 250 attachments are parsed and indexed, and the message is marked as partially indexed because it had additional attachments that weren't parsed.
Maximum attachment depth	30	The maximum number of nested attachments that are parsed. For example, if an email message has another message attached to it and the attached message has an attached Word document, the Word document and the attached message will be indexed. This behavior will continue for up to 30 nested attachments.
Maximum number of attached images	0	An image that's attached to an email message is skipped by the parser and isn't indexed.

Indexing limit	Maximum value	Description
Maximum time spent parsing an item	30 seconds	A maximum of 30 seconds is spent parsing an item for indexing. If the parsing time exceeds 30 seconds, the item is marked as partially indexed.
Maximum parser output	2 million characters	The maximum amount of text output from the parser that's indexed. For example, if the parser extracted 8 million characters from a document, only the first 2 million characters are indexed.
Maximum annotation tokens	2 million	When an email message is indexed, each word is annotated with different processing instructions that specify how that word should be indexed. Each set of processing instructions is called an annotation token. To maintain the quality of service in Office 365, there's a limit of 2 million annotation tokens for an email message.
Maximum body size in index	67 million characters	The total number of characters in the body of an email message and all its attachments. When an email message is indexed, all text in the body of the message and in all attachments is concatenated into a single string. The maximum size of this string that is indexed is 67 million characters.
Maximum unique tokens in body	1 million	As previously explained, tokens are the result of extracting text from content, removing punctuation and spaces, and then dividing it into words (called tokens) that are stored in the index. For example, the phrase `"cat, mouse, bird, dog, dog"` contains 5 tokens. But only 4 of these are unique tokens. There's a limit of 1 million unique tokens per email message, which helps prevent the index from getting too large with random tokens.

More information about partially indexed items

- As previously stated, because message and document properties and their metadata are indexed, a keyword search might return results if that keyword appears in the indexed metadata. However, that same keyword search might not return the same item if the keyword only appears in the content of an item with an unsupported file type. In this case, the item would be returned as a partially indexed item.
- If a partially indexed item is included in the search results because it matched the search query criteria, it isn't included with partially indexed items when you export search results.
- Although a file type is supported for indexing and is indexed, there can be indexing or search errors that will cause a file to be returned as a partially indexed item. For example, searching a large Excel file might be partially successful (because the first 4 MB are indexed), but then fails because the file size limit is exceeded. In this case, it's possible that the same file is returned with the search results and as a partially indexed item.

- Files that are encrypted with Microsoft encryption technologies and are attached to an email message that matches the criteria of a search can be previewed and will be decrypted when exported. At this time, files that are encrypted with Microsoft encryption technologies (and stored in SharePoint or OneDrive for Business) are partially indexed.
- Email messages encrypted with S/MIME are partially indexed. This includes encrypted messages with or without file attachments.
- Email messages protected using Azure Rights Management are indexed and will be included in the search results if they match the search query. Rights-protected email messages are decrypted and can be previewed and exported. This functionality requires that you're assigned the RMS Decrypt role, which is assigned by default to the eDiscover Manager role group.
- If you create a query-based hold that's associated with an eDiscovery case, all partially indexed items are placed on hold. This includes partially indexed items that don't match the search query criteria for the hold. For more information about creating query-based eDiscovery holds, see Create an eDiscovery hold.

Last updated on 10/31/2024

Investigating partially indexed items in eDiscovery

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

An eDiscovery search that you run from the Microsoft Purview portal automatically includes partially indexed items in the estimated search results when you run a search. Partially indexed items are Exchange mailbox items and documents on SharePoint and OneDrive for Business sites that for some reason weren't completely indexed for search. Most email messages and site documents are successfully indexed because they fall within the indexing limits for email messages. However, some items may exceed these indexing limits, and will be partially indexed. Here are other reasons why items can't be indexed for search and are returned as partially indexed items when you run an eDiscovery search:

- Email messages have an attached file that can't be opened; this is the most common cause of partially indexed email items.
- Too many files attached to an email message.
- A file attached to an email message is too large.
- The file type is supported for indexing but an indexing error occurred for a specific file.

Although it varies, most organizations customers have less than 1% of content by volume and less than 12% of content by size that is partially indexed. The reason for the difference between the volume versus size is that larger files have a higher probability of containing content that can't be completely indexed.

For more information about partially index items in Content search, see Investigating partially indexed items in Content search.

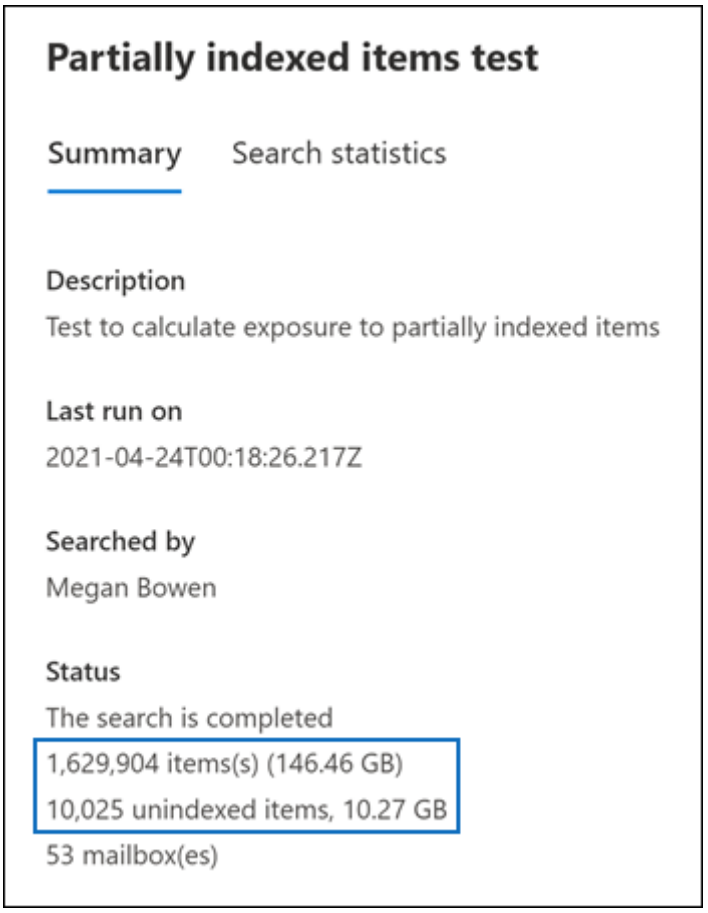
Why does the partially indexed item count change for a search?

After you run an eDiscovery search, the total number and size of partially indexed items in the locations that were searched are listed in the search result statistics that are displayed in the detailed statistics for the search. Note these are called *unindexed items* in the search statistics. Here are a few things that will affect the number of partially indexed items that are returned in the search results:

- If an item is partially indexed and matches the search query, it's included in both the count (and size) of search result items and partially indexed items. However, when the results of that same search are exported, the item is included only with set of search results; it's not included as a partially indexed item.
- Partially indexed items located in SharePoint and OneDrive sites *aren't* included in the estimate of partially indexed items that's displayed in the detailed statistics for the search. However, partially indexed items can be exported when you export the results of an eDiscovery search. For example, if you only search sites, the estimated number partially indexed items will be zero.

Calculating the ratio of partially indexed items in your organization

To understand your organization's exposure to partially indexed items, you can run a search for all content in all mailboxes (by using a blank keyword query). In the following example, there are 1,629,904 (146.46 GB) fully indexed items and 10,025 (10.27 GB) partially indexed items.



You can determine the percentage of partially indexed items by using the following calculations.

To calculate the ratio of partially indexed items in your organization:

```
(Total number of partially indexed items/Total number of items) x 100
```

```
(10025/1629904) x 100 = 0.62%
```

By using the search results from the previous example, 0.62% of all mailboxes items are partially indexed.

To calculate the percentage of the size of partially indexed items in your organization:

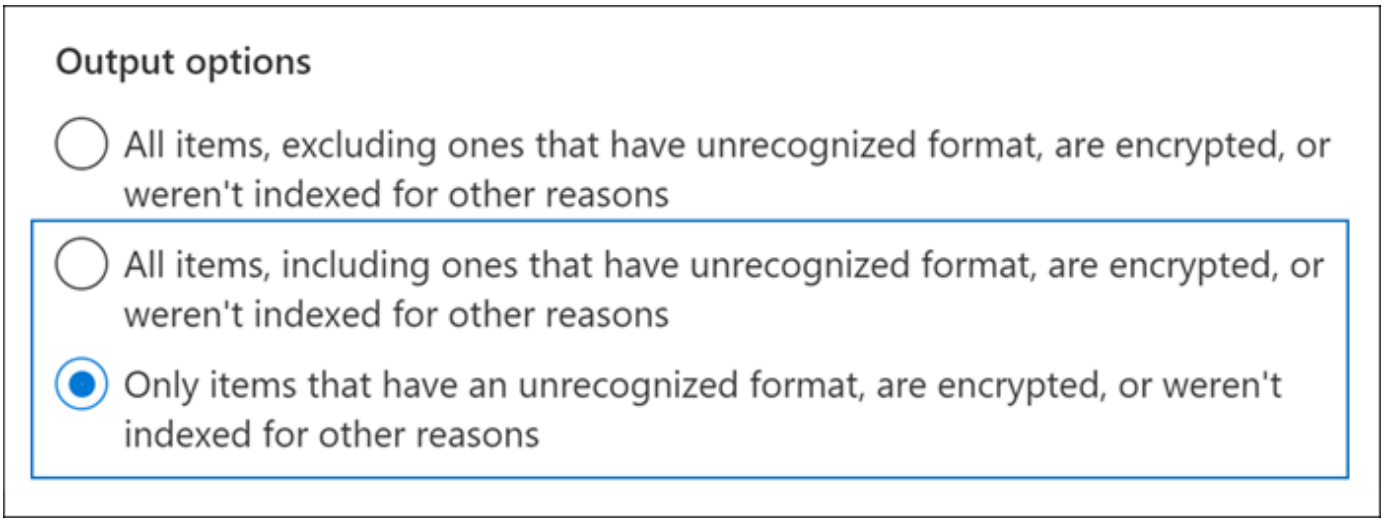
```
(Size of all partially indexed items/Size of all items) x 100
```

```
(10.27 GB/146.46 GB) x 100 = 7.0%
```

So in the previous example, 7% of the total size of mailbox items are from partially indexed items. As previously stated, most organizations customers have less than 1% of content by volume and less than 12% of content by size that is partially indexed.

Working with partially indexed items

In cases when you need to examine partially indexed items to validate that they don't contain relevant information, you can export a content search report that contains information about partially indexed items. When you export a content search report, be sure to choose one of the export options that includes partially indexed items.



When you export eDiscovery search results or a search report using one of these options, the export includes a report named Unindexed Items.csv. This report includes most of the same information as the ResultsLog.csv file; however, the Unindexed Items.csv file also includes two fields related to partially indexed items: **Error Tags** and **Error Properties**. These fields contain information about the indexing error for each partially indexed item. Using the information in these two fields can help you determine whether or not the indexing error for a particular impacts your investigation.

> ⓘ **Note**
>
> The Unindexed Items.csv file also contains fields named **Error Type** and **Error Message**. These are legacy fields that contain information that is similar to the information in the **Error Tags** and **Error Properties** fields, but with less detailed information. You can safely ignore these legacy fields.

Errors related to partially indexed items

Error tags are made up of two pieces of information, the error and the file type. For example, in this error/file-type pair:

```text
parseroutputsize_xls
```

`parseroutputsize` is the error and `xls` is the file type of the file the error occurred on. In cases where the file type wasn't recognized or the file type didn't apply to the error, you will see the value `noformat` in place of the file type.

The following is a list of indexing errors and a description of the possible cause of the error.

Expand table

Error tag	Description
`attachmentcount`	An email message had too many attachments, and some of these attachments weren't processed.
`attachmentdepth`	The content retriever and document parser found too many levels of attachments nested inside other attachments. Some of these attachments weren't processed.
`attachmentrms`	An attachment failed decoding because it was RMS-protected.
`attachmentsize`	A file attached to an email message was too large and couldn't be processed.
`indexingtruncated`	When writing the processed email message to the index, one of the indexable properties was too large and was truncated. The truncated properties are listed in Error Properties field.
`invalidunicode`	An email message contained text that couldn't be processed as valid Unicode. Indexing for this item may be incomplete.
`parserencrypted`	The content of attachment or email message is encrypted, and Microsoft 365 couldn't decode the content.
`parsererror`	An unknown error occurred during parsing. This typically results from a software bug or a service crash.
`parserinputsize`	An attachment was too large for the parser to handle, and the parsing of that attachment didn't happen or wasn't completed.
`parsermalformed`	An attachment was malformed and couldn't be handled by the parser. This result can be due to old file formats, files created by incompatible software, or viruses pretending to be something other than claimed.
`parseroutputsize`	The output from the parsing of an attachment was too large and had to be truncated.
`parserunknowntype`	An attachment had a file type that Microsoft 365 couldn't detect.
`parserunsupportedtype`	An attachment had a file type that Office 365 could detect, but parsing that file type isn't supported.
`propertytoobig`	The value of an email property in Exchange Store was too large to be retrieved and the message couldn't be processed. This typically only happens to the body

Error tag	Description
	property of an email message.
`retrieverrms`	The content retriever failed to decode an RMS-protected message.
`wordbreakertruncated`	Too many words were identified in the document during indexing. Processing of the property stopped when reaching the limit, and the property is truncated.

Error fields describe which fields are affected by the processing error listed in the Error Tags field. If you're searching a property such as `subject` or `participants`, errors in the body of the message won't impact the results of your search. This can be useful when determining exactly which partially indexed items you might need to further investigate.

Last updated on 09/02/2025

Disable reports when you export Content Search results

> **Important**
>
> The eDiscovery Export tool was permanently **retired on August 31, 2025**, this functionality isn't available in the new eDiscovery experience.

When you use the eDiscovery Export tool to export the results of a Content Search in the Microsoft Purview portal, the tool automatically creates and exports two reports that contain additional information about the exported content. These reports are the Results.csv file and the Manifest.xml file (see the Frequently asked questions about disabling export reports section in this article for detailed descriptions of these reports). Because these files can be very large, you can speed up the download time and save disk space by preventing these files from being exported. You can do this by changing the Windows Registry on the computer that you use to export the search results. If you want to include the reports at a later time, you can edit the registry setting.

Create registry settings to disable the export reports

Perform the following procedure on the computer that you'll use to export the results a content search.

1. Close the eDiscovery Export tool if it's open.
2. Perform one or both of the following steps, depending on which export report you want to disable.

   - **Results.csv**

     Save the following text to a Windows registry file by using a filename suffix of .reg; for example, DisableResultsCsv.reg.

     text

     ```
     Windows Registry Editor Version 5.00
     reg add HKLM\SOFTWARE\Microsoft\Exchange\Client\eDiscovery\ExportTool /v
     ResultCsvEnabled /t REG_SZ /d False
     ```

- **Manifest.xml**

  Save the following text to a Windows registry file by using a filename suffix of .reg; for example, DisableManifestXml.reg.

  ```text
  Windows Registry Editor Version 5.00
  reg add HKLM\SOFTWARE\Microsoft\Exchange\Client\eDiscovery\ExportTool /v
  ResultEdrmEnabled /t REG_SZ /d False
  ```

3. In Windows Explorer, select or double-click the .reg file that you created in the previous steps.
4. In the User Access Control window, select **Yes** to let the Registry Editor make the change.
5. When prompted to continue, select **Yes**.

   The Registry Editor displays a message saying that the setting was successfully added to the registry.

Edit registry settings to re-enable the export reports

If you disabled the Results.csv and Manifest.xml reports by creating the .reg files in the previous procedure, you can edit those files to re-enable a report so that it's exported with the search results. Again, perform the following procedure on the computer that you'll use to export the results a content search.

1. Close the eDiscovery Export tool if it's open.
2. Edit one or both of the .reg edit files that you created in the previous procedure.

   - **Results.csv**

     Open the DisableResultsCsv.reg file in Notepad, change the value `False` to `True`, and then save the file. For example, after you edit the file, it looks like this:

     ```text
     Windows Registry Editor Version 5.00
     reg add HKLM\SOFTWARE\Microsoft\Exchange\Client\eDiscovery\ExportTool /v
     ResultCsvEnabled /t REG_SZ /d True
     ```

   - **Manifest.xml**

Open the DisableManifestXml.reg file in Notepad, change the value `False` to `True`, and then save the file. For example, after you edit the file, it looks like this:

```text
Windows Registry Editor Version 5.00
reg add HKLM\SOFTWARE\Microsoft\Exchange\Client\eDiscovery\ExportTool /v
ResultEdrmEnabled /t REG_SZ /d True
```

3. In Windows Explorer, select or double-click a .reg file that you edited in the previous step.
4. In the User Access Control window, select **Yes** to let the Registry Editor make the change.
5. When prompted to continue, select **Yes**.

   The Registry Editor displays a message saying that the setting was successfully added to the registry.

Frequently asked questions about disabling export reports

What are the Results.csv and Manifest.xml reports?

The Results.csv and Manifest.xml files contain additional information about the content that was exported.

- **Results.csv** An Excel document that contains information about each item that is download as a search result. For email, the result log contains information about each message, including:
  - The location of the message in the source mailbox (including whether the message is in the primary or archive mailbox).
  - The date the message was sent or received.
  - The Subject line from the message.
  - The sender and recipients of the message.
  - Whether the message is a duplicate message if you enabled de-duplication when exporting the search results. Duplicate messages will have a value in the **Parent ItemId** column that identifies the message as a duplicate. The value in the **Parent ItemId** column is the same as the value in the **Item DocumentId** column of the message that was exported.

For documents from SharePoint and OneDrive for Business sites, the result log contains information about each document, including:
    - The URL for the document.
    - The URL for the site collection where the document is located.
    - The date that the document was last modified.
    - The name of the document (which is located in the Subject column in the result log).

- **Manifest.xml** A manifest file (in XML format) that contains information about each item included in the search results. The information in this report is the same as the Results.csv report, but it's in the format specified by the Electronic Discovery Reference Model (EDRM). For more information about EDRM, go to https://www.edrm.net .

When should I disable exporting these reports?

It depends on your specific needs. Many organizations don't require additional information about search results, and don't need these reports.

What computer do I have to do this on?

You have to change the registry setting on any local computer that you run the eDiscovery Export tool on.

After I change this setting, do I have to restart the computer?

No, you don't have to restart the computer. But if the eDiscovery Export tool is running, you have to close it and then restart it after you change the registry setting.

Does an existing registry key get edited or does a new key get created?

A new registry key is created the first time you run the .reg file that you created in the procedure in this article. Then the setting is edited each time you change and rerun the .reg edit file.

Last updated on 09/02/2025

Check your search query for errors

> **⊗ Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

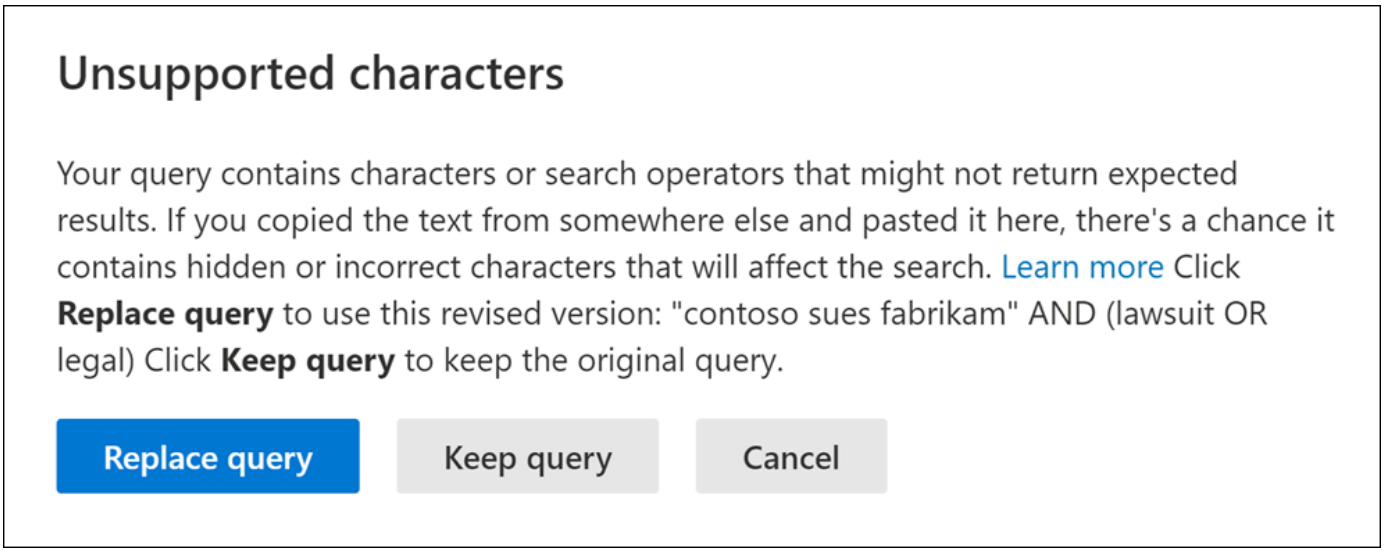
Here's a list of the unsupported characters that we check for in search queries for Content search and Microsoft Purview eDiscovery (Standard). Unsupported characters are often hidden, and they typically cause a search error or return unintended results.

- **Smart quotation marks** - Smart single and double quotation marks (also called curly quotes) aren't supported. Only straight quotation marks can be used in a search query.
- **Non-printable and control characters** - Non-printable and control characters don't represent a written symbol, such as an alpha-numeric character. Examples of non-printable and control characters include characters that format text or separate lines of text.
- **Left-to-right and right-to-left marks** - These marks are control characters used to indicate text direction for left-to-right languages (such as English and Spanish) and right-to-left languages (such as Arabic and Hebrew).
- **Lowercase Boolean operators** - If you use a Boolean operator, such as **AND**, **OR**, and **NOT** in a search query, it must be uppercase. When we check a query for typos, the query syntax will often indicate that a Boolean operator is being used even though lowercase operators might be used; for example, `(WordA or WordB) and (WordC or WordD)`.

What happens if a query has an unsupported character?

If unsupported characters are found in your query, a warning message is displayed that says unsupported characters were found and suggests an alternative. You then have the option keep the original query or replace it with the suggested revised query.

Here's an example of the warning message that's displayed after you select **Check query for typos** for the search query in the previous screenshot. Note the original query used smart quotes and lowercase Boolean operators.



How to prevent unsupported characters in your search queries

Unsupported characters are typically added to a query when you copy the query or parts of the query from other applications (such as Microsoft Word or Microsoft Excel) and paste them in the keyword box on the query page of a Content Search. The best way to prevent unsupported characters is to just type the query in the keyword box. Or you can copy a query from Word or Excel, and then paste it in a plain text editor, such as Microsoft Notepad. Save the text file and select **ANSI** in the **Encoding** drop-down list. This will remove any formatting and unsupported characters. Then you can copy and paste the query from the text file to the keyword query box.

Last updated on 11/21/2025

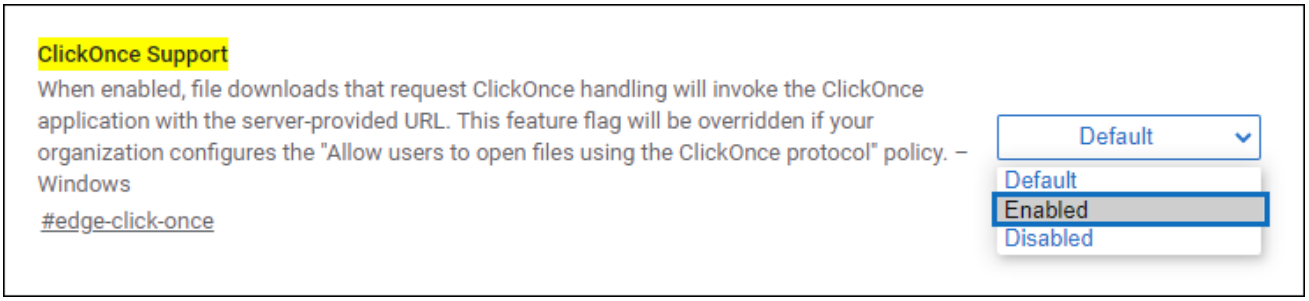
Use the eDiscovery Export Tool in Microsoft Edge

> **Important**
>
> The eDiscovery Export tool was permanently **retired on August 31, 2025**, this functionality isn't available in the new eDiscovery experience.

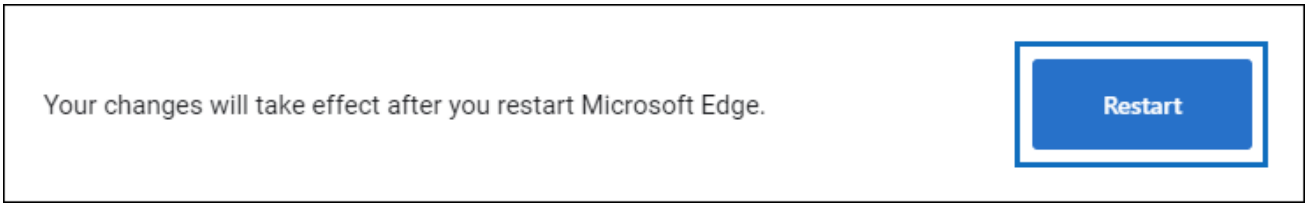
As a result of recent changes to the newest version of Microsoft Edge, ClickOnce support is no longer enabled by default. To continue using the eDiscovery Export Tool to download Content Search or eDiscovery search results, you either need to use Microsoft Internet Explorer or enable ClickOnce support in the newest version of Microsoft Edge.

Enable ClickOnce support in Microsoft Edge

1. In Microsoft Edge, go to **edge://flags/#edge-click-once**.
2. If the existing value is set to **Default** or **Disabled** in the dropdown list, change it to **Enabled**.

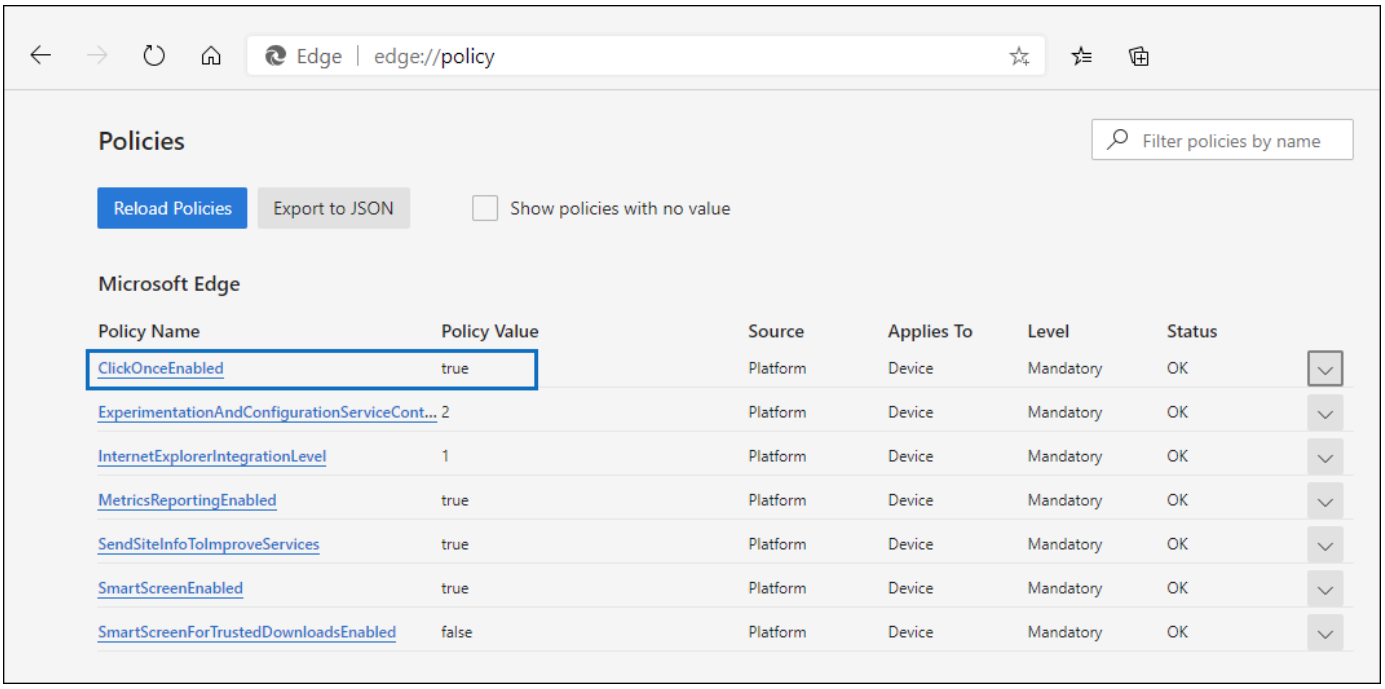


3. Scroll down to the bottom of the browser window and select **Restart** to restart Microsoft Edge.



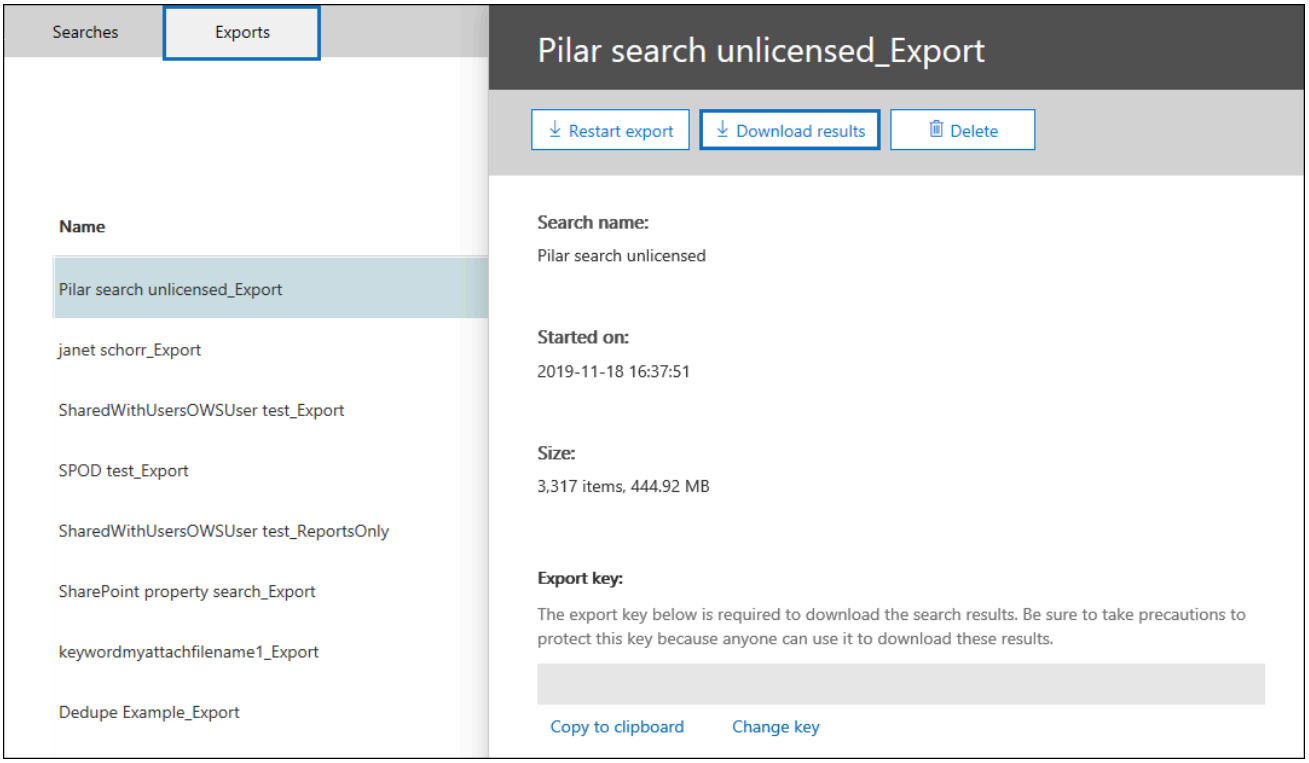
> **Note**

Organizations can use Group Policy to disable ClickOnce support. To check if there's an organizational policy for ClickOnce support, go to **edge://policy**. The following screenshot shows that ClickOnce is enabled across the entire organization. If this policy value is set to **false**, you'll need to contact an admin in your organization.

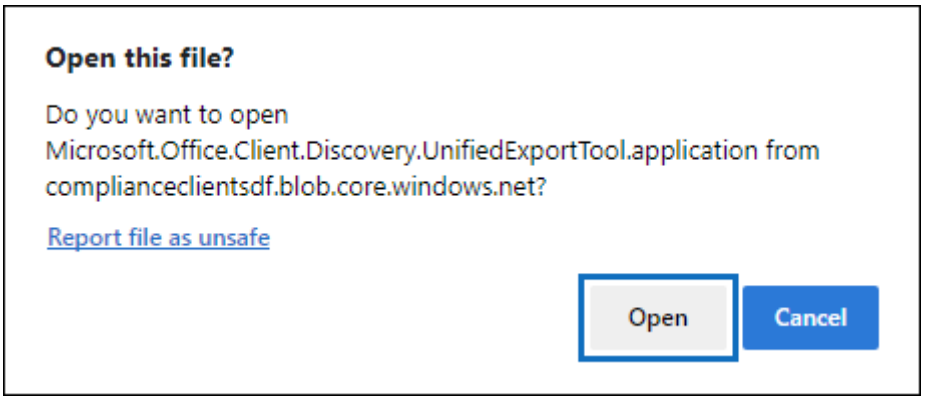


Install and run the eDiscovery Export Tool

1. Select **Download results** on the flyout page of an export in Content Search or an eDiscovery case.



2. You'll be prompted with a confirmation to launch the tool, select **Open**. If the eDiscovery Export Tool isn't installed, you'll be prompted with a Security Warning.



3. Select **Install**. After it's installed, the export tool will launch automatically.

For more information, see the following articles:

- Export Content Search results
- How to enable experiment flags in Microsoft Edge

Last updated on 09/02/2025

Use Content Search to search the mailbox and OneDrive site for a list of users

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

Security & Compliance PowerShell provides numerous cmdlets that let you automate time-consuming eDiscovery-related tasks. Currently, creating a Content search in the Microsoft Purview portal to search a large number of custodian content locations takes time and preparation. Before you create a search, you have to collect the URL for each OneDrive site and then add each mailbox and OneDrive site to the search. This script prompts you for the name of your organization's MySite domain (for example, **contoso** in the URL `https://contoso-my.sharepoint.com`), a list of user email addresses, the name of the new Content Search, and the search query to use. The script gets the OneDrive URL for each user in the list, and then it creates and starts a Content Search that searches the mailbox and OneDrive site for each user in the list, using the search query that you provide.

Permissions and script information

> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

- You have to be a member of the eDiscovery Manager role group in the Microsoft Purview portal or the Microsoft Purview portal and a SharePoint global administrator to run the script in Step 3.

- Be sure to save the list of users that you create in Step 2 and the script in Step 3 to the same folder. This makes it easier to run the script.
- The script includes minimal error handling. Its primary purpose is to quickly and easily search the mailbox and OneDrive site of each user.
- The sample scripts provided in this article aren't supported under any Microsoft standard support program or service. The sample scripts are provided AS IS without warranty of any kind. Microsoft further disclaims all implied warranties including, without limitation, any implied warranties of merchantability or of fitness for a particular purpose. The entire risk arising out of the use or performance of the sample scripts and documentation remains with you. In no event shall Microsoft, its authors, or anyone else involved in the creation, production, or delivery of the scripts be liable for any damages whatsoever (including, without limitation, damages for loss of business profits, business interruption, loss of business information, or other pecuniary loss) arising out of the use of or inability to use the sample scripts or documentation, even if Microsoft has been advised of the possibility of such damages.

Step 1: Install the SharePoint Management Shell

The first step is to install the SharePoint Management Shell. You don't have to use the shell in this procedure, but you have to install it because it contains prerequisites required by the script that you run in Step 3. These prerequisites allow the script to communicate with SharePoint to get the URLs for the OneDrive sites.

Go to Set up the SharePoint Management Shell environment and perform Step 1 and Step 2 to install the SharePoint Management Shell.

Step 2: Generate a list of users

The script in Step 3 will create a Content Search to search the mailboxes and OneDrive accounts for a list of users. You can just type the email addresses in a text file, or you can run a command in PowerShell to get a list of email addresses and save them to a file (located in same folder that you'll save the script to in Step 3).

Here's an Exchange Online PowerShell command that you can run to get a list of email addresses for all users in your organization and save it to a text file named `Users.txt`.

PowerShell

```
Get-Mailbox -ResultSize unlimited -Filter { RecipientTypeDetails -eq 'UserMailbox'} |
```

```
Select-Object PrimarySmtpAddress > Users.txt
```

After you run this command, be sure to open the file and remove the header that contains the property name, `PrimarySmtpAddress`. The text file should just contain a list of email addresses, and nothing else. Make sure there are no blank rows before or after the list of email addresses.

Step 3: Run the script to create and start the search

When you run the script in this step, it prompts you for the following information. Be sure to have this information ready before you run the script.

- **Your user credentials** - The script uses your credentials to access SharePoint to get the OneDrive URLs and to connect to Security & Compliance PowerShell.
- **Name of your MySite domain** - The MySite domain is the domain that contains all the OneDrive sites in your organization. For example, if the URL for your MySite domain is `https://contoso-my.sharepoint.com`, then you would enter `contoso` when the script prompts you for the name of your MySite domain.
- **Pathname of the text file from Step 2** - The pathname of the text file that you created in Step 2. If the text file and the script are located in the same folder, then enter the name of the text file. Otherwise, enter the complete pathname for the text file.
- **Name of the Content Search** - The name of the Content Search that will be created by the script.
- **Search query** - The search query that is used with the Content Search is created and run. For more information about search queries, see Keyword queries and search conditions for eDiscovery.

To run the script:

1. Save the following text to a Windows PowerShell script file by using a filename suffix of .ps1; for example, `SearchEXOOD4B.ps1`. Save the file to the same folder where you saved the list of users in Step 2.

   PowerShell

   ```
   # This PowerShell script will prompt you for the following information:
   #    * Your user credentials
   #    * The name of your organization's MySite domain
   #    * The pathname for the text file that contains a list of user email
   addresses
   ```

```
#    * The name of the Content Search created
#    * The search query string
# The script will then:
#    * Find the OneDrive site for each user in the text file
#    * Create and start a Content Search using the previous information
# Get user credentials
if (!$credentials)
{
    $credentials = Get-Credential
}
# Get the user's MySite domain name.  We use this to create the admin URL and
root URL for OneDrive
$mySiteDomain = Read-Host "What is your organization's MySite domain?  For
example,  'contoso' for 'https://contoso-my.sharepoint.com'"
$AdminUrl = "https://$mySiteDomain-admin.sharepoint.com"
$mySiteUrlRoot = "https://$mySiteDomain-my.sharepoint.com"
# Get other required information
$inputfile = read-host "Enter the file name of the text file that contains the
email addresses for the users you want to search"
$searchName = Read-Host "Enter the name for the new search"
$searchQuery = Read-Host "Enter the search query you want to use"
$emailAddresses = Get-Content $inputfile | where {$_ -ne ""}  | foreach{
$_.Trim() }
# Connect to Security & Compliance PowerShell
if (!$s -or !$a)
{
    Import-Module ExchangeOnlineManagement
    Connect-IPPSSession
}

# Load the SharePoint assemblies from the SharePoint Management Shell
# To install, go to https://go.microsoft.com/fwlink/p/?LinkId=255251
if (!$SharePointClient -or !$SPRuntime -or !$SPUserProfile)
{
    $SharePointClient =
[System.Reflection.Assembly]::LoadWithPartialName("Microsoft.SharePoint.Client")
    $SPRuntime =
[System.Reflection.Assembly]::LoadWithPartialName("Microsoft.SharePoint.Client.Ru
ntime")
    $SPUserProfile =
[System.Reflection.Assembly]::LoadWithPartialName("Microsoft.SharePoint.Client.Us
erProfiles")
    if (!$SharePointClient)
    {
        Write-Error "SharePoint Management Shell isn't installed, please install
from: https://go.microsoft.com/fwlink/p/?LinkId=255251 and then run this script
again"
        return;
    }
}
if (!$spCreds)
{
    $spCreds = New-Object
Microsoft.SharePoint.Client.SharePointOnlineCredentials($credentials.UserName,
$credentials.Password)
```

```
}
# Add the path of the User Profile Service to the SPO admin URL, then create a
new webservice proxy to access it
$proxyaddr = "$AdminUrl/_vti_bin/UserProfileService.asmx?wsdl"
$UserProfileService= New-WebServiceProxy -Uri $proxyaddr -UseDefaultCredential
False
$UserProfileService.Credentials = $credentials
# Take care of auth cookies
$strAuthCookie = $spCreds.GetAuthenticationCookie($AdminUrl)
$uri = New-Object System.Uri($AdminUrl)
$container = New-Object System.Net.CookieContainer
$container.SetCookies($uri, $strAuthCookie)
$UserProfileService.CookieContainer = $container
Write-Host "Getting each user's OneDrive URL"
$urls = @()
foreach($emailAddress in $emailAddresses)
{
    try
    {
        $prop =
$UserProfileService.GetUserProfileByName("i:0#.f|membership|$emailAddress") |
Where-Object { $_.Name -eq "PersonalSpace" }
        $url = $prop.values[0].value
        $furl = $mySiteUrlRoot + $url
        $urls += $furl
        Write-Host "-$emailAddress => $furl"
    }
    catch
    {
        Write-Warning "Could not locate OneDrive for $emailAddress"
    }
}
Write-Host "Creating and starting the search"
$search = New-ComplianceSearch -Name $searchName -ExchangeLocation
$emailAddresses -SharePointLocation $urls -ContentMatchQuery $searchQuery
# Finally, start the search and then display the status
if($search)
{
    Start-ComplianceSearch $search.Name
    Get-ComplianceSearch $search.Name
}
```

2. Open Windows PowerShell and go to the folder where you saved the script and the list of users from Step 2.

3. Start the script; for example:

PowerShell

```
.\SearchEXOOD4B.ps1
```

4. When prompted for your credentials, enter your email address and password, and then select **OK**.

5. Enter following information when prompted by the script. Type each piece of information and then press **Enter**.

    - The name of your MySite domain.
    - The pathname of the text file that contains the list of users.
    - A name for the Content Search.
    - The search query (leave this blank to return all items in the content locations).

    The script gets the URLs for each OneDrive site and then creates and starts the search. You can either run the **Get-ComplianceSearch** cmdlet in Security & Compliance PowerShell to display the search statistics and results, or you can go to the **Content search** page in the Microsoft Purview portal to view information about the search.

Last updated on 09/02/2025

Create, report on, and delete multiple Content Searches

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

Quickly creating and reporting discovery searches is often an important step in eDiscovery and investigations when you're trying to learn about the underlying data, and the richness and quality of your searches. To help you do this, Security & Compliance PowerShell offers a set of cmdlets to automate time-consuming Content Search tasks. These scripts provide a quick and easy way to create many searches, and then run reports of the estimated search results that can help you determine the quantity of data in question. You can also use the scripts to create different versions of searches to compare the results each one produces. These scripts can help you to quickly and efficiently identify and cull your data.

Before you create a Content Search

- You have to be a member of the eDiscovery Manager role group in the Microsoft Purview portal to run the scripts that are described in this article.
- To collect a list of the URLs for the OneDrive for Business sites in your organization that you can add to the CSV file in Step 1, see Create a list of all OneDrive locations in your organization.
- Be sure to save all the files that you create in this article to the same folder. That makes it easier to run the scripts.
- The scripts include minimal error handling. Their primary purpose is to quickly create, report on, and delete multiple Content Searches.
- The sample scripts provided in this article aren't supported under any Microsoft standard support program or service. The sample scripts are provided AS IS without warranty of any

kind. Microsoft further disclaims all implied warranties including, without limitation, any implied warranties of merchantability or of fitness for a particular purpose. The entire risk arising out of the use or performance of the sample scripts and documentation remains with you. In no event shall Microsoft, its authors, or anyone else involved in the creation, production, or delivery of the scripts be liable for any damages whatsoever (including, without limitation, damages for loss of business profits, business interruption, loss of business information, or other pecuniary loss) arising out of the use of or inability to use the sample scripts or documentation, even if Microsoft has been advised of the possibility of such damages.

Step 1: Create a CSV file that contains information about the searches you want to run

The comma separated value (CSV) file that you create in this step contains a row for each user that you want to search. You can search the user's Exchange Online mailbox (which includes the archive mailbox, if it's enabled) and their OneDrive site. Or you can search just the mailbox or the OneDrive site. You can also search any site in your SharePoint Online organization. The script that you run in Step 3 will create a separate search for each row in the CSV file.

1. Copy and paste the following text into a .txt file using NotePad. Save this file to a folder on your local computer. You save the other scripts to this folder as well.

   text

   ```
   ExchangeLocation,SharePointLocation,ContentMatchQuery,StartDate,EndDate
   sarad@contoso.onmicrosoft.com,https://contoso-
   my.sharepoint.com/personal/sarad_contoso_onmicrosoft_com,(lawsuit OR
   legal),1/1/2000,12/31/2005
   sarad@contoso.onmicrosoft.com,https://contoso-
   my.sharepoint.com/personal/sarad_contoso_onmicrosoft_com,(lawsuit OR
   legal),1/1/2006,12/31/2010
   sarad@contoso.onmicrosoft.com,https://contoso-
   my.sharepoint.com/personal/sarad_contoso_onmicrosoft_com,(lawsuit OR
   legal),1/1/2011,3/21/2016
   ,https://contoso.sharepoint.com/sites/contoso,,,3/21/2016
   ,https://contoso-
   my.sharepoint.com/personal/davidl_contoso_onmicrosoft_com,,1/1/2015,
   ,https://contoso-
   my.sharepoint.com/personal/janets_contoso_onmicrosoft_com,,1/1/2015,
   ```

   The first row, or header row, of the file lists the parameters that will be used by **New-ComplianceSearch** cmdlet (in the script in Step 3) to create a new Content Searches. Each parameter name is separated by a comma. Make sure there aren't any spaces in the header

row. Each row under the header row represents the parameter values for each search. Be sure to replace the placeholder data in the CSV file with your actual data.

2. Open the .txt file in Excel, and then use the information in the following table to edit the file with information for each search.

Expand table

Parameter	Description
`ExchangeLocation`	The SMTP address of the user's mailbox.
`SharePointLocation`	The URL for the user's OneDrive for Business site or the URL for any site in your organization. For the URL for OneDrive for Business sites, use this format: `https://<your organization>-my.sharepoint.com/personal/<user alias>_<your organization>_onmicrosoft_com`. For example, `https://contoso-my.sharepoint.com/personal/sarad_contoso_onmicrosoft_com`.
`ContentMatchQuery`	The search query for the search. For more information about creating a search query, see Keyword queries and search conditions for Content Search.
`StartDate`	For email, the date on or after a message was received by a recipient or sent by the sender. For documents on SharePoint or OneDrive for Business sites, the date on or after a document was last modified.
`EndDate`	For email, the date on or before a message was sent by a sent by the user. For documents on SharePoint or OneDrive for Business sites, the date on or before a document was last modified.

3. Save the Excel file as a CSV file to a folder on your local computer. The script that you create in Step 3 will use the information in this CSV file to create the searches.

Step 2: Connect to Security & Compliance PowerShell

The next step is to connect to Security & Compliance PowerShell for your organization. For step-by-step instructions, see Connect to Security & Compliance PowerShell.

Step 3: Run the script to create and start the searches

The script in this step creates a separate Content Search for each row in the CSV file that you created in Step 1. When you run this script, you're prompted for two values:

- **Search Group ID** - This name provides an easy way to organize the searches that are created from the CSV file. Each search that's created is named with the Search Group ID, and then a number is appended to the search name. For example, if you enter **ContosoCase** for the Search Group ID, then the searches are named **ContosoCase_1**, **ContosoCase_2**, **ContosoCase_3**, and so on. The name you type is case sensitive. When you use the Search Group ID in Step 4 and Step 5, you have to use the same case as you did when you created it.

- **CSV file** - The name of the CSV file that you created in Step 1. Be sure to include the use the full filename and include the .csv file extension; for example, `ContosoCase.csv`.

To run the script:

1. Save the following text to a Windows PowerShell script file by using a filename suffix of .ps1; for example, `CreateSearches.ps1`. Save the file to the same folder where you saved the other files.

Powershell

```
# Get the Search Group ID and the location of the CSV input file
$searchGroup = Read-Host 'Search Group ID'
$csvFile = Read-Host 'Source CSV file'

# Do a quick check to make sure our group name will not collide with other
searches
$searchCounter = 1
import-csv $csvFile |
  ForEach-Object{

 $searchName = $searchGroup +'_' + $searchCounter
 $search = Get-ComplianceSearch $searchName -EA SilentlyContinue
 if ($search)
 {
    Write-Error "The Search Group ID conflicts with existing searches.  Please
choose a search group name and restart the script."
    return
 }
 $searchCounter++
}

$searchCounter = 1
import-csv $csvFile |
  ForEach-Object{

 # Create the query
 $query = $_.ContentMatchQuery
 if(($_.StartDate -or $_.EndDate))
 {
       # Add the appropriate date restrictions.  NOTE: Using the Date condition
```

```
property here because it works across Exchange, SharePoint, and OneDrive for
Business.
       # For Exchange, the Date condition property maps to the Sent and Received
dates; for SharePoint and OneDrive for Business, it maps to Created and Modified
dates.
       if($query)
       {
           $query += " AND"
       }
       $query += " ("
       if($_.StartDate)
       {
           $query += "Date >= " + $_.StartDate
       }
       if($_.EndDate)
       {
           if($_.StartDate)
           {
               $query += " AND "
           }
           $query += "Date <= " + $_.EndDate
       }
       $query += ")"
 }

  # -ExchangeLocation can't be set to an empty string, set to null if there's no
location.
  $exchangeLocation = $null
  if ( $_.ExchangeLocation)
  {
        $exchangeLocation = $_.ExchangeLocation
  }

 # Create and run the search
 $searchName = $searchGroup +'_' + $searchCounter
 Write-Host "Creating and running search: " $searchName -NoNewline
 $search = New-ComplianceSearch -Name $searchName -ExchangeLocation
$exchangeLocation -SharePointLocation $_.SharePointLocation -ContentMatchQuery
$query

 # Start and wait for each search to complete
 Start-ComplianceSearch $search.Name
 while ((Get-ComplianceSearch $search.Name).Status -ne "Completed")
 {
    Write-Host " ." -NoNewline
    Start-Sleep -s 3
 }
 Write-Host ""

 $searchCounter++
}
```

2. In Windows PowerShell, go to the folder where you saved the script in the previous step, and then run the script; for example:

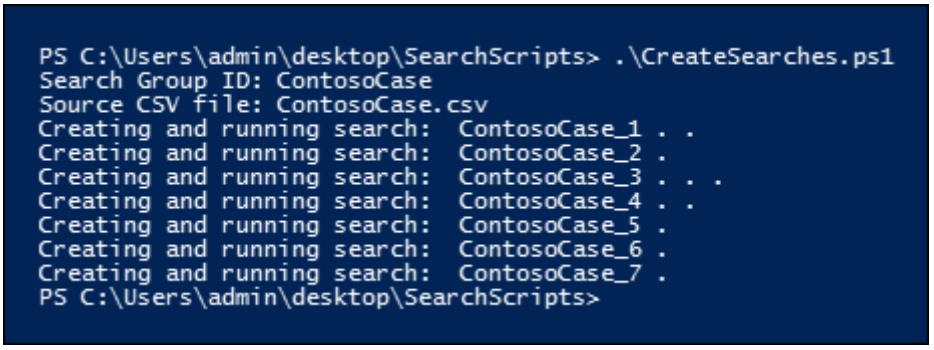
```
Powershell
.\CreateSearches.ps1
```

3. At the **Search Group ID** prompt, type a search group name, and then press **Enter**; for example, `ContosoCase`. Remember that this name is case sensitive, so you have to type it the same way in the subsequent steps.

4. At the **Source CSV file** prompt, type the name of the CSV file, including the .csv file extension; for example, `ContosoCase.csv`.

5. Press **Enter** to continue running the script.

   The script displays the progress of creating and running the searches. When the script is complete, it returns to the prompt.



Step 4: Run the script to report the search estimates

After you create the searches, the next step is to run a script that displays a simple report of the number of search hits for each search that was created in Step 3. The report also includes the size of results for each search, and the total number of hits and total size of all searches. When you run the reporting script, you are prompted for the Search Group ID, and a CSV filename if you want to save the report to a CSV file.

1. Save the following text to a Windows PowerShell script file by using a filename suffix of .ps1; for example, `SearchReport.ps1`. Save the file to the same folder where you saved the other files.

```
Powershell
```

```
$searchGroup = Read-Host 'Search Group ID'
$outputFile = Read-Host 'Enter a file name or file path to save the report to a
.csv file. Leave blank to only display the report'
$searches = Get-ComplianceSearch | ?{$_.Name -clike $searchGroup + "_*"}
$allSearchStats = @()
foreach ($partialObj in $searches)
{
   $search = Get-ComplianceSearch $partialObj.Name
   $sizeMB = [System.Math]::Round($search.Size / 1MB, 2)
   $searchStatus = $search.Status
   if($search.Errors)
   {
       $searchStatus = "Failed"
   }elseif($search.NumFailedSources -gt 0)
   {
       $searchStatus = "Failed Sources"
   }
   $searchStats = New-Object PSObject
   Add-Member -InputObject $searchStats -MemberType NoteProperty -Name Name -
Value $search.Name
   Add-Member -InputObject $searchStats -MemberType NoteProperty -Name
ContentMatchQuery -Value $search.ContentMatchQuery
   Add-Member -InputObject $searchStats -MemberType NoteProperty -Name Status -
Value $searchStatus
   Add-Member -InputObject $searchStats -MemberType NoteProperty -Name Items -
Value $search.Items
   Add-Member -InputObject $searchStats -MemberType NoteProperty -Name "Size" -
Value $search.Size
   Add-Member -InputObject $searchStats -MemberType NoteProperty -Name "Size(MB)"
-Value $sizeMB
   $allSearchStats += $searchStats
}
# Calculate the totals
$allItems = ($allSearchStats | Measure-Object Items -Sum).Sum
# Convert the total size to MB and round to the nearst 100th
$allSize = ($allSearchStats | Measure-Object 'Size' -Sum).Sum
$allSizeMB = [System.Math]::Round($allSize  / 1MB, 2)
# Get the total successful searches and total of all searches
$allSuccessCount = ($allSearchStats |?{$_.Status -eq "Completed"}).Count
$allCount = $allSearchStats.Count
$allStatus = [string]$allSuccessCount + " of " + [string]$allCount
# Totals Row
$totalSearchStats = New-Object PSObject
Add-Member -InputObject $totalSearchStats -MemberType NoteProperty -Name Name -
Value "Total"
Add-Member -InputObject $totalSearchStats -MemberType NoteProperty -Name Status -
Value $allStatus
Add-Member -InputObject $totalSearchStats -MemberType NoteProperty -Name Items -
Value $allItems
Add-Member -InputObject $totalSearchStats -MemberType NoteProperty -Name
"Size(MB)" -Value $allSizeMB
$allSearchStats += $totalSearchStats
# Just get the columns we're interested in showing
$allSearchStatsPrime = $allSearchStats | Select-Object Name, Status, Items,
```

```
"Size(MB)", ContentMatchQuery
# Print the results to the screen
$allSearchStatsPrime |ft -AutoSize -Wrap
# Save the results to a CSV file
if ($outputFile)
{
   $allSearchStatsPrime | Export-Csv -Path $outputFile -NoTypeInformation
}
```

2. In Windows PowerShell, go to the folder where you saved the script in the previous step, and then run the script; for example:

Powershell

```
.\SearchReport.ps1
```

3. At the **Search Group ID** prompt, type a search group name, and then press **Enter**; for example `ContosoCase`. Remember that this name is case sensitive, so you have to type it the same way you did when you ran the script in Step 3.

4. At the **File path to save the report to a CSV file (leave blank to just display the report)** prompt, type a file name of complete filename path (including the .csv file extension) if you want to save the report to a CSV file. name of the CSV file, including the .csv file extension. For example, you could type `ContosoCaseReport.csv` to save it to the current directory or you could type `C:\Users\admin\OneDrive for Business\ContosoCase\ContosoCaseReport.csv` to save it to a different folder. You can also leave the prompt blank to display the report but not save it to a file.

5. Press **Enter**.

   The script displays the progress of creating and running the searches. When the script is complete, the report is displayed.

```
PS C:\Users\admin\desktop\SearchScripts> .\SearchReport.ps1
Search Group ID: ContosoCase
Enter a file name or file path to save the report to a .csv file. Leave blank to only display the report:

Name          Status       Items  Size (MB)    ContentMatchQuery
----          ------       -----  ---------    -----------------
ContosoCase_1 Completed      865      23.24    (lawsuit OR legal) AND (Date >= 1/1/2000 AND Date <= 12/31/2005)
ContosoCase_2 Completed      229       8.81    (lawsuit OR legal) AND (Date >= 1/1/2006 AND Date <= 12/31/2010)
ContosoCase_3 Completed      249       9.86    (lawsuit OR legal) AND (Date >= 1/1/2011 AND Date <= 12/31/2015)
ContosoCase_4 Completed     1600     663.43    (Date <= 12/31/2015)
ContosoCase_5 Completed      160      62.70    (Date >= 1/1/2015)
ContosoCase_6 Completed      140      65.20    (Date >= 1/1/2015)
ContosoCase_7 Completed       90       2.60    (Date >= 1/1/2015)
Total         7 of 7        3333     835.84


PS C:\Users\admin\desktop\SearchScripts>
```

> **Note**

If the same mailbox or site is specified as a content location in more than one search in a search group, the total results estimate in the report (for both the number of items and the total size) might include results for the same items. That's because the same email message or document will be counted more than once if it matches the query for different searches in the search group.

Step 5: Run the script to delete the searches

Because you might be creating numerous searches, this last script just makes it easy to quickly delete the searches you created in Step 3. Like the other scripts, this one also prompts you for the Search Group ID. All searches with the Search Group ID in the search name will be deleted when you run this script.

1. Save the following text to a Windows PowerShell script file by using a filename suffix of .ps1; for example, `DeleteSearches.ps1`. Save the file to the same folder where you saved the other files.

   Powershell

   ```
   # Delete all searches in a search group
   $searchGroup = Read-Host 'Search Group ID'
   Get-ComplianceSearch |
      ForEach-Object{
      # If the name matches the search group name pattern (case sensitive), delete
   the search
      if ($_.Name -cmatch $searchGroup + "_\d+")
      {
          Write-Host "Deleting search: " $_.Name
          Remove-ComplianceSearch $_.Name -Confirm:$false
      }
   }
   ```

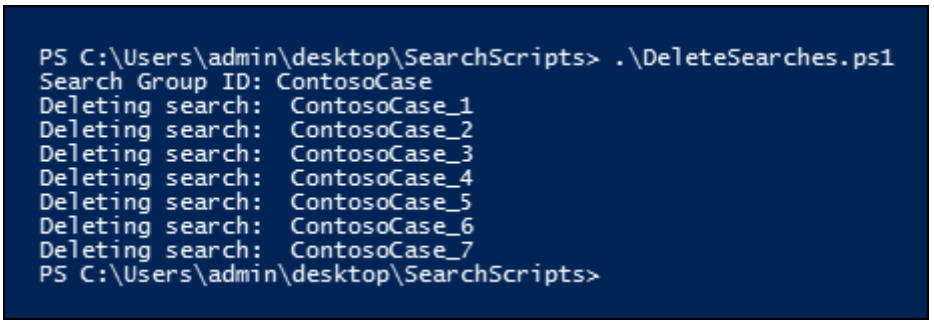
2. In Windows PowerShell, go to the folder where you saved the script in the previous step, and then run the script; for example:

   Powershell

   ```
   .\DeleteSearches.ps1
   ```

3. At the **Search Group ID** prompt, type a search group name for the searches that you want to delete, and then press **Enter**; for example, `ContosoCase`. Remember that this name is case sensitive, so you have to type it the same way you did when you ran the script in Step 3.

The script displays the name of each search that's deleted.



Last updated on 03/31/2025

Clone a Content Search

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

Creating a Content Search in the Microsoft Purview portal in Microsoft 365 that searches many mailboxes or SharePoint and OneDrive for Business sites can take a while. Specifying the sites to search can also be prone to errors if you mistype a URL. To avoid these issues, you can use the Windows PowerShell script in this article to quickly clone an existing Content Search. When you clone a search, a new search (with a different name) is created that contains the same properties (such as the content locations and the search query) as the original search. Then you can edit the new search by changing the keyword query or the date range, and run it.

Why clone Content Searches?

- To compare the results of different keyword search queries run on the same content locations.
- To save you from having to reenter a large number of content locations when you create a new search.
- To decrease the size of the search results. For example, if you have a search that returns too many results to export, you can clone the search and then add a search condition based on a date range to reduce the number of search results.

Script information

- You need to install the Exchange Online PowerShell module. For instructions, see Install and maintain the Exchange Online PowerShell module.

- You have to be a member of the eDiscovery Manager role group in the Microsoft Purview portal to run the script described in this topic.
- The script includes minimal error handling. The primary purpose of the script is to quickly clone a content search.
- The script creates a new Content Search, but doesn't start it.
- This script takes into account whether the Content Search that you're cloning is associated with an eDiscovery case. If the search is associated with a case, the new search will also be associated with the same case. If the existing search isn't associated with a case, the new search will be listed on the **Content search** page in the Microsoft Purview portal.
- The sample script provided in this topic isn't supported under any Microsoft standard support program or service. The sample script is provided AS IS without warranty of any kind. Microsoft further disclaims all implied warranties including, without limitation, any implied warranties of merchantability or of fitness for a particular purpose. The entire risk arising out of the use or performance of the sample script and documentation remains with you. In no event shall Microsoft, its authors, or anyone else involved in the creation, production, or delivery of the scripts be liable for any damages whatsoever (including, without limitation, damages for loss of business profits, business interruption, loss of business information, or other pecuniary loss) arising out of the use of or inability to use the sample scripts or documentation, even if Microsoft has been advised of the possibility of such damages.

Step 1: Run the script to clone a search

The script in this step will create a new Content Search by cloning an existing one. When you run this script, you'll be prompted for the following information:

- **Your user credentials** - The script will use your credentials to connect to Security & Compliance PowerShell. As previously stated, you have to be a member of the eDiscovery Manager role group in the Microsoft Purview portal to run the script.
- **The name of the existing search** - This is the Content Search that you want to clone.
- **The name of the new search that will be created** - If you leave this value blank, the script will create a name for the new search that is based on the name of the search that you're cloning.

To clone a search:

1. Save the following text to a Windows PowerShell script file by using a filename suffix of .ps1; for example, `CloneSearch.ps1`.

PowerShell

```
# This PowerShell script clones an existing content search in Microsoft Purview
compliance.

# Ask for the name of the search you want to clone
$searchName = Read-Host 'Enter the name of the search that you want to clone'
# Ask for the name of the new search
$newSearchName = Read-Host 'Enter a name for the new search [leave blank to
automatically generate a name]'
$originalSearch = Get-ComplianceSearch $searchName -EA SilentlyContinue
# Make sure we have a valid search before continuing
if(!$originalSearch)
{
    Write-Error "Couldn't find search: $searchName"
    return
}
$searchNameCounter = 1
# Find a suitable name for the new search
while(!$newSearchName)
{
    $newSearchName = $originalSearch.Name + "_" + $searchNameCounter
    $tempSearch = Get-ComplianceSearch $newSearchName -EA SilentlyContinue
    if ($tempSearch)
    {
        $newSearchName = $null
        $searchNameCounter++
    }
}
$caseName
# Determine if the search is part of a case; if so get the case name
if ($originalSearch.CaseId)
{
    $searchCase = Get-ComplianceCase $originalSearch.CaseId
    $caseName = $searchCase.Name
}
# Need to cast this value as a Boolean the old fashion way
$allowNotFoundExchangeLocationsEnabled = $false
if ($originalSearch.AllowNotFoundExchangeLocationsEnabled)
{
    $allowNotFoundExchangeLocationsEnabled = $true
}
$newSearch = New-ComplianceSearch -Name $newSearchName -
AllowNotFoundExchangeLocationsEnabled $allowNotFoundExchangeLocationsEnabled -
Case $caseName -ContentMatchQuery $originalSearch.ContentMatchQuery -Description
$originalSearch.Description -ExchangeLocation $originalSearch.ExchangeLocation -
ExchangeLocationExclusion $originalSearch.ExchangeLocationExclusion -Language
$originalSearch.Language -SharePointLocation $originalSearch.SharePointLocation -
SharePointLocationExclusion $originalSearch.SharePointLocationExclusion -
PublicFolderLocation $originalSearch.PublicFolderLocation
```

```
if ($newSearch)
{
    Write-Host $newSearch.Name "was successfully created" -ForegroundColor Yellow
}
```

2. Connect to Security & Compliance PowerShell. In the same PowerShell window, go to the folder where you saved the script.

3. Run the script; for example:

   PowerShell

   ```
   .\CloneSearch.ps1
   ```

4. Enter following information when prompted by the script. Type each piece of information and then press **Enter**.

   - The name of the existing search.
   - The name of the new search.

   The script creates the new Content Search, but doesn't start it. This gives you a chance to edit and run the search in the next step. You can view the properties of the new search by running the **Get-ComplianceSearch** cmdlet or by going to the **Content search** or **eDiscovery** page in the Microsoft Purview portal, depending on whether the new search is associated with a case.

Step 2: Edit and run the cloned search in the Microsoft Purview portal

After you run the script to clone an existing Content Search, the next step is to go to the Microsoft Purview portal to edit and run the new search. As previously stated, you can edit a search by changing the keyword search query and adding or removing search conditions. For more information, see:

- Content Search in Office 365
- Keyword queries and search conditions for Content Search
- eDiscovery cases

Last updated on 11/21/2025

Get started with eDiscovery (Standard)

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

Microsoft Purview eDiscovery (Standard) in Microsoft Purview provides a basic eDiscovery tool that organizations can use to search and export content in Microsoft 365 and Office 365. You can also use eDiscovery (Standard) to place an eDiscovery hold on content locations, such as Exchange mailboxes, SharePoint sites, OneDrive accounts, and Microsoft Teams. Nothing is needed to deploy eDiscovery (Standard), but there are some prerequisite tasks that an IT admin and eDiscovery manager have to complete before your organization can start using eDiscovery (Standard) to search, export, and preserve content.

This article discusses the steps necessary to set up eDiscovery (Standard). This includes ensuring the proper licensing required to access eDiscovery (Standard) and place an eDiscovery hold on content locations, as well as assigning permissions to your IT, legal, and investigation team so they can access and manage cases. This article also provides a high-level overview of using cases to search for and export content.

Step 1: Verify and assign appropriate licenses

Microsoft Purview eDiscovery capabilities are included with Microsoft Purview. The licensing requirements may vary even within capabilities, depending on configuration options. For licensing requirements, guidance, and options, see the Microsoft 365 guidance for security & compliance.

For information about how to assign licenses, see Assign licenses to users.

Step 2: Verify that required eDiscovery apps are enabled

eDiscovery (Standard) requires the following Enterprise apps to be enabled in your Microsoft 365 or Office 365 organization. If these apps aren't enabled, you won't be able to access eDiscovery (Standard) view, filter, and search features.

Expand table

App	App ID
Built-in security add-on for on-premises mailboxes (formerly Exchange Online Protection)	00000007-0000-0ff1-ce00-000000000000
ComplianceWorkbenchApp	92876b03-76a3-4da8-ad6a-0511ffdf8647
Office365Zoom	0d38933a-0bbd-41ca-9ebd-28c4b5ba7cb7

For more information about how to view and enable apps, see:

- Quickstart: View enterprise applications
- Quickstart: Add an enterprise application

Step 3: Assign eDiscovery permissions

To access eDiscovery (Standard) or be added as a member of an eDiscovery (Standard) case, a user must be assigned the appropriate permissions. Specifically, a user must be added as a member of the eDiscovery Manager role group in the Microsoft Purview portal. Members of this role group can create and manage eDiscovery (Standard) cases. They can add and remove members, place an eDiscovery hold on users, create and edit searches, and export content from an eDiscovery (Standard) case.

For step-by-step guidance to configure permissions for eDiscovery, see Assign eDiscovery permissions.

Step 4: Create an eDiscovery (Standard) case

The next step is to create a case and start using eDiscovery (Standard). Complete the following steps to create a case and add members. The user who creates the case is automatically added as a member.

1. Go to Microsoft Purview portal and sign in using the credentials for a user account that has been assigned the appropriate eDiscovery permissions. Members of the Organization Management role group can also create eDiscovery (Standard) cases.
2. In the left navigation pane of the Microsoft Purview portal, select **Show all**, and then select **eDiscovery (Standard)**.
3. On the **eDiscovery (Standard)** page, select **Create a case**.
4. On the **New case** flyout page, give the case a name (required) and then type an optional description. The case name must be unique in your organization.
5. Select **Save** to create the case.

The new case is created and displayed on the eDiscovery (Standard) page. You may have to select **Refresh** to display the new case.

Step 5 (optional): Add members to an eDiscovery (Standard) case

If you create a case in Step 3 and you're the only person who will use the case, then you don't have to perform this step. You can start using the case to create eDiscovery holds, search for content, and export search results. Perform this step if you want to give other users (or roles group) access to the case.

1. On the **eDiscovery (Standard)** page in the Microsoft Purview portal, select the name of the case that you want to add members to.

2. On the case home page, select the **Settings** tab, and then select **Access & permissions**.

3. On the **Access & permissions** flyout page, under **Members**, select **Add** to add members to the case.

   You can also choose to add role groups as members of a case. Under **Role groups**, select **Add**. You can only assign the role groups that you're a member of to a case. That's because role groups control who can assign members to an eDiscovery case.

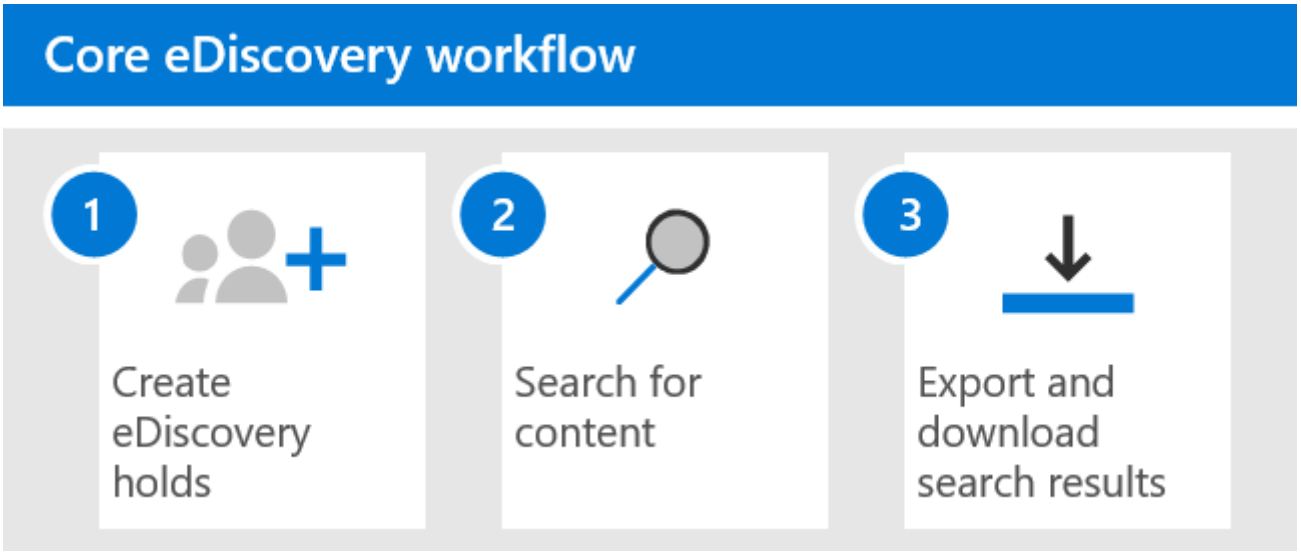
4. In the list of people or role groups that can be added as members of the case, select to the left of the name of the people (or role groups) that you want to add. If you have a large list of people or role groups who can be added as members, use the **Search** box to search for a specific person or role group in the list.

5. After you select the people or role groups to add as members of the case, select **Save** to save the new members or role groups.

> **Important**
>
> - If a role is added or removed from a role group that you've added as a member of a case, then the role group will be automatically removed as a member of the case (or any case the role group is a member of). The reason for this is to protect your organization from inadvertently providing additional permissions to members of a case. Similarly, if a role group is deleted, it will be removed from all cases it was a member of. For more information, see Assign eDiscovery permissions.
> - Only an eDiscovery Administrator can remove members from a case. Users who are members of the eDiscovery Manager subgroup can't remove members from a case, even if the user created the case.

Explore the eDiscovery (Standard) workflow

To get you started using eDiscovery (Standard), here's a simple workflow of creating eDiscovery holds for people of interest, searching for content that relevant to your investigation, and then exporting that data for further review. In each of these steps, we'll also highlight some extended eDiscovery (Standard) functionality that you can explore.



1. Create an eDiscovery hold. The first step after creating a case is placing a hold (also called an *eDiscovery hold*) on the content locations of the people of interest in your investigation. Content locations include Exchange mailboxes, SharePoint sites, OneDrive accounts, and the mailboxes and sites associated with Microsoft Teams and Microsoft 365 Groups. While this step is optional, creating an eDiscovery hold preserves content that may be relevant to the case during the investigation. When you create an eDiscovery hold you can preserve all content in specific content locations or you can create a query-based hold to preserve only

the content that matches a hold query. In addition to preserving content, another good reason to create eDiscovery holds is to quickly search the content locations on hold (instead of having to select each location to search) when you create and run searches in the next step. After you complete your investigation, you can release any hold that you created.

2. **Search for content**. After you create eDiscovery holds, use the built-in search tool to search the content locations on hold. You can also search other content locations for data that may be relevant to the case. You can create and run different searches that are associated with the case. You use keywords, properties, and conditions to build search queries that return search results with the data that's most likely relevant to the case. You can also:

    - View search statistics that may help you refine a search query to narrow the results.
    - Preview the search results to quickly verify whether the relevant data is being found.
    - Revise a query and rerun the search.

3. **Export and download search results**. After you search for and find data that's relevant to your investigation, you can export it out of Office 365 for review by people outside of the investigation team. Exporting data is a two-step process. The first step is to export the results of a search in the case out of Office 365. This is accomplished by copying the results of a search to a Microsoft-provided Azure Storage location. The next step is to use the eDiscovery Export tool to download the content to a local computer. In addition to the exported data files, the export package contains an export report, a summary report, and an error report.

> **Important**
>
> The eDiscovery Export tool was permanently **retired on August 31, 2025**, this functionality isn't available in the new **eDiscovery experience**.

Last updated on 02/12/2026

Search for content in an eDiscovery (Standard) case

> **⊗ Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

After a Microsoft Purview eDiscovery (Standard) case is created and people of interest in the case are placed on hold, you can create and run one or more searches for content relevant to the case. Searches associated with an eDiscovery (Standard) case aren't listed on the **Content search** page in the Microsoft Purview portal. These searches are listed on the **Searches** page of the eDiscovery (Standard) case the searches are associated with. This also means that searches associated with a case can only be accessed by case members.

For organizations that have EU General Data Protection Regulation (GDPR) requirements to protect and enable individuals' privacy rights inside the European Union (EU), you can also manage investigations in response to a Data Subject Requests (DSRs) submitted by a person in your organization. The User Data Search case tool has been retired, and its functionality has merged with eDiscovery (Standard). You can now use content to search for content to support DSRs all locations supported by eDiscovery (Standard) searches.

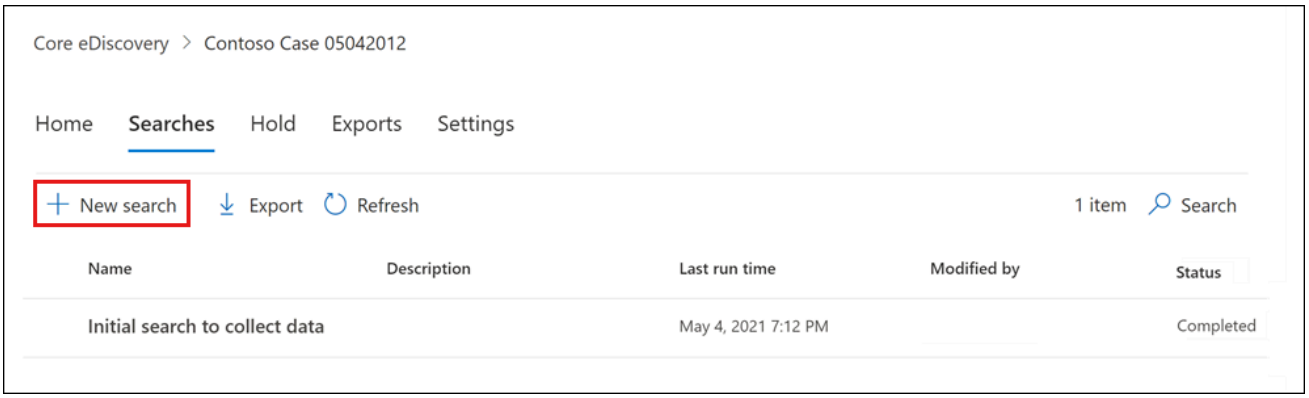
To create an eDiscovery (Standard) search:

1. Go to https://compliance.microsoft.com and sign in using the credentials for user account that has been assigned the appropriate eDiscovery permissions and is a member of the case.

2. In the left navigation pane of the Microsoft Purview portal, select **Show all**, and then select **eDiscovery (Standard)**.

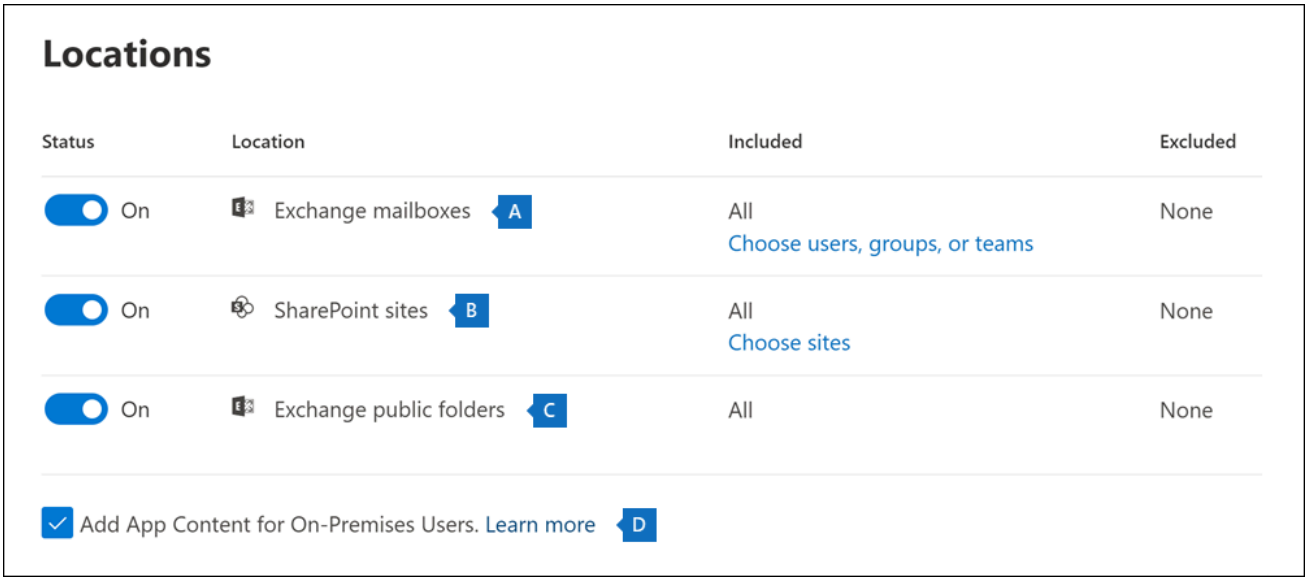
3. On the **eDiscovery (Standard)** page, select the case that you want to create an associated search, and then select **Open case**.

4. On the **Home** page for the case, select the **Searches** tab, and then select **New search**.



5. In the **New search** workflow, type a name for the search, and an optional description that helps identify the search. The name of the search must be unique in your organization.

6. On the **Locations** page, choose the content locations that you want to search. You can search mailboxes, sites, and public folders.



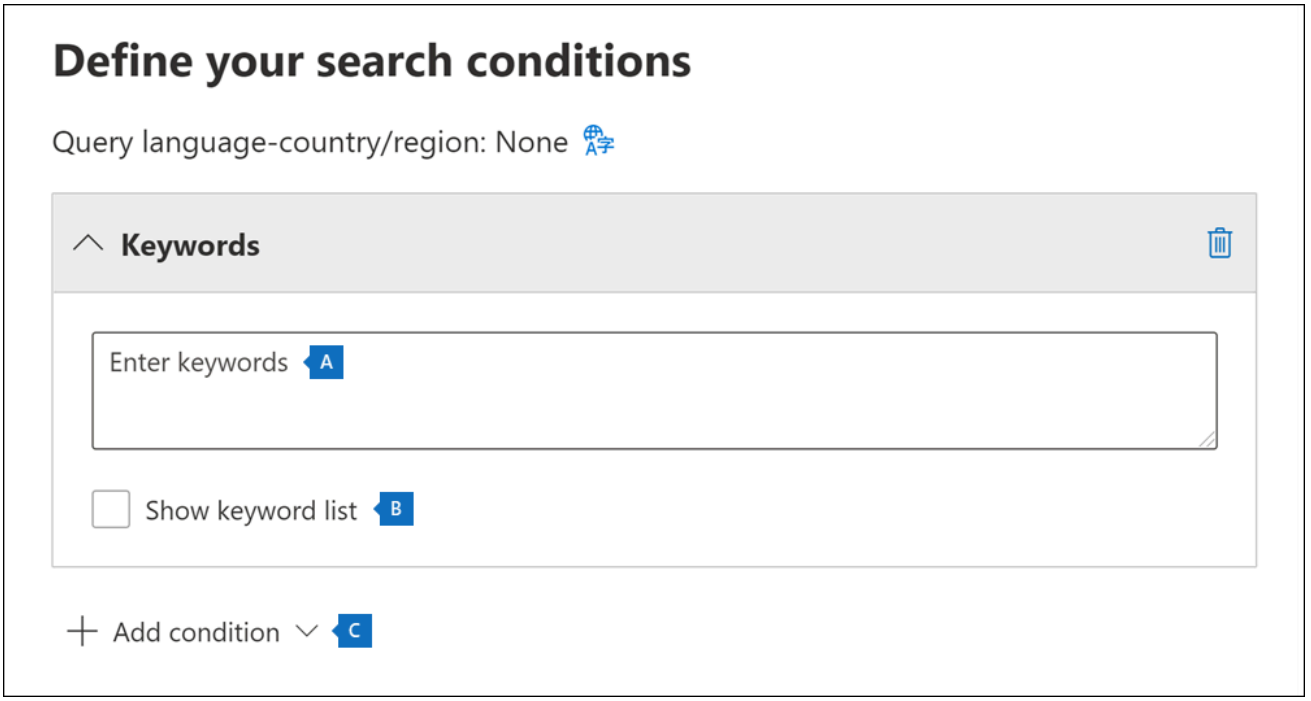
   a. **Exchange mailboxes**: Set the toggle to **On** and then select **Choose users, groups, or teams** to specify the mailboxes to place on hold. Use the search box to find user mailboxes and distribution groups (to place a hold on the mailboxes of group members) to place on hold. You can also search the mailbox associated with a Microsoft Team (for channel messages), Office 365 Group, and Viva Engage Group. All Copilot activity data (user prompts and Copilot responses) generated in supported Microsoft 365 apps and services is stored in custodian mailboxes. For more information about the application data stored in mailboxes, see Content stored in mailboxes for eDiscovery.

b. **SharePoint sites**: Set the toggle to **On** and then select **Choose sites** to specify SharePoint sites and OneDrive accounts to place on hold. Type the URL for each site that you want to place on hold. You can also add the URL for the SharePoint site for a Microsoft Team, Office 365 Group, or Viva Engage Group. The Recycle Bin in SharePoint sites isn't indexed and therefore unavailable for searching. As a result, eDiscovery searches can't find any Recycle Bin content to place holds.

c. **Exchange public folders**: Set the toggle to **On** to put all public folders in your Exchange Online organization on hold. You can't choose specific public folders to put on hold. Leave the toggle switch off if you don't want to put a hold on public folders.

d. Keep this checkbox selected to search for Teams content for on-premises users. For example, if you search all Exchange mailboxes in the organization and this checkbox is selected, the cloud-based storage used to store Teams chat data for on-premises users is included in the scope of the search. For more information, see Search for Teams chat data for on-premises users.

7. On the **Define your search conditions** page, type a keyword query and add conditions to the search query if necessary.



a. Specify keywords, message properties such as sent and received dates, or document properties such as file names or the date that a document was last changed. You can use more complex queries that use a Boolean operator, such as **AND**, **OR**, **NOT**, and **NEAR**. If you leave the keyword box empty, all content located in the specified content locations is

included in the search results. For more information, see Keyword queries and search conditions for eDiscovery.

b. Alternatively, you can select the **Show keyword list** checkbox and the type a keyword in each row. If you do this, the keywords on each row are connected by a logical operator (**c:s**) that is similar in functionality to the **OR** operator in the search query that's created.

   Why use the keyword list? You can get statistics that show how many items match each keyword. This can help you quickly identify which keywords are the most (and least) effective. You can also use a keyword phrase (surrounded by parentheses) in a row. For more information about the keyword list and search statistics, see Get keyword statistics for searches.

   > **Note**
   >
   > To help reduce issues caused by large keyword lists, you're limited to a maximum of 20 rows in the keyword list.

c. You can add search conditions to narrow a search and return a more refined set of results. Each condition adds a clause to the search query that is created and run when you start the search. A condition is logically connected to the keyword query (specified in the keyword box) by a logical operator (**c:c**) that is similar in functionality to the **AND** operator. That means that items have to satisfy both the keyword query and one or more conditions to be included in the results. This is how conditions help to narrow your results. For a list and description of conditions that you can use in a search query, see Search conditions.

8. Review the search settings (and edit if necessary), and then submit the search to start it.

After the search is completed, you can preview the search results. If necessary, select **Refresh** on the **Searches** page to display the search you created.

More information about searching content locations

- When you select **Choose users, groups, or teams** to specify mailboxes to search, the mailbox picker that's displayed is empty. This is by design to enhance performance. To add recipients to this list, select **Choose users, groups, or teams**, type a name (a minimum of three characters) in the search box, select the check box next to the name, and then select **Choose**.

- You can add inactive mailboxes, Microsoft Teams, Viva Engage Groups, Office 365 Groups, and distribution groups to the list of mailboxes to search. Dynamic distribution groups aren't supported. If you add Microsoft Teams, Viva Engage Groups, or Office 365 Groups, the group or team mailbox is searched; the mailboxes of the group members aren't searched.
- To add sites to the search, turn on the toggle and then select **Choose sites**. Type the URL for each site that you want to search. You can also add the URL for the SharePoint site for a Microsoft Team, a Viva Engage Group, or an Office 365 Group.

Last updated on 09/02/2025

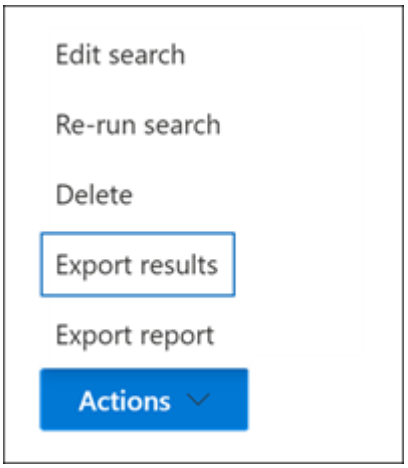
Export content from an eDiscovery (Standard) case

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

After a search associated with a Microsoft Purview eDiscovery (Standard) case is successfully run, you can export the search results. When you export search results, mailbox items are downloaded in PST files or as individual messages. When you export content from SharePoint and OneDrive sites, copies of native Office documents and other documents are exported. A Results.csv file that contains information about every item that's exported and a manifest file (in XML format) that contains information about every search result is also exported.

Export search results

1. Go to the Microsoft Purview portal and sign in using the credentials for user account that has been assigned the appropriate eDiscovery permissions.
2. In the left navigation pane of the Microsoft Purview portal, select **Show all**, and then select **eDiscovery** > **eDiscovery (Standard)**.
3. On the **eDiscovery (Standard)** page, select the name of the case that you want to create the hold in.
4. On the **Home** page for the case, select the **Searches** tab.
5. On the **Actions** menu at the bottom of the flyout page, select **Export results**.

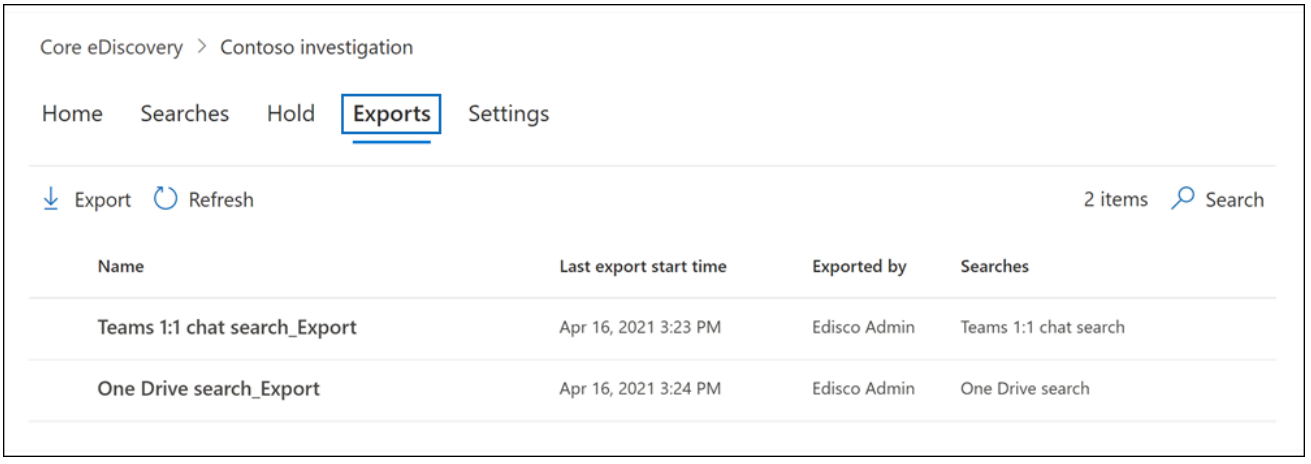


The workflow to export the results of a search associated with an eDiscovery (Standard) case is that same as exporting the search results for a search on the **Content search** page. For step-by-step instructions, see Export content search results.

> **Note**
>
> When you export search results, you have the option to enable deduplication so that only one copy of an email message is exported even though multiple instances of the same message might have been found in the mailboxes that were searched.

After you start the export, the search results are prepared for downloading, which means they're transferred to a Microsoft-provided Azure Storage location in the Microsoft cloud.

6. Select the **Exports** tab in the case to display the list of export jobs.



You may have to select **Refresh** to update the list of export jobs so that it shows the export job you created. Export jobs have the same name as the corresponding search with **_Export** appended to the search name.

7. Select the export job you created to display status information on the flyout page. This information includes the percentage of items that have been transferred to the Azure Storage location.

8. After all items have been transferred, select **Download results** to download the search results to your local computer. For more information downloading search results, see Step 2 in Export content search results

> **Note**
>
> The exported search results must be downloaded within 14 days after you created the export job.

More information about exporting searches from a case

- If you restart the export, any changes to the queries of the searches that make up the export job won't affect the search results that are retrieved. When you restart an export, the same combined search query job that was run when the export job was created will be run again.
- Also, if you restart an export, the search results that are copied to the Azure Storage location overwrites the previous results. The previous results that were copied won't be available to be downloaded.

Last updated on 11/21/2025

Close, reopen, or delete eDiscovery (Standard) cases

> **⊗ Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic **Content Search**, classic **eDiscovery (Standard)**, and classic **eDiscovery (Premium)**.
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the **new eDiscovery experience** in the **Microsoft Purview portal**.

This article describes how to close, reopen, delete, or upgrade a Microsoft Purview eDiscovery (Standard) cases in Microsoft Purview.

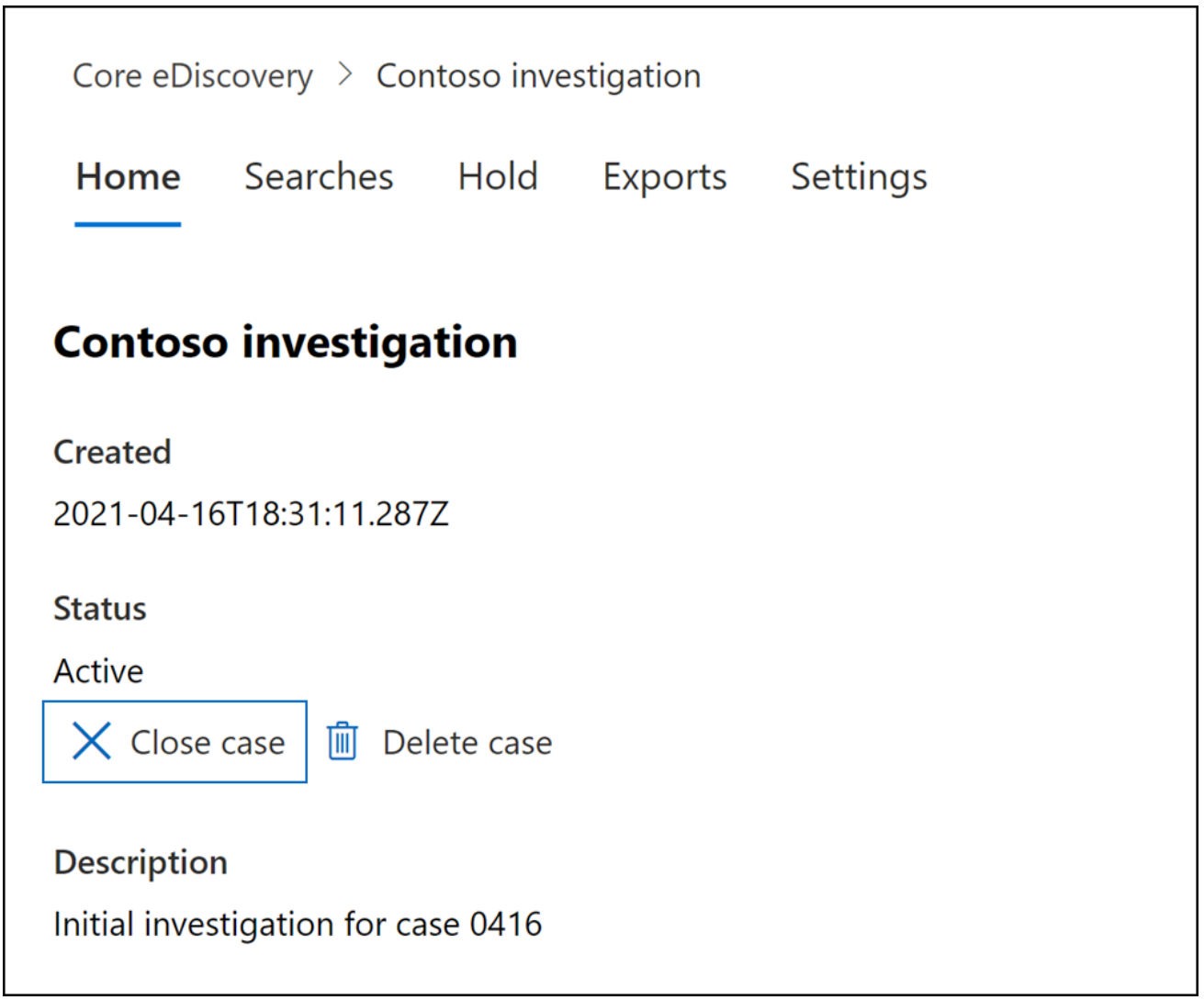
Close a case

When the legal case or investigation supported by an eDiscovery (Standard) case is completed, you can close the case. Here's what happens when you close a case:

- If the case contains any eDiscovery holds, they'll be turned off. After the hold is turned off, a 30-day grace period (called a *delay hold*) is applied to content locations that were on hold. This helps prevent content from being immediately deleted and provides admins the opportunity to search for and restore content before it may be permanently deleted after the delay hold period expires. For more information, see Removing content locations from an eDiscovery hold.
- Closing a case only turns off the holds that are associated with that case. If other holds are placed on a content location (such as a Litigation Hold, a retention policy, or a hold from a different eDiscovery (Standard) case) those holds will still be maintained.
- The case is still listed on the eDiscovery (Standard) page in the Microsoft Purview portal. The details, holds, searches, and members of a closed case are retained.
- You can edit a case after it's closed. For example, you can add or remove members, create searches, and export search results. The primary difference between active and closed cases is that eDiscovery holds are turned off when a case is closed.

To close a case:

1. In the Microsoft Purview portal, select **eDiscovery** > **eDiscovery (Standard)** to display the list of eDiscovery (Standard) cases in your organization.
2. Select the name of the case that you want to close.



3. On the home page, under **Status**, select **Close case**.

   A warning is displayed saying that the holds associated with the case will be turned off.

4. Select **Yes** to close the case.

   The status on the case home page is changed from **Active** to **Closing**.

5. On the **eDiscovery (Standard)** page, select **Refresh** to update the status of the closed case. It might take up to 60 minutes for the closing process to complete.

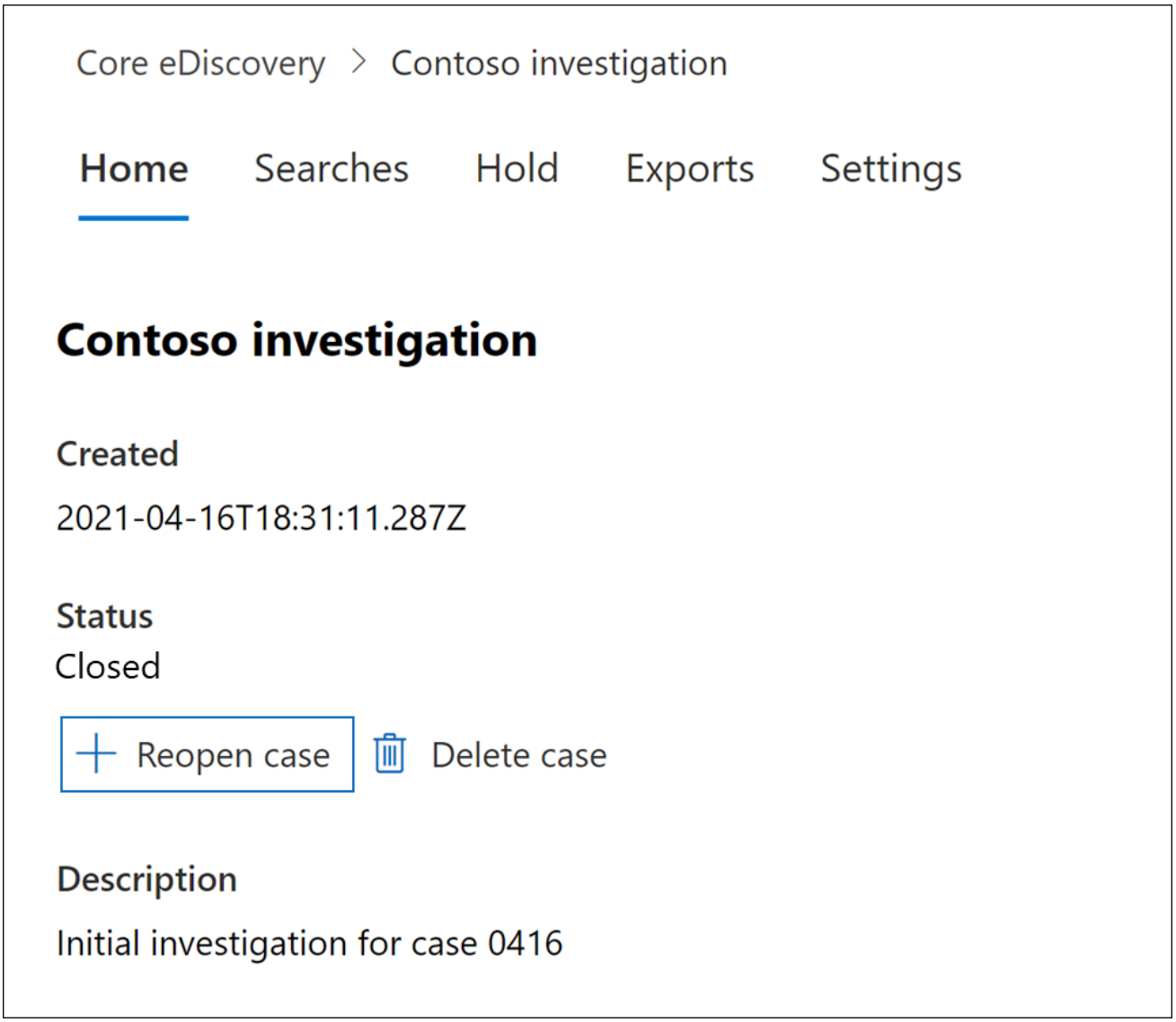
   When the process is complete, the status of the case is changed to **Closed** on the **eDiscovery (Standard)** page.

Reopen a closed case

When you reopen a case, any eDiscovery holds that were in place when the case was closed won't be automatically reinstated. After the case is reopened, you'll have to go to the **Holds** page and turn on the previous holds. To turn on a hold, select it to display the flyout page, and then set the **Status** toggle to **On**.

1. In the Microsoft Purview portal, select **eDiscovery (Standard)** to display the list of eDiscovery (Standard) cases in your organization.

2. Select the name of the case that you want to reopen.



3. On the home page, under **Status**, select **Reopen case**.

   A warning is displayed saying that the holds that were associated with the case when it was closed won't be turned on automatically.

4. Select **Yes** to reopen the case.

The status on the case home page flyout page is changed from **Closed** to **Active**.

5. On the **eDiscovery (Standard)** page, select **Refresh** to update the status of the reopened case. It might take up to 60 minutes for the reopening process to complete.

   When the process is complete, the status of the case is changed to **Active** on the **eDiscovery (Standard)** page.

6. (Optional) To turn on any holds associated with the reopened case, go to **Holds** tab, select a hold, and then select the checkbox under **Status** on the hold flyout page.

Delete a case

You can also delete active and closed eDiscovery (Standard) cases. When you delete a case, all searches and exports in the case are deleted, and the case is removed from the list of cases on the **eDiscovery (Standard)** page in the Microsoft Purview portal. You can't reopen a deleted case.

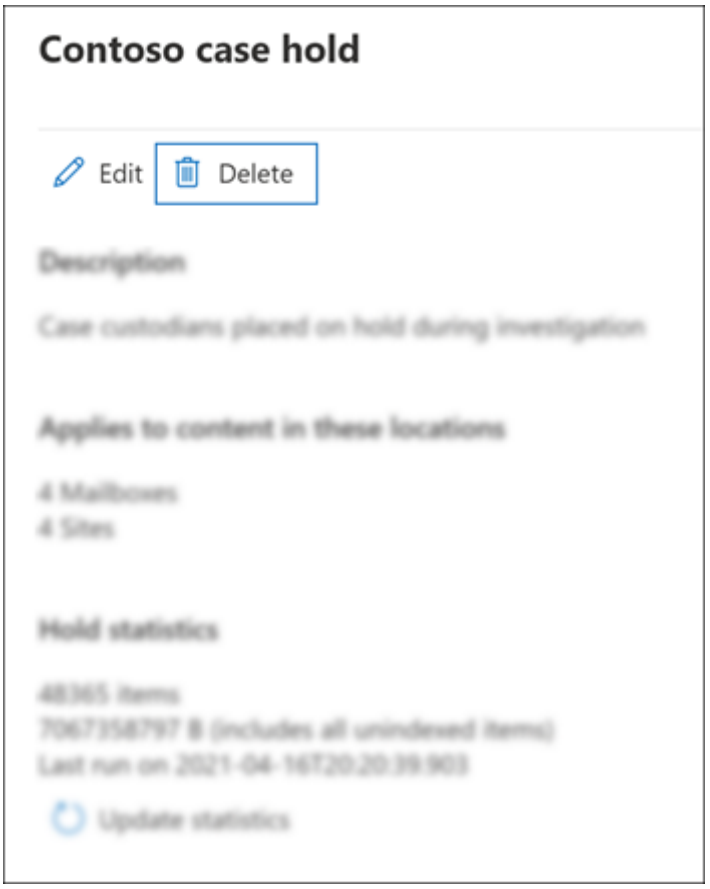
Before you can delete a case (whether it's active or closed), you must first delete *all* eDiscovery holds associated with the case. That includes deleting holds with a status of **Off**.

To delete an eDiscovery hold:

1. Go to the **Holds** tab in the case that you want to delete.

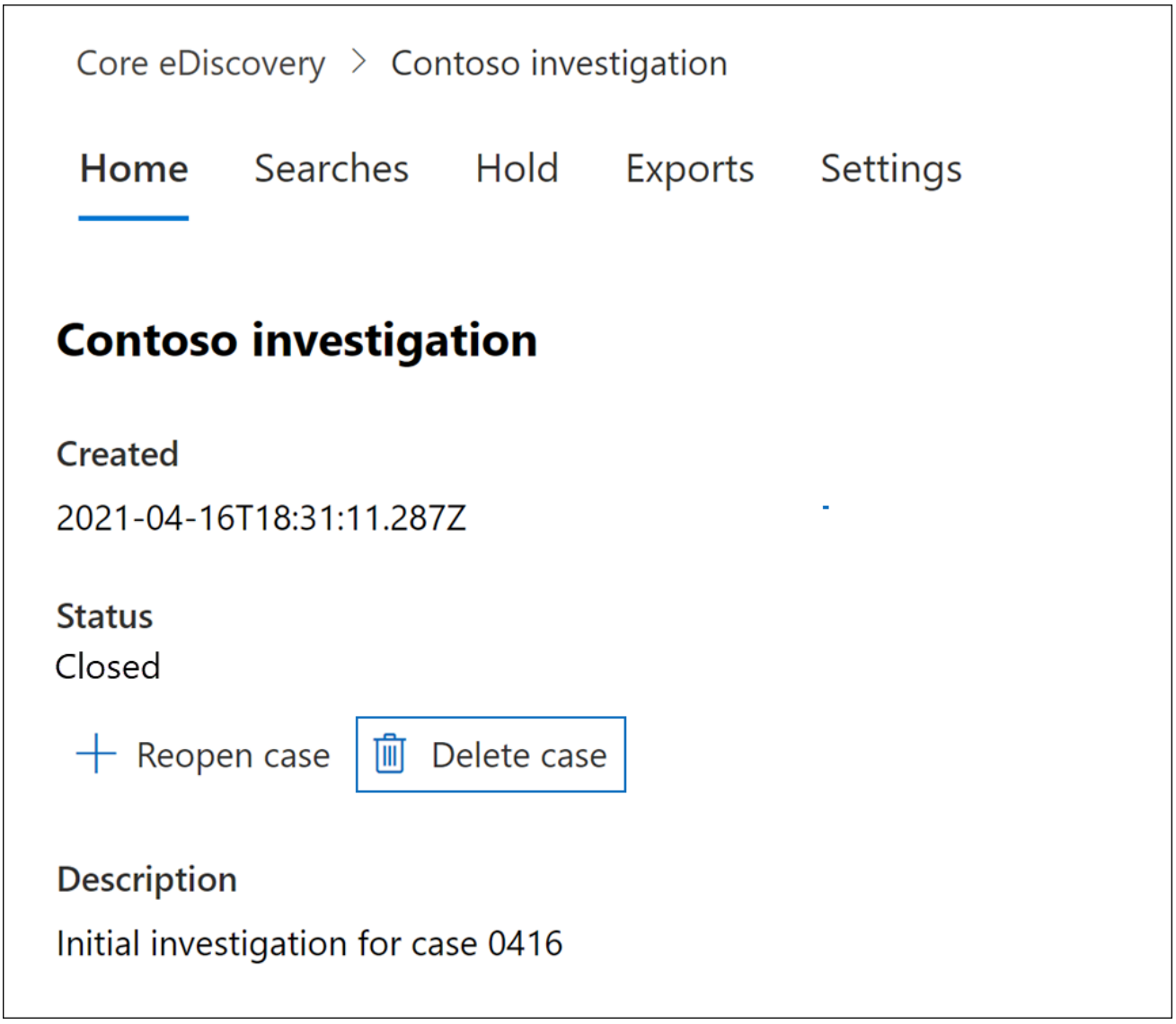
2. Select the hold that you want to delete.

3. On the flyout page, select **Delete**.



To delete a case:

1. In the Microsoft Purview portal, select **eDiscovery** > **eDiscovery (Standard)** to display the list of eDiscovery (Standard) cases in your organization.
2. Select the name of the case that you want to delete.
3. On the case home page, under **Status**, select **Delete case**.



If the case you're trying to delete still contains eDiscovery holds, you'll receive an error message. You'll have to delete all holds associated with the case and then try again to delete the case.

Last updated on 11/21/2025

Overview of Microsoft Purview eDiscovery (Premium)

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic **Content Search**, classic **eDiscovery (Standard)**, and classic **eDiscovery (Premium)**.
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the **new eDiscovery experience** in the **Microsoft Purview portal**.

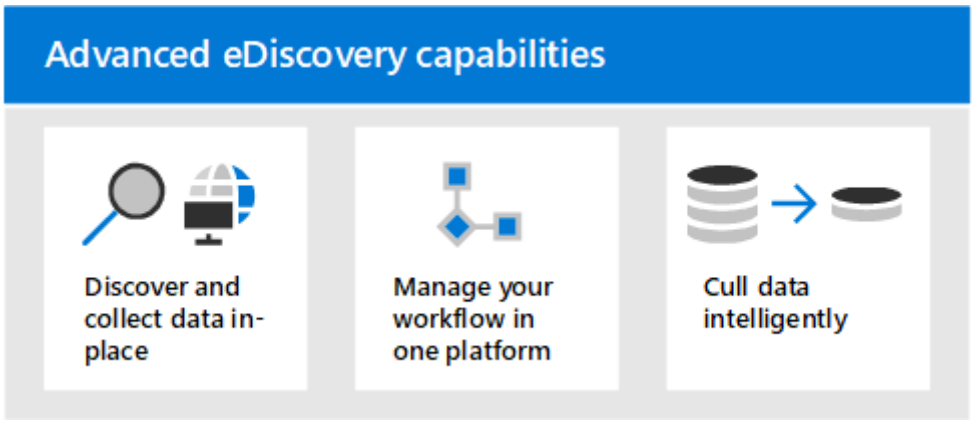
The Microsoft Purview eDiscovery (Premium) solution builds on the existing Microsoft eDiscovery (Standard) and analytics capabilities. eDiscovery (Premium) provides an end-to-end workflow to preserve, collect, analyze, review, and export content that's responsive to your organization's internal and external investigations across Microsoft 365 and Office 365.

You can also use eDiscovery (Premium) to place an eDiscovery hold on content locations, such as Exchange mailboxes, SharePoint sites, OneDrive accounts, and Microsoft Teams. It also lets legal teams manage the entire legal hold notification workflow to communicate with custodians involved in a case.

eDiscovery (Premium) capabilities

eDiscovery (Premium) can help your organization respond to legal matters or internal investigations by discovering data where it lives. You can seamlessly manage eDiscovery workflows by identifying persons of interest and their data sources, seamlessly apply holds to preserve data, and then manage the legal hold communication process. By collecting data from the source, you can search the live Microsoft 365 platform to quickly find what you need. Intelligent, machine learning capabilities such as deep indexing, email threading, and near duplicate detection also help you reduce large volumes of data to a relevant data set.

The following sections describe how these eDiscovery (Premium) capabilities can help your organization.



Discover and collect data in-place

Traditionally, organizations that rely on multiple third-party eDiscovery solutions require copying large volumes of data out of Microsoft 365 to process and having to host duplicate data. This necessity increases the time to find relevant data and the risk, cost, and complexity of managing multiple solutions.

eDiscovery (Premium) in Microsoft 365 lets you discover data at the source and staying within your Microsoft 365 security and compliance boundary. By collecting data in-place from the live system, eDiscovery (Premium) reduces the friction of going back to the source and reduces unnecessary work of having to find missing content, which often happens when journaling lags in traditional eDiscovery solutions.

Native search and collection capabilities for data in Teams, Viva Engage, SharePoint Online, OneDrive for Business, and Exchange Online further enhances data discovery. For example, eDiscovery (Premium):

- Reconstructs Teams conversations (instead of returning individual messages from conversations).
- Collects cloud-based content shared with users by use of links or modern attachments in email message and Teams chats.
- Has built-in support for hundreds of non-Microsoft 365 file types.
- Collects data from third-party sources (such as Bloomberg, Facebook, Slack, and Zoom Meetings) that's imported and archived in Microsoft 365 by data connectors.

Manage eDiscovery workflow in one platform

eDiscovery (Premium) can help you reduce the number of eDiscovery solutions you need to rely on. It provides a streamlined, end-to-end workflow, all which occurs within Microsoft 365. eDiscovery (Premium) helps reduce the friction of identifying and collecting potential sources of relevant information by automatically mapping unique and shared data sources to the person of

interest (known as a *custodian*), and by providing reporting and analytics on potentially relevant data prior to collecting it for analysis and review.

Additionally, Microsoft Graph APIs can help you automate the eDiscovery workflow and extend eDiscovery (Premium) for custom solutions.

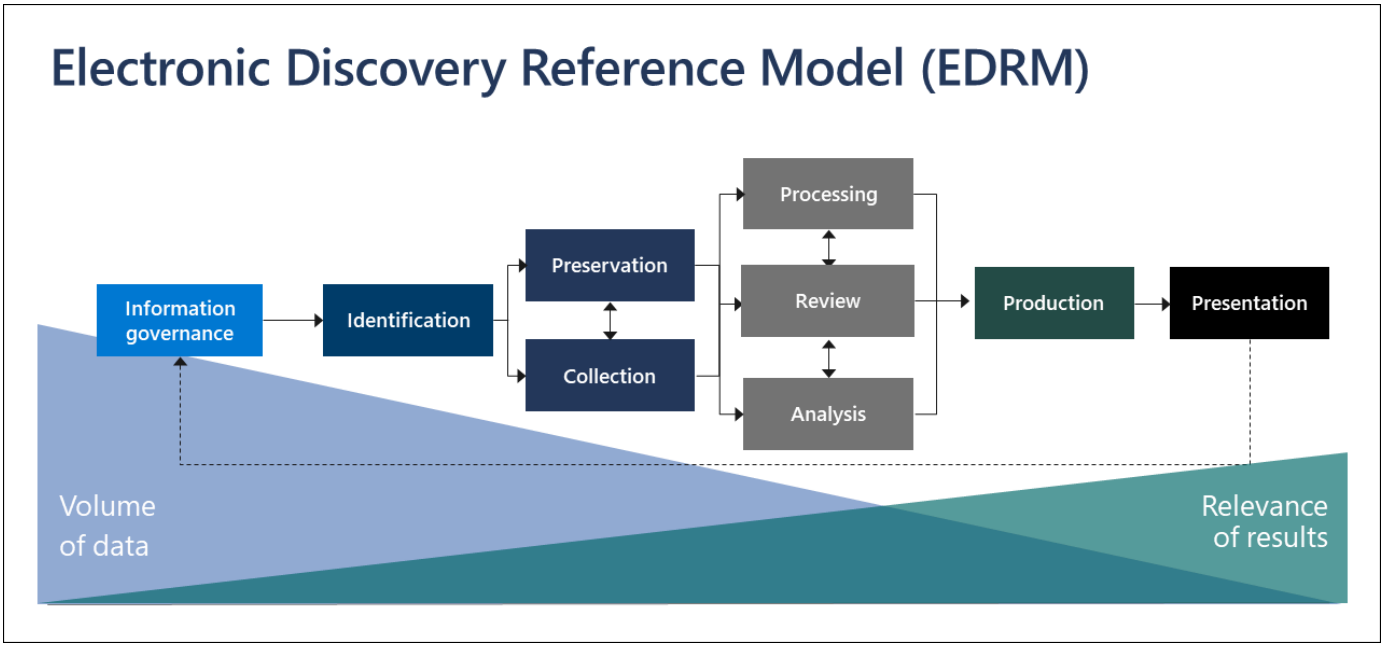
Cull data intelligently

Intelligent, machine learning capabilities in eDiscovery (Premium) help you reduce the amount of data to review. These intelligent capabilities help you reduce and cull large volumes of data to a relevant set. For example, a built-in review set query helps filter only for unique content by identifying near duplicates. This capability can substantially reduce the amount of data to review.

Additional machine learning capabilities can further refine and identify relevant data using smart tags and technology assisted review tools like the Relevance modules.

eDiscovery (Premium) alignment with the Electronic Discovery Reference Model

The built-in workflow of eDiscovery (Premium) in Microsoft 365 aligns with the eDiscovery process outlined by the Electronic Discovery Reference Model (EDRM).



(Image based on the EDRM model on edrm.net)

At a high level, here's how eDiscovery (Premium) supports the EDRM workflow:

- **Identification.** After you identify potential persons of interest in an investigation, you can add them as custodians (also called *data custodians*, because they may possess information that's relevant to the investigation) to an eDiscovery (Premium) case. After users are added as custodians, it's easy to preserve, collect, and review custodian documents.
- **Preservation.** To preserve and protect data that's relevant to an investigation, eDiscovery (Premium) lets you place a legal hold on the data sources associated with the custodians in a case. You can also place non-custodial data on hold. eDiscovery (Premium) also has a built-in communications workflow so you can send legal hold notifications to custodians and track their acknowledgments.
- **Collection.** After you identified (and preserved) the data sources relevant to the investigation, you can use the built-in search tool in eDiscovery (Premium) search for and collect live data from the custodial data sources (and non-custodial data sources, if applicable) that may be relevant to the case.
- **Processing.** After you've collected all data relevant to the case, the next step is process it for further review and analysis. In eDiscovery (Premium), the in-place data that you identified in the collection phase is copied to an Azure Storage location (called a *review set*), which provides you with a static view of the case data.
- **Review.** After data has been added to a review set, you can view specific documents and run additional queries to reduce the data to what is most relevant to the case. Also, can annotate and tag specific documents.
- **Analysis.** eDiscovery (Premium) provides integrated analytics tool that helps you further cull data from the review set that you determine isn't relevant to the investigation. In addition to reducing the volume of relevant data, Advance eDiscovery also helps you save legal review costs by letting you organize content to make the review process easier and more efficient.
- **Production** and **Presentation.** When you're ready, you can export documents from a review set for legal review. You can export documents in their native format or in an EDRM-specified format so they can be imported into third-party review applications.

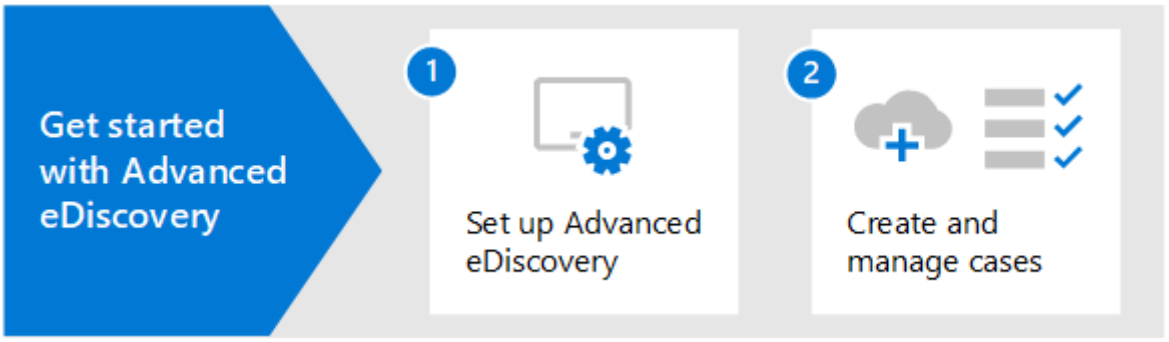
Subscriptions and licensing

Microsoft Purview eDiscovery capabilities are included with Microsoft Purview. The licensing requirements may vary even within capabilities, depending on configuration options. For licensing requirements, guidance, and options, see the subscription requirements for eDiscovery (Standard) and eDiscovery (Premium).

For information about how to assign licenses, see Assign licenses to users.

Get started with eDiscovery (Premium)

There are two quick and easy steps to get started with eDiscovery (Premium).



Expand table

Steps	Description
Set up eDiscovery (Premium)	After verifying the subscription and licensing requirements, you can assign permissions and configure organization-wide settings to get started using eDiscovery (Premium).
Create and manage cases	Create cases to manage the eDiscovery (Premium) workflow for all legal and other types of investigations in your organization.

Integration with Insider Risk Management

Cases in Microsoft Purview Insider Risk Management can be quickly escalated to new cases in Microsoft Purview eDiscovery (Premium) when additional legal review is needed for potentially risky user activity. The tight integration between these solutions can help your risk and legal teams work more efficiently and can help provide a complete end-to-end view of user activities under review. Check out how to get started with Insider Risk Management and how to easily escalate an Insider Risk Management case to an eDiscovery (Premium) case.

Training

Training your IT administrators, eDiscovery managers, and compliance investigation teams in the basics for eDiscovery (Premium) can help your organization get started more quickly using Microsoft 365 eDiscovery tools. Microsoft 365 provides the following resource to help these users in your organization getting started with eDiscovery: Describe the eDiscovery and audit capabilities of Microsoft 365.

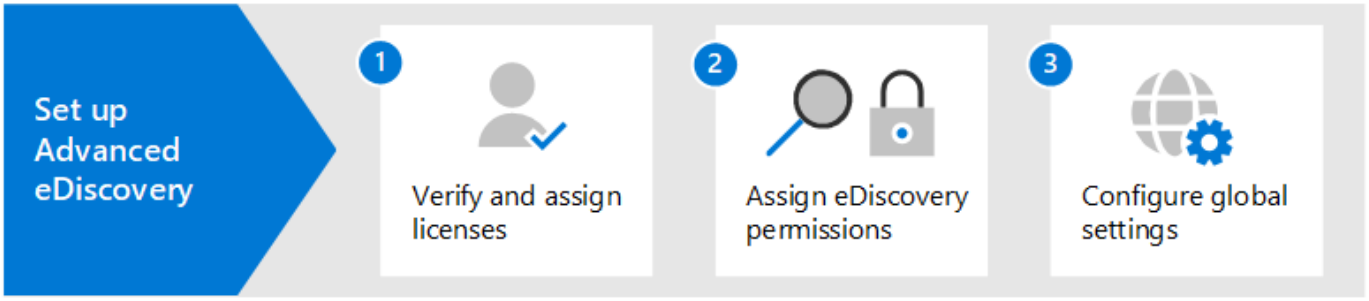
Last updated on 11/21/2025

Get started with eDiscovery (Premium)

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

Microsoft Purview eDiscovery (Premium) provides an end-to-end workflow to preserve, collect, review, analyze, and export data that's responsive to your organization's internal and external investigations. Nothing is needed to deploy eDiscovery (Premium), but there are some prerequisite tasks that an IT admin and eDiscovery manager have to complete before your organization can start to create and use eDiscovery (Premium) cases to manage your investigations.

This article discusses the following steps necessary to set up eDiscovery (Premium).



This includes ensuring the proper licensing required to access eDiscovery (Premium) and add custodians to cases, and assigning permissions to your legal and investigation team so they can access and manage cases.

Step 1: Verify and assign appropriate licenses

Microsoft Purview eDiscovery capabilities are included with Microsoft Purview. The licensing requirements may vary even within capabilities, depending on configuration options. For licensing requirements, guidance, and options, see the subscription requirements for eDiscovery (Premium).

Step 2: Assign eDiscovery permissions

To access eDiscovery (Premium) or added as a member of an eDiscovery (Premium) case, a user must be assigned the appropriate permissions. Specifically, a user must be added as a member of the eDiscovery Manager role group in the Microsoft Purview portal. Members of this role group can create and manage eDiscovery (Premium) cases. They can add and remove members, place custodians and content locations on hold, manage legal hold notifications, create and edit searches associated in a case, add search results to a review set, analyze data in a review set, and export and download from an eDiscovery (Premium) case.

For step-by-step guidance to configure permissions for eDiscovery, see Assign eDiscovery permissions.

Step 3: Configure global settings for eDiscovery (Premium)

The last step to complete before people in your organization start to create and use cases is to configure global settings that apply to all cases in your organization. At this time, the only global setting is *attorney-client privilege detection* (more global settings will be available in the future). This setting enables the attorney-client privilege model to run when you analyze data in a review set. The model uses machine learning to determine the likelihood that a document contains content that is legal in nature. It also compares the participants of documents with an attorney list (that you submit when setting up the model) to determine if a document has at least one participant who is an attorney.

For more information about setting up and using the attorney-client privilege detection model, see Set up attorney-client privilege detection in eDiscovery (Premium).

> **Note**
>
> This is an optional step that you can perform anytime. Not implementing the attorney-client privilege detection model doesn't prevent you from creating and using eDiscovery (Premium) cases.

Step 4: Verify that required eDiscovery apps are enabled

eDiscovery (Premium) requires the following Enterprise apps to be enabled in your Microsoft 365 or Office 365 organization. If these apps aren't enabled, you won't be able to access all eDiscovery (Premium) view, filter, and search features.

Expand table

App	App ID
Built-in security add-on for on-premises mailboxes (formerly Exchange Online Protection)	00000007-0000-0ff1-ce00-000000000000
ComplianceWorkbenchApp	92876b03-76a3-4da8-ad6a-0511ffdf8647
MicrosoftPurviewEDiscovery	b26e684c-5068-4120-a679-64a5d2c909d9
Office365Zoom	0d38933a-0bbd-41ca-9ebd-28c4b5ba7cb7

> **Note**
>
> The *MicrosoftPurviewEDiscovery* app provides support for exporting data via scripts in eDiscovery (Premium). For more information about exporting data via scripts, see Get caseOperation.

For more information about how to view and enable apps, see:

- Quickstart: View enterprise applications
- Quickstart: Add an enterprise application

Next steps

After you set up eDiscovery (Premium), you're ready to create a case.

Last updated on 02/18/2026

eDiscovery (Premium) workflow for content in Microsoft Teams

> **⊗ Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

This article provides a comprehensive set of procedures, guidelines, and best practices for using Microsoft Purview eDiscovery (Premium) to preserve, collect, review, and export content from Microsoft Teams. The goal of this article is to help you optimize your eDiscovery workflow for Teams content.

There are seven categories of Teams content that you can collect and process using eDiscovery (Premium):

- **Teams 1:1 chats**. Chat messages, posts, and attachments shared in a Teams conversation between two people. Teams 1:1 chats are also called *conversations*.
- **Teams group chats**. Chat messages, posts, and attachments shared in a Teams conversation between three or more people. Also called *1:N* chats or *group conversations*.
- **Teams reactions**. Reactions applied to chat messages, posts, and attachments in a Teams conversation.
- **Teams channels**. Chat messages, posts, replies, and attachments shared in a standard Teams channel.
- **Teams meetings**. Audio and transcripts from recorded Teams meetings.
- **Private channels**. Message posts, replies, and attachments shared in a private Teams channel.
- **Shared channels**. Message posts, replies, and attachments shared in a shared Teams channel.

Where Teams content is stored

A prerequisite to managing Teams content in eDiscovery (Premium) is to understand the type of Teams content that you can collect, process, and review in eDiscovery (Premium) and where that content is stored in Microsoft 365.

Teams data is stored in Azure Cosmos DB. Teams compliance records captured by the substrate are in Exchange Online and are available for ediscovery.

The following table lists Teams content type and where each is stored for compliance purposes. The data stored in Exchange online is hidden from clients. eDiscovery never operates against the real Teams message data, which remains in Azure Cosmos DB.

Expand table

Teams category	Location of chat messages/posts	Location of files/attachments	Location of meeting recordings
Teams 1:1 chats	Messages in 1:1 chats are stored in the Exchange Online mailbox of all chat participants.	Files shared in a 1:1 chat are stored in the OneDrive for Business account of the person who shared the file.	N/A
Teams group chats	Messages in group chats are stored in the Exchange Online mailbox of all chat participants.	Files shared in group chats are stored in the OneDrive for Business account of the person who shared the file.	N/A
Teams reactions	Messages in group chats are stored in the Exchange Online mailbox of all chat participants.	Files shared in group chats are stored in the OneDrive for Business account of the person who shared the file.	N/A
Teams channels	All channel messages and posts are stored in the Exchange Online mailbox associated with the team.	Files shared in a channel are stored in the SharePoint Online site associated with the team.	N/A
Teams meetings	Chats in recorded meetings are stored in the OneDrive for Business account for the user recording the Teams meeting.	Files and attachments shared in recorded meetings are stored in the OneDrive for Business account for the user recording the Teams meeting.	Meeting recordings are stored in the OneDrive for Business account for the user recording the Teams meeting.
Private channels	Messages sent in a private channel are stored in the Exchange Online mailboxes of all members of the private channel.	Files shared in a private channel are stored in a dedicated SharePoint Online site associated with the private channel.	N/A

Teams category	Location of chat messages/posts	Location of files/attachments	Location of meeting recordings
Shared channels	Messages sent in a shared channel are stored in a system mailbox associated with the shared channel.[1]	Files shared in a shared channel are stored in a dedicated SharePoint Online site associated with the shared channel.	N/A

> **Note**
>
> [1] To search for (and preserve) messages sent in a shared channel, you have to search or specify the Exchange Online mailbox for the parent Team.

Create a case for Teams content

The first step to managing Teams content in eDiscovery (Premium) is to create a case using the new case format that's optimized for managing Teams content. The new case format helps accommodate significant increases in case size, both for total data volume and the total number of items in cases.

For step-by-step guidance on how to create a case, see Create and manage an eDiscovery (Premium) case.

Add Teams custodial data sources and preserve Teams content

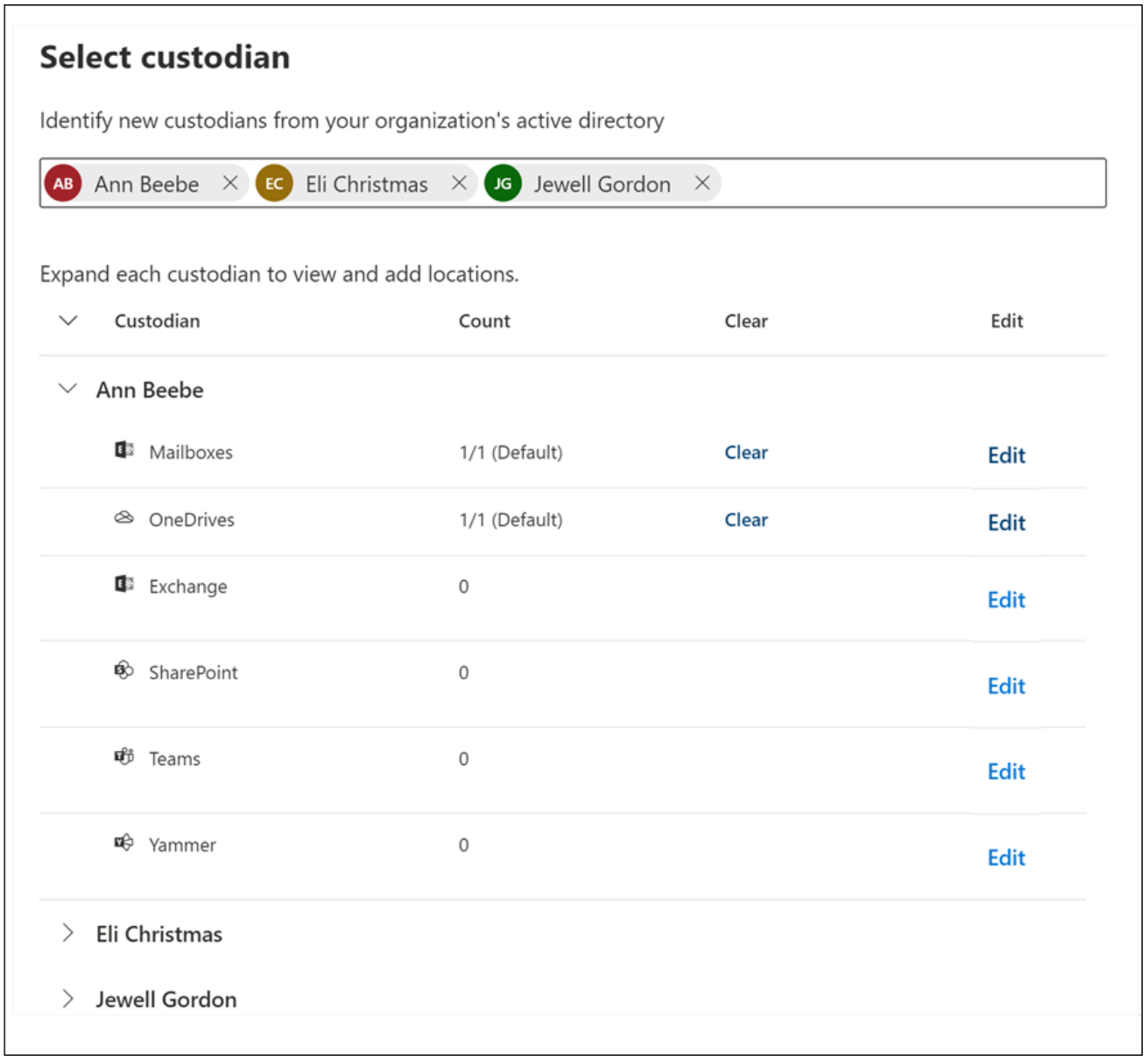
The next step is to identify the users who are the data custodians in your investigation and add them and their content locations as custodians to the case you created in the previous section. When you add custodians, you can specify their mailbox and OneDrive account as custodial data sources. You can also specify Teams content locations as custodian data sources to quickly place these locations on legal hold to preserve content during your investigation. It also makes it easy to collect content and add it to a review set.

To add custodians to a case and preserve custodial data sources:

1. Go to the eDiscovery (Premium) case that you created in the previous section, and then select **Data sources**.
2. On the **Data sources** page, select **Add data source** > **Add new custodians**.

3. In the **New custodian** workflow, add one or more users as custodians to the case by typing the first part of the user's name or alias. After you find the correct person, select their name to add them to the list.

4. Expand each custodian to view the primary data sources that have been automatically associated to the custodian, and to select other locations to associate to the custodian.



5. Follow these guidelines to add custodial data sources for Teams content. Select **Edit** to add a data location.

   - **Mailboxes**. The custodian's mailbox is selected by default. Keep this selected to add (and preserve) 1:1 chats, group chats, and private channel chats as custodial data.
   - **OneDrives**. The custodian's OneDrive account is selected by default. Keep this selected to add (and preserve) files shared in 1:1 chats, group chats, and Teams meetings as custodial data.

- **SharePoint**. Add the SharePoint site associated with any private or shared channel the custodian is a member of to add (and preserve) as custodial data the files shared in a channel. Select **Edit** and then add the URL for the SharePoint site associated with a private or shared channel. To learn how to locate the private and shared channels a user is a member of, see eDiscovery of private and shared channels.
- **Teams**. Add the teams that the custodian is a member of to add (and preserve) as custodial data all channel messages and all files shared to a Teams channel. This recommendation includes adding the mailbox for the parent team of a shared channel the custodian is a member of. When you select **Edit**, the mailbox and site associated with each team the custodian is a member of are displayed in a list. Select the teams that you want to associate to the custodian. You have to select both the corresponding mailbox and site for each team.

> ⓘ **Note**
>
> You can also add the mailbox and site of Teams that custodians aren't members of as a custodian data location. You do this by clicking **Edit** next to **Exchange** and **SharePoint** and then adding the mailbox and site associated with the team.

6. After you add custodians and configure the custodial data sources, select **Next** to display the **Hold settings** page. A list of the custodians is displayed and the checkbox in the **Hold** column is selected by default. This checkbox indicates that a hold will be placed on the data sources that you associated with each custodian. Leave these checkboxes selected to preserve this data.

7. On the **Hold settings** page, select **Next** to review the custodians settings. Select **Submit** to add the custodians to the case.

For more information about adding and preserving data sources in eDiscovery (Premium) cases, see:

- Add custodians to an eDiscovery (Premium) case
- Add non-custodial data sources to an eDiscovery (Premium) case

Collect Teams content and add to review set

After adding custodians to the case and preserving content in custodian data sources, the next step in the workflow is to search for Teams content that's relevant to your investigation and add it to a review set for further review and analysis. Though it's typical to collect Teams content

together with content from other Microsoft 365 services such as email in Exchange and documents in SharePoint, this section will specifically focus on collecting Teams content in a collection. You can create additional collections that collect non-Teams content to add to a review set.

When you collect Teams content for a case, there are two steps in the workflow:

1. **Create a collection estimate**. The first step is to create a *collection estimate*, which is an estimate of the items that match your search criteria. You can view information about the results that matched the search query, such as the total number and size of items found, the different data sources where they were found, and statistics about the search query. You can also preview a sample of items returned by the collection. Using these statistics, you can change the search query and rerun the collection estimate as many times as is necessary to narrow the results until you're satisfied you're collecting the content relevant to your case.
2. **Commit a collection estimate to a review set**. Once you're satisfied with the results of a collection estimate, you can commit the collection to a review set. When you commit a collection estimate, the items returned by the collection are added to a review set for review, analysis, and export.

You also have the option of not running a collection estimate and adding the collection results directly to a review set when you create and run the collection.

To create a collection of Teams content:

1. Go to the eDiscovery (Premium) case that you added the custodians to in the previous section, and then select **Collections**.

2. On the **Collections** page, select **New collection**.

3. On the **Name and description** page, enter a name (required) and description (optional) for the collection. Select **Next**.

4. On the **Custodial data sources** page, select **Select custodians** to select the custodians that you added to the case. The list of the case custodians is displayed on the **Select custodians** flyout page.

5. Select one or more custodians and then select **Add**. After you add specific custodians to the collection, a list of specific data sources for each custodian is displayed. These data sources are the ones that you configured when you added the custodian to the case. All custodian data sources are selected by default. This includes any Teams or channels that you associated with a custodian.

We recommend doing the following things when collecting Teams content:

- Remove custodians' OneDrive accounts from the collection scope (by unselecting the checkbox in the **Custodian's OneDrive** column for each custodian). This prevents the collection of duplicate files that were attached to 1:1 chats and group chats. Cloud attachments are automatically collected from each conversation found in the collection when you commit the collection estimate to the review set. By using this method (instead of searching OneDrive accounts as part of the collection), files attached to 1:1 and group chats are grouped in the conversation they were shared in.
- Unselect the checkbox in the **Additional site** column to remove the SharePoint sites containing files shared in private or shared channels. Doing this eliminates collecting duplicate files that were attached to private or shared channel conversations because these cloud attachments are automatically added to the review set when you commit the collection estimate and grouped in the conversations they were shared in.

6. If you previously followed the steps to add Teams content as custodian data sources, you can skip this step and select **Next**. Otherwise, on the **Non-custodial data sources** page, you can choose non-custodial data sources that contain Teams content that you may have added to the case to search in the collection.

7. If you previously followed the steps to add Teams content as custodian data sources, you can skip this step and select **Next**. Otherwise, on the **Additional locations** page, you can add other data sources to search in the collection. For example, you could add the mailbox and site for a team that wasn't added as a custodial or non-custodial data source. You can also select the **Shared Teams channels** option to include shared channels during tenant-wide searches. Otherwise, select **Next** and skip this step.

8. On the **Conditions** page, configure the search query to collect Teams content from the data sources that you specified on the previous pages. You can use various keywords and search conditions to narrow the scope of the collection. For more information, see Build search queries for collections.

    - To help ensure the most comprehensive collection of Teams chat conversations (including 1:1, group, and channel chats) use the **Type** filter condition and select the **Instant messages** option.
    - To help ensure you add recorded Teams meeting information to the collection, use the **File type** filter condition and include *.mp4* as contained value.
    - We also recommend including a date range or several keywords to narrow the scope of the collection to items relevant to your investigation.

9. On the **Review your collection** page, review the collection settings and select **Submit** to create a collection estimate or **Save and close** to save the collection settings to complete later.

When the process of adding the collection to the review set is completed, the **Status** value for the collection on the **Collections** tab is set to **Estimated**.

Commit a collection estimate to a review set

When you're satisfied with the items you've collected in a collection estimate and are ready to analyze, tag, and review them, you can commit a collection to a review set in the case.

For step-by-step guidance on how to commit a collection, see Commit a collection estimate to a review set in eDiscovery (Premium).

Review Teams content in a review set

After you add collections of Teams content to a review set, the next step is to review the content for its relevance to your investigation and cull it if necessary. An important prerequisite to reviewing Teams content is understanding how eDiscovery (Premium) processes Teams meeting, chat conversations, and attachments when adding them to a review set. This processing of Teams content results in the following three things:

- **Grouping**. How messages, posts, and replies Teams conversations are grouped together and presented in the review set. This also includes attachments in chat conversations are extracted and group within the conversation.
- **Conversation transcript threading**. How eDiscovery (Premium) determines what additional content from a conversation to collect to provide context around items that matched the collection criteria.
- **Deduplication**. How eDiscovery (Premium) handles duplicate Teams content.
- **Metadata**. Metadata properties that eDiscovery (Premium) adds to Teams content after it's collected and added to a review set.

> ⓘ **Note**
>
> The date and time for all conversations in Teams are shown in Coordinated Universal Time (UTC).

Understand grouping, conversation threading, deduplication, and Teams metadata will help you optimize the review and analysis of Teams content. This section also has tips for viewing Teams content in a review set.

Grouping

When content from Teams chat conversations is added to a review set, messages, posts, and replies from conversations are aggregated in HTML transcript files. A single chat conversation can have multiple transcript files. An important function of these transcript files is to present Teams content as continuous conversations and not as individual (or separate) messages. This helps provides context for items that matched the search criteria of your collections in the previous step and reduce the number of items collected into the review set. Transcripts and associated items can be grouped by either *family* or *conversation*. Items in the same family will have the same value for the **FamilyId** metadata property. Items in the same conversation will have the same value for the **ConversationId** metadata property.

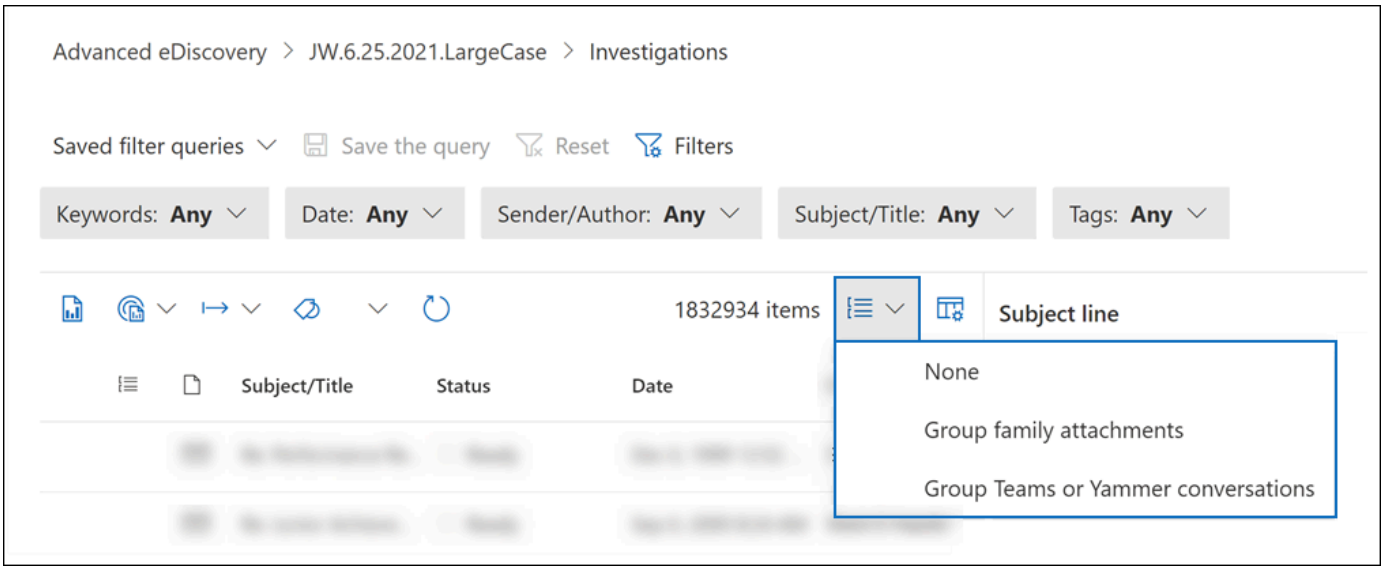
The following table describes how the different types of Teams content are grouped by family and conversation.

Expand table

Teams content type	Group by family	Group by conversation
Teams 1:1 and group chats	A transcript and all of its attachments and extracted items share the same **FamilyId**. Each transcript has a unique **FamilyId**.	All transcript files and their family items within the same conversation share the same **ConversationId**. This includes the following items: • All extracted items and attachments of all transcripts that share the same **ConversationId**. • All transcripts for the same chat conversation • All custodian copies of each transcript • Transcripts from subsequent collections from the same chat conversation For Teams 1:1 and group chat conversations, you might have multiple transcript files, each one corresponding to a different time frame within the conversation. Because these transcript files are from the same conversation with the same participants, they share the same **ConversationId**.
Standard, private, and	Each post and all replies and attachments are saved to its	Each post and its attachments and extracted items have a unique **ConversationId**. If there are subsequent

Teams content type	Group by family	Group by conversation
shared channel chats	own transcript. This transcript and all of its attachments and extracted items share the same **FamilyId**.	collections or new replies from the same post, the delta transcripts resulting from those collections will also have the same **ConversationId**.
Teams meetings	Each meeting	

Use the **Group** control in the command bar of a review set to view Teams content grouped by family or conversation.



- Select **Group family attachments** to view Teams content grouped by family. Each transcript file is displayed on a line in the list of review set items. Attachments are nested under the item.
- Select **Group Teams or Viva Engage conversations** to view Teams content grouped by conversation. Each conversation is displayed on a line in the list of review set items. Transcript files and attachments are nested under the top-level conversation.

> **Note**
>
> Cloud attachments are grouped with the conversations they appear in. This grouping is accomplished by assigning the same **FamilyId** as the transcript file of the message the file was attached to and the same **ConversationId** as the conversation the message appeared in. This means multiple copies of cloud attachments may be added to the review set if they were attached to different conversations.

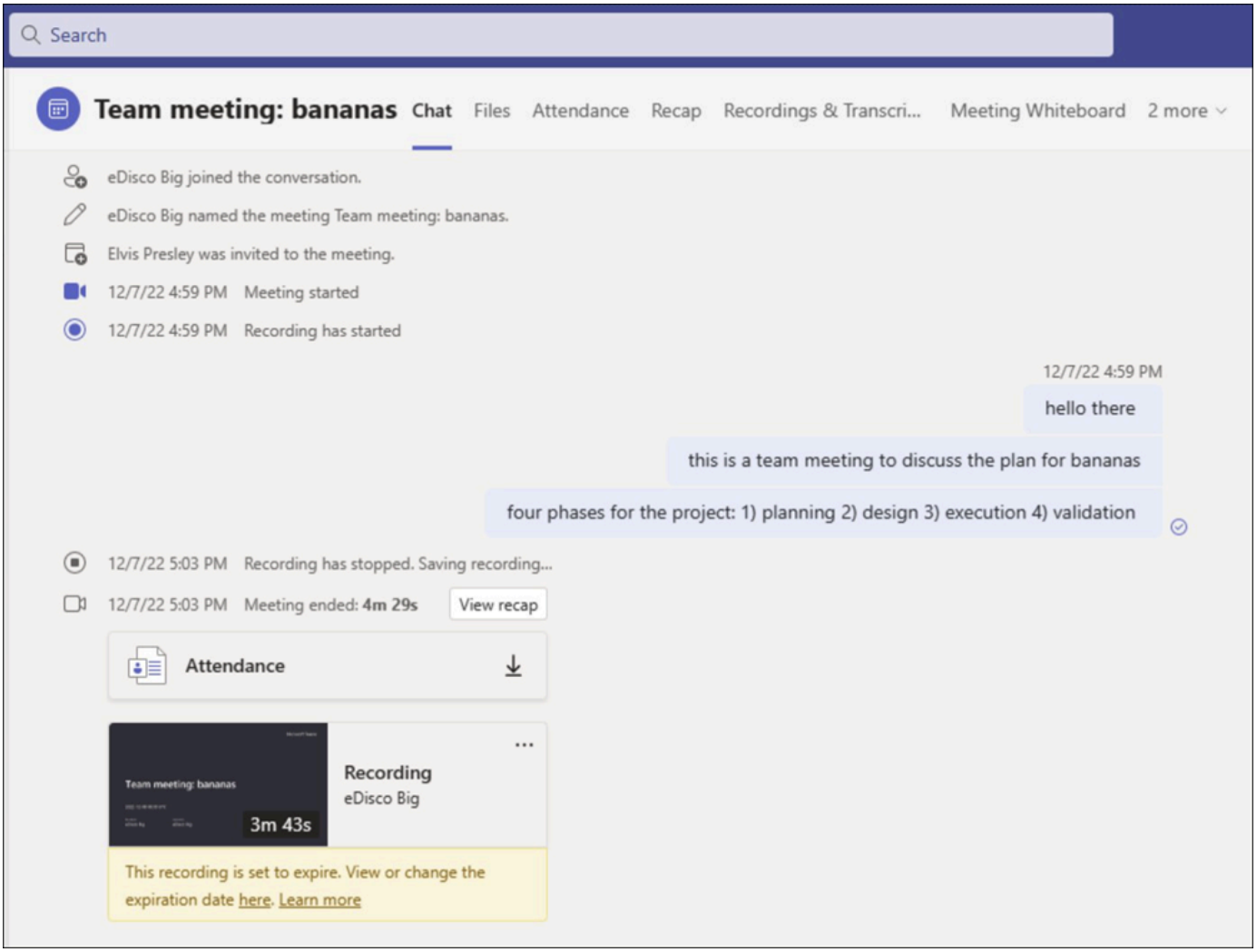
Viewing recorded meeting transcripts

The transcript of audio of the recorded meeting is captured as a separate file and indexed automatically for search. Recorded meetings in a review set are stored as a .zip file that contains the following files:

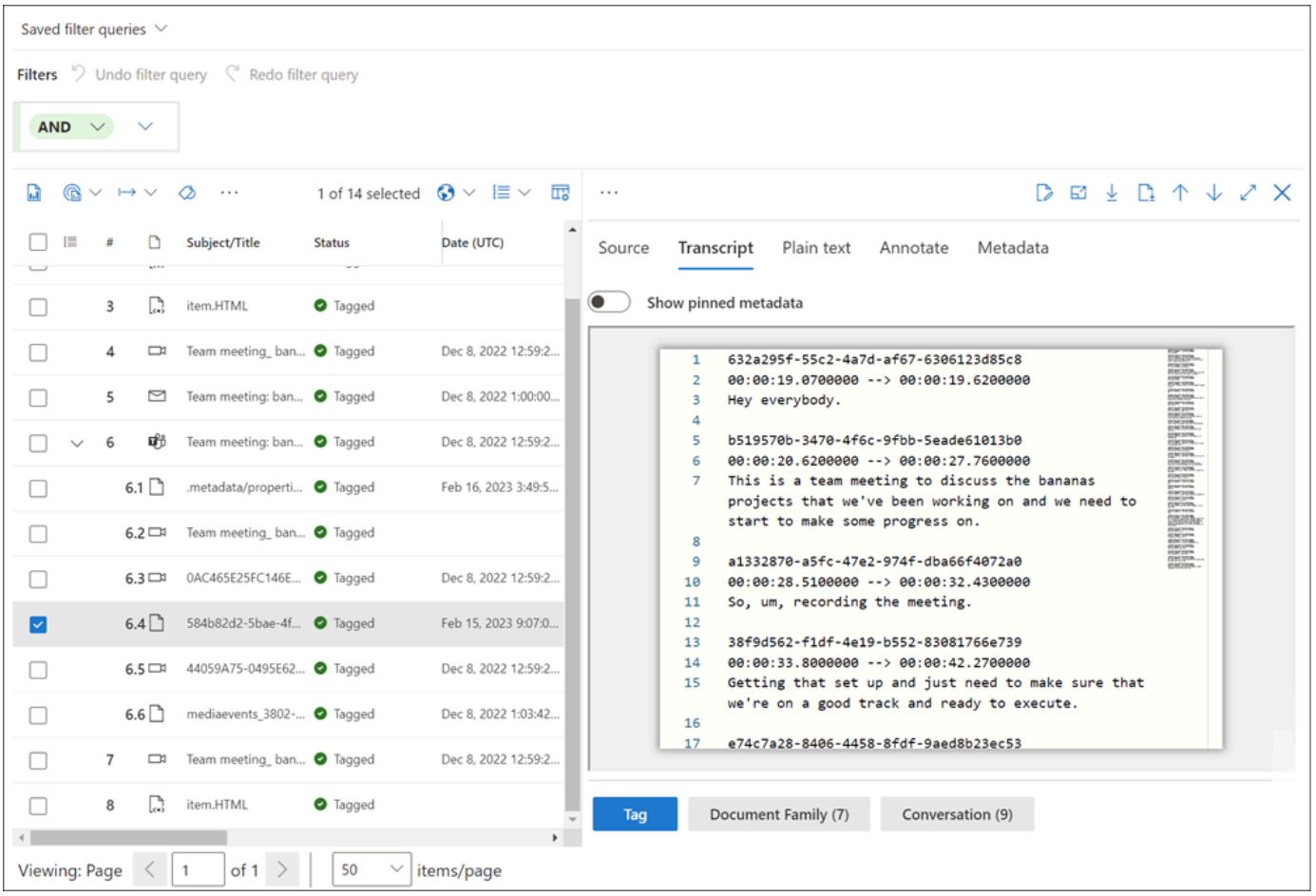
- The transcript of the meeting audio in .txt format
- The video recording of the meeting in .mp4 format
- The thumbnail image of the meeting in .jpg format
- Meeting metadata and meeting chapters (as applicable) in .json format

To view meeting audio transcript files in a review set, you'll select the meeting and the **Transcript** viewer on the meeting details pane. The following screenshots show an example of a meeting in the Teams client and the meeting transcript file of the same meeting in the review set.

Meeting in Teams client



Meeting transcript file in review set

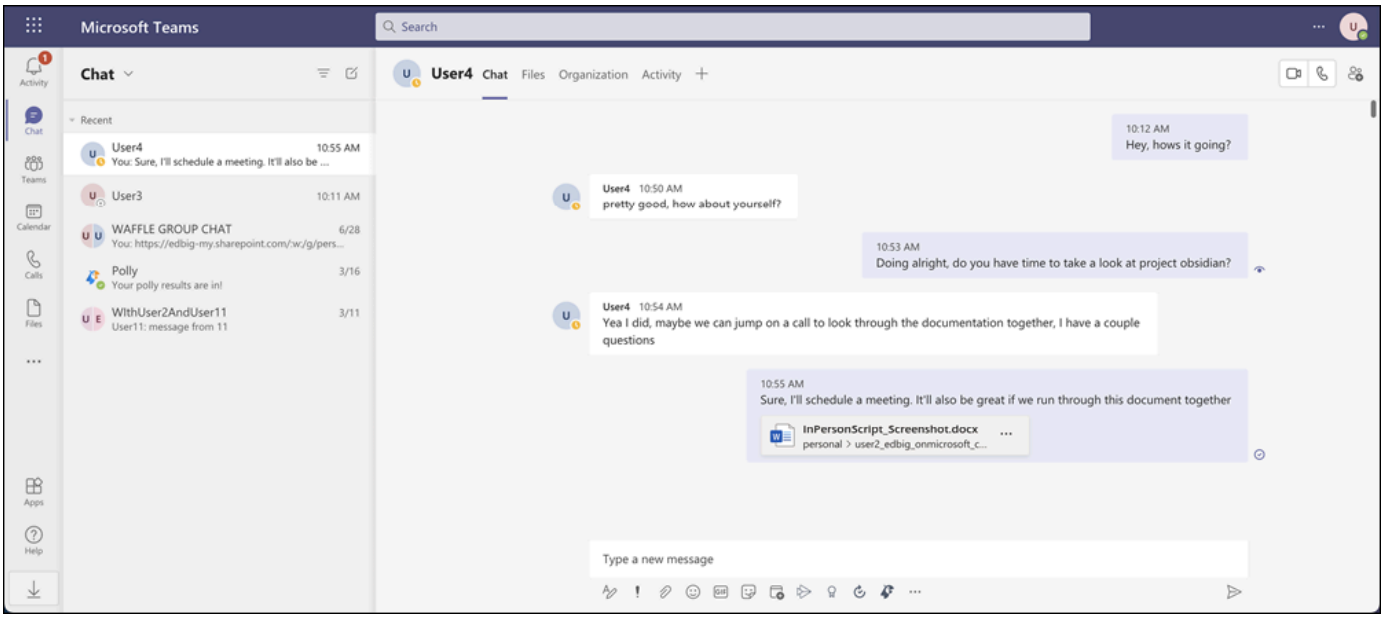


Viewing conversation transcript files

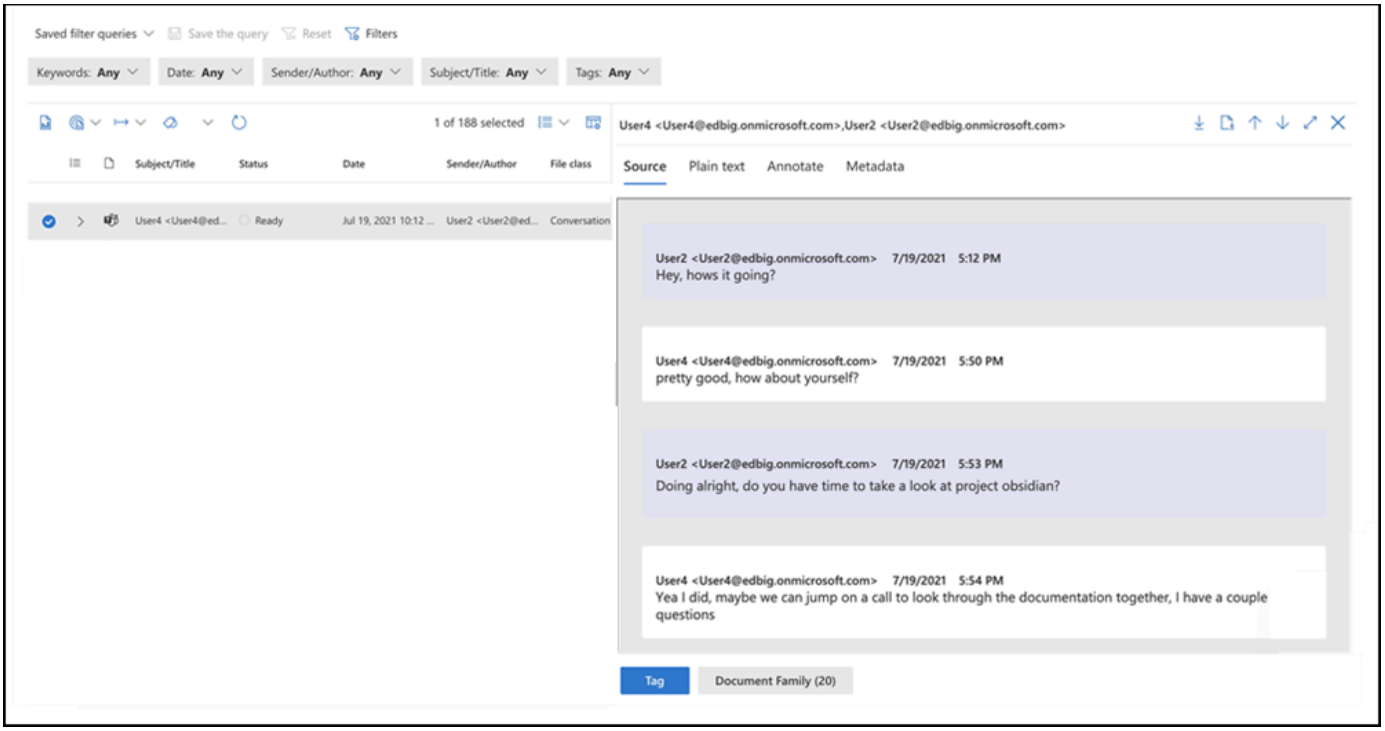
When viewing transcript files in a review set, some of the messages are highlighted in purple. The messages that are highlighted depend on which custodian copy of the transcript you're viewing. For example, in a 1:1 chat between User4 and User2, the messages posted by User4 are highlighted in purple when you view the transcript collected from User4's mailbox. When you view User2's transcript of the same conversation, messages posted by User2 are highlighted in purple. This highlighting behavior is based on the same Teams client experience, where a user's posts are highlighted in purple in the Teams client.

The following screenshots show an example of conversation in the Teams client and the transcript file of the same conversation in the review set. The purple highlighting in the transcript file indicates that the transcript was collected from User2's mailbox.

Conversation in Teams client



Conversation in transcript file



Conversation transcript threading

Conversation threading functionality in the new case format in eDiscovery (Premium) helps you identify contextual content related to items that may be relevant to your investigation. This feature produces distinct conversation views that include chat messages that precede and follow the items match the search query during collection. This capability allows you to efficiently and rapidly review complete chat conversations (called *threaded conversations*) in Microsoft Teams. As previous explained, chat conversations are reconstructed in HTML transcript files when eDiscovery (Premium) adds Teams content to a review set.

Here's the logic used by eDiscovery (Premium) to include additional messages and replies transcript files that provide context around the items match the collection query (called *responsive items*) you used when collecting Teams content. Different threading behaviors are based on the types of chats and the search query used to collect the responsive items. There are two common collection scenarios:

- Queries that use search parameters, such as keywords and property:value pairs
- Queries that only use date ranges

Expand table

Teams content type	Queries with search parameters	Queries with date ranges
Teams 1:1 and group chats	Messages that were posted 12 hours before and 12 hours after responsive items are grouped with the responsive item in a single transcript file.	Messages in a 24-hour window are grouped in a single transcript file.
Standard, private, and shared Teams channel chats	Each post that contains responsive items and all corresponding replies are grouped in a single transcript file.	Each post that contains responsive items and all corresponding replies are grouped in a single transcript file.

Deduplication of Teams content

The following list describes the deduplication (and duplication) behavior when collecting Teams content into a review set.

- Each transcript file added to a review set should be a one-to-one mapping to content stored in data locations. That means eDiscovery (Premium) doesn't collect any Teams content that has already been added to the review set. If a chat message is already collected in a review set, eDiscovery (Premium) doesn't add the same message from the same data location to the review set in subsequent collections.

- For 1:1 and group chats, copies of messages are stored in the mailbox of each conversation participant. Copies of the same conversation that exist in different participants' mailboxes are collected with different metadata. As a result, each instance of the conversation is treated as unique and brought into the review set in separate transcript files. So if all participants of a 1:1 or group chat are added as custodians in a case and included in the scope of a collection, then copies of each transcript (for the same conservation) are added to the review set and will be grouped together with the same **ConversationId**. Each of these

copies is associated with a corresponding custodian. **Tip**: The **Custodian** column in the review set list identifies the custodian for the corresponding transcript file.

- In subsequent collections of items from the same conversation, only the delta content that wasn't previously collected previously is added to the review set and grouped (by sharing the same **ConversationId**) with the previously collected transcripts from the same conversation. Here's an example of this behavior:

    1. Collection A collects messages in a conversation between User1 and User2 and adds to review set.
    2. Collection B collects messages from the same conversation, but there are new messages between User1 and User2 since Collection A was run.
    3. Only the new messages in Collection B are added to the review set. These messages are added to a separate transcript file, but the new transcript is grouped with the transcripts from Collection A by the same **ConversationId**.

    This behavior applies to all the types of Teams chats.

Metadata for Teams content

In large review sets with thousands or millions of items, it can be difficult to narrow the scope of your review to Teams content. To help you focus your review on Teams content, there are metadata properties that are specific to Teams content. You can use these properties to organize the columns in the review list and configure filters and queries to optimize the review of Teams content. These metadata properties are also included when you export Teams content from eDiscovery (Premium), to help you organize and view content post-export or in third-party eDiscovery tools.

The following table describes metadata properties for Teams content.

Expand table

Metadata property	Description
ContainsEditedMessage	Indicates whether a transcript file contains an edited message. Edited messages are identified when viewing the transcript file.
ConversationId	A GUID that identifies the conversation that the item is associated with. Transcript files and attachments from the same conversation have the same value for this property.
Conversation name	The name of the conversation the transcript file or attachment is associated with. For Teams 1:1 and group chats, the value of this property is the UPN of all

Metadata property	Description
	participants of the conversation are concatenated. For example, `User3 <User3@contoso.onmicrosoft.com>,User4 <User4@contoso.onmicrosoft.com>,User2 <User2@contoso.onmicrosoft.com>`. Teams channel (standard, private, and shared) chats use the following format for conversation name: `<Team name>,<Channel name>`. For example, `eDiscovery vNext, General`.
ConversationType	Indicates the type of Team chat. For Teams 1:1 and group chats, the value for this property is `Group`. For standard, private, and shared channel chats, the value is `Channel`.
Date	The time stamp of the first message in the transcript file in UTC.
FamilyId	A GUID that identifies the transcript file for a chat conversation. Attachments will have the same value for this property as the transcript file that contains the message the file was attached to.
FileClass	Indicates that type of content. Items from Teams chats have the value `Conversation`. In contrast, Exchange email messages have the value `Email`.
MessageKind	The message kind property. Teams content has the value `microsoftteams , im`.
Recipients	A list of all users who received a message within the transcript conversation.
TeamsChannelName	The Teams channel name of the transcript.

For descriptions of other eDiscovery (Premium) metadata properties, see Document metadata fields in eDiscovery (Premium).

Export Teams content

After you have reviewed and culled Teams content in a review set, you can export the transcript files that contain content that's responsive to your investigation. There aren't any specific export settings for Teams content. Each transcript file is exported as an HTML message file. The metadata properties discussed in the previous section are included when Teams content is exported.

Each transcript file is referenced in the load file and can be located using the relative path in the Export_native_path field in the load file. Transcript files are located in the Conversations folder in the root export folder.

Tips for viewing Teams content in a review set

Here are some tips and best practices for viewing Teams content in a review set.

- Use the **Customize columns** control in the command bar to add and organize columns to optimize the review of Teams content. You can add and remove columns that are useful for Teams content. You can also sequence the order of columns by dragging and dropping them in the **Edit column** flyout page. You can also sort on columns to group Teams content with similar values for the column you sort on.
- Useful columns that to help you review Teams content include **Custodian**, **Recipients**, and **File type** or **Message kind**.
- Use filters for Teams-related properties to quickly display Teams content. There are filters for most of the metadata properties described in the previous section.

Deleting Teams chat messages

You can use eDiscovery (Premium) and the Microsoft Graph Explorer to respond to data spillage incidents, when content containing confidential or malicious information is released through Teams chat messages. Admins in your organization can search for and delete chat messages in Microsoft Teams. This feature can help you remove sensitive information or inappropriate content in Teams chat messages. For more information, see Search and purge chat messages in Teams.

Last updated on 10/31/2024

Conduct an eDiscovery investigation of content in Microsoft Teams

Applies to: Microsoft Teams

> **⊗ Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic **Content Search**, classic **eDiscovery (Standard)**, and classic **eDiscovery (Premium)**.
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the **new eDiscovery experience** in the **Microsoft Purview portal**.

Large enterprises are often exposed to high penalty legal proceedings that demand submission of all Electronically Stored Information (ESI). Microsoft Teams content can be searched and used during eDiscovery investigations.

Overview

All Microsoft Teams 1:1 or group chats are journaled through to the respective users' mailboxes. All standard channel messages are journaled through to the group mailbox representing the team. Files uploaded in standard channels are covered under the eDiscovery functionality for SharePoint Online and OneDrive for Business.

eDiscovery of messages and files in private channels works differently than in standard channels. To learn more, see eDiscovery of private channels.

Recorded Teams meetings are stored in the OneDrive for Business account of the meeting organizer or initiator, depending on the *MeetingRecordingOwnership* configuration. Attendee identification information when using the **Hide Attendee Names** functionality for Teams meetings is stored in the user mailbox of the meeting organizer. To learn more, see eDiscovery (Premium) workflow for content in Microsoft Teams.

Not all Teams content is eDiscoverable. The following table shows the Teams content types that you can search for using Microsoft eDiscovery tools:

Expand table

Content type	Notes
Audio recordings	Audio calls between Teams user and external contacts
Card content	See Search for card content for more information.
Chat links	
Chat messages	This includes content in standard Teams channels, 1:1 chats, 1:N group chats, chats with yourself, and chats with guests.
Code snippets	
Edited messages	If the user is on hold, previous versions of edited messages are also preserved.
Emojis, GIFs, and stickers	
Inline images	
Loop components	Content in a loop component is saved in a .fluid file that's stored in the OneDrive for Business account of the user who sends the loop component. That means you have to include OneDrive as a data source when searching for content in loop components.
Meeting IM conversations	
Meeting metadata[1]	
Meeting recordings and transcripts	Transcripts of the meeting audio are extracted and provided as a separate file. Maximum supported recorded meeting .mp4 file size is 350 MB. If the recorded meeting file size is greater than 350 MB, a processing error occurs and the file is available for download.
Name of channel	
Quotes	Quoted content is searchable. However, search results don't indicate that the content was quoted.
Reactions (such as likes, hearts, and other reactions)	Reactions are supported for all commercial customers after June 1, 2022. Reactions before this date aren't available for eDiscovery. Expanded reactions are now supported. To understand reaction history, the content must be on legal hold.
Subject	
Tables	
Teams Video Clip (TVC)	Search TVC with "Video-Clip" keyword and "save as" a .mp4 file for each TVC attachment by right-clicking the preview.

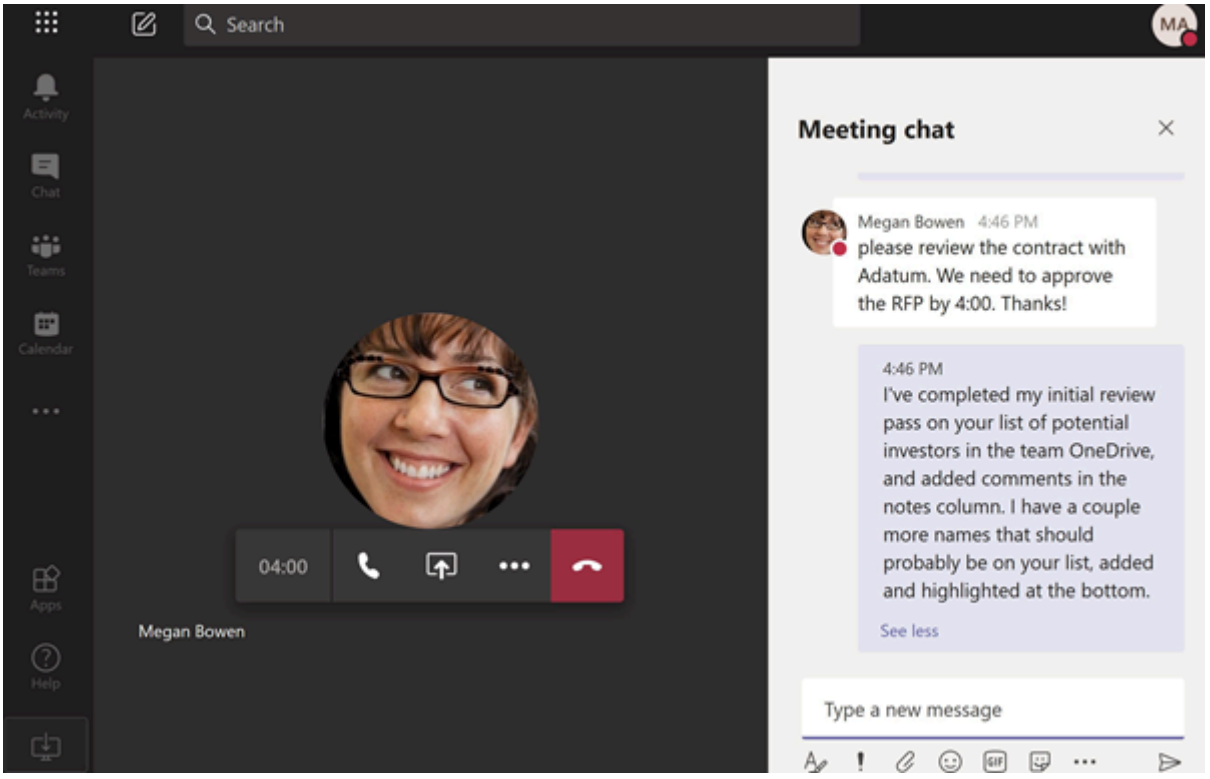
Content type	Notes
	TVCs are collected as Teams conversation attachments (if smaller than 200 MB) and separate .mp4 files. TVC file data is discoverable in eDiscovery review sets and can be exported. Preview of video clips isn't currently supported.

[1] Meeting (and call) metadata includes the following:

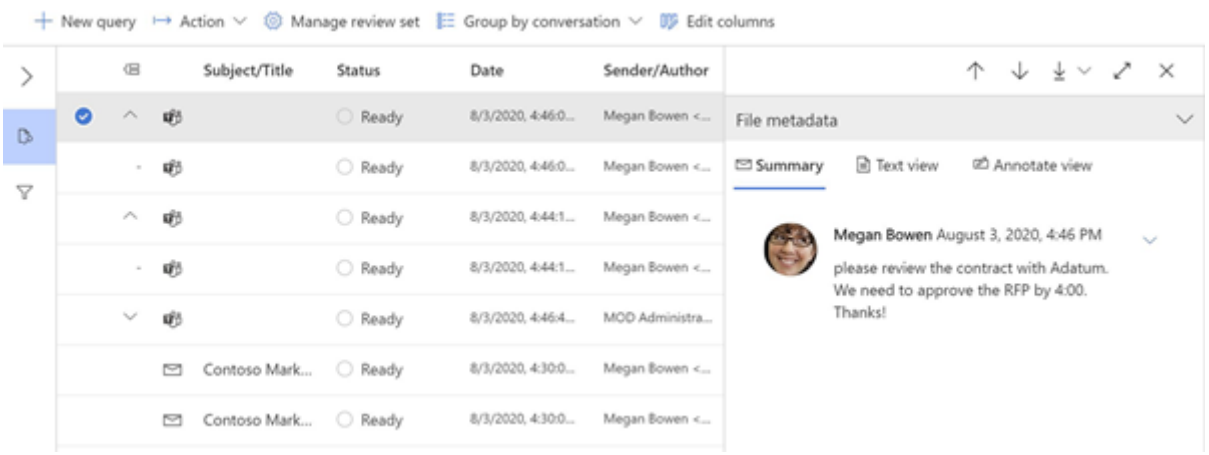
- Meeting start and end time, and duration
- Meeting join and leave events for each participant
- VOIP joins/calls
- Federated user joins
- Guest joins

> **ⓘ Important**
>
> Anonymous users joining meetings and calls aren't currently supported in eDiscovery search queries.

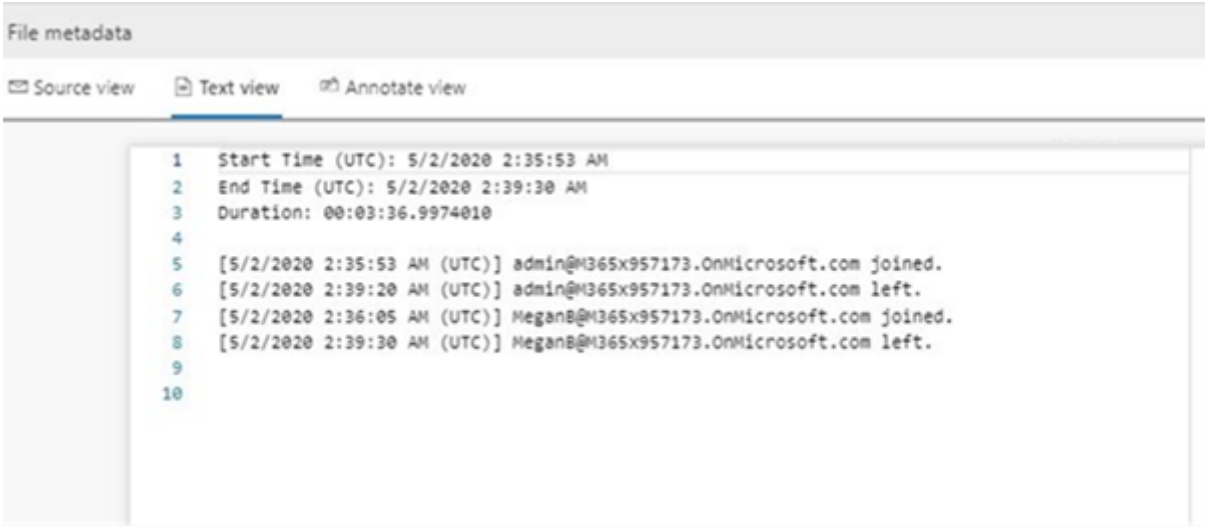
Here's an example of a chat conversation between participants during a meeting.



Here's an example of the compliance copy of the same chat conversation viewed in an eDiscovery tool.



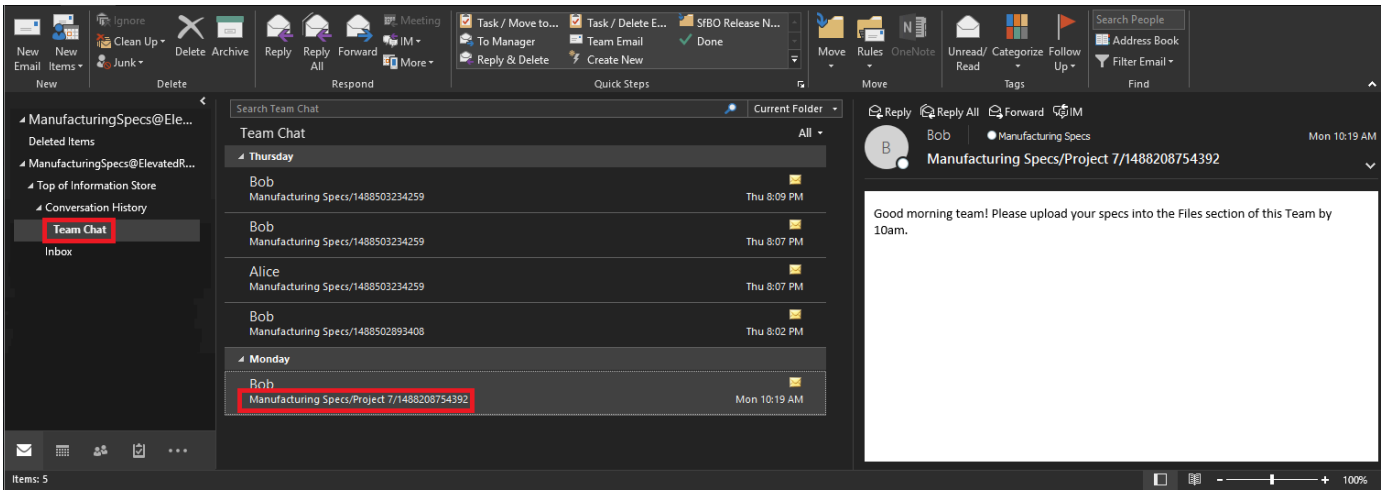
Here's an example of the meeting metadata.



For more information about conducting an eDiscovery investigation, see Get started with eDiscovery (Standard).

Microsoft Teams data appears as IM or Conversations in the Excel eDiscovery export output. You can open the `.pst` file in Outlook to view those messages after you export them.

When viewing the .pst file for the team, all conversations are located in the Team Chat folder under Conversation History. The title of the message contains the team name and channel name. For example, the image below shows a message from Bob who messaged the Project 7 standard channel of the Manufacturing Specs team.



Private chats in a user's mailbox are stored in the Team Chat folder under Conversation History.

eDiscovery of private and shared channels

Compliance copies of messages in private and shared channels are sent to different mailboxes depending on the channel type. That means you have to search different mailbox locations based on the type of channel a user is a member of.

- **Private channels**. Compliance copies are sent to the mailbox of all members of the private channel members. That means you have to search the user mailbox when searching for content in private channel messages.
- **Shared channels**. Compliance copies are sent to a system mailbox that's associated with the parent team. Because Teams doesn't support an eDiscovery search of a single system mailbox for a shared channel, you have to search the mailbox for the parent team (by selecting the name of the Team mailbox) when searching for message content in shared channels.

Each private and shared channel has its own SharePoint site that's separate from the parent team site. That means files in private and shared channels are stored in its own site and managed independently of the parent team. This means you must identify and search the specific site associated with a channel when searching for content in files and channel message attachments.

Use the following sections to help identify the private or shared channel to include in your eDiscovery search.

Identifying the members of a private channel

Use the procedure in this section to identify members of a private channel so that you can use eDiscovery tools to search the member's mailbox for content in private channel messages.

Before you perform these steps, make sure you have the latest version of the Teams PowerShell module installed.

1. Run the following command to get the group ID of the team that contains the shared channels you want to search.

   PowerShell

   ```
   Get-Team -DisplayName <display name of the the parent team>
   ```

   > **Tip**
   >
   > Run the **Get-Team** cmdlet without any parameters to display a list of all Teams in your organization. The list contains the group Id and DisplayName for every team.

2. Run the following command to get a list of private channels in the parent team. Use the group ID for the team that you obtained in step 1.

   PowerShell

   ```
   Get-TeamChannel -GroupId <parent team GroupId> -MembershipType Private
   ```

3. Run the following command to get a list of private channel owners and members for a specific private channel.

   PowerShell

   ```
   Get-TeamChannelUser -GroupId <parent team GroupId> -DisplayName "Partner Shared Channel"
   ```

4. Include the mailboxes of owners and members of a private channel as part of your eDiscovery search query in eDiscovery (Standard) or when identifying and collecting custodian content in eDiscovery (Premium).

Identifying the SharePoint site for private and shared channels

As previously explained, files shared in private and shared channels (and files attached to channel messages) are stored in the site collection associated with the channel. Use the procedure in this section to identify the URL for the site associated with a specific private or shared channel. Then you can use eDiscovery tools to search for content in the site.

Before you perform these steps, install the SharePoint Online Management Shell and connect to SharePoint Online.

1. Optionally, run the following to get a list of all SharePoint site collections associated with shared channels in the parent team.

   PowerShell

   ```
   Get-SPOSite
   ```

   > **Tip**
   >
   > The naming convention of the URL for a site that's associated with private and shared channels is `[SharePoint domain]/sites/[Name of parent team]-[Name of private or shared channel]`. For example, the URL for the shared channel named "Partner Collaboration", which is located in the "Engineer Team" parent team in the Contoso organization is `https://contoso.sharepoint.com/sites/EngineeringTeam-PartnerCollaboration`.

2. Run the following PowerShell commands to display the URL for all SharePoint sites associated with the private and shared channels in your organization. The output of the script also includes the group ID of the parent team, which you need to run the commands in step 3.

   PowerShell

   ```
   $sites = Get-SPOSite -Template "TEAMCHANNEL#1"
   foreach ($site in $sites) {$x= Get-SPOSite -Identity $site.url -Detail;
   $x.relatedgroupID; $x.url}
   ```

   > **Note**
   >
   > SharePoint sites for private channels created before June 28, 2021 use the value `"TEAMCHANNEL#0"` for the custom template ID. To displays private channels created after this date, use the value `"TEAMCHANNEL#1"` when running the previous two scripts. Shared channels only use the value of `"TEAMCHANNEL#1"`.

3. For each parent team, run the following PowerShell commands to identify the private and shared channel sites, where `$groupID` is the group ID of the parent team.

PowerShell

```
$sites = Get-SPOSite -Template "TEAMCHANNEL#1"
$groupID = "<group ID of parent team)"
foreach ($site in $sites) {$x= Get-SpoSite -Identity $site.url -Detail; if
($x.RelatedGroupId -eq $groupID) {$x.RelatedGroupId;$x.url}}
```

4. Include the site associated with a private or shared channel as part of your eDiscovery search query in eDiscovery (Standard) or when identifying and collecting custodian content in eDiscovery (Premium).

Search for content for guests

You can use eDiscovery tools to search for Teams content related to guests in your organization. Teams chat content that's associated with a guest is preserved in a cloud-based storage location and can be searched for using eDiscovery. This includes searching for content in 1:1 and 1:N chat conversations in which a guest is a participant with other users in your organization. You can also search for private channel messages in which a guest is a participant and search for content in *guest:guest* chat conversations where the only participants are guests.

To search for content for guests:

1. Connect to Microsoft Graph PowerShell. For more information, see the Microsoft Graph PowerShell overview. Be sure to complete Step 1 and Step 2 in the previous article.

2. After you successfully connect to Microsoft Graph PowerShell, run the following command to display the user principal name (UPN) for all guests in your organization. You have to use the UPN of the guest when you create the search in step 4.

   PowerShell

   ```
   Get-MgUser -Filter "userType eq 'Guest'" -All $true | FL UserPrincipalName
   ```

   > **Tip**
   >
   > Instead of displaying a list of user principal names on the computer screen, you can redirect the output of the command to a text file. You can do this by appending `> filename.txt` to the previous command. The text file with the user principal names will be saved to the current folder.

3. In a different Windows PowerShell window, connect to Security & Compliance PowerShell. For instructions, see Connect to Security & Compliance PowerShell. You can connect with or without using multi-factor authentication.

4. Create a content search that searches for all content (such as chat messages and email messages) in which the specified guest was a participant by running the following command.

   PowerShell

   ```
   New-ComplianceSearch <search name> -ExchangeLocation <guest UPN>  -
   AllowNotFoundExchangeLocationsEnabled $true -IncludeUserAppContent $true
   ```

   For example, to search for content associated with the guest Sara Davis, you would run the following command.

   PowerShell

   ```
   New-ComplianceSearch "Sara Davis Guest" -ExchangeLocation
   "sara.davis_hotmail.com#EXT#@contoso.onmicrosoft.com" -
   AllowNotFoundExchangeLocationsEnabled $true -IncludeUserAppContent $true
   ```

   For more information about using PowerShell to create content searches, see New-ComplianceSearch.

5. Run the following command to start the content search that you created in step 4:

   PowerShell

   ```
   Start-ComplianceSearch <search name>
   ```

6. Go to https://compliance.microsoft.com and then select **Show all** > **Content search**.

7. In the list of searches, select the search that you created in step 4 to display the flyout page.

8. On the flyout page, you can do the following things:

   - Select **View results** to view the search results and preview the content.
   - Next to the **Query** field, select **Edit** to edit and then rerun the search. For example, you can add a search query to narrow the results.
   - Select **Export results** to export and download the search results.

Search for card content

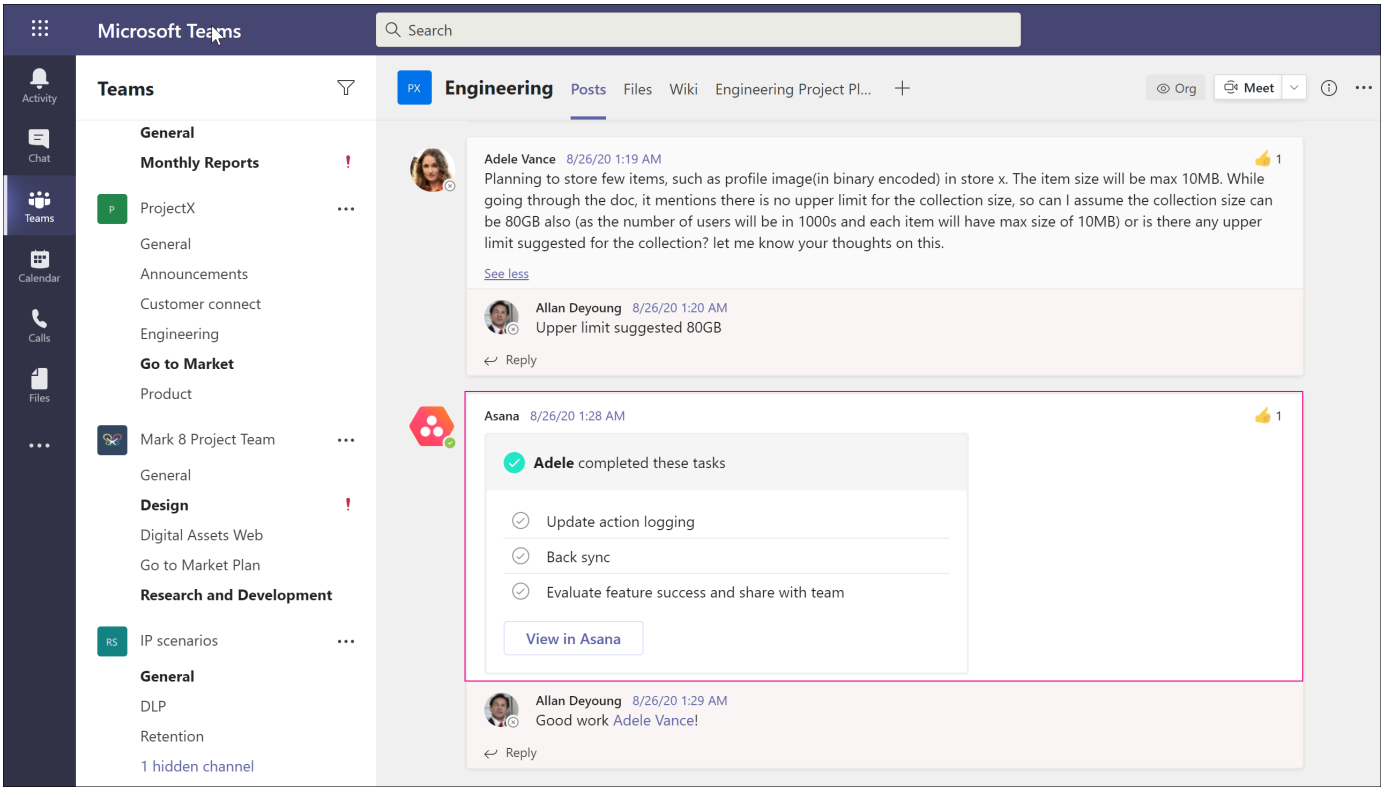
Card content generated by apps in Teams channels, 1:1 chats, and 1xN chats are stored in mailboxes and can be searched. A *card* is a UI container for short pieces of content. Cards can have multiple properties and attachments, and can include items that can trigger card actions. For more information, see Cards

Like other Teams content, where card content is stored is based on where the card was used. Content for cards used in a Teams channel is stored in the Teams group mailbox. Card content for 1:1 and 1xN chats are stored in the mailboxes of the chat participants.

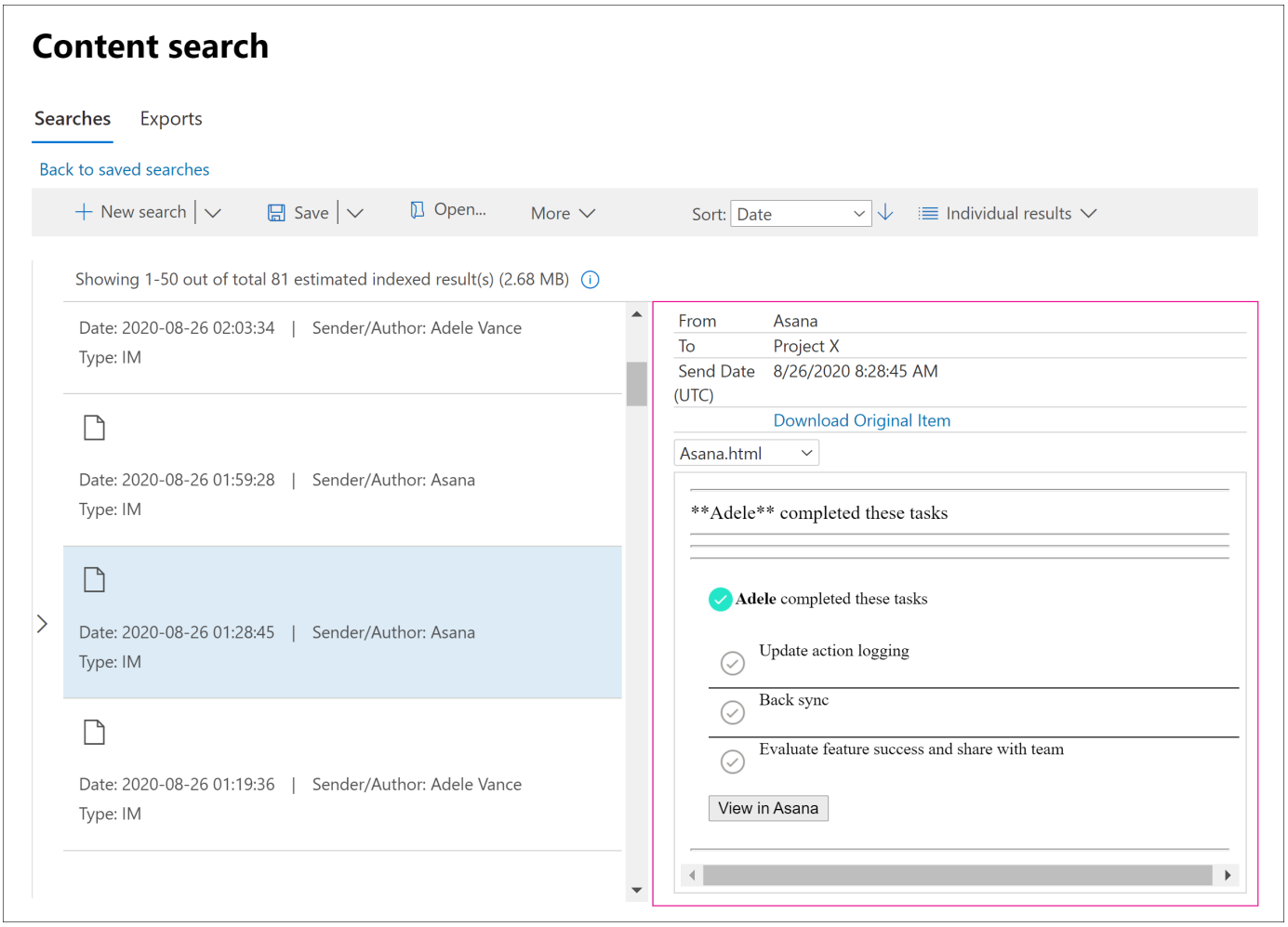
To search for card content, you can use the `kind:microsoftteams` or `itemclass:IPM.SkypeTeams.Message` search conditions. When reviewing search results, card content generated by bots in a Teams channel has the **Sender/Author** email property as `<appname>@teams.microsoft.com`, where `appname` is the name of the app that generated the card content. If card content was generated by a user, the value of **Sender/Author** identifies the user.

When viewing card content in Content search results, the content appears as an attachment to the message. The attachment is named `appname.html`, where `appname` is the name of the app that generated the card content. The following screenshots show how card content (for an app named Asana) appears in Teams and in the results of a search.

Card content in Teams



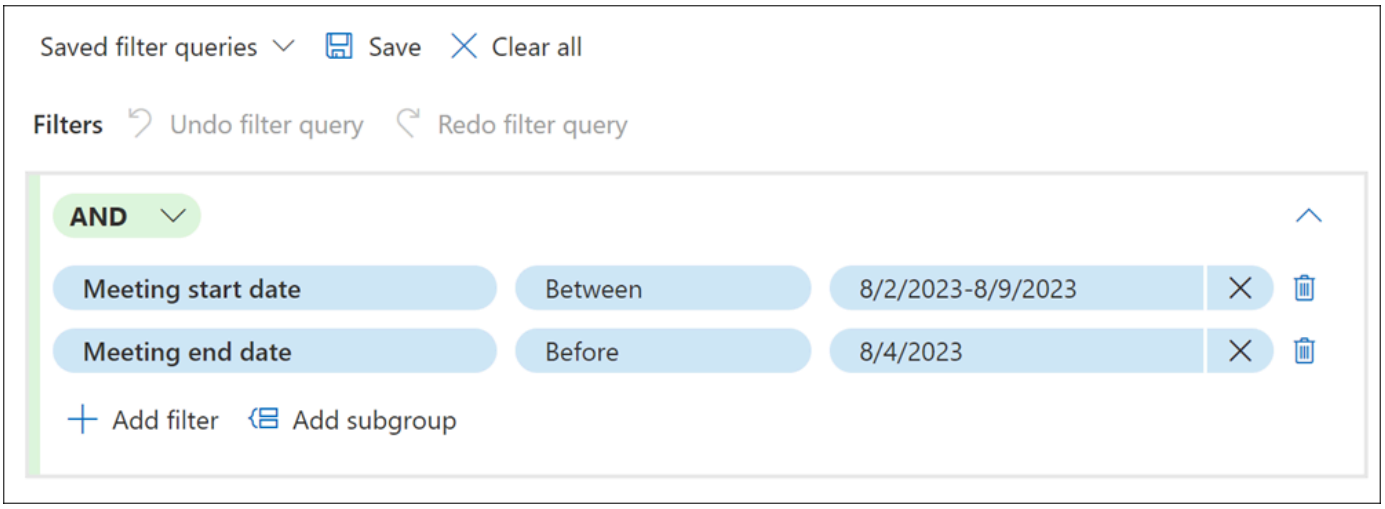
Card content in search results



> **Note**
>
> To display images from card content in search results at this time (such as the checkmarks in the previous screenshot), you have to be signed into Teams (at https://teams.microsoft.com) in a different tab in the same browser session that you use to view the search results. Otherwise, image placeholders are displayed.

Search for Teams meetings by date

Admins can search for Teams meeting content based on the meeting start or end dates. To filter review set items by specific Teams meeting dates, you can use the *Meeting start date* and *Meeting end date* properties in the advanced filtering options in eDiscovery (Premium).



Search for hidden attendees in Teams meetings

The **Hide Attendee Names** feature in Microsoft Teams hides the name of attendees from other attendees so only organizers can see attendees names. This feature may be used for meetings or events with external businesses, vendors, confidential meetings, or other meetings where personal privacy between attendees is important. When this feature is enabled, the names of attendees are hidden in the meeting roster, meeting chat, and meeting recordings.

To search for the names of attendees in Teams meetings where the **Hide Attendee Names** feature was enabled, you must include the organizer's mailbox in the scope for the search. Hidden attendee names are only available in the organizer's mailbox.

eDiscovery in external access and guest environments

Admins can use eDiscovery to search for content in chats messages in a Teams meeting in external access and guest access environments based on the following restrictions:

- **External access**: In a Teams meeting with users from your organization and users from an external organization where external attendees are using external access, admins in both organizations can search for content in chat messages from the meeting.
- **Guest**: In a Teams meeting with users from your organization and guests, only admins in the organization who hosts the Teams meeting can search for content in chat messages from the meeting.

Related articles

- Microsoft 365 eDiscovery solutions
- Get started with eDiscovery (Standard)
- Teams workflow in eDiscovery (Premium)

- Teams PowerShell Overview

Last updated on 10/31/2024

Search for and delete chat messages in Teams

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

You can use eDiscovery (Premium) and the Microsoft Graph Explorer to search for and delete chat messages in Microsoft Teams. This feature can help you find and remove sensitive information or inappropriate content. This search and delete workflow will also help you respond to a data spillage incident, when content containing confidential or malicious information is released through Teams chat messages.

> **Tip**
>
> You can also use the purge feature in Data Security Investigations to search for and permanently delete Microsoft Teams items in your organization. Data Security Investigations provides a built-in search and purge workflow with a purge queue dashboard to help you manage and track purge queries. For more information, see Purge in Data Security Investigations.

Before you search and delete chat messages

- To create an eDiscovery (Premium) case and use collections to search for chat messages, you have to be a member of the *eDiscovery Manager* role group in the Microsoft Purview portal. To delete chat messages, you have to be assigned the *Search And Purge* role. This role is assigned to the *Data Investigator* and *Organization Management* role groups by default. For more information, see Assign eDiscovery permissions.

- Search and purge are supported for most conversations within your tenant. Search and purge for Teams Connect Chat (External Access or Federation) conversations isn't supported.

  > **Important**
  >
  > Chats with yourself (or chats by users with themselves) aren't supported for search and delete.

- A maximum of 100 items per mailbox can be removed at one time. Because the capability to search for and remove chat messages is intended to be an incident-response tool, this limit helps ensure that chat messages are quickly removed.

- Deleting chat messages in federated environments isn't supported.

Search and delete workflow

Here's the process to search for and delete Teams chat messages:



Step 1: Create a case in eDiscovery (Premium)

The first step is to create a case in eDiscovery (Premium) to manage the search and delete process. For information about creating a case, see Use the new case format.

Step 2: Create a collection estimate

After you create a case, the next step is to create a collection estimate to search for the Teams chat messages that you want to delete. The delete process you perform is Step 5 deletes all items that are found in the collection estimate (within the 10 item per location limit).

In eDiscovery (Premium), a *collection* is an eDiscovery search of the Teams content locations that contain the chat messages that you want to delete. Create the collection estimate in the case that you created in the previous step. For more information, see Create a collection estimate.

Data sources for chat messages

Use the following table to determine which data sources to search depending on the type of chat message you need to delete.

Expand table

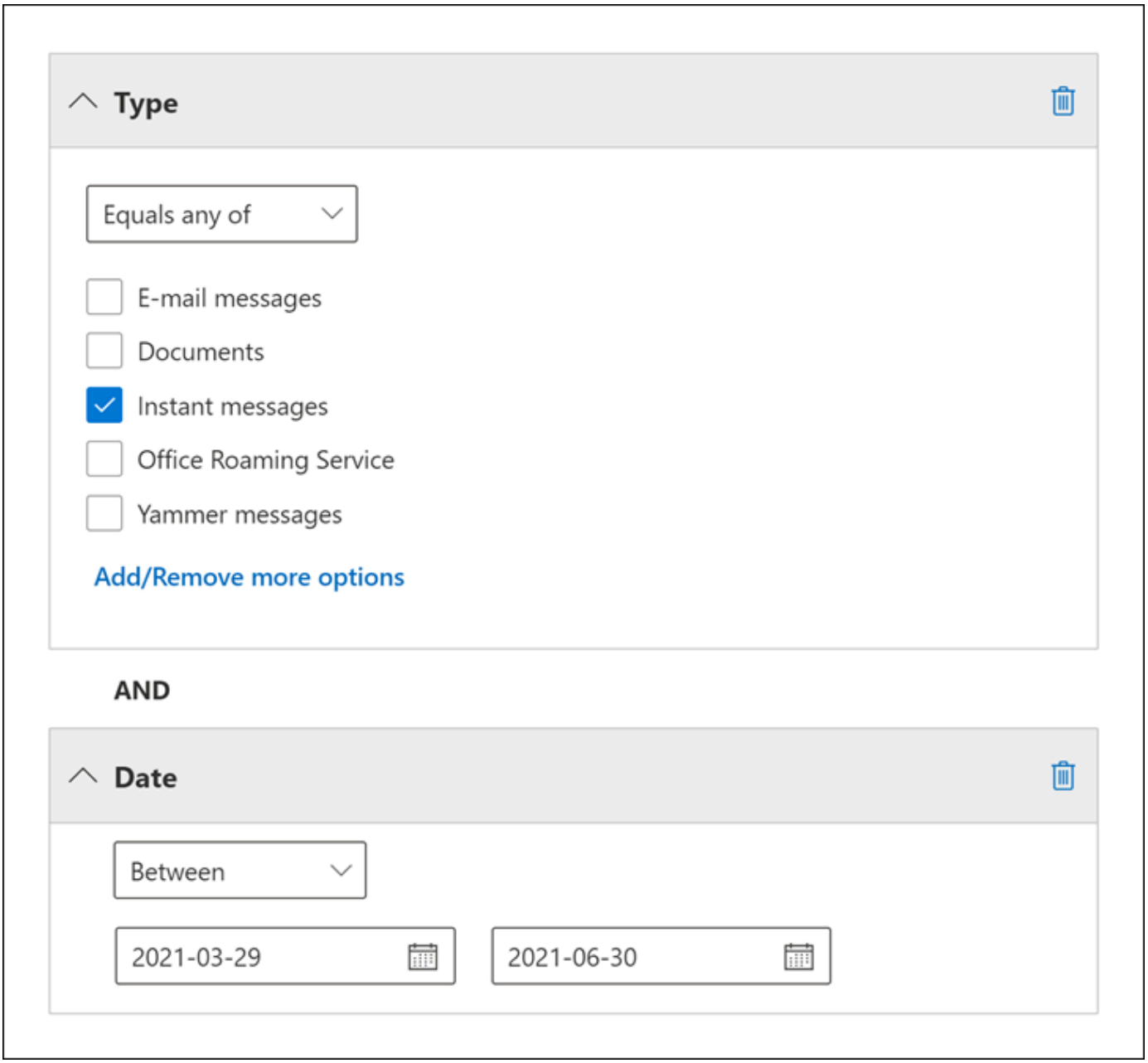
For this type of chat...	Search this data source...
Teams 1:1 chats	The mailbox of chat participants.
Teams group chats	The mailboxes of chat participants.
Teams channels (standard and shared)	The mailbox associated with the parent team.
Teams private channels	The mailbox of the private channel members.

> **Note**
>
> In Step 4, you also have to identify and remove any holds and retention policies assigned to the mailbox that contains the type of chat messages that you want to delete.

Tips for searching for chat messages

To help ensure the most comprehensive collection of Teams chat conversations (including 1:1 and group chats, and chats from standard, shared, and private chats) use the **Type** condition and select the **Instant messages** option when you build the search query for the collection estimate. We also recommend including a date range or several keywords to narrow the scope of the collection to items relevant to your search a delete investigation.

Here's an example of a sample query using the **Type** and **Date** options:



For more information, see Build search queries for collections.

Step 3: Review and verify chat messages to delete

The deletion process in Step 5 deletes the items returned by the collection. It's important that you review the collection estimate results to ensure that the collection only returns the items that you want to delete. To review a sample of items in a collection estimate, see the "Next steps after a collection estimate is complete" section in Create a collection estimate.

Additionally, you can use the collection statistics (specifically the Top Locations statistics) to generate a list of the data sources that contain items returned by the collection. Use this list in the next step to remove hold and retention policies from the data sources that contain search results. For more information, see Collection statistics and reports.

Step 4: Remove all holds and retention policies from data sources

Before you can delete chat messages from a mailbox, you have to remove all organization-wide holds, site holds, or retention policy holds that are assigned to a target mailbox. If not, the chat you're trying to delete is retained.

Use the list of mailboxes that contain the chat messages that you want to delete and determine if there's a hold or retention policy assigned to those mailboxes, and then remove the hold or retention policy. Be sure to identify the hold or retention policy that you remove so that you can reassign to the mailboxes in Step 7.

For instructions about how to identify and remove holds and retention policies, see "Step 3: Remove all holds from the mailbox" in Delete items in the Recoverable Items folder of cloud-based mailboxes on hold.

Step 5: Delete chat messages from Teams

> **Note**
>
> Because Microsoft Graph Explorer is not available in some US Government clouds (GCC High and DOD), you must use PowerShell to accomplish these tasks. See the **Delete chat messages with PowerShell** for details.

Now you're ready to actually delete chat messages from Teams. Use the Microsoft Graph Explorer to perform the following three tasks:

1. Get the ID of the eDiscovery (Premium) case that you created in Step 1. This is the case that contains the collection created in Step 2.
2. Get the ID of the collection that you created in Step 2 and verified the search results in Step 3. The search query in this collection returns the chat messages that will be deleted.
3. Delete the chat messages returned by the collection.

For information about using Graph Explorer, see Use Graph Explorer to try Microsoft Graph APIs.

> **Important**
>
> To perform these three tasks in Graph Explorer, you may have to consent to the eDiscovery.Read.All and eDiscovery.ReadWrite.All permissions. For more information, see the

"Consent to permissions" section in Working with Graph Explorer.

Get the case ID

1. Go to https://developer.microsoft.com/graph/graph-explorer and sign in to the Graph Explorer with an account that's assigned the *Search And Purge* role in the Microsoft Purview portal.

2. Run the following GET request to retrieve the ID for the eDiscovery (Premium) case. Use the value `https://graph.microsoft.com/v1.0/security/cases/ediscoveryCases` in the address bar of the request query. Be sure to select **v1.0** in the API version dropdown list.

   This request returns information about all cases in your organization on the **Response preview** tab.

3. Scroll through the response to locate the eDiscovery (Premium) case. Use the **displayName** property to identify the case.

4. Copy the corresponding ID (or copy and paste it to a text file). You'll use this ID in the next task to get the collection ID.

> **Tip**
>
> Instead of using the previous procedure to obtain the case Id, you can open the case in the Microsoft Purview portal and copy the case Id from the URL.

Get the eDiscoverySearchID

1. In Graph Explorer, run the following GET request to retrieve the ID for the collection that you created in Step 2, and contains the items you want to delete. Use the value `https://graph.microsoft.com/v1.0/security/cases/ediscoveryCases/{ediscoveryCaseID}/searches` in the address bar of the request query, where *{ediscoveryCaseID}* is the CaseID that you obtained in the previous procedure.

2. Scroll through the response to locate the collection that contains the items that you want to delete. Use the *displayName* property to identify the collection that you created in Step 3.

   In the response, the search query from the collection is displayed in the *contentQuery* property. Items returned by this query is deleted in the next task.

3. Copy the corresponding ID (or copy and paste it to a text file). You'll use this ID in the next task to delete the chat messages.

> **Tip**
>
> Instead of using the previous procedure to obtain the search Id, you can open the case in the Microsoft Purview portal. Open the case and navigate to the Jobs tab. Select the relevant collection and under Support information, find the job ID (the job ID displayed here is the same as the collection ID).

Delete the chat messages

1. In Graph Explorer, run the following POST request to delete the items returned by the collection that you created in Step 2. Use the value `https://graph.microsoft.com/v1.0/security/cases/ediscoveryCases/{ediscoveryCaseID}/searches/{ediscoverySearchID}/purgeData` in the address bar of the request query, where *{ediscoveryCaseID}* and *{ediscoverySearchID}* are the IDs that you obtained in the previous procedures.

   If the POST request is successful, an HTTP response code is displayed in a green banner stating that the request was accepted.

For more information on purgeData, see sourceCollection: purgeData.

Delete chat messages with PowerShell

You can also delete chat messages using PowerShell. For example, to delete messages in the US Government cloud you could use a command similar to:

```
Connect-MgGraph -Scopes "ediscovery.ReadWrite.All" -Environment USGov
```

```
Invoke-MgGraphRequest -Method POST -Uri
'/v1.0/security/cases/ediscoveryCases/<ediscoverySearchID>/searches/<search ID>/purgeData'
```

For more information on using PowerShell to delete chat messages, see ediscoverySearch: purgeData.

Step 6: Verify chat messages are deleted

After you run the POST request to delete chat messages, these messages are removed from the Teams client and replaced with an automatically generated stating that an admin has removed the message. For an example of this message, see the End-user experience section in this article.

If you need to confirm a chat message has been removed and don't have access to the end-user message in Teams, re-run the search and verify if any matching messages are found. If there aren't any results for the message, the message has been removed.

Deleted chat messages are moved to the *SubstrateHolds* folder, which is a hidden mailbox folder. Deleted chat messages are stored there for at least 1 day, and then are permanently deleted the next time the timer job runs (typically between 1-7 days). For more information, see Learn about retention for Microsoft Teams.

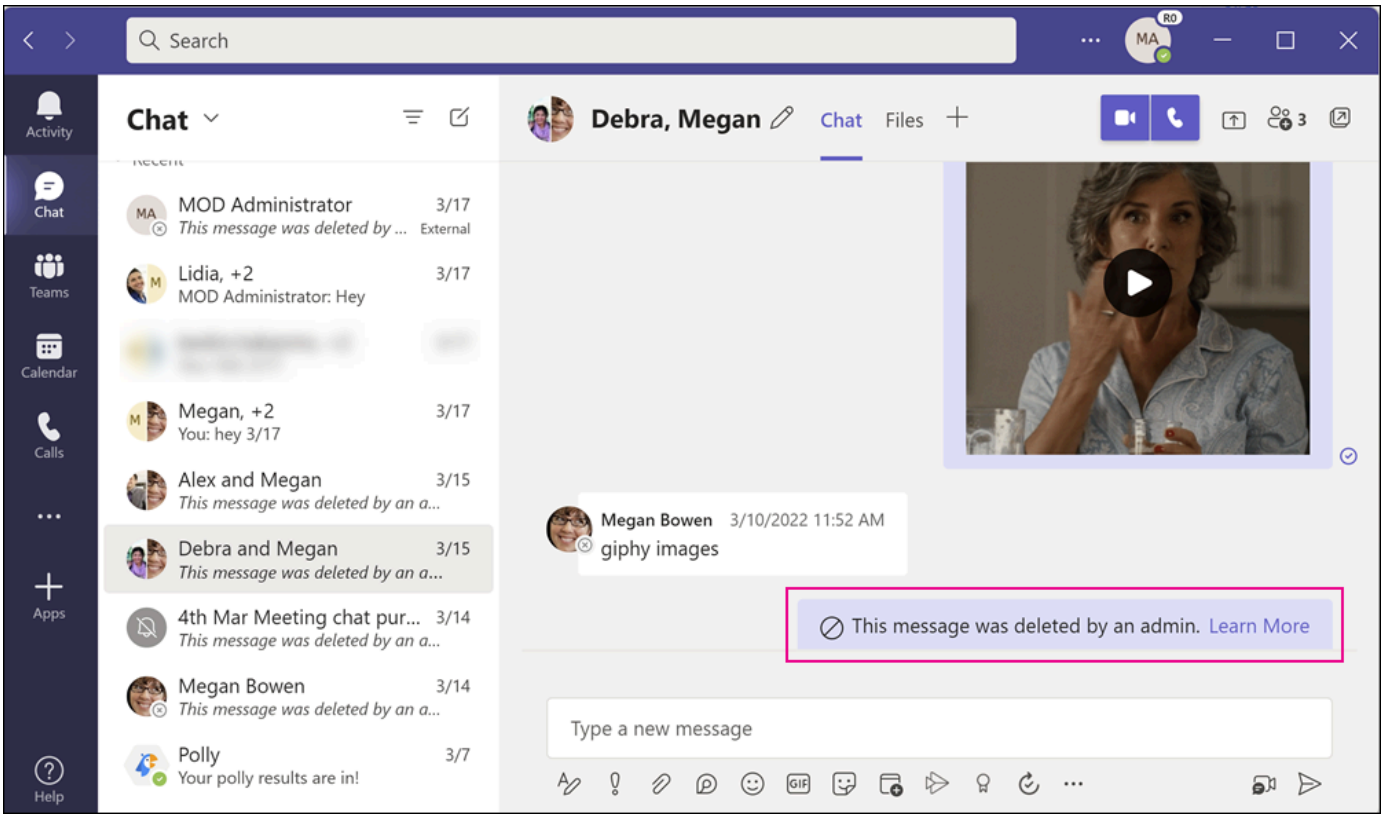
> **Note**
>
> Because Microsoft Graph Explorer is not available in the US Government cloud (GCC, GCC High, and DOD), you must use PowerShell to accomplish these tasks.

Step 7: Reapply holds and retention policies to data sources

After you verify that chat messages are deleted and removed from the Teams client, you can reapply the holds and retention policies that you removed in Step 4.

End-user experience

For deleted chat messages, users see an automatically generated message stating "This message was deleted by an admin".



The message in the previous screenshot replaces the chat message that was deleted.

> **Note**
>
> If you're an end-user and a chat message was deleted, contact your admin for more information.

Last updated on 03/09/2026

Create and manage an eDiscovery (Premium) case

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

After setting up Microsoft Purview eDiscovery (Premium) and assigning permissions to eDiscovery managers in your organization that will manage cases, the next step is to create and manage a case.

This article also provides a high-level overview of using cases to manage the eDiscovery (Premium) workflow for a legal case or other types of investigations.

Create a case

Complete the following steps to create a case and configure case settings. The user who creates the case is automatically added as a member. Members of the case can access the case in the Microsoft Purview portal and perform eDiscovery (Premium) tasks.

1. Go to the Microsoft Purview portal and sign in using the credentials for user account that has been assigned eDiscovery permissions. Members of the *Organization Management* role group can also create eDiscovery (Premium) cases.
2. In the left navigation pane of the Microsoft Purview portal, select **Show all**, and then select **eDiscovery** > **Premium**, and then select the **Cases** tab.
3. Select **Create a case**.
4. On the **Name and description** page, complete the following fields:

- **Name**: give the case a name (required). The case name must be unique in your organization
- **Description**: Add an optional description to help others understand this case.
- **Number**: Enter an optional docket number or other numeric identifier.
- **Case format**: The **New (recommended)** option is automatically selected.

> ⓘ Note
>
> The legacy **Classic** format is no longer available when creating new cases. This format is now retired for all new cases.

5. Select **Next**.

6. On the **Members and settings** page, complete the following fields as applicable:

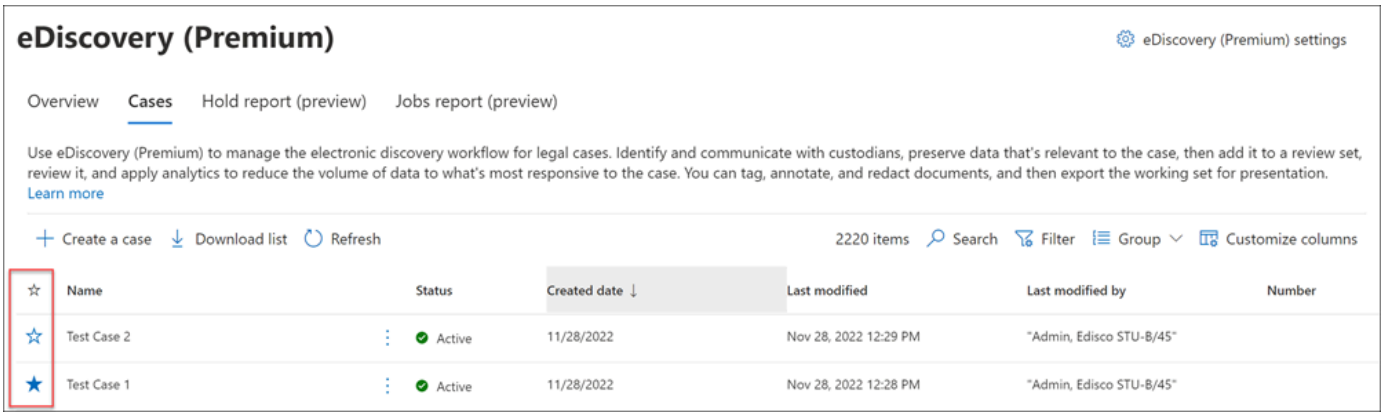
   - **Team members**: Select users and groups that should be assigned to the case. Make sure that users and groups assigned here have been assigned the appropriate eDiscovery permissions.
   - **Search and analytics**: Select the options to configure the case. You can skip this section and configure these settings after the case is created if needed.
   - **Text to ignore**: Add text or regex expressions to define text to ignore in the case. You can apply this to *Near-duplicates*, *Email threads*, or *Themes* modules.
   - **Optical character recognition (OCR)**: Configure the option and settings for finding text contained in images during advanced indexing.

7. Select **Next**.

8. On the **Summary** page, review the settings for the case and edit the settings if needed. Select **Submit** to create the new case and start your investigation.

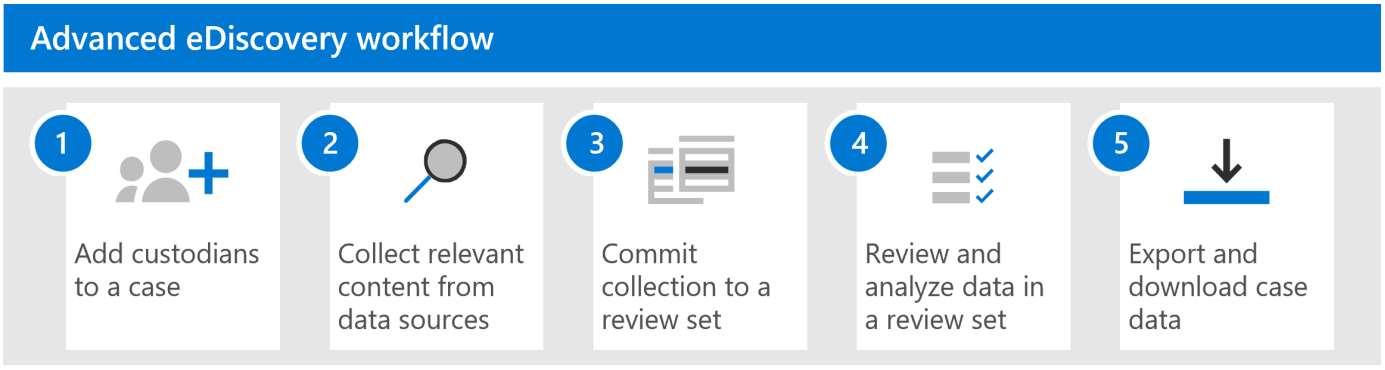
Mark a case as a favorite

You can mark an eDiscovery (Premium) case as a favorite for quicker access to cases you want to prioritize. Cases marked as favorites can be accessed quickly via the eDiscovery (Premium) **Overview** page or can be sorted to be shown at the top of the **Cases** tab for ease of access. You can mark a case as a favorite in the case list on the **Cases** tab or on the top right of the case **Overview** tab for each case. The **Recent favorite cases** card on the **Overview** tab displays all the cases marked as favorites in your organization.



Manage the workflow

To get you started using eDiscovery (Premium), here's a basic workflow that aligns with common eDiscovery practices. In each of these steps, we'll also highlight some extended eDiscovery (Premium) functionality that you can explore. ediscovery-overview



1. **Add custodians and non-custodial data sources to the case**. The first step after creating a case is to add custodians. A *custodian* is a person having administrative control of a document or electronic file that may be relevant to the case. Additionally, you can add data sources that aren't associated with a specific user but may be relevant to the case. These are *non-custodial data sources*.

   Here are some things that happen (or that you can do) when you add custodians to a case:

   - Data in the custodian's Exchange mailbox, OneDrive account, and any Microsoft Teams or Viva Engage groups that the custodian is a member of can be "marked" as custodial data in the case.
   - Custodian (and non-custodial) data is reindexed (by a process called *Advanced indexing*). This helps optimize searching for it in the next step.
   - You can place a hold on custodian and non-custodial data. This hold preserves data that may be relevant to the case during the investigation. To learn more about managing holds, see Manage holds in eDiscovery (Premium).

- You can associate other data sources with a custodian (for example, you can associate a SharePoint site or Microsoft 365 Group with a custodian) so this data can be reindexed, placed on hold, and searched, just like the data in the custodian's mailbox or OneDrive account.

- You can use the communications workflow in eDiscovery (Premium) to send a legal hold notification to custodians.

2. **Collect relevant content from data sources**. After you add custodians and non-custodial data sources to a case, use the built-in collections tool to search these data sources for content that may be relevant to the case. You use keywords, properties, and conditions to build search queries that return search results with the data that's most likely relevant to the case. You can also:

   - View collection statistics that may help you refine a collection to narrow the results.
   - Preview a sample of the collection to quickly verify whether the relevant data is being found.
   - Revise a query and rerun the collection.

3. **Commit collection to a review set**. Once you've configured and verified that a search returns the desired data, the next step is to add the search results to a review set. When you add data to a review set, items are copied from their original location to a secure Azure Storage location. The data is reindexed again to optimize it for thorough and fast searches when reviewing and analyzing items in the review set. Additionally, you can also add non-Office 365 data into a review set.

   There's also a special kind of review set that you can add data to, called a *conversation review set*. These types of reviews sets provide conversation reconstruction capabilities to reconstruct, review, and export threaded conversations like those in Microsoft Teams. For more information, see Review conversations in eDiscovery (Premium).

4. **Review and analyze data in a review set**. Now that data is in a review set, you can use a wide-variety of tools and capabilities to view and analyze the case data with the goal of reducing the data set to what is most relevant to the case you're investigating. Here's a list of some tools and capabilities that you can use during this process.

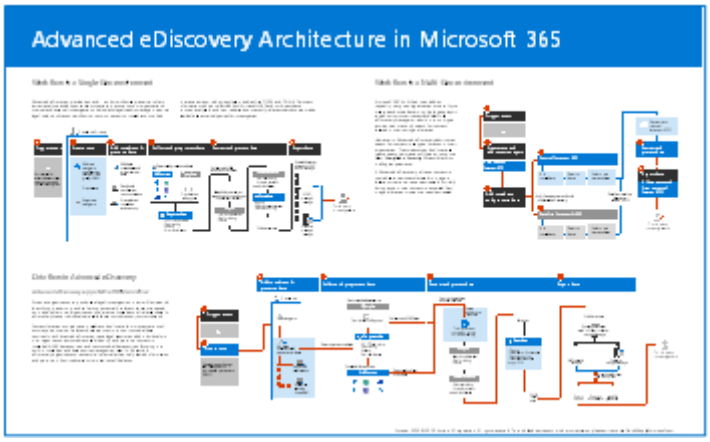
   - Group and view documents. This includes selecting the group options for review sets in your cases, viewing the metadata for each document in a review set, and viewing the document in its native version or text version.

- Create queries and filters. You create search queries using various search criteria (including the ability to search all file metadata properties to further refine and cull the case data to what is most relevant to the case. You can also use review set filters to quickly apply other conditions to the results of a search query to further refine those results.
- Create and use tags. You can apply tags to documents in a review set to identify which are responsive (or non-responsive to the case) and then use those tags when creating search queries to include or exclude the tagged documents. You can also tagging to determine which documents to export.
- Annotate and redact documents. You can use the annotation tool in a review to annotate documents and redact content in documents as work product. We generate a PDF version of an annotated or redacted document during review to reduce the risk of exporting the unredacted native version of the document.
- Analyze case data. The analytics functionality in eDiscovery (Premium) is powerful. After you run analytics on the data in review set, we perform analysis such as near duplicate detection, email threading, and themes that can help reduce the volume of documents that you have to review. We also generate an Analytics reports that summarize the result of running analytics. As previously explained, running analytics also runs the attorney-client privilege detection model.

5. **Export and download case data**. A final step after collecting, reviewing, and analyzing case data is to export it out of eDiscovery (Premium) for external review or for review by people outside of the investigation team. Exporting data is a two-step process. The first step is to export data out of the review set and copy it to a different Azure Storage location (one provided by Microsoft or one managed by your organization). Then you use Azure Storage Explorer to download the data to a local computer. In addition to the exported data files, the contains of the export package also contains an export report, a summary report, and an error report.

eDiscovery (Premium) architecture

Here's an architecture diagram that shows the eDiscovery (Premium) end-to-end workflow in a single-geo environment and in a multi-geo environment, and the end-to-end data flow that's aligned with the Electronic Discovery Reference Model.



View as an image

Download as a PDF file

Download as a Visio file

Last updated on 09/02/2025

Use the new case format in eDiscovery (Premium)

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

More organizations are using the eDiscovery (Premium) solution in Microsoft Purview for critical eDiscovery processes. This includes responding to regulatory requests, investigations, and litigation. As usage of eDiscovery (Premium) increases, a common customer requirement is to expand the total amount of content that can be managed in a single eDiscovery (Premium) case. To help accommodate significant increases in case size, both for total data volume and the total number of items, you can choose the new case format when you create an eDiscovery (Premium) case.

Benefits of the new case format

The new case format allows you to manage cases that contain over 40 million items. This capability helps you effectively manage large volumes of case data through the entire eDiscovery workflow.

Here's a list of other benefits of large cases in eDiscovery (Premium) workflow.

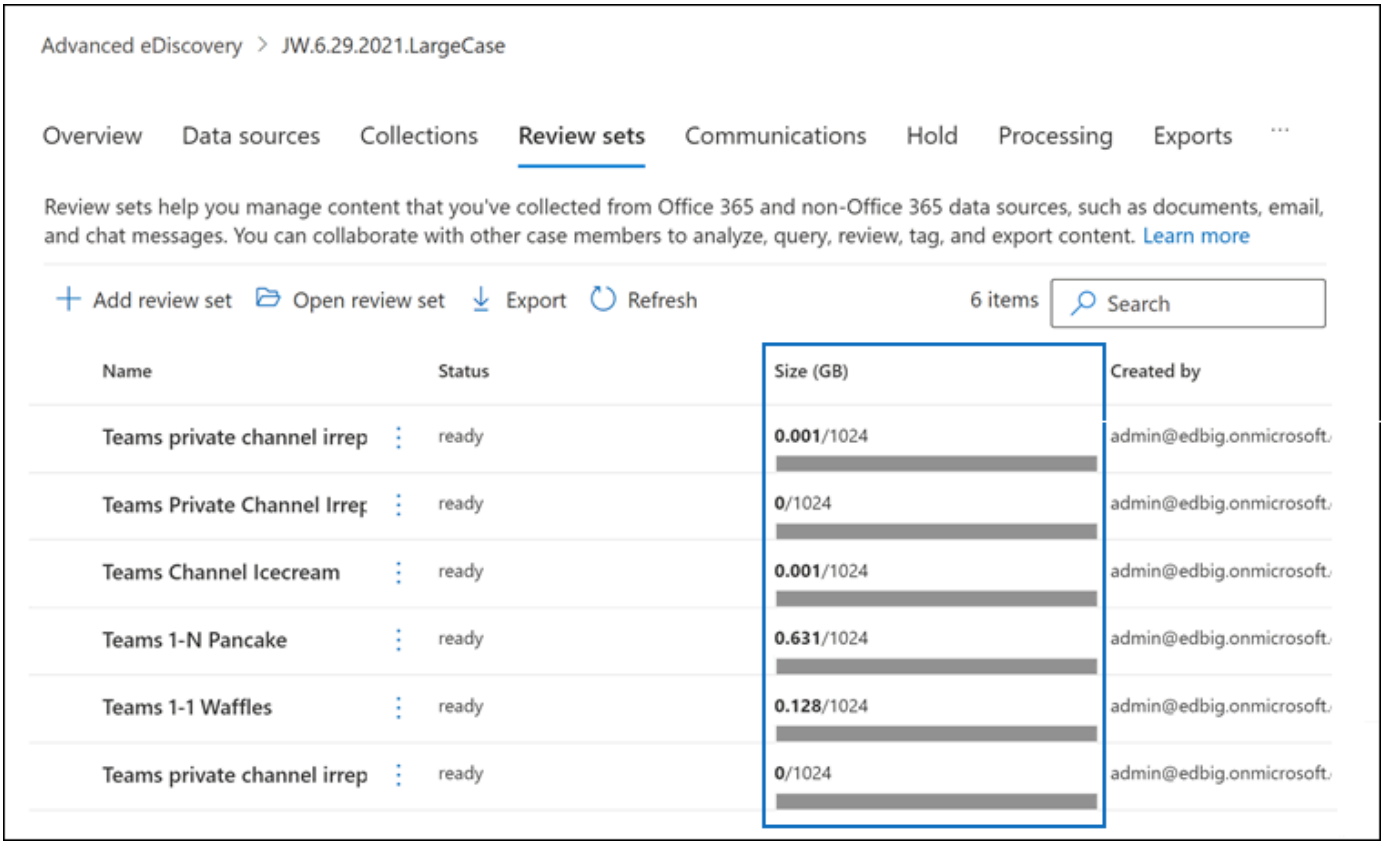
- **Collection**: In the new case format, you can collect up to 1 TB of data for a single collection.

  For each case, the collection settings will default to collect cloud Attachments and contextual Teams and Viva Engage content. These settings help to collect the full picture of digital communications within an investigation. For Teams and Viva Engage contextual conversations, the new case format will convert time-based snapshots of 1:1, 1: N and Channel conversations into HTML transcripts to help provide context of conversations and reduce the total number of items produced by chat-based content.

- **Review**: Each review set will support up to 1 TB of pre-expansion content. Additional metadata will be available for filters and queries including Team name, channel name and conversation name for Teams content. Each transcript will include time-based content for before and after the responsive item. For Channel conversations, the root post and all replies will be collected for responsive content. For more information, see eDiscovery (Premium) workflow for content in Microsoft Teams

- **Export**: You can export large sets of content in a single export job. The new case format lets you can export 5 million documents or 500 GB, whichever is smaller in an export job.

Additionally, the new case format includes an updated user interface that displays the total size of each review set in the case. Review set sizes are displayed in a column on the **Review sets** tab.



Create a case in the new format

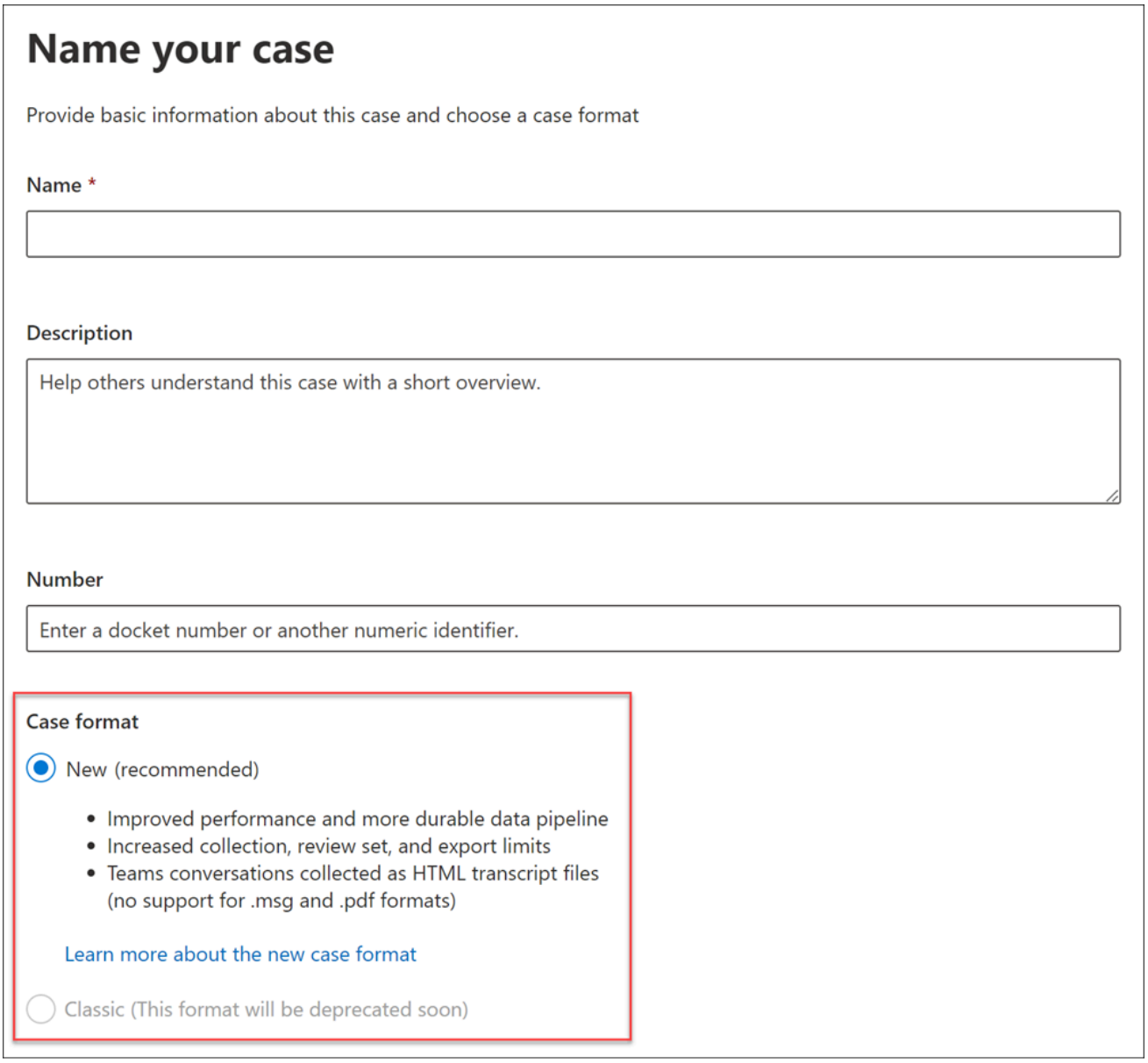
To create an eDiscovery (Premium) case using the new case format, complete the following steps:

1. Go to https://compliance.microsoft.com and sign in.

2. In the left navigation pane of the Microsoft Purview portal, select **eDiscovery** > **Premium**.

3. On the **eDiscovery (Premium)** page, select the **Cases** tab, and then select **Create a case**.

The **New eDiscovery case** flyout page is displayed. The **Case format** section on the **Name and description** page provides the option to create a case using the *New (recommended)* or *Classic* format.



4. After naming the case, select the **New (recommended)** option, and then select **Next** to continue in the new case workflow to finish creating your case.

After you've created your case, the case format information is displayed on the **Case format** card on the case **Overview** tab and in the **Case format** section on the case settings page for each case.

Frequently asked questions

If I attempt to collect over 1 TB in a single collection, will it work?

The performance will be negatively impacted and may cause instability in some instances.

If cloud attachments are included by default in the large case format; how can I remove that content from my review experience?

Use review set filters to filter by message kind or to exclude cloud attachments by using the HasAttachment filter. For more information, see Query and filter content in a review set.

When exporting chat conversation transcripts, will the load file contain all of the expanded metadata and a single item for each transcript?

All metadata for a conversation is embedded in the HTML transcript file. Many of the common fields are available in the load file. For more information about exported metadata, see Document metadata fields in eDiscovery (Premium).

Last updated on 11/21/2025

Work with custodians and non-custodial data sources in eDiscovery (Premium)

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic **Content Search**, classic **eDiscovery (Standard)**, and classic **eDiscovery (Premium)**.
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the **new eDiscovery experience** in the **Microsoft Purview portal**.

When an organization responds to a legal investigation, the workflow around identifying, preserving, and collecting potentially relevant content is based on the people in the organization who are the custodians of relevant data. In eDiscovery, these individuals are called *data custodians* (or just *custodians*) and are defined as "persons having administrative control of a document or electronic file." For example, the custodian of an email message could be the owner of the mailbox that contains the relevant message.

Additionally, there might be content located in mailboxes and sites that aren't associated with a custodian but that's relevant to the case. Content locations where case custodians don't have administrative control but might be owners of relevant data, are known as *non-custodial data sources*.

In an eDiscovery (Premium) case, legal teams can add individuals in their organization as custodians, and identify and preserve custodial data sources such as Exchange mailboxes, OneDrive accounts, and SharePoint and Teams sites. They can also identify and preserve non-custodial data sources. By using the built-in custodian and data source management tool in eDiscovery (Premium), organizations can secure electronically stored information from inadvertent (or intentional) deletion. This tool lets you eliminate the time-consuming and error-prone process of manually having to perform the legal hold processes.

For more information about working with custodians, see the following articles:

- Add custodians to a case
- Bulk-add custodians to a case

- Manage custodians in a case
- View custodian activity
- Add non-custodial data sources to a case

Last updated on 04/01/2026

Add custodians to an eDiscovery (Premium) case

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

Use the built-in custodian management tool in Microsoft Purview eDiscovery (Premium) to coordinate your workflows around managing custodians and identifying relevant, custodial data sources associated with a case. When you add a custodian, the system can automatically identify and place a hold on their Exchange mailbox and OneDrive for Business account. During the discovery process of your investigation, you might also identify other data sources (such as mailboxes, sites, or Teams) that a custodian accessed or contributed to. In this situation, you can use the custodian management tool to associate those data sources will a specific custodian. After you add custodians to a case and associate other data source with them, you can quickly preserve data and search the custodial data.

You can add and manage custodians in eDiscovery (Premium) cases in four steps:

1. Identify the custodians.
2. Choose custodian data locations.
3. Configure hold settings.
4. Review the custodians and complete the process.

Make sure you have the necessary permissions

To add custodians to a case, you must be a member of the eDiscovery Manager role group. This provides you with the necessary permissions to add custodians to a case and place a hold on the custodial data sources. For more information, see Assign eDiscovery permissions.

Step 1: Identify custodians

1. Go to https://compliance.microsoft.com and sign in with a user account that has been assigned the appropriate eDiscovery permissions.
2. In the left navigation pane, select **eDiscovery** > **eDiscovery (Premium)**, and select the **Cases** tab.
3. Select the case that you want to add custodians to.
4. Select the **Data sources** tab, and then select **Add data source** > **Add new custodians**.
5. Add one or more users in your organization as custodians to the case by typing the first part of a person's name or alias. Both active and inactive users are searched. After you find the correct person, select their name to add them to the list.

Step 2: Choose custodian data locations

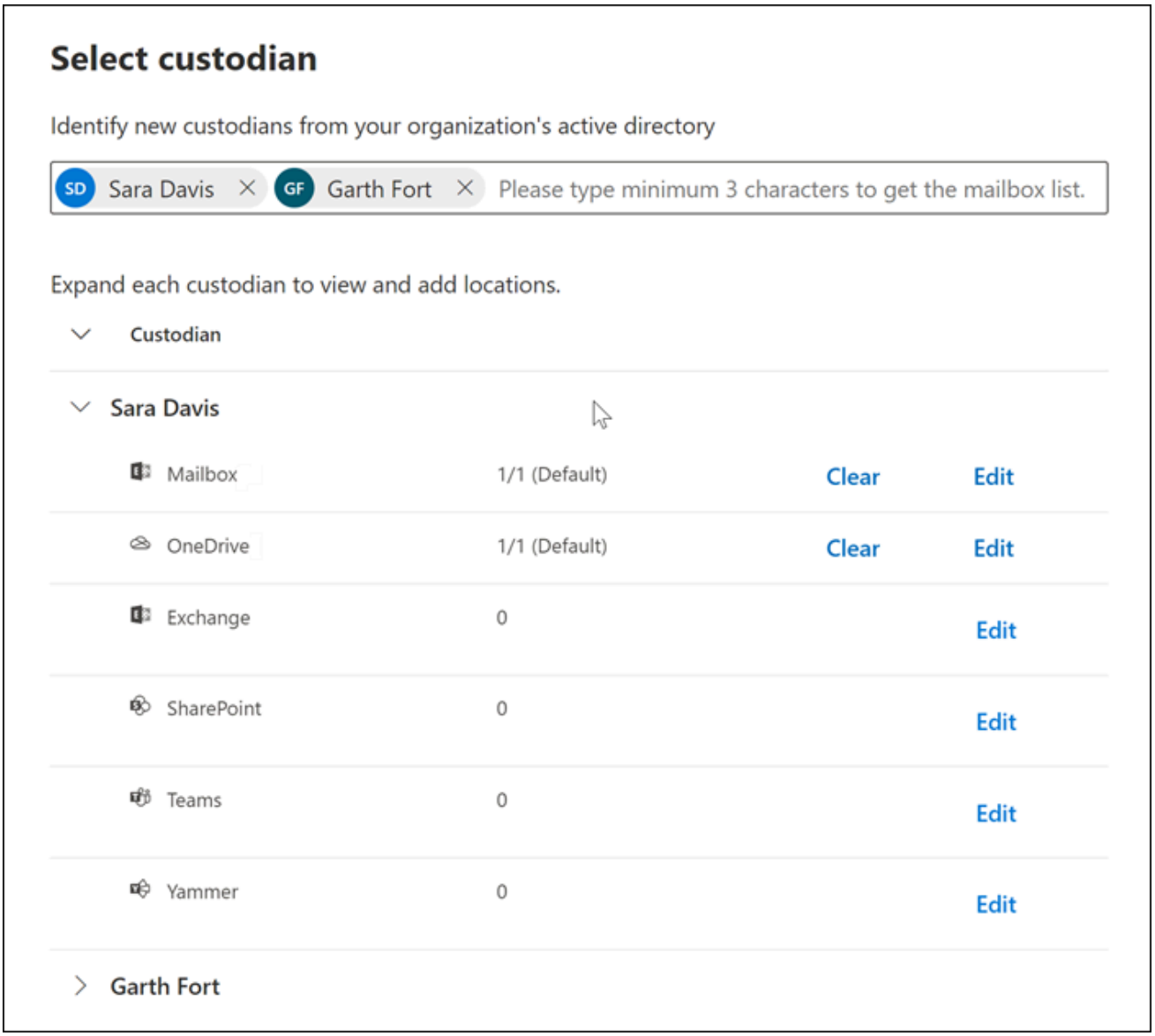
After you select custodians, the system automatically attempts to identify and verify these users and their data sources. After adding custodians to the list, the tool automatically includes the primary mailbox and OneDrive account for each active user that has been added as a custodian. If the user is inactive, the tool will only identify the primary mailbox. You can choose not to include these data sources when adding custodians to the case.

In addition to a custodian's mailbox and OneDrive account, you can also associate other data locations to a custodian, such as SharePoint site or a Microsoft Team the custodian is a member of. This allows you to preserve, collect, analyze, and review content in other data sources associated with the custodians of the case.

To deselect the primary mailbox and OneDrive account for a custodian:

1. Expand the custodian to view the primary data locations that have been automatically associated to each custodian.

2. Select **Clear** next to **Mailbox** or **OneDrive** to remove a custodian's mailbox or OneDrive account from being associated as a data location for this custodian.



To associate other mailboxes, sites, Teams, or Viva Engage groups to a specific custodian:

1. Expand a custodian to display the following services to associate data locations with the custodian. Select **Edit** next to a service to add a data location.

    - **Exchange**: Use to associate other mailboxes to the custodian. Type into the search box the name or alias (a minimum of three characters) of user mailboxes or distribution groups. Select the mailboxes to assign to the custodian and then select **Add**.
    - **SharePoint**: Use to associate SharePoint sites to the custodian. Select a site in the list or search for a site by typing a URL in the search box. Select the sites to assign to the custodian and then select **Add**. If a user is inactive, their OneDrive site will need to be added as an additional SharePoint location here.
    - **Teams**: Use to assign the Microsoft Teams the custodian is currently a member of. Select the teams to assign to the custodian and then select **Add**. After you add a team, the system automatically identifies and locates the SharePoint site and group mailbox associated to that team and assigns them to the custodian.

- **Viva Engage**: Use to assign the Viva Engage groups the custodian is currently a member of. Select the groups to assign to the custodian and then select **Add**. After you add a team, the system automatically identifies and locates the SharePoint site and group mailbox associated to that group and assigns them to the custodian.

> ⓘ Note
>
> You can use the **Exchange** and **SharePoint** location pickers to associate any mailbox or site in your organization to a custodian. , This includes associating the mailbox and site for a Microsoft Team or Viva Engage group that a custodian is not a member of. To do this, you have to add both the mailbox and site associated with each team or Viva Engage group.

2. You can view the total number of mailboxes, sites, Teams, and Viva Engage groups assigned to each custodian by expanding each custodian in the table. When you've finalized the assigned data locations for each custodian, these associations will be maintained and used during the collection, processing, and review stages in the eDiscovery (Premium) workflow.

3. After adding custodians and configuring their data locations, select **Next** to go to the **Hold settings** page.

Step 3: Configure hold settings

After you've finalized the custodians and their data locations, you can place some or all of the custodians on hold. When you place a custodian on hold, all content in all content locations that are associated with the custodian is preserved until you remove the hold or release the custodian from the hold. In some cases, you may want to add custodians to a case without placing them on hold. For more information about managing holds, see Manage holds in eDiscovery (Premium).

To place the custodians and data sources on hold:

1. On the **Hold settings** page, you can apply a hold to individual custodians by selecting the checkbox under the **Hold** column. Alternatively, you can place all custodians on hold by selecting the **Hold** checkbox at the top of the column.

2. Verify the custodian hold selections and then select **Next**.

> ⓘ Note

If you don't place a hold on a custodian, the custodian and their associated data sources will be added to the case but the content in those data sources won't preserved by the hold that associated with the case.

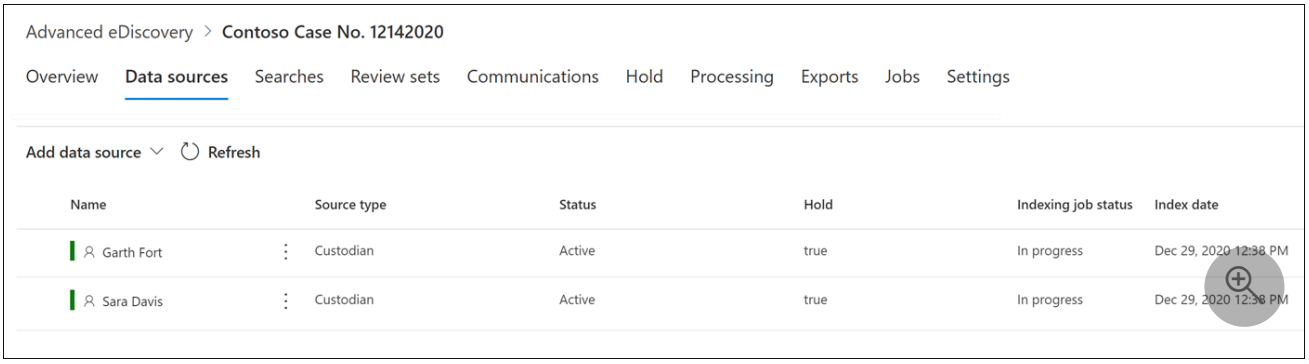
Step 4: Review the custodians and complete the process

Before you actually add the custodians to the case, you can review the list of custodians, the data locations assigned to them, and the hold settings.

1. Verify and review all the data sources count and the hold setting associated with each custodian in the table. If necessary, go back to the **Identify custodian** or **Hold settings** pages to make any changes.

2. Select **Submit** to add custodians and their data locations to the case and apply all custodial hold settings.

   The new custodians are added to the case and displayed on the **Data sources** tab.



Retry hold action

Once a custodian is placed on hold, you can validate the status in the **Hold** column of the list page. If the hold status isn't equal to *True* (when hold is enabled) or *False* (when hold is disabled), you can multi-select the relevant custodians and select *Retry hold action* at the top of the page. *Retry hold action* is only available for custodial sources and when multiple custodians are selected. If *Retry hold action* is performed for data sources successfully placed or released from hold, no action is triggered.

Last updated on 09/02/2025

Import custodians to an eDiscovery (Premium) case

> ⊗ **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic **Content Search**, classic **eDiscovery (Standard)**, and classic **eDiscovery (Premium)**.
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the **new eDiscovery experience** in the **Microsoft Purview portal**.

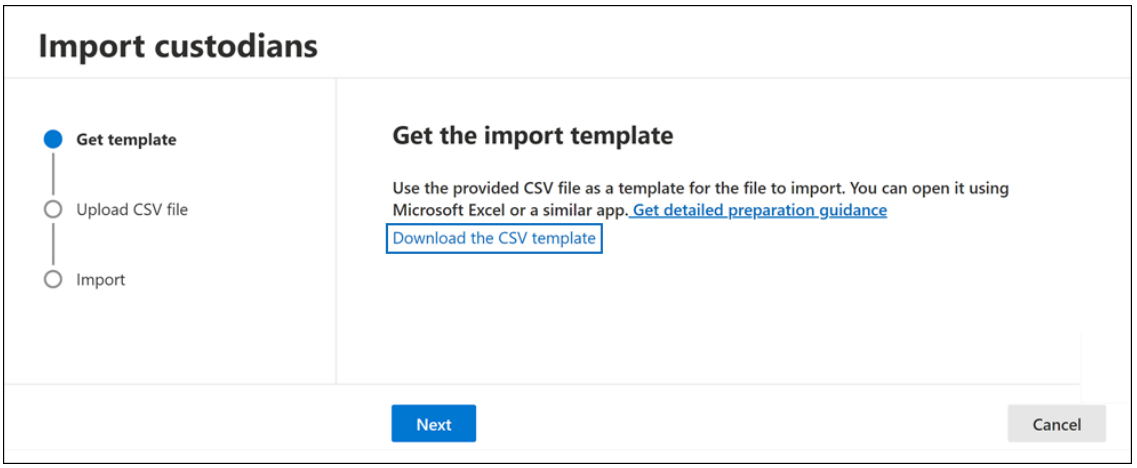
For Microsoft Purview eDiscovery (Premium) cases that involve many custodians, you can import multiple custodians at once by using a CSV file that contains the information necessary to add them to a case. The import custodians tool will also validate the CSV file before the import job is created. This means you can fix any errors in the CSV file instead of having to wait until the import job is complete before learning there are errors that prevent a custodian from being added to the case.

Before you import custodians

- You can import a maximum of 1,000 custodians (rows) per CSV file. Note that importing 1,000 custodians at the same time might result in timeout errors and some custodians might fail the import. To remediate this, repeat the import and the failed custodians should be imported. To avoid timeouts we recommend importing 200 custodians at a time.
- You can associate up to 500 data sources for each custodian.
- You can only import custodians that are part of your organization's Microsoft Entra ID.
- Each custodian must have a unique email address.
- To import an inactive mailbox as a custodian or to associate an inactive mailbox with another custodian, add a "." prefix to the email address of the inactive mailbox (for example, .sarad@contoso.onmmicrosoft.com).

Import custodians

1. Open the eDiscovery (Premium) case and select the **Data sources** tab.
2. Select **Add data source** > **Import custodians**.
3. On the **Get template** workflow page, select **Download the CSV template** to download a custodian template CSV file.

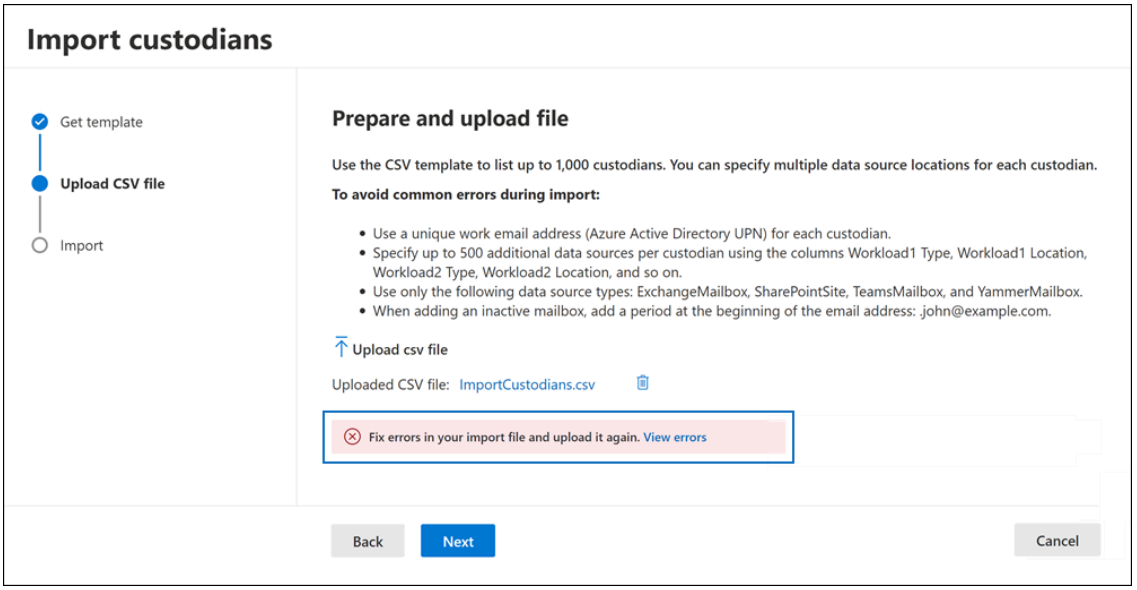


4. Add the custodial information to the CSV file and save it to your local computer. See the Custodian CSV file section for detailed information about the required properties in the CSV file.

5. After you've prepared the CSV file with the custodian information, go back to the **Data sources** tab, and select **Add data source** > **Import custodians** again.

6. On the **Upload CSV file** workflow page, select **Upload csv file** and then upload the CSV file that contains the custodian information.

   After you upload the CSV file, the import workflow validates the CSV file. If any validation errors exist, the workflow displays an error banner with a link to view the errors.



   The error information identifies the row and column of the cell that contains the error, and suggests a remediation action. You have to fix any validation error and then reupload the fixed CSV file. The CSV file must be successfully validated before you can create the import custodian job.

7. Once the CSV file has been successfully validated, select **Next** and then select **Import** to start the import job.

After you start the import job, eDiscovery (Premium) does the following things:

- Creates a job named *Import custodians* on the **Jobs** tab of the case. You can select the *Import custodians* job on the **Jobs** tab to view the status and details of the import custodian job under the **Support information** section. For more information, see Manage jobs.
- Performs Advanced indexing of all data sources for each custodian.
- Places all custodian data sources on hold (if the **Is OnHold** property in the CSV file is set to TRUE)

When the import custodian job is complete, the custodians and their associated data sources are added to the **Data sources** page of the case.

Custodian CSV file

After you download the CSV custodian template, you can add custodians and their data sources in each row. Be sure not to change the column names in the header row. Use the workload type and workload location columns to associate other data sources to a custodian.

Expand table

Column name	Description
Custodian contactEmail	The custodian's UPN email address. For example, sarad@contoso.onmicrosoft.com.
Exchange Enabled	TRUE/FALSE value to include or not include the custodian's mailbox.
OneDrive Enabled	TRUE/FALSE value to include or not include the custodian's OneDrive for Business account.
Is OnHold	TRUE/FALSE value to indicate whether to place the custodian data sources on hold. [1]
Workload1 Type	String value indicating the type of data source to associate with the custodian. Possible values include: - ExchangeMailbox - SharePointSite - TeamsMailbox[2] - Viva EngageMailbox[2]. The previous values for these workload types are case sensitive. The CSV file contains columns for three workload types and their corresponding workload locations. You can add a total of 500 workload types and locations.
Workload1 Location	Depending on your workload type, this would be the location of the data source. For example, the email address for an Exchange mailbox or the URL for a SharePoint site.

> **Note**
>
> [1] If you put more than 1,000 mailboxes or 100 sites on hold in a case, the system will automatically scale the eDiscovery hold as needed. This means the system automatically adds data locations to multiple hold policies, instead of adding them to a single policy. However, the limit of 10,000 case hold policies per organization still applies. For more information about hold limits, see Limits in eDiscovery (Premium).

[2] When you include TeamsMailbox and Viva EngageMailbox workloads in the CSV file, the group site (TeamSite and Viva EngageSite) are automatically added by default. You don't need to specify TeamsSite and Viva EngageSite separately in the CSV file.

Here's an example of a CSV file with custodian information:

Expand table

Custodian contactEmail	Exchange Enabled	OneDrive Enabled	Is OnHold	Workload1 Type	Workload1 Location
robinc@contoso.onmicrosoft.com	TRUE	TRUE	TRUE	SharePointSite	`https://contoso.sharepoint.c`
pillarp@contoso.onmicrosoft.com	TRUE	TRUE	TRUE		
johnj@contoso.onmicrosoft.com	TRUE	TRUE	TRUE		
sarad@contoso.onmicrosoft.com	TRUE	TRUE	TRUE	ExchangeMailbox	.saradavis@contoso.onmicros

Note

As previously explained, add a "." prefix to the UPN address of an inactive mailbox to import an inactive mailbox as a custodian or to associate an inactive mailbox with another custodian.

Retry hold action

Once a custodian is placed on hold, you can validate the status in the **Hold** column of the list page. If the hold status isn't equal to *True* (when hold is enabled) or *False* (when hold is disabled), you can multi-select the relevant custodians and select *Retry hold action* at the top of the page. *Retry hold action* is only available for custodial sources and when multiple custodians are selected. If *Retry hold action* is performed for data sources succesfully placed or released from hold, no action is triggered. For more information about managing holds, see Manage holds in eDiscovery (Premium).

Last updated on 04/22/2026

Manage custodians in an eDiscovery (Premium) case

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

The **Data sources** tab in a Microsoft Purview eDiscovery (Premium) case contains a list of all custodians that have been added to the case. After you add custodians to a case, details about each custodian are automatically collected from Microsoft Entra ID and are viewable in eDiscovery (Premium).

View custodian details

The **Data Sources** list view can be filtered by *Source name*, *Source type* (custodian/data location), *Locations*, *Source status*, *Indexing status*, *Hold status*, and *Index date*. The list view can also be grouped by *Data source*, *Source type*, *Source status*, *Hold status*, and *Indexing status*.

To view the details about a custodian, select the custodian from the list on the **Data sources** tab. A flyout page is displayed with the following tabs and information:

- Summary tab
  - Employee information
  - Statuses
- Locations tab
- Import errors

Additionally, the command bar for the custodian flyout page enables the following actions for the custodian:

- **Edit source**: Opens the **Edit custodian** page where you can add or remove data sources associated with the custodian.
- **Update index**: Updates the index for the custodian's data sources. This action is useful when you want to reindex the custodian's data sources to address partially indexed items.
- **Release**: Releases the custodian from the case. This action releases any holds placed on the custodian's data sources through the default hold policies for custodian/non-custodian sources. Any holds placed on the custodian's data sources in other cases or through manual policies (Query base holds) still apply.
- **View activity**: Opens the **View activity** page where you can search for specific activity for the custodian in the audit log.
- **Retry hold action**: Retries the hold for the selected custodians and locations. If the hold option was selected while adding/editing the source, the hold is applied again. If the hold option was not selected while adding/editing the source, the hold is removed.

Summary tab

The **Summary** tab shows the following sections and information about the custodian:

Employee information

- **Title**: The custodian's job title.
- **User principal name**: The user principal name (UPN) for the custodian. For example, user1@contoso.onmicrosoft.com.
- **Manager**: The custodian's manager. The designated manager receives any escalation communications for this custodian.
- **Mailbox**: The custodian's primary mailbox.
- **Location**. The office location in the custodian's place of business.
- **City**: The city in which the custodian works.
- **State**: The state in which the custodian works.
- **Country**: The country/region in which the custodian works.
- **Department**: The name for the department in which the custodian works.

> **Important**
>
> If the UPN of a custodian changes after the custodian is added to a case, the custodian information (Title, Manager, Location, etc.) isn't retained and is displayed in the custodian

summary pane as *No data to show*.

Statuses

- **Hold status**: The hold status of the locations added to the custodian. The hold statuses include:
  - *On hold*: Indicates the custodian and associated locations is successfully placed on hold.
  - *Not on hold*: Indicates the custodian and associated data locations are either not placed on hold or successfully released from hold.
  - *Applying hold*: Indicates the application of hold to the custodian and associated data locations is in progress. Wait until the hold job completes and verify that the hold is placed successfully.
  - *Releasing hold*: Indicates that the release of hold from the custodian and associated data locations is in progress. Wait until the hold job completes and verify if the hold is released successfully.
  - *Partial hold*: Indicates errors for some locations associated with the custodian couldn't be placed on hold.
  - *Partial hold released*: Indicates some locations associated with the custodian weren't released from hold.
  - *Hold failed*: Indicates the hold couldn't be placed on the custodian and any of the associated locations.
  - *Release failed*: Indicates the hold couldn't be released from the custodian and any of the associated locations.

    If the hold couldn't be placed or released successfully, the total count of failed locations with associated error categories and their resolution are displayed in the **Summary** tab. Select **Show locations** to view the full list of failed locations or select **Export list** to download a list of the failed locations by error type. For more information about data source errors, see Manage hold status and errors.

    > ⓘ **Note**

> The improved hold error details and resolution are available only for newly added custodians. A generic error message is displayed for older custodial locations that had previously failed. Users can get the improved error messaging and resolution information for these custodians by selecting **Retry hold action**. This retries the hold and generates the improved error messaging and resolution if the error persists.

- **Indexing status**: Indicates the indexing state, including the last updated date and time (UTC).

Locations tab

This tab shows the count and type of data sources (mailboxes, sites, and Teams) and the hold status for each location associated with the custodian and are part of the case.

- Each data location shows its hold status.
- If you don't see a hold status for a data source, check the status of the custodian hold that's listed on the **Hold** tab for the case. The custodian hold identifies the specific data sources place on hold.

Import errors tab

This tab shows the count and list of locations associated with the custodian that errored during the import custodian process and allows user to remove invalid locations from the sources.

Edit a custodian

As your case progresses, you may discover that there may be additional data sources relevant to a specific custodian and the case. In other scenarios, you may want to remove certain data sources that have been reviewed and deemed as not relevant.

To update the data sources that are associated with a custodian:

1. Go to **eDiscovery** > **eDiscovery (Premium)** and open the case.
2. Select the **Data Sources** tab.
3. Select a custodian from the list, and then select **Edit source** on the flyout page command bar.
4. To add or remove the primary mailbox and OneDrive account for the custodian:

- Expand the custodian to view the primary data locations that have been previously associated with the custodian.
- Select **Edit** next to **Mailboxes** or **OneDrive** to add the custodian's mailbox or OneDrive location.
- Select **Clear** next to **Mailboxes** or **OneDrive** to remove the custodian's mailbox or OneDrive account from being associated as a data location for this custodian.

5. To add or remove other mailboxes, sites, Teams, or Viva Engage groups to a specific custodian, select **Edit** next to the service to add a data location.

   - **Exchange**: Use to associate other mailboxes to the custodian. Type into the search box the name or alias (a minimum of three characters) of user mailboxes or distribution groups. Select the mailboxes to assign to the custodian and then select **Add**.
   - **SharePoint**: Use to associate SharePoint sites to the custodian. Select a site in the list or search for a site by typing a URL in the search box. Select the sites to assign to the custodian and then select **Add**.
   - **Teams**: Use to assign the Microsoft Teams the custodian is currently a member of. Select the teams to assign to the custodian and then select **Add**. After you add a team, the system automatically identifies and locates the SharePoint site and group mailbox associated to that team and assigns them to the custodian.
   - **Viva Engage**: Use to assign the Viva Engage groups the custodian is currently a member of. Select the groups to assign to the custodian and then select **Add**. After you add a team, the system automatically identifies and locates the SharePoint site and group mailbox associated to that group and assigns them to the custodian.

   > ⓘ **Note**
   >
   > You can use the **Exchange** and **SharePoint** location pickers to associate any mailbox or site in your organization, including teams or Viva Engage groups that a custodian is not a member of, to a custodian. To do this, you have to add both the mailbox and site associated with each team or Viva Engage group.

6. After editing the data locations for the custodian, select **Next** to go to the **Hold settings** page.

7. On the **Hold settings** page, update the hold settings for the custodian. Select **Next**.

8. On the **Review** page, review the changes and select **Submit** to save the changes.

Reindex custodian data

In most eDiscovery workflows for legal investigations, a subset of a custodian's data is searched after the custodian is added to a legal case. Because of large file sizes or possible data corruption, some items in the data sources associated with a custodian may be partially indexed. Using the Advanced indexing capability in the eDiscovery (Premium), most partially indexed items can be automatically remediated by reindexing these items on demand.

When a custodian is added to a case, the data located in the data sources associated with the custodian is automatically reindexed (by the Advanced indexing process). This means you can leave the data in-place instead of having to download and remediate it and then search it offline. However, during the lifecycle of a legal case new data sources might be associated with a custodian. In this case, you can reindex the custodian's data by rerunning the Advanced indexing process to remediate any partially indexed items and update the index for the custodian's data.

To trigger the reindexing process to address partially indexed items:

1. Go to **eDiscovery** > **eDiscovery (Premium)** and open the case.
2. Select the **Data sources** tab.
3. On the **Data sources** page, select a custodian whose data must be reindexed.
4. On the flyout page, select **Update index**.A dialog is displayed to confirm the index update, select **Yes** to continue.

Reindexing custodian data is a long-running process; the corresponding job that's created is named **Re-indexing custodian data**. You can track the progress on the **Jobs** tab or on the **Custodians** tab by monitoring the status in the **Indexing job status** column.

For more information, see:

- Work with processing errors
- Manage jobs

Release a custodian from a case

A custodian is released in situations where a case is closed, the custodian is no longer under obligation to preserve content for a case, or when the custodian is deemed to no longer be relevant to the case.

If you release a custodian after a hold notice was published, a release notice will be sent to the custodian. Additionally, any holds placed on data sources that were associated with the custodian

are removed. If the custodian was placed on a *silent hold*, where they weren't issued any legal hold notifications, a release notice won't be sent but any holds placed on data sources that were associated with that custodian are removed.

To release a custodian:

1. Go to **eDiscovery** > **eDiscovery (Premium)** and open the case.
2. Select the **Data sources** tab.
3. On the **Data sources** page, and then select the custodian who is being released from the case.
4. On the flyout page, select **Release**.

   A warning page is displayed explaining that if a hold is placed on a data source associated with the custodian, the hold is removed, and that any other hold associated with a different eDiscovery (Premium) case still apply. That includes other types of preservation and retention features (such as a Microsoft 365 retention policy).
5. Select **Yes** to confirm that you want to release the custodian.

   The status for this user on the **Custodians** tab is set to **Released** and the **Hold status** on the flyout page is changed to **False**.

> **Note**
>
> A custodian may be simultaneously involved in several legal cases. When a custodian is released from a specific case, the holds and notifications across other cases won't be impacted.

Last updated on 09/02/2025

View custodian audit activity

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

Need to find if a user viewed a specific document or purged an item from their mailbox? Microsoft Purview eDiscovery (Premium) is now integrated with the existing audit log search tool in the Microsoft Purview portal. Using this embedded experience, you can use the eDiscovery (Premium) Custodian Management tool to facilitate your investigation by easily accessing and searching the activity for custodians within your case.

Get permissions

You have to be assigned the *View-Only Audit Logs* or *Audit Logs* role in Exchange Online to search or export the audit log. By default, these roles are assigned to the *Compliance Management* and *Organization Management* role groups on the **Permissions** page in the Exchange admin center. To give a user the ability to search the eDiscovery (Premium) audit log with the minimum level of privileges, you can create a custom role group in Exchange Online, add the *View-Only Audit Logs* or *Audit Logs* role, and then add the user as a member of the new role group. For more information, see Manage role groups in Exchange Online.

> **Important**
>
> If you assign a user the *View-Only Audit Logs* or *Audit Logs* role on the Permissions page in the Microsoft Purview portal, they won't be able to search or export the audit log. You have to assign the permissions in Exchange Online. This is because the underlying cmdlet used to search the audit log is an Exchange Online cmdlet.

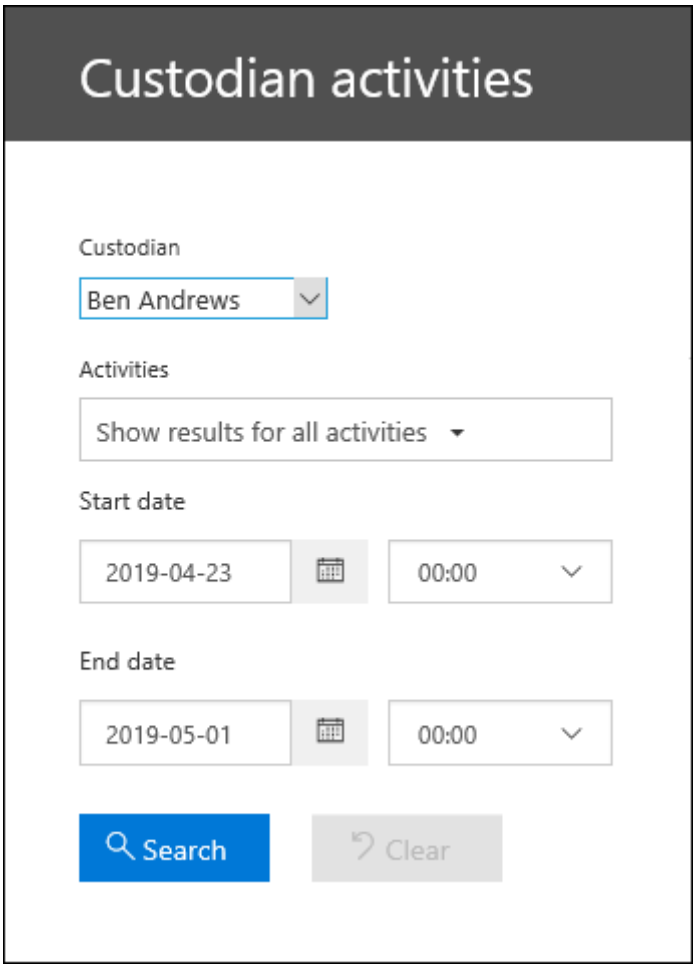
Step 1: Search the audit log for activities performed by a custodian

1. Go to **eDiscovery** > **eDiscovery (Premium)** and open the case.

2. Select the **Sources** tab.

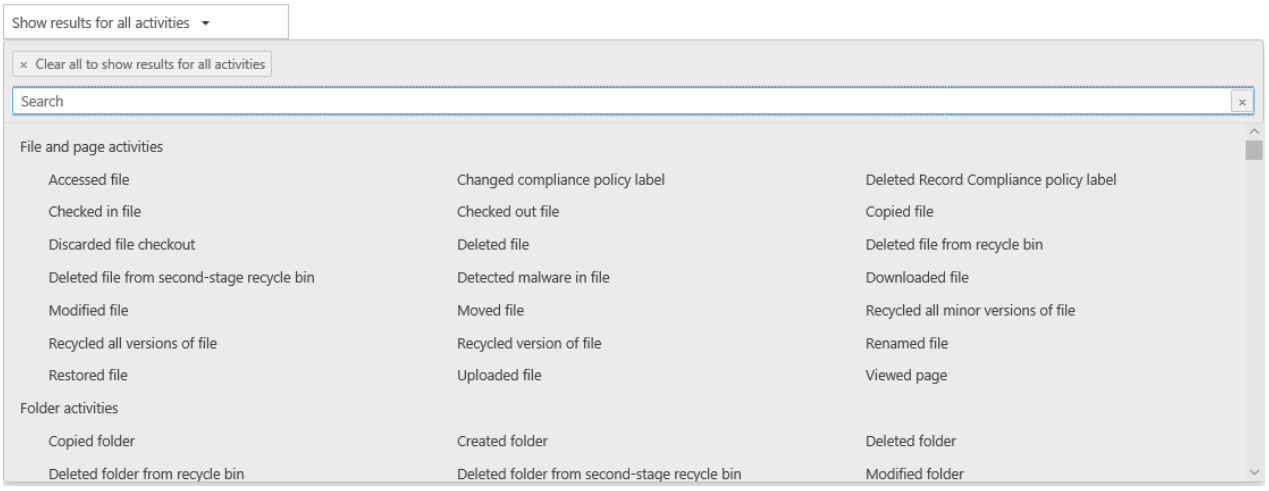
3. On the **Custodians** page, select a custodian from the list, and then select **View custodian activity** on the flyout page.

   The Custodian activities search page is displayed. Note the custodian you selected in the previous step is displayed in the **Custodian** drop-down box. You can select different custodians in the drop-down box, but you can only search for activities for one custodian at a time.



4. Configure the following search criteria:

   a. **Activities** - Select the drop-down list to display the activities that you can search for. After you run the search, only the audit records for the selected activities are displayed. Selecting **Show results for all activities** will display results for all activities performed by the custodian that match the other search criteria.



b. **Start date and End date** - Select a date and time range to display the events that occurred within that period. The last seven days are selected by default. The date and time are presented in Coordinated Universal Time (UTC) format. The maximum date range that you can specify is one year.

c. **Custodians** - Select in this box and then select a specific custodian to display search results for. Audit records for the selected activity performed by the users you select in this box are displayed in the list of results.

5. Select 

to run the search using your search criteria. The search results are loaded, and after a few moments they are displayed under Results on the Custodian Activities search page.

Step 2: View the audit log search results

The results of an audit log search are displayed under Results on the Custodian Audit log page. A maximum of 5,000 (newest) events are displayed in increments of 150 events. To display more events you can use the scroll bar in the Results pane, or you can press Shift + End to display the next 150 events.

The results contain the following information about each event returned by the search.

- **Date**: The date and time (in UTC format) when the event occurred.
- **IP address**: The IP address of the device that was used when the activity was logged. The IP address is displayed in either an IPv4 or IPv6 address format.
- **User**: The user (or service account) who performed the action that triggered the event.
- **Activity**: The activity performed by the user. This value corresponds to the activities that you selected in the Activities drop down list. For an event from the Exchange admin audit log, the value in this column is an Exchange cmdlet.

- **Item**: The object that was created or modified as a result of the corresponding activity. For example, the file that was viewed or modified or the user account that was updated. Not all activities have a value in this column.
- **Detail**: Additional detail about an activity. Again, not all activities will have a value.

Step 3: Filter the search results

In addition to sorting, you can also filter the results of an audit log search. This can help you quickly filter the results for a specific user or activity.

To filter the results:

1. Create and run an audit log search.
2. When the results are displayed, select **Filter results**.
3. Keyword boxes are displayed under each column header.
4. Select one of the boxes under a column header and type a word or phrase, depending on the column you're filtering on. The results will dynamically readjust to display the events that match your filter.
5. To clear a filter, select the **X** in the filter box or just select **Hide filtering**.

Export the search results to a file

You can export the results of an audit log search to a comma separated value (CSV) file on your local computer. You can open this file in Microsoft Excel and use features such as search, sorting, filtering, and splitting a single column (that contains multi-value cells) into multiple columns.

1. Run an audit log search, and then revise the search criteria until you have the desired results.
2. Select Export results and select one of the following options:
    - **Save loaded results:** Choose this option to export only the entries that are displayed under **Results** on the **Custodian Audit log search** page. The CSV file that is downloaded contains the same columns (and data) displayed on the page (Date, User, Activity, Item, and Details). An additional column (titled **More**) is included in the CSV file that contains more information from the audit log entry. Because you're exporting

the same results that are loaded (and viewable) on the Audit log search page, a maximum of 5,000 entries are exported.

- **Download all results:** Choose this option to export all entries from the audit log that meet the search criteria. For a large set of search results, choose this option to download all entries from the audit log in addition to the 5,000 results that can be displayed on the **Custodian Audit log** search page. This option will download the raw data from the audit log to a CSV file, and contains additional information from the audit log entry in a column named AuditData. It may take longer to download the file if you choose this export option because the file may be much larger than the one that's downloaded if you choose the other option.

  > **Important**
  >
  > You can download a maximum of 50,000 entries to a CSV file from a single audit log search. If 50,000 entries are downloaded to the CSV file, you can probably assume there are more than 50,000 events that met the search criteria. To export more than this limit, try using a date range to reduce the number of audit log entries. You might have to run multiple searches with smaller date ranges to export more than 50,000 entries.

3. After you select an export option, a message is displayed at the bottom of the window that prompts you to open the CSV file, save it to the Downloads folder, or save it to a specific folder

For more information about viewing, filtering, or exporting audit log search results, see Search the audit log.

Last updated on 09/02/2025

Add non-custodial data sources to an eDiscovery (Premium) case

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

In Microsoft Purview eDiscovery (Premium) cases, it doesn't always meet your needs to associate a Microsoft 365 data source with a custodian in the case. But you may still need to associate that data with a case so that you can search it, add it to a review set, and analyze and review it.

Non-custodial data sources in eDiscovery (Premium) let you add data to a case without having to associate it to a custodian. It also applies the same eDiscovery (Premium) functionality to non-custodial data that's available for data associated with custodian. Two of the most useful things that you can apply to non-custodial data is placing it on hold and processing it using Advanced indexing.

Add a non-custodial data source

Complete the following steps to add and manage non-custodial data sources in an eDiscovery (Premium) case.

1. In the Microsoft Purview portal, select **eDiscovery** > **Premium**.
2. Select the **Cases** tab and select the case that you want to add the data to.
3. Select the **Data sources** tab and then select **Add data source** > **Add data locations**.
4. On the **New non-custodial data locations** flyout page, choose the data sources that you want to add to the case. Select one or more of the following data source locations:

- **SharePoint**: Expand this section and select **Edit** to add SharePoint sites. Complete the following steps to add sites and configure holds:
  a. Select a site in the list or you can search for a site by typing the URL of the site in the search bar. Select the sites that you want to add as non-custodian data sources and select **Add**.
  b. After you add non-custodial data sources, you can place those locations on hold or not. Select or unselect the **Hold** checkbox next to the data source to place it on hold.
  c. Select **Add** at the bottom of the **New non-custodial data locations** flyout page to add the data sources to the case.

- **Exchange**: Expand this section and select **Edit** to add one or more Exchange mailboxes. Complete the following steps to add mailboxes and configure holds:
  a. Type a name or alias (a minimum of three characters) in the search box for mailboxes or distribution groups. Select the mailboxes that you want to add as non-custodian data sources and select **Add**.
  b. After you add non-custodial data sources, you can place those locations on hold or not. Select or unselect the **Hold** checkbox next to the data source to place it on hold.
  c. Select **Add** at the bottom of the **New non-custodial data locations** flyout page to add the data sources to the case.

  > **Note**
  >
  > You can use the **SharePoint** and **Exchange** sections to add sites and mailboxes associated with a Team or Viva Engage group as non-custodial data sources. You have to separately add the mailbox and site associated with a Team or Viva Engage group.
  >
  > Also, adding a root site URL (such as `https://contoso-my.sharepoint.com/personal/` or `https://contoso-my.sharepoint.com/`) as a SharePoint data source isn't supported. You have to add specific sites.

- **Teams**: Expand this section and select **Add Teams locations** to add one or more Microsoft Teams channels. Complete the following steps to add channels and configure holds:
  a. On the **Channels** page, enter a team name and select from the channels associated with the team. You can add up to 10 teams at one time and all the associated

private, standard, and shared channels. Select **Next**. This automatically selects the [storage locations] (/purview/ediscovery-teams-workflow#where-teams-content-is-stored) where the channel messages and files are stored.

b. On the **Configure holds** page, select the locations associated with each Team and channel to place on a hold. Select **Next**.

c. On the summary page, review the selected channel locations and hold configuration. Select **Add channels** to add the channels and configure holds.

d. The **Status** page confirms if the locations were added successfully or if the location already exists in the case. Select **Done**.

Each non-custodial data source that you added is listed on the **Data sources** page. Non-custodial data sources are identified by the *Data location* value in the **Source type** column.

After you add non-custodial data sources to the case, a job named *Reindexing non-custodial data* is created and displayed on the **Jobs** tab of the case. After the job is created, the Advanced indexing process in initiated and the data sources are reindexed. For more information about managing holds, see Manage holds in eDiscovery (Premium).

Release a hold for non-custodial data sources

After you place a hold on a non-custodial data source, a hold policy that contains the non-custodial data sources for the case is automatically created. When you place other non-custodial data sources on hold, they're added to this hold policy.

To release a hold on a non-custodial data source, complete the following steps:

1. In the Microsoft Purview portal, select **eDiscovery** > **Premium**.
2. Select the **Cases** tab and select the case that you want to add the data to.
3. Select the **Data sources** tab and the locations to release from hold.
4. Select **Release**.
5. Select **Yes** to release the hold for the location. Any holds placed on the data sources in other cases still apply.

Last updated on 11/21/2025

Work with communications in eDiscovery (Premium)

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

Microsoft Purview eDiscovery (Premium) allows legal departments to simplify their processes around tracking and distributing legal hold notifications. The custodian communications tool enables legal departments to manage and automate the entire legal hold process, from initial notifications, to reminders, and to escalations, all in one location.

What is a legal hold notification?

A legal hold (also known as a *litigation hold*) notice is a notification sent from an organization's legal department to employees, contingent staff, or custodians of data that may be relevant to a legal investigation. These notifications instruct custodians to preserve electronically stored information and any content that may be relevant to an active or impending legal matter. Legal teams must know that each custodian has received, read, understood, and has agreed to comply with the given instructions.

The legal hold notification process

An organization has a duty to preserve relevant information when it learns about an impending litigation or regulatory investigation. To comply with the preservation requirements of an investigation, the organization should immediately inform potential custodians about their duty to preserve relevant information.

With eDiscovery (Premium), legal teams can create and customize their legal hold notification workflow. The custodian communications tool lets legal teams configure the following notices

and workflows:

1. **Issuance notice:** A legal hold notice is issued (or initiated) by a notification from the legal department to custodians who may have relevant information about the case matter. This notice instructs the custodians to preserve any information that may be needed for discovery.

2. **Re-Issuance notice:** During a case, custodians may be required to preserve more content (or less content) than was previously requested. For this scenario, you can update the existing hold notice and reissue it to custodians.

3. **Release notice:** Once a matter is resolved and the custodian is no longer subject to a preservation requirement, the custodian can be released from the case. Additionally, you can notify the custodian that they're no longer required to preserve content, and provide instructions about how to resume their normal work activity and their data.

4. **Reminders and escalations:** In some instances, just issuing a notice isn't enough to satisfy legal discovery requirements. With each notification, legal teams can schedule a set of reminder and escalation workflows to automatically follow up with unresponsive custodians.

    - **Reminders:** After a legal hold notice has been issued or reissued to a set of custodians, an organization can set up reminders to alert unresponsive custodians.

    - **Escalations:** In some cases, if a custodian remains unresponsive even after a set of reminders over a period of time, the legal team can set up an escalation workflow to notify unresponsive custodians and their manager.

For more information about managing the custodian communication process, see the following:

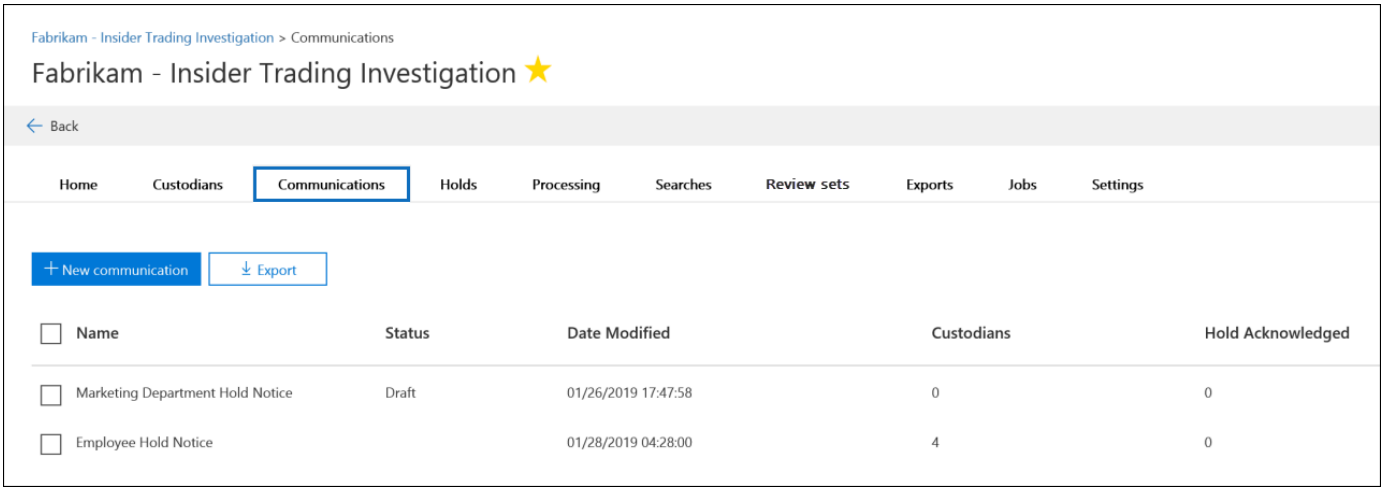
- Create a legal hold notice
- Use the communications editor
- Manage hold notifications
- Acknowledge a hold notification

Last updated on 11/21/2025

Create a legal hold notice

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

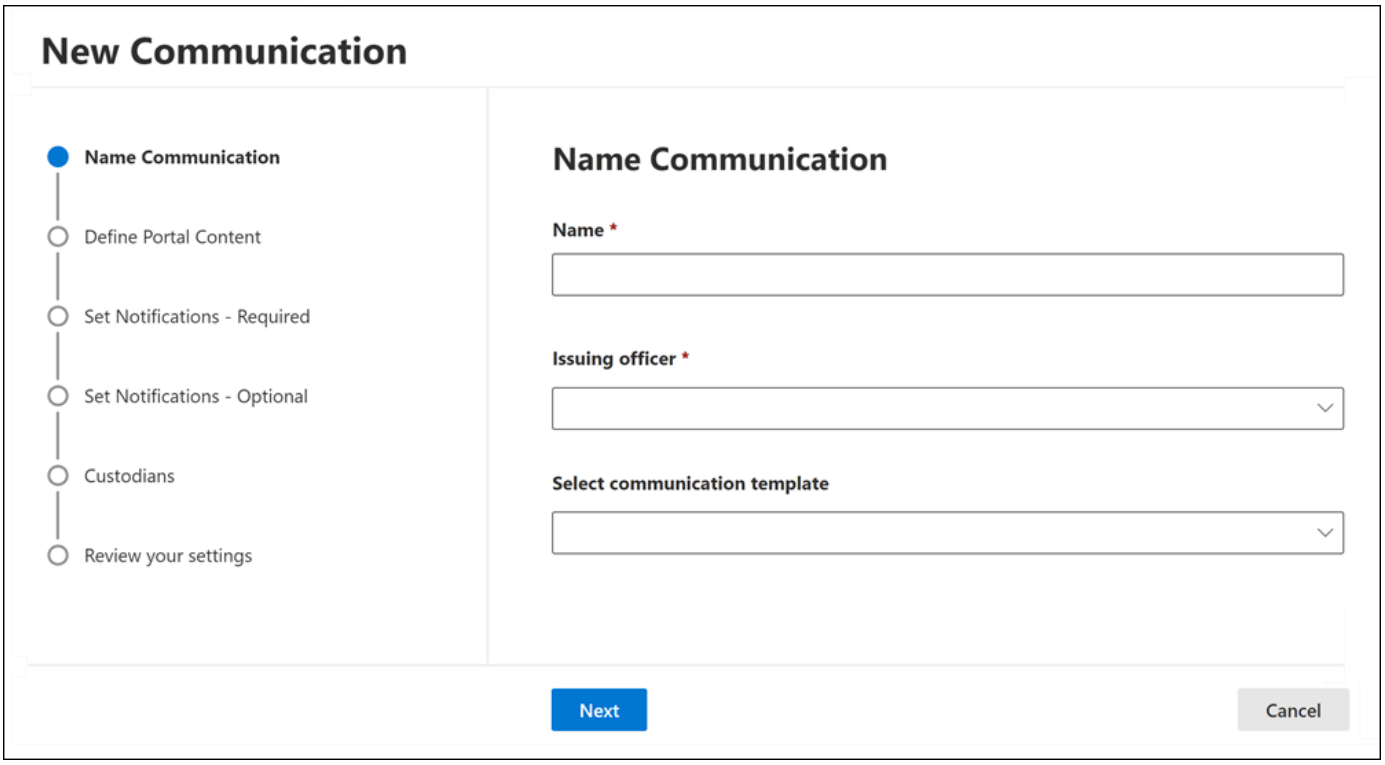
Using eDiscovery (Premium) custodian communications, organizations can manage their workflow around communicating with custodians. Through the Communications tool, legal teams can systematically send, collect, and track legal hold notifications. The flexible creation process also allows teams to customize the hold notification workflow and the content in the notices sent to custodians.



The article outlines the steps in the hold notification workflow.

Step 1: Specify communication details

The first step is to specify the appropriate details for legal hold notices or other custodian communications.



1. In the Microsoft Purview portal, go to **eDiscovery** > **Premium** to display the list of cases in your organization.

2. Select a case, select the **Communications** tab, and then select **New communication**.

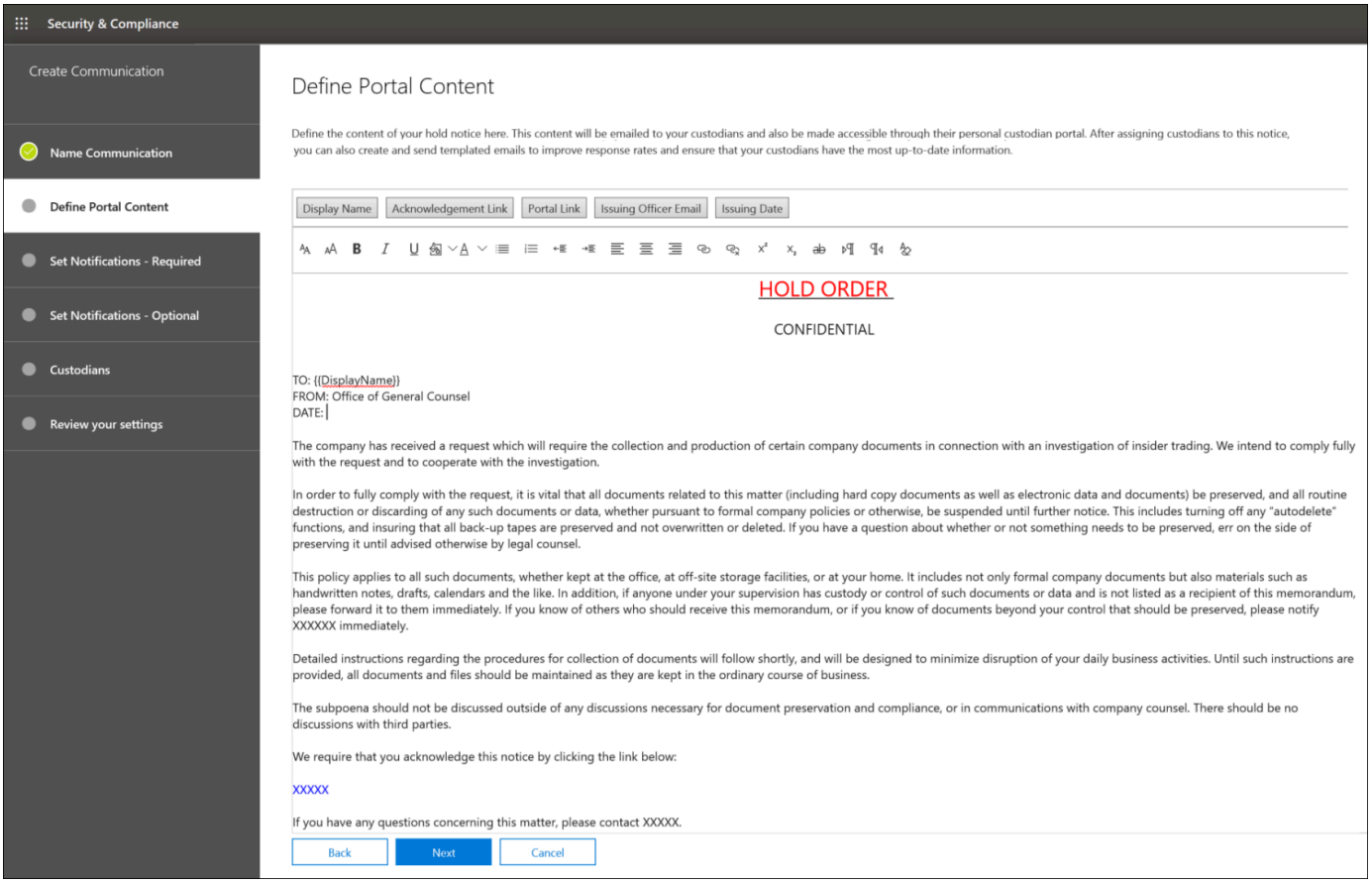
3. On the **Name communication** page, specify the following communication settings.

    - **Name**: This is the name for the communication.
    - **Issuing officer**: The drop-down list displays users in your organization who can be selected as the issuing officer for the communication. Each communication sent to custodians will be sent on behalf of the selected issuing officer. The list of users in the drop-down consists of the members of the case and the organization-wide issuing officers. These issuing officers are added by an eDiscovery Administrator, and are available in all eDiscovery (Premium) cases in your organization. For more information, see Manage issuing officers.
    - **Select communication template**: The drop-down list displays the templates from the Communications library on the eDiscovery (Premium) settings page. If you select a template, it will be displayed on the **Define portal content** as a starting point for the text of the notification that you're creating. If you don't select a template, then you'll have to create the notice yourself from scratch. For more information about communication templates, see Manage custodian communications templates.

4. Select **Next**.

Step 2: Define the portal content

Next, you can create and add the content of the hold notice. On the **Define portal content** page in the **Create communication** workflow, specify the contents of the hold notice. This content will be automatically appended to the Issuance, Re-Issue, Reminder, and Escalation notices. Additionally, this content will appear in the custodian's Microsoft Purview portal. If you selected a template from the Communications library, it will be displayed and provide a starting point for the notice you're creating.



To create the portal content:

1. Type (or cut and paste from another document) your hold notice in the textbox for the portal content. If you selected a communications template on the previous workflow page, the template is displayed. You can edit the template content as necessary.
2. Insert merge variables into your notice to customize the notice and share the Custodian in the Microsoft Purview portal.
3. Select **Next**.

> **Tip**

To learn more about how to can customize the content and format of the portal content, see **Use the Communications Editor**.

Step 3: Set the required notifications

After you've defined the contents of the hold notice, you can set up the workflows around sending and managing the notification process. Notifications are email messages that are sent to notify and follow up with custodians. Every custodian added to the communication will receive the same notification.

To set up and send a hold notice, you must include Issuance, Re-Issuance, and Release notifications.

Issuance notification

After the communication is created, the **Issuance Notification** is initiated by the specified Issuing Officer. The Issuance notification is the first communication sent to the custodian to inform them about their preservation obligations.

To create an issuance notification:

1. In the **Issuance** tile, select **Edit**.
2. If necessary, add additional case members or staff to the **Cc** and **Bcc** fields. To add multiple users to these fields, separate email addresses with a semi-colon and without spaces between the addresses. For example, *user1@contoso.com;user2@contoso.com;user3@contoso.com*.
3. Specify the **Subject** for the notice (required).
4. Specify the contents or additional instructions that you would like to provide to the custodian (required). The portal content you defined in Step 2 is added to the end of the issuance notice.
5. Select **Save**.

Re-Issuance notification

As the case progresses, custodians may be required to preserve additional or less data than was previously instructed. After you update the portal content, the reissuance notification is sent and alerts custodians about any changes to their preservation obligations.

To create a reissuance notification:

1. In the **Reissue** tile, select **Edit**.
2. If necessary, add additional case members or staff to the **Cc** and **Bcc** fields. To add multiple users to these fields, separate email addresses with a semi-colon and without spaces between the addresses. For example, *user1@contoso.com;user2@contoso.com;user3@contoso.com*.
3. Specify the **Subject** for the notice (required).
4. Specify the contents or additional instructions that you would like to provide to the custodian (required). The portal content you defined in Step 2 is added to the end of the reissuance notice.
5. Select **Save**.

> **Note**
>
> If the portal content is modified (on the **Define Portal Content** page in the **Edit communication** workflow), the re-issuance notification will be automatically sent to all custodians assigned to the notice. After the notification is sent, custodians will be asked to re-acknowledge their hold notice. If you have set up any reminder or escalation workflows, these will also re-start. For more information about what other case management events trigger communications, see Events that trigger notifications.

Release notification

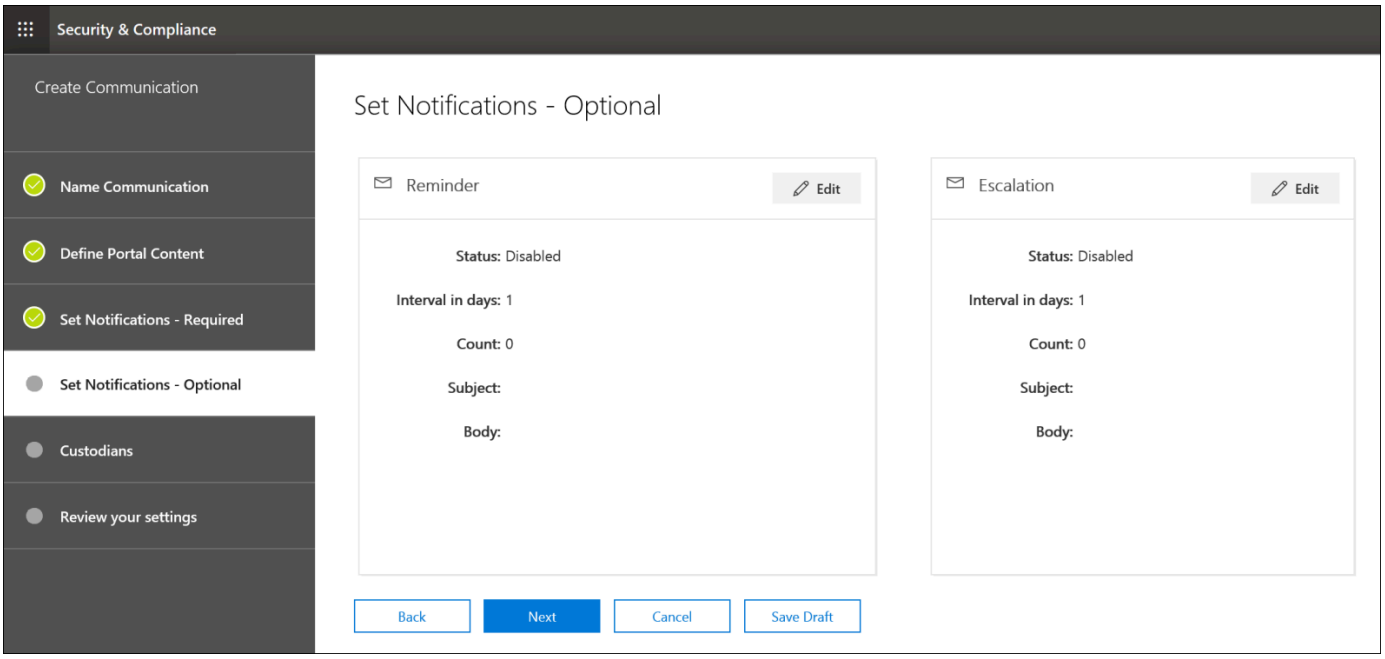
After a matter is resolved or if a custodian is no longer subject to preserve content, you can release the custodian from a case. If the custodian was previously issued a hold notice, the release notification can be used to alert custodians that they've been released from their obligation.

To create a release notification:

1. In the **Release** tile, select **Edit**.
2. If necessary, add additional case members or staff to the **Cc** and **Bcc** fields. To add multiple users to these fields, separate email addresses with a semi-colon and without spaces between the addresses. For example, *user1@contoso.com;user2@contoso.com;user3@contoso.com*.
3. Specify the **Subject** for the notice (required).
4. Specify the contents or additional instructions that you would like to provide to the custodian (required).
5. Select **Save** and go to the next step.

(Optional) Step 4: Set the optional notifications

Optionally, you can simplify the workflow for following up with unresponsive custodians by creating and scheduling automated reminder and escalation notifications.



Reminders

After you've sent a hold notification, you can follow up with unresponsive custodians by defining a reminder workflow.

To schedule reminders:

1. In the **Reminder** tile, select **Edit**.
2. Enable the **Reminder** workflow by turning on the **Status** toggle (required).
3. Specify the **Reminder interval (in days)** (required). This is the number of days to wait before sending the first and follow-up reminder notifications. For example, if you set the reminder interval to seven days, then the first reminder would be sent seven days after the hold notification was initially issued. All subsequent reminders would also be sent every seven days.
4. Specify the **Number of reminders** (required). This field specifies how many reminders to send to unresponsive custodians. For example, if you set the number of reminders to 3, then a custodian would receive a maximum of three reminders. After a custodian acknowledges the hold notification, reminders will no longer be sent to that user.
5. Specify the **Subject** for the notice (required).
6. Specify the contents or additional instructions that you would like to provide to the custodian (required). The portal content you defined in Step 2 is added to the end of the

reminder notice.
7. Select **Save** and go the next step.

Escalations

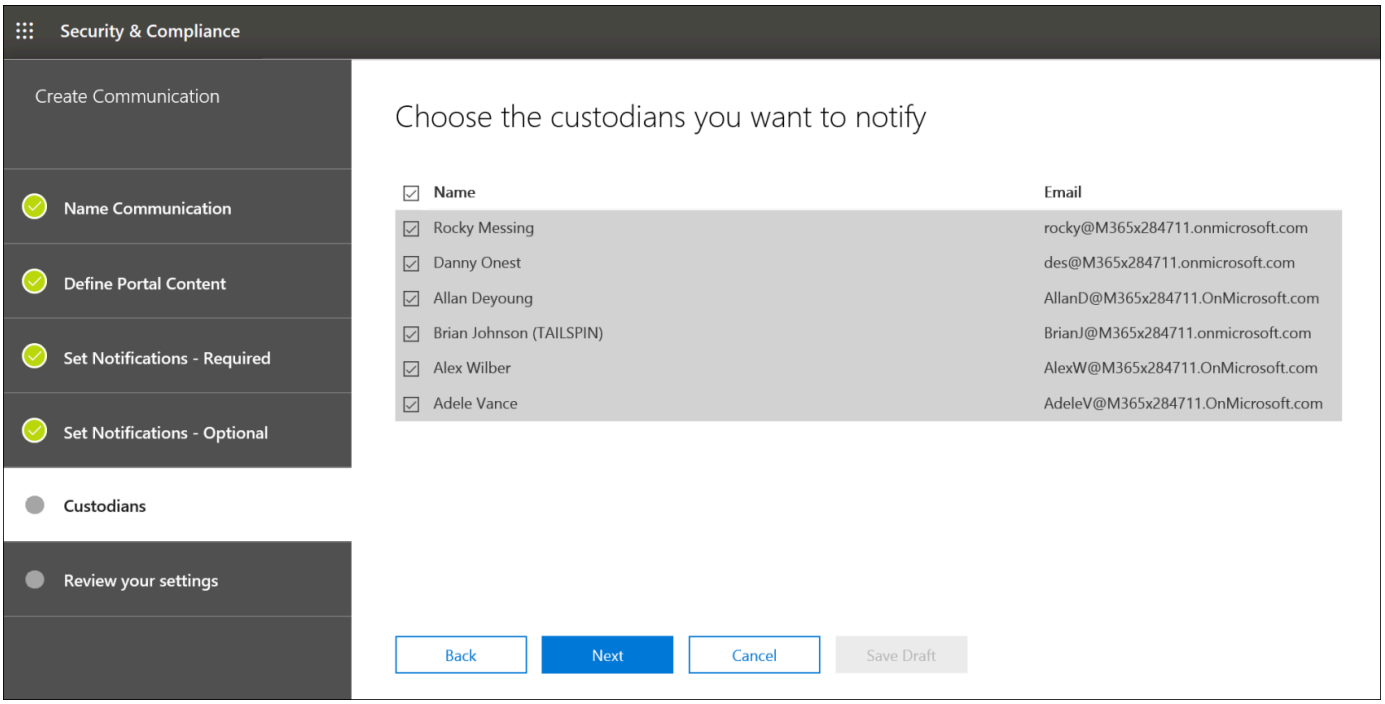
In some situations, you may need additional ways to follow up with unresponsive custodians. If a custodian doesn't acknowledge a hold notification after receiving the specified number of reminders, the legal team can specify a workflow to automatically send an escalation notice to the custodian and their manager.

To schedule escalations:

1. In the **Escalation** tile, select **Edit**.
2. Enable the **Escalation** workflow by turning on the **Status** toggle.
3. Specify the **Escalation interval (in days)** (required).
4. Specify the **Number of escalations** (required). This field specifies how many escalations to send to unresponsive custodians. For example, if you set the number of escalations to 3, then an escalation notice would be sent to the custodian and their manager a maximum of three times. After a custodian acknowledges the hold notification, escalations will no longer be sent.
5. Specify the **Subject** for the notice (required).
6. Specify the contents or additional instructions that you would like to provide to the custodian (required). The portal content you defined in Step 2 is added to the end of the escalation notice.
7. Select **Save** and go the next step.

Step 5: Assign custodians to receive notifications

After you've finalized the content for notifications, select the custodians that you would like to send notifications to.



To add custodians:

1. Assign custodians to the communication by selecting the checkbox next to their name.

   After the communication is created, the notification workflow will automatically apply to the selected custodians.

2. Select **Next** to review the communication settings and details.

> **Note**
>
> You can only add custodians who have been added to the case and haven't been sent another notification within the case. If the custodian is inactive or released, they can't be added to any new communications.

Step 6: Review settings

After you review the settings and select **Send** to complete the communication, the system will automatically start the communication workflow by sending the issuance notice.

Events that trigger notifications

The following table describes events in the case management process that trigger when the different types of notifications are sent to custodians.

Expand table

Type of communication	Trigger
Issuance notices	The initial creation of the notification. You can also manually resend a hold notification.
Reissuance notices	Updating the portal content on the **Define Portal Content** page in the **Edit communication** workflow.
Release notices	The custodian is released from the case.
Reminders	The interval and number of reminders configured for the reminder.
Escalations	The interval and number of reminders configured for the escalation.

Last updated on 09/02/2025

Use the communications editor

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

As you define the content of your portal content, legal hold notifications, and related reminders/escalations, you can use the Communications Editor to format and dynamically customize your content.

Rich text editor

The Communications Editor allows user to customize the text using the editor options. For example, users can change font types, create bulleted lists, highlight content, and more.

Merge field variables

You can use email merge variables from the Communications Editor to embed customized custodian attributes into a communication's body text. When sent to the custodian, the merge field will be populated with the corresponding field. For example, when sent to custodian John Smith, the merge field [Custodian Name] would be translated with the corresponding name.

You can use email merge fields by selecting the **Merge field** icons on the top of the rich-text editor control. The placeholder will be added based off the location of the users' cursor.

List of merge field variables

Expand table

Field name	Field details
Display Name	The custodian's first and last name.
Acknowledgment Link	A customized link to record each custodian's acknowledgment.
Portal Link	A customized link for the custodian's Microsoft Purview portal.
Issuing Officer	The email address of the specified issuing officer.
Issuing Date	The date that the notice was issued (UTC).

Last updated on 11/21/2025

Manage hold notifications

> **Important**
>
> Legal hold communications with custodians was permanently **retired on August 31, 2025**, this functionality isn't available in the new eDiscovery experience.

After you've initiated your legal hold notification workflow, you can use the communications workflow in Microsoft Purview eDiscovery (Premium) to track the status of your communications. The Communications tab contains a list of all notifications within your eDiscovery (Premium) case. You can see details such as the number of custodians that have been assigned or have acknowledged the notice.

Monitor acknowledgments

After you select a communication from the **Communications** tab, you can view a list of custodians that have acknowledged a hold notice.

1. In the Microsoft Purview portal, go to **eDiscovery** > **eDiscovery (Premium)**.
2. Select a case and then select the **Communications** tab.
3. Select a communication to display the **Custodian communication** flyout page.

A list of custodians associated with the selected communication is displayed on the communication flyout page. This page also displays insights and about how many acknowledgments were received and how many are outstanding. The page also shows which custodians have sent an acknowledgment that they received the hold notification.

Resend a hold notice

Occasionally, custodians lose track of email messages in their day-to-day work. Or for a long-running litigation case, a custodian may contact you or others and request that you resend a notice. As you manage the communications workflow for legal hold notices, you may need to resend a notice to bring it back to the "top of a user's mailbox".

To resend a hold notice to a custodian:

1. In eDiscovery (Premium), select a case and then select the **Communications** tab.

2. Select a communication to display the **Custodian communication** flyout page.
3. Select **More** > **Re-send hold notice**.
4. On the **Re-send hold notice** flyout page, select the custodians that you what to resend the notice and type an optional reason.
5. Select **Re-send** to send the notice to the selected custodians.

If a custodian hasn't acknowledged the hold notification, the reminder and escalation workflow is restarted. If a custodian has acknowledged the hold notice, the custodian will receive a copy of the original hold notice.

> ⓘ **Note**
>
> You can only resend a legal hold notification to custodians that are assigned to the communication.

Update preservation requirements

As the case progresses, custodians may be required to preserve additional or less data than was previously instructed. In eDiscovery terms, you need to reissue the hold notice with updated content.

To update the contents of the initial hold notice:

1. In eDiscovery (Premium), select a case and then select the **Communications** tab.
2. Select the hold notice that you want to update and select **Edit** on the **Custodian communication** flyout page.
3. In the **Edit Communication** workflow, select **Define Portal Content** in the left pane of the workflow, and update the contents of the notice.
4. Select **Save**.

The reissuance notice will be sent to all the custodians assigned to the legal hold notification. In addition, if the Reminder or Escalation notice is enabled, then the workflows for those types of notices will restart.

Update legal hold notifications and settings

When you update the content or settings of the Issuance, Release, Reissue, Reminder, or Escalation notice, these changes will apply to all future communications generated by the workflow.

More information

- Add custodians to a case
- Create a legal hold notice
- Acknowledge a hold notification

Last updated on 09/02/2025

Acknowledge a hold notification

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

When responding to a regulatory request or investigation, you might need to inform custodians of their obligation to preserve electronically stored information (ESI) and any material that might be relevant to an active or imminent legal matter. After you send the notification, legal teams need to know that each custodian receives, reads, understands, and agrees to follow the given instructions.

To help reduce the time, cost, and effort of following up with your custodians, eDiscovery (Premium) enables you to send and follow up on legal hold notifications through email. In addition to email notices, each custodian can access an individualized Microsoft Purview portal, so custodians stay informed of changes to their obligation status.

Email notifications

After you issue a Legal Hold Notification, each custodian receives a unique and personalized email containing your defined legal hold notice and added instructions.

> **Tip**
>
> To allow your custodians to acknowledge their notice or access their Microsoft Purview portal directly from their email, see how you can use the built-in Communication Editor.

Based on the configuration of your legal hold notification, your custodians might receive the following notices:

- **Issuance notice:** The first notice you send to your custodian. This notice contains your issuance instructions and the hold notice appended to the end of your message.
- **Reminder notice:** If you enable this notice, your custodians receive a reminder notice based on the specified frequency and interval. The reminders continue to be sent either until the custodian acknowledges their notice or until the number of reminders is exhausted.
- **Escalation notice:** If you enable this notice, your custodian and their manager receive an escalation notice after the reminder notices are exhausted. The system automatically sends escalation notices until the specified number of escalations is completed or until the custodian acknowledges their hold notification.
- **Reissue notice:** During an investigation, if you update the contents of the hold notice, the updated notice is automatically sent to the custodian.
- **Release notice:** When you release a custodian from the case, you send the release notice to the custodian.

Microsoft Purview portal

In addition to the email notifications, each custodian can access a unique Microsoft Purview portal. Through the Microsoft Purview portal, each custodian can view, access, and acknowledge their active hold notifications.

Last updated on 11/21/2025

Manage issuing officers in eDiscovery (Premium)

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

When you or others create a hold notification or other type of communication that is sent to a user who is a custodian in case, you have to specify an issuing officer. The notification is sent to the custodian on behalf of the specified issuing officer. For example, a paralegal in your organization might be responsible for creating and sending hold notifications to custodians in a case. In this scenario, the paralegal can specify an attorney in the organization as the issuing officer. Who can be specified as an issuing officer? There are two types of users who can be selected as an issuing officer for a custodian communication:

- Any member of the specific case the communication is being sent in behalf of.
- Any user who is added to a list of organization-wide issuing officers. Users from this list can be added an issuing officer to any case in your organization.

This article explains how to add and remove users to the list of organization-wide issuing officers.

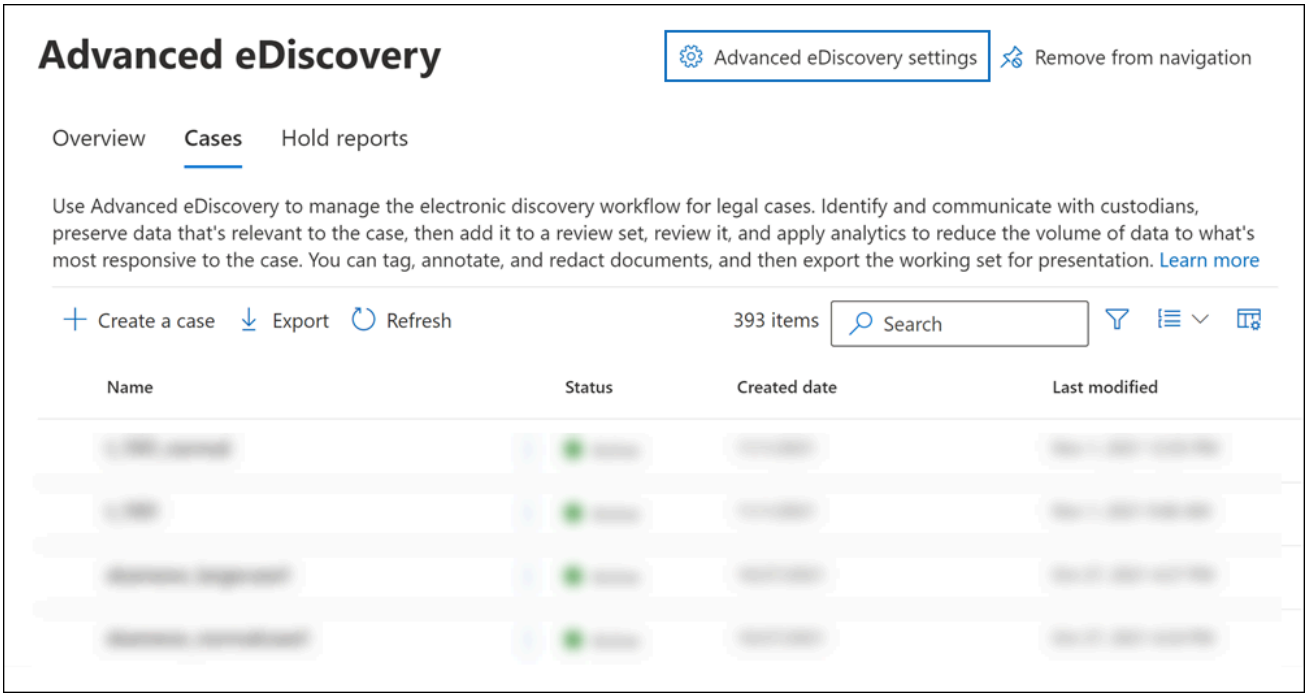
Before you add an issuing officer

- You must be an eDiscovery Administrator in your organization to add or remove issuing officers. For more information, see Assign eDiscovery permissions in the Microsoft Purview portal
- The user who is added as an issuing officer must have an active mailbox in your Microsoft 365 organization.
- Your organization can have a maximum of 15 issuing officers. Members of a case who can be specified as an issuing officer aren't counted toward this limit. This limit only applies to

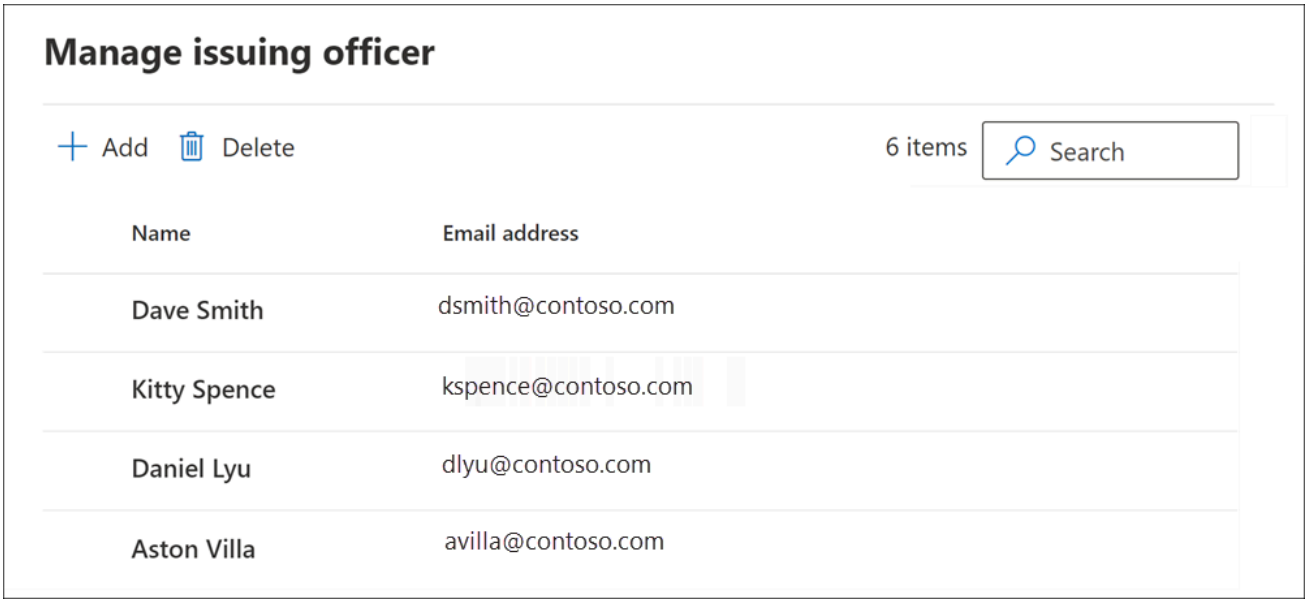
the number of users that can be added to the **Issuing officers** page in eDiscovery (Premium).

Add an issuing officer

1. In the Microsoft Purview portal, go to eDiscovery (Premium), and then select **eDiscovery (Premium) settings**.



2. On the **Settings** page, select the **Issuing officers** tab to display the **Manage issuing officers** page.



3. Select **Add** and then search for and add one or more users to the list of issuing officers.

After you add users as issuing officers, you or other users will be able to specify these users as an issuing officer for custodian communications for any case in your organization. For more information about creating custodian communications, see Create a legal hold notice.

Remove an issuing officer

1. In the Microsoft Purview portal, go to eDiscovery (Premium), and then select **eDiscovery (Premium) settings**.
2. On the **Settings** page, select the **Issuing officers** tab.
3. Select one or more users in the issuing officers list, and then select **Delete**.

After you delete users from the list of issuing officers, those users can no longer be specified as an issuing officer in new custodian communications, unless the user is a member of the specific case the communication is being issued from. Also, removing an issuing officer won't affect any communications that were sent before the user was removed as an issuing officer.

Last updated on 11/21/2025

Manage custodian communications templates in eDiscovery (Premium)

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

When you or other users create a hold notification or other types of custodian communications, you had to create the communication document from scratch by using the communications editor on the **Communications** tab in an eDiscovery (Premium) case. Now, we've released a new feature that lets you create communications templates that can be used to create communications in any case in your organization. After communication templates are created, they're available to be used in a case. This means that paralegals or other users who create custodian communications don't have to start from scratch to build a notification. Instead, they can select a template to build the notification that is sent to a custodian.

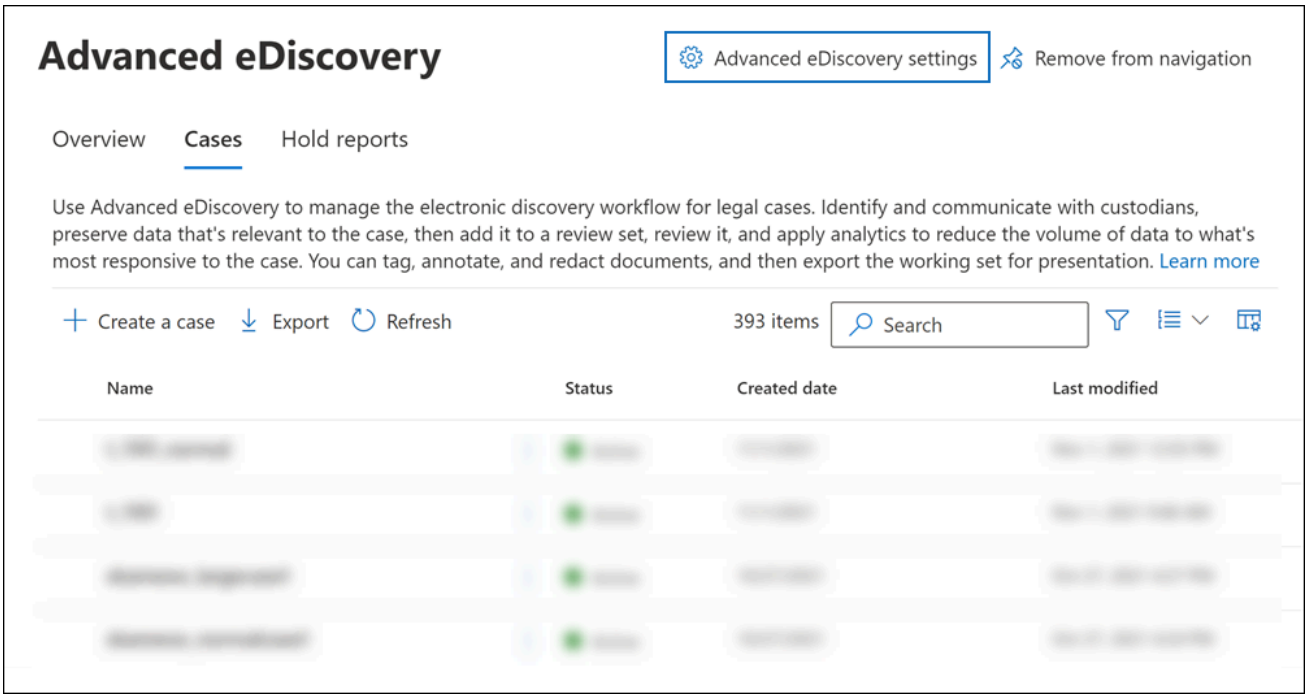
This article explains how to create organization-wide communications templates and select them when creating a new custodian notification for a specific eDiscovery (Premium) case.

Before you create templates in the Communications library

- You must be an eDiscovery Administrator in your organization to add or remove templates in the Communications library in eDiscovery (Premium). For more information, see Assign eDiscovery permissions in the Microsoft Purview portal
- Your organization can have a maximum of 50 templates in the Communications library.

Create a communications template

1. In the Microsoft Purview portal, go to eDiscovery (Premium), and then select **eDiscovery (Premium) settings**.



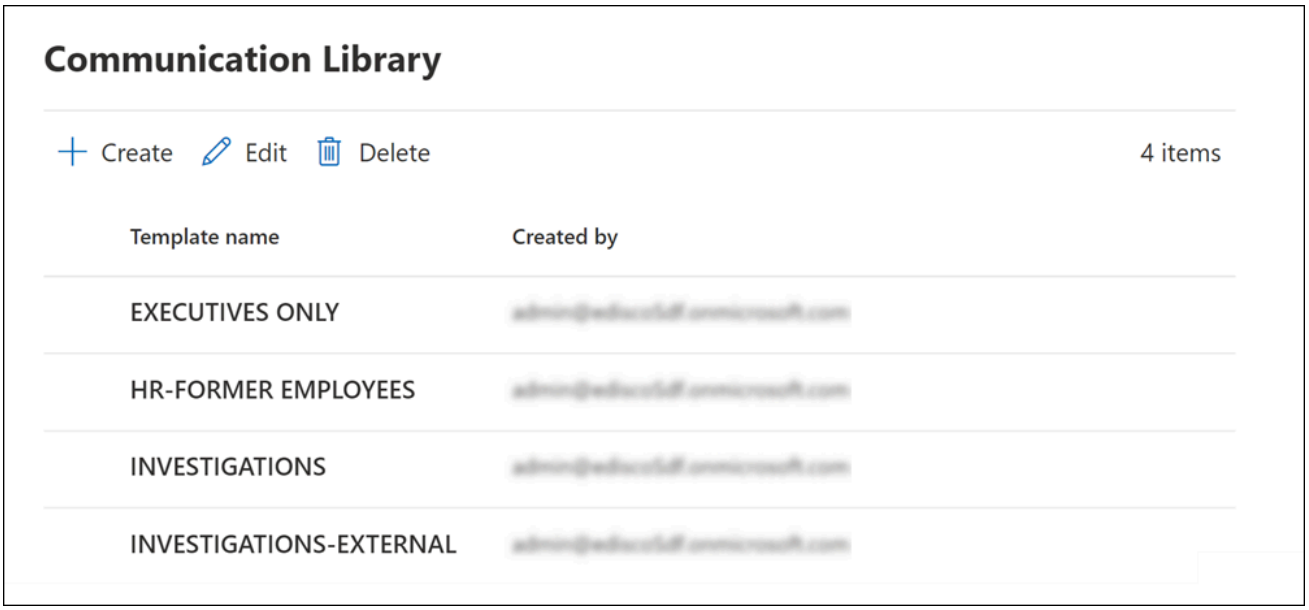
2. On the **Settings** page, select the **Communications library** tab.

3. On the **Communications library** page, select **Create**.

4. Follow the procedure to create a custodian communication. For step-by-step instructions, see Create a legal hold notification.

> **Note**
>
> The steps to create a communications template are the same as the workflow to create a notification within a case. The only difference is that when you create a template, you don't specify an issuing officer and you don't assign custodians. Specifying an issuing officer and assigning custodians is done when you use a communications template to create a custodian notification for a case.

5. After you create a template, it's displayed on the **Communications library** page.



You or other eDiscovery Administrators can edit a communications template. Any changes that you make to a template don't affect or modify any notifications that were previously created using that template. These changes will only apply to new notifications that created using the updated template.

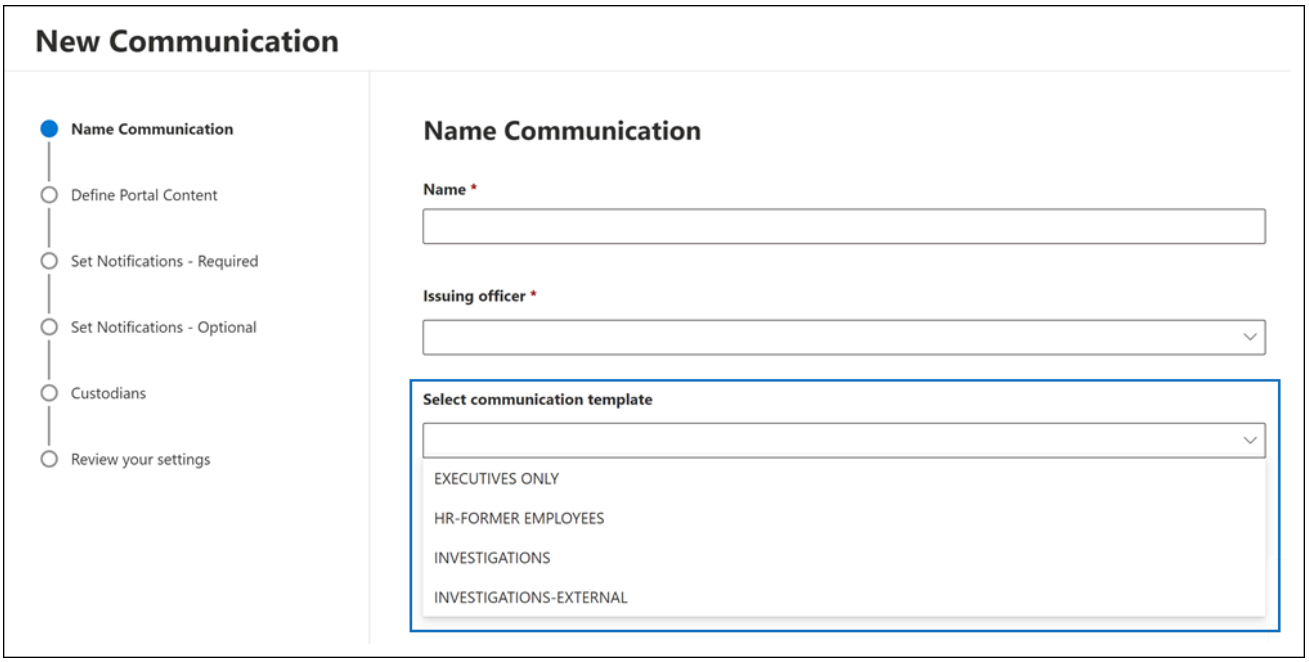
Use a communications template to create a custodian notification

After one or more communications templates are created in the Communications library, these templates can be selected to create a custodian notification in a case.

To select a template:

1. In the Microsoft Purview portal, go to **eDiscovery** > **Premium** to display the list of cases in your organization.
2. Select a case, select the **Communications** tab, and then select **New communication**.
3. On the **Name communication** page, use the **Select communication template** drop-down list to select a communications template to use to create the custodian notification.

   The list of templates in your organization's Communication library is displayed in the drop-down list.



Last updated on 11/21/2025

Manage holds in eDiscovery (Premium)

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

You can use a Microsoft Purview eDiscovery (Premium) case to create holds to preserve content that might be relevant to your case. Using the eDiscovery (Premium) hold capabilities, you can place holds on custodians and their data sources. Additionally, you can place a noncustodial hold on mailboxes and OneDrive for Business sites. You can also place a hold on the group mailbox, SharePoint site, and OneDrive for Business site for a Microsoft 365 group. Similarly, you can place a hold on the mailbox and site that are associated with Microsoft Teams. When you place content locations on hold, content is held until you release the custodian, remove a specific data location, or delete the hold policy entirely. After you create an eDiscovery hold, it may take up to 24 hours for the hold to be applied.

> **Important**
>
> For long term data retention not related to eDiscovery investigations, it is strongly advised to use retention policies and retention labels. For more information, see Learn about retention policies and retention labels.

View custodian-based holds

In some cases, you may have a set of custodians that you've identified and decided to preserve their data during the case. In eDiscovery (Premium), when these custodians are placed on hold, the user and their selected data sources are automatically added to a custodian hold policy.

Points to remember about hold policies:

- The best policy status is fetched when a location is viewed in a data source
- If the existing applied policy is deleted or released, the system replaces the hold policy status and name with next best applied or active policy.
- Only 1 policy per location is reported. However, there can be more than 1 policy impacting the location within the case. View the hold report to determine all applicable policies.
- The status for a location is fetched only from the policies applicable to the case. A policy from another case won't be referred to highlight the Hold status of the location.
- The hold status of a location is displayed in the *Hold status* column in the hold report.

To view the custodian hold policy:

1. In the Microsoft Purview portal, select **eDiscovery** > **Premium** to display the list of cases in your organization.
2. Go to the **Sources** tab to add custodians within your case. To learn how you can add and place custodians on hold within an eDiscovery (Premium) case, see Add Custodians to a case. If you have already added custodians and placed them on hold, go to step 3.
3. Go to the **Holds** tab and select **CustodianHold(HoldId)** and ensure the hold is placed successfully.

The hold status shown for a custodian in the data source tab is the most restrictive status for all holds in the underlying locations within the custodian or noncustodial data source. For example, if a custodian is added and includes their mailbox, OneDrive, and several Teams channels, all these data sources are placed on hold. However, if the hold for OneDrive returns an error but hold on all other data sources is successful, the hold status for the custodian is *Not on Hold* because the most restrictive status also applies to all of the underlying holds.

For an individual location within a data source, the hold status of that location is determined using all the hold policies applied on that location and the distribution status. The hold status of any successfully applied hold is applied. For example, a location is part of three hold policies and two policies have a *Failed/error* hold status. If the third policy has a hold successfully applied, then the hold status of the location would be *On Hold* based on the success of the hold on the third policy.

View noncustodial holds

In some cases, you may have a set of data that isn't tied to a specific custodian, but you need to identify the data as relevant to the case and preserve it. When these noncustodial sources are placed on hold in eDiscovery (Premium), the selected data sources are automatically added to a noncustodial hold policy.

To view a noncustodial hold for an eDiscovery (Premium) case:

1. In the Microsoft Purview portal, select **eDiscovery** > **Premium** to display the list of cases in your organization.
2. Select the **Sources** tab to add noncustodial locations within your case. To learn how you can add and place noncustodial data on hold within an eDiscovery (Premium) case, see Add noncustodial data sources to an eDiscovery (Premium) case. If you've already added noncustodial location and placed them on hold, go to step 3.
3. Select the **Holds** tab and select **NCDSHold(HoldId)** and ensure hold is placed successfully.

The source policy for the noncustodial location is shown below the hold status. The hold status of that location is determined using all the hold policies applied to the location and the distribution status. The least restrictive hold status of all the statuses is applied.

Placing custodial and noncustodial locations on hold

Placing custodial and noncustodial locations on hold preserves all the content in the locations. For steps to place query-based hold in eDiscovery (Premium), see Create eDiscovery holds in an eDiscovery case.

> **Note**
>
> When you create a query-based hold, all content from selected locations is initially placed on hold. After the timer job in either Exchange or SharePoint runs, any content that doesn't match the specified query is cleared from the hold. After the character count across all queries on a single location exceeds 10,000 characters, the entire location is placed on hold.

> **Note**
>
> If the SMTP address of the user changes after you place the user's mailbox on hold, the mailbox remains on hold. To use the new SMTP address to place hold, create a new hold.

Place a hold on Microsoft Teams and Microsoft 365 groups

Microsoft Teams is built on Microsoft 365 groups. Therefore, placing them on hold in eDiscovery (Premium) is similar. Keep the following things in mind when placing Microsoft 365 groups and Microsoft Teams on hold:

- To place content located in Microsoft 365 groups and Microsoft Teams on hold, you have to specify the mailbox and SharePoint site that associated with a group or team.

- Run the **Get-UnifiedGroup** cmdlet in Exchange Online to view properties for a Microsoft 365 group or Microsoft Team. This method is a good way to get the URL for the site that's associated with a Microsoft 365 group or a Microsoft Team. For example, the following command displays selected properties for a Microsoft 365 group named Senior Leadership Team:

  Console

  ```
  Get-UnifiedGroup "Senior Leadership Team" | FL
  DisplayName,Alias,PrimarySmtpAddress,SharePointSiteUrl
  DisplayName            : Senior Leadership Team
  Alias                  : seniorleadershipteam
  PrimarySmtpAddress     : seniorleadershipteam@contoso.onmicrosoft.com
  SharePointSiteUrl      :
  https://contoso.sharepoint.com/sites/seniorleadershipteam
  ```

  > **Note**
  >
  > To run the Get-UnifiedGroup cmdlet, you have to be assigned the View-Only Recipients role in Exchange Online or be a member of a role group that's assigned the View-Only Recipients role.

- When a user's mailbox is searched, any Microsoft 365 group or Microsoft Team that the user is a member of isn't searched. Similarly, when you place a Microsoft 365 group or Microsoft Team hold, only the group mailbox and group site are placed on hold; the mailboxes and OneDrive sites of group members aren't placed on hold unless you explicitly add them as custodians or place their data sources hold. Therefore, if you need to place a Microsoft 365 group or Microsoft Team on hold for a specific custodian, consider mapping the group site and group mailbox to the custodian (See Managing Custodians in eDiscovery (Premium)). If the Microsoft 365 group or Microsoft Team isn't attributable to a single custodian, consider adding the source to a noncustodial hold.

- To get a list of the members of a Microsoft 365 group or Microsoft Team, you can view the properties on the **Home** > **Groups** page in the Microsoft 365 admin center. Alternatively, you can run the following command in Exchange Online PowerShell:

  PowerShell

```
Get-UnifiedGroupLinks <group or team name> -LinkType Members | FL
DisplayName,PrimarySmtpAddress
```

> **Note**
>
> To run the **Get-UnifiedGroupLinks** cmdlet, you have to be assigned the View-Only Recipients role in Exchange Online or be a member of a role group that's assigned the View-Only Recipients role.

- Channel conversations that are part of a Microsoft Teams channel are stored in the mailbox that's associated with the Team. Similarly, files that team members share in a channel are stored on the team's SharePoint site. Therefore, you have to place the Microsoft Team mailbox and SharePoint site on hold to retain conversations and files in a channel.

- Alternatively, conversations that are part of the Chat list in Microsoft Teams are stored in the mailbox of the user's who participate in the chat. Files that a user shares in Chat conversations are stored in the OneDrive site of the user who shares the file. Therefore, you have to place the individual user mailboxes and OneDrive sites on hold to retain conversations and files in the Chat list.

- Every Microsoft Team or team channel contains a Wiki for note-taking and collaboration. The Wiki content is automatically saved as a .mht file. This file is stored in the Teams Wiki Data document library on the team's SharePoint site. You can place the content in the Wiki on hold by placing the team's SharePoint site on hold.

  > **Note**
  >
  > The capability to retain Wiki content for a Microsoft Team or team channel (when you place the team's SharePoint site on hold) was released on June 22, 2017. If a team site is on hold, the Wiki content will be retained starting on that date. However, if a team site is on hold and the Wiki content was deleted before June 22, 2017, the Wiki content was not retained.

Hold status

The following table lists the status of holds for items:

Expand table

Hold Status	Definition
Applied	A hold is applied on the location. Content is preserved based on the hold policy conditions.
Apply Failed	The hold policy trying to apply Hold on the location has failed. Check the distribution errors in the hold policy.
Applying	A hold is being applied on the location. The hold job is in progress.
Not on Hold	A hold isn't applied on the location. Content preservation isn't force.
Released	A hold is released on the location. Content preservation is disabled.
Release failed	The hold policy releasing a hold on the location failed. Check the distribution errors in the hold policy.
Releasing	A hold release for the location is in progress.
Unknown	The hold status is unknown for the location. Check hold policies, distribution status, and the hold report. This status may appear in cases where the system isn't able to map a hold policy to the location, indicating there isn't a hold present or applied on the location.

Manage hold status errors

You may encounter errors while placing a hold on custodial or noncustodial data sources. The following table lists the errors that you may encounter and the recommended resolution.

Expand table

Hold error types	Description	Resolution
Policy deployment interrupted	A system error indicating a problem was encountered while applying the hold.	Select **Retry hold action** on the custodial/non-custodial data source flyout page command bar to retry the hold application.
Site inaccessible	Indicates the SharePoint location associated with the requested hold request isn't accessible and may be read only.	Contact your SharePoint site administrator to configure the site as writable and retry the hold with **Retry hold action** on the custodial/non-custodial data source flyout page command bar.
Site not found.	Indicates the SharePoint location associated with the requested hold may have been	Check the site URL and confirm if the SharePoint site exists. Once confirmed,

Hold error types	Description	Resolution
	moved, deleted, or the site URL may not exist.	edit the custodial/non-custodial data source for the site and retry the hold action.
Mailbox not found	Indicates the mailbox associated with the requested hold isn't a valid mailbox.	Verify the email address and check that it's a valid Exchange Online mailbox. Once confirmed, edit the custodial/non-custodial data source for the mailbox and then retry the hold action.
Distribution group has too many members	Indicates the distribution group associated with the requested hold has more than 1,000 email addresses. Currently, a distribution group having more than 1,000 email addresses can't be expanded to be placed on hold.	Add the individual email addresses as custodial or noncustodial data sources or split the distribution group into groups with less than 1,000 email addresses and retry the hold action.
Invalid email address or URL	Indicates the location associated with the requested hold has an invalid email address or site URL.	Specify a valid email address or URL that exists within your organization.

Additional resources

- eDiscovery hold limits

Last updated on 09/02/2025

Create eDiscovery holds in an eDiscovery case

> **⊗ Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

You can use a Microsoft Purview eDiscovery (Premium) or (Standard) case to create holds to preserve content that might be relevant to the case. You can place a hold on the Exchange mailboxes and OneDrive for Business accounts of people you're investigating in the case. You can also place a hold on the mailboxes and sites that are associated with Microsoft Teams, Microsoft 365 groups, and Viva Engage Groups. When you place content locations on hold, content is preserved until you remove the content location from the hold or until you delete the hold.

> **ⓘ Important**
>
> For long term data retention not related to eDiscovery investigations, it is strongly advised to use retention policies and retention labels. For more information, see Learn about retention policies and retention labels.

After you create an eDiscovery hold, it may take up to 24 hours for the hold to take effect.

When you create a hold, you have the following options to scope the content that's preserved in the specified content locations:

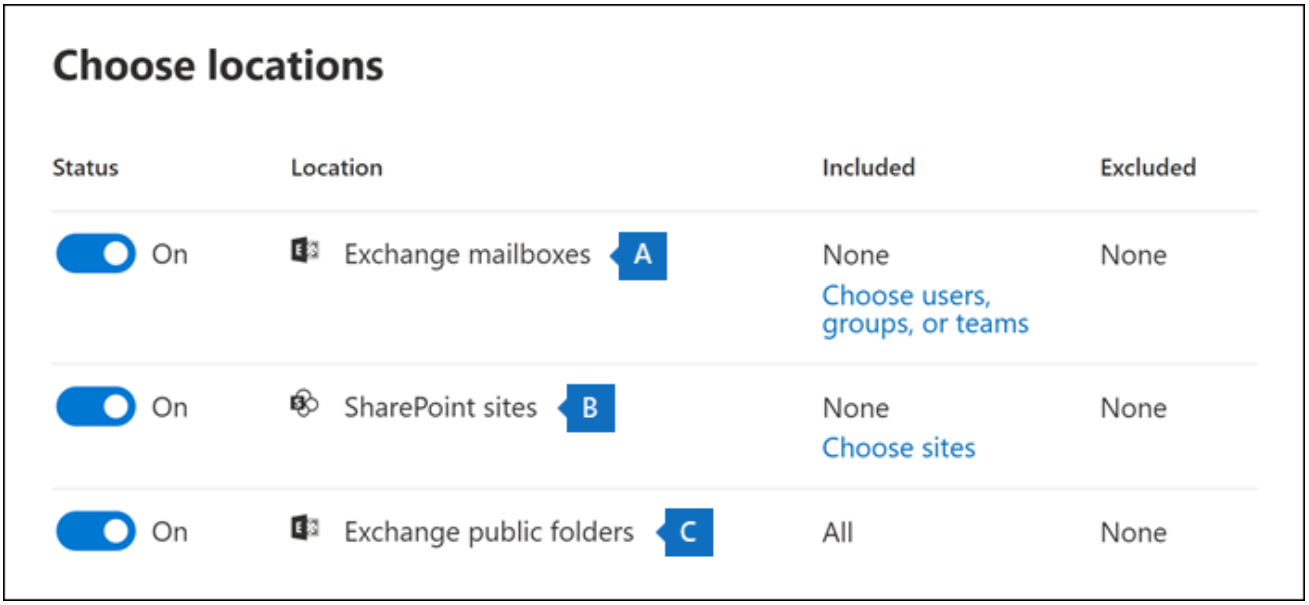
- Create an infinite hold where all content in the specified locations is placed on hold. Alternatively, you can create a query-based hold where only the content in the specified locations that matches a search query is placed on hold.
- Specify a date range to preserve only the content that was sent, received, or created within that date range. Alternatively, you can hold all content in specified locations regardless of

when sent, received, or created.

How to create an eDiscovery hold

To create an eDiscovery hold that's associated with an eDiscovery (Premium) or (Standard) case:

1. Go to Microsoft Purview portal and sign in using the credentials for user account with the appropriate eDiscovery permissions.
2. In the left navigation pane, select **Show all**, and then select **eDiscovery** > **Premium** or **eDiscovery** > **Standard**.
3. On the **eDiscovery** > **Premium** or **eDiscovery (Standard)** page, select the name of the case that you want to create the hold in.
4. On the **Home** page for the case, select the **Hold** tab.
5. On the **Hold** page, select **Create**.
6. On the **Name your hold** workflow page, give the hold a name and add an optional description, and then select **Next**. The name of the hold must be unique in your organization.
7. On the **Choose locations** workflow page, choose the content locations that you want to place on hold. You can place mailboxes, sites, and public folders on hold.



   a. **Exchange mailboxes**: Set the toggle to **On** and then select **Choose users, groups, or teams** to specify the mailboxes to place on hold. Use the search box to find user mailboxes and distribution groups (to place a hold on the mailboxes of group members) to place on hold. You can also place a hold on the associated mailbox for a Microsoft

Team, Microsoft 365 group, and Viva Engage Group. For more information about the application data that is preserved when a mailbox is placed on hold, see Content stored in mailboxes for eDiscovery.

> **Important**
>
> When you select a distribution list to be placed on hold, the hold is placed on each of the member mailboxes in the distribution list when the policy is created. Subsequent changes in the distribution list do not change or update the holds or the policy.

b. **SharePoint sites**: Set the toggle to **On** and then select **Choose sites** to specify SharePoint sites and OneDrive accounts to place on hold. Type the URL for each site that you want to place on hold. You can also add the URL for the SharePoint site for a Microsoft Team, Microsoft 365 group or a Yammer Group.

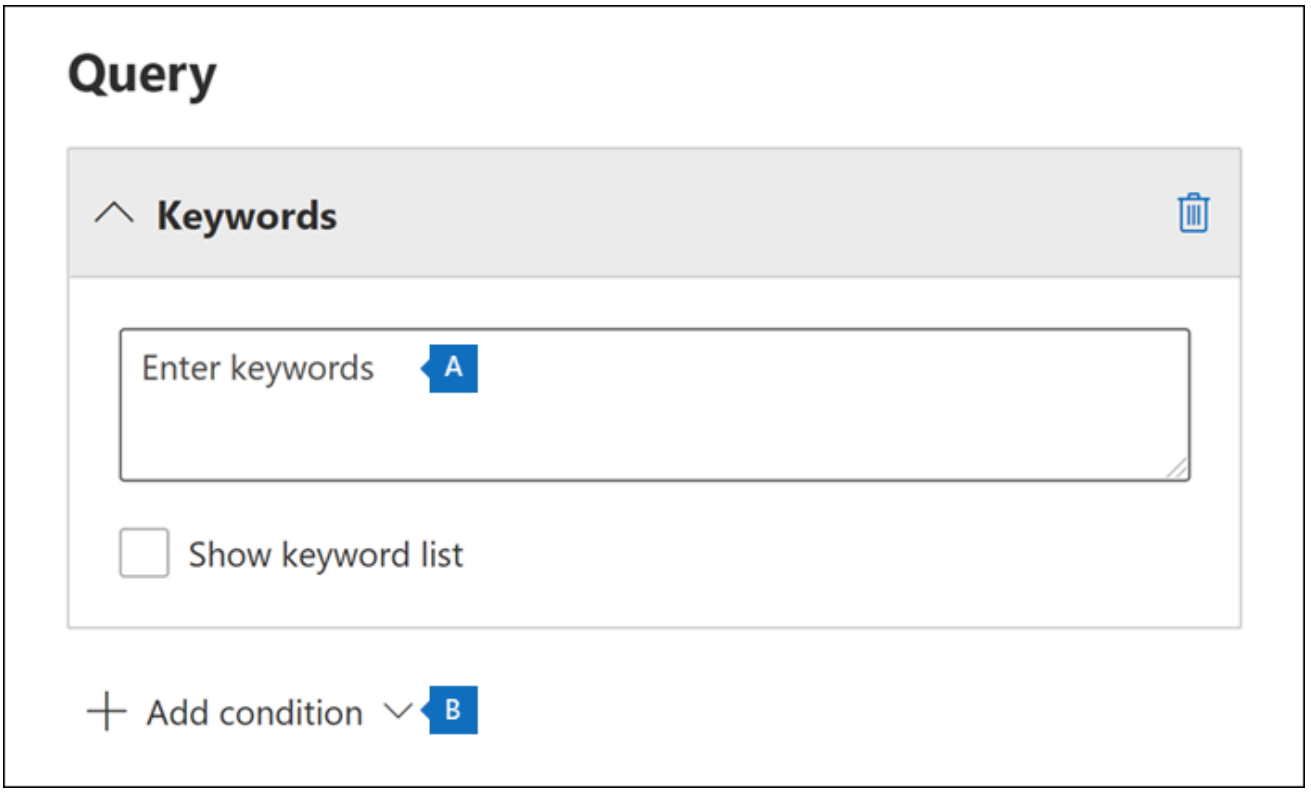
> **Important**
>
> To create a hold for a subsite related to a SharePoint Online site, you must use the *Path* property in a query filter to select a specific subsite.

c. **Exchange public folders**: Set the toggle to **On** to put all public folders in your Exchange Online organization on hold. You can't choose specific public folders to put on hold. Leave the toggle switch off if you don't want to put a hold on public folders.

> **Important**
>
> When adding Exchange mailboxes or SharePoint sites to a hold, you must explicitly add at least one content location to the hold. In other words, if you set the toggle to **On** for mailboxes or sites, you must select specific mailboxes or sites to add to the hold. Otherwise, the eDiscovery hold will be created but no mailboxes or sites will be added to the hold.

8. When finished adding locations to the hold, select **Next**.

9. To create a query-based hold using keywords or conditions, complete the following steps. To preserve all content in the specified content locations, select **Next**.



a. In the box under **Keywords**, type a query to preserve only the content that matches the query criteria. You can specify keywords, email message properties, or site properties, such as file names. You can also use more complex queries that use a Boolean operator, such as **AND**, **OR**, or **NOT**.
b. Select **Add condition** to add one or more conditions to narrow the query for the hold. Each condition adds a clause to the KeyQL search query that is created and run when you create the hold. For example, you can specify a date range so that email or site documents that were created within the date ranged are preserved. A condition is logically connected to the keyword query (specified in the **Keywords** box) and other conditions by the **AND** operator. That means items have to satisfy both the keyword query and the condition to be preserved.

For more information about creating a search query and using conditions, see Keyword queries and search conditions for eDiscovery.

10. After configuring a query-based hold, select **Next**.

11. Review your settings (and edit them if necessary), and then select **Submit**.

After creating a hold, check that the hold is applied successfully by navigating to the **Hold** tab in the case and selecting the hold policy. For more information about troubleshooting holds with errors, see Resolve eDiscovery hold errors.

> ⓘ Note

When you create a query-based hold, all content from selected locations is initially placed on hold. Subsequently, any content that doesn't match the specified query is cleared from the hold every seven to 14 days. However, a query-based hold won't clear content if more than five holds of any type are applied to a content location, or if any item has indexing issues.

Query-based holds placed on sites

Keep the following things in mind when you place a query-based eDiscovery hold on documents located in SharePoint sites:

- A query-based hold initially preserves all documents in a site for a short period of time after they're deleted. That means when a document is deleted, it is moved to the Preservation Hold library even if it doesn't match the criteria of the query-based hold. However, deleted documents that don't match a query-based hold will be removed by a timer job that processes the Preservation Hold library. The timer job runs periodically and compares all documents in the Preservation Hold library to your query-based eDiscovery holds (and other types of holds and retention policies). The timer job deletes the documents that don't match a query-based hold and preserves the documents that do.
- Query-based holds shouldn't be used to perform targeted preservation, like preserving documents in a specific folder or site or by using other location-based hold criteria. Doing so may have unintended results. We recommend using non-location based hold criteria such as keywords, date ranges, or other document properties to preserve site documents.

Search locations on eDiscovery hold

When you search for content in an eDiscovery (Standard) case, you can quickly configure the search to only search the content locations that have been placed on a hold associated with the case.

Select the **Locations on hold** option to search all the content locations that have been placed on hold. If the case contains multiple eDiscovery holds, the content locations from all holds are searched when you select this option. Additionally, if a content location was placed on a query-based hold, only the items that match the hold query are searched when you run the search. In other words, only the content that matches both the hold criteria and the search criteria is returned with the search results. For example, if a user was placed on query-based case hold that preserves items that were sent or created before a specific date, only those items would be

searched. This is accomplished by connecting the case hold query and the search query by an **AND** operator.

Here are some other things to keep in mind when searching locations on eDiscovery hold:

- If a content location is part of multiple holds within the same case, the hold queries are combined by **OR** operators when you search that content location using the all case content option. Similarly, if a content location is part of two different holds, where one is query-based and the other is an infinite hold (where all content is placed on hold), then all content is search because of the infinite hold.
- If a search is configured it to search locations on hold and then you change an eDiscovery hold in the case (by adding or removing a location or changing a hold query), the search configuration is updated with those changes. However, you have to rerun the search after the hold is changed to update the search results.
- If multiple eDiscovery holds are placed on a single location in an eDiscovery case and you select to search locations on hold, the maximum number of keywords for that search query is 500. That's because the search combines all the query-based holds by using the **OR** operator. If there are more than 500 keywords in the combined hold queries and the search query, then all content in the mailbox is searched, not just that content that matches the query-based case holds.
- If an eDiscovery hold has a status of **On (Pending)**, you can still search the locations on hold while the hold is being turned on.

Preserve content in Microsoft Teams

Conversations that are part of a Microsoft Teams channel are stored in the mailbox that's associated with the Microsoft Team. Similarly, files that team members share in a channel are stored on the team's SharePoint site. Therefore, you have to place the Team mailbox and SharePoint site on eDiscovery hold to preserve conversations and files in a channel.

Alternatively, conversations that are part of the Chat list in Teams (called *1:1 chats* or *1:N group chats*) are stored in the mailboxes of the users who participate in the chat. And files that users share in chat conversations are stored in the OneDrive account of the user who shares the file. Therefore, you have to add the individual user mailboxes and OneDrive accounts to an eDiscovery hold to preserve conversations and files in the chat list. It's a good idea to place a hold on the mailboxes of members of a Microsoft Team in addition to placing the team mailbox and site on hold.

> **Note**
>
> If your organization has an Exchange hybrid deployment (or your organization synchronizes an on-premises Exchange organization with Office 365) and has enabled Microsoft Teams, on-premises users can use the Teams chat application and participate in 1:1 chats and 1:N group chats. These conversations are stored in cloud-based storage that's associated with an on-premises user. If an on-premises user is placed on an eDiscovery hold, the Teams chat content in the cloud-based storage will be preserved. For more information, see **Search for Teams chat data for on-premises users**.

For more information about preserving Teams content, see Place a Microsoft Teams user or team on legal hold.

Preserve card content

Similarly, card content generated by apps in Teams channels, 1:1 chats, and 1:N group chats are stored in mailboxes and is preserved when a mailbox is placed on an eDiscovery hold. A *card* is a UI container for short pieces of content. Cards can have multiple properties and attachments, and can include items that trigger card actions. For more information, see Cards. Like other Teams content, where card content is stored is based on where the card was used. Content for cards used in a Teams channel is stored in the Teams group mailbox. Card content for 1:1 and 1xN chats are stored in the mailboxes of the chat participants.

Preserve meeting and call information

Summary information for meetings and calls in a Teams channel is also stored in the mailboxes of users who dialed into the meeting or call. This content is also preserved when an eDiscovery hold is placed on user mailboxes.

Preserve content in private channels

Starting in February 2020, we also turned on the ability to preserve content in private channels. Because private channel chats are stored in the mailboxes of the chat participants, placing a user mailbox on eDiscovery hold preserves private channel chats. Also, if a user mailbox was placed on an eDiscovery hold prior to February 2020, the hold will now automatically apply to private channel messages stored in that mailbox. Preserving files shared in private channels is also supported.

Preserve wiki content

Every Team or team channel also contains a Wiki for note taking and collaboration. The Wiki content is automatically saved to a file with a .mht format. This file is stored in the Teams Wiki Data document library on the team's SharePoint site. You can preserve the wiki content by adding the team's SharePoint site to an eDiscovery hold.

> **Note**
>
> The capability to preserve Wiki content for a Team or team channel (when you place the team's SharePoint site on hold) was released on June 22, 2017. If a team site is on hold, the Wiki content will be retained starting on that date. However, if a team site is on hold and the Wiki content was deleted before June 22, 2017, the Wiki content was not preserved.

Microsoft 365 groups

Teams is built on Microsoft 365 groups. Therefore, placing Microsoft 365 groups on eDiscovery hold is similar placing Teams content on hold.

Keep the following things in mind when placing both Teams and Microsoft 365 groups on an eDiscovery hold:

- As previously explained, to place content located in Teams and Microsoft 365 groups on hold, you have to specify the mailbox and SharePoint site that associated with a group or team.
- Run the **Get-UnifiedGroup** cmdlet in Exchange Online PowerShell to view properties for Teams and Microsoft 365 groups. This is a good way to get the URL for the site that's associated with a Team or Microsoft 365 group. For example, the following command displays selected properties for a Microsoft 365 group named Senior Leadership Team:

  text

  ```
  Get-UnifiedGroup "Senior Leadership Team" | FL
  DisplayName,Alias,PrimarySmtpAddress,SharePointSiteUrl

  DisplayName            : Senior Leadership Team
  Alias                  : seniorleadershipteam
  PrimarySmtpAddress     : seniorleadershipteam@contoso.onmicrosoft.com
  SharePointSiteUrl      :
  https://contoso.sharepoint.com/sites/seniorleadershipteam
  ```

> **Note**
>
> To run the **Get-UnifiedGroup** cmdlet, you have to be assigned the View-Only Recipients role in Exchange Online or be a member of a role group that's assigned the View-Only Recipients role.

- When a user's mailbox is searched, any Team or Microsoft 365 group that the user is a member of won't be searched. Similarly, when you place a Team or Microsoft 365 group on eDiscovery hold, only the group mailbox and group site are placed on hold. The mailboxes and OneDrive for Business sites of group members aren't placed on hold unless you explicitly add them to the eDiscovery hold. So if you have to place a Team or Microsoft 365 group on hold for a legal reason, consider adding the mailboxes and OneDrive accounts of team or group members on the same hold.

- To get a list of the members of a Team or Microsoft 365 group, you can view the properties on the **Groups** page in the Microsoft 365 admin center. Alternatively, you can run the following command in Exchange Online PowerShell:

  PowerShell

  ```
  Get-UnifiedGroupLinks <group or team name> -LinkType Members | FL DisplayName,PrimarySmtpAddress
  ```

  > **Note**
  >
  > To run the **Get-UnifiedGroupLinks** cmdlet, you have to be assigned the View-Only Recipients role in Exchange Online or be a member of a role group that's assigned the View-Only Recipients role.

Preserve content in OneDrive accounts

> **Important**
>
> The retention period for deleted OneDrive accounts is different from the retention period for mailboxes. For more information about the retention period for deleted OneDrive accounts, see OneDrive retention and deletion.

To collect a list of the URLs for the OneDrive for Business sites in your organization so you can add them to a hold or search associated with an eDiscovery case, see Create a list of all OneDrive locations in your organization. The script in this article creates a text file that contains a list of all OneDrive sites in your organization. To run this script, you have to install and use the SharePoint Online Management Shell. Be sure to append the URL for your organization's MySite domain to each OneDrive site that you want to search. This is the domain that contains all your OneDrive; for example, `https://contoso-my.sharepoint.com`. Here's an example of a URL for a user's OneDrive site: `https://contoso-my.sharepoint.com/personal/sarad_contoso_onmicrosoft.com`.

> **Important**
>
> The URL for a user's OneDrive account includes their user principal name (UPN) (for example, `https://alpinehouse-my.sharepoint.com/personal/sarad_alpinehouse_onmicrosoft_com`). In the rare case that a person's UPN is changed, their OneDrive URL will also change to incorporate the new UPN. If a user's OneDrive account is part of an eDiscovery hold, and their UPN is changed, you need to update the hold by adding the user's new OneDrive URL and removing the old one. If the URL for the OneDrive site changes, previously placed holds on the site remain effective and content is preserved. For more information, see **How UPN changes affect the OneDrive URL**.

Removing content locations from an eDiscovery hold

After a mailbox, SharePoint site, or OneDrive account is removed from an eDiscovery hold, a *delay hold* is applied. This means that the actual removal of the hold is delayed for 30 days to prevent data from being permanently deleted (purged) from a content location. This gives admins an opportunity to search for or recover content that will be purged after an eDiscovery hold is removed. The details of how the delay hold works for mailboxes and sites are different.

- **Mailboxes:** A delay hold is placed on a mailbox the next time the Managed Folder Assistant processes the mailbox and detects that an eDiscovery hold was removed. Specifically, a delay hold is applied to a mailbox when the Managed Folder Assistant sets one of the following mailbox properties to **True**:
  - **DelayHoldApplied:** This property applies to email-related content (generated by people using Outlook and Outlook on the web) that's stored in a user's mailbox.
  - **DelayReleaseHoldApplied:** This property applies to cloud-based content (generated by non-Outlook apps such as Microsoft Teams, Microsoft Forms, and Microsoft Viva Engage)

that's stored in a user's mailbox. Cloud data generated by a Microsoft app is typically stored in a hidden folder in a user's mailbox.

When a delay hold is placed on the mailbox (when either of the previous properties is set to **True**), the mailbox is still considered to be on hold for an unlimited hold duration, as if the mailbox was on Litigation Hold. After 30 days, the delay hold expires, and Microsoft 365 will automatically attempt to remove the delay hold (by setting the DelayHoldApplied or DelayReleaseHoldApplied property to **False**) so that the hold is removed. After either of these properties are set to **False**, the corresponding items that are marked for removal are purged the next time the mailbox is processed by the Managed Folder Assistant.

For more information, see Managing mailboxes on delay hold.

- **SharePoint and OneDrive sites:** Any SharePoint or OneDrive content that's being retained in the Preservation Hold library isn't deleted during the 30-day delay hold period after a site is removed from an eDiscovery hold. This is similar to what happens when a site is released from a retention policy. Additionally, you can't manually delete this content in the Preservation Hold library during the 30-day delay hold period. To release a site from the 30-day delay hold/grace period hold, see the Can't delete a site because of an invalid retention policy or eDiscovery hold troubleshooting article.

  For more information, see Releasing a policy for retention.

A delay hold is also applied to content locations on hold when you close an eDiscovery (Standard) case because holds are turned off when a case is closed. For more information about closing a case, see Close, reopen, and delete an eDiscovery (Standard) case.

eDiscovery hold limits

The following table lists the limits for eDiscovery cases and case holds.

Expand table

Description of limit	Limit
Maximum number of cases for an organization.	No limit
Maximum number of eDiscovery hold policies for an organization. This limit includes the combined total of hold policies in eDiscovery (Standard) and eDiscovery (Premium) cases.	10,000[1]
Maximum number of mailboxes in a single eDiscovery hold. This limit includes the combined total of user mailboxes, and the mailboxes associated with Microsoft 365	1,000

Description of limit	Limit
groups, Microsoft Teams, and Viva Engage Groups.	
Maximum number of sites in a single eDiscovery hold. This limit includes the combined total of OneDrive for Business sites, SharePoint sites, and the sites associated with Microsoft 365 groups, Microsoft Teams, and Viva Engage Groups. <br/	100
Maximum number of cases displayed on the eDiscovery home page, and the maximum number of items displayed on the Holds, Searches, and Export tabs within a case.	1,000[1]
Hold limits for SharePoint sites and OneDrive	For details, see SharePoint limits.

Note

[1] To view a list of more than 1,000 cases, holds, searches, or exports, you can use the corresponding Security & Compliance PowerShell cmdlet:

- Get-ComplianceCase
- Get-CaseHoldPolicy
- Get-ComplianceSearch
- Get-ComplianceSearchAction

Last updated on 02/21/2025

Hold report in eDiscovery (Premium) (preview)

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic **Content Search**, classic **eDiscovery (Standard)**, and classic **eDiscovery (Premium)**.
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the **new eDiscovery experience** in the **Microsoft Purview portal**.

The hold report in eDiscovery (Premium) lets users with *eDiscovery Administrator* or *eDiscovery Manager (preview)* permissions access a built-in report with information on all holds associated with eDiscovery cases in the Microsoft Purview portal. This report includes eDiscovery holds associated with eDiscovery (Standard) and eDiscovery (Premium) cases. The hold report currently doesn't show custodian association for each data source, but will show the location.

The hold report is a report on all hold policies within the tenant and lists all locations that are part of those hold policies (whether enabled or disabled). Holds from eDiscovery (Standard) cases are identified by the **Standard** value in the **Case type** column. You can use the **Case type** group option to group the eDiscovery holds from eDiscovery (Standard) cases.

Step 1: Verify your subscription and permissions

- Your organization must have an organization subscription that supports eDiscovery (Premium) to generate and access the hold report.
- Users accessing the hold report must be assigned the *eDiscovery Administrator* or *eDiscovery Manager (preview)* role. All eDiscovery Administrators and eDiscovery Managers in your organization can access it on the **Hold report (preview)** tab in **Reports** section of eDiscovery (Premium). eDiscovery Managers (preview) can only view hold reports which are related to the cases they are a member of. For more information about eDiscovery permissions and the eDiscovery Administrator and eDiscovery Manager roles, see Assign eDiscovery permissions.

Step 2: View the hold report

1. Go to **eDiscovery** > **Premium** > **Reports** > **Hold report (preview)** to view your hold report.
2. Sort, filter, group by, and customize columns to customize your hold report.
3. Download the hold report to CSV file for further analysis. The entire report is downloaded regardless of any filtering or grouping option applied to the report.

The item count at the top of the hold report is total count of locations within the hold policies in the tenant. If the same location is placed on hold under different policies, then the same location is displayed multiple times in the report under the respective case and hold policy and contribute to multiple item counts.

When you search the report, only the **Case name** and **Location** columns are searched. For example, if you search for the keyword *Test*, the search only returns items that are associated with a case or location that contains the word Test.

Default hold report columns

- **Case Name**: Name of the case that the hold policy is associated with. If the case name changed and a hold policy associated with the case is also updated, the case name change for the associated row is immediately reflected in the report. When a full sync of the report occurs (currently every 3 to 7 days), the case name is also updated.
- **Hold name**: Name of the Hold Policy that contains the associated locations.
- **Last modified**: The date when the eDiscovery hold policy was last updated.
- **Location**: The exchange mailbox or SharePoint site in the hold policy. If this field displays *None*, it indicates that the hold policy has no active locations within it. Any previous locations that were part of the policy are released from hold and removed from the policy.
- **Location error**: Displays the error message associated with the location within the hold policy. If this field displays *None*, it indicates that associated location doesn't have errors, even if it's part of a *Partially Applied/Released* policy.
- **Location type**: Indicates if the location is a SharePoint site or Exchange mailbox.
- **Policy status**: The following policy status values are available:
    - *Applied successfully*: Indicates the location of the hold policy is turned on (enabled), a rule is enabled for the policy and the policy is applied successfully. This status means the

location is placed on hold.

- *Applying*: Indicates that the hold policy that the location is part of is set to be turned on (enabled) and is in progress to apply holds to the locations in it. This job is an asynchronous job, which means some locations may have been placed on hold while others may still be in progress. Wait for the hold job to complete checking the final status.

- *Applied Partially*: – Indicates that the hold policy that the location is part of is set to be turned on (enabled) and has some locations where the hold may haven't applied. Check the 'Location error' column. If the associated result is 'None' then the location is placed on hold. If there's an associated error, then the location isn't placed on hold and the error message is displayed here. Check the hold policy for more details to remediate and retry the hold.

- *Released successfully*: Indicates that the hold policy that the location is part of is set to be turned off (disabled) and has completed successfully. This status means the associated location is successfully released from hold.

- *Releasing*: Indicates that the hold policy that the location is part of is set to be turned off (disabled) and is in progress to release holds from the locations in it. This job is an asynchronous job, which means some locations may have been removed from hold while others may still be in progress. Wait for the hold job to complete checking the final status.

  > ⓘ **Note**
  >
  > System default policies (Custodian/NCDS) can only turned off by closing the case. The locations within these policies can still be removed from hold in an active case by releasing them from the data source.

- *Released partially*: Indicates that the hold policy that the location is part of is set to be turned off (disabled) and has some locations where the hold may haven't released. Check the 'Location error' column. If the associated result is 'None' then the location is successfully released from hold. If there's an associated error, then the location isn't removed from hold and the error message is displayed here. Check the hold policy for more details to remediate and retry the action.

- *Policy failed*: Indicates that the policy is set up incorrectly. The policy doesn't have a rule or was set up with rule set as disabled. These policies won't be enforced even if the policy status is successful. To ensure these policies are enforced, create a rule and resync the policy. For more information, see New-CaseHoldRule.

Other available hold report columns

- **Case type**: Indicates if the case is an eDiscovery (Standard) or eDiscovery (Premium case).
- **Case status**: Shows the status of case. Values are *Active*, *Closing*, *Closed with error*, *Pending delete*, or *Closed*. If the case status changed and a hold policy associated with the case is also updated, the case status change for the associated row is immediately reflected in the report. When a full sync of the report occurs (currently every 3 to 7 days), all case status changes are also updated.
- **Policy created by**: Shows the user that created the initial hold policy.
- **Policy last modified by**: Shows the user that last modified the hold policy.
- **Policy last modified**: Shows the last date that the policy was updated on. This date is the policy update date and doesn't necessarily match the date the location was placed on/released from hold.
- **Report last fetched**: Displays the date and time stamp of the latest fetch of policy status and location information by the report for the specific policy.
- **Policy created date**: Shows the first date that the policy was created on. This date is the policy creation date and doesn't necessarily match the date the location was placed on/released from hold.
- **Hold ID**: The unique ID of the hold policy.
- **Case ID**: The unique ID of the case associated with the hold policy.
- **Rule query**: This displays any content match query that is applied to the policy. If there's no content match query applied to the policy, this field is empty.
- **Rule ID**: The unique ID for the rule associated with the hold policy. This status is useful for sending information to Microsoft when reporting any hold errors. If there's no rule associated with the policy, *Rule not available* is displayed in this field. Policies where a rule isn't available aren't set up correctly and not enforced. Create rules and resync the policy. For more information, see New-CaseHoldRule.

Last updated on 10/31/2024

Use a script to create a report on holds in eDiscovery cases

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

The script in this article lets eDiscovery administrators and eDiscovery managers generate a report that contains information about all holds that are associated with eDiscovery (Standard) and eDiscovery (Premium) cases in the Microsoft Purview portal. The report contains information such as the name of the case a hold is associated with, the content locations that are placed on hold, and whether the hold is query-based. If there are cases that don't have any holds, the script creates an additional report with a list of cases without holds.

See the More information section for a detailed description of the information included in the report.

Admin requirements and script information

- To generate a report on all eDiscovery cases in your organization, you have to be an eDiscovery Administrator in your organization. If you're an eDiscovery Manager, the report will only include information about the cases that you can access. For more information about eDiscovery permissions, see Assign eDiscovery permissions.
- The script in this article has minimal error handling. The primary purpose is to quickly create report about the holds that are associated with the eDiscovery cases in your organization.
- The sample scripts provided in this article aren't supported under any Microsoft standard support program or service. The sample scripts are provided AS IS without warranty of any kind. Microsoft further disclaims all implied warranties including, without limitation, any implied warranties of merchantability or of fitness for a particular purpose. The entire risk arising out of the use or performance of the sample scripts and documentation remains

with you. In no event shall Microsoft, its authors, or anyone else involved in the creation, production, or delivery of the scripts be liable for any damages whatsoever (including, without limitation, damages for loss of business profits, business interruption, loss of business information, or other pecuniary loss) arising out of the use of or inability to use the sample scripts or documentation, even if Microsoft has been advised of the possibility of such damages.

Step 1: Connect to Security & Compliance PowerShell

The first step is to connect to Security & Compliance PowerShell for your organization. For step-by-step instructions, see Connect to Security & Compliance PowerShell.

Step 2: Run the script to report on holds associated with eDiscovery cases

After you've connected to Security & Compliance PowerShell, the next step is to create and run the script that collects information about the eDiscovery cases in your organization.

1. Save the following text to a Windows PowerShell script file by using a filename suffix of .ps1; for example, CaseHoldsReport.ps1.

PowerShell

```
 #script begin
 " "
 write-host "**********************************************"
 write-host "Security & Compliance Center   " -foregroundColor yellow -
backgroundcolor darkgreen
 write-host "eDiscovery cases - Holds report         " -foregroundColor yellow -
backgroundcolor darkgreen
 write-host "**********************************************"
 " "

 #prompt users to specify a path to store the output files
 $time = get-date -Format dd-MM-yyyy_hh.mm
 $Path = Read-Host 'Enter a folder path to save the report to a .csv file
(filename is created automatically)'
 $outputpath = $Path + '\' + 'CaseHoldsReport' + ' ' + $time + '.csv'
 $noholdsfilepath = $Path + '\' + 'CaseswithNoHolds' + $time + '.csv'

 #add case details to the csv file
 function add-tocasereport {
     Param([string]$casename,
         [String]$casetype,
         [String]$casestatus,
         [datetime]$casecreatedtime,
```

```
          [string]$casemembers,
          [datetime]$caseClosedDateTime,
          [string]$caseclosedby,
          [string]$holdname,
          [String]$Holdenabled,
          [string]$holdcreatedby,
          [string]$holdlastmodifiedby,
          [string]$ExchangeLocation,
          [string]$sharePointlocation,
          [string]$ContentMatchQuery,
          [datetime]$holdcreatedtime,
          [datetime]$holdchangedtime,
          [string]$holdstatus,
          [string]$holderror
     )

     $addRow = New-Object PSObject
     Add-Member -InputObject $addRow -MemberType NoteProperty -Name "Case name" -
Value $casename
     Add-Member -InputObject $addRow -MemberType NoteProperty -Name "Case type" -
Value $casetype
     Add-Member -InputObject $addRow -MemberType NoteProperty -Name "Case status"
-Value $casestatus
     Add-Member -InputObject $addRow -MemberType NoteProperty -Name "Case
members" -Value $casemembers
     Add-Member -InputObject $addRow -MemberType NoteProperty -Name "Case created
time" -Value $casecreatedtime
     Add-Member -InputObject $addRow -MemberType NoteProperty -Name "Case closed
time" -Value $caseClosedDateTime
     Add-Member -InputObject $addRow -MemberType NoteProperty -Name "Case closed
by" -Value $caseclosedby
     Add-Member -InputObject $addRow -MemberType NoteProperty -Name "Hold name" -
Value $holdname
     Add-Member -InputObject $addRow -MemberType NoteProperty -Name "Hold
enabled" -Value $Holdenabled
     Add-Member -InputObject $addRow -MemberType NoteProperty -Name "Hold created
by" -Value $holdcreatedby
     Add-Member -InputObject $addRow -MemberType NoteProperty -Name "Hold last
changed by" -Value $holdlastmodifiedby
     Add-Member -InputObject $addRow -MemberType NoteProperty -Name "Exchange
locations" -Value  $ExchangeLocation
     Add-Member -InputObject $addRow -MemberType NoteProperty -Name "SharePoint
locations" -Value $sharePointlocation
     Add-Member -InputObject $addRow -MemberType NoteProperty -Name "Hold query"
-Value $ContentMatchQuery
     Add-Member -InputObject $addRow -MemberType NoteProperty -Name "Hold created
time (UTC)" -Value $holdcreatedtime
     Add-Member -InputObject $addRow -MemberType NoteProperty -Name "Hold changed
time (UTC)" -Value $holdchangedtime
     Add-Member -InputObject $addrow -MemberType NoteProperty -Name "Hold Status"
-Value $holdstatus
     Add-Member -InputObject $addrow -MemberType NoteProperty -Name "Hold Error"
-Value $holderror

     $allholdreport = $addRow | Select-Object "Case name", "Case type", "Case
```

```
status", "Hold name", "Hold enabled", "Case members", "Case created time", "Case
closed time", "Case closed by", "Exchange locations", "SharePoint locations",
"Hold query", "Hold created by", "Hold created time (UTC)", "Hold last changed
by", "Hold changed time (UTC)", "Hold Status", "Hold Error"
     $allholdreport | export-csv -path $outputPath -notypeinfo -append -Encoding
ascii
 }

 #get information on the cases and pass values to the case report function
 " "
 write-host "Gathering a list of eDiscovery (Standard) cases and holds..."
 " "
 $edc = Get-ComplianceCase -ErrorAction SilentlyContinue
 foreach ($cc in $edc) {
     write-host "Working on case :" $cc.name
     if ($cc.status -eq 'Closed') {
         $cmembers = ((Get-ComplianceCaseMember -Case $cc.name).windowsLiveID) -
join ';'
         add-tocasereport -casename $cc.name -casetype $cc.casetype -casestatus
$cc.Status -caseclosedby $cc.closedby -caseClosedDateTime $cc.ClosedDateTime -
casemembers $cmembers
     }
     else {
         $cmembers = ((Get-ComplianceCaseMember -Case $cc.name).windowsLiveID) -
join ';'
         $policies = Get-CaseHoldPolicy -Case $cc.Name | % { Get-CaseHoldPolicy
$_.Name -Case $_.CaseId -DistributionDetail }
         if ($policies -ne $NULL) {
             foreach ($policy in $policies) {
                 $rule = Get-CaseHoldRule -Policy $policy.name
                 add-tocasereport -casename $cc.name -casetype $cc.casetype -
casemembers $cmembers -casestatus $cc.Status -casecreatedtime $cc.CreatedDateTime
-holdname $policy.name -holdenabled $policy.enabled -holdcreatedby
$policy.CreatedBy -holdlastmodifiedby $policy.LastModifiedBy -ExchangeLocation
(($policy.exchangelocation.name) -join ';') -SharePointLocation
(($policy.sharePointlocation.name) -join ';') -ContentMatchQuery
$rule.ContentMatchQuery -holdcreatedtime $policy.WhenCreatedUTC -holdchangedtime
$policy.WhenChangedUTC -holdstatus $policy.DistributionStatus -holderror
$policy.DistributionResults
             }
         }
         else {
             Write-Host "No hold policies found in case:" $cc.name -
foregroundColor 'Yellow'
             " "
             [string]$cc.name | out-file -filepath $noholdsfilepath -append
         }
     }
 }

 #get information on the cases and pass values to the case report function
 " "
 write-host "Gathering a list of eDiscovery (Premium) cases and holds..."
 " "
 $edc = Get-ComplianceCase -CaseType Advanced -ErrorAction SilentlyContinue
```

```
 foreach ($cc in $edc) {
     write-host "Working on case :" $cc.name
     if ($cc.status -eq 'Closed') {
         $cmembers = ((Get-ComplianceCaseMember -Case $cc.name).windowsLiveID) -
join ';'
         add-tocasereport -casename $cc.name -casestatus $cc.Status -casetype
$cc.casetype -caseclosedby $cc.closedby -caseClosedDateTime $cc.ClosedDateTime -
casemembers $cmembers
     }
     else {
         $cmembers = ((Get-ComplianceCaseMember -Case $cc.name).windowsLiveID) -
join ';'
         $policies = Get-CaseHoldPolicy -Case $cc.Name | % { Get-CaseHoldPolicy
$_.Name -Case $_.CaseId -DistributionDetail }
         if ($policies -ne $NULL) {
             foreach ($policy in $policies) {
                 $rule = Get-CaseHoldRule -Policy $policy.name
                 add-tocasereport -casename $cc.name -casetype $cc.casetype -
casemembers $cmembers -casestatus $cc.Status -casecreatedtime $cc.CreatedDateTime
-holdname $policy.name -holdenabled $policy.enabled -holdcreatedby
$policy.CreatedBy -holdlastmodifiedby $policy.LastModifiedBy -ExchangeLocation
(($policy.exchangelocation.name) -join ';') -SharePointLocation
(($policy.sharePointlocation.name) -join ';') -ContentMatchQuery
$rule.ContentMatchQuery -holdcreatedtime $policy.WhenCreatedUTC -holdchangedtime
$policy.WhenChangedUTC -holdstatus $policy.DistributionStatus -holderror
$policy.DistributionResults

             }
         }
         else {
             write-host "No hold policies found in case:" $cc.name -
foregroundColor 'Yellow'
             " "
             [string]$cc.name | out-file -filepath $noholdsfilepath -append
         }
     }
 }

 " "
 Write-host "Script complete! Report files saved to this folder: '$Path'"
 " "
 #script end
```

2. In the Windows PowerShell session that opened in Step 1, go to the folder where you saved the script.

3. Run the script; for example:

PowerShell

```
.\CaseHoldsReport.ps1
```

The script prompts for a target folder to save the report to.

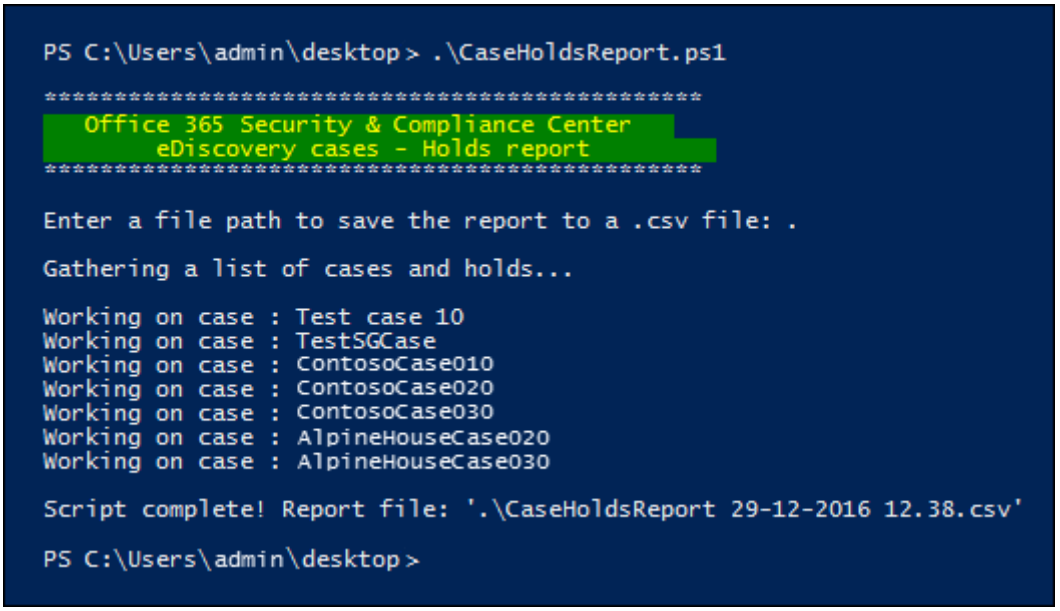
4. Type the full path name of the folder to save the report to, and then press **Enter**.

> **Tip**
>
> To save the report in the same folder that the script is located in, type a period (".") when prompted for a target folder. To save the report in a subfolder in the folder where the script is located, just type the name of the subfolder.

The script starts to collect information about all the eDiscovery cases in your organization. Don't access the report file while the script is running. After the script is complete, a confirmation message is displayed in the Windows PowerShell session. After this message is displayed, you can access the report in the folder that you specified in Step 4. The file name for the report is `CaseHoldsReport<DateTimeStamp>.csv`.

Additionally, the script also creates a report with a list of cases that don't have any holds. The file name for this report is `CaseswithNoHolds<DateTimeStamp>.csv`.

Here's an example of running the CaseHoldsReport.ps1 script.



More information

The case holds report that's created when you run the script in this article contains the following information about each hold. As previously explained, you have to be an eDiscovery Administrator to return information for all holds in your organization. For more information about case holds, see eDiscovery cases.

- The name of the hold and the name of the eDiscovery case that the hold is associated with.
- Whether the hold is associated with an eDiscovery (Standard) or eDiscovery (Premium) case.
- Whether or not the eDiscovery case is active or closed.
- Whether or not the hold is enabled or disabled.
- The members of the eDiscovery case that the hold is associated with. Case members can view or manage a case, depending on the eDiscovery permissions they've been assigned.
- The time and date the case was created.
- If a case is closed, the person who closed it and the time and date it was closed.
- The Exchange mailboxes and SharePoint sites locations that are on hold.
- If the hold is query-based, the query syntax.
- The time and date the hold was created and the person who created it.
- The time and date the hold was last changed and the person who changed it.

Last updated on 11/21/2025

Place a Microsoft Teams user or team on legal hold

Applies to: Microsoft Teams

> **⊗ Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

When a reasonable expectation of litigation exists, organizations are required to preserve electronically stored information (ESI), including Teams chat messages that are relevant to the case. Organizations may need to preserve all messages related to a specific investigation or for a specific person. This article will discuss using a legal hold to preserve content in Microsoft Teams. To preserve content in other services in Microsoft 365, see Create an eDiscovery hold.

> **ⓘ Note**
>
> Private channel chats are stored in user mailboxes, while standard channel chats are stored in the mailbox associated with the parent team. If there is already a legal hold in place for a user mailbox, the hold policy will now automatically apply to private channel messages stored in that mailbox. There is no further action needed for an admin to turn this on. Legal hold of files shared in private channels is also supported.

Within Microsoft Teams, an entire team or select users can be put on legal hold. Doing that will make sure that all messages that were exchanged in those teams (including private and shared channels) or messages exchanged by those individuals are discoverable by the organization's compliance managers or Teams Admins.

> **ⓘ Note**

Placing a user on hold does not automatically place a group on hold or vice-versa. Notifications sent in activity feeds can't be placed on hold.

To put a user or a team on legal hold, use the following guidance:

- Create a hold in eDiscovery (Standard)
- Create and manage a hold in eDiscover (Premium)

Important

When a user or group is placed on hold, all compliance copies of messages are preserved. For example, if a user posts a message in a channel and then modifies the message, both copies of the message are preserved. Without the hold, only the latest message is preserved.

Content locations to place on hold to preserve Teams content

As a helpful guide, use the following table to understand what content locations (such as a mailbox or a site) to place on hold to preserve different types of Teams content.

Expand table

Scenario	Content location
Chat messages for a user (for example, 1:1 chats, 1:N group chats, and private channel conversations)	User mailbox
Chat messages in standard and shared channels	Mailbox associated with the parent team
Files in standard channels (for example, Wiki content and files)	SharePoint site associated with the parent team
Files in private and shared channels	Dedicated SharePoint site associated with the channel
User's private content	The user's OneDrive for Business account
Card content in chats	User mailbox for 1:1 chats, 1:N group chats, and private channel conversations; the parent team mailbox for card content in standard

Scenario	Content location
	and shared channel messages. For more information, see the "Preserve card content" section in Create an eDiscovery hold.

> **Note**
>
> To retain message content in private channels, you need to put the user mailboxes (of the members of a private channel) on hold. and when using eDiscovery tool to search private channel messages, you have to search the user's mailbox. As was stated earlier, private channel chats are stored in user mailboxes, not in the group mailbox associated with the parent team.

Last updated on 11/21/2025

Use a script to add users to a hold in an eDiscovery (Standard) case

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

Security & Compliance PowerShell provides cmdlets that let you automate time-consuming tasks related to creating and managing eDiscovery cases. Currently, using the Microsoft Purview eDiscovery (Standard) case in the Microsoft Purview portal to place a large number of custodian content locations on hold takes time and preparation. For example, before you create a hold, you have to collect the URL for each OneDrive for Business site that you want to place on hold. Then for each user you want to place on hold, you have to add their mailbox and their OneDrive for Business site to the hold. You can use the script in this article to automate this process.

The script prompts you for the name of your organization's My Site domain (for example, `contoso` in the URL `https://contoso-my.sharepoint.com`), the name of an existing eDiscovery case, the name of the new hold that associated with the case, a list of email addresses of the users you want to put on hold, and a search query to use if you want to create a query-based hold. The script then gets the URL for the OneDrive for Business site for each user in the list, creates the new hold, and then adds the mailbox and OneDrive for Business site for each user in the list to the hold. The script also generates log files that contain information about the new hold.

Here are the steps to make this happen:

Step 1: Install the SharePoint Online Management Shell Step 2: Generate a list of users Step 3: Run the script to create a hold and add users

Before you add users to a hold

- You have to be a member of the eDiscovery Manager role group in the Microsoft Purview portal and a SharePoint Online administrator to run the script in Step 3. For more information, see Assign eDiscovery permissions in the Microsoft Purview portal.
- A maximum of 1,000 mailboxes and 100 sites can be added to a hold that's associated with an eDiscovery case in the Microsoft Purview portal. Assuming that every user that you want to place on hold has a OneDrive for Business site, you can add a maximum of 100 users to a hold using the script in this article.
- Be sure to save the list of users that you create in Step 2 and the script in Step 3 to the same folder. That will make it easier to run the script.
- The script adds the list of users to a new hold that is associated with an existing case. Be sure the case that you want to associate the hold with is created before you run the script.
- The script in this article supports modern authentication when connecting to Security & Compliance PowerShell and SharePoint Online Management Shell. You can use the script as-is if you are a Microsoft 365 or a Microsoft 365 GCC organization. If you are an Office 365 Germany organization, a Microsoft 365 GCC High organization, or a Microsoft 365 DoD organization, you will have to edit the script to successfully run it. Specifically, you have to edit the line `Connect-IPPSSession` and use the *ConnectionUri* and *AzureADAuthorizationEndpointUri* parameters (and the appropriate values for your organization type) to connect to Security & Compliance PowerShell. For more information, see the examples in Connect to Security & Compliance PowerShell.
- The script automatically disconnects from Security & Compliance PowerShell and SharePoint Online Management Shell.
- The script includes minimal error handling. Its primary purpose is to quickly and easily place the mailbox and OneDrive for Business site of each user on hold.
- The sample scripts provided in this topic aren't supported under any Microsoft standard support program or service. The sample scripts are provided AS IS without warranty of any kind. Microsoft further disclaims all implied warranties including, without limitation, any implied warranties of merchantability or of fitness for a particular purpose. The entire risk arising out of the use or performance of the sample scripts and documentation remains with you. In no event shall Microsoft, its authors, or anyone else involved in the creation, production, or delivery of the scripts be liable for any damages whatsoever (including, without limitation, damages for loss of business profits, business interruption, loss of business information, or other pecuniary loss) arising out of the use of or inability to use the sample scripts or documentation, even if Microsoft has been advised of the possibility of such damages.

Step 1: Install the SharePoint Online Management Shell

The first step is to install the SharePoint Online Management Shell if it's not already installed on your local computer. You don't have to use the shell in this procedure, but you have to install it because it contains pre-requisites required by the script that you run in Step 3. These prerequisites allow the script to communicate with SharePoint Online to get the URLs for the OneDrive for Business sites.

Go to Set up the SharePoint Online Management Shell environment and perform Step 1 and Step 2 to install the SharePoint Online Management Shell on your local computer.

Step 2: Generate a list of users

The script in Step 3 will create a hold that's associated with an eDiscovery case, and the add the mailboxes and OneDrive for Business sites of a list of users to the hold. You can just type the email addresses in a text file, or you can run a command in PowerShell to get a list of email addresses and save them to a file (located in same folder that you'll save the script to in Step 3).

Here's an Exchange Online PowerShell command to get a list of email addresses for all users in your organization and save it to a text file named HoldUsers.txt.

PowerShell

```
Get-Mailbox -ResultSize unlimited -Filter { RecipientTypeDetails -eq 'UserMailbox'} |
Select-Object PrimarySmtpAddress > HoldUsers.txt
```

After you run this command, open the text file and remove the header that contains the property name, `PrimarySmtpAddress`. Then remove all email addresses except the ones for the users that you want to add to the hold that you'll create in Step 3. Make sure there are no blank rows before or after the list of email addresses.

Step 3: Run the script to create a hold and add users

When you run the script in this step, it will prompt you for the following information. Be sure to have this information ready before you run the script.

- **Your user credentials:** The script will use your credentials to connect to Security & Compliance PowerShell. It will also use these credentials to access SharePoint Online to get the OneDrive for Business URLs for the list of users.
- **Name of your SharePoint domain:** The script prompts you to enter this name so it can connect to the SharePoint admin center. It also uses the domain name for the OneDrive URLs in your organization. For example, if the URL for your admin center is

`https://contoso-admin.sharepoint.com` and the URL for OneDrive is `https://contoso-my.sharepoint.com`, then you would enter `contoso` when the script prompts you for your domain name.

- **Name of the case:** The name of an existing case. The script will create a new hold that is associated with this case.
- **Name of the hold:** The name of the hold the script will create and associate with the specified case.
- **Search query for a query-based hold:** You can create a query-based hold so that only the content that meets the specified search criteria is placed on hold. To place all content on hold, just press **Enter** when you're prompted for a search query.
- **Turning on the hold or not:** You can have the script turn on the hold after it's created or you can have the script create the hold without enabling it. If you don't have the script turn on the hold, you can turn it on later in the Microsoft Purview portal or by running the following PowerShell commands:

```
PowerShell
Set-CaseHoldPolicy -Identity <name of the hold> -Enabled $true
```

```
PowerShell
Set-CaseHoldRule -Identity <name of the hold> -Disabled $false
```

- **Name of the text file with the list of users** - The name of the text file from Step 2 that contains the list of users to add to the hold. If this file is located in the same folder as the script, just type the name of the file (for example, HoldUsers.txt). If the text file is in another folder, type the full pathname of the file.

After you've collected the information that the script will prompt you for, the final step is to run the script to create the new hold and add users to it.

1. Save the following text to a Windows PowerShell script file by using a filename suffix of `.ps1`. For example, `AddUsersToHold.ps1`.

```
PowerShell
#script begin
" "
```

```
write-host "*********************************************"
write-host "   Security & Compliance PowerShell  " -foregroundColor yellow -
backgroundcolor darkgreen
write-host "   eDiscovery (Standard) cases - Add users to a hold   " -
foregroundColor yellow -backgroundcolor darkgreen
write-host "*********************************************"
" "
# Connect to Security & Compliance PowerShell using modern authentication
if (!$SccSession)
{
  Import-Module ExchangeOnlineManagement
  Connect-IPPSSession
}

# Get the organization's domain name. We use this to create the SharePoint admin
URL and root URL for OneDrive for Business.
""
$mySiteDomain = Read-Host "Enter the domain name for your SharePoint
organization. We use this name to connect to SharePoint admin center and for the
OneDrive URLs in your organization. For example, 'contoso' in 'https://contoso-
admin.sharepoint.com' and 'https://contoso-my.sharepoint.com'"
""

# Connect to PnP Online using modern authentication
Import-Module PnP.PowerShell
Connect-PnPOnline -Url https://$mySiteDomain-admin.sharepoint.com -UseWebLogin

# Load the SharePoint assemblies from the SharePoint Online Management Shell
# To install, go to https://go.microsoft.com/fwlink/p/?LinkId=255251
if (!$SharePointClient -or !$SPRuntime -or !$SPUserProfile)
{
    $SharePointClient =
[System.Reflection.Assembly]::LoadWithPartialName("Microsoft.SharePoint.Client")
    $SPRuntime =
[System.Reflection.Assembly]::LoadWithPartialName("Microsoft.SharePoint.Client.Ru
ntime")
    $SPUserProfile =
[System.Reflection.Assembly]::LoadWithPartialName("Microsoft.SharePoint.Client.Us
erProfiles")
    if (!$SharePointClient)
    {
        Write-Error "The SharePoint Online Management Shell isn't installed.
Please install it from: https://go.microsoft.com/fwlink/p/?LinkId=255251 and then
re-run this script."
        return;
    }
}

# Get other required information
do{
$casename = Read-Host "Enter the name of the case"
$caseexists = (get-compliancecase -identity "$casename" -erroraction
SilentlyContinue).isvalid
if($caseexists -ne 'True')
{""
```

```
write-host "A case named '$casename' doesn't exist. Please specify the name of an existing case, or create a new case and then re-run the script." -foregroundColor Yellow
""}
}While($caseexists -ne 'True')
""
do{
$holdName = Read-Host "Enter the name of the new hold"
$holdexists=(get-caseholdpolicy -identity "$holdname" -case "$casename" -erroraction SilentlyContinue).isvalid
if($holdexists -eq 'True')
{""
write-host "A hold named '$holdname' already exists. Please specify a new hold name." -foregroundColor Yellow
""}
}While($holdexists -eq 'True')
""
$holdQuery = Read-Host "Enter a search query to create a query-based hold, or press Enter to hold all content"
""
$holdstatus = read-host "Do you want the hold enabled after it's created? (Yes/No)"
do{
""
$inputfile = read-host "Enter the name of the text file that contains the email addresses of the users to add to the hold"
""
$fileexists = test-path -path $inputfile
if($fileexists -ne 'True'){write-host "$inputfile doesn't exist. Please enter a valid file name." -foregroundcolor Yellow}
}while($fileexists -ne 'True')
#Import the list of addresses from the txt file.  Trim any excess spaces and make sure all addresses
    #in the list are unique.
  [array]$emailAddresses = Get-Content $inputfile -ErrorAction SilentlyContinue | where {$_.trim() -ne ""}  | foreach{ $_.Trim() }
  [int]$dupl = $emailAddresses.count
  [array]$emailAddresses = $emailAddresses | select-object -unique
  $dupl -= $emailAddresses.count
#Validate email addresses so the hold creation does not run in to an error.
if($emailaddresses.count -gt 0){
write-host ($emailAddresses).count "addresses were found in the text file. There were $dupl duplicate entries in the file." -foregroundColor Yellow
""
Write-host "Validating the email addresses. Please wait..." -foregroundColor Yellow
""
$finallist =@()
foreach($emailAddress in $emailAddresses)
{
if((get-recipient $emailaddress -erroraction SilentlyContinue).isvalid -eq 'True')
{$finallist += $emailaddress}
else {"Unable to find the user $emailaddress"
[array]$excludedlist += $emailaddress}
```

```
}
""
#Find user's OneDrive account URL using email address
Write-Host "Getting the URL for each user's OneDrive for Business site." -
foregroundColor Yellow
""
$AdminUrl = "https://$mySiteDomain-admin.sharepoint.com"
$mySiteUrlRoot = "https://$mySiteDomain-my.sharepoint.com"
$urls = @()
foreach($emailAddress in $finallist)
{
try
{
$url=Get-PnPUserProfileProperty -Account $emailAddress | Select PersonalUrl
$urls += $url.PersonalUrl
       Write-Host "- $emailAddress => $url"
       [array]$ODadded += $url.PersonalUrl
       }catch {
 Write-Warning "Could not locate OneDrive for $emailAddress"
 [array]$ODExluded += $emailAddress
 Continue }
}
$urls | FL
if(($finallist.count -gt 0) -or ($urls.count -gt 0)){
""
Write-Host "Creating the hold named $holdname. Please wait..." -foregroundColor
Yellow
if(($holdstatus -eq "Y") -or ($holdstatus -eq  "y") -or ($holdstatus -eq "yes") -
or ($holdstatus -eq "YES")){
New-CaseHoldPolicy -Name "$holdName" -Case "$casename" -ExchangeLocation
$finallist -SharePointLocation $urls -Enabled $True | out-null
New-CaseHoldRule -Name "$holdName" -Policy "$holdname" -ContentMatchQuery
$holdQuery | out-null
}
else{
New-CaseHoldPolicy -Name "$holdName" -Case "$casename" -ExchangeLocation
$finallist -SharePointLocation $urls -Enabled $false | out-null
New-CaseHoldRule -Name "$holdName" -Policy "$holdname" -ContentMatchQuery
$holdQuery -disabled $false | out-null
}
""
}
else {"No valid locations were identified. Therefore, the hold wasn't created."}
#write log files (if needed)
$newhold=Get-CaseHoldPolicy -Identity "$holdname" -Case "$casename" -erroraction
SilentlyContinue
$newholdrule=Get-CaseHoldRule -Identity "$holdName" -erroraction SilentlyContinue
if(($ODAdded.count -gt 0) -or ($ODExluded.count -gt 0) -or ($finallist.count -gt
0) -or ($excludedlist.count -gt 0) -or ($newhold.isvalid -eq 'True') -or
($newholdrule.isvalid -eq 'True'))
{
Write-Host "Generating output files..." -foregroundColor Yellow
if($ODAdded.count -gt 0){
"OneDrive Locations" | add-content .\LocationsOnHold.txt
"==================" | add-content .\LocationsOnHold.txt
```

```
$newhold.SharePointLocation.name | add-content .\LocationsOnHold.txt}
if($ODExluded.count -gt 0){
"Users without OneDrive locations" | add-content .\LocationsNotOnHold.txt
"================================" | add-content .\LocationsNotOnHold.txt
$ODExluded | add-content .\LocationsNotOnHold.txt}
if($finallist.count -gt 0){
" " | add-content .\LocationsOnHold.txt
"Exchange Locations" | add-content .\LocationsOnHold.txt
"==================" | add-content .\LocationsOnHold.txt
$newhold.ExchangeLocation.name | add-content .\LocationsOnHold.txt}
if($excludedlist.count -gt 0){
" "| add-content .\LocationsNotOnHold.txt
"Mailboxes not added to the hold" | add-content .\LocationsNotOnHold.txt
"================================" | add-content .\LocationsNotOnHold.txt
$excludedlist | add-content .\LocationsNotOnHold.txt}
$FormatEnumerationLimit=-1
if($newhold.isvalid -eq 'True'){$newhold|fl >.\GetCaseHoldPolicy.txt}
if($newholdrule.isvalid -eq 'True'){$newholdrule|Fl >.\GetCaseHoldRule.txt}
}
}
else {"The hold wasn't created because no valid entries were found in the text
file."}
""
#Disconnect from SCC PowerShell and PnPOnline

Write-host "Disconnecting from SCC PowerShell and PnP Online" -foregroundColor
Yellow
Get-PSSession | Remove-PSSession
Disconnect-PnPOnline

Write-host "Script complete!" -foregroundColor Yellow
""
#script end
```

2. On your local computer, open Windows PowerShell and go to the folder where you saved the script.

3. Run the script; for example:

   PowerShell

   ```
   .\AddUsersToHold.ps1
   ```

4. Enter the information that the script prompts you for.

   The script connects to Security & Compliance PowerShell, and then creates the new hold in the eDiscovery case and adds the mailboxes and OneDrive for Business for the users in the list. You can go to the case on the **eDiscovery** page in the Microsoft Purview portal to view the new hold.

After the script is finished running, it creates the following log files, and saves them to the folder where the script is located.

- **LocationsOnHold.txt:** Contains a list of mailboxes and OneDrive for Business sites that the script successfully placed on hold.
- **LocationsNotOnHold.txt:** Contains a list of mailboxes and OneDrive for Business sites that the script did not place on hold. If a user has a mailbox, but not a OneDrive for Business site, the user would be included in the list of OneDrive for Business sites that weren't placed on hold.
- **GetCaseHoldPolicy.txt:** Contains the output of the **Get-CaseHoldPolicy** cmdlet for the new hold, which the script ran after creating the new hold. The information returned by this cmdlet includes a list of users whose mailboxes and OneDrive for Business sites were placed on hold and whether the hold is enabled or disabled.
- **GetCaseHoldRule.txt:** Contains the output of the **Get-CaseHoldRule** cmdlet for the new hold, which the script ran after creating the new hold. The information returned by this cmdlet includes the search query if you used the script to create a query-based hold.

Last updated on 09/02/2025

How to identify the type of hold placed on an Exchange Online mailbox

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

This article explains how to identify holds placed on Exchange Online mailboxes in Microsoft Purview and Microsoft 365.

Microsoft Purview offers several ways that your organization can prevent mailbox content from being permanently deleted. This allows your organization to retain content to meet compliance regulations or during legal and other types of investigations. Here's a list of the retention features (also called *holds*) in Microsoft Purview and Microsoft 365:

- **Litigation Hold**: Holds that are applied to user mailboxes in Exchange Online.
- **eDiscovery hold**: Holds that are associated with a Microsoft Purview eDiscovery (Standard) case in the Microsoft Purview portal. eDiscovery holds can be applied to user mailboxes and to the corresponding mailbox for Microsoft 365 Groups and Microsoft Teams.
- **In-Place Hold**: Holds that are applied to user mailboxes by using the In-Place eDiscovery & Hold tool in the Exchange admin center in Exchange Online.

  > **Note**
  >
  > In-Place Holds have been retired and you can no longer create In-Place Holds or apply them to mailboxes. However, In-Place Holds might still be applied to mailboxes in your organization, which is why they are included in this article. For more information, see Retirement of legacy eDiscovery tools.

- **Microsoft Purview retention policies:** Can be configured to retain (or retain and then delete) content in user mailboxes in Exchange Online and in the corresponding mailbox for Microsoft 365 Groups and Microsoft Teams. You can also create a retention policy to retain Skype for Business Conversations, which are stored in user mailboxes.

  There are two types of Microsoft Purview retention policies that can be assigned to mailboxes.
  - **Specific location retention policies:** These are policies that are assigned to the content locations of specific users. You use the **Get-Mailbox** cmdlet in Exchange Online PowerShell to get information about retention policies assigned to specific mailboxes. For more information about this type of retention policy, see the section A policy with specific inclusions or exclusions from the retention policy documentation.
  - **Organization-wide retention policies:** These are policies that are assigned to all content locations in your organization. You use the **Get-OrganizationConfig** cmdlet in Exchange Online PowerShell to get information about organization-wide retention policies. For more information about this type of retention policy, see the section A policy that applies to entire locations from the retention policy documentation.

- **Microsoft Purview retention labels:** If a user applies a Microsoft Purview retention label (one that's configured to retain content or retain and then delete content) to *any* folder or item in their mailbox, a hold is placed on the mailbox as if the mailbox was placed on Litigation Hold or assigned to a Microsoft Purview retention policy. For more information, see the Identifying mailboxes on hold because a retention label has been applied to a folder or item section in this article.

To manage mailboxes on hold, you may have to identify the type of hold that's placed on a mailbox so that you can perform tasks such as changing the hold duration, temporarily or permanently removing the hold, or excluding a mailbox from a Microsoft Purview retention policy. In these cases, the first step is to identify the type of hold placed on the mailbox. And because multiple holds (and different types of holds) can be placed on a single mailbox, you have to identify all holds placed on a mailbox if you want to remove or change a hold.

Step 1: Obtain the GUID for holds placed on a mailbox

You can run the following two cmdlets in Exchange Online PowerShell to get the GUID of the holds that are placed on a mailbox. After you obtain a GUID, you use it to identify the specific hold in Step 2. A Litigation Hold isn't identified by a GUID. Litigation Holds are either enabled or disabled for a mailbox.

- **Get-Mailbox:** Use this cmdlet to determine whether Litigation Hold is enabled for a mailbox and to get the GUIDs for eDiscovery holds, In-Place Holds, and Microsoft Purview retention policies that are assigned to a mailbox. The output of this cmdlet will also indicate if a mailbox has been explicitly excluded from an organization-wide retention policy.
- **Get-OrganizationConfig:** Use this cmdlet to get the GUIDs for organization-wide retention policies.

To connect to Exchange Online PowerShell, see Connect to Exchange Online PowerShell.

Get-Mailbox

Run the following command to get information about the holds and Microsoft Purview retention policies applied to a mailbox.

PowerShell

```
Get-Mailbox <username> | FL LitigationHoldEnabled,InPlaceHolds
```

> **Tip**
>
> If there are too many values in the InPlaceHolds property and not all of them are displayed, you can run the `Get-Mailbox <username> | Select-Object -ExpandProperty InPlaceHolds` command to display each GUID on a separate line.

The following table describes how to identify different types of holds based on the values in the *InPlaceHolds* property when you run the **Get-Mailbox** cmdlet.

Expand table

Hold type	Example value	How to identify the hold
Litigation Hold	`True`	Litigation Hold is enabled for a mailbox when the *LitigationHoldEnabled* property is set to `True`.
eDiscovery hold	`UniH7d895d48-7e23-4a8d-8346-533c3beac15d`	The *InPlaceHolds property* contains the GUID of any hold associated with an eDiscovery case in the Microsoft Purview portal. You can tell this is an eDiscovery hold because the GUID starts with the `UniH` prefix (which denotes a Unified Hold).

Hold type	Example value	How to identify the hold
In-Place Hold	`c0ba3ce811b6432a8751430937152491` or `cld9c0a984ca74b457fbe4504bf7d3e00de`	The *InPlaceHolds* property contains the GUID of the In-Place Hold that's placed on the mailbox. You can tell this is an In-Place Hold because the GUID either doesn't start with a prefix or it starts with the `cld` prefix.
Microsoft Purview retention policy applied to the mailbox	`mbxcdbbb86ce60342489bff371876e7f224:1` or `skp127d7cf1076947929bf136b7a2a8c36f:3`	The *InPlaceHolds* property contains GUIDs of any specific location retention policy that's applied to the mailbox. You can identify retention policies because the GUID starts with the `mbx` or the `skp` prefix. The `skp` prefix indicates that the retention policy is applied to Skype for Business conversations in the user's mailbox.
Excluded from an organization-wide Microsoft Purview retention policy	`-mbxe9b52bf7ab3b46a286308ecb29624696`	If a mailbox is excluded from an organization-wide Microsoft Purview retention policy, the GUID for the retention policy that the mailbox is excluded from is displayed in the *InPlaceHolds* property and is identified by the `-mbx` prefix.

Get-OrganizationConfig

If the *InPlaceHolds* property is empty when you run the **Get-Mailbox** cmdlet, there still may be one or more organization-wide Microsoft Purview retention policies applied to the mailbox. Run the following command in Exchange Online PowerShell to get a list of GUIDs for organization-wide Microsoft Purview retention policies.

PowerShell

```
Get-OrganizationConfig | FL InPlaceHolds
```

> **Tip**
>
> If there are too many values in the InPlaceHolds property and not all of them are displayed, you can run the `Get-OrganizationConfig | Select-Object -ExpandProperty InPlaceHolds` command to display each GUID on a separate line.

The following table describes the different types of organization-wide holds and how to identify each type based on the GUIDs contained in *InPlaceHolds* property when you run the **Get-**

OrganizationConfig cmdlet.

Expand table

Hold type	Example value	Description
Microsoft Purview retention policies applied to Exchange mailboxes, Exchange public folders, and Teams chats	`mbx7cfb30345d454ac0a989ab3041051209:2`	Organization-wide retention policies applied to Exchange mailboxes, Exchange public folders, and 1xN chats in Microsoft Teams are identified by GUIDs that start with the `mbx` prefix. Note 1xN chats are stored in the mailbox of the individual chat participants.
Microsoft Purview retention policy applied to Microsoft 365 Groups and Teams channel messages	`grp1a0a132ee8944501a4bb6a452ec31171:3`	Organization-wide retention policies applied to Microsoft 365 groups and channel messages in Microsoft Teams are identified by GUIDs that start with the `grp` prefix. Note channel messages are stored in the group mailbox that is associated with a Microsoft Team.

For more information about retention policies applied to Microsoft Teams, see Learn about retention policies for Microsoft Teams.

Understanding the format of the InPlaceHolds value for retention policies

In addition to the prefix (mbx, skp, or grp) that identifies an item in the InPlaceHolds property as a Microsoft Purview retention policy, the value also contains a suffix that identifies the type of retention action that's configured for the policy. For example, the action suffix is highlighted in bold type in the following examples:

`skp127d7cf1076947929bf136b7a2a8c36f`:1

`mbx7cfb30345d454ac0a989ab3041051209`:2

`grp1a0a132ee8944501a4bb6a452ec31171`:3

The following table defines the three possible retention actions:

Expand table

Value	Description
1	Indicates that the retention policy is configured to delete items or the retention policy is used to publish retention labels. If policy is publishing retention labels, retention configuration for each published label have to be checked.
2	Indicates that the retention policy is configured to hold items. The policy doesn't delete items after the retention period expires.
3	Indicates that the retention policy is configured to hold items and then delete them after the retention period expires.

> **Note**
>
> Because retention label policies publish or auto-apply labels that apply item-level actions, they will always show an action value of 1 within the InPlaceHolds property of the mailbox.
>
> To identify whether a hold has been applied to a folder or item within the mailbox, refer to Identifying mailboxes on hold because a retention label has been applied to a folder or item.

For more information about retention actions, see the Retaining content for a specific period of time section.

Step 2: Use the GUID to identify the hold

After you obtain the GUID for a hold that is applied to a mailbox, the next step is to use that GUID to identify the hold. The following sections show how to identify the name of the hold (and other information) by using the hold GUID.

eDiscovery holds

Run the following commands in Security & Compliance PowerShell to identify an eDiscovery hold that's applied to the mailbox. Use the GUID (not including the UniH prefix) for the eDiscovery hold that you identified in Step 1.

To connect to Security & Compliance PowerShell, see Connect to Security & Compliance PowerShell.

The first command creates a variable that contains information about the hold. This variable is used in the other commands. The second command displays the name of the eDiscovery case the

hold is associated with. The third command displays the name of the hold and a list of the mailboxes the hold applies to.

PowerShell

```
$CaseHold = Get-CaseHoldPolicy <hold GUID without prefix>
```

PowerShell

```
Get-ComplianceCase $CaseHold.CaseId | FL Name
```

PowerShell

```
$CaseHold | FL Name,ExchangeLocation
```

In-Place Holds

Run the following command in Exchange Online PowerShell to identify the In-Place Hold that's applied to the mailbox. Use the GUID for the In-Place Hold that you identified in Step 1. The command displays the name of the hold and a list of the mailboxes the hold applies to.

PowerShell

```
Get-ComplianceSearch -Identity <hold GUID> | FL Name,SourceMailboxes
```

If the GUID for the In-Place Hold starts with the `cld` prefix, be sure to include the prefix when running the previous command.

> **ⓘ Important**
>
> As we continue to invest in different ways to preserve mailbox content, we're announcing the retirement of In-Place Holds in the Exchange admin center (EAC). Starting July 1, 2020 you won't be able to create new In-Place Holds in Exchange Online. But you'll still be able to manage In-Place Holds in the EAC or by using the **Set-MailboxSearch** cmdlet in Exchange Online PowerShell. However, starting October 1, 2020, you won't be able to manage In-Place Holds. You'll only be remove them in the EAC or by using the **Remove-MailboxSearch** cmdlet. For more information about the retirement of In-Place Holds, see Retirement of legacy eDiscovery tools.

Microsoft Purview retention policies

Connect to Security & Compliance PowerShell and run the following command to identity the Microsoft Purview retention policy (organization-wide or specific location) that's applied to the mailbox. Use the GUID (not including the mbx, skp, or grp prefix or the action suffix) that you identified in Step 1.

PowerShell

```
Get-RetentionCompliancePolicy <hold GUID without prefix or suffix> -DistributionDetail
| FL Name,*Location
```

New retention policies that apply to the following locations don't stamp directly on organization configuration or Exchange Online objects:

- Enterprise AI apps
- Microsoft Copilot experiences
- Teams chats
- Teams private channel messages
- Viva Engage community messages
- Viva Engage user messages
- Other AI apps

To determine whether these retention compliance policies apply to a mailbox during an investigation, use Policy Lookup or Retention cmdlets for newer locations.

PowerShell

```
Get-AppRetentionCompliancePolicy <hold GUID without prefix or suffix> -
DistributionDetail  | FL Name, Applications, *Location
```

Identifying mailboxes on hold because a retention label has been applied to a folder or item

Whenever a user applies a retention label that's configured to *retain* or *retain and then delete* content to any folder or item in their mailbox, the *ComplianceTagHoldApplied* mailbox property is set to **True**. When this happens, the mailbox is treated similarly to if it was placed on hold, such as when assigned to a Microsoft Purview retention policy or placed on Litigation Hold, however with some caveats. When the *ComplianceTagHoldApplied* property is set to **True**, the following things occur:

- If the mailbox or the user's Microsoft 365 account is deleted, the mailbox becomes an inactive mailbox.
- You aren't able to disable the mailbox (either the primary mailbox or the archive mailbox, if it's enabled).
- Items that have been deleted from the mailbox will follow one of two paths depending on if they're labeled or not:
  - **Unlabeled items** will follow the same path deleted items take when no holds apply to the mailbox. The time that it takes for these items to be permanently deleted is determined by the deleted item retention configuration and whether single item recovery is enabled for the mailbox or not.
  - **Labeled items** will be retained within the recoverable items folder in the same way they would be if a Microsoft Purview retention policy applied, but at the individual item level. If multiple items have different labels that are configured to *retain* or *retain and then delete* content at different intervals, each item will be retained based on the configuration of the applied label.
- Other holds, such as Microsoft Purview retention policies, eDiscovery holds or litigation hold can extend how long labeled items are retained based on the principles of retention.

To view the value of the *ComplianceTagHoldApplied* property for a single mailbox, run the following command in Exchange Online PowerShell:

PowerShell

```
Get-Mailbox <username> | FL ComplianceTagHoldApplied
```

For more information about retention labels, see retention labels.

Managing mailboxes on delay hold

After any type of hold is removed from a mailbox, a *delay hold* is applied. This means that the actual removal of the hold is delayed for 30 days to prevent data from being permanently deleted (purged) from the mailbox. This gives admins an opportunity to search for or recover mailbox items that will be purged after a hold is removed. A delay hold is placed on a mailbox the next time the Managed Folder Assistant processes the mailbox and detects that a hold was removed. Specifically, a delay hold is applied to a mailbox when the Managed Folder Assistant sets one of the following mailbox properties to **True**:

- **DelayHoldApplied:** This property applies to email-related content (generated by people using Outlook and Outlook on the web) that's stored in a user's mailbox.

- **DelayReleaseHoldApplied:** This property applies to cloud-based content (generated by non-Outlook apps such as Microsoft Teams, Microsoft Forms, and Microsoft Viva Engage) that's stored in a user's mailbox. Cloud data generated by a Microsoft app is typically stored in a hidden folder in a user's mailbox.

When a delay hold is placed on the mailbox (when either of the previous properties is set to **True**), the mailbox is still considered to be on hold for an unlimited hold duration, as if the mailbox was on Litigation Hold. After 30 days, the delay hold expires, and Microsoft 365 will automatically attempt to remove the delay hold (by setting the *DelayHoldApplied* or *DelayReleaseHoldApplied* property to **False**) so that the hold is removed. After either of these properties are set to **False**, the corresponding items that are marked for removal are purged the next time the mailbox is processed by the Managed Folder Assistant.

> ⓘ **Note**
>
> If the user account for the mailbox is disabled, the mailbox isn't processed by the Managed Folder Assistant and the delay hold remains after the 30 days have expired. For more information, see Delay hold considerations.

To view the values for the DelayHoldApplied and DelayReleaseHoldApplied properties for a mailbox, run the following command in Exchange Online PowerShell.

PowerShell

```
Get-Mailbox <username> | FL *HoldApplied*
```

To remove the delay hold before it expires, you can run one (or both) the following commands in Exchange Online PowerShell, depending on which property you want to change:

PowerShell

```
Set-Mailbox <username> -RemoveDelayHoldApplied
```

Or

PowerShell

```
Set-Mailbox <username> -RemoveDelayReleaseHoldApplied
```

You must be assigned the Legal Hold role in Exchange Online to use the *RemoveDelayHoldApplied* or *RemoveDelayReleaseHoldApplied* parameters.

To remove the delay hold on an inactive mailbox, run one of the following commands in Exchange Online PowerShell:

PowerShell

```
Set-Mailbox <DN or Exchange GUID> -InactiveMailbox -RemoveDelayHoldApplied
```

Or

PowerShell

```
Set-Mailbox <DN or Exchange GUID> -InactiveMailbox -RemoveDelayReleaseHoldApplied
```

> **Tip**
>
> The best way to specify an inactive mailbox in the previous command is to use its Distinguished Name or Exchange GUID value. Using one of these values helps prevent accidentally specifying the wrong mailbox.

For more information about using these parameters for managing delay holds, see Set-Mailbox.

Delay hold considerations

Keep the following things in mind when managing a mailbox on delay hold:

- If either the *DelayHoldApplied* or *DelayReleaseHoldApplied* property is set to **True** and a mailbox (or the corresponding user account) is deleted, the mailbox becomes an inactive mailbox. That's because a mailbox is considered to be on hold if either property is set to **True**, and deleting a mailbox on hold results in an inactive mailbox. To delete a mailbox and not make it an inactive mailbox, you have to set both properties to **False**.
- A mailbox is considered to be on hold for an unlimited hold duration if either the *DelayHoldApplied* or *DelayReleaseHoldApplied* property is set to **True**. However, that doesn't mean that *all* content in the mailbox is preserved. It depends on the value that's set to each property. For example, let's say both properties are set to **True** because holds are removed from the mailbox. Then you remove only the delay hold that's applied to non-Outlook cloud data (by using the *RemoveDelayReleaseHoldApplied* parameter). The next time the Managed Folder Assistant processes the mailbox, the non-Outlook items marked for removal are

purged. Any Outlook items marked for removal won't be purged because the DelayHoldApplied property is still set to **True**. The opposite would also be true: if *DelayHoldApplied* is set to **False** and *DelayReleaseHoldApplied* is set to **True**, then only Outlook items marked for removal would be purged.

How to confirm that an organization-wide retention policy is applied to a mailbox

When an organization-wide retention policy is applied or removed to a mailbox, exporting the mailbox diagnostics logs can help you be certain that Exchange Online has applied or removed the retention policy to the mailbox. To view this information, you first need to validate a few things using Exchange Online PowerShell.

Obtain the GUIDs for any retention policies explicitly applied to a mailbox

PowerShell

```
Get-Mailbox <username> | Select-Object -ExpandProperty InPlaceHolds
```

Obtain the GUIDs for any organization-wide retention policies applied to mailboxes

PowerShell

```
Get-OrganizationConfig | Select-Object -ExpandProperty InPlaceHolds
```

Get the Mailbox Diagnostics for HoldTracking

The Hold Tracking Mailbox Diagnostics logs maintain a history of the holds applied to a user mailbox. If the object is inactive, include `-IncludeInactiveMailboxes` parameter.

PowerShell

```
$ht = Export-MailboxDiagnosticLogs <username> -ComponentName HoldTracking
$ht.MailboxLog | Convertfrom-Json
```

Get the Mailbox Diagnostics for Substrate HoldTracking

The Substrate Hold Tracking Mailbox Diagnostics logs maintain a history of the holds applied to a user mailbox. If the object is inactive, include `-IncludeInactiveMailboxes` parameter.

PowerShell

```
$sht = Export-MailboxDiagnosticLogs <username> -ComponentName SubstrateHoldTracking
$sht.MailboxLog | Convertfrom-Json
```

Review the results of the Mailbox Diagnostics logs

If you gather data from the previous step, the resulting data may look something like this:

> **ed** `: 0001-01-01T00:00:00.0000000` **hid** `: mbx7cfb30345d454ac0a989ab3041051209:1` **ht** `: 4`
> **lsd** `: 2020-03-23T18:24:37.1884606Z` **osd** `: 2020-03-23T18:24:37.1884606Z`

Use the following table to help you understand each of the previous values listed in the diagnostics log.

Expand table

Value	Description
ed	Indicates the End date, which is the date the retention policy was disabled. MinValue means the policy is still assigned to the mailbox.
hid	Indicates the GUID for the retention policy. This value will correlate to the GUIDs that you collected for the explicit or organization-wide retention policies assigned to the mailbox.
ht	Indicates the hold type. Values are *0* for LitigationHold, *1* for InPlaceHold, *2* for ComplianceTagHold, *3* for DelayReleaseHold, *4* for OrganizationRetention, *5* for CompliancePolicy, *6* for SubstrateAppPolicy, and *7* for SharepointPolicy.
lsd	Indicates the Last start date, which is the date the retention policy was assigned to the mailbox.
osd	Indicates the Original start date, which is the date that Exchange first recorded information about the retention policy.

When a retention policy is no longer applied to a mailbox, we'll place a temporary delay hold on the user to prevent purging content. A delay hold can be disabled by running the `Set-Mailbox -RemoveDelayHoldApplied` command.

Next steps

After you identify the holds that are applied to a mailbox, you can perform tasks such as changing the duration of the hold, temporarily or permanently removing the hold, or excluding an inactive mailbox from a Microsoft Purview retention policy. For more information about performing tasks related to holds, see one of the following articles:

- Run the Set-RetentionCompliancePolicy -Identity <Policy Name> -AddExchangeLocationException <user mailbox> command in Security & Compliance PowerShell to exclude a mailbox from an organization-wide Microsoft Purview retention policy. This command can only be used for retention policies where the value for the *ExchangeLocation* property equals `All`.
- Run the Set-AppRetentionCompliancePolicy -Identity <Policy Name> -AddExchangeLocationException <user mailbox> command in Security & Compliance PowerShell to exclude a mailbox from a Teams private channel messages, Viva Engage user messages, Viva Engage community messages, Teams chats, Microsoft Copilot experiences, Enterprise AI apps, and other AI apps. You can only use this command for retention policies where the value for the *ExchangeLocation* property equals `All`.
- Run the Set-Mailbox <\DN or Exchange GUID> -RemoveDelayHoldApplied if needed to remove DelayHoldApplied if that's True. If the object is inactive, include the `-InactiveMailbox` parameter.
- Run the Set-Mailbox <\DN or Exchange GUID> -RemoveDelayReleaseHoldApplied if needed to remove DelayReleaseHoldApplied if that's True. If the object is inactive, include the `-InactiveMailbox` parameter.
- Run the Set-MailUser <\DN or Exchange GUID> -RemoveDelayHoldApplied if needed to remove DelayReleaseHoldApplied if that's True.
- Run the Set-MailUser <\DN or Exchange GUID> -RemoveDelayReleaseHoldApplied if needed to remove DelayReleaseHoldApplied if that's True.
- Run the Set-Mailbox <\DN or Exchange GUID> -LitigationHoldEnabled $False if needed to remove LitigationHold. If the object is inactive, include the `-InactiveMailbox` parameter.
- Run the Set-MailUser <\DN or Exchange GUID> -RemoveLitigationHoldEnabled if needed to remove LitigationHold.
- Run the Set-Mailbox <\DN or Exchange GUID> -InactiveMailbox -RemoveComplianceTagHoldApplied if needed to remove ComplianceTagHoldApplied from inactive mailbox. If the object is active, include the `-ProvideConsent` parameter Set-Mailbox <\DN or Exchange GUID> -RemoveComplianceTagHoldApplied -ProvideConsent
- Run the Set-MailUser <\DN or Exchange GUID> -RemoveComplianceTagHoldApplied if needed to remove ComplianceTagHoldApplied from inactive mailuser.

- Run the Set-Mailbox <\DN or Exchange GUID> -ExcludeFromAllHolds if needed to remove holds from inactive mailbox. Check the Set-Mailbox | ExcludeFromAllHolds to understand which holds this parameter can remove and which it cannot.
- Run the Set-MailUser <\DN or Exchange GUID> -ExcludeFromAllHolds if needed to remove holds from inactive mailuser. Check the Set-MailUser | ExcludeFromAllHolds to understand which holds this parameter can remove and which it cannot.
- Change the hold duration for an inactive mailbox
- Delete an inactive mailbox
- Delete items in the Recoverable Items folder of cloud-based mailboxes on hold

---

Last updated on 02/18/2026

Delete items in the Recoverable Items folder of cloud-based mailboxes on hold

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

The Recoverable Items folder for an Exchange Online mailbox exists to protect from accidental or malicious deletions. It's also used to store items that are retained and accessed by compliance features, such as holds and eDiscovery searches. However, in some situations organizations might have data that's been unintentionally retained in the Recoverable Items folder that they must delete. For example, a user might unknowingly send or forward an email message that contains sensitive information or information that may have serious business consequences. Even if the message is permanently deleted, it might be retained indefinitely because a legal hold has been placed on the mailbox. This scenario is known as *data spillage* because data has been unintentionally *spilled* into Office 365. In these situations, you can delete items in a user's Recoverable Items folder for an Exchange Online mailbox, even if that mailbox is placed on hold with one of the different hold features in Office 365. These types of holds include Litigation Holds, In-Place Holds, eDiscovery holds, and retention policies created in the Microsoft Purview portal.

> **Tip**
>
> You can also use the purge feature in Data Security Investigations to search for and permanently delete Exchange mailbox and Microsoft Teams items in your organization. Data Security Investigations provides a streamlined search and purge workflow that doesn't require you to manually remove holds before deleting items. For more information, see Purge in Data Security Investigations.

This article explains how admins can delete items from the Recoverable Items folder for cloud-based mailboxes that are on hold. This procedure involves disabling access to the mailbox and disabling single item recovery, disabling the Managed Folder Assistant from processing the mailbox, temporarily removing the hold, deleting items from the Recoverable Items folder, and then reverting the mailbox to its previous configuration. Here's the process:

Step 1: Collect information about the mailbox

Step 2: Prepare the mailbox

Step 3: Remove all holds from the mailbox

Step 4: Remove the delay hold from the mailbox

Step 5: Delete items in the Recoverable Items folder

Step 6: Revert the mailbox to its previous state

> ⊗ **Caution**
>
> The procedures outlined in this article will result in data being permanently deleted (purged) from an Exchange Online mailbox. That means messages that you delete from the Recoverable Items folder can't be recovered and won't be available for legal discovery or other compliance purposes. If you want to delete messages from a mailbox that's placed on hold as part of a Litigation Hold, In-Place Hold, eDiscovery hold, or retention policy created in the Microsoft Purview portal, check with your records management or legal departments before removing the hold. Your organization might have a policy that defines whether a mailbox on hold or a data spillage incident takes priority.

Before you delete items

- To create and run a Content Search, you have to be a member of the eDiscovery Manager role group or be assigned the Compliance Search management role. To delete messages, you have to be a member of the Organization Management role group or be assigned the Search And Purge management role. For information about adding users to a role group, see Assign eDiscovery permissions.
- If a mailbox is assigned to an organization-wide retention policy, you have to exclude the mailbox from the policy before you can delete items from the Recoverable Items folder. It may take up to 24 hours to synchronize the policy change, and remove the mailbox from

the policy. For more information, see "Organization-wide retention policies" in the Remove all holds from the mailbox section in this article.

- You can't perform this procedure for a mailbox that has been assigned retention settings with a retention policy that's locked by using Preservation Lock. That's because this lock prevents you from removing or excluding the mailbox from the policy and from disabling the Managed Folder Assistant on the mailbox. For more information about locking policies for retention,see Use Preservation Lock to restrict changes to retention policies and retention label policies.
- The procedure described in this article isn't supported for inactive mailboxes. That's because you can't reapply a hold (or retention policy) to an inactive mailbox after you remove it. When you remove a hold from an inactive mailbox, it's changed to a normal soft-deleted mailbox and will be permanently deleted from your organization after it's processed by the Managed Folder Assistant.
- If a mailbox isn't placed on hold (or doesn't have single item recovery enabled), you can delete the items from the Recoverable Items folder. For more information about how to do this, see Search for and delete email messages in your organization.

Step 1: Collect information about the mailbox

This first step is to collect selected properties from the target mailbox that will affect this procedure. Be sure to write down these settings or save them to a text file because you'll change some of these properties and then revert back to the original values in Step 6, after you delete items from the Recoverable Items folder. Here's a list of the mailbox properties you need to collect.

- *SingleItemRecoveryEnabled* and *RetainDeletedItemsFor*. If necessary, you'll disable single recovery and increase the deleted items retention period in Step 3.
- *LitigationHoldEnabled* and *InPlaceHolds*. You need to identify all the holds placed on the mailbox so that you can temporarily remove them in Step 3. See the More information section for tips about how to identify the type hold that might be placed on a mailbox.

Additionally, you need to get the mailbox client access settings so you can temporarily disable them so the owner (or other users) can't access the mailbox during this procedure. Finally, you can get the current size and number of items in the Recoverable Items folder. After you delete items in the Recoverable Items folder in Step 5, you'll use this information to verify that items were removed.

1. Connect to Exchange Online PowerShell. Be sure to use a user name and password for an administrator account that's been assigned the appropriate management roles in Exchange

Online.

2. Run the following command to get information about single item recovery and the deleted item retention period.

```
PowerShell
Get-Mailbox <username> | FL SingleItemRecoveryEnabled,RetainDeletedItemsFor
```

If single item recovery is enabled, you'll have to disable it in Step 2. If the deleted item retention period isn't set for 30 days (the maximum value in Exchange Online), then you can increase it in Step 2.

3. Run the following command to get the mailbox access settings for the mailbox.

```
PowerShell
Get-CASMailbox <username> | FL
EwsEnabled,ActiveSyncEnabled,MAPIEnabled,OWAEnabled,ImapEnabled,PopEnabled
```

You'll disable all of these access methods in Step 2.

4. Run the following command to get information about the holds and retention policies applied to the mailbox.

```
PowerShell
Get-Mailbox <username> | FL LitigationHoldEnabled,InPlaceHolds
```

> **Tip**
>
> If there are too many values in the *InPlaceHolds* property and not all of them are displayed, you can run the `Get-Mailbox <username> | Select-Object -ExpandProperty InPlaceHolds` command to display each value on a separate line.

5. Run the following command to get information about any organization-wide retention policies.

```
PowerShell
Get-OrganizationConfig | FL InPlaceHolds
```

If your organization has any organization-wide retention policies, you'll have to exclude the mailbox from these policies in Step 3. It may take up to 24 hours to replicate the change.

> **Tip**
>
> If there are too many values in the *InPlaceHolds* property and not all of them are displayed, you can run the `Get-OrganizationConfig | Select-Object -ExpandProperty InPlaceHolds` command to display each value on a separate line.

6. Run the following command to determine if a delay hold is applied to the mailbox.

```
PowerShell
Get-Mailbox <username> | FL DelayHoldApplied,DelayReleaseHoldApplied
```

If the value of the *DelayHoldApplied* or *DelayReleaseHoldApplied* property is set to **True**, a delay hold is applied to the mailbox and must be removed. For more information about delay holds, see Step 4: Remove the delay hold from the mailbox.

If the value of either properties is set to **False**, a delay hold isn't applied to the mailbox, and you can skip Step 4.

7. Run the following command to get the current size and total number of items in folders and subfolders in the Recoverable Items folder in the user's primary mailbox.

```
PowerShell
Get-MailboxFolderStatistics <username> -FolderScope RecoverableItems | FL
Name,FolderAndSubfolderSize,ItemsInFolderAndSubfolders
```

If the user's archive mailbox is enabled, run the following command to get the size and total number of items in folders and subfolders in the Recoverable Items folder in their archive mailbox.

```
PowerShell
Get-MailboxFolderStatistics <username> -FolderScope RecoverableItems -Archive |
FL Name,FolderAndSubfolderSize,ItemsInFolderAndSubfolders
```

When you delete items in Step 5, you can choose to delete or not delete items in the Recoverable Items folder in the user's primary archive mailbox. If auto-expanding archiving

is enabled for the mailbox, items in an auxiliary archive mailbox won't be deleted.

Step 2: Prepare the mailbox

After collecting and saving information about the mailbox, the next step is to prepare the mailbox by performing the following tasks:

- **Disable client access to mailbox** so that the mailbox owner can't access their mailbox and make any changes to the mailbox data during this procedure.
- **Increase the deleted item retention period** to 30 days (the maximum value in Exchange Online) so that items aren't purged from the Recoverable Items folder before you can delete them in Step 5.
- **Disable single Item recovery** so that items won't be retained (for the duration of the deleted item retention period) after you delete them from the Recoverable Items folder in Step 5.
- **Disable the Managed Folder Assistant** so that it doesn't process the mailbox and retain the items that you delete in Step 5.

Perform the following steps in Exchange Online PowerShell.

1. Run the following command to disable all client access to the mailbox. The command syntax assumes that all client access methods were enabled on the mailbox.

   PowerShell

   ```
   Set-CASMailbox <username> -EwsEnabled $false -ActiveSyncEnabled $false -
   MAPIEnabled $false -OWAEnabled $false -ImapEnabled $false -PopEnabled $false
   ```

   > **Note**
   >
   > It might take up to 60 minutes to disable all client access methods to the mailbox. Note that disabling these access methods won't disconnect the mailbox owner if they are currently signed in. If the owner isn't signed in, they won't be able to access their mailbox after these access methods are disabled.

2. Run the following command to increase the deleted item retention period the maximum of 30 days. This assumes that the current setting is less than 30 days.

   PowerShell

```
Set-Mailbox <username> -RetainDeletedItemsFor 30
```

3. Run the following command to disable single item recovery.

```
PowerShell
Set-Mailbox <username> -SingleItemRecoveryEnabled $false
```

> ⓘ Note
>
> It might take up to 240 minutes to disable single item recovery. Don't delete items in the Recoverable Items folder until this period has elapsed.

4. Run the following command to prevent the Managed Folder Assistant from processing the mailbox. As previously explained, you can disable the Managed Folder Assistant only if a retention policy with a Preservation Lock isn't applied to the mailbox.

```
PowerShell
Set-Mailbox <username> -ElcProcessingDisabled $true
```

Step 3: Remove all holds from the mailbox

The last step before you can delete items from the Recoverable Items folder is to remove all holds (that you identified in Step 1) placed on the mailbox. All holds must be removed so that items won't be retained after you delete them from the Recoverable Items folder. The following sections contain information about removing different types of holds on a mailbox. See the More information section for tips about how to identify the type hold that might be placed on a mailbox. For more information, see How to identify the type of hold placed on an Exchange Online mailbox.

> ⊗ Caution
>
> As previously stated, check with your records management or legal departments before removing a hold from a mailbox.

Litigation Hold

Run the following command in Exchange Online PowerShell to remove a Litigation Hold from the mailbox.

```
PowerShell
Set-Mailbox <username> -LitigationHoldEnabled $false
```

> **Note**
>
> Similar to disabling single item recovery, it might take up to 240 minutes to remove the Litigation Hold. Don't delete items from the Recoverable Items folder until this period has elapsed.

In-Place Hold

Run the following command in Exchange Online PowerShell to identify the In-Place Hold that's placed on the mailbox. Use the GUID for the In-Place Hold that you identified in Step 1.

```
PowerShell
Get-ComplianceSearch -Identity <hold GUID> | FL Name
```

After you identify the In-Place Hold, you can use the Exchange admin center (EAC) or Exchange Online PowerShell to remove the mailbox from the hold. For more information, see Create or remove an In-Place Hold.

Retention policies applied to specific mailboxes

Run the following command in Security & Compliance PowerShell to identify the retention policy that is applied to the mailbox. This command will also return any Teams conversation retention policies applied to a mailbox. Use the GUID (not including the `mbx` or `skp` prefix) for the retention policy that you identified in Step 1.

```
PowerShell
Get-RetentionCompliancePolicy <retention policy GUID without prefix> | FL Name
```

After you identify the retention policy, go to the **Data lifecycle management** > **Microsoft 365** > **Retention** page in the Microsoft Purview portal, edit the retention policy that you identified in the

previous step, and remove the mailbox from the list of recipients that are included in the retention policy.

Organization-wide retention policies

Organization-wide, Exchange-wide, and Teams-wide retention policies are applied to every mailbox in the organization. They're applied at the organization level (not the mailbox level) and are returned when you run the **Get-OrganizationConfig** cmdlet in Step 1. Run the following command in Security & Compliance PowerShell to identify the organization-wide retention policies. Use the GUID (not including the `mbx` prefix) for the organization-wide retention policies that you identified in Step 1.

PowerShell

```
Get-RetentionCompliancePolicy <retention policy GUID without prefix> | FL Name
```

After you identify the organization-wide retention policies, go to the **Data lifecycle management** > **Microsoft 365** > **Retention** page in the Microsoft Purview portal, edit each organization-wide retention policy that you identified in the previous step, and add the mailbox to the list of excluded recipients. Doing this removes the user's mailbox from the retention policy.

Organization-wide Viva Engage and private channel retention policies

Organization-wide Yammer/Private Channel retention policies apply to mailboxes that match the workload type of the hold. You can check these policies in your organization with the following cmdlet:

PowerShell

```
Get-AppRetentionCompliancePolicy | fl name,guid,mode,enabled
```

A mailbox needs to be excluded from all organization-wide retention policies, regardless of type (Microsoft Teams, Viva Engage, Exchange) or scope (keeping data, deleting data).

If the mailbox isn't soft deleted or inactive, you can exclude the mailbox from the retention policy in PowerShell by adding an Exchange location exception. For more information, see the AddExchangeLocationException parameter. You can't exclude inactive or soft deleted mailboxes or mail users from these retention policies. To exclude inactive or soft deleted mailboxes or mail users from these retention policies, contact the Microsoft Support team.

> **Important**
>
> After you exclude a mailbox from an organization-wide retention policy, it can take up to 24 hours to synchronize this change and remove the mailbox from the policy.

Retention labels

Whenever a user applies a label that's configured to retain content or retain and then delete content to any folder or item in their mailbox, the *ComplianceTagHoldApplied* mailbox property is set to **True**. When this happens, the mailbox is considered to be on hold, as if it was placed on Litigation Hold or assigned to a retention policy.

To view the value of the *ComplianceTagHoldApplied* property, run the following command in Exchange Online PowerShell:

```
PowerShell
Get-Mailbox <username> |FL ComplianceTagHoldApplied
```

After you've identified that a mailbox is on hold because a retention label is applied to a folder or item, you can use the Content search tool in the Microsoft Purview portal to search for labeled items by using the **Retention label** condition. For more information, see:

- The "Using Content Search to find all content with a specific retention label" section in Learn about retention policies and retention labels
- The "Search conditions" section in Keyword queries and search conditions for Content Search.

For more information about labels, see Learn about retention policies and retention labels.

eDiscovery holds

Run the following commands in Security & Compliance PowerShell to identify the hold associated with an eDiscovery case (called *eDiscovery holds*) that's applied to the mailbox. Use the GUID (not including the `UniH` prefix) for the eDiscovery hold that you identified in Step 1. The second command displays the name of the eDiscovery case the hold is associated with; the third command displays the name of the hold.

```
PowerShell
```

```
$CaseHold = Get-CaseHoldPolicy <hold GUID without prefix>
```

PowerShell

```
Get-ComplianceCase $CaseHold.CaseId | FL Name
```

PowerShell

```
$CaseHold.Name
```

After you've identified the name of the eDiscovery case and the hold, go to the **eDiscovery** > **eDiscovery** page in the Microsoft Purview portal, open the case, and remove the mailbox from the hold. For more information about identifying eDiscovery holds, see the "eDiscovery holds" section in How to identify the type of hold placed on an Exchange Online mailbox.

Step 4: Remove the delay hold from the mailbox

After any type of hold is removed from a mailbox, the value of the *DelayHoldApplied* or *DelayReleaseHoldApplied* mailbox property is set to **True**. This occurs the next time the Managed Folder Assistant processes the mailbox and detects that a hold has been removed. This is called a *delay hold* and means the actual removal of the hold is delayed for 30 days to prevent data from being permanently deleted from the mailbox. (The purpose of a delay hold is to give admins an opportunity to search for or recover mailbox items that will be purged after a hold is removed.) When a delay hold is placed on the mailbox, the mailbox is still considered to be on hold for an unlimited duration, as if the mailbox was on Litigation Hold. After 30 days, the delay hold expires, and Microsoft 365 will automatically attempt to remove the delay hold (by setting the *DelayHoldApplied* or *DelayReleaseHoldApplied* property to **False**) so that the hold is removed. For more information about a delay hold, see the "Managing mailboxes on delay hold" section in How to identify the type of hold placed on an Exchange Online mailbox.

If the value of the *DelayHoldApplied* or *DelayReleaseHoldApplied* property is set to **True**, run one of the following commands to remove the delay hold:

PowerShell

```
Set-Mailbox <username> -RemoveDelayHoldApplied
```

Or

```
```

PowerShell

```
Set-Mailbox <username> -RemoveDelayReleaseHoldApplied
```

You must be assigned the Legal Hold role in Exchange Online to use the *RemoveDelayHoldApplied* or *RemoveDelayReleaseHoldApplied* parameter.

Step 5: Delete items in the Recoverable Items folder

Now you're ready to actually delete items in the Recoverable Items folder by using the New-ComplianceSearch and New-ComplianceSearchAction cmdlets in Security & Compliance PowerShell.

> **Note**
>
> If the mailbox quota is met and the user mailbox is declining emails, you may receive a 554 5.2.0 error when deleting recoverable items. For more information, see "554 5.2.0 STOREDRV.Deliver.Exception" when sending emails in Exchange Online.

Here's an overview of the process to search for and delete items in a user's Recoverable Items folder:

1. Copy the folder IDs for all subfolders in the Recoverable Items folder. Alternatively, you can redirect the output of the script to a text file.

   Here's a list and description of the subfolders in the Recoverable Items folder that you can search and delete items from:

   - **Deletions**: Contains soft-deleted items whose deleted item retention period hasn't expired. Users can recover soft-deleted items from this subfolder using the Recover Deleted Items tool in Outlook.
   - **DiscoveryHolds**: Contains hard-deleted items that have been preserved by an eDiscovery hold or a retention policy. This subfolder isn't visible to end users.
   - **SubstrateHolds**: Contains hard-deleted items from Teams and other cloud-based apps that have been preserved by a retention policy or other type of hold. This subfolder isn't visible to end users.

2. Use the **New-ComplianceSearch** cmdlet (in Security & Compliance PowerShell) or use the Content search tool in the Microsoft Purview portal to create a content search that returns items from the target user's Recoverable Items folder. You can do this by including the

FolderId in the search query for all subfolders that you want to search. For example, the following query returns all messages in the Deletions and eDiscoveryHolds subfolders:

```text
folderid:<folder ID of Deletions subfolder> OR folderid:<folder ID of
DiscoveryHolds subfolder>
```

> ⓘ Note
>
> If you use the **New-ComplianceSearch** cmdlet to search the Recoverable Items folder, be sure to use **Start-ComplianceSearch** cmdlet to run the search.

3. After you've created a content search and validated that it returns the items that you want to delete, use the `New-ComplianceSearchAction -Purge -PurgeType HardDelete` command (in Security & Compliance PowerShell) to permanently delete the items returned by the content search that you created in the previous step. For example, you can run a command similar to the following command:

```PowerShell
New-ComplianceSearchAction -SearchName "RecoverableItems" -Purge -PurgeType
HardDelete
```

4. A maximum of 10 items per mailbox are deleted when you run the previous command. That means you may have to run the `New-ComplianceSearchAction -Purge` command multiple times to delete all the items that you want to delete in the Recoverable Items folder. To delete additional items, you first have to remove the previous compliance search purge action. You do this by running the `Remove-ComplianceSearchAction` cmdlet. For example, to delete the purge action that was run in the previous step, run the following command:

```PowerShell
Remove-ComplianceSearchAction "RecoverableItems_Purge"
```

After you do this, you can create a new compliance search purge action to delete more items. You'll have to delete each purge action before creating a new one.

To get a list of the compliance search actions, you can run the `Get-ComplianceSearchAction` cmdlet. Purge actions are identified by `_Purge` appended to the search name.

Verify that items were deleted

To verify that you've successfully deleted items from the Recoverable Items folder of a mailbox, use **Get-MailboxFolderStatistics** cmdlet in Exchange Online PowerShell to check the size and number of items in Recoverable Items folder. You can compare these statistics with the ones you collected in Step 1.

Run the following command in to get the current size and total number of items in folders and subfolders in the Recoverable Items folder in the user's primary mailbox.

PowerShell

```
Get-MailboxFolderStatistics <username> -FolderScope RecoverableItems | FL
Name,FolderAndSubfolderSize,ItemsInFolderAndSubfolders
```

Run the following command to get the size and total number of items in folders and subfolders in the Recoverable Items folder in the user's archive mailbox.

PowerShell

```
Get-MailboxFolderStatistics <username> -FolderScope RecoverableItems -Archive | FL
Name,FolderAndSubfolderSize,ItemsInFolderAndSubfolders
```

Step 6: Revert the mailbox to its previous state

The final step is to revert the mailbox back to its previous configuration. This means resetting the properties that you changed in Step 2 and reapplying the holds that you removed in Step 3. This includes:

- Changing the deleted item retention period back to its previous value. Alternatively, you can just leave this set to 30 days, the maximum value in Exchange Online.
- Re-enabling single Item recovery.
- Re-enabling the client access methods so that the owner can access their mailbox.
- Reapplying the holds and retention policies that you removed.
- Re-enabling the Managed Folder Assistant to process the mailbox.

> ⓘ **Important**

> We recommend that you wait 24 hours after re-applying a hold or retention policy (and verifying that it's in place) before you re-enable the Managed Folder Assistant to process the mailbox.

Perform the following steps (in the specified sequence) in Exchange Online PowerShell.

1. Run the following command to change the deleted item retention period back to its original value. This assumes that the previous setting is less than 30 days; for example, 14 days.

   ```
   PowerShell
   Set-Mailbox <username> -RetainDeletedItemsFor 14
   ```

2. Run the following command to re-enable single item recovery.

   ```
   PowerShell
   Set-Mailbox <username> -SingleItemRecoveryEnabled $true
   ```

3. Run the following command to re-enable all client access methods to the mailbox.

   ```
   PowerShell
   Set-CASMailbox <username> -EwsEnabled $true -ActiveSyncEnabled $true -MAPIEnabled $true -OWAEnabled $true -ImapEnabled $true -PopEnabled $true
   ```

4. Reapply the holds that you removed in Step 3. Depending on the type of hold, use one of the following procedures.

   Litigation Hold

   Run the following command to re-enable a Litigation Hold for the mailbox.

   ```
   PowerShell
   Set-Mailbox <username> -LitigationHoldEnabled $true
   ```

   In-Place Hold

   Use the EAC (or Exchange Online PowerShell) to add the mailbox back to the In-Place Hold.

   Retention policies applied to specific mailboxes

Use the Microsoft Purview portal to add the mailbox back to the retention policy. Go to the **Data lifecycle management** > **Microsoft 365** > **Retention** page in the Microsoft Purview portal, edit the retention policy, and add the mailbox back to the list of recipients that the retention policy is applied to.

Organization-wide retention policies

If you removed an organization-wide or Exchange-wide retention policy by excluding it from the policy, then use the Microsoft Purview portal to remove the mailbox from the list of excluded users. Go to the **Data lifecycle management** > **Microsoft 365** > **Retention** page in the Microsoft Purview portal, edit the organization-wide retention policy, and remove the mailbox from the list of excluded recipients. Doing this reapplies the retention policy to the user's mailbox.

eDiscovery case holds

Use the Microsoft Purview portal to add the mailbox back the hold that's associated with an eDiscovery case. Go to the **eDiscovery (Standard)** page, open the case, and add the mailbox back to the hold.

5. Run the following command to allow the Managed Folder Assistant to process the mailbox again. As previously stated, we recommend that you wait 24 hours after reapplying a hold or retention policy (and verifying that it's in place) before you re-enable the Managed Folder Assistant.

```
PowerShell
Set-Mailbox <username> -ElcProcessingDisabled $false
```

6. To verify that the mailbox has been reverted back to its previous configuration, you can run the following commands and then compare the settings to the ones that you collected in Step 1.

```
PowerShell
Get-Mailbox <username> | FL
ElcProcessingDisabled,InPlaceHolds,LitigationHoldEnabled,RetainDeletedItemsFor,Si
ngleItemRecoveryEnabled
```

```
PowerShell
```

```
Get-CASMailbox <username> | FL
EwsEnabled,ActiveSyncEnabled,MAPIEnabled,OWAEnabled,ImapEnabled,PopEnabled
```

More information

Here's a table that describes how to identify different types of holds based on the values in the *InPlaceHolds* property when you run the **Get-Mailbox** or **Get-OrganizationConfig** cmdlets. For more detailed information, see How to identify the type of hold placed on an Exchange Online mailbox.

As previously explained, you have to remove all holds and retention policies from a mailbox before you can successfully delete items in the Recoverable Items folder.

Expand table

Hold type	Example value	How to identify the hold
Litigation Hold	`True`	The *LitigationHoldEnabled* property is set to `True`.
In-Place Hold	`c0ba3ce811b6432a8751430937152491`	The *InPlaceHolds* property contains the GUID of the In-Place Hold that's placed on the mailbox. You can tell this is an In-Place Hold because the GUID doesn't start with a prefix. You can use the Get-ComplianceSearch cmdlet in Exchange Online PowerShell to get information about the In-Place Hold on the mailbox.
Retention policies in the Microsoft Purview portal applied to specific mailboxes	`mbxcdbbb86ce60342489bff371876e7f224` or `skp127d7cf1076947929bf136b7a2a8c36f`	When you run the **Get-Mailbox** cmdlet, the *InPlaceHolds* property also contains GUIDs of retention policies applied to the mailbox. You can identify retention policies because the GUID starts with the `mbx` prefix. If the GUID of the retention policy starts with the `skp` prefix, that indicates that the retention policy is applied to Skype for Business conversations. To identity the retention policy that's applied to the mailbox, run the following command in Security & Compliance PowerShell: `Get-RetentionCompliancePolicy <retention policy GUID without prefix> | FL Name` Be sure to remove the `mbx` or `skp` prefix when you run this command.

Hold type	Example value	How to identify the hold
Organization-wide retention policies in the Microsoft Purview portal	No value or `-mbxe9b52bf7ab3b46a286308ecb29624696` (indicates that the mailbox is excluded from an organization-wide policy)	Even if the *InPlaceHolds* property is empty when you run the **Get-Mailbox** cmdlet, there still might be one or more organization-wide retention policies applied to the mailbox. To verify this, you can run the `Get-OrganizationConfig | FL InPlaceHolds` command in Exchange Online PowerShell to get a list of the GUIDs for organization-wide retention policies. The GUID for organization-wide retention policies applied to Exchange mailboxes starts with the `mbx` prefix; for example, `mbxa3056bb15562480fadb46ce523ff7b02`. To identity the organization-wide retention policy that's applied to the mailbox, run the following command in Security & Compliance PowerShell: `Get-RetentionCompliancePolicy <retention policy GUID without prefix> | FL Name` If a mailbox is excluded from an organization-wide retention policy, the GUID for the retention policy is displayed in the *InPlaceHolds* property of the user's mailbox when you run the **Get-Mailbox** cmdlet; it's identified by the prefix `-mbx`; for example, `-mbxe9b52bf7ab3b46a286308ecb29624696`
eDiscovery case hold in the Microsoft Purview portal	`UniH7d895d48-7e23-4a8d-8346-533c3beac15d`	The *InPlaceHolds* property also contains the GUID of any hold associated with an eDiscovery case in the Microsoft Purview portal that might be placed on the mailbox. You can tell this is an eDiscovery case hold because the GUID starts with the `UniH` prefix. You can use the `Get-CaseHoldPolicy` cmdlet in Security & Compliance PowerShell to get information about the eDiscovery case that the hold on the mailbox is associated with. For example, you can run the command `Get-CaseHoldPolicy <hold GUID without prefix> | FL Name` to display the name of the case hold that's on the mailbox. Be sure to remove the `UniH` prefix when you run this command. To identity the eDiscovery case that the hold on the mailbox is associated with, run the following commands:

Hold type	Example value	How to identify the hold

		`$CaseHold = Get-CaseHoldPolicy <hold GUID without prefix>`
		
		`Get-ComplianceCase $CaseHold.CaseId | FL Name`

Last updated on 10/31/2024

Increase the Recoverable Items quota for mailboxes on hold

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

The default Exchange retention policy—named *Default MRM Policy*—that is automatically applied to new mailboxes in Exchange Online contains a retention tag named Recoverable Items 14 days move to archive. This retention tag moves items from the Recoverable Items folder in the user's primary mailbox to the Recoverable Items folder in the user's archive mailbox after the 14-day retention period expires for an item. Emails in Deletions folder will retain based on **RetainDeletedItemsFor** parameter and move to other folders in recoverable deleted items and then to archive mailbox. For this to happen, the user's archive mailbox must be enabled. If the archive mailbox isn't enabled, no action is taken, which means that items in the Recoverable Items folder for a mailbox on hold aren't moved to the archive mailbox after the 14-day retention period expires. Because nothing is deleted from a mailbox on hold, it's possible that the storage quota for the Recoverable Items folder might be exceeded, especially if the user's archive mailbox isn't enabled.

To help reduce the chance of exceeding this limit, the storage quota for the Recoverable Items folder is automatically increased from 30 GB to 100 GB when a hold is placed on a mailbox in Exchange Online. If the archive mailbox is enabled, the storage quota for the Recoverable Items folder in the archive mailbox is also increased from 30 GB to 100 GB. If archive is enabled on a mailbox that is on hold with auto-expanding archive disabled, the storage quota of the Recoverable Items folder quota is increased from 100 GB to 105 GB. If the auto-expanding archiving feature in Exchange Online is enabled, the total storage quota for the user's archive mailbox, including the Recoverable Items folder, is 1.5 TB.

The following table summarizes the storage quota for the Recoverable Items folder.

Expand table

Location of Recoverable Items folder	Mailboxes not on hold	Mailboxes on hold
Primary mailbox	30 GB	100 GB
Archive mailbox, including Recoverable Items folder *	1.5 TB	1.5 TB

> **Note**
>
> * The initial storage quota for the archive mailbox is 100 GB for users with an Exchange Online (Plan 2) license. However, when auto-expanding archiving is turned on for mailboxes on hold, the storage quota for both the archive mailbox and the Recoverable Items folder is increased to 110 GB. Additional archive storage space (which includes the Recoverable Items folder) up to 1.5 TB will be provisioned when necessary. For more information about auto-expanding archiving, see Learn about auto-expanding archiving.

> **Note**
>
> When archiving is enabled for a specific user whose mailbox is on hold or assigned to a retention policy, the storage quota for the user's primary mailbox's Recoverable Items folder is raised from 100 GB to 105 GB and Recoverable items warning quota is raised from 90 GB to 95 GB. This additional space supports an efficient archiving experience for users, even if auto-expanding archive isn't enabled.

When the storage quota for the Recoverable Items folder in the primary mailbox of a mailbox on hold is close to reaching its limit, you can do the following things:

- **Enable the archive mailbox and turn on auto-expanding archiving.** You can enable an additional storage capacity for the Recoverable Items folder simply by enabling the archive mailbox and then turning on the auto-expanding archiving feature in Exchange Online. This results in 110 GB for the Recoverable Items folder in the primary mailbox and up to 1.5 TB of combined storage capacity for the archive and Recoverable Items folder. For more information, see Enable archive mailboxes and Enable auto-expanding archiving.

  > **Note**
  >
  > After you enable the archive for a mailbox that's close to exceeding the storage quota for the Recoverable Items folder, you might want to run the Managed Folder Assistant

to manually trigger the assistant to process the mailbox so that expired items are moved to the Recoverable Items folder in the archive mailbox. See Step 4 for instructions. Note that other items in the user's mailbox might be moved to the new archive mailbox. Consider telling the user that this may happen after you enable the archive mailbox.

- **Create a custom Exchange retention policy for mailboxes on hold.** In addition to enabling the archive mailbox and auto-expanding archiving for mailboxes on Litigation Hold or In-Place Hold, you might also want to create a custom Exchange retention policy for mailboxes on hold. This lets you apply a retention policy to mailboxes on hold that's different from the Default MRM Policy that's applied to mailboxes that aren't on hold, and lets you apply retention tags that are designed for mailboxes on hold. This includes creating a new retention tag for the Recoverable Items folder.

The remainder of this topic describes the step-by-step procedures to create a custom Exchange retention policy for mailboxes on hold.

Step 1: Create a custom retention tag for the Recoverable Items folder

Step 2: Create a new Exchange retention policy for mailboxes on hold

Step 3: Apply the new Exchange retention policy to mailboxes on hold

(Optional) Step 4: Run the Managed Folder Assistant to apply the new retention settings

Step 1: Create a custom retention tag for the Recoverable Items folder

The first step is to create a custom retention tag (called a retention policy tag or RPT) for the Recoverable Items folder. As previously explained, this RPT moves items from the Recoverable Items folder in the user's primary mailbox to the Recoverable Items folder in the user's archive mailbox. You have to use PowerShell to create an RPT for the Recoverable Items folder. You can't use the Exchange admin center (EAC).

1. Connect to Exchange Online PowerShell

2. Run the following command to create a new RPT for the Recoverable Items folder:

```
PowerShell
```

```
New-RetentionPolicyTag -Name <Name of RPT> -Type RecoverableItems -
AgeLimitForRetention <Number of days> -RetentionAction MoveToArchive
```

For example, the following command creates an RPT for the Recoverable Items folder named "Recoverable Items 30 days for mailboxes on hold", with a retention period of 30 days. This means that after an item has been in the Recoverable Items folder for 30 days, it will be moved to the Recoverable Items folder in the user's archive mailbox.

PowerShell

```
New-RetentionPolicyTag -Name "Recoverable Items 30 days for mailboxes on hold" -
Type RecoverableItems -AgeLimitForRetention 30 -RetentionAction MoveToArchive
```

> **Tip**
>
> We recommend that the retention period (defined by the *AgeLimitForRetention* parameter) for the Recoverable Items RPT is the same as the deleted item retention period for the mailboxes that the RPT will be applied to. This allows a user the entire deleted item retention period to recover deleted items before they are moved to the archive mailbox. In the previous example, the retention period was set to 30 days based on the assumption that the deleted item retention period for mailboxes is also 30 days. An Exchange Online mailbox is configured to retain deleted items for 14 days, by default. But you can change this setting to a maximum of 30 days. For more information, see **Change the deleted item retention period for a mailbox in Exchange Online**.

Step 2: Create a new Exchange retention policy for mailboxes on hold

The next step is to create a new retention policy and add retention tags to it, including the Recoverable Items RPT that you created in Step 1. This new policy will be applied to mailboxes on hold in the next step.

Before you create the new retention policy, determine the additional retention tags that you want to add. For a list of the retention tags that are added to the Default MRM Policy and for information about creating new retention tags, see the following:

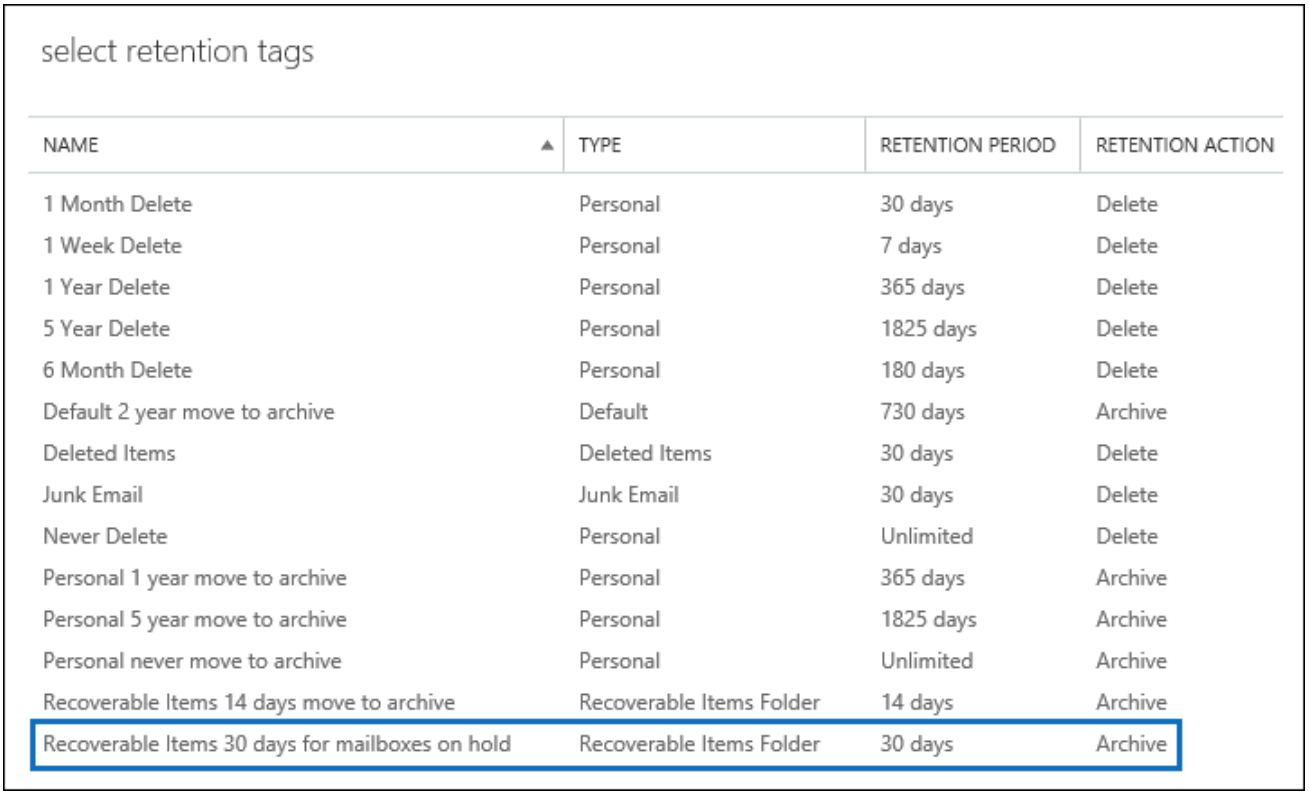
- Default Retention Policy in Exchange Online

- Default folders that support Retention Policy Tags
- The "Create a retention tag" section in the Create a Retention Policy topic.

You can use the EAC or Exchange Online PowerShell to create a retention policy.

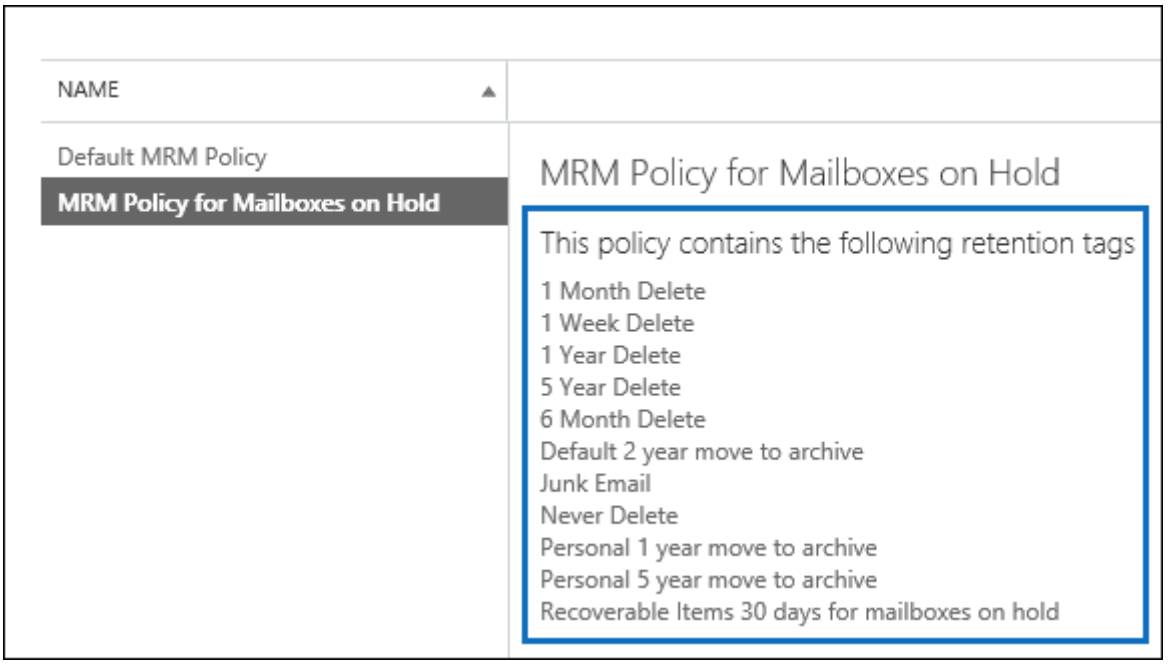
Use the EAC to create a retention policy

1. In the EAC, go to **Compliance management** > **Retention policies**, and then select **Add +**.
2. On the **New retention policy** page, under **Name**, type a name that describes the purpose of the retention policy; for example, **MRM Policy for Mailboxes on Hold**.
3. Under **Retention tags**, select **Add +**.
4. In the list of retention tags, select the Recoverable Items RPT that you created in Step 1, and then select **Add**.



5. Select additional retention tags to add to the retention policy. For example, you might want to add the same tags that are included in the Default MRM Policy.
6. When you're finished adding retention tags, select **OK**.
7. Select **Save** to create the new retention policy.

   Notice that the retention tags linked to the retention policy are displayed in the details pane.



Use Exchange Online PowerShell to create a retention policy

Run the following command to create new retention policy for mailboxes on hold.

PowerShell

```
New-RetentionPolicy <Name of retention policy>  -RetentionPolicyTagLinks <list of retention tags>
```

For example, the following command creates the retention policy and linked retention tags that are displayed in the previous illustration.

PowerShell

```
New-RetentionPolicy "MRM Policy for Mailboxes on Hold"  -RetentionPolicyTagLinks "Recoverable Items 30 days for mailboxes on hold","1 Month Delete","1 Week Delete","1 Year Delete","5 Year Delete","6 Month Delete","Default 2 year move to archive","Junk Email","Never Delete","Personal 1 year move to archive","Personal 5 year move to archive"
```

Step 3: Apply the new Exchange retention policy to mailboxes on hold

The last step is to apply the new retention policy that you created in Step 2 to mailboxes on hold in your organization. You can use the EAC or Exchange Online PowerShell to apply the retention policy to a single mailbox or to multiple mailboxes.

Use the EAC to apply the new retention policy

1. Go to **Recipients** > **Mailboxes**.
2. In the list view, select the mailbox you want to apply the retention policy to, and then select **Edit** ✎.
3. On the **User Mailbox** page, select **Mailbox features**.
4. Under **Retention policy**, select the retention policy that you created in Step 2, and then select **Save**.

You can also use the EAC to apply the retention policy to multiple mailboxes.

1. Go to **Recipients** > **Mailboxes**.
2. In the list view, use the Shift or Ctrl keys to select multiple mailboxes.
3. In the details pane, select **More options**.
4. Under **Retention Policy**, select **Update**.
5. On the **Bulk assign retention policy** page, select the retention policy that you created in Step 2, and then select **Save**.

Use Exchange Online PowerShell to apply the new retention policy

You can use Exchange Online PowerShell to apply a new retention policy to a single mailbox. But the real power of PowerShell is that you can use it to quickly identify all the mailboxes in your organization that are on either Litigation Hold or In-Place Hold, and then apply the new retention policy to all mailboxes on hold in a single command. Here are some examples of using Exchange PowerShell to apply a retention policy to one or more mailboxes. All of the examples apply the retention policy that was created in Step 2.

This example applies the new retention policy to Pilar Pinilla's mailbox.

PowerShell

```
Set-Mailbox "Pilar Pinilla" -RetentionPolicy "MRM Policy for Mailboxes on Hold"
```

This example applies the new retention policy to all mailboxes in the organization that are on Litigation Hold.

PowerShell

```
$LitigationHolds = Get-Mailbox -ResultSize unlimited | Where-Object
{$_.LitigationHoldEnabled -eq 'True'}
```

PowerShell

```
$LitigationHolds.DistinguishedName | Set-Mailbox -RetentionPolicy "MRM Policy for
Mailboxes on Hold"
```

This example applies the new retention policy to all mailboxes in the organization that are on In-Place Hold.

PowerShell

```
$InPlaceHolds = Get-Mailbox -ResultSize unlimited | Where-Object {$_.InPlaceHolds -ne
$null}
```

PowerShell

```
$InPlaceHolds.DistinguishedName | Set-Mailbox -RetentionPolicy "MRM Policy for
Mailboxes on Hold"
```

You can use the **Get-Mailbox** cmdlet to verify that the new retention policy was applied.

Here are some examples to verify that the commands in the previous examples applied the "MRM Policy for Mailboxes on Hold" retention policy to mailboxes on Litigation Hold and mailboxes on In-Place Hold.

PowerShell

```
Get-Mailbox "Pilar Pinilla" | Select RetentionPolicy
```

PowerShell

```
Get-Mailbox -ResultSize unlimited | Where-Object {$_.LitigationHoldEnabled -eq
'True'} | FT DisplayName,RetentionPolicy -Auto
```

PowerShell

```
Get-Mailbox -ResultSize unlimited | Where-Object {$_.InPlaceHolds -ne $null} | FT
DisplayName,RetentionPolicy -Auto
```

(Optional) Step 4: Run the Managed Folder Assistant to apply the new retention settings

After you apply the new Exchange retention policy to mailboxes on hold, it can take up to 7 days in Exchange Online for the Managed Folder Assistant to process these mailboxes using the settings in the new retention policy. Instead of waiting for the Managed Folder Assistant to run, you can use the **Start-ManagedFolderAssistant** cmdlet to manually trigger the assistant to process the mailboxes that you applied the new retention policy to.

Run the following command to start the Managed Folder Assistant for Pilar Pinilla's mailbox.

PowerShell

```
Start-ManagedFolderAssistant "Pilar Pinilla"
```

Run the following commands to start the Managed Folder Assistant for all mailboxes on hold.

PowerShell

```
$MailboxesOnHold = Get-Mailbox -ResultSize unlimited | Where-Object {($_.InPlaceHolds
-ne $null) -or ($_.LitigationHoldEnabled -eq "True")}
```

PowerShell

```
$MailboxesOnHold.DistinguishedName | Start-ManagedFolderAssistant
```

More information

- After you enable a user's archive mailbox, consider telling the user that other items in their mailbox (not just items in the Recoverable Items folder) might be moved to the archive mailbox. This is because the Default MRM Policy that's assigned to Exchange Online mailboxes contains a retention tag (named Default 2 years move to archive) that moves items to the archive mailbox two years after the date the item was delivered to the mailbox or created by the user. For more information, see Default Retention Policy in Exchange Online
- After you enable a user's archive mailbox, you might also tell the user that they can recover deleted items in the Recoverable Items folder in their archive mailbox. They can do this in Outlook by selecting the **Deleted Items** folder in the archive mailbox, and then selecting **Recover Deleted Items from Server** on the **Home** tab. For more information about recovering deleted items, see Recover deleted items in Outlook for Windows.

Last updated on 10/31/2024

Place a mailbox on Litigation Hold

Article • 04/30/2025

APPLIES TO: 2016 2019 Subscription Edition

Place a mailbox on Litigation Hold to preserve all mailbox content, including deleted items and original versions of modified items. When you place a mailbox on Litigation Hold, the user's archive mailbox (if it's enabled) is also placed on hold. Deleted and modified items are preserved for a specified period or until you remove the mailbox from Litigation Hold. All such mailbox items are returned in an In-Place eDiscovery in Exchange Server search.

Before you begin

- You need to be assigned permissions before you can perform this procedure or procedures. To see what permissions you need, see the "In-Place Hold" entry in the Messaging policy and compliance permissions in Exchange Server topic.
- The Litigation Hold setting may take up to 60 minutes to take effect.
- Litigation Hold preserves items in the Recoverable Items folder in the user's mailbox. The default size for this folder is 30 GB. Depending on number and size of items deleted or modified, the size of the Recoverable Items folder of the mailbox may increase quickly. The Recoverable Items folder is configured with a high quota by default. We recommend that you monitor mailboxes that are placed on Litigation Hold on a weekly basis to ensure they don't reach the limits of the Recoverable Items quotas.
- As previously explained, when you place a user's mailbox on Litigation Hold, the user's archive mailbox is also placed on hold.
- Litigation Hold preserves deleted items and also preserves original versions of modified items until the hold is removed. You can optionally specify a hold duration, which preserves a mailbox item for the specified duration period. If you specify a hold duration period, it's calculated from the date a message is received or a mailbox item is created.
- See the More information section for a description of the Litigation Hold workflow process.

Use the EAC to place a mailbox on Litigation Hold

1. Go to **Recipients** > **Mailboxes**.

2. In the list of user mailboxes, click the mailbox that you want to place on Litigation Hold, and then click **Edit** .

3. On the mailbox properties page, click **Mailbox features**.

4. Under **Litigation hold: Disabled**, click **Enable** to place the mailbox on Litigation Hold.

5. On the **Litigation Hold** page, enter the following optional information:

    - **Litigation hold duration (days)**: Use this box to specify how long mailbox items are held when the mailbox is placed on Litigation Hold. The duration is calculated from the date a mailbox item is received or created. If you leave this box blank, items are held indefinitely or until the hold is removed. Use days to specify the duration.

    - **Note***: Use this box to inform the user their mailbox is on Litigation Hold. The note will appear on the **File** tab in Outlook 2010 or later.

    - **URL***: Use this box to direct the user to a website for more information about Litigation Hold. This URL appears on the **File** tab Outlook 2010 or later.

    *If you leave the **Note** and **URL** values blank, the user isn't notified that you placed a litigation hold on their mailbox.

6. Click **Save** on the **Litigation Hold** page, and then click **Save** on the mailbox properties page.

Use the Exchange Management Shell to place a mailbox on Litigation Hold indefinitely

This example places the mailbox bsuneja@contoso.com on Litigation Hold. Items in the mailbox are held indefinitely or until the hold is removed.

PowerShell

```
Set-Mailbox bsuneja@contoso.com -LitigationHoldEnabled $true
```

> **Note**
>
> When you place a mailbox on Litigation Hold indefinitely (by not specifying a duration period), the value for the *LitigationHoldDuration* property mailbox is set to `Unlimited`.

Use the Exchange Management Shell to place a mailbox on Litigation Hold and preserve items for a specified duration

This example places the mailbox bsuneja@contoso.com on Litigation Hold and preserves items for 2555 days (approximately 7 years).

PowerShell

```
Set-Mailbox bsuneja@contoso.com -LitigationHoldEnabled $true -
LitigationHoldDuration 2555
```

Use the Exchange Management Shell to place all mailboxes on Litigation Hold

Your organization may require that all mailbox data be preserved.

This example places all user mailboxes in the organization on Litigation Hold and sets the hold duration for one year (365 days).

PowerShell

```
Get-Mailbox -ResultSize Unlimited -Filter "RecipientTypeDetails -eq 'UserMailbox'"
| Set-Mailbox -LitigationHoldEnabled $true -LitigationHoldDuration 365
```

The example uses the Get-Mailbox cmdlet to retrieve all mailboxes in the organization, specifies a recipient filter to include all user mailboxes, and then pipes the list of mailboxes to the Set-Mailbox cmdlet to enable the Litigation Hold and set the hold duration.

To place all user mailboxes on an indefinite hold, run the previous command but don't include the *LitigationHoldDuration* parameter.

See the More information section for examples of using other recipient properties in a filter to include or exclude one or more mailboxes.

Use the Exchange Management Shell to remove a mailbox from Litigation Hold

This example removes the mailbox bsuneja@contoso.com from Litigation Hold.

PowerShell

```
Set-Mailbox bsuneja@contoso.com -LitigationHoldEnabled $false
```

How do you know this worked?

To verify that you have successfully placed a mailbox on Litigation Hold, do the one of the following:

- In the EAC:

    1. Go to **Recipients** > **Mailboxes**.
    2. In the list of user mailboxes, click the mailbox that you want to verify Litigation Hold settings for, and then click **Edit** ✎.
    3. On the mailbox properties page, click **Mailbox features**.
    4. Under **Litigation hold**, verify that hold is enabled.
    5. Click **View details** to verify when the mailbox was placed on Litigation Hold and by whom. You can also verify or change the values in the optional **Litigation hold duration (days)**, **Note**, and **URL** boxes.

- In the Exchange Management Shell, run one of the following commands:

    PowerShell

    ```
    Get-Mailbox <name of mailbox> | Format-List LitigationHold*
    ```

    or

    PowerShell

    ```
    Get-Mailbox -ResultSize Unlimited -Filter "RecipientTypeDetails -eq
    'UserMailbox'" | Format-List Name,LitigationHold*
    ```

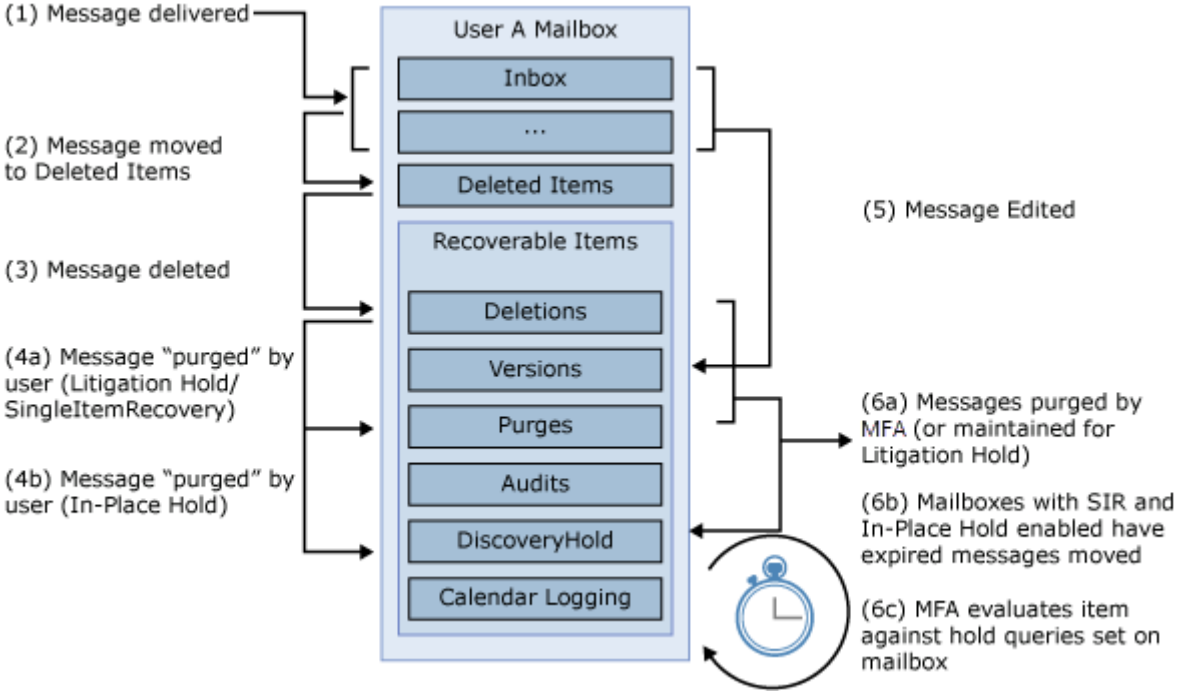
    If a mailbox is placed on Litigation Hold indefinitely, the value for the *LitigationHoldDuration* property mailbox is set to `Unlimited`.

More information

- **How does Litigation Hold work?** In the normal deleted item workflow, a mailbox item is moved to the Deletions subfolder in the Recoverable Items folder when a user permanently deletes it (Shift + Delete) or deletes it from the Deleted Items folder. A deletion policy (which is a retention tag configured with a Delete retention action) also moves items to the Deletions subfolder when the retention period expires. When a user purges an item in the Recoverable Items folder or when the deleted item retention period expires for an item, it's moved to the Purges subfolder in the Recoverable Items folder and marked for permanent deletion. It will be purged from Exchange the next time the mailbox is processed by the Managed Folder Assistant (MFA).

  When a mailbox is placed on Litigation Hold, items in the Purges subfolder are preserved for the hold duration specified by the Litigation Hold. The hold duration is calculated from the original date an item was received or created, and defines how long items in the Purges subfolder are held. When the hold duration expires for an item in the Purges subfolder, the item is marked for permanent deletion and will be purged from Exchange the next time the mailbox is processed by the MFA. If an indefinite hold is placed on a mailbox, items will never be purged from the Purges subfolder.

  The following illustration shows the subfolders in the Recoverable Items folders and the hold workflow process.



> **Note**
>
> If an In-Place Hold is placed on a mailbox, purged items are moved from the Deletions subfolder to the DiscoveryHolds subfolder and are preserved for the hold duration for the In-Place Hold.

- If your organization requires that all mailbox data has to preserved for a specific period of time, consider the following before you place all mailboxes in an organization on Litigation Hold.

  - When you use the previous command to place a hold on all mailboxes in an organization (or a subset of mailboxes matching a specified recipient filter) only mailboxes that exist at the time that you run the command are placed on hold. If you create new mailboxes later, you have to run the command again to place the new mailboxes on hold. If you create new mailboxes often, you can run the command as a scheduled task as frequently as required.

  - Placing all mailboxes on Litigation Hold can significantly impact mailbox sizes. In an Exchange 2016 or Exchange 2019 organization, plan for adequate storage to meet your organization's preservation requirements.

  - The Recoverable Items folder has its own storage limit, so items in the folder don't count towards the mailbox storage limit. As previously explained, preserving mailbox data for a long period of time will result in growth of the Recoverable Items folder in a user's mailbox and archive. We recommend that you periodically monitor the size of this folder by using the **Get-MailboxFolderStatistics** cmdlet to ensure it doesn't reach the limit. For more information, see:

    - Get-MailboxFolderStatistics

    - Clean up or delete items from the Recoverable Items folder.

- The previous command to place a hold on all mailboxes uses a recipient filter that returns all user mailboxes. You can use other recipient properties to return a list of specific mailboxes that you can then pipe to the **Set-Mailbox** cmdlet to place a Litigation Hold on those mailboxes.

  Here are some examples of using the **Get-Mailbox** and **Get-Recipient** cmdlets to return a subset of mailboxes based on common user or mailbox properties. These examples assume that relevant mailbox properties (such as *CustomAttributeN* or *Department*) have been populated.

  PowerShell

  ```
  Get-Mailbox -RecipientTypeDetails UserMailbox -ResultSize unlimited -Filter
  'CustomAttribute15 -eq "OneYearLitigationHold"'
  ```

  PowerShell

```
Get-Recipient -RecipientTypeDetails UserMailbox -ResultSize unlimited -Filter
'Department -eq "HR"'
```

PowerShell

```
Get-Recipient -RecipientTypeDetails UserMailbox -ResultSize unlimited -Filter
'PostalCode -eq "98052"'
```

PowerShell

```
Get-Recipient -RecipientTypeDetails UserMailbox -ResultSize unlimited -Filter
'StateOrProvince -eq "WA"'
```

PowerShell

```
Get-Mailbox -ResultSize Unlimited -Filter "RecipientTypeDetails -ne
'DiscoveryMailbox'"
```

You can use other user mailbox properties in a filter to include or exclude mailboxes. For details, see Filterable Properties for the -Filter Parameter.

Preserve Bcc and expanded distribution group recipients for classic eDiscovery

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

Litigation holds, eDiscovery holds, and Microsoft 365 retention policies created in the Microsoft Purview portal allow you to preserve mailbox content to meet regulatory compliance and eDiscovery requirements. Information about recipients directly addressed in the To and Cc fields of a message is included in all messages by default. But your organization may require the ability to search for and reproduce details about all recipients of a message. This includes:

- **Recipients addressed using the Bcc field of a message:** Bcc recipients are stored in the message in the sender's mailbox, but not included in headers of the message delivered to recipients.
- **Expanded distribution group recipients:** Recipients who receive the message because they're members of a distribution group to which the message was addressed, either in the To, Cc or Bcc fields.

Exchange Online and Exchange Server 2013 (Cumulative Update 7 and later versions) retain information about Bcc and expanded distribution group recipients. You can search for this information by using an eDiscovery tool in the Microsoft Purview portal.

How Bcc recipients and expanded distribution group recipients are preserved

> **Important**

Only the sender's mailbox's sent items will contain distribution group expansion

Information about Bcc'ed recipients is stored with the message in the sender's mailbox. This information is indexed and available to eDiscovery searches and holds.

Information about expanded distribution group recipients is stored with the message after you place a mailbox on In-Place Hold or Litigation Hold. In Microsoft 365, this information is also stored when a Microsoft Purview retention policy is applied to a mailbox. Distribution group membership is determined at the time the message is sent. The expanded recipients list stored with the message isn't impacted by changes to membership of the group after the message is sent.

Expand table

Information about...	Is stored in...	Is stored by default?	Is accessible to...
To: and Cc: recipients	Message properties in the sender and recipients' mailboxes.	Yes	Sender, recipients, and compliance officers
Bcc: recipients	Message property in the sender's mailbox.	Yes	Sender and compliance officers
Expanded distribution group recipients	Message properties in the sender's mailbox.	No. Expanded distribution group recipient information is stored after a mailbox is placed on In-Place Hold or Litigation Hold, or assigned to a Microsoft Purview retention policy.	Compliance officers

Searching for messages with expanded distribution group recipients

When searching for messages sent to a recipient, eDiscovery search results now include messages sent to a distribution group that the recipient is a member of but only in the sender mailbox and if the sender is on hold.

Important

Distribution group expansion details are only available in the sender's mailbox under Sent Items. If you search the recipient's mailbox, messages sent to distribution groups will not show the expanded list of group members.

Scenario: John is a member of the US-Sales distribution group. This table shows eDiscovery search results when Bob sends a message to John directly or indirectly via a distribution group.

Expand table

When you search Bob's mailbox for messages sent...	And the message is sent with...	Results include message?
To: John	John on To:	Yes
To: John	US-Sales on To:	Yes
To: US-Sales	US-Sales on To:	Yes
Cc: John	John on Cc:	Yes
Cc: John	US-Sales on Cc:	Yes
Cc: US-Sales	US-Sales on Cc:	Yes
When you search John's mailbox for message...		
To/Cc	US-Sales on To/Cc:	No. Only sender mailbox can be searched.

Searching for messages with Bcc

Scenario: Bob sends an email to John (To/Cc) and Jack (Bcc directly, or indirectly via a distribution group). The tables below show eDiscovery search results.

Important

Distribution group expansion detail will only be available if the sender is on hold.

Expand table

When you search Bob's mailbox...	For messages sent...	Results include message?	Notes
Bob's mailbox	To:/Cc: John	Yes	Presents an indication that Jack was Bcc'ed.
Bob's mailbox	Bcc: Jack	Yes	Presents an indication that Jack was Bcc'ed.
Bob's mailbox	Bcc: Jack (via distribution group)	Yes	List of members of the Bcc'ed distribution group, expanded when the message was sent, is visible in eDiscovery search preview, export, and logs.
When you search John's mailbox...			
John's mailbox	To:/Cc: John	Yes	No indication of Bcc recipients.
John's mailbox	Bcc: Jack (directly or via distribution group)	No	Bcc: information isn't stored in the message delivered to recipients. You must search the **sender's mailbox**.
When you search Jack's mailbox...			
Jack's mailbox	To:/Cc: John (directly or via distribution group)	Yes	To:/Cc: information is included in message delivered to all recipients.
Jack's mailbox	Bcc: Jack (directly or via distribution group)	No	Bcc: information isn't stored in the message delivered to recipients. You must search the **sender's mailbox**.

Frequently asked questions

When and where is Bcc recipient information stored?

Bcc recipient information is preserved by default in the original message in sender's mailbox. If the Bcc recipient is a distribution group, distribution group membership is only expanded if the sender's mailbox is on hold or assigned to a Microsoft 365 retention policy.

When and where is the list of expanded distribution group recipients stored?

Group membership is expanded at the time the message is sent. The list of expanded distribution group members is stored in the original message in the sender's mailbox. The sender's mailbox must be on In-Place Hold, Litigation Hold, or assigned to a Microsoft 365 retention policy.

Can the To/Cc recipients see which recipients were Bcc'ed?

No. This information isn't included in message headers, and isn't visible to To/Cc recipients. The sender can see the Bcc field stored in the original message stored in their mailbox. Compliance officers can see this information when searching the sender's mailbox.

How can I ensure that expanded distribution group recipients are always preserved?

To ensure that expanded distribution group members are always preserved with a message, Place all mailboxes on hold or create an organization-wide Microsoft 365 retention policy.

Which types of groups are supported?

Distribution groups, mail-enabled security groups, and dynamic distribution groups are supported.

Is there a limit on the number of distribution group recipients that are expanded and stored in the message?

Up to 10,000 members of a distribution group is preserved.

Are nested distribution groups supported?

Yes, 25 levels of nested distribution groups are expanded.

Where is the Bcc and expanded distribution group recipient information visible?

Bcc and expanded distribution group recipients information is visible to Compliance officers when performing an eDiscovery search. Bcc and expanded distribution group recipients are included in search results copied to a Discovery mailbox or exported to a PST file and in the eDiscovery log included in search results. Bcc recipient information is also available in search preview.

What happens if a member of a distribution group is hidden from the organization's global address list (GAL)?

There's no impact. If recipients are hidden from the GAL, they're still included in the list of recipients for the expanded distribution group.

Last updated on 06/05/2025

Work with processing errors in eDiscovery (Premium)

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

Processing is the process of file identification, expansion of embedded documents and attachments, text extraction, and Optical Character Recognition (OCR) of image files and the subsequent indexing of that content.

When you add custodians and non-custodian data sources to a case on the **Sources** tab, all partially indexed items from Microsoft 365 are processed to make them fully searchable. Likewise, when content is added to a review set from both Microsoft 365 and non-Microsoft 365 data sources, this content is also processed.

The **Processing** tab in eDiscovery (Premium) provides insight into the status of Advanced indexing for different processing scenarios.

For more information, see the following articles:

- Advanced indexing of custodian data
- Error remediation when processing data
- Single item error remediation

Last updated on 11/21/2025

Advanced indexing of custodian and noncustodial data sources

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

When a custodian or noncustodial data source is added to an eDiscovery (Premium) case, any content that was deemed as partially indexed or had indexing errors is reindexed. The reindexing process is called *Advanced indexing*. There are many reasons that content is partially indexed or has indexing errors. This includes image files or the presence of images in a file, unsupported file types, or file sized indexing limits.

For SharePoint files, Advanced indexing only runs on items marked as partially indexed or items with indexing errors. In Exchange, email messages with image attachments aren't marked as partially indexed or with indexing errors. This means that those files won't be reindexed by the Advanced indexing process.

To learn more about processing support and partially indexed items, see:

- Supported file types in eDiscovery (Premium)
- Partially indexed items in eDiscovery
- File formats indexed by Exchange Search
- Default crawled file name extensions and parsed file types in SharePoint Server

Viewing Advanced indexing results

After the Advanced indexing process is completed, you can get an understanding of the effectiveness of reprocessing. In the Advanced indexing results view on the **Processing** tab for a case, the graph lists the number of items added to the *hybrid index*. The hybrid index is where eDiscovery (Premium) stores the reprocessed content.

This view also includes the number of items that require remediation and another graph of errors by file type. For more information, see:

- Error remediation when processing data
- Single item error remediation

Updating the Advanced index

When a custodian or noncustodial data source is added to an eDiscovery (Premium) case, all partially indexed items are reprocessed. However, as time passes, more partially indexed items may be added to a user's mailbox or OneDrive account. If necessary, you can update the index for specific custodian or noncustodial data source. For more information, see Manage custodians in an eDiscovery (Premium) case. You can also update the index for all custodians and noncustodial data sources in a case by selecting the **Update index** on the **Processing** tab.

> **Note**
>
> Updating custodian and noncustodial indexes is a long running process. It's recommended that you don't update indexes more than once a day in a case.

Last updated on 11/21/2025

Error remediation when processing data

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

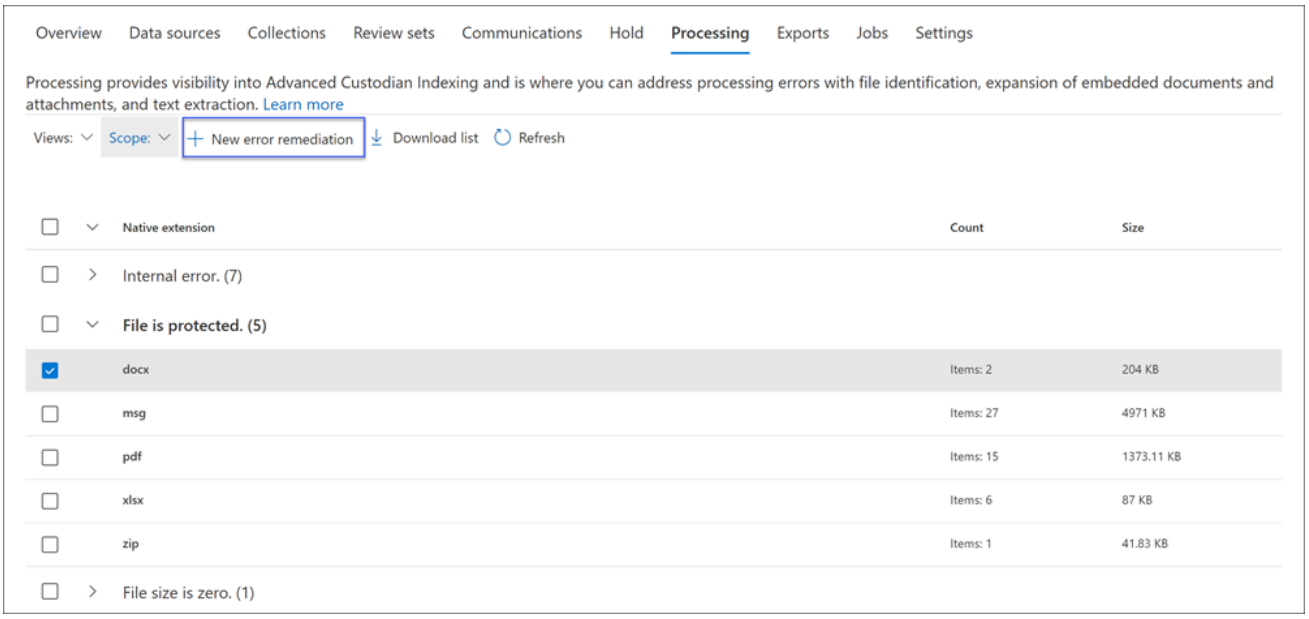
Error remediation allows eDiscovery administrators the ability to rectify data issues that prevent Microsoft Purview eDiscovery (Premium) from properly processing the content. For example, files that are password protected can't be processed since the files are locked or encrypted. Using error remediation, eDiscovery administrators can download files with such errors, remove the password protection, and then upload the remediated files.

Use the following workflow to remediate files with errors in eDiscovery (Premium) cases.

Create an error remediation session to remediate files with processing errors

> **Note**
>
> If the error remediation workflow is closed at any time during the following procedure, you can return to the error remediation session from the **Processing** tab by selecting **Remediations** in the **View** drop-down menu.

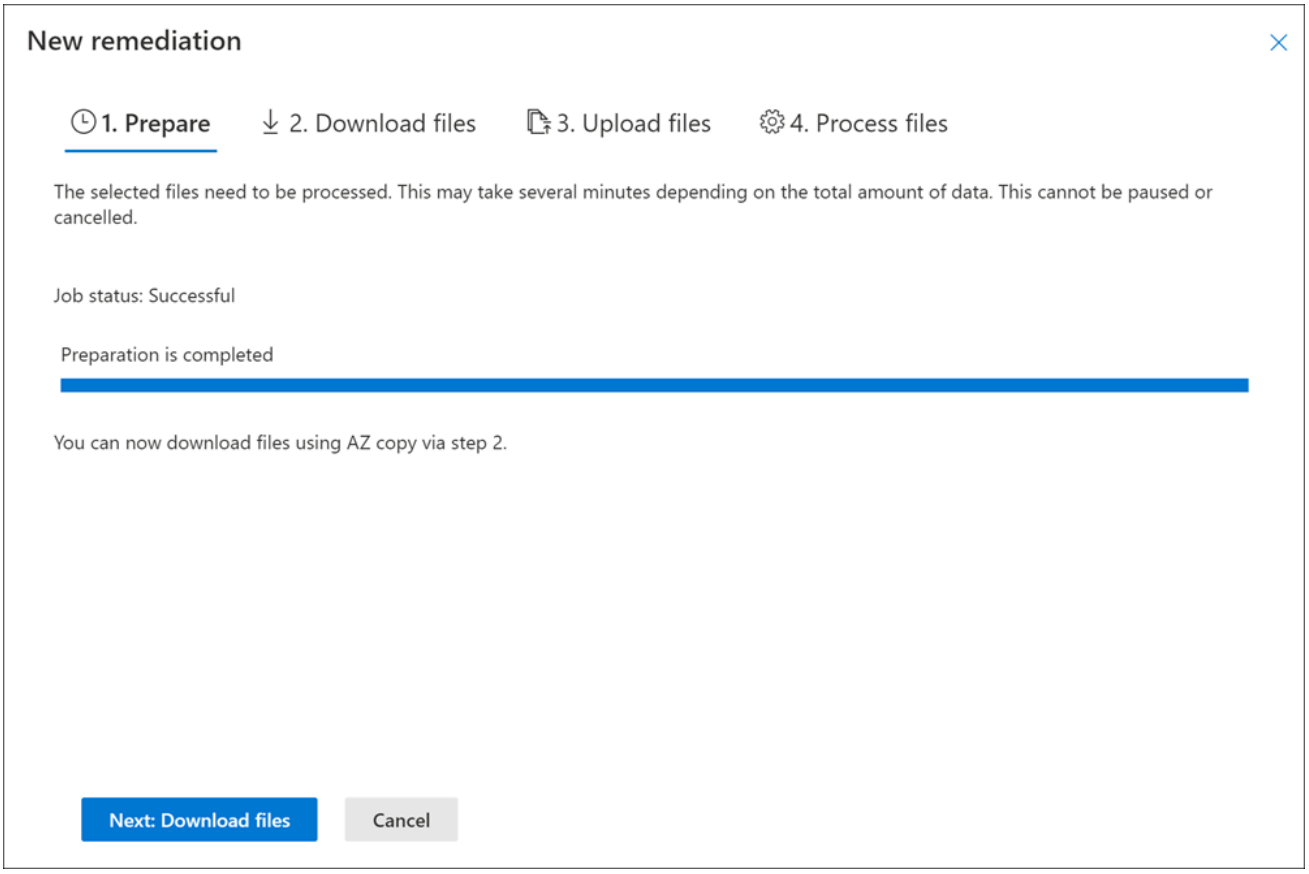
1. On the **Processing** tab in the eDiscovery (Premium) case, select **Errors** in the **View** drop-down menu and then select a review set in the **Scope** drop-down menu. You must select a review set in the scope to start the remediation of any errors. After you've selected the scope, this section displays errors from the specific review set.



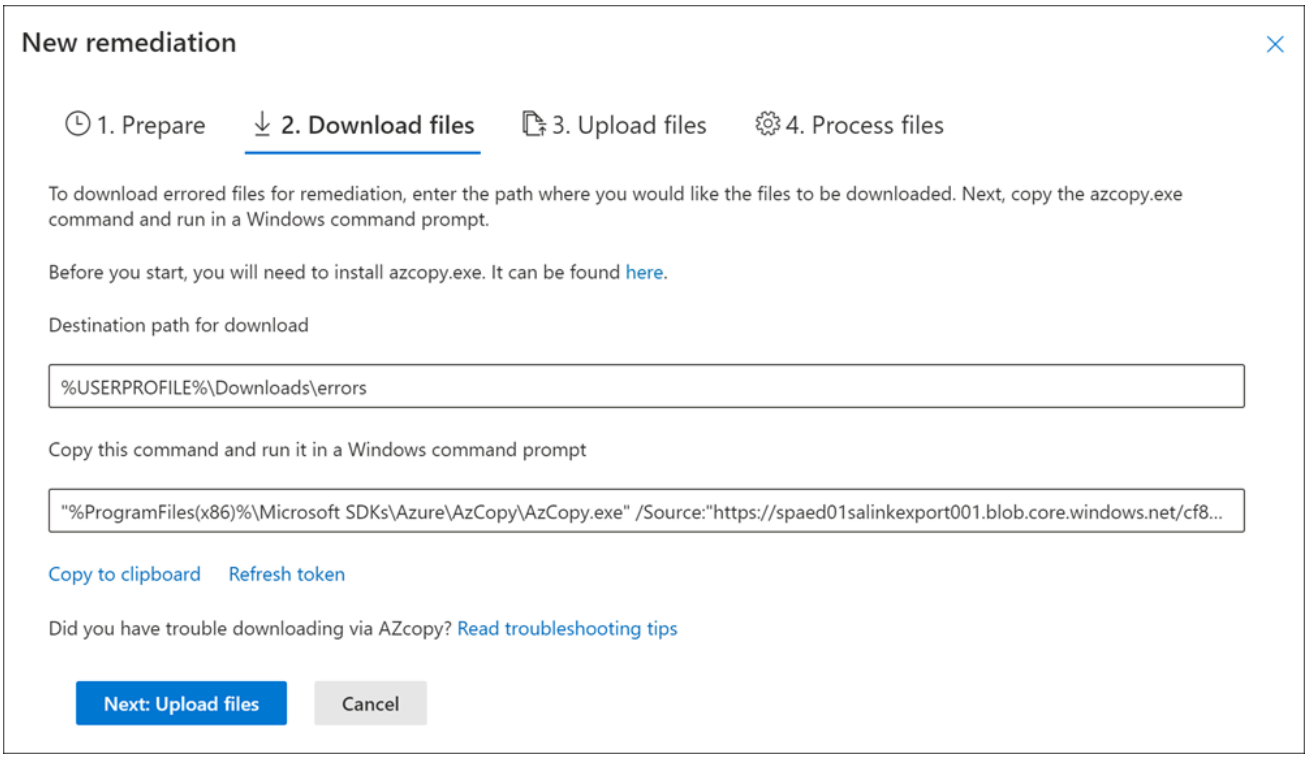
2. Select the errors you want to remediate by selecting either the error type or file type option. In the following example, we're remediating a password protected file.

3. Select **New error remediation**.

   The error remediation workflow starts with a preparation stage where the files with errors are copied to a Microsoft-provided Azure Storage location so that you can download them to your local computer to remediate.



4. After the preparation is complete, select **Next: Download files** to proceed with download.



5. To download files, specify the **Destination path for download**. This is a path to the parent folder on your local computer where the file will be downloaded. The default path, %USERPROFILE%\Downloads\errors, points to the logged-in user's downloads folder. You can change this path if desired. If you do change it, we recommend that you use a local file path for the best performance. Don't use a remote network path. For example, you could use the path **C:\Remediation**.

   The path to the parent folder is automatically added to AzCopy command (as the value of the **/Dest** parameter).

6. Copy the predefined command by selecting **Copy to clipboard**. Open a Windows Command Prompt, paste the AzCopy command, and then press **Enter**.

   > **Note**
   >
   > You should use the latest version of AzCopy to successfully use the command that's provided on the **Download files** page. You also must use AzCopy to upload the files in step 10. To install the latest version of AzCopy, see Get started with AzCopy. If the supplied AzCopy command fails, see Troubleshoot AzCopy in eDiscovery (Premium).

   The files that you selected are downloaded to the location that you specified in step 5. In the parent folder (for example, **C:\Remediation**), the following subfolder structure is automatically created:

`<Parent folder>\Subfolder 1\Subfolder 2\<file>`

- *Subfolder 1* is named with the ID for the case or the review set, depending on the scope that you selected in step 1.
- *Subfolder 2* is named with the file ID of the downloaded file
- The downloaded file is located in *Subfolder 2* and is also named with the file ID.

Here's an example of the folder path and error file name that's created when items are downloaded to the **C:\Remediation** parent folder:

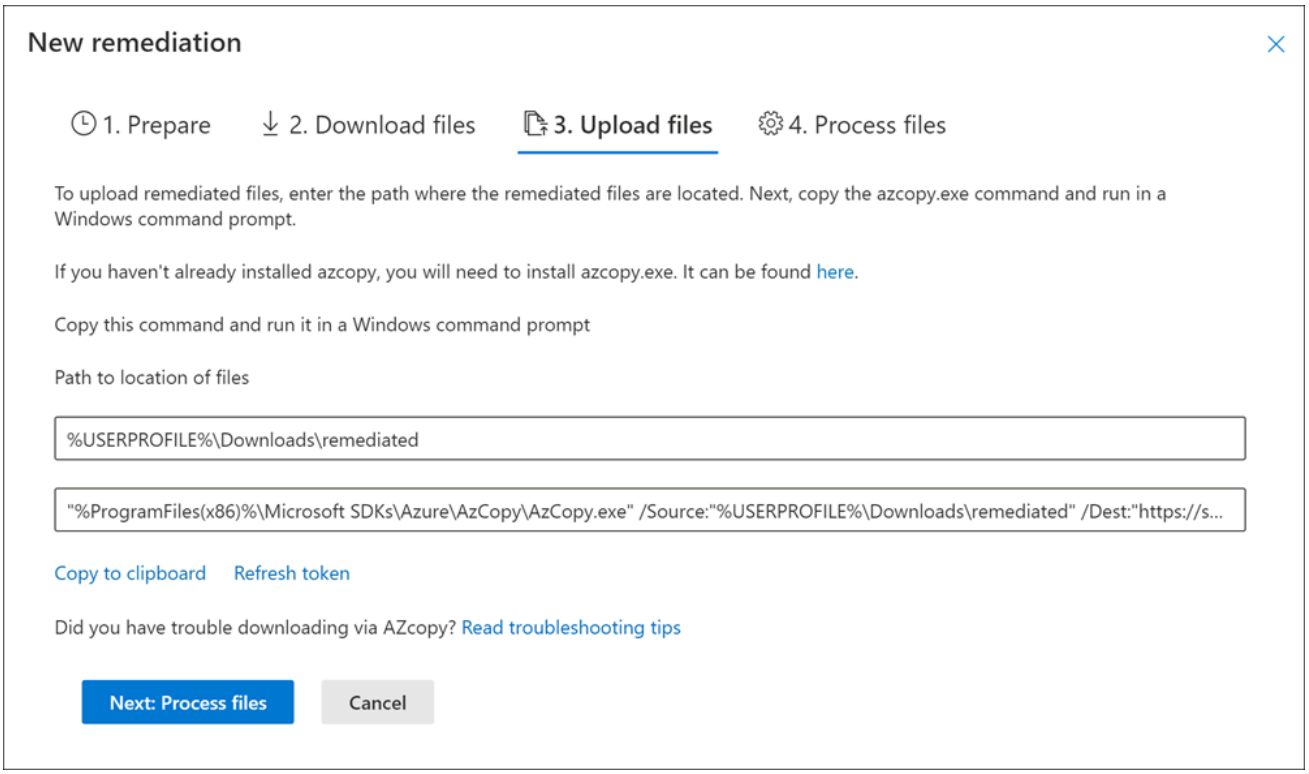
`C:\Remediation\232f8b7e-089c-4781-88c6-210da0615d32\d1459499146268a096ea20202cd029857d64087706e6d6ca2a224970ae3b8938\d1459499146268a096ea20202cd029857d64087706e6d6ca2a224970ae3b8938.docx`

If multiple files are downloaded, each one is downloaded to a subfolder that's named with the file ID.

> **ⓘ Important**
>
> When you upload files in step 9 and step 10, the remediated files must have that same filename and be located in the same subfolder structure. The subfolder and file names are used to associated the remediated file with the original error file. If the folder structure or file names are changed, you'll receive the following error: `Cannot apply Error Remediation to the current Workingset`. To prevent any issues, we recommend that keep the remediated files in the same parent folder and subfolder structure.

7. After downloading the files, you can remediate them with an appropriate tool. For password-protected files, there are several password cracking tools you can use. If you know the passwords for the files, you can open them and remove the password protection.

8. Return to eDiscovery (Premium) and the error remediation workflow and then select **Next: Upload files**. This moves to the next page where you can now upload the files.



9. Specify the parent folder where the remediated files are located in the **Path to location of files** text box. Again, the parent folder must have the same subfolder structure that was created when you downloaded the files.

   The path to the parent folder is automatically added to AzCopy command (as the value of the **/Source** parameter).

10. Copy the predefined command by selecting **Copy to clipboard**. Open a Windows Command Prompt, paste the AzCopy command, and then press **Enter**. upload the files.

11. After you run the AzCopy command, select **Next: Process files**.

    When processing is complete, you can go to review set and view the remediated files.

Remediating errors in container files

In situations when the contents of a container file (such as a .zip file) can't be extracted by eDiscovery (Premium), the containers can be downloaded and the contents expanded into the same folder in which the original container resides. The expanded files will be attributed to the parent container as if it was originally expanded by eDiscovery (Premium). The process works as described previously except for uploading a single file as the replacement file. When you upload remediated files, don't include the original container file.

Remediating errors by uploading the extracted text

Sometimes it's not possible to remediate a file to native format that eDiscovery (Premium) can interpret. But you can replace the original file with a text file that contains the original text of the native file (in a process called *text overlay*). To do this, follow the steps described in this article but instead of remediating the original file in the native format, you would create a text file that contains the extracted text from the original file, and then upload the text file using the original filename appended with a .txt suffix.

For example, you download a file during error remediation with the filename 335850cc-6602-4af0-acfa-1d14d9128ca2.abc. You open the file in the native application, copy the text, and then paste it into a new file named 335850cc-6602-4af0-acfa-1d14d9128ca2.abc.txt. When you do this, be sure to remove the original file in the native format from the remediated file location on your local computer before uploading the remediated text file to eDiscovery (Premium).

> **Note**
>
> The number of files used in error remediation upload must match the number of files downloaded from the review set. If the number of uploaded files doesn't match the number of downloaded files, the matching process will fail (even if AzCopy reports that the upload was successful).

What happens when files are remediated

When remediated files are uploaded, the original metadata is preserved except for the following fields:

- ExtractedTextSize
- HasText
- IsErrorRemediate
- LoadId
- ProcessingErrorMessage
- ProcessingStatus
- Text
- WordCount
- WorkingsetId

For a definition of all metadata fields in eDiscovery (Premium), see Document metadata fields.

Last updated on 11/21/2025

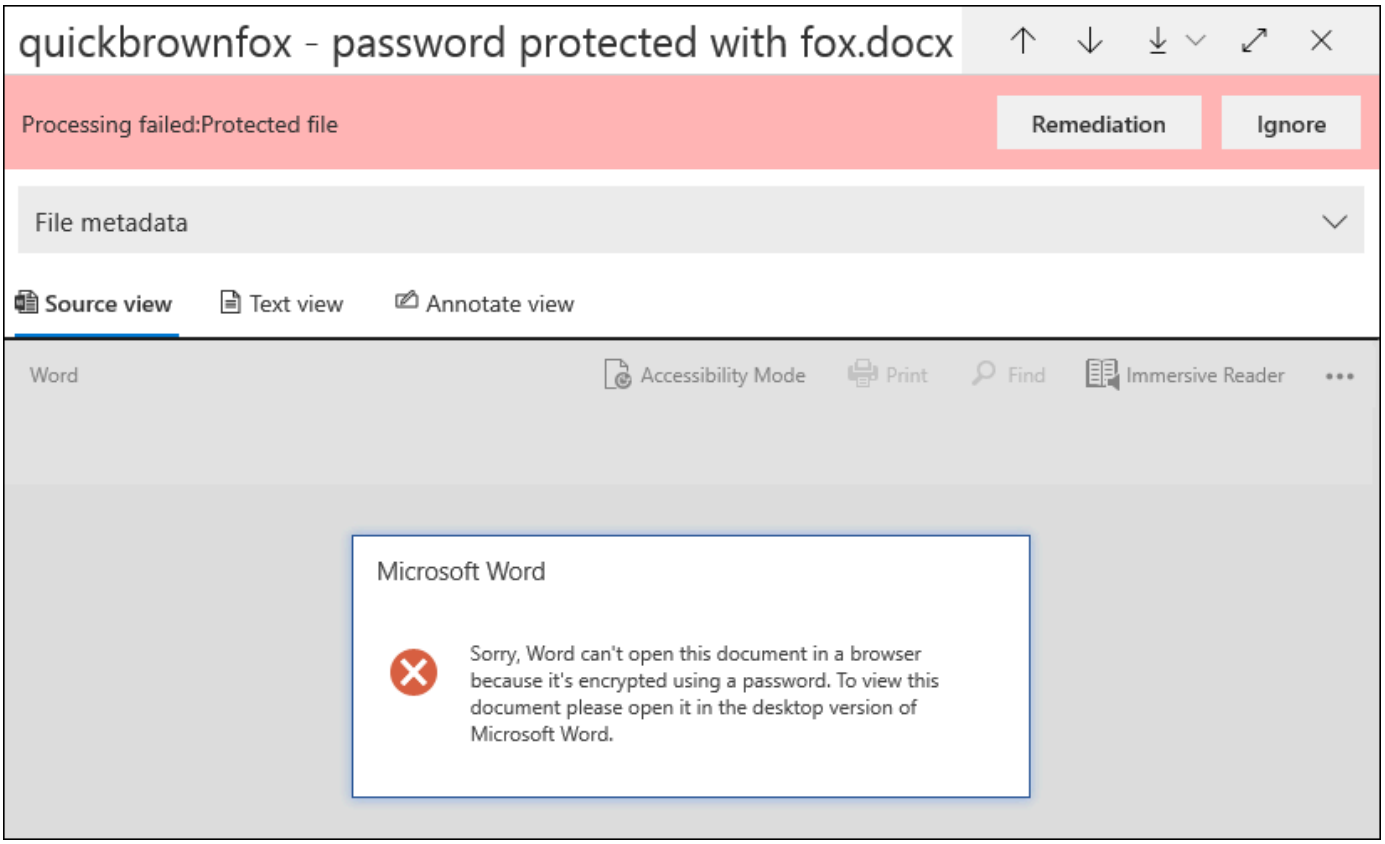
Single item error remediation in eDiscovery (Premium)

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

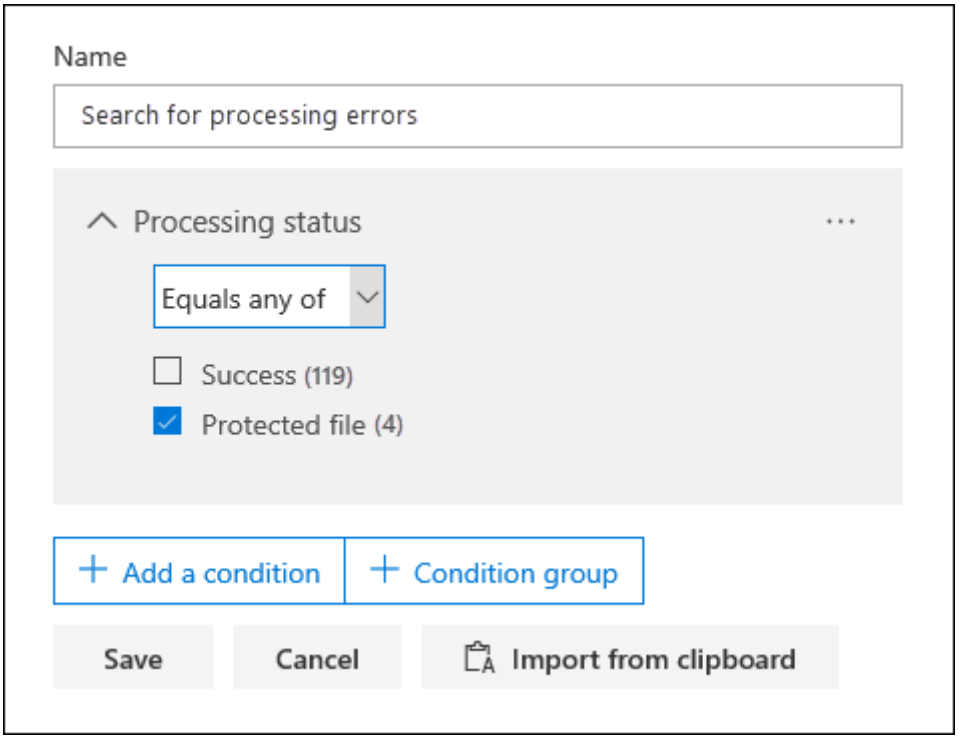
Error remediation gives Microsoft Purview eDiscovery (Premium) users the ability to rectify data issues that prevent eDiscovery (Premium) from properly processing the content. For example, files that are password protected can't be processed because those files are locked or encrypted. Previously, you could only remediate errors in bulk by using this workflow. But sometimes, it doesn't make sense to remediate errors in multiple files when you're unsure if any of those files are responsive to the case you're investigating. It also might not make sense to remediate errors before you've had a chance to review the file metadata (such as file location or who had access) to help you make up-front decisions about responsiveness. A new feature called *single item error remediation* gives eDiscovery managers the ability to view the metadata of files with a processing error and if necessary remediate the error directly in the review set. The article discusses how to identify, ignore, and remediate files with processing errors in a review set.

Identify documents with errors

Documents with processing errors in a review set are now identified (with a banner). You can remediate or ignore the error. The following screenshot shows the processing error banner for a Word document in a review set that is password-protected. Also notice that you can view the file metadata of documents with processing errors.



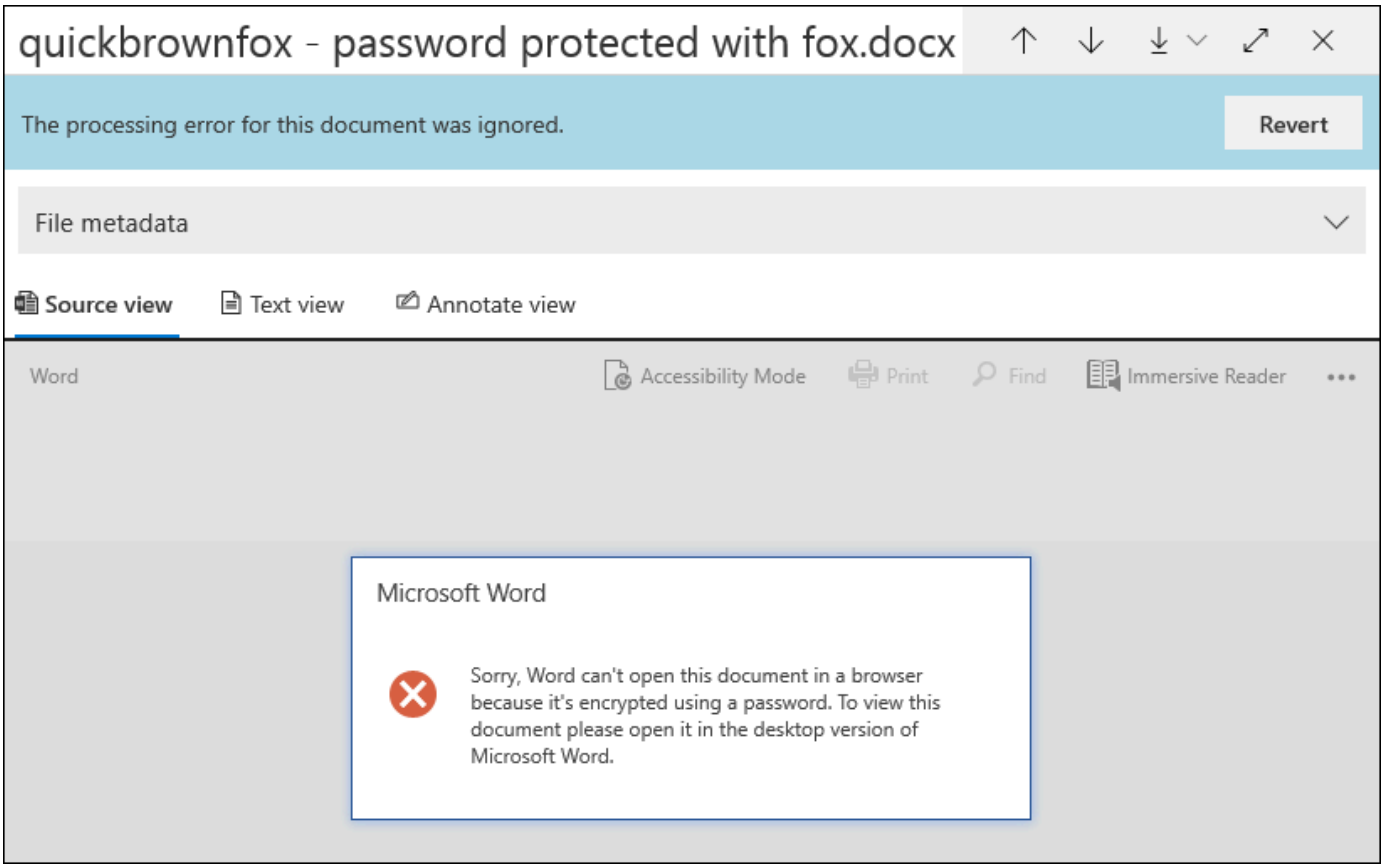
You can also search for documents that have processing errors by using the **Processing status** condition when querying the documents in a review set.



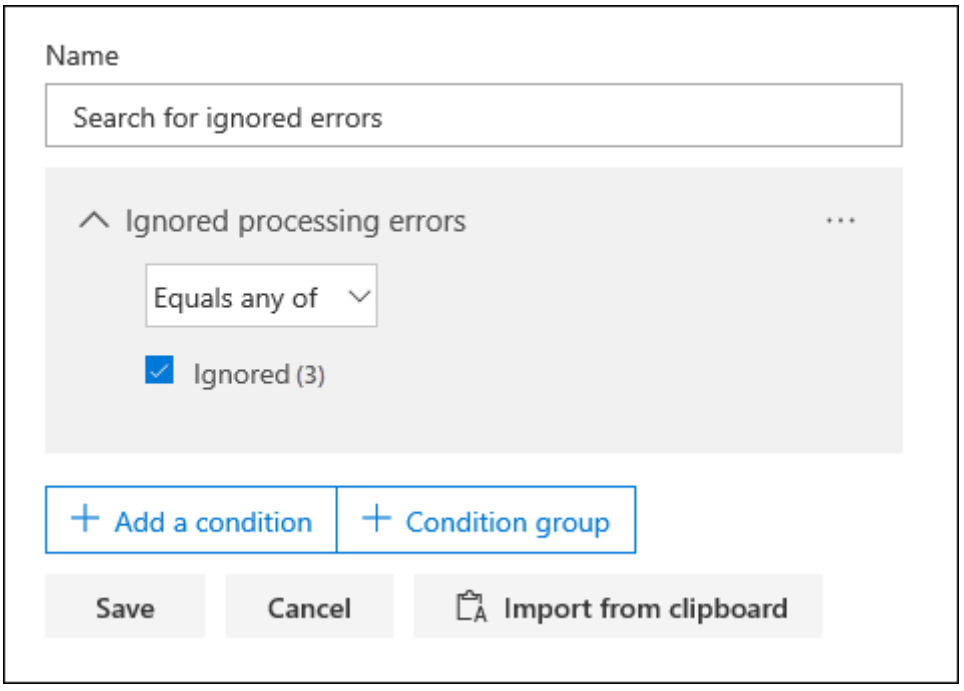
Ignore errors

You can ignore a processing error by selecting **Ignore** in the processing error banner. When you ignore an error, the document is removed from the bulk error remediation workflow. After an

error is ignored, the document banner changes color and indicates that the processing error was ignored. At any time, you can revert the decision to ignore the error by selecting **Revert**.



You can also search for all documents that had a processing error that was ignored by using the *Ignored processing errors* condition when querying documents in a review set.

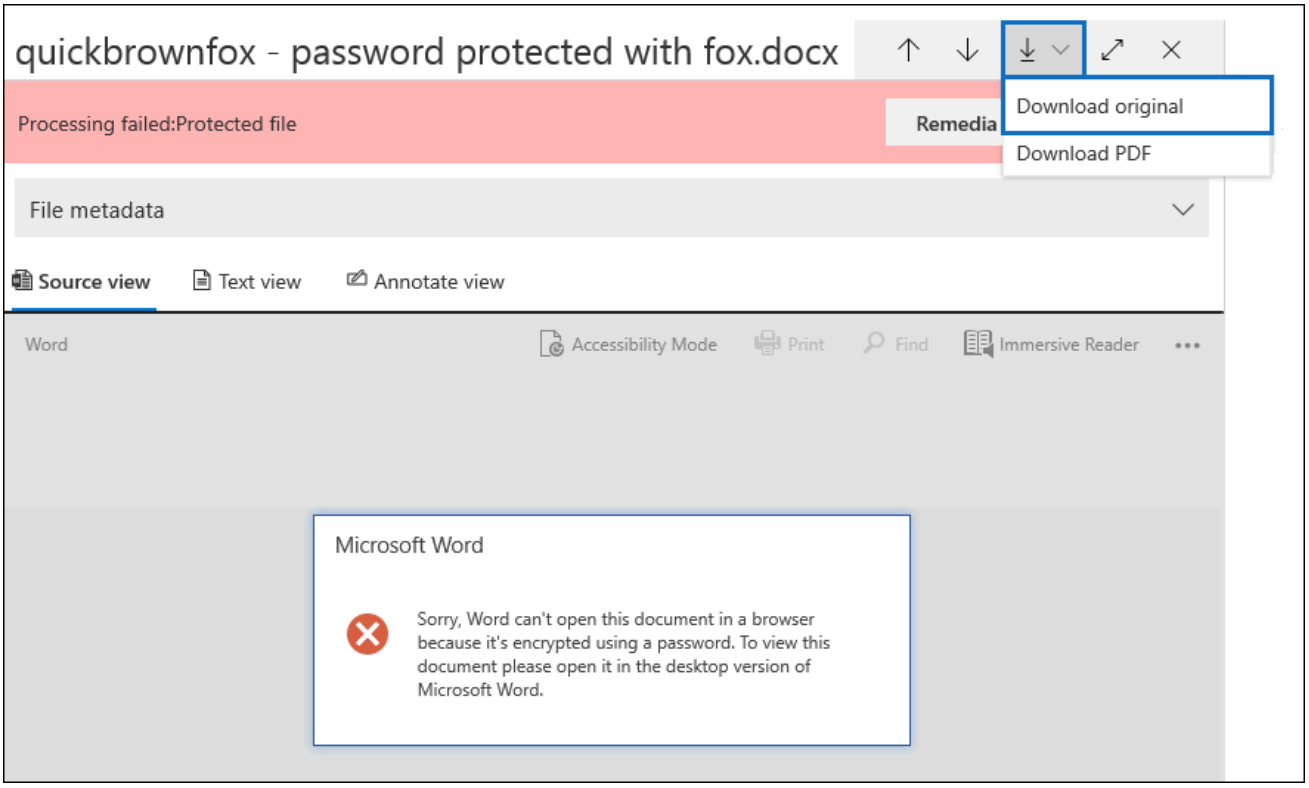


Remediate a document with errors

Sometimes you may be required to remediate a processing error in documents (by removing a password, decrypting an encrypted file, or recovering a corrupted document) and then add the remediated document to the review set. This allows you to review and export the error document together with the other documents in the review set.

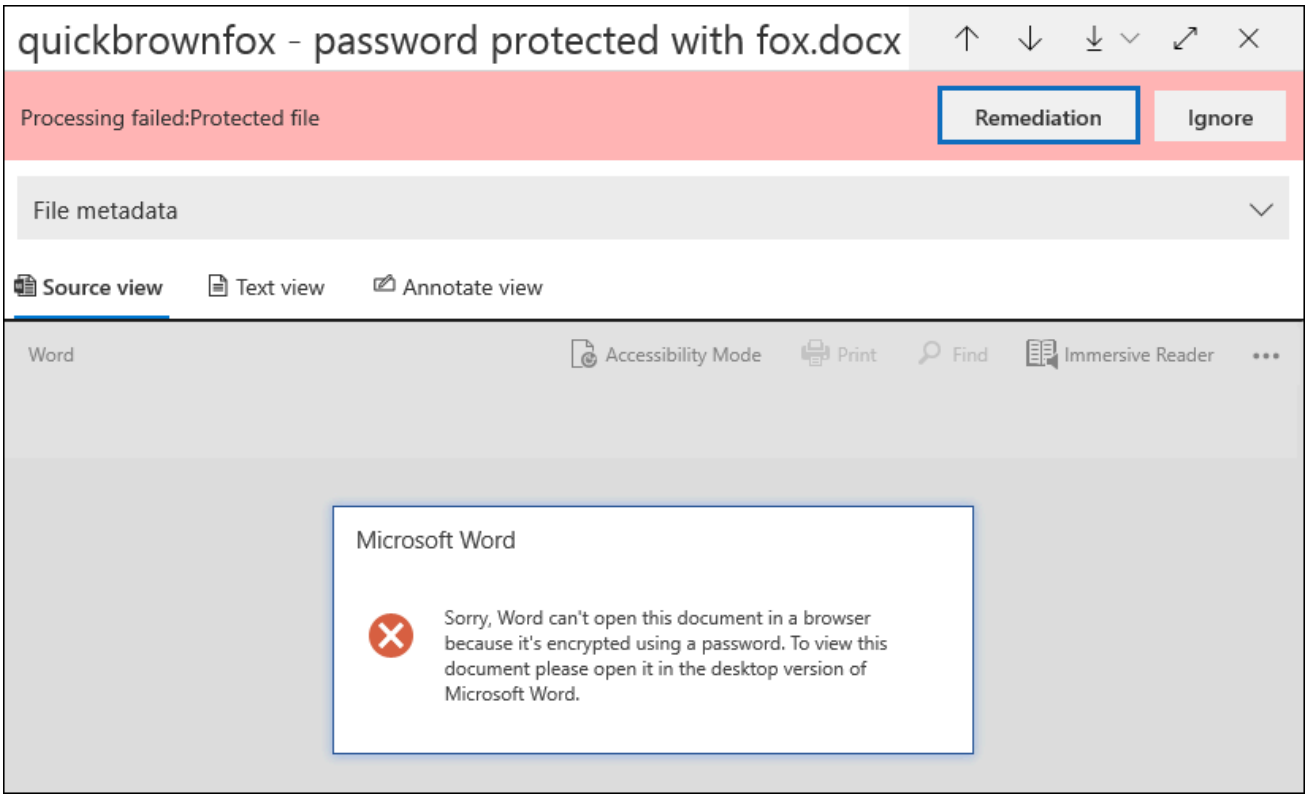
To remediate a single document, follow these steps:

1. Select **Download** > **Download original** to download a copy of the file to a local computer.

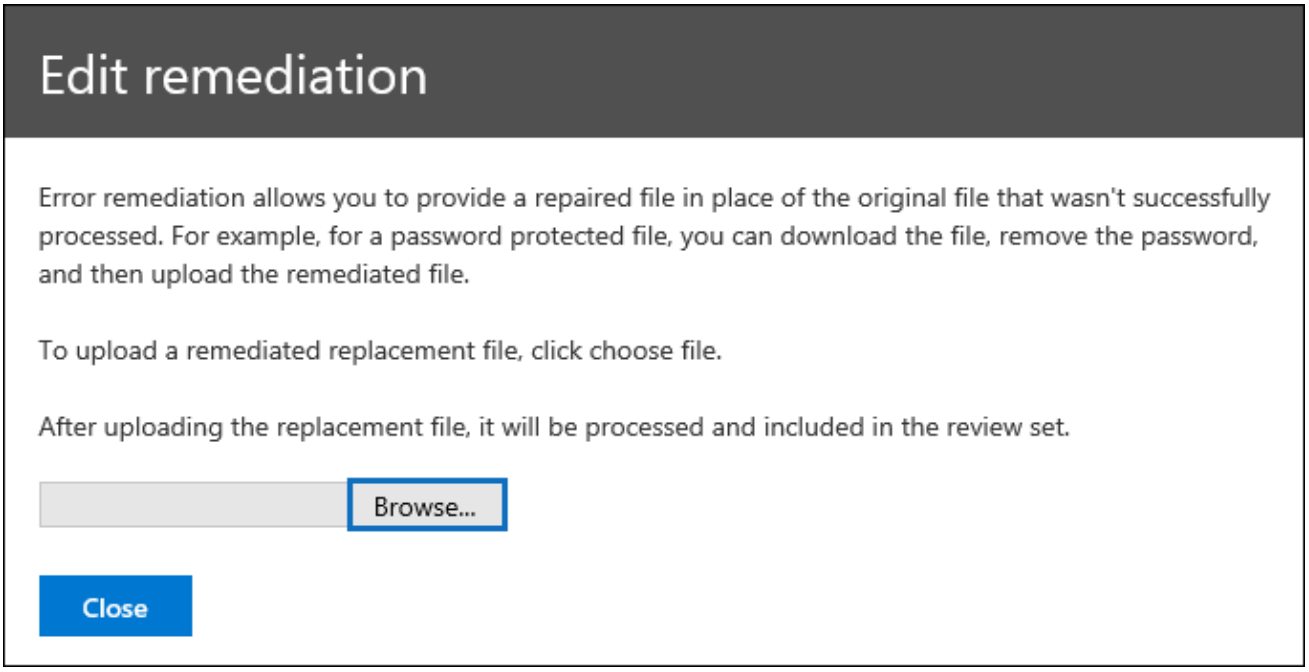


2. Remediate the error in the file offline. For encrypted files, that would require decryption software, to remove password protection, either provide the password and save the file or use a password cracker. After you remediate the file, go to the next step.

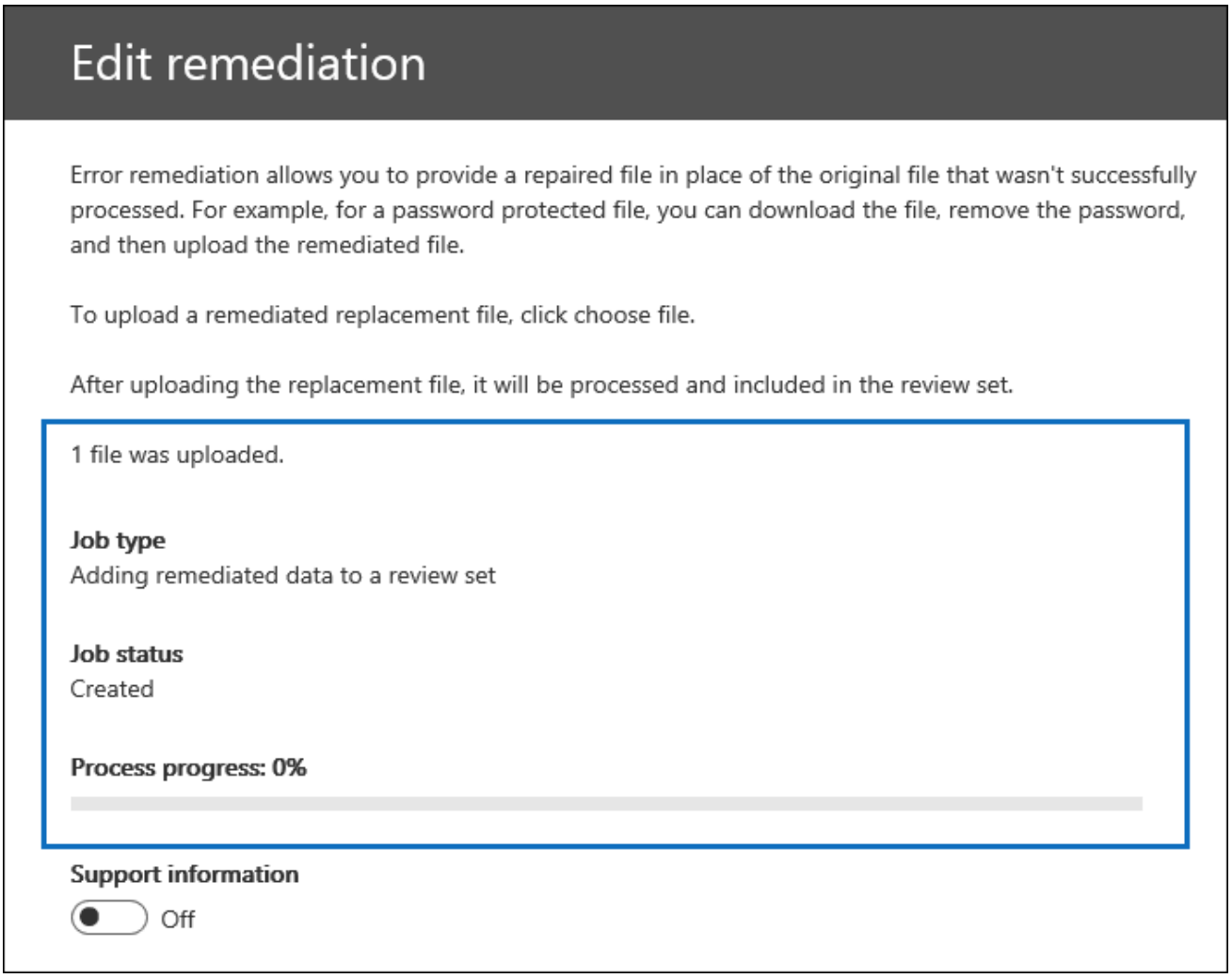
3. In the review set, select the file with the processing error that you remediated, and then select **Remediation**.



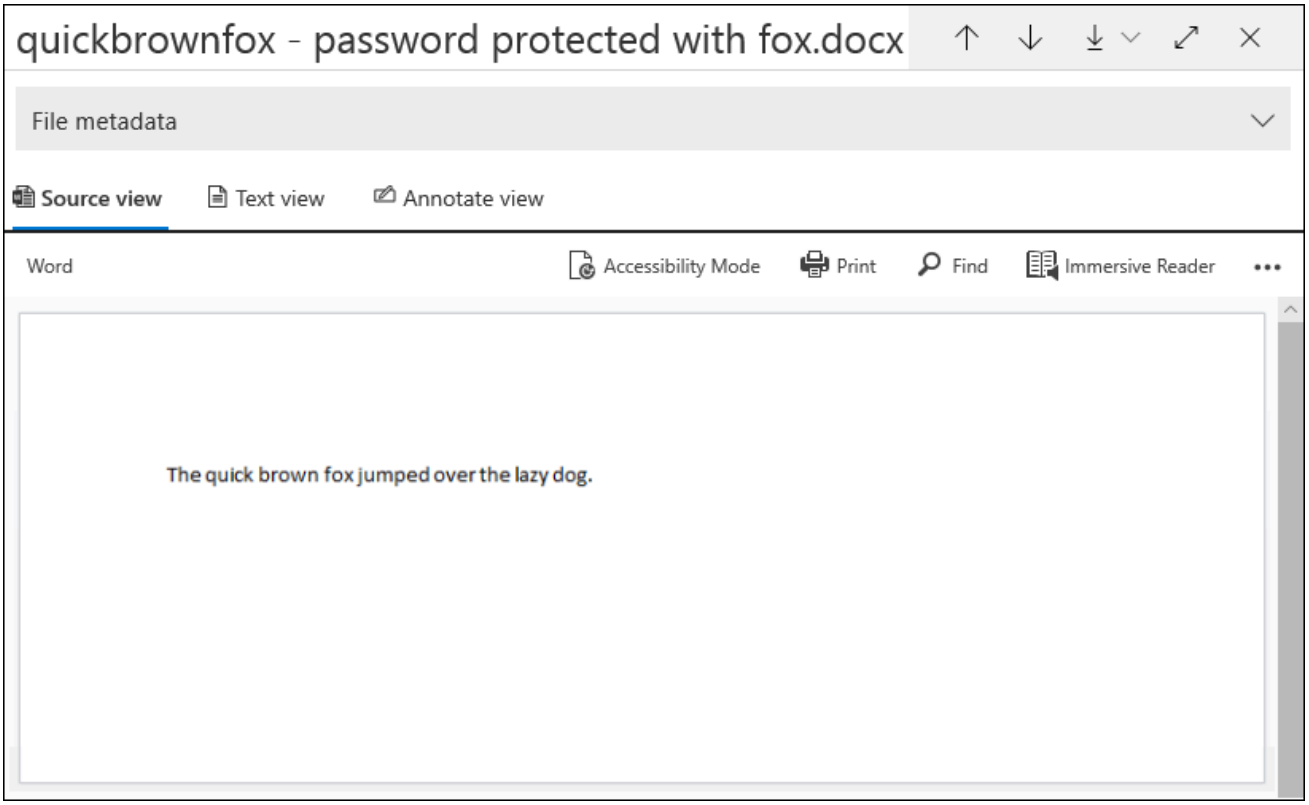
4. Select **Browse**, go to the location of the remediated file on your local computer, and then select the file.



After selecting the remediated file, it is automatically uploaded to the review set. You can track the processing status of the file.



After processing is completed, you can view the remediated document.



For more information about what happens when a document is remediated, see What happens when files are remediated.

Search for remediated documents

You can search for all documents in a review set that were remediated by using the **Keywords** condition and specifying the following property:value pair: **IsFromErrorRemediation:true**. This property is also available in the export load file when you export documents from a review set.

Last updated on 11/21/2025

Learn about collections in eDiscovery (Premium)

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

When organizations are faced with gathering the communications and content that may be relevant to an investigation or potential litigation, they face a significant challenge under the best of circumstances. In today's modern workplace, the volume, variety, and velocity of content is enabling innovation and remote work, while also expanding the requirements and process for managing collections for eDiscovery investigations.

Collections in eDiscovery (Premium) help eDiscovery managers quickly scope a search for content across email, documents, Teams reactions, and other content in Microsoft 365. Collections provide managers with an estimate of the content that may be relevant to the case. This estimate allows managers to make quick, informed decisions about the size and scope of content relevant to a case. eDiscovery managers can create a collection to search custodial data sources (such as mailboxes and SharePoint sites), non-custodial data sources (such as SharePoint sites, Group mailboxes, and Microsoft Teams), and by using specific search criteria (such as keywords and date ranges) to quickly define the scope of their collection.

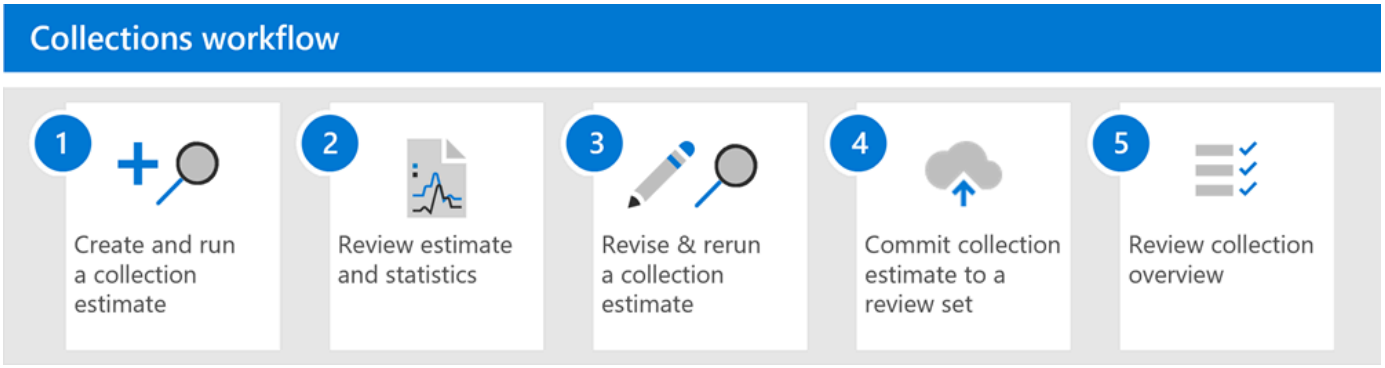
After the collection is defined, eDiscovery managers must first submit the collection to create an estimate. The estimate includes estimates for data volume, the content locations that contain results, and the number of hits for search query condition, and more. These insights can help to inform if the collection should be revised to narrow or expand the scope of the collection before moving on the review and analyze stages in the eDiscovery workflow.

When the manager is satisfied with the scope of the collection and the estimated amount of content that's likely to be responsive, the manager can add or *commit* the content to a review set.

When committing a collection to a review set, that manager also has the options to include chat conversations, cloud attachments, partially indexed items, and document versions. The content in the collection also goes through another level of processing during ingestion into the review set. and the collection will be updated with the final collection summary. After content is added to the review set, eDiscovery managers can continue to query, group, and refine the content in to help with minimization and review. Additionally, the collection is updated with information and statistics about the content committed to the review set. This information provides a historical reference about the content in the collection.

Collections workflow

To get started using collections in eDiscovery (Premium), here's a basic workflow and descriptions of each step in the process.



1. **Create and run a collection estimate**. The first step is to create a collection estimate and define the custodial and non-custodial data sources to search. You can also search other data sources that haven't been added to the case. After you add the data sources, you configure the search query to search the data sources for content relevant to the case. You can use keywords, properties, and conditions to build search queries that return content that's likely most relevant to the case. For more information, see Create a collection estimate.

2. **Review estimate and statistics**. After you create a collection estimate and run it, the next step is to view collection statistics to help you verify whether relevant content is being found and the content locations with the most hits. You can also review a sample of the search results to further help you determine if the content is within scope of your investigation. For more information, see Search statistics for collection estimates.

3. **Revise & rerun a collection estimate**. Based on the estimates and statistics returned by the collection, you can edit the collection estimate by changing the data sources that are searched and the search query to expand or narrow the collection. You can update and

rerun the collection estimate until you're confident that collection contains the content that's most relevant to your case.

4. **Commit a collection estimate to a review set**. When you're satisfied that the collection returns the type content that is relevant to the case, you can commit the collection to the review set. When you commit a collection, you have the option to add conversation threads, cloud attachments, and document versions to the review set, all of which might be relevant to the case.

   When you commit a collection, child items such as email signatures and images are extracted from a parent item (such as an email message, chat message, or document) and then processed by Optical Character Recognition (OCR) to extract any text from the child item. Text extracted from child items is then added to its parent item so you can view it in the review set. By not adding child items to the review set as a separate file, eDiscovery (Premium) helps limit the number of potentially immaterial items added to the review set. For more information about how child items are handled, see Collection statistics and reports.

   For more information, see Commit a collection estimate to a review set.

5. **Review collection overview**. After you commit a collection to a review set, information about the collection is retained, such as statistics about extracted items, deep indexing, the search query used for the collection, and the content locations that items were collected from. Also, committed collections can't be edited or rerun. You can only copy or delete them. Preserving collections provides a historical record of the collected items that were added to a review set. For more information, see Statistics and reports for committed collections.

Last updated on 09/02/2025

Create a collection estimate in eDiscovery (Premium)

> **⊗ Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

After you've identified custodians and any non-custodian data sources for the case, you're ready to identify and locate a set of documents that are relevant by using the **Collections** tool to search data sources for relevant content. You can create a collection that searches specified data sources for content that matches your search criteria. You must create a *collection estimate*, which is an estimate of the items found.

When you create a collection estimate, you can view information about the estimated results that matched the search query, such as the total number and size of items found, the different data sources for the found items, and statistics about the search query. You can also review a sample of items that were returned by the collection. Using these statistics, you can change the search query and rerun the collection estimate to narrow your results. Once you're satisfied with the collection results, you can commit the collection to a review set. When you commit a collection estimate, the items returned by the collection are added to a review set for review, analysis, and export.

Recommendations when creating collection estimates

- Consider adding custodians and non-custodial data sources to the case before you create a collection estimate. Adding these sources allows you to quickly select the data sources when you create a collection estimate in the workflow. If you choose to skip adding custodians or non-custodial data sources, you can search for these sources and any additional data sources (ones that haven't been added to the case as custodial or non-custodial locations) in a collection estimate for content that may be relevant to the case.

These data sources might include mailboxes, SharePoint sites, and Teams. If this situation is applicable to your case, use additional locations to add these sources to your collection.

- Verify that you have permissions to create a collection estimate for a case.
- When configuring a new collection, you can save your progress in the workflow as a draft collection and complete it later if needed. Just select **Save and close** on any page in the collection workflow to save your progress as a draft. The collection is added to the collection list with a status of *Draft*. To resume working on the collection, select the draft collection, select **Actions**, then select **Edit collection**.

Create a collection estimate

1. In the Microsoft Purview portal, navigate to **eDiscovery** > **Premium**.
2. Select the **Cases** tab and open an eDiscovery (Premium) case, and then select the **Collections** tab.
3. On the **Collections** page, select **New collection**.
4. On the **Name and description** page, enter a name (required) and description (optional) for the collection. After the collection is created, you can't change the name, but you can modify the description.
5. On the **Custodial data sources** page, select one of the following options to identify the custodial data sources to collect content from:
    - Select **Select custodians** to search specific custodians that were added to the case. If you use this option, a list of the case custodians is displayed. Select one or more custodians. After you select and add the custodians, you can also select the specific data sources to search for each custodian. Data sources that are displayed are specified when the custodian is added to the case.
    - Select the **Select all** toggle to search all custodians that were added to the case. When you select this option, all data sources for all custodians are searched.
6. On the **Non-custodial data sources** page, select one of the following options to identify the non-custodial data sources to collect content from:
    - Select **Select non-custodial data sources** to select specific non-custodial data sources that were added to the case. If you use this option, a list of data sources displayed. Select one or more of these data sources.

- Select the **Select all** toggle to select all non-custodial data sources that were added to the case.

7. On the **Additional locations** page, you can select other mailboxes and sites to search as part of the collection. You can also choose additional search options for additional search locations (defined in eDiscovery settings) to include in the collection.

   Use the **Additional locations** options to choose the types of data sources weren't added as custodial or non-custodial data locations in the case. You also have two options when searching additional data sources:

   - To search all content locations for a specific service (Exchange mailboxes, SharePoint and OneDrive sites, or Exchange public folders), select the corresponding **Select all** toggle in the **Status** column. This option will search all content locations in the selected service.

8. On the **Search query** page, you can create the search query that is used to collect items from the data sources that you've identified in the previous workflow pages. You can search for keywords, property:value pairs, or use a keyword list. You can also add various search conditions to narrow the scope of the collection or user the query builder to define query operators, filters, and conditions. For more information, see:

   - Build search queries using keywords and conditions for collections
   - Use the KeyQL editor to build search queries
   - Use the query builder to build search queries
   - Use the Natural Language Query builder to build search queries (preview)

   Use the **Additional search options** options to include other search locations. The options available here are set by your admin in eDiscovery global settings. Search options that aren't defined by your eDiscovery admin aren't selectable or available. Search options enabled by your eDiscovery admin can be selected and included in the collection as applicable. You have two choices for additional search options:

   - **Guest mailboxes**: Select to include guest mailboxes during tenant-wide searches.
   - **Shared Teams channels**: Select to include shared channels during tenant-wide searches.

9. On the **Conditions** page, you can create the search query that is used to collect items from the data sources that you've identified in the previous workflow pages. You can search for keywords, property-value pairs, or use a keyword list. You can also add various search

conditions to narrow the scope of the collection. For more information, see Build search queries for collections.

10. On the **Review your collection and get estimates** page, you'll see a summary of the collection estimate process. The collection estimate is saved for further review and refinement, but the results won't be automatically committed to a review set. You'll need to decide to commit to a review set later.

11. On the **Review your collection** page, you can review and update the collection selection, data sources, and conditions that you configured on the previous pages. Select **Edit** for any of the selections and update if needed.

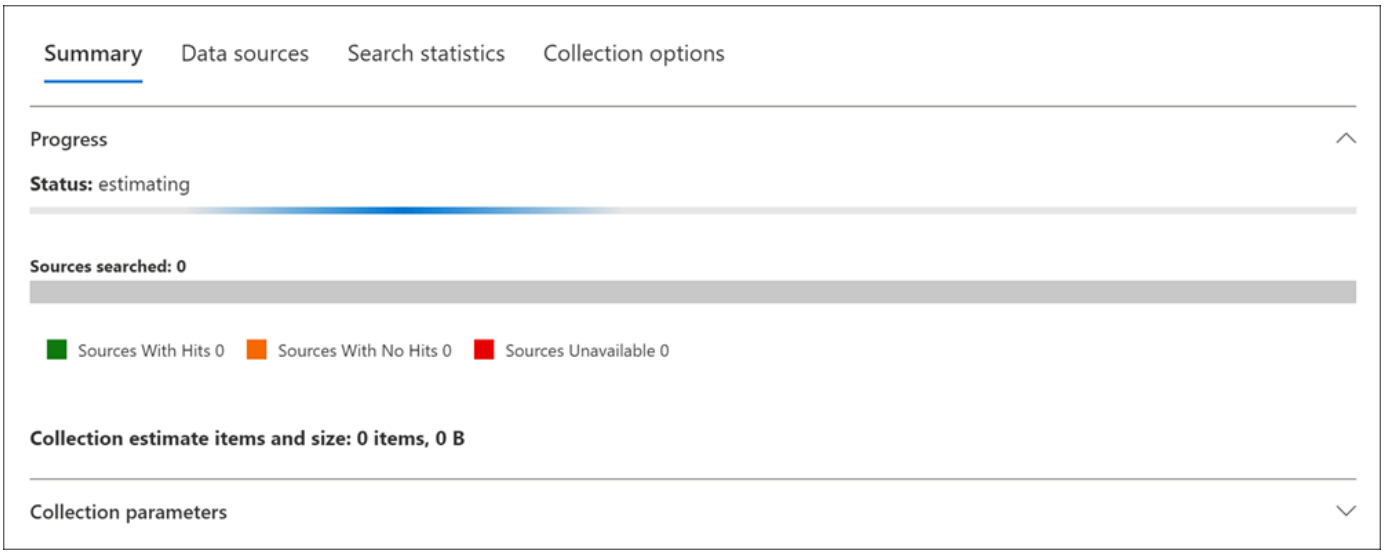
12. Select **Submit** to create the collection estimate. A page is displayed confirming that the collection was created.

What happens after you create a collection estimate

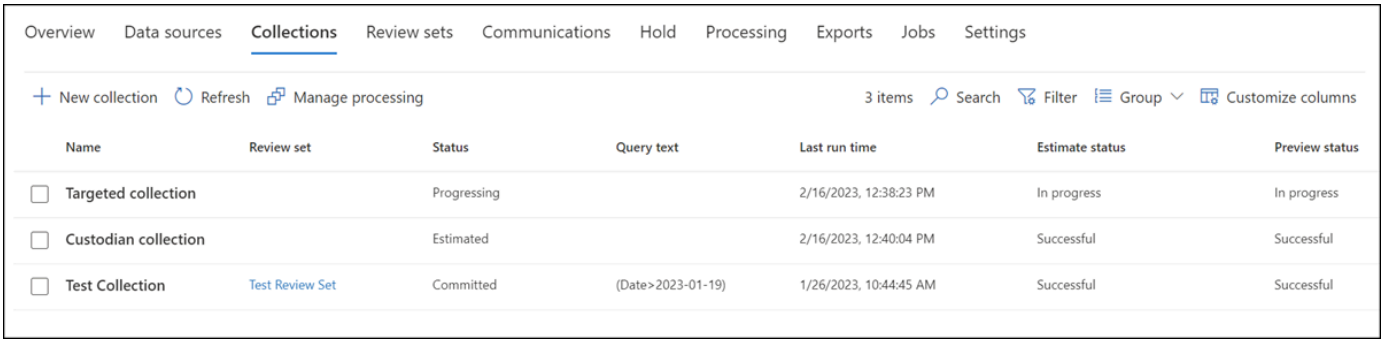
After you create a collection estimate, it listed on the **Collections** page in the case and the status shows that it's in progress. A job named **Preparing search preview and estimates** is also created and displayed on the **Jobs** page in the case. During the collection estimate process, eDiscovery (Premium) performs a search estimate using the search criteria and data sources that you specified in the collection. eDiscovery (Premium) also prepares a sampling of items that you can preview.

Additionally, you can now monitor the progress of the estimate that shows how many sources and items have been completed when building the estimate. As part of the estimate progress, you can see statistics for the following areas:

- **Status**: The status of the estimate. Values are *In progress*, *Successful*, or *UNSUCCESSFUL VALUE?*
- **Sources searched**: The number of sources searched in the estimate.
  - **Sources with Hits**: Sources in the estimate that meet the conditions configured for the collection.
  - **Sources with No Hits**: Sources in the estimate that don't have any items that meet the conditions configured for the collection.
  - **Sources Unavailable**: Sources that were unavailable during the period that the estimate was in progress.
- **Collection estimate items and size**: The number of items and size of the collection estimate.



When the collection is complete, the following columns and corresponding values on the **Collection** page are updated:



- **Status**: Indicates the status and type of collection. A value of **Estimated** indicates that a collection estimate is complete. This same value also indicates that the collection is a collection estimate, and that it hasn't been added to a review set. A value of **Committed** in the **Status** column indicates that the collection has been added to a review set.
- **Estimate status**: Indicates the status of the estimated search results and whether or not the search estimates and statistics are ready for review. A value of **Successful** indicates the results of the collection estimate are ready for review. After you first submit a collection estimate, a value of **In progress** is displayed to indicate the collection is still running
- **Preview status**: Indicates the status of the sample items that you can preview. A value of **Successful** indicates the items are ready for preview. After you first submit a collection estimate, a value of **In progress** is displayed to indicate that the collection estimate is still running.

Next steps after a collection estimate is complete

After the collection estimate is successfully completed, you can perform various tasks. To perform most of these tasks, go the **Collections** tab and select the name of the collection estimate to

display the flyout page.

Summary | Data sources | Search statistics | Collection options

Collection overview (preview)

Understand how the items were found, retrieved, and processed as part of this collection.

Locations with hits

We **searched 7,238 locations,** and found **5,815 locations with hits.**

7,309 mailboxes and 0 sites were searched.

1 locations failed to be searched during estimate.

Retry failed locations

Pre-collection estimates

2,549,828 items (~378.26 GB) had search hits.

11,600 items were partially indexed and weren't fully searched. You can add these items to a review set for inspection.
Learn more about partial indexing.

Collection parameters

Additional locations

Exchange

EXCHANGE_ALL

Created

2/8/2023, 5:10:32 PM

Last modified

2/8/2023, 5:10:32 PM

Estimate

2,549,828 item(s) (378.26 GB)

11,600 unindexed item(s), 6.29 GB

5,815 mailboxes

0 sites



> **Note**
>
> You may see differences between the number of *locations with hits* and the total number of Exchange mailboxes, SharePoint and OneDrive sites, or Exchange public folders searched. Empty archived mailboxes are searched, but aren't included in the *location with hits* total.

Here's a list of things you can do from the collection flyout page:

- Select the **Summary** tab to view summary information about the collection and the estimated search results returned by the collection. This includes:
  - **Collection overview** section: This section includes the **Locations with hits** subsection that provides data on the total number of locations searched and the number of locations that contained items resulting in hits. For example, a collection scoped to Exchange Online mailboxes and SharePoint sites may summarize results by the number of mailboxes and sites searched, and the combined number of mailboxes and sites that had hits. This section also includes the **Pre-collection estimate** subsection that displays the total number of items and size of the estimated search results. You have the option to add these items to the review set for inspection.
  - **Collection parameters** section: This section includes information about additional locations, when the collection was created and modified, and a summary of the estimate items, size, and any unindexed items and size.

- Select the **Data sources** tab to view a list of custodians and non-custodial data sources) that were searched in the collection. Any additional content locations that were search are listed under **Locations** on the **Summary** tab.

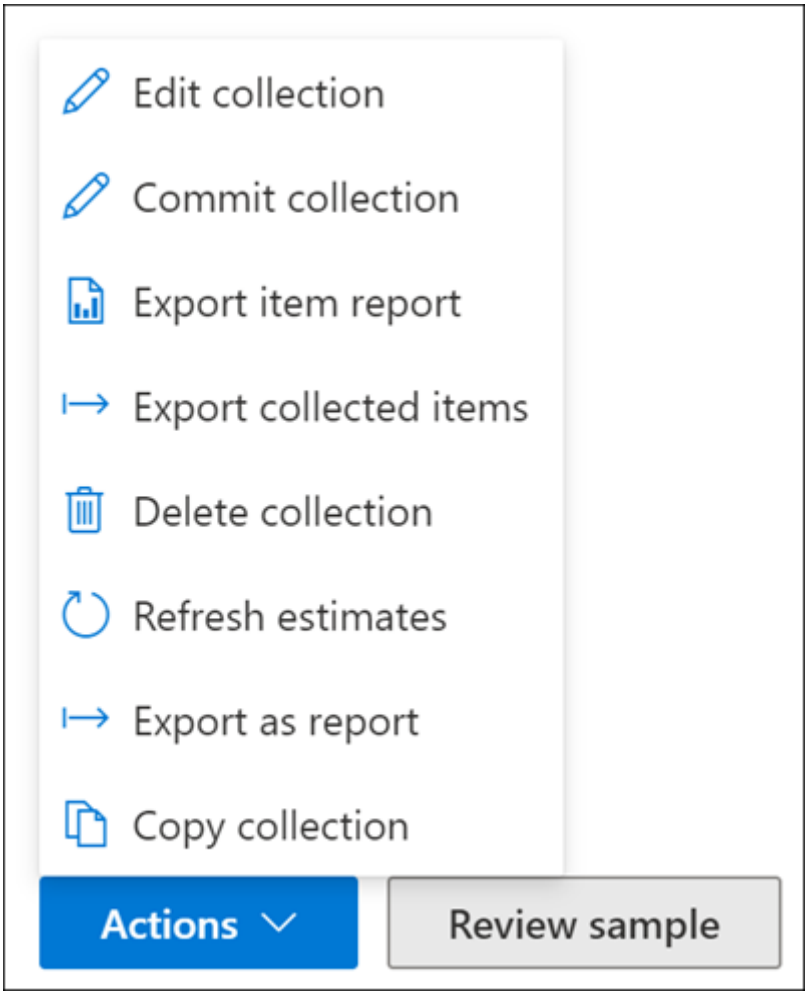
- Select the **Search statistics** tab to view statistics about the collection. This includes the total number and size of items found in each service (for example, Exchange mailboxes or SharePoint sites) and a condition report that displays statistics about the number of items returned by different components of the search query used by the collection. For more information, see Collection statistics and reports.

- Select the **Collection options** tab to view information about the included collection options. This includes a summary of the additional locations and retrieval options chosen in the collection workflow.

- Select **Review sample** (located at the bottom of the flyout page) to preview a sample of the items returned by the collection.
- Take actions on the collection estimate by selecting choices in the **Actions** menu. For more information, see the following section about the available management options.

Manage a collection estimate

You can use the options in the **Actions** menu on the flyout page of a collection estimate to perform various management tasks.



Here's the descriptions of the pre-collection estimate management options.

- **Edit collection**: Change the settings of the collection estimate. After you make changes, you can rerun the collection and update the search estimates and statistics. As previously explained, you use this option to commit a collection estimate to a review set.
- **Commit collection**: Commit a collection to a review set. This means that you rerun the collection (using the current settings) and add the items returned by the collection to a review set. As previously explained, you can also configure additional settings (such as conversation threading and cloud-based attachments) when you add the collection to a

review set. For more information and step-by-step instructions, see Commit a collection estimate to a review set.

- **Export item report**: Similar to the exporting items in Content search, you can choose this option to export the results of the report that is based on the actual items that can be retrieved from the source. After selecting, you have the following export options for collected items:

  - **Types of collected items to include in the export**: Choose to export collected items with search hits, items with search hits and partially indexed items without hits, or only partially indexed items without search hits. You can also choose to one or more of the following options for collected items:
    - Include Microsoft Teams and Viva Engage conversations
    - Include cloud attachments
    - Include all existing versions of Microsoft 365 documents on SharePoint
    - Include subfolder contents (insider subfolders of a matched folder)
    - Include files in SharePoint lists (and their child items)

- **Export collected items**: Export the collected items without adding the items to the review set. This option is useful in scenarios where data residency requirements associated with data storage may be prohibitive and you need collected data as a download. After selecting, you have the following export options for collected items:

  - **Types of collected items to include in the export**: Choose to export collected items with search hits, items with search hits and partially indexed items without hits, or only partially indexed items without search hits. You can also choose to one or more of the following options for collected items:
    - Include Microsoft Teams and Viva Engage conversations
    - Include cloud attachments
    - Include all existing versions of Microsoft 365 documents on SharePoint
    - Include subfolder contents (inside subfolders of a matched folder)
    - Include files in SharePoint lists (and their child items)

    To export Teams messages as individual messages when exporting collected items, unselect **Teams and Viva Engage conversations**. This exports conversations as individual messages and not as conversation threads.

    > ⓘ **Note**

The **Teams and Viva Engage conversations** setting is selected by default. If selected, exported items contain Teams conversation threads related to the matched search.

  - **How to format emails**: Choose an option of how collected emails should be formatted:
    - Individual .pst files for each mailbox
    - Individual .msg files for each message
    - Individual .eml files for each message

- **Delete collection**: Delete a collection estimate. After you commit a collection estimate to a review set, it can't be deleted.

- **Refresh estimates**: Rerun the query (against the data sources) specified in the collection estimate to update the search estimates and statistics.

- **Export as report**: Exports information about the collection estimate to a CSV file that you can download to your local computer. The export report contains the following information:
  - The identity of each content location that contains items that match the search query in the collection estimate. These locations are typically mailboxes or sites.
  - The total number of items in each content location.
  - The total size (in bytes) of the items in each content location.
  - The service (such as Exchange or SharePoint) in which the content location is located.

  For more information about the limits for **Export as report**, see Collection limits.

- **Copy collection**: Create a new collection estimate by copying the settings from an existing collection. You have to use a different name for the new collection. You also have the option to modify the settings before you submit the new collection. After you submit it, the search query runs and new estimates and statistics are generated. This is a good way to quickly create additional collection estimate and then modify selected settings as necessary while still preserving information in the original collection. This also lets you easily compare the results of two similar collections.

> **Note**
>
> After a collection estimate is committed to a review set, you can only select **Copy collection** and **Export as report**.

Last updated on 10/31/2024

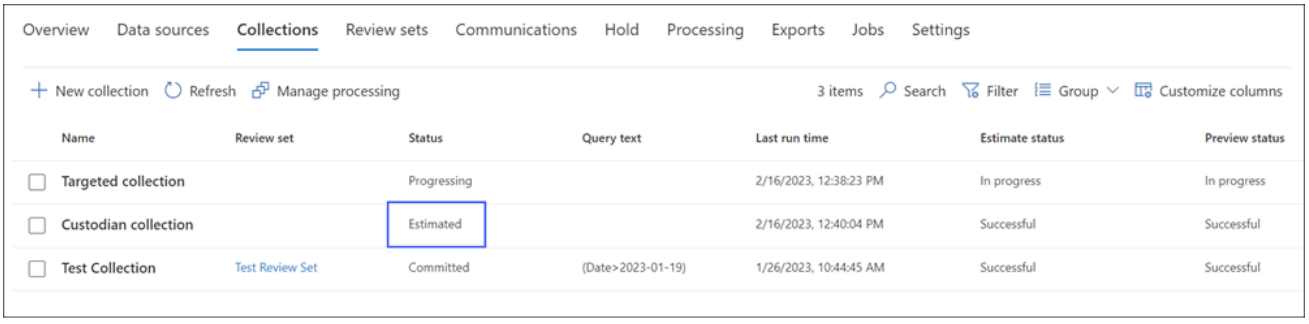
Commit a collection estimate to a review set in eDiscovery (Premium)

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

When you're satisfied with the items you've collected in a collection estimate and are ready to analyze, tag, and review them, you can commit a collection to a review set in the case. When you commit a collection estimate to a review set, collected items are copied from their original content location in Microsoft 365 services and added to a review set. A review set is a secure, Microsoft-provided Azure Storage location in the Microsoft cloud.

Commit a collection estimate to a review set

1. In the Microsoft Purview portal, navigate to **eDiscovery** > **Premium**.
2. Select the **Cases** tab and open an eDiscovery (Premium) case, and then select the **Collections** tab.



> **Tip**

A value of `Estimated` in the **Status** column identifies the collection estimates that can be added to a review set. A status of `Committed` indicates that a collection has already been added to a review set.

3. On the **Collections** page, select the collection estimate that you want to commit to a review set.

4. On the bottom of the flyout page, select **Commit collection**.

5. Configure the following settings:

   a. Decide whether to **Add to a new review set** (which is created after you submit the collection) or **Add to an existing review set**. Complete this section based on your decision.

   b. Configure the **Retrieval** settings:

      A. **Teams and Viva Engage conversations**: Select this option to add conversation threads to the collection that include the chat items returned by the search query in the collection. This means that the chat conversation that contains items that match the search criteria is reconstructed. This lets you review chat items in the context of the back and forth conversation. Collect up to 12 hours of related conversations when a message matches a search. For more information, see Conversation threading in eDiscovery (Premium).

      To export Teams messages as individual messages when committing to review set, unselect **Teams and Viva Engage conversations**. This commits your review set with individual Teams messages. Once the collection is committed, you can export the documents and messages in the review set. The exported results contain the individual messages for Teams and Viva Engage instead of conversation threads.

> ⓘ **Note**
>
> The **Teams and Viva Engage conversations** setting is selected by default. If selected, review sets contain Teams conversation threads related to the matched search.

- **Cloud attachments**: Select this option to include modern attachments or linked files when the collection results are added to the review set. This means the target file of a modern attachment or linked file is added to the review set.

- **All document versions**: Select this option to enable the collection of all versions of a SharePoint document per the version limits and search parameters of the collection. Selecting this option will significantly increase the size of items that are added to the review set.

- **Partially indexed items**: Select this option to add partially indexed items from additional data sources to the review set. If the collection searched additional data sources (as specified on the **Additional locations** page in the collections workflow), there may be partially indexed items from these locations that you want to add to the review set. Custodial and non-custodial data sources typically don't have partially indexed items. That's because the Advanced indexing process reindexes items when custodial and non-custodial data sources are added to a case. Also, Adding partially indexed items will increase the number of items added to the review set.

> ⓘ **Note**
>
> The options to select retrieval options are configured on the global eDiscovery **Settings** > **Collections** page. Retrieval options selected on the **Collections** setting page will be selected as the default retrieval settings for all collections.

- **Subfolder contents**: Collect items inside subfolders of a matched folder. This option applies only to SharePoint and OneDrive data sources.
- **List attachments**: Collect files attached to SharePoint lists and their child items.

6. Configure settings in the **Collection ingestion scale** section to define the scale of the collection to add to the review set:

   - **Add all of collection to review set**: Select this option to add all the items that match the search criteria of the collection to the review set.

   - **Add only collection sample to review set**: Select this option to add a sample of the collection results to the review set instead of adding all results. If you select this option, select **Edit sample parameters** and choose one of the following options:
     - **Sample based on confidence**: Items from the collection are added to the review set will be determined by the statistical parameters that you set. If you typically use a confidence level and confidence interval when sampling results, specify them in the drop-down boxes. Otherwise, use the default settings.
     - **Randomly sample**: Items from the collection are added to the review set based on a random selection of the specified sample size percentage of the total number of

items returned by the search.

7. Select **Commit** to commit the collection to the review set.

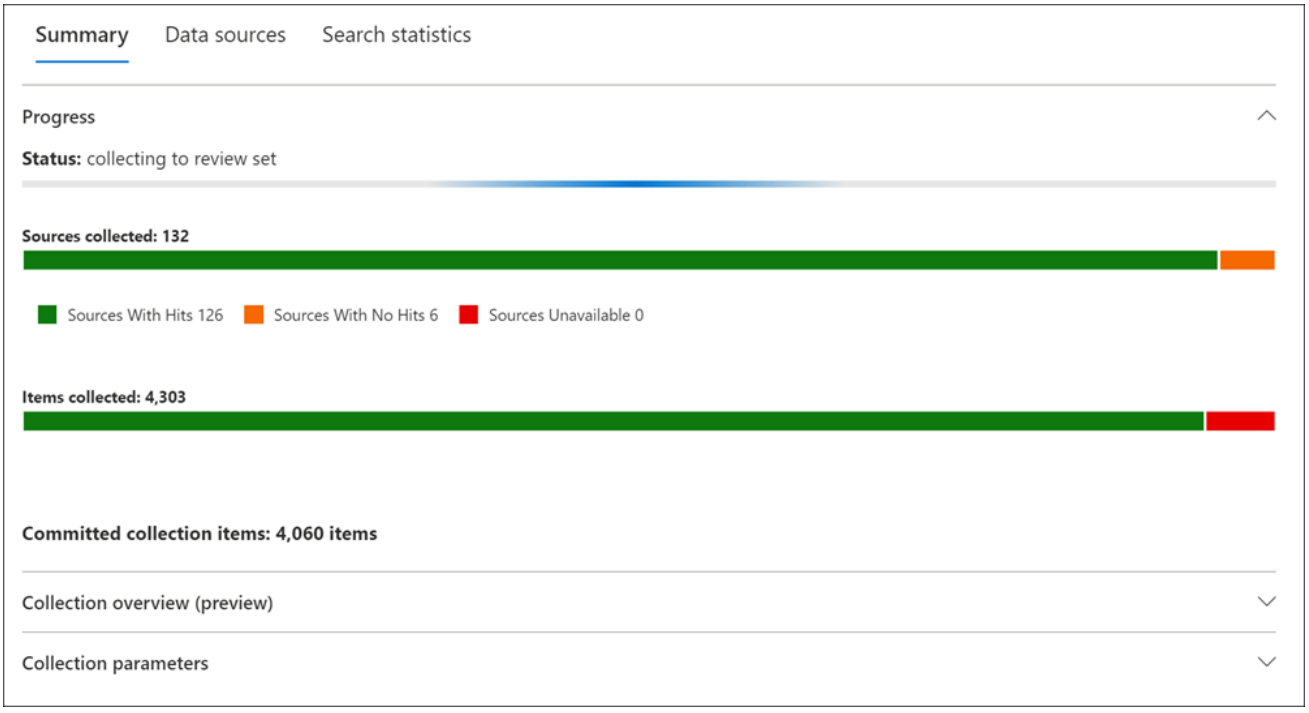
What happens after you commit a collection estimate

When you commit a collection estimate to a review set, the following things happen:

- If you created a new review set to commit the collection to, the review set is created and displayed on the **Review sets** tab in the case. This status value means the review set has been created; it doesn't mean that the collection has been added to the review set. The status of adding items in the collection to the review set is displayed on the **Collections** tab and in the **Summary** tab on the collection flyout page.

- The collection search query is run again. This means the actual search results copied to the review set may be different than the estimated results that were returned when the collection search was last run.

- The progress of the collection commitment is displayed on the collection estimate flyout page. From here, you can monitor the progress and status of the collection, including information about sources and items collected. During the collection process, the status of the collection is shown as *Adding to review set* on the **Collection** tab.



.

- All items in the search results are copied from the original data source in the live service, and copied to a secure Azure Storage location in the Microsoft cloud.

- Encrypted SharePoint and OneDrive documents and encrypted files attached email messages that's returned in the search results are decrypted when you commit the collection to a review set. You can review and query the decrypted files in the review set. For more information, see Decryption in Microsoft 365 eDiscovery tools.

- Optical character recognition (OCR) functionality extracts text from images, and includes the image text with the content that's added to a review set. For more information, see the Optical character recognition section in this article.

- After the commit is successfully completed, the value of the status column of on the **Collections** tab is changed to *Committed* and the **Summary** tab on the collection flyout page is updated with the full results of the committed collection. A link to the review set for the collection is added to the **Review set** column for the collection. Select this link to open the review set results.

Optical character recognition

When you commit a collection to a review set, optical character recognition (OCR) functionality in eDiscovery (Premium) automatically extracts text from images, and includes the image text with the content that's added to a review set. You can view the extracted text in the Text viewer of the selected image file in the review set. This lets you conduct further review and analysis on text in images. OCR supports loose files, email attachments, and embedded images. For a list of OCR-supported image file formats, see Supported file types in eDiscovery (Premium).

You have to enable OCR functionality for each case that you create in eDiscovery (Premium). For more information, see Configure search and analytics settings.

Last updated on 11/21/2025

Collection statistics and reports in Microsoft Purview eDiscovery (Premium)

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

After you create a collection estimate, you can view statistics on the retrieved items, such as the content locations that contain the most items that matched the search criteria and the number of items returned by the search query. You can also preview a subset of the results.

When you've identified the set of documents you want to further examine, you can add the search results to a review set to collect and process.

Search statistics for collection estimates

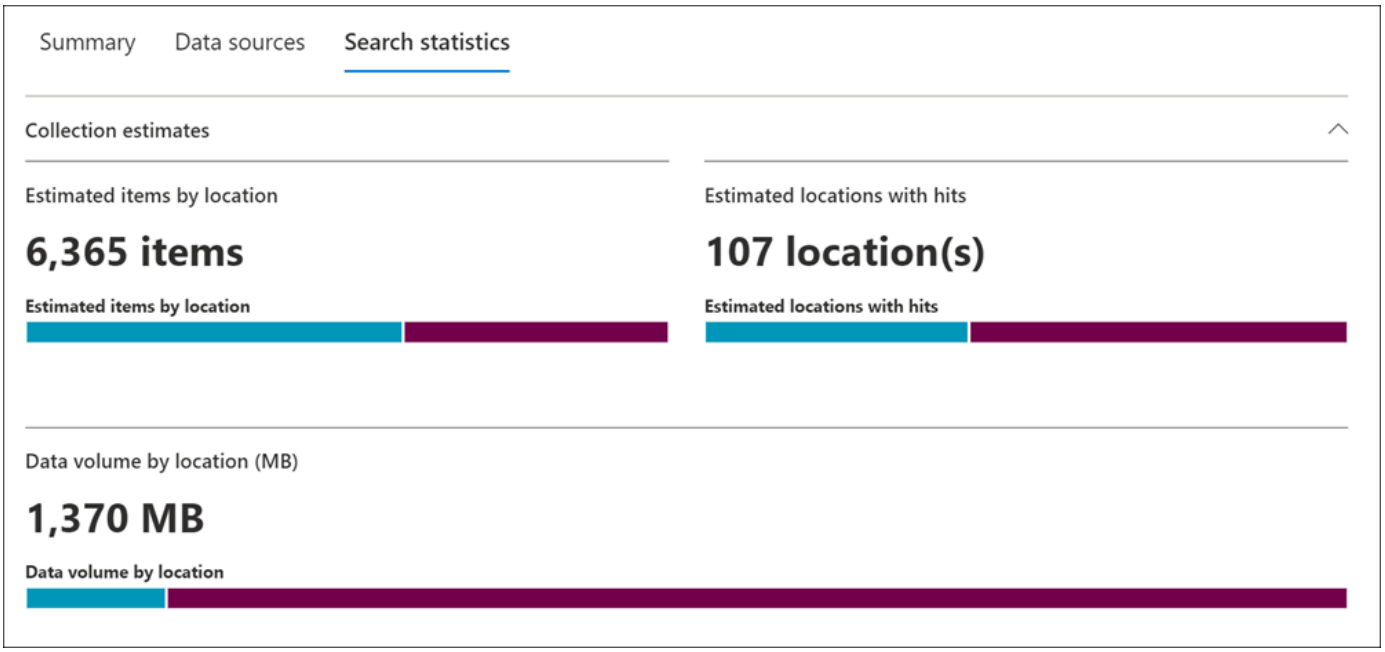
This section describes the statistics that are available for collection estimates. These statistics are available on the **Search statistics** tab on the flyout page of a collection estimate.

Collection estimates

This section displays a graphical summary of the estimated items returned by the collection. This indicates the number of items that match the search criteria of the collection. This information gives you an idea about the estimated number of items returned by the collection.

- **Estimated items by locations**: The total number of estimated items returned by the collection. The specific number of items located in mailboxes and located in sites is also displayed.
- **Estimated locations with hits**: The total number of content locations that contain items returned by the collection. The specific number of mailbox and site locations is also displayed.

- **Data volume by location (in MB)**: The total size of all estimated items returned by the collection. The specific size of mailbox items and site items is also displayed.

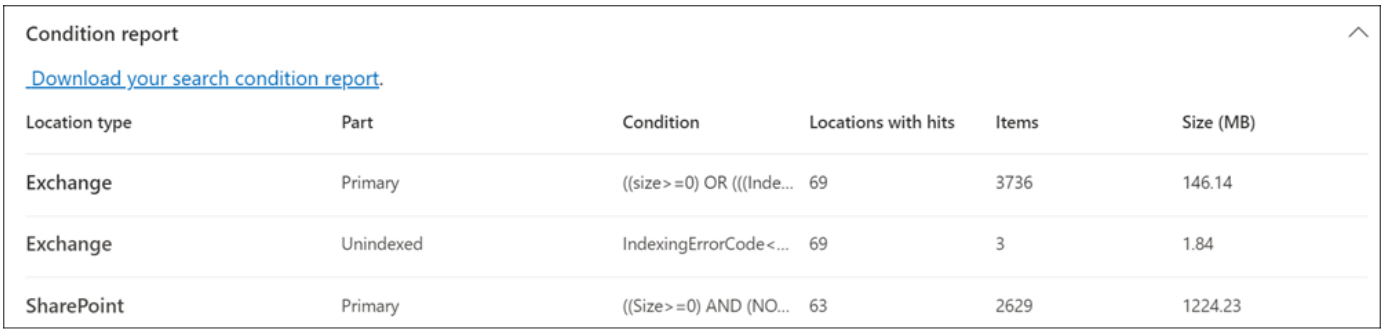


Condition report

This section displays statistics about the collection search query and the number of estimated items that matched different parts of the search query. You can use these statistics to analyze the number of items that match each component of search query. This can help you refine the search criteria for the collection and if necessary narrow the scope of the collection.

- **Location type**: The type of content location that the query statistics are applicable to. The value of **Exchange** indicates a mailbox location; a value of **SharePoint** indicates a site location.
- **Part**: The part of the search query the statistics are applicable to. **Primary** indicates the entire search query. **Keyword** indicates the statistics in the row are for a specific keyword. If you use a keyword list when for the search query in the collection, statistics for each component of the query are included in this table.
- **Condition**: The actual component (keyword or condition) of the search query that was run for the collection estimate that returned the statistics displayed in the corresponding row.
- **Locations with hits**: The number of the content locations (specified by the **Location type** column) that contain items that match the primary or keyword query listed in the **Condition** column.
- **Items**: The number of items (from the specified content location) that match the query listed in the **Condition** column. As previously explained, if an item contains multiple instances of a keyword that is being searched for, it's only counted once in this column.

- **Size (MB)**: The total size of all items that were found (in the specified content location) that match the search query in the **Condition** column.


Condition report

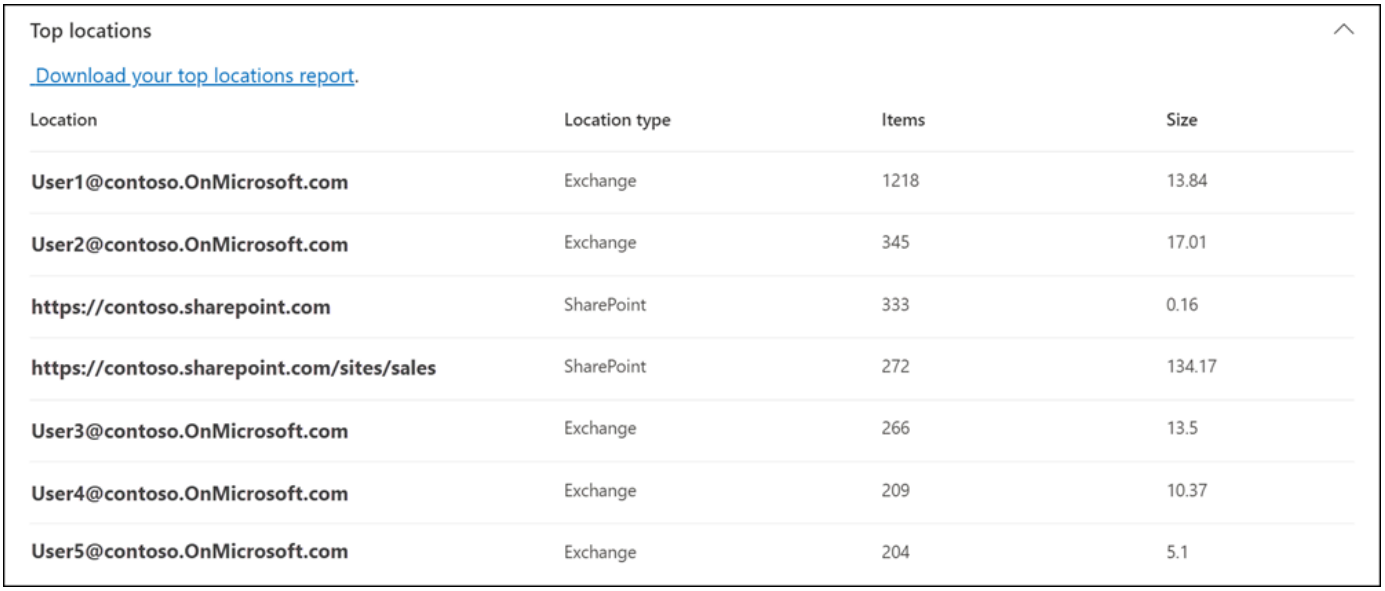
Download your search condition report.

Location type	Part	Condition	Locations with hits	Items	Size (MB)
Exchange	Primary	((size>=0) OR (((Inde...	69	3736	146.14
Exchange	Unindexed	IndexingErrorCode<...	69	3	1.84
SharePoint	Primary	((Size>=0) AND (NO...	63	2629	1224.23

Top locations

This section displays statistics about the specific content locations with the most items returned by the collection.

- **Location**: The name of the location (the email address of mailboxes and the URL for sites).
- **Location type**: A mailbox will have an *Exchange* type and SharePoint sites will have a *SharePoint* type.
- **Items**: Estimated number of items in the content location returned by the collection.
- **Size**: The total size of estimated items in each content location.


Top locations

Download your top locations report.

Location	Location type	Items	Size
User1@contoso.OnMicrosoft.com	Exchange	1218	13.84
User2@contoso.OnMicrosoft.com	Exchange	345	17.01
https://contoso.sharepoint.com	SharePoint	333	0.16
https://contoso.sharepoint.com/sites/sales	SharePoint	272	134.17
User3@contoso.OnMicrosoft.com	Exchange	266	13.5
User4@contoso.OnMicrosoft.com	Exchange	209	10.37
User5@contoso.OnMicrosoft.com	Exchange	204	5.1

Statistics and reports for committed collections

This section describes the statistics that are available after you commit a collection to a review set, including the actual number of items added to the review set. These statistics (in addition to load set information) provide historical information about content added to a case.

After you commit a collection to a review set, the following tabs are displayed on the flyout page of the committed connection. Each of these tabs contains different types of information about the collection.



Summary tab for collections

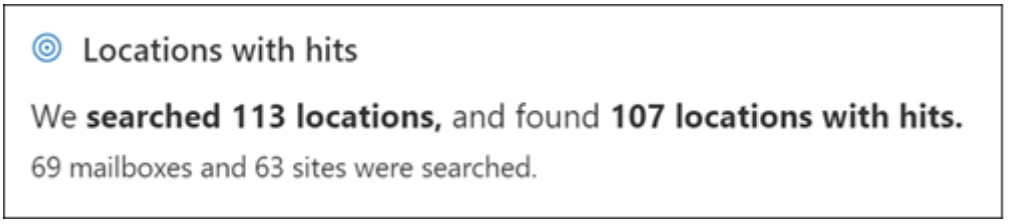
The **Summary** tab for committed collections contains several sections that provide information about the collection items, item estimates and locations, indexing, file types, and more.

Collection overview (preview)

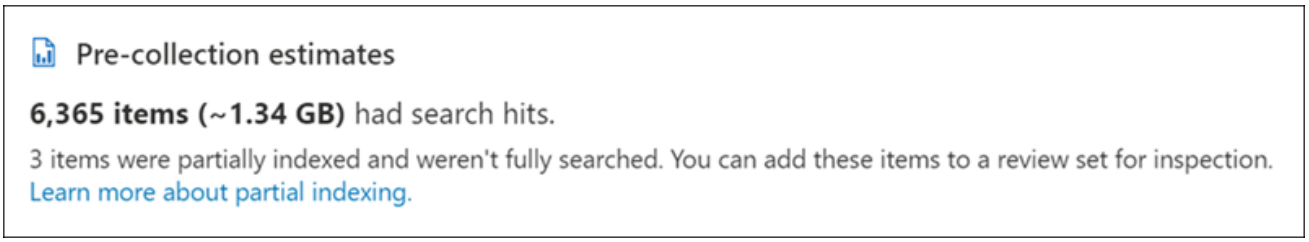
The **Collection overview** section on the **Summary** tab of a committed review set contains information about how items were found, retrieved, and processed as part of the collection. The information in this section is a summary of the information gathered and displayed during the initial collection estimate and includes information gathered and processed during the collection committal to the review set. You can use the information in this section to help improve understanding on how the final collection data evolves from the estimate data during collection committal and processing.

This information is segmented and displayed in the following sub-sections:

- **Locations with hits**: This section summarizes the number of locations searched and the number of location with hits based on the conditions defined in the collection estimate. It also displays the number and type of locations searched.



- **Pre-collection estimates**: This section summarizes number of items and size of the total items calculated for the collection estimate (before the collection is committed). A count of any items that were partially indexed is also displayed.

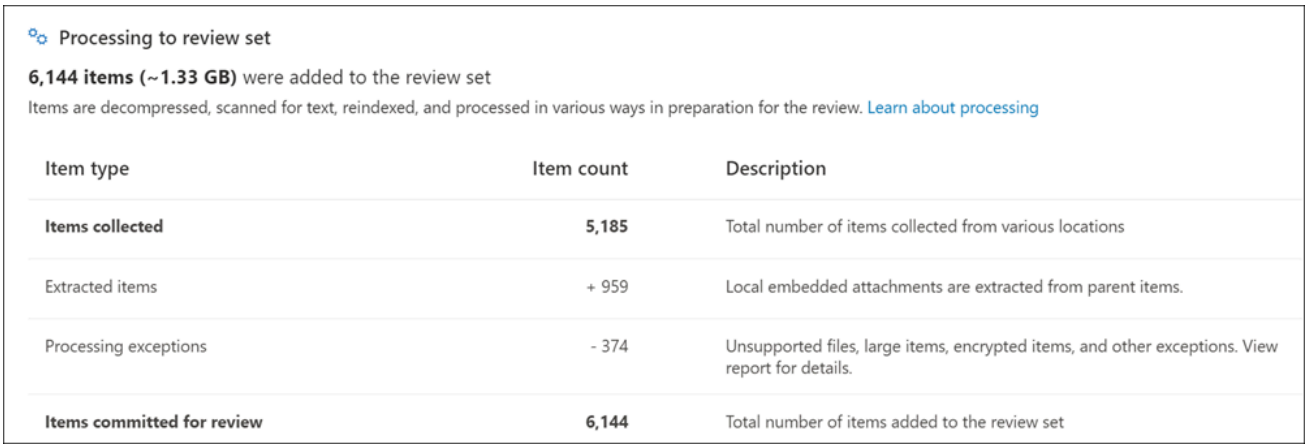


- **Collection items**: This section displays the type, count, and a description of all items retrieved from locations with hits during the committal of the collection to the review set. During this process, there may be some addition and consolidation of items during processing. Additions and consolidations may include types included in the following table:

Expand table

Item type	Item type description
Items with hits	Estimated number of items found by search.
Review set duplicates	Items that are already in the same review set aren't collected.
Search duplicates	Duplicate instances of the same items aren't collected.
All versions	All versions of items in SharePoint are collected.
Historical versions	Historical versions maintained by SharePoint are collected.
Cloud attachments	Cloud attachments are identified and collected from SharePoint links.
Version as shared	The shared versions of items are collected in addition to their latest versions.
Contextual Teams conversations	Additional chat messages that might be relevant are collected.
Consolidated chat transcripts	Collected chat messages are consolidated into fewer HTML transcripts.
Retrieval exceptions	Access time outs, empty files, and other exceptions.
Total items collected	Total number of items collected from various locations.

- **Processing to review set**: This section displays the type, count, and description of the items that were added to the review set from the collection estimate. Items are decompressed, scanned for text, reindexed, and processed in preparation for the review.



Collection contents

The **Collection contents** section on the **Summary** tab of a committed review set contains statistics and other information about the items that were collected from the data sources in the collection and added to the review set.

- **Total extracted items**. The total number of items added to the review set. This number indicates the sum of parent items and child items added to the review set.

  > **Tip**
  >
  > Hover the cursor over the parent or child item bars to display the total number of parent or child items.

- **Parent items**. The number of items returned by the collection that was used to collect the items that were added to the review set. This number corresponds to (and is equal to) the estimated number of items that is displayed in the **Collection parameters** section. The number of parent items he collection information that was used to collect the items that were added to the review set.

  A parent item might contain multiple child items. For example, an email message is a parent item if it contains an attached file or has a cloud attachment. In this case, the attached file or the target file of the cloud attachment is considered a child item. When you commit a collection, parent items and any corresponding child items (like attached files and cloud attachments) are added to the review set as individual items or files.

- **Child items**. The number of child items added to the review set. Only child items that are file attachments and cloud attachments are added to the review set as individual files. Other types of child items, such as email signatures and images, are extracted from a parent item and then processed by Optical Character Recognition (OCR) to extract any text from the

child item. Text extracted from these types of child items is then added to its parent item so you can view it in the review set. By not adding child items to the review set as a separate file, eDiscovery (Premium) helps streamline the review process by limiting the number of potentially immaterial items in the review set.

- **Unique items**. The number of unique items added to the review set. Unique items are unique to the review set. All items are unique when the first collection is added to a new review set because there were no previous items in the review set.

- **Identified duplicate items**. The number of items from the collection that weren't added to the review set because the same item already exists in the review set. Statistics about duplicate items can help explain the differences between the number of estimated items from a collection estimate and the actual number of items added to the review set.

Indexing

The **Indexing** section on the **Summary** tab of a committed review set contains indexing information about the items added to the review set.

- **New indexed items**. The number of items that were newly indexed before they were added to the review set. Examples of a newly indexed item are child items extracted from a parent item and then indexed before they're added to the review set. Also, items that aren't located in custodial data sources and non-custodial content locations listed on the **Data sources** tab in the case are indexed before they're added to the review. For example, newly indexed items would include items collected from additional locations.

- **Updated indexed items**. The number of partially indexed items that were successfully indexed and added to the review set. This statistic indicates the partially indexed items from custodial and non-custodial content locations **Data sources** tab that were successfully indexed when the collection was committed to the review set.

- **Indexing errors**. The number of partially indexed items that couldn't be indexed before they were added to the review set. These items might require error remediation.

Collection file types

The **Collection file types** section on the **Summary** tab of a committed review set displays the item count for each file type included in the collection. File types include email, image, document, audio, video file types and more. For a complete list of the supported file types, see Supported file types in eDiscovery (Premium).

Errors

The **Errors** section on the **Summary** tab of a committed review set displays the type and number of errors encountered during the collection committal. Errors may include files with a zero size, protected files, malformed files, internal errors, and more. An error report in .csv format can be downloaded by selecting **Download your errors report** in this section.

Collection parameters

The **Collection parameters** section on the **Summary** tab of a committed review set displays the collection information that was used to collect the items that were added to the review set. This tab displays information that is similar to the information on the **Search statistics** tab. This section provides a quick snap shot of the search query used by the collection, the content locations that were searched, and the estimated collection results. As previously explained, the number of estimated items in this section would be equal to the number of parent items shown in the **Collection contents** section.

Data sources tab for collections

The **Data sources** tab contains information about custodial and non-custodial data sources. The **Custodial data sources** section displays the all custodial data sources for each custodian included in the collection. The **Non-custodial data sources** section displays the all non-custodial data sources for the collection. The data source information displayed is also available on the main **Data sources** tab for each case.

Search statistics tab for collections

The statistics displayed on the **Search statistics** tab are the same statistics from the last time that a collection estimate was run. This includes collection estimates, condition report, and top locations. This information is preserved from the collection estimate for historical reference, and can be compared to the actual collection that was committed to the review set.

Differences between collection estimates and the actual committed collection

When you run a collection estimate, an estimate of the number of items (and their total size) that meet the collection criteria is displayed on the **Summary** tab and in **Collection estimates** section of the **Search statistics** tab. After you commit a collection estimate to a review set, the actual number of items (and their total size) added the review set are often different from the estimates.

In most cases, more items are added to the review set than were estimated from the collection estimate. The following list describes the most common reasons for these differences and tips for identifying them:

- **Child items**. Child items (such as files attachments and cloud attachments) that are extracted from their parent items and added as individual files. The number of child items may increase the number of items that are added to the review set. In general, the number of parent items identified in the **Collection contents** section on the **Summary** tab of a committed collection should be equal to the number of estimated items from the collection estimate.

- **Duplicate items**. Items from the collection estimate that have already been added to the review set in a previous collection won't be added. As previously explained, the number of duplicate items in the collection is displayed in the **Collection contents** section on the **Summary** tab.

- **Collection configuration options**. When you commit a collection estimate to a review set, you have to option to include conversation threads, cloud attachments, and document versions. Any of these items that are added to the review set aren't included in the estimates of the collection estimate. They're identified and collected only when you commit the collection. Selecting these options will most likely increase the number of items added to the review set.

  For example, multiple versions of SharePoint documents aren't included in the estimate for the collection estimate. But if you select the option to include all document versions when you commit a collection estimate, the actual number (and total size) of items added to the review set will increase.

  For more information about these options, see Commit a collection estimate to a review set.

Here are other reasons why the estimated results from a collection estimate can be different that the actual committed results.

- **The way results are estimated for collection estimates**. An estimate of the search results returned by a collection estimate is just that, an estimate (and not an actual count) of the items that meet the collection query criteria. To compile the estimate of email items, a list of the message IDs that meet the search criteria is requested from the Exchange database. But when you commit the collection to a review set, the collection is rerun and the actual messages are retrieved from the Exchange database. So differences might result because of how the estimated number of items and the actual number of items are determined.

- **Changes that happen between the time when estimating and committing collection estimates**. When you commit a collection estimate to a review set, the search is rerun to collect that most recent items in the search index that meet the search criteria. It's possible that additional items were created, sent, or deleted that meet the search criteria in the time between when the collection estimate was last run and when the collection estimate is committed to a review set. It's also possible that items that were in the search index when the collection estimate results were estimated are no longer there because they were purged from a data source before committing the collection. One way to mitigate this issue is to specify a date range for a collection. Another way is to place a hold on content locations so that items are preserved and can't be purged.

- **Unindexed items**. If the collection estimate included searching all Exchange mailboxes or all SharePoint sites, then only unindexed items from content locations that contain items that match the collection criteria will be added to the review set. In other words, if no results are found in a mailbox or site, then any unindexed items in that mailbox or site won't be added to the review set. However, unindexed items from all content locations (even those that don't contain items that match the collection query) will be included in the estimated collection results.

- **Partially indexed items**: Selection of this option adds partially indexed items from additional data sources to the review set. If the collection searched additional data sources (as specified on the **Additional locations** page in the collections workflow), there may be partially indexed items from these locations that you want to add to the review set. Custodial and non-custodial data sources typically don't have partially indexed items. That's because the Advanced indexing process reindexes items when custodial and non-custodial data sources are added to a case. Also, Adding partially indexed items will increase the number of items added to the review set.

  After partially indexed items are added to the review set, you can apply a filter to specifically view these items. For more information, see Scenario example: Filter partially indexed items

  Alternatively, if the collection estimate included specific content locations (which means that specific mailboxes or sites where specified on the **Additional locations** page in the collection estimate workflow), then unindexed items (that aren't excluded by the collection criteria) from the content locations specified in the search will be exported. In this case, the estimated number of unindexed items and the number of unindexed items that are added to the review set should be the same.

Last updated on 10/31/2024

Build search queries for collections

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

When configuring the search query when creating a collection in an eDiscovery (Premium) case, you can use keywords to find specific content and conditions to narrow the scope of the search to return items that are most relevant to your legal investigation.

If you prefer to build a search query with the Keyword Query Language (KeyQL) editor, Query builder, or the Natural Language Query tools, see:

- Use the KeyQL editor to build search queries
- Use the query builder to build search queries
- Use the Natural Language Query builder to build search queries (preview)

Keyword searches

Type a keyword query in the **Keywords** box in the search query. You can specify keywords, email message properties, such as sent and received dates, or document properties, such as file names or the date that a document was last changed. You can use more complex queries that use a Boolean operator, such as **AND**, **OR**, **NOT**, and **NEAR**. You can also search for sensitive information (such as social security numbers) in documents in SharePoint and OneDrive (not in email messages), or search for documents that have been shared externally. If you leave the **Keywords** box empty, all content located in the specified content locations is in the search results.

Keyword list

Alternatively, you can select the **Show keyword list** check box and the type a keyword or keyword phrase in each row. The keywords in each row are connected by a logical operator (which is represented as *c:s* in the search query syntax) that is similar in functionality to the **OR** operator in the search query that's created. This means items that contain any keyword in any row are in the search results. You can add up to 180 rows in the keyword list in eDiscovery (Premium) search queries.

Why use the keyword list? You can get statistics that show how many items match each keyword in the keyword list. This can help you quickly identify the keywords that are the most (and least) effective. You can also use a keyword phrase (surrounded by parentheses) in a row in the keywords list. For more information about search statistics, see Collection statistics and reports

Conditions

You can add search conditions to narrow the scope of a search and return a more refined set of results by selecting **Add condition**.

Each condition adds a clause to the search query that is created and run when you start the search. A condition is logically connected to the keyword query specified in the keyword box by a logical operator (which is represented as *c:c* in the search query syntax) that is similar in functionality to the **AND** operator. That means items have to satisfy both the keyword query and one or more conditions to be included in the search results. This is how conditions help to narrow your results.

For a list and description of conditions that you can use in a search query, see the "Search conditions" section in Keyword queries and search conditions.

Last updated on 11/21/2025

Use the KeyQL editor to build search queries

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

The Keyword Query Language (KeyQL) query option in Microsoft Purview eDiscovery tools search provides feedback and guidance when you build search queries in Content search, Microsoft Purview eDiscovery (Standard), and eDiscovery (Premium). When you enter queries in the editor, it provides autocompletion for supported searchable properties and conditions and provides lists of supported values for standard properties and conditions. For example, if you specify the `kind` email property in your query, the editor will present a list of supported values that you can select. The KeyQL editor also displays potential query errors in real time that you can fix before you run the search. Best of all, you can paste complex queries directly into the editor without having to manually build queries using the keywords and conditions cards in the standard condition builder.

Here are the key benefits to using the KeyQL editor:

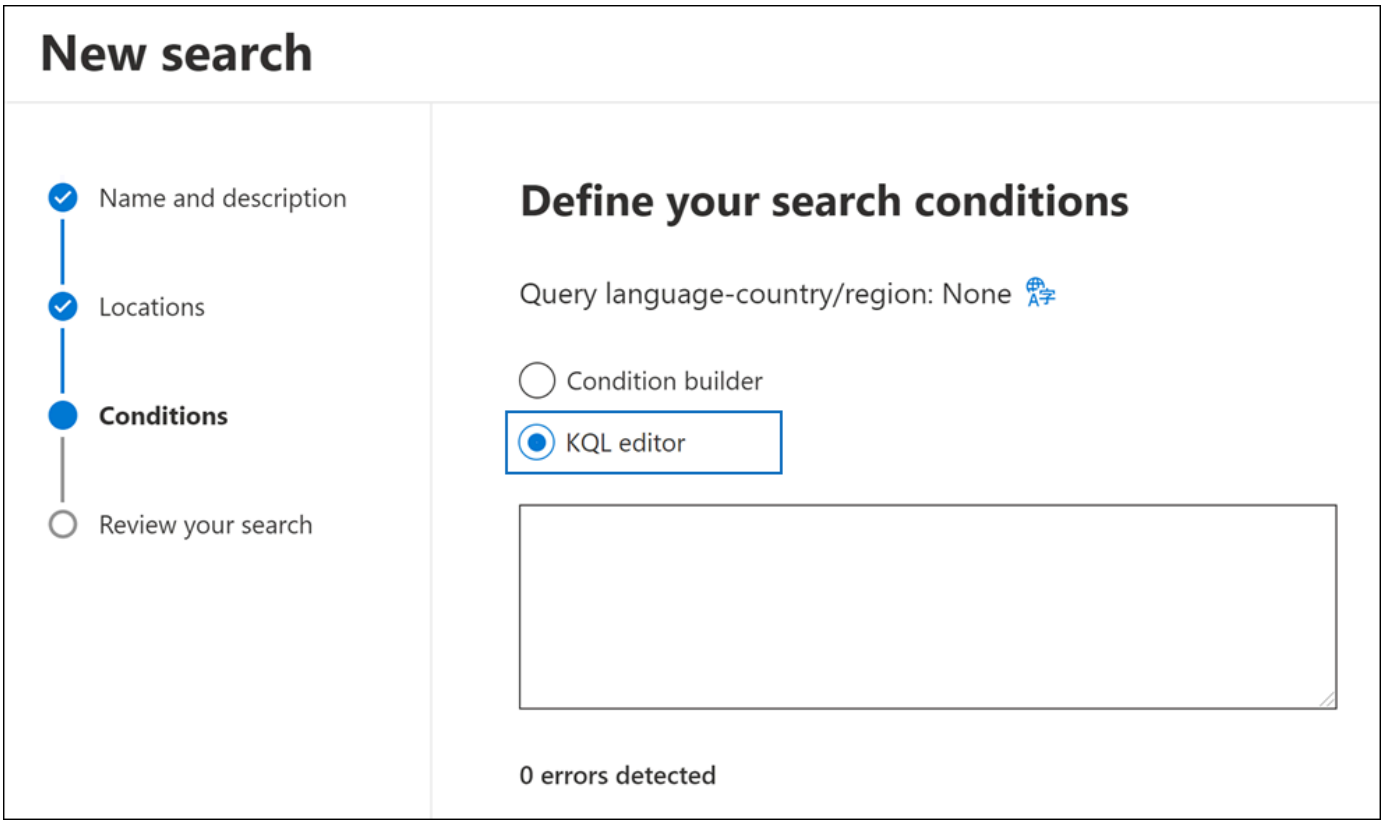
- Provides guidance and helps you build search queries from scratch.
- Lets you quickly paste long, complex queries directly into the editor. For example, if you receive a complex query from opposing counsel, you can paste that into the KeyQL editor instead of having to use the condition builder.
- Quickly identifies potential errors and displays hints about how to resolve issues.

The KeyQL editor is also available when you create query-based holds in eDiscovery (Standard) and eDiscovery (Premium).

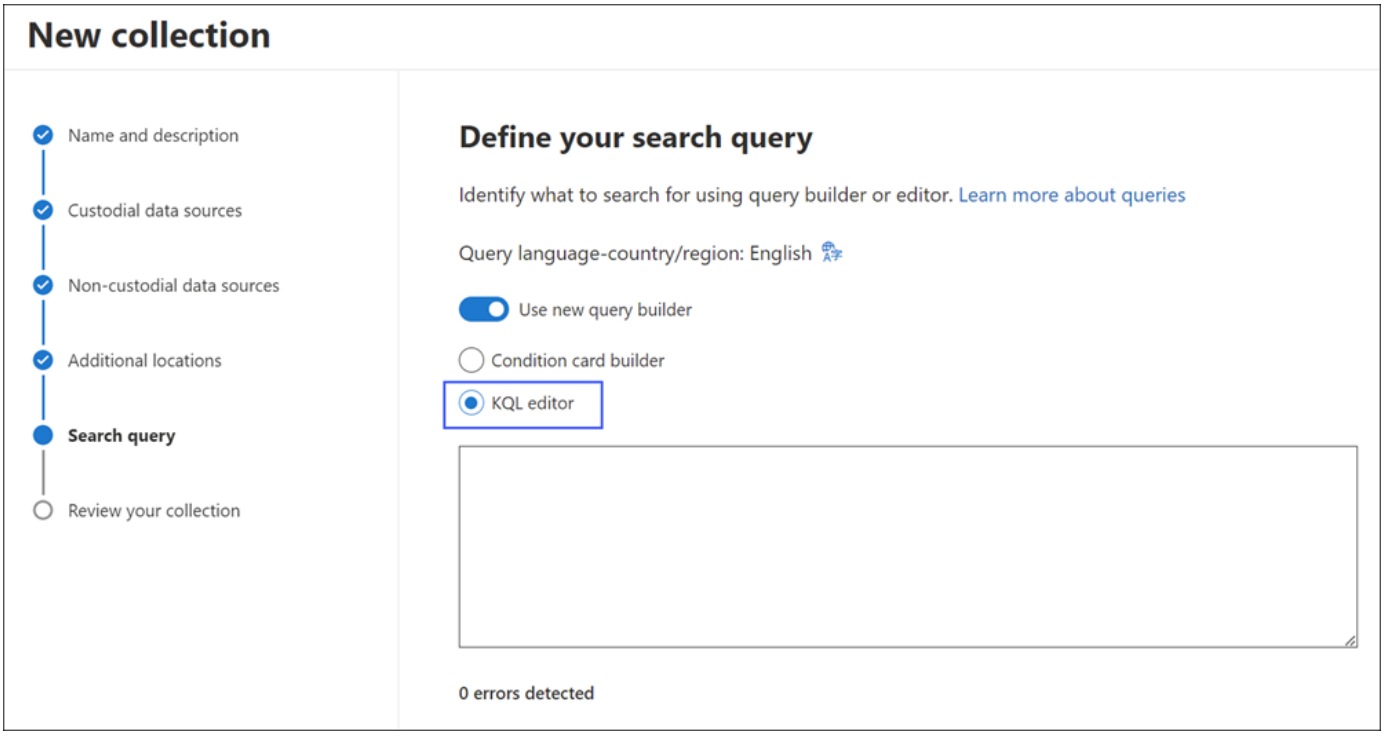
Displaying the KeyQL editor

When you create or edit an eDiscovery search, the option to display and use the KeyQL editor is located on the **Conditions** page in the search or collections workflow.

KeyQL editor in Content search and eDiscovery (Standard)



KeyQL editor in eDiscovery (Premium)

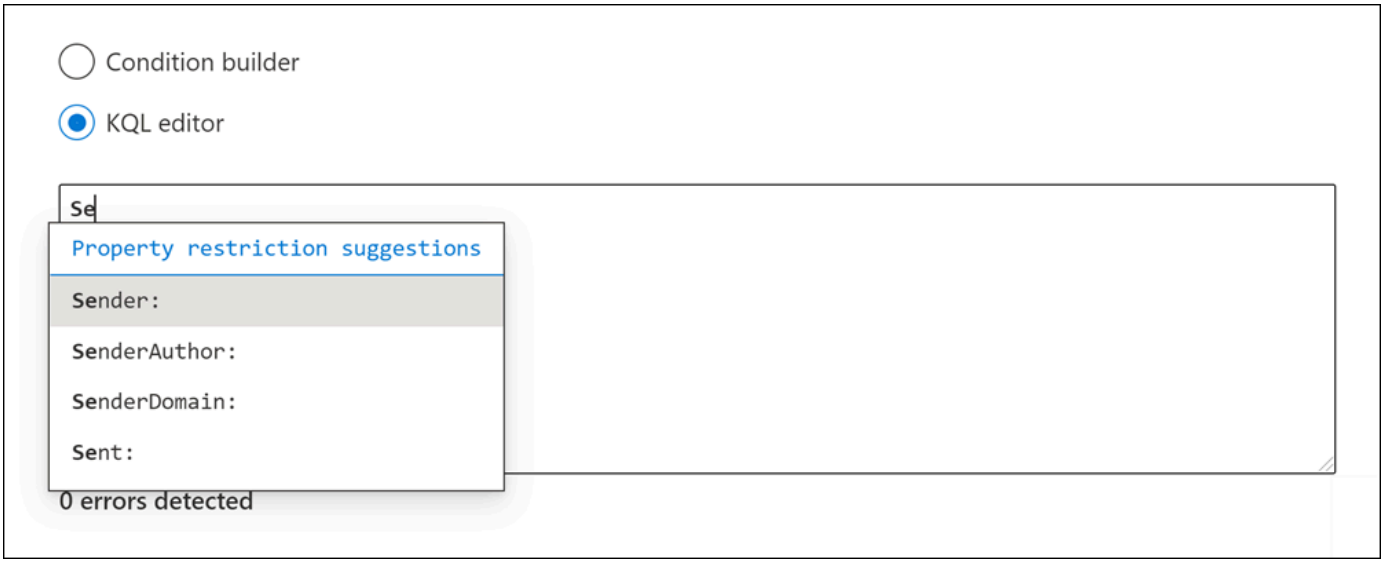


Using the KeyQL editor

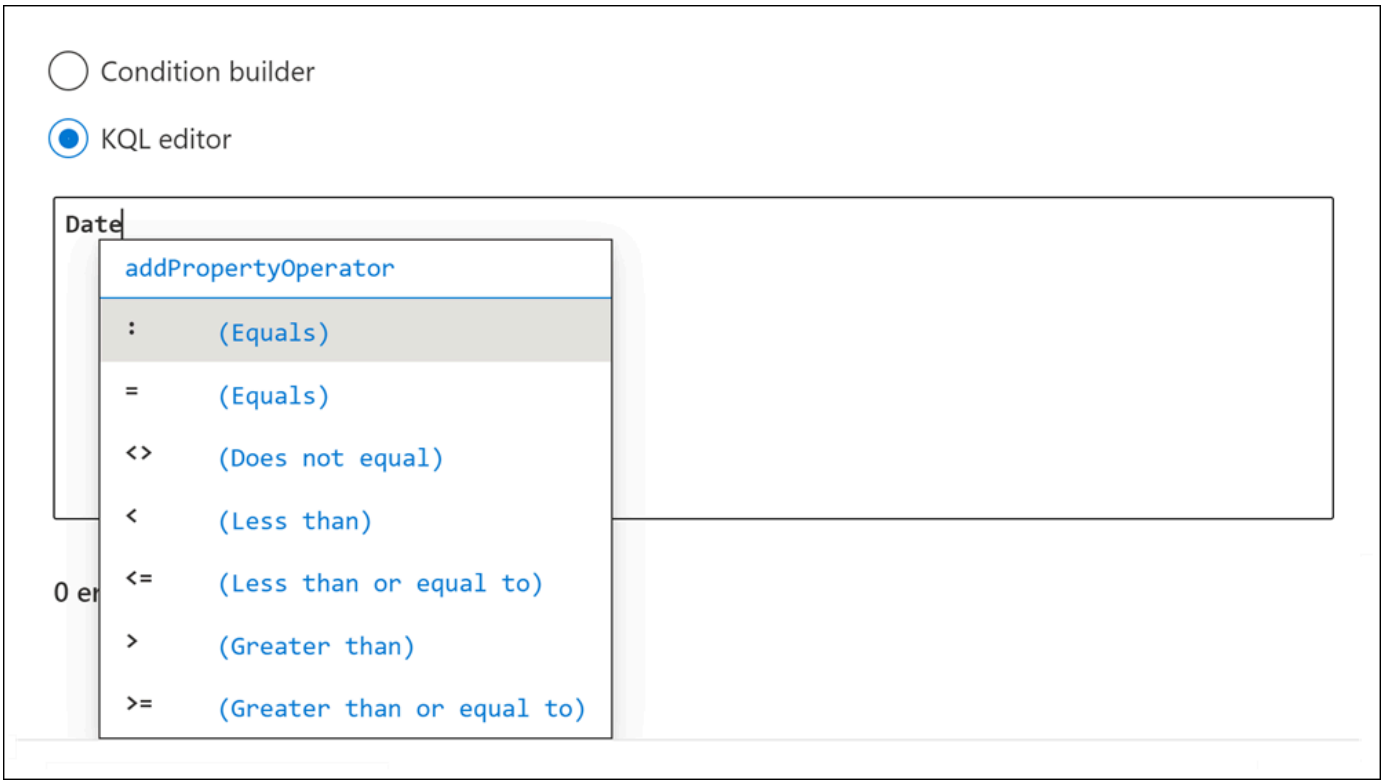
The following sections show examples of how the KeyQL editor provides suggestions and detects potential errors.

Autocompletion of search properties and operators

When you start to enter a search query in the KeyQL editor, the editor displays suggested autocompletion of supported search properties (also called *property restrictions*) that you can select. You have to type a minimum of two characters to display a list of supported properties that begin with those two characters. For example, the following screenshot shows the suggested search properties that begin with `Se`.



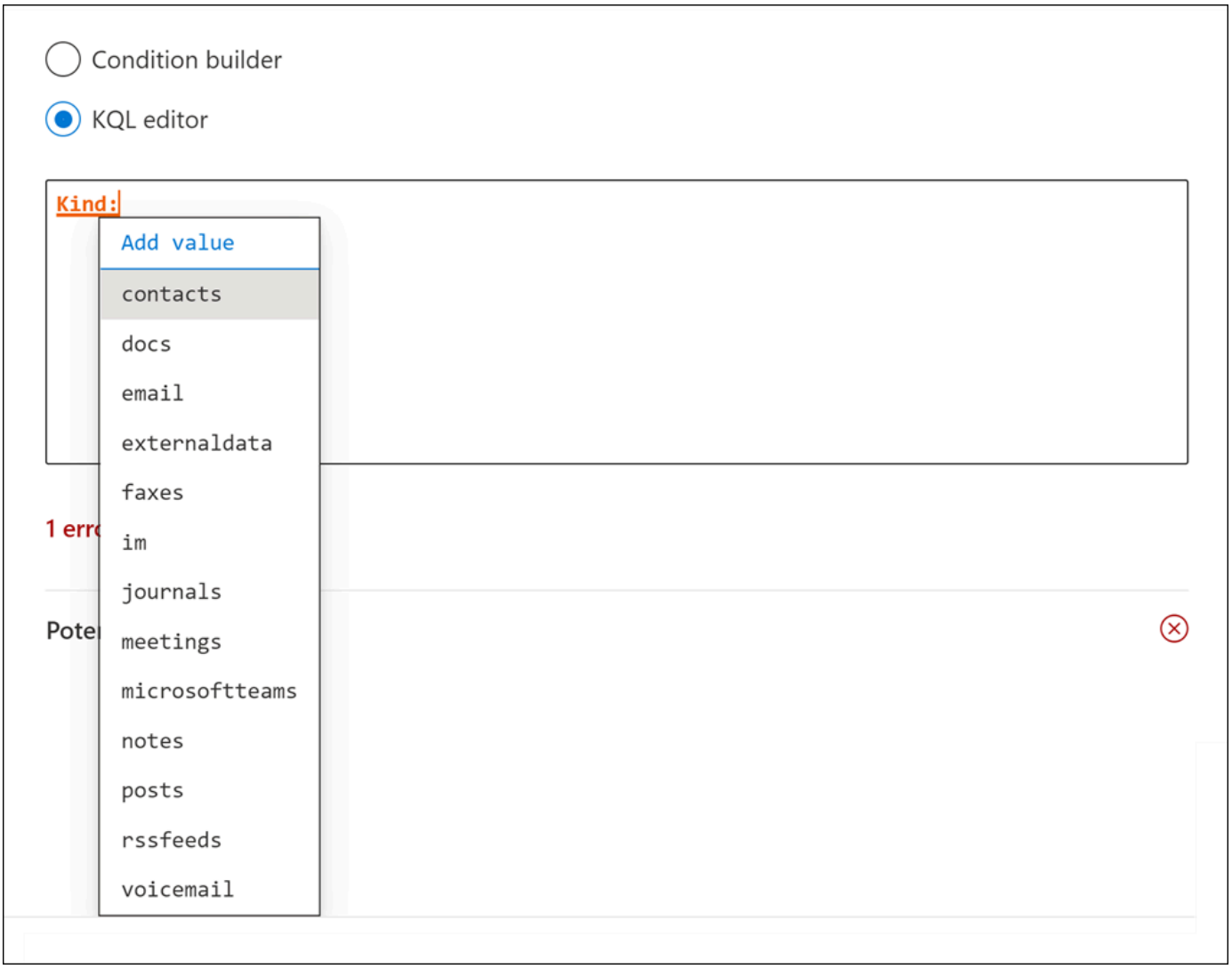
Additionally, the editor also suggests provides a list of supported operators (such as `:`, `=` and `<>`) when you type in a complete property name. For example, the following screenshot shows the suggested operators for the `Date` property.



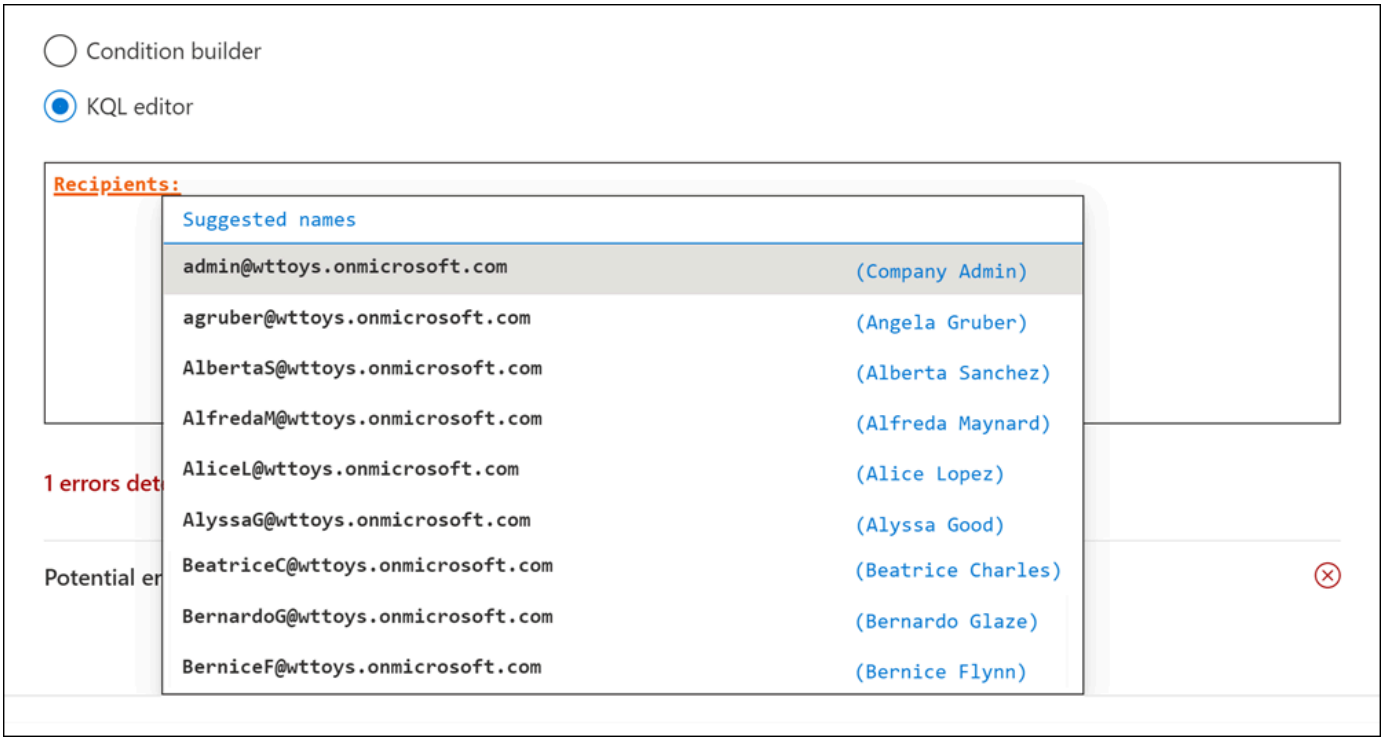
For more information about the supported search properties and operators, see Keyword queries and search conditions for eDiscovery.

Property value suggestions

The KeyQL editor provides suggestions for possible values of some properties. For example, the following screenshot shows the suggested values for the `Kind` property.



The editor also suggests a list of users (in UPN format) when you type email recipient properties, such as `From`, `To`, `Recipients` and `Participants`.

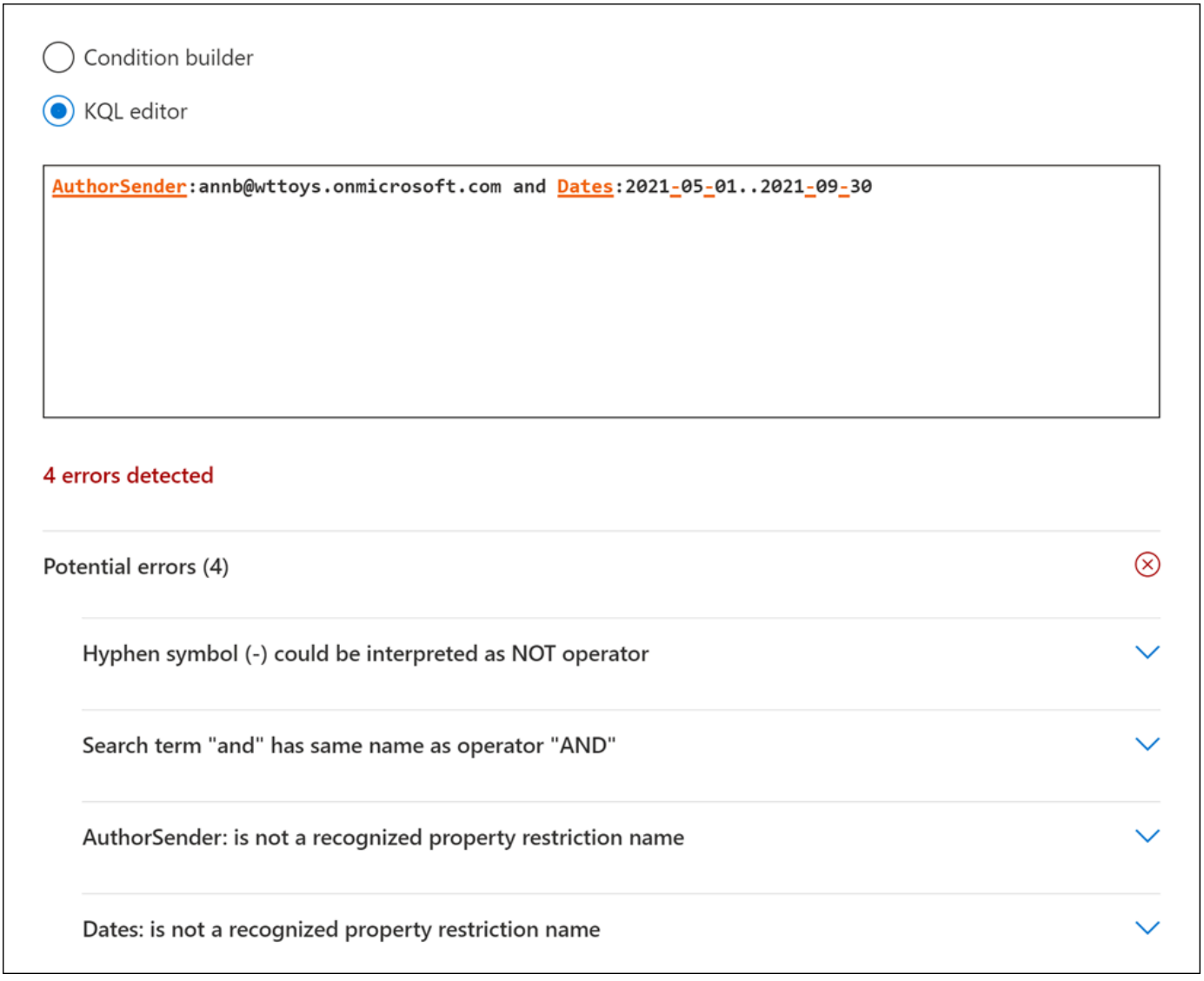


Detection of potential errors

The KeyQL editor detects potential errors in search queries, and provides a hint of what is causing the error to help you resolve the error. The editor also indicates a potential error when a property doesn't have a corresponding operation or value. Potential errors in the query are highlighted in red text, and explanations and possible fixes for the error are displayed in the **Potential errors** drop-down section.

> **Important**
>
> Nested quotation marks aren't supported in the KeyQL editor.

For example, if you pasted the following query into the KeyQL editor, four potential errors would be detected.



In this case, you can use the potential error hints to troubleshoot and fix the query.

More information

- You can toggle between the condition builder and the KeyQL editor. For example, if you use the condition builder to configure a query using the Keywords box and multiple condition cards, you can display the resulting query in the KeyQL editor. However, if you create a complex query (with keywords and conditions) in the KeyQL editor, the resulting query is displayed only in the Keywords box when you view it in the condition builder.
- If you paste a complex query into the KeyQL editor, the editor detects potential errors and suggests possible solutions to resolve errors.

Last updated on 11/21/2025

Use the query builder to create search queries

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

The query builder option in collection search tool provides a visual filtering experience when you build search queries in Microsoft Purview eDiscovery (Premium). Use the new query builder to construct complex queries with additional functionality, including AND, OR, and grouping of conditions. These features in the query builder help you build queries more effectively, provide a visual interface for grouping sub-queries, and provide additional space for complex keyword queries to be constructed and reviewed.

Displaying the query builder

When you create or edit an eDiscovery search for a collection in eDiscovery (Premium), the option to display and use the query builder is located on the **Search query** page in the collections workflow. Select **Use new query builder** to display and use the query builder.

Using the query builder

After you've selected **Use new query builder**, you're ready to get started. To create a query and custom filtering for your search, use the following controls:

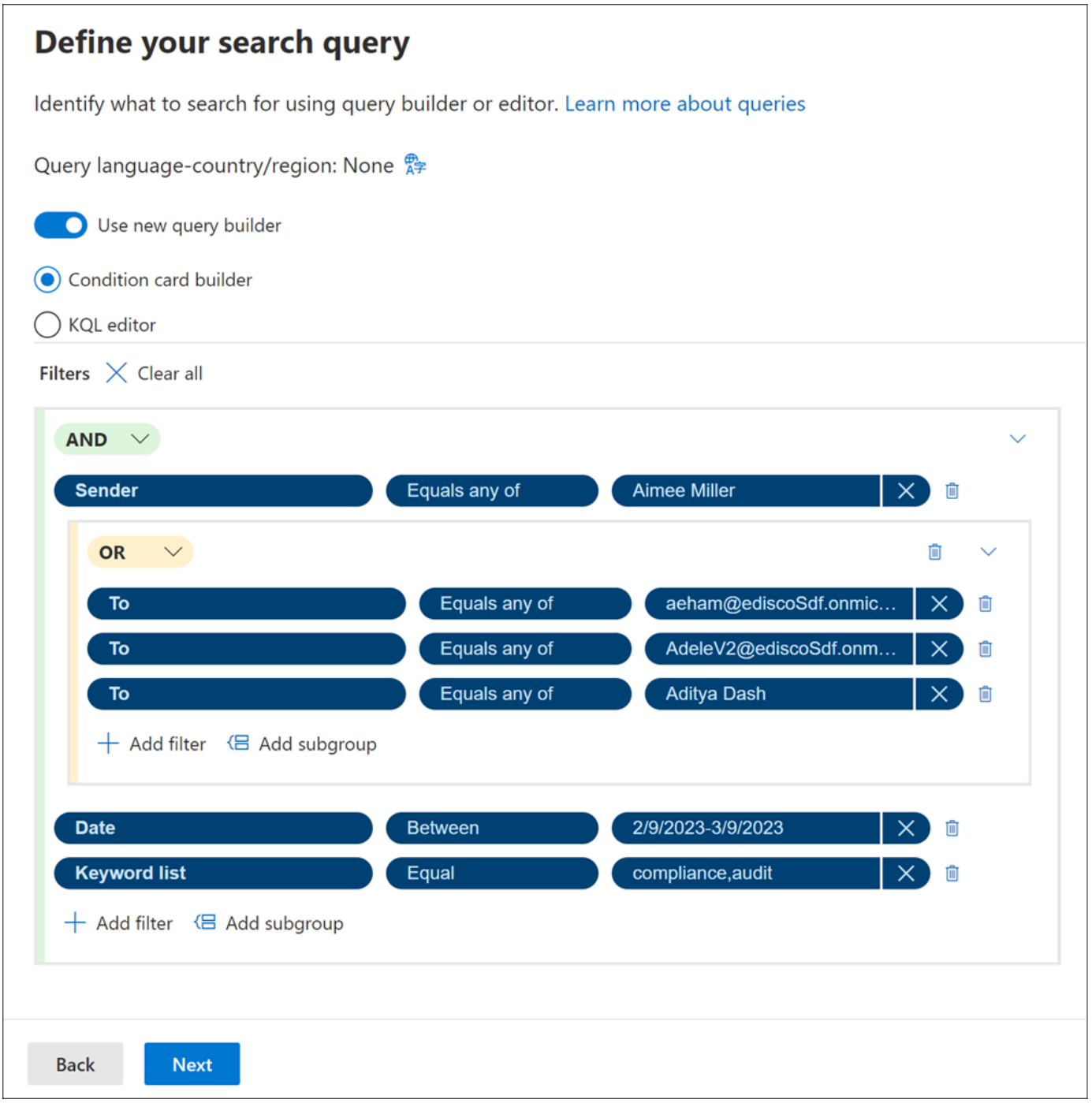
- **AND/OR**: These conditional logical operators allow you to select the query condition that applies to specific filters and filter subgroups. These operators allow you to use multiple filters or subgroups connected to a single filter in your query.
- **Select a filter**: Allows you to select filters for the specific data sources and location content selected for the collection.

- **Add filter**: Allows you to add multiple filters to your query. Is available after you've defined at least one query filter.
- **Select an operator**: Depending on the selected filter, the operators compatible for the filter are available to select. For example, if the *Date* filter is selected, the available operators are *Before*, *After*, and *Between*. If the *Size (in bytes)* filter is selected, the available operators are *Greater than*, *Greater or equal*, *Less than*, *Less or equal*, *Between*, and *Equal*.
- **Value**: Depending on the selected filter, the values compatible for the filter are available. Additionally, some filters support multiple values and some filters support one specific value. For example, if the *Date* filter is selected, select date values. If the *Size (in bytes)* filter is selected, select a value for bytes.
- **Add subgroup**: After you've defined a filter, you can add a subgroup to refine the results returned by the filter. You can also add a subgroup to a subgroup for multi-layered query refinement.
- **Remove a filter condition**: To remove an individual filter or subgroup, select the remove icon to the right of each filter line or subgroup.
- **Clear all**: To clear the entire query of all filters and subgroups, select **Clear all**.

Scenario example

The eDiscovery administrator needs to create a query to find emails sent from Aimee Miller to Adam Eham, Adele Vance, or Aditya Dash that were sent between February 9, 2023 and March 9, 2023 that contains the keywords *compliance* and *audit*. For this example, the administrator creates the following query using the new query builder:

1. For the first filter, the administrator selects *Sender*, then selects the *Equals any of* operator, then selects *Aimee Miller* from the list of users available in the **Value** control.
2. Next, the administrator selects **Add subgroup** and the **OR** operator to define the other users that Aimee may have sent an email to about the compliance audit.
3. In the subgroup, the administrator selects the *To* filter, the *Equals any of* operator, and the *Value (user)* for each of the other users that Aimee may have sent email to about the compliance audit. In this example, the administrator creates a filter in the subgroup for Adam Eham, Adele Vance, and Aditya Dash.
4. To define the date range, the administrator selects **Add filter** and selects the *Date* filter, the *Between* operator, and start and ending dates for the *Value*.
5. Finally, the administrator selects the *Keyword list* filter, the *Equal* operator, and *compliance, audit* as the keyword *Value*.



Last updated on 11/21/2025

Use the Natural Language Query builder to create search queries (preview)

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic **Content Search**, classic **eDiscovery (Standard)**, and classic **eDiscovery (Premium)**.
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the **new eDiscovery experience** in the **Microsoft Purview portal**.

The Natural Language Query (preview) builder option in collection search tool allows you to use natural language and Microsoft Security Copilot to quickly generate a Keyword Query Language (KeyQL) statement when you build search queries in Microsoft Purview eDiscovery (Premium).

Use the Natural Language Query (preview) builder to construct complex queries with additional functionality, including AND, OR, and grouping of conditions, all while using natural language prompts. This feature helps you build queries more easily using predefined prompts for example scenarios and allows you to refine and enhance custom prompts for more accurate search queries.

Prerequisites to use the Natural Language Query (preview) builder

You must have Microsoft Security Copilot to use the Natural Language Query (preview) builder. For Security Copilot guidance, see the following:

- Microsoft Security Copilot in Microsoft Purview
- Get started with Microsoft Security Copilot

Microsoft Purview eDiscovery capabilities are included with Microsoft Purview. For more information, see Subscriptions and licensing.

For information about permissions you need to use the Natural Language Query (preview) builder, see Understand authentication in Microsoft Security Copilot.

Displaying the Natural Language Query (preview) builder

When you create or edit an eDiscovery search for a collection in eDiscovery (Premium), the option to display and use the Natural Language Query builder is located on the **Search query** page in the collections workflow. Select **Natural Language Query** to display and use the query builder.

Using the Natural Language Query (preview) builder

After you've selected data sources for a search query, you can create a customized KeyQL search query from a natural language prompt using Security Copilot. You can also choose to use prompt suggestions as a starting point to create and refine KeyQL queries for common or custom search scenarios.

After you've selected data sources for a search query, you can create a customized KeyQL search query from a natural language prompt using Microsoft Security Copilot. You can also choose to use prompt suggestions as a starting point to create and refine KeyQL queries for common or custom search scenarios.

1. In the **Natural language prompt** pane, choose one of the following options:

    - Enter your search query question. You can include user, data source, and other content details as applicable.
    - Select **View prompts** to select one of the following prompt suggestions:
        - *Find all emails containing the words budget and finance and have attachments*
        - *Search all chats in the month of January 2020 that contain the word 'financial year'*
        - *Search for files of type .docx that contain the words confidential and budget*

2. Review the natural language prompt. To refine the prompt with Copilot, select **Refine**.
3. When the prompt is finalized, select **Generate KeyQL**.
4. Review the KeyQL query in the **Keyword Query Language (KeyQL) result** pane. If you need to refine the KeyQL query results, you can update the prompt in the **Natural language prompt** pane and select **Generate KeyQL** again. 1. When the KeyQL results are finalized, select **Copy KeyQL**.
5. Paste the KeyQL results in the query field on the **Keyword query language (KeyQL)** tab. You can close **Draft a query with Copilot**.
6. Select **Next**.

7. On the **Review your collection** page, you can review and update the collection selection, data sources, and conditions that you configured on the previous pages. Select **Edit** for any of the selections and update if needed.
8. Select **Submit** to create the collection estimate. A page is displayed confirming that the collection was created.

Last updated on 11/21/2025

Keyword queries and search conditions for eDiscovery

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

This article describes the properties available to help find content across email and chat in Exchange Online and documents and files stored on SharePoint and OneDrive using the eDiscovery search tools in the Microsoft Purview portal.

This includes Content search, Microsoft Purview eDiscovery (Standard), and Microsoft Purview eDiscovery (Premium) (eDiscovery searches in eDiscovery (Premium) are called *collections*). You can also use the ***-ComplianceSearch** cmdlets in Security & Compliance PowerShell to search for these properties.

This article also describes:

- Using Boolean search operators, search conditions, and other search query techniques to refine your search results.
- Searching for communications of various types related to specific users and projects during a specific time frame.
- Searching for site content that is related to a specific project, users and/or subjects during a specific time period.

For step-by-step instructions on how to create different eDiscovery searches, see:

- Content search
- Search for content in eDiscovery (Standard)
- Create a collection estimate in eDiscovery (Premium)

> **Note**
>
> eDiscovery searches in the Microsoft Purview portal and the corresponding ***-ComplianceSearch** cmdlets in Security & Compliance PowerShell use the Keyword Query Language (KeyQL). For more detailed information, see Keyword Query Language syntax reference.

Search tips and tricks

- The time zone for all searches is Coordinated Universal Time (UTC). Changing time zones for your organization isn't currently supported. Time zone display settings in the search view are only for applicable for values in *Data* column and don't affect time stamps on collected items.
- Keyword searches aren't case-sensitive. For example, **cat** and **CAT** return the same results.
- The Boolean operators **AND**, **OR**, **NOT**, and **NEAR** must be uppercase.

- Using quotes stops wild cards and any operations inside the quotes.
- A space between two keywords or two `property:value` expressions is the same as using **OR**. For example, `from:"Sara Davis" subject:reorganization` returns all messages sent by Sara Davis or messages that contain the word reorganization in the subject line. However, using a mix of spaces and **OR** conditionals in a single query might lead to unexpected results. We recommend using either spaces or **OR** in a single query.
- Use syntax that matches the `property:value` format. Values aren't case-sensitive, and they can't have a space after the operator. If there's a space, your intended value is a full-text search. For example `to: pilarp` searches for "pilarp" as a keyword, rather than for messages sent to pilarp.
- When searching a recipient property, such as To, From, Cc, or Recipients, you can use an SMTP address, alias, or display name to denote a recipient. For example, you can use pilarp@contoso.com, pilarp, or "Pilar Pinilla."
- You can use only prefix searches; for example, **cat*** or **set***. Suffix searches (***cat**), infix searches (**c*t**), and substring searches (***cat***) aren't supported.
- When searching a property, use double quotation marks (" ") if the search value consists of multiple words or special characters. For example, `subject:budget Q1` returns messages that contain **budget** in the subject line and that contain **Q1** anywhere in the message or in any of the message properties. Using `subject:"budget Q1"` returns all messages that contain **budget Q1** anywhere in the subject line.
- To exclude content marked with a certain property value from your search results, place a minus sign (-) before the name of the property. For example, `-from:"Sara Davis"` excludes any messages sent by Sara Davis.
- You can export items based on message type. For example, to export Skype conversations and chats in Microsoft Teams, use the syntax `kind:im`. To return only email messages, you would use `kind:email`. To return chats, meetings, and calls in Microsoft Teams, use `kind:microsoftteams`.
- When searching sites, you have to add the trailing `/` to the end of the URL when using the `path` property to return only items in a specified site. If you don't include the trailing `/`, items from a site with a similar path name are also returned. For example, if you use `path:sites/HelloWorld` then items from sites named `sites/HelloWorld_East` or `sites/HelloWorld_West` would also be returned. To return items only from the HelloWorld site, you have to use `path:sites/HelloWorld/`.
- The **Query language-country/region** must be defined in your search query prior to collecting content.
- When searching the **Sent** folders for emails, using the SMTP address for the sender isn't supported. Items in the **Sent** folder contain only display names.

Finding content in Exchange Online

Admins are often charged with finding out who knew what when in the most efficient and effective way possible to respond to requests concerning ongoing or potential litigation, internal investigations, and other scenarios. These requests are often urgent, involve multiple stakeholder teams, and have significant impact if not completed in a timely manner. Knowing how to find the right information is critical for admins to complete searches successfully and help their organizations to manage the risk and cost associated with eDiscovery requirements.

When an eDiscovery request is submitted, often there's only partial information available for the admin to start to collect content that might be related to a particular investigation. The request might include employee names, project titles, rough date ranges when the project was active, and not much more. From this information, the admin needs to create queries to find relevant content across Microsoft 365 services to determine the information needed for a particular

project or subject. Understanding how information is stored and managed for these services help admins more efficiently find what they need quickly and in an effective manner.

Email, chat, meeting, and Microsoft 365 Copilot and Microsoft 365 Copilot Chat activity data (user prompts and Copilot responses) are all stored in Exchange Online. Many communication properties are available for searching items included in Exchange Online. Some properties such as *From*, *Sent*, *Subject*, and *To* are unique to certain items and aren't relevant when searching for files or documents in SharePoint and OneDrive for Business. Including these types of properties when searching across workloads can sometimes lead to unexpected results.

For example, to find content related to specific employees (*User 1* and *User 2*), associated with a project called *Tradewinds*, and during January 2020 through January 2022, you might use a query with the following properties:

- Add User 1 and User 2's Exchange Online locations as data sources to the case
- Select User 1 and User 2's Exchange Online locations as collection locations
- For **Keyword**, use *Tradewinds*
- For **Date Range**, use the *January 1, 2020* to *January 31, 2022* range

> **Important**
>
> For emails, when a keyword is used, we search subject, body, and many properties related to the participants. However, due to recipient expansion, search might not return expected results when using the alias or part of the alias. Therefore we recommend using the full UPN.

Searchable email properties

The following table lists the email message properties that can be searched by using the eDiscovery search tools in the Microsoft Purview portal or by using the **New-ComplianceSearch** or the **Set-ComplianceSearch** cmdlet.

> **Important**
>
> While email messages might have other properties supported in other Microsoft 365 services, only the email properties listed in this table are supported in eDiscovery search tools. Attempting to include other email messages properties in searches isn't supported.

The table includes an example of the *property:value* syntax for each property and a description of the search results returned by the examples. You can enter these `property:value` pairs in the keywords box for an eDiscovery search.

> **Note**
>
> When searching email properties, it's not possible to search for message headers. Header information isn't indexed for collections. Additionally, items in which the specified property is empty or blank aren't searchable. For example, using the *property:value* pair of **subject:""** to search for email messages with an empty subject line return zero results. This also applies when searching site and contact properties.

Expand table

Property	Property description	Examples	Search results returned by the examples

AttachmentNames	The names of files attached to an email message.	`attachmentnames:annualreport.ppt`
		
		`attachmentnames:annual*`
			Messages that have an attached file named *annualreport.ppt*. In the second example, using the wildcard character (*) returns messages with the word *annual* in the file name of an attachment.[1]

Bcc	The Bcc field of an email message.[1]	`bcc:pilarp@contoso.com`
		
		`bcc:pilarp`
		
		`bcc:"Pilar Pinilla"`
			All examples return messages with *Pilar Pinilla* included in the Bcc field. (See Recipient Expansion)

Category	The categories to search. Categories can be defined by users by using Outlook or Outlook on the web (formerly known as Outlook Web App). The possible values are: • blue • green • orange • purple • red • yellow	`category:"Red Category"`	Messages that have been assigned the *red* category in the source mailboxes.

Cc	The Cc field of an email message.[1]	`cc:pilarp@contoso.com`
		
		`cc:"Pilar Pinilla"`
			In both examples, messages with *Pilar Pinilla* specified in the Cc field. (See Recipient Expansion)

Folderid	The folder ID (GUID) of a specific mailbox folder in 48-character format. If you use this property, be sure to search the mailbox that the specified folder is located in. Only the specified folder is searched. Any subfolders in the folder won't be searched. To search subfolders, you need to use the *Folderid* property for the subfolder you want to search.	`folderid:4D6DD7F943C29041A65787E30F02AD1F00000000013A0000`
		
		`folderid:2370FB455F82FC44BE31397F47B632A70000000001160000 AND participants:garthf@contoso.com`
			The first example returns all items in the specified mailbo folder. The second example returns all items in the specifie mailbox folder that were sent or receive by *garthf@contoso.com*

From	The sender of an email message.[1]	`from:pilarp@contoso.com`	Messages sent by th specified user. (See Recipient Expansion)
HasAttachment	Indicates whether a message has an attachment. Use the values **true** or **false**.	`from:pilar@contoso.com AND hasattachment:true`	Messages sent by th specified user that have attachments.

Importance	The importance of an email message, which a sender can specify when sending a	`importance:high`
		
		`importance:medium`
			Messages that are marked as high importance, mediun

Property	Property description	Examples	Search results returned by the examples
	message. By default, messages are sent with normal importance, unless the sender sets the importance as **high** or **low**.	`importance:low`	importance, or low importance.

IsRead	Indicates whether messages have been read. Use the values **true** or **false**.	`isread:true`
		
		`isread:false`
			The first example returns messages with the IsRead property set to **True**. The second example returns messages with the IsRead property set to **False**

ItemClass	Use this property to search specific third-party data types that your organization imported to Office 365. Use the following syntax for this property: `itemclass:ipm.externaldata.<third-party data type>*`	`itemclass:ipm.externaldata.Facebook* AND subject:contoso`
		
		`itemclass:ipm.externaldata.Twitter* AND from:"Ann Beebe" AND "Northwind Traders"`
			The first example returns Facebook items that contain the word "contoso" in the Subject property. The second example returns Twitter items that were posted by Ann Beebe and that contain the keyword phrase "Northwind Traders".
			For a complete list of values to use for third-party data types for the ItemClass property, see Use Content search to search third-party data that was imported to Office 365.

Kind	The type of email message to search for. Possible values:
	contacts
	
	docs
	
	email
	
	externaldata
	
	faxes
	
	im
	
	journals
	
	meetings
	
	microsoftteams (returns items from chats, meetings, and calls in Microsoft Teams)
	
	notes
	
	posts
	
	rssfeeds
	
	tasks
		`kind:email`
		
		`kind:email OR kind:im OR kind:voicemail`
		
		`kind:externaldata`
			The first example returns email messages that meet the search criteria. The second example returns email messages, instant messaging conversations (including Skype for Business conversations and chats in Microsoft Teams), and voice messages that meet the search criteria. The third example returns items that were imported to mailboxes in Microsoft 365 from third-party data sources, such as Twitter, Facebook, and Cisco Jabber that meet the search criteria. For more information, see

Property	Property description	Examples	Search results returned by the examples
	voicemail		Archiving third-party data in Office 365 .
Participants	All the people fields in an email message. These fields are From, To, Cc, and Bcc.[1]	`participants:garthf@contoso.com` `participants:contoso.com`	Messages sent by or sent to garthf@contoso.com The second example returns all messages sent by or sent to a user in the contoso.com domain. (See Recipient Expansion)
Received	The date that an email message was received by a recipient.	`received:2021-04-15` `received>=2021-01-01 AND received<=2021-03-31`	Messages that were received on April 15, 2021. The second example returns all messages received between January 1, 2021 and March 31, 2021.
Recipients	All recipient fields in an email message. These fields are To, Cc, and Bcc.[1]	`recipients:garthf@contoso.com` `recipients:contoso.com`	Messages sent to garthf@contoso.com The second example returns messages sent to any recipient in the contoso.com domain. (See Recipient Expansion)
Sent	The date that an email message was sent by the sender.	`sent:2021-07-01` `sent>=2021-06-01 AND sent<=2021-07-01`	Messages that were sent on the specifiec date or sent within the specified date range.
Size	The size of an item, in bytes.	`size>26214400` `size:1..1048567`	Messages larger tha 25 MB. The second example returns messages from 1 through 1,048,567 bytes (1 MB) in size.
Subject	The text in the subject line of an email message. **Note:** When you use the Subject property in a query, the search returns all messages in which the subject line contains the text you're searching for. In other words, the query doesn't return only those messages that have an exact match. For example, if you search for `subject:"Quarterly Financials"`, your results include messages with the subject "Quarterly Financials 2018".	`subject:"Quarterly Financials"` `subject:northwind`	Messages that contain the phrase "Quarterly Financials anywhere in the text of the subject line. The second example returns all messages that contain the word northwind in the subject line.
To	The To field of an email message.[1]	`to:annb@contoso.com`	All examples return messages where Anr

Property	Property description	Examples	Search results returned by the examples
		`to:annb` `to:"Ann Beebe"`	Beebe is specified in the To: line.

> **Note**
>
> [1] For the value of a recipient property, you can use email address (also called *user principal name* or UPN), display name, or alias to specify a user. For example, you can use annb@contoso.com, annb, or "Ann Beebe" to specify the user Ann Beebe.

Recipient expansion

> **Tip**
>
> Use the new **eDiscovery** experience and condition builder to use common mailbox and site **properties** like *MessageIDs* and *ChatThreadIds* when searching for specific recipients or messages.

When searching any of the recipient properties (From, To, Cc, Bcc, Participants, and Recipients), Microsoft 365 attempts to expand the identity of each user by looking them up in Microsoft Entra ID. If the user is found in Microsoft Entra ID, the query is expanded to include the user's email address (or UPN), alias, display name, and LegacyExchangeDN. For example, a query such as `participants:ronnie@contoso.com` expands to `participants:ronnie@contoso.com OR participants:ronnie OR participants:"Ronald Nelson" OR participants:"<LegacyExchangeDN>"`.

To prevent recipient expansion, add a wild card character (asterisk) to the end of the email address and use a reduced domain name; for example, `participants:"ronnie@contoso*"` Be sure to surround the email address with double quotation marks.

However, be aware that preventing recipient expansion in the search query might result in relevant items not being returned in the search results. Email messages in Exchange can be saved with different text formats in the recipient fields. Recipient expansion is intended to help mitigate this fact by returning messages that might contain different text formats. So preventing recipient expansion might result in the search query not returning all items that might be relevant to your investigation.

> **Note**
>
> If you need to review or reduce the items returned by a search query due to recipient expansion, consider using eDiscovery (Premium). You can search for messages (taking advantage of recipient expansion), add them to a review set, and then use review set queries or filters to review or narrow the results. For more information, see Collect data for a case and Query the data in a review set.

Finding content in SharePoint and OneDrive

> **Tip**
>
> Use the new **eDiscovery** experience and search **export options** to download all list attachments for SharePoint sites.

When searching for documents and files located in SharePoint or OneDrive for Business, it might make sense to adjust the query approach based on the metadata for the documents and files of interest. Files and documents have relevant properties like *Author*, *Created*, *CreatedBy*, *FileName*, *LastModifiedTime*, and *Title*. Most of these proprieties aren't relevant when searching for communications content in Exchange Online, and using these properties might lead to unexpected results if used across both documents and communications. Additionally, *FileName* and *Title* of a document might not be the same and using one or the other to try to find a file with specific content might lead to different or inaccurate results. Keep these properties in mind when searching for specific document and file content in SharePoint and OneDrive for Business.

For example, to find content related to documents created by User 1, for a project called *Tradewinds*, for specific files named *Financials*, and from January 2020 to January 2022, you might use a query with the following properties:

- Add User 1's OneDrive for Business site as a data sources to the case
- Select User 1's OneDrive for Business site as a collection location
- Add additional SharePoint site locations related to the project as collection locations
- For **FileName**, use *Financials*
- For **Keyword**, use *Tradewinds*
- For **Date Range**, use the *January 1, 2020* to *January 31, 2022* range

Searchable site properties

The following table lists the SharePoint and OneDrive for Business properties that can be searched by using the eDiscovery search tools in the Microsoft Purview portal or by using the **New-ComplianceSearch** or the **Set-ComplianceSearch** cmdlet.

> **Important**
>
> While documents and files stored on SharePoint and OneDrive for Business might have other properties supported in other Microsoft 365 services, only the document and file properties listed in this table are supported in eDiscovery search tools. Attempting to include other document or file properties in searches isn't supported.

The table includes an example of the *property:value* syntax for each property and a description of the search results returned by the examples.

Expand table

Property	Property description	Example
Author	The author field from Office documents, which persists if a document is copied. For example, if a user creates a document and the emails it to someone else who then uploads it to SharePoint, the document will still retain the original author. Be sure to use the user's display name for this property.	`author:"Garth Fort"`
ContentType	The SharePoint content type of an item, such as Item, Document, or Video.	`contenttype:document`
Created	The date that an item is created.	`created>=2021-06-01`

Property	Property description	Example
CreatedBy	The person that created or uploaded an item. Be sure to use the user's display name for this property.	`createdby:"Garth Fort"`
DetectedLanguage	The language of an item.	`detectedlanguage:english`

DocumentLink	The path (URL) of a specific folder on a SharePoint or OneDrive for Business site. If you use this property, be sure to search the site that the specified folder is located in. We recommend using this property instead of the *Site* and *Path* properties.
	
	To return items located in subfolders of the folder that you specify for the documentlink property, you have to add /* to the URL of the specified folder; for example, `documentlink: "https://contoso.sharepoint.com/Shared Documents/*"`
		`documentlink:"https://contoso-my.sharepoint.com/personal/garthf_contoso_com/Documents/Priva`
		
		`documentlink:"https://contoso-my.sharepoint.com/personal/garthf_contoso_com/Documents/Share with Everyone/*" AND filename:confidential`

FileExtension	The extension of a file; for example, docx, one, pptx, or xlsx.	`fileextension:xlsx`

FileName	The name of a file.	`filename:"marketing plan"`
		
		`filename:estimate`

LastModifiedTime	The date that an item was last changed.	`lastmodifiedtime>=2021-05-01`
		
		`lastmodifiedtime>=2021-05-01 AND lastmodifiedtime<=2021-06-01`

ModifiedBy	The person who last changed an item. Be sure to use the user's display name for this property.	`modifiedby:"Garth Fort"`

SharedWithUsersOWSUser	Documents that have been shared with the specified user and displayed on the **Shared with me** page in the user's OneDrive for Business site. These are documents that have been explicitly shared with the specified user by other	`sharedwithusersowsuser:garthf`
		
		`sharedwithusersowsuser:"garthf@contoso.com"`

Property	Property description	Example
	people in your organization. When you export documents that match a search query that uses the SharedWithUsersOWSUser property, the documents are exported from the original content location of the person who shared the document with the specified user. For more information, see Searching for site content shared within your organization.	
Size	The size of an item, in bytes.	`size>=1` `size:1..10000`
Title	The title of the document. The Title property is metadata that's specified in Microsoft Office documents. It's different from the file name of the document.	`title:"communication plan"`

Searchable contact properties

The following table lists the contact properties that are indexed and that you can search for using eDiscovery search tools. These are the properties that are available for users to configure for the contacts (also called personal contacts) that are located in the personal address book of a user's mailbox. To search for contacts, you can select the mailboxes to search and then use one or more contact properties in the keyword query.

> **Tip**
>
> To search for values that contain spaces or special characters, use double quotation marks (" ") to contain the phrase; for example, `businessaddress:"123 Main Street"`.

Expand table

Property	Property description
BusinessAddress	The address in the **Business Address** property. The property is also called the **Work** address on the contact properties page.
BusinessPhone	The phone number in any of the **Business Phone** number properties.
CompanyName	The name in the **Company** property.
Department	The name in the **Department** property.
DisplayName	The display name of the contact. This is the name in the **Full Name** property of the contact.
EmailAddress	The address for any email address property for the contact. Users can add multiple email addresses for a contact. Using this property would return contacts that match any of the

Property	Property description
	contact's email addresses.
FileAs	The **File as** property. This property is used to specify how the contact is listed in the user's contact list. For example, a contact could be listed as *FirstName,LastName* or *LastName,FirstName*.
GivenName	The name in the **First Name** property.
HomeAddress	The address in any of the **Home** address properties.
HomePhone	The phone number in any of the **Home** phone number properties.
IMAddress	The IM address property, which is typically an email address used for instant messaging.
MiddleName	The name in the **Middle** name property.
MobilePhone	The phone number in the **Mobile** phone number property.
Nickname	The name in the **Nickname** property.
OfficeLocation	The value in **Office** or **Office location** property.
OtherAddress	The value for the **Other** address property.
Surname	The name in the **Last** name property.
Title	The title in the **Job title** property.

Search operators

Boolean search operators, such as **AND**, **OR**, and **NOT**, help you define more-precise searches by including or excluding specific words in the search query. Other techniques, such as using property operators (such as `>=` or `..`), quotation marks, parentheses, and wildcards, help you refine a search query. The following table lists the operators that you can use to narrow or broaden search results.

Expand table

Operator	Usage	Description
AND	keyword1 AND keyword2	Returns items that include all of the specified keywords or `property:value` expressions. For example, `from:"Ann Beebe" AND subject:northwind` would return all messages sent by Ann Beebe that contained the word northwind in the subject line. [2]
+	keyword1 + keyword2 + keyword3	Returns items that contain *either* `keyword2` or `keyword3` *and* that also contain `keyword1`. Therefore, this example is equivalent to the query `(keyword2 OR keyword3) AND keyword1`. The query `keyword1 + keyword2` (with a space after the + symbol) isn't the same as using the **AND** operator. This query would be equivalent to `"keyword1 + keyword2"` and return items with the exact phase `"keyword1 + keyword2"`.
OR	keyword1 OR keyword2	Returns items that include one or more of the specified keywords or `property:value` expressions. [2]
NOT	keyword1 NOT keyword2 NOT from:"Ann Beebe" NOT kind:im	Excludes items specified by a keyword or a `property:value` expression. In the second example excludes messages sent by Ann Beebe. The third example excludes any instant messaging conversations, such as Skype for Business conversations that are saved to the Conversation History mailbox folder. [2]
NEAR	keyword1 NEAR(n) keyword2	Returns items with words that are near each other. In the *keyword1 NEAR(n) keyword2* syntax, *n* equals the number of words exclusive of *keyword1* and *keyword2*. For example, to identify instances where the term *best* is within 3 words of *worst* (example sentence, 'Best is opposite of worst.'), you would use **best NEAR(3) worst**. This returns

Operator	Usage	Description
		any items where there are 3 or fewer words between *best* (keyword1) and *worst* (keyword2). If no number is specified, the default *n* value is 8. [2]
:	property:value	The colon (:) in the `property:value` syntax specifies that the value of the property being searched for contains the specified value. For example, `recipients:garthf@contoso.com` returns any message sent to garthf@contoso.com.
=	property=value	The same as the `:` operator.
<	property<value	Denotes that the property being searched is less than the specified value. [1]
>	property>value	Denotes that the property being searched is greater than the specified value.[1]
<=	property<=value	Denotes that the property being searched is less than or equal to a specific value.[1]
>=	property>=value	Denotes that the property being searched is greater than or equal to a specific value.[1]
..	property:value1..value2	Denotes that the property being searched is greater than or equal to value1 and less than or equal to value2.[1]
" "	"fair value" subject:"Quarterly Financials"	In a keyword query (where you type the `property:value` pair in the **Keyword** box), use double quotation marks (" ") to search for an exact phrase or term. However, if you use the **Subject** or **Subject/Title** search condition condition, don't add double quotation marks to the value because quotation marks are automatically added when using these search conditions. If you do add quotation marks to the value, two pairs of double quotations are added to the condition value, and the search query will return an error.
*	cat* subject:set*	Prefix searches (also called *prefix matching*) where a wildcard character (*) is placed at the end of a word in keywords or `property:value` queries. In prefix searches, the search returns results with terms that contain the word followed by zero or more characters. For example, `title:set*` returns documents that contain the word "set", "setup", and "setting" (and other words that start with "set") in the document title. **Note:** You can use only prefix searches; for example, **cat*** or **set***. Suffix searches (***cat**), infix searches (**c*t**), and substring searches (***cat***) aren't supported. Also, adding a period (.) to a prefix search changes the results that are returned. That's because a period is treated as a stop word. For example, searching for **cat*** and searching for **cat.*** will return different results. We recommend not using a period in a prefix search.
( )	(fair OR free) AND (from:contoso.com) (IPO OR initial) AND (stock OR shares) (quarterly financials)	Parentheses group together Boolean phrases, `property:value` items, and keywords. For example, `(quarterly financials)` returns items that contain the words quarterly and financials.

> **Note**
>
> [1] Use this operator for properties that have date or numeric values.
> [2] Boolean search operators must be uppercase; for example, **AND**. If you use a lowercase operator, such as **and**, it is treated as a keyword in the search query.

Search conditions

You can add conditions to a search query to narrow a search and return a more refined set of results. Each condition adds a clause to the KeyQL search query that is created and run when you start the search.

- Conditions for common properties
- Conditions for mail properties
- Conditions for document properties
- Operators used with conditions
- Guidelines for using conditions
- Examples of using conditions in search queries

Conditions for common properties

Create a condition using common properties when searching mailboxes and sites in the same search. The following table lists the available properties to use when adding a condition.

Expand table

Condition	Description
Date	For email, the date a message was created or imported from a PST file. For documents, the date a document was last modified. If you're searching for email messages for a specific time period, you should use the message *Received* and *Sent* conditions if you're unsure if the email messages might have been imported instead of natively created in Exchange.
Sender/Author	For email, the person who sent a message. For documents, the person cited in the author field from Office documents. You can type more than one name, separated by commas. Two or more values are logically connected by the **OR** operator. (See Recipient Expansion)
Size (in bytes)	For both email and documents, the size of the item (in bytes).
Subject/Title	For email, the text in the subject line of a message. For documents, the title of the document. As previously explained, the Title property is metadata specified in Microsoft Office documents. You can type the name of more than one subject/title values, separated by commas. Two or more values are logically connected by the **OR** operator. **Note**: Don't include double quotation marks to the values for this condition because quotation marks are automatically added when using this search condition. If you add quotation marks to the value, two pairs of double quotations are added to the condition value, and the search query will return an error.
Retention label	For both email and documents, retention labels that can be automatically or manually applied to messages and documents. Retention labels can be used to declare records and help you manage the data lifecycle of content by enforcing retention and deletion rules specified by the label. You can type part of the retention label name and use a wildcard or type the complete label name. For more information about retention labels, see Learn about retention policies and retention labels.

Conditions for mail properties

Create a condition using mail properties when searching mailboxes or public folders in Exchange Online. The following table lists the email properties that you can use for a condition. These properties are a subset of the email properties that were previously described. These descriptions are repeated for your convenience.

Expand table

Condition	Description
Message kind	The message type to search. This is the same property as the Kind email property. Possible values:

Condition	Description
	• contacts • docs • email • externaldata • fax • im • journals • meetings • microsoftteams • notes • posts • rssfeeds • tasks • voicemail
Participants	All the people fields in an email message. These fields are From, To, Cc, and Bcc. (See Recipient Expansion)
Type	The message class property for an email item. This is the same property as the ItemClass email property. It's also a multi-value condition. So to select multiple message classes, hold the **CTRL** key and then select two or more message classes in the drop-down list that you want to add to the condition. Each message class that you select in the list is logically connected by the **OR** operator in the corresponding search query. For a list of the message classes (and their corresponding message class ID) that are used by Exchange and that you can select in the **Message class** list, see Item Types and Message Classes.
Received	The date that an email message was received by a recipient. This is the same property as the Received email property.
Recipients	All recipient fields in an email message. These fields are To, Cc, and Bcc. (See Recipient Expansion)
Sent	The date that an email message was sent by the sender. This is the same property as the Sent email property.
Subject	The text in the subject line of an email message. **Note**: Don't include double quotation marks to the values for this condition because quotation marks are automatically added when using this search condition. If you add quotation marks to the value, two pairs of double quotations are added to the condition value, and the search query will return an error.
To	The recipient of an email message in the To field.

Conditions for document properties

Create a condition using document properties when searching for documents on SharePoint and OneDrive for Business sites. The following table lists the document properties that you can use for a condition. These properties are a subset of the site properties that were previously described. These descriptions are repeated for your convenience.

Expand table

Condition	Description
Author	The author field from Office documents, which persists if a document is copied. For example, if a user creates a document and the emails it to someone else who then uploads it to SharePoint, the document will still retain the original author.
Title	The title of the document. The Title property is metadata that's specified in Office documents. It's different than the file name of the document.
Created	The date that a document is created.
Last modified	The date that a document was last changed.

Condition	Description
File type	The extension of a file; for example, docx, one, pptx, or xlsx. This is the same property as the FileExtension site property. **Note:** If you include a File type condition using the **Equals** or **Equals any of** operator in a search query, you can't use a prefix search (by including the wildcard character (*) at the end of the file type) to return all versions of a file type. If you do, the wildcard is ignored. For example if you include the condition `Equals any of doc*`, only files with an extension of `.doc` are returned. Files with an extension of `.docx` won't be returned. To return all versions of a file type, used the *property:value* pair in a keyword query; for example, `filetype:doc*`.

Operators used with conditions

When you add a condition, you can select an operator that is relevant to type of property for the condition. The following table describes the operators that are used with conditions and lists the equivalent that is used in the search query.

Expand table

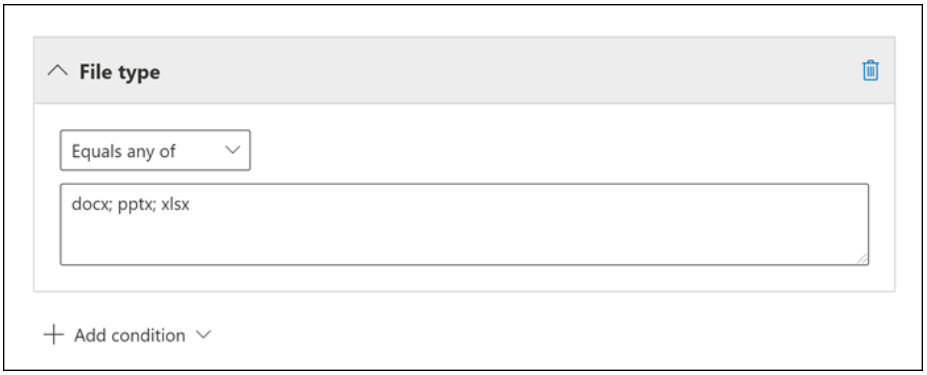
Operator	Query equivalent	Description
After	`property>date`	Used with date conditions. Returns items that were sent, received, or modified after the specified date.
Before	`property<date`	Used with date conditions. Returns items that were sent, received, or modified before the specified date.
Between	`date..date`	Use with date and size conditions. When used with a date condition, returns items there were sent, received, or modified within the specified date range. When used with a size condition, returns items whose size is within the specified range.
Contains any of	`(property:value) OR (property:value)`	Used with conditions for properties that specify a string value. Returns items that contain any part of one or more specified string values.
Doesn't contain any of	`-property:value` `NOT property:value`	Used with conditions for properties that specify a string value. Returns items that don't contain any part of the specified string value.
Doesn't equal any of	`-property=value` `NOT property=value`	Used with conditions for properties that specify a string value. Returns items that don't contain the specific string.
Equals	`size=value`	Returns items that are equal to the specified size.[1]
Equals any of	`(property=value) OR (property=value)`	Used with conditions for properties that specify a string value. Returns items that are a match of one or more specified string values.
Greater	`size>value`	Returns items where the specified property is greater than the specified value.[1]
Greater or equal	`size>=value`	Returns items where the specified property is greater than or equal to the specified value.[1]
Less	`size<value`	Returns items that are greater than or equal to the specific value.[1]
Less or equal	`size<=value`	Returns items that are greater than or equal to the specific value.[1]
Not equal	`size<>value`	Returns items that don't equal the specified size.[1]

> **Note**
>
> [1] This operator is available only for conditions that use the Size property.

Guidelines for using conditions

Keep the following in mind when using search conditions.

- A condition is logically connected to the keyword query (specified in the keyword box) by the **AND** operator. That means that items have to satisfy both the keyword query and the condition to be included in the results. This is how conditions help to narrow your results.
- If you add two or more unique conditions to a search query (conditions that specify different properties), those conditions are logically connected by the **AND** operator. That means only items that satisfy all the conditions (in addition to any keyword query) are returned.
- If you add more than one condition for the same property, those conditions are logically connected by the **OR** operator. That means items that satisfy the keyword query and any one of the conditions are returned. So, groups of the same conditions are connected to each other by the **OR** operator and then sets of unique conditions are connected by the **AND** operator.
- If you add multiple values (separated by commas or semi-colons) to a single condition, those values are connected by the **OR** operator. That means items are returned if they contain any of the specified values for the property in the condition.
- Any condition that uses an operator with **Contains** and **Equals** logic will return similar search results for simple string searches. A simple string search is a string in the condition that doesn't include a wildcard). For example, a condition that uses **Equals any of** will return the same items as a condition that uses **Contains any of**.
- The search query that is created by using the keywords box and conditions is displayed on the **Search** page, in the details pane for the selected search. In a query, everything to the right of the notation `(c:c)` indicates conditions that are added to the query. `(c:c)` shouldn't be used in manually entered queries and isn't equal to **AND** or **OR**.
- Conditions only add properties to the search query; they don't add operators. This is why the query displayed in the detail pane doesn't show operators to the right of the `(c:c)` notation. KeyQL adds the logical operators (according to the previously explained rules) when the executing the query.
- You can use the drag and drop control to resequence the order of conditions. Select the control for a condition and move it up or down.
- As previously explained, some condition properties allow you to type multiple values (separated by semi-colons). Each value is logically connected by the **OR** operator, and results in the query `(filetype=docx) OR (filetype=pptx) OR (filetype=xlsx)`. The following illustration shows an example of a condition with multiple values.



> **Note**
>
> You can't add multiple conditions (by selecting **Add condition** for the same property). Instead, you have to provide multiple values for the condition (separated by semi-colons), as shown in the previous example.

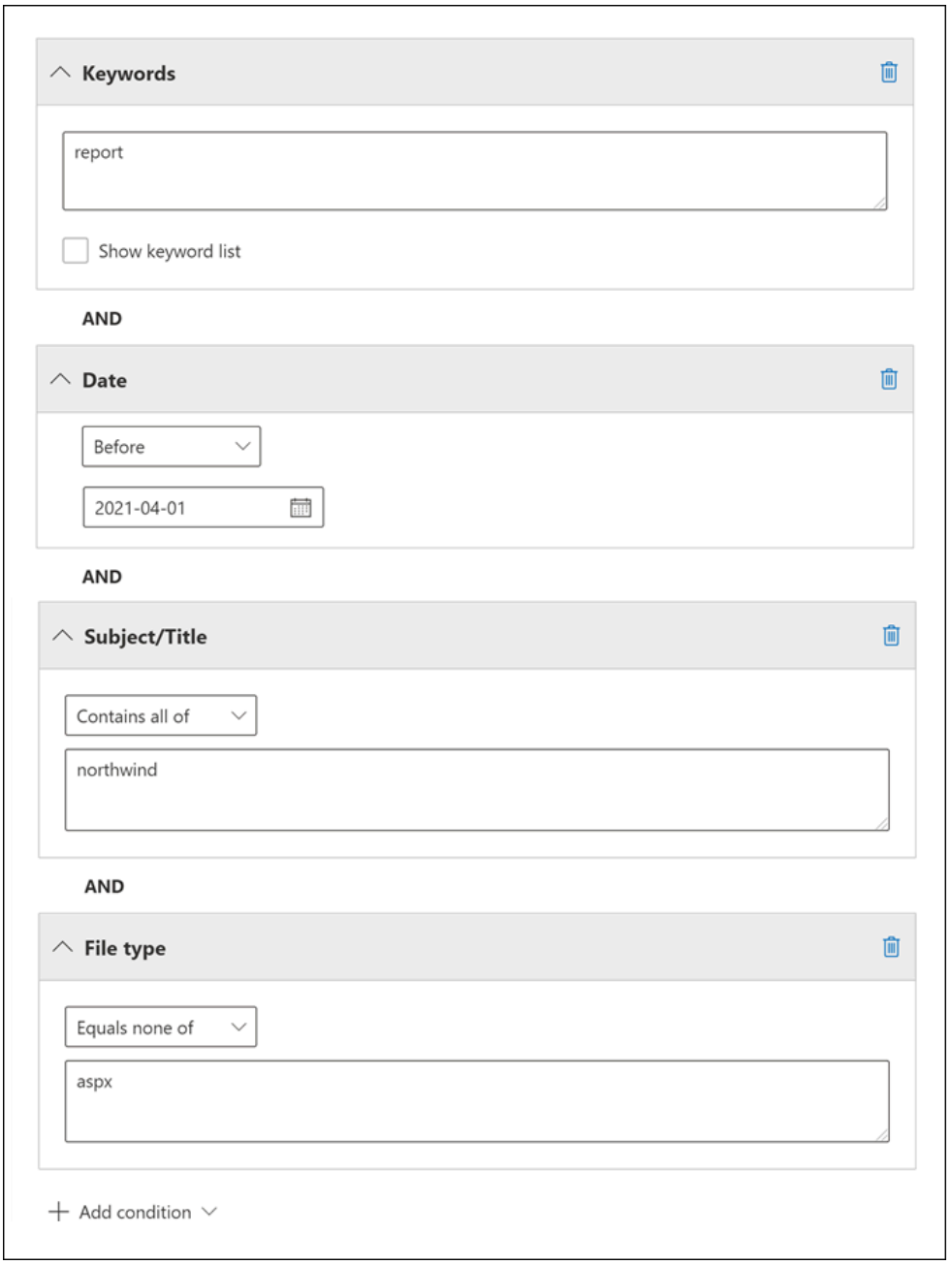
Examples of using conditions in search queries

The following examples show the GUI-based version of a search query with conditions, the search query syntax that is displayed in the details pane of the selected search (which is also returned by the **Get-ComplianceSearch** cmdlet), and the logic of the corresponding KeyQL query.

Example 1

This example returns email items or documents that contain the keyword "report", that were sent or created before April 1, 2021, and that contain the word "northwind" in the subject field of email messages or in the title property of documents. The query excludes Web pages that meet the other search criteria.

GUI:



Search query syntax:

```
report(c:c)(date<2021-04-01)(subjecttitle:"northwind")(-filetype:aspx)
```

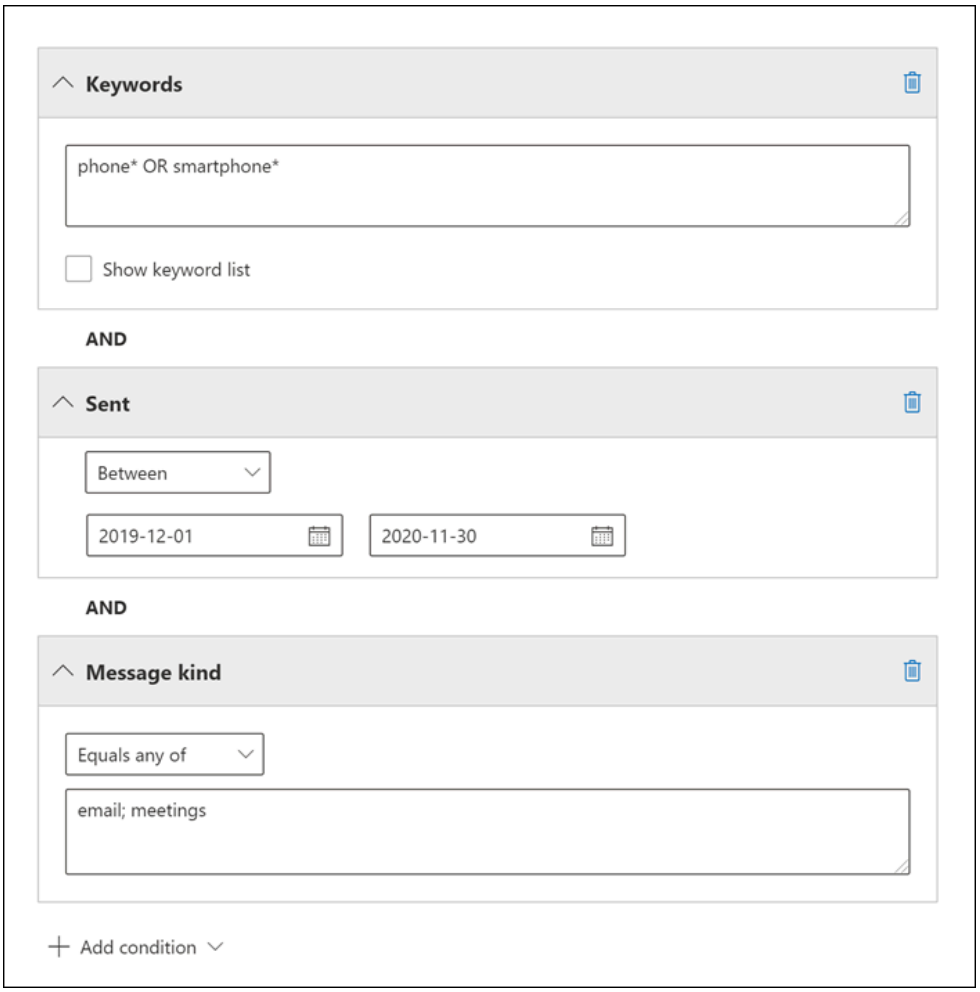
Search query logic:

```
report AND (date<2021-04-01) AND (subjecttitle:"northwind") NOT (filetype:aspx)
```

Example 2

This example returns email messages or calendar meetings that were sent between December 1, 2019 and November 30, 2020 and that contain words that start with "phone" or "smartphone".

GUI:



Search query syntax:

```
phone* OR smartphone*(c:c)(sent=2019-12-01..2020-11-30)(kind="email")(kind="meetings")
```

Search query logic:

```
phone* OR smartphone* AND (sent=2019-12-01..2020-11-30) AND ((kind="email") OR (kind="meetings"))
```

Special characters

Some special characters aren't included in the search index and therefore aren't searchable. This also includes the special characters that represent search operators in the search query. Here's a list of special characters that are either replaced by a blank space in the actual search query or cause a search error.

```
+ - = : ! @ # % ^ & ; _ / ? ( ) [ ] { }
```

Searchable sensitive data types

You can use eDiscovery search tools in the Microsoft Purview portal to search for sensitive data, such as credit card numbers or social security numbers, that is stored in documents on SharePoint and OneDrive for Business sites. You can do this by using the `SensitiveType` property and the name (or ID) of a sensitive information type in a keyword query. For example, the query `SensitiveType:"Credit Card Number"` returns documents that contain a credit card number. The query `SensitiveType:"U.S. Social Security Number (SSN)"` returns documents that contain a U.S. social security number.

To see a list of the sensitive information types that you can search for, go to **Data classifications** > **Sensitive info types** in the Microsoft Purview portal. Or you can use the **Get-DlpSensitiveInformationType** cmdlet in Security & Compliance PowerShell to display a list of sensitive information types.

Limitations for searching sensitive data types

- To search for custom sensitive information types, you have to specify the ID of the sensitive information type in the `SensitiveType` property. Using the name of a custom sensitive information type (as shown in the example for built-in sensitive information types in the previous section) will return no results. Use the **Publisher** column on the **Sensitive info types** page in the Microsoft Purview portal (or the **Publisher** property in PowerShell) to differentiate between built-in and custom sensitive information types. Built-in sensitive data types have a value of `Microsoft Corporation` for the **Publisher** property.

  To display the name and ID for the custom sensitive data types in your organization, run the following command in Security & Compliance PowerShell:

  ```
  PowerShell
  ```

  ```
  Get-DlpSensitiveInformationType | Where-Object {$_.Publisher -ne "Microsoft Corporation"} | FT Name,Id
  ```

  Then you can use the ID in the `SensitiveType` search property to return documents that contain the custom sensitive data type; for example, `SensitiveType:7e13277e-6b04-3b68-94ed-1aeb9d47de37`

- You can't use sensitive information types and the `SensitiveType` search property to search for sensitive data at-rest in Exchange Online mailboxes. This includes 1:1 chat messages, 1:N group chat messages, and team channel conversations in Microsoft Teams because all of this content is stored in mailboxes. However, you can use data loss prevention (DLP) policies to protect sensitive email data in transit. For more information, see Learn about data loss prevention and Search for and find personal data.

Searching for site content shared with external users

You can also use eDiscovery search tools in the Microsoft Purview portal to search for documents stored on SharePoint and OneDrive for Business sites that have been shared with people outside of your organization. This can help you identify sensitive or proprietary information that's being shared outside your organization. You can do this by using the `ViewableByExternalUsers` property in a keyword query. This property returns documents or sites that have been shared with external users by using one of the following sharing methods:

- A sharing invitation that requires users to sign in to your organization as an authenticated user.
- An anonymous guest link, which allows anyone with this link to access the resource without having to be authenticated.

Here are some examples:

- The query `ViewableByExternalUsers:true AND SensitiveType:"Credit Card Number"` returns all items that have been shared with people outside your organization and contain a credit card number.
- The query `ViewableByExternalUsers:true AND ContentType:document AND site:"https://contoso.sharepoint.com/Sites/Teams"` returns a list of documents on all team

sites in the organization that have been shared with external users.

> **Tip**
>
> A search query such as `ViewableByExternalUsers:true AND ContentType:document` might return numerous .aspx files in the search results. To eliminate these (or other types of files), you can use the `FileExtension` property to exclude specific file types; for example `ViewableByExternalUsers:true AND ContentType:document NOT FileExtension:aspx`.

What is considered content that is shared with people outside your organization? Documents in your organization's SharePoint and OneDrive for Business sites that are shared by sending a sharing invitation or that are shared in public locations. For example, the following user activities result in content that is viewable by external users:

- A user shares a file or folder with a person outside your organization.
- A user creates and sends a link to a shared file to a person outside your organization. This link allows the external user to view (or edit) the file.
- A user sends a sharing invitation or a guest link to a person outside your organization to view (or edit) a shared file.

Issues using the ViewableByExternalUsers property

While the `ViewableByExternalUsers` property represents the status of whether a document or site is shared with external users, there are some caveats to what this property does and doesn't reflect. In the following scenarios, the value of the `ViewableByExternalUsers` property won't be updated, and the results of a search query that uses this property might be inaccurate.

- Changes to sharing policy, such as turning off external sharing for a site or for the organization. The property will still show previously shared documents as being externally accessible even though external access might have been revoked.
- Changes to group membership, such as adding or removing external users to Microsoft 365 Groups or Microsoft 365 security groups. The property won't automatically be updated for items the group has access to.
- Sending sharing invitations to external users where the recipient hasn't accepted the invitation, and therefore doesn't yet have access to the content.

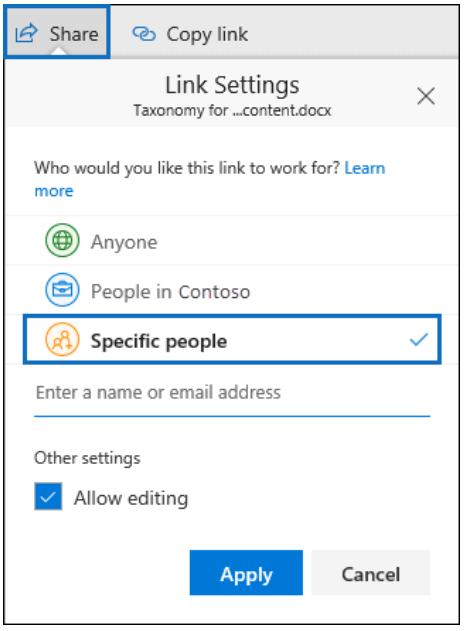
In these scenarios, the `ViewableByExternalUsers` property won't reflect the current sharing status until the site or document library is recrawled and reindexed.

Searching for site content shared within your organization

As previously explained, you can use the `SharedWithUsersOWSUser` property so search for documents that have been shared between people in your organization. When a person shares a file (or folder) with another user inside your organization, a link to the shared file appears on the **Shared with me** page in the OneDrive for Business account of the person who the file was shared with. For example, to search for the documents that have been shared with Sara Davis, you can use the query `SharedWithUsersOWSUser:"sarad@contoso.com"`. If you export the results of this search, the original documents (located in the content location of the person who shared the documents with Sara) are downloaded.

Documents must be explicitly shared with a specific user to be returned in search results when using the `SharedWithUsersOWSUser` property. For example, when a person shares a document in their OneDrive account, they have the option to share it with anyone (inside or outside the

organization), share it only with people inside the organization, or share it with a specific person. Here's a screenshot of the **Share** window in OneDrive that shows the three sharing options.



Only documents that are shared by using the third option (shared with **Specific people**) are returned by a search query that uses the `SharedWithUsersOWSUser` property.

Searching for Skype for Business conversations

You can use the following keyword query to specifically search for content in Skype for Business conversations:

PowerShell

```
kind:im
```

The previous search query also returns chats from Microsoft Teams. To prevent this, you can narrow the search results to include only Skype for Business conversations by using the following keyword query:

PowerShell

```
kind:im AND subject:conversation
```

The previous keyword query excludes chats in Microsoft Teams because Skype for Business conversations are saved as email messages with a Subject line that starts with the word "Conversation".

To search for Skype for Business conversations that occurred within a specific date range, use the following keyword query:

PowerShell

```
kind:im AND subject:conversation AND (received=startdate..enddate)
```

Character limits for searches

There's a 4,000 character limit for search queries when searching for content in SharePoint sites and OneDrive accounts. Here's how the total number of characters in the search query are calculated:

- The characters in keyword search query (including both user and filter fields) count against this limit.
- The characters in any location property (such as the URLs for all the SharePoint sites or OneDrive locations being searched) count against this limit.
- The characters in all the search permissions filters that are applied to the user running the search count against the limit.

> **Note**
>
> The 4,000 character limit applies to Content search, eDiscovery (Standard), and eDiscovery (Premium).

Last updated on 10/31/2024

Collect cloud attachments in Microsoft Purview eDiscovery (Premium)

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

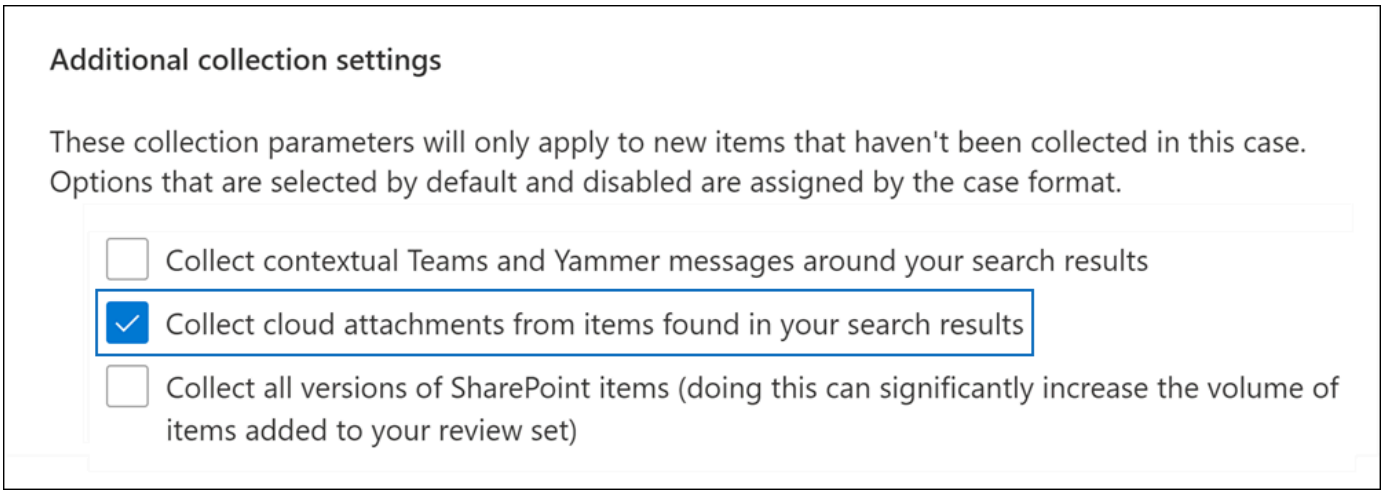
Cloud attachments are links to documents that are typically stored in SharePoint site and OneDrive. So instead of attaching an actual copy of a document in an email message or a Teams chat conversation, you have the option of sharing a link to the file. Cloud attachments are an effective way to share documents and collaborate with other people in your organization. But cloud attachments present challenges during the eDiscovery workflow because only the cloud attachment link and not the actual content in the shared document are returned in an eDiscovery search. To address this challenge, eDiscovery (Premium) provides two solutions for collecting cloud attachments:

- Collecting the live version of a document that is linked to in a cloud attachment.
- Collecting the version of the document at the time it was shared in a cloud attachment.

Collecting cloud attachments

When you create a collection estimate and the search results contain items that include cloud attachments, you have to the option of collecting the target of the cloud attachment when you commit the collection estimate to a review set. When you select this option, eDiscovery (Premium) adds the documents that are linked to in the cloud attachment to the review set. This option allows you to review the target documents and determine if the document is relevant to your case or investigation.

The following screenshot shows the option to include the targets of cloud attachments when you commit a collection to a review set.



> **Note**
>
> You have the option to include all versions (in addition to the version that was shared) of cloud attachments in the review set.

When working with collections and cloud attachments, the following applies:

- Cloud attachments aren't represented or included in collection estimates. Only direct hits are included in a collection.
- If cloud attachments are moved or no longer available at the location from where it was shared, it isn't included in a collection.
- If the cloud attachment is under retention of legal hold policy, the cloud attachment is included in collections even if deleted by a user.
- For cloud attachments that have been renamed or moved to a different container, it may or may not be included in a collection depending on how it was shared. If the cloud attachment was shared via a link and the link is still active, it's retrieved and included in the collection. If the sharing link to the location for the cloud attachment isn't available, the cloud attachment isn't retrieved or included in the collection.
- For cloud attachments with parent items placed on hold, the hold applies to cloud attachment in certain situations. This depends on the source location of the cloud attachment and the hold policy applied to that location. The policies applicable to the cloud attachment recipient won't necessarily apply to the location of the shared cloud attachment.

For instructions committing a collection to a review set, see Commit a collection estimate to a review set.

Collecting the version shared in a cloud attachment (preview)

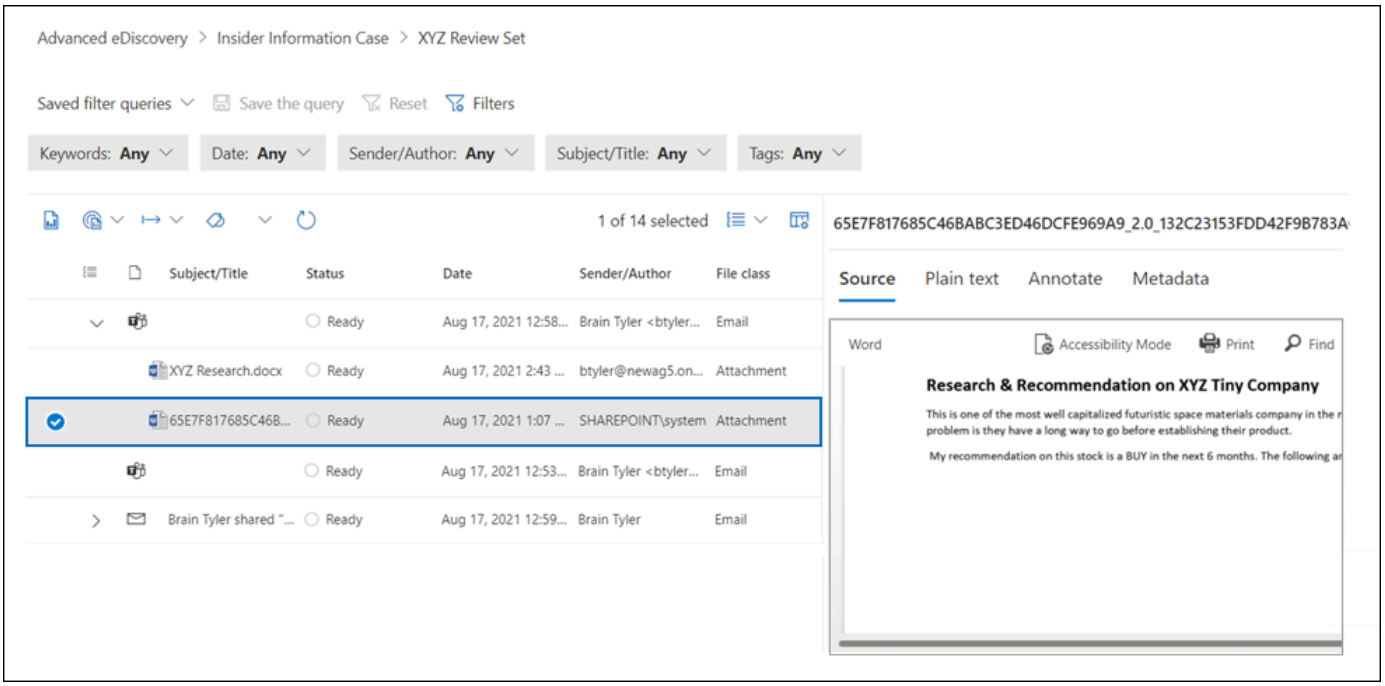
The eDiscovery (Premium) workflow for collecting cloud attachments only includes adding the most current version of a cloud attachment to a review set. This means the version that's collected and added to a review set could be different than the version that was originally shared in the cloud attachment. So it's possible that content that was present in the cloud attachment at the time it was shared might have been removed and doesn't exist in the current version that's added to the review set.

Organizations can use Microsoft Purview retention labels to preserve the version of a document at the time when it was shared as a cloud attachment. To do this, your organization can create a retention label, choose the option apply the label to cloud attachments, and then automatically apply the label to documents stored in SharePoint and OneDrive. After you set up this configuration, a copy of a document is created at the time when the file is shared. Also, if the document is modified and shared again as a cloud attachment, the modified version is also preserved. If the file is modified and shared again, a new copy of the file as a new version is preserved.

Preserving the shared versions of cloud attachments can help your organization scope the preservation and collection of potentially relevant content to the specific version of the document that was shared rather than the current live version. After you implement this retention solution, both the current live version of a cloud attachment and the version that was shared in the cloud attachment are collected and added to a review set.

For instructions on setting up a retention label and automatically applying it to cloud attachments, see Auto-apply labels to cloud attachments.

The following screenshot shows a cloud attachment document, named *XYZ Research.docx*, that was added to a review set. The document was shared as a cloud attachment in a Teams chat conversation. The review set also contains the version that was originally shared in the cloud attachment. Notice the name of this version of the cloud attachment is generated by the system and the author is identified as **SharePoint**.



Additionally, the current live version and the version that was shared have the same **FamilyId** property value, which is the same as the **FamilyId** for the parent object (such as an email message or a Teams chat conversation). This lets you group cloud attachments with the item in which they were shared.

After you've implemented the retention label and auto-apply the label to SharePoint documents, you still select the option to collect cloud attachments when committing a collection estimate to a review set. When the cloud attachments are collected, both the current live version and the version that was originally shared are added to the review set.

Last updated on 09/02/2025

Manage review sets in eDiscovery (Premium)

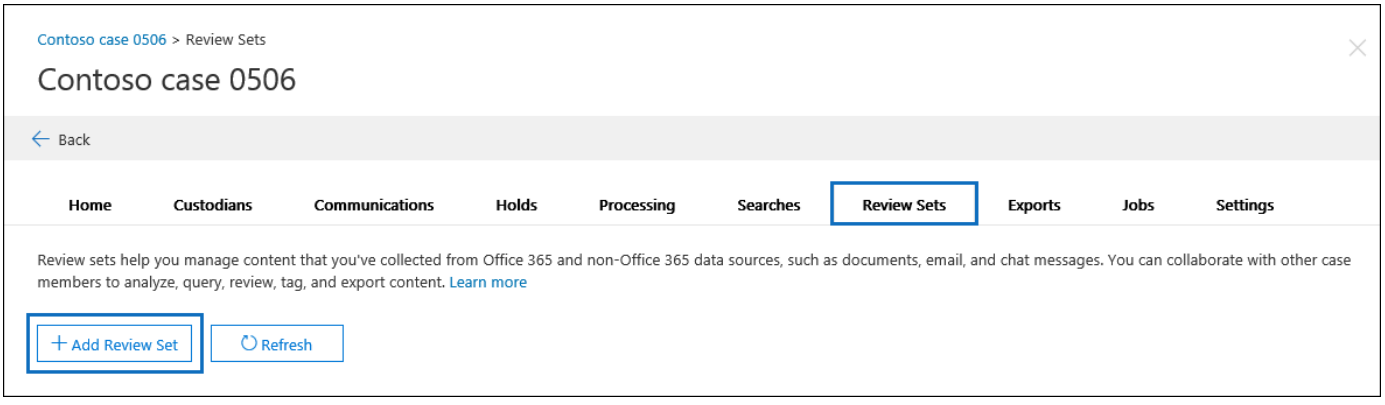
> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

Review sets are a static set of documents where you can analyze, query, view, tag, and export data in a case. For more information about performing these tasks, see:

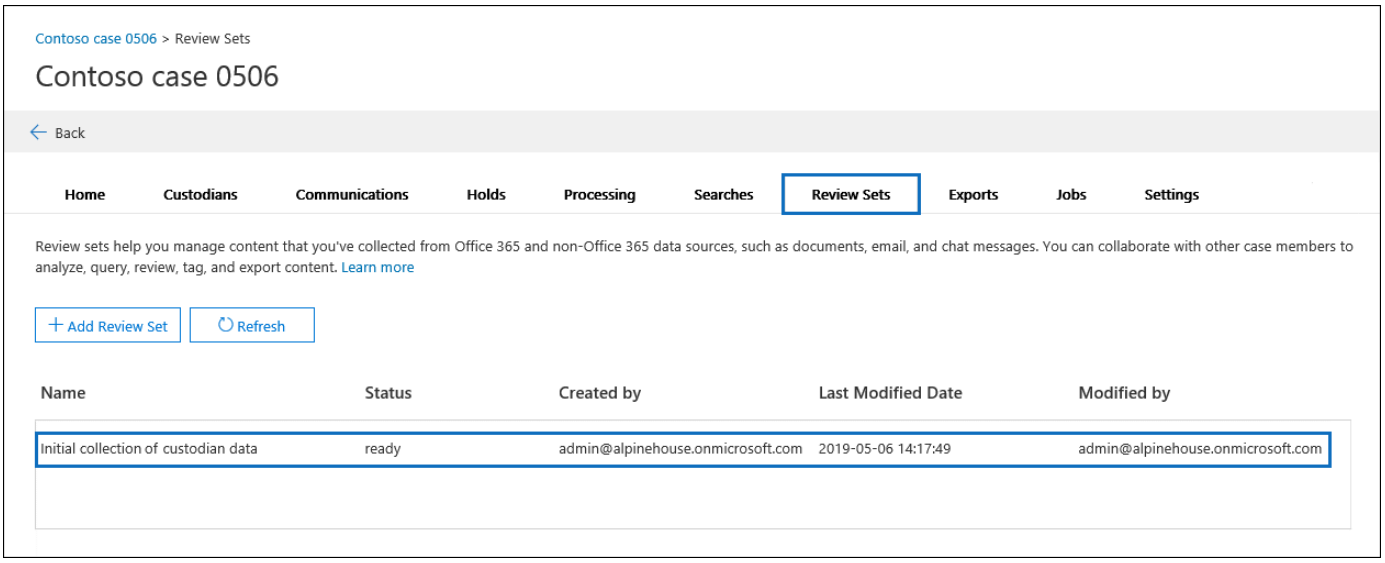
- Analyze data in a review set
- Query the data in a review set
- View documents in a review set
- Tag documents in a review set
- Export case data

Create a review set

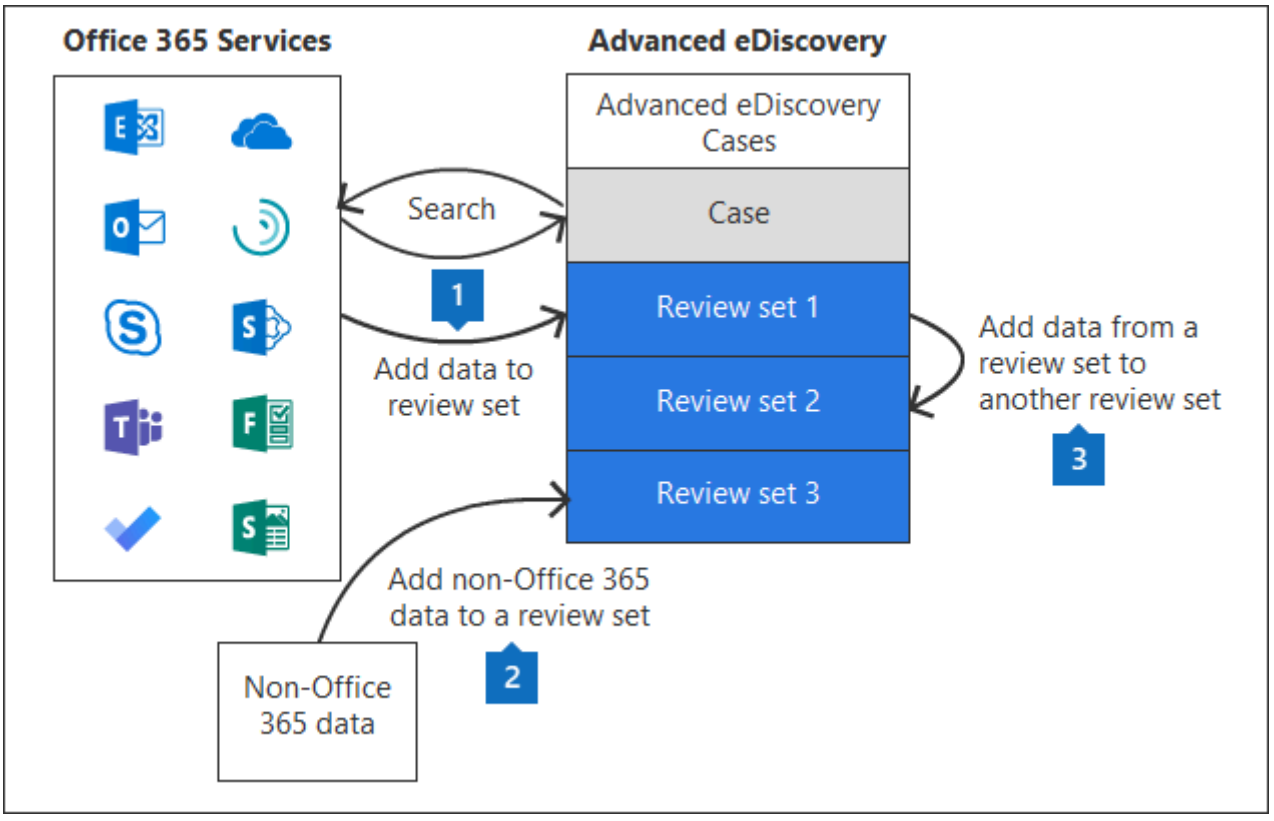
Review sets can be created on the **Review sets** tab by selecting **+ Add review set**.



On the **Add review set** flyout page, type a name for the review set and then select **Save**. The new review set is displayed in the list on the **Review sets** tab.



There are three different ways to add data to a review set in an eDiscovery (Premium) case.



1. Add search results to a review set
2. Load non-Microsoft 365 data into a review set
3. Add data to a review set from another review set

> **Note**
>
> You can't delete items from a review set and you can't delete review sets from a case. To delete a review set (and delete the data in it), you have to delete the eDiscovery (Premium)

case the review set is located in. For more information, see Close or delete an eDiscovery (Premium) case.

Last updated on 11/21/2025

Load non-Microsoft 365 data into a review set

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

Not all documents that you need to analyze in Microsoft Purview eDiscovery (Premium) are in Microsoft 365. With the non-Microsoft 365 data import feature in eDiscovery (Premium), you can upload documents that aren't in Microsoft 365 to a review set. This article shows you how to bring your non-Microsoft 365 documents into eDiscovery (Premium) for analysis.

Requirements to upload non-Office 365 content

Using the upload non-Microsoft 365 feature described in this article requires that you've the following:

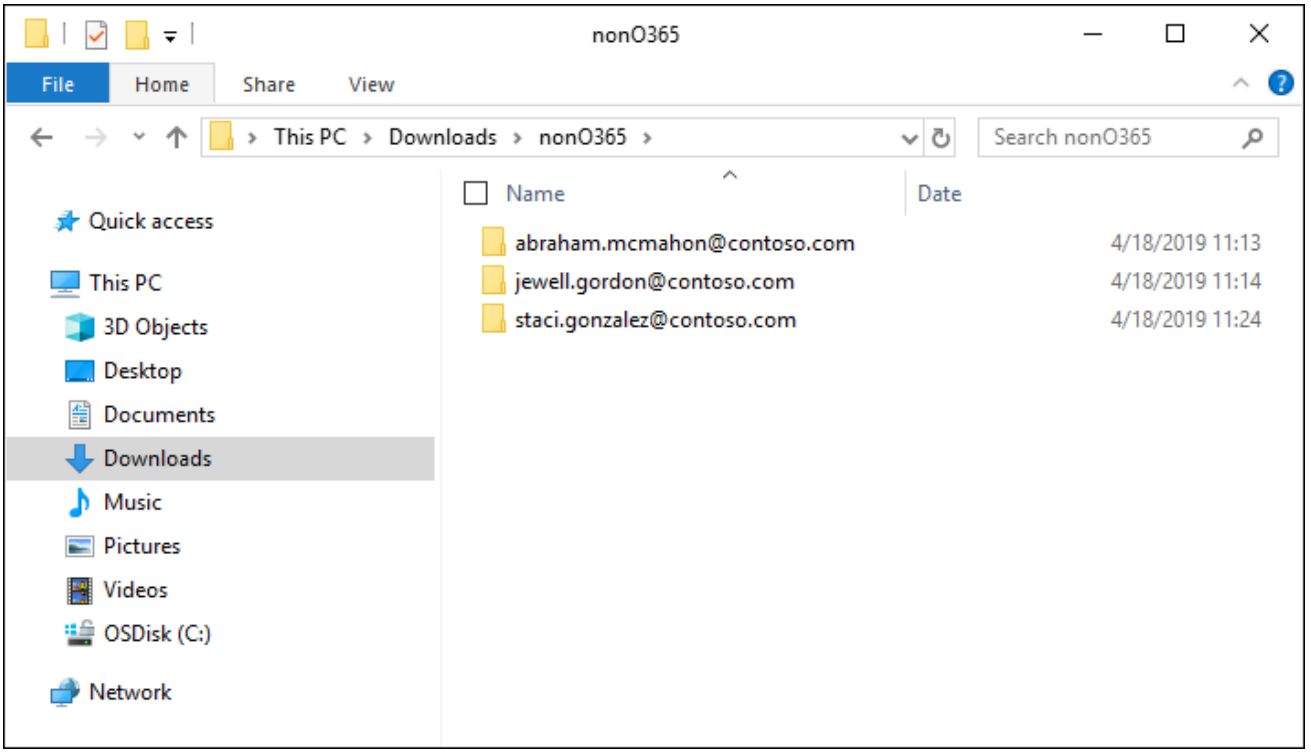
- All custodians that you want to associate non-Microsoft 365 content to must have the appropriate license. For more information, see Get started with eDiscovery (Premium).
- An existing eDiscovery (Premium) case.
- Custodians must be added to the case before you can upload and associate the non-Microsoft 365 data to them.
- Non-Microsoft 365 data must be a file type that's supported by eDiscovery (Premium). For more information, see Supported file types in eDiscovery (Premium).
- All files that are uploaded to a review set must be located in folders, where each folder is associated with a specific custodian. The names for these folders must use the following naming format: *alias@domainname*. The alias@domainname must be the user's Microsoft

365 alias and domain. You can collect all the alias@domainname folders in a root folder. The root folder can only contain the alias@domainname folders. Loose files in the root folder aren't supported.

The folder structure for the non-Microsoft 365 data that you want to upload would be similar to the following example:

- c:\nonO365\abraham.mcmahon@contoso.com
- c:\nonO365\jewell.gordon@contoso.com
- c:\nonO365\staci.gonzalez@contoso.com

Where abraham.mcmahon@contoso.com, jewell.gordon@contoso.com, and staci.gonzalez@contoso.com are the SMTP addresses of custodians in the case.



- An account that is assigned to the eDiscovery Manager role group (and added as eDiscovery Administrator).

- The AzCopy tool installed on a computer that has access to the non-Microsoft 365 content folder structure. To install AzCopy, see Get started with AzCopy. Be sure to install AzCopy in the default location, which is **%ProgramFiles(x86)%\Microsoft SDKs\Azure\AzCopy**. You should use the latest version of AzCopy. Other versions of AzCopy may not work when loading non-Microsoft 365 data in eDiscovery (Premium).

Upload non-Microsoft 365 content into eDiscovery (Premium)

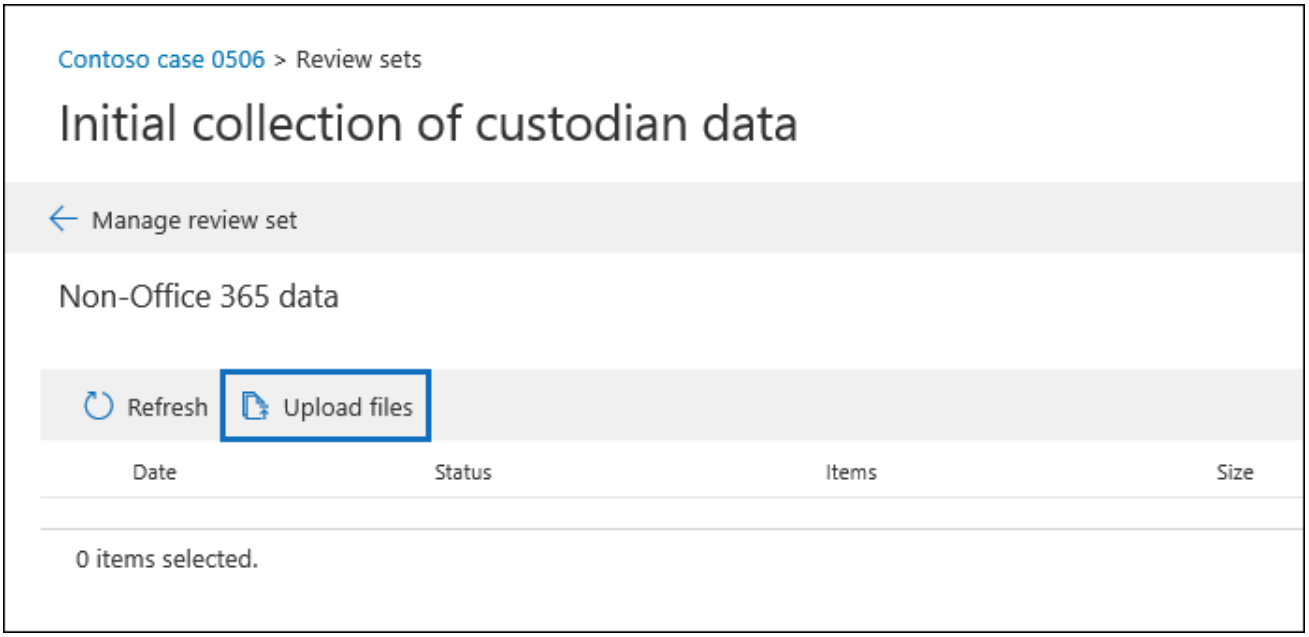
1. As an eDiscovery Manager or eDiscovery Administrator, open eDiscovery (Premium), and go to the case that the non-Microsoft 365 data will be uploaded to.

2. Select **Review sets**, and then select the review set to upload the non-Microsoft 365 data to. If you don't have a review set, you can create one.

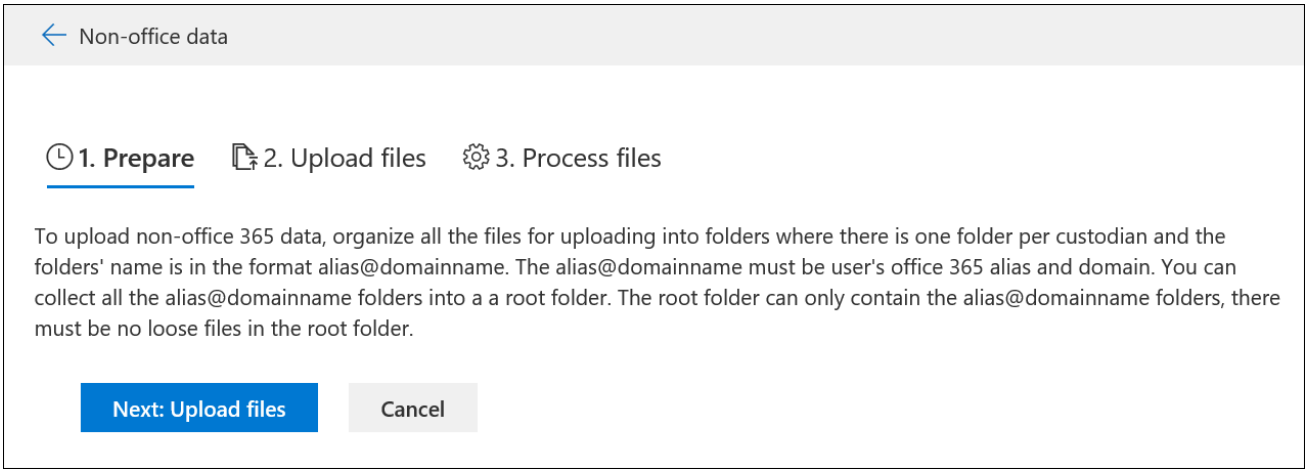
3. Open the review set by either selecting on it or selecting it and selecting **Open review set**.

4. In the review set, select **Manage review set** (the down arrow just after the **Actions** option), and then select the **Non-Office 365 data** option.

5. Select **Upload files** to start the data import workflow.

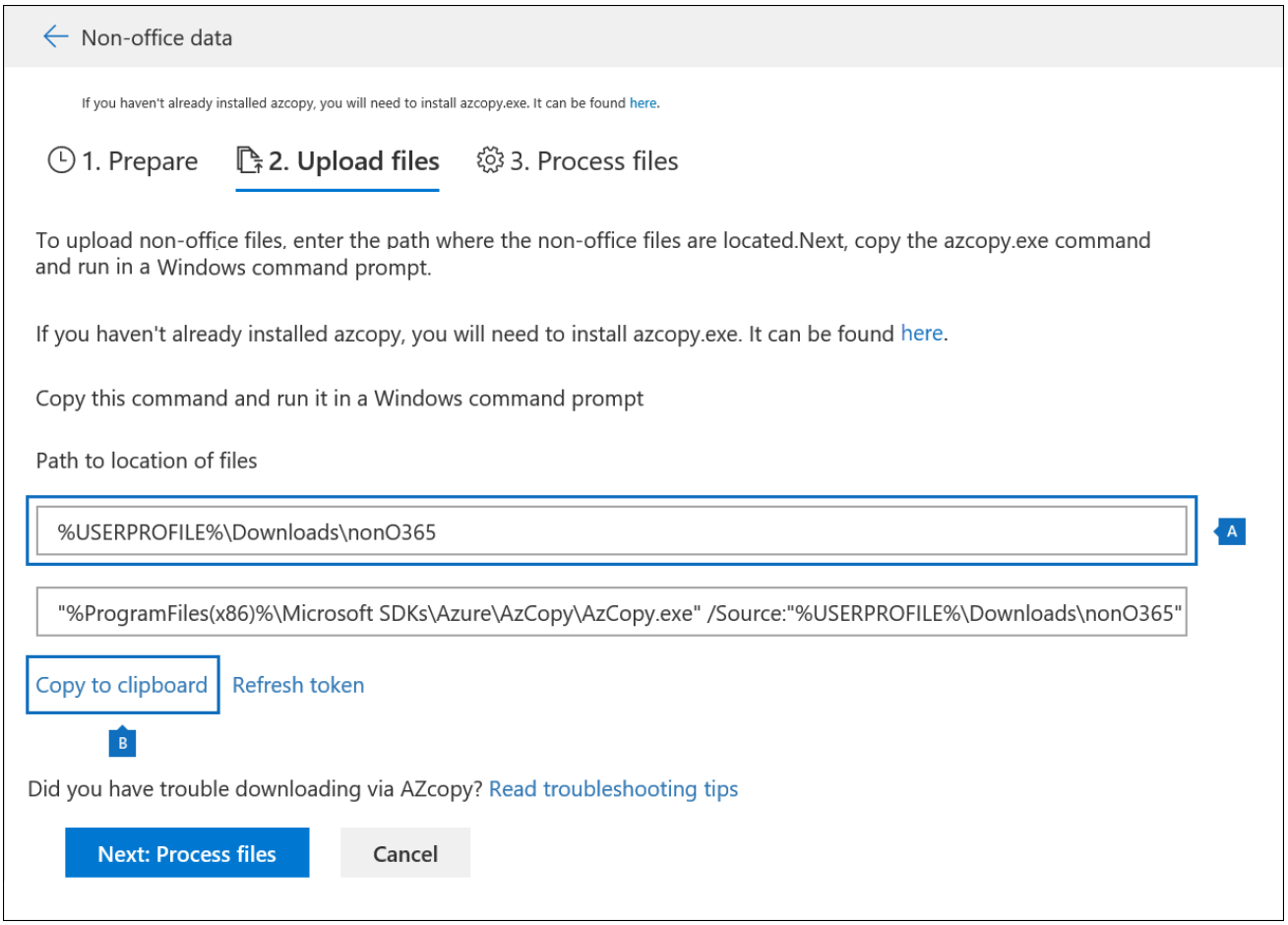


The first step in the workflow prepares a secure Microsoft-provided Azure Storage location to upload the files to. When the preparation is completed, **Next: Upload files** becomes active.



6. Select **Next: Upload files**.

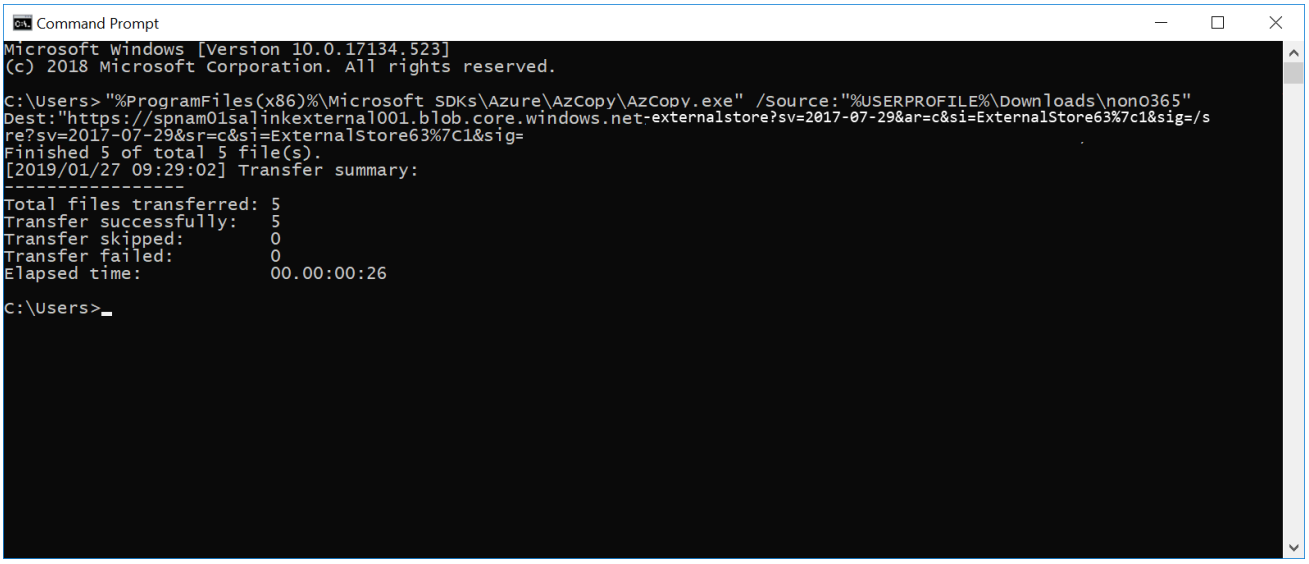
7. On the **Upload files** page, do the following:



a. In the **Path to location of files** box, verify or type the location of the root folder where you've stored the non-Microsoft 365 data you want to upload. For example, for the location of the example files shown in the **Before you begin section**, you would type **%USERPROFILE\Downloads\nonO365**. Providing the correct location ensures the AzCopy command displayed in box under the path is properly updated.

b. Select **Copy to clipboard** to copy the command that is displayed in the box.

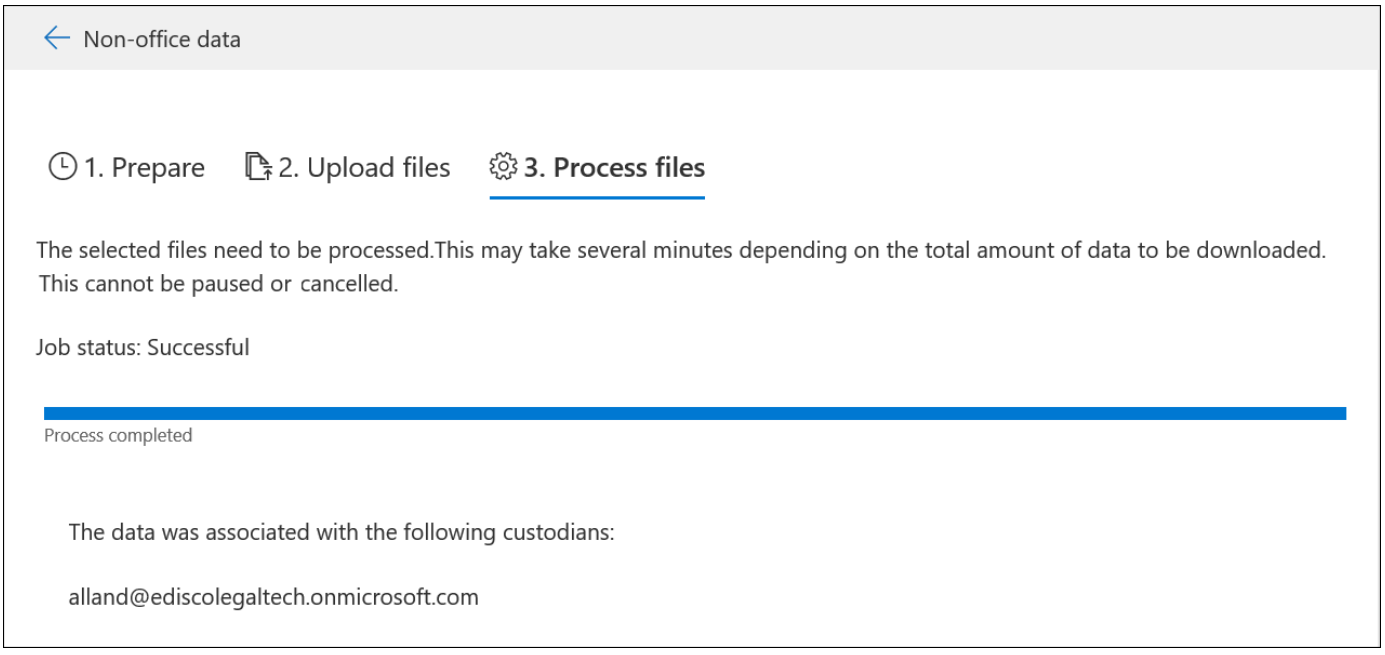
8. Start a Windows command prompt, paste the command that you copied in the previous step, and then press **Enter** to start the AzCopy command. After you start the command, the non-Microsoft 365 files will be uploaded to the Azure Storage location that was prepared in step 4.



> **Note**
>
> You should use the latest version of AzCopy to successfully use the command that's provided on the **Upload files** page. If the supplied AzCopy command fails, see Troubleshoot AzCopy in eDiscovery (Premium).

9. Go back to the Microsoft Purview portal, and select **Next: Process files** in the workflow. This initiates processing, text extraction, and indexing of the non-Microsoft 365 files that were uploaded to the Azure Storage location.

10. Track the progress of processing the files on the **Process files** page or on the **Jobs** tab by viewing a job named **Adding non-Microsoft 365 data to a review set**. After the job is finished, the new files will be available in the review set.



11. After the processing is finished, you can close the workflow.

Last updated on 11/21/2025

Add data to a review set from another review set

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

In some cases, it may be necessary to select documents from one review set and work with them individually in another review set. This is especially useful if you've culled content in a review set and want to run analytics on the subset of data.

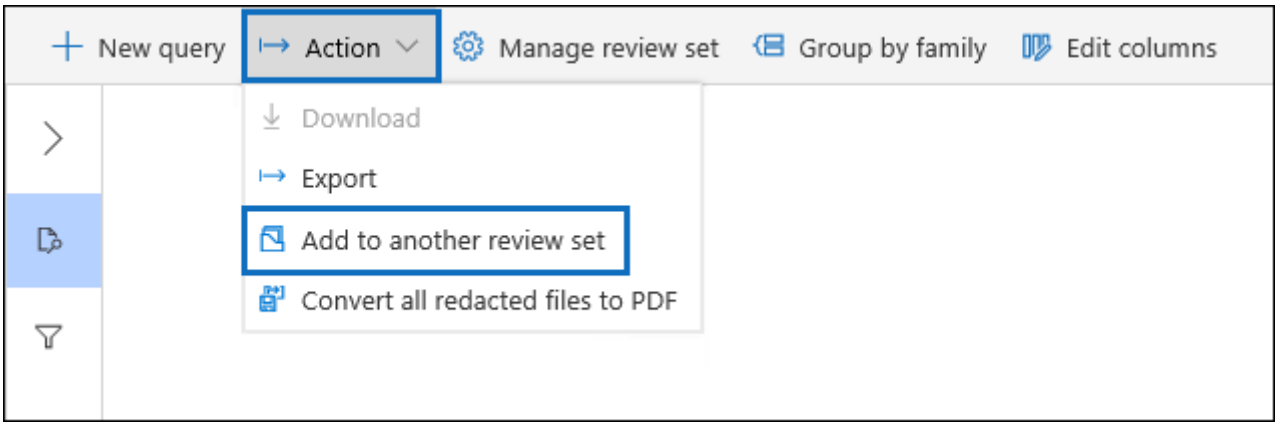
Follow the workflow in this article to add content from one review set to another.

Create a review set

Before you start, you'll need to create a review set to add the data to. A new review set can be added on the **Review sets** tab of the case. For more information, see Create a review set.

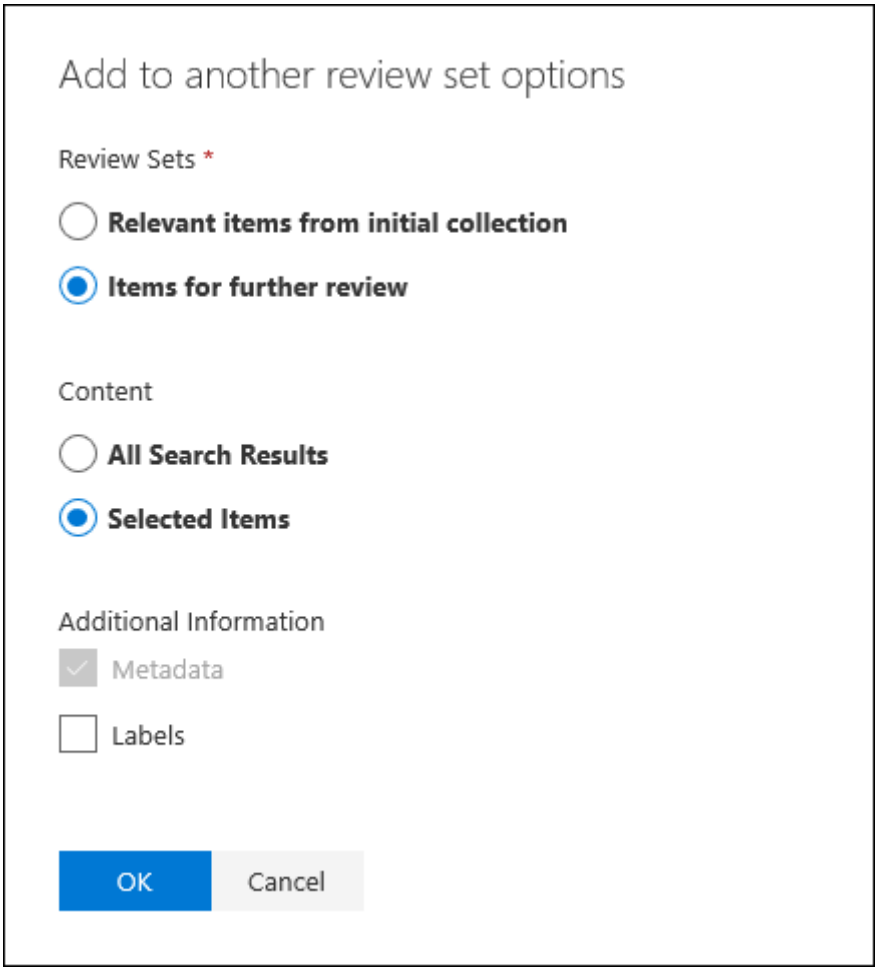
Step 1: Identify content to add to another review set

You can add content from one review set to another one by selecting specific documents in the source review set or by selecting all items returned by review set query. If you're adding selected items, select the items, select **Action**, and then select **Add to another review set**.



Step 2: Specify options for adding to another review set

In the **Add to another review set options** flyout page, choose the review set you want to add the items to. Choose whether to add **All search results** or **Selected items**. **Additional information** provides options to include all metadata from the items and whether to include the tags (by selecting the **Labels** check box) from the source review set when the documents are added to the new review set.



After you select **Ok**, a new job (named **Adding data to another review set**) is created to add the content to another review set. You can go to the **Jobs** tab and monitor the progress of this job. For more information, see Manage jobs.

Last updated on 11/21/2025

Configure review set grouping settings for eDiscovery (Premium) cases

> **⊗ Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

You can configure grouping settings for each Microsoft Purview eDiscovery (Premium) case to control how the data in a review set is grouped and displayed. Turning on the **Enable group** setting is only available for review sets in cases created in your organization after March 15, 2023.

Depending when the case is created, items in a review set can be grouped in two ways:

- **Group review set items by Group ID and Thread ID**: When enabled, review set items are grouped by *Group ID* and *Thread ID*. This setting is the default for all cases created after March 15, 2023.
- **Group review set items by Family ID and Conversation ID**: When disabled, review set items are grouped by *Family ID* and *Conversation ID*. This setting is the default for all cases created before March 15, 2023.

For more information about review set grouping, see Group and view documents in a review set.

Configure review set grouping settings for a case

To configure review set grouping settings for a case:

1. On the **eDiscovery (Premium)** page, select the case.
2. On the **Settings** tab, under **Review sets**, choose **Select**.
3. On the **Review sets** page, select the **Enable group** toggle to enable grouping review set items by *Group ID* and *Thread ID*.

> **Note**
>
> For cases created before March 15, 2023, the **Enable group** toggle is disabled. Item grouping in review sets for these cases is based on *Family ID* and *Conversation ID*.

Last updated on 11/21/2025

Query and filter content in a review set

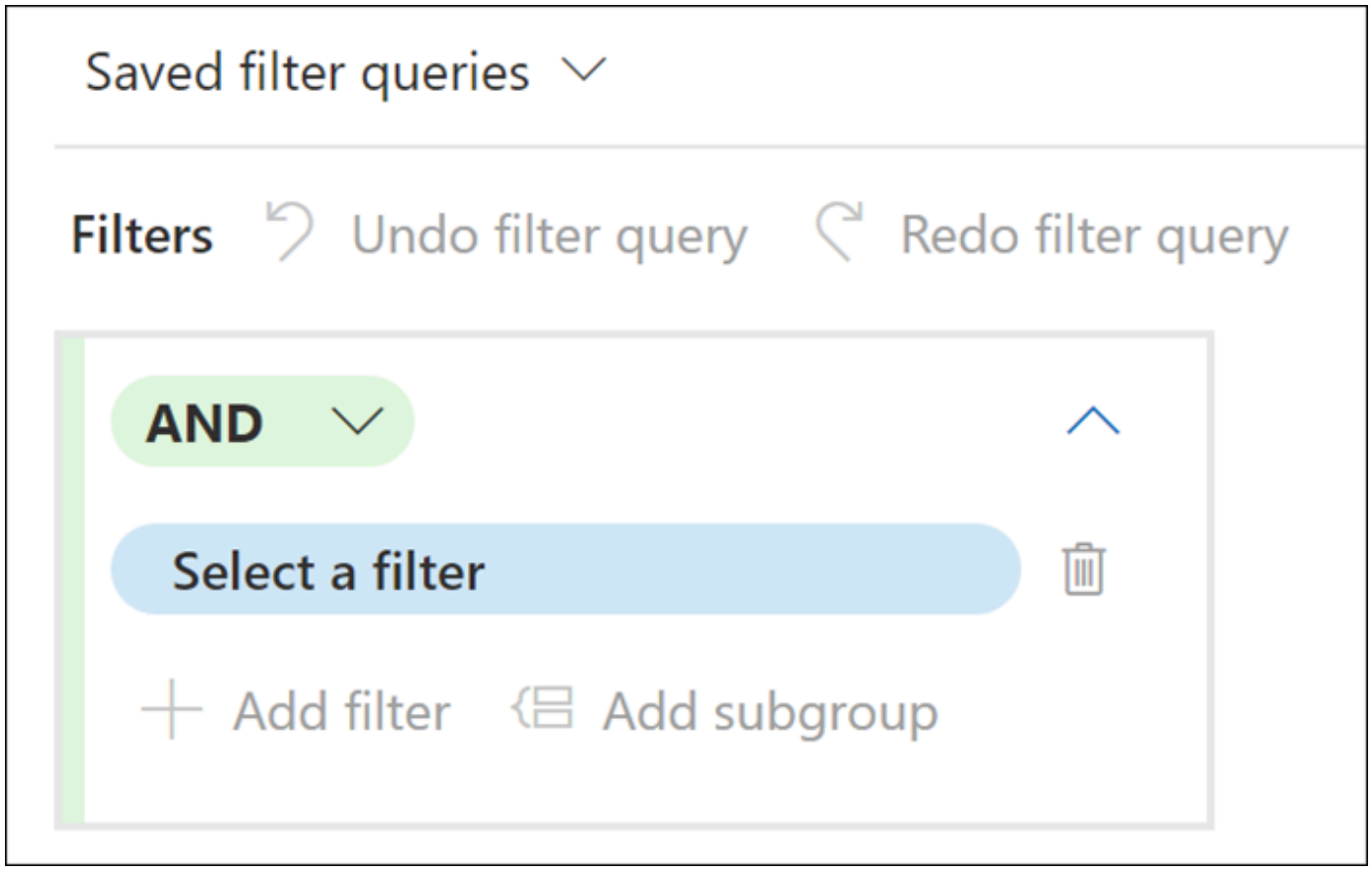
> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic **Content Search**, classic **eDiscovery (Standard)**, and classic **eDiscovery (Premium)**.
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the **new eDiscovery experience** in the **Microsoft Purview portal**.

In most cases, it's useful to dig deeper into the content in a review set and organize it to facilitate a more efficient review. Using filters and queries in a review set helps you focus on a subset of documents that meet the criteria of your review. For more information about available search conditions, see Keyword queries and search conditions for eDiscovery.

Advanced filters (preview)

Improving on the default filters in previous versions, eDiscovery (Premium) now provides advanced filters that let you build more flexible and advanced filters for review sets. Similar to the collection query builder experience, the advanced filtering capability enables you to:

- Quickly search for filter conditions.
- Create complex filters using subgroups, *AND*, or *OR* conditions.
- Easily change your queries with **Undo filter query** and **Redo filter query** controls.
- Manage saved filters without having to navigate to another area.
- Use *Is empty* and *Is not empty* conditions for each filter.



> **Important**
>
> A review set displays a maximum of 1,000 items per page and up to 10 pages (for a total of 10,000 items displayed per review set). Use default or custom filters to adjust the displayed items as needed. Matched item counts are estimates.

Advanced filter controls (preview)

To create and custom filtering for your review set, use the following controls:

- **AND/OR**: These conditional logical operators allow you to select the query condition that applies to specific filters and filter subgroups. These operators allow you to use multiple filters or subgroups connected to a single filter in your query.
- **Select a filter**: Allows you to select filters for the specific data sources and location content selected for the collection.
- **Add filter**: Allows you to add multiple filters to your query. Is available after you've defined at least one query filter.
- **Select an operator**: Depending on the selected filter, the operators compatible for the filter are available to select. For example, if the *Date* filter is selected, the available operators are *Before*, *After*, and *Between*. If the *Size (in bytes)* filter is selected, the available operators are *Greater than*, *Greater or equal*, *Less than*, *Less or equal*, *Between*, and *Equal*.

- **Value**: Depending on the selected filter, the values compatible for the filter are available. Additionally, some filters support multiple values and some filters support one specific value. For example, if the *Date* filter is selected, select date values. If the *Size (in bytes)* filter is selected, select a value for bytes.
- **Add subgroup**: After you've defined a filter, you can add a subgroup to refine the results returned by the filter. You can also add a subgroup to a subgroup for multi-layered query refinement.
- **Remove a filter condition**: To remove an individual filter or subgroup, select the remove icon to the right of each filter line or subgroup.
- **Clear all**: To clear the entire query of all filters and subgroups, select **Clear all**.

Filter types

Every searchable field in a review set has a corresponding filter that you can use for filter items based on a specific field.

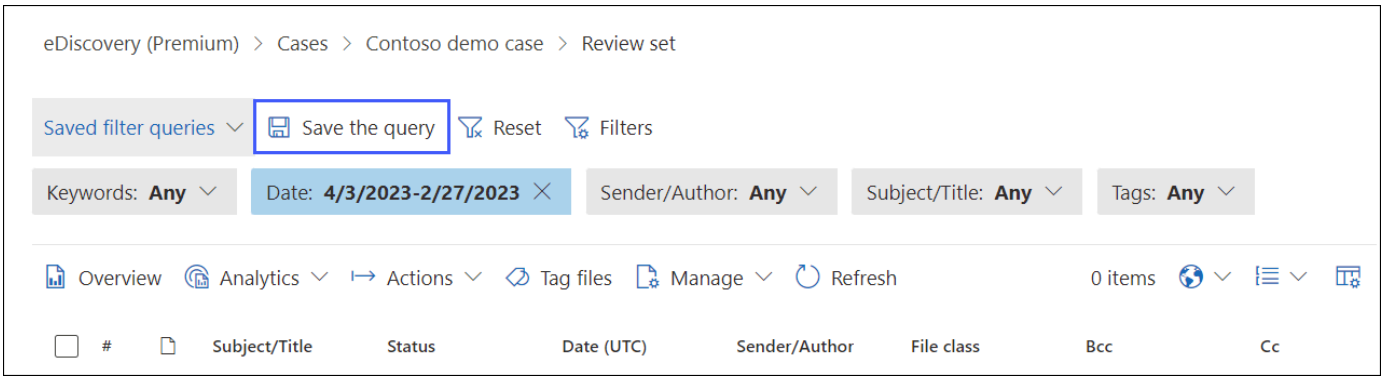
There are multiple types of filters:

- **Freetext**: A freetext filter is applied to text fields such as *Subject*. You can list multiple search terms by separating them with a comma.
- **Date**: A date filter is used for date fields such as *Last modified date*.
- **Search options**: A search options filter provides a list of possible values (each value is displayed with a checkbox that you can select) for particular fields in the review. This filter is used for fields, such as *Sender*, where there's a finite number of possible values in the review set.
- **Keyword**: A keyword condition is a specific instance of freetext condition that you can use to search for terms. You can also use KeyQL-like query language in this type of filter. For more information, see the Query language and Advanced query builder sections in this article. Keyword searching in a review set only applies to document/email content and does not extend to metadata fields (such as *Filename*).

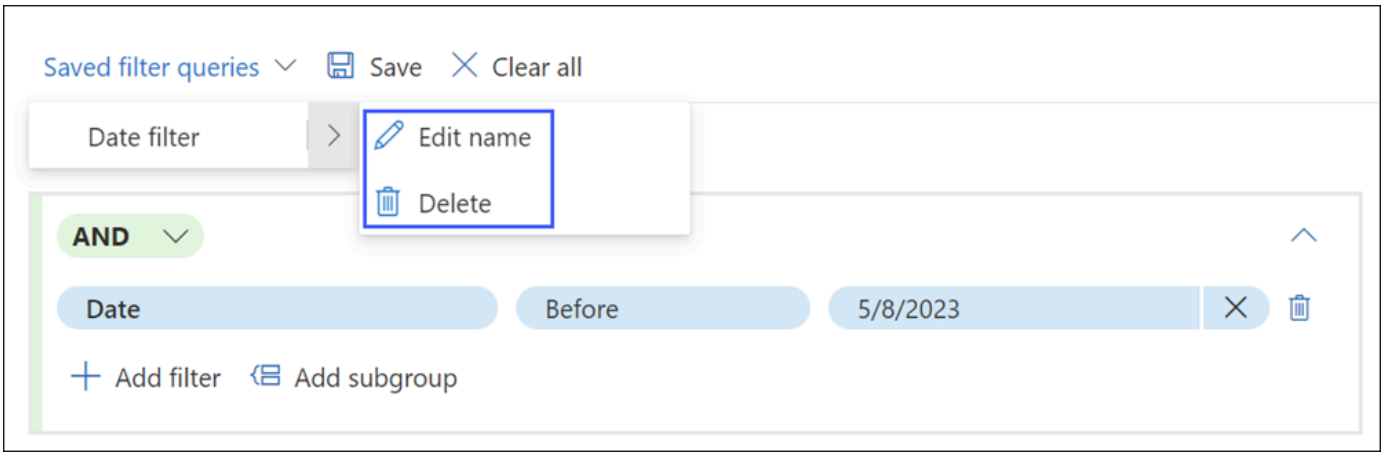
Save and manage filter queries

After you're satisfied with your filters, you can save the filter combination as a filter query. This saved filter query lets you apply the filter in the future review sessions.

To save a filter, select **Save** on the **Save filter queries** command bar and name it. You or other reviewers can run previously saved filter queries by selecting the **Saved filter queries** dropdown and selecting a filter query to apply to review set documents.



To edit or delete a saved filter query, select **Saved filter queries** and expand the filter properties to display the **Edit** and **Delete** options for the saved filter query.



Use query language support for KeyQL and Keyword filters

When using the *KeyQL* or *Keyword* filters, you can use a KeyQL-like query language to build your review set search query. The query language for these two filters supports standard Boolean operators, such as **AND**, **OR**, **NOT**, and **NEAR**. It also supports a single-character wildcard (?) and a multi-character wildcard (*).

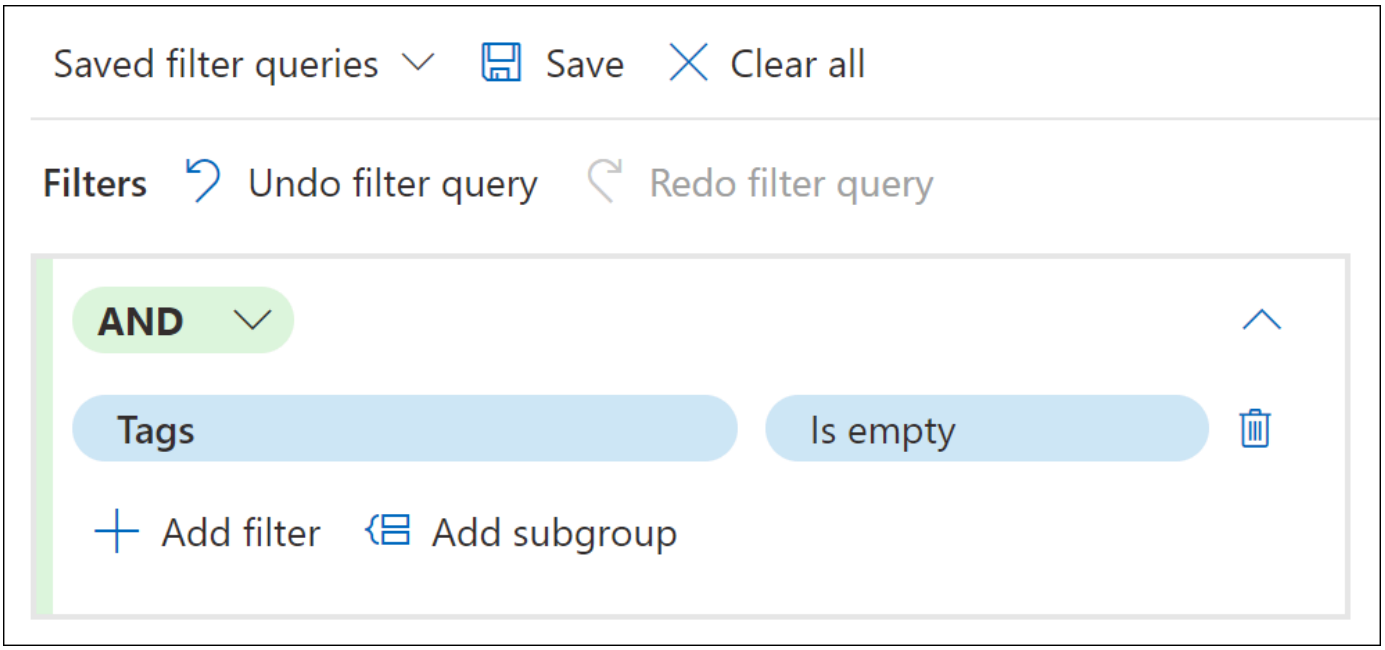
> **Note**
>
> Review filters only support wildcards (? or *) on a single term. Using wildcards in searches on phrases that consist of multiple terms aren't supported.

Scenario example: Filter for untagged items in a review set

An eDiscovery administrator needs to create a query to find all items in the review set that haven't had any tagging applied. For this example, the administrator creates the following review set filter query:

1. For the first filter, the administrator selects the filter and types *tag* in filter search. The filter *Tags* is displayed as a matching option, and the administrator selects it.
2. The administrator then selects **Select an operator** and selects the *Is empty* operator. This operator returns all items that don't have any tags applied.

The review set is immediately updated and the only the items that aren't tagged are displayed.

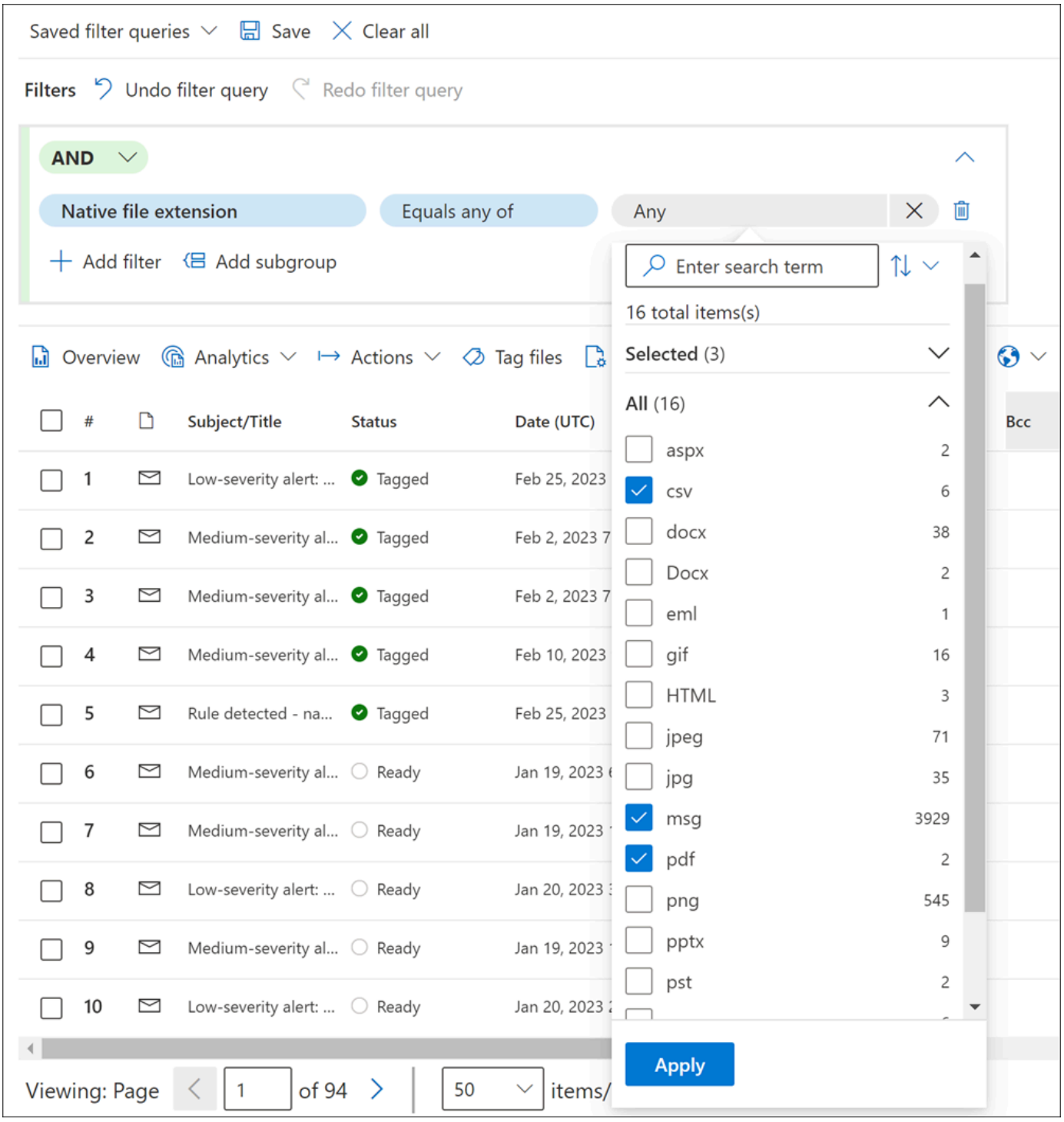


Scenario example: Filter for native file type items in a review set

An eDiscovery administrator needs to create a query to find all items in the review set that are a certain type, such as .csv, .msg, or .pdf. For this example, the administrator creates the following review set filter query:

1. For the first filter, the administrator selects the filter and types *file* in filter search. The filter *Native file extension* is one of the filter options displayed in the search results, and the administrator selects it.
2. The administrator then selects **Select an operator** and selects the *Equals any of* operator. The administrator selects the *Any* field and selects the checkboxes for the file types to include in the filter query.

The review set is immediately updated and the only the items that match the selected file types are displayed.



Scenario example: Filter partially indexed items

If you selected the option to add partially indexed items from additional data sources when you committed the collection estimate to a review set. You'll probably want to identify and view those items to determine if an item might be relevant to your investigation and whether you need to remediate the error that resulted in the item being partially indexed.

At this time, there isn't a filter option in a review set to display partially indexed items. But we're working on it. Until then, here's a way you can filter and display the partially indexed items that you added to a review set.

1. Create a collection and commit it to a new review set *without* adding partially indexed items from the additional data sources.
2. Create a new collection by copying the collection from step 1.
3. Commit the new collection to the same review set. But this time, add the partially indexed items from the additional data sources. Because items from the collection you created in step 1 have already been added to the review set, only the partially indexed items from the second collection are added to the review set.
4. After both collections are added to the review set, select the review set, and select **Load sets**.
5. Copy or make note of the **Load Id** for the second collection (the one you created in step 2). The collection name is identified in the **Source info** column.
6. Back in the review set, select **Filter**, expand the **IDs** section, and then select the **Load Id** checkbox.
7. Expand the **Load Id** filter, and then select the checkbox for the load ID that corresponds to the second collection to display the partially indexed items.

Last updated on 09/02/2025

Tag documents in a review set in eDiscovery (Premium)

> **⊗ Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

Organizing content in a review set is important to complete various workflows in the eDiscovery process. This workflow includes:

- Culling unnecessary content
- Identifying relevant content
- Identifying content that must be reviewed by an expert or attorney

When experts, attorneys, or other users review content in a review set, their opinions related to the content can be captured by using tags. For example, if the intent is to cull unnecessary content, a user can tag documents with a tag such as "non-responsive". After content has been reviewed and tagged, a review set search can be created to exclude any content tagged as "non-responsive". This process eliminates the non-responsive content from the next steps in the eDiscovery workflow.

You can use *tag templates* to define tags and review workflow across all cases and review sets in your organization. This method is the recommeded approach and helps provide structure and standardization for tagging to ensure searches identify similar content across your organization. You can also use the tagging panel in an individual review set to customized tags for each case when tag templates don't meet your needs, but tags created here are specifically for the individual case and review set.

Tag types

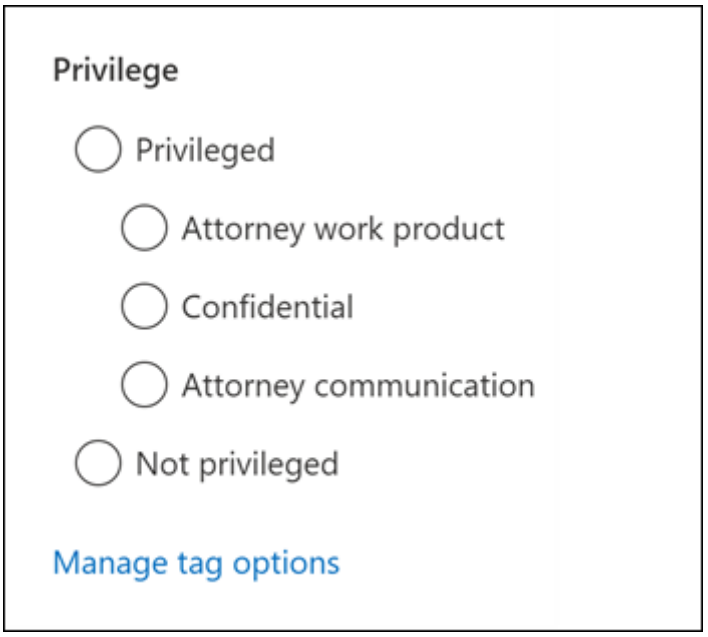
eDiscovery (Premium) provides two types of tags:

- **Single-select tags**: Restricts reviewers to selecting a single tag within a tag group. These types of tags can be useful to ensure that reviewers don't select conflicting tags such as "responsive" and "non-responsive". Single choice tags appear as radio buttons.
- **Multi-select tags**: Allow reviews to select multiple tags within a tag group. These types of tags appear as checkboxes.

Tag groups

In addition to the tag types, the structure of how tags groups are organized in the tag panel can be used to make tagging documents more intuitive. Tags are organized in tag groups. Review set search supports the ability to search by tag and by tag groups. This means you can create a review set search to retrieve documents tagged with any tag in a group or the group itself.

For example, if you need to identify and tag privileged content, grouping can be used to make it clear that a reviewer can tag a document as *Privileged* (the tag group) and select *Attorney work product* (specific tag) as the type of privilege within the tag group.



Tag templates (preview)

Tag templates are used to define and standardize tags and the review workflow across all cases and review sets in your organization. Each tag template contains tag groups where you can create a grouping structure for individual tags. You can create tag templates in the **Tag templates** section in **eDiscovery (Premium) settings**.

Create tag templates

To create a new tag template, complete the following steps:

1. In the Microsoft Purview portal, select **eDiscovery** > **Premium** and then select **eDiscovery (Premium) settings**.
2. Select **Tag templates**.
3. On the **Manage tag templates** page, select **Create**.
4. On the **New tag template** page, enter a name and description for the tag template.
5. In the initial tag group, enter a tag group name and use the dropdown to define if this as a **Multi-select** or **Single-elect** tag group.
6. Enter at least one tag name in the tag group. Select **Add tag** to add additional tags to the tag group.
7. If needed, select **Add tag group** to add additional tag groups and tags to the tag template.
8. Preview the tag template group and tag structure in the **Preview** pane on the right side of the **New tag template** page.
9. When complete, select **Save**.

The tag template is now available for use in all cases and review sets in your organization.

Manage tag templates

After you create a tag template, you can edit or delete it. You can also view information about the tag template, including the dates when it was created or last modified.

To edit a tag template, select the tag template name on the **Manage tag templates** page and select **Edit**. On the **Edit tag template** page, you can edit the tag template name, description, and tag groups and tags. When you're done, select **Save**.

To delete a tag template, select the tag template name on the **Manage tag templates** page and select **Delete**. Select **Yes** to confirm the deletion or **No** to cancel the deletion.

Apply tag templates to a case and review sets

After you've created a tag template, you can apply it to a review set. You can apply a tag template when you create a new review set or when you edit an existing review set.

To apply a tag template to a new review set, complete the following steps:

1. In the Microsoft Purview portal, select **eDiscovery** > **Premium**.
2. Select the **Cases** tab and select a case.
3. In the case, select **Review sets**, select a review set, and then select **Open review set**.

4. On the review set command bar, select **Manage** > **Tags**.
5. On the **Manage tags** page, delete the blank tag group then select **Import tag templates**.
6. On the **Import tag templates** flyout page, select the tag template that you want to apply to the review set and then select **Import**.
7. If needed, you can edit the tag template as it applies to the case and review set.
8. Select **Save** to apply the tag templates to the case and review set.

The applied tag template is now available for use in the case and all included review sets.

Create tags for a specific case or review set

If you choose not to use tag templates for a specific case or review set, you can create tags directly in the review set. This is useful when you need to create tags that are specific to a case or review set. However, these tags are unique to the case or review set and can't be used in other cases or review sets.

Before applying tags to items in a review set, you need to create a tag structure.

1. Open a review set and on the review set command bar, select **Manage** > **Tags**.
2. On the **Manage tags** flyout page, enter a *Tag group name* and select either **Multi-select** or **Single-select** for the tag group.
3. Enter one or more tags in the tag group. Select **Add tag** to add additional tags to the tag group.
4. Select **Add tag group** if more tag groups and tags are needed.
5. Preview the tag template group and tag structure in the **Preview** pane on the right side of the **New tag template** page.
6. Select **Save** to apply the tag group and tags to the case and review set.

Apply tags

With the tag structure from tag templates or custom tagging in place, reviewers can apply tags to items in a review set by configuring tagging settings.

1. In the review set command bar, select **Tag files** to display the **Tag files** flyout page (also called the *tagging panel*).

2. On the **Tag files** flyout page, you can set the following options to configure how to tag items displayed in the review set. The filters or filter queries currently applied to the review set determine which items are displayed and therefore the items that you can apply tags to. For more information, see Query and filter content in a review set.

- **Choose selection**. Choose one the following options to determine the scope of items to apply tags to.
  - **Tag selected items**: This option applies tags to the items that you select. You can select items before or after launching the tagging panel. This option displays (in real time) the number of selected items that are tagged.
  - **Tag all items in list**: This option applies tags to all items displayed in the review set. This option displays the total number of items that are tagged.

- **Expand selection**: Use the following options to tag additional items that are related to tagged items in the review set. Depending on the **Enable group** option enabled in the case settings, you'll see the following options:

  With *Enable group* in settings disabled:
  - **None**: This option doesn't apply tags to associated family items or associated conversation items. It only applies tags to the items that are selected or to all items in the review set list.
  - **Include family groups**: This option applies the same tag to the associated family items of items that are tagged. *Family items* are items that share the same **FamilyId** metadata property value. For example, a document that's attached to an email message shares the same **FamilyId** as the email message. So if this option is selected for this example, the email message and the document are tagged, even though the document might not be included in the list of review set items.
  - **Include conversation groups**: This option applies the same tag to all items that are in the same Teams or Viva Engage conversation as the items that are tagged. *Conversation items* are items that share the same **ConversationId** metadata property value. All messages, posts, and corresponding transcript file of a conversation share the same **ConversationId**. If this option is selected, then all items in the same conversation (and transcript file) are tagged, even though some of those conversation items might not be included in the list of review set items. For more information about conversation items, see the "Grouping" section in eDiscovery (Premium) workflow for content in Microsoft Teams.

  > ⓘ **Note**
  >
  > Including family or conversation groups doesn't change the count of items shown in the **Tag selected items** or **Tag all items in list** options. In other words, the number of grouped items that are tagged is not displayed.

With *Enable group* in settings enabled:

- **None**: This option doesn't apply tags to family group items or conversation group items. It only applies tags to the items that are selected or to all items in the review set list.
- **Include family groups**: This option applies the same tag to the associated family items of items that are tagged. *Family group* items are items that share the same **GroupId** metadata property value. For example, a document that's attached to an PowerPoint file shares the same **GroupId** as the document. So if this option is selected for this example, the document and the PowerPoint file are both tagged. Both are grouped together in the list of review set items.
- **Include conversation groups**: This option applies the same tag to all items that are in the same email, Teams, or Viva Engage conversation as the items that are tagged. *Conversation group* items are items that share the same **ThreadId** metadata property value. All messages, posts, and corresponding transcript files of a conversation share the same **ThreadId**. If this option is selected, then all items in the same conversation (and transcript file) are tagged. All items are grouped together in the list of review set items.

- **Assign tags**: This section displays the tags (organized by tag groups) that you can apply to documents. These are either the tag groups and tags assigned using tag templates or custom tag groups and tags that you've created for this case. You can only apply one single-choice tag (identified by a radio button) per tag group. However, you can apply multiple multi-choice tags (identified by a checkbox).

3. Select **Apply tags** to apply the tags based on your settings.

   The **Applying tags** status message is displayed for each tag group on the tagging panel to indicate a tagging job has been started. Tags for each tag group in the **Assign tags** section are greyed out until the job is completed.

> **Tip**
>
> If you're in the process of configuring the settings on the tagging panel, but want to start over, select **Reset tag assignment** to clear the current setting. This control doesn't apply to items that are already tagged, and it doesn't change or remove tags from previously tagged items.

Monitor tagging jobs

When you tag a large number of items (or select the **Tag all items in list**) option, a **Tagging documents** job is created. You view the status of this job on the **Jobs** tab in the case. This view helps you track large tagging jobs that may take a long time to complete. In some cases, a tagging job might be complete, but the **Applying tags** status message in the tagging panel is still displayed. To update the status of tagging jobs, select **Refresh** in the review set command bar.

Removing tags

You can remove tags from items in a review set. However, you can't remove a single-choice tag that's been applied to a review set item. You can only change a single-choice tag to another single-choice tag within the same tag group.

To remove a tag:

1. Select the items that you want to remove the tag from.
2. Select **Tag files** to display the tagging panel.
3. Under **Assign tags**, unselect the tag, and then select **Apply tags**.

You can also use the previous procedure to change the tag applied to selected items. After unselecting the current tag, you can select a different one.

Last updated on 09/02/2025

Group and view documents in a review set in eDiscovery (Premium)

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

eDiscovery (Premium) review sets display content using different grouping options and include specialized viewers that you can use to examine details about individual items.

Grouping

Use the **Group** control in the command bar of a review set to view review content grouped by the following options:

Option 1: Group items with Group ID and Thread ID

If you're new to eDiscovery or are an existing eDiscovery customer with a steady stream of new cases, grouping review set items using Group IDs and Thread IDs is the recommended option. This grouping option makes it easier to find and review related items in a review set by grouping them together in the list view. With the **Enable group** option enabled in the case settings, you'll see the following grouping options on the **Group** control in the command bar of a review set:

- **Group by families**: All items related to a specific file are grouped together using the same Group ID. For example, if you have a PowerPoint file in the review set that includes imbedded images or .zip files, these images and files are grouped with the PowerPoint file and shown as nested items with the file in the item list view. Microsoft 365 Copilot and Microsoft 365 Copilot Chat user prompts and responses are also grouped together using the same Group ID when there are multiple prompts and responses associated with a file.

- **Group by conversations**: All email messages, Teams conversations, and Viva Engage conversations are grouped using the same Thread ID and appear as nested items. Additionally, all associated content for these messages and conversations is also grouped together. For example, if you have an email conversation that includes several email messages, some of which include attachments and some that include embedded images, all of the email messages, attachments, and images are grouped together in the review set list view under an applicable item.

> ⓘ Note
>
> For cases created before March 15, 2023, the **Enable group** toggle is disabled and these grouping options aren't available. Item grouping in review sets for these cases is based on *Family ID* and *Conversation ID* described in Option 2.

Option 2: Group items with Family ID and Conversation ID

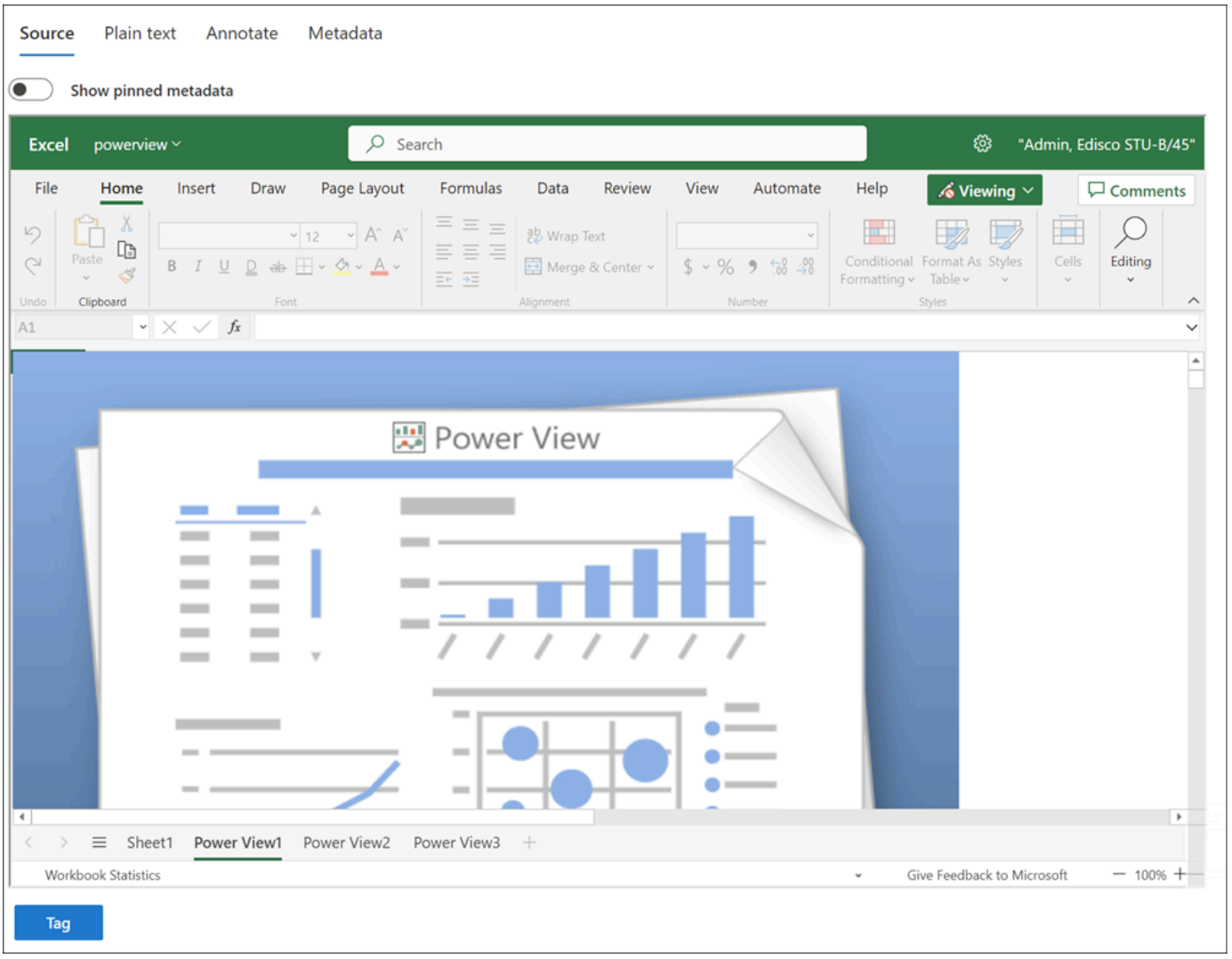
If you're an existing eDiscovery with a large number of existing cases or use existing internal or third-party automation to help process review set items, you may want to continue to group review set items with Family IDs and Conversation IDs. With the **Enable group** option disabled in the case settings, you'll see the following grouping options on the **Group** control in the command bar of a review set:

- **Group family attachments**: View review set content grouped by family. Each transcript file is displayed on a line in the list of review set items. Attachments are nested under the item.
- **Group Teams or Viva Engage conversations**: View Teams and Viva Engage content grouped by conversation. Each conversation is displayed on a line in the list of review set items. Transcript files and attachments are nested under the top-level conversation.

> ⓘ Note
>
> Cloud attachments are grouped with the conversations they appear in. This grouping is accomplished by assigning the same **FamilyId** as the transcript file of the message the file was attached to and the same **ConversationId** as the conversation the message appeared in. This means multiple copies of cloud attachments may be added to the review set if they were attached to different conversations.

Source view

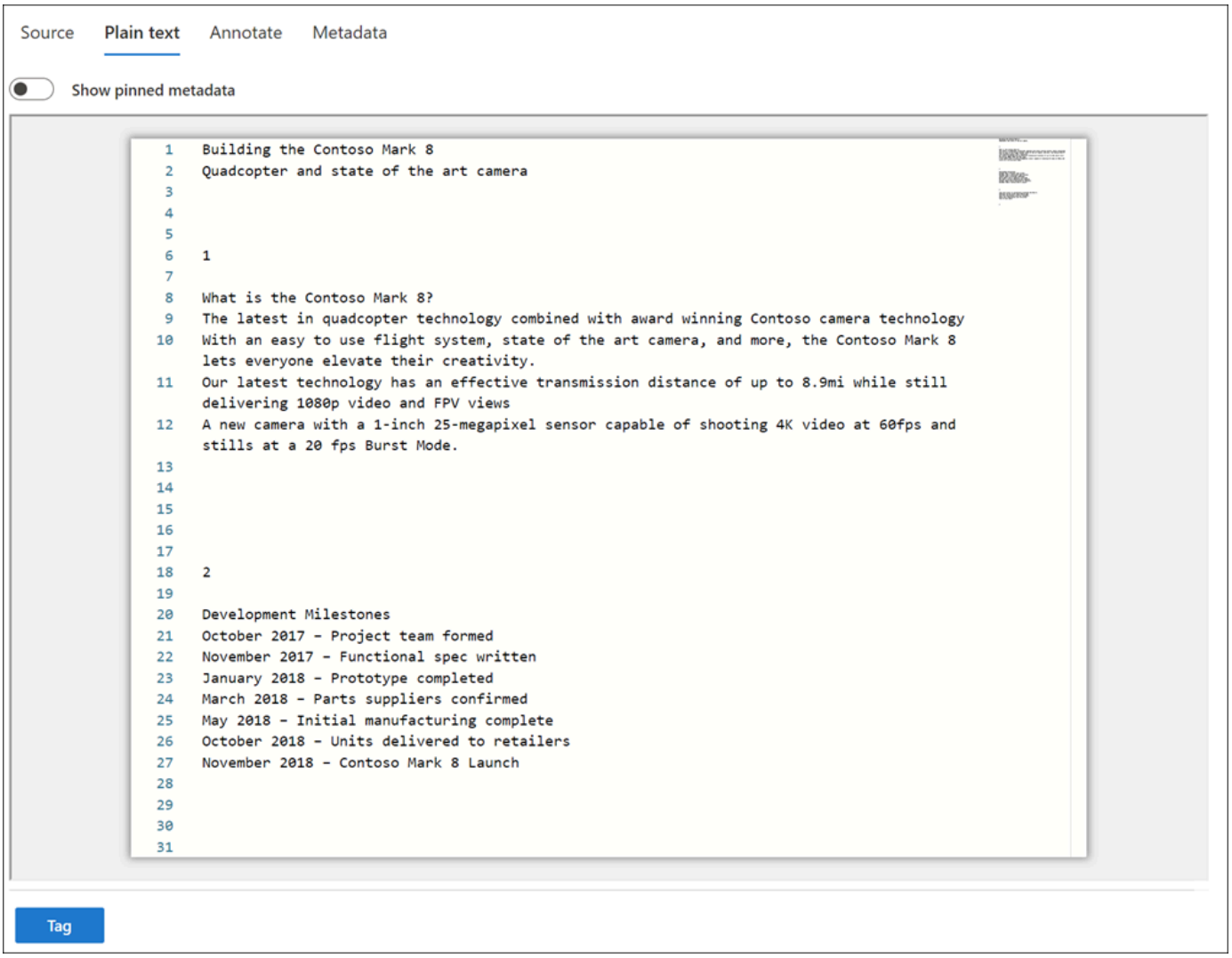
The **Source** viewer displays the richest view of a selected document. It supports hundreds of file types and is meant to display the truest to native experience possible. For Microsoft Office files, the viewer uses the web version of Office apps to display content such as document comments, Microsoft Teams chats, Excel formulas, hidden rows/columns, and PowerPoint notes.



Plain text view

The **Plain text** viewer provides a view of the extracted text of a selected file. It ignores any embedded images and formatting but is useful if you're trying to understand the content quickly. Text view also includes these features:

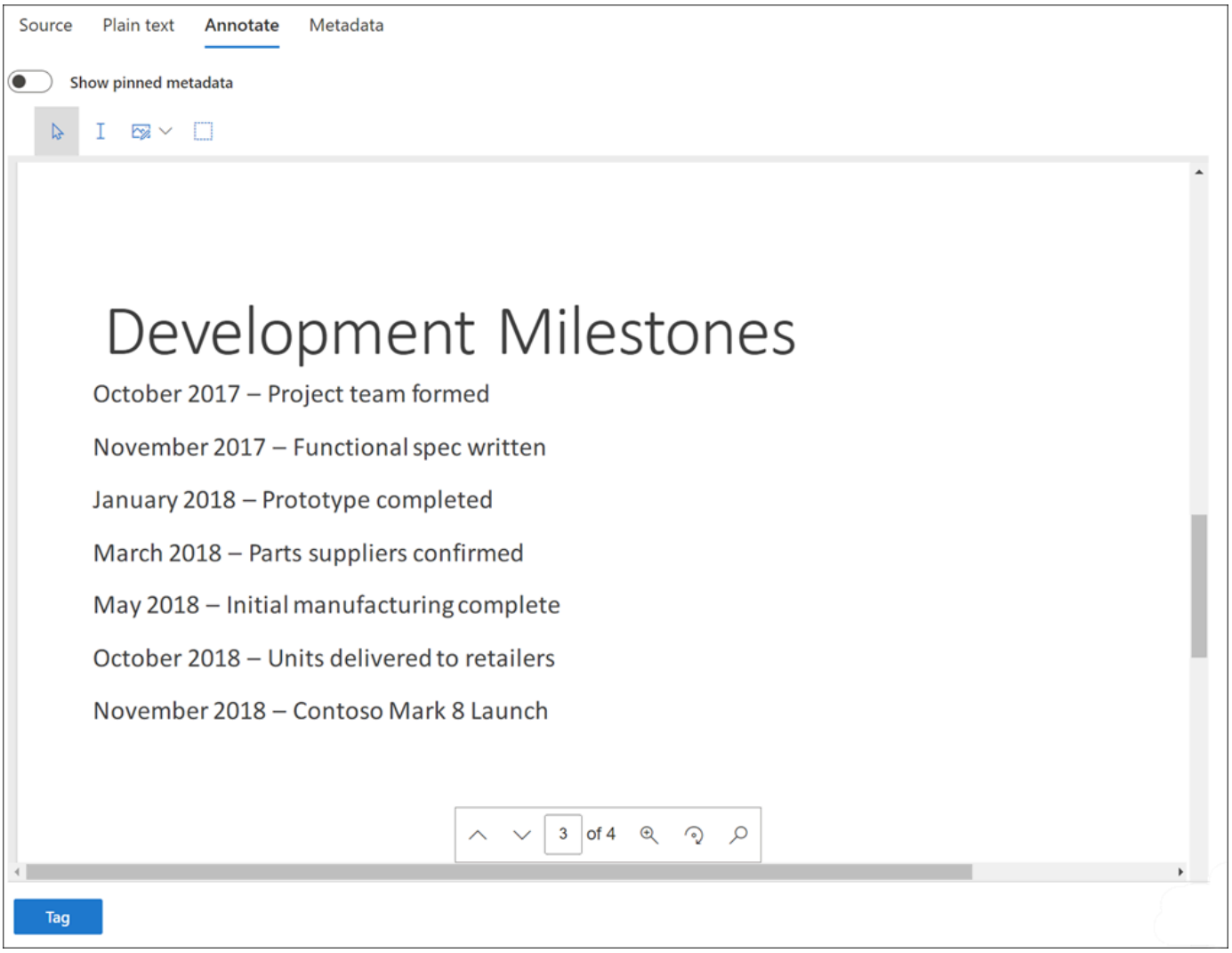
- Line counter makes it easier to reference specific portions of a document
- Search hit highlighting that highlights terms within the document and in the scrollbar



Annotate view

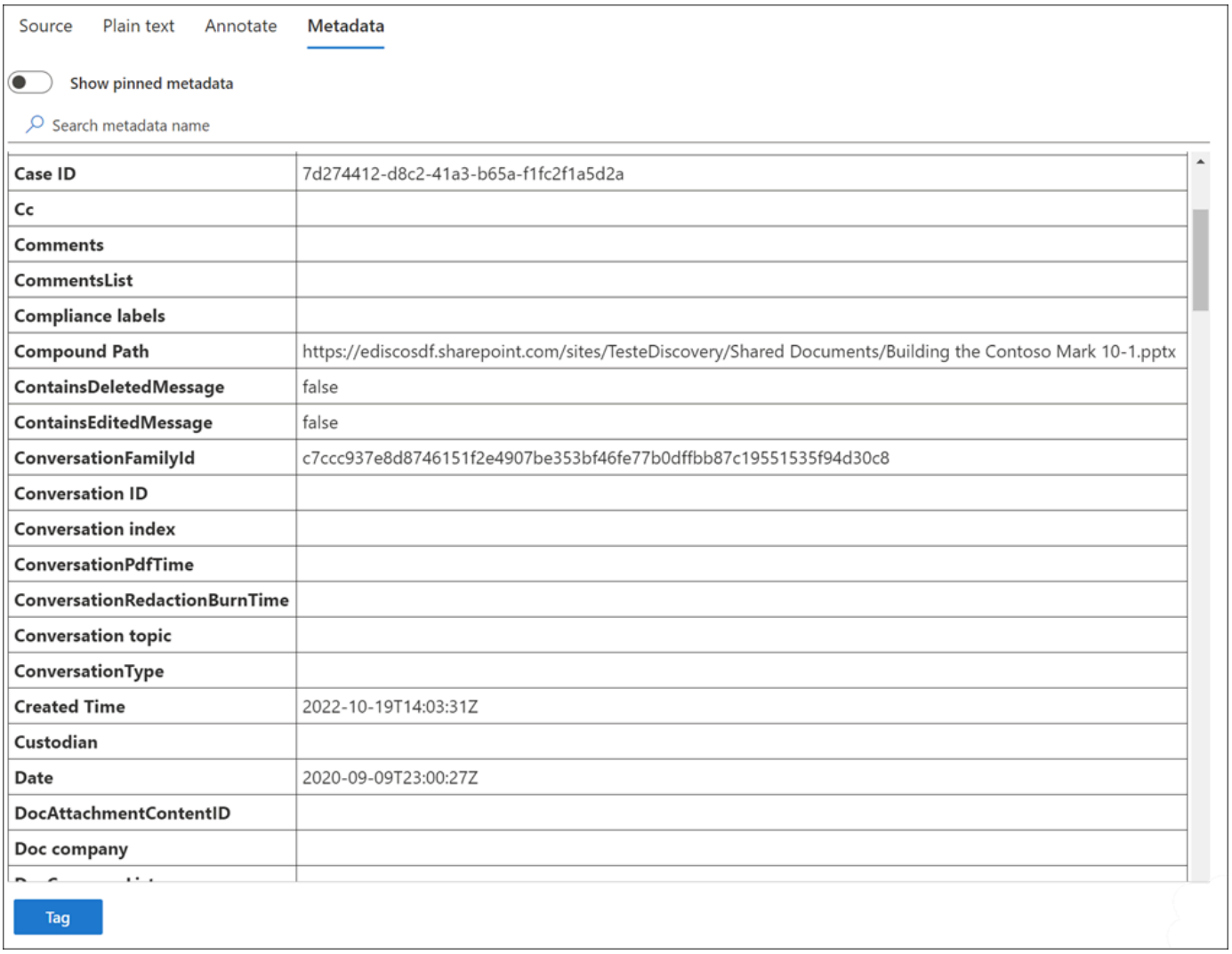
The **Annotate** view provides features that allow users to apply markup on a selected document including:

- **Select annotations**: Select annotations on a document to delete.
- **Select text**: Select text on the document to delete.
- **Area redactions**: Draw a box on the document to hide sensitive content.
- **Pencil**: Free-hand draw in selectable colors on a document in order to bring attention to certain portions of a document.
- **Toggle annotation transparency**: Make annotations semi-transparent in order to view the content behind the annotation.
- **Previous page**: Navigates to previous page.
- **Next page**: Navigates to the next page.
- **Go to page**: Enter a specific page number to navigate to.
- **Zoom**: Set the zoom level for annotate view.
- **Rotate**: Rotate the document clockwise.



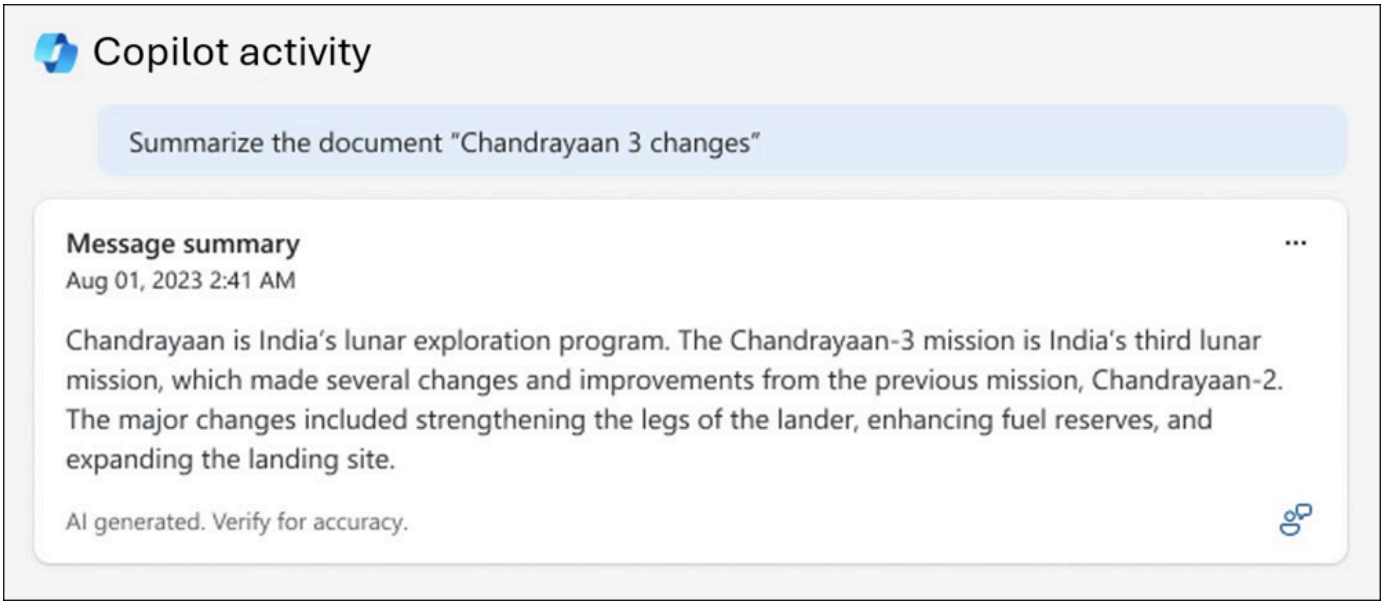
Metadata view

The panel in the **Metadata** view can be toggled on/off to display various metadata associated with the selected document. Although the search results grid can be customized to display specific metadata, there are instances where scrolling horizontally can be difficult while reviewing data. The File metadata panel allows a user to toggle on a view within the viewer.



Copilot activity view

The **Copilot activity** view displays the Microsoft 365 Copilot and Microsoft 365 Copilot Chat activity associated with the selected item. User prompts and Copilot responses are displayed in the order they were generated.



Summarize an item by using Microsoft Security Copilot

You can use Security Copilot in Microsoft Purview to provide a contextual summary of most items in a review set. The summary provided is in the context of text included in a selected item. This summary can save time for reviewers by quickly identifying information helpful when tagging or exporting items.

Security Copilot summarizes the entire item, including any documents, meetings transcripts, or attachments. Most all of the common document file types are supported.

Limitations

- The *Plain text* view must be available for an item for Security Copilot availability. For more information on which file types support plain text views, see Supported file types in eDiscovery (Premium).
- The item content must be at least 100 words and fewer than 15,000 words. If the item content is fewer than 100 words or more than 15,000 words, Security Copilot doesn't summarize it.
- Currently, only English language messages are supported for summarization.
- Currently, only individual Teams messages can be summarized. For example. if you summarize an individual Teams message, Copilot doesn't pull in the surrounding conversation (family or conversation groups).

For more information about Microsoft Security Copilot and Microsoft's artificial intelligence (AI) responsibilities, see Responsible AI FAQ.

Prerequisites for using Security Copilot

To summarize an item in a review set in eDiscovery using Security Copilot, you must have:

- Specific licenses and you must onboard your organization to Security Copilot. For more information, see Learn more about Security Copilot licensing requirements and onboarding.
- Your admin assigned at least one of the required roles to access the Security Copilot portal. For more information about these required roles, see Understand authentication in Security Copilot.

> **Note**

Currently, eDiscovery only supports Security Copilot in the embedded experience (not the stand-alone experience).

Summarize an item in a review set

1. In eDiscovery, open a review set and filter the items as needed.
2. Select an item in the list. You can select a parent item or a child item.
3. In an item in any item view, then select **Summarize**. This opens the **Copilot** panel on the right side of the screen and displays a summary of the item.

   > **Note**
   >
   > If the plain text view isn't available for an item, **Summarize** won't be available to select.

4. To use Copilot to create the summary, select one of the suggested prompts at the bottom of the panel. For example, if the item is a legal contract in the Microsoft Word format, select the *Identify key participants*, *Identify key topics*, or *Identify action items* prompt.

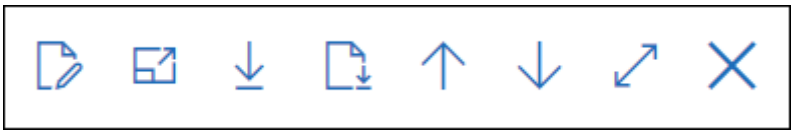
   You can also enter an open-ended question in the *Ask questions and get insights about this item* field. Copilot responses are generated without using information from prior messages. For more information about prompts and promptbooks, see Microsoft Copilot in Microsoft Purview prompts and promptbooks.

   > **Note**
   >
   > Only item content is summarized. If you ask a question relevant to the item content, Copilot provides a result. If the question isn't relevant to the item content, Copilot instructs you to ask a different question.

5. If you want to provide feedback on the summarized content, select the drop-down arrow in the lower-right corner below the summary, and then enter your feedback.
6. After reviewing the summary, you can take other eDiscovery actions connected with review set items.

Viewer and management tools

For selected content, there are additional view and management tools to help you work with documents.



You can select the following actions for selected content:

- **View or update notes**: Opens the **Notes** pane where you can view or add notes to the selected document. You can display the notes for the document in a column in the dashboard list by customizing your displayed columns.
- **Show in a new window**: Displays the selected document in a new browser window/tab.
- **Download PDF**: Downloads the selected document as a .pdf file.
- **Move up**: Selects the previous review set item and displays the item in the viewer.
- **Move down**: Selects the next review set item and displays the item in the viewer.
- **Maximize/Minimize**: Maximizes the viewer pane to full screen and minimizes the viewer pane.
- **Close**: Closes the selected item from the viewers.

> **Important**
>
> To download a PDF containing items from a review set, you must be assigned the *Export* role or be assigned to a role group that includes the *Export* role. For more information about eDiscovery permissions, see Assign eDiscovery permissions in the Microsoft Purview portal.

Last updated on 09/02/2025

eDiscovery (Premium) dashboard for review sets

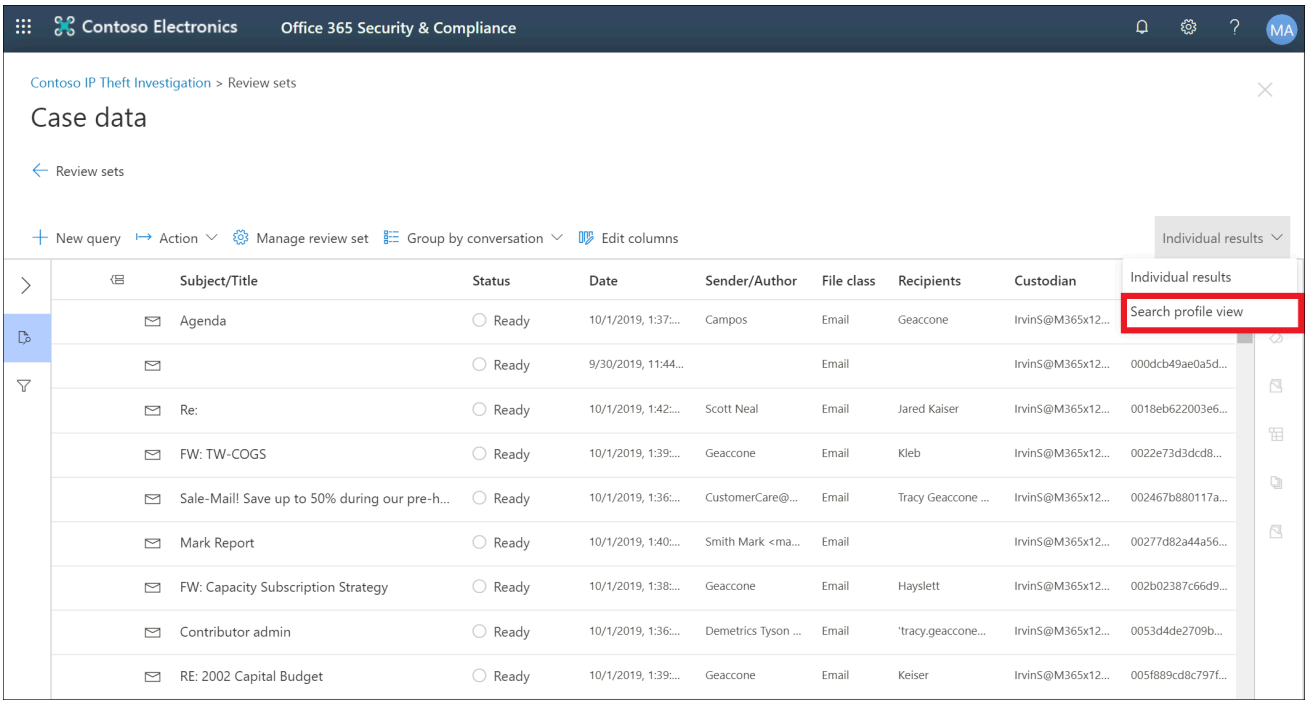
> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

> **Important**
>
> Search profile view has been retired as of March 31, 2024 and is not available in new eDiscovery cases.

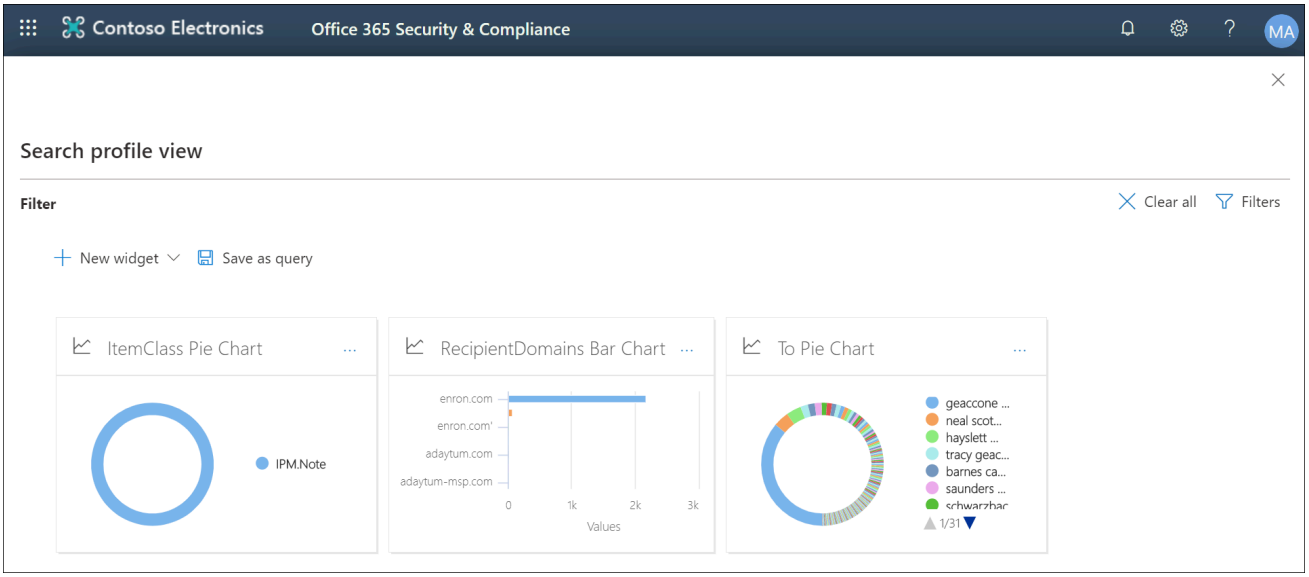
For some cases in Microsoft Purview eDiscovery (Premium), you may have a large volume of documents and email messages that need to be reviewed. Before you start the review process, you may want to quickly analyze your corpus to identify trends or key statistics that will help you develop your review strategy. To do this, you can use the eDiscovery (Premium) dashboard for review sets to quickly analyze your corpus.

Step 1: Create a widget on the review set dashboard

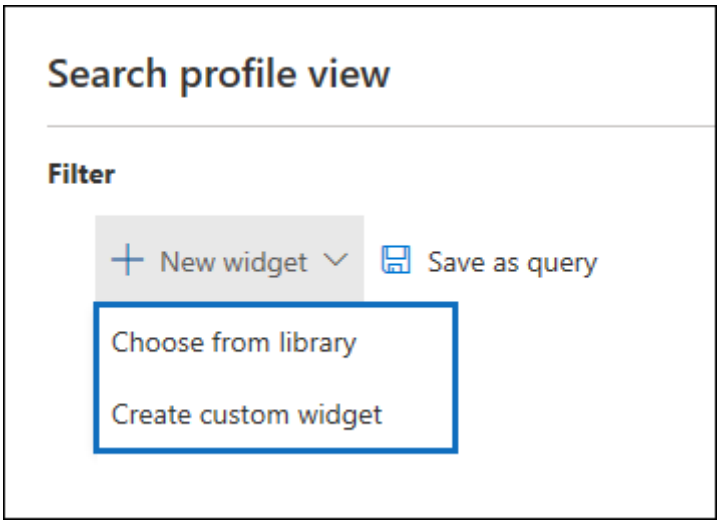
1. In the Microsoft Purview portal, go to **eDiscovery** > **eDiscovery (Premium)** to display the list of cases in your organization.
2. Select an existing case.
3. Select the **Review Set** tab, and then select a review set.
4. In the **Individual results** dropdown list, select **Search profile view**.



The **Search profile view** page is displayed; the first time you display this page, three default widgets are displayed.

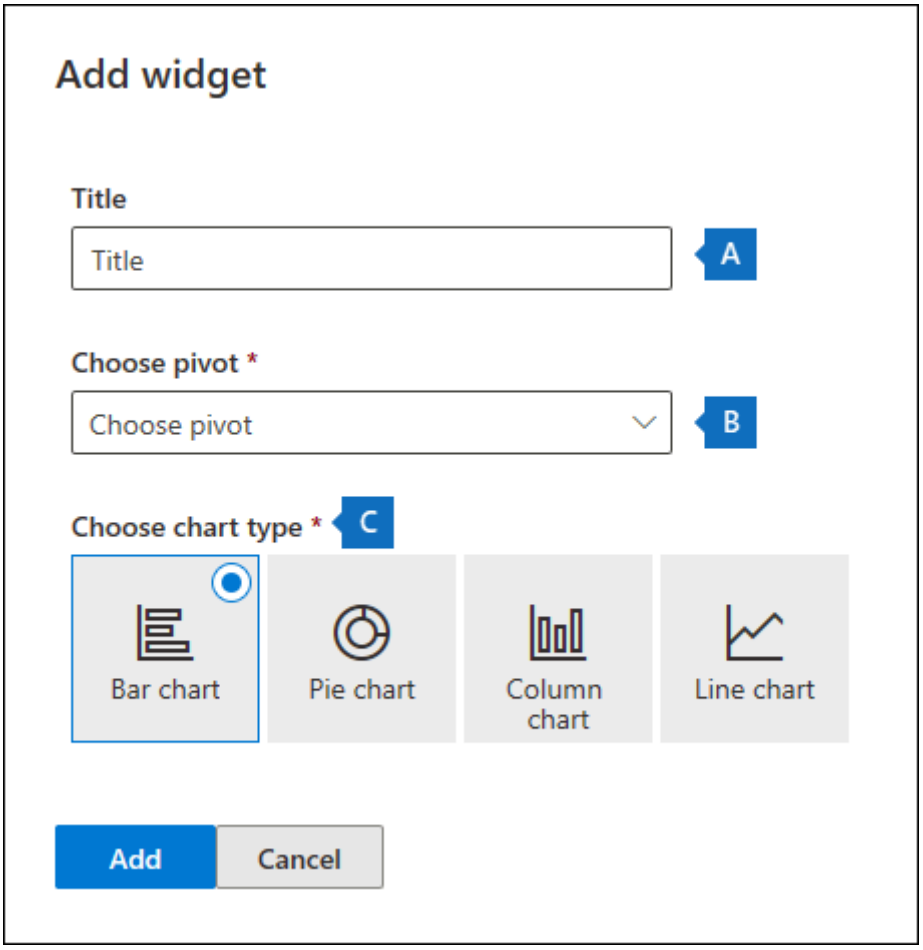


5. Select the **New widget** and then select one of the following items:



- **Choose from library:** Displays a default library of widgets. You select a widget and then select **Add** to add it to the widgets on the **Search profile view** page.
- **Create custom widget:** Displays a flyout page that you can use to set up a custom widget.

6. To create a custom widget, do the following on the **Add widget** flyout page:



a. Type a name for the widget, which is displayed in the widget title bar. Naming a widget is required, but it's helpful to identify the widget data.

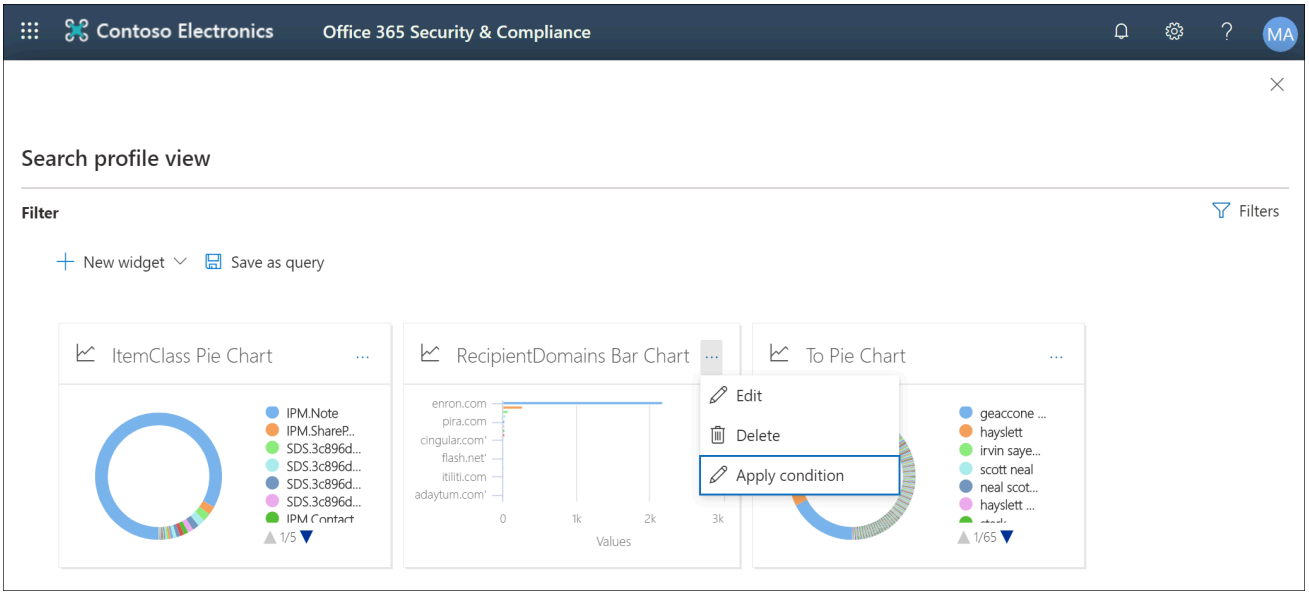
b. Select a property in the **Choose pivot** dropdown list that will be used for the widget data. The items in this list are the searchable properties for the items in the review set. For a description of these properties, see Document metadata fields in eDiscovery (Premium). The pivot options for the widget are listed in the **Searchable field name** column in this topic.

c. Select a chart type to display the data from the selected pivot property.

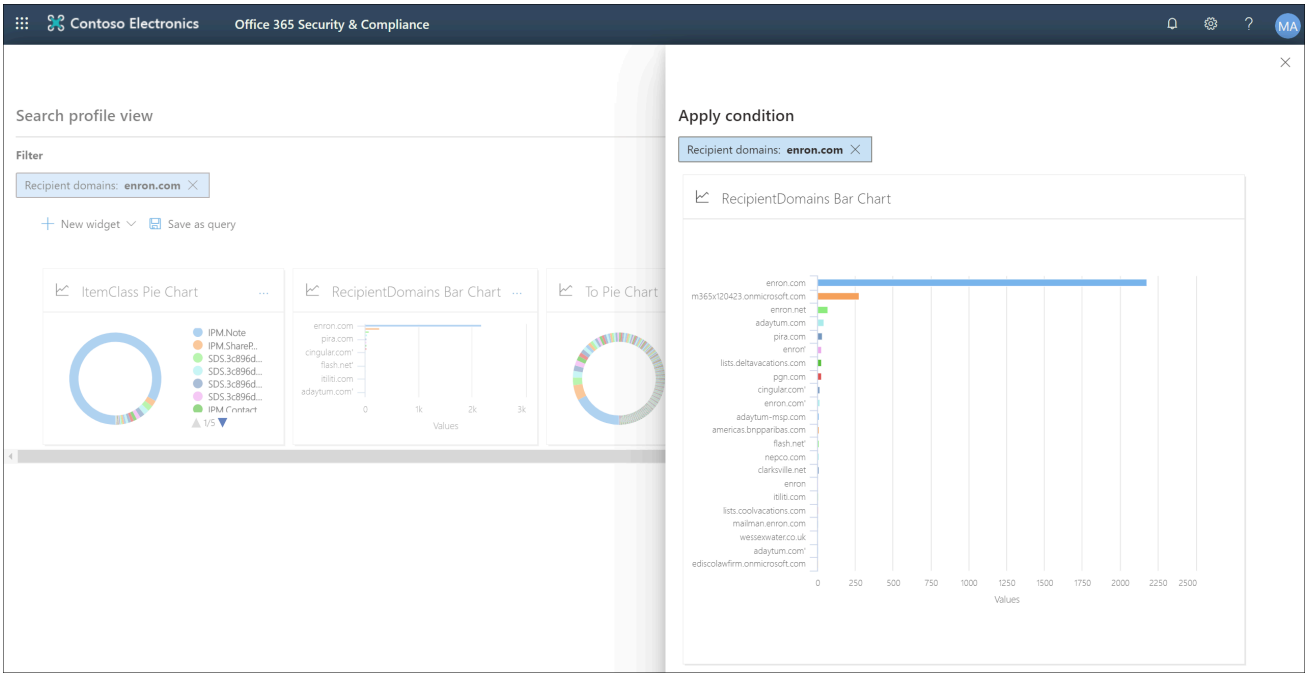
7. Select **Add** to create the custom widget and display it on the **Search profile view** page.

Step 2: Create a review set search query

1. Select **...** in the widget title bar, and then select **Apply condition**.

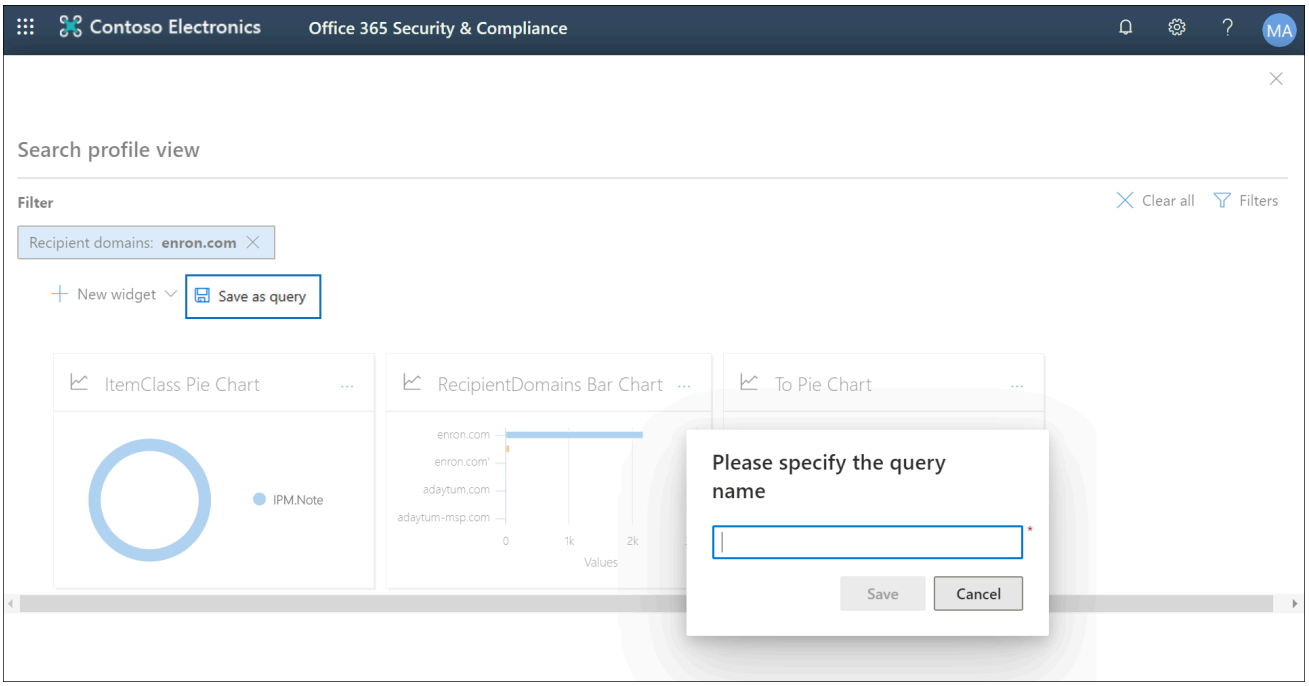


2. On the flyout page, select an element on the widget key or widget chart to create a filter.



3. Repeat steps 1-2 for other widgets multiple widgets.

4. When you're done, select **Save as query** to save your conditions as a new search query for the review set.



5. Close the **Search profile view** to return to the search results view.

   If you have created any visual filters, the resulting query is applied to the search results that are displayed, and the search query that you saved in step 4 is displayed under **Saved queries**. For more information about review set queries, see Query the data in a review set.

Last updated on 11/21/2025

Guest access in eDiscovery (Premium)

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

With guest access, you can provide access to an eDiscovery Premium case to people outside your organization. You can invite guests users to eDiscovery (Premium) cases just like you can invite guests into your Teams environment.

A guest is someone who doesn't have a work or school account with your organization. For example, guests may include external partners, vendors, suppliers, or consultants. Anyone who isn't part of your organization can be added as guest for an eDiscovery (Premium) case. This access means that anyone with a Microsoft Entra account or personal email account (Outlook, Gmail, or others) can participate as a guest in eDiscovery and have access to specific eDiscovery (Premium) cases.

Guest access in eDiscovery uses Azure B2B technology, which depends on an organization-wide Azure collaboration setting. When you invite a guest to eDiscovery, a guest account is created for them in Microsoft Entra ID. They're covered by the same compliance and auditing protection as other Microsoft 365 users. Guest access is subject to Microsoft Entra ID and Microsoft 365 service limits.

> **Note**
>
> eDiscovery guests are currently limited to *Reviewer* role group permission only. For more information about the reviewer role, see Assign eDiscovery permissions in the Microsoft Purview portal.

Set up guest access

Setting up guest access in eDiscovery requires configuring other settings in Microsoft Entra ID. To enable access to invite guests to eDiscovery cases, complete the following steps:

1. Check the B2B external collaboration settings to ensure that collaboration with guests isn't blocked. If the **Guest invite settings** is set to *No one in the organization can invite guest users including admins* then inviting guest into eDiscovery won't work. Contact your system administrator to ensure this setting is properly configured. You can view and update the collaboration setting using the *Authorization Policy* API.

2. Depending on the specific configuration in Step 1, eDiscovery administrators may or may not need additional roles to approval a guest access request submitted from the eDiscovery user.

    - If the *Anyone in the organization can invite guest users including guests and non-admins (most inclusive)* or *Member users and users assigned to specific admin roles can invite guest users including guests with member permission* options are selected, then eDiscovery administrators can approve invitation requests without additional roles.
    - If *Only users assigned to specific admin roles can invite guest users* is selected, then the eDiscovery administrator who would approve the invitation request requires the *Invite guest* role to approve and trigger the invitation email to the guest. See the Assign Microsoft Entra roles to users article for guidance on how to add the *Guest Inviter* role to users. If you prefer to use PowerShell to add this role, see Assign the Guest Inviter role to a user.

3. To enable collaboration with guest reviewers from other Microsoft 365 cloud environments, including if you're a Microsoft Azure Government organization, both your organization and the guest user's organization must enable specific connection settings. For detailed configuration guidance, see Collaborate with guests from other Microsoft 365 cloud environments.

Enable guest access in eDiscovery (Premium)

eDiscovery admins must enable guest access in eDiscovery (Premium) before they can invite users as guests.

To enable guest access for guests in eDiscovery (Premium), complete the following steps:

1. In the Microsoft Purview portal, navigate to **eDiscovery** > **Premium** > **eDiscovery (Premium) settings** > **Guest users**.
2. Toggle **Guest access** to *On*.

Now eDiscovery managers and admins can initiate the guest invitation process within any eDiscovery (Premium) case.

Invite guests to an eDiscovery (Premium) case

To invite guests to an eDiscovery (Premium) case, complete the following steps:

1. In the Microsoft Purview portal, navigate to **eDiscovery** > **Premium** > **eDiscovery (Premium) cases** and select the case you want to invite guests to.
2. In the selected case, select the **Settings** tab and then select **Access & permissions**.
3. In the **Guest users** section, select **Add**.
4. On the **Invite guest** flyout pane, complete the following fields:
    - **Full name**: Enter the full name of the guest you want to invite.
    - **Email address**: Enter the email address of the guest you want to invite.
    - **Organization**: Enter the organization of the guest you want to invite.
    - **Guest role**: Select the role you want to assign to the guest. The guest is currently limited to *Reviewer* role group permission only.
    - **Justification**: Explain why this person needs access to the case.
5. Select **Invite** to submit the invitation request for review and approval by an eDiscovery admin.

Approve guest access requests

After guest invitations are submitted, eDiscovery admins must approve the requests before the guest can access the case. eDiscovery admins must have the *Role Management* role assigned to approve guest access requests.

To approve guest access requests, eDiscovery managers must complete the following steps:

1. eDiscovery managers are notified of pending invitation requests on the **Overview** tab in eDiscovery (Premium).
2. To review and approve or deny the request, navigate to **eDiscovery** > **Premium** > **eDsicovery (Premium) settings** > **Guest users**. Select the guest request and select **Create account** or **Deny**.
3. If the guest is being added to the tenant for the first time, eDiscovery admins need to perform a two-step approval process. They first need to approve creation of an account,

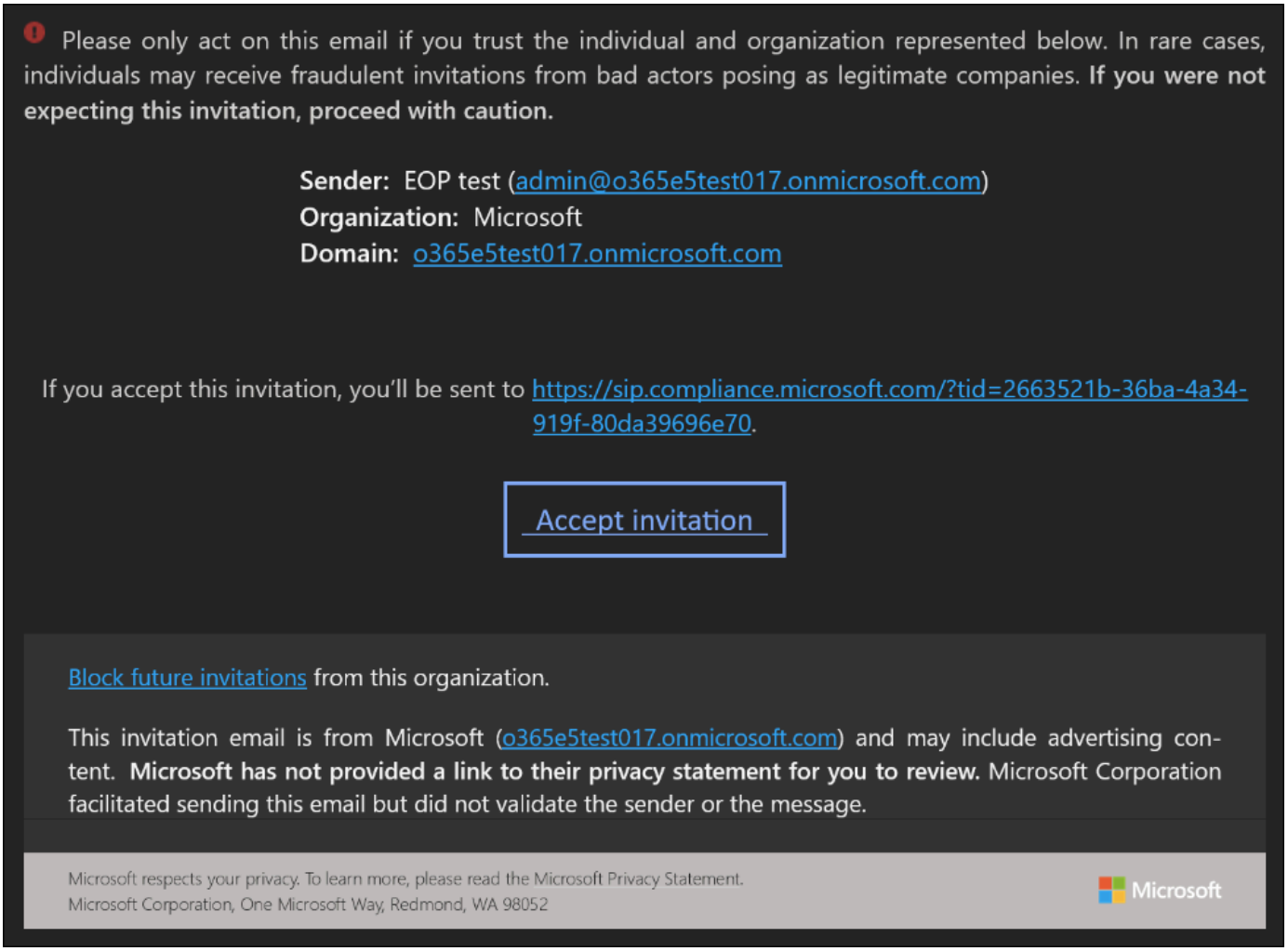
which may take up to 24 hours. After 24 hours, eDiscovery admins can approve guest access to the case. For future access requests for access to other cases for the same guest, only one-step approval is required.

In you encounter an error when approving the guest access to a case, check the following:

- The guest's email isn't disabled, blocked or deleted by the tenant admin.
- If there are any members in the *Reviewer* role group whose Microsoft Entra account has been deleted. Currently, an error occurs when adding a new member to a role group that has any deleted users. If the *Reviewer* role group has any deleted users, remove them from the role group and retry the guest approval.

Accepting guest invitations

After the request is approved by the eDiscovery admin, the guest receives a welcome email from *Microsoft Invitations*. This email contains information about the eDiscovery invitation and a link where the guest can accept the invite and sign in. The guest must select the unique link provided in the email or copy/paste the link in a browser to log into the specific eDiscovery organization.



Guests that have a work or school account in Microsoft Entra ID can accept the invitation and authenticate immediately. Other users are sent a one-time passcode to validate their identity. You must ensure that the one-time passcode setting is enabled in your organization.

> **Note**
>
> If the guest is logged into another eDiscovery organization in their browser, they need to sign out from the organization and clear their browser cache before selecting the link in the invitation email.

After accepting the invitation and authenticating, the guest can participate in the eDiscovery case like an internal reviewer of the case and perform review activities such as querying, filtering, tagging, redaction, downloading single items for review, and more.

Removing guest access

An eDiscovery admin can remove a guest's access to a case from eDiscovery Premium settings page. To remove guest access for a user, eDiscovery admins must complete the following steps:

1. To review and approve or deny the request, navigate to **eDiscovery** > **Premium** > **eDsicovery (Premium) settings** > **Guest users**.
2. Select the guest and select **Remove** from the flyout pane.

The guest continues to have access to other cases that they have been added to. When a guest is removed from the last case they have access to, they're removed from the *Reviewer* role group and won't have access to any cases in eDiscovery (Premium).

The guest continues to have a Microsoft Entra guest account in the organization. A regular review and clean-up of guest accounts from Microsoft Entra ID is recommended.

Last updated on 09/02/2025

Analyze data in a review set in eDiscovery (Premium)

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

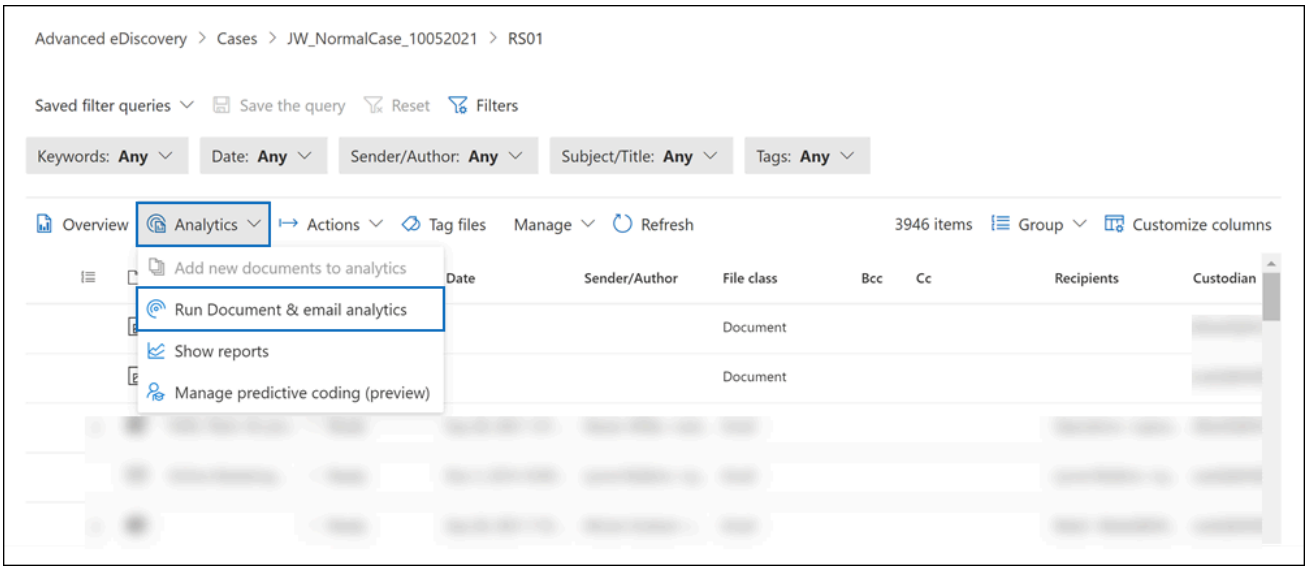
When the number of collected documents is large, it can be difficult to review them all. Microsoft Purview eDiscovery (Premium) provides a number of tools to analyze the documents to reduce the volume of documents to be reviewed without any loss in information, and to help you organize the documents in a coherent manner. To learn more about these capabilities, see:

- Near duplicate detection
- Email threading
- Themes

Run analytics for a review set

To analyze data in a review set:

1. Configure analytics settings for your case. For more information, see Configure search and analytics settings.
2. Open the review set you want to analyze.
3. Select **Analytics** > **Run document & email analytics**.



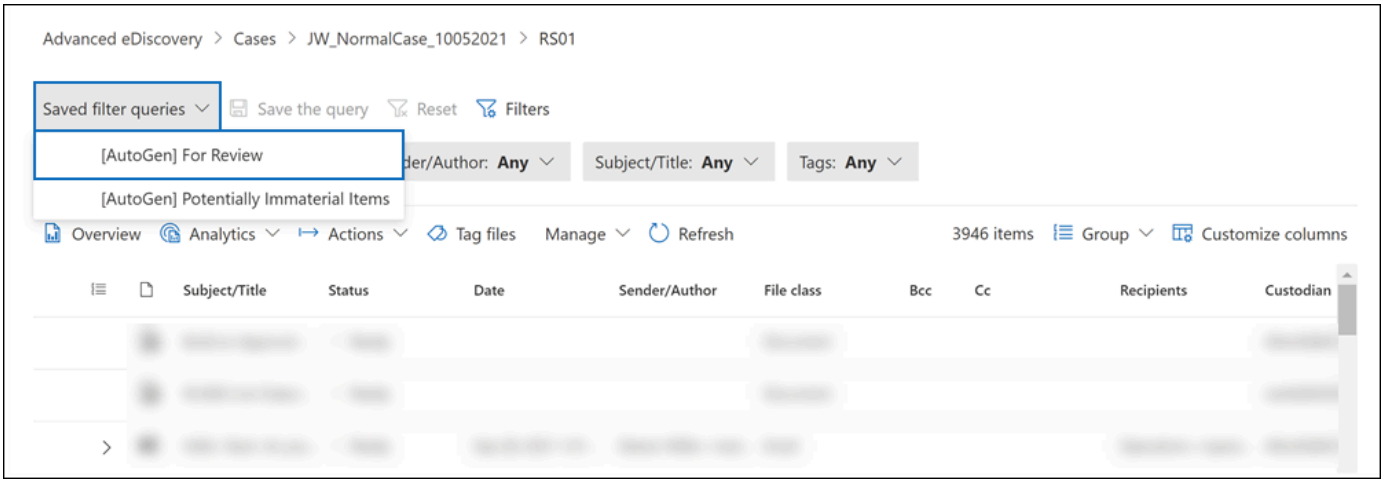
You can check the progress of analysis on the **Jobs** tab of the case.

After analysis is completed, you can view the analytics report, run queries within your review set on outputs of the analysis (see Query within your review set, and see related documents of a given document (see Reviewing data in review set.

Using the For Review filter query

After running analytics for the review set, you can use an automatically generated filter query (called *For Review*) that filters your review to exclude immaterial, duplicate, or non-inclusive items. This leaves you with only the items that are representative, unique, and inclusive in the review set.

To apply the **For Review** filter query to a review set, select the **Saved filter queries** dropdown list, and then select **[AutoGen] For Review**.



Here's the syntax for the **For Review** filter query:

```
(((FileClass="Email") AND (IsInclusive="True") AND (MarkAsRepresentative="Unique")) OR ((FileClass="Attachment") AND (MarkAsRepresentative="Unique")) OR ((FileClass="Document")
```

```
AND (MarkAsRepresentative="Unique")) OR ((FileClass="Conversation") AND
(MarkAsRepresentative="Unique")) OR ((FileClass="Attachment" OR FileClass="Conversation" OR
FileClass="Document" OR FileClass="Email") AND NOT (Exists:MarkAsRepresentative)))
```

The following list describes the result of the filter query in terms of what content is displayed after you apply it to the review set.

- **Email**: Displays items that are marked as both **Inclusive** and **Representative**. An email is marked as IsInclusive True when it contains all the unique content from a thread, including all previous replies. It ensures that only the most comprehensive email in a thread is reviewed, which is essential for understanding the full context of the conversation without having to review each individual reply. An email is MarkedAsRepresentative Unique when it is the single representative copy among duplicates.
- **Attachments**: Filters out duplicate attachments in the review set. Only attachments that are unique in the review set are displayed.
- **Documents**: Filters out duplicate documents. Only documents that are unique in the review set are displayed.
- **Teams conversations**: Filter out duplicate Teams (and Viva Engage) conversations. Only conversations that are unique in the review set are displayed.
- **Other**: Any emails, attachments, documents and conversations in the review set where MarkedAsRepresentative field contains no value. This means analytics was not able to generate a value for it, so the reviewer should review them and assess themselves.

For more information about inclusive types and document uniqueness, see Email threading in eDiscovery (Premium).

Analytics report

To view the analytics report for a review set:

1. Open the review set.
2. Select **Analytics** > **Show reports**.

The **Analytics** report has seven components from the analysis:

- **Target population:** The number of email messages, attachments, and loose documents found in the review set.

- **Documents (excluding attachments):** The number of loose documents that are pivots, unique near duplicates of a pivot, or an exact duplicate of another document.
- **Emails:** The number of email messages that are marked as *inclusive* vs not.
- **Attachments:** The number of email attachments that are unique or duplicates of another email attachment in the review set.
- **Number documents by file type:** The number of files, identified by file extension.
- **Documents by source:** A summary of content by its original data source.
- **Documents aggregated by process:** A summary of content by review set processes.

Last updated on 11/21/2025

Near duplicate detection in eDiscovery (Premium)

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

Consider a set of documents to be reviewed in which a subset is based on the same template and has mostly the same boilerplate language, with a few differences here and there. If a reviewer could identify this subset, review one of them thoroughly, and review the differences for the rest, they would not have missed any unique information while taking only a fraction of time that would have taken them to read all documents cover to cover. Near duplicate detection groups textually similar documents together to help you make your review process more efficient.

How does it work?

When near duplicate detection is run, the system parses every document with text. Then, it compares every document against each other to determine whether their similarity is greater than the set threshold. If it is, the documents are grouped together. Once all documents have been compared and grouped, a document from each group is marked as the "pivot"; in reviewing your documents, you can review a pivot first and review the other documents in the same near duplicate set, focusing on the difference between the pivot and the document that is in review.

Last updated on 11/21/2025

Email threading in eDiscovery (Premium)

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

Consider an email conversation that has been going on for a while. In most cases, the last message in the email thread includes the contents of all the preceding messages. Therefore, reviewing the last message gives a complete context of the conversation that happened in the thread. Email threading identifies such messages so that reviewers can review a fraction of collected documents without losing any context.

What does email threading do?

Email threading in eDiscovery refers to the process of organizing a sequence of related emails that are part of the same conversation. This includes the initial email and all subsequent replies and forwards linked to that original email. By grouping these emails into threads, reviewers can see the entire context of a conversation, making it easier to understand the flow of communication. This approach helps identify relevant information more efficiently, as it eliminates the need to review each email individually. Email messages post analytics process can have the following metadata populated:

- **Is Inclusive**: This field identifies whether an email contains all the unique content from a thread, including all previous replies. It ensures that only the most comprehensive email in a thread is reviewed, which is essential for understanding the full context of the conversation without having to review each individual reply.
- **Has Unique Attachments**: This field marks emails that contain attachments not found in other emails within the same thread. Even if the email content is duplicated, unique attachments are flagged to ensure that all relevant documents are reviewed. This is

important in the legal review process to ensure that no unique evidence is overlooked, even if the email body itself is not unique.

How is it different from conversations in Outlook?

At a glance, this sounds similar to conversation groupings in Outlook. However, there are some important distinctions. Consider an email conversation that got forked into two conversations; for instance, someone responded to an email that isn't the latest in the conversation so the last two emails in the conversation both have unique content.

Outlook would still group the emails into a single conversation; reading only the last email would mean missing the context of the second-to-last email, which also contains unique content. Because email threading parses out each email into individual components and compares them, email threading would mark both of the last two emails as inclusive, ensuring that you won't miss any context as long as you read all emails marked as inclusive.

Let's also consider an email thread with multiple replies, where some replies include inline responses that modify the quoted content. If an inline reply alters part of the previous email, the latest reply doesn't fully encompass the content of the earlier email. Both the latest reply and the earlier email with unique content are marked as inclusive. This approach ensures that any unique information from the inline reply is preserved and not overlooked.

Last updated on 11/21/2025

Themes in eDiscovery (Premium)

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

When you create a new document, you generally start with one or more ideas that you want to convey in the document, and then compose the document using words that align with these ideas. The more prevalent an idea is, the more frequent the words that are related to that idea tend to be. This method also aligns to how readers consume documents. The important things to understand from reading a document are the main ideas that the document is trying to convey. This also includes which ideas appear where and what the relationships between the ideas are.

This process can be extended to how an eDiscovery reviewer wants to consume a set of documents in a case. They want to see which ideas are present in the review sets and which documents are talking about those ideas. If they find a particular document of interest, they want to be able to see documents that discuss similar ideas.

The **Themes** functionality in eDiscovery (Premium) attempts to mimic how humans reason about documents, by analyzing the *themes* that are discussed in a review set and assigning a theme to documents in the review set. In eDiscovery (Premium), Themes goes one step further and identifies the *dominant theme* in each review set and document. The *dominant theme* is the one that appears the most often in a document.

How does themes work?

The **Themes** functionality analyzes documents with text in a review set to parse out common themes that appear across all the documents in the review set. eDiscovery (Premium) assigns those themes to the documents in which they appear. It also labels each theme with the words used in the documents that are representative of the theme. Because a document can contain various types of subject matter, eDiscovery (Premium) often assigns multiple themes to review

sets and documents. This is referred to as the *Themes list*. The theme that appears most prominently in a review set or document is designated as its *dominant theme*.

Configuring Themes

Themes are supported for cases and apply to all the review sets within them. You can configure the settings for themes when you create a new case or you can update the theme settings for an existing case.

Configure Themes when creating a new case

To configure themes when you create a new case, complete the following steps:

1. In the Microsoft Purview portal, navigate to **eDiscovery** > **Premium**.
2. Select the **Cases** tab and open an eDiscovery (Premium) case, and then select **Create a case**.
3. On the **Name and description** page, name the case and complete the **Description**, **Number**, and **Case format** options as applicable. Select **Next** to continue.
4. On the **Members and settings** page, select **Group items by theme**. Select the following theme options as applicable:

    - **Maximum number of themes:** Specifies the maximum number of themes that can be generated when you run analytics on the data in review sets included in a case. For more information on limits, see Limits in eDiscovery (Premium).
    - **Add numeric identifiers to themes:** Numbers (that identify a theme) are included when generating themes.
    - **Automatically set maximum number of themes:** In certain situations, there may not be enough documents in a review set to produce the desired number of themes for the case. When this setting is enabled, eDiscovery (Premium) adjusts the maximum number of themes dynamically rather than attempting to enforce the maximum number of themes.

5. If you need to exclude keywords associated with themes, enter the text or regular expression needed in the **Text to ignore** field. In the **Apply to** field, select **Themes** to apply the text or regular expression to all themes.
6. Complete the other field on the **Members and settings** page as applicable, then select **Next**.
7. On the **Summary** page, review the case settings, then select **Submit** to create the case.

After a new case is created, analytics are automatically run on the data when the review sets are added to the case. Themes for the review sets are generated as part of the analytics processing.

Configure Themes for an existing case

To configure themes for an existing case, complete the following steps:

1. In the Microsoft Purview portal, navigate to **eDiscovery** > **Premium**.
2. Select the **Cases** tab and open an eDiscovery (Premium) case.
3. Select the **Settings** tab, and then choose **Select** on the **Search & analytics** card.
4. On the **Search & analytics** page, select the **Themes** checkbox and configure the following options as applicable:

    - **Max number of themes:** Specifies the maximum number of themes that can be generated when you run analytics on the data in review sets included in a case. For more information on limits, see Limits in eDiscovery (Premium).
    - **Include numbers themes:** Numbers (that identify a theme) are included when generating themes.
    - **Adjust maximum number of themes dynamically:** In certain situations, there may not be enough documents in a review set to produce the desired number of themes for the case. When this setting is enabled, eDiscovery (Premium) adjusts the maximum number of themes dynamically rather than attempting to enforce the maximum number of themes.

5. If you need to exclude keywords associated with themes, select **Edit** in the **Ignore text** section. Enter the text or regular expression needed in the **Text (Regular expressions are supported)** field. In the **Apply to** field, select **Themes** to apply the text or regular expression to all themes.
6. Select **Save** to save the theme settings.
7. Navigate to the review set in the case that you want to apply the new theme settings to, and select **Analytics** > **Run Document & email analytics**.
8. Select **Yes** to confirm that you want to run analytics on the review set.

After the analytics job completes, you can view the themes that were generated for the review set by filtering the documents in the review set by theme.

Filtering documents by Theme

Filtering documents by theme can significantly save time when reviewing documents. For example, if you're looking for documents that discuss a particular topic, you can filter the documents by the dominant theme that is related to that topic. You can also filter documents by other themes in the theme list to find documents that are similar to a document that you're interested in. To display the themes for a document as a column in the document list for the review set, select **Customize columns** and select **Dominant theme** and **Themes list**.

To filter documents by theme, complete the following steps:

1. In a review set, choose **Filters** and expand the options for **Analytics & predictive coding**.
2. Select **Dominant theme** and **Themes list**, and then select **Done** to add these filters to the review set.
3. Use the conditional controls as needed to filter the documents by specific themes for each of these filters.

To filter documents by theme using advanced filters (preview), complete the following steps:

1. In a review set, choose **Select a filter** and select **Dominant theme**.
2. Select an operator to use with the **Dominant theme** and define the value to use with the operator.
3. Use an addition **Themes list** filter and the operator and values to applicable to this filter. You can configure the **AND** and **OR** operators to filter documents by a combination of the **Dominant theme** and **Themes list** values.

Last updated on 09/02/2025

Quick start: Predictive coding in eDiscovery (Premium) (preview)

> **⊗ Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

This article presents a quick start for using predictive coding in Microsoft Purview eDiscovery (Premium). The predictive coding module uses intelligent, machine learning capabilities to help you cull large volumes of case content that's not relevant to your investigation. This is accomplished by creating and training your own predictive coding models that help you prioritize the most relevant items for review.

Here's an a quick overview of the predictive coding process:



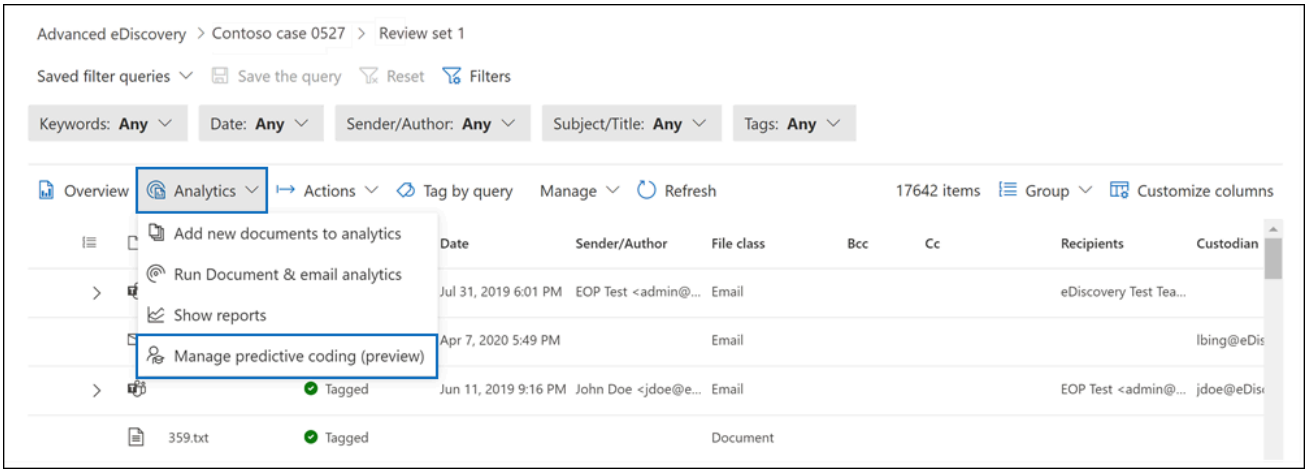
To get started, you create a model, label as few as 50 items as relevant or not relevant. The system then uses this training to apply prediction scores to every item in the review set. This lets you filter items based on the prediction score, which allows you to review the most relevant (or non-relevant) items first. If you want to train models with higher accuracies and recall rates, you can continue labeling items in subsequent training rounds until the model stabilizes. Once the model is stabilized, you can apply the final prediction filter to prioritize items to review.

For a detailed overview of predictive coding, see Learn about predictive coding in eDiscovery (Premium).

Step 1: Create a new predictive coding model

The first step is to create a new predictive coding model in the review set:

1. In the Microsoft Purview portal, open an eDiscovery (Premium) case and then select the **Review sets** tab.
2. Open a review set and then select **Analytics** > **Manage predictive coding (preview)**.



3. On the **Predictive coding models (preview)** page, select **New model**.
4. On the flyout page, type a name for the model and an optional description.
5. Select **Save** to create the model.

   It will take a couple minutes for the system to prepare your model. After it's ready, you can perform the first round of training.

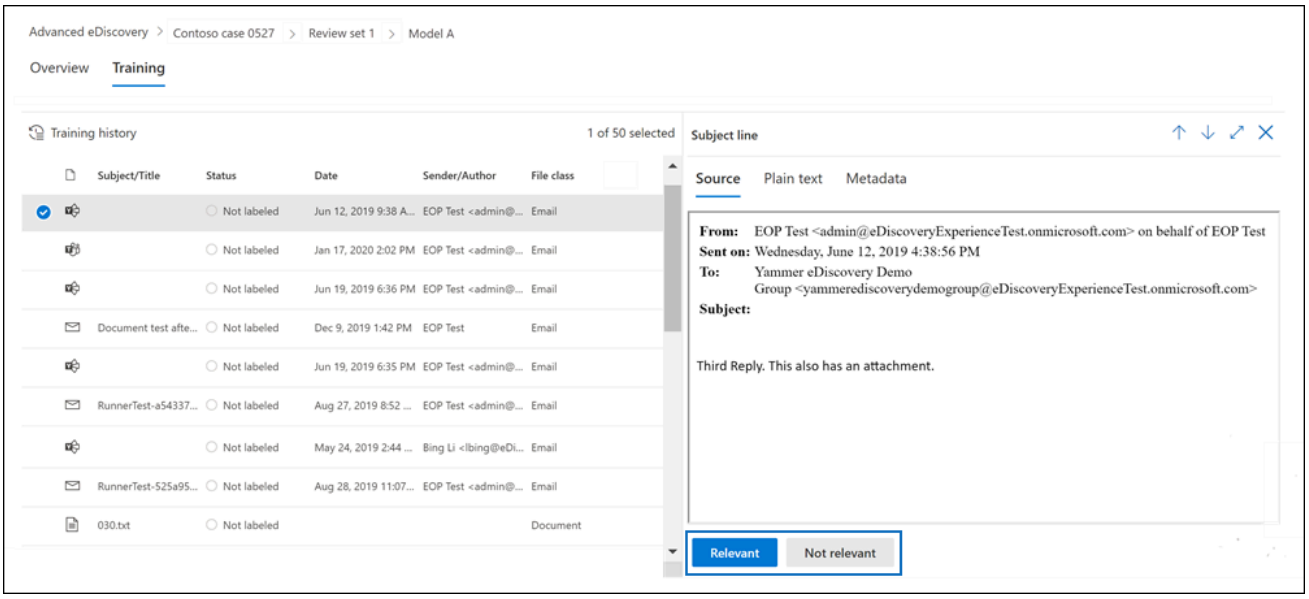
For more detailed instructions, see Create a predictive coding model.

Step 2: Perform the first training round

After you create the model, the next step is to complete the first training round by labeling items as relevant or not relevant.

1. Open the review set and then select **Analytics** > **Manage predictive coding (preview)**.
2. On the **Predictive coding models (preview)** page, select the model that you want to train.
3. On the **Overview** tab, under **Round 1**, select **Start next training round**.

   The **Training** tab is displayed and contains 50 items for you to label.
4. Review each document and then select **Relevant** or **Not relevant** at the bottom of the reading pane to label it.



5. After you've labeled all 50 items, select **Finish**.

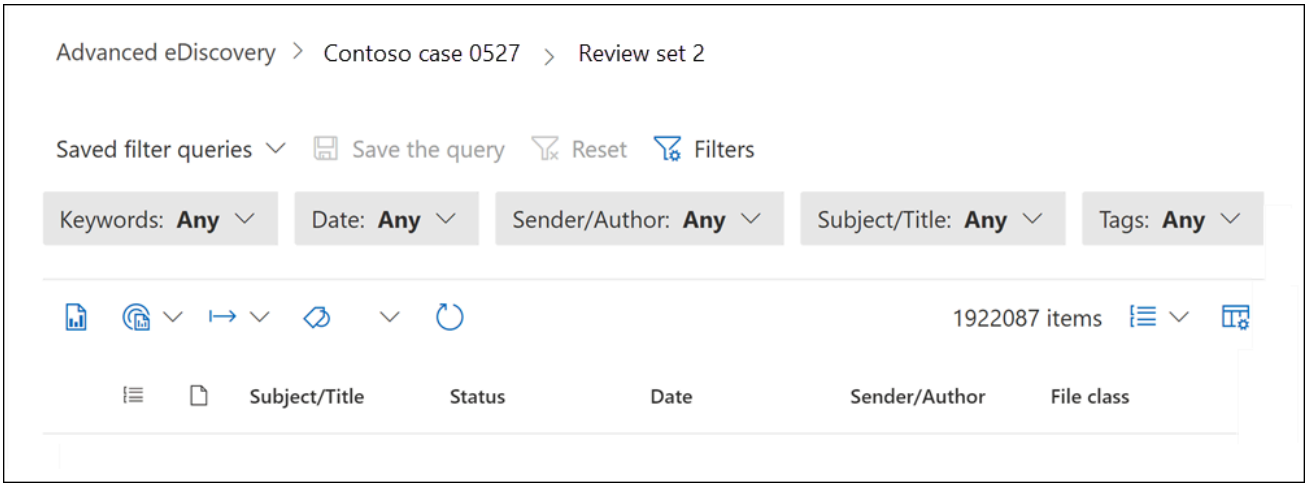
   It will take a couple minutes for the system to "learn" from your labeling and update the model. When this process is complete, a status of **Ready** is displayed for the model on the **Predictive coding models (preview)** page.

For more detailed instructions, see Train a predictive coding model.

Step 3: Apply the prediction score filter to items in review set

After you perform at lease one training round, you can apply the prediction score filter to items in review set. This lets you review the items the model has predicted as relevant or not relevant.

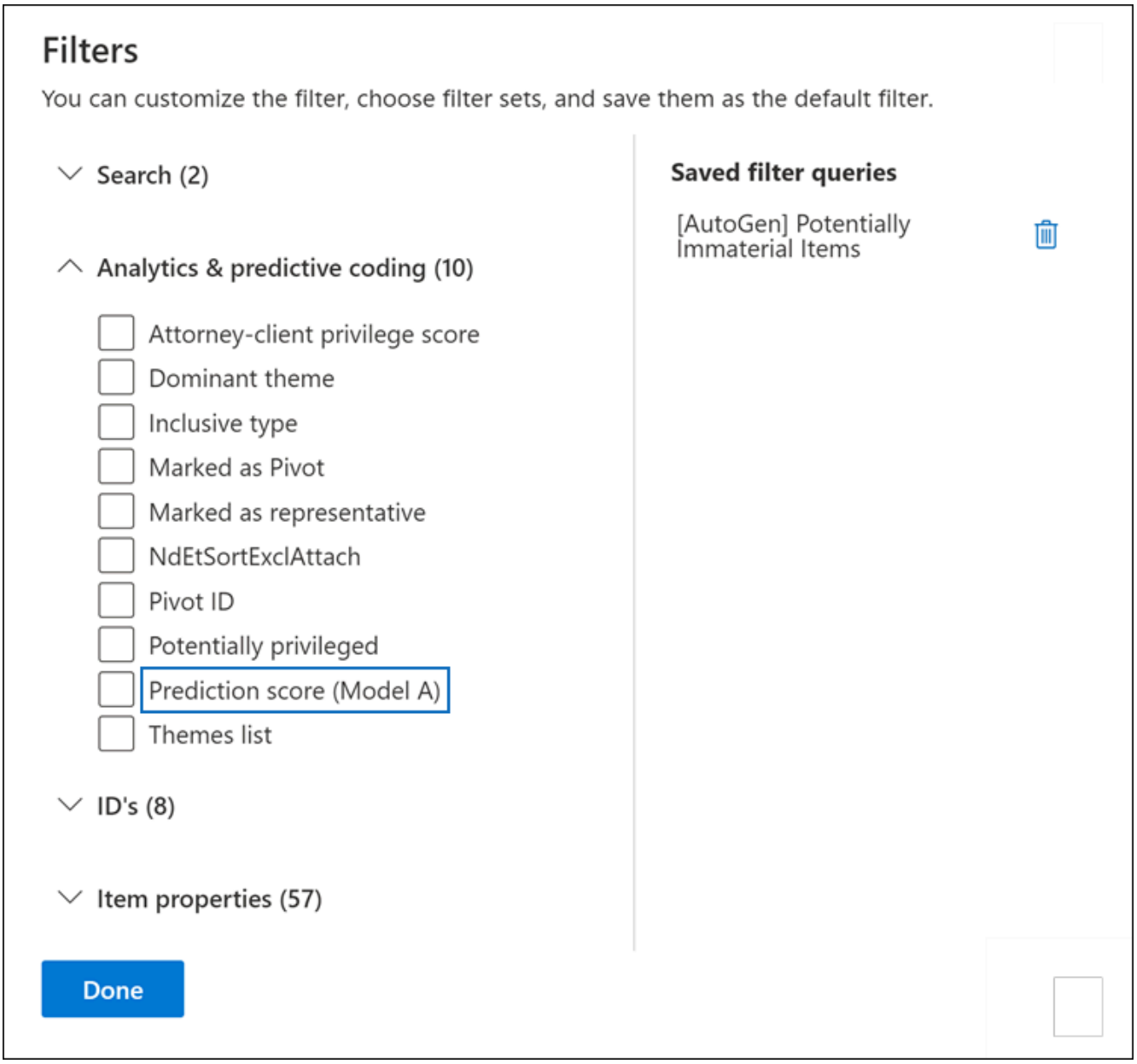
1. Open the review set.



   The pre-loaded default filters are displayed at the top of the review set page. You can leave these set to **Any**.

2. Select **Filters** to display the **Filters** flyout page.

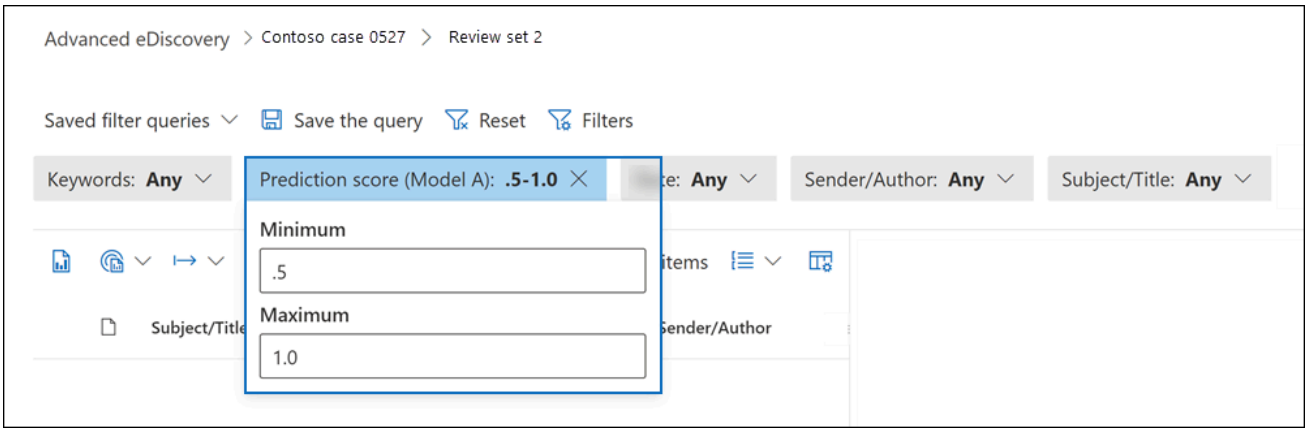
3. Expand the **Analytics & predictive coding** section to display a set of filters.



The naming convention for prediction score filters is **Prediction score (model name)**. For example, the prediction score filter name for a model named **Model A** is **Prediction score (Model A)**.

4. Select the prediction score filter that you want to use and then select **Done**.

5. On the review set page, select the dropdown for the prediction score filter and type minimum and maximum values for the prediction score range. For example, the following screenshot shows a prediction score range between **.5** and **1.0**.



6. Select outside the filter to automatically apply the filter to the review set.

A list of documents with a prediction score within the range you specified is displayed on the review set page.

For more detailed instructions, see Apply a prediction filter to a review set.

Step 4: Perform more training rounds

More than likely, you'll have to perform more training rounds to train the module to better predict relevant and non-relevant items in the review set. In general, you'll train the model enough times until it stabilizes enough to meet your requirements.

For more information, see Perform additional training rounds

Step 5: Apply the final prediction score filter to prioritize review

Repeat the instructions in Step 3 to apply the final prediction score to the review set to prioritize the review of relevant and non-relevant items after you complete all the training rounds and stabilize the model.

Last updated on 04/01/2026

Learn about predictive coding in eDiscovery (Premium) (preview)

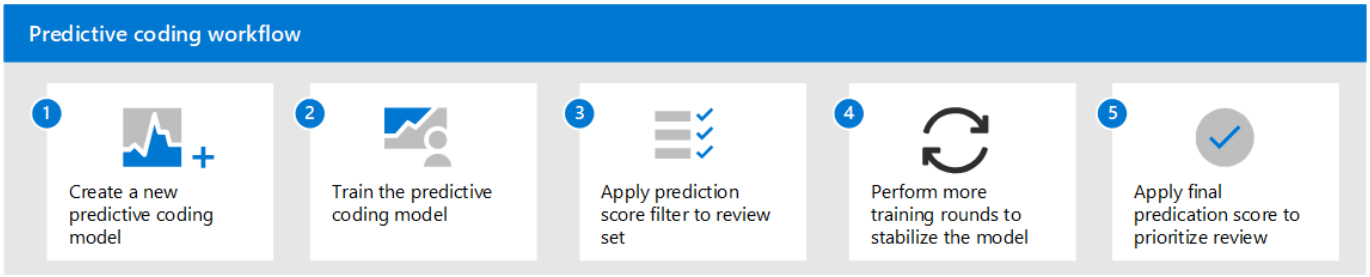
> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

The predictive coding module in eDiscovery (Premium) uses the intelligent, machine learning capabilities to help you reduce the amount of content to review. Predictive coding helps you reduce and cull large volumes of case content to a relevant set of items that you can prioritize for review. This is accomplished by creating and training your own predictive coding models that help you prioritize the review of the most relevant items in a review set.

The predictive coding module is designed to streamline the complexity of managing a model within a review set and provide an iterative approach to training your model so you can get started faster with the machine learning capabilities in eDiscovery (Premium). To get started, you can create a model, label as few as 50 items as relevant or not relevant. The system uses this training to apply prediction scores to every item in the review set. This lets you filter items based on the prediction score, which allows you to review the most relevant (or non-relevant) items first. If you want to train models with higher accuracies and recall rates, you can continue labeling items in subsequent training rounds until the model stabilizes.

The predictive coding workflow

Here's an overview and description of each step predictive coding workflow. For a more detailed description of the concepts and terminology of the predictive coding process, see Predictive coding reference.



1. **Create a new predictive coding model in the review set**. The first step is to create a new predictive coding model in the review set. You must have at least 2,000 items in the review set to create a model. After you create a model, the system will determine the number of items to use as a *control set*. The control set is used during the training process to evaluate the prediction scores that the model assigns to items with the labeling that you perform during training rounds. The size of the control set is based on the number of items in the review set and the confidence level and margin of error values that are set when creating the model. Items in the control set never change and aren't identifiable to users.

   For more information, see Create a predictive coding model.

2. **Complete the first training round by labeling items as relevant or not relevant**. The next step is to train the model by starting the first round of training. When you start a training round, the model randomly selects additional items from the review set, which is called the *training set*. These items (both from the control set and the training set) are presented to you so that you can label each one as either "relevant" or "not relevant". Relevancy is based on the content in the item and not any of the document metadata. After you complete the labeling process in the training round, the model will "learn" based on how you labeled the items in the training set. Based on this training, the model will process the items in the review set and apply a prediction score to each one.

   For more information, see Train a predictive coding model.

3. **Apply the prediction score filter to items in review set**. After the previous training step is completed, the next step is to apply the prediction score filter to the items in the review to display the items that the model has determined are "most relevant" (alternatively, you could also use a prediction filter to display items that are "not relevant"). When you apply the prediction filter, you specify a range of prediction scores to filter. The range of prediction scores fall between **0** and **1**, with **0** being "not-relevant" and **1** being relevant. In general, items with prediction scores between **0** and **0.5** are considered "not-relevant" and items with prediction scores between **0.5** and **1** are considered relevant.

   For more information, see Apply a prediction filter to a review set.

4. **Perform more training rounds until the model stabilizes**. You can perform additional rounds of training if you want to create a model with a higher accuracy of prediction and increased recall rates. *Recall rate* measures the proportion of items the model predicted were relevant among items that are actually relevant (the ones you marked as relevant during training). The recall rate score ranges from **0** to **1**. A score closer to **1** indicates the model will identify more relevant items. In a new training round, you label additional items in a new training set. After you complete that training round, the model is updated based on new learning from your most recent round of labeling items in the training set. The model will process the items in the review set again, and apply new prediction scores. You can continue performing training rounds until your model stabilizes. A model is considered stabilized when the churn rate after the latest round of training is less than 5%. *Churn rate* is defined as percentage of items in a review set where the prediction score changed between training rounds. The predictive coding dashboard displays information and statistics that help you assess the stability of a model.

5. **Apply the "final" prediction score filter to review set items to prioritize review**. After you complete all the training rounds and stabilize the model, the last step is to apply the final prediction score to the review set to prioritize the review of relevant and non-relevant items. This is the same task that you performed in step 3, but at this point the model is stable and you don't plan on running any more training rounds.

Last updated on 09/02/2025

Create a predictive coding model (preview)

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

The first step in using the machine learning capabilities of predictive coding in eDiscovery (Premium) is to create a predictive coding model. After you create a model, you can train it identify the relevant and non-relevant content in a review set.

To review the predictive coding workflow, see Learn about predictive coding in eDiscovery (Premium)

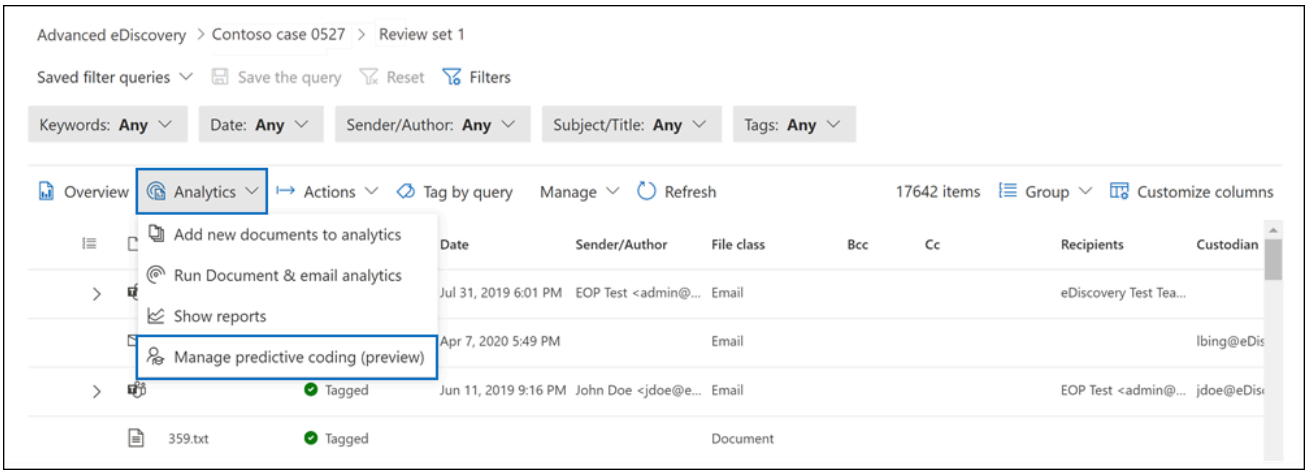
Before you create a model

- There must be a minimum of 2,000 items in a review set to create a predictive coding model.
- Be sure to commit all collections to the review set before you create a model. Items added to a review set after the model is created won't be processed and assigned a prediction score that generated by the model.
- Any item in the review set that doesn't contain text would won't be processed by the model or assigned a prediction score. Items with text will be included in the control set or a training set.

Create a model

1. In the Microsoft Purview portal, open an eDiscovery (Premium) case and then select the **Review sets** tab.

2. Open a review set and then select **Analytics** > **Manage predictive coding (preview)**.



3. On the **Predictive coding models (preview)** page, select **New model**.

4. On the flyout page, type a name for the model and an optional description.

5. Optionally, you can configure advanced settings (by selecting **Advanced options** on the flyout page) related to the confidence level and margin of error. These settings affect the number of items included in the control set. The *control set* is used during the training process to evaluate the prediction scores that the model assigns to items with the labeling that you perform during the training rounds. If your organization has guidelines about confidence level and margin of error for document review, specify them in the appropriate boxes. Otherwise, use the default settings.

6. Select **Save** to create the model.

   It takes a couple minutes for the system to prepare your model. After it's ready, you can perform the first round of training.

What happens after you create a model

After you create a model, the following things occur in the background during the creation and preparation of the model:

- The system calculates the number of items for the control set. This size is based on the number of items in the review set and the settings for the confidence level and the margin of error. Items for the control set are randomly selected and designated as control set items. The system includes 10 items from the control set in the first round of training.
- The system randomly selects 40 items from the review set to be included in the training set for the first round of training. Therefore, the first round of training includes 50 items for labeling: 40 items from the training set and 10 items from the control set.

Next steps

After you create a model for a review set, the next step is performing training rounds to "teach" the model to identify content that is relevant to your investigation. For more information, see Train a predictive coding model.

Last updated on 11/21/2025

Train a predictive coding model (preview)

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

After you create a predictive coding model in Microsoft Purview eDiscovery (Premium), the next step is to performing the first training round to train the model on what is relevant and non-relevant content in your review set. After you complete the first round of training, you can perform subsequent training rounds to improve the model's ability to predict relevant and non-relevant content.

To review the predictive coding workflow, see Learn about predictive coding in eDiscovery (Premium)

Before you train a model

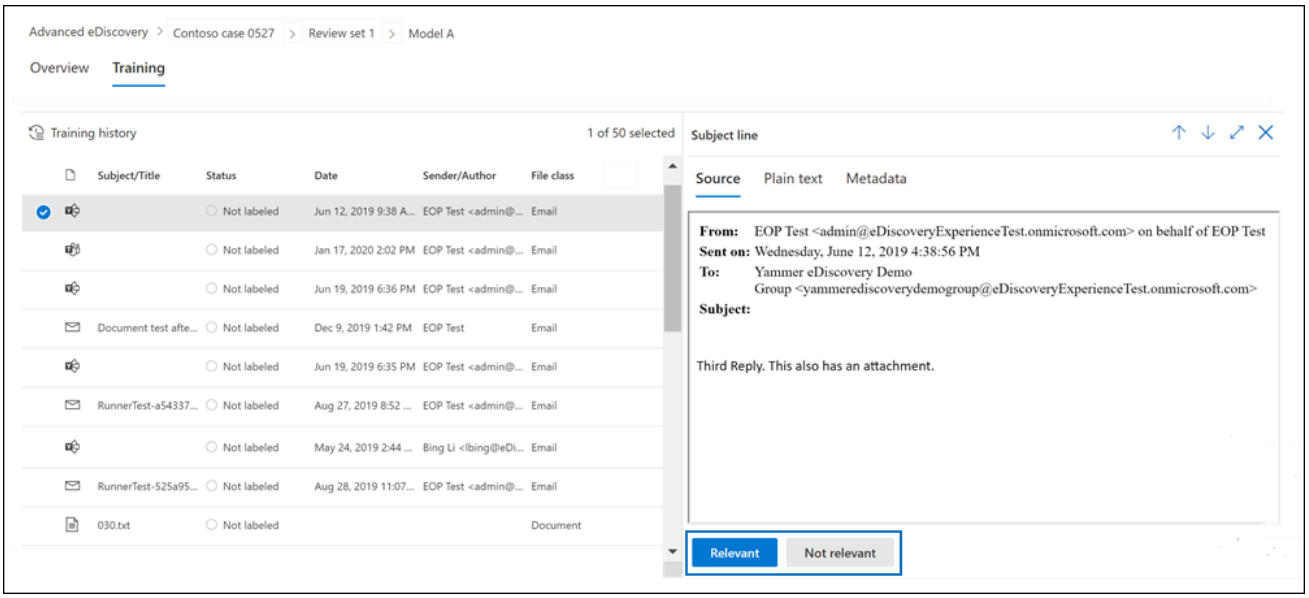
- During a training round, label items as **Relevant** or **Not relevant** based on the relevancy of the content in the document. Don't base your decision on the values in the metadata fields. For example, for email messages or Teams conversations, don't base your labeling decision on the message participants.

Train a model for the first time

1. In the Microsoft Purview portal, open an eDiscovery (Premium) case and then select the **Review sets** tab.
2. Open a review set and then select **Analytics** > **Manage predictive coding (preview)**.
3. On the **Predictive coding models (preview)** page, select the model that you want to train.
4. On the **Overview** tab, under **Round 1**, select **Start next training round**.

The **Training** tab is displayed and contains 50 items for you to label.

5. Review each document and then select **Relevant** or **Not relevant** at the bottom of the reading pane to label it.



6. After you've labeled all 50 items, select **Finish**.

   It will take a couple minutes for the system to "learn" from your labeling and update the model. When this process is complete, a status of **Ready** is displayed for the model on the **Predictive coding models (preview)** page.

Perform additional training rounds

After you perform the first round of training, you can perform subsequent training rounds by following the steps in the previous section. The only difference is the number of the training round will be updated on the model **Overview** tab. For example, after performing the first training round, you can select **Start next training round** to start the second round of training. And so on.

Each round of training (both those in progress and those that are complete) is displayed on the **Training** tab for the model. When you select a training round, a flyout page with information and metrics for the round is displayed.

What happens after you perform a training round

After you perform the first training round, a job is started that does the following things:

- Based on how you labeled the 40 items in the training set, the model learns from your labeling and updates itself to become more accurate.
- The model then processes each item in the entire review set and assigns a prediction score between **0** (not relevant) and **1** (relevant).
- The model assigns a prediction score to the 10 items in the control set that you labeled during the training round. The model compares the prediction score of these 10 items with the actual label that you assigned to the item during the training round. Based on this comparison, the model identifies the following classification (called the *Control set confusion matrix*) to assess the model's prediction performance:

Expand table

Label	Model predicts item is relevant	Model predicts item is not relevant
Reviewer labels item as relevant	True positive	False positive
Reviewer labels item as not relevant	False negative	True negative

Based on these comparisons, the model derives values for the F-score, precision, and recall metrics and the margin of error for each one. Scores for these model performance metrics are displayed on a flyout page for the training round. For a description of these metrics, see Predictive coding reference.

- Finally, the model determines the next 50 items that will be used for the next training round. This time, the model might select 20 items from the control set and 30 new items from the review set and designate them as the training set for the next round. The sampling for the next training round is not uniformly sampled. The model will optimize the sampling selection of items from the review set to select items where the prediction is ambiguous, which means the prediction score is in the 0.5 range. This process is known as *biased selection*.

What happens after you perform subsequent training rounds

After you perform subsequent training rounds (after the first training round), the model does the following things:

- The model is updated based on the labels that you applied to the training set in that round of training.
- The system evaluates the model's prediction score on the items in the control set and check whether the score aligns with how you labeled items in the control set. The evaluation is performed on all labeled items from control set for all training rounds. The results of this evaluation are incorporated in the dashboard on the **Overview** tab for the model.
- The updated model reprocesses every item in the review set and assign each item an updated prediction score.

Next steps

After you perform the first training round, you can perform more training rounds or apply the model's prediction score filter to the review set to view the items the model has predicted as relevant or not relevant. For more information, see Apply a prediction score filter to a review set.

Last updated on 11/21/2025

Apply a prediction score filter to a review set (preview)

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

After you create a predictive coding model in Microsoft Purview eDiscovery (Premium) and train it to the point where it's stable, you can apply the prediction score filter to display review set items that the model has determined are relevant (or not relevant). When you create a model, a corresponding prediction score filter is also created. You can use this filter to display items assigned a prediction score within a specified range. In general, prediction scores between **0** and **.5** are assigned to items that model has predicted are not relevant. Items assigned prediction scores between **.5** and **1.0** are items the model has predicted are relevant.

Here are two ways you can use the prediction score filter:

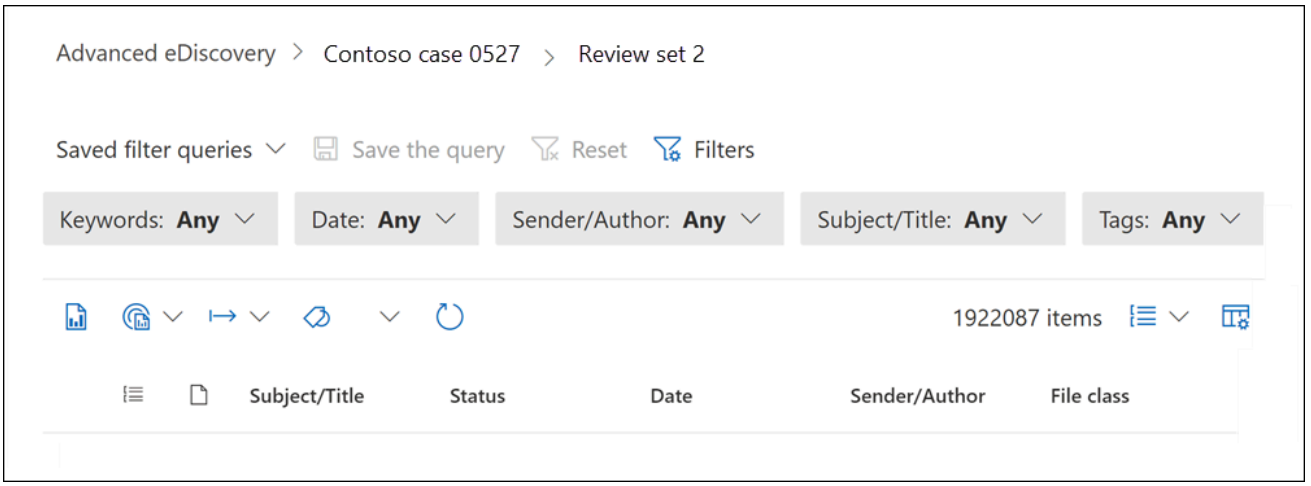
- Prioritize the review of items in a review set that the model has predicted are relevant.
- Cull items from the review set that the model has predicted aren't relevant. Alternative, you can use the prediction score filter to deprioritize the review of nonrelevant items.

Before you apply a prediction score filter

- Create a predictive coding model so that a corresponding prediction score filter is created.
- You can apply a prediction score filter after any of the training rounds. But you may want to wait after performing several rounds or until the model is stable before using the prediction score filter.

Apply a prediction score filter

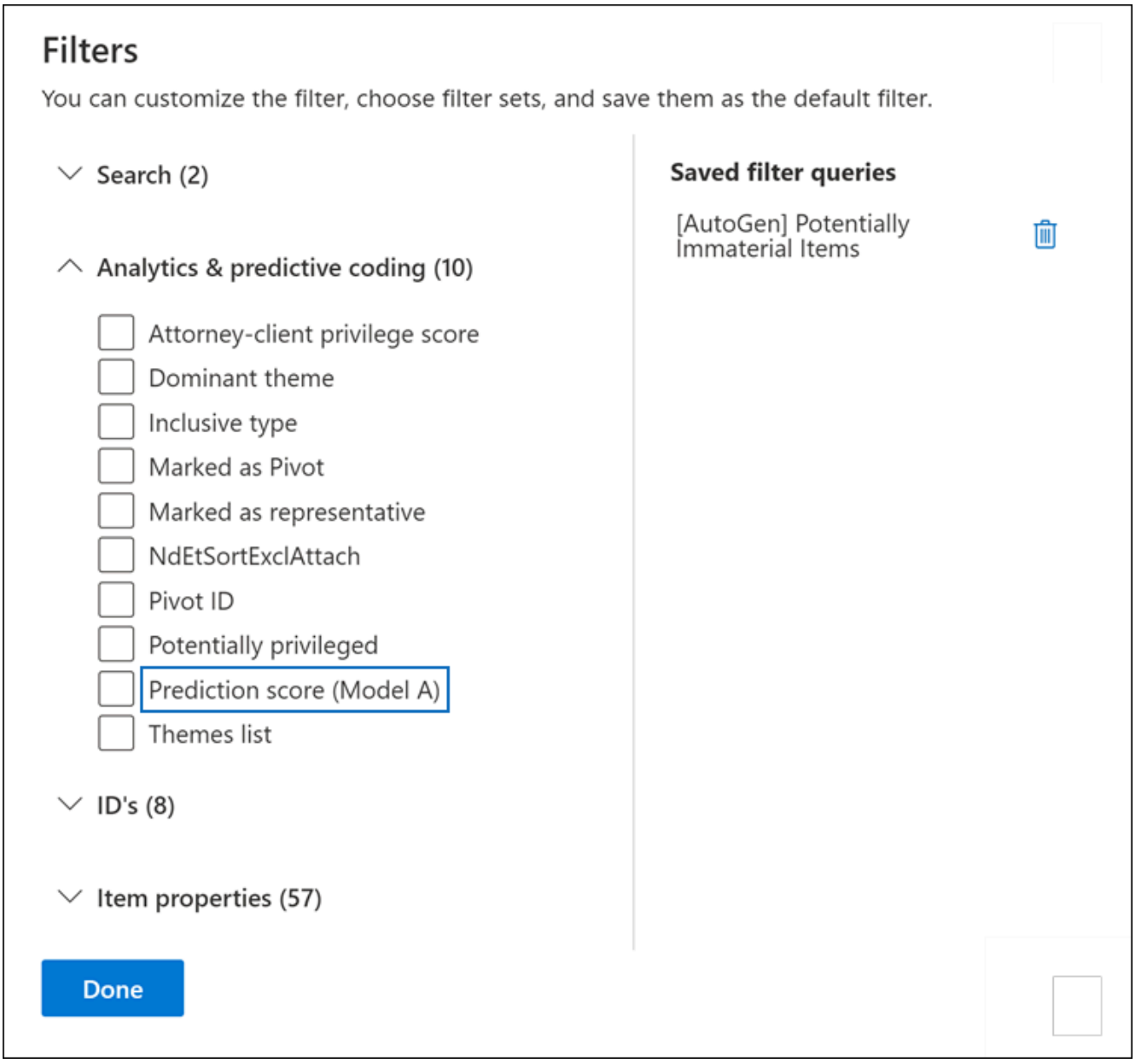
1. In the Microsoft Purview portal, open the eDiscovery (Premium) case, select the **Review sets** tab, and then open the review set.



   The preloaded default filters are displayed at the top of the review set page. You can leave these set to **Any**.

2. Select **Filters** to display the **Filters** flyout page.

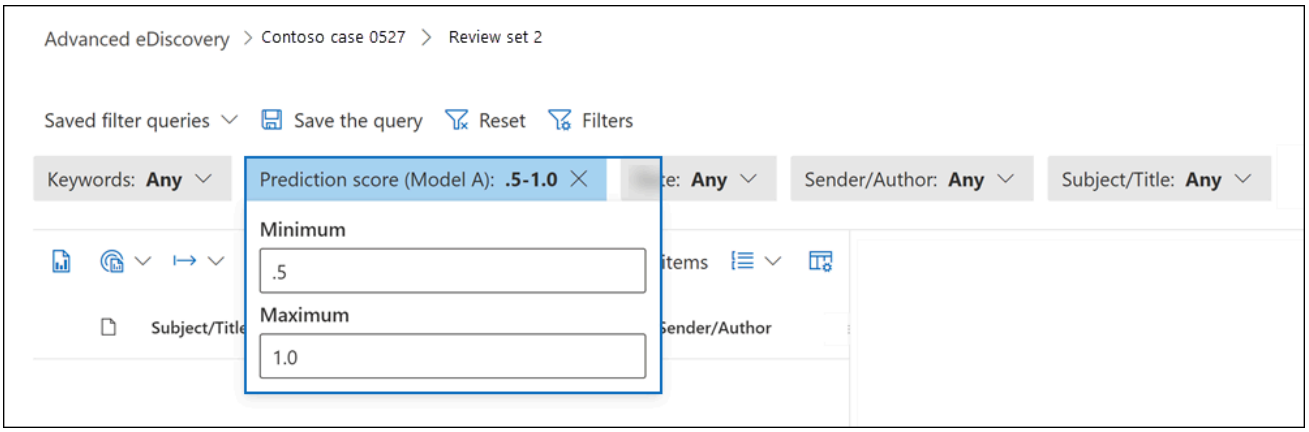
3. Expand the **Analytics & predictive coding** section to display a set of filters.



The naming convention for prediction score filters is **Prediction score (model name)**. For example, the prediction score filter name for a model named **Model A** is **Prediction score (Model A)**.

4. Select the prediction score filter that you want to use and then select **Done**.

5. On the review set page, select the dropdown for the prediction score filter and type minimum and maximum values for the prediction score range. For example, the following screenshot shows a prediction score range between **.5** and **1.0**.



6. Select outside the filter to automatically apply the filter to the review set.

A list of documents with a prediction score within the range you specified is displayed on the review set page.

> **Tip**
>
> To view the actual prediction score assign to a document, you can select the **Metadata** tab in the reading pane. The prediction scores for all models in the review set are displayed in the **RelevanceScores** metadata property.

More information

- For more information about using filters, see Query and filter content in a review set.

Last updated on 11/21/2025

Predictive coding reference (preview)

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

This article describes the key concepts and metrics of the predictive coding tool in Microsoft Purview eDiscovery (Premium). The sections in the article are listed in alphabetical order.

Confidence level

The confidence level is an advanced setting when you create a predictive coding model. It defines that the model's performance metrics (for example, richness, precision, and recall) fall within a specified range (that's determined the margin of error defined for the model) that's representative of the true values of the prediction scores the model assigns to items in the review set. The values for the confidence level and margin of error also help determine how many items are included in the control set. The default value for the confidence level is 0.95% or 95%.

Control set

A control set is used during the training process of a predictive coding model. The control set is to evaluate the prediction scores that the model assigns to items with the labeling that you perform during training rounds. The size of the control set is based on the number of items in the review set and the confidence level and margin of error values that are set when creating the model. Items in the control set never change and aren't identifiable to users. The total number of items in the control set is displayed on the flyout page for a training round.

Control set confusion matrix

After you complete a training round, the model assigns a prediction score to the 10 items in the control set that you labeled during the training round. The model compares the prediction score of these 10 items with the actual label that you assigned to the item during the training round. Based on this comparison, the model identifies the following classifications to assess the model's prediction performance:

Expand table

Label	Model predicts item is relevant	Model predicts item isn't relevant
Reviewer labels item as relevant	True positive	False positive
Reviewer labels item as not relevant	False negative	True negative

Based on these comparisons, the model derives values for the F-score, precision, and recall metrics and the margin of error for each one. The number of each of the confusion types from the matrix is displayed on the flyout page for a training round.

F-score

The F-score is a weighted average of the scores for the precision and recall metrics. The range of scores for this metric is from **0** to **1**. A score closer to **1** indicates the model will more accurately detect relevant items. The F-score metric is displayed on the model dashboard and on the flyout page for each training round.

Margin of error

The margin of error is an advanced setting when you create a predictive coding mode. It specifies the degree of error in performance metrics (for example, richness, precision, and recall) that's derived from the random sampling of items in your control set. A lower margin of error requires a larger control set to ensure that the model's performance metrics fall within a smaller range. The values for the margin of error and confidence level also help determine how many items are included in the control set. The default value for the margin of error is 0.05% or 5%.

Model stability

Model stability indicates the model's ability to accurately predict whether a document in a review set is relevant or not relevant. When a model is unstable, more training rounds may need to be performed to include the model's stability. When the model is stable, no more training rounds may need to be performed. The model dashboard indicates the current state of the model's stability. When a model is stable, the performance metrics have reached a level that matches the settings for the confidence level and margin of error.

Overturn rate

The overturn rate is the percentage of items in the review set where the prediction score changed between training rounds. A model is considered stable when the overturn rate is less than 5%. The overturn rate metric is displayed on the model dashboard and on the flyout page for each training round. The overturn rate for the first training round is zero because there isn't a previous prediction score to overturn.

Precision

The precision metric measures the proportion of items that are relevant among the items the model predicted were relevant. This means that items in the control set where label as relevant by the reviewer and predicted as relevant by the model. The range of scores for this metric is from **0** to **1**. A score closer to **1** indicates the model will identify fewer non-relevant items. The precision metric is displayed on the model dashboard and on the flyout page for each training round.

Prediction score

This is the score that a model assigns to each document in a review set. The score is based on the document's relevance compared to model's learning from the training rounds. In general, items with prediction scores between **0** and **0.5** are considered not relevant and items with prediction scores between **0.5** and **1** are considered relevant. The prediction score is contained in a document metadata field. You can use a prediction filter to display the items in a review set that fall within a specified prediction range.

Recall

The recall metric measures the proportion of items the model predicted were relevant among items that are relevant. This means that relevant items in the control set that the model predicted were also labeled as relevant by the reviewer. The range of scores for this metric is from **0** to **1**. A

score closer to **1** indicates the model will identify a larger portion of relevant items. The recall metric is displayed on the model dashboard and on the flyout page for each training round.

Review set

A review set provides the scope of a predictive coding model. When you create a new model for the review set, items for the control set and training sets are selected from the review set. When the model assigns prediction scores, it assigns those scores the items in the review. You have to add all items to the review set before you create a predictive coding model. If you add items after you create a model, those items won't be assigned a prediction score.

Richness

The richness metric measures the percentage of review set items the model predicts as relevant. The range of scores for this metric is from **0** to **1**. The richness metric is displayed on the model dashboard.

Sampled items

The term *sampled items* is a reference to random sample of items in a review set (that contains text) that are selected and associated with the control set when you create a predictive coding model. A random sample of items is also selected for each training round. Items selected for the control set of a model are never included in a training set for that same model. The reverse is also true: training set items are never included in the control set.

Training set

The model randomly selects items from the review set and adds them to a training set. During a training round, items from the training set (in addition to items from the control set) are presented to you so that you can label each one as either "relevant" or "not relevant". This labeling or "training" process helps the model learn how to predict which items in the review are relevant or not relevant. Each time you perform a training round, the model selects more items from the review and adds them to the training set for that training round. Items from the control set are never selected for a training set.

Last updated on 09/02/2025

Export case data in eDiscovery (Premium)

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

There are three ways to export data from a review set:

Download: Download (by using a browser) a small set of native files. This is the quickest way to export a small set of data. This method preserves the native file names. For more information, see Download documents from a review set.

Export: Customize what data is exported. This includes exporting file metadata, native files, text files, and redacted documents that have been saved to a PDF file. Once the exported data is prepared, you can download the data from the **Export** tab to a local computer. For more information, see Export documents from a review set

Add to another review set: Copy data from one review set to a different review set. For more information, see Add data from one review set to another review set.

> **Note**
>
> In Microsoft 365, data is hashed and those hashes are provided in the output file for verification purposes. This is supplemented by audit logs and reporting functionality, such as collection statistics, load set reports, and export reports (including the export load file).

Last updated on 11/21/2025

Download selected documents from a review set

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

Download offers a simple way to download content from a review set in native format. The download tool in eDiscovery (Premium) uses the browser's data transfer features. A browser prompt appears when a download is ready.

Files downloaded using this method are zipped in a container file, and this file contains item-level files. It means that if you select to download an attachment, you receive the email message with the attachment included. Similarly, if you export an Excel spreadsheet that is embedded in a Word document, the Word document and the embedded Excel spreadsheet are included in the download.

When you download items, the *Last Modified Data* property is preserved and can be viewed as a file property.

> **Important**
>
> Downloading an item from a review set does not preserve the item metadata. Item metadata is only preserved when exporting items from a review set.

To download a review set item from a review set in an eDiscovery (Premium) case, start by selecting the item you want to download and select **Download original** from the item menu.

> **Important**

To download original items from a review set, you must be assigned the *Export* role or be assigned to a role group that includes the *Export* role. For more information about eDiscovery permissions, see **Assign eDiscovery permissions in the Microsoft Purview portal**.

To download multiple items from a review set at the same time, select the items you want to download then export the selected documents from the review set.

Last updated on 11/21/2025

Export documents from a review set in eDiscovery (Premium)

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

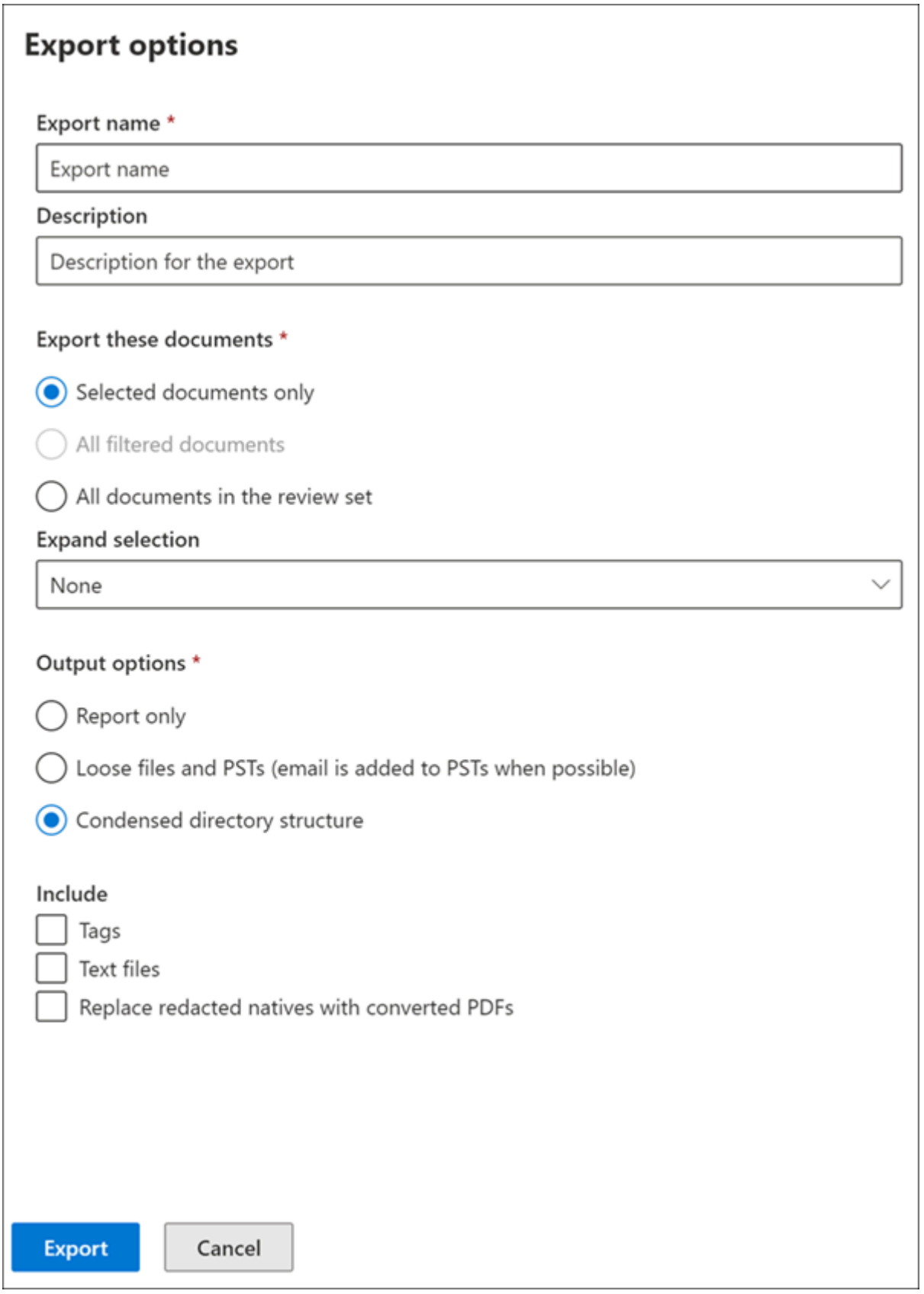
Export allows users to customize the content that is included in the download package when you export document from a review set in eDiscovery (Premium).

To export documents from a review set:

> **Important**
>
> Exporting documents to your Azure Storage locations in eDiscovery was permanently **retired on August 31, 2025**, this functionality isn't available in the new eDiscovery experience.

1. In the Microsoft Purview portal, open the eDiscovery (Premium) case, select the **Review sets** tab, and then select the review set that you want to export.
2. In the review set, select the items to export.
3. Select **Action** > **Export**.

   The Export tool displays the flyout page with the settings to configure the export. Some options are selected by default, but you can change these. See the following section for descriptions of the export options that you can configure.



4. After you configure the export options, select **Export** to start the export process. Depending on the option that you selected in **Output options** section, you can access the export files by direct download or in your organization's Azure Storage account.

> ⓘ **Note**

Export jobs are retained for the life of the case. However, you must download the content from an export job within 30 days after the export job is complete.

Export options

Use the following options to configure the export. Not all options are allowed for some output options, most notably, the export of text files and redacted PDFs aren't allowed when exporting to the PST format.

- **Export name**: Name of the export job. This is used to name the ZIP files that will be downloaded.
- **Description**: Free-text field for you to add a description.
- **Export these documents**
  - *Selected documents only*: This option exports only the documents that are currently selected. This option is only available when items are selected in a review set.
  - *All filtered documents*: This option exports the documents in an active filter. This option is only available when a filter is applied to the review set.
  - *All documents in the review set*: This option exports all documents in the review set.
- **Expand selection**: Use the following options to export additional items that are related to filtered/selected items in the review set:
  - *None*: This option doesn't export the family items or conversation items. It only exports the items that are selected or to all items in the review set list.
  - *Include associated family items*: This option includes associated family items. Family items are items that share the same *FamilyId* metadata property value. For example, a document that's attached to an email message shares the same *FamilyId* as the email message. If this option is selected, the email message and the document are exported in this example, even though the document might not be included in the list of review set items.
  - *Include associated conversation items*: This option includes associated items that are in the same Teams or Viva Engage conversation. Conversation items are items that share the same *ConversationId* metadata property value. All messages, posts, and corresponding transcript file of a conversation share the same *ConversationId*.

  If any option other than *None* is selected, all items that share the same association ID are exported, even though some of those items might not be in the results of the current review

set query. For more information about conversation items, see eDiscovery (Premium) workflow for content in Microsoft Teams.

- **Output options**: Exported content is either available for download directly through a web browser or can be sent to an Azure Storage account. The first two options enable direct download.
  - *Reports only*: Only the summary and load file are created.
  - *Loose files and PSTs (email is added to PSTs when possible)*: Files are exported in a format that resembles the original directory structure seen by users in their native applications. For more information, see the Loose files and PST export structure section.
  - *Condensed directory structure*: Files are exported and included in the download.
  - *Condensed directory structure exported to your Azure Storage account*: Files are exported to your organization's Azure Storage account. For this option, you have to provide the URL for the container in your Azure Storage account to export the files to. You also have to provide the shared access signature (SAS) token for your Azure Storage account. For more information, see Export documents in a review set to an Azure Storage account.

    > **Note**
    >
    > Condensed output options include all extracted items. The item count for condensed output is always more than the *Loose files and PSTs* option because email and attachments are counted differently. For example, for an email with two attachments, the condensed output results in 3 items (1 item for the email and 1 item for each attachment) while the *Loose files and PSTs* option would result in 1 item for the email and all attachments.

- **Include**
  - *Tags*: When selected, tagging information is included in the load file.
  - *Text files*: This option includes the extracted text versions of native files in the export.
  - *Replace redacted natives with converted PDFs*: If redacted PDF files are generated during review, these files are available for export. You can choose to export only the native files that were redacted (by not selecting this option) or you can select this option to export the PDF files that contain the actual redactions.

The following sections describe the folder structure for loose files and condensed directory structure options. Exports are partitioned into ZIP files with a maximum size of uncompressed

content of 75 GB. If the export size is less than 75 GB, the export consists of a summary file and a single ZIP file. For exports larger than 75 GB of uncompressed data, multiple ZIP files are created. Once downloaded, the ZIP files can be uncompressed into a single location to recreate the full export.

Loose files and PST export structure

If you select this export option, the exported content is organized in the following structure:

- **Summary.csv**: Includes a summary of the content exported from the review set. This file is structured with the following columns:
  - *Total*: The number of the items found in the review set.
  - *Actual*: he number of items fetched in the package for the export.
  - *Errors*: The number of errors and warnings. More information can be found in the errors and warnings file. The warning and error file contains important information about the items that weren't exported. Check this file to make sure you understand which files have issues.
  - *Skipped Processing*: The number of items skipped for processing.

    > ⓘ **Note**
    >
    > Skipped processing also contains the number of extracted child items for PST format. For example, a review set has two items, one corresponding to an email and one corresponding to an attachment. When the file is converted to the PST, it is mappped to one item only (the mail item). This results in skipping the child attachment.

  - *Export Containers*: This contains the list of .zip and .pst files created for the export.
- Root folder: This folder in named [Export Name] x of z.zip and is repeated for each .zip file partition. The root folder contains the following:
  - *Export_load_file_x of z.csv*: The metadata file.
  - *Warnings and errors x of z.csv*: This file includes information about errors encountered when trying to export from the review set.
- **Exchange**: This folder contains all content from Exchange stored in PST files. Redacted PDF files can't be included with this option. If an attachment is selected in the review set, the parent email message is exported with the attachment attached.

  The Exchange folder may also contain a subfolder named mailboxname_loosefiles.zip, which contains the following items:

- Information Rights Management (IRM) protected messages that have been decoded.
- Error-remediated messages.
- Modern attachments or links referenced in messages.
- Encrypted items (which aren't included in the PST files in the Exchange folder).

- **SharePoint**: This folder contains all native content from SharePoint in a native file format. Redacted PDF files can't be included with this option.

> ⓘ **Note**
>
> The *Group* option selected in the review set determines which items are included in the content exported from the review set. If *None* is selected, the exported content only includes the selected or filtered items. If *Group by family* is selected, the exported content includes all items in the family.

Condensed directory structure

- **Summary.csv**: Includes a summary of the content exported from the review set. This file is structured with the following columns:
  - *Total*: The number of the items found in the review set.
  - *Actual*: he number of items fetched in the package for the export.
  - *Errors*: The number of errors and warnings. More information can be found in the errors and warnings file.
  - *Skipped Processing*: The number of items skipped for processing.

    > ⓘ **Note**
    >
    > Skipped processing also contains the number of extracted child items for PST format. For example, a review set has two items, one corresponding to an email and one corresponding to an attachment. When the file is converted to the PST, it is mappped to one item only (the mail item). This results in skipping the child attachment.

  - *Export Containers*: This contains the list of .zip and .pst files created for the export.

- **Root folder**: This folder in named [Export Name] x of z.zip and will be repeated for each ZIP file partition.
  - *Export_load_file_x of z.csv*: The metadata file and also includes the location of each file that is stored in the ZIP file

- *Warnings and errors x of z.csv*: This file includes information about errors encountered when trying to export from the review set.
- *NativeFiles*: This folder contains all the native files that were exported. Natives files are replaced with redacted PDFs if you selected the *Replace redacted natives with converted PDFs* option.
- *Error_files*: This folder contains files that had either extraction or other processing error. The files are placed into separate folders, either ExtractionError or ProcessingError. These files are listed in the load file.
- *Extracted_text_files*: This folder contains all of the extracted text files that were generated at processing.

Condensed directory structure exported to your Azure Storage Account

This option uses the same general structure as the *Condensed directory structure*, the contents are zipped and the data is saved to your Azure Storage account. This option is used when working with a third-party eDiscovery provider. For details about how to use this option, see Export documents in a review set to an Azure Storage account.

Last updated on 09/02/2025

Export documents in a review set to an Azure Storage account

When you export documents from a review set in an eDiscovery (Premium) case, you have the option to export them to an Azure Storage account managed by your organization. If you use this option, the documents are uploaded to your Azure Storage location. After they're exported, you can access the documents (and download them to a local computer or other location) by using the Azure Storage Explorer.

This article provides instructions for how to export documents to your Azure Storage account and the use the Azure Storage Explorer to connect to an Azure Storage location to download the exported documents. For more information about Azure Storage Explorer, see Use Azure Storage Explorer.

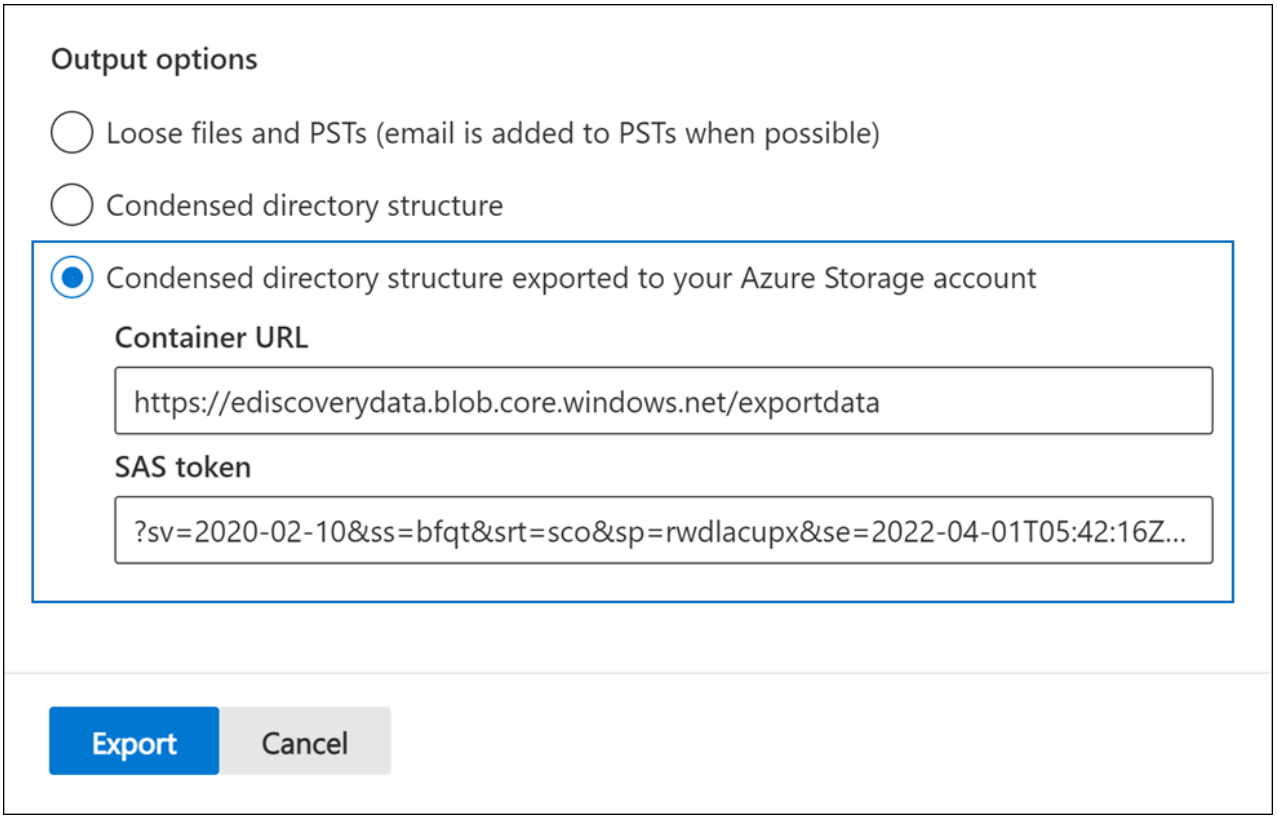
Before you export documents from a review set

- You need to provide a shared access signature (SAS) token for your Azure Storage account and the URL for a specific container in the storage account to export documents from a review set. Be sure to have these at hand (for example, copied to a text file) when you perform Step 2.
    - **SAS token**: Be sure to get the SAS token for your Azure Storage account (and not for the container). You can generate an SAS token for your account in Azure Storage. To do this, go to the Azure Storage account, and select **Share access signature** under the **Settings** settings in the storage account blade. Make sure you've configured *Add*, *Create*, *Delete*, *Read*, and *Write* permissions and allow all resource types when you generate the SAS token.
    - **Container URL**: You need to create a container to upload the review set documents to, and then get a copy of the URL for the container; for example, `https://ediscoverydata.blob.core.windows.net/exportdata`. To get the URL, go to the container in Azure Storage, and select **Properties** under the **Settings** section in the container blade.
- Download and install the Azure Storage Explorer. For instructions, see Azure Storage Explorer tool. You use this tool to connect to the container in your Azure Storage account and download the documents that you exported in Step 1.

Step 1: Export the documents from a review set

The first step is to create an export job to export documents out of a review set. For more detailed instructions about all the export options, see Export documents from a review set. The following procedure highlights the settings to export documents to your organization's Azure Storage account.

1. In the Microsoft Purview portal, open the eDiscovery (Premium) case, select the **Review sets** tab, and then select the review set that you want to export.
2. In the review set, select **Action** > **Export**.
3. On the **Export options** flyout page, type a name (required) and description (optional) for the export.
4. Configure the settings in the documents, metadata, content, and options sections. For more information about these settings, see Export documents from a review set.
5. In the **Output options** section, select the **Condensed directory structure exported to your Azure Storage account** option.
6. Paste the container URL and the SAS token for your storage account in the corresponding fields.



> **Important**
>
> If the container URL contains unsupported or invalid characters, you'll receive an error when creating the export job. For more information about these characters, see

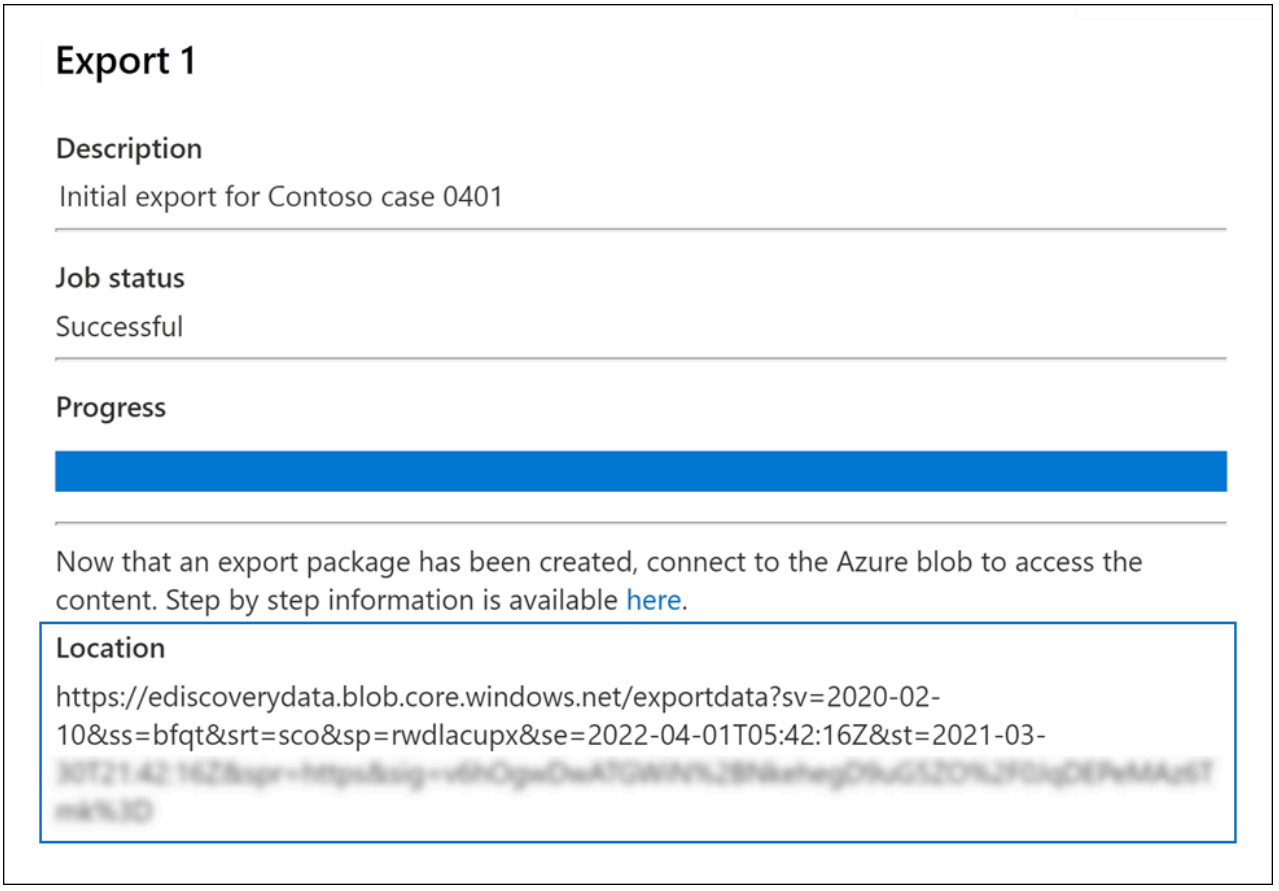
Keyword queries and search conditions for eDiscovery.

7. Select **Export** to create the export job.

Step 2: Obtain the SAS URL from the export job

The next step is to obtain the SAS URL that's generated after you create the export job in Step 1. You use the SAS URL to connect to the container in your Azure Storage account that you exported the review set documents to.

1. On the **eDiscovery (Premium)** page, go to the case, and then select the **Exports** tab.
2. On the **Exports** tab, select the export job that you want to download. This is the export job that you created in Step 1.
3. On the flyout page, under **Locations**, copy the SAS URL that's displayed. If necessary, you can save it to a text file so you can access it in Step 3.



Tip

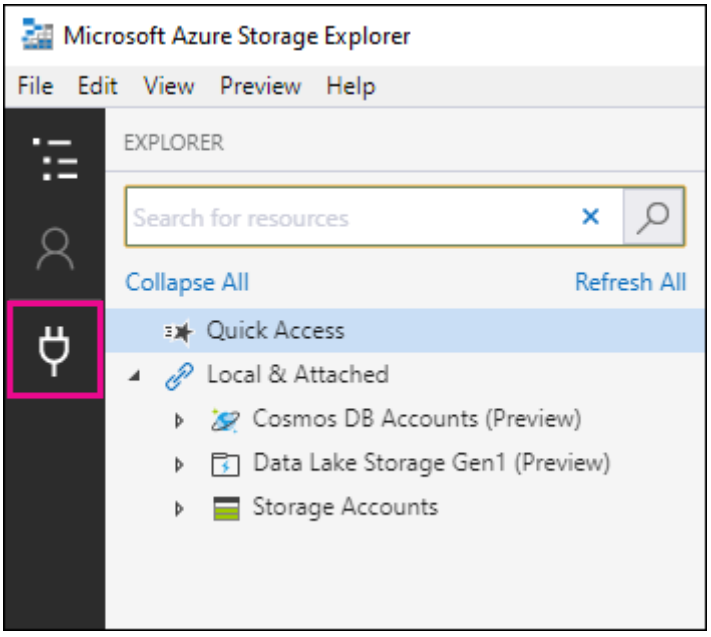
> The SAS URL that's displayed in the export job is a concatenation of the container URL and the SAS token for your Azure Storage account. You can copy it from the export job or create it yourself by combining the URL and the SAS token.

Step 3: Connect to the Azure Storage container

The final step is to use the Azure Storage Explorer and the SAS URL to connect to the container in your Azure Storage account and download the exported documents to a local computer.

1. Start the Azure Storage Explorer that you downloaded and installed.

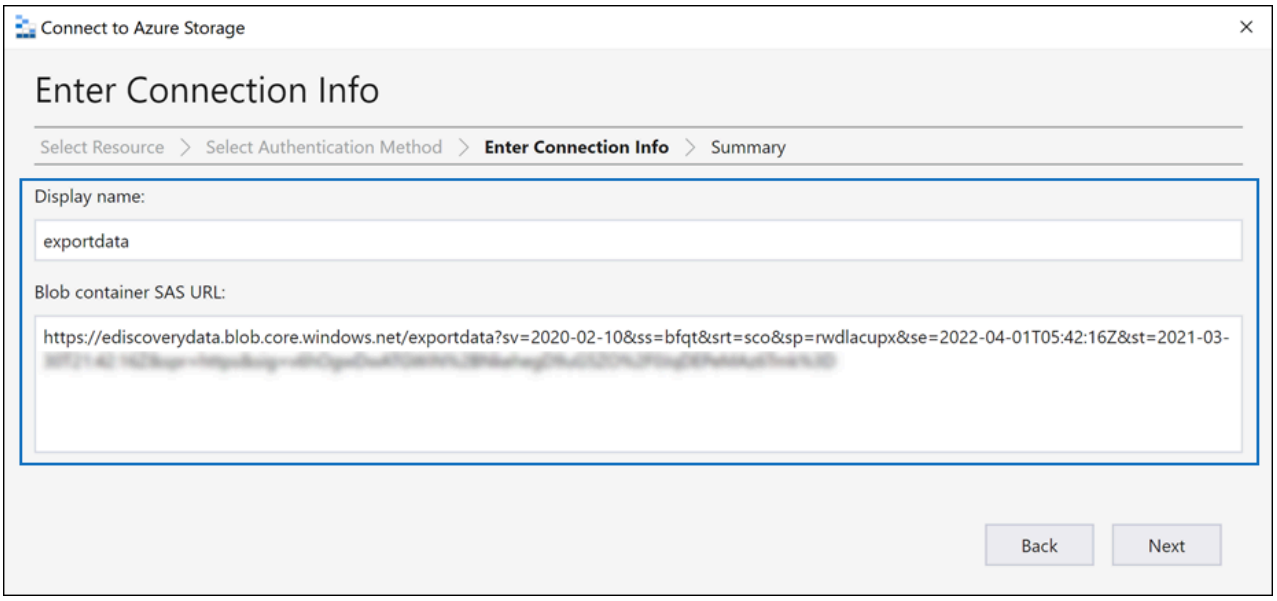
2. Select the **Open Connect Dialog** icon.



3. On the **Connect to Azure Storage** page, select **Blob container**.

4. On the **Select Authentication Method** page, select the **Shared access signature (SAS)** option, and then select **Next**.

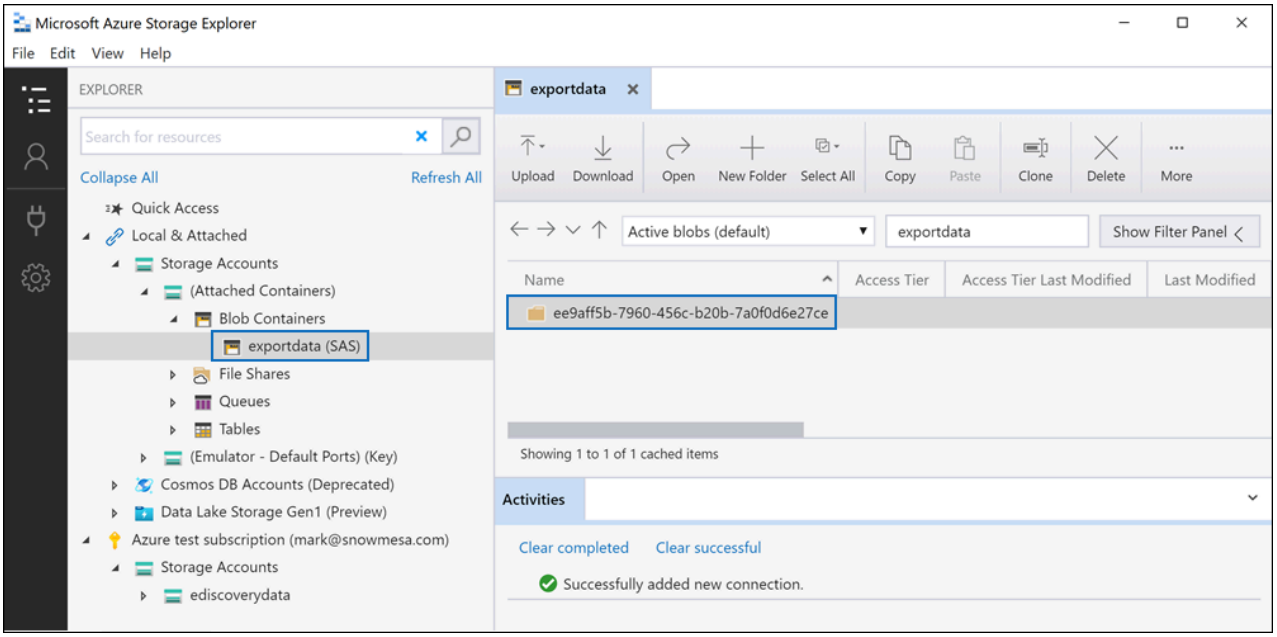
5. On the **Enter Connection Info** page, paste the SAS URL (that you obtained in the export job in Step 2) in the **Blob Container SAS URL** box.



Notice that the container name is displayed in the **Display name** box. You can edit this name.

6. Select **Next** to display the **summary** page and then select **Connect**.

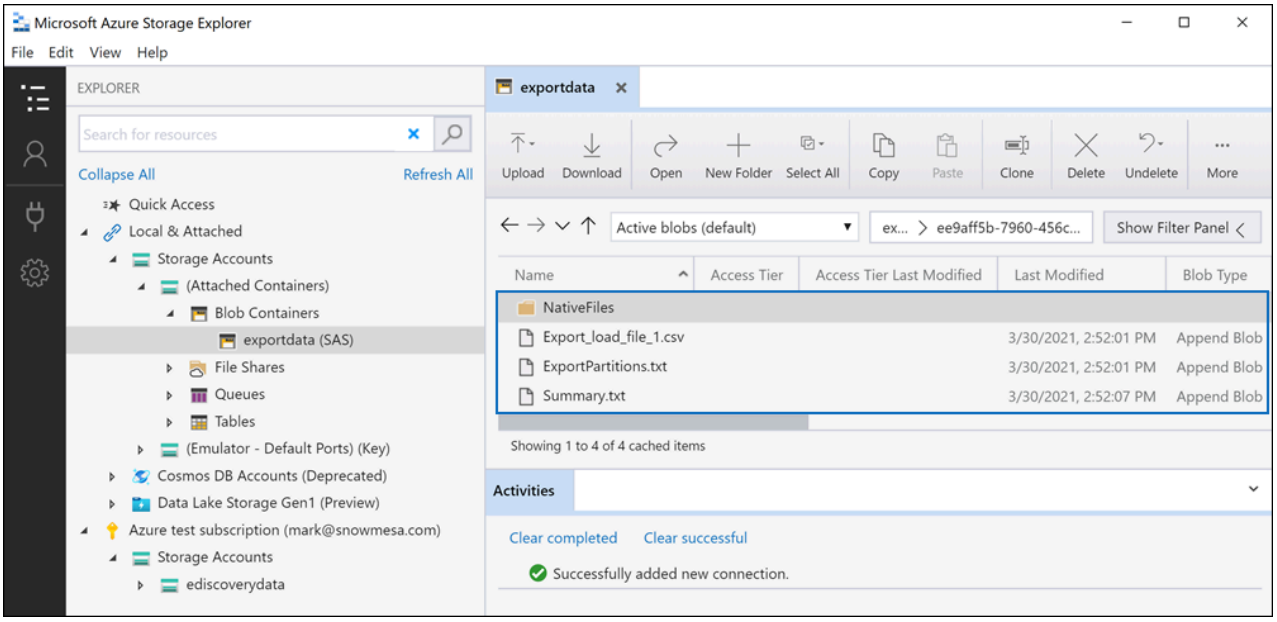
   The **Blob containers** node (under **Storage Accounts** > **(Attached Containers)** > is opened.



It contains a container named with the display name from Step 5. This container contains a folder for each export job that you've downloaded to the container in your Azure Storage account. These folders are named with an ID that corresponds to the ID of the export job. You can find these export IDs (and the name of the export) under **Support information** on the flyout page for each **Preparing data for export** job listed on the **Jobs** tab in the eDiscovery (Premium) case.

7. Double-click the export job folder to open it.

A list of folders and export reports is displayed.



8. To export all contents from the export job, select the **Up** arrow to go back to the export job folder, and then select **Download**.

9. Specify the location where you want to download the exported files, and then select **Select folder**.

   The Azure Storage Explorer starts the download process. The status of the downloading the exported items is displayed in the **Activities** pane. A message is displayed when the download is complete.

> ⓘ **Note**
>
> Instead of downloading the entire export job in Azure Storage Explorer, you can select specific items to download and view.

More information

- The export job folder contains the following items. The actual items in the export folder are determined by the export options configured when the export job was created. For more information about these options, see Export documents from a review set.
  - Export_load_file.csv: This CSV file is a detail export report that contains information about each exported document. The file consists of a column for each metadata property for a document. For a list and description of the metadata that's included in this report, see the **Exported field name** column in the table in Document metadata fields in eDiscovery (Premium).

- Summary.txt: A text file that contains a summary of the export including export statistics.
- Extracted_text_files: This folder contains a text file version of each exported document.
- NativeFiles: This folder contains a native file version of each exported document.
- Error_files: This folder includes the following items when the export job contains any error files:
  - ExtractionError.csv: This CSV file contains the available metadata for files that weren't properly extracted from their parent item.
  - ProcessingError: This folder contains documents with processing errors. This content is at an item level, which means if an attachment had a processing error, the document that contains the attachment will also be included in this folder.

Last updated on 03/24/2026

Manage jobs in eDiscovery (Premium)

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

Jobs report

eDiscovery (Premium) includes a jobs report that lists all jobs that count towards case jobs concurrency and daily limits in eDiscovery for a defined time period. *eDiscovery administrators* and *eDiscovery Managers* (preview) can access this report on the **Reports** > **Jobs report** tab and can use this report to see the job activities across Content Search activities and eDiscovery (Standard and Premium) cases. *eDiscovery Managers* (preview) can view the jobs usage cards at the top of the report, but the jobs list is limited to the cases to which they're assigned.

> **Important**
>
> Jobs listed on the report can't be deleted or removed. Long-running jobs automatically time out after 7 days.

The job report summarizes the following key information for jobs up to the last 30 days:

- All jobs that are in progress or completed (successfully or with error) in a specified time range across eDiscovery Standard and Premium.
- Insights into how far away your organization is from reaching job-related organization-wide limits.
- Creates a .csv report for the list of jobs.
- Filters for job types, statuses, and for hours/days.
- Quick access to case/job pages by selecting case name links.

The following eDiscovery (Premium) job types currently don't count towards the concurrency/daily limits and aren't shown in the report:

- Prepare for export
- Run analytics
- Tag Items
- Compare load sets
- Convert redacted items to PDF
- Relevance train and Relevance load
- Prepare remediation

> **Note**
>
> Any jobs included in the list view of the report and that have run for more than 24 hours do not contribute towards the concurrency/daily limits for your organization.

The following eDiscovery (Premium) jobs started in classic cases don't contribute to throttling limits and aren't reflected in the report. However, these jobs started for cases in the New case format are included in the report:

- Adding non-Microsoft 365 data to a review set
- Adding remediated data to a review set
- Re-indexing custodian data
- Adding data to a review set
- Adding data to another review set

Job types and descriptions

Long-running processes are associated with specific jobs that support cases, collections, and reporting in Microsoft Purview eDiscovery (Premium). These jobs are triggered by user actions when using and managing cases and collections.

Expand table

Job type	Description
Add external data	A user uploads non-Microsoft 365 data to a review set. The data is also indexed during this process. For example, files from an on-premises file server or a client computer are uploaded to a review set. For more information, see Load non-Microsoft 365 data into a review set.

Job type	Description
Add items to another review set	A user adds documents from one review set to a different review set in the same case. For more information, see Add data to a review set from another review set.
Add items to review set	A user adds a collection to a review set. This job consists of two sub jobs: - **Export** - A list of items in the collection is generated. - **Ingestion & Indexing** - The items in the collection that match the search query are copied to an Azure Storage location (in a process called *ingestion*) and then those items in the Azure Storage location are reindexed. This new index is used when querying and analyzing items in the data set. For more information, see Add search results to a review set.
Adding remediated items to a review set	Data with processing errors is remediated and loaded back into a review set. For more information, see: • Error remediation when processing data • Single item error remediation
Apply holds	A user adds custodial and non-custodial sources on the Data sources page and places them on hold. This triggers an Apply holds job. For more information, see, Work with custodians in eDiscovery (Premium).
Compare load sets	A user looks at the differences between different load sets in a review set. A load set is an instance of adding data to a review set. For example, if you add the results of two different searches to the same review set, each would represent a load set.
Converting redacted items to PDF	After a user annotates a document in a review set and redacts a portion of it, they can choose to convert the redacted document to a PDF file. This ensures that the redacted portion won't be visible if the document is exported for presentation. For more information, see View documents in a review set.
Create conversation PDFs	When a user adds the results of a search to a conversation review set, instant message conversations (also called threaded conversations) in services like Microsoft Teams are reconstructed in a PDF file. This job is also triggered when a user selects **Action** > **Create conversation PDFs in a review set**. For more information, see Review conversations in eDiscovery (Premium).
Prepare search estimates	After a user creates and runs or reruns a collection estimate, the search tool searches the index for items that match the search query and prepares an estimate that includes the number and total size of all items by the search, and the number of data sources searched. For more information, see Collect data for a case.
Import custodians	User imports multiple custodians at once by uploading a CSV file that contains the information necessary to add them to a case. For more information, see Import custodians to an eDiscovery (Premium) case.
Preparing for export	A user exports documents from a review set. When the export process is complete, they can download the exported data to a local computer. For more information, see Export

Job type	Description
	case data.
Prepare remediation	When a user selects a file and creates a new error remediation in the Error view on the Processing tab of a case, the first step in the process is to upload the file that has the processing error to an Azure Storage location in the Microsoft cloud. This job tracks the progress of the upload process. For more information about the error remediation workflow, see Error remediation when processing data.
Prepare search preview	After a user creates and runs a new collection estimate (or reruns an existing collection estimate), the search tool prepares a sample subset of items (that match the search query) that can be previewed. Previewing search results help you determine the effectiveness of the search. For more information, see Learn about collections in eDiscovery.
Purge data	Use eDiscovery (Premium) and the Microsoft Graph Explorer to search for and delete chat messages in Microsoft Teams. This feature can help you find and remove sensitive information or inappropriate content. This search and purge workflow will also help you respond to a data spillage incident, when content containing confidential or malicious information is released through Teams chat messages.
Reindex data sources	When you add custodial or non-custodial data sources to a case, all partially indexed items in the selected data source locations are reindexed by a process called *Advanced indexing*. This job is also triggered when you select **Update index** on the **Processing** tab of a case, and when you update the index for a specific data source on the properties flyout page. For more information, see Advanced indexing of custodian data.
Relevance load	This job is triggered when user starts a "new model" - Review set > Manage predictive coding models (preview). It sets up the machine learning model. After you create a model, the following things occur in the background during the creation and preparation of the model: - The system calculates the number of items for the control set. This size is based on the number of items in the review set and the settings for the confidence level and the margin of error. Items for the control set are randomly selected and designated as control set items. The system includes 10 items from the control set in the first round of training. - The system randomly selects 40 items from the review set to be included in the training set for the first round of training. Therefore, the first round of training includes 50 items for labeling: 40 items from the training set and 10 items from the control set.
Run analytics	A user analyzes data in a review set by running eDiscovery (Premium) analytics tools such as near duplicate detection, email threading analysis, and themes analysis. For more information, see Analyze data in a review set.
Tag items	This job is triggered when a user selects **Start tagging job** in the **Tagging panel** when reviewing documents in a review set. A user can start this job after tagging documents in a review set and then bulk-selecting them in the view document panel. For more information, see Tag documents in a review set.

Job status

The following table describes the different status states for jobs:

Expand table

Status	Description
Submitted	A new job was created. The date and time that the job was submitted is displayed in the **Created** column on the **Jobs** tab.
Submission failed	The job submission failed. You should attempt to rerun the action that triggered the job.
In progress	The job is in progress, you can monitor the progress of the job in the **Jobs** tab.
Successful/Completed	The job was successfully completed. The date and time that the job completed is displayed in the **Completed** column on the **Jobs** tab.
Partially successful	The job was successful. This status is typically returned when the job didn't find any partially indexed data (also called *unindexed data*) in some of the custodian data sources.
Failed/Completed with errors	The job failed. You should attempt to rerun the action that triggered the job. If the job fails a second time, we recommend that you contact Microsoft Support and provide the support information from the job.

Last updated on 09/02/2025

Close or delete an eDiscovery (Premium) case

> ⊗ **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic **Content Search**, classic **eDiscovery (Standard)**, and classic **eDiscovery (Premium)**.
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the **new eDiscovery experience** in the **Microsoft Purview portal**.

When the legal case or investigation supported by a Microsoft Purview eDiscovery (Premium) case is completed, you can close or delete a case. You can also reopen a closed case.

Close a case

Here's what happens when you close an eDiscovery (Premium) case:

- If the case contains any content locations on hold, these holds are turned off. After the hold is turned off, a 30-day grace period (called a *delay hold*) is applied to content locations that were on hold. This helps prevent content from being immediately deleted and gives admins an opportunity to search for or recover content that will be permanently deleted after the delay hold period expires. For more information, see Removing content locations from an eDiscovery hold.
- Closing a case only turns off the holds that are associated with that case. If other holds are place on a content location (such as a Litigation Hold, Microsoft Purview eDiscovery (Standard) hold, or a hold from a different eDiscovery (Premium) case) those holds will still be maintained.
- The case is still listed on the eDiscovery page in the Microsoft Purview portal. The details, holds, searches, and members of a closed case are retained.
- You can edit a case after it's closed. For example, you can add or removing members, create searches, export search results, and prepare search results for analysis in eDiscovery (Premium). The primary difference between active and closed cases is that holds are turned off when a case is closed.

> **Important**
>
> If a case has more than 100 hold policies, you must manually remove the policies using PowerShell. For more information, see Remove-CaseHoldPolicy.

To close a case:

1. On the **eDiscovery (Premium)** page, select the case that you want to close.
2. On the **Settings** tab, under **Case Information**, choose **Select**.
3. At the bottom of the **Case Information** flyout page, select **Actions**, and then select **Close case**. It may take up to 60 minutes for the closing process to complete.

When closing a case, you'll see the status of the case changing as the case is processed for closing. The following table describes the case status messages that you'll see when closing a case:

Expand table

Status	Description
Active	The case is active.
Closed	The case is closed without errors and the case can be reopened.
Closed with Errors	The case is closed, but there are errors. Errors are most likely associated with holds that were not released during closing. To resolve, check policies with errors, resolve the errors, and redistribute the policies.
Closing	The case is being closed. This status is usually seen only for case with active holds. Cases without holds are immediately closed.
Pending delete	The case likely has a hold in the pending deletion state. Check the hold policy and redistribute to resolve the error. Afterwards, cases without any holds should be deleted successfully.

Reopen a closed case

When you reopen an eDiscovery (Premium) case, any holds that were in place when the case was closed won't be automatically reinstated. After the case is reopened, you'll have to go to the **Holds** tab and turn on the previous holds. To turn on a hold, select it to display the flyout page, and then set the **Status** toggle to **On**.

To reopen a closed case:

1. On the **eDiscovery (Premium)** page, select the case that you want to reopen.
2. On the **Settings** tab, under **Case Information**, select **Select**.
3. At the bottom of the **Case Information** flyout page, select **Actions**, and then select **Reopen case**.

It might take up to 60 minutes for the reopening process to complete.

Delete a case

You can delete both active and closed eDiscovery (Premium) cases. When you delete a case, all components associated with the case, such as the list of custodians, communications, searches, review sets, and export job are deleted. The case is removed from the list of cases on the **eDiscovery (Premium)** page in the Microsoft Purview portal. You can't recover or reopen a deleted case.

> **Note**
>
> In data spillage scenarios, the only way to remove items in a review set is to delete the eDiscovery (Premium) case. Other "search and purge" methods don't remove items from a review set.

Before you can delete a case, you must first delete *all* holds listed on the holds page of the case. That includes deleting holds with a status of **Off**. Default hold policies can only be deleted when the hold is turned off. You must close an active case to turn off any default hold policies in the case. Once the holds are turned off for default hold policies, they can be deleted.

To delete holds associated with a case:

1. Go the **Holds** tab in the eDiscovery (Premium) case that you want to delete.
2. Select the hold that you want to delete.
3. On the flyout page, select **Delete**. Ensure that the hold is turned off.

To delete a case:

1. On the **eDiscovery (Premium)** page, select the case that you want to delete.
2. On the **Settings** tab, under **Case Information**, choose **Select**.
3. At the bottom of the **Case Information** flyout page, select **Actions**, and then select **Delete case**.

Last updated on 09/02/2025

Add or remove members from an eDiscovery (Premium) case

> **⊗ Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

You can add or remove members to manage who can access the case. However, before a member can access an eDiscovery (Premium) case (and perform tasks in the case), you must add the user to the eDiscovery Manager role group on the **Permissions** page in the Microsoft Purview portal. For more information, see Assign eDiscovery permissions.

1. On the **eDiscovery (Premium)** page, go to the case that you want to add a member to.
2. Select the **Settings** tab and then choose **Select** in the **Access & permissions** tile.
3. Under **Manage members**, select **Add** to add members to the case. You can also choose to add a role group to the case by selecting **Add** under **Manage role groups**.
4. In the list of people or role groups that can be added as members of the case, select the check box next to the names of the people or role groups that you want to add.

   > **ⓘ Note**
   >
   > When adding a role group to a case, you can only add the role groups that you are a member of.

5. After you've selected the people or role groups to add as members of the case, select **Add**.
6. In the **Manage this case** flyout page, select **Save** to save the new list of case members.

> **Important**
>
> If a role is added or removed from a role group that you've added as a member of a case, then the role group will be automatically removed as a member of the case (or any case the role group is a member of). The reason for this is to protect your organization from inadvertently providing additional permissions to members of a case. Similarly, if a role group is deleted, it will be removed from all cases it was a member of. For more information, see Assign eDiscovery permissions.

Removing members from a case

Only an eDiscovery Administrator can remove members from a case. Even if you're assigned to the eDiscovery Manager role group or initially created the case, you won't be able to remove yourself or other members from a case unless you're also an eDiscovery Administrator. To remove yourself or other members from a case, contact an eDiscovery Administrator in your organization.

Last updated on 11/21/2025

Configure search and analytics settings for eDiscovery (Premium) cases

> **⊗ Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

You can configure settings for each Microsoft Purview eDiscovery (Premium) case to control the following functionality:

- Near duplicates and email threading
- Themes
- Autogenerated review set query
- Ignore text
- Optical character recognition

Configure analytics settings for a case

To configure search and analytics settings for a case:

1. On the **eDiscovery (Premium)** page, select the case.
2. On the **Settings** tab, under **Search & analytics**, choose **Select**. The case settings page is displayed. These settings are applied to all review sets in a case.

The following sections in this article describe the analytics settings that you can configure for a case.

Near duplicates and email threading

In this section, you can set parameters for duplicate detection, near duplicate detection, and email threading. For more information, see Near duplicate detection and Email threading.

- **Near duplicates/email threading:** When turned on, duplicate detection, near duplicate detection, and email threading are included as part of the workflow when you run analytics on the data in a review set.
- **Document and email similarity threshold:** If the similarity level for two documents is over the threshold, both documents are put in the same near duplicate set.
- **Minimum/maximum number of words:** These settings specify that near duplicates and email threading analysis are performed only on documents that have at least the minimum number of words and at most the maximum number of words.

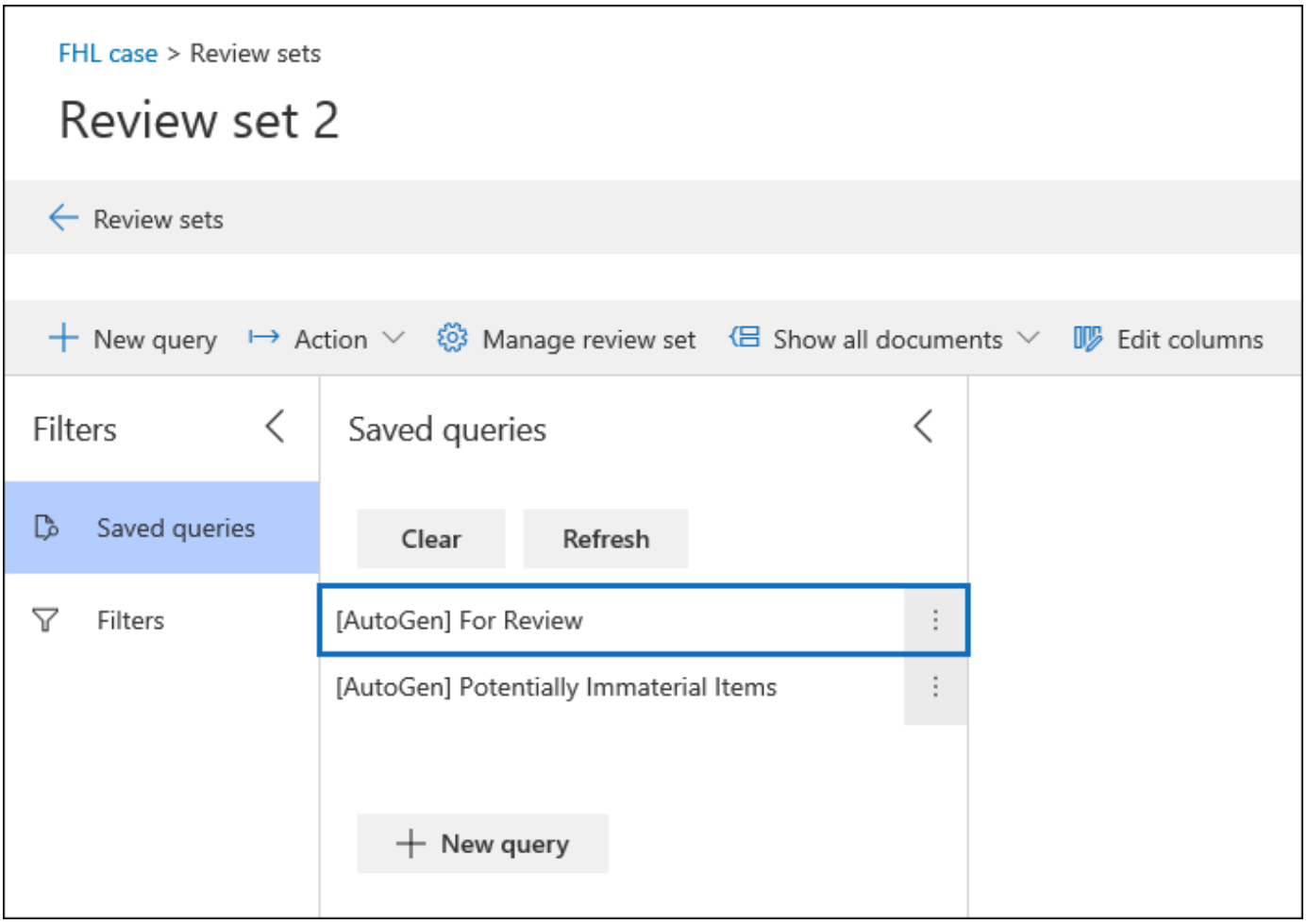
Themes

In this section, you can set parameters for themes. For more information, see Themes.

- **Themes:** When turned on, themes clustering is performed as part of the workflow when you run analytics on the data in a review set.
- **Maximum number of themes:** Specifies the maximum number of themes that can be generated when you run analytics on the data in a review set.
- **Include numbers in themes:** When turned on, numbers (that identifies a theme) are included when generating themes.
- **Adjust maximum number of themes dynamically:** In certain situations, there may not be enough documents in a review set to produce the desired number of themes. When this setting is enabled, eDiscovery (Premium) adjusts the maximum number of themes dynamically rather than attempting to enforce the maximum number of themes.

Review set query

If you select the **Automatically create a For Review saved search after analytics** checkbox, eDiscovery (Premium) autogenerates review set query named **For Review.**



This query basically filters out duplicate items from the review set. This lets you review the unique items in the review set. This query is created only when you run analytics for a review set in the case. For more information, about review set queries, see Query the data in a review set.

Ignore text

There are situations where certain text will diminish the quality of analytics, such as lengthy disclaimers that get added to email messages regardless of the content of the email. If you know of text that should be ignored, you can exclude it from analytics by specifying the text string and the analytics functionality (Near-duplicates, Email threading, Themes, and Relevance) that the text should be excluded for. Using regular expressions (RegEx) as ignored text is also supported.

Optical character recognition (OCR)

When this setting is turned on, OCR processing will be run on image files. OCR processing is run in the following situations:

- When custodians and non-custodial data sources are added to a case. When OCR is applied to image files, the text in those files will be searchable during a collection. OCR processing is performed during the Advanced indexing process. OCR is only run on items that are

processed during Advanced indexing. For example, if a large PDF file that is partially indexed or had other indexing errors is processed during Advanced indexing, the file will also have OCR applied. In other words, OCR processing only occurs on files that are reindexed during the Advanced indexing process.

This means there may be situations where custodians are added to a case but some email attachments won't be processed for OCR because those files aren't processed during Advanced indexing.

- When content from other data sources(that aren't associated with a custodian and added to the case in a non-custodial data source) is added to a review set.
- KeyQL search (collection estimates) don't run OCR on an item unless an item had indexing errors and these items must be Advanced indexed.

After data is added to a review set, image text can be reviewed, searched, tagged, and analyzed. You can view the extracted text in the Text viewer of the selected image file in the review set. For more information, see:

- Advanced indexing of custodian data
- Add search results to a review set
- Supported image file types

Last updated on 11/21/2025

Limits in eDiscovery (Premium)

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

This article describes the limits in the Microsoft Purview eDiscovery (Premium) solution in Microsoft Purview.

Case and review set limits

The following table lists the limits for cases and review sets in eDiscovery (Premium).

Expand table

Description of limit	Classic Case Limit	New Case Limit
Total number of documents that can be added to a case (for all review sets in a case).	3 million	40 million
Total file size per load set. This includes loading non-Office 365 into a review set.	300 GB	1 TB
Maximum number of custodians per case	Not applicable	1000
Maximum number of load sets per case.	200	200
Maximum number of review sets per case.	20	20
Maximum number of tag groups per case.	1,000	1,000
Maximum number of unique tags per case.	1,000[1]	1,000[1]

Collection limits

Expand table

Description of limit	Limit
Maximum number of locations for a single collection. This limit includes the combined total of user mailboxes, and the mailboxes associated with Microsoft 365 Groups, Microsoft Teams, and Viva Engage Groups.	1000

The limits for the **Export as report** and the export results are as follows:

Expand table

Description of limit	Limit
Total items	1 million
Total locations	500,000
Exported results (items)	500,000
Exported results (locations)	100,000

Export limits

The limits described in this section are related to exporting documents out of a review set.

Expand table

Description of limit	Limit
Maximum size of a single export.	5 million documents or 500 GB, whichever is smaller

Hold limits

The following table lists the limits for holds associated with an eDiscovery (Premium) case.

Expand table

Description of limit	Limit
Maximum number of hold policies for an organization. This limit includes the combined total of hold policies in Microsoft Purview eDiscovery (Standard) and Microsoft Purview eDiscovery (Premium) cases.	10,000[2]
Maximum number of mailboxes in a single case hold. This limit includes the combined total of user mailboxes, and the mailboxes associated with Microsoft 365 Groups, Microsoft Teams, and	2,000

Description of limit	Limit
Viva Engage Groups.	
Maximum number of sites in a single case hold. This limit includes the combined total of OneDrive sites, SharePoint sites, and the sites associated with Microsoft 365 Groups, Microsoft Teams, and Viva Engage Groups.	2,000

> **Note**
>
> If your license is downgraded or if the **Case** setting is changed to *Standard*, existing holds aren't removed but the policy is now subject to the *Standard* case limits (1,000 locations). For example, if a *Premium* case has a hold policy with 1,500 mailboxes and 1,500 sites and the case setting is changed to *Standard*, you can't add additional locations to the hold policy because it's now based on the *Standard* limits. However, you can remove locations from the hold policy.

Indexing limits

The following table lists the indexing limits in eDiscovery (Premium).

Expand table

Description of limit	Limit
Maximum number of characters extracted from a single file.	10 million[3]
Maximum size of a single file.	150 MB[3]
Maximum depth of embedded items in a document.	25[3]
Maximum size of files processed by Optical Character Recognition (OCR).	24 MB[3]
Maximum Advanced indexing throughput	2 GB per hour

Jobs limits

Expand table

Description of limit	Limit
Maximum number of concurrent jobs in your organization.	100
Maximum number of concurrent jobs that a single user can start at one time.	50

Description of limit	Limit
Maximum number of concurrent tenant-wide jobs (for example, tenant-wide searches) in your organization.	50
Maximum number of concurrent tenant-wide jobs (for example, tenant-wide searches) that a single user can start at one time.	25

Review set download limits

Expand table

Description of limit	Limit
Total file size or maximum number of documents downloaded from a review set.	3 MB or 50 documents[6]

Review set viewer limits

Expand table

Description of limit	Limit
Maximum number of items displayed per page in a review set.	1,000

> **Note**
>
> Use default or custom filters to adjust the displayed items in a review set as needed.

Search limits

The limits described in this section are related to using the search tool on the **Searches** tab to collect data for a case. For more information, see Collect data for a case in eDiscovery (Premium).

Expand table

Description of limit	Limit
Maximum number of mailboxes or sites that can be searched in a single search.	No limit
Maximum number of searches that can run at the same time.	No limit
Maximum number of characters for a search query (including operators and conditions).	10,000[4]

Description of limit	Limit
Maximum number of characters for a search query for SharePoint and OneDrive for Business sites (including operators and conditions).	10,000 4,000 with Wildcards[4]
Minimum number of alpha characters for prefix wildcards; for example, **one*** or **set***.	3
Maximum variants returned when using prefix wildcard to search for an exact phrase or when using a prefix wildcard and the **NEAR** Boolean operator.	10,000[5]
Maximum number of items per user mailbox that are displayed on preview page for searches. The newest items are displayed.	100
Maximum number of items from all mailboxes displayed on preview page for searches.	1,000
Maximum number of mailboxes that can be previewed for search results. If there are more than 1,000 mailboxes that contain items that match the search query, only the top 1,000 mailboxes with the most results are available for preview.	1,000
Maximum number of items from SharePoint and OneDrive for Business sites displayed on preview page for searches. The newest items are displayed.	200
Maximum number of SharePoint and OneDrive for Business sites that can be previewed for search results. If there are more than 200 sites that contain items that match the search query, only the top 200 sites with the most results are available for preview.	200
Maximum number of items per public folder mailbox displayed on preview page for searches.	100
Maximum number of items found in all public folder mailbox items displayed on preview page for searches.	200
Maximum number of public folder mailboxes that can be previewed for search results. If there are more than 500 public folder mailboxes that contain items that match the search query, only the top 500 mailboxes with the most results are available for preview.	500
The maximum size of an item that can be viewed on the sample page of a collection estimate.	10,000,000 bytes (approximately 9.5 MB)

Search times

Microsoft collects performance information for searches run by all organizations. While the complexity of the search query can impact search times, the biggest factor that affects how long searches take is the number of mailboxes searched. Although Microsoft doesn't provide a Service

Level Agreement for search times, the following table lists average search times for collection searches based on the number of mailboxes included in the search.

Expand table

Number of mailboxes	Average search time
100	30 seconds
1,000	45 seconds
10,000	4 minutes
25,000	10 minutes
50,000	20 minutes
100,000	25 minutes

Themes limits

The limits described in this section are related to manually defining the maximum number of themes in settings for a case.

Expand table

Description of limit	Limit
Minimum number of the maximum number of defined themes.	3
Maximum number of defined themes.	10,000

Viewer limits

Expand table

Description of limit	Limit
Maximum size of Excel file that can be viewed in the native viewer.	4 MB

Need help with eDiscovery issues?

Check out the Resolve common eDiscovery issues article for basic troubleshooting steps that you can take to identify and resolve issues that you might encounter during an eDiscovery search or

elsewhere in the eDiscovery process.

Reference notes

[1] This is the maximum number of tags that you can create in a case. This limit isn't related to the number of documents that can be tagged.

[2] When you put more than 1,000 mailboxes or 100 sites on hold in a single hold policy, the system will automatically scale the hold as needed. This means the system will automatically add data locations to multiple hold policies, instead of adding them to a single hold policy. However, the limit of 10,000 case hold policies per organization still applies.

[3] Any item that exceeds a single file limit shows up as a processing error.

[4] When searching SharePoint and OneDrive for Business locations, the characters in the URLs of the sites being searched count against this limit. The total number of characters consists of:

- All characters in both the Users and Filters fields.
  - All search permissions filters that apply to the user.
  - The characters from any location properties in the search, including ExchangeLocation, PublicFolderLocation, SharPointLocation, ExchangeLocationExclusion, PublicFolderLocationExclusion, SharePointLocationExclusion, and OneDriveLocationExclusion. For example, including all SharePoint sites and OneDrive accounts in the search counts as six characters, as the word "ALL" will appear for both the SharePointLocation and OneDriveLocation field.

[5] For non-phrase queries (a keyword value that doesn't use double quotation marks) we use a special prefix index. This tells us that a word occurs in a document, but not where it occurs in the document. To do a phrase query (a keyword value with double quotation marks), we need to compare the position within the document for the words in the phrase. This means that we can't use the prefix index for phrase queries. In this case, we internally expand the query with all possible words that the prefix expands to; for example, **time*** can expand to "**time OR timer OR times OR timex OR time boxed OR ...**". The limit of 10,000 is the maximum number of variants the word can expand to, not the number of documents matching the query. There's no upper limit for non-phrase terms.

[6] This limit applies to downloading selected documents from a review set. It doesn't apply to exporting documents from a review set. For more information about downloading and exporting documents, see Export case data in eDiscovery (Premium).

Last updated on 09/02/2025

Supported file types in eDiscovery (Premium)

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

Microsoft Purview eDiscovery (Premium) supports many file types at many different levels. The following tables in this article list the currently supported files types.

These tables indicate if a file type is supported for text extraction (and Optical Character Recognition or OCR text extraction for image files), viewable in the native viewer, and also supported in the Annotate viewer in eDiscovery (Premium).

Archive / Container

Expand table

Mime type	File identification	Metadata extraction	Container extraction	Possible Extensions
application/x-7z-compressed	Yes	Yes	Yes	.7z
application/x-rar-compressed	Yes	Yes	Yes	.rar[1]
application/x-tar	Yes	Yes	Yes	.tar
application/zip	Yes	Yes	Yes	.zip

> **Note**
>
> [1] Version 5 .rar files aren't currently supported.

Audio / Video

Expand table

Mime type	File identification	Metadata extraction	Text extraction	Native viewer	Annotate viewer	Possible Extensions
application/mp4	Yes	Yes	No	Yes	No	.f4v; .m4a; .m4v; .mp4; .mp4v; .mpeg; .mpeg4
audio/mpeg	Yes	Yes	No	Yes	No	.mpeg
video/3gpp	Yes	Yes	No	Yes	No	.3gp
video/3gpp2	Yes	Yes	No	Yes	No	.3g2; .3gp2
video/quicktime	Yes	Yes	No	Yes	No	.moov; .mov; .qt
video/x-m4v	Yes	Yes	No	Yes	No	.m4v

Database

Expand table

Mime type	File identification	Metadata extraction	Text extraction	Native viewer	Annotate viewer	Possible Extensions
application/x-msaccess	Yes	Yes	Yes	No	No	.mdb

Email

Expand table

Mime type	File identification	Metadata extraction	Text extraction	Native viewer	Annotate viewer	Possible Extensions
application/vnd.ms-outlook	Yes	Yes	Yes	Yes	Yes	.msg

Mime type	File identification	Metadata extraction	Text extraction	Native viewer	Annotate viewer	Possible Extensions
message/rfc822	Yes	Yes	Yes	Yes	Yes	.eml
text/vcard-contact	Yes	Yes	Yes	Yes	Yes	.vcf

Email Container

Expand table

Mime type	File identification	Metadata extraction	Container extraction	Possible Extensions
application/mbox	Yes	Yes	Yes	.mbox
application/vnd.ms-outlook-pst	Yes	Yes	Yes	.pst

HTML

Expand table

Mime type	File identification	Metadata extraction	Text extraction	Native viewer	Annotate viewer	Possible Extensions
application/xhtml+xml	Yes	Yes	Yes	Yes	Yes	.xhtml
application/xml	Yes	Yes	Yes	Yes	Yes	.xml
text/html	Yes	Yes	Yes	Yes	Yes	.htm; .html; .shtml

Image

Expand table

Mime type	File identification	Metadata extraction	OCR text extraction	Native viewer	Annotate viewer	Possible Extensions
image/bmp	Yes	Yes	Yes	Yes	Yes	.bmp
image/emf	Yes	Yes	Yes	Yes	Yes	.emf
image/gif	Yes	Yes	Yes	Yes	Yes	.gif
image/jpeg	Yes	Yes	Yes	Yes	Yes	.jpeg; .jpg
image/png	Yes	Yes	Yes	Yes	Yes	.png
image/svg+xml	Yes	Yes	Yes	Yes	No	.svg
image/tiff	Yes	Yes	Yes	Yes	Yes	.tif
image/vnd.dwg	Yes	Yes	Yes	Yes	Yes	.dwg; .dxf
image/wmf	Yes	Yes	Yes	Yes	Yes	.wmf

Microsoft Excel

Expand table

Mime type	File identification	Metadata extraction	Text extraction	Native viewer	Annotate viewer	Possible Extensions
application/vnd.ms-excel	Yes	Yes	Yes	Yes	Yes	.dat; .xls
application/vnd.ms-excel.sheet.binary.macroenabled.12	Yes	Yes	Yes	Yes	No	.xlsb
application/vnd.ms-excel.sheet.macroenabled.12	Yes	Yes	Yes	Yes	Yes	.xlsm
application/vnd.ms-excel.template.macroenabled.12	Yes	Yes	Yes	No	No	.xltm
application/vnd.openxmlformats-officedocument.spreadsheetml.sheet	Yes	Yes	Yes	Yes	Yes	.xlsx
application/vnd.openxmlformats-officedocument.spreadsheetml.template	Yes	Yes	Yes	Yes	Yes	.xltx

Microsoft OneNote

Expand table

Mime type	File identification	Metadata extraction	Text extraction	Native viewer	Annotate viewer	Possible Extensions
application/onenote	Yes	Yes	Yes	No	No	.one

Microsoft PowerPoint

Expand table

Mime type	File identification	Metadata extraction	Text extraction	Native viewer	Annotate viewer	Possi Extens
application/vnd.ms-powerpoint	Yes	Yes	Yes	Yes	Yes	.pot; . .pp
application/vnd.openxmlformats-officedocument.presentationml.presentation	Yes	Yes	Yes	Yes	Yes	.ppt
application/vnd.openxmlformats-officedocument.presentationml.slideshow	Yes	Yes	Yes	Yes	Yes	.pps
application/vnd.openxmlformats-officedocument.presentationml.template	Yes	Yes	Yes	Yes	Yes	.pot

Microsoft Project

Expand table

Mime type	File identification	Metadata extraction	Text extraction	Native viewer	Annotate viewer	Possible Extensions
application/vnd.ms-project	Yes	Yes	Yes	No	Yes	.mpp

Microsoft Publisher

Expand table

Mime type	File identification	Metadata extraction	Text extraction	Native viewer	Annotate viewer	Possible Extensions
application/x-mspublisher	Yes	Yes	Yes	Yes	Yes	.pub

Microsoft Visio

Expand table

Mime type	File identification	Metadata extraction	Text extraction	Native viewer	Annotate viewer	Possible Extensions
application/vnd.ms-visio.drawing	Yes	Yes	Yes	Yes	No	
application/vnd.visio	Yes	Yes	Yes	Yes	Yes	.vsd

Microsoft Word

Expand table

Mime type	File identification	Metadata extraction	Text extraction	Native viewer	Annotate viewer	Poss Exten
application/msword	Yes	Yes	Yes	Yes	Yes	.dat;
application/rtf	Yes	Yes	Yes	Yes	Yes	.doc
application/vnd.ms-word.document.macroenabled.12	Yes	Yes	Yes	Yes	Yes	.do
application/vnd.ms-word.template.macroenabled.12	Yes	Yes	Yes	Yes	Yes	.do
application/vnd.openxmlformats-officedocument.wordprocessingml.document	Yes	Yes	Yes	Yes	Yes	.dc
application/vnd.openxmlformats-officedocument.wordprocessingml.template	Yes	Yes	Yes	Yes	Yes	.dc

Microsoft Works

Expand table

Mime type	File identification	Metadata extraction	Text extraction	Native viewer	Annotate viewer	Possible Extensions
application/vnd.ms-works-ss	Yes	Yes	No	No	No	.wps
application/vnd.ms-works-wp	Yes	Yes	No	No	No	.wps

Open Document Format

Mime type	File identification	Metadata extraction	Text extraction	Native viewer	Annotate viewer	Possible Extension
application/vnd.oasis.opendocument.text	Yes	Yes	Yes	Yes	Yes	.odt

Other

Mime type	File identification	Metadata extraction	Text extraction	Native viewer	Annotate viewer	Possible Extensions
application/json	Yes	Yes	Yes	Yes	Yes	n/a
application/octet-stream	Yes	No	No	No	No	.fluid
application/vnd.ms-graph	Yes	Yes	No	No	No	
application/winhlp	Yes	Yes	No	No	No	.hlp
application/x-tnef	Yes	Yes	No	No	No	

Plain Text

Mime type	File identification	Metadata extraction	Text extraction	Native viewer	Annotate viewer	Possible Extensions
text/csv	Yes	Yes	Yes	Yes	Yes	.csv
text/plain	Yes	Yes	Yes	Yes	Yes	.con; .css; .csv; .dat; .pl; .txt

Portable Document Format

Mime type	File identification	Metadata extraction	Text extraction	Native viewer	Annotate viewer	Possible Extensions
application/pdf	Yes	Yes	Yes	Yes	Yes	.pdf

Word Perfect

Mime type	File identification	Metadata extraction	Text extraction	Native viewer	Annotate viewer	Possible Extensions
application/vnd.wordperfect; version=5.0	Yes	Yes	Yes	No	No	.wpd
application/vnd.wordperfect; version=5.1	Yes	Yes	Yes	No	No	.wpd
application/vnd.wordperfect; version=6.x	Yes	Yes	Yes	No	No	.wpd

Word Pro

Mime type	File identification	Metadata extraction	Text extraction	Native viewer	Annotate viewer	Possible Extensions
application/vnd.lotus-wordpro	Yes	Yes	No	No	No	.lwp

Last updated on 09/02/2025

Document metadata fields in eDiscovery (Premium)

> **⊗ Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

The following table lists the metadata fields for documents in a review set in a case in Microsoft Purview eDiscovery (Premium). For more information about searchable properties when searching Microsoft 365 content locations when you're collecting data for an eDiscovery (Premium) case, see Keyword queries and search conditions for Content Search.

This table provides the following information:

- **Field name** and **Display field name:** The name of the metadata field and the name of the field that's displayed when viewing the file metadata of a selected document in a review set. Some metadata fields aren't included when viewing the file metadata of a document. These fields are highlighted with an asterisk (*).
- **Searchable field name:** The name of the property that you can search for when running a review set query.
- **Exported field name:** The name of the metadata field that included when documents are exported.
- **Description:** A description of the metadata field.

> **ⓘ Note**
>
> The **Keywords** field in review set search uses Keyword Query Language (KeyQL). The fields listed in the **Searchable field name** column can be used in the **Keywords** field in a review set search to form complex queries without you having to use the query builder. For more information about KeyQL, see Keyword Query Language syntax reference.

Expand table

Field name and Display field name	Searchable field name	Exported field name	Description
Attachment Content ID	AttachmentContentId	*Not exported*	Attachment content ID the item.
Attorney client privilege score	AttorneyClientPrivilegeScore	*Not exported*	Attorney-client privileg model content score.
Author	Author	Doc_authors	Author from the docun metadata.
BCC	Bcc	Email_bcc	Bcc field for message t The format is *DisplayN* *<SMTPAddress>*.
CC	Cc	Email_cc	Cc field for message ty The format is *DisplayN* *<SMTPAddress>*.
Channel Name	Channel	ChannelName	This field is the Teams channel name. Only ap to Microsoft Teams cor
Compliance labels	ComplianceLabels	Compliance_labels	Retention labels applie content in Office 365.
Compound Path	CompoundPath	Compound_path	Human readable path t describes the source of item.
Content*	Content	*Not exported*	Extracted text of the ite
Conversation Body	ConversationBody	*Not exported*	Conversation body of t item.
Conversation ID	ConversationId	Conversation_ID	Conversation ID from t message. For Teams 1:1 group chats, all transcr files and their family ite within the same conversation share the Conversation ID. For m information, see eDisc (Premium) workflow fo content in Microsoft Te
Conversation Family ID	ConversationFamilyID	ConversationFamilyID	The ID that identifies individual elements of conversation and the re items in the conversati
Conversation Index	*Not searchable*	Conversation_index	Conversation index fro message.
Conversation Name	*Not searchable*	ConversationName	This field depends on content type. **Teams 1:1 chat:** first 40

Field name and Display field name	Searchable field name	Exported field name	Description
			characters of first mess **Teams 1:N chat:** Name group chat; if not availa the first 40 characters c first message. **Teams Channel Post:** P title or announcement subhead; if not availabl first 40 characters of th message.
Conversation Pdf Time	ConversationPdfTime	*Not exported*	Date when the PDF ver of the conversation wa created.
Conversation Redaction Burn Time	ConversationRedactionBurnTime	*Not exported*	Date when the PDF ver of the conversation wa created for Chat.
Conversation Topic	ConversationTopic	*Not exported*	Conversation topic of t item.
Conversation Type	ConversationType	ConversationType	The type of chat conversation. Values ar **Teams 1:1 and group cl** **and all Viva Engage** **conversations:** Group **Teams channels and pr** **channels:** Channel
Contains Deleted Message	ContainsDeletedMessage	ContainsDeletedMessage	Indicates if the chat transcript includes a de message
Contains Edited Message	ContainsEditedMessage	ContainsEditedMessage	Indicates if the chat transcript includes an e message
Teams Announcement Title	TeamsAnnouncementTitle	TeamsAnnouncementTitle	Title from a teams announcement.
N/A	*N/A*	Converted_file_path	The path of the conver export file. For internal Microsoft use only.
Custodian	Custodian	Custodian	Name of the custodian item was associated wi
Date	Date	Date	Date is a computed fiel that depends on the fil type. **Email**: Sent date **Email attachments**: Las modified date of the document; if not availa

Field name and Display field name	Searchable field name	Exported field name	Description
			the parent's sent date **Embedded documents** modified date of the document; if not availa the parent's last modifi date **SPO documents (inclu** **modern attachments)**: modified date of the document; if not availa SharePoint last modifie date **Non-Office 365 docun** Last modified date **Meetings**: Meeting sta date **VoiceMail**: Sent date **IM**: Sent date **Teams**: Sent date
Document comments	DocComments	Doc_comments	Comments from the document metadata.
Document company	*Not searchable*	Doc_company	Company from the document metadata.
Document date created	CreatedTime	Doc_date_created	Create date from docu metadata.
DocIndex*	*Not searchable*	*Not exported*	The index in the family *0* means it's the root.
Document keywords	*Not searchable*	Doc_keywords	Keywords from the document metadata.
Document modified by	*Not searchable*	Doc_modified_by	The user who last mod the document from document metadata.
Document revision	Doc_Version	Doc_Version	Revision from the docu metadata.
Document subject	*Not searchable*	Doc_subject	Subject from the docur metadata.
Document template	*Not searchable*	Doc_template	Template from the document metadata.
DocLastSavedBy	*Not searchable*	Doc_last_saved_by	The name of the user w last saved the documer
Dominant theme	DominantTheme	Dominant_theme	Dominant theme as calculated for analytics
Duplicate subset	*Not searchable*	Duplicate_subset	Group ID for exact

Field name and Display field name	Searchable field name	Exported field name	Description
			duplicates.
EmailAction*	*Not searchable*	Email_action	Values are *None*, *Reply*, *Forward*; based on the subject line of a messa
Email Delivery Receipt Requested	*Not searchable*	Email_delivery_receipt	Email address supplied Internet Headers for de receipt.
Importance	EmailImportance	Email_importance	Importance of the mes *0* - Low; *1* - Normal; *2* -
Ignored processing errors	ErrorIgnored	Error_Ignored	Error was ignored and remediated.
EmailInternetHeaders	EmailInternetHeaders	Email_internet_headers	The full set of email he from the email messag
EmailLevel*	*Not searchable*	Email_level	Indicates a message's l within the email thread belongs to; attachment inherit its parent messa value.
Email Message ID	*Not searchable*	Email_message_ID	Internet message ID fro the message.
EmailReadReceiptRequested	*Not searchable*	Email_read_receipt	Email address supplied Internet Headers for re receipt.
Email Security	EmailSecurity	Email_security	Security setting of the message: *0* - None; *1* - Signed; *2* - Encrypted; . Encrypted and signed.
Email Sensitivity	EmailSensitivity	email_sensitivity	Sensitivity setting of th message: *0* - None; *1* Personal; *2* - Private; *3* CompanyConfidential.
Email set	EmailSet	Email_set	Group ID for all messag the same email set.
EmailThread*	*Not searchable*	Email_thread	Position of the messag within the email set; co of node IDs from the ro the current message ar separated by periods (.
N/A	*N/A*	Export_native_path	The path of the exporte file.
Extracted content type	*Not searchable*	Native_type	Extracted content type, the form of mime type;

Field name and Display field name	Searchable field name	Exported field name	Description
			example, *image/jpeg*
N/A	*N/A*	Extracted_text_path	The path to the extract text file in the export.
ExtractedTextLength*	*Not searchable*	Extracted_text_length	Number of characters i extracted text.
FamilyDuplicateSet*	*Not searchable*	Family_duplicate_set	Numeric identifier for families that are exact duplicates of each othe (same content and all t same attachments).
Family ID	FamilyId	Family_ID	Groups together attachments and extrac items from email and c with its parent item. Th includes the chat or em and all attachments an extracted items.
Family Size	*Not searchable*	Family_size	Number of documents the family.
File class	FileClass	File_class	For content from Share and OneDrive: *Docume* For content from Excha *Email* or *Attachment*. For content from Team Viva Engage: *Conversat*
File ID	FileId	File_ID	Document identifier un within the case.
File system date created	*Not searchable*	File_system_date_created	Created date from file system (only applies to Office 365 data).
File system date modified	*Not searchable*	File_system_date_modified	Modified date from file system (only applies to Office 365 data).
File Type	FileType	*Not exported*	File type of the item ba on file extension.
Group ID	GroupId	Group_ID	Groups together all ite email and documents. email, this includes the message and all attach and extracted items. Fo documents, this includ document and any embedded items.

Field name and Display field name	Searchable field name	Exported field name	Description
Has attachment	EmailHasAttachment	Email_has_attachment	Indicates whether or nc message has attachme
Has attorney	HasAttorney	*Not exported*	*True* when at least one the participants is foun the attorney list; otherv the value is *False*.
HasText*	*Not searchable*	Has_text	Indicates whether or nc item has text; possible values are *True* and *Fal*
Immutable ID	*Not searchable*	Immutable_ID	This ID is used to uniqu identify a document wi review set. This field ca used in a review set sea and the ID can't be use access a document in i native location.
Inclusive type	InclusiveType	Inclusive_type	Inclusive type calculate analytics: *0* - not inclus - inclusive; *2* - inclusive minus; *3* - inclusive cop
In Reply To ID	*Not searchable*	In_reply_to_ID	In reply to ID from the message.
InputFileExtension	*Not searchable*	Original_file_extension	The original file extensi the file.
InputFileID	*Not searchable*	Input_file_ID	The file ID of the top le item in the review set. I attachment, this ID will the ID of the parent. Th be used to group famil together.
Is modern attachment	IsModernAttachment	*Not exported*	This file is a modern attachment or linked fi
Is from document version	IsFromDocumentVersion	*Not exported*	Current document is fr different version of anc document.
Is email attachment	IsEmailAttachment	*Not exported*	This item is from an em attachment that shows an attached item to the message.
Is inline attachment	IsInlineAttachment	*Not exported*	This was attached inlin shows up in the body c message.

Field name and Display field name	Searchable field name	Exported field name	Description
Is Representative	IsRepresentative	Is_representative	One document in every of exact duplicates is m as representative.
Item class	ItemClass	Item_class	Item class supplied by exchange server; for example, *IPM.Note*
Last modified date	LastModifiedDate	Doc_date_modified	Last modified date fror document metadata.
Load ID	LoadId	Load_ID	The ID of the load set i which the item was adc a review set.
Location	Location	Location	String that indicates th of location that docum were sourced from. **Imported Data** - Non- 365 data **Teams** - Microsoft Tear **Exchange** - Exchange mailboxes **SharePoint** - SharePoir sites **OneDrive** - OneDrive accounts
Location name	LocationName	Location_name	String that identifies th source of the item. For exchange, this will be t SMTP address of the mailbox; for SharePoint OneDrive, the URL for t site collection.
N/A	*N/A*	Marked_as_pivot	This file is the pivot in a duplicate set.
Marked as representative	MarkAsRepresentative	*Not exported*	One document from ea set of exact duplicates marked as representati
Meeting End Date	MeetingEndDate	Meeting_end_date	Meeting end date for meetings.
Meeting Start Date	MeetingStartDate	Meeting_start_date	Meeting start date for meetings.
Message kind	MessageKind	Message_kind	The type of message tc search for. Possible valu *contacts* *docs* *email*

Field name and Display field name	Searchable field name	Exported field name	Description
			externaldata *faxes* *im* *journals* *meetings* *microsoftteams* (returns items from chats, meet and calls in Microsoft T *notes* *posts* *rssfeeds* *tasks* *voicemail*
Modern Attachment Parent ID	*Not searchable*	ModernAttachment_ParentId	The Immutable ID of th document's parent.
Native Extension	NativeExtension	Native_extension	Native extension of the
Native file name	NativeFileName	Native_file_name	Native file name of the
Native file size	Size	Native_size	Number of bytes of the native item.
NativeMD5	*Not searchable*	Native_MD5	MD5 hash (128-bit has value) of the file stream
NativeSHA256	*Not searchable*	Native_SHA_256	SHA256 hash (256-bit l value) of the file stream
ND/ET Sort: Excluding attachments	NdEtSortExclAttach	ND_ET_sort_excl_attach	Concatenation of the e thread (ET) set and Nea duplicate (ND) set. This is used for efficient sor at review time. A *D* is prefixed to ND sets and is prefixed to ET sets.
ND/ET Sort: Including attachments	NdEtSortInclAttach	ND_ET_sort_incl_attach	Concatenation of an er thread (ET) set and nea duplicate (ND) set. This is used for efficient sor at review time. A *D* is prefixed to ND sets and is prefixed to ET sets. E email item in an ET set followed by its appropr attachments.
Near Duplicate Set	*Not searchable*	ND_set	Items that are similar to pivot document share t same ND_set.
O365 authors	*Not searchable*	O365_authors	Author from SharePoin

Field name and Display field name	Searchable field name	Exported field name	Description
O365 created by	*Not searchable*	O365_created_by	Created by from Sharel
O365 date created	*Not searchable*	O365_date_created	Created date from SharePoint.
O365ModifiedDate	*Not searchable*	O365_date_modified	The date a document (document version) coll from SharePoint or On for Business was modif This is the same modifi date as the one display the version history in th SharePoint and OneDri user experience.
O365 modified by	*Not searchable*	O365_modified_by	Modified by from Shar or OneDrive.
Other custodians	DedupedCustodians	Deduped_custodians	List of custodians of documents that are exa duplicates (for email, b on content; for docume based on hash).
Other file IDs	DedupedFileIds	Deduped_file_IDs	List of file IDs of docum that are exact duplicate email, based on conter documents, based on h
Other paths	Dedupedcompoundpath	Deduped_compound_path	List of compound path documents that are exa duplicates (email: base content, documents: ba on hash).
Parent ID	ParentId	Parent_ID	ID of the item's parent.
ParentNode	*Not searchable*	Parent_node	The closest preceding e message in the email th
Participant domains	ParticipantDomains	Email_participant_domains	List of all domains of participants of a messa
Participants	Participants	Email_participants	List of all participants c message; for example, *Sender*, *To*, *Cc*, *Bcc*.
Pivot ID	PivotId	Pivot_ID	The ID of a pivot.
Potentially privileged	PotentiallyPrivileged	Potentially_privileged	True if attorney-client privilege detection mo considers the documer potentially privileged

Field name and Display field name	Searchable field name	Exported field name	Description
Processing status	ProcessingStatus	Error_code	Processing status after item was added to a re set.
Read percentile	ReadPercentile	*Not exported*	Read percentile for the document based on Relevance.
Received	Received	Email_date_received	The date and time the was received in UTC.
Recipient Count	*Not searchable*	Recipient_count	Number of recipients i message.
Recipient domains	RecipientDomains	Email_recipient_domains	List of all domains of recipients of a message
Recipients	Recipients	Email_recipients	List of all recipients of message (To, Cc, Bcc).
N/A	*N/A*	Redacted_file_path	The path of the redacte replacement file in the export.
N/A	*N/A*	Redacted_text_path	The path of the redacte text file replacement in export. For internal Microsoft use only.
Relevance tag Case issue 1	*Not searchable*	Relevance_tag_case_issue_1	Relevance tag Case issu from Relevance.
Relevance score	RelevanceScore	*Not exported*	Relevance score of a document based on Relevance.
Relevance tag	RelevanceTag	*Not exported*	Relevance score of a document based on Relevance.
Representative ID	RepresentativeId	*Not exported*	Numeric identifier of e set of exact duplicates.
N/A	*N/A*	Row_number	The row number of the in the load file.
Sender	Sender	Email_sender	Sender (From) field for message types. The for *DisplayName <SmtpAddress>*.
Sender/Author	SenderAuthor	*Not exported*	Calculated field compri of the sender or autho the item.

Field name and Display field name	Searchable field name	Exported field name	Description
Sender domain	SenderDomain	Email_sender_domain	Domain of the sender.
Sent	Sent	Email_date_sent	Sent date of the messa Chats: Beginning date the transcript
Set ID	*Not searchable*	Set_ID	Documents of similar content (ND_set) or em within the same email (Email_set) share the sa Set_ID.
Set Order: Inclusive First	SetOrderInclusivesFirst	Set_order_inclusives_first	Sorting field - email an attachments: counter- chronological; docume pivot first then by descending similarity s
SimilarityPercent	*Not searchable*	Similarity_percent	Indicates how similar a document is to the piv the near duplicate set.
Subject	Subject	Email_subject	Subject of the message
Subject/Title	SubjectTitle	*Not searchable*	Calculated field compri of the subject or title o item.
Tags	Tags	Tags	Tags applied in a reviev
Team Name	TeamName	TeamName	**Teams:** Name of team **Viva Engage:** Commun name
Themes list	ThemesList	Themes_list	Themes list as calculate analytics.
Thread ID	ThreadId	Thread_ID	The Thread ID from em messages, Teams conversations, and Viva Engage conversations. email messages, all rep messages and attachm share the same Thread For Teams 1:1 and grou chats, all transcript files their associated items the same conversation the same Thread ID. Fo more information, see documents in a review
Title	Title	Doc_title	Title from the documer metadata. Title from th document metadata. F

Field name and Display field name	Searchable field name	Exported field name	Description
			Teams and Viva Engage content, this is the valu from the ConversationI property.
To	To	Email_to	To field for message ty The format is *DisplayName<SmtpAd*
Unique in email set	UniqueInEmailSet	*Not exported*	*False* if there's a duplic the attachment in its e set.
Version Group ID	*Not searchable*	Version_Group_Id	Groups together the different versions of th same document.
VersionNumber	*Not searchable*	Version_Number	The version number of document collected fro SharePoint or OneDrive Business. This is the sa version number as the displayed in the versio history in the SharePoi OneDrive user experie
Was Remediated	WasRemediated	Was_Remediated	*True* if the item was remediated, otherwise
Word count	WordCount	Word_count	Number of words in th item.

Last updated on 11/21/2025

Search for and delete Copilot data

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

You can use eDiscovery (Premium) and the Microsoft Graph Explorer to search for and delete user prompts and Microsoft 365 Copilot and Microsoft 365 Copilot Chat responses in supported applications and services. This feature can help you find and remove sensitive information or inappropriate content included in Copilot activities. This search and deletion workflow can also help you respond to a data spillage incident, when content containing confidential or malicious information is released through Copilot-related activity.

Before you search and delete Copilot data

- To create an eDiscovery (Premium) case and use collections to search for Copilot activity data, you have to be a member of the *eDiscovery Manager* role group in the Microsoft Purview portal. To delete Copilot data, you have to be assigned the *Search And Purge* role. This role is assigned to the *Data Investigator* and *Organization Management* role groups by default. For more information, see Assign eDiscovery permissions.
- A maximum of 100 items per mailbox can be removed at one time. Because the capability to search for and remove Copilot data is intended to be an incident-response tool, this limit helps ensure that this data is quickly removed.

Step 1: Create a case in eDiscovery (Premium)

The first step is to create a case in eDiscovery (Premium) to manage the search and deletion process. For information about creating a case, see Use the new case format.

Step 2: Create a collection estimate

After you create a case, the next step is to create a collection estimate to search for the Copilot data that you want to delete. The deletion process you perform is Step 5 deletes all Copilot-related items that are found in the collection estimate (within the 10 item per location limit).

In eDiscovery (Premium), a *collection* is an eDiscovery search of the content locations that contain Copilot data that you want to delete. Create the collection estimate in the case that you created in the previous step. For more information, see Create a collection estimate.

Data sources for Copilot data

The following table lists the applications and services that are sources for Copilot data. All user prompts to Copilot and responses from Copilot are stored in a user's mailbox.

Expand table

For this type of Microsoft Copilot data...	Search this item class...
Excel	IPM.SkypeTeams.Message.Copilot.Excel
Facilitator AI notes	IPM.SkypeTeams.Message.TeamCopilot.AiNotes.Teams
Loop	IPM.SkypeTeams.Message.Copilot.Loop
Microsoft 365 App	IPM.SkypeTeams.Message.Copilot.M365App
Microsoft Copilot for Bing (Microsoft 365 Copilot Chat)	IPM.SkypeTeams.Message.Copilot.BizChat
Microsoft Forms	IPM.SkypeTeams.Message.Copilot.Forms
OneNote	IPM.SkypeTeams.Message.Copilot.OneNote
Outlook	IPM.SkypeTeams.Message.Copilot.Outlook
PowerPoint	IPM.SkypeTeams.Message.Copilot.Powerpoint
SharePoint	IPM.SkypeTeams.Message.Copilot.SharePoint
Teams Channel	IPM.SkypeTeams.Message.Copilot.Teams
Teams Chat	IPM.SkypeTeams.Message.Copilot.Teams
Teams Copilot Chat (Microsoft 365 Copilot Chat)	IPM.SkypeTeams.Message.Copilot.BizChat
Teams Meeting	IPM.SkypeTeams.Message.Copilot.Teams
Teams Microsoft 365 Chat (BF)	IPM.SkypeTeams.Message
WebChat	IPM.SkypeTeams.Message.Copilot.WebChat

For this type of Microsoft Copilot data...	Search this item class...
Whiteboard	IPM.SkypeTeams.Message.Copilot.Whiteboard
Word	IPM.SkypeTeams.Message.Copilot.Word

> **Note**
>
> In Step 4, you also have to identify and remove any holds and retention policies assigned to the mailbox that contains the type of Copilot data that you want to delete.

Tips for searching for Copilot data

To help ensure the most comprehensive collection of Copilot data, use the **Type** condition and select the **Copilot activity** option when you build the search query for the collection estimate. We also recommend including a date range or several keywords to narrow the scope of the collection to items relevant to your search and delete investigation.

For more information, see Build search queries for collections.

Identifying web queries in Microsoft 365 Copilot usage

With the web search enabled for Microsoft 365 Copilot or Microsoft Copilot to include the latest data from the web, the web search queries sent to Microsoft Bing are searchable in eDiscovery. For more information about web search, see Data, privacy, and security for web search in Microsoft 365 Copilot and Microsoft 365 Copilot Chat.

Complete the following steps for finding these web queries:

1. Using the Condition builder in eDiscovery, search for Copilot activity using the filter *Type*, *Equals any of*, and *Copilot activity*.
2. In the query results, download any single item.
3. Open up the downloaded item in a text editor.
4. Look for *WebSearchQuery*
5. If the Copilot activity is involved in a Bing search query, *WebSearchQuery* is present in the downloaded file. It's followed by the specific query sent in the Microsoft Bing search query.

Step 3: Review and verify Copilot data to delete

The deletion process in Step 5 will delete the items returned by the collection. It's important that you review the collection estimate results to ensure that the collection only returns the items that you want to delete. To review a sample of items in a collection estimate, see the **Next steps after a collection estimate is complete** section in Create a collection estimate.

Additionally, you can use the collection statistics (specifically the *Top Locations* statistics) to generate a list of the data sources that contain items returned by the collection. Use this list in the next step to remove hold and retention policies from the user mailboxes that contain search results. For more information, see Collection statistics and reports.

Step 4: Remove holds and retention policies from data sources

Before you can delete Copilot data from a mailbox, you have to remove any hold or retention policy that is assigned to a target mailbox. If not, then the data you're trying to delete is retained.

Use the list of mailboxes that contain the Copilot data that you want to delete and determine if there's a hold or retention policy assigned to those mailboxes, and then remove the hold or retention policy. Be sure to identify the hold or retention policy that you remove so that you can reassign to the mailboxes in Step 7.

For instructions about how to identify and remove holds and retention policies, see **Step 3: Remove all holds from the mailbox** in Delete items in the Recoverable Items folder of cloud-based mailboxes on hold.

Step 5: Delete Copilot data

> **Note**
>
> Because Microsoft Graph Explorer is not available in some US Government clouds (GCC High and DOD), you must use PowerShell to accomplish these tasks. See the **Delete Copilot data with PowerShell** for details.

Now you're ready to delete Copilot data from user mailboexes. Use the Microsoft Graph Explorer to perform the following three tasks:

1. Get the ID of the eDiscovery (Premium) case that you created in Step 1. This is the case that contains the collection created in Step 2.

2. Get the ID of the collection that you created in Step 2 and verified the search results in Step 3. The search query in this collection returns the Copilot data to be deleted.
3. Delete the Copilot data returned by the collection.

For information about using Graph Explorer, see Use Graph Explorer to try Microsoft Graph APIs.

> **Important**
>
> To perform these three tasks in Graph Explorer, you may have to consent to the eDiscovery.Read.All and eDiscovery.ReadWrite.All permissions. For more information, see the "Consent to permissions" section in **Working with Graph Explorer**.

Get the case ID

1. Go to https://developer.microsoft.com/graph/graph-explorer and sign in to the Graph Explorer with an account that's assigned the *Search And Purge* role in the Microsoft Purview portal.

2. Run the following GET request to retrieve the ID for the eDiscovery (Premium) case. Use the value `https://graph.microsoft.com/v1.0/security/cases/ediscoveryCases` in the address bar of the request query. Be sure to select **v1.0** in the API version dropdown list.

   This request returns information about all cases in your organization on the **Response preview** tab.

3. Scroll through the response to locate the eDiscovery (Premium) case. Use the **displayName** property to identify the case.

4. Copy the corresponding ID (or copy and paste it to a text file). You'll use this ID in the next task to get the collection ID.

> **Tip**
>
> Instead of using the previous procedure to obtain the case Id, you can open the case in the Microsoft Purview portal and copy the case Id from the URL.

Get the eDiscoverySearchID

1. In Graph Explorer, run the following GET request to retrieve the ID for the collection that you created in Step 2, and contains the items you want to delete. Use the value `https://graph.microsoft.com/v1.0/security/cases/ediscoveryCases/{ediscoveryCaseID}/searches` in the address bar of the request query, where *{ediscoveryCaseID}* is the CaseID that you obtained in the previous procedure.

2. Scroll through the response to locate the collection that contains the items that you want to delete. Use the *displayName* property to identify the collection that you created in Step 3.

   In the response, the search query from the collection is displayed in the *contentQuery* property. Items returned by this query are deleted in the next task.

3. Copy the corresponding ID (or copy and paste it to a text file). You'll use this ID in the next task to delete Copilot data.

> **Tip**
>
> Instead of using the previous procedure to obtain the search Id, you can open the case in the Microsoft Purview portal. Open the case and navigate to the Jobs tab. Select the relevant collection and under Support information, find the job ID (the job ID displayed here is the same as the collection ID).

Delete Copilot data

1. In Graph Explorer, run the following POST request to delete the items returned by the collection that you created in Step 2. Use the value `https://graph.microsoft.com/v1.0/security/cases/ediscoveryCases/{ediscoveryCaseID}/searches/{ediscoverySearchID}/purgeData` in the address bar of the request query, where *{ediscoveryCaseID}* and *{ediscoverySearchID}* are the IDs that you obtained in the previous procedures.

   If the POST request is successful, an HTTP response code is displayed in a green banner stating that the request was accepted.

For more information on purgeData, see sourceCollection: purgeData.

Delete Copilot data with PowerShell

> **Note**

Because Microsoft Graph Explorer is not available in the US Government cloud (GCC, GCC High, and DOD), you must use PowerShell to accomplish these tasks.

You can also delete Copilot data using PowerShell. For example, to delete Copilot data in the US Government cloud you could use a command similar to:

```
Connect-MgGraph -Scopes "ediscovery.ReadWrite.All" -Environment USGov
```

```
Invoke-MgGraphRequest -Method POST -Uri
'/v1.0/security/cases/ediscoveryCases/<ediscoverySearchID>/searches/<search ID>/purgeData'
```

For more information on using PowerShell to delete Copilot data, see ediscoverySearch: purgeData.

Step 6: Verify Copilot data is deleted

After you run the POST request to delete Copilot data, this data is removed from the user's mailbox. There isn't any visible notification or confirmation for the user that the data has been deleted.

Deleted Copilot data is moved to the *SubstrateHolds* folder, which is a hidden mailbox folder. Deleted Copilot data is stored there for at least 1 day and then are permanently deleted the next time the timer job runs (typically between 1-7 days).

Step 7: Reapply holds and retention policies to user mailboxes

After you verify that the Copilot data is deleted, you can reapply the holds and retention policies to user mailboxes that you removed in Step 4.

Last updated on 09/02/2025

Find sensitive data stored on sites with eDiscovery

> **⊗ Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

Users often store sensitive data, such as credit card numbers, social security numbers, or personal, on their sites, and over time this can expose an organization to significant risk of data loss. Documents stored on sites—including OneDrive for Business sites—could be shared with people outside the organization who shouldn't have access to the information.

With Microsoft Purview eDiscovery, you can discover documents that contain sensitive data throughout your tenant. After discovering the documents, you can work with the document owners to protect the data. This article can help you form a query to search for sensitive data.

Forming a basic query

There are three parts that make up a basic query: SensitiveType, count range, and confidence range. For example, **SensitiveType:"<type>"** is required, and both **|<count range>** and **|<confidence range>** are optional.

Sensitive type - required

Queries typically begin with the property `SensitiveType:"` and an information type name from the sensitive information types inventory, and end with a `"`. To search for custom sensitive information types, you have to specify the ID of the sensitive information type in the `SensitiveType` property. Using the name of a custom sensitive information type won't return any results.

Ranges - optional

Both of the next two parts are ranges, so let's quickly examine what a range looks like. In SharePoint queries, a basic range is represented by two numbers separated by two periods, which looks like this: `[number]..[number]`. For instance, if `10..20` is used, that range would capture numbers from 10 through 20. There are many different range combinations and several are covered in this article.

Let's add a count range to the query. You can use count range to define the number of occurrences of sensitive information a document needs to contain before it's included in the query results. For example, if you want your query to return only documents that contain exactly five credit card numbers, use this: `SensitiveType:"Credit Card Number|5"`. Count range can also help you identify documents that pose high degrees of risk. For example, your organization might consider documents with five or more credit card numbers a high risk. To find documents fitting this criterion, you would use this query: `SensitiveType:"Credit Card Number|5.."`. Alternatively, you can find documents with five or fewer credit card numbers by using this query: `SensitiveType:"Credit Card Number|..5"`.

Confidence range

Finally, confidence range is the level of confidence that the detected sensitive type is actually a match. The values for confidence range work similarly to count range. You can form a query without including a count range. For example, to search for documents with any number of credit card numbers—as long as the confidence range is 85 percent or higher—you would use this query: `SensitiveType:"Credit Card Number|*|85.."`.

> **Important**
>
> The asterisk (`*`) is a wildcard character that means any value works. You can use the wildcard character (`*`) either in the count range or in the confidence range, but not in a sensitive type.

Additional query properties and search operators available in the eDiscovery Center

Queries in SharePoint also contains the LastSensitiveContentScan property, which can help you search for files scanned within a specific timeframe. For query examples with the `LastSensitiveContentScan` property, see the Examples of complex queries in the next section.

You can use SharePoint eDiscovery search properties such as `Author` or `FileExtension`. You can use operators to build complex queries. For the list of available properties and operators, see the Using Search Properties and Operators with eDiscovery blog post.

Examples of complex queries

The following examples use different sensitive types, properties, and operators to illustrate how you can refine your queries to find exactly what you're looking for.

Expand table

Query	Explanation
`SensitiveType:"International Banking Account Number (IBAN)"`	The name might seem strange because it's so long, but it's the correct name for that sensitive type. Make sure to use exact names from the sensitive information types inventory. You can also use the name of a custom sensitive information type that you created for your organization.
`SensitiveType:"Credit Card Number|1..4294967295|1..100"`	This returns documents with at least one match to the sensitive type "Credit Card Number." The values for each range are the respective minimum and maximum values. A simpler way to write this query is `SensitiveType:"Credit Card Number"`, but where's the fun in that?
`SensitiveType:"Credit Card Number|5..25" AND LastSensitiveContentScan:"8/11/2018..8/13/2018"`	This returns documents with 5-25 credit card numbers that were scanned from August 11, 2018 through August 13, 2018.
`SensitiveType:"Credit Card Number|5..25" AND LastSensitiveContentScan:"8/11/2018..8/13/2018" NOT FileExtension:XLSX`	This returns documents with 5-25 credit card numbers that were scanned from August 11, 2018 through August 13, 2018. Files with an XLSX extension aren't included in the query results. `FileExtension` is one of many properties that you can include in a query. For more information, see Using Search Properties and Operators with eDiscovery.
`SensitiveType:"Credit Card Number" OR SensitiveType:"U.S. Social Security Number (SSN)"`	This returns documents that contain either a credit card number or a social security number.

Examples of queries to avoid

Not all queries are created equal. The following table gives examples of queries that don't work in SharePoint and describes why.

Expand table

Unsupported query	Reason
`SensitiveType:"Credit Card Number|.."`	You must add at least one number.
`SensitiveType:"NotARule"`	"NotARule" isn't a valid sensitive type name. Only names in the sensitive information types inventory work in eDiscovery queries.
`SensitiveType:"Credit Card Number|0"`	Zero isn't valid as either the minimum value or the maximum value in a range.
`SensitiveType:"Credit Card Number"`	It's might be difficult to see, but there's extra white space between "Credit" and "Card" that makes the query invalid. Use exact sensitive type names from the sensitive information types inventory.
`SensitiveType:"Credit Card Number|1. .3"`	The two-period portion shouldn't be separated by a space.
`SensitiveType:"Credit Card Number| |1..|80.."`	There are too many pipe delimiters (|). Follow this format instead: `SensitiveType: "Credit Card Number|1..|80.."`
`SensitiveType:"Credit Card Number|1..|80..101"`	Because confidence values represent a percentage, they can't exceed 100. Choose a number from 1 through 100 instead.

For more information

- Sensitive information type entity definitions
- Run a Content Search
- Keyword queries and search conditions for Content Search

Last updated on 09/02/2025

Set up historical versions in eDiscovery (Premium) (preview)

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

The historical versions feature in eDiscovery (Premium) lets eDiscovery managers in your organization search for and collect content from all versions of documents stored in SharePoint Online and OneDrive for Business. Then you can add that content to a review set for analysis and review. This helps you find and review content from a specific version of a document that may be relevant to a case or investigation, even if the latest version of the same document doesn't contain the relevant information.

To support the historical versions capability in eDiscovery (Premium), SharePoint administrators must enable versioning for sites in their organization. Then, when users modify documents in SharePoint or OneDrive, implicit regular versions are created when document is saved (or autosaved). SharePoint versioning allows for tracking of the activity performed on SharePoint items (including documents, events, and tasks). This versioning capability leaves an audit trail that can provide evidence in legal investigations. These older versions of a document are available to the organization, who may be required to share such versions that have sensitive or relevant content during court discovery in a legal matter.

After an eDiscovery administrator turns on historical versions for the organization and then activates it for specific SharePoint sites, the SharePoint content push service crawls all major and minor versions of documents on the activated sites and then sends those versions for indexing. After the crawling and indexing process is complete, documents and their versions are available for eDiscovery search. As long as a specific version can be accessed (by version history), then that version will be discoverable in an eDiscovery (Premium) collection search.

Set up historical versions

To enable historical versions in eDiscovery (Premium), your organization has to turn it on and then activate specific sites so that all versions of documents stored on those sites are indexed for search. Before you set up eDiscovery (Premium) for historical versions, you have to enable versioning support in SharePoint.

Step 1: Turn on versioning in SharePoint

The first step is to turn on versioning in SharePoint Online so that all versions of a document are preserved. For instructions, see Versioning in SharePoint.

> **Important**
>
> Adding SharePoint sites to historical versions isn't currently supported in multi-geo environments. Only the primary domain of your organization is enabled for historical versions.

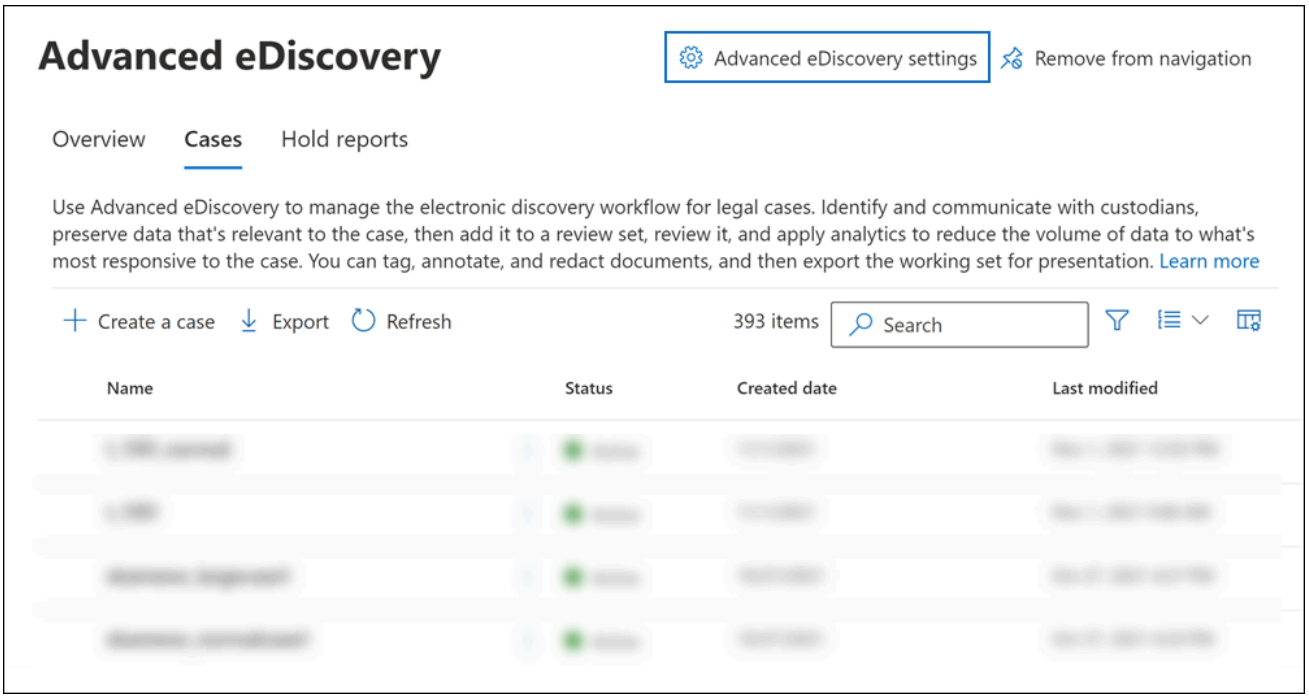
Step 2: Turn on historical versions

The next step is to turn on historical versions in eDiscovery (Premium). To turn on historical versions for your organization, a person must be a global administrator or an eDiscovery Administrator (a member of the eDiscovery Administrator subgroup in the eDiscovery Manager role group). After historical versions is turned on, it will apply to your entire organization.

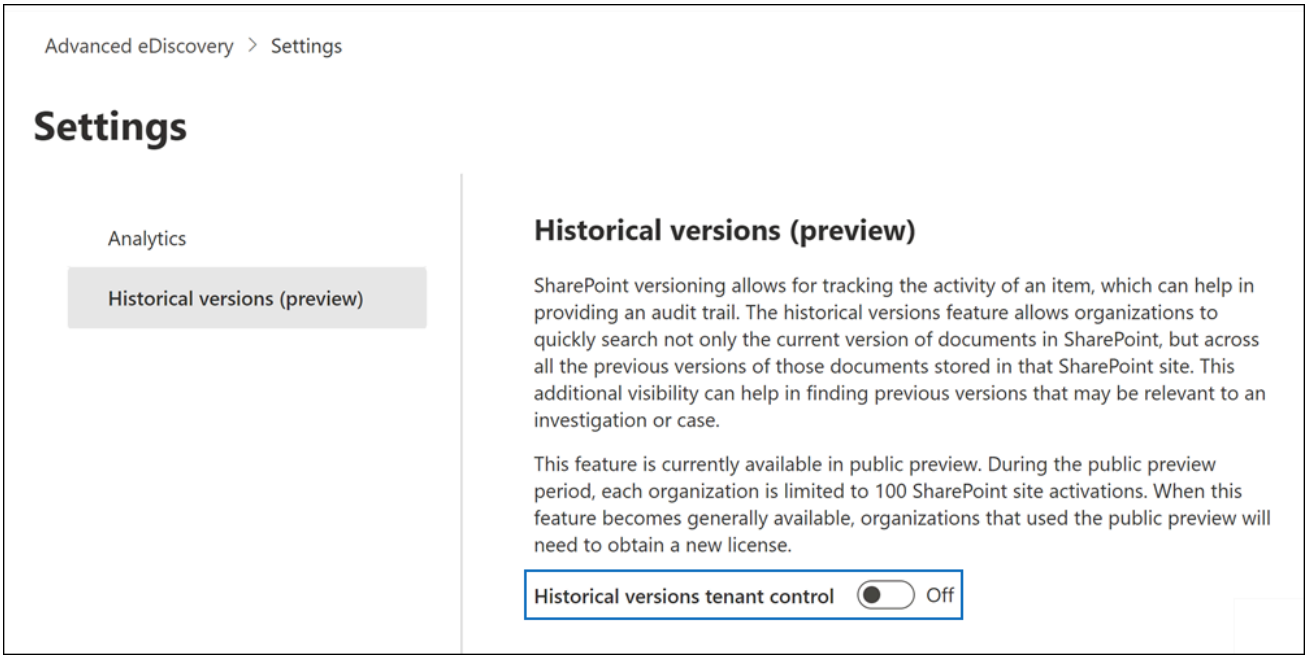
> **Important**
>
> Microsoft recommends that you use roles with the fewest permissions. Minimizing the number of users with the Global Administrator role helps improve security for your organization. Learn more about Microsoft Purview roles and permissions.

> **Important**
>
> After you turn on historical versions, you won't be able to turn it off during the public preview. You will be able to turn it off after historical versions is released for general availability.

1. In the Microsoft Purview portal, go to eDiscovery (Premium), and then select **eDiscovery (Premium) settings**.



2. On the **Settings** page, select the **Historical versions (preview)** tab, and then switch the **Historical versions tenant control** toggle to on.



   A warning is displayed that says you won't be able to turn off historical versions. As previously stated, you won't be able to turn off historical versions until the feature is released for general availability.

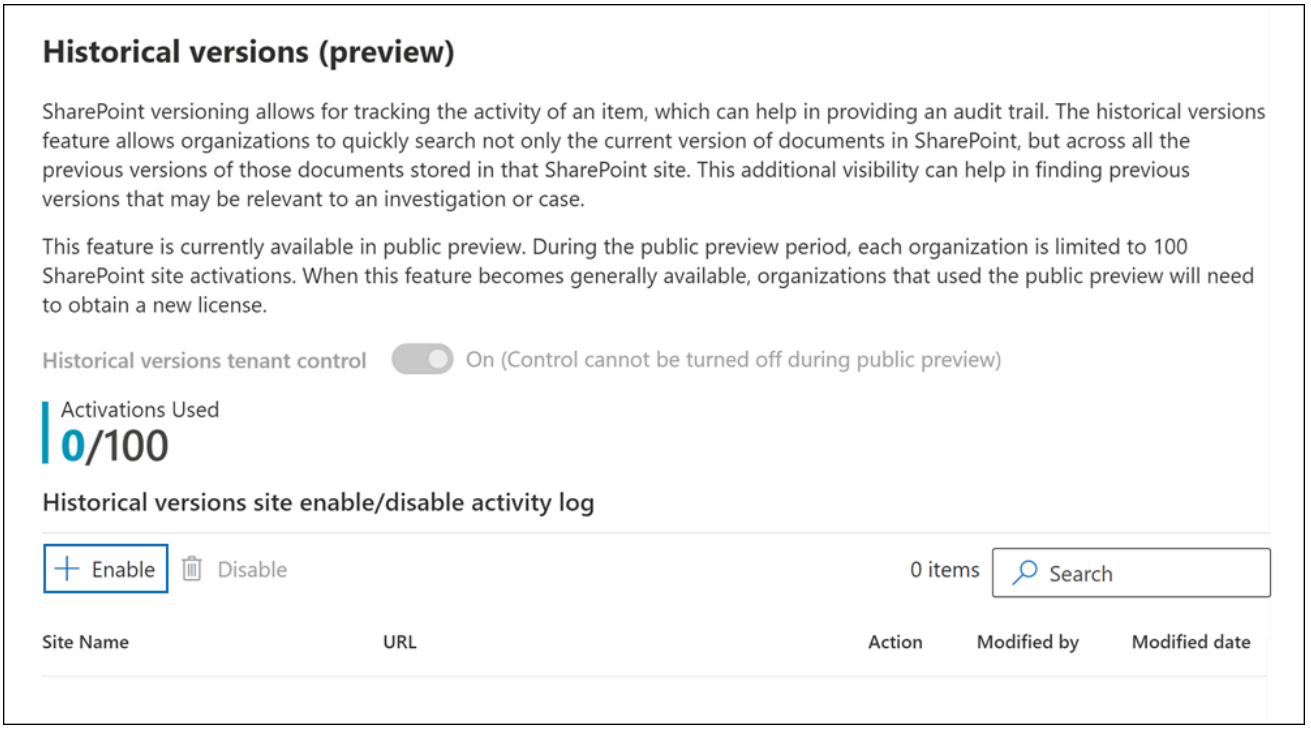
3. Select **Yes** to turn on historical versions.

Step 3: Activate SharePoint sites

After you turn on historical versions for your organization, the last step is to activate SharePoint sites to support historical versions. When you activate a site (by adding it to a list of sites on the **Historical versions** tab), the site is recrawled and all versions of documents stored on that site are indexed for search.

> **Note**
>
> There is a limit of 100 site activations per organization during the public preview of historical versions. An activation is counted against this limit whenever you enable or disable a site for historical versions. If you enable multiple sites, each site is counted as a single activation. The total number of activations is displayed on the **Historical versions** tab.

1. On the **Historical versions** tab on the eDiscovery (Premium) **Settings** page, select **Enable** to activate sites for historical versions.

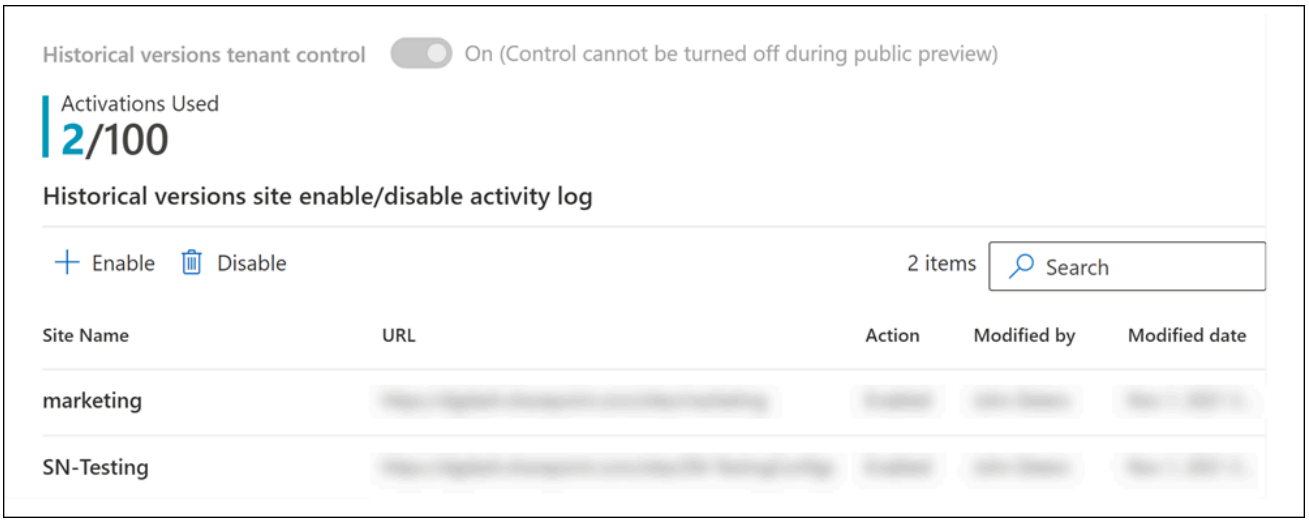


   A flyout out page is displayed that contains a list of all SharePoint sites in your organization.

2. Select a site to activate, and then select **Enable** to activate it for historical versions. You can use the search box to search for a specific site.

   A warning is displayed that says document versions on the site will be indexed and this indexing process will take some time before the versions will be ready to be searched. The warning also states that you won't be able to disable historical versions for the selected site for 30 days.

3. Select **Yes** to activate the site for historical versions.



The site is added to the list of activated sites. The counter showing the number of site activations is also updated.

4. Repeat the previous steps for each site that you want to activate for historical versions.

Frequently asked questions

How are historical versions different that the option to "collect all versions" when you commit a collection estimate to a review set?

Currently, only the latest version of documents is indexed for search. That means when you run a collection estimate, only the latest versions of documents are searched. If a document matches the keyword query for the collection, it is returned in the collection results. However, if the latest version of a document doesn't match a search query, the document won't be returned event if older versions of the document contain the keyword. To help mitigate this situation, eDiscovery (Premium) gives you the ability to collect all versions of the document when you commit a collection to a review set. That means any older version that may contain the keyword will be added to the review set.

Historical versions are different and more efficient than "collecting all versions" because when you activate a site, all versions of a document (and not just the last version) are indexed for search. The result is that if an older version of a document contains a keyword that matches the search query, it will be returned by the collection.

When historical versions is enabled for a site, does it impact the performance of the site?

No. Once historical versions is enabled for a site, the performance for the site will the same as it was before the site was enabled. The crawling and indexing processes that are performed on the

site after it's enabled will occur at a slower rate and be performed during off-peak hours. Enabling historical versions for a site will kick off a backfill process, which finds all the versions of documents on the site and then sends those versions to the index. Depending on the number of document versions for site, this backfill process might impact service health. We have mitigated this potential impact in the following ways:

- We make the best effort to process these versions during off-peak hours.
- We process document versions in our lowest priority queues, which allows most service resources to be delegated to user changes.

How long do I have to wait after a site is activated until all historical versions of documents on that site are all indexed and available for eDiscovery search?

Based on the number of documents for a site and the average number of versions per document, we try to estimate the total number of files per site. Based on this, an estimate of how long it may take to index is as follows:

`Number of versions / (Processing rate of 100,000 files per day ) + .5 days = Total number of days to process a site`

The half-day is added as a buffer since indexing of versions on site is optimized to run during off-peak hours.

For example, if the total number of documents and all versions for a site is 150,000 then it will take at least two days to process the site for historical versions:

`150,000 files/100,000 files per day + .5 days = 2 days`

Why is it not recommended to frequently activate or deactivate sites for historical versions?

When you deactivate a previously activated site, a clean-up process is triggered. This process will take time to complete. If that same site is activated again, the backfill process of reindexing the site has to be rerun. Both the clean-up and backfill processes are time- and resource-intensive. Therefore we recommended that you carefully consider and plan which sites you want to activate for historical version so you can avoid repeatedly activating and deactivating historical versions for a site.

Last updated on 09/02/2025

Set up attorney-client privilege detection in eDiscovery (Premium)

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

A major and costly aspect of the review phase of any eDiscovery process is reviewing documents for privileged content. Microsoft Purview eDiscovery (Premium) provides machine learning-based detection of privileged content to make this process more efficient. This feature is called *attorney-client privilege detection*.

How does it work?

When attorney-client privilege detection is enabled, all documents in a review set will be processed by the attorney-client privilege detection model when you analyze the data in the review set. The model looks for two things:

- **Privileged content**: The model uses machine learning to determine the likelihood that the document contains content that is legal in nature.
- **Participants**: As part of setting up attorney-client privilege detection, you have to submit a list of attorneys for your organization. The model then compares the participants of the document with the attorney list to determine if a document has at least one attorney participant.

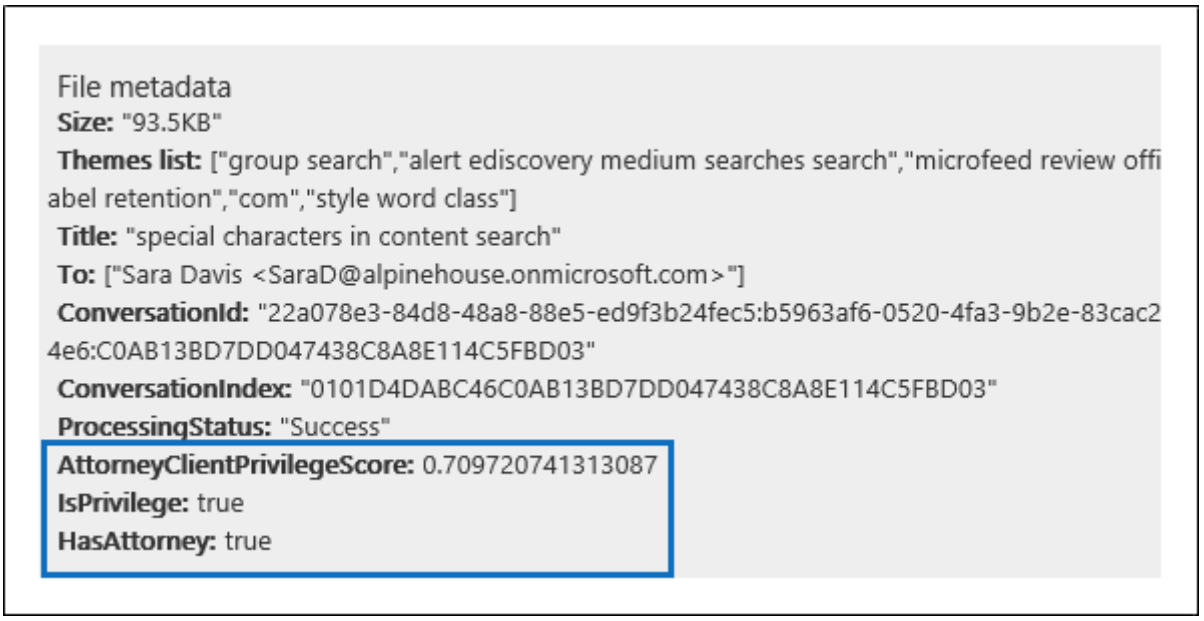
The model produces the following three properties for every document:

- **AttorneyClientPrivilegeScore:** The likelihood the document is legal in nature; the values for the score are between **0** and **1**.
- **HasAttorney:** This property is set to **true** if one of the document participants is listed in the attorney list; otherwise the value is **false**. The value is also set to **false** if your organization

didn't upload an attorney list.

- **IsPrivilege:** This property is set to **true** if the value for **AttorneyClientPrivilegeScore** is over the threshold *or* if the document has an attorney participant; otherwise the value is set to **false**.

These properties (and their corresponding values) are added to the file metadata of the documents in a review set, as shown in the following screenshot:



These three properties are also searchable within a review set. For more information, see Query the data in a review set.

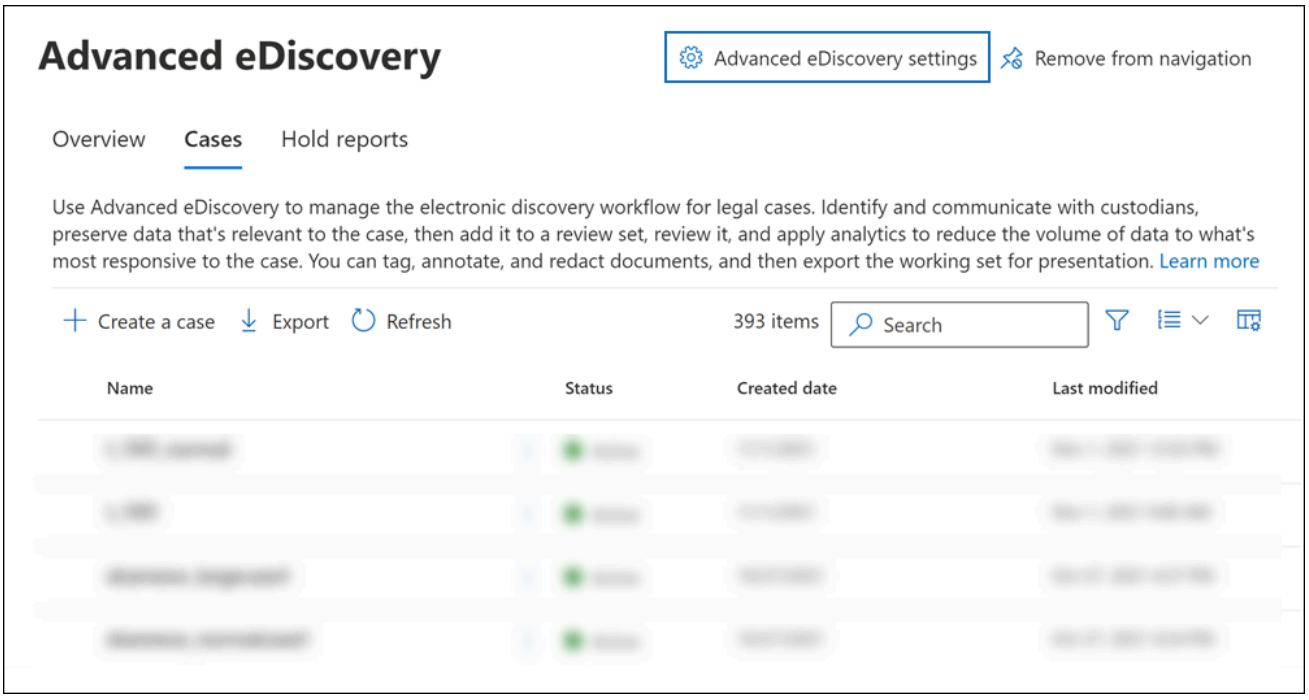
Set up the attorney-client privilege detection model

To enable the attorney-client privilege detection model, your organization has to turn it on and then upload an attorney list.

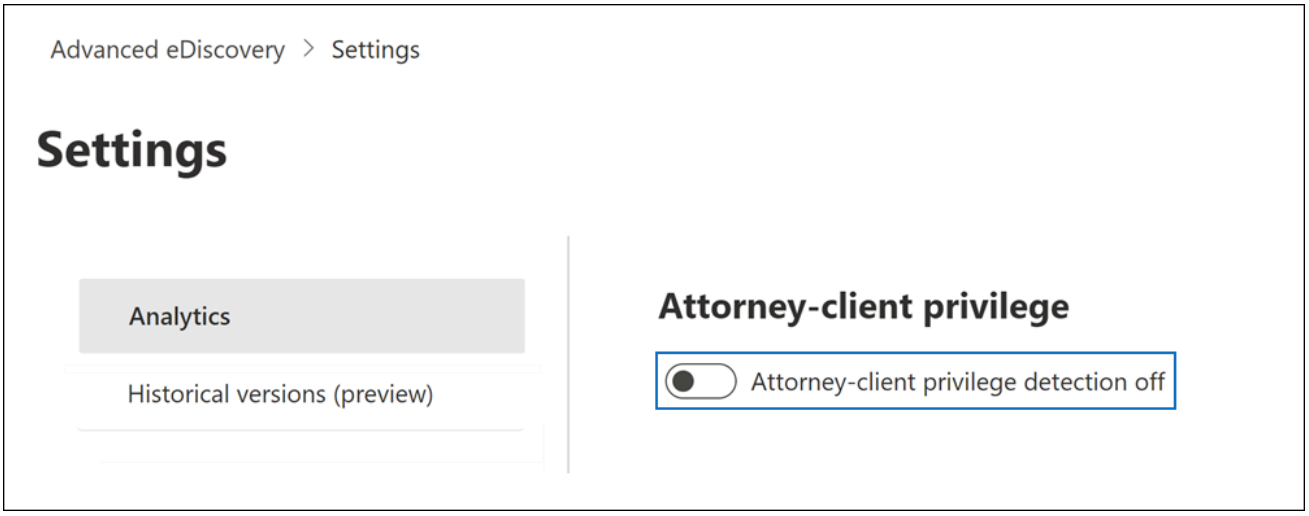
Step 1: Turn on attorney-client privilege detection

A person who is an eDiscovery Administrator in your organization (a member of the eDiscovery Administrator subgroup in the eDiscovery Manager role group) must make the model available in your eDiscovery (Premium) cases.

1. In the Microsoft Purview portal, go to eDiscovery (Premium), and then select **eDiscovery (Premium) settings**.



2. On the **Settings** page, select the **Analytics** tab, and then switch the **Attorney-client privilege detection** toggle to on.



3. Select **Save** to save the change.

Step 2: Upload a list of attorneys (optional)

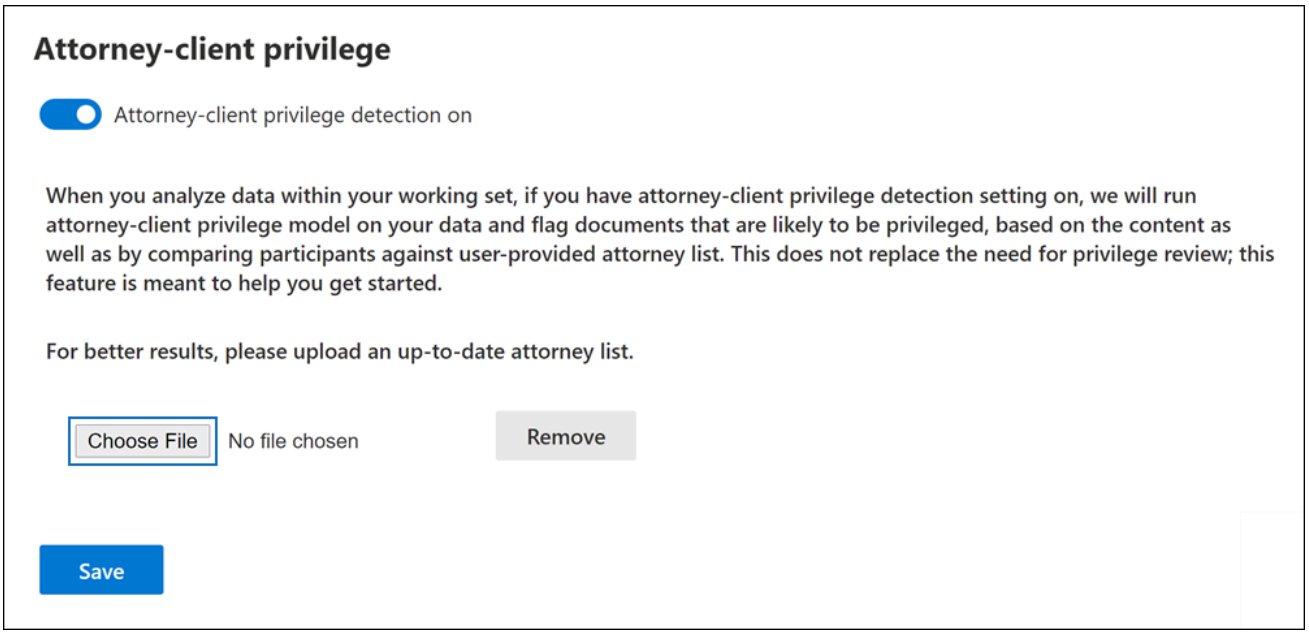
To take full advantage of the attorney-client privilege detection model and use the results of the **Has Attorney** or **Potentially Privileged** detection that was previously described, we recommend that you upload a list of email addresses for the lawyers and legal personnel who work for your organization.

To upload an attorney list for use by the attorney-client privilege detection model:

1. Create a .csv file (without a header row) and add the email address for each appropriate person on a separate line. Save this file to your local computer.

2. On the eDiscovery (Premium) **Settings** page, select the **Analytics** tab.

   The **Attorney-client privilege** page is displayed, and the **Attorney-client privilege detection** toggle is turned on.



3. Select **Choose file** and then find and select the .csv file that you created in step 1.

4. Select **Save** to upload the attorney list.

Use the attorney-client privilege detection model

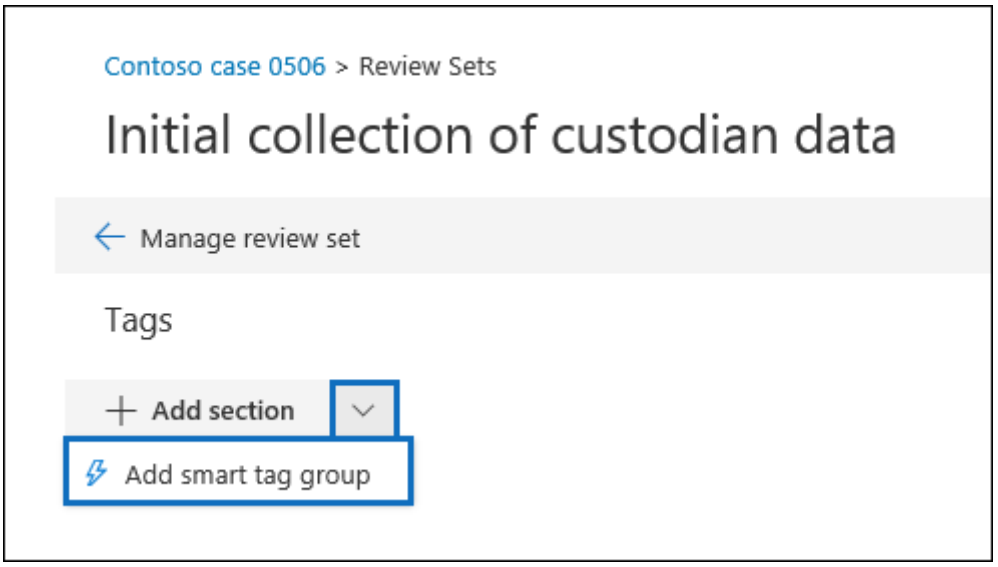
Follow the steps in this section to use attorney-client privilege detection for documents in a review set.

Step 1: Create a smart tag group with attorney-client privilege detection model

One of the primary ways to see the results of attorney-client privilege detection in your review process is by using a smart tag group. A smart tag group indicates the results of the attorney-client privilege detection and shows the results in-line next to the tags in a smart tag group. This lets you quickly identify potentially privileged documents during document review. Additionally, you can also use the tags in the smart tag group to tag documents as privileged or non-privileged. For more information about smart tags, see Set up smart tags in eDiscovery (Premium).

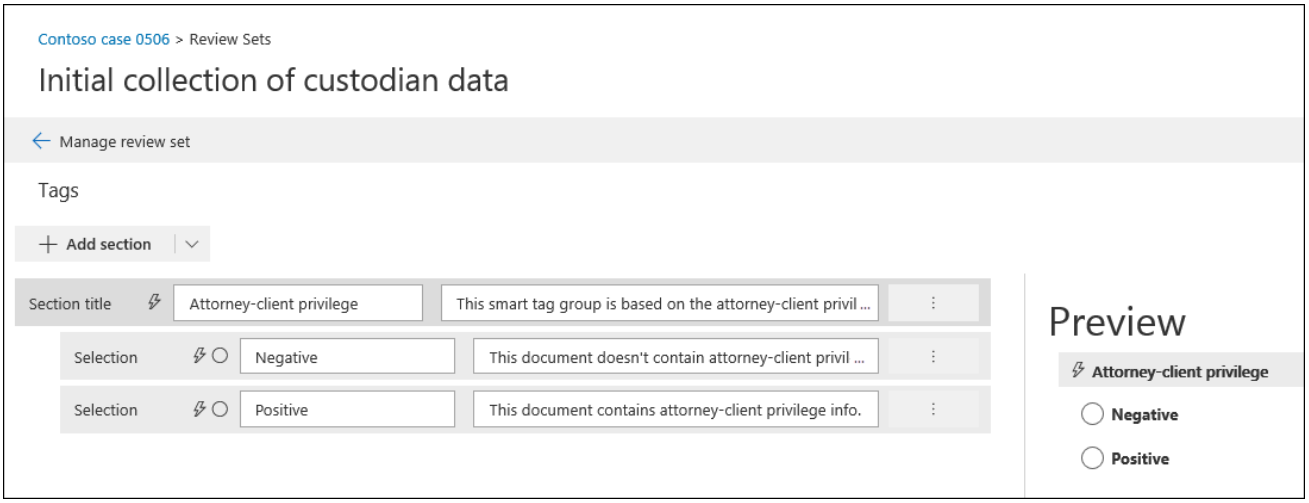
1. In the review set that contains the documents that you analyzed in Step 1, select **Manage review set** and then select **Manage tags**.

2. Under **Tags**, select the pull-down next to **Add group** and then select **Add smart tag group**.



3. On the **Choose a model for your smart tag** page, choose **Select** next to **Attorney-client privilege**.

   A tag group named **Attorney-client privilege** is displayed. It contains two child tags named **Positive** and **Negative**, which correspond to the possible results produced by the model.



4. Rename the tag group and tags as appropriate for your review. For example, you can rename **Positive** to **Privileged** and **Negative** to **Not privileged**.

Step 2: Analyze a review set

When you analyze the documents in a review set, the attorney-client privilege detection model will also run and the corresponding properties (described in How does it work?) will be added to every document in the review set. For more information about analyzing data in review set, see Analyze data in a review set in eDiscovery (Premium).

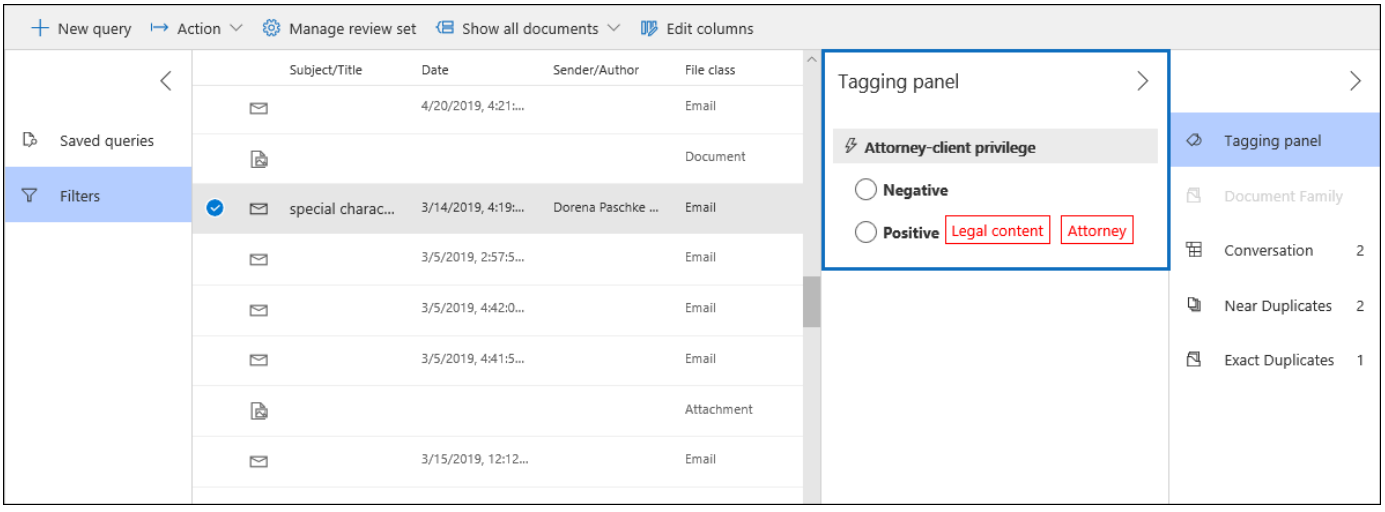
Step 3: Use the smart tag group for review of privileged content

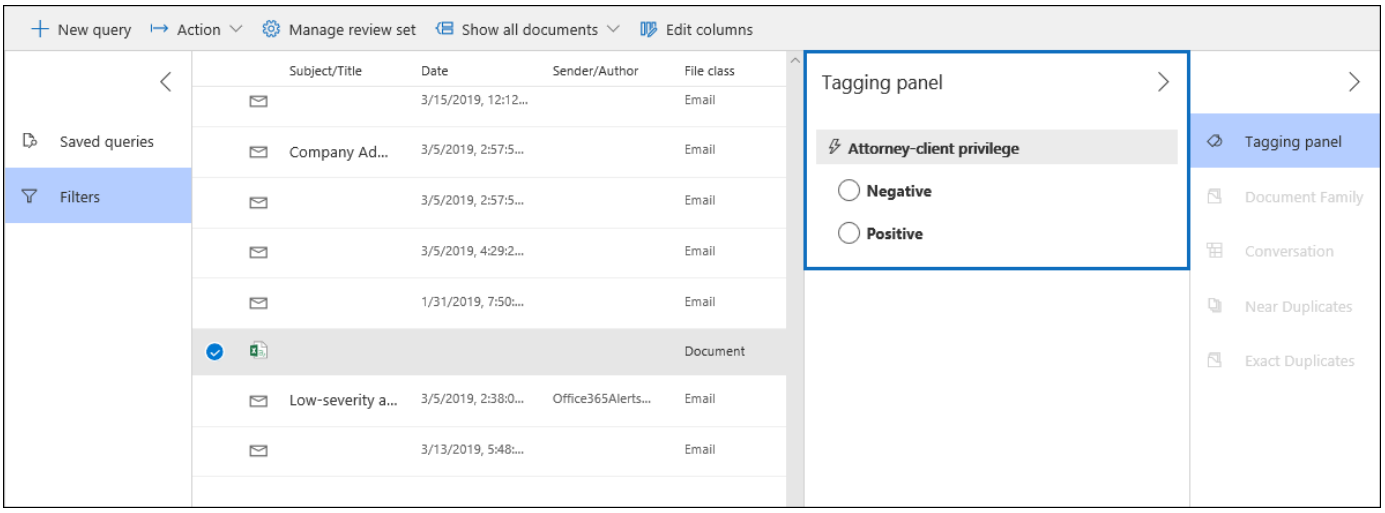
After analyzing the review set and setting up smart tags, the next step is to review the documents. If the model has determined the document is potentially privileged, the corresponding smart tag in the **Tagging panel** will indicate the following results produced by the attorney-client privilege detection:

- If the document has content that may be legal in nature, the label **Legal content** is displayed next to the corresponding smart tag (which in this case is the default **Positive** tag).
- If the document has a participant who is found in your organization's attorney list, the label **Attorney** is displayed next to the corresponding smart tag (which in this case is also the default **Positive** tag).
- If the document has content that may be legal in nature *and* has a participant found in the attorney list, both the **Legal content** and **Attorney** labels are displayed.

If the model determines that a document doesn't contain content that is legal in nature or doesn't contain a participant from the attorney list, then neither label is displayed in the tagging panel.

For example, the following screenshots show two documents. The first one contains content that is legal in nature and has a participant found in the list of attorneys. The second contains neither and therefore doesn't display any labels.





After you review a document to see if it contains privileged content, you can tag the document with the appropriate tag.

Last updated on 09/02/2025

Assign eDiscovery permissions in the Microsoft Purview portal

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

To use any of the eDiscovery-related tools in the Microsoft Purview portal, assign the appropriate permissions. The easiest way to assign roles is to add the person to the appropriate role group on the **Permissions** page in the Purview portal. This article describes the permissions required to perform these tasks.

> **Tip**
>
> You can view your own permissions on the eDiscovery (Premium) overview page in the Purview portal. You must have at least one role assigned for your permissions to be displayed.

The primary eDiscovery-related role group in Purview portal is called **eDiscovery Manager**. Two subgroups exist within this role group:

- **eDiscovery Manager** - An eDiscovery Manager can use eDiscovery search tools to search content locations in the organization, and perform various search-related actions such as preview and export search results. Members can also create and manage cases in Microsoft Purview eDiscovery (Premium), add and remove members to a case, create case holds, run searches associated with a case, and access case data. eDiscovery Managers can only access and manage the cases they create. They can't access or manage cases created by other eDiscovery Managers.

- **eDiscovery Administrator** - An eDiscovery Administrator is a member of the eDiscovery Manager role group, and can perform the same content search and case management-related tasks that an eDiscovery Manager can perform. An eDiscovery Administrator can also:
  - Access all cases that are listed on the **eDiscovery (Premium)** page in the Purview portal.
  - Access case data in eDiscovery (Premium) for any case in the organization.
  - Manage any eDiscovery case after they add themselves as a member of the case.
  - Remove members from an eDiscovery case. Only an eDiscovery Administrator can remove members from a case. Users who are members of the eDiscovery Manager subgroup can't remove members from a case, even if the user created the case.

  For reasons why you might want eDiscovery Administrators in your organization, see eDiscovery Administrator considerations.

> **Note**
>
> eDiscovery (Premium) licensing requirements:
>
> - The custodian of the data must be assigned an Office 365 E5 or Microsoft 365 E5 license.
> - Users with an Office 365 E1 or an Office 365 or Microsoft 365 E3 license can be assigned a Microsoft Purview Suite (formerly known as Microsoft 365 E5 Compliance) or Microsoft 365 eDiscovery and Audit add-on license.
> - Administrators, compliance officers, or legal personnel who are assigned to cases as members and use eDiscovery (Premium) to collect, view, and analyze data don't need an E5 license.
>
> For more information about eDiscovery (Premium) licensing, see **Subscriptions and licensing in eDiscovery (Premium)**.

Prerequisites

- You must be a member of the *Organization Management* role group or be assigned the *Role Management* role to assign eDiscovery permissions in the Purview portal.
- You can use the Add-RoleGroupMember cmdlet in Security & Compliance PowerShell to add a mail-enabled security group as a member of the *eDiscovery Managers* subgroup in the *eDiscovery Manager* role group. However, you can't add a mail-enabled security group

to the *eDiscovery Administrators* subgroup. For details, see eDiscovery Administrator considerations.

Assign eDiscovery permissions

To assign eDiscovery permissions to users in the Microsoft Purview portal, follow these steps:

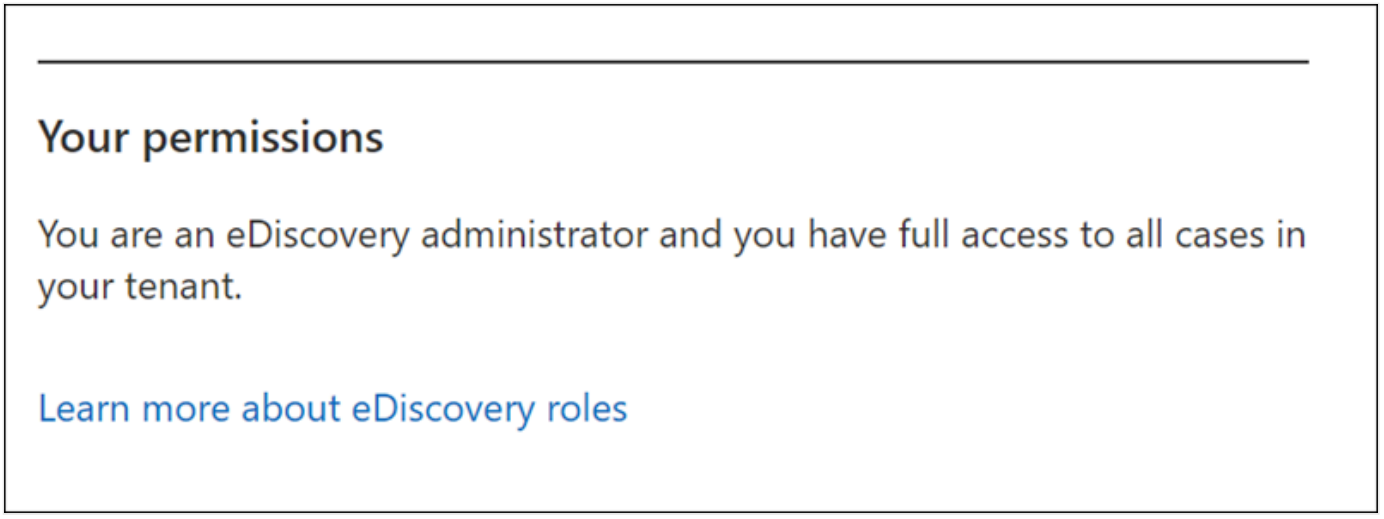
1. Go to the Purview portal and sign in with an account that can assign permissions.
2. In the left pane, select **Roles & scopes** > **Permissions**.
3. On **Permissions**, under **Microsoft Purview solutions**, select **Roles**.
4. On **Role groups for Microsoft Purview solutions**, select **eDiscovery Manager**.
5. On the **eDiscovery Manager** flyout, choose one of the following options based on the eDiscovery permissions that you want to assign.
   a. Select **Edit**.
   b. On **Manage eDiscovery Manager**, select **Choose users**.
   c. Search for and select the users you want to add as *eDiscovery Manager*, then select **Select**.
   d. Select **Next**.
   e. To assign users to the *eDiscovery Administrator* role group, select **Choose users**.
   f. Search for and select the users you want to add as *eDiscovery Administrator*, then select **Select**.
   g. Select **Next**.
   h. On **Review the role group and finish**, review the role group changes. Select **Save** to save the changes to the eDiscovery role groups.

> **Note**
>
> You can also use the **Add-eDiscoveryCaseAdmin** cmdlet to make a user an eDiscovery Administrator. However, you must assign the *Case Management* role to the user before you can use this cmdlet to make them an eDiscovery Administrator. For more information, see Add-eDiscoveryCaseAdmin.

On **Permissions** in the Microsoft Purview portal, you can also assign users eDiscovery-related permissions by adding them to the *Compliance Administrator*, *Organization Management*, and *Reviewer* role groups. For a description of the eDiscovery-related role-based access control roles assigned to each of these role groups, see Role-based access control roles related to eDiscovery.

View assigned permissions

When eDiscovery users sign in to Discovery in the Microsoft Purview portal, the **Your permissions** card on the eDiscovery **Overview** tab displays their assigned permissions. This card outlines the user's access and roles, including any limitations to eDiscovery cases.



RBAC roles related to eDiscovery

The following table lists the eDiscovery-related role-based access control roles in the Microsoft Purview portal, and shows the built-in role groups that each role belongs to by default.

Expand table

Role	Compliance Administrator	eDiscovery Manager & Administrator	Organization Management	Reviewer
Case Management	✓	✓	✓	
Communication		✓		
Compliance Search	✓	✓	✓	
Custodian		✓		
Export		✓		
Hold	✓	✓	✓	
Manage review set tags		✓		
Preview		✓		

Role	Compliance Administrator	eDiscovery Manager & Administrator	Organization Management	Reviewer
Review		✓		✓
RMS Decrypt		✓		
Search And Purge			✓	

> **Note**
>
> To view the list of eDiscovery cases, a user must have at least one of the roles listed in the previous table. For users who aren't eDiscovery Administrators, the cases shown are limited to those where the user is a member.

Run the following diagnostic test to check if the *Export*, *Preview*, or *Search* roles are assigned to the designated admin account.

1. Select the **Help** control in the top right of the Microsoft Purview portal. Enter *Diag:edisRBACdiag* in search (or select the eDiscovery RBAC Check diagnostic) to run the **eDiscovery RBAC Check** test.
2. In the **Run diagnostics** section, enter the UPN or email address of the user trying to run an export, preview, or search task.
3. Select **Run Tests**. If the user doesn't have the necessary eDiscovery roles, assign the roles to perform the desired task.

The following sections describe each of the eDiscovery-related role-based access control roles listed in the previous table.

Case Management role

This role lets users create, edit, delete, and control access to eDiscovery (Premium) cases in the Microsoft Purview portal. You must assign the *Case Management* role before you can use the **Add-eDiscoveryCaseAdmin** cmdlet to make a user an eDiscovery Administrator.

For more information, see Get started with eDiscovery (Premium).

Communication role

This role lets users manage all communications with the custodians identified in an eDiscovery (Premium) case. This management includes creating hold notifications, hold reminders, and

escalations to management. The user can also track custodian acknowledgment of hold notifications and manage access to the custodian portal that each custodian uses to track communications for the cases where they were identified as a custodian.

For more information, see Work with communications in eDiscovery (Premium).

Compliance Search role

This role lets users run the Content Search tool in the Microsoft Purview portal to search mailboxes and public folders, SharePoint Online sites, OneDrive for Business sites, Skype for Business conversations, Microsoft 365 groups, Microsoft Teams, and Viva Engage groups. This role allows a user to get an estimate of the search results and create export reports, but other roles are needed to initiate content search actions such as previewing, exporting, or deleting search results.

In Content search and eDiscovery (Standard), users who are assigned the *Compliance Search* role but don't have the *Preview* role can preview the results of a search when a user assigned the *Preview* role initiates the preview action. The user without the *Preview* role can preview results for up to two weeks after the initial preview action was created.

Similarly, users in Content search and eDiscovery (Standard) who are assigned the *Compliance Search* role but don't have the *Export* role can download the results of a search when a user assigned the *Export* role initiates the export action. The user without the *Export* role can download the results of a search for up to two weeks after the initial export action was created. After that, they can't download the results unless someone with the *Export* role restarts the export.

The two-week grace period for previewing and exporting search results (without the corresponding search and export roles) doesn't apply to eDiscovery (Premium). Users must be assigned the *Preview* and *Export* roles to preview and export content in eDiscovery (Premium).

Custodian role

This role lets users identify and manage custodians for eDiscovery (Premium) cases and use the information from Microsoft Entra ID and other sources to find data sources associated with custodians. The user can associate other data sources such as mailboxes, SharePoint sites, and Teams with custodians in a case. The user can also place a legal hold on the data sources associated with custodians to preserve content in the context of a case.

For more information, see Work with custodians in eDiscovery (Premium).

Export role

This role lets users export the results of a Content Search to a local computer. It also lets them prepare search results for analysis in eDiscovery (Premium).

For more information about exporting search results, see Export search results from Microsoft Purview portal.

Hold role

This role lets users place content on hold in mailboxes, public folders, sites, Skype for Business conversations, and Microsoft 365 groups. When content is on hold, content owners can still modify or delete the original content, but the content is preserved until the hold is removed or until the hold duration expires.

For more information about holds, see:

- Create a hold in eDiscovery (Standard)
- Create a hold in eDiscovery (Premium)

Manage review set tags role

This role lets users create, edit, and delete review set tags for cases they can access. Users need to at least have the *Review* role and this role to manage tags during reviews.

Preview role

This role lets users view a list of items that a Content Search returns. They can also open and view each item from the list to view its contents.

Review role

This role lets users access review sets in eDiscovery (Premium). Users who are assigned this role can see and open the list of cases on the **eDiscovery** > **Premium** page in the Microsoft Purview portal that they're members of. After the user accesses an eDiscovery (Premium) case, they can select **Review sets** to access case data. This role doesn't allow the user to preview the results of a collection search that's associated with the case or do other search or case management tasks. Users with this role can only access the data in a review set.

RMS Decrypt role

This role lets users view rights-protected email messages when previewing search results and export decrypted rights-protected email messages. This role also lets users view and export a file that's encrypted with a Microsoft encryption technology when the encrypted file is attached to an email message that's included in the results of an eDiscovery search. Additionally, this role lets users review and query encrypted email attachments that are added to a review set in eDiscovery (Premium). For more information about decryption in eDiscovery, see Decryption in Microsoft 365 eDiscovery tools.

Search And Purge role

This role lets users perform bulk removal of data matching the criteria of a content search. For more information, see Search for and delete email messages in your organization.

Add role groups as members of eDiscovery cases

You can add role groups as members of eDiscovery (Standard) and eDiscovery (Premium) cases so that members of the role groups can access and perform tasks in the assigned cases. The roles assigned to the role group define what members of the role group can do. Adding a role group as a member of the case lets members access and perform those tasks in a specific case. For more information about adding role groups as members of cases, see:

- Get started with eDiscovery (Standard)
- Add or remove members from an eDiscovery (Premium) case

If you add or remove a role from a role group, the role group is automatically removed as a member of any case it belongs to. This behavior protects your organization from inadvertently providing extra permissions to members of a case. Similarly, if you delete a role group, you remove it from all cases it was a member of.

Before you add or remove roles to a role group that might be a member of an eDiscovery case, run the following commands in Security & Compliance PowerShell to get a list of cases the role group is a member of. After you update the role group, add the role group back as a member of those cases.

Get a list of eDiscovery (Standard) cases a role group is assigned to

PowerShell

```
Get-ComplianceCase -RoleGroup "Name of role group"
```

Get a list of eDiscovery (Premium) cases a role group is assigned to

PowerShell

```
Get-ComplianceCase -RoleGroup "Name of role group" -CaseType AdvancedEdiscovery
```

eDiscovery Administrator considerations

- **Why create an eDiscovery Administrator?** An eDiscovery Administrator is a member of the *eDiscovery Manager* role group who can view and access all eDiscovery cases in your organization. This ability to access all the eDiscovery cases has two important purposes:
  - If a person who is the only member of an eDiscovery case leaves your organization, no one (including members of the *Organization Management* role group or another member of the *eDiscovery Manager* role group) can access that eDiscovery case because they aren't a member of a case. In this situation, there's no way to access the data in the case. But because an eDiscovery Administrator can access all eDiscovery cases in the organization, they can view the case and add themselves or another eDiscovery manager as a member of the case.
  - Because an eDiscovery Administrator can view and access all eDiscovery (Standard) and eDiscovery (Premium) cases, they can audit and oversee all cases and associated compliance searches. This functionality can help to prevent any misuse of compliance searches or eDiscovery cases. And because eDiscovery Administrators can access potentially sensitive information in the results of a compliance search, you should limit the number of people who are eDiscovery Administrators.

- **Can I add a group as a member of the eDiscovery Manager role group?** As described in the Prerequisites section, you can add a mail-enabled security group as a member of the eDiscovery Managers subgroup in the *eDiscovery Manager* role group by using the **Add-RoleGroupMember** cmdlet in Security & Compliance PowerShell. For example, you can run the following command to add a mail-enabled security group to the *eDiscovery Manager* role group.

  PowerShell

  ```
  Add-RoleGroupMember "eDiscoveryManager" -Member <name of security group>
  ```

  Exchange distribution groups and Microsoft 365 Groups aren't supported. You must use a mail-enabled security group, which you can create in Exchange Online PowerShell by

running `New-DistributionGroup -Type Security`. You can also create a mail-enabled security group (and add members) in the Exchange admin center or in the Microsoft 365 admin center. It might take up to 60 minutes after you create it for a new mail-enabled security group to be available to add to the eDiscovery Managers role group.

You can't make a mail-enabled security group an eDiscovery Administrator by using the **Add-eDiscoveryCaseAdmin** cmdlet in Security & Compliance PowerShell. You can only add individual users as eDiscovery Administrators.

You also can't add a mail-enabled security group as a member of a case.

Last updated on 08/28/2025

Set up compliance boundaries for eDiscovery investigations

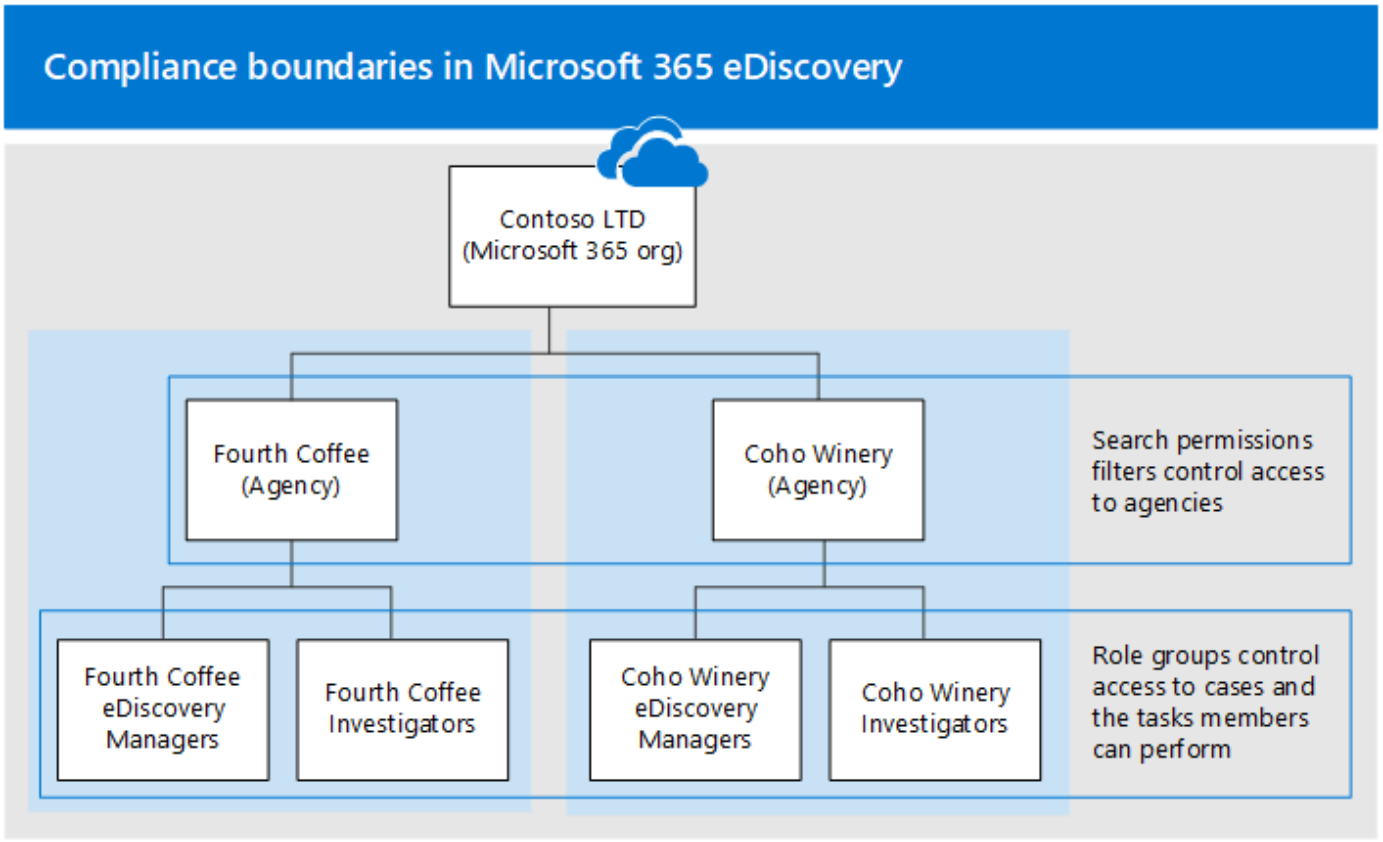
> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

The guidance in this article can be applied when using either Microsoft Purview eDiscovery (Standard) or Microsoft Purview eDiscovery (Premium) to manage investigations.

Compliance boundaries create logical boundaries within an organization that control the user content locations (such as mailboxes, OneDrive accounts, and SharePoint sites) that eDiscovery managers can search. Compliance boundaries also control who can access eDiscovery cases used to manage the legal, human resources, or other investigations within your organization.

The need for compliance boundaries is often necessary for multi-national corporations that have to respect geographical boarders and regulations and for governments, which are often divided into different agencies. In Microsoft 365, compliance boundaries help you meet these requirements when performing content searches and managing investigations with eDiscovery cases.

We'll use the example in the following illustration to explain how compliance boundaries work.



In this example, Contoso LTD is an organization that consists of two subsidiaries, Fourth Coffee and Coho Winery. The business requires that eDiscovery managers and investigators can only search the Exchange mailboxes, OneDrive accounts, and SharePoint sites in their agency. Also, eDiscovery managers and investigators can only see eDiscovery cases in their agency, and they can only access the cases that they're a member of. Additionally in this scenario, investigators can't place content locations on hold or export content from a case. Here's how compliance boundaries meet these requirements.

- The search permissions filtering functionality for eDiscovery controls the content locations that eDiscovery managers and investigators can search. This control means eDiscovery managers and investigators in the Fourth Coffee agency can only search content locations in the Fourth Coffee subsidiary. The same restriction applies to the Coho Winery subsidiary.

- Role groups provide the following functions for compliance boundaries:
  - Control who can see the eDiscovery cases in the Microsoft Purview portal. This control means that eDiscovery managers and investigators can only see the eDiscovery cases in their agency.
  - Control who can assign members to an eDiscovery case. This control means eDiscovery managers and investigators can only assign members to cases that they themselves are a member of.
  - Control the eDiscovery-related tasks that members can perform by adding or removing roles that assign specific permissions.

- When a search permissions filter is applied to a role group, members of the role group can perform the following search-related actions as long as the permissions to perform an action is assigned to the role group:
  - Search for content
  - Preview search results
  - Export search results
  - Purge items returned by a search

Before you set up compliance boundaries

- Users must be assigned an Exchange Online license. To verify assignments, use the Get-User cmdlet in Exchange Online PowerShell.

Setting up compliance boundaries

Here's the steps for setting up compliance boundaries:

- Step 1: Identify a user attribute to define your agencies
- Step 2: Create a role group for each agency
- Step 3: Create a search permissions filter to enforce the compliance boundary
- Step 4: Create an eDiscovery case for an intra-agency investigations

Step 1: Identify a user attribute to define your agencies

The first step is to choose an attribute to use that will define your agencies. This attribute is used to create the search permissions filter that limits an eDiscovery manager to search only the content locations of users who are assigned a specific value for this attribute. For example, let's say Contoso decides to use the **Department** attribute. The value for this attribute for users in the Fourth Coffee subsidiary would be `FourthCoffee` and the value for users in Coho Winery subsidiary would be `CohoWinery`. In Step 3, you use this `attribute:value` pair (for example, *Department:FourthCoffee*) to limit the user content locations that eDiscovery managers can search.

Here are some examples of user attributes that you can use for compliance boundaries:

- Company
- CustomAttribute1 - CustomAttribute15
- Department
- Office

- CountryOrRegion (Two-letter country code)

Step 2: Create a role group for each agency

The next step is to create the role groups in the Microsoft Purview portal that will align with your agencies.

To create the role groups, go to the **Permissions** page in the Microsoft Purview portal and create a role group for each team in each agency that will use compliance boundaries and eDiscovery cases to manage investigations.

We recommend that the role groups created for the compliance boundary don't have any roles attached to it. This role group should only be used to assign users to the role group. Separate built-in (eDiscovery Manager) or custom role groups should be used to assign roles to members. For more information about eDiscovery-related roles, see Assign eDiscovery permissions.

> **Note**
>
> To ceate and update role groups with empty roles, you must use PowerShell cmdlets. For details, see New-RoleGroup and Add-RoleGroupMember.

Using the Contoso compliance boundaries scenario, four role groups need to be created and the appropriate members added to each one.

- Fourth Coffee eDiscovery Managers
- Fourth Coffee Investigators
- Coho Winery eDiscovery Managers
- Coho Winery Investigators

> **Important**
>
> If a role is added or removed from a role group that you've added as a member of a case, then the role group is automatically removed as a member of the case (or any case the role group is a member of). The reason for this is to protect your organization from inadvertently providing additional permissions to members of a case. If a role group is deleted, it is removed from all cases it was a member of. We recommend that the role groups created for compliance boundaries don't have any roles assigned to them. Use separate built-in/custom role groups to assign roles to members.

Step 3: Create a search permissions filter to enforce the compliance boundary

After you've created role groups for each agency, the next step is to create the search permissions filters that associate each role group to its specific agency and defines the compliance boundary itself. You need to create one search permissions filter for each agency. For more information about creating security permissions filters, see Configure permissions filtering for Content Search.

Here's the syntax that's used to create a search permissions filter used for compliance boundaries for the scenario in this article.

PowerShell

```
New-ComplianceSecurityFilter -FilterName <name of filter> -Users <role groups> -
Filters "Mailbox_<MailboxPropertyName>  -eq '<Value> '", "SiteContent_Path -like
'<SharePointURL>' -or SiteContent_Path -like '<OneDriveURL>'"
```

Here's a description of each parameter in the command:

- `FilterName`: Specifies the name of the filter. Use a name that describes or identifies the agency that the filter is used in.
- `Users`: Specifies the users or groups who get this filter applied to the search actions they perform. For compliance boundaries, this parameter specifies the role groups (that you created in Step 2) in the agency that you're creating the filter for. This parameter is a multi-value parameter so you can include one or more role groups, separated by commas.
- `Filters`: Specifies the search criteria for the filter. For compliance boundaries, you define the following filters. Each one applies to different content locations.
  - `Mailbox`: Specifies the mailboxes or OneDrive accounts that the role groups defined in the `Users` parameter can search. This filter allows members of the role group to search only the mailboxes or OneDrive accounts in a specific agency; for example, `"Mailbox_Department -eq 'FourthCoffee'"`.
  - `SiteContent`: This filter includes two separate filters. The first `SiteContent_Path` specifies the SharePoint sites in the agency that the role groups defined in the `Users` parameter can search. For example, `SiteContent_Path -like 'https://contoso.sharepoint.com/sites/FourthCoffee*'`. The second `SiteContent_Path` filter (connected to the first `SiteContent_Path` filter by the `or` operator) specifies the

agency's OneDrive domain (also called the *MySite* domain). For example, `SiteContent_Path -like 'https://contoso-my.sharepoint.com/personal*'`. You can also use the `Site_Path` filter in place of the `SiteContent` filter. The `Site` and `SiteContent` filters are interchangeable, and don't affect search permissions filters described in this article.

> **Important**
>
> Why is the `SiteContent` filter for OneDrive included in the previous search permissions filter? Although the `Mailbox` filter applies to *both* mailboxes and OneDrive accounts, the inclusion of the SharePoint filter would exclude OneDrive accounts if you didn't also include the OneDrive `Site` filter. If the search permissions filter didn't include a SharePoint filter, then you wouldn't have to include a separate OneDrive filter because the Mailbox filter would include OneDrive accounts in the scope of the compliance boundary. In other words, a search permissions filter with only the `Mailbox` filter would include both mailboxes and OneDrive accounts.

Here are examples of the two search permissions filters that would be created to support the Contoso compliance boundaries scenario. Both of these examples include a comma-separated filters list, in which the mailbox and site filters are included in the same search permissions filter and are separated by a comma.

Fourth Coffee

PowerShell

```
New-ComplianceSecurityFilter -FilterName "Fourth Coffee Security Filter" -Users
"Fourth Coffee eDiscovery Managers", "Fourth Coffee Investigators" -Filters
"Mailbox_Department -eq 'FourthCoffee'", "SiteContent_Path -like
'https://contoso.sharepoint.com/sites/FourthCoffee*' -or SiteContent_Path -like
'https://contoso-my.sharepoint.com/personal*'"
```

Coho Winery

PowerShell

```
New-ComplianceSecurityFilter -FilterName "Coho Winery Security Filter" -Users "Coho
Winery eDiscovery Managers", "Coho Winery Investigators" -Filters "Mailbox_Department
-eq 'CohoWinery'", "SiteContent_Path -like
'https://contoso.sharepoint.com/sites/CohoWinery*' -or SiteContent_Path -like
'https://contoso-my.sharepoint.com/personal*'"
```

> **Note**
>
> The syntax for the `Filters` parameters in the previous examples includes a *filters list*. A filters list is a filter that includes a mailbox filter and a site path filter separated by a comma. In the previous example, notice that a comma separates `Mailbox` and `SiteContent` filters: `-Filters "Mailbox_<MailboxPropertyName> -eq '<Value> '", "SiteContent_Path -like '<SharePointURL>' -or SiteContent_Path -like '<OneDriveURL>'"`. When this filter is processed during the running of an eDiscovery search, two search permissions filters are created from the filters list: one mailbox filter and one SharePoint/OneDrive filter. An alternative to using a filters list would be to create two separate search permissions filters for each agency: one search permissions filter for the mailbox attribute and one filter for the SharePoint and OneDrive site attributes. In either case, the results will be the same. Using a filters list or creating separate search permissions filters is a matter of preference.

How do the search permissions filters work in this scenario?

Here's how the search permission filters are applied for each agency in this scenario.

1. The `Mailbox` filter is first applied to define the content locations that eDiscovery managers can search. In this case, Coho Winery eDiscovery managers can only search the mailboxes and OneDrive accounts of users whose *Department* mailbox property has a value of **FourthCoffee**; Coho Winery eDiscovery managers can only search the mailboxes and OneDrive accounts of users whose *Department* mailbox property has a value of **CohoWinery**. The `Mailbox` filter is a *content location filter*, because it specifies the content locations that eDiscovery managers can search. In both filters, eDiscovery managers can only search content locations with a specific mailbox property value.

2. After the content locations that can be searched are defined, the next part of the filter defines the content that eDiscovery managers can search. The first `SiteContent` filter lets Fourth Coffee eDiscovery managers only search for documents that have a site path property that contains (or starts with) `https://contoso.sharepoint.com/sites/FourthCoffee`; Coho Winery eDiscovery managers can only search documents that have a site path property that contains (or starts with) `https://contoso.sharepoint.com/sites/CohoWinery`. Therefore, the two `SiteContent` filters are *content filters* because they define the content that can be searched for. In both filters, eDiscovery managers can only search for documents with a specific document property value. All SharePoint-related filters are content filters because searchable site properties are stamped on all documents. For more information, see Configure permissions filtering for eDiscovery.

> **Note**
>
> Although the scenario in this article doesn't use them, you can also use mailbox content filters to specify the content that eDiscovery managers can search for. The syntax for mailbox content filters is `"MailboxContent_<property> -<comparison operator> '<value>'"`. You can create content filters based on date ranges, recipients, and domains or any searchable email property. For example, this filter would allow eDiscovery managers to only search for mail items sent or received by users in the contoso.com domain: `"MailboxContent_Participants -like 'contoso.com'"`. For more information about mailbox content filters, see Configure search permissions filtering.

3. The search permissions filter is joined to the search query by the **AND** Boolean operator. That means when an eDiscovery manager in one of the agencies runs an eDiscovery search, the items returned by the search must match the search query and the conditions defined in the search permissions filter.

Step 4: Create an eDiscovery case for intra-agency investigations

The final step is to create an eDiscovery (Standard) case or eDiscovery (Premium) case in the Microsoft Purview portal and then add the role group that you created in Step 2 as a member of the case. This results in two important characteristics of using compliance boundaries:

- Only members of the role group added to the case will be able to see and access the case in the Microsoft Purview portal. For example, if the Fourth Coffee Investigators role group is the only member of a case, then members of the Fourth Coffee eDiscovery Managers role group (or members of any other role group) won't be able to see or access the case.
- When a member of the role group assigned to a case runs a search associated with the case, they'll only be able to search the content locations within their agency (which is defined by the search permissions filter that you created in Step 3.)

To create a case and assign members:

1. Go to the **eDiscovery (Standard)** or **eDiscovery (Premium)** page in the Microsoft Purview portal and create a case.
2. In the list of cases, select the name of the case you created.

3. Add role groups as members to the case. For instructions, see the one of the following articles:

    - Add members to an eDiscovery (Standard) case
    - Add members to an eDiscovery (Premium) case

> **Note**
>
> When adding a role group to a case, you can only add the role groups that you are a member of.

Searching and exporting content in multi-geo environments

Search permissions filters also let you control where content is routed for export and which datacenter can be searched when searching content locations in a SharePoint Multi-Geo environment.

- **Export search results:** You can export the search results from Exchange mailboxes, SharePoint sites, and OneDrive accounts from a specific datacenter. This means that you can specify the datacenter location that search results are exported from.

  Use the *Region* parameter for **New-ComplianceSecurityFilter** or **Set-ComplianceSecurityFilter** cmdlets to create or change which datacenter the export is routed through.

  Expand table

Parameter value	Datacenter location
NAM	North American (datacenters are in the US)
EUR	Europe
APC	Asia Pacific
CAN	Canada

- **Route content searches:** You can route the content searches of SharePoint sites and OneDrive accounts to a satellite datacenter. This means you can specify the datacenter location where searches are run.

Use one of the following values for the *Region* parameter to control the datacenter location that searches will run in when searching SharePoint sites and OneDrive accounts.

Expand table

Parameter value	Datacenter routing locations for SharePoint
NAM	US
EUR	Europe
APC	Asia Pacific
CAN	US
AUS	Asia Pacific
KOR	The organization's default datacenter
GBR	Europe
JPN	Asia Pacific
IND	Asia Pacific
LAM	US
NOR	Europe
BRA	North American datacenters

If you don't specify the *Region* parameter for a search permissions filter, the organization's primary SharePoint region is searched. Search results are exported to the closest datacenter.

To simplify the concept, the *Region* parameter controls the datacenter that is used to search for content in SharePoint and OneDrive. This doesn't apply to searching for content in Exchange because Exchange content searches aren't bound by the geographic location of datacenters. Also, the same *Region* parameter value may also dictate the datacenter that exports are routed through. This is often necessary to control the movement of data across geographic boarders.

Note

If you're using eDiscovery (Premium), the *Region* parameter doesn't control the region that data is exported from. Data is exported from the organization's central location. Also, searching for content in SharePoint and OneDrive isn't bound by the geographic location of

datacenters. All datacenters are searched. For more information about eDiscovery (Premium), see Overview of the eDiscovery (Premium) solution in Microsoft 365.

Here are examples of using the *Region* parameter when creating search permission filters for compliance boundaries. This assumes that the Fourth Coffee subsidiary is located in North America and that Coho Winery is in Europe.

PowerShell

```
New-ComplianceSecurityFilter -FilterName "Fourth Coffee Security Filter" -Users
"Fourth Coffee eDiscovery Managers", "Fourth Coffee Investigators" -Filters
"Mailbox_Department -eq 'FourthCoffee'", "SiteContent_Path -like
'https://contoso.sharepoint.com/sites/FourthCoffee*' -or SiteContent_Path -like
'https://contoso-my.sharepoint.com/personal*'" -Region NAM
```

PowerShell

```
New-ComplianceSecurityFilter -FilterName "Coho Winery Security Filter" -Users "Coho
Winery eDiscovery Managers", "Coho Winery Investigators" -Filters "Mailbox_Department
-eq 'CohoWinery'", "SiteContent_Path -like
'https://contoso.sharepoint.com/sites/CohoWinery*' -or SiteContent_Path -like
'https://contoso-my.sharepoint.com/personal*'" -Region EUR
```

Keep the following things in mind when searching and exporting content in multi-geo environments.

- The *Region* parameter doesn't control searches of Exchange mailboxes. All datacenters will be searched when you search mailboxes. To limit the scope of which Exchange mailboxes are searched, use the *Filters* parameter when creating or changing a search permissions filter.
- If it's necessary for an eDiscovery Manager to search across multiple SharePoint regions, you need to create a different user account for that eDiscovery manager to use in the search permissions filter to specify the region where the SharePoint sites or OneDrive accounts are located. For more information about setting this up, see the "Searching for content in a SharePoint Multi-Geo environment" section in Content Search.
- When searching for content in SharePoint and OneDrive, the *Region* parameter directs searches to either the primary or satellite location where the eDiscovery manager will conduct eDiscovery investigations. If an eDiscovery manager searches SharePoint and OneDrive sites outside of the region that's specified in the search permissions filter, no search results are returned.

- When exporting search results from eDiscovery (Standard), content from all content locations (including Exchange, Skype for Business, SharePoint, OneDrive, and other services that you can search by using the Content Search tool) are uploaded to the Azure Storage location in the datacenter that's specified by the *Region* parameter. This helps organizations stay within compliance by not allowing content to be exported across controlled borders. If no region is specified in the search permissions filter, content is uploaded to the organization's primary datacenter.

  When exporting content from eDiscovery (Premium), you can't control where content is uploaded by using the *Region* parameter. Content is uploaded to an Azure Storage location in a datacenter in your organization's central location. For a list of geo locations based on your central location, see Microsoft 365 Multi-Geo eDiscovery configuration.

- You can edit an existing search permissions filter to add or change the region by running the following command:

  PowerShell

  ```
  Set-ComplianceSecurityFilter -FilterName <Filter name>  -Region <Region>
  ```

Using compliance boundaries for SharePoint hub sites

SharePoint hub sites often align with the same geographical or agency boundaries that eDiscovery compliance boundaries follow. That means you can use the site ID property of the hub site to create a compliance boundary. To do this, use the Get-SPOHubSite cmdlet in SharePoint Online PowerShell to obtain the SiteId for the hub site, and then use this value for the department ID property to create a search permissions filter.

Use the following syntax to create a search permissions filter for a SharePoint hub site:

PowerShell

```
New-ComplianceSecurityFilter -FilterName <Filter Name> -Users <User or Group> -Filters
"Site_Departmentid -eq '{SiteId of hub site}'"
```

Here's an example of creating a search permissions filter for a hub site for the Coho Winery agency:

PowerShell

```
New-ComplianceSecurityFilter -FilterName "Coho Winery Hub Site Security Filter" -Users
"Coho Winery eDiscovery Managers", "Coho Winery Investigators" -Filters
"Site_Departmentid -eq '44252d09-62c4-4913-9eb0-a2a8b8d7f863'"
```

Compliance boundary limitations

Keep the following limitations in mind when managing eDiscovery cases and investigations that use of compliance boundaries.

- When creating and running a search, you can select content locations that are outside of your agency. However, because of the search permissions filter, content from those locations isn't included in the search results.
- Compliance boundaries don't apply to holds in eDiscovery cases. That means an eDiscovery manager in one agency can place a user in a different agency on hold. However, the compliance boundary will be enforced if the eDiscovery manager searches the content locations of the user who was placed on hold. That means the eDiscovery manager won't be able search the user's content locations, even though they were able to place the user on hold.
- If you're assigned a search permissions filter (either a mailbox or a site filter) and you try to export unindexed items for a search that includes all SharePoint sites in your organization, you'll receive the following error message: `Unable to execute the task. Reason: The scope options UnindexedItemsOnly or BothIndexedandUnindexedItems are not allowed when the executing user has a compliance security filter applied`. If you're assigned a search permissions filter and you want to export unindexed items from SharePoint, you'll have to rerun the search and include specific SharePoint sites to search. Otherwise, you'll only be able to export indexed items from a search that includes all SharePoint sites. For more information about the options when you export search results, see Export Content search results.
- Search permissions filters aren't applied to Exchange public folders.
- Supported search filters include `-and`, `-or`, `-like`, `-not`, `-eq`, and `-ne`. We don't recommend using other exclusion filters (such as using `-notlike)`, `inotlike`, etc. in a search permissions filter) for a content-based compliance boundary. Using these exclusion filters can have unexpected results if content with recently updated attributes hasn't been indexed.
- Respect for compliance boundaries for OneDrive sites in searches may be impacted when:

- The sites (or associated mailboxes) are moved or re-homed
- OneDrive ownership is re-assigned, or more than one owner is added
- The OneDrive site is renamed/re-issued

Moving a OneDrive site can change the ownership of the content and may include creating an arbitration mailbox. When these resources are moved, there's a chance that content search results may be available across compliance boundaries. When a OneDrive site is re-assigned, re-named or assigned to more than one owner, it's possible that the content from these sites may be available across compliance boundaries.

- Compliance boundaries for mailboxes may be impacted when:
- The mailbox isn't associated with a licensed user (includes disabled or deleted users)
  - A mailbox isn't managed or synced from Microsoft Entra ID

    Additionally, there are several types of mailboxes that may produce content during search, regardless of their compliance boundaries. These mailbox types include:
    - EquipmentMailbox
    - GuestMailUser
    - LinkedMailbox
    - RoomList
    - RoomMailbox
    - SchedulingMailbox
    - SharedMailbox
    - SystemMailbox
    - TeamMailbox

  To make sure that the search respects the compliance boundaries as expected, users can set attributes on shared mailboxes, equipment mailboxes, and other mailboxes using the Set-Mailbox cmdlet or by using Invoke-ComplianceSecurityFilterAction to ensure the attribute is set correctly. If the attribute is properly configured and aligned with the compliance boundary attribute, the mailbox respects the compliance boundary.

More information

- If a mailbox is de-licensed or soft-deleted, the user will no longer be considered within the compliance boundary. If a hold was placed on the mailbox when it was deleted, the content preserved in the mailbox is still subject to a compliance boundary or search permissions filter.

- If compliance boundaries and search permissions filters are implemented for a user, then we recommend that you don't delete a user's mailbox and not their OneDrive account. In other words, if you delete a user's mailbox, you should also remove the user's OneDrive account since mailbox_RecipientFilter is used to enforce search permission filter for OneDrive.
- Compliance boundaries and search permissions filters depend on attributes being stamped on content in Exchange, OneDrive, and SharePoint and the subsequent indexing of this stamped content.

Frequently asked questions

Who can create and manage search permissions filters (using New-ComplianceSecurityFilter and Set-ComplianceSecurityFilter cmdlets)?

To create, view, and modify search permissions filters, you have to be a member of the Organization Management role group in the Microsoft Purview portal.

If an eDiscovery manager is assigned to more than one role group that spans multiple agencies, how do they search for content in one agency or the other?

The eDiscovery manager can add parameters to their search query that restrict the search to a specific agency. For example, if an organization has specified the **CustomAttribute10** property to differentiate agencies, they can append the following to their search query to search mailboxes and OneDrive accounts in a specific agency: `CustomAttribute10:<value>`.

What happens if the value of the attribute that's used as the compliance attribute in a search permissions filter is changed?

It takes up to three days for a search permissions filter to enforce the compliance boundary if the value of the attribute that's used in the filter is changed. For example, in the Contoso scenario let's say that a user in the Fourth Coffee agency is transferred to the Coho Winery agency. As a result, the value of the **Department** attribute on the user object is changed from *FourthCoffee* to *CohoWinery*. In this situation, Fourth Coffee eDiscovery and investors will get search results for that user for up three days after the attribute is changed. Similarly, it takes up to three days before Coho Winery eDiscovery managers and investigators get search results for the user.

Can an eDiscovery manager see content from two separate compliance boundaries?

Yes, this can be done when searching Exchange mailboxes by adding the eDiscovery manager to role groups that have visibility to both agencies. However when searching SharePoint sites and OneDrive accounts, an eDiscovery manager can search for content in different compliance

boundaries only if the agencies are in the same region or geo location. **Note:** This limitation for sites doesn't apply in eDiscovery (Premium) because searching for content in SharePoint and OneDrive isn't bound by geographic location.

Do search permissions filters work for eDiscovery case holds, Microsoft 365 retention policies, or DLP?

No, not at this time.

If I specify a region to control where content is exported, but I don't have a SharePoint organization in that region, can I still search SharePoint?

If the region specified in the search permissions filter doesn't exist in your organization, the default region is searched.

What is the maximum number of search permissions filters that can be created in an organization?

There's no limit to the number of search permissions filters that can be created in an organization. However, a search query can have a maximum of 100 conditions. In this case, a condition is defined as something that's connected to the query by a Boolean operator (such as **AND**, **OR**, and **NEAR**). The limit of the number of conditions includes the search query itself plus all search permissions filters that are applied to the user who runs the search. Therefore, the more search permissions filters you have (especially if these filters are applied to the same user or group of users), the better the chance of exceeding the maximum number of conditions for a search.

To understand how this limit works, you need to understand that a search permissions filter is appended to the search query when a search is run. A search permissions filter is joined to the search query by the **AND** Boolean operator. The query logic for the search query and a single search permissions filter would look like this:

```text
<SearchQuery> AND <PermissionsFilter>
```

Multiple search permissions filters are combined together by the **OR** Boolean operator, and then those conditions are connected to the search query by the **AND** operator.

The query logic for the search query and multiple search permissions filters would look like this:

text

```
<SearchQuery> AND (<PermissionsFilter1> OR <PermissionsFilter2> OR
<PermissionsFilter3>...)
```

It's possible the search query itself may consist of multiple conditions connected by Boolean operators. Each condition in the search query would also count against the 100-condition limit.

Also, the number of search permissions filters appended to a query depends on the user who is running the search. When a specific user runs a search, the search permissions filters that are applied to the user (which is defined by the *Users* parameter in the filter) are appended to the query. Your organization could have hundreds of search permissions filters, but if more than 100 filters are applied to the same users, then it's likely the 100-condition limit will be exceeded when those users run searches.

There's one more thing to keep in mind about the condition limit. The number of specific SharePoint sites that are included in the search query or search permissions filters also count against this limit.

To prevent your organization from reaching the conditions limit, keep the number of search permissions filters in your organization to few as possible to meet your business requirements.

Last updated on 10/31/2024

Configure permissions filtering for eDiscovery

> **⊗ Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

You can use search permissions filtering to let an eDiscovery manager search only a subset of mailboxes and sites in your organization. You can also use permissions filtering to let that same eDiscovery manager search only for mailbox or site content that meets a specific search criteria.

For example, you might want an eDiscovery manager to search only the mailboxes of users in a specific location or department. You do this by creating a filter that uses a supported recipient filter to limit which mailboxes a specific user or group of users can search. You can also create a filter that specifies what mailbox content a user can search for. This is done by creating a filter that uses a searchable message property. Similarly, you can let an eDiscovery manager search only specific SharePoint sites in your organization. You do this by creating a filter that limits which site can be searched. You can also create a filter that specifies what site content can be searched. This is done by creating a filter that uses a searchable site property.

Search permissions filters are applied when you search for content using Content search, Microsoft Purview eDiscovery (Standard), and Microsoft Purview eDiscovery (Premium) in the Microsoft Purview portal. When a search permissions filter is applied to a specific user, that user can perform the following search-related actions:

- Search for content
- Preview search results
- Export search results
- Purge items returned by a search

You can also use search permissions filtering to create logical boundaries (called *compliance boundaries*) within an organization that control the user content locations (such as mailboxes, SharePoint sites, and OneDrive accounts) that specific eDiscovery managers can search. For more information, see Set up compliance boundaries for eDiscovery investigations.

The following four cmdlets in Security & Compliance PowerShell let you configure and manage search permissions filters:

- New-ComplianceSecurityFilter
- Get-ComplianceSecurityFilter
- Set-ComplianceSecurityFilter
- Remove-ComplianceSecurityFilter

Requirements to configure permissions filtering

- To run the compliance security filter cmdlets, you have to be a member of the *Organization Management* role group in the Microsoft Purview portal. For more information, see Permissions in the Security & Compliance Center.
- You have to connect to both Exchange Online and Security & Compliance PowerShell to use the compliance security filter cmdlets. This is necessary because these cmdlets require access to mailbox properties, which is why you have to connect to Exchange Online PowerShell. See the steps in the next section.
- See the More information section for additional information about search permissions filters.
- Search permissions filtering is applicable to inactive mailboxes, which means you can use mailbox and mailbox content filtering to limit who can search an inactive mailbox. See the More information section for additional information about permissions filtering and inactive mailboxes.
- Search permissions filtering can't be used to limit who can search public folders in Exchange.
- There's no limit to the number of search permissions filters that can be created in an organization. However, a search query can have a maximum of 100 conditions. In this case, a condition is defined as something that's connected to the query by a Boolean operator (such as **AND**, **OR**, and **NEAR**). The limit for the number of conditions includes the search query itself plus all search permissions filters that are applied to the user who runs the search. Therefore, the more search permissions filters you have (especially if these filters are applied to the same user or group of users), the better the chance of exceeding the maximum number of conditions for a search. To prevent your organization from reaching

the conditions limit, keep the number of search permissions filters in your organization to few as possible to meet your business requirements. For more information, see Set up compliance boundaries for eDiscovery investigations.

Connect to Exchange Online and Security & Compliance PowerShell in a single session

Before you can successfully run the script in this section, you have to download and install the Exchange Online PowerShell module. For information, see Install and maintain the Exchange Online PowerShell module.

1. Save the following text to a Windows PowerShell script file by using a filename suffix of **.ps1**. For example, you could save it to a file named **ConnectEXO-SCC.ps1**.

   PowerShell

   ```
   Import-Module ExchangeOnlineManagement
   $UserCredential = Get-Credential
   Connect-ExchangeOnline -Credential $UserCredential -ShowBanner:$false
   Connect-IPPSSession -Credential $UserCredential
   $Host.UI.RawUI.WindowTitle = $UserCredential.UserName + " (Exchange Online + Security & Compliance)"
   ```

2. On your local computer, open Windows PowerShell, go to the folder where the script that you created in the previous step is located, and then run the script; for example:

   PowerShell

   ```
   .\ConnectEXO-SCC.ps1
   ```

How do you know if this worked? After you run the script, cmdlets from Exchange Online PowerShell and Security & Compliance PowerShell are available. If you don't receive any errors, you connected successfully. A quick test is to run Exchange Online PowerShell and Security & Compliance PowerShell cmdlets. For example, you can run and **Get-Mailbox** and **Get-ComplianceSearch**.

For troubleshooting PowerShell connection errors, see:

- Connect to Exchange Online PowerShell
- Connect to Security & Compliance PowerShell

New-ComplianceSecurityFilter

The **New-ComplianceSecurityFilter** cmdlet is used to create a search permissions filter. Here's the basic syntax for this cmdlet:

```
PowerShell
New-ComplianceSecurityFilter -FilterName <name of filter> -Users <user or role group>
-Filters <filter>
```

The following sections describe the parameters for this cmdlet. All parameters are required to create a search permissions filter.

FilterName

The *FilterName* parameter specifies the name of the permissions filter. This name is used to identity a filter when using the **Get-ComplianceSecurityFilter**, **Set-ComplianceSecurityFilter,** and **Remove-ComplianceSecurityFilter** cmdlets.

Filters

The *Filters* parameter specifies the search criteria for the compliance security filter. You can create three different types of filters:

- **Mailbox or OneDrive filtering:** This type of filter specifies the mailboxes and OneDrive accounts the assigned users (specified by the *Users* parameter) can search. This type of filter is called a *content location* filter because it defines the content locations that a user can search. The syntax for this type of filter is **Mailbox_** *MailboxPropertyName*, where *MailboxPropertyName* specifies a mailbox property used to scope the mailboxes and OneDrive accounts that can be searched. For example, the mailbox filter `"Mailbox_CustomAttribute10 -eq 'OttawaUsers'"` would allow the user assigned this filter to search only the mailboxes and OneDrive accounts that have the value "OttawaUsers" in the CustomAttribute10 property.

  Any supported filterable recipient property can be used for the *MailboxPropertyName* property in a mailbox or OneDrive filter. The following table list four commonly used recipient properties used to create a mailbox or OneDrive filter. The table also includes an example of using the property in a filter.

Expand table

Property name	Example
Alias	`"Mailbox_Alias -like 'v-'"`
Company	`"Mailbox_Company -eq 'Contoso'"`
CountryOrRegion	`"Mailbox_CountryOrRegion -eq 'United States'"`
Department	`"Mailbox_Department -eq 'Finance'"`

> **Note**
>
> The Mailbox filter applies to both user and group mailboxes. The associated OneDrive site for the user mailbox and SharePoint sites associated to the group mailbox satisfying the mailbox_ filter are allowed.

- **Mailbox content filtering:** This type of filter is applied on the content that can be searched. This type of filter is called a *content filter* because it specifies the mailbox content or searchable email properties the assigned users can search for. The syntax for this type of filter is **MailboxContent**_*SearchablePropertyName*, where *SearchablePropertyName* specifies a Keyword Query Language (KeyQL) property that can be specified in a search. For example, the mailbox content filter `"MailboxContent_Recipients -like 'contoso.com'"` would allow the user assigned this filter to only search for messages sent to recipients in the contoso.com domain. For a list of searchable email properties, see Keyword queries and search conditions for eDiscovery.

  > **Important**
  >
  > A single search filter can't contain a mailbox filter and a mailbox content filter. To combine these in a single filter, you have to use a filters list. But a filter can contain a more complex query of the same type. For example, `"Mailbox_CustomAttribute10 -eq 'FTE' -and Mailbox_MemberOfGroup -eq '$($DG.DistinguishedName)'"`

- **Site and site content filtering:** There are two SharePoint- and OneDrive-related filters that you can use to specify what site or site content the assigned users can search.
  - **Site**_*QueryableSiteProperty*
  - **SiteContent**_*QueryableSiteProperty*

  These two filters are interchangeable. For example, `"Site_Path -like 'https://contoso.sharepoint.com/sites/doctors'"` and `"SiteContent_Path -like`

`'https://contoso.sharepoint.com/sites/doctors'"` return the same results. For a list of searchable site properties, see Keyword queries and search conditions for eDiscovery For a more complete list, see Overview of crawled and managed properties in SharePoint. Properties marked with a **Yes** in the **Queryable** column can be used to create a site or site content filter.

> **ⓘ Important**
>
> Setting up a site filter with one of the supported properties doesn't mean the site property in the filter will propagate to all documents on that site. This means the user is still responsible for populating the specific property fields associated with the documents on that site in order for the site filter to work and capture the right content. For example, if the user has a security filter "Site_RefineableString00 -eq 'abc'" applied and then the user runs a search using keyword query "xyz". The security filter gets appended to the query and the actual query running would be "xyz **AND RefineableString0:'abc'**". The user needs to ensure that documents on the site indeed have values in the RefineableString00 field as"abc". If not, the search query won't return any results.

Keep the following considerations in mind when configuring the *Filters* parameter for search permissions filters:

- Unlike mailboxes, there isn't a content location filter for sites even though the *Site* filter looks like a location filter. All filters for SharePoint and OneDrive are content filters (which is also why *Site_* and *SiteContent_* filters are interchangeable) because site-related properties like *Path* are stamped directly on the documents. Why is this? It's a result of the way that SharePoint is designed. In SharePoint, there isn't a "site object" with properties, like there is with Exchange mailboxes. Therefore, the *Path* property is stamped on the document and contains the URL of the site where the document is located. This is why a *Site* filter is considered a content filter and not a content location filter.
- You have to create a search permissions filter to explicitly prevent users from searching content locations in a specific service (such as preventing a user from searching any Exchange mailbox or any SharePoint site). In other words, creating a search permissions filter that allows a user to search all SharePoint sites in the organization doesn't prevent that user from searching mailboxes. For example, to allow SharePoint admins to only search SharePoint sites, you have to create a filter that prevents them from searching mailboxes. Similarly, to allow Exchange admins to only search mailboxes, you have to create a filter that prevents them from searching sites.

Users

The *Users* parameter specifies the users who get this filter applied to their searches. Identify users by their alias or primary SMTP address. You can specify multiple values separated by commas, or you can assign the filter to all users by using the value **All**.

You can also use the *Users* parameter to specify a Microsoft Purview portal role group. This lets you create a custom role group and then assign that role group a search permissions filter. For example, let's say you have a custom role group for eDiscovery managers for the U.S. subsidiary of a multi-national corporation. You can use the *Users* parameter to specify this role group (by using the Name property of the role group) and then use the *Filter* parameter to allow only mailboxes in the U.S. to be searched. You can't specify distribution groups with this parameter.

Using a filters list to combine filter types

A *filters list* is a filter that includes a mailbox filter and a site filter separated by a comma. This comma also functions as an **OR** operator. Using a filters list is the only supported method for combining different types of filters. In the following example, notice that a comma separates the **Mailbox** and **Site** filters:

PowerShell

```
-Filters "Mailbox_CustomAttribute10 -eq 'OttawaUsers'", "SiteContent_Path -like
'https://contoso.sharepoint.com/sites/doctors'"
```

When a filter that contains a filters list is processed during the running of a search, two search permissions filters are created from the filters list: One for each filter that's separated by a comma. So in the previous example, one mailbox search permissions filter and one site search permissions filter would be created. These filters are connected by the **OR** operator.

An alternative to using a filters list would be to create two separate search permissions filters. So in the previous example, you'd create one filter for the mailbox attribute and one filter for the site attribute. In either case, the results are the same. Using a filters list or creating separate search permissions filters is a matter of preference.

> ⓘ **Note**
>
> For OneDrive sites, the comma works as AND. For SharePoint sites, the comma works as OR.

Keep the following things in mind about using a filters list:

- You have to use a filters list to create a filter that includes a **Mailbox** filter and a **MailboxContent** filter.
- Each component of a filters list can contain a complex filter syntax. For example, the mailbox and site filters can contain multiple filters separated by an **-or** operator:

```
PowerShell
-Filters "Mailbox_Department -eq 'CohoWinery' -or Mailbox_CustomAttribute10 -eq
'CohoUsers'", "SiteContent_Path -like
'https://contoso.sharepoint.com/sites/CohoWinery*'"
```

Examples of creating search permissions filters

Here are examples of using the **New-ComplianceSecurityFilter** cmdlet to create a search permissions filter.

This example allows members of the "US Discovery Managers" role group to search only the mailboxes and OneDrive accounts in the United States.

```
PowerShell
New-ComplianceSecurityFilter -FilterName USDiscoveryManagers  -Users "US Discovery
Managers" -Filters "Mailbox_CountryOrRegion  -eq 'United States'"
```

This example allows the user annb@contoso.com to perform search actions only for mailboxes and OneDrive accounts in Canada. This filter contains the three-digit numeric country code for Canada from ISO 3166-1.

```
PowerShell
New-ComplianceSecurityFilter -FilterName CountryFilter  -Users annb@contoso.com -
Filters "Mailbox_CountryCode  -eq '124'"
```

This example allows the users *donh* and *suzanf* to search only the mailboxes and OneDrive accounts that have the value 'Marketing' for the CustomAttribute1 mailbox property.

```
PowerShell
New-ComplianceSecurityFilter -FilterName MarketingFilter  -Users donh,suzanf -Filters
"Mailbox_CustomAttribute1  -eq 'Marketing'"
```

This example allows members of the "Fourth Coffee eDiscovery Managers" role group to search only the mailboxes and OneDrive accounts that have the value 'FourthCoffee' for the Department mailbox property. The filter also allows the role group members to search for documents in the Fourth Coffee SharePoint site.

PowerShell

```
New-ComplianceSecurityFilter -FilterName "Fourth Coffee Security Filter" -Users
"Fourth Coffee eDiscovery Managers", "Fourth Coffee Investigators" -Filters
"Mailbox_Department -eq 'FourthCoffee'", "SiteContent_Path -like
'https://contoso.sharepoint.com/sites/FourthCoffee*' -or SiteContent_Path -like
'https://contoso-my.sharepoint.com/personal*'"
```

> **Note**
>
> In the previous example, an additional site content filter (`SiteContent_Path -like 'https://contoso-my.sharepoint.com/personal*'`) has to be included so that role group members can search for documents in OneDrive accounts. If this filter isn't included, the filter would only allow role group members to search for documents located in `https://contoso.sharepoint.com/sites/FourthCoffee`.

This example allows members of the eDiscovery Manager role group to search only the mailboxes and OneDrive accounts of members of the Ottawa Users distribution group. The Get-DistributionGroup cmdlet in Exchange Online PowerShell is used to find the members of the Ottawa Users group.

PowerShell

```
$DG = Get-DistributionGroup "Ottawa Users"
```

PowerShell

```
New-ComplianceSecurityFilter -FilterName DGFilter  -Users eDiscoveryManager -Filters
"Mailbox_MemberOfGroup -eq '$($DG.DistinguishedName)'"
```

This example prevents any user from performing search actions on the mailboxes and OneDrive accounts of members of the Executive Team distribution group. That means users can delete content from these mailboxes. The Get-DistributionGroup cmdlet in Exchange Online PowerShell is used to find the members of the Executive Team group.

PowerShell

```
$DG = Get-DistributionGroup "Executive Team"
```

PowerShell

```
New-ComplianceSecurityFilter -FilterName NoExecutivesPreview  -Users All -Filters
"Mailbox_MemberOfGroup -ne '$($DG.DistinguishedName)'"
```

This example allows members of the OneDrive eDiscovery Managers custom role group to only search for content in OneDrive accounts in the organization.

PowerShell

```
New-ComplianceSecurityFilter -FilterName OneDriveOnly  -Users "OneDrive eDiscovery
Managers" -Filters "SiteContent_Path -like 'https://contoso-
my.sharepoint.com/personal*'"
```

This example restricts the user to performing search actions only on email messages sent during the calendar year 2015.

PowerShell

```
New-ComplianceSecurityFilter -FilterName EmailDateRestrictionFilter -Users
donh@contoso.com -Filters "MailboxContent_Received -ge '01-01-2015' -and
MailboxContent_Received -le '12-31-2015'"
```

Similar to the previous example, this example restricts the user to performing search actions only on documents that were last changed sometime in the calendar year 2015.

PowerShell

```
New-ComplianceSecurityFilter -FilterName DocumentDateRestrictionFilter -Users
donh@contoso.com -Filters "SiteContent_LastModifiedTime -ge '01-01-2015' -and
SiteContent_LastModifiedTime -le '12-31-2015'"
```

This example prevents members of the "OneDrive Discovery Managers" role group from performing search actions on any mailbox in the organization.

PowerShell

```
New-ComplianceSecurityFilter -FilterName NoEXO -Users "OneDrive Discovery Managers" -
Filters "Mailbox_Alias -notlike '*'"
```

This example prevents anyone in the organization from performing search actions on email messages that were sent or received by janets or sarad.

PowerShell

```
New-ComplianceSecurityFilter -FilterName NoSaraJanet -Users All -Filters
"MailboxContent_Participants -notlike 'janets@contoso.onmicrosoft.com' -and
MailboxContent_Participants -notlike 'sarad@contoso.onmicrosoft.com'"
```

This example uses a filters list to combine mailbox and site filters. In this example, the mailbox filter is a content location filter and the site filter is a content filter.

PowerShell

```
New-ComplianceSecurityFilter -FilterName "Coho Winery Security Filter" -Users "Coho
Winery eDiscovery Managers", "Coho Winery Investigators" -Filters "Mailbox_Department
-eq 'CohoWinery'", "SiteContent_Path -like
'https://contoso.sharepoint.com/sites/CohoWinery*'"
```

Get-ComplianceSecurityFilter

The **Get-ComplianceSecurityFilter** is used to return a list of search permissions filters. Use the *FilterName* parameter to return information for a specific search filter.

Set-ComplianceSecurityFilter

The **Set-ComplianceSecurityFilter** is used to modify an existing search permissions filter. The following sections describe the parameters for this cmdlet. The only required parameter is *FilterName*.

FilterName

The *FilterName* parameter specifies the name of the permissions filter.

Users

The *Users* parameter specifies the users who get this filter applied to their searches. Because this is a multi-value property, specifying a user or group of users with this parameter overwrite the existing list of users. See the following examples for the syntax to add and remove selected users.

You can also use the *Users* parameter to specify a Microsoft Purview portal role group. This lets you create a custom role group and then assign that role group a search permissions filter. For

example, let's say you have a custom role group for eDiscovery managers for the U.S. subsidiary of a multi-national corporation. You can use the *Users* parameter to specify this role group (by using the Name property of the role group) and then use the *Filter* parameter to allow only mailboxes in the U.S. to be searched. You can't specify distribution groups with this parameter.

Filters

The *Filters* parameter specifies the search criteria for the compliance security filter. You can create three different types of filters:

- **Mailbox and OneDrive filtering:** This type of filter specifies the mailboxes and OneDrive accounts the assigned users (specified by the *Users* parameter) can search. The syntax for this type of filter is **Mailbox**_ *MailboxPropertyName*, where *MailboxPropertyName* specifies a mailbox property used to scope the mailboxes that can be searched. For example, the mailbox filter `"Mailbox_CustomAttribute10 -eq 'OttawaUsers'"` would allow the user assigned this filter to search only the mailboxes that have the value "OttawaUsers" in the CustomAttribute10 property. Any supported filterable recipient property can be used for the *MailboxPropertyName* property. For a list of supported properties, see Filterable properties for the -RecipientFilter parameter.

- **Mailbox content filtering:** This type of filter is applied on the content that can be searched. It specifies the mailbox content the assigned users can search for. The syntax for this type of filter is **MailboxContent**_*SearchablePropertyName*, where *SearchablePropertyName* specifies a Keyword Query Language (KeyQL) property that can be specified in a search. For example, the mailbox content filter `"MailboxContent_Recipients -like 'contoso.com'"` would allow the user assigned this filter to only search for messages sent to recipients in the contoso.com domain. For a list of searchable email properties, see Keyword queries and search conditions for eDiscovery.

- **Site and site content filtering:** There are two SharePoint and OneDrive for Business site-related filters that you can use to specify what site or site content the assigned users can search:
  - **Site**_ *SearchableSiteProperty*
  - **SiteContent**_*SearchableSiteProperty*

  These two filters are interchangeable. For example, `"Site_Path -like 'https://contoso.spoppe.com/sites/doctors*'"` and `"SiteContent_Path -like 'https://contoso.spoppe.com/sites/doctors*'"` return the same results. For a list of searchable site properties, see Overview of crawled and managed properties in SharePoint.

Properties marked with a **Yes** in the **Queryable** column can be used to create a site or site content filter.

Examples of changing search permissions filters

These examples show how to use the **Get-ComplianceSecurityFilter** and **Set-ComplianceSecurityFilter** cmdlets to add or remove a user to the existing list of users that the filter is assigned to. When you add or remove users from a filter, specify the user by using their SMTP address.

This example adds a user to the filter.

PowerShell

```
$filterusers = Get-ComplianceSecurityFilter -FilterName OttawaUsersFilter
```

PowerShell

```
$filterusers.users.add("pilarp@contoso.com")
```

PowerShell

```
Set-ComplianceSecurityFilter -FilterName OttawaUsersFilter -Users $filterusers.users
```

This example removes a user from the filter.

PowerShell

```
$filterusers = Get-ComplianceSecurityFilter -FilterName OttawaUsersFilter
```

PowerShell

```
$filterusers.users.remove("annb@contoso.com")
```

PowerShell

```
Set-ComplianceSecurityFilter -FilterName OttawaUsersFilter -Users $filterusers.users
```

Remove-ComplianceSecurityFilter

The **Remove-ComplianceSecurityFilter** is used to delete a search filter. Use the *FilterName* parameter to specify the filter you want to delete.

More information

How does search permissions filtering work? The permissions filter is appended to the search query when a search is run. The permissions filter is joined to the search query by the **AND** Boolean operator. The query logic for the search query and the permissions filter would look like this:

```text
<SearchQuery> AND <PermissionsFilter>
```

For example, you have a permissions filter that allows Bob to perform all search actions on the mailboxes of members of the Workers distribution group. Then Bob runs a search on all mailboxes in the organization with the search query `sender:jerry@adatum.com`. Because the permissions filter and the search query are logically combined by an **AND** operator, the search returns any message sent by jerry@adatum.com to any member of the Workers distribution group.

What happens if you have multiple search permissions filters? In a search query, multiple permissions filters are combined by **OR** Boolean operators. So results will be returned if any of the filters are true. In a search, all filters (combined by **OR** operators) are then combined with the search query by the **AND** operator.

```text
<SearchQuery> AND  (<PermissionsFilter1> OR <PermissionsFilter2> OR
<PermissionsFilter3>)
```

Let's take the previous example, where a search filter allows Bob to search only the mailboxes of the members of the Workers distribution group. Then we create another filter that prevents Bob from searching Phil's mailbox ("Mailbox_Alias -ne 'Phil'"). And let's also assume that Phil is a member of the Workers group. When Bob runs a search (from the previous example) on all mailboxes in the organization, search results are returned for Phil's mailbox even though you applied filter to prevent Bob from searching Phil's mailbox. This is because the first filter, which allows Bob to search the Workers group, is true. And because Phil is a member of the Workers group, Bob can search Phil's mailbox.

Does search permissions filtering work for inactive mailboxes? Yes, you can use mailbox and mailbox content filters to limit who can search inactive mailboxes in your organization. Like a regular mailbox, an inactive mailbox has to be configured with the recipient property that's used to create a permissions filter. If necessary, you can use the **Get-Mailbox -InactiveMailboxOnly** command to display the properties of inactive mailboxes. For more information, see Create and manage inactive mailboxes.

Does search permissions filtering work for public folders? No. As previously explained, search permissions filtering can't be used to limit who can search public folders in Exchange. For example, items in public folder locations can't be excluded from the search results by a permissions filter.

Does allowing a user to search all content locations in a specific service also prevent them from searching content locations in a different service? No. As previously explained, you have to create a search permissions filter to explicitly prevent users from searching content locations in a specific service (such as preventing a user from searching any Exchange mailbox or any SharePoint site). In other words, creating a search permissions filter that allows a user to search all SharePoint sites in the organization doesn't prevent that user from searching mailboxes. For example, to allow SharePoint admins to only search SharePoint sites, you have to create a filter that prevents them from searching mailboxes. Similarly, to allow Exchange admins to only search mailboxes, you have to create a filter that prevents them from searching sites.

Do search permissions filters count against search query character limits? Yes. Search permissions filters count against the character limit for search queries. For more information, see Limits in eDiscovery (Premium).

What is the maximum number of search permissions filters that can be created in an organization?

There's no limit to the number of search permissions filters that can be created in an organization. However, a search query can have a maximum of 100 conditions. In this case, a condition is defined as something that's connected to the query by a Boolean operator (such as **AND**, **OR**, and **NEAR**). The limit of the number of conditions includes the search query itself plus all search permissions filters that are applied to the user who runs the search. Therefore, the more search permissions filters you have (especially if these filters are applied to the same user or group of users), the better the chance of exceeding the maximum number of conditions for a search.

To understand how this limit works, you need to understand that a search permissions filter is appended to the search query when a search is run. A search permissions filter is joined to the

search query by the **AND** Boolean operator. The query logic for the search query and a single search permissions filter would look like this:

```text
<SearchQuery> AND <PermissionsFilter>
```

Multiple search permissions filters are combined together by the **OR** Boolean operator, and then those conditions are connected to the search query by the **AND** operator.

The query logic for the search query and multiple search permissions filters would look like this:

```text
<SearchQuery> AND (<PermissionsFilter1> OR <PermissionsFilter2> OR
<PermissionsFilter3>...)
```

It's possible the search query itself may consist of multiple conditions connected by Boolean operators. Each condition in the search query would also count against the 100-condition limit.

Also, the number of search permissions filters appended to a query depends on the user who is running the search. When a specific user runs a search, the search permissions filters that are applied to the user (which is defined by the *Users* parameter in the filter) are appended to the query. Your organization could have hundreds of search permissions filters, but if more than 100 filters are applied to the same users, then it's likely the 100-condition limit will be exceeded when those users run searches.

There's one more thing to keep in mind about the condition limit. The number of specific SharePoint sites that are included in the search query or search permissions filters also count against this limit.

To prevent your organization from reaching the conditions limit, keep the number of search permissions filters in your organization to few as possible to meet your business requirements.

Last updated on 10/31/2024

Content stored in Exchange Online mailboxes for eDiscovery

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

A mailbox in Exchange Online is primarily used to store email-related items such as messages, calendar items, tasks, and notes. But that's changing as more cloud-based apps also store their data in a user's mailbox. One advantage of storing data in a mailbox is that you can use the search tools in content search, Microsoft Purview eDiscovery (Standard), and Microsoft Purview eDiscovery (Premium) to find, view, and export the data from these cloud-based apps.

The data from some of these apps is stored in hidden folders located in a non-interpersonal message (non-IPM) subtree in the mailbox. Data from other cloud-based apps might not be stored *in* the mailbox, but it's *associated with* the mailbox, and is returned in searches (if that data matches the search query). Regardless of whether cloud-based data is stored in or associated with a user mailbox, the data is typically not visible in an email client when a user opens their mailbox.

The following table lists the apps that either stores or associates data with a cloud-based mailbox. The table also describes the type of content that each app produces.

Expand table

Microsoft 365 app	Description
Forms*	Forms and responses to a form are stored in files that are attached to email messages and stored in a hidden folder in the mailbox of the user who created the form. Forms created before April 2020 are stored as a PDF file. Forms created after 2020 are stored as a JSON file. Responses to a form are stored in a CSV file. When you export content from Forms in a PST file, this data is located in the

Microsoft 365 app	Description
	ApplicationDataRoot folder in a subfolder named with the following globally unique identified (GUID): **c9a559d2-7aab-4f13-a6ed-e7e9c52aec87**.
Microsoft 365 Groups	Email messages, calendar items, contacts (People), notes, and tasks are stored in the mailbox that's associated with a Microsoft 365 group.
Microsoft 365 Copilot and Microsoft 365 Copilot Chat	All Copilot activity data (user prompts and Copilot responses) generated in supported Microsoft 365 apps and services is stored in custodian mailboxes.
Outlook/Exchange Online	Email messages, calendar items, contacts (People), notes, and tasks are stored in a user's mailbox.
People	Contacts in the People app (which are the same contacts as the ones accessible in Outlook) are stored in a user's mailbox.
Class Schedule	Plans created in Class Schedule are stored in the mailbox of the corresponding Microsoft 365 Group that is provisioned when a new plan is created. The alias for the group mailbox is the name of the plan.
Skype for Business	Conversations in Skype for Business are stored in the Conversation History folder in a user's mailbox. If the mailbox of a participant of a Skype meeting is placed on Litigation Hold or assigned to a retention policy, files attached to a meeting are retained in the participants mailbox.
Sway*	Sways are stored as an HTML file that is attached to an email message and stored in a hidden folder in the mailbox of the user who created the sway. When you export content from Sway in a PST file, this data is located in the **ApplicationDataRoot** folder in a subfolder named with the following GUID: **905fcf26-4eb7-48a0-9ff0-8dcc7194b5ba**.
Tasks	Tasks in the Tasks app (which are the same tasks as the ones accessible in Outlook) are stored in a user's mailbox.
Teams	Conversations that are part of a Teams channel are associated with the Teams mailbox. Conversations that are part of the Chat list in Teams (also called *1 x N chats*) are associated with the mailbox of the users who participate in the chat. Also, summary information for meetings and calls in a Teams channel are associated with mailboxes of users who dialed into the meeting or call. So when searching for Teams content, you would search the Teams mailbox for content in channel conversations and search user mailboxes for content in 1 x N chats.
To-Do	Tasks (called *to-dos*, which are saved in to-do lists) in the To-Do app are stored in a user's mailbox.
Viva Engage	Conversations and comments within a Viva Engage community are associated with the Microsoft 365 group mailbox, as well as the user mailbox of the author and any named recipients (@ mentioned or Cc'ed users). Private messages sent outside of a

Microsoft 365 app	Description
	Viva Engage community are stored in the mailbox of the users who participate in the private message.

> **Note**
>
> * At this time, if a hold is placed on a mailbox (by using holds in eDiscovery (Standard) or eDiscovery (Premium) cases), content from this app will not be preserved by the hold.

Last updated on 11/21/2025

Decryption in Microsoft Purview eDiscovery tools

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

Encryption is an important part of your file protection and information protection strategy. Organizations of all types use encryption technology to protect sensitive content within their organization and ensure that only the right people have access to that content.

To run common eDiscovery tasks on encrypted content, eDiscovery managers are required to decrypt email message content when exported from content searches, Microsoft Purview eDiscovery (Standard) cases, and Microsoft Purview eDiscovery (Premium) cases. Content encrypted with Microsoft encryption technologies wasn't available for review until after export.

To make it easier to manage encrypted content in the eDiscovery workflow, Microsoft Purview eDiscovery tools now incorporate the decryption of encrypted files attached to email messages and sent in Exchange Online.[1] Additionally, encrypted documents stored in SharePoint Online and OneDrive for Business are decrypted in eDiscovery (Premium)[2].

Prior to this new capability, only the content of an email message protected by rights management (and not attached files) were decrypted. Encrypted documents in SharePoint and OneDrive couldn't be decrypted during the eDiscovery workflow. Now, files that are encrypted with a Microsoft encryption technology is located on a SharePoint or OneDrive account are searchable and decrypted when the search results are prepared for preview, added to a review set in eDiscovery (Premium), and exported. Additionally, encrypted documents in SharePoint and OneDrive that are attached to an email message (as a copy) are searchable. This decryption capability allows eDiscovery managers to view the content of encrypted email attachments and

site documents when previewing search results, and review them after they've been added to a review set in eDiscovery (Premium).

Requirements for decryption in eDiscovery

- **Permissions**: You have to be assigned the *RMS Decrypt* role to preview, review, and export files encrypted with Microsoft encryption technologies. You also have to be assigned this role to review and query encrypted files that are added to a review set in eDiscovery (Premium). This role is assigned by default to the eDiscovery Manager role group on the **Permissions** page in the Microsoft Purview portal. For more information about the RMS Decrypt role, see Assign eDiscovery permissions.
- **Run the Inbox Repair tool on exported PST files**: After you've exported PST files, we recommend that you run the Inbox Repair tool (ScanPST.exe) to diagnose and repair any errors in the PST files.

Supported encryption technologies

For Exchange, Microsoft Purview eDiscovery tools support items encrypted with Microsoft encryption technologies. These technologies are Azure Rights Management (Azure RMS)[3] and Microsoft Purview Information Protection (specifically sensitivity labels). For more information about Microsoft encryption technologies, see Encryption and the various email encryption options available. Content encrypted by S/MIME or third-party encryption technologies isn't supported. For example, previewing or exporting content encrypted with non-Microsoft technologies isn't supported.

> **Note**
>
> The decryption of email messages sent with an Microsoft Purview Message Encryption custom branding template is not supported by Microsoft eDiscovery tools. When using an OME custom branding template, email messages are delivered to the OME portal instead of the recipient's mailbox. Therefore, you won't be able to use eDiscovery tools to search for encrypted messages because those messages are never received by the recipient's mailbox.

For SharePoint, content labeled with SharePoint online service are decrypted. Items labeled or encrypted in the client before uploading to SharePoint, legacy document library RMS templates or settings and S/MIME or other standards aren't supported[2].

eDiscovery activities that support encrypted items

The following table identifies the supported tasks that can be performed in Microsoft Purview eDiscovery tools on encrypted files attached to email messages and encrypted documents in SharePoint and OneDrive. These supported tasks can be performed on encrypted files that match the criteria of a search. A value of `N/A` indicates the functionality isn't available in the corresponding eDiscovery tool.

Expand table

eDiscovery task	Content search	eDiscovery (Standard)	eDiscovery (Premium)
Search for content in encrypted files in sites and email attachments[1]	No	No	Yes
Preview encrypted files attached to email	No	No	Yes
Preview encrypted documents in SharePoint and OneDrive	No	No	Yes
Review encrypted files in a review set	N/A	N/A	Yes
Export encrypted files attached to email	Yes	Yes	Yes
Export encrypted documents in SharePoint and OneDrive	No	No	Yes

Supported decryption

The following table describes the decryption supported by eDiscovery (Standard) and eDiscovery (Premium) for email, email with attachments, and files hosted by SharePoint.

Expand table

Item type	Task	eDiscovery (Standard)	eDiscovery (Premium)
Encrypted email	Search	Yes	Yes
Encrypted email	Decryption to .pst	No	Yes
Encrypted email	Decryption to file	Yes	Yes
Encrypted mail and attachment	Search	No	Yes (with Advanced indexing)[1]

Item type	Task	eDiscovery (Standard)	eDiscovery (Premium)
Encrypted mail and attachment	Decryption to .pst	No	Yes
Encrypted mail and attachment	Decryption to file	No	Yes
File in SharePoint with MIP label	Search	No	Yes
File in SharePoint with MIP label	Decryption	No	Yes
File in SharePoint with other encryption[2]	Search, Decryption	No	No

> **Important**
>
> eDiscovery (Standard) doesn't support legacy encryption protocols.

Decryption limitations with email and attachments

eDiscovery support for decryption of email messages and attachments is subject to the following limitations:

- Decryption isn't supported when the email or attachment encryption is applied in an external organization. eDiscovery only supports decryption for email and attachments that are encrypted in your organization.
- When decrypting emails or attachments, the owner of the mailbox where the emails and attachments are included in eDiscovery activities must have access to view the encrypted content. Decryption for emails or attachments isn't supported if they're sent or forwarded other recipients that can't view the encrypted content. Changes in the owner's groups or other organization permissions may also affect decryption support.

Decryption limitations with sensitivity labels in SharePoint and OneDrive

eDiscovery doesn't support encrypted files in SharePoint and OneDrive when a sensitivity label that applied the encryption is configured with either of the following settings:

- Users can assign permissions when they manually apply the label to a document. This is sometimes referred to as *user-defined permissions*.

- User access to the document has an expiration setting that is set to a value other than **Never**.

For more information about these settings, see the "Configure encryption settings" section in Restrict access to content by using sensitivity labels to apply encryption.

Documents encrypted with the previous settings can still be returned by an eDiscovery search. This result may happen when a document property (such as the title, author, or modified date) matches the search criteria. Although these documents might be included in search results, they can't be previewed or reviewed. These documents will also remain encrypted when they're exported in eDiscovery (Premium).

> **Important**
>
> Decryption isn't supported for files that are locally encrypted and then uploaded to SharePoint or OneDrive. For example, local files that are encrypted by the Microsoft Purview Information Protection client and then uploaded to Microsoft 365 aren't supported. Only files that are encrypted in the SharePoint or OneDrive service are supported for decryption.

Notes

[1] Encrypted files located on a local computer and copied to an email message aren't decrypted and indexed for eDiscovery. For eDiscovery (Premium), encrypted email and attachments in recipient mailbox needs to be advanced indexed to be decrypted. For more information about Advanced indexing, see Advanced indexing of custodian data.

[2] Only items labeled in SharePoint (or uploaded to SharePoint after integration with sensitivity labels are enabled) and that have labels with admin-defined permissions and no expiration are decrypted. All other encrypted files in SharePoint aren't decrypted. For more information, see Enable sensitivity labels for files in SharePoint and OneDrive.

Other documents aren't decrypted, including:

- Files encrypted in the client and uploaded before sensitivity labels were integrated with SharePoint Online.
- Documents encrypted with legacy RMS templates and not labeled.
- Documents with user-defined permissions or with expiration settings (SMIME or other standards).

[3] Only content encrypted with RMS keys hosted in Microsoft 365 is transparently decrypted by eDiscovery (Premium). Double Key Encryption (DKE), Hold Your Own Key (HYOK), on-premises RMS, etc. aren't supported. For more information, see Key types for the service.

Last updated on 10/31/2024

Conversation threading in eDiscovery (Premium)

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

Instant messaging is a convenient way to ask questions, share ideas, or quickly communicate across large audiences. As instant messaging platforms, like Microsoft Teams and Viva Engage groups, become core to enterprise collaboration, organizations must evaluate how their eDiscovery workflow addresses these new forms of communication and collaboration.

The conversation reconstruction feature in Microsoft Purview eDiscovery (Premium) is designed to help you identify contextual content and produce distinct conversation views. This capability allows you to efficiently and rapidly review complete instant message conversations (also called *threaded conversations*) that are generated in platforms like Microsoft Teams.

With conversation reconstruction, you can use built-in capabilities to reconstruct, review, and export threaded conversations. Use eDiscovery (Premium) conversation reconstruction to:

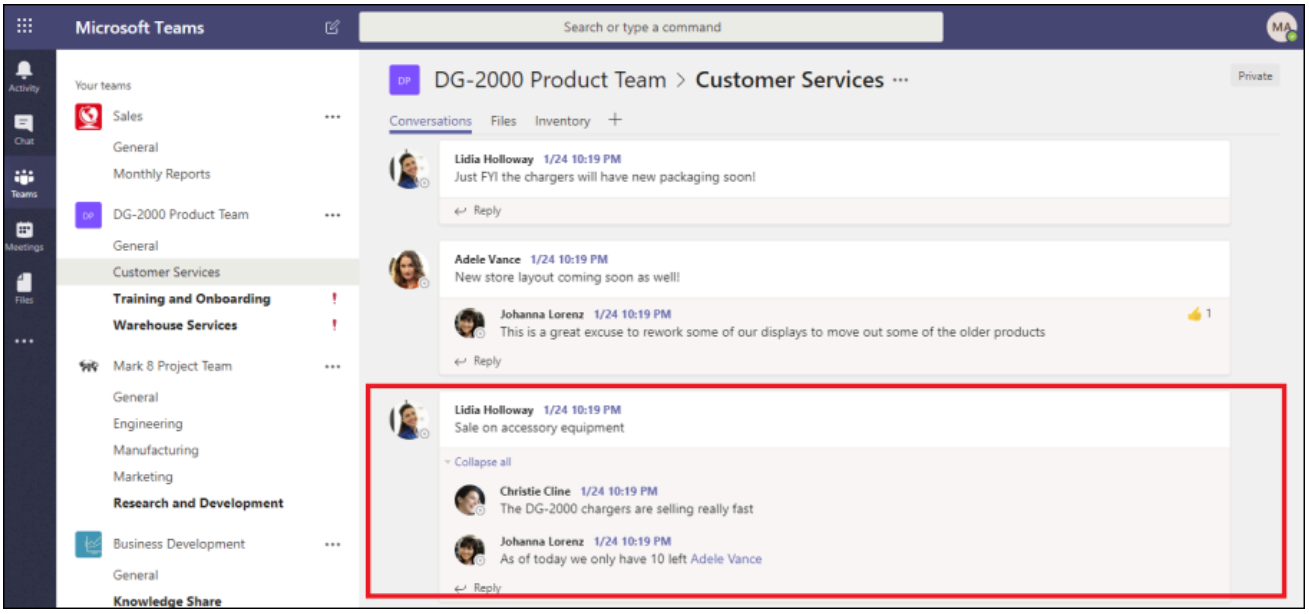
- Preserve unique message-level metadata across all messages within a conversation.
- Collect contextual messages around your search results.
- Review, annotate, and redact threaded conversations.
- Export individual messages or threaded conversations

Terminology

Here are few definitions to help you get start using conversation reconstruction.

- **Messages:** Represent the smallest unit of a conversation. Messages may vary in size, structure, and metadata.

- **Conversation:** Represents a grouping of one or more messages. Across different applications, conversations may be represented in different ways. In some applications, there is an explicit action that results from replying to an existing message. Conversations are formed explicitly as a result of this user action. For example, here is a screenshot of a channel conversation in Microsoft Teams.



  In other apps (such as group chat messages in Teams), there is not a formal reply chain and instead messages appear as a "flat river of messages" within a single thread. In these types apps, conversations are inferred from a group of messages that occur within a certain time. This "soft-grouping" of messages (as opposed to a reply chain) represent the "back and forth" conversation about a specific topic of interest.

Step 1: Create a collection estimate

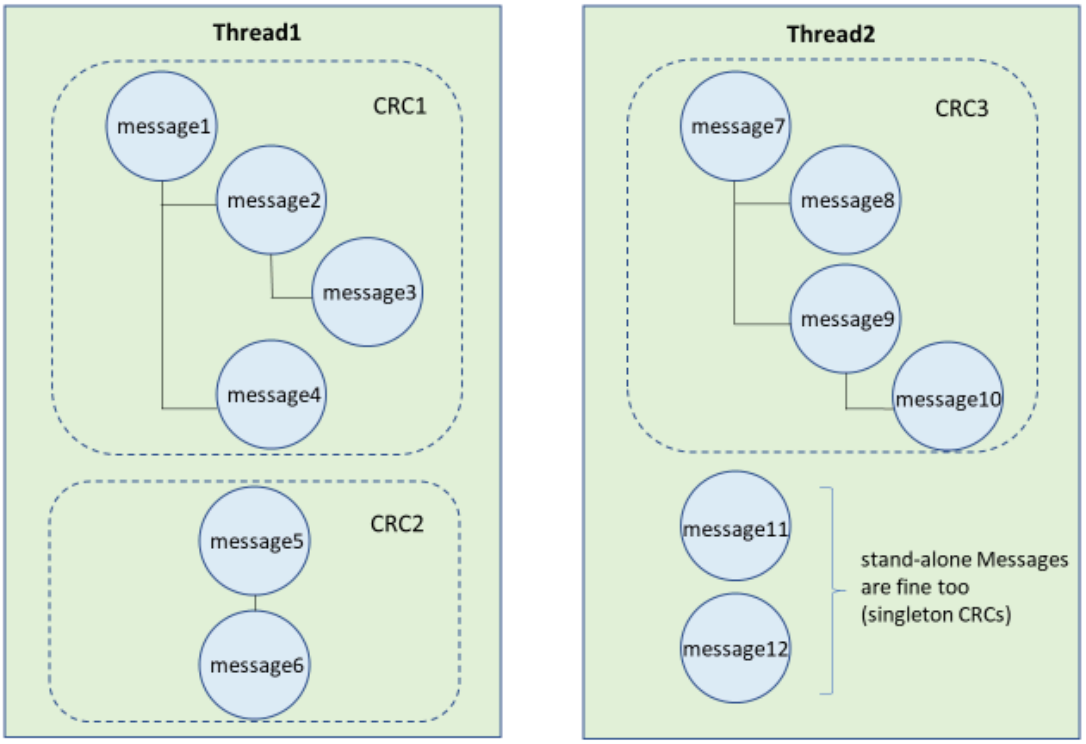
After you have identified relevant custodians and content locations, you can create a search to find potentially relevant content. On the **Collections** tab in the eDiscovery (Premium) case, you can create a collection by selecting **New collection** and following the workflow. For information about how you can create a collection, build a search query, and preview the search results, see Create a collection estimate.

Step 2: Commit a collection estimate to a review set

After you have reviewed and finalized the search query in a collection, you can add the search results to a review set. When you add your search results into a review set, the original data is copied to an Azure Storage area to facilitate the review and analysis process. For more

information about adding search results to a review set, see Commit a collection estimate to a review set.

When you add items from conversations to a review set, you can use the threaded conversations option to collect contextual messages from conversations that contain items that match the search criteria of the collection. After you select the thread conversations option, the following things can happen:

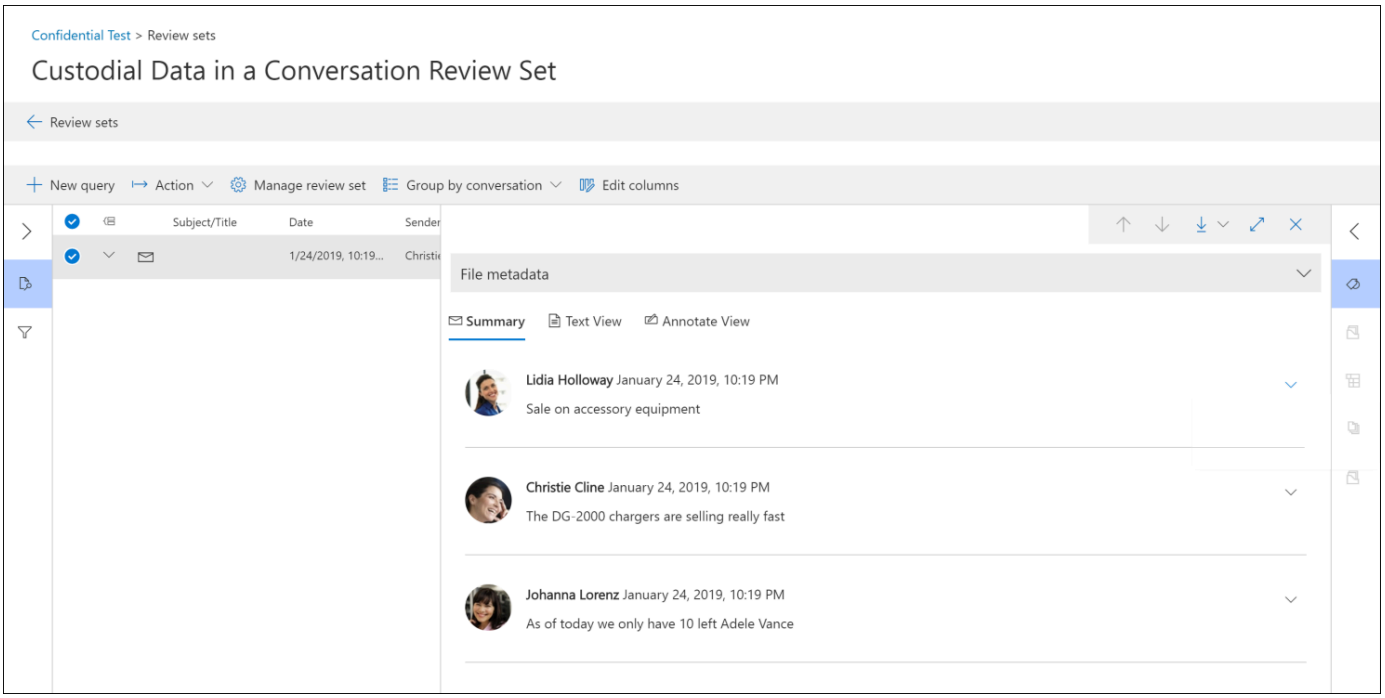


1. Using a keyword and date range query, the search returned a hit on *Message 3*. This message was part of a larger conversation, illustrated by *CRC1*.
2. When you add the data into a review set and enable the conversation retrieval options, eDiscovery (Premium) will go back and collect other items in *CRC1*.
3. After the items have been added to the review set, you can review all the individual messages from *CRC1*.

To enable the threaded conversations option, see Commit a collection estimate to a review set.

Step 3: Review and export threaded conversations

After the content has been processed and added to the review set, you can start reviewing the data in the review set. Individual messages are threaded together and presented as conversations. This lets you review and export contextual conversations.



The following sections describe reviewing and exporting conversations.

Reviewing conversations

In a review set, you can use the following options to facilitate the review process.

- **Group by conversation:** Groups messages within the same conversation together to help users simplify and expedite their review process.
- **Summary view:** Displays the threaded conversation. In this view, you can see the entire conversation and also access the metadata for each individual message.
  - View metadata for individual messages
  - Download individual messages
- **Text view:** Provides the extracted text for the entire conversation.
- **Annotate view:** Lets you markup a threaded view of the conversation. All messages in the conversation share the same annotated document.
- **Tagging:** When viewing conversations in a review set, you can view and apply tags by selecting **Tagging panel** in the Coding panel.
- **Rerun conversation conversion:** When messages are added to a conversation review set, a conversion job is automatically run to create the threaded summary and annotate views. If the Conversation Reconstruction job fails, you can rerun this job by selecting **Action** > **Create conversation PDFs** in the review set.

> **Note**
>
> The date and time for all conversations in Teams are shown in Coordinated Universal Time (UTC).

Exporting conversations

For the options you can select when exporting conversations from a review set, see Export documents from a review set.

Specifically, you can export entire chat conversations in a single PDF file or you can export each chat message in a conversation as an individual file.

More information

To learn more about how to review case data in eDiscovery (Premium), see the following articles:

- Query and filter content in a review set
- Tag documents in a review set
- View case data
- Analyze case data
- Export case data

Last updated on 09/02/2025

CJK language support for eDiscovery (Premium)

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

Microsoft Purview eDiscovery (Premium) supports double-byte character set languages (these include Simplified Chinese, Traditional Chinese, Japanese, and Korean, which are collectively known as *CJK* languages) for the following advanced scenarios in a review set:

- When you query the data in a review set.
- When you tag documents in a review set.
- When you analyze case data in a review set by using near duplicate detection, email threading, and themes analytics.

Frequently asked questions

How do I create a search to collect items that contains CJK characters?

You can use CJK characters for keyword searches, keyword queries and search conditions when searching for content in eDiscovery (Premium). Searching for CJK characters is also supported when searching for content in Microsoft Purview eDiscovery (Standard) and Content Search.

We provide CJK support for all search operators and search conditions, including the boolean operators **AND**, **OR**, **NOT**, and **NEAR**.

If you're certain that content locations or items contain CJK characters, but searches aren't returning any results, select the query language-country/region icon and select the corresponding language-country/region culture code value for the search. The default language/region is neutral.

Can I search for multiple languages at once?

It depends on your search scenario.

- When you query data in a review set in eDiscovery (Premium), you can search for multiple languages.
- When you create a search to collect data, create separate collections for each language you're targeting. For example, if you are searching for a document that contains both Chinese and Korean, select Chinese for your first collection and select Korean for your second collection.

I don't see the query language-country/region icon to select a language for queries in a review set. How can I specify a query language in a review set search?

For review set queries, you don't need to specify a document language. eDiscovery (Premium) automatically detects document languages when you add content to a review set. This helps you optimize your query results in a review set.

Can I see detected languages in file metadata?

No, you can't see detected languages in file metadata.

Can I filter by document languages in a review set?

No, you can't filter, sort, or search by document languages in a review set.

Will this CJK release for review set scenarios affect any of my existing searches and review sets?

No, none of your existing searches and review sets will change. You don't need to reindex existing data, and search results for English text will be the same.

How do I change my display language to Chinese, Japanese, or Korean?

For information about how to change display language and time zone, see How to set language and region settings for Office 365.

Limits

- OCR doesn't support CJK characters from image files
- Email files (such as *.eml and *.msg) in Annotate view aren't supported for CJK languages.
- Search hit highlighting in Text view isn't supported for CJK languages.

- The Relevance module used to analyze data doesn't support CJK languages.
- Query-based holds aren't supported for CJK languages.

Last updated on 11/21/2025

Collect eDiscovery diagnostic information

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

Occasionally Microsoft Support engineers require specific information about your issue when you open a support case related to Microsoft Purview eDiscovery (Standard) or Microsoft Purview eDiscovery (Premium). This article provides guidance on how to collect diagnostic information to help support engineers investigate and resolve issues. Typically, you don't need to collect this information until asked to do so by a Microsoft Support engineer.

> **Important**
>
> The output from the cmdlets and diagnostic information described in this article may include sensitive information about litigation or internal investigations in your organization. Before sending the raw diagnostic information to Microsoft Support, you should review the information and redact any sensitive information (such as names or other information about parties to litigation or investigation) by replacing it with `xxxxxxx`. Using this method will also indicate to the Microsoft Support engineer that information was redacted.

Collect diagnostic information for eDiscovery (Standard)

Collecting diagnostic information for eDiscovery (Standard) is cmdlet-based, so you'll have to use Security & Compliance PowerShell. The following PowerShell examples will run cmdlets and then save the output to a specified text file. In most support cases, you should only have to run one of these commands.

To run the following cmdlets, connect to Security & Compliance PowerShell. After you're connected, run one or more of the following commands and be sure to replace placeholders with the actual object names.

After reviewing the generated text file and redacting sensitive information, send it to the Microsoft Support engineer working on your case.

> **Note**
>
> You can also run the commands in this section to collect diagnostic information for the searches and exports listed on the **Content search** page in the Microsoft Purview portal.

Collect information about searches

The following command collects information that's helpful when investigating issues with a Content search or a search associated with an eDiscovery (Standard) case.

PowerShell

```
Get-ComplianceSearch "<Search name>" | FL > "ComplianceSearch.txt"
```

Collect information about search actions

The following command collects information to investigate problems with previewing, exporting, or purging the results of a Content search or a search associated with an eDiscovery (Standard) case. You can identify the name of the search action by selecting an export that's listed on the **Exports** tab. To identify the names of preview and purge actions, you can run the **Get-ComplianceSearchAction** cmdlet to display a list of all actions. The format for the search action name is constructed by appending `_Preview`, `_Export`, or `_Purge` to the name of the corresponding search.

PowerShell

```
Get-ComplianceSearchAction "<Search action name>" | FL > "ComplianceSearchAction.txt"
```

Collect information about eDiscovery holds

When an eDiscovery hold associated with an eDiscovery (Standard) case isn't functioning as expected, run the following command to collect information about the Case Hold Policy and

associated Case Hold Rule for the eDiscovery hold. The *Case hold policy name* in the following command is the same as the name of the eDiscovery hold. You can identify this name on the **Holds** tabs in the eDiscovery (Standard) case.

PowerShell

```
Get-CaseHoldPolicy "<Case hold policy name>" | %{"--CaseHoldPolicy--";$_|FL;"--
CaseHoldRule--";Get-CaseHoldRule -Policy $_.Name | FL} > "eDiscoveryCaseHold.txt"
```

Collect all case information

Sometimes, it's not apparent what information is required by Microsoft Support to investigate your issue. In this situation, you can collect all of the diagnostics information for an eDiscovery (Standard) case. The *eDiscovery (Standard) case name* in the following command is the same as the name of a case that's displayed on the **eDiscovery (Standard)** page in the Microsoft Purview portal.

PowerShell

```
Get-ComplianceCase "<eDiscovery (Standard) case name>"| %{$_|fl;"`t==Searches==";Get-
ComplianceSearch -Case $_.Name | FL;"`t==Search Actions==";Get-ComplianceSearchAction
-Case $_.Name |FL;"`t==Holds==";Get-CaseHoldPolicy -Case $_.Name | %{$_|FL;"`t`t
==$($_.Name) Rules==";Get-CaseHoldRule -Policy $_.Name | FL}} > "eDiscoveryCase.txt"
```

Collect diagnostic information for eDiscovery (Premium)

Collect case information

The **Settings** tab in an eDiscovery (Premium) case lets you quickly copy the diagnostic information for the case. The diagnostic information is saved to the clipboard so you can paste it to a text file and send to Microsoft Support.

1. Go to the Microsoft Purview portal, and select **eDiscovery** > **Advanced**.
2. Select a case and then select the **Settings** tab.
3. Under **Case Information**, select **Select**.
4. On the flyout page, select **Actions** > **Copy support information** to copy the information to the clipboard.
5. Open a text file (in Notepad) and then paste the information in the text file.

6. Save the text file and name it something like `eDP Diagnostic Info YYYY.MM.DD` (for example, `eDP Diagnostic Info 2020.11.03`).

After reviewing the file and redacting sensitive information, send it to the Microsoft Support engineer working on your case.

Collect job information

1. While in a case, select the **Jobs** tab.
2. Scroll down to the Job that failed and select it to open the job information pane.
3. Select **Copy support information** to copy the job diagnostics information to the clipboard.
4. Open a text file (in Notepad) and then paste the information into the text file.
5. Save the text file and name it something like `eDP Job Diagnostic Info YYYY.MM.DD` (for example, `eDP Diagnostic Info 2020.11.03`).

After reviewing the file and redacting sensitive information, send it to the Microsoft Support engineer working on your case.

Last updated on 03/16/2026

Search for eDiscovery (classic) activities in the audit log

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

Content Search and eDiscovery-related activities (for Microsoft Purview eDiscovery (Standard) and Microsoft Purview eDiscovery (Premium)) that are performed in Microsoft Purview portal or by running the corresponding PowerShell cmdlets are logged in the audit log. Events are logged when administrators or eDiscovery managers (or any user assigned eDiscovery permissions) perform the following Content Search and eDiscovery (Standard) tasks in the Microsoft Purview portal:

- Creating and managing eDiscovery (Standard) and eDiscovery (Premium) cases.
- Creating, starting, and editing Content searches.
- Performing search actions, such as previewing, exporting, and deleting search results.
- Managing custodians and review sets in eDiscovery (Premium).
- Configuring permissions filtering for Content search.
- Managing the eDiscovery Administrator role.

For more information about searching the audit log, the permissions that are required, and exporting search results, see Search the audit log.

How to search for and view eDiscovery activities

Currently, you have to do a few specific things to view eDiscovery activities in the audit log. Here's how:

1. Go to the Microsoft Purview portal and sign in using your work or school account.

2. In the left navigation pane of the Microsoft Purview portal, select **Audit**.
3. In the **Activities** drop-down list, under **eDiscovery activities** or **eDiscovery (Premium) activities**, select one or more activities to search for.

   > ⓘ Note
   >
   > The **Activities** drop-down list also includes a group of activities named **eDiscovery cmdlet activities** that will return records from the cmdlet audit log.

4. Select a date and time range to display eDiscovery events that occurred within that period.
5. In the **Users** box, select one or more users to display search results for. Leave this box blank to return entries for all users.
6. Select **Search** to run the search using your search criteria.
7. After the search results are displayed, you can select **Filter results** to filter or sort the resulting activity records. Unfortunately, you can't use filtering to explicitly exclude certain activities.
8. To view details about an activity, select the activity record in the list of search results.

   A **Details** fly out page is displayed that contains the detailed properties from the event record. To display additional details, select **More information**. For a description of these properties, see the Detailed properties for eDiscovery activities section.
9. If desired, you can export the audit log search results to a CSV file, and then use the Excel Power Query feature to format and filter these records. For more information, see Export, configure, and view audit log records.

eDiscovery activities

The following table describes the Content Search and eDiscovery (Standard) activities that are logged when an administrator or eDiscovery manager performs an eDiscovery-related activity using the Microsoft Purview portal. Some activities performed in eDiscovery (Premium) may be returned when you search for activities in this list.

> ⓘ Note

The eDiscovery activities described in this section provide similar information to the eDiscovery cmdlet activities described in the next section. We recommend that you use the eDiscovery activities described in this section because they will appear in the audit log search results within 30 minutes. It may take up to 24 hours for eDiscovery cmdlet activities to appear in audit log search results.

Expand table

Friendly name	Operation	Corresponding cmdlet	Description
Added member to eDiscovery case	CaseMemberAdded	Add-ComplianceCaseMember	A user was added as a member of an eDiscovery case. As a member of a case, a user can perform various case-related tasks depending on whether they've been assigned the necessary permissions.
Changed content search	SearchUpdated	Set-ComplianceSearch	An existing content search was changed. Changes can include adding or removing content locations or editing the search query.
Changed eDiscovery administrator membership	CaseAdminUpdated	Update-eDiscoveryCaseAdmin	The list of eDiscovery Administrators in your organization was changed. This activity is logged when the list of eDiscovery Administrators is replaced with a group of new users. If a single user is added or removed, the CaseAdminAdded operation is logged.
Changed eDiscovery case	CaseUpdated	Set-ComplianceCase	An eDiscovery case was changed. Changes include closing an open case or reopening a closed case.

Friendly name	Operation	Corresponding cmdlet	Description
Changed eDiscovery case membership	CaseMemberUpdated	Update-ComplianceCaseMember	The membership list of an eDiscovery case was changed. This activity is logged when all members are replaced with a group of new users. If a single member is added or removed, CaseMemberAdded or CaseMemberRemoved operation is logged.
Changed search permissions filter	SearchPermissionUpdated	Set-ComplianceSecurityFilter	A search permissions filter was changed.
Changed search query for eDiscovery case hold	HoldUpdated	Set-CaseHoldRule	A query-based hold associated with an eDiscovery case was changed. Possible changes include editing the query or date range for a query-based hold.
Content search preview item downloaded	PreviewItemDownloaded	N/A	A user downloaded an item to their local computer (by selecting the **Download original item** link) when previewing search results.
Content search preview item listed	PreviewItemListed	N/A	A user selected **Preview search results** to display the preview search results page, which lists up to 1,000 items from the results of a search.
Created content search	SearchCreated	New-ComplianceSearch	A new content search was created.
Created eDiscovery administrator	CaseAdminAdded	Add-eDiscoveryCaseAdmin	A user was added as an eDiscovery Administrator in the organization.

Friendly name	Operation	Corresponding cmdlet	Description
Created eDiscovery case	CaseAdded	New-ComplianceCase	An eDiscovery case was created. When a case is created, you only have to give it a name. Other case-related tasks such as adding members, creating holds, and creating content searches associated with the case result in additional events being logged.
Created search permissions filter	SearchPermissionCreated	New-ComplianceSecurityFilter	A search permissions filter was created.
Created search query for eDiscovery case hold	HoldCreated	New-CaseHoldRule	A query-based hold associated with an eDiscovery case was created.
Deleted content search	SearchRemoved	Remove-ComplianceSearch	An existing content search was deleted.
Deleted eDiscovery administrator	CaseAdminRemoved	Remove-eDiscoveryCaseAdmin	An eDiscovery Administrator was deleted from your organization.
Deleted eDiscovery case	CaseRemoved	Remove-ComplianceCase	An eDiscovery case was deleted. Any hold associated with the case has to be removed before the case can be deleted.
Deleted search permissions filter	SearchPermissionRemoved	Remove-ComplianceSecurityFilter	A search permissions filter was deleted.
Deleted search query for eDiscovery case hold	HoldRemoved	Remove-CaseHoldRule	A query-based hold associated with an eDiscovery case was deleted. Removing the query from the hold is often the result of

Friendly name	Operation	Corresponding cmdlet	Description
			deleting a hold. When a hold or a hold query is deleted, the content locations that were on hold are released.
Downloaded export of content search	SearchExportDownloaded	N/A	A user downloaded the results of a content search to their local computer. A **Started export of content search** activity has to be initiated before search results can be downloaded.
Previewed results of content search	SearchPreviewed	N/A	A user previewed the results of a content search.
Purged results of content search	SearchResultsPurged	New-ComplianceSearchAction	A user purged the results of a Content search by running the **New-ComplianceSearchAction -Purge** command.
Removed analysis of content search	RemovedSearchResultsSentToZoom	Remove-ComplianceSearchAction	A content search prepare action (to prepare search results for eDiscovery (Premium)) was deleted. If the preparation action was less than two weeks old, the search results that were prepared for eDiscovery (Premium) were deleted from the Microsoft Azure storage area. If the preparation action was older than 2 weeks, then this event indicates that only the corresponding preparation action was deleted.
Removed export of	RemovedSearchExported	Remove-ComplianceSearchAction	A content search export action was deleted. If the export action was less

Friendly name	Operation	Corresponding cmdlet	Description
content search			than two weeks old, the search results that were uploaded to the Microsoft Azure storage area were deleted. If the export action was older than 2 weeks, then this event indicates that only the corresponding export action was deleted.
Removed member from eDiscovery case	CaseMemberRemoved	Remove-ComplianceCaseMember	A user was removed as a member of an eDiscovery case.
Removed preview results of content search	RemovedSearchPreviewed	Remove-ComplianceSearchAction	A content search preview action was deleted.
Removed purge action performed on content search	RemovedSearchResultsPurged	Remove-ComplianceSearchAction	A content search purge action was deleted.
Removed search report	SearchReportRemoved	Remove-ComplianceSearchAction	A content search export report action was deleted.
Started analysis of content search	SearchResultsSentToZoom	New-ComplianceSearchAction	The results of a content search were prepared for analysis in eDiscovery (Premium).
Started content search	SearchStarted	Start-ComplianceSearch	A content search was started. When you create or change a content search by using the Microsoft Purview portal, the search is automatically started.
Started export of	SearchExported	New-ComplianceSearchAction	A user exported the results of a content search.

Friendly name	Operation	Corresponding cmdlet	Description
content search			
Started export report	SearchReport	New-ComplianceSearchAction	A user exported a content search report.
Stopped content search	SearchStopped	Stop-ComplianceSearch	A user stopped a content search.
(none)	CaseViewed	Get-ComplianceCase	A user viewed an eDiscovery (Standard) case in the Microsoft Purview portal. The audit record for this event includes the name of the case that was viewed.
(none)	SearchViewed	Get-ComplianceSearch	A user viewed a Content search in the Microsoft Purview portal by accessing the search on the **Searches** tab in an eDiscovery (Standard) case or accessing it on the **Content search** page. The audit record for this event includes the identity of the search that was viewed.
(none)	ViewedSearchPreviewed	Get-ComplianceSearchAction -Preview	A user previewed the results of a Content search in the Microsoft Purview portal. This activity is also logged when a user previews the results of a search associated with an eDiscovery (Standard) case.

eDiscovery (Premium) activities

The following table describes the eDiscovery (Premium) activities logged in the audit log. These activities can be used to help you track the progression of activity in an eDiscovery (Premium) case.

Expand table

Friendly name	Operation	Description
Added data to another review set	AddWorkingSetQueryToWorkingSet	User added documents from one review set to a different review set.
Added data to review set	AddQueryToWorkingSet	User added the search results from a content search associated with an eDiscovery (Premium) case to a review set.
Added non-Microsoft 365 data to review set	AddNonOffice365DataToWorkingSet	User added non-Microsoft 365 data to a review set.
Added remediated documents to review set	AddRemediatedData	User uploads documents that had indexing errors that were fixed to a review set.
Analyzed data in review set	RunAlgo	User ran analytics on the documents in a review set.
Annotated document in review set	AnnotateDocument	User annotated a document in a review set. Annotation includes redacting content in a document.
Compared load sets	LoadComparisonJob	User compared two different load sets in a review set. A load set is when data from a content search that associated with the case is added to a review set.
Converted redacted documents to PDF	BurnJob	User converted all the redacted documents in a review set to PDF files.
Created review set	CreateWorkingSet	User created a review set.
Created review set search	CreateWorkingSetSearch	User created a search query that searches the documents in a review set.
Created tag	CreateTag	User created a tag group in a review set. A tag group can contain one or more child tags. These tags are then used to tag documents in the review set.
Deleted review set search	DeleteWorkingSetSearch	User deleted a search query in a review set.

Friendly name	Operation	Description
Deleted tag	DeleteTag	User deleted a tag or a tag group in a review set.
Downloaded document	DownloadDocument	User downloaded a document from a review set.
Edited tag	UpdateTag	User changed a tag in a review set.
Exported documents from review set	ExportJob	User exported documents from a review set.
Modified case setting	UpdateCaseSettings	User modified the settings for a case. Case settings include case information, access permissions, and settings that control search and analytics behavior.
Modified review set search	UpdateWorkingSetSearch	User edited a search query in a review set.
Previewed review set search	PreviewWorkingSetSearch	User previewed the results of a search query in a review set.
Remediated error documents	ErrorRemediationJob	User fixes files that contained indexing errors.
Tagged document	TagFiles	User tags a document in a review set.
Tagged results of a query	TagJob	User tags all of the documents that match the criteria of search query in a review set.
Viewed document in review set	ViewDocument	User viewed a document in a review set.

eDiscovery cmdlet activities

The following table lists the cmdlet audit log records that are logged when an administrator or user performs an eDiscovery-related activity by using the Microsoft Purview portal or by running the corresponding cmdlet in Security & Compliance PowerShell. The detailed information in the audit log record is different for the cmdlet activities listed in this table and the eDiscovery activities described in the previous section.

As previously stated, it may take up to 24 hours for eDiscovery cmdlet activities to appear in the audit log search results.

> **Tip**
>
> The cmdlets in the **Operation** column in the following table are linked to the corresponding cmdlet help topic on TechNet. Go to the cmdlet help topic for a description of the available parameters for each cmdlet. The parameter and the parameter value that were used with a cmdlet are included in the audit log entry for each eDiscovery cmdlet activity that's logged.

Expand table

Friendly name	Operation (cmdlet)	Description
Created hold in eDiscovery case	New-CaseHoldPolicy	A hold was created for an eDiscovery case. A hold can be created with or without specifying a content source. If content sources are specified, they'll be identified in the audit log entry.
Deleted hold from eDiscovery case	Remove-CaseHoldPolicy	A hold that is associated with an eDiscovery case was deleted. Deleting a hold releases all of the content locations from the hold. Deleting the hold also results in deleting the case hold rules associated with the hold (see **Remove-CaseHoldRule** below).
Changed hold in eDiscovery case	Set-CaseHoldPolicy	A hold that is associated with an eDiscovery was changed. Possible changes include adding or removing content locations or turning off (disabling) the hold.
Created search query for eDiscovery case hold	New-CaseHoldRule	A query-based hold associated with an eDiscovery case was created.
Deleted search query for eDiscovery case hold	Remove-CaseHoldRule	A query-based hold associated with an eDiscovery case was deleted. Removing the query from the hold is often the result of deleting a hold. When a hold or a hold query is deleted, the content locations that were on hold are released.
Changed search query for eDiscovery case hold	Set-CaseHoldRule	A query-based hold associated with an eDiscovery case was changed. Possible changes include editing the query or date range for a query-based hold.
Created eDiscovery case	New-ComplianceCase	An eDiscovery case was created. When a case is created, you only have to give it a name. Other case-related tasks such as adding members, creating holds, and creating content searches associated with the case result in additional events being logged.

Friendly name	Operation (cmdlet)	Description
Deleted eDiscovery case	Remove-ComplianceCase	An eDiscovery case was deleted. Any hold associated with the case has to be removed before the case can be deleted.
Changed eDiscovery case	Set-ComplianceCase	An eDiscovery case was changed. Changes include closing an open case or reopening a closed case.
Added member to eDiscovery case	Add-ComplianceCaseMember	A user was added as a member of an eDiscovery case. As a member of a case, a user can perform various case-related tasks depending on whether they've been assigned the necessary permissions.
Removed member from eDiscovery case	Remove-ComplianceCaseMember	A user was removed as a member of an eDiscovery case.
Changed eDiscovery case membership	Update-ComplianceCaseMember	The membership list of an eDiscovery case was changed. This activity is logged when all members are replaced with a group of new users. If a single member is added or removed, the **Add-ComplianceCaseMember** or **Remove-ComplianceCaseMember** operation is logged.
Created content search	New-ComplianceSearch	A new content search was created.
Deleted content search	Remove-ComplianceSearch	An existing content search was deleted.
Changed content search	Set-ComplianceSearch	An existing content search was changed. Changes can include adding or removing content locations that are searched and editing the search query.
Started content search	Start-ComplianceSearch	A content search was started. When you create or change a content search by using the Microsoft Purview portal GUI, the search is automatically started. If you create or change a search by using the **New-ComplianceSearch** or **Set-ComplianceSearch** cmdlet, you have to run the **Start-ComplianceSearch** cmdlet to start the search.
Stopped content search	Stop-ComplianceSearch	A content search that was running was stopped.
Created content search action	New-ComplianceSearchAction	A content search action was created. Content search actions include previewing search results, exporting search results, preparing search results for analysis in eDiscovery (Premium), and permanently deleting items that match the search criteria of a content search.

Friendly name	Operation (cmdlet)	Description
Deleted content search action	Remove-ComplianceSearchAction	A content search action was deleted.
Created search permissions filter	New-ComplianceSecurityFilter	A search permissions filter was created.
Deleted search permissions filter	Remove-ComplianceSecurityFilter	A search permissions filter was deleted.
Changed search permissions filter	Set-ComplianceSecurityFilter	A search permissions filter was changed.
Created eDiscovery administrator	Add-eDiscoveryCaseAdmin	A user was added as an eDiscovery Administrator in your organization.
Deleted eDiscovery administrator	Remove-eDiscoveryCaseAdmin	An eDiscovery Administrator was deleted from your organization.
Changed eDiscovery administrator membership	Update-eDiscoveryCaseAdmin	The list of eDiscovery Administrators in your organization was changed. This activity is logged when the list of eDiscovery Administrators is replaced with a group of new users. If a single user is added or removed, the **Add-eDiscoveryCaseAdmin** or **Remove-eDiscoveryCaseAdmin** operation is logged.
(none)	Get-ComplianceCase	This activity is logged when a user viewed a list of eDiscovery (Standard) or eDiscovery (Premium) cases. This activity is also logged when a user views a specific case in eDiscovery (Standard). When a user views a specific case, the audit record includes the identity of the case that was viewed. If the user only viewed a list of cases, the audit record doesn't contain a case identity.
(none)	Get-ComplianceSearch	This activity is logged when a user viewed a list of Content searches or searches associated with an eDiscovery (Standard) case. This activity is also logged when a user views a specific Content search or views a specific search associated with an eDiscovery (Standard) case. When a user views a specific search, the audit record includes the identity of the search that was viewed. If the user only viewed a list of searches, the audit record doesn't contain a search identity.

Friendly name	Operation (cmdlet)	Description
(none)	Get-ComplianceSearchAction	This activity is logged when a user viewed a list of compliance search actions (such as exports, previews, or purges) or actions associated with an eDiscovery (Standard) case. This activity is also logged when a user views a specific compliance search action (such as an export) or views a specific action associated with an eDiscovery (Standard) case. When a user views a search action, the audit record includes the identity of the search action that was viewed. If the user only viewed a list of actions, the audit record doesn't contain an action identity.

Detailed properties for eDiscovery activities

The following table describes the properties that are included on the flyout page for an eDiscovery activity listed in the search results. These properties are also included in the CSV file when you export the audit log search results. An audit log record for an eDiscovery activity won't include every detailed property listed below.

> **Tip**
>
> When you export the search results, the CSV file contains a column named **AudtiData**, which contains the detailed properties described in the following table in a multi-value property. You can use the Power Query feature in Excel to split this column into multiple columns so that each property will have its own column. This will let you sort and filter on one or more of these properties. For more information, see Search the audit log.

Expand table

Property	Description
Case	The identity (GUID) of the eDiscovery case that was created, changed, or deleted.
ClientApplication	eDiscovery cmdlet activities have a value of **EMC** for this property. This indicates the activity was performed by using the Microsoft Purview portal GUI or running the cmdlet in PowerShell.
ClientIP	The IP address of the device that was used when the activity was logged. The IP address is displayed in either an IPv4 or IPv6 address format.

Property	Description
ClientRequestId	For eDiscovery activities, this property is typically blank.
CmdletVersion	The build number for the version of the Microsoft Purview portal running in your organization.
CreationTime	The date and time in Coordinated Universal Time (UTC) when the eDiscovery activity was completed.
EffectiveOrganization	The name of the Microsoft 365 organization.
ExchangeLocations	The Exchange Online mailboxes that are included in a content search or placed on hold in an eDiscovery case.
Exclusions	Mailbox or site locations that are excluded from a content search or a hold in an eDiscovery case.
ExtendedProperties	Additional properties from a content search, a content search action, or hold in an eDiscovery case, such as the object GUID and the corresponding cmdlet and cmdlet parameters that were used when the activity was performed.
Id	The ID of the report entry. The ID uniquely identifies the audit log entry.
NonPIIParameters	A list of the parameters (without any values) that were used with the cmdlet identified in the Operation property. The parameters listed in this property are the same as those listed in the Parameters property.
ObjectId	The GUID or name of the object (for example, a Content search or an eDiscovery (Standard) case) that was created, accessed, changed, or deleted by the activity listed in the Operation property. This object is also identified in the Item column in the audit log search results.
ObjectType	The type of eDiscovery object that the user created, deleted, or modified; for example, a content search action (preview, export, or purge), an eDiscovery case, or a content search.
Operation	The name of the operation that corresponds to the eDiscovery activity that was performed.
OrganizationId	The GUID for your Microsoft 365 organization.
Parameters	The name and value for the parameters that were used with the corresponding cmdlet.
PublicFolderLocations	The public folder locations in Exchange Online that are included in a content search or placed on hold in an eDiscovery case.
Query	The search query associated with the activity, such as a content search or a query-based hold.

Property	Description
RecordType	The type of operation indicated by the record. The value of **18** indicates an event related to an activity listed in the eDiscovery cmdlet activities section. A value of **24** indicates an event related to an activity listed in the How to search for and view eDiscovery activities section.
ResultStatus	Indicates whether the action (specified in the Operation property) was successful or not.
SecurityComplianceCenterEventType	Indicates that the activity was a Microsoft Purview portal event. All eDiscovery activities will have a value of **0** for this property.
SharepointLocations	The SharePoint Online sites that are included in a content search or placed on hold in an eDiscovery case.
StartTime	The date and time in Coordinated Universal Time (UTC) when the eDiscovery activity was started.
UserId	The user who performed the activity (specified in the Operation property) that resulted in the record being logged. Records for eDiscovery activity performed by system accounts (such as NT AUTHORITY\SYSTEM) are also included in the audit log.
UserKey	An alternative ID for the user identified in the UserId property. For eDiscovery activities, the value for this property is typically the same as the UserId property.
UserServicePlan	The subscription used by your organization. For eDiscovery activities, this property is typically blank.
UserType	The type of user that performed the operation. The following values indicate the user type. 0 A regular user. 2 An administrator in your organization. 3 A Microsoft datacenter administrator or datacenter system account. 4 A system account. 5 An application. 6 A service principal.
Version	Indicates the version number of the activity (identified by the Operation property) that's logged.
Workload	The service where the activity occurred. For eDiscovery activities, the value is **SecurityComplianceCenter**.

Last updated on 02/12/2025

eDiscovery solution series: Data spillage scenario - Search and purge

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

What is data spillage and why should you care? Data spillage is when a confidential document is released into an untrusted environment. When a data spillage incident is detected, it's important to quickly assess the size and locations of the spillage, examine user activities around it, and then permanently purge the spilled data from the system.

> **Tip**
>
> You can also use the purge feature in Data Security Investigations to identify and permanently delete Exchange mailbox and Microsoft Teams items associated with data spillage incidents. Data Security Investigations provides a streamlined search and purge workflow with a purge queue dashboard to help you manage and track purge queries across your organization. For more information, see Purge in Data Security Investigations.

Data spillage scenario

You're a lead information security officer at Contoso. You are informed of a data spillage situation where an employee unknowingly shared a highly confidential document with multiple people through email. You want to quickly assess who received this document internally and externally. Once identified, you would like to share case findings with other investigators to review, and then permanently remove the data from Office 365. After the investigation is complete, you want to generate a report with the evidence of permanent removal and other case details for any future reference.

Scope of this article

This document provides a list of instructions on how to permanently remove a message from Microsoft 365 so that it's not accessible or recoverable. To delete a message and make it recoverable until the deleted item retention period expires, see Search for and delete email messages in your organization.

Workflow for managing data spillage incidents

Here's a how to manage a data spillage incident:



Step 1: Manage who can access the case and set compliance boundaries (Optional)
Step 2: Create an eDiscovery case
Step 3: Search for the spilled data
Step 4: Review and validate case findings
Step 5: Use message trace log to check how spilled data was shared
Step 6: Prepare the mailboxes
Step 7: Permanently delete the spilled data
Step 8: Verify, provide a proof of deletion, and audit

Things to know before you start

- The data spillage workflow described in this article doesn't delete chat messages in Microsoft Teams. To search for and delete Teams chat messages, see Search and purge chat messages in Teams.
- When a mailbox is on hold, a deleted message remains in the Recoverable Items folder until the retention period expires or the hold is released. Step 6 describes how to remove hold from the mailboxes. Check with your records management or legal departments before removing the hold. Your organization might have a policy that defines whether a mailbox on hold or a data spillage incident takes priority.
- To control which user mailboxes a data spillage investigator can search and manage who can access the case, you can set up compliance boundaries and create a custom role group, which is described in Step 1. To do this, you have to be a member of the Organization Management role group or be assigned the role management role. If you or an

administrator in your organization has already set compliance boundaries, you can skip Step 1.

- To create a case, you must be a member of the eDiscovery Manager role group or be a member of a custom role group that's assigned the Case Management role. If you're not a member, ask a Microsoft 365 administrator to add you to the eDiscovery manager role group.
- To create and run a Content Search, you have to be a member of the eDiscovery Manager role group or be assigned the Compliance Search management role. To delete messages, you have to be a member of the Organization Management role group or be assigned the Search And Purge management role. For information about adding users to a role group, see Assign eDiscovery permissions.
- To search the audit log eDiscovery activities in Step 8, auditing must be turned on for your organization. You can search for activities that were performed within the last 90 days. To learn more about how to enable and use auditing, see the Auditing the data spillage investigation process section in Step 8.

(Optional) Step 1: Manage who can access the case and set compliance boundaries

Depending on your organizational practice, you need to control who can access the eDiscovery case used to investigate a data spillage incident and set up compliance boundaries. The easiest way to do this is to add investigators as members of an existing role group in the Microsoft Purview portal and then add the role group as a member of the eDiscovery case. For information about the built-in eDiscovery role groups and how to add members to an eDiscovery case, see Assign eDiscovery permissions.

You can also create a new role group that aligns with your organizational needs. For example, you might want a group of data spillage investigators in the organization to access and collaborate on all data spillage cases. You can do this by creating a "Data Spillage Investigator" role group, assigning the appropriate roles (Export, RMS Decrypt, Review, Preview, Compliance Search, and Case Management), adding the data spillage investigators to the role group, and then adding the role group as a member of the data spillage eDiscovery case. See Set up compliance boundaries for eDiscovery investigations in Office 365 for detailed instructions on how to do this.

Step 2: Create an eDiscovery case

An eDiscovery case provides an effective way to manage your data spillage investigation. You can add members to the role group that you created in Step 1, add the role group as a member of

new an eDiscovery case, perform iterative searches to find the spilled data, export a report to share, track the status of the case, and then refer back to the details of the case if needed. Consider establishing a naming convention for eDiscovery cases used for data spillage incidents, and provide as much information as you can in the case name and description so you can locate and refer to in the future if necessary.

To create a new case, you can use eDiscovery in the Microsoft Purview portal. See "Create a new case" in Get started with eDiscovery (Standard).

Step 3: Search for the spilled data

Now that you've created a case and managed access, you can use the case to iteratively search to find the spilled data and identify the mailboxes that contain the spilled data. You will use the same search query that you used to find the email messages to delete those same messages in Step 7.

To create a content search associated with an eDiscovery case, see Search for content in an eDiscovery (Standard) case.

> **Important**
>
> The keywords that you use in the search query may contain the actual spilled data that you're searching for. For example, if you searching for documents containing a social security number and you use the it as search keyword, you must delete the query afterwards to avoid further spillage. See **Deleting the search query** in Step 8.

Step 4: Review and validate case findings

After you create a content search, you need to review and validate that the search results and verify that they consist only of the email messages that must be deleted. In a content search, you can preview a random sampling of 1,000 email messages without exporting the search results to avoid further data spillage.

If you have more than 1,000 mailboxes or more than 100 email messages per mailbox to review, you can divide the initial search into multiple searches by using additional keywords or conditions such as date range or sender/recipient and review the results of each search individually. Make sure to note down all search queries to use when you delete messages in Step 7.

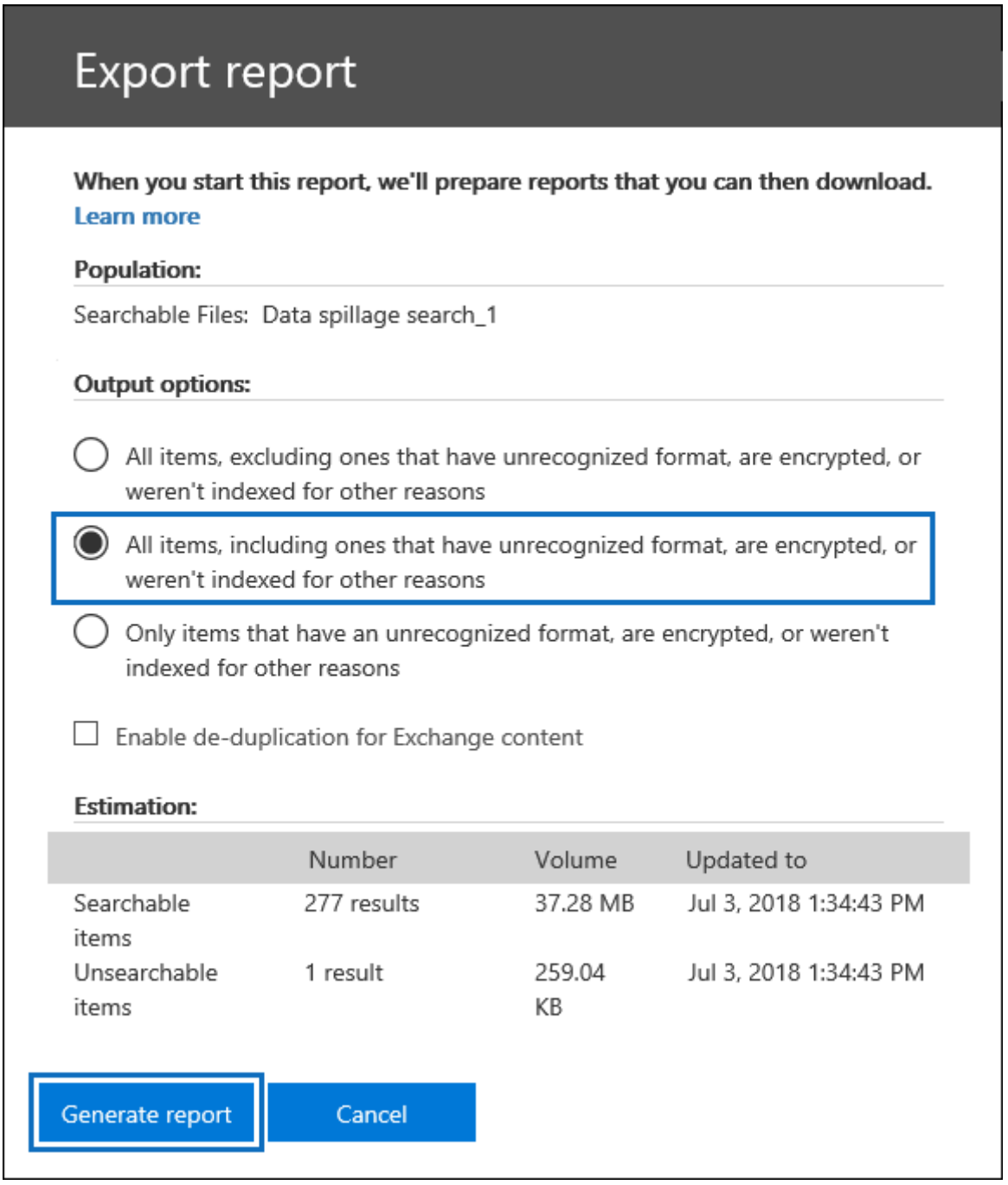
When you find an email message that contains spilled data, check the recipients of the message to determine if it was shared externally. To further trace a message, you can collect sender information and date ranges so you can use the message trace logs. This process is described in Step 5.

After you verified the search results, you may want to share your findings with others for a secondary review. People who you assigned to the case in Step 1 can review the case content in both eDiscovery and Microsoft Purview eDiscovery (Premium) and approve case findings. You can also generate a report without exporting the actual content. You can also use this same report as a proof of deletion, which is described in Step 8.

To generate a statistical report:

1. Go to the **Search** page in the eDiscovery case, and select the search that you want to generate a report for.
2. On the flyout page, select **More** > **Export report**.

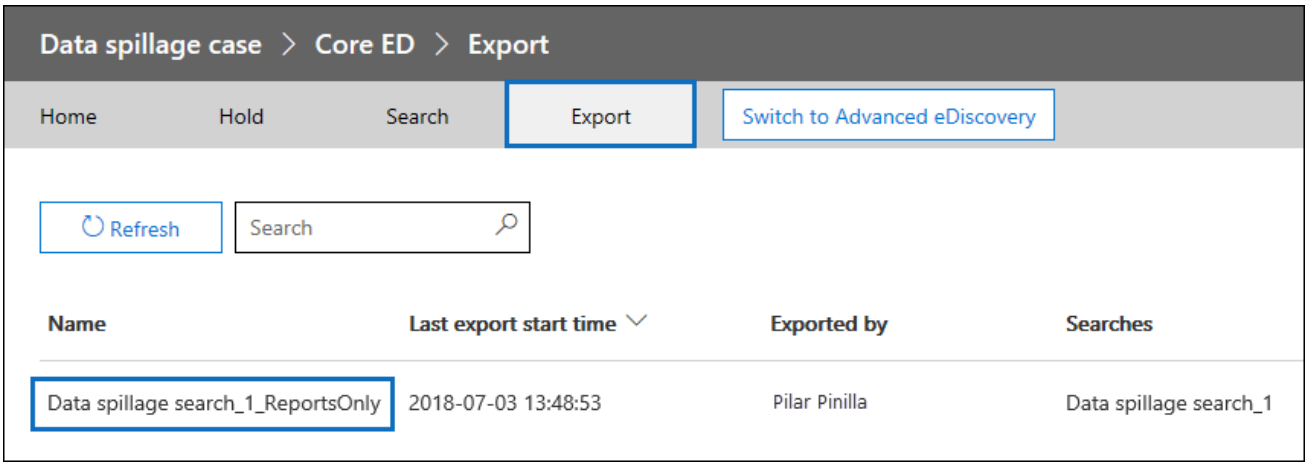
   The Export report page is displayed.



3. Select **All items, including ones that have unrecognized format, are encrypted, or weren't indexed for other reasons** and then select **Generate report**.

4. In the eDiscovery case, select **Export** to display the list of export jobs. You may have to select **Refresh** to update the list to display the export job you created.

5. Select the export job, and then select **Download** report on the flyout page.



The **Export Summary** report contains the number of locations found with results and the size of the search results. You can use this to compare with the report generated after deletion and provide as a proof of deletion. The **Results** report contains a more detailed summary of the search results, including the subject, sender, recipients, if the email was read, dates, and size of each message. If any of the details in this report contains that actual spilled data, be sure to permanently delete the Results.csv file when the investigation is complete.

Step 5: Use message trace log to check how spilled data was shared

To further investigate if email with spilled data was shared, you can optionally query the message trace logs with the sender information and the date range information that you gathered in Step 4. The retention period for message trace is 30 days for real-time data and 90 days for historical data.

You can use Message trace in the Microsoft Purview portal or use the corresponding cmdlets in Exchange Online PowerShell. It's important to note that message tracing doesn't offer full guarantees on the completeness of data returned. For more information about using Message trace, see:

- Message trace in the Security & Compliance Center
- New Message Trace in Security & Compliance Center

Step 6: Prepare the mailboxes

After you review and validate that the search results contain only the messages that must be deleted, you need to collect a list of the email addresses of the impacted mailboxes to use in Step 7 when you delete the spilled data. You may also have to prepare the mailboxes before you can

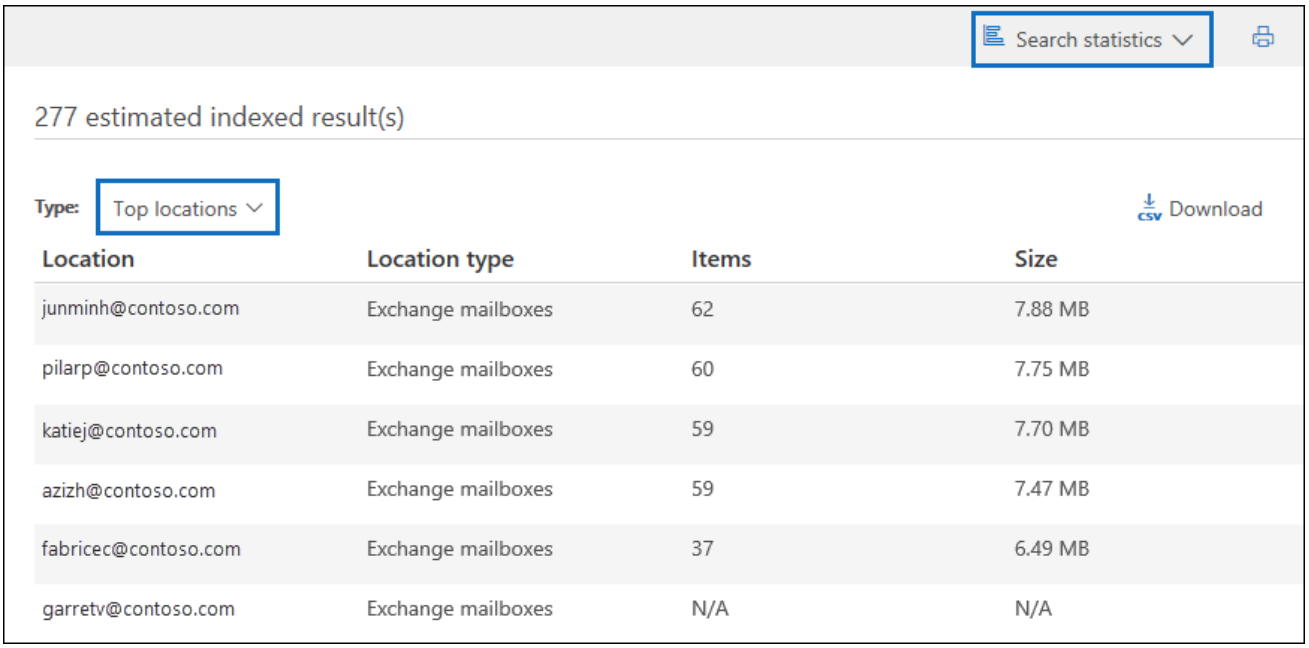
permanently delete email messages depending on whether single item recovery is enabled on the mailboxes that contain the spilled data or if any of those mailboxes are on hold.

Get a list of addresses of mailboxes with spilled data

There are two ways to collect a list of email addresses of mailboxes with spilled data.

Option 1: Get a list of addresses of mailboxes with spilled data

1. Open the eDiscovery case, go to the **Search** page and select the appropriate content search.
2. On the flyout page, select **View results**.
3. In the **Individual results** drop down list, select **Search statistics**.
4. In the **Type** drop down list, select **Top locations**.



Location	Location type	Items	Size
junminh@contoso.com	Exchange mailboxes	62	7.88 MB
pilarp@contoso.com	Exchange mailboxes	60	7.75 MB
katiej@contoso.com	Exchange mailboxes	59	7.70 MB
azizh@contoso.com	Exchange mailboxes	59	7.47 MB
fabricec@contoso.com	Exchange mailboxes	37	6.49 MB
garretv@contoso.com	Exchange mailboxes	N/A	N/A

   A list of mailboxes that contain search results is displayed. The number of items in each mailbox that match the search query is also displayed.

5. Copy the information in the list and save it to a file or select **Download** to download the information to a CSV file.

Option 2: Get mailbox locations from the export report

Open the Export Summary report that you downloaded in Step 4. In the first column in the report, the email address of each mailbox is listed under **Locations**.

Prepare the mailboxes so you can delete the spilled data

If single item recovery is enabled or if a mailbox is placed on hold, a permanently deleted (purged) message will be retained in Recoverable Items folder. So before you can purge spilled data, you need to check the existing mailbox configurations and disable single item recovery and remove any hold or retention policy. Keep in mind that you can prepare one mailbox at a time, and then run the same command on different mailboxes or create a PowerShell script to prepare multiple mailboxes at the same time.

- See "Step 1: Collect information about the mailbox" in Delete items in the Recoverable Items folder of cloud-based mailboxes on hold for instructions about how to check if single item recovery is enabled or if the mailbox is placed on hold or it's assigned to a retention policy.
- See "Step 2: Prepare the mailbox" in Delete items in the Recoverable Items folder of cloud-based mailboxes on hold for instructions about disabling single item recovery.
- See "Step 3: Remove all holds from the mailbox" in Delete items in the Recoverable Items folder of cloud-based mailboxes on hold for instructions about how to remove a hold or retention policy from a mailbox.
- See "Step 4: Remove the delay hold from the mailbox" in Delete items in the Recoverable Items folder of cloud-based mailboxes on hold for instructions about removing the delay hold that is placed on the mailbox after any type of hold is removed.

> **Important**
>
> Check with your records management or legal departments before removing a hold or retention policy. Your organization may have a policy that defines whether a mailbox on hold or a data spillage incident takes priority.

Be sure to revert the mailbox to previous configurations after you verify that the spilled data has been permanently deleted. See the details in Step 7.

Step 7: Permanently delete the spilled data

Using the mailbox locations that you collected and prepared in Step 6 and the search query that was created and refined in Step 3 to find email messages that contain the spilled data, you can now permanently delete the spilled data. As previously explained, to delete messages, you have to be a member of the Organization Management role group or be assigned the Search And Purge management role. For information about adding users to a role group, see Assign eDiscovery permissions.

To delete the spilled messages, see Search for and delete email messages.

Keep the following limits in mind when deleting spilled data:

- The maximum number of mailboxes in a search that you can use to delete items by doing a search and purge action is 50,000. If the search that you create in Step 3 searches more than 50,000 mailboxes, the purge action will fail. Searching more than 50,000 mailbox in a single search might typically happen when you configure the search to include all mailboxes in your organization. This restriction still applies even when less than 50,000 mailboxes contain items that match the search query.
- A maximum of 10 items per mailbox can be removed at one time. Because the capability to search for and remove messages is intended to be an incident-response tool, this limit helps ensure that messages are quickly removed from mailboxes. This feature isn't intended to clean up user mailboxes.

> **Important**
>
> Email items in a review set in an eDiscovery (Premium) case can't be deleted by using the procedures in this article. That's because items in a review set are copies of items in the live service that are copied and stored in an Azure Storage location. This means they won't be returned by a content search that you create in Step 3. To delete items in a review set, you have to delete the eDiscovery (Premium) case that contains the review set. For more information, see Close or delete an eDiscovery (Premium) case.

Step 8: Verify, provide a proof of deletion, and audit

The final step in the workflow to manage a data spillage incident is to verify that the spilled data was permanently removed from the mailbox by going to the eDiscovery case and rerunning the same search query that was used to delete that data to confirm that no results are returned. After you confirm the spilled data has been permanently removed, you can export a report and include it (along with the original report) as a proof of deletion. Then you can close the case which will allow you to reopen it if you have to refer to it in the future. Additionally, you can also revert mailboxes to their previous state, delete the search query used to find the spilled data, and search for auditing records of tasks performed when managing the data spillage incident.

Reverting the mailboxes to their previous state

If you changed any mailbox configuration in Step 6 to prepare the mailboxes before the spilled data was deleted, you will need to revert them to their previous state. See "Step 6: Revert the

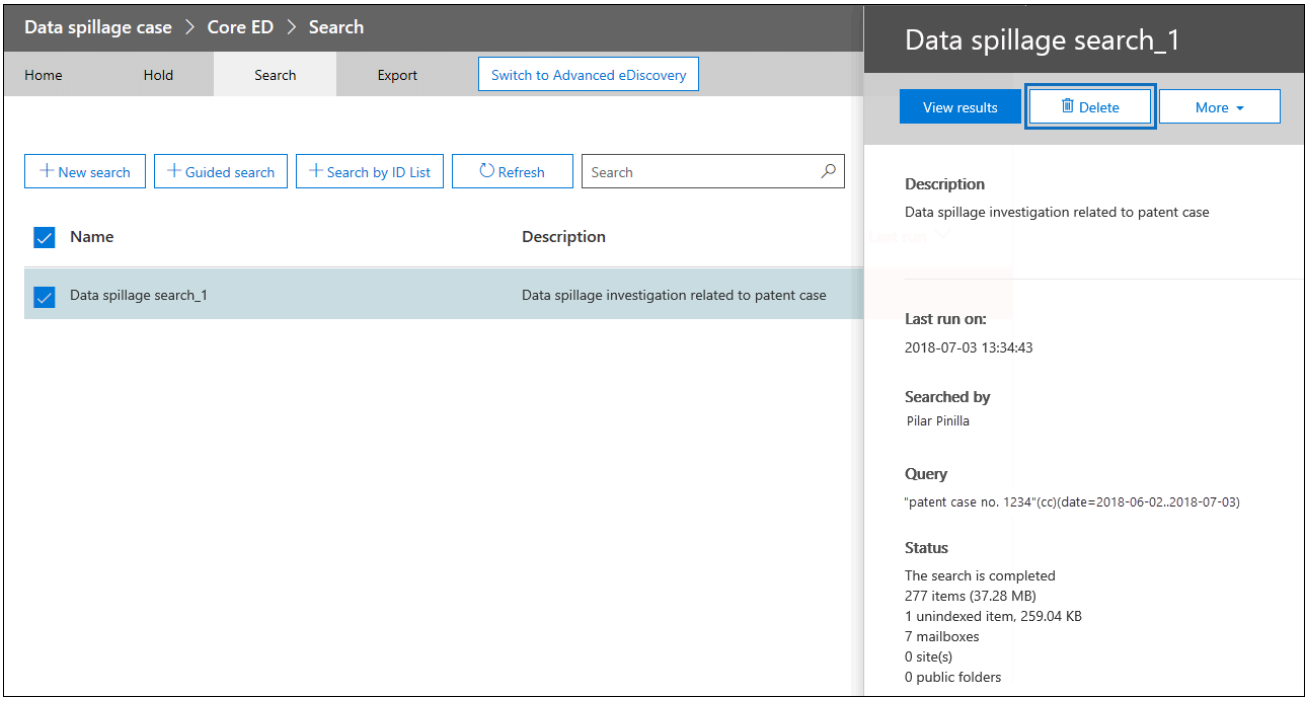
mailbox to its previous state" in Delete items in the Recoverable Items folder of cloud-based mailboxes on hold.

Deleting the search query

If the keywords in the search query that you created and used in Step 3 contains some of all of the actual spilled data, you should delete the search query to prevent further data spillage.

1. In the Microsoft Purview portal, open the eDiscovery case, go to the **Search** page, and select the appropriate content search.

2. On the flyout page, select **Delete**.



Auditing the data spillage investigation process

You can search the audit log for the eDiscovery activities that were performed during the investigation. You can also search the audit log to return the audit records for the **New-ComplianceSearchAction -Purge** command that you ran in Step 7 to delete the spilled data. For more information, see:

- Search the audit log
- Search for eDiscovery activities in the audit log

Last updated on 09/02/2025

Troubleshoot AzCopy in eDiscovery (Premium)

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

When loading non-Microsoft 365 data or documents for error remediation in Microsoft Purview eDiscovery (Premium), the user interface supplies an Azure AzCopy command that contains parameters with the location of where the files that you want to upload are stored and the Azure storage location that the files will be uploaded to. To upload your documents, you copy this command and then run it in a Command Prompt on your local computer.

Usually the command that's provided works when you run it. However, there may be cases when the command that's displayed won't run successfully. Here's a few possible reasons.

The supported version of AzCopy isn't installed on the local computer

You should use the latest version of AzCopy to load non-Microsoft 365 data in eDiscovery (Premium). To install the latest version, see Get started with AzCopy.

AzCopy isn't installed on the local computer or it's not installed in the default location

If AzCopy isn't installed or it's installed in a location other than the default install location (which is `%ProgramFiles(x86)%`), you may receive the following error when you run the AzCopy command:

> The system cannot find the path specified.

If AzCopy isn't installed on the local computer, see Get started with AzCopy. Be sure to install it in the default location.

If AzCopy is installed, but it's installed in a location different than the default location, you can copy the command, paste it to a text file, and then change the path to the location where AzCopy is installed. For example, if Azcopy is located in `%ProgramFiles%`, then you can change the first part of the command from `%ProgramFiles(x86)%\Microsoft SDKs\Azure\AzCopy.exe` to `%ProgramFiles%\Microsoft SDKs\Azure\AzCopy`. After you make this change, copy it from the text file, and then run it a Command Prompt.

> **Tip**
>
> If AzCopy is installed in a location other then the default install location, consider uninstalling it and then re-installing it in the default location. This will help prevent this issue in the future.

Last updated on 11/21/2025

Retirement of legacy eDiscovery tools

> **Caution**
>
> Microsoft retired all classic eDiscovery experiences on August 31, 2025. This retirement includes classic Content Search, classic eDiscovery (Standard), and classic eDiscovery (Premium).
>
> The guidance in this article only applies to organizations hosted in Microsoft 365 operated by 21Vianet (China). If your organization isn't hosted by 21Vianet, use the guidance for the new eDiscovery experience in the Microsoft Purview portal.

> **Important**
>
> The functionality of the legacy eDiscovery tools described in this article has either been removed from the Microsoft 365 service or is still available, but no longer supported. Any functionality that's still available may be removed without notice. If you're still using any of these legacy tools, consider migrating to the eDiscovery tools in the Microsoft Purview portal or one of the alternatives described in this article.

Over the years, Microsoft has provided eDiscovery tools that let you search, preview, and export email content from Exchange Online. However, these tools no longer offer an effective way to search for non-Exchange content in other Microsoft 365 services, such as SharePoint and Microsoft 365 Groups.

To address this, Microsoft offers other eDiscovery tools that help you search for a wide variety of Microsoft 365 content. And we've been working hard to incorporate the most current and powerful eDiscovery functionality in the Microsoft Purview portal. This allows organizations to respond to legal, internal, and other document requests for content across many Microsoft 365 services, including Exchange Online.

As a result of this new and improved eDiscovery functionality in the Microsoft Purview portal, the following eDiscovery-related features and functionality related to searching for email content in Exchange Online and Microsoft 365 are retired:

- In-Place eDiscovery and In-Place Holds in the Exchange admin center.

- The Exchange Online PowerShell cmdlets that support In-Place eDiscovery and In-Place Holds (these cmdlets are collectively identified as **-MailboxSearch* cmdlets). This includes the following cmdlets:
  - New-MailboxSearch
  - Start-MailboxSearch
  - Stop-MailboxSearch
  - Set-MailboxSearch

  > **Note**
  >
  > The **Remove-MailboxSearch** cmdlet will be available after the other ****-MailboxSearch*** cmdlets are retired so that you can use it to help in your transition to other eDiscovery and hold tools. However, after a certain date (cited below) Microsoft Support will no longer supports this cmdlet.

- The Search-Mailbox cmdlet in Exchange Online PowerShell.

- The following operations in the Exchange Web Services API:
  - GetSearchableMailboxes
  - SearchMailboxes
  - SetHoldOnMailboxes
  - GetHoldOnMailboxes

- Microsoft Purview eDiscovery (Premium) v1.0, which is the first version of eDiscovery (Premium) that's accessed through a Microsoft Purview eDiscovery (Standard) case in the Microsoft Purview portal. The retirement of eDiscovery (Premium) v1.0 doesn't impact your ability to create and manage eDiscovery (Standard) cases.

> **Note**
>
> The eDiscovery functionality being retired only applies to cloud-based versions of Microsoft 365 and Office 365. eDiscovery functionality in on-premises versions of Exchange and SharePoint will still be supported until further notice.

The following sections in this article provide guidance about each feature being retired. This information including timelines and alternative tools that you can use instead of the retired tool.

In-Place eDiscovery and In-Place Holds in the Exchange admin center

As per the original announcement on July 1, 2017, the In-Place eDiscovery & Hold functionality in the Exchange admin center (EAC) is being retired. The In-Place eDiscovery & Holds page in the EAC allowed you to search, hold, and export content from Exchange Online. In-Place eDiscovery also let you copy search results to a discovery mailbox so that you or other eDiscovery managers could review content and make it available for legal, regulatory, and public requests.

Because all of these capabilities (except for copying search results to a discovery mailbox) are now available in the content search, eDiscovery and eDiscovery (Premium) tools in the Micrsoft Purview portal (with improved functionality, reliability, and support for a wide range of Microsoft 365 services), we recommend that you start using these tools as soon as possible. To help you in the transition to these other eDiscovery tools, the table below lists the tools you can use instead of In-Place eDiscovery and In-Place Hold.

> **Note**
>
> When migrating mailboxes from on-premises Exchange to Exchange Online, you need to create new holds in the Microsoft Purview portal for mailboxes with holds to preserve the data. Legacy holds in on-premises Exchange organizations aren't migrated.

Scope of affected organizations

- Office 365 and Microsoft 365 Enterprise organizations
- Office 365 and Microsoft 365 Education organizations
- Office 365 and Microsoft 365 Government organizations; this includes GCC, GCC High, and DoD
- Office 365 and Microsoft 365 operated by 21Vianet (China)
- Office 365 and Microsoft 365 Germany

Timeline for retirement

- July 1, 2020: You won't be able to create new searches and holds, but you can still run, edit, and delete existing searches at your own risk. Microsoft Support will no longer In-Place eDiscovery & Holds in the EAC.
- October 1, 2020: The In-Place eDiscovery & Holds functionality in the EAC was placed in a read-only mode. This means you'll only be able to remove existing searches and holds.

Alternative tools

The following table describes other tools that you can use to replace the existing functionality that's being retired.

Expand table

Functionality	Alternative tool	Comments
Search, export, and hold for legal purposes	eDiscovery (Standard) cases in the Microsoft Purview portal	Using the capabilities of eDiscovery (Standard) cases provide the functional parity to In-Place eDiscovery and In-Place Holds. This includes the following: • Search scales to millions of locations • Higher reliability for searching, exporting, and placing content on hold • Searching for content in for Exchange Online, SharePoint, OneDrive, Skype for Business, Microsoft Teams, Viva Engage Groups, Microsoft 365 Groups, and other content stored in Office 365 applications
Hold for retention purposes	Retention policies in the Microsoft Purview portal	You can use Retention policies to retain content and, if desired, delete it after the retention period expires. Other capabilities include: • Applying policies to your entire organization • Applying policies to specific content locations such as Exchange Online, SharePoint, OneDrive, Skype for Business, Microsoft Teams, and Office 365 Groups • Applying policies to specific users For more information, see Learn about retention policies and retention labels.
Copy email search results to a discovery mailbox for review	Review sets in eDiscovery (Premium) v2.0	Reviewing content in Microsoft 365 has never been easier. Review sets have many great capabilities to help reduce the amount of time and data needed to review, including: • Perform fast search queries and filter content in a review set • Analyze content in a review set; this includes email threading, near-duplicate detection, Themes analysis, and Predictive coding • Tag documents in a review set • Tagging suggestions to help identify attorney client privilege content For more information, see Overview of the eDiscovery (Premium) solution in Microsoft 365.

Functionality	Alternative tool	Comments
		Alternatively, you can export search results to PST files and then use Microsoft 365 Import service to import the PSTs to a discovery mailbox. For step-by-step instruction, see Use network upload to import PST files to Office 365.
Copy messages from one mailbox to a different mailbox	Assign permissions to a mailbox	To give a person access to another user's email (such as when an employee leaves your organization and you need to give another person access to the former employee's email), we recommended that you assign that person permissions to access the former employee's mailbox. So instead of copying mailbox items to another user mailbox or a shared mailbox, just assign a user the permissions necessary to access the source mailbox.
Restore items from the Recoverable Items folder	Restore-RecoverableItems	You can restore permanently deleted items (also known as *soft-deleted* items) in mailboxes, as long as the deleted item retention period for an item hasn't expired. For more information, see Recoverable Items folder in Exchange Online.

FAQs about In-Place eDiscovery and In-Place Holds

I use the copy search results functionality of In-Place eDiscovery & Holds in the EAC to copy search results to a discovery mailbox for review by attorneys. What options do I have now?

There are two ways to replicate this functionality today. The first is to use review sets in eDiscovery (Premium) v2.0. Review sets have many of the same capabilities you see in a traditional review tool like fast search of documents, tagging, email threading, near duplicate grouping, themes analysis, and predictive coding. If you still want to use discovery mailboxes for review, the second option is to export search results to PST files and then import the PST files to a discovery mailbox by using the PST import feature in the Microsoft compliance center.

How do I control which content locations (such as mailboxes or sites) that can an eDiscovery manager can search using the new tools?

The Microsoft Purview portal also uses compliance boundaries to control which content locations an eDiscovery Manager can search. Compliance boundaries are useful in government entities that need to stay within agency boundaries or multi-national corporations required to respect geographical boarders.

*-MailboxSearch cmdlets

As per the original notice announced on July 1, 2017 in the Exchange admin center, the In-Place eDiscovery & Hold functionality and the corresponding ***-MailboxSearch** cmdlets are being retired. These cmdlets provide users the ability to search, hold, and export mailbox content for legal, regulatory, and public requests.

Because these capabilities are now available in the Microsoft Purview portal and Office 365 Security & Compliance PowerShell with improved performance and scalability, you should using these improved cmdlets. These cmdlets include *-ComplianceCase, *-ComplianceSearch, *-CaseHoldPolicy, *-CaseHoldRule, and *-ComplianceSearchAction.

Scope of affected organizations

- Office 365 and Microsoft 365 Enterprise organizations
- Office 365 and Microsoft 365 Education organizations
- Office 365 and Microsoft 365 Government organizations; this includes GCC, GCC High, and DoD
- Office 365 and Microsoft 365 operated by 21Vianet (China)
- Office 365 Germany

Timeline

- July 1, 2020: You won't be able to use **New-MailboxSearch** to create new In-Place eDiscovery searches and In-Place Holds, but you can still use cmdlets to run, edit, and delete existing searches and holds at your own risk. Microsoft Support will no longer provide assistance for these types of searches and holds.
- October 1, 2020: The In-Place eDiscovery & Holds functionality in the EAC was placed in a read-only mode. You won't be able to use the **New-MailboxSearch**, **Start-MailboxSearch**, or **Set-MailboxSearch** cmdlets. You'll only be able to get and remove existing searches and holds.

Alternative tools

The following table describes other tools that you can use to replace the existing functionality that's being retired.

Expand table

Functionality	Alternative tools	Comments
Search and export	*-ComplianceSearch *-ComplianceSearchAction *-ComplianceCase	The ComplianceSearch and ComplianceSearchAction cmdlets work together to help you search and export content. You can create a new search and view the search estimate by using the **New-**, **Get-**, and **Start-ComplianceSearch** cmdlets. Then you can use the **New-ComplianceSearchAction** cmdlet to export the search results. You'll still have to use the eDiscovery (Standard) tool in the Microsoft Purview portal to download those search results to your local computer. **Note:** If you use these cmdlets to create searches that aren't associated with an eDiscovery (Standard) case, these searches will be located on the **Content search** page in the Microsoft Purview portal.
Hold content in a mailbox	*-CaseHoldPolicy *-CaseHoldRule *-ComplianceCase	Holds in the Microsoft Purview portal must be associated with a ComplianceCase. First, create the compliance case, and then create a CaseHoldPolicy and a CaseHoldRule. **Note:** Creating a CaseHoldPolicy without a creating CaseHoldRule will render the hold inoperable until the CaseHoldRule is created and associated to the CaseHoldPolicy. See the cmdlet documentation for more information.
Copy search results to a discovery mailbox	None	There's no direct replacement for this functionality because it does not provide access to all Microsoft 365 services. See the following FAQ below for alternative solutions.
Copy messages from one mailbox to a different mailbox	Assign permissions to a mailbox	To give a person access to another user's email (such as when an employee leaves your organization and you need to give another person access to the former employee's email), we recommended that you assign that person permissions to access the former employee's mailbox. So instead of copying mailbox items to another user mailbox or a shared mailbox, just assign a user permissions to access the source mailbox.

FAQs about *-MailboxSearch cmdlets

We use Copy Search to export email messages or instant Messages for purposes other eDiscovery and legal investigations. What other options do we have after these cmdlets are retired?

The Microsoft Graph APIs provide a number of methods for extracting data for analysis and other purposes that are far more resilient and scalable than the using the ***-MailboxSearch** cmdlets.

After the cmdlets are retired, will I still be able to remove or retrieve searches?

Yes, although we're removing the ability to create and modify searches, you'll still be able to use the **Remove-MailboxSearch** until further notice. However, the use of this cmdlet will be at your own risk after the retirement dates and Microsoft Support will no longer be able to provide assistance.

Search-Mailbox cmdlet

The **Search-Mailbox** cmdlet in Exchange Online PowerShell has been retired as originally announced in a warning in the cmdlet output starting back in 2018. The **Search-Mailbox** cmdlet was originally used to search a user's mailbox and purge malicious content. We recommend that you start using the **New-ComplianceSearch** and **New-ComplianceSearchAction** cmdlets in Office 365 Security & Compliance PowerShell to search for and purge content. For a built-in security experience, the <Microsoft 365 security features provide robust threat protection for email and many other Microsoft services.

Scope of affected organizations

- Office 365 and Microsoft 365 Enterprise organizations
- Office 365 and Microsoft 365 Education organizations
- Office 365 and Microsoft 365 Government organizations; this includes GCC, GCC High, and DoD
- Office 365 and Microsoft 365 operated by 21Vianet (China)
- Office 365 Germany

Timeline

April, 2024: The **Search-Mailbox** cmdlet is no longer be available and Microsoft Support will no longer provide assistance.

Alternative tools

The following table describes other tools that you can use to replace the existing functionality that's being retired.

Expand table

Functionality	Alternative tools	Comments
Search a mailbox	*-ComplianceSearch *-ComplianceSearchAction	The ComplianceSearch and ComplianceSearchAction cmdlets work together to help you search content. You can create a new search and view the search estimate by using the **New-**, **Get-**, and **Start-ComplianceSearch** cmdlets.
Delete bulk email from a mailbox	Set up an archive and deletion policy for mailboxes	Admins can create an archiving and deletion policy that automatically moves items to a user's archive mailbox and automatically deletes items from the mailbox.
Copy search results to a discovery mailbox		There's no direct replacement for this functionality because it does not provide access to all Microsoft 365 services. See the FAQs in the ***-MailboxSearch cmdlets** section for alternative solutions.
Copy messages from one mailbox to a different mailbox	Assign permissions to a mailbox	To give a person access to another user's email (such as when an employee leaves your organization and you need to give another person access to the former employee's email), we recommended that you assign that person permissions to access the former employee's mailbox. So instead of copying mailbox items to another user mailbox or a shared mailbox, just assign a user permission to access the source mailbox.
Purge messages from a mailbox	*-ComplianceSearch *-ComplianceSearchAction	The ComplianceSearch and ComplianceSearchAction cmdlets work together to help you search and purge content. You can create and run a search with **New-ComplianceSearch** and **New-ComplianceSearch** cmdlets, and then you can purge the content by using **New-ComplianceSearchAction -Purge -PurgeType** command. For more information, see Search for and delete messages.
Purge messages from a mailbox	Assign permissions to a mailbox	To purge messages from a mailbox, assign an administrator permissions to access the employee's mailbox. Messages can be deleted and recycled as needed taking advantage of the built-in search and viewing capabilities in Outlook.

Exchange Web Services API operations

These operations in the Exchange Web Services API are used by the In-Place eDiscovery & Holds feature in the Exchange admin center and the corresponding ***-MailboxSearch** cmdlets in

Exchange Online PowerShell. They will also be retired as part of retiring the other legacy eDiscovery tools.

Scope of affected organizations

- Office 365 and Microsoft 365 Enterprise organizations
- Office 365 and Microsoft 365 Education organizations
- Office 365 and Microsoft 365 Government organizations; this includes GCC, GCC High, and DoD
- Office 365 Germany

Timeline

July 1, 2020: The GetSearchableMailboxes, SearchMailboxes, SetHoldOnMailboxes, and GetHoldOnMailboxes operations will no longer be available, and Microsoft Support will no longer provide assistance.

eDiscovery (Premium) v1.0

eDiscovery (Premium) v1.0, which is the version of eDiscovery (Premium) available in an eDiscovery (Standard) case by selecting **Switch to eDiscovery (Premium)**, is being retired. Its functionality has been replaced by the new eDiscovery (Premium) solution in the Microsoft Purview portal.

To determine if your organization is using eDiscovery (Premium) v1.0:

1. Go to the Microsoft Purview portal, select **eDiscovery (Standard)**, and open an eDiscovery (Standard) case.

2. If you see **Switch to eDiscovery (Premium)**, selecting it takes you to the 1.0 version of eDiscovery (Premium), which is being retired. The ability to create and manage cases in eDiscovery (Standard) won't be affected. Only the ability to add and analyze case data in eDiscovery (Premium) v1.0 (by selecting **Switch to eDiscovery (Premium)**) is being retired.

The new eDiscovery (Premium) solution in Microsoft 365 (also known as *eDiscovery (Premium) v2.0*) provides all of the capabilities of the original solution, but now includes a custodian-based approach of identifying content in other Microsoft 365 services, collecting that content, and then adding it to a review set where reviewers can take advantage of fast search queries, tagging, and analytics features to help cull relevant documents. eDiscovery (Premium) now includes improved processing and native viewers for both Microsoft and non-Microsoft file types, a full list of file

types is here and supported metadata fields are here. Also, the new eDiscovery (Premium) solution provides a powerful custodian holds management feature that lets you apply holds to content in different services, notify users of the holds, and track custodian responses, all within an eDiscovery (Premium) case.

To access eDiscovery (Premium) v2.0:

Go to the Microsoft Purview portal, select **eDiscovery** > **Advanced**, and open an eDiscovery (Standard) case.

At this time, we recommend that you begin to transition your eDiscovery workflow to the new eDiscovery (Premium) functionality. If necessary, you can archive your eDiscovery (Premium) 1.0 cases by exporting the content and storing it offline. Although you'll still be able to access eDiscovery (Premium) v1.0 in existing cases until December 31, 2020, Microsoft Support won't provide support after October 1, 2020. See the following timeline for more details.

Scope of affected organizations

- Office 365 and Microsoft 365 Enterprise organizations
- Office 365 and Microsoft 365 Education organizations
- Office 365 and Microsoft 365 Government organizations; this includes GCC, GCC High, and DoD
- Office 365 Germany

Timeline

- July 1, 2020: You won't be able to create new eDiscovery (Premium) v1.0 cases.
- October 1, 2020: You won't be able to add new data (Prepare search results for eDiscovery (Premium)) to any cases. You'll be able to continue working with data in existing cases at your own risk. Microsoft Support will no longer provide assistance.
- December 31, 2020: You won't be able to access eDiscovery (Premium) v1.0 cases.

Alternative tools

The eDiscovery (Premium) solution in the Microsoft Purview portal.

Last updated on 02/12/2025

Exchange Online mail encryption with AD RMS

> *Microsoft Purview service description*

To help prevent information leakage, Exchange Online includes Information Rights Management (IRM) functionality that provides online and offline protection of email messages and attachments. You can configure Exchange Online IRM to use on-premises Active Directory Rights Management Service (AD RMS), if needed, to satisfy your organization requirements. This is not common. If you do not have a requirement to use AD RMS, use Microsoft Purview Message Encryption instead.

IRM protection can be applied by users in Microsoft Outlook or Outlook on the web, and it can be applied by administrators using transport protection rules or Outlook protection rules. IRM helps you and your users control who can access, forward, print, or copy sensitive data within an email.

Changes to how IRM works with message encryption and Microsoft Entra ID

As of September 2017, when you set up Microsoft Purview Message Encryption for your organization, you also set up IRM for use with the Azure Rights Management service. You no longer set up IRM with Rights Management service separately. Instead, message encryption and rights management work seamlessly together. For more details about Microsoft Purview Message Encryption, see Message Encryption FAQ. If you're ready to get started using Microsoft Purview Message Encryption within your organization, see Set up Microsoft Purview Message Encryption.

How IRM works with Exchange Online and Active Directory Rights Management Services

Exchange Online IRM uses on-premises Active Directory Rights Management Services (AD RMS), an information protection technology in Windows Server 2008 and later. IRM protection is applied to email by applying an AD RMS rights policy template to an email message. Rights are attached to the message itself so that protection occurs online and offline and inside and outside of your organization's firewall.

Users can apply a template to an email message to control the permissions that recipients have on a message. Actions, such as forwarding, extracting information from a message, saving a message or printing a message can be controlled by applying an AD RMS rights policy to the message.

You can configure IRM to use an AD RMS server running Windows Server 2008 or later. You can use this AD RMS server to manage the AD RMS rights policy templates for your cloud-based organization. Outlook also relies on the AD RMS server to enable users to apply IRM protection to messages they send. For details, see Configure IRM to use an on-premises AD RMS server.

After it's enabled, IRM protection can be applied to messages as follows:

- **Users can manually apply a template using Outlook and Outlook on the web.** Users can apply an AD RMS rights policy template to an email message by selecting the template from the **Set permissions** list. When users send an IRM-protected message, any attached files that use a supported format also receive the same IRM protection as the message. IRM protection is applied to files associated with Word, Excel, and PowerPoint, as well as .xps files and attached email messages.

- **Administrators can use transport protection rules to apply IRM protection automatically to both Outlook and Outlook on the web.** You can create transport protection rules to IRM-protect messages. Configure the transport protection rule action to apply an AD RMS rights policy template to messages that meet the rule condition. After you enable IRM, your organization's AD RMS rights policy templates are available to use with the transport protection rule action called **Apply rights protection to the message with**.

- **Administrators can create Outlook protection rules.** Outlook protection rules automatically apply IRM-protection to messages in Outlook 2010 (not Outlook on the web) based on message conditions that include the sender's department, who the message is sent to, and whether recipients are inside or outside your organization. For details, see Create an Outlook Protection Rule.

Last updated on 05/14/2026

Configure IRM to use an on-premises AD RMS server

> *Microsoft Purview service description*

For use with on-premises deployments, Information Rights Management (IRM) in Exchange Online uses Active Directory Rights Management Services (AD RMS), an information protection technology in Windows Server 2008 and later. IRM protection is applied to email by applying an AD RMS rights policy template to an email message. Rights are attached to the message itself so that protection occurs online and offline and inside and outside of your organization's firewall.

This topic shows you how to configure IRM to use an AD RMS server. For information about using Microsoft Purview Message Encryption with Microsoft Entra ID and the Azure Rights Management service, see the Message encryption FAQ.

To learn more about IRM in Exchange Online, see Information Rights Management in Exchange Online.

What do you need to know before you begin?

- Estimated time to complete this task: 30 minutes
- You need to be assigned permissions before you can perform this procedure or procedures. To see what permissions you need, see the "Information Rights Management" entry in the Messaging policy and compliance permissions topic.
- The AD RMS server must be running Windows Server 2008 or later. For details about how to deploy AD RMS, see Installing an AD RMS Cluster.
- For details about how to install and configure Windows PowerShell and connect to the service, see Connect to Exchange Online PowerShell.
- For information about keyboard shortcuts that may apply to the procedures in this topic, see Keyboard shortcuts for the Exchange admin center in Exchange Online.

How do you do this?

Complete the following steps to export the trusted publishing domain, import it into Exchange Online, distribute a rights policy template, and enable IRM.

Step 1: Use the AD RMS console to export a trusted publishing domain (TPD) from an AD RMS server

The first step is to export a trusted publishing domain (TPD) from the on-premises AD RMS server to an XML file. The TPD contains the following settings needed to use RMS features:

- The server licensor certificate (SLC) used for signing and encrypting certificates and licenses
- The URLs used for licensing and publishing
- The AD RMS rights policy templates that were created with the specific SLC for that TPD

When you import the TPD, it's stored and protected in Exchange Online.

1. Open the Active Directory Rights Management Services console, and then expand the AD RMS cluster.
2. In the console tree, expand **Trust Policies**, and then click **Trusted Publishing Domains**.
3. In the results pane, select the certificate for the domain you want to export.
4. In the **Actions** pane, click **Export Trusted Publishing Domain**.
5. In the **Publishing domain file** box, click **Save As** to save the file to a specific location on the local computer. Type a file name, making sure to specify the `.xml` file name extension, and then click **Save**.
6. In the **Password** and **Confirm Password** boxes, type a strong password that will be used to encrypt the trusted publishing domain file. You will have to specify this password when you import the TPD to your cloud-based email organization.

Step 2: Use the Exchange Management Shell to import the TPD to Exchange Online

After the TPD is exported to an XML file, you have to import it to Exchange Online. When a TPD is imported, your organization's AD RMS templates are also imported. When the first TPD is imported, it becomes the default TPD for your cloud-based organization. If you import another TPD, you can use the `-Default` switch on the `Import-RMSTrustedPublishingDomain` cmdlet to make that TPD the default for your organization, so its templates are available to users.

To import the TPD into Exchange Online so the associated AD RMS templates become available to your organization, run the following command in Exchange Online PowerShell:

PowerShell

```
Import-RMSTrustedPublishingDomain -FileData ([System.IO.File]::ReadAllBytes('<path to exported TPD file>')) -Name "<name of TPD>" -ExtranetLicensingUrl <URL> -IntranetLicensingUrl <URL>
```

The following table describes each parameter:

Expand table

Parameter	Value to provide
FileData	The byte content of the exported TPD XML file. Specify the full file path in the `ReadAllBytes` method.
Name	A display name for the TPD in Exchange Online.
ExtranetLicensingUrl	The external licensing URL of the AD RMS cluster.
IntranetLicensingUrl	The internal licensing URL of the AD RMS cluster.

You can obtain the values for the *ExtranetLicensingUrl* and *IntranetLicensingUrl* parameters in the Active Directory Rights Management Services console. Select the AD RMS cluster in the console tree. The licensing URLs are displayed in the results pane. These URLs are used by email clients when content has to be decrypted and when Exchange Online needs to determine which TPD to use.

When you run this command, you'll be prompted for a password. Enter the password that you specified when you exported the TPD from your AD RMS server.

The following example imports an exported TPD file and specifies the internal and external licensing URLs required for AD RMS integration. In this case, the TPD XML file was exported from the AD RMS server and saved to the desktop of the Administrator account. The Name parameter is used to specify a name for the imported TPD.

PowerShell

```
Import-RMSTrustedPublishingDomain -FileData ([System.IO.File]::ReadAllBytes('C:\Users\Administrator\Desktop\ExportTPD.xml')) -Name "Exported TPD" -ExtranetLicensingUrl https://corp.contoso.com/_wmcs/licensing -IntranetLicensingUrl https://rmsserver/_wmcs/licensing
```

For detailed syntax and parameter information, see Import-RMSTrustedPublishingDomain.

How do you know that you successfully imported the TPD?

To verify that you have successfully imported the TPD, run the **Get-RMSTrustedPublishingDomain** cmdlet to retrieve TPDs in your Exchange Online organization. For details, see the examples in Get-RMSTrustedPublishingDomain.

Step 3: Use the Exchange Management Shell to distribute an AD RMS rights policy template

After you import the TPD, you must make sure an AD RMS rights policy template is distributed. A distributed template is visible to Outlook on the web (formerly known as Outlook Web App) users, who can then apply the templates to an email message.

To verify which RMS templates are available in the default TPD after the import, list all templates by running the following command:

PowerShell

```
Get-RMSTemplate -Type All | fl
```

If the value of the *Type* parameter is `Archived`, the template isn't visible to users. Only distributed templates in the default TPD are available in Outlook on the web.

To make a specific RMS template available to users in Outlook on the web, set its type to Distributed by running the following command:

PowerShell

```
Set-RMSTemplate -Identity "<name of the template>" -Type Distributed
```

For example, to distribute the Company Confidential template so that users can apply it to messages, run the following command:

PowerShell

```
Set-RMSTemplate -Identity "Company Confidential" -Type Distributed
```

For detailed syntax and parameter information, see Get-RMSTemplate and Set-RMSTemplate.

The Do Not Forward template

When you import the default TPD from your on-premises organization into Exchange Online, one AD RMS rights policy template named **Do Not Forward** is imported. By default, this template is distributed when you import the default TPD. You can't use the **Set-RMSTemplate** cmdlet to modify the **Do Not Forward** template.

When the **Do Not Forward** template is applied to a message, only the recipients addressed in the message can read the message. Additionally, recipients can't do the following:

- Forward the message to another person.
- Copy content from the message.
- Print the message.

> **Important**
>
> The **Do Not Forward** template can't prevent information in a message from being copied with third-party screen capture programs, cameras, or users manually transcribing the information

You can create additional AD RMS rights policy templates on the AD RMS server in your on-premises organization to meet your IRM protection requirements. If you create additional AD RMS rights policy templates, you have to export the TPD from the on-premises AD RMS server again and refresh the TPD in the cloud-based email organization.

How do you know that you successfully distributed the AD RMS rights policy template?

To verify that you have successfully distributed and AD RMS rights policy template, run the **Get-RMSTemplate** cmdlet to check the template's properties. For details, see the examples in Get-RMSTemplate.

Step 4: Use the Exchange Management Shell to enable IRM

After you import the TPD and distribute an AD RMS rights policy template, enable internal licensing so that Exchange Online can apply and consume IRM-protected content. Run the following command to enable IRM for your cloud-based email organization.

PowerShell

```
Set-IRMConfiguration -InternalLicensingEnabled $true
```

For detailed syntax and parameter information, see Set-IRMConfiguration.

How do you know that you successfully enabled IRM?

To verify that you have successfully enabled IRM, run the Get-IRMConfiguration cmdlet to check IRM configuration in the Exchange Online organization.

Verify that IRM is working

Verify the full IRM configuration

To verify that you have successfully imported the TPD and enabled IRM, do the following:

- Use the **Test-IRMConfiguration** cmdlet to test IRM functionality. For details, see the Test-IRMConfiguration examples.
- Compose a new message in Outlook on the web and IRM-protect it by selecting **Set permissions** option from the extended menu (•••).

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/26/2026

Microsoft Purview troubleshooting

This library of articles provides guidance to troubleshoot and fix errors that you may experience when using Microsoft Purview. For a complete list of articles, browse the navigation pane on the left, or use the search box above the pane to filter by keywords and find help with a specific issue.

Archive mailboxes

HOW-TO GUIDE

Messages aren't moved to archive mailboxes

"Mailbox is full" warning

Archive mailbox issues

eDiscovery

HOW-TO GUIDE

Resolve eDiscovery issues

Audit logs

HOW-TO GUIDE

Mailbox audit logs

Issue when searching audit log

UserAgent property

Message encryption

HOW-TO GUIDE

Can't display message

Disable Encrypt-Only

Data Loss Prevention (DLP)

HOW-TO GUIDE

Common scenarios for tips

DLP not working

Retention

HOW-TO GUIDE

MRM and retention policies

Managed folders disappear

Can't manage disposition reviews

Microsoft Compliance

Microsoft documentation and resources you need to get started with compliance for your business or organization.

Regulatory & industry compliance

Microsoft guidance on compliance with industry areas and international & domestic standards and regulations.



General Data Protection Regulation (GDPR)

Guidance to help you honor rights and fulfill obligations under the GDPR when using Microsoft products and...



Microsoft compliance offerings

Microsoft offers a comprehensive set of compliance offerings to help your organization comply wit...



Microsoft Cloud for industries

Learn how Microsoft Azure, Dynamics 365, Microsoft 365, and Microsoft Power Platform can support compliance need...



Australia and New Zealand compliance

Learn how Microsoft can help support compliance requirements in Australia and New Zealand.



Digital Operational Resilience Act (DORA)

Learn how Microsoft is preparing for the upcoming EU Digital Operational Resilience Act and supporting our FSI...



Microsoft and US government compliance

Learn how Microsoft is supporting compliance with US government requirements.

Service assurance

Use expertly curated compliance resources and documentation to support your risk assessments of Microsoft online services.

Get Started

- Foundational training on how Microsoft safeguards data
- Risk Assessment Guide for Microsoft Cloud

Artificial Intelligence (AI)

- AI overview
- Microsoft Copilot ISO 42001 Report
- Microsoft Copilot ISO 42001 Alignment
- 2024 RAI Transparency Report
- Copilot for Microsoft 365 data, privacy, and security
- Copilot for Microsoft 365 Risk Assessment QuickStart

Infrastructure

- Architecture
- Datacenter security
- Resiliency and continuity
- Network security

Operations

- Audit logging & monitoring
- Encryption and key management
- Identity and access management
- Incident management
- Privacy and data management
- Security development operation
- Threat and vulnerability management

Planning and Personnel

- Governance, risk, and compliance
- Human resources
- Supplier management

Reports and Assessments

- SOC Reports
- ISO Reports
- BCDR
- Penetration Tests
- Get started with the Service Trust Portal

Get started with compliance in Microsoft online services

If you're new to compliance and wondering where to start with Microsoft online services, this section provides links to key compliance areas for each service.

Microsoft Azure

Dynamics 365

Microsoft Purview

Azure compliance provides you with compliance offerings, blueprints, compliance guides, and resources to help you…

Business Central contains functionality and is built on a platform that makes it easy for customers to comply with…

If your organization needs to comply with legal or regulatory standards, start here to learn about governance and…

Learn about Microsoft Priva

Privacy is top of mind for organizations and consumers today, and concerns about how personal data is handled are steadily increasing. To meet regulatory requirements and build customer trust, organizations need to take a "privacy by default" stance and empower individuals to have control over their data. Robust privacy solutions are essential not only in ensuring regulatory adherence, but also in cultivating a culture of transparency, accountability, and respect for user privacy.

Microsoft Priva supports privacy operations across an organization's data landscape. Priva Privacy Risk Management, located in the Microsoft Priva portal, provides visibility into your organization's data. Customizable and easy-to-create policies help you identify privacy risks in your data estate and enable easy remediation to help reduce risks.

The Microsoft Priva portal

The Microsoft Priva portal is designed to give you streamlined access to all Priva privacy protection and privacy risk management capabilities. Learn more about accessing and navigating the Priva portal.

How and where Priva identifies items with personal data

Personal data is typically personal information that is related to a living person that can be used to identify that person. It may be a data type that can directly identify the individual, such as a name, passport number, social security number, and so on, or combinations of different data types that can be used to identify the individual. The definition of personal data or personal information may vary under applicable law, so make sure you understand the types of data for which you have legal obligations.

Priva utilizes foundational capabilities of Microsoft Purview to help you identify these personal data types based on your settings by using data classifications and sensitive information types (SITs). Organizations that create custom classifications and SITs can use them in Priva.

Priva evaluates your organization's data stored in the following Microsoft 365 services within your Microsoft 365 tenant:

- Exchange Online
- SharePoint Online
- OneDrive for Business
- Microsoft Teams

Priva evaluates data that is only within your organization's Microsoft 365 environment. It doesn't access personal data that isn't part of the organization's Microsoft 365 environment. For example, it doesn't access a user's personal Microsoft 365 account.

Priva also evaluates your organization's data sources that are registered through Microsoft Purview. Learn more about Microsoft Purview solutions for data governance.

Resources

- Priva Privacy Risk Management eBook
- Manage data privacy and data protection with Microsoft Priva and Microsoft Purview
- Microsoft Privacy

Next steps

To begin using Priva, see Get started with Priva.

Legal disclaimer

Microsoft Priva legal disclaimer

Last updated on 05/11/2026

Microsoft Privacy

Microsoft documentation and resources you need to get started with privacy for your business or organization.



TRAINING

Learn how Microsoft safeguards customer privacy



TRAINING

Learn about Microsoft 365 privacy standards



TRAINING

Learn about data protection, privacy regulation, and data processing



GET STARTED

Get started with privacy in Microsoft cloud services

EU Data Boundary

Understand the technical and organizational measures Microsoft has taken to provide the EU Data Boundary and deliver enhanced data residency solutions to our customers in Europe.

Get started

What is the EU Data Boundary?

Continuing data transfers for all services

Learn about continuing data transfers that apply to all EU Data Boundary services

Continuing data transfers for specific services

Learn about data transfers that apply to specific services

Configuring Azure non-regional services

Learn about configuring Azure non-regional services for the EU Data Boundary

Learn more about privacy in Microsoft services

If you're new to privacy and wondering where to start with Microsoft services, this section provides links to key privacy areas for each service.

Microsoft Azure

Learn how the privacy and confidentiality of the data you store and process in Azure services is protected and use...

Dynamics 365

Microsoft Dynamics 365 is committed to helping customers meet their privacy, personal data, and General...

Microsoft 365 Apps for enterprise

Microsoft is committed to providing you with the information and controls you need to make choices about...

Microsoft Windows

Learn how to meet General Data Protection Regulation (GDPR requirements by viewing and configuring...

Microsoft Priva

Get started with Priva to help your organization safeguard personal data across your data landscape and build a privacy-resilient workplace.

Learn about Priva

Priva offers organizations tools to proactively identify and protect against privacy risks and help you manage the...

Get started with Priva

Follow these steps to set up prerequisites and get started exploring the family of Priva solutions.

Manage risks with Privacy Risk Management

Priva gives you the capability to set up policies that identify privacy risks and enable easy remediation.

Microsoft Purview deployment models

Microsoft Purview provides integrated solutions and features that complement each other. For new Microsoft Purview administrators, it can be overwhelming to know where to start. But even existing administrators might be unfamiliar with how some Microsoft Purview features can enhance their data security.

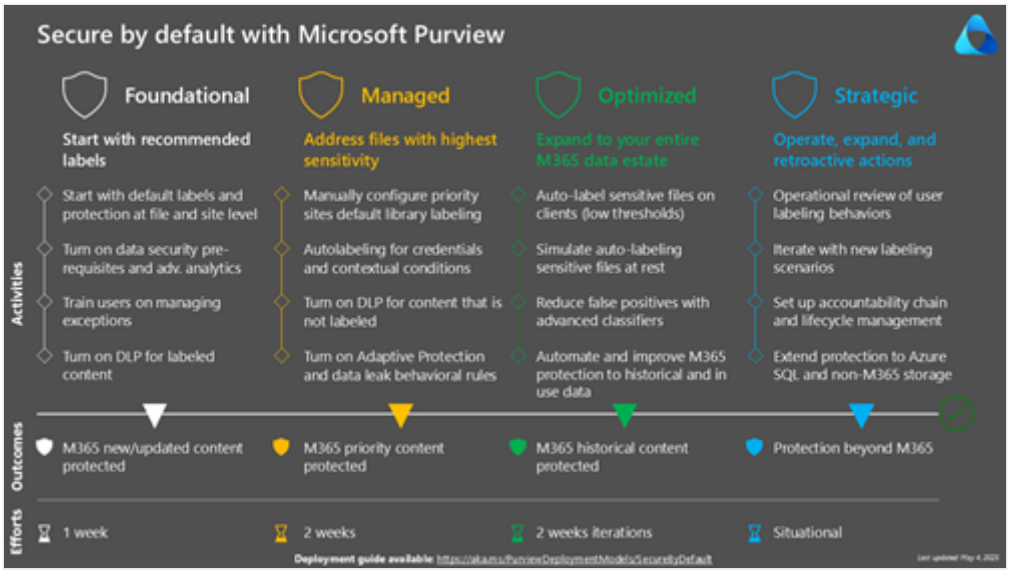
To address the need for shorter, actionable, and prescriptive guidance, the Product Engineering team is releasing a new series: Microsoft Purview Deployment Models. Deployment models are scenario-based and include a comprehensive step-by-step guide to jumpstart your deployment.

Audience

IT professionals, security professionals, and compliance officers who are responsible for deploying and managing Microsoft Purview in their organization. They don't provide detailed technical guidance but rather a high-level overview of the deployment process for a configuration that meets a specific business need. If you're new to Microsoft Purview technologies, concepts, and planning, deployment, and administration, see Learn about Microsoft Purview.

Deployment models come from our engineering teams. They're based on proven deployments with organizations. The objective of deployment models is to enable organizations to accelerate securing content and deploy faster.

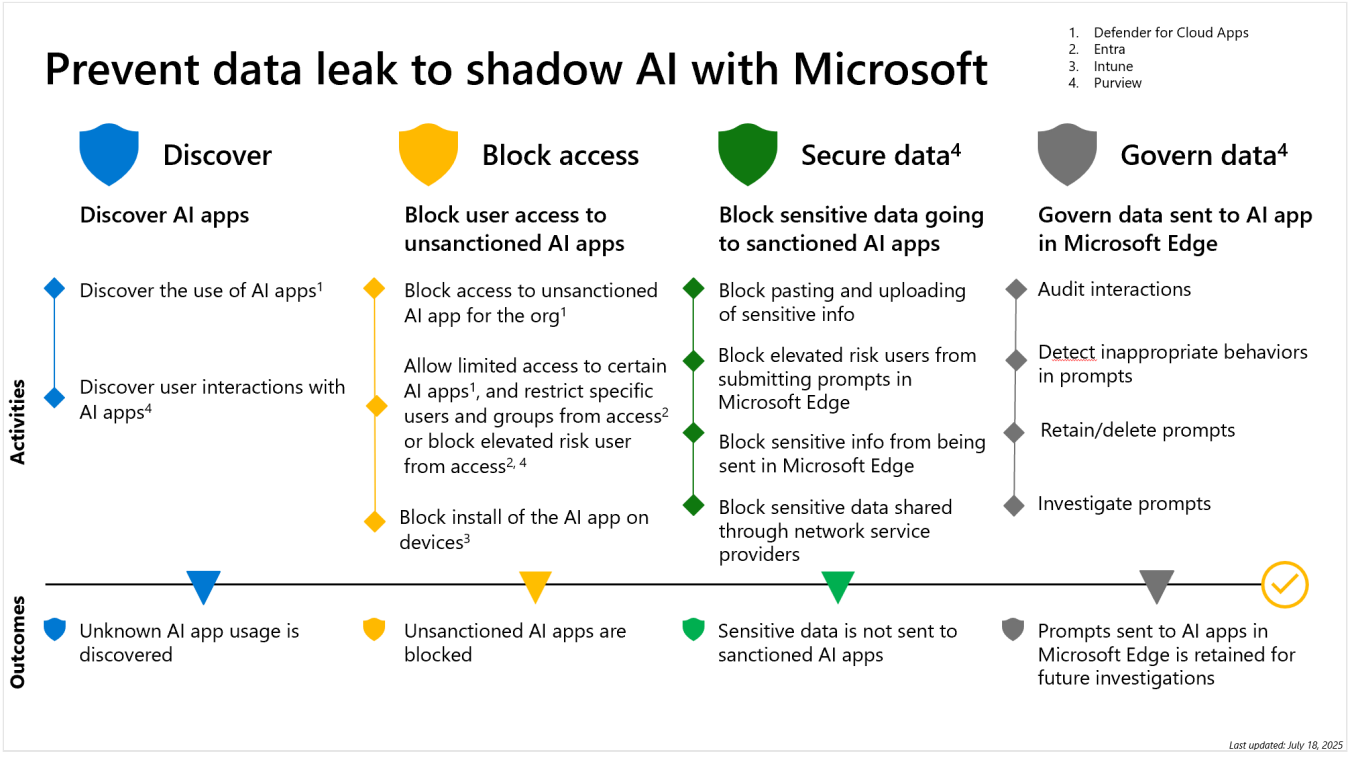
Secure by default with Microsoft Purview



Use this deployment model to help rapidly implement a secure by default configuration with Microsoft Purview Information Protection, Data Loss Prevention, and Insider Risk Management.

- Introduction
- Step 1: Start with default labeling
- Step 2: Address files with the highest sensitivity
- Step 3: Expand protection to your entire Microsoft 365 data estate
- Step 4: Operate, expand, and retroactive actions

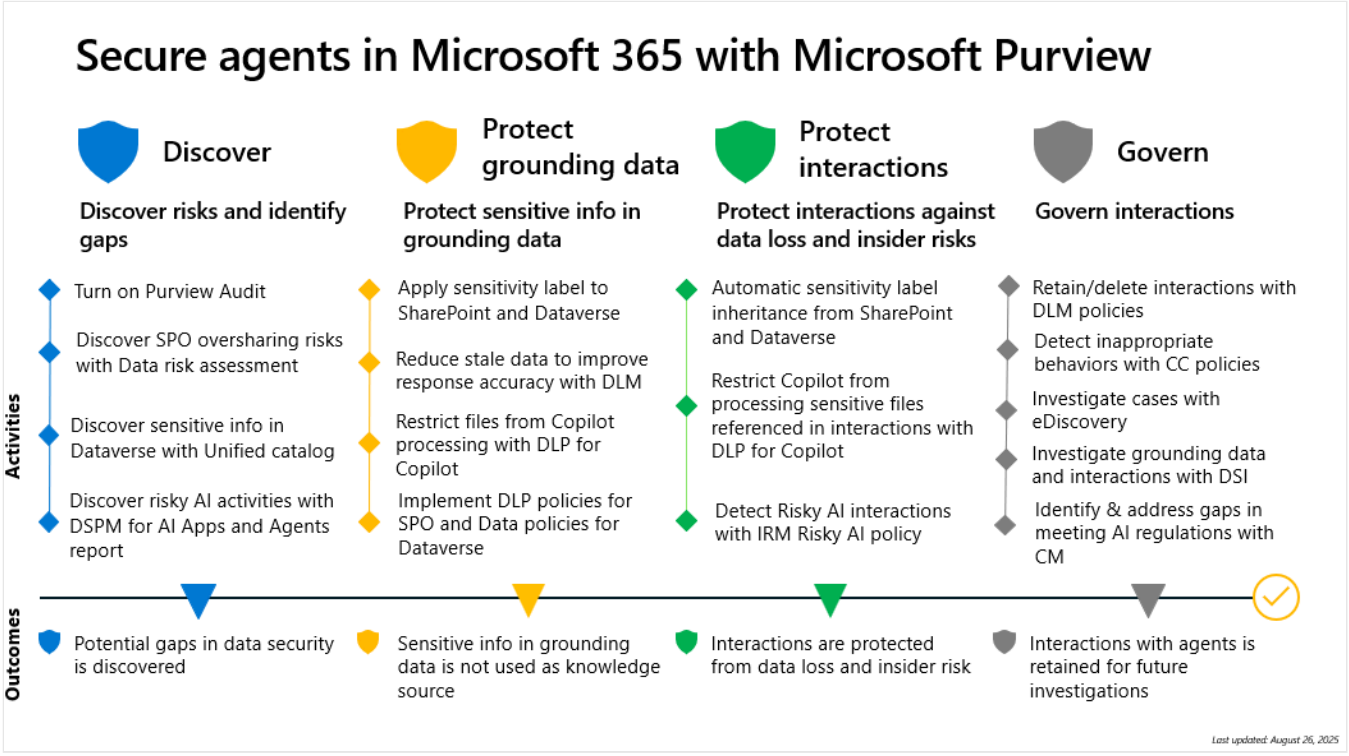
Prevent data leak to shadow AI



Use this deployment model to help your organization identify and prevent data leaks to shadow AI by using Microsoft Purview, Microsoft Defender for Cloud Apps, Microsoft Entra, and Microsoft Intune.

- Introduction
- Step 1: Discover AI apps
- Step 2: Block access to unsanctioned AI apps
- Step 3: Block sensitive data going to sanctioned AI apps
- Step 4: Govern data sent to AI apps

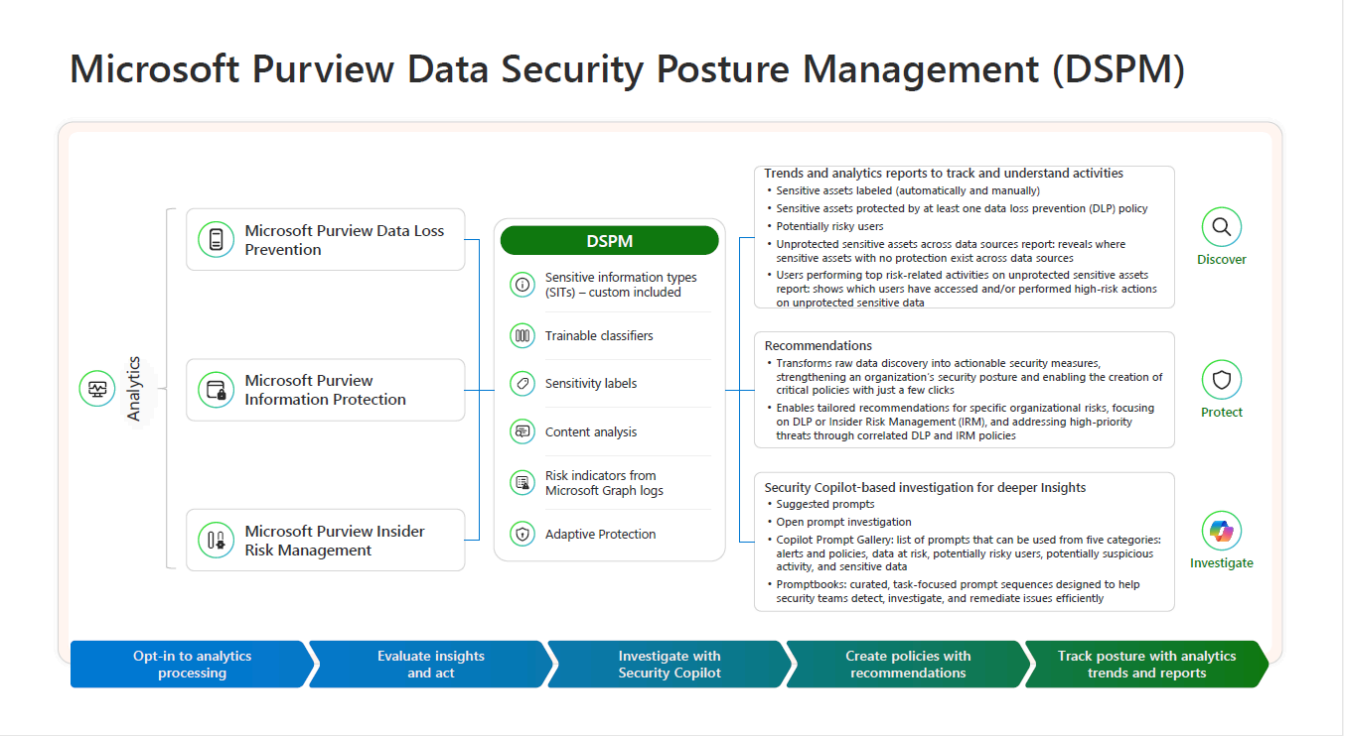
Secure and govern Microsoft 365 Copilot agents



Use this deployment model for Microsoft 365 Copilot agents by securing both the grounding data and interactions with Microsoft Purview.

- Introduction
- Step 1: Discover risks and identify gaps
- Step 2: Protect sensitive info in grounding data
- Step 3: Protect interactions
- Step 4: Govern interactions

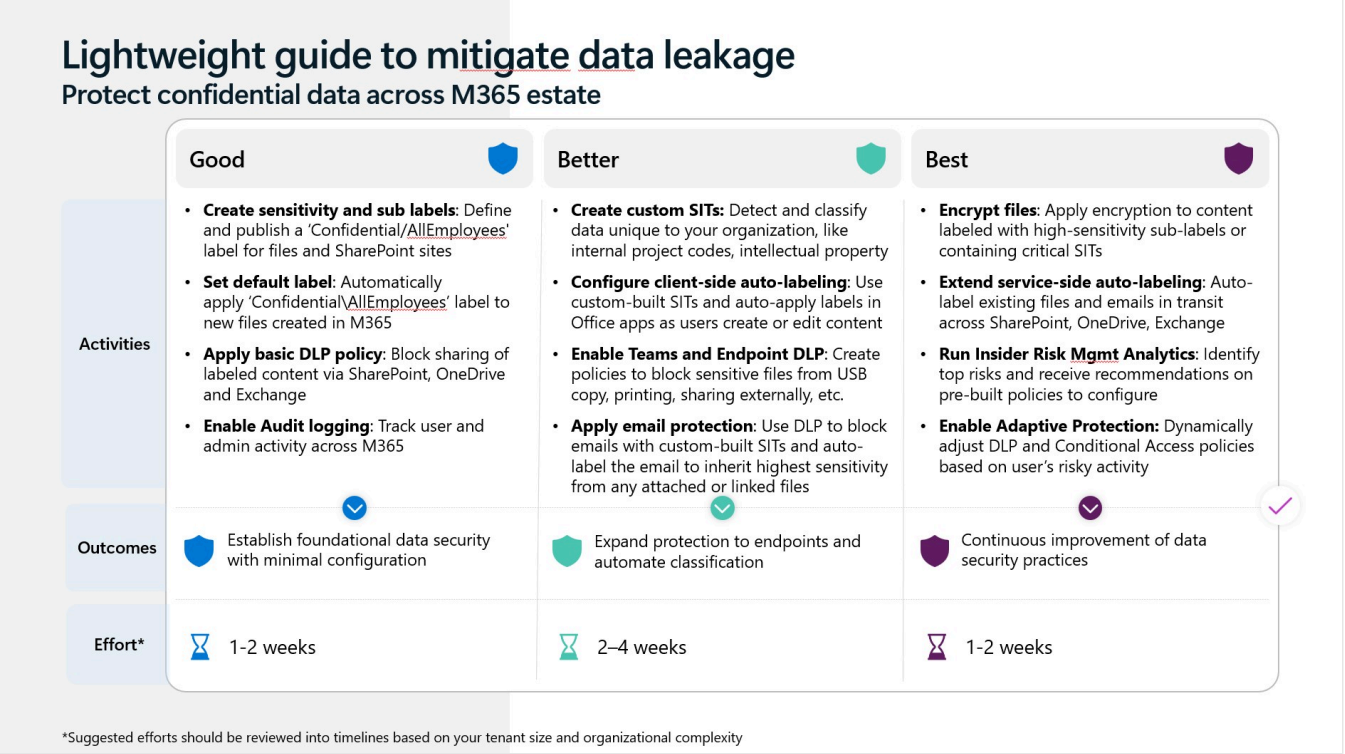
Deploy and use Data Security Posture Management



Use this deployment model to prepare for a Data Security Posture Management (DSPM) deployment with step-by-step guidance.

- Introduction
- Step 1: Establish foundational elements
- Step 2: Configure access and analytics
- Step 3: Understand data landscape and risks
- Step 4: Take action and investigate with Security Copilot

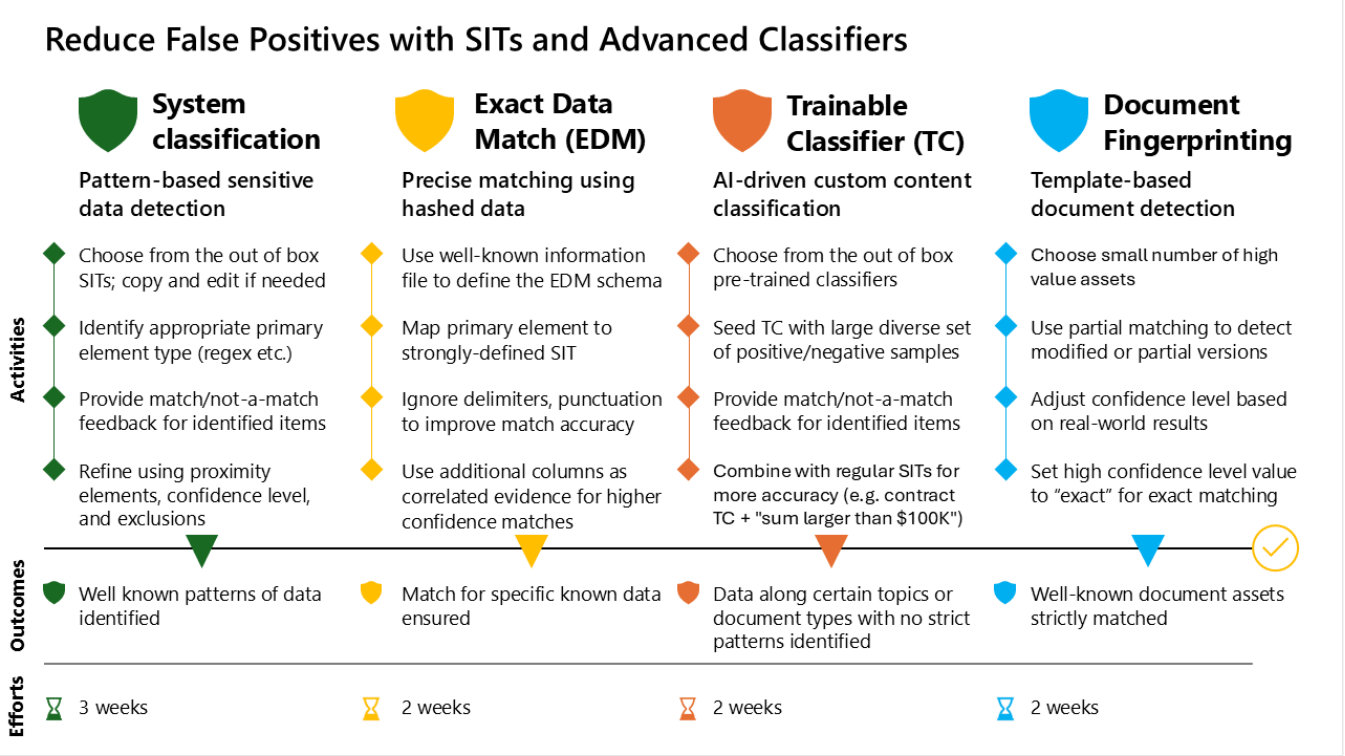
Lightweight guide to mitigate data leakage



Use this deployment model to enable core data security capabilities with minimal configuration, focusing on essential steps to get started quickly.

- Introduction
- Step 1: Establish foundational data security
- Step 2: Expand protection to endpoints
- Step 3: Continuous improvement

Reduce false positives by using SITs and advanced classifiers



Use this deployment model to minimize false positives by effectively leveraging the classification capabilities of Microsoft Purview.

- Introduction
- Step 1: Configure pattern-based sensitive data detection
- Step 2: Implement Exact Data Match for precise matching
- Step 3: Use trainable classifiers for complex content
- Step 4: Apply document fingerprinting for templates

ⓘ **Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 04/03/2026

Privacy in Microsoft Purview

Microsoft is committed to providing the information you need to manage privacy concerns in Microsoft Purview. All aspects of Microsoft Purview adhere to Microsoft's privacy governance model, which includes controls on data minimization, retention and deletion, location and transfer, sharing, and more.

For information about Microsoft's privacy governance model, see Privacy & data management overview - Microsoft Service Assurance.

Data security and privacy

Microsoft Purview processes and stores the following content types:

- Customer data
- End User Identifiable Information (EUII)
- Administrator data
- End User Pseudonymized Identifiers (EUPI)
- Feedback data

These types of data typically come from other Microsoft 365 services, such as Exchange, SharePoint, Microsoft Teams, and Copilot. Microsoft Purview also sources data when tenant admins configure policies and define data classification labels that can be applied to the organization's assets. This data includes emails, files, chats, messages, or Copilot interactions. EUII, administrator data, and EUPI data come from Microsoft Purview tenant administrators or users. Tenant administrators can also bring data from third-party sources so that the solutions offered by Microsoft Purview can extend to these sets of organizational data.

Customer data

You can categorize Microsoft Purview customer data as follows:

- Custom configuration settings.
- Search queries: In e-discovery, administrators can perform search and export results. This action could contain a portion of the data impacted by the user action, and the system retains it until the user deletes it. This category includes all Microsoft 365 user data, such as emails, files, Teams chat, and Copilot & AI interactions.
- Alerts

- Reports
- Tenant data that connectors ingest from third-party data sources.
- Data classification labels.

End User Identifiable Information (EUII)

End User Identifiable Information (EUII) includes a user's email address, User Principal Name (UPN), end user's IP address, address book data, user name, and user display name. Tenant admins choose how to process EUII data and decide which solutions to apply to these end users' data.

Administrator data

Administrator data includes the tenant admin's email address, UPN, IP address, user name, and display name. These types of data are specifically processed around tenant admin's activities when they perform administrative duties, such as:

- Set up policies, configurations, or connectors.
- Assign data classification label or policies to users.
- Put a user's data on legal hold.
- Assign compliance tasks to users.

End user pseudonymized identifier (EUPI)

End User Pseudonymized Identifier (EUPI) data includes UserID, UserKey, and Mailbox GUID. EUPI and EUII are used for user activities, audit logs, eDiscovery search, and service logs.

Feedback data

Feedback data is the feedback about the compliance solutions that a tenant administrator submitted through Unified Feedback Platform within Microsoft Purview.

How are these content types processed?

In the solutions that Microsoft Purview provides, customer data, EUII, administrator data, EUPI, and feedback data are applied to data assets such as emails, files, chats, messages, or Copilot and AI app interactions. For users managed by tenant admins, Microsoft Purview data processing can include monitoring for risky activities, preserving accounts and data, assigning tasks, and more. Microsoft Purview solutions handle these data assets as follows:

Expand table

Data asset	Processing
Exchange emails and calendar items	Microsoft Purview solutions process this data to enable: - Data classification labels - Legal hold - Data Loss Prevention policies - Encryption with customer's own encryption key
SharePoint and OneDrive files	
Messages from Teams, Skype for Business, and Viva Engage	
Copilot and AI app interactions.	
Email subject & metadata	Microsoft Purview solutions process this data for audit that tracks activities of tenant admins, analysts, and users.
Calendar item data	
List, filename, & metadata from SharePoint and OneDrive.	

Retention

Tenant administrators primarily control the retention of these types of data, except for the following cases:

- **eDiscovery:** Tenant administrators retain data that is in scope for eDiscovery case investigation until they delete the case when the legal investigation is complete, and the case is closed. A former employee's data isn't deleted until the investigation case is closed.
- **Audit:** User activity logs are retained for 180 days by default, 365 days for E5 licensed customers, and 10 years if customers purchase an additional license.
- **Communication policy hits:** Tenant administrators retain these as configured. The default policy retention is one year.
- **Configuration and policy data:** Tenant administrators retain this data until they delete configuration settings or policies.
- **Connector data:** Tenant administrators retain all data that is brought into the in-scope Microsoft Purview solutions through connectors until they remove the connectors and purge the data.

- **Administrator data:** Administrator data is retained until the person is no longer a tenant administrator.
- **Feedback data:** Feedback data is retained for 180 days.

When a user or the tenant leaves Microsoft 365 services, their data is retained for 90 days and no longer than 180 days to allow recovery or export of the data, except where specified by policy. For example, for eDiscovery, the data is retained until the tenant admin releases the user's account from legal hold.

Data residency

Customer data is stored based on where customer accounts are provisioned and in accordance with the commitments in the Online Service Terms. For more information, see Where your Microsoft 365 customer data is stored

Data sharing

Microsoft Purview solutions don't share data outside of Microsoft or across Microsoft without customer consent or authorization. Microsoft Purview compliance solutions share data with Microsoft Security Copilot as a plugin to Security Copilot. Microsoft Security Copilot processes the following data to provide the Purview Copilot capability:

- Microsoft 365 audit logs collected through the Microsoft Purview Audit pipeline
- eDiscovery and Communication Compliance data, which could include:
  - Emails
  - Teams conversations
  - Copilot and AI app interactions
  - Microsoft Purview alerts or insights generated by Data Loss Prevention (DLP) and Insider Risk Management (IRM)
  - Underlying data that led to such alerts or insights, such as audit logs and telemetry data collected by Microsoft Defender for Office related to emails and Teams, DLP, or IRM policies

For customers in the Government Community Cloud (GCC), data sharing between government and commercial cloud environments may occur, depending on the location of the service offering.

Data Subject Requests

For information about how to find, access, and act on personal data that resides in Microsoft IT systems, see: Introduction to DSRs.

Data privacy and protection – Respond to incidents and subject requests

Regulatory compliance

As a product, Microsoft Purview provides solutions to help manage risk and compliance for your organization. For more information about these product capabilities, see: Microsoft Purview risk and compliance solutions.

As an online service, Microsoft Purview aligns to the Microsoft Security Policy (MSP). For more information, see: Governance, risk, and compliance overview - Microsoft Service Assurance.

For more information about regulatory compliance in Microsoft Purview, see: Regulatory & industry compliance.

Resources

For more information about Microsoft privacy and compliance, see the following resources:

- Microsoft Trust Center
- Microsoft Privacy Statement
- General Data Protection Regulation
- The EU Data Boundary for the Microsoft Cloud
- Microsoft Privacy

ⓘ**Note:** The author created this article with assistance from AI. Learn more

---

Last updated on 06/04/2026